Salvador Pueyo Puntí

Irreversibility and Criticality in the Biosphere
Irreversibilitat i criticalitat a la Biosfera

Departament d'Ecologia
Universitat de Barcelona

Programa de Doctorat d'Ecologia de la Universitat de Barcelona, bienni 1993 – 1995. Memòria presentada per Salvador Pueyo Puntí per optar al títol de Doctor en Biologia, sota la direcció del Dr. Jordi Flos i Bassols.

Vist-i-plau del director de la tesi:

Jordi Flos i Bassols
Departament d'Ecologia
Universitat de Barcelona

Barcelona, gener del 2003

Als meus pares

Brief Index/*Índex abreujat*

Index/*Índex*

2. Physics of directionality in succession and evolution 91

3. Catastrophes, succession and criticality: the case of wildland fires 139

Introducció i visió general

Introducció

Si haguéssim estat a la Lluna (literalment parlant) ens n'hauríem adonat, i potser ens hauríem espantat. En un racó del nostre planeta hi havia una taca, força visible, que no parava de créixer. En cosa de mesos s'havia fet més gran que, diguem, Suïssa (amb perdó), i seguia creixent. Però de moment va deixar de créixer, i no estàvem a la Lluna, i no era a Suïssa, i la majoria no ens en vam adonar, o no ens vam espantar. Era l'any 1998. Succeïa en el nostre món, sí, però en una de tantes terres d'oblidades i oblidats, que no existeixen perquè no tenen. Era una d'aquelles regions on hi havia vida en una multiplicitat de formes excepcional, on hi havia moltes cultures. I no obstant, fora de qui hi vivia, quasi ningú no n'hauria sabut res si no hagués estat perquè els esquitxos de la catàstrofe tocaren les parts vitals de varis dels "dracs", de les "economies emergents" que han començat a existir perquè han començat a tenir. Els carrers de Singapur, Kuala Lampur, Jakarta... s'havien enfosquit, hi havia gent amb màscares anti-gas, aeroports tancats, un avió estavellat... La "taca" en si estava a l'est de Borneo. Era foc i cendres, però no qualsevol foc. Era el pitjor episodi de la història escrita dels incendis forestals, i era justament, misteriosament, en una regió de selva tropical. Misteriosament perquè no fa tants anys que deien els entesos/es que la selva tropical no es cremava, que era immune al foc. Els intents experimentals de calar foc a algun rodal de selva madura havien fracassat sistemàticament, la selva es feia el seu microclima humit i el foc no hi tenia entrada. Però quelcom ha canviat, i els incendis gegantins han proliferat a les selves de l'Insulíndia i de l'Amazònia. El misteri no està en els factors que ho han provocat, que es coneixen bé. Quan els tentacles del mercat global, escudats en les prescripcions econòmiques de les institucions globals, s'enduen les fustes més preuades, que hi ha esparses per aquí i per allà, foraden la selva i n'esguerren el microclima, i a llavors tot crema quan arriba un "Niño" tirant a sever. "El Niño" és una modificació en les condicions meteorològiques i oceàniques que afecta tot el món, especialment l'oceà Pacífic i les terres que l'envolten, i que es repeteix a intervals irregulars, d'uns tres a set anys. Mentre la fusta se'n va i la selva s'afebleix, El Niño es fa cada cop més agressiu a causa del canvi climàtic... Els factors que han capgirat les coses no són un misteri, el misteri està en la radicalitat del capgirament, en què en un tres i no res els rècords de no-foc hagin esdevingut rècords de foc.

Una proposta agosarada que faig en aquest llibre és que el que va succeir a Borneo, i a d'altres llocs, va ser una transició de fase. Una transició de fase és el que li passa a l'aigua quan es fa gel i al gel quan es fa aigua i, encara que sembli mentida, a priori és tan concebible per a una selva que tingui transicions de fase com ho és per a un got d'aigua. Les dades sobre terres "oblidades" són escasses i per tant només en dono evidència indirecta: indirecta, però evidència. L'aigua es fa gel a 0^oC (en condicions estàndard) i, si la meva proposta és encertada, la selva també canvia qualitativament quan se la força fins a un punt igual de concret, que per seguir la terminologia física haurem d'anomenar punt crític. La "criticalitat" del títol del llibre és aquesta, és la condició d'estar en un punt crític, en un sentit que la física defineix amb precisió. En el llibre hi faig moltes altres propostes agosarades, a més d'aquesta, i, si jo i/o alguns altres autors/es anem per bon camí, la idea de criticalitat no interessa tan sols quan pensem en respostes radicals de sistemes pertorbats, sinó que també serveix per comprendre aspectes bàsics de la naturalesa dels ecosistemes en el seu "dia a dia".

La "irreversibilitat" dels esdeveniments de Borneo és més òbvia que la seva "criticalitat". Les ferides de Borneo cicatritzaran si se les deixa, però el temps que fa falta perquè les coses tornin a ser comparables a grans trets (mai no ho seran en el detall) és extraordinàriament llarg en comparació amb el breu lapse en què s'han perdut: en aquest sentit, i també en d'altres, es tracta d'un procés irreversible. La naturalesa irreversible dels processos ecològics és evident, però tot i així hi ha algunes coses a aclarir en relació amb aquest tema.

El cas de Borneo no tan sols il·lustra els conceptes que encapçalen aquest llibre, d'irreversibilitat i criticalitat (junt amb "Biosfera", que és la capa viva del planeta) sinó també el "per què" del seguit de recerques teòriques en què m'he embarcat i que aquí recullo. El foc a la selva és un exemple de qüestió gravíssima, però que es coneix molt poc i no hi ha gaire interès a conèixer. Després d'haver fet aquest estudi considero que cal esmerçar tots els esforços per reunir-ne dades i aplicar-los les eines matemàtiques que es troben en aquest llibre, i d'altres que anem trobant. Potser així podrem confirmar amb més solidesa (o rebutjar) la proposta que he explicat, i potser fins i tot podrem quantificar els llindars crítics que no hem de violar. És clar que tot i així és probable que seguim sense respectar-los, però conèixer-los és una passa important perquè no sigui així, i una passa important per conèixer-los era conceptualitzar la possibilitat que existissin, reunir-ne evidències i posar les eines matemàtiques per anar més lluny, que és el que he fet per ara. En alguns altres temes que he investigat, l'aplicació ja es pot dur a terme en l'estat actual de la recerca (des de la predicció d'incendis en altres biomes fins a la quantificació de la diversitat), mentre d'altres estan en un estadi molt més especulatiu (com els que fan referència a la planificació de l'agricultura). El gruix de la meva feina ha estat d'immersió en qüestions molt abstractes, però sempre amb el propòsit d'acabar emergint amb eines per resoldre problemes pràctics i importants. No pas importants per al mercat ni per als cercles acadèmics, sinó aquells que percebo com a veritablement importants.

Els fets de Borneo es poden considerar una de les catàstrofes ecològiques més greus dels temps recents, però representen tan sols un episodi més en tot un entramat de canvis que estem vivint en aquest moment. Fa uns anys, vuit persones es van recloure en un recinte d'1.3 hectàrees sense cap intercanvi de matèria amb el món exterior, recinte en què es va intentar reproduir les funcions ecològiques que fan possible la continuïtat de la vida humana sobre el planeta. D'aquesta instal·lació, ubicada a Arizona, se'n va dir Biosfera II (Cohen i Tilman 1996). Els plans no van funcionar, tot se'n va anar en orris malgrat les despeses d'un milió de dòlars a l'any tan sols en energia. La composició de gasos es va destirotar, de manera que féu falta, primer, una injecció d'oxigen des de l'exterior i, al cap de no gaire temps, abandonar el projecte. Quasi totes les espècies que s'hi havia introduït es van extingir, els conreus es van quedar sense pol·linitzadors, els arbres es van tornar fràgils i trencadissos, hi va haver una explosió poblacional de formigues "boges" (és el nom de l'espècie)... La conclusió és que som incapaços/ces de reproduir un seguit de funcions a les quals no es dóna importància, perquè ja ens les vam trobar en marxa i mai no ens han demanat cap esforç. Ara la humanitat ha esdevingut un element de gran pes en el funcionament de la Biosfera (la de debò): s'estima que actualment marca el destí del 40% de la producció primària neta de les plantes de terres emergides (Vitousek et al. 1997) i del 25% de les aigües que hi precipiten (Postel et al. 1996), que és responsable del 50% de la fixació de nitrogen (Vitousek et al. 1997), que ha augmentat la concentració atmosfèrica de CO_2 en un 30% (WHO i UNEP 1996), que ha acabat amb el 50% de les selves tropicals (Myers 1995a), que ha multiplicat la taxa d'extincions d'espècies per un factor entre 1000 i 10000 (Pimm 1995)... Les principals funcions del planeta ja no ens són tan alienes, però no hem assumit les responsabilitats de la nostra posició. Altres civilitzacions havien estat prou importants per poder fer palesa la seva incompetència ecològica i desaparèixer, però mai com ara la incompetència no s'havia globalitzat.

Sóc dels qui pensem que, per garantir una vida digna a la generació actual i a les que vindran cal replantejar-se a fons la societat que hem fet o s'ha fet (i encara més, però no només, si també ens preocupa el benestar dels qui no són de la nostra espècie, com em passa a mi). No tan sols s'ha de replantejar per motius ecològics, però el coneixement ecològic ha de ser clau a l'hora de reorientar-la. Ja ho pensava abans de començar aquest projecte, i per això em vaig endinsar en els textos de teoria ecològica, esperant aprendre sobre les qüestions més profundes de l'ecologia, per saber com havíem d'enfocar un canvi social en profunditat. Però no hi vaig trobar el que hi buscava, no vaig veure-hi gran cosa que em resultés realment convincent, enlloc de treure'n respostes el que em van quedar van ser més preguntes... però també idees sobre com es podrien abordar. I a partir d'aquí la roda ha estat imparable. Per petita que fos la probabilitat que realment pogués resoldre alguna de les grans qüestions de fons de l'ecologia fent cas de les idees que se m'anaven acudint, la importància d'aconseguir-ho era tan gran que no he deixat de seguir el fil d'aquestes idees, em duguessin on em duguessin, sense respectar plans previs, procurant no deixar-me condicionar per les possibilitats de finançament ni de publicació ni per si la feina era més o menys agradable, i sense escatimar-hi temps, amb una certa actitud d'emergència. El fruit ha estat una acumulació considerable de resultats de tota mena, cobrint àrees molt variades, que tenen tanmateix una coherència de fons i esbossen una certa visió d'allò que se'n podria anomenar "ecologia fonamental", junt amb un seguit d'aplicacions concretes. En tots els casos, tenen a veure amb regularitats a nivell d'ecosistema. El títol d'"Irreversibilitat i criticalitat a la Biosfera" no reflecteix cap pla inicial, sinó un intent a posteriori de posar alguna etiqueta a les coses que he fet, a base de triar uns conceptes que han resultat tenir-hi un protagonisme especial, però no exclusiu. El concepte d'irreversibilitat fou el que va orientar més les fases inicials del meu treball, en mirar d'aprofundir en les propostes dels autors/es que volien fonamentar la teoria ecològica en la termodinàmica. A mesura que vaig anar tenint la impressió que aquesta aproximació tocava sostre i vaig cercar noves vies d'avanç, vaig anar constatant com la idea de criticalitat adquiria cada cop més força, hom diria que de manera espontània, sense esperar-m´ho, i m'anava donant eines per fer prediccions contrastables en molts camps. Em fa l'efecte que "criticalitat" serà un dels conceptes clau per a la ciència ecològica del futur.

Un dels motius de la meva insatisfacció amb l'ecologia que es fa avui dia és que s'ha transformat bàsicament en un seguit d'"estudis de casos", estudis d'una qüestió molt concreta en una espècie o unes poques espècies, en un indret particular... Els estudis de casos són indispensables, però una ciència no es pot fer tan sols a base d'estudis de casos, el que dóna cos a una ciència és allò comú i essencial que apareix un i altre cop quan furguem en profunditat rera cada cas concret. No sé per què l'ecologia actual és així. Potser per inèrcia acadèmica, per manca de preparació per emprar les eines adients, per l'entorn en què es mou l'activitat científica, que afavoreix fortament aquells treballs que es poden planificar amb detall abans de rebre el finançament, que es duen a terme en un temps preestablert i moderat i que és segur que donaran uns resultats publicables... Essent més mal pensat, fins i tot s'hi podria entreveure un rerafons ideològic, ja que entre els qui maneguen els diners per a recerca hi pot haver hagut molt més interès en l'estudi de qüestions concretes que permetin resoldre problemes petits però aparents (fer cosmètica ambiental, podríem dir-ne) o simplement no tinguin cap conseqüència pràctica, que no pas a abordar qüestions de fons, que ens interessen o ens haurien d'interessar més als qui pensem en propiciar transformacions de fons. Això pot haver canviat una mica arran de les cimeres internacionals en les quals s'ha negociat tractats sobre el medi ambient global, i que han resultat en la disponibilitat d'un cert finançament adreçat a què la "comunitat científica" engendri dictàmens consensuats sobre els grans problemes del planeta. Tinc la impressió que els resultats evidencien un cop més l'estat relativament precari de la ciència ecològica, i que les contribucions que han donat més força a aquests dictàmens no han sortit tant de l'ecologia com d'altres ciències relacionades que tenen una base teòrica molt més forta, com és el cas de la climatologia.

D'altra banda, ni es tracta que tothom es dediqui a fer teoria ni el temps ja emprat en estudis de casos és temps perdut, és temps que ha servit per acumular dades que poden fornir un bon brou de conreu en el qual podríem fer créixer el coneixement ecològic de fons força depressa, un cop ens hi posem, a més de tenir, és clar, l'interès que tinguin els "casos" en si mateixos. En altres ciències, com la física, sembla que hi hagi una interacció continuada entre avanços teòrics que suggereixen experiments i experiments que permeten contrastar els avanços teòrics i en suggereixen de nous. En ecologia podríem fer força via si tenim en compte que molts dels experiments o observacions que suggereixi la teoria ja s’hauran fet en algun moment, i només caldrà tornar-ne a tractar les dades de la manera oportuna per respondre les preguntes que ens anem fent (per a això aniria bé, és clar, facilitar molt més l’accés a les dades en brut, alhora que s’atorgui el just reconeixement a qui les hagi obtingut; val a dir que ja s’ha donat algunes passes en aquest sentit). En aquest llibre es poden trobar alguns exemples d’aquesta manera de fer.

De fet hi ha un volum apreciable de treballs emmarcats sota l’epígraf d’“ecologia teòrica”, però, contràriament a allò que hom podria esperar, rarament entren de veritat en qüestions de “teoria ecològica”, entesa com a marc conceptual en el qual encaixar el coneixement ecològic. La major part dels treballs d’“ecologia teòrica” estan completament dissociats del que fa la gent que es dedica a l’ecologia experimental o descriptiva i es limiten a fer també un “estudi de casos”, en aquest cas un estudi, sobretot per simulació d’ordinador (més que no pas teòric), de models inventats, més o menys inspirats en l’ecologia però que no tenen per què tenir-hi una relació real. Acostumen a quedar-se en un estadi més propi d’una “ecologia hipotètica” que no d’una “ecologia teòrica”. Està clar que aquesta recerca també és útil, perquè ajuda a explorar l’espai de possibilitats disponible a l’hora d’interpretar la realitat ecològica i per tant contribueix al “brou de conreu” del que parlàvem, però amb això tampoc no n´hi ha prou. Val a dir que una línia de recerca necessària perquè aquesta aportació al brou sigui profitosa és el desenvolupament de l’utillatge estadístic per estudiar amb dades empíriques les coses de les que parlen els ecòlegs/ogues teòrics/ques. En un cert moment (inclòs en el període d’elaboració de la meva tesi doctoral) em vaig dedicar a aquesta qüestió, i els resultats es poden trobar a Pueyo (1997). En aquest llibre només n’hi he inclòs un breu extracte (Ap. 4A).

Junt amb tot això també hi ha alguns treballs que veritablement aborden qüestions fonamentals. No obstant, aquest front de recerca té un pes molt petit actualment comparat amb ara fa trenta anys, i no és pas perquè en aquella època hagués quedat tot resolt. En aquells moments hi havia una efervescència de treballs que intentaven trobar regularitats ecològiques generals i emmarcar-les en un context físic més ampli. En el mateix període sorgiren treballs que atragueren l’atenció sobre els problemes ambientals globals i treballs que encetaren l’etapa actual de l’economia ecològica, alhora que creixia el moviment ecologista, en part nodrint-se d’aquests avenços intel·lectuals. Les grans aportacions a la teoria ecològica que hi hagué aleshores estaven ben fornides d’intuïcions profundes que les feien fascinants, però eren lluny de generar uns esquemes coherents i contrastats. Això és inevitable en qualsevol cosa que comença, i en aquests trenta anys la teoria ecològica hauria tingut temps de sobres per madurar i superar aquestes limitacions. Però no ha estat així. El llibre d'en Robert H. Peters (1991) "A critique for ecology" és representatiu del canvi que hi hagué en els punts de vista dominants. Peters mostra que l'ecologia ha estat i està poblada de conceptes no definits adequadament i que, per tant, no es poden estudiar de manera objectiva a partir de les dades, de manera que molt treball de recerca i moltes discussions han estat en el fons espúries. Però el que aquest autor encoratjava - i el que ha succeït en bona mesura - és que, en nom del rigor, l'ecologia renunciés a la profunditat i a la complexitat i esdevingués una massa desestructurada de relacions empíriques senzilles, necessàriament superficials.

Hi ha una minoria que no ha renunciat a intentar respondre qüestions ecològiques de fons, però, malauradament, només una minoria dins d'aquesta minoria es pot dir que hagi fet feina veritablement rigorosa. Si preveure la trajectòria d'una pilota llençada a l'aire ja té suficient complicació matemàtica per repel·lir la majoria de mortals, no es pot esperar menys de sistemes tan terriblement complexos com els ecosistemes. Per tant, les aportacions a la teoria ecològica solen justificar-se amb alguna seqüència paorosa d'equacions, o un seguit de raonaments obscurs, o apel·lacions a resultats de ciències que resulten remotes per a l'"ecòleg/oga mitjà/ana". No crec que erri gaire si dic que la repulsió que generen aquests ingredients és normalment suficient perquè les persones que avaluen, filtren, llegeixen, apliquen i propaguen els treballs quedin dissuadides de penetrar-hi a fons, de manera que els criteris que en dicten l'impacte acaben tenint poca relació amb la seva credibilitat real. D'això n'hauria de quedar convençut qualsevol que llegeixi amb detall aquest llibre (incloent les meves pròpies equacions paoroses, raonaments obscurs i apel·lacions a ciències remotes), ja que la meva dèria per tenir alguna idea clara sobre teoria ecològica m'ha dut a fer tants esforços com calguessin per mirar d'entendre la "lletra petita" de molts treballs, i el que he anat trobant deixava força a desitjar. De resultes, una part força gran de les contribucions novedoses en aquest llibre consisteixen a desmentir resultats que es troben a la bibliografia, cobrint des dels clàssics de l'ecologia fins a articles recents d'impacte, incloent-ne uns quants a Nature i Science. Ara mateix no em sembla que sigui gaire difícil trobar errades bàsiques fins i tot a aquests nivells, sempre i quan s'hi posi prou paciència.

Reconec que és "lleig" dedicar-se a criticar la feina d'altri, però si es vol que l'ecologia tingui capacitat predictiva és indispensable impedir que es rebleixi de resultats espuris. Una de les claus de l'avanç científic (o la clau de l'avanç científic, per als seguidors de Karl Popper) està en tenir i exercir la capacitat de desmentir suposicions falses, i aquí l'exerceixo abastament. En això m'acosto a en Peters, però no pas en la mena d'alternatives que proposem un i altre. S'ha de deixar clar que rarament els desmentits són desfetes totals, que molt sovint les propostes fallides tenen al darrera intuïcions molt interessants que poden sobreviure en una certa mesura a l'esfondrament de les argumentacions racionals en què han volgut convertir-se, i que, sigui o no així, poden haver estat passes necessàries en el que la ciència té de "prova i error". No sé si serà molt gran la part de les meves pròpies propostes que acabarà corrent la mateixa sort (si arriben a tenir prou ressò perquè algú es molesti a buscar-ne les falles), però podré estar ben content si el resultat net de la meva feina acaba sent tan positiu com el de la de molts/es dels qui critico. Val a dir que, donat que aquest llibre pretén donar a conèixer resultats nous més que no pas reiterar les coses que ja se sabien, corro el risc que sembli que subvaloro els treballs d'altres autors/es que sí que contenen aportacions correctes i interessants. En demano disculpes per avançat.

Amb aquest llibre faig aportacions de diferents tipus. D'algunes tinc la pretensió que puguin esdevenir bocins de la "ciència dura", o tirant a dura, que m'agradaria que l'ecologia fonamental arribés a ser algun dia. Entre aquestes s'inclouen deduccions a partir de "primers principis" i induccions posades a prova amb anàlisis estadístiques de dades d'accés més o menys públic, procedents de treballs d'altres persones. D'altres són simples especulacions, o fins i tot opinions personals, i en aquests casos he procurat que quedés ben clar. Cada especulació és alhora una proposta de recerca per a qui vulgui assumir-la.

Com explico amb un cert detall més avall, la via d'avanç que em desperta més expectatives, i la que més he fet servir en aquest treball, és la que es basa en la força dels "grans nombres", en la idea que, quan tenim molts elements junts però no gaire coordinats, hi ha certes propietats del conjunt que són bastant previsibles, encara que no sapiguem gaire sobre el detall de les parts (o que no n'hi hagi gran cosa a saber, si és que els detalls no paren de canviar). Tota la física estadística es basa en aquest principi, mentre que l'ecologia està molt lluny d'haver-ne tret el suc que en podria treure. Del que parlo no és senzillament d'aplicar la física estadística a l'ecologia, parlo de fer ecologia, de fer una ciència a la mesura de la realitat ecològica, que és molt diferent de la realitat dels sistemes més tradicionalment estudiats per la física estadística, però de fer-la amb una mentalitat molt més pròxima a la d'aquests disciplina i, és clar, aprofitant-ne alguns mètodes i conceptes quan estigui justificat (per exemple, el concepte de criticalitat, i uns quants més que poso a treballar al llarg del llibre). La importància ecològica de les regularitats basades en grans nombres ja va ser percebuda per Patten (1975), i hi ha alguns treballs que tenen bastant en comú amb aquest punt de vista, com els de "macroecologia" (p. ex. Brown 1995, Maurer 1999, Gaston i Blackburn 2000), els que han mirat d'introduir tècniques com l'ús d'equacions de difusió (Engen and Lande 1996a, b) o el MAXENT (Lurié and Wagensberg 1983) per analitzar l'estructura de les comunitats, o els que suggereixen exemples ecològics del que s'ha anomenat "criticalitat autoorganitzada" (p. ex. Solé et al. 1999, Solé i Alonso 1998, Drossel i Schwabl 1992), i d'altres.

Suposo que no sóc el jutge més imparcial per decidir l'encert del tipus d'aproximació a l'ecologia que faig ni de les meves troballes concretes, però en conjunt em crec el que he escrit i m'anima molt que un seguit de persones, el punt de vista de les quals valoro molt, hagin fet comentaris ben favorables sobre alguns retalls del meu treball que els he anat passant. Em va animar molt que el propi Ramon Margalef, una de les principals figures de l'ecologia del segle XX i, segurament, l'autor que més m'ha inspirat, em passés un comentari manuscrit sobre una versió prèvia del capítol 5, que tracta de patrons de diversitat, segons el qual "és un estudi excel· lent, per tant fa pensar en les mancances de la major part de les aproximacions a la diversitat". També m'han animat molt els comentaris d'en Jordi Flos, professor d'ecologia de la Universitat de Barcelona, pensador molt profund però poc conegut i director de la meva tesi, d'en David Jou, un dels pares de la termodinàmica estesa, professor de la Universitat Autònoma de Barcelona i poeta, i d'altres científics intel· ligents però poc coneguts com en Miquel Àngel Rodríguez-Arias, l'Emilia Gutiérrez, en Joan Lluís Pretus o en David Alonso.

Essent conseqüent amb les finalitats d'aquest treball, cloc el llibre amb algunes reflexions sobre com s'ha de repensar l'economia, a la llum d'una ecologia que m'he esforçat molt a repensar. Són tan sols reflexions d'algú que no és economista però a qui li agradaria ajudar, en la mesura de les seves possibilitats, a què acabi de madurar una economia ecològica com cal, que desbanqui els mites de l'economia convencional abans no sigui massa tard.

Resultats

Amb aquesta secció pretenc fer accessibles el gruix del contingut d'aquest treball sense que el lector/a s'hagi de submergir en el detall tècnic, que era inevitable consignar en el cos del llibre, i que allà queda registrat per a qui vulgui aprofundir més en algun dels temes. Hi explico les troballes que trobo més importants amb un llenguatge més planer i informal que no el dels altres capítols, incorporant-hi a més l'explicació dels principals conceptes previs que cal conèixer per poder comprendre les aportacions que faig en aquesta obra. Les parts d'aquesta secció no es corresponen amb capítols del llibre.

Somnis i desenganys termodinàmics

Vaig encetar la sèrie de recerques que han donat lloc a aquest llibre sota la poderosa influència del seguit d'autors que han optat per una aproximació termodinàmica a l'ecologia o a l'economia ecològica. En cas de ser cert, el que diuen aquests autors tindria unes implicacions extraordinàries.

La termodinàmica permet preveure en quina direcció canviaran irreversiblement certs sistemes en certes condicions. El postulat més conegut de la termodinàmica és sens dubte la Segona Llei, que diu que qualsevol sistema aïllat tendeix a assolir el màxim d'una certa variable que s'anomena entropia i que té una certa relació intuïtiva amb el desordre. Per exemple, si tenim un recipient aïllat amb dos gasos, posem oxigen i nitrogen, amb les molècules ben ordenades, cadascuna de les d'oxigen a una banda i cadascuna de les de nitrogen a una altra, està clar que acabarà tot ben barrejat... ben entròpic. Si el sistema no està aïllat, sinó que té alguna mena d'intercanvi amb l'exterior, les coses són més complicades i ja no té per què assolir la màxima entropia. No obstant, en aquest sistema encara hi pot haver canvis irreversibles, i si no fan augmentar l'entropia en el propi sistema la faran augmentar en alguna altra banda, de manera que es pot dir que qualsevol procés irreversible va associat a una "producció d'entropia". Per exemple, el lector/a no té per què augmentar la seva entropia al llarg del temps, però si pensa en tot allò que ha ingerit al llarg de la seva vida i què n'ha quedat després estarà d'acord que, en conjunt, ha estat un gran productor/a d'entropia. Si ha estat possible que aparegués i es mantingués la vida sobre la Terra ha estat perquè contínuament rebem llum del Sol, que després se'n torna cap a l'espai exterior en forma d'infraroigs, cosa que implica un augment d'entropia.

L'interès que l'economia ecològica ha dipositat en la termodinàmica es deu principalment al fet que qualsevol activitat econòmica té un "cost" termodinàmic inexcusable, ja que demanarà algun augment d'entropia en algun lloc o altre. Per tant, els sistemes econòmics estan restringits per la capacitat del seu entorn (en el sentit més ampli de la paraula) per absorbir l'entropia que produeixen, com ho estan els propis ecosistemes. Es pot dir que d'això no n'hi ha dubte i, en tot cas, el que s'ha d'acabar d'aclarir és el grau de rellevància pràctica que té.

Els estudiosos/es de la teoria ecològica que s'han aproximat a la termodinàmica han volgut anar encara més lluny. S'ha de dir que, tot i que el principi de la màxima entropia només serveix per a sistemes isolats, es poden definir altres variables, anomenades potencials termodinàmics, que tendeixen a un màxim o un mínim en alguns sistemes senzills que no estan aïllats. Per exemple, un cristall no està pas en un estat de màxima entropia, però si, a més de l'entropia del cristall, també tenim en compte la calor que ha desprès en formar-se i que ha fet augmentar l'entropia del seu entorn, podem definir un potencial que s'anomena energia lliure i que haurà assolit el seu mínim. La gran aposta dels ecòlegs/ogues més atrets per la termodinàmica ha estat postular que també en els sistemes tan complexos com els ecosistemes o, pel cas, els sistemes socio-econòmics, es pot definir algun potencial termodinàmic... tot i que cadascun dels autors d'impacte n'ha triat un de diferent (exergia, "emergia" amb "m", producció d'entropia...). Si això fos cert, les conseqüències serien extraordinàries, i no sé si els propis autors que ho defensen se n'han adonat del tot. Podríem predir totes les característiques d'un ecosistema un cop hagués passat prou temps sense pertorbacions externes, perquè serien les d'un estat molt ben definit, el que correspon al màxim o al mínim del potencial. I qui diu un ecosistema diu la Biosfera en conjunt. I, per què no, el mateix seria aplicable als sistemes socio-econòmics, en tant que part de la Biosfera i en tant que sistemes complexos per dret propi. Qualsevol discrepància política de fons esdevindria supèrflua, ja que el sistema econòmic i polític del futur estaria predeterminat per les lleis de la física. Els postulats que defensen l'existència de potencials d'aquesta mena van tenir un gran predicament cap als anys 70, i encara el tenen entre la minoria d’autors/es que segueix treballant pròpiament en teoria ecològica. En un moment donat jo també m’ho vaig creure força, que algun d’aquests potencials funcionava de veritat, pel prestigi de l’autor que ho proposava i la solvència de la ciència física en la qual, presumptament, es fonamentava. Però l’intent d’assimilar-ho fins a les darreres conseqüències i de fer-ho compatible amb els fets quotidians que m’envoltaven se’m feia molt dur. Dono la raó a Stefan Baumgärtner (1996) quan diu que “potser el concepte d’entropia hauria d’anar amb un avís de risc per a la salut!”. Tot i així, vaig voler aprofundir-hi més i vaig fer l’esforç de comprendre fins el darrer detall els arguments en els quals es basaven aquestes asseveracions. El resultat va ser que es van esvair. Ara mateix tinc força clar que no hi ha justificació per a cap dels potencials que s’han proposat (sec. 2.2). A sobre, he pogut constatar que la variable que actualment gaudeix d’una atenció més intensa com a suposat potencial, que és l’anomenada “exergia”, es mesura aplicant a les dades de camp unes equacions incorrectes, que introdueixen un biaix que arriba a ser de més del 70000% (subsec. 1.3.1).

La il· lusió dels potencials termodinàmics va sorgir de l’estudi de la successió ecològica. “Successió” es podria definir com al seguit de canvis que acusa un ecosistema mentre no està subjecte a cap pertorbació. Pensem, per exemple, en tot allò que s’esdevé en un bosc tallat fins que no arriba a recuperar-se, si abans no pateix cap maltempsada comparable. En els anys 70, Ramon Margalef per una banda i Edward P. Odum per una altra van proposar que hi havia un seguit de tendències que es posarien de manifest a tota mena d’ecosistemes al llarg de la successió, com ara un decreixement en la relació entre producció i biomassa, un major control dels cicles de nutrients per part dels éssers vius, unes mides més grans dels organismes, una major especialització, heterogeneïtat espacial, etc. La idea més general i intuïtiva és, segurament, que els ecosistemes esdevenen més complexos al llarg de la successió. La manera més elemental d’explicar aquestes tendències era atribuir-les a algun potencial universal, i així ho feren de manera més o menys temptativa varis autors, incloent el propi Margalef. En el llibre mostro que les justificacions concretes per a cadascun dels potencials fallen ostensiblement (sec. 2.2), però hi ha alguna altra manera d’explicar aquesta mena d’observacions entorn de la successió? (en la mesura que es puguin considerar acurades, cosa que no tots els autors/es accepten). Certament hi és (subsec. 2.3.1-2.3.2), i la clau ens la dóna el propi Margalef, en una observació posterior, aparentment elemental però, de fet, molt profunda.

Margalef diu que hi ha una asimetria en el temps tal que, en qualsevol sistema, la complexitat augmenta a poc a poc i decreix sobtadament. Pensant una mica en les realitats de la vida, sembla clar que aquesta regularitat es compleix en termes generals, tot i que podríem imaginar excepcions. El que sí que sembla innegable és que un impacte "massa fort" de qualsevol mena actua com a una pertorbació, destrueix (l'etiqueta d'un conegut rom nicaragüenc compleix hàbilment amb els requeriments de sanitat informant-nos que "todo en exceso perjudica"). Dit d'una manera més acurada (i menys tautològica), sempre hi ha alguna velocitat dels canvis tal que, si es força un sistema complex a excedir-la, el resultat tan sols pot ser una pèrdua de complexitat. I això, per què? (subsec. 2.3.1) Una resposta com cal demana aclarir què entenem per complexitat. La majoria de definicions formals de les quals disposem provenen de la teoria de la computació i tenen a veure amb el nombre de passes lògiques que calen per aconseguir l'element del qual estiguem parlant. Si mai tinguéssim la capacitat de simular amb tot detall un arbre fent servir un ordinador, sens dubte l'ordinador s'hauria d'estar molta més estona fent operacions de tota mena que si ens conforméssim simulant una pedra. Si la complexitat de la qual parlem és aquesta, i segurament sí que l'és, està clar que per generar-se de manera natural demanarà temps, i en canvi la podrem perdre tan de pressa com vulguem. D'altra banda, cal tenir en compte que les "passes lògiques" de la natura són de moltes menes i poden tenir velocitats molt variades, però sempre podrem dir que en el límit de la velocitat molt alta només hi pot haver destrucció.

Què té a veure l'asimetria temporal d'en Margalef amb els suposats potencials termodinàmics de la successió? Posem que un ecosistema tendeixi, per la seva pròpia dinàmica interna, a assolir un cert nivell de complexitat, que no té per què ser la màxima concebible ni la màxima sota cap criteri en particular, és la que tendeix a assolir i ja està. Com a conseqüència de l'asimetria temporal abans esmentada, les pertorbacions només podran reduir la complexitat, no pas augmentar-la (la caiguda d'un asteroide mai no farà sorgir dinosaures instantanis...). La tendència entre pertorbació i pertorbació serà a recuperar el nivell de complexitat característic del sistema, i en concret serà una tendència a augmentar la complexitat, però no pas perquè la complexitat final hagi de ser un màxim de cap mena, sinó perquè la pertorbació anterior tan sols ha pogut reduir-la, fos quina fos la complexitat que hi havia abans. Per tant, l'existència de "tendències" és previsible a priori, i no queda res que calgui justificar amb cap potencial termodinàmic (subsec. 2.3.2). Un cop explicat, això sembla trivial, però no ho és. La prova és la gran talla intel·lectual dels ecòlegs de primera línia que han dedicat els seus esforços a desenvolupar interpretacions termodinàmiques de la successió, i el fet que mai ningú no els hagi plantejat abans aquesta alternativa (que jo sàpiga).

Abans d'arribar aquesta conclusió en la seva forma actual vaig tenir temps de fer una bona immersió en la termodinàmica dels sistemes molt complexos i que s'autoorganitzen, com els ecosistemes, i de fer algunes troballes addicionals (cap. 1). El que de moment en sabem, d'aquest tema, és que els sistemes en qüestió només poden existir si de manera continuada o reiterada produeixen entropia que expulsen cap al seu entorn; per tant, es diu que estan fora de l'equilibri (a diferència de, per exemple, els cristalls, que poden mantenir-se indefinidament en un estat d'equilibri un cop s'han format, moment en el qual ja s'ha produït tota l'entropia que s'havia de produir). La seva producció d'entropia per unitat de temps no pot ser arbitràriament baixa, i per tant es diu que estan no només "fora" sinó també "lluny" de l'equilibri. A més, experiments senzills han mostrat de manera frapant la facilitat amb què, en establir condicions allunyades de l'equilibri (per exemple en escalfar un fluid des de baix i deixar que se'n vagi la calor cap amunt, o en mantenir certes reaccions químiques actives subministrant reactius i enretirant-ne els productes) poden aparèixer estructures ordenades, lògicament no gaire complexes (per a tenir-ne de complexes caldria temps: el que no s'ha assolit ni sembla que es pugui assolir és la generació espontània). Un cop establert aquest marc tan genèric, la termodinàmica no ha aportat fins ara gran cosa que ens orientés sobre com evolucionen aquests sistemes. Per si no n'estava prou, això ha quedat encara més clar amb el fracàs dels intents de trobar-hi potencials termodinàmics per part d'autors procedents de l'ecologia. Però en aquest llibre proposo que alguna cosa més sí que en podem treure, de la termodinàmica.

Una cosa que semblen haver oblidat molts/es dels/les qui s'han dedicat a la termodinàmica dels ecosistemes i de sistemes comparables és que es tracta de sistemes històrics, amb memòria, en els quals la informació hi té un paper essencial. Ja siguin els filaments d'ADN que es copien i es tradueixen a polipèptids, els senyals captats pels sentits dels organismes, les seves respostes, les funcions neurobiològiques, l'adaptació de les espècies a l'entorn i la coadaptació entre diferents espècies, els canvis en la composició dels ecosistemes com a resposta als canvis ambientals... són ben inherents a la vida la còpia, el processament, la conservació i les pèrdues d'informació. Quan, com a resultat dels fluxos d'informació, dues coses tenen unes certes correspondències entre elles (per exemple, les correspondències que hi ha entre una molècula d'ADN i una proteïna que s'ha format a partir d'aquesta, o, és clar, entre dues molècules d'ADN germanes), es diu que tenen "informació mútua" (subsec. 1.2.2). La informació mútua no pertany a cap de les dues coses en particular, és quelcom compartit: el que la proteïna "sap" sobre l'ADN no és ni més ni menys que el que l'ADN "sap" sobre la proteïna. Un punt important és que la informació mútua té significació termodinàmica. Tornem al recipient amb oxigen a una banda i nitrogen a l'altra. Posem que no només tenim un recipient, sinó que en tenim molts de semblants. En primera aproximació, l'entropia de tots els recipients junts és la suma de les entropies de cadascun, i la cosa se sol deixar així. Però no és exacte: si, a sobre de tenir l'oxigen separat del nitrogen, disposem que tots els recipients el tinguin al mateix costat, podrem dir que hi ha una informació mútua entre ells, i l'entropia del conjunt serà més petita que la suma de les entropies. En aquest cas la diferència és totalment negligible i podem ignorar aquesta qüestió, però la cosa canvia quan entrem en el domini de l'alta complexitat dels processos biològics (sec. 1.3). Per varis motius: no tan sols la informació és la que dirigeix els processos irreversibles que tenen lloc en els sistemes vius, sinó que el seguit d'informacions mútues que contenen té per si mateix un pes no negligible en els balanços d'entropia (tot i que petit, si més no en les seves components més evidents). A més, si la vida només pot existir lluny de l'equilibri és justament perquè tan sols en aquestes condicions hi pot haver els fluxos d'informació de la mena que la defineixen, i la Segona Llei en determina algunes propietats bàsiques, com veurem. Una de les coses que faig en aquest llibre és esbossar la incorporació de termes d'informació mútua en la termodinàmica dels sistemes allunyats de l'equilibri (subsec. 1.2.4), incloent els ecosistemes, cosa que es pot dir que obre una porta a l'holisme sense abandonar la "ciència dura". Quan hi ha informació mútua entre les parts d'un sistema, aquesta informació mútua pertany al conjunt, no a les parts, i no obstant està subjecta a una llei objectiva, predictiva i ben establerta com és la Segona Llei de la termodinàmica. A més, mostro que això és cert no només si mesurem la informació mútua de variables que tinguin una significació directa en termes d'entropia, com són les posicions i velocitats de les molècules, sinó per qualsevol variable que mesurem, per molt macroscòpica que sigui (Ap. 1B). Per exemple, fa temps que alguns autors han sospitat que certs aspectes de l'evolució biològica es poden reduir a termodinàmica. Amb el formalisme de la informació mútua algunes d'aquestes sospites es poden confirmar de manera molt directa (subsec. 1.3.2). Posem que dues poblacions d'una espècie queden totalment isolades en termes de reproducció. Sabem que aquestes poblacions divergiran al llarg del temps en moltes característiques, especialment aquelles que no venen molt marcades per la selecció natural, i acabaran esdevenint espècies diferents. Doncs aquest fenomen és una manifestació de la Segona Llei, en virtut de la qual la informació mútua entre les dues poblacions es dissipa al llarg del temps. Naturalment, d'informació mútua també se'n pot generar a més de perdre-se'n, però per a això cal el que anomeno un "acoblament actiu" (subsec. 1.2.5): cal que hi hagi un mecanisme que estableixi una interacció entre els dos elements entre els quals hagi de sorgir la informació mútua i l'associï a un increment d'entropia en algun tercer element o en algun dels propis elements implicats, superior a la reducció d'entropia que representa la generació d'informació mútua, de manera que en quedi "pagat" el "cost" termodinàmic. Qualsevol procés de mesura, observació, còpia, inscripció, etc. és un acoblament actiu i requereix un consum d'energia. Ja a començaments del segle XX hi va haver autors que postulaven, correctament, que prendre una mesura de qualsevol tipus tenia un cost termodinàmic. Després s'han enrevessat les coses i ara se sol pensar que el que té un cost és esborrar la informació enlloc d'adquirir-la, però fent servir el formalisme de la informació mútua mostro que això no és exacte, que en el fons el que té un cost és adquirir-la (per exemple, en prendre una mesura), tal com ens diu inicialment la intuïció (subsec. 1.2.3). I encara podem anar molt més enllà si ens fixem bé en l'altra cara d'aquesta regla, és a dir, que mentre no hi hagi cap acoblament actiu, la informació mútua només es pot perdre, no pas guanyar. Com a resultat, tota informació mútua (fora de l'equilibri termodinàmic) és conseqüència d'algun acoblament actiu que hi va haver en el passat. Sembla clar que aquest fet no és ni més ni menys que allò que, en les nostres vides quotidianes, expressem dient que recordem esdeveniments del passat, però en canvi no podem "recordar" esdeveniments del futur (subsec. 1.2.5). Del fet que hi hagi memòria però no precognició se'n diu que és una "sageta del temps", perquè estableix una diferència entre passat i futur, justament la més òbvia per als humans. La Segona Llei, que diu que l'entropia futura de l'Univers mai no serà menor que la passada, és una altra sageta del temps. Fa temps que se sospita que hi ha una connexió entre les dues sagetes, però fins ara els intents de formular-la no eren prou satisfactoris i donaven lloc a paradoxes. Fent servir el concepte d'informació mútua, la connexió queda clara i les paradoxes desapareixen. Lògicament, tot això serveix tant per a la memòria pròpiament dita, la dels sistemes nerviosos, com per a la munió de "memòries" que s'acumulen arreu, ja sigui en forma d'anells de creixement, d'espècies indicadores, de característiques genètiques que reflecteixen la història evolutiva o d'empremtes, traces, marques o indicis de qualsevol mena que permetin reconstruir la història d'un sistema.

Anteriorment he parlat d'una altra sageta del temps, segons la qual quan un sistema es força a canviar molt de pressa el resultat serà una pèrdua de complexitat, mai un guany. L'he justificat parlant de processos de computació, de processament d'informació. En el darrer paràgraf ha quedat clar que el processament d'informació és asimètric en el temps, té lloc "de passat a futur", com a conseqüència de la Segona Llei. Alhora, l'asimetria temporal del processament d'informació estableix l'orientació temporal de l'asimetria en les variacions de complexitat, que per tant també és una conseqüència de la Segona Llei de la termodinàmica. Fet i fet, els autors/es que pensen que les tendències de la successió ecològica tenen una base termodinàmica no és pas en això en el que s'equivoquen. Fa temps que es creu que la termodinàmica és l'únic que introdueix una asimetria temporal en el rang d'escales que ens interessa, i per tant qualsevol fenomen que trobem asimètric en el temps hi ha de tenir alguna mena de relació. El problema és que aquests autors/es s'han quedat en la forma més trivial d'establir-hi la connexió, que és pensant en termes de potencials termodinàmics. En realitat la connexió hi és, però és terriblement més subtil, com hem vist (subsec. 2.3.1-2.3.2).

Malauradament, tota aquesta feina entorn de la termodinàmica dels ecosistemes quasi no m'ha servit per fer prediccions noves. Bàsicament, pel que ha servit és per lligar coses que ja sabíem però no sabíem que tenien relació entre elles, per expressar formalment i justificar coses que en certa manera sabíem, però no sabíem que sabíem, i per rebutjar un munt de teories errònies i amb un impacte considerable i tancar les portes a d'altres teories semblants. D'aquestes teories sí que se'n pretenia extreure prediccions noves, que han quedat automàticament desautoritzades, la qual cosa no deixa de ser una manera de fer prediccions noves (la predicció que les prediccions de les teories en qüestió no seran encertades). Potser sí que queden grans avanços a fer dins d'aquest terreny que finalment ens duran sorpreses interessants i clarament útils, però en això no sóc particularment optimista. De fet, molts dels plantejaments de les pròximes seccions, que sí que permeten fer prediccions noves (i, al meu parer, interessants), tenen fortes connexions conceptuals amb la termodinàmica, però ja no són pròpiament termodinàmica.

A banda d'això, la qüestió de les tendències de la successió i d'altres de relacionades es deu poder refinar analitzant quina mena d'organismes estan "optimitzats" per a cada etapa de la successió, en base a la selecció natural (sec. 2.4). Un resultat (relativament nou) que mostro (subsec. 2.4.1) és que, fins i tot quan hi ha xarxes d'interaccions complicades, es pot preveure, sota condicions molt àmplies, que els organismes tendiran a ubicar-se en un màxim relatiu de la taxa de creixement poblacional en el "mapa" de possibilitats genètiques, donades les condicions de contorn en què es trobin. És clar que aquestes condicions de contorn canviaran com a fruit de la pròpia evolució dels organismes i la taxa de creixement mitjà sempre acabarà sent zero (tant si ja és un màxim com si encara no), i de tant en tant, a més, algun màxim relatiu deixarà de ser-ho (cosa que pot generar dinàmiques interessants), però el cas és que aquests màxims relatius segueixen sent útils com a referents i que es reforça l'interès dels molts treballs que s'han fet amb la idea d'optimitzacions evolutives. Aquests "potencials parcials" actuant sobre els organismes (o, si es vol, les espècies) no es tradueixen de cap manera genèrica en potencials actuant sobre l'ecosistema en conjunt, però poden ajudar a comprendre millor les tendències de la successió. Només en condicions molt restringides (subsec. 2.4.2), en què pràcticament l'únic tipus d'interacció rellevant sigui la competència basada en l'eficiència en la utilització dels recursos un cop absorbits pels organismes, es pot esperar que l'ecosistema tendeixi a assolir l'extrem d'una variable, en concret tendeixi al màxim de biomassa (o d'alguna mesura més apropiada de la "magnitud" del sistema biològic).

Tot i que aquesta aproximació més convencional a la qual duen aquests darrers comentaris té una utilitat força clara, considero que la idea de Margalef que he mirat de refinar, de l'asimetria temporal entre guanys i pèrdues de complexitat, és quelcom que mereix molta reflexió i que podria ser encara més útil, a la seva manera. La percepció qualitativa de la natura afecta la nostra forma d'interaccionar-hi, i pot acabar tenint tantes o més conseqüències pràctiques que les receptes per fer prediccions quantitatives. En la darrera subsecció d'aquesta "visió general" espero deixar clar per què dic que aquest principi (modificat) d'en Margalef té tant d'interès.

Quan més serem, més riurem: els “grans nombres” com a clau de l’ecologia?

No sembla que la termodinàmica dels ecosistemes sigui capaç de complir les grans promeses que s’han fet en nom seu, però el que sí que fa és donar pistes molt interessants. Tornem a l’arquetípic recipient amb dos gasos a punt de barrejar-se. La quantitat de molècules que hi pot haver en un recipient com aquest està més enllà de la imaginació. Si enlloc de tenir tantes molècules en tinguéssim tan sols unes poques i volguéssim preveure detalladament què farà cadascuna en el futur (aquí no val la pena entrar en les esmenes que es podrien fer a aquest plantejament des del punt de vista quàntic), trobaríem que és extremadament difícil. Si el que tenim és una infinitat de molècules, cadascuna tirant per la seva banda, una previsió així de detallada és impossible. Però, justament perquè n’hi ha tantes i cadascuna “tira per la seva banda”, el que sí que sabem és que seria impensable que no acabessin ben barrejades. Explicada d’una manera un pèl barroera, aquesta és la justificació de la Segona Llei. En qüestions de detall, tenir més elements fa les prediccions més difícils, però en qüestions globals les fa sovint més fàcils. I, si més no en el cas del recipient amb gasos, són justament aquests aspectes globals els que ens interessen: no en faríem res de conèixer el destí que li espera a cada molècula concreta, el que volem saber és si l’oxigen i el nitrogen es barrejaran o no. La ciència anomenada física estadística és capaç de fer prediccions molt acurades i, sovint, de molta utilitat pràctica, referides a sistemes amb molts graus de llibertat (és a dir, amb molts elements que “tiren cadascun per la seva banda”). Entre moltes altres coses, permet donar una interpretació microscòpica a la Segona Llei.

Una peculiaritat de l’ecologia en comparació amb altres ciències de la vida és que tracta amb sistemes integrats per grans quantitats d’elements no gaire coordinats entre si: qualsevol ecosistema digne d’anomenar-se’n inclou molts individus i sovint moltes espècies, i fins els factors abiòtics poden ser prou variats i heterogenis per fugir d’una descripció senzilla. A primera vista, això condemna l’ecologia a ser una “ciència” entre cometes, molt tova, destinada tan sols a seguir acumulant “estudis de casos” un darrera l’altre per sempre més. Però allò que la realitat ecològica té de difús i múltiple, que ara per ara és la debilitat de la ciència ecològica, bé en podria esdevenir la força. És una càrrega insuportable per a l’aproximació mecanicista a l'ecologia, per a l'intent de registrar cadascun dels engranatges d'un ecosistema com si fos una gran màquina, perquè no ho és pas. Alhora, és justament el que fa plausible que hi hagi regularitats generals que sorgeixin de manera comparable a les que estudia la física estadística, i que es puguin estudiar semblantment. No és que un ecosistema sigui com un gas, segurament està a mig camí entre el gas i la màquina (entre l'ordre i el caos, com dirien alguns estudiosos/es de la teoria de sistemes complexos). Tampoc és que els organismes o les espècies siguin com molècules: afortunadament, els éssers vius som molt més espavilats i amb molta més capacitat de burlar-nos de qualsevol expectativa teòrica... Per tant, no crec que es puguin esperar lleis tan exactes i generals com algunes de les que deriven de la física estadística. Tot i així, crec que la línia d'avanç que promet més rau en trobar resultats teòrics que siguin molt robustos en front de canvis en el detall dels models dels quals deriven, gràcies a efectes de "grans nombres", i a posar a prova empíricament si es compleixen a la pràctica. És clar que finalment tornem a l'"estudi de casos", però tindrem una idea clara i ben fonamentada de quins resultats podem esperar i com posar-los a prova, i si les regularitats més previsibles es van confirmant posaran un marc general per als estudis de detall, i justament ens indicaran en quins aspectes val la pena fer-ne. En aquest llibre presento un seguit de resultats teòrics obtinguts amb aquesta filosofia, no pas fent un simple transvasament de resultats de la física estadística, sinó treballant en termes genuïnament ecològics però inspirant-me força en aquesta branca de la física i abordant els problemes amb mentalitat de "grans nombres". A més mostro que aquests resultats passen el filtre del contrast amb les dades empíriques i que són útils per a gestió. Hi ha alguns treballs d'altres autors/es que, en certa manera, també s'acosten a aquesta manera de fer, com comento a la introducció. Qui es miri per sobre aquest treball, per exemple els gràfics que hi ha, es trobarà sovint amb variables que no tenen un significat evident d'entrada. No és estrany: els ecosistemes pertanyen a un nivell d'integració diferent al de la nostra vida quotidiana com a organismes individuals, i els conceptes rellevants hi són uns altres. Arribar-hi demana un esforç d'abstracció, però el que queda acaba tenint utilitats ben concretes.

Tornem al recipient amb gas. Mirem la concentració d'oxigen que tenim en cadascuna de les posicions d'un regle que va de banda a banda del recipient. Si trobem qualsevol diferència entre un punt i un altre, ben segur que serà transitòria, que finalment la concentració d'oxigen serà la mateixa a tot arreu, per una qüestió de grans nombres (a no ser, és clar, que mantinguem activament el gas fora d'equilibri).

Doncs ara fem el salt (sec. 5.4) i passem de molècules a espècies, i passem de la posició de cada molècula a l'abundància de cada espècie. Una diferència entre un i altre cas és que les espècies poden ocupar "nínxols ecològics" més o menys separats (alimentar-se de coses diferents, niar a diferents tipus d'indrets...) i, de resultes, tenir les seves abundàncies regulades per separat. Però una certa superposició entre nínxols sempre hi és, i a més deu ser molt gran en els casos en què trobem una gran quantitat d'espècies en ambients relativament homogenis, com ara entre els arbres de la selva tropical, els coralls dels tròpics o el fitoplàncton. Si més no en aquests casos, hi haurà un marge ampli d'abundàncies en què aquestes no s'explicaran per la "grandària" de cada nínxol, sinó que seran el resultat d'un cúmul de variacions imprevisibles en el nombre de naixements i morts en el si de cada espècie al llarg del temps. Dins d'aquest marge, és molt immediat aplicar una aproximació de física estadística semblant a la de l'oxigen en el nostre recipient. És clar que no podem traslladar directament els resultats d'un cas a l'altre sense introduir-hi algunes modificacions. Una molècula en el nostre recipient anirà fent moviments cap a dreta i cap a esquerra a mesura que vagi col.lisionant amb altres molècules, i aquesta dinàmica serà exactament igual estigui on estigui del recipient. Prenguem, en canvi, una espècie amb 100 exemplars. Si globalment es troba en un estat estacionari, l'abundància mitjana esperada al cap d'un temps seguirà sent de 100, però el valor real se'n desviarà. Posem que la desviació típica sigui de ±10. Perquè les coses fossin ben bé com amb la molècula, la desviació hauria de ser també de ±10 si l'abundància de partida és de 1000, 10000... Però no és així. Per exemple, si l'abundància de partida és de 10000, haurem d'esperar com a mínim una desviació de ±100. Aquesta diferència en la dinàmica té un impacte molt fort sobre la distribució d'abundàncies finalment esperable (subsec. 5.4.1). No serà pas uniforme com en el cas de les molècules, és a dir, no esperarem pas el mateix nombre d'espècies amb 100 individus que amb 200, 300 o 400. Per contra, el nombre esperat $s(n)$ d'espècies amb n exemplars tindrà la forma

$$s(n) = an^{-b} \qquad (0.1)$$

per a algunes constants a i b. Aquesta distribució estadística s'anomena distribució potencial (Ap. 2A), i veurem com ens reapareix un i altre cop, parlant de temes molt diferents a banda de les abundàncies d'espècies. El significat d'aquesta equació es pot percebre més fàcilment si, per exemple, prenem b=1, que és l'esperable en el cas més senzill, en què no hi ha influència de fluctuacions ambientals que tinguin efectes diferents per a diferents espècies (subsec. 5.4.1). Aleshores, l'Eq. 0.1 es redueix a

$$s(n) = \frac{a}{n} \qquad (0.2).$$

Per exemple, posem que a la comunitat que estudiem hi ha una espècie amb uns 100000 exemplars. Aleshores, n'hi haurà unes 10 que en tindran cap a 10000, unes 100 que en tindran cap a 1000, unes 1000 que en tindran cap a 100, unes 10000 que en tindran cap a 10... Val a de dir que, encara que tinguem juntes espècies amb diferents taxes de reproducció i mortalitat, aquest resultat seguirà essent vàlid.

Això d'utilitzar equacions basades en la difusió de partícules per explicar les distribucions d'abundàncies ja ho han fet alguns altres autors, però enlloc de quedar-se amb aquest plantejament tan senzill però robust i, a partir d'aquí, emprar altres eines, han volgut introduir en els models de difusió premisses més complicades i artificioses, com ara que la difusió estigui parcialment contrarestada per una tendència, igual d'intensa en totes les espècies, a assolir una certa abundància de referència, la mateixa per a totes. No hi ha dubte que aquesta mena de tendències hi són, però no són pas tan iguals per a cada espècie, i per tant en assumir-les així perdem la robustesa que dóna la màgia dels grans nombres (sec. 5.8). S'aplica una tècnica que prové de la física estadística, però sense la filosofia en què es basa la física estadística. Aquests treballs han aportat molt (ni més ni menys que el mètode de difusió), però trobo que avançarem encara més si intentem mantenir-nos dins dels límits d'allò que no depèn dels detalls. Això és el que he mirat de fer jo.

Comencem per una situació molt senzilla però que pot ser prou realista en alguns casos (subsec. 5.4.2). Posem que les espècies són tan semblants ecològicament que donarien lloc a l'Eq. 0.2 si no fos per un fet addicional que no podem ignorar: per molt que puguin variar al llarg del temps les abundàncies de cada espècie concreta, la biomassa total estarà molt més restringida i dependrà fortament dels recursos disponibles. Si fem una passa arriscada, que és assimilar biomassa a abundància (és arriscada per les possibles diferències de mida entre organismes de diferent espècies), aquesta restricció esdevé molt fàcil d'incorporar. En física estadística, fa temps que s'ha resolt la qüestió de què passa amb els conjunts de partícules que tendeixen a la màxima entropia subjecta a restriccions com ara que l'energia total del sistema estigui prefixada, i que s'ha suggerit que els mateixos procediments matemàtics podrien tenir aplicacions en altres camps. Doncs bé, aquest és un cas en què podem fer servir aquesta aproximació (subsec. 5.4.2, Ap. 5B), i el resultat que ens dóna és

$$s(n) = a\frac{e^{-fn}}{n} \qquad (0.3).$$

Per a valors realistes de f, el terme exponencial que ens ha aparegut en el denominador quasi no introdueix cap canvi en l'Eq. 0.2 per a un ampli rang d'abundàncies, però, a efectes pràctics, estableix un límit superior (d'una forma concreta). Quan vaig trobar l'Eq. 0.3 vaig pensar que em sonava d'alguna cosa... fins que me'n vaig adonar que era justament la primera distribució estadística que es va fer servir per ajustar a dades empíriques d'abundàncies! (que jo sàpiga). En concret, és la distribució anomenada "de les sèries logarítmiques", o simplement "logsèries", que Fisher va utilitzar ens anys 40 per aplicar a dos conjunts força amplis de dades de papallones nocturnes, l'un de Malàisia i l'altre d'Anglaterra.

De moment la situació que ens hem plantejat és massa senzilla, i per tant cal generalitzar-la. Per arribar a les equacions 0.2 i 0.3 hem suposat que totes les espècies reaccionaven semblantment a les fluctuacions ambientals. Si no és així, sinó que les espècies són prou diferents perquè allò que vagi bé per a una vagi malament per a una altra, ens quedem bàsicament amb l'Eq. 0.1, només que b es farà més gran que 1 (sense ultrapassar b=2) (subsec. 5.4.5). L'Eq. 0.1 sempre tindrà, és clar, alguna fita superior (algun límit superior), que en els casos més senzills tindrà la mateixa forma exponencial que en l'Eq. 0.3. És a dir, el que tenim de moment és una distribució potencial fitada, i aquest és un punt de partida molt bo per seguir generalitzant el model, ja que es tracta d'una distribució molt robusta. No és estrany que alguns autors/es hagin trobat que una distribució d'aquesta mena, amb algun b entre 1 i 2, és apropiada per ajustar les seves dades. Entre d'altres, és el cas d'un inventari molt extens d'arbres tropicals i també d'una mostra d'artròpodes (insectes i similars) d'un prat, que es veu que és la mostra d'organismes més extensa de la qual s'havia estudiat la distribució d'abundàncies fins ara (en aquest cas els autors no van identificar que es tractava d'una distribució potencial, però van fer servir una expressió matemàticament equivalent) (subsec. 5.4.5).

Vegem per què una distribució potencial és tan robusta, tan immune als detalls dels nostres models. Ja he esmentat que, per arribar-hi tal com hi hem arribat aquí, no fa falta que les diferent espècies tinguin les mateixes taxes de reproducció i de mortalitat. El que sí que hem assumit és que hi havia molta superposició en els seus nínxols ecològics, però hi haurà casos en què això serà totalment inversemblant. Posem que tenim una mostra de plàncton, en què hi ha fitoplàncton i zooplàncton barrejats. Podem aventurar que les diferents espècies de fitoplàncton tinguin molta superposició de nínxols entre si, i que les diferent espècies de zooplàncton també, però està clar que fitoplàncton i zooplàncton no se superposaran pas gaire. Doncs una de les propietats interessants de la distribució potencial és que, si tenim varis “gremis” (conjunts d’espècies amb grans semblances ecològiques) ben diferenciats i cadascun presenta una distribució potencial d’abundàncies, amb b no gaire diferent, el conjunt també tindrà una distribució potencial (subsec. 5.4.3). La forma exponencial que potser tenia la fita superior de la distribució per a cadascun dels gremis sí que es pot perdre fàcilment, però el més important és que seguirem tenint una distribució potencial fitada.

Amb això ja hem avançat molt, però encara estem amb models massa idealitzats. Ben segur que hi ha molts més factors que no hem tingut en compte i que són necessaris si realment volem reflectir la realitat. El més remarcable potser sigui la segregació de nínxols que hi pugui haver entre espècies individuals, no tan sols entre gremis amb moltes espècies cadascun (suposo que, un cop llegides les altres seccions, el lector/a no se sorprendria si fins i tot les grandàries dels nínxols presentessin una distribució potencial, però en aquesta possibilitat no hi entrarem). A vegades poden ser importants certes influències humanes, i moltes coses més. Podríem seguir afegint factors i complicant el model, però a llavors ja tornaríem a estar amb els problemes de sempre. Afortunadament, no fa falta. Encara que sembli increïble, si sabem que tenim un seguit de factors rellevants però sense un impacte gaire fort, podem incorporar-los en el model, així a l’engròs, sense necessitat de saber ni tan sols de quins factors estem parlant (la tècnica matemàtica que ho permet és el desenvolupament en sèries de Taylor) (subsec. 5.4.4, Ap. 5D). En fer-ho, la primera conclusió que n’extraiem és que, si la desviació respecte al model de partida és molt petita, en general es podrà expressar senzillament com a un canvi en el paràmetre b a l’Eq. 0.1 i, per tant, seguirem dins del domini de les distribucions potencials. La segona conclusió és que, si la desviació és petita però no tant, en general les abundàncies es podran ajustar bé a una distribució lognormal (la distribució tal que, en fer el logaritme de les dades, esdevé una distribució normal o “campana de Gauss”, en aquest cas espapçada per un extrem). Aquest resultat és molt significatiu: he esmentat que la distribució potencial fitada , i la logsèries en particular, s'han fet servir per ajustar les distribucions d'abundàncies, però resulta que la més emprada és justament la lognormal, des que, ja en els anys 40, en Preston va observar que anava bé per ajustar una gran quantitat de mostres, per a les quals la logsèries no sempre era adequada. La utilitat de la lognormal s'ha explicat fins ara amb certs raonaments estadístics elementals... però inaplicables en aquest context, com mostro en el llibre (Ap. 5E). Per contra, les troballes que he resumit suggereixen fermament que la lognormal funciona perquè és la distribució que va millor per ajustar dades provinents de distribucions semblants a la potencial però que no siguin ben bé la potencial. I encara hi ha un element més que hi quadra, amb aquesta interpretació. El propi Preston, que va introduir la lognormal en ecologia, va observar que sempre trobava una certa relació entre els paràmetres d'aquesta distribució quan mirava d'ajustar-la a dades d'abundàncies. D'aquesta versió restringida de la lognormal en va anomenar distribució canònica. Així com la lognormal semblava (només semblava) explicable fàcilment, aquest no era el cas de la canònica, que sempre ha quedat com a un misteri sense acabar de resoldre. Però el fet és que, si la lognormal ha de ser pròxima a una potencial fitada, tal com implica la meva proposta, justament ha de tenir les característiques de la distribució canònica (Ap. 5F).

Cal tenir en compte que en cap moment no tanquem la porta a que hi hagi factors addicionals que sí que tinguin molt impacte i, per tant, puguin donar qualsevol mena de resultat, encara que s'allunyi de les distribucions anteriors. No obstant, els elements que afavoreixen les distribucions de les quals he parlat són suficientment forts perquè a la pràctica siguin les primeres que cal prendre en consideració, i el que sabem fins ara d'ecologia suggereix que, en la majoria dels casos, no fa falta anar més enllà.

Aquesta història no s'acaba amb les distribucions d'abundàncies. Una de les regularitats ecològiques més freqüentment citades, i que sovint es veu com a una de les poques que realment tenen interès, és la relació entre la superfície A d'un territori i el nombre d'espècies S que s'hi troba, a la qual típicament s'atribueix la forma (ara també es coneixen alguns casos que se'n desvien significativament)

$$S = cA^z \qquad (0.4),$$

on c i z són constants, i z sempre és bastant menor que 1. És a dir, el nombre d'espècies augmenta amb la superfície menys que proporcionalment, i sovint en aquesta forma concreta. Hi ha molts mecanismes que poden prendre part en aquesta relació, però n'hi ha un d'especialment bàsic i inevitable. Encara que els organismes estiguessin col·locats en l'espai de manera totalment aleatòria, el nombre d'espècies que trobaríem aniria creixent a mesura que augmentéssim la superfície mostrejada. Em vaig preguntar què hauríem de trobar en aquest darrer cas si la distribució d'abundàncies és potencial (Eq. 0.1)... doncs justament l'Eq. 0.4, amb $z=b-1$ (subsec. 5.4.6). Les desviacions respecte a l'atzar es manifestaran primàriament fent desviar z d'aquest valor. Per tant, $\Delta z=z-(b-1)$ és una mesura del grau de contagi espacial entre organismes de les mateixes espècies. A la bibliografia es troba algun altre intent de relacionar la distribució d'abundàncies amb l'Eq. 0.4 que ha tingut força impacte, però que és clarament incorrecte (Ap. 5G).

En definitiva, gràcies a aquest seguit de raonaments basats en els "grans nombres", han adquirit sentit les principals distribucions d'abundàncies que hom ha emprat i també la principal relació biogeogràfica (l'Eq. 0.4). Un dels capítols importants de l'ecologia del segle XX ha cristal·litzat en una estructura lògica coherent, si més no de manera temptativa (Fig. 5.1).

El conèixer aquests aspectes tan bàsics de la relació entre dinàmica subjacent i patrons de diversitat ens dóna les eines de partida per esbrinar coses sobre l'una a partir de l'altra (sec. 5.5), i també un toc d'alerta (subsec. 5.5.2). En la bibliografia hi ha un seguit de models dinàmics molt concrets, la validesa dels quals pretén recolzar-se en la seva capacitat per explicar distribucions d'abundància observades empíricament. Els resultats d'aquest treball indiquen que una gran quantitat de models diferents donaran lloc al mateix resultat qualitatiu i, per tant, el trobar-lo no és pas un argument gaire fort a favor del model que es tracti.

Per complementar aquestes troballes teòriques, vaig analitzar dos extensos inventaris de fitoplàncton marí (sec. 5.6) publicats per en Ramon Margalef, un del Mediterrani i un del Carib (ambdós ultrapassen de llarg la grandària de la mostra d'artròpodes que abans he esmentat i que es considera la més gran mai analitzada). En els dos casos, la distribució és molt pròxima a una potencial fitada amb $b \approx 1.2$ ("$\approx$" significa "aproximadament igual a") (Fig. 5.2, 5.3). Però el més interessant es troba en estudiar per separat els dos grups més àmpliament representats, les diatomees i les dinoflagel·lades. Cal tenir en compte que diatomees i dinoflagel·lades no són tan sols dos grans grups taxonòmics, sinó també dos grans grups ecològics. Les diatomees se solen considerar més “oportunistes”, mentre que les dinoflagel·lades són més característiques d’estadis més madurs de la successió, en què se suposa que l’ecosistema està més estructurat. Podem assumir que les diatomees se semblem més a un únic gremi amb molta superposició ecològica que no pas les dinoflagel·lades. Doncs resulta que, sigui o no per aquest motiu (si una cosa deixen clara els raonaments anteriors és que, en aquest terreny, s’ha de tenir molta cura amb les interpretacions), en ambdós casos les diatomees mostren $b \approx 1.0$ i les dinoflagel·lades $b \approx 1.45$. En concret, en el cas del Mediterrani les diatomees s’ajusten amb una gran exactitud a la distribució logsèries (Fig. 5.4, 5.5).

L’aplicació més immediata dels coneixements sobre patrons de diversitat rau en la seva quantificació i seguiment o “monitoratge” amb propòsits de conservació (sec. 5.7). Diversitat és un concepte que té a veure tant amb el nombre d’espècies (o varietats, o grups...) com amb el fet que tinguin una representació més o menys desigual. No és un concepte prou concret ni unidimensional com per correspondre’s clarament amb una mesura numèrica determinada que permeti comparacions directes, i el resultat és que n’hi ha moltes de disponibles, i ben poca orientació a l’hora d’usar-les, i no obstant s’empren abastament. D’aquest seguit d’índexs de diversitat hi ha dues qüestions a avaluar (sec. 5.7): el seu interès conceptual i la seva solvència estadística. Els problemes estadístics sorgeixen del fet que les espècies poc abundants només comencen a estar representades a les mostres a mesura que aquestes es fan grans. Per tant, a mesura que creixen, la panoràmica que obtenim no és tan sols més acurada, sinó també més àmplia, i enlloc de limitar-nos a reduir l’error en l’estimació dels índexs ens podem trobar que el valor a estimar vagi canviant. Justament el fet que les distribucions d’abundàncies siguin pròximes a potencials magnifica aquest inconvenient, perquè, dins del rang en què efectivament ho siguin, per cada espècie amb una certa abundància en tindrem moltes més amb molta menys abundància, i per molt que augmentem la mostra el problema no decreixerà. En estudiar la conducta dels principals índexs en front de distribucions d’abundàncies realistes (sec. 5.7), he pogut constatar que alguns dels més emprats (com els índexs a i Q, l’índex anomenat de Margalef i el de Mehinick, i també l’”índex d’equitabilitat” més habitual) acaben anant-se’n cap a zero o infinit en augmentar la mostra, a vegades després d’haver derivat en direcció contrària... Per tant, no serveixen, fora de casos molt concrets (per exemple, l'a és útil si i només si tenim estrictament una logsèries). Altres índexs, com el de Shannon (que de fet va ser proposat com a tal per en Margalef i és el més emprat per aquest autor) o el de Simpson, acaben convergint cap a un valor concret, però a un ritme lentíssim. Per exemple, en el cas de les diatomees mediterrànies (Fig. 5.7) cal identificar milers d'exemplars per poder dir que hem assolit el valor definitiu de qualsevol d'aquests dos índexs, i en altres casos la cosa és pitjor. Pel que fa a la simple estimació de quantes espècies hi ha en un lloc a partir d'una mostra gaire parcial, no té fiabilitat quan ens les havem amb distribucions pròximes a la potencial, a no ser que fem assumpcions força arriscades. En algun cas, no obstant, pot estar justificat, i en algun cas podria valer la pena fer servir l'índex de Shannon o el de Simpson, però aleshores hauríem de posar un rang de valors enlloc d'un valor concret (detalls a la subsec. 5.7.1), a no ser que disposem de mostres molt grans. D'altra banda, tenint en compte que les distribucions d'abundàncies són tan pròximes a una potencial, trobo que el que tindria més sentit és estimar directament els dos paràmetres de la distribució potencial, és a dir, b i a a l'Eq. 0.1 (subsec. 5.5.3). Millor dit, no estrictament a, que depèn de la grandària de la mostra, sinó la mesura invariant que determina a, i que jo anomeno D_b. Aleshores ens quedem amb dos números enlloc d'un, però amb un contingut força clar. Es pot dir que la diversitat augmenta en augmentar b i en augmentar D_b. b està determinat pel tipus de dinàmica subjacent i es pot fer servir com a mesura de la distància respecte al patró esperable en models senzills amb una gran semblança ecològica entre les espècies, que correspon a b=1. Quan volem comparar mostres amb una b semblant, D_b és l'índex de diversitat idoni. En el llibre explico com aplicar-lo a la tria d'àrees a conservar en els casos, molt limitats, en què tingui sentit aplicar els plantejaments semi-automatitzats de la "planificació sistemàtica de la conservació" (subsec. 5.7.2).

Quan els "grans nombres" s'encongeixen

A hores d'ara deu haver quedat clar que la gran esperança que li veig a l'ecologia com a ciència digna d'anomenar-se'n és la possibilitat, que ja hem començat a concretar, de trobar regularitats basades en la força estadística dels grans nombres, grans nombres d'individus, grans nombres d'espècies, i altres possibles "grans nombres" d'interès, igual que la física estadística ha trobat lleis potents que permeten preveure la dinàmica de conjunts de partícules, com ara les molècules de gas en un recipient, amb la condició indispensable que n'hi hagi moltes. Ho he il·lustrat amb el seguit de troballes sobre patrons de diversitat que tot just he resumit. Però hi ha una diferència inquietant entre l'ecòleg/oga que estudia la diversitat i el físic/a que estudia un gas en un recipient: el físic/a no forma part del gas ni està constatant com el seu "gran nombre" de molècules es fa cada cop més petit! Al mateix temps que anem descobrint aspectes de la natura que són predictibles gràcies a un efecte de grans nombres, ens podem trobar que aquesta predictibilitat s'esvaeix a mesura que el "gran nombre" s'enxiqueix irreversiblement. És essencial que ens fem una idea de les conseqüències d'aquest canvi i, és clar, que l'aturem! Hi ha molts motius per conservar la diversitat apart del seu presumible efecte estabilitzador sobre la dinàmica dels ecosistemes, però aquest efecte, si es confirma encara més i es troba que és prou fort com per tenir un impacte important sobre els humans, és el que pot arribar a resultar més convincent de cara a què es prenguin mesures serioses de conservació.

Recentment s'ha parat molta atenció a l'efecte de la diversitat sobre l'estabilitat (cap. 4). És un dels pocs temes en què la ciència ecològica ara dominant es posa una qüestió de fonament, i també un dels pocs temes en què ha penetrat un plantejament rudimentari de "grans nombres", tot i que s'intenta encaixar dins dels esquemes mecanicistes habituals.

La màxima atenció s'ha centrat en un seguit d'experiments amb ecosistemes artificials o artificialitzats (parcel·les de prat, mesocosmos amb espècies terrestres, d'altres amb invertebrats intermareals, microcosmos amb espècies d'aigua dolça) en els quals s'ha manipulat directament el nombre d'espècies i se n'ha quantificat l'efecte sobre la variabilitat en paràmetres globals com ara la producció, la biomassa o el flux de CO_2. El resultat general que sembla emergir és que la variabilitat en aquesta mena de paràmetres decreix amb la diversitat. En aquesta secció, per "diversitat" em referiré bàsicament a nombre d'espècies (o varietats, grups, etc.), com s'ha fet habitual en la bibliografia sobre aquest tema. La mesura de variabilitat que s'ha fet servir més és el coeficient de variació (*CV*) de la biomassa (el *CV* és el tant per cent que la desviació típica representa en relació a la mitjana).

Aquests experiments han motivat una sèrie de treballs teòrics per intentar explicar-ne els resultats. Una de les aportacions que ha tingut més impacte proposa que són conseqüència directa d'una regla que pertany a la teoria dels grans nombres. És allò que els autors d'aquesta aportació anomenen "efecte mitjana", que no és més que el decreixement de la variància quan es fa la mitjana entre diferents variables, fins i tot si estan correlacionades (de fet, aquest "efecte mitjana" ja s'havia posat en relació amb la diversitat en els anys 60, cosa que els autors/es que ara treballen en aquest camp semblen desconèixer) (sec. 4.1). Un dels principals investigadors en aquesta àrea, en David Tilman, acceptà la importància de l'efecte mitjana, però replicà que també hi ha altres factors que intervenen i, per tant, no és suficient per esperar un efecte positiu de la diversitat sobre l'estabilitat. En concret, assenyalà que, en variar la diversitat, també pot canviar el *CV* de cada espècie individual, així com les covariàncies entre espècies (és a dir, el grau en què estan sincronitzades les variacions de diferents espècies). De resultes, aquest intent de generalització teòrica ha quedat reemplaçat per un seguit d'estudis de models mecanicistes de comunitats senzilles en les quals l'única interacció entre espècies és la competència pels recursos. Normalment donen un resultat favorable a l'efecte estabilitzador, però no pas sempre. En tots aquests models, un augment de diversitat produeix un augment en el *CV* de cada espècie concreta, ja que quan més espècies hi ha més extenses són les superposicions entre nínxols i cada abundància concreta passa a estar menys regulada. L'estructura de covariàncies també canvia. A hores d'ara no es veu evident que la conjunció d'aquests factors hagi de donar un resultat universal.

En el llibre mostro que totes aquestes provatures de models es podrien haver estalviat si s'haguessin fet uns pocs raonaments genèrics. Sota condicions molt generals, podem esperar que la diversitat estabilitzi. En la subsec. 4.2.2 mostro amb detall que els models que semblen contradir aquest resultat és perquè tenen assumpcions totalment inversemblants. El motiu pel qual podem esperar que la diversitat estabilitzi es pot veure amb un senzill experiment mental (subsec. 4.2.2). Posem que tenim un ecosistema amb un cert nombre d'espècies (per simplicitat, les considerarem asexuals i sense diversitat intraespecífica, cosa que no afecta les conclusions). Repartim els individus de cada espècie en un seguit de subpoblacions o "pseudoespècies", i seguim-ne la dinàmica per separat (per assegurar-ne la persistència, sempre que una subpoblació s'extingeixi podem triar a l'atzar un individu de la mateixa espècie i assignar-lo a la subpoblació). Formalment haurem augmentat el nombre d'"espècies", que de fet són pseudoespècies, i cada pseudoespècie tindrà un *CV* molt elevat, ja que tindrà una superposició total de nínxols amb altres pseudoespècies. No obstant, el *CV* del conjunt no pot haver canviat, ja que les transformacions que hem fet han estat tan sols formals, no pas reals. Per tant, paral·lelament a l'augment en el *CV* de cada pseudoespècie hi haurà hagut un canvi en l'estructura de les covariàncies que l'haurà compensat perfectament. Ara transformem les pseudoespècies en espècies. Aquesta operació consisteix a donar a cada pseudoespècie un valor diferent en els paràmetres que defineixen la seva resposta en front de diferents situacions. En principi, això hauria de decréixer les correlacions entre les biomasses de cada pseudoespècie (ara, espècie) i, de resultes, decréixer el *CV* de la biomassa conjunta. Com a conclusió, l'expectativa d'una reducció en el *CV* global esdevé més consistent si, enlloc de veure la diversificació com a un augment en el nombre de variables de les quals fem la mitjana (que també es correcte), la veiem com a una minva en la correlació entre variables.

Cal remarcar que l'efecte estabilitzador es fonamenta en les respostes diferencials de cada espècie a factors ambientals o interns al propi ecosistema, i que hem vist en la secció anterior que aquestes diferències tenen un impacte sobre els patrons de diversitat que observem (en particular, sobre el paràmetre b en l'Eq. 0.1). Per exemple, tenim més base per esperar un efecte estabilitzador de la diversitat de dinoflagel·lades que no pas de la diversitat de diatomees, amb totes les reserves que he expressat sobre la interpretació d'aquests patrons.

En principi, els raonaments anteriors no serveixen tan sols per a comunitats competitives senzilles, sinó també per a sistemes amb tipus més variats d'interaccions, fins i tot per a sistemes arbitraris, sempre i quan tinguin moltes espècies. S'ha de matisar que la diversificació pot tenir efectes no trivials sobre algunes interaccions no competitives, però les més previsibles a priori no van pas en contra de l'efecte estabilitzador (subsec. 4.2.2).

En el cas de les comunitats competitives senzilles, l'efecte estabilitzador de la diversitat es pot esperar en tot el rang possible de nombres d'espècies. En canvi, quan tenim xarxes complicades d'interaccions tan sols és possible generalitzar en el límit dels grans nombres (subsec. 4.2.3). És ben conegut que un sistema amb poques variables pot presentar dinàmiques de tota mena: pot tendir a un certs valors estables de les variables, amb més o menys intensitat, pot mostrar cicles periòdics, o fins i tot les variacions imprevisibles pròpies de la dinàmica anomenada "caos determinista". Se sap que en ecosistemes senzills es pot presentar tot aquest ventall de possibilitats, i seria inútil intentar trobar alguna relació general de la variabilitat amb el nombre d'espècies. Tan sols més enllà d'aquell nombre de variables que en física estadística s'anomena "límit termodinàmic", en què l'estadística esdevé més important que la dinàmica, podem esperar una relació general d'aquest tipus. Amb això tenim, de moment, dues formes en què podem esperar que la diversitat estabilitzi: d'una banda, tenim una relació regular entre diversitat i variabilitat, esperable tan sols per damunt del límit termodinàmic (fora del cas concret, però important, de les comunitats competitives senzilles, en què es pot esperar per a tot el rang de diversitats possibles); d'altra banda, el fet que fa falta una certa diversitat per mantenir-se per damunt d'aquest límit termodinàmic. Si anem a parar per sota d'aquest límit, trobem un problema d'estabilitat força pitjor que el simple augment regular de variabilitat, ja que el sistema esdevé tan fràgil que pot canviar radicalment la seva dinàmica per qualsevol petita modificació en la composició d'espècies o en altres atributs. La necessitat d'estar més enllà del límit termodinàmic per trobar regularitats generals no ha estat copsada pels ecòlegs/ogues actuals, que cerquen relacions regulars entre diversitat i estabilitat fins i tot en el límit de molt poques espècies, sense restringir-se a l'únic cas en què això té sentit, que és quan tan sols hi ha interaccions competitives. El més sorprenent és que de vegades semblen trobar-les! Aquest és el cas d'un dels experiments més coneguts sobre diversitat i estabilitat, que es va fer sobre microcosmos amb plàncton d'aigua dolça. N'he reanalitzat les dades (Ap. 4B), i he trobat que aquesta aparença de resposta regular no és més que un resultat espuri de la forma en què els autors/es n'havien fet el tractament. En realitat, les dades procedents d'aquest experiment estan d'acord amb les expectatives que aquí manifesto.

A més de l'efecte immediat sobre la variabilitat en els paràmetres globals, la diversitat és fonamental per a un altre aspecte de l'estabilitat, de manera encara molt més evident: forneix el recurs bàsic per al procés històric d'evolució d'espècies i ecosistemes, fa possible l'adaptabilitat (subsec. 4.3.1). Podem suposar que la diversitat també contribueix a l'adaptabilitat d'una manera més subtil (subsec. 4.3.2). Si acceptem que, en termes generals, a més espècies (o varietats, etc.) més superposició de nínxols, la biomassa de cada espècie estarà menys estretament regulada i el sistema tindrà globalment més flexibilitat per adaptar-se als canvis. En un sistema complex com la Biosfera, tots els detalls canvien, tard o d'hora, i això és bàsic per a la pròpia persistència del sistema. La modalitat d'estabilitat que hi té més rellevància és la que podem definir senzillament com a la capacitat d'un sistema complex per seguir sent un sistema complex, i això sembla exigir que els detalls concrets hi siguin modificables.

No podem abandonar el tema de la diversitat i l'estabilitat sense donar una ullada a les grans discussions que n'hi havia hagut fa dècades, en el decurs de les quals es van plantejar qüestions importants que encara no s'han resolt (sec. 4.1 i 4.4). A mitjans del segle XX ja hi havia la idea general entre els ecòlegs/ogues que la diversitat estabilitzava. Molts dels arguments que es feien servir tenien a veure amb el que ara s'anomena "efecte mitjana", només que no tan enfocat als paràmetres globals en tant que mitjana de les aportacions de cada espècie, com a la capacitat de cada espècie de reduir la seva pròpia variabilitat a base d'interaccionar amb moltes altres espècies, fent-ne la "mitjana". La diversitat es veia com a una condició per poder tenir xarxes tròfiques amb moltes connexions, en les quals els depredadors estabilitzarien llurs poblacions a base de fer la mitjana entre vàries preses, les preses farien la mitjana entre varis depredadors, etc. Hom esperava fluctuacions més violentes en sistemes amb poques espècies i xarxes tròfiques senzilles, i així semblaven confirmar-ho les oscil· lacions poblacionals de certes espècies boreals. La relació entre diversitat i estabilitat es va començar a usar com a un argument conservacionista. Però tot plegat va quedar estroncat amb una aportació molt important d'en Robert May. Aquest autor va estudiar les propietats de certs models matemàtics molt senzills d'ecosistemes, amb connexions entre espècies col· locades a l'atzar (connexions de tot tipus, no tan sols alimentàries), i va observar que hi havia un llindar en el producte del nombre d'espècies per la proporció de connexions realitzades d'entre les possibles (aquest producte no és més que el nombre de connexions per espècie) tal que, en excedir-lo, el sistema quasi sempre passava d'estable a inestable. Tot i que en May va ser prudent sobre la relació entre aquesta troballa i el món real, la idea que es va imposar és que la diversitat, per si mateixa, el que feia era desestabilitzar en comptes d'estabilitzar. Durant dues dècades, això va desencoratjar tant la recerca en aquest camp com l'ús d'arguments ecològics per a la conservació de la biodiversitat, la qual cosa pot haver tingut conseqüències molt greus. Les coses no van canviar fins que no van arribar els experiments que semblaven indicar que la diversitat sí que estabilitzava, i va fer falta trobar la manera de conciliar aquest resultat amb el dogma teòric aleshores imperant. Tilman va mostrar que el problema estava en què s'estaven barrejant diferents facetes de l'estabilitat. Per a May, un sistema era estable quan totes i cadascuna de les espècies recuperaven la seva població inicial després que el sistema hagués estat lleugerament pertorbat. Per tant, es tractava d'una estabilitat a nivell d'espècie, no pas a nivell de paràmetres globals. El propi May havia assenyalat que els seus resultats no implicaven un efecte negatiu de la diversitat sobre els paràmetres globals, però aquesta apreciació va passar quasi desapercebuda fins que en Tilman la va desenterrar. Però amb això no queda tot resolt. Els grans autors anteriors a May, com E. P. Odum o Robert MacArthur, esperaven que un gran nombre d'espècies i de connexions entre elles tingués un efecte estabilitzador sobre cadascuna de les espècies, no tan sols sobre el conjunt. Els raonaments que ho recolzaven eren ben clars, i ara els podem expressar amb l'"efecte mitjana". A llavors, per què els models matemàtics d'en May no satisfeien les expectatives d'aquests autors? Que jo sàpiga, abans de les meves recerques no s'havia trobat el motiu d'aquesta contradicció. Abans de res, s'ha de dir que no és cert que dels resultats d'en May se'n pugui extreure que un gran nombre de connexions en un sistema sempre desestabilitzi: en el llibre mostro (subsec. 4.4.2) que diferents plantejaments duen a diferents conclusions, i fins dono una recepta per construir sistemes simulats que necessiten moltes connexions per ser estables (Ap. 4C). Però no crec que estigui aquí la clau, ja que aquestes variacions semblen allunyar-se encara més de la realitat ecològica que el propi plantejament d'en May. Segons el meu parer, hi ha dues claus en aquest enigma (subsec. 4.4.3). Una de les claus està en que l'efecte mitjana coexisteix amb un altre efecte que va en direcció contrària, que és l'efecte de superposició de nínxols. Quan augmentem el nombre de connexions en un sistema simulat, cada espècie de depredador tindrà més espècies de preses entre les quals fer la mitjana, però moltes d'aquestes espècies de preses les haurà de compartir amb altres espècies de depredadors perquè les preses també hauran augmentat el seu nombre de connexions, de manera que els nínxols de diferents espècies de depredadors se superposaran i les seves poblacions passaran a estar menys estretament regulades. Semblantment passarà amb les pròpies preses, així com amb simbionts, paràsits, competidors, comensals, etc. La prova d'aquest punt està en el fet que, si extreiem "cirúrgicament" dels models d'en May l'efecte de la superposició de nínxols, el resultat trobat per en May es capgira totalment. És a dir que, tot i l'"efecte mitjana", podem esperar que la superposició de nínxols produïda per un major nombre de connexions per espècie decreixi la intensitat de la regulació sobre cada espècie concreta suficientment perquè la seva població acusi unes fluctuacions cada cop més àmplies, fins al punt en què simplement no tingui capacitat per retornar a la població de referència o fins i tot tendeixi a allunyar-se'n (la qual cosa farà que s'extingeixi o provoqui una reorganització del sistema). Una segona clau està, de nou, en la faceta de l'estabilitat de la qual estiguem parlant. Si deixem apart les espècies que tendeixin a allunyar-se de la seva "població de referència" enlloc d'acostar-s'hi (que només poden existir de manera molt transitòria), l'efecte d'unes superposicions importants de nínxols és que hi hagi uns rangs amplis d'abundàncies en les quals les espècies derivin amunt i avall sense que se'n pugui discernir clarament les poblacions de referència. Un cop estem en aquest límit, la capacitat que cada espècie tingui per fer la mitjana entre d'altres espècies no farà pas que deixi de derivar, però sí que farà que rebi impactes més febles i, per tant, les oscil·lacions siguin més lentes i graduals. Per tant, els resultats d'en May no són pas incompatibles amb la idea de n'E. P. Odum que en sistemes amb moltes connexions no hi hagi d'haver oscil·lacions tan "violentes". Val a dir que aquesta darrera faceta de l'estabilitat va associada amb una major predictibilitat en el subministrament d'aliments i en altres aspectes a nivell dels organismes individuals, la qual cosa té una gran significació biològica.

Amb aquestes troballes teòriques no es resol la qüestió de la diversitat i l'estabilitat, perquè per a això fa falta molta més investigació empírica. Però espero haver contribuït apreciablement a fer una passa prèvia indispensable, com és resoldre les confusions d'ordre conceptual que entorpien la recerca en aquest camp. S'ha de remarcar que, de moment, tot suggereix que la diversitat té un impacte positiu sobre l'estabilitat dels ecosistemes.

Hi ha dos aspectes més de gran importància pràctica relacionats amb diversitat, xarxes d'interaccions i estabilitat per als quals dono alguna idea temptativa en el treball: la relació entre l'estructura de les xarxes i la dinàmica d'extincions (subsec. 4.4.4), i l'efecte de la diversitat de plantes sobre la propagació de plagues i patògens en agricultura i en ecosistemes naturals (sec. 4.5). Però per fer-ne cinc cèntims calen abans alguns conceptes que explico en les pròximes seccions.

Criticalitat

Dins del marc de l'enfocament "físic estadístic" de l'ecologia, les prediccions més frapants ara per ara semblen provenir d'aplicar-hi la idea de criticalitat, línia que s'ha encetat des de la pròpia literatura física i a la qual afegeixo un seguit de contribucions en aquest llibre. Aplicar el concepte de criticalitat implica ni més ni menys que dur l’analogia del recipient ple de molècules fins a l’extrem d’acceptar que els ecosistemes poden patir transicions de fase (com el pas de gas a líquid o de líquid a sòlid, per exemple).

Quan he fet alguna xerrada explicant aquestes idees, he començat per convidar el públic a jugar al joc de l’esquirol (subsec. 3.2.1). Aquest joc s’inspira en la llegenda que diu que, en uns altres temps, la Península Ibèrica estava tan forestada que un esquirol podia recórrer-la de punta a punta sense baixar dels arbres. Doncs cada jugador/a fa d’esquirol en un “bosc” abstracte i simplificat, que és com un tauler d’escacs en què cada casella pot estar ocupada o no per un arbre. L’esquirol entra per un costat del tauler i ha d’arribar al cantó oposat, però només pot passar per caselles contigües que continguin arbres (sense incloure desplaçaments en diagonal, tot i que això només tindria conseqüències de tipus quantitatiu, no pas qualitatiu). En les xerrades demano al públic que faci dos equips i els dono a cadascun deu taulers de 30x30 caselles, cadascun amb una configuració diferent dels arbres. Els arbres hi estan col· locats a l’atzar, de manera que cada casella té una probabilitat aproximada de 0.6 que n’hi toqui un. Aleshores, resulta que l’equip B troba camins per creuar 8 dels 10 taulers, mentre que l’equip A només en troba la meitat. No és que a l’equip B hi hagi millors esquirols: és que hi ha una diferència subtil entre els dos conjunts de taulers. El truc està en l’“aproximada” de la “probabilitat aproximada de 0.6”. Per als taulers de l’equip A faig servir una probabilitat r de 0.57, mentre que per a l’equip B faig servir r =0.61. Sorprenentment, amb aquestes centèssimes n’hi ha prou per produir aquesta gran diferència en els resultats. I la diferència hauria estat molt més radical si el tauler hagués estat més gran, és a dir, si hagués estat en el límit dels “grans nombres” (ara no pas d’espècies ni de molècules, sinó d’“arbres”). Aleshores, els esquirols de l’equip B haurien aconseguit creuar tots els taulers, mentre que els de l’equip A no n’haurien aconseguit creuar ni un de sol. Aquest canvi dràstic és tècnicament una transició de fase, com ho és la congelació de l‘aigua a 0^oC (en condicions estàndard). En aquest cas també hi ha un llindar molt concret en el qual té lloc la transició, l’anomenat “punt crític”: la probabilitat “crítica” és de r_c=0.5927...

Tot i que de moment estem a un nivell de “model de joguina” molt allunyat de la realitat, immediatament suggereix aplicacions. Els ecòlegs/ogues del paisatge fa temps que s’hi han fixat en plantejar-se la qüestió de si els indrets que serveixen d’hàbitat potencial per a una espècie estan prou interconnectats perquè l’espècie pugui persistir (subsec. 6.3.2).

També fa temps que alguns físics/ques s’han plantejat que aquest model pugui servir per estudiar la propagació del foc (subsec. 3.2.1). Posem que cau un “llamp” en algun punt del tauler triat a l’atzar, i que si hi ha un arbre el crema i es propaga a tots els arbres contigus, i d’aquests als contigus, etc., fins que no troba cap via per seguir. Està clar que aquest model és totalment irreal, però veurem que haurà estat bo com a punt de partida. En aquest cas, la transició de fase es manifesta pel fet que, per damunt del punt crític ($r > r_c$), hi ha una certa probabilitat que el foc mai no deixi de trobar vies per on propagar-se, de manera que es mantingui indefinidament (millor dit, podria arribar a mantenir-se indefinidament si el tauler fos infinit). En canvi, per sota del punt crític ($r < r_c$) aquesta possibilitat no hi és. I no cal estar gaire per sota per poder garantir, a més, que el nombre de cel·les que cremarà serà petit.

Estudiem amb més detall com canvia la grandària mitjana $\bar{s}$ dels “focs” en modificar r. Poc o molt, el lector/a deu estar familiaritzat/da amb la idea de creixement exponencial. $\bar{s}$ creixeria de manera exponencial amb r si, cada cop que incrementéssim r en una certa mesura, $\bar{s}$ es dupliqués. Doncs quan estem a la vora d’una transició de fase com aquesta, la resposta es fa més que exponencial. Quan més gran és r més gran és el factor pel qual es multiplica $\bar{s}$ com a resposta a un augment donat de r, seguint un patró determinat, fins que ultrapassem el punt crític. Aleshores, podem dir que $\bar{s}$ s’ha fet infinita, perquè hi ha focs que poden mantenir-se indefinidament.

Les coses més sorprenents succeeixen si estem just en el punt crític. D’entrada, les propietats del punt crític no semblen tenir més interès que l’acadèmic, perquè el punt crític és tan sols un punt d’entre una infinitat. Però, com veurem després, hi ha bons motius pels quals el punt crític sí que té molta importància pràctica, i per tant l’estudiarem una mica. Una propietat universal dels sistemes en estat crític és el seu caràcter “escalant”, amb patrons que es repeteixen a totes les escales. El fet de ser escalant és típic de les formes que en Benoît Mandelbrot anomena fractals. Hi ha molts elements de la natura que són fractals: muntanyes, núvols, línies de costa, arxipièlags, grups de llacs, rius, llamps, arbres, bronquis, bròquils... Qualsevol muntanya té el seu propi relleu: hi podem distingir muntanyes més petites i, formant part d’aquestes, muntanyes encara més petites... Un arbre gran té branques, i cada branca té les seves pròpies branques, i aquestes en tenen d'altres... Dins d'uns límits més o menys amplis, trobem que en aquests casos el tot se sembla a les parts, que no hi ha cap escala característica, perquè tant si mirem les coses a una escala petita com a una escala gran veiem més o menys el mateix (no tots són pròpiament escalants, però van bé per il·lustrar el concepte). En el cas del "bosc" simulat que estigui just en el punt crític, el caràcter escalant se li pot posar al descobert de vàries maneres, però n'hi ha una que ens interessa especialment: les grandàries s dels incendis. Resulta que, en el punt crític, les grandàries tenen una distribució estadística potencial, igual que la que havíem vist amb les abundàncies de les espècies:

$$f(s) = as^{-b} \qquad (0.5),$$

on $f(s)$ és la densitat de probabilitat de s (una mesura proporcional al nombre d'incendis de grandària s) i a i b són constants. La distribució potencial és una distribució escalant, sense cap valor particular de s que tingui més significació que els altres. Per exemple, sabem que per cada incendi d'una certa mida n'hi haurà aproximadament 2^b amb la meitat de superfície, per cadascun d'aquests n'hi haurà uns 2^b que medeixin la meitat, i així successivament, cobrint la major part del marge des de la mida del tauler fins a la mida de la cel·la, siguin les que siguin.

Hi ha alguna "connexió profunda" entre aquests incendis simulats i les troballes anteriors sobre les abundàncies d'espècies, que justifiquin aquesta coincidència de patrons? Doncs sí (sec. 5.3), i es diu criticalitat. Posem que introduïm un o uns pocs individus d'una espècie (asexual, per simplificar) en un indret determinat. Cada individu donarà lloc, com a mitjana, a m individus a la següent generació. "Com a mitjana" perquè no tots els individus deixaran el mateix nombre de descendents i, per tant, aquest nombre és una variable aleatòria. Doncs bé, en variar m tenim una transició molt semblant a la del foc. Per a $m<1$, és segur que l'espècie s'extingirà (encara que, si té sort, podrà fer la viu-viu durant una temporadeta). Si $m>1$ no podem pas garantir que no s'extingeixi, perquè partim d'uns pocs individus i pot ser que, per mala sort, no deixin gens de descendència, però també hi ha la possibilitat que l'espècie s'expandeixi "per sempre" o, millor dit, fins que sigui tan abundant que les coses canviïn perquè ha entrat en joc algun mecanisme de regulació (aquest cas és semblant al del foc, que també es "reprodueix", amb la diferència que l'espècie vivent no sempre està tan estretament condicionada com el foc pel caràcter bidimensional del "tauler", cosa que fa la dinàmica del foc menys intuïtiva). En cas que hi hagi una forta superposició de nínxols amb altres espècies, la dinàmica a llarg termini seguirà dominada per les fluctuacions aleatòries, malgrat els mecanismes de regulació (que seran compartits), però inevitablement tindrà $m=1$. Per tant, d'una manera trivial, hem anat a parar a un estat crític, i ens trobem amb la signatura de la criticalitat, que és l'absència d'escales característiques: distribució d'abundàncies escalant, relació espècies-àrea escalant... (cap. 5) i també una flexibilitat justa que deu facilitar l'adaptabilitat (subsec. 4.3.2). Sense haver-ho fet explícit, ja feia estona que parlàvem sobre criticalitat.

Actualment hi ha un gran interès en uns models que, suposadament, tendeixen a assolir un estat crític com a conseqüència de la seva pròpia dinàmica, no pas d'aquesta forma tan elemental que acabo d'explicar, sinó per uns altres mecanismes que s'engloben sota el terme de "criticalitat autoorganitzada" i que s'ha plantejat per explicar l'abundància de formes fractals en el món en què vivim (des d'ara ens referirem a aquest fenomen amb les sigles SOC, de l'anglès "self-organized criticality"). Com explico en la propera secció, en aquest llibre dono bons motius per pensar que aquest interès no és gens exagerat, ja que la SOC sembla tenir una extraordinària importància ecològica (com la té en altres camps). Ara bé, també dono evidència que, tot i que hi ha una forta relació entre SOC i criticalitat, en realitat els sistemes SOC no estan en un estat crític! (subsec. 3.4.3).

Recapitulant una mica, tenim tres motius per interessar-nos en la criticalitat des de l'ecologia:

- Perquè hi pot haver sistemes que, en ser manipulats des de l'exterior (per exemple, per una intervenció humana) sofreixin un canvi de fase, passant per un estat crític.
- Perquè hi pot haver sistemes amb "criticalitat autoorganitzada" (SOC), que té una estreta relació amb la criticalitat... tot i que no és criticalitat.
- Perquè hi pot haver sistemes que, d'una manera força trivial que no té relació amb la SOC, tendeixin per si mateixos a un estat crític (possible cas dels patrons de diversitat).

