UNIVERSITAT DE BARCELONA

Climate change and precipitation trends in the northern Mediterranean

Didac Fortuny

Climate change and precipitation trends in the northern Mediterranean

Didac Fortuny

Supervised by

Dr. Ileana Bladé

Departament d'Astronomia i Meteorologia

Universitat de Barcelona

Programa de doctorat en Física

A thesis submitted for the degree of

Doctor of Philosophy

Barcelona, summer 2015

Contents

Resum de la tesi

Segons les projeccions dels principals models climàtics utilitzats en l'elaboració del cinquè informe del Grup Intergovernamental d'Experts sobre el Canvi Climàtic, la regió mediterrània nord patirà un fort descens de precipitació al llarg del segle XXI. Però, com de realistes són aquestes projeccions? Fins a quin punt són robustes?

Per tal d'avaluar la versemblança d'aquestes projeccions, molts autors han estudiat el comportament de la precipitació mediterrània en les últimes dècades, buscant-hi indicis de que l'assecament projectat ja és detectable. No obstant, degut a que la precipitació al Mediterrani nord presenta una forta variabilitat interna, no és senzill discriminar els canvis de precipitació naturals d'aquells associats a un senyal persistent d'origen forçat. Per aquest motiu, en la primera part de la tesi proposem un mètode per detectar quines tendències de precipitació observades són fàcilment atribuïbles a la variabilitat natural i quines poden ser considerades respostes a forçaments externs. Aquest mètode consisteix en estudiar com varia la magnitud d'aquestes tendències en aplicar lleugers canvis en el període pel qual estan estimades. Degut a que la variabilitat natural és pràcticament aleatòria, un petit canvi en el període de càlcul pot comportar passar d'una tendència forta i estadísticament significativa a una tendència feble i no significativa. En canvi, en presència d'un senyal persistent (resposta a forçaments externs), les tendències mostren una magnitud similar en períodes de càlcul consecutius. Fent ús d'aquest tipus d'anàlisi, els nostres resultats indiquen que l'evolució de la precipitació al Mediterrani nord és compatible amb la presència d'un senyal d'assecament forçat només a l'hivern i des de mitjan segle XX ençà. Per la resta d'estacions, en canvi, les tendències de precipitació observades són compatibles amb la variabilitat natural. Paral·lelament, també hem estudiat quin paper ha tingut la Oscil·lació de l'Atlàntic Nord (NAO), que és el principal mode de variabilitat de la precipitació mediterrània, en les tendències de precipitació observades. Els nostres resultats indiquen que els canvis de precipitació observats a l'hivern durant les últimes dècades són en gran mesura explicables per una tendència positiva de l'índex NAO durant el mateix període de temps.

L'objectiu de la segona part de la tesi ha estat avaluar si, utilitzant simulacions climàtiques, es pot detectar un senyal d'assecament al Mediterrani nord i, en cas afirmatiu, si aquest senyal és atribuïble a forçaments externs. Un senyal climàtic es considera detectable si no és reproduïble amb simulacions climàtiques que no incloguin cap tipus de forçament extern (anomenades simulacions preindustrials) i es considera atribuïble a forçaments externs si, a més a més, es pot reproduir amb simulacions climàtiques que sí que incorporin aquests forçaments (anomenades simulacions històriques). Desafortunadament, en el cas de la precipitació d'hivern mitjanada sobre el Mediterrani nord, hem trobat que ni les simulacions preindustrials ni les històriques inclouen tendències tan intenses com les observades en les últimes dècades. Per explicar aquesta inconsistència, hem plantejat dues hipòtesis: (1) els canvis de precipitació observats a l'hivern estan principalment associats a variabilitat natural, però aquesta està subestimada en els models climàtics; i (2) els canvis observats tenen un origen forçat, però el forçament (o la resposta al forçament) no està ben reproduït en les simulacions històriques. No obstant, segons els nostres resultats, aquesta conclusió només és vàlida per intervals que mostren una forta tendència en la NAO i en les regions en les quals la resposta de la NAO és més gran. Per tant, hem analitzat si la inconsistència en les tendències de precipitació entre models i observacions són degudes a una mala representació per part dels models de la variabilitat natural o del senyal forçat en la NAO. Degut a que tendències observades de la NAO en altres moments del segle XXI, clarament no atribuïbles a forçaments (degut a que són de signe contrari a la tendència recent i no han persistit), tampoc no són compatibles amb les tendències observades, hem conclòs que és plausible que la variabilitat de la NAO estigui subestimada en els models. Aquest resultat reforça la hipòtesi que les tendències observades de precipitació d'hivern a la regió mediterrània són degudes a variabilitat natural i que les inconsistències amb els models són degudes a que aquesta variabilitat està subestimada en les simulacions climàtiques.

Finalment, en la tercera part d'aquesta tesi hem avaluat fins a quin punt són robustes les projeccions de canvis de precipitació al Mediterrani nord durant el segle XXI. Si bé és cert que la majoria de models coincideixen en el signe del canvi de precipitació projectat (tots ells mostren assecament), fins ara no s'havia examinat el grau d'acord entre models en referència a altres aspectes del canvi de precipitació com la seva magnitud total, la seva distribució estacional i geogràfica o com de gradual serà la transició cap a les condicions més seques de finals del segle XXI. No obstant, avaluar el consens entre models en referència a aquests aspectes de l'assecament és crucial si volem descriure quin serà el comportament de la precipitació en les properes dècades. Utilitzant en cada cas la manera més adequada de comparar els resultats dels diferents models, hem trobat que les úniques característiques de les projeccions que presenten un bon grau d'acord entre models són el signe negatiu del canvi de precipitació i la seva distribució estacional, en que l'assecament a l'estiu és més intens que a l'hivern (a diferència del què passa en les observacions). En canvi, aspectes en

els quals els models difereixen considerablement són la magnitud de la tendència de precipitació futura (que a l'estiu va des de -10% per segle fins a -80% per segle), l'amplitud de la variabilitat multi-decennal superposada al senyal d'assecament, i l'any aproximat en que aquest senyal forçat es farà evident per sobre de la variabilitat interna. A més a més, també hem determinat un alt grau de desacord entre models en quant a com de gradual s'espera que sigui l'assecament. Mentre que en alguns models l'assecament es produeix de manera brusca en algun moment del segle XXI, en altres models l'assecament es produeix de manera continuada al llarg de tot el segle.

Acknowledgements

I acknowledge the providers of all observational data sets and the developers of all climate models used in this thesis for sharing their data. I thank my supervisor, Dr. Ileana Bladé, for proposing the topic of this thesis and for all her comments and corrections. The obtaining and arrangement of CMIP5 data has been carried out with help from Pedro Gámez. I also acknowledge Dr. Yolanda Castro-Díez, Dr. Carmen Llasat and Dr. Belén Rodríguez-Fonseca for being part of my thesis committee and for their useful remarks and suggestions. This work has been partially supported by the CONSOLIDER project *Supercomputación y eCiencia* (Ref. CSD2007-0050) from the Spanish Ministerio de Economía y Competitividad.

Moltes gràcies als meus companys de departament, amb qui sempre he pogut comentar resultats i figures, discutir articles, contrastar idees i agafar-ne de noves. Gràcies també a tots els que m'heu fet de correctors voluntaris mirant-vos abstracts, pòsters o capítols d'aquesta tesi. Gràcies per haver-me fet de grup de recerca.

En especial li dono les gràcies als meus companys de despatx. Yolanda Sola, després de *maDAMlenes* i *calippos*, un parell de tisores trencades o desaparegudes, 126 *pritis* i uns quants centenars de *dunkin coffees*, estic convençut que la vida al DAM ha estat més agradable gràcies a tu. Et trobaré a faltar, i tu trobaràs a faltar la meva calculadora. També vull agrair la seva ajuda al Vicent Altava (atsí!), primer referent, mestre de l'IDL i millor pastisser. Gràcies també a l'Albert Ossó, *llum de la meva vida*, per acompanyar-me durant quatre anys tot el dia i tota la nit (a partir de les 12 del migdia, és clar).

Muchas mercedes también a Don Pedro Gámez por haber sido un buen compadre de obcecaciones. Y, por encima de todo, muchas mercedes por tu favor a la hora de procurarme los datos de los modelos y por haber compartido tus escrips y haber estado siempre dispuesto a enmendarlos cuando no procedían acertadamente. Gracias, *scriptator*!

Gràcies a les mireies, que m'han acompanyat des del màster (menys quan han marxat a Boulder, és clar). Mireia Udina, gràcies per les estones de desconnexió al servei d'esports, en especial per deixar-te guanyar a tennis i fer-me fingir una lesió a la cama per evitar

tornar a perdre contra tu. Mireia Mateu, gràcies per descobrir-me les galetes holandeses i donar-me una excusa per pujar a prendre l'aire al terrat.

Gràcies al Club de los Viernes i, sobretot, als seus membres titulars i amics invisibles. Miriam Olid, patinadora, pastissera i pingüina d'honor. *Honor!* Gràcies per donar el teu vistiplau a figures i pòsters i viceversa. Josep Manel Carrasco, segur que no t'agrada que et donin les gràcies, així que *für den Schmerz!* En tot cas, m'ha agradat poder mantenir viu d'alguna manera l'astrònom amateur que encara porto dins.

M'agradaria també donar les gràcies a l'Adolfo Magaldi per la seva predisposició a ajudar en qualsevol cosa i pel seu afecte. En particular, gràcies per haver-me guiat al principi, ajudant-me a redactar textos en anglès i introduint-me al meravellós món del Linux quan jo encara n'era un principiant. 42 abraçades ben fortes, noi! Gràcies també a l'Aitor Atencia que encara ara, des de 5.896 km de distància (si no és que ha tornat a canviar d'adreça sense avisar, és clar), està sempre disposat a resoldre'm dubtes i donar-me consells. Que el blending t'acompanyi, Aitor.

Vull agrair també la seva ajuda i la seva companyia al Toni Barrera, la Mercè Barnolas, l'Albert Saragossa, el Marco Turco, el Raül Marcos, l'Àlex Montornés i el Bernat Jiménez. Gràcies a tots! I també a aquells que han vingut de fora, la Sarah Lindberg, el Johan Liakka, la Leticia Palazzi, el Diego Fustes, l'Adrian Matei, la Silvia Calò i el Richard Caceres. Gràcies també a la gent de secretaria, el J.R. Rodríguez, la Montse Solé i la Rosa Iborra. Que "fàcils" es tornen els tràmits administratius quan hi ha persones competents al darrere.

També vull recordar la gent que m'he trobat arreu del món, la Irene Polo, la Nube González, la Cécile Agosta, el Navid Constantinou, l'Mxolisi Shongwe, l'Andreas Safron, la Ieva Nariūnaitė, la Joanna Jędruszkiewicz, la Ildikó Pieczka, l'Andreas Hoy, la Hadas Reiser, la Jelena Lukovic, el Muhammad Hassim, el Karim Houchi, el Samuele del Bianco, el Daniel Hernandez-Deckers, el Christian Sippel, el Pedro Sousa, el Colin Kelley, el Luis Rodriguez, el Pere Esteban, la Irena Balog i el Gernot Vogt, entre d'altres. Gracias en especial a Javi García por su ayuda y sus consejos.

He de donar les gràcies també als amics que m'han ajudat amb la tesi des de fora del departament. Sergi Gonzàlez, company meteoròleg (o diplomat per ser precisos), gràcies per ajudar-me de bon grat sempre que t'ho he demanat. Gràcies Daniel Roig per la teva empenta i la teva generositat. Gràcies David Diego per les estones de treball al ⟨café|. Josep Mas, gràcies pel teu bon esperit i per caminar amb mi en la travessa del desert (ànims, *que te'n* sortiràs tu també!). Albert Benseny, gràcies pel teu assessorament en Latex i gnuplot, per les teves correccions (fins que sortia el sol), pel teu exemple, pels teus ànims, pel teu afecte.

Per anar acabant, gràcies als amics que sempre m'han donat el seu suport i que m'han animat en temps difícils. Gràcies al Pau, l'Ana, la Joana, la Sílvia, la Raquel, la Inés,

l'Andreu, el Lluís, la Marta, la Blanca i l'Elisa. Gràcies també al Cor Iolis per les bones estones que m'ha fet passar a la facultat.

Als meus pares, Francesc i Teresa, vull agrair-los haver-me re-acollit en l'últim any de la tesi. Gràcies per preocupar-vos per mi, per animar-me, per fer-me saber que podia comptar amb vosaltres (tot i que ja ho sabia), per la vostra paciència en moments complicats.

Per acabar, gràcies Marina. Gràcies per distreure'm i per fer-me treballar en proporcions òptimes. Gràcies per fer-me d'hoste. Gràcies per haver-me donat altes dosis d'*atopetil* sense recepta, per haver-me fet de valeriana i de salbutamol. Això no hauria estat possible sense tu.

Abstract

According to climate projections for the end of the twenty-first century, regional responses to global warming include a strong and widespread drying over the northern Mediterranean, particularly intense in summer. Many authors have reported observed negative precipitation trends in the last decades and have considered them evidences that the projected drying is already detectable. Because of the strong decadal variability of precipitation series, however, it is not straightforward to discern whether observed trends are attributable to an external signal or whether they are consistent with internal variability alone. In order to assess if precipitation series in the northern Mediterranean region are compatible with the presence of a persistent signal, we examine the sensitivity of trend estimations to the choice of time interval: since internal variability is expected to be random, the magnitude of natural precipitation changes varies greatly when estimated for intervals with different arbitrary initial and final years. Instead, series under the influence of a persistent signal exhibit trends of similar magnitude in a wide range of consecutive intervals. In this regard, our results indicate that only in winter and from mid-twentieth century onwards the evolution of precipitation series is consistent with what one would expect if a persistent signal was superimposed on internal variability. Trends for other seasons and periods, instead, are sensitive to the choice of time interval. Moreover, we find that the recent observed winter drying trends are largely explainable by an upward trend in the North Atlantic Oscillation (NAO) observed in the last decades of the twentieth century.

The role played by climate change and internal variability in the observed winter northern Mediterranean drying is examined using historical (forced) and pre-industrial (unforced) CMIP5 climate simulations. For area-averaged precipitation we find that neither historical nor pre-industrial simulations contain trends as strong as the ones observed in the recent decades. We consider two explanations for this result: (1) observed trends have a forced origin but the forcing (or the response to the forcing) is underestimated in historical simulations, and (2) observed trends are mainly natural but climate models underestimate the range of internal variability of northern Mediterranean precipitation. We find that this

result is consistent with the fact that climate simulations do not produce NAO trends as strong as those observed in the last decades.

Regarding future projections of drying in the northern Mediterranean region, many studies focus on the multi-model ensemble mean without always paying attention to results for individual climate models. Evaluating inter-model consensus, however, is crucial to assess whether projections are robust or not. We investigate the extent to which individual CMIP5 projections agree with each other about future precipitation changes over the northern Mediterranean in terms of the seasonality of the changes, their magnitude and geographical distribution, the transition towards the driest conditions and the time of emergence of the forced signal. For each analysis we find the most appropriate way to compare results for single models and illustrate their similarities and discrepancies. Aspects of the projections with good inter-model agreement are the sign of the projected northern Mediterranean area-averaged drying and its seasonal distribution: models agree that the strongest drying will occur in summer (which is opposite to the seasonal distribution of observed recent precipitation trends). Instead, aspects with weak inter-model agreement are the magnitude of the trend (ranging from about -10%/century to about -80%/century in summer), the amplitude of the decadal to multi-decadal variability and the time of emergence of the forced signal, with a inter-model spread for summer covering the entire twenty-first century.

Chapter

ONE

Introduction

1.1 Global climate change

According to the fifth assessment report (AR5[1]) of the Intergovernmental Panel on Climate Change (IPCC), since the industrial revolution the surface of the Earth has never been as warm as in the recent decade (Fig. 1.1a). The global mean temperature increase between 1850-1900 and 2003-2012 is estimated as 0.78 °C [IPCC, 2013], even though the magnitude of the warming varies from region to region (Fig. 1.1b). Our confidence in the global warming signal is high for several reasons. First, the general evolution of global temperature in the last century largely coincides among independent data sets. Second, the physical mechanism behind the atmospheric temperature increase is well understood.

The main driver of the global warming is the increase in the radiative forcing observed from the industrial revolution onwards. Figure 1.2, extracted from the IPCC AR5, illustrates the role of different agents in the total effective radiative forcing [Myhre et al., 2013]. The general evolution of the radiative forcing follows the increase of greenhouse-gas (GHG) concentrations, partially compensated by the negative effect of aerosols. GHG absorb part of the long-wave radiation emitted by the Earth, warming the atmosphere. As the atmosphere warms, it emits more long-wave radiation, increasing the total incident upon the surface. Regarding natural forcings, while volcanoes cause a reduction in the radiative forcing in the years after the eruption due to the large amount of sulphate aerosols ejected into the stratosphere which absorb and reflect short-wave radiation from the Sun, the influence of the 11-year solar cycle is very weak. We can affirm, therefore, that global warming is mainly a consequence of the anthropogenic increase in the concentration of GHG gases in the atmosphere. Since the emission of GHG keeps growing, the response of the climate in the future is expected to become larger as well. In particular, climate models

[1] All acronyms and abbreviations are listed at the end of the thesis.

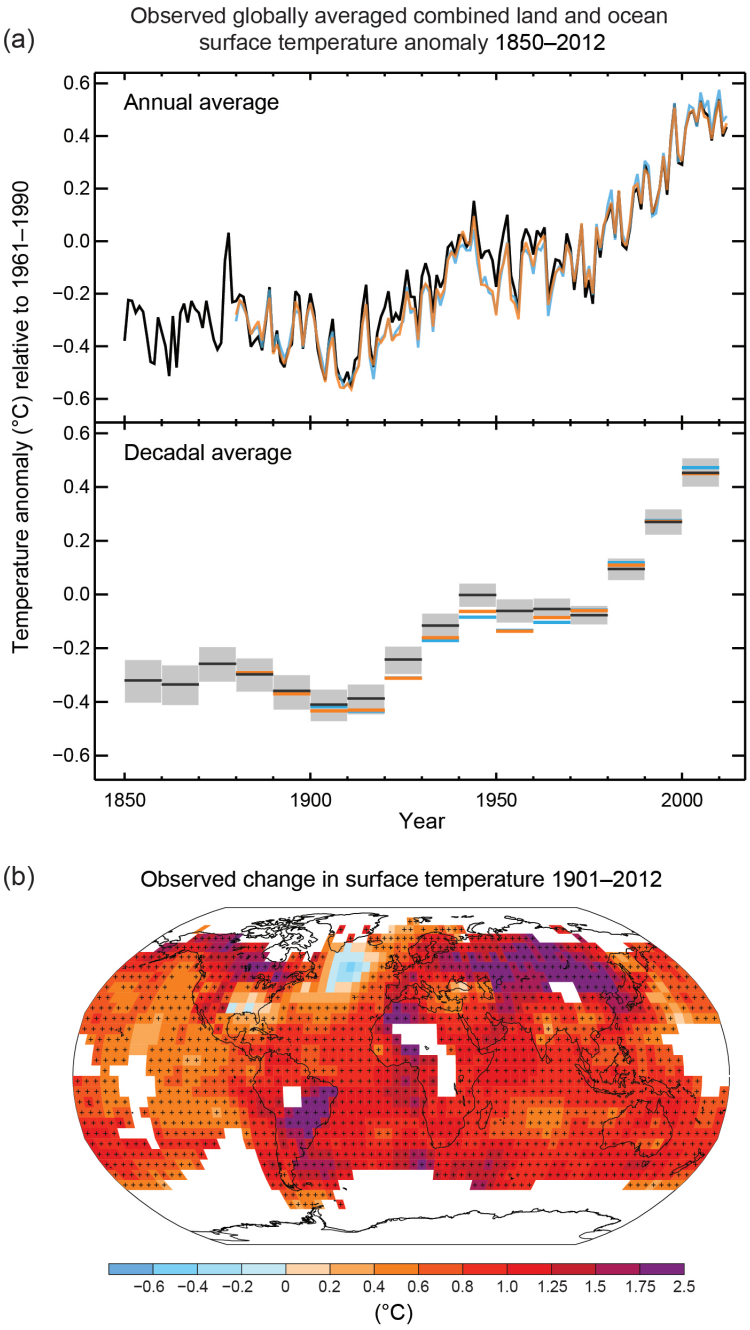


Figure 1.1: *(a) Observed global mean combined land and ocean surface temperature anomalies, from 1850 to 2012 from three data sets. Top panel: annual mean values. Bottom panel: decadal mean values including the estimate of uncertainty for one dataset (black). Anomalies are relative to the mean of 1961-1990. (b) Map of the observed surface temperature change from 1901 to 2012 derived from temperature trends determined by linear regression from one dataset (orange line in panel a). Trends have been calculated where data availability permits a robust estimate (i.e., only for grid boxes with greater than 70% complete records and more than 20% data availability in the first and last 10% of the time period). Other areas are white. Grid boxes where the trend is significant at the 10% level are indicated by a + sign. For a listing of the datasets and further technical details see the Technical Summary Supplementary Material.* {*Figures 2.19–2.21; Figure TS.2*}. Figure extracted from IPCC (2013) (their Fig. SPM.1).

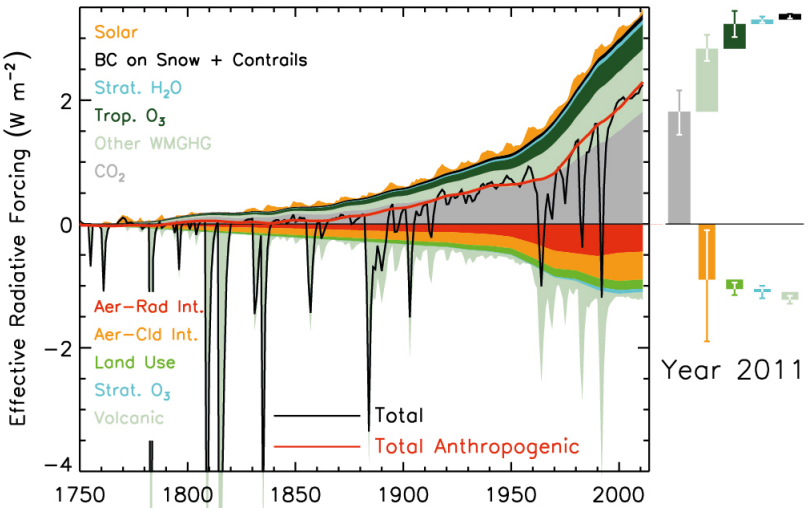


Figure 1.2: *Time evolution of forcing for anthropogenic and natural forcing mechanisms. Bars with the forcing and uncertainty ranges (5 to 95% confidence range) at present are given in the right part of the figure. For aerosol the ERF due to aerosol–radiation interaction and total aerosol ERF are shown. The uncertainty ranges are for present (2011 versus 1750) and are given in Table 8.6. For aerosols, only the uncertainty in the total aerosol ERF is given. For several of the forcing agents the relative uncertainty may be larger for certain time periods compared to present. See Supplementary Material Table 8.SM.8 for further information on the forcing time evolutions. Forcing numbers provided in Annex II. The total anthropogenic forcing was 0.57 (0.29 to 0.85) W m^{-2} in 1950, 1.25 (0.64 to 1.86) W m^{-2} in 1980 and 2.29 (1.13 to 3.33) W m^{-2} in 2011.* Figure extracted from Myhre et al. (2013) (their Fig. 8.18).

project for the 2081-2100 interval a global surface warming ranging from 1.0 °C to 3.7 °C (with respect of the 1986-2005 mean), depending on the future scenario (Fig. 1.3).

Contrary to the uniform sign of observed regional temperature changes, the water cycle responses to global warming are non homogeneous neither in sign nor in magnitude, and they may be masked by decadal and multi-decadal variability. According to the IPCC AR5 [Hartmann et al., 2013] globally averaged precipitation over land exhibits a weak positive trend within the 1901-2008 interval (Fig. 1.4, bottom panel). Confidence in this result, however, is low because of the discrepancy among different data sets regarding the magnitude of the trend and its statistical significance, together with the fact that precipitation estimates for the first half of the twentieth century have large uncertainty associated. Figure 1.5 represents the global distribution of observed annual mean precipitation trends (1901-2010 and 1951-2010) according to different data sets. In general, all maps indicate that the strongest precipitation trends are positive and found in mid-latitudes. An exception to the global tendency is, as we shall see later, the Mediterranean region. Observed precipitation trends over the Mediterranean region are non-significant according to all data sets, and their sign is not uniform for 1901-2010 although it is mainly negative for the recent period 1951-2010.

In contrast with the weak observed trends, global precipitation is projected to increase

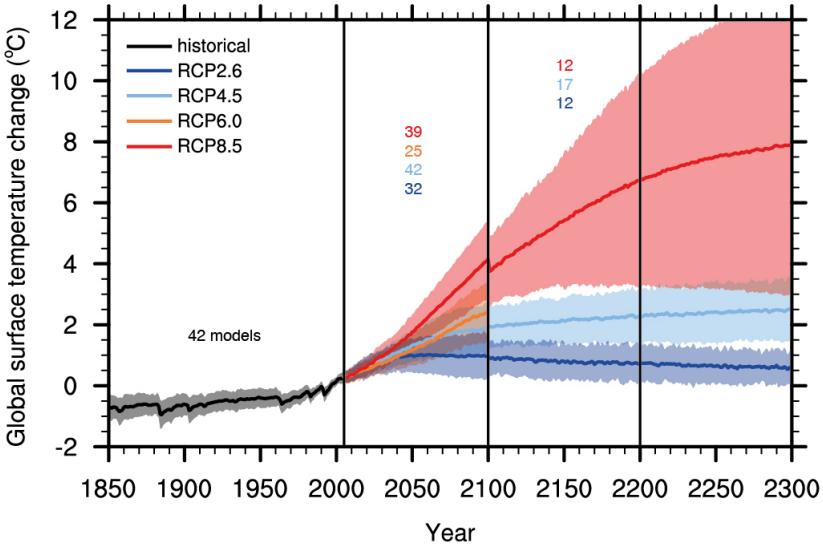


Figure 1.3: *Time series of global annual mean surface air temperature anomalies (relative to 1986–2005) from CMIP5 concentration-driven experiments. Projections are shown for each RCP for the multi-model mean (solid lines) and the 5 to 95% range (±1.64 standard deviation) across the distribution of individual models (shading). Discontinuities at 2100 are due to different numbers of models performing the extension runs beyond the 21st century and have no physical meaning. Only one ensemble member is used from each model and numbers in the figure indicate the number of different models contributing to the different time periods. No ranges are given for the RCP6.0 projections beyond 2100 as only two models are available.* Figure extracted from Collins et al. (2013) (their Fig. 12.5).

during the twenty-first century almost linearly with global temperature in all scenarios, at a rate ranging from 1%/°C to 3%/°C, according to the models of the phase 5 of the Coupled Model Intercomparison Project (CMIP5) used in the IPCC AR5 [Collins et al., 2013]. Still, the sign and the magnitude of precipitation projections exhibit large regional and seasonal variability (Fig. 1.6), generally following the "dry gets drier, wet gets wetter" mechanism derived from the Clausius–Clapeyron equation. This mechanism states that, as a result of the increase in the saturated vapour pressure of the atmosphere associated with the warming, the moisture transport from the subtropics to subpolar and equatorial regions increases, intensifying the difference between precipitation and evaporation ($P-E$) [Held and Soden, 2006]. As a result, current wet regions ($P-E>0$) become wetter and current dry regions ($P-E<0$) become drier. Nevertheless, according to Scheff and Frierson (2012), the largest future drying is not expected over the subtropical dry regions but poleward of the current subtropical precipitation minima. They associate this shift in the distribution of mid-latitudinal dry regions to dynamical mechanisms such as the poleward expansion of the Hadley cell [Lu et al., 2007]. In particular, the Mediterranean region has been identified as one of the world's regions in which the responses to climate change on regional climate are most likely to be most severe Diffenbaugh and Giorgi (2012).

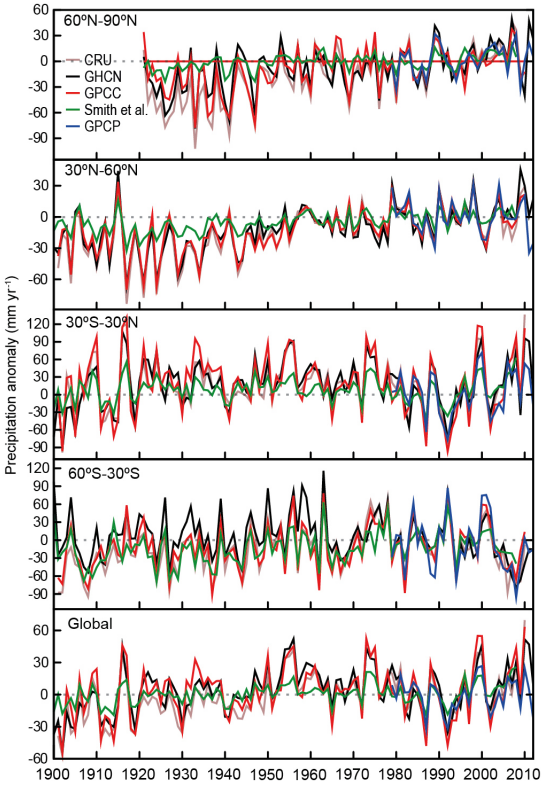


Figure 1.4: *Annual precipitation anomalies averaged over land areas for four latitudinal bands and the globe from five global precipitation data sets relative to a 1981–2000 climatology.* Figure extracted from Hartmann et al. (2013) (their Fig. 2.28).

1.2 Climatology of the northern Mediterranean region

According to the Köppen classification, the Mediterranean climate is characterized by a seasonal behaviour with wet and mild winters and dry and warm summers [Köppen, 1936]. The observed annual cycle of land precipitation averaged over the northern Mediterranean region (from now on *NMed* region, 10°W-40°E/36°N-45°N[2]), estimated with E-OBS data for the 1950-2014 interval, is consistent with Köppen's description (Fig. 1.7[3]). The driest months are July and August, with less than half the precipitation of the wettest months, November and December. The transition from wet to dry is more gradual in spring than in autumn.

Apart from seasonal variability, precipitation in the northern Mediterranean region also exhibits large spatial variations, as illustrated by the annual climatology map computed by averaging the entire record of E-OBS precipitation (1950-2014) at each grid-point (Fig.

[2]All regions for which we average data are indicated at the end of the thesis.

[3]Note that we plot results placing December at the beginning of the horizontal axis and not at its end. We made this choice for consistency with future plots in which December shares more seasonal characteristics in terms of precipitation changes with January, February and March than with November. For this reason we place December in the first place in all plots that show results for individual months.

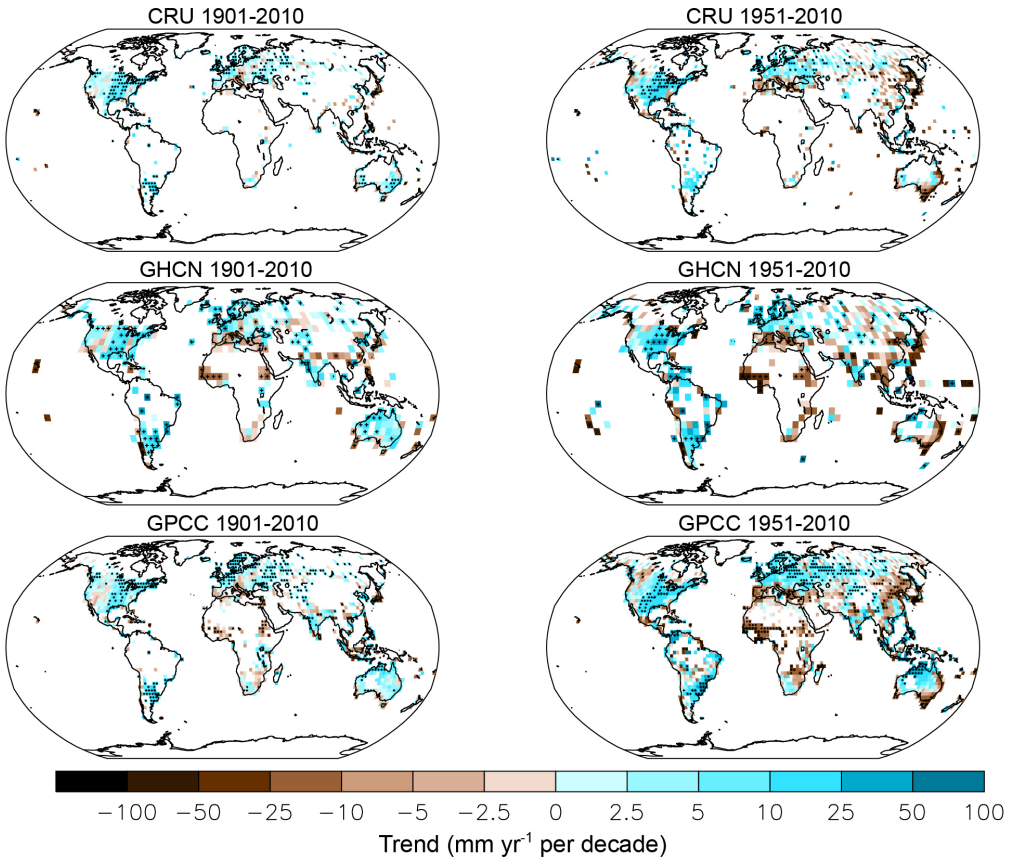


Figure 1.5: *Trends in annual precipitation over land from the CRU, GHCN and GPCC data sets for 1901–2010 (left-hand panels) and 1951–2010 (right-hand panels). Trends have been calculated only for those grid boxes with greater than 70% complete records and more than 20% data availability in first and last decile of the period. White areas indicate incomplete or missing data. Black plus signs (+) indicate grid boxes where trends are significant (i.e., a trend of zero lies outside the 90% confidence interval).* Figure extracted from Hartmann et al. (2013) (their Fig. 2.29).

1.8a). Because most of the annual precipitation takes place in winter, the annual climatological precipitation pattern is generally very similar to the one for winter (Fig. 1.8c).

Winter precipitation in the northern Mediterranean is mainly caused by storms originating as low pressure systems from the north Atlantic that reach the Iberian peninsula and then zonally cross the Mediterranean sea. For this reason, western coasts exhibit, in general, higher climatological winter precipitation than eastern coasts in all Mediterranean peninsulas [e.g. Bolle, 2003]. The effect of the topography is also noticeable, with climatological precipitation being higher over mountain regions, particularly the Pyrenees and the Alps [e.g. Lionello et al., 2006]. Thus, the regions of the northern Mediterranean exhibiting largest climatological winter precipitation are north-western Iberia (where the maximum values in the entire basin are found), the French Riviera (in particular the Pyrenees), the north of Italy (in particular the Alps), and the western half of the Balkan Peninsula. The lowest values are found in south-eastern Iberia, the eastern Balkans and the south of the

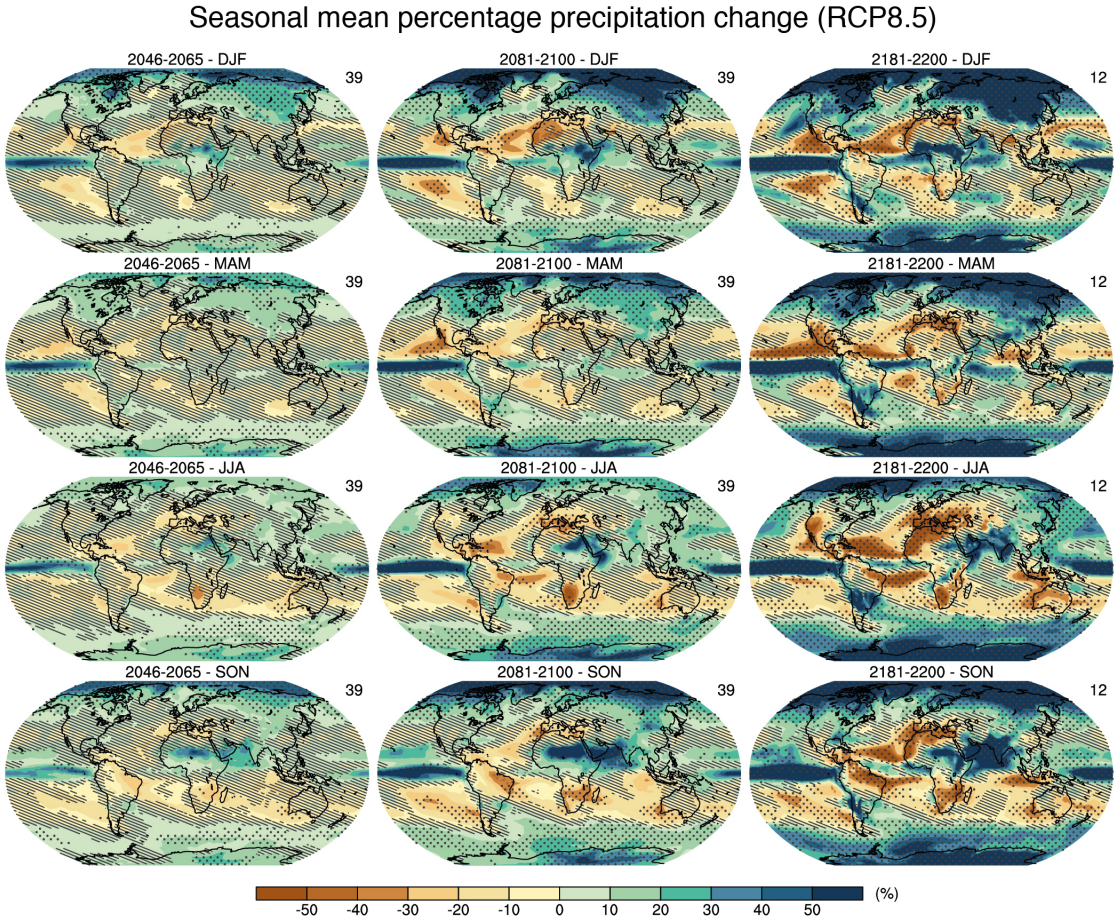


Figure 1.6: *Multi-model CMIP5 average percentage change in seasonal mean precipitation relative to the reference period 1986–2005 averaged over the periods 2045–2065, 2081–2100 and 2181–2200 under the RCP8.5 forcing scenario. Hatching indicates regions where the multi-model mean change is less than one standard deviation of internal variability. Stippling indicates regions where the multi-model mean change is greater than two standard deviations of internal variability and where at least 90% of models agree on the sign of change (see Box 12.1).* Figure extracted from Collins et al. (2013) (their Fig. 12.22).

Italian Peninsula.

In summer (Fig. 1.8b), a meridional gradient dominates the climatological precipitation pattern. Climatological precipitation is very low across the whole basin, being minimum in the south of all Mediterranean peninsulas and maximum over mountain regions. The reason for such low values is the strong subsidence that prevails in the Mediterranean region during summer. While in the western part of the basin the subsidence is mainly associated with the subtropical climatological anticyclone over the Atlantic ocean located to the west of the Mediterranean region [Rodwell and Hoskins, 2001], in the eastern Mediterranean the subsidence is mainly driven by the Indian summer monsoon [Rodwell and Hoskins, 1996; Ziv et al., 2004; Rizou et al., 2015]. Moreover, Simpson et al. (2015) have recently found that about 30% of the summer subsidence over the Mediterranean region is due to

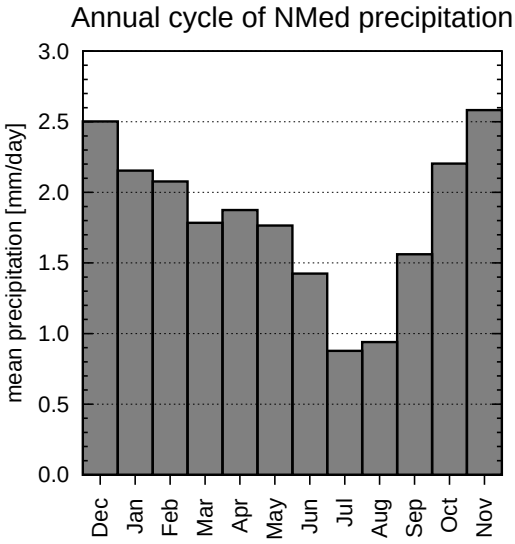


Figure 1.7: Annual cycle of E-OBS precipitation (1950-2014) averaged over land in the NMed region.

the topography of north Africa and the Middle East.

In all seasons, the inter-annual variability of northern Mediterranean precipitation is mainly modulated by the North Atlantic Oscillation (NAO), an atmospheric pattern that reflects the fluctuations in the sea level pressure (SLP) difference between northern and mid-latitudes in the north Atlantic area. Circulation changes associated with the NAO explain a large fraction of the precipitation in the northern Mediterranean region in all seasons [Hurrell et al., 2003], as discussed in the following section.

1.2.1 The North Atlantic Oscillation

The NAO is a large scale phenomenon consisting of an exchange of atmospheric mass between two centres of action located over the Azores and over Iceland. The projection of this phenomenon onto the SLP field is a meridional dipole located in the north Atlantic region. Although this pattern is present year-round [Barnston and Livezey, 1987], it is more intense in winter. It is also in winter when the NAO explains the largest fraction of inter-annual SLP variability in the north Atlantic and the largest fraction of northern Mediterranean winter precipitation variability [Hurrell et al., 2003]. From now on, when we talk about the NAO we will refer to the winter pattern.

In order to assess the state of the oscillation at a given instant, climatologists use different indices. Some of these indices are based on the difference in SLP between the two NAO centres of action [e.g. Jones et al., 1997]. The NAO index obtained with this method, however, does not capture displacements of the centres of action. Alternatively, since the leading empirical orthogonal function (EOF) of winter north Atlantic SLP corresponds to the NAO, we can define the inter-annual NAO index as the first principal component (PC)

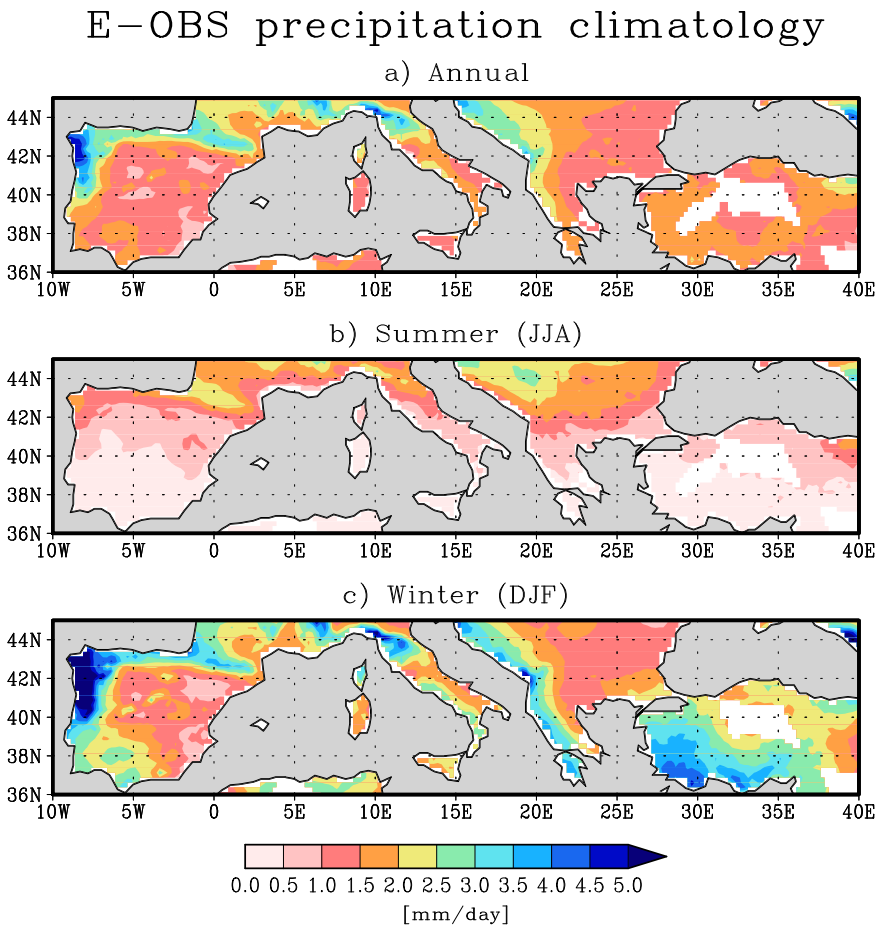


Figure 1.8: Climatology of E-OBS (a) annual (b) June-August summer and (c) December-February winter precipitation (1950-2014) in the northern Mediterranean region. White regions indicate grid-points with more than 30% of missing values in their time series.

of this SLP field. In this thesis we use this method because it allows us to project the EOF obtained from observational data onto the SLP field of climate model simulations. For convention, a positive value of the NAO index indicates an accumulation of mass over the southern center of action and a depletion of mass over the northern center of action (Iceland).

The direct linear relation between NAO and precipitation is shown in Fig. 1.9 in terms of the regression between both variables. Precipitation over southern Europe and the north-western African coast is strongly anticorrelated with the NAO, particularly in the Iberian Peninsula. On the other hand, regressions are positive and strong over northern Europe, in particular over the north of the British Isles and the Scandinavian Peninsula.

The meridional contrast in the winter precipitation signature of the NAO is due to the circulation anomalies associated with each phase of the NAO, as the diagram shown in Fig. 1.10 illustrates. The redistribution of mass in positive NAO winters strengthens both the Azores hight and the Icelandic low, intensifying the velocity of the winds over Europe

[Hurrell et al., 2003]. In these situations, the moisture transport in the north Atlantic region develops a small south-southwest component that shifts the area of maximum moisture transport to northern Europe, creating an area of moisture convergence and favouring precipitation there [Hurrell, 1995]. In negative NAO winters, instead, the moisture transport is more zonal and becomes maximum over Mediterranean latitudes. The northward shift of the moisture transport during the positive NAO phase also modifies the balance between evaporation and precipitation [Hurrell et al., 2003]. Relative to negative NAO situations, positive NAO winters exhibit negative evaporation-precipitation balances (more precipitation than evaporation) over the British Isles and the Scandinavian Peninsula. Over central and southern Europe, instead, balances are positive, indicating a decrease of precipitation relative to evaporation.

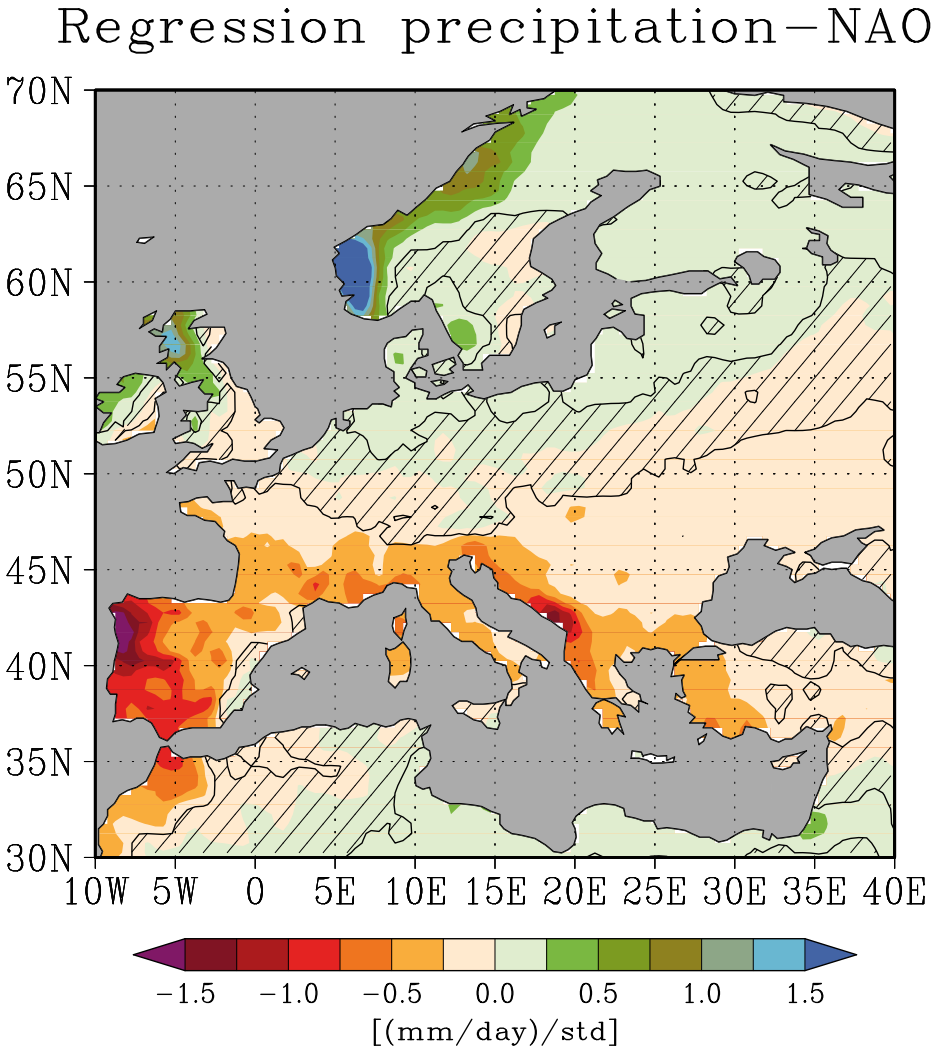


Figure 1.9: Regression of winter (December-March) GPCC precipitation onto the NAO index (estimated as the first PC of the Trenberth SLP field in the 90°W-30°E/40°N-70°N region) for the 1902-2010 interval. Series have been detrended before the calculation. Hatched regions indicate regions where the regression is not statistically significant with a 95% confidence level.

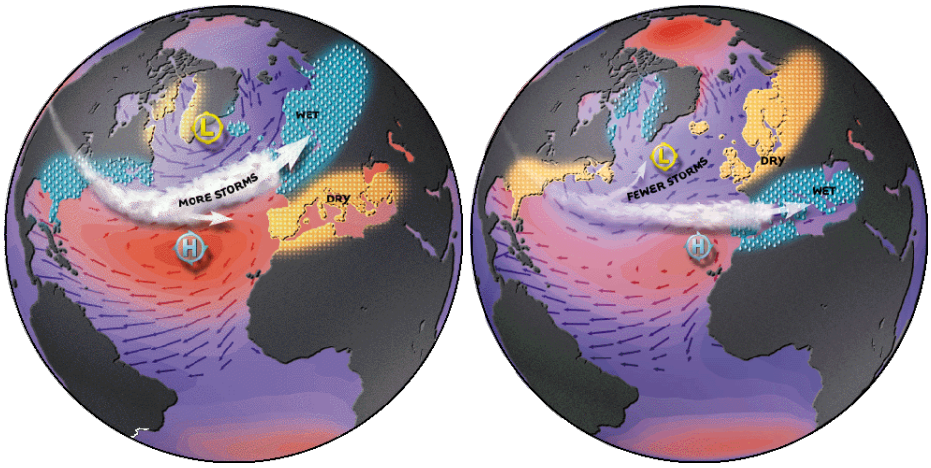


Figure 1.10: Diagram of weather anomalies associated with the positive (left) and the negative (right) phases of the NAO. Figures by Martin Visbeck and Heidi Cullen (www.ldeo.columbia.edu/res/pi/NAO).

1.3 Climate change in the Mediterranean region

According to the last IPCC reports, climate simulations from CMIP3 and CMIP5 multi-model ensembles project a strong drying over the northern Mediterranean region (Fig. 1.11) [Christensen et al., 2007; Christensen et al., 2013]. But, are these projections realistic? To what extent are they consistent across models? The goal of this thesis is to find an answer to these two key questions.

To answer the first question and increase our confidence in the projections of drying over the northern Mediterranean region, we look for a comparable drying signal in observations. A climate signal is considered detectable if it is not consistent with internal variability alone [e.g. Knutson et al., 2013]. In addition, a signal can be attributed to anthropogenic forcings if it is also consistent with historical forced simulations. Another way to increase our confidence in the projections of drying is to examine whether climate models reproduce the NAO well, as it is the main driver of Mediterranean precipitation as discussed above [Hurrell et al., 2003].

Looking for evidence that the projected northern Mediterranean drying signal is already visible in observations, many authors have reported precipitation decreases in the last decades. For instance, Giorgi and Lionello (2008) report area-averaged precipitation reductions (9.5°W-38.5°E/28°N-48°N) in all seasons but autumn, based on the difference between 1981-2000 and 1961-1980 means. Consistently with this result, van Haren et al. (2012) find 1961-2000 winter drying trends (October-March) across the northern Mediterranean basin using several data sets and re-analysis data, and Philandras et al. (2011) report negative trends at individual gauge stations in the northern Mediterranean in annual and winter precipitation for the period 1900-2010. Focusing on the drying in summer, Pal et al. (2004) find precipitation decreases (June-August) in some regions of the Mediterranean

basin, estimated as the difference between 1976-2000 and 1951-1975 means. In order to take any of these results as possible responses of northern Mediterranean precipitation to the increase in GHG concentrations, however, one should first reject the hypothesis that the reported precipitation changes are associated with internal variability alone. Yet, many studies do not subject these results to such careful scrutiny.

An approach which guarantees that internal variability does not play a major role in the observed precipitation changes is the one followed by Mariotti et al. (2008), who estimate the drying as the precipitation trend for a century-long time interval (1900-2007). In general, we can approximate a particular precipitation change to the sum of the contribution of a persistent drying or wetting signal and the contribution of random internal variability. Therefore, calculating long-term trends reduces the contribution of internal variability by increasing the signal (trend) to noise (internal variability) ratio of the change. As a result, Mariotti et al. (2008) find a statistically significant trend of −0.048 (mm/day)/century

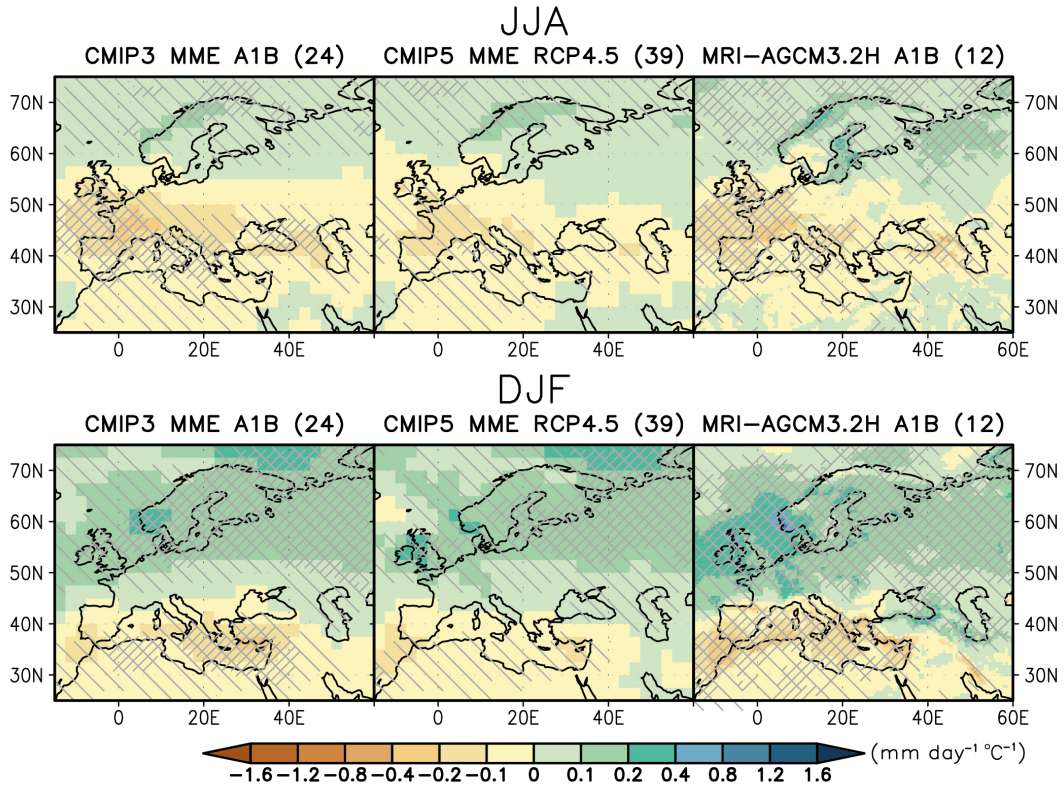


Figure 1.11: *Projected summer (June-August) and winter (December-February) precipitation changes between 2080-2099 and 1986-2005 according the the CMIP3 (left; SRESA1B scenario) and CMIP5 (middle; RCP4.5 scenario) multi-model ensemble means. Right maps are 2075-2099 minus 1979-2003 precipitation changes under the SRESA1B according to the MRI-AGCM3.2 model. Precipitation changes are normalized by the global annual mean surface air temperature change in each scenario. Simple (double) hatching indicates regions where more then 66% (90%) of models coincide with the sign of the multi-model ensemble mean.* Figure extracted from Christensen et al. (2013) (their Fig. 14.22).

in Mediterranean area-averaged (10°W-40°E/28°N-47°N) annual mean precipitation with GHCN data. Nevertheless, as we shall examine in Chap. 2, we have low confidence in observational precipitation changes estimated for periods longer than 60 years due to the scarcity of gauge stations in the first half of the twentieth century. In fact, the statistically significant trend found by Mariotti et al. (2008) is not reproducible using other data sets, as pointed out by themselves.

Another way to verify that an observed precipitation change is not explainable with internal variability only is to check that results are not sensitive to the choice of time interval. While the contribution of the drying or wetting signal persists for an extended period of time, the contribution of internal variability varies between consecutive time intervals due to its random nature. To carry out such analysis, in Chap. 4 we propose the use of two two-dimensional diagrams in which we plot all possible trends of any length and all possible decadal and multi-decadal anomalies in a given time series. With these diagrams, we are able to distinguish which of the reported precipitation changes cited above are easily attributable to internal variability alone. Finally, finding that recent observed changes in the NAO are not explainable by internal variability alone would also indicate that northern Mediterranean precipitation is somewhat under the influence of an external forcing.

Even though only with observational data we can hypothesise that a particular precipitation change is inconsistent with internal variability alone, in order to conclude that a drying signal is detectable we must use model simulations. In Chap. 5 we crate distributions of possible precipitation trends according to all models within the CMIP5 ensemble. On the one hand, we determine the inter-model spread of changes in simulations without any imposed anomalous external forcing. This corresponds to the range of values that a change associated only with internal variability can take. On the other hand, we calculate precipitation changes in simulations that incorporate the evolution of the anomalous external forcings observed throughout the twentieth century. The obtained range of changes corresponds to the possible magnitudes that a precipitation change can take if it is under the influence of external forcings. By calculating where observed precipitation trends fall with respect to these two distributions we can test whether a persistent drying signal is detectable in observational precipitation and whether this is attributable to external forcings [Knutson et al., 2013].

A few authors have examined detection and attribution of the Mediterranean drying signal [van Haren et al., 2012; Barkhordarian et al., 2013; Hoerling et al., 2012; Kelley et al., 2012b]. Nevertheless, results are sometimes contradictory. For instance, Hoerling et al. (2012) find the observed area-averaged (10°W-40°E/30°N-45°N) winter (November-April) precipitation change 1971-2000 minus 1902-1970 within the spread of that in CMIP3 model simulations with historical forcings but at the very edge of the spread of unforced model simulations. Therefore, they conclude that it is unlikely that the observed drying

is attributable to internal variability alone. However, Kelley et al. (2012b) conclude that the contribution of internal variability to the observed 1965-1995 winter (November-April) precipitation trend is dominant because they find that the multi-model trend in CMIP3 forced simulations is very weak in comparison with the observed one, and that unforced simulations are able to exhibit 30-year trends stronger than the observed one. Unfortunately, the reconciliation of different conclusions in not straightforward due to the fact that different studies use different observational data sets, different simulations, average different months over a different domain, use different time intervals and estimate precipitation changes with different methods. To solve this problem, in Chap. 5 we detail which particular conclusions can be drawn for each time interval and each part of the basin and check that they are not sensitive to the data set or the months included in season definitions. We also discuss the differences between results obtained with CMIP5 and CMIP3 models.

The second goal of this thesis is to evaluate whether projections are consistent across models. In general, studies describing the future behaviour of Mediterranean precipitation base their analyses only on the multi-model mean, without paying attention to possible inter-model discrepancies. Although examining trends in the multi-model mean is a first step towards characterizing the most likely evolution of northern Mediterranean precipitation in projections, it may mask large discrepancies across models. To evaluate to what extent are projections of drying consistent across models, in Chap. 6 we evaluate the inter-model agreement regarding the seasonal distribution of the projected precipitation changes, their magnitude and their geographical distribution. We also determine the range of different behaviours that precipitation from different projections follow in their transition towards drier conditions and the spread in the projected time of emergence of the forced signal (i.e. the approximate year in which the drying signal stands above internal variability).

Finally, in Chap. 6 we also examine which mechanisms lead to the projected strong northern Mediterranean drying in projections. In particular, we examine whether precipitation reductions projected for the twenty-first century are driven by the NAO or whether they are driven by other mechanisms response to the global warming such as the "dry gets drier, wet gets wetter" mechanism derived from the Clausius–Clapeyron equation or the poleward shift of the subtropical precipitation minima described by Scheff and Frierson (2012) discussed above. In addition to these processes that could drive future precipitation changes in many parts of the world, Rowell and Jones (2006) propose other mechanisms that could explain the strong projected summer drying over Europe. First, they argue that precipitation decreases in spring would lead to drier soils in summer, which could reduce convective rainfall in summer. Second, they suggest that the reduction of soil moisture in summer, product of the warming, could create a positive feedback mechanism and reduce the soil moisture further, which would explain why the drying is particularly intense in

summer. Third, they hypothesise that, because land regions warm faster than sea regions (due to the greater thermal inertia of water), the increase of sea evaporation associated with the warming could be unable to compensate for the increase in atmospheric moisture capacity over land, which would also reduce land precipitation [Manabe et al., 1992; Gibelin and Déqué, 2003]. Finally, they suggest that large scale circulation patterns could also play a major role in explaining the strong projected summer drying.

To summarize, this thesis is organized as follows: In Chap. 2 we describe the observational data sets and model simulations used in the rest of the thesis and discuss their strengths and weaknesses. Next, in Chap. 3 we explain the statistical methods in which our calculations are based and provide an explanation about how they are performed and how they should be interpreted. In Chap. 4 we evaluate the hypothesis that the recent evolution of observational northern Mediterranean precipitation (and the NAO index) is consistent with what one would expect if a persistent trend was superimposed on internal variability. Then, in Chap. 5 we use model simulations in order to check whether a drying (upward NAO) signal is detectable in observations and whether it is attributable to external forcings. Finally, in Chap. 6 we study similarities and discrepancies among CMIP5 models regarding projections of future drying in the northern Mediterranean region. In Chap. 7 we discuss the main conclusions of the thesis.

Chapter

TWO

Data

In this chapter we detail the data sources for this thesis. In Sec. 2.1 we list the available observed precipitation data sets and discuss which ones are more suitable for our purposes. We also describe the SLP and temperature observational data sets used in this thesis. In Sec. 2.2 we specify which models and simulations we use and the differences between climate experiments.

2.1 Observational gridded data sets

As we focus on the climate of the Mediterranean region at the regional scale, the best choice is to use gridded data sets. Unlike data from individual stations, gridded data sets provide a continuous field and do not bias the analysis towards regions with a large density of measurements available. These products are created by interpolating the available data from single gauge stations at a given time instant onto a common lon×lat grid, and facilitates the comparison of results from observational data with those obtained with model outputs. Mainly because the number of stations is not constant in time (as we shall see below), time series constructed from such gridded products are not homogeneous and therefore they should be used with caution in trend estimations [i.e. Harris et al., 2014]. To avoid drawing erroneous conclusions arising from this non-homogeneity, we test the reproducibility of our results by using different data sets.

2.1.1 Precipitation

Precipitation is extremely variable at short scales, both spatially and temporally. When the density of stations is scarce, interpolation methods are used to fill a regular data grid. As actual precipitation fields are not smooth, however, these grids can miss local scale

phenomena. These are the main characteristics that we need to consider regarding observational data bases:

- Resolution: Since we work with large scale phenomena, an extremely high resolution should not be required. Nevertheless, because of the mesoscale nature of Mediterranean topography (which includes mountain chains, enclosed seas, archipelagos and small peninsulas), this region requires a more fine resolution than other regions of the world.

- Length: Because we perform climate studies we look for series as long as possible. In general, however, gridded precipitation data prior to 1950 is based on a limited number of stations (as we shall see below). For this reason, definite conclusions cannot be drawn for earlier periods.

- Station coverage: More important than the resolution of the data grid is the density of stations used in its development. Some regions, in particular desert regions and developing countries, generally have few measuring stations. In our case, as we shall see below, north Africa lacks measurements because it includes large non-inhabited areas.

In table 2.1 we list the most common used gridded precipitation data sets covering the Euro-Mediterranean region. Except PREC, all these products cover only land regions. In Sec. 2.1.1.1 we show the station coverage of the main data sets. Because we focus on seasonal to multi-decadal scales, monthly means are used. Even though finding similar results using different data sources would increase the robustness of our conclusions, we must keep in mind that the data sets in table 2.1 are far from being independent, as many gauge measurements are shared by several data sets [Harris et al., 2014]. Next we present the main characteristics of all these data sources.

	data set	interval	resolution
1	E-OBS v10.0*	1950-2014	0.25°×0.25°
2	GPCC v6	1901-2010	0.5°×0.5°
3	CRU TS v3.22	1901-2013	0.5°×0.5°
4	UDel v3.02	1900-2010	0.5°×0.5°
5	GHCN v2	1900-2014	5°×5°
6	Hulme v1.0	1900-1998	3.75°×2.50°
7	PREC**	1948-2014	2.5°×2.5°

Table 2.1: Observational data sets of gridded monthly land precipitation. For each of them we indicate its interval coverage and its spatial resolution (lon×lat). *E-OBS data is only available over Europe and some regions in the north African coast. **PREC data is also available over sea.

1. E-OBS: This data set belongs to the project ENSEMBLES of the European Climate Assessment and Dataset (ECA&D) and includes daily data for precipitation, minimum, maximum and mean temperature, and SLP. Its developers have focused on creating a high resolution data set over Europe (0.25°×0.25°, the highest resolution available for gridded precipitation) using the best methods for spatial interpolation. The current version of the E-OBS data set is v10.0, for which records begin in 1950 and are monthly updated. The main reference of the E-OBS data set, where interpolation methods and other technical characteristics are explained, is Haylock et al. (2008). Data can be downloaded from www.ecad.eu/download/ensembles/download.php.

2. GPCC: The precipitation data set developed by the Global Precipitation Climatology Centre (GPCC) is a German contribution to the World Climate Research Programme (WCRP) and to the Global Climate Observation System (GCOS). From a variety of precipitation products, we use the Full Data Reanalysis Product which is the one that includes measurements from a largest number of stations all around the world and has the highest spatial resolution 0.5°×0.5°. The current version is GPCC v6, which begins in 1901 and ends in 2010 (unfortunately it has not been updated in recent years). The gauge data treatment and the interpolation techniques used to create this gridded data set are explained in detail by Schneider et al. (2011), Becker et al. (2013) and Schneider et al. (2014). GPCC data can be downloaded from ftp-anon.dwd.de/pub/data/gpcc/html/fulldata_v6_download.html.

3. CRU: Developed by the Climate Research Unit (CRU) of the University of East Anglia (UEA), this data set provides precipitation anomalies (among other field variables) for land areas all over the world. Time series begin in 1901 and are regularly updated. Its spatial resolution is 0.5°×0.5°. The current version of this data set is CRU TS v3.22 and its main reference is Harris et al. (2014) (or Mitchell and Jones (2005) for older versions). CRU data can be downloaded from badc.nerc.ac.uk/browse/bad/cru/data or from www.cru.uea.ac.uk/cru/data/hrg/.

4. UDel: This data set is developed by the Department of Geography of the University of Delaware (UD). With a resolution of 0.5°×0.5° it covers all land regions of the world during the time interval 1900-2010. The current version is v3.02, described by Matsuura and Willmott (2012). UDel data can be downloaded from climate.geog.udel.edu/~climate/html_pages/download.html#P2011.

