

UNIVERSIDAD NACIONAL DE CÓRDOBA

APLICACIÓN DE MODELOS MIXTOS NO
LINEALES EN EL CRECIMIENTO DE FRUTOS
DE PERAL (PYRUS COMMUNIS) Cv.
WILLIAM'S

PARA OPTAR AL GRADO DE: MAGISTER EN ESTADÍSTICA APLICADA

AUTOR: VALENTÍN TASSILE

INGENIERO AGRÓNOMO

2020



APLICACIÓN DE MODELOS MIXTOS NO LINEALES EN EL CRECIMIENTO DE FRUTOS DE PERAL (PYRUS COMMUNIS) CV. WILLIAM'S by VALENTÍN TASSILE is licensed under a [Creative Commons Reconocimiento-NoComercial-SinObraDerivada 4.0 Internacional License](https://creativecommons.org/licenses/by-nc-nd/4.0/).

COMISIÓN ASESORA DE TESIS

Director: Dr. Bramardi Sergio

Co-Directora: Dra. María del Pilar Diaz

Fecha de aprobación de de tesis: 2 de Noviembre de 2015

Agradecimientos

Este trabajo no habría sido posible sin el estímulo de mi familia, que siempre me apoyó en mis búsquedas.

A Omar Alvarez y Sergio Bramardi, mis grandes mentores en esta hermosa profesión.

A la Comisión Asesora y al Tribunal Evaluador por los enormes aportes realizados en este trabajo.

A mis amigos y compañeros de trabajo por alentarme y escucharme en este transitar por la Academia.

Palabras Clave:

Modelos No Lineales Mixtos, Pronósticos de Producción, Efectos Aleatorios, Predicción de Distribuciones de Tamaños

Resumen:

El conocimiento del peso final de frutos con antelación a la cosecha es una información de vital importancia por las consecuencias que tiene sobre las distintas etapas del proceso de producción y comercialización en las regiones de producción. A efectos de obtener mejoras en la capacidad predictiva de los modelos sobre la distribución final del tamaño de frutos, es necesario que los mismos contemplen los diversos factores que afectan el proceso de crecimiento. Para el caso de los frutos de pepita el proceso de crecimiento del diámetro ecuatorial de los mismos responde a un modelo sigmoideo en forma de S, un comportamiento claramente no lineal en los parámetros. La presente tesis se propone lograr un modelo mixto no lineal con alta capacidad predictiva para frutos de peras cv. William's cultivados en huertos comerciales de los Valles Irrigados de la Provincia de Río Negro, Argentina. Se consideran, para ello, aquellas fuentes de mayor relevancia para la variabilidad de la variable de respuesta, mancomunando información de naturaleza longitudinal y transversal e incorporando covariables ambientales representativas de distintos momentos críticos del crecimiento de los frutos. El cribado final de modelos considera métricas adecuadas para que los candidatos propuestos posean un alto poder predictivo en situaciones no contempladas en la etapa de modelado.

Application of Non-Linear Mixed Models in the Growth of Pear Fruits(*Pyrus Communis*). Cv. William 's

Keywords:

Mixed Nonlinear Models, Production Forecasts, Random Effects, Prediction of Size Distributions

Summary:

The knowledge of the final weight of fruits before the harvest is vitally important information due to the consequences it has on the different stages of the production and marketing process in the production regions. In order to obtain improvements in the predictive capacity of the models on the final distribution of fruit size, it is necessary that they consider the various factors that affect the growth process. In the case of pome fruits, the growth process of their equatorial diameter responds to an S-shaped sigmoid model, a clearly non-linear behavior in the parameters. The present thesis sets out to achieve a nonlinear mixed model with high predictive capacity for cv. William 's cultivated in commercial gardens of the Irrigated Valleys of the Province of Río Negro, Argentina. For this, those sources of greatest relevance for the variability of the response variable are considered, pooling information of a longitudinal and cross-sectional nature and incorporating representative environmental covariates of different critical moments of fruit growth. The final screening of models considers adequate metrics so that the proposed candidates have a high predictive power in situations not contemplated in the modeling stage.

Índice general

I INTRODUCCIÓN Y MOTIVACIÓN	19
1. INTRODUCCIÓN	20
1.1. PRESENTACIÓN DEL PROBLEMA	20
1.2. MOTIVACIÓN	22
1.3. DESCRIPCIÓN DEL TRABAJO	25
1.4. OBJETIVO GENERAL	28
1.5. OBJETIVOS ESPECÍFICOS	28
1.6. METODOLOGÍA	28
1.7. ALCANCES DEL TRABAJO	29
2. DESCRIPCIÓN DEL CULTIVO Y EL ÁREA DE ESTUDIO	30
2.1. LOCALIZACIÓN GEOGRAFICA	30
2.2. SUELO, CLIMA Y MEDIO AMBIENTE	31
2.3. EVOLUCIÓN HISTORICA DEL CULTIVO	32
2.3.1. ORIGEN E HISTORIA DE LA PERA WILLIAM'S	32
2.3.2. EL CULTIVO EN LA REGIÓN	33
II REVISIÓN DE LITERATURA	35
3. ASPECTOS AGRÓNOMICOS DEL CRECIMIENTO	36