Al llarg d'aquest llibre dono evidència que les tres possibilitats es compleixen i són ben importants per als ecosistemes.

"Criticalitat" autoorganitzada

En parlar sobre “somnis i desenganys termodinàmics” havia comentat la constatació d’en Margalef que, en els sistemes complexos com els ecosistemes, hi ha una asimetria temporal entre els períodes en què la complexitat s’acumula gradualment (successió, si parlem d’ecosistemes) i els episodis puntuals de destrucció ràpida. Margalef va completar aquest esquema amb una observació misteriosa: que sol haver-hi una gran quantitat de pertorbacions de molt poca intensitat i unes poques d’una gran intensitat, de manera que si representem el logaritme de les intensitats vs. el logaritme de les freqüències ens queda una línia recta. Això és equivalent a dir que les magnituds de les pertorbacions tenen típicament una distribució estadística potencial (Eq. 0.5). Fa temps que se sap que és així en el cas concret dels terratrèmols, i el que féu Margalef és generalitzar-ho temptativament a tota mena de pertorbacions. Aquestes idees d’en Margalef no són gaire conegudes, tot i la seva relació amb les aportacions cabdals fetes posteriorment per en Per Bak i col· laboradors, que sí que tingueren molt impacte, merescudament. Aquests autors van descobrir uns mecanismes senzills i creïbles que podien explicar la generació de catàstrofes amb una distribució potencial, i els batejaren amb el nom de “criticalitat autoorganitzada” (SOC) (subsec. 2.3.3, 3.2.2). Des de llavors han proliferat els treballs que plantegen models SOC inspirats de forma més o menys llunyana en diferents tipus de catàstrofes, i també estudis empírics que suporten, de manera més o menys rigorosa, el caràcter potencial de les distribucions de diferents tipus de catàstrofes. Es parla de SOC en terratrèmols, tempestes, esllavissades, extincions en massa, incendis forestals, epidèmies, crisis econòmiques... i fins i tot guerres. El que jo he fet (cap. 3) és introduir noves idees en la teoria de la SOC, desenvolupar el model SOC d’incendis forestals mirant d’aproximar-lo a la realitat, contrastar-lo exitosament amb dades empíriques en diferents aspectes (sense limitar-me tan sols a mirar si hi ha una distribució potencial), extreure’n receptes de gestió... i adonar-me’n de la urgència d’aplicar-les. A més, en el llibre comento una mica les grans conseqüències que tindria si es constatés semblantment l’adequació d’una aproximació SOC a certes plagues i malalties agrícoles (sec. 4.5) i a la dinàmica d’extincions (subsec. 4.4.4).

El model SOC d’incendis forestals que es troba actualment a la bibliografia física va ser desenvolupat per Barbara Drossell i Franz Schwabl (subsec. 3.2.2). Aquests autors/es introduïren dinamisme en el model senzill i “estàtic” d’incendis forestals que explico a la secció anterior. Al llarg del temps es van col· locant “arbres” en el “tauler” a l’atzar, de manera que cada cel· la buida té una certa probabilitat, prefixada i comú a totes, d'aconseguir un arbre en un cert període. Semblantment, van caient "llamps" a l'atzar, amb una freqüència donada. Cada cop que cau un llamp en una cel· la ocupada, crema l'arbre que hi ha i el foc es propaga d'acord amb les regles que he explicat en la secció anterior. S'estableix que el procés des que cau el llamp fins que s'apaga el foc té una duració zero en comparació amb el procés d'acumulació d'arbres, reconeixent de manera implícita l'asimetria temporal de la qual parlava al començament, que és ben clara en el cas dels incendis forestals. Al cap d'un temps resulta que els arbres han passat a estar col· locats de manera fractal en el tauler i que les grandàries dels incendis han adquirit una distribució potencial (però fitada). Recordem que, en el model estàtic de la secció anterior, la grandària mitjana dels incendis passa d'insignificant a "infinita" en una franja molt estreta a la vora del punt crític. Si això es pot estendre al model dinàmic, serà dins d'aquesta franja on la taxa de combustió s'igualarà a la taxa a la qual s'afegeixen arbres (ja que la taxa global de combustió és proporcional a la grandària mitjana dels focs, cosa que no succeiria si els incendis no fossin "instantanis") i el sistema hi quedarà atrapat. Així se sol explicar que el sistema tingui les propietats escalants pròpies de la criticalitat: es considera que s'ha col· locat per si sol a la vora d'un punt crític, i per això se'n diu "crític autoorganitzat". De fet, no és evident que hagi de ser ben bé així, ja que, un cop en marxa, el sistema ja no és el mateix sistema construït aleatòriament del qual parlàvem en la secció anterior, sinó que és el resultat d'una història prèvia que n'ha modificat algunes característiques importants. En particular, els arbres tendeixen a estar més a prop els uns dels altres que si estiguessin disposats a l'atzar (sense perdre el caràcter fractal) i la densitat d'arbres és molt inferior a la que hom troba en el punt crític del model estàtic. Malgrat aquests canvis, se sol assumir que el sistema està efectivament a prop d'un punt crític. Com veurem, això no sembla que sigui cert.

Els conjunts de dades que s'han analitzat d'incendis forestals en el món real mostren distribucions potencials, semblantment al model (sec. 3.3). Per exemple, les dades dels boscos boreals a l'estat d'Alberta (Canadà), que són les que jo he treballat més, mostren unes grandàries d'incendis nítidament escalants des d'1 ha (com a mínim) fins a una mica més de 100000 ha (Fig. 3.1). La distribució potencial dels incendis és un dels molts fets que els estudiosos/es d'avui dia ignoren que ja se sabien fa 30 anys. Fa poc es va redescobrir (això sí, a partir d'anàlisis molt més esteses geogràficament que les que s'havien fet abans) i atribuir a la SOC, en un moment en què jo ja m'havia posat a analitzar en detall si de veritat s'hi pot establir una connexió, tenint en compte, per una banda, que hi ha molt poca relació entre aquests models senzills i els boscos reals i, per altra, que la distribució potencial per si mateixa n'és una prova molt parcial, i encara més si pensem que els valors que hom troba a la pràctica per al paràmetre b (Eq. 0.5) estan típicament algunes dècimes pel damunt que els que pot generar el model de Drossell i Schwabl (que no passa de 1.2).

Una esmena evident al model de Drossell i Schwabl si es vol aplicar a incendis reals és que no és cert que els arbres apareixin de sobte, per aquí i per allà. El que hi ha a la natura és una acumulació i modificació gradual del "combustible" a cada punt. Fa temps que s'ha postulat, per a alguns tipus de bosc, que passa un cert temps des que un rodal ha estat cremat fins que torna a tenir una probabilitat gran de cremar-se, com a conseqüència de l'anomenada "successió del combustible". He modelitzat la successió del combustible (subsec. 3.4.1) reemplaçant les presències/absències d'arbres a cada casella per un número r indicant la probabilitat que es cali foc a la casella quan una de les que l'envolten s'està cremant. Aquest número es fa zero quan una casella s'ha cremat, i després canvia de manera gradual i determinista fins que és afectada per un nou incendi. En si mateixa, aquesta modificació no té cap conseqüència seriosa per a les propietats del model dinàmic, ni de l'estàtic. Tenint en compte això, ja no és tan estrany que els resultats de Drossell i Schwabl siguin aplicables als boscos reals. Això sí, la part determinista de l'evolució del combustible ha de satisfer unes mínimes condicions perquè sorgeixi SOC: el valor final r_∞ que assoleix r en absència de foc ha d'estar per damunt del valor crític de r en el model estàtic, r_c, i el temps que passa fins que r arriba a ultrapassar r_c ha de ser molt gran en comparació amb la duració d'un incendi. No fa falta que r sempre creixi, en absència de foc. S'ha dit que alguns tipus de boscos són més refractaris al foc quan són molt madurs que quan tenen una maduresa intermèdia (subsec. 3.4.6): això és compatible amb una dinàmica SOC sempre i quan r no arribi a descendir per sota de r_c en madurar.

Una clau per anar més enllà de la simple observació que hi ha una distribució potencial està en comparar la resposta dels sistemes reals i simulats a canvis en les condicions de contorn: és com experimentar amb els sistemes, cosa que es pot fer sense embrancar-se amb grans projectes, tan sols extraient la informació apropiada dels "experiments naturals" que han quedat recollits en el registre històric.

El model de Drossell i Schwabl tan sols permet manipular un paràmetre ambiental, la freqüència d’ignicions, i ells/es mateixos/es en van estudiar les conseqüències (subsec. 3.4.5). La resposta del sistema és sorprenent: quan la freqüència d’ignicions augmenta, la grandària mitjana dels incendis tendeix gradualment a decréixer, de manera que, a la llarga, la taxa global de combustió no depèn gaire del nombre d’incendis que hi ha. Parlar de grandària mitjana no tindria sentit si no fos perquè, a diferència del que succeeix quan s’està exactament en un punt crític (recordem que se sol interpretar que el model de foc forestal es col· loca a la vora del punt crític, però no ben bé en el punt), en aquest model les grandàries dels focs no són indefinidament escalants, sinó que tenen un límit superior s_{max} (que no té res a veure amb la limitació òbvia posada per la mida del “tauler” si és finit). He pogut observar (sec. 3.3) que les dades empíriques de focs mostren també una fita superior estadísticament significativa (cosa ben important de cara a gestió). A més, a la bibliografia sobre incendis forestals fa temps que s’ha proposat que, efectivament, les grandàries dels incendis decreixen quan es fan més freqüents, en base a diverses observacions empíriques (subsec. 3.4.5). Es tracta d’una bona prova a favor de l’enfocament SOC. Cal aclarir en aquest punt que això no vol dir que un increment ràpid en la freqüència d’incendis sigui innocu, ja que, si veritablement el bosc és capaç de reaccionar-hi tal com he explicat, serà d’una manera molt lenta i, mentrestant, pot haver patit un excés de foc que li hagi deixat empremtes irreversibles (subsec. 3.7.3).

El gran avantatge del meu model és que, a més de la freqüència d’ignicions, permet introduir sense cap dificultat les condicions ambientals (sec. 3.4) (i les intervencions humanes, sec. 3.5) que modulen la propagació del foc un cop s’ha calat. Se sap bé que la propagació dels incendis no és tan sols cosa de combustible, sinó que la meteorologia i altres factors hi pinten molt. Una manera senzilla d’abordar aquesta qüestió, i que no sembla que aporti menys que d’altres opcions més complicades, consisteix a tractar la r de cada casella com a la suma de dues components (subsec. 3.4.1):

$$r=r_i+r_e \qquad (0.6).$$

r_i es refereix al combustible i funciona tal com funcionava la pròpia r en l’aproximació anterior. r_e és la component que captura els factors externs: podem establir que sigui la mateixa per a totes les caselles i a llavors fer-la canviar al llarg del temps com ens plagui.

Al seu torn, d'entre les variacions de r_e podem distingir les que són a curt (subsec. 3.4.2-3.4.4) de les que són a llarg (subsec. 3.4.5) termini. Per exemple, si parlem de la component atmosfèrica de r_e, les de curt termini són les variacions meteorològiques de dia a dia, d'any en any, o qualsevol període curt comparat amb la recurrència dels incendis; les de llarg termini corresponen a canvis climàtics un cop hagi passat molt temps des que s'hagin produït i estabilitzat. Les simulacions mostren que els canvis a llarg termini de r_e actuen semblantment als canvis en el nombre d'incendis: el bosc es reorganitza de manera que les compensa parcialment, i les coses no canvien gaire (subsec. 3.4.5).

Els canvis a curt termini són més reveladors. El resultat immediat de reduir r_e (subsec. 3.4.2, Fig. 3.3A-C) és un augment de b (Eq. 0.5) i una disminució de la fita superior s_{max}, de manera que els incendis es fan més petits, i quan augmenta r_e el resultat és el contrari. Si deixem que el bosc s'autoorganitzi subjecte a fluctuacions ambientals continuades (tal com suceeix a la pràctica), b i s_{max} van ballant al ritme que els marca r_e. Per tant, quan prenem totes les dades juntes, el que tenim és una superposició de distribucions potencials amb paràmetres una mica diferents, però prou pròxims perquè el conjunt també sigui a tots els efectes una distribució potencial fitada (Fig. 3.2). La b conjunta pot assolir valors superiors als del model de Drossell i Schwabl, cosa que resol les discrepàncies amb les dades reals. Però el resultat més interessant va aparèixer en prendre les dades d'Alberta i separar-les en grups en funció d'un índex meteorològic (Fig. 3.3D-F). Va resultar que cada subgrup mostrava una distribució potencial amb valors diferents de b i s_{max}, i que canviaven amb l'índex meteorològic d'una manera molt semblant a com canvien amb r_e en el model. Aquest resultat no tan sols dona un recolzament molt fort al model, sinó que a més té una utilitat pràctica molt directa. Un cop analitzades les dades històriques tal com ho he fet en el cas d'Alberta, permet preveure la distribució estadística complerta de les grandàries dels focs en funció d'unes variables meteorològiques per a les quals s'obtenen prediccions rutinàriament.

També és interessant veure què succeeix si, enlloc d'analitzar les respostes de b i s_{max} a r_e, mirem directament la resposta de la grandària mitjana dels focs $\bar{s}$ (subsec. 3.4.3). Quan això ho fem amb un bosc SOC simulat, trobem que $\bar{s}$ creix més que exponencialment amb r_e (Fig. 3.4A), d'una manera característica que sembla delatar l'aproximació gradual a un punt crític, tal com succeeix en variar r en un model "estàtic" amb el "combustible" repartit homogèniament. No podem observar la transició crítica cap a un $\bar{s}$ infinit perquè el bosc que simulem serà finit (com qualsevol bosc real, d'altra banda), però la forma de la resposta fa pensar que existeix aquest punt crític, per damunt del qual la grandària dels focs tan sols està limitada per l'extensió total que hi ha disponible o per altres factors externs. A més, permet estimar-ne la posició, que no sembla gaire diferent de la que té en un "bosc" homogeni. En aquesta qüestió, l'única diferència important entre un bosc SOC i un bosc homogeni és que la resposta del bosc SOC a r_e és molt més suau, molt més pròxima a una resposta exponencial, que no pas la resposta del bosc homogeni, que és molt abrupta. Això és conseqüència que, en el primer cas, el combustible està ubicat d'una manera més "agrumollada", cosa que afavoreix la propagació del foc. Per això hi ha focs molt grans en els boscos SOC a pesar que el seu valor mitjà de r és tan baix que faria negligibles els focs en un bosc homogeni. Tenint en compte tot això, no sembla correcte dir que un bosc SOC simulat sigui crític ni estigui a la vora d'un punt crític. El que succeeix és que té una resposta molt més suau a variacions de r, cosa que fa que, per una banda, no li calgui estar a prop del punt crític per arribar a taxes de combustió prou altes per assolir l'estat estacionari i, per altra, que tot i estar lluny del punt crític mostri propietats semblants a les que tindria si n'estigués a prop. He optat per mantenir la denominació de "criticalitat autoorganitzada", tot i que potser fóra més correcte fer servir algun altre terme, com ara "paracriticalitat autoorganitzada"...

La variació de les grandàries dels focs a Alberta no es pot distingir d'una resposta exponencial (Fig. 3.5), cosa que no és sorprenent. Les coses canvien, però, si parem atenció al cas més ben documentat d'incendis que es van percebre com a anòmals, és a dir, els que van tenir lloc en el parc de Yellowstone el 1988 (subsec.3.4.4). Aquest episodi va cloure un període, iniciat el 1972, en el qual es va permetre la lliure propagació d'una bona part dels incendis, després d'una llarga història prèvia de supressió, en què probablement s'havia acumulat un excés de combustible. Els focs de 1972-1988 semblen mostrar una resposta més que exponencial a un índex meteorològic (Fig. 3.6A), però no hi ha prou dades per tenir un resultat estadísticament acurat. Sí que hi ha un element molt suggestiu: $\boldsymbol{b}$ varia en funció de l'índex meteorològic (Fig. 3.6B) de manera semblant al $\boldsymbol{b}$ d'Alberta, només que en aquest cas va arribar a un valor excepcionalment baix de $\boldsymbol{b} \approx 1$ el 1988 (Fig. 3.7A). Es pot demostrar que, si existeix un estat crític ben definit, no pot tenir una $\boldsymbol{b}$

per sota d'aquest valor, cosa que em fa tenir una certa sospita que l'estiu del 1988 es va assolir o ultrapassar l'estat crític. Si és així, l'únic que limitava les grandàries dels focs eren les restriccions externes. En concret, va ser la duració de l'estació favorable al foc, ja que els van apagar les nevades de la tardor.

Per a boscos com els de Yellowstone, la qüestió de si es va ultrapassar o no el punt crític té un interès pràctic limitat, ja que per sota d'aquest punt les mides dels incendis segueixen sent apreciables (per la resposta gradual dels sistemes SOC) i per damunt segueixen estant limitades (per les condicions de contorn). Però la possibilitat d'ultrapassar-lo esdevé paorosa si hi ha boscos que no siguin SOC perquè tinguin $r_\infty < r_c$. En un cas així es pot esperar una transició molt abrupta de foc negligibles a focs extraordinaris, per als quals la flora, fauna, sòl, etc. no estaran preparats, ni tampoc els humans que habitin en el bosc o a prop. Doncs probablement aquests boscos existeixen, probablement sigui el cas de les selves tropicals (sec. 3.6, subsec. 3.7.4). La informació que n'hi ha a la bibliografia mostra una gran concordança qualitativa amb les expectatives del model, com explicava al començament de la introducció. Malauradament, no he tingut accés a dades quantitatives adequades per posar-ho a prova més adequadament (he tingut accés a informació de satèl·lit de l'Amazònia d'un parell d'anys, però no n'he pogut treure cap conclusió clara), però és extremadament important estudiar aquesta qüestió. Cal analitzar si efectivament hi ha punts crítics, quantificar-los i, un cop coneguem amb precisió aquests llindars tan vitals a no ultrapassar, fer tots els esforços imaginables perquè els impactes antropogènics sobre l'entorn global i regional no desencadenin gaires més catàstrofes com les que ja hem començat a veure els darrers anys.

No crec que la gestió dels boscos presumptament SOC (sec. 3.5, subsec. 3.7.2) sigui una qüestió tan urgent com la política a seguir amb les selves tropicals, però té la seva importància i planteja reptes encara més complexos a nivell teòric, als quals probablement també ens hàgim d'enfrontar quan vulguem gestionar altres tipus de catàstrofes apart del foc. La gran dificultat de gestionar un sistema SOC està en què tendeix a anar en contra de qualsevol cosa que li facis i, si finalment aconsegueixes "sotmetre'l", el pots haver transformat en un polvorí. Si aquest model és correcte, dedicar-se a apagar focs (a no ser que sigui com a resposta a un augment recent en el nombre d'ignicions o en la facilitat de propagació) o esquarterar el territori amb "perímetres de protecció prioritària" (com els que es plantegen actualment a

Catalunya) tindrà una efectivitat molt reduïda a llarg termini, en comparació amb la que tingui a curt termini. A més, un cop n'hagi minvat l'efectivitat, el bosc hi haurà esdevingut "addicte": si s'hi deixa d'intervenir de sobte (sia de manera planificada, com a Yellowstone, o com a conseqüència d'una crisi econòmica o social, com ha succeït en gran part en els territoris de l'antiga Unió Soviètica després del seu esfondrament) els focs esdevindran molt pitjors que no eren abans de començar a intervenir. La tendència a la minva en l'efectivitat tan sols desapareix si l'esforç d'apagar focs és tan intens com perquè r_∞ caigui per sota de r_c i el bosc deixi de ser SOC, però a llavors el nivell d'addicció es fa extrem, les conseqüències de deixar d'intervenir esdevenen encara molt pitjors.

Aleshores, què podem fer amb un bosc SOC? (subsec. 3.7.2, sec. 3.5). Per començar, prevenir o neutralitzar qualsevol augment ràpid en la taxa d'ignicions o en els factors que afavoreixen la propagació del foc, ja que la velocitat dels canvis és sovint més important que no la seva magnitud. També mirar de reparar qualsevol modificació antropogènica (en la composició d'espècies, en les condicions ambientals...) que hagi pogut causar un cert augment sostingut en les taxes de combustió. Poden emprar-se intensivament bombers/es i tallafocs en àrees reduïdes que calgui mantenir lliures de foc, immerses en un entramat que funcioni de manera més natural. A la llarga, només serveix d'alguna cosa estendre la supressió a tot el territori si és prou intensa per eliminar la dinàmica SOC... i si s'assumeix el risc de daltabaix general que se'n deriva. Si no es considera acceptable aquest risc ni tampoc una dinàmica de focs natural, es pot afavorir que, de manera molt gradual, l'estructura del bosc derivi cap a una situació d'incendis freqüents però petits. Cal dir que la prescripció d'incendis és una manera d'aconseguir-ho, però no és l'única, que si s'aplica cal que sigui de manera gradual, que no té raó de ser allà on la freqüència d'incendis ja està augmentant massa depressa per altres motius, i que segons com s'apliqui pot ser contraproduent (quan els incendis prescrits s'adrecen a crear discontinuïtats en grans masses de combustible, com és sovint el cas, són comparables a l'apagament de focs, ja que afavoreixen l'acumulació de combustible enlloc de limitar-la com fa la pròpia autoorganització del bosc). Allà on sigui adient ecològicament i econòmica, i anant amb moltíssima cura, fins es pot reemplaçar el bosc SOC per un mosaic forestal comparable però mantingut per procediments aliens al foc: així semblen haver estat tradicionalment molts paisatges mediterranis, i en aquest cas la biota ja està amotllada a una situació així. En canvi, en enretirar combustible de manera homogènia, es corre el risc d'afavorir una transició abrupta cap a grans incendis en algun moment del futur, a més de mantenir una estructura del territori que pot ser menys apropiada en termes de biodiversitat.

L'interès d'estudiar els incendis no s'acaba en els propis incendis: aquestes troballes obren una porta a la comprensió de la dinàmica de les catàstrofes que segurament tingui un abast molt més general, ja que les assumpcions en què ens hem basat eren molt àmplies. Aprendre a predir i gestionar les catàstrofes és sovint qüestió de vida o mort, de moltes vides o moltes morts. Per anar més enllà de les aproximacions convencionals, el primer que podem fer és posar a prova si tenen una distribució estadística potencial i, en cas afirmatiu, estimar-ne els paràmetres. Ha quedat clar que la distribució potencial és importantíssima, no tan sols pel que fa a catàstrofes, sinó també per a molts altres aspectes de l'ecologia i d'altres ciències. Malauradament, la formació estadística que es dóna habitualment a les universitats no inclou les eines per treballar-hi, ja que se centra en la distribució normal o gaussiana, que té propietats molt diferents, i no conec cap text sintètic que supleixi aquesta mancança. Com a resultat, els pocs treballs en què es té en compte aquesta distribució acostumen a ser bastant fluixos des del punt de vista estadístic. He intentat cobrir aquest buit amb l'apèndix 2A, on ajunto informació sobre la distribució potencial procedent de la bibliografia amb receptes pròpies i ensenyances que he extret de la meva experiència tractant dades. Té especial importància el posar a prova si hi ha un límit superior a la magnitud de les catàstrofes. Un cop establert que hi ha una distribució potencial i estimats els paràmetres, una altra passa fàcil de dur a terme en molts casos serà esbrinar si aquests paràmetres varien en funció de factors externs, tal com he fet en el cas dels incendis a Alberta (subsec. 3.4.2), i construir així un sistema de pronòstic de catàstrofes. I es pot anar encara molt més enllà. S'ha de remarcar que una de les facetes més incertes de les projeccions de canvi climàtic global, i alhora una de les més importants, és l'evolució de tota mena d'esdeveniments extrems: espero que aquesta incertesa minvi apreciablement en aplicar aquests mètodes que proposo.

Des del punt de vista socioecològic, seria especialment important esbrinar si, en el cas de l'agricultura, hi ha plagues i/o malalties amb una dinàmica SOC (sec. 4.5). N'hi ha que són de tipus epidèmic, de manera que es pot dir que es presenten en episodis ben definits que comencen i s'acaben, potser de manera ràpida en comparació amb el temps que fa falta perquè els camps tornin a estar en condicions que afavoreixin noves proliferacions. Si és així, tenim els requisits bàsics per a una dinàmica com la dels incendis. De fet, un article en una revista d'impacte va proposar una dinàmica SOC per a algunes epidèmies humanes, donant-li un recolzament amb dades empíriques, tot i que força feble. En el cas agrícola, la recuperació de les condicions de propagació ve mediada per les decisions dels agricultors/es, però això no és incompatible amb la dinàmica SOC si cada individu planifica pensant en la seva pròpia parcel·la enlloc del territori en conjunt. Les condicions locals de propagació no depenen tan sols de la biomassa de plantes, sinó també de la seva homogeneïtat genètica i de qualsevol mètode que s'empri per al control de les plagues o els patògens. La hipòtesi d'una dinàmica SOC explicaria l'observació que la proporció de la producció agrícola nord-americana consumida per plagues i patògens no va variar gaire entre els anys 40 i els 80, malgrat un augment de més del 30000% en la quantitat de plaguicides sintètics aplicats multiplicada pel seu grau de toxicitat, i que aquesta proporció no sembli ser gaire diferent de la de l'Europa medieval. Si decobríssim que aquesta hipòtesi és vàlida, tindríem una bon punt de partida per eliminar o reduir substancialment l'aplicació de plaguicides i al mateix temps augmentar la producció agrícola. Està clar que els problemes d'alimentació no es poden resoldre augmentant tan sols la producció: és indispensable un accés més igualitari als aliments, desincentivar el creixement demogràfic i assegurar la sostenibilitat d'aquesta producció, però un augment de producció a mig termini bé podria esdevenir indefugible. Doncs vegem com es podria reestructurar l'agricultura mundial si trobéssim que la hipòtesi SOC es compleix. S'hauria de fer una divisió del treball en grans àrees geogràfiques. En algunes àrees, els agricultors/es es coordinarien per augmentar el combat de plagues i patògens apreciablement per damunt dels seus òptims locals, servint-se de la biodiversitat vegetal, dissenyant apropiadament la disposició en l'espai de cada varietat de plantes a totes les escales, i aplicant altres mètodes de control ambientalment responsables. Idealment, s'hauria d'arribar a superar la pròpia dinàmica SOC. D'aquesta manera, s'aconseguiria una reducció dràstica en la incidència de plagues i malalties i un gran augment en la producció aprofitable. Si això fos tot, la selecció natural faria més pressió sobre plagues i patògens per adaptar-se a la nova situació que no pas sobre les plantes per seguir sent-hi resistents, i a partir de cert moment n'hi podria haver una expansió abrupta. Per evitar-ho, caldria dedicar altres territoris a agricultura tradicional, també baixa en plaguicides però sense coordinació entre agricultors/es, amb una dinàmica SOC i coevolució en marxa entre plantes i llurs paràssits. D'aquests territoris sortirien noves varietats de plantes per anar renovant les dels territoris d'alta resistència. De moment tot això és molt fantasiós, perquè ni tan sols sabem si hi ha plagues o malalties amb una dinàmica SOC. Però el cas és que tenim eines per esbrinar-ho i cal fer-ho ràpidament, ja que les conseqüències serien extraordinàries. Sigui com sigui, el que està clar és que fa falta aturar la pèrdua irreversible de biodiversitat agrícola, ara magnificada per la globalització en la seva forma actual, les pressions del mercat i la introducció de plantes genèticament modificades.

Un altre terreny de gran importància ecològica en què hi podria haver SOC és en les xarxes de relacions alimentàries i d'altres tipus d'interaccions entre espècies, tenint en compte que a vegades hi ha una propagació d'espècie a espècie de canvis importants com ara la pròpia extinció de les espècies implicades. La idea d'una dinàmica SOC a aquest nivell ha estat proposada per altres autors/es i pot ser molt important a l'hora d'entendre els trets estructurals d'aquestes xarxes i de preveure les conseqüències per als ecosistemes d'introduir o extingir espècies, així com de canvis ambientals que puguin modular la propagació de les fluctuacions. L'evidència, però, és encara minsa. No, obstant, m'he dedicat a introduir algunes variacions sobre els models existents i explorar-ne les conseqüències (subsec. 4.4.4, Ap. 4D). Una idea que aporto és que, encara que els nexes entre espècies són la via per on es propaguen les fluctuacions, una multiplicitat de nexes també podria amortir-les. Per tant, un dels motors de la dinàmica SOC podria ser la tendència a l'especialització perdent nexes enlloc de la tendència a afegir nexes en què s'ha pensat a vegades. A llavors, encara que, seguint un plantejament força estès, protegíssim selectivament les espècies que tenen més números per desencadenar canvis ecològics d'importància en extingir-se, ens podríem trobar que les baixes en la resta d'espècies agreugessin les conseqüències en el moment en què s'arribés a extingir alguna de les espècies clau. Cal insistir que aquests comentaris sobre xarxes tròfiques són tan sols especulacions, esperant a ser contrastades amb informació empírica.

De l'ecologia a l'economia

La humanitat ha arribat al punt perillós de tenir una capacitat per modificar el seu propi món que excedeix de lluny la comprensió que en té. Cada cop hi ha més gent conscient del fet que, per donar resposta als immensos reptes que tenim, ja no serveixen els conceptes que, encara avui dia, informen la presa de decisions als més alts nivells, si és que algun cop han servit. Un seguit d'autors/es insatisfets amb l'economia que s'ensenya han mirat a l'ecologia buscant nous camins, i d'aquesta mirada ha nascut l'economia ecològica. A l'economia ecològica li falta bastant per arribar a ser el que pot arribar a ser, perquè també a l'ecologia li falta. Una de les meves motivacions per intentar com pogués donar una empenta intel· lectual a l'ecologia ha estat l'expectativa que aquesta empenta també fos empenta per a l'economia ecològica. Des de l'economia ecològica s'ha parlat molt de termodinàmica, de llindars "crítics" que cal no superar, de l'estabilitat dels sistemes i com els afecta la pèrdua de diversitat... S'ha parlat molt de la mena de temes als quals m'he dedicat, i com que els meus resultats hi poden aportar alguna cosa, i com que també he donat algunes voltes a qüestions més pròpies de l'economia que l'ecologia, he pensat que valia la pena deixar sobre el paper algunes reflexions entorn d'aquests assumptes (cap. 6).

La recerca d'un marc termodinàmic per a l'economia ecològica ha estat un tema recurrent al qual s'ha parat molta atenció (sec. 6.2). Hi fou clau el llibre "The Entropy Law and the Economic Process", de Nicholas Georgescu-Roegen, un dels pares de l'economia ecològica moderna. Aquest autor va intuir que el valor econòmic implica alguna dosi de "baixa entropia", tot i que no hi hauria cap correspondència quantitativa entre ambdues mesures. En economia se sol considerar que, perquè quelcom tingui valor, cal que s'ajusti a les pretensions d'algun agent econòmic (per exemple, uns pantalons vermells no tindrien valor si no hi hagués gent que aspira a tenir uns pantalons vermells). Aquesta correspondència entre l'objecte valorat i les expectatives de qui valora és informació mútua, i per tant és completament correcte dir que el valor implica "baixa entropia", d'acord amb els resultats que resumeixo a "somnis i desenganys termodinàmics". A més, n'hi ha prou amb les propietats bàsiques de la variable "valor" perquè el procés econòmic només sigui concebible lluny de l'equilibri termodinàmic. En canvi, no es pot justificar la idea d'en Georgescu-Roegen que la Segona Llei imposaria una data de caducitat al sistema econòmic. La termodinàmica no fa impossible la sostenibilitat, tan sols en restringeix les condicions (com ja assumeixen, correctament, la majoria d'economistes ecològics). Ni tan sols és certa la idea implícita o explícita de molts autors, segons la quals els problemes de sostenibilitat es poden reduir a un afer d'augments d'entropia a l'entorn. Per a la sostenibilitat és importantíssim fer quadrar els balanços termodinàmics, però amb això no n'hi ha prou. Tot problema de sostenibilitat resulta de canvis irreversibles produïts per l'activitat econòmica, però la majoria dels que ens preocupen més avui dia tenen a veure amb pèrdues de complexitat, no amb acumulació d'entropia. És el cas de la degradació dels ecosistemes, la pèrdua de sòls, el col· lapse de les pesqueries, la minva de la biodiversitat, etc., i, tot i que no sigui tan immediat d'entendre, també és el cas del canvi climàtic. Com abans ha quedat clar, la irreversibilitat en les pèrdues de complexitat és en darrer terme un resultat de la irreversibilitat termodinàmica, però són dues coses diferents. Conseqüentment, encara erren més els autors que (a diferència de Georgescu-Roegen) pretenen arribar a quantificar el valor de totes les coses amb una única mesura de tipus termodinàmic (subsec. 6.2.2). La complexitat de la Biosfera és inherentment multidimensional. La degradació dels sòls no es pot resoldre pescant menys, encara que una cosa sembli compensar l'altra quan només en mirem l'energia, emergia, exergia o el que sigui.

L'economia ecològica no pot reduir-se a economia energètica, ha de ser veritablement ecològica, s'ha d'interessar per l'encaix de l'economia humana en l'entramat biosfèric en tota la seva complexitat. Però no és gens senzill establir quina mena d'encaix és l'encaix apropiat (subsec. 6.3.1). En altres èpoques, els ecòlegs/ogues creien que hi havia un "equilibri ecològic", i els/les qui defensaven prudència i contenció a l'hora de transformar l'entorn argumentaven que no l'havíem de trencar. Però amb el temps s'ha anat percebent cada cop més que no hi ha aital equilibri, que el canvi a totes les escales d'espai i de temps és inherent a la natura tal com la coneixem. Aleshores, a sant de què hauríem d'anar els humans amb tanta cura? Doncs perquè, d'acord amb la "sageta ecològica del temps" que implica el principi (modificat) de Margalef que he formulat en parlar de "somnis i desenganys termodinàmics", quan forcem les coses a canviar per damunt d'un cert ritme el resultat tan sols pot ser destructiu, tant si partim d'un "equilibri" com si no. I, en principi, destruir no és desitjable ni des d'un punt de vista biocèntric ni des d'un punt de vista antropocèntric, per molt que sempre hi hagi hagut episodis de destrucció i que la Biosfera que coneixem no es pugui entendre sense aquests episodis. Des d'un punt de vista biocèntric, no és ètic que la humanitat afegeixi més destrucció per molt que també n'hi hagi d'altres orígens, tal com no és ètic fer una mala passada a una persona encara que aquesta persona pugui tenir també molts altres maldecaps. Des d'un punt de vista antropocèntric, no ens preocupa si hi ha hagut altres episodis de destrucció massiva, el que ens preocupa és que no n'hi hagi cap ara que ens pugui arrossegar a nosaltres, o que ens pugui deixar sense recursos importants per al nostre benestar d'una manera irreversible a escala humana.

Potser les extincions en massa formen part d'una dinàmica adaptativa a una escala de temps gran, potser podríem veure l'actual tan sols com a la primera passa (tot i que la més ràpida i irreversible) en l'adaptació de la Biosfera al nou món que els humans estem fent. No obstant, abans de seguir endavant amb la desfeta hauríem de preguntar-nos si de veritat és aquest el món que volem, si de veritat serem capaços de formar-ne part, i si val la pena que les altres espècies n'hagin de pagar un preu tan alt. I també hauríem d'assegurar-nos que no estem forçant la Biosfera més enllà de la seva capacitat d'adaptació.

Tot i que no hi hagi pròpiament un "equilibri" per trencar, el que sí que succeeix és que un canvi determinat pot desencadenar una avalanxa d'altres canvis, i també que hi ha, segons sembla, llindars ben definits en l'impacte sobre els ecosistemes a partir dels quals es produeixen canvis radicals. El problema està en tenir mètodes per quantificar la magnitud de les avalanxes i la posició dels llindars, i en això crec haver fet aportacions substancials en aquest llibre. Algunes de les principals receptes proposades per economistes ecològics per fer operacional la idea de sostenibilitat consisteixen a establir un seguit de llindars "crítics" que cal no excedir. Empren el terme "crítics" en un sentit més genèric que no el sentit físic que faig servir en aquest llibre, però està clar que la quantificació dels llindars crítics en el segon sentit és indispensable per decidir alguns dels llindars "crítics" en el primer.

Les coses es compliquen si tenim en compte que les intervencions sobre l'entorn no tan sols poden desencadenar catàstrofes d'una manera directa, sinó que deixen empremtes que poden acabar modificant els règims de catàstrofes (i de fluctuacions no pròpiament catastròfiques) en el futur. Hem vist, en primer lloc, que la pèrdua de diversitat sembla fer que els ecosistemes esdevinguin globalment més fluctuants i menys adaptables. També hem vist que, en molts casos, els règims naturals de catàstrofes semblen estar determinats per una dinàmica de "criticalitat autoorganitzada" (SOC), cosa que els fa respondre de formes sorprenents i no gens trivials a la càrrega de catàstrofes anteriorment sofertes pel sistema i a les intervencions externes. Espero haver contribuït a fer més predictibles aquestes respostes. Això és encara més important si tenim en compte que bé podria haver-hi una dinàmica SOC rera els flagells clàssics de la humanitat, com les fams, les epidèmies i àdhuc les guerres.

Està clar que amb el coneixement de com funcionen els sistemes no n'hi ha prou per prendre decisions polítiques. Aquest coneixement pot servir per preveure les conseqüències dels nostres actes, en una certa mesura, però cal algun criteri per avaluar aquestes conseqüències (els punts de vista ètics biocèntric i antropocèntric, que he esmentat, són diferents opcions, però només com a punt de partida molt bàsic). Actualment el criteri estàndard és el de cost-benefici en termes monetaris, i per tant val la pena analitzar-lo en amb un cert detall (subsec. 6.4.1). El valor monetari té interès per a objectes dels quals disposa un individu o un conjunt d'individus i que pot intercanviar en un mercat, però no està tan clar que permeti anar més enllà, com quan es fa servir com a guia per a les institucions públiques, i fins i tot s'aplica a elements de l'entorn per als quals no hi ha cap mercat (aire net, clima, biodiversitat...), a base d'intentar estimar el preu que tindrien si fossin comercialitzables (això en el millor dels casos, el més habitual és que simplement s'ignorin). L'ús de la mesura monetària es fonamenta en dos factors: la intercanviabilitat i la desitjabilitat. La intercanviabilitat és real quan tens algú a qui comprar i vendre, però, com hem vist en parlar de termodinàmica i complexitat, no hi ha una intercanviabilitat física entre les coses a escala global, que les faci comparables. Això sembla de calaix, però després trobem que un equip d'economistes del Panell Intergovernamental sobre el Canvi Climàtic (incloent dos premis Nobel) ens dóna a entendre que les generacions futures tindran el màxim marge de decisió si els deixem el màxim de riquesa mesurada monetàriament, com si d'aquesta manera poguessin decidir si volen "comprar" una atmosfera en condicions encara que nosaltres ja l'hàgim espatllat... En canvi, si la mesura monetària s'utilitza tan sols per comparar les coses en funció que siguin més o menys desitjables, no hi ha cap incorrecció física, però sí, com ja han manifestat molts altres autors/es, unes decisions polítiques de base molt fortes que es volen dissimular quan es presenta aquesta mesura com a objectiva, fins i tot a treballs pretesament científics. És cert que els preus de les coses resulten d'integrar les decisions de moltes persones i reflecteixen llurs preferències, però és un reflex esbiaixat: com diuen molts autors/es, és com una votació en què el nombre de vots és proporcional als diners que té cadascú. Aquest biaix es posa de manifest de manera ben gràfica quan es vol posar preu a elements externs al mercat, com la pròpia vida humana. Segons els càlculs "científics" d'uns altres economistes del Panell Intergovernamental, la vida d'un ciutadà d'un país "ric" val 15 cops més que la d'un ciutadà d'un país "pobre"... El fet és que no pot existir una mesura objectiva del valor. Qualsevol mesura de valor demana, com a mínim, una decisió que no és ni pot ser científica, la del valor per a qui: per a tothom? només per als rics? només per als de cert país? només homes i no dones? només la generació actual i no les futures? només per a humans i no per a éssers d'altres espècies? Tan sols si ja s'ha pres aquesta decisió hom pot fer valoracions. A més, caldria tenir en compte altres biaixos de la valoració monetària, com els que tenen a veure amb addicció, publicitat, efectes posicionals, etc.

Es pot pensar que unes valoracions "alternatives" podrien ser formalment anàlogues a les monetàries, però fins i tot això és discutible. Les valoracions monetàries són linears i additives. Per a qualsevol funció matemàtica que reuneixi unes mínimes condicions, l'esquema lineal i additiu és una bona aproximació per a canvis molt petits. Però ens podem trobar que també hàgim de valorar canvis grans, perquè estiguem pensant en lapses de temps llargs, o en transicions de fase, o en remodelacions substancials de l'economia. Aleshores ens cal una aproximació no-lineal. En principi, aquesta aproximació és matemàticament immediata (prenent més termes en una sèrie de Taylor), però aplicar-la a la pràctica pot ser tremendament difícil. Com diuen alguns economistes ecològics/ques, potser sigui una complicació innecessària el voler prendre les decisions per procediments totalment formalitzats i quantitatius, com pretén l'economia convencional, i valgui més recórrer a procediments no-algorítmics de decisió multi-criteri.

Un dels aspectes de l'economia convencional que més s'ha discutit és el "descompte del futur" (subsec. 6.4.2). Si tenim un euro i l'ingressem al banc, se'ns multiplicarà exponencialment al ritme que marquin els tipus d'interès. A més de posar un valor monetari al medi ambient, se sol considerar que només està justificat deixar de guanyar ara una certa quantitat per obtenir un benefici ambiental futur si el valor monetari d'aquest benefici ha de superar el valor d'aquella quantitat un cop s'hagi multiplicat exponencialment en el banc, suposant que els tipus d'interès actuals es poden extrapolar indefinidament en el futur. El resultat és que res no importa més enllà d'una certa finestra temporal, que no sol anar gaire més enllà de quinze o vint anys. Aquest punt de vista ha estat molt criticat des de l'economia ecològica, amb raó. Les justificacions per a aquesta política es basen en part en criteris que impliquen una discriminació envers les generacions futures i en part en l'assumpció que hi haurà un progrés indefinit, de manera que els beneficis que puguem deixar per al futur seran cada cop més innecessaris. Aquesta és una predicció fortíssima amb uns fonaments feblíssims, tenint en compte que encara no se sap quasi res sobre la dinàmica de les fluctuacions econòmiques i socials i poca cosa sobre les limitacions ambientals (i que no hi ha res que duri per sempre), per no parlar del poc que se sap sobre la relació entre guanys monetaris i benestar real. Contrasta de manera frapant amb l'extremada certitud que s'exigeix a les prediccions ambientals abans d'incórrer en qualsevol cost monetari en el seu nom. A més, si prenem uns criteris de valor més igualitaris i, per tant, més centrats en les necessitats bàsiques, no fóra gens inversemblant un retrocés en un futur proper, ja que estem malmetent de manera molt greu les bases ecològiques de llur satisfacció, alhora que es manté a moltes regions un entorn sòcio-econòmic favorable al creixement de la població i es desmantella qualsevol mecanisme de redistribució de la riquesa. Per tant, si fem servir criteris més igualitaris no hi ha justificació per a un descompte del futur d'aquesta forma i, dins d'una certa finestra temporal, fins es podria justificar un descompte negatiu. D'altra banda, les incerteses sobre les conseqüències futures d'allò que fem avui dia podrien justificar un cert descompte en alguns casos, tot i que també forneixen arguments addicionals contra el descompte en altres casos. Ara bé, si hem d'aplicar un descompte per motius d'incertesa, hauria de ser molt petit comparat amb l'actual i mai seguint un esquema exponencial que assenyali una sola escala temporal. En el món en què vivim se superposen moltes escales temporals, i com a conseqüència la predictibilitat decreix de manera menys que exponencial, de manera més aviat potencial: qui sap si és per això que els experiments amb humans i altres animals mostren que en les nostres vides quotidianes descomptem efectivament de manera menys que exponencial, cosa que no concorda amb els models de l'economia convencional. A més, l'escala temporal a considerar per a cada procés ecològic (o geològic, o social, etc.) hauria de dependre en bona part de l'escala característica del procés en qüestió. Per moltes incerteses que hi hagi, segueix essent raonable assumir una probabilitat no gaire petita que les pèrdues de biodiversitat que hi ha ara o les deixalles radioactives que estem produint segueixin sent un problema seriós d'aquí cent mil o un milió d'anys. Una conseqüència important d'incloure en els nostres comptes horitzons temporals grans és que necessàriament hem de prendre en consideració termes no-lineals dels que abans he esmentat.

La qüestió del valor que es dóna als esdeveniments futurs està fortament associada a la idea de sostenibilitat (subsec. 6.4.3), idea poc clara i que ha generat moltes discussions. En un article força encertat, Costanza i Patten assenyalen que un sistema sostenible és un sistema que persisteix, però que no té sentit afavorir la persistència sense més, ja que el que trobem a la natura és una jerarquia de sistemes dins d'altres sistemes, que es renoven a diferents escales temporals, i que la renovació dels subsistemes és necessària per a l'adaptabilitat dels sistemes dels quals formen part. Per tant cal decidir què és el que volem que persisteixi i per quant temps, tenint en compte les seves interrelacions amb altres elements. Podem afegir que tan sols té sentit posar-se la sostenibilitat com a fita per si mateixa en termes de característiques molt globals, no pas de detall, i ens cal reprendre la qüestió dels criteris de valor per decidir quines són les característiques globals que volem mantenir. Una "sostenibilitat" que podríem triar és la sostenibilitat en la capacitat del nostre món com a suport per al benestar humà (o del conjunt éssers dels quals ens preocupi el benestar). A continuació veurem com es pot concretar millor aquesta idea, prenent com a punt de partida l'anàlisi crítica de l'anomenada "sostenibilitat feble", de la qual el Banc Mundial publica estimacions anuals. La recepta de la sostenibilitat feble procedeix de l'ortodòxia econòmica i consisteix en que cada generació deixi a la generació següent una quantitat de capital no inferior a la que li ha deixat la generació precedent, prenent com a "quantitat de capital" la suma de tots els tipus de capital, incloent el "capital natural", mesurats en unitats monetàries. Aquesta idea ha estat criticada no tan sols des de l'economia ecològica sinó també per algun representant de l'ala més ortodoxa de l'ortodòxia econòmica, assenyalant que el criteri de decisió ha de ser l'optimització econòmica convencional, no pas la sostenibilitat de cap mena. Des de l'economia ecològica s'acostuma a optar per la "sostenibilitat forta", que es basaria en una multiplicitat de variables enlloc d'una sola. La forma més habitual en que es presenta és com a un conjunt de llindars "crítics" que cal no superar, però no hi ha una sola definició tan concreta com la de la sostenibilitat feble. Des del meu punt de vista, les principals crítiques que es poden fer a la sostenibilitat feble són els fets que empri unitats monetàries i que es plantegi en base a una sola escala temporal, que és la d'una generació. Enlloc de pensar en termes monetaris hauríem de pensar en alguna mesura més real de benestar mitjà $\overline{w}(t)$ a temps t. Com que no és gaire versemblant pretendre tenir una mesura acceptable de $\overline{w}(t)$ i menys fer-ne prediccions, ja pot quedar clar de bell antuvi que el criteri de sostenibilitat que proposo només pretén servir de referent conceptual, encara que l’expressi en llenguatge matemàtic. El següent punt a considerar és que no n’hi ha prou amb què cada generació es preocupi de la que ve a continuació, ja que el que faci també en pot afectar d’altres, a causa de la multiplicitat d’escales temporals. No només pot posar més difícil a les generacions següents que mantinguin els nivells de benestar no decreixents, sinó que la pròpia idea dels nivells de benestar no decreixents no és prou satisfactòria. Prenguem el cas dels residus radioactius, que poden durar centenars, milers o milions d’anys. Com que el problema queda tan repartit al llarg del temps, el fet que temporalment s’empri energia nuclear no dificulta gaire el manteniment d’uns nivells no decreixents de benestar. Però, d’altra banda, la generació que l’ha emprat pot haver estat destruint molt més benestar (futur) que el que produïa, i en aquest sentit es pot dir que ha seguit una estratègia insostenible. Aleshores la variable que cal sostenir no és el benestar a cada generació, sinó una variable f que expressi el benestar mitjà esperat per a la generació i totes les següents (que és una manera d’entendre la “capacitat del nostre món com a suport per al benestar”, a totes les escales temporals):

$$f(t)=\lim_{c\to\infty} E\left[\frac{\int_t^{t+c}\left(N(\boldsymbol{t})\overline{w}(\boldsymbol{t})\right)d\boldsymbol{t}}{\int_t^{t+c}N(\boldsymbol{t})d\boldsymbol{t}}\right],$$

on N és la població i E és l’esperança matemàtica. La condició de sostenibilitat que proposo és, per tant

$$\frac{df(t)}{dt}\geq 0 \qquad (0.7).$$

Com que això no es pot quantificar, a la pràctica s’haurà de treballar amb moltes variables i, per tant, anirem a parar en darrer terme a una “sostenibilitat forta”, però concretant-la en criteris que puguem suposar, per sentit comú, que estan d’acord amb l’esperit de l’Eq. 0.7. Tot i que fa estona que ens hem allunyat de l’ortodòxia econòmica, ens podem plantejar de nou, ara en termes de benestar real enlloc de termes monetaris, el per què parlar de sostenibilitat enlloc d’optimització. Doncs resulta que l’optimització consistiria en:

$$f=\max \qquad (0.8).$$

A diferència del que succeeix amb altres definicions de sostenibilitat, en aquest cas la sostenibilitat (Eq. 0.7) és ordinàriament una condició necessària per a l'optimització (Eq. 0.8). És clar que el millor sempre és optimitzar, però podem decidir que la sostenibilitat és el mínim acceptable, així com, en un context ètic més ampli, sovint s'admet que el millor és fer el màxim de bé però el mínim acceptable és no fer més mal que bé. Estem tan lluny d'assolir la sostenibilitat que podríem estar prou contents/es d'assolir, si més no, aquest mínim acceptable.

Fins ara hem estat parlant de com s'hauria de plantejar una economia molt més orientada al bé comú, és a dir, amb uns objectius molt més igualitaris que els de l'economia convencional, que està feta a la mesura dels interessos particulars. Ara bé, és realment viable assolir i mantenir una economia gaire igualitària? (sec. 6.5). Els sistemes econòmics són sistemes complexos que s'autoorganitzen subjectes a restriccions físiques que tot just comencem a entendre, i no són totalment maleables, no podrem organitzar el sistema econòmic de qualsevol manera que puguem imaginar. Hem vist la gran importància per als ecosistemes de la criticalitat i la SOC, i s'ha començat a acumular evidència que són igualment importants en els sistemes econòmics. El propi descobridor de la SOC, Per Bak, ha escrit que potser no tenim cap opció millor que el capitalisme, per la seva presumpta dinàmica SOC. Una de les propietats escalants que fan pensar en criticalitat o SOC és justament la distribució potencial dels ingressos, que ja fou observada per Vilfredo Pareto el segle XIX, i que posa de manifest les immenses desigualtats que hi ha. Afortunadament, no crec que en l'estadi actual de coneixements es pugui asseverar que la desigualtat sigui una conseqüència inevitable de les lleis de la física. Cert que sembla haver-hi molts exemples importants de criticalitat/SOC i que són interessants perquè permeten fer moltes prediccions, però també hi ha molts exemples de sistemes que no són SOC. És clar que, si fem cas d'una proposta de Chris G. Langton que va tenir molta repercussió, podem pensar que els sistemes capaços de processar informació i d'adaptar-se tan sols poden existir "entre l'ordre i el caos", a la vora d'algun estat de criticalitat, SOC o similar, estat amb la flexibilitat justa, en què els senyals tant poden propagar-se molt com poc. Però es pot afegir immediatament que un sistema pot ser crític/SOC en un aspecte i no ser-ho en un altre (per exemple, pot ser que la selva tropical sigui crítica/SOC en termes de diversitat però no de foc). Si la distribució d'ingressos és crítica, hi haurà un cert dinamisme en l'accés dels individus a diferents nivells de control sobre l'economia, en funció de les seves actuacions, que facilitarà l’adaptabilitat del sistema, però l’adaptabilitat també pot tenir altres orígens. Prenguem per exemple el model econòmic imaginat per l’economista ecològic Hermann Daly, en el qual hi hauria un lliure mercat però també un límit inferior i un de superior sobre els ingressos, prou separats perquè hi hagués un incentiu a funcionar segons els senyals del mercat (amb els mecanismes adients per corregir els forts biaixos entre els senyals del mercat i allò que és desitjable per a la col· lectivitat), però també prou junts perquè no arribessin a haver-hi diferències de classe. Hem de suposar que aquestes restriccions no acabarien amb l’adaptabilitat, sinó que aquesta es canalitzaria d’altres formes. Per exemple, les estratègies i pautes de conducta econòmica de cada individu es copiarien de l’un a l’altre i, si miréssim la distribució d’abundàncies d’individus que s’ajustessin a diferents estratègies, segurament trobaríem una distribució potencial com la de les abundàncies d’espècies, indicant criticalitat. Aquestes abundàncies serien flexibles i sotmeses a “pressions selectives”, de manera que farien el sistema adaptable. Les institucions públiques i altres circuits de decisió externs al mercat també tenen els seus propis mecanismes de processar informació que, poc o molt, funcionen.

Un aspecte en que tant el sistema capitalista com el sistema d’en Daly podrien ser SOC és el de les crisis econòmiques. Si és així, les formes convencionals de tractar-les (ja siguin neoliberals o keynesianes) no sembla que hagin de ser gaire efectives a llarg termini. Les crisis podrien ser com un filtre que s’endú empreses i conjunts d’empreses entrelligades que han esdevingut massa fràgils, i que altrament podrien fer que el conjunt del sistema econòmic es tornés també trencadís. L’únic que sembla que s’hi pugui fer és evitar d’engrandir-les més del compte (amb polítiques “procícliques”, com les que l’FMI ha estat acusat d’aplicar a alguns països d’ingressos baixos), posar unes bones xarxes socials de seguretat pels afectats/des (en el sistema d’en Daly està clar que hi serien) i reemplaçar en una certa mesura la funció de les crisis a base de promoure, en el dia a dia, característiques de les empreses i conjunts d’empreses que les faci més robustes.

Si volem avançar en la justícia social i la sostenibilitat ambiental ens calen uns bons mecanismes socials de regulació, aliens al mercat, que facin que el sistema s’acosti a un òptim entre igualtat i incentiu (que, personalment, imagino molt decantat cap a la igualtat, en comparació amb el panorama actual) i garanteixin que els incentius afavoreixin les conductes més favorables al bé comú. No crec que puguin prendre la forma d’institucions monolítiques, perquè això els faria inestables i corruptibles. El que caldria és una multiplicitat de circuits de decisió de diferents formes, a totes les escales, amb una forta base democràtica, molt dinamisme, funcions redundants, i capacitat per controlar-se mútuament. No tenim uns motius sòlids per pensar que no valgui la pena lluitar per objectius socials i ambientals (que "un altre món" no sigui possible), però sospito que s'haurà de fer sense pretendre posar punt i final al joc inacabable que té lloc "entre l'ordre i el caos".

Ara mateix, un dels esculls més evidents que tenim al davant és l'avanç de la globalització neoliberal. Cloc el llibre amb uns pensaments sobre la globalització (sec. 6.6), no sense abans donar una breu ullada a la magnitud actual i la naturalesa de les desigualtats internacionals en l'apropiació dels recursos naturals i els "serveis" dels ecosistemes. Malgrat que té moltes limitacions, faig servir com a indicador la "petjada ecològica". A efectes pràctics, la petjada ecològica ve a ser una mesura de la part de la producció primària de la Biosfera que cada persona o grup humà consumeix, ja sigui producció actual o acumulada en forma de combustibles fòssils. Si mirem la probabilitat que una persona escollida a l'atzar visqui en un país amb una certa petjada ecològica (plantejament que deixa de banda la desigualtat interna de cada país i en manté tan sols la component internacional) ens trobem, un cop més, amb una distribució estadística pròxima a la potencial (Fig. 6.1A). Les diferències en el consum d'energia externa són les que més contribueixen a que la distribució prengui aquesta forma, però la magnitud de la desigualtat no és gaire menor per a les altres components de la petjada. Aquestes desigualtats posen el marc de la globalització actual i tenen molt a veure amb un procés històric anterior que també podríem anomenar de "globalització", en un sentit ampli.

A més de la globalització per la via del mercat, ja fa temps que hi ha en marxa una altra vessant de la globalització, que és el pas d'unes activitats econòmiques amb impactes ambientals locals a d'altres per als quals els impactes són inherentment globals (com és el pas d'usar carbó vegetal, que repercuteix sobretot en els boscos locals, a usar combustibles fòssils, amb les conseqüències que tenen per al clima global).

Com diu Margalef, el transport permet combinar recursos que estaven segregats en l'espai: això pot fer augmentar el potencial econòmic (que és bo o dolent, segons com es faci servir), però també fa un curt-circuit sobre els mecanismes de regulació local, tan ecològics com socials (subsec. 6.6.3). Malauradament, això succeeix sense que s'hagin desenvolupat uns mecanismes socials de regulació global per garantir la justícia social i per anticipar els mecanismes ecològics de regulació global. A més, la superposició en els recursos accessibles per a cada individu, en desaparèixer la barrera de la distància, és comparable a la superposició de nínxols entre diferents espècies, de manera que deu fer eixamplar les fites a la distribució presumptament potencial en els recursos usats per cada individu (com també les de la distribució potencial de les grandàries de les ciutats), i per tant la desigualtat. Quan algun cop s'ha proposat alguna fórmula per recuperar, si més no, un cert poder dels estats (que és quasi l'únic nivell al qual hi ha a vegades algun mecanisme democràtic), a base de permetre'ls rebutjar la importació de productes fets sota estàndards socials o ambientals inferiors als del país destinatari, s'ha temut que esdevingués una excusa per a un proteccionisme esbiaixat sistemàticament contra els països de pocs ingressos. Però aquest problema es pot resoldre si es reconeixen i es miren de separar dos efectes que aquí es barregen: l'efecte de les limitacions que depenen directament dels ingressos per càpita i l'efecte d'interferència entre circuits de regulació (que també apareixerà enter països amb el mateix nivell d'ingressos). És sobre el segon, que els estats haurien de tenir capacitat d'intervenir.

Una altra faceta de la globalització és que els intercanvis internacionals condueixen a un augment en el grau d'especialització de les persones, i permeten que també els països s'especialitzin (subsec. 6.6.4). Això també ajuda a augmentar el potencial econòmic, per bé o per mal, i també té conseqüències no trivials pel que fa a desigualtat i implicacions incertes per a l'estabilitat global. Tot i l'optimisme d'alguns models econòmics simplistes, hi ha bons motius per pensar que aquests intercanvis tendeixen a generar desigualtat i perjudiquen netament una part dels/les participants. Per aquests i altres motius (com ara el "deute ecològic" que els països d'ingressos alts ja han acumulat envers els d'ingressos baixos), els acords econòmics internacionals només poden ser justos si inclouen provisions que afavoreixin per norma els països d'ingressos baixos, en particular mantenint-los un bon marge per adoptar un proteccionisme selectiu en front dels països d'ingressos alts.

Encara hi ha més problemes: el procés d'uniformització ràpida en molts aspectes (subsec. 6.6.1), com ara en afavorir-se les invasions biològiques, així com els costos ambientals del transport en si (subsec. 6.6.5): cost energètic, dificultat per tancar els cicles de materials i fragmentació de l'hàbitat per part de les infrastructures que el fan possible. Això no vol dir que la millor escala espacial sigui la mínima. En una economia més orientada a la sostenibilitat, l'especialització que sorgeix quan l'economia funciona a una escala espacial gran permetria un guany en l'eficiència física en l'ús d'alguns recursos, que compensaria els problemes anteriors fins a cert punt. És a dir, cal buscar una escala òptima per a cada activitat econòmica, que no té per què ser ni la màxima ni la mínima possible. Això sí, cal una bona dosi de prudència en front dels augments excessius d'escala, ja que poden generar dependències que facin més difícil la marxa enrera. I, sobretot, calen unes fortes pressions socials per aconseguir uns mecanismes socials de regulació global que orientin el món per la via de la justícia social, la sostenibilitat ambiental i el respecte per la resta de la Biosfera.

Agraïments

A nivell crematístic no tinc gran cosa a agrair. Un requeriment habitual per a les beques és l'integrar-se en algun grup de recerca finançat, cosa que limita molt la possibilitat de fer treballs gaire innovadors, sobretot si van a contracorrent. Malgrat la manca de mecanismes que poguessin compensar-me, he pensat que aquesta feina tenia prou interès col·lectiu per tirar-la endavant apanyant-me-les com pogués, i així ho he fet. L'únic ajut oficial que vaig tenir va ser una modesta beca d'"acabament de tesi" (ejem...) de la Universitat de Barcelona el curs 1997-98, una part de la qual he retornat en forma de "matrícules de vinculació", que graven en concret la gent sense beca.

Per sort, a nivell intel·lectual i a nivell afectiu sí que tinc moltíssim a agrair. I començaré per aquelles i aquells amb qui he tingut interaccions directament relacionades amb aquest treball.

Vull fer arribar el meu agraïment al meu director de tesi, Jordi Flos. Hauria estat difícil trobar-ne un de més apropiat. Poques persones haurien estat capaces, com ell, d'entendre sense problemes i acceptar sense manies fins els meus esborranys més enrevessats, amb raonaments complicats, a vegades mal explicats, equacions, salts amunt i avall en l'arbre del coneixement i violacions flagrants de les ortodòxies que el poblen, i després sorprendre'm amb comentaris imaginatius, profunds i totalment adients. I poques persones haurien estat capaces d'afinar tan bé la mesura en què havien d'intervenir com a director: de seguir el que feia, fer bones aportacions, donar-me ànims, recordar-me que la tesi s'ha d'acabar, però no posar cap mena d'obstacle a què tirés endavant amb les meves pròpies idees.

Estic agraït també a les altres persones que han llegit esborranys preliminars de parts de la meva tesi i me n'han fet comentaris interessants i, quasi sempre, animadors. De moment (per ordre alfabètic): David Alonso, Oriol Bosch, Juan Camacho, Emilia Gutiérrez, David Jou, Ramon Margalef, Joan Lluís Pretus, Miquel Àngel Rodríguez-Arias i Francesc Sagués. Al Miquel Àngel també li agraeixo la seva disposició permanent a donar un cop de ma en el que calgués, i la seva capacitat inigualable per insuflar entusiasme. Moltes gràcies al Quimet Colom i a la Sílvia per haver-se ofert a aportar el seu treball a aquest projecte, si hagués calgut. Gràcies a l'Ester Clavero per resoldre'm uns dubtes taxonòmics. Agraeixo també l'interès i els comentaris de (ordre alfab.) Joan Armengol, Jordi Bigues, Joan Garcia, Carles Gràcia, Josep Anton Morguí, Bruce D. Malamud, Ricard V. Solé, Mikel Zavala, i de molta altra gent que esmento en els darrers paràgrafs (o no). A en Joan Armengol també li agraeixo el fet d'haver estat el primer a tenir-me la confiança d'encarregar-me un treball d'"ecologia teòrica aplicada", gràcies al qual vaig poder posar un peu en el Dept. d'Ecologia de la Universitat de Barcelona, on després he seguit treballant en aquesta línia. Estic agraït, lògicament, a les persones i institucions que han cedit dades que he emprat en aquest treball, o que m'han ajudat a aconseguir-les: Oriol Bosch, Manoel Cardoso, Peter J. Morin, Joan Lluís Pretus, William H. Romme, la Direcció General de Biodiversitat del Dept. de Medi Ambient del Govern Balear, en la persona de Gabriel Bardi, el Servei d'Agents Rurals i Prevenció d'Incendis de la Generalitat de Catalunya, en la persona d'Esteve Canyameres, el Yellowstone Center for Resources (Yellowstone National Park, NPS), en la persona d'Ann Deutch. I també a les persones que han organitzat, assistit i fet aportacions a les xerrades que he fet donant a conèixer alguns punts de la tesi: al Dept. d'Ecologia de la Universitat de Barcelona, a classe d'Ecologia Teòrica de la Fac. de Biologia de la Universitat de Barcelona, a la Universitat Internacional de Menorca – Illa del Rei, al Centre de Recerca Ecològica i Aplicacions Forestals, al Dept. de Física Estadística de la Universitat Autònoma de Barcelona i al centre social okupat de Can Masdéu.

Als meus pares els tinc tanta cosa a agrair que no acabaríem. Per començar, suposo que sense ells mai no hauria arribat a estar tan sonat com per embarcar-me en tot això. També tinc molt a agrair a la resta de la família, vius/ves i difunts/es. Cal fer una menció especial a la gata Psiquis, amb qui m'unien uns lligams emocionals molt forts en els primers temps de la tesi, i que aleshores era la qui feia un seguiment més estret de l'avanç de la meva feina.

En els darrers temps d'aquest treball, he gaudit d'un ambient de feina que no podia ser més engrescador, un ambient de feina inoblidable i irrepetible, perquè ha estat molt més que un ambient de feina, gràcies a les riualleres (per ordre alfab.): Blanca, Caro, Mireia, Núria, Rosa i Tura; als riuallers: Cesc, Marc i Toni; i l'altra gent amb qui també he compartit darrerament la sala de treball, reminiscent del cambrot dels Germans Marx, en la qual s'ha gestat majoritàriament aquest llibre (per ordre alfab.): Ester, Juanma, Fiona, Marta, Olga, Pere, Sergi i Xavi; així com altres persones del Dept. amb qui he anat interaccionant i que encara no he esmentat (per ordre alfab.): Ainhoa, Andrea, Bernat, Bet, Biel, Chechu, Cristina, Eugènia, Fede, Guillermo, Helena, Jaime, Jaume, Javier, John, Jordi, Jose, Laia, Maria, Marc, Miguel Ángel, Montse, Narcís, Nuri, Oliver, Paco, Quico, Quique, Rafa, Rita, Sam, Sergi, Susana, Teresa A., Teresa B., Vicenç, Willy i Xavi.

I a totes les persones que han format part del meu entorn afectiu d'una manera més estreta, mentre feia aquest treball, amb les quals he tingut sovint intercanvis actius d'idees, a més d'emocions, i que han estat vitals perquè pogués tirar endavant. Unes quantes ja estan escampades pels paràgrafs anteriors, però també voldria esmentar (per ordre alfab.): Ana, Andreu, Anna, Bego, Carles, Carlos, David, Eli, Elisenda, Eulàlia, Eva (gràcies per ses ales de canya!), Ferran, Hannah, Iola, Israel, Joan, Joann, Jordi, Laia, Laura, Lluís, Manu, Montse, Maria, MªJosep, Núria, Nurieta, Olga, Òskar, Paulo, Ricard, Roger, Roser, Sílvia B. i Sílvia E. He estat a la vora d'allargar molt aquesta llista amb altres persones per qui també sento afecte, però ho he deixat bàsicament amb les que crec que són o han estat més íntimes (per ara).

I, especialment... ευχαριστω τη Ναντια, γιατι ο, τι μου εδωσε, ειναι καλυτερο απο ολες τις υποτροφιες του κοσμου (espero que tingui sentit).

Chapter 1

Irreversibility in biological processes

1.1. Introduction

Some of the most ambitious theories on the development of ecosystems across time, on biological evolution and also on ecological economics rely on the tentative application of thermodynamic evolution criteria to large-scale high-complexity self-organized systems (see chapt. 2 and 6). If such theories were to hold, many features of the evolution of the Biosphere would turn out to be an unavoidable consequence of the laws of physics, so they would be inescapable and would rule in any imaginable biosphere in any other planet. The future of our own Biosphere, human society included, would become predictable.

It should be stressed that such applications of thermodynamics are just tentative, because they are never the result of a rigorous deductive chain starting from the laws of thermodynamics. One of my main purposes when beginning to work in this book was to follow such a deductive chain to see which was the result. This chapter offers the first steps in the chain. They lead to a physical justification for some of the most basic features that make biological processes time-asymmetric (i. e. it explains in physical terms why a movie with living characters does not look the same when projected backwards). The chain is followed up in chapter 2 (see in particular the end comment in subsect. 3.2.1), which deals with directionality in ecological succession and evolution, and chapter 6 (sect. 6.2), which deals with ecological economics. The bulk of these chapters (and also of the others) can be understood without previously reading chapt. 1, so this can be skipped by readers mainly interested in ecological or economic applications.

It is generally accepted that the Second Law of thermodynamics is the ultimate source of time asymmetry in macroscopic systems, so it is completely reasonable to search for a thermodynamic explanation for any irreversible process observed in biological systems. However, the connection between thermodynamic and biological irreversibility is far from trivial, except for gross energetic balances and for particular biochemical reactions. An excellent introduction to these simple applications of thermodynamics to biology is offered by Jou (1985).

Long ago, a paradox was perceived between the Second Law and the observed increases of complexity across development of single organisms and evolution of life as a whole. This happened because the Second Law postulates that isolated systems can only increase their entropy across time, and entropy can be taken as a measure of disorder. Schrödinger (1944) made it clear that there was no such paradox, because living systems are not isolated and their decreases in entropy are more than compensated by coupled entropy increases in their environment. This same explanation applies to other ordered structures, such as crystals.

When systems are not isolated, their state of equilibrium does not correspond to an extreme value for its own entropy, but an extreme value for some potential that captures the entropy of both system and environment: minimum Gibbs free energy, minimum Helmholtz free energy, etc., depending on contour conditions. The order of crystals can be explained because minimum free energy does not coincide with maximum entropy. The case of life is more involved, because living systems are never in a state of thermodynamic equilibrium, due to the exchanges of energy and matter with the environment.

There is no consensus on the definition of entropy and other thermodynamic variables out of equilibrium. The usual solution is assuming local equilibrium (Nicolis and Prigogine 1977): the studied system is assumed to be made up of small parts close enough to equilibrium for the thermodynamic variables to be locally well-defined. Global entropy is then taken to be the sum of local entropies.

For each small part of the system, which here I call cell, a local balance of entropy can be established:

$$dS = d_i S + d_e S \qquad (1.1).$$

d_iS is a term for "entropy production". It consists of the variation of entropy in the cell plus the coupled variation of entropy in its environment. d_eS is a "flow of entropy", which captures the variation of entropy in the environment coupled to the processes in the cell (with changed sign). When both are summed up, we thus have the cell's local entropy variation dS. The Second Law is expressed as

$$\frac{d_i S}{dt} \geq 0 \qquad (1.2),$$

so $\frac{dS}{dt}$ can either be positive or negative.

The production of entropy satisfies:

$$\frac{d_i S}{dt} = \sum_i \boldsymbol{h}_i \Gamma_i \ .$$

$\{\Gamma_i\}$ are “thermodynamic forces”, which are related to the distance to equilibrium for different variables (e. g. the affinity of a given chemical reaction), and $\{\boldsymbol{h}_i\}$ are the conjugate “thermodynamic flows”, which measure the rate of change in these same variables (e. g. the rate of the chemical reaction). At thermodynamic equilibrium, all the forces and flows are null, and so are entropy production, entropy flow and local entropy variation.

A way to take a system out of equilibrium consists of imposing from outside a non-null value for one or several of the forces. By application of a well-known mathematical technique called Taylor series expansion (see e. g. App. 5D), it can be expected that, if the distance from equilibrium is small, the relations between forces and flows will be linear:

$$\boldsymbol{h}_i = \sum_j L_{ij} \Gamma_j \quad \forall i \qquad (1.3).$$

Furthermore, since at equilibrium everything is time-symmetric, close to equilibrium the interactions are expected to be reciprocal:

$$L_{ij}=L_{ji} \quad \forall i, j \qquad (1.4)$$

(Onsager reciprocity relations, see de Groot and P. Mazur 1962, Nicolis and Prigogine 1977).

Prigogine (see Nicolis and Prigogine 1977) proved that, if a system is kept out of equilibrium by fixing one or several of the forces, but close enough to equilibrium for Eqs. 1.3 and 1.4 to hold, the system will tend to achieve a minimum entropy production. Entropy production thus plays close to equilibrium a role similar to free energy at equilibrium.

If the distance to equilibrium is increased enough, the relation between forces and fluxes will become non-linear and then the system may acquire a non-trivial spatio-temporal organization (Nicolis and Prigogine 1977). Well-know examples are the convective cells of Bénard and the oscillating reaction of Belousov-Zhabotinsky. The patterns that result from self-organization far from equilibrium are called dissipative structures. Life can be considered an extremely complex instance of dissipative structure in a planet that is kept far from equilibrium mainly by the input of solar radiation and the output of longer wave radiation to the outer space, which, together, represent a permanent outflow of entropy.

While irreversible thermodynamics has allowed to better understand the origin and the conditions for self-organization, it has rendered few predictions on the properties of self-organized systems. The predictive capacity of far-from-equilibrium thermodynamics is still far from comparable to that of equilibrium or close-to-equilibrium thermodynamics. Some of the main advances in this sense have been accomplished by the school of “extended thermodynamics” (Jou et al. 1993), which has found ways to go beyond the stringent assumption of local equilibrium.

Here I explore another path to further advances. The basic idea is that some general features of the complex structures that arise far from equilibrium can be expressed in information theoretic terms, and that information processing is subject to constraints that derive from the laws of thermodynamics. So the processes of self-organization far from equilibrium have in fact some general features that are imposed by the laws of thermodynamics.

I develop this approach in sect. 1.2 and I put forward some biological applications in sect. 1.3.

1.2. Thermodynamics of mutual information

1.2.1. Preliminaries

"Let us suppose that a vessel is divided into two portions, *A* and *B* by a division in which there is a small hole, and that a being who can see the individual molecules opens and close this hole, so as to allow only the swifter molecules to pass from *A* to *B*, and only the slower ones to pass from *B* to *A*. He will, thus, without expenditure of work raise the temperature of *B* and lower that of *A*, in contradiction to the second law of thermodynamics" (Maxwell 1871, quoted in Brillouin 1962). Under the charm of the Demon of Maxwell (see Leff and Rex 1990), Szilard (1929) opened the path toward a thermodynamics of the flows of information. He proposed that the Demon's trick lies in that the observation of the molecules has a thermodynamic cost and it is high enough to 'rescue' the Second Law. In addition to its acknowledged interest for other reasons, I consider that a deep understanding of this matter is of fundamental importance if irreversible thermodynamics is to fully embrace biological phenomena, where flows of information are the essential ingredient.

Szilard's idea was further developed by Brillouin (1962) but later denied by a set of authors who assert that it is erasure of information and not measurement that has a thermodynamic cost. The last is considered to be the contemporary view (Landauer 1961, 1988, 1997; Bennet 1982, 1987; Zurek 1989a, b; Caves 1993). It is a result of the application of algorithmic information theory.

The algorithmic information content $K(s)$ of a string of data s is defined as the size $|s^*|$ of the minimal program that, when executed on a universal computer, yields output s (Zurek 1989a, b):

$$K(s) \equiv |s^*| \qquad (1.5).$$

This information measure is an alternative to the better-known Shannon "entropy" (Shannon 1948, reprinted in Shannon and Weaver 1963),

$$H(X) = -\sum_i P(x_i)\log_2 P(x_i) \qquad (1.6),$$

where X is a random variable and $\{P(x_i)\}$ is a set of probabilities. This is not a measure of entropy in the physical sense of the word, but a statistical measure that can be applied to any random variable, irrespective of its meaning.

Consider a variable Ξ whose possible values $\{x_i\}$ correspond to each possible microstate of a physical system given its macrostate. Then Eq. 1.6 displays an immediate relation with a well-known measure of physical entropy S, Gibbs entropy:

$$S_{(G)} = k\log(2)H(\Xi) \qquad (1.7),$$

where k is Boltzmann's constant.

For a large class of distributions $\{P(x_i)\}$, the average algorithmic information $\sum_i K(x_i)P(x_i)$ closely matches Shannon entropy (Eq. 1.6). Bennet (1982) took advantage of this fact and the relation that I expressed through Eq. 1.7 and proposed a measure of physical entropy that we could express as

$$S_{(a\lg)} = k\log(2)K(x).$$

A peculiarity of $S_{(\mathrm{alg})}$ as compared to other measures of entropy is that it is well defined for single microstates x instead of ensembles of microstates. Zurek (1989 a, b) showed how a reversible deterministic evolution could lead to an increase of $S_{(\mathrm{alg})}$.

Bennet (1982) translated the Maxwell Demon challenge to a simpler setting: the measured system has two possible states, *R* and *L* (from 'right' and 'left'). In correspondence, the measuring device has the possible states *R* and *L* plus an erased standard state *O*. He established the following cycle:

1. The measuring device passes from *O* to either *R* or *L* depending on the state of the measured system. Neither before nor after this step the device contains any additional information (that was not already in the measured system), so there is no change in entropy.
2. The measured system may change its state. The information in the device ceases to be sufficient to know the state of the measured system, hence the information content of device and system taken together increases. This is an increase of entropy that can be used to extract work.
3. The device passes to *O* in order to be able to perform a new measurement. There is a loss of information, because in this case the initial state of the device did not depend on the observed system and it cannot be known any more. This is a decrease of entropy and has a thermodynamic cost.

After the study of this cycle, Bennet asserts that it is erasure (step 3) and not measurement (step 1) that has a thermodynamic cost, in consonance with earlier findings by Landauer (1961).

As will become more apparent in subsect. 1.2.3, both the Szilard-Brillouin and the algorithmic approach require that the joint entropy of device and measured system be not calculated as a sum of their entropies. On this grounds I explore the thermodynamic meaning of mutual information, the information-theoretic measure that explicitly captures such discrepancy with the simple extensive case.

With this tool on our hands, we may try to answer several questions. I study again where the true thermodynamic cost lies in a cycle like Bennet's and reach a conclusion different than the accepted at present. I also analyze how to introduce mutual information terms into irreversible thermodynamics. I finally provide an explanation for the coincidence between the arrows of time of thermodynamics on the one hand and memory and causality on the other. In sect. 1.3 I discuss the biological consequences.

1.2.2. Mutual information

Two well-known properties of Shannon "entropy" (Eq. 1.6) are (Wherl 1978):

1. Subadditivity: the joint entropy of two random variables cannot exceed the sum of their entropies,

$$H(X,Y) \le H(X)+H(Y) \qquad (1.8).$$

2. Additivity: in the particular case of independent random variables, the joint entropy is the sum of each variable's entropy,

$$H(X,Y) = H(X)+H(Y) \qquad (1.9).$$

These two properties allow to quantify the degree of correspondence between two variables by means of mutual information *I* (Shannon 1948, reprinted in Shannon and Weaver 1963; for examples with ecological variables see Flos and Carbonell 1991, Wagensberg et al. 1991):

$$I(X:Y) = H(X)+H(Y)-H(X,Y),$$

It can also be expressed as

$$I(X:Y) = H(X)-H(X|Y) = H(Y)-H(Y|X) \qquad (1.10),$$

where the conditional entropy $H(Y|X)$ has the form

$$H(Y|X) = -\sum_{ij} P(x_i)P(y_j|x_i)\log_2 P(y_j|x_i) \qquad (1.11).$$

Since Gibbs entropy is formally a particular case of Shannon entropy, it has the same properties above stated and the measure of mutual information can be defined. On the other hand, physical entropy is usually taken as extensive. Were this strictly right, Eq. 1.8 would always reduce to Eq. 1.9. However it has been claimed since long ago that it does not (Wherl 1978). As shown in subsect. 1.2.3, the last must be the case if the Demon of Maxwell is to keep exorcised in any of the possible ways we know to the date.

Thus we can introduce a thermodynamic version *J*(*A*:*B*) for the mutual information between two systems *A* and *B*, such that

$$S(A,B) = S(A)+S(B)-J(A:B), \qquad (1.12),$$

where

$$J(A,B) = k\log(2)I(\Xi_A : \Xi_B) \qquad (1.13).$$

In this way a nonlocal term explicitly enters the measure of entropy.

The relation of Eq. 1.12 with Zurek (1989a, b) equation for physical entropy is analyzed in App. 1A.

Eq. 1.12 can be applied to other definitions of physical entropy than Gibbs's. E. g.:

- The measure of mutual information is not only well defined for the Shannon but also for the algorithmic (Zurek 1989a, b) and the quantum (Cerf and Adami 1997) versions of information theory. It can be used for their corresponding physical entropies.
- We can use Eq. 1.12 in terms of the Boltzmann expression

  $$S_{(B)} = k\log(W),$$

  where W is the number of accessible microstates. It suffices to consider that the number of microstates W_{AB} accessible to two systems A and B taken together may be less than $W_A W_B$.
- Of course the last developments concerning Gibbs's expression are also useful for any version of physical entropy that preserves its Shannon-like form. E. g. when the microstates are subject to coarse graining or their set of probabilities is someway mapped to a different space of probabilities (see Coveney 1988).

It is not the same case when we establish the state of the two systems in some terms that have nothing to do with either microstates or any other set of 'minimal states' that enter any definition of physical entropy. However if the considered states do not overlap in terms of such minimal states, their mutual information is still thermodynamically meaningful in that it sets a lower constraint over $J(A{:}B)$. This seems quite obvious, but I give a demonstratetion in App. 1B. If we thus express the state of both systems by the variables X_A and X_B, we have

$$J(A:B) \geq k\log(2)I(X_A : X_B) \qquad (1.14),$$

hence

$$S(A,B) \leq S(A) + S(B) - k\log(2)I(X_A : X_B).$$

As a principle, this has to rule even for macroscopic features.

1.2.3. The cost of measurement revisited

Take a measuring device B applied to a system A. The following are the variations for each of the terms in Eq. 1.12 along Bennet's cycle (subsect. 1.2.1):

1. $S(B)$ increases, $J(A{:}B)$ increases, $S(A,B)$ does not change.

2. $S(B)$ does not change, $J(A{:}B)$ decreases, $S(A,B)$ increases.
3. $S(B)$ decreases, $J(A{:}B)$ does not change, $S(A,B)$ decreases.

$S(A)$ does not change in any step.

In this setting it is equally valid to consider a measurement cost in terms of $J(A{:}B)$ or an erasure cost in terms of $S(B)$. The interest of the erasure cost lies in that, in this particular case, the increase in J is paid by an increase in $S(B)$, and this decreases again at some point because the cycle is assumed to go on indefinitely. Thus the erasure cost is, at most, complementary to the measurement cost, so the criticisms on early authors who talked of measurement cost are not justified. Perhaps the concept of erasure cost is more useful for computer design, but is it universally valid? Let us investigate it with a set of variations on Bennet's cycle (subsect. 1.2.1).

First of all, we may consider the case in which the standard state is not defined. Bennet (1982) advanced that, in this situation, a measurement is accompanied by an implicit erasure of any previous information, so we may still charge the last for the thermodynamic cost (this is however not so evident if we work with Gibbs entropy instead of algorithmic entropy).

Second, we may consider the objection raised by Fahn (1996), namely that there could well be a cycle in which $S(O) > S(R)$, $S(L)$, in opposition to Bennet's implicit assumption $S(O)=S(R)=S(L)$. E. g. take the simple case $S(O) = S(\{R,L\})$. I. e. the entropy at the standard state equals the entropy at the nonstandard state, when no more distinctions are made in the last case. Then there is no net change of entropy in step 3 and the only decrease takes place at step 1, during measurement. It is possible to argue that in this case what counts is not the explicit erasure at step 3 but an implicit one at step 1 and charge the last for the overall thermodynamic cost of the cycle. For $S(O) > S(\{R,L\})$ the costly erasure would also be the implicit one, while for $S(O) < S(\{R,L\})$ both would share some amount of the cost.

Even if we accept the idea of implicit erasure, there is one problem left. Bennet assumes a sustained cycle where each measurement has been preceded by some other measurement, in which B already became saturated, making implicit or explicit erasure unavoidable. However, in the real world there are not only sustained cycles, but also transitories. These are equally subject to the Second Law, whose consequence is that J will always have a thermodynamic cost. However, in this case there will not necessarily be a cost of erasure. The “measurement cost” in terms of J should be considered more general than the “erasure cost”.

Unless a completely different solution is found for the paradox, we have to conclude that any definition of entropy of physical interest should involve nonextensivity and subadditivity. Hence the mutual information term J should be well defined and physically meaningful, and the 'cost of measurement' approach turns out to be right.

1.2.4. Irreversible thermodynamics of information flows

Consider two systems with well-defined roles of measured system (A) and measuring device (B). This means that the state of B is set as a deterministic function of the state of A, while the reciprocal does not hold. This is a unidirectional flow of information and can be expressed in theoretic-information terms as follows:

$$\begin{cases} I\left(\Xi_A(t) : \Xi_B(s) \middle| \Xi_B(t), \Xi_E(t)\right) \geq 0 \\ I\left(\Xi_B(t) : \Xi_A(s) \middle| \Xi_A(t), \Xi_E(t)\right) = 0 \end{cases} \qquad \forall t, s : s > t \qquad (1.15),$$

where: t , s is time; Ξ_A, Ξ_B and Ξ_E are the random variables expressing the microstate of the systems A and B and the environment and

$$I(X : Y|Z) = H(X|Z) - H(X|Y,Z) = H(Y|Z) - H(Y|X,Z).$$

Any unidirectional flow of information (Eq. 1.15) is incompatible with Onsager's reciprocity relations (Eq. 1.14), which apply at equilibrium and also define the close-to-equilibrium region (together with linearity) (de Groot and Mazur 1962). Hence such flows are exclusive of far from equilibrium systems. Maxwell Demons and any other measuring devices belong to far from equilibrium thermodynamics.

There is no consensus on the definition of entropy out of equilibrium, but the conclusions of subsect. 1.2.3 should apply to it. I. e., provided that it is additive, it should also be nonextensive and subadditive. In any case a term J satisfying Eq. 1.12 should be well defined and not constrained to zero.

Much of the progress in irreversible thermodynamics has relied on the assumption of local equilibrium (de Groot and Mazur 1962). Let us divide a spatially-extended nonequilibrium system in N cells of linear size d. The entropy of each cell C_i is phenomenologically estimated on the basis of equilibrium criteria. The entropy of the whole system is taken as a simple sum:

$$S = \sum_{i=1}^{N} S(C_i) \qquad (1.16).$$

This approach has been progressively extended by including more phenomenological traits (Jou et al. 1993). Is it possible to predict how should its final form look like? Nonextensivity and subadditivity lead us to postulate the following eventual substitute for Eq. 1.16:

$$S = \sum_{i=1}^{N} S(C_i) - \Psi_d \qquad (1.17),$$

where

$$\Psi_d = \sum_{i=2}^{N} J\left(C_i : \underset{j=1}{\overset{i-1}{\Theta}} C_j \right) \qquad (1.18)$$

and satisfies $\Psi_d \geq 0$. $\underset{j=1}{\overset{i-1}{\Theta}} C_j$ is the joint subsystem that we get by putting together the cells C_j such that $j<i$. The order of the cells in the recursion in Eq. 1.18 is indifferent.

The measure Ψ_d is related to Chaitin (1979) view of organization. In this chapter I call it in this way, although the mentions to “organization” in the rest of the book will not refer to this narrow meaning.

The variations of entropy have the form

$$dS = \sum_{i=1}^{N} dS(C_i) - d\Psi_d .$$

Therefore, for the entropy production in each cell to be nonnegative as required by the Second Law (Eq. 1.2), the local balance in Eq. 1.1 should be enlarged as follows:

$$dS = d_i S + d_e S + d_s S \qquad (1.19),$$

where the term d_sS is a "flow through scales". It quantifies local entropy increases (or decreases) coupled to equivalent increases (decreases) in the mutual information between the cell and its environment, in the same way that the term d_eS includes local entropy increases (decreases) coupled to equivalent entropy decreases (increases) in the environment.

E. g. if our "cell" is the measuring device in Bennet's cycle (*subsect*. 1.2.1), we find at each step:

1. $dS>0$, $d_iS=0$, $d_eS=0$, $d_sS>0$.
2. $dS=0$, $d_iS=0$, $d_eS=0$, $d_sS=0$.
3. $dS<0$, $d_iS=0$, $d_eS<0$, $d_sS=0$.

At step 1 the device absorbs entropy from the joint system organization. At step 3 it bombs it to the environment. At step 2 there is a positive entropy production, but it is not considered because it occurs at the joint system scale, not the local scale.

The balance for each cell is not enough. We should add the balance for the organization Ψ_d. We can also apply Eq. 1.19 in this case. Then the "flow through scales" passes to be bidirectional: it includes both variations in Ψ_d coupled to variations in the local entropies $S(C_i)$ and variations in Ψ_d coupled to variations in the mutual information between the whole system and its environment.

It is necessary to incorporate these refinements if nonequilibrium thermodynamics is to fully embrace biological phenomena, which cannot be conceived without flows of information. The quantitative importance of biological information in thermodynamic terms is analyzed in sect. 3.1.

1.2.5. The balance of organization and the memory arrow of time

In subsect. 1.2.4 I state the possibility of expressing the balance of a system's organization Ψ_d in entropic terms, by applying Eq. 1.19 to the system's organization instead of the local entropy. We can however express the content of this same equation in a different, more useful form:

$$\frac{d\Psi_d}{dt} = C - D + F$$

F is a simple flow term. It quantifies the flow from the environment to the system of structures that either are themselves organized or share some mutual information with the rest of the system. F is related to both d_eS and d_sS in Eq. 1.19.

C is what I call active coupling. It is the production of system's organization linked to an increase in some other component of the entropy of the system and/or its environment. It satisfies

$$C = \left(-\frac{d_eS}{dt} - \frac{d_sS}{dt} \right) - F .$$

D is the spontaneous dissipation of organization (which represents a production of entropy),

$$D = \frac{d_iS}{dt} .$$

C and F can be either positive or negative. D is constrained by the Second Law:

$$D \geq 0 \qquad (1.20).$$

Hence organization can only be produced by active coupling.

In the case of unidirectional flows of information (Eq. 1.15), the resulting organization is nonequilibrium (subsect. 1.2.4) and tends to dissipate. I. e. we can assert that for some long enough time lag τ,

$$\overline{D} = \frac{1}{t}\int_{t_o}^{t_o+t} D(t)dt > 0 .$$

Then active coupling is required not only to produce but also to steadily sustain organization (unless it is just imported from outside by having $F>0$).

Eq. 1.20 leads immediately to the memory arrow of time, by virtue of which memories can only remember the past (Hawking et al. 1993). Leaving random fluctuations aside (see subsect. 1.3.1), each bit of mutual information is the result of some present or past active coupling. It cannot arise spontaneously before some future active coupling. If we read measurement, observation, copy, inscription, etc. for active coupling, we get the arrow. This can also be taken as a statement of causality.

Philosophers of time have explained the memory arrow of time in a different way (Grünbaum 1973; see also Schlick 1948, Reichenbach 1956). It has been stated that:

- A "trace" left in a system by an interaction decreases its entropy.
- Since an entropy lower than the maximum cannot have arisen spontaneously, the existence of a trace allows us to infer that there has been an interaction in the past.
- We cannot perform any similar inference concerning future interactions.
- Hence such traces act as a memory of past interactions, but not of future interaction.

This theory has lead to some paradoxes. E. g., the signal left by a bomb may well be one of increased entropy and nevertheless it is useful as a memory of a past bombardment (Earman 1974, Horwich 1987).

My proposal is different in what follows:

- I do not expect any general rule about a memory device increasing or decreasing its entropy after the interaction. The relevant changes take place in mutual information, which acts as a nonlocal term for entropy and thus does not belong to the device.

- The measure of mutual information quantifies in an objective way the 'knowledge' that the memory device has about the source of information. I do not enter into subjective concerns on whether it is reasonable from the point of view of the device to infer anything about the true meaning of the traces it may lodge.

The above developments, in addition, leave without basis Popper (1956) rejection of the Second Law as the single source of macroscopic irreversibility.

1.3. Biological consequences

1.3.1. Thermodynamic weight of biological organization

The measure Ψ_d from subsect. 1.2.4. captures the thermodynamic weight of organization. Here I discuss the order of magnitude of this term when applied to biological systems, and I compare this approach with other approaches in the literature, which I show to be incorrect. The order of magnitude of Ψ_d is not only relevant from a biological point of view, but also for irreversible thermodynamics in general, because it allows to have an idea of the quantitative importance of including organization in thermodynamic balances.

A component of biological organization that is especially easy to quantify is that given by the ordering of monomers in polymer sequences. It is by no means the only one, but is illustrative. Let us consider that the N parts in Eqs. 1.17-1.18 correspond to the N polymers in an organism. When taking successive polymers as stated in Eq. 1.18, we will find in almost all of the cases that the sequences will be redundant with other sequences previously found, for $N \to \infty$. Though not exact, a good approach to the negative contribution of their mutual information to entropy per mol of monomer is simply

$$J = Ak\log(2)\overline{H} = R\log(2)\overline{H} \qquad (1.21),$$

where A is the Avogadro number, R is the ideal gas constant and $\overline{H}$ is the application of Eq. 1.6 to the relative frequencies of each kind of monomer (the same development can be performed in terms of algorithmic information).

An equation equivalent to Eq. 1.21 was already applied to this issue in the forties (Butler 1946; see also Morowitz 1968). However, some more recent approaches with a considerable impact are quite different and clearly wrong.

Battley (1993, see also Battley 1987, 1999) works with the bacteria *Escherichia coli* grown in a substrate of succinic acid. He takes water, reserve substances and ions out of the bacteria. He estimates the entropy of the remaining organic matter just as a function of its elementary composition. He obtains an entropy per unit mass that is even higher than that of the succinic acid. He concludes that biological complexity makes no significant negative contribution to entropy. He could not have found any other result, since he has always excluded it from his calculations.

At the other extreme we find Jørgensen (1992, 1994, 1999, 2001a; Jørgensen et al. 1995), a leading theoretical ecologist who works with exergy (see chapt. 2). "Exergy" or "availability" is defined in the engineering literature as the amount of work that can be extracted from a system when it is allowed to reach the state of equilibrium with its surroundings (e. g. Wark 1995). This author uses it as the basis for a presumed Fourth Law of Thermodynamics (Jørgensen 2001b). He distinguishes chemical from informational contributions to exergy. In his works, he finds that the informational part is from 1.3 to 715 times more important than the chemical one, depending on the kind of organism.

A simple argument shows that such numbers are unbelievable. It is accepted that each consumer organism preserves as such about 10% the chemical free energy it consumes (Margalef 1989). No known metabolic pathway uses the information content of the food as a fuel, so most of the information content of the consumer could only have its origin in the remaining 90% of chemical exergy. Since a large part of this 90% is dissipated, the informational exergy of the consumer should be much lower than 9 times the chemical exergy, never 715 times. The results below indicate that it is in fact some orders of magnitude lower.

The origin of such a huge discrepancy lies in the author's criteria. The following equation synthesizes his way to quantify an organism's exergy per gram:

$$Ex = \left(m - m_{eq}\right) - RT \log\left(20^{-CG} / N\right) \qquad (1.22),$$

where m is the chemical potential, m_{eq} is the equilibrium chemical potential, R is the ideal gas constant, T is the absolute temperature, C is the number of codoms per gene, G is the number of coding genes in the genome and N is the number of cells. 20 is the usual number of types of aminoacidic residuals in a protein. He takes $m - m_{eq} = 18.8$ $kJ\ g^{-1}$, which is a figure close to Battley (1993).

The second term in Eq. 1.22 is close to $-RT \log\left(20^{-CG}\right)$. According to Eq. 1.21, this is roughly the "informational" exergy of a mol of polymers, each one with a copy of the codifying part of the genome of the organism. The division by N is a strange maneuver with little impact. A mol of such polymers has no relation to the actual number of each kind of polymer that is found in 1 g of organic matter. As shown below, even the number of monomers per gram is much lower than one mol, so the implicit assumption of one mol of the above “idealized” polymers per gram produces a strong overestimation of the thermodynamic weight of organization.

Since neither Battley’s nor J∅rgensen approaches give any credible information on the order of magnitude of the thermodynamic weight of information in biological systems, I proceed to apply Eq. 1.21.

Take the bacteria *Escherichia coli* without water, salts and reserves, like Battley (1993). From its chemical composition (Neidhart 1987) I calculate the information content of each one of the main kinds of polymer. The results appear in table 1.1 (more detailed and accurate calculations concerning proteins can be found in Dewey 1997; here we are just interested in orders of magnitude). They make up a total of about 10^{22} bits per gram, which corresponds to $\Psi \approx 0.1$ $J\ g^{-1}\ K^{-1}$.

The entropy of this same organic matter without organization has been calculated to be about 4 $J\ g^{-1}\ K^{-1}$ (Battley 1993). The part of organization that we took into account subtracts about 2.5% from this amount.

In terms of free energy, the weight of organization is $T\Psi \approx 30$ $J\ g^{-1}$ at a temperature T = 298.15 K. Let us take as a reference the free energy of oxidation of the disorganized organic matter, which is about -20 $kJ\ g^{-1}$ (Battley 1993). The

	bits/monomer	μmol monomer/ g cell	bits/g cell
Protein	3.84	5265	7.84×10^{21}
RNA	1.98	653	5.18×10^{20}
DNA	2.00	104	8.33×10^{19}

Table 1.1. *(Redundant) information in the sequences of the polymers in 1 g of* E. coli.

calculated contribution of organization is 0.15% this amount (compare with J∅rgensen's 71500%).

Is organization by itself relevant for the thermodynamic balance of biological systems? 2.5% or 0.15% are small figures but not completely negligible, and they just cover a part of it (of course it is much more relevant in another sense, in that it determines which of the thermodynamically possible processes take place at each moment).

Is thermodynamics relevant for the evolution of organization in biological systems? It is absolutely relevant. The probability of finding all the polymers in a gram of organic matter completely ordered just by chance is

$$P(\Psi) = e^{-\frac{\Psi}{k}}.$$

For the value of ψ that we found, $P(\Psi) \approx 10^{-3x10^{21}}$. Biological organization lies well beyond the realm of thermic fluctuations.

$P(\Psi)$ decreases very fast as the amount of information increases. As few as 20 bits result in $P(\Psi) \approx 10^{-6}$. This makes the Second Law really valid even for the amount of organization that can be found at macroscopic scales.

1.3.2. Irreversible flows of information in biological systems

It is in biological systems where we can find most instances of asymmetric flows of information (with well-defined source and receiver of information), information processing and accumulation of memory. Hence the conclusions of subsect. 1.2.5 have their main applications in a biological context. From subsect. 1.2.5 we can remark that:

- Asymmetric information flows and accumulation of memory take place from past to future.
- Asymmetric information flows have an energetic cost, i. e. they have to be coupled to an outflow of entropy. This should be much larger than the entropy that corresponds to the amount of flowing information.
- The mutual information generated from asymmetric flows dissipates spontaneously.

Some macroscopic biological asymmetric information flows are:

- Acquisition of information on the environment by each organism.
- Regulation of the environment by organisms.
- Reproduction.
- Adaptation (flow of information from the environment to the genetic register).

All of these processes necessarily have to be coupled to an outflow of entropy above some well-defined value, produce memories of the past and produce some mutual information that spontaneously dissipates.

The dissipation of mutual information can take place either by fluctuations in the source or in the sink of information. E. g. in the case of neurobiological memories (i. e. memory in the original sense of the word), mutual information can either be lost because the organism forgets some information (fluctuation in the sink) or because this information becomes obsolete (fluctuation in the source). In the case of adaptation, it can be lost either by deleterious genetic modifications (fluctuation in the sink) or by environmental change (fluctuation in the source). “Unprogrammed” modifications in the organism seem to play an important role in aging (Lithgow and Kirkwood 1996), so aging is in part a direct consequence of the Second Law, which produces a loss of the mutual information that has been acquired through reproduction.

When two populations in a species become reproductively isolated, their characteristics diverge and they finally become two different species. This is also an instance of spontaneous dissipation of mutual information, and hence a manifestation of the Second Law.

Genetic features not strictly determined by natural selection will change across generations, in such a way that the “memory” of the genetic features of the ancestors will be irreversibly lost. This is the basis of Dollo’s law, which states that “evolution never repeats itself”. Brooks and Willey (1986, see also Weber et al. 1988) intuited that there was some relation between Dollo’s law and the Second Law of thermodynamics. They suggested a way to establish this relation and claimed to have found a “unified theory of biology” that would allow to fusion thermodynamics with evolutionary biology. However their approach is rather involved and speculative. The developments in this chapter allow to establish a much more solid and direct connection.

1.4. Discussion

Life is not comparable to gas in a recipient. Biological systems are historical systems, with active exchanges of information and accumulation of memories at all levels. Thermodynamics began by studying simple systems like gas in a recipient and produced simple evolution criteria, centered on some potential to be maximized or minimized. Biologists who have seek inspiration in thermodynamics have often tried to apply similar schemas to living systems; this is discussed in detail in chapter 2, where I show the general lack of solid foundations for such claims.

