UNIVERSITAT DE BARCELONA

Departament d'Estadística

Nonlinear Mixed-effects Models and Nonparametric Inference.

A Method Based on Bootstrap for the Analysis of Non-normal Repeated Measures Data in Biostatistical Practice.

Rachid El Halimi

Programa de Doctorat

Probabilitats i Estadística

Bienni 1997-99

Memòria presentada per a optar al grau de

Doctor en Matemàtiques

Certifico que la present memòria

ha estat realitzada per

Rachid el Halimi,

al Departament d'Estadística

de la Universitat de Barcelona

sota la meva direcció

Prof. Jordi Ocaña Rebull

Catedràtic del Departament d'Estadística

Universitat de Barcelona

Barcelona, a 1 de juny de 2005

Acknowledgements

This work is the result of many years of effort and would not have been possible without the assistance of many people who gave their support in different ways. To them, I would like to convey my gratitude and sincere appreciation.

I have been indebted with my supervisor, **Dr. JORDI OCAÑA REBULL**, whose guidance throughout the various revisions of this thesis, his previous instruction, suggestions, encouragement, constant availability, incredible patience and support in an unusual extent, as well as his academic experience, have been invaluable to me. He will always be remembered as the key factor that geared my career towards this path. It has been an honor and a pleasure working under his guidance.

My sincere appreciation to **Dra. MARIA DEL CARME RUIZ DE VILLA** and **Dr. CARLES M. CUADRAS AVELLANA** from the *Department of Statistics, University of Barcelona*, for their recommendations and generous cooperation to obtain some real data sets. They contributed to greatly improve this work. I would like also to extend my sincere appreciation to all professors and students of the *Department of Statistics, Faculty of Biology, University of Barcelona*, for their support during the realization of this project. It is a pleasure to acknowledge their extraordinary talents.

Special thanks to **Dr. EDUARD ESCRICH**, *Department of Cellular Biology, Physiology and Immunology, Faculty of Medicine, Universitat Autónoma de Barcelona*, **Dra. GENOVEVA MONTSERRAT**, *Department of Productes Naturals, Faculty of Pharmacy, University of Barcelona*, **Dra. M.CRISTINA POBLET**, *Department of Economic Theory, University of Barcelona*, and to **MANUEL VARAS**, PhD student and friend, from the *Department of Ecology, Faculty of Biology, University of Barcelona*. Their permission to use their data sets, and their reviews and helpful comments were decisive in the realization of this work.

My colleagues from the University of Pompeu Fabra deserve special thanks for giving me access to their computers. Thanks to their collaboration, many computer-intensive simulations presented in this thesis were possible.

This thesis was possible thanks to the financial support of project FIS 00/1130 by the Instituto de Salud Carlos III and 2001/SGR00067 from the Generalitat de Catalunya,.

CONTENTS

CONTENTS

0. RESUMEN EN CASTELLANO

La memoria trata sobre el análisis de modelos lineales y/o no lineales con efectos mixtos, con matrices estructuradas de varianzas-covarianzas de los efectos aleatorios y/o de los residuos.

La modelación de datos experimentales desde el marco teórico de los modelos mixtos brinda la posibilidad de analizar datos con estructuras de dependencia, no balanceados y en ocasiones con falta de normalidad. Estos permiten contemplar la falta de cumplimiento de los supuestos tradicionales y modelar, de manera flexible, complicadas estructuras de datos. Esta memoria tiene varios objetivos. Entre ellos se destaca: a) Presentar un "taller" de análisis avanzado de datos en el contexto de los modelos mixtos, que permite incrementar la precisión de las estimaciones. b) Favorecer la conceptualización de la modelación estadística en el contexto teórico-práctico de los modelos mixtos, exponiendo diversos tipos de modelos y considerando las implicaciones prácticas de su uso. c) Poner de manifiesto ciertas preocupaciones por la sensibilidad de las inferencias respecto de las suposiciones del modelo, especialmente cuando no cumplen las hipótesis habituales sobre normalidad de residuos y de factores aleatorios. El propósito principal del trabajo es estudiar la validez del empleo de modelos mixtos no lineales para analizar datos de medidas repetidas y discutir la robustez del enfoque inferencial paramétrico, y proponer y evaluar posibles alternativas al mismo, basadas en la metodología bootstrap. Se discute además el mejor procedimiento para generar las muestras bootstrap a partir de datos longitudinales bajo modelos mixtos, y se realiza una adaptación de la metodología bootstrap a métodos de ajuste en dos etapas, como STS (Standard two-stage) y GTS (Global two-stage).

Naturalmente, como paso previo se precede a un análisis exhaustivo de la situación actual en cuanto se refiere a la modelización y métodos de estimación utilizables.

La memoria presentada está estructurada en siete capítulos que se agrupan en tres partes principales, correspondientes a otros tantos objetivos. La primera de ellas se dedica al análisis de varios conjuntos de datos, en la forma que se indicó anteriormente. En ella se pretende poner de manifiesto las ventajas e inconvenientes de este enfoque de análisis, e ilustrar que, en ocasiones,

parece que no se cumplen los supuestos habituales de validez de la inferencia paramétrica. La secunda parte se ocupa del estudio de robustez del enfoque paramétrico basado en la aproximación propuesta por Lindstrom y Bates (1990). También se aborda un estudio similar para los métodos en dos etapas. En esta segunda parte, se estudia el comportamiento de los métodos estadísticos citados, principalmente, mediante simulación. La tercera parte aborda el tema de la metodología bootstrap y plantea un posible enfoque inferencial alternativo.

0.1. Análisis de datos reales usando modelos mixtos

El capitulo 1 se dedica a presentar diferentes conjuntos de datos reales. De cada conjunto se hace una breve descripción y una representación gráfica que ayuda a hacerse una idea de su estructura. A modo de preparación de análisis posteriores, se procura resaltar los siguientes aspectos de los datos: identificación de la tendencia general de las observaciones para cada individuo o unidad experimental a lo largo del tiempo, detección de posibles dinámicas no lineales en el tiempo, variación de la variabilidad dentro de cada individuo, representando diferentes criterios de agrupación dentro de un mismo gráfico, para tratar de explicar dicha variabilidad. Finalmente se formula un modelo lineal o no lineal (según el caso) con efectos mixtos que parece representar adecuadamente el conjunto de datos.

Estudio edafológico de la influencia de contaminantes

Entrando ya más en el detalle de cada conjunto de datos, el primero de ellos corresponde a la tesis de la Dra. Genoveva Montserrat ("*Efectes de l´aportació de residus al sól sobre la mineralització del carboni i del nitrogen*", Departament de Productes Naturals, Biologia Vegetal i Edafologia, Facultat de Farmàcia, Universitat de Barcelona, 2000). Se trata de un estudio de curvas de dosis-respuesta en los que se evalúa la respuesta de cada sujeto (en este caso, una muestra de suelo) bajo diferentes dosis de residuos industriales. Las variables de respuesta observadas fueron: Nitrificación, Amonificación global, Amonificación residual, Carbono diario (CO_2) y Carbono Acumulado (CO_2).

Descripción

Los experimentos consistieron en la realización de incubaciones de dos muestras de suelo con cuatro residuos diferentes: orgánico (LS), cerámico (CR), galvánico (LG) y una ceniza volante (CC). Las mezclas se efectuaron a cuatro dosis diferentes de suelo más residuo:

- ◊ Dosis de 50 $mg.ha^{-1}$
- ◊ Dosis de 100 $mg.ha^{-1}$
- ◊ Dosis de 250 $mg.ha^{-1}$
- ◊ Dosis de 500 $mg.ha^{-1}$

En primer lugar, se realizó un seguimiento de la transformación de N-orgánico a N-amoniacal a lo largo de 62 días, determinando la cantidad de N-amoniacal residual generada por la microflora. Además se realizó la determinación de las formas de N-nítrico generadas por la microflora nitrificante a partir del N-amoniacal.

Las gráficas de la evolución del N-amoniacal total generado en cada periodo de tiempo, que corresponde a la suma del N-amoniacal residual más N-nítrico:

$$\text{N-organico} \xrightarrow{\text{Amonificadores}} \text{N - NH}_4$$
$$\underbrace{\text{N - NH}_4 \longrightarrow \text{N - NO}_2^- \rightarrow \text{N - NO}_3^-}_{\text{Nitrificadores}}$$

se muestran en el capitulo 1 (Figura 1.1-1.3). También, se efectuó el seguimiento del CO_2 acumulado desprendido por las muestras a lo largo de 100 días de incubación. Asimismo, se hizo el seguimiento del CO_2 desprendido diariamente por las mezclas (véase Figura 1.4-1.5).

Objetivos

Con la determinación del CO_2 emitido por las mezclas y la de las diferentes formas de N, se pretende analizar el efecto de estos 4 residuos en dos suelos de naturaleza diferente. Interesa especialmente caracterizar la evolución de cuatro variables independientes (dosis-respuesta) denominadas en el capitulo 1 de la presente memoria como: "*nitrification*", "*global ammonification*", "*residual ammonification*", "*daily carbon*" y "*accumulated carbon*", en función de tiempo.

Modelos

Los datos referidos a la respuesta "*nitrificación*", visualizados en la Figura 1.1 en el capitulo 1, usando S-Plus, sugieren que el modelo lineal (lineal en los parámetros y no lineal en el tiempo; esto es, la *nitrificación* no es proporcional al tiempo) es adecuado para describir la evolución de la variable independiente en función del tiempo, pero los parámetros varían con los sujetos. El análisis exploratorio preliminar conduce al siguiente modelo lineal con efectos mixtos:

Modelo 1 [Nitrification data]: $$y_{ij} = (\delta_0 + \eta_{i0}) + (\delta_1 + \eta_{i1})\, t_{ij} + (\delta_2 + \eta_{i2}) \log(t_{ij}) + e_{ij}$$

donde el subíndice i corresponde a la curva, j corresponde a cada observación para una misma curva, y es la variable dependiente, e recoge la parte de variabilidad individual de las observaciones que aún queda sin explicar con nuestro modelo (error o residuo). Se estudian m curvas, es decir, se tienen m ecuaciones, y, por lo tanto, m valores para los coeficientes δ_{i0}, δ_{i1} y δ_{i2}. Lo valores de esos coeficientes se consideran como variables aleatorias de acuerdo con las ecuaciones $\delta_{i0}=\delta+\eta_{i0}$, $\delta_{i1}=\delta+\eta_{i1}$ y $\delta_{i2}=\delta+\eta_{i2}$, donde δ es el valor medio que representa un efecto sistemático (fijo) y $\eta_i=(\eta_{i0}, \eta_{i1}, \eta_{i2})'$ representa un efecto aleatorio asociado a la curva i. El vector η_i no incluye los errores del ajuste, que son los residuos e_{ij} del modelo, sino que pretende reflejar el hecho de que las curvas pueden estar afectadas por factores aleatorios no considerados en el modelo. Esto es distinto a que un dato concreto tomado en un momento dado se aleje de la curva debido a factores aleatorios relacionados con el muestreo. En principio, se considera que los vectores de efectos aleatorios, η_i, se distribuyen de una manera independiente entre ellos según una distribución normal multivariante de media cero y matriz de covarianzas común D, y que los errores aleatorios se distribuyen según una normal con media cero y varianza constante σ^2. Pero en el análisis inicial mostrado en el capitulo 3, se observa que la varianza de los residuos aumenta con la media, y, ésta última, lo hace con arreglo a la curva de crecimiento. De hecho, el Modelo 1 es muy flexible y da lugar a un modelo alternativo donde la varianza estará determinada por los parámetros de la curva como indica el siguiente modelo $\mathrm{var}(e_{ij}) = \sigma^2 \mu\left(t_{ij}, \delta\right)^{\tau}$. También se detecta una relación entre los residuos correspondientes a

medidas consecutivas dentro de un mismo individuo, por lo que se modifica el modelo incluyendo covarianzas entre dichos residuos. Cómo hacerlo es un tanto arbitrario. Es frecuente suponer (y en el presente caso parece funcionar bien dicha suposición) que la correlación entre dos residuos sucesivos es ρ, que la correlación entre dos residuos separados por un tercero es ρ^2, que la correlación entre dos residuos separados por dos consecutivos es ρ^3, etc. De manera que se propone un modelo alternativo donde los errores están modelados de la siguiente forma: $\text{cor}(e_{ij}, e_{ik}) = \rho^{|j-k|}$; donde $|\rho| < 1$ y $j \neq k$.

Respecto de la respuesta "*Amonificación global*", visualizada en la Figura 1.2 en el capitulo 1, se observan una serie de características que dan lugar al siguiente modelo lineal mixto:

Modelo 2 [Global amonification data]: $y_{ij} = (\delta_0 + \eta_{i0}) + \delta_1 t_{ij} + e_{ij}$.

Tenemos pues un modelo que estima la "*Amonificación global*" con un solo factor aleatorio: asociado a la ordenada δ_{i0} y un factor fijo: la tendencia lineal a lo largo del tiempo, δ_{i1}. De nuevo la matriz de varianzas-covarianzas de los errores se estructura de la misma manera que para el modelo anterior (Modelo 1), pero la correlación es mejor representarla mediante una función cuadrática racional ("rational quadratic model"), descrita en Pinheiro y Bates (2000).

Asimismo, la figura 1.3 del capitulo 1 representa las observaciones de la respuesta "*Amonificación residual*" en función de tiempo. Ello sugiere el siguiente modelo mixto:

Modelo 3 [Residual ammonification data]: $y_{ij} = \left(\delta_0 + \eta_{i0}\right) + (\delta_1 + \eta_{1i}) t_{ij} + e_{ij}$.

Como se observa en la figura 3.13 del capitulo 3, la varianza de los errores aumenta con el tiempo por cada sujeto, lo que da lugar a un modelo alternativo donde la varianza se modela como, $\text{var}(e_{ij}) = \sigma^2 t_{ij}^{2\tau}$, y la correlación entre observaciones del mismo sujeto tiene estructura de

correlación espacial esférica ("*Spherical spatial correlation structure*") de acuerdo con Pinheiro y Bates (2000).

En el mismo experimento, los datos relacionados con carbono diario mostrados en la figura 1.4 del capitulo 1, sugieren el siguiente modelo no lineal con efectos mixtos siguiente:

Modelo 4 [Daily carbon data]:

$$\begin{cases} y_{ij} = \delta_{1i}.\exp(-\delta_{2i}t_j) + \delta_{3i}.\exp(-\delta_{4i}t_j) + e_{ij} \\ \delta_{1i} = \delta_1 + \eta_{1i} \\ \delta_{2i} = \delta_2 + \eta_{2i} \\ \delta_{3i} = \delta_3 + \eta_{3i} \\ \delta_{4i} = \delta_4 + \eta_{4i} \end{cases}$$

Asimismo, de la figura 1.5 del capitulo 1, es razonable suponer el siguiente modelo para el carbono acumulado:

Modelo 5 [Accumulated carbon data]:

$$\begin{cases} y_{ij} = \delta_{1i}(1 - e^{-\delta_{2i}t_{ij}}) + e_{ij}, \quad i = 1,2,...,n, \quad j = 1,2,....,n \\ \delta_{1i} = \delta + \eta_{1i}; \quad \delta_{2i} = \delta + \eta_{2i} \end{cases} .$$

Uno de los objetivos asociados al uso de estos modelos es evaluar si hay evidencia de variación diferencial a lo largo de tiempo entre las curvas-respuesta asociadas a las distintas dosis. Por ello se ha considerado también un modelo alternativo que toma en consideración el efecto del factor dosis sobre los parámetros del modelo.

Análisis

En el capítulo 3 se recogen, entre otros, los resultados obtenidos en lo que se refiere a datos de *nitrification, amonification, daily y accumulated carbon.* Para ello se ha verificado, en la primera parte, la eficacia del método mixto para estos casos. Se ha abordado, entre otros aspectos, una descripción del tratamiento estadístico arbitrado por la modelación de la heterocedasticidad de la varianza y/o correlaciones de los errores.

Los resultados son, en general, similares a los obtenidos por Montserrat (2000), con algunas evidentes discrepancias asociadas al método usado en la estimación, y a veces, relacionadas con el modelo elegido. Es el caso, por ejemplo, del modelo biexponencial usado en el análisis del carbono diario en vez del modelo mono-exponencial usado por Montserrat.

La validez de las conclusiones extraídas de un modelo estadístico depende, a su vez, de la validez del modelo. Aunque se asume que el modelo puede no ser una representación exacta de la población, sí se requiere que reproduzca sin distorsiones las características principales de la misma (Aitken et al., 1989). De esta forma, un examen pormenorizado de la correspondencia entre los datos y el modelo constituye una parte importante del modelado estadístico. La evaluación del modelo se centra, principalmente, en la evaluación de posibles errores de diferentes especificaciones. Así que, la inspección de diversas estructuras de los errores, especialmente la distribución de los errores del ajuste, ha puesto de manifiesto que la hipótesis de normalidad no se cumple en algunas ocasiones, lo cual sirve de base para análisis posteriores realizados en esta memoria.

Datos de Cáncer de Mama

Descripción

Se obtuvieron de un seguimiento a lo largo del tiempo de varias variables que caracterizan el crecimiento tumoral en ratas con tumores mamarios inducidos artificialmente. Estos resultados están integrados en una serie de estudios realizados por el equipo de Dr. Eduard Escrich durante más de 15 años, para determinar la dinámica del crecimiento del tumor de mama bajo una variedad de condiciones (Escrich, Solanas y Segura, 1994, Escrich y col., 1994). La figura 1.6 del capitulo 1 muestra datos de un experimento sobre carcinogénesis en el que sesenta ratas Sprague-Dawley hembras vírgenes de 22 días de edad, fueron puestas en jaulas y mantenidas en una habitación controlada medioambientalmente a 24 ± 1^0C, a un 50% de humedad, en ciclo de 12 horas de luz y 12 horas de oscuridad. Al llegar, las ratas fueron alimentadas *ad libitum* con una de dos diferentes dietas semi-sintéticas, 40 recibieron una dieta baja en grasas y 20 una dieta rica en grasas. A la edad de 53 días, todos los animales recibieron una dosis única de 5 mg. de carcinógeno (7, 12- dimetilbenzo(α) antraceno -DMBA-, Sigma) por rata administrado en aceite de maíz por medio de una sonda gástrica (Huggins, Grand and Brillantes, 1961, Escrich, 1987). Un día después de la administración del carcinógeno, veinte animales del grupo de la dieta baja en grasas fueron escogidos al azar y trasladados permanentemente a la dieta rica en

grasas, los otros animales continuaron con la dieta inicial hasta el final del estudio. En otras palabras, se formaron finalmente tres grupos de tratamiento de 20 ratas cada: siempre dieta baja en grasas (etiquetada como "Dieta 1" en los análisis), siempre dieta rica en grasas ("Dieta 2") y baja de grasas antes de la administración del carcinógeno ("Dieta 3"). Las ratas se examinaron y palparon una vez a la semana, para detectar y medir tumores de mama. Para cada nuevo tumor detectado, se registró la fecha de detección, su localización y una estimación de su volumen. Al final del estudio, 201 días después de la administración del carcinógeno, las ratas fueron decapitadas. En la necropsia, los tumores fueron extirpados rápidamente, medidos, aclarados en solución salina normal, y divididos para observación histopatológica. Solamente los adenocarcinomas mamarios confirmados fueron comunicados en los resultados. Los datos analizados en la presente memoria se refieren a los volúmenes totales de todos los tumores dentro de cada rata.

Objetivos

El objetivo final de este estudio es de comprender la relación entre la dinámica de crecimiento tumoral y diversos factores de riesgo, vinculados a la dieta.

Modelos

Consideraciones teóricas y empíricas condujeron a suponer que un modelo mixto adecuado para describir la evolución a lo largo del tiempo del "volumen tumoral total" es:

Modelo 6 [Breast cancer data]:

$$y_{ij} = \frac{\delta_{1i}.\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$$
$$= \frac{\delta_{1i}}{e^{-(t_{ij}-\delta_{2i})}+\exp\left(-\left[(\delta_{3i}-1)/\delta_{3i}\right](t_{ij}-\delta_{2i})\right)} + e_{ij}$$
$$\delta_{ki} = \delta_k + \eta_{ki},\ i = 1,\ldots,m \text{ and } k = 1,\ldots,3 \quad .$$

donde y_{ij} representa el volumen tumoral total para la rata i en el instante t_{ij} , δ_1, δ_2 y δ_3 representan parámetros fijos (vinculados, respectivamente, al volumen tumoral final, al momento de aparición de los tumores y a la velocidad de crecimiento tumoral) y η_{1i}, η_{2i} y η_{3i} los

correspondientes parámetros aleatorios, asociados a la variabilidad individual de cada rata. Se supone que los parámetros aleatorios y los residuos del modelo son independientes y normales.

Análisis

El capitulo 3 presenta un estudio exhaustivo realizado sobre los datos de un experimento en cáncer de mama dónde se planea el crecimiento del volumen del tumor para determinar la influencia de un factor externo, la dieta. Para estos datos, el Modelo 6 parece ser adecuado para investigar la dinámica de crecimiento tumoral. La tasa del crecimiento definida, en el modelo, por el parámetro δ_{3i}, permite controlar distintos modos del crecimiento tumoral. En algunos animales, el tumor tiende a tener una curva logística, es el caso $\delta_{3i} = 1$. En otros animales, el crecimiento parece tener una forma exponencial, este caso está adecuadamente descrito cuando $\delta_{3i} > 1$. En ocasiones los tumores entran en regresión, tendiendo a desaparecer. Este caso corresponde a $\delta_{3i} < 1$.

No se encontró ninguna relación entre los parámetros fijos y la dieta. Los análisis gráficos exploratorios hacen pensar en alguna relación entre la variación de los parámetros aleatorios, en el sentido de que las dietas hiperlipídicas tienen mayores variaciones asociadas. Pero esta relación no se ha visto confirmada por los procedimientos inferenciales basados en la comparación de modelos bajo un método de linealización condicional. La inspección de las suposiciones del modelo sugiere que, mientras que los efectos aleatorios parece que se distribuyen normalmente, ello claramente no sucede para los residuos.

Dado el gran número de observaciones por el sujeto, parece razonable emplear otras técnicas basadas en las estimaciones individuales. En esta categoría están los métodos bietápicos o en dos etapas ("two-stage") que se consideraron como una alternativa para reajustar el modelo y reanalizar los datos. Después de obtener las estimaciones de los efectos fijos y las predicciones de los aleatorios, se obtuvieron diferencias sorprendentes entre estas estimaciones y las correspondientes al procedimiento basado en la una aproximación lineal, comentado anteriormente. Esto crea incertidumbre respecto de la fiabilidad de los métodos de estimación, en cuanto a la magnitud de las estimaciones que cada uno de ellos producen, y por tanto refleja una dificultad fundamental al estimar los parámetros fijos y aleatorios cuando no se cumplen las

suposiciones de normalidad. En otras palabras, los parámetros estimados presentan valores distintos, dependiendo del método que se haya seleccionado, lo que crea incertidumbre en cuanto a la elección del método de estimación y sugiere que hay que actuar cautelosamente en la interpretación de los resultados, y que vale la pena investigar la robustez de la inferencia paramétrica normal en los modelos mixtos no lineales.

Estas discrepancias entre diferentes aproximaciones, y dado que el requisito de normalidad no siempre se cumple, son un argumento más para motivar la necesidad de un estudio de simulación para evaluar la robustez de los métodos de linealización y bietápicos frente a la no normalidad de los efectos aleatorios y/o de los residuos.

Datos FANGAR: 2 niveles

Descripción

Otra clase de estudios en los que nos encontramos con observaciones correlacionadas, no independientes, son los denominados datos agrupados (cluster data), en los que existe un diseño jerárquico. Esta estructura hace que este tipo de modelos se conozcan con el nombre de modelos multinivel (multilevel), siendo los más utilizados los de 2 niveles (en el caso de los estudios longitudinales, el nivel 2 lo constituyen los sujetos y el nivel 1 las observaciones sobre cada sujeto). Con el objetivo de ilustrar en qué se traduce este modelo a la hora de aplicarlo a datos reales, se consideran los datos mostrados en la figura 1.8 de capitulo 1 (pertenecientes al Departamento de Ecología de la Facultad de Biología de la Universidad de Barcelona, tesis doctoral titulada "*Ecología de la bahía del Fangar (delta del Ebro) y su relación con el cultivo de moluscos bivalvos",* en preparación, a cargo de Manuel Varas, que consiste en los valores numéricos de respuesta de salinidad por muestra de agua recogida en la bahía denominada El Fangar, situada en el delta del río Ebro. Las mediciones fueron hechas entre 1986 y 1987. Las muestras se recogieron en tres localizaciones (estaciones) etiquetadas "Fangar1", "Fangar2" y "Fangar3", y cada una tenía tres niveles de profundidad (1-primer nivel, 2-segundo nivel y 3-tercer nivel).

Objetivo

El interés principal es modelar la variación de la salinidad en función del tiempo.

Modelo

Consideraciones empíricas y teóricas condujeron a suponer que el siguiente modelo es adecuado para evaluar la dinámica de estos datos:

Modelo 7 [**salinidad**]:

$$y_{ijk} = \delta_0 + \delta_1 \cos(t_{ij}) + \delta_2 \sin(\delta_3 \cdot t_{ij}) + \eta_i + \eta_{ik} + e_{ijk}$$
$$\eta_i \sim N(0, \sigma_1^2),\ \eta_{ik} \sim N(0, \sigma_2^2),\ e_{ijk} \sim N(0, \sigma^2)$$

Donde y_{ijk} es la observación de salinidad j en la estación i bajo la profundidad k (k=1,2,3) en el tiempo (fecha) t_{ij}. δ_0, δ_1, δ_2, y δ_3 son los efectos fijos y η_i es el vector de los efectos aleatorios asociados a la estación i, η_{ik} denota el parámetro aleatorio asociado a la profundidad k en la estación i. Se supone que η_i, η_{ik} y e_{ik} son independientes por todas las i y k.

0.2. Métodos Establecidos

En el capítulo 2 se realiza una revisión exhaustiva del estado de la cuestión que nos ocupa. Partiendo de un análisis de los modelos lineales y no lineales mixtos, se recorren los diferentes métodos de estimación más utilizados en la literatura. En la parte 1/2 a 2/3 del capitulo se ha enfocado en el desarrollo de algunas técnicas de estimación de los efectos fijos y de los componentes aleatorios como la máxima verosimilitud (ML) y la máxima verosimilitud restringida (REML). Se refiere, con un poco de detalle, al procedimiento de máxima verosimilitud que tiene su origen con Fisher (1925) y fue aplicada por primera vez al modelo mixto general por Hartley y Rao (1967). Este método supone que los términos del error y los efectos aleatorios tienen distribución normal, de tal suerte que del modelo, $y=f(X,\delta,Z,\eta)+e$, donde y representa el vector de observaciones, f es una función conocida, X y Z son matrices de "diseño" asociadas a los efectos fijos, δ, y a los efectos aleatorios, η_i,.(Estos últimos a su vez, se consideran como N(0,D)), y e representa los errores aleatorios distribuidos como N(0, Λ); se obtiene y normal de media $E(y)=E(E(y|\eta))=\int f(X,\delta,Z,\eta)dF_\eta(\eta)$, donde F_η denotando la función de distribución de η, y de varianza $var(y)=E(var(y|\eta))+var(E(y|\eta))$. En el caso lineal donde $f(X,\delta,Z,\eta)=X\delta+Z\eta$, la verosimilitud L se expresa como: $L=(2\pi)^{-1/2n}|ZDZ'+\Lambda|^{-1/2}\exp(-1/2[(y-$

$X\delta)'(ZDZ'+\Lambda)^{-1}(y-X\delta)]$), donde n representa el tamaño muestral. Al maximizar L con respecto a δ y los componentes de varianza, se llega a las ecuaciones normales que tienen que ser resueltas para obtener los estimadores de máxima verosimilitud de δ y los elementos de D y de Λ. Las ecuaciones obtenidas se tienen que resolver para $\tilde{\delta}$ y $\tilde{\delta}^2$, siendo estos últimos los elementos implícitos en la matriz de varianzas-covarianzas total. No son lineales en los elementos de las varianzas-covarianzas, aunque una vez obtenidos se pueden usar para obtener $\tilde{\delta}$. La no-linealidad implica que, en general, no se puede resolver las ecuaciones analíticamente. Basado en maximizar la verosimilitud de la muestra, ofrece estimadores con propiedades más útiles, normalidad asintótica y matriz de dispersión muestral asintótica conocida. Sin embargo, existe una pérdida de grados de libertad, asociada a la estimación de los efectos fijos, que conduce a subestimar la varianza del error (Searle et al., 1992). Una variante de la estimación de máxima verosimilitud es la Máxima Verosimilitud Restringida o REML. Fue usada por Patterson y Thompson (1971) para diseños de bloques. REML maximiza la verosimilitud de un vector de combinaciones de $X\delta$. Supóngase que L es la matriz que produce al vector de combinaciones lineales. Entonces, Ly= L$X\delta$+L$Z\eta$+Le, es invariante a $X\delta$ si y sólo si LX=0. Pero LX=0 si y sólo si L=TM para M=I-X(X'X)$^-$X' y alguna T. Para producir combinaciones lineales que eliminan a los efectos fijos, L debe ser de rango completo. Harville (1977) llamó a los elementos Ly "contrastes de error" porque tienen un valor esperado cero, E(Ly)= L$E(X\delta+Z\eta,+e)$=L $X\delta$=0.

El modelo de especificación general, $y=f(X,\delta,Z,\eta)+e$, es una clase que incluya varios modelos considerados en la literatura como casos especiales. Por ejemplo, Vonesh y Carter (1992) y Gumpertz y Pantula (1992) han considerado casos en que los efectos aleatorios intervienen aditivamente en el modelo, $y=f(X,\delta)+Z\eta+e$. También cuando la función f es invertible, el modelo puede verse como un caso especial de GLMM (*genralized linear mixed model*) aplicado a los datos de distribución normal cuya función de enlace no es la identidad. Cuando la función $f(X,\delta,Z,\eta)$ es no lineal en los parámetros aleatorios (o en ambos parámetros fijos y aleatorios), resulta muy difícil calcular la verosimilitud directamente, y al contrario de en los modelos lineales mixtos, no hay ninguna solución firme al MLE incluso cuando los parámetros de

varianza σ y D sean conocidos. Sin embargo, se han planteado varias aproximaciones estadísticas en la literatura, que han sido aplicadas en análisis de los datos. Entre ellas, se destaca, la aproximación en la que Sheiner y Beal (1980) han sugerido que se reemplace el modelo $y=f(X,\delta,Z,\eta)+e$ por otro modelo que sea lineal en los parámetros aleatorios derivado como la aproximación de Taylor del primer-orden entorno de $\eta=0$. Entonces se considera el pseudo-modelo

$$y \approx f(X,\delta,Z,\eta)\,|_{\eta=0} + Z^{SB}(\delta,0)\eta + e\,;\ \text{donde}\ Z^{SB}(\delta,0) = \left(\left. \frac{\partial f}{\partial \eta'} \right|_{\eta=0} \right). \qquad (1)$$

Esta claro que el modelo (1) no coincide con el modelo original y, como consecuencia este estimador no es necesariamente válido para el modelo original. Lindstrom y Bates (1990) han sugerido la misma aproximación pero no en torno a $\eta=0$, como en Sheiner y Beal, sino en torno de $\tilde{\eta}$, donde $\tilde{\eta}$ es una "estimación" de η.

$$y \approx \left\{ f(X,\delta,Z,\eta)\,|_{\eta=\tilde{\eta}} - Z^{LB}(\delta,\tilde{\eta})\tilde{\eta} \right\} + Z^{LB}(\delta,\tilde{\eta})\eta + e;\ \text{donde}\ Z^{LB}(\delta,\tilde{\eta}) = \left. \frac{\partial f(X,\delta,Z,\eta)}{\partial \eta} \right|_{\tilde{\eta}}. \qquad (2)$$

Ambos procedimientos se resuelven numéricamente. Sin embargo, a pesar de la popularidad de los modelos (1) y (2), debe notarse que, en teoría, cuando se dispone de un número limitado de observaciones por individuo, estos métodos proporcionan estimaciones consistentes sólo cuando el número de individuos m tiende a infinito y la variabilidad de los efectos aleatorios tiende a cero, al mismo tiempo (Zhiyu Ge. et al, 2003).

Se continua con una exposición sobre la aplicación de los Métodos de dos-etapas (two-stage methods) y la aproximación de Laplace. Como complemento de lo anterior se efectúa una extensión al caso multinivel, y se expone criterios destinados a confirmar ciertos aspectos estadísticos del análisis y comparar diversos Modelos Mixtos.

Finalmente, se recorren los diferentes enfoques que históricamente se han dado a los aspectos teóricos hasta llegar a la situación actual, en donde la aplicación de modelos mixtos resulta relevante.

0.3. Robustez Del Enfoque Paramétrico

La inferencia paramétrica tradicional para los coeficientes en un modelo no lineal mixto, se basa en condiciones distribucionales y suposiciones que pueden, o no, ser apropiadas para un conjunto dado de datos. Sin embargo, si estas condiciones no se cumplen en un caso particular, la inferencia paramétrica puede ser inexacta. En el capitulo 4 se expone una serie de tests de robustez destinados a confirmar ciertos aspectos manifestados en el análisis antes expuesto. Por un lado, se examinó la sensibilidad del enfoque parámetrico, basado en la aproximación de Lindstrom y Bates (LB, 1990), a los supuestos en cuanto a la distribución de los efectos aleatorios y los errores. Para ello, se llevaron a cabo estudios de simulación donde se hicieron variar diferentes condiciones; algunas utilizadas como base de referencia y otras seleccionadas para ilustrar serios incumplimientos de las condiciones básicas en la aplicación del método. Las series de simulaciones trataban de emular los estudios experimentales descritos anteriormente. En ellas se utilizaron valores de los parámetros del orden de los valores observados experimentalmente. El modelo y la situación base que principalmente se simuló corresponde al estudio sobre el cáncer de mama, con simulaciones complementarias correspondientes a los otros modelos. En una primera fase se pretendió estudiar el efecto del alejamiento de las condiciones de normalidad, suponiendo siempre que el modelo básico (Modelo 6) era correcto. Por ello se generaron series de 1000 muestras (consistente en las observaciones repetidas durante 25 "semanas" del "volumen tumoral" de tres grupos experimentales de 20 ratas cada uno) de acuerdo con el modelo pero generando los parámetros aleatorios y los residuos según todas las posibles permutaciones de los cuatro siguientes distribuciones (distribución normal (N), distribución exponencial negativa (E), distribución uniforme (U) y distribución gamma (G)). Para cada una de estas distribuciones, e inspirado en los datos reales y también en ensayos piloto realizados con simulaciones previas, se emplearon los siguientes valores de los parámetros: $(\delta_1,\ \delta_2,\ \delta_3)'=(5.051236,\ 13.86703,\ 0.8486555)'$,

$$D = \begin{pmatrix} 51.88633 & -1.0498620 & -0.05460614 \\ & 15.8465000 & -0.04587260 \\ & & 0.01362791 \end{pmatrix}$$

y σ=0.939921 como parámetros "reales" en el diseño de simulación. Los efectos aleatorios simulados se generaron de acuerdo con una expresión equivalente a $\eta_i = Lh$, donde h simboliza una versión estandarizada del vector de los efectos aleatorios, generado a partir de una de las posibles distribuciones anteriores, reescalada convenientemente para que tuviese media cero y varianza 1, y L simboliza la parte inferior de la matriz triangular resultante de la descomposición de Cholesky de la matriz de varianzas-covarianzas D. Los residuos simulados para cada individuo se generaron de acuerdo con una expresión equivalente a $e_i = \sigma u_i$, donde u_i simboliza un vector estandarizado de residuos i.i.d, generados a partir de una posible distribución de media cero y de varianza 1. En todos los casos se ajustaron los parámetros para que los momentos de segundo orden fuesen los mismos de una distribución a otra (y comparables a los experimentales). A partir de cada serie de simulación, se estudió el sesgo y el error cuadrático medio de los estimadores de los parámetros fijos y de la varianza de los residuos, y de la matriz de covarianzas del vector de parámetros aleatorios, juntamente con el recubrimiento real de los intervalos de confianza de estos parámetros. Los estimadores considerados fueron los previstos en la función *nlme*, basados en máxima verosimilitud (ML) y en máxima verosimilitud restringida (REML), descritos en capitulo 2. Los programas estaban escritos en lenguaje S-Plus.

En las Tablas (4.1)-(4.3) del capitulo 4 aparecen recogidas las medidas correspondientes a los efectos fijos δ_1, δ_2 y δ_3, para cada una de las 16 condiciones manipuladas. Los resultados presentados ponen de manifiesto ciertos aspectos a resaltar. En primer lugar, la parte fija del modelo no lineal mixto ajustado con el procedimiento LB parece poco afectada (en términos del sesgo y del error cuadrático medio) por el incumplimiento de la normalidad de los efectos aleatorios y los residuos. Los parámetros estimados presentaron valores cercanos a los valores "reales" con un sesgo pequeño en la mayoría de los casos, pero significativo. En este sentido cabe recordar que, para el parámetro δ_1 (que caracteriza la media de la distribución de δ_{1i}), el mejor sesgo se obtuvo en el caso que la distribución de los efectos aleatorios fuese uniforme y la distribución de los errores fuese normal (condición UN), era aproximadamente de 0.01 y correspondió al estimador ML; mientras que el peor sesgo (en torno a 0.26) se obtuvo en el caso

de que la distribución de los efectos aleatorios fuese exponencial y la distribución de los errores normal o uniforme (EN o EU). Bajo las condiciones normales de los efectos aleatorios y los errores (NN), el sesgo era de 0.16. Los sesgos resultaron ser moderadamente más elevados en el caso del parámetro δ_2, la razón pudo deberse a la naturaleza del mismo. Para este parámetro, el mejor sesgo se obtuvo cuando los efectos aleatorios seguían una distribución uniforme y los errores seguían una distribución exponencial, y era de 0.05 aproximadamente, mientras que el peor sesgo se obtuvo bajo la condición EN (sesgo cercano a 0.9). En el caso de la condición normal, NN, el sesgo era de 0.18. La mayor "estabilidad" en los sesgos acontecía para el parámetro δ_3, con estrechas diferencias entre sesgos. Esta medida de precisión estadística era casi constante, con valor 0.02 cuando la distribución de los efectos aleatorios era normal y uniforme, y valor 0.03 (0.04) cuando la distribución de los efectos aleatorios era exponencial (gamma).

Desde esta óptica, y mientras que la estimación puntual de la parte fija del modelo se muestra relativamente robusta, los resultados presentados en las Tablas (4.4)-(4.10), ponen de relieve que éste no es el caso para los componentes aleatorios, con un elevado sesgo, tal como se apuntaba en otros trabajos, especialmente Hartford y Dividian (1999) para el caso no lineal.

A diferencia de lo que ha sucede con el sesgo y el error cuadrático medio, los intervalos de confianza asintóticos que calcula *nlme* de S-Plus, cuando se evalúan según los criterios de recubrimiento y de equilibrio (número de veces que el verdadero valor del parámetro queda fuera, a la izquierda del intervalo, comparado con el número de veces que el verdadero valor del parámetro queda a la derecha del intervalo) se revelan como muy poco robustos, con un recubrimiento real mucho menor que el nominal. De hecho, incluso bajo condiciones de normalidad (NN), estos intervalos funcionan bastante mal para los tamaños muéstrales (inter- e intra-individuos) simulados y correspondientes a los experimentos reales en los que están inspiradas las simulaciones. En concreto, tal y como se aprecia en las tablas indicadas anteriormente, los intervalos de confianza arrojan un nivel de recubrimiento diferente para cada parámetro. De hecho y dado que idealmente el recubrimiento debería ser exactamente el recubrimiento nominal del intervalo deseado (95% en este caso), la robustez es bastante aceptable para el parámetro fijo δ_1, donde el nivel de recubrimiento se mantuvo entre 88.8% y

95.6%. En cuanto a la condición NN, el nivel de recubrimiento se limita a 93.6%. Es importante resaltar que, para este parámetro, los peores niveles del recubrimiento se registraron bajo condiciones EN y GN. Otro aspecto importante a destacar es el buen nivel de equilibrio mostrado por los intervalos de confianza en la mayoría de los casos.

En cuanto al parámetro δ_2, los intervalos, por el contrario, dan peores resultados, peores incluso bajo condiciones normales (NN). De hecho el nivel de recubrimiento se mantuvo entre 89% (UE) a 50% (EN), con 86% para el caso NN. En correspondencia con los recubrimientos de los intervalos, la media de sus longitudes oscila entre 2.55 para UE y 2.03 para el caso EN. Los peores niveles de recubrimiento corresponden claramente a las distribuciones asimétricas (E y G) de los efectos aleatorios y están asociados con una tendencia fuerte de los intervalos a subestimar el verdadero valor del parámetro, que frecuentemente queda a la derecha (48.5% para el caso EN) del limite superior del intervalo.

Al igual que ocurría en el caso de δ_2, es importante resaltar que, para el parámetro δ_3, los niveles de recubrimiento son también muy inadecuados en todas las condiciones, produciendo niveles más bajos oscilando entre 65.3% (EN) y 77.5% (UN) con un 75.2% en el caso NN. Las medias de las longitudes de estos intervalos muestran una tendencia similar y, de nuevo, los peores niveles de recubrimiento están asociados con la asimetría (E y G) de las distribuciones para los efectos aleatorios, esta vez con una alta tendencia para los intervalos a sobrestimar el verdadero valor del parámetro. En cuanto al equilibrio, la proporción de rechazos por la izquierda es mayor en todos los casos. En el peor de ellos (caso EN), la proporción correspondiente está en 32.5%.

Por último, cabe notar que todos los intervalos de confianza para los parámetros fijos parecen ser afectados más por la distribución de los efectos aleatorios que por la distribución de los residuos, y además son muy sensibles a la asimetría y "skewness"de las distribuciones.

En secundo lugar, un examen más pormenorizado de los resultados (referidos a los componentes aleatorios) presentados en las Tablas (4.4)-(4.10) pone de manifiesto otra serie de aspectos que también requieren ser descritos. En términos generales, la inferencia en los componentes de la matriz D incluso es menos robusta que la inferencia en los efectos fijos, pero adecuada para el caso normal. Para todos los componentes de D, el nivel de recubrimiento en el caso de NN se

mantuvo siempre cerca del nivel nominal 95%, mientras que el nivel recubrimiento registrado bajo las condiciones que violan la normalidad es muy errático y afectado por ambas, la distribución de los efectos aleatorios y de los residuos. Esta falta de robustez es especialmente evidente para componentes de la varianza-covarianza asociados a η_{i2} y η_{i3}. De nuevo, el peor recubrimiento se ha registrado cuando la asimetría es más alta (en el caso de E y de G) para la distribución de los efectos aleatorios.

Como ya se esperaba, los resultados para el estimador de la desviación estándar de los residuos, σ, mostraron una dependencia clara en la distribución simulada de los residuos, en lugar de la de los efectos aleatorios. Por un lado, los resultados para el sesgo y MSE mostraron que es robusto, y por otro, el recubrimiento de los intervalos de confianza para este parámetro va de 99.2% en el caso de EU a 65.7% en el caso de UG, con 93.3% en el caso NN. Los peores valores del recubrimiento se han observado cuando la distribución de los residuos es E o G. Además, e independientemente del recubrimiento real, los intervalos de confianza se mostraron equilibrados en todos los casos.

A continuación y en la misma línea se investigó el impacto de las distribuciones no normales de los residuos (pero manteniendo los efectos aleatorios siempre normales) sobre los parámetros fijos y los componentes aleatorios del modelo de genotipos de Soja ("Soybean genotypes model", descrito en Davidian y Giltinan, 1995, y re-analizado por Pinheiro y Bates, 2000). Para ello se generaron series de 1000 tablas de pseudo datos, con residuos generados según las siguientes distribuciones:

- N- distribución normal: que representa el caso dónde la suposición usual de normalidad en los errores es válida.
- SN - distribución normal "skewed" descrita en Azzalini (1986).
- NPM – estimación kernel no paramétrica estimada a partir de datos reales.
- E - distribución Exponencial.
- G - distribución Gamma.

Las últimas dos distribuciones (G y E) se eligieron para representar una situación donde la distribución real de los errores no se distribuye simétricamente y tienen colas más pesadas que la normal. Los parámetros fijos y las componentes aleatorias establecidas para el proceso de simulación fueron los siguientes:

$$\delta = (\delta_1, \quad \delta_2, \quad \delta_3)' = (19.26, 55, 8.4)'$$

$$D = \begin{pmatrix} 25 & 2.50 & 4.00 \\ & 8.00 & 2.32 \\ & & 2.00 \end{pmatrix}; \qquad \sigma = 1.$$

Los resultados obtenidos dependen de cada parámetro en concreto. Para el parámetro δ_1, tal y como se aprecia en la Tabla (4.23) y en el gráfico 4.20 del capitulo 4, no hay diferencias significativas entre los resultados para las cinco distribuciones bajo consideración, con respecto al nivel de recubrimiento, sesgo y MSE. Para esta última medida, los resultados son similares con valores que mantienen entre 0.54 y 0.67. Ninguno de los niveles de recubrimiento del intervalo de confianza logra el valor nominal 95%, pero hay que destacar un robustez aceptable, con niveles de recubrimiento entre 94% bajo la distribución de residuos de sN y 91.5% bajo la distribución exponencial. En el caso de NN se obtuvo el 92.1% del recubrimiento. Para el parámetro δ_2, los resultados demostraron que los valores del sesgo y de MSE son casi idénticos para este parámetro, salvo por la distribución sN donde estas medidas son considerablemente más pequeñas. Los intervalos de confianza actuaron pobremente en todos los casos, salvo en el caso que la distribución de residuos sN. El recubrimiento se mantuvo entre 79.9% para la distribución exponencial (y 84.6% para la distribución gamma) y 93.5% para la distribución sN, con un 83.2% (valor pobre) para los residuos normales. En la correspondencia con los recubrimientos, las anchuras de los intervalos mostraron una tendencia similar, con el mejor valor (el más corto) obtenido bajo la distribución sN. Para el parámetro δ_3, se observó que la distribución de residuos de sN también dio buenos resultados, no sólo desde el punto de vista de MSE (0.06), sino también desde el punto de vista de la probabilidad del recubrimiento (90.7%) y el equilibrio, pero inferior al nivel nominal 95%. En todos los casos, las probabilidades de recubrimiento son pobres, oscilan entre 78.8% (exponencial) y 90.7% para (el caso sN), con 80% en el caso normal, indicando una robustez baja del método LB por este parámetro. De nuevo

estos resultados no se han mantenido para los componentes aleatorios (D y σ). En general, los resultados de ambas series de simulaciones concordaban e indicaban que, cuando la verdadera distribución de residuos es asimétrica, el recubrimiento real de los intervalos asintóticos es notablemente más bajo que el valor nominal fijado. Con respecto al recubrimiento de los parámetros fijos, el mejor resultado correspondió a la distribución de los errores normal "skewed" y no a la normal. Por otro lado, los resultados se invierten parcialmente para los efectos aleatorios.

Dada la relevancia de este método, es innegable el interés de investigar otros aspectos que puedan contaminar la eficiencia del mismo. Otra circunstancia que puede afectar la eficiencia del método LB puede ser el hecho de imponer restricciones erróneas sobre la forma de la matriz de varianzas y covarianzas de los parámetros aleatorios. Por esta razón, se realizó un estudio de simulación de Montecarlo adicional para investigar el impacto de suposiciones incorrectas en la verdadera estructura de la matriz de covarianzas, y el verdadero modelo de la correlación de los residuos, sobre el comportamiento del método de linealización. En la memoria se presentan dos estudios de simulación, cada uno en de los cuales a su vez consiste de dos partes diferentes. En la primera parte de cada estudio de simulación, se investigaron las consecuencias cuando se ajusta un modelo restringido inapropiado. En la segunda parte, se investigó el caso inverso, de manera que el proceso del ajuste se realizó sin restricciones pero con datos generados de acuerdo con algunas estructuras reales determinadas.

En la primera fase del estudio de simulación, se generaron series de 1000 muestras de acuerdo con el Modelo 6, ajustado para los datos de cáncer de mama, generando siempre los parámetros aleatorios y los residuos según la distribución normal. A continuación, para cada una de las tablas generadas se procesó el ajuste del Modelo 6 usando el método LB y suponiendo, cada vez, (en el proceso del ajuste) una de las estructuras de la matriz de covarianzas de los efectos aleatorios (diagonal, diagonal por bloques y no estructurada) combinada con una de tres posibles estructuras de la matriz de correlación de los residuos (residuos independientes, AR(1) y MA(1)). En la segundo fase del estudio de simulación, las series de 1000 muestras se generaron

de acuerdo con las tres "verdaderas" estructuras de la matriz de covarianzas de los efectos aleatorios (diagonal, diagonal por bloques y no-estructurada), combinada con cada una de las tres estructuras de correlación de los residuos normales (residuos independientes, AR(1) y MA(1)). El estudio permite concluir que, en términos de sesgo y de error cuadrático, y de grado de recubrimiento real de los intervalos de confianza asintóticos asociados (e ignorando algunos criterios como la conveniencia de evitar la sobreparametrización de los modelos), tiene consecuencias mucho menos deseables sobre la fiabilidad del método de ajuste y sobre la validez de la inferencia suponer alguna estructura de la matriz de covarianzas de los efectos aleatorios o de los residuos, cuando no es adecuada, que no hacer ninguna suposición de estructura cuando sería válido hacerlo.

Se llevó también a cabo un estudio de simulación para evaluar la robustez de algunas otras técnicas de estimación, como STS (Standard two stage method) y GTS (Global two stage method) (Steimer et al, 1984), que no requieren para su aplicación una estricta suposición de normalidad. Esto permitió comparar técnicas alternativas de ajuste de modelos no lineales mixtos. El método STS trata las estimaciones individuales como si ellas mismas fuesen los verdaderos valores de los parámetros individuales. Esto conduce estimadores a muy simples para los parámetros fijos y de la matriz de la varianzas-covarianzas de los parámetros aleatorios. El estimador de la covarianza D resulta ser muy sesgado. El método de GTS trata de minimizar el sesgo y tiene, en teoría, mejores propiedades estadísticas, pero su cómputo es muy costoso de tiempo; tanto que sólo fue posible realizar series de 100 réplicas de simulación, por lo que los resultados concernientes a este método son meramente ilustrativos. A raíz de lo expuesto, pensamos que la elección entre uno de los dos procedimientos puede depender de la robustez de los mismos hacia las suposiciones del modelo para detectar los efectos implicados en el diseño. Pues, como acabamos de exponer en el apartado anterior, el diseño de simulación ha sido muy similar a lo antes expuesto para investigar la robustez de LB, sólo que las estimaciones de los parámetros se obtuvieron según estos últimos procedimientos, STS y GTS, en lugar de LB. Con respecto al sesgo y MSE, los estimadores de STS para los efectos fijos están (y los estimadores

de GTS parecen ser) notablemente más sesgados y menos eficaces que los estimadores del procedimiento LB correspondiente, en todas las condiciones. Las distribuciones de los efectos aleatorios y los residuos afectan su sesgo y la varianza de la misma manera que a los estimadores de LB. Los intervalos de confianza asintóticos normales asociados al método GTS, parecen dar resultados inadecuados, con recubrimientos extremadamente bajos en muchos casos y dependientes de la distribución y de cada parámetro estudiado. De hecho para el parámetro δ_1, su nivel de recubrimiento es mucho mayor que el nominal, con intervalos de confianza muy largos. Esta tendencia se acentúa para las distribuciones asimétricas de los efectos aleatorios (E y G). Para el parámetro δ_2, esta tendencia se invierte. El nivel del recubrimiento alcanza su mejor valor en el caso de UU (83%) y el peor nivel de recubrimiento se registró en el caso de GN (35.4%). Todos los niveles del recubrimiento están por debajo de 50% para E y G en el caso de los efectos aleatorios. Los peores resultados corresponden al parámetro de δ_3, con todos los niveles de recubrimiento estimados menores de 10%.

0.4. Remuestreo Bootstrap sobre Modelos Mixtos no Lineales

Los resultados de la simulación presentados en el apartado anterior han confirmado que la aproximación paramétrica basada en la hipótesis de normalidad no es fiable cuando la distribución de la variable estudiada se aparta seriamente de la normal. En concreto, los intervalos de confianza aproximados basados en una aproximación lineal, y en general en los resultados asintóticos de la máxima verosimilitud, no son robustos frente a la desviación de la hipótesis de normalidad de los datos, incluso para tamaños muéstrales relativamente grandes. Otros métodos pueden ser más factibles, como el método basado en el perfil de verosimilitud z (Bates y Watts, 1988, Pinheiro y Bates, 1995) pero este método depende aún más de la forma de la verosimilitud (además, computacionalmente es intensivo como la aproximación bootstrap discutida a continuación) y por lo tanto es difícil aplicarlo bajo un enfoque no paramétrico. Hay aproximaciones alternativas basadas en relajar la forma de la distribución de los residuos y/o efectos aleatorios, estimándola parcialmente de los datos, como en Davidian, y Gallart (1993) y

Tao y Yandell (1999). Los métodos de Bootstrap (véaase por ejemplo Efron, 1979) pueden proporcionar otras (y quizá adecuadas) alternativas inferenciales. Cómo realizar el remuestreo bootstrap en un contexto de datos de medidas repetidas y modelos no lineales con efectos mixtos es una cuestión controvertida. Das y Krishen (1999) estudian el mejor procedimiento para generar las muestras bootstrap a partir de datos longitudinales bajo modelos mixtos. Un tema menos estudiado es el de cómo aplicar el principio "plug-in" en la construcción de los estadísticos al nivel de las muestras. En nuestra opinión, la pauta principal está en que el remuestreo bootstrap debe reproducir el "verdadero" proceso que se llevó acabo para obtener los "verdaderos" datos. La idea básica es de considerar dos etapas, con diferentes componentes aleatorias, en el proceso del remuestreo:

Primero, se han seleccionado algunas unidades experimentales, posiblemente al azar, de una población de interés. Esto da lugar a una muestra de m unidades experimentales, $i = 1, \ldots, m$. Segundo, se ha realizado un experimento sobre cada una de estas unidades, dando lugar a un conjunto de medidas repetidas para cada una, a lo largo de una dimensión adecuada como el tiempo o el espacio, $\mathbf{y}_i = \left(\mathbf{y}_{i1}, \ldots, \mathbf{y}_{in_i}\right)$, y en definitiva a una conjunto completo de datos, $\mathbf{y}$ $\mathbf{y} = \left(\mathbf{y}_1, \ldots, \mathbf{y}_m\right)'$. El remuestreo Bootstrap debe reproducir ambas etapas. Por ejemplo, omitir el segundo paso es equivalente a asumir implícitamente que las medidas observadas por cada individuo son valores fijos, que el experimento realizado sobre cada individuo siempre dará el mismo resultado. Esto sugiere que un proceso aparentemente intuitivo y simple de remuestrear como, primero, realizar ajustes individuales de los parámetros fijos y, segundo, remuestrear estos parámetros ajustados para cada individuo de la manera usual, como observaciones independientes, no dará, en general, buenos resultados.

Retomando el concepto de remuestreo bietápico sugerido en el párrafo anterioor, mientras que su primera etapa se acomoda al bootstrap usual, suponiendo que los datos son independientes e idénticamente distribuidos, la segunda etapa está relacionada con el tema más complejo del remuestreo de datos dependientes. Los métodos de remuestreo más usuales para serie temporales (vea una revisión en Alonso, Peña y Romo, 2002) no parecen adecuados en un contexto de modelos mixtos, dado que se basan en suposiciones de estacionariedad, poco adecuadas en la

mayoría de situaciones reales tratadas en esta memoria. Un método de remuestreo que parece libre de todas estas restricciones es el Bootstrap en Cadena de Markov ("Markov Chain Bootstrap") (MCB) introducido en Vasnev (2001) y Anatolyev y Vasnev (2002). Con modificaciones menores, se conjetura su validez en el contexto de este trabajo.

Cómo remuestrear depende de cada problema específico y en las suposiciones que se hacen. En el capitulo 5 se ha elaborado una esquema de diferentes métodos de remuestreo boostrap basados en dos categorías principales de suposiciones: Aquellas acerca del modelo no lineal mixto para ajustar los datos y, dependiendo del modelo no lineal mixto supuesto, aquellas acerca de los aspectos de la distribución de los factores aleatorios y de los residuos.

Concentrándonos en el caso del bootstrap no paramétrico, donde no se quiere hacer ninguna suposición sobre el modelo no lineal y tampoco acerca de las distribuciones, la manera más obvia de implementar el procedimiento basado en el remuestreo en dos-etapas, entre las unidades experimentales y dentro de cada unidad experimental, es la que figura en el esquema mostrado en la Ilustración 1, donde en el primer paso, se remuestrea al azar, y con reemplazamiento, filas completas de la matriz de datos, es decir, se construiye una nueva matriz de los datos Y* con m filas donde la fila i se define como $\mathbf{y}_i^* = \mathbf{y}_{i*}$, e i es un valor tomado con equiprobabilidad del conjunto {1,2,...,m}. En algunas situaciones, el proceso del remuestreo precedente se realizará de una manera estratificada, tomando m_k índices al azar y con reemplazamiento para cada estrato k, con $m = \sum_{k=1}^{s} m_k$. El segundo paso del remuestreo se tiene que llevar a cabo dentro de cada unidad experimental elegida en la primera etapa según un método apropiado para el remuestreo de datos correlacionados. Esto dará lugar al conjunto final de datos remuestreados $\mathbf{y}^{**} = \left(\mathbf{y}_1^{**},\ldots,\mathbf{y}_m^{**}\right)'$.

Ilustración 1 Esquema del procedimiento para la obtención del estimador bootstrap

$$\mathbf{Y} = \left(\mathbf{Y}_{ij}\right)_{\substack{1\le i\le m\\ 1\le j\le n}} \quad \text{datos originales}$$

$$\downarrow \text{Paso 1}$$

$$\{1,2,\ldots,i,\ldots,m\} \xrightarrow[P=\frac{1}{m}]{\text{remuestreo bootstrap}} Y^* = \begin{pmatrix} \mathrm{y}_{1^*1} & \cdots & \mathrm{y}_{1^*n} \\ \downarrow & \cdots & \downarrow \\ y_{m^*1} & \cdots & \mathrm{y}_{m^*n} \end{pmatrix}$$

$$\downarrow \text{Paso 2}$$

$$\Rightarrow \mathrm{y}_i^* = \mathrm{y}_{i^*} = \left(\mathrm{y}_{i^*1}, \mathrm{y}_{i^*2}, \ldots, \mathrm{y}_{i^*n}\right) \xrightarrow[p_t=\begin{cases} 0 & \text{si } y^*_{i,t-1} \notin I_k \\ \dfrac{1}{\#\left\{y^*_{il}\in I_k\right\}} & \text{si } y^*_{i,t-1}\in I_k \end{cases}]{\text{remuestreo bootstrap}} \mathbf{y}_i^{**} = \left(y_{i1}^{**},\ldots,y_{in}^{**}\right)$$

$$I_k=[a_k\text{-}1,\ a_k], a_0=y^*_{i(1)},\ \#\left\{a_{k-1}\le y^*_{ij}\le a_k\right\}=\nu,\ k=1,\ldots,r.$$

$\rightarrow$ *S*e repite el paso 2 para i=1,2,...,m

$$\downarrow \text{Paso 3}$$

$$Y^{**(b)} = \begin{pmatrix} \mathrm{y}_{11}^{**} & \cdots & \mathrm{y}_{1n}^{**} \\ \downarrow & \cdots & \downarrow \\ y_{m1}^{**} & \cdots & \mathrm{y}_{mn}^{**} \end{pmatrix}^{(b)} \quad \boxed{\text{Muestra bootstrap}} \Rightarrow \text{Valores bootstrap}$$

$\boxed{B = 1000}$ Se repite Paso 1, 2,3 para $b = 1,2,\ldots,B$.

$$\downarrow \text{Paso 4}$$

$$\Rightarrow \text{ Calculo de: } \mathcal{S}^* = \left\{\left(\hat{\delta}^*(1), \hat{D}^*(1)\right),\ldots,\left(\hat{\delta}^*(B), \hat{D}^*(B)\right)\right\},\ \ldots$$

Como se mencionó anteriormente, el método MCB (Bootstrap en Cadena de Markov) parece especialmente apropiado para nuestros propósitos. En él, cada nuevo elemento de la remuestra final se ha obtenido de acuerdo con la probabilidad siguiente:

$$p_t = \begin{cases} 0 & \text{si } y^*_{i,t-1} \notin I_k \\ \dfrac{1}{\#\left\{ y^*_{il} \in I_k \right\}} & \text{si } y^*_{i,t-1} \in I_k \\ I_k = [a_k - 1,\ a_k],\ a_0 = y^*_{i(1)}, \\ \#\left\{ a_{k-1} \le y^*_{ij} \le a_k \right\} = \nu, \\ k = 1, \ldots, r. \end{cases},$$

donde I_k son intervalos construidos de la manera indicada en el capitulo 5. En el tercer paso se crea la muestra bootstrap para todos los individuos utilizando el paso 2, y se calcula los valores bootstrap. En el cuarto paso, se repiten los pasos 1, 2, 3, en nuestro caso un número de veces igual a 1000 (o a veces 100), y se calcula las medidas especificadas.

En un contexto de ajuste mediante modelos mixtos, para aplicar en la práctica el esquema de remuestreo anterior, es necesario escoger un procedimiento de ajuste concreto, para analizar los datos originales y cada uno de los preudo-datos bootstrap. Bajo el modelo de remuestreo libre (con respecto a ambos, los modelos no lineales y la forma de las distribuciones), parece más adecuado escoger también un procedimiento de ajuste libre de suposiciones estrictas, si es posible. Ignorando para el momento sus limitaciones, el método estándar de dos etapas, STS parece una opción adecuada. En breve, el método es como sigue:

1. Como en todos los métodos en dos etapas, se obtienen las estimaciones individuales para los parámetros fijos., $\hat{\delta}_1, \hat{\delta}_2, \ldots, \hat{\delta}_m$.

2. Suponiendo que la forma general del modelo es $\begin{aligned} \mathbf{y}_i &= f(\mathbf{t}_i, \delta_i) + \mathbf{e}_i \\ \delta_i &= \mathbf{X}_i \delta + \mathbf{Z}_i \eta_i \end{aligned}$ donde la matriz de diseño $\mathbf{Z}_i$ es la identidad, es decir, los verdaderos parámetros fijos para el i-ésimo individuo o unidad experimental pueden expresarse como $\delta_i = \mathbf{X}_i \delta + \eta_i$. Entonces, tomando las estimaciones como si ellas mismas fuesen los verdaderos valores de δ_i, se obtienen los estimadores siguientes para los parámetros fijos y la matriz de covarianzas de los factores aleatorios:

$$\hat{\delta}_{STS} = \left[\sum_{i=1}^{m} (\mathbf{X}_i' \mathbf{X}_i) \right]^{-1} \left[\sum_{i=1}^{m} \mathbf{X}_i' \hat{\delta}_i \right]$$

y

$$\hat{D}_{STS} = (m-1)^{-1}\sum_{i=1}^{m}(\hat{\delta}_i - \mathbf{X}_i\hat{\delta}_{STS})(\hat{\delta}_i - \mathbf{X}_i\hat{\delta}_{STS})'.$$

Estas estimaciones no tienen propiedades muy buenas para el sesgo, y no hay una manera obvia de hacer inferencia basada en ellos. Pero se puede conjeturar que la metodología del bootstrap proporcionará las soluciones adecuadas para todos estos inconvenientes. En concreto, y completando el esquema anterior detallando el paso 3, el procedimiento del bootstrap se lleva a cabo como sigue:

◇ Obtención de las estimaciones de STS sobre los datos iniciales, $\hat{\delta}_{STS}$, para los parámetros fijos y $\hat{D}_{STS}$ para la matriz de covarianzas de los parámetros aleatorios.

◇ Uso de los valores originales de los estadísticos, y los obtenidos aplicando el bootstrap (paso 4), para calcular la corrección del sesgo de los parámetros fijos y de la matriz de covarianzas de los efectos aleatorios utilizando las siguientes formulas

$$\hat{\delta}_{BCSTS} = 2\hat{\delta}_{STS} - E^*\left(\hat{\delta}^*_{STS}\right), \text{ with } E^*\left(\hat{\delta}^*_{STS}\right) \cong \frac{1}{B}\sum_{b=1}^{B}\hat{\delta}^*_{STS}(b) \text{ y}$$

$$\hat{D}_{BCSTS} = 2\hat{D}_{STS} - E^*\left(\hat{D}^*_{STS}\right), \text{ with } E^*\left(\hat{D}^*_{STS}\right) \cong \frac{1}{B}\sum_{b=1}^{B}\hat{D}^*_{STS}(b).$$

Además se han calculado diversos intervalos de confianza bootstrap para la estimación de los efectos fijos. Se utilizaron dos procedimientos para la obtención de los intervalos de confianza: el método percentil (Efron, 1979) y bootstrap-t simetrizado (Efron y Tibshirani, 1993). El primero asigna como extremos inferior y superior del intervalo de confianza de recubrimiento nominal 1-α, los percentiles $\alpha/2$ y 1-$\alpha/2$ de la distribución bootstrap del estimador puntual. El segundo compensa las limitaciones del método percentil ante la ausencia de simetría en la distribución del estimador, y en aquellas situaciones en que la forma de la distribución cambia dependiendo de los valores del parámetro.

Para estudiar el comportamiento de las estimaciones bootstrap y los dos diferentes intervalos de confianza, se ha diseñado un estudio simulación de Montecarlo en el que el proceso de generación de los datos era el mismo que en el apartado anterior, es decir, se generaron los datos

simulados según las mismas distribuciones (Normal, Gamma, Exponencial y Uniforme) para los efectos aleatorios y los residuos, y según modelo descrito. Los parámetros de la población también eran los mismos especificados anteriormente. Tal como se aprecia en los resultados presentados en el capitulo 6 los intervalos bootstrap-t proporcionan, en la mayoría de las condiciones, mejores resultados que los intervalos del percentil, donde su nivel de recubrimiento depende de cada parámetro en concreto y en las distribuciones simuladas, con valores bajos en algunas condiciones. En cambio, los intervalos de bootstrap-t arrojan un nivel de recubrimiento bastante robusto en todas las condiciones, con valores cercanos al recubrimiento nominal de 0.95.

Para finalizar subrayemos que este estudio de simulación (preliminar, debido al reducido número de réplicas de simulación, 100) permite concluir que para poblaciones de cuya distribución no pueda afirmarse que sea cierta la hipótesis de normalidad, el método "bootstrap" proporciona un estimador de los parámetros, en términos de amplitud del intervalo y de su cobertura relativamente más adecuado que el método clásico, basado en la hipótesis de normalidad de la variable estudiada. Pero se necesita un estudio más extenso para mejorarlo y mejorar los métodos computacionales que permitan hacer el cálculo más factible y permitir estudios de la simulación más concluyentes.

Chapter 1

1. Introduction and Motivation

1.1. Introduction

The purpose of this chapter is to introduce some data sets that are extensively analyzed along this monograph, and to introduce some approaches to analyze longitudinal data, that is, data in form of repeated measurements on the same experimental unit over time or over another set of conditions, like drug concentrations. The same response is measured repeatedly on each member of a sample of m individuals. "Individuals" may correspond to diverse entities like humans, animals, plants, laboratories, or experiments. Data are routinely collected in this fashion in a broad range of applications, including agriculture, life sciences, medical and public health research, and industrial applications. For example:

- In agriculture, a measure of growth may be taken on the same plot weekly over the growing season. Plots are assigned to different treatments at the start of the season.
- In medical studies, a measure of the tumour growth may be taken at weekly intervals on rats with induced cancer. Rats are assigned to take different treatments at the start of the study.
- In industrial applications, measures are also taken on the effect of diverse industry residuals on the process of mineralization of carbon and nitrogen.

The scientific questions of interest often involve not only the usual ones, such as how the mean response differs across treatments, but also how the change in mean response over time differs and other issues concerning the relationship between response and time.

1.2. Random two-stage model

The data sets shown in Figures 1.1 to 1.5 were obtained and analyzed by Dra. Genoveva Montserrat, Departament de Productes Naturals, Biologia Vegetal i Edafologia, Facultat de Farmàcia, Universitat de Barcelona. The main goal of the experiment was to examine the effects of industrial residuals on different soil types. Each of two soil conditions, SN: "*Sol classificat com a Xerumbrept segons Sol Taxonomy System, i es troba a la reserva natural de la Font Groga, Serra de Collserola, al terme municipal se St Cugat del Vallés*" and SO: "*Sol correspond a un Xerorthent, i es troba ubicat a la Sta. Creu d´Olorda, també a la Serra de Collserola, molt a la vora d´una exploració de calcàries, al terme municipal de Molins de Rei*", was exposed to four residual types (with comparable properties to common industrial residuals) labelled (CR–Residu ceràmic: "*procedent d'una indústria de fabricació de sanitaris*", CC–Cendra Volant: "*residu procedent d'una central termoelèctrica que combustiona carbó*", LS–llot orgànic: "*Originat en la depuració d'aigües de neteja i sanitàries d'una indústria d'automació*" and LG–llot galvánic: "*obtingut de la depuració de residus líquids d` una indústria de tractaments de superfícies metàl·liques (galvanitzats, cromats, niquelats ect.) concretament de botons*". Each with five increasing levels denoted (0 (None), 50 (dose of 50 $mg.ha^{-1}$), 100 (dose of 100 $mg.ha^{-1}$), 250 (dose of 250 $mg.ha^{-1}$) and 500 (dose of 500 $mg.ha^{-1}$) according to Table (1.1).

Table (1.1): Scheme of the design experiment

Soil type	Residual type	Dose factor				
SN	CR	0	50	100	250	500
SN	CC	0	50	100	250	500
SN	LS	0	50	100	250	500
SN	LG	0	50	100	250	500
SO	CR	0	50	100	250	500
SO	CC	0	50	100	250	500
SO	LS	0	50	100	250	500
SO	LG	0	50	100	250	500

Next, during 62 days, the quantity of global ammoniac and residual, resulting from the transformation of N-organic (organic nitrogen) into N-ammoniac (ammoniacal nitrogen) was observed; and also the quantity of nitrite, resulting from transformation of the N-ammoniac into N-nitric, according to the following illustrative diagram:

$$\text{N-organic} \xrightarrow{\text{Amonification}} \text{N-NH}_4$$
$$\underbrace{\text{N-NH}_4 \longrightarrow \text{N-NO}_2^- \rightarrow \text{N-NO}_3^-}_{\text{Nitrification}}$$

Additionally, the quantity of daily and accumulated carbon (CO_2) along 100 days of incubation was observed. The main objective of the study was to: elucidate the evolution of the various responses ("dose-responses"), labelled "nitrification", "global ammonification", "residual ammonification", "daily carbon" and "accumulated carbon" in function of time.

1.2.1. Nitrification data

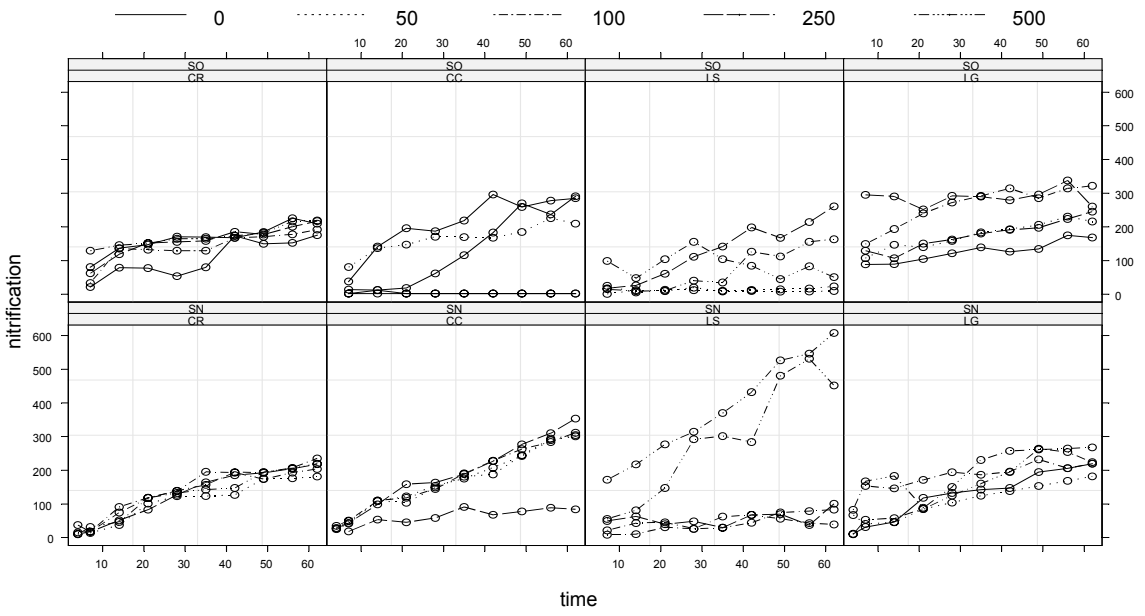


Figure 1.1: nitrification data in function of time

A good starting point is to realize that each plot has its own trajectory with its own peculiar features. If we focus on the trajectory of a particular curve, we see that, typically, the behaviour

profile varies with plots, suggesting that the relationship between the response and time for each plot may be described as follows:

INDIVIDUAL (FIRST STAGE) MODEL:

$$y_{ij} = \delta_{i0} + \delta_{i1} t_{ij} + \delta_{i2} \log(t_{ij}) + e_{ij} \tag{1.1}$$

where i refers to the plot number, j refers to the time at which the j measurement on plot i was taken, y_{ij} represents the observation j for the i–th plot. The coefficients δ_{i0}, δ_{i1} and δ_{i2} denote respectively the intercept, the "slope" and the "individuals" coefficients representing the effect of the logarithm of time. The time observations vary from these trajectories due to errors resulting form the data collection process, represented by e_{ij} for the j–th response. These within "errors" are taken to have zero mean and some variance, representing the magnitude of variation on measurements within the plots.

POPULATION (SECOND STAGE) MODEL: Model (1.1) only tells part of the history; it describes what happens at the level of an individual plot and includes explicit mention (through e_{ij}) of the within plots variation. However, it does not by itself acknowledge among-plots variation. We have recognized that the trajectories differ across plots. For now, we may think of the plot observed as arising from a population of all such plots. All plots have their own coefficients; thus, we may think abstractly in this population as the population of all vectors $\delta_i=(\delta_{i0},\ \delta_{i1},\ \delta_{i2})'$, one for each plot. It is natural to think in this population as being "centred" around a "typical" coefficients value. More formally, we may think in the mean value of each coefficient in the population of all such δ_i vectors. Individual parameter vectors vary around this mean. This way of thinking suggests a model for this population as follows:

Let $\delta=(\delta_0,\ \delta_1,\ \delta_2)'$ be the mean vector of the population of all vectors of form $\delta_i= \delta+\eta_i$ which is a shorthand way for saying

$$\begin{cases} \delta_{i0} = \delta_0 + \eta_{i0} \\ \delta_{i1} = \delta_1 + \eta_{i1} \\ \delta_{i2} = \delta_2 + \eta_{i2} \end{cases} . \tag{1.2}$$

Here, $\eta_i=(\eta_{i0}\ ,\ \eta_{i1}\ ,\ \eta_{i2})'$ denotes a vector of random effects associated to individual parameters. In model (1.2) we assume that the η_i are independent and identically distributed with common distribution $N(0,\sigma^2 D)$ and that e_{ij} are independent and identically distributed with common distribution $N(0,\sigma^2)$ and are independent of the η_i. The vector of random effects, η_i, describes how the coefficients for the *i–th* plot deviate from the mean value. Thus, (1.2) has the flavour of a regression-type model for the plot-specific regression parameters, with a systematic component, the mean, and a random component summarizing how things vary about it.

Putting the two models (1.1) and (1.2) together gives a complete description of what we believe about each plot and the population of all plots, explicitly identifying the two sources (within and among plots) of variation. We may write the model for the *i–th* plot in the alternative form

$$y_{ij} = (\delta_0 + \eta_{i0}) + (\delta_1 + \eta_{i1})t_{ij} + (\delta_2 + \eta_{i2})\log(t_{ij}) + e_{ij} \qquad (1.3)$$

1.3. Linear mixed models

Random coefficients models, where we develop an overall statistical model by thinking first about individual trajectories in a "subject-specific" fashion, are a special case of a more general model framework based on the same perspective. This model framework, known popularly as the linear mixed effect model, is still based on thinking about individual behaviour first, of course. However, the possibilities for how this is represented, and how the variation in the population is represented are broadened. The result is very flexible and result in a rich set of models for characterizing repeated measurement data.

1.3.1. Global amonification data

In this section, we attempt to model the global ammonification, shown in Figure 1.2 and already described in section 1.2 (see Montserrat, 2000, for more details)

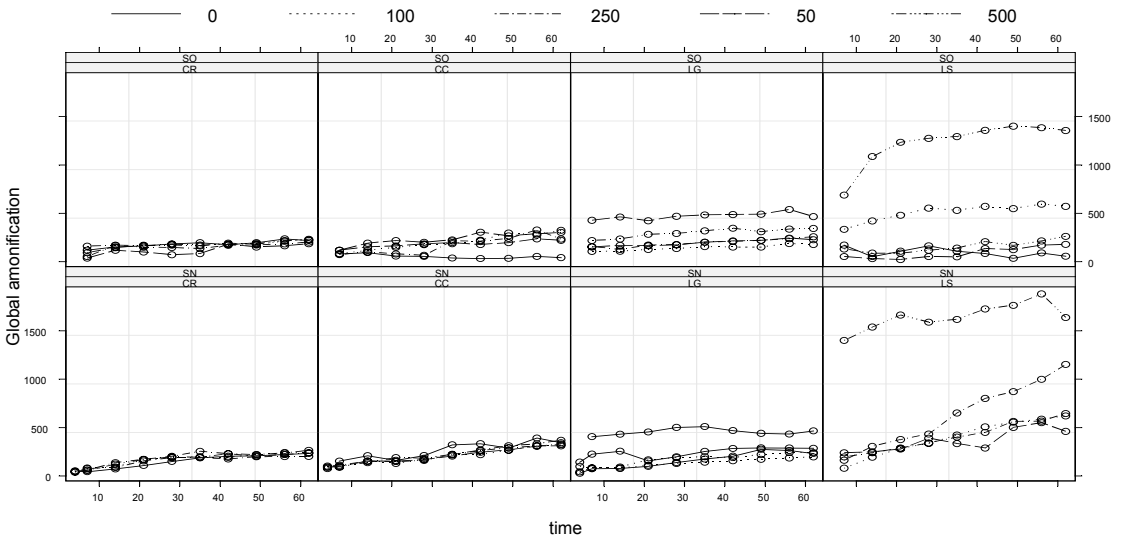


Figure 1.2: Global ammonification data profile

We propose a regression model characterizing each curve above,

$$y_{ij} = \delta_{i0} + \delta_{i1} t_{ij} + e_{ij}. \tag{1.4}$$

The rate of change over time (slope) of each profile seems very similar across residual types, except for the "LS" industrial residual, keeping in mind that there is variation within units, making the profiles not to look perfectly like straight lines.

The upshot is that the intercepts appear to definitely vary across units, while the slopes do not seem to vary much at all. There are two possibilities:

One possibility is that the "true" underlying slopes are identical for all units in the population.

A more reasonable explanation may be that, relative to how the intercepts vary across units, the variation among the slopes is much smaller, making them appear to hardly vary at all. It may be that the rate of change over time for this population is quite similar, but not exactly identical, for all units.

If the first possibility is true, we might consider a model that reflects the fact that slopes are virtually identical across units explicitly. The following "second-stage" model would accomplish this goal:

$$\begin{aligned} \delta_{i0} &= \delta_0 + \eta_{i0} \\ \delta_{i1} &= \delta_1 \end{aligned} \quad . \tag{1.5}$$

Note that in (1.5) the individual-specific slopes, δ_{i1}, have no random effects associated with them. This formally reflects the belief that the δ_{i1} do not vary in the population of units. Thus, under this population model, while the intercepts are random, with an associated random effect and thus varying in the population, the slopes are all equal to the fixed value δ_1 and do not vary at all. Thus, there is only a single, scalar random effect η_{i0}. The covariance matrix for the population, D, reduces to just a single variance.

Plugging the representations for δ_{i0} and δ_{i1} into the first stage model (1.4), we obtain

$$y_{ij} = (\delta_0 + \eta_{i0}) + \delta_{i1} t_{ij} + e_{ij}.$$

If we believe that the second possibility is true, the usual random coefficients model still applies:

$$\begin{aligned} \delta_{i0} &= \delta_0 + \eta_{i0} \\ \delta_{i1} &= \delta_1 + \eta_{i1}. \end{aligned}$$

Then, in matrix D, the value D_{11} representing the variance of η_{i0} (among intercepts) would be much larger than the value of D_{22} representing the variance of η_{i1} (among slopes).

1.3.2. Residual ammonification data

Figure 1.3 show that, the residual ammonification has the same profile than the global ammonification with the difference that both, the individual specific intercepts and the slopes, δ_{i0} and δ_{i1}, seem to be both random. This leads to the corresponding linear mixed model

$$y_{ij} = \left(\delta_0 + \eta_{i0}\right) + (\delta_1 + \eta_{1i}) t_{ij} + e_{ij}. \tag{1.6}$$

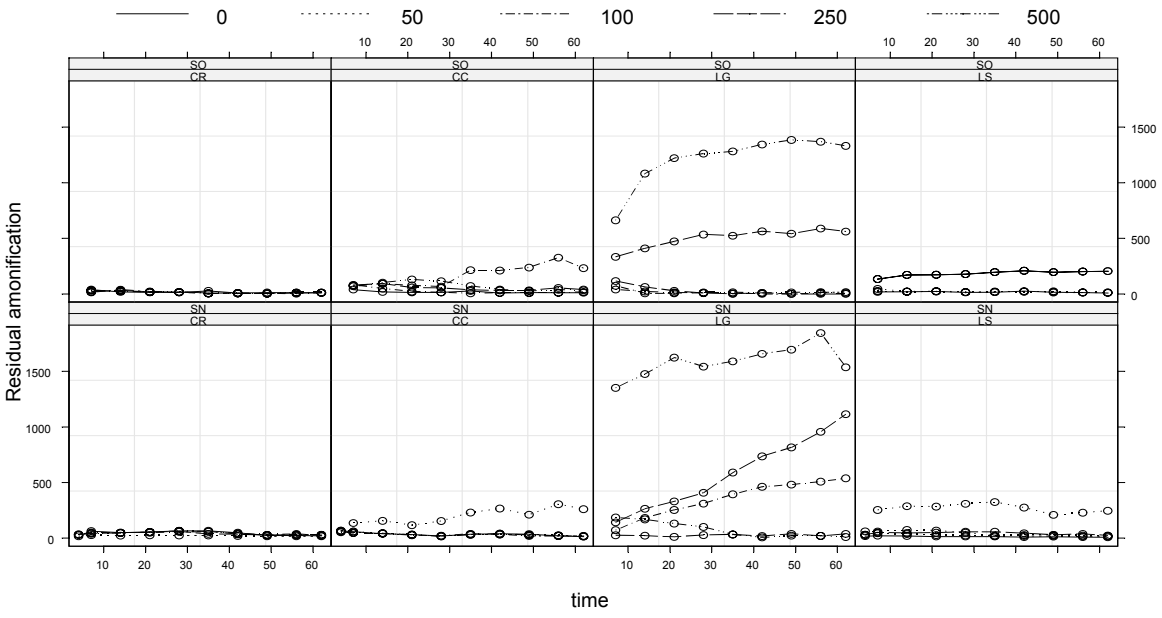


Figure 1.3: Nitrogen data – residual ammonification

1.4. Non-linear mixed effects models

Non-linear mixed models, which generalise non-linear models and linear mixed effects models, are more flexible than linear mixed models. The concept is outlined using the following example.

1.4.1. Daily carbon data

Figure 1.4 corresponds to the daily carbon profile in function of time within each combination of industrial residuals type and soil type.