3.1. INTRODUCCIÓN	36
3.2. FACTORES INTERNOS QUE AFECTAN EL CRECIMIENTO	38
3.3. FACTORES EXTERNOS QUE AFECTAN EL CRECIMIENTO	40
3.4. BASES FISIOLÓGICAS DEL CRECIMIENTO DE FRUTOS	44
3.5. PREDICCIÓN DEL TAMAÑO DEL FRUTO	45
4. ASPECTOS ESTADÍSTICOS	48
4.1. MODELO DE REGRESIÓN LINEAL CLÁSICO	48
4.2. MODELOS DE CRECIMIENTO DE FRUTOS	51
4.2.1. INTRODUCCIÓN	51
4.2.2. MODELO DE REGRESIÓN NO LINEAL	51
4.2.3. MODELOS DE CRECIMIENTO SIGMOIDAL	54
4.3. MODELOS DE RESPUESTA MULTIVARIADA	55
4.3.1. INTRODUCCIÓN	55
4.3.2. MODELO LINEAL GENERAL MULTIVARIADO	57
4.3.3. MODELOS SUJETO-ESPECÍFICOS	59
4.3.4. MODELOS SUJETOS-ESPECIFICOS VERSUS MODELOS PROME- DIO POBLACIONALES	68
4.4. MODELOS NO LINEALES DE EFECTOS MIXTOS	70
4.4.1. INTRODUCCIÓN	70
4.4.2. NLME PARAMÉTRICOS	72
4.4.3. NLME NO PARAMÉTRICOS	73
4.4.4. NLME SEMI-PARAMÉTRICOS	73
4.4.5. ENFOQUE BAYESIANO A LOS NLME	74
4.5. IMPLEMENTACIÓN DE LOS NLME	74
4.5.1. INTRODUCCIÓN	74
4.5.2. INFERENCIA A NIVEL DE SUJETO	76

4.5.3. MÉTODOS DE ESTIMACIÓN EN NLME	78
4.6. ESTRATEGIAS PARA LA CONSTRUCCIÓN DE MODELOS	86
4.6.1. INTRODUCCIÓN	86
4.6.2. PROPIEDADES ASINTÓTICAS DE LOS NLME	86
4.6.3. PRUEBAS DE HIPÓTESIS SOBRE EFECTOS FIJOS	87
4.6.4. DETERMINACIÓN DE ESTRUCTURAS EN LOS EFECTOS ALEATO- RIOS	90
4.6.5. INCORPORACIÓN DE COVARIABLES	92
4.6.6. EXAMINANDO EL MODELO AJUSTADO	94
4.7. MÉTRICAS DE COMPARACIÓN DE MODELOS	95
4.8. MÉTODOS PARA VALIDAR MODELOS	99
4.8.1. PREDICCIÓN DE NUEVOS EFECTOS ALEATORIOS	99
4.8.2. PREDICCIÓN DE FUTURAS OBSERVACIONES	101
III EXPERIMENTACIÓN	104
5. MATERIALES Y MÉTODOS	105
5.1. OBSERVACIONES LONGITUDINALES	105
5.2. OBSERVACIONES TRANSVERSALES	106
5.3. INFORMACIÓN AGROCLIMÁTICA	108
5.4. MODELACIÓN	110
5.5. CALIBRACIÓN	112
5.6. PREDICCIONES	114
6. RESULTADOS Y DISCUSIÓN	115
6.1. MODELACIÓN	115
6.1.1. INTRODUCCIÓN	115
6.1.2. MODELACIÓN EN LAS MEDICIONES LONGITUDINALES	116

6.1.3. INCORPORACIÓN DE MEDIDAS TRANSVERSALES	124
6.1.4. INCORPORACIÓN DE COVARIABLES AMBIENTALES	130
6.2. CALIBRACIÓN	137
6.3. PREDICCIONES	138
6.3.1. BASE DE VALIDACIÓN DE DATOS LONGITUDINALES	140
6.3.2. BASE DE VALIDACIÓN DE DATOS TRANSVERSALES	144
6.3.3. ESPERANZAS Y CORRELACIONES MARGINALES INDUCIDAS . . .	144
7. CONCLUSIONES	146
A. MÉTRICAS DE VALIDACIÓN.	
RESULTADOS OBTENIDOS DETALLADOS	150
B. VERIFICACIÓN GRÁFICA DE SUPUESTOS	176
C. CORRELACIONES MARGINALES	182
D. SCRIPTS: PREDICCIÓN DE EFECTOS ALEATORIOS	186
E. SCRIPTS: PREDICCIÓN DE VARIABLES DE RESPUESTA	203
F. SCRIPTS: MÉTRICAS DE EVALUACIÓN	205
Bibliografía	209