When searching for any solid consequence that thermodynamics may have for self-organized systems, the main result I have found is that it produces the memory arrow of time. The arrow by itself is not of course a novel finding, it was already well-known and is evident from everybody experience. The novelty is its derivation from the Second Law: the connection between both was already suspected by several authors but, to my knowledge, the previous formulations of this connection were more speculative than the one I present. Many readers are likely to find no interest in this, because, as the memory arrow of time was already known, it will not lead to novel predictions concerning biological systems. However, the fact that something is known does not mean that it does not have to be explained. Furthermore, if we are to follow the path from first principles to novel predictions, it is not strange that in the midway we find things known but unexplained: this is an indication that we are in the right way. Other authors interested in the thermodynamics of ecosystems missed these intermediate steps, so what they proposed were just unfounded hypotheses. These authors often put forward simplistic ideas that ignore features as essential for life as information processing, memory and history.

Perhaps the only thing I offer in this chapter for readers interested in novel biological predictions is the clarification of the thermodynamic weight that should be given to biological organization, at least because it allows to dismiss some purported predictive models by other authors.

While not so clear for predictive biology, the findings in this chapter should have interest for physics and philosophy of science, if correct. I do not only explain the memory arrow of time. I also modify the so-called “contemporary view” on the thermodynamics of measurement, erasure and computation. And I generalize the basic equations used in irreversible thermodynamics. On the other hand, the developments in subsect. 1.3.1 suggest that the new terms that I introduce into these equations might not have a large weight, so it remains to be seen whether this modification is to foster further progress.

The developments in this chapter open a door to holism, since I show that complete thermodynamic balances should necessarily include mutual information, which is nonlocal and can be defined at any scale.

1.5. Conclusions

I show that, unless a solution to the Maxwell Demon paradox completely different to the presently known is found, its "exorcism" requires any additive measure of entropy to be nonextensive and subadditive (definitions in subsect. 1.2.2 and justification in subsect. 1.2.3). A thermodynamic meaning should be given to a nonlocal term expressing the mutual information between different systems (Eq. 1.12).

Contrary to the contemporary view, any measurement does have a thermodynamic cost (subsect. 1.2.3). This is more general than the “erasure cost” with which some authors pretend to replace it.

Systems with asymmetric information flows can only exist far from equilibrium; this must be the case for Maxwell Demons and any measuring devices (subsect. 1.2.4). Hence the above results should apply out of equilibrium, irrespective of the way entropy is defined in these conditions. This carries me to modify some basic equations widely used in irreversible thermodynamics, by incorporating mutual information terms (Eqs. 1.17-1.19). It remains to be seen whether these modifications are going to be relevant for the future development of irreversible thermodynamics.

I develop some equations for the balance of organization in a system. The organizational version of the Second Law (Eq. 1.20) provides an immediate justification for the memory arrow of time (subsect. 1.2.5).

The tools I obtain are the appropriate to quantify the thermodynamic weight of biological organization (subsect. 1.3.1). Some of the main current criteria to measure it, such as Battley’s and J∅rgensen’s, are shown to be incorrect. The particular component of organization consisting of the ordering of polymeric sequences is shown to make a small but not completely negligible contribution to the entropy and free energy of living systems. It is indeed large enough to be subject to the Second Law (subsect. 1.3.1).

I consider that the thermodynamic constraints on information flows pose the main foundation for the deep time asymmetry observed in highly organized systems, such as the Biosphere. This is relevant for many processes, including organism-level information processing, aging, adaptation, and the divergences and irreversibilities of biological evolution.

I show in next chapter that the presumed trends in ecological succession and biological evolution seem to stand on this informational irreversibility rather than any other form of thermodynamic irreversibility.

Appendices

Appendix 1A. Comment on some other contributions about entropy and information

Zurek (1989a, b) defines physical entropy as a sum of "missing information" plus "known randomness":

$$S = H' + K' \qquad (1A1).$$

The "missing information" that I express as H' has a Shannon-Gibbs form, while the "known randomness" that I express as K' has an algorithmic form. Which is the relation of Eq. 1A1 with my Eq. 1.12?

Take an observed system *A* and an observer system *B*. From Eq. 1.10,

$$H(\Xi_A) = H(\Xi_A | \Xi_B) + I(\Xi_A : \Xi_B) \qquad (1A2).$$

If we give an algorithmic expression to $I(\Xi_A : \Xi_B)$ while leaving the Shannon expression for $H(\Xi_A|\Xi_B)$, what we get is equivalent to Eq. A1. This is useful from the observer point of view, whose own microstate can be taken as fully specified while the microstate of the observed system is not. For other uses, such as the equations developed in subsect. 1.2.4, which involve more than two systems, Eq. 1.12 is more appropriate.

Caves (1993) gets a puzzling result from Eq. 1A1. He claims that "a typical fine-grained phase-space pattern or a typical pure state has total entropy much greater than equilibrium". If we take Eq. 1A2 as a point of departure, what he does is:

- Expressing $I(\Xi_A : \Xi_B)$ in algorithmic terms and $H(\Xi_A|\Xi_B)$ in Shannon terms, following Zurek, though this is not essential.
- Coarse graining $I(\Xi_A : \Xi_B)$ and $H(\Xi_A|\Xi_B)$ with different criteria: coarser in the second case.
- Defining "equilibrium" as maximum $H(\Xi_A|\Xi_B)$.

As a principle, both $H(\Xi_A|\Xi_B)$ and $I(\Xi_A : \Xi_B)$ have the same range of possible values $[0, max(H(\Xi_A))]$. Of course the ranges become different if we dissociate the coarse-graining criteria. If we choose a finer coarse-graining for $I(\Xi_A : \Xi_B)$, we get a higher $H(\Xi_A)$ in Eq. 1A2 by maximizing it than by maximizing $H(\Xi_A|\Xi_B)$. The maximization of $I(\Xi_A : \Xi_B)$ is achieved by Caves's typical fine-grained phase-space patterns or typical pure states. However, the resulting $H(\Xi_A)$ has no specific coarse-graining criterion, thus losing any clear physical meaning.

Appendix 1B. Thermodynamic meaning of mutual information when measured for arbitrary variables

Here I give the demonstration of Eq. 1.14, $J(A:B) \geq k\log(2)I(X_A : X_B)$. The variables are defined in subsect. 1.2.2. This equation means that, when there is some mutual information between two systems, this will involve a decreased physical entropy, regardless of the particular descriptors of the systems that exhibit the mutual information.

According to Eq. 1.13, this is equivalent to $I(\Xi_A : \Xi_B) \geq I(X_A : X_B)$. As a first step, I demonstrate that

$$I(\Xi_A : \Xi_B) = I(X_A : \Xi_B) + I(\Xi_A : \Xi_B|X_A) \qquad \text{(1B1),}$$

where $I(\Xi_A : \Xi_B|X_A) = H(\Xi_A|X_A) - H(\Xi_A|\Xi_B, X_A)$.

Eq. 1B1 means that the mutual information between the microstates of B (Ξ_B) and X_A is fully embedded within the mutual information between the microstates of both systems. Eq. 1B1 is true if and only if:

$$H(\Xi_B)-H(\Xi_B|\Xi_A)=H(\Xi_B)-H(\Xi_B|X_A)+H(\Xi_B|X_A)-H(\Xi_B|\Xi_A,X_A),$$

i. e.

$$H(\Xi_B|\Xi_A)=H(\Xi_B|\Xi_A,X_A) \qquad \text{(1B2).}$$

According to Eq. 1.11,

$$H(\Xi_B|\Xi_A,X_A)=-\sum_{ijk}P(\boldsymbol{x}_i^A)P(x_j^A|\boldsymbol{x}_i^A)P(\boldsymbol{x}_k^B|x_j^A,\boldsymbol{x}_i^A)\log_2(P(\boldsymbol{x}_k^B|x_j^A,\boldsymbol{x}_i^A)) \qquad \text{(1B3).}$$

Since $P(x_j^A|\boldsymbol{x}_i^A)$ is either 0 or 1 and in the last case $P(\boldsymbol{x}_k^B|x_j^A,\boldsymbol{x}_i^A)$ reduces to $P(\boldsymbol{x}_k^B|\boldsymbol{x}_i^A)$, Eq. 1B3 gives rise to Eq. 1B2 and Eq. 1B1 is confirmed.

In the same way as Eq. 1B1, we can show that

$$I(X_A:\Xi_B)=I(X_A:X_B)+I(X_A:\Xi_B|X_B) \qquad \text{(1B4).}$$

From Eq. 1B1 and Eq. 1B4,

$$I(\Xi_A:\Xi_B)=I(X_A:X_B)+I(X_A:\Xi_B|X_B)+I(\Xi_A:\Xi_B|X_A).$$

Since $I(X_A:\Xi_B|X_B)\geq 0$ and $I(\Xi_A:\Xi_B|X_A)\geq 0$, it is confirmed that

$$I(\Xi_A:\Xi_B)\geq I(X_A:X_B)$$

and so is Eq. 1.14.

Chapter 2

Physics of directionality in succession and evolution

2.1. Introduction

Along the 20th century, prominent ecologists such as Lotka, H. T. Odum, Margalef, Ulanowicz, and J∅rgensen have hypothesized that there is a directionality in the change of ecosystems and the whole Biosphere across time, which is imposed by the laws of physics in the form of some thermodynamic potential that is either maximized or minimized (see e. g. Müller 1997). This thermodynamic potential is generally implied to be as universal for complex systems far from equilibrium as e. g. free energy for closed systems at equilibrium. It would rule for any biosphere in any other planet. Some authors have even labeled their own proposals as Fourth Law of Thermodynamics (J∅rgensen 1992, 2001b).

It is difficult to exaggerate the importance of any such "Fourth Law" if it were to really hold. We could predict nearly all of the features of ecosystems, because they would tend to approach a very specific situation. Some authors have begun to make quite specific predictions by introducing into model ecosystems the parameters that maximize or minimize their potential of choice (J∅rgensen 1992, 1999, 2001 a, J∅rgensen and Padisák 1996), and there is some interest in finding solutions of compromise between the predictions from the potential supported by each influential author (J∅rgensen 1992, Patten 1995, Fath et al. 2001). More importantly, the human socioeconomic system, as a part of the Biosphere and as a complex system far from equilibrium "by own right", would be subject to the same law. Hence any fundamental political discussion would be superfluous, because humanity would have no choice in the future organization of society and economy: it could be no other than the one that maximizes or minimizes the potential.

One of the foundations for these proposals is the observation of ecological succession. This is the term used by ecologists for the series of changes experienced by an ecosystem in the absence of severe disturbances. There are two main views of succession, a holistic one and a mechanistic one, historically represented by Clements and Gleason, respectively (Golley, 1977). Holistic ecologists developed the idea that succession has common features for all kinds of ecosystems, from planktonic systems to rainforests. Mechanist ecologists deny any universality and consider that it is only possible to make predictions from the detailed knowledge of the specific interactions relevant in each case.

Margalef and E. P. Odum offered two highly influential syntheses of the holistic view of succession. Margalef (1963) stated: 'Any ecosystem not subjected to strong disturbances coming from outside changes in a progressive and directional way. We say that the ecosystem becomes more mature. The two most noticeable changes accompanying this process are the increase of complexity of structure and the decrease of energy flow per unit biomass'. E. P. Odum (1969) put together a list of 24 trends in a table that became classical in textbooks (table 2.1) (this already included an explicitly thermodynamic trend, “decrease in entropy”, which is unrealistic if we take the strict physical meaning of entropy). In general terms, those trends agree with the more synthetic claims by Margalef.

Any trend in succession implies an asymmetry in time (Margalef 1991 pg. 234; see also Margalef 1997 pg. 88, 90). It is no absurdity to try to set bridges with thermodynamics, since thermodynamics is generally considered to be the only reason for any time asymmetry relevant at a biological scale. But this does not mean that the bridge should have the form of a potential, let alone of the specific potentials that have been proposed.

Out of these hypothesized potentials, the one that receives most attention nowadays is exergy, whose main proponent is Jørgensen (see e. g. Jørgensen 2001 a) (also H. T. Odum (1988, 1996) “emergy” has raised a considerable interest). Ecosystem exergy has already been the object of about 80 papers and several books, and the number of works displays an increasing trend. Most or all of this amount of work stands on an erroneous basis. This is obvious from chapter 1, where I showed that the equations used to calculate exergy in ecosystems are mistaken, and introduce a bias that in some cases amounts to more than 70000%.

There are some more reasons to doubt of all of these potentials. No universal potential has been found from studies on far from equilibrium thermodynamics (e. g. Nicolis and Prigogine 1977). Wilhelm and Brüggemann (2000) argued against any

Table 2.1. *The trends of succession, as viewed by E. P. Odum (1969).*

Ecosystem attributes	**Developmental stages**	**Mature stages**
	Community energetics	
1. Gross production / community respiration (P/R ratio)	Greater or less than 1	Approaches 1
2. Gross production / standing crop biomass (P/B ratio)	High	Low
3. Biomass supported / unit energy flow (B/E ratio)	Low	High
4. Net community production (yield)	High	Low
5. Food chains	Linear, predominantly grazing	Weblike, predominantly detritus
	Community structure	
6. Total organic matter	Small	Large
7. Inorganic nutrients	Extrabiotic	Intrabiotic
8. Species diversity – variety component	Low	High
9. Species diversity – equitability component	Low	High
10. Biochemical diversity	Low	High
11. Stratification and spatial heterogeneity (pattern diversity)	Poorly organized	Well-organized
	Life history	
12. Niche specialization	Broad	Narrow
13. Size of organism	Small	Large
14. Life cycles	Short, simple	Long, complex
	Nutrient cycling	
15. Mineral cycles	Open	Closed
16. Nutrient exchange rate, between organisms and the environment	Rapid	Slow
17. Role of detritus in nutrient regeneration	Unimportant	Important
	Selection pressure	
18. Growth form	For rapid growth ("*r*-selection")	For feedback control ("*K*-selection")
19. Production	Quantity	Quality
	Overall homeostasis	
20. Internal symbiosis	Undevelopped	Developped
21. Nutrient conservation	Poor	Good
22. Stability (resistance to external perturbations)	Poor	Good
23. Entropy	High	Low
24. Information	Low	High

potential in the Biosphere by suggesting that life could share the properties of a class of cellular automata that is unique in its capacity to generate complexity, and is subject to no potential.

In sect. 3.2 below I analyze the main potentials in the literature and show that none of them is well-founded. In subsect. 3.3.1 and 3.3.2 I show that there is however a physical basis to expect trends, but this does not consist on any potential. This allows to reconcile the holistic view of succession and evolution with the belief in an open future. In the rest of the paper, I put forward several other elements that allow to understand in some more detail the features of succession and disturbances, and the own structure of the Biosphere. In the last section I discuss the interest of these findings in front of current ecological challenges.

2.2. A critique to the main proposals of thermodynamic potentials driving the development of ecosystems

2.2.1. Principles based on Prigogine’s minimum entropy production

As explained in sect. 1.1, Prigogine found that systems that are kept slightly away from thermodynamic equilibrium by external constraints tend to achieve a minimum entropy production. This does not hold anymore far from equilibrium, where self-organization processes takes place (see Nicolis and Prigogine 1977).

Margalef (1963) signaled that one of the central trends observed in ecological succession is the decrease of energy flow per unit biomass. He mentioned the possibility that this was a consequence of Prigogine’s minimum entropy production principle (Margalef 1980 pg. 158-164, 1991 pg. 242, 1997 pg. 91). In more general terms, Prigogine principle would produce in ecosystems a trend to decrease energy flow 'per total mass with capacity of controlling the future' (Margalef, 1989, pg. 771) or 'per unit of stored and transmitted information' (Margalef, 1968, pg. 29). Since the flow of energy is constrained from the outside, this would involve an increase in the amount of organized matter. This hypothesis had an appreciable influence. It was further developed by Schneider (1988) and Matsuno (1978).

The extension of the minimum entropy production principle to the development of ecosystems cannot be accepted, for several reasons:

- Prigogine's principle is only valid close to equilibrium, where relations can be assumed linear and reciprocal, as an extension of the time-reversible state of equilibrium (see sect. 1.1, or Nicolis and Prigogine 1977). This is unbelievable for ecosystems: their complexity arises out of their nonlinearity, and their trajectories are clearly time-asymmetric.
- The deduction of the principle assumes that the system is kept out of equilibrium by constraints over one or several thermodynamic forces, while there is no constrain on thermodynamic flows (see sect. 1.1, or Nicolis and Prigogine 1977). If this were the case of ecosystems, there could be organisms that increased their own intake of energy without end, because any depletion of its source would be compensated by the external agent charged for keeping the 'forces' constant whichever the flows. In the case of ecosystems there are clearer restrictions over flows than forces.
- Margalef's 'energy flow' is in fact more related to net production than gross production (Margalef, 1991 pg. 242, 1997 pg. 91), while it is the last the most related to entropy production. Net production per unit biomass has to do with the turnover of macroscopic pieces (organisms), rather than specific entropy production.

Matsuno (1978) tried to proof theoretically the purported minimization of the relation between production P and biomass B in ecosystems, which he called Margalef principle, in a way completely unrelated to Prigogine principle. The essential hypotheses that can be extracted from his equations are that (a) the rate of variations in B is an increasing function of P/B and (b) ecosystems minimize the rate of variations in B. The first hypothesis is reasonable, but the second one is not justified.

The appeal of Prigogine principle for some ecologists may in fact be based on the idea that natural selection should favor a decrease in the energy required to sustain any given biomass, which is equivalent to an increase in the efficiency of energy use. This reasoning has no relation to Prigogine's and apply under completely different circumstances, despite the analogy in the conclusions. Neither it is related to the arguments by Matsuno. The validity of this idea is quite obvious, but it does not lead to any extreme principle, as apparent from the developments in next section.

2.2.2. Principles based on Lotka’s maximum power

Lotka (1922) proposed that the energy flow through the “system of living organisms” is maximized across evolution. This hypothesis is commonly referred as the maximum power principle, since power is energy per unit time. It was widely developed by H. T. Odum (Odum and Pinkerton 1955, Odum 1971).

The hypothetical thermodynamic goal functions that receive a highest attention in current systems ecology are mainly variations on Lotka’s principle. These goal functions consist of the maximization of some measure related to energy: either the maximization of the “stock” accumulated in ecosystems or far-from-equilibrium systems in general, or the flow through such systems. The difference between these measures and energy is that they incorporate some criterion of energy “quality”. The most popular among these are exergy, emergy and ascendency.

While the principle of minimum entropy production is a thermodynamic principle that has been incorrectly applied to biology, the presumed principles of maximization of each of these magnitudes are essentially based on biological reasoning but are claimed to have a general thermodynamic value. Some authors label their own proposals as the “Fourth Law of thermodynamics” (J∅rgensen 1992, 2001 b). H. T. Odum (1995) even complements the presumed “Fourth Law” with a “Fifth Law”.

The concept of exergy was introduced into this field by J∅rgensen and Mejer (1979) and has been widely developed by several authors, mainly J∅rgensen (see e. g. J∅rgensen 2001 a). As mentioned in subsect. 1.3.1, "exergy" or "availability" is defined in the engineering literature as the amount of work that can be extracted from a system while it is allowed to reach the state of equilibrium with its surroundings (Wark 1995). Since energy is conserved, any “energy expenditure” in biology, engineering or economics would certainly be expressed more properly as an “exergy expenditure”. Lotka’s principle is based on the assumption that living systems maximize their uptake of “available energy”, which in fact corresponds to exergy. While some other authors give more importance to the rate of exergy dissipation (Schneider and Kay 1994), Jørgensen proposes that the stock of exergy is maximized. He considers that the exergy of a biological system is not just a function of its chemical composition, but also of its informational content (the procedure of calculation can be found e. g. in Jørgensen et al. 1995). As shown in subsect. 1.3.1, a mistake in his developments produced an error of several orders of magnitude in the calculation of the contribution made by this informational part. In his calculations, the information content seems to be the main determinant of exergy, when in fact it is almost negligible. The real exergies of ecosystems have no relation with the exergies calculated with the method of J∅rgensen, and are essentially equivalent to biomasses (plus the exergy stored in the abiotic component of ecosystems, if this is included). This does not mean that the numbers obtained from his equations are useless. They have interest as a measure of complexity (unrelated to thermodynamics). They could be refined by ignoring the chemical component, but this will not change very much the results.

Emergy is a concept introduced by H. T. Odum (1988, 1996). The emergy of some element is the energy of a given kind (usually solar energy) required to obtain it. The rate of emergy use per unit time is called empower. Emergy is indeed a fine measure of energetic cost, with many possible applications. For controversies about H. T. Odum proposals, see Mansson and Glade (1993), Patten (1993) and Odum (1995).

Ascendency was introduced by Ulanowicz (1986, 1997, 1999). It is the sum of the energetic flows between the compartments of the ecosystem, multiplied by the mutual information (see subsect. 1.2.2) between the sources and sinks of these flows.

As mentioned in the introduction, there have been some recent efforts to produce hypotheses on ecosystem development that integrate these and other purported goal functions (J∅rgensen 1992, Patten 1995, Fath et al. 2001).

These measures are useful for several purposes, but there is no solid reason to assume that they act as goal functions or potentials. Recently, there has been some trend to avoid explicit references to maximization or minimization of functions, and use instead more fuzzy concepts such as “propensities” (Ulanowicz 1999) or “orientors” (Fath et al. 2001), which may be a step forward but does not solve the lack of a solid foundation.

Here I show in detail that Lotka’s principle is unjustified. The arguments can be immediately extended to the other purported principles, which are essentially based on the same reasoning.

Although the proponents of some of the above variations on Lotka principle claim their own goal functions to trascend natural selection (this is treated at the end of this section), the original proposal by Lotka was explicitly founded on it. Lotka took as good Boltzmann contention that Darwinian “struggle for life” was fundamentally a struggle for available energy. He deduced that natural selection would favor organisms able to absorb any available energy still unused by the “system of organic nature”, so its total energy flux would be maximized.

The validity of the maximum power principle would require two main premises, which can be shown to be false:

- Natural selection maximizes the uptake of available energy by organisms.
- Goal functions operating on organisms by virtue of natural selection can be extrapolated to whole ecosystems.

As discussed in detail in sect. 2.4, natural selection can be assumed to maximize the rate of increase of organisms of each species or genotype (which becomes eventually zero, nevertheless). This rate of increase depends on the uptake of available energy, the uptake of each of the other required resources, and the efficient use of the resources for the survival and reproduction of the organism. Since there are trade-offs between these functions, none of them will be maximized: each one will be optimized in such a way to maximize the rate of increase. Any innovation that relaxes the trade-offs will be favored by natural selection but, while these do not completely vanish, an uptake of available energy either below or above the optimum will be avoided by natural selection. Some of the reasons why an uptake of energy above some given rate may do more harm than good are:

- It will demand some use of the internal resources of the organism, which may be required for other functions. E. g. the assemblage of chloroplasts and mitochondria (the cell’s “energy plants”) requires the appropriate materials and keeps a part of the cell “construction machinery” occupied for some time, and these organules occupy some space in the cytoplasm once finished.
- It will demand structural or behavioral features that may be suboptimal for other functions. E. g. a high uptake of certain forms of available energy requires some organs with large and well expounded surfaces, but this may also produce an excessive loss of other resources (e. g. water in plants in dry areas), mechanical

fragility, susceptibility to changes in environmental factors such as temperature, vulnerability to consumers and pathogens, etc.

- A high energy throughput may damage biological structures. E. g. because of thermal agitation if the energy is eventually dissipated in the form of heat, release of free oxygen radicals from respiration, or the direct impact of the incoming energy on the own collecting structures when this comes in too concentrated forms, such as high frequency radiation (which plants and algae filter out instead of using).

Margalef (1989, pg. 113) rejects the maximum power principle from the simple observation of the small fraction of incoming radiation that plants use for photosynthesis. Flos (1984) suggests that the green color of plants indicates that they are not limited by energy, since it means that they do not use the band of the visible light spectrum in which this is most abundant. In fact, most of the solar energy used by terrestrial plants is spent in transpiration instead of photosynthesis, but this use is also kept well below the maximum attainable levels in areas with permanent or seasonal water limitations.

Even if we found natural selection to drive the evolution of organisms to the maximization of some thermodynamic function, this would not imply that ecosystems as a whole also maximize it. This can be seen from a simple argument. Consider N species. Consider a certain variable $\{j_i\}$, $i \in [1,N]$, which each species maximizes by virtue of Natural Selection. Consider a certain characteristic $\{c_i\}$, $i \in [1,N]$, which is genetically determined. Consider both $\{j_i\}$ and $\{c_i\}$ to be continuous and derivable. Each species will evolve to some value of c that maximizes its own j:

$$\frac{dj_i}{dc_i} = 0 \quad \forall i \, .$$

If the aggregate j for the ecosystem as a whole is taken as the sum over $\{j_i\}$, we get:

$$j = \sum_i j_i$$

$$\frac{dj}{dc_i} = \frac{dj_i}{dc_i} + \sum_{j \neq i} \frac{dj_j}{dc_i} \qquad \forall i$$

$$\frac{dj}{dc_i} = \sum_{j \neq i} \frac{dj_j}{dc_i} \qquad \forall i$$

For j to be a maximum, $\sum_{j\neq i}\frac{dj_j}{dc_i}$ must be 0 $\forall i$. There is no reason to assume this condition to generally hold.

Only for some particular schemes of interactions there is a basis to expect natural selection to lead to the maximization of some aggregate ecosystem feature. This is treated in subsect. 2.4.2. Biomass (or some more appropriate measure of the “magnitude” of the biological system) is maximized when the only relevant interaction between organisms is competition through efficiency in the use of resources (and exclusive symbiosis). This does not hold any more when either (i) differences in the capacity of resource uptake per unit biomass are relevant for competition, (ii) organisms have an effect on the supply of resources to the system (e. g. through varying mineralization rates, sedimentation of organisms, rests of organisms and fecal pellets to deep aphotic layers in aquatic systems, or bacterial transformations of nutrients), or (iii) there are interactions between organisms not mediated by their shared resources (other than exclusive symbiosis; e. g. competition by direct interference, predation, parasitism, propagation of fluctuations such as fire or epidemics from organism to organism). Therefore, a maximum biomass principle could be an acceptable approximation in some cases, but just an approximation and just in some cases.

Authors such as H. T. Odum (1988) or Ulanowicz (1999) champion the presence of extremal principles in ecosystems with complex sets of interactions, by appealing to autocatalytic cycles. Energy may follow different pathways in an ecosystem, which may include several trophic levels and also abiotic compartments. Each compartment will have an incidence on the overall flux of the pathway to which it belongs, both through its own inputs and outputs and through some possible control on other sections of the pathway. According to these authors, there will be a competition for available energy between each possible pathway, which will have as a result a selection of the one that maximizes energy flow or some related magnitude. E. g. they consider that a pathway in which some portion of producers is regularly lost to feed consumers will be selected if the consumers perform some function that enhances production. In spite of the various kinds of interactions involved, each of the models put forward by these authors can be reduced to a set of subsystems competing for a source of energy with a fixed flow. This has the essential features of a simple competitive community, therefore justifying the extremal behavior. The only difference is that the competing entities are not single species. These authors' models have an essential feature which is unrealistic: lack of interactions between pathways not mediated by their shared resource. Without this feature, we do not have anymore the conditions for an extremal behavior. E. g. consider that there are only two pathways *A* and *B*. The first one has a producer species p_A and a consumer species c_A, and the second one a producer p_B and a consumer c_B. If the only difference between them is that c_A invests some resources in performing a function that enhances the production of p_A, indeed the pathway *A* will outcompete *B* and the system will reach the state of maximum power. But this will not be the case if c_B is able to consume p_A. Then the function performed by c_A will benefit both c_A and c_B, and c_B may well outcompete c_A because it does not spend its resources in any similar function. The result will not be maximum power anymore (in this case due to a "short-circuit" between pathways driven by what economists would call a "free rider effect").

As a conclusion, there is no basis for the maximum power principle, neither for any of its variants. On the other hand, it is difficult not to be fascinated by the exuberance of life and its capacity to take advantage of the most unsuspected resources in the most remote places. It is tempting to revive Aristotle's tenet "nature abhors vacuum" in an ecological context. Lotka and the ecologists who have formulated different variants of the maximum power principle are in a position similar to Aristotle's, who proposed this tenet when fascinated by the capacity of fluids to reach any hidden corner. Torricelli showed that the idea of a high but finite atmospheric pressure offered a better description of nature than a "minimum vacuum principle", and everything indicates that the situation in ecology is rather similar.

2.3. Fluctuations as a key for directionality

2.3.1. Time scale separation in complex systems

Here I offer a justification and develop an idea proposed by Margalef, which I use in the next section as a basis to explain the existence of trends in succession and evolution. Here I refer to this idea as "Margalef principle" (although this label could be applied to many of his proposals, and has specifically been used for the hypothesis discussed in subsect. 2.2.1, by Matsuno 1978)

Margalef (1991, pg. 234; see also Margalef 1997, pg. 88, 90) postulates the existence of a fundamental time asymmetry in ecology: increases in the level of “organization” are “slow”, while “simplifications” are “fast”. I. e., he postulates a separation of time scales between increases and decreases of “organization”.

This assertion is almost self-evident if we take as “organization” some of the standard measures of complexity used in computation theory (see e. g. van Leeuwan 1990). The application of these measures to biology has been suggested by several authors (such as Bennet 1988). In general, the complexity Θ of a configuration is defined as the number of logical steps necessary to generate it, given certain conditions. It suffices the assumption of some positive relation between number of logical steps and real time in natural processes to deduce Margalef principle, though slightly modified. Some reasons to introduce modifications are:

- While the rate of increase of complexity $\frac{d\Theta}{dt}$ will have some upper constraint that will not be shared by the rate of decrease, there is no general reason to expect lower constraints in any case. Θ may decrease slowly in “senescent” systems. The only thing we can say is that this process can suffer a fast acceleration, but not such a fast reversion, by the effect of some outer disturbance.
- There is not a direct correspondence between number of logical steps and real time in nature. The speed of single “steps” (e. g. biochemical reactions) can widely change, and also the amount of steps that are taking place at the same time in a given region (the degree of “parallelism”). This second factor is specially sensitive to the own level of complexity. But these sources of error can be neglected when the amount of complexity involved is asymptotically high. E. g. forest succession after a replacement fire requires a number of logical steps many orders of magnitude above those required for the previous combustion.
- There is no such limitation on $\frac{d\Theta}{dt}$ if the complexity is “imported” instead of generated in the own system.

- There may be "qualitative innovations" that produce large increases of Θ in a single or a few logical steps. However, such innovations make part of computation processes (e. g. extreme search) that require a large number of time steps, and it is not possible to predict in which step will the innovation occur. Otherwise, such changes could not be considered to increase Θ.

All of these points should be taken into account to reformulate Margalef principle in a more rigorous way. I propose one such reformulation based on a formalism analogous to those in chapt. 1.

Complexity shares with entropy the additive property mentioned in subsect. 1.2.2: if the states of two systems A and B are independent, the joint complexity is the sum of the complexities of each of the variables:

$\Theta(A,B)= \Theta(A)+ \Theta(B)$.

If they are not independent, this does not hold anymore. In analogy with mutual information (subsect. 1.2.2), we can define a magnitude Φ that accounts for this effect:

$\Phi(A:B)= \Theta(A,B)- \Theta(A)- \Theta(B)$.

Unlike entropy, complexity is not subadditive, so Φ may be either positive or negative. The whole can be either more or less than the parts.

In analogy to Eq. 1.19 in subsect. 1.2.4, we can establish an equation for the local balance of complexity for any given site:

$$d\Theta = d_i\Theta + d_e\Theta + d_s\Theta \qquad (2.1),$$

where:

- $d_i\Theta$ is strictly the production of complexity, the generation of novel complex configurations within the site.
- $d_e\Theta$ is the flux of complexity that takes place when unique (without replicas) elements with a given complexity are transported into the site. In this case, the increase in Θ in the site is balanced by an equivalent decrease in its environment.
- $d_s\Theta$ accounts for the production or transport into the site of complex elements equal to elements from somewhere else (but different to any other element in the site). In this case, the increase in Θ in the site is balanced by an equivalent increase in Φ between the site and the environment.

Out of these terms, only $d_i\Theta$ is subject to strong limitations. Taking into account the above points, we can reformulate Margalef principle by stating that, when a system changes as a direct consequence of a modification in contour conditions, we will find:

$$\lim_{\left|\frac{d_i\Theta}{dt}\right|\to\infty} P\left(\frac{d_i\Theta}{dt} > 0\right) \approx 0 \qquad (2.2).$$

The condition that the system "changes as a direct consequence of a modification in contour conditions" is unnecessary if we exclude particular kinds of processes, such as search processes, which display "innovations" that represent fast endogenous increases in complexity.

The rate of complexity production $\frac{d_i\Theta}{dt}$ can be replaced by the rate of complexity increase $\frac{d\Theta}{dt}$ when dealing with processes essentially internal to the system (e. g. plant development in an ecosystem once the seeds are available) but not in other cases (e. g. immigration of organisms, which are complex).

The modified Margalef principle suggests that complex systems will "ordinarily" exhibit a time scale separation in which decreases in complexity are faster than increases, but does not preclude any other situation. It is only strictly predictive at the limit in which a system with an asymptotically high complexity is forced to change at an asymptotically high speed. The status of this principle should be considered analogous to that of the principles of linear nonequilibrium thermodynamics, because both hold in an asymptotic domain. This last science deals with systems asymptotically close to equilibrium: there is no general recipe to know at which distance from equilibrium the predictions begin to be valid and this should be determined empirically for each kind of system.

Eq. 2.2 is time asymmetric, thus setting an arrow of time. The orientation of this arrow is determined by the orientation of the thermodynamic one, because the first arrow depends on the time orientation of information processing and this depends on the last, for the reasons stated in chapter 1. So the modified Margalef principle is a consequence of the Second Law of thermodynamics.

2.3.2. Trends without potentials?

Take a complex system whose asymptotic dynamics is stationary and well-defined given some contour conditions (stationary in the sense used in time series theory, i. e. the marginal distribution and the structure of dependence of its successive states do not depend explicitly on time). Subject the system to a brief disturbance, i. e. to a modification of the contour conditions during a brief time span. It follows from Eq. 2.2 that, if the system is complex enough and the disturbance is strong enough, the immediate result will be a loss of complexity. It follows from the above assumptions on the system dynamics that, after the initial decrease in complexity, the most probable response of the system will be a compensating increase in complexity. This will be much slower than the initial decrease.

I. e., under wide conditions, after a complex system has been subjected to a disturbance, it will display a trend of increasing complexity. This is the same effect that would be observed if complexity acted as a potential, but the reason is completely different in each case. In the first case, the most probable asymptotic degree of complexity does not have to be the maximum attainable one, in contrast with the second one. In the first case, the directionality in the response is a consequence of the asymmetry of the disturbances, while in the second no assumption is required on the nature of the disturbances.

If we give a look at E. P. Odum’s trends of succession in table 2.1. we will realize that, for several of these trends, it is obvious that an external disturbance can only reverse them, regardless of whether or not they have previously attained their maximum possible value. So, if the ecosystem tends to recover its former state, it will display the trends, regardless of whether or not the system is searching an extreme.

Several of the observed trends of succession can be predicted for complex systems, regardless of their having or not some potential. So the trends by themselves give no information on the existence of such potentials, which may or may not hold. This is the main argument against the validity of the thermodynamic potentials reviewed in sect. 2.2: not only they lack a good theoretical justification, but the empirical observations they pretended to explain can be equally expected if there are no such potentials. This looks like being quite trivial, but does not seem to have been realized by all of the prominent scientists who have tried to justify the directionality of succession by postulating goal functions of any kind. The view here developed offers a much more open but equally holistic view of succession. According to this view, the directionality in succession and also in evolution is ultimately a consequence of the Second Law of thermodynamics, but in a much subtler way.

2.3.3. Phase transitions and criticality in ecosystems

Consider a system that is characterized by some variable, which is called "order parameter", which is a function of some "control parameter" that we manipulate from the outside. This function could be continuous and derivable, but this is not necessarily the case.

If the function has some point of discontinuity, it is said to display a first-order phase transition. E. g. when the "control parameter" temperature is driven beyond a discrete value, which is the point of fusion of water, this suffers a discontinuous change. There is evidence of drastic changes in ecosystems (Scheffer et al. 2001) and also in the coupled ocean-atmosphere system (Rahmstorf 1995) as a response to smooth changes in certain parameters, which can be interpreted as first-order phase transitions. Since these are fast changes, their immediate ecological effect can indeed be expected to be "destructive" (they decrease complexity), by virtue of the above developments.

If the function is continuous but has some point where it is not derivable, it is said to display a second-order phase transition. The point where this takes place is called a critical point. Second-order phase transitions have a great importance for complex systems such as ecosystems, because they are displayed by any simple model of propagation of fluctuations, and also of reproduction. E. g. take a virus that is introduced into a population of hosts large enough to be assumed infinite, for practical purposes. Let us ignore spatial effects. In its simplest form, the subsequent evolution of the viral population will depend just on two factors: the mean rate of increase r (mean reproduction rate minus average mortality) and the unavoidable randomness in the fate of each single individual. If the mean rate of increase is smaller than $r=1$, it is secure that the population will go extinct. If it is larger than one, there is still some probability that it goes extinct, but also some probability that it expands forever. The probability P that the virus population expands "infinitely" (i.e. until it becomes so large that this simplified model does not hold anymore) is an order parameter that displays a second order phase transition when the control parameter r surpasses the critical value r_c=1. For $r<r_c$, P=0. For $r>r_c$, P is an increasing function of r. P is thus not derivable in the critical point $r=r_c$.

When there is a second order phase transition, the global parameters of the system will change in a characteristic manner as a function of the control parameter (see the formalism in subsect. 3.2.1). This change may have a varying degree of abruptness. When the abruptness is high, the second order transition will be comparable to a first order transition, for practical purposes. As shown below, there are cases in which anthropogenic pressure may have unleashed second order phase transitions with outstanding consequences.

There is currently a great interest in second order phase transitions, but not for the above reason. Consider a system whose control parameter changes as a function of the own dynamics of the system. Further consider that this dynamics has this two features: (i) there is some "driving force" that, by itself, would produce an increase in the control parameter across time, such that this would pass through the critical point, but (ii) there is some additional mechanism that constraints the control parameter not to surpass the critical point. These systems are called self-organized critical (SOC) (see e. g. Bak 1996). It is often interpreted that they become trapped in the critical point. Things are however more complex, due to the reorganization that these systems suffer once they approach criticality, which modifies the own response of the order parameter to the control parameter. The detailed analysis of one of these systems in chapt. 3 makes clear that it does not really become trapped in any critical point, although it acquires the property most characteristic of criticality, which is scaling. Scaling means that the whole resembles the parts: e. g. there are small mountains within large mountains... (see Mandelbrot 1983). The scaling features include a fractal structure in space, a power-law statistical distribution of the sizes of the fluctuations, and what is called a $1/f$ power spectrum in the evolution across time. SOC has received a huge attention as a possible source for the scaling features often observed in nature.

The order parameter in SOC models is often the probability P that some fluctuation propagates indefinitely. In these models, the fluctuations move the control parameter down. In many such models, it is essential that there is a separation of time scales between increases and decreases in the control parameter, such that the decreases due to fluctuations are much faster (e. g. Vespigniani and Zapperi 1998). This is what ensures that the system does not surpass the point in which the fluctuations propagate indefinitely. Otherwise, the driving force could increase the control parameter at the same rate in which a permanent “fluctuation” makes it decrease, and the system could reach a steady state above criticality instead of a SOC state.

In subsect. 2.3.1 I try to offer an explanation for the frequent observation of a time scale separation in ecological systems, between the increases in several variables and their decreases under the action of certain kinds of fluctuations. This separation of time scales poses an important ingredient for SOC. Furthermore, some of these fluctuations have a multiplicative nature. So it is worth asking whether SOC dynamics is relevant for ecological succession (see Flos and Gutiérrez 1995). Before the discovery of SOC dynamics, Margalef (1986) had already postulated that ecological disturbances generally had a power law distribution. If proved, this finds an immediate explanation in SOC dynamics, although it could also have other sources (as proposed by Newman 1997 for extinction dynamics).

One of the first SOC models was someway inspired on forest fires (Drossel and Schwabl 1992). On the other hand, Robertson (1972, quoted in Strauss et al. 1989) had shown that wildland fires have a power law statistical distribution. I have made a thorough effort to test the consequent hypothesis that wildland fires do have a SOC dynamics (a hypothesis put forward independently by Malamud et al. 1998, and by Ricotta et al. 1999, while this work was in course). The results of this work, which make up chapter 3, strongly support this hypothesis.

SOC dynamics driven by fire seems to be the main determinant of the natural succession dynamics in a large portion of the terrestrial ecosystems other than tropical rainforests. Fire can be assumed to have a more important role than external disturbances, because it is regulated by the own ecosystem, producing a succession dynamics quite robust in front of environmental differences. Chapter 3 passes from theory to practical applications and gives SOC-based recipes useful for management. It offers a method for fire risk prediction as a function of meteorological previsions, which is tested with real data, and sets the basis for fire regime prediction as a function of global change and different management strategies.

The SOC model of succession vs. fire can be used to illustrate the above idea of trends without potentials. At least during some part of their development, forest (and other terrestrial ecosystem) stands increase their capacity to propagate fire. This is partially due to the increase of biomass. Biomass ordinarily displays a trend to increase, but this is often interrupted, either by some external disturbance or by fire (although the immediate effect is the same, it is not really proper to include fire as an external disturbance, since it is regulated by the own system). The overall ecosystem dynamics, fire included, favors a certain level of biomass that cannot be considered a maximum. E. g. the aggregate biomass could be highly increased if each single stand did not accumulate enough biomass to reach the critical point for replacement fires, but they do and then the system-wide biomass is smaller.

On the other hand, it has been suggested that the simulated systems in simple SOC "forest fire" models attain a maximum energy dissipation (Drossel and Schwabl 1992). At first view this could indicate that SOC models give a basis to a Lotka-like potential (subsect. 2.2.2). However, the "maximum" in the case of the forest fire model is defined in a too narrow sense to be given much importance. Simulations indicate that the average age in which stands catch fire approximately maximizes the rate of burning, where the "stands" are in this case parts of the forest large as compared to trees and small as compared to the average fire. This age corresponds to a point slightly above criticality. The maximum burn rate is also slightly above criticality: below this maximum, fire will burn a too small part of the biomass, while a position above the maximum indicates that some time has already been "wasted" since fire had the opportunity to burn almost everything. The proximity between the age of getting fire and the maximum can be simply explained because most stands are embedded within larger parts of the forest with a similar degree of maturity, so they get fire as soon as they are ready to burn almost completely because this is the moment when the rest of stands around them are ready to burn similarly, so they transmit the fire at the first opportunity.

In the case of tropical rainforests, I propose in chapter 3 that SOC dynamics does not develop because the critical threshold is not reached, thanks to their strong microclimatic regulation. However, an external tuning of the control parameters of anthropogenic origin could produce a second order phase transition, as above explained. This may have already happened in some cases, in which tropical rainforests have switched from being immune to fire to host some of the largest fires ever registered. In this case, the results are similar to those of a first order phase transition, because there is a huge change in the state of the forest for a small change in parameters, though the response function is not truly discontinuous. Fire does not only seem to be the main determinant of natural succession dynamics in terrestrial ecosystems out of tropical rainforests, but also the main threat for the future of tropical rainforests.

The importance of SOC for ecological succession is probably not restricted to fire. Solé and Manrubia (1995) showed that tree highs in a tropical rainforest have a multifractal structure, similar to that generated by a model based on tree falls when this approaches criticality. I suggested to some colleagues to perform a similar analysis in marine prairies of *Posidonia oceanica* and they found a similar result (Manzanera et al., in prep.), which could or could not have been produced by some internal SOC dynamics. They went further and discovered that some features of the multifractal spectrum capture the degree of anthropogenic disturbance on the prairie. Several authors sustain that macroevolution and extinctions have a SOC dynamics (Drossell 2001; see also subsect. 4.4.4) and “keystone” effects (e. g. Bond 1993) may lead us to hypothesize that this also holds at a successional time scale for local species assemblages, which would allow for a high amount of ecological predictions if confirmed (subsect. 4.4.4). There is also evidence of power law distributions in epidemics, at least those affecting human populations in islands (Rhodes and Anderson 1996). In sect. 4.5 I discuss the practical consequences if SOC is found to be frequent in pest dynamics, specially in agriculture. External disturbances affecting ecosystems such as floods, storms, earthquakes and volcanoes also display power law distributions, which could be the result of SOC dynamics in the processes that generate them.

If the above phenomena are confirmed to have a SOC dynamics, it will probably be straightforward to generalize to them many of the findings in chapter 3. Then SOC models will not only justify the observed self-similar patterns in fluctuations and other features, but also allow to predict the response of these patterns to external forcing by natural or anthropogenic factors. E. g. to predict how climate change will affect not only fire but also extreme rain events and floods, and even epidemics.

The theory of SOC systems offers a privileged framework for the understanding of succession and fluctuations, which seems to be much more powerful than the thermodynamic and similar approaches discussed in previous sections. However, we are far from being able to ensure that all of the ecosystems are SOC systems. In the case of forest fires, where I have performed an analysis much more detailed than has been done for any other proposed ecological instance of SOC (as far as I know), there is strong evidence for SOC in a large portion of the forests of the world but also for the existence of forests that are not SOC in terms of fire, specifically tropical rainforests, which on the other hand seem to be SOC for other reasons.

App. 2A gives some comments of the statistics of power laws, which are indeed the distributions of choice for dealing with external disturbances and also with internal fluctuations similar to disturbances, as suggested by Margalef (1986), regardless of their being or not produced by a SOC dynamics.

2.4. Natural selection and directionality

2.4.1. Natural selection within complex systems

In the previous sections in this chapter I have outlined the physical basis for some very general features of directionality in complex systems. An increase in detail in the understanding of directionality in succession and evolution requires taking natural selection into account. Therefore, I include in this section a brief discussion about natural selection and its relationship with potentials, trends, and complexity, although the novelty of most of the ideas here expounded is dubious.

There is consensus in that “natural selection” is the main “force” driving evolution. Natural selection could be interpreted at two levels. In its most basic form, it implies that sets of organisms with certain features are able to persist, while sets of organisms with other features are not. In a more specific form, it implies that the capacity to persist depends on some potential that tends to be maximized, which is called fitness.

Fitness is usually introduced as a parameter ad hoc in models of single species with several genotypes. In nature, things are more complex, because there are many kinds of interactions between organisms in the same or different species and between organisms and the environment. We can then wonder if systems with complex sets of interactions really produce some potential conducting the evolution of the organisms, and, if so, which are the consequences for the system as a whole.

In the simplest case, assume that new genotypes in an ecosystem (with several species) arise only as a result of random mutations (this is not essential, but simplifies the argument) and that, in the absence of such mutations, the system tends to approach some attractor. Then it is easy to show that evolution is driven by a potential. Close to the attractor, each genotype i has some expected biomass $E(B_i)$, and an expected rate of increase $E(r_i)$ that will be zero in all cases. While these can be taken as instantaneous figures if the ecosystem is driven by a point attractor, they should otherwise be averaged over the characteristic time scale of the attractor. The ecosystem will indeed be structurally stable in front of any new genotype i that satisfies $E(r_i)<0$ for $B_i \to 0$. If the new genotype has $E(r_i)>0$ for $B_i \to 0$, B_i will increase up to some limit set by nonlinear mechanisms, which will eventually drive the system to a new attractor with $E(r_i)=0$ and B_i either null or positive. A fictitious ecosystem structurally stable in front of any possible mutation would have each of its genotypes at some relative maximum of $E(r)$ in the map of possible genetic configurations, though all of these maxima would be identically $E(r)=0$. $E(r)$ thus acts as a potential for the characteristics of the organisms.

Allen (1976) studied analytically the structural stability of dynamic systems driven by point attractors in front of new mutants. He studied the general case of several mutants introduced at the same time, but if we restrict ourselves to introductions one per one, his result reduces to stating that the necessary condition for a new mutant i to be accepted is $\frac{\partial}{\partial B_i}\left(\frac{dB_i}{dt}\right) > 0$. This is equivalent to $r_i>0$, because $r_i = \frac{\partial}{\partial B_i}\left(\frac{dB_i}{dt}\right)$ for $B_i \to 0$ (if the system has n variables, $r_i \approx \frac{1}{dB_i}\sum_{j=1}^{n}\left\{\frac{\partial}{\partial B_j}\left(\frac{dB_i}{dt}\right)dB_j\right\} \approx \frac{\partial}{\partial B_i}\left(\frac{dB_i}{dt}\right)$, because, for $B_i=0$, $\frac{\partial}{\partial B_j}\left(\frac{dB_i}{dt}\right) = 0$ for any $j \neq i$).

Whenever a new genotype arises, the parameters of the system change. Van Valen (1973; see also Stenseth and Maynard Smith 1984) suggested that changes in the adaptative landscape of a species produced by changes in other species could keep the evolution going on forever. I. e. the system would never reach a structurally stable state. This is at the basis of current models of self-organized criticality in macroevolution (Drossel 2001). In these models, the trend for each species to maximize a fitness function is introduced a priori, like in conventional models of population genetics. The above developments imply that such fitness functions naturally arise in dynamic systems with certain features.

Lovelock (1988) proposes to enlarge the idea of natural selection by taking into account the effects of the organisms on the environment. If we take "natural selection" in a narrow sense, this means that such effects should be included within the fitness function. If these effects take place on the immediate surrounding of each organism and the possible effects at a larger scale are just an extension of the immediate effects, this may well be correct. But this does not seem to be the case for the direct effects on the environment at a large scale. The above results indicate that the "fitness" of a genotype i only depends on $\frac{\partial}{\partial B_i}\left(\frac{dB_i}{dt}\right)$, not on $\frac{\partial}{\partial B_i}\left(\frac{dB_j}{dt}\right)$ for any $j \neq i$. As B_i becomes macroscopic, it will have some impact on the other variables and its eventual value will depend on the interactions with them. It is true that the organisms may have some impact on the environment, and that in the long term the environment resulting from this impact combined with any other factor of relevance will be compatible with the persistence of any species that is not already extinct. Even in some case the impact of some species on the environment may become necessary for its own persistence. However, in no case this impact will be optimized in terms of any fitness function. So the claims by Lovelock will be correct if we take natural selection in the wider sense mentioned at the beginning of this section, not in the narrow one.

Since a fitness function can be predicted to arise for certain dynamic systems but not for other imaginable systems, the idea of fitness should not be considered a tautology. This does not mean that there is any simple method to predict in detail the genetic features that maximize it. Many simple formal problems of search for extremes are considered to be intractable. This means that the number of logical steps necessary to strictly solve them increases with the amount of possibilities of choice at a rate higher than any polynomial function of the last (Leeuwan, 1990). "Genetic algorithms" based on natural selection are known to be very efficient to obtain values close to the true extreme, so there is no method to predict their result that involves appreciably less computation than the own algorithms. If "adaptative landscapes" are intractable, the outcome of natural selection will attain a high amount of complexity, in the computational sense of the word (subsect. 2.3.1). This may precisely be the reason why natural selection is usually perceived as the main source of complexity, in the informal sense of the word. This does not impede to obtain some partial but relevant predictions from optimization theory (Parker & Maynard-Smith 1990), as currently applied to ethology and other fields.

2.4.2. System-wide potentials in simple cases

By virtue of natural selection, the evolution of organisms can be assumed to be driven by a potential, as discussed in the above subsection. As mentioned in subsect. 2.2.2, only for quite particular sets of interactions we have a clear basis to expect potentials acting also on the ecosystem as a whole. Here I analyze in some detail which are these sets.

Let us take as a point of departure the simple case of a set of species i= 1,..., n that have the same limiting resource, and that have no relevant interaction between them not mediated by their effect upon the concentration R of the resource. Tilman (1990) analyzed several models of this kind. He concluded that the result of the competition between such species would depend upon R_i^*, which is the concentration of the limiting resource at the steady-state when the only species is i: the species with the smallest R^* outcompetes the others. He successfully tested his predictions both with freshwater algae and with terrestrial plants.

In these models, if we have a species i in isolation and at steady state and we introduce an inoculum of species j, this will outcompete i if and only if $R_i^* < R_i^*$. Beyond particular models (and leaving apart the effects of random fluctuations), this is straightforward to proof under two general assumptions, which are sufficient but not necessary: (i) $\frac{dB_k}{dt} > 0$ for $R > R_k^*$ and $\frac{dB_k}{dt} < 0$ for $R < R_k^*$ given k (B is biomass), and (ii) the variation in R is monotonous. It follows that if i is alone at steady state (with $R = R_i^*$) and we introduce some j with $R_j^* > R_i^*$, then $\frac{dB_j}{dt} < 0$ and

j will not prosper. If $R_j^* < R_i^*$, then $\frac{dB_j}{dt} > 0$. Beyond some point, we will have $R < R_i^*$, and therefore $\frac{dB_i}{dt} < 0$ until i extinguishes, while $\frac{dB_j}{dt} > 0$ until $R = R_j^*$.

In this simple case, R acts as a potential. It is not however comparable to the presumed universal potentials in sect. 2.2, because this one depends on the identity of the limiting resource, which is not necessarily predictable from first principles.

Under more stringent conditions, the minimization of R is equivalent to the maximization of biomass (or some other more appropriate measure of the "magnitude" of the biological system). This can be expected to hold when (i) the results of the competitions are mainly driven by differences in the efficiency in the use of the resource, and (ii) there are no significant differences between species in their effects upon resource supply (e. g. through different rates of remineralization, or different sinking rates in planktonic systems). The second assumption could perhaps be dropped if we replace "maximization of biomass" by "minimization of production to biomass ratio", as suggested by Margalef, but not indeed the first one. Take e. g. one of Tilman's (1990) models of competition for a nutrient (here I drop the subindices because I take the case of a single species):

$$\begin{cases} \frac{dB}{dt} = rB\left(1 - \frac{h}{Q}\right) - mB \\ \frac{dQ}{dt} = \frac{vR}{r + K} - rQ\left(1 - \frac{h}{Q}\right) \\ \frac{dR}{dt} = y - \frac{vRB}{R + K} \end{cases} \tag{2.3}$$

where Q is the nutrient content per unit biomass. The meaning of the parameters is as follows: r is the maximal growth rate, h is the minimal tissue content required for plant survival, v is the maximal rate of nutrient uptake per unit plant biomass, K is the half-saturation constant for the Michaelis-Menten nutrient uptake, m is the rate of loss of plant tissue not directly due to lack of the nutrient, and y is the rate of nutrient supply.

The steady-state R, to be minimized, is

$$R^* = \frac{rhKm}{v(r - m) - rhm}.$$

This depends both on parameters related to resource uptake given biomass (v and K) and on parameters related to the efficiency in the use of resources to increase and sustain biomass (r, h, and m). The effect of these last parameters upon R^* is mediated by their effect upon biomass:

$$\frac{1}{R^*} = \frac{v}{K}\left(\frac{B^*}{y}\right) - \frac{1}{K}$$

(where B^* is the steady-state biomass, and $\frac{B^*}{y} = \frac{r-m}{rhm}$, which is independent from the other parameters in Eq. 2.3). Therefore, if we can ignore differences in uptake capacity and in effects upon supply y, biomass becomes a potential to be maximized, but not otherwise.

Even R ceases to be a potential if we take into account that competition often involves interference between species not mediated by resources (Margalef 1989, pg. 665-666, offers some amazing examples).

All of this remains valid as we pass to competitive communities with several species at steady state. A principle of “maximum biomass” is unlikely to strictly hold (as e g. expected by Nicolis and Prigogine 1977, pg. 455-458, from a simple model), because there are several other factors at work besides the differences in the amount of biomass that can be sustained with the available resources. Furthermore, some simple competitive communities do not even approach a point attractor (see the practical examples in May 1974b, Tilman and Wedin 1996, Constantino et al. 1997). Despite all of these qualifications, there is also empirical evidence that can be taken to support that a “maximum biomass” principle provides a reasonable approximation for the outcome of competition in plant communities at a small scale (see subsect. 4.3.1), and this could well extend to other systems mainly driven by competition.

However, many ecosystems are largely shaped by other interactions besides competition. Few of these will not distort the simple schema leading to the maximization of biomass (exclusive symbiosis is one of these few). Some of the interactions that will distort it are trophic interactions (there are several instances in which the biomass of an ecosystem seems to remain low because of the grazing pressure and it increases after removal of the consumers, see e. g. Matson and Hunter 1992), the propagation of fluctuations from organism to organism (such as fire or epidemics, see subsect. 2.3.3), or the multiplicity of chemical transformations of nutrients carried out by bacteria. We have no reason to think that a system with such involved interactions will maximize biomass or any other parameter (see also subsect. 2.2.2), although the maximization of biomass may still be a useful approximation for parts of this system.

Allen (1976), as a corollary of his analytic findings mentioned in last subsection, suggested a potential for prey-predator interactions: there would be a general trend toward increasing the ratio of biomasses between predator and prey. He reached this conclusion by applying the "partial potential" (in my terms) that he had found to all of the parameters of the predator-prey system except the parameter of interaction between them. Since the evolution of preys and that of predators will push this parameter in opposing directions, he presumed that it would remain roughly invariant. However, there is no strong argument for this assumption. He mentioned the example of phytoplankton, which often sustains a biomass of zooplankton larger than its own biomass, but forgot cases of organisms much more successful in their defense against potential consumers, such as trees or elephants.

2.4.3. From partial potentials to system-wide trends

Subsect. 2.4.2 suggests that only in quite narrow cases we should expect complete ecosystems to search extremes in parameters such as biomass or nutrient concentration. However, there are other ways in which the partial potentials acting on the evolution of organisms (subsect. 2.4.1) may have an incidence upon ecosystem trends. This is the issue of this subsection. Though the concept of natural selection is generally used for sets of genotypes within a species at an evolutionary time scale, it may well be extended to sets of species in an ecosystem at a successional time scale, if the ecosystem is large enough for its composition not to be mainly determined by the fluxes of organisms between the system and its exterior. The consequent species replacement may be complemented by physiological changes in the organisms that also maximize fitness because have evolved as a response to previous successional sequences with similar features.

The maximization of the rate of increase r has eventually no consequence for the own r, because this will always be null at the steady state. Neither this has any universal consequence for the amount of biomass of the genotype, the species or the ecosystem. The eventual biomass is determined once nonlinear mechanisms have reset all of the increase rates to zero, so it will depend on the nature of such mechanisms. As shown in last subsection, only in some particular cases there will be a maximization of biomass.

The trends of succession can be partially explained by direct application of the basic tenet that we can extract from subsect. 2.3.1-2.3.2: complexity needs time. However, the detail of some of these trends can be better understood if we take into account the possible diverging selection pressures that may arise along succession. These are explicitly mentioned by E. P. Odum (1969) between his trends (Table 1).

Some of the basic trade-offs for maximizing *r* include increasing resource uptake vs. increasing efficiency of use and increasing reproduction or growth rate vs. decreasing mortality. Since there are trade-offs between them, none of these parameters acts as a potential. Of course, any change that allows to relax the trade-offs will be favored by natural selection.

Elementary application of optimization theory allows to predict that, other things equal, a high resource availability will favor uptake over efficiency, and also reproduction and growth over longevity (this is the idea behind the notion of *r* and *K* strategists). These rules will often allow for system-wide predictions, at least in systems clearly dominated by a single trophic level.

Many disturbances involve an increase in the availability of resources, be it as a consequence of the destruction of biomass, or because the disturbance consisted of a direct addition of nutrients, an increase in turbulence, etc. This favors a higher turnover in the earlier stages of succession, and also a higher uptake of resources and a lower efficiency in its use. But this is not extensive to all kinds of disturbances. A disturbance may also consist of a decrease in the availability of resources, e. g. a drought. When there is a drought, many organisms pass to forms of resistance, which imply a very low turnover.

When the development of individual organisms involves many logical steps, this will require a high longevity and a low rate of increase of biomass. This may be a part of the explanation for the relation between high complexity and low turnover put by Margalef (1963), though not the complete explanation (e. g. turnover may be related to size by geometric constraints, and size to complexity). So a part of the changes in the relative rates of increase of different kinds of organisms along succession, which may be expressed as variations in selection rates, could reflect the longer time required by organisms with a higher developmental complexity to reach high biomasses, and thus be directly explainable by the developments from subsect. 2.3.2. In the case of resistance forms, the low turnover is not associate to a large amount of logical steps, but to their slowing down.

The time requirements for complex information processing will introduce a delay in the response to selection pressures, which complicates things. While the conditions favoring a high reproduction rate may decrease monotonously after a disturbance that increased the availability of resources, the actual reproduction rate will not, because an increase in reproduction rate is a complex operation. This is apparent from bacterial cultures, which are well known to display a period of latency before exponential growth.

These comments make apparent that some of the purported trends of succession are not as universal as have sometimes been thought to be. Even turnover may change in contradictory ways, which is a further argument not to interpret decreases in turnover as a direct effect of the principle of minimum entropy production or any other physical principle. The changes across succession can be better understood by thinking in terms of natural selection. The main physical constraint to take into account is the required time for complexity to develop.

Natural selection establishes a set of partial potentials acting on the evolution of organisms, which does not necessarily translate to any system-wide potential in biomass, production, turnover, etc. However, we may wonder if complexity by itself acts as a potential for the Biosphere as a whole, as a result of natural selection. Along the history of life, there have been successive originations of organisms more complex than all those that preceded them, at least in the intuitive sense of “complexity”, and probably also in the computational sense. Mainly as a result of these originations, the Biosphere can be considered to have become ever more complex. This has not been a monotonous trend, since it has been interrupted by extinction events of varying magnitudes, but usually the increases of complexity between such events have been larger than necessary to compensate for the losses that these produced. One of the reasons must have been that these ever more complex organisms required more time to evolve, and evolution has never been interrupted, because no extinction event has spanned the whole system. Some optimization problems can only be fully solved in an infinite time (Zurek 1989b), so natural selection could be able to keep on producing complexity indefinitely. The progressive approximation to the extremes of the "adaptative landscapes" will be partially compensated by the changes in the landscapes produced by the own evolutionary process, as postulated by many authors (since Van Valen 1973), but this might not affect all of the dimensions of the landscape. E. g. evolution could be pushed up and down along some trade-off, while keeping a definite direction towards relaxing the trade-off. On the other hand, we cannot rule out that the own evolutionary process produces at some point a complete extinction and interrupts itself. Wilhelm and Brüggemann (2000) compared the evolution of the Biosphere to a class of algorithms able to generate an indefinite computational complexity; one of the features of these algorithms is that it is impossible to know on advance whether they are going to halt. So it is not inconceivable that this same limit to predictability applies to the Biosphere.

2.5. The discretization of life

Margalef invites to imagine a biosphere different from ours, which would consist on a single continuous living layer. The striking contrast between this biosphere and the ours, made of discrete organisms, reminds that the reasons for such discretization are far from obvious. Here I discuss this and other related properties of our Biosphere, in the light of the previous developments.

Consider the frequency n_f of small fluctuations of a given kind affecting a complex system. Consider that there is a probability p_W that one such fluctuation becomes macroscopic and causes a loss of complexity $\Delta\Theta<0$. Consider that the system requires a time T to recover its former complexity, and that this is long enough as compared to the time of propagation of the fluctuation to neglect the last (see sect. 2.3). Then the development of the system will be subject to the constraint

$$n_f p_W < T^{-1}$$

For large $\Delta\Theta$, this reduces to

$$n_f\, p_\Omega \to 0 \qquad (2.4)$$

At least for fluctuations that propagate in agreement with simple rules such as those mentioned in sect, p_W will suffer a second order phase transition (see subsect. 2.3.3) at some critical point r_c for a control parameter r, such that for large systems we find $p_W=0$ for $r<r_c$ (subcriticality; for $r=r_c$, which corresponds to criticality, p_W also converges to 0 with increasing system size, but more slowly), and $p_W>0$ for $r>r_c$

(supercriticality). For the rest of this section, I call a system "supercritical" whenever p_w is not negligible, including the possible cases in which a second order phase transition cannot be defined.

So Eq. 2.4 roughly reduces to the condition that the system should not be supercritical, unless we are dealing with some kind of fluctuation extremely rare, with $n_f \to 0$.

This condition applies to the Biosphere, in front of many conceivable fluctuations. However, for most of these it will not apply to organisms. The main reason (a less important one is a lower n_f by the effect of size) is that, unlike the Biosphere, organisms are not "alone", so $d\Theta$ does not reduce to $d_i\Theta$ in Eq. 2.1. Most of the complexity of individual organisms is due to $d_s\Theta$ instead of $d_i\Theta$, by means of reproduction. So reproduction allows organisms to bypass the threshold in Eq. 2.4 and become supercritical, and so we do. A fluctuation such as a local lesion in a vital organ or a tumourous cell can propagate throughout the organism and eliminate virtually all of its complexity. Individuals are supercritical sets of biomass (this is the reason why individuals are indivisible). This establishes a fundamental difference between "simple mortals" and "Gaia".

To be "mortal" (from other causes than large external disturbances) is indeed a disadvantage in terms of natural selection, so it should be linked to some strong selective advantage. Here I offer a conjecture to explain which is this link. Since organisms are systems far from thermodynamic equilibrium, its maintenance is an active process (see chapt. 1). All of the parts that make up an organism have been "designed" by natural selection to have some function in this process of maintenance, in growth or/and in reproduction. A fluctuation affecting one of the parts may disrupt the function of maintenance, and in this way the fluctuation will propagate to the whole organism. The probability that this happens can be reduced by having parts of the organism with redundant functions, but this requires additional resources. For any reason, the selective optimum in this trade-off does not correspond to enough redundancy to make p_Ω negligible (except perhaps in some modular organisms). Once we have supercritical organisms with "natural" mortality, its rate could be further optimized because n_f may vary with metabolic rate. The resulting probability of failure of organism maintenance combined with the probability of death by external causes sets the scenario for the evolution of the genetic “development programs”, which could secondarily sharpen the upper limit to longevity.

If we accept that natural selection pushes organisms to be supercritical and dynamic constraints force the Biosphere not to be supercritical, the only option seems to be the division of the Biosphere into discrete pieces that are internally supercritical, but do not propagate the fluctuations to other pieces enough for the Biosphere as a whole to be also supercritical.

The Biosphere is also a far-from-equilibrium system that actively maintains itself, but it is composed of an extremely large number of organisms with highly redundant functions, which prevents the failure of its maintenance. In this prevention also seems to play an important role the diversity of organisms, which avoids their synchronous response to given changes. The relation between diversity and stability is a fairly active area of research, and is treated in chapter 4 (Naeem (1998) also uses the concept of redundancy in an ecological context, but only in relation to diversity of organisms, not to their multiplicity).

The sizes of the pieces in which the Biosphere is divided seem to roughly obey a power law. This is suggestive of SOC, but right now we have no serious basis to sustain this hypothesis. This is an area that deserves further work. An interesting point is that the simple forest fire model by Drossel and Schwabl (1992) consists of a SOC system made up of supercritical pieces (clusters of trees) that grow and replace smaller pieces (in this case by incorporating them, but this is not so different from outcompeting them), eventually “dying” and releasing space for small pieces to arise and to begin again the sequence. A model for the sizes of trees could not differ very much from this, but indeed things should be quite different e.g. for a model of a pelagic ecosystem, where size has some relation with trophic level (see e. g. Camacho and Solé 2001). The issue of size distribution is also treated in App. 5L. Some of the main results from chapter 5 should be taken into account also in this field: namely, that sets of many different superimposed or not superimposed processes could eventually produce a power law and, conversely, that a power law by itself is a weak evidence for any of these processes.

2.6. From thermodynamics to statistical physics

Chapter 1 and the current one make clear that thermodynamics is essential to understand why ecosystems are the way they are, but it does not seem to let us go much beyond the explanation of basic features that the ecologist already took for granted, in contrast to the optimistic claims by other authors. Further advance seems to require additional assumptions that cannot be justified from first principles alone, and should be tested. However, there is a large field between the above strict deductive derivation of universal laws and the accumulation of results referring to very narrow situations, which makes up most of the theoretical and empirical ecological literature. Some of the developments in this chapter pertain someway to this middle field, and so do the chapters 3 to 5 in the book.

It may be the case that many different sets of assumptions more or less realistic converge to the same results. These should be paid special attention, for two reasons. On the one hand, because such results will be quite close to general laws (though they are not), with a high (but not total) predictive power. On the other hand, because the awareness of such families of models with convergent results prevents from the erroneous predictions that will follow when a particular mechanistic model is taken for proved just because it produces these results. This is not rare, as shown in chapter 5.

The most clear source for such general results is the convergence in statistical features that arise when we have many degrees of freedom. This is the basis of statistical physics (e. g. Garrod 1995) and has allowed to uncover or justify several laws governing large sets of molecules under given conditions, including the Second Law of thermodynamics. These laws have a strong predictive power. The same may apply when we have many organisms and many species instead of many molecules, although in the first case there will always be some margin of incertitude, because biological systems are much more flexible than simple sets of molecules. This approach is analyzed in detail in chapter 5, and several results are derived and tested with empirical data. Already Patten (1975) appreciated the interest that a “large number” approach would have for ecology.

Wagensberg et al. (1988, 1991) applied to ecosystems an approach based on statistical physics, which implied a potential to be maximized. This is the Maximum Entropy formalism or MAXENT (Jaynes 1983). It is a formal generalization of the Second Law of thermodynamics to “entropies” defined from arbitrary variables, which do not have to correspond any more to the variables relevant for thermodynamics. E. g. we may take the population of a species instead of the position of a particle. The use of MAXENT as an analytic tool may be correct, but it requires the specification of some aprioristic probabilities. This is usually ignored, which may lead to erroneous results and also give a false impression of universality. This issue is also treated in chapter 5, which includes an example of (seemingly) successful application (subsect. 5.4.2, sect. 5.6, and App. 5B).

Large number effects pave the ground for the existence of emergent regularities, including potentials and, hence, directionality. But there is a more startling corollary. Features based on large numbers will weaken or even completely change as soon as the numbers are not "large" anymore. This is essential for the diversity-stability issue, so relevant in this moment of massive extinction, as has already been perceived by some authors. Some contributions to this area are presented in chapter 4.

2.7. Discussion and conclusions

The main conclusions from this chapter are:

- None of the thermodynamic potentials that have been proposed as explanations for the directionality in ecological succession is well founded. Exergy in particular, which is the one receiving more attention nowadays, is used in a clearly incorrect way in the literature.
- There is a well-founded time asymmetric feature in complex systems, which I call modified Margalef principle and capture in Eq. 2.2. This expresses that there is a limitation in the rate of production of complexity that does not hold for the rate of simplification. It is a consequence of the Second Law, but quite an indirect one.
- As a corollary of the modified Margalef principle, ecosystems will display trends without need of potentials. Any measure associate to complexity will decrease whenever there is a strong enough disturbance, and this will normally be followed by a trend to increase until recovering the former state, regardless of its being or not an extreme of any function. This simple reasoning leaves

without basis the main argument behind all of the hypotheses on ecological potentials.

- The separation of time scales between increases and decreases of complexity facilitates the development of self-organized critical dynamics. There is evidence that this phenomenon is the main determinant of succession dynamics in many ecosystems, but this could not be the case for all of the ecosystems. Where it is, a large array of predictions follows (see chapter 3).
- Natural selection produces a partial potential shaping organisms, which can be expected to hold even in the context of complex interaction webs. This does not imply any potential acting on the ecosystem as a whole. However, the consideration of natural selection allows to explain the detail of some of the purported trends of succession, and also to show that not all of them are so general.
- Only for competitive communities under quite narrow assumptions we may expect a maximization of biomass as a result of natural selection. This could still be a reasonable approximation for competitive communities in general, but not for ecosystems with more involved interaction webs.
- The interplay between the partial potential represented by natural selection and the global constraints on the propagation of fluctuations could explain the discretization of the Biosphere into individual organisms.
- When a complex system has many degrees of freedom, it acquires regularities that are based on statistics, instead of being a direct application of thermodynamics. These allow to define potentials, but they are not really universal. This is treated in detail in chapter 5.

The modified Margalef principle in Eq. 2.2 is indeed holistic, but it has two strong differences with the holistic hypotheses discussed in sect. 3.2. On the one hand, it has a more rigorous physical foundation. On the other hand, it imposes no goal to the evolution of complex systems. It thus reinforces holism while reconciling it with the belief in an open future.

The rest of developments in this paper allow for further increases in predictability, but partial and standing on additional assumptions, which means that there is always some margin for the surprise. Much of the potential increase in predictability will rely on the statistical physical approach, which is also holistic because it refers to features that are defined at a high level of integration and have little dependence on the details of the system, though this dependence is not null.

One feature that could hypothetically be universal for all complex self-organized systems is criticality. This will be the case if, as proposed by Langton (1990), complex information processing can only arise spontaneously close to criticality. In fact, we should restate this condition in terms of “criticality or self-organized criticality” because, as shown in chapt. 3, at least some of the so-called “self-organized critical” (SOC) systems are not really critical, despite sharing the main properties of critical systems. Furthermore, the requirement of criticality or SOC still leaves a huge margin of maneuver. One system may display SOC in relation to some particular variable and not in another. E. g. I give evidence in chapter 3 that many forests are SOC in relation to fire, while tropical rainforests are not, but Solé and Manrubia (1995) found evidence of SOC in these last forests, which should be related to some other variable. In addition, a complex system may approach criticality in ways unrelated to SOC. In subsect. 4.3.2 and sect. 5.3 I discuss one such path to criticality, which is quite elementary and indeed relevant for ecology (and economy, see subsect. 6.7.2).

The central result in this chapter is the modified Margalef principle, which is captured in Eq. 2.2. To some extent, what this principle expresses is well-known by (almost) everybody in an intuitive way. Margalef made it explicit and what I have done about it is to (i) modify Margalef’s formulation, (ii) express it in a more formal way, (iii) show that it follows from the laws of physics, and (iv) show that it allows to dissociate trends from potentials. The formalized principle, the causes and the consequences I postulate are falsifiable, so these are scientific findings. On the other hand, since the principle is qualitative instead of quantitative and it was already known in an intuitive way, it has not furnished novel ecological predictions to be tested. Nevertheless, I still consider it to be of interest, for three reasons:

- If correct, it allows to improve our understanding of nature, by showing that there is a coherence behind a lot of apparently dissociate facts.

- It leaves without justification a set of other theories, which did not only pretend to improve our understanding of nature in an incorrect way, but also produce novel predictions that can thus be expected to be false.
- If further developed, it could perhaps render new predictions.

The same three reasons justify the interest of the findings in chapter 1.

I do not consider that the first point have just an academic interest. The qualitative perception of nature has a strong impact on our way to interact with it, so the practical interest of theoretical findings does not reduce at all to the production of detailed predictions. In chapt. 6 (subsect. 6.3.1) I explain why I think that the awareness of the (modified) Margalef principle may have a foremost importance for our relation with the rest of the Biosphere.

Appendices

Appendix 2A

Handbook for the use of the power law statistical distribution in ecology

Statistics plays a fundamental role in modern ecology, but the statistical methods currently used are mainly based on the study of the Gaussian distribution. The developments in this book and in many other works, however, leave little doubt of the ubiquity in ecology of a fairly different distribution: the power law distribution, also called Pareto or hyperbolic distribution. The magnitudes of disturbances, the areas of each kind of patch in landscapes, the abundances of species, and the sizes of organisms are some of the many ecological variables that ordinarily resemble much more a power law than a Gaussian curve. Both distributions have quite different properties. The ultimate reason why ecologists often find useful to take logarithms of some variables is that, in this way, a power law becomes an exponential distribution, which is much closer to a Gaussian one. Hence, distributions close to power laws become close to Gaussian curves. Indeed it is often preferable to treat power laws as power laws, but ecologists (and even physicists, who are often more familiar with this distribution) usually lack the necessary statistical background. In this App. I provide a brief introduction to the statistics of power laws that includes the essentials to treat ecological data in a way different than usual. This is the result of putting together information that is quite disperse in the literature, and adding some lessons of my own experience in dealing with ecological data. In App. 5I add some more points that are required for the developments in chapt. 5, concerning the abundances of species.

Take a continuous statistical variable x with a power law distribution. The literature on power laws (see e. g. e. g. López Cachero 1990, Mandelbrot 1983) most often treats the case in which there is a lower bound x_0 to the power law and no upper bound. In this case, the probability density function (p. d. f.) will have the form:

$$f(x) = \left[(\boldsymbol{b}-1)x_0^{b-1}\right]x^{-b} \qquad \text{(2A1),}$$

where $\boldsymbol{b}>1$. The corresponding distribution function or cumulative probability function is:

$$F(x) = 1-\left[x_0^{b-1}\right]x^{-b+1} \qquad \text{(2A2).}$$

In the field of power laws, it is quite usual to work with the probabilities of excedence, i. e. the probability that the variable exceeds some given value:

$$P(X > x) = \left[x_0^{b-1}\right]x^{-b+1} \qquad \text{(2A3).}$$

Power laws are often found in connection with fractal forms (Mandelbrot 1983, 1984), i. e. forms with distinguishable features in a wide range of scales (mountains, coastlines, archipelagos, clouds, rivers, trees...), which may be “scaling”, i. e. with similar patterns in all this range of scales. E. g. the sizes of the lakes or the islands in a scaling landscape have a power law distribution. Power laws have no characteristic scale, and Mandelbrot (1983) stated that they should be considered scaling probability distributions. Note from Eq. 2A1 that if we multiply x by a constant k, we only have to rescale $f(x)$ by multiplying it by a constant $k^{-\boldsymbol{b}}$ (or rescale $P(X>x)$ by multiplying by a constant $k^{-\boldsymbol{b}+1}$, in Eq. 2A3), regardless of the value of x (while we do not collide with the lower bound x_0).

Power laws have rather special statistical features (e. g. López Cachero 1990). This can be seen if we try to calculate the moments of the distribution:

$$E\left(x^k\right) = \int_{x_0}^{\infty} x^k f(x)dx = \left[(\boldsymbol{b}-1)x_0^{b-1}\right]\int_{x_0}^{\infty} x^{k-b+1}dx \qquad \text{(2A4).}$$

If we want to know the variance, we have to calculate $E(x^2)$, but it is apparent from g 2A4 that this will diverge to infinity for $\boldsymbol{b} \leq 3$. Similarly, the expectation $E(x)$ will diverge for $\boldsymbol{b} \leq 2$. For $\boldsymbol{b} \leq 1$ we will not even have a proper statistical distribution:

any proper statistical distribution should satisfy $\int_{x_0}^{\infty} f(x)dx = 1$, while this integral diverges to infinity for any $f(x) = ax^{-b}$ with $b \leq 1$ and a>0. As apparent from subsect. 3.4.4, even this last point is much more than something of academic interest.

The reason to set always some lower bound x_0 is that $\int_{0}^{\infty} f(x)dx$ diverges to infinity for any $f(x) = ax^{-b}$ with a>0, regardless of b. This is no problem for practical applications, because any scaling feature ceases to hold below some scale. E. g. forest fires display a power law, as apparent from chapt. 3, which can be interpreted as a proof that the process that determines their size is essentially the same at any scale. But indeed this will not be the case anymore as soon as the scale is so small that the sizes of fires are comparable to the sizes of individual trees. Furthermore, the resolution of measures will always be limited. The analysis of forest fire data leaves little doubt that the registered sizes of the fires smaller than the unit of measure (e. g. smaller than 1 ha or 1 acre) are ordinarily subject to huge rounding errors. Fortunately, once we are within the scaling region we can choose any x_0 that we find convenient, without any loss of statistical rigor. Note that, regardless of the value of x_0 that we set in Eqs. 2A1-2A3, we will be equally able to estimate b and, once we have it, we will be able to deduce the complete distribution for any other x_0 within the scaling range.

On practice there will of course always be some upper cutoff x_M to the distribution. In some cases this will be large enough not to affect the empirical distribution, while in others it will be necessary to take it into account. If the last case, Eqs. 2A1-2A3 become:

$$f(x) = \left[\frac{b-1}{x_0^{-b+1} - x_M^{-b+1}}\right]x^{-b} \qquad (2A5)$$

$$F(x) = \frac{x_0^{-b+1} - x^{-b+1}}{x_0^{-b+1} - x_M^{-b+1}} \qquad (2A6)$$

$$P(X > x) = \frac{x^{-b+1} - x_M^{-b+1}}{x_0^{-b+1} - x_M^{-b+1}} \qquad (2A7).$$

Eqs. 2A5-2A7 converge with Eqs. 2A1-2A3 for $x_M \to \infty$. When there is an upper cutoff, all of the moments become finite (though their convergence will be slow, if the power law covers several orders of magnitude). Furthermore, the constraint b>1 does not hold anymore. Eqs. 2A5-2A7 are valid for any b except b=1. In this last case, we have

$$f(x)=[\log(x_M)-\log(x_0)]^{-1}x^{-1}$$

$$F(x)=\frac{\log(x)-\log(x_0)}{\log(x_M)-\log(x_0)}$$

$$P(X>x)=\frac{\log(x_M)-\log(x)}{\log(x_M)-\log(x_0)}$$

One of the most usual methods to obtain an image of the empirical distribution of a set of data is representing them in a histogram. However, when power laws are suspected, there is often a preference for representing the probabilities of excedence $P(X>x)$. This is straightforward to find, by ranking the data from larger to smaller, and representing the value x of each datum vs. its rank number (which is proportional to the estimated $P(X>x)$; if in addition the rank is divided by the number of data, we have a direct estimation of $P(X>x)$). This is put in a bilogarithmic scale, because it follows from Eq. 2A3 that $\log(x)$ vs. $\log(P(X>x))$ is a straight line with slope $-b+1$. An approximate fit to a straight line is then taken as an evidence of power law, and b is estimated by regression. The attraction of this method lies in that (i) it is easy and fast to apply, (ii) in contrast to the histogram, does not require to group the data by intervals, so we do not have the problem of choosing the intervals and all of the data are directly displayed in the figure, and (iii) the results are quite often apparently satisfactory. Unfortunately, it also has quite serious drawbacks, which are generally overlooked. On the one hand, when there is a significant upper cutoff x_M, which is often the case, the linearity of the bilogarithmic representation is lost, as apparent from Eq. 2A7, and there is some curvature that covers all of the graph, the more intense the smaller is x_M. A fair power law may thus appear to have a poor fit. On the other hand, the position of each spot in this graph is strongly constrained by the rest of spots, to the extent that the ordinates can only be a decreasing function of the abscissas, by construction. As a consequence, distributions quite different from a power law may seem to offer an acceptable fit; this may happen even when these are clearly unimodal. Some works even give the coefficient of correlation between $\log(x)$ and $\log(P(X>x))$ as an evidence for good fit, when in fact this coefficient is meaningless for this representation and will always overestimate the adjustment to the expected outcome. Therefore, this representation may have some applicability in preliminary exploratory analyses but, as a principle, it is not recommendable for other uses.

The clearest alternative consists of representing the empirical p. d. f. When the data follows other distributions closer to a Gaussian one the method of choice is the histogram, but in this other case there is quite a better procedure. We should take multiplicative intervals, estimate the density of probability for each interval, and represent x vs. $\hat{f}(x)$ (the estimated $f(x)$) in a bilogarithmic scale. I. e., the intervals should have the form $[x_A, x_B)$ with $x_B=kx_A$ for some constant k. If n_{AB} is the amount of data within this interval and N is the total data number, the estimated density of probability will be:

$$\hat{f}(x_{AB}) = \frac{n_{AB}}{N(x_B - x_A)},$$

which should be assigned to the (logarithmically) central value of x in the interval, $x_{AB} = \sqrt{x_A x_B}$. The bilogarithmic representation will be a straight line both if the power law is bounded (Eq. 2A5) and if it is not (Eq. 2A1). This has none of the problems of representing the probabilities of excedence and also has several advantages over a histogram: (i) the data are much more homogeneously distributed per intervals; instead, in the histogram of a power law there will be a huge concentration of data in the first few intervals, a large array of empty intervals, and a few intervals with one single data in very sparse positions, which is quite intractable; (ii) the spots occupy a much more regular position in the bilogarithmic plot, which allows to obtain a clearer image and to make regressions, which would not be so correct in a histogram; and (iii) the histogram introduces a spurious deviation from the expected straight line. This last point can be seen by calculating the expected result for the estimators when the width of the intervals is not negligible (both converge to the density function as this width vanishes). If we estimate the p. d. f. through a histogram per intervals $[x_A, x_B)$ with $x_B=x_A+k$, the expected value for the estimator will be:

$$\hat{f}(x) = \frac{1}{k} \int_{x-\frac{k}{2}}^{x+\frac{k}{2}} f(x)dx = \frac{\left(x - \frac{k}{2}\right)^{-b+1} - \left(x + \frac{k}{2}\right)^{-b+1}}{k\left(x_0^{-b+1} - x_M^{-b+1}\right)},$$

which only asymptotically becomes a straight line in a log-log plot. Conversely, if we take multiplicative intervals:

$$\hat{f}(x) = \frac{1}{k^{\frac{1}{2}}x - k^{-\frac{1}{2}}x} \int_{k^{-\frac{i}{2}}x}^{k^{\frac{i}{2}}x} f(x)dx = \left[\frac{k^{\frac{b-1}{2}} - k^{\frac{-b+1}{2}}}{\left(k^{\frac{1}{2}} - k^{-\frac{1}{2}}\right)\left(x_0^{-b+1} - x_M^{-b+1}\right)} \right] x^{-b},$$

which will produce a straight line of slope **b** in the log-log plot regardless of the size of the intervals. This will not however be the case if we have “genuinely” discrete data (i. e. the reason for their being discrete is not a finite resolution in the measures) which follow a discrete power law (not just a discrete approximation to a continuous power law), i. e. a set of probabilities $P(n) = \left[\sum_{i=n_0}^{n_M} i^{-b} \right]^{-1} n^{-b}$. Then the histogram will display a straight line of slope -**b**, while the representation with multiplicative intervals will suffer some deviation from the straight line for small n. However, the histogram will still have the other disadvantages above enumerated. Therefore, in this particular case the histogram is more appropriate to examine the lower end of the distribution, but the other representation is still better to obtain a global view. In the case of continuous variables (even if rounded up) this last representation can be considered superior to the histogram in all respects.

Our data will often be rounded to multiples of the unit of measure. It is thus recommendable to have the above multiplicative intervals also defined in terms of multiples of the unit of measure. So the chosen x_0 should either be the unit of measure or some multiple, and k should be an integral number. Otherwise, there would be spurious irregularities at the lower end of the empirical p. d. f. The integer number that allows for a highest resolution is k=2, so this will generally be the value of choice, unless the data are so scarce to require less resolution.

b can be estimated either by regression on the empirical p. d. f. or by maximum likelihood (m. l. e.).

The maximum likelihood estimator $\hat{b}$ of b is a function of the geometric mean g of our data. When there is no significant upper bound x_M to the power law, it reduces to:

$$\hat{b} = \frac{1}{\log(g) - \log(x_0)} + 1,$$

where "log" refers to the natural logarithm. Otherwise, it should be found iteratively by searching the value $\hat{b}$ that maximizes $\log\left(\frac{\hat{b}-1}{x_0^{-\hat{b}+1} - x_M^{-\hat{b}+1}}\right) - \hat{b}\log(g)$. In both cases it is immediate to find confidence intervals and perform contrasts of hypotheses, because the only source of variability of this estimator is the variability in log(g), which is just an average of N independent identically distributed random variables with a finite variance (the logarithm of each datum) and thus has a Gaussian distribution for large N (more exactly, it has a gamma distribution for $x_M \to \infty$, since it is a sum of exponential variables, but this readily converges to a Gaussian).

m. l. e. is statistically more powerful than a regression when our data fairly adjust to a power law. However, the estimation of $\hat{b}$ by regression on the empirical p. d. f. is much more robust in front of "imperfections" in our data set. In particular, rounding off is deadly for the m. l. e. of $\hat{b}$. Therefore, this should only be applied when x_0 is large as compared to the resolution of the measures. However, this is not always a guarantee of being free from rounding problems if the source of the data is dubious, since these could have been rounded also at higher scales. Official data sometimes include a suspiciously high frequency of values such as x=10, x=100, x=500, x=1000... as compared to e. g. x=18, x=246, x=819... Furthermore, we could be interested on estimating the parameters for the power law that best fits some set of data that deviates from a true power law to some extent. Most of the variability in a set of data might have been produced by some process with no characteristic scale, but there will always be some interference from other processes. In some cases this will be negligible, while in others it will be important but we may still be interested in fitting a power law in order to study the main process. As a conclusion, regression seems to be the best general purpose method, but it can be advantageously replaced by m. l. e. after carefully examining the data to ensure that they are well suited.

If we estimate b by regression, we can establish confidence intervals and perform contrasts of hypotheses by using the same nonparametric methods that can be applied to any other statistic (see e. g. Manly 1991). E. g. if we want to find out whether two sets of data A and B of size N_A and N_B have significantly different b, we only have to measure $\left|\hat{b}_A - \hat{b}_B\right|$ for the original data and for l surrogate pairs of sets of data generated by randomly redistributing the data from A and B in two sets that also have sizes N_A and N_B. Let l_+ be the number of surrogate pairs in which $\left|\hat{b}_A - \hat{b}_B\right|$ is equal or larger than the original. Then, the null hypothesis $\hat{b}_A = \hat{b}_B$ can be rejected with a level of confidence $e = \frac{l_+ + 1}{l}$.

Once we know b, we can test whether or not there is a significant upper bound x_M. Let max(x) be the largest value from our set of N data. Then it follows from Eq. 2A7 that we can reject the hypothesis $x_M \to \infty$ with a confidence level

$$e = \left(1 - \left[\frac{\max(x)}{x_0}\right]^{-b+1}\right)^N \qquad \text{(2A8).}$$

If we can reject, $x_M \to \infty$, we can also test whether the upper bound is higher than some given value x_M. From Eq. 2A7 we can reject this last hypothesis with a confidence interval:

$$e = \left(\frac{x_0^{-b+1} - [\max(x)]^{-b+1}}{x_0^{-b+1} - x_M^{-b+1}}\right)^N \qquad \text{(2A9).}$$

This has a huge practical importance, because it allows to estimate the worst possible disaster of some given kind in some given area, while the overall conditions do not change.

The fit to a power law can be tested with standard methods such as c^2. This requires distributing the data per intervals, with some minimum number of data per interval. Some authors recommend a minimum of 5 data per interval. One good procedure consists of taking the same multiplicative intervals used to estimate the empirical p. d. f. and joining any group with less than 5 data with some of its neighbors. If we transform logarithmically a set of data with a power law distribution, the result will have an exponential distribution. Therefore, we can also test the fit to a power law by taking advantage of some of the methods that have been specially developed to test exponentiality (e. g. Margolin and Mauer 1976, Marlin 1984).

The power law or “Pareto” distribution without upper bound can be seen as a particular case of the “generalized Pareto distribution” gPd (Pickands III 1975), whose probability of excedence is:

$$P(X > x) = \exp\left(- \int_0^{x/x_0} \left[(1 + ct)_+ \right]^{-1} dt \right).$$

This distribution can take three different forms depending on the parameter c. For c=0 we have a gPd type I, which is an exponential distribution. For c>0 we have a gPd type II, which is a power law distribution with $b = \frac{1}{c} + 1$. For c<0 we have a gPd type III, which is a distribution with an upper bound. All or almost all of the statistical distributions that have been found to have practical or theoretical interest converge with one of these three distributions for asymptotically large values of x (the typical “textbook” distributions without an upper bound, such as the Gaussian, Poisson, gamma, c^2, t, F, etc., converge with an exponential distribution for large x). The gPd is related to the extreme value generalized distribution (evgd), which is the asymptotic distribution, for $N \to \infty$, that acquires the maximum of N data distributed in any way, provided that it asymptotically acquires some definite distribution (Galambos 1987). The evgd also has three types, with a correspondence with the three types of the gPd. The maximum of a set of values whose distribution converges to a given type of gPd for large x will converge to the corresponding type of evgd for large N. Conversely, an evgd of a given type converges to the gPd of the corresponding type for a large x.

In view of this classification of asymptotic distributions, the hypothesis of an exponential distribution is the natural option to contrast with the hypothesis of a power law distribution. We will be in far more solid grounds to claim that the distribution of a given set of data is close to a power law if, in addition to pass some test of fit to a power law, it does not pass the test of fit to an exponential. While the plots of both the p. d. f. and the probabilities of excedence of a power law correspond to a straight line in a log-log scale, in the case of an exponential distribution both correspond to a straight line in a loglinear scale (log(f(x)) or log(P(X>x) vs. x). The representation of these plots is highly informative and often makes superfluous any more formal test.

If we want to formally test the power law hypothesis vs. the exponential hypothesis just for the part of a statistical distribution between some bounds (x_0,x_M), and these are not preestablished, the appropriate choice of (x_0,x_M) poses an additional problem. In Pueyo (1994) I developed a method to compare both hypotheses in the case of the upper end of a distribution (i. e. assuming $x_M \rightarrow \infty$), and applied it to the sets of streamflow data for two rivers (Cabriel, in Valencia, and Ter, in Catalonia). I sorted the data from the largest to the smallest and took successively each one as x_0. For each x_0 I tested both the exponential and the power law hypotheses. I plotted the number N of data entering each test vs. the levels of significance e that allowed to reject each of the hypotheses. Note that, if a hypothesis can be rejected, e will be small, while, if the hypothesis is correct, it will take at random any value between 0 and 1. The plots displayed three well defined regions. For small N there were oscillations between 0 and 1, as could be expected because of the lack of data. In a second region, the exponential e vanished and the power law e followed a random walk between 0 and 1. Beyond some point, both e became nearly null. This was a proof that floods in these rivers have a distribution close to a power law, which did not require to previously decide a boundary between “floods” and “baseline flow”.

The power law distribution has a rather peculiar behavior when summing up or averaging data (Mandelbrot 1983). The central limit theorem is generally taken to imply that everything will converge to a Gaussian distribution under these operations. However, one requirement for this to hold is that the variance should be finite. Instead, any distribution that, for asymptotically large x, converges to a power law with b<3, will have an infinite variance. In this case, there will still be a convergence to a well defined and predictable distribution, but this will not be Gaussian. The distribution to which such a sum or average converges is called Lévy-stable (Mandelbrot 1983, DuMouchel 1973). The Gaussian distribution is a limit case of Lévy-stable distribution, found for some given values of its parameters. This is the only Lévy-stable distribution with no power-law tail. A Lévy-stable distribution with a power-law tail is the best candidate for measures that aggregate several variables whose largest values also display a power law, at least if b is appreciably smaller than 3 (otherwise, the convergence to this distribution when summing or averaging is very slow, see DuMouchel 1973). On practice, there will always be a higher bound x_M to the distribution and thus a finite variance, but power laws sometimes span many orders of magnitude and the convergence to a Gaussian will then be really slow. Therefore, at least for b appreciably smaller than 3, there will be an intermediate range of number of summed or averaged data in which the distribution will be Lévy-stable with an upper bound at x_M, and thus have little resemblance to a Gaussian curve. Mantegna and Stanley (1995) offer an empirical example.

The mean is a measure of scale, so its meaning for a scaling distribution is not as immediate as for other distributions (if there were no upper bound x_M, this would not even converge for $b<2$). Any other distribution has some characteristic scale, and will readily converge under addition or averaging to a Gaussian distribution, in which there is a coincidence between all of the measures of scale, such as the mean, the median or the mode. Instead, for a power law the mean is far above the median (let alone the mode, which is at the lower end of the distribution) and corresponds to a value rarely attained. For example, the author (Pueyo 1994) found that, throughout the Valencian part of the Xúquer basin (a Mediterranean area), the mean streamflow typically exceeds the median by 30% (although only the upper end of the distribution approaches a power law). Therefore, the average streamflow misrepresents the typical conditions found by the river biota. Something more worrying is that it may easily lead to overestimate the potential extractions from rivers in which the water from floods is not usable, since these make up an outstanding contribution to the mean. In particular, this is one of the deficiencies of the Spanish government project of massive water transfer from the Ebre river to other basins (see Ibáñez et al. 1999).

Nevertheless, in some cases the mean is really relevant, at high spatio-temporal scales. It is e. g. amazing the great importance that the mean fire size seems to have in forest dynamics, if we take into account that some of the features of fire size distributions apparent from chapt. 3: the scaling region often spans many orders of magnitude (e. g. five in Fig. 3.1), and b is so small in all of the cases analyzed that the average fire would be infinite if there were no upper bound x_M. However, the mean fire size seems to be actively regulated by the own forest, mainly by modifying x_M, in such a way that the mean rate of burning does not change very much in front of long term changes in the frequency of ignitions (see chapt. 3).

One distribution that should be mentioned in connection with the power law is the lognormal. This shares with the power law the basic pattern with a high frequency of very small values and a low frequency of very large values, though not so exaggeratedly. Its behavior at the limit of very large x belongs to the exponential class instead of the power law class (Galambos 1987), but the convergence with the exponential distribution is extremely slow. For practical purposes, it should be considered much closer to a power law, and the developments in chapt. 5 even suggest a high proximity between both in terms of generating mechanisms in nature.

Chapter 3

Catastrophes, succession and criticality: the case of wildland fires

3.1. Introduction

It may seem surprising the inclusion of a chapter about wildland fire in a book about fundamental issues of ecological theory. This has two major justifications. On the one hand, fire can be expected to be one of the key factors for the future of the Biosphere, in contrast to the little attention it receives in the ecological literature. On the other hand, the study of fires opens a window to nothing less than the nature of catastrophes in general. This “window” is not just one of philosophical interest. Here I make the whole way from basic theory to practical tools for everyday fire prediction, which are shown to work with empirical data and must be extensible to other kinds of catastrophes. The results expounded in this chapter defy the usual perception of catastrophes as a matter of fate, independent from human decisions. Instead, human action changes the regimes of catastrophes. I show that, to some extent, these changes are predictable and, hence, potentially decidable.

Margalef (1991, 1997) postulates that there is an asymmetry of change in ecological systems and complex systems in general, between periods of succession, with a gradual and directional increase of complexity, and events of sudden destruction. The physical roots of this contrast is treated in subsect. 2.3.1. He adds another intriguing observation (Margalef 1986): large catastrophes are rarer than small catastrophes, in such a way that, if we represent the frequency of catastrophes of each magnitude vs. magnitude in a log-log scale, we find a straight line. In other words, Margalef (1986) suggests a power law statistical distribution for catastrophes in general, as well known since long ago in the case of earthquakes (Scholz 1990). The presence of power laws has been confirmed for several kinds of catastrophes, besides earthquakes (see e. g. Turcotte 1999).

Some years ago, Bak et al. (1987, 1988) (see also Bak 1996) discovered a kind of dynamics that could explain power law catastrophes and other scaling features observed in nonequilibrium systems: self-organized criticality (SOC). This is explained in subsects. 3.2.1. – 3.2.2. below. There is some evidence of SOC in many kinds of catastrophes: extreme rain events (Peters et al. 2002), landslides (Guzetti et al. 2002), wildland fires (Malamud et al. 1998, Ricotta et al. 1999), mass extinctions (Solé et al. 1997), epidemics in islands (Rhodes and Anderson 1996), wars (Turcotte 1999), economic oscillations (Bak 1996), etc.

Soon after the findings by Bak et al. (1987, 1988), Drossel and Schwabl (1992) (see also Clar et al. 1996) developed a SOC “forest fire” model. This did not pretend to accurately reproduce fire dynamics: it rather took forest fires as a metaphor to explore this model for its theoretical interest. On the other hand, Robertson (1972, quoted in Strauss et al. 1989) had already found that the sizes of forest fires do in fact approach a power law, as would be expected from SOC. My work has consisted of taking Drossel and Schwabl’s model as a point of departure to build a model that pretends to represent wildland fires with more realism, and comparing its outcome with observations from empirical data. While this project was in course, the papers by Malamud et al. (1998) and Ricotta et al. (1999) enlarged the evidence of power laws in fires and conjectured that this would be a consequence of a SOC dynamics that could be explained by Drossel and Schwabl’s model. Here I show that wildland fires agree with predictions for my more developed model not just in this feature but also in several others, and indicate the far-reaching consequences that this has for dealing with fire on practice.