5. GHCN: The Global Historical Climatology Network (GHCN) is a database of temperature, precipitation and pressure developed by the National Climate Data Center (NCDC), the Arizona State University (ASU) and the Carbon Dioxide Information

Analysis Center (CDIAC). The current version of the data set is GHCN v2 and precipitation records cover all land regions of the world beginning in 1900. Although it is monthly updated, its coarse resolution 5°×5° makes this data set not very suitable for our purposes. Therefore, we only use it to reproduce results by other authors. GHCN data and more information about the data set can be downloaded from ftp.ncdc.noaa.gov/pub/data/gnu/v2/grid/.

6. Hulme: This data set was created by Dr. Mike Hulme at the Climate Research Unit (CRU). The current version v1.0 of the data set covers all land regions of the world within the interval 1900-1998 with a spatial resolution of 3.75°×2.50°. We only use this data set to test the reproducibility of some of our results because of its low resolution and, more importantly, because it ends in 1998. Data and more information about the Hulme data set can be downloaded from iridl.ldeo.columbia.edu/SOURCES/.UEA/.CRU/.Hulme/.Global/.

7. PREC: The PRECipitation REConstruction Dataset (PREC) is developed by the National Oceanic and Atmospheric Administration (NOAA). Updated regularly, it covers the entire world since 1948 at a 2.5°×2.5° resolution. This is the only data set in table 2.1 that contains data over sea, based on an EOF reconstruction of gauge measurements over land and over islands. The main reference of PREC data set is Chen et al. (2002). Data can be downloaded from www.esrl.noaa.gov/psd/data/gridded/data.prec.html.

In Chap. 4 we take E-OBS as our reference data set because it is built with a meticulous quality control in the selection of gauge stations, uses sophisticated interpolation methods, and is updated on a monthly basis. Unfortunately, the data set begins in 1950. Therefore, for century-long studies we use GPCC data, which begins in 1901. Even though this product has not been updated since 2010, it is the one with the most stations available among the data sets beginning around 1900. CRU and UDel data sets are also used in Chap. 4 to assess the robustness of our results to the choice of data source. GHCN and Hulme have only been employed to reproduce results from other authors and we do not show any results obtained with them. Finally, the PREC data set was used in the analyses described in Bladé et al. (2012a), some of which we refer to and reproduce in this thesis.

2.1.1.1 Gauge station coverage

For the four data sets in table 2.1 with the finest spatial resolution (E-OBS, GPCC, CRU and UDel) we examine the temporal and spatial distribution of their gauge measurements. For GPCC and CRU the number of gauge stations per grid-point on every month is provided in their respective data portals. For E-OBS, a list of the coordinates and the temporal

coverage of each station used in the development of the data set is available. For UDel data, this information has been kindly provided by Dr. Kenji Matsuura.

First we represent the location of gauge stations in the Euro-Mediterranean region for each data set. Figures 2.1a-c show results for GPCC, CRU and UDel data sets for their respective entire record (which covers roughly the whole twentieth century and the first decade of the twenty-first century). Figures 2.1d-f show results if we only consider data since 1950. As we shall see below, from 1950 onwards the amount of gauge measurements increases in all data sets. Results for E-OBS, with a smaller grid size and only covering the interval from 1950 onwards are represented separately in Fig. 2.2.

In Fig. 2.1 we see that GPCC is the data set that incorporates data from a largest number of gauge stations across the Euro-Mediterranean region. The best station coverage for GPCC data is found over central and western Europe (in particular almost all of Germany is covered with gap-less stations and even several stations per grid-point, not shown), whereas over northern Africa and the Near East observations are sparse. To a lesser extent, UDel data also have good station coverage, particularly over central Europe, while the number of gauge stations used in the CRU data set is considerably smaller. Even though the CRU data set has fewer gauge stations, phenomena at a regional scale should still be captured and therefore this data set will be useful to test the reproducibility of the results obtained with other data sets. Finally, E-OBS, which is considered to be the best data set over Europe due to its high resolution and its optimal interpolation methods, has actually fewer available stations than GPCC, particularly over north Africa, Anatolia and the Near East (Fig. 2.2). Presumably, the lack of stations in E-OBS data is associated with the fact that they use a strict quality control in the selection of gauge stations.

With Figs. 2.1 and 2.2 we also assess the quality of the local precipitation series, evaluated as the percentage of the total period that is covered at each grid-point (we calculate, at each grid-point, the percentage of months that have at least 1 measure available). Low values of this magnitude can indicate a large number of gaps in the local station series or, more commonly, that measurements do not cover the entire interval. Within the Mediterranean region, GPCC data exhibit records with missing values in more than 50% of the months over some parts of the Italian, Balkan and Anatolian peninsulas (Fig. 2.1a). Nevertheless, if we remove the period 1901-1949 from the analysis (Fig. 2.1d), the percentage of gaps decreases substantially (consistent with the increase in the number of gauge measurements). For the time interval 1950-2010, regions located in the west, central and south Europe are well sampled, with less than 25% of gaps almost everywhere. Although UDel is also well station-covered, almost all grid-points with measurements exhibit more than 50% of missing values during the entire record (Fig. 2.1c). In this case, the improvement obtained using the 1950-2010 interval is generally very small except over the British Isles (Fig. 2.1f). Instead, for CRU data most of the stations contain fewer than 50% of missing

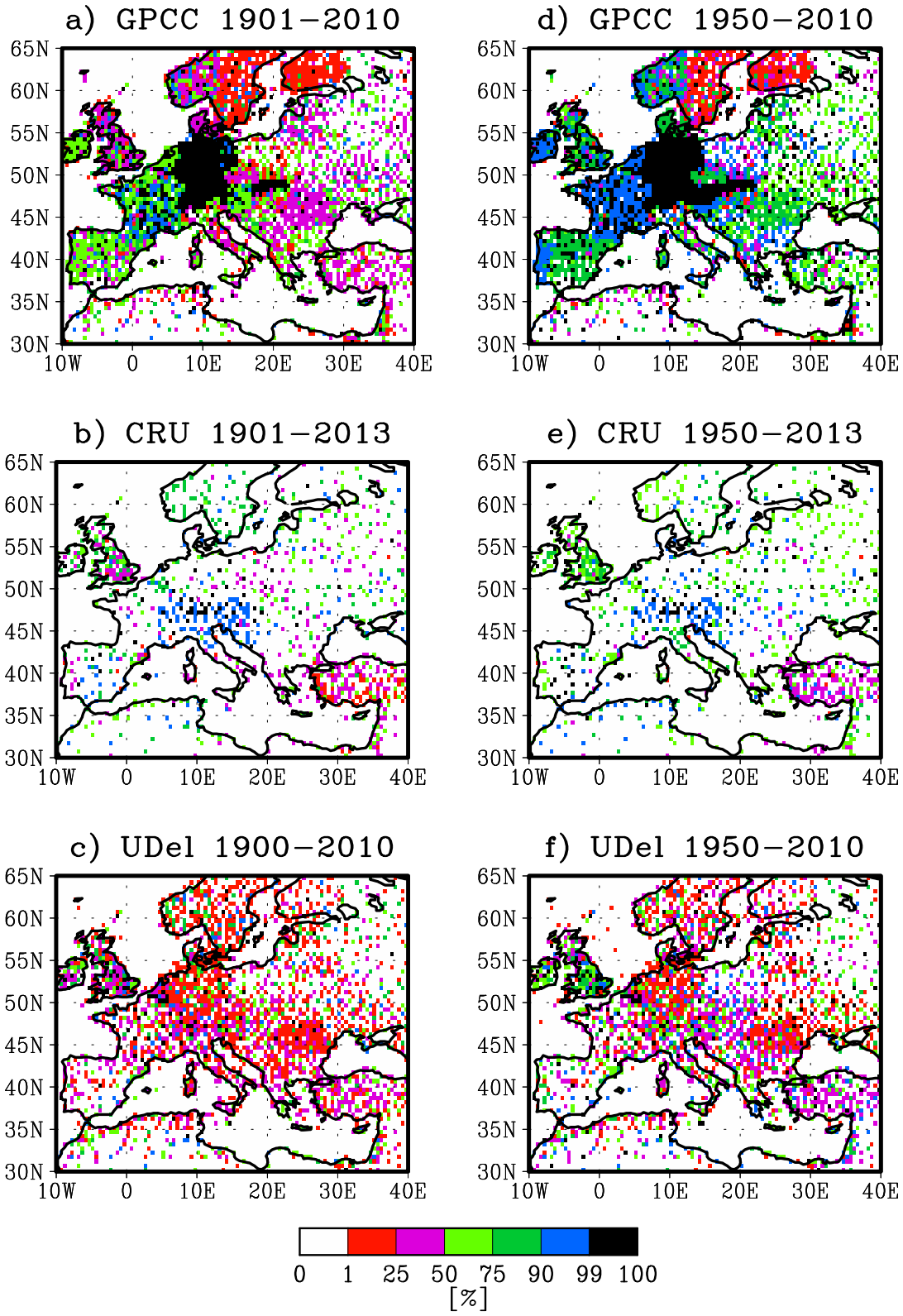


Figure 2.1: Station location and percentage of months with at least one gauge measurement at each grid-point for different data sets and time intervals. Each grid-point may contain data from several stations.

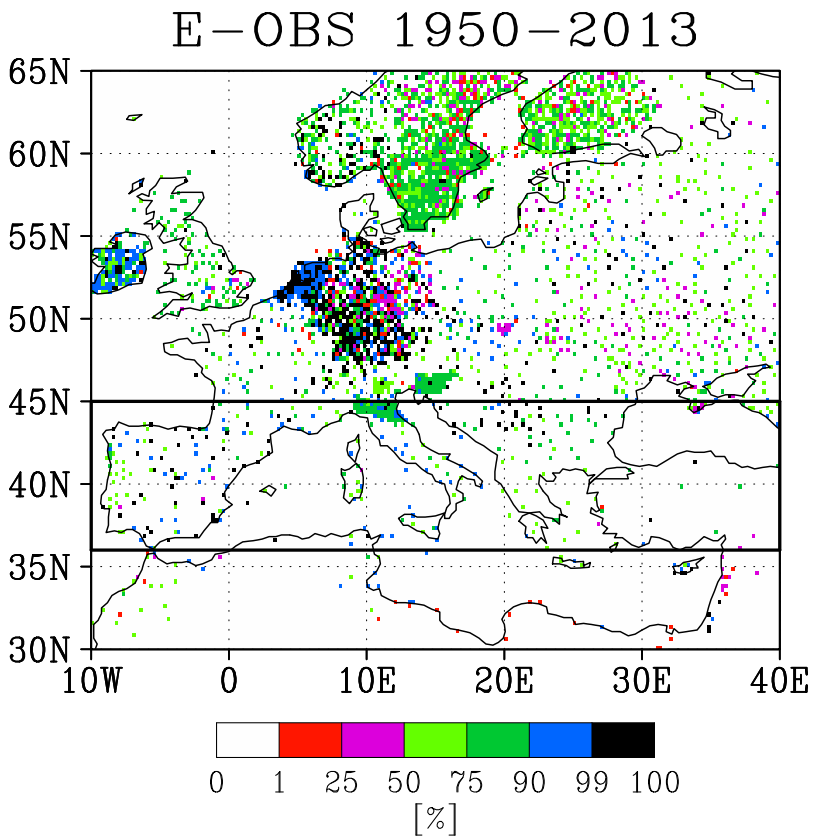


Figure 2.2: Station location and percentage of months with at least one gauge measurement at each grid-point for E-OBS data (1950-2013). Each grid-point may contain data from several stations. The box indicates the limits of the NMed region.

months for the 1950-2012 interval (Fig. 2.1e). Finally, we note that very few of the E-OBS stations exhibit more than 50% of missing values, even over southern Europe (Fig. 2.2).

Even though our original objective was to analyse precipitation changes in the entire Mediterranean region, having realized that northern Africa is not well station-covered in any of the analysed data sets (except at the coast), we have considered it more appropriate to restrict our analysis to the northern basin. Therefore, from now on our region of study will be the northern Mediterranean (NMed) region, defined as the South Europe/Mediterranean region used in IPCC AR5, removing latitudes south from the Gibraltar strait (10°W-40°E/36°N-45°N, indicated in Fig. 2.2).

The consequences of averaging data over regions with an insufficient number of stations can be illustrated in the following example. Although, as pointed out by Mariotti et al. (2008), the 1901-2010 linear trend of annual precipitation averaged over the entire Mediterranean (10°W-40°E/30°N-45°N) exhibits a statistically significant trend (−0.087 (mm/day)/century, Fig. 2.3, red line) with a 99% confidence level, this result is based partially in data from north Africa, where there are few gauge stations and most information results from interpolations. Restricting the area-average to the NMed region (Fig. 2.3, blue

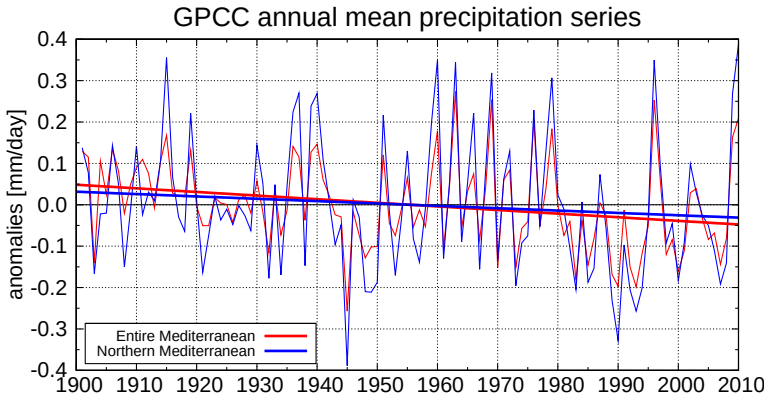


Figure 2.3: Anomalies of GPCC annual mean land precipitation averaged over the entire Mediterranean (red curve) and over the northern Mediterranean (blue curve). Anomalies are estimated with respect to the 1901-2010 climatological mean. Linear fits are also included.

line), the obtained linear trend is weaker (-0.057 (mm/day)/century) and not statistically significant with a 95% confidence level.

In Fig. 2.4 we represent the time evolution of the number of total gauge stations in the NMed region in the four considered data sets. We thus confirm that GPCC is the data set with more stations all along the record, with more than double the amount of stations than any other data set practically at any time. In particular, the E-OBS data base, which is considered to be the best precipitation data set for Europe, uses about four times fewer stations in the NMed region than GPCC during the second half of the twentieth century.

For the century-long data sets, we see the general increase of gauge stations in the second half of the twentieth century commented above. A decrease in the number of stations is also observed in the last years of each data set, associated with a delay between the time in which gauge measurements are taken by local organizations and the time they are processed and delivered to the organizations that create the large scale data sets [Schneider et al., 2011]. Finally, the annual cycle is visible in most time series, with generally more stations in the cold season and fewer in the warm season.

For the four data sets we see that, in general, increases in the number of stations are not gradual but abrupt, contributing to the non-homogeneity of the local time series commented above. Given these limitations, we will consider a result to be robust only if it is reproducible with all four precipitation data sets. In case of discrepancies, we will discuss and justify which set is more reliable in each case.

2.1.2 Sea level pressure

As we shall see in Sec. 3.2, we define the NAO index as the first principal component of the SLP field over the north Atlantic area. Following our previous works on the summer NAO [Bladé et al., 2012a;Bladé et al., 2012b], to calculate the observed NAO index we use data

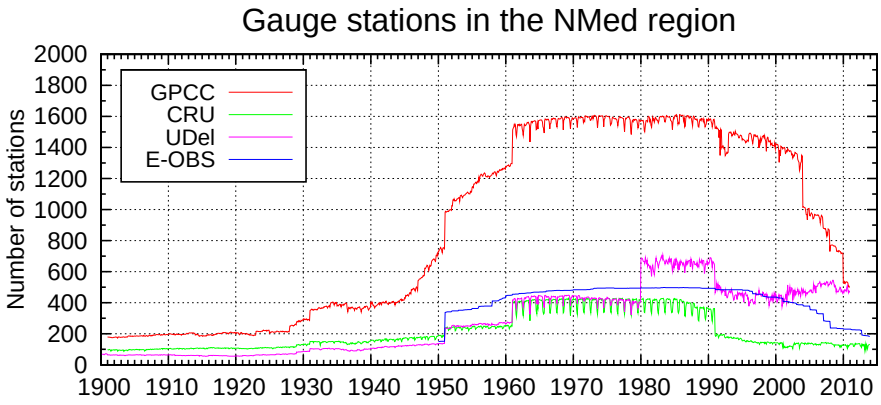


Figure 2.4: Total number of precipitation gauge stations per month in the NMed region for different data sets.

from the Trenberth SLP gridded data set (see Sec. 3.1.8). It covers the time interval 1899-2014 with a spatial resolution of 5°×5°. See Trenberth and Paolino (1980) for a complete explanation about the data set. Data can be downloaded from rda.ucar.edu/datasets/ds010.1/index.html#subset.html?_da=y.

2.1.3 Surface air temperature

As commented in Chap. 1, precipitation changes in the Mediterranean region are not as strong as temperature changes. For this reason, in Chaps. 5 and 6 we use observed surface air temperature data to illustrate the behaviour of a field variable that is clearly influenced by a long-term trend and compare it with that of precipitation. We use the HadCRUT v4.2.0.0 gridded data set from the Climate Research Unit (CRU), which combines surface air temperature over land from CRUTEM v4 and sea surface temperature from HadSST v3. It covers the interval 1850-2014 with a spatial resolution of 5°×5°. The main reference of HadCRUT set is Morice et al. (2012) and data can be downloaded from www.cru.uea.ac.uk/cru/data/temperature/.

2.2 Climate models

In Chaps. 5 and 6 we use climate model data to test the hypothesis that observed precipitation trends are attributable to external forcing and to assess which aspects of future projections are robust and reliable. In this section we describe the model ensembles and experiments used in those chapters.

The Coupled Model Intercomparison Project (CMIP) is an initiative of the Working Group of Coupled Modelling (WGCM) of the World Climate Research Programme (WCRP) that provides standards of implementation to modelling centres all around the

world since 1995. Through the web portal of the Program for Climate Model Diagnosis and Intercomparison (PCMDI; http://cmip-pcmdi.llnl.gov/), data outputs and model information is made available to the climate community. Once every few years, CMIP members propose a list of climate *experiments* for all models to run under a set of common parameters in order to make their outputs directly comparable. Generally speaking, these experiments impose particular configurations of boundary conditions, corresponding to those observed in pre-industrial times, in the twentieth century, or expected according to possible future scenarios.

Currently, the CMIP project is on its phase 5 (CMIP5) [Taylor et al., 2012b]. CMIP5 models are the core of the modelling studies of the IPCC AR5 and this thesis. Models from the phase 3 of the CMIP project (CMIP3) [Meehl et al., 2007] are also used here in order to asses the improvements in the new models and to reproduce some results found in the literature. Some of the differences between CMIP5 and CMIP3 models are a higher spatial resolution, the inclusion of new cloud and aerosol processes and a better representation of some climate processes such as a more sophisticated scheme of atmospheric chemistry [Boucher et al., 2013; Flato et al., 2013].

In this thesis we analyse climate model output from four kinds of experiment: pre-industrial control simulations, historical simulations, RCP4.5 projections and RCP8.5 projections. Taylor et al. (2012b) provide a complete explanation of how different experiments are designed. Pre-industrial simulations are realizations of climate models under boundary conditions characteristic of mid-nineteenth century and without anomalous external forcings imposed. These simulations are generally very long (few centuries) in order to detect drifts and are useful to characterize the internal variability of climate variables. Historical simulations are initialized at some point of a particular pre-industrial control simulation and from then on they incorporate observed forcings (both natural and anthropogenic) from mid-nineteenth century to the first decade of the twenty-first century.

Initialized in the final years of a particular historical simulation, projections are model representations of the future climate. In climate projections we must provide the model with those boundary conditions that we expect we will observe in the future. Because these conditions depend on many uncertain socio-economic factors, several possible future Representative Concentration Pathway (RCP) scenarios are designed. Figure 2.5 shows the evolution of the radiative forcing associated with each scenario. In Chap. 6 we use only data from RCP4.5 and RCP8.5 scenarios. These scenarios are the most studied and they are also those with more model outputs available. RCP4.5 is chosen because the evolution of the forcing is intermediate between the most optimistic and the most pessimistic scenarios, even though the radiative forcing stabilizes in mid-twenty-first century and, therefore, the forced signal is expected lose strength from then on. Instead, the RCP8.5 scenario is the

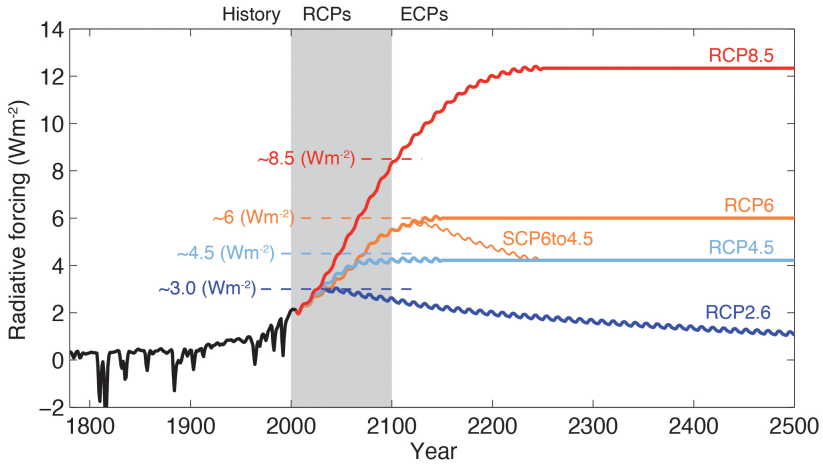


Figure 2.5: *Total RF (anthropogenic plus natural) for RCPs and extended concentration pathways (ECP)—for RCP2.6, RCP4.5, and RCP6, RCP8.5, as well as a supplementary extension RCP6 to 4.5 with an adjustment of emissions after 2100 to reach RCP4.5 concentration levels in 2250 and thereafter. Note that the stated RF levels refer to the illustrative default median estimates only. There is substantial uncertainty in current and future RF levels for any given scenario. Short-term variations in RF are due to both volcanic forcings in the past (1800–2000) and cyclical solar forcing assuming a constant 11-year solar cycle (following the CMIP5 recommendation), except at times of stabilization. (Reproduced from Figure 4 in Meinshausen et al., 2011.).* Figure extracted from Cubasch et al. (2013) (their Fig. 1 in box 1.1).

scenario with the largest increment in the radiative forcing in which we expect to find the largest anthropogenic forced signal.

CMIP5 data (table 2.2) have been downloaded using the methodology described by Gámez (2013) [personal communication]. Because not all climate simulations and projections cover the exact same time interval, we use 1902-2004 as our historical period (because it is the common period in most of the climate simulations and the observational GPCC data set used as a reference) and 2007-2099 as our main interval of study for projections. Due to these choices, a few CMIP5 runs covering shorter periods are omitted. Pre-industrial control simulations and projections without the corresponding historical simulation available have also been omitted. Finally, we have also discarded model runs in which any of the studied variables (precipitation, SLP and surface air temperature) is missing.

To facilitate inter-model comparisons and to calculate multi-model ensemble means, prior to any analysis we re-grid the model data onto a common grid using bi-linear interpolation (see Sec. 3.1.5). For precipitation we use a 0.5°×0.5° grid, the same as that of the GPCC observational gridded data set (see Sec. 3.1.5). This choice follows the methodology that we used in Bladé et al. (2012a). In the course of that study, however, we found that conclusions were not sensitive to the choice of the grid or the re-gridding method. For SLP fields used for NAO calculations, and also following Bladé et al. (2012a), we re-grid model data to a common 2.5°×2.5° grid. The methodology used to compute the NAO in each climate model is explained in Sec. 3.2. Finally, for surface air temperature data used

in Sec. 5.1 we use data re-gridded to a 1°×1° grid. In order to compare model results with those found in observational data, when computing areal averages we exclude sea grid-points (our observational data do not cover them). The land-sea mask used to distinguish between land and sea was created using the CRU data set as a reference. Note that, in order to be rigorous, the optimal approach should have been to mask out sea grid-points prior to the re-gridding using each model's own land-sea mask.

In some parts of the thesis we do not show results for individual simulations but, for simplicity, we present results for model means. In those cases, we first calculate the model means by averaging the results obtained with all available runs of the same model for the experiment we are working with. When we calculate the multi-model ensemble mean we average these models means. Doing this, we avoid that models with more runs contribute more to the mean than models with fewer runs. The multi-model mean created with all CMIP5 models will be referred as *multi5*.

Different runs of a given experiment are labelled with letters *r*, *i* and *p* (e.g. r1i1p1). This nomenclature helps to identify the historical simulation used to initialize a particular projection (we have checked that runs labelled with the same name for consecutive intervals are indeed consecutive analysing metadata). Taylor et al. (2012a) provide a complete description of names and conventions used in CMIP5 data. The "realization number" *r* indicates a particular choice of the initial conditions. The "initialization method indicator" *i* is associated with a particular method of initializing simulations with observational data. There are no *i* differences between runs in any of the models and experiments used in this

	model	resolution	pre-ind.	his.	RCP4.5	RCP8.5
1	ACCESS1.0	~1.9°×1.25°	1 (500)	1	1	1
2	ACCESS1.3	~1.9°×1.25°	1 (500)	3	1	1
3	BCC-CSM1.1	~2.8°×2.8°	1 (500)	3	1	1
4	BCC-CSM1.1(m)	~1.1°×1.1°	1 (400)	3	1	1
5	BNU-ESM	~2.8°×2.8°	1 (559)	1	1	1
6	CanESM2	~2.8°×2.8°	1 (996)	5	5	5
7	CCSM4	~1.25°×0.9°	3 (1327)	6	6	6
8	CESM1(BGC)	~1.25°×0.9°	1 (500)	1	1	1
9	CESM1(CAM5)	~1.25°×0.9°	1 (319)	3	3	3
10	CESM1(CAM5.1, FV2)	~2.5°×1.9°	-	4	-	-
11	CESM1(FASTCHEM)	~1.25°×0.9°	1 (222)	3	-	-
12	CESM1(WACCM)	~2.5°×1.9°	1 (200)	1	-	-
13	CMCC-CESM	~3.75°×3.7°	1 (277)	1	-	1
14	CMCC-CM	~0.75°×0.7°	1 (330)	1	1	1
15	CMCC-CMS	~1.9°×1.9°	1 (500)	1	1	1

Table 2.2: *Continues in the following page.*

	model	resolution	pre-ind.	his.	RCP4.5	RCP8.5
16	CNRM-CM5	~1.4°×1.4°	1 (850)	10	1	5
17	CSIRO-Mk3.6.0	~1.9°×1.9°	1 (500)	10	10	10
18	EC-EARTH	~1.1°×1.1°	1 (280*)	7	2	2
19	FGOALS-g2	~2.8°×2.8°	1 (700)	4	1	1
20	FIO-ESM	~2.8°×2.8°	1 (800)	3	3	3
21	GFDL-CM2.1	~1.25°×2°	-	10	-	-
22	GFDL-CM3	1.25°×2°	1 (500)	5	1	1
23	GFDL-ESM2G	~1.25°×2°	1 (500)	1	1	1
24	GFDL-ESM2M	~1.25°×2°	1 (500)	1	1	1
25	GISS-E2-H-CC	1.25°×2°	1 (251)	1	1	1
26	GISS-E2-H-NINT	1.25°×2°	1 (540)	6	5	2
27	GISS-E2-H-TCAD	1.25°×2°	1 (531)	5	5	1
28	GISS-E2-H-TCADI	1.25°×2°	1 (531)	6	5	2
29	GISS-E2-R-CC	1.25°×2°	1 (251)	1	1	1
30	GISS-E2-R-NINT	1.25°×2°	1 (550)	6	6	2
31	GISS-E2-R-TCAD	1.25°×2°	1 (531)	6	5	1
32	GISS-E2-R-TCADI	1.25°×2°	1 (531)	6	6	2
33	HadCM3	3.75°×2.5°	-	10	-	-
34	HadGEM2-A	~1.9°×1.25°	-	1	1	1
35	HadGEM2-CC	~1.9°×1.25°	1 (241)	1	1	1
36	HadGEM2-ES	~1.9°×1.25°	1 (576)	4	4	4
37	INM-CM4	2°×1.5°	1 (500)	1	1	1
38	IPSL-CM5A-LR	~3.75°×1.9°	1 (1000)	6	4	4
39	IPSL-CM5A-MR	~2.5°×1.3°	1 (300)	3	1	1
40	IPSL-CM5B-LR	~3.75°×1.9°	1 (300)	1	1	1
41	MIROC5	~1.4°×1.4°	1 (670)	5	3	3
42	MIROC-ESM	~2.8°×2.8°	1 (630)	3	1	1
43	MIROC-ESM-CHEM	~2.8°×2.8°	1 (255)	1	1	1
44	MPI-ESM-LR	~1.9°×1.9°	1 (1000)	3	3	3
45	MPI-ESM-MR	~1.9°×1.9°	1 (1000)	3	3	1
46	MPI-ESM-P	~1.9°×1.9°	1 (1156)	2	-	-
47	MRI-CGCM3(p1)	~1.1°×1.1°	1 (500)	3	1	1
48	MRI-CGCM3(p2)	~1.1°×1.1°	-	2	-	-
49	MRI-ESM1	~1.1°×1.1°	-	1	-	-
50	NorESM1.M	~2.5°×1.9°	1 (501)	3	1	1
51	NorESM1.ME	~2.5°×1.9°	1 (252)	1	1	1

Table 2.2: *Begins in the previous page.* List of climate models in the CMIP5 multi-model ensemble. The second column indicates the original spatial resolution of each model. The remaining columns indicate the number of available runs per model for pre-industrial and historical simulations and for RCP4.5 and RCP8.5 projections. The total length (in years) of pre-industrial control simulations is indicated in brackets. *The length of the available data in the EC-EARTH pre-industrial run is 280 years for precipitation, 451 years for SLP and 452 years for surface air temperature.

thesis. Finally, the "perturbed physics" *p* number indicates a particular configuration of the physical parameters of a given model.

While most models only have one *p* adjustment, three models have more than one (GISS-E2-H, GISS-E2-R and MRI-CGCM3). For instance, there are 25 historical simulations available for the GISS-E2-R model distributed as follows: 6 runs named r#i1p1, 6 runs named r#i1p2, 6 runs named r#i1p3 and 7 runs named r#i1p#. A simple look at the trend in global annual mean surface air temperature in each of the 25 runs (Fig. 2.6) illustrates why runs with different *p* number should not be treated as if they were the same model. In this case, following Miller et al. (2014) we have considered p1 runs to belong to the model GISS-E2-R-NINT, p2 runs to belong to GISS-E2-R-TCAD and p3 runs to belong to GISS-E2-R-TCADI. We have discarded the remaining runs. With the GISS-E2-H model we have proceeded similarly. Finally, the two versions of the MRI-CGCM3 model have been simply labelled as MRI-CGCM3(p1) and MRI-CGCM3(p2).

After these considerations, we are left with 47 pre-industrial simulations from 45 models, 179 historical simulations from 51 models, 105 RCP4.5 projections from 42 models and 84 RCP8.5 projections from 43 models. Table 2.2 contains the list of all CMIP5 models and simulations used in this thesis. The last four columns indicate the number of runs available for each experiment. The total length in years of pre-industrial control runs is indicated in brackets. The second column indicates original spatial resolutions.

For CMIP3 models, simulations for the historical period correspond to the experiment called *20C3M*. Because this multi-model ensemble was developed in mid-2000s, historical simulations generally end in 1999 or 2000 and, therefore, do not cover our entire historical period. To extend the CMIP3 historical simulations up to 2004 we use projections for the scenario A1B of the Special Report on Emissions Scenarios (*SRESA1B*) [Morita et al.,

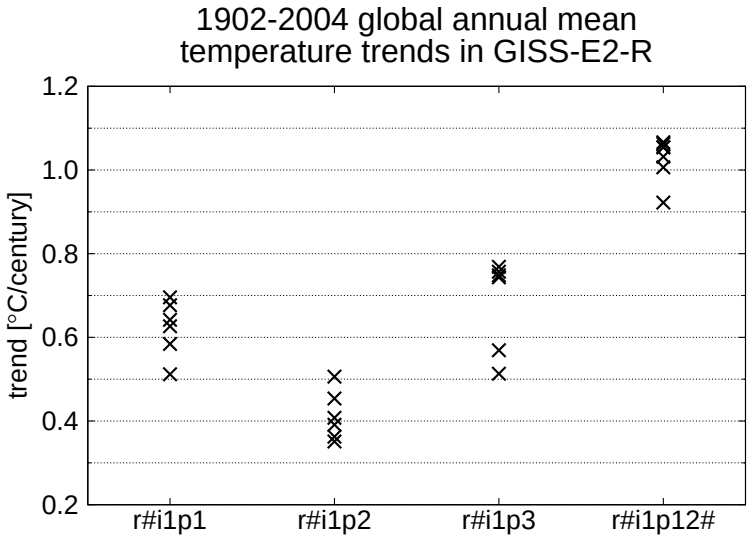


Figure 2.6: 1902-2004 annual mean global surface air temperature trend in each GISS-E2-R historical simulation. Simulations are grouped according to their perturbed physics *p* number.

	model	resolution	*20C3M+SRESA1B*
1	BCCR-BCM2.0	1.9°×1.9°	1
2	CCSM3	1.4°×1.4°	6
3	CGCM3.1(T47)	~2.8°×2.8°	5
4	CGCM3.1(T63)	~1.9°×1.9°	1
5	CNRM-CM3	~1.9°×1.9°	1
6	CSIRO-MK3.0	~1.9°×1.9°	1
7	CSIRO-MK3.5	~1.9°×1.9°	1
8	INGV-ECHAM4	~1.1°×1.1°	1
9	ECHAM5/MPI-OM	~1.9°×1.9°	4
10	ECHO-G	~3.9°×3.9°	2
11	FGOALS-g1.0	~2.8°×2.8°	3
12	GFDL-CM2.0	2.0°×2.5°	1
13	GFDL-CM2.1	2.0°×2.5°	3
14	GISS-AOM	3°×4°	2
15	GISS-EH	4°×5°	3
16	GISS-ER	4°×5°	2
17	UKMO-HadCM3	2.50°×3.75°	1
18	UKMO-HadGEM1	~1.3°×1.9°	1
19	INM-CM3.0	4°×5°	1
20	IPSL-CM4	2.50°×3.75°	1
21	MIROC3.2(hires)	~1.1°×1.1°	1
22	MIROC3.2(medres)	~2.8°×2.8°	3
23	MRI-CGCM2.3.2	~2.8°×2.8°	5
24	PCM	~2.8°×2.8°	3

Table 2.3: List of climate models in the CMIP3 multi-model ensemble. The second column indicates the original spatial resolution of each model. The last column indicates the number of available historical simulations.

2000]. As the differences between different CMIP3 scenarios are minimal in the first years of the twenty-first century, results are not sensitive to the choice of scenario. The concatenation *20C3M+SRESA1B* was carried out during the development of Bladé et al. (2012a), using metadata to associate each projection with the corresponding historical simulation. CMIP3 models are listed in table 2.3 together with their original spatial resolution and the number of available runs. A total of 53 runs from 24 models have been used. CMIP3 data were downloaded from the CMIP3 multi-model database (esg.llnl.gov:8443/index.jsp) and were interpolated following the same methodology explained for CMIP5 models. The multi-model mean created with all CMIP3 models will be referred as *multi3*.

Chapter

THREE

Methods

This thesis addresses the issue of climate change detection and attribution in the northern Mediterranean using statistics to test physically based hypotheses. To avoid using statistical methods as if they were black boxes, in Sec. 3.1 we briefly explain their basis, why they are useful, how they are applied and how they should be interpreted. We also provide references that include deeper explanations. In Sec. 3.2 we explain how we calculate the NAO index using observational data and climate model simulations.

3.1 Statistical methods

3.1.1 Linear trends

We define the linear trend of a given variable over a particular time interval as the slope of the linear fit estimated using the method of least squares. As we see in expression 3.1, we approximate the magnitude of the variable x at the instant t as the sum of an independent term (a) and a term that is linearly related to the value of t (bt).

$$\hat{x}(t) = a + bt \tag{3.1}$$

The method of least squares consists of finding the parameters a and b that minimize the quadratic difference between the actual values of x and the linear fit $\hat{x}(t)$ (expressions 3.2 and 3.3, where $\mathrm{cov}(x,t)$ is the covariance of x and t, $\mathrm{var}(t)$ the variance of t, and $\overline{x}$ and $\overline{t}$ the mean of each variable).

$$b = \frac{\mathrm{cov}(x,t)}{\mathrm{var}(t)} \tag{3.2}$$

$$a = \overline{x} - b\overline{t} \tag{3.3}$$

3.1.1.1 Detrending time series

For certain analyses, is necessary to remove the linear trend b from a given time series $x(t)$ prior to a calculation (expression 3.4). In this thesis we perform time series detrending basically in two cases:

- When we evaluate the standard deviation of a time series as a measure of the variability associated only with natural causes. Not detrending the time series would include variability associated with a potential superimposed trend.
- When we calculate the linear relation between two time series, as a strong trend in both variables (caused by a common forcing or not) would artificially inflate the magnitude of their regression or correlation.

$$x_{dt}(t) = x(t) - bt \tag{3.4}$$

3.1.2 Regressions

The term b in expression 3.2 is used to estimate the linear trend of the variable x within a particular time interval. In a more general context, this parameter is also used to estimate the linear regression between two arbitrary variables (i.e. the change in x associated with a given change of t). Although linear regression is the most used method to estimate the linear relationship between two variables, Lanzante (1996) describes other methods less sensitive to the presence of outliers in time series.

3.1.3 Correlations

To estimate statistically the correlation between two variables (x_1 and x_2) we use the Pearson correlation coefficient r (expression 3.5, where $\mathrm{cov}(x_1, x_2)$ is the covariance of x_1 and x_2, and $\sigma(x_1)$ and $\sigma(x_2)$ the standard deviations).

$$r = \frac{\mathrm{cov}(x_1, x_2)}{\sigma(x_1)\sigma(x_2)} \tag{3.5}$$

3.1.3.1 Averaging correlations

To calculate the average of a given set of correlations we cannot calculate its arithmetic mean directly. That is because the sampling distribution of Pearson correlation coefficients does not follow a normal distribution and, as a result, the direct average of a set of correlations is not a good approximation to the population mean [e.g. Wilks, 2006]. Instead, we apply the Fisher's Z transformation (expression 3.6) to each correlation in our set. Because the sampling distribution of the Z-transformed correlation coefficients does follow a

normal distribution, we can average the obtained Z values [e.g. Lane, 2013]. Finally, we inverse-transform the result to obtain the mean correlation.

$$Z = \text{atanh}(r) \tag{3.6}$$

3.1.4 Areal averages

All data used in this thesis are available as gridded data sets. This means that we have a magnitude of a particular variable at each point (λ,ϕ) of a regular grid (where λ is the longitude and ϕ is the latitude), which represents the average value of such variable in the area surrounding it. Because the longitudinal and the latitudinal grid spacings ($\Delta\lambda$ and $\Delta\phi$ respectively) are regular over the entire grid, the area over which a given grid-point is representative diminishes with latitude.

To calculate the areal average of a variable v using gridded data, we weight the magnitude of v at each grid-point by the percentage of the total area that it covers using expression 3.7, where dx_j and dy_i are the zonal and meridional sizes of each grid cell (the area of influence of a particular grid-point). Given that $dx_j = a\Delta\lambda\cos(\phi_j)$ and $dy_i = a\Delta\phi$ (with a the average radius of the Earth), expression 3.7 becomes expression 3.8. Therefore, for gridded data sets with constant longitudinal and latitudinal spacings, we must weight each value by the (normalized) cosine of its latitude when calculating areal averages.

$$\overline{v} = \frac{\sum_{i,j} v_{ij} dx_j dy_i}{\sum_{i,j} dx_j dy_i} \tag{3.7}$$

$$\overline{v} = \frac{\sum_{i,j} v_{ij} \cos(\phi_j)}{\sum_{i,j} \cos(\phi_j)} \tag{3.8}$$

3.1.5 Interpolations

In some sections of this thesis we need two or more data sets to share the same grid (for instance when we create multi-model means). Several techniques exist to re-grid and interpolate gridded data sets. Maeland (1988) and Press et al. (1992) are good references about the topic.

In this thesis, for simplicity, we only use bilinear interpolation using the four grid-points (two in each direction) closest to the one we are evaluating. This method consists of finding the magnitude of a given variable v at the point (x, y) by weighting the magnitude of v at these grid-points (expression 3.9, where d is the distance between the target point and one of the four neighbouring points, x^- and x^+ indicate the closer points in the longitudinal dimension, y^- and y^+ indicate the closer points in the latitudinal dimension, and Δx and Δy indicate the longitudinal and latitudinal grid spacing).

$$v_{x,y} = \frac{d_{x^+} d_{y^+}}{\Delta x \Delta y} v_{x^-,y^-} + \frac{d_{x^-} d_{y^+}}{\Delta x \Delta y} v_{x^+,y^-} + \frac{d_{x^+} d_{y^-}}{\Delta x \Delta y} v_{x^-,y^+} + \frac{d_{x^-} d_{y^-}}{\Delta x \Delta y} v_{x^+,y^+} \tag{3.9}$$

3.1.6 Tests of statistical significance

Each time we calculate a statistical parameter, we must assess whether the hypothesis that it has been obtained by chance (i.e. the null hypothesis) can be rejected or not. We consider a certain value as statistically significant with a certain confidence level (usually 95%) when the probability of obtaining it by chance is lower than a certain threshold (in this case 5%). Wilks (2006) and Hartmann (2014) are good references to go further on hypothesis testing.

All tests of statistical significance consist of calculating a statistical value x, comparing it with its probability distribution, and finding the probability of obtaining a value for the statistic larger than the actual one (Fig. 3.1). Depending on whether we assume that x follows a theoretical probability distribution or not, we classify a test as *parametric* or *non-parametric*.

If we do not have physical arguments to believe that x will be found at any of the two extremes of the distribution, we use *two-sided* tests. This means that we consider x as statistically significant at a 95% confidence level if it falls within any of the two 2.5% tail-end areas of the probability distribution. Instead, we use *one-sided* tests when we have an a priori reason to believe that we will find x at a particular end of the distribution. In that case, we consider x as statistically significant at a 95% confidence level if it falls within that particular 5% tail-end area of the probability distribution.

3.1.6.1 Monte Carlo tests

Monte Carlo tests are non-parametric tests in which we create the probability distribution of x by reordering the data from which x is evaluated a large number of times. Although this methodology is computationally slow, it is simple and intuitive. As an example, we

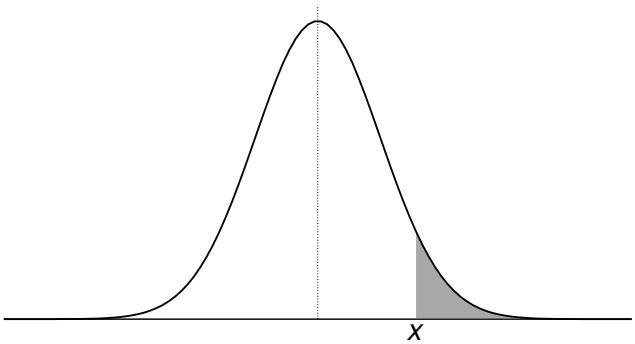


Figure 3.1: Hypothetical probability distribution of a certain statistic x. The shaded area indicates the probability of obtaining a value larger than the one we are testing.

explain in detail how we use a Monte Carlo test to evaluate the statistical significance of a linear trend b over a particular time interval of length N.

First, we calculate the linear trend using the original data (with expression 3.2), obtaining b^*. Then, we assess which is the probability of obtaining a time series with trend larger than b^* within the given range of data. To do it, we reorder the time series randomly a large number of times (we usually do 1000 re-samplings), calculating each time the trend and creating the probability distribution of b. Finally, we count which percentage of obtained trends is larger than b^*. If the obtained percentage is smaller than 2.5% (5%) if we are performing a two-sided (one-sided) test, the null hypothesis that there is no trend in the original series is rejected and b^* is considered statistically significant with a 95% confidence level.

3.1.6.2 Student's test for autocorrelated series

When a time series is autocorrelated, re-sampling significance tests overestimate the statistical significance of trends. This is because in autocorrelated series the effective number of degrees of freedom (DOF, which in non autocorrelated series coincides with the length of the series) is reduced. In general, as the number of DOF decreases, the probability distribution of trends (see Fig. 3.1) becomes wider. Viewing the reduction of the number of DOF as an effective reduction of the length of the original series, the probability of finding large trends increases while the probability of finding small trends decreases. When we reorder the original series to perform a Monte Carlo test, however, the autocorrelation is lost and, therefore, the reduction of the number of DOF is not taken into account. As a result, the obtained probability distribution of trends is narrower than the one we would obtain given the correct number of DOF. As a consequence, the probability of finding a trend that is member of the *correct* distribution under the tail of the *wrong* distributions is artificially low.

This is not usually the case for precipitation series but it is the case for temperature series because the thermal inertia of the ocean inhibits temperature changes from one year to the next. In autocorrelated series, then, we estimate the statistical significance of our trends using the parametrical test proposed by Santer et al. (2000). First we calculate the regression residuals $e(t)$, which are defined as the difference between each value of the original series $x(t)$ and the value of the linear fit at the same instant $\hat{x}(t)$. Then, the standard error of the trend s_b is defined by expression 3.10, where n_t is the length of the original series and s_e is the variance of the regression residuals s_e, estimated with expression 3.11.

$$s_b = \frac{s_e}{\sqrt{\sum_{t=1}^{n_t}(t-\bar{t})^2}} \tag{3.10}$$

$$s_e = \sqrt{\frac{\sum_{t=1}^{n_t} e(t)^2}{n_t - 2}} \tag{3.11}$$

As in the case of Monte Carlo tests, applying this test to autocorrelated series overestimates the statistical significance of the results. To account for this we replace the number of DOF n_t in expression 3.11 by the effective number of DOF n_e estimated with expression 3.12, where r_1 is the lag-1 autocorrelation of the original series. Finally, we compare the statistic $t_b = b/s_b$ with the Student's distribution with $\nu = n_e$ DOF.

$$n_e \approx \frac{1 - r_1}{1 + r_1} n_t \tag{3.12}$$

3.1.6.3 Wilcoxon-Mann-Whitney test

In order to evaluate the statistical significance of the difference between two averages, for instance the change of precipitation between two independent time intervals, we use the non-parametric Wilcoxon-Mann-Whitney test. In this case, the null hypothesis is that the two data samples belong to the same distribution. This test consists of joining data from the two samples, arrange them in ascending order, and then sum the rank of the the members of each sample obtaining R_1 and R_2. The Mann–Whitney statistic U is defined following either of the two expressions 3.13, where n_1 and n_2 are the sample sizes. The obtained statistic follows a normal distribution with mean μ and standard deviation σ as indicated in expression 3.14.

$$U_1 = R_1 - \frac{n_1}{2}(n_1 + 1) \tag{3.13a}$$

$$U_2 = R_2 - \frac{n_2}{2}(n_2 + 1) \tag{3.13b}$$

$$\mu = \frac{n_1 n_2}{2} \tag{3.14a}$$

$$\sigma = \sqrt{\frac{n_1 n_2 (n_1 + n_2 + 1)}{12}} \tag{3.14b}$$

3.1.6.4 Field significance

When working with data fields, we evaluate the statistical significance of our results at every grid-point using the tests explained above. The resulting field of p-values (which represent the probability under the tail of the distribution) indicates the local significance

of the results at each grid-point. It is also necessary, however, to calculate the collective statistical significance of the pattern as a whole, known as *field significance* [Livezey and Chen, 1983].

The field significance is the probability of having a particular proportion p of locally significant grid-points in a given field. To estimate it, we compare the portion of locally significant values in our map with the probability distribution of the percentage of area significant of our field. This distribution can be estimated as the binomial distribution (expression 3.15) of a field with N as the total number of grid-points and S_L as the probability that a given grid-point passes the local significance test (usually 5%). The field significance of our field is then defined as the area under the tail of this probability distribution function.

$$P(p) = \left(\frac{N!}{p!(N-p)!}\right) S_L^p (1 - S_L)^{N-p} \tag{3.15}$$

To illustrate this, we calculate the distribution function of a field with $N = 100$ grid-points and a threshold of local significance of $S_L = 0.05$ (Fig. 3.2). Note that the distribution function is discrete because the possible number of passed tests is discrete. As the red line indicates, we should obtain at least 10 grid-points with local significance in order to obtain a field significance over 95% ($P < 0.05$). While the probability of having at least 9 significant grid-points is still $P = 0.063$, that of 10 is $P = 0.028$.

Although this is the general methodology to estimate field significance, it is only valid if all local tests are linearly independent. In actual geographical fields this is never the case, as climate variables tend to vary smoothly between neighbouring regions, so instead of the

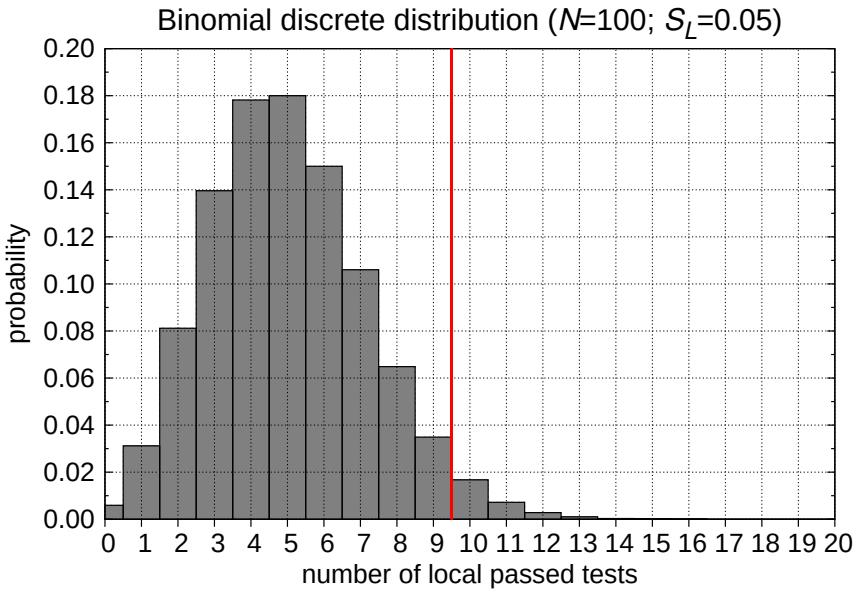


Figure 3.2: Binomial distribution function of a field with 100 DOF and a probability of 0.05 to obtain a significant local value. Red line indicates for which number of locally passed tests the area under the tail of the distribution is lower than 5%.

total number of grid-points we must use the true number of spatial DOFs considering the spatial autocorrelation of our field.

Several methods exist to calculate the DOF of a given field. We follow that used by Sterl et al. (2007) to calculate the significance of a global field. Their method is based on the calculation of the decorrelation length d, defined as the distance over which the correlation drops below $1/e$. Having found this length, they calculate the area of a circle of radius d and compare it with that of the Earth's surface.

Their approach, however, is only valid for global fields. In particular, in the case of our main region of study (the northern Mediterranean), as it is meridionally narrow and longitudinally wide, the resulting circle of radius d overflows the region limits and therefore is not a good measure of the *decorrelation surface* we are looking for. We, then, find more suitable for our situation to calculate separate longitudinal and latitudinal decorrelation lengths (d_x and d_y) and, instead of calculating the area of a circle of radius d, use the area of a rectangle $2d_x \times 2d_y$.

In order to obtain the decorrelation length in each direction we calculate for each pair of grid-points the correlation between the two time series and also the longitudinal and latitudinal distance between them. For trend and linear change calculations, we use only the time period for which the trend is calculated. In case of absolute differences between two time intervals, we use the period that results from joining both periods. As a result, we obtain the spatial correlation function of our field as a function of the distance between pairs of grid-points in both directions. Binning the results into 10 separate intervals, we obtain a monotonically decreasing function and we can estimate for which distance the correlation drops below $1/e$ in each direction.

When we work only with data over land, the calculation of the total surface of the field is not direct. Instead, we sum the area associated with each land grid-point g using expression 3.16, where $\Delta\lambda$ and $\Delta\phi$ are the longitudinal and latitudinal grid spacing and ϕ_n is the latitude at point n. This is only valid for uniform *lat*×*lon* fields.

$$S_{field} = \sum_{g} \Delta\lambda\Delta\phi \cos(\phi_g) \tag{3.16}$$

Finally, to obtain the number of DOF N' of our field we divide the total surface of the field by the decorrelation surface found (expression 3.17). Note that surfaces and distances have been considered in terms of great circle angles. Having found this number, we use it as N in the estimation of our binomial distribution function (expression 3.15) to calculate the field significance of our maps.

$$N' = \frac{S_{field}}{4d_x d_y} \tag{3.17}$$

3.1.7 Filtering time series

In this section we first describe the filtering technique that we use to investigate the low-frequency behaviour of time series. Then, we describe how we create a band-pass filter that allows us to assess the magnitude of the internal variability at decadal to multi-decadal scales, removing the influence of both high frequency variations and lower frequency fluctuations or trends. Hartmann (2014) is a good reference to go further in filtering techniques and learn other methods.

3.1.7.1 Low-pass filter

A filter is considered *non-recursive* if a given value of the filtered series depends only on its neighbour values in the original series. Instead, a filter is considered *recursive* if it also depends on the previous values in the filtered time series. While the former filters are more stable, with the latter we obtain sharper response functions [Hartmann, 2014].

For simplicity, in this thesis we use the general low-pass symmetric non-recursive filter with Lanzcos smoothing described by Hartmann (2014). To obtain the filtered time series f_{fil} with this filter we replace the value of the original series f at the instant t by a weighted average of its N surrounding values, each one with a specific w_k weight associated (expression 3.18). Note that this method inevitably excludes $N/2$ values at both edges of the original time series.

$$f_{fil}(t) = \sum_{k=-N/2-1}^{N/2-1} w_k \, (t + k\Delta t) \, f(t) \tag{3.18}$$

This methodology based on weights is used in other filtering techniques as well. For instance, in an ordinary running mean all weights would be equal. In our case, to filter frequencies higher than α we use the weights defined by expression 3.19.

$$w_k = \frac{1}{k\pi} \sin(\alpha k) \tag{3.19}$$

The response function of a filter is defined as the function that we must convolve with the Fourier transform of our series to obtain the desired distribution of frequencies. Ideally, then, it would be a step function being 1 for frequencies lower than α and 0 for frequencies higher than α. To eliminate small wiggles in the response function of the filter described here we use Lanzcos smoothing over our weights (expression 3.20).

$$\tilde{w}_k = w_k \mathrm{sinc}\left(\frac{\pi k}{N}\right) \tag{3.20}$$

Figure 3.3 represents the response function of a general low-pass symmetric non-recursive filter with Lanzcos smoothing with $N = 11$ and $\alpha = 2\pi/10$ (i.e. which filters

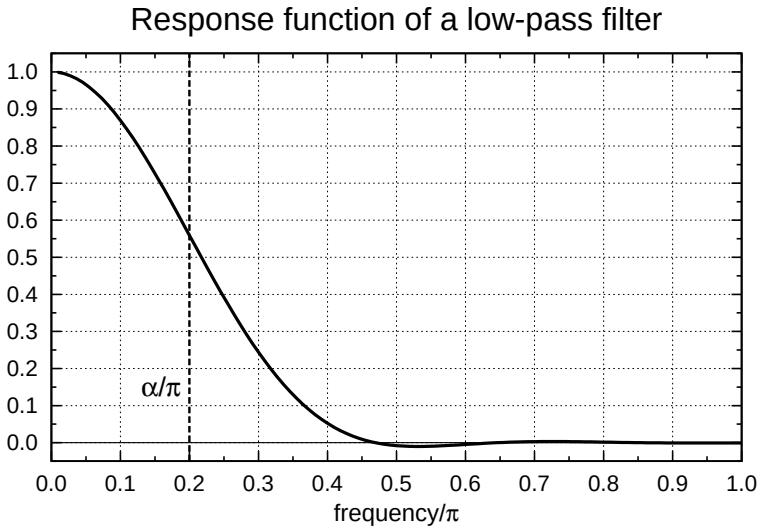


Figure 3.3: Response function of a general low-pass symmetric non-recursive filter with Lanzcos smoothing with $N = 11$ and $\alpha = 2\pi/10$ (indicated with the dashed vertical line).

out periods lower than 10 years). Using a larger N we would obtain a sharper response function but we would lose more values at the extremes of the original series.

3.1.7.2 Band-pass filter

Next we explain how we calculate a band-pass filter that, ideally, should remove frequencies larger than α_H and smaller than α_L from a given series $f(t)$. Again, we follow Hartmann (2014).

First, we apply a high-pass filter with $\alpha = \alpha_L$ to $f(t)$ to remove frequencies lower than α_L. To do it, we apply a low-pass filter with $\alpha = \alpha_L$ and subtract obtained series from $f(t)$, obtaining $f_{hp\text{-}fil}(t)$. Then, we remove frequencies larger than α_H from $f_{hp\text{-}fil}(t)$ with a low-pass filter with $\alpha = \alpha_H$, obtaining $f_{bp\text{-}fil}(t)$. If the response function of the low-pass filter is $R_L^{\alpha_L}$, the response function of the high-pass filter is $R_H^{\alpha_H} = 1 - R_L^{\alpha_H}$ and the response function of the band-pass filter is $R_B^{\alpha_L \alpha_H} = R_H^{\alpha_H} R_L^{\alpha_L}$. Figure 3.4 represents the response function of a band-pass filter with $N = 31$, $\alpha_H = 2\pi/10$ and $\alpha_L = 2\pi/40$ (i.e. which filters periods lower than 10 years and higher than 40 years).

3.1.8 Principal component analysis

The aim of principal component analysis is to decompose a given time-evolving spatial field into linearly independent patterns sorted by the amount of temporal variance of the original data that they explain. This method is particularly useful when the obtained patterns (known as empirical orthogonal functions or EOFs) can be interpreted in terms of atmospheric variability. In this cases, it allows us to study the temporal evolution of these

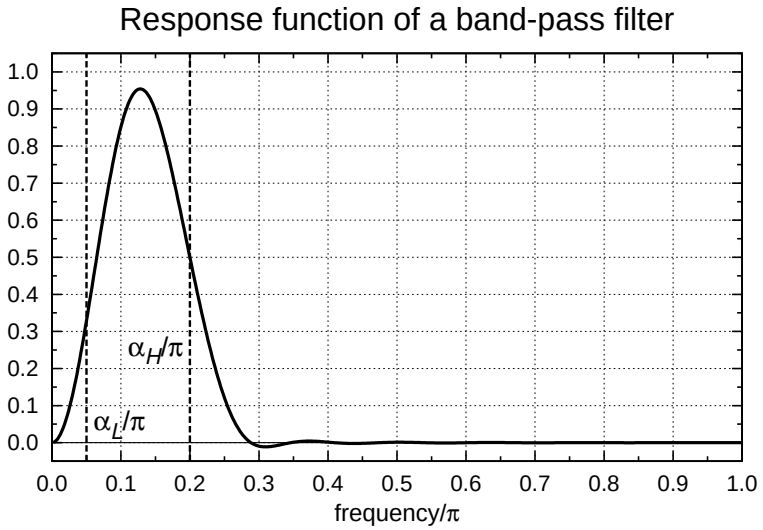


Figure 3.4: Response function of a general band-pass symmetric non-recursive filter with Lanzcos smoothing with $N = 31$, $\alpha_H = 2\pi/10$ and $\alpha_L = 2\pi/40$ (indicated with the dashed vertical lines).

patterns by projecting the original data onto the EOFs at each instant, obtaining time series known as principal components or PCs.

We use principal component analysis to study the geographical and temporal structure of the NAO, as it corresponds to the first EOF of the SLP field (or the geopotential height at a given pressure) over the north Atlantic [Hurrell et al., 2003]. Here we explain in detail the methodology we use to calculate the NAO with PC analysis. In Sec. 3.2 we indicate which data sets, regions and time periods we use to estimate the NAO with observational and model data.

First we take the time series of SLP at each one of the grid-points contained in the selected region and calculate the anomalies with respect to the mean over the entire series. Then, we weight these anomalies by the square root of the cosine of the latitude of that grid-point[1]. With the obtained time series we create a data matrix D of size $N \times T$, where N is the total number of grid-points and T the length of the time series (expression 3.21).

$$D = \begin{pmatrix} p_1(t_1) & p_1(t_2) & ... & p_1(t_T) \\ p_2(t_1) & p_2(t_2) & ... & p_2(t_T) \\ ... & ... & & ... \\ p_N(t_1) & p_N(t_2) & ... & p_N(t_T) \end{pmatrix} \tag{3.21}$$

Then, we create a symmetric matrix C containing the covariance between each pair of time series for individual grid-points (expression 3.22). Because the covariance is de-

[1]We weight for the square root of the cosine of the latitude because later we will calculate the covariance between each pair of grid-point time series and, therefore, the obtained value will be weighted by the product of both square roots and will account for the latitudinal weighting necessary when we work with gridded data sets (see Sec. 3.1.4).

fined as the average product of anomalies between two time series, we calculate the matrix product between D and its transpose D^T and divide by the length of the time series.

$$C = \frac{D\,D^T}{T} = \begin{pmatrix} \mathrm{cov}(p_1,p_1) & \mathrm{cov}(p_1,p_2) & ... & \mathrm{cov}(p_1,p_N) \\ \mathrm{cov}(p_2,p_1) & \mathrm{cov}(p_2,p_2) & ... & \mathrm{cov}(p_2,p_N) \\ ... & ... & & ... \\ \mathrm{cov}(p_N,p_1) & \mathrm{cov}(p_N,p_2) & ... & \mathrm{cov}(p_N,p_N) \end{pmatrix} \tag{3.22}$$

When we calculate the eigenvectors of the covariance matrix we obtain a set of N orthonormal vectors EOF^n (expression 3.23), sorted by the amount of variance of the original field that they explain. Moreover, the portion of variance explained by each of these EOFs is $\lambda_n / \sum \lambda_n$, where λ_n is the eigenvalue associated with the eigenvector n. When we represent each component of these eigenvectors in its grid-point of the original field we obtain the geographical patterns of these EOFs.

$$EOF^n = \begin{pmatrix} x_1^n \\ x_2^n \\ ... \\ x_N^n \end{pmatrix} \tag{3.23}$$

Finally, to obtain the PC associated with each EOF (i.e. its time evolution) we project the weighted original field onto each EOF (expression 3.24). All PCs are orthogonal and their magnitude is proportional to the number of grid-points in the original field. Here we use the PC of the first EOF of the north Atlantic SLP field to describe the evolution of the NAO index and we usually present it normalized by its (detrended) standard deviation.

$$PC^n = D\ EOF^n = \begin{pmatrix} p_1(t_1) & p_1(t_2) & ... & p_1(t_T) \\ p_2(t_1) & p_2(t_2) & ... & p_2(t_T) \\ ... & ... & & ... \\ p_N(t_1) & p_N(t_2) & ... & p_N(t_T) \end{pmatrix} \begin{pmatrix} x_1^n \\ x_2^n \\ ... \\ x_N^n \end{pmatrix} \tag{3.24}$$

There are several ways to visualize the physical patterns associated with the EOFs. In this thesis they are presented as regressions between the PCs and the original data. For instance, to represent the spatial signature of the NAO we plot the regression between the first PC of north Atlantic SLP and the local SLP time series at each grid-point. Other authors display the pattern in terms of composite anomalies at time instants in which the first PC is very positive and in which it is very negative to illustrate the differences between the two phases. Good references to learn more about PC analyses are Hartmann (2014), von Storch and Zwiers (1999), Wilks (2006) and Hannachi et al. (2007).

3.2 Estimation of the NAO index

3.2.1 The NAO in observations

Following our previous publications [Bladé et al., 2012a; Bladé et al., 2012b], in Chap. 4 we define the observed NAO index as the first PC of the winter SLP field over the north Atlantic [Hurrell et al., 2003]. We use the Trenberth SLP data set (1900-2014) over the region 90°W-30°E/40°N-70°N following Greatbatch and Rong (2006). In Fig. 3.5 we represent the obtained NAO pattern displayed in terms of the regression of the SLP field onto the PC. In brackets we indicate the percent of variance of the SLP field explained. Throughout the thesis, winter is defined as the December-March (DJFM) mean for reasons explained in Chap. 4. For convention, the year associated with a given winter is that in which winter ends.

The obtained PC has been compared with the NAO index provided by the Climate Prediction Center (CPC), which is based on a rotated principal component analysis of 700 hPa geopotential height for individual months [www.cpc.ncep.noaa.gov/data/teledoc/telecontents.shtml]. The correlation coefficient between the two series within the interval 1951-2013 is 0.90.

3.2.2 The NAO in model simulations

Following the methodology used in Bladé et al. (2012a), the NAO index in a given model is estimated as the projection of the observed EOF pattern onto the SLP field of that model (in

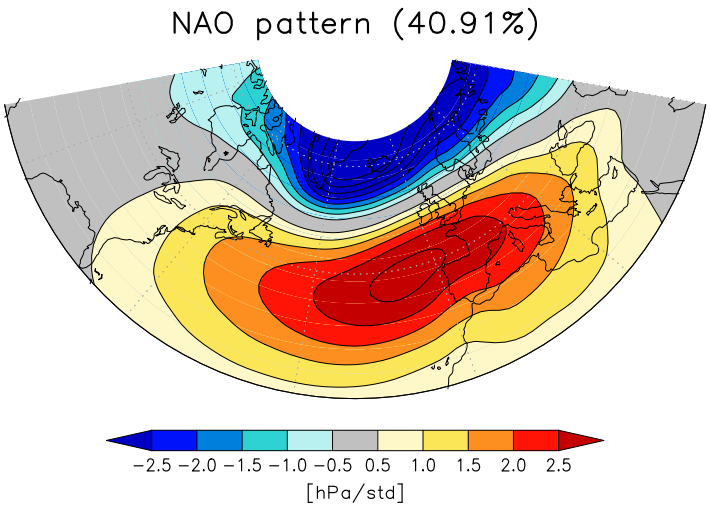


Figure 3.5: Spatial pattern of the observed NAO defined as the first PC of the winter Trenberth SLP field in the 90°W-30°E/40°N-70°N domain for the 1900-2014 interval. The pattern is displayed in terms of the detrended regression of the SLP field onto the normalized NAO index. Thus, anomalies correspond to one standard deviation of the NAO index. In brackets, the percentage of SLP variance explained by the associated EOF.

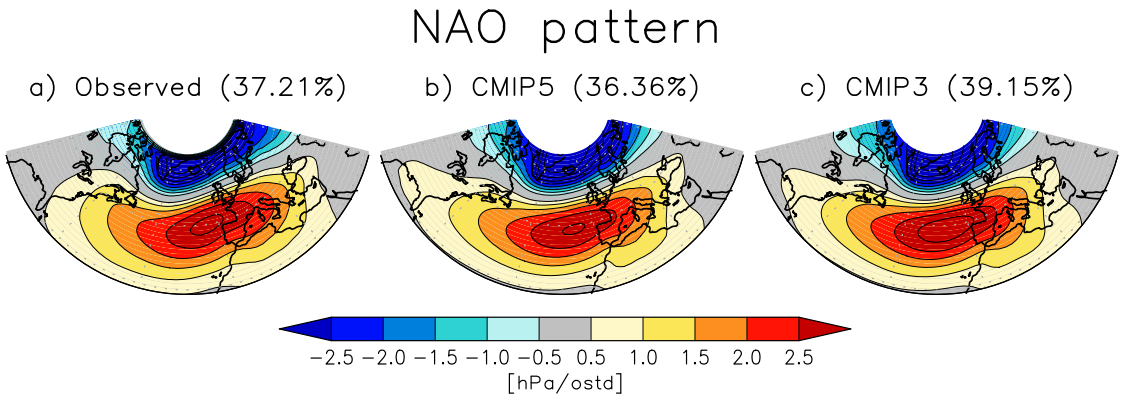


Figure 3.6: (a) Spatial pattern of the observed NAO defined as the first PC of the winter Trenberth SLP field (interpolated to a 2.5×2.5 grid) in the 90°W-30°E/40°N-70°N domain for the 1902-2004 interval. The pattern is displayed in terms of the detrended regression of the SLP field onto the normalized NAO index. In brackets, the percentage of SLP variance explained by the associated EOF. (b) Mean NAO spatial pattern of the CMIP5 multi-model ensemble. See text for a complete description of how the NAO in climate models is calculated. The pattern is displayed in terms of the multi-model weighted average of the SLP-NAO regression maps obtained with single simulations, standardized in terms of the observed standard deviation. (c) The same for the NAO in the CMIP3 multi-model ensemble.

the same 90°W-30°E/40°N-70°N domain). As a result, for each time we obtain a projection coefficient that defines the magnitude of the NAO index at that particular time. The resulting time series (or pseudo-PC), therefore, reflects the evolution of the NAO index in that model. Instead, we could have defined the NAO index of a given model as the first PC of its own SLP field. This method, however, might have led us to wrong conclusions: possible discrepancies among climate models regarding the impact of the NAO on Mediterranean precipitation could be due to the fact that we would be defining the NAO index as the time evolution of a different spatial pattern with different dynamical mechanisms associated. To illustrate this, Davini and Cagnazzo (2013) document the misrepresentation of the NAO by some CMIP5 models when estimated as the first EOF of geopotential height at 500 hPa over the north Atlantic. They describe the inability these EOF to reproduce the observed cyclonic Rossby wave breaking over Greenland associated with the negative phase of the observed NAO.