Índice de figuras

3.1. Curva de Crecimiento del Fruto.	37
4.1. Fuentes de Variación Intra-sujeto.	62
5.1. Mediciones de las Bases de Modelación LONGITUDINALES y TRANSVERSALES	108
5.2. Mediciones de las Bases de Validación LONGITUDINALES y TRANSVERSALES	109
6.1. Correlograma del MODELO 2b	122
6.2. Correlograma de MODELO 9	124
6.3. $b_{LONGI,3,SITIO_j}$ y $b_{TRANS,3,SITIO_j}$ de MODELO 20 vs ITAC propuestos en Etapas Tempranas del Ciclo de Crecimiento	132
6.4. $b_{2,SITIO_j}$ de MODELO 21 vs ITAC propuestos en Etapas Tempranas del Ciclo de Crecimiento	133
6.5. $b_{LONGI,3,SITIO_j}$, $b_{TRANS,3,SITIO_j}$ de MODELO 20 vs ITAC propuestos para el transcurso del mes de Diciembre	134
A.1. Distribución de Diámetros Obs. para Base de Validación LONGITUDINALES	173
A.2. Distribución de Diámetros Obs. para Base de Validación TRANSVERSALES	175
B.1. Verificación Gráfica de Supuestos de MODELO 9	177

B.2. Verificación Gráfica de Supuestos de MODELO 15	178
B.3. Residuales de MODELO 11	179
B.4. Verificación Gráfica de Supuestos de MODELO 17	180
B.5. Verificación Gráfica de Supuestos de MODELO 22	181
C.1. Correlación Marginal para obs. del mismo Fruto en MODELO 21	183
C.2. Correlación Marginal para obs. del mismo Fruto en MODELO 24	184
C.3. Correlación Marginal para obs. del mismo Fruto en MODELO 27	185

Índice de cuadros

5.1. Número de mediciones y frutos en Base de Modelación LONGITUDINALES	107
5.2. Número de mediciones y frutos en Base de Validación LONGITUDINALES	107
5.3. Número de frutos en Base de Modelación TRANSVERSALES	107
5.4. Número de frutos en Base de Validación TRANSVERSALES	108
6.1. CI de Modelos Propuestos estimados por ML en Base de Modelación LONGITUDINALES	118
6.2. CV del Diámetro Ecuatorial de Frutos	119
6.3. Cambios de CV de Diámetros Ecuatoriales en Frutos Simulados	120
6.4. Valores de CV_{b_i} de MODELO 4	120
6.5. CI de modelos estimados por ML considerando Correlación en los Errores	123
6.6. Parámetros estimados en Modelos estimados por REML en Base de Modelación LONGITUDINALES	125
6.7. CI en Modelos con Información Combinada estimados por ML	126
6.8. Capacidad Predictiva en Modelos con Información Combinada	127
6.9. CI en Modelos con Información Combinada II estimados por ML	129
6.10. Capacidad Predictiva en Modelos con Información Combinada II	130
6.11. Parámetros Estimados en Modelos con Información Combinada estimados por REML	130
6.12. Capacidad Predictiva en Modelos con Información Combinada e ITAC . . .	136

6.13. Parámetros Estimados en Modelos con Información Combinada e ITAC estimados por REML	137
6.14. Combinaciones Consideradas para Base de Validación LONGITUDINALES	139
6.15. Combinaciones Consideradas para Base de Validación TRANSVERSALES .	139
6.16. Capacidad Predictiva para una Obs. en Base de Validación LONGITUDINALES	141
6.17. Capacidad Predictiva en Base de Validación TRANSVERSALES según nro. de Obs.	143
6.18. Capacidad Predictiva en Base de Validación TRANSVERSALES	145
A.1. Métricas de Validación para Modelos Propuestos	150
A.2. Condición I en Base de Validación LONGITUDINALES	172
A.3. Condición II en Base de Validación LONGITUDINALES	174
A.4. Condición III en Base de Validación LONGITUDINALES	174