The main concern that pushed me to this area was the recent series of events that are taking place in some of the main tropical rainforest areas in the world, which are passing in a few years from being virtually immune to fire to host the most extreme fires ever registered, because of human impacts (Uhl 1998, Nepstad et al. 1999, Siegert et al. 2000, Page et al. 2002). It is urgent to understand what is going on in order to prevent it, and also to develop a medium-term forecasting system for minimizing damage in the cases in which it is too late for prevention. Unfortunately, tropical rainforest fire data are not as readily available as data from other kinds of forests, so the main proof of the applicability of my model to the former is that it works for the lasts. In particular, I hypothesize that these rainforests are undergoing a phase transition in the strict physical sense of the word. I support the idea that an ecosystem can suffer a phase transition with empirical evidence that Yellowstone was at the neighborhood of such a transition when the polemical 1988 fires unfolded.

Regardless of its interest for tropical rainforests, the results from this chapter are indeed useful for other kinds of wildlands. For most of them, fire can be argued to be the main determinant of successional dynamics, as justified in the discussion. In some countries, huge efforts are devoted to fire management. I consider that current management criteria should be revised in the light of the findings here expounded. I also expect these findings to be generalizable to other kinds of catastrophes. If this expectation is fulfilled, the results from this chapter could well help to save many lives and avoid many other damages, since few things have such a practical importance and are as little known as the dynamics of catastrophes. This is apparent from the forecasts on global climate change and its consequences, in which the future evolution of the extreme events of all kinds remains one of the most elusive points (see sect. 9.3.6 in IPCC 2001).

The order of magnitude of the number of simulations required for this work has been about one hundred, besides empirical data analyses, bibliographical review and theoretical work. Of course I do not present all of the simulations, but rather outline the emerging patterns and illustrate them with some instances. Some technical details are given in App. 3A.

The chapter is organized as follows:

- Sect. 3.2.: *Previous related models* introduces the forest fire models in the literature based on percolation theory and self-organized criticality, which constitute the point of departure for my developments.
- Sect. 3.3.: *From simulations to reality?* discusses the ability of current models to explain the empirical observations on fire regimes. I include a first analysis of boreal forest fire data using a method more accurate than those in the literature.
- Sect. 3.4.: *A predictive model of succession and fire dynamics* introduces a new model, analyses its main features, including its response to environmental variations, contrasts these features with empirical data from Alberta’s boreal forests and from Yellowstone, and discusses what these findings teach us about the nature of self-organized criticality and about the adequacy of this concept to capture the dynamics of wildland fires.
- Sect. 3.5.: *Anthropogenic factors* analyzes the implications of my model for the effects of human activities on fire regimes, and contrasts some of the expectations

that arise from theory and simulations with data from Yellowstone and from the Mediterranean region.

- Sect. 3.6.: *Disaster in the tropical rainforests?* raises the hypothesis that some of the main tropical rainforest areas are becoming susceptible to suffer second order phase transitions due to anthropogenic impacts.
- Sect. 3.7.: *Discussion and conclusions* as usual, with emphasis on the consequences of the findings in this chapter for fire management strategies.

3.2. Previous related models

3.2.1. The percolation model of forest fire

The percolation model of forest fire (e. g. Stauffer and Aharony 1992) was the basis for the posterior application of the self-organized criticality approach to this field.

The concept of percolation can be introduced with the game of the squirrel, which is inspired by the legend that, long ago, a squirrel could travel from end to end of the Iberian Peninsula by jumping from tree to tree. Take a square lattice such that each cell (i,j) can either contain a tree or be empty. The trees are located at random, and the probability of a cell containing a tree is r. Consider that you are a squirrel in a cell with a tree. You can only move to other cells with trees, and only if they are side-by side with the cell where you are. I. e. you can move from (i,j) to any cell with a tree out of the set $\{(i-1,j), (i+1,j), (i, j-1), (i, j+1)\}$. The game consists of finding some path to cross from one side of the lattice to the opposite one. When a lattice has any such path, it is said to percolate.

This game offers many surprises. In a speech I asked the audience to split into two teams and I distributed among them several sheets with different 30x30 lattices generated at random with $r \approx 0.6$. Team B found paths from side to side for 8 lattices out of 10. Despite its effort, team A could only find half this amount of paths. Team B was not really integrated by better “squirrels”: the key was a subtle difference between both sets of lattices. The lattices for team A were produced by taking r=0.57, while for team B I took r=0.61. These few hundredths of difference produced a radical change in the outcome, and the difference would have been even much more radical, had the lattices been larger. For large enough lattices, team A would have found all of the lattices to percolate, while team B would have found no percolating lattice.

In physical terms, this simple setting displays a phase transition, like the phase transition from ice to liquid water at a well defined temperature, which is 0^oC at standard pressure. There is an equally well defined “critical” threshold for percolation: r_c=0.5927... At the limit of infinite lattice size, this threshold strictly separates percolating from non-percolating lattices. r_c is modified if we take different designs, such as allowing fire to propagate in diagonal, taking non-square cells, etc., but in any case we have a phase transition with similar features.

The use of this model for a first approximation to simulate forest fires is obvious. We can establish that, when a lighting strikes on a cell with a tree, fire will propagate to all of the cells with trees that are side-by-side with this first one, and will proceed to successive layers of cells according to the same rules. The result is that, for low to moderate r, fires cover just a few cells. At a small distance from r_c, the sizes s grow steeply. For $r > r_c$, there is a non-vanishing probability P_∞ of a fire becoming infinite, if the lattice is infinite, because there will always be some percolating set of connected cells with trees. If the lattice is finite, the average size of fires $\bar{s}$ will depend on the size of the lattice for $r > r_c$, but will have “internal” limits for $r < r_c$.

In this model, r is a “control parameter” that we can modify, while P_∞ is an “order parameter” that changes as a function of r and displays a phase transition at $r = r_c$. If P_∞ jumped discontinuously at $r = r_c$, we would have a “first order” phase transition. In this case, there is no discontinuity. However, P_∞ is null for $r < r_c$ and an increasing function for $r > r_c$, so $\frac{dP_\infty}{dr}$ is not defined at $r = r_c$. When a function is continuous but not derivable for a given value of the control parameter, we have a “second order” phase transition. Using the standard vocabulary for second order phase transitions, we can call our system critical when $r = r_c$, subcritical when $r < r_c$ and supercritical when $r > r_c$.

The dependence of $\bar{s}$ on r for $r < r_c$ has the following form, which is characteristic for second order phase transitions:

$$\bar{s} = a\left|r - r_c\right|^c \qquad (3.1),$$

with χ<0. The abrupt transition in $\bar{s}$ above described corresponds formally to a small $|c|$, which can be better understood from subsect. 3.4.3.

The probability density function (p. d. f.) $f(s)$ of fire sizes s has the form:

$$f(s) \propto s^{-b} e^{ls} \qquad (3.2),$$

where b is a constant $b = \frac{96}{91} \approx 1.05$, while l depends on $|r - r_c|$: l>0 for $r > r_c$, l=0 for $r = r_c$ and l<0 for $r < r_c$. At supercriticality, Eq. 3.2 is not a proper density function, because it diverges for $s \to \infty$ (for any prefactor larger than zero), and so does the distribution function (the cumulative probability function). At criticality, Eq. 3.2 becomes a power law. At subcriticality, the power law is modified by the exponential term. The part of the distribution that can be assimilated to a power law readily shrinks when decreasing r below r_c, which is linked to the sharp decrease in $\bar{s}$ in Eq. 3.1.

An universal feature of systems at criticality is the lack of any characteristic scale. The power law distribution satisfies this property and is associated to a fractal structure of the set of occupied cells in the lattice, which is lost whenever $r \neq r_c$.

Besides theory and simulations, the existence of a critical percolation threshold has also been found in controlled combustion experiments (Nahmias et al. 2000). The possible importance of percolation thresholds for ecosystem features other than fire, such as biological connectivity as a factor for species persistence, has received considerable attention from landscape ecologists (see With and King (1997) for a review).

3.2.2. The standard SOC forest fire model

Bak et al. (1987, 1988; see also Bak 1996) introduced the concept of self-organized criticality (SOC). SOC takes place in systems able to undergo a second order phase transition, whose control parameter is modified by the own dynamics of the system in such a way that it tends to achieve the critical state or, perhaps more properly, a state that shares the main properties of the critical state. SOC has been proposed as a general mechanism for the origin of scaling features in systems far from thermodynamic equilibrium.

Drossel and Schwabl (1992) (see also Clar et al. 1996) proposed a SOC model of forest fires, working on an earlier model by Bak et al. (1990). This model is a dynamic version of the above percolation model. At each time step, each empty cell can become occupied by a tree with a probability $p<<1$, and each occupied cell can be ignited by lighting with a probability $f<<p$. If a tree is ignited, fire propagates according to the rules in subsect. 32.1 until it finds no more trees to burn, and all of the burnt cells become empty cells. There are neither new trees nor new lighting until the fire extinguishes, which is a simple way to represent a separation of time scales: the propagation of fire is extremely fast as compared to the growth of trees.

In the absence of fire, r would grow across time up to $r=1$. However, the sharp increase in burning close to r_c does not allow the system to surpass this point. As a first approximation, the system could be expected to keep trapped in the neighborhood of r_c, and thus share all of the features of a simple percolation system close to criticality, but in fact there are some differences because the dynamics of the system generates correlations in the positions of the trees. In particular, r is appreciably smaller than r_c (e. g. Clar et al. (1994) set $r \approx 0.4$, although the exact figure changes with $\frac{f}{p}$; compare with $r_c \approx 0.6$), and b (Eq. 3.2) is not exactly the same. The main interest of SOC systems lies in that they nearly preserve the main property of critical systems, which is scaling.

On the other hand, strict scaling is only expected at the limit $\frac{f}{p} \to 0$ (in infinite lattices). The outcome of the simulations is well approximated by the density function:

$$f(s) \propto s^{-b} C(s/s_{max}) \qquad (3.3),$$

where C is a cutoff function that sets an upper limit to scaling close to some given value s_{max} (Drossel and Schwabl 1992, Clar et al. 1996). This is just an approximation, because there is some small deviation from the power law also below this upper limit (Pastor-Satorra and Vespigniani 2000). When Eq. 3.3 becomes an exact power law, at $\frac{f}{p} \to 0$, we have $s_{\max} \to \infty$ and an exponent that has been estimated to be $b \approx 1.08$ (Pastor-Satorra and Vespigniani 2000). Above the limit

$\frac{f}{p} \to 0$, the b that best fits the density function will be a little higher. It is not possible to give an exact value for b, because this will depend slightly on $\frac{f}{p}$ and on the range of s chosen to estimate it. As far as I know, the largest estimate in the literature is $b \approx 1.2$ (Grassberger 1993).

The variations in the parameters in Eq. 3.3 as a function of $\frac{f}{p}$ affect $\bar{s}$ in such a way that the total rate of burning becomes nearly independent from the frequency of ignitions (Drossel and Schwabl 1992, Clar et al. 1996).

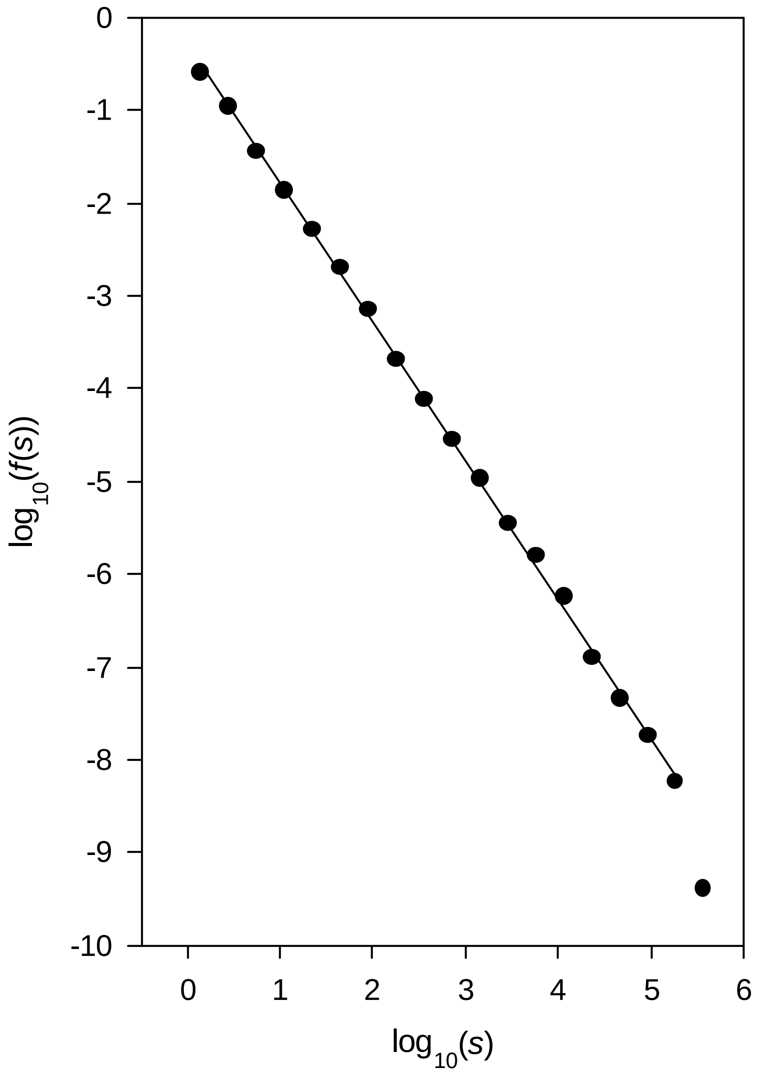
3.3. From simulations to reality?

Several authors have shown forest fire sizes to display bounded power laws (Robertson 1972 (quoted in Strauss et al. 1989), Strauss et al. 1989, Malamud et al. 1998, Ricotta et al. 1999), by using several techniques.

On the other hand, their results are not always completely clearcut, specially those by Ricotta et al. (1999, 2001), who searched for straight parts in the log-log plot of the estimated probabilities of excedence for each empirical value. As shown in App. 2A, this plot is only expected to be straight for unbounded power laws, while the plot for bounded power laws such that of Eq. 3.3 display some curvature (for all of its range of values, not just the upper end).

More clearcut results can be obtained by taking advantage of the statistical techniques put forward in App. 2A, which I use throughout this chapter. In particular, I estimated the p. d. f. of several sets of data as there described, and this left little doubt that they displayed power laws. In Fig. 3.1 I show the empirical p. d. f. for one of the largest sets, the set of wildland fire sizes in Alberta (Canada). This p. d. f. displays a precise scaling behavior for 5 orders of magnitude, which is quite a high figure, if we take into account that most of the purported "fractals" in the physical literature span little more than one order of magnitude. It is found to have a significant upper cutoff (e=4.5x10^{-6}), by application of the criterion in App. 2A.

Fig. 3.1. *Wildland fire size distribution in Alberta (Canada). Empirical probability density function (p. d. f.) for the sizes (ha) of the 9333 fires that burnt at least 1 ha from 1966 to 1999. The estimated parameters are b=1.50 and* s_{max}*=409154 ha.*



At first sight, these results could be thought to confirm that the above model is valid (as suggested by Malamud et al. 1998 and Ricotta et al. 1999). However, this is far from obvious, for several reasons:

- The simple existence of a power law is a weak proof, because it could have other origins.
- In most instances, the observed parameters do not coincide with expectations from the model. Instead, the empirical data typically display larger values of b, as mentioned by Malamud et al. (1998) for the sets they analyzed. This is also the case for Alberta, with b=1.5.
- The standard SOC "forest fire" model took forest fires just as a metaphor. The mechanisms on which it is based are not and have never pretended to be realistic for forest fires.

Ricotta et al. (1999) also mention as a problem the existence of an upper bound, but in fact this agrees with the expectations from the model.

In view of these problems, I have developed a more realistic forest fire model, which also displays SOC for a part of its parameter space, and I have compared simulated and empirical data, both for the overall p. d. f. and for other features. The results strongly support the validity of the model and furnish useful predictive tools.

3.4. A predictive model of succession and fire dynamics

3.4.1. The model

The most obvious difference between Drossel and Schwabl's model and real-world forests is that trees do not arise suddenly here and there. Instead, we may expect a gradual and ubiquitous accumulation of fuel, and a progressive modification of its features, in agreement with the "fuel succession" conceptual model (e. g. Schimmel and Granström 1997, Agee and Huff 1987, Romme 1982). In some boreal coniferous forests there is clear evidence of increasing fire hazard with stand age for several decades after fire (Niklasson and Granström 2000). Fire behavior is also strongly conditioned by other variables than fuel, in particular by weather (Bessie and Johnson 1995).

In the models in sect. 3.2 the probability of a cell catching fire when some of its neighbors is burning is either 0 or 1 (depending on whether or not it contains a tree). Instead, each cell should be assigned some probability r taking any possible value between 0 and 1.

If all of the cells have the same r, this can be taken as a control parameter in the same way as r in sect. 2.1., and the result is the same (see the comments on "bond percolation" in Stauffer and Aharony 1992). I. e., there is also a second order phase transition for a critical value r_c, and Eqs. 3.1-3.2 remain valid. For the kind of lattice we are considering, r_c=0.5, but this depends on the particular shape and has no practical significance.

In order to model fuel succession in isolation, we can take the model in subsect. 3.2.2. with some modifications:

- Lightings strike on each cell at random moments in a continuous time.
- A cell receiving a lighting burns with probability r.

- Fires evolve in a series of discrete steps, which are considered to have a null duration. At each step, fire propagates from each burning cell (i,j) to each of its unburnt neighbor cells (k,l) (4-neighborhood, i. e. (k,l) can take the values $\{(i-1,j), (i+1,j), (i, j-1), (i, j+1)\}$) with probability $r(k,l)$, and then extinguishes in (i,j). This process continues until the fire is completely out. I establish periodic boundary conditions.
- After a fire, all of the burn cells are assigned $r=0$.
- Between fires, each r evolves in agreement with some continuous function with an attractor at some value r_∞.

I tested several functions for the evolution of r in the absence of fire and all of them produced the same results, so in the rest of the paper I use the simplest one: in the absence of fire, r increases at a constant rate until r_∞, and then stops. In particular, a logistic equation gives the same results.

The simulations indicate that this model has the same outcome as the standard SOC forest fire model in subsect. 3.2.2, if and only if $r_\infty > r_c$. Otherwise, it behaves similarly to a simple percolation model. So, as a principle, all of the developments in the literature concerning the model by Drossel and Schwabl (1992) can be also applied to this model, which simulates in a simplified but robust form a process known to work in nature, fuel succession.

It is not essential that $r=0$ after the fire, for this kind of dynamics to develop. There is enough with a post-fire period with $r<r_c$, long enough for the duration of fires to be neglected as compared to this period.

The main limitation of the above model is that it does not take into account environmental variations, which are known to have a strong impact on fire dynamics (Bessie and Johnson 1995). This can be solved if r is decomposed into a component that is determined by the self-organization of the own system and another one that is controlled from outside. The outcome also appears to be robust to the particular criterion chosen to combine both components, so I take a simple addition:

$$r=r_i+r_e \tag{3.4}$$

where r_i is the internal or “fuel” component and evolves in agreement with the above model, and r_e is the external or “environmental” component that is established by the modeller.

In the rest of sect. 3.4 and in sect. 3.5 I analyze the features of this model and compare them with data from real-world forests.

3.4.2. The short-term impact of environmental variations in model and reality

I performed a simple but illustrative simulation in which r_e was constant in space and time during a fire but was chosen at random for each fire, within some bounds.

The result is again a p. d. f. close to a power law, but b increases and is affected by the precise distribution of r_e. This solves the discrepancy between simpler models and the empirical observations. E. g. Fig. 3.2 displays the results of a simulation, with $b \approx 1.4$.

If we distribute fire sizes as a function of r_e (in intervals of r_e with constant width) we obtain a set of power laws, such that b decreases and s_{max} increases when r_e increases (Fig. 3.3A-C). The result is qualitatively the same if the forest self-organizes under a null r_e, and r_e is modified only for the fire to be registered (and we continue the simulation as if this fire had never happened). I. e. the pattern in Fig. 3.3A-C does not require a previous history of environmental fluctuations. This response to the control parameter r_e is qualitatively different from the standard response in percolation theory (subsect. 3.2.1), which would lead us to expect a constant b.

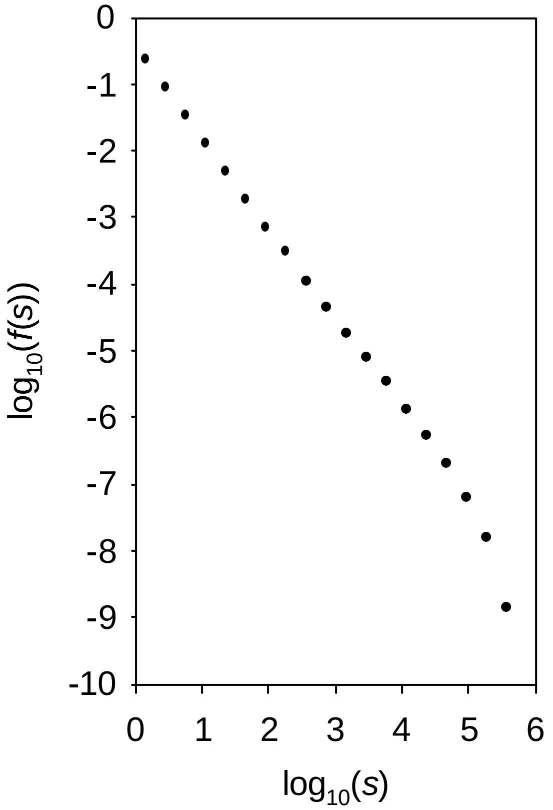


Fig. 3.2. *Simulated fire size p. d. f. under short-term environmental fluctuations. In this case,* $b \approx 1.4$.

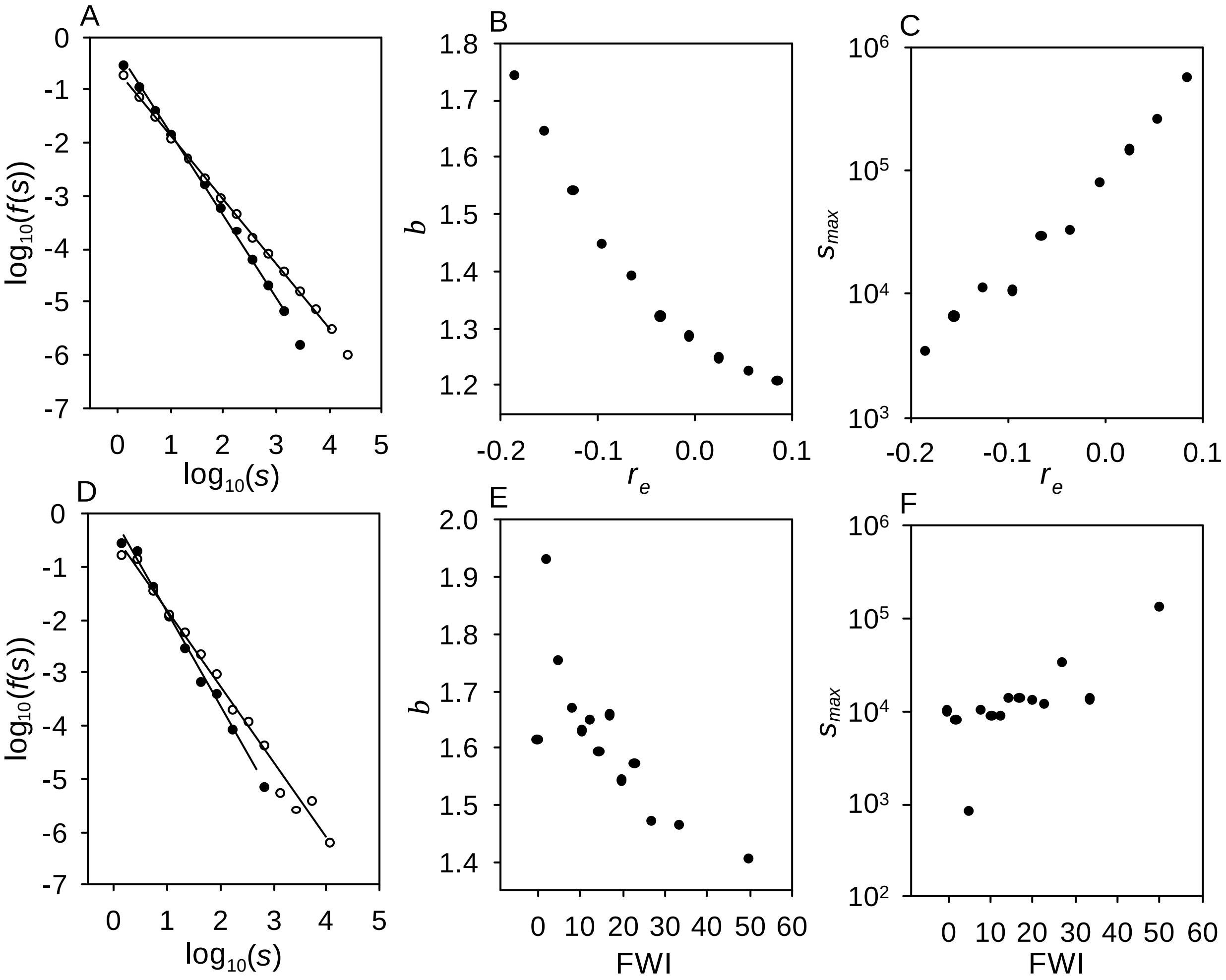


Fig. 3.3. *The short-term effect of environmental fluctuations on fire regime in a simulation and a real case. The real fires are the 3,134 fires that burnt at least 1 ha in Alberta from 1983 to 1995. The environmental term for the real fires is the Canadian Fire Weather Index (FWI).* (**A**) *p. d. f. for simulated fire sizes, with r_e from –0.14 to –0.11 (solid circles) and r_e from 0.01 to 0.04 (open circles).* (**B**) *relationship between r_e and b in simulated fires. (**C**) relationship between r_e and s_{max} in simulated fires.* (**D**) *p. d. f. for real fire sizes in ha, with FWI from 4 to 6 (solid circles) and FWI from 30 to 37 (open circles).* (**E**) *relationship between FWI and b in real fires. (**F**) relationship between FWI and s_{max} (ha) in real fires.*

I performed a similar analysis with data from the wildland fires in Alberta, by taking the Canadian Fire Weather Index (FWI) (Van Wagner 1987) as an approximation to the atmosphere-driven component of r_e (in this case, I chose the intervals of FWI in such a way to have a similar number of data per point). This index is obtained solely from atmospheric data. The result is again a set of power laws (Fig. 3.3D). The relationships between parameters (Fig. 3.3E-F) are similar to those found in the model. The precise relationships change through time, probably due to long-term changes in the forest structure or to modifications in fire suppression, so I display the results just for one of the databases in which the official Alberta fire register is divided, which spans from 1983 to 1995, instead of the complete register like in Fig. 3.1.

This result supports the model and has an immediate application. The complete fire size distribution can be predicted from a routinely calculated weather index.

3.4.3. Phase transitions in fire dynamics, and the nature of SOC

As explained in subsect. 3.2.1, simple percolation models display a phase transition for a critical value of the control parameter. If forest fires followed this model, above criticality there would be some probability of a fire able to sustain itself indefinitely and expand as far as allowed by external constraints, such as the total size of the forested area or the duration of the fire season. In subsect. 3.4.2 I show that the response to the control parameter in simulated SOC forests and in an instance of real-world set of wildlands is qualitatively different from the response in simple percolation models, but we can still wonder whether these systems can suffer a second order phase transition and which will be the features of this transition. This is useful both for fire prediction in extreme situations and for a better understanding of the nature of SOC .

A second order phase transition cannot be strictly observed in systems with finite size, because these cannot sustain an infinite fire. However, we can test whether some equation with the same form as Eq. 3.1 is satisfied while fires are still reasonably smaller than the maximum allowed by outer constraints. If this is the case, there will be some critical value r_c’ for $\bar{r} = r_e + \bar{r}_i$, where $\bar{r}_i$ is the average or r_i over space in the SOC state. Since the only parameter that we manipulate is r_e, this is equivalent to a critical value $r_c' - \bar{r}_i$ for r_e:

$$\bar{s} = a\left|\left(r_c' - \bar{r}_i\right) - r_e\right|^c \qquad (3.5).$$

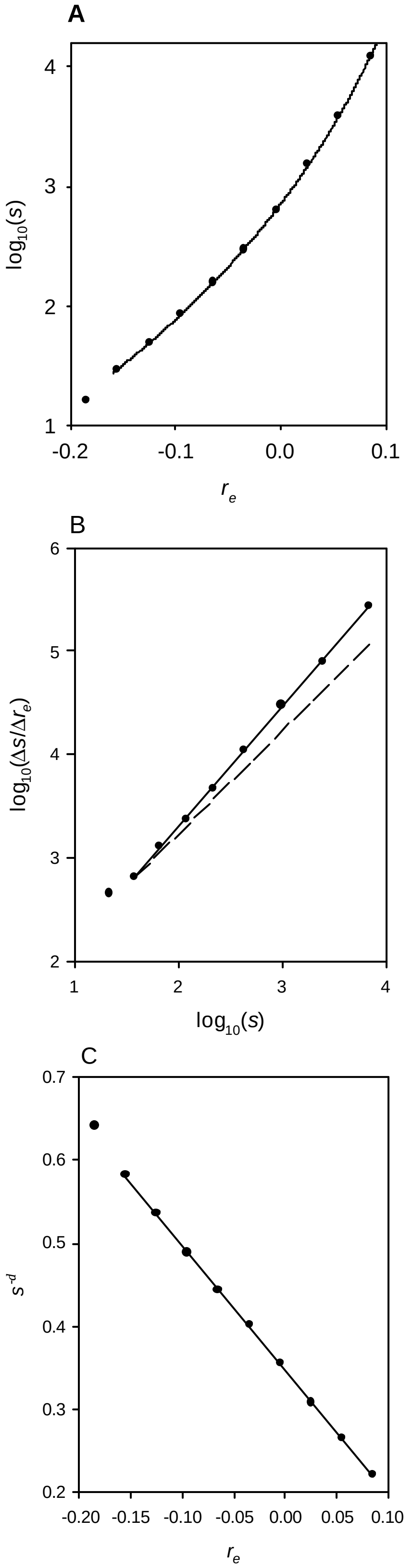


Fig. 3.4. *Evidence of phase transition in simulated forest fires. Explanation in the text.*

In Fig. 3.4A I represent r_e vs. $\log(\bar{s})$ for the same data on simulated fires that was used for Fig. 3.3A-C. The increase of $\bar{s}$ with r_e is faster than exponential (which would correspond to a straight line in Fig. 3.4A), as expected from Eq. 3.5.

From Eq. 3.5,

$$\frac{d\bar{s}}{dr} = -ac\left(\left(r_c' - \bar{r}_i\right) - r_e\right)^{c-1} \qquad (3.6).$$

If we introduce Eq. 3.5 into Eq. 3.6, we have

$$\frac{d\bar{s}}{dr_e} = -\left(a^{\frac{1}{c}} c\right)\bar{s}^{\frac{c-1}{c}}$$

We define $k = -a^{\frac{1}{c}} c$ and $d = -\frac{1}{c}$, so

$$\frac{d\bar{s}}{dr_e} = k\bar{s}^{1+d} \qquad (3.7).$$

d quantifies the difference from a simple exponential growth of $\bar{s}$ as a function of r_e. Since the data in Fig. 3.4A are distributed into groups that correspond to intervals of r_e, we can approximate the differentials in Eq. 3.7 by differences between successive groups and have:

$$\log\left(\frac{\Delta\bar{s}}{\Delta r_e}\right) = \log(k) + (1+d)\log(\bar{s}) \quad (3.8).$$

Fig. 3.4B displays $\log(\bar{s})$ vs. $\log\left(\frac{\Delta\bar{s}}{\Delta r}\right)$ for our simulated data. The result is well matched by a straight line, as expected form Eq. 3.8 (leaving apart the spot that corresponds to a smallest $\bar{s}$). Furthermore, $\boldsymbol{d}$ is clearly higher than 0 (discontinuous line), as corresponds to an increase faster than exponential. Both results together constitute a strong proof that this system is susceptible to suffer a second order phase transition. By simple regression I estimate $\boldsymbol{d} \approx 0.16$.

From Eq. 3.5,

$$\bar{s}^{-d} = \left(k\boldsymbol{d}\left(r_c' - \bar{r}_i\right)\right) - \left(k\boldsymbol{d}\right)r_e \qquad (3.9).$$

Fig. 3.4C displays r_e vs $\bar{s}^{-d}$ for the estimated $\boldsymbol{d}$ and confirms the straight line expected from Eq. 3.9. If we perform a simple regression and we divide the independent term by minus the slope, we get an estimation of the critical point for r_e, $r_c' - \bar{r}_i \approx 0.23$. Since $\bar{r}_i \approx 0.30$, this corresponds to $r_c' \approx 0.53$.

It is also immediate to extract the estimated k from the slope and $\boldsymbol{d}$, $k \approx 9.34$, so we have all of the parameters in Eq. 3.5 ($a \approx 0.08$, $\boldsymbol{c} \approx -6.23$, $r_c' - \bar{r}_i \approx 0.23$). The estimated curve has been superimposed on the empirical data in Fig. 3.4A. There is a clear agreement between both.

As a conclusion, there is evidence of a second order phase transition in the analyzed SOC system. The estimated critical r is close to the critical r for simple percolation. We are still missing confidence margins and simulations for other parameter values, but what seems clear is that the critical r is well above the range of r found by the system with some frequency. The main difference with a simple percolation model is that $\boldsymbol{d}$ is much smaller ($\boldsymbol{d} \approx 0.16$, as compared to $\boldsymbol{d} \approx 0.44$ for simple percolation), which indicates a much smoother response to changes in r.

The interpretation that emerges for the self-organized "critical" state in this model is that it is in fact quite far from the critical state, which remains well-defined once the system has self-organized. This process of self-organization consists of a generation of correlations that make the response function smoother, which decreases appreciably the variability of the system. It would otherwise suffer strong breakdowns when reiteratively approaching criticality. This smoother response is not only valid for the average size but also for other features of the system, so the main properties of criticality are found quite far from the critical point, making the system seem "critical". I maintain the denomination "self-organized criticality", though in view of these results it would perhaps be more exact to use some other expression, such as "self-organized paracriticality".

3.4.4. Was there a phase transition in Yellowstone?

The results from last section suggest that the response of forest fires to environmental forcing empirically found for Alberta (Fig. 3.3D-F) is characteristic of a kind of systems displaying a second order phase transition as forcing reaches a given critical value. Is it possible to find some example of such a phase transition in the real world? The Alberta data used for Fig. 3.3D-F displays no clear departure from an exponential response of $\bar{s}$ to FWI, as shown in Fig. 3.5 (at least in its overall trend: I ignore if the apparent "waves" are significant; I do not introduce error bars into Fig. 3.5 because in this case they would be meaningless, for the reasons explained in App. 2A). This is no evidence for a lack of a phase transition, since the data are quite noisy, d is expected to be small, the system is expected to remain ordinarily far from criticality, and fire sizes will be constrained by outer factors. What it indicates is that the answer should be searched somewhere else.

Yellowstone 1988 fires (Romme and Despain 1989 a, b) constitute one of the best documented instances of fires extreme enough to be perceived as "abnormal" by many observers, and have not ceased to be surrounded by polemics (Stone 1998, Wattenburg 1998, Anderson et al. 1999). From 1872 to 1972, there was a policy of complete fire suppression, though it is often assumed that this only became efficient after the World War II (Romme and Despain 1989 a, b).

.

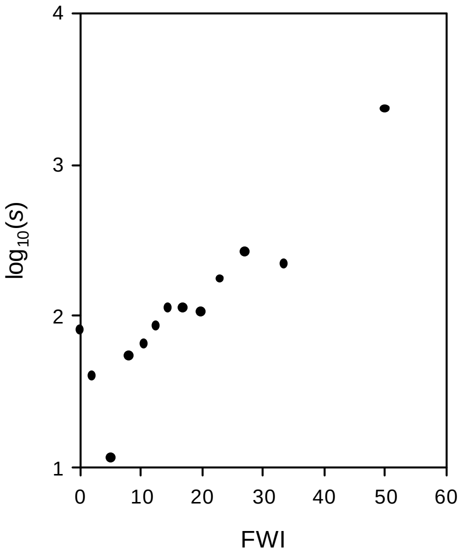


Fig. 3.5. *Effect of FWI on mean fire size (ha) for the same set of data used for Fig. 3.3C-E.*

In 1972, a prescribed natural fire program was implemented. Fires were allowed to burn whenever they were started by lighting and did not threaten some specific areas, while fighting any other fire. The program was considered to be successful until 1988. That year, the summer fire weather was particularly severe. By July 21, a large area was burning and the decision was taken to fight all of the fires. However, they soon became too extreme to be manageable, despite a suppression effort unique in history: it was too late. Many fires kept burning until they were finally extinguished by autumn snows. There has been a long discussion on whether these fires were the result of the increased accumulation of fuel during the previous suppression period. Here I do not quantify the weight that previous suppression might have had on 1988 fires, though I clarify the mechanism for this possible impact in subsects. 3.4.5 and 3.5.2 below and hence set the basis for quantifying it. In this section, I use Yellowstone data just to analyze the response of fire to weather from 1972 to 1988. Fire variations within this period should mainly be attributed to weather (only the late suppression efforts in 1988 must have produced some decoupling between weather and fire), regardless of the possible differences between this and other periods attributable to other factors.

The fire data were obtained from the December 1999 U. S. Dept. of Interior Wildland Fire Management Data CD-ROM. One problem with this data basis is that, in some cases, different registers correspond to a single fire at different stages of the fire suppression process and this is not indicated. I could edit 1988 data with the help of historical accounts (specially Barker 1996). In the few cases in which two fires converged into a single one, the area burnt after the fusion was distributed between both fires proportionally to their previous sizes. I did not have the information to edit the data from 1972 to 1987, so I only use the edited version of the 1988 set where I do not have to compare it with other years. The analysis of 1988 data suggests that this problem does not have much effect on the results.

I use the atmospheric index obtained by Balling et al. (1992) (Fig. 3 in their paper) as an approximation to r_e. This index is calculated on a yearly scale and is a linear combination of temperature and precipitation during the fire season and during the antecedent season. Here I call it Index of Balling, Meyer and Wells (IBMW).

In Fig. 3.6A I represent IBMW vs. $\log(\bar{s})$ for each year or set of years (before this step, I sorted the years according to IBMW and merged the years with consecutive IBMW in the cases in which they had less than 20 fires). The figure suggests a higher than exponential response of $\bar{s}$ to IBMW. This deviation would be even stronger, had not there been the tremendous (but late) suppression effort in 1988. However, it does not have much statistical significance. Furthermore, there is enough noise for $\bar{s}$ to decrease with IBMW at some points, which does not allow to use a representation like the one in Fig. 3.4B.

A much more suggestive result is found when analyzing the dependence of b on IBMW, like done for the case of Alberta in Fig. 3.3E. Fig. 3.6B shows that the fires approach $b \approx 1$ as IBMW increases. This approximation does not seem to be just an asymptotic convergence: instead, Fig. 3.6C suggests that it is faster than exponential. The p. d. f. of 1988 fires is shown in Fig. 3.7 A, and Fig. 3.7B displays the whole set of fires from 1972 to 1988. I analyzed more carefully the data from

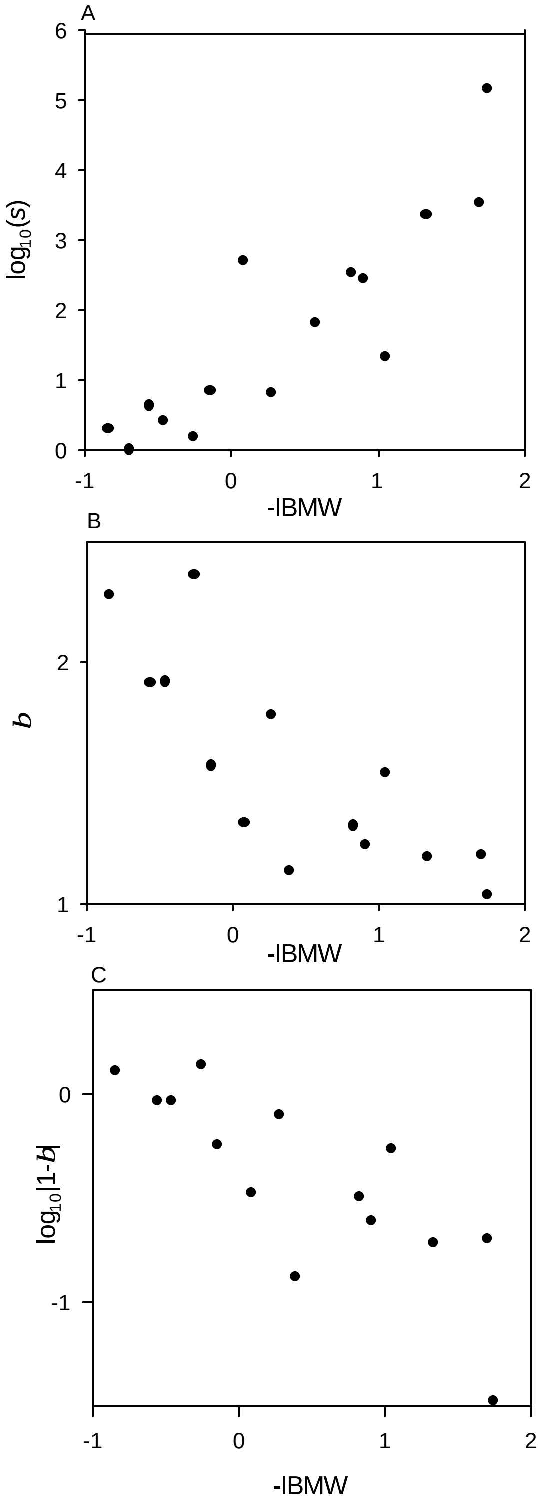


Fig. 3.6. *Yellowstone fires during the period of the prescribed natural fire policy, from 1972 to 1988 (unedited data, see text).* **(A)** *Mean fire size (acres) as a function of an environmental index, IBMW, for all of the 528 fires registered distributed per years.* **(B)** *Estimated b, as a function of IBMW, for the same data, distributed per years.* **(C)** *A different representation of the results in (B).*

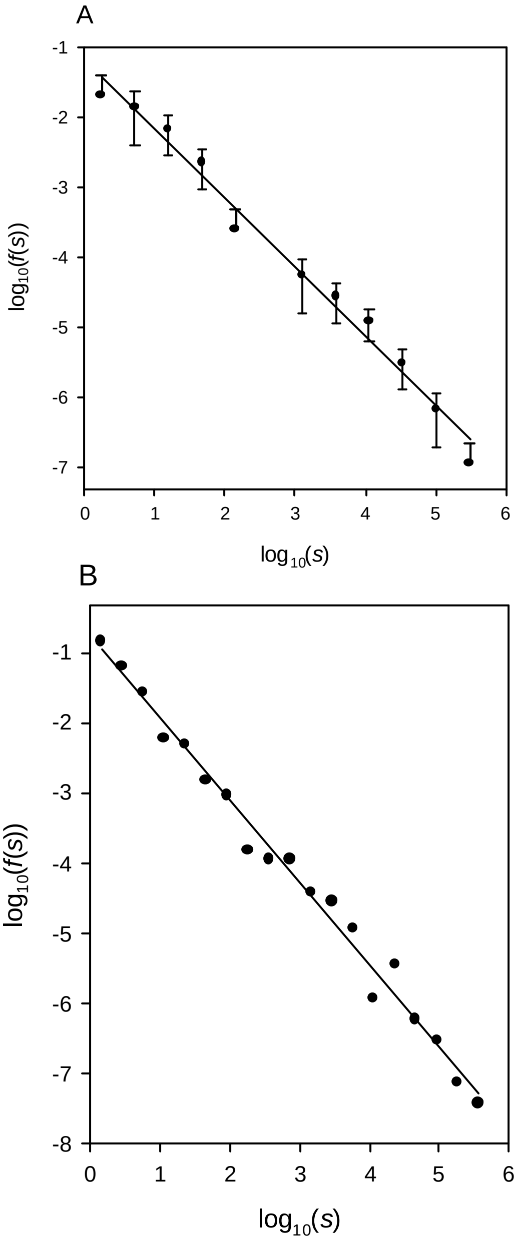


Fig. 3.7. *Empirical p. d. f. for fire sizes (acres) not smaller than 1 acre in Yellowstone.* **(A)** *year 1988, 25 fires (edited data, see text). I include error bars because the number of data per spot is very small as compared to most other figures in this chapter. The estimated parameters are b=0.98 and s_{max}=372199 acres.* **(B)** *From 1972 to 1988, 102 fires (unedited data, see text).*

1988. The few largest fires had their origin out of the park and went out with the change of weather at the end of the season. Both factors may distort the results when trying to extract inferences about fire dynamics from the fire size distribution. However, if we ignore these fires we still find that the outcome of maximum likelihood estimation is $b = 0.97$, with a 90% confidence interval (0.94, 1.08). Without eliminating the most problematic fires, the figure moves slightly downwards. Such b are smaller than any other b in the literature or in my own data analyses.

There are reasons to think that these values of b indicates a phase transition. We just require two assumptions, which seem to be backed by the previous developments: (i) b decreases and s_{max} increases monotonously with r (whenmodifying r_e, which in this case will roughly correspond to IBMW), and (ii) there is a critical point r_c' for which the system would be completely scaling if there were no finite size effects. This means that the bending function in Eq. 3.2 will vanish at criticality, and the distribution of fire sizes will become a pure power law, except for the effect of outer constraints. An unbounded power law distribution with $b \leq 1$ is an improper distribution (see App. 2A), and improper distributions are only found at supercriticality (see subsect. 3.2.1). Hence, if there is a critical point r_c', this will have b>1. From the assumption of a monotonous relation between r and b, b>1 also for $r < r_c'$. Therefore, for $r \leq r_c'$, we can only find $b \to 1$ if $r \to r_c'$.

These findings suggest that Yellowstone 1988 fires took place either beyond or very close to the critical point in which these become able to sustain themselves indefinitely, while allowed by contour conditions. The result was several months of uninterrupted burning, until such contour conditions materialized in the form of autumn snows.

3.4.5. Long-Term Response to Changes in Environment and Ignition Frequency

The above results allow to predict short-term responses to environmental changes affecting fire propagation. Short-term responses to changes in ignition frequency f are trivially linear. In both cases, long-term responses are quite different.

Consider a forest that has developed under some fixed f and r_e. Set a new f or a new r_e and conserve it for the remainder of the time. After the initial sudden modification in fire regime, this will undergo a gradual change until a new steady-state regime is reached.

It is well known that the total rate of burning in the standard SOC forest fire model has little dependence on f, so $\bar{s}$ is almost inverse to f (Drossel and Schwabl 1992, Clar et al. 1996). This also holds for the model in this paper. Malamud et al. (1998) call this phenomenon “Yellowstone effect”, by assuming that this explains 1988 events, but the results of subsects. 3.4.4 and 3.5.2 indicate that this is not the case. However, there are several instances in the literature of a seeming decrease of fire size with fire frequency, and some of them were put forward before the discovery of SOC dynamics. Minnich (1983) sustained that fires were smaller in Northern Baja California than Southern California as a consequence of a higher number of fires in the first case (see also Minnich and Chou-Yue-Hong 1997). The paper by Keeley et al. (1999), which is largely a criticism of Minnich (1983), offers however evidence of increases in fire frequency accompanied by decreases in fire sizes in several sites in

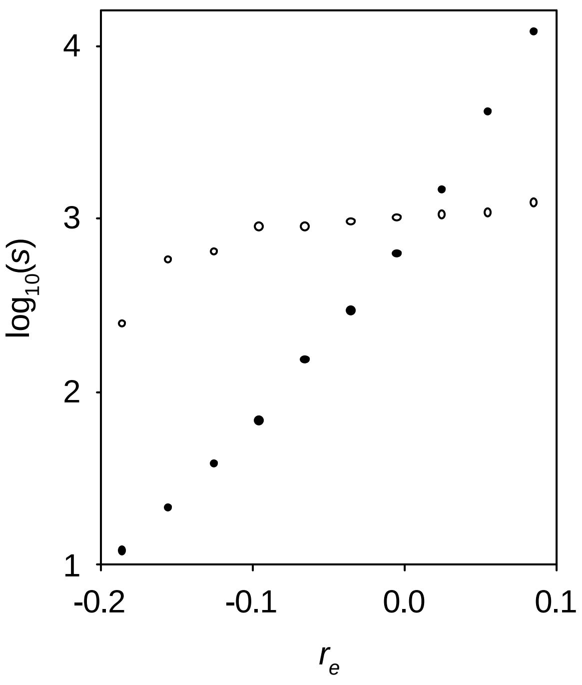


Fig. 3.8. *Average simulated fire size as a function of short-term (solid circles) and long-term (empty circles) environmental variations. Note the logarithmic scale of the sizes.*

Southern California, across the 20th century. Of course this could also be attributed to other human influences. Another instance is provided by the reconstruction of fire history in a Swedish boreal landscape from 1350 to 1850, recently performed by Niklasson and Granström (2000). Fire sizes gradually decreased following the increase in frequency after the expansion of non-Saami settlements.

Fig. 3.8 allows to compare the short term and long term effects of modifying r_e in our model forest. Model forests reorganize themselves in the long term in such a way that compensates for most of the change in the rate of burning that variations in r_e produce in the short term, whenever $r_\infty + r_e$ remains large as compared to r_c.

When $r_\infty + r_e < r_c$, the forest abandons its SOC dynamics and becomes close to a simple percolation model. Fig. 3.9A depicts this transition. While $r_\infty + r_e$ is appreciably smaller than r_c, fires are negligible. There is a sharp change in a small interval around $r_\infty + r_e = r_c$, and a nearly linear relation between r_e and $\bar{s}$ when $r_\infty + r_e$ is appreciably larger than r_c. As apparent from Fig. 3.9B, the change in r_i is roughly the specular image of the change in $\bar{s}$. Although r does not have to be proportional to the amount of biomass, Fig. 3.9B suggests that the transition around $r_\infty + r_e = r_c$ also involves a large transition in terms of biomass. In any case, if this model holds for forests in the real world, this sharp change in burning will be associate to a sharp change in many other of their characteristics.

The above results are just useful as a first approximation, because r_e is always subject to strong short term fluctuations, so we should consider long term changes in the distribution of r_e, rather than long term changes in r_e taken as a single value. One first result that takes this point into account is the following. Let us call $t(r_\infty)$ the time span that a burnt stand requires to reach $r_i = r_\infty$. If a model forest has not been subject to any $r_e > r_c - r_\infty$ for a period equal or longer than $t(r_\infty)$, its dynamics will be close to a simple percolation model instead of SOC.

It is known from the standard SOC model that, when taking finite size effects into account, SOC dynamics is abandoned when the system size decreases under a certain value given f, or f decreases under a certain value given system size (Schenk et al. 2000). This is also the case for the model in this paper, but the variable that controls the transition is not exactly f: it is the frequency of ignitions that take place under conditions satisfying $r_e > r_c - r_\infty$.

Modifications in f and r_e have also some effects that the simulations in Figs. do not incorporate. In particular, these simulations are not reliable in front of large and abrupt increases in either of both variables. The extreme rate of burning that will follow on the short term may produce irreversible changes in species composition or in soil properties, which should be introduced into the model as modifications in the parameters of the curve of growth of r_i.

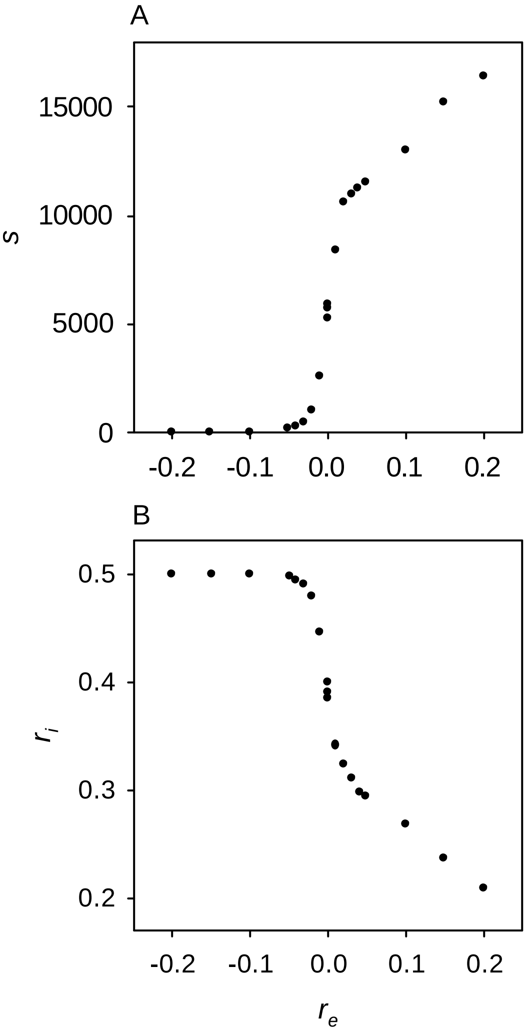


Fig. 3.9. *Transition in fire regime. Steady-state values for* (**A**) *average fire area* $\bar{s}$ *and* (**B**) *the fuel parameter* $\bar{r}_i$, *as a function of the environmental parameter* r_e, *when* r_e *is fixed, in a simulated forest.*

3.4.6. More complex models

In spite of the amazing similitude between model and real-world data, the above model is still a huge simplification. There are many factors with influence of fire dynamics that have still not been considered. Those inherently anthropogenic are treated in sect. 3.5. Here I deal with some of the most relevant between the other factors.

Topography is indeed essential when predicting the evolution of single fires (e. g. Viegas 1998). I tried whether it also affects the kind of p. d. f. shape or the parameter b. I generated fractal landscapes with realistic parameters by inverse Fourier transform (Voss 1988). Each cell (i,j) in the lattice was assigned an altitude z_{ij}. Then the probability of transmission from $(i+\Delta i,j+\Delta j)$ to (i,j) was made dependent on a factor $r_e \propto z_{ij} - z_{(i+\Delta i)(j+\Delta j)}$ (in accordance with Eq. 3.4), where $\Delta i, \Delta j = \pm 1$. No significant effect was found.

Species composition affects the evolution of r_i in each stand between fires, and hence determines fire rotation interval. The lower is r_∞ and the rate of increase of r_i, the longer will be the fire rotation interval.

Beyond a given point, r_i might decrease as stands get more mature (Aber et al. 2000, Bosch 1994). The forest will still display a SOC dynamics while $r_\infty > r_c$ (this inequality is of course modified when taking r_e into account). For a given curve of growth up to the point in which the peak value of r_i is reached, the fire rotation interval will be longer if there is a posterior decrease in r_i than if this remains constant.

Another point to consider is that fire registers often contain a mixture of different kinds of fires, from ground fires to crown fires (Viegas 1998), and also a mixture of different kinds of ecosystems. It might seem strange that these mixtures behave in the same way as models without any such heterogeneity. However, if we mix several sets of data such that each adjusts to a power law without large differences in b, the result is again a power law, as shown in subsect. 5.4.3. Their b will change in a common direction as a function of environmental parameters, so it is not surprising that the mixture displays a response similar to that of one single kind of fire and ecosystem. The changes in the overall s_{max} will mainly reflect changes in the kind of fire and ecosystem with the largest s_{max}.

3.5. Anthropogenic factors

3.5.1. Generalities

Nowadays, fire regimes cannot be decoupled from human activities. Some of the anthropogenic factors can be directly introduced to the previous formalisms, between the many factors that have an effect on f, r_e or other parameters. Others require a qualitatively different treatment, which is offered in the following subsections.

Humans indeed modify f, and are in many regions the main source of ignitions. The conclusions from subsect. 3.4.5 should as a principle be valid regardless of the origin of ignitions, except for some particular circumstances. Specifically, the location of some prescribed fires is carefully chosen to generate discontinuities in fuel, so their consequences can be better grasped from the below developments about firebreaks than the above developments about random ignitions. r_e is modified as a result of anthropogenic impacts on global and regional climate and, in some cases, by the change of water-table levels due to water consumption. Some alterations of the state of the forest, including the modification of species composition and several forms of degradation, exploitation and management will essentially act in such a way to modify the parameters of the curve of growth of r_i.

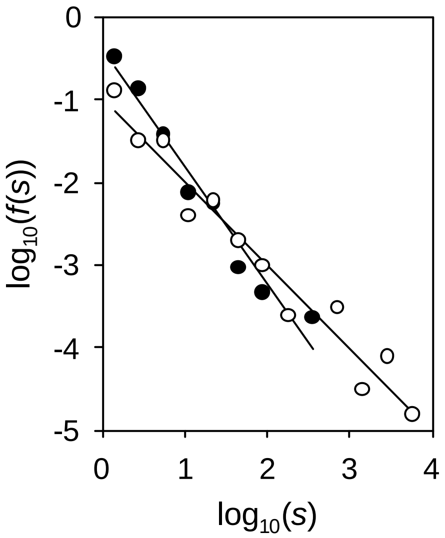
Other anthropogenic factors are treated in detail in each of the below sections.

3.5.2. Fire Suppression

The effectiveness of fire suppression is a decreasing function of fire size. Hence the effect of fire suppression lies somewhere between the effect of a decrease in f and the effect of a homogeneous decrease in r_e. The larger is the elasticity of firefighting intensity to fire size, the closer to the last case we will be.

The Yellowstone 1972-1987 fire register offers a window to the short-term effect of fire suppression. During this period, fires were allowed to burn whenever they were started by lightning and did not threaten specific areas. The rest of the fires were actively fought (Romme and Despain 1989 a, b). If we take the p. d. f. of fought

Fig. 3.10. *The short-term effect of fire suppression. Empirical p. d. f. for the size (acres) of fires that burnt at least 1 acre in Yellowstone from 1972 to 1987: fires that were fought (solid circles; 34 data) vs. fires that were allowed to burn under the prescribed natural fire policy (empty circles; 32 data).*



and unfought fires (Fig. 3.10) and compare their parameters by randomization, we find a significantly higher b in the first case (significance level $e<10^{-4}$). This suggests that fire suppression has a significant effect on the overall r_e in the case of Yellowstone (though there could be confounding factors).

When fire suppression can be assimilated to a decrease in f, in the short term it will only affect the number of fires above a given small size, but not their statistical distribution. If the long-term forest response is as suggested by the SOC model, the original fire rotation interval will be almost recovered, while s_{max} will increase above its original level. An increase in fire sizes as a consequence of fire suppression has been feared by several authors, such as Minnich (1983), and remains a controversial issue.

When suppression can be assimilated to a homogeneous decrease in r_e, all of the fire size distribution will be strongly affected in the short term, in the way explained in the subsect. 3.4.2. Fig. 3.10 seems to display a similar effect. Also in this case the long term response of the forest will offset most of the effect of suppression. However, in contrast to the above case, s_{max} will stay below its original value. On the other hand, fire intensity may increase due to increases in $\overline{r_i}$, and produce detrimental ecological effects. If the suppression-driven decrease in r_e is so strong that $r_e + r_\infty < r_c$, the response will be completely different: fire will be virtually eradicated, as seems to have happened in some real cases (Niklasson and Granström 2000).

An intermediate model can be obtained by letting r_e be gradually modified as a response to the evolution of each particular fire. On the one hand, r_e will change across time t as an increasing function of the length ℓ of the flaming front, since a finite amount of suppression resources has to be distributed throughout the fireline. On the other, the amount of mobilized resources will probably increase with fire size, but less than proportionally. This can be captured in an equation of the form

$$\Delta r_t \propto -\frac{s_t^l}{\ell_t},$$

where s_t is the burnt area at time t and $l \in (0,1)$. For some range around $l \approx 0.5$, the above fraction is roughly adimensional and the simulations show that it does not disrupt the scaling character of fire size distributions. Otherwise, the power law is distorted to some extent. If the power law is conserved, it displays an increase in b at steady-state, so fire suppression may well play a role in the observed values of b. s_{max} may either increase or decrease as compared to its original value.

Under all of the approaches, fire suppression can be expected to suffer a major loss of effectiveness across time, unless it is just applied to offset increases in f or r_e. In terms of rotation interval, the loss of effectiveness will only be partial. In terms of s_{max}, the forest response may or may not overcompensate suppression, depending on suppression intensity and its elasticity to fire size. Suppression may be counterproductive in terms of fire intensity. It may also have more direct ecological impacts. The loss of effectiveness is only avoided if suppression is strong enough for the forest to lose its SOC dynamics and acquire the features of a system in the subcritical state of a simple percolation model.

Perhaps the main problem with fire suppression is that it produces dependence. After a long period of suppression, its interruption would be equivalent to a sudden increase in f and/or r_e and would produce an amount of burning much above the original levels. The most dramatic outcome would be expected where complete eradication has been reached. In this case, there would be a transition to supercriticality.

The sudden change in fire policy in Yellowstone in 1972 was equivalent to an increase in r_e, which lasted up to 1988. This added to the variations in r_e due to atmospheric fluctuations, which have been analyzed in subsect. 3.4.4. The magnitude of the change in r_e due to fire policy depends on the extent to which the forest had approached the steady-state that corresponds to the previous suppression intensity. There is no unanimity on the impact that this had on the amount of fuel (Romme and Despain 1989 a, b), so more studies are required to determine whether or not the phase transition that seems to have taken place (or nearly so; see subsect. 3.4.4) in 1988 was anthropogenic.

3.5.3. Firebreaks

Firebreaks may have different functions. They may wrap particular sites such as e. g. buildings to protect them. In this case, they can be expected to perform their function without interfering with the overall fire dynamics. They may separate forests from sources of ignition such as high-tension cables or roads, and in this case their effect will consist of keeping f low. In some cases, the main effect of firebreaks will be to improve the access for fire suppression, and then they should be considered in the context of suppression as treated in the above section.

The least trivial case is that in which firebreaks act as fuel discontinuities within the forest. At first sight, they could be thought to be comparable to the areas with low r_i that emerge in SOC model forests as a consequence of their own dynamics and separate areas with higher r_i, thus limiting the propagation of fire. However, they are quite different. The first difference is that they are linear, which increases their effectiveness per unit surface of firebreak. Furthermore, they are often located strategically to further increase their effectiveness. The result is that they produce a lower rate of burning for a given $\bar{r}$. On the contrary, the correlations that emerge in SOC forests produce a higher rate of burning for a given $\bar{r}$, which makes the response of $\bar{s}$ to $\bar{r}$ smoother, as shown in subsect. 3.4.3. Therefore, the consequences of firebreaks will be in fact similar to those of fire suppression. Their effectiveness will decrease with time. It will completely disappear unless their location is modified to track the evolution of the forest or, alternatively, they divide the forest in small portions. They will generate dependence, since their suppression will produce an increase in $\bar{s}$ above the levels previous to the opening of the firebreak. This last effect will be stronger the more effective are the firebreaks.

The case in which firebreaks form closed loops is specially complex.

If they are completely impermeable to fire, they will set a constraint on s_{max}. If this is not smaller than the original s_{max}, fires will recover their original size distribution. The only long-term effect of firebreaks will then be to predefine the particular boundaries of the major fires, but not their extension.

If the fragments are smaller than the original s_{max}, simulations display an increase in the frequency of the largest fires compatible with the constraints, in such a way that fire rotation interval remains unchanged, though of course s_{max} decreases. As fragments get smaller, there is a transition from SOC dynamics to their quasiperiodical complete burning, as known from the standard SOC forest fire model (Schenk et al. 2000). Below certain size, fire rotation interval will increase because it will be constrained by the frequency of ignitions within the fragment, but fire intensity may then increase.

If a forest is divided by effective firebreaks into a set of fragments that are mostly of small size, the overall fire size distribution will indeed be strongly dependent on the distribution of the sizes of the fragments. When firebreaks are not completely impermeable to fire, fragments may accumulate enough fuel to ensure that this barrier is surpassed. In this case, the forest will recover an extended SOC dynamics, but rotation interval may increase. I have explored all of these effects with the appropriate simulations.

3.5.4. Anthropogenic mosaics. The case of the Mediterranean

The traditional landscapes in the Mediterranean consist of forest-agropastoral mosaics (e. g. Pyne 1997, Folch and Camarassa 1993). The structure of these mosaics is likely to directly determine the distribution of fire sizes, like the natural mosaic produced by a SOC dynamics. Fire will also have an effect on the evolution of the landscape, but it might not play such a strong role as it seems to play in less human-transformed terrestrial ecosystems.

The island of Minorca offers an instance of specially well conserved traditional landscape. Chust et al. (1999) found out that it had a fractal structure. A mosaic produced by socioecological processes other than fire will have more impact on fire regime than forest self-organizing processes when its fractal dimension is smaller than that expected from SOC. In this case, the fire size distribution will still be scaling, though with a higher b. However, some other properties of SOC forests will not hold, such as the long-term relationship between f and mean size that has been observed elsewhere. As apparent from Fig. 3.11, Minorca fires display a scaling

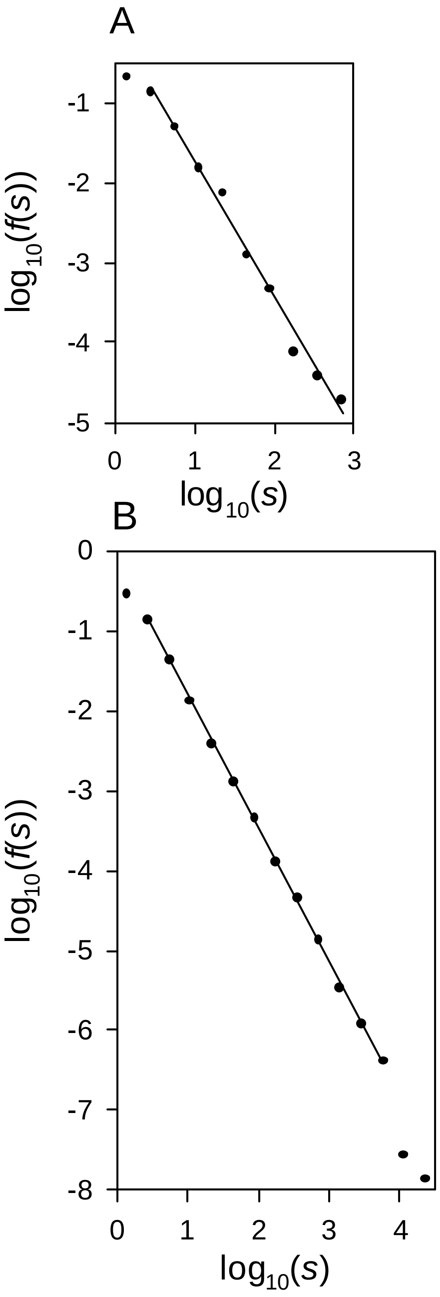


Fig. 3.11. *Fire in the Mediterranean.* (**A**) *Empirical p. d. f. for the sizes (ha) of the 102 fires not smaller than 1 ha in Minorca from 1970 to 2002.* (**B**) *Empirical p. d. f. for the sizes (ha) of the 8974 fires not smaller than 1 ha in Catalonia from 1966 to 1999.*

distribution for at least 2 orders of magnitude, with $b \approx 1.69$, which is large as compared to other kinds of wildlands.

In some Mediterranean areas, forests are expanding after a fast shrink of the area devoted to agriculture and cattle raising, which had previously undergone a fast expansion. Catalonia is one of the clearest instances. In these case, forest fires can be though to be in transition to a SOC state. If there were a direct transition from the traditional mosaic to the SOC state, this could be expected to be quite smooth, but the homogeneity introduced by the previous agricultural expansion could produce an intermediate period of highly increased burning, unless compensated by some asynchrony in the abandonment of fields. In fact, things are even much more complex, because there have been many other dramatic changes: the intensity of fuel extraction from forests fluctuated similarly to the extension of agropastoral soil, the anthropogenic sources of ignition increased really fast, and there has been a huge increase in fire suppression and firebreak opening. Furthermore, the current species composition is quite different than expected by many authors if forests become more mature. In addition, there has been a trend to increasing drought that can be expected to follow up due to global climate change.

Despite all of these factors, the p. d. f. of fires in Catalonia looks much the same as that in Minorca. It displays a power law with $b \approx 1.67$, which is not significantly different. The largest fires in Catalonia largely exceed those in Minorca, but there is not enough data to establish with statistical rigor the upper bound to Minorca's fires. It is worth mentioning that, even though the transformation of the landscape has not been as deep in Minorca as in Catalonia, the analysis of the time series of fires reveals that the fire regime is also far from stationary in the first case.

Ricotta et al. (2001) analyzed the distribution of wildland fires in several sites in the Mediterranean and found them to roughly agree with power law distributions, often with large b.

3.6. Disaster in the tropical rainforests?

One of the results of subsect. 3.4.5 is that model forests do not have a SOC dynamics if r_e has not surpassed $r_c - r_\infty$ for a given period of time. In this case, they will work like a simple percolation model. This model has two main features. On the one hand, fires will be negligible while the environmental conditions in which the forest has developed do not change. On the other, if r_e is modified the response in terms of burning intensity will be much sharper (much larger d in Eq. 3.7).

The qualitative information in the literature gives reasons to think that tropical rainforests do not display a SOC dynamics in terms of fires and are closer to this second type of behavior. On the one hand, it is usually assumed that fire does not propagate in this kind of forest in ordinary conditions (Uhl 1998), thanks to their self-regulated microclimate (indeed, world maps of burning intensity based on satellite information display a sharp contrast between tropical rainforest areas and the surrounding fire-prone biomes such as savannahs). On the other hand, tropical rainforests have suffered colossal fires when subject to certain exceptional conditions.

The view of tropical rainforests as immune to fire changed with the extreme fires in Borneo in 1982-83 (Uhl 1998). These were followed by an even stronger fire event in the same region in 1997-98, which is considered to be "the largest fire disaster ever observed" (Siegert et al. 2001), in striking contrast with the common situation in this kind of ecosystem. Rests of charcoal in the soil also offer evidence of past fire episodes in Amazonia, once in several hundred years (Uhl 1998).

The fire episodes in Amazonia in past centuries have been hypothesized to correspond to exceptional El Niño events (Meggers 1994, Uhl 1998). The recent fires in Borneo have coincided with severe but not exceptional El Niño. It is generally accepted that these were the result of the coincidence between the drought produced by El Niño and the previous disruption of the microclimatic regulation of large tracts of rainforests due to the sparse extraction of timber (Laurence 1998, Brown 1998, Siegert et al. 2001). Large areas in Amazonia are also known to have become drier and more fire-prone due to timber extraction (Nepstad et al. 1999) and have already suffered considerable fires (Cochrane 1998). In the absence of severe El Niño, tropical rainforest burning is concentrated in the most heavily degraded areas.

The frequency of strong fire events in tropical rainforest areas is feared to increase in the near future due to the combination of several factors: (i) the impact of logging on local conditions (Laurence 1998, Nepstad et al. 1999, Siegert et al. 2001, Page et al. 2002), (ii) the impact of deforestation on regional climate (Kleidon and Heimann 2000, Lawton et al. 2001), and (iii) the impact of global climate change. Global climate change is predicted to produce more El Niño-like conditions and more extreme El Niño events (IPCC 2001). El Niño is associated to droughts in most of the vast tropical rainforest areas of the Neotropics and Insulindia-Oceania, and also in a part of the smaller tropical rainforest remnants in Madagascar and India (see map in Holmgren et al. 2001).

The developments in this paper allow to hypothesize that the response of tropical rainforest fire to external factors will have the form of Eq. 3.5 for a wide range of values (beyond some point in the neighborhood of criticality, fire size will be mainly determined by outer constraints), and that the parameter c will be appreciably below the range of values found in the other kinds of forests (which corresponds to a larger d in Eq. 3.7). In other words, the response should be much more abrupt. On the other hand, c is likely to be larger than expected from simple percolation models, because there are other sources of spatial correlations (e. g. Solé and Manrubia 1995) that could perhaps have some relevance for fire, besides the own traces of previous fires. Unfortunately, the available data on tropical rainforest fires is scarce, but it is urgent to find the ways to test these hypotheses and adjust the parameters. The results should be used to set limits to the different kinds of rainforest disturbance, and also for short and medium term forecasting in order to minimize damage in the cases in which it is too late to prevent fire-prone conditions.

If a part of tropical rainforests attains a situation in which r_e surpasses $r_c - r_\infty$ with a given frequency (see subsect. 3.4.5), these will undergo a transition to a SOC dynamics. This represents a deep modification in the features of the ecosystem. Some basic aspects of this modification can be grasped from Fig. 3.9. Although there is no direct correspondence between both variables, the drastic change in $\bar{r}_i$ in Fig. 3.9 could capture the drastic difference in biomass between tropical forests and savannas. This expected long term decrease in carbon storage should be introduced into global climate models, and also the even stronger and fast release of carbon in the short term (about its magnitude, see Page et al. 2002, Schimel and Baker 2002). If it takes place, this phenomenon will reinforce the positive feed-back on climate that is already predicted (Cox et al. 2000) from other effects of drought on tropical rainforests unrelated to fire. Indeed, however, the consequences in situ are even much more worrying than the consequences for climate.

3.7. Discussion and conclusions

3.7.1. On the dynamics of fires and other catastrophes

The huge degree of coincidence between the expectations from the model developed in this paper (based on an earlier model by Drossel and Schwabl 1992) and the results obtained from the analysis of empirical data strongly supports that the model really captures the essentials of fire dynamics. The coincidence goes much beyond the simple fact that fire sizes have a power law distribution, which was the point highlighted in earlier papers by other authors Malamud et al. 1998, Ricotta et al. 1999). One of the next steps is the contrast of these results with empirical data also for other kinds of catastrophes that can be expected to behave like fire. In sect. 4.4 I outline the consequences that this may have for an extremely important issue, which is pest outbreaks in agriculture. Yellowstone provides evidence that there can be catastrophic phase transitions in ecosystems, so it is no absurdity to wonder whether these can be found for processes such as global food supply.

One of the main conclusions is that, unless human interference is really hard, the ecosystems of vegetated lands regulate their own fires. This regulation takes two completely different forms. In the case of tropical rainforests, microclimate regulation keeps conditions in which fire cannot ordinarily propagate. In contrast, wherever wildlands are often enough subject to fire prone enough conditions (both “enough” correspond to well defined values, at least in theory), which seems to be the case for most of the other wildlands, there is an intense fire dynamics but it is strongly controlled by ecosystem self-organization. Fire has a huge importance in both cases. In the first case, if the system is carried to fire-prone conditions, an abrupt onset of extreme fires is expected. This is just what seems to happen in practice, due to anthropogenic disturbances, and may well be the main threat on tropical rainforests. In the second case, the ecosystem has robust succession patterns that are maintained through fire regulation, rather independently from the presence of other kinds of disturbances such as e. g. the direct effects of meteorological events, which makes fire a more important agent.

When fire dynamics is internalized by the system, this acquires the capacity to gradually reorganize itself in such a way to compensate to a large extent for enduring changes in contour conditions (in environmental conditions, rate of ignitions or management practices). At steady state the rate of burning is quite independent from such conditions, unless these reach the threshold in which there ceases to be such an internalized fire dynamics.

Ecosystem self-organization consists of the emergence of some given correlations in the distribution of fuel in space. As a result of these correlations, the system suffers less extreme fires and less variability in the amount of biomass across time. Contrarily to what one may initially expect, these correlations produce a large increase in the rate of burning, as compared with a system with the same amount and kinds of fuel but without spatial heterogeneity. This apparent paradox can be understood if we take into account that we are dealing with a system able to undergo a second order phase transition. This means that fire size increases in an ever-accelerating way as propagation conditions become more favorable, up to a critical point in which fire is able to sustain itself indefinitely and would thus even have some probability of becoming infinite if allowed by contour conditions (which is of course not the case). This pattern of response is qualitatively maintained as the system self-organizes, but the response becomes smoother, as apparent from the fact that fires already have appreciable sizes quite below criticality. The result is that a robust balance between fuel production and burning is reached far from criticality, so it is more difficult that the system approaches this last state. Otherwise, the system would reiteratively visit the neighborhood of criticality and suffer strong fluctuations

The smoother response is not only apparent from the average fire size, but also from many other features, which remain similar to those found at criticality. This is as much as saying that they remain nearly scaling. This is the reason why the dynamics displayed by the model by Drossel and Schwabl (1992)., from which my model is just an extension, is usually put together with other models under the label "self-organized criticality" (SOC). If we accept that wildland fires have a well defined critical state but remain generally quite apart from this state, it may be misleading to call them self-organized "critical". As I put forward in subsect. 3.4.3, they could alternatively be called self-organized "paracritical", since they present characteristic properties of criticality away from criticality. However, I keep the denomination "critical".

Although the self-organized system exhibits a second order phase transition, this presents a qualitative difference as compared to homogeneous systems: the slope of the power law fire size distribution changes when modifying the control parameter. I find this effect both in the model and in empirical fire data.

We can wonder to which extent we can really ensure that this model captures fire dynamics, and that forests have a SOC dynamics. The main proof is the complete agreement in the distribution of fire sizes as a function of environmental conditions, which is far from trivial, as apparent from the above comments. This is however a partial proof. This result seems to be a consequence of a multifractal distribution of fuel amount and fuel features in space. This should be complemented with evidence that this distribution is a consequence of the own fire dynamics. As I mention in subsect. 3.5.4, I do not think that this holds for some highly humanized landscapes, but seems to hold in other cases. For some forests, there is evidence of a relationship between fire hazard and time since last burning (Niklasson and Granström 2000), and the differences in maturity between different sites can be mainly attributed to past fires. Even more definitive, if really proved, are the claims found in the literature of a gradual decrease in fire sizes following increases in the frequency of ignitions (Minnich 1983, Niklasson and Granström 2000). The coincidence between model and reality in the long-term responses to changes in this and other features of the contour conditions should be better studied.

3.7.2. On fire forecasting

The most immediate practical application of the results presented in this chapter is in fire forecasting. As apparent from Fig. 3.3, the complete fire size distribution can be predicted from readily available weather data. This does not even require the theoretical model to be correct. We just have to use the available data in each area to prepare curves like those in Fig. 3.3.

At first view, one problem with this forecasting method is the difficulty to directly estimate s_{max}, as apparent from Fig. 3.3 F. One alternative consists of fitting the curves r_e vs. b (Fig. 3.3 B and E) and r_e vs. $\bar{s}$ (Fig. 3.5), deduce from these a curve for s_{max} and refine it by taking into account the maximum observed s.

On the other hand, one strength of the methodology here described is that it allows to test whether s_{max} is smaller than any given value, by applying Eq. 2A9 from App. 2A (across this chapter, I test in several cases whether s_{max} has to be considered finite, by applying Eq. 2A8 from App. 2A, which is a particular case of Eq. 2A9). I. e., it allows to estimate the worst possible fire disaster in a given area while its overall fire conditions do not change.

One point that should be studied and is likely to easily render straightforward results is the correspondence between the presumable multifractal features of r_i and the response of fire size distribution to r_e, in model and reality. This will allow to implement a fire forecasting system also in cases in which there is no fair historical register of fire sizes or where historical information cannot be extrapolated because of readily changing fire conditions.

It is even more important to test the predictive ability of the model for long term changes, and to study the shape of the transitories following enduring changes in contour conditions. This will allow to predict future fire regimes in front of global climate change, regional or local environmental changes, changes in the frequency of anthropogenic ignitions, and changes in fire management.

Prospective in front of global climate change requires the study of transitories, because it is too fast to consider just long term effects, and too enduring to consider just short term effects. Besides the direct interest of knowing the impact of climate change, this will allow to improve the own climate predictions by using coupled climate models with realistic terms for the changes in the carbon balance of forests and other ecosystems.

The kind of model developed in this chapter may also be useful to predict the response of other kinds of catastrophes to climate change, perhaps including extreme atmospheric events, which are still poorly predicted by current models.

3.7.3. On fire management

There are strong discussions on the best recipes to manage fires, which are founded on diverging views of fire dynamics. These have been feared to lead to a “schism” in fire science, in the terms of Goldammer (1999). Some authors emphasize the importance of fuel succession, while others question it in front of evidence indicating that variations in forest fires have a stronger statistical relation to environmental variations than fuel differences (Bessie and Johnson 1995). These views are often linked to a preference for the use of prescribed fires in the first case and fire suppression in the second.

The developments in this chapter offer a solution to this divergence. In simple percolation models, fires are rather insensitive to environment except at a tiny region at the neighborhood of criticality, as corresponds to Eq. 3.5 when c is large. Self-organization appreciably decreases c and thus extends the sensitivity to environment to a much larger range of situations. Self-organization is driven by fuel succession but its outcome is precisely a high short-term sensitivity to environmental fluctuations. In other words, fuel succession builds the “eyes” to “see” the environment: both are necessary for “vision” to take place.

The model here presented allows to predict the outcome of each management option in detail. Fire management may have different purposes, which will lie somewhere between eliminating fire and recovering natural fire regimes.

One of the main conclusions of this study is that short term responses to fast changes are generally stronger than long term responses to slow changes, so the rate of change in the parameters is often more important than their point value. A transitory period of increased burning after a fast change is not only a problem while it lasts, but may also produce irreversible damages. When it is not possible to avoid a fast increase in the frequency of ignitions or a fast and enduring environmental change that favors fire, fire suppression has a role to play in slowing down the variation in overall fire conditions.

The fast change produced when suddenly halting a long suppression period is no exception. Hence, when it is planned to recover the natural fire regime, the appropriate recipe is a gradual retreat of suppression, not a sudden interruption like in Yellowstone, although I have not investigated the quantitative importance that this had in the shift of this area to the neighborhood of criticality in 1988. This is something to take into account in the Russian Federation, where fire suppression has decreased in large areas due to the precarious economic situation (UN 1998), while global climate change is also producing more fire prone conditions (Stocks et al. 1998).

Even if conveniently slowed down, it is of course recommendable to avoid anthropogenic changes that increase the rate of burning above natural levels. These include some given changes in species composition (Chapin III et al. 2000, Bosch 1995, Terradas 1996), forest structure, water table levels and, indeed, global climate change. These also include increases in the rate of ignitions, but a gradual increase in this rate could be useful in some circumstances, as justified below.

Things are more difficult in the cases in which the natural fire regime is not considered acceptable. The traditional methods of fire management, such as fire suppression and firebreaks, have some effectiveness but also some outstanding drawbacks, when they become the main management tools. The essential point to keep in mind is that the ecosystem can be expected to gradually react by becoming more fire-prone. This has two major consequences to take into account when weighting costs against benefits. First, the “benefits” cannot be evaluated from the immediate effects of these procedures, because their effectiveness will suffer a major decrease with time and, once this has happened, it will not be possible to suddenly stop spending resources in these ineffective interventions without producing extreme fire events. Second, the “costs” should include the risk of such extreme events, because there will be some probability of failure of fire management. The case of the Russian Federation illustrates that this adds additional steps to the propagation of socioeconomic crises (which could have much in common with fire propagation). There is also one third problem to consider, which is the set of ecological side effects unrelated to the fire SOC dynamics: e. g. the fragmentation of a landscape in large pieces protected by certain kinds of firebreaks may disrupt biological connectivity, while its effectiveness in terms of fire may become null in the long term (although this particular case should be better analyzed, this could be the case of the “Prioritary Protection Perimeters” in Catalonia, Peix 1999). The first of the above problems can be avoided if suppression intensity exceeds a given threshold, in which the forest loses any capacity to control its own fire regime. But the price is an exacerbation of the second problem: in case of failure of the management system, wildlands will attain supercriticality. This would represent the generalization to a whole country (if this is the level at which the management system fails) of a situation comparable to Yellowstone 1988 or worse. It is instructive that the difference between the geometry of human-made firebreaks and the geometry of the “firebreaks” that emerge in self-organizing forests is that the former decrease and the last increase the rate of burning that corresponds to a given amount and kind of fuel. In this way, natural forests keep far from criticality (on the other hand, tropical rainforests seem to have “chosen” a “suppression” approach...).

A more robust way to manage fires, when the natural fire regime is not considered acceptable, is the manipulation of the fuel mosaic without changing its essential natural features. The purpose is attaining the kind of landscape that corresponds to a relatively high frequency of fires and, in correspondence, a small maximum size and a small average size. The most evident procedure consists of increasing fire frequency. There are many precedents of an intentional increase in fire frequency, spanning from the ancestral techniques of Australian aboriginals to the prescribed fires recently introduced into Western forestry. However, there are some points to take into account before using this procedure. (i) The spontaneous reorganization of wildlands is slow, so the rate of increase in the number of ignitions should be low. Otherwise, we will have an unacceptable increase in the rate of burning in the short term, which could have irreversible consequences. (ii) For the same reason, the retreat of fire suppression while the new landscape structure is generated should be slow, if there is a previous history of fire suppression. (iii) An increase in ignition frequency produces some increase in the rate of burning also in the long term (although small as compared to the short term effect of a sudden increase), and its consequences should be assessed. (iv) In many areas the current frequency of ignitions has already escalated much above the natural one in a brief time span, without need of additional prescribed fires (Bosch 1995, Keeley et al. 1999). Instead, some fire suppression will be required where this increase is too large or too fast. (v) The increase in ignition frequency to reduce fire sizes through forest self-organization should not be confused with the prescription of fires in strategic sites to generate discontinuities between large masses of fuel, which is just a cheap way to build firebreaks, with the consequent problems.

The generation of the desired mosaic does not have to rely solely on increasing fire frequency. It can be accelerated through directed forestry interventions. Neither its maintenance has necessarily to rely on fire. Other options range from timber extraction in accordance to a natural disturbance – based management to the inclusion of extensive agropastoral areas taking the role of the burnt areas in a natural forest, much like traditional Mediterranean landscapes (subsect. 3.5.4). Each of these options will have deep ecological and socioeconomic implications, highly variable from one case to another, which should be carefully analyzed before implementing any of them. These do not have to be always negative. E. g., in the case of Catalonia, several authors claim that an increase in extensive agropastoral areas (intertwined with forests) would be positive for biodiversity conservation, which is not surprising if we take into account that humans have shaped Mediterranean landscapes for millennia.

Fire suppression and firebreaks have their own roles to play in the mosaic strategy. These include smoothing the transitories between different fire regimes, avoiding a frequency of anthropogenic ignitions above the planned one, keeping particular sites free of fire, and reacting in front of meteorological situations too exceptional to be internalized by the forest. If the level of disturbance required to maintain the desired fire sizes is very large, it could be better to combine the mosaic strategy with a wider use of these other techniques, in a hybrid approach. These will also be usable with less restrictions if the mosaic is intended to be maintained by procedures other than fire, such as timber extraction or agropastoral activities.

There are some obvious variations on the above mosaic approach, which have however too many problems. One is the generation of a mosaic with a more clearcut dissociation between forest patches than in a natural or a traditional landscape (i. e. a kind of combination of the mosaic approach and the firebreak approach). This could produce some gain in fire safety, but at the price of a disruption of biological connectivity that could be disastrous for biodiversity, and probably also other ecological side-effects. Another variation is a homogeneous reduction of fuel load, instead of a heterogeneous one. A drawback is that, if human intervention ceases for any reason, a forest mosaic will undergo a smooth transition toward a self-organized forest, while a more homogeneous landscape will suffer a problematic transition, with large continuous masses of forest attaining in a nearly synchronic way the conditions for fire propagation, after some period of time. Furthermore, this last option could also be less suitable for biodiversity conservation.

There could be many imaginative approaches to fire management. The model presented in this paper is a basic tool to asses their effects in qualitative terms, and also in quantitative terms if enough data are available.

3.7.4. On tropical rainforests

The most alarming (and potentially useful) conclusions of this work concern tropical rainforests. Unfortunately, fire data are scarce for this biome, so these conclusions are still tentative. It is urgent to gather any available data, analyze them in detail, perform any field research necessary to fill data gaps, transform the scientific conclusions into policy recipes and ensure that they are applied.

Across this chapter I have shown the striking coincidences between the expectations from the model here expounded and the results of the empirical analysis of fire data from different areas out of the humid tropics. This same model, however, predicts a different type of dynamics once the parameters surpass a given threshold. The qualitative information in the literature suggests that tropical rainforests have this other dynamics. Since I have still not been able to test this from tropical rainforest data, I rely by the moment on a combination of qualitative (appreciations from the literature) and indirect (test of the model with other kinds of forests) evidence.

It is usually assumed that fire does not propagate in ordinary conditions in tropical rainforests. Nevertheless, the worst fire episode ever registered took place in Borneo, a tropical rainforest area, in 1998. There seems to be a radical shift in fire conditions in several major tropical rainforest areas in this few last years, due to the synergy between timber extraction, deforestation and global climate change.

The model in this paper suggests that the effects of all of these factors can be reduced to variations in one single parameter, r. The explicit predictions with a largest practical relevance are:

- The response of average fire size $\bar{s}$ to r will agree with Eq. 3.5, which I assume to hold for the other kinds of forest as well. This will only cease to work at the point in which $\bar{s}$ becomes limited by outer constraints.
- The parameter c in Eq. 3.5 will be much larger for tropical rainforests than for other kinds of forest (smaller d in Eq. 3.7). This implies a much more abrupt response of $\bar{s}$ to r.
- While for other kinds forests the critical value of r_e in Eq. 3.5 will gradually change under environmental forcing, which decreases the probability of a major disaster, this will not take place in tropical rainforests.
- When certain thresholds in the intensity of the above disturbances are surpassed and maintained, the rainforest will not recover anymore and will become a qualitatively different ecosystem, with a much lower biomass and an active fire regime, as shown in Fig. 3.9 and explained in subsects. 3.4.5 and 3.6.

I interpret the sharp contrast between the usual situation in tropical rainforests and the catastrophic events in these last few years as a consequence of the large value of c (small d). The corollary is that in 1998 and, previously, in 1982, Borneo's rainforests nearly attained, attained or even surpassed a critical point r_c in which a phase transition takes place, in the strict physical sense of the word. The detailed analysis of Yellowstone data in subsect. 3.4.4 shows that this is neither an entelechia nor a simple metaphor.

The above set of anthropogenic disturbances must be pushing toward supercriticality most of the huge tropical rainforest regions in Insulindia-Oceania and in the Neotropics, and a part of the smaller rainforest remnants in Madagascar and India. In these circumstances, it is necessary to put numbers into the parameters of Eq. 3.5 as soon as possible, and to test its validity. The special shape of this equation allows to obtain well defined management targets. Targets for environmental management often have the form of particular values not to attain, such as a minimum instream flow or a maximum concentration of a given pollutant. However, in most cases these values are highly arbitrary, because they are chosen out of gradual response curves. Instead, in Eq. 3.5 almost all of the response to r is concentrated in a tiny range of values in the neighborhood of a particular value r_c (if c is large, as seems to be the case of tropical rainforests). Since the modification of r is the result of several contributions (greenhouse gas emissions, deforestation, timber extraction), there will be a well-defined multidimensional surface, encompassing several kinds of disturbances, which should not be surpassed if we want to avoid a major disaster in some given tropical rainforest region.

The required research will of course have to begin by gathering any fire data, and concomitant data on factors affecting fire. In the absence of the abundant historical data that is found for other biomes, we can take advantage of the fact that the model in this chapter establishes a connection between structure in space and dynamics in time, so we could perhaps perform fair predictions from spatial data alone. As mentioned in subsect. 3.7.2, I have not established the precise correspondence between the multifractal features of the distribution of fuel in space and the fire responses to environmental factors explained in subsects. 3.4.2-3.4.4, but this must be quite immediate. There are several works that can serve as a basis when passing from the model to the real world. Solé and Manrubia (1995) calculated the multifractal spectrum of tree heights for an instance of tropical rainforest. There are abundant satellite data on many parameters at a coarser resolution, covering every tropical rainforest area. There are well known models that allow to predict the effect of different parameters on the local conditions for fire propagation, which is a basis to transform these parameters into values of r. There are works establishing a relation between different factors and the likelihood of fire in the particular case of tropical rainforests. Nepstad et al. (1998) performed a complete map of fire risk for Brazilian Amazonia in 1998, based on the predicted amount of soil water available to plants in each site.

Once we have the parameters for Eq. 3.5 for the main tropical rainforest regions and also the equation to calculate r from the set of relevant factors, we will be able to perform detailed forecasts on the response of fire dynamics in front of each possible situation. This should be introduced into coupled climate models. The future evolution of the global climate is expected to be strongly conditioned by feedbacks involving tropical rainforests, even without taking fire into account. The appropriate consideration of fire in the humid tropics will thus modify global climate forecasts, both because of the immediate release of carbon if rainforests attain supercriticality and because of the stepwise reduction in carbon storage capacity is fire-prone conditions above a given threshold are maintained.

With all of these results in the hand, it will be possible to tentatively predict the future of tropical rainforests under different scenarios of greenhouse gas emission and other factors affecting the global climate, tropical deforestation and timber extraction, and establish the thresholds not to be surpassed.

In the cases in which it is too late for preventing fire-prone conditions, it is also necessary to predict if a tropical rainforest area will approach or surpass criticality, with some margin of time. This will allow to take measures to minimize the number of ignitions, to try to suppress fires when they are still small (though, unfortunately, the best means of fire suppression are located in areas of the world where their usefulness is more doubtful), and to minimize damages where large fires cannot be avoided.

Of course, prevention is much more desirable. Although we still do not know the maximum tolerable level for each kind of disturbance, there is little doubt that we are readily approaching it, so it is not necessary to await the precise numbers to make every pressure to revert these trends. The socioeconomic barriers to surpass are immense, but so will be the consequences if we fail to. The abruptness of fire transitions and the fact that these have already begun to take place are worrying, but also give some hope, because one consequence is that the actions taken will have returns in the short term, which seems to be a requirement for policy makers and economists to take problems seriously.

The keys to the solution lie at a global scale, to a huge extent. Current tropical countries' forest policies are largely an expression of decisions taken by multilateral institutions (for the Indonesian case, see Rowell and Moore (1999), Gautam et al. (2000)). While this situation is maintained, these decisions should take fire constraints into account: otherwise, they will be completely counterproductive. An enhancement of unsuitable logging by international trade agreements and organizations (see e. g. Amilien 1996) should also be prevented.

Industrialized countries should fully assume their responsibility in curbing climate change. When having the numbers, we could well find that current precarious compromises are not enough to avoid a large fraction of the remaining tropical rainforest to end up between flames.

Chapter 4

Diversity, stability and criticality

4.1. Introduction

Previous chapters deal with some essential features of ecosystem fluctuations. There is currently a strong interest in the role that biological diversity may play in such fluctuations (McCann 2000, Naeem et al. 1999). In particular, a recent wave of experiments seems to reveal that diversity decreases variability in aggregate ecosystem features such as biomass, productivity or CO_2 flux. It is extremely urgent to reach a deep understanding of the relation between diversity and stability, in a moment of sky-rocketing loss of both natural and domestic biodiversity. Although there are other strong arguments for conservation, the risk of destabilization of the processes that support human life on Earth may well become the most convincing one.

Biodiversity loss is an irreversible process, as is information loss in general, for the reasons analyzed in detail in the first two chapters. The paleontological register seems to indicate that diversity requires about 10 million years for recovery after extinctions (Kirchner and Weil 2000), so what we lose in the next few years may become a constrain on human development for the remaining time of existence of our species. We are just beginning to understand the nature and importance of this constraint: we should strive to go ahead in this understanding before the constraint becomes inescapable.

There is a huge literature on hypothetical models of the relation between diversity and stability. These models have had a great impact on the perception of this issue, even more than experiments. However, they have clear limitations and some of the most influential ones stand in sharp contradiction between them. What I do in this chapter is dissecting some of the main models in depth, looking at the sources of their contradictions and resolving them, and putting forward some complementary proposals. All of this collapses into a coherent framework. This is still largely hypothetical, at least in some parts, but I think that this is a necessary step to foster further advances integrating theoretical and empirical work.

One of the features of the theoretical (or hypothetical) framework that I obtain is that it contemplates the possibility of critical phenomena in several respects. If some of them is confirmed, it will allow for a considerable leap forward in terms of detailed predictions. In particular, it will have far-reaching applications for food security.

By the middle of the 20th century, diversity was generally thought to have a stabilizing effect. Outstanding ecologists such as Odum (1953), MacArthur (1955), Elton (1958) and Margalef (1966) sustained this tenet with several arguments. Elton (1958) realized the tremendous implications that this would have in terms of conservation.

This view was defied by May (1972, 1973), who analyzed the stability of randomly assembled model ecosystems. In particular, he analyzed asymptotic stability as defined in dynamic system theory, which I will refer to as “dynamic stability”. In this context, a model ecosystem is considered to be stable if it recovers a given set of species abundances after this has been disturbed. May found that, by increasing species number, connectance or interaction strength between species, model ecosystems pass from being almost certainly stable to being almost certainly unstable as soon as a threshold is surpassed.

Despite May’s own precautionary comments, his finding was generally perceived as a last word in the diversity-stability issue. For decades, diversity was assumed to decrease ecosystem stability. Some authors, such as McNaughton (1977, 1978, 1985), offered theoretical arguments and even empirical evidence pointing in the opposite direction, but they had a limited impact. During this period, functional ecological arguments were not given much weight in discussions about conservation, also with outstanding exceptions (Ehrlich & Ehrlich 1981).

Things changed at the mid nineties, when a strong attention was paid to a set of experiments with artificial ecosystems providing evidence that diversity decreases variability in aggregate ecosystem magnitudes. These took place in prairie plots (Tilman and Downing 1996, Tilman 1996, 1999), macro and mesocosms with representative species from terrestrial ecosystems (Naeem and Lee 1997, van der Heijden et al. 1998), and microcosms with representative microbial freshwater species (McGrady-Steed et al. 1997), and later in mesocosms with intertidal invertebrates (Emmerson et al. 2001) (a comprehensive survey of empirical evidence for and against diversity effects upon stability and other ecosystem features can be found in

Schläpfer and Schmid (1999); see also Díaz and Cabido (2001) for a more specific survey about plant functional diversity). Some authors raised technical criticisms against these experiments (see e. g. Huston 1997, Kaiser 2000), but these seem to be largely solved (Hughes and Petchey 2001, Loreau and Hector 2001). However, a new experiment has introduced new doubts (Pfisterer and Schmid 2002, see subsect. 4.2.2).

The upsurge of experiments prompted the development of theoretical models that could explain the new empirical observations. Tilman (1996) solved the apparent contradiction between experimental results and dominant theoretical points of view by recovering a study performed by May (1974a) on models of simple competitive communities, whose impact had been much lower than May's complex ecosystem model. In the case of simple competition, May had found that an increase in species number decreased species-level population stability but not community aggregate stability. On this basis, Tilman proposed that diversity generally increases variability at species level stability and decreases it at community level.

The main arguments for a stabilizing effect of diversity are just different versions of the old admonition not to put all the eggs in the same basket. This was already the case for Odum (1953) and MacArthur (1955) ideas, as put forward by Pimm (1991). More recently, this admonition has been restated in terms of "redundancy" as a condition for "reliability" (Naeem 1997, 1998; in fact the term "reliability" had already been used in the diversity-stability context by Cooke (1971, pg. 506)), "insurance effect" (Yachi and Loreau 1999) or "portfolio effect" (Tilman et al. 1998; this metaphor was simultaneously applied to the role of diversity in agriculture by Brummer 1998).

This has also been put in explicit statistical terms. Doak et al. (1998) attribute the ecosystem effects of biodiversity to one of the most elementary forms of "not-all-eggs-in-one-basket" effect, the "averaging effect": the average of several variables has a lower variance than the single variables, under very general conditions. Although Doak et al. contribution has received a great attention, the expected effect of different features of diversity on different features of stability for simple statistical reasons had already been explicitly put forward by den Boer (1968) and Reddingius and den Boer (1970). Yachi and Loreau (1999) analyzed from a statistical point of view the dependence upon species richness of the aggregate productivity under two different simplifying assumptions: this either corresponds to the average of single-species productivities or to the maximum of these productivities.

Tilman et al. (1998) pointed out that things are more complex than suggested by the simple idea of the averaging effect. This should be complemented by taking into account the dependence upon diversity of single-species population averages and variances and of the covariances between them. Though not stated in these terms, this invited to an integration of the dynamic system and the statistical approaches.

Several contributions have been made to the modeling of purely competitive communities initiated by May (1974a) and recovered by Tilman (1996), now taking into account the existence of statistical “averaging” effects. Different authors have found either increasing (Ives et al. 1999, Lehman & Tilman 2000) or decreasing (Hughes & Roughgarden 2000) stability with species number.

There have also been some developments about diversity and stability in trophic webs (McCann et al. 1998, Neutel et al. 2002). However, the source of the deep discrepancy between the landmark contributions by Odum (1953) and MacArthur (1955) on the one hand and May (1972, 1973) on the other remains to be addressed.

Some potential contributions to the issue of diversity and stability can be extracted from other sources besides the literature explicitly addressed to this issue. This is the case of the paper about the “River Continuum Concept” by Vanotte et al. (1980), which is considered a landmark for river ecology. These authors suggested that rivers or parts of rivers subject to a high variability in the physical environment (in particular, medium-sized streams) would have a high biotic diversity, which would on its turn have a stabilizing effect upon energy flow.