To project the observed EOF onto the SLP field of climate models, both variables must have the same grid. For this reason, in Chap. 5 we will re-calculate the observed NAO using Trenberth data interpolated to a 2.5×2.5 grid (the grid of SLP in our climate simulations). We also reduce the period used to calculate the observed NAO to the historical period of climate models (1902-2004). This "new" observed NAO is displayed in Fig. 3.6a in terms of detrended regression of the SLP field onto the principal component. We see that it is very similar to the pattern of the observed NAO used in Chap. 4 (see Fig. 3.5).

The NAO index in climate models is standardized in terms of the standard deviation of the detrended observed principal component (ostd). If we had normalized the NAO index of each model using its own standard deviation instead, the magnitude of the NAO index for different climate models would not be comparable.

In models with several runs, to obtain the model mean NAO index we calculate the pseudo-PC of each ensemble member and then average. Similarly, to obtain the NAO index for the multi-model ensemble means we average all model means. To illustrate the mean geographical pattern of anomalies associated with the NAO in the multi-model ensembles, we calculate for each individual simulation the detrended regression of its SLP field onto its NAO index. The average of the resulting regression maps (calculated first for each model) is presented in Figs. 3.6b,c for the CMIP5 and the CMIP3 multi-model ensembles.

Recently, Gonzalez-Reviriego et al. (2014) proposed to use the method of partial least squares to evaluate teleconnection based indices in climate models. This method looks for patterns in model data that maximizes the covariance between the model field and an observed reference field. In case of the NAO pattern their results generally agree with ours.

Chapter

FOUR

Recent observed precipitation changes in the northern Mediterranean

As discussed in the introduction, the Mediterranean region is one of the areas where climate models consistently project strong drying for the end of the twenty-first century as a response to the anthropogenic increase in the concentration of GHG gases in the atmosphere [Christensen et al., 2013; Diffenbaugh and Giorgi, 2012]. But, are these projections credible? How consistent are they with recent observed trends?

The goal of this chapter is to assess which seasons and regions in the northern Mediterranean exhibit robust precipitation decreases in the last decades and check whether they are consistent with the seasonal and geographical distribution of the projected drying signal. We consider a precipitation change as robust if it is not explainable only with internal variability. Although assessing the robustness of observed precipitation changes is not straightforward because of the strong multi-decadal variability of the northern Mediterranean precipitation, it is essential if we want to use such changes as evidence that the projected drying signal is already appreciable. Otherwise, we could be incorrectly basing our confidence in the drying signal on fluctuations of precipitation associated only with natural noise.

One way to assess whether an observed precipitation change is robust or not is to test if its magnitude is highly sensitive to the time period for which it is estimated, since a persistent signal should be detected in a wide range of consecutive time intervals. To that end, we present a methodology based on the analysis of two two-dimensional diagrams. In the first one we represent the linear change between each possible pair of years within the observational record, while in the second one we plot the mean anomaly of each possible sub-interval of the total record, regardless of its length, with respect to a reference mean.

Although many authors have reported significant precipitation changes in the Mediterranean region since the mid-twentieth century [Piervitali et al., 1998; Pal et al., 2004; Giorgi and Lionello, 2008; Hoerling et al., 2012; van Haren et al., 2012], most of them do not assess the robustness of their results. For instance, Pal et al. (2004) base their conclusion that the projected summer drying signal is already appreciable on the result that June-August precipitation decreased between the 1976-2000 mean and the 1951-1975 mean. Since they do not verify whether the precipitation reduction is still visible when averaging over other time intervals, however, their result does not contribute to increase our confidence in the drying signal.

Because the NAO is the main driver of precipitation variability in the northern Mediterranean region [Hurrell et al., 2003], detecting a signal in the NAO could explain the recent observed precipitation changes. However, the fact that the positive NAO trend observed in the last decades of the twentieth century [e.g. Hurrell, 1995; Gillett, 2005] has reversed in the first decade of the twenty-first century makes it difficult to assess if there is an appreciable signal superimposed on the multi-decadal internal variability of the NAO. In the second part of this chapter we use the two-dimensional diagrams to examine whether the observed changes in the NAO index are robust or not.

The chapter begins with a detailed explanation of the methodology used to assess whether precipitation changes are sensitive to the choice of time interval. Then we examine precipitation changes starting with a wide perspective (areal averages and annual means) and then narrowing down to a closer one (analysing seasonal means and the spatial distribution). We will initially be using the E-OBS data set [Haylock et al., 2008], covering the 1950-2014 interval, because it is the best data set available over Europe due to its high spatial resolution and its optimal interpolation methods (see Sec. 2.1.1). Then, we change to GPCC data [Schneider et al., 2011; Becker et al., 2013; Schneider et al., 2014], which begin in 1901, to place the results for the last decades in a century-long context. CRU [Harris et al., 2014; Mitchell and Jones, 2005] and UDel [Matsuura and Willmott, 2012] data are used as well to add robustness to our results. In the last part of the chapter, we focus on the impact of the NAO on Mediterranean precipitation. First, we examine the sensitivity of the observed changes in the NAO index to the choice of time interval. Then, we evaluate the extent to which the recent observed precipitation changes are explainable by the NAO.

4.1 Two-dimensional multi-period diagrams

To estimate the sensitivity of observed changes to the choice of time interval and assess whether observational series are consistent with a persistent trend we introduce two two-dimensional diagrams. Prior to that, to illustrate how one must interpret data plotted in

these kinds of diagram, we introduce three synthetic precipitation time series. The first is a series with a pure negative trend of −1.5 (mm/day)/century (Fig. 4.1a). The second one is created with random values normally distributed, with null mean, 1.5 times the standard deviation of the previous series, and no trend (Fig. 4.1b). Finally, we sum both series in order to create a more realistic situation, in which a persistent moderate long-term trend is superimposed on random internal variability of comparable variance (Fig. 4.1c). Each time series has been represented in terms of anomalies with respect to the 1961-1990 mean, as it is a standard climate normal of the World Meteorological Organization (WMO) [Arguez and Vose, 2011].

4.1.1 2DLC diagram

In order to evaluate the sensitivity of trend estimates to the choice of time period, Liebmann et al. (2010) designed a two-dimensional parameter diagram in which they plot the value of all possible trends of any length in a given time series as a function of the last year and the length of the segment. This diagram allows them to assess whether a result found for a particular time interval is valid for a wide range of adjacent periods. Rather than absolute trends they use the cumulative linear change (i.e. the trend multiplied by the length of its period) because, in this way, short non-persistent changes lose strength and persistent trends are emphasized. Instead of representing the last year of the segment against its length, here we choose to plot linear changes as a function of the initial and the final

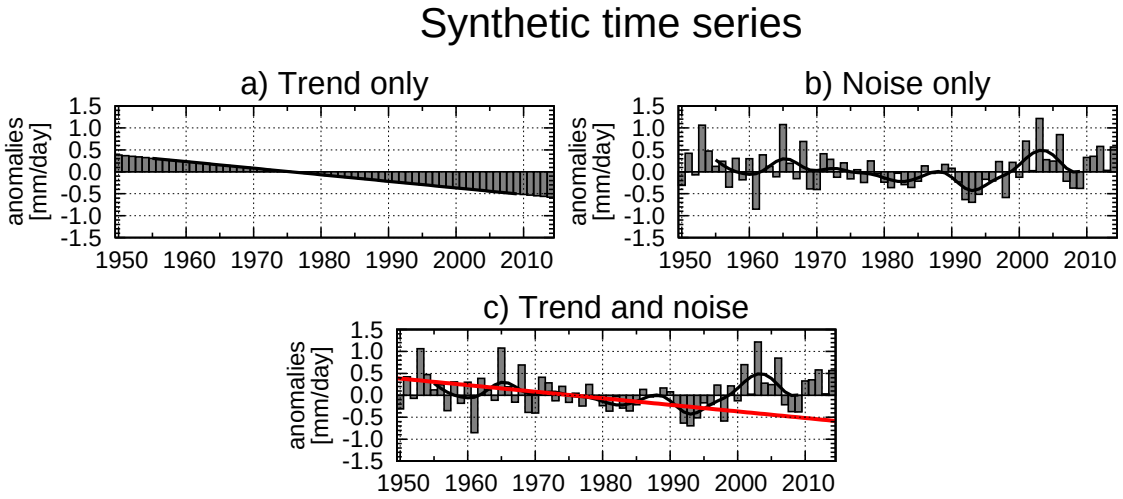


Figure 4.1: Synthetic time series created to exemplify the diagrams used in this chapter. (a) Pure trend of −1.5 (mm/day)/century. (b) Normally distributed noise with standard deviation 1.5 times larger than that of the previous series and no trend. (c) Sum of the other two time series. Anomalies are relative to the 1961-1990 mean. Black curves represent data filtered using a low-pass symmetric non-recursive filter with a cut-off period of 10 years. The red line in (c) represents the linear trend estimated for the entire record.

year of the period, as we find it eases the interpretation of our results. As a consequence, the diagonal broken lines (labelled on top) indicate time periods with the same length. Other than the differences in axis, our methodology and that of Liebmann et al. (2010) are totally equivalent. This diagram also allows us to represent which changes are statistically significant at a given confidence threshold (usually 95%) evaluated with a Monte Carlo test with 1000 iterations (see Sec. 3.1.6 for a complete explanation of statistical significance estimation techniques). Periods shorter than 10 years are hatched as they represent inter-annual variability and we are interested only in changes at decadal to multi-decadal scales. From now on, we will refer to this kind of diagram as *two-dimensional linear change* (2DLC) diagram.

In the 2DLC diagram for the synthetic series with a pure trend (Fig. 4.2a), all changes are negative and their magnitude increase with the length of period for any given initial of final year (the longest time interval is located in the upper left corner). As one may expect, all changes are statistically significant at a 95% confidence level. In the diagram for the time series of pure noise (Fig. 4.2b), instead, changes does not exhibit a dominant sign and their magnitude and statistical significance is highly sensitive to the choice of time interval. Note that, given that we are testing the statistical significance of the change at a 95% confidence level, the series of pure noise is still expected to exhibit a 5% of statistically significant changes only by chance. Finally, in the diagram for the hybrid series (Fig. 4.2c), the negative changes dominate over the positive ones, and the magnitude of the change for a given initial or final year generally increases with length but not as monotonously as in the first case due to the random noise. Even though the largest changes are not found for the longest time intervals (but for intervals beginning around 1950 and ending in the 1990s, due to a succession of wet years in the 2000s), changes for intervals longer than 50 years are generally stronger than for intervals about 30-year long.

Therefore, one indication that observed precipitation is consistent with a persistent drying trend would be to obtain a similar-looking diagram using observational data. In it, negative changes should dominate over positive changes and the longest intervals should exhibit strong and statistically significant changes for a wide range of adjacent periods (with some modulations caused by internal variability). Instead, a change for a given time interval, even if statistically significant, will be more consistent with noise only if changes for adjacent longer periods are not statistically significant or decrease substantially in magnitude.

4.1.2 2DMDA diagram

In order to assess how anomalous a particular time interval is compared with a longer reference period, we have devised another kind of two-dimensional diagram. In this new diagram, we plot every possible decadal to multi-decadal anomaly of any length, defined

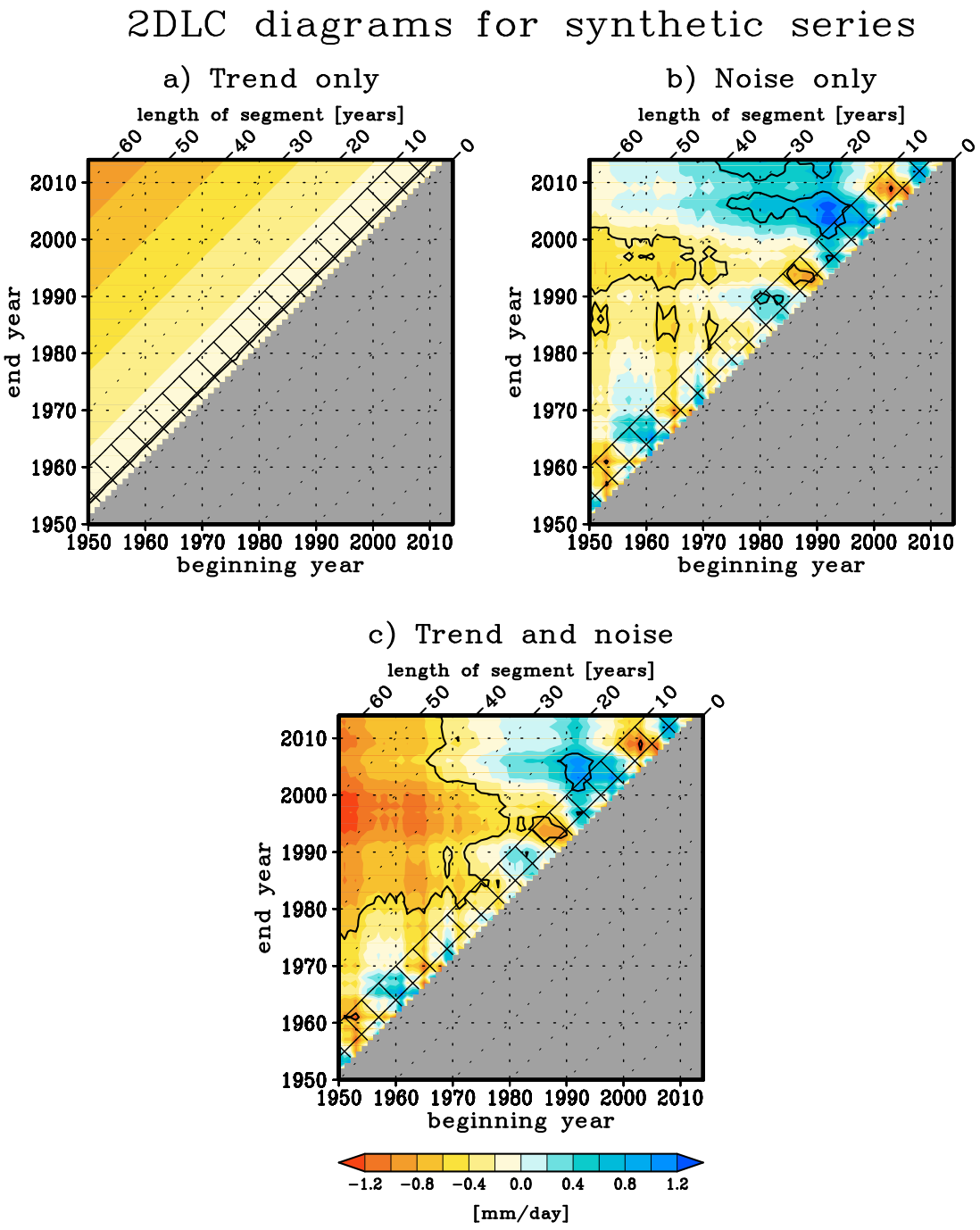


Figure 4.2: 2DLC diagram for the synthetic time series in Fig. 4.1. All cumulative linear changes between any pair of years within the 1950-2014 period are represented. Change is calculated as the multiplication of a given linear trend and the length of the corresponding period. The horizontal axis indicates the beginning year and the vertical axis the end year. Diagonal broken lines (labelled on top) indicate intervals with the same length. Contours indicate statistically significant changes at a 95% confidence level. Results for periods shorter than 10 years have been hatched as they correspond to inter-annual variability.

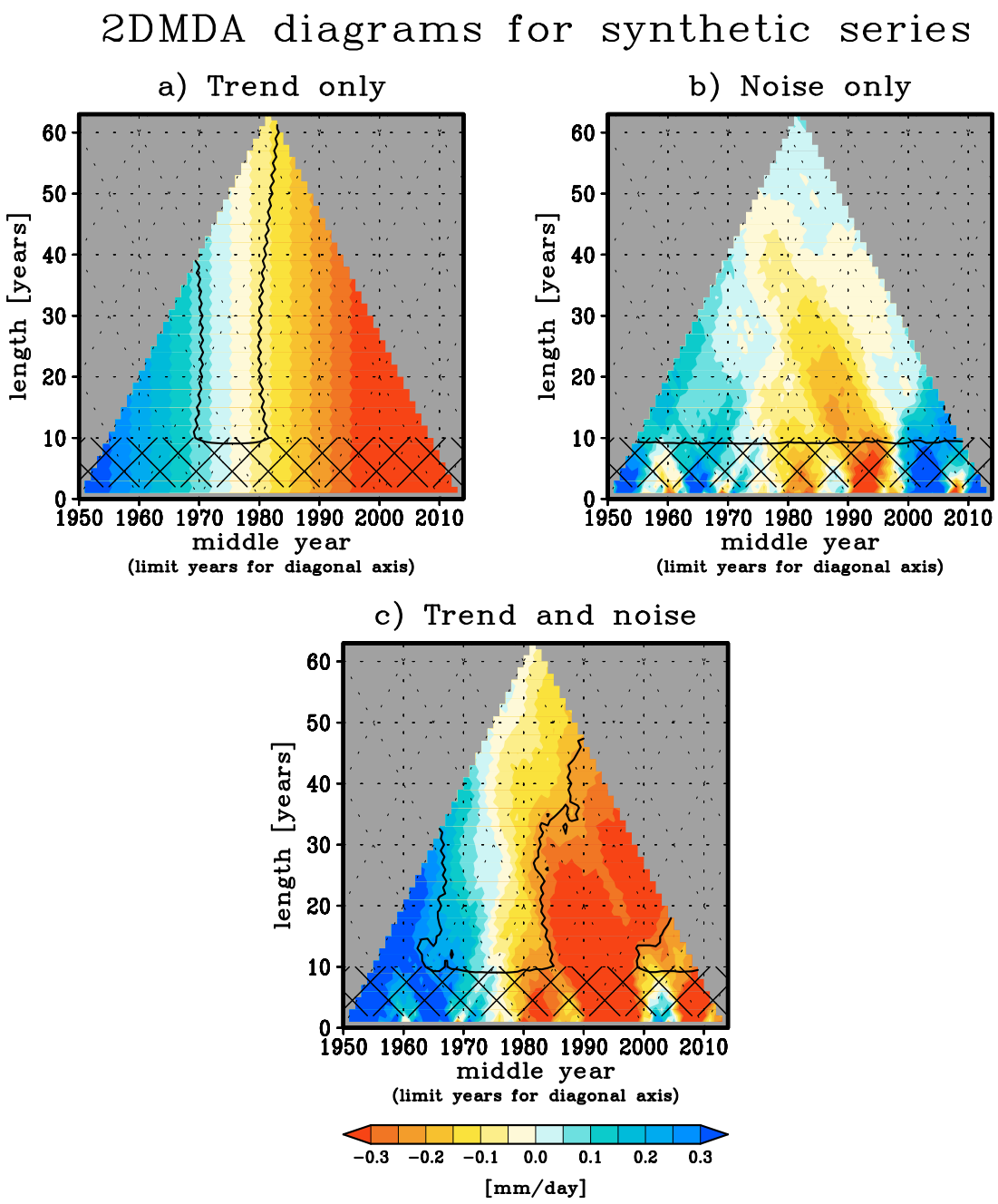


Figure 4.3: 2DMDA diagram for the synthetic time series presented in Fig. 4.1. Every annual to multi-decadal anomaly of any length within the 1950-2014 period is represented. Anomalies are defined as the difference between the mean precipitation in a given time interval and the climatological mean (1961-1990). The horizontal axis indicates the middle year of each interval and the vertical axis to its length. Diagonal broken lines indicate the first (top right to lower left) and the last (top left to lower right) year of each interval. Contours indicate statistically significant changes at a 95% confidence level. Results for periods shorter than 10 years have been hatched as they correspond to inter-annual variability.

as the average precipitation in a given time segment minus the mean reference climatology (in this case 1961-1990). Therefore, each horizontal line in this diagram represents the running mean of the original series with a different length of window. Furthermore, following the diagonal broken lines we find the initial (top right to lower left) and the final (top left to lower right) year of a particular segment. To test the statistical significance of the multi-decadal anomalies (at a 95% confidence level) we use the non-parametric Wilcoxon–Mann–Whitney test (see Sec. 3.1.6). As in the 2DLC diagram, time periods shorter than 10 years are not considered as they are dominated by short inter-annual variability. We will refer to this kind of diagram as *two-dimensional decadal to multi-decadal anomaly* (2DMDA) diagram.

In case of the pure negative trend with no short term variability (Fig. 4.3a) we find a uniform transition from wet (blue) to dry (red) anomalies for all time scales, with statistically significant decadal to multi-decadal anomalies found for intervals distant from the reference period 1961-1990. Instead, and consistent with the fact that random positive and negative anomalies compensate, weak non-significant decadal to multi-decadal anomalies of either sign characterize the pure noise series (Fig. 4.3b). Finally, in the realistic series (Fig. 4.3c), the transition from wet to dry anomalies is visible at all time scales but modulated by the noise. Statistically significant anomalies, however, are less widespread than in the case of a pure trend.

4.2 Changes in annual mean precipitation

The aim of this section is to assess whether a persistent change in northern Mediterranean annual mean precipitation is appreciable above internal variability during the last decades. In this section we use observational E-OBS data [Haylock et al., 2008], beginning in 1950, as it is the most reliable data base available and it is updated at a monthly basis. Results, however, are reproducible with other data sets.

As explained in Sec. 4.1, some conclusions on precipitation changes can be highly sensitive to slight variations of time interval. To illustrate this, we first estimate annual precipitation trends for two periods: a cherry-picked time period which we know maximizes trends (1960-2007; Fig. 4.4a), and the entire length of the E-OBS record (1950-2014; Fig. 4.4b). The trend pattern for the first interval indicates that precipitation has decreased in most parts of the basin. Statistically significant trends are even present in the Iberian and Italian peninsulas and in the south Balkans. Yet, when choosing the full period of record, all trends lose strength and the number of local statistically significant trends is greatly reduced.

In order to visualize the evolution of annual precipitation we average over the NMed region. In Fig. 4.5a we represent anomalies for the 1950-2014 period relative to the stan-

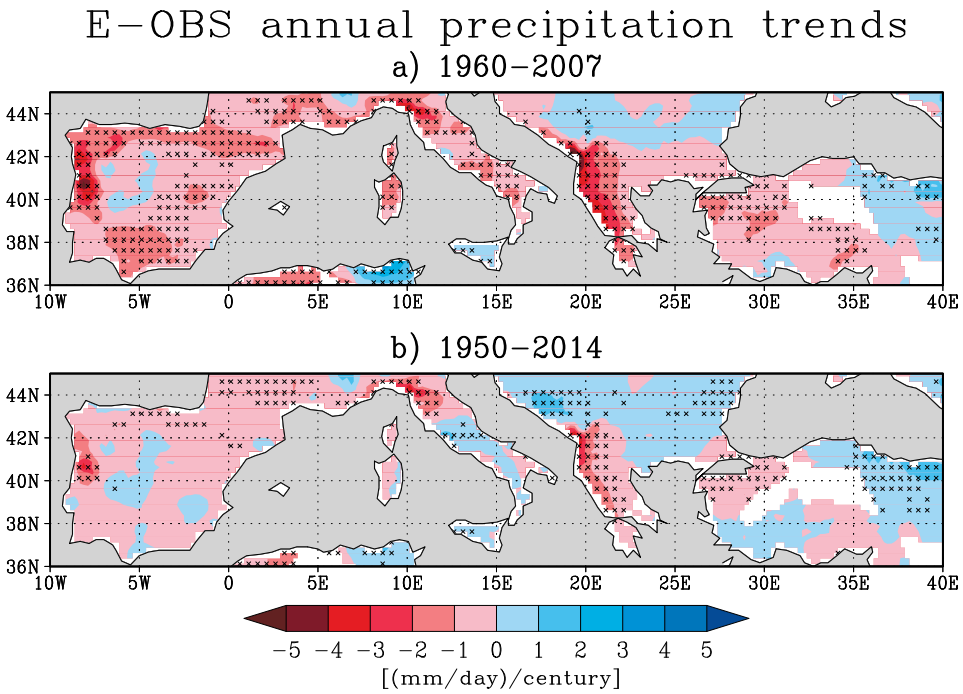


Figure 4.4: Linear trends of E-OBS annual mean precipitation in the northern Mediterranean region for the intervals 1960-2007 (a) and 1950-2014 (b). Crosses indicate regions with statistical-significant trends at a 95% confidence level. White regions indicate grid-points with more than 30% of missing values in its time series.

dard climate normal of the WMO 1961-1990 [Arguez and Vose, 2011]. At first sight we see that area-averaged precipitation exhibits a change in frequency of wet and dry years around 1980. While in the first half of the record (1950-1982) the number of wet years is almost the same as the number of dry years (17 wet vs. 16 dry), in the second half of the record (1982-2014) dry years overwhelmingly predominate (8 wet vs. 25 dry). In this second interval, moreover, we find dry periods lasting up to nine consecutive years (1988-1995), whereas before the maximum number of consecutive dry years is 3. On the other hand, the last year of the series (2014) is the wettest one in the entire record, and 3 of the 6 wettest years (1996, 2010 and 2014) are found in the second half of the record. Moreover, analysing decadal means (Fig. 4.5b), we see that the first decade of the record (1950s) is in fact drier than the last one (2000s), and that from the 1980s onwards each decade has been less dry than the previous one. In fact, the persistent 1982-1995 dry interval (with only one wet year) makes the 1980s the driest decade, followed by the 1990s. Based on this, only by exploring the time series, one cannot affirm with confidence that NMed annual mean precipitation exhibits a persistent drying trend.

The time series, however, does help us understand the differences in the pattern of trends in Fig. 4.4a (1960-2007 and 1950-2014). While the first trend begins in one of the wettest years on record and ends in a dry year (blue line in Fig. 4.5a), the second trend is much weaker because it begins in the driest year and ends in the wettest one (red line). This

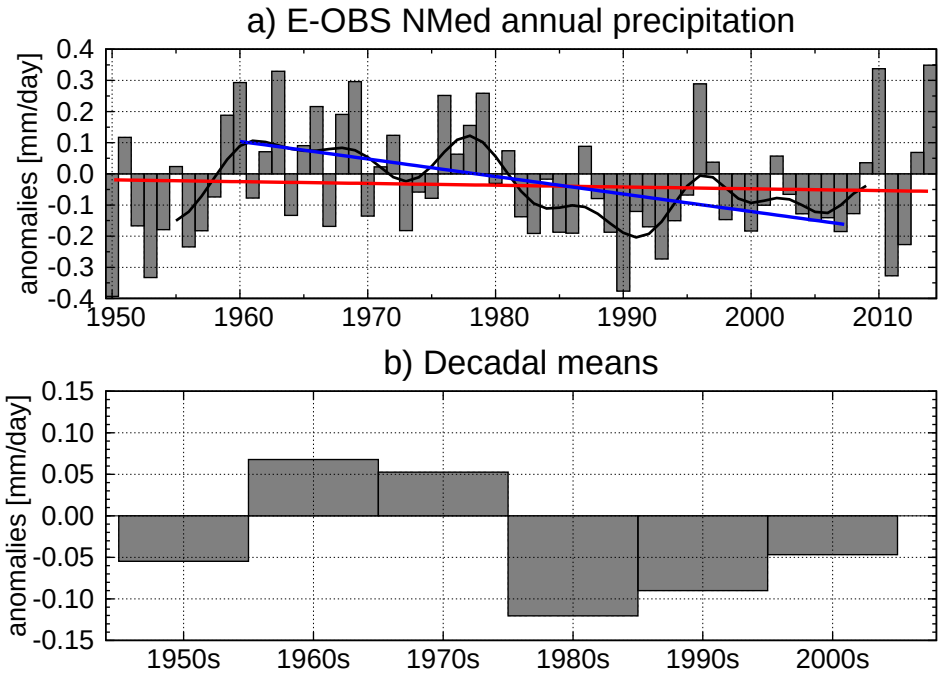


Figure 4.5: (a) Anomalies of E-OBS NMed annual mean precipitation relative to the 1961-1990 mean. The black curve represents filtered data. The blue line represents the linear fit in the 1960-2007 interval. The red line represents the linear fit in the 1950-2014 period. (b) Decadal averages.

illustrates the importance of checking whether trend estimations are sensitive to the choice of time interval if they are to be used as a tools to detect persistent signals and emphasizes the relevance of our two-dimensional diagrams.

In the 2DMDA diagram for NMed annual mean precipitation (Fig. 4.6), the transition towards dry conditions is visible for all time scales up to 35-year long. Wet non-significant multi-decadal anomalies are found before 1980 followed by a strong dry signal after 1980. In fact, Fig. 4.6 bears a good qualitative resemblance to the diagram for the realistic synthetic time series in Fig. 4.3c except for the lack of statistical significance of the wet anomalies. The two halves of the record considered in earlier paragraphs are highlighted with circles in Fig. 4.6. While the 1950-1982 multi-decadal anomaly (red circle) is not statistically significant, the 1982-2014 interval (blue circle) is located in a wide region of statistically significant dry anomalies. The conclusion that recent decades have been significantly dry compared to the reference mean is thus not sensitive to the exact choice of time period. Because the estimation of the statistical significance of anomalies, however, is sensitive to the selected reference period, in Sec. 4.6 we use century-long data to better evaluate the statistical significance of our multi-decadal anomalies.

We can now estimate for which time periods the linear decrease in NMed annual precipitation is statistically significant by analysing the 2DLC diagram (Fig. 4.7). First, we see that negative changes dominate over positive changes as in the synthetic time series with a

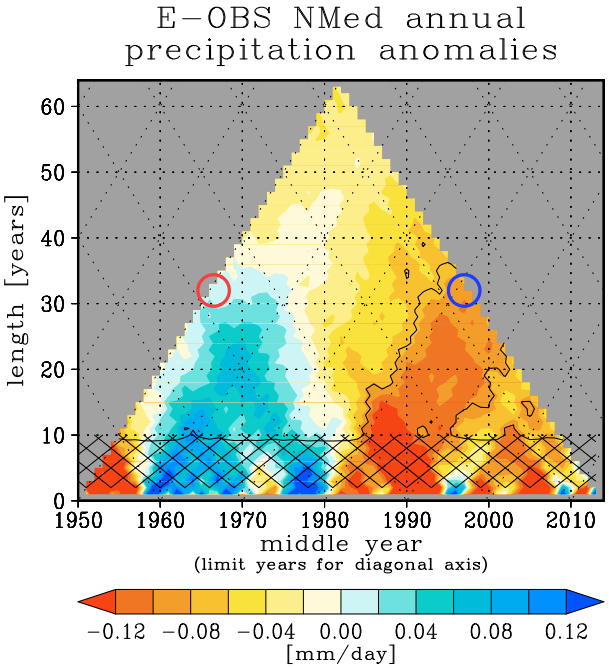


Figure 4.6: 2DMDA diagram for the E-OBS NMed annual mean precipitation. Anomalies are relative to the 1961-1990 mean. Contours indicate statistically significant changes at a 95% confidence level. For a complete explanation of the 2DMDA diagram see Sec. 4.1.2. The red circle indicates the multi-decadal anomaly of the first half of the record (1950-1982) and the blue circle the one of the second half (1982-2014).

pure trend superimposed on internal noise (see Fig. 4.2c). Focusing on periods longer than 30 years, however, we see that the diagram for NMed annual precipitation differs from the one for the synthetic time series. First, the longest time interval (and all those beginning in 1950-1955) does not exhibit statistically significant negative changes due to the fact that the first decade of the record was relatively dry (see Fig. 4.5b). Second, the magnitude of the change does not increase as clearly with the length of period. Instead, the strongest statistically significant changes are found for some time intervals beginning in the very wettest period of the record (the 1960s) and for some time intervals ending in the driest period of the record (the early 1990s). Focusing on the cherry-picked time interval of Fig. 4.4 (1960-2007, blue circle) we see that, although it indeed exhibits a statistically significant negative change, the condition that changes must remain statistically significant and generally increase its magnitude as with increase the interval length is not met.

In conclusion, evidence for a persistent trend in NMed annual precipitation is ambiguous. One way to clarify this issue would be to increase the signal-to-noise ratio of any potential trend by isolating the months in which it is strongest. Alternatively, we could increase the signal-to-noise ratio by taking longer series: while changes associated with a persistent, albeit non-monotonic, signal accumulate with time, changes associated with internal variability compensate each other.

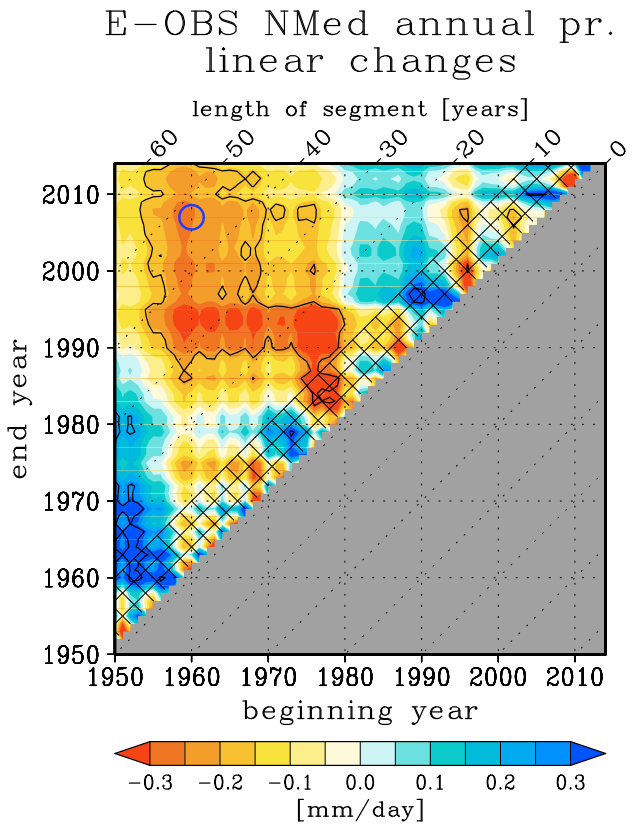


Figure 4.7: 2DLC diagram for E-OBS NMed annual mean precipitation. Contours indicate statistically significant changes at a 95% confidence level. For a complete explanation of the 2DLC diagram see Sec. 4.1.1. The blue circle highlights the change in the 1960-2007 interval.

4.3 Multi-month precipitation changes

Refining our analysis in order to allow any potential drying signal to emerge more clearly above internal variability, we analyse the seasonality of the NMed precipitation change so that we may identify the particular time of the year in which the decrease took place. We therefore compute the 2DMDA diagram for each month and show them together in Fig. 4.8. This methodology facilitates the visualisation of the evolution of precipitation in all months and its comparison; if we had represented the twelve monthly time series separately, we would not be able to see whether changes are robust across different time scales (length of segment). Periods shorter than 10 years are omitted because they are dominated by high-frequency internal noise, and periods longer than 50 years have been removed as no significant anomalies emerge when using intervals with a length close to that of the total record.

Figure 4.8 illustrates that the only robust, statistically significant, multi-month signal is found in winter (see the box). From December to March, the last decades exhibit a pronounced transition towards dry conditions regardless of the time scale. January and, particularly, February exhibit the strongest signal, with statistically significant multi-decadal

anomalies in the last decades for all lengths of interval. The multi-month nature of the decrease in winter precipitation is unique, as the only other statistically significant negative signal is found in June but, apart from being substantially weaker (consistent with the lower climatological rainfall), it is not present in adjacent months. September, while its last two decades are significantly wet, does not exhibit a continuous change but a succession of wet and dry periods about twenty-year long. Therefore, annual rainfall reductions are clearly due to winter decreases, partially compensated for by increases in autumn.

It is noticeable that in summer, the season for which the strongest decreases are projected for the future in climate models [Collins et al., 2013], there are no robust signals of change. Although some authors have reported summer precipitation decreases [Pal et al., 2004; Giorgi and Lionello, 2008], only in June is the signal statistically significant and only for the last 15 years. Therefore, alleged changes in summer are not robust and merely result from having unintentionally chosen time intervals beginning in anomalously wet years and

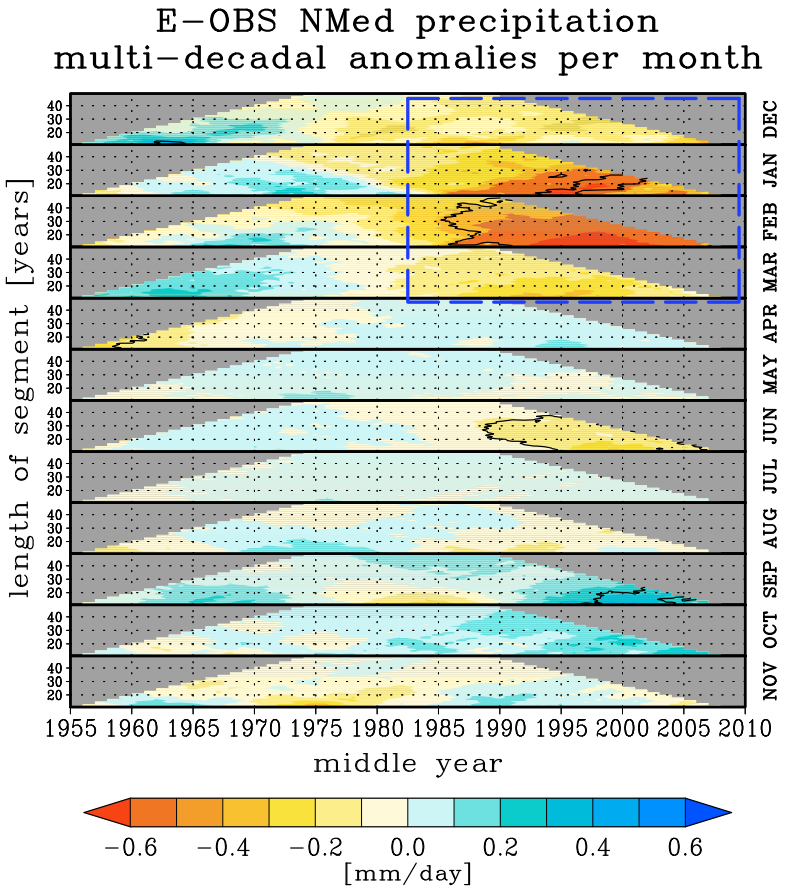


Figure 4.8: 2DMDA diagram for E-OBS precipitation for each month averaged over the NMed region. Anomalies are relative to the 1961-1990 mean. The horizontal axis indicates the middle year of the each interval and the vertical axis (in each panel) its length. Contours indicate statistically significant changes at a 95% confidence level. The box highlights the recent dry decades observed in winter months. For a complete explanation of the 2DMDA diagram see Sec. 4.1.2.

ending in anomalously dry years.

Figure 4.8 also illustrates how the use of traditional seasons can be misleading. For instance, the spring (March-May) precipitation reduction found between 1961-1980 and 1981-2000 by Giorgi and Lionello (2008) is not the result of an actual season-long behaviour but is solely due to a (non statistically significant) decrease in March. The two-dimensional diagram also shows that the use of the wet and dry halves of the year is not appropriate for our purposes, as averaging six months together dilutes strong signals that are only present in a shorter set of months. For instance, the decreases found by Hoerling et al. (2012) in the wet season (November-April) might be stronger and more statistically significant if the chosen period did not contain April and November. Likewise, the precipitation changes in the warm season found across the basin by Kelley et al. (2012a) (May-October) and van Haren et al. (2012) (April-September) are weak and without a dominant sign because they are not the result of a multi-month signal. Instead, they are a combination of decreases in May, June and August counterbalanced by increases in April, July and especially September and October. A somewhat better approach was used by Barkhordarian et al. (2013), who choose every possible overlapping 3-month season along the year. Using this methodology they show that winter changes are not sensitive to the specific winter definition as each possible 3-month season containing months from December to May exhibits negative changes in the 1966-2005 period. Also consistent with our findings, they find no robust signal in summer as half of the 3-month seasons containing July or August exhibit decreases and the other half exhibit increases. Examining precipitation anomalies in our multi-diagram, however, has an added value because it does not restrict the analysis to 3-month seasons and, more importantly, allows us to visualize precipitation changes without having to choose a time interval a priori.

4.4 Changes in winter precipitation

Having determined that the only season in which robust precipitation decreases have occurred is winter (based on Fig. 4.8), using only winter months we expect to find a drying signal stronger than at annual scale. Therefore, we repeat the multi-scale analysis in order to assess whether the evolution of NMed winter precipitation in the last decades is consistent with what one would expect if a persistent trend was superimposed on internal variability. From now on, based on the results of Fig. 4.8, we define winter as the average from December to March, but results presented here are reproducible (and enhanced) using only "high winter" months (January and February averaged, not shown).

The contrast between rainfall in the first and the second half of the record is indeed more evident in the winter time series (Fig. 4.9a) than at the annual scale, mostly due to the very dry 1989-1993 interval, which contains 4 out of the 6 driest winters in the entire

series. Looking at the evolution of decadal winter averages (Fig. 4.9b) we see that, unlike annual means, the last three decades are considerably drier than the first three decades (see Fig. 4.5b). In particular, the first decade of the record is not as dry as for the annual mean and the 1990s now appear to be the driest decade (whereas when considering annual precipitation the driest decade was the 1980s).

To assess the robustness of decadal to multi-decadal drying signals to variations in time interval, we again use the 2DMDA diagram (Fig. 4.10). The pattern of multi-decadal anomalies for winter is similar to that for the annual mean (see Fig. 4.6), except for the stronger amplitude (note the change of scale) and the greater number of intervals exhibiting statistical significant anomalies. In fact, this diagram highly resembles the one for the synthetic time series with a persistent trend superimposed on internal noise (see Fig. 4.3c). The influence of the noise is visible from the fact that anomalies beginning after 2000 are not statistically significant, and from the fact that wet anomalies for the first decades are not statistically significant either. The existence of a persistent drying signal is visible from the fact that, for interval lengths longer than 15 years, all multi-decadal anomalies centred in years after the late 1980s are statistically significant.

In winter the persistent signal is more clear compared to the ambiguous signal found for the annual mean. Because the winter 2DLC diagram (Fig. 4.11) resembles that for the *realistic* synthetic time series, we conclude that a drying signal, albeit weak, is detectable

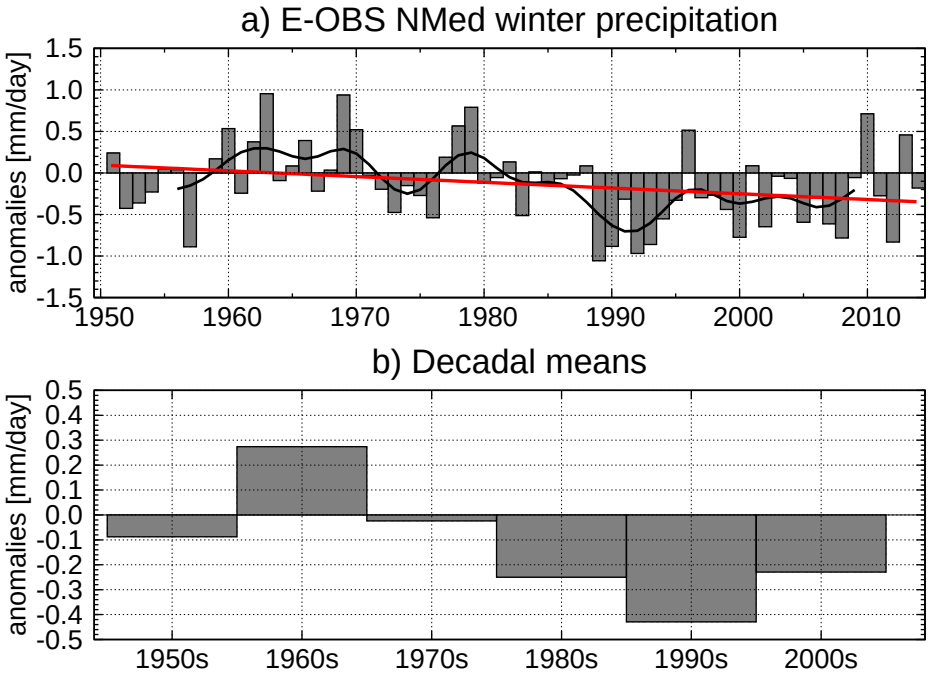


Figure 4.9: (a) Anomalies of E-OBS NMed winter precipitation relative to the 1961-1990 mean. The black curve represents filtered data. The red line represents the linear fit in the 1951-2014 period. (b) Decadal averages.

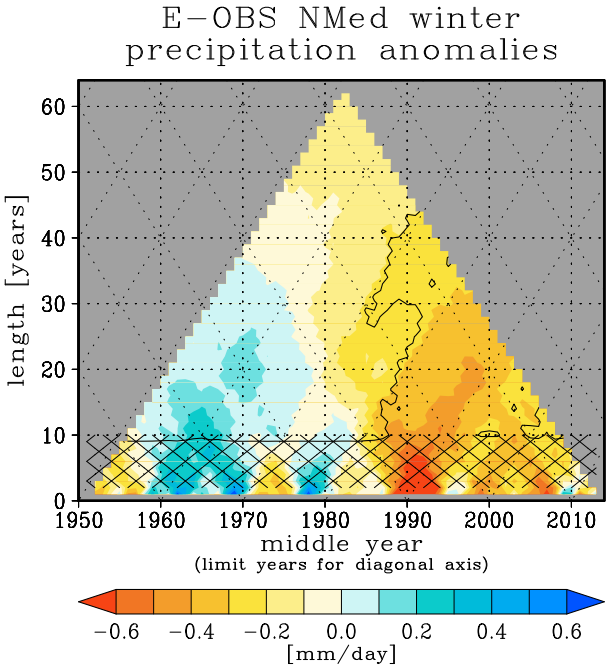


Figure 4.10: 2DMDA diagram for the E-OBS NMed winter precipitation. Anomalies are relative to the 1961-1990 mean. Contours indicate statistically significant changes at a 95% confidence level. For a complete explanation of the 2DMDA diagram see Sec. 4.1.2.

in NMed winter precipitation. Still, internal variability obscures this signal, even for winter only means. First, the strongest changes are not exactly found for the longest time intervals. Second, the magnitude does not monotonously increase with the length of time period for a given initial or final year. In particular, changes beginning in the first decade are still weaker than those beginning around 1960. Lastly, the largest changes are found for time intervals ending in the persistent dry interval around 1990 (not the recent years).

In conclusion, from 1980 onwards winters in the northern Mediterranean region have been anomalously dry. Moreover, our analysis illustrates that the evolution of NMed winter precipitation from 1950 onwards is consistent with what one would expect if a persistent trend was superimposed on multi-decadal internal variability.

4.4.1 Century-long overview of winter changes

Is the drying signal detectable already in the first half of the twentieth century? Is it evident when we compare with a longer reference period? To which extent has recent NMed precipitation been anomalously low in a century-long context? To answer these questions we examine GPCC data, which starts in 1901, keeping in mind that the scarce amount of gauge stations prior to 1950 makes data for the first half of the twentieth century less reliable (see Sec. 2.1.1). We note that all results found here are reproducible with CRU (1901-2013) and UDel (1900-2010) data sets; the correlation of NMed area-averaged winter precipita-

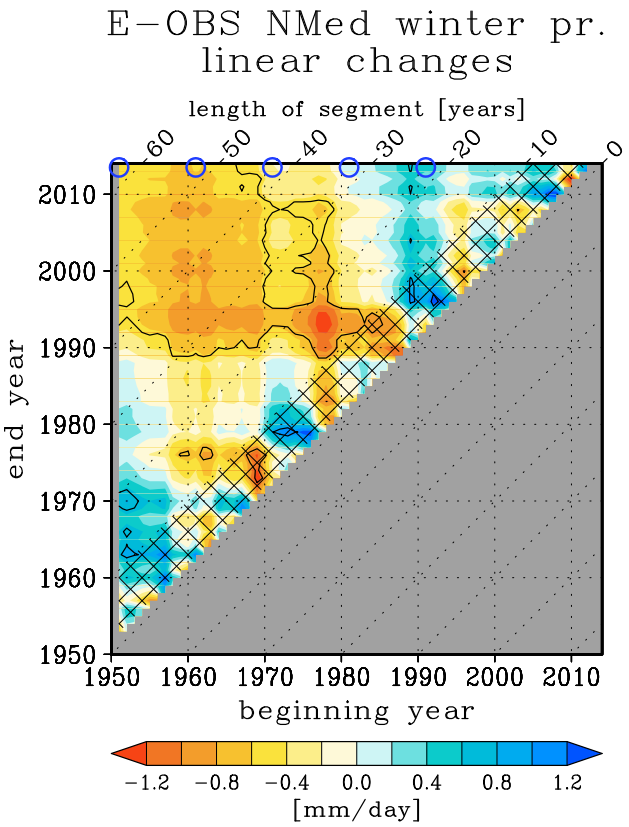


Figure 4.11: 2DLC diagram for E-OBS NMed winter precipitation. Contours indicate statistically significant changes at a 95% confidence level. For a complete explanation of the 2DLC diagram see Sec. 4.1.1 The blue circles indicate time intervals ending in 2014 and beginning every 10 years from 1951 and 1991 (used in Fig. 4.16).

tion between GPCC and those data sets in the common period is $r = 0.99$ in both cases. The correlation with E-OBS is $r = 0.96$.

To begin with, we analyse the century-long time series of GPCC NMed winter precipitation anomalies (Fig. 4.12a). Now, the average over the entire period is used as a climatological mean because, with such a long record, the unperturbed mean state of reference precipitation is better captured with the total mean than with the mean over the short 30-year period used before. Viewed in this longer context, the change around 1980 is even more clear. Even though in the 1981-2010 interval the frequency of dry years is not much larger than the frequency of wet years (17 dry vs. 13 wet), most of the dry anomalies are very large (in comparison with the entire period). Not only is such an accumulation of strong dry anomalies unprecedented, but most of the wet anomalies found in this interval are weak (with only few exceptions). The evolution of the decadal means (Fig. 4.12b) indicates that, until the 1960s, NMed winter precipitation was dominated by decadal variability. The reduction of winter precipitation between the 1960s and the 1990s, instead, is several times significantly stronger than the magnitude of the decadal variability observed in previous decades, partially due to the fact that in the 1989-1995 interval we find again

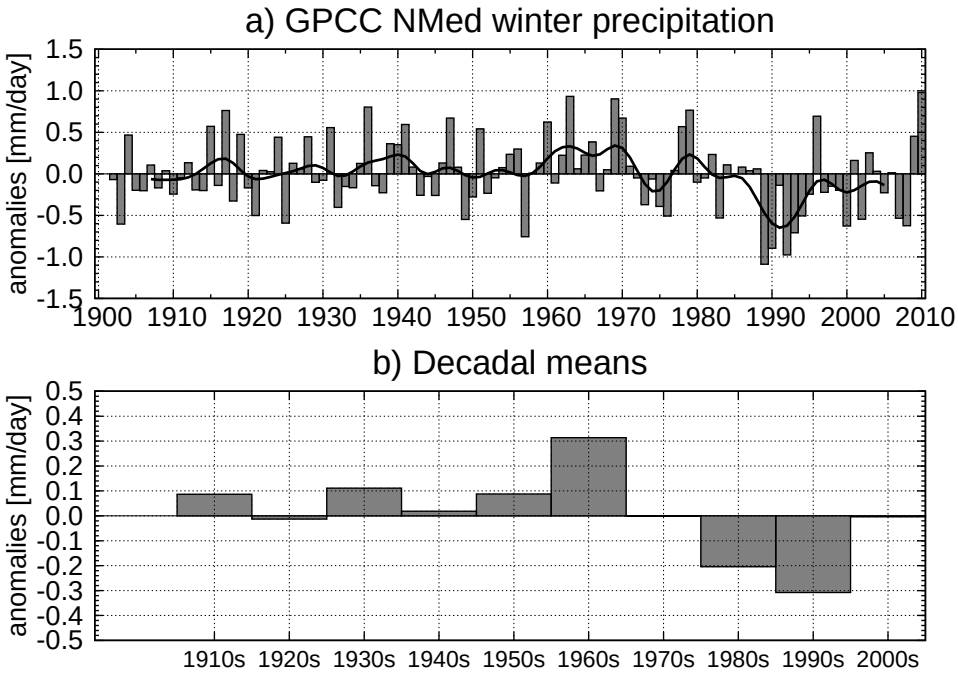


Figure 4.12: (a) Anomalies of GPCC NMed winter precipitation relative to the entire record (1902-2010). The black curve represents filtered data. (b) Decadal averages.

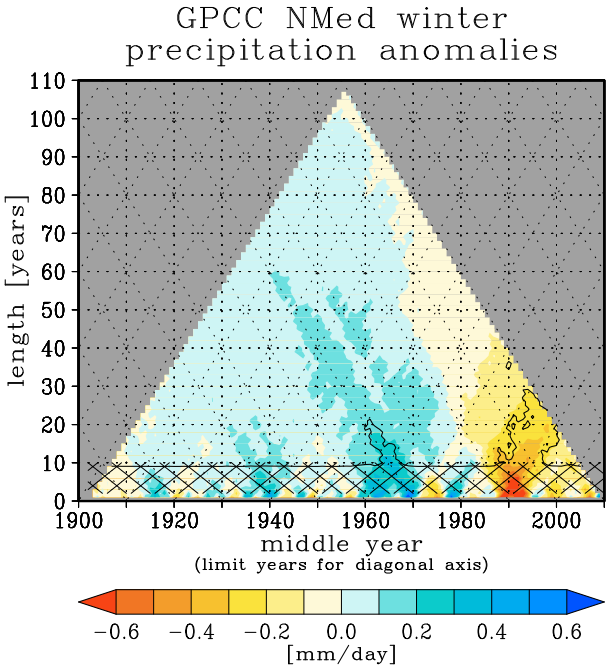


Figure 4.13: 2DMDA diagram for GPCC NMed winter precipitation. Anomalies are relative to the entire record (1902-2010). Contours indicate statistically significant changes at a 95% confidence level. For a complete explanation of the 2DMDA diagram see Sec. 4.1.2.

4 of the 5 driest winters not only since 1951 but also since 1902. The recent pattern of decrease, however, breaks down in the 2000s, as precipitation increases again reaching the same decadal value than in the 1970s. The last winter of the GPCC record (2010), in fact, exhibits the wettest anomaly from the beginning of the twentieth century. Note that in the 2000s we find considerable differences between E-OBS and GPCC data sets, presumably due to the fact that it takes some time for the recently measured data to be added in the gridded data sets and therefore our confidence in the data declines in the last years of each data set (see Sec. 2.1.1).

According to to our findings, then, the recent strong dry anomalies seem unusual compared to the entire record. The 2DMDA diagram (Fig. 4.13) confirms that every time interval centred after 1980 is drier than any other prior segment and, therefore, the last three decades are anomalously dry not only compared to the previous three decades (as seen with E-OBS) but relative to the entire twentieth century. Moreover, all anomalies for time intervals beginning in the 1980s that include the 1988-1995 persistent dry period are statistically significant. Instead, no period longer than 20 years exhibits statistically significant wet anomalies. This conclusion is consistent with the result by Hoerling et al. (2012) who, using a wider winter definition (November-April) and averaging over a larger region (10°W-40°E/30°N-45°N), interpret the recent drying as a change of regime towards drier conditions around 1970.

Yet, according to the 2DLC diagram (Fig. 4.14), the evolution of NMed winter precipitation during the 1902-2010 interval does not quite match what we one would expect if a persistent trend was superimposed on internal noise during the entire period. First, negative changes do not dominate over positive ones when considering the entire diagram. Second, even though there are statistically significant reductions for periods ending recently up to about 110-year long, the longest time intervals do not exhibit statistically significant changes for a wide range of adjacent periods. Finally, the magnitude of the change for time intervals beginning before 1960 do not generally increase with the interval length. Therefore, a persistent NMed winter drying signal is not detectable during the entire twentieth century but only within the last 60-year period (dashed box).

4.4.2 Spatial distribution of the changes

We now investigate whether the result that NMed winter precipitation has been anomalously low since about 1980 is valid for the whole basin or whether such reduction is only visible in the areal mean. For this purpose we calculate, at every grid-point in the NMed region, the difference between the mean winter precipitation after and before 1980 (1981-2010 minus 1902-1980) using GPCC data (Fig. 4.15). The choice of periods is based on the approximate timing of the change found in previous sections and in order to average at

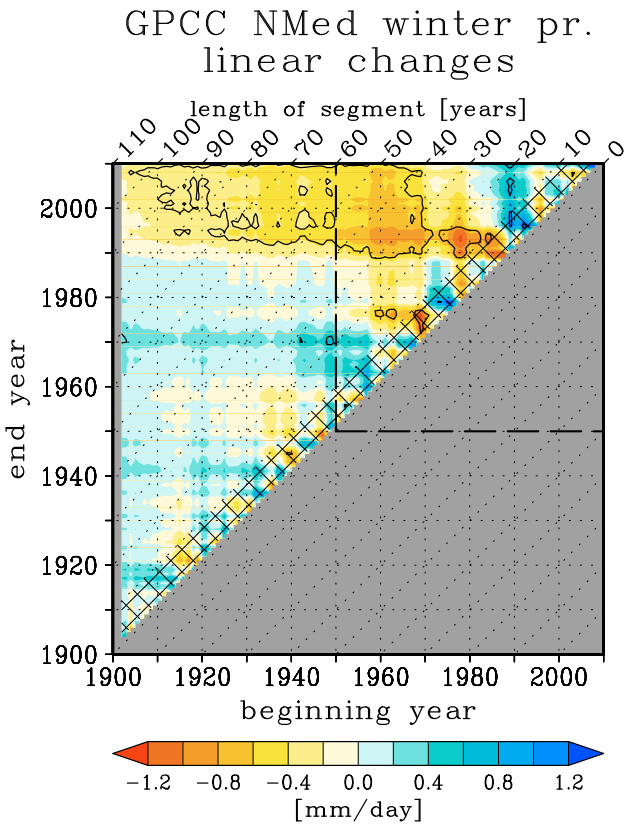


Figure 4.14: 2DLC diagram for GPCC NMed winter precipitation. Contours indicate statistically significant changes at a 95% confidence level. For a complete explanation of the 2DLC diagram see Sec. 4.1.1. The dashed box encloses results for time intervals beginning and ending in the 1950-2010 interval.

least 30 years of data. Results, however, are not sensitive to the exact choice of time intervals (not shown). We evaluate the statistical significance of the detected changes using the Wilcoxon-Mann-Whitney test (see Sec. 3.1.6). Results over 95% confidence level are indicated with small crosses.

The resulting map is characterized by a widespread pattern of precipitation decreases, considerably resembling Fig. 1 in Hoerling et al. (2012), even though they use a wider winter definition (November-April) and calculate the change before and after 1970 instead of 1980. An estimation of the field significance (see Sec. 3.1.6) reveals that the pattern of precipitation reductions over the NMed region is statistically significant at a 95% confidence level. The only region that exhibits slight increases of precipitation is the region north of the Balkan Peninsula, although these are not statistically significant. These Mediterranean reductions are, in fact, part of a large scale pattern characterized by increases in central and northern Europe, as discussed by Hoerling et al. (2012) and van Haren et al. (2012).

Although the pattern of NMed precipitation differences exhibits statistical field significance, only certain regions exhibit locally significant precipitation reductions. In particular, the Iberian Peninsula exhibits significant decreases in several regions, concentrated mainly

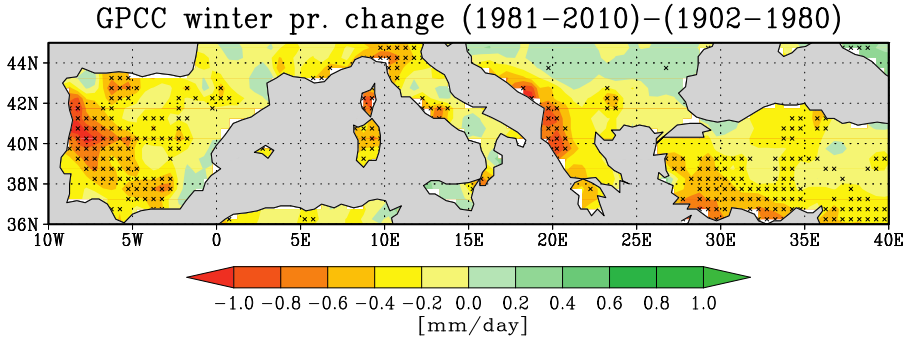


Figure 4.15: Differences in GPCC winter precipitation between 1981-2010 and 1902-1980. Crosses indicate locally significant differences at a 95% confidence level of confidence using the Wilcoxon-Mann-Whitney test. The field significance of the results is over 95%.

in its central-western part. In the Italian Peninsula only the north and parts of its centre-west exhibits significant changes. Large significant reductions are also found over Corsica and Sardinia and the western Balkan coast. Finally, a large fraction of Anatolia, mainly the south-western part of the peninsula, also exhibits large significant decreases.

We now assess whether at the local scale there is evidence of persistent trends lasting until recent years. To do so, we plot the linear change of E-OBS winter precipitation at each grid-point of the northern Mediterranean for time intervals ending in 2014 and beginning every 10 years from 1951 until 1991 (Fig. 4.16), highlighting with crosses those regions with changes statistically significant at a 95% confidence level. For each analysed time interval, the change for the areal mean is highlighted in Fig. 4.11 with blue circles. The pattern of changes for 1951-2014 (Fig. 4.16a) and 1961-2014 (Fig. 4.16b) is generally negative, the only exception being the northern Balkans and most parts of the centre-south of the Italian Peninsula. Only certain regions, however, exhibit locally significant changes: north-western Iberia, northern Italy, the south-western Balkans and Anatolia. Weaker but still significant reductions are found also in the French Riviera, Corsica and Sardinia. As was the case for the areal mean (Fig. 4.11), changes beginning in 1961 are slightly stronger than those beginning in 1951.

The negative winter changes found in Iberia are due to marked decreases in February and March as reported by Paredes et al. (2006) and discussed in detail in Fortuny (2008) and Bladé and Castro-Díez (2010). Precipitation in those months over the central-western part of the Iberian Peninsula exhibited a pronounced decrease of precipitation between the early 1960s and the late 1990s (Fig. 4.17) although it appears to have recovered since.

For the 1971-2014 period (Fig. 4.16c) results are less remarkable. While a (weak) negative pattern persists in the western half of the basin and over Anatolia, locally significant increases of precipitation now dominate the Balkan Peninsula except at the south-western coast, where significant reductions still prevail. Over the Italian Peninsula, we again find

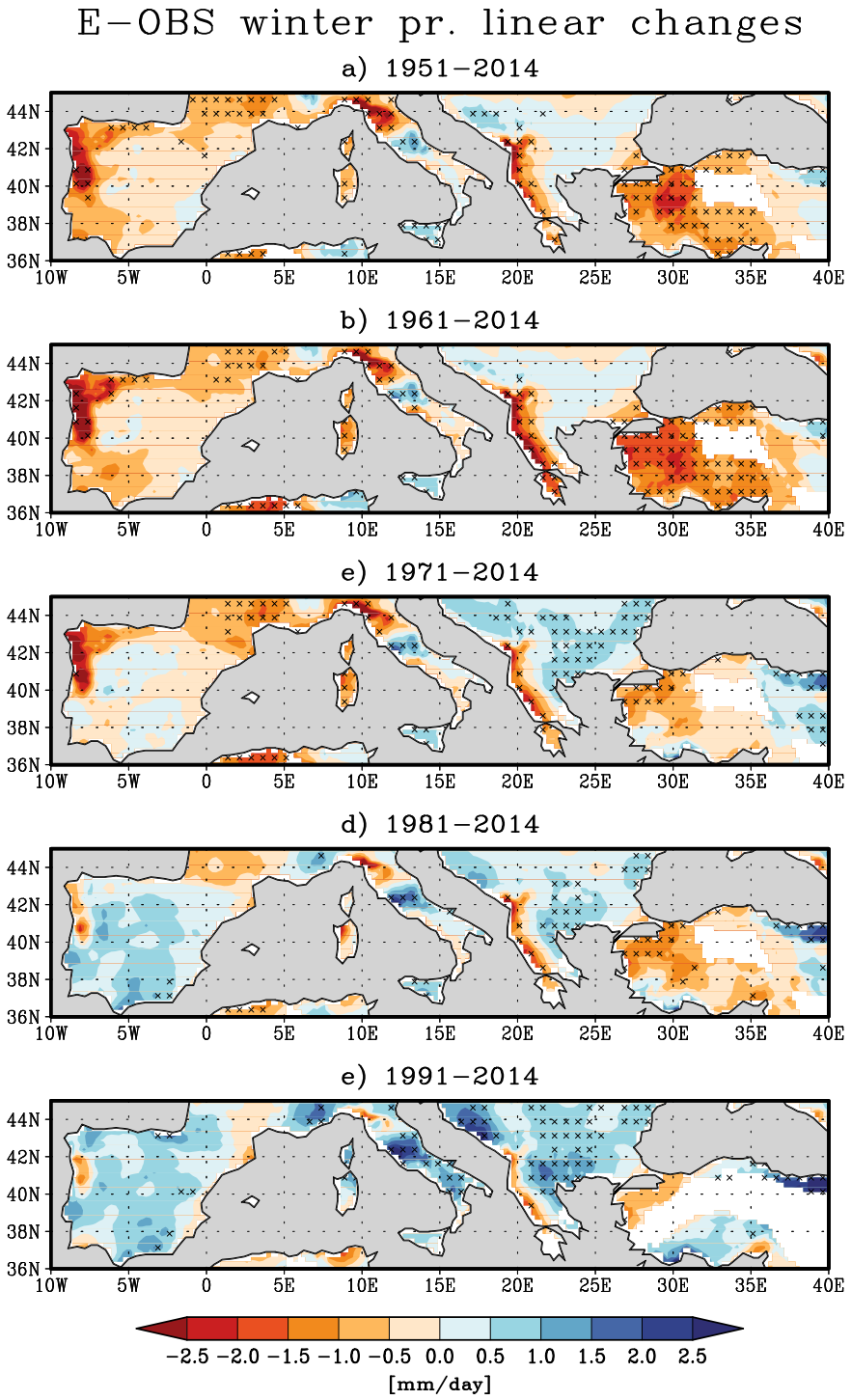


Figure 4.16: Linear change of GPCC winter precipitation for different time intervals ending in 2014. Crosses indicate changes statistically significant at a 95% confidence level. White regions indicate grid-points with more than 30% of missing values in its time series.

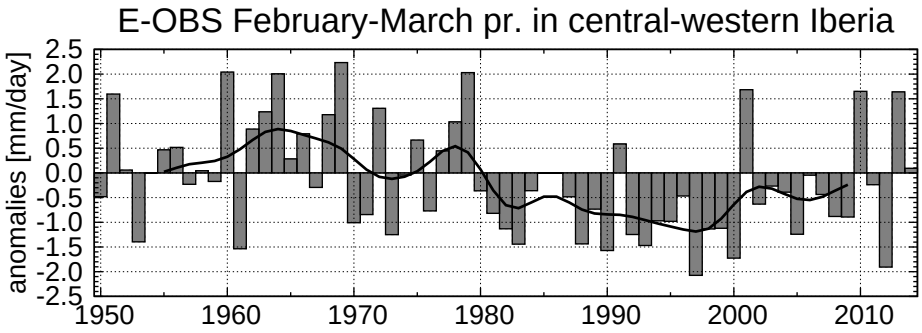


Figure 4.17: Anomalies of E-OBS February-March mean precipitation averaged over western Iberia (10°W-2°W/36°N-43°N) relative to the 1961-1990 mean. The black curve represents filtered data.

significant positive changes in the centre and in Sicily and negative changes over the northern part of the peninsula. The weak negative signal over the French Riviera and Corsica and Sardinia remains.