LISTA DE ACRÓNIMOS

AIC Criterio de Información de Akaike

ANOVA Modelo Lineal General

AR(1) Autocorrelación de Primer Orden

ASE Error Estandar Asintótico

BIC Criterio de Información Bayesiano

BLUP Mejores Predictores Lineales Insesgados

B.46.T.100.MIN.16 Covariable Ambiental Candidata Temprana para β_3

B.122.T.150.MIN.22 Covariable Ambiental Candidata Tardia para β_3

B.60.T.100.MIN.17 Covariable Ambiental Candidata para β_2

CAR(1) Autocorrelación Contínua de Primer Orden

CCC Coeficiente de Concordancia

CI Criterios de Información

CP Criterios Predictivos

CV Coeficiente de Variación

DDA Días posteriores al Primero de Agosto

DDPF Días Posteriores a Plena Floración

δ Precisión Global de los Errores

EBEs Estimador Bayesiano Empírico

EBLUP Método de Linealización de Primer Orden sobre los BLUPs

$EBLUP_b$ Predictor de la Expansión EBLUP

\bar{e} Sesgo Medio

$\bar{e}\%$ Sesgo Medio Relativo

$|\bar{e}|$ Sesgo Medio Absoluto

$|\bar{e}\%|$ Sesgo Medio Relativo Absoluto

FO Método de Linealización de Primer Orden

GGE Ecuaciones de Estimación Generalizadas

INTA Alto Valle Instituto Nacional de Tecnología Agropecuaria Alto Valle

ITAC Índices de Termoacumulación

MANOVA Modelo Lineal General Multivariado

MCMC Cadenas de Markov Monte Carlo

ML Máxima Verosimilitud

NL Mínimos Cuadrados No Lineales

NLME Modelo Mixto No Lineal

LRT Razón de Verosimilitud

PNL_y Predictor No Lineal

PL_y Predictor Lineal

PP Promedio-Poblacional

PPPRNN Programa de Pronóstico de Producción de Frutos de Pepita de las Provincias
de Rio Negro y Neuquén

PRESS Suma de Cuadrados de Errores de Predicción

REML Máxima Verosimilitud Restringida

RENSPA Registro Nacional Sanitario de Productores Agropecuarios

RMSE Raiz del Cuadrado Medio del Error

SENASA Servicio Nacional de Sanidad y Calidad Agroalimentaria Centro Regional Patagonia Norte

SS Sujeto-Específicos

SD Precisión de los Errores

UNCo Universidad Nacional del Comahue

ZEP_b Predictor de la Expansión FO

Parte I

INTRODUCCIÓN Y MOTIVACIÓN

Capítulo 1

INTRODUCCIÓN

1.1

PRESENTACIÓN DEL PROBLEMA

El conocimiento del peso final de frutos con antelación a la cosecha es una información de vital importancia por las consecuencias que esta tiene sobre las estrategias de manejo del monte frutal, las tareas de empaque y la logística de comercialización. Es por ello que la predicción de la distribución final del tamaño de frutos, a partir de muestreos previos a la cosecha, es una práctica usual en las principales regiones de producción frutícola.

Para el caso de los frutos de pepita, el proceso de crecimiento del diámetro del fruto a través del tiempo responde a un modelo sigmoideo en forma de S, que parte de algún punto fijo e incrementa su pendiente en forma monótona hasta alcanzar un punto de inflexión, en el que la pendiente comienza a decrecer y la función a aproximarse en forma asintótica al valor definitivo [88]. El modelo utilizado para describir este comportamiento ha sido ampliamente estudiado, Bramardi [114], en sus trabajos de medidas de no linealidad propone el modelo Logístico en la tercer parametrización según Ratkowski [103] para describir el crecimiento de frutos de pera de los cultivares William's y Packam's Triumph.

Existe además, un nivel de variabilidad estocástica en el crecimiento de los frutos que es gobernada por factores que no son fácilmente observables, como el sitio, la genética, las labores culturales y las condiciones climáticas. La acción simultánea de los factores antes mencionados generan la componente estocástica de la variabilidad que provee la razón fundamental para utilizar el enfoque de los modelos mixtos no lineales (NLME) los cuales permiten modelar además de los efectos fijos, los efectos aleatorios generados en el proceso. Los NLME se constituyen también como una propuesta superadora por razones de interpretabilidad, parsimonia y, por sobre todo, validación más allá del rango de datos observado [58].

Otro aspecto a considerar cuando nos encontramos con datos correspondientes a medidas repetidas, es decir a datos donde las mediciones fueron realizadas sobre un mismo individuo de manera repetida en el tiempo, es la presencia de correlaciones entre estas. Si la correlación serial es ignorada en el ajuste de los modelos, a pesar que los parámetros estimados permanecen insesgados, las varianzas estimadas del error de los parámetros resultan sesgadas. Sin embargo, estos estimadores del error sesgados son frecuentemente usados en diferentes pruebas de hipótesis para seleccionar el mejor conjunto de covariables por lo tanto, el ajuste del modelo es sensible a la especificación de la estructura de covarianza del error [76].

A través de parámetros fijos y aleatorios, los NLME tienen la ventaja de analizar estructuras de datos agrupados y modelar la matriz de covarianza de datos correlacionados, donde covariables son explicadas a través de los parámetros fijos como en los modelos no lineales de regresión tradicionales, y donde varias fuentes de heterogeneidad y datos correlacionados pueden ser considerados a través de la inclusión de efectos aleatorios.