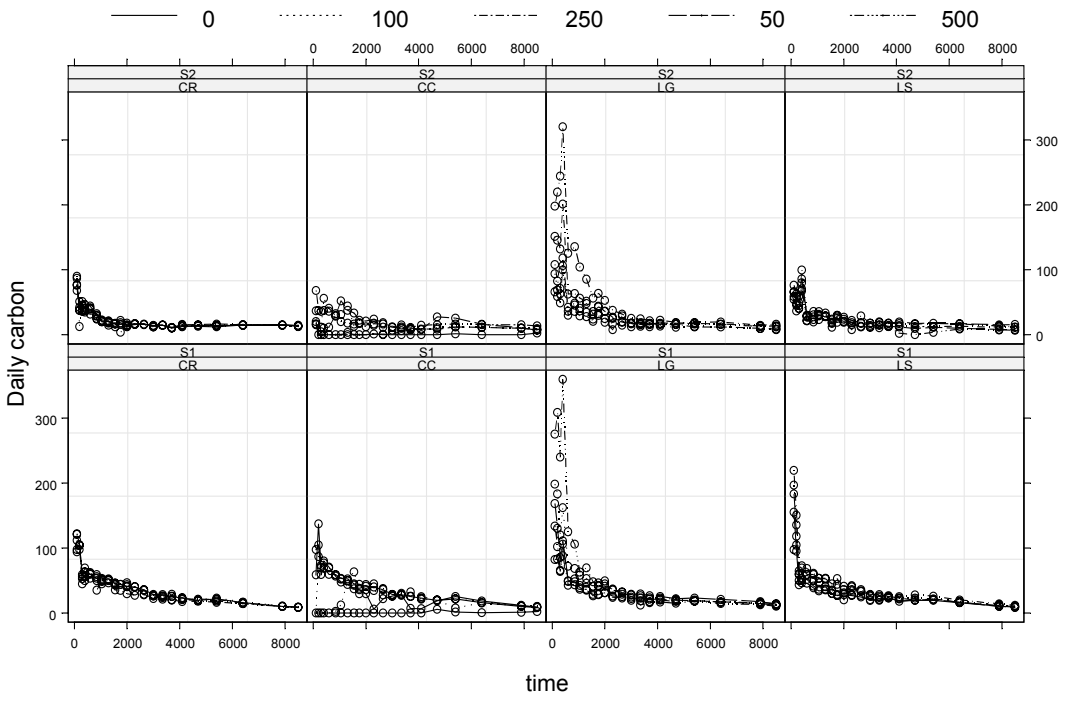


Figure 1.4 Daily carbon for each dose within each industrial residual type for the two soil types

The shapes of these trajectories suggest that a biexponential function

$$\begin{cases} y_{ij} = \delta_{1i}.\exp(-\delta_{2i}t_j) + \delta_{3i}.\exp(-\delta_{4i}t_j) + e_{ij} \\ \delta_{1i} = \delta_1 + \eta_{1i} \\ \delta_{2i} = \delta_2 + \eta_{2i} \\ \delta_{3i} = \delta_3 + \eta_{3i} \\ \delta_{4i} = \delta_4 + \eta_{4i} \end{cases} \tag{1.7}$$

used in Davidian and Giltinan (1995) and in Pinheiro and Bates (2000) is a reasonable representation of daily carbon in function of time, instead of the uniexponential function used in Montserrat (2000).

δ_1 and δ_3 may be interpreted as the initial daily carbon (at t=0) and δ_2 and δ_4 represent the rates of change over time.

1.4.2. Accumulated carbon data

These data were used in Montserrat (2000) to measure the respiratory activity of soils. They express the flux in (mg/100 g soil) liberated by the soil surfaces along 100 days of incubation. Figure 1.5 graphically displays the relation between accumulated carbon and time, according to dose and industrial residual type.

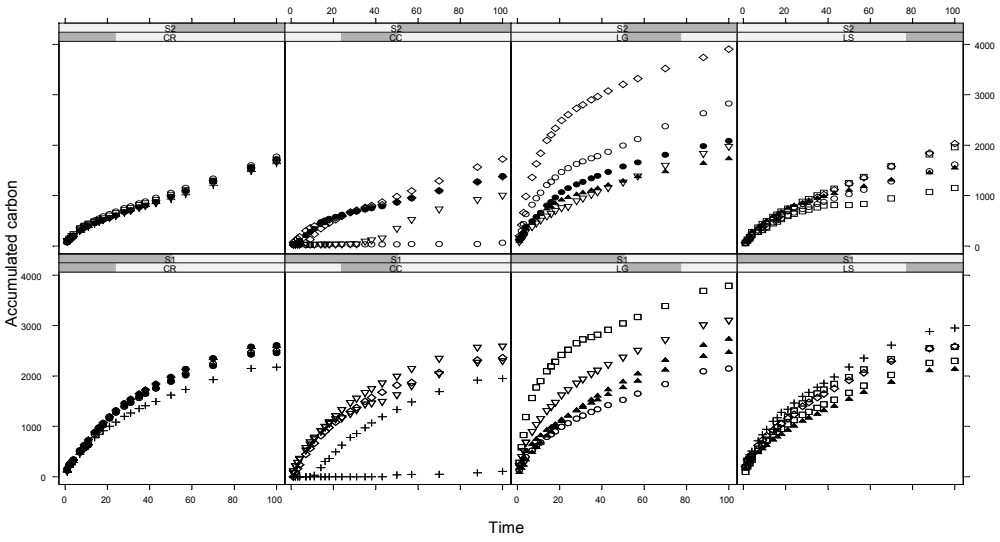


Figure 1.5 Accumulated carbon in function of time for each combination of dose and industrial residual type

The suggested model is,

$$\begin{aligned} y_{ij} &= \delta_{1i}(1-e^{-\delta_{2i}t_{ij}}) + e_{ij}, \quad i = 1,2,...,n, \quad j = 1,2,....,n \\ \delta_{1i} &= \delta + \eta_{1i}; \quad \delta_{2i} = \delta + \eta_{2i} \qquad . \end{aligned} \tag{1.8}$$

The parameter δ_{1i} represents the initial accumulation of carbon at zero time, and δ_{2i} represents the rate of change in the carbon accumulation.

1.4.3. Breast cancer data

The statistical analyses described in the next chapters are based on data from a study on the influence of lipids on the development of cancer. These results are integrated in a series of studies performed by the team of Dr. Eduard Escrich during more than 15 years, in order to

determine the dynamics of breast tumour growth under a variety of conditions (Escrich, Solanas and Segura, 1994, Escrich et al., 1994). Figure 1.6 presents data from an experiment of carcinogenesis in which sixty 22-day-old virgin female Sprague-Dawley rats were housed four per cage and maintained in an environmentally controlled room at 24 $\pm$ 1^{o} C, at a 50% of humidity, in a 12 hours light/12 hours dark cycle. Upon arrival, rats were fed *ad libitum* with one of two different semi-synthetic diets, 40 received a low-fat diet and 20 a high-fat diet. At 53 days of age, all animals received a single dose of 5 mg of carcinogen (7, 12-dimethylbenz(α) anthracene - DMBA-, Sigma) per rat administered in corn oil by means of a gastric gavage (Huggins, Grand and Brillantes, 1961, Escrich, 1987). One day after the administration of the carcinogen, twenty animals from the low-fat diet group were randomly chosen and permanently transferred to the high-fat diet, the remaining animals continued with the initial diet until the end of the study. In other words, three treatment groups of 20 rats each one were finally formed: always low-fat diet (labelled as "Diet 1" in the analyses outputs), always high-fat diet ("Diet 2") and low-fat before the carcinogen administration / high-fat after the carcinogen administration ("Diet 3").

The rats were examined and palpated for mammary tumours once per week. When a tumour was first palpated, the date, its location and an estimate of its volume were recorded. At the end of the study, 201 days after carcinogen administration, rats were decapitated. At necropsy, tumours were rapidly removed, measured, rinsed in normal saline, and divided for histopathology. Only confirmed mammary adenocarcinomas were reported in the results. The data analyzed in this section are the total volumes of the tumours in each rat.

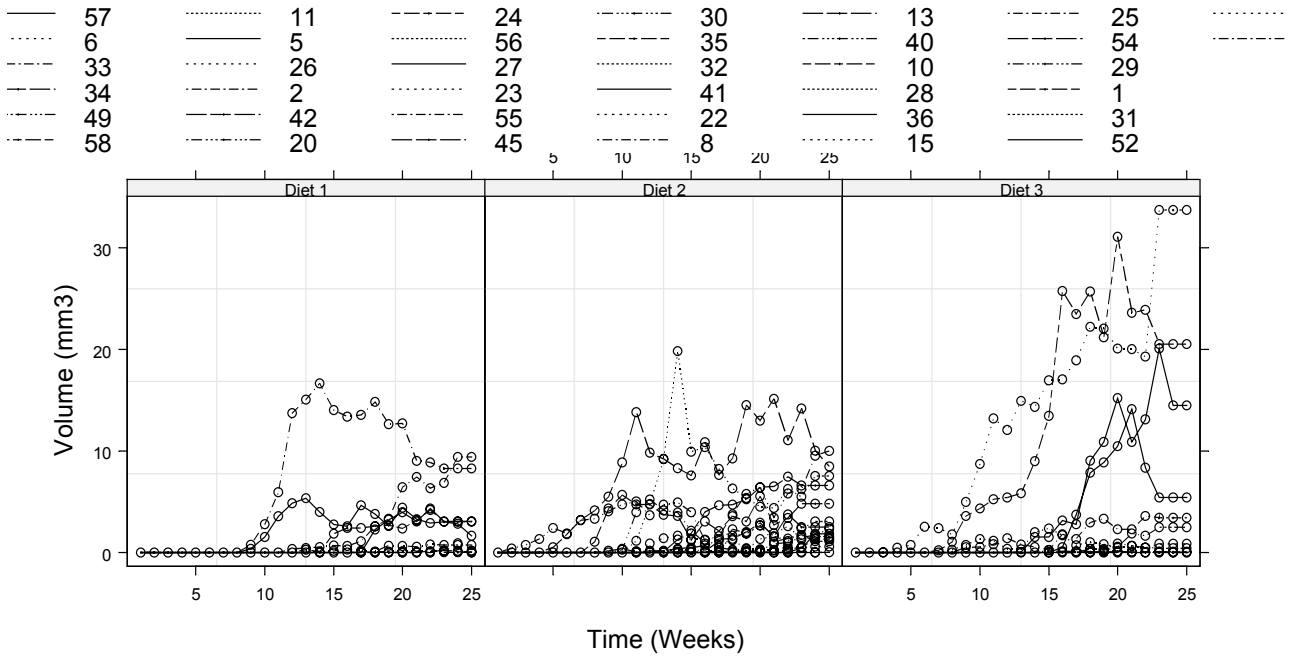


Figure 1.6: Tumoral volume of rats measured over a period of 25 weeks. The rats are divided in three groups of different diets

The objectives of the study were to:

- Establish the relationship between the growth dynamics of tumoral volume and time.
- Compare the growth characteristics of different treatment groups (diets).
- Explain the variability between rat responses.
- Examine the change between rates with respect to diets.

It's evident from Figure 1.6, that several features are notable; if we carefully look to the trajectory of each particular plot, we can see that, typically, some have flat form, others start with a low rate of growth that increases in the middle of the observation period. Overall, a straight line is not adequate to describe trajectories. Rather, the form of all trajectories seems more complicated. It seems also clear that there are similarly shaped profiles for each plot, except for some ones that have a flat form. The differences between curves suggest treatment

differences and possible interactions between time and treatment. Heterogeneous within–plot variability seems also clear, which may also include serial correlation.

Grouping all these characteristics, we can suggest the following model that represents the situation of several such plots by allowing each plot to have its own model, with its own parameters. More formally, if y_{ij} is a tumoral volume measurement at time t_{ij} for the $i-th$ rat,

$$y_{ij} = \frac{\delta_{1i}.\exp(t_{ij} - \delta_{2i})}{1 + \exp(\frac{t_{ij} - \delta_{2i}}{\delta_{3i}})} + e_{ij}$$

where:

$$i = 1,2,...,60$$
$$j = 1,2,...,25 \qquad (1.9)$$
$$\delta_{1i} = \delta_1 + \eta_{1i}$$
$$\delta_{2i} = \delta_2 + \eta_{2i}$$
$$\delta_{3i} = \delta_3 + \eta_{3i}\,.$$

(δ_1, δ_2, δ_3) represent mean values of, respectively, maximal tumoral volume, the time until the 50% maximal volume, and the rate of growth. They are usually referred as fixed effects, characterizing the systematic part of the response. On the other hand, $\eta_i=(\eta_{1i}, \eta_{2i}, \eta_{3i})'$ are the random effects associated to these fixed effects, characterizing the among unit (within-rat) variation. The model (1.9), and more general versions of it, are subject (rat) specific models.

1.5. Multilevel non-linear mixed models

In some cases, the preceding "single level" approach must be extended to a multilevel one, where multiple nested factors are considered:

Suppose that the data consists of N_1 communities and each community consists of N_2 families with N_3 children within each family. Here communities, families and children define level one, two and three for a three level data respectively. Suppose that we are interested in estimating the effect of x_{ijk}, x_{ij}, and x_i, the explanatory variables in levels one, two, and three respectively, on the response measured for each child. Moreover assume that we believe the community and the family populations are heterogeneous. To control for heterogeneity we introduce two random

effects B_{ij} and B_i at the second and first level. With these specifications the linear model will be of the form

$$\delta_1 x_{ijk} + \delta_2 x_{ij} + \delta_3 x_i + B_i + B_{ij} \text{ , i = 1, 2, ...,}N_1\text{, j = 1, 2, ...,}N_2\text{, k = 1, 2, ...,}N_3.$$

One may estimate δ_1, δ_2, δ_3 the fixed effects of x_{ijk}, x_{ij}, and x_i and the variances of B_{ij} and B_i by using appropriate estimation procedure.

Following we introduce, by the real example, a two levels mixed models that consists of a nonlinear component of which may contain fixed and random effects:

1.5.1. Fangar data: two levels

Consider the data shown in Figure 1.8 (from the Department of Ecology at the Faculty of Biology of the University of Barcelona), which consists on the response numerical values of salinity for each sample of water that was taken in the "bahia" named the "Fangar".

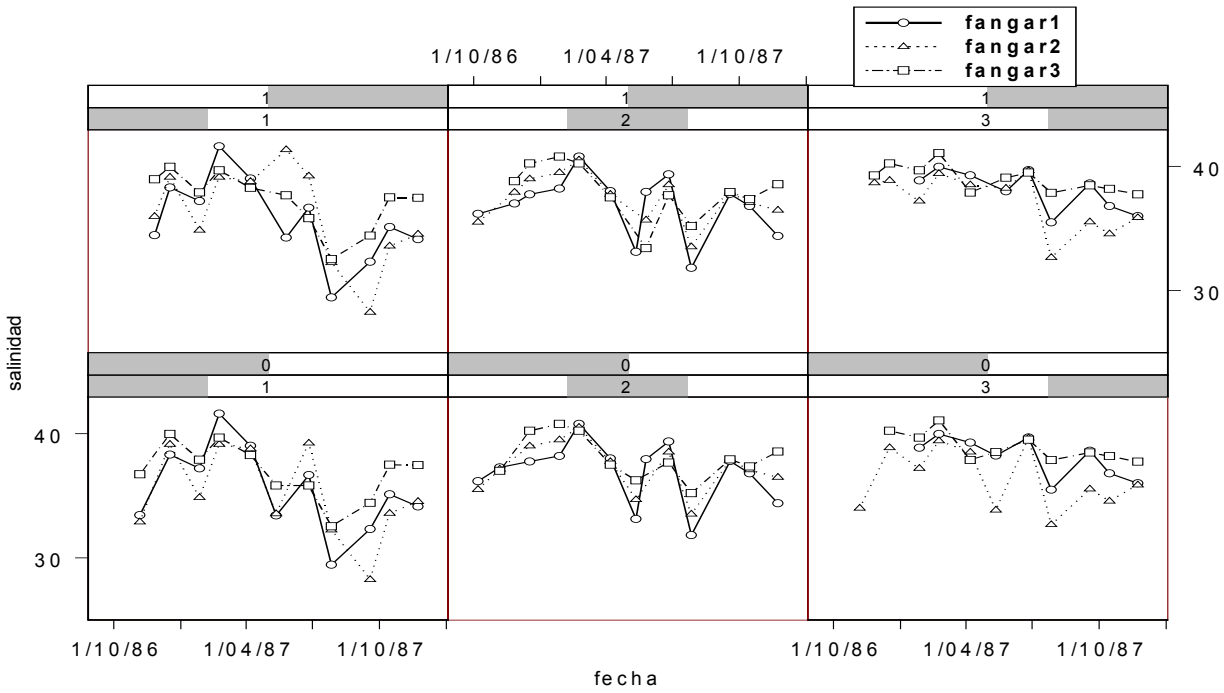


Figure 1.8: Salinity in function of time

The Fangar is situated in the delta of Ebro river. The measurements were made between 1986 and 1987. Samples were taken in three locations (stations) labelled "Fangar1", "Fangar2" and "Fangar3", and each one had three levels of profundity (1–first level, 2–second level and 3–third

level). The main interest is to model the variation of the salinity in function of time. The method used in this section applies a non-linear mixed model with two levels:

$$\begin{aligned} y_{ijk} &= \delta_0 + \delta_1 \cos(t_{ij}) + \delta_2 \sin(\delta_3 \cdot t_{ij}) + \eta_i + \eta_{ik} + e_{ijk} \\ \eta_i &\sim N(0, \sigma_1^2),\ \eta_{ik} \sim N(0, \sigma_2^2),\ e_{ijk} \sim N(0, \sigma^2) \end{aligned} \tag{1.10}$$

where y_{ijk} is the $j-th$ observation of salinity in the $i-th$ station under $k-th$ (k=1,2,3) profundity at time (date) t_{ij}. δ_0, δ_1, δ_2, and δ_3 are the fixed effects and η_i is the random effect associated to the $i-th$ station, η_{ik} denotes the random parameter associated to the profundity k for each station i. We assume that $\eta_{i,}$ η_{ik} and e_{ik} are independent for all different i, k.

Exhaustive analysis for this data, using this class of models and other techniques, is explained in the doctoral dissertation *"Ecología de la bahía del Fangar (delta del Ebro) y su relación con el cultivo de moluscos bivalvos"* inscribed in the department Ecology, Faculty of Biology University of Barcelona.

1.6. ***Discussion***

Among other methods, the most common approach to investigate the change over time in characteristics which are measured repeatedly for each unit participating in the study are multivariate models and random effects models. The multivariate model can be applied jointly with multivariate methods for missing observations (Dempster, Laird and Rubin, 1977). Because of its simplicity and connection to familiar multivariate analysis of variance techniques, this model and method are quite popular, and are often adopted by default, sometimes without proper attention to the validity of the underlying assumptions. However, when the measures are taken at arbitrary instants (e.g. not every individual has the same number of observations) or when the dimension of the dispersion matrix for each vector of responses is large, this approach becomes unattractive and difficult to apply, since a full multivariate model with an unrestricted dispersion matrix requires a profusion of variance parameters, many of which will be poorly estimated. In addition, the full multivariate model does not permit the definition and estimation of (random) individual characteristics, and takes no explicit account of where the times of observation are, chronologically. Whereas two-stage random effects models can be used easily, they are based on the explicit identification of individual and population characteristics, and their form may be naturally extended to the unbalanced situation, and also can be used for modelling complex situations as heterogeneous errors variances, removing the assumption of a common variance shared by all subjects. In our opinion, linear mixed effects (LME) models are more flexible and rich for characterizing repeated measurement data. They represent a generalization of two-stage random effects, adequate to represent the variation in the population. An extension of the linear mixed model is the generalized linear mixed model. In some cases, it may be more advantageous because it is applicable in the case where the response data are non-normal –but its distribution is known. The extension of LME to the non-linear case (NLME) is more complicated. The NLME is often more interpretable as the model parameters, generally, have a natural physical interpretation.

A major complication of NLME is to estimate the parameters of interest. While methods for solving linear random coefficients problems have been addressed in the literature and are available in many software packages, methods for non-linear random coefficients models are still being in research, and it's impossible to state any general statistical properties for non-linear mixed models in the finite sample case, even when variance parameters are known. Therefore, the only way to make any general statements on the non-linear mixed model estimation is to consider asymptotic properties of estimators.

Since the maximum likelihood estimation for NLME models leads to cumbersome integration problems, because random parameters appear inside the non-linear expectation function, several approximate methods have been proposed:

One approach is based on the first order approximation to the exact random effects from the expectation regression function. It was originally proposed by Sheiner and Beal (1980) and developed by Vonesh and Carter (1992).

Another method uses a first order approximation. It was suggested by Lindstrom and Bates (1990); they approximate non-linear functions around the means of random effects.

Wolfinger (1993) showed that the Lindstrom and Bates estimator can be obtained using the Laplace's approximation to the likelihood function.

Pocock el al. (1981) suggested an intuitively appealing two-stage estimator.

Pinheiro and Bates (1995) compared several estimators for non-linear mixed effects models via statistical simulation.

Our focus in this dissertation is to study methods for the analysis of continuous real longitudinal data such as breast cancer and others, in the first part, based on the assumption that the normal distribution is a reasonable probability model, and in the second, we relax this assumption using alternate methods such as two-stage and bootstrap methods.

Along with theoretical results, we use extensive simulation studies on a number of different models to investigate the strengths and weaknesses of the methods and to give a general comparison of them.

Chapter 2

2. Review of statistical methods

2.1. Introduction

Mixed models have been generating an increasing interest in current statistical literature in last years, because they represent a rich and powerful tool for analyzing repeated measures data. In this chapter we discus some inferential methods related to these models, with more detail.

2.2. Linear mixed model

Laird and Ware (1982) have proposed a general linear mixed model for longitudinal data,

$$y_i = X_i\delta + Z_i\eta_i + e_i \tag{2.1}$$

where y_i is an $(n_i\times 1)$ vector of responses for the $i-th$ experimental unit i=1,2,...,m, X_i is an $(n_i\times p)$ design matrix that characterizes the systematic part of the response, i.e. depending on covariates and time, δ is a $(p\times 1)$ vector of parameters corresponding to the fixed effects, that complete the characterization of the systematic part of the response, Z_i is a $(n_i\times k)$ design matrix that characterizes random variation in the response attributable to within–unit sources, η_i is a $(k\times 1)$ vector of random effects that completes the characterization of among–unit variation (k and p are not necessarily equal), and e_i is a $(n_i\times 1)$ vector of within–unit errors characterizing variation due to the way in which the responses are measured on the $i-th$ unit.

.

2.2.1. Assumptions on random variation

The model components $\eta_i(k\times 1)$ and $e_i(n_i\times 1)$ characterize the two sources of random variation, among and within units respectively. The usual assumptions are:

1. $e_i \sim N(0,\sigma^2 W_i)$, where W_i is a $(n_i \times n_i)$ covariance matrix that characterizes variance and correlation due to within–unit sources. This includes error in measuring the response and possible correlation introduced by the serial nature of data collection. The most common (and simple) choice for W_i is the model that states that variance is the same at all time points and that measurements are sufficiently far apart in time so that no correlation is induced because of data collection, i.e. $W_i = I_{n_i}$.
2. $\eta_i \sim N(0,\sigma^2 D)$, here D is a $(k\times k)$ covariance matrix that characterizes variation due to among–unit sources. The dimension of D corresponds to the number of among–unit random effects in the model. It is possible to allow D to have a particular form or to be unstructured. It is also possible to have different D matrices for different groups.
3. e_i and η_i are independent.

With these assumptions we have:

$$\begin{aligned} &E(y_i) = E(E(y_i \mid \eta_i)) = X_i\delta, \\ &\operatorname{var}(y_i) = E(\operatorname{var}(y_i \mid \eta_i)) + \operatorname{var}(E(y_i \mid \eta_i)) \\ &\qquad\qquad = \sigma^2(Z_i D Z_i' + W_i) = \sigma^2\Sigma_i, \\ &\text{where } \Sigma_i = Z_i D Z_i' + W_i \text{ and } y_i \sim N_{n_i}(X_i\delta, \sigma^2\Sigma_i) \end{aligned} \tag{2.2}$$

That is, the model with the above assumptions on e_i and η_i implies that the y_i are multivariate normal random vectors of dimension n_i with a particular form of covariance matrix. The form of Σ_i implied by the model has two distinct components, the first having to do with variation solely from among–unit sources and the second having to do with variation solely from within–unit sources.

It is sometimes more convenient to write the model for all individuals stacked in one vector as,

$$\begin{cases} y = X\delta + Z\eta + e \\ \operatorname{var}(y) = \sigma^2(ZD_\eta Z' + W) = \sigma^2\Sigma \end{cases} \tag{2.3}$$

where $y = \left[y_1^{'},...,y_m^{'}\right]^{'}$, $X^{'} = \left[X_1^{'},...,X_m^{'}\right]^{'}$, $\eta^{'} = \left[\eta_1^{'},...,\eta_m^{'}\right]^{'}$, $e^{'} = \left[e_1^{'},...,e_m^{'}\right]^{'}$

$Z = \bigoplus_{i=1}^{m} Z_i = \text{diag}(Z_1,...,Z_m)$, $W = \text{diag}(W_1,...,W_m)$, and $D_\eta = \text{diag}(D,\ldots,D)_{(km\times km)}$

2.2.2. Estimation of effects and covariance parameters

Several methods have been proposed to estimate parameters characterizing "mean", δ, and "variation", the distinct element of D_η, and the parameters that make up W. The standard frequentist approach under the assumption of multivariate normality is to use the methods of maximum likelihood and restricted maximum likelihood.

2.2.2.1. Maximum Likelihood

The basic premise of maximum likelihood is as follows. We would to estimate the parameters that characterize our model based on the data we have. One approach would be to use as the estimator a value that "best explains" the data (e.g. find the parameter values that are most likely to have given the data that were actually observed). The likelihood function is the probability density for the data given the parameters, but regarded as a function of the parameters for fixed data, instead of as a function of data for fixed parameters.

Due to the usual within-unit correlation, the likelihood function is obtained by first writing the joint probability density of the observations for each subject in matrix form. Since different subjects are independent, the joint probability density for all subjects (units) is the product of the individual probability densities. The random response variables in the joint probability distribution are replaced by the actual observations, and since the two random components, η_i and e_i, have zero means and are uncorrelated, the resulting likelihood function for all m individual, according to the formulae (2.2), is:

$$L(\delta,\sigma,\Sigma) = \prod_{i=1}^{m}(2\pi)^{-(n_i/2)}\left|\sigma^2\Sigma_i\right|^{-1/2}\exp\left\{-(y_i - X_i\delta)'(\sigma^2\Sigma_i)^{-1}(y_i - X_i\delta)/2\right\}$$

or, equivalently, the log- likelihood is

$$l(\delta,\sigma,\theta,w) = -\frac{1}{2}\left[\sum_{i=1}^{m} n_i \log(2\pi) + \log(\sigma^{2n_i}\left|\Sigma_i(\theta,w)\right|) + \sigma^{-2}\left(y_i - X_i\delta\right)' \Sigma_i^{-1}(\theta,w)\left(y_i - X_i\delta\right)\right] \quad (2.4)$$

where the intra-individual covariance parameters are represented by w and the distinct elements of the inter-individual covariance matrix D are represented by θ. Collected all variance components into a single covariance parameter vector $\vartheta=(\theta',w')'$ and maximizing l by first differentiating them with respect to σ^2 and δ and setting the result equal to zero, the following result, given ϑ, is obtained:

$$\begin{cases} \hat{\sigma}^2(\vartheta) = \sum_{i=1}^{m}(y_i - X_i\hat{\delta}(\vartheta))'\Sigma_i^{-1}(y_i - X_i\hat{\delta}(\vartheta))'/N \\ \hat{\delta}(\vartheta) = \left[\sum_{i=1}^{m}(X_i'\Sigma_i^{-1}X_i)\right]^{-1}\left[\sum_{i=1}^{m}X_i'\Sigma_i^{-1}y_i\right] \end{cases} \quad (2.5)$$

where $N = \sum_i n_i$ is the total number of observations.

Second, the ML estimation of ϑ proceeds by maximizing the profile log-likelihood function ϑ constructed by substituting $\delta = \hat{\delta}(\vartheta)$ and $\sigma^2 = \hat{\sigma}^2(\vartheta)$ into (2.4). Specifically,

$$l(\vartheta) = -\frac{1}{2}\left[\sum_{i=1}^{m} n_i \log(2\pi) + \log(\sigma^{2n_i}\left|\Sigma_i\right|) + \sigma^{-2}\left(y_i - X_i\delta\right)' \Sigma_i^{-1}\left(y_i - X_i\delta\right)\right]\Bigg|_{\delta=\hat{\delta}(\vartheta),\sigma^2=\hat{\sigma}^2(\vartheta)}. \quad (2.6)$$

The derivative of (2.6) with respect to ϑ yields the ML estimating equations for ϑ:

$$-\frac{1}{2}\operatorname{tr}\left(\Sigma^{-1}\frac{\partial\Sigma}{\partial\vartheta_j}\right) + \frac{\hat{\sigma}^{-2}}{2}\left(y - X\hat{\delta}\right)'\Sigma^{-1}\frac{\partial\Sigma}{\partial\vartheta_j}\Sigma^{-1}\left(y - X\hat{\delta}\right) = 0. \quad (2.7)$$

Equations (2.7) are non-linear in the variance components elements, ϑ_j (because depend on Σ^{-1}). Thus these components are complicated functions of the variance elements. Hence (except for the trivial cases) cannot be obtained directly. Therefore, solutions have to be obtained numerically, usually by iteration.

If we note $\hat{\vartheta}$ the ML estimate, the maximum likelihood estimates $\hat{\delta}$ and $\hat{\sigma}^2$ are then obtained by setting $\vartheta=\hat{\vartheta}$ in the equations (2.5).

For computational purposes, Pinheiro and Bates (2000) have estimated these parameters based on the orthogonal-triangular decomposition of the model matrices. This simplifies the problem of

optimizing the likelihood to get the likelihood estimates because it reduces the dimension of the optimization.

2.2.2.2. Best linear unbiased prediction

As we mentioned above, the linear mixed effect model is a subject-specific model in the sense that an individual "regression model" is characterized as having "mean" $X_i\delta+Z_i\eta_i$. Thus, if we want to characterize individual behavior in this model, we need to "estimate" both δ and η_i.

We already have discussed how to estimate δ. However, how do we "estimate" η_i?

Technically, η_i is not a fixed constant like δ, rather, it is a random effect that varies across units. Thus, when we seek for an "estimate" η_i, we try to characterize a random, not a fixed, quantity as the experimental units are randomly chosen from the population. In situations where interest focuses on characterizing a random quantity, it is customary to use a different terminology in order to preserve the notion that we are interested in something that varies. Thus, "estimation" of a random quantity is often called prediction to emphasize the fact that we are trying to get our hands on something that is not fixed and immutable, but something whose value arises in a random fashion. Thus, in order to characterize individual unit behavior, we wish to develop a method for prediction of the η_i.

In ordinary regression analysis, a prediction problem arises when one wishes to get future values of the response that might be observed; that is; it is desired to predict future y values that might be observed at certain covariate settings on the basis of the data at hand. In this case, for the value of the y at a certain covariate value, x_0 is the mean of y values that might be seen at x_0, $x_0^{'}\delta$.

As the mean is not known (because δ is not known), the usual approach is to use as the prediction the estimated mean $x_0^{'}\hat{\delta}$, where $\hat{\delta}$ is the estimate of δ.

By analogy, this may be the first approach for the prediction of η_i . However, an assumption of the model is that $\eta_i \sim N(0,\sigma^2 D)$, so that $E(\eta_i)=0$ for all i. Thus, following this logic, we would use 0 as the prediction for η_i, for any unit. This would lead to the same "estimate" for individual-specific quantities like δ_i in a random coefficient model for all units. But the whole point is that

individuals are different; thus, this tactic does not seem sensible, as it gives the same result regardless of individuals.

Additionally, this approach does not take advantage of the fact that we have some extra information available. Under model (2.1) we have $y_i=X_i\delta+Z_i\eta_i+e_i$ that is, the data y_i and the underlying random effects η_i are related. This suggests that there is information about η_i in y_i that we could exploit. More precisely, is there some function of data y_i that may be used as a predictor for η_i? Of course, this function would also be random, as it is a function of the random data y_i. If we denote by $\alpha(y_i)$ such a predictor, then one possibility would be to choose it so that the distance between $\alpha(y_i)$ and η_i, which we can measure as $(\eta_i\text{-}\alpha(y_i))^2$ is small. As both y_i and η_i are random, and hence vary in the population, we'd like the distance to be small averaged over all its possible values:

$$E(\eta_i - \alpha(y_i))^2. \tag{2.8}$$

Among all possible functions we might choose, the function $\alpha(y_i)$ minimizing this expected distance is the conditional expectation of η_i given y_i, $E(\eta_i|y_i)$.

To prove the preceding assertion, suppose that $\tilde{\eta}_i$ is some prediction of η_i. We add and subtract $E(\eta_i|y_i)$ which, for convenience, will be denoted η_i^0, i.e. $\eta_i^0 \equiv E(\eta_i \mid y_i)$:

$$\begin{aligned} E(\tilde{\eta}_i - \eta_i)'(\tilde{\eta}_i - \eta_i) &= E(\tilde{\eta}_i - \eta_i^0 + \eta_i^0 - \eta_i)'(\tilde{\eta}_i - \eta_i^0 + \eta_i^0 - \eta_i) \\ &= E(\tilde{\eta}_i - \eta_i^0)'(\tilde{\eta}_i - \eta_i^0) + 2E(\tilde{\eta}_i - \eta_i^0)'(\eta_i^0 - \eta_i) \\ &\quad + E(\eta_i^0 - \eta_i)'(\eta_i^0 - \eta_i). \end{aligned}$$

Note that the last term, $E(\eta_i^0 - \eta_i)'(\eta_i^0 - \eta_i)$, do not depends on $\tilde{\eta}_i$. The second term is null:

$$\begin{aligned} E(\tilde{\eta}_i - \eta_i^0)'(\eta_i^0 - \eta_i) &= E_{y_i}\left\{E_{\eta_i}\left[(\tilde{\eta}_i - \eta_i^0)'(\eta_i^0 - \eta_i) \mid y_i\right]\right\} \\ &= E_{y_i}\left\{\left[E_{\eta_i}(\tilde{\eta}_i) - \eta_i^0\right](\eta_i^0 - \eta_i^0)\right\} = 0. \end{aligned}$$

Therefore

$$E(\tilde{\eta}_i - \eta_i)'(\tilde{\eta}_i - \eta_i) = E(\tilde{\eta}_i - \eta_i^0)'(\tilde{\eta}_i - \eta_i^0) + \text{nonnegative terms without } \tilde{\eta}_i.$$

Thus, the minimum of this quantity can be achieved by choosing $\tilde{\eta}_i = \eta_i^0$; i.e., the best predictor is $\tilde{\eta}_i = E\left(\eta_i \middle| y_i\right)$. The expectation is itself a random quantity, it is a function of the random vector y_i. Thus, under the normality assumption

$$\begin{pmatrix}\eta_i \\ y_i\end{pmatrix} \sim N\left(\begin{pmatrix}0 \\ X_i\delta\end{pmatrix}, \sigma^2\begin{bmatrix}D & DZ_i' \\ Z_iD & \Sigma_i\end{bmatrix}\right).$$

We know that $\eta_i \mid y_i \sim N(0 + DZ_i'\Sigma_i^{-1}(y_i - X_i\delta), V_i^{-1})$ for $V_i = \sigma^{-2}\left(D - DZ_i'\Sigma_i^{-1}Z_iD\right)^{-1}$.

Thus,

$$\tilde{\eta}_i = E(\eta_i \mid y_i) = DZ_i'\Sigma_i^{-1}(y_i - X_i\delta) \ .$$

An alternative but coincident approach is based on the theory on general Bayesian linear models. Since the random effects in the model (2.1) are assumed to be random variables, it is more natural to estimate them using Bayesian techniques (see, Gelman et al. 1995). As discussed above, the distribution of the vector y_i of responses for the $i-th$ individual, conditional on that individual's specific regression coefficients η_i is multivariate normal with mean vector $X_i\delta + Z_i\eta_i$ and with covariance matrix $\sigma^2\Sigma_i$. Further, the marginal distribution of η_i is multivariate normal with mean 0 and covariance matrix, $\sigma^2 D$. In Bayesian literature, this last distribution is usually called the prior distribution of the parameters η_i. Once observed values y_i have been collected, the so-called posterior distribution η_i, defined as the distribution of η_i, conditional of y_i, can be calculated. If we denote the density function of y_i conditional on η_i, and the prior density function of η_i by $f(\eta_i|y_i)$ and $f(\eta_i)$, respectively, we have that the posterior density function of η_i given y_i is given by

$$f(\eta_i \mid y_i) = \frac{f(y_i \mid \eta_i).f(\eta_i)}{\int f(y_i \mid \eta_i).f(\eta_i)d\eta_i}.$$

Once the posterior density, $f(\eta_i|y_i)$, has been derived, it can be used for estimating η_i in any adequate way. A reasonable estimate is the expectation of the posterior density:

$$E(\eta_i \mid y_i) = \int \eta_i f(\eta_i \mid y_i)d\eta_i = \tilde{\eta}_i(\vartheta)$$

In any case, in the realistic situation where the covariance parameters ϑ are not known, one forms the "approximate" BLUP (best linear unbiased prediction) for η_i as

$$\hat{D}Z_i'\hat{\Sigma}_i^{-1}(y_i - X_i\delta) = \hat{\eta}_i(\hat{\vartheta}) \tag{2.9}$$

where $\hat{\vartheta}$ is the usual estimator of ϑ. The resulting expression is not longer exactly the BLUP, but it is an adequate approximation. It is often referred to as the estimated BLUP, or EBLUP, to acknowledge that the ϑ have been replaced by estimates.

2.2.2.3. Restricted maximum Likelihood

A widely acknowledged problem with maximum likelihood (ML) estimation has to do with the estimation of the parameters σ^2 and ϑ characterizing the covariance structure. Although the ML estimates of δ for a particular model are (approximately) unbiased, the estimates for ϑ and σ^2 are markedly biased when m is not large. For parameters that represent variances, it is usually the case that the estimated values are too small, thus giving an optimistic picture, of how variable things really are. Thompson (1962) introduced the idea of restricted maximum likelihood, REML, (also known as residual maximum likelihood) for the purpose of obtaining unbiased or less biased estimates of variances than maximum likelihood estimates. This is even more important when the number of fixed parameters is not small relative to the total number of observations. A detailed treatment of the more technical aspects may be found in Diggle, Liang, and Zegger (1994, section 4.5).

REML estimation is based on the linear combinations of elements of response y in the model (2.3), chosen in such a way that those combinations do not contain any fixed effects. These linear combinations turn out to be equivalent to residuals obtained after fitting the fixed effects. However, ϑ is now obtained from maximizing the likelihood function of a set of error contrasts chosen as $\Gamma'y=\Gamma'X\delta+\Gamma'Z\eta+\Gamma'e$, where Γ is any $(N\times(N-p))$ full rank matrix with columns orthogonal to the columns of the X, for example $\Gamma=I_N-X(X'X)^{-1}X'$, i.e. $\Gamma'X\delta=0$ for all δ, and hence $\Gamma'X=0$. As a consequence, $\Gamma'y\sim N(0,\ \Gamma'(\sigma^2\Sigma)\Gamma)$ and the log-likelihood function can be written as

$$l_R = \log(L_R) = -\frac{1}{2}(N-p)\log(2\pi) - \frac{1}{2}\log\left|(\Gamma'(\sigma^2\Sigma)\Gamma)\right| - \frac{1}{2}y'\Gamma\left[\Gamma'(\sigma^2\Sigma)\Gamma\right]^{-1}\Gamma'y.$$

Indicating $G=\sigma^2\Sigma$ we have $\Gamma[\Gamma'G\Gamma]^{-1}\Gamma'=G^{-1}-G^{-1}X(X'G^{-1}X)^{-1}X'G^{-1}$. In fact, both $\Gamma^{+}=\Gamma(\Gamma'\Gamma)^{-1}\Gamma'$ and $X^{+}=X(X'X)^{-1}X'$ are symmetric and idempotent, i.e. ($\Gamma^{+2}=\Gamma^{+}$ and $X^{+2}=X^{+}$), and $\Gamma'X=0$. Therefore $\Gamma^{+}X=0$ and $X^{+}\Gamma=0$. Hence $T=I-X^{+}-\Gamma^{+}$ is symmetric, and idempotent. Therefore

$$\begin{aligned}\mathrm{tr}(TT')=\mathrm{tr}(T^2)&=\mathrm{tr}(I)-\mathrm{tr}(X^{+})-\mathrm{tr}(\Gamma^{+})\\&=N-r_X-r_\Gamma=N-r_X-(N-r_X)\\&=0\end{aligned}$$

But T is real, so that $\mathrm{tr}(TT')=0$ implies $T=0$. Therefore $I-X^{+}=\Gamma^{+}$. Because G is assumed to be positive definite, a symmetric matrix $G^{1/2}$ always exists such that $G=(G^{1/2})^2$. Then,

since $(G^{1/2}\Gamma)'G^{-1/2}X=0$, because $\Gamma'X=0$, the preceding result applies for Γ and X replaced by $G^{1/2}\Gamma$ and $G^{1/2}X$, respectively. Making these replacements after writing $I-X^{+}=\Gamma^{+}$ as

$$I-X(X'X)^{-1}X'=\Gamma(\Gamma'\Gamma)^{-1}\Gamma'$$

gives

$$I-G^{-1/2}X(X'G^{-1}X)^{-1}X'G^{-1/2}=G^{1/2}\Gamma(\Gamma'G\Gamma)^{-1}\Gamma'G^{1/2}$$

i.e.

$$G^{-1}-G^{-1}X(X'G^{-1}X)^{-1}X'G^{-1}=\Gamma(\Gamma'G\Gamma)^{-1}\Gamma'.$$

By factorizing the exponent in the REML likelihood as

$$\begin{aligned}(y-X\delta)'G^{-1}(y-X\delta)=y'\left[G^{-1}-G^{-1}X(X'G^{-1}X)^{-1}X'G^{-1}\right]y\\+(\delta-\hat{\delta})'X'G^{-1}X(\delta-\hat{\delta})\end{aligned}$$

where $\hat{\delta}=(X'G^{-1}X)^{-1}XG^{-1}y$, the estimation is obtained by maximizing the above likelihood function. Thus the log-restricted likelihood can be obtained as,

$$l_R=\log(L_R)=l(\vartheta,\sigma^2,\hat{\delta})-\frac{1}{2}\log\left|X'G^{-1}X\right|+\frac{p}{2}\log(2\pi).$$

Additionally, from a bayesian point of view, REML can be derived by integrating out δ, using a non-informative, or flat, prior distribution. e.g.

$$y\mid\eta\sim N(X\delta+Z\eta,\sigma^2W),\ \ \eta\mid D_\eta\sim N(0,\sigma^2D_\eta),\ \delta\sim U(-\infty,+\infty).$$

As a consequence,

$$L_R = \iint L(\delta, \eta, \sigma^2, D_\eta, W \mid y) d\eta d\delta$$

where,

$$L(\delta, \eta, \sigma^2, D_\eta, W \mid y) = \frac{1}{(2\pi\sigma^2)^{\frac{N}{2}} |W|^{\frac{1}{2}}} . \exp\left\{ -\frac{1}{2\sigma^2} [y - (X\delta + Z\eta)]' W^{-1} [y - (X\delta + Z\eta)] \right\}$$
$$\times \frac{1}{(2\pi\sigma^2)^{\frac{km}{2}} |D_\eta|^{\frac{1}{2}}} \exp\left\{ -\frac{1}{2\sigma^2} \eta' D_\eta \eta \right\}.$$

2.3. Non-linear mixed effects

The non-linear mixed effect model has received substantial attention in the literature (Sheiner and Beal 1985, Lindstrom and Bates 1990, Davidian and Giltinan 1995, Vonesh and Chinchilli 1997, Pinheiro and Bates 2000, Zhiyu Ge. et al, 2003). In this section we present with more detail the general two-stage non-linear mixed model. The advantage of building a model in stages is that each stage can be relatively simple and easy to understand, while the entire model may be rather complicated, especially, when we take into account the heterogeneity of variances and/or within individual correlations.

- The first stage (individual model) consists in adjusting individual non-linear regression models,

$$\begin{gathered} y_{ij} = f(t_{ij}, \delta_{ij}) + \sigma g(E(y_{ij} \mid \delta_{ij}), z_{ij}, \tau) e_{ij} \\ \text{where} \quad i = 1, \ldots, m, \; j = 1, \ldots, n_i \end{gathered} \tag{2.10}$$

where:

- y_{ij} denotes the *j–th* response for the *i–th* individual taken at a set of conditions summarized by the vector of covariates t_i, so that a total of $N = \sum_{i=1}^{m} n_i$ responses have been observed.
- f is a known non linear function, common to all individuals, specifying the relationship between y_{ij} and t_{ij}.
- δ_{ij} is a vector whose components are the parameters of the systematic function f characterizing the *i–th* individual.

- The scalar parameter σ defines the overall level of precision in the response.
- $g(.)$ is a variance function, specifying a possible heteroscedasticity within individuals. It may depend on the mean response $E(y_{ij}|\delta_{ij})$, on the constants z_{ij} (which may include some or all of the components of t_i), and on an additional $q-$dimensional parameter vector τ, that fully specifies the functional form of the variance.

- e_{ij} is a random error term, with zero mean and a variance reflecting uncertainty in the response given the individual, and characterizing a systematic pattern of correlation. Formally, if we adopt the notation of Pinheiro and Bates (2000) that p_{ij}, $p_{ij'}$ are vector positions associated to e_{ij}, $e_{ij'}$ respectively, the correlation for the i–th individual associated with some distance, say, $d(p_{ij}, p_{ij'})$, can be expressed in the following form, $\text{cor}(e_{ij}, e_{ij'})=M(d(p_{ij}, p_{ij'}),\xi)$ where $M(.,\xi)$ is a correlation function taking values between –1 and 1, and the ξ is a vector of correlation parameters.

- The second stage (population model) consists in adjusting a more general linear model,

$$\delta_{ij} = X_{ij}\delta + Z_{ij}\eta_i \tag{2.11}$$

where X_{ij} and Z_{ij} are the "design" matrices associated with the fixed effects vector δ and random effect η_i respectively, depending on the i–th individual and possibly on the value of some covariates at the j–th observation. For example, the clearance and volume in pharmacokinetics both vary across subjects and over time, as in the IGF-I pharmacokinetic data in trauma patients cited in Davidian (1995). The random effects η_i are usually assumed to be i.i.d. with $\text{E}(\eta_i)=0$ and $\text{var}(\eta_i)=D$. The most popular assumption is that η_i are normally distributed with these moments. The second stage (2.11) allows the possibility of nonlinear dependence of the δ_{ij} on the fixed and random effects and allows the possibility that the dimensions of the random effects and the vector δ_i may not coincide.

Collecting the responses and errors for the individual into the $(n_i\times 1)$ vector,

$$y_i = \left[y_{i1},...,y_{in_i}\right]',\ f_i(\delta) = \left[f(t_{i1},\delta_{i1}),...,f(t_{in_i},\delta_{in_i})\right]',\ X_i = \left[x_{i1}',...,x_{in_i}'\right]',$$
$$Z_i = \left[Z_{i1}',...,Z_{in_i}'\right]',\ z_i = \left[z_{i1},...,z_{in_i}\right]',$$
$$W_i(\sigma,\tau) = \sigma.\mathrm{diag}\left[g(E(y_{i1} \mid \delta_{i1}),z_{i1},\tau),...,g(E(y_{in_i} \mid \delta_{in_i}),z_{in_i},\tau)\right]_{(n_i\times n_i)},$$
$$C_i(\xi) = \left[M(d(p_{ij},p_{ij'}),\xi)\right]_{jj'},\ e_i = \left[e_{i1},.....,e_{in_i}\right]'$$

we can write (2.10) and (2.11) in matrix form for the i–th individual as

$$\begin{cases} y_i = f_i(t_i,\delta_i) + e_i \\ \delta_i = X_i\delta + Z_i\eta_i,\ i = 1,...,m \\ e_i = \Lambda_i^{1/2}(\delta,\eta_i,\sigma,\tau,\xi)u_i, \text{where } u_i \text{ has zero mean and identity covariance} \\ \Lambda_i^{1/2}(\delta,\eta_i,\sigma,\tau,\xi) \text{ is the Cholesky decomposition of} \\ \Lambda_i(\delta,\eta_i,\sigma,\tau,\xi) = W_i(\delta,\eta_i,\sigma,\tau)C_i(\xi)W_i(\delta,\eta_i,\sigma,\tau) \end{cases} \quad (2.12)$$

where:

- the $(n_i\times n_i)$ diagonal matrix $W_i(\delta,\eta_i,\sigma,\tau)$ characterizes intra-individual variance; it depends on the fixed and random parameters, on the vector of parameters τ that characterizes the magnitude of variation, and on the scale parameter, σ.
- $C_i(\xi)$ is a $(n_i\times n_i)$ matrix describing the correlation pattern within the individual with total of parameter $\xi(\varsigma\times 1)$ common to all individuals.

It is sometimes more convenient to write the model for all individuals stacked in one vector with $N = \sum_{i=1}^{m} n_i$ elements, $y = \left[y_1',...,y_m'\right]'$ with this notation,

$$\begin{aligned} &y = f(t,\gamma) + e \\ &\gamma = X\delta + Z\eta \\ &e = \Lambda^{1/2}(\delta,\eta,\sigma,\tau,\xi)u \\ &\Lambda(\delta,\eta,\sigma,\tau,\xi) = \bigoplus_{i=1}^{m}\Lambda_i(\delta,\eta_i,\sigma,\tau,\xi), \\ &u = \left[u_1',...,u_m'\right]', \\ &f(t,\gamma) = \left(f_1(t_1,\gamma_1)',...,f_m(t_m,\gamma_m)'\right)', \\ &X = \left[X_1',...,X_m'\right], Z = \bigoplus_{i=1}^{m} Z_i, \eta = \left[\eta_1',...,\eta_m'\right]. \end{aligned} \quad (2.13)$$

Usually, it is assumed that the η have zero mean and common variance specified by D_η. Additionally, it is assumed that the u's have zero mean and unit block diagonal covariance matrix.

2.3.1. Estimation of effects and variance components

There are several methods for estimating the parameters in the non-linear mixed effects, but in general they are classified into two categories: exact and approximate methods. By exact methods, we mean those methods based on the exact likelihood or on the exact bayesian analysis. Both a bayesian and maximum likelihood estimation procedures often encounter significant computational difficulties with numerical integration, especially when η_i is multi-dimensional. In fact this difficulty is often so dramatic that the exact maximum likelihood procedure is seldom implemented in practice.

In order to perform maximum likelihood or bayesian analysis, one must generally evaluate a high dimensional integral, which requires numerical, usually Monte Carlo methods. For maximum likelihood estimation, we require the marginal density of y given the parameters:

$$p(y \mid \delta, D_\eta, \sigma, \tau, \xi) = \int p(y \mid \eta, \delta, \sigma, \tau, \xi) p(\eta \Big| D_\eta) d\eta. \tag{2.14}$$

Since the function of the model is non-linear in the random parameter, there is no closed expression for this integral.

Alternative approaches to nonlinear mixed effects estimation are proposed to alleviate this numerical difficulty, and they include: the EM algorithm, importance sampling, and approximate likelihood methods such as linearization algorithms, Laplacian approximation methods, and Gaussian quadrature approximation methods. The EM algorithm has successful applications in some specific classes of non-linear mixed effects models. One example was discussed by Walker (1996). In his paper he suggested Monte Carlo integration to handle conditional expectations in the E-step. For the general non-linear mixed effects models, the EM algorithm faces computational challenges, which are subject to further research, in both the E-step and the M-step. Since the EM algorithm is an attempt to get at the general MLE,

essentially it has the same difficulty that the exact maximum likelihood procedure has in this case.

Importance sampling is a common method to handle integrations numerically. When appropriate importance distributions are used and enough samples are obtained, the marginal likelihood can be approximated with high degree of numerical accuracy. However, in order to achieve satisfactory numerical stability, this method could be computationally intensive, and hence numerically less efficient than many other parametric methods. Linearization algorithms for parametric nonlinear mixed effect models are by far the most popular approaches. They are natural extensions of the linearization algorithm for classical nonlinear regression. They include the "first order algorithm" and the "conditional first order algorithm", which they are implemented in popular software packages such as NONMEM (Beal and Sheiner, 1992) and the nlme library in S-Plus and R. These methods are also available in SAS proc nlmixed. These software packages focus specifically on the case where the distributions are assumed to be normal.

2.3.1.1. Sheiner and Beal approximation

A natural mathematical approximation for the integral (2.14) is a Taylor series expansion around the mean value $E(\eta_i)=0$ applied to the model (2.12):

$$\begin{aligned} y_i &\cong f_i(\delta,\eta_i=0)+Z_i^{SB}(\delta,0)\left(\eta_i-0\right)+\Lambda_i^{1/2}(\delta,\eta_i=0,\sigma,\tau,\xi)u_i+\frac{\partial\Lambda_i^{1/2}(\delta,\eta_i=0,\sigma,\tau,\xi)}{\partial\eta_i}\left(\eta_i-0\right)u_i \\ &\approx f_i(\delta,0)+Z_i^{SB}(\delta,0)\eta_i+\Lambda_i^{1/2}(\delta,\eta_i=0,\sigma,\tau,\xi)u_i \qquad (2.15) \\ \text{where } Z_i^{SB}(\delta,0) &= \left(\left.\frac{\partial f_i}{\partial\gamma_i}\right|_{\gamma_i=X_i\delta}\right)\left(\left.\frac{\partial\gamma_i}{\partial\eta_i'}\right|_{\eta_i=0}\right)=\left(\left.\frac{\partial f_i}{\partial\gamma_i}\right|_{\gamma_i=X_i\delta}\right)Z_i\ . \end{aligned}$$

The crossproduct term involving $\eta_i u_i$ is disregarded as small relative to the leading three terms, as both η_i and u_i have mean 0. This approximation in fact corresponds to "linear" mixed effects (in the random parameters), but $Z_i^{SB}(\delta,0)$ is not a fixed design matrix, it depends on the fixed parameter δ. Additionally the fixed part of the model $f_i(\delta,0)$ is a non-linear function of δ. As a

consequence, we can write down the approximate marginal distribution of y_i, which is normal with expectation and covariance given by:

$$\begin{cases} E(y_i) \approx f_i(\delta, 0) \\ \text{cov}(y_i) \approx \Lambda_i(\delta, \eta_i = 0, \sigma, \tau, \xi) + Z_i^{SB}(\delta, 0) D Z_i^{SB}(\delta, 0)' = \Sigma_{SB,i} . \end{cases} \tag{2.16}$$

Under the usual assumptions of normality of the conditional distributions $y_i|\eta_i$ and $\eta_i \sim N(0,D)$, it follows from (2.15) that the approximate marginal distribution for y_i well be normal with mean and variance as in (2.16). Under these conditions, then, one approach would be to approximate the integral (2.14) by the normal density with these moments.

An important feature of (2.16) is that the approximate marginal covariance matrix depends on the fixed effects, δ. These features complicate the estimation of parameters based on the ML or REML. That is, the differentiation of the approximate marginal log-likelihood would result in a quadratic estimating equation for δ and a quadratic of τ and ξ. To avoid this problem, Bates and Watts (1980) suggest ignoring this dependence in differentiating the likelihood. The arguments of these authors are based on the fact that the space spanned by the columns of $Z_i^{SB}(\delta,0)$ do not depends on the parameter-effects curvature in the tangent plane. Therefore, $Z_i^{SB}(\delta,0)$ may be assumed to vary slowly with respect to δ.

Let $w=(\tau',\sigma,\xi',\theta')'$ denote a vector of parameters consisting of the intra-individual covariance parameter (τ',σ,ξ') and the distinct elements of D, θ'. Sheiner and Beal (1980) numerically optimize the likelihood function associated to (2.16), over δ and w, using their well-known first order method. Alternatively, the log-likelihood function

$$l(\delta, w) \cong -\frac{1}{2}\left[N \log(2\pi) + \sum_{i=1}^{m} \log\left|\Sigma_{SB,i}\right| + \sum_{i=1}^{m} \left(y_i - f_i\left(\delta, \eta_i = 0\right)\right)' \Sigma_{SB,i}^{-1} \left(y_i - f_i\left(\delta, \eta_i = 0\right)\right) \right]$$

may be used to derive the iterative zero expansion estimation procedure for δ, η and w. The strategy is to fix w to estimate δ by $\hat{\delta}_{SB}(w)$ which solves, for a complete model, the equation

$$A_{SB}' \Sigma_{SB}^{-1} (y - f(\hat{\delta}_{SB}(w), 0)) = 0, \text{ where } A_{SB} = \left.\frac{\partial f(\delta, 0)}{\partial \delta}\right|_{\delta = \hat{\delta}_{SB}(w)} .$$

Equivalently,

$$\hat{\delta}_{SB}(w) = (A_{SB}^{'}\Sigma_{SB}^{-1}A_{SB})^{-1}A_{SB}^{'}\Sigma_{SB}^{-1}Y_{SB}$$
$$\text{where } Y_{SB} = y - f(\hat{\delta}_{SB}(w),0) + A_{SB}\hat{\delta}_{SB}(w). \tag{2.17}$$

The next use the estimate of δ to derive an updated estimate for w, which yields the ML equations for w, constructed by substituting $\delta = \hat{\delta}_{SB}(w)$ into $l(\delta, w)$:

$$-\frac{1}{2}\mathrm{tr}(\Sigma_{SB}^{-1}\frac{\partial\Sigma_{SB}}{\partial w_j}) + \frac{1}{2}(Y_{SB} - A_{SB}\hat{\delta}_{SB}(w))^{'}\Sigma_{SB}^{-1}\frac{\partial\Sigma_{SB}}{\partial w_j}\Sigma_{SB}^{-1}(Y_{SB} - A_{SB}\hat{\delta}_{SB}(w)) = 0. \tag{2.18}$$

The dependence of Z^{SB} on w also is ignored in calculating $\frac{\partial\Sigma_{SB}}{\partial w_j}$, so

$$\frac{\partial\Sigma_{SB}}{\partial w_j} = Z^{SB}\frac{\partial D}{\partial w_j}Z^{SB'} + \frac{\partial\Lambda_{SB}}{\partial w_j},$$
$$\text{where } w_j \text{ is the } j-th \text{ element of } \Sigma_{SB}.$$

Solving (2.17) and (2.18) iteratively until convergence yields the ML zero expansion estimates which may be viewed as approximate ML estimates of (δ,w). Assuming $\hat{\delta}$ and $\hat{w}$ are set equal to their final estimates, η is predicted by the BLUP formula as in the linear mixed model,

$$\hat{\eta}_{SB} = D_{\eta}Z^{SB'}\Sigma_{SB}^{-1}(Y_{SB} - A_{SB}\hat{\delta}_{SB}).$$

In order to take into account the loss of degrees of freedom when estimating δ_{SB}, the residual profile likelihood and the corresponding likelihood equations for w are often preferred in practice (Patterson and Thompson, 1971, Harville 1977).Specifically, the REML version is:

$$l_R(w) = -\frac{1}{2}\left[N - rank(A_{SB})\right]\log(2\pi) - \frac{1}{2}\log\left|A_{SB}^{'}\Sigma_{SB}^{-1}A_{SB}\right| - \frac{1}{2}\log\left|\Sigma_{SB}\right|$$
$$-\frac{1}{2}(Y_{SB} - A_{SB}\delta)^{'}\Sigma_{SB}^{-1}(Y_{SB} - A_{SB}\delta)\,|_{\delta=\hat{\delta}_{SB}(w)} \quad . \tag{2.19}$$

Assuming that the derivative matrix A_{SB} varies slowly with respect to (δ,w) (Bates and Watts, 1980), the derivative of (2.19) with respect to w leads to the REML estimating equations for w:

$$-\frac{1}{2}tr(\Gamma_{SB}\frac{\partial\Sigma_{SB}}{\partial w_j}) + \frac{1}{2}(Y_{SB} - A_{SB}\hat{\delta}_{SB}(w))^{'}\Sigma_{SB}^{-1}\frac{\partial\Sigma_{SB}}{\partial w_j}\Sigma_{SB}^{-1}(Y_{SB} - A_{SB}\hat{\delta}_{SB}(w)) = 0 \tag{2.20}$$

where $\Gamma_{SB} = \Sigma_{SB}^{-1} - \Sigma_{SB}^{-1}A_{SB}(A_{SB}^{'}\Sigma_{SB}^{-1}A_{SB})^{-1}A_{SB}^{'}\Sigma_{SB}^{-1}$. The REML method alternates between solving (2.17) and (2.20). To take into account the derivative matrices of fixed parameters, Vonesh and

Carter (1992) advocate for the use of generalized least squares methods for inference as an alternative to Maximum (Restricted) likelihood.

2.3.1.2. Lindstrom and Bates approximation

The Lindstrom and Bates algorithm also uses a first order Taylor expansion, but around the current estimates of the random effects, $\eta_i = \tilde{\eta}_i$ and an estimate of $\delta = \tilde{\delta}$. Lindstrom and Bates argued that this might result in a more accurate approximation. These estimates are based on minimizing the penalized least square function given below. These authors assume that the inter-individual regression function is linear $\left(\gamma_i = X_i\delta + Z_i\eta_i\right)$, and also the intra-individual covariance matrix does not depend on γ_i (and hence on η_i), but rather depends on i only through its dimension. Davidian (1995), has generalized these two assumptions to account for more general intra-individual covariance structure, and a non-linear regression inter individual function, the model with general covariance matrix based on Davidian (1995) is:

$$
\begin{aligned}
y_i &\approx f_i(\delta,\tilde{\eta}_i) + Z_i^{LB}(\delta,\tilde{\eta}_i)(\eta_i - \tilde{\eta}_i) + \frac{\partial}{\partial\eta_i}\left\{\Lambda_i^{1/2}(\delta,\tilde{\eta}_i,w)\right\}\left(\eta_i - \tilde{\eta}_i\right)u_i + \Lambda_i^{1/2}(\delta,\tilde{\eta}_i,w)u_i \\
&= \left\{f_i(\delta,\tilde{\eta}_i) - Z_i^{LB}(\delta,\tilde{\eta}_i)\tilde{\eta}_i\right\} + Z_i^{LB}(\delta,\tilde{\eta}_i)\eta_i + \Lambda_i^{1/2}(\delta,\tilde{\eta}_i,w)u_i \\
&\text{where,} \\
&u_i \sim (0, \mathrm{I}_{n_i}) \\
&Z_i^{LB}(\delta,\tilde{\eta}_i) = \left.\frac{\partial f_i(\delta,\eta_i)}{\partial\eta_i}\right|_{\tilde{\eta}_i}.
\end{aligned}
\tag{2.21}
$$

By analogy to the step above, the cross-product term including $\left(\eta_i - \tilde{\eta}_i\right)u_i$ has been disregarded as negligible.

Treating $\tilde{\eta}_i$ as a fixed constant, (2.21) suggests the following approximation marginal moments:

$$
E(y_i) \approx f_i(\delta,\tilde{\eta}_i) - Z_i^{LB}(\delta,\tilde{\eta}_i)\tilde{\eta}_i, \quad \mathrm{var}(y_i) \approx Z_i^{LB}(\delta,\tilde{\eta}_i)DZ_i^{LB}(\delta,\tilde{\eta}_i)' + \Lambda_i(\delta,\tilde{\eta}_i,w).
$$

If we note that the matrices $Z_i^{LB}(\tilde{\delta},\tilde{\eta}_i)$ and $\Lambda_i(\tilde{\delta},\tilde{\eta}_i,w)$ are constant with respect to δ, the above expressions substituting δ by $\tilde{\delta}$,define a linear mixed model. Moreover, by a further approximation to the marginal mean, expanding $f_i(\delta,\tilde{\eta}_i)$ and $Z_i^{LB}(\delta,\tilde{\eta}_i)$ up to the linear term around $\tilde{\delta}$ and ignoring "negligible" terms, we obtain

$$
\begin{aligned}
E(y_i) &\approx f_i(\tilde{\delta},\tilde{\eta}_i) + A_{LB}(\tilde{\delta},\tilde{\eta}_i)(\delta-\tilde{\delta}) - Z_i^{LB}(\tilde{\delta},\tilde{\eta}_i)\tilde{\eta}_i \\
&= f_i(\tilde{\delta},\tilde{\eta}_i) - A_{LB}(\tilde{\delta},\tilde{\eta}_i)\tilde{\delta} - Z_i^{LB}(\tilde{\delta},\tilde{\eta}_i)\tilde{\eta}_i + A_{LB}(\tilde{\delta},\tilde{\eta}_i)\delta
\end{aligned}
$$

where $A_{LB}(\delta,\eta_i) = \dfrac{\partial}{\partial\delta}\left(f_i\left(\delta,\eta_i\right)\right)$.

Note that if we define the "pseudo-response vector"

$$
\begin{aligned}
s_i &= y_i - f_i(\tilde{\delta},\tilde{\eta}_i) + A_{LB}(\tilde{\delta},\tilde{\eta}_i)\tilde{\delta} + Z_i^{LB}(\tilde{\delta},\tilde{\eta}_i)\tilde{\eta}_i \\
&\cong A_{LB}(\tilde{\delta},\tilde{\eta}_i)\tilde{\delta} + Z_i^{LB}(\tilde{\delta},\tilde{\eta}_i)\tilde{\eta}_i + \tilde{e}_i, \text{ where } \tilde{e}_i = y_i - f_i\left(\tilde{\delta},\tilde{\eta}_i\right)
\end{aligned}
\tag{2.22}
$$

then, approximately, we have

$$
E(s_i) \cong A_{LB}(\tilde{\delta},\tilde{\eta}_i)\delta.
$$

The approximate model for the expectation of s_i is linear in δ; moreover, from above,

$$
\mathrm{var}(s_i) \approx Z_i^{LB}(\tilde{\delta},\tilde{\eta}_i)DZ_i^{LB}(\tilde{\delta},\tilde{\eta}_i)' + \Lambda_i(\tilde{\delta},\tilde{\eta}_i,w).
$$

Together, both expressions for the mean and the variance define an approximate linear mixed effects model with "constant" design matrices

$$
A_{LB}(\tilde{\delta},\tilde{\eta}_i) \text{ and } Z_i^{LB}(\tilde{\delta},\tilde{\eta}_i).
$$

This model may thus be fitted by techniques for linear mixed effects models. Under the normality assumption, we have a "pseudo-response vector" or working vector, which is approximated by an LME model with regression matrices $A_{LB}(\tilde{\delta},\tilde{\eta}_i)$ and $Z_i^{LB}(\tilde{\delta},\tilde{\eta}_i)$. We have

$$
s_i \sim N\left(A_{LB}(\tilde{\delta},\tilde{\eta}_i)\delta,\ \Lambda_i(\tilde{\delta},\tilde{\eta}_i,w) + Z_i^{LB}(\tilde{\delta},\tilde{\eta}_i)DZ_i^{LB}(\tilde{\delta},\tilde{\eta}_i)'\right) \tag{2.23}
$$

and then the problem is converted to that of fitting a "linear mixed model" of the type discussed in Laird and Ware (1982).

Based on the same strategy used in the Sheiner and Beal approximation, ignoring the dependencies on the δ, the algorithm proposed by Lindstrom and Bates alternates between two steps:

1. ***Penalized non-linear least squares***:

Given the current estimate of w, $\widehat{w}$, maximize the log likelihood with respect to δ and η:

$$l(\delta,\eta \mid \widehat{w}) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\left\{\sum_{i=1}^{m}\log\left|\hat{D}\right| + \eta_i^{'}\hat{D}^{-1}\eta_i^{'} + \log\left|\hat{\Lambda}_i(\delta,\eta_i)\right| + \left[y_i - f_i(\delta,\eta_i)\right]^{'}\hat{\Lambda}_i^{-1}(\delta,\eta_i)\left[y_i - f_i(\delta,\eta_i)\right]^{'}\right\}.$$

This is accomplished simultaneously by specifying an augmented non-linear least squares problem for "data" $\tilde{y}_i = \begin{bmatrix}\Lambda_i^{-1/2}y_i \\ 0\end{bmatrix}$ and a non-linear regression function $\tilde{f}_i(\delta,\eta_i) = \begin{bmatrix}\Lambda_i^{-1/2}f(\delta,\eta_i) \\ D^{-1/2}\eta_i\end{bmatrix}, i = 1,2,...,m.$ The objective function can be converted in the standard non-linear problem, which can be maximized by minimizing:

$$\sum_{i=1}^{m}(\tilde{y}_i - \tilde{f}_i(\delta,\eta_i))^{'}(\tilde{y}_i - \tilde{f}_i(\delta,\eta_i)).$$

2. ***Linear mixed effect***:

Given the "estimates" of random and fixed effects η_i and δ, derived in the previous step, fit the linear mixed–effects model (2.23) to the pseudo-observation. Estimate δ and w by maximizing the log-likelihood function

$$l(\delta,w) = -\frac{N}{2}\log(2\pi) - \frac{1}{2}\left\{\sum\nolimits_{i=1}^{m}\log\left|\Sigma_i^{LB}\right| + \left(s_i - A_{LB}\delta\right)^{'}\Sigma_i^{LB^{-1}}\left(s_i - A_{LB}\delta\right)\right\}$$

or by maximizing the REML,

$$l_R(\delta,w) = l(\delta,w) - \frac{1}{2}\log\left|A_{LB}^{'}\Sigma^{LB^{-1}}A_{LB}\right| + \frac{p}{2}\log(2\pi) \quad .$$

This step is accomplished by a differentiation process analogous to the previous approximation (Beal and Sheiner) regarding that $\mathrm{cov}(y_i)$ is constant with respect to δ, or via the GLS method (Vonesh Carter, 1992) for the sake of taking into account the possible dependence of the covariances on fixed parameters (Davidian, 1995).

2.3.1.3. STS- Standard two stage approximation

When fitting the model (2.10)-(2.11) the interest usually focuses on estimation on the δ and D parameters characterizing the mean values of γ_i and the random inter-individual variation among them. When the number of observations n_i on each individual is sufficient to allow

credible estimation of the individual regression parameters γ_i, Pocock et al. (1981) proposed a two-stage procedure to estimate parameters δ:

i. Individual estimates of the γ_i are obtained via usual regression methods (non-linear least squares), "pooling" residuals over i to derive estimates $\hat{\sigma}$ and $\hat{\tau}$ of the common σ and τ.

ii. The individual estimates are taken as "data" (Steimer et al. 1984, Davidian and Giltinan 1995) to make inferences on δ and D by fitting the approximate linear regression model $\hat{\gamma}_i = X_i\delta + \eta_i$ by generalized least squares (GLS):

$$\hat{\delta}_{STS} = \left[\sum_{i=1}^{m}(X_i'X_i)\right]^{-1}\left[\sum_{i=1}^{m}(X_i'X_i)\tilde{\gamma}_i\right] \text{ and}$$
$$\hat{D}_{STS} = (m-1)^{-1}\sum_{i=1}^{m}(\tilde{\gamma}_i - X_i\hat{\delta}_{STS})(\tilde{\gamma}_i - X_i\hat{\delta}_{STS})'.$$

These estimates have the advantage that they do not require the assumption of normality, but because of the failure in taking into account the uncertainty in estimating γ_i, the estimate of D is markedly biased. For example, in the simple cases when $X_i = \mathrm{I}_{n_i}$ the coefficients vectors, $\hat{\gamma}_i$ are averaged over individuals to obtain the estimates $\hat{\delta}_{STS} = m^{-1}\sum_{i=1}^{m}\tilde{\gamma}_i$ and $\hat{D}_{STS} = (m-1)^{-1}\sum_{i=1}^{m}\left(\hat{\gamma}_i - \hat{\delta}_{STS}\right)\left(\hat{\gamma}_i - \hat{\delta}_{STS}\right)'$. For this simple case Davidian and Giltinan(1995) prove that:

$$E(\hat{D}_{STS}) = D + m^{-1}\sigma^2\sum_{i=1}^{m}\Omega_i$$

where $\sigma^2\Omega_i$ is the asymptotic covariance matrix of the individual estimate $\hat{\gamma}_i$.

2.3.1.4. GTS- Global two stage approximation

To avoid the problem of uncertainty in the estimation of γ_i, it is assumed that the approximate (for large n_i) distribution of $\hat{\gamma}_i$ conditional on $\gamma_1,\ldots,\gamma_m$ is normal with mean γ_i and covariance matrix Δ_i

$$\hat{\gamma}_i \big| \gamma_i \sim \mathrm{N}(\gamma_i, \Delta_i) \tag{2.24}$$

where Δ_i must be estimated (with substituted $\tilde{\gamma}_i$, $\hat{\sigma}$ and $\hat{\tau}$). So we may write $E(\hat{\gamma}_i \big| \gamma_i) \approx \gamma_i$ and $\text{var}(\hat{\gamma}_i \big| \gamma_i) \approx \Delta_i$.

Thus, at the second stage, δ and D may be estimated by standard techniques applied to this approximate model, assuming independent "data", e.g. by fitting a linear mixed model $\hat{\gamma}_i \approx \gamma_i + v_i = X_i\delta + \eta_i + v_i$ where v_i has zero mean and covariance matrix Δ_i. The assumption (2.24) is justified by theorem 2.1 from Carrol and Ruppert (1988) when common methods such as generalized least squares are used to obtain individual $\hat{\gamma}_i$. We have therefore $\hat{\gamma}_i \sim \text{N}(X_i\delta, \Delta_i + D)$. Thus, we have

$$E(\hat{\gamma}_i) = E\left\{E(\hat{\gamma}_i \big| \gamma_i)\right\} \approx X_i\delta$$

and

$$\begin{aligned}\text{var}(\hat{\gamma}_i) = \text{var}\left\{E\left(\hat{\gamma}_i \big| \gamma_i\right)\right\} + E\left\{\text{var}\left(\hat{\gamma}_i \big| \gamma_i\right)\right\} &\approx \text{var}(\gamma_i) + E(\Delta_i) \\ &\approx D + \Delta_i.\end{aligned}$$

As a consequence, the inference on δ and D can be carried out using the usual expressions for a linear mixed effects model or via the EM (Expectation-Maximization) algorithm based on the maximum likelihood applied to "data" $\hat{\gamma}_i$. The EM algorithm is a computational method to maximize an objective function. In this specific case can be stated as follows:

(a) Obtain starting values as

$$\hat{\delta}_{(0)} = m^{-1}\sum_{i=1}^{m}\hat{\gamma}_i, \;\; \hat{D}^{(0)} = (m-1)^{-1}\sum_{i=1}^{m}(\hat{\gamma}_i - X_i\hat{\delta}_{(0)})(\hat{\gamma}_i - X_i\hat{\delta}_{(0)})'$$

set $k = 0$.

(b) E- step: At iteration $(k+1)$ compute

$$\hat{\gamma}_{i,k+1} = \left(\Delta_i^{-1} + \hat{D}_k^{-1}\right)^{-1}\left[\Delta_i^{-1}\hat{\gamma}_i + \hat{D}_k^{-1}X_i\hat{\delta}_k\right] \quad .$$

(c) M- step: Obtain updated estimates as

$$\begin{cases}\hat{\delta}_{k+1} = \sum\nolimits_{i=1}^{m} W_{i,k}\hat{\gamma}_{i,k+1}, \text{ where } W_{i,k} = \left(\sum\limits_{i=1}^{m} X_i'\hat{D}_k^{-1}X_i\right)^{-1} X_i'\hat{D}_k^{-1} \\ \hat{D}_{k+1} = m^{-1}\sum\limits_{i=1}^{m}(\Delta_i^{-1} + \hat{D}_k^{-1})^{-1} + m^{-1}\sum\limits_{i=1}^{m}(\hat{\gamma}_{i,k+1} - X_i\hat{\delta}_{k+1})(\hat{\gamma}_{i,k+1} - X_i\hat{\delta}_{k+1})'\end{cases}$$

set $k = k + 1$ and return to (b).

The algorithm is iterated until convergence. We notice that $\hat{D}_{GTS}$ will be a biased estimator for D (Davidian and Giltinan, 1995).

2.3.1.5. Laplace approximation

A potential drawback of the two-stage approach is that some or all individuals may not have sufficient data to support individual estimation of γ_i. Moreover, the simplicity of the second stage estimation of δ and D that follows from the assumption of approximate (large-n_i) independence of the $\tilde{\gamma}_i$ is only relevant when scale variance σ is "small". An alternative strategy that may bypass these difficulties is to base inference on δ and D in model (2.21)-(2.23) by approximating the marginal likelihood contribution of the data $y_i = (y_1, \ldots, y_{n_i})'$ on individual i directly.

Generally speaking, a Laplace approximation is defined as follows:

Consider the integral $\mathrm{I} = \int h(x)dx$, where $h(x) = \exp(l(x))$ and $l(x)$ is a real valued function of x. Let $\hat{x}$ be the unique minimum of l. Then,

$$l'(\hat{x}) = 0 \text{ and } \quad l(x) \cong l(\hat{x}) + \frac{1}{2}(x - \hat{x})' l''(\hat{x})(x - \hat{x}).$$

Using this approximation, we have:

$$\mathrm{I} = e^{l(\hat{x})} \int \exp\left[\frac{1}{2}(x - \hat{x})' l''(\hat{x})(x - \hat{x})\right] dx$$

and the integrand can be rewritten as the density of a normal random variable, to obtain:

$$\mathrm{I} \cong (2\pi)^{k/2} \exp\left[l(\hat{x})\right] \left| -l''(\hat{x}) \right|^{-1/2}.$$

This is the Laplace approximation to the integral I.

Wolfinger (1993) and Vonesh (1996) discuss how the Laplace approximation may be applied in nonlinear mixed-effects models when $p(y_i | \eta_i, \delta)$ and $p(\eta_i | D)$ are normal densities. Both authors consider the specific situation in which $\Lambda_i(\delta, \eta_i, , \sigma, \tau', \xi')$ does not depend on δ_i, which we write

as $\Lambda_i(\sigma,\tau',\xi')$. In this case, and for the non-linear mixed model given in (2.12) we have the marginal density of the response $y_i(n_i\times 1)$ in the obvious notation:

$$p(y_i \mid \delta, D, \omega) = \int p(y_i \mid \eta_i, \delta, \omega) p(\eta_i \mid D) d\eta_i \tag{2.25}$$

where

$$\begin{cases} p(y_i \mid \eta_i, \delta, \omega) p(\eta_i \mid D) = (2\pi)^{-n_i/2} (2\pi)^{-k/2} \left|\Lambda_i(\omega)\right|^{-1/2} |D|^{-1/2} \\ \exp(-\frac{1}{2}\left[y_i - f_i(\delta,\eta_i)\right]' \Lambda_i^{-1}(\omega)\left[y_i - f_i(\delta,\eta_i)\right] - \frac{1}{2}\eta_i' D^{-1}\eta_i) \\ \omega = (\sigma,\tau',\xi')' \quad . \qquad\qquad\qquad\qquad\qquad\qquad . \end{cases}$$

Consider approximating the integral in (2.25) by the Laplace approximation to the integral. We may identify

$$l(\eta_i) = -\frac{1}{2}\left\{\left(y_i - f_i(\delta,\eta_i)\right)' \Lambda_i^{-1}(\omega)\left(y_i - f_i(\delta,\eta_i)\right) + \eta_i' D^{-1}\eta_i\right\}$$

Then, using results for matrix differentiation, it is straightforward to show that

$$l'(\eta_i) = \frac{\partial}{\partial \eta_i} l(\eta_i) = A_{LP,i}'(\delta,\eta_i)\Lambda_i^{-1}(\omega)\left[y_i - f_i(\delta,\eta_i)\right] - D^{-1}\eta_i$$

and

$$l''(\eta_i) = -D^{-1} - A_{i,LP}'(\delta,\eta_i)\Lambda_i^{-1}(\omega)A_{i,LP}(\delta,\eta_i) + \frac{\partial}{\partial \eta_i'}\left\{A_{i,LP}'(\delta,\eta_i)\Lambda_i^{-1}(\omega)(y_i - f_i(\delta,\eta_i))\right\}$$

where $A_{i,LP} = \dfrac{\partial f_i(\delta,\eta_i)}{\partial \eta_i}$. The last term in this equality has null conditional expectation. Disregarding this term and replacing $l''(\eta_i)$ by its conditional expectation on the right hand side of the above Laplace approximation, we obtain

$$\begin{aligned} p(y_i \mid \delta, D, \omega) \cong\ & (2\pi)^{-n_i/2} (2\pi)^{-k/2} \left|\Lambda_i(\omega)\right|^{-1/2} |D|^{-1/2} (2\pi)^{k/2} \left|D^{-1} + A_{i,LP}'(\delta,\tilde{\eta}_i)\Lambda_i^{-1}(\omega)A_{i,LP}(\delta,\tilde{\eta}_i)\right|^{-1/2} \\ & \times \exp\left\{-\frac{1}{2}\left[\left(y_i - f_i(\delta,\tilde{\eta}_i)\right)' \Lambda_i^{-1}\left(y_i - f_i(\delta,\tilde{\eta}_i)\right)\right] - \frac{1}{2}\tilde{\eta}_i' D^{-1}\tilde{\eta}_i\right\} \end{aligned}$$

where $\tilde{\eta}_i$ maximizing $l(\eta_i)$ must satisfy $\tilde{\eta}_i = DA_{i,LP}'\Lambda_i^{-1}(\omega)\left[y_i - f_i(\delta,\tilde{\eta}_i)\right]$.

Writing $K_i(\delta,\eta_i) = f_i(\delta,\eta_i) - A_{i,LP}(\delta,\eta_i)\eta_i$, the density $p(y_i \mid \delta, D, \omega)$ may be rewritten as

$$p(y_i|\delta, D, \omega) \cong (2\pi)^{-n_i/2} \left|\Lambda_i(\omega) + A_{i,LP}(\delta, \tilde{\eta}_i) D A'_{i,LP}(\delta, \tilde{\eta}_i)\right|^{-1/2}$$
$$\times \exp\left\{-\frac{1}{2}\left(y_i - K_i\left(\delta, \tilde{\eta}_i\right)\right)' \left[\Lambda_i(w) + A_{i,LP}(\delta, \tilde{\eta}_i) D A'_{i,LP}(\delta, \tilde{\eta}_i)\right]^{-1} \left(y_i - K_i\left(\delta, \tilde{\eta}_i\right)\right)\right\}$$

which has the form of a normal density with mean

$$K_i(\delta, \tilde{\eta}_i) = f_i(\delta, \tilde{\eta}_i) - A'_{i,LP}(\delta, \tilde{\eta}_i)\tilde{\eta}_i$$

and covariance matrix

$$\Lambda_i(\omega) + A_{i,LP}(\delta, \tilde{\eta}_i) D A'_{i,LP}(\delta, \tilde{\eta}_i).$$

According to the inter-individual independence assumption, the joint density is finally given by

$$p(y|\delta, D, \omega) = \prod_{i=1}^{m} p(y_i|\delta, D, \omega). \tag{2.26}$$

For fixed w, the estimation $\hat{\delta}(\omega)$ for δ maximizing (2.26) is obtained first, and later the approximate profile likelihood of ω constructed by substituting $\delta = \hat{\delta}(\omega)$ in (2.26) is maximized, using the same strategy used in the approximation of Sheiner and Beal.

2.4. Multilevel non-linear mixed models

Several clinical, social and educational data, to cite some application fields, have a hierarchical organization in which units at one level are grouped within units at the next levels. An appropriate way to understand the relationship between measurements made on the different levels, and variability at different grouping levels, is by using multilevel modeling.

In the non-linear mixed model, multiple nested levels represent a natural extension for the single level (individual-subject). They have recently received a great deal of attention in statistical literature, due to their flexibility in modeling complex data structures. This is based on a modified model for random effects (Pinheiro, 2000, Goldestein, 1995).