In contrast, the recent periods 1981-2014 (Fig. 4.16d) and 1991-2014 (Fig. 4.16e) are dominated by positive changes, although they are weak and only significant in central Italy and the central and northern Balkans, with the exception of a narrow region of the south-western Balkans and north-western Anatolia, where we still find large and statistically significant negative changes. As can be seen in Fig. 4.18, however, these negative signals are not robust as they are not present in other data sets (note that in this figure we use the common period 1981-2010). Thus, although most regions of the northern Mediterranean experienced a decrease in winter precipitation between 1950 and 1980 (with the noticeable exception of the northern Balkans), this decrease has substantially slowed down or reversed since 1980.

4.5 Changes in the NAO

Because the NAO is the main driver of inter-annual and decadal variations in northern Mediterranean precipitation [Hurrell et al., 2003; Mariotti and Dell'Aquila, 2012], detecting a persistent signal in the NAO series may help account for the NMed drying observed in the last decades. We start by carrying out a multi-scale analysis of the NAO index, examining whether its observed evolution in the last century is inconsistent with internal variability alone. Then, we evaluate what fraction of the precipitation decrease is linearly related to changes in the NAO. The estimation of the NAO index is explained in detail in Sec. 3.2.1.

The time evolution of the standardized NAO index since 1900 to 2014 is presented in Fig. 4.19. In it we detect an increase in the frequency of winters with positive NAO since the 1980s. In particular, 21 out of the last 34 winters exhibited a positive NAO, with a

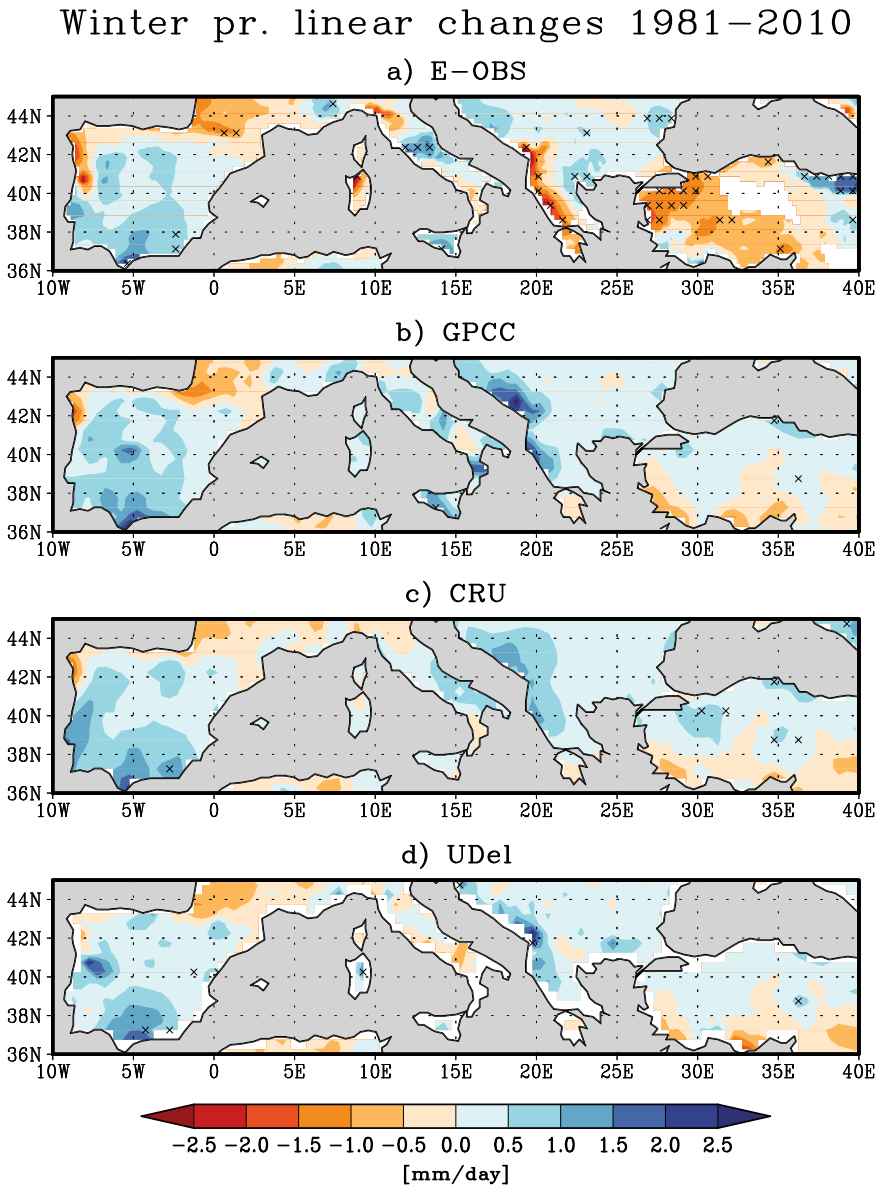


Figure 4.18: Linear change 1981-2010 of winter precipitation for E-OBS (a), GPCC (b), CRU (c) and UDel (d) data. Crosses indicate changes above the 95% confidence level.

noticeable succession of 7 consecutive positive NAO winters in 1989-1995. Before 1981, the number of positive (40) and negative (41) NAO winters was the same.

The behaviour of the NAO index highly resembles the evolution of winter precipitation averaged over the northern Mediterranean region (see Fig. 4.12). Indeed, the correlation coefficient between the two detrended time series is $r = -0.83$ in the common period 1902-2010. Moreover, both variables exhibit a change around the 1980s and an unprecedented accumulation of large anomalies in the 1989-1995 interval. The NAO index, however, seems to exhibit a stronger tendency to persist in the same phase for several years than winter precipitation. For instance, in the 1950s and 1960s, the NAO index displayed a

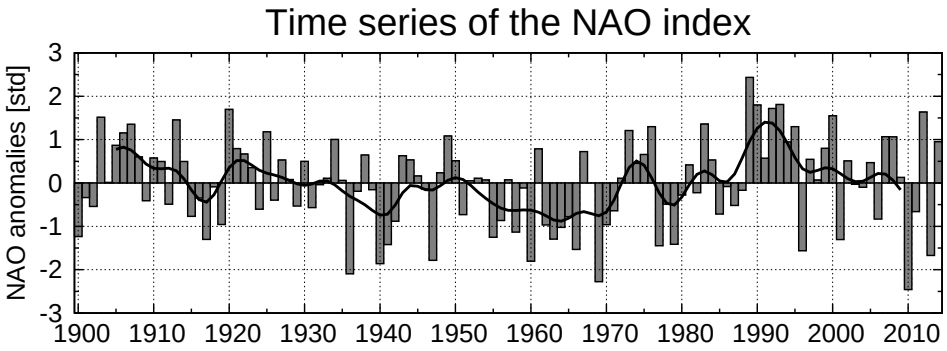


Figure 4.19: Standardized winter NAO index anomalies, relative to the entire period (1900-2014). The NAO index has been defined as the first principal component of the Trenberth SLP field in the north Atlantic (90°W-30°E/40°N-70°N). The black curve represents filtered data.

preference for the negative phase, with no corresponding tendency in precipitation.

We next use the 2DMDA diagram (Fig. 4.20) to evaluate how anomalous the last decades have been in comparison to the whole record (1900-2014). For time intervals centred after 1980, the behaviour of the NAO index is similar to that of winter precipitation (see Fig. 4.13), but for a change of sign. Recall that during positive NAO winters the maximum moisture transport is found over northern Europe, whereas the Mediterranean region experiences moisture divergence that reduces precipitation. The vast majority of time intervals centred after 1980, then, exhibit positive anomalies; these are statistically significant if they begin in the 1980s, end before the 2010, and include the 1989-1995 period. But, as was the case for precipitation, anomalies for periods beginning after 1990 are no longer statistically significant. Before 1980, positive anomalies for periods longer than 10 years are only detected in the first decades of the twentieth century, and they are considerably weaker than those for the recent decades. On the other hand, the last decades are not the only statistically significant anomalous period in the NAO time series. Periods centred in the 1950s and the 1960s, when the NAO had a special preference for the negative phase, exhibit statistically significant anomalies for periods up to 50-year long. This result indicates that strong and persistent anomalous states of the NAO have not been confined to recent decades (such as was the case for precipitation), but there are precedents of signals even longer than, and as intense as, the recently observed. Another noticeable result is that anomalies of any interval length ending in the last four winters in record (2011-2014) are no longer statistically significant, indicating that the previous increase between the 1960s and the 1990s has been offset by recent changes.

The robustness and the statistical significance of NAO long-term changes is now evaluated using the 2DLC diagram (Fig. 4.21). Analysing first those time intervals beginning after 1950 (dashed box), we see that (1) positive changes dominate over negative ones, (2) the longest time intervals exhibit statistically significant changes for a wide range of adjacent periods, and (3) the magnitude of the change generally increases with the length

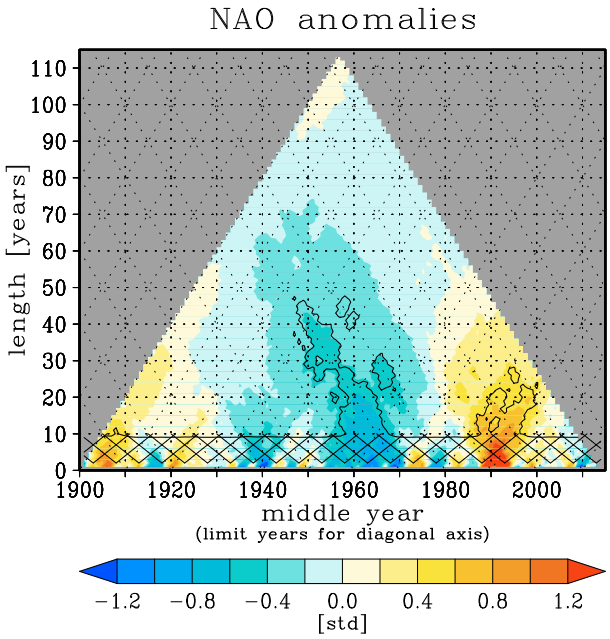


Figure 4.20: 2DMDA diagram for the NAO index plotted in Fig. 4.19. Anomalies are relative to the entire record (1900-2014). Contours indicate statistically significant changes at a 95% confidence level. For a complete explanation of the 2DMDA diagrams see Sec. 4.1.2.

of period. In particular, as was the case for precipitation (see Figs. 4.11 and 4.14; note the sign change), since 1950 onwards, most time intervals ending after 1990 exhibit statistically significant changes. Moving to longer time intervals, the magnitude of the changes drops when including years prior to 1950, and becomes non-significant for time intervals beginning before the mid-1920s. Therefore, a century-long signal is not detectable. Moreover, apart from those statistically significant positive changes in periods ending in the recent decades, we also find strong statistically significant negative changes for time intervals beginning in the 1910-1920s and ending in the 1960-1980s, associated with the NAO preference for the negative phase during the central decades of the twentieth century.

4.5.1 Role of the NAO in winter precipitation changes

Having determined that NMed winter precipitation and the NAO index have experienced statistically significant trends in the last decades, now we examine the extent to which the signal in both variables is related. In other words, we examine in detail whether the NAO is able to explain the observed northern Mediterranean winter precipitation decreases. The use of the two-dimensional diagrams will allow us to assess whether conclusions like those by Hoerling et al. (2012) and Barkhordarian et al. (2013) are justified. In the first case, Hoerling et al. (2012) conclude that the NAO has played a minor role in the observed precipitation changes in the last decades. They base their conclusion on the lack of a long-term (1902-2010) NAO trend and on the result that the NAO has not exhibited a dramatic change

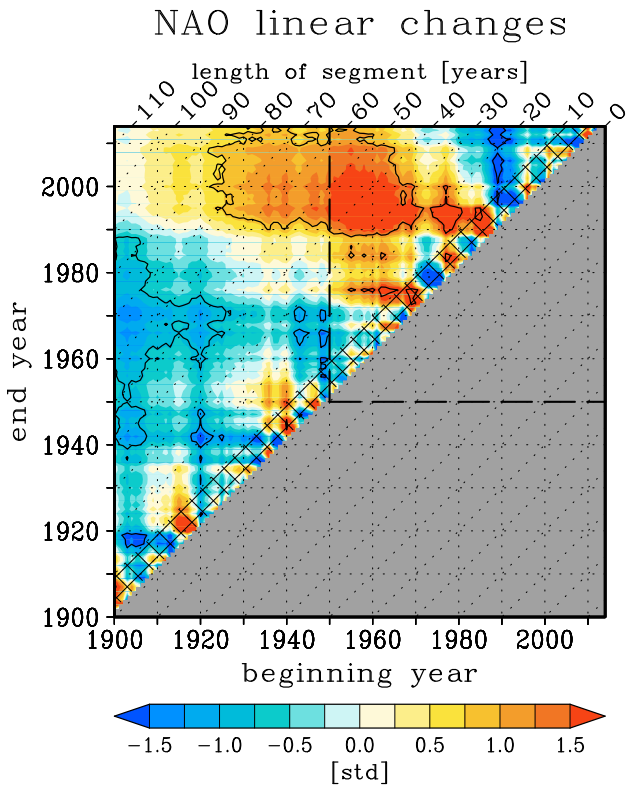


Figure 4.21: 2DLC diagram for the NAO index plotted in Fig. 4.19. Contours indicate statistically significant changes at a 95% confidence level. For a complete explanation of the 2DLC diagram see Sec. 4.1.1. The dashed box encloses results for time intervals beginning and ending in the 1950-2014 interval.

of behaviour since 1970 like the one experimented by precipitation. We have demonstrated, however, that the NAO anomalies for the last decades of the twentieth century are actually significantly anomalous. In the second case, Barkhordarian et al. (2013) also conclude that the contribution of the NAO to the northern Mediterranean precipitation trend for the period 1979-2008 is almost negligible. They use the 1979-2008 period because it corresponds to the entire record of the Global Precipitation Climatology Project (GPCP) data set but, according to Fig. 4.21, in this time period the NAO trend is very weak as it is located in the transition between intervals with positive NAO change and intervals with negative NAO change. In this section we evaluate whether their conclusion can be extrapolated to time intervals in which the NAO trend is stronger.

To calculate the fraction of area-averaged NMed precipitation explained by the NAO in each individual winter, we use the regression of this precipitation onto the standardized NAO index (which is −0.35 (mm/day)/std). We compute the product of this regression and the NAO index at every instant to create a time series representing the component of precipitation that is linearly explained by the NAO (hereafter *pr-nao*). The difference between the actual precipitation and pr-nao is, therefore, the component that is not related

to the NAO (hereafter *pr-nonao*).

In Fig. 4.22 we represent the time series of pr-nao and pr-nonao, together with the total NMed winter precipitation (previously shown in Fig. 4.12). Consistent with the large correlation between NAO and precipitation, pr-nao accounts for most of the precipitation anomalies. In particular, the accumulation of dry years in the 1989-1995 interval is seen to be clearly due to the persistent positive NAO signal in those years. The pr-nonao time series, instead, exhibits weaker anomalies without any apparent change in the frequency of dry or wet anomalies.

Figures 4.23a and 4.23b show the 2DMDA diagram for pr-nao and pr-nonao, respectively. As the time series of pr-nao is the product of the NAO time series (Fig. 4.19) and a constant negative value (the regression coefficient between precipitation and NAO), the behaviour of both time series is identical, except for the sign. As a consequence, the analysis of the pr-nao time series and its two-dimensional diagrams is the same than that of the NAO index carried out in section above. Still, here we plot these diagrams in order to compare them with those of pr-nonao. Unlike the diagram for pr-nao, the diagram for pr-nonao does

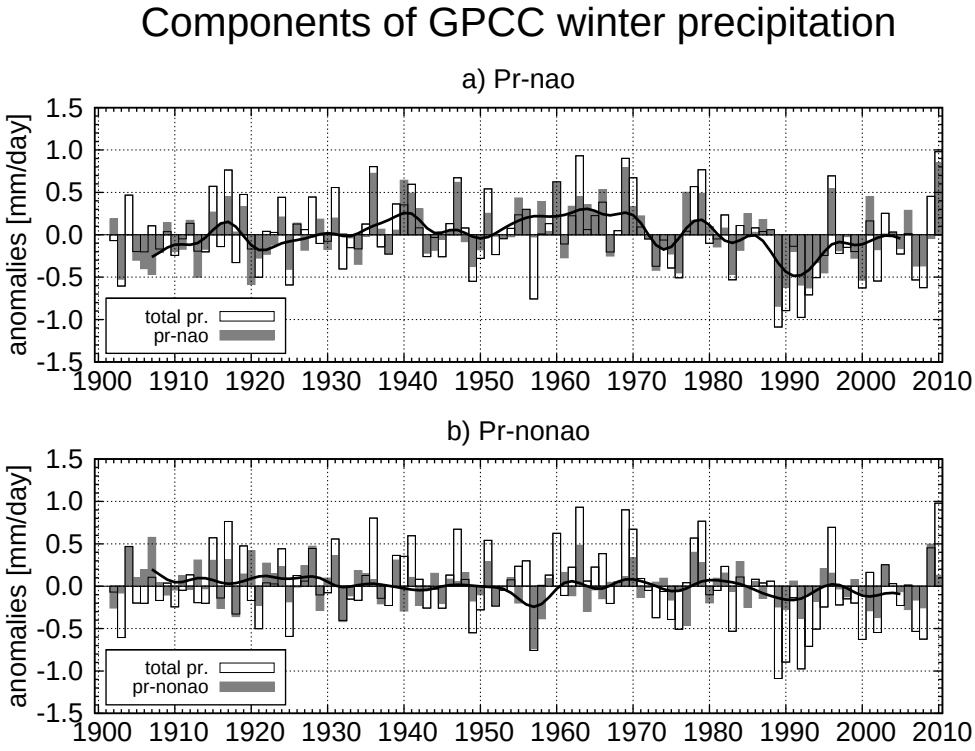


Figure 4.22: (a) Filled boxes: Time series of the component of GPCC NMed winter precipitation that is linearly related to the NAO (pr-nao). It is defined as the product of the linear regression of the precipitation onto NAO time series and the normalized value of the index in each winter. (b) Filled boxes: Time series for the fraction of precipitation that is not consistent with the NAO (pr-nonao), defined as the difference between the total precipitation and pr-nao. Unfilled boxes in both plots represent the time series of GPCC NMed precipitation shown in Fig. 4.12a.

not exhibit any statistically significant anomalies and is consistent with internal variability alone. Yet, the fact that all multi-decadal anomalies of pr-nonao post-1980 are weakly negative may indicate that other mechanisms slightly reinforced the negative impact of the NAO on precipitation in those particular decades. Instead, the negative (non-significant) anomalies product of the internal variability found for pr-nonao in periods centred around 1960 may indicate that these NAO-unrelated mechanisms partially compensated the strong NAO-induced precipitation increase in those years, explaining why in the precipitation diagram (Fig. 4.13) there is not a statistically significant wet period matching the negative NAO signal.

This is also visible in Figs. 4.23c-d, in which we plot the percent contribution of both components to the total change to make more evident our results. These figures allow us to objectively evaluate the conclusion drawn by Hoerling et al. (2012) that the NAO played a minor role in the precipitation change around 1970. In the diagram for pr-nao we see that, for time periods beginning after 1960, pr-nao is the largest component of NMed winter precipitation, accounting for more than 50% of almost all possible multi-

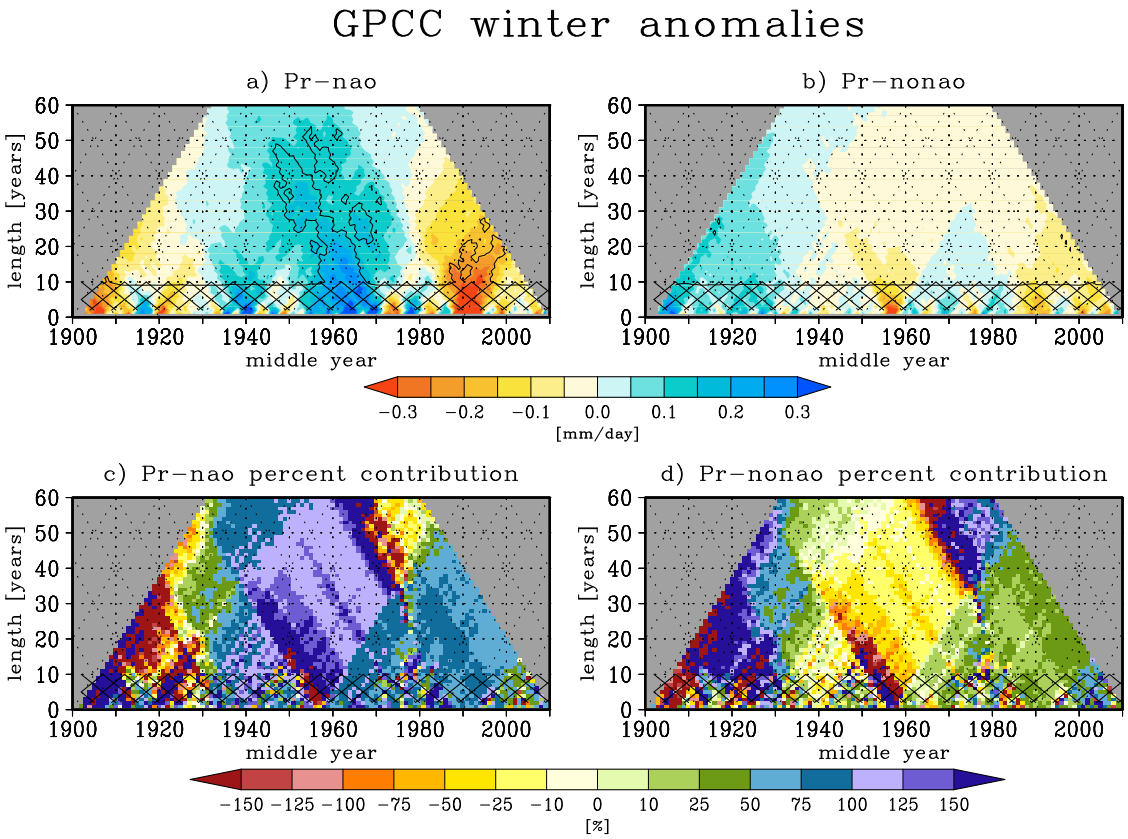


Figure 4.23: (a-b) 2DMDA diagrams for pr-nao and pr-nonao. Anomalies are relative to the entire record (1902-2010). Contours indicate statistically significant changes at a 95% confidence level. Results for periods longer than 60 years have been omitted because no significant results are found for such long time intervals. For a complete explanation of the 2DMDA diagram see Sec. 4.1.2. (c-d) Percent contribution of each component to the anomalies of total precipitation (see Fig. 4.13).

decadal anomalies. Instead, in the first decades of the twentieth century NAO-unrelated mechanisms more than compensated for the effect of the NAO; note that the precipitation anomalies were positive (see Fig. 4.13) despite the positive NAO.

To evaluate the contribution of pr-nao and pr-nonao to the winter precipitation linear changes observed in Fig. 4.14 we use the 2DLC diagrams (Figs. 4.24a-b). Despite the lack of statistically significant changes in the diagram for pr-nonao, only this component exhibits weak but statistically significant century-long trends. This result is indeed consistent with the conclusion drawn by Hoerling et al. (2012) that the NAO played a minor role in the long-term 1902-2010 wet-season (November-April) precipitation trend. Nevertheless, according to the diagrams of percent contributions (Figs. 4.24c-d), when focusing on changes within the 1950-2010 interval (dashed box) we see that they are mainly dominated (between 50% and 100%) by the pr-nao component. These two-dimensional diagrams also allow us to conclude that the conclusion drawn by Barkhordarian et al. (2013) that the NAO played no role in the 1979-2008 winter precipitation trends is not valid for time intervals in which the NAO trend is stronger.

Finally, we assess whether the conclusion that the NAO has played a major role in the precipitation changes observed in the last decades is valid for each region across the basin. To compute the spatial signature of the precipitation response to NAO variations, we perform a linear regression between the two variables at every grid-point of the basin using GPCC data (Fig. 4.25a). Very similar results are found with E-OBS, CRU and UDel data (Figs. 4.25b-d). As discussed in Sec. 1.2.1, NAO and precipitation are inversely related in all the northern Mediterranean and exhibit statistically significant regressions everywhere except in small regions of the eastern coast of Iberia, Sicily, and some parts of Anatolia. The largest regressions are found in Iberia, particularly in its western half, and in the western Balkan coast.

The pattern of regressions largely resembles the first principal component of winter Mediterranean precipitation presented by Kelley et al. (2012b), who also use GPCC data (their Fig. 2). The low values found in the eastern part of the basin (west of 20°E) lead them to conclude that the NAO is not able to completely explain the observed drying in those regions. Similarly, Hoerling et al. (2012) base their conclusion that the observed Mediterranean precipitation decreases (their Fig. 1) cannot be attributed to the NAO on the result that the change in wet season NAO-related precipitation (November-April) between 1971-2010 and 1902-1970 is too weak (their Fig. 6). The change of scale between the two figures by Hoerling et al. (2012), however, makes it difficult to accurately see how they compare. In order to quantify the NAO influence on winter precipitation across the basin and be able to contrast these conclusions, we estimate the geographical signature of pr-nao and pr-nonao changes and their percent contribution to the total change.

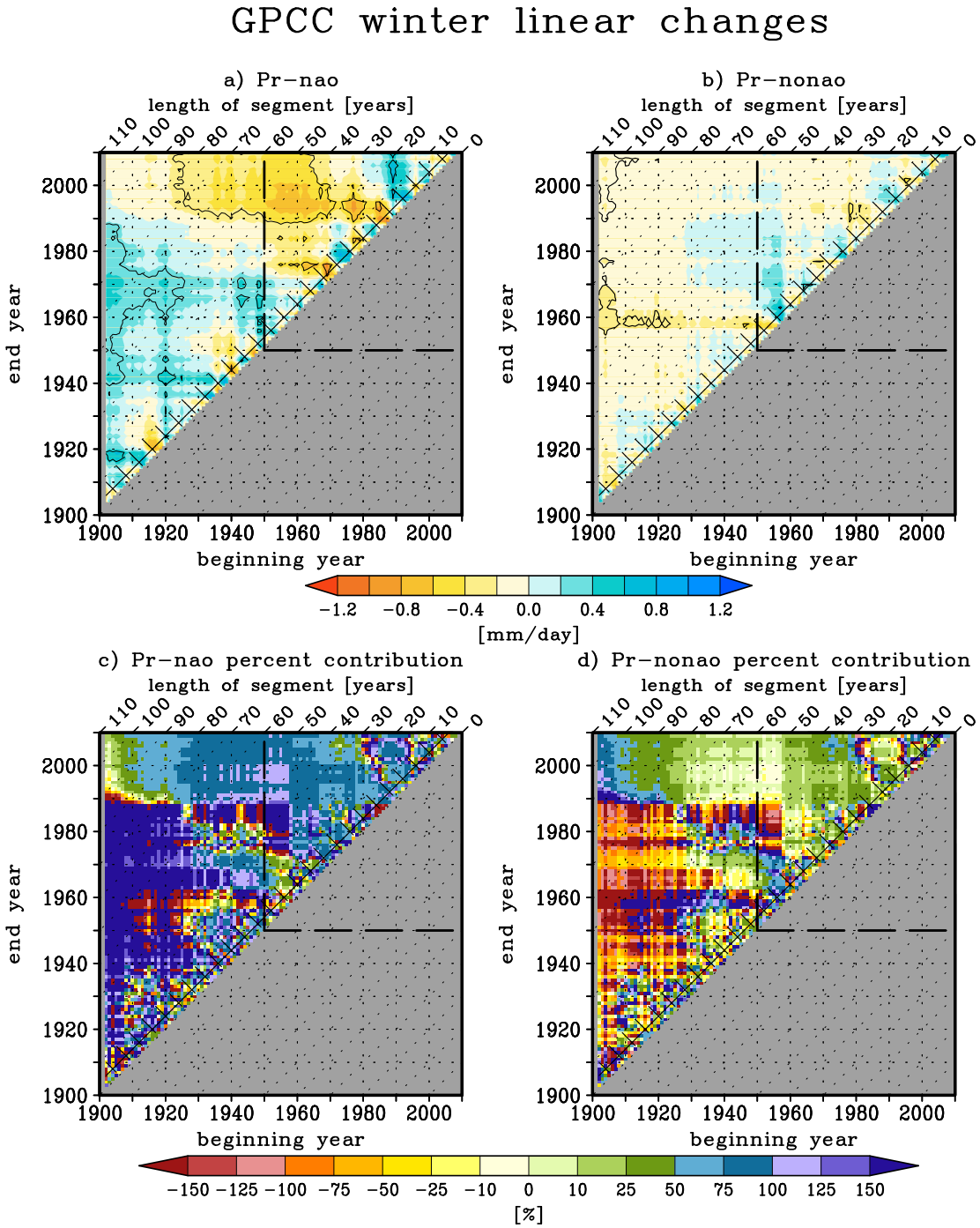


Figure 4.24: (a-b) 2DLC diagram for pr-nao and pr-nonao. Contours indicate statistically significant changes at a 95% confidence level. For a complete explanation of the 2DLC diagram see Sec. 4.1.1. The dashed box encloses results for time intervals beginning and ending in the 1950-2010 interval. (c-d) Percent contribution of each component to the total precipitation changes (see Fig. 4.14).

Figure 4.26a displays the difference in the winter pr-nao component between 1981-2010 and 1902-1980. The largest decreases in pr-nao coincide with those found for total precipitation (see Fig. 4.15), and the magnitude of precipitation and pr-nao changes is comparable. Instead, the signature of pr-nonao changes is generally non-significant. The only significant signals are found in small isolated regions of northern Italy, Corsica and Sardinia, Anatolia and also in extreme north-western Iberia and the northern Balkans, where

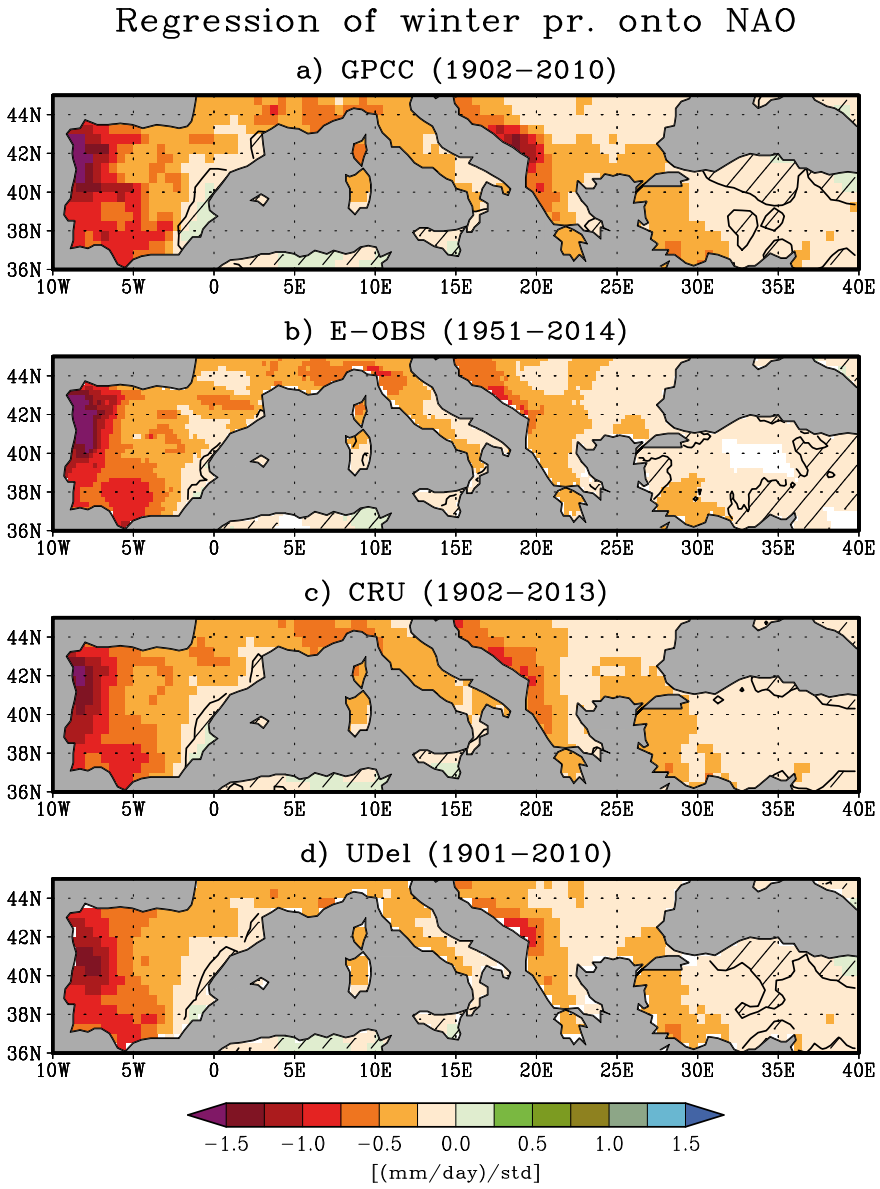


Figure 4.25: (a) Regression of GPCC winter precipitation onto the NAO index shown in Fig. 4.19 calculated for the 1902-2010 interval. Dashed regions indicate changes under the 95% confidence level of statistical significance. White regions indicate grid-points with more than 30% of missing values in its time series. All time series have been detrended prior to the calculation. (b-d) The same for E-OBS, CRU and UDel winter precipitation. The time interval for which each regression is estimated is indicated in brackets.

positive changes are found in total precipitation, contrary to the general decrease.

Looking at the percent contribution maps (Fig. 4.26c and 4.26d), we find that the role of the pr-nao component of winter precipitation is clearly dominant in the western part of the basin, in particular Iberia and the French Riviera. In the central part of the basin (Italian Peninsula and south of Balkan Peninsula) both pr-nao and pr-nonao components contribute equally to the total precipitation. The only regions where the NAO played a minor role in

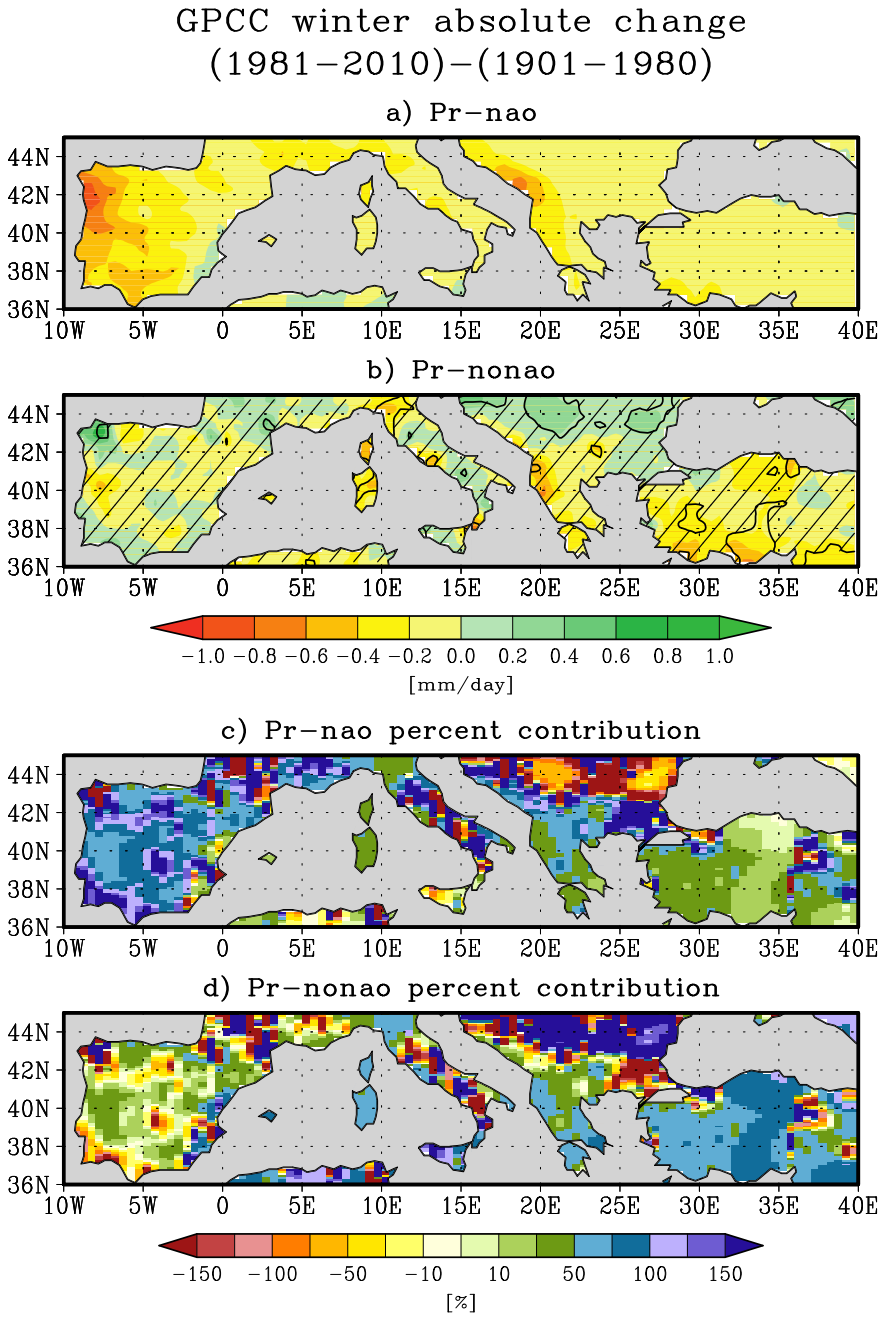


Figure 4.26: (a-b) Differences in pr-nao and pr-nonao between 1981-2010 and 1902-1980 at every grid-point. Hatched regions indicate locally non-significant differences at 95% confidence level. (c-d) Percent contribution of each component to the absolute precipitation change (see Fig. 4.15).

the winter precipitation change are the Anatolian Peninsula and the north of the Balkan Peninsula. Also in Corsica, Sardinia and the north of the Italian Peninsula the pr-nonao component dominates, although the NAO is still able to explain between 25% and 50% of the observed winter precipitation changes.

Our conclusions are in agreement with Hoerling et al. (2012) and Kelley et al. (2012b) in as much as we find the largest impact of the NAO in the western half of the basin. Our results, however, contradict the conclusion by Hoerling et al. (2012) that NAO-related precipitation anomalies are too weak even in the western half of the basin, since we instead find that they account for more than 50% (75% in most areas) of the total observed precipitation decrease.

4.6 Conclusions and discussion

With the use of our two-dimensional diagrams we have determined for which seasons and periods the evolution of northern Mediterranean precipitation is consistent with the presence of a persistent signal and verified that the results are not sensitive to the choice of time interval. Our conclusions are robust as they are reproducible with different data sets and are not sensitive to the choice of months included in season definitions.

In this way we have shown that the weak annual precipitation decrease observed in the last 30-year period is the result of strong decreases in winter partially compensated by weak increases in April, October and mainly September. Focusing on the winter decreases we have concluded that the last three decades have been significantly drier than earlier ones, and the evolution of area-averaged winter precipitation in the northern Mediterranean since mid-twentieth century onwards is compatible with what one would expect if a persistent drying trend was superimposed on internal noise. Moreover, we have found that the decrease in winter precipitation is not only visible in the areal average but is present in most of the northern Mediterranean basin. The regions in which the strongest decreases have been observed are the western Iberia, the northern Italy, Corsica and Sardinia, the southwestern Balkans and Anatolia. The northern half of the Balkan Peninsula is an exception as precipitation increases have occurred in that region since mid-twentieth century.

In the last decades of the twentieth century an upward trend in the NAO index has been observed, causing NAO values for the last decades of the twentieth century to be significantly anomalous also with respect to the entire 110-year observational period. By computing the fraction of winter precipitation linearly related to the NAO, we conclude that the upward NAO trend explains most of the northern Mediterranean winter precipitation decreases. This conclusion is also valid for regional changes in most of the basin, with the exception of the northern Balkans and Anatolia, where the main contribution to the change appears to come from other mechanisms.

Whether the detected signals in winter precipitation and the NAO index are attributable to anthropogenic forcings is not assessable with observational data only. For this reason, in Chap. 5 we compare the largest observed winter precipitation and NAO changes with associated changes in forced and unforced climate simulations. Nevertheless, some of the results of this chapter already call into question this hypothesis. First, in the last decade the pattern of winter precipitation decreases and NAO increases has broken down. Of course, the lack of large anomalies in the 2000s could be due to internal variability superimposed on the drying signal. But it could also be the succession of wet (positive NAO) years in the 1960s and dry (negative NAO) years in the 1980s and 1990s what was the result of low-frequency internal variability. Second, the large positive winter NAO anomalies observed in the recent 30-year period, while highly significant with respect to the entire 110-year period, are not unprecedented. Earlier intervals up to 50-year long also exhibited large statistically significant anomalies of opposite sign (in the 1950-1960s). In fact, in the 1900-1975 interval the evolution of the NAO index was compatible with the presence of a downward signal superimposed on internal variability. Finally, the seasonal distribution of observed precipitation changes in the northern Mediterranean does not match that of the future projections of drying. According to CMIP5 projections, the largest northern Mediterranean drying response to climate change is expected in summer. According to observations, however, only in winter have we observed robust precipitation decreases in the last decades.

Chapter

FIVE

Comparing recent precipitation trends in the northern Mediterranean in models and observations

Are the recent observed precipitation decreases in the northern Mediterranean found in Chap. 4 robust evidences that the strong drying signal projected by climate models [Christensen et al., 2013] is already detectable? Considering only winter months, not only has the last 30-year period been found significantly drier than the entire twentieth century, but the negative winter precipitation trends in the last decades also stand out in a context of strong multi-decadal variability. In order to determine whether these observed decreases indicate that the projected anthropogenically-forced drying signal has begun to emerge it is essential to compare them with forced and unforced model simulations.

In this chapter we use pre-industrial and historical simulations from the models of the CMIP5 multi-model ensemble [Taylor et al., 2012b]. While pre-industrial simulations are run only with natural processes allowing us to estimate the range of changes associated with internal variability, historical simulations incorporate the anomalous forcings observed from the industrial revolution onwards and, therefore, allow us to determine the responses to these external forcings. Following Knutson et al. (2013), a signal will be considered detectable only if it is inconsistent with internal variability alone. In addition, a detected signal will be considered attributable to external forcings only if it is consistent with forced simulations. Even though without examining the range of inter-model variability one cannot do detection and attribution of climate signals, most authors who analyse the response of European/Mediterranean precipitation to climate change have only focused on the study of multi-model mean trends as a measure of the forced signal present in climate simulations [e.g. Mariotti et al., 2008; Giorgi and Lionello, 2008; van Haren et al., 2012].

We have already noted in Chap. 4 that the seasonal distribution of the drying is not consistent between models and observations: while the strongest drying is projected for summer, the only robust decreases so far have been observed in winter. This discrepancy leads us to first assessing whether the relatively strong changes in observed winter precipitation are consistent with the relatively weak winter signal in models. In other words: would the observed winter precipitation change not be an outlier if considered a member of the model ensemble distributions?

In the previous chapter we also concluded that recent observed winter precipitation trends in the northern Mediterranean region are largely explainable by an upward trend observed in the NAO during the last decades of the twentieth century [e.g. Hurrell, 1995; Gillett, 2005]. This begs the question: do models support the hypothesis that a forced signal is detectable in the NAO? And, if so, is this signal attributable to external forcings? To find the answers to these questions we calculate the NAO index in pre-industrial and historical simulations.

This chapter begins with an illustrative explanation about how we estimate the contribution of external forcings and internal variability in climate simulations and which conclusions can be drawn depending on the difference between observed and simulated trends. Then, in Sec. 5.2 we check that CMIP5 models are able to reproduce the main aspects of northern Mediterranean precipitation climatology. In Sec. 5.3 we describe the seasonal distribution of precipitation changes in historical simulations. Next, in Sec. 5.4 and 5.5 we describe separately the simulated precipitation trends in winter and in summer. For winter, moreover, we compare observed precipitation and NAO trends with forced and unforced simulations and assess whether they are consistent, whether a drying signal (and an upward NAO signal) is detectable and whether it is attributable to external forcings. For summer, we describe the strong summer trends present in historical simulations and summarize our articles about the impact of the summer NAO (SNAO) on Mediterranean precipitation [Bladé et al., 2012a; Bladé et al., 2012b].

5.1 Comparing observed and simulated trends

The change in precipitation ΔP (or any other climatic variable) during a given time interval, whether it is evaluated as a difference of means or as the linear trend over the entire period, can be approximated as the sum of a persistent signal superimposed on internal variability $\Delta P = \Delta P_{sig} + \Delta P_{int}$. The persistent signal is assumed to be the response of precipitation to external forcings (either natural or anthropogenic), whereas internal variability is random, result of non-linear dynamical processes intrinsic to the climate system [Deser et al., 2012a]. Given an observed change of precipitation in a given time interval, however, it is not possible to discern the contribution of ΔP_{sig} and ΔP_{int}. The only hypothesis we can

assess on the basis of observations alone is whether the evolution of precipitation in that interval is consistent with what we would expect if a persistent trend was superimposed on random noise (see Chap. 4). Using historical simulations and pre-industrial simulations, instead, we can assess the hypotheses that a persistent signal is detectable in observational series and that it is attributable to external forcings.

In Fig. 5.1 we illustrate possible probability distribution functions of precipitation changes obtained with independent model simulations assuming gaussianity. Magenta distributions, centred in 0, illustrate the range of possible changes in absence of forcings ΔP_{int} and from now on they will be called *natural distributions*. Blue distributions represent the range of possible changes in presence of a superimposed forced signal $\Delta P = \Delta P_{int} + \Delta P_{sig}$ and, therefore, they are centred in ΔP_{sig}. From now on, these distributions will be called *forced distributions*. The position of the observed precipitation trend relative to these two distributions will determine whether the hypotheses that a signal is detectable and that it is attributable to external forcings are supported or not. We consider the observed value to belong to these distributions if it is inside their central 95% range, indicated in the figure. This analysis is based on the methodology described by Knutson et al. (2013).

The first diagram in Fig. 5.1 symbolizes a situation in which the forced signal is large in comparison with internal changes. This example is applicable, for instance, to surface air temperature averaged over Europe as we shall see later. In this case, we identify five distinct regions in which the observed change can be found, indicated with colours in the horizontal axis:

• **Green region:** The observed trend is outside the natural distribution but inside the forced distribution. In that case, models support the hypotheses that a signal is detectable and that it is attributable to external forcings.

• **Yellow region:** The observed trend is inside both distributions. Because the observed trend is not inconsistent with the natural distribution, we consider that no signal is detectable.

• **Black region:** The observed trend is outside the forced distribution but inside the natural distribution. Again, the hypothesis that a signal is detectable is rejected.

• **Red region:** The observed trend is larger than the whole range of possible changes in presence of a forced signal. We then interpret that a signal is detected and in part attributable to an external forcing. Since the observed trend is not fully compatible with model simulations, we can conclude that either the forced signal or internal variability are underestimated as discussed below.

• **Purple region:** The observed trend and the forced signal have opposite sign. Moreover, the observed trend is outside the natural distribution. A signal is detected, but again it is not compatible with model simulations. Either the forced signal is unrealistic or the range of internal variability is underestimated.

If the observed change of precipitation falls in one of the last two regions, models are partially misrepresenting reality. On the one hand, the misrepresentation could be related to the magnitude of the forced signal. As indicated in Fig. 5.1b, with a stronger forced signal any given observed change can fall inside the forced distribution (note the shift in the forced distribution from the broken one to the solid one), making models and observations compatible. On the other hand, the misrepresentation could be associated with the range of internal changes, i.e. the width of the distributions. As illustrated in Fig. 5.1c, any given observed change can fall inside one of the distributions if these become wider.

Because not all variables exhibit separate natural and forced distributions, the diagram in Fig. 5.1d symbolizes a situation in which the forced signal is negligible in comparison with the width of the distributions. This is the case, as we shall see later, of northern Mediterranean area-averaged winter precipitation. Under these conditions, the effective number of conclusions that we can draw is reduced. Now green and black regions are almost non-existent and we only have to test whether the observed change is inside or outside the yellow region. In the first case, we will conclude that climate simulations are consistent with observations but no signal is detectable. In the second case, even if the observed trend falls within the red region, the conclusion that the signal is attributable in part to an external forcing becomes irrelevant given the weakness of the forced signal. In this situation, even though the condition of detection is fulfilled (the observed change is inconsistent with internal variability alone), we will conclude that simulations underestimate the magnitude of the forced signal (or even miss it completely) or the range of internal variability.

In addition, the diagram in Fig. 5.1e exemplifies a situation in which the forced signal is very strong and the natural distribution is completely separated from the natural distribution. This is the case of the surface air temperature averaged over the northern hemisphere. Now, we find another region:

• **Gray region:** The observed trend is larger than the range of natural changes but smaller than all possible changes in presence of a forcing. The signal is detected and partially attributable to external forcings. However, the observed trend is not fully compatible with model simulations and either the forced signal is overestimated or internal variability is underestimated.

5.1.1 Estimation of forced signal and internal variability

In this chapter we use two kind of climate simulations from CMIP5 models to estimate the natural and forced distributions of changes (see Sec. 2.2). First we use pre-industrial control simulations, which are run under fixed natural forcings characteristic of pre-industrial times for long time intervals (generally several centuries) and do not include anomalous external forcings. Secondly we work with historical simulations, which are initialized in a selected instant of the control runs and from then on they incorporate the new forcings

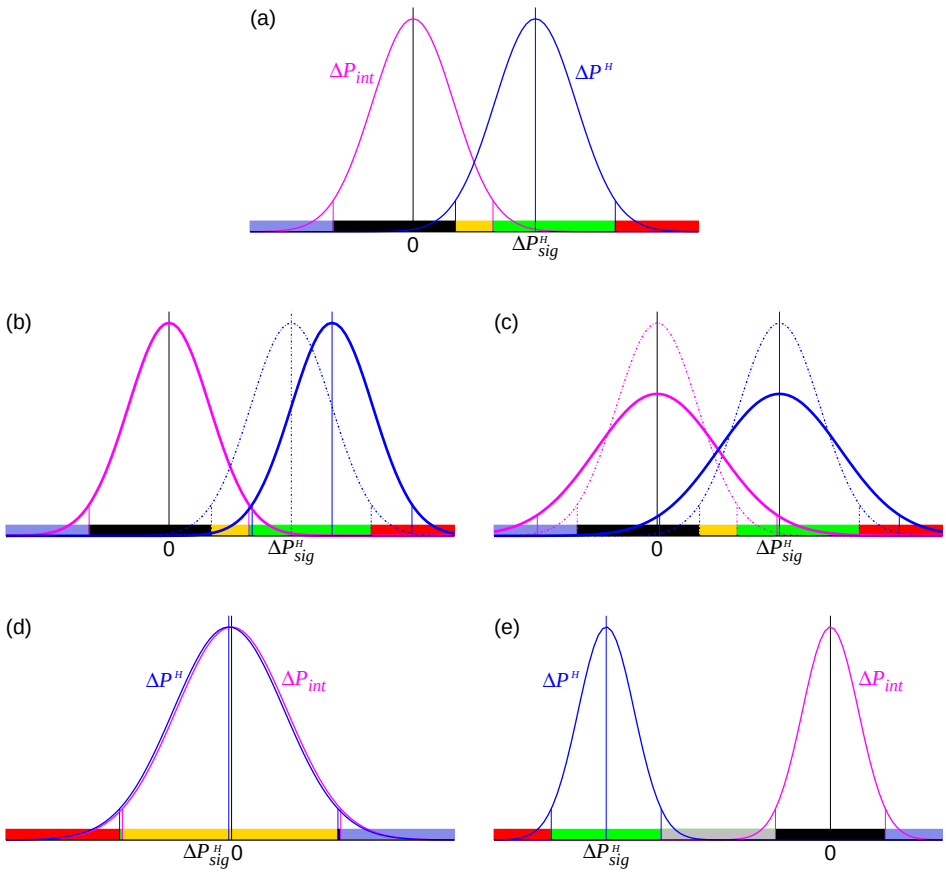


Figure 5.1: Example of distributions of trends obtained with model simulations. Magenta curves represent trends only associated with internal variability and blue curves represent trends influenced by a forced signal. For all distributions the 2.5th and the 97.5th percentile are indicated. Colours in the horizontal axes represent those regions in which an observed trend can be found relatively to the two represented distributions. See text for a complete explanation of each diagram.

observed since mid-nineteenth century, both anthropogenic (e.g. the increase in the concentration of GHG and aerosols) and natural (e.g. volcanic aerosols). The precipitation change for a given time interval within a given historical run can be expressed, then, as $\Delta P^H = \Delta P^H_{sig} + \Delta P^H_{int}$. Again, for a single simulation it is not possible to evaluate the contribution of each component.

Now suppose we have many historical simulations available for a given climate model and the only difference among them is the selected instant in the pre-industrial control run used to initialize them (i.e. the initial conditions). In this situation, ΔP^H_{sig} is expected to be common in all those runs, as they all share the same forcing and the same potential response to the given forcing. Instead, the difference in initial conditions among simulations should cause ΔP^H_{int} to be unrelated from one run to another. Therefore, if we average the change in precipitation for a given time interval in all those many historical realizations of the same model, we expect $\overline{\Delta P^H_{int}} \rightarrow 0$ and, therefore, the magnitude of the change associated with the forced signal for that model would arise as $\Delta P^H_{sig} \approx \overline{\Delta P^H}$. Unlike ΔP^H_{sig}, the change

associated with internal variability ΔP^H_{int} for a given time interval and for a given model is not a single magnitude but the whole range of values that each one of the single historical simulations takes. To estimate this range, we must remove the change associated with the forced signal from all those historical simulations $\Delta P^H - \Delta P^H_{sig}$ and represent the obtained distribution of values. This methodology is used by Deser et al. (2012a), Deser et al. (2012b) and Deser et al. (2014) to separate the contribution of internal variability from the contribution of external forcings in future temperature and precipitation projections.

One can also estimate the magnitude of internal variability using pre-industrial control runs [i.e. Deser et al., 2012b]. Note that, as these simulations are run with no external forcings, the magnitude of the change for a given time interval is entirely due to internal variability: $\Delta P^{PI} = \Delta P^{PI}_{int}$. To estimate the range of internal variability associated with a given time scale in a climate model we calculate all possible trends of the same fixed length in pre-industrial simulations. Because these simulations are generally very long, this method gives a better estimate of internal variability than the spread of trends in historical simulations. This methodology, however, is only useful to estimate the internal variability present in historical times if one assumes that the forcing present in historical simulations has not substantially modified internal variability ($\Delta P^{PI}_{int} = \Delta P^H_{int}$).

Although ideally long-term changes in pre-industrial control simulations ΔP^{PI}_{sig} should be null, some models exhibit small drifts in some variables. These drifts are commonly associated with models not being in equilibrium, discontinuities in fluxes between model components, or errors regarding the conservation of heat and moisture [Sen Gupta et al., 2013]. For instance, small drifts in globally averaged precipitation related to drifts in globally averaged sea surface temperature have been reported in CMIP5 models by Sen Gupta et al. (2013). Because of these drifts, the changes in pre-industrial control runs are not just due to internal variability. Moreover, drifts add a spurious signal to forced historical simulations. For this reason we assess the magnitude of the pre-industrial drift in all variables analysed in this chapter. If it is small enough compared with the magnitude of the forced change and the range of internal variability, we consider it negligible; if not, we remove the drift from precipitation series and repeat our calculations.

Using the methodology described here we are able to obtain an estimate for the fraction of a given precipitation change associated with internal variability and the fraction related to a forced signal. What fraction of this forced signal is anthropogenic, however, cannot be assessed without analysing historical simulations including only anthropogenic or only natural forcings, as Knutson et al. (2013) do for regional surface temperature. The discussion about whether the forced signals found in this chapter have a natural or anthropic origin is beyond the scope of this thesis.

Example: forced signal and internal variability in two models

To illustrate our method to estimate the magnitude of the forced signal and the range of internal variability we analyse the 1902-2004 trend of annual mean surface air temperature averaged over the Euro-Mediterranean region (from now on *EuroMed* region, 10°W-40°E/30°N-75°N). We use two models with several historical simulations available: CSIRO-Mk3.6.0 (10 runs) and EC-EARTH (7 runs). In this example we choose to analyse temperature because, unlike precipitation, it exhibits a marked forced signal and its study will be more didactic.

To begin with, in Fig. 5.2 we represent together the time series of all historical simulations available for each model (grey lines). Black curves represent data filtered using a low-pass symmetric non-recursive filter with a cut-off period of 10 years (see Sec. 3.1.7). At first sight we see that the behaviour of the temperature in the two models is clearly different. The CSIRO-Mk3.6.0 low-pass filtered time series seems dominated by multi-decadal variability (Fig. 5.2a), whereas EC-EARH exhibits a clear warming signal (Fig. 5.2b). As explained above, we estimate the magnitude of the 1902-2004 forced signal in a given model as the mean trend across all available simulations ($\Delta T^H_{sig} \approx \overline{\Delta T^H}$). Here we

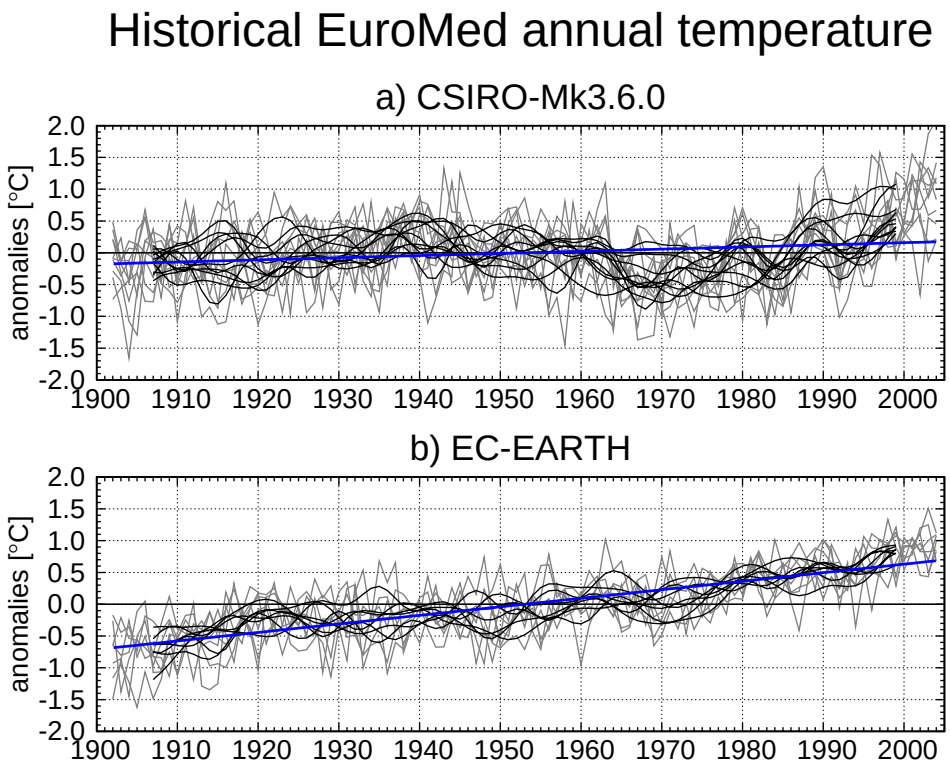


Figure 5.2: Anomalies of annual surface air temperature averaged over the EuroMed region relative to the 1902-2004 mean for all available historical simulations of the CSIRO-Mk3.6.0 (a) and EC-EARTH (b) models. Gray curves indicate raw data and black curves filtered data. Blue lines indicate the average linear fit.

represent the forced signal as the mean linear regression line (blue lines; 0.34 °C/century for CSIRO-Mk3.6.0 and 1.34 °C/century for EC-EARTH). Given that the prescribed external forcing is supposed to be the same in both models, the difference in the magnitude of the signal between the models must be associated with differences in their response to the forcing or to the fact that the number of available ensemble members is not enough to represent model noise.

In Fig. 5.3 we show in descending order the magnitude of the 1902-2004 EuroMed temperature trend in individual simulations of each model (black bars) together with the model means (blue bars) and the observed trend (red bars, estimated with HadCRUT data). For single runs and observations, solid bars indicate statistically significant trends (with a 95% confidence level evaluated using a Student's test taking into account the lag-1 autocorrelation of the series, see Sec. 3.1.6). For model means, the filled fraction of the bar is proportional to the percentage of runs with statistically significant trends.

As expected, in both cases EuroMed temperature trends exhibit a certain spread around the model mean as a result of internal variability. From now on we will refer to this spread in historical trends as the *historical spread* and it will be defined as half the difference between the maximum and the minimum value. While this spread in CSIRO-Mk3.6.0 is about 0.5 °C/century, in EC-EARTH it is only about 0.2 °C/century. Discrepancies in the

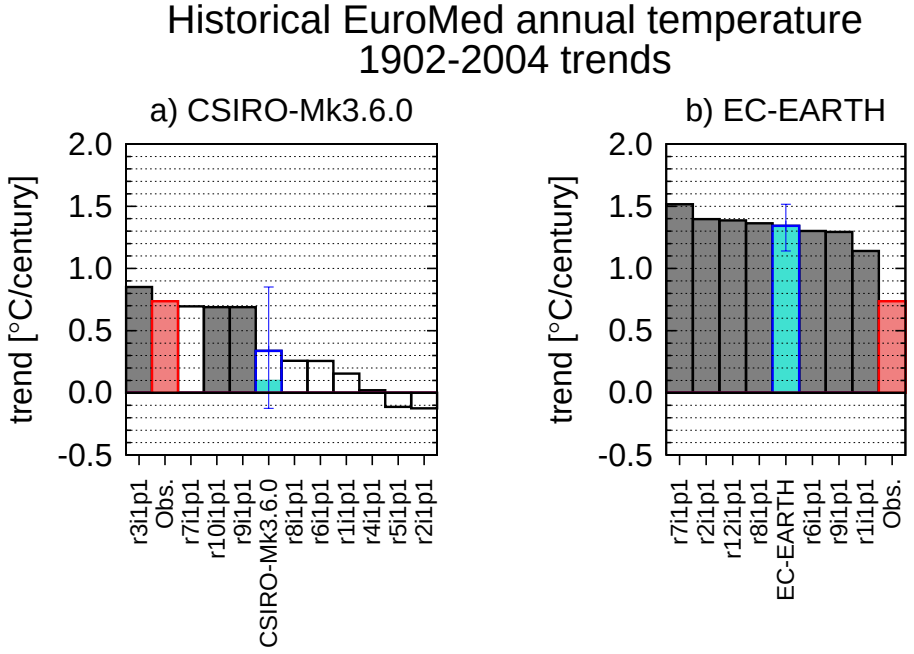


Figure 5.3: Linear trend (1902-2004) of annual surface air temperature averaged over the EuroMed region in all available historical simulations of the CSIRO-Mk3.6.0 (a) and EC-EARTH (b) models. Blue bars indicate model means and red bars observed trends (HadCRUT). For single runs and observations, full bars indicate statistically significant trends with a 95% confidence level. For model means, the filled fraction of the bar is proportional to the number of runs with statistically significant trends. The blue whisker indicates the minimum and the maximum trends.

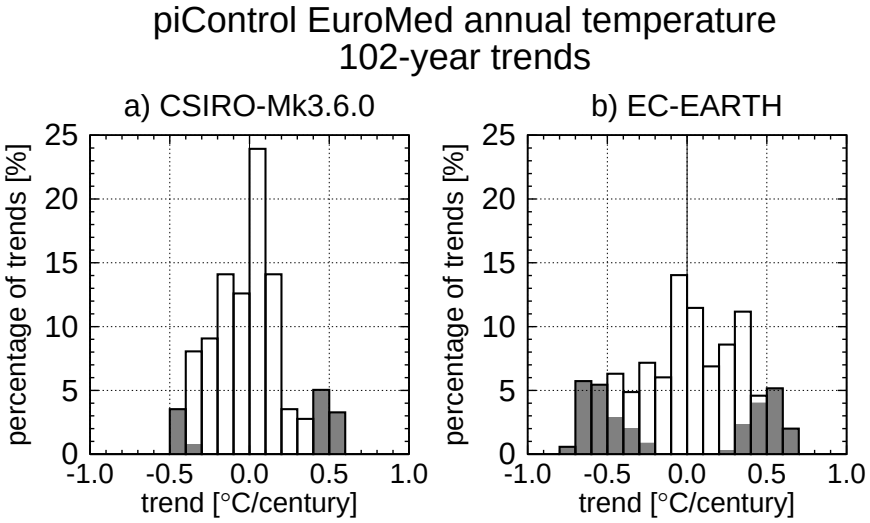


Figure 5.4: (a) Distribution of all 102-year EuroMed annual temperature trends present in CSIRO-Mk3.6.0 pre-industrial control simulations. Shaded boxes indicate statistically significant trends. (b) The same for EC-EARTH.

estimation of the historical spread are in part due to intrinsic model differences and in part due to uncertainties in the estimation due to the limited number of runs available.

To increase the number of samples in the estimation of the range of internal variability and determine which of the two explanations is more plausible we use pre-industrial control simulations, calculating all possible 102-year trends (regardless of the initial and final years) in each model's pre-industrial control run. In Fig. 5.4 we represent the obtained distribution of trends. The half-width of this distribution will be called from now on the *pre-industrial spread*. Shaded boxes include only statistically significant trends[1].

Comparing Fig. 5.4b with Fig. 5.3b we see that for EC-EARTH the pre-industrial spread (0.68 °C/century) is substantially larger than the historical spread (0.19 °C/century). Presumably, this is not due to a substantial change in the internal variability between pre-industrial and historical periods, but to the larger number of samples. For CSIRO-Mk3.6.0, instead, the spreads are similar (0.54 °C/century and 0.49 °C/century). In Fig. 5.4 we also see that the pre-industrial spread is larger in EC-EARTH than in CSIRO-Mk3.6.0. This result illustrates that each model has its own intrinsic range of internal variability as a result of its internal physics and parametrizations. Because in nature we only dispose of

[1]Note that the fraction of statistically significant trends is surprisingly large for pre-industrial control simulations which should only include noise. This overestimation of statistical significance is associated with the large autocorrelation present in temperature series. While we have taken into consideration the lag-1 autocorrelation in the significance test (see Sec. 3.1.6), the influence of autocorrelation is clearly not totally removed. Nevertheless, as temperature series are only used as an example and we do not draw conclusions from them, we do not go further on the correct estimation of statistical significance in autocorrelated series. Precipitation series are barely autocorrelated and, therefore, do not present this drawback.

one *realization*, it is not possible to assess which model estimates better the *real* internal variability. Given that the range of internal variability is better estimated with pre-industrial simulations than with historical simulations, from now on we will only use the pre-industrial spread and we will refer to it as simply *natural spread*. Because in general the pre-industrial spread is wider than the historical spread, with this approach we obtain an upper bound for internal variability.