MOTIVACIÓN

La información sobre los que se desarrolla la presente tesis es parte de una extensa labor desarrollada por los integrantes del Área de Estadística de la Facultad de Ciencias Agrarias de la Universidad Nacional del Comahue (UNCo). Los datos recopilados han sido material de estudio de varios proyectos de investigación y algunos de ellos, de la tesis de Maestría del Dr. Sergio Bramardi. Posteriormente estos se han integrado a la base de datos del Programa de Pronóstico de Producción de Frutos de Pepita de las Provincias de Río Negro y Neuquén (PPPRNN).

El PPPRNN consiste en un programa interinstitucional que tiene como objeto la estimación anticipada del volumen de producción potencial de los montes frutales implantados en las provincias de Río Negro y Neuquén. Dicho programa es parte de las actividades que el Área de Estadística, antes citado, realiza en convenio con la Secretaria de Fruticultura del Gobierno de la Provincia de Río Negro y el Ministerio de Desarrollo Territorial de la Provincia del Neuquén. El mismo cuenta, además, con la colaboración de instituciones gubernamentales como el Instituto Nacional de Tecnología Agropecuaria Alto Valle (INTA Alto Valle), el Servicio Nacional de Sanidad y Calidad Agroalimentaria Centro Regional Patagonia Norte (SENASA) y el apoyo de entidades no gubernamentales como la Cámara Argentina de Fruticultores Integrados (CAFI), la Cámara Argentina de la Industria y Exportaciones de Jugos de manzana, pera y afines (CINEX) y la Federación de Productores de las Provincias de Río Negro y Neuquén.

El aporte de SENASA se materializa a través de las fechas de plena floración y cosecha de las variedades en consideración dentro del PPPRNN. SENASA, a su vez participa con la información censal del Registro Nacional Sanitario de Productores Agropecuarios (RENSPA), una declaración jurada que realizan anualmente los productores frutícolas en relación al número y la composición de sus plantaciones.

El Programa en mención comprende tres áreas de trabajo en las que se releva información de carácter longitudinal y transversal de las variedades de interés, a partir de establecimientos productivos de la región. La información longitudinal y parte de la transversal proviene de chacras experimentales en las que se registra semanalmente el diámetro ecuatorial de frutos seleccionados a lo largo de ciclo de crecimiento de los frutos. La información transversal restante se obtiene en un operativo a campo que comprende una ventana de observación de dos semanas. La extensión territorial del operativo alcanza los 300 kilómetros e involucra la labor de 36 relevadores sobre 326 chacras y 2.782 unidades muestrales (cuadros).

Las predicciones del Programa se realizan para las variedades de manzanas Red Delicious (con sus clones mejorados), Granny Smith y Gala (con sus clones mejorados), y los cultivares de peras William's, Packham's Triumph y Beurre D'Anjou. Para las variedades tradicionales el relevamiento y la estimación de producción se realiza por los estratos de edad: 10 a 19 años, 20 a 29 años, 30 a 39 años y 40 o más años, mientras que para el caso de Gala y sus clones los estratos de edad comprenden los agrupamientos: 5 a 9 años, 10 a 19 años y 20 o más años. Se considera, además, el sistema de conducción de la parte aérea de las plantas (espaldera y monte libre).

Para la ejecución de las predicciones es necesario conocer para cada estrato tres elementos fundamentales: (i) la carga al momento del pronóstico de los árboles expresada en número de frutos por planta, (ii) el tamaño a cosecha del fruto medio, (iii) y la cantidad de plantas reales del estrato.

Para el primer requerimiento, carga de frutos en los árboles, se realiza el conteo total en dos árboles por unidad muestral (cuadros). El número de unidades muestrales dentro de cada estrato se determina en función de su variabilidad interna, permitiendo conocer la carga media por árbol. A efectos de la consideración final de la carga frutal se considera, además, el error incurrido por los relevadores al momento de contabilizar visualmente la carga a través de un factor denominado "fruto oculto", obtenido experimentalmente.

La segunda información referida al tamaño del fruto a la fecha de cosecha se obtiene, desde el año 2005, a partir de curvas de crecimiento especialmente desarrolladas para las condiciones de la región. En años anteriores esta estimación se realizaba aplicando un modelo probabilístico que relacionaba el peso de los frutos con la carga, que para las características de nuestra región, no resultaba ser el método más adecuado.

La tercer información: número de plantas reales de cada estrato o superficie efectiva ocupada por el estrato, se obtiene a partir los RENSPA suministrados por SENASA.

Con la información precedente, la estimación de producción de la variedad A en cada estrato región-sistema de conducción-edad está dada por:

$$P_A = \sum_{i=1}^n (NF_i PFM_i PL_i) / k \quad (1.1)$$

donde:

- *n*: Número de estratos
- *NF*: Número medio de frutos por árbol
- *PFM*: Peso estimado del fruto medio a cosecha
- *PL*: Número de árboles del estrato
- *k*: Factor menor que la unidad que depende de cada variedad y tiene por finalidad corregir el error que se produce durante el conteo de cargas por fruto oculto.