The multilevel version of model (2.10)-(2.11) for two levels of nesting is written as:

$$\begin{cases} y_{ilj} = f(t_{ilj}, \gamma_{ilj}) + \sigma g(E(y_{ilj} \mid \gamma_{ilj}), z_{ilj}, \tau).u_{ilj} \\ i = 1,...,m,\ l = 1,\ldots,m_i,\ j = 1,...,n_{il} \\ \gamma_{ilj} = X_{ilj}\delta + Z_{(i)lj}\eta_i + Z_{(il)j}\eta_{il} \\ \eta_i \sim \mathrm{N}(0, \sigma^2\Sigma_1), \eta_{il} \sim \mathrm{N}(0, \sigma^2\Sigma_2) \quad . \end{cases} \tag{2.27}$$

where we assume that there are m first-level, m_i second-level groups within the *i-th* first-level group, and n_{il} repeated measures on the *l-th* second-level group within the *i-th* first-level group. The first-level random effects η_i are independently distributed with covariance matrix $\sigma^2\Sigma_1$. The second-level random effects, η_{il}, are independently distributed with covariance $\sigma^2\Sigma_2$ and assumed to be independent of the first level random effects. Sometimes it is more convenient to rewrite (2.27) in vectorial form as:

$$\begin{aligned} & y_{il} = f(t_{il}, \gamma_{il}(\delta, \eta_i, \eta_{il})) + \sigma\Lambda_{il}^{1/2}(\delta, \eta_i, \eta_{il}, \tau)u_{il} \\ & i = 1,...,m,\ l = 1,\ldots,m_i \\ & \gamma_{il}(\delta, \eta_i, \eta_{il}) = X_{il}\delta + Z_i\eta_i + Z_{il}\eta_{il} \\ & \eta_i \sim N(0, \sigma^2\Sigma_1),\ \eta_{il} \sim N(0, \sigma^2\Sigma_2) \\ & y_{il} = \left(y_{il1},\ldots,y_{iln_i}\right)',\ f_{il}(t_{il}, \gamma_{il}) = \left(f(t_{il1}, \gamma_{il1}),\ldots,f(t_{iln_i}, \gamma_{iln_i})\right)' \\ & X_{il} = \left(X'_{il1},\ldots,X'_{iln_i}\right)',\ Z_{(i)l} = \left(Z'_{(i)l1},\ldots,Z'_{(i)ln_i}\right)',\ Z_{il} = \left(Z'_{il1},\ldots,Z'_{iln_i}\right)' \\ & u_{il} = \left(u_{il1},\ldots,u_{il\,n_i}\right)'. \end{aligned} \tag{2.28}$$

In general, we put no constraints on Σ_1 and Σ_2 other than assuming that they are variance-covariance matrices. However, in specific model building, much of the work lies in specifying structures for these matrices. For reasons of computational stability and speed, as well as parsimony, it is advantageous to reduce the number of parameters estimated in these matrices by setting certain covariances to zero. However, given the structure of the model, there is often no a priori reason to impose independence conditions on the random effects operating on a given level.

2.4.1. Likelihood function

The likelihood function for the multilevel case can be calculated in the same way as in the single level case. For the first level random effects, η_i, and all the second–level random effects, η_{il},

pertaining to first level i, the likelihood function is defined integrating over both levels of random effects as follows,

$$L(\delta,\sigma^2,\Sigma_1,\Sigma_2,\Lambda \mid y) = p(y \mid \delta,\sigma^2,\Sigma_1,\Sigma_2,\Lambda) = \prod_{i=1}^{m} p(y_i \mid \delta,\sigma^2,\Sigma_1,\Sigma_2,\Lambda_i) \tag{2.29}$$

Where the conditional density of y_i is multivariate normal

$$p(y_i \mid \delta,\sigma^2,\Sigma_1,\Sigma_2,\Lambda_i) = \int \prod_{l=1}^{m_i} A_{il}.p(\eta_i \mid \sigma^2,\Sigma_1)d\eta_i$$
$$\text{where, } A_{il} = \int p(y_{il} \mid \eta_i,\eta_{il},\delta,\sigma^2,\Lambda_{il})\cdot p(\eta_{il} \mid \sigma^2,\Sigma_2)d\eta_{il} \ .$$

- The second level effects conditioned to the first level i, $y_{il}, l = 1,\ldots,m_i$, are multivariate normal and mutually independent,

$$p(y_{il} \mid \eta_{il},\eta_i,\delta,\sigma^2,\Lambda_{il}) = (2\pi\sigma^2)^{-(n_{il}/2)}\left|\Lambda_{il}\right|^{-\frac{1}{2}}.\exp\left\{\frac{-1}{2\sigma^2}\cdot\left(y_{il}-f_{il}(\delta,\eta_i,\eta_{il})\right)'\Lambda_{il}^{-1}\left(y_{il}-f_{il}(\delta,\eta_i,\eta_{il})\right)\right\}.$$

- The marginal density of all second-level effects pertaining to the first level i is also multivariate normal and independent

$$p(\eta_{il} \mid \sigma^2,\Sigma_2) = (2\pi\sigma^2)^{-(k_2/2)}\left|\Sigma_2\right|^{-1/2}\exp\left\{\frac{-1}{2\sigma^2}\eta_{il}'\Sigma_2^{-1}\eta_{il}\right\}.$$

- The marginal density of the first level random effects η_i, is also multivariate normal

$$p(\eta_i \mid \sigma^2,\Sigma_1) = (2\pi\sigma^2)^{-(k_1/2)}\left|\Sigma_1\right|^{-1/2}\exp\left\{\frac{-1}{2\sigma^2}\eta_i'\Sigma_1^{-1}\eta_i\right\}.$$

Thus,

$$\begin{cases} p(y_i \mid \delta,\sigma^2,\Sigma_1,\Sigma_2,\Lambda_i) = \int \prod_{l=1}^{m_i} A_{il}.p(\eta_i \mid \sigma^2,\Sigma_1)d\eta_i \\ \text{where: } \prod_{l=1}^{m_i} A_{il} = \prod_{l=1}^{m_i}\int p(y_{il} \mid \eta_i,\eta_{il},\delta,\sigma^2,\Lambda_{il})p(\eta_{il} \mid \sigma^2,\Sigma_2)d\eta_{il} \\ \text{and } \prod_{l=1}^{m_i} A_{il} = (2\pi\sigma^2)^{-(\sum_{l=1}^{m_i} n_{il}+m_i k_2)/2}\left|\Sigma_2\right|^{-m_i/2}\prod_{l=1}^{m_i}\int B_{il}d\eta_{il} \\ \text{where: } B_{il} = \left|\Lambda_{il}\right|^{-1/2}\exp\left[-(2\sigma^2)^{-1}\left\{\left(y_{il}-f_{il}(\delta,\eta_{il},\eta_i)\right)'\Lambda_{il}^{-1}\left(y_{il}-f_{il}(\delta,\eta_{il},\eta_i)\right)+\eta_{il}'\Sigma_2^{-1}\eta_{il}\right\}\right]. \end{cases}$$

that conducts to:

$$p(y_i \mid \delta,\sigma^2,\Sigma_1,\Sigma_2,\Lambda_i = \left(2\pi\sigma^2\right)^{-(k_1+n_i+m_i k_2)/2}\left|\Sigma_1\right|^{-\frac{1}{2}}\left|\Sigma_2\right|^{-\frac{m_i}{2}}\int\left(\prod_{l=1}^{m_i}\int B_{il}d\eta_{il}\right)\exp\left\{-\left(2\sigma^2\right)^{-1}\eta_i'\Sigma_1^{-1}\eta_i\right\}d\eta_i$$

where $B_{il} = \left|\Lambda_{il}\right|^{-1/2} \exp\left[-(2\sigma^2)^{-1}\left\{\left(y_{il} - f_{il}(\delta,\eta_{il},\eta_i)\right)' \Lambda_{il}^{-1}\left(y_{il} - f_{il}(\delta,\eta_{il},\eta_i)\right) + \eta_{il}'\Sigma_2^{-1}\eta_{il}\right\}\right]$.

Writing $\dot{\eta}_i = (\underbrace{\eta_{i1}',\ldots,\eta_{im_i}'}_{\eta_i^c},\underbrace{\eta_i'}_{\eta_i})'$ and $n_i = \sum_{l=1}^{m_i} n_{il}$, we finally have:

$$p(y_i \mid \delta,\sigma^2,\Sigma_1,\Sigma_2,\Lambda_i) = (2\pi\sigma^2)^{-(n_i+m_i k_2+k_1)/2}\left|\Sigma_1\right|^{-1/2}.\left|\Sigma_2\right|^{-(m_i/2)}$$
$$\int\prod_l^{m_i}\left|\Lambda_{il}\right|^{-\frac{1}{2}}.\exp\left[-\frac{1}{2\sigma^2}\left[\left\{\sum_{l=1}^{m_i}\left(y_{il} - f_{il}(\delta,\eta_{il},\eta_i)\right)' \Lambda_{il}^{-1}\left(y_{il} - f_{il}(\delta,\eta_{il},\eta_i)\right) + \eta_{il}'\Sigma_2^{-1}\eta_{il}\right\} + \eta_i'\Sigma_1^{-1}\eta_i\right]\right] d\dot{\eta}_i .$$

2.4.2. LB-Approximation

In the same way as described in the basic single level NLME model, the estimation algorithm based on the extended multilevel "LB-approximation" that includes the heteroscedasticity and correlation within subject errors, can be formulated using the transformation given by Thisted (1988), for positive definite Λ_{il}:

$$y_{il}^+ = \left(\Lambda_{il}^{-½}\right)' y_{il}$$
$$f_{il}^+ = \left(\Lambda_{il}^{-½}\right)' f_{il}$$
$$u_{il}^+ = \left(\Lambda_{il}^{-½}\right)' u_{il}$$

and applied to the model (2.28). The estimation algorithm described by Lindstrom and Bates (1990) alternates between two steps. A penalized nonlinear least squares (PNLS) step consists on minimizing, over δ , η_i and η_{il} , $i = 1,\ldots,m$, $l = 1,\ldots,m_i$, the nonlinear least square function

$$\sum_{i=1}^{m}\left[\sum_{l=1}^{m_i}\left[\left\|y_{il}^+ - f_{il}^+(t_{il},\gamma_{il}(\delta,\eta_i,\eta_{il}))\right\|^2 + \eta_{il}'\left(\sigma^2\Sigma_2\right)^{-1}\eta_{il}'\right] + \eta_i'\left(\sigma^2\Sigma_1\right)^{-1}\eta_i'\right].$$

This step is completed by the efficient Gauss-Newton algorithm described and implemented en the S-plus function *nlme* from Pinheiro.

The LME step updates the estimates of Σ_1, Σ_2 and Λ based on a first order Taylor expansion of the model function f^+ around the current estimates of δ and the conditional modes of the random effects η_i and η_{il} , which will be denoted by $\widehat{\delta}^{(k)}$, $\widehat{\eta}_i^{(k)}$ and $\widehat{\eta}_{il}^{(k)}$, respectively, derived in the previous step. The algorithm alternates between these two steps until convergence.

Pinheiro and Bates (2000, chapter 7) have presented theory, computational methods, and software for LME and NLME models. Their methodology is based on that of Lindstrom and Bates (1990). Their formulation is presented and illustrated exclusively as a univariate model. However, a mutivariate case can be formulated as a special case of their model and, therefore, can be handled in the framework of their methodology.

2.4.3. Asymptotic properties of mixed models

For linear mixed models, when the number of individuals, m, tends to infinity and the number of observations per individual, n_i, remains finite, i.e. uniformly bounded, the restricted maximum likelihood and maximum likelihood LME are consistent and asymptotically normal, and efficient (Vonesh and Carter, 1987, Pinheiro 1994, Stukel and Demidenko, 1997).

Little is known about the statistical properties of estimates for non-linear mixed effects models even in large samples (Davidian and Giltinan, 1995, Demidenko, 1997). The only way to make any general statements on the NLME model estimation is to consider asymptotic properties of estimators, i.e., when the number of observations tends to infinity. Vonesh and Carter (1992) distinguish two situations:

- When the number of individuals, m tends to infinity and the number of observations per individual, n_i remains finite –i.e. the uniformly bounded case.
- The number of individuals m tends to infinity along with $\min n_i$.

Demidenko (1997) has established some asymptotic properties of four classes of estimators: maximum likelihood estimator (MLE), an estimator based on the first order approximation to the expectation function (Vonesh and Carter, 1992), the two-stage estimator and the Lindstrom and Bates (1990) estimator.

2.4.4. Methods for models comparison

LIKELIHOOD RATIO TEST: the standard test to investigate the difference between strata groups is Wald test, but it has been observed that the likelihood ratio test tends to be more reliable than the Wald approach when m (number of individuals) is not too large.

The likelihood ratio test is applicable in the situation in which we wish to test what are often called "reduced" versus "full" model hypotheses. The "reduced" model is just a special instance of the "full" model. Thus, the "reduced" model, and the null hypothesis corresponding to it, are said to be nested within the "full" model, defining the alternative hypothesis. When the hypothesis are nested in this way, so that we may naturally think in a "full" (H_1) and "reduced" (H_0) model, a fundamental result of statistical theory is that one may construct an approximate test of H_0 vs. H_1 based on the likelihood for the two nested models under considerations.

Suppose now that we fit the "full" model by the method of maximizing the likelihood $L_{full}(\delta, \vartheta)$ and we fit the "reduced" model by maximizing the likelihood function $L_{red}(\delta_0, \vartheta)$ (the likelihood is the same, except that the mean of the data vector is restricted to have the form specified by H_o). Let $\hat{L}_{full}$ and $\hat{L}_{red}$ denote the values of the likelihood with estimates plugged in:

$$\hat{L}_{full} = L_{full}(\hat{\delta}, \hat{\vartheta}) \text{ and } \hat{L}_{red} = L_{red}(\hat{\delta}_0, \hat{\vartheta}).$$

Then the likelihood ratio statistic is given by

$$T_{LRT} = -2\{\log(\hat{L}_{red}) - \log(\hat{L}_{full})\} \quad .$$

When m tents to infinity, the asymptotic distribution of T_{LRT} is a χ^2 distribution with degrees of freedom equal to the difference in the number of parameters in the two models. Thus, if this difference is equal, say, to r then we reject H_0 in favor of H_1 at an α-level of significance if

$$T_{LRT} > \chi^2_{r,(1-\alpha)}$$

where $\chi^2_{r,(1-\alpha)}$ stands for the $1-\alpha$ quantile of a chi-square distribution with r degrees of freedom.

AKIKE'S INFORMATION CRITERION [*Akaike, H. (1974)*] (AIC): One drawback of the likelihood ratio test is that it requires the model under the null hypothesis to be nested within that of the alternative.

Other approaches to compare models have been proposed that do not require this restriction. These are based on the notion of comparing penalized versions of the logarithm of the likelihood obtained under H_0 and H_1, where that "penalty" adjusts each log-likelihood according to the number of parameters that must be fitted. It is a fact that, more parameters we add to a model,

the larger the (log) likelihood becomes. Thus, if we wish to compare two models with different numbers of parameters, and favoring the model that gives the large log-likelihood value, one would prefer the model that has the lowest AIC, defined as

$$AIC = \text{-}2 \log \hat{L} + 2s \quad .$$

where s is the number of parameters in the model.

SCHWARZ'S BAYESIAN INFORMATION CRITERION [*Schwarz, G. (1978)*] (BIC): In this approach, the penalty is to subtract the number of fitted parameters adjusted by the number of observations. If, as before, N is the number of observations, we have

$$BIC = -2 \log \hat{L} + s \log(N).$$

One would prefer the model with the lowest BIC value.

2.5. Bibliographic Review

A simple version of the random coefficient regression model was first proposed, for the analysis of growth curves, by Elston and Grizzly (1962) who applied this model to the growth curve in orthodontics. Rao (1965) derived the best linear estimator for the mean parameter vector of the random coefficient regression model in the case where the within-individual design matrix is the same for all individuals. He also derived the best linear unbiased predictor for the random coefficients, for known variance and covariance parameters. Swamy (1970, 1971,1973) proposed several estimators, including generalized least squares and maximum likelihood, and provided large sample results for the estimators and associated test statistics. Rosenberg (1973) viewed the random coefficient regression model as a special case of the mixed effects models, in which some parameters are fixed and some random. He proposed empirical Bayes estimation of random coefficients, along with maximum likelihood estimation of the variance-covariance parameters. Harville (1977), writing primarily about variance components models but putting them into the more general mixed-effects model framework, provided a comprehensive overview of methods of estimation, including maximum likelihood, restricted maximum likelihood (REML), jointly with algorithms for computing maximum likelihood estimates of variance components. Laird and Ware (1982) described the E-M algorithm for computing maximum likelihood and REML estimates in mixed models. Gelfand, et al. (1990) proposed Gibbs sampling , which is a Markov Chain Monte Carlo method, for a full Bayesian analysis of a hierarchical growth model. Applications of linear and non- linear random coefficient models to pharmacokinetics and other areas are cited in Yuh et al (1994). For the case where the random effects may not be normally distributed, Davidian and Gallan (1993) proposed a method for maximum likelihood estimation of the density of the random effects along with estimation of the fixed parameters in non-linear mixed effects models, assuming only a smooth density for the distribution of the random coefficients. Tao et al. (1999) proposed a new strategy to estimate the parameters when the random effects are not normally distributed using the estimate of the

distribution function based on an algorithm introduced by Newton and Zhang (1999), a predictive recursive algorithm for non-parametric distribution estimation.

Chapter 3

3. Some parametric biostatistical analysis

3.1. Introduction

This chapter will be devoted to a more detailed analysis of some of the data sets that have already been introduced in the first chapter. Linear and non-linear mixed models will play a central role in these analyses.

3.2. Modelization under a parametric perspective

3.2.1. Analysis of nitrification data

In this section, we will analyze the nitrification data obtained by Dra. Genoveva Montserrat. The mixed models analysis presented here complement the more classical analyses performed by this researcher, based on ANOVA and other statistical tools.

3.2.1.1. Classical versus individual models

The initial approach to model the nitrification data, displayed in the Figure 1.1 of the first chapter, is based on ignoring the grouping structures and assuming that the true regression relationship between the response y and time t is the sum of a systematic part described by a linear function $\mu(t) = \delta_0 + \delta_1 t + \delta_2 \log(t)$ and a random part described by a random error, equal by construction to the discrepancy between y and $\mu(t)$. The classical linear regression model resulting for the outcome observation $y_j\,(j = 1,\ldots,n)$ is given by

$$\begin{aligned} y_j &= \delta_0 + \delta_1 t_j + \delta_2 \log(t_j) + e_j \\ y_j &= (1, t_j, \log(t_j))\,\delta + e_j \end{aligned} \tag{3.1}$$

where t_j is the value of the predictor or regressor variables for observations j, δ is a $p\times 1$ parameter vector of regression coefficients and e_j is the error term for observation j.

It is a standard assumption in fitting these models to suppose that the e_j are independent and normally distributed.

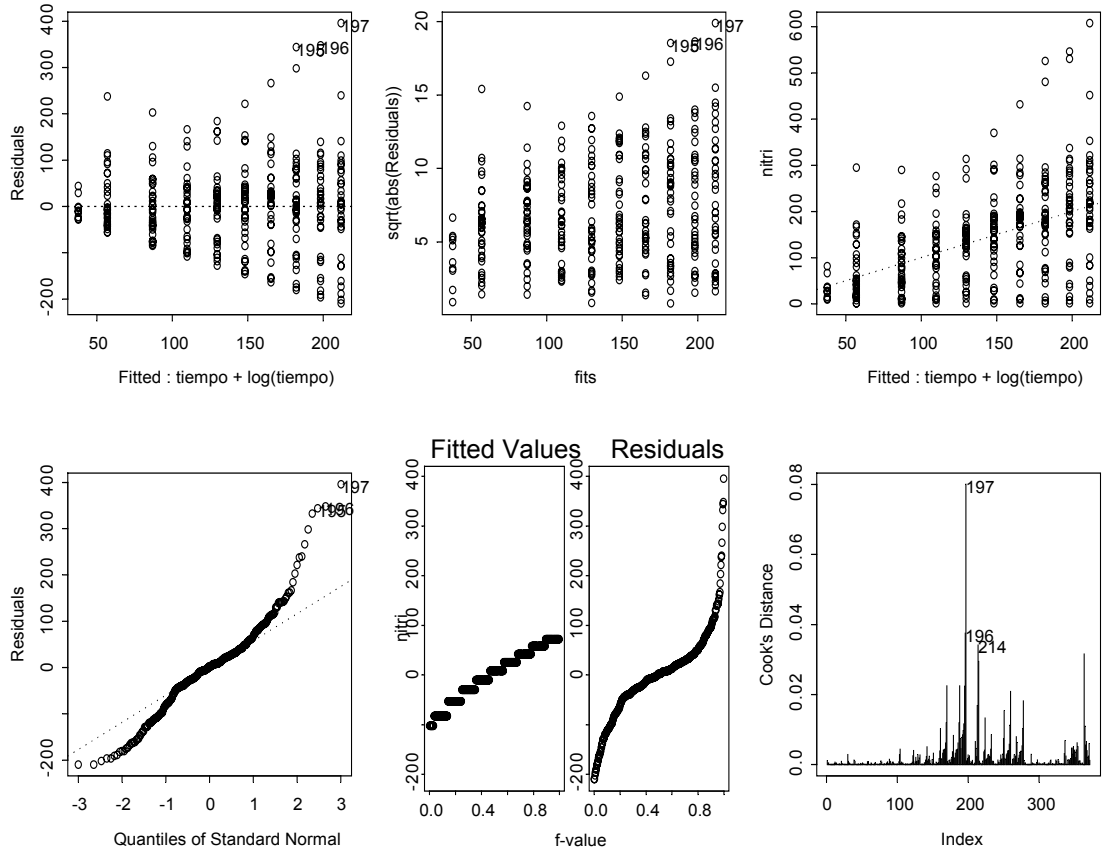


Figure 3.1: Plots for nitrification data

One of the fundamental problems is to asses how well this model fits the data, in order to modify it or even abandon it, if it doesn't satisfactory explain the data.

The simplest and most informative method for assessing the fit is to use graphical methods, using plots that reveal the strengths and weaknesses of the model. From this point of view, Figure 3.1 summarizes the essential results of fitting the model (3.1) to the full set of data. The dashed line in the third plot gives a good idea of how well the model fits the data. It seems clear that it doesn't explain them satisfactorily. The plot of the residuals against the fitted values

reveals unexplained structures in the residuals, which in an adequate model should appear as nothing but noise. The normal probability plot, particularly with medium or small samples, provides a visual test of the assumption that the model's errors are normally distributed (Chambers et al. 1983). The Cook's distance plots (the measure of the influence of individual observations on the regression coefficients) show some influential observations. Hence, we do not feel justified to use the simple regression line as a way to estimate and explain the evolution of the nitrification data as a function of time. The residuals-fit spread plot is particularly appealing. It relates the spread of the fitted values with the spread of the residuals. Since the model is an attempt to explain the variation in the data, one hopes that the spread in fitted values will be much greater than in residuals. Actually, the spread of the residuals is greater than the observed spread of original data. In addition, the regression line explains only about 25.2% of the variation in the data. Finally, the analysis of the residuals by subject number, displayed in the left panel of Figure 3.2, illustrates the fundamental problem when the "effects" for each subject are ignored. Moreover, this method does not provide an estimate of the between–subject variability which has a great interest in the nitrification data analysis.

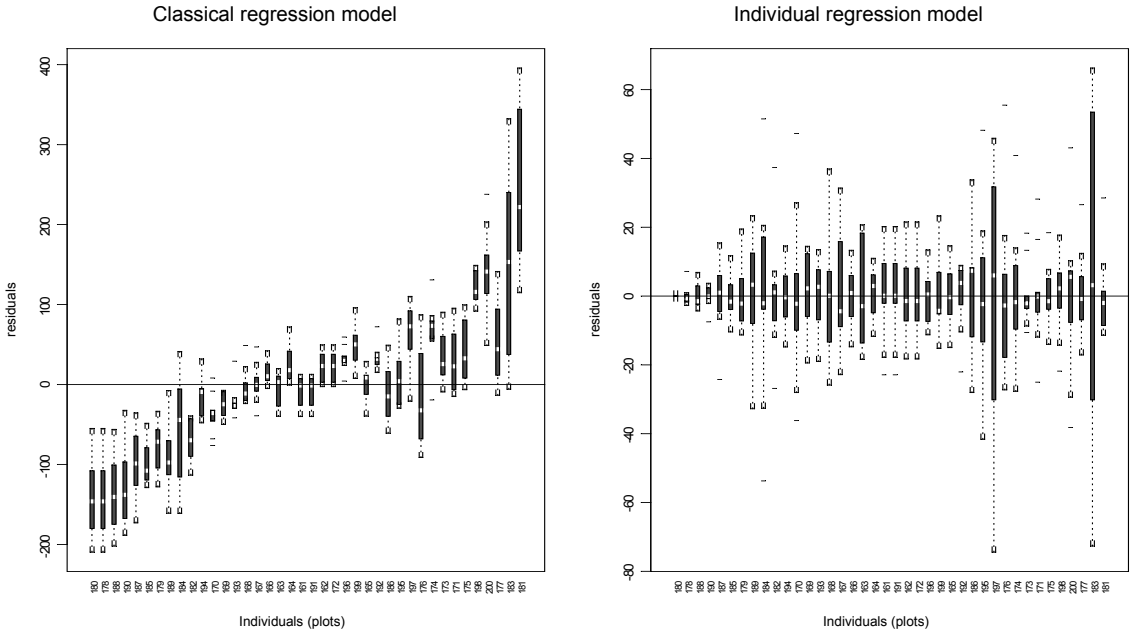
An alternative approach is to construct linear models for each subject (i.e. the "curve effects" indicated in Figure 3.2 may be incorporated into the model by allowing the mean of each subject to be represented by separate parameter). This leads to the so called "individual model":

$$\begin{aligned} y_{ij} &= \delta_{0i} + \delta_{1i} t_j + \delta_{2i} \log\left(t_j\right) + e_{ij} \\ y_{ij} &= \left(1, t_j, \log(t_j)\right) \delta_j + e_{ij} \end{aligned} \tag{3.2}$$

where j is the observation index, $j = 1,\ldots,N$, for the i–th subject, $i = 1,\ldots,m$, and $\delta_i = \left(\delta_{0i}, \delta_{1i}, \delta_{2i}\right)'$ is the parameter vector, while e_{ij} is the error term for observation j in the i–th subject.

The use of model (3.2) yields to a substantial reduction in the estimation of σ, $\hat{\sigma} = 20.40235$ instead of $\hat{\sigma} = 89.4$ of model (3.1). The residuals (displayed in the right panel of Figure 3.2) are now centered on zero and have a considerably smaller magnitude.

In order to assess how well model (3.2) fits to data, a coefficient of multiple determination for each subject i, R_i^2, is plotted for each subject separately in Figure 3.3.



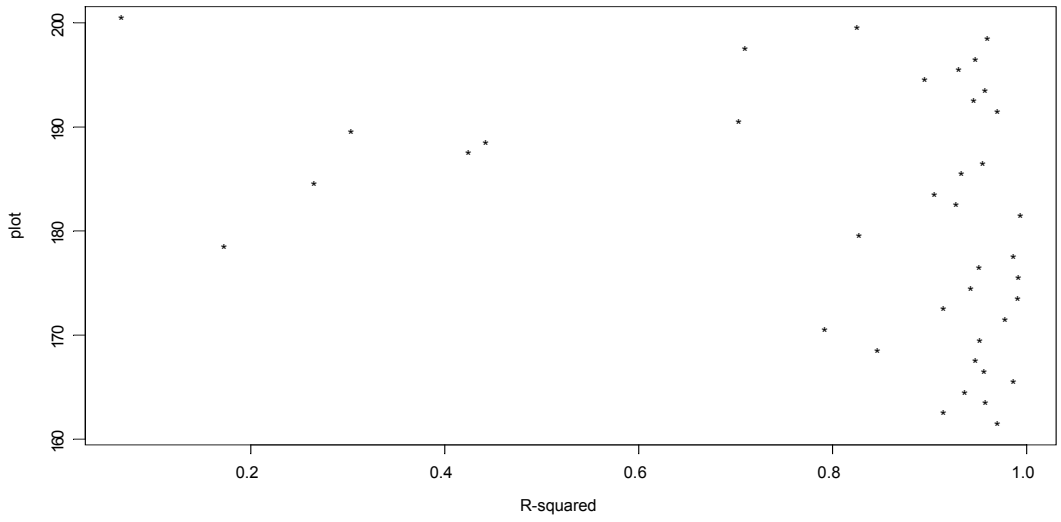


Figure 3.2 Box plots of residuals by subject (curve) tend to be centred on zero

Figure 3.3: Coefficients R_i^2 of multiple determinations for each subject

As is expected, the overall mean of the R^2 coefficients indicates that around 82.2% of the observed variance is now explained by the separate linear regression models, instead of a 25.2% for the global model (3.1). But its interpretation is quite difficult because the number of parameters in model (3.2) increases linearly with the number of subjects.

3.2.1.2. A compromise between global and individual models: mixed models

In this section we consider a model similar to (3.2) with the difference that the coefficients δ_i are split into two parts, one fixed part expressing its mean δ, and an other part expressing the random deviation of the *i–th* individual coefficient vector with respect to the mean, $\eta_i = \delta_i - \delta$. This model is also useful in studies in which the levels of some (or all) explanatory variables or covariates vary within individuals and are not the same across individuals.

The next question to be addressed in this modeling process is the determination of which parameters are random and which ones are purely fixed across individuals. Probably the best way to achieve this is by analyzing a plot of the confidence intervals for the individual parameters of model (3.2).

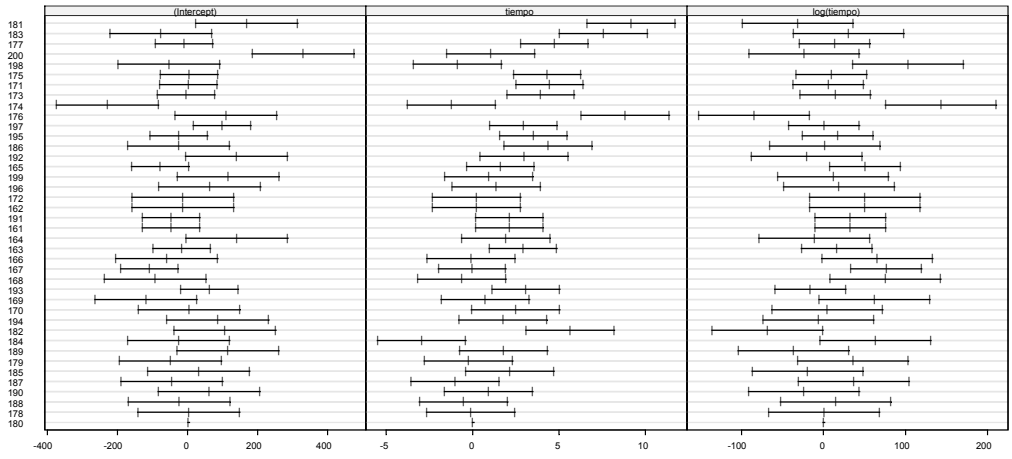


Figure 3.4: Confidence intervals for individual parameters used to identify the variability between subjects effects.

It is clear from Figure 3.4 that a strong variability between the individual parameters is present, suggesting that all parameters should be treated (almost in a first approach) as mixed with a fixed and random part.

When we fit this model using restricted maximum likelihood (REML), we obtain that the p-value associated to the F-test for the intercept, δ_0, slope, δ_1, and the coefficient of the log-time, δ_2, have very small values ($<.0001$, $<.0001$ and 0.0026 respectively). That confirms that this model is linear in the parameters and non-linear in time. Additionally, the large magnitude of the REML estimates of the variances in the diagonal of the covariance matrix of random parameters, $\hat{\sigma}_0 = 63.142534$, $\hat{\sigma}_1 = 2.305624$ and $\hat{\sigma}_2$=26.474684 respectively, indicate that all random parameters should be included in the model. In other words, the estimated parameters in the computer output may reflect the actual variability introduced by this step. Figure 3.5 displays the superposed plot of the individual coefficient estimates for model (3.2) and those of the mixed model,

$$y_{ij} = (\delta_0 + \eta_{i0}) + (\delta_1 + \eta_{i1})t_{ij} + (\delta_2 + \eta_{i2})\log(t_{ij}) + e_{ij} \tag{3.3}$$

(that is, the fixed parameters plus the random individual effects). With model (3.3), an improved fit is reached with less parameter variability.

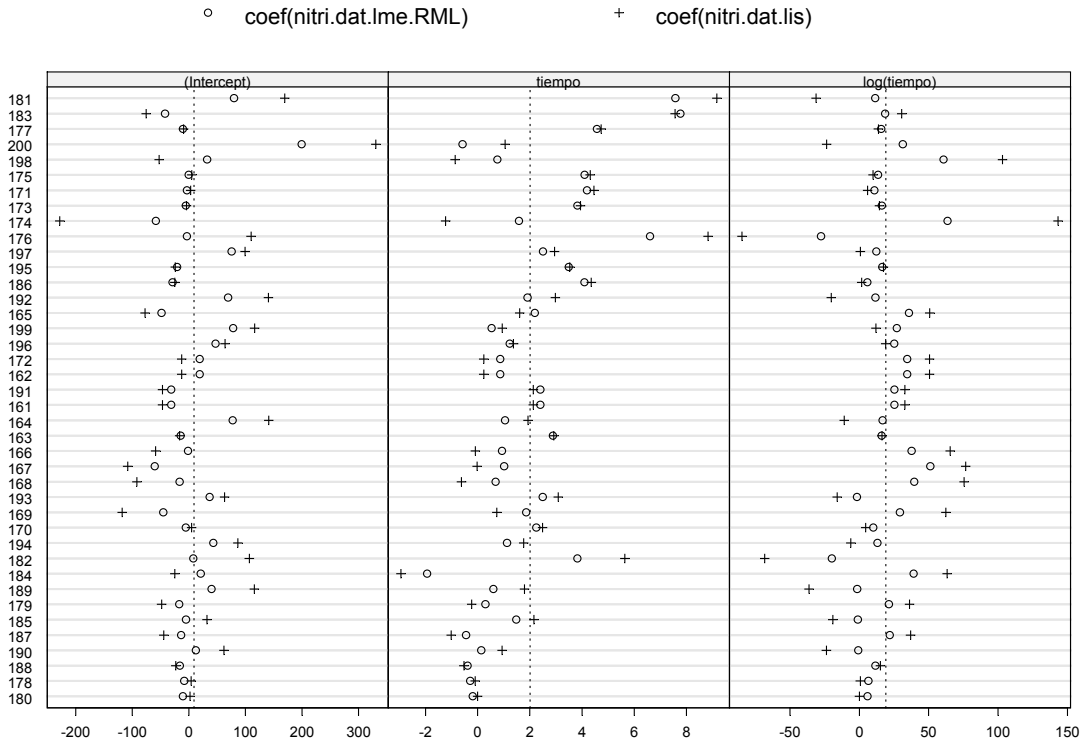


Figure 3.5: Individual estimates from separate fit (+ model (3.2)) and from mixed fit (o model (3.3))

The most relevant information about the fitted models, such as parameter estimates and their standard errors, as well as the values in the diagonal of the covariance matrix, obtained using maximum likelihood (ML) and restricted maximum likelihood (REML) is listed in Table (3.1). For the sake of the comparison, the random effects distribution for the two methods is displayed in Figure 3.6.

Table (3.1): Estimate of mixed effects of the model (3.3)

	Model-Method	
	(3.3)-ML	(3.3)-REML
δ_0	8.99639	9.13212
SE	13.85158	14.02322
δ_1	2.00666	2.00853
SE	0.40547	0.40934
δ_2	19.16049	19.10164
SE	6.21545	6.29228
D_{00}	2283.39724	2385.78563
D_{11}	4.612647	4.65126
D_{22}	321.62220	344.24724
σ	20.544899	20.535733

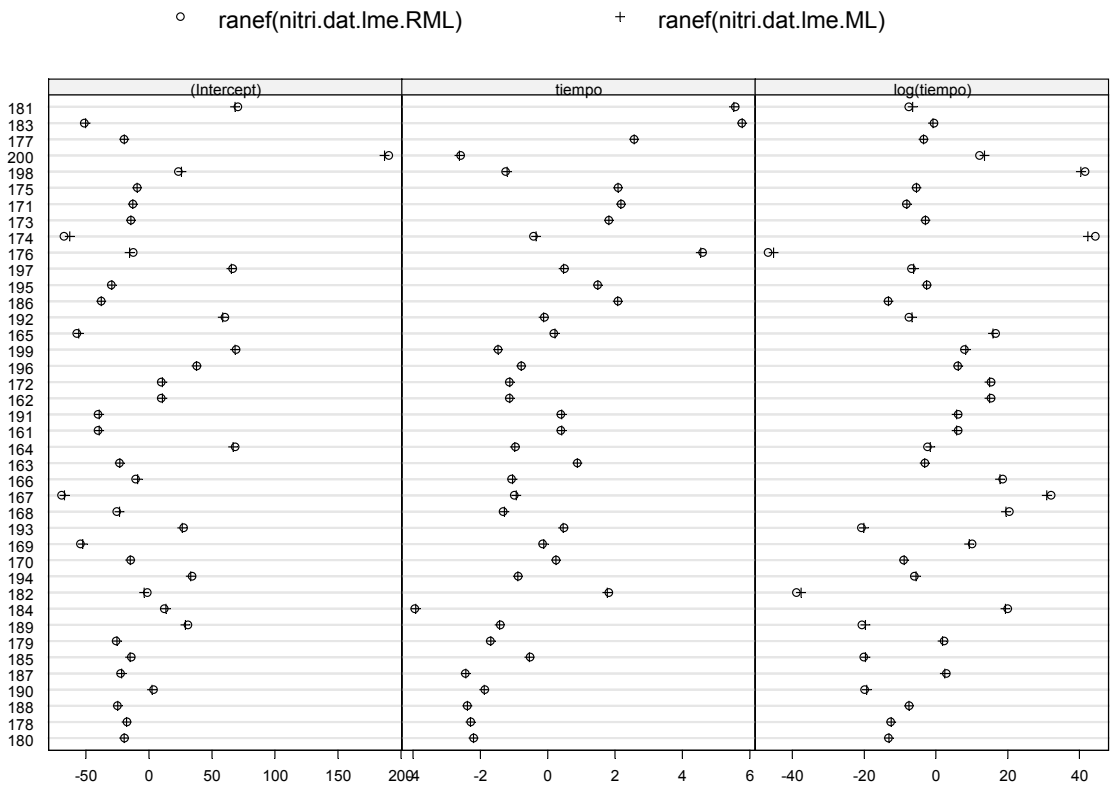


Figure 3.6: Scatter plots of random effects of model (3.3)

In addition, the scatter plot for the estimated random effects corresponding to model (3.3), shown in Figure 3.7, indicates that there is no substantial correlation between any of these random effects and hence no random effects could be dropped from the model.

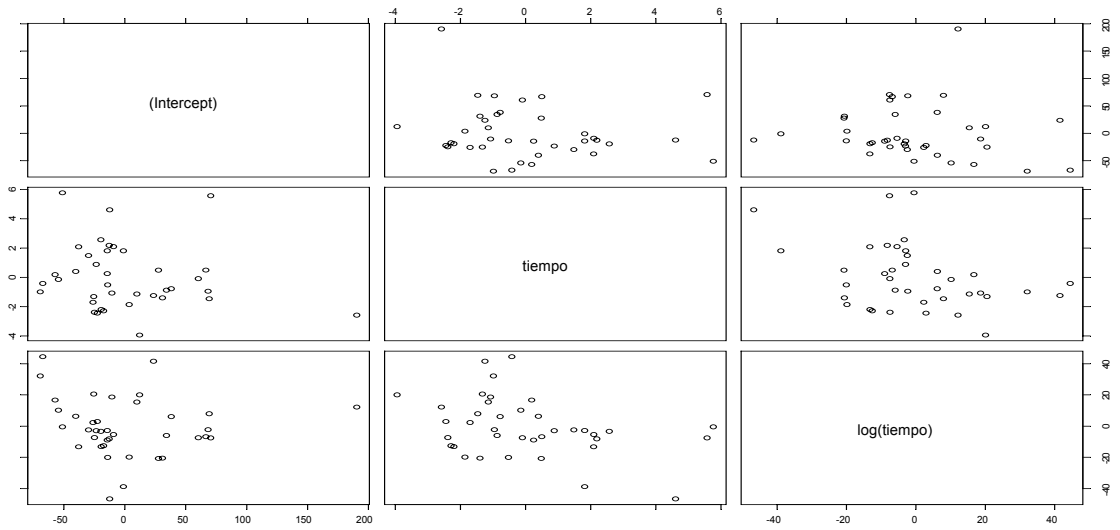


Figure 3.7: Matrix plots of random effects

The estimation of the parameters is only the first stage in any analysis. The next step is an extensive examination of the fit and the residuals.

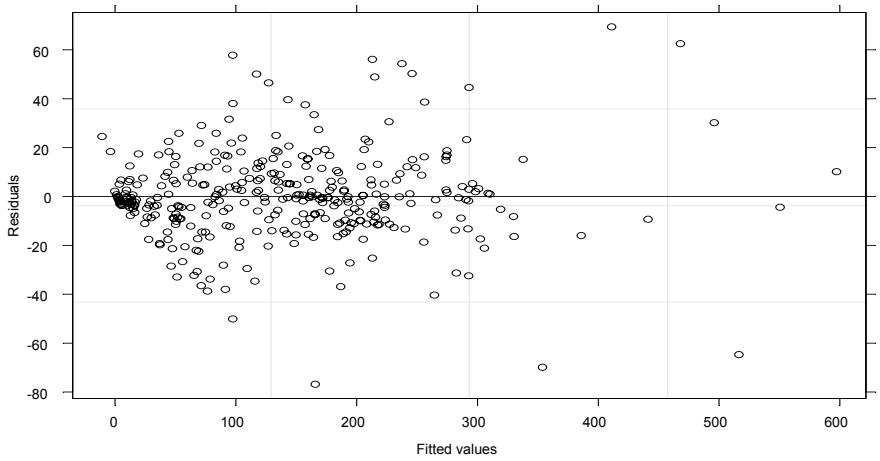


Figure 3.8: Residuals versus fitted model (3.3) for each subject

Careful examination of the plot involving standardized residuals versus fitted values (Figure 3.8), suggests an heterogeneous intra-individual variance for some subjects. The power of mean seems to be an adequate function to model the structure of this variance heterogeneity. The proposed model is:

$$\begin{aligned} y_{ij} &= \left(\delta_{0i} + \eta_{0i}\right) + \left(\delta_{1i} + \eta_{1i}\right) t_j + \left(\delta_{2i} + \eta_{2i}\right) \log(t_j) + e_{ij} \\ \text{var}(e_{ij}) &= \sigma^2 \mu\left(t_{ij}, \delta\right)^{\tau} \end{aligned} \tag{3.4}$$

where τ is a parameter describing the spread of change of the variance.

Table (3.2) includes the estimate of this new parameter, τ. We next compare the accuracy of model (3.3) with the accuracy of model (3.4), obtained by considering the above structure in the variance of residuals. The AIC of model (3.4) is 3571.661, smaller than the AIC of model (3.3), 3643.334, suggesting that it should be preferred over the previous models. These results give some support to the convenience of considering some heteroscedastiticity in the residuals and to model it as in (3.4).

Since the responses in the nitrification study are repeated measurements over time, we need to make some kind of exploratory time series analysis to check for serial correlation within subjects. One might be tempted to introduce both, heterogeneity of variance and intra-individual correlation. The plot of the empirical autocorrelation within subject residuals shown in Figure 3.9 provides a first visual assessment.

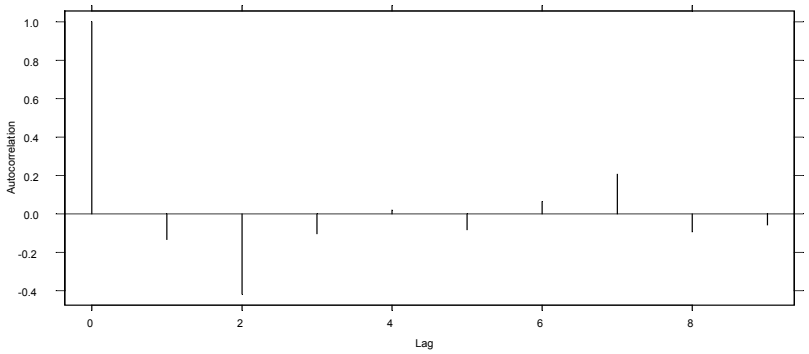


Figure 3.9: The autocorrelation plot of residuals of nitrification data

It is reasonable to suppose that adjacent measurements may be associated. Thus, we also explore the convenience of introducing an auto-regressive (AR) model of order $p = 1$, with a correlation parameter ϕ, to describe the residuals structure:

$$\begin{aligned} y_{ij} &= \left(\delta_{0i} + \eta_{0i}\right) + \left(\delta_{1i} + \eta_{1i}\right) t_j + \left(\delta_{2i} + \eta_{2i}\right) \log(t_j) + e_{ij} \\ \text{var}(e_{ij}) &= \sigma^2 \mu\left(t_{ij}, \delta\right)^{\tau} \\ \text{cor}(e_{ij}, e_{ik}) &= \phi^{|j-k|}; \text{ where } |\phi| < 1 \text{ and } j \neq k \end{aligned} \tag{3.5}$$

The resulting model has the smallest AIC, 3565.658, and also the smallest BIC, 3612.587. For model (3.4) these values are 3571.661, and 3614.680 respectively. Therefore, model (3.5) seems

to give a better representation of the nitrification data. Table (3.2) summarizes the parameter estimates as well as measures of fit of models (3.4) and (3.5).

Table (3.2): Restricted maximum likelihood (REML) fits of the nitrification data

	Fitted Models	
	Model (3.4)	Model (3.5)
δ_0	9.34932	4.81243
SE	11.89616	11.13176
δ_1	2.07647	1.99241
SE	0.43710	0.42575
δ_2	18.06124	19.94213
SE	6.79376	6.70211
D_{00}	2541.46382	1427.434135
D_{11}	5.800428	4.514748
D_{22}	992.75815	712.93484
σ	1.809026	2.393094
τ	0.5013103	0.4746697
ϕ	–	0.3407895
AIC	3571.661	3565.658
BIC	3614.68	3612.587
$\log - Lik$	-1774.83	-1770.829

The plots shown in the Figure 3.10 capture, reasonably well, the main trends of model (3.5). The plot of observed versus fitted values indicates that, using restricted maximum likelihood as fitting a method, model (3.5) fits the data reasonably well (most points fall relatively close to the $y = x$ line). Also, the plot of residuals versus fitted values and "qq-norm" plots do not indicate any departures from the typical assumptions of the model.

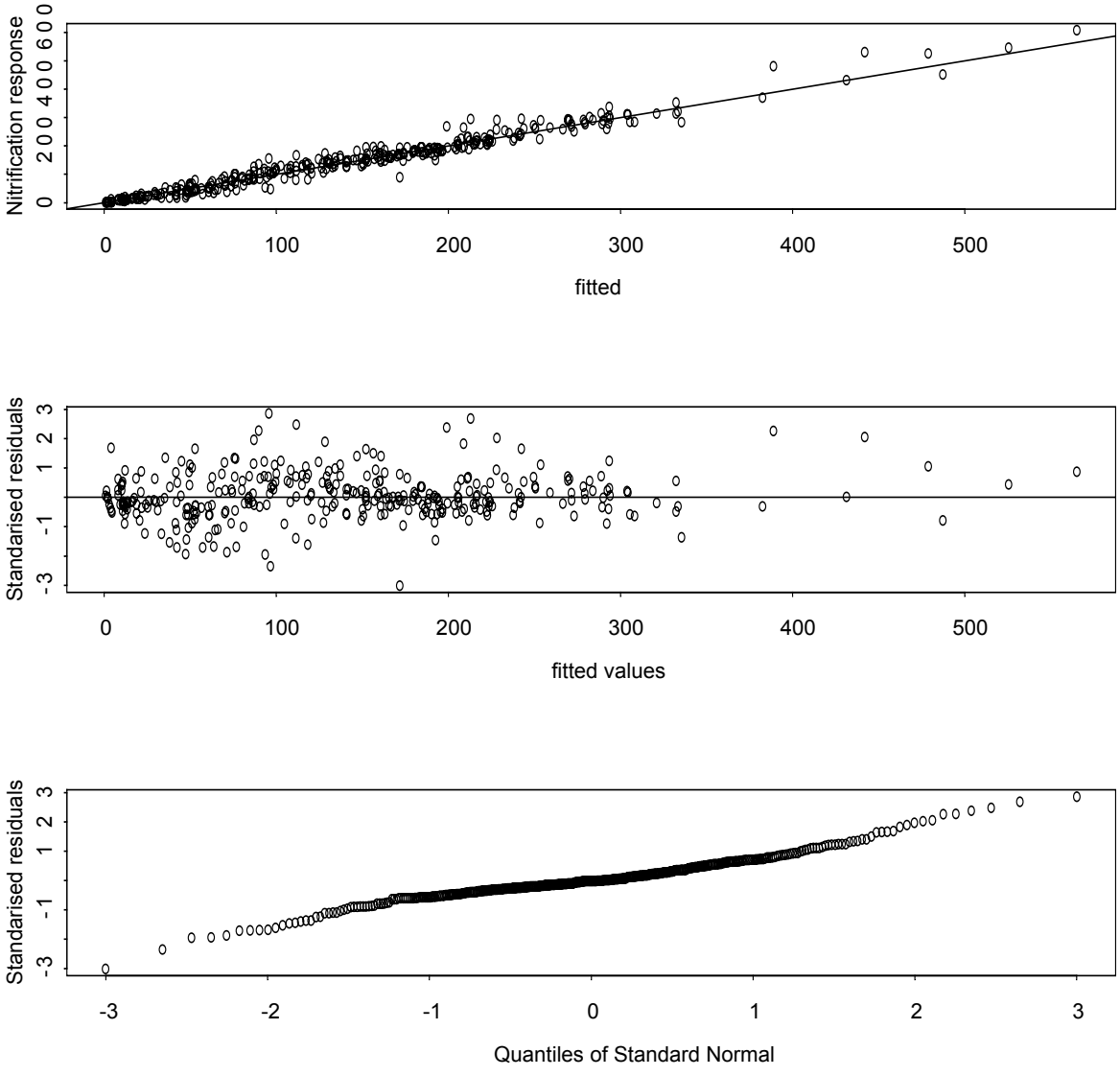


Figure 3.10: plots of fit of model (3.5)

3.2.2. Analysis of ammonification data

3.2.2.1. Ammonification global analysis

As was mentioned in the first chapter, the primary analytic goal of these experiments is to describe the relationship between the response processes of data and the time of application. While the starting models (1.4) to (1.6) can be a good choice, the preliminary analyses also indicate that some modelisation of the heterogeneity of variance would be convenient. The alternative model would be:

$$\begin{aligned} y_{ij} &= \delta_0 + \left[\delta_1 + \eta_{1i}\right] t_{ij} + \eta_{0i} + e_{ij} \\ \text{var}(e_{ij}) &= \sigma^2 \mu(t_{ij}, \delta)^{\tau}. \end{aligned} \tag{3.6}$$

The fit results of this model, constraining the covariance matrix D to be diagonal in all cases, are summarized in Table (3.3) and also displayed in Figure 3.11. It seems clear that the estimates of the variance components have a very high magnitude. This suggest that some covariate should be incorporated into the model, to explain the systematic variation among subjects (doses).

Table (3.3): Restricted maximum likelihood (REML) fits of the amonification data

	Models	
	Models: (1.4)-(1-6)	Model (3.6)
δ_0	168.0247	164.7841
SE	41.23775	39.82926
δ_1	3.5631	3.6356
SE	0.51099	0.50630
D_{00}	66639.2556	62495.8079
D_{11}	9.52621	9.547842
σ	38.305255	2.937196
τ	–	0.4464635
AIC	4134.233	4026.357
BIC	4157.714	4053.751
$\log - Lik$	-2061.116	-2006.178

In addition, the possible presence of a correlation structure in the ammonification data is suggested by the sample semivariogram of the residuals of the model (3.6).

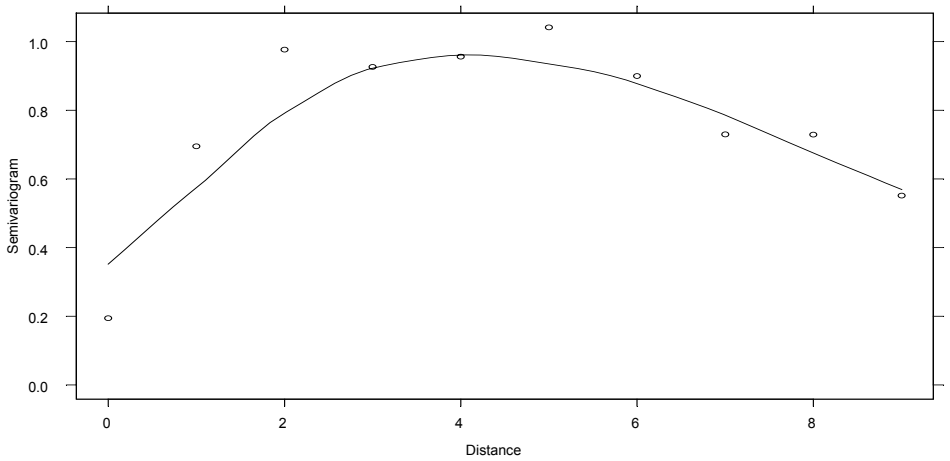


Figure 3.11: Sample semivariogram corresponding to the standardized residuals, including a "loess" smoother to enhance the visualization of semivariogram patterns

The resulting plot, shown in Figure 3.11, suggests a possible rational quadratic model, described in Pinheiro and Bates (2000), to model this correlation structure. Table (3.4) summarizes the comparison among a model where the errors have the structure mentioned above and model (3.6). The small value of the p-value associated to the likelihood ratio test indicates that the rational quadratic model fits the data significantly better than the independent error model.

Table (3.4): Likelihood ratio test of the correlation model

	DF	AIC	BIC	Log-Lik	L.Ratio	P_value
Independent model (model (3.6))	7	4026.357	4053.751	-2006.178		
Rational quadratic model	8	4001.483	4032.791	-1992.741	26.87361	$<$. 0001

Also from Figure 3.12, displaying the most important features of this fit, it is clear that the estimated values of the response are in better concordance with the observed values and do not indicate any departure from the usual assumptions of the model.

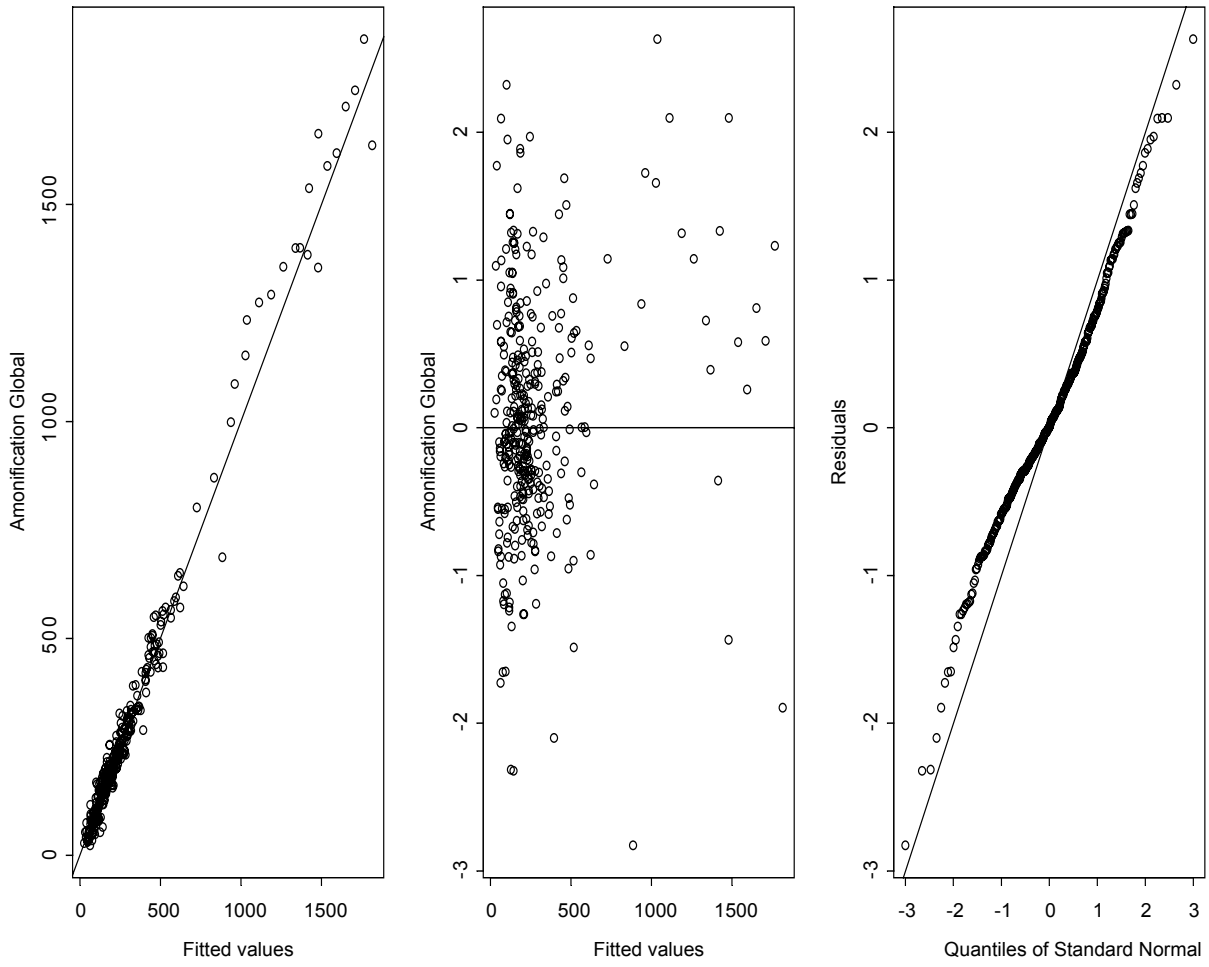


Figure 3.12: Fitted plots of amonification data

3.2.2.2. Ammonification residual

The relationship between the ammonification model residuals and time can also be described by the same linear mixed model (1.4) to (1.6) formulated in the first chapter, but the diagnostic plots of the model residuals versus time indicate that the assumption of constant variance is violated and suggest that

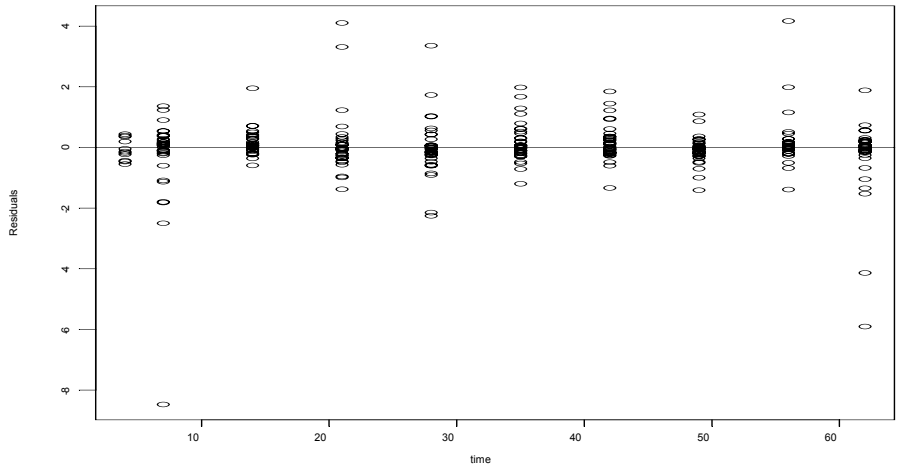


Figure 3.13: residual vs. time for amonification residual data

the variance function is a power function of time, defined as $t^{2\tau}$ where τ is the variance function coefficient. On the other hand, Figure 3.14 suggests that the correlation between two observations can be modeled by the spherical spatial correlation structure also described in Pinheiro and Bates (2000).

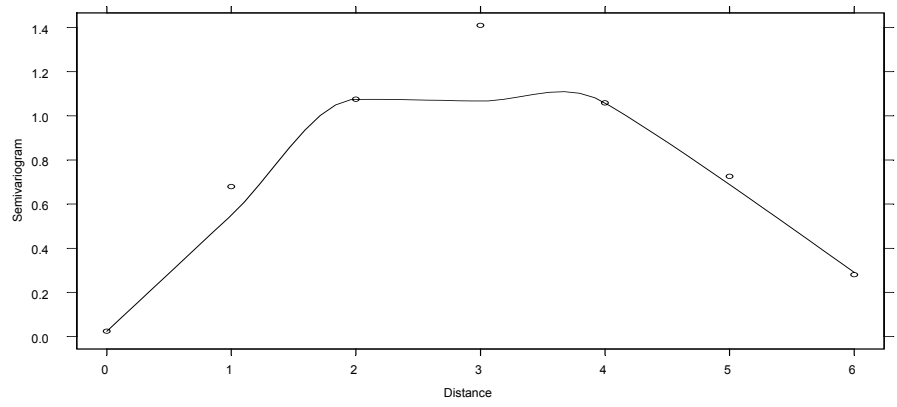


Figure 3.14: Variogram for the amonification residuals

One way to test whether the preceding suggestions are adequate to model the correlation structure is by fitting the "reduced" model (model (3.6) with $\text{var}(e_{ij}) = \sigma^2 t_{ij}^{2\tau}$) and comparing it with the "full" model, with additional spherical spatial structure for the model residuals. The results are summarized in Table (3.5). It seems clear that the spherical spatial correlation model produces a significantly better fit.

Table (3.5): Likelihood ratio test of the spherical spatial correlation model

Model	DF	AIC	BIC	Log-Lik	L.Ratio	P_value
Power function structure model	7	4049.954	4077.349	-2017.977		
Spherical spatial correlation structure	8	3994.489	4025.797	-1989.245	57.46499	<. 0001

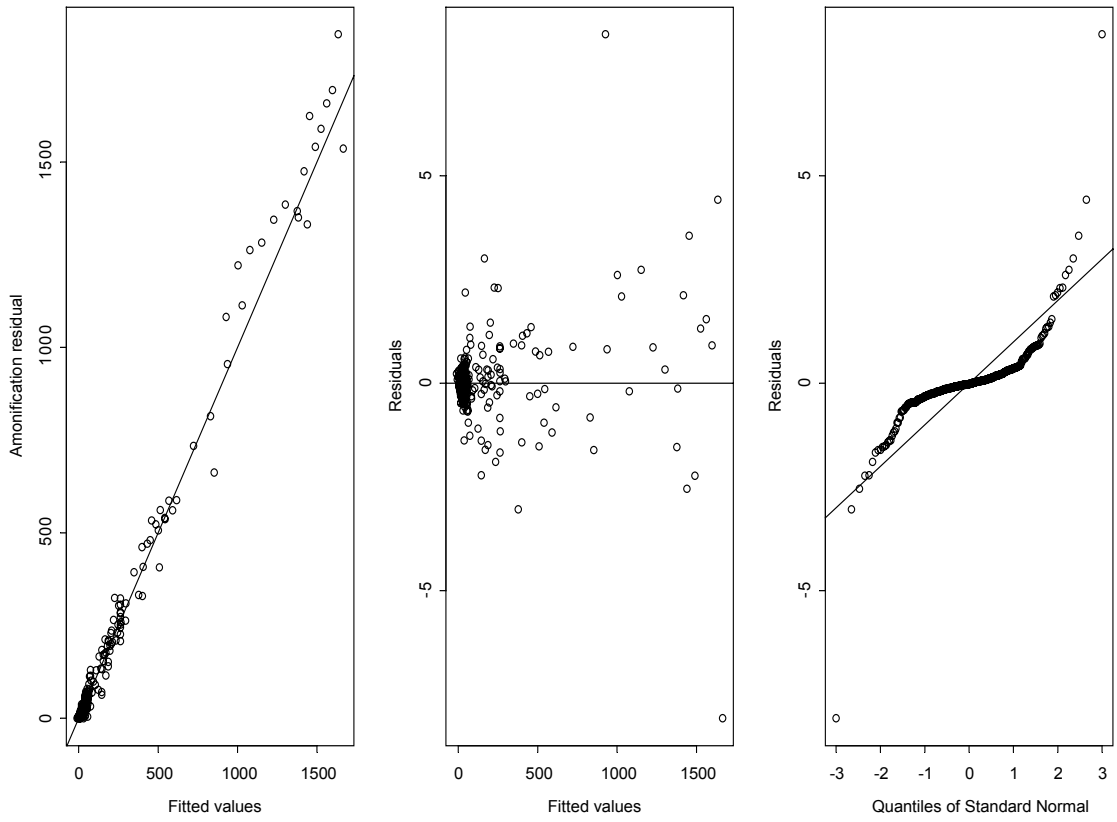


Figure 3.15: Plot of the fit model of the amonifcation residual data

The next step is to asses the validity of the fitted model by examining Figure 3.15. The good agreement between fitted and observed values indicates that the model mentioned above is adequate to represent the data. On the other hand, the graphics suggest the presence of some outliers and a clear departure from normality of the model residuals. Additional analysis of these data (based on the classical methods) and discussion of formal testing procedures that address the comparisons of interest are given in the Montserrat (2000).

From the discussion above one might conclude that the main role of mixed estimates is to help to study the inter-subject variability. and to model the complex situations such as heterogeneous errors variances, removing the assumption of a common variance shared by all subjects.

3.2.3. Analysis of daily carbon (CO_2) respiration data

The preliminary exploratory analyses of these data suggest a clear non-linear dynamics. Continuing with the same strategy used in the previous analysis, the following non-linear mixed model may be used to investigate the relationship between the daily carbon and time

$$\begin{aligned} & y_{ij} = \left(\delta_{i1} + \eta_{i1}\right)\exp(\left(\delta_{i2} + \eta_{i2}\right)\cdot t_{ij}) + \left(\delta_{i3} + \eta_{i3}\right)\exp(\left(\delta_{i4} + \eta_{i4}\right)\cdot t_{ij}) + e_{ij} \\ & \text{var}(e_{ij}) = \sigma^2 \cdot t_{ij}^{2\tau_r}, r = 1,\ldots,4 \\ & cor(e_{ij}, e_{ik}) = \phi^{|j-k|},\ |\phi| < 1 \text{ and } j \neq k. \end{aligned} \tag{3.7}$$

This model incorporates the heterogeneity of variance for each type of industrial residual and a common correlation structure. The principal features of the model fit are displayed in the Figure 3.16 and listed in the Table (3.6). The residuals seem to be clearly not normal.

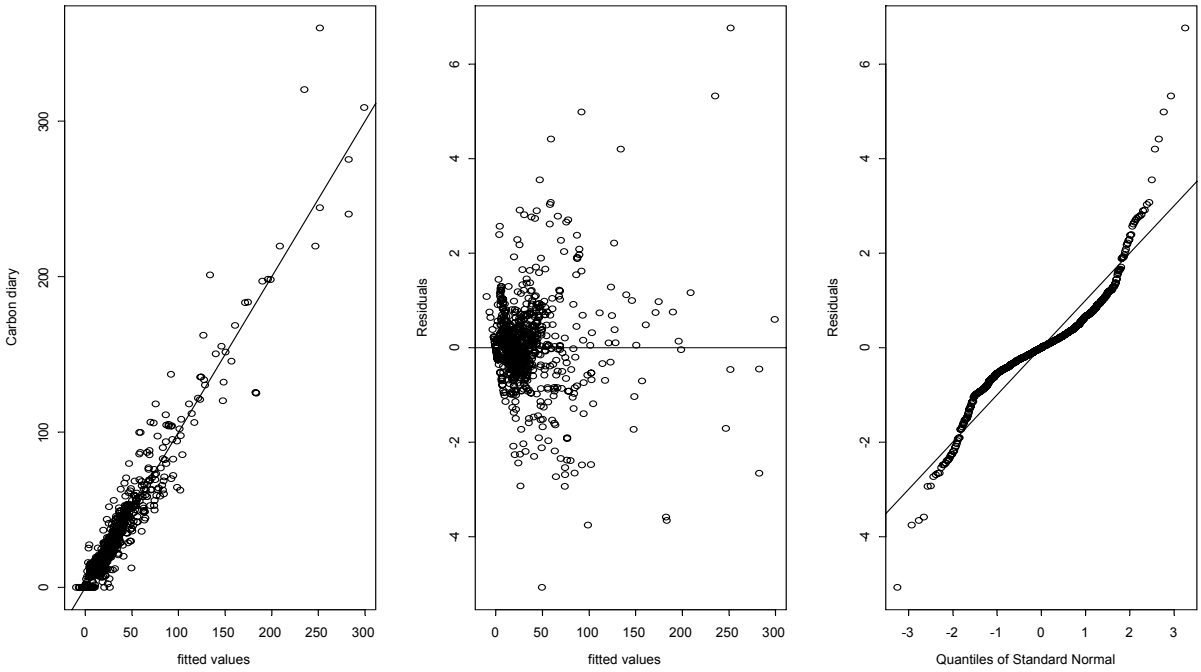


Figure 3.16: Essential features of the fit of the model (3.8)

In the next sections, we address the question on how these departures from the assumption of normality do affect the validity of the parametric inferential methods used in this section.

Table (3.6): summary results of model (3.7) for daily carbon data

Fixed effects		Random effects	
Parameters	Values	Parameters	Values
δ_1	70.33666	D_{11}	5642.753918
SE	12.61296	D_{12}	-9.4398874
δ_2	-5.60778	D_{22}	0.2631310
SE	0.03888	D_{13}	436.2364906
δ_3	41.00208	D_{23}	0.7790953
SE	2.87760	D_{33}	246.0313316
δ_4	-8.49721	σ	9.02382
SE	0.05206	τ_1	0.8064208
		τ_2	1
		τ_3	1.769822
		τ_4	1.013621
		ϕ	-0.08103743

Because interest focuses on identification of associations between daily carbon model parameters and the measured subject attributes dose effects for each industrial residual and each soil type, we incorporate the dose effects as covariates in each one of the parameters of the non-linear model, as follows:

$$
\begin{aligned}
y_{ij} &= \delta_{i1}\exp(\delta_{i2}\cdot t_{ij}) + \delta_{i3}\exp(\delta_{i4}\cdot t_{ij}) + e_{ij} \\
\delta_{i1} &= \delta_1 + \lambda_{12}dS_{2i} + \lambda_{13}dS_{3i} + \lambda_{14}dS_{4i} + \lambda_{15}dS_{5i} + \eta_{i1} \\
\delta_{i2} &= \delta_2 + \lambda_{22}dS_{2i} + \lambda_{23}dS_{3i} + \lambda_{24}dS_{4i} + \lambda_{25}dS_{5i} + \eta_{i2} \\
\delta_{i3} &= \delta_3 + \lambda_{32}dS_{2i} + \lambda_{33}dS_{3i} + \lambda_{34}dS_{4i} + \lambda_{35}dS_{5i} + \eta_{i3} \\
\delta_{i4} &= \delta_4 + \lambda_{42}dS_{2i} + \lambda_{43}dS_{3i} + \lambda_{44}dS_{4i} + \lambda_{45}dS_{5i} + \eta_{i4}
\end{aligned}
\tag{3.8}
$$

where $dS_{mi}, m = 1,\ldots,5$, is a binary variable taking the value 1 if the $i-th$ subject has been treated with dose m and 0 otherwise.

Tables (3.7) to (3.10) summarize the comparison results of the F-tests based on the maximum likelihood. It is clear that the dose effects do not affect the coefficient δ_{4i} (ratio of growth).

Table (3.7): Test of the dose effects

Soil type : SN					
Residuals	H_0	df_1	df_2	F-value	p-value
CR	$\delta_1 = 0$	1	86	22392.71	<.0001
CR	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	86	413.19	<.0001
CR	$\delta_2 = 0$	1	86	7575.57	<.0001
CR	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	86	50.30	<.0001
CR	$\delta_3 = 0$	1	86	49459.00	<.0001
CR	$\lambda_{32} = \lambda_{33} = \lambda_{34} = \lambda_{35} = 0$	4	86	125.68	<.0001
CR	$\delta_4 = 0$	1	86	13369.32	<.0001
CR	$\lambda_{42} = \lambda_{43} = \lambda_{44} = \lambda_{45} = 0$	4	86	0.22	0.9284
Soil type: SO					
CR	$\delta_1 = 0$	1	90	9771.469	<.0001
CR	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	90	962.119	<.0001
CR	$\delta_2 = 0$	1	90	2946.030	<.0001
CR	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	90	7.314	0.2956
CR	$\delta_3 = 0$	1	90	8051.971	<.0001
CR	$\lambda_{32} = \lambda_{33} = \lambda_{34} = \lambda_{35} = 0$	4	90	41.500	<.0001
CR	$\delta_4 = 0$	1	90	2307.871	<.0001
CR	$\lambda_{42} = \lambda_{43} = \lambda_{44} = \lambda_{45} = 0$	4	90	0.312	0.8694

Table (3.8): Test of the dose effects

Soil type : SN					
Residuals	H_0	df_1	df_2	F-value	p-value
CC	$\delta_1 = 0$	1	90	708.5883	<.0001
CC	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	90	18.7095	<.0001
CC	$\delta_2 = 0$	1	90	319.2148	<.0001
CC	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	90	2.7808	0.0314
CC	$\delta_3 = 0$	1	90	312.5887	<.0001
CC	$\lambda_{32} = \lambda_{33} = \lambda_{34} = \lambda_{35} = 0$	4	90	14.5650	<.0001
CC	$\delta_4 = 0$	1	90	69.6212	<.0001
CC	$\lambda_{42} = \lambda_{43} = \lambda_{44} = \lambda_{45} = 0$	4	90	0.0248	0.9988
Soil type: SO					
CC	$\delta_1 = 0$	1	90	723.224	<.0001
CC	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	90	2566.934	<.0001
CC	$\delta_2 = 0$	1	90	1810.788	<.0001
CC	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	90	414.047	<.0001
CC	$\delta_3 = 0$	1	90	200.098	<.0001
CC	$\lambda_{32} = \lambda_{33} = \lambda_{34} = \lambda_{35} = 0$	4	90	383.569	<.0001
CC	$\delta_4 = 0$	1	90	1.998	0.161
CC	$\lambda_{42} = \lambda_{43} = \lambda_{44} = \lambda_{45} = 0$	4	90	0.002	1.000

Table (3.9): Test of the dose effects

Soil type : SN					
Residuals	H_0	df_1	df_2	F-value	p-value
LG	$\delta_1 = 0$	1	90	3640.244	<.0001
LG	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	90	105.901	<.0001
LG	$\delta_2 = 0$	1	90	2151.809	<.0001
LG	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	90	135.663	<.0001
LG	$\delta_3 = 0$	1	90	877.636	<.0001
LG	$\lambda_{32} = \lambda_{33} = \lambda_{34} = \lambda_{35} = 0$	4	90	69.249	<.0001
LG	$\delta_4 = 0$	1	90	142.572	<.0001
LG	$\lambda_{42} = \lambda_{43} = \lambda_{44} = \lambda_{45} = 0$	4	90	0.029	0.9983
Soil type: SO					
LG	$\delta_1 = 0$	1	90	7252.241	<.0001
LG	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	90	192.773	<.0001
LG	$\delta_2 = 0$	1	90	4200.599	<.0001
LG	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	90	14.661	<.0001
LG	$\delta_3 = 0$	1	90	1041.979	<.0001
LG	$\lambda_{32} = \lambda_{33} = \lambda_{34} = \lambda_{35} = 0$	4	90	8.935	<.0001
LG	$\delta_4 = 0$	1	90	151.923	<.0001
LG	$\lambda_{42} = \lambda_{43} = \lambda_{44} = \lambda_{45} = 0$	4	90	0.015	0.9995

Table (3.10): Test of the dose effects

Soil type : SN					
Residuals	H_0	df_1	df_2	F-value	p-value
LS	$\delta_1 = 0$	1	86	63864.63	<.0001
LS	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	86	4011.56	<.0001
LS	$\delta_2 = 0$	1	86	14117.50	<.0001
LS	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	86	108.39	<.0001
LS	$\delta_3 = 0$	1	86	30350.03	<.0001
LS	$\lambda_{32} = \lambda_{33} = \lambda_{34} = \lambda_{35} = 0$	4	86	563.17	<.0001
LS	$\delta_4 = 0$	1	86	8927.35	<.0001
LS	$\lambda_{42} = \lambda_{43} = \lambda_{44} = \lambda_{45} = 0$	4	86	0.50	0.7327
Soil type: SO					
LS	$\delta_1 = 0$	1	90	2473.570	<.0001
LS	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	90	70.598	<.0001
LS	$\delta_2 = 0$	1	90	760.640	<.0001
LS	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	90	94.641	<.0001
LS	$\delta_3 = 0$	1	90	1186.032	<.0001
LS	$\lambda_{32} = \lambda_{33} = \lambda_{34} = \lambda_{35} = 0$	4	90	378.481	<.0001
LS	$\delta_4 = 0$	1	90	556.169	<.0001
LS	$\lambda_{42} = \lambda_{43} = \lambda_{44} = \lambda_{45} = 0$	4	90	0.637	0.6372

3.2.4. Analysis of accumulated carbon data

The same idea used in the previous analysis can be also used here to characterize the process of the accumulated carbon data analysis. This relationship was also characterized by modifying the model (1.10), already introduced in the first chapter, incorporating the heterogeneity of variance model. The essential information on the fit results is displayed in Figure 3.17 and listed in Table (3.11). It is again obvious that this model adequately describes the evolution of data, but again with some departures over the usual assumptions of normal i.i.d distribution of the model residuals.

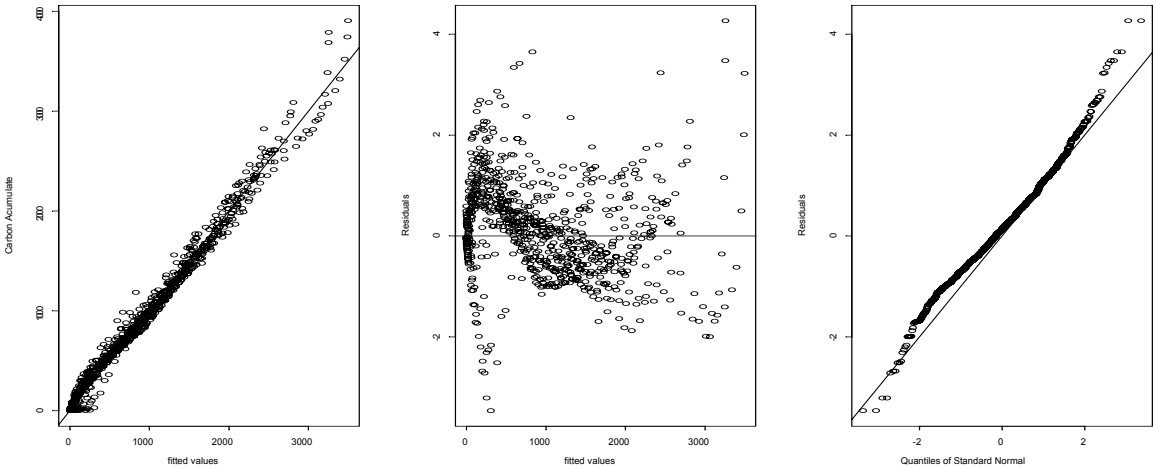


Figure 3.17: Essential features of the fit of the accumulated carbon model

Table (3.11): summary results of the accumulated carbon model

Fixed effects		Random effects	
Parameters	Values	Parameter	Values
δ_1	2319.503	D_{11}	292152.537711
SE	96.34060	D_{12}	2.3390893026
δ_2	0.026	D_{22}	0.0001960094
SE	0.00225	σ	28.20303149
		τ	0.1831883

To investigate the effects of different industrial residuals the carbon accumulated data, we relate the individual parameters to a covariate, doses, as follows[1]:

$$\begin{aligned} y_{ij} &= \delta_{i1} \cdot (1 - \exp(-\delta_{i2} \cdot t_{ij})) + e_{ij} \\ \delta_{i1} &= \delta_1 + \lambda_{12} dS_{2i} + \lambda_{13} dS_{3i} + \lambda_{14} dS_{4i} + \lambda_{15} dS_{5i} + \eta_{i1} \\ \delta_{i2} &= \delta_2 + \lambda_{22} dS_{2i} + \lambda_{23} dS_{3i} + \lambda_{24} dS_{4i} + \lambda_{25} dS_{5i} + \eta_{i2} \end{aligned} \tag{3.9}$$

where $dS_{mi}, m = 1,\ldots,5$ is again a binary variable taking the value 1 if the $i-th$ subject has received the level m of the dose factor and 0 otherwise.

Table (3.12): Test of the dose effects

Soil type : SN					
Residuals	H_0	df_1	df_2	F-value	p-value
CR	$\delta_1 = 0$	1	206	441236.8	<.0001
CR	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	206	622.3	<.0001
CR	$\delta_2 = 0$	1	206	8021.1	<.0001
CR	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	206	2.9	0.0232
CC	$\delta_1 = 0$	1	210	97504.44	<.0001
CC	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	210	1890.33	<.0001
CC	$\delta_2 = 0$	1	210	1335.88	<.0001
CC	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	210	70.06	<.0001
LG	$\delta_1 = 0$	1	206	60157.48	<.0001
LG	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	206	226.07	<.0001
LG	$\delta_2 = 0$	1	206	1149.48	<.0001
LG	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	206	59.08	<.0001
LS	$\delta_1 = 0$	1	206	104135.2	<.0001
LS	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	206	233.3	<.0001
LS	$\delta_2 = 0$	1	206	1919.1	<.0001
LS	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	206	1.0	0.4282

The parameters $\delta_i = \left(\delta_{i1},\ldots,\delta_{i4}\right)'$ for model (3.8) have the interpretation described in the first chapter, but the parameters $\lambda_{uv}, u = 1,\ldots,4;\ v = 2,\ldots,5$ and those for model (3.9) are the difference between groups under level of control dose (dose 0) and groups under other dose levels. This convenient codification of the dose factor is a special type of "contrast" (see. Venables and Ripley, 1997, §6.2, for other possible codification choices and more information on contrasts).

Table (3.13): Test of the dose effects

Soil type : SO					
Residuals	H_0	df_1	df_2	F-value	p-value
CR	$\delta_1 = 0$	1	96	31882.46	<.0001
CR	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	96	36.28	<.0001
CR	$\delta_2 = 0$	1	96	244.06	<.0001
CR	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	96	0.79	0.537
CC	$\delta_1 = 0$	1	100	7150.565	<.0001
CC	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	100	868.466	<.0001
CC	$\delta_2 = 0$	1	100	118.223	<.0001
CC	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	100	2.287	0.0652
LG	$\delta_1 = 0$	1	96	33319.84	<.0001
LG	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	96	497.53	<.0001
LG	$\delta_2 = 0$	1	96	763.23	<.0001
LG	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	96	28.21	<.0001
LS	$\delta_1 = 0$	1	96	52496.17	<.0001
LS	$\lambda_{12} = \lambda_{13} = \lambda_{14} = \lambda_{15} = 0$	4	96	1857.15	<.0001
LS	$\delta_2 = 0$	1	96	687.45	<.0001
LS	$\lambda_{22} = \lambda_{23} = \lambda_{24} = \lambda_{25} = 0$	4	96	20.21	<.0001

CONCLUSIONS: It is clear from Table (3.7) to (3.10) that the second term representing the ratio of growth of daily carbon is independent of doses under all industrial residuals and soil types, while other parameters are related. But in the accumulated carbon data, the ratio of growth, corresponding to parameter δ_{i2}, is highly significant except for the industrial residual "LS" under soil type "SN" and for the industrial residuals "CR" and "CC" for soil type "SO". The interpretation of the results of daily and accumulated carbon confirm the results reported in the Ph.D. dissertation of Dra. Montserrat.

3.2.5. Analysis of breast cancer data

Model (1.9), formulated in the first chapter, tries to reflect the primary interest of these studies, that is, to establish the tumoral growth dynamics and to relate it with risk factors in diet. The model is adequate to reflect the high observed variability in the dynamics of tumor growth: they can grow logistically, exponentially, or even flat and finally regress,

$$
\begin{aligned}
y_{ij} &= \frac{\delta_{1i}.\exp(t_{ij}-\delta_{2i})}{1+\exp\left(\frac{t_{ij}-\delta_{2i}}{\delta_{3i}}\right)} + e_{ij} \\
&\text{where:} \\
\delta_{1i} &= \delta_1 + \eta_{1i} \\
\delta_{2i} &= \delta_2 + \eta_{2i} \\
\delta_{3i} &= \delta_3 + \eta_{3i} \\
i &= 1,2,...,38 \\
j &= 1,2,...,25 \qquad .
\end{aligned}
\tag{3.10}
$$

where y_{ij} is the total tumor volume in rat i at time t_{ij}, δ_1, δ_2 and δ_3 are the coefficients of the model that express, respectively, an hypothetical maximum or asymptotic tumoral volume, the time of tumor appearance and the rate tumoral growth, and e_{ij} are residuals. The parameter δ makes the model sufficiently flexible in shape. When $\delta > 1$ it describes exponential growth. When $\delta < 1$ it is adequate to describe regressing tumors. $\delta = 1$ values define a logistic-like dynamics. The meaning of these parameters is better understood under mixed-effects model perspective. In El Halimi et al (2003) there is a detailed description of the initial steps of the modeling process. Next we introduce some tables to summarize some characteristics of the options finally taken as the most adequate to fit model (3.10), using the *nlme* function in S-plus. The "information table" lists some basic informations on the raw data and the model, including the number of observations, subjects, parameters, dependent and subject variables, random effects, relevant distributions, and type of optimization. They are useful to check that we have specified our data set and model correctly.