In order to assess where the observed 1902-2004 EuroMed annual temperature trend (0.74 °C/century) is found with respect to the natural and forced distributions, in Fig. 5.5 we represent together the distribution of 102-year EuroMed annual temperature trends obtained with pre-industrial control simulations (magenta) and the distribution of 1902-2004 historical trends (blue). For the later, instead of using the trends obtained with individual historical runs, we create a new distribution combining the mean forced signal (estimated as the model mean trend) and the natural spread obtained with pre-industrial control runs. In both cases the observed trend falls inside the forced distribution and outside the natural distribution (green region described in Sec. 5.1). Therefore, according to these two models, a warming signal is detectable in 1902-2004 EuroMed annual temperature and, moreover, it is attributable to an external forcing. Note that, for EC-EARTH, if we had considered the historical spread as the range of trends obtained with historical simulations (see Fig. 5.3b), the conclusion would be different because, with a narrower range of internal noise,

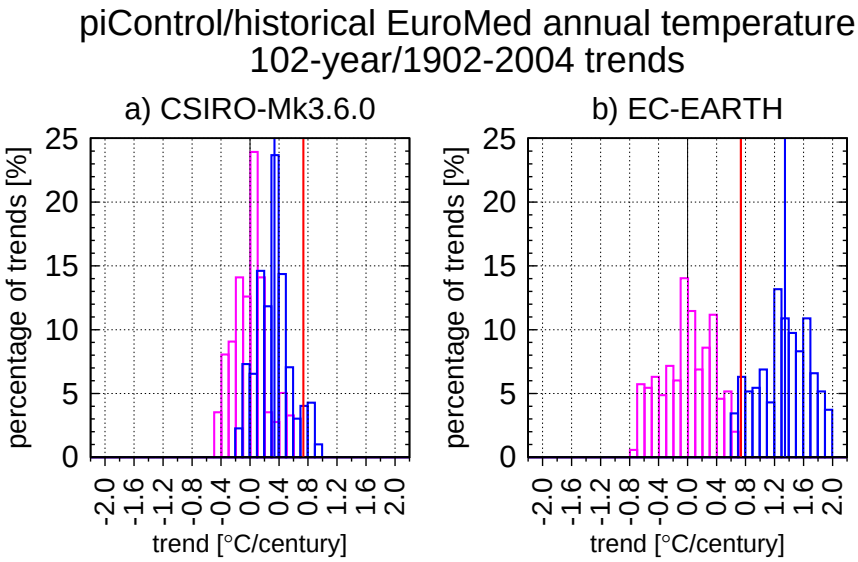


Figure 5.5: (a) Distribution of 102-year trends only due to internal variability (natural distribution; magenta), and with the distribution of trends under the influence of the 1902-2004 forced signal (forced distribution; blue) for CSIRO-Mk3.6.0 EuroMed annual temperature. The blue vertical line indicates the model mean and the red vertical line the observed trend obtained with HadCRUT data. (b) The same for EC-EARTH.

the observed trend would fall outside the distribution of forced trends. This methodology, then, is benevolent with climate models. We will keep this in mind when analysing results for precipitation in following sections.

To make sure that our conclusions are not affected by model drifts, following Knutson et al. (2013) we calculate the linear trend of EuroMed annual temperature for each control run along its entire length. Choosing the longest possible time intervals maximizes the signal-to-noise ratio of the trend and, therefore, gives a magnitude as unperturbed by internal variability as possible. As a result, the obtained drift is -0.003 °C/century for CSIRO-Mk3.6.0 and -0.021 °C/century for EC-EARTH. Since these values are very small compared with the forced signals and the range of internal variability, we consider them negligible and skip the drift removal.

Example with all CMIP5 models

Ideally, we would repeat this analysis for each CMIP5 model and we would discuss how many and which models support the different conclusions listed in Sec. 5.1. Unfortunately, for most climate models the number of historical simulations available is very limited (see table 2.2). Moreover, pre-industrial control simulations are not available for all CMIP5 models. As an alternative, we use the whole CMIP5 multi-model ensemble as a *megamodel*.

There are two main ways of working with such megamodel. The first choice is to consider all 179 available simulations as equal members of the CMIP5 megamodel. This approach is useful because it provides a large number of historical samples. Nevertheless, to consider all these simulations comparable we must assume the hypothesis that all climate models share the same forced signal ΔP^H_{sig} and the same range of internal variability ΔP^H_{int}, which we just showed not to be true for the models used in the previous example. Another commonly used approach is to weight the contribution of each simulation by the inverse of the total number of simulations of the same model. In other words, this methodology consists of first performing the analysis separately for each model and then averaging the obtained results. Consequently, we obtain measures of the multi-model mean forced signal ΔP^{HM}_{sig} in which each model contributes equally. Here we combine both approaches, indicating which one we are following in each case and checking that results are not sensitive to the choice.

We begin the analysis of the complete CMIP5 ensemble by representing the evolution of the EuroMed annual temperature during the 1902-2004 interval in all available historical runs (Fig. 5.6). Even though there is large inter-model variability, a superimposed warming trend is easily detectable. Moreover, the warming seems to increase its rate in the last decades of the twentieth century.

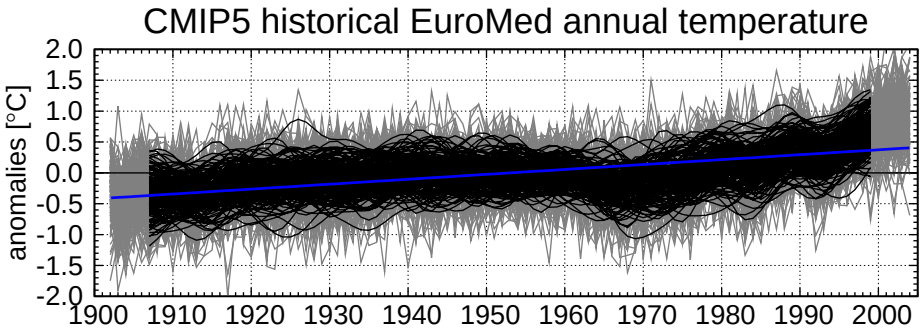


Figure 5.6: Anomalies of annual surface air temperature averaged over the EuroMed region relative to the 1902-2004 mean for all available historical simulations of the CMIP5 multi-model ensemble. Gray curves indicate raw data and black curves filtered data. The blue line indicates the multi-model average linear fit.

To estimate the magnitude of the mean forced signal ΔT_{sig}^{HM} in the 1902-2004 interval we first calculate ΔT_{sig}^{H} for each single model by averaging the 1902-2004 EuroMed annual temperature trend of all its simulations. Note that, the larger the number of simulations available for a given model, the better the achieved estimate of its forced signal ΔT_{sig}^{H}. The obtained trends are shown in Fig. 5.7 in descending order (grey bars), together with the multi-model mean (blue bar) and the observed trend (red bar). The filled fraction of each bar indicates the fraction of individual runs of the model that exhibit statistically significant trends. The magnitude of the multi-model mean forced signal is 0.80 °C/century, with an historical spread of about the same magnitude.

To obtain the natural spread associated with 1902-2004 EuroMed annual temperature trends we calculate all possible 102-year trends in all pre-industrial control simulations. In order to give to each model the same weight in the estimation of internal variability, the bin contribution of the trends of a given model to the multi-model distribution has been weighted by the inverse of the number of possible 102-year trends within its pre-industrial series. In Fig. 5.8 we represent the obtained natural distribution, in which we see that the magnitude of the pre-industrial spread is about 0.9 °C/century.

Having estimated the mean forced signal and the spread in the internal changes, in Fig. 5.9 we represent together the distribution of 1902-2004 EuroMed annual temperature trends under the influence of forcing (forced distribution; blue boxes) and the distribution of 102-year trends with no external forcing (natural distribution; magenta boxes). The blue vertical line indicates the multi-model mean and the red vertical line the observed trend. As in the previous example, the blue distribution has been created by adding the forced mean signal (i.e. the multi-model average trend) to the distribution of trends in pre-industrial control simulations. Because the observed trend falls within the forced distribution (very close to the multi-model mean) and outside the 95% central fraction of the natural distribution, we conclude again that according to the CMIP5 ensemble a warming signal is detectable

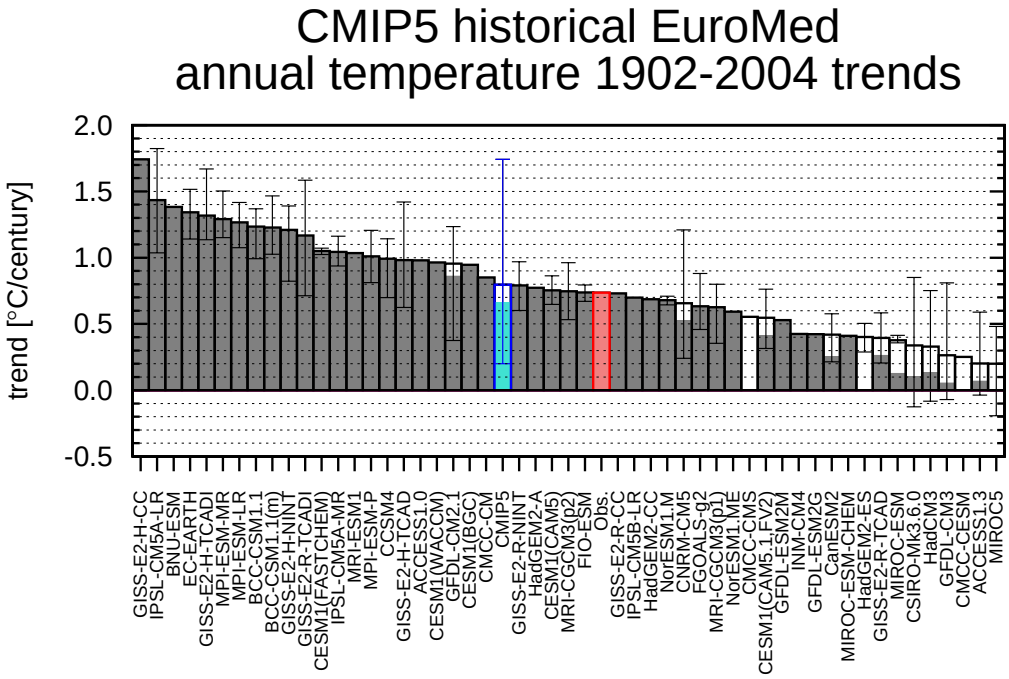


Figure 5.7: Linear trend (1902-2004) of annual surface air temperature averaged over the EuroMed region in all CMIP5 models. The blue bar corresponds to the multi5 mean and the red bar to the observed trend (HadCRUT). For models with several runs, the mean trend is represented, the filled fraction of the bar is proportional to the number of runs with statistically significant trends, and black whiskers indicate the minimum and the maximum trend in individual runs. The blue whisker indicates the minimum and the maximum model-mean trends.

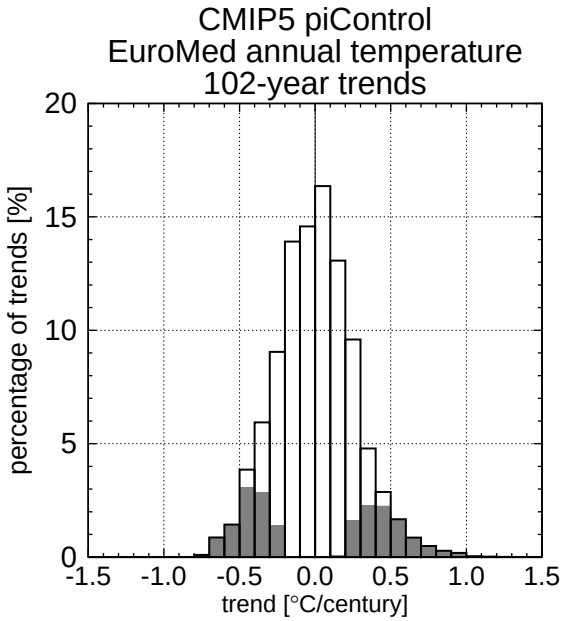


Figure 5.8: Distribution of all 102-year EuroMed annual temperature trends present in the CMIP5 multi-model ensemble pre-industrial control simulations. Shaded boxes include only statistically significant trends.

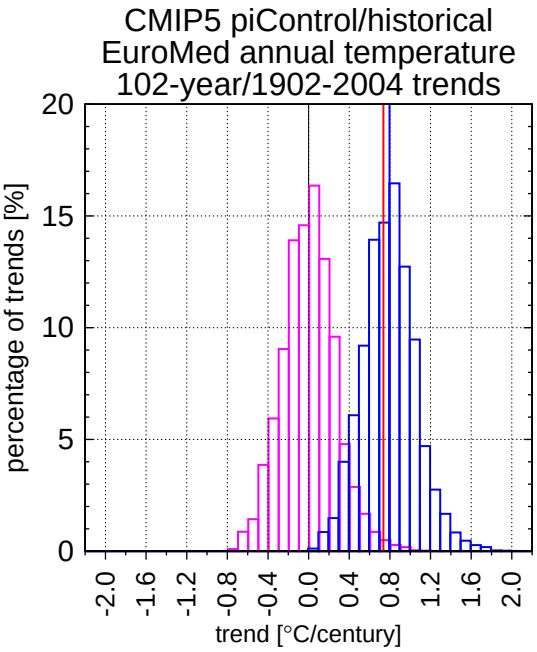


Figure 5.9: Distribution of 102-year trends only due to internal variability (natural distribution; magenta), and with the distribution of trends under the influence of the 1902-2004 forced signal (forced distribution; blue) for CMIP5 EuroMed annual temperature. The blue vertical line indicates the model mean and the red vertical line the observed trend obtained with HadCRUT data.

in 1902-2004 EuroMed annual temperature and it is attributable to an external forcing. We have checked that the range of pre-industrial drifts is small enough in comparison with the obtained forced and internal values (not shown) that a drift removal would not change our results. Just to illustrate the strength of the forced signal in temperature, we also present the distributions obtained for 1902-2004 surface air temperature trends averaged over the entire northern hemisphere (Fig. 5.10), for which the distributions barely overlap.

5.2 Northern Mediterranean climatology

Before using climate models to analyse precipitation changes in the northern Mediterranean, we must verify that they correctly reproduce the climatology of this region. For this purpose we calculate the annual cycle of land precipitation averaged over the 10°W-40°E/36°N-40°N region (hereafter *NMedLand* region) during the 1902-2004 period using the CMIP5 multi-model ensemble mean. We also analyse models from the old CMIP3 multi-model ensemble [Meehl et al., 2007] (see Sec. 2.2) to assess how conclusions change with the new models and to reproduce results from other studies. We refer to the multi-model ensemble mean created with all CMIP5 models (averaging first all individual runs for each model) as *multi5* mean and to the one created with CMIP3 models as *multi3* mean.

The annual cycle obtained with the multi5 mean (Fig. 5.11a), featuring wet winters and

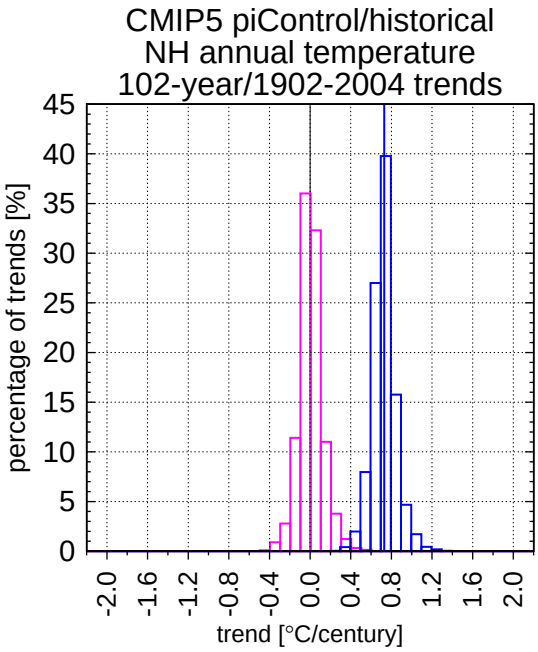


Figure 5.10: Distribution of 102-year trends only due to internal variability (natural distribution; magenta), and with the distribution of trends under the influence of the 1902-2004 forced signal (forced distribution; blue) for CMIP5 annual temperature averaged over the northern hemisphere. The blue vertical line indicates the model mean and the red vertical line the observed trend obtained with HadCRUT data.

dry summers (July and August are particularly dry), correctly reproduces the observed one calculated with GPCC data for the same period (blue squares, very similar to the annual cycle calculated with E-OBS data in 1950-2013, see Fig. 1.7). The multi3 mean exhibits the same annual cycle (Fig. 5.11b), even though its climatological mean precipitation is slightly weaker than both multi5's and observations in all months (in agreement with results by Giorgi and Lionello (2008), their Fig. 9, who report a small negative bias in precipitation in CMIP3 models).

We also represent the spatial distribution of climatological annual mean precipitation across the basin in models, averaging multi5 precipitation at each grid-point in the 1902-2004 interval (Fig. 5.12a). In general, the multi-model ensemble mean captures well the large scale pattern of observed climatological precipitation (Fig. 5.12c, calculated with GPCC data for the same period). In particular, the meridional gradient of precipitation (less precipitation in the south, more in the north) and the west-east gradient in Iberia and the Balkans are well reproduced. Due to the coarse resolution of climate models, however, observed local maximums of orographic origin, in particular the Balkans, the Pyrenees, and the Appenines are not fully captured by climate simulations. This is also true for the observed precipitation maximum located in the north-western Iberian coast, which is correctly located but too weak. CMIP5 models, which in general have higher resolution,

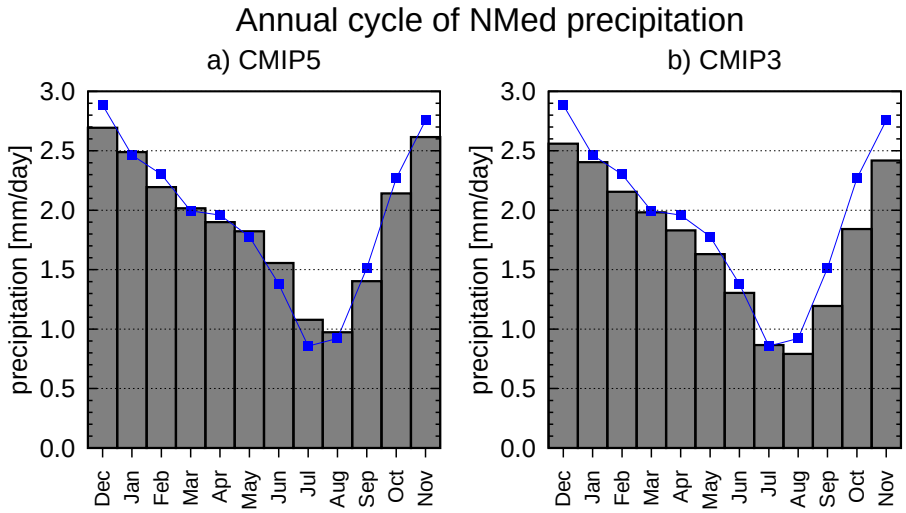


Figure 5.11: (a) Annual cycle of precipitation averaged over the NMedLand region (1902-2004) obtained using the multi5 (a) and the multi3 (b) ensemble means. The blue squares represent the observed monthly climatology of the GPCC data in the same period.

do a better job than CMIP3 models (Fig. 5.12b) in those regions. This improvement is also visible in the climatology calculated at a seasonal scale (not shown).

5.3 Seasonal distribution of precipitation changes

In Chap. 4 we found that only in winter does NMedLand precipitation exhibit coherent precipitation multi-month decreases (see Fig. 4.8). In this section we evaluate the seasonal distribution of NMedLand precipitation changes in historical simulations to compare them with observations. Following the same methodology used in Sec. 4.3, we compute the 2DMDA diagram of multi5 NMedLand precipitation for each month during the 1902-2004 interval (Fig. 5.13). The multi-diagram shows that all months except January exhibit a gradual precipitation decrease, as there is a continuous transition from wet to dry conditions at all time scales. In winter months (from December to March), however, anomalies are generally not statistically significant.

Since precipitation decreases gradually in all months (except for January), we calculate monthly trends using the longest time interval 1902-2004 (Fig. 5.14a). First, we see that the largest long-term trends are found in May-June and October. Kelley et al. (2012a), however, found the largest decreases (estimated for the 1950-2004 interval) during the traditional spring (March-May). As seen in Fig. 5.13, the use of a different time interval is not the source of the discrepancy (for time intervals beginning in 1950 the decrease in traditional summer is clearly stronger than in spring). The disparity between results must come from the fact that they average over the entire Mediterranean or that Kelley et al. (2012a) use a smaller subset of the CMIP5 ensemble.

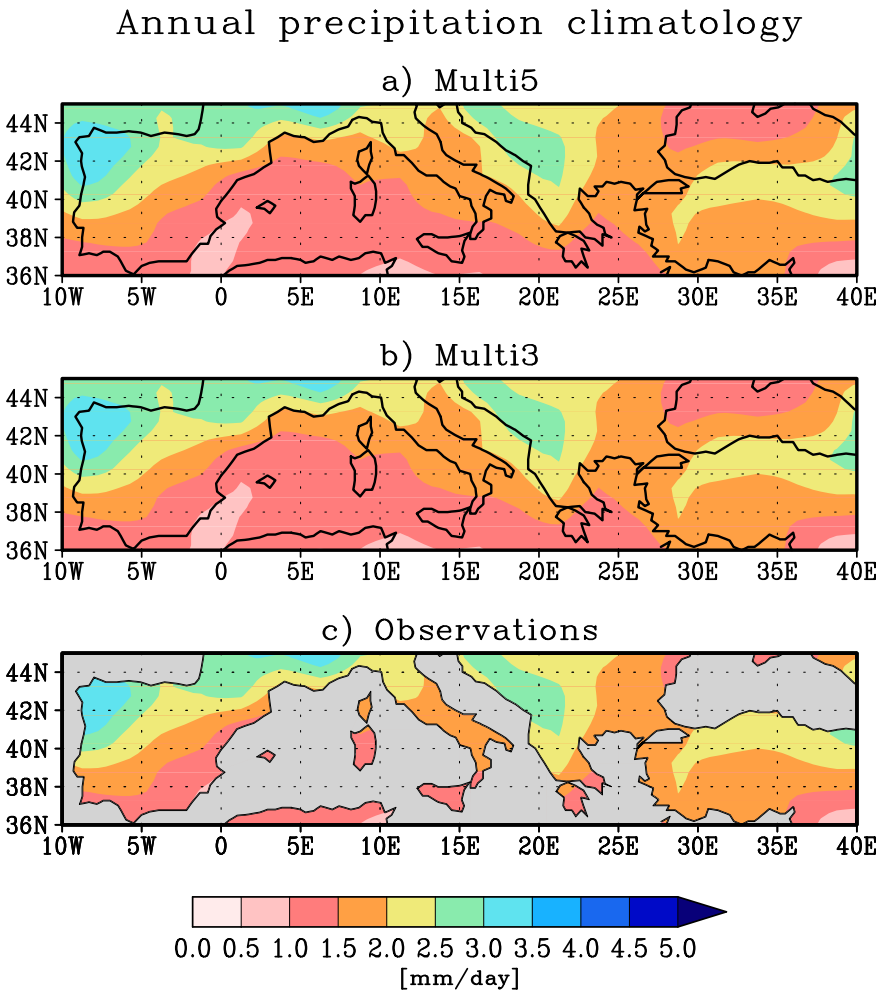


Figure 5.12: Climatology of annual mean precipitation (1902-2004) obtained using the multi5 (a) and the multi3 (b) ensemble means and the GPCC observational data set (c).

Contrasting with the strong summer trends, simulated winter long-term trends are weak and not even statistically significant in January (indicated by a red asterisk). Contrary to observations, then, the annual drying in historical simulations is barely influenced by winter precipitation trends. This conclusion is consistent with results by Pan (2013) and Mariotti et al. (2015), who conclude that the annual decrease of precipitation in CMIP5 models is mainly due to precipitation changes in the northern half of the Mediterranean (over 38°N) whereas it is dominated by winter precipitation changes in the southern half.

Because NMedLand precipitation exhibits a strong seasonal behaviour (see Fig. 5.11), drying trends of the same magnitude imply different relative changes in each time of the year: precipitation reductions in summer (with small rainfall amounts) have larger consequences than in winter (when the precipitation flux is largest). To illustrate this we calculate relative monthly trends by dividing each model's absolute trend by its 1902-2004 monthly climatology. In this way we can see that high-summer months (July and August) exhibit

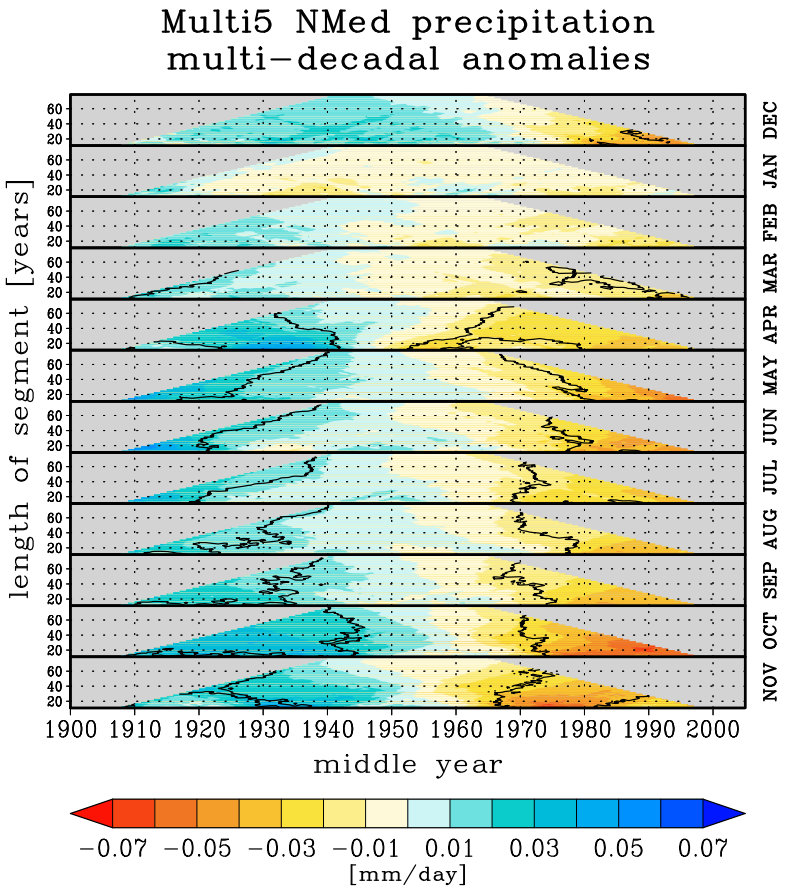


Figure 5.13: 2DMDA diagram (see Sec. 4.1) for each month using multi5 mean precipitation averaged over the NMedLand region. Anomalies are relative to the 1902-2004 mean. See Sec. 4.3 for a complete explanation of how this diagram is created.

the largest relative decreases.

Using multi3 historical simulations we find the same seasonal distribution (Figs. 5.14c-d) but trends are stronger in all months but October. In the cold half of the year, discrepancies between the two ensembles could be associated with differences in simulated NAO trends (we check this in Sec. 5.4.3). Instead, in the warm half of the year discrepancies between ensembles could be associated with the stronger Mediterranean temperature trends found in CMIP3 models (not shown). As proposed by Rowell and Jones (2006), an increase in summer temperature could contribute to the depletion in precipitation through soil moisture feedbacks. The fastest warming over land than over sea (associated with the thermal inertia of the water) also implies that the saturated vapour pressure over land increases faster than the amount of water vapour supplied by the sea, further reducing precipitation over land. Using these arguments, a stronger warming in CMIP3 models could cause a larger drying in comparison with CMIP5 models. This relation, however, is very complex and should be studied carefully, which is beyond the scope of this thesis.

The seasonal distribution of long-term trends in historical simulations, then, does not

match that of observations. In the following sections we examine separately precipitation changes in winter (the season in which observed precipitation changes are strongest) and summer (the season in which we find the strongest trends in historical simulations). First, in Sec. 5.4 we focus on the study of historical trends in winter and examine whether historical simulations are consistent with observed changes described in Chap. 4. For consistency with that chapter, we define winter as the December-March mean. Next, in Sec. 5.5 we describe the strong summer trends present in historical simulations. To be able to compare our results with those obtained in Bladé et al. (2012a) and Bladé et al. (2012b), from now on we define summer as the July-August average.

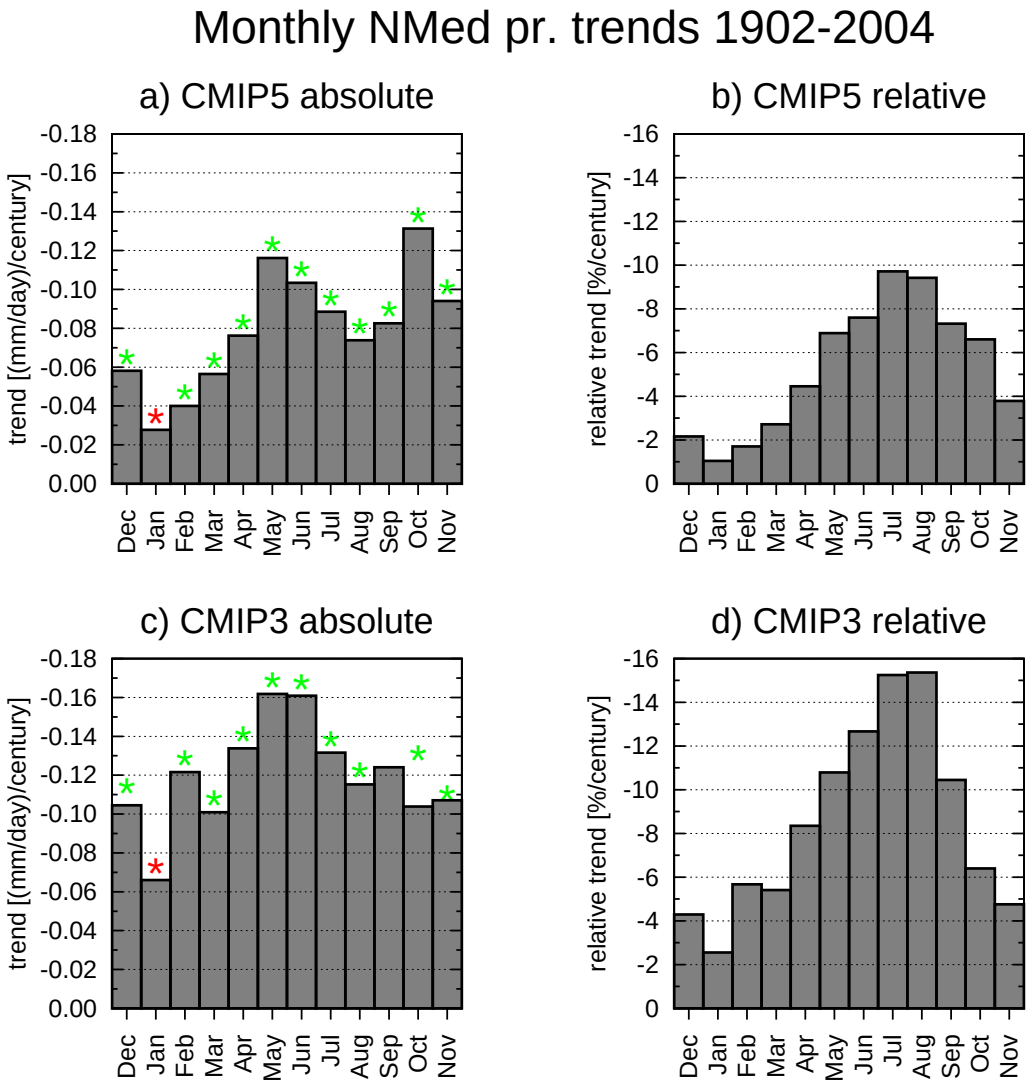


Figure 5.14: Monthly 1902-2004 NMedLand precipitation trends for CMIP5 (a) and CMIP3 (c) multi-model ensembles. Green asterisks indicate statistically significant trends; red asterisks indicate non-statistically significant trends. (b) and (d): Multi-model average of relative trends with reference of each model's 1902-2004 climatology.

5.4 Trends in winter

The aim of this section is to assess whether the strong observed winter precipitation trends are compatible with the weak simulated trends (see Sec. 5.1). We use two complementary approaches, first analysing winter precipitation itself and then analysing the winter NAO (hereafter only NAO), as it is the main driver of winter precipitation in the northern Mediterranean [Hurrell et al., 2003].

5.4.1 Winter precipitation trends

As seen in Sec. 5.1, we can express precipitation changes in historical simulations as $\Delta P^H = \Delta P^H_{sig} + \Delta P_{int}$, where ΔP^H_{sig} describes the forced part of the trend, evaluated as the multi-model mean change $\overline{\Delta P^{HM}}$, and ΔP_{int} the range of internal variability, which can be estimated either with historical simulations (assuming that all models share the same forced signal) or with control simulations (assuming that internal variability has not changed since pre-industrial times). In Fig. 5.15a we represent together the 1902-2004 NMedLand winter precipitation time series of all available historical runs listed in table 2.2 (grey lines; black lines represent filtered data). The red curve indicates the observational filtered series (GPCC). Unlike in case of EuroMed annual temperature (see Fig. 5.6), the long-term multi-model mean trend of NMedLand winter precipitation (blue line) is very weak in comparison with multi-decadal variability. Moreover, filtered anomalies in observations around 1990 are larger than those present in any CMIP5 climate simulation.

Exploring the evolution of the multi5 ensemble mean series during the historical period (Fig. 5.15b) we see that while in the first half of the 1902-2004 interval wet anomalies predominate (31 wet vs. 21 dry winters in 1902-1953), in the second half the situation is reversed (21 wet vs. 31 dry winters in 1953-2004). We also see that the reduction in multi5 NMedLand winter precipitation is quite uniform, without abrupt changes of rate like those present in observations (see Fig. 4.12). This is confirmed by the 2DLC diagram (Fig. 5.15d; see Sec. 4.1.1 for a complete explanation of this diagram), which illustrates that the magnitude and statistical significance of linear changes generally increase gradually with the length of segment. The 2DLC diagram also indicates that the drying signal is weak (in comparison with the noise that remains even after the multi-model average), because in order to find coherent statistically significant changes we must consider periods longer than 70-80 years. The weakness of the winter signal is also visible in the 2DMDA diagram (Fig. 5.15c; see Sec. 4.1.2 for a complete explanation of this diagram), in which a transition from wet to dry conditions is visible at all time-scales but very few of the multi-decadal anomalies are statistically significant.

As seen in Fig. 5.14, the multi3 ensemble mean exhibits larger trends than the multi5 ensemble in all months. Consistently, the 2DLC diagram (Fig. 5.16) shows that the winter

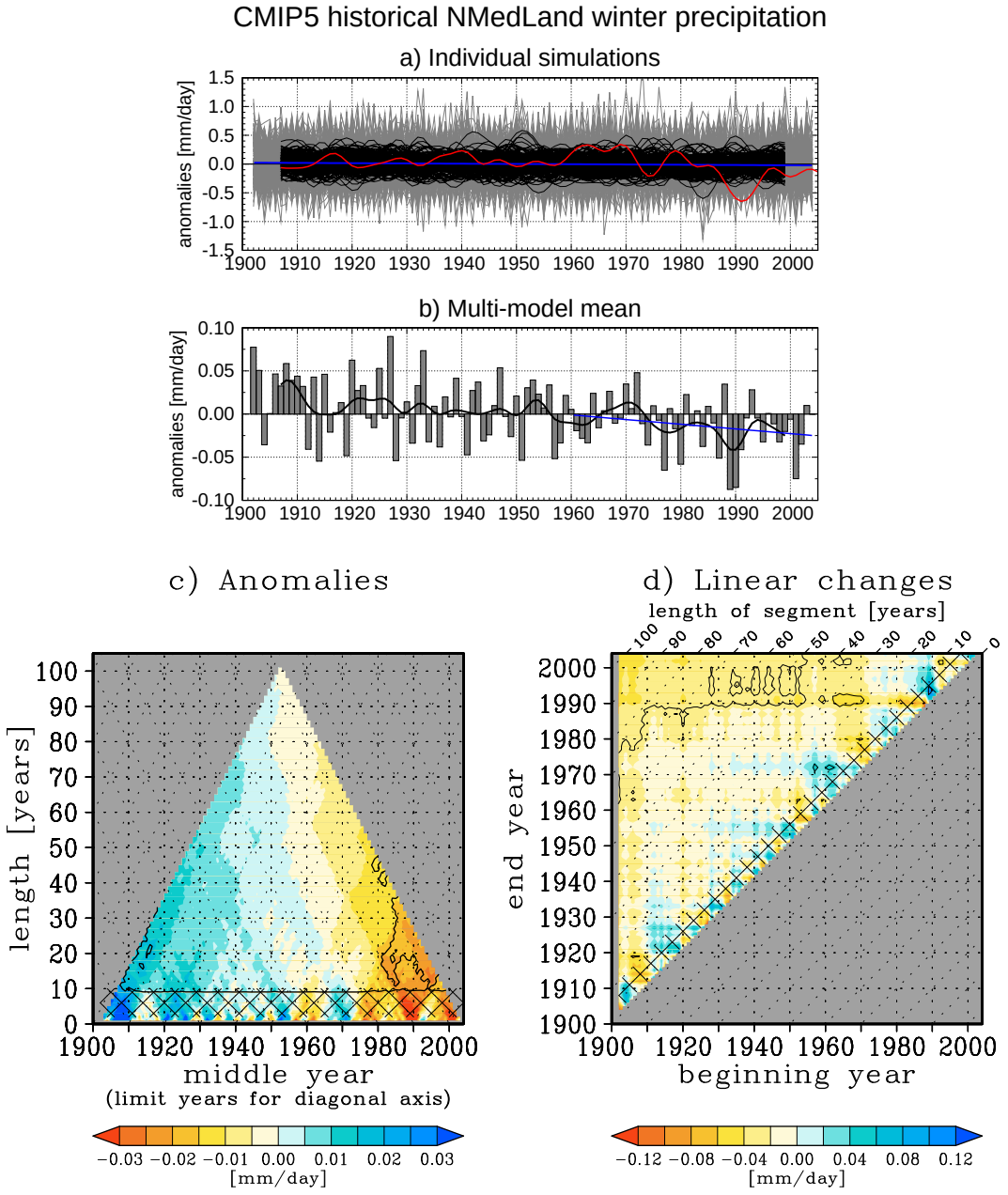


Figure 5.15: (a) NMedLand winter precipitation anomalies relative to the 1902-2004 mean in all the CMIP5 historical runs (grey curves). Black curves represent filtered data. The red curve represents observed precipitation anomalies (GPCC, filtered). The blue line represents the linear fit in the 1902-2004 interval. (b) The same for the multi5 ensemble mean. The blue line represents the linear fit in the 1960-2004 interval. (c-d) 2DMDA and 2DLC diagrams (see Sec. 4.1) for multi5 series.

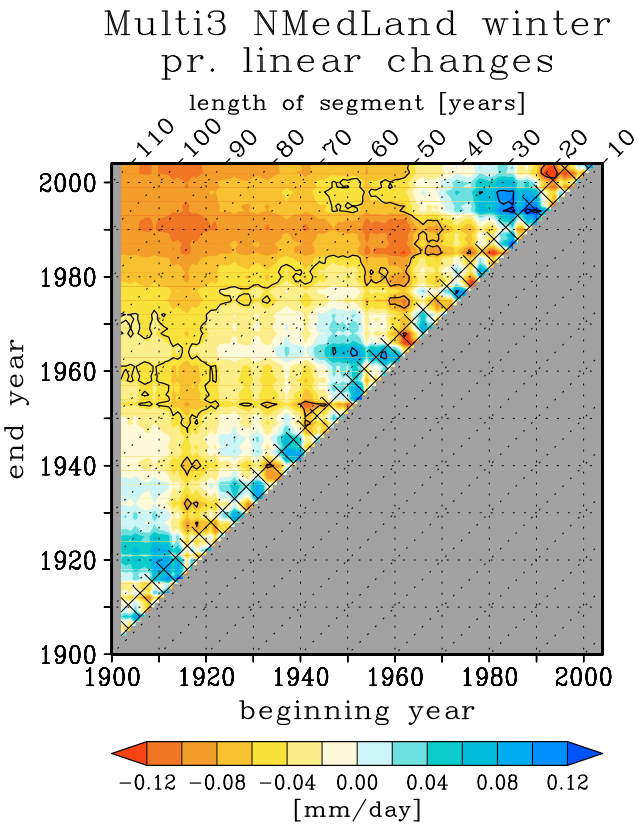


Figure 5.16: 2DLC diagram of multi3 NMedLand winter precipitation in the 1902-2004 interval. Contours indicate intervals with statistically significant trends.

signal in the multi3 ensemble mean is substantially stronger than in CMIP5, as the range of statistically significant intervals covers periods longer than 40-50 years.

5.4.1.1 1960-2004 area-averaged winter precipitation trends

In order to assess whether the strong observed winter precipitation changes are consistent with the simulated trends we select the time interval 1960-2004, the one ending in the last year of our historical period (2004) with the largest observed change. If observations and simulations are compatible for this time period we may assume that they would also be compatible in time intervals with a weaker observed change. In any case, later we check whether conclusions drawn for this period are sensitive to the choice of time interval.

In Fig. 5.17, in which we represent the magnitude of individual CMIP5 1960-2004 winter NMedLand precipitation trends ΔP^{HM} for individual models, we see that the historical spread is about 0.6 (mm/day)/century, about 12 times the multi-model mean trend (−0.05 (mm/day)/century). There is also a lack of inter-model agreement regarding the sign of the change, what indicates that the mean forced signal is very weak in comparison with inter-model variability. Moreover, the fact that very few runs exhibit statistically significant trends and that 31 out of the 33 models with more than one historical simula-

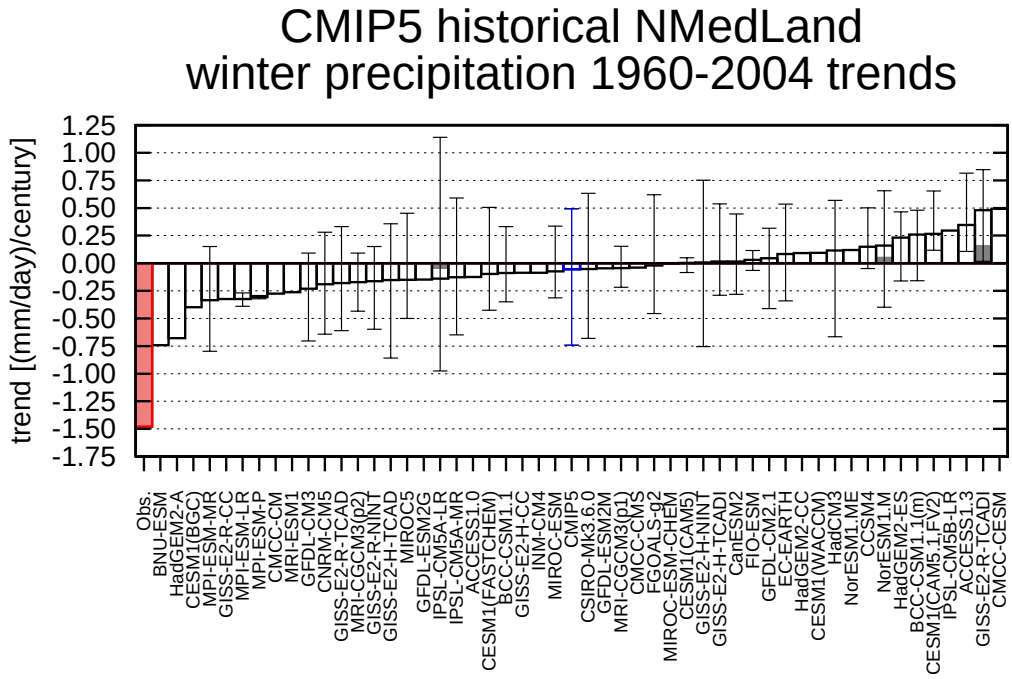


Figure 5.17: Linear trend (1960-2004) of winter precipitation averaged over the NMedLand region in all CMIP5 models. The blue bar corresponds to the multi5 mean and the red bar to the observed trend (GPCC). The filled fraction of each bar is proportional to the number of runs of the model with statistically significant trends. In those models with several runs, black whiskers indicate the maximum and the minimum trend in individual runs. The blue whisker indicates the minimum and the maximum model-mean trends.

tion available have runs with both positive and negative changes indicates that the forced signal is also weak in comparison with internal variability in single realizations. Presumably, the discrepancies in the range of internal variability displayed by models with more than one simulation (black whiskers) are not due to intrinsic differences between models but associated with the poor sampling caused by the low number of available runs. For this reason, we will use pre-industrial control simulations to evaluate the natural spread. We also see that the mean forced trend $\overline{\Delta P^{HM}}$ in the 1960-2004 interval is about 30 times smaller than the observed trend for the same period (-1.47 (mm/day)/century). This result indicates that, even if the forced signal was found compatible with the observed trend, its contribution would be very weak.

In Fig. 5.18 (magenta boxes) we represent the distribution of 44-year winter precipitation trends in pre-industrial control simulations (i.e. our natural distribution). Climate drifts in these simulations have been calculated and considered negligible (not shown). Even though, as shown in table 2.2, for some models we do not have pre-industrial simulations available, we consider that we have enough of them to sample the range of internal changes correctly. Comparing the natural distribution with the forced distribution (blue boxes; created adding the magnitude of the multi-model mean trend to each mem-

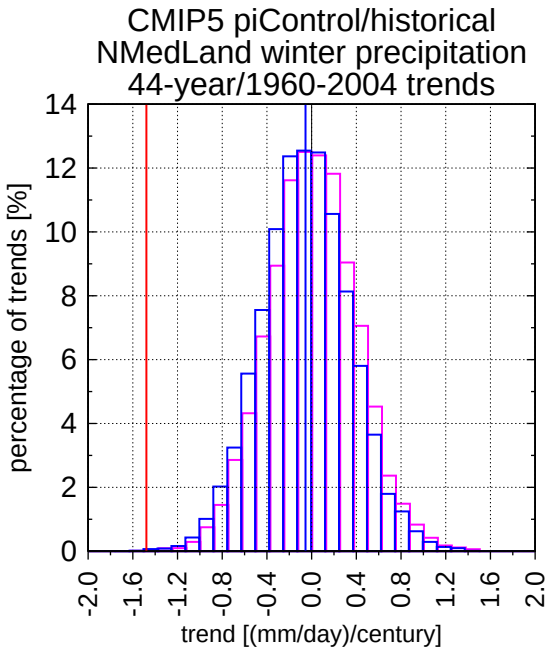


Figure 5.18: Distribution of 44-year trends only due to internal variability (natural distribution; magenta), and with the distribution of trends under the influence of the 1960-2004 forced signal (forced distribution; blue) for CMIP5 NMedLand winter precipitation. The blue vertical line indicates the model mean and the red vertical line the observed trend obtained with GPCC data.

ber of the natural distribution) we see again that the magnitude of the forced signal (−0.05 (mm/day)/century) is negligible in comparison with the width of the distribution of internal variability (∼2 (mm/day)/century) so that the distributions are very similar as in the idealized example in Fig. 5.1d. The fact that the observed trend (red vertical line in Fig. 5.18) is outside both the natural and the forced distributions (red region in Fig. 5.1) indicates that model simulations are not consistent with observations. If we create the forced distribution by calculating all possible 44-year trends in all available historical simulations we find a very similar result: there is no model simulation with any 44-year trend as large as the observed one in the 1960-2004 interval. We may only conclude, then, that either the forced signal is underestimated in historical simulations (because the external forcing is too weak or because the response to the forcing is too weak or not well captured), or the range of natural variability is underestimated by the models (see Sec. 5.1).

Our results generally agree with those by Barkhordarian et al. (2013) for CMIP3. Also examining the role of external forcings in observed winter precipitation changes, they calculate 1966-2005 Mediterranean precipitation trends for each overlapped 3-month season using precipitation averaged over the entire Mediterranean basin. As in our analysis, observed precipitation trends in the winter seasons are not within the range of trends calculated with pre-industrial simulations. Moreover, using the multi-model CMIP3 mean from models forced with only anthropogenic forcings, they find weak wetting instead of drying

trends in December-February and January-March. Given this difference of sign between the observed and the forced trends, they also conclude that observations are not consistent with forced simulations.

On the other hand, our results are in contradiction with Hoerling et al. (2012), who carry out an analysis similar to ours in order to assess whether recent precipitation changes in the Mediterranean region (evaluated as 1971-2010 minus 1902-1970 relative anomalies) are within the simulated range of variability using cold-season (November-April) precipitation in CMIP3 pre-industrial and historical simulations averaged over the entire Mediterranean area. With those parameters they find the observed signal to be within the range of the simulated forced trends, with the magnitude of the multi-model mean forced signal about one half of the observed change, and their natural and forced distributions statistically different at a 99.9% confidence level. Presumably, the discrepancy with our results comes from the fact that the old CMIP3 models exhibited a stronger forced signal, and could be also associated with the fact that they average precipitation over a larger area and using more months. Moreover, only 1 out of 22 historical simulations exhibits a precipitation decrease as large as or larger than the observed one, so their conclusion that the observed change is within the historical spread is marginal.

5.4.1.2 Decadal standard deviation

An approach to checking whether natural variability of NMedLand winter precipitation is underestimated in climate models is to compare the standard deviation present in pre-industrial control simulations with the observed one. To remove the influence of year-to-year variations, we apply a low pass filter with a cut-off period of 10 years (see Sec. 3.1.7) to the series. While pre-industrial control simulations provide a very good tool to evaluate the variance of precipitation in unforced conditions, the evaluation of the observed variance has some difficulties associated with the fact that observed records are short and potentially contaminated by external forcings [Imbers et al., 2014].

First, we cannot easily remove the part of the variance associated with the forced signal because we do not know its magnitude. Knutson et al. (2013) propose an approach that consists of removing the multi-model mean trend from the observed time series (as it is the best estimate of the magnitude of the forced change) prior to the calculation of the observed standard deviation. They then apply the same procedure to each individual historical simulation in order to estimate the *historical* internal variability. By calculating the difference between this *historical* standard deviation and the *pre-industrial* standard deviation estimated from control simulations, they obtain an estimate of the increase in internal variability between the two epochs. Removing this value from the corrected observed standard deviation they obtain an estimate of the *observed pre-industrial* standard deviation, which now can be compared directly with pre-industrial control simulations. In our case,

however, due to the fact that the forced signal in winter precipitation is negligible in comparison with the magnitude of the observed trend, the first correction does not produce any effect. Regarding the second correction, it is also small and we obtain very similar results when using historical simulations instead of the pre-industrial ones (not shown).

Second, the estimation of the observed standard deviation is very sensitive to the choice of time interval. To illustrate this sensitivity we describe the results obtained using three different time intervals. First, the entire historical interval 1902-2004 was used and the obtained detrended standard deviation was found larger than any in the pre-industrial distribution. However, this value is clearly affected by the strong negative trend that NMedLand winter precipitation exhibits during the second half of the 20th century (see Fig. 4.12), which is not fully removed by a linear detrending over the entire record. Then we used the 1902-1960 interval, for which the obtained standard deviation was found lower than that of any model. As winter precipitation did not exhibit significant trends in that period, this could have been the better choice if it were not for the fact that we have low confidence in data for the first half twentieth century (see Sec. 2.1.1). Finally, the standard deviation obtained for the 1960-2004 trend was again found larger than that of all pre-industrial simulations.

As results obtained with that methodology are interval-sensitive, we choose to apply a high-pass filter to the series in order to filter out observed trends more than 40-year long. The band-pass filter result of combining the 10-year low-pass filter and the 40-year high-pass filter (see Sec. 3.1.7), gives an estimation of the variance at a decadal and multi-decadal scale without being perturbed by long-term trends. In Fig. 5.19 we represent the obtained distribution of standard deviations. All time series have been detrended prior to the calculation.

The spread in Fig. 5.19 indicates that there is little agreement among models regarding the magnitude of the standard deviation, even though it appears that the simulated range of internal decadal variability is not widely different from the observed one (0.099 mm/day), as this is located in the center of the distribution. This conclusion, however, is not very solid due to the commented problems when estimating the magnitude of the observed variance.

5.4.1.3 Sensitivity to time interval

The 1960-2004 interval was intentionally chosen because it exhibits the largest observed NMedLand winter precipitation trend. It is necessary, therefore, to test whether the conclusion that model simulations are inconsistent with observations is valid for time intervals in which the observed trend is not so extreme. For this purpose, we calculate the natural and the forced distribution for each time period within the 1902-2004 interval. The natural distribution for a given period is obtained by calculating all possible pre-industrial precipitation trends for the corresponding interval length. Forced distributions are obtained by

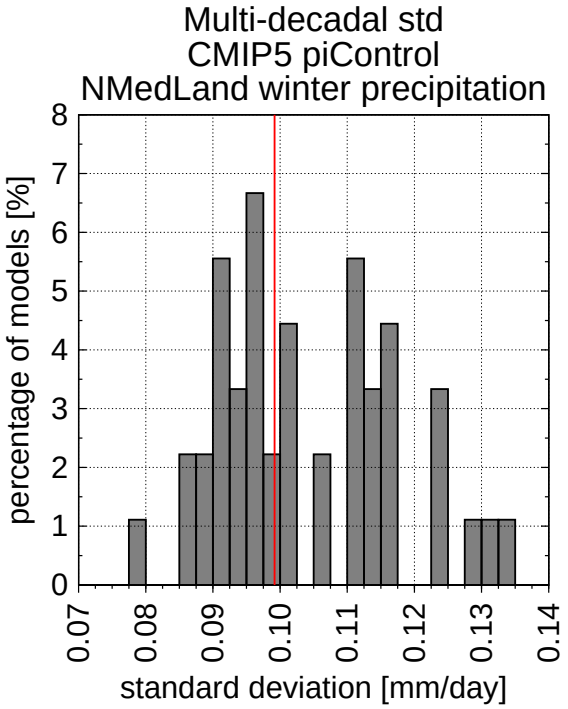


Figure 5.19: Distribution of standard deviation of CMIP5 pre-industrial precipitation series. All time series have been detrended prior the calculation and a band-pass symmetric non-recursive filter with cut-off periods of 10 and 40 years has been applied. The red vertical line indicates the observed standard deviation (GPCC).

adding the multi-model mean precipitation trend of each time interval (Fig. 5.15c) to the natural distribution of the appropriate length. We have checked that very similar results are found when estimating the natural and forced distributions using historical simulations (not shown), which adds robustness to our results.

As a measure of how both distributions are separated in each time interval, in Fig. 5.20b we represent the ratio between the magnitude of the forced signal and the half width of the natural distribution (defined as half the difference between the 97.5th and the 2.5th percentiles). For the period 1960-2004 studied before (see Fig. 5.18), in which both distributions are very similar, this magnitude is 0.064. Given that the signal-to-noise ratio never exceeds 0.25, we conclude that most part of the natural and the forced distributions are overlapped for all time intervals.

Then, we assess in which *region* of those described in Sec. 5.1 the observed NMedLand winter precipitation changes fall with respect to the two distributions for all intervals within the historical period (Fig. 5.20a). For most time intervals longer than 30 years beginning after 1920 and ending after 1990 (those in which the observed trend is largest), the conclusion drawn is the same as that for the 1960-2004 period: the observed NMedLand winter precipitation trend is not compatible with model simulations. For other time intervals within the historical period, the observed trend is within both the natural and the

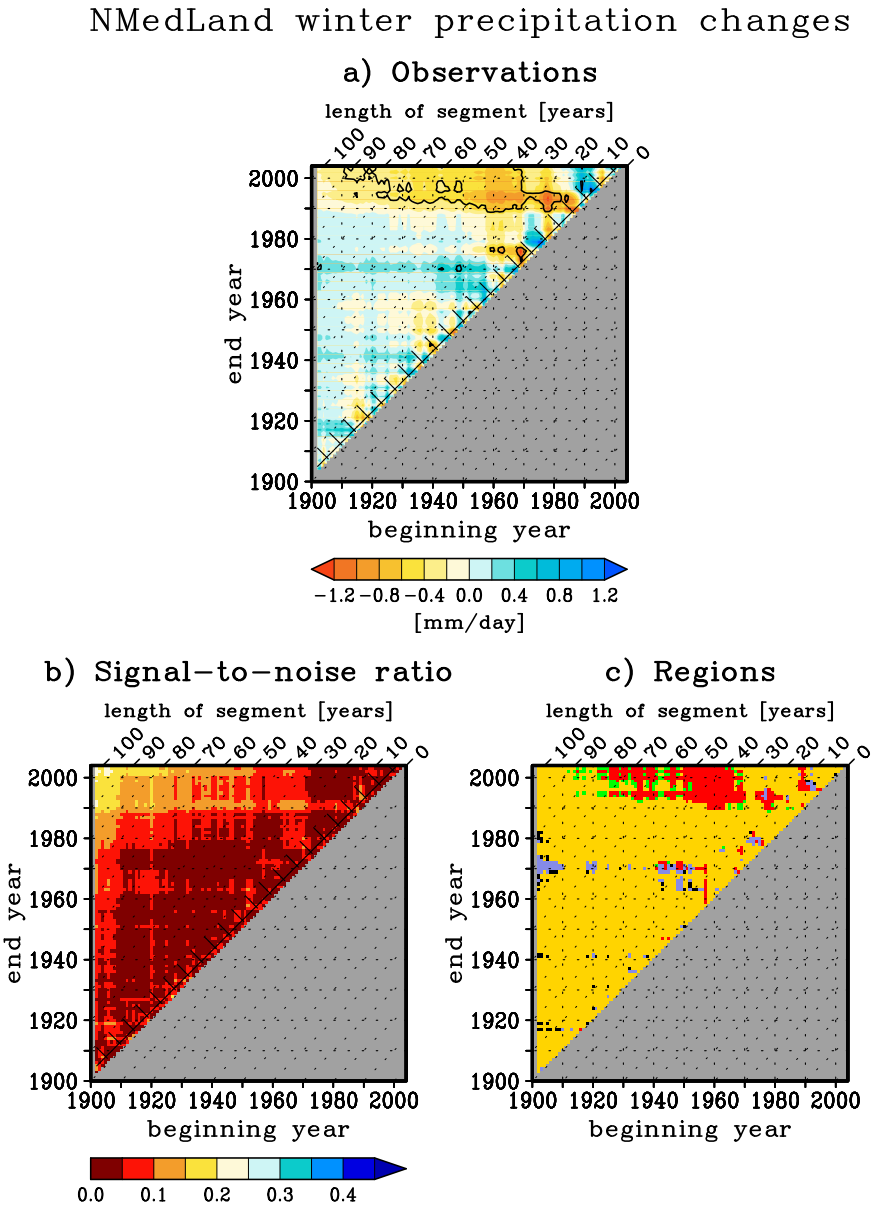


Figure 5.20: (a) 2DLC diagram (see Sec. 4.1.1) of observed (GPCC) NMedLand winter precipitation. Contours indicate statistically significant changes (b) Signal-to-noise ratio, defined as the ratio between the multi-model mean trend and the half-width of the distributions in each time interval. (c) Two-dimensional diagram in which we compare, for each possible time interval within the 1902-2004 period, the observed NMedLand winter precipitation trend with the natural distribution and forced distribution of models simulations. Colours correspond to the regions described in Sec. 5.1.

forced distributions and, therefore, simulations are compatible with observed changes but no signal is detectable. In conclusion, then, climate simulations are only inconsistent with observations in those time intervals in which the observed NMedLand winter precipitation trend is strongest.

5.4.1.4 Geographical distribution

Next, we evaluate whether the conclusions drawn for area-averaged NMedLand winter precipitation in the 1960-2004 interval are valid at a regional scale. First, as a measure of the forced signal we calculate the multi5 trend at each grid-point (Fig. 5.21b). At first sight we see that the pattern of simulated trends does not coincide with that of observed trends (Fig. 5.21a). In particular, the obtained geographical pattern exhibits a marked meridional gradient, with large negative trends in the south and positive trends in the north (with a strip of non-significant trends crossing the basin at central latitudes), not present in the observed pattern. Even in those regions where the sign of the multi5 and the observed trends coincide, the latter are much larger (note the $\times 10$ change of scale). The largest multi5 precipitation decreases are found in the south of the Italian, the Balkan and the Anatolian peninsulas and the south-east of Iberia. Instead, precipitation increases are particularly large in north-western Iberia. The fact that about half grid-points exhibit negative trends and the other half exhibit positive trends partly explains the weakness of the signal in the multi5 areal mean. Results are very similar when calculating trends for the entire historical period (1902-2004), for which the response is better estimated as the influence of internal variability is lower (not shown).

To evaluate to what extent different models agree in the sign of the change, in Fig. 5.21c we represent the portion of climate models in which the 1960-2004 trend is negative. The agreement regarding the sign of the change is low in the entire basin and only reaches a 70% in the southernmost and the northernmost parts of the basin. But, even in the regions were the agreement is largest, fewer than 15% of the single trends are statistically significant (not shown), indicating that in individual simulations the local forced signal generally does not stand above natural multi-decadal variability. Kelley et al. (2012a), using a smaller subset of CMIP5 models, find a similar pattern when calculating linear trends in cold season (November-April) precipitation during the 1950-2004 interval (their Fig. 3c).

Next, we calculate the amplitude of the natural distribution in each grid-point, estimated as half the difference between the 97.5th percentile and the 2.5th percentile of the local distributions of 44-year winter precipitation trends in pre-industrial control simulations. The multi-model average natural spread is represented in 5.22a. To avoid results to be influenced by regional drifts, the long-term winter precipitation trend in pre-industrial control simulations has been removed at each grid-point prior the calculations. The obtained pattern, with an absolute maximum in north-western Iberia and weaker relative maximums in western Balkans, south-western Anatolia and over the French Riviera, largely resembles the observed regression of winter precipitation over the NAO index shown in Fig. 4.25. This result is consistent with the fact that the NAO is the main source of winter precipitation variability in the northern Mediterranean and so is also the main source of spread in regional trends.

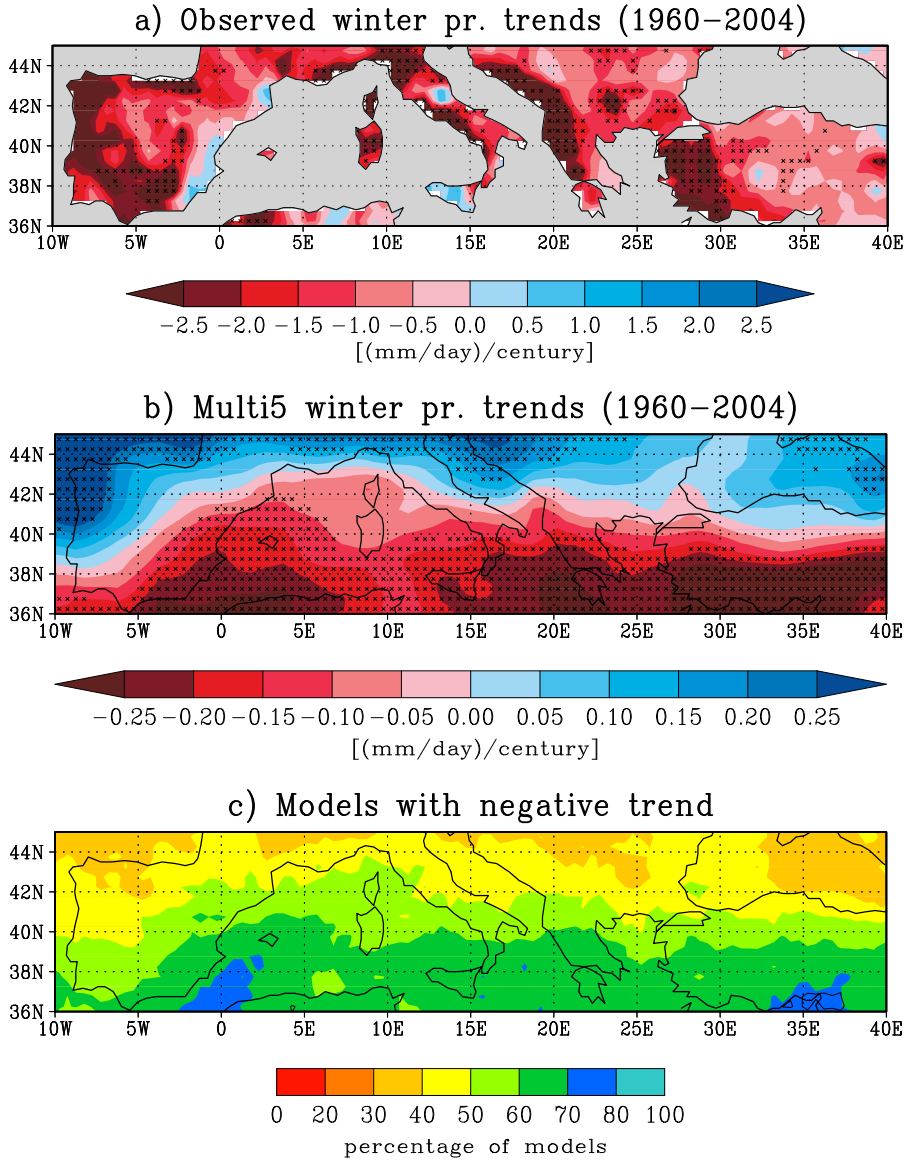


Figure 5.21: (a) Observed (GPCC) winter precipitation trends during the 1960-2004 interval. Crosses indicate statistically significant trends with a 95% confidence level. (b) The same for the multi5 ensemble mean (note the change of scale). (c) Percentage of climate models in the CMIP5 multi-model ensemble with negative trend at each grid-point.

Comparing Figs. 5.22a and 5.21b we see that the natural spread is one order of magnitude larger than the multi-model mean trend in all the basin. As a result, the signal-to-noise ratio (5.22b, evaluated as the multi-model mean trend divided by the natural spread), which follows a similar pattern than the multi-model mean trend, does not exceed 0.3 in almost any grid-point. This result indicates that, for the most part, the local forced and natural distributions overlap in all the basin.

In Fig. 5.22c we indicate where the observed local trend falls with respect to the simulated local distributions (see Sec. 5.1). In most parts of the basin the observed trend is inside both distributions, even in those regions where the observed trend has an opposite

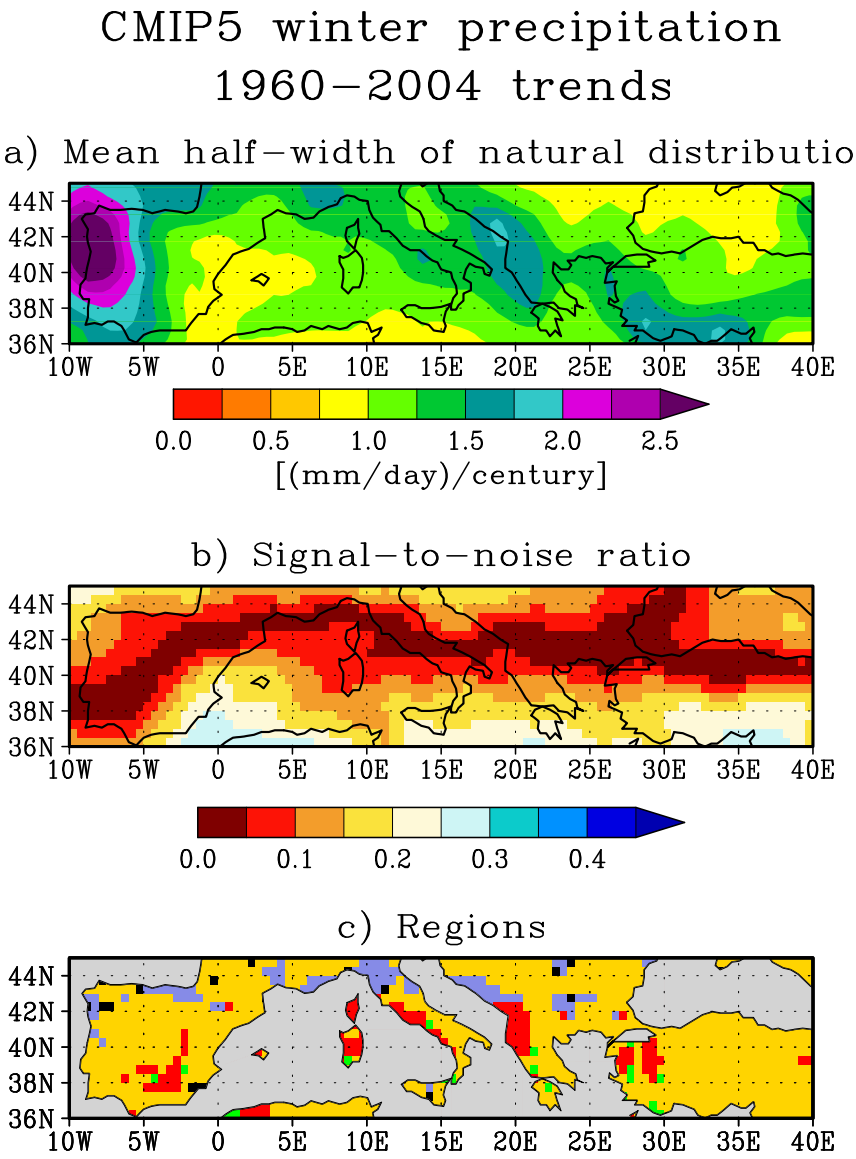


Figure 5.22: (a) Mean half-width of the local distribution of CMIP5 44-year winter precipitation trends associated only with internal variability, calculated first for each model and then averaged. Local distributions are estimated with all possible 44-year trends in pre-industrial control simulations. Half-width estimated as half the difference between the 97.5th and the 2.5th percentile of local distributions. (b) Signal-to-noise ratio of winter precipitation calculated dividing the 1960-2004 multi-model mean trend and the half-width of the natural distribution per grid-point. (c) Map in which we compare, at each grid-point, the observed 1960-2004 winter precipitation trend with the local natural and forced distributions. Colours correspond to the regions described in Sec. 5.1.

sign with respect to the multi-model mean, and therefore no drying signal is detectable. Only in regions in which the observed trends are strongest, which are the regions in which the response to the NAO is largest, is the observed trend outside the model distributions. These regions include the central/western Balkan coast, the west of the Italian Peninsula and some isolated regions in Iberia and Anatolia. The conclusion drawn for the areal mean

(i.e. that simulated 1960-2004 winter precipitation trends are not compatible with observations) is therefore not valid in most parts of the basin. This is consistent with the fact that the multi-model areal average include trends of both signs (see Fig. 5.22b) and therefore its magnitude is weaker than most local trends. This result is also consistent with the fact that the regional average filters part of the internal variability and makes the width of the areal-mean natural distribution smaller than that for individual grid-points.