El resultado obtenido se extiende luego a todos los estratos de la variedad para obtener la estimación total de producción para la región. Esta información generada es acompañada con estimaciones de la distribución de tamaños de fruta a cosecha para cada una de las variedades.

DESCRIPCIÓN DEL TRABAJO

La predicción de los tamaños finales en cuanto a valor medio y su distribución se constituyen en una herramienta de vital importancia dentro de los elementos a considerar en los procedimientos predictivos. En el estado actual del arte, esta se obtiene a partir del comportamiento medio que presentan a lo largo de su desarrollo, frutos correspondientes a un mismo peso a cosecha a lo largo de las temporadas consideradas. La obtención de este valor requiere el ajuste de modelos correspondientes a funciones sigmoideas, no lineales en sus parámetros [114] para ciertas clases de categorías de tamaño. Estas categorías de tamaño generalmente se corresponden a clases comerciales de tamaño que actúan como criterio de estratificación a partir de las cuales se obtienen los comportamientos medios.

Una alternativa más abarcativa de estimación al método antes descrito consiste en un NLME, el cual, además de caracterizar el crecimiento típico de los frutos, contempla la posibilidad de incorporar efectos aleatorios a varios niveles de agrupamiento en los coeficientes, la modelación de la correlación del error y la incorporación de covariables que caracterizan la temporada de crecimiento.

En los NLME la modelación de los efectos fijos de la componente estructural se vale de medidas LONGITUDINALES que consisten en mediciones de diámetro ecuatorial obtenidas en los mismos frutos a lo largo de su ciclo de crecimiento. Sin embargo, las estimaciones de los efectos aleatorios así obtenidas están condicionadas a las restricciones impuestas por los procedimientos establecidos en la selección inicial de los frutos. Si los NLME tienen propósitos predictivos, es necesario generalizar a una población mas amplia. Por ello, además de las medidas LONGITUDINALES, es necesario considerar medidas TRANSVERSALES definidas como muestras aleatorias de diámetros ecuatoriales de frutos obtenidas en un reducido numero de ocasiones de ciclo de crecimiento [126].

Se ha comprobado que la inclusión de parámetros aleatorios en varios niveles de agrupamiento permite mejorar significativamente el ajuste en comparación a modelos que consideran sólo uno de ellos. Además, magnitudes significativas de variabilidad residual estocástica indican, cuando estas se encuentran presentes, que algunos efectos fijos son ignorados por el modelo, como factores relacionados con la genética, la topografía o el clima [96].

La disponibilidad de la información sobre estos factores es poco común en datos de crecimiento, por ello, los efectos aleatorios pueden actuar como una aproximación a estos efectos fijos que, o bien son desconocidos, o resultan difíciles de cuantificar.

Un modelo mixto con efectos aleatorios modelado a través de parámetros aleatorios puede aliviar, además, los problemas de heterogeneidad de varianza y autocorrelación residual presente en datos medidos repetidamente a lo largo del tiempo. Esto se debe indirectamente a las complejidades y compensaciones del modelado de la estructura de varianza-covarianza de los efectos aleatorios y, directamente, a través de una estructura más compleja de varianzas-covarianzas del término del error [30].

Sin embargo, los efectos aleatorios por si solos no pueden aliviar los problemas de heterocedasticidad y autocorrelación serial. Un motivo posible, de esta incapacidad radica en una ligera falta de flexibilidad en la forma del modelo adoptado para algunos frutos, por lo que es necesario además proponer y modelar una correcta estructura de la varianza residual [58].

Los efectos aleatorios son útiles además, no sólo para cuantificar la variabilidad en una población, sino también para obtener Estimadores Bayesianos Empíricos (EBEs) para cada uno de los efectos aleatorios en algún parámetro. Estos últimos pueden utilizarse para examinar grupos de subpoblaciones que pueden ser de interés o a los efectos de revelar la existencia de efectos fijos cuantificables no considerados y apreciados a través de asociaciones entre estos y posibles fuentes de variabilidad en alguno de los niveles considerados de los efectos aleatorios.

Las covariables ambientales surgen aquí como candidatas para caracterizar la combinación del efecto de sitio y ciclo de crecimiento, con un criterio de inclusión relacionado a factores que inciden en la fisiología del crecimiento de los frutos como son distintos rangos de acumulación de temperaturas o evapotranspiración, surgiendo así un gran número de posibles covariables a considerar.

Debido a que la evaluación de cada posible covariable en un modelo se torna prohibitiva cuando tenemos un gran número de candidatas, es necesario recurrir a métodos de cribado de variables. La metodología propuesta en la presente tesis, consiste en tratar a los EBEs de los efectos aleatorios como “datos” y utilizar métodos exploratorios para evaluar la relación entre los efectos aleatorios y las covariables de interés a través de métodos manuales y automáticos.