Table (3.14): The information table from the analysis

Information on data	
Data set	palpSZ
Dependent variable	Volume
Observations Used	950
Observations Not Used	550 (all nulls)
Total observations	1500
Subject variable	Rat
Subjects Used	38
Subjects Not Used	22
Total subjects	60
Max Obs. Per subject	25
Distribution of Dep. variable	Normal
Random effects	$\eta_i = \left(\eta_{i1}, \eta_{i2}, \eta_{i3}\right)'$
Distribution of random effects	Normal
Optimization Technique	Newton Raphson
Estimation method	EM algorithm LB-method Two-stage methods

.

Table (3.15):

Initial values of the fixed parameters in the fitting process		
δ_1	δ_2	δ_3
1.9	9.9	1.1

Table (3.16):

Goodness of fit information	
Log Likelihood	-1483.757
AIC (smaller is better)	2987.514
BIC (smaller is better)	3036.079

The "goodness of fit information" table lists the final maximized value of the log likelihood as well as the information criteria of Akaike and Schwarz. These statistics can be used to compare different non-linear mixed models.

Table (3.17):

Fixed Parameter Estimates					
Parameter	Estimate	Std.Error	DF	t- value	Pr>|t|
δ_1	4.7272642	1.094135	910	4.32055	<.0001
δ_2	14.5404748	0.939590	910	15.47535	<.0001
δ_3	0.9825508	0.018387	910	53.43605	<.0001
Confidence intervals					
Parameter		Alpha	Lower		Upper
δ_1		0.05	2.5833368		6.871192
δ_2		0.05	12.6993741		16.381576
δ_3		0.05	0.9465211		1.018580

The "Fixed Parameter Estimates" table lists the maximum likelihood estimates of three parameters and their approximate standard errors computed using the Newton algorithm. Approximate t-values and confidence intervals limits are also provided, jointly with degrees of freedom. All parameter are highly significant in the model, that is, they should be included.

Table (3.18):

Random Components Estimates						
D_{11}	D_{22}	D_{33}	D_{12}	D_{13}	D_{23}	σ_e
45.241877	15.293208	0.002880	-2.084413	-0.120253	0.000666	0.93970098

The relative magnitude of the estimates of D_{11} and D_{22} indicates that the rat to rat variation is considerable while the estimated correlation between the random effects associated with δ_1 and δ_2, $\frac{D_{12}}{\sqrt{D_{11}D_{22}}}$, is very small (0.003). Thus a covariance matrix with a more elaborated structure can be useful for the interpretation of the random variation and to improve the model, avoiding the use of unnecessary parameters. The scatter plot matrix displayed in the Figure 3.18 suggests that all random effects are uncorrelated (between individuals).

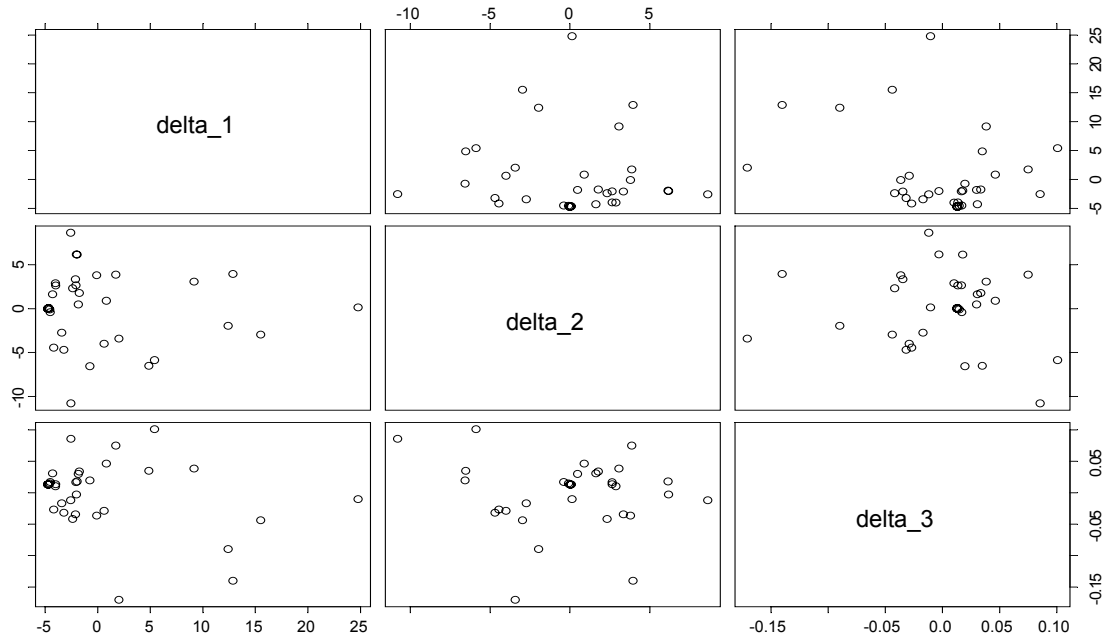


Figure 3.18: Scatter plot of estimated random effects of breast cancer data

As a consequence, we assume that the $\eta_i = (\eta_{i1}, \eta_{i2}, \eta_{i3})'$ are independent and identically distributed across individuals, with common distribution $N(0, \sigma^2 D)$. Thus, the covariance matrix for model (3.10) can be modified as follows:

$$D^{(d)} = \begin{pmatrix} D_{11} & 0 & 0 \\ 0 & D_{22} & 0 \\ 0 & 0 & D_{33} \end{pmatrix}. \tag{3.11}$$

Also we observe from Table (3.18) that the parameters η_{1i} and η_{2i} have approximately equal variance and are independent of the η_{3i}. This can be rephrased by saying that the covariance matrix of the random effects D may be partitioned in four blocks as follows:

$$D^{(B)} = \left[\begin{array}{cc|c} D_{11} & 0 & 0 \\ 0 & D_{11} & \\ \hline 0 & & D_3 \end{array}\right]. \tag{3.12}$$

We can test the adequacy of this structured covariance matrix by comparing the accuracy of the different fits, as shown in the following table:

Table (3.19): Likelihood ratio test of the different structure of matrix D

	DF	AIC	BIC	Log-Lik	L.Ratio	P_value
[Model (3.10): D]	10	2986.050	3034.615	-1483.025		
[Model (3.11): $D^{(d)}$]	7	2979.358	3013.354	-1482.679	0.6917187	0.8752
[Model (3.12): $D^{(B)}$]	6	2978.922	3008.061	-1483.461	1.563813	0.2111

The large p-values for the likelihood ratio test and the smaller AIC, BIC and DF (degrees of freedom) values for model (3.12) suggest that it should be preferred: the more parsimonious structure (less parameters) of the covariance matrix $D^{(B)}$ is not associated to a loss in fit and to a significantly different model.

The second question of interest for this data is whether the main response profiles and the variability among subjects (rats) differ significantly among diets. This was done by testing the differences between Diets 1 and 2 $\left(\delta_{1,2},\ \delta_{2,2} \text{ and } \delta_{3,2} \text{ in (3.13)}\right)$ and between Diets 1 and 3 $\left(\delta_{1,3},\ \delta_{2,3} \text{ and } \delta_{3,3} \text{ in (3.13)}\right)$ for each one of the fixed parameters using both, the t-test and the F-test.

$$
\begin{aligned}
y_{ij} &= \frac{\delta_{i1}\cdot\exp(t_{ij}-\delta_{i2})}{1+\exp(\frac{t_{ij}-\delta_{i2}}{\delta_{i3}})}+e_{ij} \\
&\text{where:} \\
\delta_{i1} &= \delta_1+\delta_{1,2}d_{i,2}+\delta_{1,3}d_{i,3}+\eta_{i1} \\
\delta_{i2} &= \delta_2+\delta_{2,2}d_{i,2}+\delta_{2,3}d_{i,3}+\eta_{i2} \\
\delta_{i3} &= \delta_3+\delta_{3,2}d_{i,2}+\delta_{3,3}d_{i,3}+\eta_{i3} \\
\eta_i &\sim \mathrm{N}(0,D),\ D=\mathrm{diag}(D_{11},D_{11},D_{33})
\end{aligned}
\qquad (3.13)
$$

In model (3.13) $d_{i,2}$ is the binary variable taking the value 1 if the $i-th$ individual receives treatment 2 (Diet2), $d_{i,3}$ is a binary indicator variable for treatment 3 (Diet3), and δ_1, δ_2 and δ_3 are respectively fixed effects for rats under treatment group control (Diet1). The other parameters, $\delta_{1,2}$, $\delta_{2,2}$ and $\delta_{3,2}$, represent average effects between rats under Diet 2 and

rats under Diet 1 and $\delta_{1,3}$, $\delta_{2,3}$ and $\delta_{3,3}$ represent average effects between rats under Diet 3 and rats under Diet 1.

Table (3.20):

Fixed Parameter Estimates					
Parameter	Estimate	Std.Error	DF	t- value	Pr>|t|
δ_1	4.06110	1.875074	904	2.16583	0.0306
$\delta_{1,2}$	-0.17801	2.392791	904	-0.07439	0.9407
$\delta_{1,3}$	2.70799	2.540790	904	1.06580	0.2868
δ_2	15.36709	2.406930	904	6.38452	<.0001
$\delta_{2,2}$	-0.75437	2.917864	904	-0.25853	0.7961
$\delta_{2,3}$	-0.04564	3.135983	904	-0.01455	0.9884
δ_3	0.97900	0.039056	904	25.06648	<.0001
$\delta_{3,2}$	-0.01999	0.048646	904	-0.41103	0.6811
$\delta_{3,3}$	-0.02116	0.051783	904	-0.40867	0.6829

The large p-value for the t-tests associated with terms involving the Diet suggest that they are not needed in the model, that is, Diet is a poor explanatory variable for the fixed effects.

Table (3.21): F-test of the Diets effects

Hypothesis H_0	df_1	df_2	F-value	$p-value$
$\delta_1 = 0$	1	904	65.042	<.0001
$\delta_{1,2} = \delta_{1,3} = 0$	2	904	0.850	0.4277
$\delta_2 = 0$	1	904	282.518	<.0001
$\delta_{2,2} = \delta_{2,3} = 0$	2	904	0.115	0.8913
$\delta_3 = 0$	1	904	2515.231	<.0001
$\delta_{3,2} = \delta_{3,3} = 0$	2	904	0.105	0.9007

These results are in agreement with the large p-values for the F-test. Given the sample sizes and the variability in these experiments, there are no evidences supporting the existence of a possible relation between diet and the fixed effects parameters representing the dynamics of tumour growth. Possibly, these terms could be eliminated from the model.

We can also investigate the distribution of random effects between treatment groups by means of the box plots of random effects for each treatment group:

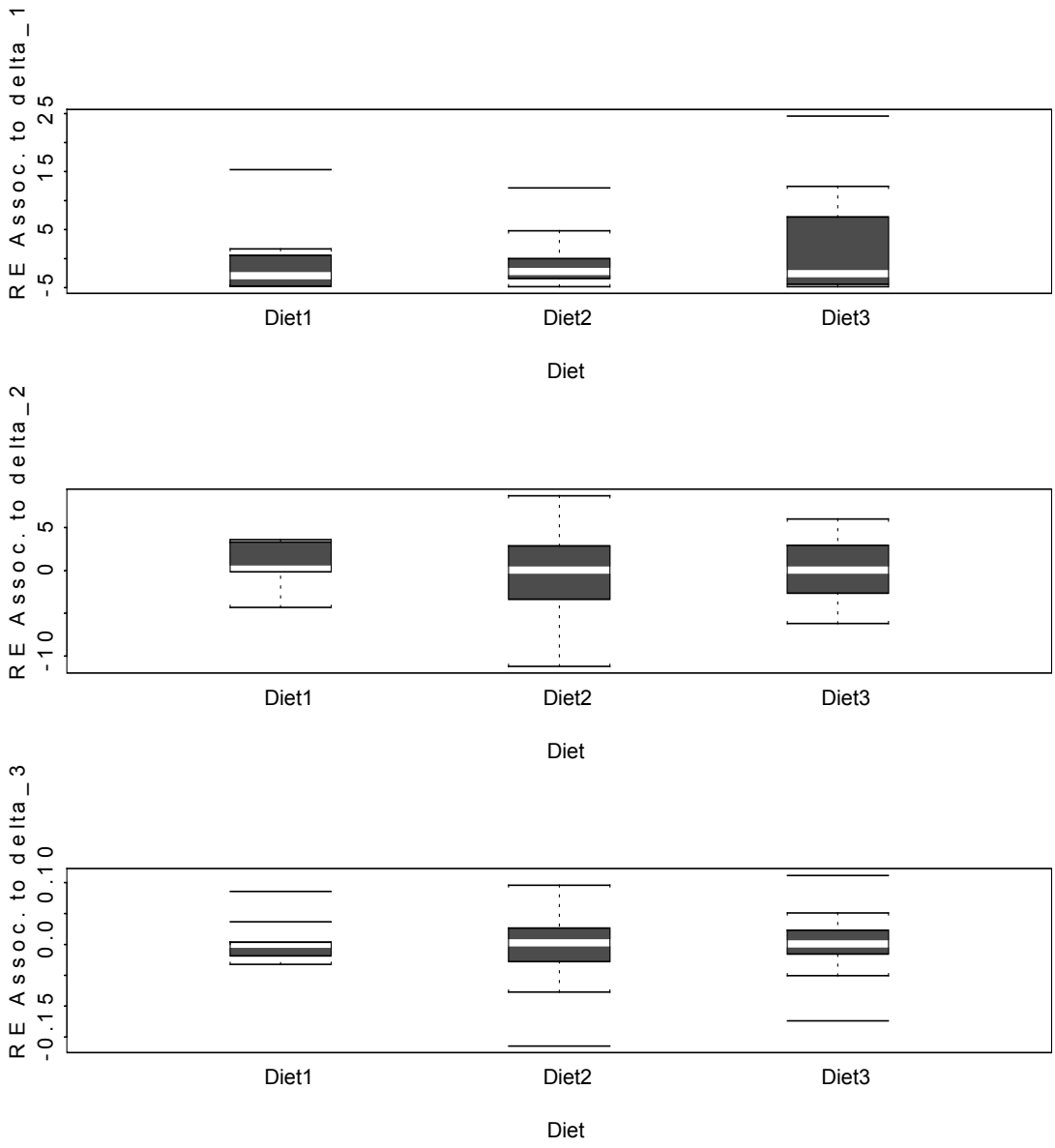


Figure 3.19: boxplots of estimated random effects versus the diet factor

Figure 3.19 suggests that the variability of the random effects corresponding to the "maximal volume" parameter and to the "time to the 50% of the maximal volume" depend on the treatment factor (Diet). This suggests an alternative model (3.14) in which the random effects, η_{1i}, η_{2i}, η_{3i}, can be reparameterised in the following form:

$$\begin{cases} \eta_{1i} = \varphi_{0,i} + \varphi_{02,i} d_{2i} + \varphi_{03,i} d_{3i} \\ \eta_{2i} = \gamma_{0,i} + \gamma_{02,i} d_{2i} + \gamma_{03,i} d_{3i} \\ \eta_{3i} = \lambda_{0i} \end{cases} \tag{3.14}$$

Where $\varphi_{0,i}$, $\gamma_{0,i}$, and $\lambda_{0,i}$, are the random parameter "intercepts" under treatment group 1 (Diet1) for individual i, and $\varphi_{0q,i}$, $\gamma_{0q,i}$ and $\lambda_{0q,i}$, $q = 2,3$ are a difference in "intercept" between rat under treatment group q (Diet q) and rat under treatment group 1 (Diet 1), for the individual i.

Incorporation of the treatment group as a covariate in the random effects associated to δ_1 and δ_2 of model (3.14), shows that the systematic among-rat variability considerably reduces the random component ($D_{11} = 43.98973808$ to 17.87568847 in model (3.12)).

Table (3.22) reflects the result of the comparison test between model (3.12) and model (3.14). The results indicate that the preferred model, according to the AIC, BIC and log likelihood criteria is (3.14).

Table (3.22): Likelihood ratio test of the model (3.12)

	DF	AIC	BIC	Log-Lik
Model (3.13)	6	2978.922	3008.061	-1483.461
Model (3.15)	6	2978.413	3007.551	-1483.206

The next step in this analysis should be an examination of the residuals from model (3.14), conditionally on a set of random effects, to assess how well the model fits the data. The pattern of increasing variability in the residuals, displayed in the Figure 3.20, suggests that a possible heteroscedasticity between Diets should be considered (e.g. conditionally to treatment group, Diet, the variance is proportional to the power of the time, but with different power parameters according to different diets, $\tau = (\tau_1, \tau_2, \tau_3)'$) as shown in the model (3.15).

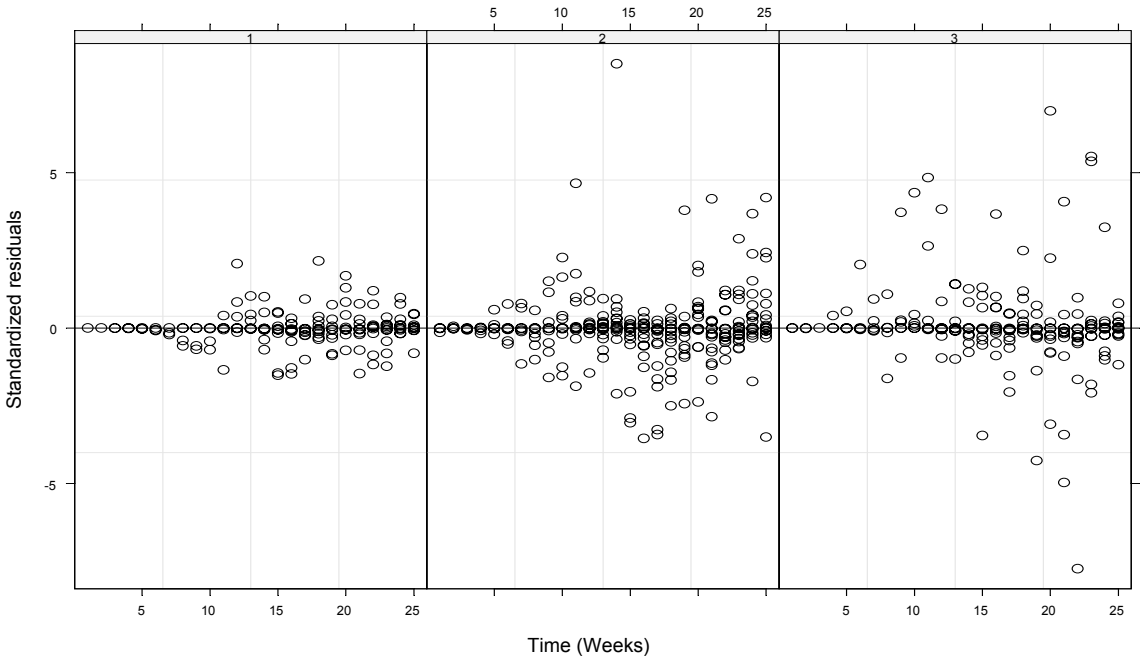


Figure 3.20: Plots of standardized residuals versus time by different Diets.

$$\text{var}(e_{ij}) = \begin{cases} \sigma^2 t^{2\tau_1} & \text{for treatment group 1 } (\textit{Diet 1}) \\ \sigma^2 t^{2\tau_2} & \text{for treatment group 2 } (\textit{Diet 2}) \\ \sigma^2 t^{2\tau_3} & \text{for treatment group 3 } (\textit{Diet 3}) \end{cases} \tag{3.15}$$

The adequacy of model (3.15) is indirectly confirmed by the results shown in Table (3.23)

Table (3.23): Likelihood ratio test of heteroscedasticity

	DF	AIC	BIC	Log-Lik	L.Ratio	P_value
Model (1.15)	6	2978.413	3007.551	-1483.206		
Model (3.16)	9	1713.787	1757.496	-847.894	1270.625	<.0001

As expected, there is a highly significant increase in the likelihood associated to the inclusion of the variance function model (3.15).

On other hand, the model of the variance function leads to a dramatic reduction of the standard deviation of residuals, $\hat{\sigma} = 0.01085069$ versus $\hat{\sigma} = 0.9387914$ in model (3.14).

Table (3.24) indicates that the parameters of the variance function are very precisely estimated.

Table (3.24): Confidence intervals of the variance function parameters

Parameter	Lower	Estimation	Upper
τ_1	1.277690	1.375705	1.473720
τ_2	1.584872	1.683407	1.781942
τ_3	1.697739	1.803426	1.909112

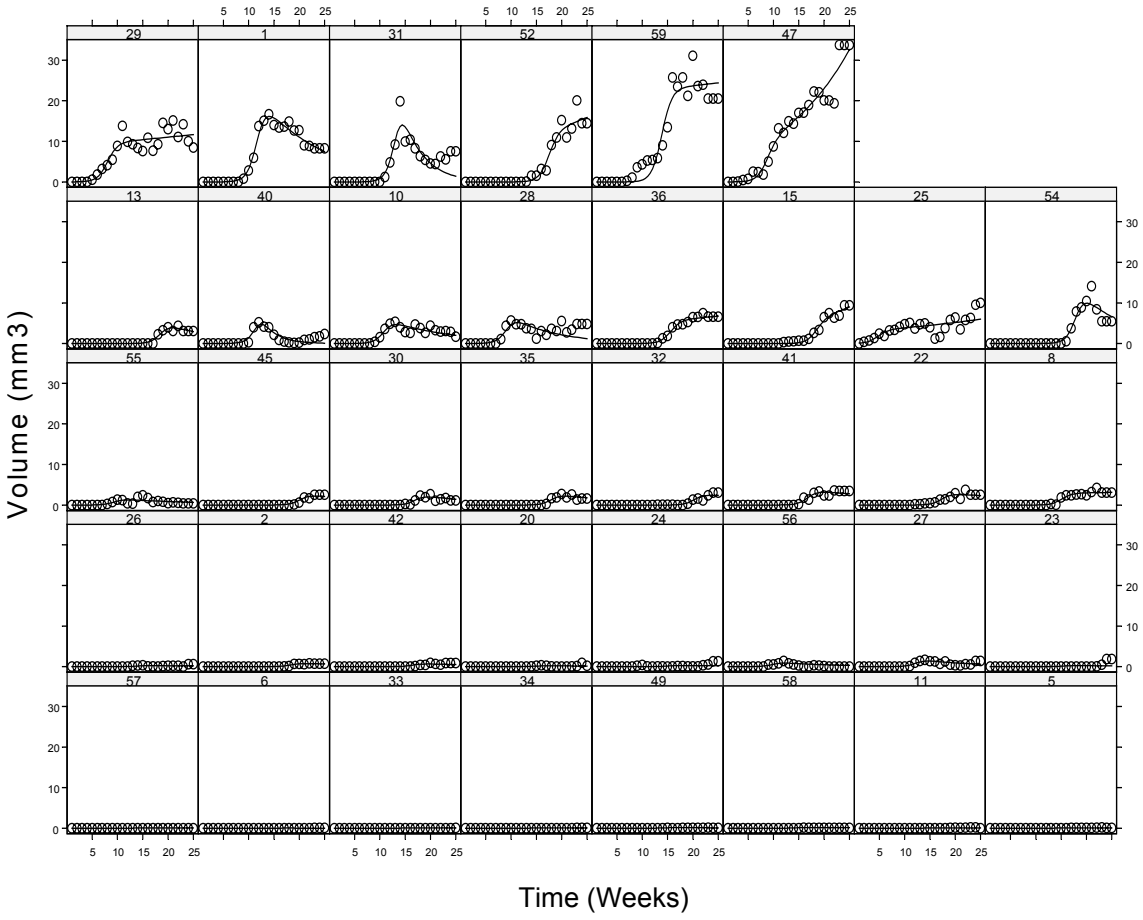


Figure 3.21: Plot of predicted versus time by subject corresponding to the fit of the model (3.15)

A final assessment of the heteroscedastic version of the non-linear mixed effects model (3.15) is provided by the plot of the predictions by subject displayed in Figure 3.21. The predicted values closely match the observed tumoral volume, attesting the adequacy of the model.

Because the tumours are measured sequentially over time on the same subjects (rats), the within–subjects observations are likely to be correlated. Thus, it may be necessary to introduce some correlation structure to obtain a satisfactory fit of the model. One candidate correlation structure for modeling the within group (inside rat) error covariance structure, suggested by the empirical autocorrelation function displayed in the Figure 3.22, is a moving average model of order 1, which is denoted MA (1), in which it is assumed that the current observation is a linear

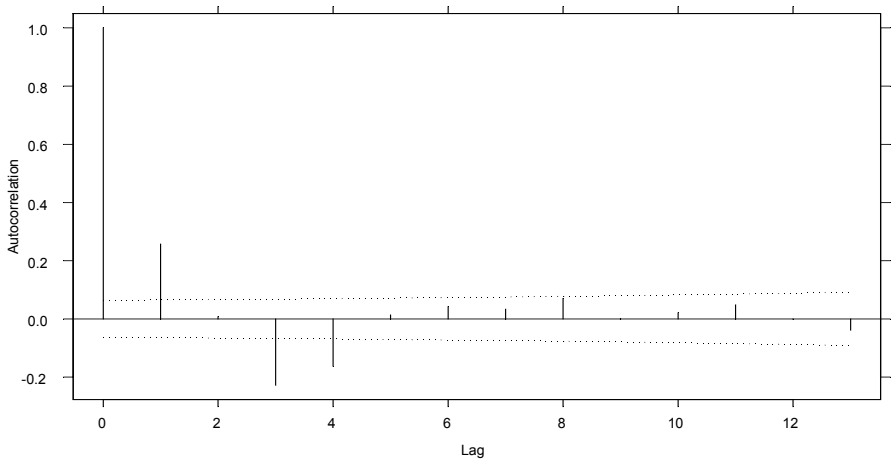


Figure 3.22: Empirical autocorrelation function corresponding to the standardized residuals.

function of independent and identically distributed noise terms,

$$\rho = cor(e_t, e_{t+1}) = \frac{\theta}{\theta^2 + 1} \tag{3.16}$$

The confidence interval on the correlation parameter θ, (0.2055158, 0.3678353), is bounded away from zero, indicating that the MA (1) model provides a significantly better fit. We can confirm it with the likelihood ratio test:

Table (3.25): Likelihood ratio test of heteroscedasticity

	DF	AIC	BIC	Log-Lik	L.Ratio	P_value
Model (1.15)	9	1713.787	1757.496	-847.8937		
Model (3.16)	10	1634.799	1683.363	-807.3993	80.98881	<.0001

As expected, the likelihood ratio test clearly rejects the assumption of independence.

The last part of this analysis is to check whether the underling distributional assumptions appear to be valid for the data.

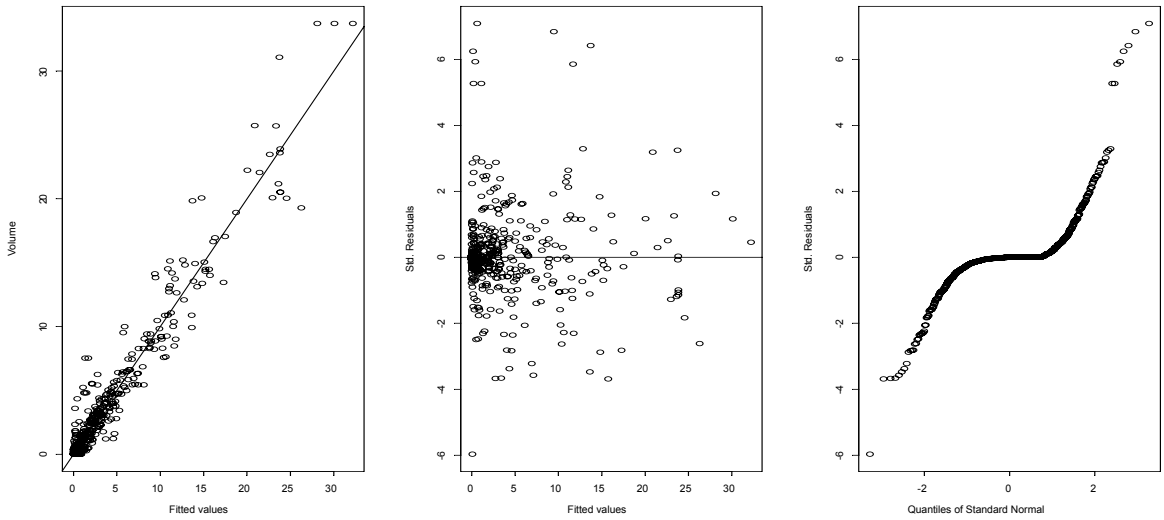


Figure 3.23: Normal plot of residuals for the NLME fit of breast cancer data

The first plot, observed versus fitted values, indicates that the linear model does a reasonable job in explaining the growth volume. The points fall relatively close to the $y = x$ line, indicating a reasonable agreement between the fitted and observed values. The second plot, residual versus fitted values, suggests the presence of some outliers in the data. The remaining residuals appear to be homogeneously scattered around the $y = 0$ line. The final plot, indicates clearly a departure from the assumption of normality of the within-group errors. That makes questionable the parametric statistical inference of the NLME models. To avoid this problem, we can use other alternative methods that are not based on the above assumption, for example the two-stage methods like the standard two-stage method (STS) and global two-stage (GTS) described in Davidian and Giltinan (1995). The parameter estimates obtained from refitting the model with these last methods, for each treatment group (Diet), separately and without any restrictions on the distributional assumptions are reported together with those of the LB-method in the following table:

Table (3.26) Summarizes results for fitting the breast cancer model via LB, STS and GTS methods.

Fixed effects				
Treatment group (Diets)	Parameters	LB-Method	STS Method	GTS Method
Diet 1 (Control)	$\delta_1^{(d1)}$	4.070928	5.149776	5.1674889
	$\delta_2^{(d1)}$	15.36477	17.36185	16.5434671
	$\delta_3^{(d1)}$	0.9808802	0.8504108	0.9315751
Diet 2	$\delta_1^{(d2)}$	3.868072	5.00464	5.0236048
	$\delta_2^{(d2)}$	14.63375	14.92942	14.2788336
	$\delta_3^{(d2)}$	0.9582356	0.8141487	0.9296444
Diet 3	$\delta_1^{(d3)}$	6.658124	7.3835	7.3986007
	$\delta_2^{(d3)}$	15.75488	16.30708	15.7686305
	$\delta_3^{(d3)}$	0.9717476	0.8908591	0.9508061

For these data, the LB- method based on the maximum likelihood procedure and the two-stage methods (STS and GTS) give a surprising discrepancy between estimates in all parameters (fixed and random components). These results illustrate that non-linear mixed models are very sensitive to the choice of the estimation method, specially in their random components.

In chapters 6 and 7 we compare some resampling inferential methods, adequate for the two stage methods, with the parametric methods introduced and applied in the first part of this monography.

Table (3.27): covariance matrix for each Diet separately using LB, STS, GTS.

Variance covariance matrix: LB- method

Diet 1

$$\begin{pmatrix} 38.821 & -7.539 & -0.089 \\ -7.539 & 7.085 & 0.050 \\ -0.089 & 0.050 & 0.002 \end{pmatrix}$$

Diet 2

$$\begin{pmatrix} 17.836 & -6.309 & -0.066 \\ -6.309 & 29.216 & -0.058 \\ -0.066 & -0.058 & 0.004 \end{pmatrix}$$

Diet 3

$$\begin{pmatrix} 87.802 & -1.723 & -0.063 \\ -1.723 & 8.403 & -0.064 \\ -0.063 & -0.064 & 0.003 \end{pmatrix}$$

Variance covariance matrix: STS- method

Diet 1

$$\begin{pmatrix} 43.720 & -20.517 & 0.585 \\ -20.517 & 21.938 & -0.771 \\ 0.585 & -0.771 & 0.058 \end{pmatrix}$$

Diet 2

$$\begin{pmatrix} 28.618 & -12.381 & 0.476 \\ -12.381 & 29.866 & -0.800 \\ 0.476 & -0.800 & 0.117 \end{pmatrix}$$

Diet 3

$$\begin{pmatrix} 98.041 & -5.560 & 0.523 \\ -5.560 & 16.521 & -0.460 \\ 0.523 & -0.460 & 0.070 \end{pmatrix}$$

Variance covariance matrix: GTS- method

Diet 1

$$\begin{pmatrix} 50.296 & -18.250 & 0.112 \\ -18.250 & 19.647 & -0.306 \\ 0.112 & -0.306 & 0.031 \end{pmatrix}$$

Diet 2

$$\begin{pmatrix} 30.002 & -8.865 & -0.194 \\ -8.865 & 31.848 & -0.295 \\ -0.194 & -0.295 & 0.047 \end{pmatrix}$$

Diet 3

$$\begin{pmatrix} 106.852 & -0.822 & -0.028 \\ -0.822 & 20.316 & -0.334 \\ -0.028 & -0.334 & 0.037 \end{pmatrix}$$

Chapter 4

4. Validation of the parametric Approach

4.1. Introduction

Some simulation studies were carried out with the purpose of testing the robustness of parametric inference based on linearization in non-linear mixed models. In a first series of simulations, the consequences of an inappropriate assumption of normality were investigated. Another set of simulations was devoted to study the consequences of inappropriate assumptions on the structure of the covariance matrix of the random effects and/or on the correlation structure of the residuals. Two models or "scenarios" were simulated. The first one is based on the data and experiments on breast cancer, introduced in chapters 1 and 3. The other is based on a classical data set in mixed-models applications: the Soybean genotypes data described in Davidian and Giltinan (1995) and Pinheiro and Bates (2000).

4.2. Simulation study on distributional assumptions

The parametric approach to non-linear mixed-effects modeling using the LB-method (Lindstrom and Bates method) is, essentially, based on the standard assumption of normality of the errors and random effects. But, as some of the analyses presented in the preceding chapter suggest, this assumption may not always be realistic. In fact, the residuals and random effects are not directly observed, which makes difficult to verify any hypothesis concerning their distribution. These violations of the assumptions of the model pose questions on the validity of the inferences made during the modeling process.

4.2.1. Breast cancer model simulation

We generated series of 1000 data tables according to the model described in chapters 1 and 3 and given by:

$$y_{ij} = \frac{\delta_{1i} \cdot \exp(t_{ij} - \delta_{2i})}{1 + \exp((t_{ij} - \delta_{2i}) / \delta_{3i})} + e_{ij} \tag{4.1}$$

$$\begin{cases} \delta_{1i} = \delta_1 + \eta_{1i} \\ \delta_{2i} = \delta_2 + \eta_{2i} \\ \delta_{3i} = \delta_3 + \eta_{3i} \end{cases} \tag{4.2}$$

where the observations y_{ij} correspond to the simulated tumoral volume of subject i (i=1,...,m=38) at time t_{ij} (j=1,...,n_i=25), the random effects $\eta_i = (\eta_{1i}, \eta_{2i}, \eta_{3i})'$ are i.i.d. $(0,D)$ and the e_{ij} are independent of the η_i, with zero mean and variance σ^2.

Based on the results of the real experiments, the "known" population values for the fixed effects were taken as $(\delta_1, \delta_2, \delta_3)' = (5.051236, 13.86703, 0.8486555)'$. These values were chosen near to the estimated values given by the S-plus 2000 implementation of the LB-method (*nlme* function) and according to the maximum likelihood (ML) and restricted maximum likelihood (REML) variants of the estimation procedure. The random factor vectors, $(\eta_{1i}, \eta_{2i}, \eta_{3i})'$, were generated according to four possible marginal distributions (always the same for all three components): Normal (N), Uniform (U), Exponential (E) or Gamma (G). Its dependence structure was always characterized by the following covariance matrix:

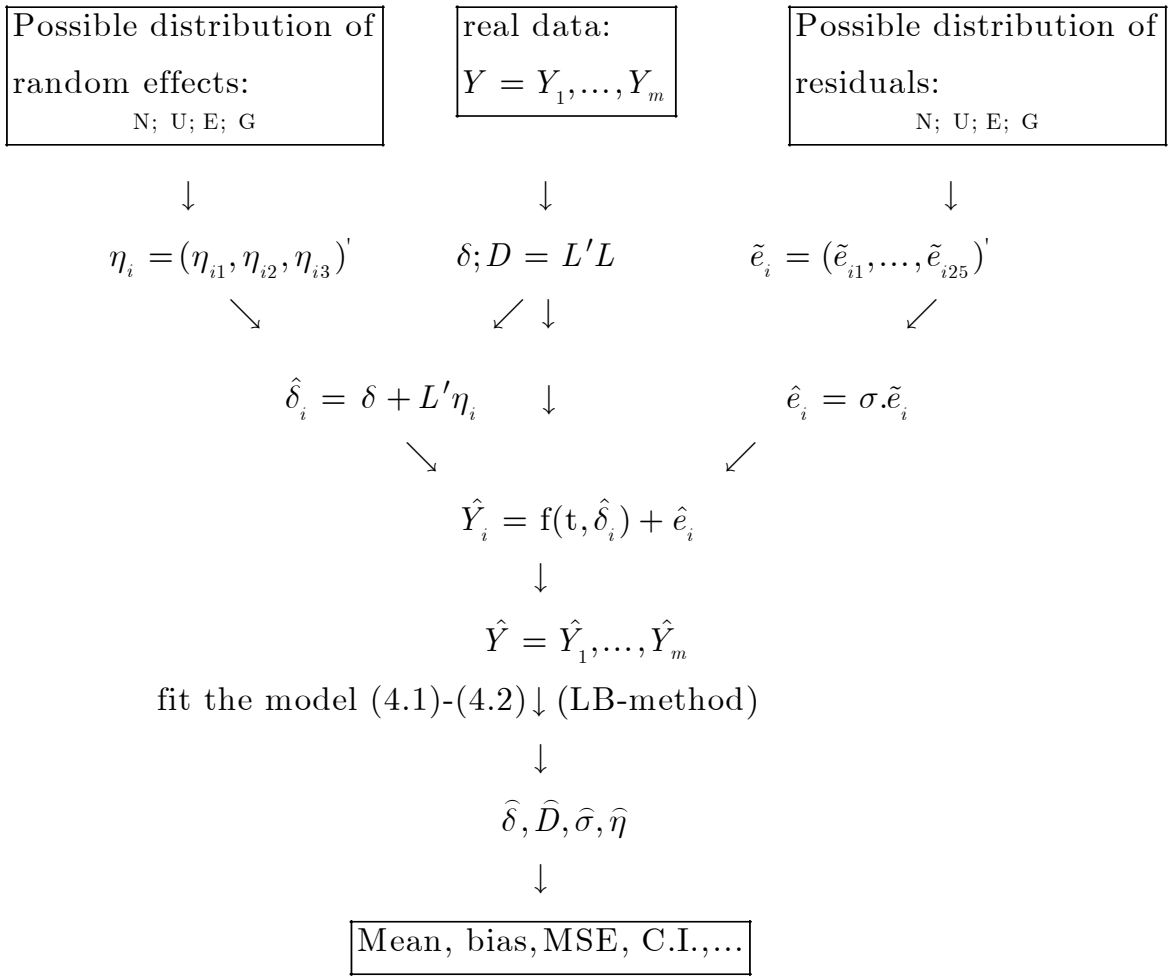
$$D = \begin{pmatrix} 51.88633 & -1.0498620 & -0.05460614 \\ & 15.8465000 & -0.04587260 \\ & & 0.01362791 \end{pmatrix}$$

also estimated from experimental data. The errors e_{ij} were also generated according to the same four possible distributions: Normal (N), Uniform (U), Exponential (E) or Gamma (G), always characterized by a dispersion parameter of σ=0.939921.

Each one the generated data tables was processed to fit a nonlinear mixed model like (4.1)-(4.2) via the LB-method, maintaining the same starting values for both, the ML and the REML estimation procedures. The resulting set of (simulated) parameter estimates for each series was used to compute the summaries described in the next section and used to evaluate the

performance of the inferential methods, like the true coverage of the confidence intervals or the true estimation biases.

The following diagram illustrates the process of generation of each series of 1000 pseudo data:



Inspired in the real data, the values described above for the parameters δ, D and σ were established as the "population" or "true" parameters of the simulation. L stands for the Cholesky decomposition of the "true" covariance matrix D. The transformation $L'\eta$ transformed the random effects generated as three initially independent values, $\eta = (\eta_1, \eta_2, \eta_3)'$, according to an standardized version (zero mean, unit variance) of one of the possible distributions (N, U, E or G), in values sampled from a random vector with covariance matrix D. The addition of the preceding random effects to the fixed effects produced a simulated set of parameters of the non-linear model. The residuals or errors were generated in a similar way, first as i.i.d. standardized

values and subsequently converted to values with standard deviation σ. The final transformation $\hat{Y} = \mathrm{f}(\mathrm{t}, \hat{\delta}) + \hat{e}$, where f stands for the model function (4.1), provided each simulated data set.

4.2.1.1. Results for the estimations of the fixed effects parameters: ML- Method

Figure 4.2 contains box-plots for the fixed effect parameter estimates and their corresponding standard deviations, from all simulation runs. For each box-plot, the first character in the abscises axis represents the simulated distribution for the random effects, and the second character the simulated distribution for the residuals, e.g. NN stands for "normal random effects /normal residuals", UE "uniform random effects/ uniform residuals", EE "exponential random effects/ exponential residuals" and so on. The horizontal continuous straight line represents the true value of the corresponding parameter. Thus, in this way, the box-plots give a graphical idea of the bias, standard deviation and dispersion of the fixed effects estimators, under different possible distributions.

If $\hat{\delta}_k$ stands for a fixed effect or parameter estimate at the k-th simulation replication and δ_T for the true value of the parameter, the summary measures for the estimators of the fixed effects are defined as:

MEAN: denotes the average of each series of 1000 estimates, $\frac{1}{1000}\sum_{k=1}^{1000}\hat{\delta}_k$,

BIAS: corresponds to the simulation estimate of bias, that is, MEAN – δ_T, the bias of a statistic indicates, on average, how much the estimator will over- or underestimate the “true” parameter value.

$\mathrm{C.I}_{\mathrm{Bias}}$: denotes approximate confidence intervals for bias, $\mathrm{BIAS} \pm \mathrm{Z}_\alpha \sqrt{\frac{S^2_{\hat{\delta}}}{1000}}$, where Z_α is the normal critical value at a 95% confidence level and $S^2_{\hat{\delta}} = \sum_{k=1}^{1000}\left(\hat{\delta}_k - \mathrm{MEAN}\right)^2 / 999$,

MSE: is the mean squared error, $\sum_{k=1}^{1000}\left(\hat{\delta}_k - \delta_T\right)^2 / 1000$, is a measure of its accuracy that takes into account both bias and the standard error.

COVERAGE-PROBABILITY: denotes the observed coverage of t-based 95% confidence intervals computed using the model-based standard errors and t-distribution critical values based on 910 degrees of freedom (the intervals provided by the *nlme* function),

PROB. LOW (PROB. UPP): denotes the proportions of δ_T lower (upper) than the lower (upper) bound of the confidence intervals (that is, they correspond to non-coverage probabilities), and <<Average width>> denotes the arithmetic average of the 1000 observed lengths of the asymptotic intervals described above. The precision of the Monte Carlo estimates made with these measures will increase with an increasing number of simulation repetitions.

Let us comment the Tables (4.1) to (4.3) and Figure 4.2. The means over 1000 simulations of δ_1 are close to the true value for all combinations of distributions, the biases are considerably small and significantly different from 0, and the Mean Square Error (MSE) is nearly always the same under all conditions, according to Table (4.1). The variability pattern shown in the left-hand plot of Figure 4.2 is quite similar for all distributions, even under large departures from normality. On the other hand, the measures of precision of the δ_1 estimate (right part of the plot), $\hat{\sigma}_{\hat{\delta}_1}$, indicated as "sd_delta_1" seem to be more erratic when the random effects deviate from normality. The true coverage probabilities, compared with the nominal value of 0.95 confirm that they are acceptably robust for the δ_1 parameter, with coverages ranging from 95.6% to 88.8%. The maximum coverage corresponds to the UG case and the worst coverage to the EN and the GN cases. The NN case has 93.6% coverage. In all cases, we remark that all simulation runs produce a very similar results in terms of the width of the intervals. Additionally, the intervals are appreciably equitailed (in the sense that the non coverage cases are nearly symmetrically distributed, approximately one half of the times the confidence interval is at left of the true parameter value, and the other half of the times it is at right), with asymmetric random effects (exponential or gamma) propitiating the last type of asymmetry (intervals underestimating the true parameter value).

Table (4.2) provides a summary of simulation experimental results for the sampling distribution characteristics of the starting moment for tumour growth parameter, δ_2 . In contrast with the

previous parameter, the values presented in this Table and in Figure 4.2 demonstrate a negative bias, indicating that on average they will underestimate the "true" value. The observed bias when the random effects are generated from asymmetric distributions (Exponential or Gamma) is much greater (around four times in absolute value) than the observed bias for the case of symmetric distributions (Normal or Uniform). The same tendency is observed for the MSE values. For this parameter, the parametric confidence intervals are clearly not robust, and even not adequate for the normal case, with coverages ranging from 89% (UE) to 50.2% (GU), with a 86% of coverage for the NN case. In correspondence with the coverages, their width ranges from 2.56 for the UN case to 2.03 for the EN case. The worst coverages clearly correspond to the asymmetric (E and G) distributions for the random effects, and are associated to a strong tendency for the intervals to underestimate the true parameter value, many times its upper bound lies at left of the true parameter value (47% for the GN case).

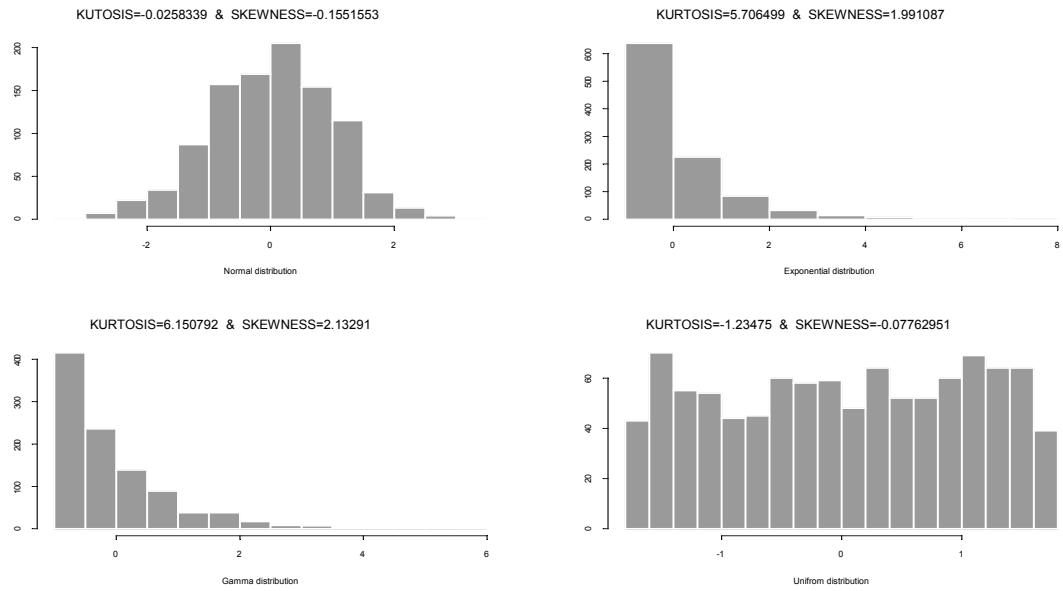
Table (4.3) contains simulation summary statistics for the speed in tumor growth parameter, δ_3. Again, the bias and MSE values are higher when the distribution of the random effects is exponential or gamma. The true confidence interval coverages demonstrate a poor degree of robustness (and poor adequacy even in the normal case), with coverages from 77.5% (UN) to 65.3% (EN) with 75.2% for the NN case. The width of these intervals shows a similar tendency and, again, the worst coverages are associated to the asymmetrical E and G distributions for the random effects, now with a high tendency of the intervals to overestimate the true parameter value. In the worst case (EN), 32.5% of the times the lower bound of the intervals lies at right of the true parameter value.

Finally, all the confidence intervals for the fixed-effects parameters seem to be more affected by the distribution of the random effects than by the distribution of the residuals, with a tendency according to the skewness and asymmetry of these distributions: from best to worse, the performance is associated to the order U, N, G, E for the distribution of the random effects.

Another interesting aspect of the simulation studies is the possible relationship between some common statistical measures, such as skewness (measure of symmetry) and kurtosis (measure of

a type of departure from normality) of the distribution of random effects and errors, and the coverage probabilities of the interval estimates of parameters.

In statistical literature, the notion of skewness of a distribution is related to a symmetry property. The most commonly used measure of skewness is the standardized third cumulant. In fact, the classical tests of symmetry make use of the standardized third sample cumulant measure and a departure from the normal value of zero then indicates skewness. Intuitively, we think in a distribution as being skewed if it systematically deviates from a symmetrical form. Kurtosis is a measure of whether the data are peaked or flat relative to a normal distribution. That is, data sets with high kurtosis tend to have a distinct peak near the mean, decline rather rapidly, and have heavy tails. Data sets with low kurtosis tend to have a flat top near the mean rather than a sharp peak. Is given by the standardized fourth cumulant which equals zero for any normal distributions. Positive kurtosis indicates a "peaked" distribution and negative kurtosis indicates a "flat" distribution.



The following example shows histograms for 1000 random numbers generated from a Normal, a Exponential, a Gamma, and a Uniform distribution. The first histogram is a sample of a normal distribution, is a symmetric distribution with well behaved tails. This is indicated by kurtosis

and skewness near to zero. The second histogram is a sample of exponential distribution, is a skewed distribution with value equal to 1.991087 (near to the true value) and kurtosis equal to 5.706499 (also near to the true value). The third histogram is a sample from a gamma distribution, also a skewed distribution (2.13291) with considerable kurtosis (6.150792). The fourth histogram is a sample from a uniform distribution, a symmetric distribution with kurtosis (-1.23475) and skewness near to zero.

From Tables (4.1) to (4.3), and not too surprisingly, we find that the coverage probabilities of the confidence intervals are much below the nominal values for all fixed parameters, when the kurtosis and skewness of random distribution are significantly large (i.e. correspond to an exponential and gamma distribution), especially for parameter δ_2 with coverage probabilities ranging from approximately 50% to 54.6%. The value of the coverage probabilities in the upper one-side (PROB. UPP) increase with the values of kurtosis and skewness. The same conclusion can be deduced for the diagonal elements of the covariance matrix of the random effects. For the parameter σ, the standard error of residuals, it is observed that the coverage varies greatly across the values of kurtosis and skewness of the errors distribution, with poor coverage probabilities when heavily tailed error distributions are used. It is also interesting to note that (approximately) the same pattern is also seen for the linear mixed model case, see Sanchez and Ocaña (2000) and Sanchez and Ocaña (2001) for more details.

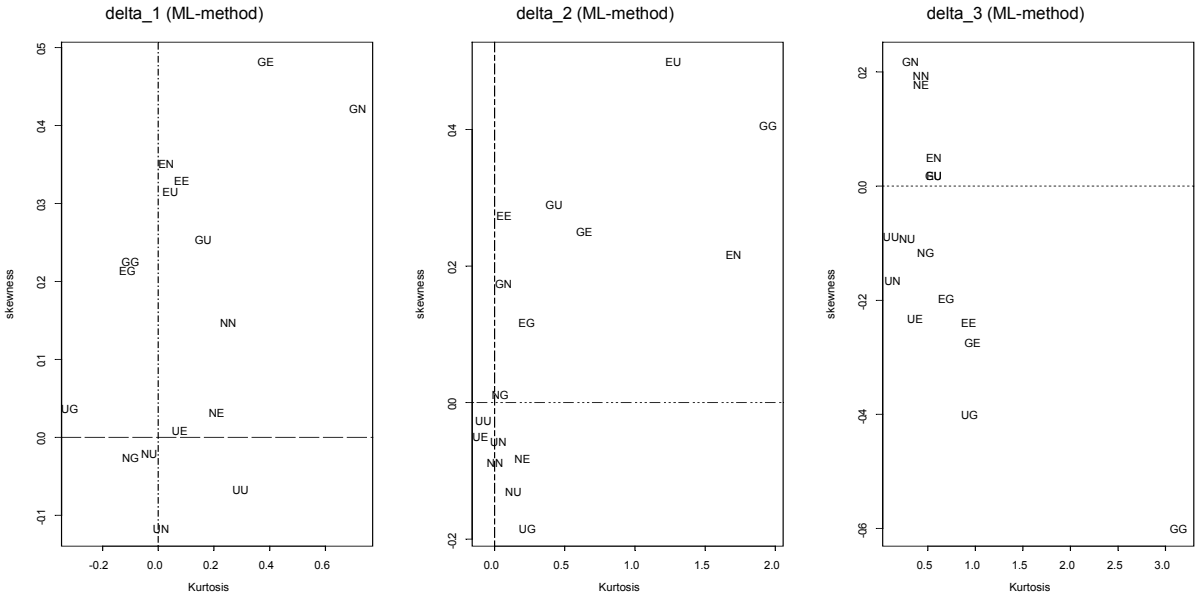


Figure 4.1 : A chart relating the distribution of fixed parameters to the values of kurtosis and skewness.

In the last paragraphs, we have discussed the effect on the coverage probabilities of departures from normality of the random effects and the residuals. These departures from normality have been quantified by means of the kurtosis and the skewness of the corresponding simulated distributions. In fact, the poor observed coverage probabilities are associated to an observable lack of normality of the parameter estimates which, in turn, may be quantified by means of their sample kurtosis and skewness.

For each fixed parameter, the sample coefficients of kurtosis, *Kurt*, and skewness, *Skw*, are defined as

$$Kurt = \frac{1}{M-1}\frac{\sum_i \left(\widehat{\delta}_i - \overline{\widehat{\delta}}\right)^4}{\widehat{\sigma}_{\widehat{\delta}}^4} - 3 \quad \text{and} \quad Skw = \frac{1}{M-1}\frac{\sum_i \left(\widehat{\delta}_i - \overline{\widehat{\delta}}\right)^3}{\widehat{\sigma}_{\widehat{\delta}}^3},$$

where $\widehat{\sigma}_{\widehat{\delta}}$ is the square root of $\sum_i \left(\widehat{\delta}_i - \overline{\widehat{\delta}}\right)^2 \Big/ (M-1)$ and M= 1000 is number of simulated parameter estimations. These simulation summaries were calculated and plotted in Figure 4.1, for each series of $M = 1000$ (simulated) estimates. As in the preceding discussions, NN stands for normal random effects and residuals, NU for normal random effects and uniform residuals,

and so on. Good agreement with normal asymptotic theory happens when the point (*Kurt*, *Skw*) is close to the origin. Departure of *Kurt* and *Skw* from the normal value of zero is an indication of non-normality in the distribution of the NLME estimates of the parameters. Contradicting the asymptotic theory, for both parameters, δ_1 and δ_2, the departure from normality of their estimations is significant, in the sense of kurtosis and skewness. These plots indicate that, for all fixed parameters and specially for δ_2, and for gamma and exponential random effects, the distribution of the estimates of the fixed effects is clearly not normal, in correspondence with the very poor coverages of the confidence intervals under these simulated distributions. In these cases, the upper one-side coverage probabilities increase together with the coefficients of kurtosis and skewness. This indicates a possible relationship between the asymmetry of the distributions of the random effects and the upper one-side coverage probabilities. On the other hand, the simulations with observed coverage probabilities in best agreement with the nominal ones, are associated to symmetric normal or uniform random effects. Otherwise, the coverage tends to be lower than the nominal level.

In conclusion, the simulations suggest that the non-linear mixed model estimators based on the LB-method do not perform adequately when the model assumptions are violated, especially under random effects distributions with large kurtosis and skewness. In these cases, the robust approaches discussed in the next chapters, may provide a good alternative.

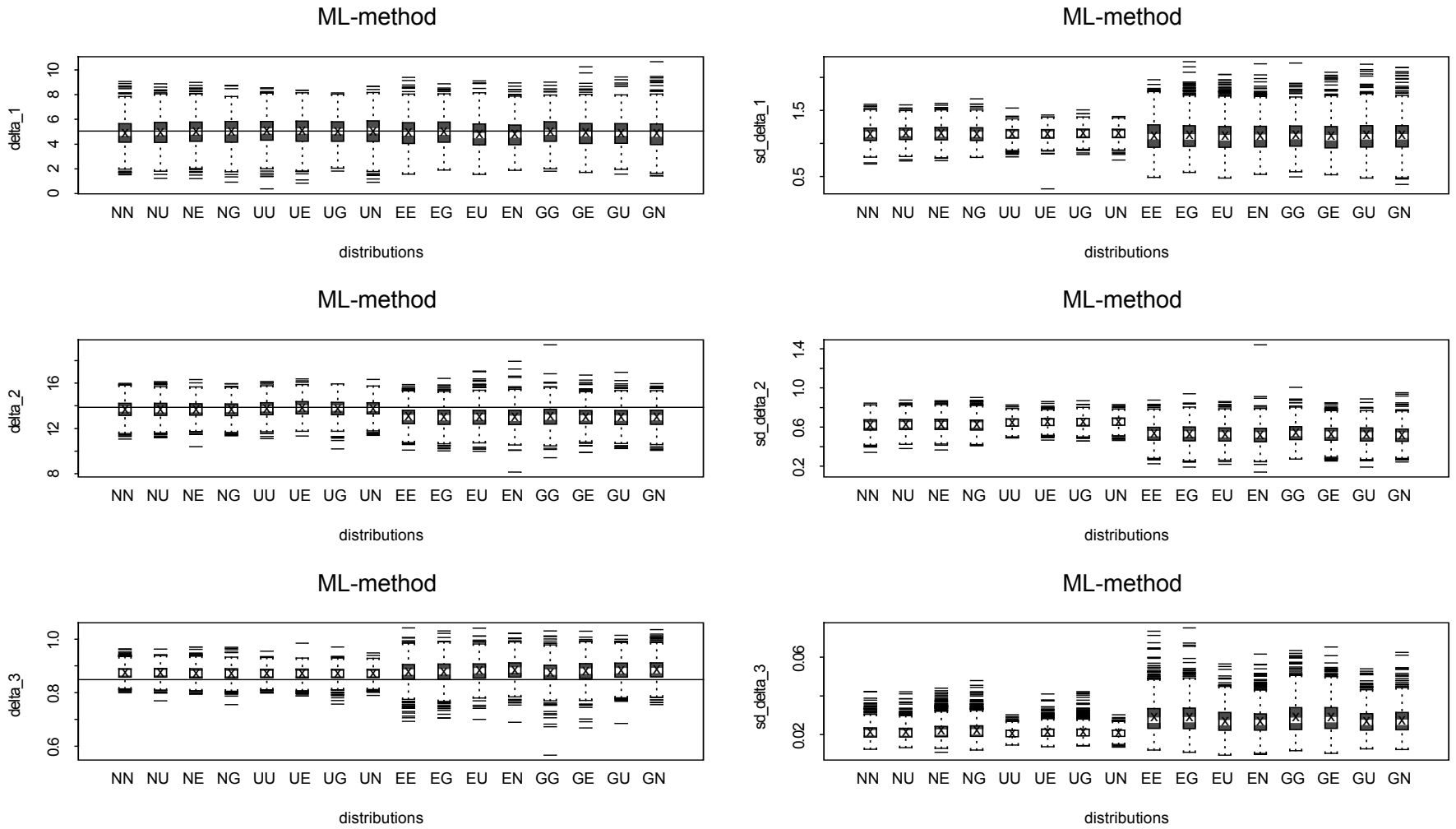


Figure 4.2: Box-plots for estimates the fixed effects in the first column in the figure and his corresponding standard deviation in the second column

Table (4.1): summary results for fixed effect parameter estimate δ_1 of 1000 simulations

Maximum likelihood (ML)									
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{Bias}$	MSE	COVERAGE- PROBABILITY (ASYMPTOTIC INTERVALS)			Average width
						PROB. LOW	PROB.	PROB. UPP	
	N	**N**	4.886202	-0.1650338±0.07285519	1.40892	2.1%	93.6%	4.3%	4.47681
	N	**U**	4.944577	-0.1066586±0.07229706	1.370611	1.8%	94.2%	4%	4.509656
	N	**E**	4.989871	-0.06136455±0.07269087	1.377849	2.5%	94.4%	3.1%	4.493222
	N	**G**	5.006408	-0.04482788±0.07484748	1.458836	3.2%	92.9%	3.9%	4.474741
	U	**U**	5.0624	0.01116402±0.07638769	1.355001	3.5%	92.9%	3.6%	4.481386
	U	**E**	5.077319	0.02608274±0.07227156	1.358957	3.3%	94.2%	2.5%	4.474291
	U	**G**	5.025815	-0.02542078±0.0684192	1.217979	2%	95.6%	2.4%	4.506321
	U	**N**	5.040766	-0.01047035±0.0743065	1.435953	2.3%	93.7%	4%	4.502533
	E	**E**	4.945365	-0.1058706±0.07475541	1.464453	1.4%	89.2%	9.4%	4.384257
	E	**G**	4.995135	-0.0561009±0.07417945	1.434085	1%	91.9%	7.1%	4.432266
	E	**U**	4.787922	-0.2633136±0.07377517	1.484717	0.6%	89%	10.4%	4.365749
	E	**N**	4.787817	-0.2634191±0.07195203	1.415682	0.8%	88.8%	10.4%	4.383538
	G	**G**	5.035892	-0.0153443±0.07393566	1.360595	0.8%	92.6%	6.6%	4.423033
	G	**E**	4.932696	-0.1185403±0.07446576	1.456056	1%	91.1%	7.9%	4.368033
	G	**U**	4.882763	-0.1684727±0.07320609	1.422014	0.4%	90.1%	9.5%	4.420608
	G	**N**	4.849067	-0.2021688±0.07697968	1.581883	1%	88.9%	10.1%	4.398509
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.2): summary results for fixed effect parameter estimate δ_2 of 1000 simulations

Maximum likelihood (ML)									
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE- PROBABILITY (ASYMPTOTIC INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	13.68003	-0.1869998±0.04997716	0.6844947	4%	86%	10%	2.434305
	N	**U**	13.66495	-0.2020765±0.0486743	0.6569372	3.2%	87.7%	9.1%	2.454789
	N	**E**	13.65974	-0.2072862±0.04740306	0.6273083	2.9%	87.9%	9.2%	2.455891
	N	**G**	13.64766	-0.2193667±0.04754954	0.6360792	3.6%	87%	9.4%	2.446186
	U	**U**	13.76617	-0.1008636±0.05228649	0.6449659	4.9%	88.1%	6.9%	2.533951
	U	**E**	13.8117	-0.05532674±0.0481583	0.6061715	4%	89%	7%	2.547763
	U	**G**	13.76261	-0.1044175±0.04970701	0.6534258	4.2%	88.9%	6.9%	2.546479
	U	**N**	13.76965	-0.09738244±0.0474913	0.5960001	4%	88.8%	7.2%	2.556021
	E	**E**	13.04766	-0.8193652±0.0569269	1.514089	2%	53.5%	44.5%	2.081576
	E	**G**	12.98718	-0.8798521±0.05773234	1.640885	1.4%	53%	45.6%	2.077891
	E	**U**	13.02594	-0.8410943±0.05879794	1.606477	2.5%	52.3%	45.2%	2.052472
	E	**N**	12.9656	-0.9014302±0.05855666	1.70425	1.5%	50%	48.5%	2.031885
	G	**G**	13.04935	-0.8176776±0.06117519	1.599911	1.9%	54.6%	43.5%	2.113754
	G	**E**	12.99329	-0.8737422±0.05505199	1.551558	1.6%	52.1%	46.3%	2.062114
	G	**U**	12.97432	-0.8927112±0.0573862	1.653317	1.7%	50.2%	48.1%	2.050767
	G	**N**	13.01355	-0.8534754±0.05730351	1.582337	1.5%	51.5%	47%	2.04065
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.3): summary results for fixed effect parameter estimate δ_3 of 1000 simulations

Maximum likelihood (ML)									
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	0.8737723	0.0251168±0.001494758	0.001211879	24.6%	75.2%	0.2%	0.08369995
	N	**U**	0.8741647	0.02550918±0.001493094	0.001230451	25.9%	73.6%	0.5%	0.08322564
	N	**E**	0.8723808	0.02372535±0.001568866	0.001202958	23.1%	76.7%	0.2%	0.08647451
	N	**G**	0.8712071	0.02255159±0.001660905	0.001225943	23.2%	76.5%	0.3%	0.08721456
	U	**U**	0.8713741	0.02271856±0.001513452	0.001047984	24.1%	75.6%	0.3%	0.0808331
	U	**E**	0.8707938	0.02213835±0.001528409	0.001097587	24.6%	75%	0.4%	0.08342959
	U	**G**	0.8701139	0.02145838±0.00152868	0.001068158	21.9%	77.4%	0.7%	0.08386012
	U	**N**	0.870907	0.02225153±0.00141523	0.001015975	22.1%	77.5%	0.4%	0.08113575
	E	**E**	0.8771704	0.02851486±0.002790856	0.002838578	25.8%	71.6%	2.6%	0.1143967
	E	**G**	0.8780229	0.02936741±0.002740271	0.002815166	27.7%	69.9%	2.4%	0.1141702
	E	**U**	0.8817126	0.03305709±0.002535862	0.002765034	29.1%	68.7%	2.2%	0.1056396
	E	**N**	0.8849053	0.03624979±0.002503449	0.002943834	32.5%	65.3%	2.2%	0.1058578
	G	**G**	0.8749357	0.02628024±0.002843073	0.002702158	24.6%	72.6%	2.8%	0.1141582
	G	**E**	0.8793333	0.03067775±0.00269119	0.002824523	28.6%	68.9%	2.5%	0.1143106
	G	**U**	0.8841023	0.03544684±0.002408504	0.002764988	31.6%	66.9%	1.5%	0.1054696
	G	**N**	0.8856528	0.03699725±0.002502829	0.002997777	31.7%	66.4%	1.9%	0.1074295
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

4.2.1.2. Results for the estimations of the variance parameters: ML- method

A summary of statistical measures of the variance components D and σ are presented in Tables (4.4) to (4.10), with corresponding box-plots presented in Figures 4.3 to 4.4. Theoretically, the use of the ML estimation method will tend to yield biased estimates of D. This is confirmed by the mean observed biases, that were significantly different from 0. In terms of mean square error, the MSEs generally have the largest magnitude for diagonal elements, D_{11} and D_{22}, and become larger when the law of the random effects is not normal or uniform. In these cases, the mean values become slightly less precise in the sense of large intervals for C.I_{Bias}.

In the *nlme* library there are methods to compute confidence intervals for the standard deviations of the random effects (that is, for the square root of the elements in the diagonal of D, $\sqrt{D_{ii}}$) and for the correlation terms, $D_{ij}/(\sqrt{D_{ii}}\sqrt{D_{jj}})$, and not directly for the elements of D. According to this fact, we also discuss the simulation results on the true coverage of the confidence intervals for the standard deviations and correlations. In general, the coverage probabilities are lower than the nominal values, except under normal or uniform distribution of the random effects, specially for $\sqrt{D_{11}}$ and to a lesser extent for $\sqrt{D_{22}}$. In all cases, the poor coverages correspond to inflated values of PROB.UPP, that is, there is a marked tendency to the confidence intervals being lower than the true parameter value. For $\sqrt{D_{33}}$, the coverage probabilities are even more perturbed and, in contrast, the low errors are much more frequent than the UPPER errors, specially when the random effects are exponential or gamma. For the correlation parameter $D_{12}/(\sqrt{D_{11}}\sqrt{D_{22}})$, the best coverage occurs under normality conditions (i.e. the NN combination). In general terms, the inference on the variance parameters is even less robust than the inference on the fixed effects, but adequate for the normal case. For all standard deviations and correlations, the coverage in the NN case always lies near the nominal 95% level, while the coverages under departures of the normality assumption are very erratic and affected by both, the distribution of the random effects and the residuals. Again, the worst coverage results and the highest asymmetry in the no coverage cases correspond to the asymmetric E and G distributions for the random effects. Otherwise the coverage is, always lower than its nominal

value. For the residuals parameter σ it appears that, as expected, the coverage probabilities depend on the error distribution more than that of the random effects. The results for the bias and MSE show that it is robust. On the other hand, the coverage of the asymptotic confidence intervals for this parameter ranges from 99.2% in the EU case to a 65.7% in the UG case, with a 93.3% of coverage when both, the simulated random effects and residuals, are normal. The worst coverage values correspond to the case of the E and G distributions for the residuals. Independently of the true coverage, in all cases the confidence intervals are quite equitailed.

The sampling distributions of the estimators are skewed in many cases except when the errors are normal.

A general conclusion of this part of the simulation study is that the LB-method performs very poorly under departures from normality in the sense of kurtosis and skewness of random effects distribution in all parameters. Under these conditions the coverage probabilities become worst and very low compared with the nominal ones.

In order to confirm the results on fixed effects previously obtained, based on ML-method, and with the purpose to investigate the inference statistics for covariance components when using LB-method based on an alternative approach, additional simulation series were carried out, and presented in the next section, for the restricted maximum likelihood estimates (REML).

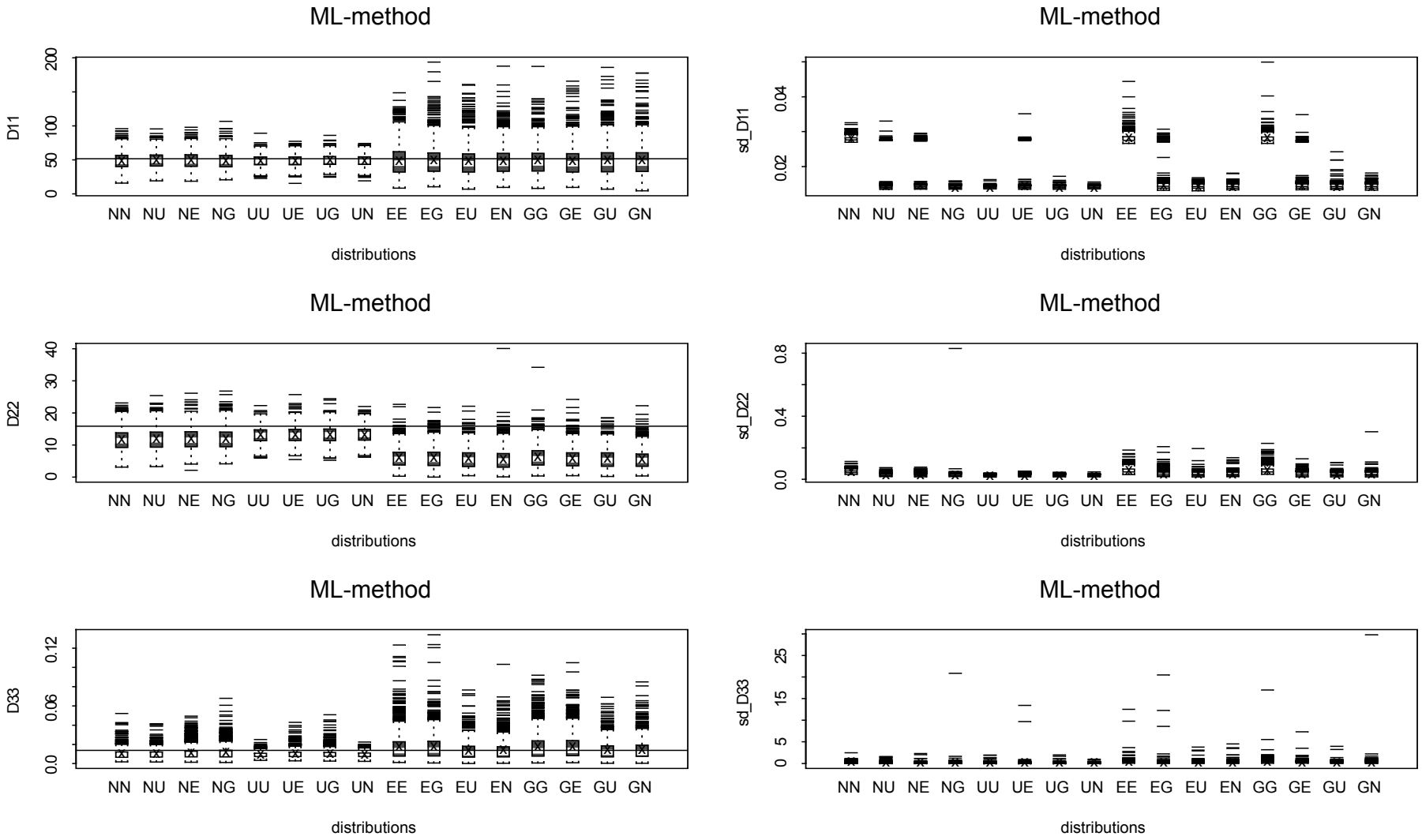


Figure 4.3: Box-plots of covariance parameter estimates from simulation settings of the breast cancer model using ML-method.

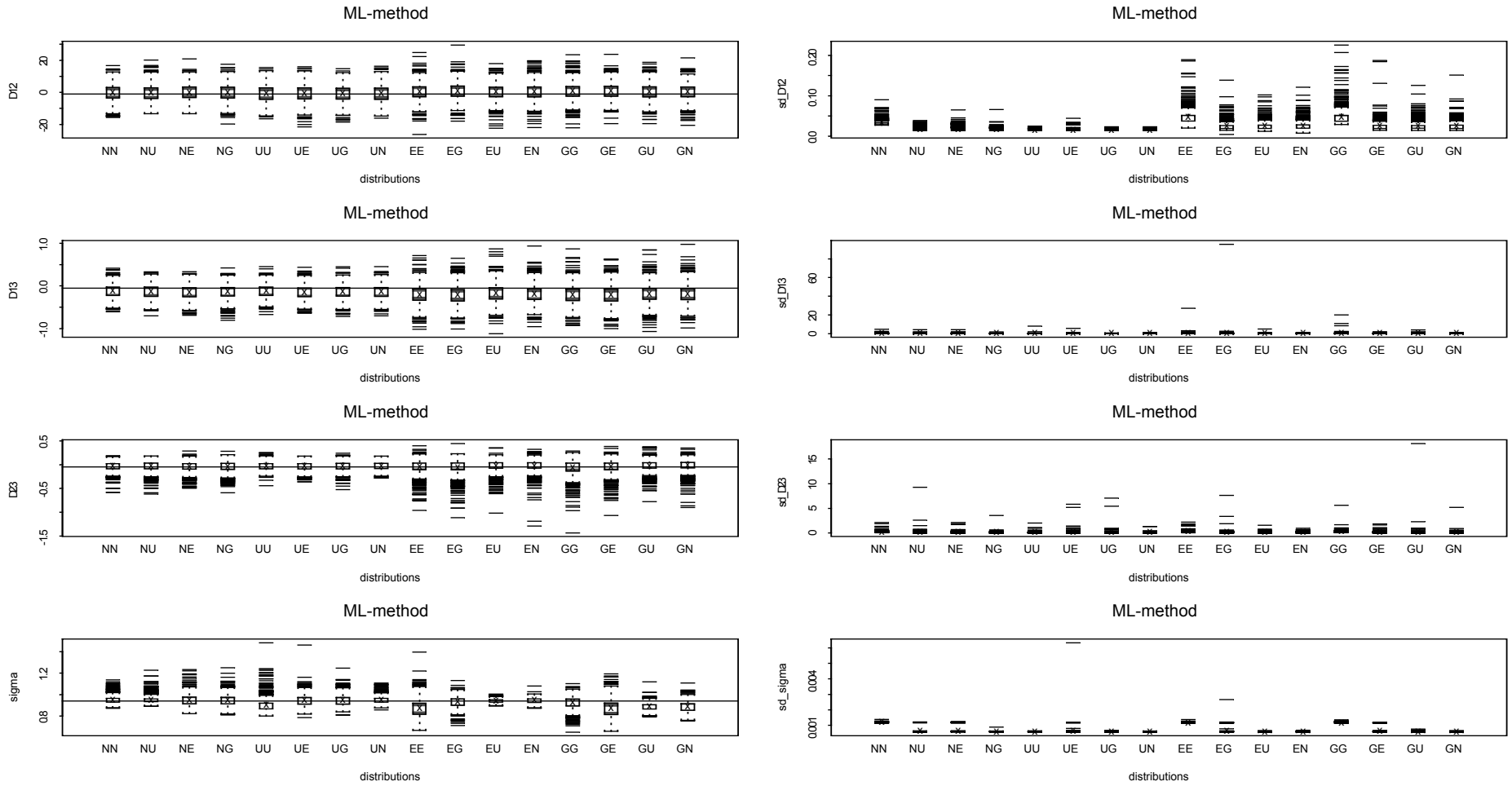


Figure 4.4: Box-plots of deviation standard estimates of covariance components D and σ from simulation settings of the breast cancer model using ML-method

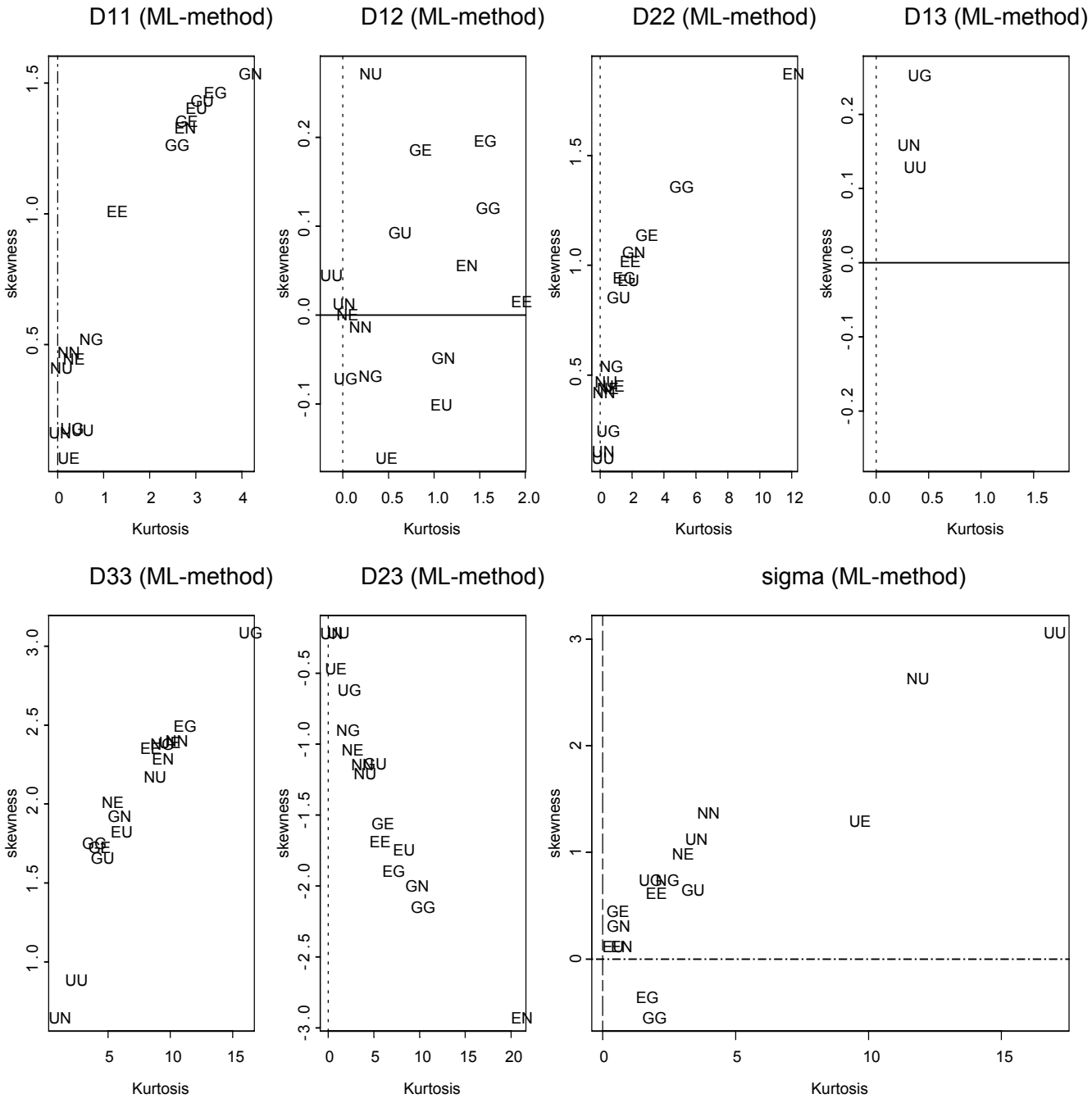


Figure 4.5: A chart relating the distribution of components to the values of kurtosis and skewness.