5.4.2 Impact of the NAO on winter precipitation

In the former section we have found that the regions in which the observed trend is inconsistent with model simulations are those in which the response of winter precipitation to NAO changes is strongest. For this reason, in this section we assess whether the inconsistency between models and observations in these regions and in the areal-mean can be explained by inconsistencies in the simulated NAO trends. But first we must assess whether the observed linear relation between these variables is well captured by climate models. We calculate the correlation of NMedLand winter precipitation and the NAO index using each CMIP5 pre-industrial control simulation (see Sec. 3.2.2 for a description of how we calculate the NAO index in models). To average correlations, we apply a Fisher's Z transformation to each correlation value before the calculation and transform inversely the result afterwards (see Sec. 3.1.3). Each time series has been detrended prior to the calculation. Using historical simulations we obtain very similar results (not shown).

The histogram of correlations (Fig. 5.23a) shows that, even though the sign of the correlation is well captured by all the ensemble members, its magnitude is underestimated by practically all of them. While the observed correlation in the historical period is $r = -0.82$, the multi-model average is only $r = -0.63$ and more than a half of the simulations exhibit lower values (down to $r = -0.34$). The linear regression between the two variables is also underestimated by the vast majority of the climate models (Fig. 5.23b). While the observed linear relation is -0.33 (mm/day)/ostd, the multi-model mean is -0.23 (mm/day)/ostd and some models reach -0.13 (mm/day)/ostd.

In Chap. 4 the linear relation between NAO and winter precipitation was found statistically significant in most parts of the basin, with the few exceptions of the eastern Iberian coast, Sicily and some Anatolian regions (see Fig. 4.25). The largest negative regressions were found in north-western Iberia and over the western Balkan coast. To evaluate whether this pattern is reproduced by climate models, we calculate the mean CMIP5 linear regression at each grid-point by averaging the regression found for each pre-industrial control simulation (Fig. 5.24a; using historical simulations we obtain very similar results, not shown). We observe that the regression is negative in all grid-points and that its magnitude is particularly large in the west Balkans, the central and northern parts of the Italian Peninsula, the French Riviera and the central and western half of Iberia. The largest regressions

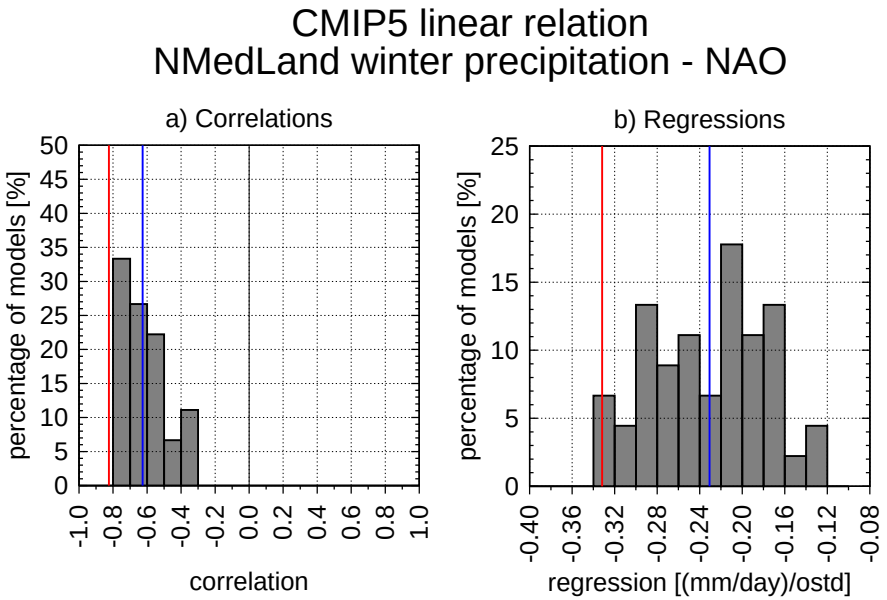


Figure 5.23: (a) Histogram of correlations (1902-2004) between NMedLand winter precipitation and the NAO index using pre-industrial simulations of all models in the CMIP5 ensemble. Time series have been detrended prior to the calculation. The red line indicates the observed correlation (GPCC) and the blue line the average of all CMIP5 models. (b) The same for the regression of NMedLand winter precipitation onto the NAO index.

are found in the western coast of Iberia. Qualitatively, the CMIP5 winter precipitation-NAO regression pattern coincides with that obtained with observations (Fig. 5.24b). In the western coast of Iberia and the western Balkans, however, regressions are too weak. The fact that the maximums in observed regressions cover small areas could explain why climate models, with coarser resolution, are not able to capture them.

In order to assess the degree of agreement among the members of the CMIP5 ensemble, we calculate the relative inter-model standard deviation of regressions at each grid-point (Fig. 5.24c). The fact that almost the entire basin exhibits standard deviations much lower than 100% (except the regions located more to the south) indicates that there is a general agreement regarding the sign and magnitude of the regression everywhere. Generally speaking, in regions where the model regression is largest, the agreement among models is also largest.

Moreover, we calculate at each grid-point the number of individual runs (taking all available runs, not model averaging) exhibiting a regression larger (in absolute value) than the observed one (Fig. 5.24d). Even though in the areal average almost all models exhibited a regression smaller than observations (see Fig. 5.23b), in most grid-points the observed regression lies inside the 5% – 95% of the inter-model distribution. Only in those regions where observed regressions are largest (some grid-points in the north-western Iberia and the western Balkans) are these completely outside the inter-model range of variability

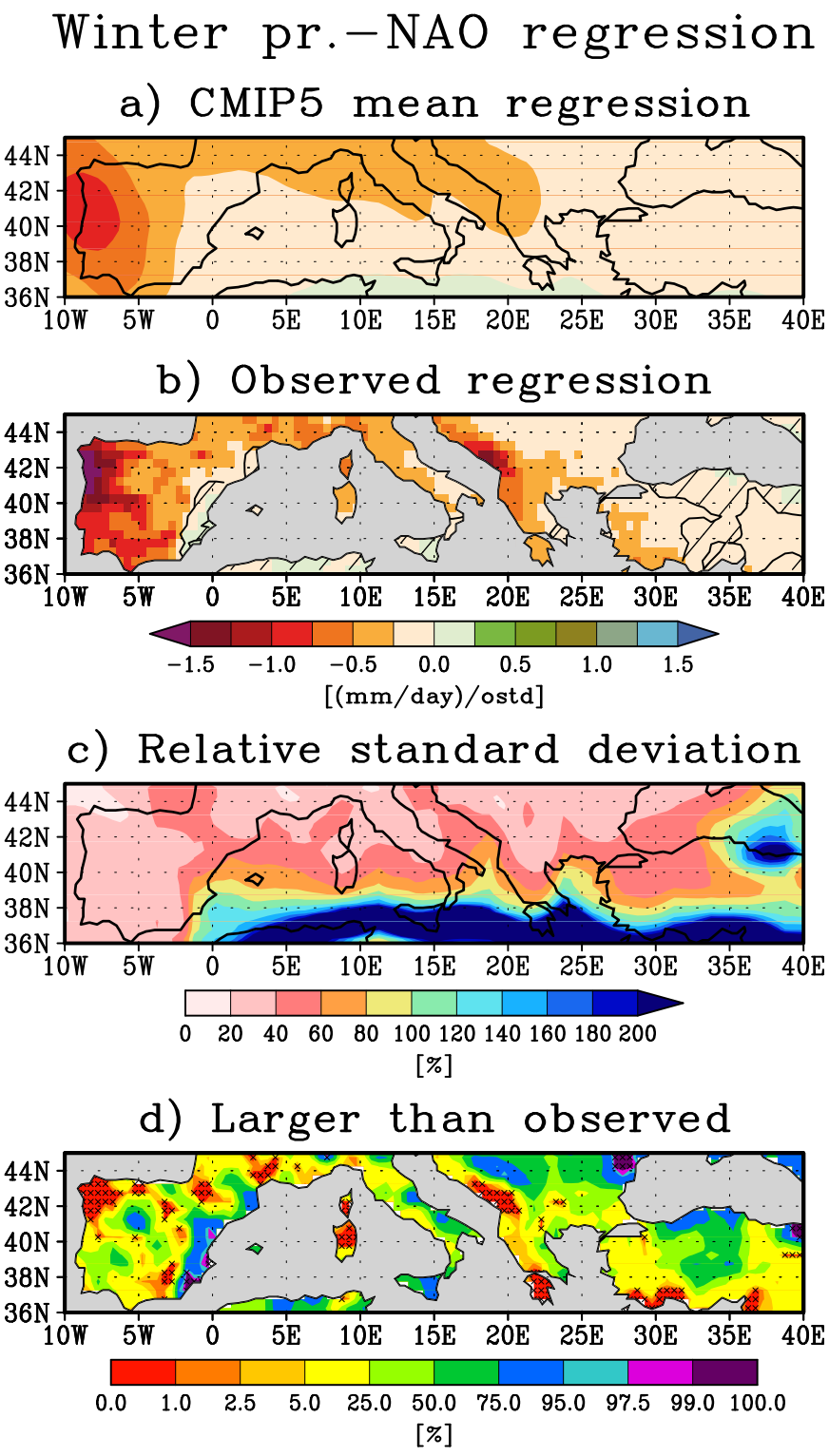


Figure 5.24: (a) Average CMIP5 detrended regression of winter precipitation onto the NAO (1902-2004). CMIP5 mean calculated averaging regressions in pre-industrial runs. (b) The same for GPCC observations. Hatching indicates non-significant regressions with a 95% confidence level. (c) Inter-model standard deviation of the regression field. (d) At each grid-point, number of individual runs with a regression more negative that the observed one. Crosses indicate grid-points in which the observed regression is completely outside the inter-model distribution.

(indicated by crosses), with the exception of Sardinia and the Peloponnese. Thus, broadly speaking, climate models do an adequate job of reproducing the observed linear relation between winter precipitation and the NAO, although correlations and regressions are slightly underestimated in the areal mean and in some specific regions.

5.4.3 NAO trends

Although the linear relation between area-averaged NMedLand winter precipitation and the NAO index tends to be somewhat underestimated in climate models, the NAO still explains a large fraction of the simulated precipitation variance. In this section we assess whether NAO simulations are compatible with observations and whether they support the hypothesis that NAO trends observed in recent decades have a forced origin.

The time series of the NAO index is represented for all available historical simulations in Fig. 5.25a. Like in case of precipitation (see Fig. 5.15), multi-decadal variability masks any long-term forced signal. However, when we focus on the time series of the multi5 NAO index (Fig. 5.25b) we see that in the second half of the 1902-2004 period the NAO mostly exhibits positive values, whereas in the first half negative anomalies predominate. In particular, in the 1902-1953 interval we find 30 (22) winters with negative (positive) NAO; while in the 1953-2004 interval we find 18 (34) winters with negative (positive) NAO. Extending the analysis back to 1862, we see that only 8 winters in the 1862-1901 interval exhibited positive NAO values (with reference to the 1902-2004 mean). The 2DMDA diagram (Fig. 5.25c) illustrates that the upward transition of the NAO during the twentieth century is visible regardless of the selected time scale. Nevertheless, as was the case of winter precipitation, only few multi-decadal anomalies are statistically significant inside the historical period (indicated with a dashed triangle) and we must extend our analysis back to the nineteenth century to find robust statistically significant multi-decadal anomalies.

In the 2DLC diagram (Fig. 5.25d) we detect a forced signal in the NAO index, although it is so weak compared to internal variability that within the baseline period (indicated with a dashed box) some intervals longer than 75-80 years still exhibit non-significant changes. One must extend the analysis back to 1862 for a consistent forced signal to emerge from the noise that still remains after the multi-model average and the diagram to resemble that of a synthetic series with a persistent signal superimposed (see Fig. 4.2c). Focusing on linear trends ending in the last year of the baseline period (2004, diagram top), we see that all time periods ending in 2004 and beginning after 1950 exhibit negative NAO changes (most of them non-significant), indicating that in the last 60 years this NAO mean forced signal has not continued increasing.

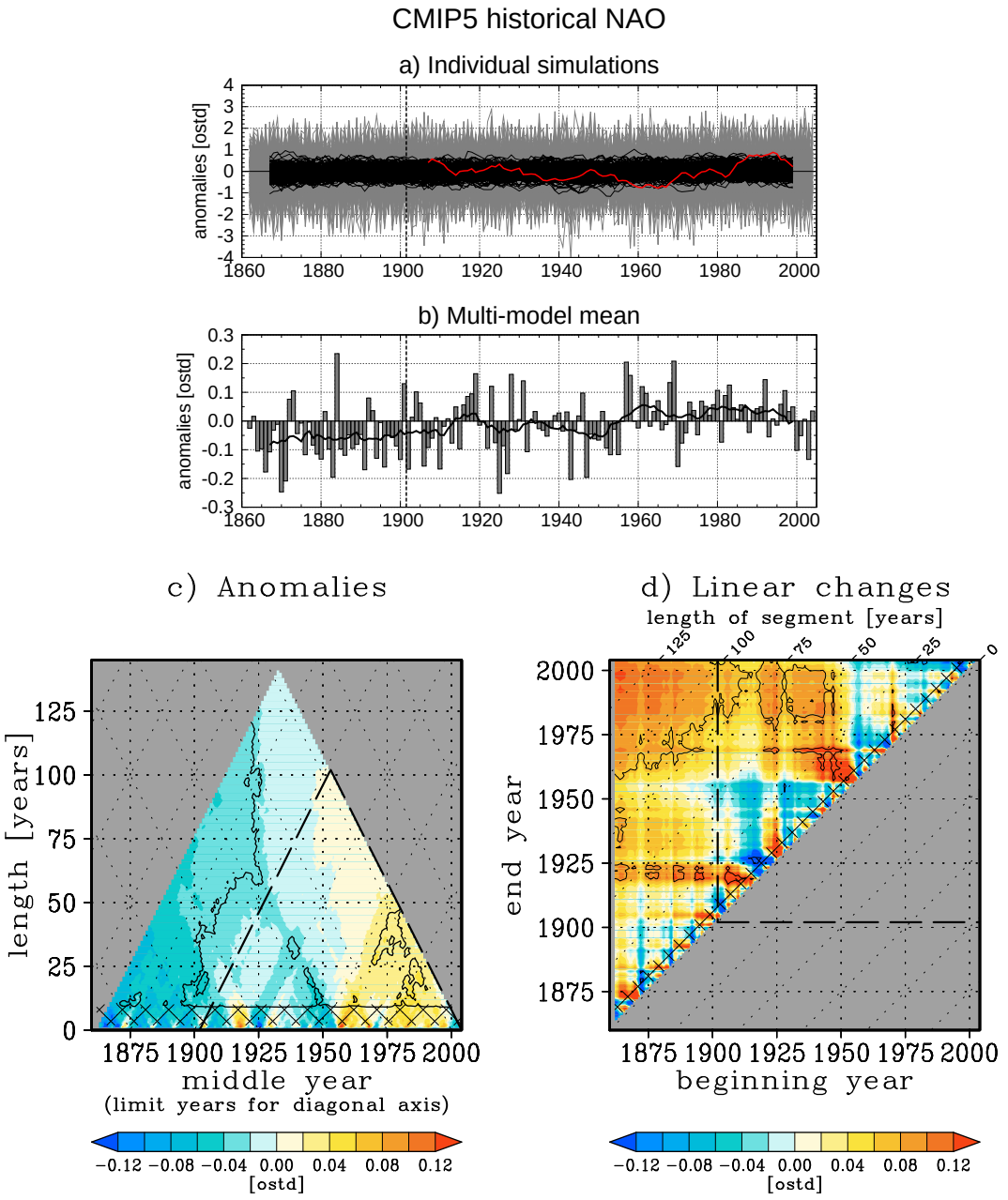


Figure 5.25: (a) NAO anomalies relative to the 1902-2004 mean in all the CMIP5 historical runs (grey curves), expressed relatively to the standard deviation of the observed NAO index (ostd). Black curves represent filtered data. The red curve represents observed NAO anomalies (Trenberth, filtered). The blue line represents the linear fit in the 1960-2004 interval. The dashed vertical line indicates the beginning of the 1902-2004 period. (b) The same for the multi5 ensemble mean. (c-d) 2DMDA and 2DLC diagrams (see Sec. 4.1) for multi5 series. The dashed boxes include time intervals in the 1902-2004 period.

5.4.3.1 1960-2004 NAO trends

In Fig. 5.26 we represent the 1960-2004 NAO trend for individual models, for the multi5 ensemble mean and for observations. The fact that very few simulations exhibit statistically significant trends indicates that the forced signal is weak in comparison with internal variability in almost any single simulation. Moreover, the fact that there is no inter-model agreement regarding the sign of the change (the blue whisker covers trends from −1.89 ostd/century to 2.92 ostd/century) indicates that the forced signal is also very weak in comparison with inter-model variability.

Consistent with the weakness of the forced signal, the natural distribution of 1960-2004 trends virtually coincides with the forced distribution (Fig. 5.27). Because the observed trend is outside both distributions and it has opposite sign than the multi-model mean trend, we conclude that historical simulations are not consistent with model simulations and that either the range of internal variability is underestimated or the forced response of the NAO is not well represented (largely underestimated and opposite to the observed one). Our results are not affected by drift in pre-industrial control simulations (not shown). Using historical simulations to estimate the natural and forced distributions we find very similar results. In particular, any model simulation exhibits a NAO trend as strong as the observed

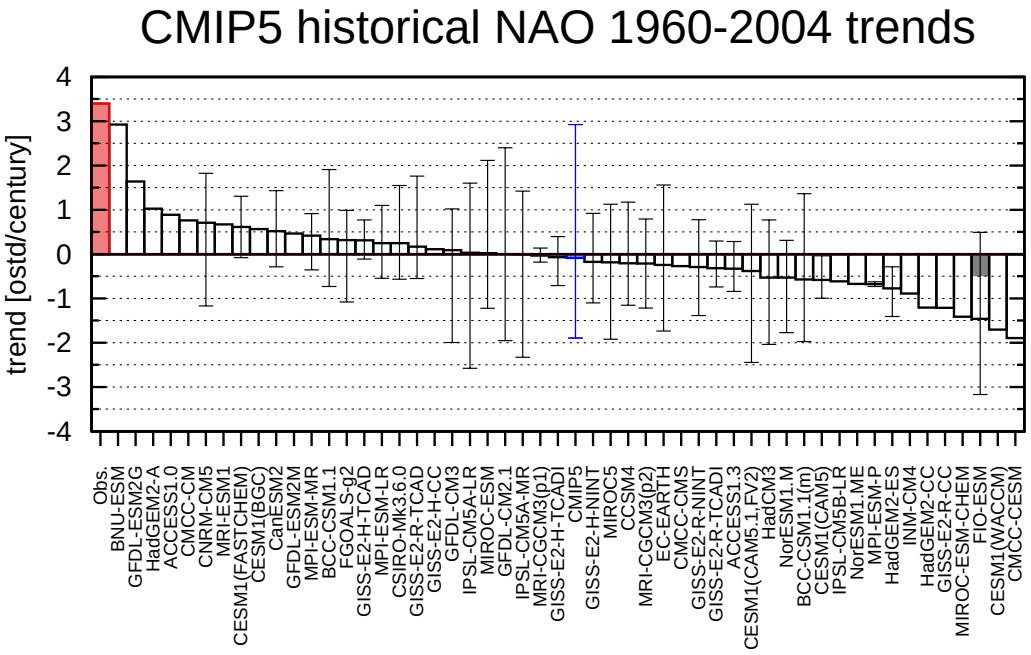


Figure 5.26: Linear trend (1902-2004) of the NAO index in all CMIP5 models. The blue bar corresponds to the multi5 mean and the red bar to the observed trend (Trenberth). The filled fraction of each bar is proportional to the number of runs of the model with statistically significant trends. In those models with several runs, black whiskers indicate the maximum and the minimum trend in individual runs. The blue whisker indicates the minimum and the maximum model-mean trends.

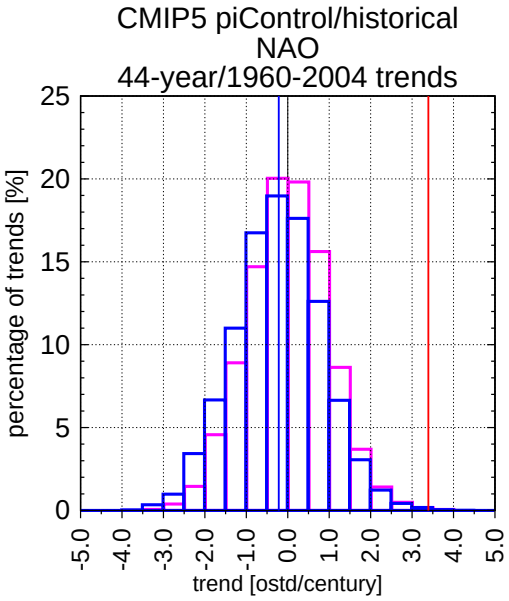


Figure 5.27: Distribution of 44-year trends only due to internal variability (natural distribution; magenta), and with the distribution of trends under the influence of the 1960-2004 forced signal (forced distribution; blue) for CMIP5 NAO index. The blue vertical line indicates the model mean and the red vertical line the observed trend obtained with Trenberth data.

one in the 1960-2004 interval. Osborn (2004), with a similar analysis carried out with a small subset of CMIP3 models, concludes that the large upward NAO trend observed from the 1960s until the 1990s did not match neither simulated natural variability nor a green-house gas forced signal. He argued that only those single models exhibiting both the largest variability and the largest forced signal are consistent with observations.

Then, we conclude that, like was the case of area-averaged NMedLand winter precipitation, NAO recent trends are not compatible with the internal variability range and the forced response of CMIP5 simulations. Kelley et al. (2012b) find similar results for 1965-1995 November-April trends of the NAO and of the first PC of Mediterranean precipitation using pre-industrial and historical simulations of CMIP3 models. Because they find models able to produce trends as large as the observed ones as very unusual events, they conclude that winter precipitation and NAO forced responses are beginning to emerge from internal variability. Their conclusion, however, is not valid when using CMIP5 models, as the mean forced signal is weaker and, therefore, models are unable to produce NAO trends stronger or as strong as the observed ones.

5.4.3.2 Sensitivity to time interval

Next, we assess whether the conclusions drawn for the 1960-2004 trend are robust to the choice of time interval. First, we calculate for each time interval the signal-to-noise ratio defined as before (Fig. 5.28b). Because it is lower than 0.15 in all periods, we conclude

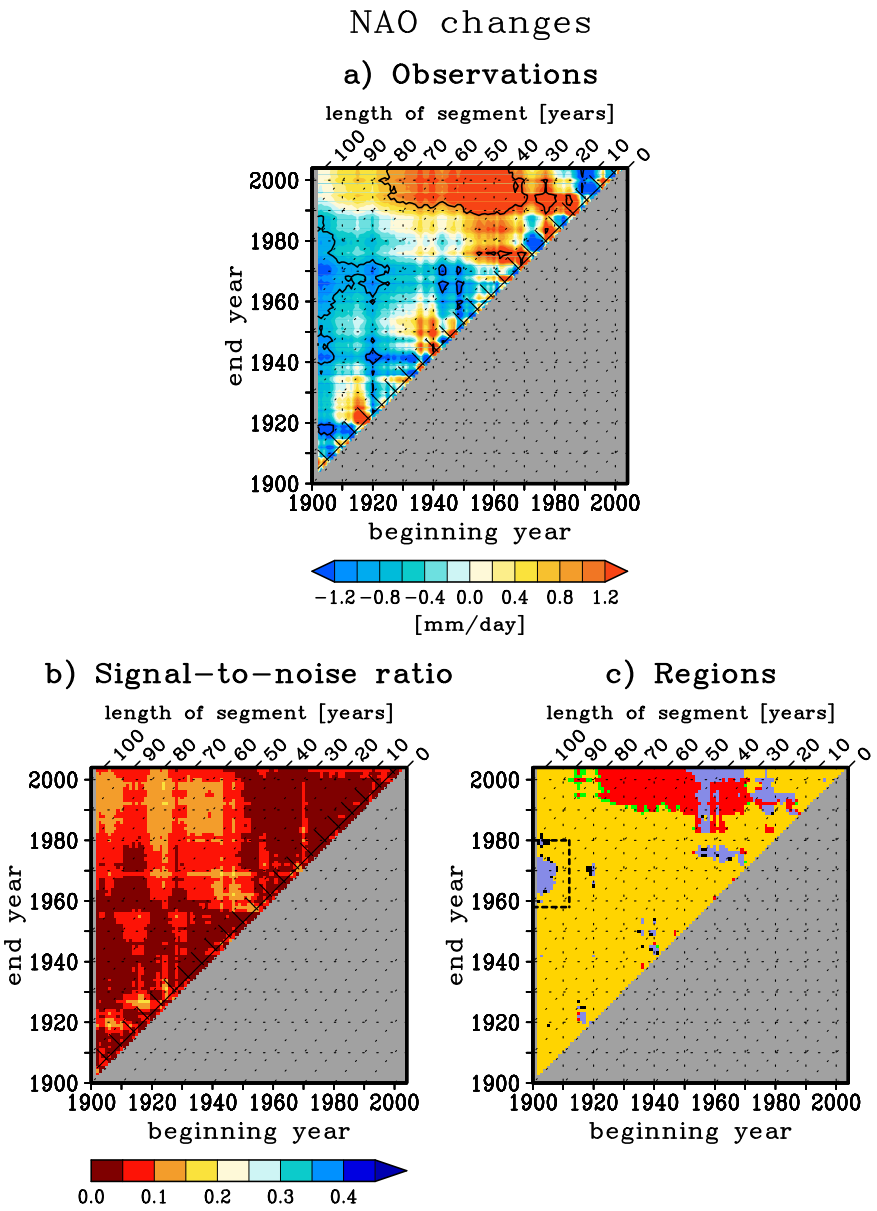


Figure 5.28: (a) 2DLC diagram of observed (Trenberth) NAO. Contours indicate statistically significant trends. (b) Signal-to-noise ratio, defined as the ratio between the multi-model mean trend and the width of the distributions in each time interval. (c) Two-dimensional diagram in which we compare, for each possible time interval within the 1902-2004 period, the observed NAO trend with the natural distribution and forced distribution of models simulations. Colours correspond to the regions described in Sec. 5.1.

that the forced distribution is very similar to the natural distribution in all sub-intervals of the historical period.

Next, we analyse in which *region* of the distributions the observed NAO trend falls for each time interval within the historical period. Generally speaking, in time intervals with the strongest changes (see Fig. 5.28a) these changes are outside both the natural and the forced distributions, in some cases even with the opposite sign than the multi-model mean.

This is the case of NAO changes estimated for time intervals beginning in 1920-1970 and ending after 1990, in which the observed NAO exhibits very strong and statistically significant positive trends. It is also the case of NAO changes estimated for time intervals beginning in 1900-1910 and ending around 1970, in which the observed NAO exhibits strong and statistically significant changes. In both situations, observed NAO changes are not compatible with model simulations and either the range of internal variability is underestimated or the response to the forcing is not well captured. In time intervals in which observed changes are not statistically significant, instead, these are generally consistent with natural and forced trends and, therefore, no signal is detected.

Because models are not able to reproduce the strongest NAO trends for the 1900-1970 interval (dashed box), which are clearly natural (they have not persisted), the hypothesis that the inconsistency between model and observed precipitation trends is associated with an underestimation of the range of internal variability gains strength. Given the impact of the NAO on winter precipitation, the inability of the CMIP5 models to correctly simulate the NAO changes observed in the last decades, then, is consistent with their inability to reproduce the observed NMedLand winter precipitation changes.

5.4.3.3 Comparison with CMIP3

Looking at the 2DLC diagram of multi3 NAO (Fig. 5.29) we see that, in general, for any time interval within the historical period, the CMIP3 forced NAO signal is stronger than the CMIP5 signal (see Fig. 5.25c). Given the strong influence of the NAO on winter precipitation, this result could explain why the NMedLand winter drying signal is stronger in CMIP3 than in CMIP5 (compare Figs. 5.15c and 5.16). The origin of the discrepancies in the magnitude of NAO trends between these multi-model ensembles is beyond the scope of this thesis, but it could be related to improvements in the coupling of oceanic and atmospheric models that would affect sea surface temperature, or to a better simulation of stratospheric processes due to the increase of vertical resolution and height of model top in the CMIP5 ensemble.

As a first approach to study differences between CMIP5 and CMIP3 trends, we select a subset of models of each ensemble, pairing models that include an old version in CMIP3 and a new version in CMIP5 (Fig. 5.30). With them, we calculate the difference between old models and new models in the 1902-2004 NAO trend and in the NMedLand winter precipitation drying (estimated as the precipitation trend with the sign changed). The scatter plot in Fig. 5.31 confirms that the difference in NAO trend is well correlated with the difference in drying trend, supporting the hypothesis that the weaker CMIP5 NMedLand winter precipitation trend of the multi-model ensemble mean with respect to CMIP3 is associated with the fact that the NAO trend is weaker as well.

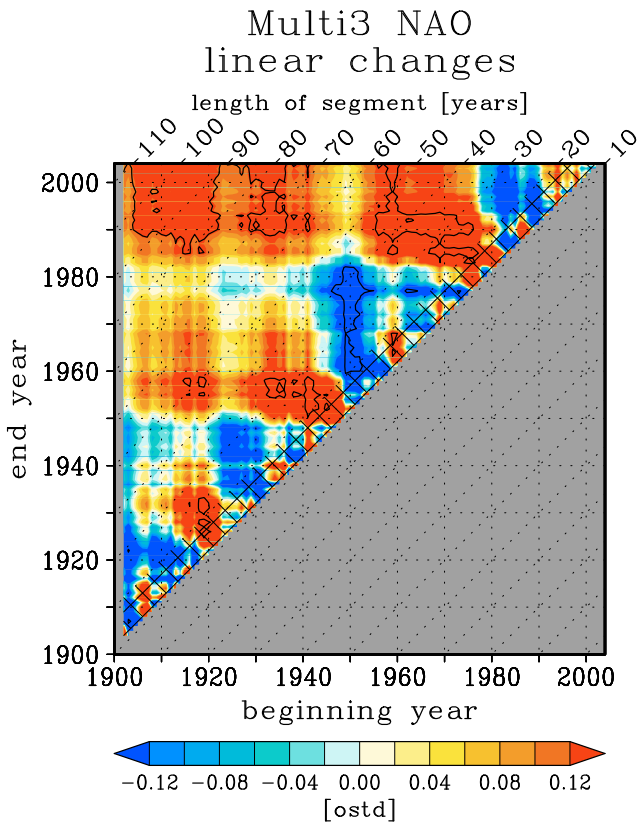


Figure 5.29: 2DLC diagram of multi3 NAO in the 1902-2004 interval. Contours indicate intervals with statistically significant trends.

5.5 Summer precipitation trends

In Sec. 5.3 we found that summer months are those in which historical simulations exhibit the strongest precipitation decreases (see Figs. 5.13 and 5.14). According to results found in Chap. 4 we know that a comparable signal is not evident in observations, as NMedLand summer precipitation changes in the last decades are for now weak and not statistically significant (see Fig. 4.8). Because summer precipitation changes are projected to be very large in the future (see Chap. 6), in this section we explore the summer drying signal in historical simulations.

In Fig. 5.32a, in which we represent the evolution of NMedLand summer precipitation in historical simulations within the historical period, we see that, as was the case for winter precipitation, multi-decadal variability and inter-model variability are very strong in comparison with the superimposed drying signal. In this case, however, observed filtered anomalies never exceed the range of filtered anomalies in historical simulations.

When analysing the time series of the multi5 ensemble mean (Fig. 5.32b), we see that the forced signal in summer is stronger (−0.08 (m/day)/century; blue line) and that the inter-annual variability is lower than in winter. In fact, the 2DMDA and the 2DLC diagrams (Fig. 5.32c-d) illustrate that NMedLand summer precipitation in historical sim-

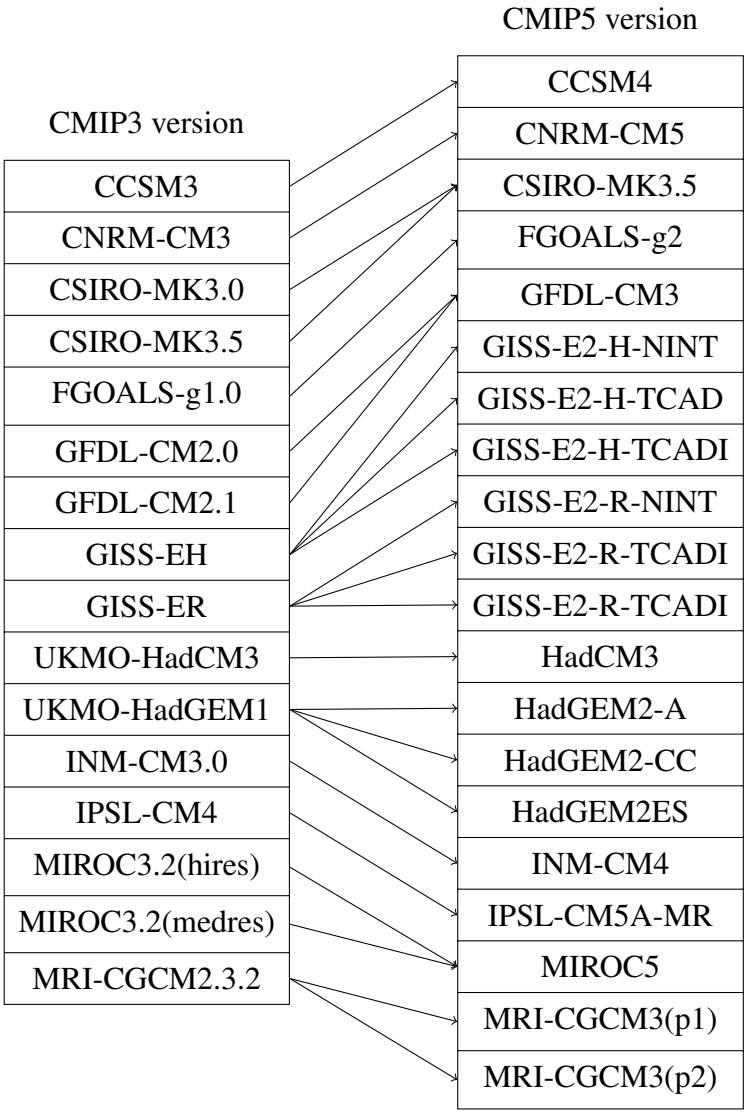


Figure 5.30: Versions of the same model in the CMIP3 and the CMIP5 multi-model ensembles.

ulations behave like synthetic time series in which we superimpose an external trend to normally distributed noise (see Sec. 4.1).

In the two-dimensional diagram of Fig. 5.33b, we represent the signal-to-noise ratio between the multi-model mean trend and the half-width of the natural distribution in each time interval within the 1902-2004 period. Even though the signal-to-noise ratio is generally larger than in winter, the fact that it never exceeds 0.5 indicates that the forced distribution is never completely separated from the natural distribution. Because observed NMedLand summer precipitation changes are weak in the vast majority of time intervals (Fig. 5.33a), they are generally contained within both distributions (Fig. 5.33c) and, thus, a drying signal is not detectable.

In order to analyse the inter-model dispersion regarding the magnitude of the forced signal, we calculate the linear trend in each historical simulation for the entire period 1902-2004 (Fig. 5.34). The entire historical period has been chosen in order to increase

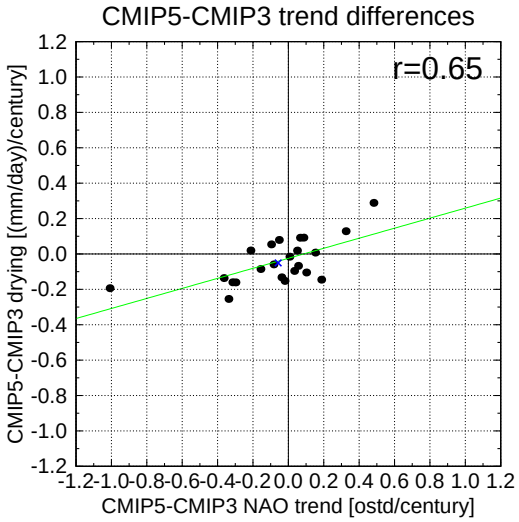


Figure 5.31: Scatter plot of model-mean CMIP5-CMIP3 difference in 1902-2004 NAO trends against CMIP5-CMIP3 difference in NMedLand winter precipitation drying (estimated as the trend with the sign changed). We only include models with versions in both examples. The blue cross indicates the difference in trend in the multi-model ensemble mean (including all models in each ensemble). The green line represents the linear fit using a linear regression, with the associated correlation coefficient indicated in the top right corner.

the signal-to-noise ratio and remove as much internal variability as possible. The fact that about a 80% of models exhibit negative trends indicates that the forced signal is strong in comparison with inter-model variability. Nevertheless, even though the number of individual simulations with statistically significant trends is larger than in winter (see Fig. 5.17), most models still exhibit non-significant trends, what means that internal variability is still strong in comparison with the forced signal. The signal in NMedLand summer precipitation, then, is still far from being as strong as that in EuroMed annual temperature (see Fig. 5.7) in which all models coincide in the sign of the change and almost all single runs exhibit statistically significant trends. In Fig. 5.35 we represent the natural and the forced distributions of 1902-2004 NMedLand summer precipitation trends. Unlike in the studied winter trends, both distributions are not totally overlapped but in this case the observed trend is so weak that it is inside both distributions.

Finally, we represent the geographical pattern of forced decreases evaluated using the multi5 NMedLand summer precipitation 1902-2004 trend at each grid-point (Fig. 5.36a). Unlike in winter, summer precipitation long-term trends are negative everywhere and statistically significant in most part of the basin. In general, largest decreases are found in the north of the basin and they are particularly intense in the Balkan region. The distribution of the largest forced changes shown in Fig. 5.36a does not coincide with that found by Kelley et al. (2012a) (see their Fig. 3f), as they find the largest trends in northern Iberia. The discrepancies could be due to the fact that they calculate the forced signal over a shorter period (1950-2004) and using a wider summer definition (May-October). When calculat-

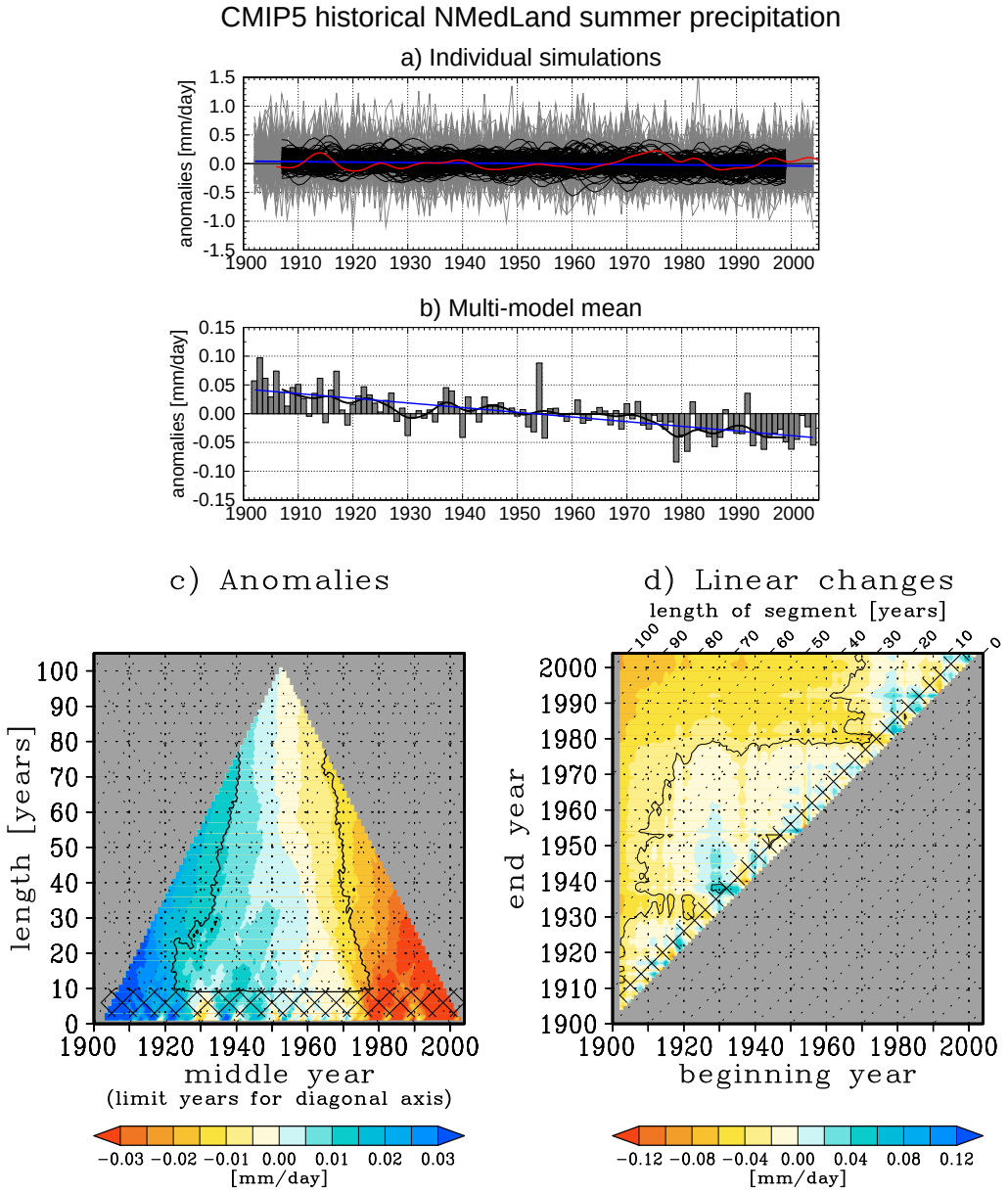


Figure 5.32: (a) NMedLand summer precipitation anomalies relative to the 1902-2004 mean in all the CMIP5 historical runs (grey curves). Black curves represent filtered data. The red curve represents observed precipitation anomalies (GPCC, filtered). The blue line represents the linear fit in the 1902-2004 interval. (b) The same for the multi5 ensemble mean. (c-d) 2DMDA and 2DLC diagrams (see Sec. 4.1) for multi5 series.

ing the multi-model average relative summer decreases (dividing the absolute trends by the summer climatological mean per model at each grid-point; Fig. 5.36b), changes in the north are de-emphasized and the meridional gradient disappears. Instead, the largest summer precipitation decreases are found in the eastern half of the basin. Using the old CMIP3 multi-model ensemble we obtain a very similar pattern, with slightly stronger decreases (not shown).

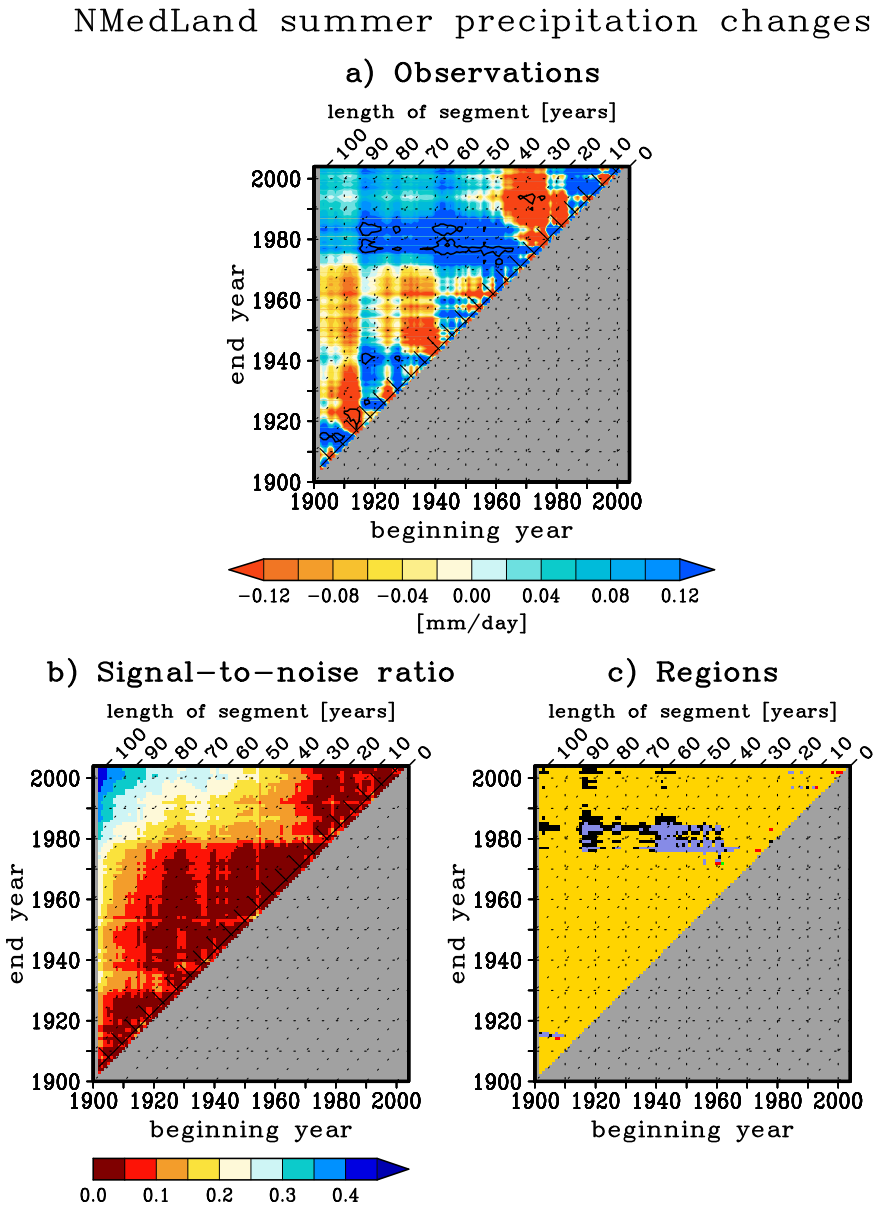


Figure 5.33: (a) 2DLC diagram of observed (GPCC) NMedLand summer precipitation. Contours indicate statistically significant trends. (b) Signal-to-noise ratio, defined as the ratio between the multi-model mean trend and the width of the distributions in each time interval. (c) Two-dimensional diagram in which we compare, for each possible time interval within the 1902-2004 period, the observed NMedLand summer precipitation trend with the natural distribution and forced distribution of models simulations. Colours correspond to the regions described in Sec. 5.1.

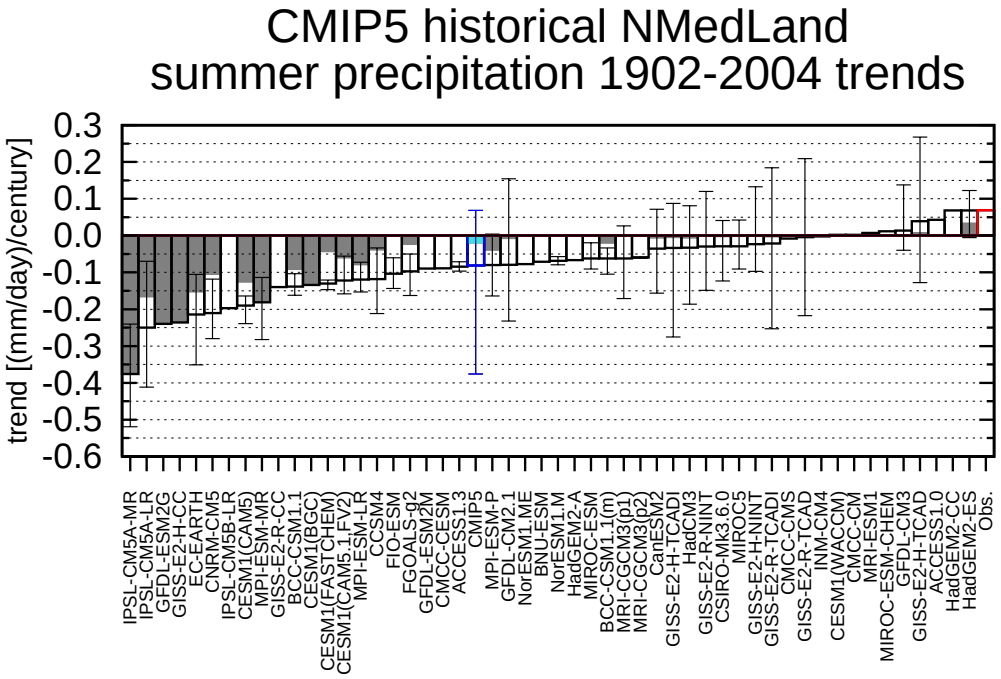


Figure 5.34: Linear trend (1902-2004) of summer precipitation averaged over the NMedLand region in all CMIP5 models. The blue bar corresponds to the multi5 mean and the red bar to the observed trend (GPCC). The filled fraction of each bar is proportional to the number of runs of the model with statistically significant trends.

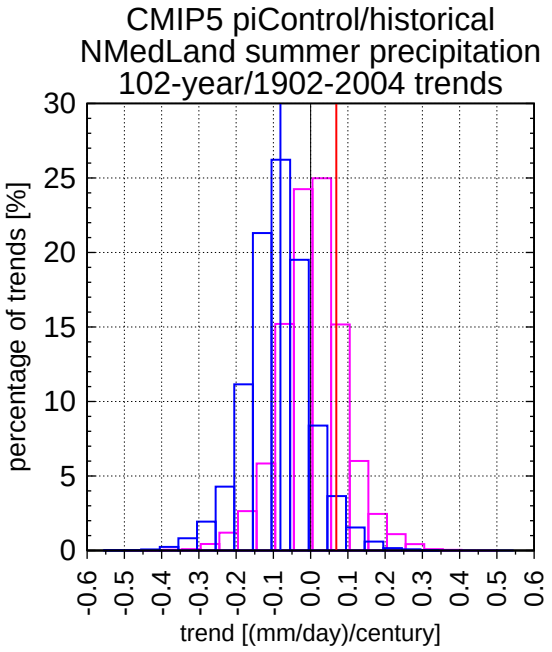


Figure 5.35: Distribution of 102-year trends only due to internal variability (natural distribution; magenta), and with the distribution of trends under the influence of the 1902-2004 forced signal (forced distribution; blue) for CMIP5 NMedLand summer precipitation. The blue vertical line indicates the model mean and the red vertical line the observed trend obtained with GPCC data.

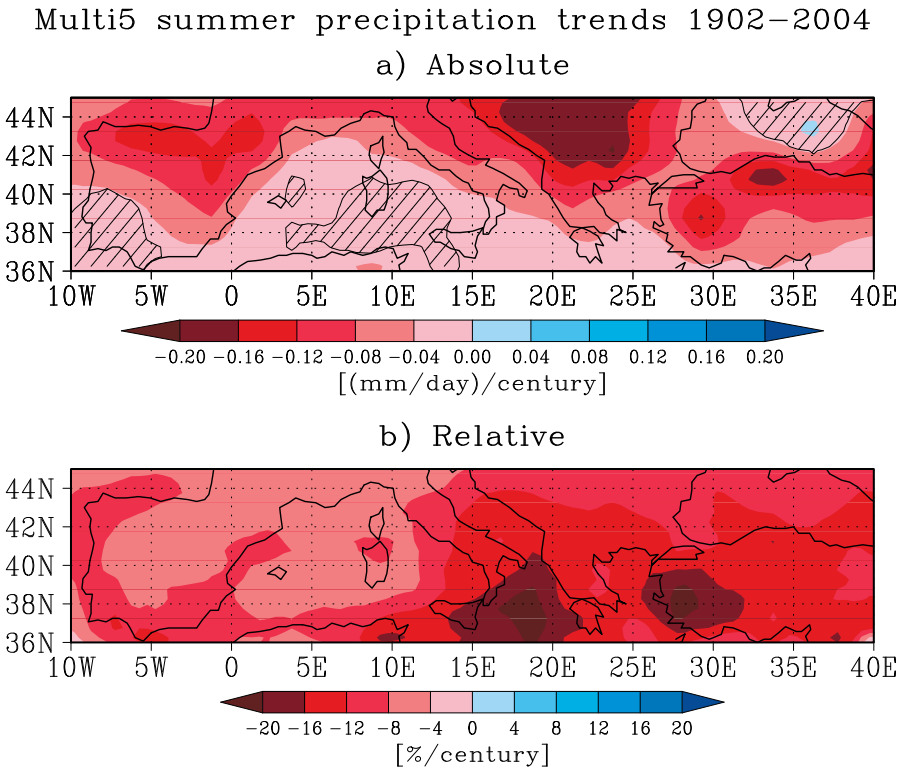


Figure 5.36: (a) Map of 1902-2004 linear trends in multi5 summer precipitation. Hatching indicates non-significant trends with a 95% of confidence. (b) Relative trends, calculated dividing the total trend by the summer climatological mean at each grid-point.

5.5.1 Impact of the SNAO on summer precipitation

The role of the SNAO in summer precipitation trends was already assessed in Bladé et al. (2012a) and Bladé et al. (2012b). In this section we provide a summary of these works. First, as a background, we describe the observed relation between SNAO and summer precipitation in the northern Mediterranean. Then, we discuss the inability of climate models to correctly reproduce such relation and the consequences of this drawback.

5.5.1.1 Observations

In the first work we studied the observed impact of the SNAO on summer precipitation in Europe, revisiting and completing the work previously done by Folland et al. (2009). In Fig. 5.37 we see that summer precipitation is positively correlated with the SNAO index in the northern Mediterranean. Correlations, however, are on average smaller than those observed in winter due to the fact that, whereas in winter the southern lobe of the NAO is located over the Azores and the associated SLP changes directly influence the wind circulation over the Mediterranean region (see Fig. 3.6a), the southern lobe of the SNAO is centred over the British Isles and the associated SLP variations do not alter directly the advection of Atlantic air over the Mediterranean (Fig. 5.38). Moreover, even if the southern

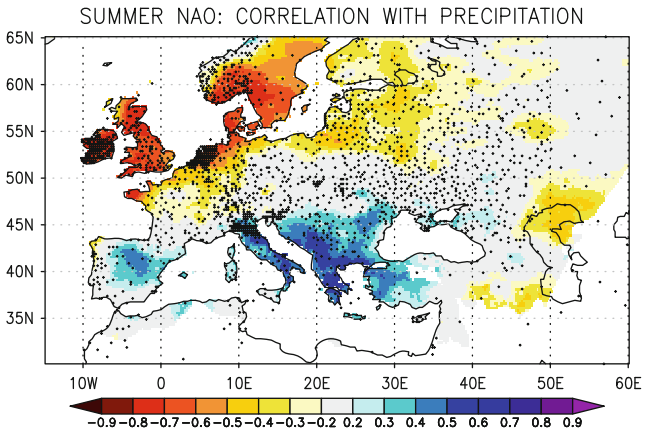


Figure 5.37: Correlation of observed summer E-OBS precipitation and the SNAO index. Grey regions are those with correlations lower than $r = 0.2$, with a confidence level lower than 90% estimated using a two-sided Student-t test. Blank regions are those with missing data. Dots indicate station locations. This figure corresponds to Fig. 5a in Bladé et al. (2012b).

lobe of the SNAO was able to influence the advection of wet air over southern Europe, we would expect negative correlations (as in winter) instead of positive. Looking for a dynamic explanation for the positive correlations in the northern Mediterranean, in Bladé et al. (2012b) we found negative anomalies of 500-hPa and 200-hPa geopotential height over the Balkan region and, less intense, over west of Iberia associated with positive SNAO summers. Then, we concluded that the mid-level cooling associated with these upper-level troughs would increase the potential instability in that region and enhance convective precipitation.

Even though the correlation between the SNAO index and NMedLand summer precipitation is low ($r = 0.43$), Mariotti and Dell'Aquila (2012) claim that the SNAO largely explains the evolution of the summer (June-August) precipitation averaged over the entire Mediterranean at decadal scale, as they find a correlation coefficient of $r = 0.79$ between

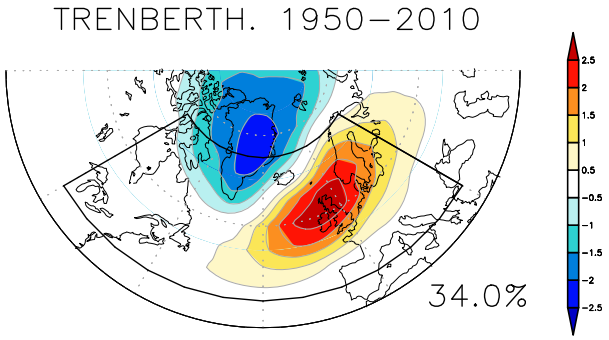


Figure 5.38: Spatial pattern of the observed SNAO defined as the first PC of the summer Trenberth SLP field in the 90°W-30°E/40°N-70°N domain, for the period 1950-2010. The pattern is displayed in terms of the regression of the SLP field onto the normalized NAO index (units are hPa/std). In the bottom-right corner, the percentage of SLP variance explained by the associated EOF. This figure corresponds to Fig. 1b in Bladé et al. (2012b).

the two time series after detrending them and applying a 10-year running mean (see their Fig. 3f). When reproducing such calculation using our own parameters (July-August mean averaged over the northern Mediterranean) here we obtain a smaller correlation $r = 0.64$.

The largest correlations across the basin were found in the Balkan and Italian Peninsulas (up to $r = 0.6$) and in central Iberia (up to $r = 0.5$). In particular, the area-mean precipitation over the Italian/Balkan region (10°E-30°E/37.5°N-45°N) was found linearly related to the SNAO index with a correlation of $r = 0.64$. However, even though decadal fluctuations of summer precipitation over this region were found modulated by the SNAO, an upward trend present in the SNAO until the year 2000 had no Italian/Balkan summer precipitation wetting trend associated (see Fig. 8 in Bladé et al. (2012b)).

Using the 2DMDA and the 2DLC diagrams for Italian/Balkan summer precipitation and the SNAO index we reinforce the conclusions obtained in Bladé et al. (2012b). First, in the 2DMDA diagrams (Figs. 5.39a and 5.39c) we see that, for periods longer than 30 years, the distribution of SNAO multi-decadal anomalies generally matches the distribution of Italian/Balkan precipitation anomalies, with negative anomalies for time intervals ending before 1970 and positive anomalies from then on. However, while in the 2DLC diagram for the SNAO (Fig. 5.39b) all time intervals longer than 70 years exhibit statistically significant upward SNAO trends, in the diagram for Italian/Balkan precipitation (Fig. 5.39d) we see that these long-term SNAO upward trends have no Italian/Balkan wetting trends associated. In Bladé et al. (2012b) we hypothesized that the SNAO-associated wetting could have been partially compensated by soil moisture negative feedbacks or aerosol effects.

5.5.1.2 Models

In our second work [Bladé et al., 2012a], we studied whether the observed impact of the SNAO on northern Mediterranean summer precipitation was correctly captured by CMIP3 models. Our results indicate that, even though the general pattern of positive correlations in the northern Mediterranean is captured by most models (Fig. 5.40), the magnitude of the simulated correlations is about half the observed one. In particular, in the Italian/Balkan region, where the observed correlations are largest, all single simulations exhibit a precipitation-SNAO correlation lower than the observed one and about one third of them exhibit values lower than $r = 0.2$ or even negative (Fig. 5.41). We attributed this weak simulated response of summer Mediterranean precipitation to the SNAO to the inability of climate models to correctly reproduce the upper level through centred over the Italian/Balkan region associated with positive SNAO summers that creates propitious conditions for the development of convective precipitation in that area. Results found in Bladé et al. (2012a) have been recently repeated using the new CMIP5 models by I. Bladé and P. Gámez [personal communication], obtaining very similar results (not shown). In conclusion, then, because the observed link between summer precipitation in the northern

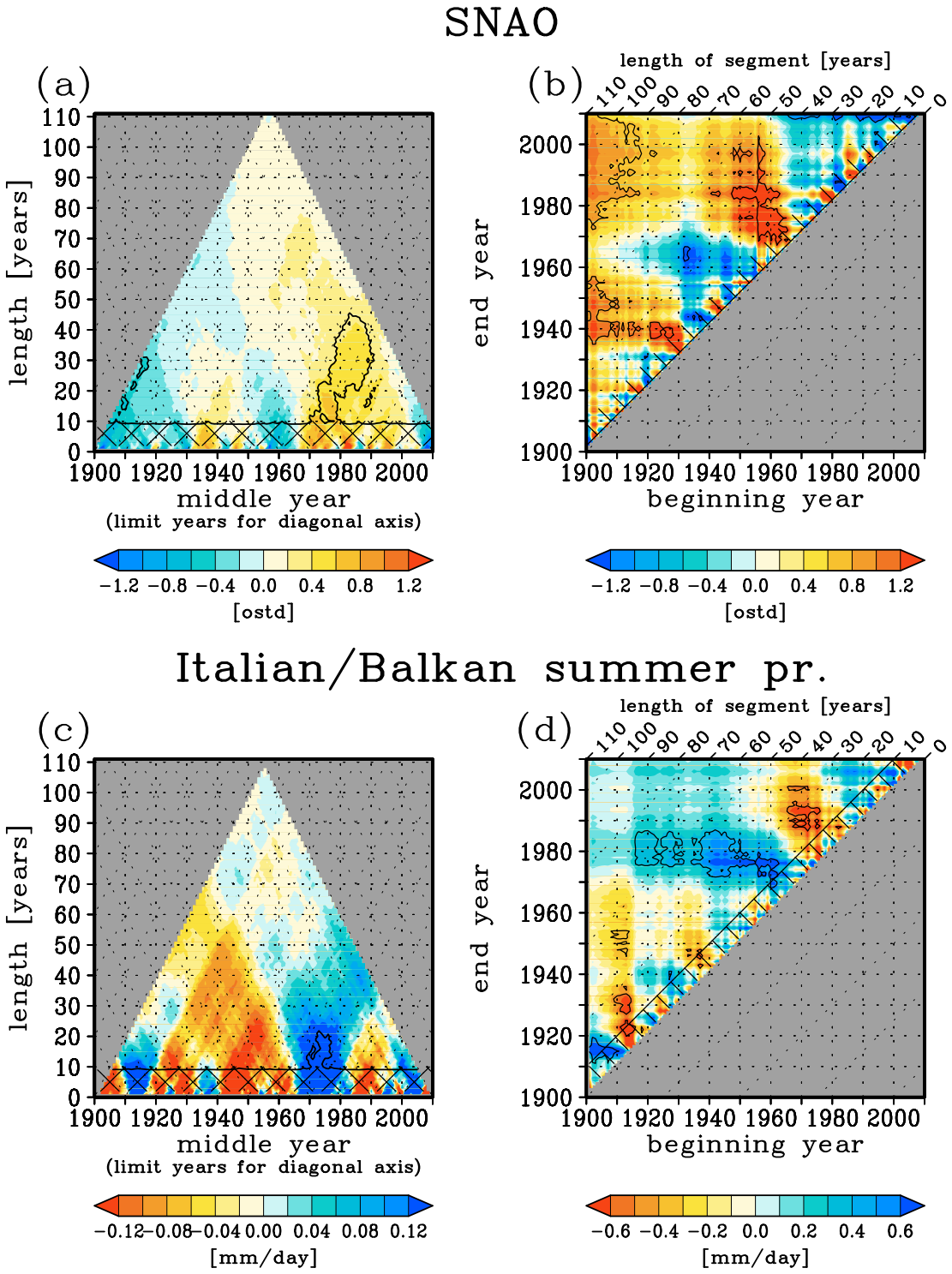


Figure 5.39: (a-b) 2DMDA diagram (see Sec. 4.1) and 2DLC diagram (see Sec. 4.1.1) for the observed SNAO index defined as the first PC of the July-August mean SLP pattern in the north Atlantic in the 1902-2004 interval. (c-d) The same for observed Italian/Balkan summer precipitation (GPCC) relative to the 1902-2004 mean.

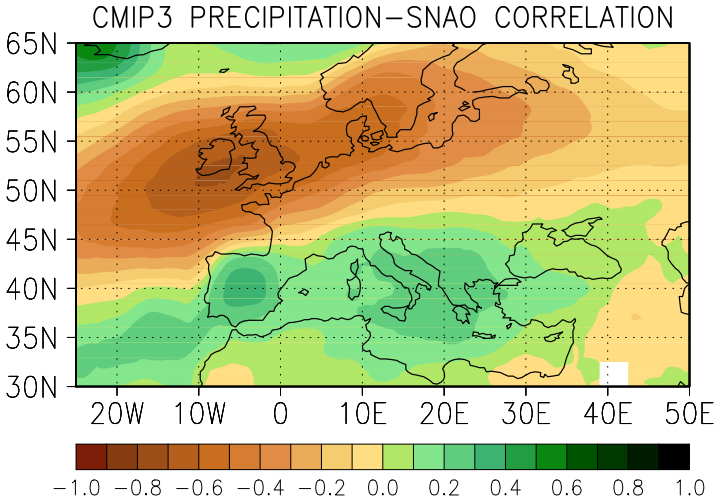


Figure 5.40: Map of correlations between summer precipitation and the SNAO index in the CMIP3 multi-model ensemble mean. For each simulation, the SNAO is defined as the projection of the observed SNAO pattern onto the model SLP field. Correlations for single runs are averaged using the Fisher's Z transformation (see Sec. 3.1.3). This figure corresponds to Fig. 3d in Bladé et al. (2012a).

Mediterranean and the SNAO is not well captured by climate models, the lack of SNAO-related summer precipitation in models could partially explain the strong drying projected for the end of the twenty-first century.

5.6 Conclusions and discussion

By comparing observations with forced and unforced model simulations, in this chapter we have examined whether the strong winter precipitation and NAO changes observed in the last decades are compatible with model simulations, whether they evidence that an anomalous signal is detectable and whether this signal is attributable to external forcings.

The forced signal in winter northern Mediterranean precipitation and the NAO during the historical period, estimated as the evolution of the CMIP5 multi-model ensemble mean, is weak in comparison with internal variability and with inter-model variability. For this reason, the inter-model distribution of forced trends (estimated with historical simulations) is very similar to the inter-model distribution of internal trends (estimated with pre-industrial simulations) for any time period within the 1902-2004 interval. As a result, the observed change for a given time interval is either inside or outside both distributions at the same time.