Long before the onset of modern ecology, traditional farmers made a thorough use of biodiversity to stabilize their crops. Despite the spectacular drop of diversity in modern agriculture, this still makes frequent use of marginal plant varieties with genes of resistance to the pests and pathogens that emerge one after the other. Agroecologists claim for a return to a more intensive use of biodiversity. The phytopathological literature has recently offered some advances on the function of biodiversity (e. g. Garrett and Mundt 1999).

Sect. 4.2 in this chapter deals with the “not-all-eggs-in-one-basket” principle as a large number effect. I propose a modification on Doak et al. (1998) statistical argument to expect a decrease in aggregate biomass variability with increasing diversity. The main criticisms against Doak et al.'s argument do not hold anymore against this modified one. I show that the purported destabilizing effect of diversity in some models (Hughes & Roughgarden 2000) stands on clearly unrealistic assumptions, and that neither the empirical results by Pfisterer and Schmid (2002) contradict the expectations about a positive effect of diversity on stability. I also indicate that there are reasons to expect decreasing variability with diversity in much more general models than those of simple competitive communities, but only at the limit of many species. Few-species systems cannot be expected to display any definite relation between diversity and variability, but should generally be more fragile than many-species systems, in the sense that a change of species composition can produce a qualitative change in the aggregate dynamics. In App. 4A, which is complementary to this section, I reanalyze the data from the experiment by McGrady-Steed et al. (1997) and show that they agree with the above predictions, contrary to what their paper may suggest at first view.

In sect. 4.3 I note that, besides the short-term effect of diversity on biomass variability, there is a long-term need of diversity for adaptability. This is the most evident and the most important effect of diversity on stability. It is necessary for the persistence of complex systems as complex systems despite their changes in the details, which is the main kind of stability. I propose that diversity does not only provide the raw material for adaptation, but may also enhance it by pushing ecosystems to the proximity of a critical state, which allows them to react in front of external fluctuations. The assumption of criticality is the basis for the developments in chapter 5, where I show that it offers a coherent explanation for much of our current knowledge of diversity patterns and for the results I obtain from my own analyses of diversity data, and I give corollaries for diversity data treatment and interpretation and for management.

Sect. 4.4 deals with some special features of webs with a complex set of connections, which have to do with the propagation of fluctuations and with species-level stability. As a principle, this does not affect the global stability of biomass and related variables, which is discussed in the previous sections. I dissect in depth May's criterion (May 1973, 1974) and the ideas by Odum (1953) and MacArthur (1954), show the reasons why they contradict each other, discuss the credibility of each one (with the help of some Montecarlo experiment), and reach a synthesis between both. I also discuss the possibility of a self-organized critical extinction dynamics and the consequences that this would have.

In sect. 4.5 I discuss the relation between diversity the propagation of pests and pathogens, paying special attention to agriculture. I put forward and develop the hypothesis of a self-organized critical dynamics. Although the proposals in this section are highly speculative, I carry them to their last consequence, which is a tentative recipe to reorganize agricultural production if these were found to hold.

I finally offer a synthesis of my tentative conclusions.

Only in sect. 4.5 there is an explicit consideration of space. In this respect, other parts of this chapter (at least sect. 4.4) should be considered less advanced than the rest of the book.

4.2. Large number effects in the relation between diversity and stability

4.2.1. Precedents

Here I summarize some previous contributions that make up the context for the developments in this section. They all refer to the variability in the aggregate biomass of communities as a function of the number of species n.

Variability is usually expressed as the coefficient of variation (CV) of the aggregate community biomass across time (CV is 100 times the relation between standard deviation and average).

Doak et al. (1998) asserted that CV should decrease with n for the simple statistical fact that the variance of the average of several variables decreases with the number of variables (they essentially rediscovered the forgotten findings by den Boer (1968) and Reddingius and den Boer (1970), and introduced them into the modern discussion on diversity and stability). As stressed by these authors, this “averaging effect” is affected quantitatively but not qualitatively by the covariances between variables (it vanishes at the limit of covariance 1, but in no case it is reversed). However, as a principle it could be affected both quantitatively and qualitatively by the possible dependence on n of the distributions of population fluctuations at species level (Tilman et al. 1998).

Tilman et al. (1998) thus replied to Doak et al. (1998) that the averaging effect does not make a positive relation between diversity and stability statistically inevitable. In fact, both modeling and empirical data indicate that single species population dynamics changes with n and, more specifically, displays an increase in CV (Tilman 1996, Tilman et al. 1998). Lehman and Tilman (2000) studied by simulation specific models of simple competitive communities. They found CV to decrease with n in all cases, but the reason for this result was not obvious. Single species CV increased with n, while summed variances and summed covariances displayed a nonmonotonous response to n and the overall biomass increased.

Several authors have studied analytically the stability of models of simple competitive communities. They use Lotka-Volterra models with quadratic competition terms, either in continuous or discrete time. In discrete time and in a deterministic approach we have

$$B_{i,t+1} = B_{i,t} + r_i B_{i,t} \frac{K_i - B_{i,t} - \sum_{j=1(j \neq i)}^{n} h_{ij} B_{j,t}}{K_i} \qquad (4.1).$$

Hughes & Roughgarden (2000) analyzed this model, and Ives et al. (1999) analyzed a roughly equivalent one. Both teams considered the simple case of homogeneous parameters $r_i=r \ \forall i$, $K_i=K \ \forall i$ and $h_{ij}=h \ \forall \ i, j$. Once they had characterized its dynamic stability properties, they passed to introduce stochastic terms.

The analysis of dynamic stability is achieved by eigenvalue decomposition, which is the same technique already used by May (1974a) for a similar model. Each possible steady state of a dynamic system is characterized by some set of eigenvalues $\{l\}$. Of course we are interested in the particular steady state in which all of the species have a non-null biomass.

Take a situation in which the system has been slightly separated from its steady state (in this case “slightly” just means that no species has gone extinct, but for other systems this is more restrictive). The system tends to recover the steady state if and only if $|l_{max}|<1$ (in the continuous case, $l_{max}<0$), where l_{max} is the maximum eigenvalue. In general terms, the strength of fluctuation dampening at the single-species level is highly influenced by l_{max}. In this particular case, where all of the species have the same parameters, it is completely determined by it. The dampening of fluctuations at this level is stronger the smaller is | l_{max} | (or l_{max} instead of | l_{max} | in the continuous case). This particular case offers a further advantage, initially noticed by May (1974a): l_{min} completely determines the strength of fluctuation dampening for the aggregate biomass.

Hughes & Roughgarden (2000) found the following eigenvalues:

$$\begin{cases} l_1 = 1 - r \\ l_{rest} = 1 - \dfrac{r(1-h)}{(n-1)h+1} \end{cases} \qquad (4.2),$$

where the second through the n^{th} eigenvalues each equal l_{rest}. Note that $l_{max} = l_{rest}$ and $l_{min} = l_1$. The value of l_{min} is the same that was found from the model by Ives et al. (1999).

This means that l_{max} increases and the dampening capacity of single species decreases as either n or h increase, until reaching the threshold of instability | l_{max} |=1 for $n \to \infty$ or h =1. This is consistent with the findings by other authors (May 1974a, Tilman 1996, Ives et al. 1999, Lehman & Tilman 2000) , who have found species-level variability to increase with species richness or niche overlapping.

Conversely, l_{min} is unaffected by either n or h, which means that the dampening capacity at an aggregate level will not change with species richness. As a principle, this implies that there is no obstacle to the averaging effect, which will come into force once environmental fluctuations are taken into account. So Ives et al. (1999) concluded that increasing the number of species will in most cases decrease variability.

Surprisingly, Hughes & Roughgarden (2000) found the aggregate CV to increase with species number when calculating it explicitly from l_1. If each species biomass is subject to a perturbation of variance s_z^2 at each time step, the aggregate variance is

$$\operatorname{var}\left(\sum_{i=1}^{n} B_i\right) = \frac{s_z^2 n}{1-(1-r)^2} \qquad (4.3),$$

The average biomass B of each single species is

$$B = \frac{K}{1+(n-1)h} \qquad (4.4).$$

The total biomass $B_T = nB$ increases with n, but this does not suffice to counteract the increase of variance with n given s_z^2, when obtaining CV:

$$CV = 100\frac{s_z\left[1+h(n-1)\right]}{K\sqrt{n\left[1-(1-r)^2\right]}} \qquad (4.5).$$

In spite of so many interesting insights, we are thus left with contradictory results even at the level of specific simple models of competitive communities, and a lack of results for more complex and general models.

4.2.2. Statistical foundations for a stabilizing effect of diversity

Here I put forward a general argument to expect a stabilizing effect of diversity under very wide conditions. This is a statistical argument related to that of Doak et al. (1998), but immune to most of the criticism that this has received. It is not restricted to simple competitive communities: as a principle, it can be applied for any realistic pattern of interactions between species.

Take an ecosystem with an asymptotically high number of species and individuals. Without loss of generality, consider that we are just dealing with asexual species devoid of intraspecific diversity (it is straightforward to see that this is unnecessary for the argument to hold, but it allows to explain it more clearly). Let us classify the organisms in each species into a set of arbitrary subpopulations or “pseudospecies” and track their evolution separately (to ensure persistence, whenever a subpopulation goes extinct one can choose at random one organism in the species and assign it to that subpopulation). This will change some formal descriptors (n will increase, some strong interaction terms will arise due to the perfect competition between subpopulations in the same species) but there will be changes neither in the features of individual organisms nor in the dynamics of aggregate ecosystem quantities such as biomass. The pseudospecies will exhibit a larger biomass CV across time than the species, but this will be perfectly compensated by the covariances between them. The true diversification comes when the organisms belonging to different subpopulations are given different values of the parameters that define their response in front of different situations. As a principle, this should lead to a decrease in the correlations between the biomasses of the subpopulations and, on its turn, this should decrease the CV in aggregate biomass.

This imaginary experiment illustrates that an increase of the number of species in a model consist of two different operations:

- Increasing the resolution in the grouping of organisms into variables that are considered independently.
- Increasing the differences between organisms.

These two operations are the responsible for each of the two effects of diversity observed in the models: the first produces the increase in single-species *CV*, while the second produces the decrease in aggregate *CV*. Objectively, an increase of diversity consists just of the second operation, so the effects of the first one should not be considered to pose any conceptual problem.

When taking the average of several variables, both an increase in the number of variables and a decrease in the correlation between them will lead to a decrease of variance, other things equal. This allows to approach the diversity-stability issue in two different ways.

Doak et al. (1998) asserted that a positive diversity-stability relationship was statistically inevitable by thinking in terms of number of variables, and taking species as variables. The problem of this approach is that it is obvious that changing species number does not leave "other things equal" because, in general, the more species we have, the smaller and more variable will be the population of each one. This is the reason why posterior studies abandoned the search for general arguments and passed to analyze very specific models, to see which of the different effects of diversity dominated in each one.

Here I take correlations between variables instead of number of variables. This allows to eliminate the main "side effects" that appeared in the first case, because they were due to the first operation above, not to the second one. If there is any "side effect" left, it will also be a problem for the first approach, because it implicitly includes the second operation. The reason for taking into account the separation into species in the above thought experiment was for comparison with earlier approaches, but otherwise this would have been unnecessary. Diversity means differences between organisms, and covers from intraspecific variety to the variety of large taxa. There is a direct relation from difference to decorrelation and from this to decreased aggregate variance, unless there is some important side effect.

Is there any side effect left? In the case of simple competitive models, the only one is the increase in aggregate biomass, which will further decrease *CV*. It is due to the fact that, when organisms differ between them, they can cover a wider niche space. I address a subtler possible side effect in subsect. 4.2.3. In more complex communities, the only additional side effect that seems to be well-founded is a decrease in the intensity of trophic interactions (at least between hosts and parasites), which as a principle should further increase biomass stability. The most obvious reason to expect this change in interactions is that, when there is a high diversity, the rate of encounters of a consumer with potential preys with the particular features it requires will be low. See sect. 4.5 for a deeper analysis of this phenomenon.

The approach proposed in this section can be used to understand the origin of the contradictions that have arisen from the analysis of models of simple competitive communities. We can apply the above thought experiment to Eq. 4.1 As a first approximation, we take the case $h \approx 1$.

The first step is pseudodiversification, i. e. increasing n without increasing the real differences between organisms. This should not change the aggregate biomass variability. From Eq. 4.3, this requires

$$s_z^2 \propto n^{-1} \qquad (4.6).$$

The second step will be loss of correlation between species, if there is any, and this will increase biomass stability.

The source of the apparent destabilization found by Hughes and Roughgarden (2000) was taking a constant s_z^2 instead of something like Eq. 4.6. In so doing, they introduced an increasing noise input to the whole system with increasing n, which was not justified. This explains why an invariant aggregate dampening capacity seemed to coexist with an increasing aggregate variability.

The input of noise to systems like this one can be studied in more general terms. Two kinds of noise are usually recognized in population dynamics: demographic noise and environmental noise (see e. g. May 1973).

Demographic noise consists of random terms in individual organism reproduction and death events that are not correlated between individuals. Its variance is proportional to population, since the variances of independent variables are additive (e. g. May 1973).

Environmental noise consists of random terms that are correlated between organisms, because they depend on features of their common environment. In general it is assumed that, in terms of environmental noise, each organism has a correlation 1 with the rest of organisms in the same species and a correlation 0 with the organisms in other species (e. g. May 1973). This is an extreme assumption, but a weaker one will not qualitatively affect the following results, provided that the correlation within species is stronger than the correlation between organisms of different species.

The terms s_z^2 used by Hughes & Roughgarden (2000) are independent between species. Both demographic noise and simple environmental noise satisfy this condition. Indeed environmental noise does not satisfy it in the case of pseudospecies, so at this stage of our argument we will have to consider that Eqs. 4.3-4.5 refer to demographic noise.

If we assume the biomass B of each single species to be proportional to its population, both demographic and environmental noise can be seen to scale with B in the form:

$$s_z^2 = kB^q \qquad (4.7),$$

with some constant k and satisfying q=1 for demographic noise and q =2 for simple environmental noise (e. g. May 1973).

So, in the case of demographic noise,

$$s_z^2 = kB = B_T n^{-1},$$

where B_T is the total biomass and n is species number. Though obtained in a completely different way, this agrees with Eq. 4.6, whose validity is thus not restricted to the particular model we are dealing with (this result also holds on average if each species has a different k, provided that there is no covariance between k and B).

If we consider species instead of pseudospecies, environmental noise will also have to be taken into account. This will correspond to q >1 in Eq. 4.7, and will allow for compensation effects between the inputs of noise to each species. Since the aggregate dampening capacity does not change with n in this model, such compensations between noise inputs will effectively decrease the variability in the aggregate biomass.

From Eqs. 4.3, 4.4 and 4.7 we have

$$CV = xn^{-\frac{1}{2}}\left[1+h(n-1)\right]^{1-\frac{q}{2}},$$

where

$$x = \frac{100K^{\frac{q-1}{2}}}{\sqrt{1-(1-r)^2}} > 0.$$

It we derive CV about n as if n were continuous, we find

$$\frac{dCV}{dn} = -\frac{x}{2}\left[\frac{1-h}{n} + (q-1)h\right],$$

which satisfies $\frac{dCV}{dn} \leq 0$ for $q \geq 1$ and $0 \leq h \leq 1$. It follows that CV is a nonincreasing function of n also for discrete n (because the set of discrete numbers is just a subset of the set of continuous numbers).

When Hughes & Roughgarden (2000) found an increasing variability with diversity by keeping s_z^2 constant, they were unconsciously assuming that the responses of organisms to environmental fluctuations are negatively correlated for organisms in the same species and uncorrelated between species, which is senseless. In a paper that appeared once I had finished this chapter, Ives and Hughes (2002) also found that the contradictions in the results of their respective previous papers were due to differing assumptions about s_z^2, but took each option as equally valid. The above developments indicate that, in fact, there are conceptual inconsistencies behind the choice of a constant s_z^2.

Even if we introduce positive correlations between the effects of environmental noise on each species, we should expect a nondecreasing CV as n increases. In the limit case of q=2, correlation 1 between species and B_T increasing when n increases, we will find CV not to decrease, but neither increase, as n increases.

On the other hand, it should be stressed that a nondecreasing variability in relative terms, as measured by CV, does not necessarily imply a nondecreasing variability in absolute terms, as e. g. measured by the variance. Therefore, there is no contradiction between the expectations from these simple models and the findings by Pfisterer and Schmid (2002), which are perceived as puzzling by several authors (Pfisterer and Schmid 2002, Naeem 2002, Cameron 2002). Pfisterer and Schmid assembled a series of grassland artificial communities with different numbers of species, and subjected them to a drought. They found that, the richer the communities, the higher the production but also the higher the decrease of production under the drought. Diversity did not increase relative variability. In this case it just had no effect on it, since the production under the drough was a more or less constant proportion of the production without the drought.

The pattern of increasing stability in the aggregate community properties when diversity increases, as observed in most of the reported experiments, has a strong theoretical basis that does not depend on details of the models used.

4.2.3. The few species limit vs. the many species limit

The general assertions about diversity and aggregate biomass variability in the above section referred to ecosystems with an asymptotically high number of species and individuals. There is no basis to expect them to generally hold if we have just a few species.

As a principle, few-species communities can be modeled by means of low-dimensional dynamic systems. These systems are well-known to present a large array of possible dynamics, from strongly dampened point attractors to deterministic chaos (e. g. Schuster (1988), Wiggins (1990), Solé and Manrubia (1996); for application to ecology, see this last reference and Flos (1995), Hastings et al. (1993)). Different species compositions (given contour conditions) will produce different kinds of dynamics, and the resulting aggregate variances will be radically different. So in general species number will be much less relevant than species composition, and the effect of species number, if any, should not necessarily be one of decreasing aggregate variability. We can e. g. take into account that continuous systems require a minimum of 2 dimensions to pass from point attractors to limit cycles, and a minimum of 3 to pass from any of these to deterministic chaos (e. g. Wiggins 1990), so species number could even have a destabilizing effect.

Only when we have a large number of species we will attain high-dimensional systems, in which a statistical approach will become more useful than a dynamic-system approach. As quoted by Loreau (2000), Patten (1975) asserted that ecosystem complexity would confer behavioral regularity, because the many processes occurring simultaneously would average out in their net system effects. This is well-known from physics: statistical physics has rendered or justified a set of powerful laws of wide application, but they are only valid “at the thermodynamic limit”, i. e. for an asymptotically large number of particles (e. g. Garrod 1995, pg. 51). When the number of particles is not large enough to assume that the system is at the thermodynamic limit, the particular interactions between them have to be taken into account.

In the many species limit it is thus plausible the emergence of general regularities that do not hold for few species limit. In particular, one regularity likely to emerge is a decreasing aggregate biomass variability with increasing species number (however, the variability may not ever become as small as that of some few-species system with a strong point attractor). In the next chapter, I show that many different models can be expected to converge and produce some other regularities that are well documented to hold in real-world ecosystems, such as patterns of species abundance distribution or species-area relation, also by large number effects.

Some particular patterns of interactions may allow the effect of diversity on biomass variability to extend to the few species limit. This is the case of the models of simple competitive communities mentioned in subsections 4.2.1 and 4.2.2, which is the kind of model that has been studied until now to search for this effect. I have tested the consequences of introducing heterogeneous and asymmetric competition coefficients, by performing simulations of Lotka-Volterra competitive communities with coefficients chosen at random, and the effect still extends to the few species limit. This result is thus quite robust for competitive communities in general, although not necessarily universal (if the models in subsect. 4.2.2 are analyzed in detail allowing the parameters to differ for each species, it is easy to see that strong correlations between some of these parameters, such as a positive one between k in Eq. 4.7 and K in Eq. 4.1, would make the average CV increase as diversity increases; I know of no reason for these correlations when there are many species, but they could arise by chance if we have just a few).

So we have three well-founded hypotheses for statistical effects of diversity on the stability of aggregate properties:

- At the limit of few species, ecosystems are fragile, since a change of species composition can produce a drastic modification in their dynamics and structure. At the limit of many species, ecosystems acquire robust and predictable features.
- At the limit of many species, increasing diversity decreases the variability in aggregate community biomass.

- Where the only interactions between species are competitive, increasing diversity generally decreases the variability in aggregate community biomass for all the range of diversity values.

The empirical evidence in support of these hypotheses is still partial:

- There is evidence of fairly different kinds of dynamics in few-species systems, including deterministic chaos (Tilman and Wedin 1991, Hanski et al. 1993, Constantino et al. 1997). In this last case, the main limitation for empirical studies is that the standard methods to diagnose chaos require quite longer time series than are usually available in ecology. The author has developed one method that may help to overcome this limitation (Pueyo 1997).
- Next chapter provides evidence that many-species ecosystems display predictable features in their patterns of diversity, in agreement with predictions based on large-number effects.
- There is still no solid evidence of decreasing biomass variability with increasing species number in the case of ecosystems with multiple kinds of interactions at the limit of many species.
- There is evidence of decreasing biomass variability with increasing species number in competitive communities. This consists of observations for different numbers of plant species in savannahs (McNaughton 1985) or in experimental plots in temperate prairies (Tilman & Downing 1996).

Out of the experiments on diversity and stability that have recently appeared in high-impact journals, there is a specially worth-considering one in this context: the one by McGrady-Steed et al. (1997), with freshwater microcosms with several trophic levels. At first view, this seems to support a positive diversity-stability relation, but not in the way that we should expect from the above hypotheses. In App. 4A I reanalyze the data and show that in fact they agree with the above hypotheses and provide a nice illustration of them, though they cannot be used to proof them. Besides its direct relevance for this discussion, a reanalysis of McGrady-Steed et al. (1997) data is useful to see some of the problems that can arise when experimentally testing the relation between diversity and stability, and to plan future experiments.

4.3. Diversity and adaptability

4.3.1. Diversity as information

It is trivial that adaptation is not possible without variety. This is the most unquestionable way in which diversity can be said to have a stabilizing effect. In the long term, life could not even persist without diversity, though we do not know how much diversity is necessary to reasonably ensure persistence.

As a first approximation, this can be expressed by simple statistical reasoning, related to that treated in the above section. If each variety, species or higher-rank taxon has a given probability of going extinct in a given period of time, the more varieties, species or higher-rank taxa we have, the least probability that everything goes extinct. Things are however much more complex.

Adaptation is a flux of information from the environment to the genetic register. It increases the mutual information between organisms and their physical and biological environment (see chapt. 1). We could generalize the term adaptation to flows of information from the environment to sets of organisms, to species, or even to ecosystems. At least, we can talk of "adaptation" when the flux of information increases the probability of persistence of the biological system with which we are dealing. For a complex system, to "persist" should just mean to stay as a complex system, while details can and do change across time. This is indeed the main kind of "stability" (like, at the species level, already Preston (1969) noted that the most relevant stability is the persistence of the species, rather than the constancy of its numbers).

Margalef (1968) related diversity to the capacity of an information channel (this was the justification for his introduction of Shannon index as a measure of diversity), and this is fully pertinent in this context. Diversity has a direct relation with the maximum amount H_{max} of information that can be stored in a system, so it also increases the maximum possible amount I_{max} of mutual information between the environment and the system. However, for diversity to favor adaptation, the realized mutual information I should increase when I_{max} does. I. e. there should be some mechanism favoring an increase in I, at least up to some limit, for the constraint set by

H_{max} to have some effect. Ecosystems or parts of them should tend to achieve some predictable properties dependent on their environment, for diversity to increase adaptability (all of this is valid both if we consider diversity as number of different kinds of organisms and if we consider some combined measure of richness and equitability, such as the index of Shannon, but it is the first kind of diversity the most likely to act as a constraint on the dynamics, since equitability could as a principle change quite fast as a result of the own dynamics).

The most elementary case in which this can be expected to happen is the case in which a system dynamics is driven by some potential. Then the system will approach the configuration that maximizes or minimizes the potential, which on its turn may depend on external factors, so there will be a well-defined correspondence between external factors and properties of the system. I will thus tend to grow until some value (it will not be the maximum possible, because for this the system should not only approach an extreme of the potential, but also approach it infinitely, which would only happen for a fully deterministic, noiseless system).

Chapt. 2 makes clear that there is no basis to assume that complete ecosystems are driven by some potential. But this is not the case of single species. These tend to have the maximum rate of increase, which will be null at the steady state but still an extreme of a function, involving a high I. If we take in isolation species or sets of competing species under some assumptions (which are unlikely to strictly hold but could be taken as an approximation), the approach to these extremes involves increasing biomass or some related magnitude. Even though these potentials do not extend to the whole ecosystem, they are presumably the main responsible for the increase and maintenance of complexity. Adaptation can be understood in terms of them.

MacArthur (1969, 1970, 1972) analyzed in much detail the features of Lotka-Volterra symmetric competition models, when explicitly introducing the amount of available resources of different kinds and making their properties vary along an unidimensional continuum. Each species was given some “utilization function”, reflecting its capacity to use resources with different properties. The system was found to minimize the squared difference between actual and available production along the resource continuum. May (1973) compared the species in this model with the components in a Fourier decomposition as used in time series analysis. The weights of these “generalized Fourier components” are those that maximize the fit to the resource continuum. Indeed, the more components we have, the better will be the fit. I. e., the more species the higher the mutual information with the environment and the higher the biomass. The more species, the more adapted the system will be.

Experiments with competitive communities (Tilman 1996, Hector et al. 1999, Loreau and Hector 2001) seem to uncover not only an increasing stability but also an increasing resource use, production or biomass, which supports that diversity increases adaptability in a wide sense. Not only intraspecific variety is necessary for species adaptability at evolutionary time scales, but also species diversity favors ecosystem adaptability at successional time scales.

A decrease in diversity makes a system both less adapted and less adaptable. The second cannot be completely captured by any simple statistical reasoning like that in the second paragraph of this subsection. An efficient search of extremes of a potential is not a matter of independent trials. Experience with computer algorithms (Kirkpatrick et al. 1983) indicates that it is a process that can be neither completely exploratory nor completely conservative. Most “trials” should take place close to the already “occupied” sites in the “adaptative landscape” (concept introduced by Wright 1932), though there should also be some trials at long distances from them. Biodiversity loss often implies a loss of memory of some of the previously occupied areas in the adaptative landscape, so recovering an equivalent degree of both adaptation and adaptability after the loss is computationally complex, and this is probably the reason why it takes so long, according to the fossil register (see Kirchner and Weil 2000). If adaptative landscapes were unimodal, like e. g. those in Eigen et al. (1988) model, the recovery of diversity would be a fast and potentially rapid operation just requiring an increase in the rate of genetic modifications. However, adaptative landscapes are “degenerate”, i. e. they have many extremes. There are two kinds of biodiversity: the dispersion of organisms around peaks, and the spread between peaks. The first is indeed necessary for adaptability, but has a cost in terms of current degree of adaptation and is easy to recover once it has been lost (unless the own peak has already been lost because of the lack of adaptability). The second does not necessarily have a cost in terms of current degree of adaptation and makes a proportionally larger contribution to adaptability, and also to the kind of stability treated in section 2. Furthermore, a part of the “degeneracy” corresponds to complementarity in terms of function in the ecosystem. It is precisely this second kind of diversity that is difficult to recover.

These comments are still qualitative and heuristic, and we are very far from passing to quantitative predictions for the relation between diversity and adaptability. But these are required. The biodiversity loss driven by humanity is really unlikely to threaten the capacity of the Biosphere to persist, but it may well threaten some of its parts. It is not absurd to wonder whether one of the threatened parts is the own food production system that sustains humanity. Humanity nowadays depends essentially on a few plant species, and their intraspecific diversity is shrinking at a formidable rate, so we are close to the lower end in the range of possible diversities, just where qualitative changes are most likely to take place. The species that take part in our food production system may require adaptability in front of natural and anthropogenic environmental variations, possible socioeconomic changes with an impact on agricultural procedures, and new pests and pathogens. This last component could well be the most demanding one in terms of diversity, because the biological environment is far more complex than the abiotic one. Sect. 4.5 puts forward some hypotheses that could lead to applicable conclusions in this important field.

4.3.2. One role for criticality

It is evident that biodiversity is necessary for adaptability because it is the pool of information that offers the possibilities to be filtered by natural selection. Here I suggest that diversity favors adaptability also in a different, complementary way.

As explained in sect. 4.2, models suggest that diversity decreases variability in the aggregate biomass of competitive communities, but increases the variability in the biomass of the individual species. This is because these models assume that a higher diversity is linked to a higher niche overlap, which makes regulation on individual species less intense. Indeed, beyond some point, diversity requires either a decrease in niche width or an increase in niche overlap. In many situations the last may well take place.

Is the decrease in species-level variability just a side effect of increasing diversity, unrelated to aggregate stability, or it has a direct relationship with the last? Tilman (1996) suggested that the second point of view could be the right one: “When climatic variations harm one species, unharmed competitors increase. Such compensatory increases stabilize global biomass, but cause abundances to be more variable”. But this point of view is not backed by the analysis of the simple competition models in subsect. 4.2.1. Eq. 4.2 shows that species-level regulation strength depends on niche overlap (l_{rest} is a function of n and h), but community-level regulation does not. If an increase of diversity is linked to an increase of specialization instead of a higher overlap, single species will not become less regulated but the expected effect on aggregate stability will be the same. It is true that, if each single species remains strongly regulated, the biomasses of Tilman’s “unharmed competitors” will not increase so much, but neither the biomass of the “harmed species” will decrease so much. These two partial effects cancel each other, so the net effect on aggregate variability remains the same. Otherwise, it would not have been possible to dissociate the effects at single-species and community level in subsect. 4.2.2.

However, some level of deregulation of individual species (let alone genotypes) must be necessary for adaptability as defined in subsect. 4.3.1. By definition, ecosystems should be able to respond to external changes, if they are to adapt. E. g. if two species (or genotypes) occupy a different position in MacArthur (1969, 1970, 1972) resource continuum, they should be able to change their proportions when the continuum of resources changes. Lack of niche overlap is not really a problem in MacArthur’s model, because of its simplicity, but it will be a problem when there are several independent regulation mechanisms. The proportions of the above species will not change so much in response to e. g. a change in the available kinds of food if there is a high niche segregation in terms of shelter or nesting sites, which keeps them strongly regulated. Neither e. g. if each one is strongly regulated by its own predators.

A system cannot be rigid if it is to be adaptable. As stated in last subsection, the essential kind of stability for a complex system is its capacity to persist as a complex system, and this requires the details of the system to be modifiable in response to external changes.

In simple competition models, any force that favors diversification (without enough increase in specialization to avoid niche overlapping) pushes species to the limit of their dynamic stability. This corresponds to total niche overlapping in simple deterministic models, and a smaller but still well-defined degree of overlapping when taking noise into account (May 1973). We could find systems permanently above the "noisy" limit if there is a high enough rate of originations to counteract the extinctions, which will be frequent at least for different genotypes within a species. The deterministic limit will indeed not be surpassed, because there is no region defined beyond it in these models, unlike other models. This partially reflects that the carrying capacity is finite.

I show in chapter 5 that, at the limit of very weak regulation, the steady state of each species population is a critical point in the physical sense of the word. So competitive communities subject to diversification approach criticality, by a mechanism unrelated to self-organized criticality (SOC). SOC models are currently the object of a high interest, and are treated in other parts of this book (specially chapter 3). In sections and I discuss the role that diversity could play in SOC dynamics.

Langton (1990) suggested that complex information processing can only emerge spontaneously close to criticality, because, below criticality, signals would only propagate to a limited extent, while above criticality we would reach a "chaotic" situation where signals propagate without limits. Only close to criticality signals would propagate to a finite but arbitrary extent, as required for complex information processing. I show in chapt. 3 that, in fact, some of the so called self-organized "critical" systems are not really critical, while sharing the main properties of criticality, including the intermediate propagability of signals. This property seems to be also required for ecosystem adaptability, if we take as "signal propagation" the growth of populations of organisms with given features. The above developments suggest that, in this case, this property really arises due to the proximity to a critical point.

Systems close to criticality have well-defined and testable properties. I show in chapter 5 that, as a whole, the body of knowledge about diversity patterns (species abundance distributions and species area relations) accumulated along the 20th century agrees with the expectations from assuming that ecosystems are close to a critical state, and this is reinforced by my own analysis of a large set of marine phytoplankton data. Furthermore, if we assume that the mechanism that leads to criticality is the one proposed in this section, it is possible to extract useful information from the analysis of abundance distributions. E. g., my results for marine phytoplankton give a stronger basis to expect a stabilizing effect from dinoflagellate diversity than diatom diversity.

4.4. Complex interaction webs

4.4.1. Introduction to May's criterion

Robert May was the author of one of the findings with a highest impact on the diversity-stability debate (May 1972, 1973). As explained in the introduction, he analyzed the stability of randomly assembled model ecosystems and found that, by increasing species number, connectance or interaction strength between species, model ecosystems pass from being almost certainly stable to being almost certainly unstable as soon as a threshold is surpassed. In the terms of subsect. 2.3.3, this corresponds to a first order phase transition. The kind of stability he analyzed was what here I call dynamic stability, which consist on the capacity of an ecosystem to recover a given set of species abundances after being disturbed (see Svirezhev and Logofet 1983 for more detail on the dynamic stability of ecosystem models). May advocated for a high prudence in the translation of his findings to real-world ecosystems, which has not always been taken into account (see Taylor 2000, May 2000).

Tilman (1996) solved the apparent discrepancy between May's theoretical findings and the outcome of recent experiments on diversity and aggregate community stability by stressing that they referred to stability at different levels. He remarked that the own May (1974a) developed a model of a simple competitive community and found that the destabilization at the single-species level was not accompanied by a destabilization at the aggregate level. May (1972, 1973) complex web model does not allow for predictions concerning aggregate stability, but the reasoning in sect. 4.2 suggests that aggregate stability should asymptotically increase with diversity even when we have a complex interaction web, and May's criterion does not contradict this result. So here I discuss May's criterion for its interest in the compositional stability of communities, leaving apart the issue of aggregate biomass stability.

May (1972, 1973) studied the properties of random model ecosystems as a function of species number n, connectance C and a measure of interaction strength α. C is the proportion of realized connections between species, out of the possible connections.

Consider an ecosystem with n species i=1,...,n, whose biomasses are $\{B_i\}$. May used populations instead of biomasses, but this does not affect his results, and using biomass is more convenient for my developments in the next section.

Assume that the ecosystem is at steady state, with $B_i = B_i^0$ given i. Now introduce a fluctuation in the form of a small deviation of some species' biomass about its equilibrium value,

$$x_i = B_i - B_i^0 \qquad (4.8).$$

This fluctuation may propagate to other species. If the ecosystem's initial steady state is dynamically stable, it will return to this state. Otherwise, the fluctuation may amplify and the initial steady state will be lost. If we want to know which is the case, we only need to have a linear approximation to the direct influence of each species on each other one in the neighborhood of the steady state. I. e. we need the set of interaction strengths

$$a_{ij} = \frac{\partial}{\partial x_j}\left(\frac{dx_i}{dt}\right) \qquad (4.9)$$

for each i and j. Close to the steady state,

$$\frac{dx_i}{dt} = \sum_{j=1}^{n} a_{ij} x_j \ .$$

May produced sets $\{a_{ij}\}$ according to the following procedure. In a first step, each coefficient a_{ij} is set to 0 with probability $1-C$. Whenever $a_{ij} \neq 0$, a_{ij} is taken out of a continuous distribution with mean 0 and standard deviation a. In a second step, a unit is subtracted from each of the self-interaction terms a_{ii}.

By direct application of Wigner's “semi-circle” law (Wigner 1959, Mehta 1991), he found the condition of stability (for large n):

$$a(nC)^{1/2} < 1 \qquad (4.10),$$

which can also be expressed as

$$(nC)^{1/2} < 1/a \qquad (4.11).$$

The steady state is almost certainly stable if Eq. 4.11 is satisfied, and almost certainly unstable if it is not.

E. g. Fig. 4.1 displays the variation in stability with C for a set of matrices build as proposed by May (1973, 1974). Fig. 4.1A displays the probability of stability, and Fig. 4.1B. displays the average of l_{max}, which is the maximum eigenvalue of the

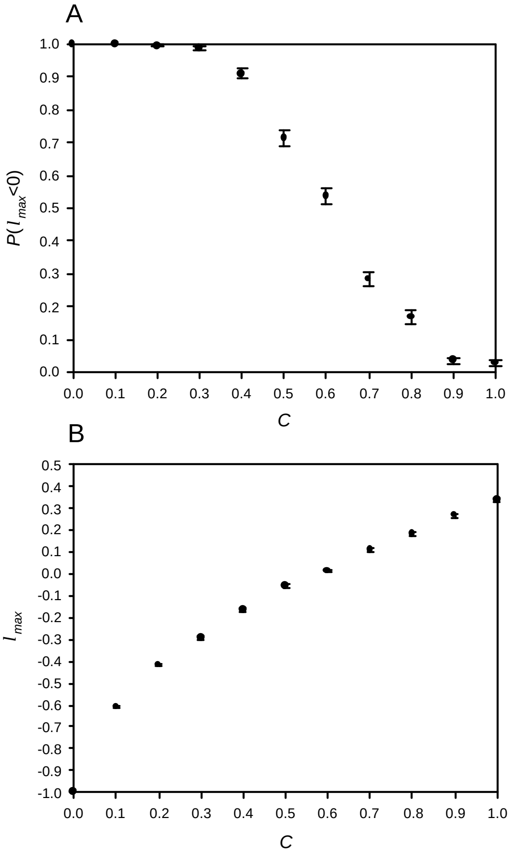


Fig. 4.1. *The destabilization effect found by May. Relation of the connectance C with the probability of stability P(*l_{max} >0) (**A**) *and the average of the maximum eigenvalue* l_{max} (**B**) *for* a=0.25 *and* n=32, *from a sample of 400 matrices per point. The error bars indicate standard error.*

matrix of interaction coefficients. A system is stable if and only if l_{max} <0. If this condition is accomplished, the dampening of fluctuations is stronger the smaller is l_{max}.

Note that in the figure P(l_{max} <0) does not change abruptly, since an abrupt change is only expected at the limit $n \rightarrow \infty$. Gardner and Ashby (1970) findings (which were the basis of May's work) indicated that this would also be found for small n, but Solow et al. (1999) have recently shown that this was incorrect.

4.4.2. Some limitations of May's criterion

Since nC is roughly the number of connections per species and a is a measure of their intensity, what Eq. 4.11 tells us is that connections destabilize and that the destabilization is rather stepwise. In this section I show that, at least to some extent, this result is due to features of May (1972, 1973) approach that do not reflect reality. As a result, (i) the negative relation between number of connections and dynamic stability is strongly relativized, although I maintain the hypothesis that it still provides a good approximation for ecological purposes, and (ii) the idea that this relation takes the form of a first order phase transition is rejected.

One of the steps in the process of building May's matrices consists of subtracting an unit from all of the diagonal terms. This step was taken for convenience instead of realism, but it has a major effect on the outcome.

Considering that (*i*) a matrix with n rows and n columns has n eigenvalues $\{l\}$, and (ii) the sum of the eigenvalues equals the trace (the sum of the diagonal elements of the matrix), it follows that, given a set of self-interaction terms $\{a_{ii}\}$, the maximum stability corresponds to $l_{\max} = \overline{a}_{ii}$. When a_{ij}=0 $\forall i \neq j$, this value of l_{max} is attained if and only if all of the diagonal terms are identical.

So, by construction, when there are no interactions May's matrices reach the maximum stability compatible with their trace (which does not change on average when modifying C or a). Then adding connections can only decrease stability. But we can also build fully connected matrices "prefabricated" to have maximum stability. App. 4B gives the recipe. For these matrices, cutting interactions down has an effect similar to adding interactions to May's matrices, as shown in Fig. 4.2A.

If a set of subsystems have the same "degree of stability", linear interactions between them can only decrease it. Conversely, if they are different, the appropriate linear interactions can "redistribute" the stability in such a way that l_{max} decreases. In particular, if some subsystems are unstable by themselves, only interactions can make the overall system stable. This is found in any model more elaborate than May's,

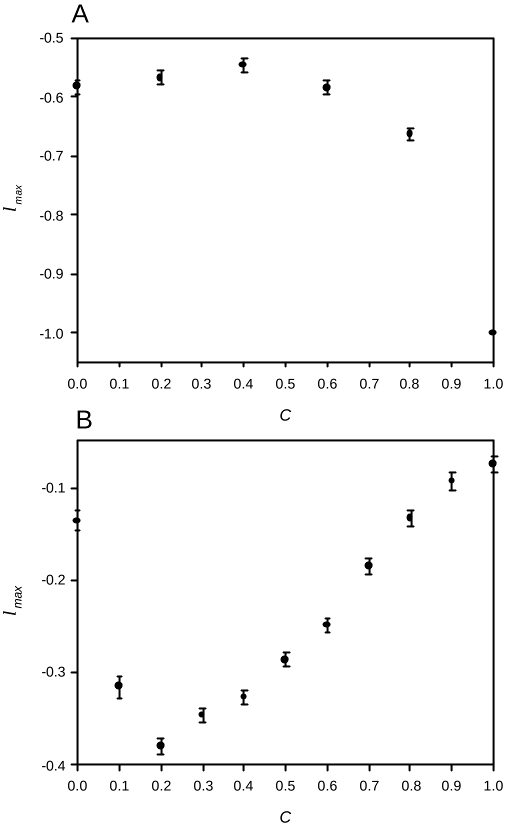


Fig. 4.2. *Response of the average of the maximum eigenvalue* l_{max} *in two different Montecarlo experiments (subsect. 4.4.2), with n=32 in both cases. The error bars indicate standard error.* **A:** *response to C in the case of matrices built to have their maximum stability when fully connected, a=0.25, from a sample of 400 matrices per point.* **B:** *response to C in the case of random diagonal terms and only interactions* +-, *a*=0.425, *from a sample of 400 matrices per point.*

where species will have different self-interaction terms and some of them may rely on others for their stability (e. g. predators on their preys). So May's point of departure is highly artificial.

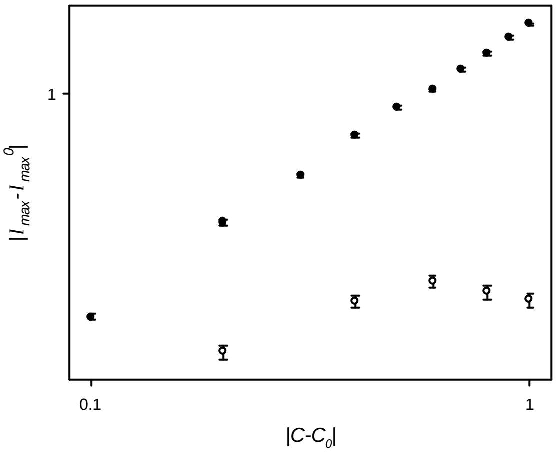


Fig. 4.3. *Comparison of structural stability for May's unconnected matrices (open circles) and fully connected matrices with the same eigenvalues (solid circles) (see subsect. 4.4.2). Variation in* $|\ l_{max}$ - $l_{max}^0|$ *with* $|C\text{-}C_0|$, *where* C_0 *is the initial connectance (*C_0*=0 in the first case and* C_0*=1 in the second) and* l_{max}^0 *is the corresponding eigenvalue. In both cases, n=32 and a =0.25. A sample of 400 matrices per point was used. Error bars indicate standard error.*

May (1973) recognized that it was possible to find dynamically stable matrices that would not satisfy his criterion, but presumed that they would be structurally unstable. He thought that any small change in parameters would make them lose their dynamic stability. As pointed out by Haydon (1994), this has not been demonstrated. Once we have our fully connected matrices with maximum dynamic stability, we can compare the effect of a structural change such as varying *C* with the effect of this change for May's matrices with *C*=0. Fig. 4.3 plots $|C\text{-}C_0|$ vs. $|\ l_{max}$ - $l_{max}^0|$ for both approaches, where "0" refers to the initial matrices with maximum stability. The changes in l_{max} with *C* in the initially unconnected matrices exceed that of initially connected matrices, which is a symptom of higher structural stability in the last case.

One possible answer to this criticism is that matrices built as in App. 4B may be even more artificial than May's matrices. While straightforward to synthesize by matrix operations, it is not so clear that they could be reached through natural self-organization processes. Perhaps dynamic stability dependent on interactions can evolve for small sets of species while in asymptotic terms adding interactions destabilize. A simple model with no realism at all but useful to illustrate this idea is presented in Fig. 4.2B. There I take a diagonal matrix with random terms (whose average is -1) and then increase *C*, but only with interactions of the type “+-“(a interaction +- between *i* and *j* is an interaction that satisfies $a_{ij}a_{ji} < 0$). These interactions are the most likely to increase stability (May 1973) (which can be increased because the diagonal terms are not equal). The result is first an increase of stability with *C*, but beyond some point a decrease. So asymptotically this converges with May’s model.

The fact that the threshold of instability in May’s model has a general well-defined value in terms of *n*, *C* and a is also a consequence of having equal diagonal terms of departure. So, even if we accept interactions to have an asymptotically destabilizing effect, we should not expect a clear-cut relation like Eq. 4.11, not even for large *n*. Furthermore, even if the diagonal terms could be taken as homogeneous (which is difficult to conceive), we would still have to find out whether Wigner semi-circle law holds for highly leptokurtic coefficients $\{a_{ij}\}$ (which would be realistic to expect, since they partially depend on population numbers, whose statistical distribution is very leptokurtic, close to a power law, as treated in chapt. 5). And also try the effect of other properties of ecosystems that differ from those of random matrices.

4.4.3. May vs. Odum and MacArthur: who was right?

Both E. P. Odum (1953, quoted in MacArthur 1955) and MacArthur (1955) on the one hand and May (1972, 1973) on the other made landmark contributions to the debate on diversity and stability, standing on solid theoretical grounds. However, they were contradictory and, to my knowledge, the source of this contradiction has still not been found until now. This should be solved if the theory of diversity and stability is to become coherent.

MacArthur (1955) paper was an elaboration of Odum (1953) idea that complex food webs make communities more stable. MacArthur viewed a stable community as one in which “the abundance of species tends to stay quite constant”, in contrast to others in which “the abundances vary greatly”. From this concept, he made the non-obvious step of defining that a community is unstable when an “abnormal abundance” in a species produces marked changes in the abundances of other species. He offered an intuitive explanation for the relationship between web complexity and stability thus defined and, as a consequence of this purported relationship, proposed to use a measure of web complexity (Shannon entropy of the magnitudes of the energy flows between species) as a measure of stability (unlike often assumed, he did not derive from information theory the idea that complexity stabilizes).

The intuitive explanation by MacArthur captures the ecological consequences of the statistical properties of the average of several variables, as explicitly put forward by Margalef (1989, pg. 698), after the work by den Boer (1968) and Reddingius and den Boer (1970). I. e. it is an expression of the "averaging effect" that has received so much attention since rediscovered by Doak et al. (1998). The only difference is that Doak et al. look at the ecosystem from outside and realize the statistical advantage of putting many species together, while Odum and MacArthur took the point of view of one of the species within the ecosystem and realized the statistical advantage of interacting with many other species. Recent experimental evidence supports that predators can increase their stability by diversifying their preys (Petchey 2000).

Let us take a closer look at the averaging operation. It is a double operation: we sum up n variables and divide the result by n. The sum increases the variance but the division decreases it in a greater amount and thus the net effect is variance-decreasing. It is important to note that there should be a division for the variance to decrease. In the case of trophic webs, we can safely assume that, if one organism feeds on many species, each species' contribution to its diet will be smaller than it would be if its diet was less diverse. It is a matter of physiological constraints. It is also a matter of definition, provided that we are working on the number of species and connections and not on the total amount of energy that flows through the system. In the same way, we may consider that in the many species and many connections case an organism will have more potential predator species but a smaller probability of being eaten by each specific one. Following Odum (1953), MacArthur (1955) proposed to measure stability as "the amount of choice that the energy has in following the paths up through the food web", which is inverse to the amount of energy circulating through each path.

As a principle, this could be introduced into May's formalism by taking a to be roughly inverse to nC. However, things are not so simple. Here I justify and look at the consequences of taking a inverse to nC, and after that I show why "things are not so simple” and which is the final picture if we also take into account some other essential factors.

The exact meaning of $\boldsymbol{a}$ and the set $\{a_{ij}\}$ is model-dependent. As a first approximation, let us take as a model the standard Lotka-Volterra trophic interactions:

$$\frac{dB_i}{dt} = f(B_i) + \sum_{j\in\Phi_i} \boldsymbol{h}_{ij} B_i B_j \qquad (4.12),$$

with a set of interaction parameters $\{\boldsymbol{h}_{ij}\}$. Φ_i is the set of species j such that $\boldsymbol{h}_{ij} \neq 0$. From Eqs. 4.8, 4.9 and 4.12,

$$a_{ij} = \boldsymbol{h}_{ij} B_i^0 \qquad (4.13),$$

given i, j.

We split each set Φ_i in the subset of species j that are consumed by i ($\boldsymbol{h}_{ij}>0$) and the subset of species j that consume i ($\boldsymbol{h}_{ij}<0$). Let us call the first Φ_i^+ and the second Φ_i^-:

$$\frac{dB_i}{dt} = f(B_i) + \sum_{j\in\Phi_i^+} \boldsymbol{h}_{ij} B_i B_j + \sum_{k\in\Phi_j^-} \boldsymbol{h}_{ij} B_i B_j \ .$$

At the steady state, $h_i^+ = \sum_{j\in\Phi_i^+} \boldsymbol{h}_{ij} B_j^0$ is proportional to the total food intake of species i per unit biomass and unit time, while $h_i^- = \sum_{j\in\Phi_i^-} \boldsymbol{h}_{ij} B_j^0$ is the fraction of i's biomass consumed by other species per unit time. From Eq. 4.13,

$$\begin{cases} h_i^+ = \displaystyle\sum_{j\in\Phi_i^+} a_{ij} \frac{B_j^0}{B_i^0} \\ h_i^- = \displaystyle\sum_{j\in\Phi_i^-} a_{ij} \frac{B_j^0}{B_i^0} \end{cases} \qquad (4.14)$$

A constant $\boldsymbol{a}$ means that the distribution of the set $\{a_{ij}\}$ in Eq. 4.14 is independent from nC. In randomly connected systems, we should not either expect the relations B_j^0 / B_i^0 to change in any preferential direction when modifying nC. As a result, the average values of h_i^+ and h_i^- will be proportional to the number of trophic links per species, which is nC. As above justified, this is not acceptable.

In order to make h_i^+ and h_i^- independent of nC, May's interaction strength parameter $\boldsymbol{a}$ should have the form

$$a = \frac{b}{nC}.$$

Once we have this relation, we can forget a and substitute Eq. 4.11 by the following equation:

$$(nC)^{1/2} > b \qquad (4.15).$$

Eqs. 4.11 and 4.15 are formally equivalent, there is no mathematical criterion to choose between them (b is a function of a and nC as much as a is a function of b and nC). The criteria for the choice should be ecological. Eq. 4.15 is obtained from MacArthur's premises and, at first view, seems to confirm his conclusion, which is just the opposite than Eq. 4.11 seemed to communicate: simple model ecosystems are not expected to become unstable above a certain value of nC but below a certain value of nC.

This is the result of introducing in isolation an internal "averaging effect" in the system. Once the initial fluctuation is introduced, May's model behaves in a deterministic way. However, for each species, the trajectories followed by each of the other species that interact with it have the properties of a random sample, so the averaging operation dampens the propagation of the oscillations, despite its deterministic nature. This is the reason why the averaging effect is reflected in the eigenvalues, unlike the case of the competition models in sect. 4.2.

The interactions in May's model are not just trophic interactions. However, for all kinds of interactions it seems more reasonable to take a constant b than a constant a. As a principle, an increase in diversity or connectance does not mean that each individual is going to interact with more individuals, but that the individuals with which it will interact will be more diverse. This is quite intuitive for mutualism, competition, etc. McNaughton (1978) had already found that the empirically estimated coefficients of interaction between competing plants in savannahs decreased with diversity, and remarked that this should be taken into account when taking conclusion for the real world from May's criterion.

The problem with MacArthur ideas and with the above apparent reversion of May's criterion is that diversifying the connections between species has other effects besides introducing this operation of averaging. When predators increase the diversity of prey species they consume, they will be able to average out the oscillations in the preys, which increases their stability. But each predator species will increase its niche overlap with other predator species, so the predator-prey regulation loops will be weakened. This is just an application of the same rule that holds for simple competitive communities, where l_{max} decreases as the overlap between competing species increases (sect. 4.2). May's model includes competition for resources not included in the model (similarly to simple competitive models), competition for resources within the model (preys), and "apparent competition" (in the terms of Holt 1977) for the predators. In all cases, increasing the number of connections per species nC will increase the overlap between regulation loops and thus weaken them.

It is not immediate to introduce all of these effects into May's formalism because, as stated in the previous section, it makes no explicit recognition that some species depend on interactions within the system for their stability. But we can get some orientation from other models. Take the case of a simple competitive community with equal parameters for all of the species:

$$\frac{dx_i}{dt} = r\left(\frac{K - x_i - \sum_{j \neq i} \boldsymbol{h} x_j}{K}\right)$$

The terms $\{a_{ij}\}$ of the linearized interaction matrix (at the steady state with non-null populations) have the form:

$$a_{ij} = \frac{-\boldsymbol{h}r}{1 + (n-1)\boldsymbol{h}},\ \forall i, j : i \neq j$$

$$a_{ii} = \frac{-r}{1 + (n-1)\boldsymbol{h}},\ \forall i\,.$$

Note that the interaction terms a_{ij} weaken with n (they would weaken with nC instead of n if the system were not fully connected), as expected from the above reasoning and empirically measured by McNaughton (1978). Also the self-regulation terms a_{ii} weaken because of niche overlapping. The interesting point is that both change proportionally, so the qualitative dynamic stability properties of the system are conserved in a normalized matrix with terms independent from n:

$$a_{ij} = -\boldsymbol{h},\ \forall i, j : i \neq j$$

$$a_{ii} = -1,\ \forall i\,.$$

We cannot perform the same development with all of the interactions in May's model, but we can make the assumption that the consequences will be comparable.

Then it turns right to use a matrix with a independent from nC and we are back to the original May's model.

As a conclusion, once we put together all of the above elements we find that May's criterion cannot be rejected in favor of the postulates by earlier authors. This is a reason to support that a high number of connections per species decreases stability. However, this conclusion requires at least four amendments:

First, the above destabilization is expected just because a multiplicity of parallel interactions weakens regulation mechanisms. However, for the same reasons it will also weaken any deterministic mechanism that destabilizes instead of regulating. This is the reason why McCann et al. (1998) found that the addition of some particular links to some particular webs stabilized them by changing their dynamics from deterministic chaos to limit cycles, and from these to point attractors. However, the weakening of stabilizing loops is likely to have a more general interest.

Second, it is well known that species-level instability does not imply system-level instability (Tilman 1996, and sect. 4.2 in this chapter) and, according to the developments in subsect. 4.3.2, in some measure it is necessary for adaptability.

Third, a low biomass variability "in ordinary conditions" is just one aspect of stability. In subsect. 4.4.4 below I put forward the hypothesis that a high number of interactions per species favors another facet of stability, by making the system more robust in front of structural changes.

Fourth, even species-level biomass variability has different facets. Here I have only treated the variance of biomass when measured across an asymptotically long time span, like most of the recent literature on diversity and stability. However, it is not less relevant the velocity of the oscillations between the different levels of biomass that a species can potentially attain. This is likely to be the facet of stability to which E. P. Odum (1971, pg. 69) referred when putting forward his expectation of "violent fluctuations in numbers" in simple trophic webs. This is directly linked to the conditions found by individual organisms across their lives, which has a foremost biological importance. At the asymptotic limit of a weak species-level regulation, an increase in the number of interactions per species can be expected to make biomass oscillations more gradual (with a smaller variance in the relative rate of change, and a reddened power spectrum, as often found in nature, according to Pimm and Redfearn 1988), as a direct consequence of the "averaging effect". This can be compared with the case of the portfolio (like Tilman et al. (1998) do with aggregate community biomass), replacing species biomass by individual income: in the absence of market regulations, the income oscillates and could attain a wide range of values if enough time were left (which is not the case, because individuals have a finite life span) with or without diversified portfolios, but these slow down the oscillations.

MacArthur stated that natural selection should find an optimum degree of specialization, taking into account that this increases efficiency and decreases stability. This could well hold, if we take into account that (i) variability in the relative rate of change has a major biological importance and is asymptotically reduced with increasing number of connections and (ii) when a species diversifies its connections, it will generally increase its own stability both in terms of biomass and rate of biomass change, regardless of the impact that this has on the system as a whole. On the other hand, the number of connections per species could have a lower limit for reasons that are pointed out in next subsection.

My tentative conclusion is that both May on the one hand and Odum and MacArthur on the other captured different facets of reality, which can be integrated into a coherent framework.

Despite integrating so different points of view, my approach in this subsection is still simplistic and leaves several factors out. One of the most important is the “dilution effect” that may be produced by a low connectance, which is treated in sect. 4.5.

The next subsection deals with a complementary issue, which is the relationship between the features of the interaction web and extinction dynamics.

4.4.4. Extinction dynamics

Several authors have developed models of self-organized criticality (SOC) in extinction dynamics (see Drossell 2001 for a review). In such models, the webs of interactions between species self-organize in such a way that leads to avalanches of extinctions of all sizes, with a power law distribution, like the model of forest fire worked out in detail in chapt. 3. Here I do not express any position on whether or not extinctions really have a SOC dynamics. However, I find it opportune to make clear the relationship between this hypothesis and the developments in the rest of this section, and also outline some possible consequences of the findings in chapt. 3 if the hypothesis is eventually found to hold.

In subsect. 4.4.1-4.4.3 and I work with models of communities with a constant species composition, interaction parameters and contour conditions, except for some stationary noise sources. Although the real world is not like this, this can suffice to analyze the main features of biomass fluctuations at some time scales. At a higher time scale, it will not be possible anymore to ignore “structural” changes in the above features. Extinctions are prominent among such changes. A high biomass variability at the species level (in the absence of structural changes) may indeed increase the frequency of extinctions, but this is not the whole story. The own structural changes, such as the extinction or the addition of one given species, will modify the population dynamics of other species and may lead them to extinction. I. e., there may be avalanches of extinctions, which have no direct relationship with biomass variability. There are several well known instances in the literature of major ecological changes when some given “keystone” species is eliminated, or as a consequence of some biological invasions (Chapin III et al. 2000).

In most simple models, propagation phenomena can be studied with percolation theory (see e. g. Stauffer and Aharony 1992). Plotnick (1993) applied percolation theory to extinction dynamics, by modeling bidimensional networks of species, in which the extinction of the species in a node could propagate with varying probabilities to each of the species in the 4 nearest nodes. He allowed the probabilities to evolve across time and developed essentially the same model that I use in chapt. 3 as a point of departure for studying forest fires, before introducing external parameters.

At the other extreme, we could allow each species to have connections at random with any other species, as May assumed when developing his criterion of stability. This leads to another criterion of stability, which refers to a different kind of “stability” but has the same form. Let r_{ij} be the probability of an extinction propagating from species i to species j, and let r be the average of $\{r_{ij}\}$ for $r_{ij} \neq 0$ and $i \neq j$ (similarly to a). If we take a large number of species n, a connectance C, and a given r (which may vary as a function of n and C), we find the criterion by direct application of percolation theory

$$r(n,C)nC<1 \qquad (4.16)$$

(the analogy to May's criterion is even more apparent if we rewrite Eq. 4.16 as $t(nC)^{1/2}<1$, where $t=r^{1/2}$). In this case, the meaning of "stability" has to do with the size of fluctuations. Above the critical threshold, an avalanche of extinctions may propagate indefinitely. Eq. 4.16 is in fact more solid than May's criterion, because it has no requirements on the diagonal terms, which arguably constitute the most artificial premise for his results, as shown in subsect. 4.4.2 (in this case, the diagonal terms play no role).

Realistic web architectures seem to lie someway between Plotnick's and May's webs. The position in this gradient does not change the essential conclusions from percolation theory, although these would indeed be useless for specific nontrivial web designs. If this is not the case, we should expect a well defined percolation threshold similar to that in Eq. 4.16.

Take the number of connections per species $f=nC$. Since r will change as a function of f, it is not obvious whether f stabilizes or destabilizes. It seems intuitive that if each species has links with just a few other species, it will strongly depend on some of them and the avalanches of extinctions will propagate more easily. If this is the case, Eq. 4.16 (or its equivalent in a more realistic web design) will pose a lower limit on f (instead of an upper limit), although this requires a deeper analysis (the developments in last subsection suggest that a high f makes each species less strongly regulated, which may indeed contribute to extinctions but also favor a pattern of random disconnected extinctions over one of synchronized avalanches).

As shown in last subsection, the elaboration of MacArthur's ideas leads to expect natural selection to favor one given level of f, which we may call f_s. On the other hand, Eq. 4.16 establishes a critical f given r, which we may call f_c. Then two things may happen. If $f_s > f_c$ (or $f_s < f_c$, if f in fact destabilizes) the trade-off will be solved just in terms of natural selection and we will have $f = f_s$, as a principle. Otherwise, system-wide mechanisms will come into force and the system will attain a self-organized critical (SOC) state. Since f is not necessarily the most relevant control parameter, we can state more generally that the system will attain a SOC state if, in the absence of extinction avalanches, it would tend to achieve some value of rf above the critical value $(rf)_c$ that can be extracted from Eq. 4.16.

This is not really the path to SOC proposed in the models in the literature (this may or may not be more realistic). In App. 4C I analyze one of these models (by Solé and Manrubia 1996), pimpoint the differences and show how to modify it to partially illustrate the above proposal.

A general characteristic of SOC dynamics is the generation of scaling features. In the case of SOC, the own web of interactions will be modified and acquire a scaling structure, there will be scaling avalanches of extinctions, and their disposition across time may also be scaling. There is empirical evidence of all of these features (see Solé et al. 1997). On the other hand, there could also be scaling avalanches of extinctions without need of SOC. The areas affected by some external disturbances (earthquakes, volcanoes, colliding asteroids, etc.) may have a power law distribution and produce a power law distribution in the number of affected species. Amazingly, Newman (1997) showed that not even this condition is required for external disturbances to produce scaling avalanches of extinctions. He presented a model in which these can be generated by "smooth" environmental fluctuations (Gaussian or other with exponential tails, besides those with power law tails) without need of connections between species. The rationale behind it is that any infrequent event, regardless of its magnitude, leaves enough time between occurrence and occurrence for the development of many species unable to survive it. Both in Newman's model and in a SOC model, spontaneous mass extinctions could be seen as a part of an adaptative dynamics, at a large time scale, which complements gradual adaptation by removing species and (in the case of SOC) sets of interconnected species that are not robust enough in front of anomalous situations, and ensures the robustness of the system as a whole (Marín and Solé 1999 even designed an optimization algorithm based on the model by Solé and Manrubia 1997).

In the view of Newman's findings, the main difference between SOC and non-SOC webs will not necessarily be their statistical distribution of extinctions. For practical purposes, the main difference will lie in the consequences for the rest of the system of eliminating or adding an arbitrary species, which is not a negligible issue. In a SOC web, the number of extinctions produced by such a modification will have a power law distribution. I. e., in the majority of cases it will have no appreciable consequence, but in a few cases it will produce major breakdowns, and all the gradation of possible effects will be captured by a simple mathematical rule. As a corollary of the developments in chapt. 3, we can further assert that environmental stress will affect the parameters of the power law distribution in a predictable manner. Instead, in simple non-SOC webs generated with the above criteria, an extinction or an addition will never unleash a large avalanche of extinctions. If we have a non-SOC web with a non-trivial design, no generalization is possible.

What if the species that we eliminate is not “arbitrary”? It has been suggested that selective protection should be given to keystone species (see e. g. Bond 1993, Power and Mills 1995). This is equivalent to increasing the fraction of extinctions that are not expected to produce large avalanches of other extinctions, or other major ecological changes. This is a sensible recipe, but its indirect consequences are not obvious. If the approach here discussed is correct, a selective elimination of species is equivalent to tuning the control parameter of the system, and on its turn this will modify the size distribution of future avalanches.

One consequence of extrapolating what we know from SOC in wildland fires is the expectation that the system will react in front of a high frequency of avalanches (or a modification in the contour conditions that increases their sizes) by decreasing their sizes and thus approaching again the former extinction rate. However, the time scale of the reorganization of the Biosphere will be so large for this response to be completely irrelevant in front of the massive amount of extinction avalanches initiated by humanity.

Another consequence is that, in systems with few species, such as some islands, we would pass from SOC to a quasi-periodic regime, with periods of rf growing above the critical $(rf)_c$ (which corresponds to a scenario of “hypermaturity”), interrupted by massive extinctions affecting almost all of the species. Above some n, neither rf nor the sizes of extinctions will depend anymore on n (if the extinction dynamics can be assumed to have reached a steady state). Therefore, the larger n the larger the global stability of the ecosystem (because the average avalanche will affect a smaller proportion of species), and the lower the connectance (which may be interpreted as a higher degree of organization, with more selective connections).

I insist that these developments concerning the hypothesis of SOC extinction dynamics do not represent an endorsement of this hypothesis, since the evidence is still weak. However, they highlight the many doors that a strong positive evidence would open, which is a good reason for further research in this field.

Despite these precautionary comments, it is remarkable that the own Charles Darwin (1859, pg. 71-74), noting that “so profound is our ignorance, and so high our presumption, that we marvel when we hear of the extinction of an organic being; and as we do not see the cause, we invoke cataclysms to desolate the world", explained extinctions by stating that "the face of nature remains uniform for long periods of time, though assuredly the merest trifle would often give the victory of one living being over another" and reported several examples of avalanches of changes propagating from species to species.

4. 5. Diversity and host-parasite dynamics

4. 5. 1. Generalities

As mentioned in sect. 4.3, the adaptation to evolving parasites (here I use the term "parasites" to lump together from pathogens to any consumer small as compared to the kind of organism it consumes) may well be one of the processes that most heavily relies on biodiversity. This gives serious reasons for concern, as agricultural diversity has reached a historical minimum and keeps falling down at full speed. It is indispensable to have some idea of the outcome to expect from this global experiment before it is complete.

The flow of information between hosts and parasites is two-way, so this is not a simple adaptative dynamics. The literature offers many suggestive models of coevolution between them (see Drossel 2001). Although both hosts and parasites use their own diversities as a "weapon", diversity as a whole seems to benefit hosts more than parasites, because this decreases the "reactivity" by a "dilution effect". Evidence of dilution effect range from cultivated plants (e. g. Garret and Mundt 1999) to wild vertebrates (Schmidt and Ostfeld 2001). The use of diversity as a tool against pests and pathogens is well established in traditional agriculture all over the world (UNEP 1995), and its effectiveness is widely accepted by modern phytopathologists. See Garret and Mundt (1999) for a review of theoretical and experimental evidence, and Wilkes (1992) about practical issues at a global level.

The dilution effect is one feature of the relation between structure and dynamics of interaction webs that was not treated in sect. 4.4, because this was centered upon the simple but unrealistic case of random webs like May's. In random webs, the proportions between the biomasses of different groups of species cannot be expected to change in any definite way as a function of diversity or connectance, because there are no such groups of species. Instead, when there are well defined trophic levels, as is the case for terrestrial plants vs. their parasites, a low connectance should as a principle be detrimental for consumers. This can be expected even from models as simple as Lotka-Volterra's (Eq. 4.12) (things become more complex when space is taken into account, which is the central theme of this section). Since parasites often have quite a narrow specificity, increases in diversity generally lead to decreases in connectance and enhance the dilution effect. The weakening of the interactions between plants and parasites will prevent the oscillations that these produce and hence can be presumed to increase stability in aggregate ecosystem measures.

The effectiveness of diversity against parasites is not only attributed to the direct decrease in their propagation by dilution effect, but also to less evident mechanisms such as (i) plants' development of immunity against certain kinds of pathogens thanks to the contact with varieties of these pathogens that only infect other varieties of plants, (ii) a slower adaptation of pathogens to a mixture than to a pure strand, (iii) effects of particular plant combinations on microclimate or (iv) on the populations of the consumers of some insect pests (see e. g. Zhu et al. (2000) and Wolfe (2000) for the first three points, and Andow (1991) for the fourth).

In subsect. 4.5.2. I briefly comment the mainstream models on the effects of plant diversity on the spread of pathogens, and propose some modifications related to spatial effects. In subsect. 4.5.3. I discuss the conditions for self-organized criticality in host-parasite dynamics and the implications that this will have if found to hold in some case. In subsect. 4.5.4. I explore the possible consequences for agriculture of non-trivial spatial effects in the relation between diversity and the dynamics of pests and pathogens, and outline a tentative proposal for reorganizing agricultural production.

4. 5. 2. Diversity effects on the spread of plant pathogens

The decrease in the propagation of pathogens by dilution effect is the most widely studied mechanism by which diversity affects plant pathogens, and it may well be the most important one. As above mentioned, a review of current models and experimental evidence can be found in Garret and Mundt (1999). There is an increasing recognition of the importance of taking space into account for the analysis of the relation between plant diversity and pathogen spread. Unfortunately, mainstream models are completely deterministic, thus leaving apart issues as fundamental as the possible existence of critical thresholds in pathogen propagation. Stochastic models in which such effects arise have also been used in phytopathology (see Shaw (1994) for a review, and more recent contributions such as Gibson (1997) and Newman et al. (2002)), but these are the exception rather than the rule and, to my knowledge, have not been applied to study the role of diversity.

Current models of host-diversity effect on plant disease have their point of departure on Leonard (1969), who just postulated an exponential increase of the infected host tissue at a rate proportional to the proportion of susceptible tissue. More recent models (Garret and Mundt 1999) allow to refine this result by taking into account (i) varying degrees of susceptibility for different plant varieties, (ii) the effects of the relative position of the different kinds of tissue in a 2-D lattice, using realistic models of average spore dispersion as a function of distance, and (iii) local saturation. These features make the models highly realistic in many respects, but they still lack one fundamental feature: stochasticity in spore dispersion and local infection growth. By construction, in these models any infection is ensured to eventually affect the totality of susceptible tissue, and the only thing that can be modified is the timing. Therefore, they may offer a reasonable predictability for the rate of growth of an epidemic, but only under the assumption that the growth proceeds. They may even give a credible portrait of the complete development of an epidemic in the particular cases in which this can be safely expected to either grow until saturation or until halted by some change in the contour conditions, e. g. due to seasonality or to the application of some treatment, but not otherwise.

The inclusion of stochasticity in the above models would allow for the chance of epidemics stopping “by themselves”, and would improve appreciably the predictions about their incidences (“incidence” is the spatial extension of the epidemic, in phytopathological terms). The qualitative features of these predictions can be clearly anticipated (see sects.) in the particular case in which (i) epidemics have a single initial focus, (ii) we can ignore the changes in contour conditions (which may be reasonable for epidemics whose duration is not a direct function of seasonality), (iii) we can ignore finite size effects, and (iv) death host tissue is not replaced by new susceptible tissue before the epidemic is over. In this case, the probability of an epidemic stopping “by itself” will become unity above some critical host diversity. The statistical distribution of the incidences of epidemics that stop by themselves will agree with Eq. 3.2, except at the limit of small epidemics or for some non-trivial spatial arrangements of the different kinds of host tissue. Both the parameters b and l in Eq. 3.2 will depend on several features of the model. Diversity will have a direct effect on l, and may or may not affect b, depending on its spatial pattern (see subsect. 3.4.3).

Not only in the above narrow case will critical thresholds matter. These could be important even for endemic pathogens (in epidemiology, unlike biogeography, “endemic” means non-epidemic), as suggested by the overwhelming importance of habitat continuity for species persistence in other kinds of organisms (Kareiva and Wennergren 1995, With and King 1997, Bascompte and Solé 1996). At some critical host diversity, there will be a qualitative change in the degree of fragmentation of the habitat of the pathogen (see sect. 3.2.1, 3.4.3).

One conclusion from the use of mainstream models (Garret and Mundt 1999), supported by experiments, is that the effect of diversity on plant diseases will be enhanced for (i) small size of the minimal units with a given genotype (individual plants in a field, monocultures in a landscape), (ii) shallow dispersal gradient of the spores of the pathogen, and (iii) small lesion sizes. All of these features can be assimilated to a small grain size as compared to the typical distance traveled by the spores, which on its turn is equivalent to a small spatial correlation in the location of susceptible tissues at some given scales. The enhancing of propagation due to correlation agrees with more general knowledge about propagation phenomena (see subsect. 3.4.3). When the spatial arrangement of host tissue is scaling, a higher set of correlations at all scales corresponds to a smaller fractal dimension. As noted in subsect. 3.4.3, in this case the enhanced propagation translates to a smoother critical transition. For a small fractal dimension, even quite above the critical host diversity there will be appreciable epidemics, but also an appreciable effect of diversity on incidence, unlike the case of more homogeneous mixtures.

Another outstanding conclusion from mainstream models (Mundt and Brophy 1988, Garret and Mundt 1999), also with experimental support, is that host diversification over huge extensions will be more effective than the diversification in a single field. This is attributed to the direct effect of spores that travel long distances (and to a possible increase of velocity of epidemic fronts across time due to air turbulence, which would amplify local effects). There are quite stronger reasons to expect this effect is we take stochasticity into account. As apparent from chapt. 3, in this case there will be a clear decoupling between local and global effects. When close to the critical propagation threshold (not necessarily so if there are fractal correlations), a small change in local propagation capacity will produce a huge change in the incidences of the epidemics.

The consequences for management of all of these theoretical developments are treated in subsect. 4. 5. 4.

4. 5. 3. Critical pests and diseases?

Rhodes and Anderson (1996) found some evidence of a power law distribution in the size of measles epidemics in the Feroe islands and concluded that epidemics in islands can be explained by Drossel and Schwabl (1992) self-organized critical (SOC) forest fire model (see subsect. 3.2.2), with the 2-D lattice in which fires take place replaced by a "social network", and lightings replaced by unpredictable inputs of viruses from continental areas where these are endemic. Grenfell et al. (2002) suggested this approach to be also valid for small towns in the continent, besides islands. In the cases in which epidemics or pest outbreaks can really be compared to forest fires, a huge predictability will be gained from the developments in chapt. 3. Unlike the more general developments in the above section, in this case it will not be necessary to specify the arrangement of hosts, because this will be a predictable outcome of the own dynamics. The direct application of the conclusions from chapt. 3 will allow to easily predict the complete statistical distribution of incidences as a function of environmental variations, which may be useful for application to epidemics (and outbreaks) affecting either humans, cultivated plants, cattle, or wildlife.

However, this approach is only valid under very restrictive assumptions. Host-parasite interaction could lead to many different dynamics, potentially very complex, let alone when there are more elements involved, such as consumers of the pest or intermediate hosts. The main requirement to suspect SOC in some given host-parasite system are: (i) the population dynamics of the parasites should be characterized by epidemics or outbursts with one or a few initial foci, (ii) the spatio-temporal location of the initial foci should not be directly determined by earlier epidemics or outbursts, in such a way that each epidemic or outburst can be really be treated as a distinct event, (iii) in each affected site there should be a refractory period (due to loss or immunization of susceptible hosts or host tissue) of high duration as compared to the duration of the epidemic or outburst, (iv) epidemics or outbreaks should not be typically halted by seasonal changes or by finite size effects, (v) additional regulation loops, such as consumers of the parasite, should either be negligible or have some non-arbitrary features, and (vi) the arrangement of susceptible hosts or host tissue (either in explicit space, in the case of sessile organisms, or in some effective space such as the hypothetical Rhodes and Anderson's "social network") should be highly dependent upon previous epidemics or outbursts.

There are several mechanisms that could allow to fulfill the second point above, besides Rhodes and Anderson's island effect. Parasites (mainly in the form of spores) may spend long periods before having contact with any susceptible host, and sometimes are transported over long distances, even at a global scale (e. g. Campbell and Madden 1990, Brown and Hovmøller 2002). These may also have bistable population dynamics: some parasites could ordinarily oscillate around a low-population attractor while, when surpassing a given threshold by chance, be pushed to a high-population attractor by mechanisms such as overwhelming the defensive capacity of the hosts or escaping from predators (). *Vibrio cholerae* bacteria have recently been observed to undergo a dramatic change in their infective capacity once they have colonized a human subject (Merrell et al. 2002) (M. A. Rodríguez-Arias, pers. com., suggested that my wildland fire model in chapt. 3 could explain the climate-driven oscillations in cholera found by Pascual et al. 2000; the main drawback to this hypothesis was the lack of a clear equivalent to "lightings" initiating the epidemics, but the finding by Merrell et al. fills this gap). There could also be parasites that ordinarily persist in small contingents in sparse focuses and only become epidemic in infrequent environmental situations (in the terms of the model in chapt. 3, this corresponds to a very low frequency of re such that $r_e + r_\infty > r_c$, but larger than the time required for an "empty" cell to attain r_∞).

The refractory period after an infection should not necessarily be determined at an individual scale: it could also be a property of the local species assemblage. Hypothetically, the competitive features of some plant communities could be such that, in the absence of parasites, one or a few species would gradually become so abundant and widespread that, if some parasite of one of these species were then introduced at some point, its descendents would be able to propagate indefinitely. Competition would then be the driving force allowing self-organized criticality to develop, in the form of epidemics or outbreaks, which would thus actively maintain diversity. Take e. g. the case of the spruce budworm (*Choristoneura fumiferana*) in the spruce-fir forests of eastern Canada. According to Holling (1973), balsam fir gradually outcompetes spruce and white birch across succession. However, spruce budworm outbreaks produce a major destruction of balsam fir in mature forests, do not affect spruce so much, and have no effect over birch. In immature stands, not even fir is as affected as in the mature ones. This does not mean that spruce budworm has a SOC dynamics, because to propose this we should at the very least analyze whether the other above conditions are also met. However, this provides a fine illustration of a mechanism based on diversity that could lead to SOC. SOC outbreaks could be integrated with SOC fires in coniferous forests, just like natural SOC fire dynamics in fact integrates fire events under different environmental conditions, and hence with different parameters (see chapt. 3).

In this model, the larger the number of species or varieties, and the smaller their differences in fitness in the absence of parasites, the weaker will be the driving force (which is equivalent to r_i in chapt. 3; also r_∞ will be smaller the larger is the asymptotic diversity in the absence of parasites) and, thus, the smaller will be the epidemics or outbreaks. Hence, diversity will stabilize. Furthermore, the biomass will be much higher with diversity than in models with a single host and a single parasite: in the forest fire model in chapt. 3, at the point in which we attain parameter values compatible with the onset of a SOC dynamics, there is a drastic reduction in biomass (see Fig. 3.9B), but in this other case this does not hold anymore, because the sites left by one species are readily occupied by other that are not susceptible to the same parasites (but are otherwise less competitive).

This adds one more option to the series of models used to illustrate the idea that parasite-host interactions may actively promote diversity (e. g.). Janzen (1970) suggested that the diversity of tropical rainforests is maintained by the interaction between plants and insects. In particular, he thought of a continuous interaction between both, which would favor a non-contagious arrangement of trees of different species in space. Instead, He et al. (1997) found a contagious arrangement in a tropical rainforest in Pasoh (Malaysia). If we have SOC outbreaks (or epidemics) instead of (or in addition to) a continuous interaction, we keep the essentials of Janzen’s hypothesis, but we should expect a contagious arrangement, which would agree with empirical observations.

In chapt. 3, I suggested that tropical rainforests have a fire regime qualitatively different from other ecosystems. In the humid tropics, heterotrophic respiration takes the biogeochemical role of fire, and biological interactions could also replace fire as the key factor for spatio-temporal dynamics. It is tempting to think that this is what explains the huge difference in diversity levels between tropical rainforests and any other terrestrial ecosystem, added to the lack of a seasonal interruption of epidemics and pest outbreaks at low latitudes (about this last point as perceived from the agricultural experience, see e. g. Wilkes 1992). However, in the absence of detailed data analyses, this is just one more hypothesis to add to the many that can be found in the literature (see e. g. Gaston 2000).

4.5.4. Speculations on the management of pests and diseases in agriculture

The suggestions in subsect. 4.5.4. could be applied with relative easiness and would perhaps render some significant improvement in predicting the spread of plant pathogens.

The model in last section could hypothetically be extended to the agricultural case. Competition between plant varieties and species would be mediated by farmers, their “fitness” would largely depend upon their market profitability, and the variations in local susceptibility would also depend on other methods applied by the farmer besides the choice of plant varieties and species. However, the model could potentially remain the same. If this model is found to reflect reality, predictability will become quite larger and easier, taking advantage of the developments in chapt. 3. However, this does not necessarily have to be the case.

While the above particular model requires stringent assumptions, quite weaker conditions would suffice to expect a steep response of the global dynamics of pests and pathogens to local conditions, as follows from subsect. 4.5.4. If this is the case, there will be a clear dissociation between the optimum level of pest and disease control from the point of view of each individual farmer and the global optimum. Changes in either the average intensity of parasite suppression by the farmers in a given region or in any other factor affecting the average local propagation of pests or pathogens (such as development of pesticide resistance by pests) will have global effects that will modify the cost-benefit terms for each single farmer, leading them to act in such a way that the incidence of pests and diseases will remain quite stable (like the rate of burning remains stable in the long term in SOC forests, see chapt. 3). This is a risky hypothesis that remains to be tested, but could explain the striking constancy in the fraction of agricultural production lost to pests and diseases, despite huge changes in the factors supposed to be important for their dynamics. Pimentel et al. (1991) signal that, from 1945 to 1985, the use of synthetic pesticides in the U. S. grew 33-fold and their toxicity increased at least ten-fold. Meanwhile, the crop losses due to insects increased from 7% in the forties to 13% in the fifties and remained constant for the following decades. The losses due to diseases increased from 10.5% in the forties to 12% in the fifties and remained constant for the following decades. May (1977) noted that the current fraction of crops lost to pests and diseases does not seem to differ much from that in Mediaeval Europe. Many things changed while the amount x toxicity of pesticides increased by more than 3.10^4%: there has been a huge reduction in plant diversity favoring just a few profitable varieties, together with the loss of other cultural practices of pest and disease control, the increase in pest resistance to pesticides, the loss of natural predators, and more stringent “cosmetic” demands on agricultural products (Pimentel et al. 1991). The precise compensation in the effects of all of these factors could be due to the regulation mechanism above hypothesized. If this is the case, the loss of diversity in favor of just a few specially profitable varieties once pesticides are applied will be comparable to the increase in fuel accumulation expected as a consequence of fire suppression (subsect. 3.5.2). One result is an increase in productivity measured in market terms, partially due to an increase in yields, and partially because of the use of less labor-intensive techniques and a stronger adjustment to market preferences. Other results are the huge environmental damages produced by pesticides and an irreversible loss of plant biodiversity at the same time that pests are readily acquiring resistance (see Lean et al. 1990).

Biodiversity loss is indeed enhanced by the acceleration in the diffusion of a few highly profitable varieties due to globalization, the acceleration in their development due to genetic engineering, and the homogenizing pressures of the global market, which have not been accompanied by a comparable conservation effort. Furthermore, germplasm as usually kept in genebanks has a limited usefulness on practice (Wilkes 1992). Not all of the effects of globalization are however so trivial. Plant breeding increasingly relies on a globalized gene pool, which increases the effectiveness in the use of biodiversity. However, if the above hypothesis holds, the result could be a further discouragement of biodiversity conservation, rather than a decrease in the incidences of pests and diseases.

As noted in subsect. 4.5.4., the phytopathological literature offers some arguments to expect a stronger effect of treatments if applied over large areas, although these will be secondary arguments if the above model is found to hold. Whichever is the right model, the effectiveness of diversification methods applied at a large spatial scale is supported by the brilliant results of an experiment recently carried out in Yunnan Province (China), one of the most productive areas in the world (Zhu 2000; see also Wolfe 2000). This involved thousands of rice farmers. In the view of the results, they voluntarily ceased to use fungicides.

If there is really a strong dissociation between local and global optima, the coordination among farmers could potentially allow for significant increases in yields while leaving pesticides apart, by using diversity more effectively than traditional agriculture (which is likely to be driven by local optimization as much as modern conventional agriculture). The coordination should serve to increase diversity and other cultural controls above the levels locally perceived as optimal, and generate appropriate spatial patterns at all scales. As a principle, the arrangements that best allow to take advantage of differential resistance can be expected to be those enhancing antipersistence at all scales (Wilkes 1992 already proposed to disrupt the genetic continuity of cultures between different countries). Varieties or species with a high absolute resistance may be better used in the form of belts fragmenting more susceptible plantations. Interweaved natural areas can also be expected to be of benefit (not only by their contribution to cutting ways of propagation of pests and pathogens, but also by their other well known “services” such as lodging predators and pollinators, protection from erosive agents, etc.). Global screening of resistant varieties as a response to pest outbreaks and epidemics would be widely used, and in this context would not lead to biodiversity loss.

This strategy would perhaps allow to abandon or strongly reduce pesticide use while increasing productivity, but is not necessarily justified where and when high productivity is not really necessary. Before applying it in this last situation, we should study if it has some drawback for the part of wildlife that takes advantage from agricultural fields, as compared to more traditional forms of low-input agriculture.

While the above spatial arrangement minimizes the propagation of pests and pathogens, it also makes more abrupt the response to other factors affecting it (see subsect. 3.4.3). This has worth considering consequences. There should be a thorough monitoring of these other factors to avoid the approximation to dangerous thresholds, because otherwise we would have no warning before a general breakdown. Special attention should be paid to coevolution. The evolution of pathogens is thought to be slowed down by host diversification (Garrett and Mundt 1999). However, if the suppression of pests and pathogens is really effective, selection pressures will be stronger on parasites than on hosts and the firsts could have a faster evolution. If this is found to be the case, the appropriate preventions should be taken to avoid parasites to reach a final “victory”. This would be a particular form of a problem shared with modern conventional agriculture: dependence upon biodiversity combined with inability to generate it. This problem of conventional agriculture has lead to the suggestion of in situ conservation of biodiversity in areas devoted to traditional agriculture (e. g. Altieri and Merrick 1987), which is a sensible proposal.

Therefore, an optimized agricultural organization (if a high production is required) could require a “division of labor” between some large regions with diversity patterns explicitly designed at all scales to minimize the incidence of pests and diseases, and other regions devoted to traditional agriculture, both sets covering similar environmental and biogeographical ranges. Of course, regions currently rich in landraces would be the best suited to integrate the second block. The lower yields in this second block would be the price to pay to go ahead with the coevolutionary history between plants and parasites, which would continuously produce new plant varieties to be used by the other regions. Fair compensation mechanisms should be established.

These proposals are highly speculative, because their scientific basis is weak. However, I considered convenient to state them because we have the tools to test the underlying hypotheses (largely the tools deployed in chapt. 3), and, if found to be correct, will be useful for choices that have an extraordinary practical importance.

No such strategy will however suffice by itself to avoid hunger. Nowadays, the cause of hunger is inequality, not lack of food production. It is difficult to conceive the elimination of hunger without stopping and reversing the current trend of increasing inequality. However, food production per capita is likely to become limiting in a near future, due to sustainability problems in production on the one hand (Brown 1997) and demographic growth on the other. Any strategy for increasing productivity can only be useful as a way to gain time to correct these problems. If it is instead used to retard their correction, it could do more harm than good.

4.6. Conclusions

A major component of this chapter consists of modifications and remarks on earlier results by other authors. Some of my main findings in this respect are:

- Doak et al. (1998) argument for an "averaging effect" as a mechanism for a stabilizing effect of diversity becomes more robust if diversity is introduced as decorrelation between sets of organisms than if it is introduced as number of sets of organisms.
- Hughes & Roughgarden (2000) claim for a destabilizing effect of diversity in a model of a simple competitive community stands on implicit unrealistic premises. Once these are corrected, the opposite result is found.
- Pfisterer and Schmid (2002) empirical results indicating an increase in variability with increasing diversity do not contradict the expectations from simple models (as implied by the comments by these same authors, and also by Naeem 2002 and Cameron 2002), because they observe such increase only in terms of absolute variability, not of relative variability.
- Arguments for a stabilizing effect of diversity on aggregate ecosystem biomass can be extended to models more complex than the simple competitive communities that have been modeled until now. However, this only holds at the limit of a high diversity, so this cannot be tested with current designs for experiments on complex ecosystems (this problem does not affect experiments with competitive communities, which can be expected to have a well-defined response at any value of diversity).

- MacGrady-Steed et al. (1997) experimental results seem to indicate at first view that diversity has a monotonous effect on stability also in complex ecosystems, which is at variance with my theoretical expectations. However, the reanalysis of their data shows that in fact they do agree with my expectations, although they cannot be used to proof them.
- I synthesize interaction matrices of dynamically stable random complex ecosystems not satisfying May (1972) criterion, and I give evidence that they do not have less structural stability that those satisfying it, in contrast to May expectations.
- I show that May (1972) criterion depends on a set of nontrivial assumptions, rather than being a general result for all kinds of conceivable systems. For ecosystems, the hypothesis of a decreasing species-level stability with increasing number of connections per species still seems the most realistic, although there is no reason to expect a threshold in the response.
- I dissect the views on web complexity and stability by May (1972, 1973) on the one hand and by MacArthur (1955) and Odum (1953) on the other. I uncover the source of their contradictions and synthesize their points of view into a single proposal. I expect a high number of connections per species to produce wider but (asymptotically) slower oscillations in the biomass of each single species, and it is plausible that it also makes the system more robust in front of structural changes. The more recent findings by McCann et al. (1998) can be embedded as a particular case within this framework.
- I put forward some proposals to increase the detail in the models of self-organized critical extinction dynamics. This is highly speculative, but produces a large set of testable predictions.
- I show how current models of diversity effects on plant pathogen propagation can be improved by including stochastic effects. I discuss the possibility of critical thresholds and the consequences these would have. On this basis, I put forward a tentative proposal for the management of plant diseases and pests.

The issue of diversity and stability is still full of uncertainty. Nevertheless, there is quite a well-defined picture that seems to emerge from the above developments:

1) Diversity increases the stability at the aggregate ecosystem level (not only of simple competitive communities, but also of more complex ecosystem). This seems to hold at least for four facets of stability (more speculative links with other facets of stability are mentioned in subsect. 4.4.4):
    a) At the limit of low diversity, ecosystems may present completely different dynamics, both in quantitative and in qualitative terms, and may change from one to another as a result of disturbances. At the limit of high diversity, the main features of ecosystems become relatively predictable.
    b) At the limit of high diversity, there is a negative relation between diversity and aggregate biomass variability, as a result of large number effects. In the case of simple competitive communities, this holds for the whole range of possible diversities.
    c) Diversity is necessary for adaptability and the own persistence of ecosystems as complex systems, regardless of the changes in their details. Diversity enhances adaptability in two different ways:
        i) By providing the raw material for adaptation, which consist of prefiltered information. This is the most obvious point that makes diversity indispensable (at least some amount of it).
        ii) By favoring criticality in the way signaled in point 2 below. This enables ecosystems to respond in front of external fluctuations.
    d) In communities integrated by hosts and parasites with a high specificity, diversity decreases the propagation of parasites. Besides its interest for other reasons, this also contributes to increasing global stability.
2) When diversity is associated to niche overlapping, it decreases stability at the species level (or at the level of genotypes within species) by weakening regulation loops. This is an elementary path to criticality.
3) A multiplicity of interactions between species weakens both the noise and the regulation loops that act upon each of them. Therefore, it has contrasting effects upon different facets of stability:
    a) The biomass of each species will have wider oscillations (except for particular interaction patterns).
    b) The oscillations will however slow down, at the limit of a weak regulation.
    c) The system is likely to become more robust in front of structural changes.

4) Diversity could be involved with criticality or self-organized criticality (SOC) at least in three respects:
    a) An increase in diversity makes the system approach criticality by the elementary mechanism mentioned in point 3 above, which is unrelated to SOC. This provides the simplest explanation for the scaling features of diversity patterns treated in chapt. 5.
    b) In communities integrated by hosts and parasites with a high specificity, the amount and spatial arrangement of diversity could hypothetically display a SOC dynamics.
    c) There could hypothetically be a SOC extinction dynamics. If this is found to hold, it will give the keys to deal with keystone species and other factors with nontrivial effects upon the rates of extinction.

This theoretical set supports as a whole the idea that diversity increases the stability of ecological systems at an aggregate level. If this is correct, ecosystem function is going to be more variable as a consequence of the anthropogenic loss of diversity, and this will have a detrimental effect on the global environment and the many benefits that humanity obtains from them. Unfortunately, at this stage we are unable to perform quantitative predictions on the magnitude of these effects.

On the other hand, natural ecosystems are not systems assembled at random, but the result of a long history (as stressed by J. Flos, pers. comm.). In contrast, the anthropogenic decrease in diversity is the result of a set of recent disturbances, so the resulting composition of the systems will have qualitative differences with the original one that cannot be completely captured by any unidimensional measure of diversity.

In particular, it should be stressed that not only species suppressions but also species additions are disturbances that may have a disruptive effect on ecosystems. This is captured by the SOC model in subsect. 4.4.4 (and App. 4D), which, unlike other models (and experimental settings), takes into account the historical nature of ecosystems. This model is still highly speculative, but the main conclusions coincide qualitatively with empirical observations. These are that both suppressions and additions will have very small effects in the great majority of cases, and very large effects in a few cases.

Despite these last points, diversity decrease can be expected to have an outstanding impact by itself.

One of the most dramatic effects could be upon agricultural production, whose biodiversity basis is particularly eroded. We do not even know at which point we should fear a general breakdown of the food production system due to some pandemic or some other fluctuation, and it is time to study this issue seriously. I consider of foremost importance to test the hypotheses in sect. 4.5, because they would give specific guidelines to restructure agricultural production to ensure food security.

We do not know what is going to happen as a result of the current anthropogenic mass extinction, so we are in a situation of high risk. We should make a high effort both to acquire the capacity to predict its effects and, most importantly, to stop it. Both things are extremely urgent and none of them is costly. The required research is very cheap as compared to the magnitude of the potential losses by this problem, and will be much more cheap and efficient if the experiments are orientated by a thorough previous theoretical work. Up to some level, conservation of biodiversity is more than cheap, because it gives so many benefits, aside from the presumed stabilizing function, that it would completely justified even in the improbable case that the last were finally found not to be so relevant. The obstacle to the conservation of biodiversity is not a lack of benefits for humanity, but the lack of awareness of these benefits and the structural obstacles posed by our current economic system. Every effort should be done to surmount these obstacles and mobilize the required resources to stop the mass extinction. The costs should be shared in a fair way, taking into account both the capacity of each human group to contribute and its part of responsibility in the problem, which are correlated.

Appendices

Appendix 4A

Chaos in ecology? A method of detection

In spite of the simplicity of their premises, deterministic dynamic systems with just a few variables can display a wide array of qualitatively different types of dynamics. Between others, these include point attractors, limit cycles and deterministic chaos (Solé and Manrubia 1996, Schuster 1988, Wiggins 1990; about chaos in ecology, Flos 1995, Hastings et al. 1993). Deterministic chaos (or just "chaos") is analogous to noise: despite its deterministic nature, chaotic systems are inherently unpredictable beyond some time window. In subsect. 4.2.3 I make reference to this variety of possible dynamics in relation to the issue of diversity and stability (and so do, in a different way, MacCann et al. 1998). The type of dynamics is indeed one of the most essential features to know about any system in which we are interested.

There are some instances of chaos in ecology relatively well established (Tilman and Wedin 1991, Hanski et al. 1993, Constantino et al. 1997). However, in most cases the amount and quality of ecological data does not suffice to apply the standard methods to uncover a chaotic dynamics. These methods have been established to be applied to simple physical systems in which it is much easier to obtain abundant and reliable data. I worked for some time with the purpose of filling this gap, and developed a method that I explain and justify in detail in Pueyo (1997). Here I just summarize it briefly.

The unpredictability of chaotic systems is due to the fact that nearby trajectories, even if they are arbitrarily close, diverge exponentially, but never leave a bounded region of the phase space ("phase space" is defined in App. 4B). The rate of exponential separation l is called Lyapunov exponent. One of the standard methods to analyse whether a time series has been produced by a chaotic system was provided by Wolf et al. (1985) and consists of a procedure to estimate l in order to know whether or not it is positive, as would correspond to chaos. The first step is the obtainment of the time-delayed coordinates from the time series, as explained in App. 4B. Then Wolf et al. select couples i of data (in time-delayed coordinates) that are close to one another (in the phase space) with a given criterion, and register their initial distance D_0^i and their distance after t time steps, D_t^i. The estimator of the Lyapunov exponent that these authors use is:

$$\hat{l} = \frac{1}{Nt}\sum_{i=1}^{N}\log\left(\frac{D_t^i}{D_0^i}\right) \qquad (4A1).$$

Unfortunately, this method does not allow to distinguish chaos from noise. Since the couples of data are expressly chosen for their proximity, in a random series D_0^i will be smaller than D_t^i more often than not, and l will appear to be positive when applying Eq. 4A1. Furthermore, no statistical criterion is given to reject the hypothesis of $l \le 0$. This is indeed a problem for ecological time series, which, in most cases, are not only noisy, but also short.

My method in Pueyo (1997) solves these two problems. Take ΔD=Dt-D_0. If the trajectories diverge exponentially, we will have

$$\Delta D \approx \left(l\, t \right) D_0$$

Therefore, in chaotic systems there will be a positive relationship between D_0 and ΔD. By contrast, in random series this relationship will be negative. Furthermore, it is easy to test statistically if the relationship is really positive by nonparametric methods.

In my paper I show that, at least for some simple simulated systems, as few as 20 data are more than enough to unambiguously distinguish chaos from noise. This suggests that this method is much more powerful than more widely known methods designed with the same purpose, such as the one by Sugihara and May (1990). It is possible to gain even more power by integrating spatial information, in a way analogous to the proposed by Solé and Bascompte (1995).

When applying the method on practice, great care should be taken with the choice of the null hypothesis. Indeed, simple noise will rarely be the most appropriate for ecological applications. Furthermore, it is recommendable to look at other traits of the time series, besides the exponential separation of trajectories. I suggest some in my paper.

Appendix 4B

Reanalysis of MacGrady-Steed et al. data from a freshwater microcosm

The experiment by McGrady-Steed et al. (1997, see additional details in McGrady-Steed and Morin 2000) seemed to support a positive diversity-stability relation, but not in the way that we should expect from the developments in subsect. 4.2.3. Here I show that this discrepancy can be solved by appropriately reanalyzing the data, which were kindly handed over by Peter J. Morin.

These authors worked with microcosms with varying species numbers of freshwater protists and small metazoans (plus bacteria). The species they used belong to different trophic levels, so the ecosystems they studied were not simple competitive communities. They introduced from 0 to 31 eukaryotic species in each flask, but posterior extinctions and uncontrolled invasions rendered time averages from 1.7 to

18 species. Their microcosms thus cover the few-species limit and may or may not reach the many-species limit. As a measure of variability, they used the standard deviation of the CO_2 flux out of the microcosms, $s(CO_2)$ (in µl per 18h).

There is something surprising with McGrady-Steed et al. results: in spite of including noncompetitive interactions and covering the few-species end of the spectrum of species richness, they found $s(CO_2)$ to decrease with species number in a nearly monotonous way, as can be seen in Fig. 4.4A. This supports a stabilizing effect of diversity, but in a different way than we should expect from subsect. 4.2.3, as a principle. From that section we can expect two effects: a pass from qualitatively different kinds of dynamics when we have few species (which makes the system fragile) to a well defined and robust dynamics at the limit of many species, and a decreasing variability in CO_2 flux with increasing diversity only when we have reached the upper region of the spectrum of diversity, not for all the spectrum.

The initial numbers of eukaryotic species n_0 were 0, 3, 5, 10, 15, 20, 25 and 31, chosen from a pool of 32 species. For n_0 from 3 to 10, four different treatments were prepared in each case, each treatment consisting of a different initial species composition. In the other cases, there was one single treatment. For each treatment there were five replicas. For each replica, the authors obtained time series with a length of six data each one, after measurements made at intervals of one week. The data were grouped in accordance to the number of eucaryotic species n in the flask and time of the measurement, and $s(CO_2)$ was calculated for each of these sets of data. In this way, the authors claimed to reach an overall measure of ecosystem variability encompassing both temporal and spatial variation.

There is one point in this setting that should be remarked. Since n is to a large extent a function of n_0, the fact that for small n_0 (from 3 to 10) there are four different treatments, while for big n_0 (15 to 31) there is only one, introduces an obvious source of variation for small n that is not introduced for large n. The below reanalysis of data indicates that this had a high impact on the result in Fig. 4.4A. Note that the change in $s(CO_2)$ from small n to large n in Fig. 4.4A takes place in a rather stepwise manner, reflecting the two different criteria that were used. Furthermore, even if there were the same number of treatments per n_0, the data with high n would present a higher similitude in species composition, because of the limited pool of species that were used.