Table (4.4): summary results for random effect parameter estimates D_{11} of 1000 simulations

D_{11}						$\sqrt{D_{11}}$			
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{Bias}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	49.12197	-2.764363±0.7636382	159.2869	0%	98.8%	1.2%	4.633831
	N	**U**	49.82366	-2.06267±0.7476408	149.6127	0.7%	93.1%	6.2%	3.295812
	N	**E**	49.44254	-2.443788±0.7526413	153.2812	0.8%	92.8%	6.4%	3.320996
	N	**G**	49.02589	-2.860445±0.7570901	157.2378	1%	91.8%	7.2%	3.243518
	U	**U**	48.79587	-3.090455±0.5379606	76.74848	0.1%	98%	1.9%	3.248054
	U	**E**	48.66453	-3.221798±0.5200195	80.7022	0%	97.7%	2.3%	3.268257
	U	**G**	49.35561	-2.53072±0.526287	78.43208	0.1%	97.8%	2.1%	3.266399
	U	**N**	49.27168	-2.61465±0.5150737	75.82735	0%	98.4%	1.6%	3.262939
	E	**E**	49.00393	-2.882396±1.392285	512.3999	3.3%	83.7%	13%	4.545643
	E	**G**	50.22051	-1.66582±1.503549	590.6544	9.7%	68%	22.3%	3.259239
	E	**U**	48.75185	-3.134485±1.447256	554.5084	7.7%	67.1%	25.2%	3.172032
	E	**N**	49.10945	-2.776883±1.432559	541.3883	9.1%	67%	23.9%	3.186562
	G	**G**	49.72573	-2.160596±1.425858	510.6072	3.2%	85.7%	11.1%	4.58357
	G	**E**	48.71184	-3.174492±1.439096	548.6362	8.2%	65.9%	25.9%	3.201626
	G	**U**	50.15517	-1.731159±1.533679	614.6738	9.4%	67.1%	23.5%	3.214977
	G	**N**	49.52359	-2.362743±1.475912	572.049	8.5%	67.4%	24.1%	3.195702
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.5): summary results for random effect parameter estimate D_{12} of 1000 simulations

			D_{12}			$D_{12}/\sqrt{D_{11}}\sqrt{D_{22}}$			
			MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE-PROBABILITY (ASYMP. INTERVALS)			Average width
	RANDOM EFFECTS	ERRORS				PROB LOW.	PROB.	PROB. UPP	
DISTRIBUTIONS	**N**	**N**	-0.2518046	0.7980574±0.3126105	26.05016	4.9%	95%	0.1%	0.7923766
	N	**U**	-0.3978791	0.6519829±0.3169908	26.55552	16.4%	83.1%%	0.5%	0.5784795
	N	**E**	-0.0798979	0.9699641±0.3030612	24.82522	17.2%	82.3%	0.5%	0.5900325
	N	**G**	-0.2234888	0.8263732±0.3082055	25.385	18.4%	80.8%	0.8%	0.5788606
	U	**U**	-0.8205849	0.2292771±0.3494239	28.40288	14.6%	84%	1.5%	0.576965
	U	**E**	-0.7195836	0.3302784±0.3279258	28.07342	14.2%	84.4%	1.4%	0.5833068
	U	**G**	-0.8362431	0.2136189±0.3200907	26.68963	15.2%	83.5%	1.3%	0.5746026
	U	**N**	-0.7365703	0.3132917±0.3186744	26.50689	14.4%	84.5%	1.1%	0.5657772
	E	**E**	0.5003593	1.550221±0.3185833	28.79683	10.7%	89.2%	0.1%	0.8551481
	E	**G**	0.809861	1.859723±0.3145599	29.18978	26%	72.9%	1.1%	0.6263893
	E	**U**	0.1560926	1.205955±0.3147653	27.21914	22.7%	75.2%	2.1%	0.618492
	E	**N**	0.3281777	1.37804±0.3202606	28.57129	23.2%	75.7%	1.1%	0.6264209
	G	**G**	0.7429011	1.792763±0.3289412	30.14064	11.2%	88.7%	0.1%	0.8524822
	G	**E**	0.7346797	1.784542±0.2999999	26.58888	24.7%	74.8%	0.5%	0.6322905
	G	**U**	0.3349302	1.384792±0.3163595	27.94411	25.3%	72.9%	1.8%	0.6260145
	G	**N**	0.1792292	1.229091±0.3016069	25.16637	22.9%	75.5%	1.6%	0.6411854
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.6): summary results for fixed effects parameter estimate D_{22} of 1000 simulations

			D_{22}			$\sqrt{D_{22}}$			
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	11.67781	-4.168688±0.2102284	28.87103	0%	95.3%	4.7%	2.579742
	N	**U**	11.90277	-3.94373±0.2034323	26.31501	0.1%	87.1%	12.8%	1.827461
	N	**E**	11.9102	-3.936304±0.2118475	27.16528	0.1%	86.6%	13.3%	1.852461
	N	**G**	11.89818	-3.948319±0.218857	28.0451	0.2%	84.4%	15.4%	1.804916
	U	**U**	13.10089	-2.745614±0.1765797	14.77832	0%	96.4%	3.6%	1.849672
	U	**E**	13.26961	-2.576892±0.1616895	13.43893	0%	97.9%	2.1%	1.872346
	U	**G**	13.22271	-2.623793±0.1695726	14.36193	0%	96.5%	3.5%	1.857041
	U	**N**	13.30909	-2.537409±0.158347	12.95882	0%	97.4%	2.6%	1.861902
	E	**E**	6.010085	-9.836415±0.1958436	106.7291	0%	57.2%	42.8%	2.354282
	E	**G**	6.052599	-9.793901±0.2007665	106.4023	0%	35.4%	64.6%	1.655071
	E	**U**	5.730879	-10.11562±0.1927222	111.9844	0%	32.4%	67.6%	1.638215
	E	**N**	5.5799	-10.2666±0.2061625	116.4559	0.1%	30%	69.9%	1.623207
	G	**G**	6.330407	-9.516093±0.2190076	102.4922	0.1%	60.3%	39.6%	2.402462
	G	**E**	5.82745	-10.01905±0.1905557	109.8241	0%	31.2%	68.8%	1.638975
	G	**U**	5.741611	-10.10489±0.1886042	111.3591	0%	32.4%	67.6%	1.635632
	G	**N**	5.612023	-10.23448±0.1874726	113.8842	0%	29.6%	70.4%	1.62804
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.7): summary results for fixed effect parameter estimate D_{13} of 1000 simulations

			D_{13}			$D_{13}/\sqrt{D_{11}}\sqrt{D_{33}}$			
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{Bias}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	-0.1283528	-0.0737467±0.00978824	0.03035367	0%	95.8%	4.2%	0.8524842
	N	**U**	-0.1347431	-0.08013693±0.009833985	0.03157045	0.3%	79%	20.7%	0.6289445
	N	**E**	-0.1513652	-0.09675906±0.01030077	0.03695494	0.7%	80.4%	18.9%	0.6323777
	N	**G**	-0.143367	-0.0887609±0.01005309	0.03416015	0.7%	82%	17.3%	0.6200693
	U	**U**	-0.1243997	-0.06979353±0.00986075	0.02744847	0.4%	83.7%	15.9%	0.6357837
	U	**E**	-0.148455	-0.09384881±0.009734705	0.03345091	0.6%	79.5%	19.9%	0.6402495
	U	**G**	-0.1391199	-0.08451377±0.01013298	0.03384361	0.4%	80.9%	18.7%	0.6327893
	U	**N**	-0.1396192	-0.08501306±0.009900535	0.03271727	0.4%	80.4%	19.2%	0.6306309
	E	**E**	-0.2073742	-0.1527681±0.01351359	0.07082725	0.1%	92.6%	7.3%	0.867063
	E	**G**	-0.2151377	-0.1605315±0.01343951	0.07274036	0.4%	72.8%	26.8%	0.6420624
	E	**U**	-0.1733657	-0.1187596±0.01329306	0.06005569	0.9%	74.3%	24.8%	0.6489391
	E	**N**	-0.1959947	-0.1413886±0.01275682	0.06230997	23.2%	75.7%	1.1%	0.6264209
	G	**G**	-0.2120879	-0.1574817±0.01421434	0.07508092	0.2%	91.4%	8.4%	0.8699468
	G	**E**	-0.2146674	-0.1600613±0.01391164	0.07594765	0.6%	72.2%	27.2%	0.6419619
	G	**U**	-0.1323687	0.01336876±0.06399821	0.06399821	0.9%	73%	26.1%	0.6434339
	G	**N**	-0.1932153	-0.1386092±0.01339042	0.06583997	1%	72.5%	26.5%	0.6411854
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.8) summary results for fixed effect parameter estimate D_{23} of 1000 simulation

			D_{23}			$D_{23}/\sqrt{D_{22}}\sqrt{D_{33}}$			
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE-PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	-0.03998685	0.005885753±0.006026178	0.0094782	0.1%	96.9%	3%	0.8994339
	N	**U**	-0.03595269	0.009919913±0.006090153	0.00974357	1.1%	86.4%	12.5%	0.6794363
	N	**E**	-0.04513761	0.0007349925±0.006431441	0.01075702	1%	85.4%	13.6%	0.6748734
	N	**G**	-0.04317801	0.002694587±0.006516454	0.01104998	1.7%	83.9%	14.4%	0.6634598
	U	**U**	-0.03145884	0.01441376±0.005580134	0.00743782	1.5%	87%	11.5%	0.6485216
	U	**E**	-0.03867888	0.007193715±0.005465864	0.00719371	0.7%	87.7%	11.6%	0.6545583
	U	**G**	-0.03600132	0.009871281±0.0054846	0.00791990	1.5%	88.7%	9.8%	0.6504467
	U	**N**	-0.03082106	0.01505154±0.00507595	0.00692675	0.8%	89%	10.2%	0.6448514
	E	**E**	-0.05308781	0.008975154±0.0209998	0.0209998	1%	91.4%	7.6%	0.9918798
	E	**G**	-0.05705459	-0.01118199±0.009428597	0.02324289	3.7%	76%	20.3%	0.7417546
	E	**U**	-0.0261368	0.0197358±0.007271963	0.01414121	3.9%	82.3%	13.8%	0.7644016
	E	**N**	-0.02837074	0.01750186±0.007708895	0.0157602	3.1%	83.9%	13%	0.7784046
	G	**G**	-0.06547542	-0.01960282±0.009798675	0.02427781	0.7%	90.7%	8.6%	0.9985969
	G	**E**	-0.00904083	-0.009040833±0.008619223	0.01940095	3%	77.5%	19.5%	0.2210729
	G	**U**	-0.02347535	0.02239725±0.006906841	0.01290708	4%	85.8%	10.2%	0.7739817
	G	**N**	-0.02043374	0.02543886±0.007406503	0.0149124	3.9%	86.5%	9.6%	0.7658838
	Breast cancer model: $y_{ij} = \dfrac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.9): summary results for fixed effect parameter estimate D_{33} of 1000 simulations

			D_{33}			$\sqrt{D_{33}}$			
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB. LOW	PROB.	PROB. UPP	
	N	**N**	0.01000946	-0.003618449±0.0003260	0.000040743	1.6%	97.1%	1.3%	0.0992976
	N	**U**	0.00992189	-0.003706±0.0003130365	0.000039217	3.6%	87.7%	8.7%	0.06877083
	N	**E**	0.01127979	-0.002348124±0.0004303	0.000053644	9%	83.3%	7.7%	0.07412037
	N	**G**	0.01160907	-0.002018844±0.0004564	0.000058243	9.9%	82.5%	7.6%	0.07391395
	U	**U**	0.00911190	-0.004516009±0.0055801	0.000027645	0.3%	96.9%	2.8%	0.06496839
	U	**E**	0.0101881	-0.003439812±0.0002806	0.000032318	3.4%	92%	4.6%	0.06860242
	U	**G**	0.0102506	-0.003377314±0.0003005	0.00003489	3.6%	92.2%	4.2%	0.06836718
	U	**N**	0.00925345	-0.004374456±0.0001774	0.000027322	0.3%	95.6%	4.1%	0.06505494
	E	**E**	0.01797212	0.004344213±0.00096711	0.000262094	22.2%	76.8%	1%	0.1426767
	E	**G**	0.01797659	0.004348677±0.00091490	0.000236584	34.5%	61.3%	4.2%	0.100379
	E	**U**	0.01363981	0.00001189559±0.007272	0.000092106	24.9%	68.7%	6.4%	0.0878535
	E	**N**	0.01402233	0.0003944152±0.0006539	0.000111355	23.2%	69.8%	7%	0.08907202
	G	**G**	0.01801382	0.004385905±0.00094184	0.000239985	24.1%	74%	1.9%	0.1427349
	G	**E**	0.01801355	0.004385642±0.00087094	0.000216490	35.9%	59.8%	4.3%	0.09946679
	G	**U**	0.01384016	0.0002122487±0.0005889	0.000090239	25.4%	68.2%	6.4%	0.08757906
	G	**N**	0.01473304	0.001105134±0.00068228	0.000122275	27.3%	66%	6.7%	0.08991439
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.10): summary results for fixed effect parameter estimate σ of 1000 simulations

Maximum likelihood (ML)									
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{Bias}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	0.9513554	0.01143439±0.002085509	0.001261784	6.5%	93.3%	0.2%	0.1286382
	N	**U**	0.9523404	0.01241937±0.002073453	0.00127224	9.7%	90.1%	0.2%	0.09162094
	N	**E**	0.9488324	0.008911373±0.00324329	0.002814846	19.7%	66.9%	13.4%	0.09257411
	N	**G**	0.9455525	0.005631508±0.00310463	0.002538247	18.6%	67.3%	14.1%	0.0904925
	U	**U**	0.9029918	-0.03692915±0.00388731	0.004872505	7.4%	92.5%	0.1%	0.09075843
	U	**E**	0.9451742	0.005253153±0.0032294	0.002739641	18.6%	67%	14.4%	0.09140316
	U	**G**	0.9457767	0.005855668±0.00312599	0.002575436	19.4%	65.7%	14.9%	0.09046874
	U	**N**	0.9487785	0.008857527±0.00194164	0.001058825	9.3%	88.3%	2.4%	0.090652
	E	**E**	0.8714248	-0.06849617±0.00503193	0.01127623	5.2%	84%	10.8%	0.1255392
	E	**G**	0.9308426	-0.00907841±0.00287986	0.002239149	10.7%	68.9%	20.4%	0.08993602
	E	**U**	0.9410298	0.001108841±0.00103892	0.000281909	0.7%	99.2%	0.1%	0.08940592
	E	**N**	0.9432262	0.003305231±0.00152507	0.000615759	3.8%	93.7%	2.5%	0.0896321
	G	**G**	0.9252021	-0.01471885±0.00332354	0.002965468	5.5%	82.8%	11.7%	0.1254316
	G	**E**	0.8708419	-0.06907907±0.00509376	0.01151923	11%	68.4%	20.6%	0.08948928
	G	**U**	0.8855721	-0.0543489±0.001961499	0.003954332	0.8%	98.8%	0.4%	0.08943087
	G	**N**	0.8852012	-0.05471978±0.00286439	0.005127885	3.8%	93.2%	3%	0.08939784
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

4.2.1.3. Results for the estimations fixed effects: REML-Method

In this phase of the simulation study, we generated series of 1000 data according to the same non-linear mixed model (4.1)-(4.2) and using the same parameters as before, and the same starting values for the estimations than in the previous section.

The results are summarized in Figures 4.6 to 4.10 and Tables (4.11) to (4.20). As expected, this approximation reduces bias in estimating the variance components, confirming the theoretical results. Also gives larger values for the estimates of variance components, especially for D_{11} and D_{22} than the ML procedure. The higher mean square error of D_{11}, D_{12} and D_{22} is appreciable in the plots, especially when the law of the random effects is exponential or gamma. Despite of these facts, the conclusions of the simulation study are the same than for the ML method.

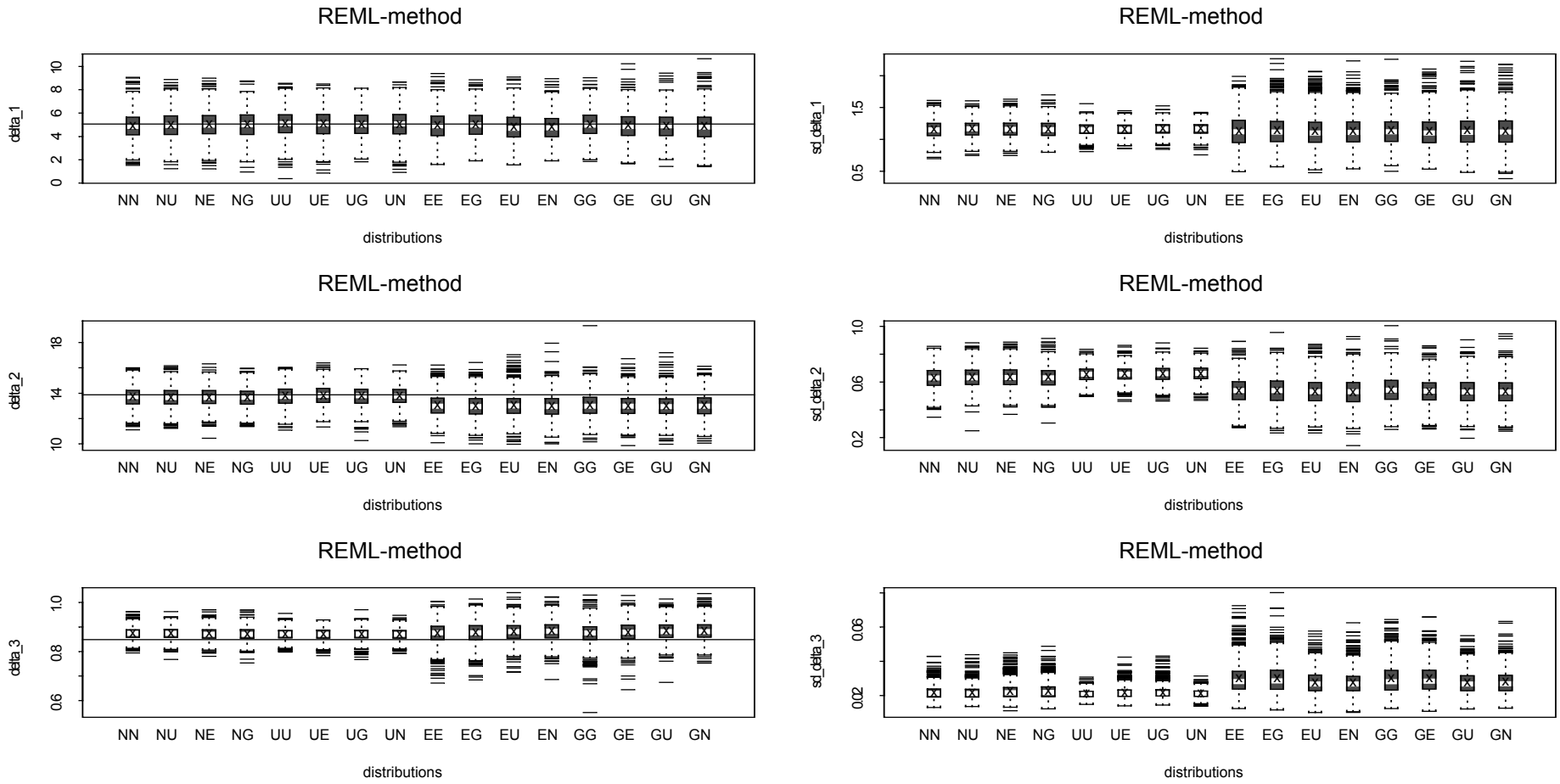


Figure 4.6: Box-plots of covariance parameter estimates from simulation settings of the breast cancer model using REML-method.

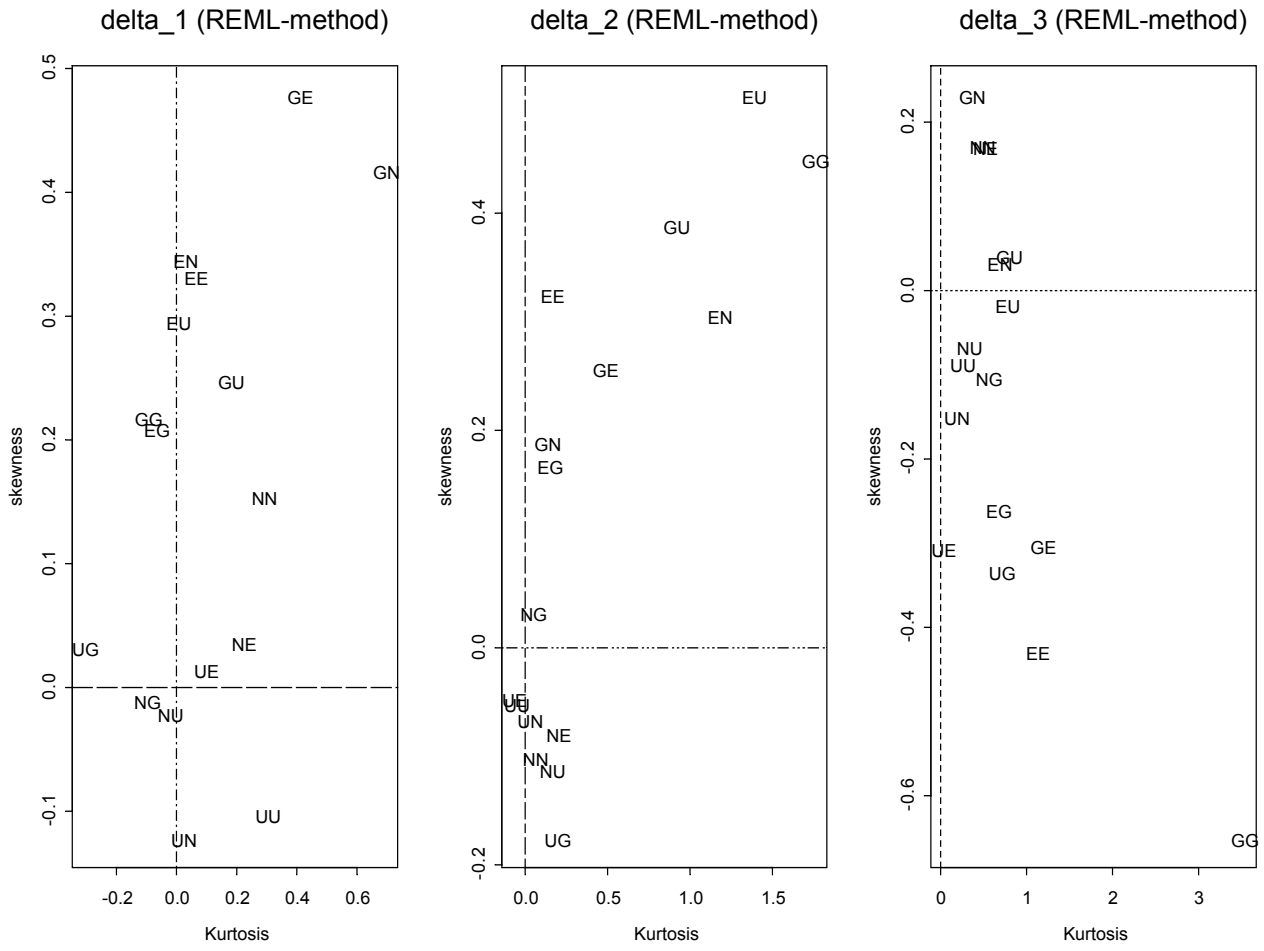


Figure 4.7: A char relating the distribution of fixed parameters to the values of kurtosis and skewness.

Table (4.11): summary results for fixed effect parameter estimate δ_1 of 1000 simulations

Restricted maximum likelihood (REML)									
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE- PROBABILITY (ASYMPTOTIC INTERVALS)			Average width
						PROB. LOW	PROB.	PROB. UPP	
	N	**N**	4.89163	-0.1596057±0.07277707	1.402818	2.1%	93.7%	4.2%	4.540506
	N	**U**	4.948345	-0.1028908±0.07245559	1.375789	1.6%	94.6%	3.8%	4.573463
	N	**E**	4.994979	-0.0562569±0.07305699	1.391125	2.4%	94.7%	2.9%	3.412721
	N	**G**	5.011928	-0.03930785±0.07481481	1.4571	3%	93.5%	3.5%	4.538401
	U	**U**	5.077169	0.02593315±0.08032211	1.38619	3.5%	93.2%	3.3%	4.541769
	U	**E**	5.080581	0.02934486±0.07210951	1.353053	3%	94.7%	2.3%	4.541181
	U	**G**	5.030605	-0.02063057±0.06847171	1.219628	1.9%	95.9%	2.2%	4.569877
	U	**N**	5.04441	-0.006825608±0.07440853	1.439835	2.3%	93.8%	3.9%	4.56763
	E	**E**	4.959712	-0.09152433±0.07497205	1.470056	1.4%	89.9%	8.7%	4.449963
	E	**G**	5.006709	-0.0445267±0.07424096	1.435294	1%	92.1%	6.9%	4.49404
	E	**U**	4.793905	-0.2573312±0.07356791	1.47366	0.4%	89.8%	9.8%	4.427747
	E	**N**	4.796197	-0.2550391±0.07188964	1.409004	0.8%	89.3%	9.9%	4.447861
	G	**G**	5.044785	-0.006450956±0.0737836	1.354811	0.8%	92.7%	6.5%	4.485408
	G	**E**	4.942322	-0.108914±0.07462648	1.460098	0.9%	91.4%	7.7%	4.429204
	G	**U**	4.886934	-0.1643016±0.07306906	1.415413	0.4%	90.9%	8.7%	4.482713
	G	**N**	4.858402	-0.1928345±0.07705762	1.581318	1.1%	89.3%	9.6%	4.461565
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.12): summary results for fixed effect parameter estimate δ_2 of 1000 simulations

Restricted maximum likelihood (REML)									
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE- PROBABILITY (ASYMPTOTIC INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	13.6896	-0.1774337±0.04990324	0.6790888	4.1%	86.3%	9.6%	2.467604
	N	**U**	13.66931	-0.1977175±0.04837813	0.6477196	3.1%	88.2%	8.7%	2.489247
	N	**E**	13.66968	-0.197353±0.04747166	0.6249813	3%	88.5%	8.5%	2.488511
	N	**G**	13.65235	-0.2146848±0.04716476	0.62457	3.7%	87.5%	8.8%	2.482292
	U	**U**	13.76485	-0.1021756±0.05493028	0.6584256	4.7%	87.8%	7.5%	2.569341
	U	**E**	13.81685	-0.05018497±0.0481563	0.6055771	3.9%	89.6%	6.5%	2.581954
	U	**G**	13.7648	-0.1022274±0.04975353	0.6541766	4.3%	89.2%	6.5%	2.582267
	U	**N**	13.77355	-0.09348465±0.0476959	0.6003228	4%	89.6%	6.4%	2.59347
	E	**E**	13.05767	-0.8093638±0.05664472	1.489466	2%	54.7%	43.3%	2.115946
	E	**G**	12.99099	-0.8760359±0.05724812	1.619706	1.3%	53.3%	45.4%	2.109074
	E	**U**	13.03448	-0.832546±0.05812583	1.571734	2.3%	54.2%	43.5%	2.083141
	E	**N**	12.98201	-0.8850216±0.05809592	1.66096	1.4%	51.5%	47.1%	2.071086
	G	**G**	13.0492	-0.8178266±0.06031339	1.5741	1.8%	55.1%	43.2%	2.143113
	G	**E**	13.00693	-0.8600956±0.05491506	1.523982	1.4%	53.8%	44.8%	2.092467
	G	**U**	12.99137	-0.8756637±0.05783555	1.636634	1.8%	51%	47.2%	2.087181
	G	**N**	13.03293	-0.8341005±0.05781278	1.564886	1.7%	52.8%	45.5%	2.080956
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.13): summary results for fixed effect parameter estimate δ_3 of 1000 simulations

Restricted maximum likelihood (REML)									
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	0.8731036	0.0244481±0.00149189	0.001176507	22.2%	77.7%	0.1%	0.08522075
	N	**U**	0.8735583	0.02490278±0.001496292	0.001202367	23.5%	76%	0.5%	0.08483158
	N	**E**	0.8716616	0.02300609±0.001594483	0.00119042	21.5%	78.2%	0.3%	0.08827654
	N	**G**	0.8706317	0.02197625±0.00165731	0.001197223	22.4%	77.2%	0.4%	0.08891658
	U	**U**	0.8707942	0.02213873±0.001567941	0.001018083	21.5%	78.1%	0.4%	0.08223334
	U	**E**	0.8701177	0.02146216±0.001511409	0.001054666	22.9%	76.8%	0.3%	0.08502393
	U	**G**	0.8695239	0.02086844±0.001523973	0.001039451	20.2%	79.2%	0.6%	0.08547874
	U	**N**	0.8702308	0.02157529±0.001418308	0.000988604	21.1%	78.6%	0.3%	0.08253104
	E	**E**	0.874889	0.02623354±0.002840129	0.002785832	24%	73.1%	2.9%	0.1173599
	E	**G**	0.8764657	0.02781023±0.002767141	0.002764614	25.9%	72%	2.1%	0.1175264
	E	**U**	0.8805018	0.03184634±0.00256148	0.00272041	27.1%	70.2%	2.7%	0.1081432
	E	**N**	0.8834488	0.03479331±0.00252527	0.002868897	30.4%	67.3%	2.3%	0.1080846
	G	**G**	0.8739253	0.02526984±0.002879656	0.002702171	22%	74.8%	3.1%	0.1168256
	G	**E**	0.8779107	0.02925522±0.002694412	0.002743778	25.8%	72%	2.2%	0.1171633
	G	**U**	0.8831964	0.03454093±0.002394667	0.002684303	27.8%	70.6%	1.6%	0.1076999
	G	**N**	0.8842994	0.03564391±0.002504351	0.00290145	29%	69.1%	1.9%	0.1092632
	Breast cancer model: $y_{ij} = \dfrac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

4.2.1.4. Results for the estimations of the variance parameters: REML- Method

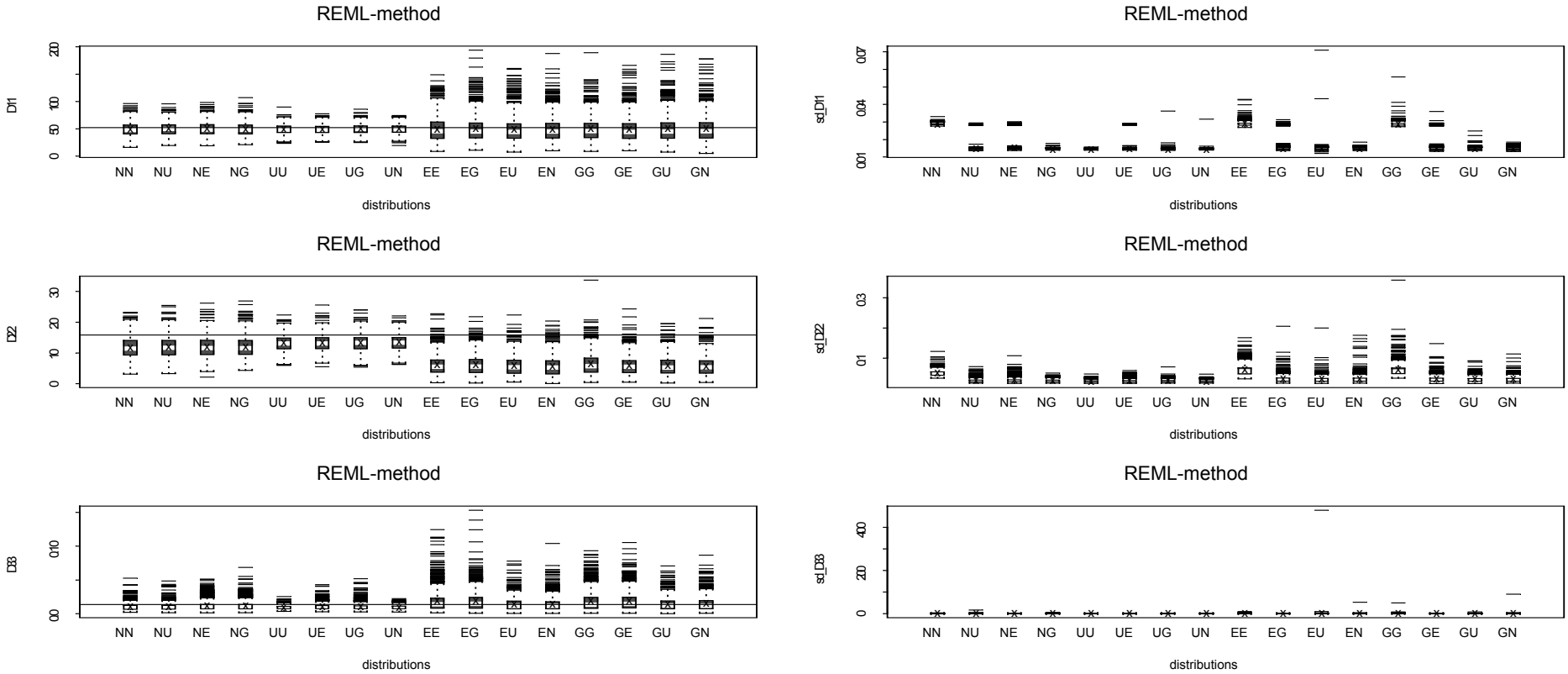


Figure 4.8: Box-plots of covariance parameter estimates from simulation settings of the breast cancer model using REML-method

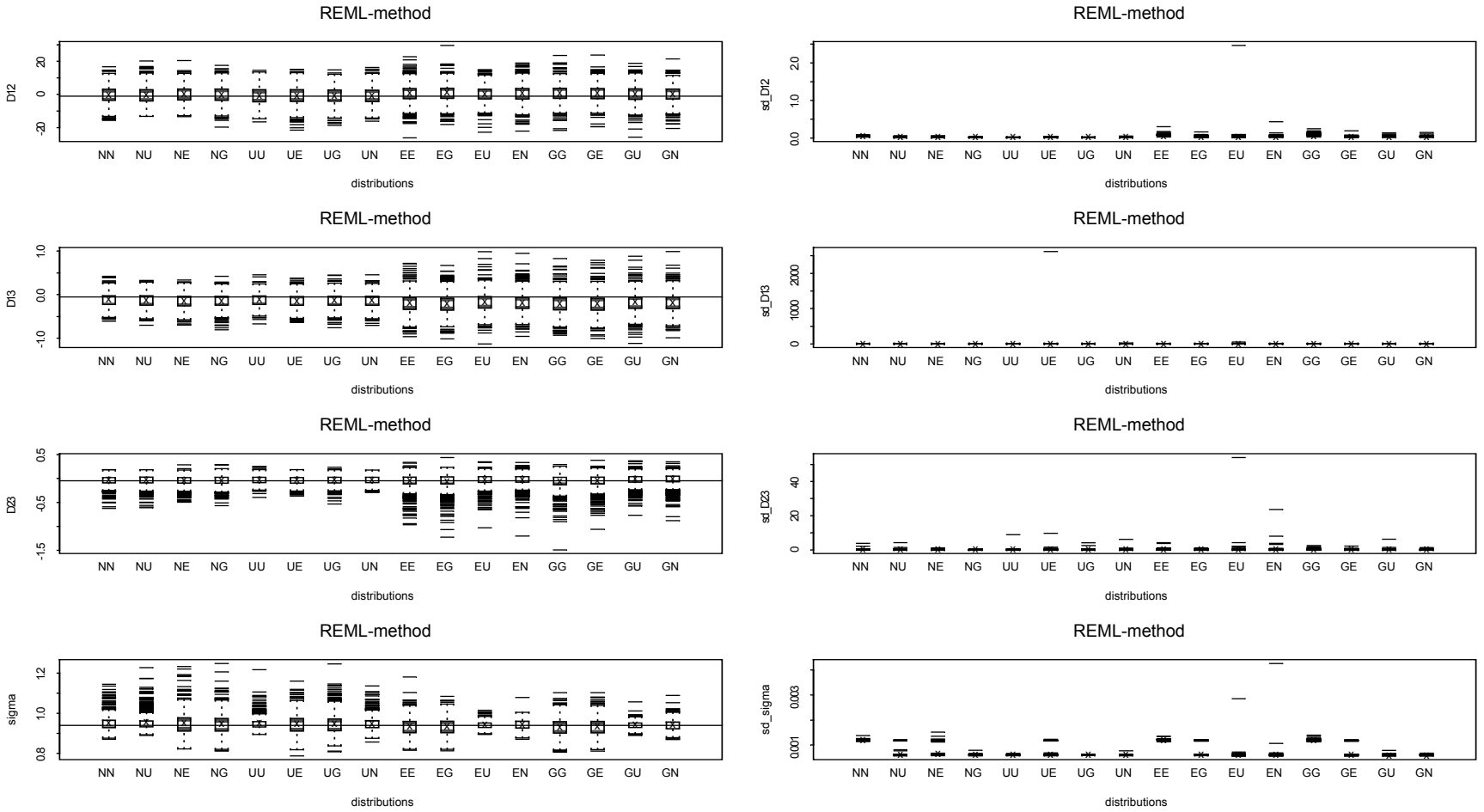


Figure 4.9: Box-plots of standard deviation of covariance parameter estimates from simulation settings of the breast cancer model using REML-method

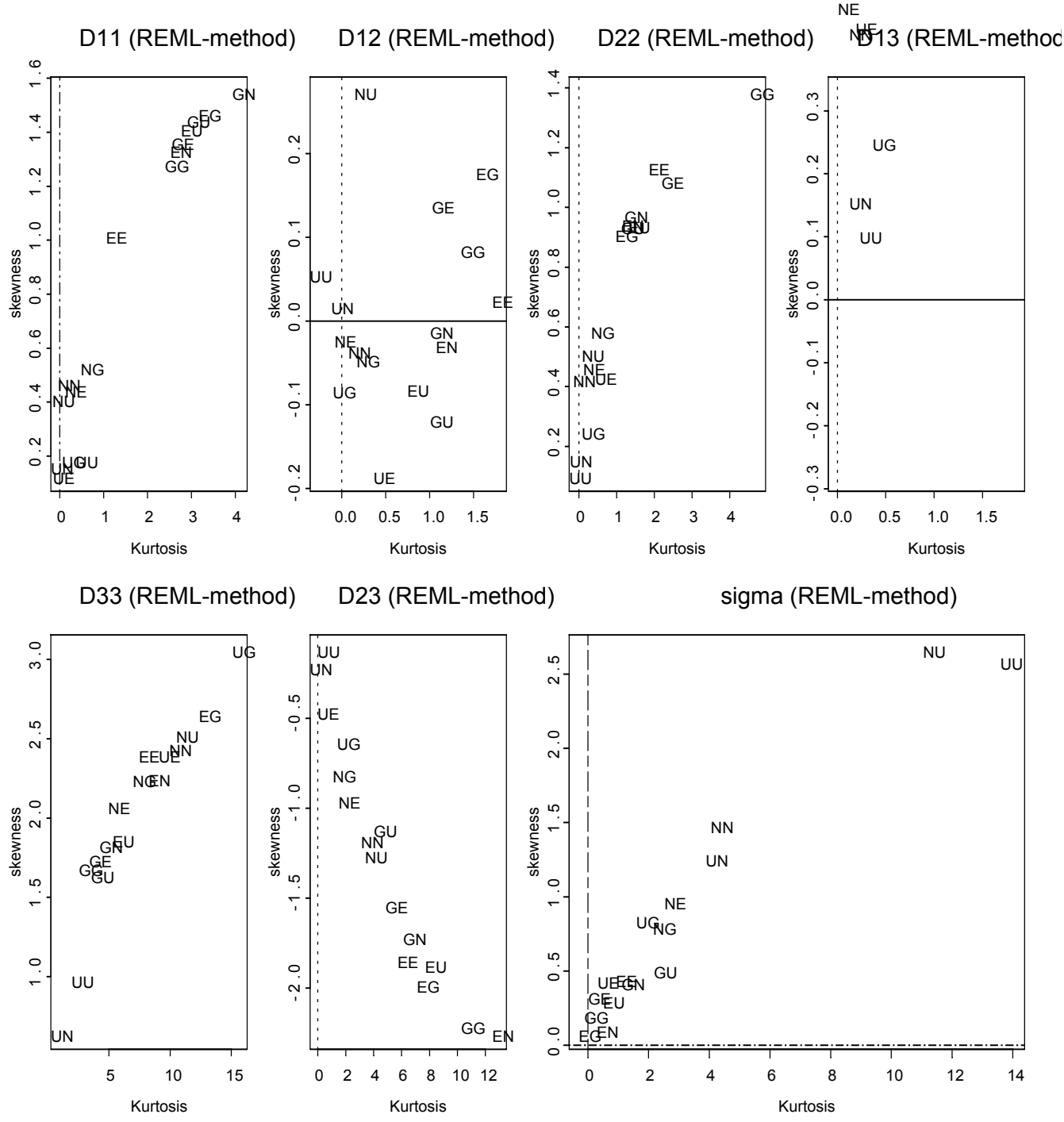


Figure 4.10: A char relating the distribution of random components to the values of kurtosis and skewness.

Table (4.14): summary results for fixed effect parameter estimate D_{11} of 1000 simulations

			D_{11}			$\sqrt{D_{11}}$			
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UP	
	N	**N**	49.23787	-2.648464±0.7629306	158.3786	0%	99%	1%	4.760906
	N	**U**	49.93401	-1.95232±0.7474479	149.0947	0.8%	95%	4.2%	3.383517
	N	**E**	49.56004	-2.32629±0.7528778	152.8133	1%	94.1%	4.9%	3.412721
	N	**G**	49.14441	-2.741918±0.7566096	156.3846	1.2%	93.2%	5.6%	3.333405
	U	**U**	48.8459	-3.040429±0.5636765	77.47834	0.1%	98.7%	1.2%	3.334254
	U	**E**	48.81395	-3.072379±0.5153007	78.49129	0%	98.4%	1.6%	3.360073
	U	**G**	49.46401	-2.422322±0.5280825	78.3875	0.1%	99.1%	0.8%	3.358366
	U	**N**	49.41447	-2.471861±0.5144745	74.94061	0%	98.9%	1.1%	3.355222
	E	**E**	49.17253	-2.713802±1.396441	514.4704	3.5%	84.5%	12%	4.675285
	E	**G**	50.28595	-1.600381±1.502268	589.4399	10.4%	69.9%	19.7%	3.346285
	E	**U**	48.84909	-3.037242±1.448406	554.7743	8%	69.8%	22.2%	3.267081
	E	**N**	49.24858	-2.637754±1.43447	542.0597	9.9%	68%	22.1%	3.27604
	G	**G**	49.81146	-2.07487±1.425349	509.8828	3.9%	87%	9.1%	4.710951
	G	**E**	48.79219	-3.094139±1.441068	549.6093	9.4%	67.4%	23.2%	3.286913
	G	**U**	50.23792	-1.648405±1.534719	615.2243	10%	68.3%	21.7%	3.300942
	G	**N**	49.62805	-2.258276±1.477045	572.4361	8.8%	69.9%	21.3%	3.283468
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.15): summary results for fixed effect parameter estimate D_{12} of 1000 simulations

			D_{12}			$D_{12}/\sqrt{D_{11}}\sqrt{D_{22}}$			
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	-0.271199	0.778663±0.3140402	26.25256	4.8%	95.1%	0.1%	0.7921005
	N	**U**	-0.4013559	0.6485061±0.3177944	26.68366	16.5%	83%	0.5%	0.575105
	N	**E**	-0.1220211	0.9278409±0.3022544	24.61827	15.8%	83.8%	0.4%	0.5865713
	N	**G**	-0.2254774	0.8243846±0.3099987	25.67	18.7%	80.4%	0.9%	0.5737349
	U	**U**	-0.8752175	0.1746445±0.3590649	27.71826	14.9%	83.9%	1.2%	0.5737438
	U	**E**	-0.731289	0.318573±0.3278987	28.0612	14.1%	84.5%	1.4%	0.5748307
	U	**G**	-0.802965	0.246897±0.319655	26.63247	13.7%	85.2%	1.1%	0.580137
	U	**N**	-0.7498625	0.2999995±0.3196626	26.66278	14.5%	84.1%	1.4%	0.5722661
	E	**E**	0.5326599	1.582522±0.3193108	29.0187	10.8%	89.1%	0.1%	0.8408088
	E	**G**	0.8179329	1.867795±0.3138666	29.10657	26.7%	72.1%	1.2%	0.6205535
	E	**U**	0.1322301	1.182092±0.3098958	26.37115	21.4%	76.8%	1.8%	0.6179933
	E	**N**	0.312494	1.362356±0.3225345	28.90841	23.7%	75%	1.3%	0.6280545
	G	**G**	0.7525964	1.802458±0.3249275	29.5224	10.9%	89%	0.1%	0.8492151
	G	**E**	0.6959339	1.745796±0.3009459	26.59993	24.7%	74.7%	0.6%	0.6278792
	G	**U**	0.2642966	1.314159±0.3180693	28.03557	23.3%	75.3%	1.4%	0.6194184
	G	**N**	0.1339154	1.183777±0.3048379	25.56657	23.6%	74.7%	1.7%	0.1279085
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.16): summary results for fixed effect parameter estimate D_{22} of 1000 simulations

			D_{22}			$\sqrt{D_{22}}$			
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	11.72086	-4.125635±0.2101231	28.50243	0%	96.5%	3.5%	2.650685
	N	**U**	11.96026	-3.886241±0.2036087	25.88355	0.2%	90.2%	9.6%	1.880518
	N	**E**	11.92206	-3.924442±0.2104575	26.91938	0.2%	89%	10.8%	1.901901
	N	**G**	11.97132	-3.875179±0.2168608	27.2467	0.2%	87.7%	12.1%	1.855918
	U	**U**	13.15851	-2.687993±0.1841406	14.50714	0%	96.7%	3.3%	1.899926
	U	**E**	13.30634	-2.540161±0.1618683	13.26602	0%	98.2%	1.8%	1.922787
	U	**G**	13.26464	-2.581859±0.1699902	14.18051	0.1%	97.6%	2.3%	1.908536
	U	**N**	13.39123	-2.455272±0.1595293	12.64648	0%	98.5%	1.5%	1.917351
	E	**E**	6.085665	-9.760835±0.2006676	105.7454	0%	62.1%	37.9%	2.421567
	E	**G**	6.083312	-9.763188±0.1998614	105.7073	0%	39.3%	60.7%	2.068211
	E	**U**	5.776094	-10.07041±0.3098958	110.9124	0%	36.2%	63.8%	1.638215
	E	**N**	5.6685	-10.178±0.1955757	113.5385	0%	34.9%	65.1%	1.687317
	G	**G**	6.350061	-9.496439±0.2182953	102.041	0.1%	64.2%	35.7%	2.479009
	G	**E**	5.883994	-9.962506±0.191578	108.7958	0.1%	35.4%	64.5%	2.092467
	G	**U**	5.818552	-10.02795±0.1895518	109.9032	0%	36.7%	63.3%	1.692269
	G	**N**	5.740848	-10.10565±0.1888611	111.3997	0%	35.2%	64.8%	1.682573
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.17): summary results for fixed effect parameter estimate D_{13} of 1000 simulations

			D_{13}			$D_{13}/\sqrt{D_{11}}\sqrt{D_{33}}$			
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROW LOW.	PROB.	PROB. UPP	
	N	**N**	-0.1276125	-0.0730064±0.009793162	0.03027009	0%	96.3%	3.7%	0.8558263
	N	**U**	-0.1340218	-0.0794157±0.009875651	0.03166893	0.2%	79.5%	20.3%	0.6258447
	N	**E**	-0.1516023	-0.09699614±0.01039358	0.0375003	0.5%	80.9%	18.6%	0.6300281
	N	**G**	-0.1441389	-0.08953279±0.00998768	0.0339569	0.6%	83.1%	16.3%	0.6212637
	U	**U**	-0.1239224	-0.06931627±0.01018017	0.02706101	0.4%	84.6%	15%	0.633535
	U	**E**	-0.1481042	-0.09349808±0.00964981	0.03295728	0.4%	81.2%	18.4%	0.6331063
	U	**G**	-0.1393729	-0.08476672±0.01015761	0.03401635	0.6%	81.8%	17.6%	0.630799
	U	**N**	-0.1386078	-0.08400162±0.00994372	0.0327692	0.3%	80.5%	19.2%	0.6320126
	E	**E**	-0.2043566	-0.1497505±0.01372025	0.07137798	0.1%	93.7%	6.2%	0.8774471
	E	**G**	-0.2157093	-0.1611032±0.01362487	0.07422874	0.4%	73.3%	26.3%	0.6205535
	E	**U**	-0.1711549	-0.1165488±0.05987566	0.6563194	0.8%	75.9%	23.3%	0.6563194
	E	**N**	-0.192846	-0.1382398±0.01290879	0.06244383	1.6%	71.7%	26.7%	0.6475584
	G	**G**	-0.2119881	-0.1573819±0.01419712	0.07492777	0.2%	92.8%	7%	0.8776281
	G	**E**	-0.2139915	-0.1593854±0.01390626	0.07569282	0.8%	72.6%	26.6%	0.1180483
	G	**U**	-0.1856996	-0.1310934±0.01334697	0.06351086	0.8%	74.2%	25%	0.6414068
	G	**N**	-0.1895675	-0.1349613±0.01331982	0.06435166	1%	74.9%	24.1%	0.6438032
	Breast cancer model: $y_{ij} = \dfrac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.18): summary results for fixed effect parameter estimate D_{23} of 1000 simulation

			D_{23}			$D_{23}/\sqrt{D_{22}}\sqrt{D_{33}}$			
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{Bias}$	MSE	COVERAGE-PROBABILITY (ASYMP. INTERVALS)			Average width
						PROW LOW.	PROB.	PROB. UPP	
	N	**N**	-0.04148868	0.004383925±0.006143342	0.009833595	0.1%	96.9%	3%	0.9018851
	N	**U**	-0.03834038	0.007532218±0.006200897	0.01005587	0.8%	87.1%	12.1%	0.6738794
	N	**E**	-0.04667984	-0.0008072438±0.006298924	0.01031843	1%	86.7%	12.3%	0.6802092
	N	**G**	-0.04493325	0.0009393517±0.006549933	0.01115736	1.4%	84.2%	14.4%	0.6590821
	U	**U**	-0.03201972	0.01385288±0.005805318	0.00742949	1.5%	87.8%	10.8%	0.637856
	U	**E**	-0.04022707	0.00564553±0.005522042	0.007961498	0.4%	88.3%	11.3%	0.6477995
	U	**G**	-0.03842681	0.007445795±0.005548065	0.00805998	1%	88.7%	10.3%	0.653118
	U	**N**	-0.03211615	0.01375645±0.00513234	0.00703914	0.6%	88.9%	10.5%	0.6444822
	E	**E**	-0.05827333	-0.01240073±0.009250401	0.02240606	0.7%	92%	7.3%	0.9934135
	E	**G**	-0.0620221	-0.0161495±0.009740625	0.02493409	3.5%	76.6%	19.9%	0.7334609
	E	**U**	-0.1294606	0.01554298±0.007599968	0.01526182	3.6%	83.3%	13.1%	0.7667045
	E	**N**	-0.03169916	0.01417344±0.007525869	0.01492966	2.6%	85%	12.4%	0.7745956
	G	**G**	-0.06662204	-0.02074944±0.009746867	0.02407208	0.8%	91.7%	7.4%	0.9965582
	G	**E**	-0.05834277	-0.01247017±0.008646944	0.0195992	2.9%	76.8%	20.3%	0.7352493
	G	**U**	-0.02451789	0.02135471±0.006962812	0.01306334	3.4%	86.1%	10.5%	0.7628167
	G	**N**	-0.02241842	0.02345418±0.007461753	0.01502898	3.4%	86.4%	10.2%	0.7669228
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.19): summary results for fixed effect parameter estimate D_{33} of 1000 simulations

			D_{33}			$\sqrt{D_{33}}$			
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{Bias}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	0.01016107	-0.003466843±0.0003285	0.000040087	1.7%	97.5%	0.8%	0.102524
	N	**U**	0.01011391	-0.003514±0.00032875	0.00004045	4%	89.1%	6.9%	0.07100997
	N	**E**	0.01152335	-0.002104563±0.0004417	0.000055197	10.1%	83.8%	6.1%	0.07690407
	N	**G**	0.01182387	-0.001804038±0.0004625	0.00005890	11.6%	81.8%	6.6%	0.07586607
	U	**U**	0.00924998	-0.004377926±0.0001853	0.000026543	0.6%	97.5%	1.9%	0.06693267
	U	**E**	0.0103687	-0.00325921±0.00028563	0.000031838	4.3%	92.5%	3.2%	0.07090216
	U	**G**	0.01044391	-0.003184004±0.0003074	0.00003478	4.1%	93.4%	2.5%	0.07083294
	U	**N**	0.00937070	-0.004257201±0.0001772	0.000026293	0.3%	96.8%	2.9%	0.06706406
	E	**E**	0.01859911	0.004971198±0.00101038	0.000290187	23.1%	76.1%	0.8%	0.1490149
	E	**G**	0.0187623	0.005134393±0.00096029	0.000266167	38.3%	58.2%	3%	0.1048519
	E	**U**	0.01401668	0.0003887741±0.0006128	0.000097806	26.9%	68.4%	4.7%	0.09184795
	E	**N**	0.01429697	0.000669061±0.00066301	0.000114760	25.4%	69.1%	5.5%	0.09296233
	G	**G**	0.01845523	0.004827315±0.00095031	0.000248043	25%	73.7%	1.4%	0.1491286
	G	**E**	0.01859519	0.004967275±0.00089811	0.000234431	38.6%	58.2%	3.2%	0.1037875
	G	**U**	0.01412249	0.0004945802±0.0005964	0.000092748	28%	66.6%	5.4%	0.09107981
	G	**N**	0.01483412	0.00120621±0.000678158	0.000121050	28.4%	67.1%	4.5%	0.09266399
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

Table (4.20): summary results for fixed effect parameter estimate σ of 1000 simulations

Restricted maximum likelihood (REML)									
DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{\text{Bias}}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW.	PROB.	PROB. UPP	
	N	**N**	0.9516251	0.01170407±0.002129576	0.001316327	7%	92.7%	0.3%	0.1286425
	N	**U**	0.9526507	0.01272971±0.002129301	0.001341083	9.9%	89.9%	0.2%	0.09164872
	N	**E**	0.9488619	0.008940869±0.00322782	0.002789332	19.9%	66.6%	13.5%	0.0926115
	N	**G**	0.9454825	0.005561527±00.0031124	0.002549922	18.7%	67.2%	14.1%	0.09048876
	U	**U**	0.9491631	0.009242132±0.00193376	0.000888479	7.3%	92.6%	0.1%	0.09067223
	U	**E**	0.9452451	0.005324082±0.00309168	0.002514011	19.2%	66.4%	14.4%	0.09105571
	U	**G**	0.9461909	0.006269855±0.00318916	0.002684205	19.5%	65.9%	14.6%	0.09046926
	U	**N**	0.9485474	0.008626433±0.00196061	0.001074032	9.6%	87.9%	2.5%	0.0906258
	E	**E**	0.9321388	-0.06918362±0.00503364	0.01137535	4.9%	84.2%	10.9%	0.1255092
	E	**G**	0.9325082	-0.007412844±0.0026347	0.001860196	10.2%	69.2%	20.6%	0.08978356
	E	**U**	0.9409825	0.001061541±0.00105172	0.000288770	0.9%	99%	0.1%	0.08953451
	E	**N**	0.9428831	0.002962113±0.00151227	0.000603496	3.5%	94%	2.5%	0.08978629
	G	**G**	0.9310265	-0.00889454±0.00281702	0.002053931	5.6%	82.8%	11.6%	0.1254111
	G	**E**	0.9321772	-0.00774385±0.0027160	0.001978303	11.3%	68.4%	20.3%	0.08948759
	G	**U**	0.9407801	0.0008591176±0.0010343	0.000278938	0.9%	98.6%	0.5%	0.08938983
	G	**N**	0.9405068	0.0005857754±0.0015446	0.000620806	3.7%	93%	3.3%	0.08936629
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$								

4.2.1.5. Robustness of the LB method under misspecification of the structure of the covariance matrix of random effects and/or the correlation structure of the residuals

The modeling process using mixed models is highly interactive and conducted in a step-by-step way, progressively incorporating hypotheses about the covariance structure, among other refinements. Crowder (1995, 2001) points that the use of an incorrect correlation matrix could cause an inconsistency problem, when using the generalized estimating equations (GEE) for analyzing repeated measures. His results are based on asymptotic theory, but Park and Yun Shin (1999) showed by simulation that the GEE estimation method is robust under misspecifications of the correlation structure, even for moderate sample sizes. Sutradhar and Das (1999) obtain similar results and suggest a more robust estimation framework. Similarly, Chaganty (1997) introduces an alternative method in order to overcome these pitfalls of the GEE approach.

In non-linear mixed effects models, the LB method is based on a Taylor-series expansion of the likelihood around successive approximations to the values of the fixed parameters and random effects. It consists in a two-step iterative process alternating between maximization of the log likelihood, given the above-mentioned values, and the generation of new parameter values using the linear mixed-effects methodology. The algorithm is based on the assumption that the intra-individual covariance matrix does not depend on the fixed parameters. This assumption can be justified by the fact that the matrix obtained by differentiation of the objective function respect to the random effects, varies somewhat slowly with respect to fixed effects (Bates and Watts, 1980). Then, it is reasonable to pose the question on the performance of these methods when the assumed covariance and/or correlation structures are inappropriate. This is a recurring question in the generalized linear models and generalized linear mixed models literature. In this section, we try to make some similar insights on the performance of a nonlinear mixed effects fitting procedure, the LB method, that is widely used by practitioners.

We present a simulation study, consisting in two differentiated parts. In the first part, the consequences of fitting inappropriate restricted models were investigated. In the second part, the reverse case was investigated: the fitting process was performed under unrestricted conditions but data were generated according to some true structuring patterns. More precisely, the simulation scenarios were:

In the first series of simulations, data were generated from the breast cancer model in the same way as designed in section 4.2.1, and with the same parameter values, but each one of generated data tables was processed to fit a nonlinear mixed model like (4.1)-(4.2) via the LB-method assuming in the fitting process one of the following structures of the covariance matrix of random effects:

i. Independent random effects: Thus, their covariance matrix is diagonal, and is denoted as D_ind in Table (4.21) and (4.22).

ii. Block-diagonal: The first two random effects, $(\eta_{1i}, \eta_{2i})'$, have equal variance and are independent of the third one, η_{3i}. This can be rephrased by saying that the covariance matrix of the random effects D can be partitioned into four blocks as follows:

$$D = \begin{bmatrix} D_{11} & D_{12} & 0 \\ D_{21} & D_{11} & 0 \\ 0 & 0 & D_{33} \end{bmatrix}.$$

The $(\eta_{1i}, \eta_{2i})'$ form a 2×2 block of random effects with compound symmetry covariance matrix, and the η_{3i} forms another 1×1 block with an unstructured covariance matrix. The corresponding covariance matrix is denoted as D_bck in the results.

iii. Unstructured covariance matrix: a general covariance matrix for the random effects (i.e. random effects have an arbitrary covariance structure) is assumed. This case is denoted as D_unst.

Each one of the preceding assumptions on the covariance structure of the random effects was combined, in the fitting process, with one of the following assumptions on the residuals correlation structure, used to model the within-group correlation:

i. Independent residuals,

ii. Autoregressive, AR(1), residuals, and

iii. Autoregressive moving average of order 1, MA(1), residuals.

The fitting conditions defined by D_unst and independent residuals were introduced to have a reference condition.

In the second part of this simulation study, the simulated data sets were generated according to three possible "true" structures of the random effects covariance matrix (instead of the matrix D defined in section 4.2.1) that were combined with each one of three possible correlation structures for the residuals.

The possible structures for the covariance matrix were:

i. Uncorrelated random effects:

$$D\_ind = \begin{pmatrix} 41.889 & 0.00000 & 0.000000000 \\ 0.000 & 17.25313 & 0.000000000 \\ 0.000 & 0.00000 & 0.004827908 \end{pmatrix}.$$

ii. Block-diagonal:

$$D\_bck = \begin{pmatrix} 20.25 & 0.00 & 0.0000 \\ 0.00 & 20.25 & 0.0000 \\ 0.00 & 0.00 & 0.0064 \end{pmatrix}.$$

iii. Unstructured covariance matrix:

$$D\_Unst = \begin{pmatrix} 48.00 & -1.00 & -0.05 \\ -1.00 & 16.00 & -0.05 \\ -0.05 & -0.05 & 0.02 \end{pmatrix}.$$

The residuals were always generated according to a normal distribution with zero mean, but with correlation specified by one of the following structures:

i. Uncorrelated: for each individual i,

$$cor(e_{ij}, e_{ij+1}) = 0,\ j = 1,\ldots,n_i - 1$$

ii. Auto-regressive of order 1, AR (1):

$$cor(e_{ij}, e_{ik}) = \phi^{|j-k|},\ \ j,k = 1,\ldots,n_i; \text{ with } \phi = 0.2740697$$

iii. Moving average of order 1, MA(1):

$cor(e_{ij}, e_{ij+1}) = \alpha_1 \big/ (1+\alpha_1^2)$, where $e_{ij} = \alpha_1 e_{ij-1} + u_{ij}$; with $\alpha_1 = 0.2095072$.

To sum up, there were 18 simulation experimental conditions. In the first part of the simulation study, data always generated according to the same parameters (defining an unstructured covariance matrix for the random factors and uncorrelated residuals) were analyzed according to 3x3=9 possible structure assumptions. In the second part of the study, data were generated according to 9 possible dependency structures for random factors and residuals, but always analyzed assuming an unstructured covariance matrix and no residuals correlation.

The simulated random effects were generated according to an expression equivalent to $\eta_i = LH$, where H stands for a standardised version of the vector of random effects, generated from a normal distribution with zero mean and unit variance, and L stands for the lower triangular matrix resulting from the Cholesky decomposition of a covariance matrix D with one of the possible structures cited previously. The simulated residuals for each individual were generated by an expression equivalent to $e_i = \sigma \Gamma^{1/2} u_i$, where u_i stands for an standardized vector of i.i.d residuals, generated from the normal distribution with zero mean and variance I_{n_i}; and Γ is the correlation matrix with the possible structures also cited and discussed above. In order to obtain each new simulated response in accordance with model (4.1)-(4.2), each generated vector of random factors and residuals was added to the expected response. For each one of the 18 experimental conditions, the entire process was repeated $N = 1000$ times.

For each generated data table, a nonlinear mixed model was fitted using the *nlme* S-Plus library, providing a set of (simulated) estimated values $\hat{\delta}, \hat{\eta}, \hat{D}$ and $\hat{\sigma}$ obtained according to the LB method. The bias and the mean square error (MSE) of these estimators was estimated from each series of 1000 simulated values, together with the true coverage of the associated asymptotic confidence intervals, always computed at a nominal 95% level.

For some simulated data tables the LB estimation methods did not converge. In these cases, we discarded the problematic pseudodata set and generated a completely new table, until convergence was reached. The possible consequences of this strategy are discussed in El Halimi et al.(2004).

1.2.1.2.1. Results for fixed effects

Table (4.21) and Figure 4.11 display the simulation results for fixed effects in the first series of simulations, where the means over 1000 simulations of δ_1 are close to the true value for all simulations runs. The bias is negative but considerably small (especially when the residuals are independent) and significantly different from 0. This suggests that, on average, they will slightly underestimate the true parameter value, and shows a similar tendency under each one of the covariance structure of random effects. The same patterns are also seen for the coverage probability measures and the average widths of the intervals. All simulations produced nearly identical results in terms of MSEs. With respect to its true coverage probability, the results are quite similar when the model is fitted under the D_unst and D_ind model assumptions; both achieving the same maximum coverage, 93.8%, when the residuals are assumed independent, but the observed coverages are also quite similar under the other assumptions for the residuals. The coverages for D_bck demonstrate the imprecise nature of the corresponding intervals, with a maximum coverage of 85.5% under the assumption of independent residuals and similar but lower coverages under the remaining assumptions on the residuals. In all conditions the intervals are appreciably equitailed, in the sense that the non coverage cases are nearly symmetrically distributed, approximately one half of the times the true parameter value is at left of the confidence interval, and the other half of the times it is at right. Thus, it is clear that the performance of the parametric intervals depends on the assumed structure for the matrix of the random effects, more than on assumptions for the residuals correlation structure.

For the δ_2 parameter the results shown in Table (4.21) suggest also a negative bias. Again, the bias and the MSE values are nearly identical, except for (D_bck, AR (1)) conditions, where the MSE is roughly 50% larger than under other conditions. The coverage results are nearly inverted: the confidence intervals have its lowest coverage when the covariance matrix of random effects has the D_ind and D_unst structures, and the largest for the D_bck structure, achieving maximum coverage, 98.2%, under (D_bck, Independent) and (D_bck, MA(1))

conditions. The confidence intervals are highly asymmetric, and the interval widths are similar in all conditions, except under D_bck, characterized by large interval lengths.

With respect to bias, the estimates of δ_3 are good and not affected by wrong assumptions on the structures of random effects and residuals. On the other hand, the confidence intervals are very sensible (and always very incorrect) to these assumptions. The interval widths are similar, and the intervals exhibit considerable asymmetry. The maximum coverage, 79.3%, occurs at D_bck / MA(1) conditions.

In general, all the confidence intervals for the fixed-effects parameters seem to be more affected by the structure of the covariance matrix random effects than by the correlation structure of the residuals.

The results of the second series of simulations are summarized in Table (4.22) and in Figure 4.12. In all cases, the point estimates of δ remain virtually unaffected by the choice of the residuals correlation structure and are slightly affected by the choice of the random effects covariance structure. Table (4.22) confirms that the averages of the fixed effects estimates are close to the true parameters for all simulations.

In accordance with the results of the simulation studies described in the preceding sections of this chapter, the true coverage of the confidence intervals based on the LB-method is not always adequate, even under a correct specification of the covariance structure of random effects and residuals. It depends, mainly, on the specific model parameters. In any case, regarding each one of the assumed random effects covariance structures, we remark that the confidence interval coverage probabilities change with the structure of the matrix D and perform poorly when D_bck is used to generate data. On the other hand, it seems not to be affected by the correlation structure of residuals.

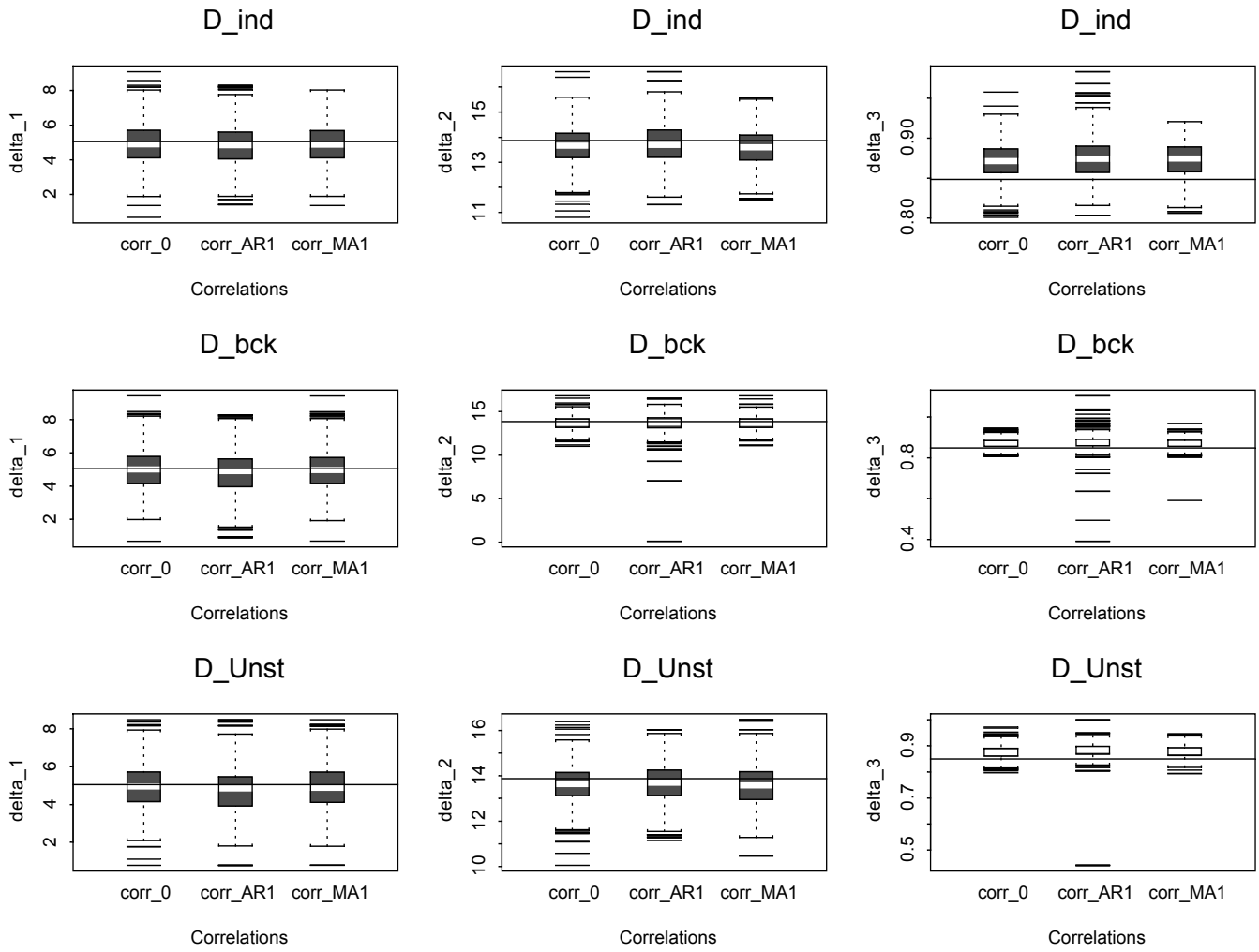


Figure 4.11: Breast cancer data. Box-plots of 1000 simulation replicates of fixed effect estimates under different structures of D and correlation structures (First part of the simulation study))

Table (4.21): Simulation results for the *breast-cancer model* under different dependence structures for random effects and residuals (simulation study 1st part)

		δ_1						
Structure of D	*Working*	*Mean*	*Bias±C.I$_{Bias}$*	*MSE*	*PROB. LOW*	*COV. PROB*	*PROB. UPP*	*Average width*
D_ind	Independent	4.890293	-0.160943±0.07338365	1.426302	2.1%	93.8%	4.1%	4.470603
	AR(1)	4.836839	-0.2143969±0.07536222	1.522899	3.1%	92.1%	4.8%	4.407850
	MA(1)	4.891355	-0.1598807±0.07399343	1.446481	3.6%	92.3%	4.1%	4.490843
D_bck	Independent	4.981798	-0.06943841±0.0750465	1.469406	6.1%	85.5%	8.4%	3.605085
	AR(1)	4.820172	-0.2310643±0.08055745	1.740972	6.8%	81.2%	12%	3.515189
	MA(1)	4.973567	-0.07766946±0.07423818	1.439237	7.1%	85.2%	7.7%	3.607577
D_Unst	Independent	4.959736	-0.09150034±0.07311737	1.398627	2.9%	93.8%	3.3%	4.469127
	AR(1)	4.743657	-0.3075795±0.07441145	1.534507	2.6%	92.1%	5.3%	4.376982
	MA(1)	4.882728	-0.168508±0.07377075	1.443608	2.3%	93.3%	4.4%	4.443382
				δ_2				
D_ind	Independent	13.6556	-0.2114322±0.04648119	0.6065372	2.9%	86.9%	10.2%	2.419676
	AR(1)	13.7077	-0.1593291±0.0499392	0.6739253	3.3%	87.7%	9%	2.387804
	MA(1)	13.5767	-0.2903252±0.04663594	0.6487374	3%	86.5%	10.5%	2.415387
D_bck	Independent	13.68257	-0.1844598±0.04613696	0.5875683	0.5%	98.2%	1.3%	3.661344
	AR(1)	13.61481	-0.2522202±0.06957238	1.322329	1.4%	92%	6.6%	3.599156
	MA(1)	13.67776	-0.1892706±0.04545654	0.5731596	0.3%	98.2%	1.5%	3.66568
D_Unst	Independent	13.62828	-0.2387503±0.04928767	0.6887295	2.9%	86.6%	10.5%	2.461421
	AR(1)	13.64761	-0.2194167±0.05337408	0.7889661	4.2%	83.3%	12.5%	2.372432
	MA(1)	13.56978	-0.2972548±0.05398211	0.8461577	2.3%	84.6%	13.1%	2.474254
				δ_3				
D_ind	Independent	0.8713735	0.02271801±0.001403027	0.001028008	22.3%	77.3%	0.4%	0.08177437
	AR(1)	0.8749249	0.02626941±0.001611147	0.001365113	26.4%	73.6%	0%	0.08163308
	MA(1)	0.8736767	0.02502116±0.00139684	0.001132438	29.2%	70.8%	0%	0.08218339
D_bck	Independent	0.8697888	0.02113329±0.001365946	0.0009318155	20.8%	79%	0.2%	0.08095256
	AR(1)	0.8763528	0.02769726±0.002709488	0.002676234	29.5%	69.6%	0.9%	0.08205351
	MA(1)	0.8697713	0.02111575±0.001477302	0.001013409	20.6%	79.3%	0.1%	0.08339257
D_Unst	Independent	0.8751281	0.02647256±0.001485831	0.001274902	27%	72.9%	0.1%	0.08439246
	AR(1)	0.8813267	0.03267124±0.002452305	0.002631287	36.8%	62.7%	0.5%	0.08965067
	MA(1)	0.8784579	0.02980238±0.001465699	0.001446835	31.3%	68.5%	0.2%	0.08541293

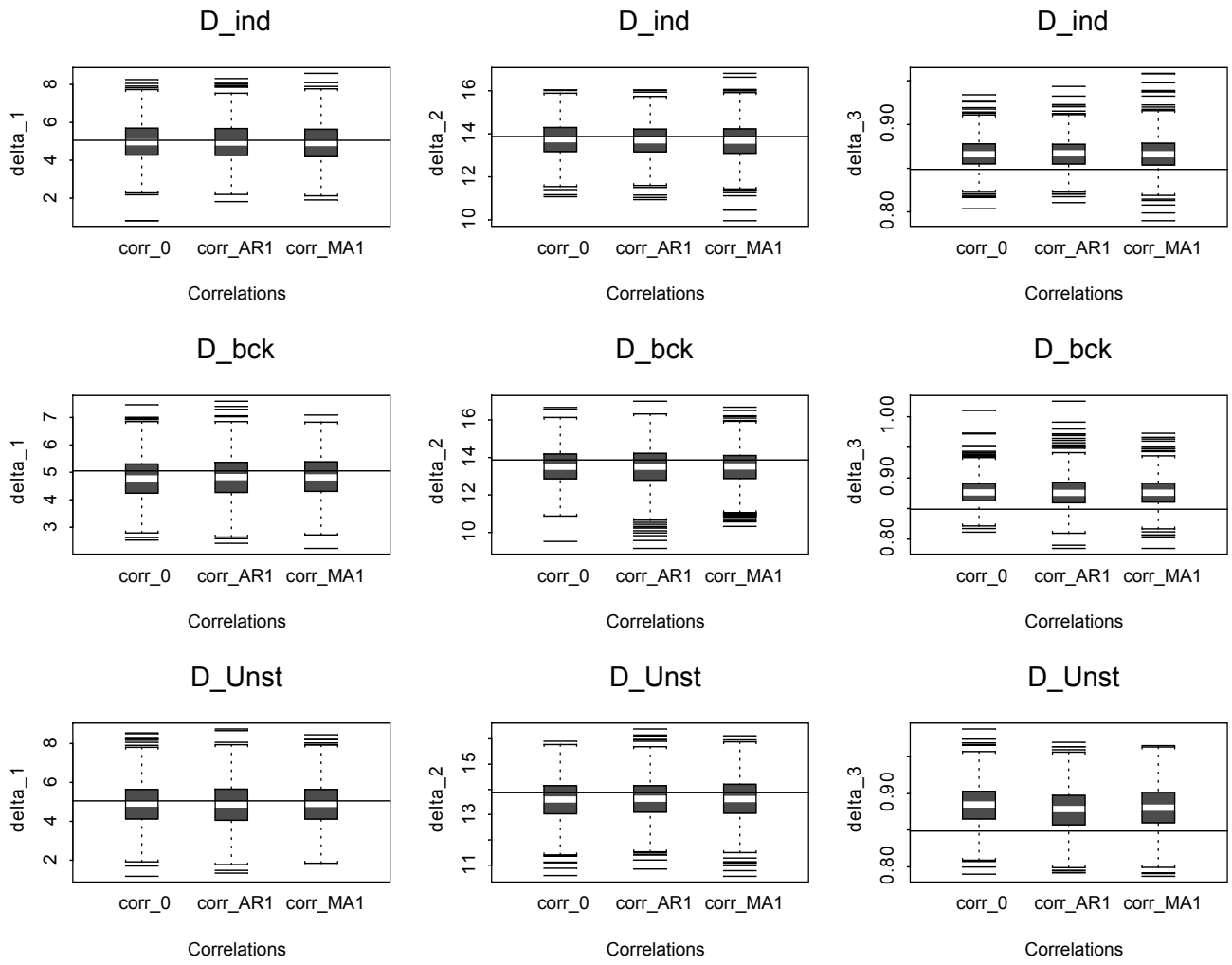


Figure 4.12: Breast cancer data. Box-plots of 1000 simulation replicates of fixed effect estimates under different structures of D and correlation structures (Second part of the simulation study).

Table (4.22): Simulation results for the *breast-cancer model* under different dependence structures for random effects and residuals (simulation study 2nd part)

		δ_1						
Structure of D	*Working*	*Mean*	*Bias±C.I$_{Bias}$*	*MSE*	*PROB. LOW*	*COV. PROB*	*PROB. UPP*	*Average width*
D_ind	Independent	4.972995	-0.0782413±0.06598653	1.138428	2.9%	93.6%	3.5%	4.055087
	AR(1)	4.950308	-0.100928±0.06366188	1.064118	2.2%	94%	3.8%	4.022737
	MA(1)	4.902534	-0.1487024±0.06559291	1.14095	2.2%	93%	4.8%	4.03939
D_bck	Independent	4.78057	-0.2706657±0.04801458	0.672774	1.7%	91.1%	7.2%	2.871522
	AR(1)	4.809881	-0.2413553±0.04955735	0.696912	1.4%	91.3%	7.3%	2.919448
	MA(1)	4.821342	-0.2298941±0.0468499	0.623634	0.6%	92.7%	6.7%	2.896072
D_Unst	Independent	4.887868	-0.1633685±0.07022588	1.309161	2%	93.2%	4.8%	4.305583
	AR(1)	4.852252	-0.1989843±0.0699953	1.313658	1.9%	93.7%	44%	4.342266
	MA(1)	4.904378	-0.1468583±0.06974602	1.286572	2.7%	93%	43%	4.352609
				δ_2				
D_ind	Independent	13.86703	-0.1537563±0.04941836	0.6587233	3.9%	87.4%	8.7%	2.531632
	AR(1)	13.69579	-0.1712411±0.04902582	0.6543568	3.8%	87.7%	8.5%	2.524353
	MA(1)	13.67516	-0.1918734±0.05365122	0.7853511	5.2%	84.9%	9.9%	2.533162
D_bck	Independent	13.5189	-0.3481297±0.06002156	1.05804	4.5%	79.6%	15.9%	2.691222
	AR(1)	13.48332	-0.3837145±0.06705982	1.316677	5.2%	76.9%	17.9%	2.679371
	MA(1)	13.48008	-0.3869503±0.06071217	1.108259	3.6%	81%	15.4%	2.689977
D_Unst	Independent	13.57039	-0.2966377±0.05084926	0.760386	3.3%	82.8%	13.9%	2.447999
	AR(1)	13.60609	-0.260937±0.05064817	0.7351727	3.6%	84.2%	12.2%	2.458311
	MA(1)	13.61227	-0.2547586±0.05246764	0.7807756	3.3%	83.9%	12.8%	2.448161
				δ_3				
D_ind	Independent	0.8665674	0.017912±0.001100081	0.000636	29.3%	70.1%	0.6%	0.05623834
	AR(1)	0.8666122	0.017959±0.001068346	0.000619	26.5%	72.9%	0.6%	0.05705633
	MA(1)	0.8662308	0.017575±0.001220561	0.000696	25.4%	74.1%	0.5%	0.06123209
D_bck	Independent	0.8779196	0.029264±0.001440842	0.001396	40.1%	59.8%	0.1%	0.06746345
	AR(1)	0.8766831	0.028028±0.001481441	0.001481	37%	62.4%	0.6%	0.07307422
	MA(1)	0.8486555	0.028395±0.001471719	0.001369	38.3%	61.2%	0.5%	0.07119195
D_Unst	Independent	0.8842841	0.035628±0.03562864	0.002110	32.4%	67.5%	0.1%	0.09821503
	AR(1)	0.8779992	0.029343±0.001854973	0.001755	24.6%	75.3%	0.1%	0.1018569
	MA(1)	0.8486555	0.031735±0.001841488	0.001887	26.8%	72.9%	0.3%	0.09997221

1.2.1.2.2. Results for variance components

Figure 4.13 summarizes the main results of the first series of simulations, those testing the impact on the LB method on the random components when different, inadequate, structures of the covariance matrix of random effects and/or correlation structures of the residuals are imposed. Figure 4.13 displays the results for the ML estimations of the terms of the random factors covariance matrix, D, and for the standard deviation of the residuals, σ. The bias and the overall distribution of the estimates of the components of D, and the coverage of the corresponding asymptotic confidence intervals, depends on the parameter in question and on the covariance of random effects imposed, more than on the correlation structure of the residuals. In general terms, the inference on the components of D is even less robust than the inference on the fixed effects, but adequate under the (correct) assumption of unstructured covariance matrix. For all components of D, the coverage in the D_unst case always lies near the nominal 95% level, while the coverages under D_bck are very erratic and affected by both, the covariance of the random effects and the correlation of residuals. Clearly, this assumption on the structure of the covariance matrix is too strong when it is incorrect.

As expected, the results for the estimator of the standard deviation of the residuals, σ, show a clear dependency on the correlation structure of the residuals, rather than that of the random effects. The coverage of the asymptotic confidence intervals for this parameter are always very inadequate and range from 49.3% in the (D_bck, MA(1)) case to 42.9% in the (D_Unst, AR(1)) case, with 45.3% of coverage under (D_Unst, independent) conditions.

The results corresponding to the estimation of the random part in the second series of simulations are displayed in Figures 4.14 to 4.17, which are box-plots of the estimates of the variance components (D and σ) for all simulation runs. They can be interpreted in the same way as Figure 4.13. As expected, the median values of all the components of D change considerably with the true values of D used in the simulations, and the variability of the estimations of the D components changes considerably with the different assumed structures of D, especially for D_{23} and D_{33}, but both the point estimates and the confidence intervals are

correct when the process of estimation is performed assuming an unstructured covariance matrix and not its exact structure. The precision of the estimates of the covariance matrix is nearly not affected by incorrectly assuming independent residuals instead of the adequate correlation model for the residuals.

Finally, the variability of the point estimates of the parameter σ is affected by the true values and the dependence structure of both, the random effects and the residuals, but not its bias nor the true coverage of the asymptotic confidence intervals, that are quite correct in all conditions. That is, again the results are acceptable when the estimation process is performed under the assumption of unstructured covariance matrix and independent residuals, instead of assuming its correct nature.