Cuando una covariable es introducida en un modelo, la variabilidad asociada con el parámetro afectado seguramente decrece debido a que los cambios en los valores de este son predichos con base a la nueva covariable. Ello obliga a reconsiderar la estructura de todo el modelo obtenido en cuanto a la estructura de varianza de los efectos aleatorios y el error.

Una vez construidos los modelos candidatos es posible calibrarlos para obtener predicciones de frutos no considerados en la etapa de modelado. Para cumplir con ese objetivo se deben obtener predicciones de los EBEs a partir de los parámetros estimados en los modelos propuestos e información disponible en los nuevos frutos. La predicción del diámetro final se obtiene a partir de los EBEs predichos, las estimaciones de los efectos fijos obtenidos de los modelos candidatos y los modelos en cuestión.

Finalmente es necesario evaluar la performance predictiva de los modelos candidatos a partir de métricas de validación externa a los efectos de seleccionar el más adecuado para los objetivos planteados.

1.4

OBJETIVO GENERAL

El objetivo central de la presente tesis radica en obtener un Modelo Mixto No Lineal con alta capacidad predictiva. Las medidas sobre las cuales se evalúa el modelo se basan en mediciones de diámetro ecuatorial de frutos de Peras Cv. William's obtenidos en fechas previas a la cosecha comercial.

1.5

OBJETIVOS ESPECÍFICOS

- Obtener modelos que consideren las fuentes de variabilidad identificadas de mayor relevancia.
- Mancomunar en un mismo modelo información de naturaleza LONGITUDINAL y TRANSVERSAL.
- Evaluar la incidencia de factores ambientales en la respuesta diferencial de los frutos a lo largo de las temporadas de crecimiento consideradas.
- Evaluar métricas adecuadas que posibiliten proponer modelos candidatos con un alto poder predictivo.

1.6

METODOLOGÍA

Metodológicamente esta tesis descansa en el desarrollo de un plan de actividades contenidas en tres fases. La primera consiste en la descripción del cultivo y el área de estudio, junto al análisis de los fundamentos matemáticos y estadísticos de los modelos mixtos no lineales; la segunda, en la descripción de la información histórica disponible ya sea de carácter longitudinal o transversal, su adecuación y caracterización; la tercera se

refiere específicamente a la utilización de las técnicas propuestas en los datos recopilados a través de los entornos computacionales adecuados con propósitos predictivos.

1.7

ALCANCES DEL TRABAJO

Este trabajo propone obtener modelos mixtos no lineales con el objeto de utilizarlos en predicciones anticipadas del tamaño de frutos de Pepita cultivar William's en los Valles Irrigados de la provincia de Rio Negro en nuevos ciclos productivos. La información disponible consiste en diámetros de frutos identificados a lo largo de varios ciclos de crecimiento en diversas localidades de la provincia de Rio Negro. La evaluación de los modelos obtenidos a través de métricas de validación externa posibilita seleccionar el modelo más adecuado para propósitos de pronóstico en ciclos de crecimiento no considerados en la etapa de modelado.

Capítulo 2

DESCRIPCIÓN DEL CULTIVO Y EL ÁREA DE ESTUDIO

2.1

LOCALIZACIÓN GEOGRAFICA

El territorio Nordpatagónico comprende una considerable extensión de geografía de las provincias de Río Negro y Neuquén.

Se trata de una planicie aluvial disectada con una suave pendiente hacia el Este, una zona de poca altura con una media de 300 a 600 metros. Su génesis es aluvial y está relacionada con superficies de erosión, como terrazas fluviales, sedimentos o ambos en conjunto. Los materiales se caracterizan como conglomerados sueltos de sedimentos continentales, sedimentos marinos, rodados patagónicos y depósitos aluviales de origen Terciario.

Este paisaje se caracteriza por la presencia de dos valles de origen fluvial, recorridos con orientación Oeste–Sureste por los ríos Negro y Colorado, que atraviesan la planicie aluvial disectada.

La región del Alto Valle de Río Negro y Neuquén, que forma parte de uno de esos valles,

se extiende por ambos márgenes de los ríos Neuquén, Limay y Negro, con un potencial de tierras factibles de riego de 1.824.000 hectáreas.

El Alto Valle se halla ubicado en los departamentos de General Roca, Provincia de Río Negro, y Confluencia, Provincia de Neuquén. El ancho del valle es variable en toda su extensión, oscilando entre los dos y los ocho kilómetros, y posee una extensión de aproximadamente ciento veintitrés kilómetros.