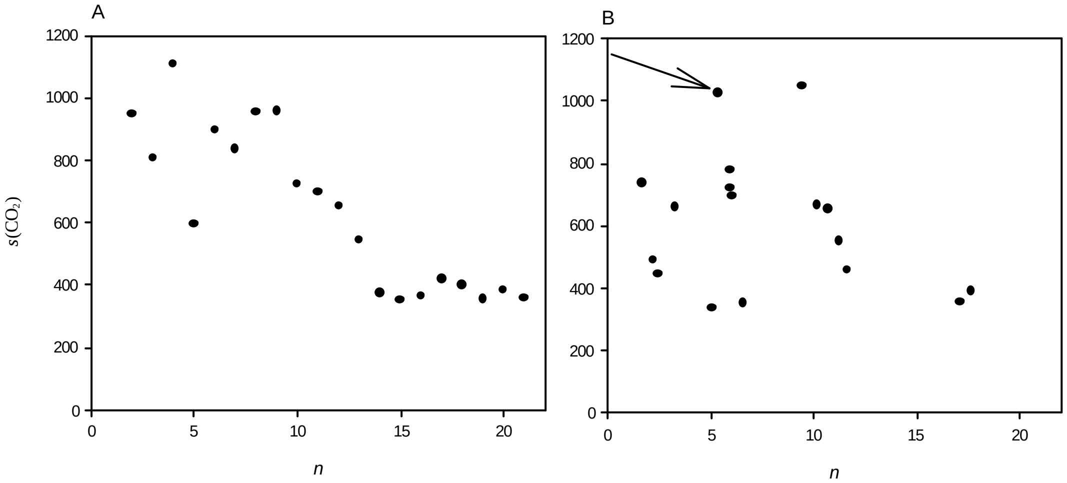


Fig. 4.4. *Relationship between species number n and standard deviation of CO_2 flux, s (CO_2), in the freshwater microcosms studied by McGrady-Steed* et al. *(1997)* **A:** *the results as presented by the authors.* **B:** *the results when extracting the effect of the change in the number of different treatments as a function of n. Fig. 4.4B agrees with theoretical expectations and, despite its appearance, is consistent with a stabilizing effect of diversity. The spot signaled by the arrow corresponds to the treatment in Fig. 4.5.*

The authors wanted to mimic a sequence of species loss across all trophic levels of an initially diverse community. Their results do indicate that, the more species we lose, the more unpredictable will the fluxes through the system be. As a principle, this can be expected for two different reasons: because the incertitude in which particular species will be lost makes the properties of the system also uncertain a priori, and because a lower diversity could produce a higher variability across time. Fig. 4.4A does not allow to disentangle the two effects, but the reanalysis of the data shows that it essentially reflects the first one. Here I analyze whether the same data allows to extract some information on the effects of diversity per se.

I calculated $s(CO_2)$ for the data grouped by treatment, and then represented it as a function of the average realized species number for each treatment, $\bar{n}$ .The result is shown in Fig. 4.4B. The monotonous relation between n and $s(CO_2)$ vanishes.

Fig. 4.4B agrees qualitatively with theoretical expectations. For few-species treatments there is a wide scatter of results, with no trend. The results for a larger number of species are compatible with an asymptotic negative relation between $\bar{n}$ and $s(CO_2)$, but this is not statistically significant. The test used to reach this result is explained at the end of the appendix.

An interesting feature of these data is that different few-species treatments really seem to give rise to qualitatively different dynamics. Note that two of the treatments produce a $s(CO_2)$ quite higher than the others. At least in one of these, this seems to be the result of low-dimensional nonlinear oscillations. This is the treatment with n_0=3, $\overline{n} = 5.4$, and label "b" in McGrady-Steed et al., which here I call 3b (following McGrady-Steed and Morin 2000). The initial species were a bacterivore microflagellate, an autotrophous chlorophyte of the genus *Chlamydomonas* and a consumer of the genus *Hypostome*.

I analyzed the data 3b by means of a well-known technique, the method of Takens (1981), which is based on the theorem of Withney. Assume that the oscillations in the flux of CO_2 depend mainly on a few variables that evolve and interact between them in a roughly deterministic way, such as e. g. the populations of 2 or 3 of the species in the microcosm. This is the same as saying that we can reasonably predict the flux of CO_2 from the state of a dynamic system that has a deterministic trajectory in a space of D dimensions, which correspond to the relevant variables. This space is called phase-space. When we perform the measurements, the system may be already close to some attractor defined in the phase-space, such as a limit cycle, a thorus or a strange (chaotic) attractor. We want to know the properties of the attractor, which is defined in D dimensions, but we only have information in one dimension, because we have measured a single variable. The solution is passing from the unidimensional time-series $f_1, f_2, ..., f_T$ to the multidimensional series $(f_1, f_2, ..., f_D)$, $(f_2, f_3, ..., f_{D+1}), ..., (f_{T-D+1}, f_{T-D+2}, ... f_T)$. Then we can represent the trajectory of the system in a graph with D-dimensions, which are called time-delayed coordinates. The trajectory in the space of time-delayed coordinates will maintain the topological properties that we would see if we could observe the original phase-space. So we will obtain a reconstruction of the attractor.

In fig. 4.5A we can see the joint representation of the trajectories of the five microcosms with the treatment 3b in time-delayed coordinates, in 2 dimensions. They display a nontrivial structure, which suggests some kind of nonlinear oscillation. This can be seen more clearly by comparing with a set of five time series obtained by reordering at random the data from the original series, in Fig. 4.5B. This structure is not observed for the other treatments, at least no so clearly. It has not been possible to better characterize the dynamics, not even by the method in Pueyo (1997), since the series are really short (6 data per series), and this can be only partially compensated by making a joint analysis of the five series with the same treatment..

This suggests several lines of advance. On the one hand, this experiment can be seen as a preliminary test to choose some appropriate species assemblage to work with nonlinear population oscillations. The assemblage of choice is indeed the used for the treatment 3b. More importantly, further experiments should be done to better test the asymptotic relation between diversity and variability. The most practical procedure must be the generation of diversity gradients from initially diverse microbial communities by dilution, a technique that has been used by Franklin et al. (2001). The results of chapter 5 give the basis to predict the variation of diversity with dilution. This experiment could be carried out in parallel with completely different microbial communities in order to have true replicas.

The possible relation between diversity and variability for asymptotically high

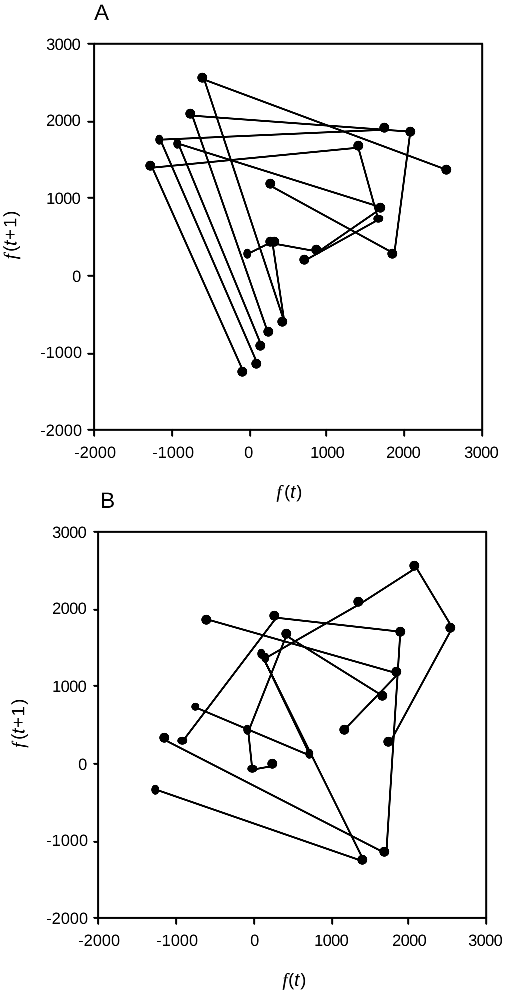


Fig. 4.5. *Nonlinear oscillations in the flux f of* CO_2 *in a set of few-species freshwater microcosms, assembled by McGrady-Steed* et al. *(1997). These are the five microcosms in which the eukaryotes initially inoculated were* Chlamydomonas, Hypostome, *and a bacterivore microflagelate.* **A:** *joint reconstruction of the trajectories of the five microcosms in time-delayed coordinates, in two dimensions.* **B:** *The same, after the values of f have been randomized both across time and between microcosms.*

diversities (such that we do not know a priori at which point we can say that we are within the “asymptotic region”) can be tested as follows: (i) take the set of numbers of species and standard deviations (π_i, s_i) for each treatment *i*; (ii) take the subset of the *m* treatments with the largest π, for each *m* from 2 or 3 up to the total number of treatments; (iii) for each *m* measure a nonparametric correlation (such as Kendall (1970) correlation) between π and s, and keep the signification level e_m that allows to reject a lack of relationship between π and s; (iv) keep the smallest of the significance levels min(e_m); (v) redistribute randomly the sets $\{\pi_i\}$ and $\{s_i\}$ by pairs (π_i, s_i), *l* times, in each case repeat all of the previous steps, and count the number of cases l_+ in which the value of min(e) is smaller or equal than for the original pairs; (vi) the global significance level that allows to reject the null hypothesis that there is no asymptotic relation is $e = \frac{l_+ + 1}{l + 1}$.

In the case of McGrady-Steed et al. data I found $e \approx 0.01$ (with l=9999; unilateral test, $e \approx 0.02$ for a bilateral test). This seems enough to accept an asymptotic relation, but a careful analysis shows that this is mainly due to the nicely ordered position of the five spots between n=9.45 and n=11.65 in Fig. 4.4B, which is obviously fortuitous, since this range of variation of n is tiny as compared to the own variation of n across time in each treatment. If we simplify the data in Fig. 4.4B by giving the same value of π to treatments in which this is similar, thus using just four values of π, we obtain $e \approx 0.08$ ($e \approx 0.16$ in a bilateral test). This does not allow to ensure that there is any asymptotic relation, but is an encouraging result, if we take into account that the procedure to obtain it has been extremely conservative.

Appendix 4C

A recipe for stability beyond May’s threshold

May (1972, 1973) used sets of matrices that had homogeneous eigenvalues when connectance was null. This is the reason why they attained in this last case the maximum dynamic stability compatible with their trace. However, it is equally possible to use matrices that have homogeneous eigenvalues in the case of full connectance C=1, whichever the value of a and n. Here I give a recipe to generate them.

We wish to generate a matrix **A** all of whose eigenvalues are l. If $\mathbf{D}_l$ is the diagonal matrix that contains the eigenvalues of **A**, this condition means that $\mathbf{D}_l = l\mathbf{I}$, where **I** is the identity matrix.

By the definition of eigenvalue, there will be some eigenvector matrix **T** satisfying

$$\mathbf{AT}=\mathbf{TD}_l \qquad (4C1).$$

The most immediate way to obtain a matrix with the eigenvalues in $\mathbf{D}_l$ is by generating a matrix **T** at random and performing the operation

$$\mathbf{A}=\mathbf{TD}_l\mathbf{T}^{-1} \qquad (4C2).$$

However, it is easy to see that in the case of homogeneous eigenvalues this operation will always produce the most trivial result, $\mathbf{A}=\mathbf{D}_l$. This means that the **T** that we require should be non-invertible.

A non-invertible **T** is easy to generate, we just have to replace one of the rows or columns by a linear combination of the rest (if we chose for the nxn elements t_{ij} of **T** a zero-centered Gaussian distribution, we can preserve the same distribution for their linear combination if its coefficients c have the distribution $c \approx N\left(0,(n-1)^{-1/2}\right)$, when calculating $t_{ij} = \sum_{k \neq i} c_{ikj} t_{kj}$ for each j and the chosen i).

Of course we cannot apply Eq. 4C2 to a non-invertible **T**. However, things are easier if we begin by searching a matrix **B** with l=0 instead of the matrix **A** with arbitrary l. It follows from Eq. 4C1 that **BT**=0. As a consequence, $\mathbf{T}^{\mathrm{T}}\mathbf{B}^{\mathrm{T}}=0$. Each of the columns of $\mathbf{B}^{\mathrm{T}}$ can thus be obtained in the form of vectors **v** by solving n times the linear system

$$\mathbf{T}^{\mathrm{T}}\mathbf{v}=0$$

Since this is an indeterminate system, for each vector we will decide a priori one of the coefficients at random. Once we have n solutions, we join them to obtain $\mathbf{B}^{\mathrm{T}}$ and thus **B**.

From **B** it is immediate to pass to **A**:

$$\mathbf{A}=\mathbf{B}+l\mathbf{I}.$$

Before this last step, it is possible to rescalate the coefficients of **B** to assign it a given a. In my experiment, I also filtered out any matrix with some value exceeding a given threshold, which seemed to generally be the result of numerical “accidents”. The eigenvalues were also verified.

Appendix 4D

Comments and variations on a self-organized critical extinction dynamics model

Subsect. 4.4.4 deals with the possibility of self-organized criticality (SOC) in extinction dynamics. One of the main models of SOC extinction dynamics in the literature is the one by Solé and Manrubia (1996). This is however different from the kind of model with which subsect. 4.4.4 deals. Here I show why it is different and how to modify it to make it closer to the proposals in that section. This does not imply a position about which of both approaches is the most realistic, if any. The purpose of this appendix is just taking advantage of Solé-Manurubia model to illustrate my own approach in subsect. 4.4.4, and making clearer the set of possible options (furthermore, I should stress that all of the developments in subsect. 4.4.4 were largely inspired by these authors’ model).

In Solé-Manrubia model, S_i defines the state of species i, for i from 0 to n. This has two possible values: S_i=0 if the species is extinct, and S_i =1 otherwise. The state of this species evolves in time t according to:

$$S_i(t+1) = \Phi\left(\sum_{j=1}^{n} g_{ij}(t) S_j(t) \right),$$

where $F(z)$=1 if z>0 and zero otherwise, and $\{g_{ij}\}$ is a set of connection terms between species. The dynamics has three steps:

1- For each species i, we chose another species j at random and replace the current value of g_{ij} by a random value between –1 and 1.

2- We update synchronously all of the species, by computing $F_i = \sum_{j=1}^{n} g_{ij}(t) S_j(t)$ for each species i, and setting $g_{kj}=g_{jk}=0\ \forall j$ for each species k such that $F_k<0$, which corresponds to its extinction.

3- One of the remaining species is chosen at random to originate all of the species that will replace those that went extinct. The new species will have connections g similar to those of its mother species, just with small modifications. If we generate a new species k from a mother species l, the new connections will have the form $g_{kj}=g_{lj}+h_{kj}$ and $g_{jk}=g_{jl}+h_{jk}$ given j, where h is a random variable with average 0 and a small standard deviation s.

Solé and Manrubia obtain a power-law distribution of the number of extinctions per time step, as characteristic for SOC dynamics. This is shown in Fig. 4.6A (full spots; with a community size of 128 species, s=0.01, a transitory of $2\text{x}10^5$ cycles, and a duration of the experiment of $5\text{x}10^5$ cycles; the regression renders $b \approx 2.5$; this differs from the value given by Solé and Manrubia, $b \approx 2$, which arises when b is estimated by regression on a conventional histogram: about the implications of each method of estimation, see App. 2A).

These avalanches of extinctions do not have the same origin as the percolation-based SOC dynamics hypothesized in subsect. 4.4.4. In that subsection, I considered the case in which each "initial" extinction produces a set of extinctions with a power-law distribution. For this distribution to be directly observable, the complete avalanche should take place in a single time step (which corresponds to a strong separation of time scales between extinction avalanches and the process of differentiation between species). Solé-Manrubia model does not share this premise, so, if there are large avalanches of propagating extinctions, it just captures a part of them in each time step, perhaps mixed with parts of other avalanches, and it is not obvious how this should translate to the observed distribution.

The power-law distribution observed from Solé-Manrubia model seems to reflect in fact the size distribution of the sets of species that are likely to get extinct at the same time because they are similar to one another. These sets may attain large sizes because of the rule of replacing all of the extinct species in a given time step by small variations on a single species. The evidence for this is the fact that, if we change this last rule and choose separately the mother species for each of the species that have to be replaced, the power law disappears and we find an exponential extinction size distribution, as shown in Fig. 4.6A (empty spots). As a side comment, let me note that this provides a nice instance of the stabilizing effect of diversity. The realistic pattern of species replacement after mass extinctions will probably lie someway between these two rules, but at this point I am interested in isolating the second one. A few small modifications on Solé-Manrubia model will suffice for this purpose:

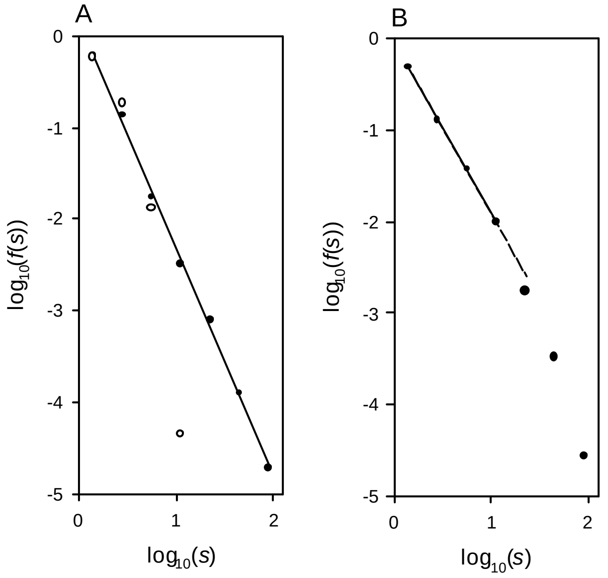


Fig. 4.6. *Outcome of two different models of extinction dynamics, by computer simulation. Density of probability f(s) of the sizes s of the extinctions that take place in one single time step, provided that s>0. The represented part of the x-axis corresponds to the total number of species in the simulated communities.* **A:** *model by Solé and Manrubia (1997).* **B:** *model proposed in App. 4D in this chapter.*

1- We chose one connection (*i*,*j*) at random and replace the current value of g_{ij} by a random value between $q-1$ and $q+1$, with $-1<q<0$ (this provides the driving force that pushes the system toward the state in which large avalanches take place).

2- We cycle through the following two steps, until there is no extinction in step 2.1:

2.1.- We update synchronously all of the species, by computing $F_i = \sum_{j=1}^{n} g_{ij}(t) S_j(t)$ for each species *i*, and setting $g_{kj}=g_{jk}=0\ \forall j$ for each species *k* such that $F_k<0$, which corresponds to its extinction. We repeat these operations until $F_i \geq 0$ for all *i*.

2.2.- We replace each extinct species *k* with a new species that will have the same connections as a mother species *l*, chosen at random separately for each *k*.

All of the extinctions that take place while not returning to step 1 sum up to a single number, excluding the extinctions of "species" that have temporarily replaced other species extinguished in the same avalanche.

This model includes a time scale separation between evolution and avalanches of extinctions. The value $q<0$ is the driving force that leads the system toward the threshold in which large avalanches take place. The size of an avalanche does not include the "species" that arise during the avalanche because these temporal replacements are interpreted as species increasing their population as a response to the extinction of other species, rather than true speciations, which would violate the separation of time scales.

This model also generates a extinction size distribution close to a power law, as apparent from Fig. 4.6B (with a community size of 128 species, q=-0.8, a transitory of 2x10^5 cycles, and a duration of the experiment of 5x10^5 cycles; there were avalanches of extinctions for 1.75% of the cycles). The regression line for the four left-most spots in the figure renders $b \approx 1.85$. The deviation for large sizes could be due to finite size effects (the upper end of the x-axis in Fig. 4.6 has been chosen to coincide with the size of the community, which allows to see the limitation that this poses).

This is still a partial illustration because it does not include the particular kind of driving force that is given more attention in subsect. 4.4.4, which is the increase in specialization across time.

Chapter 5

Scaling diversity patterns

5.1. Introduction

The relative abundance distribution of species has been claimed to be "the most fundamental measure in ecology" (Pachepsky et al. 2001). On its turn, scaling seems to be a key concept to understand abundance distributions and other traits of diversity patterns. A fair characterization of such patterns is important for establishing conservation priorities. A clear knowledge of the relation between patterns and dynamics is essential to develop a solid theory of the causes and consequences of biodiversity loss.

Scientists from many areas are currently paying a strong attention to scaling patterns. Scaling became a focus of interest with the development and wide application of fractal geometry (Mandelbrot 1983). This interest increased when self-organized criticality was identified as a credible mechanism for the generation of scaling in many kinds of systems (Bak 1996).

For the science of biological diversity, such patterns had an important role since quite earlier. One of its classical postulates is the existence of a scaling relation between area and species number. Some authors have also used scaling statistical distributions to adjust species abundance data. These authors have considered that the probability $P(n)$ that an arbitrary species is represented by n individuals in a sample approaches a power law

$$P(n) \approx xn^{-b} \qquad (5.1).$$

It is worth mentioning that Corbet (1942, quoted in Corbet 1943), who seems to have been the first to empirically quantify the relation between number of species and number of individuals in a sample of organisms (Malayan Rhopalocera butterflies), chose a power law relation. The scaling nature of power law distributions is treated in Mandelbrot (1983).

Solé and Alonso (1998) suggested that the seeming presence of power laws in abundance distributions could be a symptom of criticality.

On the other hand, the apparent finding of scaling patterns may sometimes reflect fashion rather than a property of natural systems. A survey showed that most of the studies under the heading of "fractals" in the leading group of journals *Physical Review* in 1990-1996 were restricted to about 1.3 orders of magnitude (Avnir et al. 1998). In this chapter I present evidence that this is not the case of biological diversity. I analyze one of the largest datasets in the literature with especially suitable procedures, and the result is a series of power laws spanning 3-4 orders of magnitude.

I also show that ecosystems should be considered to be close to criticality in quite a trivial sense. This involves the presence of mechanisms that inject scaling into the abundance distributions, thus explaining the empirical observations in this sense.

Indeed these are not the only mechanisms at work. I develop a global framework that allows to take into account each of the possible factors of relevance in the prediction of the eventual abundance distribution. In particular, I present some general rules that allow to ignore much of the detail that could have been considered important. Many of the additional factors can be assumed to be "well-behaved" in that either (a) leave the scaling behavior unaffected or (b) produce a deviation in a simple an predictable manner, with little dependence on the specific set of factors.

The most surprising finding is that the progressive inclusion of additional factors carries us to a tour through most of the abundance distributions that have played a central role in the history of diversity research. This is synthesized in Fig. 5.1, which is the map for our "tour".

The tour also carries us to the standard species-area relation. I show that its true link to the displayed family of abundance distributions does not take place through the lognormal, as usually assumed. Another recent outstanding proposal (Harte et al. 1999) is shown not to be either correct.

Once we accept that diversity patterns are overall scaling with well-defined deviations and we have a wide picture of the relation between patterns and dynamics, we obtain a solid basis for several applications. We can re-evaluate the works in which the observed patterns were used as evidence for specific models. We can establish which measures capture most of the information in the patterns and evaluate the performance of classical diversity indices. We can use these measures to study the impact of different factors on biodiversity and to select priority areas for conservation. All of these applications are either developed or at least outlined in this chapter.

The plan of the chapter is:

- Sect. 5.2: *Preliminary issues* reminds some elements of sampling theory and introduces the basic notation.
- Sect. 5.3: *Critical ecosystems?* establishes a basic bridge between criticality and ecosystem dynamics.
- Sect. 5.4: *A possible origin for abundance distributions and species area-relation* presents the core theoretical results, which are synthesized in Fig. 5.1. A general model is proposed for the emergence of diversity patterns from the underlying dynamics.
- Sect. 5.5: *Interpretation of abundance distributions* works out the opposite step: the induction of dynamics from patterns, in the light of the results in the previous section.
- Sect. 5.6: *Case study: marine phytoplankton* illustrates the theoretical developments with the analysis of a large set of data.
- Sect. 5.7: *Measuring diversity?* explores the practical consequences of scaling for diversity quantification, with emphasis on conservation planning.
- Sect. 5.8: *Discussion* and *Conclusions* as usual.

The chapter contains many new contributions covering a wide area, often standing on quite complex developments. It is not easy to put all of this information in a form that does not demand an excessive effort on the part of the reader. Therefore, I offer a large part of the information within a series of 12 appendices, comparable to brief papers, in such a way to keep the main text as synthetic as possible. The text thus includes many results briefly enunciated, accompanied by the reference to the appendix containing the complete developments. The last appendix is the only one untied to the main text; it is a comment on a recently proposed "synthetic theory of biodiversity" (Ritchie and Olff 1999).

5.2. Preliminary issues

This section reminds a few well-known facts about sampling theory that should be taken into account for the below developments, and introduces the basic notation. A deeper understanding of sampling can be gained from reviews such as May (1975), Engen (1978) or Pielou (1975, 1977).

I consider samples with N individuals and S species, taken from communities with M individuals and T species. Each species i has n_i representatives in the sample and m_i representatives in the community. The probability for one species of having n representatives in the sample is $P(n)$, while the number of species with n representatives is $s(n)$. The probability of having m representatives in the community is $P(m)$.

Both $\{P(n)\}$ and $\{P(m)\}$ are discrete distributions. However it is usual to use a continuous approximation to $\{P(m)\}$. I call x the continuous counterpart of m, and its probability density function (p. d. f.) is $f(x)$.

The ecological hypotheses to be analyzed will give rise to some $\{f(x)\}$, while only $\{P(n)\}$ is observable. The expected $\{P(n)\}$ is usually obtained by assuming that the sample is random, i. e. that any individual in the community has the same chance of joining it (see the cited reviews above).

It is noteworthy that the sampling distribution will not just be a blurred image of the community distribution, but also a biased image. The lower end of the community distribution is not captured by finite samples, which exclude most rare species. Preston (1948) compared this with the effect of running a veil. Note that, if a species i has a population m_i, its expected representation in the sample is

$$E(n_i|m_i) = m_i \frac{N}{M} \qquad (5.2).$$

Since n_i is discrete, a species is only expected to join the sample when $E(n_i|m_i) \geq 1$, i. e.,

$$m_i \geq \frac{M}{N} \qquad (5.3).$$

This is thus the position of Preston's veil. The veil approach is not the appropriate to obtain the sampling distribution from the community distribution (Dewdney 1998), but it is a rule of thumb useful for some of the below developments.

5.3. Critical ecosystems?

Scaling is characteristic of critical systems. Thus the observation of scaling patterns in ecosystems immediately leads to wonder whether or not they are critical as well (Solé and Alonso 1998; see also Solé et al. 2002).

One of the simplest models displaying criticality has a demographic inspiration: branching processes (e. g. Harris 1963). In an ordinary branching process, we have a set of individuals whose reproduction and death events are independent identically distributed random variables, both over time and between individuals. One of the main concerns of branching process theory is the total number of descendents left by a single individual after a time $t \to \infty$. Let r be the expected population rate of increase. If $r>0$ there is some probability P_∞ that the total progeny diverges to infinity and never becomes extinct, which increases with r. Otherwise, $P_\infty = 0$: the extinction is granted, the sooner or the later. Note that P_∞ is not derivable about r for $r=0$, but it is for any other value of r. In physical terms, this means that $r=0$ is the critical point where a second order phase transition takes place (e. g. Stauffer and Aharony 1992). A universal property of critical points in second order phase transitions is scaling. In the case of branching processes, a well-known result is that the total progeny left by an individual follows a power law distribution at $r=0$ (Harris 1963).

If the reproduction and death events are synchronized by "well-behaved" random environmental fluctuations, we still have a second order phase transition at $r=0$.

Branching process theory, even with added environmental fluctuations, is an incomplete description of population dynamics. In its simplest form, it does not include migration or age classes although these can be added (Harris 1963). Much more importantly, it does not include denso-dependence in any case. However, at least in systems with many species, it may be reasonable to expect an appreciable margin for the population of each single species in which denso-dependent forces are weak. This is an immediate corollary of the widely accepted decrease in species-level stability with increasing diversity (while the aggregate ecosystem stability increases), which is treated in sect. 4.2. Within the range of populations in which this holds, population dynamics will be well approximated by a critical branching process ($\frac{dx}{dt} \approx rx$, with $r \approx 0$).

These simple arguments render the criticality approach applicable to ecology. As commented in subsect. 4.3.2, this may help to understand the adaptability of ecosystems. The most immediate consequence of criticality, however, is that we can expect scaling features at all levels. The specific sources of scaling and the factors that may affect it are treated in sect. 5.4.

There are other not-so-trivial mechanisms that may lead a system to a state comparable to criticality. These have been recently disclosed, under the label of self-organized criticality (Bak 1996), and are treated mainly in chapt. 3, and also in sect. 4.5 and subsect. 4.4.4 in chapt. 4. It is plausible that such mechanisms have a role in the process of community assemblage and that this is also produces scaling diversity patterns. This will not be treated here, but can be easily integrated within the proposed framework.

5.4. A possible origin for abundance distributions and species-area relation

Fig. 5.1 summarizes this section. It displays a proposed set of links between several of the commonly used abundance distributions (there are several reviews available on abundance distributions, such as May (1975), Frontier (1985) or Gray (1987); see also Engen (1978)). It identifies some groups of factors that are expected either to produce a move from one distribution to another, to stay in the same distribution or to enter the circuit. It includes one path to the standard species-area distribution. In the following subsections I treat each of the elements in Fig. 5.1 and add some remarks.

5.4.1. Demographic noise

This is one of the entrances to the circuit in Fig. 5.1. It is the one that we will use to begin our trip through it.

Let us consider a simplistic model where demographic noise is the only relevant factor. I. e. each species population fluctuates in accordance with the simple branching process explained in sect. 5.3 (this is the simplest and most usual meaning of "demographic noise"; a more complex one is used by Engen et al. (1998)). In addition, there is some small constant probability that an individual belonging to a new species joins the community.

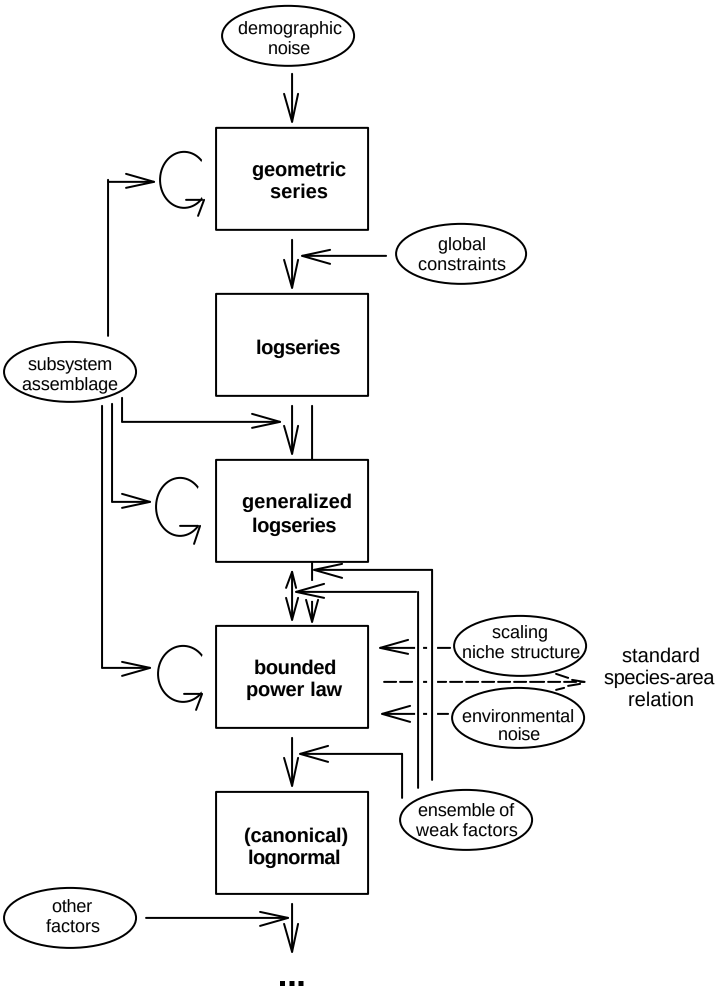


Fig. 5.1. *Synthesis of the model for the origin of abundance distributions and species-area relation. Extended explanation in sect. 5.4.*

The deduction of a steady-state statistical distribution from its generating mechanism is straightforward by applying the "diffusion" approach, which is well-known in physics. Engen and Lande (1996a, b) introduced it to the prediction of species abundance distributions. A technical point to add to their approach is put forward in App. 5A. If we apply their equations to the simple case I consider (with App. 5A in mind), we find (in continuous terms)

$$f(x) = xx^{-1} \qquad (5.4).$$

As a principle, this is unaffected by the specific distribution of reproduction and death events. The result is also the same if we have a mixture of species with different demographic features.

This distribution is a particular case of power law distribution (Eq. 5.1), with exponent b=1. As expected, from a critical dynamics we obtain a scaling outcome. (In fact, unbounded power law distributions are constrained to b>1, because their distribution function diverges at $b=1$; this demands some upper bound in our case, which is indeed biologically realistic).

Eq. 5.4 is equivalent to the geometric series distribution, which is perhaps the first species abundance distribution ever proposed. I do not expect any sample to completely adjust to Eq. 5.4, but the empirically observed distributions can be interpreted as small and senseful variations about it.

5.4.2. Global constraints

The minimum factor that should be added to the above model in the name of realism is the existence of physical limits to the total resource use that a region can sustain. This means that some species' population may have to decrease if another one is to increase. At this stage I still consider species to have similar ecological features. In this first approximation I assimilate resource use to number of individuals, though this is not completely realistic (Pagel et al. 1991, Taper and Marquet 1996).

Global constraints will modify Eq. 5.4 to some extent. The result will depend on the specific shape of the denso-dependent responses. However we may wonder if there is some "preferential" distribution, reasonably close to the outcome of most of the possible models.

If we invent a set of abundances, we can quantify its likelihood (without constraints) with Eq. 5.4. From all of the possible sets of abundances compatible with the constraints, we can thus choose which one is the most likely in what refers to Eq. 5.4. In this way we can obtain the distribution that deviates the least from Eq. 5.4 in probabilistic terms, while fulfilling the constraints. This distribution can be considered a good candidate. This approach is a transposition of the "maximum entropy" criterion of thermodynamics, which has already been used in other fields besides thermodynamics (references in App. 5B). I apply it to our case and discuss some points in App. 5B, where I also give the general equation that allows to add more constraints. The result for our minimal constraint is:

$$f(x) = xx^{-1}e^{-fx} \qquad (5.5).$$

The discrete distribution most akin to Eq. 5.5 is

$$P(n) = xn^{-1}e^{-fn} \qquad (5.6).$$

This is essentially scaling. The exponential term is just a bending function that sets the bound to scaling at the upper end of the distribution.

Eq. 5.6 happens to be the logseries distribution, which Fisher (1943) developed on empirical grounds when studying the samples of moths collected by Corbet (1943) and Williams (1943). It became one of the standard distributions in this field, though until now there has been no consensus on its ecological meaning (e. g. Death 1996).

It is usually assumed that Eq. 5.6 is the sampling distribution of Eq. 5.5, but there are some minor problems in the step from Eq. 5.5 to Eq. 5.6. A solution is given in App. 5C.

Eq. 5.6 is a good point of reference to use as a null hypothesis. Several different "neutral" models in the literature result in this distribution (Watterson 1974). In view of this coincidence, Caswell (1976) concluded that such results "seem to be inherently connected with neutrality". The above results suggest that this impression must be correct.

5.4.3. Assemblage of subsystems

The above developments refer to sets of species with highly overlapping ecological requirements. This could be a reasonable approximation within a single guild or functional group, but not for complete communities, which are assemblages of different functional groups (see Smith and Huston 1989).

Here I study the effect of joining in a single sample species from different sets with some given distribution within each set, in the case that this distribution is a bounded power law (Eq. 5.1; this includes the logseries, Eq. 5.6). This distribution has an special behavior in this respect, due to its scaling character.

The scaling character of power laws can be illustrated by multiplying *n* by a constant *k* in Eq. 5.1. We have:

$$k^b P(kn) = P(n) \qquad (5.7).$$

We only need to rescale the new probability through a constant k^b to recover the former one (of course this does not extend to the values of n close to the upper bound; neither to small *n*, where Eq. 5.1 may not hold due to the effect of sampling, if $b \neq 1$).

This makes power laws quite well-behaved in front of sampling. As we increase the size of a sample, we observe a larger portion of the underlying distribution, according to Eq. 5.3, but the scaling property (Eq. 5.7) guarantees that we still have a power law with the same exponent b. The main change takes place in the position of the bounds to scaling and in x (Eq. 5.1), which depends on them.

Take a sample made up of a set of subsamples. Consider that (a) each one contains a large number of species and (b) each one presents a power law with a similar b. A subsample i will have the distribution

$$P_i(n) \approx x_i n^{-b} \ .$$

x_i depends on both the underlying distribution and the subsample size. If the fraction of species that belong to the subsample i is r_i, the ensemble distribution is

$$P(n) = \sum_i r_i P_i(n) \approx \left(\sum_i r_i x_i \right) n^{-b} \qquad (5.8).$$

The summation above is a weighted average on x_i, which is independent of n and can be called x, so Eq. 5.8 reduces to Eq. 5.1.

This property can be called invariance under assemblage. It is reminiscent of the invariance under addition that the Central Limit Theorem postulates for Gaussian variables, but both properties should not be confused.

When the assembled subsamples have slightly different b, the result can still be well assimilated to a power law with an intermediate b. On practice, the range of b that have been found when adjusting a power law to real samples is quite small.

We can make no general prediction for the region where a significant portion of the guilds reach their bounds to scaling. If the guilds follow a logseries, the joint distribution may lose its exponential part and substitute it by some other bending function $C(n/n_M)$, whose effect will also be restricted to the upper end ($\approx n_M$):

$$P(n) = x n^{-1} C(n/n_M) \qquad (5.9).$$

This is what I call generalized logseries in Fig. 5.1.

As we pass from wide functional groups to a more "straitjacket" niche structure, with a number of well-defined niches approaching the number of species, the invariance under assemblage is lost and diversity patterns increasingly reflect the organizational patterns of ecosystems.

5.4.4. Ensemble of weak factors

In the above subsections I consider sets of factors that either produce or preserve scaling with b=1 (except at the upper bound). Of course there are many factors at work that may disrupt it. The effect of each possible factor can be predicted with standard methods (Engen and Lande 1996a, b). However, when we have a large set of possible factors and we do not know them with much detail, this does not work.

If the resulting deviation is not large as compared to sampling error, the outcome of many possible factors may become indistinguishable. We may thus try to find some simple function able to capture any such deviation irrespective of its origin. The method that allows to find a function like this is Taylor series expansion. However its application is not straightforward, as it requires some previous choices that have an effect on the final result. In App. 5D I give a detailed justification of my own choices and apply the technique.

As a result I find two distributions that are well-known in the field of diversity.

The first order approximation, for very small deviations, is a power law not constrained anymore to b=1:

$$f(x) = xx^{-b} \qquad (5.10).$$

The second order approximation, for slightly larger deviations, is a lognormal

$$f(x) \propto x^{-1} e^{-\frac{1}{2}\left(\frac{\log(x)-m}{s}\right)^2} \qquad (5.11).$$

The discrete counterpart for Eq. 5.10 is indeed Eq. 5.1. Distributions equal or similar to Eq. 5.1 with $b \neq 1$ have often been found to be of use for diversity studies, though they are sometimes presented under other denominations, such as Pareto, Zipf, Zipf-Mandelbrot, hyperbolic or extended negative binomial distribution. In particular, the distribution found by Siemann et al. (1996, 1999) for several sets of grassland arthropods can be easily demonstrated to be equivalent to Eq. 5.1 with $b \approx 1.5$. These authors claimed their overall sample to be the most thorough sample of an ecological community to date (the sets corresponded to different size classes). Solé and Alonso (1998) also found a power law with $b \approx 1.5$ for a large inventory of tropical rainforest trees.

The lognormal (Eq. 5.11) has probably been the most used distribution for species abundances, since first proposed by Preston (1948). The above developments explain why it provides a rough adjustment in a wide range of cases. In particular, they explain why samples that are well adjusted by a logseries are often found to be also acceptably adjusted by a lognormal (e. g. Death 1996, DeVries et al. 1997).

There are other explanations for the origin of the lognormal. The most widely accepted seems to be the one given by MacArthur (1960). The choice of the one or the other explanation substantially changes both the interpretation and the extrapolability of the lognormal curve. In App. 5E I show some drawbacks of MacArthur's hypothesis.

Preston (1948, 1962) found an apparent constraint on the relationship between parameters of the lognormal in the samples he analyzed. He suggested to reduce the two parameters to a single one and called the outcome canonical distribution. Other authors have confirmed the canonical effect and it is still under research (Sugihara 1980, Tokeshi 1996). I show in App. 5F that Preston's finding may be expected for lognormals obtained in the way I propose.

5.4.5. Environmental noise and scaling niche structure

We began our tour from the upper entrance in Fig. 5.1, demographic noise. This corresponds to the most trivial critical process to expect in an ecosystem and is thus the most elementary source of scaling. However there may be others, which are represented by the side entrances in Fig. 5.1.

We can also expect criticality if a population is driven by environmental noise rather than demographic noise. If the species in our ecosystem differ in their response to environmental variations, the abundance distribution will not be Eq. 5.4 anymore. Environmental noise is usually represented in a crude form, such that the correlation between the individuals in a species is 1 and between different species is 0. According to the diffusion equations in Engen and Lande (1996a and b) and taking App. 5A into account, we should expect a power law with b=2.

Some authors have stressed that niche structure may reflect fractal properties of the environment (Morse et al. 1985).

One open area of research is the possible role of criticality in community assemblage, which could also produce a scaling niche structure.

5.4.6. Species-area relation

One of the central postulates of biogeography is the standard relation between area A and species number S in a given region:

$$S = cA^z \qquad (5.12).$$

Preston (1962) suggested that its origin was a sample size effect to be expected from his canonical lognormal distribution. He stated that this would produce Eq. 5.12 with $z \approx 0.262$. This was the point of departure of MacArthur and Wilson (1967) theory of island biogeography. Though most of their developments refer to mechanisms other than sample size, they proposed that their action would consist of determining c and producing some deviation about the expected value of z, as observed in real data. However, some limitations have been found to the lognormal approach (May 1975, Leitner and Rosenzweig 1997).

There is an essential point that has been missed in this discussion, to my knowledge. In strict terms, a lognormal cannot explain by itself Eq. 5.12 because Eq. 5.12 is a scaling relation, while a lognormal is not scaling. The only scaling distribution is a power law (Eq. 5.1) without upper bound.

From Eq. 5.3, the number of species in a sample is roughly $S \propto TP\left(m > \frac{M}{N}\right)$ (see notation in sect. 5.2), where the probability P refers in this case to the abundance of a species in the community. For an unbounded power law distribution, $P(X > x) \propto x^{-b+1}$. So we find the scaling relation

$$S = c'N^{b-1} \qquad (5.13).$$

If the exponent of Eq. 5.12 depended solely on sample-size effects, we would thus have

$$z = b - 1 \qquad (5.14).$$

For bounded power laws, Eqs. 5.13-5.14 are still to asymptotically hold. Engen (1978) advanced the exact solution for the case of a power law with an exponential bending (which he calls extended negative binomial) and found Eq. 5.13 as an asymptotic result.

What Preston (1962) did was equivalent to approximating his canonical distribution with a power law. If my approach is correct, this was possible because the usefulness of this distribution stems precisely from its proximity to a power law.

In the case of random locations of the organisms, Eq. 5.14 would produce a value of z close to Preston's for the marine phytoplankton case study below, while we would find z=0.5 for the two examples of power law abundance distribution mentioned in subsect. 5.4.4 above.

It seems correct to assume that the additional factors to take into account when passing from random samples to biogeography may primarily express themselves through a change in z. With this reasoning, MacArthur and Wilson (1967) were implicitly using Taylor series. Some other shapes have been proposed instead of Eq. 5.12; He and Legendre (1996) explicitly interrelated them by Taylor series. In a recent analysis of several large sets of tropical forest tree data, Plotkin et al. (2000) found significant deviations from Eq. 5.12, which were also fitted by Taylor expansion. This is similar to some of the developments for abundance distributions in the present chapter.

Like the case of abundance distributions, fractal features of the environment or critical processes in the ecosystem may favor by themselves a shape like Eq. 5.12.

Harte et al. (1999) (see also Rosenzweig 1999) tried to extract an abundance distribution out of Eq. 5.12 in a different way. Their result is different from the mine and has no relation with any of the distributions that have proved useful to fit real data. I show in App. 5G that this is a consequence of some hidden incorrect assumptions in their developments.

5.4.7. The collapse of distributions

One distinctive feature of the theoretical framework I present is its collection of recipes to ignore details. This is necessary to justify any general rule, given the complexity and idiosyncrasy of ecosystems. The steps where we "lose" details are:

- I take demographic noise as a source of scaling. Demographic noise is inherent to reproduction and has the same effect for many possible demographic features and mixtures of demographic features.
- Besides demographic noise, there are other plausible mechanisms able to equally inject scaling.
- I introduced the global constraint without specifying the mechanisms involved. These could alter the result, but there are reasons to assume it to be robust.

- Scaling is unaffected by niche apportioning, at least while we can assume a large number of species per niche.
- I expect most of the possible additional factors to have a common well-defined effect on the distribution, while this effect is not really strong.

The main conclusion is that the results of many different models can be expected to collapse to a few distributions, which precisely coincide with the distributions that have been found useful on practice.

5.5. Interpretation of abundance distributions

The above section proposes a model for the emergence of diversity patterns from the underlying dynamics. Here I build on this model to analyze just the opposite: the induction of dynamics from patterns.

Besides adjustment tests, it is essential to have a graphical representation of the abundance distribution in each set of data. The above developments suggest that some distributions will fairly synthesize most of the information in our data in a few parameters, but we can never guarantee that the match will be complete. A graphical representation allows to notice mismatches between data and standard expectations, which may well be our main clues to nontrivial features of the dynamics. One optimized representation criterion is given in App. 5H.

5.5.1. The meaning of each abundance distribution

- The logseries distribution is confirmed as a fair null hypothesis, to test whether the underlying dynamics is compatible with a wide ensemble of neutral models. The main feature of these models is a lack of ecological differentiation between species; they also neglect the effect of space. On the other hand, a logseries is also compatible with other models. It will not change very much in case of "soft" niche segregation, due to the property of invariance under assemblage.
- The lognormal (in its discrete form) will reasonably match the outcome of many possible models. MacArthur (1960) interpretation is not strictly correct. Other models could be correct, but a lognormal distribution does not suffice to prove any of them.

- Bounded power laws represent another good option to fit distributions produced by many possible models, or at least parts of the distributions. They are less flexible than the lognormal but the meaning of the parameters is clearer, as shown below.
- App. 5D gives the recipe to obtain a sequence of distributions of increasing complexity. If a set of data is not well fitted by a power law and/or a lognormal, it will be fitted by some other distribution in the sequence. The higher the required number of steps to fit a set of real data, the higher the credibility of any simple model able to explain the chosen distribution.

5.5.2. The limits to induction

From the results in sect. 5.4, I expect many different dynamic to produce the same diversity patterns. This sets a severe limit on the induction of dynamics from patterns. Often an outcome like any of the distributions in Fig. 5.1 has been considered a strong proof for some particular model. Any conclusion obtained on this basis should now be reconsidered.

One of the last outstanding examples is the contribution by Pachepsky et al. (2001). These authors present a simulation model for a plant community, which produces a lognormal abundance distribution. They first introduce a complex setting, which is later simplified to single out the mechanisms that suffice to explain their results. Their simplified model is not much different from standard neutral models (Watterson 1974), except in that it excludes immigration/speciation. The point n=0 then acts as a sink where all the species should eventually be attracted, never as a source of new species. After some time of simulation, they find a distribution like a geometric series, as should be expected from neutral models. This distribution is later replaced by a lognormal. This is probably due to a depression in probabilities close to the sink, as extinctions proceed without replacement. This model may be a step in the right direction, but does not really explain the patterns observed in nature, because it requires a permanent species decline. Without the findings in sect. 5.4, it would have been hard to explain the similarity between model results and real-world observations, unless we had attributed the ubiquity of lognormal distributions to anthropogenic extinctions.

Still some induction is possible from abundance distributions:

- We can test whether or not a set of data is compatible with simple neutral models. We can measure an effective distance from neutrality.
- If the set displays idiosyncratic features that cannot be fitted by common-use distributions (we require many terms in the sequence in App. 5E) and we know some explanatory model, we do have a strong proof of the model.
- Empirical distributions can be used to test models predicting the values of the parameters.

The same applies to species-area relations.

The collapse of diversity patterns involves limitations but also advantages. Completely different situations can be described in terms of the same few parameters. The effect of each intervening factor can be expressed as a modification in the value of these few parameters: this simplifies the analysis of the factors and makes them comparable.

The following section identifies the relevant parameters when our data are well fitted by a power law.

5.5.3. Three central measures

Abundance distributions and species-area relations are not completely scaling. However, if deviations are tiny enough to be neglected, diversity patterns can be expressed in terms of three meaningful measures:

- $\Delta b = b - 1$,

  where b is the power law exponent. This has both a structural and a functional meaning. In structural terms, Δb (or simply b) is related to the amount of diversity. This relationship becomes completely clear-cut under the hypothesis of indefinite extrapolability of the power law behind Preston's veil. In this case, a larger Δb (or b) is a sufficient condition to declare a community more diverse than another one. In functional terms, Δb captures the incidence on abundance distribution of any mechanism other than the minimal ecosystem features considered in simple neutral models, when these mechanisms have some effect. It is thus a measure of effective distance from neutrality.
- $D_b = \left(\dfrac{b-1}{n_0^{1-b} - n_M^{1-b}} \right) \dfrac{s(n_0 \le n < n_M)}{N^{b-1}}$, $b>1$.

$$D_1 = \frac{s(n_0 \le n < n_M)}{\log(n_M/n_0)},\ \boldsymbol{b}=1.$$

where N is total sample size, $s(n_0 \le n < n_M)$ is the number of species whose representation in the sample lies in the interval $[n_0, n_M)$, and $\boldsymbol{b}$ has been estimated within this same interval. In the case of extrapolability, if two communities have the same $\boldsymbol{b}$, the larger is $D_{\boldsymbol{b}}$ the larger is the diversity (demonstration in App. 5I).

- $\Delta z = z - (\boldsymbol{b} - 1)$,

  where z is the species-area exponent. This is a measure of spatial segregation, which captures the incidence on species-area relation of any mechanism other than sampling effect.

I. e if we wish to ordinate communities as a function of their level of diversity (if well fitted by power laws), we have the following criteria:

- For two samples with the same $\boldsymbol{b}$, the larger is $D_{\boldsymbol{b}}$ the larger is the diversity.
- If $\boldsymbol{b}$ are different and the sample with the largest $\boldsymbol{b}$ also has the largest number of species (once sample sizes have been equaled by rarefaction, see Hurlbert 1971), this one is the most diverse.
- If $\boldsymbol{b}$ are different but the sample with the largest $\boldsymbol{b}$ has less species than the other (equaled sample sizes), nothing can be said with certitude.

In this last case, if the power law is indefinitely extrapolable, the sample with the largest $\boldsymbol{b}$ will contain the largest number of species above some well-defined sample size. This is one reason to declare it more diverse, but we cannot ensure extrapolability. We are left with this margin of indefinition.

In the above equation for $D_{\boldsymbol{b}}$ I assume that we are working within an interval $[n_0, n_M)$, where the distribution is strictly scaling. In this case, $\Delta\boldsymbol{b}$ can be estimated by maximum likelihood. However, these measures may have a wider use to approximately match complete distributions, even if they display some deviation from scaling. Then it seems more robust to extract $D_{\boldsymbol{b}}$ and $\Delta\boldsymbol{b}$ from a simple regression on the empirical probability "density" function. Details are given in App. 5I.

5.6. Case study: marine phytoplankton

5.6.1. Introduction

Margalef (1994) offers an ensemble inventory of Mediterranean and another one of Caribbean phytoplankton, by putting together more than 1,000 samples in each case. This includes 162,478 identified cells in the first case and 883,352 in the second. This is indeed an exceptional amount of data. When putting together all of the samples in each set, they both surpass the sample size in Siemann et al. (1996, 1999). Here I analyze them under the light of the above findings.

5.6.2. Methods

The data sets used in this section were obtained from Margalef (1994) and consist of Mediterranean and Caribbean phytoplankton. Some of their features are compiled in the first part of Table 5.1. Both sets are the result of putting together many samples obtained at different points, always at the photic zone (down to 110 m). More specifically, all the Mediterranean sampling sites were within the Catalan Sea, while all the Caribbean sites were located along the eastern coast of Venezuela. The first set comprises a total of 195,983 cells from 1,388 samples The second comprises 1,111,859 cells from 1,144 samples.

Only the cells identified to species were taken into account. For each sea, three sets were studied: (i) all the groups of phytoplankton together, (ii) diatoms and (iii) dinoflagellates. Most cells belong to one of these two last groups, which are also the only that were exhaustively identified to species level. It is of interest to analyze them separately because their differences are not only taxonomic but also ecological (e. g. Margalef 1978, 1980).

The data were grouped and displayed according to the criteria in App. 5H.

Power laws were first roughly fitted by simple regression over the complete distributions; D_b (sect. 5.5) were then calculated with Eqs. 5I5-5I6. A more rigorous fit was performed within the bounds $[10,1000)$. The values of b were then obtained by maximum likelihood estimation (m. l. e). The 90% confidence intervals were retained. D_b was calculated for the value of b found in each case, for a rounded b that allows for some comparisons and for $b=1.2$.

Logseries were also fitted. The parameter a was numerically estimated by Fisher (1943) method, which is equivalent to m. l. e. The only inputs for the estimation are the number of individuals and species in the sample. In the case of Mediterranean

	Mediterranean			**Caribbean**		
	All	**Diatoms**	**Dinofl.**	**All**	**Diatoms**	**Dinofl.**
N_T	197535	116409	14055	1113581	779347	102558
N	162478	112352	10874	883352	759794	60851
S	353	107	209	257	118	122
Power law in the interval $[10{,}1000)$, by m. l. e.						
b	1.23	1.02	1.46	1.20	1.02	1.44
90% c. i.	(1.19,1.26)	(0.91,1.07)	(1.40,1.51)	(1.15,1.24)	(0.92,1.07)	(1.35,1.50)
c^2	6.12	0.45	18.0	2.23	4.47	6.83
d. f.	5	5	6	5	5	4
e	0.29	0.99	0.21	0.82	0.48	0.15
D_b	6.24	10.82	1.90	4.60	12.92	0.66
		D_1=12.16	$D_{1.45}$=1.95		D_1=14.77	$D_{1.45}$=0.60
$D_{1.2}$	7.98	2.88	7.22	4.70	2.39	3.43
Power law in the whole range, by regression						
b	1.31	1.13	1.40	1.24	1.14	1.34
D_b	4.99	10.39	3.25	5.24	7.52	1.71
r^2	0.99	0.98	0.98	0.99	0.98	0.99
Logseries						
a	42.8	11.7	36.7	24.5	10.6	14.6
c^2	52.5	3.9	24.3	51.7	28.1	31.0
d. f.	11	11	7	14	14	10
e	$<10^{-6}$	0. 972	0.001	3×10^{-5}	0.014	6×10^{-4}
Poisson lognormal						
m_0	-10.9	-8.6	-8.4	-10.80	-9.5	-9.9
s	3.5	3.4	2.5	3.0	3.0	3.1
c^2	16.7	9.3	15.3	38.7	25.6	5.9
d. f.	10	9	6	10	10	7
e	0.082	0.41	0.018	2.8×10^{-5}	0.004	0.555

Table 5.1. *Statistics of Margalef data on marine phytoplankton. N_T is the total sample size, including both cells identified and not identified to species level, while N does not include unidentified cells. d. f. means degrees of freedom. e is the maximum significance level that allows to reject each the corresponding distribution. c. i. means confidence interval. Rest of symbols as in the text.*

diatoms, the differences about a simple "continuous" geometric series distribution was also represented (i. e. log(xf(x)) was plotted vs. log(x)).

Poisson lognormal was also fitted by m. l. e.

A c^2 test was performed for each distribution. The data were grouped as recommended in App. 5H. The groups for which the expected number of species was lower than 5 were merged to adjacent groups. A reasonable independence between the

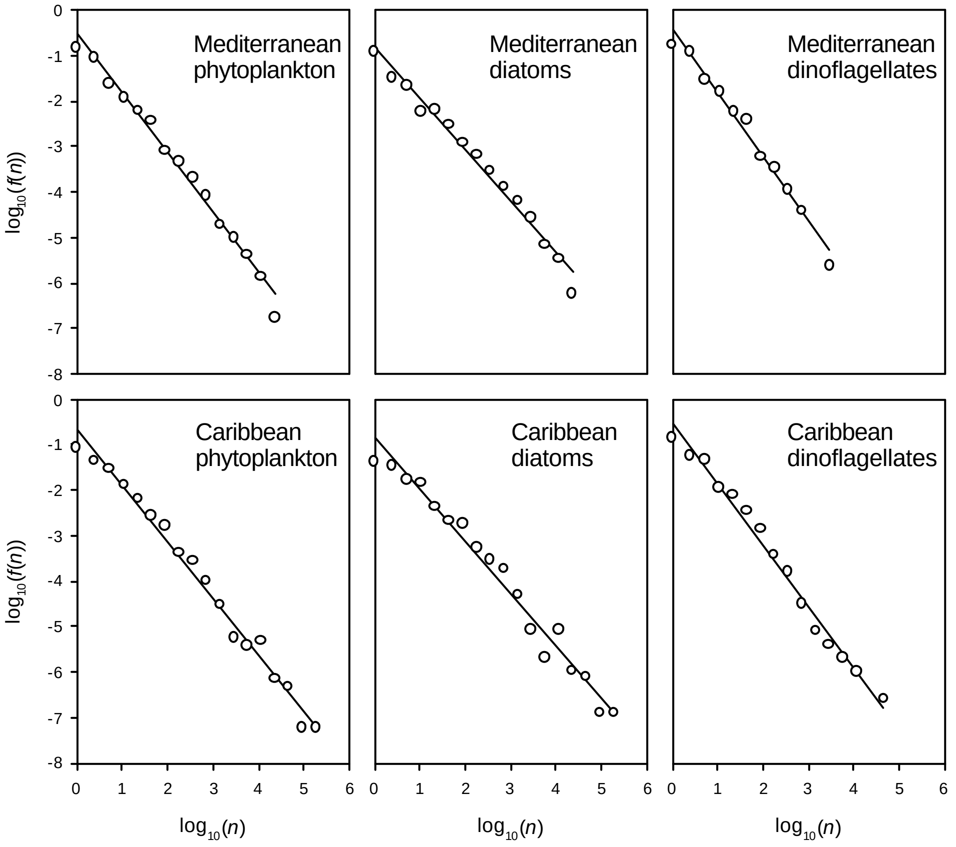


Fig. 5.2. *Scaling in marine phytoplankton. Species abundance distributions, from Margalef (1994) data. n is the number of individuals in the sample and f(n) the probability density. Overall power laws fitted by regression.*

groups was assumed. The final result was the signification level that would allow to reject the null hypothesis in each case.

5.6.3. Results

The results can be seen in Table 5.1 and Figs. 5.2-5.6.

The simple regressions in Fig. 5.2 revealed all of the analyzed sets of phytoplankton to be roughly scaling.

Power laws in the interval $[10,1000)$ (Fig. 5.3) were not rejected by c^2 (e<0.1) in any case. Diatom b are not significantly different from b =1 (90% c. i.), which is the b of geometric series and logseries and corresponds to Δb =0. Dinoflagellate b are significantly higher. The ensemble distributions display intermediate b, significantly

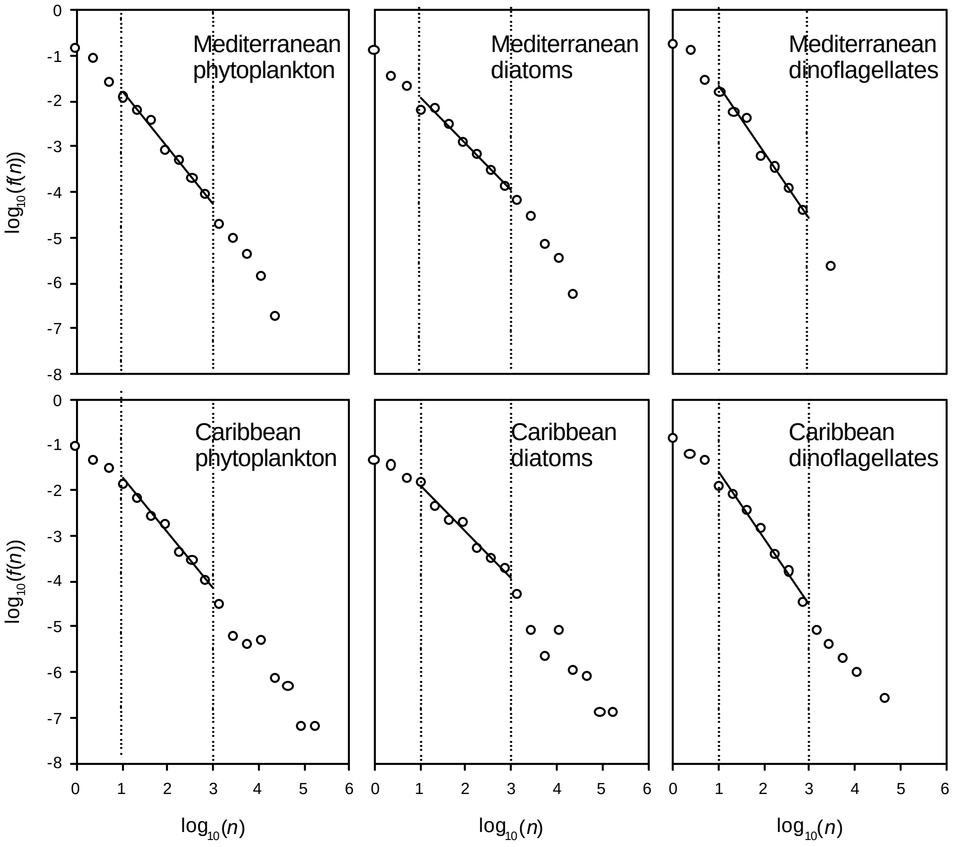


Fig. 5.3. *Fitting power laws to marine phytoplankton. Species abundance distributions, from Margalef (1994) data. n is the number of individuals in the sample and f(n) the probability density. A power law has been fitted in each case for the window from 10 to 1000 individuals, by maximum likelihood estimation.*

higher than b =1. This set of results is found both in the Mediterranean and in the Caribbean.

D_b can be appreciated to strongly depend on the assumed b. When taking rounded b, within confidence intervals, both seas can be compared for each group. The main result is a higher ensemble diversity for the Mediterranean than for the Caribbean; this is however doubtful, as justified below. Different groups cannot be compared through D_b, but dinoflagellates can be labeled as more diverse than diatoms because they exhibit more species in a smaller sample and a higher rate of increase of richness with sample size (higher b).

Logseries (Fig. 5.4) generally offer a reasonable approximation, but can be rejected for all of the sets except Mediterranean diatoms. These completely match the distribution (in this case the fit is so strong to be improbable even for a logseries),

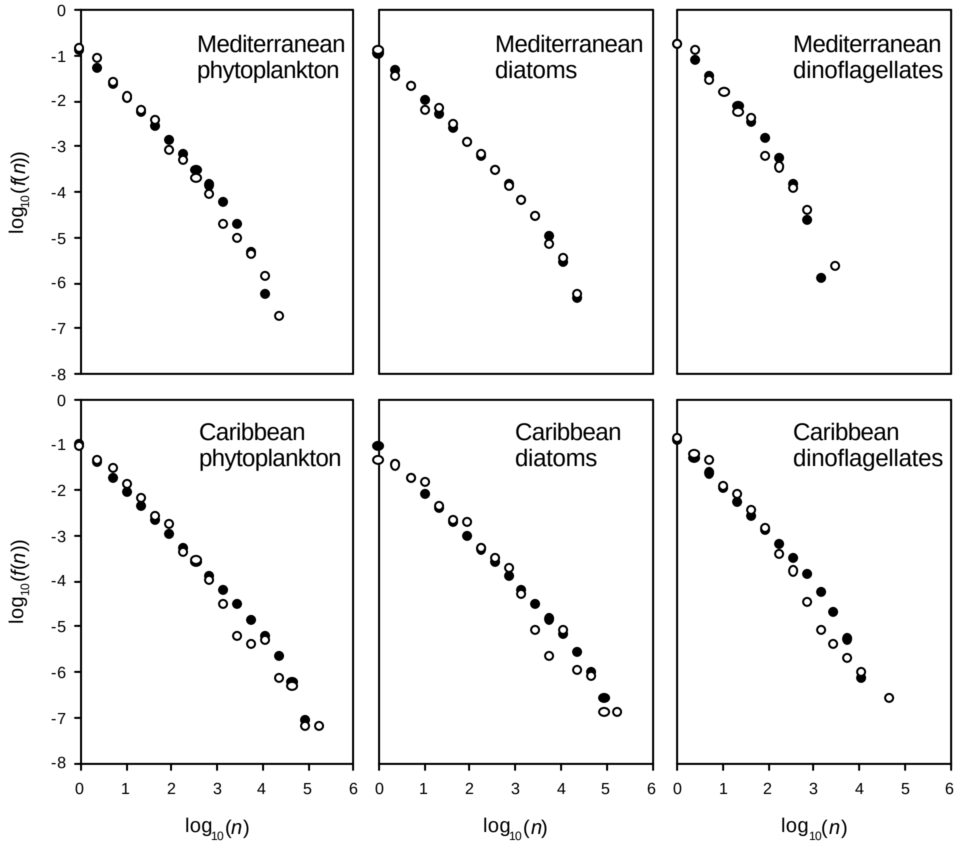


Fig. 5.4. *Testing the neutral model distribution on marine phytoplankton. Species abundance distributions. n is the number of individuals in the sample and f(n) the probability density. Empty spots: data by Margalef (1994). Full spots: neutral model prediction, which corresponds to a logseries. The prediction is made solely on the basis of the number of species and individuals in each case. Mediterranean diatoms completely fulfill the prediction, while the other groups diverge to some extent.*

which can be better appreciated in Fig. 5.5. Caribbean diatoms also have b =1, but diverge from the logseries at the bending function.

Poisson lognormals (Fig. 5.6) offer an acceptable approximation for all of the sets, but in most cases they are rejected by c^2, pointing to even higher order terms. Mediterranean diatoms, which display a complete fit to a logseries, are also especially well fitted by a lognormal.

5.6.4. Discussion and conclusions

The main points I wish to highlight about the results are:

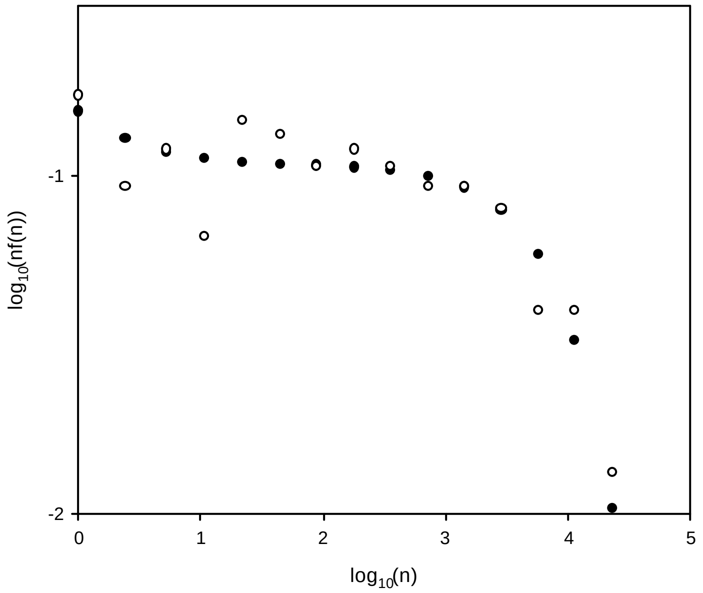


Fig. 5.5. *Residual for Mediterranean diatoms and for the neutral model distribution. Residual of the abundance distribution, after extracting the basal –1 power law. Comparison between empirical results (empty spots) and the neutral model prediction (full spots).*

- All of the sets display an overall scaling distribution in 3-4 orders of magnitude, which is a high figure (see Avnir et al. 1998).
- Mediterranean diatoms cleanly adjust to the predictions for simple neutral models, as can be seen in Fig. 5.4 and Table 5.1. Note that Fig. 5.4 does not consist of regressions, but predictions that only take into account the number of species and individuals in each sample. Predicted and observed data do not only coincide in the overall straight line with a –1 slope, but also in the deviation from the line, as can be seen in Fig. 5.5.
- Caribbean diatoms diverge from the neutral distribution in terms of bending function, without losing its characteristic slope.
- Dinoflagellates unambiguously depart from the neutral model distribution both in the Mediterranean and in the Caribbean. Their scaling slope b is significantly more pronounced, close to 1.45 (Δb=0.45), which indicates the presence of additional mechanisms that seem to favor a higher diversity.
- The observed difference between diatoms and dinoflagellates seems to be congruent with previous knowledge. Diatoms are considered "opportunistic", "r-strategists", while "K-strategist" dinoflagellates are more characteristic of mature stages of succession, in which the ecosystem is assumed to be more structured

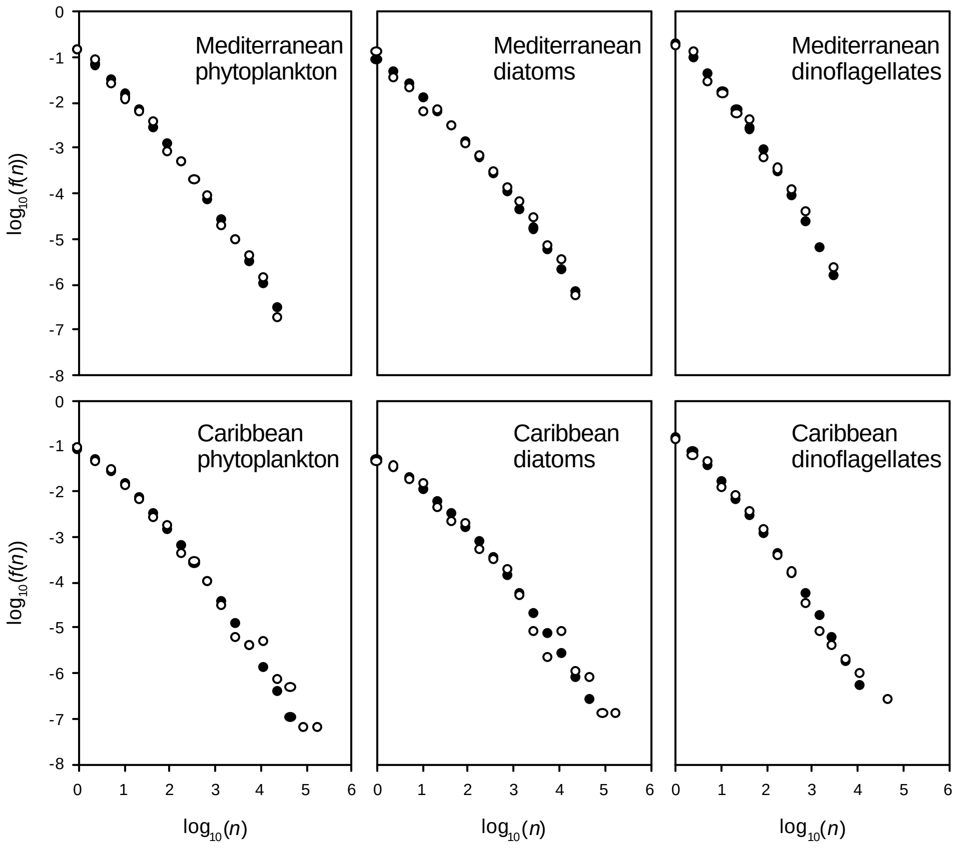


Fig. 5.6. *Fitting the lognormal to marine phytoplankton. Species abundance distributions. n is the number of individuals in the sample and f(n) the probability density. Empty spots: data by Margalef (1994). Full spots: fitted lognormal.*

(Margalef 1978, 1980). Diatoms might really be more like a single guild with a neutral dynamics.

- The overall phytoplankton distribution for each sea remains scaling in the same window where diatoms and dinoflagellates do. This illustrates the principle of invariance under assemblage in a case with a slight difference in b.
- The application of the index D_b to this case makes it clear that it is very sensitive to b. We cannot compare diatoms with dinoflagellates through D_b but we can assert that the last are more diverse, since they have a higher richness in a smaller sample combined with a higher b. According to $D_{1.2}$, Mediterranean phytoplankton is more diverse than Caribbean. The last should not at all be taken as definitive because the Caribbean samples have a much larger proportion of unidentified dinoflagellates.

- The Poisson lognormal model offers a reasonable approach to the observed distributions but not a rigorous adjustment, as expected from the theoretical framework I present. This indicates the presence of even higher order terms.

5.7. Measuring diversity?

Diversity is a multi-dimensional concept. Several levels of diversity are recognized, though most often only species diversity is quantified. Even then, we are left with several dimensions. This is clear from the above developments: a scaling abundance distribution has more than one parameter, and further parameters may be required to capture deviations from scaling. However, there has been a long historical effort to find one-dimensional indices of diversity, for a variety of purposes. This is a timely issue: the success of current conservation efforts is strongly dependent on reliable criteria to set priorities to maximize biodiversity conservation within existing constraints, and this requires some form of comparability.

The first subsection below deals with the behavior of classical indices in front of scaling. The results I find suffice to make it clear that several of the indices are useless and the rest have strong limitations. In the second subsection below I give some corollaries for systematic conservation planning.

5.7.1. Classical diversity indices and scaling

In sect. 5.5 above I give some recipes to quantify the main features of diversity patterns. However, we can find many other recipes in the literature. Besides abundance distribution fitting, there are a set of well-known indices that pretend to be more or less distribution-independent (reviews in Peet 1974, Magurran 1988). The main strength and weakness of these indices is their unidimensionality. Unidimensional indices neglect much information, but will be useful if they (a) capture the relevant information for some given purpose and (b) are statistically sound. In this subsection I analyze this second point. In the following I make some proposals for one specific purpose.

A fundamental trait of diversity indices is their possible dependence upon sample size. While other statistics always converge with sample size, this may not be

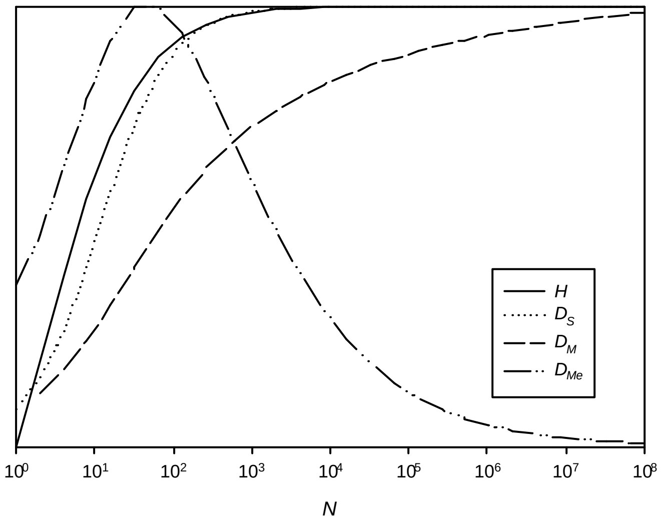


Fig. 5.7. *Diversity indices and sample size. Expected variation of diversity indices for a logseries distribution, with the same parameters that we find for Mediterranean diatoms. The equations in Table 5.2 were used. The indices have been standarized to make them cover the same range of values.*

the case for diversity indices. It is well-known that a larger sample does not only mean less error but also a larger portion of the distribution entering the estimation process.

Fig. 5.7 displays the behavior of some of the most used indices with sample size, in the case of a logseries distribution with the value of the parameter a that we find for Mediterranean diatoms. The indices can be found in Table 5.2 (from Magurran 1988). Some other indices are discussed below. Note that none of the indices in Fig. 5.7 has a fast convergence and one does not even converge.

Among the indices in Fig. 5.7 we can distinguish two groups. On the one side, the so-called Margalef (D_M) and Menhinick (D_{Me}). On the other, Shannon (H) and Simpson (D_S) (Shannon index was also introduced as a diversity index by Margalef (1956) and is the most used by this author).

Shannon	$H = -\sum_{i=1}^{S} \frac{n_i}{N} \log_2\left(\frac{n_i}{N}\right)$
Simpson	$D_S = \frac{1}{\sum_{i=1}^{S} \left(\frac{n_i}{N}\right)^2}$
Margalef	$D_M = \frac{S-1}{\log(N)}$
Menhinick	$D_{Me} = \frac{S}{\sqrt{N}}$

Table 5.2. *Diversity indices*

D_M and D_{Me} assume patterns of variation of species number with sample size that happen to require scaling abundance distributions (see *Species-area relation* above). When this is the case, D_b is an optimum index. D_M and D_{Me} can be taken as rough approximations to D_b, but only when assuming specific slopes: $b=1$ for D_M and $b=1.5$ for D_{Me}. In the example in Fig. 5.7, the first case is the asymptotically correct. Thus D_M converges, though it does extremely slowly, while D_{Me} does not converge at all. Had I chosen a distribution further from the geometric series one (such as what I find for dinoflagellates), neither D_M would have converged. This kind of index is of little use unless generalized and optimized, as I do by defining D_b . Even then we are restricted to the cases in which a power law approximation with constant b is appropriate. The last is important because D_b is very sensitive to b, as can be observed with our case study (Table 5.1). Still if our samples comply with these restrictions we cannot ensure that they are extrapolable to larger samples.

H and D_S can be seen in Fig. 5.7 to behave quite similarly and to converge. They converge as well for $b \neq 1$, provided that $b<2$ (for H) or $b<3$ (for D_S) (App. 5J), which appears to be a realistic condition. The trick is that both are quite insensitive to rare species. This is to some extent a vaccine against the potential non-stationarity introduced by Preston's veil. On the other hand, the convergence is really slow, enough to take with skepticism most of the numbers in the literature. In the cases where H or D_S are considered useful, an interval of possible values should be given, instead of a single one. A honest choice is taking the empirically measured index as a lower bound and, as an upper bound, its predicted value for an asymptotically large sample under the assumption of scaling (i. e. invariant b).

H is often accompanied by the measure of equitability $E = H/\log_2(S)$. In the general case of H converging and S unboundedly increasing with sample size, E will vanish for $N \to \infty$. It is not a fair statistic, as already noted by Peet (1975).

Several other indices are found in the literature such as the a, Q, U and Berger-Parker indices (Magurran 1988), as well as the simple measure of species richness.

One index is the a parameter of the logseries distribution (a is the parameter that determines x and f in Eq. 5.5 as a function of sample size, Fisher 1943). Indeed it will readily converge if the sample is logseries but may not converge at all if it is not.

The Q statistic by Kempton and Taylor is proportional to the slope of the central region of the empirical cumulative probability function of abundances when represented vs. $\log(n)$. This measure will only converge if a large part of the distribution is well fitted by a power law with b=1. It is thus akin to D_M but, when the condition for convergence is satisfied, this is faster because it avoids the problems with the bending function. It is also likely to be less variable with sample size when the condition is not satisfied.

The diversity index based on McIntosh U measure is a simple function of Simpson index, in asymptotic terms.

Berger-Parker index consists of the proportion of individuals in the sample that belong to the most abundant species. This is immune to the effect of rare species, so it converges and does so very fast as compared to H and D_S, but of course incorporates less information.

The index with a clearest meaning is the total species richness in a community. It is also the most difficult to apply. While Shannon, Simpson, U and Berger-Parker's indices escape to some extent Preston's veil by giving a small weight to rare species, in the case of richness a single-individual species and the most abundant one have the same weight. Several methods have been proposed to surmount this difficulty (see reviews by Bunge and Fitzpatrick (1993), Colwell and Coddington (1994)), but they usually have a poor performance when carried to the field (Palmer 1990, 1991).

One way to estimate richness involves the extrapolation of some specified pattern, and is thus conditioned to the actual extrapolability. The pattern may consist of some relation between sample size and species number. It is also common to estimate the parameters of the Poisson lognormal and then calculate the expected number of species in the hidden part of the lognormal curve. If the origin of the lognormal were as proposed by MacArthur (1960), this could be reasonable. If it is as I suggest, there is no reason to think that it will be extrapolable to the lower end of the distribution.

Other strategies take advantage of nonparametric estimators. In general they do not work if the number of undetected species is large. They may thus be useful for other kinds of distributions, but not for power laws (App. 5K).

As a summary:

- The estimation of total species richness from a finite sample is quite unreliable for distributions nearly scaling. If performed, it should be preceded by a careful analysis of the abundance distribution.

- Margalef, Mehinick, α and Q indices and equitability do not converge in general conditions. They should not be used as general indices.
- Shannon, Simpson and U indices do generally converge but slowly, which is a serious problem. A range of possible values should be given instead of a single one, as above explained.
- Berger-Parker index has no problem of convergence. It can be used as a general index, but does not capture much information from the sample.

The main limitation of my own proposals in sect. 5.5 is their reliance upon a good fit to power laws, which may not always be the case. So they cannot completely substitute the use of some of the above indices, in spite of their limitations.

5.7.2. Systematic conservation planning and scaling

The conservation of biodiversity should never be reduced to a matter of protecting limited areas; too narrowly defined "protected areas" might not even be an appropriate recipe in all contexts. However, the effectiveness of paying special attention to particular sites is generally acknowledged.

One central issue in systematic conservation planning is the selection of the set of areas to be protected, in order to conserve the maximum amount of biodiversity within given constraints. One important point to consider is the amount of biodiversity currently represented in each possible set of areas. This depends on two features: the amount of biodiversity in each single area and the complementarity between areas (Margules and Pressey 2000). This is far from sufficient, because the persistence of diversity across time is not guaranteed and, between other things, will strongly depend on the spatial pattern of the reserve network, as stressed by Cabeza and Moilanen (2001). Furthermore, standarized criteria of any kind will rarely suffice for practical decision making (Martinez-Alier et al. 1998). Although all of these considerations indicate that this is just a part of the story, here I focus on the particular issue of the amount of biodiversity in sets of areas, which presents many problems by itself.

Some operational measure of biodiversity is necessary to provide the required comparability. In general, species richness within some set of taxonomic groups is considered appropriate enough. However, even this simple measure presents major limitations, as above explained. This is analogous to the problem of taking a linear measure of a fractal object. Furthermore, available information is ordinarily the result of highly unequal sampling (different "rod sizes"). There is much literature with useful recipes (see Sutherland 2000), but the issue is far from closed.

Some advance has been achieved in comparing diversities when sampled areas are different. Since the scaling relation in Eq. 5.12 is widely accepted, it has been suggested to perform the comparisons on the grounds of the parameter c in this last equation, by assuming a "reasonable" $z \approx 0.25$ (UNEP 1995, sect. 2.3.2.1.2). (In the case of tropical forest trees, the deviations from Eq. 5.12 have been shown to be significant, but homogeneous enough to preserve comparability (Plotkin et al. 2000)). Once we accept that abundance distributions are overall scaling, we are led to the same kind of solution when the problem is sample size.

I first treat the case of samples as depicted in sect. 5.2. Then I pass to a more realistic case.

In the sect. 5.5 above I give a recipe to compare diversities but, in the case of conservation planning, complemetarity should also be taken into account.

The first thing we can do in front of a set of samples is testing whether they satisfy the following conditions (the tests do not require equal sample sizes):

- The deviations from scaling are small for a wide range of abundances.
- b is not significantly different between samples.
- The above properties are maintained when putting different samples together.

If any of the conditions is strongly violated, there is no simple answer to our problem. It is encouraging that b is not significantly different when comparing the same taxonomic groups in different seas in sect. 5.6.

If all of the points are satisfied, the samples should be rarified to make them proportional to the total amount of individuals in the area they represent. They can then be put together and the overall D_b will be the appropriate measure to compare each possible selection (the operation should be repeated enough times to neutralize the random factor that arises when rarifying the samples).

A simple count of species in rarified and aggregated samples may be a good approximation to the above procedure, but it also requires the above conditions to be fulfilled.

One problem with this approach is that the required kind of sample may be unpractical in many cases. Nowadays there is an increasing preference for "rapid biodiversity assessment" methods (Sutherland 2000). Areas candidate to protection are compared on the grounds of the lists of species obtained by experienced naturalists after a pre-established search time.

The above criteria can be adapted to the rapid assessment context provided that the time of first encounter with each species be registered (as done by Pomeroy and Tengencho (1986)). This allows to build a species accumulation curve. If we assume that the search time is proportional to an implicit sample size, the shape of the curve will have a direct correspondence with the abundance distribution and so will its parameters, as expressed through Eqs. 5.13-5.14.

In practice, the curve will also be affected by the relative positions of the organisms. Furthermore, the relation between search time and implicit sample size may not be linear. However, the overall shape is still likely to be maintained, and the changes in the parameters are not such a problem if all the areas have been surveyed in a similar way.

In strict terms, Eq. 5.13 requires abundances to be upperly unbounded, which is not realistic. The first part of the accumulation curve will depend on the bending function instead of the power law. Once we enter the power law region of the abundance distribution, the species accumulation curve will have the form

$$S = \Psi + c' N^{b-1} \qquad (5.15).$$

(with log(N) instead of N^{b-1} at the limit b=1). Eq. 5.15 asymptotically converges to Eq. 5.13.

From Eq. 5.15 follows

$$\frac{\Delta N}{\Delta S} \approx \left[c'(b-1)\right]^{-1} N^{2-b} ,$$

where ΔN is the number of additional individuals we have to identify to find ΔS new species (e. g. ΔS=1). Whenever Ψ cannot be neglected, this equation will be more practical to use than Eq. 5.15. We just have to plot N vs. $\frac{\Delta N}{\Delta S}$ in a bilogarithmic scale and we will have a straight line of slope (2-b). Search time will substitute N in the case we are treating. In any case, once we know b by simple regression we can replace it in Eq. 5.15 and find Ψ and c' through another regression.

Once we have the species accumulation curve, we must verify the above conditions. As a principle, if the conditions are fulfilled, diversities can be compared by means of the parameter c' in Eqs. 5.13 or 5.15, which is proportional to D_b. Then we require some procedure equivalent to the aggregation of rarified samples. Otherwise, we could not take complementarity into account. Fortunately, this is immediate.

The time of search for each site and species should be re-scaled by dividing it by the size of the possible area to be protected in each site. They should also be multiplied by some factor capturing differences in methods, staff or difficulty, if possible. Each site will then have some re-scaled total search time. The part of the search histories that exceed the smallest re-scaled total should be suppressed. This is the equivalent to rarifying. Then we can put together the re-scaled times to encounter for each species and each site in a possible set of areas to be protected. Whenever a species appears several times, we should only leave the register that corresponds to a smallest re-scaled search time. In this way, we obtain an ensemble species accumulation curve for each candidate set of areas and can compare their overall c'.

The set of c' obtained in a research project will not be immediately comparable to sets of c' in other projects. This can be solved if a sample as considered in sect. 5.2 is obtained in one of the sites in each project, for calibration purposes (i. e. N organisms are classified, after choosing them in a reasonably random manner). This will allow to find the constant of proportionality between c' and D_b and express all of the results in terms of D_b, which is a transportable measure.

Simple species counts treated as usual can be a reasonable approximation provided that (a) the surveys are performed in such a way that "rarefaction" is not required, (b) inter-project transportability is not pretended or methodological differences are not important and (c) the test of the above conditions for strict comparability is deemed unnecessary.

5.8. Discussion

Ecology has much in common with statistical physics. Both sciences study systems with so many degrees of freedom that it is impossible to perform predictions or even measurements for each of them. Nevertheless, in the case of statistical physics it is feasible to reach a strong predictability at a global level. The trick is to find aggregate features that are unaffected by the details, thanks to the statistical advantages of large numbers. The present chapter makes part of an effort to find analogous features in the case of ecology (Patten 1975 already expected “large numbers” to have a foremost importance for ecology). The main criteria for each of the steps in the theoretical developments summarized in Fig. 5.1 have been their ecological relevance and their arguable independence from details. I have imported some concepts and mathematical techniques from statistical physics, but I have worked in genuinely ecological terms rather than performing a simple translation from one science to the other. It is encouraging that the results closely agree with previous ecological knowledge and my own observations on one of the largest data sets in the literature. On the other hand, organisms are fortunately much more tricky than molecules. In spite of the strength of large numbers, we can never be sure that they are not going to give us some surprise. So we might reach highly useful results but not the degree of security that statistical physics can sometimes offer.

One statistical physical technique of much interest for the study of diversity patterns was already imported by Engen and Lande (1996a, b): the diffusion method (see Diserud and Engen (2000) for further developments). These authors worked out in detail the distributions generated by several dynamic models and found paths to several distributions of interest, including the lognormal. The use of these models requires to specify the expected rate of population growth as a function of instantaneous population values and then incorporate demographic and environmental noise. In the present chapter's approach, I extract from these models the effect of noise, but nothing else. I show that many possible population dynamics and mixtures of population dynamics will lead essentially to the same result. Otherwise, it would be difficult to explain the wide usefulness of these distributions, because population dynamics may be quite different for different species in the same and different ecosystems. However it is important to know which is the specific dynamics that will produce each distribution with exactitude, as can be obtained from Engen and Lande (1996a, b) and Diserud and Engen (2000). This is especially useful in case we observe some atypical feature in an empirical distribution. One technical point that I suggest to add to Engen and Lande approach is commented in App. 5A.

MacArthur (1960) approach to the lognormal is also related to the one in this chapter, in that relies in large numbers rather than details. However I give some shortcomings of his proposal in App. 5E.

The old set of neutral models giving rise to logseries (Watterson 1974) were also based on population dynamics. Their relation with the approach in this chapter is mentioned in subsect. 5.4.2. Hubbell (1997, 2001) extended one of them with the inclusion of space. This is important to be able to know in which direction will migration push the distributions. However, when passing to more specific issues, we should not forget that the parameters will be different for different species and that there are other factors at work. There is an ongoing polemic on the value of neutral models (Whitfield 2002, Levine 2002). We can extract from the developments in this chapter that neutral models capture some of the essential features of diversity patterns, because these patterns are fairly robust, but that this does not mean that the dynamics of ecosystems is “neutral”. It is rather likely to lie at some point “between order and disorder” and, contrary to naive expectations, it is precisely its “disordered” component that injects a largest predictability, thanks to the large number effects.

Solé and Alonso (1998) and Pachepsky et al. (2001) use simulation to link modelled plant space-extended dynamics with abundance distribution. In the first case the reported outcome is a power law, while in the second it is a lognormal. This is not surprising in the view of the developments in the present chapter. I show in sect. 5.5 that the model by Pachepsky et al. (2001) does not offer a credible explanation for lognormals in nature, but illustrates the fact that many different dynamic give rise to the same final result.

A completely different approach to abundance distributions is niche apportioning models (Tokeshi 1993). I give some general results of interest when niche segregation is weak as compared to population fluctuations, but niche structure may well be the dominating factor in some communities' abundance distributions. Then apportioning models become unavoidable. What most models do is expressing the same information that a distribution contains in a different way (like e. g. a Fourier transform). They allow to know which sequence of breakages is able to produce a given distribution, which can indeed give clues about its actual origin. Further progresses are possible in several ways. On the one hand, in finding solid justifications for breakage patterns. On the other, by producing models with both breakages and re-assemblages (extinctions), able to reach a steady state. It is possible that we are able to map such models to critical branching processes and they lead to scaling distributions irrespective of the details. In this case they will make part of the entrance "scaling niche structure" in Fig. 5.1.

As above mentioned, Solé and Alonso (1998) suggested that diversity patterns could be produced by some self-organized critical process. My approach relies on a "minimal" form of criticality, while leaving the doors open to other critical processes. It is important that the features produced by this minimal criticality be not taken as evidence for more involved models.

Diversity patterns could also reflect the structure of the environment (Morse et al. 1985).

I include in Fig. 5.1 an exit under the impact of possible "other factors" that cannot be included in the sets of well-behaved factors.

Fig. 5.1 links several outstanding distributions from the diversity literature, but not all of them. Perhaps the main absentee is MacArthur broken-stick distribution. This is essentially a geometric distribution (not to be confused with a geometric series distribution), which is the discrete version of an exponential distribution. It could only have been added to Fig. 5.1 as a limit for small samples of a logseries (or other bounded power laws when the bending is exponential). The broken stick is of interest as a metaphor for niche apportioning, but it is doubtful that it really holds in nature for any significant range of abundances.

In what concerns species-area relation, it becomes clear in this chapter that some of the mainstream interpretations are not correct. The other important conclusion is that the exponent has a clear meaning, as explained in sect. 5.5. On the other hand, we should bear in mind that species-area relations, like abundance distributions, are scaling as a first approximation but present significant higher order terms (see Plotkin et al. 2000).

Besides abundance distribution and species-area relation, there are other possible scaling features related to diversity that have not been treated in this chapter. These include at least range sizes, organism sizes, trophic web structure and taxonomic structure. There is abundant literature about all of them, but the methods used in the present chapter could also be useful to improve our understanding of these issues.

The analysis of Margalef's large set of phytoplankton data, which allows to include a wide range of population values, supports the idea that abundances are essentially scaling and deviations from scaling are well fitted by the distributions in Fig. 5.1. This strengthens the theoretical framework as a whole. The results for Mediterranean diatoms are especially suggestive, and so is the difference between diatoms and dinoflagellates. My lecture of these points should however be taken with caution: if something is clear from the present analysis, it is the strong limitation on dynamic interpretations of patterns.

Further understanding of the relationship between dynamics and patterns will be gained from the study of long multispecific time series. Among other things, this will allow to (i) characterize the "niche structure" (mixed with other factors) in terms of the statistical distribution of central tendencies for each species, if there are such tendencies; (ii) put apart the noise in population variations that is common to several species from the part of noise that is independent for each species; (iii) investigate the dependence of the variance of population variations on the initial population. The conclusions we obtain, if extrapolated, may be useful both for data interpretation and for management directed to conservation. We could also gain some insight on the effect of diversity loss (perhaps in terms of D_b or b) on ecosystem function.

Regardless of the underlying mechanisms, the fact that the analyzed data remain overall scaling despite the large sample size has practical consequences. This confirms that the set of three measures I suggest captures much of the relevant information in diversity data. Other measures of diversity are useful as well, but before using them it is important to take into account the problems detected in the subsect. 5.7.1.

The selection of areas to be protected in systematic conservation planning may be improved by taking scaling into account. I give practical recipes in subsect. 5.7.2 above. Of course the protection of particular sites with outstanding contributions to biodiversity does not guarantee that biodiversity is really mantained. Diversity patterns are the result of some underlying dynamics, extended over large areas, which may be affected by many factors. Some of them will indeed be relevant for mangement purposes. There is a huge literature about this issue. Some more progress might be achieved by taking into account the limited set of preferential ways in which the intervening factors can affect diversity patterns, as depicted in this chapter.

5.9. Conclusions

- There are good theoretical reasons to expect abundance distributions to be overall scaling.

- There are good theoretical reasons to expect deviations from scaling to take place in predefined ways, which give rise to some of the main distributions in the diversity literature (Fig. 5.1). This includes the logseries and the lognormal. Preston's canonical effect is to be expected.
- Scaling species-area relation can be explained as a function of scaling abundance distribution. Other proposed relations to abundance distributions are incorrect.
- The proposed origin of diversity patterns is highly insensitive to model details. This sets a severe constraint on the induction of dynamics from patterns.
- I put forward three measures able to capture most of the relevant information in diversity patterns, while the scaling approach is appropriate.
- The analysis of one of the largest sets of diversity data in the literature, consisting of marine phytoplankton data, confirms the expected features for abundance distributions.
- Most indices of diversity have problems of convergence for the observed abundance distributions.
- The quantification of diversity and complementarity for systematic conservation planning can be improved by taking scaling into account. I give practical recipes for this endeavour.

Appendices

Appendix 5A

A technical point on diffusion equations

Diffusion equations usually have the Fokker-Plank form (e. g. Nicolis and Prigogine 1977)

$$\frac{df_t(x)}{dt} = -\frac{\partial}{\partial x}\left(m(x)f_t(x)\right) + \frac{\partial^2}{\partial x^2}\left(n(x)f_t(x)\right) \qquad (5A1),$$

where t is time and $f(x)$ is the p.d.f. of x. Engen and Lande (1996 a and b) introduced their use for species abundance distributions, in two versions. They consider both the direct solution of Eq. 5A1 for a continuous "abundance" x and the solution obtained after substituting $y = \log(x)$ for x. This choice affects the final results.

One point to remember when applying Eq. 5A1 is that populations are discrete instead of continuous. As a principle, we should use a discrete birth-and-death equation rather than a continuous diffusion equation. Birth-and-death processes are more difficult to deal with, though they have also been used to calculate abundance distributions (Solé and Alonso 1998). In this approach there is no option for logarithmic or any other transformation. It is correct to approximate a birth-and-death equation with a diffusion equation, as e. g. shown in Nicolis and Prigogine (1977). However, in this case the result is unambiguously Eq. 5A1, not an analogous equation applied to y instead of x. In order to obtain an intuitive grasp of the problem with the logarithm, note that the true "available states" (the realizable population numbers, which correspond to the set of natural numbers) are evenly spaced in the natural scale, but not in the logarithmic scale. This makes a differentiation approach too artificial in the last case, since differentiation is the infinitesimal limit for evenly spaced differences.

The fact is that well-known neutral models (Watterson 1974) produce b=1, as expected from diffussion when no transformation is performed.

As a conclusion, out of the two options suggested by Engen and Lande, there are good reasons to choose the one without transformations.

Appendix 5B

Application of the Maximum Entropy formalism to abundance distributions

The Maximum Entropy formalism (MAXENT) was first proposed by Jaynes (see Jaynes 1983) and is widely used in statistical physics. MAXENT allows to find the most probable set of values $\{x\}$, under the condition that they fulfill some constraints i of the form

$$\int_{-\infty}^{\infty} h_i(x) f(x) dx = k_i \qquad (5B1).$$

They are the "most probable" in accordance with some "prior" p. d. f. $f_p(x)$. The result is expressed as a new p. d. f. $f(x)$. The role of prior probabilities is e. g. treated by Frieden (1985). The general form of the result is

$$f(x) = f_p(x) e^{\sum_i l_i h_i(x) - 1} \qquad (5B2).$$

The constants $\{l_i\}$ are found by introducing Eq. 5B2 into Eq. 5B1. Since the result should have the properties of a valid p. d. f., the 0th constraint has to be $k_0=1$, $h_0(x)=1$. We are free to choose the rest of constraints.

In the case of abundance distributions, if we take geometric series as the prior distribution, Eq. 5B2 becomes

$$f(x) \propto x^{-1} e^{\sum_i l_i h_i(x) - 1} \qquad \text{(5B3).}$$

If we take as a constraint the sum of each species' population, we have $h_1(x)=x$ and find Eq. 5.5.

The best known use of the Maximum Entropy formalism is as a justification for the Second Law of Thermodynamics, but it has been used in many other fields (Jaynes 1983). I know of no previous use in the prediction of abundance distributions, but it has been applied to other issues in ecology. While statistical entropy had already been used for descriptive purposes by Margalef (1956) (Shannon diversity index) and MacArthur (1955), MAXENT was introduced to ecology by Lurié and Wagensberg (1983) (see Wagensberg et al. 1988, 1991, for further developments). Though undoubtedly important, the main limitation of these pioneering uses of MAXENT was the lack of an explicit prior distribution. These authors implicitly used a uniform distribution, by analogy to the thermodynamic case, but there is no reason to take this distribution as a universal prior. In the case of thermodynamics, its usefulness can be considered a consequence of the theorem of Liouville (Schlögl 1989). In our case we have another justified prior, the geometric series distribution.

Appendix 5C

A technical point on logseries as a sampling distribution

Pielou (1977) notes that there are some technical problems in the step from Eq. 5.5 to Eq. 5.6. They were not completely solved in the literature at that moment, and I do not know them to have been solved thereafter. Here I show a development slightly different from the classical one in Fisher (1943), which offers a solution to the problems.