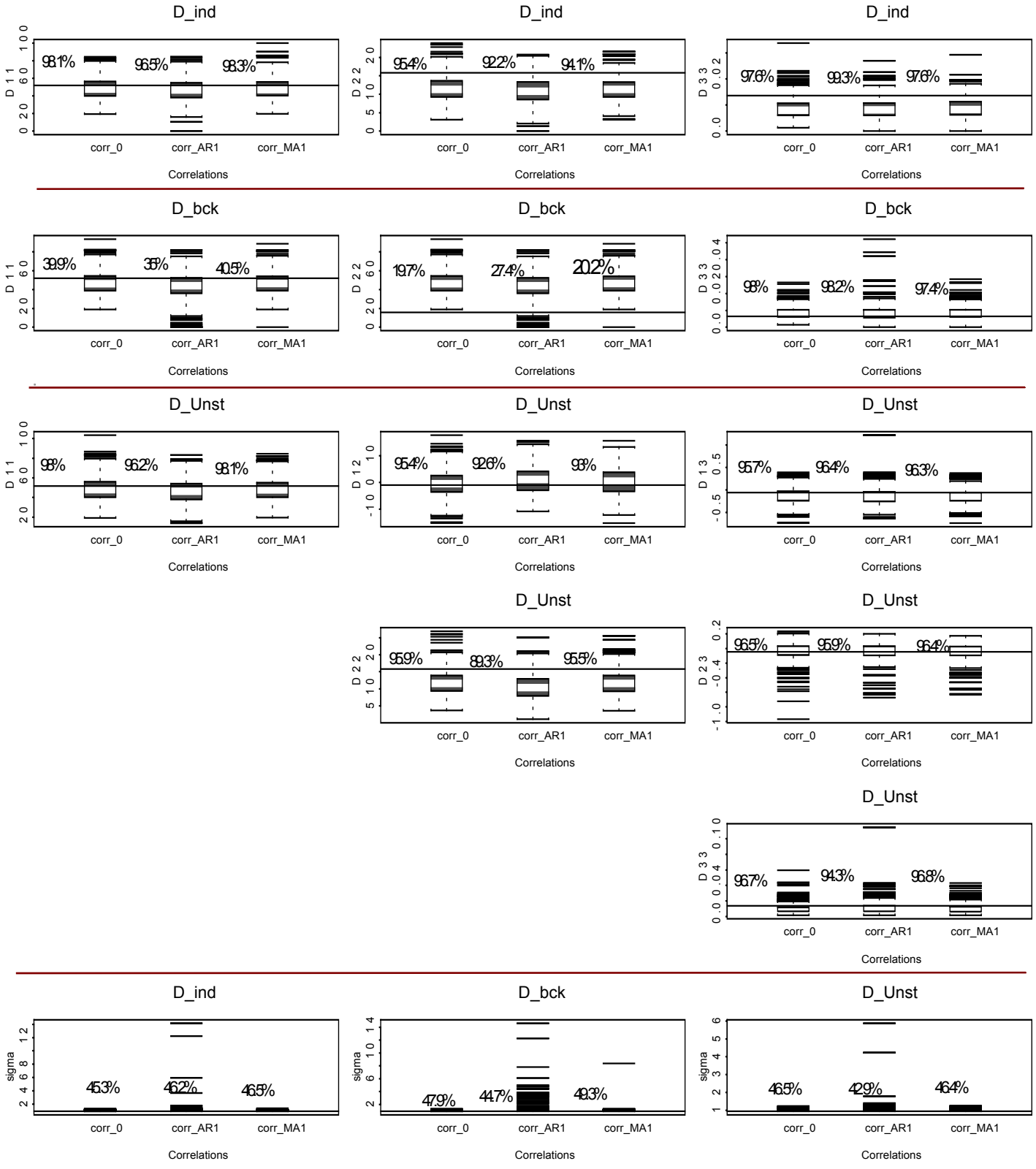


Figure 4.13: Breast cancer data. Box-plots of 1000 simulation replicates of random components estimates under different structures of D and correlation structures (first part of the simulation study)

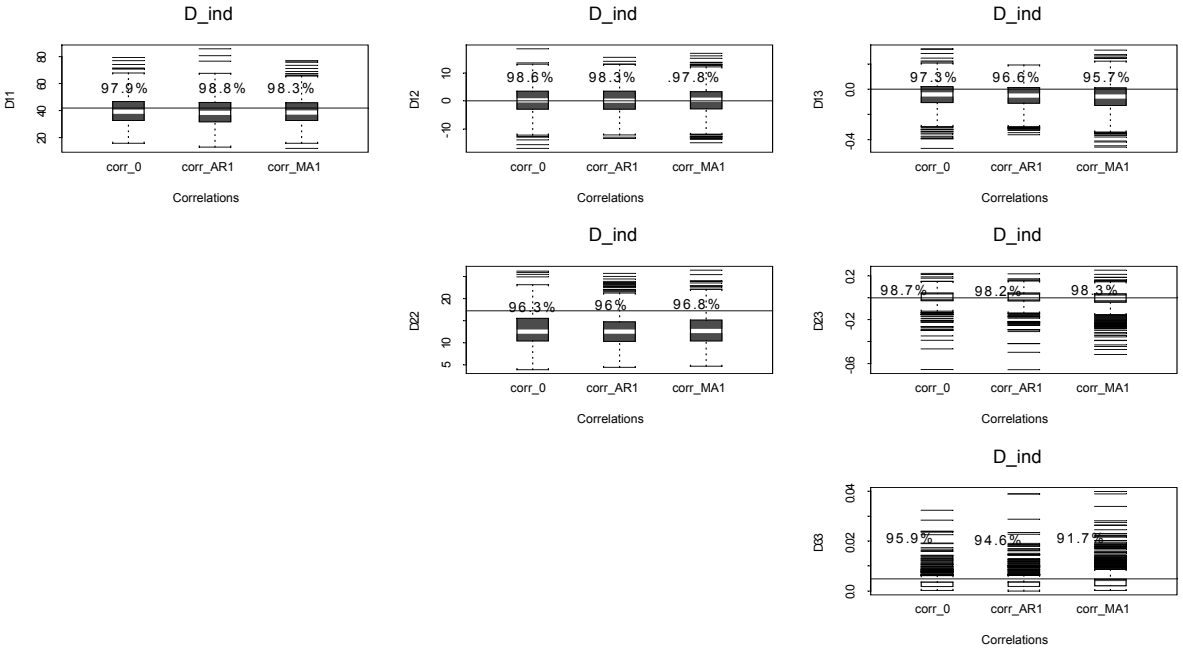


Figure 4.14: Breast cancer data. Box-plots of 1000 simulation replicates of random effects covariance estimates under D_ind and different correlation structures (simulation study 2" part)

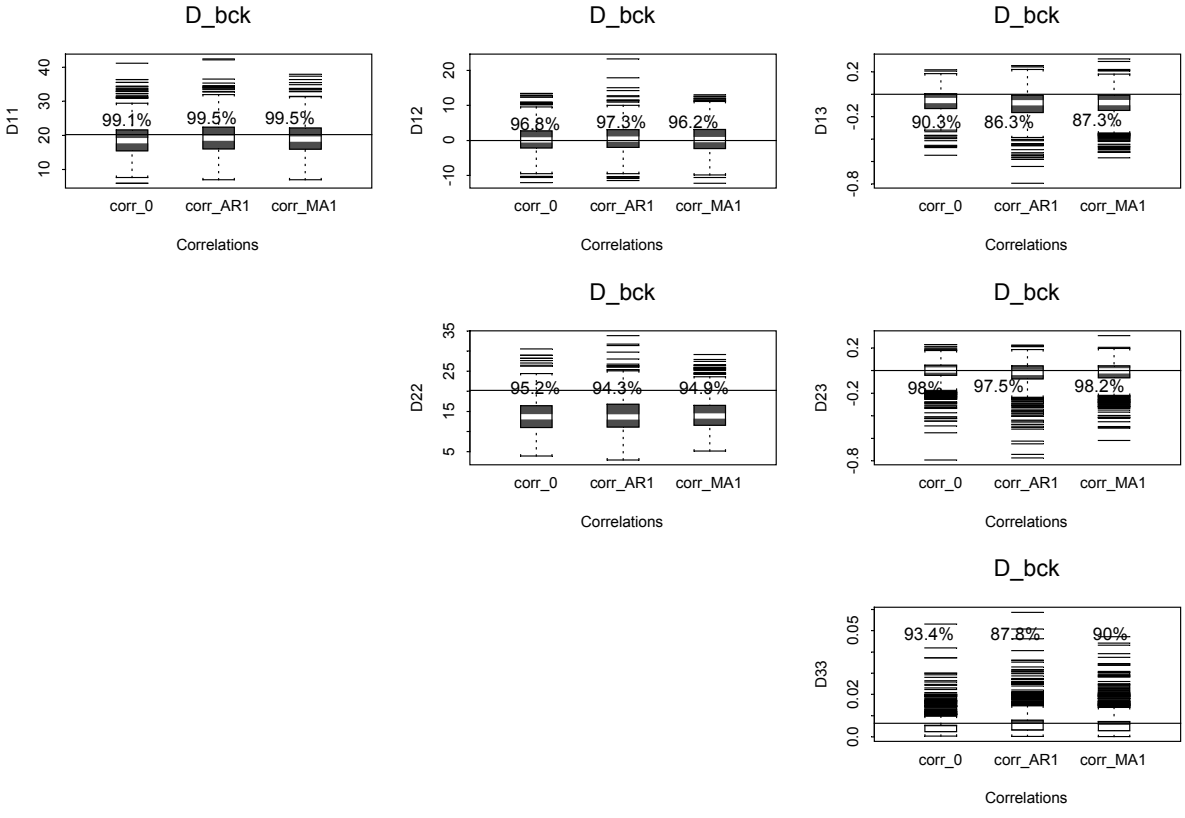


Figure 4.15: Breast cancer data. Box-plots of 1000 simulation results of random effects covariance estimates under D_bck and different structures of correlation (simulation study 2" part)

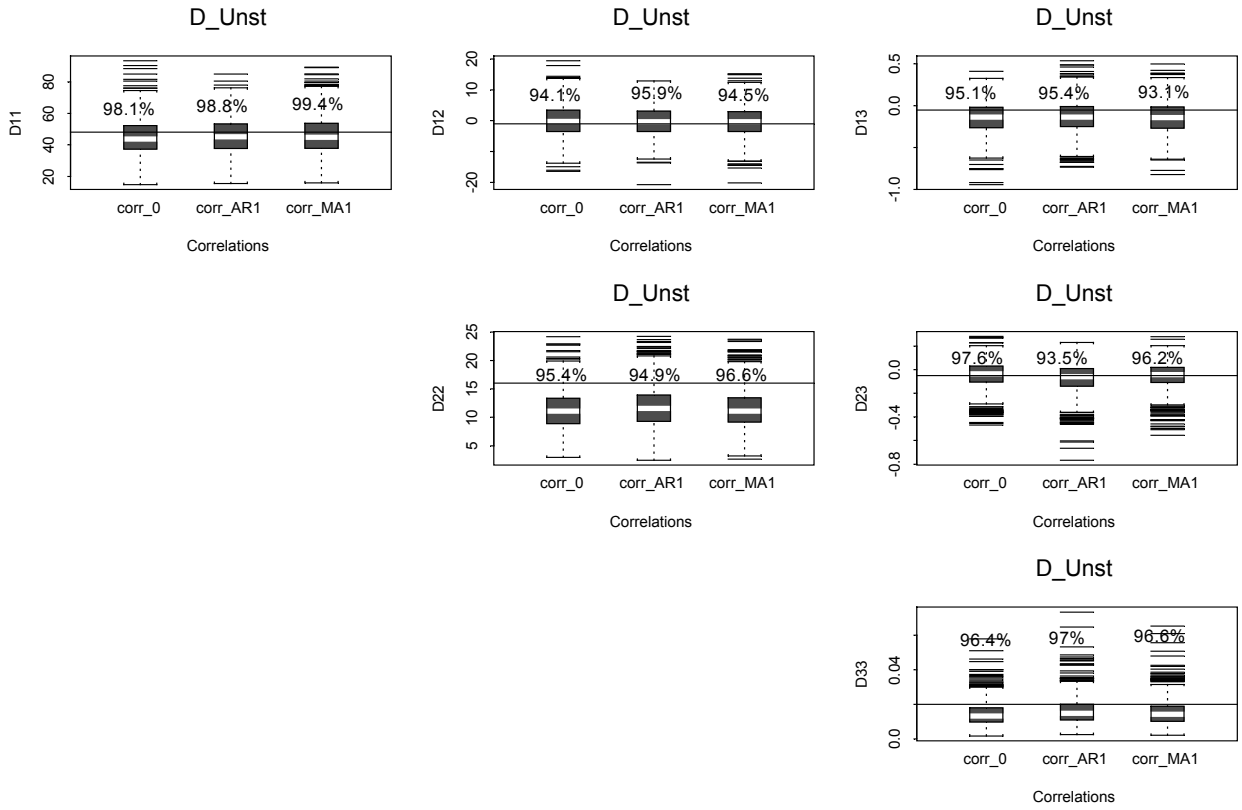


Figure 4.16: Breast cancer data. Box-plots of 1000 simulation replicates of random effects covariance estimates under D_unst and different structures of correlation (simulation study 2nd part)

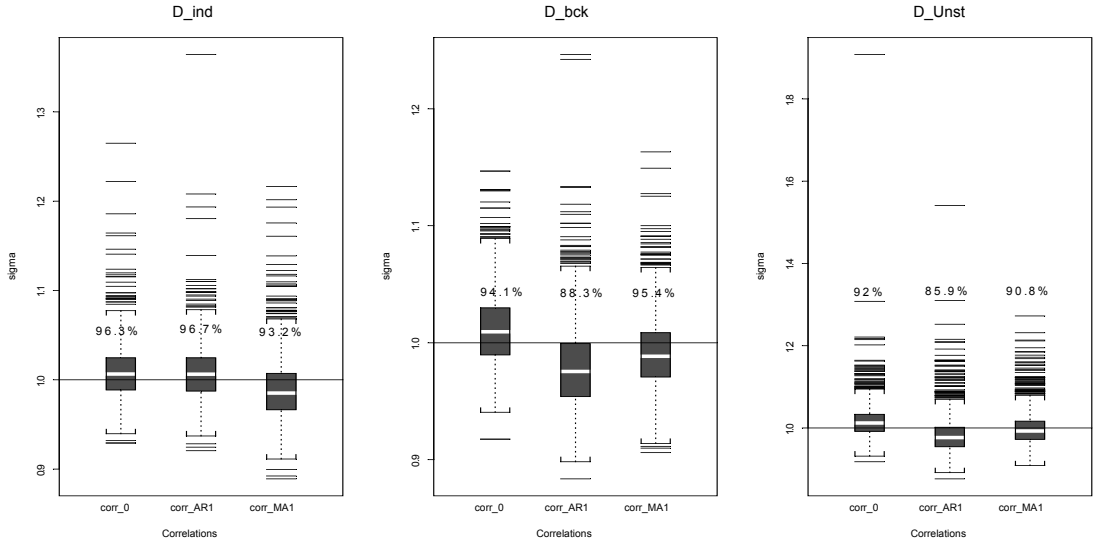


Figure 4.17: Breast cancer data. Box-plots of 1000 simulation estimates under different variance-covariance structures of random effects and different residual correlation structures (simulation study 2nd part)

4.2.2. Soybean genotypes model simulation

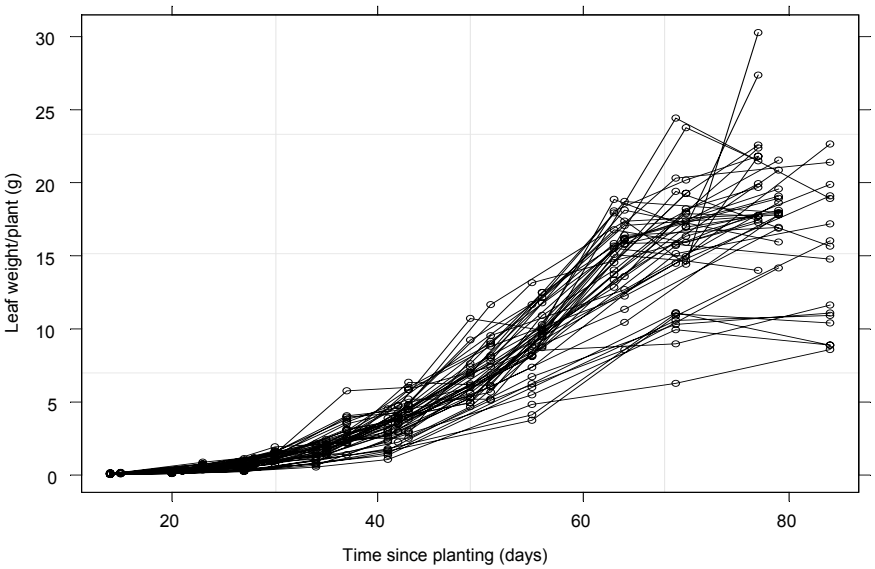


Figure 4.18: Growth curves in Soybean genotypes data

The model an data used in this simulation are reported in Davidian and Giltinan (1995, p.7) and reanalysed by Pinheiro and Bates (2000). The goal was to compare the growth patterns of two soybean genotypes, a commercial variety, Forrest (F) and an experimental strain, Plan Introduction #416937 (P). Data were collected during three years, from 1988 to 1990. At the beginning of the growing season in each year, 16 plots were planted with seeds; 8 plots with each genotype. Each plot was sampled eight to ten times at approximately weekly intervals. At each sampling time, six plants were randomly selected from each plot, leaves from this plant were weighted, and the average leaf weight per plant (in g) was calculated for each plot. Different plots in different sites were used in different years. The data are included in the S-Plus software. The objective was to model the growth pattern in terms of average leaf weight.

Examining carefully the results obtained by J.C. Pinheiro, based on normality assumption, we conclude that it is likely to think that the random effects follow a normal distribution but the errors present a clear departure from normality (see Figure 4.19). Due to this, we can pose the question on the validity of the inferences made during the modeling process.

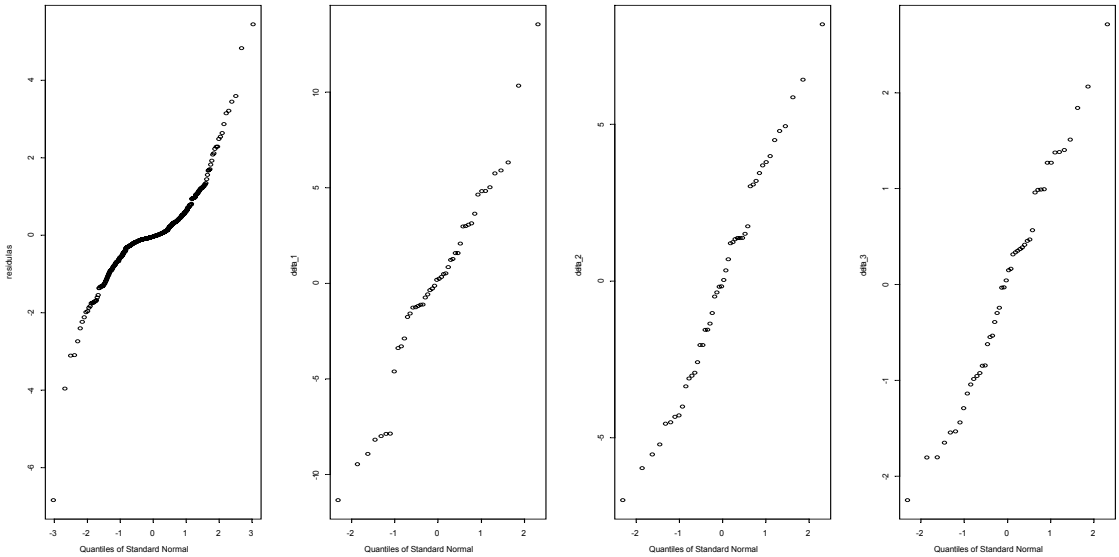


Figure 4.19: qq-norm of residuals and random effects for Soybean data

In the present section, we examine the effects of different distributions via simulation results, in the same way as used in the breast cancer case. But here we select five error distributions. Jointly with N, E and G (see section 4.2.1), we introduce a skew normal distribution with location parameter (a typical or central value that best describes the data, its effect is to translate the graph, relative to the standard normal distribution) $\zeta = -0.05$, scale parameter (its effect is to stretch out the graph) $\psi = 0.2$ and shape parameter (its effect is to allow the distribution to take a variety of shapes, depending on its value) $\varphi = 0.02$, as described in Azzalini (1986). This distribution will be designated as "sN". Finally, we also consider a semi-empirical distribution, a non-parametric gaussian kernel estimation from experimental data with the optimal bandwidth given by Sheather and Jones (1991) and Wand and Jones (1995), and implemented in S for the gaussian kernel by Venables (1997) (function width.SJ). This last distribution will be designated as "NPM".

We simulated observations from the soybean genotypes model

$$y_{ij} = \frac{\delta_{1i}}{1 + \exp\left[(\delta_{2i} - t_{ij})/\delta_{3i}\right]} + e_{ij} \quad (4.3)$$

where y_{ij} represents the average leaf weight/plant in subject i, $i = 1, \cdots, 48$, at time t_{ij}. The random effects $\eta_i = (\eta_{1i}, \eta_{2i}, \eta_{3i})'$ are $(0, D)$ and the e_{ij} are $(0, \sigma^2)$ and are independent of η_i. For this simulation we used the following "population" setting for the parameter values,

$$\delta = (\delta_1, \ \delta_2, \ \delta_3)' = (19.26, 55, 8.4)'$$

$$D = \begin{pmatrix} 25 & 2.50 & 4.00 \\ & 8.00 & 2.32 \\ & & 2.00 \end{pmatrix}; \qquad \sigma = 1.$$

The subject-specific parameters $\delta_{1i}, \delta_{2i}, \delta_{3i}$ were generated according to the model given by

$$\begin{cases} \delta_{1i} = \delta_1 + \eta_{1i} \\ \delta_{2i} = \delta_2 + \eta_{2i} \\ \delta_{3i} = \delta_3 + \eta_{3i} \end{cases} . \qquad (4.4)$$

where $\eta_i = (\eta_{1i}, \eta_{2i}, \eta_{3i})'$ follows a multivariate normal distribution with zero mean and covariance matrix D. On the other hand, an additional simulation was carried out to investigate the effect of covariance and/or correlation structure on the parametric inference when the LB-method is used. This simulation study was also divided in the same two parts (in the same way as in the breast cancer model described in section 4.2.1.5), the first one devoted the consequences of fitting under unrestricted conditions when data were generated according to some true structuring patterns, and the second one to the consequences of fitting inappropriate restricted models. More concretely, the simulated datasets were generated according to each one of the nine possible combinations of the covariance matrices for the random effects and correlation structures for the residuals described in the next section 4.2.2.3.

The results from these simulations are summarized in the following sections.

4.2.2.1. Results for the fixed effects parameters

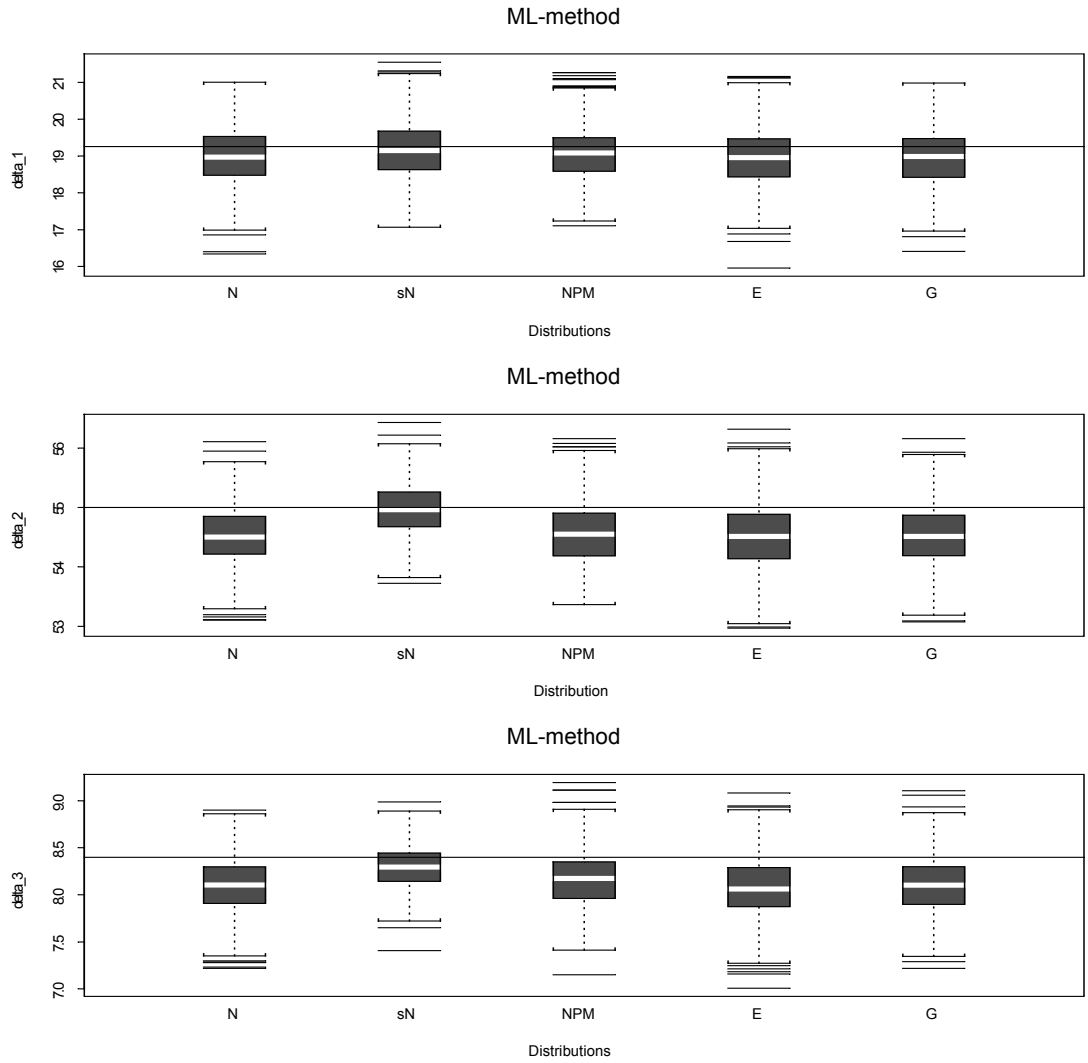


Figure 4.20: Box- plot of fixed effects of simulation results for model (4-3)-(4.4)

Under the non-linear Soybean genotypes mixed model, from Figure 4.20 and Table (4.23), there is virtually no difference between the five distributions assumed for the residuals, with respect to the average of δ_1 estimates. Its bias is significant and the MSE values tend to be similar, with values ranging between 0.54 and 0.67. The results are comparable to those in Table (4.1) when the random effects were generated according to a normal distribution. On the other hand and as in the breast cancer model simulation, the robustness of the confidence intervals depends on the concrete parameter of the model. In any case the observed coverages are near the nominal 95%.

For the specific case of the δ_1 parameter, coverages range from 94% under the sN residuals distribution to 91.5% under the exponential distribution. The NN case has 92.1% coverage. The last column of Table (4.23) shows that the confidence intervals have very similar mean lengths under all conditions. Additionally, the intervals are appreciably equitailed. We also note that, again, the behavior of the confidence intervals for this parameter is similar to those of parameters of the breast cancer model.

For the δ_2 parameter, the results shown in Table (4.23) demonstrate a negative bias indicating that, on average, they will underestimate the true parameter value. The bias and MSE statistics are nearly identical for this parameter, except for sN distribution in which these measures are the smallest. Under this model, confidence intervals for the δ_2 parameter perform poorly in all cases, except for the sN residuals distribution. The coverage ranges from 79.9% under the E distribution to 84.6% for the G case, with 93.5% and 83.2% values for the sN and N cases, respectively. In correspondence with the coverages, the widths of the intervals show a similar tendency, with the best (shortest) value obtained under the sN distribution. The intervals are always very asymmetrical, with inflated PROB. UPP. probabilities.

The worst results correspond to the δ_3 parameter, with appreciable negative bias and very poor coverage probabilities. Again the sN residuals distribution gives the best results, not only from the point of view of the of MSE, but also from the point of view of the coverage probability, a 90.7% and the best equitaildness. The coverage probabilities range from 78.8% (E) to 90.7% (sN) with 80% for the N case.

These simulations confirm that the LB-method is not very efficient even under normal conditions. But, in contrast with the previous simulations for the breast cancer model, a surprising result for this model is that, in terms of 95% nominal confidence coverages, the best results for fixed effects are obtained when the errors are skew normal (sN). On the other hand, this result is not maintained for random components (D and σ). Concretely, the results from

this phase of the simulation are displayed in Figure 4.21 to 4.22. Estimation under the nonlinear Soybean genotypes model tells a vastly different story. Here, all simulation conditions underestimate the random components of D, except under the sN distribution which provides reasonable unbiased estimates, but performs considerably worse in terms of coverage probabilities, especially for the $\sqrt{D_{33}}$ parameter (24.5% of coverage probability). In almost all cases the coverage probabilities are lower than their nominal value for $D_{12}/\sqrt{D_{11}}\sqrt{D_{22}}$, $D_{13}/\sqrt{D_{11}}\sqrt{D_{33}}$ and $\sqrt{D_{33}}$ and upper than their nominal value otherwise.

Figure 4.22 displays the results for the σ parameter. Its estimates are less precise for the NPM and specially for sN residuals, while the best results correspond to the E and G cases. The coverage probabilities range between 99% for N residuals to 0% for sN residuals.

Table (4.23): summary results for fixed effect parameters estimates for 1000 simulations

DISTRIBUTIONS	RANDOM EFFECTS	ERRORS	MEAN	Bias±C.I$_{Bias}$	MSE	COVERAGE- PROBABILITY (ASYMP. INTERVALS)			Average width
						PROB LOW	PROB.	PROB. UPP	
	Fixed effect parameter δ_1: Maximum likelihood (ML)								
	N	**N**	18.99208	-0.2679243±0.04776559	0.665096	1%	92.1%	6.9%	2.876276
	N	**sN**	19.15177	-0.1082278±0.04559323	0.552286	1.9%	94%	4.1%	2.798492
	N	NPM	19.04983	-0.2101692±0.04359019	0.5382893	1.2%	93.8%	5%	2.87612
	N	**E**	18.95573	-0.3042746±0.04281623	0.6537758	0.9%	91.5%	7.6%	2.897845
	N	**G**	18.96339	-0.2966082±0.04650596	0.6504091	0.5%	91.6%	7.9%	2.899373
	Fixed effect parameter δ_2: Maximum likelihood (ML)								
	N	**N**	54.52249	-0.4775082±0.03030467	0.4668351	0.1%	83.2%	16.7%	1.95098
	N	**sN**	54.96844	-0.03156185±0.02606712	0.1776974	2.7%	93.5%	3.8%	1.591223
	N	NPM	54.56497	-0.4350315±0.03086829	0.4370394	0.1%	83.7%	16.2.9%	1.927866
	N	**E**	54.50822	-0.491783±0.03020465	0.5211324	0.3%	79.9%	19.8%	1.974908
	N	**G**	54.53612	-0.4638765±0.0309442	0.4641885	0.2%	84.6%	15.2%	1.97972
	Fixed effect parameter δ_3: Maximum likelihood (ML)								
	N	**N**	8.100563	-0.2994368±0.01805612	0.1744442	0%	80%	20%	1.117007
	N	**sN**	8.290249	-0.1097514±0.0133505	0.05839522	0.6%	90.7%	8.7%	0.8053575
	N	NPM	8.166408	-0.2335917±0.0178202	0.1371466	0%	84.5%	15.5%	1.124151
	N	**E**	8.079644	-0.3203559±0.01745259	0.1958706	0.1%	78.8%	21.1%	1.140723
	N	**G**	8.100581	-0.2994189±0.01805035	0.1743792	0.1%	80.3%	19.6%	1.137076
	Soybean genotypes model: $y_{ij} = \frac{\delta_{1i}}{1+\exp\left[(\delta_{2i} - t_{ij})/\delta_{3i}\right]} + e_{ij}$								

4.2.2.2. Results for the covariance parameters

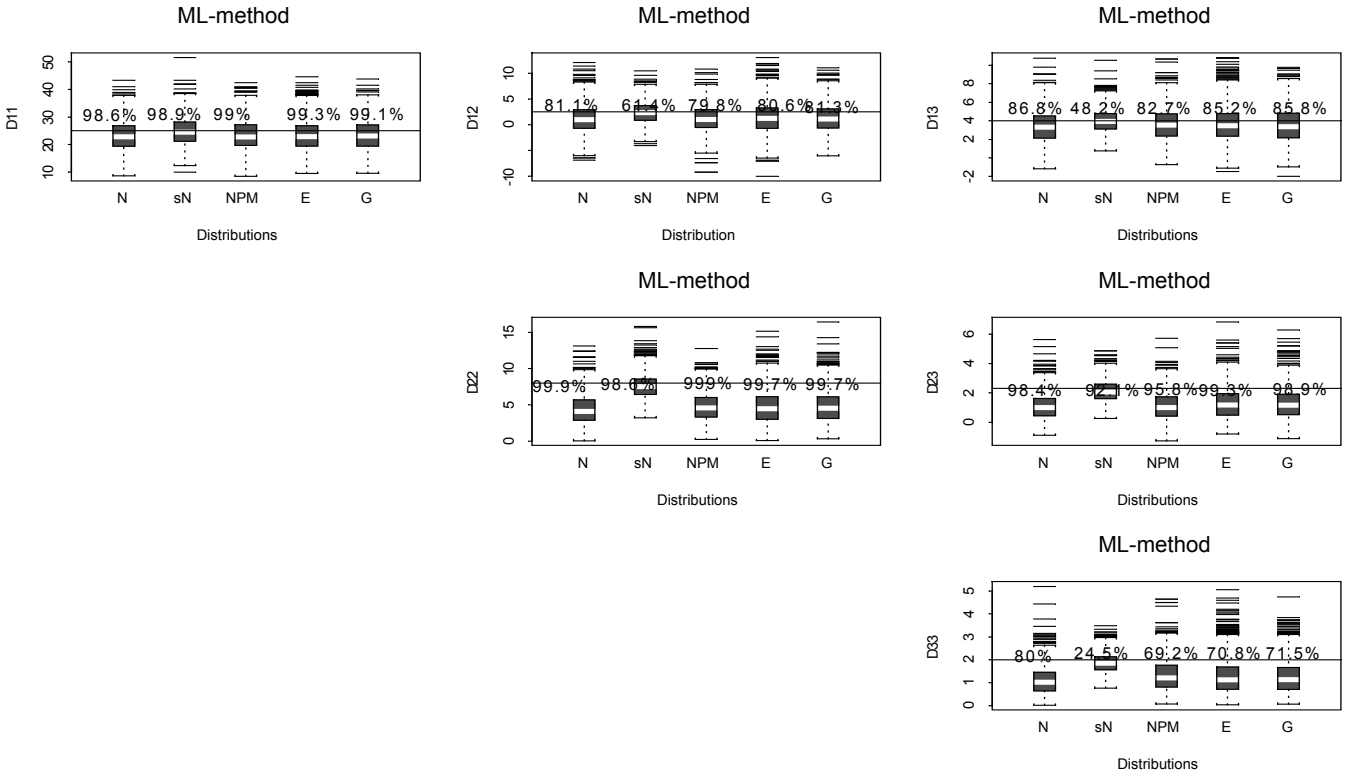


Figure 4.21: Box-plots for variance-covariance parameters of Soybean genotypes model

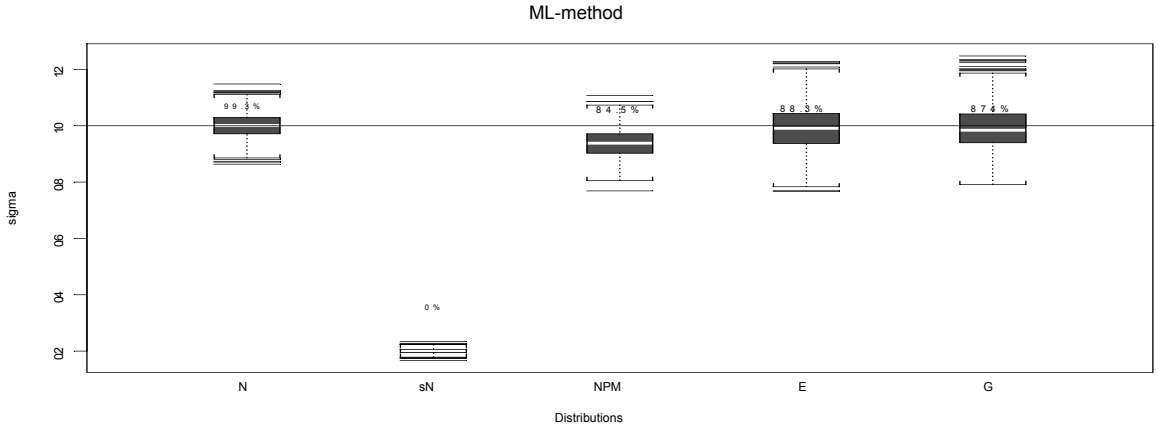


Figure 4.22: Box-plots for sigma parameter of genotype model

4.2.2.3. Robustness of LB- method in front of misspecifications of the structure of the covariance matrices and/or on residuals correlation

To close this chapter, and again under the Soybean data and logistic model (4.3)-(4.4) framework, we report a simulation study comparable to that described in section 4.2.1.5, that is, some simulations to:

a) test the consequences of fitting inappropriate restricted models, and

b) test the reverse situation: the consequences of fitting the model under unrestricted conditions when data were generated according to some true structuring patterns.

For the "a" case, the simulated datasets were generated using the same parameter values described at the beginning of section 4.2.2. For the "b" case, the simulated datasets were generated according to each one of the nine (3×3) possible combinations of the following covariance matrices for the random effects and correlation structures for the residuals:

i. Independent random effects:

$$D\_ind = \begin{pmatrix} 25 & 0 & 0 \\ 0 & 7.749999 & 0 \\ 0 & 0 & 0.8843354 \end{pmatrix}$$

ii. Block-diagonal: independent subgroups of random effects:

$$D\_bck = \begin{pmatrix} 9 & 0 & 0 \\ 0 & 9 & 0 \\ 0 & 0 & 9 \end{pmatrix}$$

iii. Unstructured covariance matrix:

$$D\_Unst = \begin{pmatrix} 31.5574859 & 0.2019679 & 0.1767219 \\ 0.2019679 & 20.1967910 & 0.2272130 \\ 0.1767219 & 0.2272130 & 5.0491977 \end{pmatrix}$$

and

i. Independent residuals: for each individual i, $cor(e_{ij}, e_{ij+1}) = 0, j = 1,\ldots,n_i$

ii. Auto-regressive process of order1, AR(1):

$$cor(e_{ij}, e_{ik}) = \phi^{|j-k|}, j,k = 1,\ldots,n_i; \text{ where } \phi = -0.6195506$$

iii. Moving average of order 1, MA (1):

$cor(e_{ij}, e_{ij+1}) = \alpha_1 / (1 + \alpha_1^2)$, where $e_{ij} = \alpha_1 e_{ij-1} + u_{ij}$; and α_1 = -0.9942504.

D_unst and independent residuals are included here to have a reference case, in fact these conditions have been studied in the preceding sections. For each one of the 18 simulation experimental conditions, 1000 data sets where generated according to the nonlinear mixed model (4.3)-(4.4).

1.2.1.2.3. Results for fixed effects

Table (4.24) tabulates the estimated values of the first series of simulations ("a") under the nonlinear mixed soybean genotypes model. The results depend on each specific parameter.

For the first parameter, δ_1, the differences between all simulation conditions with respect to mean, bias and mean square error (MSE) of the point estimates are small. The biases are negative and significantly different from zero, though small. The MSE values are also similar and small.

The main factor affecting the observed coverage of the confidence intervals is the (erroneously) assumed structure of the random effects covariance matrix. As may be expected, the best values are obtained for *D_Unst*, with a maximum of 93.4% for *D_unst* combined with independent residuals. The worst observed coverages correspond to *D_bck*, with the lowest coverage, a 82%, corresponding to *D_bck* and MA(1) residuals. The confidence intervals are always markedly asymmetrical, with an inflated PROB.UPP value (that is, there is a clear tendency for the intervals to be located entirely at right of the true parameter value). The mean width of the intervals is in accordance with the coverage, with the shortest intervals corresponding to the *D_bck* case.

For the δ_2 and δ_3 parameters, the results on the mean value, bias and MSE of the point estimates are very similar to those reported for δ_1, with small, though significant, negative bias in all conditions. On the other hand, and not in complete agreement with intuition, the results on confidence intervals are nearly completely reversed. The best coverages are obtained under the *D_bck* assumption, with a maximum for *D_bck* and AR(1) residuals (97.2% for δ_2 and 95% for δ_3). The lowest coverages correspond to *D_Unst*, with the minimum coverage associated to *D_unst* and independent residuals. The intervals are still markedly asymmetrical, with PROB.UPP inflated values.

The results of the second series ("b") of simulations are summarized in Table (4.25) and Figures 4.26 to 4.29. As a general rule, the point estimates of the fixed effects remain virtually unaffected by the fact of performing the analysis under the assumption of independent residuals when, in fact, there is some kind of residuals correlation structure, and are only slightly affected by the assumption of unstructured random effects covariance, when in fact random effects have a given structure. Table (4.25) confirms that the averages of the fixed effects estimates are close to the true parameters for all simulations. The true coverage of the confidence intervals based on the LB-method is not always adequate, even under a correct specification of the covariance structure of random effects and residuals. It depends on the specific model parameters, much more than in the true covariance structure of the random effects and even less in the correlation structure of the residuals.

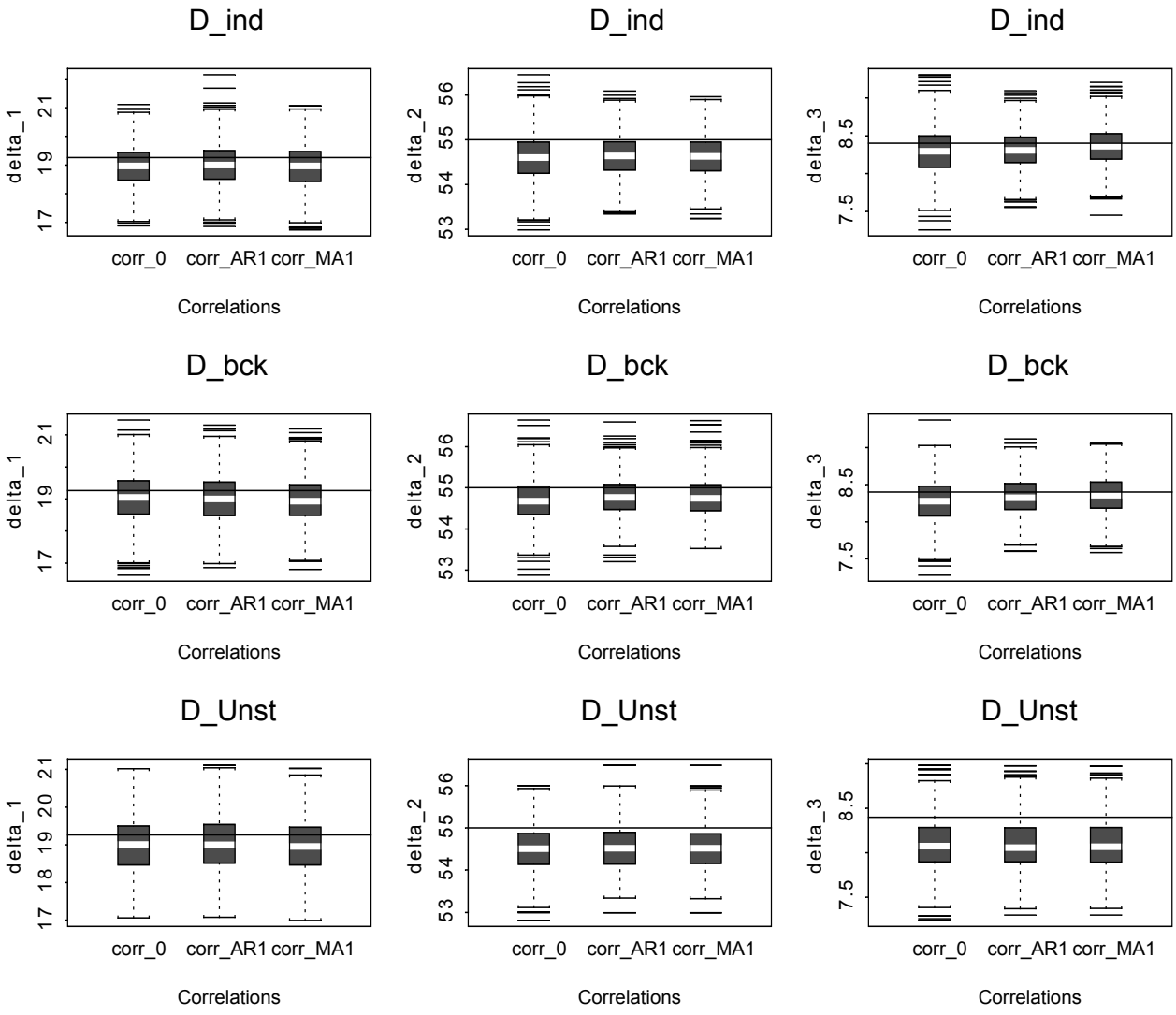


Figure 4.23: Box-plots for the results of 1000 simulations. Fixed effects estimates under different structures of the covariance matrix, D, and residuals correlation, for Soybean genotypes data ("a" simulations series)

Table (4.24): Simulation results for the *Soybean genotypes model* under different dependence structures for random effects and residuals (First case)

		δ_1						
Structure of D	*Working*	*Mean*	*Bias±C.I$_{Bias}$*	*MSE*	*PROB. LOW*	*COV. PROB*	*PROB. UPP*	*Average width*
D_ind	Independent	18.96579	-0.2942132±0.04566504	0.628838	1.4%	90.3%	8.3%	2.71004
	AR(1)	18.98863	-0.2713718±0.0459879	0.630425	1.8%	89.5%	7.3%	2.677696
	MA(1)	18.93715	-0.3228456±0.04659189	0.6687423	1.1%	89.3%	9.6%	2.69196
D_bck	Independent	19.04051	-0.2194925±0.04767421	0.6392217	2.6%	84.3%	13.1%	2.27268
	AR(1)	19.00413	-0.2558707±0.04589219	0.6131549	3.9%	82.3%	13.8%	2.201651
	MA(1)	18.96176	-0.2982434±0.04525658	0.6199689	3.4%	82%	14.5%	2.174162
D_Unst	Independent	18.9922	-0.2678027±0.04541789	0.6081411	0.9%	93.4%	5.7%	2.914135
	AR(1)	19.01046	-0.249541±0.04591368	0.6104688	1.8%	92.1%	61%	2.915253
	MA(1)	18.9591	-0.3009036±0.04629004	0.6477652	1.1%	91.5%	7.4%	2.919399
		δ_2						
D_ind	Independent	54.6048	-0.3951971±0.03267328	0.4337932	0.7%	82.4%	16.9%	1.810568
	AR(1)	54.59309	-0.4069145±0.02901113	0.4402066	0.5%	82.2%	18.9%	1.580207
	MA(1)	54.63169	-0.3683053±0.03034018	0.3750298	0.6%	81.3%	18.1%	1.631354
D_bck	Independent	54.69015	-0.3098519±0.03242608	0.3694356	0.2%	95.7%	4.1%	2.498229
	AR(1)	54.77096	-0.2290404±0.02927588	0.2753406	0.3%	97.2%	2.5%	2.332616
	MA(1)	54.7555	-0.244497±0.03025544	0.2971096	0.5%	96.1%	3.4%	2.386578
D_Unst	Independent	54.49707	-0.5029284±0.03486202	0.5689889	0.3%	79.5%	20.2%	1.951757
	AR(1)	54.5315	-0.4684997±0.03326956	0.5073296	0.7%	81.5%	17.8%	1.970419
	MA(1)	54.52909	-0.4709094±0.03280285	0.5015743	0.6%	82.3%	17.1%	1.969598
		δ_3						
D_ind	Independent	8.29996	-0.100040±0.01870675	0.10101	3.3%	87.2%	9.5%	0.9827692
	AR(1)	8.307381	-0.092619±0.01537698	0.106427	2.8%	85.3%	9.8%	0.8019704
	MA(1)	8.360465	-0.03953548±0.01578645	0.06637007	3.7%	89%	7.3%	0.8353948
D_bck	Independent	8.277938	-0.1220619±0.01854038	0.1042895	1.8%	87.3%	10.9%	0.9880616
	AR(1)	8.336482	-0.0635182±0.01532168	0.0650818	2.8%	89.5%	7.7%	0.816669
	MA(1)	8.350087	-0.04991292±0.01623579	0.07083428	3.9%	86.3%	9.8%	0.8251872
D_Unst	Independent	8.089291	-0.310709±0.01931481	0.193554	0.6%	77.6%	21.8%	1.131231
	AR(1)	8.090309	-0.3096906±0.01855465	0.1854363	0%	80.7%	19.3%	1.129665
	MA(1)	8.09188	-0.3081204±0.01824214	0.1814758	0%	81.5%	18.5%	1.13002

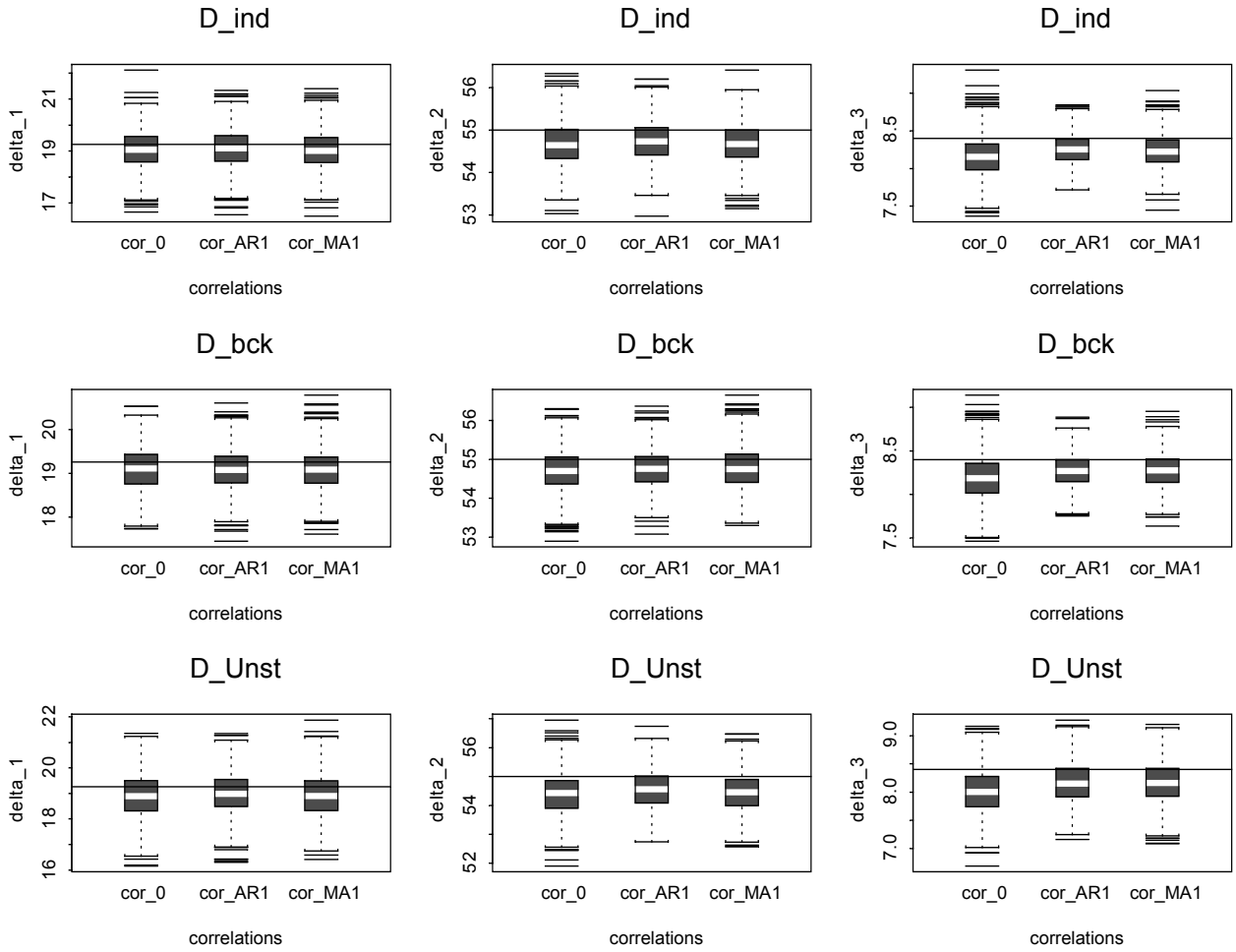


Figure 4.24: Box-plots for the results of 1000 simulations. Fixed effects estimates under different structures of the covariance matrix, D, and residuals correlation, for Soybean genotypes data ("b" simulations series)

Table (4.25): Simulation Results for *Soybean genotypes model* of different structure of covariance matrix and correlation (Second phase)

		δ_1						
Structure of D	*Working*	*Mean*	*Bias±C.I$_{Bias}$*	*MSE*	*PROB. LOW*	*COV. PROB*	*PROB. UPP*	*Average width*
D_ind	Independent	19.05448	-0.2055249±0.04558582	0.582638	1.1%	94%	4.9%	2.909964
	AR(1)	19.09658	-0.1634161±0.04548594	0.564198	1.7%	94%	4.3%	2.843851
	MA(1)	19.0362	-0.2238025±0.04533821	0.557341	1.2%	94.1%	4.7%	2.855609
D_bck	Independent	19.09566	-0.1643445±0.0300166	0.261311	1%	92.8%	6.2%	1.868798
	AR(1)	19.08624	-0.1737575±0.02823867	0.237560	1.3%	92.7%	6%	1.752299
	MA(1)	19.08849	-0.1715104±0.02834895	0.238407	1.9%	92.4%	5.7%	1.782298
D_Unst	Independent	18.91455	-0.3454477±0.05340997	0.861153	1.2%	91.4%	7.4%	3.247077
	AR(1)	19.00475	-0.2552507±0.05040407	0.725822	1.3%	93.6%	5.1%	3.202265
	MA(1)	18.92533	-0.3346746±0.05407979	0.7842382	0.9%	92.2%	6.9%	3.183286
				δ_2				
D_ind	Independent	54.67918	-0.3208238±0.03182381	0.3662928	1%	88.9%	10.1%	1.9651
	AR(1)	54.73897	-0.2610341±0.02950031	0.2942236	0.5%	88.4%	11.1%	1.74265
	MA(1)	54.68726	-0.3127378±0.03052397	0.3277257	0.2%	89%	10.7%	1.8092
D_bck	Independent	54.69736	-0.302641±0.03325201	0.3791255	0.9%	89%	10.1%	2.040034
	AR(1)	54.75664	-0.2433551±0.02997396	0.2928587	1%	90.7%	8.3%	1.830022
	MA(1)	54.7616	-0.2384041±0.03278227	0.3363042	1.6%	88.6%	9.8%	1.908499
D_Unst	Independent	54.396	-0.6039978±0.0450599	0.8928133	0.4%	84.1%	15.5%	2.78491
	AR(1)	54.57017	-0.4298255±0.04226079	0.6491889	0.4%	88%	11.6%	2.619695
	MA(1)	54.45269	-0.5473101±0.04403028	0.7451547	0.5%	86.8%	12.8%	2.654754
				δ_3				
D_ind	Independent	8.161801	-0.23820±0.01635497	0.126298	0.6%	80.9%	18.5%	0.9480518
	AR(1)	8.256304	-0.143697±0.01252	0.06137	0.9%	85%	14.1%	0.7293738
	MA(1)	8.232626	-0.167374±0.01353964	0.073253	0.7%	83.8%	15.5%	0.7712214
D_bck	Independent	8.190489	-0.209511±0.01543366	0.105838	0.7%	83.7%	15.6%	0.9447332
	AR(1)	8.270736	-0.129264±0.01142019	0.050625	0.2%	87.2%	12.6%	0.706781
	MA(1)	8.274229	-0.125771±0.012046	0.053553	0.5%	89%	10.5%	0.7573112
D_Unst	Independent	8.007475	-0.392525±0.02390434	0.302671	0.1%	79.5%	20.4%	1.46326
	AR(1)	8.170378	-0.229622±0.0220047	0.178643	0.3%	86.9%	12.8%	1.343985
	MA(1)	8.166625	-0.2333746±0.02343736	0.1807237	0.5%	89%	10.5%	1.357662

1.2.1.2.4. Results for the variance covariance parameters

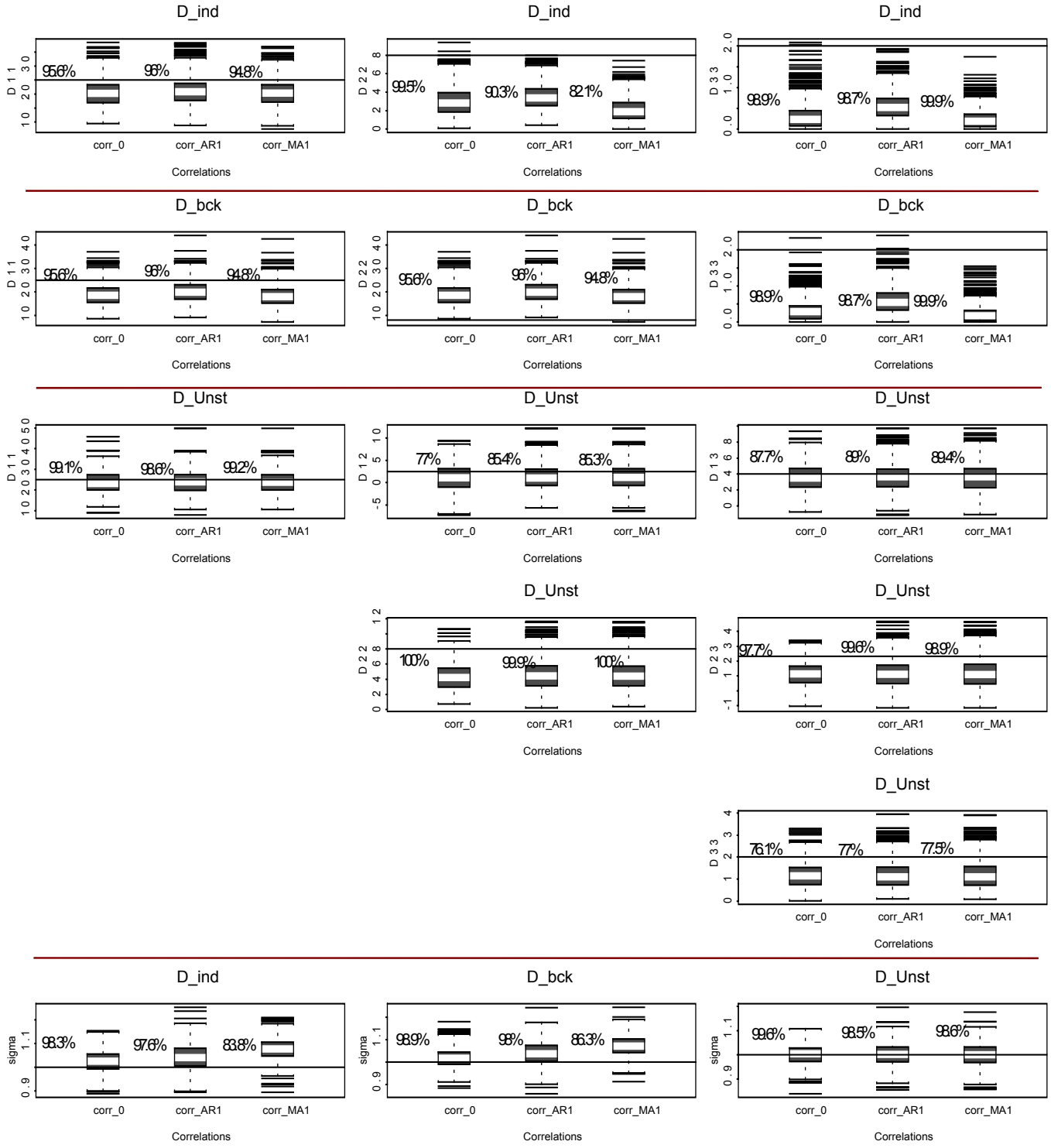


Figure 4.25: Box-plots for the results of 1000 simulations. Random components estimates under different structures of the covariance matrix, D, and residuals correlation, for Soybean genotypes data ("a" simulations series)

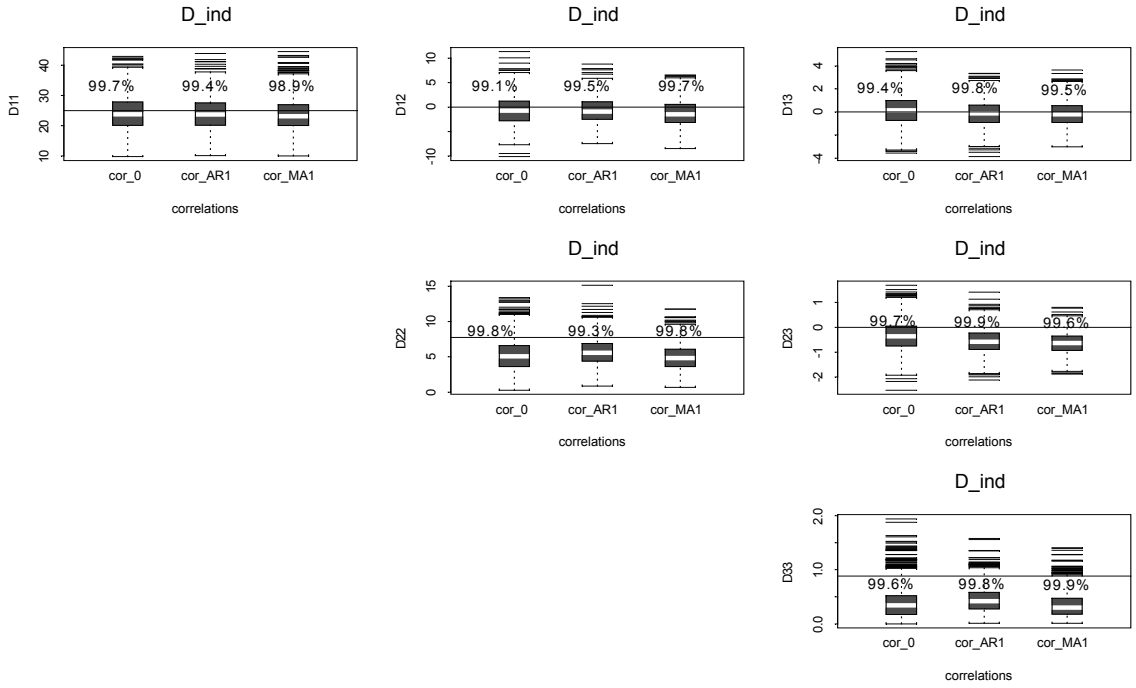


Figure 4.26: Box-plots of 1000 simulations results of random effects covariance estimates under independent random effects and different error correlation structures for Soybean genotypes data ("b" simulations series)

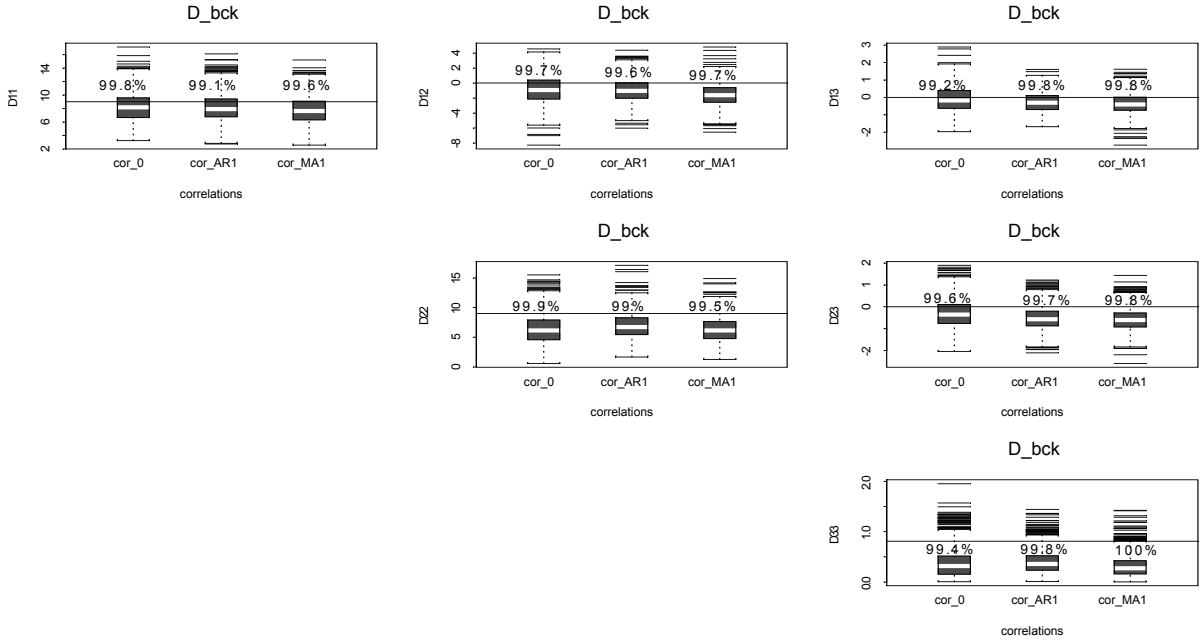


Figure 4.27: Box-plots of 1000 simulation results of random effects covariance estimates under Blocked structure of D and different error correlation structures for Soybean genotypes data ("b" simulations series)

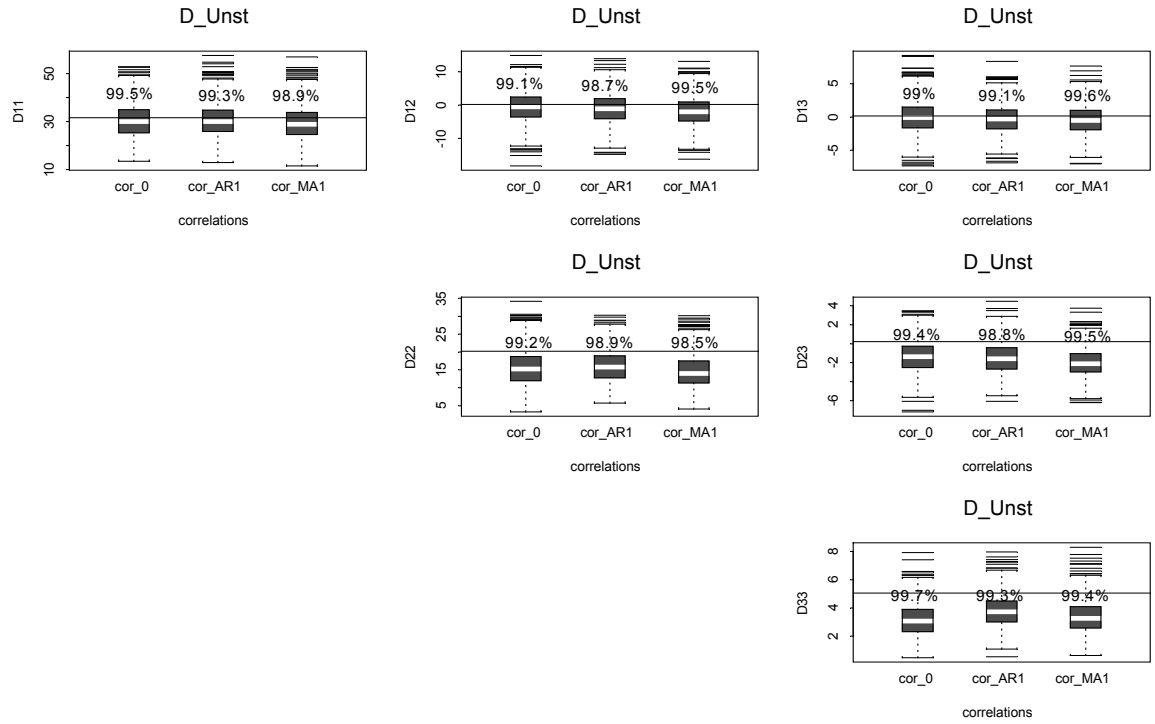


Figure 4.28: Box-plots of 1000 simulation results of random effects covariance estimates under Blocked structure of D and different error correlation structures for Soybean genotypes data ("b" simulations series)

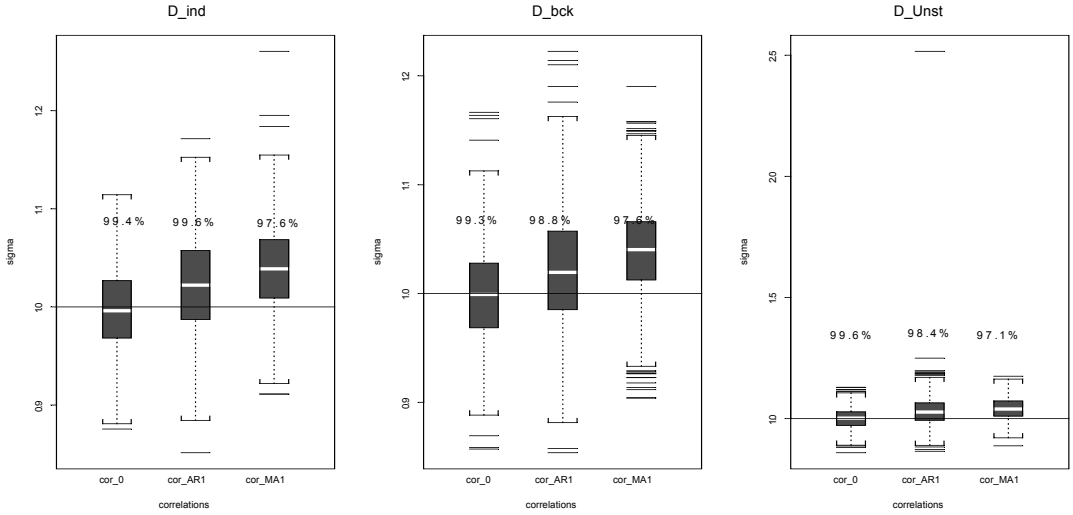


Figure 4.29: Box-plots of 1000 simulation σ estimates under different structures of D and different correlation's for soybean genotypes data ("b" simulations series)

4.3. Conclusions

Under the model (4.1)-(4.2) and according to the observed coverages for confidence intervals with 95% nominal coverage, the results of sections 4.2.1.1 to 4.2.1.4 indicate that the LB-method is not robust when the normality assumptions of random effects and/or errors are violated. In fact, both procedures (ML and REML) may not perform well even under normal conditions and a large number of observations per subject and a large number of subjects. Additionally, under this model, the results of section 4.2.1.5 indicate that the coverage of the confidence intervals for fixed effects is poor when the covariance matrix and correlation structure of errors are wrongly specified. These results lead to the conclusion that the performance of the LB-method not only depends on the true distributions of random effects and errors, but also depends on the structure of the covariance matrix and/or on the correlation structure of errors. Additionally, the robustness (and the adequacy, even under the correct methods specifications) seem to greatly depend on the specific parameter models.

The results of section 4.2.2 for the non-linear mixed model (4.3)-(4.4) suggest that the preceding conclusions may be general, and not associated to a specific model.

Chapter 5

5. A resampling inferential approach on mixed models

5.1. Introduction

Bootstrap methods applied to non-linear mixed models may give an alternative inferential approach to the parametric methods applied and analyzed in the previous chapters. We explore how to perform bootstrap resampling in a context of repeated measure data and non-linear mixed effects.

5.2. Bootstrap procedures under nonlinear mixed models

The statistical bootstrap method is a non-parametric alternative for finding error and bias estimates in situations where the assumption of a Gaussian distribution is not satisfied and where it is difficult or even impossible to develop an analytic solution. In this section we will show the reasoning behind the bootstrap method and how it is applied in the context of repeated measures and nonlinear mixed models.

Let us assume that we are interested in estimating a certain parameter δ and let us also assume that we have observations $Y_1,...,Y_n$ from a distribution F that depends on δ. Furthermore we have a method for finding an estimate $\hat{\delta}$ of δ, say $\hat{\delta} = T(Y_1,...,Y_n)$. The estimator T might be as simple as computing the mean of the observations or as complicated as fitting a breast cancer model.

Now, in addition to $\hat{\delta}$ we will also need an error estimate as well as an idea of the bias in the estimator T, among other possible measures of statistical precision. If T is fairly simple we might be able to find its distribution and get an error and a bias estimate analytically. If the situation is more complicated we might instead perform a Monte Carlo study, assuming that the distribution of the observations is completely specified. To do this, we would simulate sampling from the distribution F, generating many (say k) independent samples of size n, apply the estimator T to each sample and thereby get a sample of estimators $\hat{\delta}_1,\ldots,\hat{\delta}_k$. Then we can look at a histogram of the estimators, compute their standard deviation, and so on.

But what can we do if we do not know the distribution F? In that case the data $Y_1,\ldots,Y_n$ is all we have, and any analysis has to be based on these observations. The most simple estimate of the distribution function $F(y)$ is the empirical distribution function defined as $\hat{F}(y)=\frac{1}{n}\#\{Y_i \le y\}$, that is, the relative frequency of observations that are less or equal than y.

The basic idea of the bootstrap is to replace the distribution F in the Monte Carlo (MC) study above by its empirical distribution function $\hat{F}$.

It can be shown that sampling from the empirical distribution function means sampling *with replacement* from the observations $Y_1,\ldots,Y_n$. A bootstrap sample has the same sample size as the original data. It is made up out of the original observations, some of which might appear more than once whereas others might not be included at all. As in the MC study, we will draw many (say B) of these bootstrap samples, apply the estimator T to each one of them and thereby get bootstrap estimates $\hat{\delta}_1,\ldots,\hat{\delta}_B$ of δ. We can then study these bootstrap estimates to get an idea of the error and the bias of T, or in general of its sample distribution.

The bootstrap method as described above was first developed in Efron (1979) and is known to be a good general procedure for estimating the sampling distribution of an statistic under i.i.d conditions, see e.g. Bickel and Freedman (1981) and Singh(1981).

In practical situations, the i.i.d setup is often violated. The repeated measures design typically suggests some more structure for the errors. Since there are m subjects and n_i data are obtained

from the *i-th* subject, we have a bidimensional (possibly jagged) array of residuals or errors (or their estimates)

$$e_{ij}\,(or\,\hat{e}_{ij})\;1 \le j \le n_i;\quad 1 \le i \le m.$$

Thus, it is more reasonable to allow for dependency across the rows (i.e., errors for the same subject) although errors from different rows (corresponding to different subjects) may still be independent. Such complexity usually makes the problem of statistical analysis more difficult and requires a modification of the classical bootstrap. Das and Krishen (1999) discuss the possible ways to perform bootstrap resampling under a nonlinear mixed-effects (NLME) set-up, when the methods of estimation are based on linearization such as LB (Lindstrom and Bates 1990) and VC (Vonesh and Carter 1992) methods. In these cases, the parametric normal bootstrap approach (i.e. to estimate the parametric distribution and generate resamples according to this estimated distribution) is very natural, when applicable. On the other hand, according to Das and Krishen (1999), it is unclear how to perform non-parametric bootstrap.

In our opinion, the main guideline is that bootstrap resampling should reproduce the "true" sampling process that conducted to the "true" dataset. The basic underlying assumption concerning this sampling process is that it involves two steps, with different random components:

First, some experimental units are selected, possibly at random, from a population. This gives rise to a sample of m experimental units, $i = 1,\ldots, m$.

Second, an experiment is conducted over each one of these experimental units, giving rise to a set of random repeated measures for each experimental unit, over an adequate dimension like time or space, $\mathbf{y}_i = \left(y_{i1},\ldots,y_{in_i}\right)$, and to the full dataset, $\mathbf{y} = \left(y_1,\ldots,y_m\right)'$.

Bootstrap resampling should reproduce both steps. For example, omitting the second step is equivalent to implicitly assuming that the observed measures over each individual are fixed mean values, that they will be always the same.

How to resample depends on each specific problem and in the assumptions we make. There are two main categories of assumptions: those concerning the nonlinear model to fit to data and,

depending on the assumed nonlinear model, those concerning distributional aspects of the random factors and the residuals.

5.2.1. Bootstrapping when no assumptions on the nonlinear model are made

If no assumptions are made concerning the nonlinear model (and, thus, no assumptions are made concerning the distributions) the most obvious way to implement the first step is to sample at random, and with replacement, complete rows of the data matrix, that is, to construct a new data matrix $\mathbf{y}^*$ with m rows, where row i is defined as $\mathbf{y}_i^* = \mathbf{y}_{i^*}$ and i^* is a value taken with equiprobability from $\{1,2,\ldots,m\}$. In some situations, the preceding resampling process will be performed in a stratified way, not taking at random rows in the full data matrix (that is, indexes in {1, …, m}) but separately in s strata, taking m_k indexes in stratum k, with $m = \sum_{k=1}^{s} m_k$.

Under the same conditions, the second resampling step should be implemented, inside each experimental unit (that is, inside each $\mathbf{y}_i^*$ vector), according to an appropriate method for bootstrapping correlated data, like Nonparametric Bootstrap, Hansen (1999) or Markov Chain Bootstrap (MCB), Vasnev (2001). This will give rise to the final resampled dataset, $\mathbf{y}^{**} = \left(\mathbf{y}_1^{**},\ldots,\mathbf{y}_m^{**}\right)'$. MCB seems especially appropriate for our purposes. Assume that we are going to resample inside the *i-th* experimental unit, represented by the vector $\mathbf{y}_i^* = \left(y_{i1}^*,\ldots,y_{in}^*\right)$, (for simplicity in the notation, we assume a constant length, $n_i = n$) to generate the final resampling vector $\mathbf{y}_i^{**} = \left(y_{i1}^{**},\ldots,y_{in}^{**}\right)$. If $y_{i(1)}^*,\ldots,y_{i(n)}^*$ stand for the ordered values in $\mathbf{y}_i^*$, assume that the interval $\left[y_{i(1)}^*, y_{i(n)}^*\right]$ has been divided in r intervals, I_1, …, I_r –more details on the intervals are given below. More precisely, assume that we are in the *j-th* step of this process, that is, we are going to generate the observation y_{ij}^{**}, $2 \le j \le n$ given the last resampled value,

$y^{**}_{i,j-1}$. This resampling step is performed as follows: find the interval I_k with $y^{**}_{i,j-1}$ inside and take $y^{**}_{ij} = y^{*}_{ij^*}$ where j^* is chosen in $\{1,\ldots,n\}$ with probabilities $p_1,\ldots,p_n$ defined as:

$$p_t = \begin{cases} 0 & \text{if } y^{*}_{i,t-1} \notin I_k \\ \dfrac{1}{\#\left\{y^{*}_{il} \in I_k\right\}} & \text{if } y^{*}_{i,t-1} \in I_k \end{cases} \tag{5.1}$$

that is, if the last resampled value $y^{**}_{i,j-1}$ is in I_k, choose, with equiprobability, a value also following a point in this neighborhood.

The first resampled point, y^{**}_{i1}, deserves special attention. The approach taken in, Vasnev (2001) is to draw it at random from $\mathbf{y}^{*}_{i}$. This seems appropriate only under some stationarity assumptions frequently not met in a nonlinear mixed-models context. One possible approach is simply to make $y^{**}_{i1} = y^{*}_{i1}$, that is, always to start with the same initial points. Another possibility is to draw it at random in the interval I_k containing y^{*}_{i1}, (likely) avoiding a sharp difference between the observed and the resampled value.

Among the strategies to construct the intervals $I_1,\ldots, I_r$ discussed in, Vasnev (2001), we opt for the "uniform empirical grid", that is, to take intervals containing a fixed number of values, ν, (assuming that n is a multiple of ν, $r = n$): I_k=[a_{k-1}, a_k], $a_0 = y^{*}_{i(1)}$, $\#\left\{a_{k-1} \le y^{*}_{ij} \le a_k\right\} = \nu$, $k = 1,\ldots,r$. It avoids the possible degenerate case where the last point y^{*}_{in} is the unique element of the interval where it falls.

In MCB, all points with more recent history in the same interval, I_k, as the last resampled value, $y^{**}_{i,j-1}$, have the same probability to be chosen to continue the resample, while the remaining values have no chance. An alternative approach is to define the probabilities p_t as a decreasing function of the distance between the $y^{*}_{i,t-1}$ and $y^{**}_{i,j-1}$. A way to achieve this goal is by means of an appropriate kernel function, $p_t = K_{h\left(y^{**}_{i,j-1}\right)}\left(y^{**}_{i,j-1} - y^{*}_{i,t-1}\right)$, with calibration parameter $h\left(y^{**}_{i,j-1}\right)$. This is the Smoothing Markov Chain Bootstrap (SMCB) method, Vasnev (2001).

Note that residuals depend on the model and thus can't be used in this context. Note also that, provided that the analysis will be based on the mixed-models approach, possibly the statistics

computed over the original data and over the resampled data will assume a concrete nonlinear model, but in this approach resampling should be previous to model fitting, and independent of it.

5.2.2. Nonparametric bootstrap under a given functional model

If a given nonlinear model (like the breast cancer model shown in chapter 4º, equation (4.1)-(4.2)) is assumed, but no distributional assumptions are made, the first resampling step may still be performed as in the preceding section (sampling at random, with replacement, complete rows of the data matrix), as sampling the experimental units is previous and independent of the experimental process that will conduct to the repeated measures, concerned with the nonlinear fitting process.

Alternatively, the mixed-models fitting process may give estimates for the fixed and random effects. The random parameters will be (possibly) dependent inside each experimental unit but independent between experimental units. Then, the first step of resampling (generating experimental units) will be adequately performed by taking at random and with replacement complete random effect vectors, corresponding to a complete experimental unit, and adding them to the fixed parameters. These fixed parameters will be used to evaluate the nonlinear model to produce the expected observations for each concrete resampling experimental unit. More precisely, assume that the model is specified by the function *f*, with functional form common to all experimental units:

$$\begin{aligned} \mathbf{y}_i &= f(\mathbf{t}_i, \delta_i) + \mathbf{e}_i \\ \delta_i &= \mathbf{X}_i \delta + \mathbf{Z}_i \eta_i \end{aligned} \tag{5.2}$$

where $\mathbf{y}_i$ denotes the vector of observed responses for the *i-th* experimental unit taken at a set of conditions summarized by the vector of covariates $\mathbf{t}_i$, δ_i is a vector whose components are the parameters of the model function *f* characterizing the response in the *i-th* experimental unit and $\mathbf{e}_i$ is the corresponding vector of residuals. The model parameters are themselves modeled as

$\delta_i = \mathbf{X}_i\delta + \mathbf{Z}_i\eta_i$ where $\mathbf{X}_i$ and $\mathbf{Z}_i$ are design matrices for the fixed effects, δ, and the random effects associated to the *i-th* experimental unit, η_i, respectively.

If $\hat{\eta}_1,\ldots,\hat{\eta}_m$ stand for the estimates of the random effects in the m experimental units, $\hat{\delta}$ stands for the estimate of the fixed effects and $\hat{\mathbf{e}}_1,\ldots,\hat{\mathbf{e}}_m$ stand for the estimates of the residuals, the first resampling step is implemented as follows:

1. Generate $\hat{\eta}_i^*$, where $\hat{\eta}_i^* = \hat{\eta}_{i^*}$ and i^* is a value draw with equiprobability and with replacement from {1, ..., m},
2. Evaluate the parameters of the *i-th* resampled experimental unit: $\hat{\delta}_i^* = \mathbf{X}_{i^*}\hat{\delta} + \mathbf{Z}_{i^*}\hat{\eta}_i^*$.
3. Evaluate the resampling expected response for the *i-th* experimental unit:

   $\tilde{\mathbf{y}}_i^* = E_*\left(\hat{\mathbf{y}}_i^*\right) = f(\mathbf{t}_{i^*}, \hat{\delta}_i^*)$.
4. Repeat steps 1 to 3 to produce the expected responses $\tilde{\mathbf{y}}_1^*$, $\tilde{\mathbf{y}}_2^*$, ..., $\tilde{\mathbf{y}}_m^*$.

The algorithm is displayed in the preceding way to emphasize its applicability to the parametric bootstrap approach described below. Obviously, steps 1 to 4 may be performed more efficiently first evaluating once all the estimated predicted values, $\tilde{\mathbf{y}}_1,\ldots,\tilde{\mathbf{y}}_m$, (and the estimated residuals, $\hat{\mathbf{e}}_1,\ldots,\hat{\mathbf{e}}_m$, where $\hat{\mathbf{e}}_i = \mathbf{y}_i - \tilde{\mathbf{y}}_i$) with $\tilde{\mathbf{y}}_i = f(\mathbf{t}_i, \hat{\delta}_i)$ and $\hat{\delta}_i = \mathbf{X}_i\hat{\delta} + \mathbf{Z}_i\hat{\eta}_i$ for $i = 1, \ldots, m$, and then resampling the $\tilde{\mathbf{y}}_i$'s in the usual way, $\tilde{\mathbf{y}}_i^* = \tilde{\mathbf{y}}_{i^*}$. Note also that the design matrices $\mathbf{X}_i$ and $\mathbf{Z}_i$ will usually specify groups of experimental units defined by varying experimental conditions like different diets in the breast cancer experiments cited in the first section. In this case, resampling will be probably performed separately inside each one of these "cells", with fixed design matrices inside the *k-th* cell: $\mathbf{X}_k = \mathbf{X}_{i^*}$.

A possible drawback of this approach is that many methods to estimate the random effects depend on distributional assumptions on them, thus contradicting the first statement of this section.

In any case, the second step of resampling (repeated measures inside the experimental units) is associated, in a natural way, to sampling the residuals. It will consist in obtaining the resampled residual vectors $\mathbf{e}_1^*,\ldots,\mathbf{e}_m^*$ to be added to the expected responses vectors, $\tilde{\mathbf{y}}_1^*,\ldots,\tilde{\mathbf{y}}_m^*$, to finally generate the resampled dataset $\mathbf{y}_1^{**},\ldots,\mathbf{y}_m^{**}$, with $\mathbf{y}_i^{**} = \tilde{\mathbf{y}}_i^* + \mathbf{e}_i^*$. Depending on the assumptions

(or on the results of an exploratory analysis) made on the residuals, there may be some adequate resampling schemas.

If all the estimated residuals $\hat{\mathrm{e}}_{ij}$ are considered *iid*, an adequate resampling strategy will be to take all the estimated residuals at random and with replacement:

$$\text{for } 1 \leq i, i' \leq m,\ 1 \leq j \leq n_i^*,\ 1 \leq j' \leq n_{i'}$$

$$P_*\left\{\mathrm{e}_{ij}^* = \hat{\mathrm{e}}_{i'j'}\right\} = \frac{1}{\sum_{i'=1}^{m} n_{i'}}.$$

When the residuals are considered independent and identically distributed between experimental units but possibly dependent within experimental units, an adequate strategy will be to draw complete residual vectors, corresponding to full experimental units (implicitly assuming equal lengths n_i and equally spaced observation times): $P_*\left\{\mathbf{e}_i^* = \hat{\mathbf{e}}_{i'}\right\} = \frac{1}{m}$, for all, $1 \leq i, i' \leq m$.

Das and Krishen (1999) discuss some resampling schemas to cope with possible heterogeneities in the distribution of the residuals across experimental units, e.g., when a possible heterocedasticity between experimental units is assumed: $\mathrm{cov}\left(\mathbf{e}_i\right) = \Lambda_i$. Under this setting, an alternative approach is to generate the $\mathbf{e}_i^*$ according to a method like MCB, discussed in the preceding subsection. Remember that each vector of expected values, $\tilde{\mathbf{y}}_i$ is naturally associated to a vector of residuals, $\hat{\mathbf{e}}_i$. Thus, each $\tilde{\mathbf{y}}_i^* = \tilde{\mathbf{y}}_{i^*}$ is naturally associated to a residuals vector $\hat{\mathbf{e}}_{i^*}$ that may be resampled to finally produce $\mathbf{e}_i^*$, for any number of observations and observation times in the corresponding experimental unit, and (possibly) reproducing any particularities of its residuals distribution.