In case of the strong 1960-2004 winter precipitation and NAO observed trends, they are outside both natural and forced distributions. In this situation, we consider two hypotheses: (1) observed trends are mainly associated with natural variability, which is underestimated by climate models; and (2) observed trends have a forced origin, but the response to the

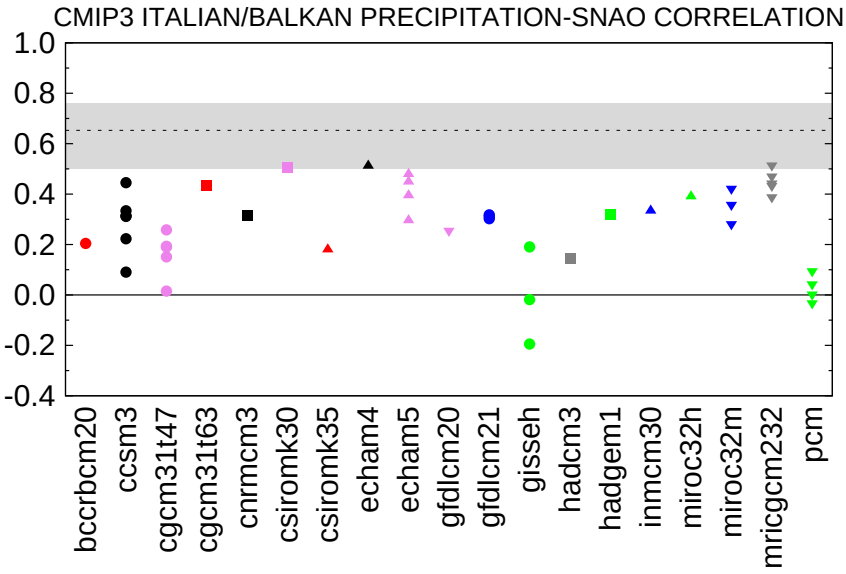


Figure 5.41: Correlation between Italian/Balkan summer precipitation and the SNAO index in each CMIP3 simulation. The dotted line indicates the observed correlation, and the gray shading the 95% confidence limits estimated with a non-parametric bootstrap test. This figure corresponds to Fig. 5 in Bladé et al. (2012a).

forcing is underestimated or misrepresented in historical simulations. These results are valid for observed trends estimated for any interval longer than 30 years beginning after 1910 and ending after 1990 (periods in which the observed changes are strongest). For sub-intervals within the 1902-1990 period, instead, observed winter precipitation and NAO trends are inside both forced and unforced distributions and, therefore no drying or upward NAO trend are detected. At a more regional level, in most parts of the northern Mediterranean basin, the observed trend is compatible with both internal variability and climate change. Only in some of the regions in which the response of winter precipitation to the NAO is strongest, including the west Balkans, the west of the Italian Peninsula and some spread regions in Iberia and Anatolia, the observed trend is outside both forced and unforced distributions.

In summer, the area-averaged drying signal is stronger than in winter. For this reason, the distribution of trends under the presence of an external forcing and the distribution of internal trends are not completely overlapped. Because observed summer precipitation trends are very weak, however, they fall within both distributions in almost any time period within the 1902-2004. Thus, even though the projected summer drying (examined in Chap. 6) is already visible in simulations for the historical period, according to observations it has yet not emerged from internal variability.

Chapter

SIX

Assessing the robustness of projected precipitation trends in the northern Mediterranean

Having examined precipitation trends in observations and historical simulations, we now turn our attention to precipitation changes in projections for the future. According to the IPCC AR5, climate models project drying in the northern Mediterranean area for the end of the twenty-first century [Christensen et al., 2013]. But, to which extent are these projections robust?

Generally, climate projections are only examined through the multi-model ensemble mean [e.g. Mariotti et al., 2008; Giorgi and Lionello, 2008]. Even though this is a good approach to estimate the most likely future behaviour of northern Mediterranean precipitation, it says nothing about whether individual models agree with each other and may mask strong discrepancies between them. In addition to examine the degree of agreement among models regarding precipitation changes, other questions of interest are: Is the amplitude of the multi-decadal variability comparable among different model runs? Do all models exhibit a gradual drying throughout the twenty-first century or is the precipitation decrease abrupt at some point of the twenty-first century? Is the time of emergence of the forced signal, i.e. the approximate year in which the forced signal stands above internal variability, common across models?

The only aspect of the projections of precipitation changes in the northern Mediterranean whose inter-model robustness has been evaluated is the sign of the change. According to Christensen et al. (2013) and Mariotti et al. (2015), the agreement in the sign of the projected precipitation change in summer is high, with more than 66% of individual models projecting negative changes in most of the basin. In winter, instead, the inter-model

agreement is weaker, as generally fewer than 66% of models coincide in the sign of the change.

In addition to studying the robustness of precipitation trends, we also examine the projected future behaviour of the NAO. As seen in previous chapters, the NAO has played the major role in driving the evolution of northern Mediterranean precipitation in the past. In this chapter we discuss whether the projected evolution of northern Mediterranean precipitation is also mainly driven by the NAO or if it is modulated by other mechanisms response to the global warming.

Among the large scale responses of the hydrologic cycle to the global warming, the most fundamental one is based on the thermodynamic relation derived from the Clausius–Clapeyron equation. This relation states that, as the atmosphere warms, its saturated vapour pressure increases and, under assumed constant relative humidity, the amount of water vapour in the atmosphere increases as well. As a result, the net moisture transport from the subtropics to subpolar and equatorial regions intensifies, causing the difference between precipitation and evaporation ($P - E$) to grow everywhere. According to Held and Soden (2006), changes in $P - E$ are mainly modulated by precipitation changes. Therefore, while current wet regions ($P - E > 0$) are expected to exhibit precipitation increases, current dry regions ($P - E < 0$) are expected to suffer drying. Scheff and Frierson (2012), however, studying precipitation change patterns in CMIP5 projections, conclude that the largest future drying is not expected over the subtropical dry regions but poleward of the current subtropical precipitation minima, implying that dynamical mechanisms such as the poleward expansion of the Hadley cell also play a major role in defining the distribution of mid-latitudinal dry regions in projections for the future.

Focusing on the Euro-Mediterranean region, Rowell and Jones (2006) propose other mechanisms to explain the strong projected summer drying. First, they suggest that precipitation reductions in spring could lead to a reduction in the soil moisture available to produce convective rainfall in summer. Second, they argue that the precipitation decrease in summer could contribute to reduce the soil moisture further, creating a positive feedback mechanism. Third, they argue that the slower warming over sea than over land (associated with the larger thermal inertia of the ocean), could cause the increase in evaporation over sea to be unable to compensate for the faster increase in atmospheric moisture capacity over land, which would reduce the relative humidity and contribute to inhibit land precipitation [Manabe et al., 1992; Gibelin and Déqué, 2003]. Finally, they point out that the projected European drying could be associated with large scale circulation patterns.

The goal of this chapter is to examine the robustness of future Mediterranean drying projections in CMIP5 models, checking whether results obtained with different climate models are robust in terms of the sign, the magnitude, the seasonality and the geographical distribution of the change. We also assess which aspects of the transition towards drier

conditions, such as the time of emergence of the forced signal or the amplitude of the multi-decadal variations, are robust across models. Finally, we examine the role of the NAO and other mechanisms described above in the projected drying trends. For all these purposes, we use projections under the RCP8.5 scenario, the one with strongest forcings, and discuss some discrepancies with projections under the intermediate RCP4.5 scenario (see Sec. 2.2 for a complete list of models and simulations).

6.1 Seasonal distribution of projected trends

We begin by assessing the seasonal distribution of projected NMedLand precipitation changes. The multi-diagram in Fig. 6.1 (see Sec. 4.3 for a complete description of how this diagram is created and how it should be interpreted) shows the evolution of decadal to multi-decadal area-averaged NMedLand precipitation anomalies in the multi5 ensemble mean during the twenty-first century under the RCP8.5 scenario. The diagram illustrates a seasonal contrast: while summer months exhibit statistically significant multi-decadal

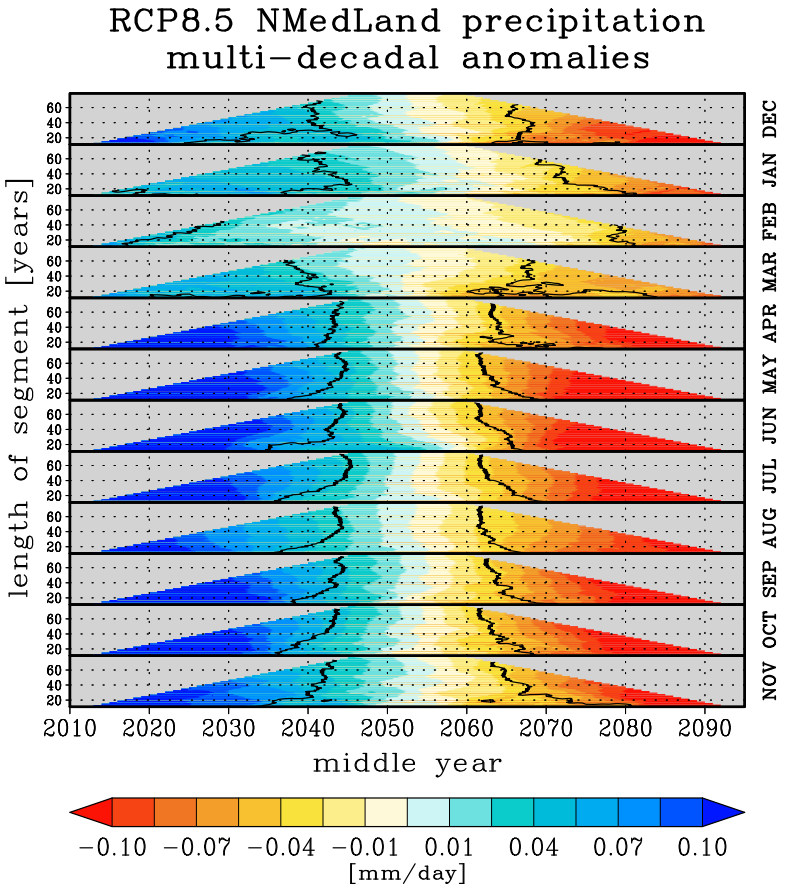


Figure 6.1: 2DMDA diagram (see Sec. 4.1) for each month using multi5 mean precipitation averaged over the NMedLand region under the RCP8.5 scenario. Anomalies are relative to the 2007-2099 mean.

anomalies (estimated with respect to the 1907-2099 mean) during the first and last decades of the series, in winter months these anomalies are weaker and even barely statistically significant in January and February. This seasonal distribution of precipitation changes matches that found in historical simulations in Chap. 5 (see Fig. 5.13).

To facilitate inter-month comparisons, in Fig. 6.2a we represent the multi-model NMed-Land precipitation trends per month for the 2007-2099 interval. The figure illustrates again the seasonal differences between winter (from December to March) and an extended warm season (from April to November). This seasonal contrast is stronger in RCP4.5 projections, for which January and February even exhibit weak wetting trends (not shown). This result is in agreement with Mariotti et al. (2015) (see their table 5), who calculate RCP4.5 2005-2098 precipitation trends over the northern Mediterranean (10°W-40°E/38°N-47°N) and find wetting trends in winter (December-February) and drying trends in summer (June-August). Looking at the multi-model average of relative trends (dividing the trend by each model's 1902-2004 climatology), the contrast is even clearer as a result of the strong annual cycle of precipitation (Fig. 6.2b).

To test whether this seasonal distribution of precipitation trends is a common feature of individual climate models we calculate, for each model, the three months with the strongest drying (defined as the largest 2007-2099 negative relative trend) and the three months with the strongest wetting (defined as the largest positive trend). In Fig. 6.3 we represent the results, indicating with red boxes the months with strongest drying and with blue boxes the months with the strongest wetting. For those models without three months with wetting, we indicate those months exhibiting the weakest drying (green boxes). For climate models with several runs available, we use the model ensemble mean. In general, months with the

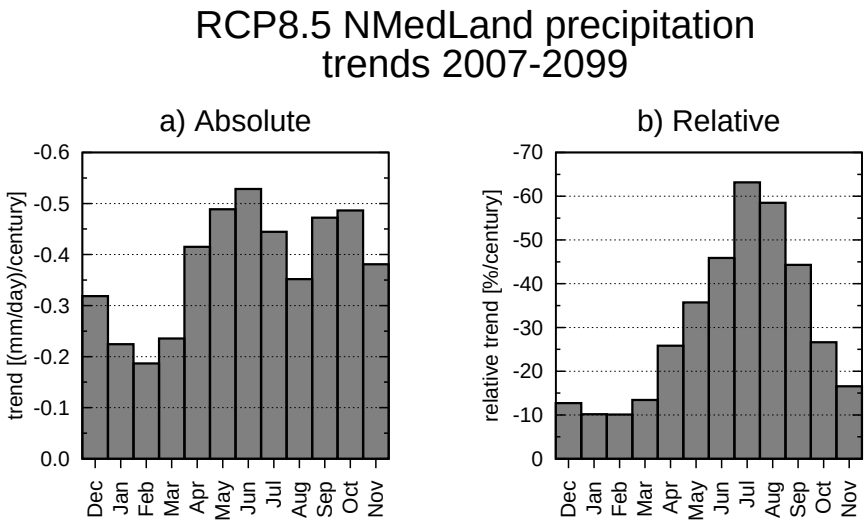


Figure 6.2: (a) 2007-2099 NMedLand precipitation trends per month for the CMIP5 multi-model ensemble under the scenario RCP8.5. (b) Average relative trends with reference of each model's 1902-2004 climatology.

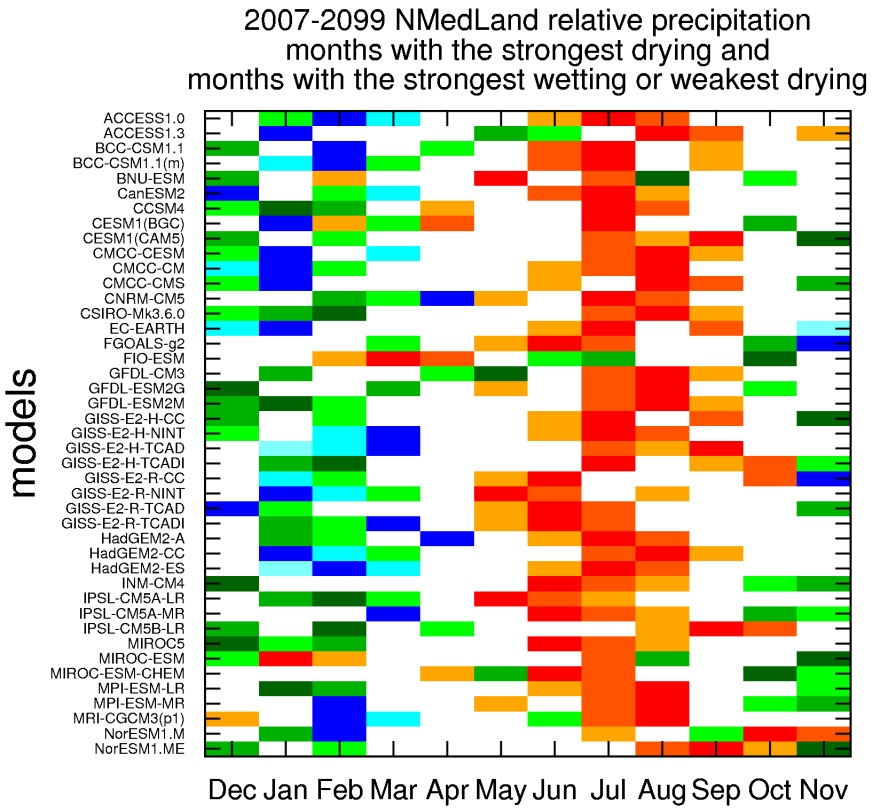


Figure 6.3: For each climate model, the 3 months in which the 2007-2099 NMedLand drying (estimated as the negative linear trend) is strongest are indicated in red (darker red means stronger drying) and the 3 months in which the wetting (estimated as the positive linear trend) is strongest are indicated in blue (darker blue means stronger wetting). When fewer than 3 models exhibit wetting, the months with weakest drying are indicated in green (lighter green means weaker drying). For models with more than one simulation available, results correspond to the model mean.

strongest drying are found in the warm season (from June to September) and months with the strongest wetting or the weakest drying are found in the cold season (from November to March), except for the ACCESS1.3 and the FIO-ESM models. These two models still exhibit weak drying in summer months, but the strongest decreases are found in autumn for ACCESS1.3 and in spring for FIO-ESM. We thus conclude that there is a good agreement among models regarding the seasonal distribution of the projected NMedLand precipitation relative trends. Using absolute trends we find similar results.

To emphasize the strong inter-model agreement regarding future drying in summer, in Fig. 6.4 we now represent the percentage of individual projections exhibiting a positive 2007-2099 NMedLand precipitation trend per month. Under the RCP8.5 scenario, the drying signal in summer is so strong that each of the 84 available runs exhibits negative trends in May, June, July and August. Instead, in January and February about one in every three projections exhibits wetting trends.

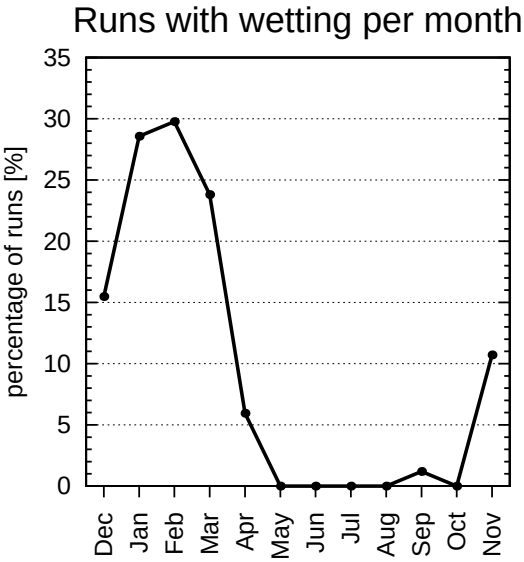


Figure 6.4: Percentage of individual runs within the CMIP5 ensemble with positive NMedLand winter precipitation trends under the RCP8.5 scenario for each month.

6.2 Geographical distribution and magnitude of the trends

Next we investigate whether climate models agree in the magnitude and spatial distribution of the projected changes. Because conclusions are different for summer and for winter, we study trends in these two seasons separately. Summer will be defined as the June-August mean and winter as the December-March mean for consistency with previous chapters. The year designating a particular winter is the one in which it ends.

6.2.1 Summer trends

According to the multi5 mean, all grid-points in the northern Mediterranean region exhibit future summer precipitation negative trends during the 2007-2099 interval (Fig. 6.5a). This widespread drying is robust among models (Fig. 6.5b), as the inter-model agreement regarding the sign of the change is larger than 90% in most of the basin (except in the south-eastern limit of the region, in which the inter-model agreement is still larger than 70%). The agreement is stronger than that presented in the AR5 [Christensen et al., 2013], in which the agreement in the sign of the change (estimated as the difference between the 2080-2099 and 1986-2005 means) is lower than 90% in virtually all the Mediterranean region (see Fig. 1.11). We have checked that this discrepancy is associated with the fact that we are examining projections for the RCP8.5 scenario, the one with the largest forcings prescribed, whereas results in AR5 are based on the intermediate RCP4.5 scenario (not shown).

Even though the sign of future trends is uniform in all the basin, multi5 trends exhibit a land-sea contrast, with strongest trends over land, in particular over mountain regions. To

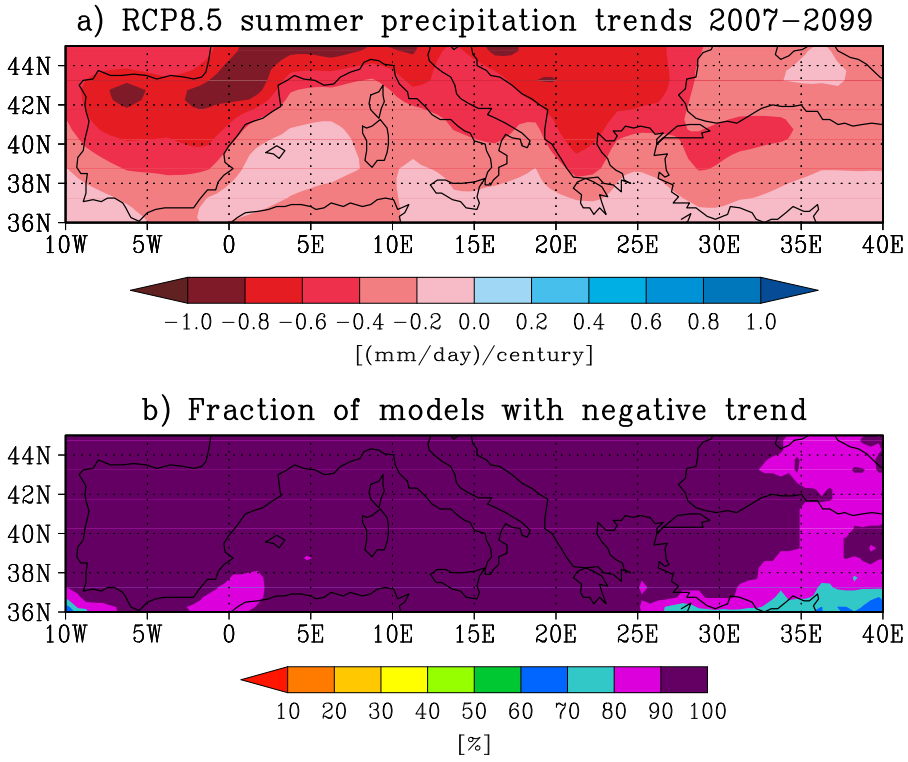


Figure 6.5: (a) 2007-2099 summer (JJA) precipitation trends in the CMIP5 multi-model ensemble under the RCP8.5 scenario. Trends are statistically significant in all grid-points. (b) Percentage of runs with negative precipitation trend per grid-point.

assess whether this contrast is also present in single projections, in Fig. 6.6 we represent the trend over land against the trend over sea for each individual projection. Because in all simulations the drying over land is stronger than over sea, we conclude that the inter-model agreement regarding the land-sea contrast of future summer precipitation trends is large. Consistent with the fact that in the Mediterranean region summer precipitation climatological amounts are larger over land than over sea (Fig. 6.7a), when we calculate relative trends the land-sea contrast vanishes and the largest drying is found in southern Iberia and in the central Mediterranean sea including the south of the Italian and Balkan peninsulas (Fig. 6.7b).

Figure 6.6 also illustrates that there is a large inter-model spread in the magnitude of the area-averaged summer precipitation trends. Over land, the area-averaged drying ranges from −0.16 (mm/day)/century to −1.00 (mm/day)/century; over sea, it ranges from −0.03 (mm/day)/century to −0.66 (mm/day)/century. In Fig. 6.8 we represent the area-averaged NMedLand summer precipitation relative trend in the 2007-2099 interval for each CMIP5 model (averaging results of projections of the same model). For models with several runs available, black whiskers indicate the minimum and the maximum trend, and the filled fraction of the bar is proportional to the fraction of runs with statistically significant trend (with a 95% confidence level). Again, the spread regarding the magnitude of the drying

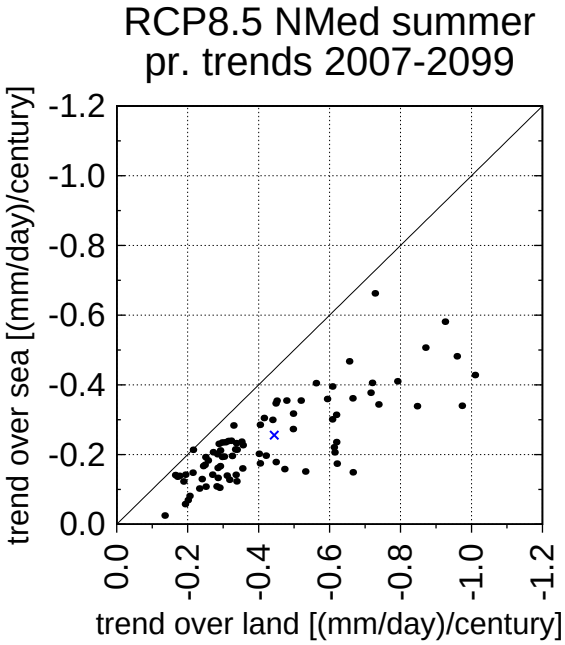


Figure 6.6: Scatter plot of 2007-2099 NMed summer precipitation trends over land and over sea in CMIP5 RCP8.5 projections. The blue cross indicates the multi-model mean.

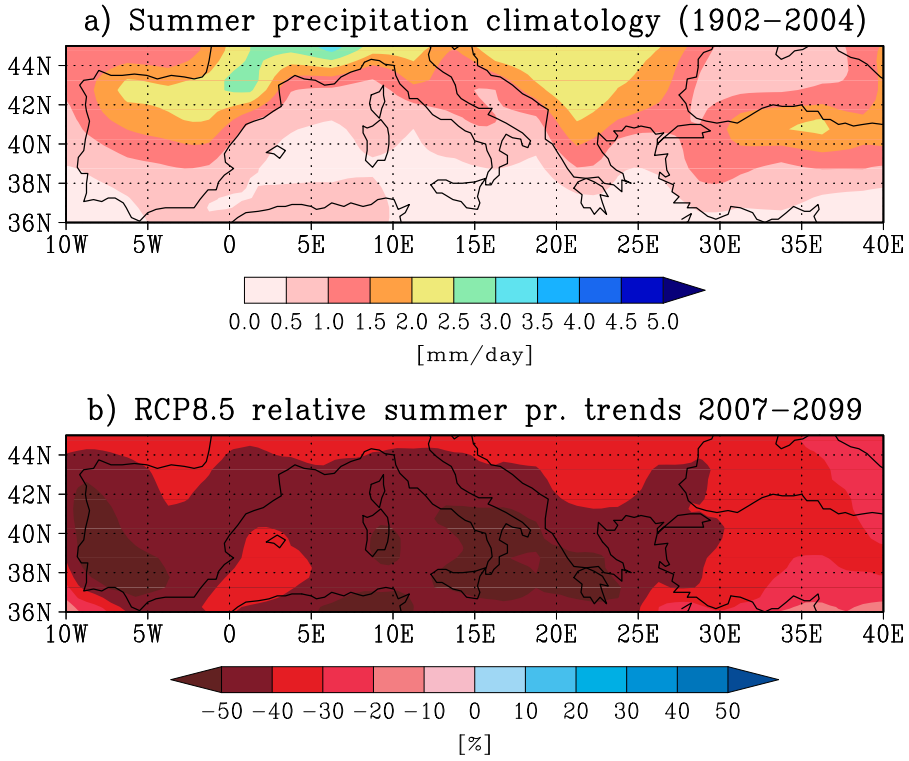


Figure 6.7: (a) 1902-2004 climatological summer precipitation according to historical CMIP5 simulations (multi-model mean). (b) 2007-2099 summer precipitation relative trends averaged for all the CMIP5 models under the RCP8.5 scenario (with respect to the climatological field).

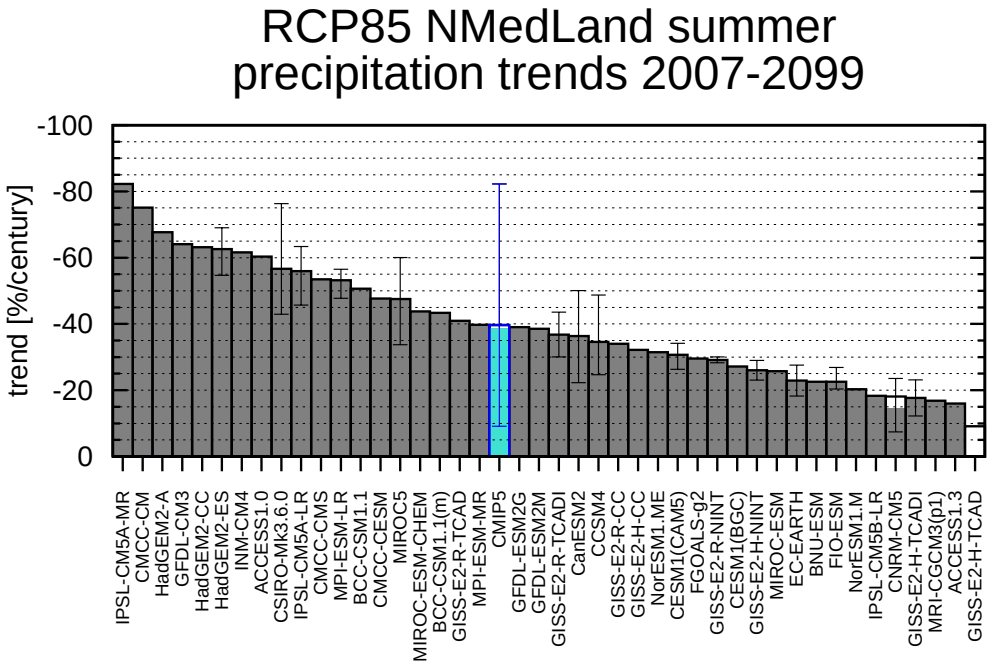


Figure 6.8: 2007-2099 NMedLand summer precipitation trends in all RCP8.5 projections. Relative trends estimated with reference to the 1902-2004 mean in historical simulations. The blue bar indicates the multi5 mean. Filled bars correspond to statistically significant trends. In those models with several runs, the model-mean is represented, the filled fraction of each bar is proportional to the fraction of runs with statistically significant trends, and black whiskers indicate the range between the minimum and the maximum trend in individual runs. The blue whisker indicates the minimum and the maximum model-mean trends.

in individual models is large (ranging from −82.2%/century to −9.1%/century). We also find that a vast majority of individual projections exhibit statistically significant trends, indicating than the magnitude of the forced signal is strong with respect to the magnitude of internal variability. In conclusion, while there is a complete inter-model agreement regarding the sign of the change, uncertainty in the magnitude is large.

In order to explore whether the inter-model spread in the magnitude of the precipitation change is related to disagreement among models in the magnitude of the projected warming we calculate 2007-2099 NMedLand summer temperature trends in individual CMIP5 projections. Since the correlation between precipitation and temperature trends (having extracted the global temperature trend) is very low (r=−0.16), we conclude that discrepancies among models regarding the magnitude of the warming are not directly responsible for the wide range of projected precipitation trends.

As discussed above, Rowell and Jones (2006) suggest that the fact that surface air warming is greatest over land than over sea (due to the large thermal inertia of water) contributes to the summer drying over land. To test this hypothesis we calculate, for each individual run, the differential land-sea summer warming and the summer precipitation

trend over land. Even though the obtained correlation between precipitation and differential warming trends is very low (−0.09), the scatter plot (Fig. 6.9) shows that projections exhibiting strongest drying are generally among those with large differential warming. This result, however, must be considered as only preliminary as it could be influenced by the fact that, instead of separating land grid-points from sea grid-points using each model's own land-sea mask, we use the same mask in all runs after having interpolated them to a common grid, as we pointed out in Chap. 2.

To evaluate our confidence in the summer drying, in Bladé et al. (2012a) we explored the role of the SNAO, the main driver of summer precipitation in the northern Mediterranean, in the projected precipitation trends. As discussed in detail in Sec. 5.5.1, the amount of SNAO-related precipitation in positive SNAO summers is underestimated in climate models. The consequence of this misrepresentation is that, given that models generally agree in projecting an upward SNAO trend for the twenty-first century, the SNAO-related increase in precipitation that should partially compensate the projected drying due to other mechanisms is not present in projections. This result, therefore, decreases our confidence in the correct representation of the magnitude of the drying trends.

6.2.1.1 Precipitation trends in central Europe

The future projected drying is not limited to the Mediterranean region but also covers central Europe. According to Haarsma et al. (2009), the projected drop of SLP in the Mediterranean area associated with the increase of temperature leads to enhanced dry easterly winds over central Europe and causes precipitation reductions there. As we discussed

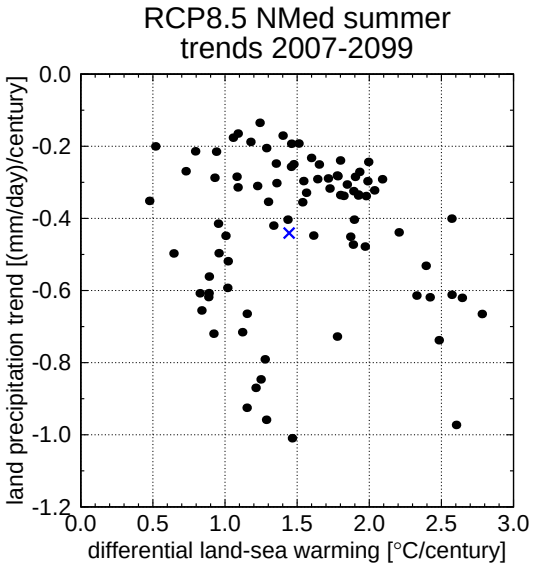


Figure 6.9: Scatter plot of the difference of 2007-2099 NMed summer temperature trends between land and sea against the total precipitation trend over land in individual CMIP5 RCP8.5 projections. The blue cross indicates the multi-model mean.

in Bladé et al. (2012b), the thermal low in the models could be exacerbated by the lack of the SNAO-related precipitation. If the link between SNAO and precipitation was well captured, therefore, precipitation decreases over central Europe could be weaker.

To explore this hypothesis, in Fig. 6.10a we represent the multi5 mean 2007-2099 summer SLP projected trend. The obtained pattern corresponds to the SNAO signature (see Fig. 5.38), with SLP decreases over the Mediterranean region, particularly strong in the eastern basin, and over latitudes higher than 50°N, and SLP increases in between, over the British Isles. Consistent with the changes in the wind field associated with this pattern, which bring dry air over central Europe, in the map of projected 2007-2099 summer precipitation trends (Fig. 6.10b) the maximum drying is found in this region. The scatter plot represented in Fig. 6.11 contains, for each available CMIP5 run, the 2007-2099 trend of summer SLP averaged over the eastern Mediterranean (10°E-35°E/30°N-45°N, box in Fig. 6.10a) and the associated 2007-2099 trend of summer precipitation averaged over central Europe (0°-20°E/40°N-50°N, box in Fig. 6.10b). Even though the correlation between the two variables is low (r=0.18), we observe a weak tendency for the runs that exhibit the largest SLP decreases over eastern Mediterranean to also exhibit strong drying over central Europe (this tendency is more visible if we remove the CanESM2 model from the analysis). Removing the hemispherical mean from SLP trends does not change our result. Therefore, even though our results are only preliminary, we conclude that the strong central Europe drying expected for the end of the twenty-first century may be another bias of the CMIP5 models associated with their misrepresentation of the SNAO responses.

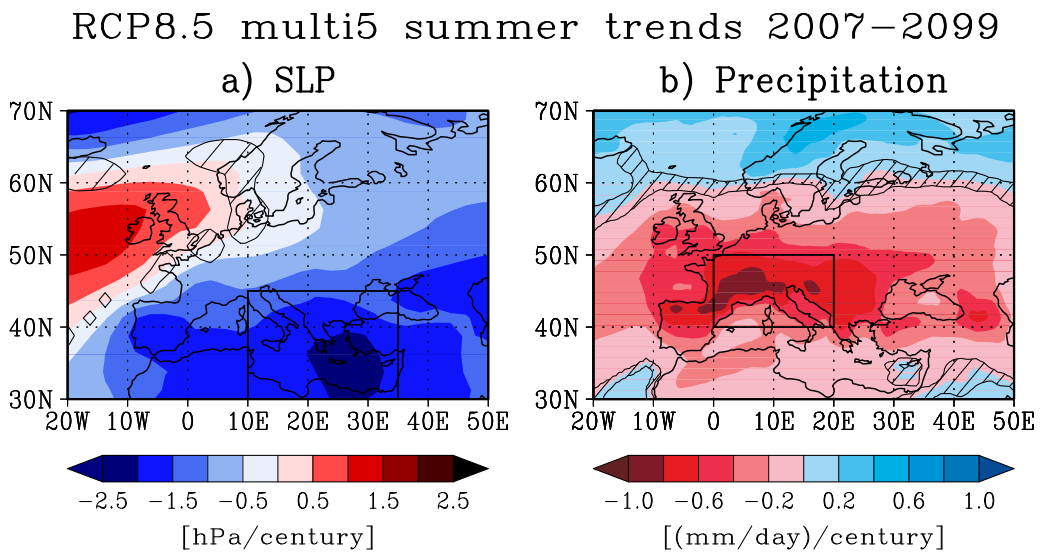


Figure 6.10: Multi5 mean summer SLP (a) and summer precipitation (b) trends 2007-2099. Dashed regions indicate non-significant trends. The box in (a) indicates the eastern Mediterranean region over which we average SLP, and the box in (b) indicates the central European region over which we average precipitation.

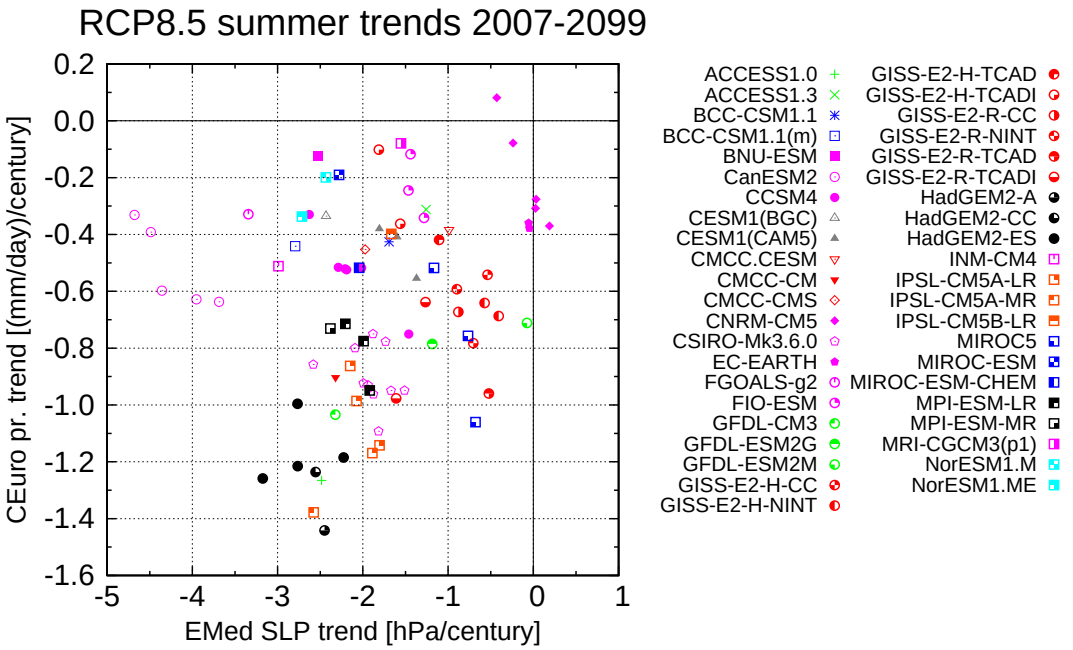


Figure 6.11: Scatter plot including 2007-2099 trends in summer SLP averaged over the eastern Mediterranean region (see figure 6.10a) and 2007-2099 trends in summer precipitation averaged over the central European region (see figure 6.10b) in all available CMIP5 runs under the RCP8.5 scenario.

6.2.2 Winter trends

Unlike in summer, the sign of the multi5 precipitation trend in winter varies meridionally across the basin (Fig. 6.12a). Projected trends are generally negative and statistically significant over regions located south of 42°N, being strongest in north Africa and in the southern part of the Iberian, Balkan and Anatolian peninsulas. Instead, regions located in the northernmost part of the basin exhibit statistically significant future wetting trends. In between, regions located in a band of about 2° of latitude exhibit non-statistically significant trends. Relative trends exhibit a similar structure, in which the meridional gradient dominates (not shown). The inter-model agreement regarding the sign of the changes is weaker in winter than in summer (Fig. 6.12d), except for the south of the basin. Over the band of non-significant trends, about half of the models project drying and half of the models project wetting. In Figs. 6.12b-c we display two extreme cases. While run r1i1p1 of the BNU-ESM model exhibits strong and statistically significant winter precipitation decreases all across the basin, in run r4i1p1 of the HadGEM2-ES model positive winter precipitation trends, generally non-significant, prevail.

The NMedLand areal average of projected winter trends, therefore, arises from the negative trends that occur over the southern and central parts of the basin, partially compensated by the positive trends in the north in some models. As a result, unlike in summer, winter area-averaged trends are generally weak with respect to internal variability, shown

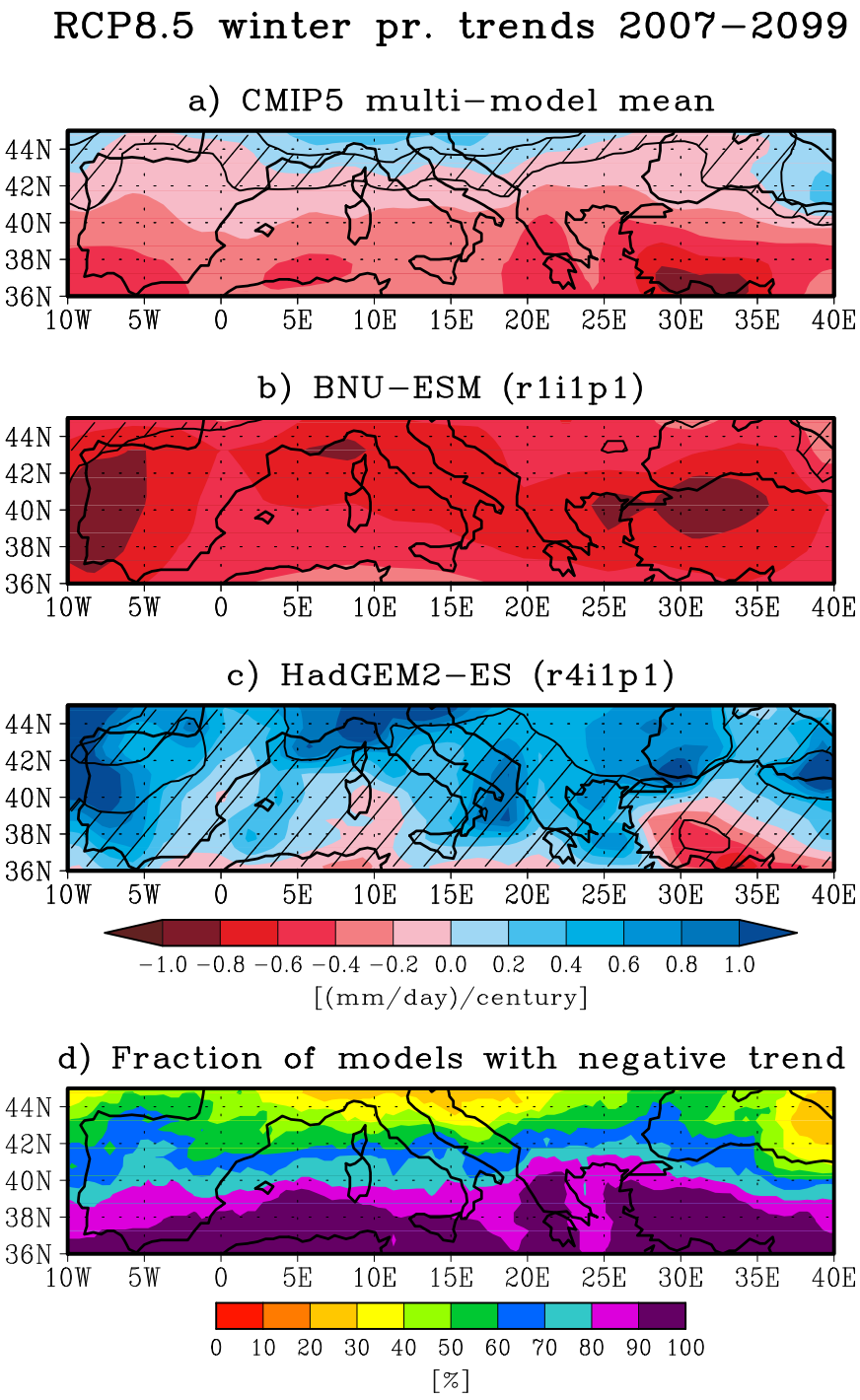


Figure 6.12: (a) 2007-2099 winter (DJFM) precipitation trends in the CMIP5 multi-model ensemble under the RCP8.5 scenario. Hatching indicates non statistically significant trends. (b-c) The same for the run r1i1p1 of the BNU-ESM model and the run r4i1p1 of the HadGEM2-ES model. (d) Percentage of runs with negative precipitation trends per grid-point.

by the fact that most individual runs exhibit non-significant trends (Fig. 6.13). The inter-model agreement regarding the sign of the winter area-averaged trends is also weaker than in summer, with 6 models exhibiting positive instead of negative projected precipitation trends. In RCP4.5 projections, in which the forcing is weaker, the amount of statistically significant winter trends is even lower and the inter-model agreement regarding the sign of the change is weaker (not shown).

Because the NAO is the main modulator of precipitation variability in the northern Mediterranean region, we assess whether the weakness of winter precipitation trends is consistent with a lack of strong NAO trends in projections. Following the methodology described in previous chapters, the NAO is calculated as the first EOF of the observed winter north Atlantic SLP projected onto the SLP field of each run (see Sec. 3.2). Studies about the NAO in projections under the RCP8.5 scenario conclude that future NAO trends are expected to be positive but weak and with poor inter-model agreement [Gonzalez-Reviriego et al., 2012; Cattiaux et al., 2013; Cattiaux and Cassou, 2013]. These results are confirmed by Fig. 6.14, in which we present 2007-2099 NAO trends for all models. First, although the multi-model mean trend is positive, 11 out of 43 models project downward trends (although none of them is statistically significant) and 10 out of the 15 models with

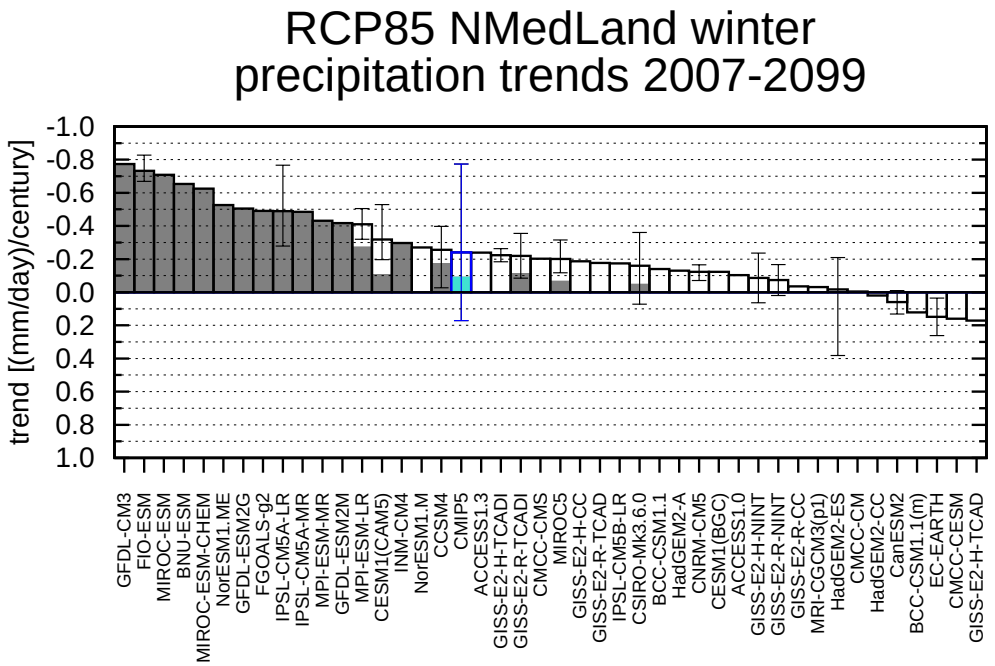


Figure 6.13: 2007-2099 NMedLand winter precipitation trends in all RCP8.5 projections. The blue bar indicates the multi5 mean. Filled bars correspond to statistically significant trends. In those models with several runs, the model-mean is represented, the filled fraction of each bar is proportional to the fraction of runs with statistically significant trends, and black whiskers indicate the range between the minimum and the maximum trend in individual runs. The blue whisker indicates the minimum and the maximum model-mean trends.

several runs exhibit both positive and negative trends. Second, even the positive projected trends are generally not statistically significant (except for the strongest ones).

The weakness of winter precipitation trends, therefore, is consistent with the lack of strong NAO trends. In addition, the spatial pattern of multi5 precipitation trends does not match the pattern of variability associated with the NAO (Fig. 6.15a). To illustrate this and assess the role of non-NAO mechanisms in the projected winter precipitation trends, we separate precipitation series for each individual run in its component linearly related to the NAO (*pr-nao*) and the residual (*pr-nonao*). As we did in Sec. 4.5, to estimate the first component we calculate the regression between precipitation and NAO and calculate its product with the value of the NAO index at each time instant. To estimate the second component we remove pr-nao from the total precipitation series.

Figs. 6.15b-c show the 2007-2099 multi-model average trends of pr-nao and pr-nonao (having averaged runs of the same model first). Because pr-nao is the component of precipitation linearly related to the NAO, the pattern of pr-nao trends coincides with that of the precipitation-NAO regression. Comparing Fig. 6.15b and Fig. 6.12a we see that the magnitude of pr-nao trends is much weaker than the magnitude of full precipitation trends

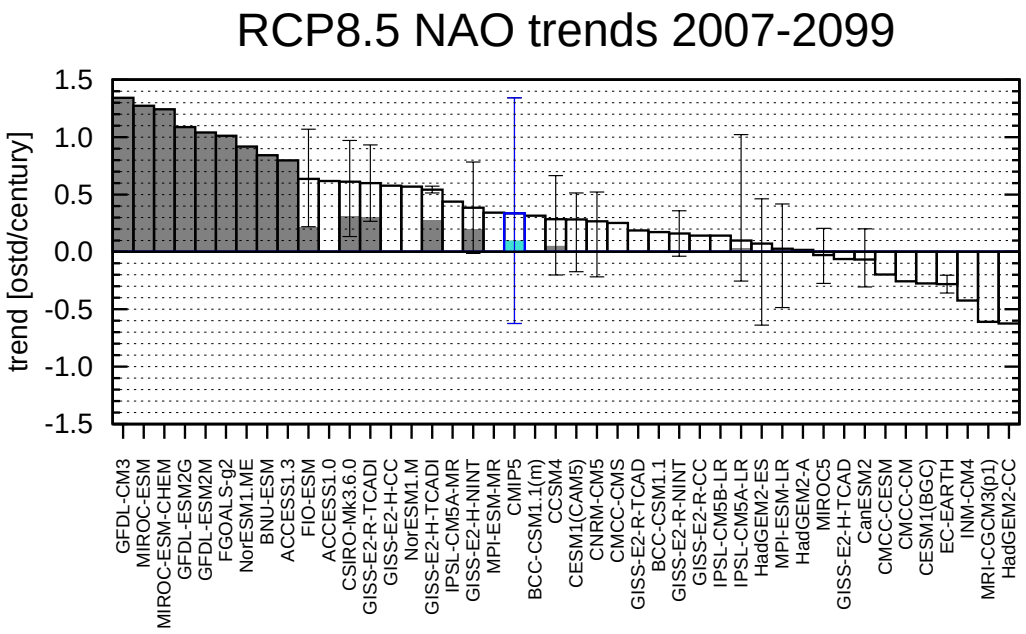


Figure 6.14: Linear trend (2007-2099) of the NAO index in all CMIP5 models under the RCP8.5 scenario. The blue bar indicates the multi5 mean. The blue bar indicates the multi5 mean. Filled bars correspond to statistically significant trends. In those models with several runs, the model-mean is represented, the filled fraction of each bar is proportional to the fraction of runs with statistically significant trends, and black whiskers indicate the range between the minimum and the maximum trend in individual runs. The blue whisker indicates the minimum and the maximum model-mean trends.

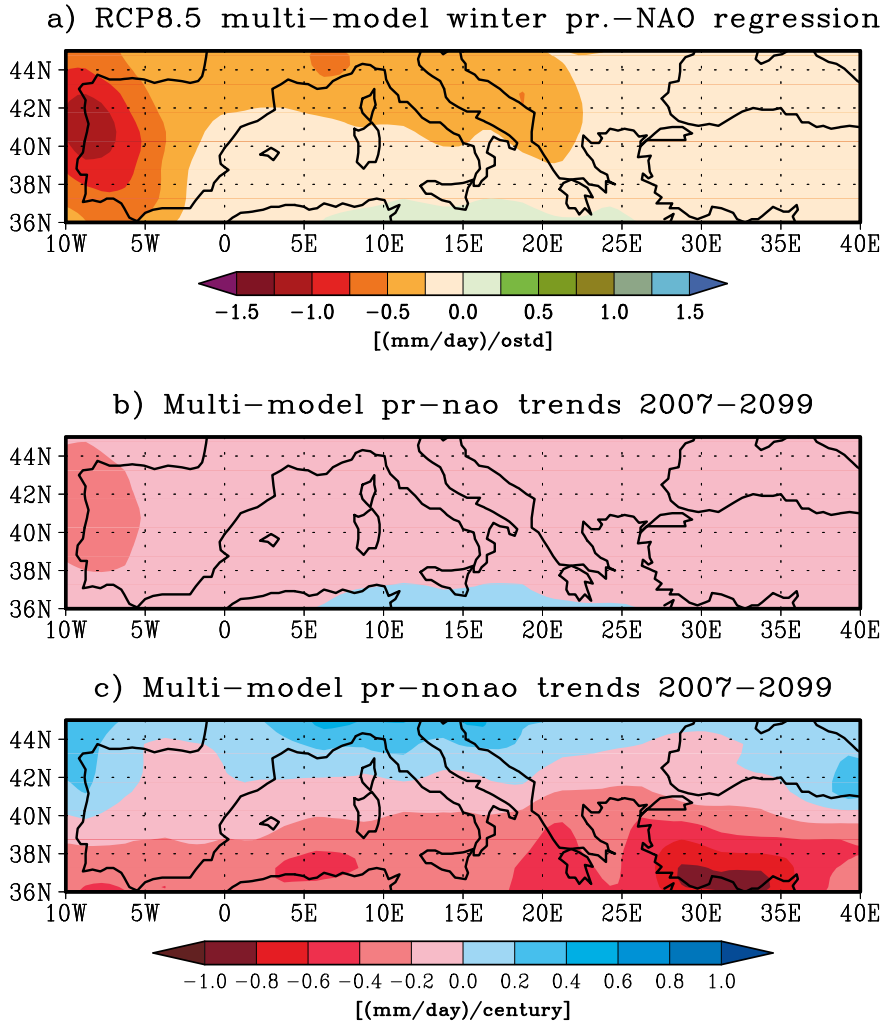


Figure 6.15: (a) Multi-model average of regressions between winter precipitation and the NAO index under the RCP8.5 scenario. Regressions have been calculated first for each available projection and model averaged. (b) 2007-2009 multi-model average of trends of the component of winter precipitation linearly related to the NAO. (c) The same for the component not linearly related to the NAO.

and that the geographical patterns do not coincide. Instead, the full pattern of projected precipitation trends highly resembles the pattern of pr-nonao trends (Fig. 6.15c).

Therefore, we conclude that the weak future winter precipitation trends in the northern Mediterranean obtained under the RCP8.5 scenario are mostly due to mechanisms not linearly related to the NAO. This may include the poleward expansion of the subtropical precipitation minima described by Scheff and Frierson (2012) associated with the expansion of the Hadley cell [Lu et al., 2007].

6.3 Transition towards dry conditions

In a previous section we have found that, even though all models project a future summer drying the northern Mediterranean region, there is a large spread in the magnitude of the change. As important as the magnitude of the change is how precipitation will behave during the twenty-first century and in particular how fast precipitation will decrease beyond the range of natural variability. In this section we focus on the transition towards dry conditions, analysing which aspects of individual projections are robust across models. Because changes in summer are stronger and more robust across models, we generally focus on summer projections. Those aspects with weak inter-model agreement in summer are expected to have even weaker inter-model agreement in seasons in which the signal is not so strong.

First, we study the amplitude of decadal to multi-decadal variability of summer precipitation. Then, we examine whether the transition towards drier conditions is gradual or not, analysing whether each decade within the twenty-first century is drier than the former one. Finally, we estimate for each projection the time of emergence of the forced signal, i.e. the year in which the forced signal will stand above internal variability.

6.3.1 Amplitude of multi-decadal variability

To estimate the amplitude of the multi-decadal variability of NMedLand summer precipitation we use the detrended standard deviation of the filtered series in the 2007-2099 interval (using a low-pass symmetric non-recursive filter with a cut-off period of 10 years, see Sec. 3.1.7). In Fig. 6.16a we represent the histogram of standard deviations obtained with all individual projections. In Fig. 6.16b we represent the relative standard deviations, calculated dividing each standard deviation by the climatological 1902-2004 mean obtained from the corresponding historical simulation. In Fig. 6.16a we see that there is a large inter-model spread in the magnitude of the multi-decadal variability, which ranges from 0.041 mm/day (for the run r1i1p1 of the CMCC-CM model, see its time series in Fig. 6.17a) to 0.176 mm/day (run r1i1p1 of the GISS-E2-R-CC model, Fig. 6.17b). Relative standard deviations range from 2.8% to 16.6%.

To check whether the variability of NMedLand precipitation in climate models has realistic magnitude, we compare the standard deviation of observed filtered summer precipitation (GPCC) with the standard-deviation of pre-industrial control simulations (Fig. 6.18). Because the observed standard deviation (0.077 mm/day; red line) is within the range of simulated standard deviations, we conclude that the order of magnitude of the simulated multi-decadal variability of summer precipitation is realistic. Moreover, even though in some models the amplitude of the multi-decadal variability seems to change between historical simulations and projections (as is the case of the example in Fig. 6.17b), comparing

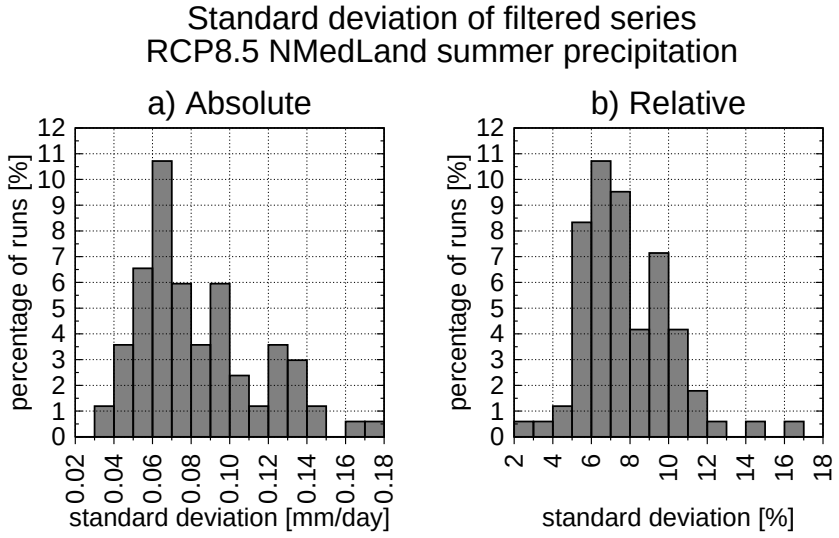


Figure 6.16: (a) Histogram of standard deviations of NMedLand summer precipitation series in individual RCP8.5 projections during the 2007-2099 interval. Series have been detrended and filtered (cut-off period of 10 years) before the calculation. (b) Relative standard deviations, obtained dividing the actual standard deviations by the climatological mean NMedLand summer precipitation (1902-2004) of each run.

Fig. 6.16 and Fig. 6.18 we see that the total inter-model range of standard deviations is about the same in both sets of experiments (examining historical standard deviations we obtain a similar histogram, not shown).

As commented in detail in Sec. 5.4.1.2, assessing whether this is also true for winter is not straightforward. Due to the strong winter precipitation trend observed in the last decades of the twentieth century, the estimation of the observed standard deviation is highly sensitive to the choice of time interval. With this caveat in mind, we used a band-pass filter to remove high-frequency variability and long-term trends and found the observed standard deviation inside the distribution of pre-industrial standard deviations (see Fig. 5.19).

6.3.2 Gradualness of the drying

In order to complement the former result, we estimate how gradual is the transition towards drier conditions in individual projections. We consider a drying as gradual if each decade within the twenty-first century is drier than the former one. An example of a gradual change is that of the observed global mean surface air temperature, in which each of the the last three decades has been warmer than the previous one (see Fig. 1.1).

Figure 6.19 shows the time series of decadal NMedLand summer precipitation in each available RCP8.5 projection. Each curve consists of nine values corresponding to the nine decades between 2010 and 2099. Because we find as gradual only one projection of the HadGEM2-ES model (r3i1p1) and one projection of the MIROC5 (r2i1p1) model, while

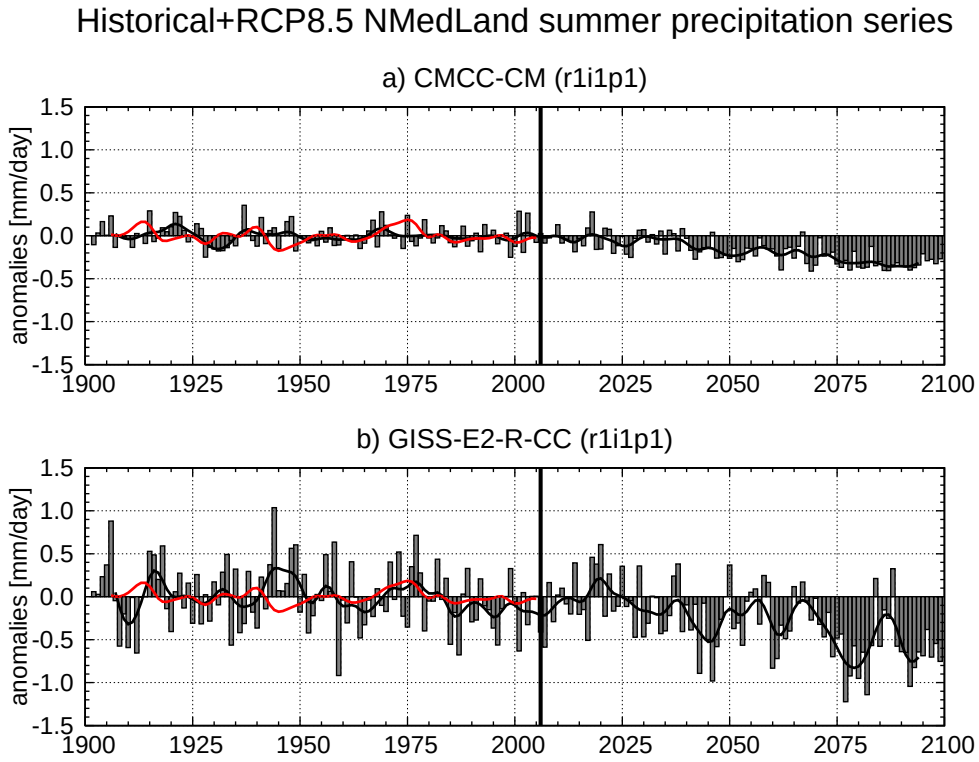


Figure 6.17: Anomalies of summer NMedLand precipitation in the historical simulations and RCP8.5 projections relative to the 1902-2004 interval according to the run r1i1p1 of the CMCC-CM model (a) and the run r1i1p1 of the GISS-E2-R-CC model. The black line indicates data filtered with a cut-off period of 10 years.

the remaining 82 projections (including all other runs from HadGEM2-ES and MIROC5) exhibit at least one decade wetter than the previous one, we conclude that models agree in projecting a non-gradual drying. If we restrict our analysis to the 2010-2050 interval, covering only four decades, we still find only 20 out of 84 runs with gradual drying.

Consistent with the inter-model spread found in decadal standard deviations (Fig. 6.16), the magnitude of the decadal fluctuations differs from one model to another. For instance, while decadal precipitation increases superimposed on the general drying signal are very weak in FGOALS-g2 or GISS-E2-H-CC, in other models such as GISS-E2-R-CC or CESM1(CAM5) they are on the order of 20% of the 1902-2004 climatological mean.

Looking at Fig. 6.19 we also see that future NMedLand winter precipitation exhibits a wide range of behaviours. First, in those models with the strongest projected trend (see Fig. 6.8) the future evolution of decadal-mean weak anomalies consists of decadal fluctuations superimposed on a persistent drying trend. This is the case, for instance, of runs from IPSL-CM5A-MR, CMCC-CM, HadGEM2-A, GFDL-CM3 and HadGEM2-CC. Second, some runs display an abrupt drying in a short lapse of two or three decades at some point of the twenty-first century. This rapid precipitation decrease can be found at the beginning of the

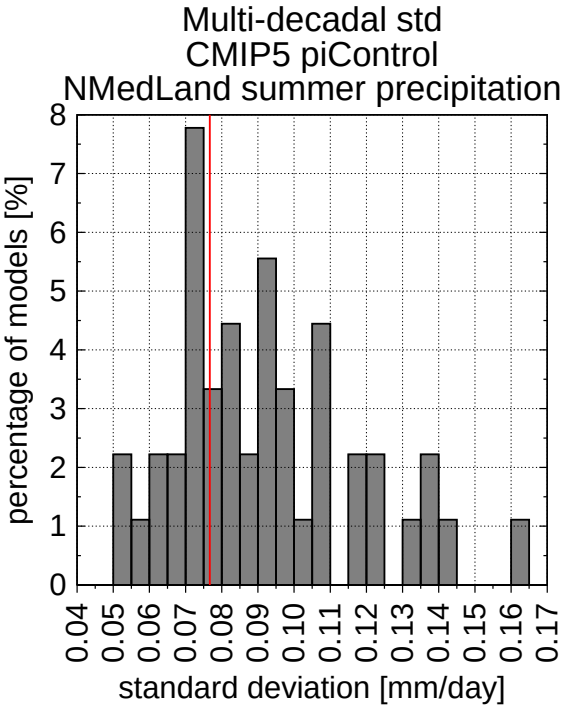


Figure 6.18: Distribution of standard deviation in CMIP5 pre-industrial control simulations. All time series have been detrended prior the calculation and a high-pass symmetric non-recursive filter with cut-off period of 10 years has been applied. The red vertical line indicates the observed standard deviation in the 1902-2004 interval (GPCC).

twenty-first century (NorESM1.ME) in the central decades (GISS-E2-R-TCAD) or at the end of the century (GFDL-ESM2G). Finally, the decadal evolution of summer precipitation in models with the weakest projected trend is generally dominated by an alternation of dry and wet decades with a very weak negative trend superimposed (ACCESS1.3; CNRM-CM5).

The wide spectrum of possible future precipitation behaviours contrasts with the uniformity of NMedLand summer temperature projections (Fig. 6.20). In addition, unlike for precipitation (as we shall see later), the signal in temperature is so strong that all decadal means within the twenty-first century are warmer than the 1902-2004 climatological mean in all models.