The usual way to predict the sampling distribution $\{P(n)\}$ for a continuous $\{f(x)\}$ is by introducing it to a Poisson distribution. In accordance with the notation in sect. 5.2 we have

$$P(n)=\frac{Q(n)}{1-Q(0)} \qquad \text{(5C1)},$$

where

$$Q(n)=\int_1^M \frac{1}{n!}\left(x\frac{N}{M}\right)^n e^{-x\frac{N}{M}} f(x)dx=\frac{M}{N}\int_{N/M}^{N}\frac{x^n}{n!}e^{-x}f\left(\frac{M}{N}x\right)dx \qquad \text{(5C2)}.$$

The limits $\left[\frac{N}{M},N\right]$ are generally assumed to be replaceable by $(0,\infty)$

$$Q(n)\approx\frac{M}{N}\int_0^\infty\frac{x^n}{n!}e^{-x}f\left(\frac{M}{N}x\right)dx \qquad \text{(5C3)}.$$

For some variable y and the corresponding change in the parameters of f, this reduces to

$$Q(n)\approx\int_0^\infty\frac{y^n}{n!}e^{-y}f(y)dy \qquad \text{(5C4)}.$$

The combination of Eq. 5C1 and Eq. 5C4 is the usual recipe to obtain sampling distributions. The problem in the case of the logseries is that, if we introduce Eq. 5.5 to Eq. 5C4, it diverges for n=0.

Let us come back from Eq. 5C4 to Eq. 5C2. By introducing Eq. 5.5 into Eq. 5C2 we have

$$Q(n)=\frac{x}{n!}\int_{N/M}^{N}x^{n-1}e^{-\left(f\frac{M}{N}+1\right)x}dx\,.$$

If we substitute $y=\left(f\frac{M}{N}+1\right)x$, we find

$$Q(n)=\frac{x}{n!\left(f\frac{M}{N}+1\right)^n}\int_{f+\frac{N}{M}}^{fM+N}y^{n-1}e^{-y}dy \qquad \text{(5C5)}.$$

If we take $fM+N\to\infty$ and $f+\frac{N}{M}\to 0$, Eq. 5C5 becomes equivalent to the usual approximation in Eq. 5C4 and presents the same problem for n=0. We should analyze if these limits are appropriate.

There is nothing to say for $fM+N\to\infty$. In the case of $f+\frac{N}{M}\to 0$, taking this limit is equivalent to adding $\int_0^{f+\frac{N}{M}}y^{n-1}e^{-y}dy$ to the integral in Eq. 5C5. This term vanishes for $n\geq 1$ but diverges to infinity for n=0. This means that the simplified limits are appropriate if and only if $n \geq 1$. So there is no contradiction in taking $Q(0)$ as a finite constant while using the limits $(0,\infty)$ for the $Q(n)$ term in Eq. 5C1. The problem is thus solved and we can obtain Eq. 5.6 from Eqs. 5.5, 5C1 and 5C4.

Appendix 5D

Application of Taylor expansion to abundance distributions

Take the generalized logseries distribution (Eq. 5.9). Consider the region below the bending function *C*, which corresponds to Eq. 5.4. Assume that a set of factors produces some deviation from Eq. 5.4.

We could only fully appreciate the effect of these factors if we completely sampled an infinite community. In practice, we will only be able to observe a part of the distribution (see sect. 5.2) and we will have some sampling error.

We may express the deviations by means of a Taylor series expansion. If the deviations are small, a few terms of the expansion could well suffice to capture all of the significant information in a sample.

We will accept new terms in our Taylor expansion as soon as they become significant as compared to the sampling error. However, the sampling error is not homogeneous along x. The calculations are strongly simplified if we use some transformation to homogenize it. This is achieved by taking both x and $f(x)$ in a logarithmic scale, as shown in App. 5H (the error is also homogeneous for the Preston (1948) representation, but in this case there is a lower bound to the y-axis that prevents a simple expansion).

Let us perform the expansion around some value $\log(x_0)$ and set $\Delta\log(x) = \log(x) - \log(x_0)$ and $\Delta\log[f(x)] = \log[f(x)] - \log[f(x_0)]$. We have from Eq. 5.4:

$$\Delta\log[f(x)] + \Delta\log(x) = 0$$

As soon as we have some non-null deviation, we can express it by means of a function h:

$$\Delta\log[f(x)] + \Delta\log(x) = h[\Delta\log(x)].$$

We do not know h, but if it is indefinitely derivable it will satisfy

$$h[\Delta\log(x)] = \sum_{i=1}^{\infty} a_i \{\Delta\log(x)\}^i$$

for a series of constants $a_1, a_2, ..., a_\infty$.

As we consider successive terms, we have:

$$\Delta \log[f(x)] + \Delta \log(x) = a_1 \Delta \log(x) \qquad (5D1),$$

$$\Delta \log[f(x)] + \Delta \log(x) = a_1 \Delta \log(x) + a_2 \{\Delta \log(x)\}^2 \qquad (5D2),$$

$$\Delta \log[f(x)] + \Delta \log(x) = a_1 \Delta \log(x) + a_2 \{\Delta \log(x)\}^2 + a_3 \{\Delta \log(x)\}^3 \qquad (5D3),$$

etc.

From Eq. 5D1 we obtain a power law distribution, not constrained anymore to b=1:

$f(x) \propto x^{-b}$,

where $b = -a_1 + 1$.

From Eq. 5D2 we obtain the lognormal distribution (while a_2<0):

$f(x) \propto x^{-1} e^{-\frac{1}{2}\left(\frac{\log(x)-m}{s}\right)^2}$,

where $m = -\frac{a_1}{2a_2} + \log(x_0)$ and $s^2 = -\frac{1}{2a_2}$.

Lognormals could fit complete distributions, including their bending functions, which is indeed not the case for power laws.

In intuitive terms, what we do with these steps is taking the straight line in a log-log plot and introducing first a change in its slope (Eq. 5D1) and then some curvature (Eq. 5D2), which are the simplest possible modifications.

It is immediate to obtain more complex functions as soon as the lognormal is unable to fit an empirical distribution.

Appendix 5E

Problems with MacArthur explanation of the lognormal

MacArthur (1960) gave an explanation for lognormal abundance distributions different than mine. This difference is important because it affects both the interpretation and the extrapolability of the lognormal. MacArthur's was an important step forward, as it linked dynamics to diversity patterns. However, it should not be taken anymore in strict terms as a plausible explanation. Engen and Lande (1996) pointed out the lack of intraspecific density dependence as a default of his model. Here I show some even more basic problems.

MacArthur defined a rate of increase of a species *i* at time *t* by

$$r_i(t) = \frac{1}{x_i(t)}\frac{dx_i(t)}{dt}.$$

By integrating over time,

$$\log(x_i(t)) = \log(x_i(0)) + \int_0^t r_i(t)dt \qquad (5E1).$$

He considered two possibilities: either the first or the second term in Eq. 5E1 is the most important. In the first case x_i will remain close to a population of reference, $x_i(0)$, as a consequence of some regulations acting on each species separately. In the second case, the pattern of diversity will be mainly a result of random fluctuations.

He assumed the rates r to vary independently over time. This corresponds to a community driven solely by environmental noise. By applying standard diffusion equations (Engen and Lande 1996a, b) plus App. 5A) this results in a power law with b=2. MacArthur used a different approach. By applying the Central Limit Theorem to the integral in Eq. 5E1, he stated that it would converge to a Gaussian distribution and thus $f(x)$ would be lognormal.

This difference of results is a product of a difference in the implicit premises.

The diffusion approach requires to specify some boundary conditions. There is indeed a lower limit to population at x=1. This can be represented by a lower absorbing barrier such that species become extinct when crossing it, and new species are introduced close to it. We can assume that introductions take place at random over time. There is also an upper limit to population, which can be represented by a reflecting barrier but can also be ignored. In any case the eventual steady-state distribution is a power law with b=2.

The application of the Central Limit Theorem requires different premises. All of the species should be introduced at the same time. All of them should have the same initial population $x(0)$, which may be much larger than 1 (since it has to coincide with the central tendency observed in empirical distributions). They should not reach any barrier: this means that either there is no bound to population (neither inferior nor superior) or the massive species introduction took place just a little time ago. In either case the result is an unsteady distribution, with increasing variance over time.

The last set of premises is indeed much more unrealistic than the former. What it represents is like the spread of a single drop of ink in a glass, while it has still not reached the walls of the glass. If we have one or several drops and enough time passes, the ink will reach a homogeneous distribution in the glass. This last will be the eventual result of diffusion, like the power law in the case of diversity.

In spite of its formulation, MacArthur probably did not think of $x(0)$ just as a point of departure, but implicitly assumed that there were some slight forces pointing to it. Then it is true that the result will be close to a lognormal. In this sense MacArthur's intuition was correct, but the formal demonstration should not be based on the Central Limit Theorem but on Taylor series, as performed in this chapter.

This divergence of explanations has at least two practical consequences.

Under MacArthur view, any deviation from lognormality would denounce the existence of mechanisms other than random fluctuations. In our view, it is the deviation from power laws that counts, so if the use of lognormals is required we have an evidence for the presence of such mechanisms.

Furthermore, the Central Limit theorem produces a lognormal as an exact outcome, so it gives a basis for an indefinite extrapolation behind Preston's veil (e. g. to infer the total species number). This is not the case if the lognormal is only a good approximation when we have just a part of the actual distribution.

Appendix 5F

A possible origin of the canonical distribution

The idea of the "canonical lognormal" was put forward by Preston (1948, 1960) as a result of an empirical observation that nowadays is still considered an enigma. This has been widely studied by Sugihara (1980), has been confirmed by several authors and is still under research (Tokeshi 1996). The framework exposed in this chapter allows to offer a new explanation to this issue.

Preston distributed the species in a sample in a series of groups that he called octaves. Although his exact definition was a little more complicated, for our purposes it is a good approximation to consider that an octave o includes all of the species whose abundances lie in the interval $\left[2^{o}, 2^{o+1}\right)$. For several samples, he represented o vs. the sum of the abundances of the species in each octave, and found a shape like the lower half of a Gauss bell. Assuming that the abundance distribution is lognormal, Preston noted that this result implied a constrain in the relationship between its two parameters and reduced them to a single one. He called canonical this particular case of lognormal.

I obtain the lognormal as a small modification over a power law, whose slope b is not much different from b=1. Consider Preston's graph in the case of a power law. In a continuous approximation, the number of individuals in the octave o is $S\int_{2^o}^{2^{o+1}} xxx^{-b}\,dx = Sx\frac{2^{2-b}-1}{2-b}e^{(2-b)\log(2)o}$, where S is the number of species in the sample. This function increases exponentially with o if $b < 2$, as is the case for b close to 1. This will rule for all of the range of the distribution except the upper end, where the bending function will produce a downward inflexion. This result is similar to half a Gauss bell. Hence actual lognormals may be canonical due to their proximity to bounded power laws.

Preston (1962) intuited some relationship between his canonical distribution and the distribution used by Pareto (1897) in a different context. The last is a power law, upperly bounded in some case.

Appendix 5G

Why the Harte-Kinzig-Green distribution is not realistic

Harte et al. (1999) developed an abundance distribution out of a scaling assumption linked to the standard species-area relation (Eq. 5.12). The distribution is essentially a bounded power law like Eq. 5.1, but has an exponent b<0. Neither my case study nor any of the cases that I know from the literature leave any room for b<0. Maddux and Athreya (1999) signaled some weaknesses of their model. Here I show where is the problem.

Harte et al. take a rectangle of area A_0, inhabited by N_0 organisms belonging to S_0 species. They divide it through successive bisections. After the ith bisection, the initial rectangle has been divided in 2^i rectangles of area $A_i = A_0/2^i$ and number of individuals $N_i = N_0/2^i$, belonging to S_i species.

Take a rectangle of area A_i and its two halves of area A_{i+1}. Each of the S_i species in the first should be found in at least one of its halves. There is some probability x that it is found only in one of the halves, and the complementary probability 1-x that it is found in both halves. Under the assumption of scaling, x

should be the same at each bisection. This is linked to the species-area relation (Eq. 5.12).

From this Harte et al. propose the recursion

$$P_i(n) = xP_{i+1}(n) + (1-x)\sum_{k=1}^{n-1} P_{i+1}(n-k)P_{i+1}(k) \qquad \text{(5G1)},$$

where $\{P_i(n)\}$ is the abundance distribution for the S_i species in an area A_i. By numerically solving Eq. 5G1 they obtain their "turned" power law distribution.

While at first sight Eq. 5G1 seems to be the logical consequence of the stated premises, in fact it requires some hidden assumptions that are not realistic.

Take the case in which the organisms are located at random. Then the (hypergeometric) probability that a specific one of the halves contains no individual is

$$P_{i+1}(0) = \frac{\binom{n}{0}\binom{N_i - n}{N_{i+1} - 0}}{\binom{N_i}{N_{i+1}}} = \frac{(N_i - n)!\,(N_i/2)!}{N_i!\,(N/2 - n)!}\,.$$

The probability that all of the individuals are concentrated in an unspecified one of the halves is

$$x = P_{i+1}(0) + (1 - P_{i+1}(0))\frac{P_{i+1}(0)}{1 - P_{i+1}(n)} = P_{i+1}(0)\left(1 + \frac{1 - P(0)}{1 - P(n)}\right).$$

Since $P(n)=P(0)$, this is an indetermination for $n=0$ (then $P(0)=1$) and otherwise reduces to

$$x = 2P_{i+1}(0) = 2\frac{(N_i - n)!\,(N_i/2)!}{N_i!\,(N/2 - n)!}\,.$$

This is 1 for $n=1$ and vanishes quite fast when increasing n. Conversely, Harte et al. take x to be independent of n in Eq. 5G1. This is the reason why they do not find a power law with a positive $\boldsymbol{b}$ that agrees with Eq. 5.1.

If the organisms belonging to the same species are clumped in any simple way, the values of x will increase and their dependence on n decrease, but it will not disappear and another thing will change in Eq. 5G1: when a species is found in both halves, the abundance in each one will tend to be unequal, while Eq. 5G1 assumes that they are independent from one another.

It is difficult to find any justification for such strange assumptions. Indeed the result is not a simple consequence of the explicit premises. It is not at all a necessary consequence of scaling.

Another misleading result in the same paper is a "endemics-area relationship". Let a be the probability that a species found in a rectangle is also in a specific one of its halves ($a = 1 - x/2$). 1-a is thus the probability that it is only in a specific half and not in the other. By recursively applying this probability, the authors conclude that the number of species in an area A_i that are found nowhere else is

$$E_i = (1-a)^i S_0 \qquad \text{(5G2).}$$

Since, as stated by the authors, $a = 2^{-z}$ (z as in Eq. 5.12), this can be also expressed as

$$E_i = \left(1-2^{-z}\right)^i S_0 \qquad \text{(5G3).}$$

This involves again hidden assumptions. Eq. 5G2 takes the result for a species splitting event as independent of the results for the rest of events. This is not the case for independent or simply clumped locations of the organisms in each species. More abundant species will have higher a (lower x). A species that is found in many rectangles A_i will be likely to be abundant and thus have a high probability of being represented in both halves A_{i+1} of any of these rectangles. At least for the case of random locations, we can deduce from the species-area relation (Eq. 5.12)

$$E_i = S_0 - c(A_0 - A_i)^z = S_0\left(1-\left(1-2^{-i}\right)^z\right),$$

which is contradictory with Eq. 5G3 for $i>1$.

Appendix 5H

An optimized representation of empirical abundance distributions

We have no way to guarantee that an arbitrary sample will be completely fitted by any theoretical distribution. Furthermore, the deviations from the main theoretical distributions may reveal the underlying mechanisms. Thus it is important to represent the complete empirical distribution in some way that is as informative as possible.

The simplest way to represent data is in the form of a histogram. However common histograms give large problems in the case of power laws, because most data are then concentrated in a few intervals while many intervals remain completely empty. They have also other limitations. Here I propose a modification of the histogram that is optimized for the geometric series distribution and thus appropriate for its relatives in Fig. 5.1.

First of all it is important to use multiplicative intervals of the form

$$[n_A, kn_A) \qquad \text{(5H1)},$$

as suggested by Preston (1948). In the case of geometric series (Eq. 5.4), the distribution of the data by intervals becomes roughly homogeneous

$$P([n_A, kn_A)) = \sum_{n=n_A}^{kn_A} P(n) \approx \int_{n_A}^{kn_A} \frac{x}{n} dn = x \log(k)$$

Preston (1948) used k=2 and called his intervals "octaves". This is the constant that allows for a highest detail while keeping discrete intervals. Preston octaves are not exactly as Eq. 5H1, since they allow for some overlapping. As noted by DeVries et al. (1997), this is a problem for tests such as c^2. So there are reasons to prefer Eq. 5H1.

Let us call octave o to the interval of abundances $[2^o, 2^{o+1})$. The (logarithmically) central abundance in an octave is $n_o = \sqrt{2^o\left(2^{o+1}-1\right)}$. Let s_o be the amount of species in the octave.

The estimator of the aggregate probability P_o for all the abundances in an octave o is $\hat{P}_o = s_o/S$. This is proportional to interval length, so we can divide it by the length and obtain the estimated probability density $\hat{p}_o = \frac{1}{2^o}\frac{s_o}{S}$. "Probability density" is a term that belongs to the domain of continuous variables, but it is also useful in our case in spite of the actual discreteness of abundances.

The best way to represent the set $\{(n_o, \hat{p}_o)\}$ is in a bilogarithmic scale. In this way the chosen $\{n_o\}$ are regularly spaced and the confidence intervals for $\{\hat{p}_o\}$ are roughly homogeneous in the case of geometric series, as shown below.

If we approximate the abundances as independent, s_o has a binomial distribution. So for large S the approximate distribution of $\hat{p}_o$ is $\hat{p}_o \approx N\left(p_o, \sqrt{\frac{p_o}{2^o S}}\right)$. The confidence intervals for $\hat{p}_o$ will thus be proportional to $\left(2^{-o}\left(P_o - \sqrt{\frac{P_o}{S}}\right) 2^{-o}\left(P_o + \sqrt{\frac{P_o}{S}}\right)\right)$. In the case of geometric series, where P_o is roughly constant, this interval will shrink with o. However, if we represent it in a logarithmic scale, its length will be $\log\left(2^{-o}\left(P_o + \sqrt{\frac{P_o}{S}}\right)\right) - \log\left(2^{-o}\left(P_o - \sqrt{\frac{P_o}{S}}\right)\right) = \log\left(\frac{P_o + \sqrt{\frac{P_o}{S}}}{P_o - \sqrt{\frac{P_o}{S}}}\right)$, which remains constant while P_o does.

Besides the small change in the bounds of the octaves, this representation is different from Preston's in that he represented $\hat{P}_o$ in a linear scale instead of $\hat{p}_o$ in a logarithmic scale. This was equivalent to transforming the data in a way that allows to contemplate the Gauss curve in case of lognormality. Preston representation is thus useful as a way to compare with the lognormal distribution, rather than as a general-use technique.

Appendix 5I

The quantification of diversity in power law distributions

Out of the two parameters b and x in a bounded power-law, only b is independent of sample size. If we want to compare communities with the same b, we require some size-independent counterpart for x.

From Eqs. 5.2 and 5.10, the expected number of species in an interval $[n_0, n_M)$ will be

$$s(n_0 \le n < n_M) = Tx \int_{n_0 \frac{M}{N}}^{n_M \frac{M}{N}} m^{-b}\, dm \qquad (5I1)$$

(notation as introduced in sect. 5.2 above; in this case x refers to the underlying distribution). We can extract from Eq. 5I1 the measure $D_b = xTM^{1-b}$, which is independent of sample size and the chosen interval $[n_0, n_M)$. From Eq. 5I1, its value is

$$D_b = \left(\frac{b-1}{n_0^{1-b} - n_M^{1-b}} \right) \frac{s(n_0 \le n < n_M)}{N^{b-1}},\ b > 1 \qquad (5I2).$$

The corresponding measure in the limit $b=1$ is

$$D_1 = \frac{s(n_0 \le n < n_M)}{\log(n_M / n_0)} \qquad (5I3).$$

D_b is a fair way to compare diversities under a given b because the knowledge of b and D_b suffices to know the expected number of species for any N and n, while within the scaling range:

$$E(s(n)) \approx D_b N^{b-1} n^{-b} \qquad (5I4).$$

If the species-area relation were just a sampling effect, this equation would be a fair synthesis of the abundance distribution (Eq. 5.1) and the species-area relation (Eq. 5.12).

In the case of a logseries, a and D_1 converge, so D_b is a generalization of a.

It is easy to see from Eq. 5I2 that, when joining several sets of species with the same b, the ensemble index is

$$D_b = \sum_i r_i^{b-1} D_b^i \ ,$$

where r_i is the proportion of individuals in group i and D_b^i is its index.

The combined use of b and D_b in quantifying diversity is commented in sect. 5.5.

D_b can be obtained from the direct application of Eqs. 5I2-5I3 once b is known from maximum likelihood estimation. In sect. 5.5 I state that in some cases it may be better to estimate these parameters by regression. The regression line will have the form

$$\log(f(n)) = q - b\log(n) \qquad (5I5),$$

where $f(n)$ is the probability "density" function (see App. 5H). D_b is then a function of q.

$$D_b = \exp(q) S N^{1-b} \qquad (5I6).$$

Appendix 5J

Convergence of Shannon and Simpson indices under power laws

Since we are not interested in the exact variation of diversity indices with sample size but in their convergence or divergence, we may take a continuous approach for simplicity.

The expected value of Shannon index (Table 5.2) for a sample of N individuals is

$$H \approx -\int_1^{n_M} s(m)\frac{m}{N}\log_2\left(\frac{m}{N}\right)dm \qquad (5J1),$$

where $s(m)$ is the number of species with abundance m and n_M is the upper bound to the abundance distribution. From Eq. 5.2,

$$n_M = gN \qquad (5J2)$$

for a constant g (which is a function of D_b and b). For a power law with $b \neq 1$, from Eqs. 5J1, 5J2 and 5I4,

$$H \approx -\int_{1}^{gN} D_b N^{b-1} m^{-b} \frac{m}{N} \log_2\left(\frac{m}{N}\right) dm,$$

$$H \approx \frac{D_b}{(2-b)\log(2)} \left\{ -g^{2-b} \log(g) + \frac{g^{2-b}}{2-b} - \left[\frac{1}{2-b} + \log(N)\right] N^{b-2} \right\}.$$

For $b \geq 2$ this diverges to infinity. Otherwise converges to

$$H \approx \frac{D_b g^{2-b}}{(2-b)\log(2)} \left\{ -\log(g) + \frac{1}{2-b} \right\}.$$

While the convergence is already slow for b=1, it is expected to worsen when approaching the limit $b = 2$.

For the Simpson index (Table 5.2),

$$D_S \approx \int_{1}^{gN} D_b N^{b-1} m^{-b} \left(\frac{m}{N}\right)^2 dm,$$

$$D_S \approx \frac{D_b}{3-b} \left\{ g^{3-b} - N^{b-3} \right\}.$$

For b>3 this diverges to infinity, while for b<3 converges to

$$D_S \approx \frac{D_b g^{3-b}}{3-b}.$$

Since actual communities do not seem to display any $b \geq 2$, the wider margin where the Simpson index converges does not automatically transform it into a better index than Shannon's. The higher the exponent for n in the index (which at the limit measures the insensitivity to rare species), the higher b it can tolerate (i. e. a higher rate of increase of rare species when running Preston's veil), but it would be senseless to state that the higher is the exponent the better is the index. Provided that they converge and do it at a reasonable speed, different indices must be compared on the basis of their meaning.

Appendix 5K

Reliability of nonparametric richness estimators under power laws

The general problem with nonparametric estimators of the total species richness in a community is that they require the number of undetected species not to be very large. To the extent that actual abundance distributions resemble power laws, this is indeed not the case. Here I show some examples.

Harris (1959) studied the variation in the lower and upper bound to the number of kinds of elements (species in our case) that one can expect when enlarging a known sample. While the lower bound converges, the upper bound readily diverges to infinite. Chao (1984) offered a nonparametric estimator for the lower bound and proposed it as a reasonable approach to the actual number of kinds of elements. She claimed good results when applying this criterion to some instances of numismatic diversity. However, in distributions close to power laws, the actual number of species that we find when enlarging a sample is much further from the lower bound than would be the case in a textbook distribution.

Heltshe and Forrester (1983) applied a jackknife procedure to remove the bias in the number of species that is estimated from a finite sample. However, the jackknife method can only cope with small biases (Manly 1991). It is thus useless in the neighborhood of a power law distribution. Smith and van Belle (1984) applied a modified version of jackknife that allows to work with higher biases, but the price is a huge error in the estimates.

In spite of the level of technical optimization reached by these and other authors, we cannot obtain from a sample any information that it does not contain. The above methods are only useful for communities that are small in either number of individuals or species.

Appendix 5L

Comment on Ritchie and Olff's "synthetic theory of biodiversity"

Ritchie and Olff (1999) recently published a purported "synthetic theory of biodiversity". This mainly relies on the role of the sizes of organisms for biodiversity. Although I have not dealt with this point in this chapter, it has a direct connection with the issues here treated, as it pretends to set a general framework for the understanding of biodiversity. This paper was published in a high impact journal but is clearly wrong. Therefore, I judged opportune to include a brief comment. I do not expound their theory, so this comment can only be understood after reading their paper. Although their formal developments are not correct, their basic insights are interesting and could well render worthwhile results if their models are conveniently rebuilt.

These authors assert that the resource requirements per individual are roughly size-invariant. This is obviously false. In spite of their faster turnover, it is cheaper to maintain a population of *N* hamsters than a population of *N* elephants. The authors misinterpret the results by Damuth (1981), who found a size-invariant energy flow per species in a community, but accompanied by huge differences in population size.

They introduce some scaling relations for food-patch size P, food concentration Q and resource concentration R. These relations make the three variables dependent on a linear measure w, hence dependent on one another. After a series of incongruent steps where the dependences are either considered or ignored with arbitrariness, they obtain a niche partition in terms of ranges of P and R. The last seem to be taken as independent variables in the final outcome. The resulting minimal P for a certain species is smaller than w^D for a large enough w (whenever $mkrw^Q > L$, i. e. $E(dN/dt|w) > 0$), while w^D was previously defined as the minimal P.

Since the authors equate resource uptake per individual and unit time to $kRBP$, the uptake in a niche defined by the ranges (P_k^*, P_j^*) and (R_i^*, R_j^*) amounts to $kB\int_{P_k^*}^{P_j^*}\int_{R_i^*}^{R_j^*} PRf(P,R)\,dP\,dR$, where f is some probability density function. Instead, they state that it is $k(P_k^* - P_i^*)(R_i^* - R_j^*)$.

The authors find their model to reasonably fit empirical observations on East African grazing mammals and North American “savannah” plants. However, different models may give rise to equations with similar structures. As an instance, I offer a model that adheres to their overall approach and does really result in such an equation.

Consider that an organism of size w scans an area $\propto w^{E1}$ per unit time. Assume a density $\propto W^{-E2}$ of resource patches of size W and energy content $\propto W^{E3}$. We take the minimum patch size W available to the organism as $\propto w$ and the energy requirement of the last as $\propto w^{E4}$. If niches do not overlap, the minimum size ratio g between species of adjacent size is given by $w^{E1}\int_{w}^{gw} W^{E3}W^{-E2}\,dW \propto w^{E4}$. The solution is $g(w) \cong \{1 + aw^{E4-E1-E3+E2-1}\}^{\frac{1}{E3-E2+1}}$, where a is a constant. This indeed has the same structure as the authors' Eq. (2) and can fit the same data. (Note that Damuth (1981) observations suggest $E4 - E1 = 0$, but there could be some other exponent in their place if e. g. the location of the organisms is dependent on resource density).

Chapter 6

Reflections on irreversibility, criticality and sustainability

6.1. Introduction

There is an increasing number of authors who think that current economic theory has to be deeply modified, and that ecology should play a key role in this endeavor. Perhaps the main weakness of this approach, called ecological economics (see e. g. Martínez Alier 1999), is the weakness of the own ecological theory, and this was one of the main reasons for my trying to contribute to its improvement. As a natural corollary, in this final chapter I put forward some opinions on the orientation that ecological economics should take and remark the ways in which my own ecological findings might help. While the other chapters are or pretend to be scientific works, this one has admittedly an appreciable ideological component. It also includes a larger proportion of ideas that are not new, for reasons of completeness and coherence.

The interface between ecology and economics has two quite different aspects. On the one hand, there is a tangible interface. The human economic system is just one subsystem within the Biosphere, albeit one with much impact on it. Therefore, the dynamics of the one cannot be understood without the dynamics of the other. On the other hand, there is a purely conceptual interface. Both ecosystems and economic systems are enormously complex self-organized systems, which pose comparable theoretical problems, and the theoretical tools necessary to solve them could also be quite similar. Until now, ecological economics has dealt essentially with the first aspect. The second aspect has only begun to attract a strong attention quite recently, but essentially in journals of physics (e. g. Bouchaud et al. 2001; see however Krugman 1997), under the label of econophysics. Here I try to clarify some issues related to the first aspect that have been considered to be important for ecological economics (about thermodynamics, ecological disturbances, critical thresholds, the concept of sustainability in a changing Biosphere...) and discuss the importance of taking into account the second aspect as a source of “internal sustainability” constraints when thinking of strong modifications in the economic system, like those that some ecological economists (and me too) might suggest. I conclude with a preliminary analysis of the ecological inequality between countries and some thoughts about globalization, which integrate both aspects.

6.2. Irreversibility and the economic process

6.2.1. Economics, ecology and thermodynamics: Georgescu-Roegen's rights and wrongs

One of the main fields of research within ecological economics is the possible role of thermodynamic irreversibility as a constraint on economic development (e. g. Binswanger 1993). Contemporary interest in this issue was triggered by the developments in the 10th chapter of the book *The Entropy Law and the Economic Process*, by Nicholas Georgescu-Roegen (1971), although many other authors had previously analyzed economic systems from an energetic point of view (see Martínez Alier 1984, Martínez Alier and Schlüpmann 1990).

Georgescu-Roegen (1971) sustained that "low entropy" is a necessary condition for economic value but there is no quantitative relationship between both, just like value is a condition for price but neither in this case there is a quantitative relationship. He though that in the environment there is nothing else than increases of entropy, while economic agents generate order, but this last process feeds on environmental "low entropy" and accelerates the "continuous and irrevocable" increase of entropy in our environment. Our species would only exist until depleting the reserves of "low entropy" in the Earth crust, and any superfluous consume would shorten its life span. He recognized that there is a continuous input of "low entropy" from the sun, but he though that this was negligible as compared to the requirements if we were to concentrate again the materials that we use and eventually disperse in the environment. Georgescu-Roegen has many followers between ecological economists, who do not however share in general his pessimism on the impossibility of a sustainable economy.

Here I begin by discussing Georgescu-Roegen ideas (many other points of view on his contributions can be found in Daly 1997) and then introduce some more points, in order to present a coherent view on the role of physical irreversibility in the interactions between economic systems and their environment.

Kaberger and Mansson (2001) criticized the idea that “low entropy” is necessary for value, by noting that everything whose entropy is not maximum has some “low entropy” and, hence, this is almost a meaningless assertion. However, the developments in chapt. 1 in this book support a relation between value and “low entropy”. According to most definitions, things are not considered to have economic value unless desired by some economic agent. Therefore, value implies mutual information (see subsect. 1.2.2) between a valuing subject and a valued object (the reverse does not hold, mutual information does not imply value). I show in chapt. 1 that mutual information at any level makes a negative contribution to physical entropy. As a corollary, value does imply “low entropy”. On the other hand, this negative contribution to entropy is fairly small. A worth considering point is that the flows of information between valuing subject and valued object are asymmetric (even the flows between two individuals who value each other can be shown to be asymmetric, if we do not aggregate artificially the different “paths” by which information circulates between them), so value is a far-from-equilibrium feature (see subsect. 1.2.4). Some consequences are (i) a huge difference between the “low entropy” implied in value and the much larger “low entropy” that has to be spent to obtain it, (ii) the requirement of a permanent or recurrent expenditure of “low entropy” if value is to be sustained across time, and (iii) the lack of a more specific quantitative relationship between expenditure of “low entropy” and gain of value (or even gain of mutual information) of general validity. The most trivial practical corollary of this set of results is that economies require and will always require an inflow of energy or, more properly, exergy (see subsect. 1.3.1), as an irreplaceable input.

Georgescu-Roegen’s idea on the fatal consequences of the dispersal of materials in the environment has been rejected by several authors (e. g. H. T. Odum 1996, Kaberger and Mansson 2001, Cleveland and Ruth 1997, and references in this last paper). H. T. Odum (1996, pg. 21) remarks that, in most natural ecosystems, the biota obtains its nutrients by re-concentrating them once they are more or less disperse in the environment, which is possible thanks to the expenditure of solar exergy. At some point in the future, seawater could well become one of the main sources of materials also for our society. The current inefficiency in the retention of materials in the economic circuit will not necessarily advance the halting of the economic process, but represents a transfer of huge exergetic costs to the future. Our descendents could well have to devote a substantial portion of their exergy expenditures to the re-concentration of some chemical elements of importance.

The particular form in which the wastes are released will have consequences for the future availability of resources, which cannot be explained with simple thermodynamic rules. This can be illustrated with the flows of nutrients in planktonic systems (see e. g. Margalef 1976). A routine function in these systems is the pass of several chemical elements from a dissolved to a particulate form, coupled to the dissipation of solar exergy (like we could imagine for a future economy with no more outstanding mineral deposits to tap). However, there is a considerable loss of available nutrients from the initial "wasteful" phases of succession until eventually reaching "sustainable" fluxes. A part of the nutrients absorbed by the biota is later released in forms that have a higher thermodynamic cost of retrieval: nitrogen is often reduced to N_2, phosphorous and other nutrients are displaced to deeper layers than the point where they were taken. In these cases, the "wastes" contain more exergy than the initial resources (due to its chemical form in the case of nitrogen and its spatial distribution in the case of phosphorous), so their increased thermodynamic cost of retrieval is not a direct reflex of the entropy production that took place when these were used. This is not strange if we take into account that we are dealing with nonequilibrium systems.

One fundamental error of Georgescu-Roegen was asserting that "while in the material environment there is only shuffling, in the economic process there is also sorting". All biological systems and, to a lesser extent, many other systems display a process of "sorting" coupled to increases of entropy in their environment, and have produced much higher levels of complexity than human economies, at least by now.

While the rate at which new exergy becomes available (as a principle, from solar radiation) sets a basic constraint on economies, and even the issue of material resources can be largely reduced to an issue of exergetic costs (despite the dissociation between exergetic costs and exergetic contents), this is by no means the only biophysical constraint, perhaps not even the main one. The main ecological problems that are currently unfolding have nothing to do with exergy shortages. The rates at which a complex system can be acted upon in any given way (including exploitation, pollution, etc.) without loss of complexity is limited, as explained in detail in chapt. 2. The time required to recover the former complexity (as a principle, through a “sorting” process external to the economic system) is ordinarily much longer than the time required to lose it, so losses are in this sense irreversible. I show in chapt. 2 that this is ultimately a consequence of the Second Law of thermodynamics or “entropy law” (see subsect. 2.3.1), but an indirect one. Losses of complexity are not necessarily solvable by using additional exergy. Instead, an injection of exergy in an ecosystem often acts as a disturbance producing further loss of complexity.

Loss of biodiversity, shrinks in the populations of particular species of interest according to some criterion (e. g. the collapse of fisheries), degradation of ecosystems, soil degradation and even the loss of traditional knowledge all belong to the category of irreversible information loss. Less obviously, climate change also does. The problem with the current anthropogenic forcing of climate is that it produces a loss in the mutual information between the Biosphere (human societies included) and the physical features of the fluid layers of the Earth. This mutual information is the result of a long evolutionary process with huge flows of information in both directions, and involves much complexity.

The recovery of substances dispersed in the solid matrix of ecosystems or in biomass is a much more complex operation than its re-concentration from water, if it is to be done without heavily disturbing ecosystems. This point may have a limited relevance in terms of resource recovery, but it has a major importance in terms of the abatement of some kinds of persistent pollutants once disperse in the environment. This involves informational irreversibility, in addition to exergetic costs.

6. 2. 2. Thermodynamic measures of value?

Some authors have tried to go much further than Georgescu-Roegen, by pretending to quantify the value of everything in thermodynamic terms (see references in Puntí 1988). Some of these proposals (not all) are a corollary of the hypotheses that pretend to explain the evolution of ecosystems as the search for the extreme of some thermodynamic function, which I reject in chapt. 2. These proposals stand in sharp contrast with the points of view of many other ecological economists, who consider that the value of different elements (here I use “elements” in the wide sense of the word, not in the chemical one) is incommesurable (see Martínez Alier et al. 1998).

In subsect. 6.2.1 I establish a qualitative difference between irreversible loss of exergy and irreversible loss of complexity. As I show in chapt. 2, both types of irreversibility are ultimately a consequence of the Second Law, but the one cannot be reduced to the other. This is a first reason to reject a thermodynamic measure of value.

Not only complexity cannot be reduced to exergy, but neither different complex features are exchangeable. In computational terms (see subsect. 2.3.1), the complexity of two states A and B is measured in relation to some state of reference C, and corresponds to the number of logical steps required to obtain each of them from C; the number of steps required to obtain A from C and B from C has no general relationship with the number of steps required to obtain B from A, so the fact that a system is in a complex state does not necessarily make it easier to attain another complex state. Not even the logical steps are always comparable, since a “logical step” is just an abstraction to be applied to qualitatively different processes. In other words, there is no physical measure that allows to compare fishes and trees, it is not possible to increase indefinitely the consumption of fish just by “proportionally” decreasing the consumption of timber, as would follow from the points of view expressed by some authors. Lack of exchangeability is not always associated to irreversibility, but it is in this last case when it poses major dilemmas, because then some exchanges are possible but only in one direction, and hence have long lasting consequences.

The authors who support thermodynamic measures of value should think, to be consequent, that all ecological problems would become superfluous if the obtainment of exergy by nuclear fusion surpassed the formidable scientific and technical obstacles that now faces (e. g. Parkins 1996) and this really meant an abundant and virtually inexhaustible stream of exergy. Given the above developments, this cannot be expected to hold. This can be illustrated with the spectacular failure to reproduce the ecological functions required for human life maintenance in the 1.27 ha enclosure called Biosphere 2, despite spending about $1 million in energy inputs per year (Cohen and Tilman 1996). Furthermore, given the interests that currently drive the economic system, we may wonder whether humanity is ready to assimilate a huge input of additional exergy, or would rather drive too large a portion of it to destructive activities enhancing environmental and social problems. In this case, the result would be comparable and as unsatisfactory as the result of injecting a large amount of nutrients into an aquatic ecosystem.

Puntí (1988) criticizes energetic values by noting that different forms of energy require different time spans to accumulate, but then suggests to use a single measure of energetic value that incorporates these differences, which fails to recognize that different kinds of elements are fundamentally inexchangeable.

A corollary of the pretension of a thermodynamic quantification of value is the practical proposal to establish prices solely determined by the thermodynamic features of the exchanged goods or services. E. g. H. T. Odum (1988) proposed to use prices proportional to "emergy" (see subsect. 2.2.2) in international transactions. Such proposals do not only ignore the lack of physical exchangeability between different elements, but also that prices should have other functions besides indicating the scarcity of natural resources. At the very least, these should also reward the individuals involved in the production process in a socially fair manner, which has no relationship with the energetic cost of their activities (some authors propose to value also human labor in thermodynamic terms, see Puntí 1988). Furthermore, prices can be used to direct consumption in some socially agreed manner, e. g. by making drinking water cheaper than water for swimming pools, while keeping an average price of water that reflects its scarcity (and not the scarcity of resources that are not exchangeable with water).

Although no aggregate biophysical measure can be used as a measure of value, some of them may still have some interest as gross indicators of the quantitative importance of the economic system as a subsystem of the Biosphere, and the magnitude of ecological inequalities between different human groups. A measure designed with indicative purposes and that currently receives much attention is the "ecological footprint" (Wackernagel and Rees 1996, Wackernagel et al. 1999, WWF 2002). As usually applied, it is quantitatively equivalent to an energetic measure of value. Its initial purpose was measuring the land surface required to sustain a given pattern of consumption. However, the inconvenience of summing up lands with different primary production per unit area PP/s has lead to give each kind of land a weight proportional to their PP/s, in more recent works. As a result, it has become an energetic measure, although expressed in units of surface area. It measures the primary production required to sustain a given supply of products of agriculture, cattle raising, forestry and fishing, plus the primary production that would naturally take place in areas occupied by human infrastructures (i. e. an energetic opportunity cost) and the primary production that would be required in order to satisfy the current consumption of exosomatic energy with fuel plantations. This last item is alternatively (and more often) quantified as the primary production of the area of forest required to absorb the CO_2 released if the exosomatic energy is obtained from fossil fuels, which gives roughly the same result. However, this last option is not recommendable. As noted by Herendeen (2000), while fuel plantations can be assumed to have a constant yield, the amount of carbon that forests can store is limited, unless these are reiteratively harvested and the timber preserved somewhere else. We can add that this amount is expected to decrease due to global climate change, mainly because the higher temperatures will enhance respiration and because large areas of tropical rainforest are predicted to become drier. Cox et al. (2000) forecast that terrestrial vegetation and soils, taken together, will become a net source instead of a net sink of CO_2 by 2050 (furthermore, these calculations do not include the fire transitions that I postulate in chapt. 3, which are likely to make things even quite worse). The current coincidence between both ways of measuring the ecological footprint is therefore contingent and transitory, and the fuel plantation criterion is more appropriate. In fact, as put by Ayres (2000), it would also be possible to take the surfaces required for renewable energy sources other than biomass: therefore, the aggregation of real consumption of current primary production with the consumption of exergy from other sources is the most clearly artificial step in the calculation of the ecological footprint.

6.3. System dynamics and organization in ecological economics

6. 3. 1. Beyond the balance of nature

The main consequence of the developments in sect. 6.2 is that ecological economics cannot be reduced to an energetic economics, it should be fully “ecological”. It should deal with the integration of the economic system in the Biosphere, with all of its irreducible complexity. Nevertheless, the Biosphere has never ceased to change, with or without humans: why then should we care for anthropogenic impacts?

Long ago, ecologists used to think that there was a “balance of nature”. One of the main roots of the concern for environment was the belief that humanity could break the balance when disturbing it. Currently, the idea of a balance of nature is not scientific anymore, but is still a popular one, and plays a role (Westoby 1997). The “ecological arrow of time” implied in the modified Margalef principle, as I formulate it in chapt. 2, could play the same role, while lying on more solid grounds. It means that, if we force the rest of the Biosphere to change too fast, the result will be destructive, regardless of the existence or not of any previous “balance”. As a principle, a destructive outcome is not desirable, neither from a biocentric point of view nor from an anthropocentric one, even accepting that some frequency of episodes of massive destruction is inherent to the natural dynamics of the Biosphere. From a biocentric point of view, it is unethical to add further destruction, regardless of whether or not we are the only source of destruction, in the same way as it is unfair to hit a person, even though he or she may receive other blows from elsewhere, and will indeed have other problems, and some day will even die. From an anthropocentric point of view, it is not relevant whether there have been and will be massive destructions in other moments of the history of Earth, we just do not want one now, not to make part of it. Even if our survival were ensured, we would not like to irreversibly lose resources that would otherwise contribute to our future welfare.

Perhaps mass extinctions make part of an adaptative dynamics at a large time scale (see subsect. 4.4.4), perhaps we could see the current one just as a first step (albeit the fastest and most irreversible one, and the only one our generation will see) in the adaptation of the rest of the Biosphere to the new world that humans are building. However, before keeping on in this way, we should wonder if this is really the world we want, if we will really be able to make part of it, and if it is worth such a price for the rest of species, and also make sure that we are not in fact exceeding the adaptative capacity of the Biosphere.

A recurrent question about conservation of nature (sometimes posed as an anti-conservationist rhetoric question) is how far in the past should we search the “natural” referent, which is difficult to answer in regions in which humans have been an active part of ecological processes for millennia (this is obvious e. g. in the Mediterranean, and less obvious but also true in many other areas). When there are no particular factors that make some option definitely preferable (see sect. 6.4 about criteria of desirability), a general rule could consist of making the least irreversible choice (are we still on time to recover pre-human "pristine" ecosystems but are going to lose the chance if we do not act now? is it instead the loss of traditional humanized landscapes the most clearly irreversible? which is the point between both options that leaves us with a most open future?).

Although there is no balance to break, some consequences of our interventions on nature have strong parallelisms with the intuitive idea of breaking a balance. A given modification in the environment may unleash an avalanche of other changes (often rapid and destructive). Furthermore, there are discrete thresholds such that, when surpassed, major ruptures take place. It is indeed essential to make the step from these qualitative appreciations to the development of practical tools to be able to perform quantitative predictions on the magnitude of the avalanches and the position of the thresholds. Some of the main contributions in this book refer to such tools (chapt. 3 and App 2A).

Things are in fact more complex, because catastrophes are not merely additive. Our disturbances on the rest of the Biosphere cannot be evaluated just in terms of their immediate consequences. In some cases, the traces left by current disturbances have an effect on the future regimes of catastrophes (or on the future regimes of fluctuations not strictly catastrophic), either enhancing or decreasing their severity. In chapt. 4 I give theoretical evidence that supports the idea that biodiversity losses make ecosystems more fluctuating and less adaptable. In chapts. 3 (and subsect. 4.4.4 and sect. 4.5) I give both theoretical and empirical evidence that some kinds of catastrophes display a self-organized critical dynamics (SOC), which make them sensitive to the record of past catastrophes and to external interventions, in a rather involved but predictable manner.

In chapt. 3 I give strong evidence of SOC in the particular case of wildland fires. As far as I know, there is no such strong evidence in any other field, but there are good reasons to suspect a SOC dynamics behind many other kinds of natural and socioeconomic catastrophes. There is a potential for a fast progress in this area by applying in these other cases the same tools that I have used for fires. Even the series of ecological changes that take place when producing the extinction or introduction of a species in an ecosystem could be an instance of SOC, although this is highly hypothetical (subsect. 4.4.4). Other possible instances of SOC (see sect. 3.1) can be found in earthquakes, landslides, storms... It is likely that the main calamities that have traditionally flagellated humanity – famines, epidemics, perhaps even wars –

have a SOC dynamics, so few things could be as important for global welfare as understanding it. A better understanding of the emergent phenomena covered by SOC theory may be much more useful in this respect than the theoretical constructs that are currently at the epicenter of the dominant discourse on welfare, such as the “Pareto efficiency” of the “economic equilibrium” in a “free market”... which is even more clear if we take into account that, as put by several authors, economic crises give reasons to suspect that also markets have a SOC dynamics rather than any “equilibrium” (see sect. 6.5).

The many papers that deal with the possible presence of SOC dynamics behind different kinds of catastrophes present it just as a post facto explanatory mechanism for the proportions of catastrophes of different magnitudes. Instead, in this book I put forward the idea that SOC systems can be managed (or mismanaged), and give some keys to deal with them. I treat in detail the instance of wildland fires (chapt. 3), and also outline some ambitious management proposals for the case that some of the main agricultural pests or diseases were found to be SOC, based on an intensive use of biodiversity (sect. 4.5).

As shown above, there are promising paths to readily advance our understanding of how the Biosphere will react in front of our interventions, and it is urgent to go ahead in such researches to be able to act in consequence. Nevertheless, the main recipe is indeed prudence: the “precautionary principle” is quite immediate once we acknowledge the modified Margalef principle. As noted by Myers (1995b), we could well find that some of the main environmental problems of our times are still unknown and unsuspected. What we can take for granted is that, if we pretend to modify our environment very fast, the results will be disastrous, unless we make an equally fast and effective effort to understand it and we find the ways to regulate the economic system in accordance to such knowledge. This stands in sharp contrast to the dominant point of view, which assumes that economic growth is inherently good and should be attained as fast as possible. Economic growth has to do with an increase in the average capacity of the economic agents to modify their environment (in a wide sense, including their social and economic environment). This seems good when thinking in simple additive terms, as conventional economics does, but not if we take into account the full web of interactions.

Besides the many injuries of all kinds that our society has already inflicted to the Biosphere, it has compromised its stability in its most essential facet, which is captured by Eq. 2.4. I postulate in sect. 4.3 that the really important facet of stability in any complex system is the capacity to sustain its complexity per se, while any particular features will change with time. This requires an organization such that there is no chance for a local fluctuation (with a non-negligible probability) to rapidly propagate throughout the whole system generating a major loss of complexity. The state of the system that would allow for such extreme fluctuations, a state that we may call “supercritical” (at least for some kinds of fluctuations), is rarely attained, because of the own impact of the large fluctuations that develop as the system approaches it. It is only attained if the system evolves toward supercriticality (or an equivalent state) very fast as compared to the frequency of the fluctuations, in such a way that these have no chance to perform their preventive function. In a period of rapid “growth” such as the current one, there is an increased chance of this outcome. For at least one feature of our world, it is obvious that this has already happened: the accumulation of nuclear weapons. A small local fluctuation (which reaches or has its origin in some minds or devices in a few decision centers) can produce the destruction of the whole world as we know it. There may well be much subtler features in which the Biosphere or, at least, the socioeconomic subsystem, has attained or may attain a comparable state. It is indeed necessary to detect all of them and revert their evolution toward a safer situation.

6.3.2. Critical thresholds

Some recent discussions on ecology and economics have paid a strong attention to discrete thresholds in the response of ecosystems to anthropogenic pressures (see e. g. Arrow et al. 1995, Pearce 1998). Several of the proponents of the so-called “strong sustainability” (see subsect. 6.4.3) consider that the recipe for a sustainable economy should consist of a set of “critical” thresholds that should never be surpassed, giving to the term “critical” a wider meaning than the one I use throughout this book. For the decision of such normative thresholds, it is indeed essential to know any objective threshold that would give rise to major ecological changes if surpassed, including any critical threshold in the physical sense of the word “critical”. Therefore, the methods to study criticality on practice, such as those that I develop in this book, are essential for “strong sustainability” to be operational, and also have a foremost importance from other points of view.

Scheffer et al. (2001) offer several instances of ecological processes in which a discontinuous change is expected as a given threshold is surpassed, in which the ecosystem suddenly loses its former resilience. In some cases the ecosystem is expected to gradually recover its former state: then, the resilience is only lost at a small time scale. In other cases, the former state is lost indefinitely. In physical terms, such discontinuous changes are called first order phase transitions (subsect. 2.3.3). Perhaps the best known instance of first order phase transition in the global environment is the discontinuous change in ocean circulation patterns (Rahmstorf 1995) that is expected as a result of global climate change.

In this book I pay more attention to second order phase transitions (see subsect. 2.3.3, 3.2.1, 3.4.4 and 3.4.4), which have the form:

$$q = a\left| r - r_c \right|^c \qquad (6.1)$$

in which q (q <0) is the environmental parameter that changes as r is modified. Although in Eq. 6.1 q diverges as r approaches the critical value r_C, this effect is sometimes limited by external constraints: then, q will remain finite at the limit $r \to r_c$. The parameter c establishes the degree of abruptness of the transition, so we have a continuous set of possibilities from $c \to -\infty$, which corresponds to a simple exponential response of q to r, to $c \to 0$, which is equivalent to a first order phase transition. As shown throughout the book, second order phase transitions have a major ecological importance.

While the point in which a first order phase transition takes place can only be predicted from a detailed knowledge of the system or from the experience of previous phase transitions, the parameters of Eq. 6.1 can as a principle be obtained phenomenologically without need of a second order phase transition having ever taken place. This requires the analysis of the relationship between r and q in historical records (of course there is never full guarantee of extrapolability). In subsect. 3.4.3 (complemented by App. 2A) I put forward statistical tools for determining these parameters.

Unfortunately, the prediction of second order phase transitions from phenomenological information has one inherent limitation. The closer is c to 0, the higher is the need of data to obtain reliable predictions, while the worse will be the consequences of prediction failure. The limit case is that of first order phase transitions, which cannot be predicted phenomenologically. Therefore, phenomenological analysis will often have to be complemented with a detailed knowledge of the underlying processes.

In sect. 3.6 and subsect. 3.7.4 I treat in detail the case of the second order phase transitions that seems to take place in tropical rainforest regions under anthropogenic pressures, which are major catastrophes with a foremost importance. The same developments could be useful in very different fields. E. g. it would not be strange to find that the extinction of some given species subject to external pressures has the form of a first or a second order phase transition. In this case, the above formalism could perhaps be more useful than more detailed population dynamic models, when setting limits to the tolerable impacts on the species.

Margalef (2001, and pers. comm; see also Margalef 1991 pg. 152) states that one of the most important impacts of humans on the rest of the Biosphere, if not the most important, is the “topological inversion” of landscapes, from a matrix of mature ecosystems with disturbed patches to patches of mature ecosystems in a matrix of anthropogenic infrastructures. Habitat fragmentation is thought to be one of the main causes of biodiversity loss (e. g. UNEP 1995), besides causing many other problems (Brokaw 1998). It is well known that, at least in random landscapes, the effect of habitat loss on fragmentation has the form of a second order phase transition (see e. g. Stauffer and Aharony 1992 about the physical basis, and e. g. With and King 1997 or Bascompte and Solé 1996 about ecological applications). When the fraction of habitable area descends below a critical threshold, the largest patch of continuous habitat passes from having the form of a matrix that spans the whole landscape, regardless of its size, to having a limited size (in the most usual models, the “topological inversion” is not compressed to this single point, because for them the point at which the habitat loses its continuity is not the same as the point at which non-habitable areas become continuous). Beyond the threshold, the average and maximum size of the fragments readily decline when decreasing the portion of habitable surface. For the sake of realism, authors such as O’Neill et al. (1992) or With et al. (1999) have tried to extend this result to landscapes with fractal correlations by means of simulations. My findings in subsect. 3.4.3 provide a more complete picture of what happens in this case, and some tools to quantify it on practice. The variation in the sizes of fragments will be more gradual, but there will still be a critical threshold. Due to the less abrupt transition, there will be a significant difference between the critical point and the point at which the largest fragment completely embraces a large but finite landscape. Therefore, the size of a each particular simulation model will have an effect on the results, which should be taken into account in future works in this field.

In fact, fragmentation is largely due to transportation infrastructures and the activities that develop around them. Due to their elongated form, these infrastructures constitute a heavy factor of fragmentation unrelated to the above critical transition in any of its forms. However, there are cases in which we should pay attention to the effects of habitat loss by itself on fragmentation. An illustrative instance is provided by the discussion in Brazil on the fraction of their properties that landowners should keep forested by law, in Amazonia (Rohter 2001).

6.4. Ultimate goals and criteria

6.4.1. The monetary measure of value, and other options

Economic decisions cannot be taken just on the basis of scientific knowledge. This may allow to predict the consequences of our actions, but such consequences should be evaluated in some way. The biocentric and anthropocentric criteria that I mention in subsect. 6.3.1 are two of the options, but just as basic points of departure. Nowadays, the orthodox method of economic decision is the cost-benefit analysis in monetary terms, which is a particular way to make more specific the anthropocentric criterion. Many elements and features of the environment are not exchanged in markets and, hence, have no market price, but environmental economists try to extend to them the conventional economic criteria, either by proposing ways to create markets for them or by trying to estimate the prices they would have if there were such markets and designing policies consequent with such estimations. Environmental economists differ from ecological economists because the former try to tackle ecological problems on the basis of conventional economic theory, while the lasts try to reformulate the economic theory by taking ecology into account. The monetary values estimated by environmental economists (ecological economists have also occasionally entered the game of putting prices, see Costanza et al. (1997) and the ensuing debate in Costanza 1998, and also Balmford et al. 2002) are presented as if they were objective measures of value, and are even published in scientific journals. In fact, there are strong political choices behind the use of the monetary value, which are rarely made explicit. These are sometimes accompanied by scientific misunderstandings. Here I briefly comment the assumptions behind the monetary measure and the consequences of introducing modifications into these assumptions.

Consider a single individual that has a given amount of money. She spends it in such a way to obtain some quantities x_i of different elements *i*, where each x_i may be measured in different units. Each *i* has a given monetary price p_i per unit of measure. One of the foundations of neoclassical economics is the mathematical finding that, if the individual distributes her money in such a way to maximize some (continuous and derivable) magnitude *U*, which is called utility, then

$$\frac{\partial U}{\partial x_i} \propto p_i \,,$$

where $\frac{\partial U}{\partial x_i}$ is called marginal utility. Prices thus allow to measure every element available in the market with a single unit, which is the monetary one. From the point of view of a single individual whose behavior maximizes something, the monetary measure allows to compare marketable things in two different ways: (i) in terms of exchangeability in the market place, and (ii) in terms of their marginal contribution to fulfilling a given goal.

Neoclassical economists take monetary values as an universal measuring rod, which is not only applied to decisions that concern one single individual but also for aggregates of individuals and even the whole world, and even to elements external to the market (in the best case; most often, these elements are simply ignored). This requires the extrapolation of one or both of the meanings that the monetary measure has for the individual utility maximizer.

The meaning in terms of exchangeability may be applied to some extent for individuals or for partial aggregates of individuals such as countries, which can participate in a market, but not for the world as a whole, exactly for the same reasons that carried me to reject the thermodynamic measures of value in subsect. 6.2.2. The failure to recognize the lack of exchangeability will pose major problems when it is associated to irreversibility, because it will lead to an underestimation of the future consequences of current decisions. Physical irreversibility, either in terms of entropy or complexity, poses no general constraint on single subsystems that have exchanges with other subsystems (Eqs. 1.19 and 2.1), but is an inescapable constraint for the set of interconnected subsystems as a whole. The Earth is nearly closed in terms of information (and matter), and has a constrained rate of entropy outflow. Therefore, there are relevant irreversibilities for the Earth as a whole, which do not necessarily rule also for its subsystems. Nevertheless, the monetary measure is used for global issues not just because of its relationship with some purported goal, but also in terms of exchangeability.

The authors of the chapter about intertemporal equity (Arrow et al. 1996) of the 1995 report of the Intergovernmental Panel on Climate Change distinguish two approaches to the issue of the discount of future (see subsect. 6.4.2), which they call “descriptionist” and “prescriptionist”. “Descriptionists” prescript a valuation of the consequences of climate change applying a discount of future equal to current interest rates in the market, while “prescriptionists” give a higher weight to climate problems. According to the authors (two of them are currently Nobel laureates), “the descriptive approach can be interpreted as maximizing the economic resources available for future generations and allowing them to decide how to use the resources”. This criterion could make sense for an individual or an enterprise, but not for the world as a whole. It is equivalent to saying that we should ensure that future generations owe many goods and services and it is up to them to decide whether they want to “sell” these goods and services to “buy” an undisturbed planet.

While irreversibility disqualifies the use of monetary or any other unit in terms of exchangeability, it is not as a principle an obstacle to the use of measures of value in terms of desirability. However, this poses serious problems, in particular when the chosen measure is the monetary one. This issue has absorbed tones of ink and there is not much to add, but it is worth including a few comments here.

Certainly, the price of things integrates the decisions of many people and reflects their preferences, but it is a biased reflex. The purported relationship between price and marginal utility of single individuals cannot be extrapolated to a whole society unless one particular criterion is chosen to aggregate utilities, which roughly consists of assigning to each person a number of “votes” equivalent to his or her income (see e. g. Brendt 1998). As a result, e. g. one human life in a high income country has been valued 15 times as much as another one in a low income country (see Pearce 1995). The use of neoclassical economic criteria for political decisions is equivalent to breaking down any separation between political and market powers, for the benefit of the last. Some economists and institutions recognize the option of using weighting factors to compensate for the differences in income when carrying out a valuation (e. g. Brendt 1998), but this is not performed on practice (Drèze 1998). Indeed, unequal participation in economic decisions within households is neither taken into account. The choice of the weight given to the preferences or the welfare of each individual is not a scientific issue, so neither the choice of the monetary measure is. Science cannot decide for whom we care when assigning values, cannot decide the weight that we give to the interests of individuals of different social classes, genders, generations, nationalities, races, species...

Furthermore, there is little scientific evidence of a strong relation between prices and contribution to an individual's welfare in real psychological terms, which is the implicit or explicit basis to use it as a criterion of desirability. It is well known, that choices in the market do not only depend on effectiveness in producing welfare, but also on factors such as publicity, addictive effects, positional effects, etc., which many people would recognize when dealing with collective decisions, despite being individually affected by these biases.

Once we have decided for whom we really care, it is as a principle possible to build some measure analogous to the monetary one, but capturing different interests and correcting its multiple biases. Some authors have studied how the relation between monetary value and individual "utility" changes as a function of income, or have proposed weighting factors to take inequality into account in practical decision making. If we added some corrections for the psychological biases at the individual level, and paid the due attention to the future consequences of irreversible losses, we would have an alternative to monetary value sharing its essential structure. However, not even the kind of structure of the monetary unit is always appropriate.

When performing the monetary valuation of the outcome of some given action, the result is assumed to have the form

$$dU = \sum_i p_i dx_i \qquad (6.2),$$

i. e. each element i is assigned a given value p_i per physical unit x_i of the element, and the increases of value due to the variation dx_i of each element i are added up. Both if the measure of value is monetary or it has other units related to welfare, this linear and additive schema will not always be realistic. I wish to highlight that this schema can be seen as the result of taking only the first-order terms in a Taylor series expansion, which is likely to be often (although not necessarily always) valid for asymptotically small changes in the set $\{x_i\}$, while it has no basis when applied to large changes in $\{x_i\}$. One given choice could involve large changes in $\{x_i\}$ when (i) there are large time scales involved (see subsect. 6.4.2), (ii) there is a high sensitivity of $\{dx_i/dt\}$ to our choices, such as in the neighborhood of a phase transition (see subsect. 6.3.2), or (iii) the choice directly involves a huge socioeconomic rearrangement. It would not be surprising to find Eq. 6.2 to fail without need of looking at such extreme cases. The idea of a Taylor series expansion allows to formalize, on the one hand, the conventional economic procedure, and on the other, the intuitive view by Kapp (1983, quoted in Martínez Alier et al. 1998), who says to be "inclined to consider the attempt at measuring social costs and social values simply in terms of monetary or market values as doomed to failure", between other reasons because "they are borne and accrue to society as a whole". We could try to go beyond and estimate dU (after the due ethical choices) by explicitly taking into account higher order terms in the Taylor series. This extension is mathematically straightforward, but adjusting the parameters will often be exceedingly complex.

While a corrected welfare predictor may work as a practical measure of value in some cases, in others it should be taken just as a conceptual referent that expresses the ultimate goal of the decision process. When a decision maker or set of decision makers (which may even consist of the whole community) has to integrate much information of qualitatively different kinds in order to reach an optimum decision, trying to formalize the whole decision process in explicit numbers will very often be a burden and a source of biases instead of a help, overall when Eq. 6.2 does not hold. In these situations, the best option seems to be the use of non-algorithmic multicriteria decision methods, as advocated by Martínez Alier et al. (1998).

The exercise of putting monetary prices to nature should not necessarily be abandoned, but should be seen just as a way to show that, even under the rules of conventional economics, the natural environment is much more relevant than assumed by the own conventional economists, and by most policy makers.

A major consequence of not discriminating between people when assigning values (both if we use some quantitative procedure or if we do not) would be giving the maximum priority to the maintenance of the biophysical support systems necessary to satisfy basic needs, and to their effective satisfaction, while current conventional valuations give an enormous importance to consumerist goods and services largely addressed to a minority. Therefore, an economic science addressed to satisfy the social needs in a more equitable way than the current one would necessarily be more “ecological”. Even if the current and future satisfaction of basic needs were ensured, consumerist goods and services would probably be less important for welfare than a healthy and pleasant natural and psycho-social environment. The untested neoclassical dogmas should be replaced by a serious research on the determinants of welfare, free from preconceptions. Some authors are already working in this line, from different points of view (see Ackerman et al. 1997).

6.4.2. Caring for the future

Many of our impacts on the environment are irreversible, in the sense that the time required for their “healing” largely exceeds the time required to inflict them. The time scales of our environment span many orders of magnitude. The characteristic times of individual organisms range from hours to hundreds of years or even millennia (in the case of *Sequoiadendrum giganteum*), gorgonias such as *Paramuricea clavata* have a growth rate of 15 mm per year (Ros 1989), we can talk of plankton succession at a scale of months (Margalef 1989 pg. 791) while forest succession often requires several hundred years (e. g. Romme 1982), we will require centuries to recover the agricultural soil that we lose in a decade (if the numbers by Pimentel et al. 1995 are correct), the main greenhouse gases remain from ten to hundreds of years in the atmosphere, ice sheets will keep on melting down for millennia once the global climate is stabilized (IPCC 2001a), the half life of different radioactive isotopes produced by nuclear plants range from fractions of second to millions of years (e. g. Lenssen 1992), the recovery of global biodiversity losses may require about ten million years (Kirchner and Veil 2000)...

There is a long-lasting discussion on the way we should value the long-term consequences of our current actions. The answer of conventional economics is the “discount of future” (e. g. Arrow et al. 1996, Hanley and Spash 1993). If we deposit an euro in a bank, it will multiply exponentially, in agreement with the current interest rates. An investment addressed to obtain some monetary gain in the future is only profitable if this gain exceeds the result of the simple multiplication of money in the bank. Therefore, the comparison of monetary benefits expected at different moments in the future is performed by weighting them with a term that decreases exponentially with time, at a rate that is called "discount rate" and equals the current interest rate. The discount of future is the fact of giving a lower weight to gains that require a higher waiting time. As a corollary of their assigning monetary values to the environment, environmental economists stand for the application of the conventional discount of future when valuing the future environmental consequences of our current actions. It is alleged that, under the usual neoclassical assumptions for a "perfect" market, the interest rate should capture the marginal value that the society assigns to the future. The result is that, as a principle, nothing is considered to matter beyond some decades, at the very most (e. g. a discount rate of 5% a year means that things lose half their value in 14 years). As a principle, the multiplicity of time scales of natural processes is replaced by a single time scale, determined by the current interest rate. On the other hand, within this logic, the relevant time horizon should be allowed to increase if there is some well-defined rate of variation of prices: the environmental economist Krutilla (1967) suggested that this is the case of "natural amenities", which would systematically increase their "price" as economic growth proceeds, but even this moderate amendment is rarely taken into account.

Ecological economists have criticized the discount of future for a number of reasons (e. g. Martínez Alier 1999, Howarth and Norgaard 1993, Martínez Alier and O'Connor 1996). The presumed determinants of the societal time preference, which would ideally be captured by interest rates (which, in fact, are also impacted by financial policies) are (i) a pure preference for the present or "impatience", (ii) a probability, increasing with time, that the consumption is not realized, together with the limit set by one's finite life span, and (iii) a marginal utility of income decreasing as income increases with economic growth (more often, the second point is not stated as an independent one). The legitimacy of each of these components in terms of environmental policy has been rejected by a number of authors. These have remarked that the first and, largely, the second, imply a discrimination between generations, and have denied the relevance of the third in the case of environment.

The idea of an exponential economic growth keeping on indefinitely in the future is a bold prediction rarely recognized as such. It ignores the environmental feed-backs and also our little knowledge on endogenous economic fluctuations (some of their main features are just beginning to be uncovered by econophysicists). Generally, future variations in environmental “prices” are also ignored, which is not strange if we take into account the little awareness on the limitations to exchangeability. And indeed the huge biases and uncertainties in the relationship between monetary measures and real welfare are also ignored. This enormous accumulation of overlooked biases and uncertainties of all kinds on the economic side sharply contrasts with the extreme reliability that is required to environmental predictions for these to be accepted as a justification for monetary expenditures or even “opportunity costs”.

The lack of reliability does not really justify an outright rejection of the prediction of economic growth as usually measured remaining positive (although necessarily at variable rates) for a fairly long period. However, there are clear reasons to fear a decrease in human welfare in a near future, if we take into account that its main material condition is the fulfillment of basic needs. Despite the fast economic growth that began with the onset of industrialization by the middle of the 18th century, anthropometric studies (Komlos 1998) indicate that this was accompanied for a long time by a worsening in the conditions of life, in the then industrializing countries. According to these studies, the preindustrial level of satisfaction of physiological needs was only recovered after the World War I. We may wonder if this late improvement would have ever taken place without the strong social fights that preceded and accompanied it. The only important gain in human biological indicators during the early industrialization was a decrease in infant and child mortality due to a retreat of smallpox, of uncertain causes (Bengtsson 1998). Later, during a large part of the 20th century, economic growth was globally accompanied by an increase in the standards of life (although there are strong local decouplings between both, see Sen 1993), which fed the current faith in conventional “progress”. Unfortunately, mortality has recently raised due to the AIDS pandemic and the collapse of Eastern countries, between other factors (e. g. Myers 2001). Furthermore, the increases in global inequality and the dismantlement of social services pose an immediate threat to the fulfillment of basic needs for a large portion of the world population. A little less immediately but more irremissibly, basic needs are threatened for ecological reasons (e. g. Brown 1997, 2002). As a result of policies addressed to the short term, the improvement in the satisfaction of basic needs across most of the 20th century has been accompanied by a series of changes that will make really difficult the continuity of such improvements: huge losses of arable soil (there is an outrageous uncertainty in this vital issue, as apparent from the wildly diverging estimates by e. g. Pimentel 1995a and b, Crosson 1995, Trimble and Crosson 2000a and b, Nearing et al. 2000, Wicherek 1994, Wood et al. 2000; about the relationship between soil loss and warfare, see Bächler 1994), exhaustion and salination of aquifers and pollution of water (e. g. Prat 1993), sweeping aside of agricultural biodiversity while pests become increasingly resistant to pesticides (e. g. UNEP 1995), collapse of fisheries (Botsford et al. 1997, Watson and Pauly 2001), loss of the products and "services" of natural ecosystems and natural biodiversity (e. g. UNEP 1995), climate change (with consequences such as (IPCC 2001b): shrink in potential crop yields in most tropical and sub-tropical regions, and also at higher latitudes beyond some level of warming; increasing drought in several regions already subject to water stress; expansion of tropical diseases; degradation of natural ecosystems; 70-200 million people per year suffering coastal storm surges by the 2080s, for a mid-range scenario; increases in several kinds of extreme events; chance of some major rupture in the global atmospheric-oceanic system), increasing resistance of bacteria to antibiotics (Cohen 1992), a high dependence on oil (several authors warn on a turning point in oil supply in a few years (Kerr 1998, Young 2001); its progressive exhaustion, combined with its heterogeneous distribution in space, has already motivated major wars), and an enormous and still growing human population, due to the maintenance of socioeconomic conditions that encourage a high birth rate in large regions of the world (in part, by the blockage of basic social services and of other social improvements, see e. g. Sarkar 1993). Furthermore, inequality may produce a competition for scarce resources such as fertile soils between satisfaction of basic needs and the production of luxury goods such as livestock feed (which currently absorbs 40% of grain production, WRI 1997) or biofuel (see the criticisms by Martínez Alier 1984 to Brazil's biofuel program). This accumulation of self-posed handicaps does not suffice to ensure that we are going to be worse off in a middle term, but gives good reasons to suspect that the only way to avoid this outcome is to readily stop environmental degradation and to establish a vigorous redistributive policy (about the relation between environmental limits and distributive issues, see also Martínez Alier 1984; about the ensuing conflicts, see Martínez Alier 1994). The currently dominating political winds do not blow in this direction.

Since human welfare may well recede in a near future largely as a consequence of ecological problems, many of the components of the environment are likely to increase their marginal value for human welfare instead of decreasing it. Therefore, there is no reason for applying a positive discount rate to the environment if our concern is human welfare without distinctions between humans, and in fact there are reasons to apply a negative discount within some time window (the period for which we expect welfare to decrease). In accordance with Krutilla argument, even “environmental amenities” that do not contribute to the satisfaction of basic needs can be expected to increase their marginal value across time as compared to consumerist goods and services.

On the other hand, we should also take into account the uncertainty that surrounds our distant future and, therefore, the long term consequences of our current actions. There are indeed huge uncertainties about the remote future of the Biosphere, in particular those introduced by the evolution of the own human society. It is not impossible that this finds cost-effective ways to accelerate some features of environmental regeneration or to replace environmental functions, neither that this generates such a burden of problems to make environmental “services” more necessary than ever before. Neither we know how much will humanity last (which is not so important if we do not just care for human welfare). Some aspects of uncertainty represent a further argument against discounting, while others could justify some kind of discounting in some case. Here I will not study in detail which are the aspects that belong to each category. However, even if we accept some discounting, this should never represent a vanishing of value at a finite time scale. It is perfectly reasonable to assign an appreciable probability to e. g. current losses of biodiversity or current radioactive pollution still having a cost 10^5 or 10^6 years from now, while for conventional economic criteria nothing matters beyond a few decades. If we are to apply some discount of future on the basis of uncertainty, not only should it be much lower than pretended by conventional economists but also non-exponential. Due to the variety of time scales involved in natural, social, and even individual evolution, it is often recognized that any such evolution is generally closer to a $1/f$ noise than white noise (Gardner 1978). This makes me think that a discount pattern close to a power law will be more appropriate than an exponential one, which would do for a Markovian world. Although we cannot ensure that this is the explanation, it is interesting that behavioral experiments show that both humans and other animals display non-exponential patterns of discount of the future (with a fatter tail) (Henderson and Sutherland 1996), in dissonance with the expectations of conventional economists, who have generally assumed that "rational" individuals would discount exponentially, because otherwise they have to reconsider their past decisions across time.

In the line of the Taylor series reasoning in subsect. 6.4.1, one consequence of taking long time horizons is that it is not possible to ignore nonlinearities anymore, neither in the prediction of the objective outcomes of current actions nor in their valuation. In fact it is often incorrect to ignore nonlinearities even at conventional time scales, most obviously where there are phase transitions involved. Indirect, nonlinear effects in the outcomes of current actions are not only due to ecological nonlinearities, but also to nonlinearities in the interaction between ecological and economic systems and to nonlinearities within the own economic systems (see sect. 6.5).

It has been considered "inconsistent" to invest in the avoidance of environmental damages by applying a discount lower than the market's one, when the same amount could be invested in sectors with a higher return rate. Such "paradoxes" arise from ignoring the nonlinearities and the multiplicity of time scales (when they do not result just from an undervaluation of the environment due to the use of monetary units). A given expanding industry will have high return rates, but these will decrease as soon as the market is saturated (as would be captured by a second-order term in a Taylor series). A lack of investment in this industry may mean that it will require some more years to reach the saturation level. Our additional investments in this industry will provide benefits, but only for a limited time. Meanwhile, the lack of environmental investment may have caused environmental losses that will require centuries to be recovered. While the rate of return of the industry was larger than the environmental one when measured within a given time window, it could be smaller if measured within a larger time window. The own idea of a single return rate for a given project, like a single discount rate, is senseless unless we assume a linear model that will only be realistic if we just do not care for the long term. Otherwise, our models will have to be nonlinear, and we will have o be very careful when comparing phenomena that take place at different time scales.

We can further illustrate the issue of nonlinearities with the case of climate change. A conventional cost-benefit analysis of an investment to reduce greenhouse gas emissions will weight (a conservative estimate of) the direct (monetary) benefits in terms of avoided climate damage against the (purportedly secure) benefits of an equal deposit in the bank. The nonlinearities due to phase transitions in the climate system are ignored, because none is considered to be really likely before one century (after 2100, the whole marine circulation pattern could well suffer a drastic reorganization, with huge climatic consequences, IPCC 2001a), so the discount of future allows conventional economists to minimize their importance (the fire transition in tropical rainforests treated in chapt. 3 has still not entered the debates, and other major nonlinearities are still not seen as well known enough to be included in the analyses). It is also ignored that the marginal environmental cost of adding further greenhouse gases to the atmosphere will increase across time, as their concentration increases (at least before and between major transitions), so a huge reduction in greenhouse gas emissions may be eventually unavoidable: linear cost-benefit analyses do not count with the future costs of restructuring the economy if now we chose a high-emission development path (or the future environmental costs of not restructuring it because this becomes too costly). Neither they count with the fact that an increase in such emissions in a high income country increases the marginal environmental cost of a comparable increase in a low income country, where it would be likely to produce higher social benefits (this is another way to say that one of the reasons for high income countries to decrease their emissions is leaving a margin for low income countries to carry out some increase in their own emissions, as put e. g. by Goodland and Daly 1993).

6.4.3. Sustainability: which?

There is an increasing perception that an improvement in welfare at a given time scale may be associated to a worsening at a higher time scale, if the first involves too many irreversible losses of natural resources and environmental “services”. This has lead to an apparent consensus that special care should be taken for “development” to be “sustainable”. The expression “sustainable development” was introduced by IUCN (1980) (Becker 1997) and developed and popularized through the Bruntland Report (World Comission on Environment and Development 1987). The need of a “sustainable development” has been recognized by most of the states of the world in the Summits of Rio and Johannesburg. For most governments, however, “sustainable” is just an adjective that is added to the word “development” (or even “growth”) in economic programs that do not reflect any special concern for the future of the environment. Then, “sustainable” becomes an empty word used for propagandistic purposes. Furthermore, the authors who take sustainable development seriously have found great trouble when trying to find out a satisfactory definition more concrete than Brundtland’s. Most of these authors are polarized around two points of view: those who stand for “weak sustainability” (e. g. Pearce and Atkinson 1993) and those who stand for “strong sustainability” (e. g. Faucheux and O’Connor 1998). Weak sustainability is supported by some environmental economists, and regularly assessed at a country-by-country level by the World Bank. For a country (or the world as a whole) to be weakly sustainable, each generation should leave to the following one a nondecreasing amount of capital per capita, where the “capital” is measured by aggregating in a single monetary variable all kinds of natural and human-made capital. In contrast, the proponents of “strong sustainability”, who are mainly ecological economists, support a multivariate approach to sustainable development. There is not a single criterion of strong sustainability as specific as that of weak sustainability, but the particular recipe most quoted in association with the term “strong sustainability” consists of the establishment of a set of “critical thresholds” that should not be surpassed. Although weak sustainability is largely rooted in conventional economics, both weak and strong sustainability have been criticized from the strict economic orthodoxy: Beckermann (1994) heavily argued that sustainability, however defined, cannot replace optimality as the ultimate criterion of economic decision, and is often incompatible with it (see the ensuing debate in Jacobs 1995, Skolimowski 1995, Daly 1995, Beckermann 1995, Common 1996, Holland 1997).

In a highly insightful paper, Costanza and Patten (1995) put forward an original point of view about sustainability. For them, a sustainable system is one which survives or persists, but a simple quest for persistence is senseless, because we should take into account “that one must look at systems and subsystems as hierarchically interconnected over a range of time and spatial scales, and that each of these systems and subsystems has a necessarily finite time span”. In particular, “larger systems have longer time spans because their component parts have shorter time spans which allows the system to adapt to changing conditions”. Therefore, instead of sustaining “per se”, we should decide what we find and we do not find desirable to sustain, with the condition that our choices about each subsystem have to be “consistent with the relationship of these subsystems with other subsystems in the hierarchy” and taking into account that their “sustainability cannot mean “maintenance forever””. The view of nature expressed in my own book is close to the view put forward by Costanza and Patten, and I agree with the main conclusions of their paper. On the other hand, I find a more dubious interest in their assertion that “a sustainable system in this context is one that attains its full expected life span within the nested hierarchy of systems and subsystems within which it is embedded”, because, in fact, systems of a given kind and scale do not only have some given expected “life span”, but also a great variability in this “life span”, which also has a role in terms of adaptability.

Costanza and Patten view of sustainability is interesting but still too fuzzy to be considered operational. We should have some way to decide what we would like to sustain and to know whether this is “consistent with the relationship of these subsystems with other subsystems in the hierarchy”. Some of the developments in the rest of this book are useful in terms of the second part. In chapt. 5 I give some clues on the relation between system function and the diversity of their components. In chapt. 4 I show that some aspects of the dynamics of ecosystems (and, probably, of economic systems too) should be understood in terms of SOC, which involves a given pattern of relationships across different spatial and time scales. I give tools to predict how SOC systems will react in front of interventions at different spatio-temporal scales and to manage them. In other cases, the fate of different subsystems could be less interrelated than in a SOC system, but we still have to take into account which are their characteristic time scales, in order to know how far into the future will last our interventions on them.

When deciding what to sustain, it would indeed be useless to try to sustain indefinitely every detail of a system (what Daly (1995) calls “absurdly strong sustainability”), but it is no absurd to try to sustain as long as possible some global features. At first view, weak sustainability may seem to be consistent with Costanza and Patten point of view, because it just pretends to sustain one aggregate feature, “capital”, while allowing for every change in the details of this “capital”. However, weak sustainability fails to recognize the multiplicity of time scales involved in nature and society, because it is measured at a single time scale, the time scale of a human generation. In what follows I give more detail on the failure of the weak sustainability concept in this and other respects, and depart from this criticism to propose a new criterion of sustainable development.

Before any other thing, it should be clear that a huge problem with weak sustainability is the fact of being based on a monetary measure. As apparent from subsect. 6.4.1, this implies a particular set of political choices. It implies that the "capital" it measures will not be the most relevant for human welfare. Since the monetary measure is acceptable as a measure of exchangeability between different elements from the point of view of single countries but not at all for the world as a whole, the index of weak sustainability (the variation of the total "capital" per capita from one generation to the next) has interest to have some idea of the "sustainability" of one country's comparative economic position in relation to the others, but not of the contribution of the country to global sustainable welfare. The index of weak sustainability is an extension of the net domestic product to include natural assets and other assets that this does not take into account, which indeed improves it: this index may have some use in allowing countries to perceive that caring for environment and natural resources usually renders benefits specific to the own country and not only benefits shared with the rest of the world, which is often appreciated. However, it will be enormously misleading if used as an orientation to reach global sustainability, since the countries with the heaviest impacts on the global environment are those that have the best scores in weak sustainability, due to their high accumulation of human-made capital (see Pearce and Atkinson 1993).

Pezzey (1989) generalizes the concept of weak sustainability by referring to "welfare" instead of capital. This does not suffice to solve all of the problems of weak sustainability, but it is a first step. If we want to work out a definition of sustainable development addressed to human welfare, one of its ingredients will have to be an unbiased measure of average human welfare $\overline{w}(t)$ at time t (or the average welfare of the set of organisms for whose welfare we care, if it is not just humans). In fact we have no good measure of $\overline{w}(t)$ and, even if we had one, our capacity to perform quantitative predictions about it would be minimal. Therefore, I do not expect to ever feed any number to the equations that I develop below departing from $\overline{w}(t)$. I will use these equations just as a way to express a conceptual proposal in a rigorous way.

The fact that weak sustainability is measured at the time scale of a single generation is a strong limitation of this concept (either in terms of capital or welfare), because the consequences of irreversible changes of different kinds span very different time scales (in other words, the world is not Markovian). The own Pearce (1999), who can be considered the main proponent of weak sustainability, recognizes that irreversibilities should be treated "with considerable precaution", but does not seem to be aware that this fundamentally impairs the usefulness of his criterion, since several of the most important environmental problems do involve long time scales.

It is indeed possible to ensure that "capital" will not decline in a given period, while sowing its posterior decline. E. g. the responsibility for a phase transition should be equally attributed to each of the previous steps that allowed to approach the transition point, but only the last one will have short-term consequences. The long-term effects of the current loss of nonrenewable resources are not necessarily captured in their current prices (see Howarth and Norgaard 1993). Such delayed outcomes in the environment and the natural resources will in addition represent delayed effects on the value of human-made capital: the accumulation of forms of human-made capital whose function depends on the availability of oil and of sink capacity for CO_2 cannot compensate for the future exhaustion of these forms of "natural capital".

Not only the legacy of one generation may pose serious obstacles to some other generation to satisfy the rule of nondecreasing capital, or nondecreasing welfare. The own idea of a nondecreasing welfare (let alone capital) is not satisfactory enough. Take the case of radioactive wastes, which may last for hundreds, thousands or millions of years. Since the problem is distributed across such a long period, a temporary use of nuclear energy by, say, a single generation, may not pose a great obstacle to the attainment of a nondecreasing welfare. However, the use of nuclear energy by this single generation could have destroyed much more (future) welfare than the welfare it has produced. If this is the case, the "production" of welfare by using nuclear energy should be considered unsustainable. The point is that, in order to have a fair definition of sustainable development, we should consider that the variable to be sustained is not welfare per se, but a variable f that integrates the expected welfare for the current generations and all of the generations that will come after it:

$$f(t) = \lim_{c\to\infty} E\left[\frac{\int_t^{t+c}\left(N(t)\overline{w}(t)\right)dt}{\int_t^{t+c} N(t)dt}\right] \qquad (6.3),$$

where E means statistical expectation (after assigning some probability to each possible "future", conditioned to our current decisions) and $N(t)$ is the population at time t. f captures the consequences of our current acts at all time scales, not just the time scale of one generation. Therefore, I propose that, for "development" to be considered "sustainable", it should satisfy

$$\frac{df(t)}{dt} \geq 0 \qquad (6.4).$$

This is an accurate way to express the idea that, for "development" to be considered "sustainable", there should be no decrease in the capacity of our world to support welfare at any time scale (we can talk of "development" if, in addition, $\frac{d\overline{w}(t)}{dt} > 0$; otherwise, we can still take Eq. 6.4 as a condition of, say, "sustainable lifestyle", to put another widespread expression) (Flos, pers. comm., has suggested to rewrite the integrals in Eq. 6.3 in the frequency domain instead of the time domain, in order to treat each of the intervening processes in terms of its characteristic time scale).

We may still wonder why to talk of sustainability instead of just optimality. When Beckermann (1994) attacked the concept of sustainability in the name of optimality, he referred to the optimality of orthodox economics, which is measured in monetary terms, and hence we cannot agree with him if we do not agree with the interest of the monetary measure. However, what about another "optimality" that would consist of a maximization of future welfare? The condition of optimality would then be

$$f = \max \qquad (6.5).$$

If we accept these definitions, sustainability (Eq. 6.4) will ordinarily be a condition for optimality (Eq. 6.5), which allows to overcome the conflict that arises between sustainability and optimality when taking other definitions. I submit that both sustainability and optimality are useful criteria. The best option is always maximizing f, but not diminishing it may be considered the minimum acceptable. This is related (but not exactly equivalent) to a more general ethical tenet: the best is to do as much good as you can, but the minimum acceptable is not to do more harm than good.

There are good reasons to think that we are far from a path of sustainable development, so we could be happy if we attained at least this partial target.

The value of Eqs. 6.3-6.4 (and also of Eq. 6.5) is just conceptual. As stated in subsect. 6.4.1, when such levels of complexity and uncertainty are involved, it is pointless to try to quantify everything. These equations should just serve to express in precise terms an ultimate goal, but, once the goal has been understood, the actual decision process will necessarily look much more like the non-algorithmic multicriteria methods discussed by Martínez-Alier et al. (1998) than a mathematical calculation. Therefore, any practical recipe emanating from this criterion will have a multivariate form and will thus be classifiable under the label of “strong sustainability”.

6.5. Dealing with inner constraints

In sect. 6.4 I remark that the proper economic policies to deal with both environmental and social issues cannot be established without some previous ethical judgments, which necessarily have an extra-scientific nature. The developments in this chapter are essentially addressed to inform policies with an ethical background quite less discriminatory than the ethical background implicit in conventional economic theory. We should however wonder if it is really possible to have policies dominated by such principles, or, instead, economic and political inequality are inescapable. A necessary requirement for reducing inequality is a high level of complicity between large numbers of people, with the purpose to counteract the dominating interests and overcome huge historical inertias, with all the strength and all the carefulness that this requires. Here I will not discuss the possible obstacles and opportunities that “human nature” may pose to this endeavor. Besides wondering about human nature, we should make sure that there is no fundamental incompatibility between equality and the global stability of the economic system, which would eventually make useless any such effort. In this section I briefly discuss this last point, and some other aspects of the sustainability of the internal structure of economies. This has an uppermost importance not only to know how far can we go in the pursuit of social and environmental goals, but also to know in which way.

Several authors have put forward the idea that a capitalist system displays a SOC dynamics, instead of having a tendency to attain an “economic equilibrium”. One of these authors is the own father of SOC theory, Per Bak, who adds the strong conjectures that “the most robust state for an economy could be the decentralized self-organized critical state of capitalistic economics” and this “is the best state that is dynamically achievable” (Bak 1996, pg. 198). Bak uses SOC theory to support capitalism, despite the deep differences between “SOC-theory economics” and conventional economics. Societies anteceding capitalism may also have been SOC in different ways, e. g. in terms of armed conflicts. This suggests several questions. Is a SOC (or a critical) state the most robust or even the only attainable for any complex system, like a society? If it is not, is it still optimal in some way? Is SOC (or criticality) necessarily associated with inequality? About this last question, a striking fact is the scaling nature of inequality in capitalist societies, as apparent from the findings of Vilfredo Pareto in the 19th century (Pareto 1897). This economist found the incomes x of individuals in several towns to display a power law distribution

$$f(x) = ax^{-b} \qquad (6.6).$$

for some constants a, b. Analyses of more recent data also reveal a power law distribution for high incomes (e. g. Dragulescu and Yakovenko 2001) (income distributions have also been fitted with a lognormal (May 1975), which is not strange in the view of my own findings in subsect. 5.4.4 and App. 5D). In other words, income distribution seems to have the fingerprint of criticality or SOC.

Throughout this book I offer abundant evidence of the importance of the concepts of criticality and SOC when studying complex systems. The reasons are that these seem to be quite widespread phenomena and that they confer a strong predictability, which is welcome in this rather involved area. This does not mean that they are universal phenomena. For example, I also give some evidence that some kinds of forests are SOC and others are not, in terms of fire. Avnir et al. (1998) found that most of the presumed “fractals” cited in the physical literature do not cover much more than an order of magnitude (in contrast to several of the instances put forward in this book), so scaling is far from universal.

On the other hand, following the suggestion by Langton (1990), we may think that the systems able to process information and to adapt can only exist “at the edge of chaos”, close to criticality or in a state with similar characteristics, such as a SOC state. This confers the right flexibility, such that signals can either propagate very much or very little. If this holds, any sustainable economic system should remain "at the edge of chaos". However, a system may well be critical/SOC in one facet of its dynamics and not in another one (e. g. tropical rainforests may be non-SOC in terms of fire but critical in terms of diversity).

Is the possible critical or SOC dynamics behind Pareto's Law necessary for adaptability? The explanation of Pareto's Law given by several authors (e. g. Champernowne 1953, Biham et al. 2001) is similar to the explanation for scaling in species abundance distributions given in sect. 5.4 in this book. I show in sect. 5.3 and subsect. 4.3.2 that this implies criticality and may well contribute to adaptability. A similar reasoning can be extended to the case of incomes. Criticality will confer flexibility in the access of individuals to different levels of control on the system, which allows selective mechanisms to act. Individuals can increase (decrease) their influence by multiplying (shrinking) their income, and they have a higher chance if they adjust (do not adjust) in given ways to market signals, and to signals external to the market. This is a major mechanism for information processing, but it is not the only possible mechanism. Take e. g. the alternative economic system imagined by the ecological economist Hermann Daly (1991, 1980). This system would consist of a free market system with an upper and a lower limit to income, and an upper limit to wealth. These constraints would be narrow enough to impede a differentiation in social classes, but there would still be some margin of variation, for individuals to have an incentive to attend market signals. In addition, there would be a set of mechanisms to correct the bias between market signals and social optima in issues such as resource consumption and pollution. In this system, incomes would not have a critical dynamics anymore, but this does not necessarily mean that the system would lose its adaptability: adaptability would probably find other channels. Strategies and patterns of economic behavior would be copied from individual to individual. If we analyzed the abundances of individuals following different strategies, we would probably find a power law distribution (like the power law distribution of species abundances, chapt. 5), indicating a critical dynamics that would give room for selective pressures to act and would confer adaptability. Furthermore, every society has some decision mechanisms external to the market that, despite their limitations, do indeed involve a capacity for information processing.

The capitalist system displays recurrent crises that have been proposed to be the result of a SOC dynamics (Scheinkman and Woodford 1994, Bak 1996; Lux and Marchesi 1999). Also a system like Daly's could well be SOC in this facet: simple models of SOC dynamics in economics ignore differences in income, so SOC crises and scaling incomes are not necessarily linked. How should we deal with such crises? Neoliberals consider that a capitalist market free from external interventions is the best recipe against crises, but in fact crises will not cease if capitalist economies tend to achieve SOC instead of equilibrium. Keynesians advocate for an anticyclic state policy and this is widely applied in high income countries; however, if crises are a result of SOC dynamics, anticyclic policies will be analogous to fire suppression, with its long-term limitations. In fact, the limitations could be even stronger. In the case of fire dynamics, the driving force is rather simple (the accumulation of fuel) and it is relatively easy to abandon the SOC regime (if one accepts the risks involved in this step). A reasonable model for an economy would be more similar to the model for interactions between species in subsect. 4.4.4 and App. 4D. Here the driving force may involve a multiplicity of features of all kinds, which contribute to the vulnerability of each economic agent and each set of interlinked economic agents in case of crisis. The disappearance of a set of enterprises during a crisis would thus be a selective mechanism complementary to the gradual replacement of enterprises in the absence of crises. This would select robust enterprises and sets of interlinked enterprises and would have as a result the robustness of the system as a whole. If this is the case, a simple increase of public expenditure during crises (as advocated by Keynesians) would not suffice to abandon the SOC dynamics. Perhaps, instead of maintaining indefinitely such an anticyclic policy, we should allow crises to eliminate the enterprises that contribute to the system's fragility while having a strong safety net for the harmed people (as there would certainly be in a system with bounded incomes, like Daly's). Of course, we should also avoid procyclic policies (like those that the IMF has been charged to apply in front of some recent crises in "developing" countries (Stiglitz 2002)). It would perhaps also be recommendable to set strong limitations on the speculative sector, because it has no obvious usefulness for the society but seems to produce fluctuations with a higher recurrence than we should expect from the "real" economy by itself. Perhaps the only way to go beyond these partial solutions is actively favoring features of the enterprises and webs of interacting enterprises that increase their robustness, which would partially replace the function likely to be played by crises, and would, at least, decrease their recurrence.

In an economy oriented to the "common good", as the one conceived by Daly and any other that we could think of, decision mechanisms external to the market have a strong role to play. They should search and maintain an optimum between equality and incentive (which I personally perceive to be very close to equality as compared to the current situation in most countries, and indeed as compared to the international situation) and correct the strong biases between market signals and social optima. This bias correction function clearly includes setting strong limits to environmental damage (at the very least to attain a "sustainable development" as defined in last section) and ensuring the availability of basic social services (including a thorough service of assessment to ease the integration of each person in the economic system). From my point of view, it should also include the neutralization of discriminations of all kinds (in addition to the discrimination due to social class), strongly limiting advertising, and more subtle things such as setting some limit to the transfer of decision power to automatic devices (which may become inconvenient beyond some point, see subsect. 6.6.4). On the other hand, it should not obstruct the economic system's adaptability. E. g., in a highly egalitarian economy there would be less need of interventions addressed to maintain the volume of activity of some given sectors above the levels justifiable from the functions that these sectors have. Such interventions make currently a strong contribution to the environmental burden and, perhaps, to the system's fragility.

Such decision mechanisms external to the market should be highly robust, in correspondence with their importance. I do not think that they should take the form of monolithic institutions, since these would be fragile (see chapt. 4) and corruptible. Instead, there should be a multiplicity of decision circuits of different forms at all scales, with a strong democratic basis, a high dynamism, redundant functions, and capacity to control each other.

We have no solid basis to think that the pressures for social justice and environmental sustainability are useless, but I suspect that they should be carried out without pretending to set an endpoint to the never-end game that takes place "between order and chaos".

6.6. Some thoughts on globalization

6.6.1. General issues

This book has been written in a moment in which the world is undergoing a fast process of economic globalization. The term "globalization" can be either understood as an increase in the spatial scale of the economic activity in abstract, or as the increase in this scale that is currently taking place, in the particular form in which it takes place. Some of the main features of this particular form of globalization are: (i) it occurs in an economic context that is essentially capitalist, (ii) there are no democratic mechanisms operating at a spatial scale comparable to the scale attained by the economic activity, able to control it, (iii) it is largely driven by multilateral institutions in which each country's power depends mainly on its current income and on the political weight it attained after the World War II, (iv) the theoretical basis for the policies undertaken by such institutions is provided by the neoliberal ideology, which essentially advocates for the elimination of any regulation on markets except those addressed to guarantee and expand private property, (v) there are some significant deviations from a pure neoliberal policy, aimed to satisfy even more the aspirations of some economic pressure groups in high income countries.

Before dealing with the effects of globalization, I present a preliminary analysis of the international component of the current inequality in the appropriation of natural resources and ecosystem "services", using the ecological footprint (subsect. 6.2.2) as an indicator (data corresponding to the year 1999, taken from WWF 2002). Fig. 6.1A represents in a log-log scale the density of probability $f(\zeta)$ for a person chosen at random to live in a country with an average ecological footprint per capita ζ. This does not capture the internal inequality within each country: if this were also included, Fig. 1.6A would display even larger disparities in ζ. As apparent from Fig. 6.1A, the distribution of ζ is quite close to a power law, with $b \approx 1.5$. Although it does not cover much more than one order of magnitude, the comparison with the log-linear representation in Fig. 6.1B leaves little doubt that a power law is a much better approximation than an exponential (an exponential probability density function corresponds to a straight line in a log-linear representation, while a power law corresponds to a straight line in a log-log representation, see App. 2A). This empirical distribution is however the result of putting together a set of variables with quite different features. In view of the comments in subsect. 6.2.2, the consumption of exosomatic energy ζ_E is perhaps the least amenable to aggregation. It makes up however about half the world ecological footprint. ζ_E seems to be the main responsible for the power law behavior of ζ, although in isolation it displays a distribution that rather approaches a lognormal. As I show in chapt. 5, the lognormal can be considered to be the non-power-law distribution closest to a power law. Fig. 6.1C represents ζ_E vs. $\zeta_E\ f(\zeta_E)$, which would give rise to a Gaussian bell shape were the distribution exactly lognormal (App. 5H). The main deviation from a power law takes place for small ζ_E (smaller than $\zeta_E \approx 1$ ha) and is masked by the other fractions of ζ when calculating the whole footprint, as in Fig. 6.1A. The ensemble of the other

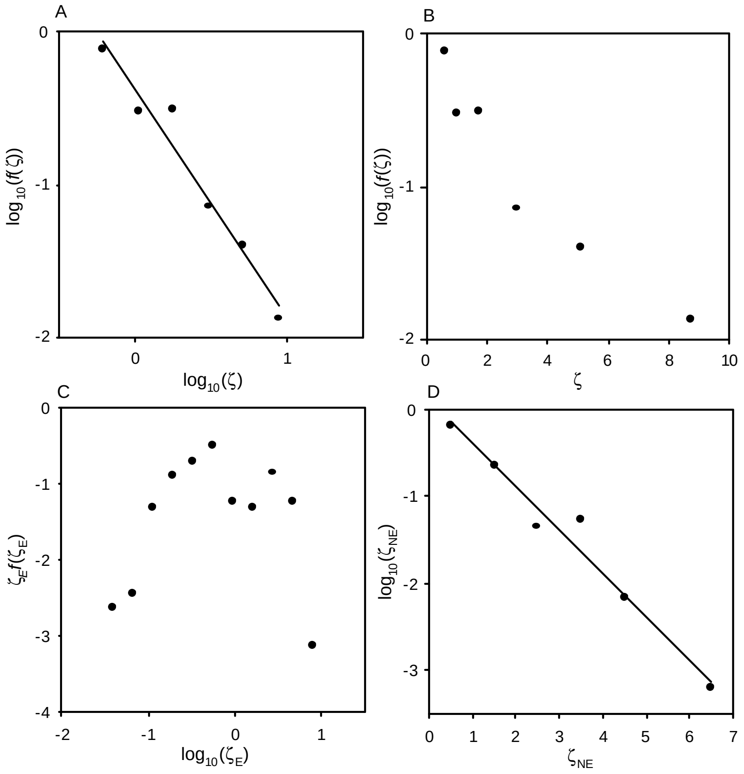


Fig. 6.1. *Snapshot of the international component of ecological inequality, using the ecological footprint as an indicator. Data from 1999, in ha.* **(A)** *Log-log plot of the density of probability for a person chosen at random to live in a country with an average footprint* ζ. **(B)** *The same, but in a log-linear plot.* **(C)** *Preston-like plot of the fraction of* ζ *that refers to use of exosomatic energy,* ζ_E. **(D)** *Log-linear plot of the other parts of* ζ *taken together,* ζ_{NE}.

fractions (ζ_{NE}) has a distribution close to exponential, as apparent from Fig. 6.1D. This suggests that the consumption of exosomatic energy is potentially much more variable than the bulk of the other fractions of the ecological footprint, and reflects more clearly the multiplicative processes that generate inequality. Nowadays there is not however a very large difference between the inequality in terms of ζ_E and in terms of ζ_{NE} (ζ_E has a coefficient of variation $CV(z_E) \approx 350$, while $CV(z_{NE}) \approx 290$). It is worth noting that the largest amount of people is concentrated in the lower end of the distribution of ζ_{NE}. The inequality in the control over resources can indeed be expected to be even larger than the inequality in their consumption, and to ordinarily approach a power law distribution (although it is an intra-country instance, think e. g. of the proverbial disparities in land tenure in several Latin American countries).