Sánchez and Ocaña (2001) analyse some of the preceding resampling approaches under a linear mixed models context.

5.2.3. Parametric bootstrap

In a fully parametric setting, assuming a given model to fit the data and the distributional form of the random effects and the residuals, both steps of resampling may be implemented using parametric bootstrap.

The procedure starts by estimating the random effects, $\hat{\eta}_1,\ldots,\hat{\eta}_m$, the fixed effects, $\hat{\delta}$, and the residuals, $\hat{\mathbf{e}}_1,\ldots,\hat{\mathbf{e}}_m$.. Now, an appropriate parametric estimation method may be fully applicable. Next, the parameters of the distribution of the random effects are estimated from the original data –for example, assuming a multivariate normal distribution, $\mathcal{N}(\mathbf{0},D)$, and estimating its covariance matrix D. In a similar way, the parameters of the residuals distribution will be estimated.

To produce the bootstrap expected responses, $\tilde{\mathbf{y}}_i^*$, the procedure in four steps, described in subsection 5.2.2, is applicable if we substitute its first step by the generation of the random effects, $\hat{\eta}_i^*$, according to its previously fitted parametric distribution. Similarly, the residuals $\mathbf{e}_i^*$ may be generated according to the assumed form of its distribution. In this process, again the possibilities range from independent and identically distributed residuals for all experimental units and observation times, to dependent residuals inside each experimental unit. The most complex bootstrapping strategy for the residuals is to make its distribution dependent on the generated random effects, that is, non-homogeneous across concrete experimental units. Sanchez and Ocaña (2001, 2002) show that this last approach is the best alternative for the linear mixed models case, provided that all distributions are correctly specified.

5.2.4. Bootstrap bias correction in two-stage methods

In the preceding subsections, we have presented a general framework for bootstrapping repeated measures data in a nonlinear mixed-models context. More restricted bootstrapping strategies may also be helpful to improve specific fitting procedures. An obvious application concerns two-stage methods, introduced in chapter 2. Remember that under this approach, we first fit the chosen model like (5.2) to the original sample of observations by a suitable method, like

generalized least squares (GLS) to get individual estimates $\tilde{\delta}_i$ for the parameters of the nonlinear model (one fit for each individual). Next, we estimate the residual values and $\hat{\sigma}^2$. We use these individual estimates as "data" to make inferences on fixed effects, δ, and on random inter-individual variation, D. In the simplest method, *STS*, the individual estimates are identified with the true parameters values, δ_i. Then, according to the model:

$$\tilde{\delta}_i \cong \mathrm{X}_i\delta + \eta_i \tag{5.3}$$

where X_i is a design matrix specifying the "treatment" conditions, very simple estimates for fixed effects:

$$\hat{\delta}_{STS} = (\sum_{i=1}^{m} \mathrm{X}_i'\mathrm{X}_i)^{-1}(\sum_{i=1}^{m} \mathrm{X}_i'\tilde{\delta}_i) \tag{5.4}$$

and for the covariance matrix of the random effects:

$$\hat{D}_{STS} = (m-1)^{-1}\sum_{i=1}^{m}(\tilde{\delta}_i - \mathrm{X}_i\hat{\delta}_{STS})(\tilde{\delta}_i - \mathrm{X}_i\hat{\delta}_{STS})'. \tag{5.5}$$

are derived. The form of these estimators is independent of any distributional assumption, but they are markedly biased.

Now, assume (mainly for notational simplicity) that the parameters in each experimental unit are modeled as $\delta_i = \delta + \eta_i$ -that is, that the design matrix $\mathbf{X}_i$ is the identity. If, in the first stage of a two-stage method, the parameters δ_i were truly observable, it will be true that $E\{\overline{\delta}\} = \delta$ and $\mathrm{cov}(\delta_i) = D$, with $\overline{\delta} = m^{-1}\sum\delta_i$. In other words, the STS estimators of the fixed effects and the covariance matrix D of the random effects will be unbiased. In fact, instead of the true δ_i values we have their estimates, $\hat{\delta}_i$, with $\hat{\delta}_i = \delta_i + \varepsilon_i$. Depending on the estimation procedure used in the first stage, these estimates may be biased, that is, we will have $E(\varepsilon_i) \neq 0$. Even if the first stage estimators are unbiased, the estimators of the covariance matrix may be biased, see below.

The distribution of the random vector δ_i depends on the sampling process over (*between*) the experimental units (concerning the random effects) while the distribution of the random vector ε_i depends on the sampling process *within* each fixed experimental unit, given δ_i. Designate as P and U the true probabilistic model associated to the sampling process over the population of

experimental units and inside each experimental unit, respectively. Then, if $\overline{\hat{\delta}} = m^{-1}\sum \hat{\delta}_i$, we have:

$$\begin{aligned} E_{P,U}\left(\overline{\hat{\delta}}\right) &= \frac{1}{m}\sum_{i=1}^{m} E_{P,U}\left(\delta_i + \varepsilon_i\right) \\ &= \frac{1}{m}\sum_{i=1}^{m} E_P\left(\delta_i\right) + \frac{1}{m}\sum_{i=1}^{m} E_U\left(\varepsilon_i \middle| \delta_i\right) \\ &= \delta + \frac{1}{m}\sum_{i=1}^{m} E_U\left(\varepsilon_i \middle| \delta_i\right) \end{aligned}$$

and

$$\begin{aligned} & E_{P,U}\left\{\frac{1}{m-1}\sum_{i=1}^{m}\left(\hat{\delta}_i - \overline{\hat{\delta}}\right)\left(\hat{\delta}_i - \overline{\hat{\delta}}\right)'\right\} \\ &= \frac{1}{m-1} E_{P,U}\left\{\sum_{i=1}^{m}\left(\delta_i + \varepsilon_i - m^{-1}\sum \hat{\delta}_i\right)\left(\delta_i + \varepsilon_i - m^{-1}\sum \hat{\delta}_i\right)'\right\} \\ &= \frac{1}{m-1} E_{P,U}\left\{\sum_{i=1}^{m}\left(\left(\delta_i - \overline{\delta}\right) + \left(\varepsilon_i - \overline{\varepsilon}\right)\right)\left(\left(\delta_i - \overline{\delta}\right) + \left(\varepsilon_i - \overline{\varepsilon}\right)\right)'\right\} \\ &= \frac{1}{m-1}\sum_{i=1}^{m} E_P\left(\left(\delta_i - \overline{\delta}\right)\left(\delta_i - \overline{\delta}\right)'\right) \\ &\qquad + \frac{1}{m-1}\sum_{i=1}^{m} E_U\left(\left(\varepsilon_i - \overline{\varepsilon}\right)\left(\varepsilon_i - \overline{\varepsilon}\right)\middle|\delta_i\right) \\ &= D + \frac{1}{m-1}\sum_{i=1}^{m} E_U\left(\left(\varepsilon_i - \overline{\varepsilon}\right)\left(\varepsilon_i - \overline{\varepsilon}\right)'\middle|\delta_i\right). \end{aligned}$$

The preceding discussion reflects commonly known facts, see for example Davidian and Giltinan, (1995). Its purpose is to emphasize that the *STS* estimates are in general biased, and their bias depends on the within experimental units sampling process. This will be the main guideline for bootstrap bias correction.

In the "bootstrap world" the role of the "true" parameters inside each experimental unit, $_i$, is assumed by their estimates, $\hat{\delta}_i$, and the role of these estimates is assumed by the same estimates computed over a resample $\hat{\delta}^*_{ib}$, for $b = 1,\ldots,B$. Here we are resampling B times over each observed experimental unit, that is, over each observed data vector $\mathbf{y}_i$, as we are emulating the second step of the sampling process, to correct the bias associated to the "E_U" expectations. Any of the bootstrapping methods discussed in the preceding subsections may be appropriate, depending on the concrete assumptions. The difference $\varepsilon_i = \hat{\delta}_i - \delta_i$ will be likely approximated by the values $\varepsilon^*_{ib} = \hat{\delta}^*_{ib} - \hat{\delta}_i$, for $b = 1,\ldots,B$. Thus, in order to correct the bias, we can use the approximations:

$$\frac{1}{m}\sum_{i=1}^{m} E_U\left(\varepsilon_i \middle| \delta_i\right) \approx \frac{1}{m}\sum_{i=1}^{m} E_*\left(\varepsilon_i^* \middle| \hat{\delta}_i\right)$$

where

$$E_*\left(\varepsilon_i^* \middle| \hat{\delta}_i\right) \approx \overline{\varepsilon}_i^* = \frac{1}{B}\sum_{b=1}^{B} \varepsilon_{ib}^*$$

and

$$\frac{1}{m-1}\sum_{i=1}^{m} E_U\left(\left(\varepsilon_i - \overline{\varepsilon}\right)\left(\varepsilon_i - \overline{\varepsilon}\right)' \middle| \delta_i\right) \approx \frac{1}{m-1}\sum_{i=1}^{m} E_*\left(\left(\varepsilon_i^* - \overline{\varepsilon}_i^*\right)\left(\varepsilon_i^* - \overline{\varepsilon}_i^*\right)' \middle| \hat{\delta}_i\right)$$

where

$$E_*\left(\left(\varepsilon_i^* - \overline{\varepsilon}_i^*\right)\left(\varepsilon_i^* - \overline{\varepsilon}_i^*\right)' \middle| \hat{\delta}_i\right) \approx \frac{1}{B-1}\sum_{b=1}^{B}\left(\varepsilon_{ib}^* - \overline{\varepsilon}_i^*\right)\left(\varepsilon_{ib}^* - \overline{\varepsilon}_i^*\right)'.$$

The preceding bootstrap procedure is just a bias correction device for the STS method.

To perform a bootstrap bias correction to the standard two-stage covariance estimator we have to evaluate

$$\hat{D}_{STS,BMC} = \hat{D}_{STS} - \widehat{bias}_{BMC} \qquad (5.6)$$

where $\widehat{bias}_{BMC}$ is an adequate bootstrap estimate of bias. As a consequence of the previous discussion, an adequate choice may be $\widehat{bias}_{BMC} = \frac{1}{m-1}\sum_{i=1}^{m}\left(\frac{1}{B-1}\sum_{b=1}^{B}\left(\varepsilon_{ib}^* - \overline{\varepsilon}_i^*\right)\left(\varepsilon_{ib}^* - \overline{\varepsilon}_i^*\right)'\right)$.

The preceding procedure is based on an adequate rationale, but is fairly computer intensive. An alternative approach in a two-stage context is as follows:

As the true bias is $bias(\hat{D}_{STS}, D) = E(\hat{D}_{STS}) - D$, an obvious and more direct estimate of it is $\widehat{bias}_B = E^*(\hat{D}_{STS}^*) - \hat{D}_{STS}$ where $E^*(\hat{D}_{STS}^*) \cong B^{-1}\sum \hat{D}_{STS}^{*,b}$ is the bootstrap estimate of the expectation, obtained from B resamples, where each resample is a random sample of size m, $\tilde{\delta}_1^*,\ldots,\tilde{\delta}_m^*$, taken with replacement, from the initial set of individual estimations, $\tilde{\delta}_1,\ldots,\tilde{\delta}_m$. Then, the bootstrap bias-corrected estimate of D becomes:

$$\hat{D}_{STS,BC} = \hat{D}_{STS} - \widehat{bias}_B \quad (5.7)$$

or, in other words:

$$\hat{D}_{STS,BC} = \hat{D}_{STS} - E^*(\hat{D}^*_{STS}) + \hat{D}_{STS}$$
$$= 2\hat{D}_{STS} - E^*(\hat{D}^*_{STS})$$

where $E^*(\hat{D}^*_{STS}) \cong B^{-1}\sum \hat{D}^{*,b}_{STS}$.

In this methodology, the bootstrap samples of $\hat{\delta}$ and $\hat{D}$ are obtained through the following iterative process:

Fit the model (5.1)-(5.2) to the original sample of observations to get $\tilde{\delta}_i$, $1,\ldots,m$

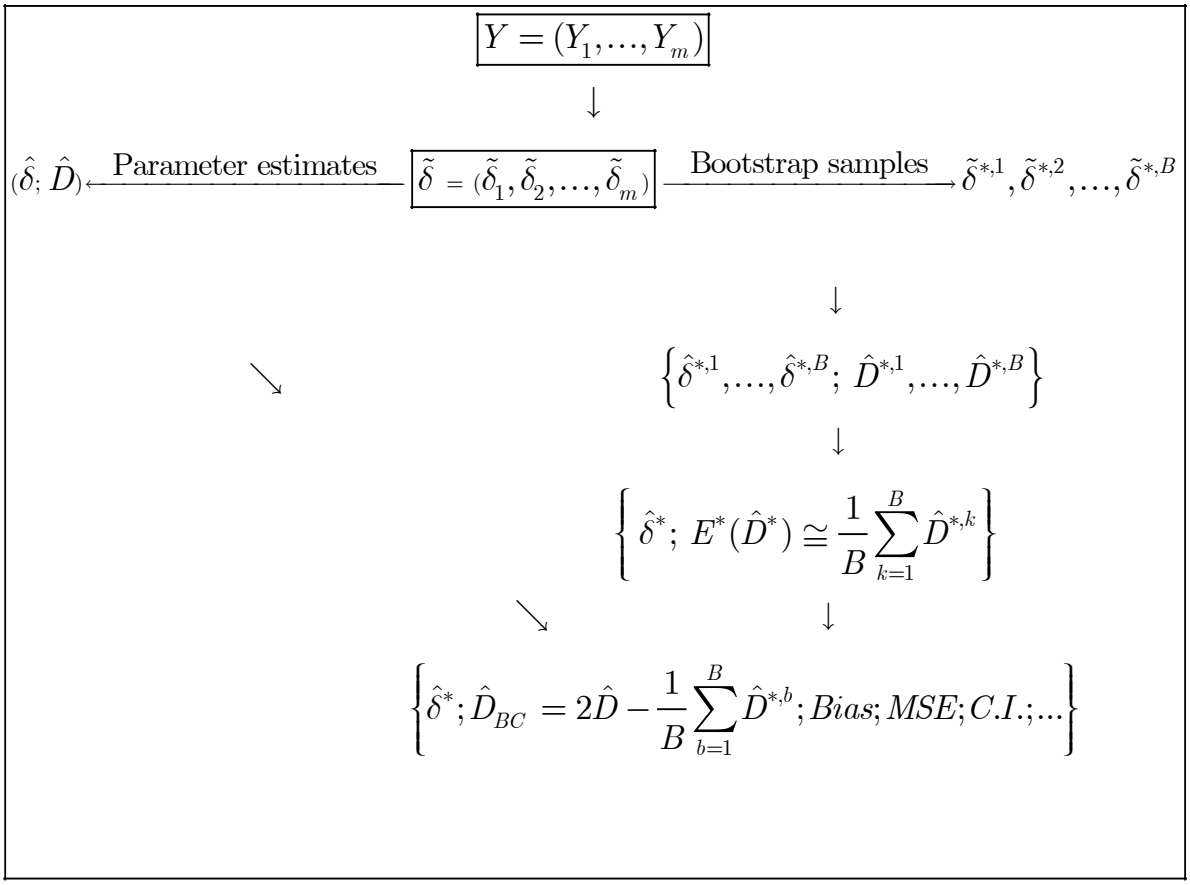
Generate a bootstrap sample of "observations" $\tilde{\delta}^* = (\tilde{\delta}^*_1,\ldots,\tilde{\delta}^*_m)$ from the first stage sample $(\tilde{\delta}_1,\ldots,\tilde{\delta}_m)$.

For each bootstrap sample we compute the $\hat{\delta}^*, \hat{D}^*$ by the standard technique applied in the equation (5.6)

The above steps (2)-(3) generate B values of $\tilde{\delta}^*$ and $\tilde{D}^*$, denoted as $\tilde{\delta}^*_1,\ldots,\tilde{\delta}^*_B$ and $\tilde{D}^*_1,\ldots,\tilde{D}^*_B$, respectively.

Calculate the mean, bias, MSE, C.I,...

The following diagram summarizes all these steps:



Despite its apparent correctness and simplicity –it apparently follows the most strict "bootstrap orthodoxy"– this "naïve" method performs very poorly in correcting the bias of the *STS* covariance estimates, as is illustrated by the simulation results. This is because it is exclusively based on resampling at a population level: taking the first step estimates as resampling units completely ignores the variability at the within-experimental-units level. On the other hand, it performs better for other purposes, like to construct confidence intervals for the fixed parameters.

Up to this point, we have presented some possible alternatives for bootstrapping in a nonlinear mixed models context. Some of them (in principle) reproduce accurately the true sampling process but are very computer intensive, while fast alternatives based on resampling only at a sample-population level (ignoring the within experimental units variability) will presumably perform inadequately. A compromise solution may be provided by the global two-stage method

(GTS) that incorporates uncertainty in the estimation of the individual parameters. Conditionally on the $\delta_1,\ldots,\delta_m$, the individual estimators $\tilde{\delta}_i, i=1,\ldots,m$, are independent. The GTS method is based on the assumption (usually made for a large number of observations for the $i-th$ individual, n_i) that the individual parameter estimations, $\tilde{\delta}_i, i=1,\ldots,m$, where the $\tilde{\delta}_i$ conditional to the $\delta_1,\ldots,\delta_m$ only depend on δ_i are normally distributed, more precisely:

$$\tilde{\delta}_i \mid \delta_i \sim N(\delta_i, \Delta_i). \qquad (5.8)$$

The assumption (5.8) is justified when common methods such as generalized least squares (Carroll and Ruppert, 1988) are used to obtain the individual estimates $\tilde{\delta}_i$.

It then follows under (5.3) that, approximately for large n_i, $\tilde{\delta}_i$ are unconditionally independent with $\tilde{\delta}_i \sim N(X_i\delta, \Delta_i + D)$. Thus, δ and D may be estimated by standard techniques applied to this approximate model $\tilde{\delta}_i \cong X_i\delta + \eta_i + \tilde{u}_i$, where $\tilde{u}_i$ has mean zero and known covariance matrix Δ_i. This may be carried out either via an EM algorithm as described by (Davidian and Giltinan, 1995) or standard linear mixed effects software (e.g. Pinheiero and Bates, 1995, Littel et al. 1996). Also, according to Davidian and Giltinan (1995), the GTS estimator for D is likely to be biased, and the bootstrap bias corrected can be obtained in the same way as above

$$\hat{D}_{GTS,BC} = 2\hat{D}_{GTS} - E^*(\hat{D}^*_{GTS}), \text{where } E^*(\hat{D}^*_{GTS}) \cong B^{-1}\sum \hat{D}^{*,b}_{GTS} \qquad (5.9)$$

Similarity as described in the above sections, under the GTS procedure, for each sample bootstrap obtained using the same process as above, the iterative EM algorithm can be used to obtain the bootstrap replicate of $\hat{\delta}_{GTS}$ and $\hat{D}_{GTS}$, based on the following model:

$$\tilde{\delta}^*_i \cong X_i\hat{\delta}_{GTS} + \eta^*_i + u^*_i \quad (5.10)$$

where u^*_i has zero mean and covariance matrix Δ^*_i.

5.2.5. Resampling when covariance matrices are parameterized

A last problem to be considered, concerning the bias corrected estimators, $\hat{D}_{BC,STS}$ (based on any flavor of bias estimation, within or between experimental units), is that they are not always positive semi definite, that is, that not always conduct to true covariance matrices. An alternative approach that ensures positive semi definiteness but not full bias correction, is to resample over adequate parameterizations of the covariance matrix. This approach is described in the following subsections.

5.2.5.1. Cholesky parameterization

Suppose that $\hat{D}_{STS}$ is positive definite. It may be factorized as $\hat{D}_{STS} = L'(\lambda_{STS})L(\lambda_{STS})$, where L is an upper triangular matrix obtained from a $n(n+1)/2-$dimensional vector of unconstrained parameters λ_{STS}. If $\hat{\lambda}^{*,k}_{STS}$ stands for the Cholesky parameterization of the covariance matrix $\hat{D}^{*,k}_{STS}$ obtained in the $k-th$, $k=1,\ldots,B$, bootstrap replication, we have

$$\hat{\lambda}_{BC,STS} = 2\hat{\lambda}_{STS} - \frac{1}{B}\sum_{k=1}^{B}\hat{\lambda}^{*,k}_{STS}$$

and a bootstrap- corrected variance covariance matrix for the standard two-stage method:

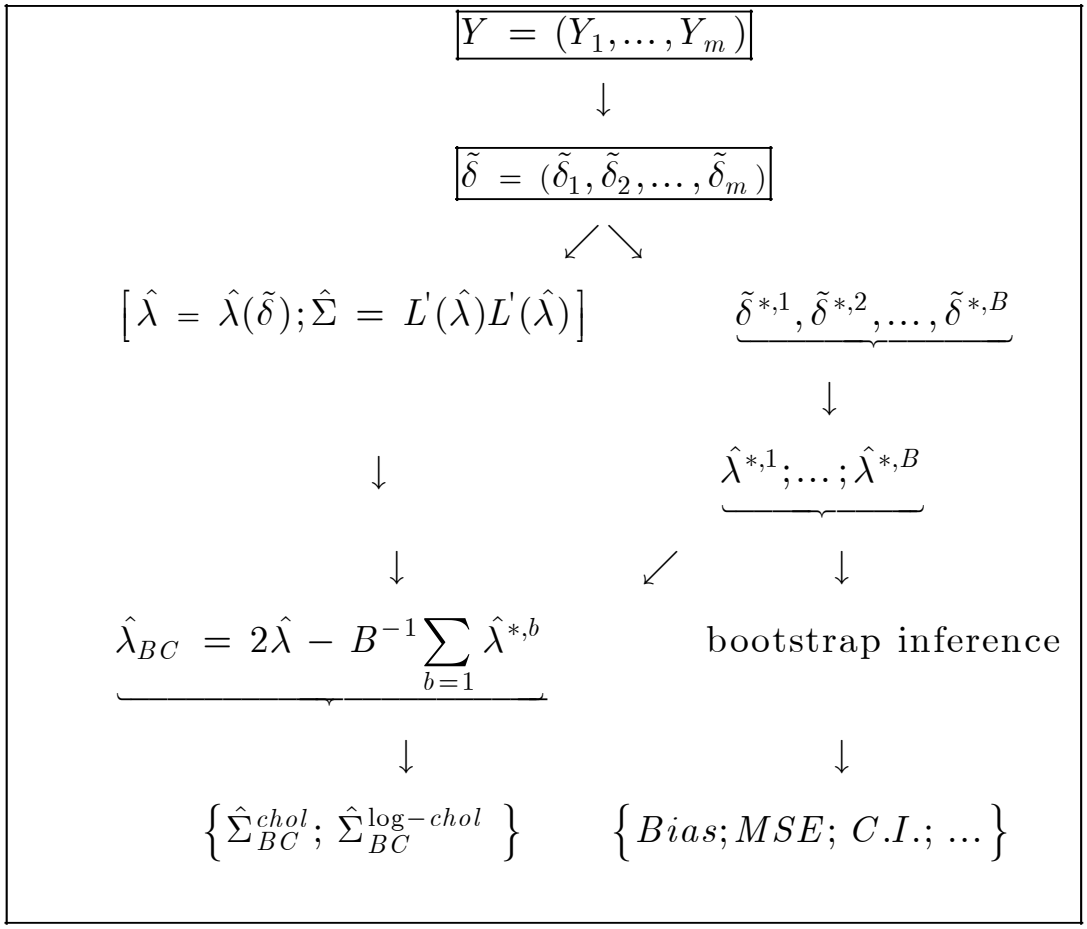
$$\hat{\Sigma}^{chol}_{BC,STS} = L'(\hat{\lambda}_{BC,STS})L(\hat{\lambda}_{BC,STS}) .$$

A problem with the Cholesky parameterization is that the Cholesky factor is not unique. In fact, if L is a Cholesky factor of $\hat{D}_{STS}$, then so is any matrix obtained by multiplying a subset of the rows of L by -1. This conducts to another matrix $\hat{\Sigma}^{chol}_{BC,STS}$. To avoid this problem we can use the called Log-Cholesky parameterization, as follows:

5.2.5.2. Log-Cholesky parameterization

The log-Cholesky is the same as the Cholesky parameterization, except that the logarithms of the diagonal elements of L are used.

The following diagram schematically illustrates the resampling procedure



5.2.6. Confidence Intervals for fixed effects

In this section we introduce some confidence intervals for fixed effects. They are based on approximating the distribution of $\hat{\delta}$ either by a gaussian distribution (the usual approach for GTS methods) or by the bootstrap distribution. The bootstrap methods will be designated as B_STS ("Naïve" bootstrap based on the STS method), B_GTS ("Naïve" bootstrap based on the GTS method), B_STS_2levs (two stage bootstrap method based on STS assuming a given

model –section 5.2.2) and B_STS_MCB (bootstrap method under STS method based on the Markov Chain bootstrap process and not assuming any model –section 5.2.1).

5.2.6.1. Bootstrap percentile confidence intervals (BPI)

Standard parametric confidence intervals can provide a measure of significance for fixed coefficients. They require the acceptance of normality assumptions regarding the estimates of coefficients. In the data analysis described in preceding sections, diagnostics did not always support these assumptions. Alternatives to the standard parametric confidence intervals are nonparametric methods using bootstrap estimates of the distribution of the coefficient estimates. As a first, simple approach, we have considered nonparametric bootstrap percentile confidence intervals. This method uses the $\alpha/2$ and 1-$\alpha/2$ quantiles of the empirical distribution of the bootstrap estimates as a 1-α level confidence interval. An approximate $1-\alpha$, equally-tailed double-sided bootstrap percentile confidence interval for a fixed parameter δ is given by $\left(\hat{\delta}^{(\alpha/2)}, \hat{\delta}^{(1-\alpha/2)}\right)$, where $\hat{\delta}^{(\alpha/2)}$ is the $\alpha/2-$quantile of the resampled bootstrap estimates $\left\{\hat{\delta}^{*,b}, b=1,\ldots,B\right\}$ (see, for example, Efron, and Tibshirani,. 1993).

The percentile method described here is the simplest and most popular bootstrap confidence interval, and is often used in practice. This is because it is easy to compute, simple to motivate, and was popularized by Efron early in the history of the bootstrap. Additionally, it is translation invariant, that is, if we define $f(\hat{\delta})$ as the parameter of interest for a monotonic function f, then the percentile method applied to this problem will produce the confidence interval $\left(f\left[\hat{\delta}^{(\alpha/2)}\right], f\left[\hat{\delta}^{(1-\alpha/2)}\right]\right)$. This is a desirable property, but percentile intervals may not be the best in all cases. In particular, they may not have the claimed coverage. Confidence interval coverage is the probability that the confidence interval includes the true parameter, under repeated sampling from the same underling population. When the coverage is the same as the stated size of the confidence interval (e.g. true coverage =95% for a nominally 95% confidence interval), the intervals are accurate. There are methods to improve the percentile approach, but

they are not explored here. Provided that adequate estimators of the standard error of the fixed effects estimates exist, we only consider studentized bootstrap intervals as an alternative to plain percentile intervals.

5.2.6.2. Symmetrized bootstrap studentized confidence intervals (SBI)

The problems with the percentile method outlined above can be circumvented by the alternative approach provided by the studentized bootstrap or bootstrap-t intervals (see, for example, Hall, 1988). Suppose that the standard error of $\hat{\delta}$ is $se(\hat{\delta})$, and denote its corresponding estimate as $\hat{\sigma}_{\hat{\delta}}$. The studentized bootstrap is based on approximating the distribution function H_F of $(\hat{\delta}-\delta)/\hat{\sigma}_{\hat{\delta}}$ by the distribution, $\hat{H}^* = H_{F_n}$, of $(\hat{\delta}^*-\hat{\delta})/\hat{\sigma}^*_{\hat{\delta}}$. The 1-$\alpha/2$ and $\alpha/2$-quantiles of H_{F_n} may be used to calculate the confidence interval. That is to say, if t^*_L and t^*_U are defined as $P_{\hat{H}^*}\left\{(\hat{\delta}^*-\hat{\delta})/\hat{\sigma}^*_{\hat{\delta}} \le t^*_L\right\} = \alpha/2$ and $P_{\hat{H}^*}\left\{(\hat{\delta}^*-\hat{\delta})/\hat{\sigma}^*_{\hat{\delta}} \ge t^*_U\right\} = \alpha/2$, the bootstrap-t interval is defined as $\left[\hat{\delta} - t^*_U\hat{\sigma}_{\hat{\delta}}, \hat{\delta} - t^*_L\hat{\sigma}_{\hat{\delta}}\right]$.

The symmetrized bootstrap-t interval, is defined as

$$\left(\hat{\delta} - t^*_{(1-\alpha)}\hat{\sigma}_{\hat{\delta}}, \hat{\delta} + t^*_{(1-\alpha)}\hat{\sigma}_{\hat{\delta}}\right)$$

where $t^*_{(1-\alpha)}$ stands for the quantile estimate given by $P_{\hat{H}^*}\left\{\left|(\hat{\delta}^*-\hat{\delta})/\hat{\sigma}^*_{\hat{\delta}}\right| \le t^*_{1-\alpha}\right\} = 1-\alpha$.

5.2.6.3. Asymptotic confidence intervals: GTS method (ASI)

For the purpose of comparability with the other methods, the asymptotic confidence intervals based on GTS are also considered. From this point of view and according to the results obtained in the above sections, standard formulae may then be used to construct standard error estimates of $\hat{\delta}$; i.e. using the approximate, for large m, covariance matrix:

$$\mathrm{var}(\hat{\delta}) \approx \left(\sum_{i=1}^{m} X_i^{'} \left(C_i^{'} + D \right)^{-1} X_i \right)^{-1} .$$

Assuming approximate normality, the asymptotic 95% confidence interval can be calculated as $\hat{\delta} \pm 1.96S$ where S is the estimated standard error of $\hat{\delta}$.

5.2.7. Bootstrap test of comparison of covariance matrices

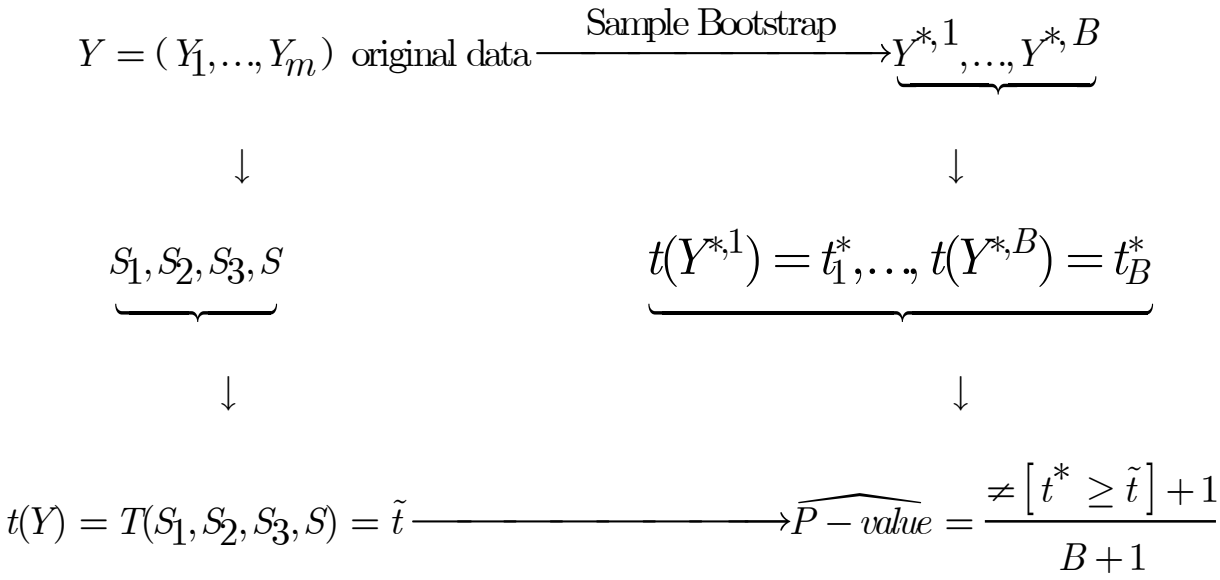
The Bootstrap provides a method for performing the null hypothesis test for the covariance matrices problem. In this section we use the bootstrap to test the validity of the assumption that the covariance matrices are the same over the various "treatments groups". The bootstrap procedure for testing this assumption has been based on the statistic given by Box(1950) as an extension to the multivariate case of Bartlett´s homogeneity of variances test (see Box, 1949):

$$T = (m - I)\log\left(|S|\right) - \sum_{i} (m_i - 1)\log\left|S_i\right|$$

where $m = \sum_{i} m_i, i = 1,2,\ldots,I, j = 1,\ldots,m_i$ and $S = (m-I)^{-1} \sum_{i} \sum_{j} (y_{ij} - y_{i\cdot})(y_{ij} - y_{i\cdot})^{'}$ is the bias corrected pooled covariance estimator, and $S_i = \left(m_i - 1\right)^{-1} \sum_{j} (y_{ij} - y_{i\cdot})(y_{ij} - y_{i\cdot})^{'}$ is the unbiased estimate, based on $m_i - 1$ degrees of freedom, of the covariance for the $i-th$ treatment group. $y_{i\cdot}$ stands for the mean in the i th group:

$$y_{i\cdot} = m_i^{-1} \sum_{j} y_{ij} \ .$$

The bootstrap test procedure for testing $H_0 : \Sigma_1 = \Sigma_2 = \ldots = \Sigma_I$ where Σ_i corresponds to the true covariance matrix in the $i-th$ treatment group (e.g. Diet i, i=1...3 for the breast cancer data) can be summarized in the following diagram:



5.2.8. Computational aspects

Macros and functions in various statistical packages include bootstrap confidence interval calculations. These include various libraries of S-plus functions Efron and Tibshirani(1993), Venables and Ripley(1994) and Davison.and Hinkley(1997). All macros and libraries can bootstrap a single sample of observations and compute bias, standard error and a variety of confidence intervals. Some packages (e.g. the "boot" library in Davison.and Hinkley, 1997) can be easily extended for multiple sample problems. In the example below, some new S-plus function such as *boot.sts*, *boot.sts.2levs, boot.sts.MCB* and *boot.gts* are used in the computations described below. All these functions are available on request to the author.

EXPERIMENTAL RESULTS ON BREAST CANCER DATA

As can be seen from Figure 5.1, the estimated values of the "individual" coefficients are not conditionally normal. Therefore the usual two stage analysis conditions are not satisfied. A common practice is to take the natural logarithms of the skewed variables ("individual" parameters) and use these transforms in the "linear" regression model in the second step.

This procedure, however, is rarely justified with theoretical arguments and furthermore twists the interpretation of the estimated coefficients.

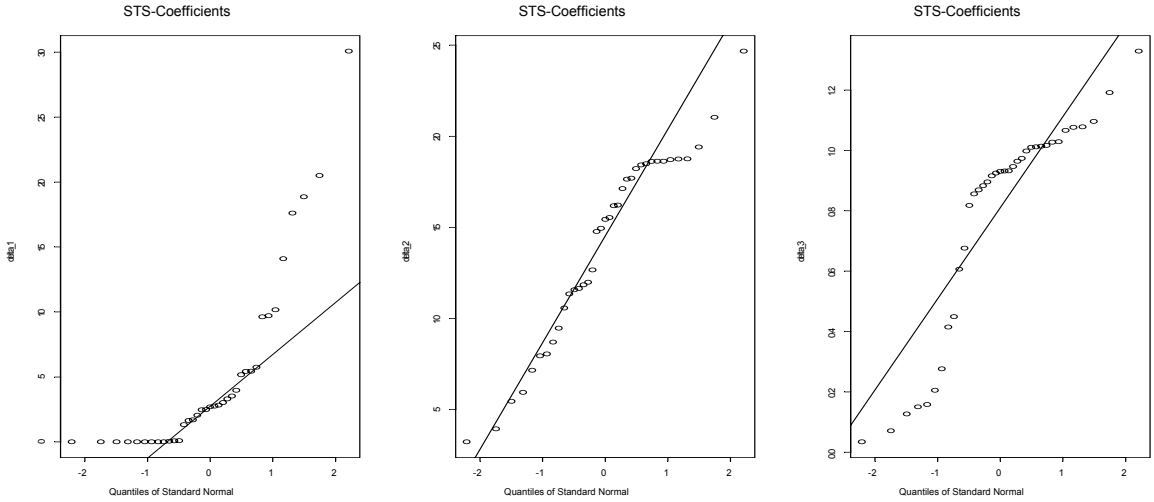


Figure 5.1: qqnorm of GLS-estimates of individual parameters

The bootstrap techniques suggested above could be applied with no assumptions about the underling distributions. We used bootstrap techniques for two purposes. First, we used the bootstrap to obtain the parameter estimates of the breast cancer model. This allows us to graphically view these distributions. Secondly, we used bootstrap techniques to evaluate confidence intervals. All bootstrap computations are based on 1000 resampling coefficients (fixed parameters and variance components) (B=1000). This is generally considered sufficient for the estimation of the confidence intervals and the bootstrap biases. For the purpose of comparability with the two-stage method, the "original" parameter estimates are added.

RESULTS FOR VARIANCE COMPONENTS

The bootstrap bias correction procedure over the breast cancer data produces the following estimations of D:

Table (5.1): Bootstrap bias corrections of the unstructured covariance matrix.

	STS method				GTS method: EM algorithm	
	STS-est.	$D^{*}_{BC,STS}$	D^{*}_{BC,STS_2levs}	D^{*}_{BC,STS_MCB}	GTS-est.	$D^{*}_{BC,GTS}$
D_{11}	48.71213519	50.0523616	39.69755492	31.4246281	48.91287792	49.44036719
D_{12}	0.90561130	0.9069809	-2.95941305	16.9362445	-0.88839969	-0.74447771
D_{22}	27.69682549	28.3448530	33.88644542	2.2404023	25.85852581	27.11840588
D_{13}	0.74762816	0.7552587	1.51554059	1.4435605	0.08340262	0.05823888
D_{23}	-0.09962465	-0.1060989	-0.01186686	0.4501587	-0.09952978	-0.08903528
D_{33}	0.12779023	0.1297800	0.22712654	0.1901365	0.05746800	0.05690517

Table (5.2): Bootstrap bias corrections based on the Cholesky decomposition

	STS method				GTS method: EM algorithm	
	STS-est.	$\hat{\Sigma}^{chol}_{BC,STS}$	$\hat{\Sigma}^{chol}_{BC,STS\_2levs}$	$\hat{\Sigma}^{chol}_{BC,STS\_MCB}$	GTS-est.	$\hat{\Sigma}^{chol}_{BC,GTS}$
D_{11}	48.71213519	52.0274329	42.4385817	44.1838183	48.91287792	51.31949942
D_{12}	0.90561130	1.7144244	-4.4475002	23.3636652	-0.88839969	-0.97640158
D_{22}	27.69682549	29.0039745	36.3885507	19.4323503	25.85852581	27.88261836
D_{13}	0.74762816	0.7805081	1.5270503	1.4682314	0.08340262	0.05668925
D_{23}	-0.09962465	-0.1216076	-0.1097385	0.6681299	-0.09952978	-0.10561462
D_{33}	0.12779023	0.1368401	0.3281264	0.2653427	0.05746800	0.06011372

Table (5.3): Bootstrap bias corrections based on the log-Cholesky decomposition

	STS method				GTS method: EM	
	STS-est.	$\hat{\Sigma}^{\log-chol}_{BC,STS}$	$\hat{\Sigma}^{\log-chol}_{BC,STS\_2levs}$	$\hat{\Sigma}^{\log-chol}_{BC,STS\_MCB}$	GTS-est.	$\hat{\Sigma}^{\log-chol}_{BC,GTS}$
D_{11}	48.71213519	54.4990901	44.7017828	49.2999821	48.91287792	53.57189128
D_{12}	0.90561130	1.7413381	-4.5716948	25.1805709	-0.88839969	-0.99158009
D_{22}	27.69682549	29.3067149	38.5639424	22.3703361	25.85852581	28.28460577
D_{13}	0.74762816	0.7983215	1.6228989	1.5722753	0.08340262	0.05787030
D_{23}	-0.09962465	-0.1223314	-0.1119759	0.6901127	-0.09952978	-0.10639233
D_{33}	0.12779023	0.1386600	0.6873935	0.3832489	0.05746800	0.06157452

The "direct" bias corrected bootstrap for two methods, B_STS and B_GTS, were in close agreement with the parametric estimations described in chapter 3. However, the B_STS_2levs and B_STS_MCB methods produced estimates of different magnitude compared to other methods, including estimation under different parameterizations (the Cholesky and log-Cholesky

parameterizations). From these single values, it is not possible to decide which estimations are more correct.

RESULTS FOR FIXED PARAMETERS

The results seem more interesting for the fixed parameter estimates. Table (5.4) lists fixed parameters estimates, with standard errors and confidence intervals using both the Bootstrap and "conventional" (two stage) statistical methods. B_STS produces the lowest standard error for the fixed parameter estimate of δ_1. Except for the B_STS_MCB method, all other methods yield nearly identical standard errors for δ_2. B_GTS and GTS methods give smaller standard errors than the others for δ_3. In addition, the 95% percentile and symmetrized bootstrap–t confidence intervals are displayed, jointly with the normal theory confidence intervals in Tables (5.5) to (5.7).

Table (5.4):

B=1000						
	δ_1		δ_2		δ_3	
Method	Estimate	SE	Estimate	SE	Estimate	SE
STS	5.036219	1.147408	14.035850	0.8651953	0.7823709	0.05876896
GTS	5.017261	1.358182	14.391253	0.78424459	0.9152000	0.002004236
B_STS	5.000351	1.131513	14.017100	0.8550138	0.7785176	0.05830000
B_STS_2lev	7.094566	1.477435	13.280970	0.8992713	0.9295607	0.03281113
B_STS_MC	5.463674	1.594982	10.767760	1.4293270	0.8811188	0.05038420
B_GTS	5.052136	1.437375	14.446860	0.7652393	0.9098835	0.0021300

Table (5.5):

B=1000				
δ_1 (α=0.05) : *ASI* [2.355223, 7.679299]				
	BIP		*SBI*	
Method	Lower	Upper	Lower	Upper
B_STS	2.934145	7.440812	2.325412	7.747027
B_STS_2levs	4.453764	10.20931	1.923673	8.148766
B_STS_MCB	2.707738	8.940004	2.325351	7.747087
B_GTS	3.130019	7.387432	1.694911	8.337461

Table (5.6):

B=1000				
δ_2 (α=0.05) : *ASI* [12.85413, 15.92837]				
	BIP		*SBI*	
Method	Lower	Upper	Lower	Upper
B_STS	12.32816	15.73007	12.23996	15.83174
B_STS_2levs	11.42171	15.15443	11.74568	16.32602
B_STS_MCB	7.958572	13.67785	10.37627	17.69543
B_GTS	12.68496	16.25886	12.29062	16.49302

Table (5.7):

B=1000				
δ_3 (α=0.05) : *ASI* [0.9112717, 0.9191283]				
	BIP		*SBI*	
Method	Lower	Upper	Lower	Upper
B_STS	0.6513871	0.8915256	0.6511329	0.913609
B_STS_2levs	0.8502308	0.9859209	0.2334169	1.331325
B_STS_MCB	0.7733006	0.9697785	0.3618236	1.202918
B_GTS	0.8530017	0.9677997	0.8467297	0.9836403

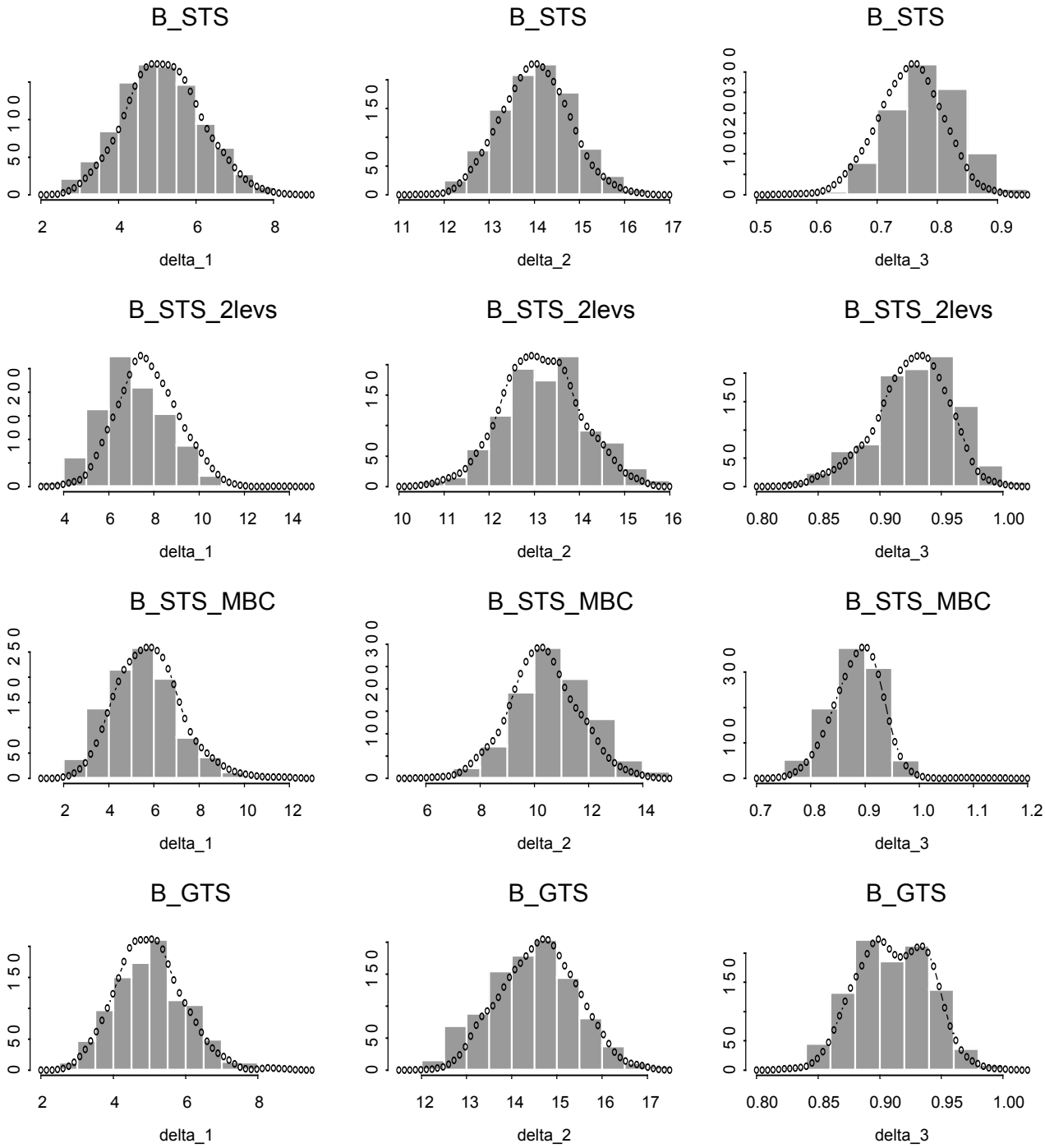
The intervals are narrower for B_STS than for the other methods except for δ_3 where *ASI* is suspiciously narrow. The B_STS_2levs and B_STS_MCB produce much wider intervals compared to the other methods, for all fixed parameters.

In general, the preceding tables are merely illustrative and it is difficult to decide if the bootstrap methods give better o worse results compared with the asymptotic methods. In the next chapter, we will try to illustrate the behavior of these intervals and compare the methods, using simulation.

For illustrative purposes, we can use the 95% confidence intervals to test whether the population parameter has a specific value, 0, against the two-sided alternative that it does not have a 0 value. The null hypothesis $H_0 : \delta = 0$ is rejected if and only if the corresponding confidence interval does not contain 0. From this point of view, the preceding null hypothesis is rejected in all cases.

Additionally, the bootstrap procedure test based on the statistic given by Box(1950) described in section 5.3.2 for testing the null hypothesis that no treatment groups effects for the breast cancer data accept this hypothesis with $\widehat{P-value} = 0.41758$, suggesting that there are no differences between the covariances of treatment groups.

Another indication about the characteristic of the bootstrap estimates can be displayed in Figure 5.2 in which the bootstrap estimates of the fixed effects under B_STS are unimodally distributed, centered about the parameter estimates and nearly Gaussian, but for other methods, the corresponding bootstrapped coefficient seems to be skewed distributed, more clearly for δ_1 and δ_3.


B_STS
0 50 100
2 4 6 8
delta_1
B_STS
0 50 150
11 12 13 14 15 16 17
delta_2
B_STS
0 100 200 300
0.5 0.6 0.7 0.8 0.9
delta_3
B_STS_2levs
0 100 200
4 6 8 10 12 14
delta_1
B_STS_2levs
0 50 150
10 11 12 13 14 15 16
delta_2
B_STS_2levs
0 50 150
0.80 0.85 0.90 0.95 1.00
delta_3
B_STS_MBC
0 50 150 250
2 4 6 8 10 12
delta_1
B_STS_MBC
0 100 200 300
6 8 10 12 14
delta_2
B_STS_MBC
0 100 300
0.7 0.8 0.9 1.0 1.1 1.2
delta_3
B_GTS
0 50 150
2 4 6 8
delta_1
B_GTS
0 50 150
12 13 14 15 16 17
delta_2
B_GTS
0 50 150
0.80 0.85 0.90 0.95 1.00
delta_3

5.3. Summary and conclusions

After introducing some asymptotic and bootstrap inferential methods for repeated measures data under a nonlinear mixed models context, we have analyzed the breast cancer data, for illustrative purposes. This is motivated by the fact that some of the assumptions of the nonlinear mixed model do not seem to be satisfied in these data.

The results for the fixed and random components of the corresponding mixed models are very different, depending on the specific inferential procedures and resampling strategies. Its relative accuracy is the subject of the next section.

Chapter 6

6. Validity of the resampling approach on mixed models

6.1. Introduction

For non-linear mixed models, popular estimation and inference methods are ML, REML and two stage methods, with errors modelled for serial correlation and sometimes conditional heteroskedasticity. However, the simulation results in previous chapters of this monograph, and in related literature, suggest that the asymptotic approximation works poorly in typically available samples, even under normal conditions with large sizes. Thus, it is important to find more reliable inference tools for this class of models. A natural candidate to achieve this goal is the bootstrap. In this chapter we investigate the performance of some of bootstrap resampling schemes discussed in the preceding chapter, with the help of Monte Carlo simulations, on the non-linear mixed models and repeated measures context. In concrete, the resampling methods under study are the "naïve" methods described at the end of section 5.2.4, based on directly resampling the parameter sets of each individual or unit, fitted in the first step of a two-stage method. These methods were designated as B_STS and B_GTS. These simplistic methods are compared with the method designated as B_STS_MCB in section 5. It is based on the two-step resampling schema described in section 5.2.1, taking at random and with replacement complete individuals or units (their associated data vectors) in the first step, and performing a Markov Chain bootstrap inside each one of these pseudodata sets in the second step. Supposedly, this method is the most robust with respect to the choice of the nonlinear model and in front of the possible true distribution of the data; it provides construction of the

bootstrap pseudo samples from the original sample without considerable break of the "individual" series structure and in a way that does not depend on distributional and model assumptions.

As all the preceding bootstrap methods may be considered as bootstrap improvements of the STS and GTS two-stage methods, described in section 2, these simulations were complemented with simulations directly concerning the STS and GTS estimators of the fixed effects, the residual variance and the covariance matrix of the random effects, and the true coverage of an asymptotic confidence interval associated to the GTS method, described below.

Several parameter settings were investigated, all based on true experiments. We focus on the use of this improved bootstrap approach to sharpen the properties of the point estimates of fixed parameters and the point estimates of covariance components of the random effects in terms of bias, mean squared error and coverage probability of two-sided 95% level intervals. Coverage of confidence intervals is defined as the relative frequency with which the final confidence interval contains the true value. Ideally, a 95% confidence interval would not include the true value of a parameter δ, due to overestimation, only a 2.5% of the times and, likewise, due to underestimation a 2.5% of the times. Here, underestimation is taken to mean that the confidence interval lies entirely below (*i.e.* includes only values less than) the true value of the parameter, and the opposite for overestimation.

Experimental analysis of coverage is essential for assessing the quality of practical implementations of bootstrap methods. Two types of intervals where included: the first one has been constructed using the basic asymptotic results for GTS (global two stage) method, based on its associated asymptotic standard error, and assuming that the estimate is normally distributed. The second type is based on bootstrap methods for confidence interval construction, concretely the percentile intervals and the symmetric bootstrap-t confidence intervals. The breast cancer model used to describe the (real) growth dynamic of tumour was selected to illustrate the proposed methods and to motivate the simulation scenarios.

Additionally, we developed some theoretical results for the bootstrap estimators.

6.2. Simulations studies

In this section, we describe the Monte Carlo simulations to examine and test the validity and performance of the bootstrap procedures introduced in the previous chapter and mentioned above in the introduction. The design of the simulation is the same than that reported in chapter 4, with the same starting parameters, mainly suggested by the analysis of breast cancer data. The procedure is as follows:

I. Choose N, the number of simulations. A common choice is $N = 10000$. However, since estimating a non-linear regression model with random and fixed effects is a very computer intensive process, and very time consuming, $N = 1000$ might be a more realistic choice. It is also possible to start with a small value like 100 and then increase it if the outcome of the results is not clear. Some of the results in this section are based on just 100 simulation replicates and must be considered just as initial prospective results. Its non-conclusive character is clear in view of the low precision of the corresponding simulation results, associated to wide confidence intervals.

II. Generate N simulated samples according to a non-linear mixed model, like the breast cancer model. Random effects and errors were generated from variety of distributions to verify if the nonparametric bootstrap procedures were not influenced by these distributions. In particular, both random effects and errors were generated from:

    N- normal distribution

    U- uniform distribution

    E- exponential distribution

    G- gamma distribution.

    The normal and uniform distributions were used as reference symmetric distributions, commonly used in the literature. The gamma and exponential distributions were chosen because they are skewed distributions.

III. In any single simulation loop, B repetitions were used to construct a bootstrap distribution, and to compute the measures of interest like parameter estimates, (1-α)% confidence intervals and so on, for the different bootstrap resampling approaches

(B_STS_MCB, B_STS and B_GTS). Note that the B_STS_MCB approach is very computer intensive, as it involves repeatedly solving non-linear optimisation problems (re-estimating the parameters) for each new resampled pseudo individual or unit.

IV. Steps II-III were repeated N times, and the estimates of statistical performance like bias, MSE, average width of the intervals, true coverage and so on (the same measures detailed in chapter 4) were computed.

The N and B sizes were dependent on each concrete scenario. They are given in Table (6.1). The results with small (100) sample sizes may be taken as preliminary but informative results.

Table (6.1): Input size values for "star" design simulation scenarios.

Scenario (method)	Simulation size	Bootstrap size
B_STS_MCB	N =100	B=1000
B_STS	N =1000	B=1000
B_GTS	N =100	B=100
STS	N=1000	–
GTS	N =100	–

The simulations were conducted on several different PC environments. The majority of the simulations were completed in Pentium III and few simulations where also performed on Pentium IV and Pentium IV Xeon computers. The full simulations are very computer-intensive and very time consuming, due to various factors considerations such as: Iterative method used to fit the NLME model, large number of data sets processed per scenario, 16 combinations of random effects and errors distributions and so on. The computer intensive character was also due to the software tool used in the simulations, S-plus (version 2000).

6.1.1. Results for fixed effects

We start with the results concerning Markov chain bootstrap method (B_STS_MCB) for the breast cancer model. The corresponding results for fixed effects are presented in Figures 6.1 and Table (6.2) to (6.4). In these Tables, the column labelled "MEAN (bias-cb)" contains the bias-corrected bootstrap estimate $2\hat{\delta} - \overline{\delta}^*$, where $\overline{\delta}^* = B^{-1}\sum \hat{\delta}_j^*$, obtained by subtracting the

bias estimate $\overline{\delta}^* - \hat{\delta}$ from $\hat{\delta}$. The remaining columns have the same meaning as in chapter 4. Tables (6.5) to (6.7) contain similar results for the same bias-corrected estimates, for the STS and GTS estimators, but when bootstrap is based on the "naïve" strategy designated as B_STS and B_GTS, respectively, in chapter 5. Finally, Tables (6.8) to (6.10) give the results for the STS and GTS estimators, directly, with no bias correction. The bias results for STS, GTS (without bias correction) and B_STS and B_GTS are also graphically summarized in Figure 6.2.

The results on the bias of the STS and GTS methods and their correction by the bootstrap methods are not conclusive, especially due to the lack of precision of the simulation results for the GTS-based methods and for the B_STS_MCB bootstrap procedure. As is predicted by theory, the STS and GTS estimators for the fixed effects are considerably biased, and this bias is affected by the choice of the simulated distribution for the random effects and for the residuals. The naïve bootstrap methods B_STS and B_GTS completely fail in correcting this bias, and should not be recommended as a bias correction procedure. Their bias is quite similar to the bias of the original estimators, but with a considerable increase in the MSE. These results agree with the other work of Sanchez (University of Barcelona), in which under the linear mixed model fashion all simulation results exhibit a bias. The author also reveals that the bias can be reduced by "adjusting" the covariance matrices of the random effects and residuals samples to equal the covariance matrices in the original sample (see Sanchez and Ocaña, (2001)). Further investigation is needed to determine the extent to which this procedure affects the general application of the non-linear mixed model.

Bootstrap bias-correction based on the B_STS_MCB procedure seems to yield estimates of fixed effects with slightly reduced bias. Their expectation seems also more stable in front of the distribution of the random effects and of the residuals, especially when the results in Figure 6.1 and Figure 6.2 are compared. On the other hand, their MSE is also clearly higher than the MSE of the initial STS and GTS estimators. This makes very difficult to recommend these bootstrap procedures as a bias correction device: they provide not clear bias reduction associated to a clear MSE increase.

Table (6.2): summary results for fixed effect parameter estimate δ_1 of 1000 simulations

B_STS_MCB- METHOD			
Random effects/ Residuals distribution	MEAN (bias-cb)	BIAS $\pm$C.I$_{\text{Bias}}$	MSE
NN	5.492216	0.4409804 ± 0.5569311	8.10702
NU	4.980802	-0.0704343 ± 0.4224747	4.558146
NE	6.027204	0.9759683 ± 0.5254446	7.92381
NG	6.005741	0.9545048 ± 0.5195447	7.726702
UU	5.178007	0.1267705 ± 0.608664	9.370446
UE	5.20605	0.1548142 ± 0.5887898	8.0555
UG	4.636939	-0.4142974 ± 0.6214919	9.92447
UN	5.178694	0.1274579 ± 0.4976337	6.269111
EE	6.538958	1.487722 ± 0.6192685	11.89649
EG	5.992006	0.9407701 ± 0.5468846	8.436857
EU	5.915088	0.8638523 ± 0.6120571	10.2052
EN	6.513922	1.462686 ± 0.5446181	9.628795
GG	6.133922	1.082686 ± 0.5694178	9.19034
GE	5.595645	0.5444094 ± 0.5694998	8.485681
GU	6.343787	1.292551 ± 0.4846859	7.602404
GN	5.752155	0.7009195 ± 0.5311182	7.613944
Breast cancer model: $y_{ij} = \frac{\delta_{1i} \exp(t_{ij} - \delta_{2i})}{1 + \exp((t_{ij} - \delta_{2i}) / \delta_{3i})} + e_{ij}$			

Table (6.3): summary results for fixed effect parameter estimate δ_2 of 1000 simulations

B_STS_MCB- METHOD			
Random effects/ Residuals distribution	MEAN (bias-cb)	BIAS $\pm$C.I$_{\text{Bias}}$	MSE
NN	14.18376	0.3167331 ± 0.2745375	2.023045
NU	14.16097	0.2939429 ± 0.2773216	2.048323
NE	14.05808	0.1910501 ± 0.2427168	1.524008
NG	14.09512	0.2280946 ± 0.2441882	1.557624
UU	14.58369	0.7166569 ± 0.2928523	2.679091
UE	14.26194	0.394914 ± 0.2467041	1.565995
UG	14.02861	0.1615841 ± 0.2925511	2.187151
UN	14.20974	0.3427131 ± 0.2984908	2.367135
EE	12.77231	-1.094716 ± 0.2785328	3.157304
EG	12.74971	-1.117315 ± 0.2226059	2.499611
EU	13.16928	-0.6977527 ± 0.2638041	2.244065
EN	13.11887	-0.7481602 ± 0.2622835	2.29675
GG	12.70996	-1.157069 ± 0.2432679	2.802268
GE	12.73231	-1.134725 ± 0.2491455	2.854948
GU	12.99361	-0.8734164 ± 0.2248572	2.03951
GN	12.99794	-0.8690927 ± 0.2705857	2.604034
Breast cancer model: $y_{ij} = \frac{\delta_{1i} \exp(t_{ij} - \delta_{2i})}{1 + \exp((t_{ij} - \delta_{2i})/\delta_{3i})} + e_{ij}$			

Table (6.4): summary results for fixed effect parameter estimate δ_3 of 1000 simulations

B_STS_MCB- METHOD			
Random effects/ Residuals distribution	MEAN (bias-cb)	BIAS $\pm$C.I$_{\text{Bias}}$	MSE
NN	0.8420027	-0.006652756 $\pm$ 0.01063725	0.002930764
NU	0.8342563	-0.01439918 $\pm$ 0.01075492	0.00315806
NE	0.830311	-0.01834454 $\pm$ 0.01011937	0.002922151
NG	0.8474471	-0.001208395 $\pm$ 0.01023377	0.002645884
UU	0.8340125	-0.01464296 $\pm$ 0.01114056	0.003348233
UE	0.8475641	-0.001091445 $\pm$ 0.01079652	0.0027017
UG	0.841077	-0.007578529 $\pm$ 0.009413824	0.002295082
UN	0.8317882	-0.01686732 $\pm$ 0.01085094	0.003257502
EE	0.8168466	-0.03180894 $\pm$ 0.01334297	0.005507172
EG	0.8603258	0.01167025 $\pm$ 0.01289654	0.004335773
EU	0.8583597	0.009704157 $\pm$ 0.01303928	0.004387231
EN	0.8459545	-0.002701031 $\pm$ 0.0139102	0.004892992
GG	0.8396211	-0.009034364 $\pm$ 0.01241393	0.003892541
GE	0.8507041	0.002048583 $\pm$ 0.01353109	0.00462721
GU	0.7983597	-0.05029584 $\pm$ 0.01303928	0.006822732
GN	0.7859545	-0.06270103 $\pm$ 0.0139102	0.008817115
Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$			

Table (6.5): summary results for fixed effect parameter estimate δ_1 under B_STS and B_GTS

DISTRIBUTIONS	RANDOM-EFFECTS	ERRORS	B_STS method			B_GTS method		
			MEAN (bias-cb)	BIAS $\pm$C.I$_{\text{Bias}}$	MSE	MEAN	BIAS $\pm$C.I$_{\text{Bias}}$	MSE
	N	**N**	5.97113	0.91989 ± 0.18634	9.876	5.865769	0.81453 ± 0.5663	8.9271
	N	**U**	6.030751	0.97951 ± 0.1844	9.7978	6.052808	1.00157 ± 0.5329	8.3216
	N	**E**	5.985496	0.93426 ± 0.1844	9.7107	5.670251	0.61902 ± 0.6109	10.0013
	N	**G**	5.973124	0.92189 ± 0.1897	10.2115	6.103062	1.05183 ± 0.5427	8.6956
	U	**U**	6.098975	1.04774 ± 0.089	3.1562	5.784428	0.73319 ± 0.453	5.8261
	U	**E**	6.086178	1.03494 ± 0.1833	9.8058	6.077438	1.0262 ± 0.4999	7.4933
	U	**G**	6.02414	0.9729 ± 0.1721	8.652	5.924831	0.8736 ± 0.5326	8.0741
	U	**N**	6.137514	1.08628 ± 0.0832	2.9797	5.842696	0.79146 ± 0.2584	2.3465
	E	**E**	6.393524	1.34229 ± 0.2091	13.1727	6.248457	1.19722 ± 0.7092	14.3935
	E	**G**	6.331789	1.28055 ± 0.1808	10.1404	6.182964	1.13173 ± 0.6233	11.2926
	E	**U**	6.461471	1.41023 ± 0.1955	11.9252	6.150373	1.09914 ± 0.6488	12.0568
	E	**N**	6.386428	1.33519 ± 0.1869	10.871	6.424975	1.37374 ± 0.6468	12.6673
	G	**G**	6.64454	1.5933 ± 0.1971	12.6377	6.1801	1.12886 ± 0.6585	12.448
	G	**E**	6.428031	1.3768 ± 0.1947	11.7517	6.423666	1.37243 ± 0.5858	10.7269
	G	**U**	6.500926	1.44969 ± 0.1987	12.3727	6.336393	1.28516 ± 0.5687	9.9864
	G	**N**	6.471372	1.42014 ± 0.2031	12.7474	6.351777	1.30054 ± 0.6185	11.5483
	Breast cancer model: $y_{ij} = \dfrac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$							

Table (6.6): summary results for fixed effect parameter estimate δ_2 under B_STS and B_GTS.

DISTRIBUTIONS	RANDOM- EFFECTSS	ERRORS	B_STS method			B_GTS method		
			MEAN (bias-cb)	BIAS $\pm$I.C$_{Bias}$	MSE	MEAN	BIAS $\pm$I.C$_{Bias}$	MSE
	N	**N**	13.48469	-0.38234 $\pm$ 0.0961	2.5467	13.50138	-0.36565 $\pm$ 0.3425	3.1569
	N	**U**	13.4743	-0.39273 $\pm$ 0.0899	2.2582	13.6368	-0.23023 $\pm$ 0.2659	1.8754
	N	**E**	13.4826	-0.38443 $\pm$ 0.0949	2.4910	13.54676	-0.32027 $\pm$ 0.2681	1.9543
	N	**G**	13.46729	-0.39974 $\pm$ 0.0953	2.5224	13.52795	-0.33908 $\pm$ 0.3081	2.5614
	U	**U**	13.48374	-0.38329 $\pm$ 0.0488	0.7667	13.56324	-0.30379 $\pm$ 0.1998	1.1209
	U	**E**	13.52093	-0.3461 $\pm$ 0.1022	2.8343	13.55655	-0.31048 $\pm$ 0.3251	2.8208
	U	**G**	13.47381	-0.39322 $\pm$ 0.1053	3.0371	13.53218	-0.33485 $\pm$ 0.3385	3.0649
	U	**N**	13.50038	-0.36665 $\pm$ 0.045	0.6613	13.65278	-0.21425 $\pm$ 0.1287	0.4725
	E	**E**	12.82566	-1.04137 $\pm$ 0.094	3.3821	13.00583	-0.8612 $\pm$ 0.2965	3.0075
	E	**G**	12.86248	-1.00456 $\pm$ 0.0915	3.1874	12.8852	-0.98183 $\pm$ 0.2857	3.068
	E	**U**	12.85198	-1.01505 $\pm$ 0.0899	3.1313	13.05214	-0.8149 $\pm$ 0.2613	2.4236
	E	**N**	12.81446	-1.05257 $\pm$ 0.0898	3.2051	13.02579	-0.84124 $\pm$ 0.2612	2.4664
	G	**G**	12.65634	-1.21069 $\pm$ 0.0903	3.5884	12.88812	-0.97891 $\pm$ 0.2745	2.9007
	G	**E**	12.98506	-0.88197 $\pm$ 0.0951	3.1274	12.9411	-0.92593 $\pm$ 0.2561	2.5475
	G	**U**	12.84359	-1.02344 $\pm$ 0.0898	3.1439	12.97493	-0.8921 $\pm$ 0.2486	2.389
	G	**N**	12.86188	-1.00514 $\pm$ 0.0917	3.1988	12.98342	-0.88361 $\pm$ 0.3125	3.2982
	Breast cancer model: $y_{ij} = \frac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$							

Table (6.7): summary results for fixed effect parameter estimate δ_3 under B_STS and B_GTS

DISTRIBUTIONS	RANDOM-EFFECTS s	ERRORS	B_STS method			B_GTS method		
			MEAN (bias-cb)	BIAS $\pm$I.C$_{\text{Bias}}$	MSE	MEAN	BIAS $\pm$I.C$_{\text{Bias}}$	MSE
	N	**N**	0.850718	0.00206 ± 0.004	0.0041	0.872812	0.02416 ± 0.0098	0.0031
	N	**U**	0.851723	0.00307 ± 0.0042	0.0046	0.869809	0.02115 ± 0.0107	0.0034
	N	**E**	0.85342	0.00476 ± 0.0039	0.004	0.874481	0.02583 ± 0.0115	0.0041
	N	**G**	0.853391	0.00474 ± 0.004	0.0042	0.87267	0.02401 ± 0.0113	0.0039
	U	**U**	0.854344	0.00569 ± 0.0018	0.0009	0.878285	0.02963 ± 0.0073	0.0022
	U	**E**	0.857581	0.00893 ± 0.0038	0.0039	0.877404	0.02875 ± 0.0102	0.0035
	U	**G**	0.857764	0.00911 ± 0.0038	0.0039	0.87834	0.02968 ± 0.0106	0.0038
	U	**N**	0.854173	0.00552 ± 0.0017	0.0008	0.879121	0.03047 ± 0.0046	0.0015
	E	**E**	0.842921	-0.00573 ± 0.0048	0.006	0.870314	0.02166 ± 0.0125	0.0045
	E	**G**	0.840627	-0.00803 ± 0.0049	0.0064	0.867023	0.01837 ± 0.0121	0.0041
	E	**U**	0.829022	-0.01963 ± 0.0046	0.0058	0.866004	0.01735 ± 0.0134	0.0049
	E	**N**	0.837728	-0.01093 ± 0.0048	0.006	0.865611	0.01696 ± 0.0114	0.0036
	G	**G**	0.845256	-0.0034 ± 0.0048	0.0061	0.869303	0.02065 ± 0.0151	0.0063
	G	**E**	0.834708	-0.01395 ± 0.0048	0.0061	0.864867	0.01621 ± 0.0141	0.0054
	G	**U**	0.838181	-0.01047 ± 0.0045	0.0055	0.865989	0.01733 ± 0.0114	0.0036
	G	**N**	0.837779	-0.01088 ± 0.0048	0.0061	0.864207	0.01555 ± 0.0143	0.0055
	Breast cancer model: $y_{ij} = \dfrac{\delta_{1i} \exp(t_{ij} - \delta_{2i})}{1 + \exp((t_{ij} - \delta_{2i}) / \delta_{3i})} + e_{ij}$							

Table (6.8): summary results for fixed effect parameter estimate δ_1 of 1000 simulations

DISTRIBUTIONS	RANDOM-EFFECTS	ERRORS	STS- METHOD			GTS- METHOD		
			MEAN	BIAS $\pm$C.I$_{\text{Bias}}$	MSE	MEAN	BIAS $\pm$C.I$_{\text{Bias}}$	MSE
	N	**N**	5.96901	0.91778 ± 0.08298	2.6331	5.810091	0.75885 ± 0.26437	2.3769
	N	**U**	6.031482	0.980246 ± 0.0817	2.6970	6.069306	1.01807 ± 0.2340	2.4475
	N	**E**	5.987198	0.935962 ± 0.0829	2.6623	5.616546	0.56531 ± 0.27916	2.3278
	N	**G**	5.973237	0.92200 ± 0.08545	2.7491	6.073506	1.02227 ± 0.25274	2.6912
	U	**U**	6.097702	1.04647 ± 0.0809	2.7969	5.757199	0.70596 ± 0.28365	2.5719
	U	**E**	6.086845	1.03560 ± 0.0828	2.8559	6.028003	0.976767 ± 0.2365	2.3947
	U	**G**	6.026001	0.97476 ± 0.0775	2.5100	5.889322	0.83808 ± 0.22648	2.0243
	U	**N**	6.137401	1.08616 ± 0.0831	2.9756	5.847827	0.79659 ± 0.25706	2.3375
	E	**E**	6.393896	1.34266 ± 0.0922	4.0143	6.230262	1.17903 ± 0.32294	4.0777
	E	**G**	6.336349	1.28511 ± 0.0804	3.3306	6.241867	1.19063 ± 0.26915	3.2844
	E	**U**	6.401189	1.3499 ± 0.0906	3.9561	6.225572	1.174336 ± 0.2801	3.4008
	E	**N**	6.386624	1.33538 ± 0.08588	3.7012	6.461853	1.410617 ± 0.3166	4.5727
	G	**G**	6.465001	1.41376 ± 0.08794	4.0097	6.161836	1.11060 ± 0.28018	3.2565
	G	**E**	6.412408	1.36117 ± 0.08757	3.8470	6.406675	1.355439 ± 0.2736	3.7667
	G	**U**	6.498558	1.44732 ± 0.09014	4.2075	6.348943	1.297707 ± 0.2703	3.5666
	G	**N**	6.467415	1.41618 ± 0.0938	4.2950	6.36266	1.311424 ± 0.2738	3.6519
	Breast cancer model: $y_{ij} = \dfrac{\delta_{1i}\exp(t_{ij} - \delta_{2i})}{1 + \exp((t_{ij} - \delta_{2i})/\delta_{3i})} + e_{ij}$							

Table (6.9): summary results for fixed effect parameter estimate δ_2 of 1000 simulations

DISTRIBUTIONS			STS- METHOD			GTS- METHOD		
	RANDOM-EFFECTS	ERRORS	MEAN	BIAS $\pm$C.I$_{\text{Bias}}$	MSE	MEAN	BIAS $\pm$C.I$_{\text{Bias}}$	MSE
	N	**N**	13.4831	-0.383900 $\pm$ 0.0429	0.6262	13.48798	-0.37905 $\pm$ 0.1456	0.6904
	N	**U**	13.47343	-0.393602 $\pm$ 0.0412	0.5973	13.59602	-0.27101 $\pm$ 0.1372	0.5585
	N	**E**	13.4824	-0.384627 $\pm$ 0.04224	0.61200	13.55335	-0.31368 $\pm$ 0.1246	0.49874
	N	**G**	13.46818	-0.39885 $\pm$ 0.04308	0.6419	13.52618	-0.34085 $\pm$ 0.1419	0.6355
	U	**U**	13.48327	-0.38375 $\pm$ 0.0459	0.6957	13.53476	-0.3323 $\pm$ 0.13558	0.5841
	U	**E**	13.51895	-0.34807 $\pm$ 0.0450	0.6489	13.55516	-0.31187 $\pm$ 0.1418	0.6155
	U	**G**	13.47301	-0.39402 $\pm$ 0.0464	0.7145	13.54202	-0.32501 $\pm$ 0.1512	0.6947
	U	**N**	13.50208	-0.36495 $\pm$ 0.0449	0.65872	13.63517	-0.23186 $\pm$ 0.1263	0.4648
	E	**E**	12.8274	-1.03963 $\pm$ 0.0424	1.5480	12.99985	-0.86717 $\pm$ 0.1364	1.2314
	E	**G**	12.86946	-0.99757 $\pm$ 0.0408	1.4270	12.87803	-0.98899 $\pm$ 0.1282	1.4015
	E	**U**	12.86421	-1.00282 $\pm$ 0.0399	1.4194	12.9615	-0.90553 $\pm$ 0.1226	1.2075
	E	**N**	12.81152	-1.0555 $\pm$ 0.04019	1.5343	12.98207	-0.88496 $\pm$ 0.1170	1.1362
	G	**G**	12.76954	-1.09748 $\pm$ 0.04075	1.6364	12.88803	-0.97900 $\pm$ 0.1223	1.3438
	G	**E**	12.87775	-0.98928 $\pm$ 0.04311	1.4620	12.93557	-0.93145 $\pm$ 0.1124	1.1931
	G	**U**	12.84179	-1.02523 $\pm$ 0.03965	1.4598	12.92642	-0.94061 $\pm$ 0.1086	1.1889
	G	**N**	12.86255	-1.00448 $\pm$ 0.04055	1.4365	12.93452	-0.93251 $\pm$ 0.1344	1.3349
	Breast cancer model: $y_{ij} = \dfrac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$							

Table (6.10): summary results for fixed effect parameter estimate δ_3 of 1000 simulations

DISTRIBUTIONS	RANDOM-EFFECTS	ERRORS	STS- METHOD			GTS- METHOD		
			MEAN	BIAS $\pm$C.I$_{\text{Bias}}$	MSE	MEAN	BIAS $\pm$C.I$_{\text{Bias}}$	MSE
	N	**N**	0.8508513	0.002195 ± 0.0017	0.0008	0.8719233	0.02327 ± 0.00489	0.00116
	N	**U**	0.8516477	0.00299 ± 0.00184	0.0009	0.8694072	0.02075 ± 0.00494	0.00106
	N	**E**	0.8535317	0.00487 ± 0.0017	0.0008	0.8728396	0.0242 ± 0.00519	0.00128
	N	**G**	0.8535028	0.00485 ± 0.00179	0.0009	0.8727989	0.02414 ± 0.00476	0.00116
	U	**U**	0.854316	0.00566 ± 0.00173	0.0008	0.8772324	0.02857 ± 0.00479	0.00140
	U	**E**	0.8577915	0.00913 ± 0.00169	0.0008	0.8782721	0.02962 ± 0.00465	0.00143
	U	**G**	0.8577082	0.00905 ± 0.0017	0.0008	0.8762791	0.02762 ± 0.00455	0.00129
	U	**N**	0.8541782	0.00552 ± 0.0017	0.0008	0.8782829	0.02963 ± 0.00467	0.00144
	E	**E**	0.8428415	-0.00581 ± 0.00216	0.0013	0.8688131	0.02016 ± 0.00584	0.00128
	E	**G**	0.8403665	-0.00828 ± 0.00218	0.0013	0.865968	0.01731 ± 0.00540	0.00105
	E	**U**	0.8345638	-0.01409 ± 0.00205	0.0013	0.8654588	0.01680 ± 0.00627	0.00130
	E	**N**	0.8376156	-0.01104 ± 0.0021	0.0013	0.8643507	0.015695 ± 0.0054	0.00099
	G	**G**	0.8421123	-0.00654 ± 0.00214	0.0012	0.8672022	0.01855 ± 0.00655	0.00145
	G	**E**	0.8385051	-0.01015 ± 0.00216	0.0013	0.8660464	0.01739 ± 0.00584	0.00118
	G	**U**	0.8381646	-0.01049 ± 0.00204	0.0012	0.8646812	0.01603 ± 0.00525	0.00097
	G	**N**	0.8377883	-0.01087 ± 0.00212	0.0013	0.8634392	0.01478 ± 0.00611	0.00118
	Breast cancer model: $y_{ij} = \dfrac{\delta_{1i}\exp(t_{ij}-\delta_{2i})}{1+\exp((t_{ij}-\delta_{2i})/\delta_{3i})} + e_{ij}$							

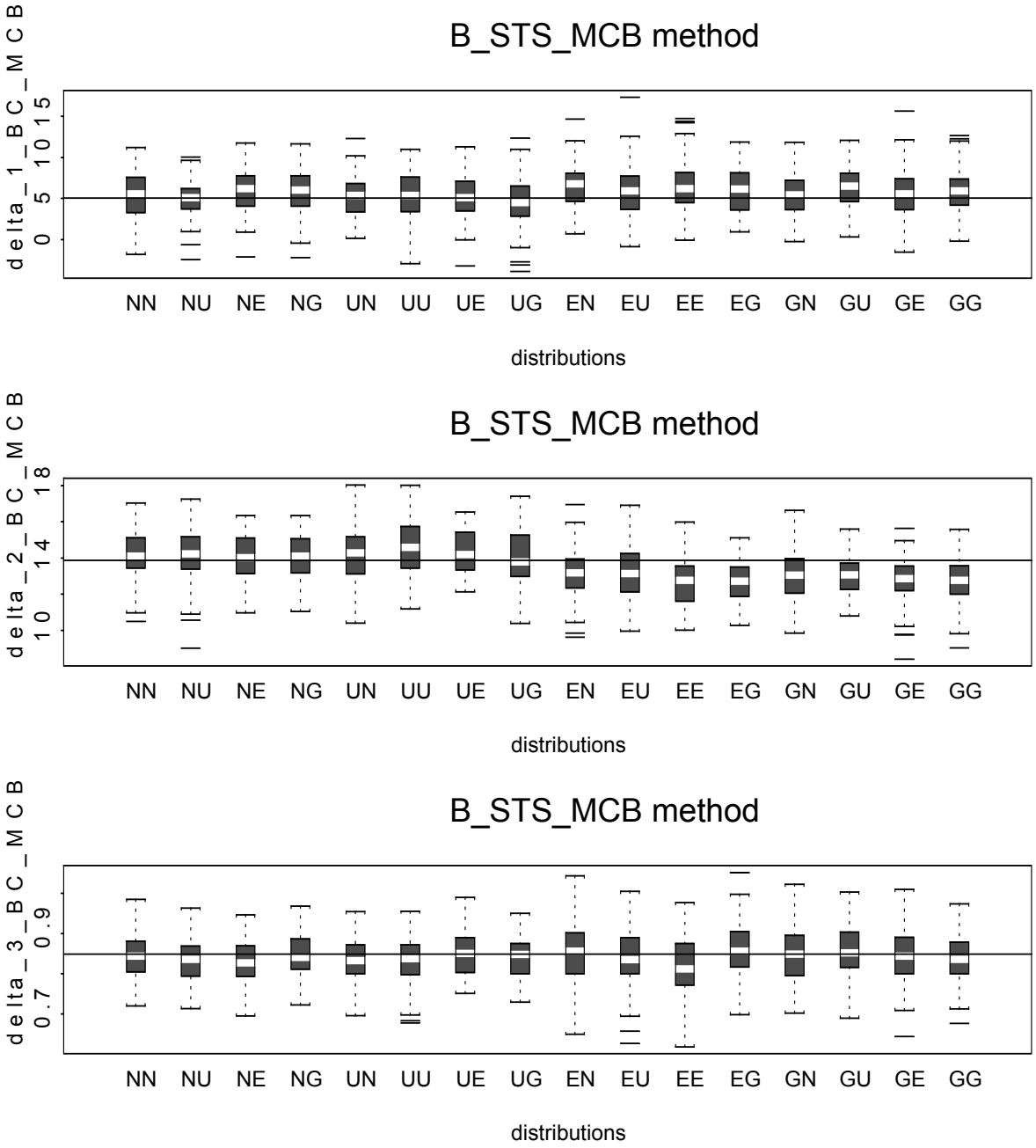


Figure 6.1: Box-plots of fixed effects under B_STS_MCB methods

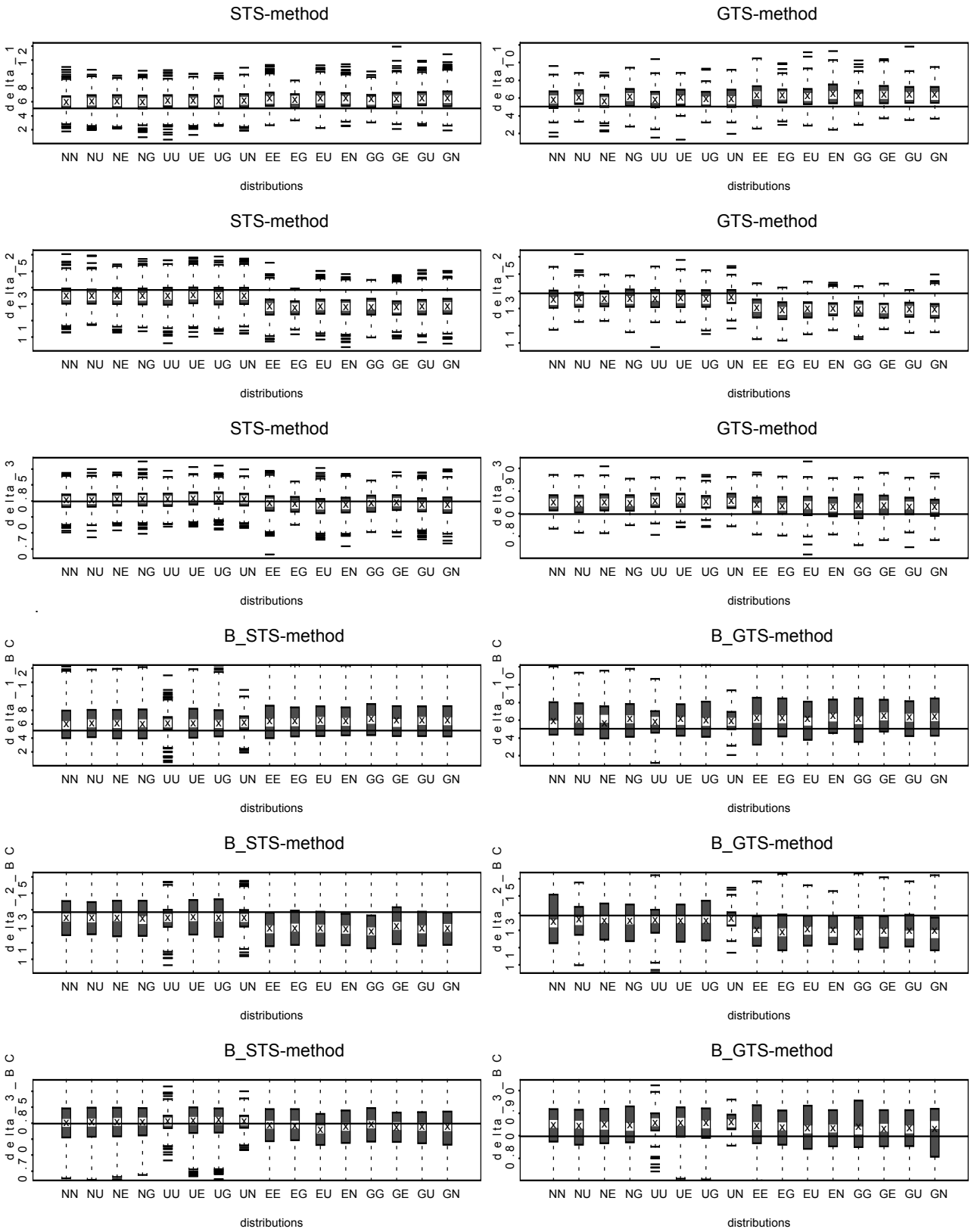


Figure 6.2: Box-plots of fixed effects under STS, GTS, B_STS and B_GTS methods

The next paragraphs are devoted to evaluate the performance of the bootstrap and asymptotic GTS confidence intervals for the fixed parameters, under the usual simulation scenarios. The different confidence intervals are compared with respect to their coverage probabilities, their length and their capability to achieve equitailedness. Results from these simulations are presented in Tables (6.11) to (6.19). In these tables, the following notation will be used to summarize the results of the simulation experiments: let $C.I$, lb and up be general symbols (for any model parameter) to represent, respectively, the observed percentage of simulation replicates with the true parameter value included in the corresponding confidence interval, with the true parameter value smaller than the lower bound of the interval, and with the true parameter value higher than the upper bound of the interval. The super-index indicates the estimation two-stage method and the sub-index indicates the type of interval –percentile bootstrap, bootstrap-t or GTS-asymptotic. For example, $C.I_P^{B\_STS\_MCB}$ denotes that the percentage of times that the true parameter value is inside the generated bootstrap percentile intervals based on B_STS_MCB bootstrap resampling, $lb_P^{B\_STS\_MCB}$ denotes the percentage of times that the true parameter is at left of the lower bound of the percentile interval based in B_STS_MCB rersampling, and $up_P^{B\_STS\_MCB}$ denotes the corresponding percentage of simulation replicates with true parameter higher than the upper bound of the interval. The "Average Width (percentile; bootstrap-t)" contains the mean length of the symmetry bootstrap-t confidence interval (bootstrap-t) and the bootstrap percentile interval (percentile).