6.3.3 Time of emergence

The time of emergence (TOE) of a climate signal is defined as the year in which a persistent forced signal is evident above internal climate variability [e.g. Giorgi and Bi, 2009; Hawkins and Sutton, 2012]. It can be estimated as the time in which the signal-to-noise ratio of a climatic series become larger than a particular threshold. Because there is not a unique method of estimating the magnitude of the signal and the amplitude of the natural noise, several authors have developed their own methodology to calculate TOEs. For

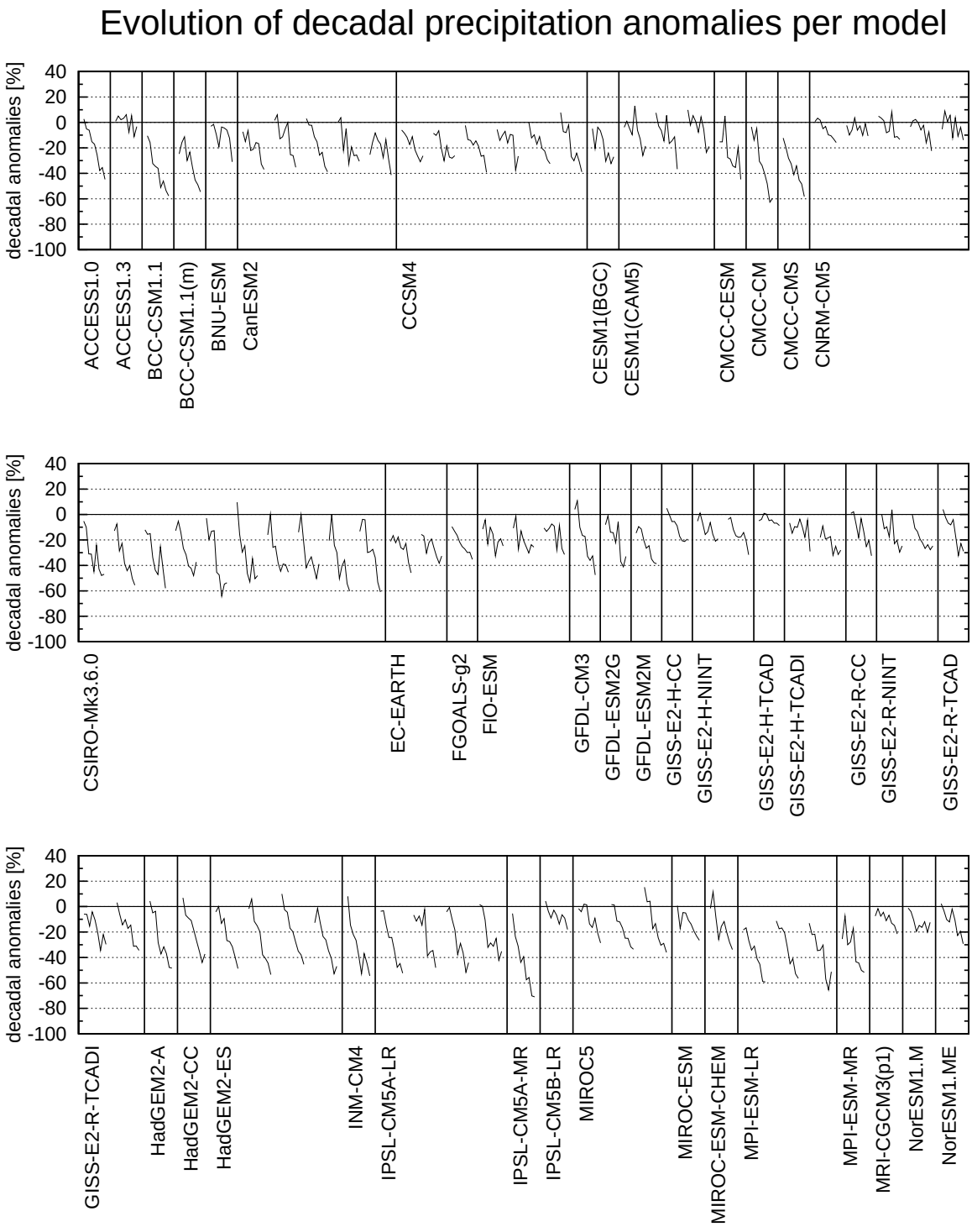


Figure 6.19: 2010-2099 relative decadal anomalies of NMedLand summer precipitation for each available RCP8.5 projection. Anomalies are relative to the 1902-2004 mean in each associated historical simulation.

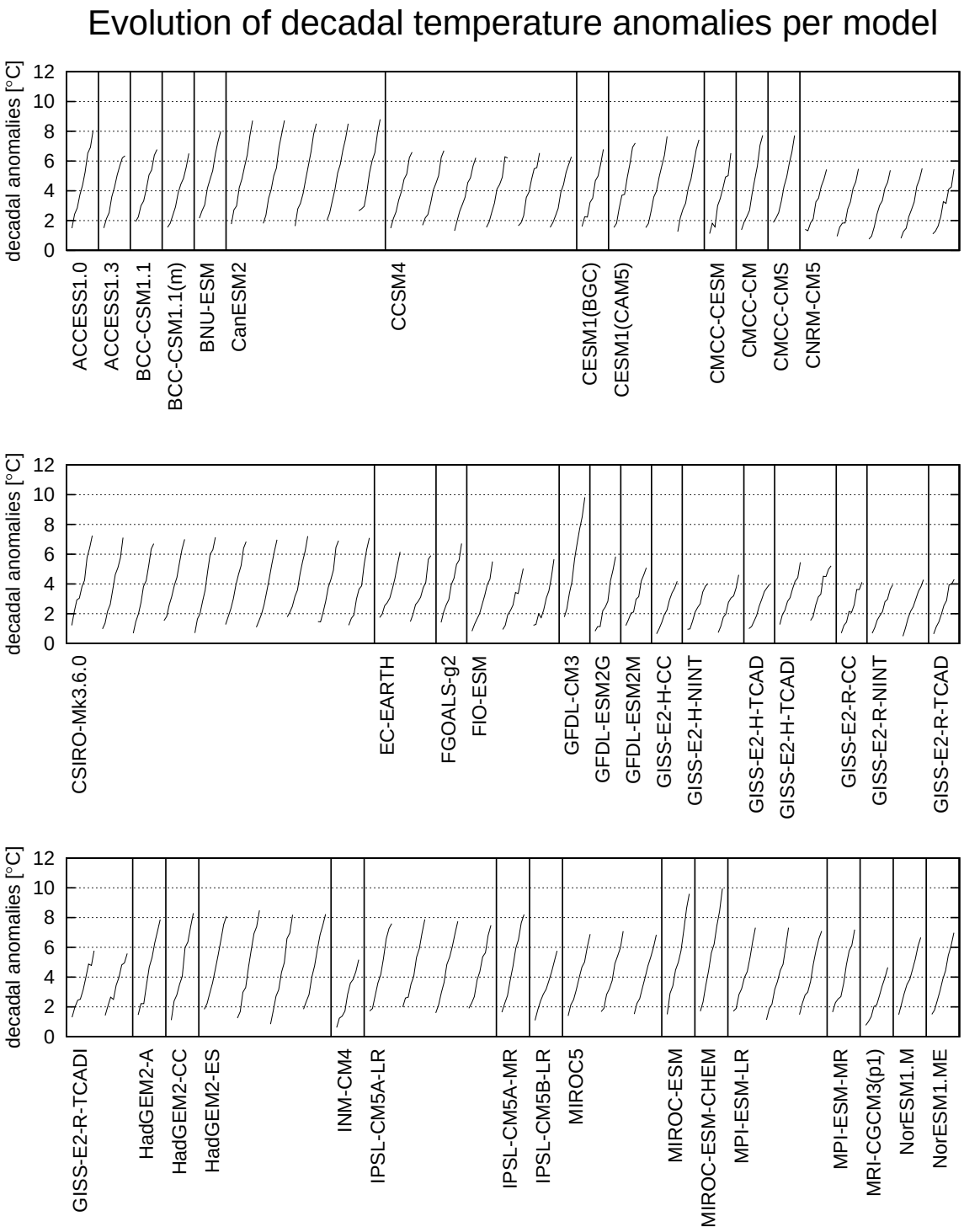


Figure 6.20: 2010-2099 decadal anomalies of NMedLand summer surface air temperature for each available RCP8.5 projection. Anomalies are relative to the 1902-2004 mean in each associated historical simulation.

instance, Giorgi and Bi (2009) define the TOE of precipitation as the first year in which the signal is larger than the noise and does not recover thereafter. They define the signal as the multi-model mean 20-year average precipitation anomaly with respect of the 1980-1999 interval. They define the amplitude of the noise as the root mean square error of the sum of the inter-model standard deviation (using model means resulting of averaging runs of the same model) and an estimation of the intra-model variability (the difference between the change in each run and the change in the model-mean in those models with several runs available). This methodology, however, does not provide a TOE for individual projections. For this reason, we follow Hawkins and Sutton (2012), whose methodology is valid for individual runs and allows us to assess the inter-model agreement of the results. This methodology consists of finding the first year in which the signal-to-noise ratio exceeds 1. The signal is defined as $S(t) = \alpha T_G(t) + \beta$ where T_G is the filtered global mean surface temperature in the model and α and β are the regression coefficients between northern Mediterranean precipitation and T_G. Noise is defined as the standard deviation of pre-industrial precipitation series.

In Fig. 6.21 we show the TOE of each available RCP8.5 projection. The TOE of the multi-model ensemble under the RCP8.5 scenario (blue line) is 2062, but the inter-model spread covers the entire twenty-first century. In Fig. 6.22 we represent the NMedLand summer precipitation series of two example models, indicating in blue the TOE year, in green the evolution of the signal and in red the amplitude of the noise. The run r1i1p1 of the FGOALS-g2 model (Fig. 6.22a) exhibits the earlier TOE, situated in 2015. Even though this run is not one of the runs with a strongest drying, it exhibits weak internal variability. Instead, according to run r10i1p1 of the CNRM-CM5 model (Fig. 6.22b) the TOE is expected in the year 2092. In this case, this model has one of the weakest drying signals. In some projections, in general those with weakest signals, the TOE is not reached in the 2007-2099 interval (indicated with red squares in the right axis of Fig. 6.21). Consistent with the general lack of strong signal in winter precipitation projections, the winter TOE is within the twenty-first century only in 11 models, those with the strongest 2007-2099 trends (not shown). Instead, because as seen in Sec. 6.3.2 the summer temperature signal is much stronger than that for summer precipitation, the TOE for summer temperature is located before 2012 in any single projection (not shown).

In conclusion, consistent with the large spread found in the magnitude of the forced signal (see Fig. 6.8) and the amplitude of internal variability (see Fig. 6.16), there is very large uncertainty regarding the time in which the northern Mediterranean summer drying signal will become evident above internal noise. Nevertheless, the fact that no run exhibits a TOE before 2015 is consistent with the lack of detection of the summer drying signal in observational series (see Chap. 5).

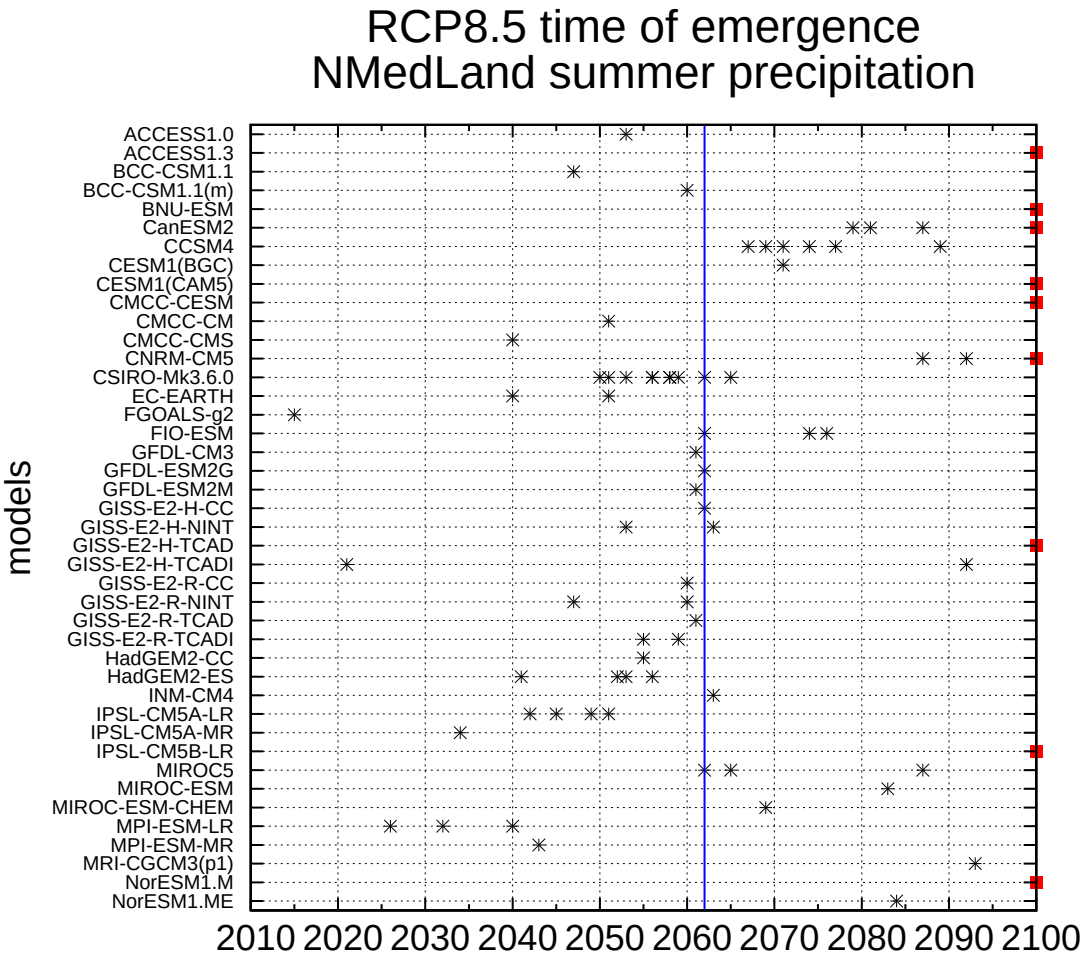


Figure 6.21: TOE of each available RCP8.5 projection in the CMIP5 ensemble (asterisks). The blue vertical line indicates the TOE of the multi-model ensemble mean. Red squares over the right axis indicate runs in which the TOE is not contained in the 2007-2099 interval. The HadGEM2-A model has been omitted due to the lack of pre-industrial control simulation.

6.4 Changes in decadal mean precipitation values

Finally we study how the magnitude of NMedLand summer precipitation decadal relative anomalies will change in the twenty-first century (2010-2099) according to the RCP8.5 scenario, taking as a reference the corresponding anomalies in historical simulations in the 1910-1999 interval. In Fig. 6.23 we represent the range between the driest and the wettest decade for each climate model (considering all its available runs) in the 2010-2099 interval (blue lines) and in the 1910-1999 interval (grey lines). A shift towards drier conditions is visible in all models, and in the majority of them the projected range of decadal precipitation is wider than the historical range (except for ACCESS1.3, EC-EARTH, GISS-E2-H-TCAD and NorESM1.M). Nevertheless, there is a large spread in the magnitude of the difference between the driest and the wettest decades in the 2010-2099 interval, ranging from 0.86 mm/day in IPSL-CM5A-LR to 0.18 mm/day in NorESM1.M. There

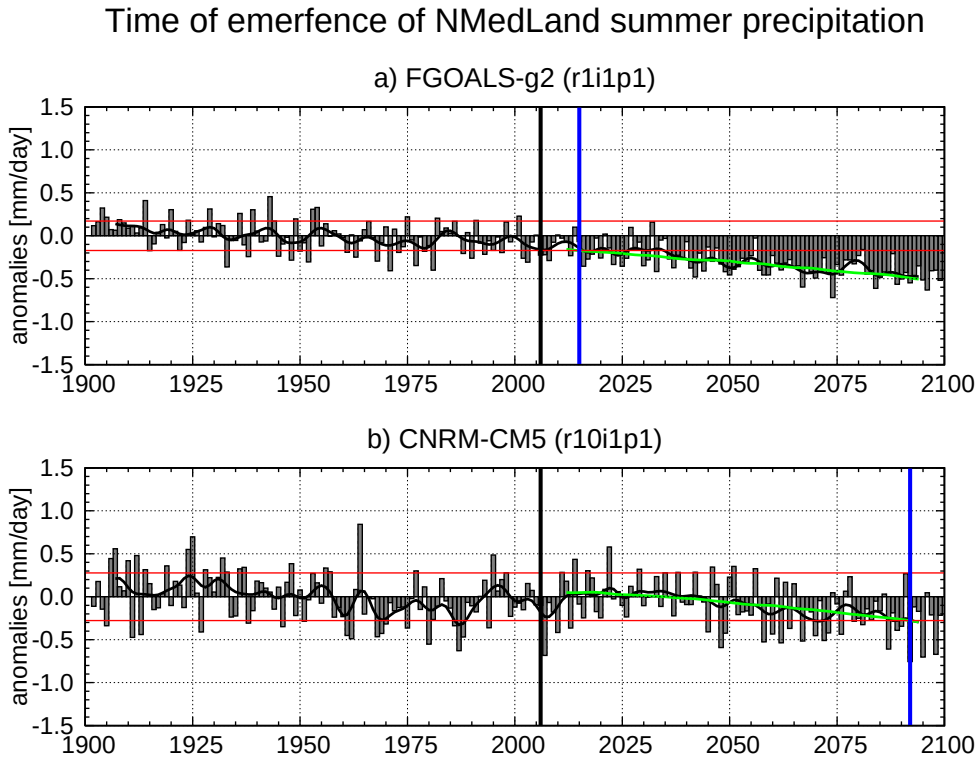


Figure 6.22: Anomalies of summer NMedLand precipitation in the historical simulations and RCP8.5 projections relative to the 1902-2004 interval according to the run r1i1p1 of the FGOALS-g2 model (a) and the run r10i1p1 of the CNRM-CM5 model. The black line indicates data filtered with a cut-off period of 10 years. Blue lines indicate the TOE year. Green curves indicate the forced signal, defined as the fraction of precipitation linearly related to the global mean surface air temperature increase. The red lines indicate the amplitude of internal noise, defined as the standard deviation of pre-industrial series.

is also large spread in the precipitation amounts in the driest decades, ranging from −0.84 mm/day in IPSL-CM5A-MR to −0.14 mm/day in ACCESS1.3. Both the total range and the minimum values are highly correlated with the magnitude of the 2007-2099 drying trend (r=0.90 and r=0.86). Moreover, in some models the projected range is barely overlapped with the historical range or they are even completely separated (in BCC-CSM1.1(m) and FGOALS-g2), whereas in other models the historical range is totally within the projected range (MIROC5 and MIROC-ESM-CHEM). Finally, while for some models all decades have lower precipitation than the climatological reference, other models exhibit both wet and dry decades. In general, the amplitude of the range between the wettest and the driest decade in the historical period is realistic, as the observational value (red bar, GPCC data) is within the spread of historical simulations.

Results for summer precipitation contrast with those for summer temperature (Fig. 6.24). Because of the strength of the summer warming signal, in particular under the RCP8.5 scenario, all twenty-first century decades are warmer than all twentieth century

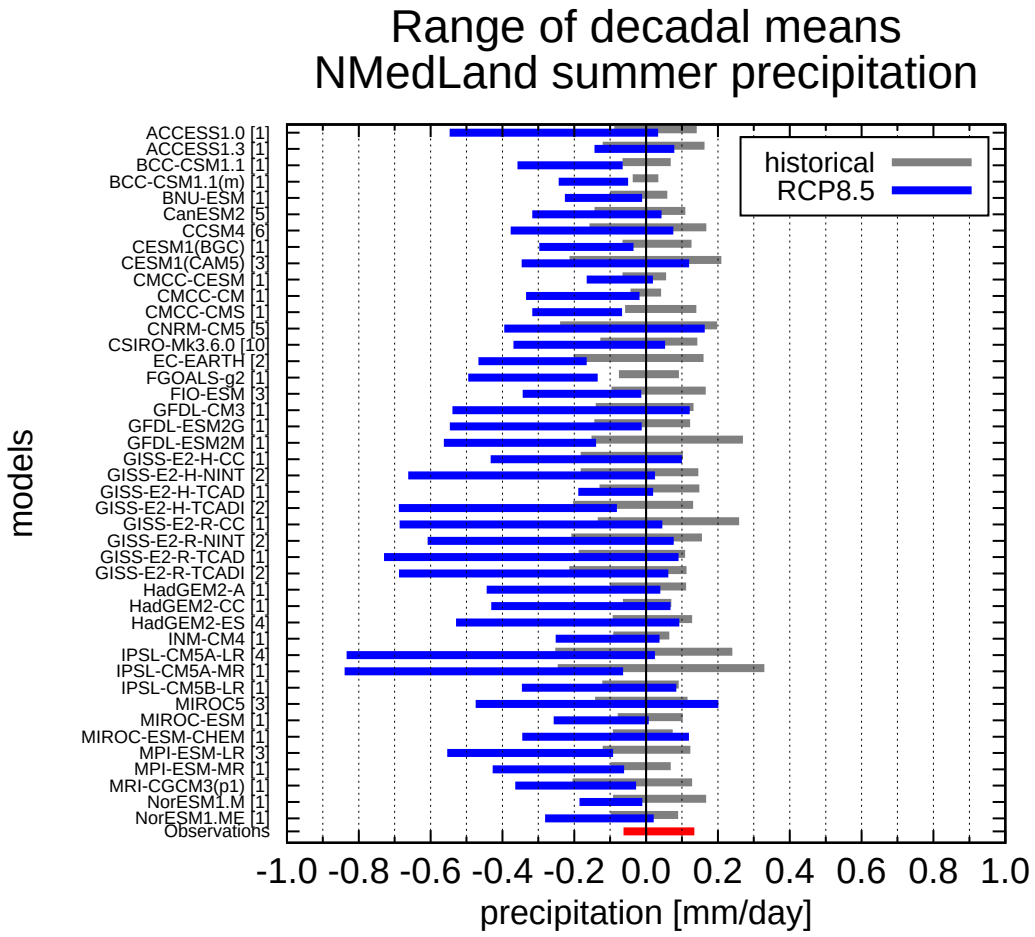


Figure 6.23: Range between the driest and the wettest decades per model. Blue lines indicate RCP8.5 projections and grey lines indicate historical simulations. Decadal anomalies relative to the 1902-2004 mean. In brackets, after the name of each model, we indicate the number of simulations available. The red bar indicates observations (GPCC).

decades in any given climate model (consistently with results shown in Fig. 6.20). Moreover, because the increase in the radiative forcing accelerates within the twenty-first century (see Fig. 2.5), the temperature range between the warmest decade (generally the last one) and the coolest decade (generally the first one) strongly increases in all models with respect to the range in the historical period.

Therefore, even though the shift of decadal NMedLand summer precipitation means towards drier values is robust across models, the magnitude of the shift and the range between the driest and the wettest decades varies from one model to another. This result is consistent with the inter-model spread observed in the magnitude of the 2007-2099 summer precipitation change (see Fig. 6.8).

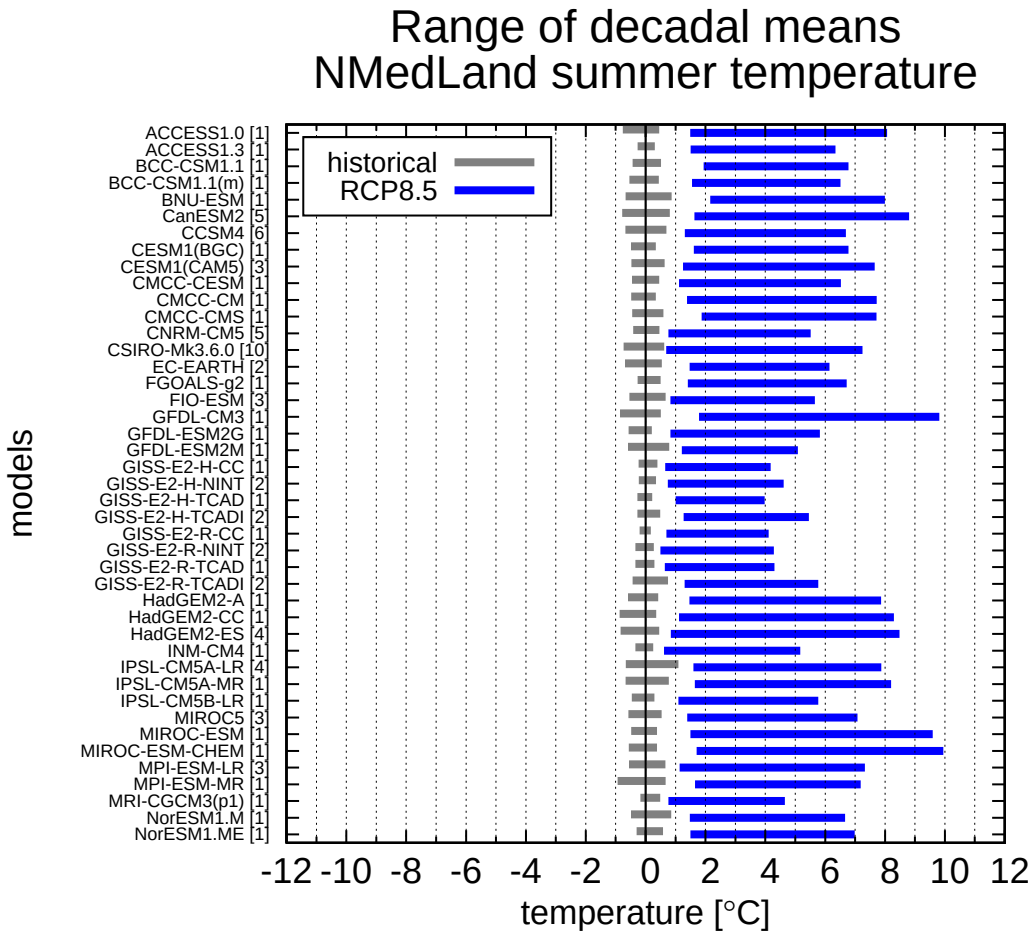


Figure 6.24: Range between the coolest and the warmest decades per model. Blue lines indicate RCP8.5 projections and grey lines indicate historical simulations. Decadal anomalies relative to the 1902-2004 mean. In brackets, after the name of each model, we indicate the number of simulations available.

6.5 Conclusions and discussion

That precipitation will generally decrease during the twenty-first century in the northern Mediterranean region is a robust result according to current climate models. The practical utility of this result for mitigation and adaptation policies, however, depends on several characteristics of this drying together with our degree of confidence in it. For this reason, in this chapter we have analysed in detail which aspects of the CMIP5 projections are robust in a multi-model sense and which aspects are realistic based on precipitation changes observed in the last decades.

First, according to our results, the future Mediterranean drying will be more intense and widespread in the warm half of the year than in the cold half. In summer, precipitation reductions are expected everywhere. There is less agreement, however, in the actual magnitude of such reductions, as area-averaged precipitation trends in the 2007-2099 interval

range from about −10%/century to about −80%/century. The lack of observed precipitation trends in summer months does not increase our confidence in such projections, but is consistent with the fact that the time of emergence of the summer drying signal takes place after 2015 in all the projections. In winter, precipitation decreases are projected over the central and southern part of the basin, whereas some northern regions are projected to experience weak precipitation increases. In addition, there is a large inter-model spread regarding the magnitude and the sign of the local and area-averaged trends.

Focusing on the summer drying, a strong decadal to multi-decadal variability is expected with a realistic order of magnitude, as the observational value falls within the inter-model spread. Our results indicate that northern Mediterranean precipitation will not decrease gradually from current values to the driest amounts of the end of the twenty-first century. In fact, models cover a wide spectrum of possible scenarios, ranging from a persistent drying signal superimposed on decadal fluctuations to a sudden precipitation decrease at some point of the twenty-first century.

Another result with good inter-model agreement is that the distribution of decadal anomalies of northern Mediterranean summer precipitation will shift towards drier conditions and become wider. Nevertheless, the extent of the shift and the amplitude of the widening is highly model-dependent, so we cannot provide a reliable description of how the distribution of decadal summer precipitation anomalies in the northern Mediterranean will look like in the future.

A following step in the analysis of future precipitation projections and their robustness involves a deep research of the drivers of the drying. Here we have examined the contribution of the NAO to the projected drying trends. In winter, we have found that projected NAO trends are weak and, therefore, the contribution of the NAO to the projected drying is minor. In summer, models miss the precipitation increases associated with the projected upward SNAO trend that could partially compensate for the drying due to other mechanisms. Isolating which other mechanisms independent of the NAO are related to northern Mediterranean precipitation reductions would allow us to assess whether these mechanisms (or their link to precipitation) are well captured by climate models. Possible misrepresentations of them could be the source of the large inter-model spread of the projected northern Mediterranean precipitation reductions.

Chapter

SEVEN

Conclusions and discussion

In this thesis we have discussed the role of climate change in precipitation trends in the northern Mediterranean region, motivated by the climate model projections of drying for the end of the twenty-first century [Christensen et al., 2013]. These projections have caught the attention of the climate community and the media, as they seem particularly consistent across models. In fact, the Mediterranean region has been described by Diffenbaugh and Giorgi (2012) as one of the world's climate change "hotspots" in which the response to global warming is expected to be strongest.

However, contrary to projections of warming, which are very reliable due to the fact that temperature increases have already been observed and also because the physical mechanisms linking the increase in the concentration of GHG and the change in temperature are well understood, assessing the realism of precipitation projections is problematic for several reasons. First, we do not have a precipitation record long enough to evaluate if current precipitation amounts are significantly different from those typical of pre-industrial times. This drawback, together with the strong multi-decadal variability of northern Mediterranean precipitation, makes it difficult to assess whether a drying signal is already detectable in observations. Moreover, even though many authors have looked for precipitation trends in the Mediterranean region, results are often contradictory or not reproducible due to the fact that trend estimations are highly sensitive to the choice of time interval, the months included in the analysis, the region of study and the source of observational data. Second, we lack a theoretical mechanism describing the response of precipitation to climate change beyond the thermodynamic argument that the difference between evaporation and precipitation will increase everywhere as a result of the increase in moisture flux divergence that follows the Clausius–Clapeyron equation. Given these problems with precipitation trends, another way of assessing whether projected precipitation changes are plausible is to examine trends in the NAO, as this circulation mode is the main driver of

precipitation variability in the northern Mediterranean. Even though we do not understand either how the NAO may respond to climate change, we have more confidence in SLP data for the first half of the twentieth century, which allows us to put the recent observed trends in a century-long context. Then, by analysing observed and simulated precipitation and NAO trends, the first question we have addressed is: Are projections of drying in the northern Mediterranean region realistic?

On the other hand, the inter-model agreement in projections of future northern Mediterranean drying has often been used to defend the reliability of the projections. But the only aspect of the drying whose consistency has been considered is the sign of the precipitation change at the end of the twenty-first century. We have explored in detail which other aspects of the projections are common across models in order to anticipate what possible scenarios we can expect for the next decades. Thus, the second question we have focused on is: Are the projections robust?

7.1 Summary of the results

In Chap. 4 we have assessed our confidence in projections by examining observational data looking for robust precipitation decreases consistent with the projected drying signal. Due to the strong multi-decadal variability of northern Mediterranean precipitation, however, the sign and the magnitude of the observed precipitation changes are very sensitive to the time period for which they are estimated. For this reason, we have employed a methodology based on two two-dimensional parameter diagrams which has allowed us to discern whether the change estimated for a given period can be tentatively attributed to internal variability. Following this method we have concluded that, contrary to projections, for which the strongest drying is found in summer, observational precipitation data exhibit no consistent signal of persistent drying in summer. Only in winter, the season in which the projected drying is weakest, has observed precipitation decreased significantly in the last decades. The evolution of winter precipitation in the second half of the twentieth century can be described as a persistent drying trend superimposed on internal multi-decadal variability. Since 2000 onwards, however, this signal has lost strength. We have also found that annual mean precipitation has also decreased in the second half of the twentieth century but to a lesser extent, as those reductions in winter have been partially compensated for by weak precipitation increases in autumn.

Because, based on observations only, it is not possible to do detection and attribution of climate signals, in Chap. 5 we have compared recent observed winter precipitation trends with those in forced and unforced CMIP5 model simulations. Little agreement has been found between observed and simulated trends. First, the multi-model mean does not reproduce the pattern of recent observed changes. In observations, the trends are negative across

the basin except for small regions in the eastern coast of Iberia, the south of the Italian Peninsula and the northern Balkans. Multi-model mean trends, instead, exhibit a meridional gradient with strong statistically significant drying trends in the south and significant wetting trends in the north. Second, multi-model mean trends for the recent decades are one order of magnitude smaller than the observed ones. Moreover, multi-model trends are negligible in comparison with inter-model variability and the simulated range of internal variability. The lack of a robust drying signal in model simulations does not support the hypothesis that recent observed winter precipitation trends have a forced origin. However, those trends are also inconsistent with internal variability, as climate models (run under either forced or unforced conditions) are not able to produce, for any time period of comparable length, winter precipitation trends as strong as those observed in the last decades.

The fact that the mismatch between the recent observed changes and the range of simulated changes is limited to those regions in which the precipitation influence of the NAO is strong has then led us to explore whether this impact is well simulated by models and whether simulations produce NAO trends as strong as the observed one. We have found that, not only is the northern Mediterranean precipitation response to NAO changes somewhat underestimated in climate models, but recent observed NAO trends are themselves stronger than any simulated trend for any period of comparable length. In addition, we have found that the strong negative NAO trends observed earlier in the twentieth century are also outside the range of simulated trends.

Assessing whether the projections of drying are reliable also requires examining the extent to which projections from different models agree with each other. In Chap. 6 we have shown that the inter-model agreement does not go much deeper than the sign and the seasonal distribution of the precipitation change (in all models the projected precipitation reduction is stronger in summer than in winter). Based on model projections we can expect precipitation decreases, but the magnitude can range from about 10%/century to about 80%/century (taking as a reference the twentieth century mean). Moreover, it is also not possible to obtain a robust estimate of when the forced signal will stand above internal variability, since the dispersion in the time of emergence of the forced signal covers the entire twenty-first century from 2015 onwards. Finally, we have also found that the pace of the projected precipitation change is very model dependent. While in some models the precipitation reduction takes place gradually over course of the twenty-first century, in other models the decrease occurs somewhat abruptly at some point of the future.

7.2 Discussion

With observations alone one cannot distinguish whether recent observed precipitation reductions are of anthropogenic or natural origin. Still, we can make a few comments. If

observed winter precipitation changes are indeed anthropogenic, the fact that the downward trend observed in the second half of the twentieth century has lost strength and even reversed in recent years means that the signal is still weak enough that it can be masked by internal variability. Alternatively, the precipitation decrease could be entirely natural. Even though the strong recent winter precipitation trends are unprecedented in the twentieth century, prior to 1950 we have only 50 years of sparse and poor quality data to compare with. Moreover, the winter precipitation reductions observed in the last decades are largely explainable by a positive trend in the NAO, which can be argued to have been natural. First, because this NAO trend has reversed in the first decade of the twenty-first century. Second, we lack a robust physical mechanism linking global warming and the NAO. Third, model projections do not exhibit a strong future NAO trend in response to increases in GHG concentrations. And finally, there are precedents of statistically significant negative NAO trends from the 1900s to the 1960-70s (Fig. 7.1a) and, compared to precipitation, we have fewer concerns about the quality of early twentieth century SLP data.

The comparison between observations and forced and unforced climate simulations could have allowed us to determine whether the recent observed winter precipitation trends are compatible to the response to external forcings or if they are consistent with internal variability. However, we have found that no forced or unforced climate model is able to produce, for any time period, precipitation trends as strong as those observed in the second half of the twentieth century. Two hypotheses could explain this mismatch: (1) the northern Mediterranean precipitation response to the forcing is too weak or misrepresented in models, and (2) the range of natural precipitation variability is underestimated in climate models.

The fact that the simulated winter precipitation response to NAO changes is too weak could explain the model underestimation of the recent winter precipitation trends (regardless of whether these are natural or forced). However, the recent observed NAO trend is itself outside the range of simulated trends (box in Fig. 7.1b and Fig. 7.1c). Therefore, we would be able to determine which of the two explanations for the lack of consistency between observations and simulated precipitation trends is more likely if we were able to gain some insight into whether the reason for the NAO trend mismatch is that (1) the NAO response to external forcings is unrealistic or that (2) the range of NAO internal variability is underestimated in models. A priori, both hypotheses seem equally possible, but the fact that the negative NAO trend observed between the 1900s and the 1960-70s is outside the range of simulated trends for any time period of comparable length (box in Fig. 7.1b and Fig. 7.1d) evidences that climate models do not represent well the range of internal variability of NAO changes. It is true that the inconsistencies are more pronounced for the recent period (in that they extend to a wider range of time intervals, see Fig. 7.1b), but the fact that none of the models simulates a range of trends wide enough to include the

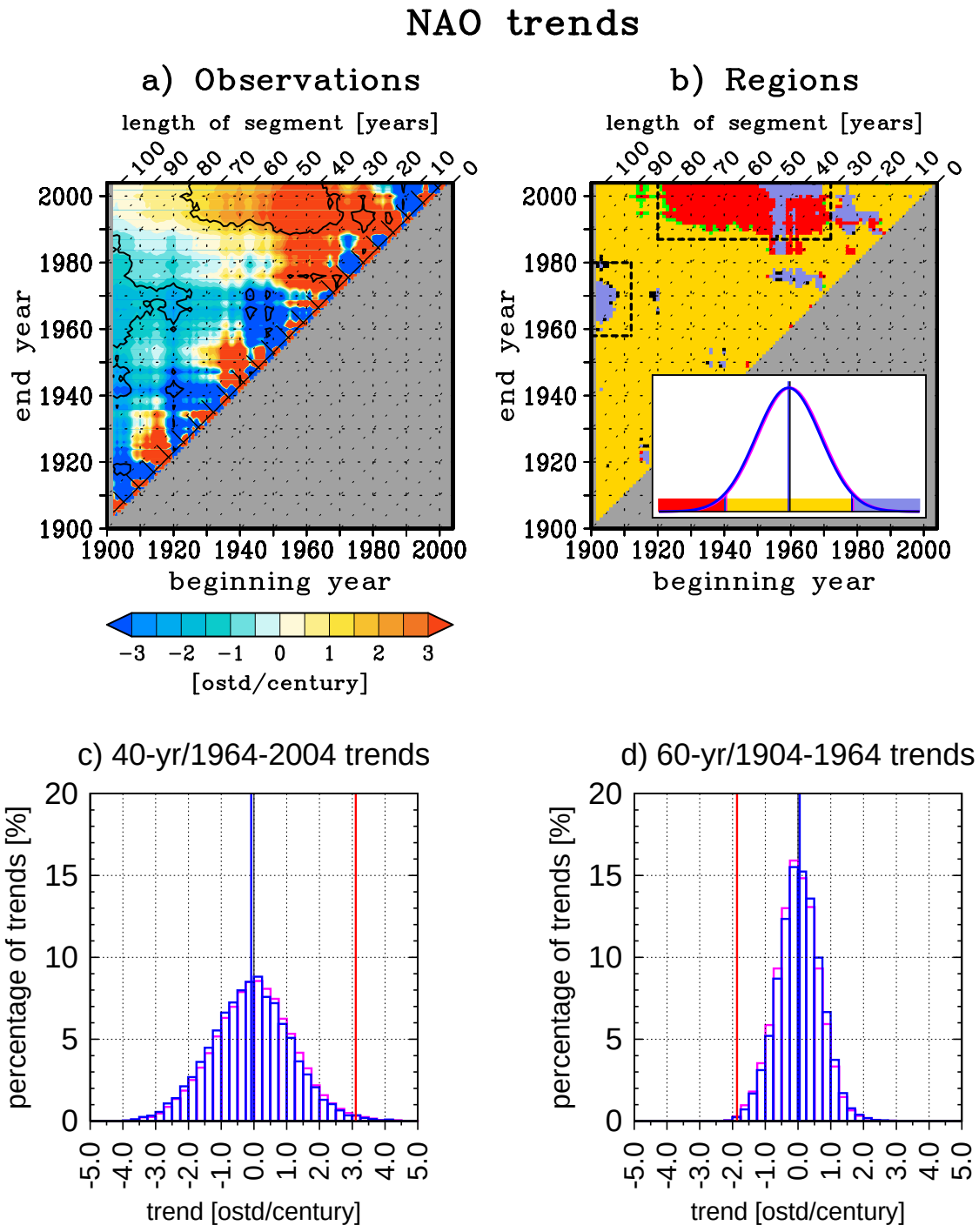


Figure 7.1: (a) Two-dimensional diagram in which we represent all observed NAO trends (Trenberth) within the 1902-2004 interval. Contours indicate statistically significant trends. (b) Two-dimensional diagram in which we compare, for each possible time interval within the 1902-2004 period, the observed trend with the natural distribution and forced distribution of CMIP5 models simulations. Colours correspond to the regions described in Sec. 5.1 (see insert). (c-d) Distribution of 40-year and 60-year trends only due to internal variability (natural distribution; magenta), and distribution of trends under the influence of the 1964-2004 and the 1904-1964 forced signal (forced distribution; blue). The blue vertical line indicates the model mean and the red vertical line the observed trend.

negative NAO trends suggests that the inconsistency for the *recent* period is at least *partly* due to unrealistically weak simulated NAO natural variability.

The most important novelty of our study is thus the suggestion that, if account is taken of model deficiencies in simulating the internal variability of NAO trends as well as the response of winter precipitation to NAO changes, recent precipitation trends could be within the range of internal variability. Moreover, this result is consistent with the fact that the time of emergence of the forced signal is found after 2015 in all CMIP5 simulations. Then, because the evolution of northern Mediterranean precipitation in recent decades seems to be still mainly modulated by natural variability, we conclude that observations cannot be used to assess whether projections of future drying are realistic.

7.3 Future work

In this thesis we have analysed climate change and precipitation trends in the northern Mediterranean from a large-scale perspective, using global climate models and focusing on general patterns. We have seen, however, that northern Mediterranean precipitation changes exhibit a complex spatial structure. For instance, the sign of the recent observational winter precipitation trends in the central and northern Balkan region has been found to be the opposite of that in the rest of the basin. For this reason, it would be interesting to expand our analysis to a more regional level, using local precipitation series and regional climate model simulations.

Furthermore, as pointed out by van Haren et al. (2015), the simulated magnitude of the projected summer drying in central Europe is sensitive to the model resolution, as a finer resolution is able to better capture the circulation anomalies that drive precipitation changes in that region. In this thesis we have used all models within the CMIP5 ensemble equally, without focusing on whether the discrepancies among them are related to their intrinsic characteristics such as their spacial resolution, how the components of each model are coupled, or which parametrizations they use. A more detailed analysis could give us insights towards identifying the sources of the inter-model discrepancy.

Another question to address in future work is whether the future evolution of precipitation will remain monotonic in the twenty-second century or if the sign of the trends will switch. Over some regions such as South America, a non-linear evolution of precipitation under climate change conditions in projections for the twenty-second century has been reported by Hawkins et al. (2014), with some models indicating that the thermodynamic precipitation increase ("wet gets wetter") will be replaced by precipitation reductions associated with a dynamical equatorward shift of the Intertropical Convergence Zone.

It should also be noted that precipitation trends alone will not define the future climate of the northern Mediterranean. The projected severe warming and increases in evapotran-

spiration will play a crucial role in modulating the impact of water deficits on the local environment [Pan, 2013; Feng and Fu, 2013; Cook et al., 2014; Mariotti et al., 2015]. In fact, evaporation increases may increment the presence of droughts even in regions where precipitation is projected to increase [Feng and Fu, 2013]. Therefore, a parallel study on evaporation trends in observations and models would complement our results.

Finally, we would like to encourage the modelling groups working in the CMIP6 project to provide a larger number of simulations for each model and for each climate experiment than that available in CMIP5. As we have pointed out in this thesis, internal variability plays a major role in the evolution of regional precipitation. With more simulations, the influence of internal variability in precipitation changes could be better estimated and we could also rigorously examine whether it has been altered as a consequence of external forcings.

References

Arguez, A. and Vose, R. S. (2011). The definition of the standard WMO climate normal: The key to deriving alternative climate normals. *Bulletin of the American Meteorological Society*, 92(6):699–704.

Barkhordarian, A., Storch, H., and Bhend, J. (2013). The expectation of future precipitation change over the Mediterranean region is different from what we observe. *Climate Dynamics*, 40(1-2):225–244.

Barnston, A. G. and Livezey, R. E. (1987). Classification, seasonality and persistence of low-frequency atmospheric circulation patterns. *Monthly Weather Review*, 115(6):1083–1126.

Becker, A., Finger, P., Meyer-Christoffer, A., Rudolf, B., Schamm, K., Schneider, U., and Ziese, M. (2013). A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901–present. *Earth System Science Data*, 5(1):71–99.

Bladé, I. and Castro-Díez, Y. (2010). Tendencias atmosféricas en la Península Ibérica durante el periodo instrumental en el contexto de la variabilidad natural. In *Clima en España: Pasado, Presente y Futuro*.

Bladé, I., Fortuny, D., van Oldenborgh, G. J., and Liebmann, B. (2012a). The summer North Atlantic Oscillation in CMIP3 models and related uncertainties in projected summer drying in Europe. *Journal of Geophysical Research*, 117(D16).

Bladé, I., Liebmann, B., Fortuny, D., and Oldenborgh, G. J. (2012b). Observed and simulated impacts of the summer NAO in Europe: Implications for projected drying in the Mediterranean region. *Climate Dynamics*, 39(3-4):709–727.

Bolle, H.-J., editor (2003). *Mediterranean climate: Variability and trends*. Regional climate studies. Springer, Berlin; New York.

Boucher, O., Randarll, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, O., Rasch, P., Satheesh, S., Sherwood, S., Stevens, B., and Zhang, X. (2013). Clouds and aerosols. In Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Cattiaux, J. and Cassou, C. (2013). Opposite CMIP3/CMIP5 trends in the wintertime Northern Annular Mode explained by combined local sea ice and remote tropical influences. *Geophysical Research Letters*, 40(14):3682–3687.

Cattiaux, J., Douville, H., and Peings, Y. (2013). European temperatures in CMIP5: Origins of present-day biases and future uncertainties. *Climate Dynamics*, 41(11-12):2889–2907.

Chen, M., Xie, P., Janowiak, J. E., and Arkin, P. A. (2002). Global land precipitation: A 50-yr monthly analysis based on gauge observations. *Journal of Hydrometeorology*, 3(3):249–266.

Christensen, J., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Held, I., Jones, R., Kolli, R., Kwon, W.-T., Laprise, R., Magaña Rueda, V., Mearns, L., Menéndez, C., Räisänen, J., Rinke, A., Sarr, A., and Whetton, P. (2007). Regional climate projections. In Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K., Tignor, M., and Miller, H., editors, *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Christensen, J., Krishna Kumar, K., Aldrian, E., An, S.-I., Cavalcanti, I., de Castro, M., Dong, W., Goswami, P., Hall, A., Kanyanga, J., Kitoh, A., Kossin, J., Lau, N.-C., Renwick, J., Stephenson, D., Xie, S.-P., and Zhou, T. (2013). Climate phenomena and their relevance for future regional climate change. In Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Collins, M., Knutti, R., Arblaster, J., Dufresne, J.-L., Fichefet, T., Friedlingstein, P., Gao, X., Gutowski, W., Johns, T., Krinner, G., Shongwe, M., Tebaldi, C., Weaver, A., and

Wehner, M. (2013). Long-term climate change: Projections, commitments and irreversibility. In Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Cook, B. I., Smerdon, J. E., Seager, R., and Coats, S. (2014). Global warming and 21st century drying. *Climate Dynamics*, 43(9-10):2607–2627.

Cubasch, U., Wuebbles, D., Chen, D., Facchini, M., Frame, D., Mahowald, N., and Winther, J.-G. (2013). Introduction. In Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Davini, P. and Cagnazzo, C. (2013). On the misinterpretation of the North Atlantic Oscillation in CMIP5 models. *Climate Dynamics*.

Deser, C., Knutti, R., Solomon, S., and Phillips, A. S. (2012a). Communication of the role of natural variability in future North American climate. *Nature Climate Change*, 2(11):775–779.

Deser, C., Phillips, A., Bourdette, V., and Teng, H. (2012b). Uncertainty in climate change projections: The role of internal variability. *Climate Dynamics*, 38(3-4):527–546.

Deser, C., Phillips, A. S., Alexander, M. A., and Smoliak, B. V. (2014). Projecting North American climate over the next 50 years: Uncertainty due to internal variability. *Journal of Climate*, 27(6):2271–2296.

Diffenbaugh, N. S. and Giorgi, F. (2012). Climate change hotspots in the CMIP5 global climate model ensemble. *Climatic Change*, 114(3-4):813–822.

Feng, S. and Fu, Q. (2013). Expansion of global drylands under a warming climate. *Atmospheric Chemistry and Physics*, 13(19):10081–10094.

Flato, G., Marotzke, J., Abiodun, B., Braconnot, P., Chou, S., Collins, W., Cox, P., Driouech, F., Emori, S., Eyring, V., Forest, C., Gleckler, P., Guilyardi, E., Jakob, C., Kattsov, V., Reason, C., and Rummukainen, M. (2013). Evaluation of climate models. In Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental*

Panel on Climate Change. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Folland, C. K., Knight, J., Linderholm, H. W., Fereday, D., Ineson, S., and Hurrell, J. W. (2009). The summer North Atlantic Oscillation: Past, present, and future. *Journal of Climate*, 22(5):1082–1103.

Fortuny, D. (2008). Tendències de precipitació a la península Ibèrica. Escalfament global o variabilitat natural?

Gibelin, A.-L. and Déqué, M. (2003). Anthropogenic climate change over the Mediterranean region simulated by a global variable resolution model. *Climate Dynamics*, 20(4):327–339.

Gillett, N. P. (2005). Detection of external influence on sea level pressure with a multi-model ensemble. *Geophysical Research Letters*, 32(19).

Giorgi, F. and Bi, X. (2009). Time of emergence (TOE) of GHG-forced precipitation change hot-spots. *Geophysical Research Letters*, 36(6).

Giorgi, F. and Lionello, P. (2008). Climate change projections for the Mediterranean region. *Global and Planetary Change*, 63(2-3):90–104.

Gámez, P. (2013). Procesamiento de datos CMIP5.

Gonzalez-Reviriego, N., Rodriguez-Puebla, C., and Rodriguez-Fonseca, B. (2012). Ensayos de simulación del índice NAO de invierno con datos de CMIP5. Salamanca. Asociación Española de Climatología.

Gonzalez-Reviriego, N., Rodriguez-Puebla, C., and Rodriguez-Fonseca, B. (2014). Evaluation of observed and simulated teleconnections over the Euro-Atlantic region on the basis of partial least squares regression. *Climate Dynamics*.

Greatbatch, R. J. and Rong, P.-p. (2006). Discrepancies between different northern hemisphere summer atmospheric data products. *Journal of Climate*, 19(7):1261–1273.

Haarsma, R. J., Selten, F., Hurk, B. v., Hazeleger, W., and Wang, X. (2009). Drier Mediterranean soils due to greenhouse warming bring easterly winds over summertime central Europe. *Geophysical Research Letters*, 36(4).

Hannachi, A., Jolliffe, I. T., and Stephenson, D. B. (2007). Empirical orthogonal functions and related techniques in atmospheric science: A review. *International Journal of Climatology*, 27(9):1119–1152.

Harris, I., Jones, P., Osborn, T., and Lister, D. (2014). Updated high-resolution grids of monthly climatic observations - the CRU TS3.10 dataset. *International Journal of Climatology*, 34(3):623–642.

Hartmann, D. L. (2014). "Objective analysis" course notes.

Hartmann, D. L., Klein Tank, A. M. G., Rusticucci, M., Alexander, L., Brönnimann, S., Charabi, Y., Dentener, F., Dlugokencky, E., Easterling, D., Kaplan, A., Soden, B. J., Thorne, P., Wild, M., and Zhai, P. (2013). Observations: Atmosphere and surface. In Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Hawkins, E., Joshi, M., and Frame, D. (2014). Wetter then drier in some tropical areas. *Nature Climate Change*, 4(8):646–647.

Hawkins, E. and Sutton, R. (2012). Time of emergence of climate signals. *Geophysical Research Letters*, 39(1).

Haylock, M. R., Hofstra, N., Klein Tank, A. M. G., Klok, E. J., Jones, P. D., and New, M. (2008). A European daily high-resolution gridded data set of surface temperature and precipitation for 1950–2006. *Journal of Geophysical Research*, 113(D20).

Held, I. M. and Soden, B. J. (2006). Robust responses of the hydrological cycle to global warming. *Journal of Climate*, 19(21):5686–5699.

Hoerling, M., Eischeid, J., Perlwitz, J., Quan, X., Zhang, T., and Pegion, P. (2012). On the increased frequency of Mediterranean drought. *Journal of Climate*, 25(6):2146–2161.

Hurrell, J. W. (1995). Decadal trends in the North Atlantic Oscillation: Regional temperatures and precipitation. *Science*, 269(5224):676–679.

Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M. (2003). An overview of the North Atlantic Oscillation. In Hurrell, J. W., Kushnir, Y., Ottersen, G., and Visbeck, M., editors, *Geophysical Monograph Series*, volume 134, pages 1–35. American Geophysical Union, Washington, D. C.

Imbers, J., Lopez, A., Huntingford, C., and Allen, M. (2014). Sensitivity of climate change detection and attribution to the characterization of internal climate variability. *Journal of Climate*, 27(10):3477–3491.

IPCC (2013). Summary for policymakers. In *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jones, P. D., Jonsson, T., and Wheeler, D. (1997). Extension to the North Atlantic oscillation using early instrumental pressure observations from Gibraltar and south-west Iceland. *International Journal of Climatology*, 17(13):1433–1450.

Kelley, C., Ting, M., Seager, R., and Kushnir, Y. (2012a). Mediterranean precipitation climatology, seasonal cycle, and trend as simulated by CMIP5. *Geophysical Research Letters*, 39(21).

Kelley, C., Ting, M., Seager, R., and Kushnir, Y. (2012b). The relative contributions of radiative forcing and internal climate variability to the late 20th century winter drying of the Mediterranean region. *Climate Dynamics*, 38(9-10):2001–2015.

Knutson, T. R., Zeng, F., and Wittenberg, A. T. (2013). Multimodel assessment of regional surface temperature trends: CMIP3 and CMIP5 twentieth-century simulations. *Journal of Climate*, 26(22):8709–8743.

Köppen, W. (1936). *Das geographische system der klimate*. Berlin.

Lane, D. M. (2013). *HyperStat online statistixs textbook*.

Lanzante, J. R. (1996). Resistant, robust and non-parametric techniques for the analysis of climate data: Theory and examples, including applications to historical radiosonde station data. *International Journal of Climatology*, 16(11):1197–1226.

Liebmann, B., Dole, R. M., Jones, C., Bladé, I., and Allured, D. (2010). Influence of choice of time period on global surface temperature trend estimates. *Bulletin of the American Meteorological Society*, 91(11):1485–1491.

Lionello, P., Malanotte-Rizzoli, P., and Boscolo, R., editors (2006). *Mediterranean climate variability*. Number 4 in Developments in Earth & environmental sciences. Elsevier, Amsterdam ; Oxford, 1st ed edition.

Livezey, R. E. and Chen, W. Y. (1983). Statistical field significance and its determination by Monte Carlo techniques. *Monthly Weather Review*, 111(1):46–59.

Lu, J., Vecchi, G. A., and Reichler, T. (2007). Expansion of the Hadley cell under global warming. *Geophysical Research Letters*, 34(6).

Maeland, E. (1988). On the comparison of interpolation methods. *IEEE Transactions on Medical Imaging*, 7(3):213–217.

Manabe, S., Spelman, M. J., and Stouffer, R. J. (1992). Transient responses of a coupled ocean-atmosphere model to gradual changes of atmospheric CO_2. Part II: Seasonal response. *Journal of Climate*, 5(2):105–126.

Mariotti, A. and Dell'Aquila, A. (2012). Decadal climate variability in the Mediterranean region: Roles of large-scale forcings and regional processes. *Climate Dynamics*, 38(5-6):1129–1145.

Mariotti, A., Pan, Y., Zeng, N., and Alessandri, A. (2015). Long-term climate change in the Mediterranean region in the midst of decadal variability. *Climate Dynamics*.

Mariotti, A., Zeng, N., Yoon, J.-H., Artale, V., Navarra, A., Alpert, P., and Li, L. Z. X. (2008). Mediterranean water cycle changes: transition to drier 21st century conditions in observations and CMIP3 simulations. *Environmental Research Letters*, 3(4):044001.

Matsuura, K. and Willmott, C. J. (2012). Terrestrial precipitation: 1900-2010 gridded monthly time series.

Meehl, G. A., Covey, C., Taylor, K. E., Delworth, T., Stouffer, R. J., Latif, M., McAvaney, B., and Mitchell, J. F. B. (2007). The WCRP CMIP3 multimodel dataset: A new era in climate change research. *Bulletin of the American Meteorological Society*, 88(9):1383–1394.

Miller, R. L., Schmidt, G. A., Nazarenko, L. S., Tausnev, N., Bauer, S. E., DelGenio, A. D., Kelley, M., Lo, K. K., Ruedy, R., Shindell, D. T., Aleinov, I., Bauer, M., Bleck, R., Canuto, V., Chen, Y., Cheng, Y., Clune, T. L., Faluvegi, G., Hansen, J. E., Healy, R. J., Kiang, N. Y., Koch, D., Lacis, A. A., LeGrande, A. N., Lerner, J., Menon, S., Oinas, V., Pérez García-Pando, C., Perlwitz, J. P., Puma, M. J., Rind, D., Romanou, A., Russell, G. L., Sato, M., Sun, S., Tsigaridis, K., Unger, N., Voulgarakis, A., Yao, M.-S., and Zhang, J. (2014). CMIP5 historical simulations (1850-2012) with GISS ModelE2. *Journal of Advances in Modeling Earth Systems*, 6(2):441–477.

Mitchell, T. D. and Jones, P. D. (2005). An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*, 25(6):693–712.

Morice, C. P., Kennedy, J. J., Rayner, N. A., and Jones, P. D. (2012). Quantifying uncertainties in global and regional temperature change using an ensemble of observational estimates: The HadCRUT4 data set. *Journal of Geophysical Research*, 117(D8).

Morita, T., Nakićenović, N., and Robinson, J. (2000). Overview of mitigation scenarios for global climate stabilization based on new IPCC emission scenarios (SRES). *Environmental Economics and Policy Studies*, 3(2):65–88.

Myhre, G., Shindell, D., Bréon, F.-M., Collins, W., Fuglestvedt, J., Huang, J., Koch, D., Lamarque, J.-F., Lee, D., Mendoza, B., Nakajima, T., Robock, A., Stephens, G., Takemura, T., and Zhang, H. (2013). Anthropogenic and natural radiative forcing. In Stocker, T., Qin, D., Plattner, G.-K., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, V., and Midgley, P., editors, *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Osborn, T. (2004). Simulating the winter North Atlantic Oscillation: The roles of internal variability and greenhouse gas forcing. *Climate Dynamics*, 22(6-7).

Pal, J. S., Giorgi, F., and Bi, X. (2004). Consistency of recent European summer precipitation trends and extremes with future regional climate projections. *Geophysical Research Letters*, 31(13).

Pan, Y. (2013). Mediterranean hydrological change projections for the 21st century assessed from CMIP5 simulations.

Paredes, D., Trigo, R. M., Garcia-Herrera, R., and Trigo, I. F. (2006). Understanding precipitation changes in Iberia in early spring: Weather typing and storm-tracking approaches. *Journal of Hydrometeorology*, 7(1):101–113.

Philandras, C. M., Nastos, P. T., Kapsomenakis, J., Douvis, K. C., Tselioudis, G., and Zerefos, C. S. (2011). Long term precipitation trends and variability within the Mediterranean region. *Natural Hazards and Earth System Science*, 11(12):3235–3250.

Piervitali, E., Colacino, M., and Conte, M. (1998). Rainfall over the Central-Western Mediterranean basin in the period 1951-1995. Part I: Precipitation trends. *Nuovo Cimento della Società Italiana di Fisica. C*, 21(3):331–344. eng.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. (1992). *Numerical recipes in C: The art of scientific computing*. Cambridge University Press, Cambridge; New York, 2nd edition.

Rizou, D., Flocas, H. A., Athanasiadis, P., and Bartzokas, A. (2015). Relationship between the Indian summer monsoon and the large-scale circulation variability over the Mediterranean. *Atmospheric Research*, 152:159–169.

Rodwell, M. J. and Hoskins, B. J. (1996). Monsoons and the dynamics of deserts. *Quarterly Journal of the Royal Meteorological Society*, 122(534):1385–1404.

Rodwell, M. J. and Hoskins, B. J. (2001). Subtropical anticyclones and summer monsoons. *Journal of Climate*, 14(15):3192–3211.

Rowell, D. P. and Jones, R. G. (2006). Causes and uncertainty of future summer drying over Europe. *Climate Dynamics*, 27(2-3):281–299.

Santer, B. D., Wigley, T. M. L., Boyle, J. S., Gaffen, D. J., Hnilo, J. J., Nychka, D., Parker, D. E., and Taylor, K. E. (2000). Statistical significance of trends and trend differences in layer-average atmospheric temperature time series. *Journal of Geophysical Research*, 105(D6):7337.

Scheff, J. and Frierson, D. M. W. (2012). Robust future precipitation declines in CMIP5 largely reflect the poleward expansion of model subtropical dry zones. *Geophysical Research Letters*, 39(18).

Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Ziese, M., and Rudolf, B. (2014). GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical and Applied Climatology*, 115(1-2):15–40.

Schneider, U., Becker, A., Meyer-Christoffer, A., Ziese, M., and Rudolf, B. (2011). Global precipitation analysis products of the GPCC.

Sen Gupta, A., Jourdain, N. C., Brown, J. N., and Monselesan, D. (2013). Climate drift in the CMIP5 models. *Journal of Climate*, 26(21):8597–8615.

Simpson, I. R., Seager, R., Shaw, T. A., and Ting, M. (2015). Mediterranean summer climate and the importance of Middle East topography. *Journal of Climate*, 28(5):1977–1996.

Sterl, A., Oldenborgh, G. J., Hazeleger, W., and Burgers, G. (2007). On the robustness of ENSO teleconnections. *Climate Dynamics*, 29(5):469–485.

Taylor, K. E., Balaji, V., Hankin, S., Juckes, M., Lawrence, B., and Pascoe, S. (2012a). CMIP5 data reference syntax (DRS) and controlled vocabularies.

Taylor, K. E., Stouffer, R. J., and Meehl, G. A. (2012b). An overview of CMIP5 and the experiment design. *Bulletin of the American Meteorological Society*, 93(4):485–498.

Trenberth, K. E. and Paolino, D. A. (1980). The Northern Hemisphere sea-level pressure data set: Trends, errors and discontinuities. *Monthly Weather Review*, 108(7):855–872.

van Haren, R., Haarsma, R. J., Vries, H. d., Oldenborgh, G. J. v., and Hazeleger, W. (2015). Resolution dependence of circulation forced future central European summer drying. *Environmental Research Letters*, 10(5):055002.

van Haren, R., Oldenborgh, G. J., Lenderink, G., Collins, M., and Hazeleger, W. (2012). SST and circulation trend biases cause an underestimation of European precipitation trends. *Climate Dynamics*, 40(1-2):1–20.

von Storch, H. and Zwiers, F. W. (1999). *Statistical analysis in climate research*. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Wilks, D. S. (2006). *Statistical methods in the atmospheric sciences*, volume 91 of *International Geophysics Series*. Academic Press, 2 edition.

Ziv, B., Saaroni, H., and Alpert, P. (2004). The factors governing the summer regime of the eastern Mediterranean. *International Journal of Climatology*, 24(14):1859–1871.

Acronyms and abbreviations

20C3M Climate of the twentieth century (see Sec. 2.2)

2DMDA Two-dimensional decadal to multi-decadal anomaly diagram (see Sec. 4.1.2)

2DLC Two-dimensional linear change diagram (see Sec. 4.1.1)

AR5 Fifth assessment report of the IPCC

CMIP3 Phase 3 of the Coupled Model Intercomparison Project

CMIP5 Phase 5 of the Coupled Model Intercomparison Project

DOF Degrees of freedom

EuroMed Euro-Mediterranean region (10°W-40°E/30°N-75°N)

GHG Greenhouse-gas

IPCC Intergovernmental Panel on Climate Change

multi3 CMIP3 multi-model ensemble mean

multi5 CMIP5 multi-model ensemble mean

NAO North Atlantic Oscillation

NMed Northern Mediterranean region (10°W-40°E/36°N-45°N)

NMedLand Land fraction of the NMed region

ostd Observed standard deviation (units of the NAO index)

pr-nao Component of precipitation series linearly explained by the NAO (see Sec. 4.5.1)

pr-nonao Component of precipitation series not related to the NAO (see Sec. 4.5.1)

RCP4.5 Representative Concentration Pathway with ~4.5 W m^2 at 2100 (see Sec. 2.2)

RCP8.5 Representative Concentration Pathway with ~8.5 W m^2 at 2100 (see Sec. 2.2)

SLP Sea level pressure

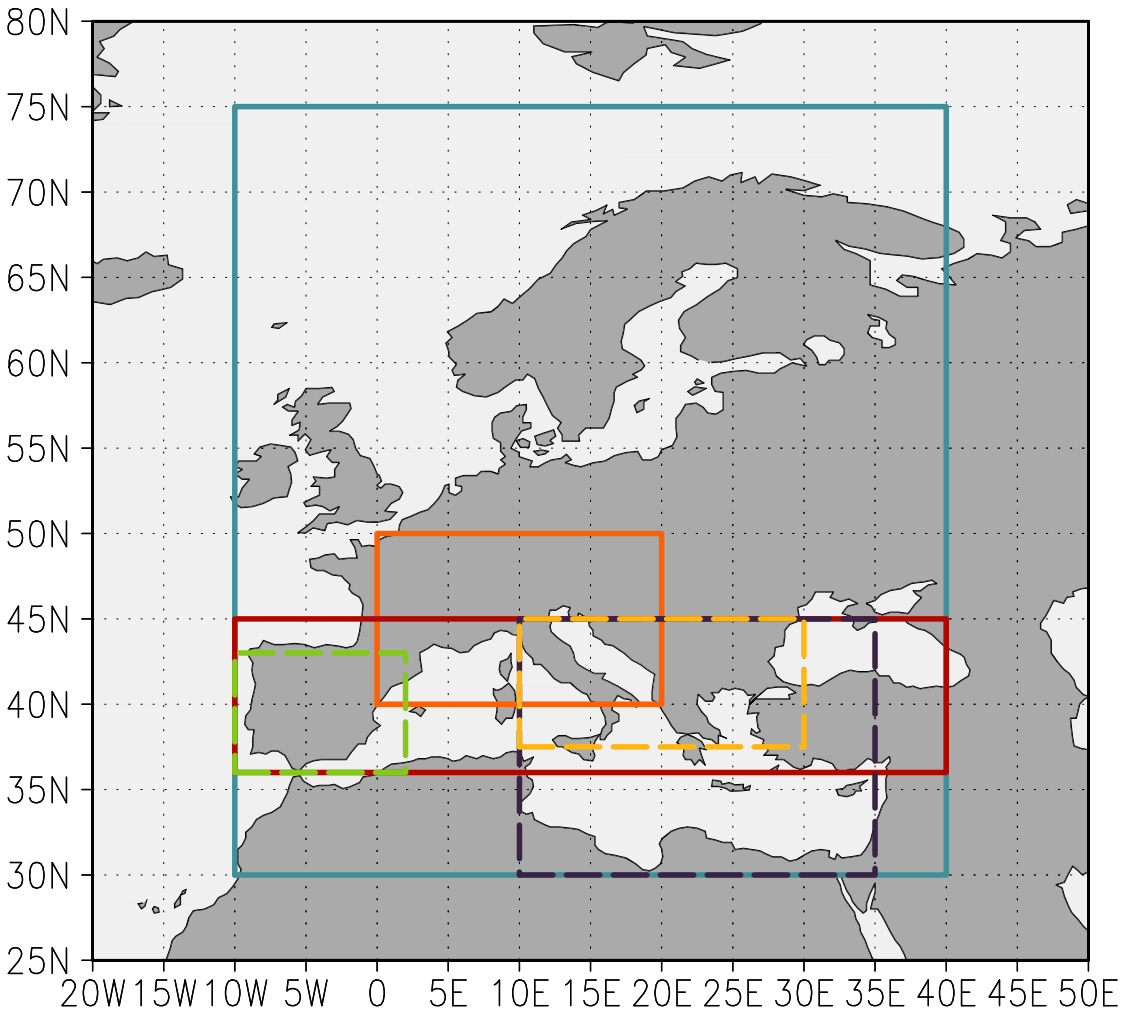
SNAO Summer North Atlantic Oscillation

SRESA1B Scenario A1B of the Special Report on Emission Scenarios (see Sec. 2.2)

std Standard deviation

TOE Time of emergence (see Sec. 6.3.3)

Regions



- Northern Mediterranean (*NMed*) region (10°W-4°W/36°N-45°N)
- Central-western part of the Iberian Peninsula from Sec. 4.4.2 (10°W-2°W/36°N-43°N)
- Euro-Mediterranean (*EuroMed*) region from Sec. 5.1 (10°W-40°E/30°N-75°N)
- Italian/Balkan region from Sec. 5.5.1 (10°E-30°E/37.5°N-45°N)
- Eastern Mediterranean region from Sec. 6.2.1.1 (10°E-35°E/30°N-45°N)
- Central Europe from Sec. 6.2.1.1 (0°-20°E/40°N-50°N)