Several of the concepts and tools that I work out in this book can be used to deal with the consequences of increasing spatial scale in complex systems such as economies. In the rest of this section I outline a conceptual framework for the problems of globalization, which could perhaps ease the application of these and other concepts and tools in a future analysis of globalization from an ecological economic and an econophysical point of view. I group my heuristic proposals in four blocks, which correspond to four facets of the increase in scale: homogenization, escape from local regulation loops, functional differentiation and transport processes. This could at least be useful as a way to organize ideas that we already had about this phenomenon. This is one of my main purposes in the next subsections, which, nevertheless, also include new ideas. See also the related discussions about international trade and the environment in Folke et al. (1994) and Muradian and Martínez Alier (2001).

6.6.2. Homogenization

One consequence of globalization in its current form is a process of mixture that erases much of the informational heterogeneity of the world. This includes biological, cultural and scientific-technological information.

The diffusion of technology is one of the few facets of globalization that could increase economic equality, but this is largely counteracted by global agreements on trade-related intellectual property rights (TRIPs).

This is not a place to treat in depth the issue of cultural exchanges, but it is worth noting that the central reasoning of chapt. 2 can be applied to cultural as well as ecological issues: when a culture is forced to change fast, the result can only be destructive (which is a consequence of the physics of irreversible processes). There are countless examples of the traumatic effects for many peoples of the contact with the currently dominating Western culture, which has lead to social disruption and marginalization, and loss of valuable knowledge. This reasoning does not extend to gradual and non-forced cultural exchanges.

The intensity of transport, combined with a lack of safety rules to avoid biological invasions, produces a worldwide biological mixture, which is extremely fast if we take the evolutionary time scale as a reference. It is thus not strange that the results are destructive. There is not just a loss of spatial heterogeneity, but also a global loss of biodiversity (while local biodiversity does not always decrease, which can be seen as an extension of the Hutchinson's paradox of plankton in turbulent waters, at least in the version by Margalef 1989, pg. 377-379) and often a disruption of ecosystem function (see e. g. Chapin III (2000) or Mooney (1998) for instances, and my sect. 4.4.4 for a hypothetical mechanism). In particular, there has been an increase in the pathways for the diffusion of pathogens, and the contacts between segments of the Biosphere that were largely isolated in the past has unleashed many "emergent diseases" (including AIDS) (e. g. Pueyo 1998).

Agricultural biodiversity is becoming specially homogeneous and poor as a consequence of globalization. This is a threat to food security, because lack of biodiversity can be expected to increase variability in aggregate production and decrease adaptability (chapt. 4). In particular, there is an increasing danger of some pest or epidemic propagating from country to country through endless fields sowed with plants with the same genotype, and a decreasing availability of genetic resources to respond in front of such an event (see Wilkes 1992). In sect. 4.5 I suggest a possible way to reorganize the agricultural diversity patterns at all scales in order to minimize the incidence of pests and pathogens through the use of biodiversity (it is still highly hypothetical). This or any comparable recipe would require large scale coordination and, hence, some form of "globalization", but this coordination should be external to the market and would be difficult to make compatible with the current form of globalization.

6.6.3. Escape from local regulation loops

The increase in the spatial scale of economies represents a short-circuit on local regulation loops, such as those set by the local availability of resources or the local social rules. This short-circuit is not just a consequence of the globalization of markets, but also of the increased use of resources and emission of pollutants with an inherently nonlocal nature: the main instance is the pass from the use of local primary production as the essential source of exergy to fossil fuels and, hence, global carbon sinks. Two of the keys for the privileged economic position of Europe are its past conquer or exploitation of other continents and its early use of fossil fuels (Lang 2000), which were two forms of eluding local limitations by passing to rely on spatial scales beyond its own political borders.

One consequence of the loss of the spatial barrier is that it becomes possible to combine different resources that were segregated in space. This is mentioned by Margalef, who points out that the same effect explains the increase in plankton productivity induced by (a moderate) turbulence. This produces an increase in economic potential, which is good news or bad news depending on the way it is used. On the other hand, there are reasons to expect some neatly negative effects in other respects.

In sect. 6.5 I mention a model for the power law distribution of incomes in capitalist economies, similar to the simple model for the power law distribution of species abundances in sect. 5.4 and subsect. 4.3.2. In the case of species abundances, the power law emerges at the limit of a high niche overlapping, and the bounds to the power law become wider the higher is the degree of overlapping. The gradual weakening of the limits on each individual’s resource use set by distance is analogous to a decrease in niche segregation. Therefore, it must have been a central factor in the emergence of the power law distribution of incomes. In a capitalist context, we can thus expect a higher inequality in resource appropriation the higher is the degree of overlapping in potential resource availability as spatial scales increase (of course this effect is mixed with many other effects). This must rule not just for individuals but also for enterprises: their sizes have also been shown to display a power law distribution (Ramsden and Kiss-Haypál 2000), and the huge transnational corporations are indeed favored by globalization.

Another consequence of overlapping in potential resource availability in the current economic context is a lack of incentive for conservation, by virtue of what has been called “free rider effect”. E. g. local traditional rules to ensure the conservation of common goods such as forests or aquifers lose support when people can obtain the same resources from elsewhere and local resources can be appropriated by outsiders. Furthermore some of the detrimental effects of human abuses on the environment could perhaps become apparent more abruptly if the spatial barriers have vanished, which would make even more difficult the generation of societal regulations that anticipate environmental feed-backs.

Even when there are the conditions for an unregulated global market to be compatible with an appropriate resource conservation (resource turnover exceeding interest rates, well-defined access rights, market demand matching the social needs), there may be problems left (in addition to the increase in inequality) due to the fact that regulation will be just global, not local. There will not be such problems if it is planed to maintain indefinitely the same spatial scale of the economy but, if it is found to have become excessive, its posterior decrease will be more difficult because some given sites will have become too deficient in some given resources. In particular, there will have been an excessive growth of urban areas, which are typically deficient in all kinds of natural resources and ecosystem “services”. The sizes of cities displays a power law distribution, with a remarkable constancy in its slope parameter b (Eq. 6.6). This distribution has been explained by mechanisms analogous to those suggested for the power law distribution of incomes (see e. g. Manrubia and Zanette 1998). The upper bound to this distribution can be expected to increase with globalization, as happens with the power law in the use of resources per capita.

One of the central problems of globalization is that, as soon as a single market is shared by several governments, their whole economic policies become subject to a kind of “free rider” effect. This imposes a strong limitation to any democratic regulation of economy, since when there is some democratic mechanism at work this rarely exceeds the scale of a country. Therefore, globalization magnifies all of the problems of capitalism, including the lack of mechanisms to counteract the other challenges posed by the increase in spatial scale. It is increasingly difficult to avoid capital outflows as an answer to policies disliked by the investors, and the General Agreement on Trade and Tariffs (GATT) forbids any rule requiring imported products to have been elaborated under social and environmental standards homologable to those in the importing country (the only exception is that governments are allowed to forbid the importation of products made by prisoners). There are some official proposals for the inclusion of social and environmental clauses in the international trade rules, but low-income countries fear that these will be used as an excuse for a biased protectionism against them. The perception of this issue is often confuse, because there are two mixed effects at work: the constraints on social and environmental standards due to each country income (which will have a clear effect on e. g. the attainable salaries and the public expenditure in absolute terms, but not necessarily on the relative weight given to each social class, to each gender or to current vs. future generations), and the genuine "free rider" effect, which takes place even if the involved countries have the same income per capita. From my point of view, there should be some agreed objective criteria to reasonably separate both effects, and states should be allowed to adopt measures in front of the second one, but not of the first.

The problems in this subsection and most of those in the other sections are not inherent to the increase in spatial scale. They could be avoided if local regulations were replaced by proper global regulations sustained by the pressure of society.

6.6.4. Functional differentiation

The main argument that is alleged in favor of globalization is that this will allow for a higher degree of specialization and hence increase "production". In a different economic context, we could replace production by physical efficiency in resource use. Certainly, we can expect specialization to increase the economic potential, which, as I put in the previous subsection, is either good or bad depending on the way this potential is used. However, specialization cannot be dissociated from issues of equity and stability.

More specifically, the pro-neoliberal-globalization ideology has a strong influence of Ricardo's theorem of comparative advantage. Roughly said, this states that, given a set of agents with some given "coefficients of production" for each economic sector, if we allow for free trade between these agents, each will specialize into the sectors in which it has a maximum "comparative advantage" and, to a higher or a lesser extent, all agents will gain from the exchanges they decide to engage on. We can obtain nearly any result from a model if we can chose at will what we take as a variable and what we take as a constant. The theorem of Ricardo takes the coefficients of production as constants and the volume of production as variable, while in fact both change. Several critics have highlighted that comparative advantages are in fact dynamic (see e. g. Ekins et al. 1994, Muradian and Martínez Alier 2001, and references therein). The theorem of Ricardo is one more instance of a linear approximation that (from Taylor-series reasoning, see sect. 6.4) might well hold when the process of economic integration is slow, the time frame for which we care is brief and the initial differences between countries are small, but not if any of these conditions fails (in particular, see subsect. 6.4.2 about reasons not to care just for the short term, from an ecological economic point of view). Some of the phenomena that may take place when leaving the linear domain can be guessed from the simple fact that both production coefficients and production volumes are complex non-equilibrium features and, as such, require an active maintenance and can be lost faster that can be gained.

A rapid economic integration transitorily acts as a disturbance that collapses some sectors faster than the people involved in them find other occupations. Stiglitz (2002) criticizes the IMF for its policy of forcing low income countries to open their frontiers to the global market as fast as possible and without paying attention to the order in their sequence of reforms, which seems to have produced disastrous results in many cases.

In the long term, the efforts necessary to develop and maintain production coefficients (either by the own agents or in the form of investments by other agents) will be guided by the same comparative advantage criteria that guides trade, so agents will irreversibly decrease their production coefficients for the sectors in which they do not participate anymore and increase them in the sectors in which they specialize. This modifies the initial "parameters", so the overall evolution will not necessarily be of benefit for everybody. Specialization will eventually do more harm than good to those agents specializing in functions whose demands decrease across time, as these functions are either replaced or used with more efficiency (in this last case demand may either increase or decrease, depending on its elasticity to price). This is essentially the mechanism behind Krugman (1981) model, in which countries specializing in primary products that serve as inputs to industries in other countries become worse off beyond some point in time (see Muradian and Martínez Alier 2001 to have a glimpse of the real importance of this process in the current world and its environmental implications). This may also be a general problem for the large fraction of people specialized in selling their work force for economic activities that they do not control (this also has an international dimension, as capital and work force often belong to different countries). Across industrialization, the functions that do not involve much complexity have been largely transferred from workers to exosomatic structures, with the double effect of higher inequality on the one hand and either a decreased physical efficiency in the use of natural resources or a genuine increase in the economic potential (which may be beneficial, but not necessarily) on the other. If these problems were addressed, such a transference would indeed be desirable, but not otherwise. (With the development of information technology, more complex functions, including important decision functions, are also transferred to exosomatic structures, as soon as these become more cost-effective than humans in market terms. This transfer is unavoidable in a system driven by competition. Up to some point, the consequences are comparable to transfer of other functions. However, carried to the extreme this process could well generate qualitatively new problems. See the warnings by the own father of cybernetics, N. Wiener 1948, and by the pioneers of "artificial life" T. S. Ray 1995 and C. G. Langton 1995).

The above "specialization trap" is indeed not irrevocable for each particular individual or country. Each economic agent can depart from its expected path either by unpredictable fluctuations or by paying an extra cost to engage in sectors in which it does not have an initial comparative advantage. When this is carried out by a country as a whole, the extra cost it is often perceived as dumping, be it economic, social or environmental.

Both for inequality between individuals and between countries, there is something even more fundamental than the partition of investments between the coefficients of production for each activity: the investment capacity of the own individual or country. This is a direct function of income, which introduces an obvious multiplicative process (also debt is multiplicative). It is also subject to unpredictable departures. To the extent that economic marginality stimulates demographic growth, this will further increase inequality. Furthermore, the sequence of liberalizing reforms in the global economy is heavily biased in favor of some sectors in high income countries, as mentioned in subsect. 6.6.1. The joint result of so many processes is a set of distributions of income and natural resource use that seem to approach power laws, which is not strange if we take into account the importance of multiplicative processes in economics and the robustness of this kind of distribution, so apparent from chapt. 5.

As a conclusion, “free markets” do not benefit everybody. They generate strong inequalities, and there may well be net losers. These problems are likely to be enhanced as markets expand, and may have a geographical projection. Assuming that international trade reinforces inequality, the rules of international trade would only be fair if they contained provisions for a systematic “positive discrimination” in favor of low income countries, strong enough to neutralize the mechanisms working against them (to be even fairer, the provisions should go beyond and actively foster equality, overall if we take into account the “ecological debt” and the many “extra-official” debts that high income countries have with low income countries, see Martínez Alier 1997). In particular, low income countries should have a wide margin for a selective protectionism in front of high income countries.

Another important issue is the stability of the web of interactions that emerges across globalization. Specialization is likely to contribute to the vulnerability mentioned in sect. 6.5, at the same time that it increases the effectiveness of economic processes. The degree of specialization attained in an economy could well result automatically from the contraposition between these two effects. This contraposition may or may not take the form of a SOC dynamics (see sect. 6.5 about SOC in economics, and subsect. 4.4.4 about specialization and SOC). If the consequence is some constancy in the number of “effective” agents with which each agent interacts, the merging of markets will produce a decrease in the proportion that this number represents as compared to the total number of participants in the market, which represents a larger specialization. This will increase the economic potential but also the degree of dependence: a posterior fragmentation of the market will generate problems not only because of the need of particular economic activities (or even natural resources, see subsect. 6.6.3) that may have been displaced abroad, but also because the economic agents will be overspecialized in relation to the size of the market (if economies are SOC, the effect on economic crises is likely to be the same as the effect of increasing r_e in subsects. 3.4.2-3.4.5). Therefore, it is imprudent to push economic integration before we have a clear idea on the optimum scale of economies (which does not necessarily have to be the maximum, if we take into account the points raised throughout this section).

In fact, the most immediate reason for concern about stability under globalization is that it is accompanied by an immense growth in the part of financial exchanges that corresponds to speculative operations. These seem to fluctuate strongly at a time scale that is short as compared to the one of the "productive" part of economy, and disturb it. Another reason is that there is a fast generation of long-range connections of all kinds and we cannot rule out that these make the system become "supercritical" (or equivalent, see sect. 2.5) in some feature, which perhaps we do not even imagine.

These concerns make irresponsible a fast process of globalization not preceded by a serious analysis of the consequences.

6.6.5. Transport processes

All of the problems put forward in the last two sections are a consequence of transport, but transport on itself also has more direct environmental effects, which are worth mentioning.

Transport requires exergy expenditure, with the consequent problems of resource needs and pollution. Transport is responsible for about one third of the world consumption of fossil fuels (Pimentel et al. 1994). The longer the transport, the higher the exergy and effort required to close material cycles. This fact, together with the loss of local regulation circuits, has already opened material cycles even at an intercontinental scale (Gardner 1998). This does not mean that transport should be minimized in a sustainable economy. A high spatial scale allows for a high specialization. In a sustainability-oriented economy, some degree of specialization would allow to increase the physical efficiency in the use of resources, which, up to some point, would compensate for the losses due to transport (e. g. larger organisms are not less efficient in the use of resources). Therefore, there must be an optimum intermediate scale. To some extent, this will be different for each economic sector. For many sectors, an increase in spatial scale is indeed useless.

Transport requires infrastructures that occupy some surface. More importantly, the infrastructures for terrestrial transport, specially roads, disrupt the topology of ecosystems. As put forward in subsect. 6.3.2, space occupation, even if random, at some point leads to fragmentation. This is much faster in the case of transport pathways, because these are elongated and precisely designed to keep anthropogenic infrastructures topologically united. Unless special measures are taken, this leads to the opposite outcome for ecosystems, which represents a major disturbance. Between other effects, fragmentation is thought to be one of the main causes of biodiversity loss.

Other problems of transport are pollution due to accidental releases of the transported commodities, and the unleashing of biological invasions (already mentioned in subsect. 6.6.2.

6.7. Conclusions

The main ideas that have been put forward in this chapter are:

- About irreversibility and the economic process:
    - Georgescu-Roegen (1971) claim that “low entropy” is a condition for economic value is backed by the results from chapt. 1. Unlike thought by this author, however, this does not imply that the life span of humanity is limited by the reserves of “low entropy” in the Earth crust.
    - Not only losses of exergy (“exergy” is a more exact word for what is usually called “energy” in an economic context) are irreversible, but also losses of complexity, which is not the same (see chapt. 2). Current environmental problems are mainly a result of the second kind of losses.
    - Different kinds of elements (in a wide, non-chemical sense of the word) are often not exchangeable. Due to irreversible complexity losses, exchangeability is often unidirectional. These limitations make inappropriate the use of any aggregate biophysical variable as a measure of value, although some of these may be useful as partial indicators.
- About system dynamics and organization in ecological economics

  - We can extract from the "ecological arrow of time" captured by the modified Margalef principle in chapt. 2 that, if we force the Biosphere to change too fast, the result can only be destructive. This is a reason to take great care in our interactions with the environment, regardless of the fact that the Biosphere has never ceased to change and we cannot talk of any "balance of nature".
  - Besides the immediate impact of anthropogenic disturbances, we should also take into account how the traces left by current disturbances will modify the future regimes of catastrophes. The results from chapt. 3 and 4 set a basis for this endeavor.
  - Some of the methods put forward in chapt. 3 can be used to quantify environmental thresholds in which phase transitions take place, which is a requisite to make operational some recipes of ecological economics.
- About ultimate goals and criteria:
  - The use of the monetary measure of value as a decision tool for public institutions has been justified on two bases, at least implicitly: on an assumption of exchangeability, and on the assumption that this measure captures the degree of "desirability". The first assumption is incorrect, so it can justify no aggregate measure. The degree of desirability cannot be established without some previous political choices. The monetary unit implies a discriminatory choice.
  - It is as a principle possible to build a non-discriminatory measure of desirability. However, this should not even share the linear and additive schema of the monetary measure. This corresponds to taking only the first-order terms in a Taylor series, which does not always suffice. Given the huge complexities involved, it could well be more useful to take advantage of non-algorithmic multicriteria decision methods, as advocated by several authors.
  - If we think in terms of welfare instead of monetary measures, there is no basis to apply a discount of future to the environment. Since there could well be a decrease in the levels of welfare in a near future largely due to ecological problems, there is even a basis for a negative discount within some time window. A positive discount at very long

terms could perhaps be justified because of the increasing uncertainty, but it should be small and non-exponential. Once we take large time scales into account, we should also consider many nonlinearities that are ordinarily neglected.

  - Sustainable development should be faced by taking into account that the persistence of complex systems requires a continuous modification in their details, and that there is a superposition of processes at very different time scales. I propose a formal criterion for development to be sustainable (Eqs. 6.3-6.4), which essentially consists on the maintenance of the capacity of our world to support welfare at all time scales. If we accept this criterion, sustainability will ordinarily be a condition for optimality (in nonmonetary terms), which allows to overcome the apparent contradiction between sustainability and optimality.
- About inner constraints:
  - It may well happen that, for the economic system to be robust and adaptable, it has to remain in a critical, self-organized critical (SOC) or analogous situation. However, it may be critical/SOC in some features and not in others. Therefore, the need of robustness and adaptability is not necessarily an obstacle to increasing equality, in spite of the fact that this would run counter the critical dynamics that seems to lie behind the distribution of incomes in capitalist economies. If economies are SOC, economic crises should be managed in different ways than those generally assumed. Furthermore, the social mechanisms charged for ensuring equality, environmental sustainability and other goals may well have to incorporate much redundancy and dynamism for them not to become unstable and corruptible.
- About globalization:
  - Inequality in the appropriation of natural resources and ecosystem services has a clear international component. The statistical distribution of the expected “ecological footprint” as a function of one’s country is

close to a power law, which reflects the multiplicative processes that generate inequality.

- The increase in the spatial scale of the economic system in a capitalist context poses a series of problems, which I have conceptualized in the terms of ecological economics and econophysics and grouped in four blocks: those related to homogenization, escape from local regulation loops, functional differentiation and transport processes. An increase in spatial scale is not necessarily negative. It may be justified in some economic sectors and not in others. If carried out, it should however be subject to strong societal regulations at a global scale, which are currently missing.

Conclusions generals

L'objectiu d'aquest treball ha estat contribuir a la recerca de regularitats generals (o quasi) al nivell d'integració d'ecosistema, amb el propòsit que el seu coneixement serveixi per orientar-nos en les nostres relacions amb l'entorn, en aquest moment de grans transformacions socioeconòmiques.

El punt de partida ha estat l'anàlisi de les propostes d'un seguit d'autors de renom que han postulat que els ecosistemes s'autoorganitzen de manera que es maximitzin o minimitzin uns certs potencials termodinàmics, cosa que explicaria les suposades tendències de la successió. La conclusió que n'he tret és que aquestes propostes no tenen cap fonament sòlid, però que sí que és cert que la Segona Llei de la termodinàmica és el motiu últim de les irreversibilitats que s'observen en els ecosistemes en termes d'informació i organització, i no tan sols d'intercanvis energètics elementals. En concret:

- La Segona Llei explica la "sageta del temps" de la memòria, en funció de la qual tan sols poden quedar registrats esdeveniments del passat i el processament d'informació té lloc de passat a futur.
- Al seu torn, la sageta del temps de la memòria justifica l'asimetria en el temps entre augments i minves de complexitat, d'acord amb la qual les minves poden tenir lloc arbitràriament depressa (catàstrofes, pertorbacions) però els augments no.
- Al seu torn, aquesta asimetria temporal fa previsible l'existència de "tendències" dissociades de qualsevol mena de potencial, deixant sense raó de ser la recerca de potencials.

Aquestes troballes permeten comprendre l'origen d'algunes regularitats molt bàsiques de la natura, però tan bàsiques que és dubtós que duguin a gaires prediccions noves, excepte la predicció de la insolvència de les prediccions que deriven dels models que pretenien explicar aquestes mateixes regularitats d'altres maneres.

La via d'avanç que em sembla més apropiada per augmentar la capacitat predictiva de l'ecologia sistèmica està en centrar l'atenció en prediccions temptatives que depenguin molt poc del detall dels models en els quals es basen, en virtut dels efectes de "grans nombres" que tan útils han estat per a la física estadística, i en posar-les a prova amb dades empíriques, fent servir eines estadístiques expressament dissenyades amb aquest propòsit. Les fites més remarcables que he assolit treballant amb aquest plantejament han estat:

- Desenvolupar un model dels règims de catàstrofes (per a tipus de catàstrofes que satisfacin certes propietats, força generals) i de com els afecten els factors externs (per exemple, el canvi climàtic). N'he constatat empíricament una forta capacitat predictiva en el cas dels incendis forestals.
- Clarificar d'un seguit d'aspectes que romanien obscurs pel que fa a la relació entre diversitat i estabilitat. El cos teòric que en resulta reforça l'expectativa que la diversitat faci decréixer la variabilitat en paràmetres globals dels ecosistemes, tal com han postulat molts autors/es i s'ha observat empíricament en alguns casos.
- Fer una teoria dels patrons de diversitat que unifica i justifica les principals distribucions estadístiques d'abundàncies d'espècies a les quals s'ha trobat utilitat empírica (logsèries, potencial, lognormal, canònica) i la relació espècies-àrea. He reforçat els resultats teòrics amb l'anàlisi d'uns inventaris molt exhaustius de fitoplàncton marí, extrets de la bibliografia.

En el decurs d'aquesta recerca s'ha mostrat especialment útil el concepte de criticalitat ("criticalitat" és la qualitat d'estar en un "punt crític" en el qual té lloc una transició de fase de segon ordre). La criticalitat ha resultat tenir interès ecològic per tres motius:

- Perquè sembla haver-hi ecosistemes que, en ser manipulats des de l'exterior (per exemple, per una intervenció humana) sofreixen una transició de fase, passant per un estat crític.
- Perquè sembla haver-hi ecosistemes amb la dinàmica anomenada "criticalitat autoorganitzada" (SOC), que té una estreta relació amb la criticalitat (tot i que mostro que no condueix els sistemes a un veritable estat crític, contràriament a allò que se sol assumir).
- Perquè sembla haver-hi ecosistemes que tendeixen per si mateixos a un estat crític d'una manera que no té relació amb la SOC, sinó amb propietats elementals de les fluctuacions en les abundàncies de les espècies.

Tot i que l'objectiu bàsic del treball era contribuir a donar uns fonaments teòrics ferms a l'ecologia, a partir dels quals poguessin sorgir aplicacions, també he fet unes quantes troballes amb una utilitat pràctica immediata. Les més remarcables són:

- He desenvolupat un sistema de pronòstic d'incendis a partir de prediccions meteorològiques rutinàries, del qual ja n'he comprovat l'eficàcia, així com un seguit de criteris de gestió del foc.
- He reunit evidència indirecta que els impactes antropogènics poden estar forçant transicions de fase de tipus catastròfic en regions de selva tropical.
- He desenvolupat eines per fer més rigorosa la quantificació de la diversitat i aplicar-la a la planificació de la conservació.

D'aquestes troballes, és la segona la que trobo que mereix una atenció més forta en aquests moments.

M'he avançat a explorar les conseqüències que tindria si trobéssim que els resultats que ja tenen un bon suport empíric en el cas dels incendis forestals també fossin extensibles a dos altres tipus de catàstrofes d'especial importància socioecològica: les plagues i epidèmies agrícoles i els efectes que resulten de les introduccions o extincions d'espècies. En concret, dono idees sobre com assolir un model agrícola productiu i sostenible, basat en l'ús intensiu de la biodiversitat vegetal, si es complís aquesta hipòtesi, que cal posar a prova quan més aviat millor.

Cloc el treball amb unes reflexions sobre com aquesta mena d'avanços en teoria ecològica s'han de reflectir en l'economia ecològica, que és l'intent de reformular la teoria econòmica parant una atenció especial a les realitats ecològiques, ja que estic convençut que l'èxit d'aquest projecte és vital. L'economia ecològica ha de defugir sobresimplificacions com ara la reducció de tot a termodinàmica o la referència a visions massa estàtiques de la natura o del propi sistema econòmic. Entre d'altres propostes, formulo un nou criteri de "sostenibilitat". També incloc alguns comentaris sobre la globalització.

References

Aber, J., N. Christensen, I. Fernandez, J. Franklin, L. Hidinger, M. Hunter, J. MacMahon, D. Mladenoff, J. Pastor, D. Perry, R. slangen, and H. van Miegroet. 2000. Applying ecological principles to management of the U. S. national forests. Issues in Ecology 6.

Ackerman, F., D. Kiron, N. R. Goodwin, J. M. Harris, and K. Gallagher (eds.). 1997. Human well-being and economic goals. Island Press, Washington D. C.

Agee, J. K., and M. H. Huff. 1987. Fuel succession in a western helmlock/Douglas-fir forest. Canadian Journal of Forest Research 17: 697-704.

Allen, P. M. 1976. Evolution, population dynamics and stability. Proceedings of the National Academy of Sciences of the USA 73: 665-668.

Altieri, M. A., and L. C. Merrick. 1987. In-situ conservation of crop genetic resources through maintenance of traditional farming systems. Economic Botany 41: 86-96.

Amilien, C. 1996. Conflicting international policies in tropical timber trade. Environmental Conservation 23: 29-33.

Anderson, J. E., W. H. Romme, G. Meyer, D. H. Knight, and L. Wallace. 1999. Yellowstone fires. Science 283: 175-176.

Andow, D. A. 1991. Vegetational diversity and arthropod population response. Annual Review of Entomology 36: 561-586.

Arrow, K., B. Bolin, R. Costanza, P. Dasgupta, C. Folke, C. S. Holling, B.-O. Jansson, S. Levin, K.-G. Mäler, C. Perrings, and D. Pimentel. 1995. Economic growth, carrying capacity, and the environment. Science 268: 520-521.

Arrow, K. J., W. R. Cline, K-G. Maler, M. Munasinghe, R. Squitieri, and J. E. Stiglitz. 1996. Intertemporal equity, discounting, and the environment. Pp. 124-144 in: Intergovernmental Panel on Climate Change. 1996. Climate Change 1995: Economic and Social Dimensions of Climate Change. Cambridge University Press.

Asselmeyer, T., and W. Ebeling. 1997. Unified description of evolutionary strategies over continuous parameter spaces. BioSystems 41: 167-178.

Avnir, D., O. Biham, D. Lidar, and O. Malcai. 1998. Is the geometry of nature fractal? Science 279: 39-40.

Ayres, R. U. 2000. Commentary on the utility of the ecological footprint concept. Ecological Economics 32: 347-349.

Bächler, G. 1994. Desertización y conflicto. La marginalización de la pobreza y de los conflictos ambientales. Ecología Política 8: 55-84.

Bak, P. 1996. How nature works. The science of self-organized criticality. Copernicus, New York.

Bak, P., C. Tang, and K. Wiesenfeld. 1987. Self-organized criticality: an explanation for 1/f noise. Physical Review Letters 59: 381-384.

Bak, P., C. Tang, and K. Wiesenfeld. 1988. Self-organized criticality. Physical Review A38: 364-374.

Bak, P., K. Chen, and C. Tang. 1990. A forest-fire model and some thoughts on turbulence. Physics Letters A147: 297-300.

Balling, R. C., Jr., G. A. Meyer, and S. G. Wells. 1992. Climate change in Yellowstone National Park: is the drought-related risk of wildfires increasing? Climatic Change 22: 35-45.

Balmford, A., A. Bruner, P. Cooper, R. Costanza, S. Farber, R. E. Green, M. Jenkins, P. Jefferiss, V. Jessamy, J. Madden, K. Munro, N. Myers, S. Naeem, J. Paavola, M Rayment, S. Rosendo, J. Roughgarden, K. Trumper, and B. K. Turner. 2002. Economic reasons for conserving wild nature. Science 297: 950-953.

Barker, R. 1996. Wildfires. Yellowstone fires and their legacy. The Post Company of Idaho Falls, Idaho Falls.

Bascompte, J., and R. V. Solé. 1996. Habitat fragmentation and extinction thresholds in spatially explicit models. Journal of Animal Ecology 65: 465-473.

Battley, E. H. 1987. Energetics of Microbial Growth. Wiley Interscience, New York.

Battley, E. H. 1993. Calculation of entropy change accompanying growth of Escherichia coli K-12 on succinic acid. Biotechnology Bioengineering 41: 422-428.

Battley, E. H. 1999. On entropy and absorbed thermal energy in biomass; a biologist's perspective. Thermochimica Acta 331: 1-12.

Baumgärtner, S. 1996. The use of the entropy concept in ecological economics. Pp. 115-135 in: M. Faber, R. Manstetten and J. Proops, eds., Edwar Elgar, Cheltenham.

Becker, B. 1997. Sustainability assessment: a review of values, concepts and methodological approaches. Issues in Agriculture 10. Consultative Group on International Agricultural Research, Washington.

Beckerman, W. 1994. Sustainable development: is it a useful concept? Environmental Values 3: 191-209.

Beckerman, W. 1995. How would you like your 'sustainability', sir? Weak or strong? A reply to my critics. Environmental Values 4 : 169-179.

Bengtsson, T. 1998. The great mortality decline : its causes and consequences. Pp. 84-99 in: C.-E. Núñez (ed.), Debates and Controversies in Economic History, Centro de Estudios Ramón Areces, Madrid.

Bennet, C. H. 1982. The thermodynamics of computation: a review. International Journal of Theoretical Physics 21: 905-940.

Bennet, C. H. 1987. Demons, engines and the second law. Scientific American 257: 88-96.

Bennet, C. H. 1988. Logical depth and physical complexity. Pp. 227-257 in: R. Herken (ed.), The Universal Turing machine. A half-century survey. Oxford University Press.

Bennet, H. 1988. Logical depth and physical complexity. Pp. 227-257 in: R. Herken (ed.), The Universal Turing machine. A half-century survey. Oxford University Press.

Bessie, W. C., and E. A. Johnson. 1995. The relative importance of fuels and weather on fire behavior in subalpine forests. Ecology 76: 747-762.

Biham, O., Z.-F. Huang, O. Malcai, and S. Solomon. 2001. Long-time fluctuations in a dynamical model of stock market indices. Physical Review E 64: 026101.

Binswanger, M. 1993. From microscopic to macroscopic theories: entropic aspects of ecological and economic processes. Ecological Economics 8: 209-234.

den Boer, P. J. 1968. Spreading of risk and stabilization of animal numbers. Acta Biotheoretica 18: 165-194.

Bond, W. J. 1993. Keystone species. Pp. 237-253 in Schulze, E.-D., and H. A. Hooney, eds. Biodiversity and Ecosystem Function. Springer-Verlag, Berlin.

Bosch, O. 1995. Bases ecològiques per a planificar la reducció del risc d'incendi. Pp. 106-109 in: Ponències i conclusions del fòrum nacional sobre els incendis forestals. Muntanya 799: 103-118.

Botsford, L. W., J. C. Castilla, and C. H. Peterson. 1997. The management of fisheries and marine ecosystems. Science 277: 509-515.

Bouchaud, J.-P., M. Marsili, and B. M. Roechner (eds.). 2001. Proceedings: Application of Physics in Economic Modelling, Prague, Czech Republic, 8-10 February 2001. Physica A 299, issues 1-2.

Brent, R. J. 1998. Cost-Benefit Analysis for Developping Countries. Edward Elgar, Cheltenham.

Brillouin, L. 1962. Science and information theory, 2nd ed. Academic Press, New York.

Brokaw, N. 1998. Fragments past, present and future. Trends in Ecology and Evolution 13: 382-383.

Brooks, D. R., and E. O. Wiley. 1986. Evolution as Entropy: Toward a Unified Theory of Biology. University of Chicago Press.

Brown, J. H. 1995. Macroecology. University of Chicago Press.

Brown, J. K. M., and M. S. Hovm∅ller. 2002. Aerial dispersal of pathogens on the global and continental scales and its impact on plant disease. Science 297: 537-541.

Brown, L. R. 1997. The Agricultural Link: How Environmental Deterioration Could Disrupt Economic Progress. WorldWatch Paper 136. WorldWatch Institute.

Brown, L. R. 2002. Ecoeconomia. Centre Unesco de Catalunya, Barcelona.

Brown, N. Out of control: fires and forestry in Indonesia. 1998. Trends in Ecology and Evolution 13: 41-41.

Brummer, E. C. 1998. Diversity, stability, and sustainable American agriculture. Agronomy Journal 90: 1-2.

Bunge , J., and M. Fitzpatrick. 1993. Estimating the number of species: a review. Journal of the American Statistical Association 88: 364-373.

Butler, J. A. V. 1946. Life and the second law of thermodynamics. Nature, issue 4005: 153-154.

Cabeza, M., and A. Moilanen. 2001. Design of reserve networks and the persistence of biodiversity. Trends in Ecology and Evolution 16: 242-248.

Camacho, J., and R. V. Solé. 2001. Scaling in ecological size spectra. Europhysics Letters 55: 774-780.

Cameron, T. 2002. 2002: the year of the "diversity-ecosystem function" debate. Trends in Ecology and Evolution 17: 495-496.

Campbell, C. L., and L. V. Madden. 1990. Introduction to Plant Disease Epidemiology. John Wiley and Sons, New York.

Caswell, H. 1976. Community structure: a neutral model analysis. Ecological Monographs 46: 327-354.

Caves, C. M. 1993. Information and entropy. Physical Review E 47: 4010-4017.

Cerf, N. J., and C. Adami. 1997. Negative entropy and information in quantum mechanics. Physical Review Letters 79: 5194-5197.

Chaitin, G. J. 1979. Toward a mathematical definition of life. In R. D. Levine, and M. Tribus (eds.), The maximum entropy formalism, MIT Press.

Champernowne, D. G. 1953. A model of income distribution. Economic Journal 63: 318-351.

Chao, A. 1984. Nonparametric estimation of the number of classes in a population. Scandinavian Journal of Statistics 11: 265-270.

Chapin III, F. S., E. S. Zavaleta, V. T. Eviner, R. L. Naylor, P. M. Vitousek, H. L. Reynolds, D. U. Hooper, S. Lavorel, O. E. Sala, S. E. Hobbie, M. C. Mack, and S. Díaz. 2000. Consequences of changing biodiversity. Nature 405: 234-242.

Chust, G., D. Ducrot, J. L. Riera, and J. L. Pretus. 1999. Characterizing human-modelled landscapes at a stationary state: a case study of Minorca, Spain. Environmental Conservation 26: 322-331.

Clar, S., B. Drossel, and F. Schwabl. 1994. Scaling laws and simulation results for the self-organized critical forest-fire model. Physical Review E50: 1009-1018.

Clar, S., B. Drossel, and F. Schwabl. 1996. Forest fires and other examples of self-organized criticality. Journal of Physics: Condensed Matter 8: 6803-6824.

Cleveland, C. J., and M. Ruth. 1997. When, where, and by how much do biophysical limits constrain the economic process? A survey of Nicholas Georgescu-Roegen contribution to ecological economics. Ecological Economics 22: 203-223.

Cochrane, M. A. 1998. Forest fires in the Brazilian Amazon. Conservation Biology 12: 948-950.

Cohen, J. E., and D. Tilman. 1996. Biosphere 2 and biodiversity: the lessons so far. Science 274: 1150-1151.

Cohen, M. L. 1992. Epidemiology of drug resistance: implications for a post-antimicrobial era. Science 257: 1050-1055.

Colwell, R. K., and J. A. Coddington. 1994. Estimating terrestrial biodiversity through extrapolation. Philosophical Transactions of the Royal Society of London, Series B 345: 101-118.

Common, M. S. 1996 Beckerman and his critics on strong and weak sustainability: confusing concepts and conditions. Environmental Values 5 : 83-87.

Connell, J. H. 1978. Diversity in tropical rain forests and coral reefs. Science 199: 1302-1310.

Constantino, R. F., R. A. Desharnais, J. M. Cushing, and B. Dennis. 1997. Chaotic dynamics in an insect population. Science 275: 389-391.

Cooke, G. D. 1971. Ecology of space travel. Pp. 498-509 in: E. P. Odum, Fundamentals of Ecology, 3th ed. Saunders, Philadelphia.

Corbet, S. 1943. Results obtained with Malayan butterflies. Pp. 42-44 in: R. A. Fisher, A. S. Corbet, and C. B. Williams, The relation between the number of species and the number of individuals in a random sample of an animal population. Journal of Animal Ecology 12: 42-58.

Costanza, R. (ed.). 1998. The value of ecosystem services. Special Issue. Ecological Economics 25, issue 1.

Costanza, R., and B. C. Patten. 1995. Defining and predicting sustainability. Ecological Economics 15: 193-196.

Costanza, R., R. d'Arge, R. de Groot, S. Farber, M. Grasso, B. Hannon, K. Limburg, S. Naeem, R. V. O'Neill, J. Paruelo, R. G. Raskin, P. Sutton, and M. van den Belt. 1997. The value of the world's ecosystem services and natural capital. Nature 387: 253-260.

Coveney, P. V. 1988. The second law of thermodynamics: entropy, irreversibility and dynamics. Nature 333: 409-415.

Cox, P. M, R. A. Betts, C. D. Jones, S. A. Spall, and I. J. Totterdell. 2000. Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. Nature 408: 184-187.

Crosson, P. 1995. Soil erosion estimates and costs. Science 269: 461-464.

Daly, H. E. (ed.). 1980. Economics, Ecology, Ethics. Essays toward a Steady-State Economy. W. H. Freeman and Co. New York [also in Spanish, Fondo de Cultura Económica, México D. F.]

Daly, H. E. 1991. Steady-State Economics, 2nd ed. Island Press, Washington.

Daly, H. E. 1995. On Wilfred Beckerman's critique of sustainable development. Environmental Values 4: 49-55.

Daly, H. E. (ed.). 1997. The contribution of Nicholas Georgescu-Roegen. Special Issue. Ecological Economics 22, issue 3.

Damuth, J. 1981. Population density and body size in mammals. Nature 290: 699-700.

Darwin, C. 1859. On the Origin of Species. John Murray, London. Facsimile: Harvard University Press, 1964.

Death, R. G. 1996. The effect of habitat stability on benthic invertebrate communities: the utility of species abundance distributions. Hydrobiologia 317: 97-107.

DeVries, P. I., D. Murray, and R. Lande. 1997. Species diversity in vertical, horizontal, and temporal dimensions of a fruit-feeding butterfly community in an Ecuadorian rainforest. Biological Journal of the Linnean Society 62: 343-364.

Dewdney, K. 1998. A general-theory of the sampling process with applications to the veil line. Theoretical Population Biology 54: 294-302.

Dewey, T. G.. 1997. Algorithmic complexity and thermodynamics of sequence-structure relationships in proteins. Physical Review E 56: 4545-4552.

Díaz, S., and M. Cabido. 2001. Vive la différence: plant functional diversity matters to ecosystem processes. Trends in Ecology and Evolution 16: 646-655.

Diserud, O. H., and S. Engen. 2000. A general and dynamic species abundance model, embracing the lognormal and the gamma models. American Naturalist 155: 497-511.

Doak, D. F., D. Bigger, E. K. Harding, M. A. Marvier, R. E. O'Malley, and D. Thomson. 1998. The statistical inevitability of stability-diversity relationships in community ecology. American Naturalist 151: 264-276.

Dragulescu, A., and V. M. Yakovenko. 2001. Exponential and power-law probability distributions of wealth and income in the United Kingdom and the United States. Physica A 299: 213-221.

Drèze, J. 1998. Distribution matters in cost-benefit analysis: Comment on K. A. Brekke. Journal of Public Economics 70: 485-488.

Drossel, B. 2001. Biological evolution and statistical physics. Advances in Physics 50: 209-295.

Drossel, B., and F. Schwabl. 1992. Self-organized critical forest-fire model. Physical Review Letters 69: 1629-1632.

DuMouchel, W. H. 1973. Stable distributions in statistical inference : 1- Symmetric stable distributions as compared to other symmetric long-tailed distributions. Journal of the American Statistical Association 68: 469-477.

Earman, J. 1974. An attempt to add a little direction to “the problem of the direction of time”. Philosophy of Science 41: 15-47.

Ehrlich, P. R, and A. H. Ehrlich. 1981. Extinction. Random House, New York.

Eigen, M., J. McCaskill, and P. Schuster. 1988. Molecular quasi-species. Journal of Physical Chemistry 92: 6881-6891.

Ekins, P., C. Folke, and R. Costanza. 1994. Trade, environment and development: the issues in perspective. Ecological Economics 9: 1-12.

Elton, C. S. 1958. The Ecology of Invasions by Animals and Plants. Methuen & Co., London.

Emmerson, M. C., M. Solan, C. Emes, D. M. Paterson, and D. Raffaelli. 2001. Consistent patterns and idiosyncratic effects of biodiversity in marine ecosystems. Nature 411: 73-77.

Engen, S. 1978. Stochastic Abundance Models. Chapman and Hall, London.

Engen, S., and R. Lande. 1996a. Population-dynamic models generating the lognormal species abundance distribution. Mathematical Biosciences 132: 169-183.

Engen, S., and R. Lande. 1996b. Population dynamic models generating species abundance distributions of the gamma type. Journal of Theoretical Biology 178: 325-331.

Engen S., O. Bakke, and A. Islam. 1998. Demographic and environmental stochasticity - Concepts and definitions. Biometrics 54: 840-846.

Fahn, P. N. 1996. Maxwell's Demon and the entropy cost of information. Foundations of Physics 26: 71-93.

Fath, B. D., B. C. Patten, and J. S. Choi. 2001. Complementarity of ecological goal functions. Journal of Theoretical Biology 208: 493-506.

Faucheux, S., and M. O'Connor. 1998. Valuation for Sustainable Development. Edward Elgar, Chentelham, UK.

Feitel, R., and W. Ebeling. 1989. Evolution of Complex Systems. Self-Organization, Entropy and Development. Kluwer Academic Publishers, Dordrecht.

Fisher, R. A. 1943. A theoretical distribution for the apparent abundance of different species. Pp. 54-57 in: R. A. Fisher, A. S. Corbet, and C. B. Williams, The relation between the number of species and the number of individuals in a random sample of an animal population. Journal of Animal Ecology 12: 42-58.

Flos, J. 1984. Ecología entre la magia y el tópico. Omega, Barcelona.

Flos, J. (ed.). 1995. Ordre i caos en ecologia. Publicacions Universitat de Barcelona.

Flos, J., and A. Carbonell. 1991. Neritic zooplankton succession studied with statistical entropy measures: the role of space and the hierarchical matching of the environment. Pp. 187-201 in: J. D. Ros and N. Prat, Homage to Ramon Margalef; or, Why is there such a pleasure in studying nature. Oecologia Aquatica 10.

Flos, J. and E. Gutiérrez. 1995. In: J. Flos (ed.). Ordre i caos en ecologia. Publicacions Universitat de Barcelona.

Folch, R., and J. M. Camarassa (eds.). 1993. Biosfera, vol. 5: Mediterrànies. Enciclopèdia Catalana, Barcelona.

Folke, C., P. Ekins, and R. Costanza (eds.). 1994. Trade and environment. Special issue. Ecological Economics 9, issue 1.

Franklin, J. F., D. R. Berg, D. A. Thornburg, and J. C. Tappeiner. 1997. Alternative silvicultural approaches to timber harvesting: variable retention harvest systems. Pp. 111-139 in: K. A. Kohm, and J. F. Franklineds (eds.), Creating a forestry for the 21st century. Island Press, Washington, DC.

Franklin, R. B., J. L. Garland, C. H. Bolster, and A. L. Mills. 2001. Impact of dilution on microbial community structure and functional potential: comparison of numerical simulations and batch culture experiments. Applied and Environmental Microbiology 67: 702-712.

Frieden, R. 1985. Estimating occurrence laws with maximum probability, and the transition to entropic estimators. Pp. 133-169 in: C. R. Smith, and W. T. Grandy Jr. (eds.), Maximum-Entropy and Bayesian Methods in Inverse Problems. Reidel, Dordrecht.

Frontier, S. 1985. Diversity and structure in aquatic ecosystems. Pp. 253-312 in: M. Bernes (ed.), Oceanography and Marine Biology, an Annual Review 23. Aberdeen University Press, Aberdeen.

Galambos, J. 1987. The Asymptotic Theory of Extreme Order Statistics (2nd. ed.). Robert E. Krieger Publishing Company, Malabar, Florida.

Gardner, G. 1998. El reciclatge dels residus orgànics. Pp. 107-126 in: Worldwatch Institute, L'Estat del Món 1998. Centre Unesco de Catalunya, Barcelona.

Gardner, M. 1978. White and Brown music, fractal curves and one-over-f fluctuations. Scientific American 238: 16-32 (April) [also in Spanish: Investigación y Ciencia n. 21 (junio 1978): 104-113].

Gardner, M. R., and W. R. Ashby. 1970. Connectance of large dynamic (cybernetic) systems: critical values for stability. Nature 228: 784-784.

Garrett, K. A., and C. C. Mundt. 1999. Epidemiology in mixed host populations. Phytopathology 89: 984-990.

Garrod, C. 1995. Statistical Mechanics and Thermodynamics. Oxford University Press.

Gaston, K. J. 2000. Global patterns in biodiversity. Nature 405: 220-227.

Gaston, K. J., and T. M. Blackburn. 2000. Pattern and Process in Macroecology. Blackwell Science Ltd., Oxford

Gautam, M., U. Lele, H. Kartodihardjo, A. Khan, I. Erwinsyah, and S. Rana. 2000. Indonesia. The challenges of the World Bank involvement in forests. The World Bank, Washington.

Georgescu-Roegen, N. 1971. The Entropy Law and the Economic Process. Harvard University Press.

Gibson, G. J. 1997. Markov chain Monte Carlo methods for fitting spatiotemporal stochastic models in plant epidemiology. Applied Statistics 46: 215-233.

Givnish, T. J. 1994. Does biodiversity beget stability? Nature 371: 113-114.

Goldammer, J. G. 1999. Forests on fire. Science 284: 1782-1783.

Golley, F. B. (ed.). 1977. Ecological Succession. Dowden, Hutchinson & Ross, Inc. Stroudburg, Pennsylvania.

Goodland, R., and H. Daly. 1993. Why Northern income growth is not the solution to Southern poverty. Ecological Economics 8: 85-101.

Grassberger, P. 1993. On a self-organized critical forest-fire model. Journal of Physics A: Mathematical and General 26: 2081-2089.

Gray, J. S. 1987. Species-abundance patterns. Pp. 53-67 in: J. H. R. Gee, and P. S. Giller (eds.), Organisation of Communities: Past and Present. Blackwell Scientific, Oxford

Grenfell, B. T., O. N. Bj∅rnstad, and J. Kappey. 2002. Travelling waves and spatial hierarchies in measles epidemics. Nature 414: 716-723.

de Groot, S. R., and P. Mazur. 1962. Non-equilibrium thermodynamics. North Holland Publishing Company, Amsterdam.

Grünbaum, A. 1973. Philosophical problems of space and time 2nd ed. D. Reidel, Dordrecht.

Guzetti, F., B. D. Malamud, D. L. Turcotte, P. Reichenbach. 2002. Power-law correlation of landslide areas in Central Italy. Earth and Planetary Science 195: 169-183.

Hall, S. J., and D. G. Rafaelli. 1993. Food webs: theory and reality. Advances in Ecological Research 24: 187-239.

Hanley, N., and C. L. Spash. 1993. Cost-Benefit Analysis and the Environment. Edward Elgar, Aldershot, Vermont.

Hanski, I., P. Turchin, E. Korpimäki, H. Henttonen. 1993. Population oscillations of boreal rodents: regulation by mustelid predators leads to chaos. Nature 364: 232-235.

Harris, B. 1959. Determining bounds on integrals with applications to cataloguing problems. Annals of Mathematical Statistics 30: 521-548.

Harte, J., A. Kinzig, and J. Green. 1999. Self-similarity in the distribution and abundance of species. Science 284: 334-336.

Hastings, A., C. L. Hom, S. Ellner, P. Turchin, H. C. J. Godfray. 1993. Chaos in ecology: is Mother Nature a strange attractor? Annual Review of Ecology and Systematics 24: 1-33.

Hawking, S. W. 1987. The direction of time. New Scientist, 9 July: 46-49.

Hawking, S. W., R. Laflamme and G. W. Lyons. 1993. Origin of time asymmetry. Physical Review D 47: 5342-5356.

Haydon, D. 1994. Pivotal assumptions determining the relationship between stability and complexity: an analytical synthesis of the stability-complexity debate. American Naturalist 144: 14-29.

He, F., and P. Legendre. 1996. On species-area relations. American Naturalist 148: 719-737.

He, F., P. Legendre, and J. V. LaFrankie. 1997. Distribution patterns of tree species in a Malaysian tropical rain forest. Journal of Vegetation Science 8: 105-114.

Hector, A., B. Schmid, C. Beierkuhnlein, M. C. Caldeira, M. Diemer, P. G. Dimitrakopoulos, J. A. Finn, H. Freitas, P. S. Giller, J. Good, R. Harris, P. Högberg, K. Huss-Danell, J. Joshi, A. Jumpponen, C. Körner, P. W. Leadley, M. Loreau, A. Minns, C. P. H. Mulder, G. O'Donovan, S. J. Otway, J. S. Pereira, A. Prinz, D. J. Read, M. Scherer-Lorenzen, E.-D. Schulze, A.-S. D. Siamantziouras, E. M. Spehn, A. C. Terry, A. Y. Troumbis, F. I. Woodward, S. Yachi, and J. W. Lawton. 1999. Plant diversity and productivity experiments in European grasslands. Science 286: 1123-1127.

Heltshe, J. F., and N. E. Forrester. 1983. Estimating species richness using the jacknife procedure. Biometrics 39: 1-11.

Henderson, N., and W. J. Sutherland. 1996. Two truths about discounting and their environmental consequences. Trends in Ecology and Evolution 11: 527-528.

Herendeen, R. A. 2000. Ecological footprint is a vivid indicator of indirect effects. Ecological Economics 32: 357-358.

Holland, A. 1997. Substitutability. Or, why strong sustainability is weak and absurdly strong sustainability is not absurd. Pp. 119-134 in: J. Foster (ed.), Valuing Nature? Routledge, London.

Holmgren, M., M. Scheffer, E. Ezcurra, J. R. Rodríguez, and G. M. J. Mohren. 2001. El Niño effects on the dynamics of terrestrial ecosystems. Trends in Ecology and Evolution 16: 89-94.

Holt, R. D. 1977. Predation, apparent competition, and the structure of prey communities. Theoretical Population Biology 12: 197-229.

Horwich P. 1987. Asymmetries in time. MIT Press, Cambridge.

Howarth, R. B., and R. B. Norgaard. 1993. Intergenerational transfers and the social discount rate. Environmental and Resource Economics 3: 337-358.

Hubbell, S. P. 1997. A unified theory of biogeography and relative species abundance and its application to tropical rain forests and coral reefs. Coral Reefs 16, Suppl.: S9-S21.

Hubbell, S. P. 2001. The Unified Neutral Theory of Biodiversity and Biogeography. Princeton University Press.

Hughes, J. B., and O. L. Petchey. 2001. Merging perspectives on biodiversity and ecosystem functioning. Trends in Ecology and Evolution 16: 222-223.

Hughes, J. B., and J. Roughgarden. Species diversity and biomass stability. American Naturalist 155: 618-627.

Hurlbert, S. H. 1971. The nonconcept of species diversity: a critique and alternative parameters. Ecology 52: 577-586.

Huston, M. A. 1997. Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. Oecologia 110: 449-460.

Ibáñez, C., N. Prat, A. Canicio, and A. Curcó. 1999. El delta del Ebro, un sistema amenazado. Bakeaz/Coagret, Bilbao.

IPCC. 2001a. Climate Change 2001: the Scientific Basis. Cambridge University Press.

IPCC. 2001b. Climate Change 2001: Impacts, Adaptation, and Vulnerability. Cambridge University Press.

IUCN. 1980. World Conservation Strategy, IUCN.

Ives, A. R., and J. B. Hughes. 2002. General relationships between diversity and stability in competitive systems. American Naturalist 159: 388-395.

Ives, A. R., K. Gross, and J. L. Klug. 1999. Stability and variability in competitive communities. Science 286: 542-544.

Jacobs, M. 1995. Sustainable development, capital substitution and economic humility : a response to Beckerman. Environmental Values 4: 57-68.

Janzen, D. H. 1970. Herbivores and the number of tree species in tropical forests. American Naturalist 104: 501-528.

Jaynes, E. T. 1983. Papers on Probability, Statistics and Statistical Physics (Rosenkratz, R. D., ed.). Reidel, Dordrecht.

J∅rgensen, S. E. 1992. Integration of Ecosystem Theories: A Pattern. Kluwer Academic Publishers, Dordrecht.

J∅rgensen, S. E. 1994. Fundamentals of Ecological Modelling, 2nd ed. Elsevier, Amsterdam.

J∅rgensen, S. E. 1999. State-of-the-art of ecological modelling with emphasis on development of structural dynamics models. Ecological Modelling 120: 75-96.

J∅rgensen, S. E. (ed.). 2001a. Thermodynamics and Ecological Modelling. Lewis Publishers, Boca Raton.

J∅rgensen, S. E. 2001b. A tentative Fourth Law of thermodynamics. Pp. 303-348 in: S. E. J∅rgensen (ed.). Thermodynamics and Ecological Modelling. Lewis Publishers, Boca Raton.

J∅rgensen, S. E., and H. Mejer. 1979. A holistic approach to ecological modelling. Ecological Modelling 7: 169-189.

J∅rgensen, S. E., and J. Padisák. 1996. Does the intermediate disturbance hypothesis comply with thermodynamics? Hydrobiologia 323: 9-12.

J∅rgensen, S. E., S. N. Nielsen, and H. Mejer. 1995. Emergy, environ, exergy and ecological modelling. Ecological Modelling 77: 99-109.

Jou, D. 1985. Introducció a la termodinàmica de processos biològics. Institut d'Estudis Catalans, Barcelona.

Jou, D., J. Casas-Vázquez and G. Lebon. 1993. Extended Irreversible Thermodynamics. Springer-Verlag, Berlin.

Kaberger, T., and B. Mansson. 2001. Entropy and economic processes - physics perspectives. Ecological Economics 36: 165-179.

Kaiser, J. 2000. Rift over biodiversity divides ecologists. Science 289: 1282-1283.

Kapp, K. W. 1983. Social costs, economic development and environmental disruption. University Press of America. Lanham, MD.

Kareiva, P., and U. Wennergren. 1995. Connecting landscape patterns to ecosystem and population processes. Nature 373: 299-302.

Kauffman, S. A., and S. Johnsen. 1991. Coevolution to the edge of chaos: coupled fitness landscapes, poised states, and coevolutionary avalanches. Journal of Theoretical Biology 149: 467-505.

Kaufman, J. H., D. Brodbeck, and O. R. Melroy. 1998. Critical biodiversity. Conservation Biology 12: 521-532.

Keeley, J. E., C. J. Fotheringham, and M. Morais. 1999. Reexamining fire suppression impacts on brushland fire regimes. Science 284: 1829-1832.

Kendall, M. G. 1970. Rank Correlation Mehods, 4th. ed. Griffin, London.

Kerr, R. A. 1998. The next oil crisis looms large - and perhaps close. Science 281: 1128-1131.

Kirchner, J. W., and A. Weil. 2000. Delayed biological recovery from extinctions throughout the fossil record. Nature 404: 177-179.

Kirkpatrick, S., C. D. Gelatt, Jr., and M. P. Vecchi. 1983. Optimization by simulated annealing. Science 220: 671-680.

Kleidon, A., M. Heimann. 2000. Assessing the role of deep rooted vegetation in the climate system with model simulations: mechanisms, comparison to observations and implications for Amazonian deforestation. Climate Dynamics 16: 183-199.

Komlos, J. 1998. Anthropometric history: the achievements of two decades of research. Pp. 219-235 in: C.-E. Núñez (ed.), Debates and Controversies in Economic History, Centro de Estudios Ramón Areces, Madrid.

Krugman, P. 1981. Trade, accumulation, and uneven development. Journal of Development Economics 8: 149-161.

Krugman, P. 1997. La Organización Espontánea de la Economía. Antoni Bosch, Barcelona.

Krutilla, A. V. 1967. Conservation reconsidered. American Economic Review 57: 777-786.

Landauer, R. 1961. Irreversibility and heat generation in the computing process. IBM Journal of Research and Development 3: 183-191.

Landauer, R. 1988. Dissipation and noise immunity in computation and communication. Nature 335: 779-784.

Lande, R. 1996. Statistics and partioning of species diversity, and similarity among multiple communities. Oikos 76: 5-13.

Lang, G. 2000. Geography as destiny? Science 288: 982-983.

Langton, C. G. 1990. Computation at the edge of chaos : phase transitions and emergent computation. Physica D 42: 12-37.

Langton, C. G. 1995. Editor's introduction. Pp. ix-xi in C. G. Langton (ed.). Artificial Life. Massachusetts Institute of Technology.

Laurance, W. F. 1998. A crisis in the making: responses of Amazonian forests to land use and climate change. Trends in Ecology and Evolution 13: 411-414.

Lawton, R. O., U. S. Nair, R. A. Pielke Sr., and R. M. Welch. 2001. Climatic impact of tropical lowland deforestation on nearby montane cloud forests. Science 294: 584-587.

Leff, H. S., and A. F. Rex (eds.). 1990. Maxwell Demon. Princeton University Press.

Lehman, C., And D. Tilman. 2000. Biodiversity, stability, and productivity in competitive communities. American Naturalist 156: 534-552.

Leitner, W. A., and M. L. Rosenzweig. 1997. Nested species-area curves and stochastic sampling: a new theory. Oikos 79: 503-512.

Lenssen, N. 1992. Haciendo frente a los residuos nucleares. Pp. 83-113 in : Worldwatch Institute, La Situación del Mundo 1992. CIP and Apóstrofe, Barcelona.

Levine, J. M. 2002. Species diversity and relative abundance in metacommunities. Trends in Ecology and Evolution 17: 99-100.

Lithgow, G. J., and T.B.L. Kirkwood. 1996. Mechanisms and evolution of aging. Science 273: 80-80.

López Cachero, M. 1990. Fundamentos y métodos de estadística, 9ª ed. Pirámide, Madrid.

Loreau, M. 2000. Biodiversity and ecosystem functioning: recent theoretical advances. Oikos 91: 3-17.

Loreau, M., and A. Hector. 2001. Partitioning selection and complementarity in biodiversity experiments. Nature 412: 72-76.

Lotka, A. 1922. Contribution to the energetics of evolution. Proceedings of the National Academy of Sciences of the U. S. A. 8: 147-150.

Lovelock, J. E. 1988. The Ages of Gaia. W. W. Norton Co., New York.

Lurié, D., and J. Wagensberg. 1983. Note on biomass diversity in ecology. Bulletin of Mathematical Biology 45: 287-293.

Lux, T., and M. Marchesi. 1999. Scaling and criticality in a stochastic multi-agent model of a financial market. Nature 498: 498-661.

MacArthur, R. 1955. Fluctuations of animal populations and a measure of community stability. Ecology 36: 533-536.

MacArthur, R. 1960. On the relative abundance of species. American Naturalist 94: 25-36.

MacArthur, R. 1969. Species packing, or what competition minimizes. Proceedings of the National Academy of Sciences of the U. S. A. 64: 1369-1375.

MacArthur, R. 1970. Species packing and competitive equilibrium for many species. Theoretical Population Biology 1: 1-11.

MacArthur, R. 1972. Geographical Ecology. Harper & Row, NewYork.

MacArthur, R., and E. O. Wilson. 1967. The Theory of Island Biogeography. Princeton University Press, Princeton

Maddux, R. D., and K. Athreya. 1999. On the distribution and abundance of species. Science 286: 1647a.

Magurran, A. E. 1988. Ecological Diversity and its Measurement. Croom Helm, London.

Malamud, B. D., G. Morein, and D. L. Turcotte. 1998. Forest fires: an example of self-organized critical behavior. Science 281: 1840-1842.

Mandelbrot, B. B. 1983. The Fractal Geometry of Nature. W. H. Freeman and co., New York.

Mandelbrot, B. B. 1984. Los objetos fractales, 2ª ed. Tusquets, Barcelona.

Manly, B. F. J. 1991. Randomization and Montecarlo Methods in Biology. Chapman and Hall, London.

Manrubia, S. C., and D. H. Zanette. 1998. Intermittency model for urban development. Physical Review E 58: 295-302.

Mansson, A., and J. M. McGlade. 1993. Ecology, thermodynamics and H. T. Odum conjectures. Oecologia 93: 582-596.

Mantegna, R. N., and H. E. Stanley. 1995. Scaling behaviour in the dynamics of an economic index. Nature 376: 46-49.

Margalef, R. 1956. Información y diversidad específica en las comunidades de organismos. Investigación Pesquera 3: 99-106.

Margalef, R. 1963. On certain unifying principles in ecology. American Naturalist 97: 357-373.

Margalef, R. 1966. Aplicació de la cibernètica a l'estudi de sistemes biològics al nivell de l'ecosistema. Treballs de la Societat Catalana de Biologia 20: 74-81.

Margalef, R. 1968. Perspectives in Ecological Theory. University of Chicago Press.

Margalef, R. 1976. Biología de los embalses. Investigación y Ciencia 1: 50-62.

Margalef, R. 1978. Life-forms of phytoplankton as survival alternatives in an unstable environment. Oceanologica Acta 1: 493-509.

Margalef, R. 1980. La Biosfera entre la termodinámica y el juego. Omega, Barcelona.

Margalef, R. 1986. Sucesión y evolución: su proyección biogeográfica. Paleontologia i Evolució 20: 7-26.

Margalef, R. 1989. Ecología Omega, Barcelona.

Margalef, R. 1991. Teoría de los Sistemas Ecológicos. Publicacions de la Universitat de Barcelona.

Margalef, R. 1994. Through the looking glass: how marine phytoplankton appears through the microscope when graded by size and taxonomically sorted. Scientia Marina 58: 87-101.

Margalef, R. 1997. Our Biosphere. Ecology Institute, Oldendorf/Luhe.

Margalef, R. 2001. La superfície del planeta i l'organització de la Biosfera: reacció als nous mecanismes afegits pel poder creixent dels humans. Lecture in Barcelona, Residència d'Investigadors del CSIC.

Margolin, B. H., and W. Maurer. 1976. Tests of the Kolmogorov-Smirnov for exponential data with unknown scale, and related problems. Biometrika 63: 149-160.

Margules, C. R., and R. L. Pressey. 2000. Systematic conservation planning. Nature 405: 243-253.

Marín, J., and R. V. Solé. 1999. Macroevolutionary algorithms: a new optimization method on fitness landscapes. IEEE Transactions on Evolutionary Computation 3: 272-286.

Marlin, P. G. 1984. Goodness-of-fit tests for the Pareto and lognormal distributions based on multiply truncated samples. Communications in Statistics A 13: 1965-1079.

Martínez Alier, J. 1984. L'ecologisme i l'economia : història d'unes relacions amagades. Edicions 62, Barcelona.

Martínez Alier, J. 1994. De la economía ecológica al ecologismo popular, 2nd ed. Icaria, Barcelona.

Martínez Alier, J. 1997. Deuda ecológica y deuda externa. Ecología Política 14: 157-173.

Martínez Alier, J. 1999. Introducció a l'economia ecològica. Rubes and Departament de Medi Ambient de la Generalitat de Catalunya, Barcelona.

Martínez Alier, J., and K. Schlüpmann. 1990. Ecological Economics: Energy, Environment and society. Basil Blackwell, Oxford.

Martínez Alier, J., and M. O'Connor. 1996. Ecological and economic distribution conflicts. Pp. 153-183 in: R. Costanza, O. Segura, and J. Martínez Alier (eds.), Getting down to Earth. Island Press, Washington D. C.

Martínez Alier, J., G. Munda, and J. O'Neill. 1998. Weak comparability of values as a foundation for ecological economics. Ecological Economics 26: 277-286.

Matson, P. A., and M. D. Hunter (eds.). 1992. The relative contributions of top-down and bottom-up forces in population and community ecology. Ecology 73: 723-765.

Matsuno, K. 1978. Evolution of dissipative system: a theoretical basis of Margalef's principle on ecosystem. Journal of Theoretical Biology 70: 23-31.

Maurer, B. A. 1999. Untangling ecological complexity. University of Chicago Press.

Maxwell, J. C. 1871. Theory of heat. Longsman, London.

May, R. M. 1972. Will a large complex system be stable? Nature 238: 413-414.

May, R. M. 1973. Stability and Complexity in Model Ecosystems. Princeton University Press.

May, R. M. 1974a. How many species: some mathematical aspects of the dynamics of populations. Lectures on Mathematics in the Life Sciences 6: 61-98.

May, R. M. 1974b. Biological populations with nonoverlapping generations:stable points, stable cycles and chaos. Science 186: 645-647.

May, R. M. 1975. Patterns of species abundance and diversity. Pp. 81-120 in: M. L. Cody, and J. M. Diamond (eds.), Ecology and Evolution of Communities. The Belknap Press of Harvard University Press, Cambridge

May, R. M. 1977. Food lost to pests. Nature 267: 669-670.

May, R. M. 2000. Relation between diversity and stability, in the real world. Science 290: 714-715.

McCallum, H . 2000. Achievement and challenge. Trends in Ecology and Evolution. 15: 352-353.

McCann, K. S. 2000. The diversity-stability debate. Nature 405: 228-233.

McCann, K., A. Hastings, and G. R. Huxel. 1998. Weak trophic interactions and the balance of nature. Nature 395: 794-798.

McGrady-Steed, J., and P. J. Morin. 2000. Biodiversity, density compensation, and the dynamics of populations and functional groups. Ecology 81: 361-373.

McGrady-Steed, J., P. M. Harris, and P. J. Morin. 1997. Biodiversity regulates ecosystem predictability. Nature 390: 162-165.

McNaughton, S. J. 1977. Diversity and stability of ecological communities: a comment on the role of empiricism in ecology. American Naturalist 111: 515-525.

McNaughton, S. J. 1978. Stability and diversity of ecological communities. Nature 274: 251-253.

McNaughton, S. J. 1985. Ecology of a grazing ecosystem: the Serengeti. Ecological Monographs. 55: 259-294.

Meggers, B. J. 1994. Archaeological evidence for the impact of mega-Niño events on Amazonia during the past two millennia. Climatic Change 28: 321-338.

Mehta, M. L. 1991. Random matrices. Revised and enlarged second edition. Academic Press, San Diego.

Merrell, D. S., S. M. Butler, F. Qadri, N. A. Dolganov, A. Alam. M. B. Cohen, S. B. Calderwood, G. K. Schoolnik, and A. Camilli. 2002. Host-induced epidemic spread of the cholera bacterium. Nature 417: 642-645.

Minnich, R. A. 1983. Fire mosaics in Southern California and Northern Baja California. Science 219: 1287-1294.

Minnich, R. A., and Chou-Yue-Hong. 1997. Wildland fire patch dynamics in the chaparral of southern California and Northern Baja California. International Journal of Wildland Fire 7: 221-248.

Mooney, H. A. 1998. The Globalization of Ecological Thought. Ecology Institute, Oldendorf/Luhe.

Morowitz, H. J. 1968. Energy flow in biology. Academic Press, New York.

Morse, D. R., J. H. Lawton, M. M. Dodson, and M. H. Williamson. 1985. Fractal dimension of vegetation and the distribution of arthropod body lengths. Nature 314: 731-733.

Müller, F. 1997. State-of-the-art in ecosystem theory. Ecological Modelling 100: 135-161.

Mundt, C. C., and L. S. Brophy. 1988. Influence of number of host genotype units on the effectiveness of host mixtures for disease control: a modeling approach. Phytopathology 78: 1087-1094.

Muradian, R., and J. Martínez Alier. 2001. Trade and the environment: from a "Southern" perspective. Ecological Economics 36: 281-297.

Myers, N. 1995a. The world's forests: need for a policy appraisal. Science 268: 162-165.

Myers, N. 1995b. Environmental unknowns. Science 269: 358-360.

Myers, N. 2001. Playing the numbers game. Nature 410: 631-632.

Naeem, S. 1998. Species redundancy and ecosystem reliability. Conservation Biology 12: 39-45.

Naeem, S. 2002. Biodiversity equals instability? Nature 416: 23-24.

Naeem, S., and S. Li. 1997. Biodiversity enhances ecosystem reliability. Nature 390: 507.

Naeem, S., F. S. Chapin III, R. Costanza, P. R. Ehrlich, F. B. Golley, D. U. Hooper, J. H. Lawton, R. V. O'Neill, H. A. Mooney, O. E. Sala, A. J. Symstad, and D. Tilman. 1999. Biodiversity and ecosystem functioning: maintaining natural life support processes. Issues in Ecology 4.

Nahmias, J., Tephany, H., Duarte, J., and S. Letaconnoux. 2000. Fire spreading experiments on heterogeneous fuel beds - Applications of percolation theory. Canadian Journal of Forest Reasearch 30: 1318-1328.

Nearing, M. A., M. J. M. Romkens, L. D. Norton, D. E. Scott, F. E. Rhoton, J. M. Laflen, D. C. Flanagan, C. V. Alonso, R. L. Binger, S. M. Dabney, O. C. Doering, C. H. Huang, K. C. McGregor, and A. Simon. 2000. Measurements and models of soil loss rates. Science 290: 1300-1301.

Neidhart, F. C.. 1987. Chemical composition of Escherichia coli. Pp. 3-6 in: F. C. Neidhart (ed.), *Escherichia coli* and *Salmonella typhimurium*. Cellular and molecular biology. Vol. 1. American Society for Microbiology, Washington.

Nepstad, D. C., A. Moreira, A. Veríssimo, P. Lefebvre, P. Schlesinger, C. Potter, C. Nobre, A. Setzer, T. Krug, A. C. Barros, A. Alencar, and J. R. Pereira. 1998. Forest fire prediction and prevention in the Brazilian Amazon. Conservation Biology 12: 951-953.

Nepstad, D. C., A. Veríssimo, A. Alencar, C. Nobre, E. Lima, P. Lefebvre, P. Schlesinger, C. Potter, P. Moutinho, E. Mendoza, M. Cochrane, and V. Brooks. 1999. Large-scale impoverishment of Amazonian forests by logging and fire. Nature 398: 505-508.

Neutel, A.-M., J. A. P. Heesterbeek, and P. C. de Ruiter. 2002. Stability in real food webs: weak links in long loops. Science 296: 1120-1123.

Newman, M. E. J. 1997. A model of mass extinction. Journal of Theoretical Biology 189: 235-252.

Newman, M. E. J., I. Jensen, and R. M. Ziff. 2002. Percolation and epidemics in a 2-dimensional small world. Physical Review E 6502: 1904.

Nicolis, G., and I. Prigogine. 1977. Self-organization in non-equilibrium systems. From dissipative structures to order through fluctuations. John Wiley & Sons, New York.

Niklasson, M., and A. Granström. 2000. Numbers and sizes of fires: long-term spatially explicit fire history in a Swedish boreal landscape. Ecology 81: 1484-1499.

Odum, E. P. 1953. Fundamentals of Ecology. Saunders, Philadelphia.

Odum, E. P. 1969. The strategy of ecosystem development. Science 164: 262-270.

Odum, E. P. 1971. Fundamentals of Ecology, 3th ed. Saunders, Philadelphia.

Odum, H. T. 1971. Environment, power and society. John Wiley & Sons.

Odum, H. T. 1988. Self-organization, transformity and information. Science 242: 1132-1139.

Odum, H. T. 1995. Energy systems concepts and self-organization: a rebuttal. Oecologia 104: 518.-522

Odum, H. T. 1996. Environmental Accounting. Emergy and Environmental Decision Making. John Wiley & Sons, New York.

Odum, H. T, and R. C. Pinkerton. 1955. Time's speed regulator: the optimum efficiency for maximum power output in physical and biological systems. American Scientist 43: 331-343.

O'Neill, R. V., R. H. Gardner, and M. G. Turner. 1992. A hierarchical neutral model for landscape analysis. Landscape Ecology 7: 55-61.

Pachepsky, E., J. W. Crawford, J. L. Bown,, and G. Squire. Towards a general theory of biodiversity. Nature 410: 923-926.

Page, S. E., F. Siegert, J. O. Rieleys, H.-D. V. Boehm, A. Jaya, and S. Limin. 2002. The amount of carbon released from peat and forest fires in Indonesia during 1997. Nature 420: 61-65.

Pagel, M. D., P. H. Harvey, and H. C. J. Godfray. 1991. Species-abundance, biomass, and resource-use distributions. American Naturalist 138: 836-850.

Palmer, M. W. 1990. The estimation of species richness by extrapolation. Ecology 71: 1195-1198.

Palmer, M. W. 1991. Estimating species richness: the second-order jacknife reconsidered. Ecology 72: 1512-1513.

Pareto, V. 1897. Cours d'Économie Politique Professé a l'Université de Lausanne. F. Rouge. Vol. 2. Lausanne. Reprinted in: V. Pareto. 1964. Cours d'Économie Politique. Droz, Geneve.

Parker, J. A., and J. Maynard Smith. 1990. Optimality theory in evolutionary biology. Nature 348: 27-33.

Parkins, W. E. 1996. Fusion prospects. Science 271: 893-893.

Pascual, M., X. Rodó, S. P. Ellner, R. Colwell, and M. J. Bouma. 2000. Cholera dynamics and El Niño. Science 289: 1766-1769.

Pastor-Satorra, R., and A. Vespignani. 2000. Corrections to scaling in the forest-fire model. Physical Review E 61: 4854-4859.

Patten, B. C. 1975. Ecosystem linearization: an evolutionary design problem. American Naturalist 109: 529-539.

Patten, C. 1993. Toward a more holistic ecology and science: the contribution of H. T. Odum. Oecologia 93: 597-602.

Patten, C. 1995. Network integration of ecological extremal principles: exergy, emergy, power, ascendency and indirect effects. Ecological Modelling 79: 75-84.

Pearce, D. W. 1998. Economic valuation and ecological economics. Pp. 40-54 in: D. W. Pearce, Economics and Environment. Edward Edgar, Cheltenham.

Pearce, D. W. 1999. Sustainable development: measurement and policy. Lecture in Barcelona, Institut Català de Tecnologia.

Pearce, D. W., and G. D. Atkinson. 1993. Capital theory and the measurement of sustainable development: an indicator of "weak" sustainability. Ecological Economics 8: 103-108.

Pearce, F. 1995. Global row over value of human life. New Scientist 147: 7-7 (19 August 1995).

Peet, R. K. 1974. The measurement of species diversity. Annual Review of Ecology and Systematics 5: 285-307.

Peet, R. K. 1975. Relative diversity indices. Ecology 56: 496-498.

Peix, J. (ed.). 1999. Foc Verd II. Departament d'Agricultura, Ramaderia i Pesca de la Generalitat de Catalunya, Barcelona.

Petchey, O. L. 2000. Prey diversity, prey composition, and predator population dynamics in experimental microcosms. Journal of Animal Ecology 69: 874-882.

Peters, O., C. Hertlein, K. Christensen. 2002. A complexity view of rainfall. Physical Review Letters 88: 018701.

Peters, R. H. 1991. A critique for ecology. Cambridge University Press.

Pezzei, J. 1989. Economic Analysis of Sustainable Growth and Sustainable Development. World Bank, Environmental Department Working Paper 15.

Pfilester, A. B., and B. Schmid. 2002. Diversity-dependent production can decrease the stability of ecosystem functioning. Nature 416: 84-86.

Pickands III, J. 1975. Statistical inference using extreme order statistics. Annals of Statistics 3: 119-131.

Pielou, E. C. 1975. Ecological Diversity. John Wiley & Sons, New York.

Pielou, E. C. 1977. Mathematical Ecology. John Wiley & Sons, New York.

Pimentel, D., L. McLaughlin, A. Zepp, B. Lakitan, T. Kraus, P. Kleinman, F. Bancini, W. J. Roach, E. Graap, W. S. Keeton, and G. Selig. 1991. Environmental and economic impacts of reducing U. S. agricultural pesticide use. Pp. 676-719 in: D. Pimentel (ed.). Handbook of Pest Management in Agriculture, vol. I, 2nd ed. CRC Press, Boca Raton.

Pimentel, D., H. Herdendorf, S. Eisenfeld, L. Olander, M. Carroquino, C. Corson, J. McDade, Y. Chung, W. Cannon, J. Roberts, L. Bluman, and J. Gregg. 1994. Achieving a secure energy future: environmental and economic issues. Ecological Economics 9: 201-219.

Pimentel, D., C. Harvey, P. Resosudarmo, K. Sinclair, D. Kurz, M. McNair, S. Crist, L. Shpritz, L. Fitton, R. Saffouri, and R. Blair. 1995a. Environmental and economic costs of soil erosion and conservation benefits. Science 267: 1117-1123.

Pimentel, D., C. Harvey, P. Resosudarmo, K. Sinclair, D. Kurz, M. McNair, S. Crist, L. Shpritz, L. Fitton, R. Saffouri, and R. Blair. 1995b. Resply. Science 269: 464-465.

Pimm, S. L. 1991. The balance of nature? University of Chicago Press, Chicago and London.

Pimm, S. L., and A. Redfearn. 1988. The variability of population densities. Nature 334: 613-614.

Pimm, S. L., G. J. Russell, J. L. Gittleman, and T. M. Brooks. 1995. The future of biodiversity. Science 269: 347-350.

Plotkin, J. B., M. D. Potts, D. W. Yu, S. Bunyavejchewin, R. Condit, R. Foster, S. Hubbell, J. LaFrankie, N. Manokaran, L. Hua-Seng, R. Sukumar, M. A. Nowak, and P. S. Ashton. 2000. Predicting species diversity in tropical forests. Proceedings of the National Academy of Sciences of the USA 97: 10850-10854.

Plotnick, R. E. 1993. Ecosystem organization and extinction dynamics. Palaios 8: 202-212.

Pomeroy, D., and B. Tengecho. 1986. Studies of birds in a semi-arid area of Kenya III - The use of "timed-species-counts" for studying regional avifaunas. Journal of Tropical Ecology 2: 231-247.

Popper, K. R. 1956. The arrow of time. Nature 177: 538-538.

Postel, S. L., G. C. Daly, and P. R. Ehrlich. 1996. Human appropriation of renewable freshwater. Science 271: 785-788.

Power, M. E., and L. S. Mills. 1995. The Keystone cops meet in Hilo. Trends in Ecology and Evolution 10: 182-184.

Prat, N. 1993. La gestión ecosistémica del agua. Ecología Política 6: 91-98.

Preston, F. W. 1948. The commoness, and rarity, of species. Ecology 29: 254-283.

Preston, F. W. 1962. The canonical distribution of commoness and rarity. Ecology 43: 185-215, 410-432.

Preston, F. W. 1969. Diversity and stability in the biological world. Pp. 1-12 in Woodwell, G. M., and H. H. Smith, eds. Diversity and Stability in Ecological Systems. Brookhaven Symposia in Biology, 22. Brookhaven National Laboratory, Upton, New York.

Pueyo, S. 1994. Alguns aspectes de la hidrologia. Pp. 83-124 in: J. Armengol, S. Pueyo, and M. A. Rodríguez-Arias. Contribucions a l'establiment de cabals de compensació per a la part valenciana de la conca del riu Xúquer. Technical report for Direcció General de Qualitat d'Aigües, Generalitat Valenciana, València.

Pueyo, S. 1997. The study of chaotic dynamics by means of very short time series. Physica D 106: 57-65.

Pueyo, S. 1998. Gérmenes sin fronteras. Integral (Barcelona) 220: 44-49.

Puntí, A. 1988. Energy accounting: some new proposals. Human Ecology 16: 79-86.

Pyne, S. J. 1997. Vestal Fire. University of Washington Press, Seattle.

Rahmstorf, S. 1995. Bifurcations of the Atlantic thermohaline circulation in response to changes in the hydrological cycle. Nature 378: 145-149.

Ramsden, J. J., and Gy. Kiss-Haypál. 2000. Company size distribution in different countries. Physica A 277: 220-227.

Ray, T. S. 1995. An evolutionary approach to synthetic biology: Zen and the art of creating life. Pp. 179-209 in: C. G. Langton (ed.). Artificial Life. Massachusetts Institute of Technology.

Reddingius, J., and P. J. den Boer. 1970. Simulation experiments illustrating stabilization of animal numbers by spreading of risk. Oecologia 5: 240-284.

Reichenbach, H. 1956. The direction of time. University of California Press, Berkeley.

Rhodes, C. J., and R. M. Anderson. 1996. Power laws governing epidemics in isolated populations. Nature 381: 600-602.

Ricotta, C., G. Avena, and M. Marchetti. 1999. The flaming sandpile: self-organized criticality and wildfires. Ecological Modelling 119: 73-77.

Ricotta, C., M. Arianoutsou, R. Díaz-Delgado, B. Duguy, F. Lloret, E. Maroudi, S. Mazzoleni, J. M. Moreno, S. Rambal, R. Vallejo, and A. Vázquez. 2001. Self-organized criticality of wildfires ecologically revisited. Ecological Modelling 141: 307-311.

Ritchie, M. E., and H. Olff. 1999. Spatial scaling laws yield a synthetic theory of biodiversity. Nature 400: 557-560.

Robertson, C. A. 1972 Analysis of forest fire data in California. Technical Report No. 11, Dep. Stat., University of California, Riverside.

Rohter, L. 2001 Bill in Brazil would allow more of jungle to be razed. The New York Times, Sept. 23.

Romme, W. H. 1982. Fire and landscape diversity in subalpine forests of Yellowstone National Park. Ecological Monographs 52: 199-221.

Romme, W. H., D. G. Despain. 1989a. Historical perspective on the Yellowstone fires of 1988. BioScience 39: 695-699.

Romme, W. H., D. G. Despain. 1989b. The Yellowstone fires. Scientific American 261: 37-46.

Ros, J. D. 1989. Els ecosistemes bentònics. Pp. 119-176 in: R. Folch (ed.), Història Natural dels Països Catalans vol. 14. Enciclopèdia Catalana, Barcelona.

Rosenzweig, M. L. 1999. Heeding the warning in biodiversity basic law. Science 284: 276-277.

Rowell, A., and P. F. Moore. 1999. Global review of forest fires. WWF/IUCN.

Sarkar, S. 1993. Una síntesis ecosocialista del problema de la sobrepoblación. Ecología Política 6: 143-152.

Scheffer, M, S. Carpenter, J. A. Foley, C. Folke, and B. Walker. 2001. Catastrophic shifts in ecosystems. Nature 413: 591-596.

Scheinkman, J. A., and M. Woodford. 1994. Self-organized criticality and economic fluctuations. American Economic Review 84: 417-421.

Schenk, K., B. Drossel, S. Clar, F. Schwabl. 2000. Finite-size effects in the self-organized critical forest-fire model. The European Physical Journal B 15: 177-185.

Schimel, D., and D. Baker. 2002. The wildfire factor. Nature 420: 29-30.

Schimmel, J., and A. Granström. 1997. Fuel succession and fire behavior in the Swedish boreal forest. Canadian Journal of Forest Research 27: 1207-1216.

Schläpfer, F., and B. Schmid. 1999. Ecosystem effects of biodiversity: a classification of hypotheses and exploration of empirical results. Ecological Applications 9: 893-912.

Schlick, M. 1948. Grundzüge der Naturphilosophie. Gerold & Co., Vienna.

Schlögl, F. 1989. Probability and Heat. Fundamentals of Thermostatistics. Vieweg, Braunschweig/Wiesbaden.

Schmidt, K. A., and R. S. Ostfeld. 2001. Biodiversity and the dilution effect in disease ecology. Ecology 82: 609-619.

Schneider, E. D. 1988. Thermodynamics, ecological succession and natural selection: a common thread. In: B. H. Weber, D. J. Depew, and J. D. Smith (eds.). Entropy, information and evolution. New perspectives on physical and biological evolution. The MIT Press.

Schneider, E. D., and J. J. Kay. 1994. Complexity and thermodynamics: towards a new ecology. Futures 24: 626-647.

Scholz, C. H. 1990. The Mechanics of Earthquakes and Faulting. Cambridge University Press.

Schrödinger, E. 1944. What is life? The physical aspect of the living cell. Cambridge University Press.

Schuster, H. G. 1988. Deterministic Chaos. An Introduction, 2nd rev. ed. VCH, Weinheim.

Sen, A. 1993. La vida y la muerte como indicadores económicos. Investigación y Ciencia, julio 1993: 6-13.

Shannon, C. E. 1948. The mathematical theory of communication. Bell System Technological Journal 27: 379.

Shannon, C. E., and W. Weaver. 1963. The mathematical theory of communication. University of Illinois Press, Urbana.

Shaw, M. W. 1994. Modeling stochastic processes in plant pathology. Annual Review of Phytopathology 32: 523-544.

Siegert, F., G. Rueker, A. Hinrichs, and A. A. Hoffmann. 2001. Increased damage from fires in logged forests during droughts caused by El Niño. Nature 414: 437-440.

Siemann, E., D. Tilman, and J. Haarstad. 1996. Insect species diversity, abundance and body size relationships. Nature 380: 704-706.

Siemann, E., D. Tilman, and J. Haarstad. 1999. Abundance, diversity and body size: patterns from a grassland arthropod community. Journal of Animal Ecology 68: 824-835.

Skolimowski, H. 1995. In defence of sustainable development. Environmental Values 4: 69-70.

Smith, E. P., and G. van Belle. 1984. Nonparametric estimation of species richness. Biometrics 40: 119-129.

Smith, T. M., and M. A. Huston. 1989. A theory of the spatial and temporal dynamics of plant communities. Vegetatio 83: 49-69.

Solé, R. V., and D. Alonso. 1998. Random walks, fractals and the origins of rainforest diversity. Advances in Complex Systems 1: 203-220.

Solé, R. V., and J. Bascompte. 1995. Measuring chaos from spatial information. Journal of Theoretical Biology 175: 139-147.

Solé, R. V., and S. C. Manrubia. 1995. Are rainforests self-organized in a critical state? Journal of Theoretical Biology 173: 31-40.

Solé, R. V., and S. C. Manrubia. 1996. Orden y caos en sistemas complejos. Edicions UPC, Barcelona.

Solé, R. V., and S. C. Manrubia. 1997. Criticality and unpredictability in macroevolution. Physical Review E 55: 4500-4507.

Solé, R. V., S. C. Manrubia, M. Benton, and P. Bak. 1997. Self-similarity of extinction statistics in the fossil record. Nature 388: 764-767.

Solé, R. V., S. C. Manrubia, M. Benton, S. Kauffman, and P. Bak. 1999. Criticality and scaling in evolutionary ecology. Trends in Ecology and Evolution 14: 159-160.

Solé, R. V., D. Alonso, and A. McKane. 2002. Self-organized instability in complex ecosystems. Philosophical Transactions of the Royal Society of London Series B - Biological Sciences 357: 617-618.

Solow, A. R, C. Costello, and A. Beet. 1999. On an early result on stability and complexity. American Naturalist 154: 587-588.

Stauffer, D., and A. Aharony. 1992. Introduction to Percolation Theory, 2nd ed. Taylor & Francis, London.

Stenseth, N. C., and J. Maynard Smith. 1984. Coevolution in ecosystems: Red Queen evolution or stasis? Evolution 38: 870-880.

Stiglitz, J. E. 2002. Globalization and its Discontents. W. W. Norton and Co., New York. [also in Catalan, ed. Empúries, Barcelona, and in Spanish, ed. Taurus, Barcelona].

Stocks, B. J. et al. 1998. Climate change and forest-fire potential in Russian and Canadian boreal forests. Climatic Change 38: 1-13.

Stone , R. 1998. Yellowstone rising again from the ashes of devastating fires. Science 280: 1527-1528.

Strauss, D., L. Bednar, and R. Mees. 1989. Do one percent of forest fires cause ninety-nine percent of the damage? Forest Science 35: 319-328.

Sugihara, G. 1980. Minimal community structure: an explanation of species abundance patterns. American Naturalist 116: 770-787.

Sugihara, G., and R. May. 1990. Nonlinear forecasting as a way of distinguishing chaos from measurement error in time series. Nature 344: 734-341.

Sutherland, W. J. 2000. The Conservation Handbook. Research, Management and Policy. Blackwell Science, Cambridge

Svirezhev, Y. M., and D. O. Logofet. 1983. Stability of Biological Communities. Mir, Moscow.

Szilard, L. 1929. Über die Entropieverminderung in einem thermodynamischen System bei Eingriffen intelligenter Wesen. Zeitschrift für Physik 53: 840-856.

Takens, F. 1981. Detecting strange attractors in turbulence. Pp. 366-381 in: D A. Rand and L. S. Young (eds.), Proceedings of the Symposium on Dynamical Systems and Turbulence, University of Warwik 1979-1980. Springer, Berlin.

Taper, M. L., and P. A. Marquet. 1996. How do species really divide resources. American Naturalist 14l: 1072-1086

Taylor, P. J. 2000. Assessing biodiversity and ecological stability. Science 290: 51-51.

Terradas, J. 1996. Mesures específiques de gestió per a la prevenció d'incendis. Pp. 209-225 in: J. Terradas (ed.), Ecologia del foc. Proa, Barcelona.

Tilman, D. 1990. Mechanisms of plant competition for nutrients: the elements of a predictive theory of competition. Pp. 117-141 in: J. B. Grace and D. Tilman (eds.). Perspectives on Plant Competition. Academic Press, San Diego.

Tilman, D. 1996. Biodiversity: population versus ecosystem stability. Ecology 77: 350-363.

Tilman, D. 1999. The ecological consequences of changes in biodiversity: a search for general principles. Ecology. 80: 1455-1474.

Tilman, D., and J. A. Downing. 1996. Biodiversity and stability in grasslands. Nature 367: 363-365.

Tilman, D., and D. Wedin. 1991. Oscillations and chaos in the dynamics of a perennial grass. Nature 353: 653-655.

Tilman, D., C. L. Lehman, and K. T. Thomson. 1997. Plant diversity and ecosystem productivity: theoretical considerations. Proceedings of the National Academy of Sciences of the U.S.A. 94: 1857-1861.

Tilman, D., C. L. Lehman, and C. E. Bristow. 1998. Diversity-stability relationships: statistical inevitability or ecological consequence? American Naturalist 151: 277-282.

Tokeshi, M. 1993. Species abundance patterns and community structure. Advances in Ecological Research 24: 111-186.

Tokeshi, M. 1996. Power fraction - a new explanation of relative abundance patterns in species-rich assemblages. Oikos 75: 543-550.

Trimble, S. W., and P. Crosson. 2000a. U. S. soil erosion rates - myth and reality. Science 289: 248-250.

Trimble, S. W., and P. Crosson. 2000b. Response. Science 290: 1301-1301.

Turcotte, D. L. 1999. Self-organized criticality. Reports on Progress in Physics 62: 1377-1429.

Uhl, C. 1998. Perspectives on wildfire in the humid tropics. Conservation Biology 12: 942-943.

Ulanowicz, R. E. 1986. Growth and Development: Ecosystems Phenomenology. Springer-Verlag, New York.

Ulanowicz, R. E. 1997. Ecology, the Ascendent Perspective. Columbia University Press.

Ulanowicz, R. E. 1999. Life after Newton: an ecological metaphysic. BioSystems 50: 127-142.

UN (Office for the Coordination of Humanitarian Affairs). 1998. Forest fires on the Island of Sakhalin and the Khabarovsk Krai - UNDAC Mission report.

UNEP. 1995. Global Biodiversity Assessment. Cambridge: UNEP & Cambridge University Press.

van der Heijden, M. G. A., J. N. Klironomos, M. Ursic, P. Moutoglis, R. Streitwolf-Engel, T. Boller, A. Wiemken, and I. R. Sanders. 1998. Mycorrhizal fungal diversity determines plant biodiversity, ecosystem variability and productivity. Nature 396: 69-72.

van Leeuwan, J. (ed.). 1990. Handbook of theoretical computer science. Vol. A: Algorithms and complexity. MIT Press Elsevier.

Van Valen, L. 1973. A new evolutionary law. Evolution Theory 1: 1-30.

Van Wagner, C. E. 1987. Development and structure of the Canadian Forest Fire Weather Index System. Canadian Forestry Service Technical Report 35.

Vanotte, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, and C. E. Cushing. 1980. The River Continuum Concept. Canadian Journal of Fisheries and Aquatic Sciences 37: 130-137.

Vespigniani, A., and S. Zapperi. 1998. How self-organized criticality works: A unified mean-field picture. Physical Review E 57: 6345-6362.

Viegas, D. X. 1998. Forest fire propagation. Philosophical Transactions of the Royal Society of London A, Mathematical, Physical and Engineering Sciences 356: 2907-2928.

Vitousek, P. M., P. R. Ehrlich, A. H. Ehrlich, and P. A. Matson. 1986. Human appropriation of the products of photosynthesis. BioScience 36: 368-373.

Vitousek, P. M., H. A. Mooney, J. Lubchenco, and J. M. Melillo. 1997. Human domination of Earth's ecosystems. Science 277: 494-499.

Voss, R. F. 1988. Fractals in nature: from characterization to simulation. Pp. 21-70 in Peitgen, H. O., and D. Saupe, eds. The science of fractal images. Springer-Verlag, New York.

Wackernagel, M., and W. Rees. 1996. Our Ecological Footprint. New Society Publishers, Gabriola Island 1996.

Wackernagel, M., L. Onisto, P. Bello, A. Callejas Linares, I. S. López Falfán, J. Méndez García, A. I. Suárez Guerrero, M. G. Suárez Guerrero. 1999. National capital accounting with the ecological footprint concept. Ecological Economics 29: 375-390.

Wagensberg, J., D. López, and J. Valls. 1988. Statistical aspects of biological organization. Journal of Physics and Chemistry of Solids 49: 695-700.

Wagensberg, J., A. García, and R. V. Solé. 1991. Energy flow-networks and the maximum entropy formalism. Pp. 253-264 in: W. T. Grandy, Jr, and L. H. Schick (eds.), Maximum entropy and Bayesian methods. Kluwer Acad. Publ.

Wark Jr, K. 1995. Advanced thermodynamics for engineering. McGraw-Hill, New York.

Watkinson, A. R., R. P. Freckleton, R. A. Robinson, and W. J. Sutherland. 2000. Predictions of biodiversity response to genetically modified herbicide-tolerant crops. Science 289: 1554-1557.

Watson, R., and D. Pauly. 2001. Systematic distortions in world fisheries catch trends. Nature 414: 534-536.

Watt, K. E. F. 1964. Comments on long-term fluctuations of animal populations and measures of community stability. Canadian Entomologist 96: 1434-1442.

Wattenburg, B. 1998. The burning of Yellowstone- another perspective. Science 282: 1051-1051.

Watterson, G. A. 1974. Models for the logarithmic species abundance distributions. Theoretical Population Biology 6: 217-250.

Weber, B. H., D. J. Depew and J. D. Smith (eds.). 1988. Entropy, information and evolution. MIT Press, Cambridge.

Westoby, M. 1997. What does ecology mean? Trends in Ecology and Evolution 12: 166-166.

Wherl, A. 1978. General properties of entropy. Reviews of Modern Physics 50: 221-260.

Whitfield, J. 2002. Neutrality versus the niche. Nature 417: 480-481.

Wicherek, S. 1994. L´érosion des grandes plaines agricoles. La Recherche 25 (setembre): 880-888.

Wiener, N. 1948. Cybernetics, or control and communication in the animal and the machine. Massachusetts Institute of Technology.

Wiggins, S. 1990. Introduction to Applied Nonlinear Dynamical Systems and Chaos. Springer, New York.

Wigner, E. P. 1959. Proceedings of the Fourth Canadian Mathematical Congress, Toronto, 174.

Wilhelm, T., and R. Brüggemann. 2000. Goal functions for the development of natural systems. Ecological Modelling 132: 231-246.

Wilkes, G. 1992. Strategies for sustaining crop germplasm preservation, enhancement, and use. Issues in Agriculture 5. Consultative Group on International Agricultural Research, Washington.

Williams, B. 1943. Results obtained by means of a light-trap at Rothamsted. Pp. 44-53 in: Fisher, R. A., A. S. Corbet, and C. B. Williams, The relation between the number of species and the number of individuals in a random sample of an animal population. Journal of Animal Ecology 12: 42-58.

With, K. A., and A. W. King. 1997. The use and misuse of neutral landscape models in ecology. Oikos 79: 219-229.

With, K. A., S. J. Cadaret, and C. Davis. 1999. Movement responses to patch structure in experimental fractal landscapes. Ecology 80: 1340-1353.

Wolf, A., J. B. Swift, H. L. Swiney, and J. A. Vastano. 1985. Determining Lyapunov exponents from a time series. Physica D 16: 285-317.

Wolfe, M. S. 2000. Crop strength through diversity. Nature 406: 681-682.

Wood, S., K. Sebastian, and S. J. Scherr. 2000. Pilot Analysis of Global Ecosystems (PAGE): Agroecosystems. International Food Policy Research Institute and World Resources Institute.

World Comission on Environment and Development. 1987. Our Common Future. Oxford University Press, Oxford.

WRI. 1997. World Resources 1996-1997. WRI, Washington.

Wright, S. 1932. The roles of mutation, inbreeding, crossbreeding and selection in evolution. Proc. XI Internat. Congr. Genetics 1: 356-366. Quoted in: D. J. Futuyma, Evolutionary biology, 2nd ed., Sineauer Ass. Inc., Sunderland, Massachusetts, 19.

WWF. 2002. Living Planet Report 2002. World Wide Fund for Nature, Gland.

Yachi, S., and M. Loreau. 1999. Biodiversity and ecosystem productivity in a fluctuating environment: the insurance hypothesis. Proceedings of the National Academy of Sciences of the USA. 96: 1463-1468.

Young, S. 2001. Global energy prospects. Nature 414: 487-488.

Zhu, Y., H. Chen, J. Fan, Y. Wang, Y. Li, J. Chen, J. X. Fan, S. Yang, L. Hu, H. Leung, T. W. Mew, P. S. Teng, Z. Wang, and C. C. Mundt. Genetic diversity and disease control in rice. Nature 406: 718-722.

Zurek, W. H. 1989a. Algorithmic randomness and physical entropy. Physical Review A 40: 4731-4751.

Zurek, W. H. 1989b. Thermodynamic cost of computation, algorithmic complexity and the information metric. Nature 341: 119-124.