Table (6.11): Estimated coverage and precision of the bootstrap confidence intervals for the δ_1 fixed parameter of Breast cancer model (B_STS_MCB method)

Random effects/ Residuals distribution	$(lb_P^{B\_STS\_MCB};\ lb_{Sm}^{B\_STS\_MCB})$	$(C.I_P^{B\_STS\_MCB};C.I_{Sm}^{B\_STS\_MCB})$	$(up_P^{B\_STS\_MCB};\ up_{Sm}^{B\_STS\_MCB})$	Average width (percentile; bootstrap-t)
NN	(0%; 0%)	(95%; 96%)	(5%; 4%)	(3.4934956; 3.16134453)
NU	(0%; 0%)	(94%; 96%)	(6%; 4%)	(4.4059059; 5.8461125)
NE	(0%; 0%)	(93%; 94%)	(7%; 6%)	(3.4732914; 3.03276862)
NG	(0%; 0%)	(94%; 95%)	(6%; 5%)	(3.4553204; 3.06517526)
UU	(0%; 0%)	(94%; 95%)	(6%; 5%)	(3.317098; 2.98247843)
UE	(0%; 0%)	(96%; 94%)	(4%; 6%)	(3.2460555; 3.0370250)
UG	(0%; 0%)	(97%; 96)	(3%; 4%)	(3.4968543; 3.07762758)
UN	(0%; 0%)	(94%; 94%)	(6%; 6%)	(3.1661659; 2.98853131)
EE	(0%; 0%)	(94%; 95%)	(6%; 5%)	(3.4736700; 3.07874738)
EG	(0%; 0%)	(98%; 94%)	(2%; 6%)	(3.2626724; 3.15658458)
EU	(0%; 0%)	(94%; 95%)	(6%; 5%)	(3.5208397; 3.0683514)
EN	(0%; 0%)	(97% ; 95%)	(3%; 5%)	(3.2014165; 3.00282160)
GG	(3%; 0%)	(89%; 92%)	(8%; 8%)	(4.1515815; 2.20053225)
GE	(0%; 0%)	(94%; 95%)	(6%; 5%)	(3.519690; 3.06777059)
GU	(0%; 0%)	(94%; 95%)	(6%; 5%)	(3.1738699; 2.98424458)
GN	(0%; 0%)	(94%; 95%)	(6%; 5%)	(3.2439030; 2.95452781)

Table (6.12): Estimated coverage and precision of the bootstrap confidence intervals for the δ_2 fixed parameter of Breast cancer model (B_STS_MCB method)				
Random effects/ Residuals distribution	$(lb_P^{B\_STS\_MCB} ;\ lb_{Sm}^{B\_STS\_MCB})$	$(C.I_P^{B\_STS\_MCB} ; C.I_{Sm}^{B\_STS\_MCB})$	$(up_P^{B\_STS\_MCB} ;\ up_{Sm}^{B\_STS\_MCB})$	Average width (percentile; bootstrap-t)
NN	(0%; 0%)	(80%; 97%)	(20%; 3%)	(4.6289267; 6.34110317)
NU	(0%; 0%)	(76%; 95%)	(24%; 5%)	(8.4194682; 6.3568938)
NE	(0%; 0%)	(84%; 98%)	(16%; 2%)	(4.6674056; 6.37701684)
NG	(0%; 0%)	(83%; 96%)	(17%; 4%)	(4.6459570; 6.33579901)
UU	(0%; 0%)	(79%; 96%)	(21%; 4%)	(4.649353; 6.37075592)
UE	(0%; 0%)	(82%; 99%)	(18%; 1%)	(4.6426161; 6.36986513)
UG	(0%; 0%)	(75%; 96%)	(25%; 4%)	(4.6539194; 6.35848184)
UN	(0%; 0%)	(81%; 98%)	(19%; 2%)	(4.6445326; 6.3594699)
EE	(0%; 0%)	(74%; 95)	(26%; 5%)	(4.6507779; 6.36786366)
EG	(0%; 0%)	(73%; 96%)	(27%; 4%)	(4.6830014; 6.3719725)
EU	(0%; 0%)	(83% ; 98%)	(17%; 2%)	(4.6315303; 6.35586926)
EN	(0%; 0%)	(77% ; 98%)	(23% ; 2%)	(4.6822401; 6.38574007)
GG	(0%; 0%)	(68%; 93%)	(32%; 7%)	(5.3498605; 5.52898469)
GE	(0%; 0%)	(83%; 95%)	(17%; 5%)	(4.620858; 6.35932324)
GU	(0%; 0%)	(58%; 98%)	(42%; 2%)	(4.6836912; 6.29007102)
GN	(0%; 0%)	(71%; 96%)	(29%; 4%)	(4.6663354; 6.32600346)

Table (6.13): Estimated coverage and precision of the bootstrap confidence intervals for the δ_3 fixed parameter of Breast cancer model (B_STS_MCB method)				
Random effects/ Residuals distribution	$(lb_P^{B\_STS\_MCB};\ lb_{Sm}^{B\_STS\_MCB})$	$(C.I_P^{B\_STS\_MCB};C.I_{Sm}^{B\_STS\_MCB})$	$(up_P^{B\_STS\_MCB};\ up_{Sm}^{B\_STS\_MCB})$	Average width (percentile; bootstrap-t)
NN	(2%; 2%)	(92%; 94%)	(6%; 4%)	(0.4943647; 0.09112892)
NU	(1%; 1%)	(93%; 95%)	(6%; 5%)	(0.6187871; 0.1947818)
NE	(2%; 2%)	(92%; 93%)	(6%; 5%)	(0.4961874; 0.0915328)
NG	(2%; 2%)	(92%; 92%)	(8%; 8%)	(0.4910545; 0.0903915)
UU	(0%; 1%)	(93.9%; 96%)	(6.1%; 3%)	(0.5017105; 0.09048858)
UE	(2%; 2%)	(92%; 92%)	(6%; 6%)	(0.5017595; 0.09007225)
UG	(3%; 3%)	(92%; 93%)	(5%; 4%)	(0.5054963; 0.09130761)
UN	(2%; 2%)	(92%; 92%)	(6 %; 6%)	(0.5006815; 0.09011907)
EE	(2%; 2%)	(92%; 92%)	(6%; 6%)	(0.4963241; 0.09254227)
EG	(2%; 2%)	(92%; 93%)	(6% ; 5%)	(0.5097649; 0.0907190)
EU	(2%; 2%)	(92%; 93%)	(6% ; 5%)	(0.4956226; 0.09100182)
EN	(3%; 3%)	(92%; 92%)	(5% ; 5%)	(0.5002988; 0.09046101)
GG	(3%; 0%)	(89%; 90%)	(8%; 10%)	(0.5334585; 0.05954418)
GE	(2%; 2%)	(92%; 93%)	(6%; 5%)	(0.5007344; 0.09046957)
GU	(2%; 2%)	(92%; 94%)	(6%; 4%)	(0.5053104; 0.08791348)
GN	(2%; 2%)	(92%; 93%)	(6%; 5%)	(0.4985269; 0.08962327)

Table (6.14): Estimated coverage and precision of the bootstrap confidence intervals for the δ_1 fixed parameter of Breast cancer model (B_STS method)				
Random effects/ Residuals distribution	$(lb_P^{B\_STS} ;\ lb_{Sm}^{B\_STS})$	$(C.I_P^{B\_STS} ; C.I_{Sm}^{B\_STS})$	$(up_P^{B\_STS} ;\ up_{Sm}^{B\_STS})$	Average width (percentile; bootstrap-t)
NN	(11.3%; 9.7%)	(88%; 89.9%)	(0.7%; 0.5%)	(5.062; 5.402)
NU	(13.6%; 10.3%)	(85.9%; 89.4%)	(0.5%; 0.3%)	(5.059; 5.399)
NE	(12.8%; 10.1%)	(86.8%: 89.8%)	(0.4%; 0.1%)	(5.103; 5.443)
NG	(11.2%; 12%)	(85.4%; 87.5%)	(0.8%; 0.5%)	(5.077; 5.412)
UU	(13.4%; 11.9%)	(86%; 87.6%)	(0.6%; 0.5%)	(5.044; 5.385)
UE	(15.2%; 12.8%)	(84.1%; 86.7%)	(0.7%; 0.5%)	(5.059; 5.403)
UG	(11.2%; 9.9%)	(88.6%; 89.9%)	(0.2%; 0.2%)	(5.121; 5.466)
UN	(14.6%; 12.4%)	(84.7%; 87.1%)	(0.7%; 0.5%)	(5.078; 5.424)
EE	(14.7%; 6.6%)	(84.5%; 92.7%)	(0.8%; 0.7%)	(5.4539; 6.45)
EG	(14.7%; 6.1%)	(84.7%; 93.5%)	(0.6%; 0.4%)	(5.4819; 6.53)
EU	(16.1%; 7.8%)	(83.2%; 91.6%)	(0.7%; 0.6%)	(5.364; 6.411)
EN	(13.4%; 4.9%)	(86.3%; 94.9%)	(0.3%; 0.2%)	(5.425; 6.453)
GG	(15.7%; 6.2%)	(83.5%; 93%)	(0.8%; 0.8%)	(5.477; 6.460)
GE	(13.9%; 5.9%)	(85.7%; 93.9%)	(0.4%; 0.2%)	(5.397; 6.407)
GU	(15.5%; 8%)	(83.9%; 91.4%)	(0.6%; 0.6%)	(5.4334; 6.55)
GN	(17.3%; 8%)	(82%; 91.3%)	(0.7%; 0.7%)	(5.4265; 6.493)

Table (6.15): Estimated coverage and precision of the bootstrap confidence intervals for the δ_2 fixed parameter of Breast cancer model (B_STS method)				
Random effects/ Residuals distribution	$(lb_P^{B\_STS}; lb_{Sm}^{B\_STS})$	$(C.I_P^{B\_STS}; C.I_{Sm}^{B\_STS})$	$(up_P^{B\_STS}; up_{Sm}^{B\_STS})$	Average width (percentile; bootstrap-t)
NN	(0.9%; 0.8%)	(90.7%; 92.2%)	(8.4%; 7%)	(2.613; 2.795)
NU	(0.6%; 0.5%)	(60.8%; 93.4%)	(8.6%; 6.1%)	(2.646; 2.828)
NE	(0.9%; 0.6%)	(89.9%; 93.1%)	(9.2%; 6.3%)	(2.614; 2.792)
NG	(0.9%; 0.5%)	(89.4%; 92.8%)	(9.7%; 6.7%)	(2.623; 2.805)
UU	(0.7%; 0.7%)	(91.2%; 92.6%)	(8.1%; 6.7%)	(2.815; 3.002)
UE	(1.4%; 1.3%)	(90.6%; 92.4%)	(8%; 6.3%)	(2.810; 2.998)
UG	(0.9%; 0.6%)	(89.9%; 92%)	(9.2%; 7.4%)	(2.807; 2.995)
UN	(1.3%; 1.1%)	(91.3%; 93.2%)	(7.4%; 5.7%)	(2.827; 3.021)
EE	(0.1%; 0%)	(59.3%; 64.4%)	(40.6%; 35.6%)	(2.506; 2.723)
EG	(0%; 0%)	(57.9%; 63.4%)	(42.1%; 36.6%)	(2.488; 2.699)
EU	(0.1%; 0%)	(64%; 67.8%)	(35.9%; 32.2%)	(2.514; 2.728)
EN	(0.1%; 0%)	(60.3%; 65.8%)	(39.6%; 34.2%)	(2.482; 2.694)
GG	(0%; 0%)	(57.9%; 63.3%)	(42.1%; 36.7%)	(2.502; 2.719)
GE	(0%; 0%)	(58.8%; 64.2%)	(41.2%; 35.8%)	(2.473; 2.682)
GU	(0.2%; 0%)	(62.4%; 67%)	(37.4%; 33%)	(2.487; 2.699)
GN	(0.1%; 0%)	(63.1%; 67.9%)	(36.8%; 32.1%)	(2.472; 2.681)

Table (6.16): Estimated coverage and precision of the bootstrap confidence intervals for the δ_3 fixed parameter of Breast cancer model (B_STS method)				
Random effects/ Residuals distribution	($lb_P^{B\_STS}$; $lb_{Sm}^{B\_STS}$)	($C.I_P^{B\_STS}$; $C.I_{Sm}^{B\_STS}$)	($up_P^{B\_STS}$; $up_{Sm}^{B\_STS}$)	Average width (percentile; bootstrap-t)
NN	(4.7%; 4%)	(93.5%; 95.2%)	(1.8%; 0.8%)	(0.108; 0.119)
NU	(6.4%; 5.8%)	(90.4%; 92.3%)	(3.2%; 1.9%)	(0.106; 0.116)
NE	(6.4%; 5%)	(92.2%; 94.2%)	(1.4%; 0.8%)	(0.107; 0.118)
NG	(7.2%; 5.6%)	(90.4%; 93.3%)	(12.4%; 1.1%)	(0.106; 0.116)
UU	(6.4%; 5.2%)	(92.1%; 94.2%)	(1.5%; 0.6%)	(0.108; 0.118)
UE	(8.4%; 6.9%)	(90.4%; 92.8%)	(1.2%; 0.3%)	(0.106; 0.116)
UG	(6.5%; 5.3%)	(92.6%; 94.3%)	(0.9%; 0.4%)	(0.106; 0.116)
UN	(5.2%; 4.1%)	(92.8%; 95.4%)	(2%; 0.5%)	(0.108; 0.118)
EE	(3.7%; 2.9%)	(91.2%; 94.1%)	(5.1%; 3%)	(0.125; 0.1377)
EG	(2.5%; 1.9%)	(93.1%; 96.2%)	(4.4%; 1.9%)	(0.127; 0.1398)
EU	(1.5%; 0.8%)	(91.8%; 95.1%)	(6.7%; 4.1%)	(0.129; 0.142)
EN	(1.9%; 1%)	(92.4%; 96.7%)	(5.7%; 2.3%)	(0.130; 0.144)
GG	(2.8%; 1.4%)	(91.4%; 95.3%)	(5.8%; 3.3%)	(0.126; 0.139)
GE	(2.6%; 1.5%)	(92.8%; 96.2%)	(4.6%; 2.3%)	(0.126; 0.138)
GU	(2%; 1.2%)	(92.9%; 96.4%)	(5.1%; 2.4%)	(0.1268; 0.139)
GN	(2.2%; 1.1%)	(93.3%; 96.6%)	(4.5%; 2.3%)	(0.1304; 0.143)

Table (6.17): Estimated coverage and precision of the bootstrap confidence intervals for the δ_1 fixed parameter of Breast cancer model (B_GTS/GTS methods)				
Random effects/ Residuals distribution	$(lb_P^{B\_GTS}; lb_{Sm}^{B\_GTS}; lb_{Asym}^{GTS})$	$(C.I_P^{B\_GTS}; C.I_{Sm}^{B\_GTS}; C.I_{Asym}^{GTS})$	$(up_P^{B\_GTS}; up_{Sm}^{B\_GTS}; up_{Asym}^{GTS})$	Average width (percentile; bootstrap-t; asymptotic)
NN	(7%; 7%; 1%)	(91%; 91%; 99%)	(2%; 2%,0%)	(4.996; 5.316; 6.901)
NU	(16%; 11%; 6%)	(84%; 89%; 94%)	(0%; 0%; 0%)	(5.127; 5.442; 7.088)
NE	(8%; 6%; 3%)	(90%; 93%; 97%)	(2%; 1%; 0%)	(5.245; 5.593; 7.321)
NG	(15%; 12%; 6%)	(85%; 88%; 94%)	(0%; 0%; 0%)	(5.116; 5.470; 7.0413)
UU	(13%; 9%; 6.1%)	(85%; 89%; 92.9%)	(2%; 2%; 1%)	(4.934; 5.296; 6.578)
UE	(14%; 11%; 6%)	(85%; 88%; 93%)	(1%; 1%; 1%)	(4.946; 5.277; 6.560)
UG	(10%; 9%; 4%)	(90%; 91%; 96%)	(0%; 0%; 0%)	(5.103; 5.456; 6.914)
UN	(9%; 7%; 5%)	(90%; 92%; 95%)	(1%; 1%; 0%)	(5.135; 5.494; 7.060)
EE	(16%; 8%; 2%)	(82%; 91%; 96%)	(2%; 1%; 2%)	(5.462; 6.4223; 8.207)
EG	(12%; 4%; 0%)	(88%; 96%; 1%)	(0%; 0%; 0%)	(5.441; 6.488; 8.361)
EU	(13%; 10%; 3%)	(86%; 89%; 96%)	(1%; 1%; 1%)	(5.290; 6.278; 8.249)
EN	(18%; 5%; 0%)	(81%; 94%; 99%)	(1%; 1%; 1%)	(5.737; 6.953; 9.371)
GG	(13%; 10%; 0%)	(86%; 89%; 99%)	(1%; 1%; 1%)	(5.195; 5.994; 7.714)
GE	(20%; 8%; 2%)	(80%; 92%; 98%)	(0%; 0%; 0%)	(5.383; 6.449; 8.010)
GU	(14%; 8%; 0%)	(86%; 92%; 99%)	(0%; 0%; 1%)	(5.444; 6.510; 8.434)
GN	(19%; 9%; 1%)	(81%; 91%; 98%)	(0%; 0%; 1%)	(5.417; 6.423; 8.034)

Table (6.18): Estimated coverage and precision of the bootstrap confidence intervals for the δ_2 fixed parameter of Breast cancer model (B_GTS/GTS methods)				
Random effects/ Residuals distribution	$(lb_P^{B\_GTS}; lb_{Sm}^{B\_GTS}; lb_{Asym}^{GTS})$	$(C.I_P^{B\_GTS}; C.I_{Sm}^{B\_GTS}; C.I_{Asym}^{GTS})$	$(up_P^{B\_GTS}; up_{Sm}^{B\_GTS}; up_{Asym}^{GTS})$	Average width (percentile; bootstrap-t; asymptotic)
NN	(2%; 2%; 6%)	(87%; 91%; 71%)	(11%; 7%; 23%)	(2.645; 2.842; 1.890)
NU	(2%; 1%; 7%)	(94%; 97%; 79%)	(4%; 2%; 14%)	(2.698; 2.871; 1.908)
NE	(2%; 2%; 5%)	(84%; 87%; 75%)	(14%; 11%; 20%)	(2.658; 2.844; 1.857)
NG	(1%; 0%; 4%)	(91%; 94%; 79%)	(8%; 6%; 17%)	(2.655; 2.837; 1.977)
UU	(1%; 1%; 2%)	(89%; 92%; 82.8%)	(10%; 7%; 15.2%)	(2.779; 2.967; 2.058)
UE	(2%; 2%; 5%)	(88%; 90%; 83%)	(10%; 8%; 12%)	(2.774; 2.950; 2.059)
UG	(1%; 0%; 5%)	(91%; 93%; 79%)	(8%; 7%; 16%)	(2.803; 2.977; 2.099)
UN	(2%; 2%; 7%)	(93%; 95%; 82%)	(5%; 3%; 11%)	(2.897; 3.111; 2.198)
EE	(0%; 0%; 0%)	(65%; 68%; 47.5%)	(35%; 32%; 52.5%)	(2.504; 2.704; 1.660)
EG	(0%; 0%; 0%)	(52%; 63%; 41.4%)	(48%; 37%; 58.6%)	(2.565; 2.777; 1.745)
EU	(0%; 0%; 0%)	(66%; 69%; 47.5%)	(34%; 31%; 25.5%)	(2.521; 2.735; 1.779)
EN	(0%; 0%; 0%)	(67%; 74%; 46.5%)	(33%; 26%; 53.5%)	(2.529; 2.737; 1.720)
GG	(0%; 0%; 0%)	(62%; 69%; 45%)	(38%; 31%; 55%)	(2.511; 2.728; 1.689)
GE	(0%; 0%; 0%)	64%; 68%; 40.4%)	(36%; 32%; 59.6%)	(2.465; 2.665; 1.567)
GU	(0%; 0%; 0%)	(62%; 67%; 43.4%)	(38%; 33%; 56.6%)	(2.473; 2.677; 1.647)
GN	(1%; 0%; 1%)	(57%; 61%; 35.4%)	(42%; 39%; 63.6%)	(2.415; 2.617; 1.539)

Table (6.19): Estimated coverage and precision of the bootstrap confidence intervals for the δ_3 fixed parameter of Breast cancer model (B_GTS/GTS methods)				
Random effects/ Residuals distribution	$(lb_P^{B\_GTS} ; lb_{Sm}^{B\_GTS} ; lb_{Asym}^{GTS})$	$(C.I_P^{B\_GTS} ; C.I_{Sm}^{B\_GTS} ; C.I_{Asym}^{GTS})$	$(up_P^{B\_GTS} ; up_{Sm}^{B\_GTS} ; up_{Asym}^{GTS})$	Average width (percentile; bootstrap-t; asymptotic)
NN	(6%; 5%; 78%)	(93%; 95%; 2%)	(1%, 0%, 20%)	(0.109; 0.120; 0.003)
NU	(6%; 6%; 78%)	(92%; 93%; 3%)	(2%; 1%; 21%)	(0.109; 0.118; 0.003)
NE	(10%; 9%; 80%)	(88%; 90%; 6%)	(2%; 1%; 14%)	(0.110; 0.120; 0.003)
NG	(11%; 9%; 81.8%)	(88%; 91%; 3%)	(1%; 0%; 15.2%)	(0.105; 0.116; 0.003)
UU	(10%; 9%; 85.9%)	(90%; 91%; 1%)	(0%; 0%; 13.1%)	(0.108; 0.119; 0.003)
UE	(9%; 5%; 85.8%)	(90%; 95%; 2%)	(1%; 0%; 12.2%)	(0.104; 0.113; 0.003)
UG	(7%; 6%; 86.8%)	(92%; 94%; 3%)	(1%; 0%; 10.2%)	(0.105; 0.115; 0.003)
UN	(7%; 6%; 88%)	(92%; 93%; 3%)	(1%; 1%; 9%)	(0.106; 0.117; 0.003)
EE	(4%; 3%; 74.7%)	(92%; 94%; 7.1%)	(4%; 3%; 18.2%)	(0.121; 0.132; 0.004)
EG	(0%; 0%; 73.7%)	(93%; 99%; 3%)	(7%; 1%; 23.2%)	(0.129; 0.144; 0.004)
EU	(1%; 1%; 66.7%)	(93%; 84%; 8%)	(6%; 5%; 52.5%)	(0.127; 0.139; 0.004)
EN	(2%; 2%; 65.7%)	(94%; 97%; 5.1%)	(4%; 1%; 29.3%)	(0.132; 0.149; 0.004)
GG	(2%; 2%; 67%)	(94%; 96%; 5%)	(4%; 2%; 28%)	(0.130; 0.142; 0.004)
GE	(3%; 1%; 72.7%)	(89%; 97%; 5.1%)	(8%; 2%; 22.2%)	(0.124; 0.136; 0.004)
GU	(0%; 0%; 72.7%)	(97%; 97%; 4%)	(3%; 3%; 23.2%)	(0.126; 0.138; 0.004)
GN	(3%; 3%; 69.7%)	(92%; 94%; 4%)	(5%; 3%; 26.3%)	(0.1307; 0.142; 0.004)

The coverage of bootstrap confidence intervals at a 95% nominal level for the B_STS_MCB method is shown in Tables (6.11) to (6.13). As shown in the columns noted ($C.I_P^{B\_STS\_MCB}; C.I_{Sm}^{B\_STS\_MCB}$), the robustness of the confidence intervals and hence of the robustness of B_STS_MCB method depend on the concrete parameter of the model. They are quite robust for the δ_1 parameter, with coverage ranging from 92% and 96% for the bootstrap-t intervals and between 89% and 98% for percentile intervals. The average width of these confidence intervals is remarkably large, but smaller than for those other methods discussed below for the same parameter. The equitailedeness (in the sense that the non coverage cases are nearly symmetrically distributed) is not improved as much as expected.

Table (6.12) display the simulation estimated true coverage for percentile and bootstrap-t confidence intervals with 95% nominal coverage, for the δ_2 parameter, jointly with their average width of these confidence intervals. As can be seen, the bootstrap-t show some tendency to overestimate the true parameter value, with coverages ranging from 93% (GG) and 99% (UE) with 97% for the NN case, while the percentile interval perform worse in all cases for this parameter and demonstrate a poor degrees of robustness. The coverage probabilities range from 58% under the combinations (GU) and 84% under (NE) with 80% for NN case. Also for this parameter, all these intervals are not equitailed as was expected.

As shown in the Table (6.13), the confidence intervals, for the parameter δ_3, tend to attain their nominal coverages in a majority of cases, but break down in the GG combination 90% for bootstrap-t interval and 89% for percentile interval. They are acceptably robust for this parameter, with coverages ranging from 90%/89% and 96%/93.9% for GG and UU respectively. The average widths of these confidence intervals are near of 0.09 for bootstrap-t type while are remarkably large for the percentile intervals (about 0.5). In all cases, the intervals are not symmetrically distributed (i.e not equitailed). These facts agree with the work of investigation along this line under linear mixed designed by Sanches, J.A., Department of Bio-statistics University of Barcelona. This provides a confirmation of the intuitive notion that the bootstrap intervals give a better results under symmetric

distributions such as normal and uniform and not satisfactory when the asymmetry distributions are used such as exponential and gamma.

In general terms it can be concluded that, the B_STS_MCB resampling method provides acceptably robust results provided it is used with an adequate inferential procedure.

The naive B_STS (or B_GTS) resampling procedure is, in general, not adequate. Table (6.14) shows that, for the δ_1 parameter, the observed coverages of the $C.I_P^{B-STS}$ intervals oscillate between 82% and 88.6% and between 86.7% and 94.9% for $C.I_{Sm}^{B-STS}$. So, it seems that the symmetric bootstrap-t intervals under B_STS method tend to perform better than the percentile intervals, but in any case they are not robust in front of departures from the normality assumption. For this parameter, the results of the percentile and symmetric bootstrap-t intervals under B_GTS method tend to be similar to those for B_STS method, but the asymptotic confidence intervals based on the GTS-method compete in these cases well with the other methods (B_STS and B_GTS methods). The best results at a nominal level of 95% are observed when the random effects and/or errors have a uniform law. Table (6.15) shows that, for the δ_2 parameter, the symmetric bootstrap-t intervals give good results in all cases, and perform well compared with the bootstrap percentile intervals and asymptotic intervals (based on GTS method). Thus B_STS and B_GTS are relatively robust for this parameter. Table (6.16) lists the coverage probabilities for the δ_3 parameter. The symmetric bootstrap and percentile intervals give quite good results at nominal 95% level for the two methods, B_STS and B_GTS, and always the symmetric bootstrap-t gives better results than the others intervals. Surprisingly, the asymptotic intervals based on the GTS-method give very bad results, due to their systematically small estimations of the standard error of the estimator of this parameter.

6.1.2. Results for random components

In the next paragraphs we describe the results on the covariance components, considering how the quality of estimation is affected by deviations from the normality assumption when the

correct model is specified. Figure 6.3 contains boxpots for the variance-covariance parameter estimates from the 16 simulation scenarios, under the B_STS_MCB procedure. In these plots, the first "column" corresponds to the "direct" bootstrap bias-corrected estimates of the covariance matrix, $\hat{D}_{TS,BC} = 2\hat{D}_{TS} - B^{-1}\sum \hat{D}^{*,b}_{TS}$ -i.e. not based on parameterisations preserving the semi definite positive character of the covariance matrices. The labels in the y-axis indicate the concrete parameter and estimate, for example "D_BC_MCB_11" denotes the estimate of the first element in the diagonal of the $\hat{D}_{TS,BC}$ estimator of D and so on. The results under B_STS_MCB indicate, as one might expect, the presence of negative elements in the diagonal of $\hat{D}_{TS,BC}$ (the second and third diagonal elements) confirming that the result of this bias correction procedure is not always a covariance matrix. All of the simulation scenarios exhibit large bias, and B_STS_MCB appears to perform poorly as a bias correction procedure. The variance and covariance estimates of the random parameters become more variable when the random effects deviate from the normality assumptions. For these reasons, an additional simulation study was carried out to investigate the effects of constraining $\hat{D}_{TS,BC}$ to be a non-negative definite matrix by parameterising it in terms of the Cholesky and log-Cholesky decompositions, already mentioned at chapter 5. The results from this analysis are displayed in the second and third "column" in Figure 6.3. For the Cholesky decomposition method, the nonnegative definite constraint does not improve the bias correction of the estimates of (D_{11}, D_{33}, D_{13} and D_{23}) while it seems to reduce the bias for the remaining parameters. When the log-Cholesky decomposition is used, the resulting estimates of the variance components exhibit small bias and small variability for all cases. In any case, the Markov chain bootstrap procedure under the STS-method seems to provide better results than the other resampling procedures, as is discussed in the next paragraphs.

Figure 6.4 displays the results for the variance components of a "direct" bias-corrected bootstrap procedure based on the B_STS and B_GTS resampling methods. We first remark the surprising values obtained in the UE series for both methods. All the simulations exhibit little bias and B_STS and B_GTS methods appear to perform relatively well, except for

$D_{33}\_BC\_MCB$ with some negative values confirming the production of non positive definite variance-covariance matrix estimates. Again, the parameterisation in terms of Cholesky or log-Cholesky tries to correct this drawback. They give robust results under the B_STS method, except for D_BC22 in which the results depend on the concrete distribution of the random effects. Additionally, the Figure 6.7 display similar results for STS and GTS estimates, directly, without bias correction, of the terms of the random factors covariance matrix, D, and for the variance the residuals, σ^2. We first remark that the presence of some outliers for D_{33}, D_{23} and σ^2 under STS method perturb the comparison. As expected, both methods exhibit a bias with height bias when the asymmetric distributions are applied in the simulation runs except for D_{11}. Quality of estimation of, D, tend to perform equal in which depend in random effects more than residuals. The estimate of σ^2 become less precise under STS method and provides a noticeable bias improvement over GTS method, and tend to depend on the residual more than the random factors.

Finally, the results obtained under the B_GTS tend to be similar to these of B_STS method, with the exception that the magnitude of the variability of the variance component parameters is larger than under the B_STS method.

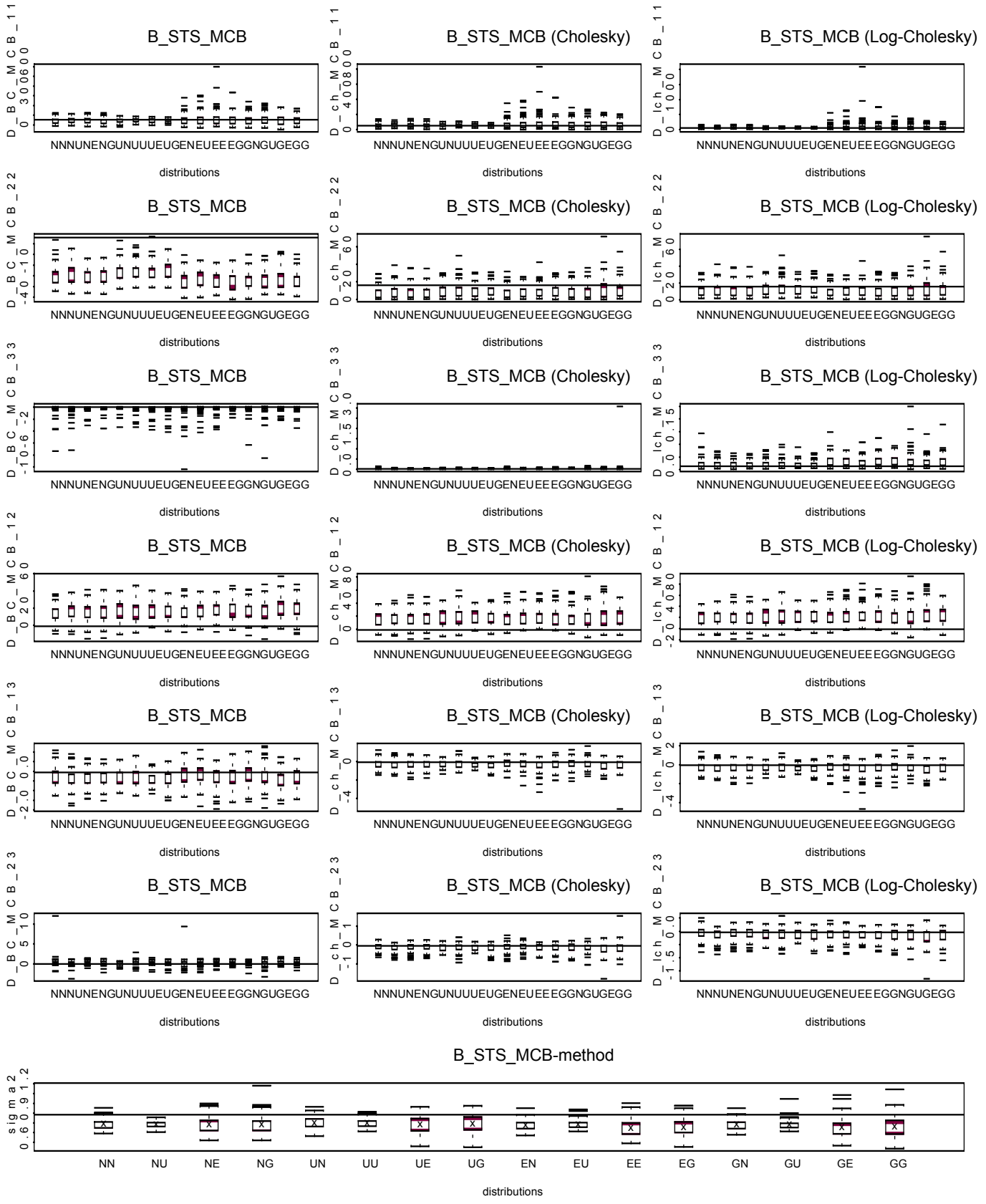


Figure 6.3: Box-plots of random components of a direct bootstrap corrected version, Cholesky parameterization and log- Cholesky parameterization under B_STS_MCB method.

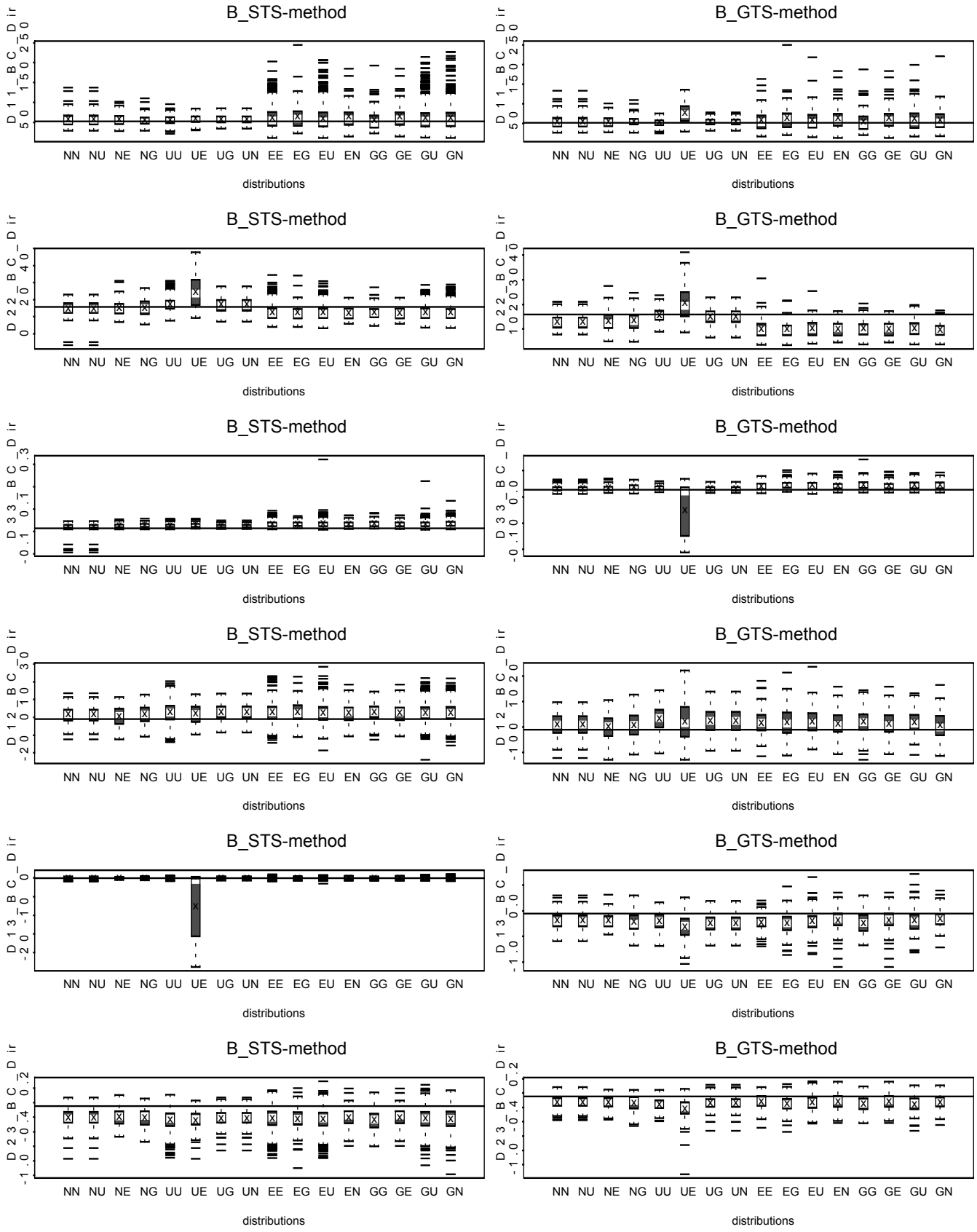


Figure 6.4: Box-plots of random components of a direct bootstrap corrected version

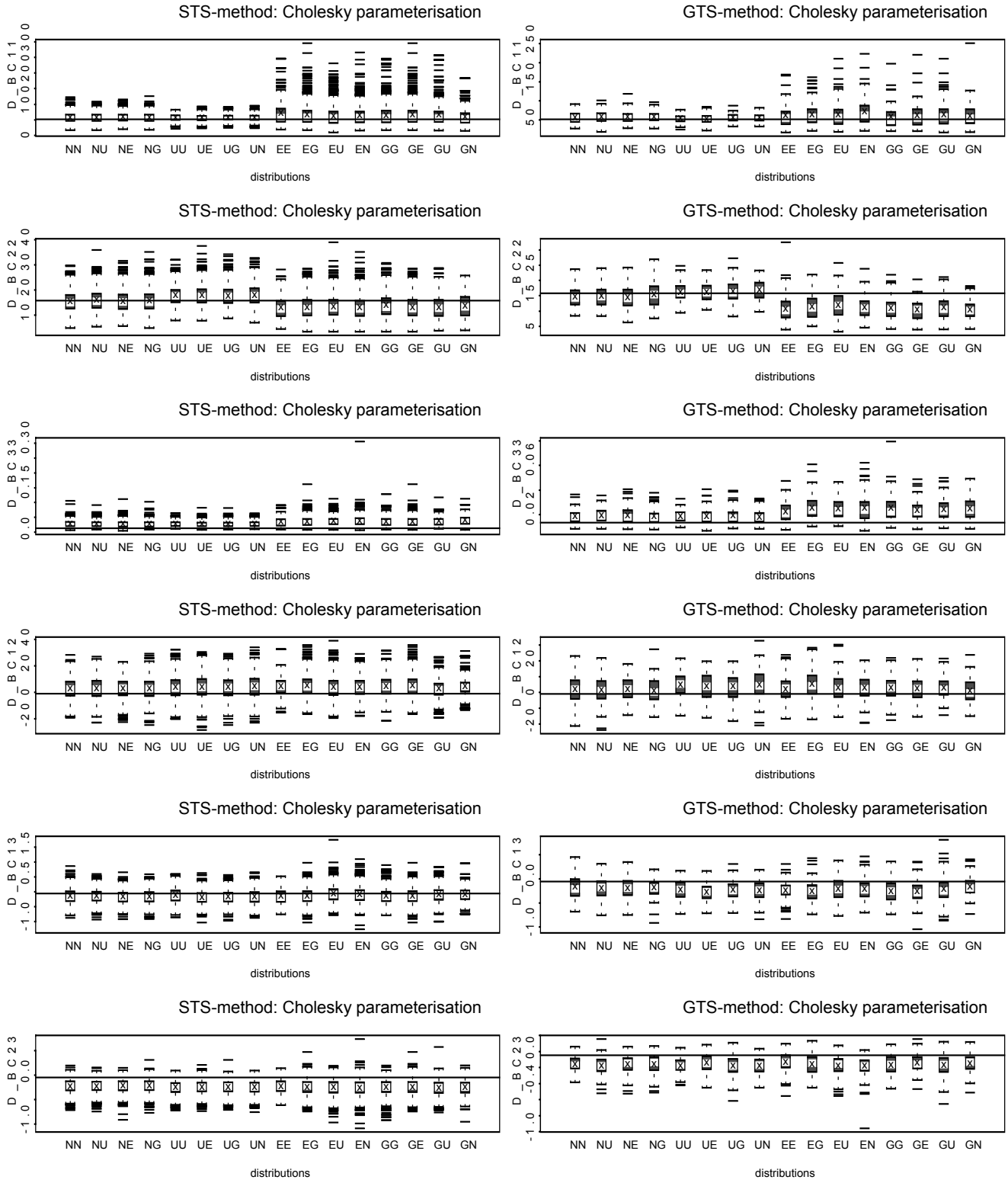


Figure 6.5: Box-plots of random components of Cholesky Bootstrap corrected version

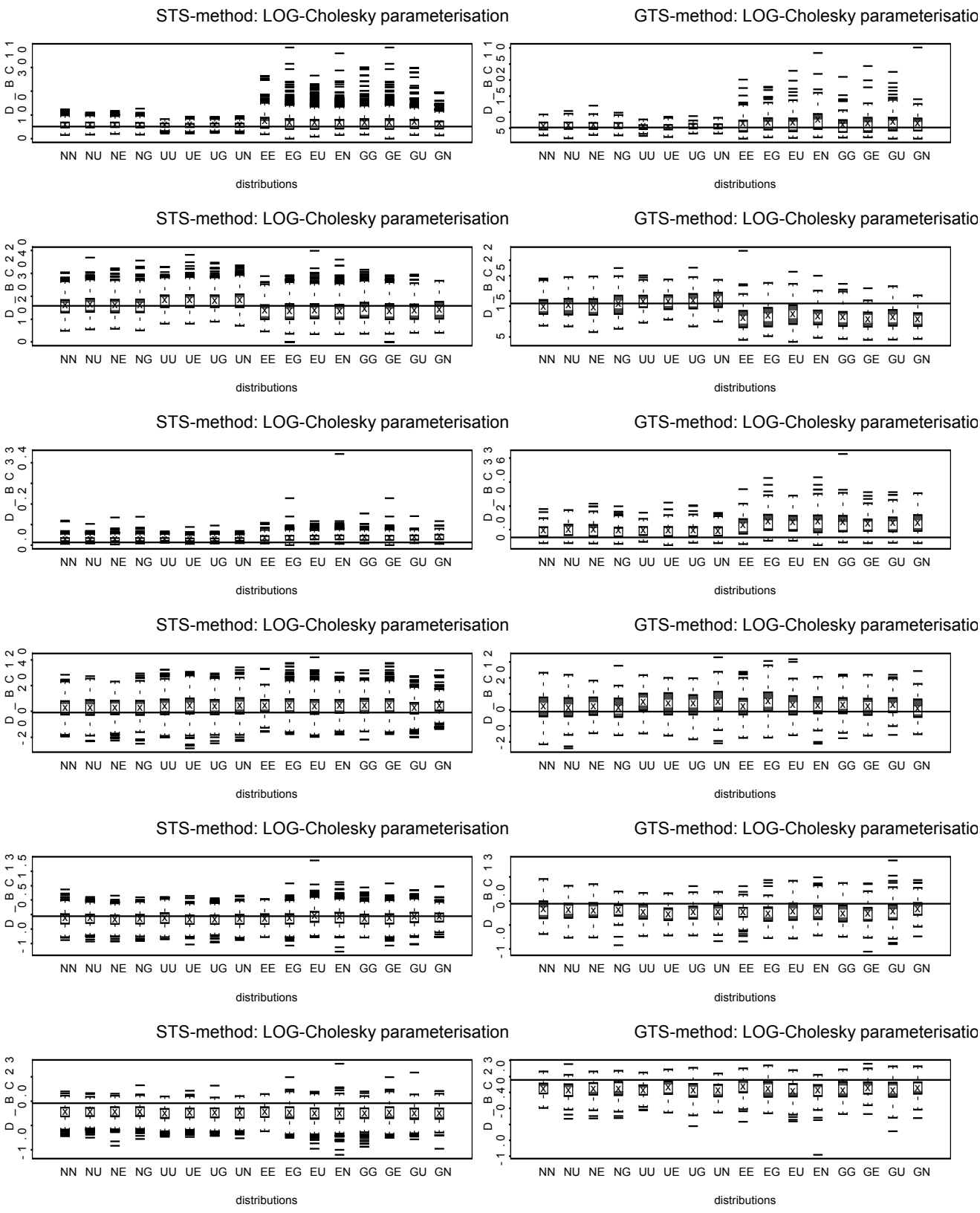


Figure 6.6: Box-plots of random components of Log-Cholesky Bootstrap corrected version

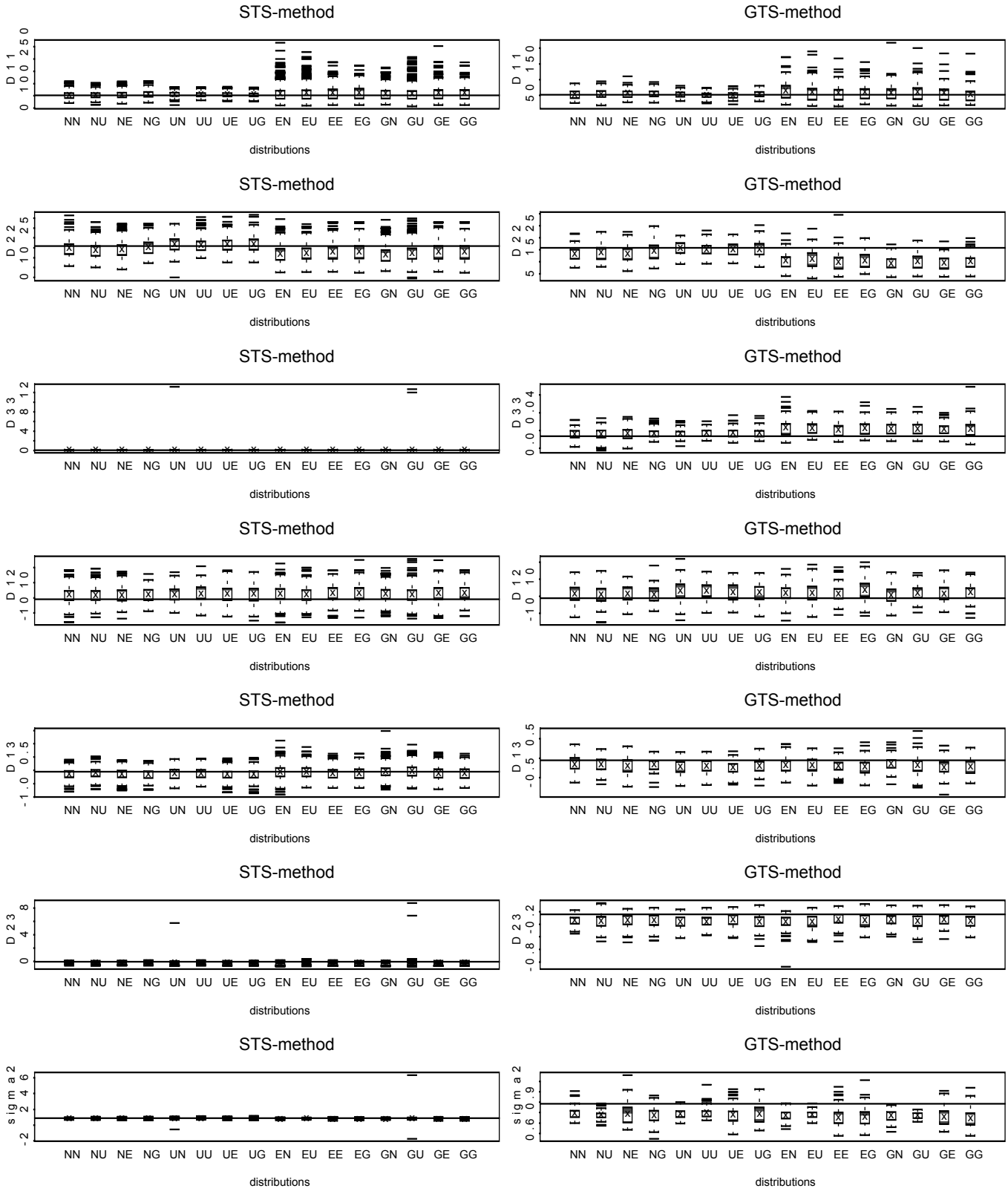


Figure 6.7: Box-plots of random components under STS and GTS methods

6.2. Some theatrical results

6.2.1. The bootstrap bias-corrected covariance matrix is asymptotically positive semi definite

The aim of this section is to investigate the asymptotic behaviour of the bootstrap bias corrected STS and GTS covariance matrices, already discussed in the previous chapter, as well as the behaviour of the bootstrap estimate of the fixed effects.

We first introduce the following notation. Let $O(x_n)$ denote the order upon which a sequence of real numbers is bounded. In a similar manner as in the work done by Carter and Yang (1986), we shall assume that

$$\left(X_i^{'}X_i\right)^{-1} = O_p(q_i^{-1}) \text{ uniformly for all elements of } \left(X_i^{'}X_i\right)^{-1} \text{ in probability} \tag{6.1}$$

where X_i is a "design" matrix corresponding to the $i-th$ individual and according to the following model

$$\tilde{\delta}_i \cong X_i\delta_i + \tilde{e}_i \tag{6.2}$$

where $\tilde{\delta}_i$ denotes the GLS estimator of δ_i and q_i is the dimension of X_i.

Let $\hat{\delta}_i = (X_i^{'}X_i)^{-1}X_i^{'}\tilde{\delta}_i$ denote the $i-th$ individual's ordinary least-squares estimator (OLS) of δ_i, thus

$$\hat{\delta}_i = \left(X_i^{'}X_i\right)^{-1}X_i\tilde{\delta}_i = \left(X_i^{'}X_i\right)^{-1}X_i^{'}(X_i\delta_i + \tilde{e}_i) = \delta_i + v_i$$

where $v_i = \left(X_i^{'}X_i\right)^{-1}X_i^{'}\tilde{e}_i$ and the "between–individual" sample covariance matrix obtained by the standard two stage method can be rewritten as:

$$\hat{D}_{STS} = (m-1)^{-1}\sum_{i=1}^{m}\left(\hat{\delta}_i - m_\delta\right)\left(\hat{\delta}_i - m_\delta\right)^{'}, \text{ where } m_\delta = \frac{\sum_{i=1}^{m}\delta_i}{m},$$

then

$$\begin{aligned}\hat{D}_{STS} &= (m-1)^{-1}\sum_{i=1}^{m}\left(\delta_i - m_\delta + v_i\right)\left(\delta_i - m_\delta + v_i\right)' \\ &= S_{\delta\delta} + (m-1)^{-1}\textstyle\sum_{i=1}^{m} v_i(\delta_i - m_\delta)' + (m-1)^{-1}\sum_{i=1}^{m}(\delta_i - m_\delta)v_i' \\ &\quad + (m-1)^{-1}\sum_{i=1}^{m} v_i v_i'\end{aligned}$$

where $S_{\delta\delta} = (m-1)^{-1}\sum_{i=1}^{m}(\delta_i - m_\delta)(\delta_i - m_\delta)'$.

Since, by (6.1), $v_i = O_p(q_i^{-1/2})$, it follows that $(m-1)^{-1}\sum v_i v_i' = O_p((m-1)^{-1}\max_i q_i^{-1/2})$, and the expected values of the second and third term are zero (Carter and Yang (1986)). However the standard two-stage estimate of D, $\hat{D}_{STS}$ can be rewritten as:

$$\hat{D}_{STS} = S_{\delta\delta} + O_p((m-1)^{-1}\max_i q_i^{-1/2}) .$$

Therefore, we have

$$\begin{aligned}\hat{D}_{STS,BC} &= 2\hat{D}_{STS} - B^{-1}\textstyle\sum_{k=1}^{B}\hat{D}_{STS}^{*,k} \\ &= 2(S_{\delta\delta} + O_p((m-1)^{-1}\max_i q_i^{-1/2})) \\ &\quad - B^{-1}\textstyle\sum_{k=1}^{B}(S_{\delta\delta} + O_p^k((m-1)^{-1}\max_i q_i^{-1/2})\end{aligned}$$

then,

$$\hat{D}_{STS,BC} = S_{\delta\delta} + O_p((m-1)^{-1}\max_i q_i^{-1/2}) .$$

When m is large ($m \to \infty$) the bootstrap corrected matrix is asymptotically semi definite. positive.

6.2.2. Moments of the bootstrap estimates

In the next result, we derive some properties of the bootstrap estimation method, by calculating the first and second moments:

In the second stage of the standard two stage methods, we choose $\tilde{\delta}_1^*,\ldots,\tilde{\delta}_m^*$ by random sampling with replacement from $\tilde{\delta}_1,\ldots,\tilde{\delta}_m$. Then, we have

$$\tilde{\delta}_i^* \simeq \delta + \eta_i^* .$$

First, the bootstrap estimators are unbiased under sampling over $(I = i^*)$: for E_I^* representing expectation over $(I = i^*)$

$$E_I^*\left(\tilde{\delta}_i^*\right) = E_I^*(\tilde{\delta}_i^* \mid I = i^*) = \tilde{\delta}_{i^*}$$

Therefore

$$E^*(\tilde{\delta}_i^*) = E^*\left\{E^*(\tilde{\delta}_i^* \mid I = i^*)\right\} = \sum_j \tilde{\delta}_j P_*(\tilde{\delta}_{i^*} = \tilde{\delta}_j) = m^{-1}\sum_j \tilde{\delta}_j = m_{\tilde{\delta}}\,.$$

As a consequence

$$E(E^*(\tilde{\delta}_i^*)) = m^{-1}\sum_j E(\tilde{\delta}_j) = \delta\,.$$

Second, the sample bootstraps have the following covariance matrix

$$\mathrm{var}^*(\tilde{\delta}_i^*) = E_I^*\left\{\mathrm{var}^*(\tilde{\delta}_i^* \mid I = i^*)\right\} + \mathrm{var}_I^*\left\{E^*\left(\tilde{\delta}_i^* \mid I = i^*\right)\right\}. \tag{6.3}$$

The second term on the right hand of the equation (6.3) can be calculated as follows

$$\begin{aligned}
\mathrm{var}_I^*\left\{E^*\left(\tilde{\delta}_i^* \mid I = i^*\right)\right\} &= \mathrm{var}^*(\tilde{\delta}_{i^*}) \\
&= E^*\left\{\left[\tilde{\delta}_{i^*} - E^*(\tilde{\delta}_{i^*})\right]\left[\tilde{\delta}_{i^*} - E^*(\tilde{\delta}_{i^*})\right]'\right\} \\
&= m^{-1}\sum_i \left(\tilde{\delta}_i - m_{\tilde{\delta}}\right)\left(\tilde{\delta}_i - m_{\tilde{\delta}}\right)' \\
&= \frac{m-1}{m}\hat{D}_{STS}
\end{aligned}$$

with $m_{\tilde{\delta}} = m^{-1}\sum \tilde{\delta}_i$

and to calculate the first term of the right hand of the equation (6.3), we need first to establish the following result

$$E_I^*(\mathrm{var}^*(\tilde{\delta}_i^* \mid I = i^*)) = \mathrm{var}^*(\tilde{\delta}_i^* - \tilde{\delta}_{i^*})\,. \tag{6.4}$$

Assuming the validity of (6.4), the fact that $E^*\left(\tilde{\delta}_i^* - \tilde{\delta}_{i^*}\right) = E^*\left\{E^*\left(\tilde{\delta}_i^* - \tilde{\delta}_{i^*} \mid I = i^*\right)\right\} = 0$ implies

$$\begin{aligned}
\mathrm{var}^*\left(\tilde{\delta}_i^* - \tilde{\delta}_{i^*}\right) &= E^*\left\{\left(\tilde{\delta}_i^* - \tilde{\delta}_{i^*}\right)\left(\tilde{\delta}_i^* - \tilde{\delta}_{i^*}\right)'\right\} \\
&= E_I^*\left\{E^*\left[\left(\tilde{\delta}_i^* - \tilde{\delta}_{i^*}\right)\left(\tilde{\delta}_i^* - \tilde{\delta}_{i^*}\right)' \mid I = i^*\right]\right\} \\
&= E^*\left[\left(\tilde{\delta}_{i^*} - \tilde{\delta}_{i^*}\right)\left(\tilde{\delta}_{i^*} - \tilde{\delta}_{i^*}\right)'\right] = 0
\end{aligned}$$

As a consequence,

$$E\left[\mathrm{var}^*(\tilde{\delta}_i^*)\right] = E\left\{\frac{m-1}{m}\hat{D}_{STS}\right\} \le E\left\{\hat{D}_{STS}\right\} = D + \sigma^2 m^{-1}\sum \Sigma_i \ge D\,.$$

we finally prove (6.4):

$$\begin{aligned}
\mathrm{var}^*\left(\tilde{\delta}_i^* - \tilde{\delta}_{i^*}\right) &= E_I^*\left\{\mathrm{var}^*\left(\tilde{\delta}_i^* - \tilde{\delta}_{i^*} \mid I = i^*\right)\right\} + \mathrm{var}_I^*\left\{E^*\left(\tilde{\delta}_i^* - \tilde{\delta}_{i^*} \mid I = i^*\right)\right\} \\
&= E_I^*\left\{\mathrm{var}^*\left(\tilde{\delta}_i^* - \tilde{\delta}_{i^*} \mid I = i^*\right)\right\} + \mathrm{var}^*(0),
\end{aligned}$$

because

$$\begin{aligned}
E^*(\tilde{\delta}_{i^*} \mid I = i^*) = \tilde{\delta}_{i^*} &= E^*(\tilde{\delta}_i^* \mid I = i^*) \\
&= E_I^*\left\{\mathrm{var}^*\left(\tilde{\delta}_i^* \mid I = i^*\right)\right\}\ .
\end{aligned}$$

On the other hand, and based on the regularity condition proposed by Freedman (1981) it can be shown that the bootstrap approximation to the distribution of STS estimates is valid, in the sense that the distribution of $\sqrt{m}\left(\hat{\delta}^{*}-\hat{\delta}\right)$ approximates the distribution of $\sqrt{m}\left(\hat{\delta}-\delta\right)$.

6.3. Conclusiones

We have studied the performance of the bootstrap inference in a large sample in non-linear mixed models and repeated measures contexts. Although it is not appropriate to draw general conclusions from a single simulation study, the results suggest some interesting features that may be worthy of further investigation. It appears that estimation of fixed parameters based on B_STS_MCB yield inferences that are substantially more accurate than B_STS and B_GTS methods and more accurate than ones based on asymptotic theory, and tends to perform robust front the departure from the normality assumption of the random effects and/or errors, even in the cases where the correlation structure and non-convergence of some subjects where expected to contaminate the inference. These results agree with the previous work of Vasnev (2000), and this provides confirmation of the intuitive notation that the Markov Chain bootstrap resampling processes for each subject do offer some improvement over the fixed effects based on two stage approximations. For variance components, this method yield relativity robust results only under the bias-corrected bootstrap under the log-Cholesky parametrization. On the other hand, one noticeable contrast in our simulation studies is that for the breast cancer model, the estimates of the fixed effects parameters and random components become slightly less precise when applying the B_STS and B_GTS procedure. One source of the explication for this discrepancy is the ignorable effect of the serial correlation between observations of subjects. Overall, the B_STS estimates are slightly better than B_GTS in terms of bias and coverage probabilities.

Finally, in terms of computation, the B_STS_MCB and B_GTS are "unfeasible" due to very intensive and very time consuming associated to non-linear optimisation problem, although the extraordinary rapid development of computing technology.

7. Summary and final conclusions

This chapter contains an outline of the results obtained in the thesis and the main conclusions, the problems that we encountered, and the directions we foresee for future work:

- Mixed models in the analysis of real data sets:

  - We have analysed some real data sets, using linear or non-linear mixed models. The mixed models were fitted with S-plus 2000, using the *lme* or *nlme* functions. The magnitudes of the variance components, heterogeneity of variance and correlation structures were estimated using maximum likelihood and/or restricted maximum likelihood. The theory of both estimation approaches for mixed models is reviewed in chapter 2.
  - An exhaustive study has been performed of the data from an experiment on breast cancer where the growth of the total tumoral volume is modeled in order to determine the influence of an external factor, diet. For these data, we suggested a general, flexible nonlinear model. The nonlinear mixed models for these data were also fitted using the *nlme* function, which uses a linearisation approach. Additionally, given the large number of observations (along time) for each subject, we also used alternative techniques based on individual estimates (two-stage methods) to refit the nonlinear mixed model to data. This allows a direct comparison between the two-stage and the linearization approaches.
  - Nonlinear mixed models (and both fitting methods) seem particularly well-suited for investigating the tumoural growth dynamics. No relation was found between the fixed parameters and diet. Exploratory graphic results suggest some relation between the variance of the random parameters and diet, indicating that hyperlipidic diets are associated with greater variances. But this relation is not confirmed by the inferential

procedures based on model comparison under a conditional linearization method, nor are confirmed by bootstrap procedures under the two-stage approach.

- Inspection of the validity of the model assumptions suggests that, while the random effects seem to be normally distributed, this is clearly not the case for the residuals.
- The large discrepancy between the linearization and the two-stage estimates is surprising and reflects a fundamental difficulty in estimating the fixed and random parameters when the normality assumption of the model is violated.
- Further investigation of this issue was required, our experience with these data suggests that caution should be exercised in the interpretation of results and that it is worthwhile to investigate the robustness of normal parametric inference in non-linear mixed models.

- Robustness of normal parametric inference in a NLME context:

- In view of the discrepancies between different fitting approaches, and provided that the normality requirement was not always fulfilled, we performed a simulation study to assess the robustness of the linearization and the two-stage methods under non-normality of random effects and/or residuals. These simulations complement similar studies, cited in this thesis.
- For the linearization approach, the main conclusion is that the parametric estimation methods based on the assumption of normality are quite robust for the fixed effects and much less robust for the random effects, with respect to bias and MSE. When the assumptions of normality of the residuals and/or the random effects are not met, some aspects of these inferential approaches are not entirely reliable, especially with respect to the coverage of the confidence intervals for the fixed effects and for the covariance estimates of the random effects. When the true distribution of the random effects is asymmetrical, the true coverage is markedly lower than the nominal one. With respect

to coverage of the true fixed effects parameters, the best results are obtained for the skew normal distribution, and not for the normal distribution. On the other hand, the results are partially reversed for the random effects.

- A complementary simulation study was devoted to the impact of incorrect assumptions on the true structure of the random effects covariance matrix and the true correlation pattern of residuals, over the performance of the linearization approach. Based on some measures such as bias, mean square error (MSE), and true coverage of the associated asymptotic confidence interval and ignoring other criteria like the convenience of avoiding over parameterized models, the main conclusion is that it seems worst to erroneously assume some structure than to not assume any structure when this would be adequate.
- The two-stage methods do not seem to be any less sensitive to violations of the distributional assumptions, and are in general worst even under normality conditions.

• Bootstrap resampling in a NLME context with application to the repeated measures:

- Different bootstrap techniques were suggested, to perform statistical inference based on the repeated measures and on non-linear mixed models. The convenience of these techniques was motivated by the results of the preceding studies with real data and by the simulations, indicating that the parametric approach may be not reliable.
- Among the suggested bootstrap approaches, the most promising alternative seems to be the method based on resampling in two-steps, between experimental units and inside each experimental unit, and on the two-stage approach to mixed models fitting –this fitting procedure was chosen by its simplicity and its distribution-free nature, not by their statistical properties, not very good. Preliminary simulation results suggest that it

may be more adequate to analyze data with large departures from normality, but further work is needed to improve it and to improve the computational methods, in order to make its calculation more feasible and to allow more concluding simulation studies.

Bibliography

Alonso, A., Peña, D. y Romo, J., 2002. Una Revisión De Los Métodos De Remuestreo En Series Temporales. *Estadística Española, Vol. 44, 133 - 159.*

Akaike, H., 1974. A new look at the statistical model identification. *IEEE Trans. On Automatic Control,* ***19:*** *716-723.*

Aitkin, M., Anderson, D. Francis, B. and Hinde, J., 1989. Statistical Modelling in GLIM. *New York: Oxford University Press.*

Anatolyev, S., Vasnev, A.., 2002. Markov chain approximation in bootstrapping autoregressions. *Economics Bulletin. 3, 1-8.*

Ann, O., Davidian, M., 1999. Estimation Data Transformations in Non-linear Mixed Effects Models, *Department of statistics, North Carolina state University.*

Azzalini, A.., 1986. Further results on a class of distributions which includes the normal ones. *Statistica, 46, 199-208.*

Bates, D.M., Watts, D.G., 1988. Nonlinear Regression Analysis and its Applications. *Wiley, New York..*

Bates, D. M. and Watts, D. G., 1980. Relative curvature measures of nonlinearity (with discussion). *J. Roy. Statist. Soc. Ser. B* ***42****, 1-25.*

Beal, S.L. and Sheiner, L.B., 1992. *NONMEM User's Guide, University of California, San Francisco, NONMEM Project Group.*

Box, G. E. P., 1949. A general distribution theory for a class of likelihood criteria. *Biometrika 36, 317-346.*

Box, G. E. P., 1950. Problems in the analysis of growth and wear curves. *Biometrics, 6, 362-389.*

Canty, A. J., 2000. An S-plus library for Resampling Methods. *Department of Mathematics Swiss Federal Institute of Technology 1015 Lausanne, Switzerland.*

Carroll, R. J., Ruppert, D., 1988. Transformation and Weighting in Regression. *London Chapman and Hall.*

Carter, R. L., Yang, M.C. K., 1986. Large sample inference in random coefficient regression models. *Communications in Statistics- Theory and Methods 15, 2507-2525.*

Chaganty, N. R., 1997. An alternative approach to the análisis of longitudinal data via generalizad estimating equations. *J. Statist. Planng. Infer.*, 63, 39-54.

Chambers, J., William C., Beat K., and Paul T., 1983. Graphical Methods for Data Analysis, *Wadsworth.*

Crowder, M., 1995. On the use of a working correlation matrix in using generalized linear models for repeated measures. *Biometrika* 82, 407-410.

Crowder, M., 2001. On repeated measures analysis

with misspecified covariance structure. *J. R. Statist. Soc.* B, 63, 55-62.

Das, S., Krishen, A., 1999. Some bootstrap methods in nonlinear mixed-effect models. *Journal of Statistical Planning and Inference. 75, 237-245.*

Davidian, M., Gallant, A.R., 1993. The nonlinear mixed effects model with a smooth random effects density. *Biometrika, 80, 475-488.*

Davidian, M.; Giltinan, D.M., 1995. Non-linear Models for Repeated Measurement Data; *Chapman & hall: London.*

Davison, A.C., D.V. Hinkley., 1997. Bootstrap Methods and Their Application. *Cambridge: Cambridge University Press.*

Demidenko, E., 1995. On the existence of the least squares estimate in nonlinear growth curve models of exponential type. *Communications in Statistics. Theory and Methods 25, 159-182.*

Dempster, A.P., Laird, N. M., and Rubin, D. B., 1977. Maximum likelihood from incomplete data via the EM algorithm (with discussion). *Journal of the Royal Statistical Society, Series B, 39, 1-38.*

Diggle, P.J., Liang, K.-Y., Zeger, S.L., 1994. Analysis of Longitudinal Data. *Oxford Science Publications. Oxford: Clarendon Press.*

Efron, B. ,1979. Bootstrap methods: another look at the jacknife. *Annals of Statistics 7, 1-26.*

Efron, B., Tibshirani, R.J., 1993. An Introduction to the Bootstrap. *Chapman & Hall, New York.*

El Halimi, R., Ocaña, J, Ruiz De Villa, M.C., Escrich, E., Solanas, M., 2003. Modeling tumor growth data using a nonlinear mixed-effects model. *InterStat*,

http://jscs.stat.vt.edu/InterStat/ARTICLES/2003/abstracts/0309002.html-ssi

El Halimi, R., Ocaña, J., Ruiz de Villa, M.C., 2004. A simulation study on the robustness of parametric inference in a nonlinear mixed modelling context. www.mathpreprints.com.

El Halimi, R. and Ocaña, J., 2004. On the consequences of misspecifing assumptions concerning residuals distribution in a repeated measures and non-linear mixed modelling context. *Preprints del Departament d'Estadística, Universitat de Barcelona.* http://www.ub.es/stat/recerca/Preprints.html

Elston, R. C., Grizzle, J. E., 1962. Estimation of time-response curves and their confidence bands. *Biometrics 18, 148-1599.*

Escrich E., Solanas M. Segura R., 1994. Experimental diets for the study of lipid influence on the induced mammary carcinoma in rats: I-diet definition. *International Journal in Vivo Research, 8, 1099-1106.*

Escrich E., Solanas M., Ruiz de Villa M. C., Ribalta Muntané J., Segura R., 1994. Experimental diets for the study of lipid influence on the induced mammary carcinoma in rats: I-diet definition. *International Journal in Vivo Research, 8, 1007-1012.*

Fisher, R. A., 1925. Statistical Methods for Research Workers. *Oliver and Boyd. Edinburgh. UK. 362 p*

Freedman, D. A., 1981. Bootstrapping regression models. *The Annals of Statistics*, 9, 1218-1228.

Gelfand, A. E., Hills, S. E., Racine-Poon, A., and Smith, A. F.M., 1990. Illustration of Bayesian inference in normal data models using Gibbs sampling. *Journal of the American Statistical Association*, **85**, 972-85.

Goldstein, H., 1995. Multilevel Statistical Models. *London, Edward Arnold; New York, Halstead Press.*

Montserrat, G., 2000. Efectes de l´aportació de residus al sól sobre la miniralització del carboni i del nitrogen. *Doctoral dissertation; Department of products Naturals. University of Barcelona.*

Gelman, A., Carlin, J.B., Stern, H.S., and Rubin, DB., 1995. Bayesian Data Analysis. *Texas in Statistical Science. London: Chapman & Hall.*

Goldstein, H., 1987. Multilevel models in educational and social research. Griffin - *Oxford University Press, London.*

Gumpertz, M.L., Pantula, S.G., 1992. Nonlinear regression with variance components. J. *Amer. Statist. Assoc. 87, 201–209.*

Hansen, B.E., 1999a. The Grid Bootstrap and Autoregressive model. *Review of Economics and Statistics 81, 594-607.*

Hansen, B.E., 1999b. Non-Parametric Dependent Data Bootstrap for Conditional Moment Models. *Working paper, University of Wisconsin-Madison.*

Hartford, A., Davidian, M., 1999. Consequences of misspecifying assumptions in nonlinear mixed effects models. *Computational Statistics & Data Analysis, 34, 139-164.*

Hartley, H., and Rao,J., 1967. Maximum-likelihood estimation for the mixed analysis of variance model. *Biometrika 54:93-108.*

Harville, D. A., 1976. Extention of the Gauss-Markov theorem to include the estimation of random effects. *Annals of Statistics 4, 384-395.*

Harville, D. A., 1977. Maximum likelihood approaches to variance component estimation and to related problems. *J. Amer. Statist. Ass. 72, 320-338.*

Huggins, C, Grand, L.C, Brillantes, F.P., 1961. Mammary cancer induced by a single feeding of polynuclear hydrocarbons, and its suppression. *Nature*, **189**: 204-207.

Hunt, C. A.; Givens, G.H.; GS., 1998. Bootstrapping for Pharmacokinetic Models: Visualization of predictive and Parameter Uncertainty. *Pharmaceutical Research, vol. 15, 5*

Jennrich, R.I., Schluchter, M.D., 1986. Unbalanced Repeated- Measures Models with Structured Covariance Matrices. *Biometrics, 42, 805-820.*

Kalsson. S, Löthgren, M., 2000. Computationally efficient double bootstrap variance estimation. *Computation statistics & data Analysis, 237-247*

Laird, N. M. and Ware, J. H., 1982. Random effects models for longitudinal data. *Biometrics 38, 963-974.*

Lindstrom, M.J., Bates, D.M., 1990. Nonlinear mixed effects models for repeated measures data. *Biometrics, 46, 673-687.*

Littell, R. C., G. A. Milliken, W. W. Stroup, and R. D. Wolfinger., 1996. SAS® system for mixed models. SAS Institute, *Inc., Cary, North Carolina, USA.*

Longford, N. T., 1993. Random Coefficients Models. *Oxford Univ. Press, Oxford.*

Newton, M.A., Zhang, Y. A, 1996. Recursive algorithm for nonparametric analysis with missing data. *Technical report 965, New York: Springer-Verlag.*

Moulton, L.H.; Zeger, S.L. 1989. Analyzing Repeated Measures on Generalized linear Models via the *Bootstrap. Biometrics, 381-394*

Park, T., Yung Shin, D., 1999. On the use of working correlation matrices in the GEE approach for longitudinal data. *Commun. Statist.- Simulation, 28 (4), 1011-1029.*

Patterson H.D, Thompson R., 1971. Recovery of inter-block information when block sizes are unequal. *Biometrika,* ***58:*** *545-554*

Pinheiro, J.C., 1994. Topics in Mixed Effects Models. *Ph.D. Thesis, University of Wisconsin-Madison.*

Pinheiro, J.C., Bates, D.M., 1995. Model Building for Nonlinear Mixed-Effects Models. *Technical Report 91; Department of Biostatistics, University of Wisconsin-Madison.*

Pinheiro, J.C., Bates, D.M., 1996. Unconstrained parameterizations for variance-covariance matrices, *Statistics and Computing, 6, 289-296*

Pinheiro, J.C., Bates, D.M., 2000. Mixed-Effects Models in S and S-Plus. *Springer-Verlag, Berlin.*

Pinhero, J.C. and D.M. Bates, 1995. Approximations to the log-likelihood function in the nonlinear mixed-effects model, *Journal of the Computational and Graphical Statistics, 4 12-35.*

Ramos, R.Q. and S.G. Pantula,1995. Estimation of nonlinear coefficient models, *Statistics and Probability Letters, 24 49-56.*

Pocock, S.J., Cook, D.G., Beresford, S.A.A., 1981. Regression of area mortality rates on explanatory variables: what weighting is appropiate. *Applied Statistics 30, 286-295.*

Rao, C. R., 1965. The theory of the least squares when the parameters are stochastic and its application to the analysis of growth curves. *Biometrika 52, 447-458.*

Rao, C. R., 1972. Estimation of variance and covariance components in linear models. *J. Amer. Statist. Ass. 67, 112-115.*

Robinson, G. K., 1991. That BLUP is a good thing: the estimation of random effects. *Statistical Science: 15-32.*

Rosenberg, B., 1973. Linear regression with randomly dispersed parameters. *Biometrika 60, 65-72*

Sánchez, J.A., Ocaña, J., 2000. Parametric bootstrap to test robustness of likelihood estimation in linear mixed models. *Workshop on Inference and Asymptotics. Ascona, Switzerland. 14-20 July 2000.*

Sánchez, J.A., Ocaña, J., 2001. Implementación de métodos de computación intensiva para modelos mixtos. *VIII Conferencia Española de Biometría. Pamplona, 28-30 Marzo 2001. Actas, 185-186.*

Sánchez, J.A., Ocaña, J., 2001. Improving resampling techniques for linear mixed models: Nested and Adjusted semiparametric bootstrap. *XXVI Congreso Nacional de Estadística e Investigación Operativa. Úbeda, Jaen, 6-9 Noviembre*

Sánchez, J.A., Ocaña, J., 2002. Computer intensive methods for mixed-effects models. In: *Härdle, W., Rönz, B. (Eds.), Proceedings in Computational Statistics. COMPSTAT 2002, Physica-Verlag, Berlin, 261-266.*

Sakamoto, Y., Ishiguro, M., and Kitagawa G., 1986. "Akaike Information Criterion Statistics", *D. Reidel Publishing Company.*

Schwarz, G., 1978. Estimating the dimension of a model. *The Annals of Statistics*, **6**: 461-464.

Searle, S., Casella, G., McCulloch, C., 1992. *Variance components. John Wiley & Sons, New York.*

Seber, G.A.F., 1984. Multivariate Observations. *New York: John Wiley and Sons.*

Shao, J., 1990. Bootstrap estimation of the Asymptotic variances of statistical functionals. *Annals of the Institute of Statistical Mathematics, 42, 737-752.*

Sheather, S. J. and Jones, M. C., 1991. A reliable data-based bandwidth selection method for kernel density estimation. Journal of the Royal Statistical Society, Series B 53, 683-690.

Sheiner, L. B., Beal, S. L., 1980. Evaluation of methods for estimating population pharmacokinetic parameters. I. Michaelis-Menten model. Routine clinical pharmacokinetic data. *Journal of pharmacokinetics and Biopharmaceutics, 8, 553-571.*

Sheiner, L.B. and Beal, S.L., 1985. "Pharmacokinetic Parameter Estimates from Several Least Squares Procedures: Superiority of Extended Least Squares," *Journal of Pharmacokinetics and Biopharmaceutics, 13, 185 -201.*

Singh, K., 1981. On the asymptotic accuracy of Efron's bootstrap, *Ann. Statist., **9**,1187-1195.*

Steimer, J.L., Mallet, A., Golmard, J.L., Boisvieux, J.F., 1984. Alternative approaches to estimation of population pharmacokinetic parameters: comparison with the nonlinear mixed effect model. *Drug Metabolism Reviews. 15, 265-269*

Stukel, T.A., Demidenko, E., 1997. Comparison of methods for general nonlinear mixed-effects models. Modelling Longitudinal and Spatially Correlated Data: Methods, Applications, and Future Directions. *Springer Lecture Notes in Statistics. Springer, New York.*

Sutradhar, B. C. and Das, K., 1999. On the efficiency of regression estimators in generalised linear models for longitudinal data. *Biometrika,* ***86****, 459–465.*

Swamy, P. A. V. B., 1970. Efficient inference in random coefficient regression model. *Econometrica 38, 311-323.*

Swamy, P. A. V. B., 1971. Statistical Inference in Random Coefficient Regression Models. *Springer-Verlag. Berlin.*

Swamy, P. A. V. B., 1973. Criteria, constraints and multicollinearity in random coefficient regression models. *Annals of Economic and Social Measurments 2, 429-450.*

S-Plus 2000 Guide to Statistics. Data Analysis Products Division, *MathSoft, Seattle, WA.*

Tao, H., Palta, M., Yandell, B.S., Newton, M.A., 1999. An estimation method for the semiparametric mixed effects model. *Biometrics. 55, 102-110.*

Thisted, R.A. 1988. Elements of statistical computing. *Chapman and Hall, New York.*

Thompson, W.A., Jr., 1962. The problem of negative estimates of variance components. *Annals of Mathematical Statistics, 33, 273-289.*

Timm, N. H., 1980. Multivariate analysis of variance of repeated measurments. *In Handbook of Statistics, Vol. I ed. By Krishnaiah, P.R. North-Holland. New York. Pp 41-87.*

Venables, W.N., and Ripley, B.D., 1997. Modern Applied Statistics with S-plus, New York: *Springer-Verlag.*

Vasnev, A., 2001. A nonparametric method for bootstrap resampling in autoregression. *Working Paper #BSP/2001/O.E., New Economic School, Moscow.*

Vonesh, E. F., Carter, R. L., 1992. Mixed effects nonlinear regression for unbalanced repeated mesures. *Biometrics, 48, 1-17.*

Vonesh, E. F., Carter, R. L., 1992. Nonlinear models for the analysis of longitudinal data. *Statistics in Medicine 11, 1929-1954.*

Vonesh, E. F., 1996. A note on the use of Laplace´s approximation for nonlinear mixed effects models. *Biometrika 83, 447-452.*

Vonesh E.F, Chinchilli V.M., 1997. Linear and Nonlinear Models for the Analysis of Repeated Measurements. *Marcel Dekker*

Walker, S., 1996. An EM algorithm for nonlinear random effects Models, *Biometrics, 52 934- 944.*

Wand, M. and Jones, M. 1995. *Kernel Smoothing*. London: *Chapman and Hall.*

Weisberg, S., 1985. Applied Linear Regression. *New York: Wiley.*

Wolfinger, R., 1993. Laplace´s approximation for non-linear mixed effects models. *Biometrika, 80, 791-795*

Wolfinger, R.D., Lin, X., 1997. Two Taylor-series approximation methods for nonlinear models. *Computational Statistics & Data Analysis, 25, 465-490.*

Yuh, L., Beal, S., Davidian, M., Harrison, F., Hester, A., Kowalski, K. Vonesh, E. and Walfinger, R., 1994. Population pharmacokinetic/pharmacodynamic methodology and applications: a bibliography. *Biometrics 50, 566-575.*

Zhiyu Ge, Peter J. Bickel, and John A. Rice, 2003. An Approximate Likelihood Approach to Nonlinear Mixed Effects Models via Spline Approximation. *Available in:* http://stat-www.berkeley.edu/users/bickel