2.2

SUELO, CLIMA Y MEDIO AMBIENTE

Los suelos del Alto Valle se caracterizan por ser áridos y alcalinos. Éstos son grises y prácticamente no han sufrido meteorización alguna, excepto aquellos que se hallan al alcance de las inundaciones de los ríos y los que han sido sometidos a cultivos. A su vez son suelos aluviales sin formación precisa de su perfil tipo: se originaron a partir de materiales aluviales y coluviales, siendo heterogéneos en cuanto a su textura y estratificación. Es posible, además, su caracterización según su distribución textural en suelos de barda, media barda, media costa y costa, con texturas que van desde una gama de arenas gruesas o limos, limosos francos, franco arenosos y arenosos respectivamente [7].

En cuanto a su fertilidad, son saturados en bases y muy pobres en materia orgánica. Las limitantes de la producción son el factor hídrico, el nitrógeno y, en menor escala, el fósforo.

La temperatura media oscila en los 14 °C y las precipitaciones se incrementan de Oeste a Este, con una mínima de 200 milímetros en la parte occidental y de 350 mm milímetros en el sector oriental.

El clima del área de estudio se clasifica como continental, templado y árido, con lluvias deficientes durante todo el año, denominado xerofítico seco [51] y de acuerdo a Thornthwaite [16] es árido mesotermal, caracterizado por escasas precipitaciones y ele-

vada evapotranspiración.

La vegetación predominante es la de tipo ribereña y la implantada por el hombre, ya que se trata principalmente, de valles irrigados destinados a la producción agrícola. En la meseta predomina la estepa arbustiva media, con especies como la jarilla (*Larrea divaricata* y *Larrea cuneifolia*) y el matasebo (*Monttea aphylla*). En algunos sectores se encuentra alpataco (*Prosopis alpataco*) y zampa (*Atriplex lampa*).

2.3

EVOLUCIÓN HISTORICA DEL CULTIVO

2.3.1

ORIGEN E HISTORIA DE LA PERA WILLIAM'S

El cultivo del peral se inicia en el oeste asiático, en los valles formados por la cadena montañosa Tian Shan en China hace unos 3.000 años, aunque escritos de Homero (700 aC) sugieren una posible procedencia europea. La primera especie domesticada fue *Pyrus pyrifolia*, peral silvestre que producía frutos comestibles. Posteriormente se realizaron hibridaciones entre *P. ussuriensis* y *P. pyrifolia*. En Occidente, *P. communis var pyraster* y *P. communis var. caucasica* fueron seguramente los ancestros del peral europeo. El género *Pyrus* incluye 22 especies divididas en occidentales y orientales según su situación geográfica. Las más utilizadas en la actualidad son *P. betulifolia*, *P. calleryana*, *P. pyrifolia* y *P. ussuriensis* (orientales), y *P. communis*, *P. communis var. pyraster*, *P. amygdaliformis*, y *P. salicifolia (occidentales)*[25]. La pera William's fue descubierta originalmente en 1765 en Inglaterra por Sir Stair. Posteriormente, un viverista llamado William's la adquirió y diseminó por toda Inglaterra. Su nombre completo es Williams Bon Chetrien, cuya traducción podría ser "Williams el buen cristiano". Hacia 1799, James Carter introdujo esta pera en los Estados Unidos estableciendo el cultivo en la propiedad de Thomas Brewer en Roxbury, Massachussets. Más tarde, Enoch Bartlett de Dorchester (Massachussets) adqui-

rió la variedad y, desconociendo su denominación, la propagó por Estados Unidos bajo su propio nombre. Fue recién en 1928 que se tomó conciencia que Bartlett y William's eran la misma entidad, cuando se introdujeron a ese país nuevas plantas procedentes de Inglaterra. El importador de esta variedad en Argentina fue probablemente el comerciante inglés James Brittain hacia el año 1817, quien la cultivó en su finca a orillas de la desembocadura del Riachuelo en Buenos Aires.

2.3.2

EL CULTIVO EN LA REGIÓN

En el Alto Valle, el cultivo y propagación del peral de esta variedad tendría sus probables orígenes en pioneros inmigrantes de principios del Siglo XX como el suizo Wilhem Gaspar Kopprio, quien instaló un vivero en la zona de Allen hacia 1910. Cuentan que Kopprio trajo de Francia un pie de membrillero de Angers en el cual injertó la mayoría de sus perales. Una crónica lo coloca entre los pioneros de la actividad, cuya especialidad fue la pera William's Bon Chretien que él propagó y llamó "la mejor pera de agua". Kopprio definía a esta variedad como "grande y mantecosa y excelente para el comercio" en su primer catálogo de plantas (1926). Es hacia la década de 1920 cuando la Estación Experimental Río Negro (transformada luego en el Instituto Nacional de Tecnología Agropecuaria Alto Valle) y la Estación Agronómica Cinco Saltos implantan la primera colección de 24 variedades de manzano sobre portainjerto Northern Spy, para contrarrestar la incidencia del pulgón lanífero, que había aniquilado las plantaciones injertadas sobre pie franco de semilla. También en ese año se implanta la primera colección de vides "de almería y americanas". En 1926 se amplía la colección de frutales de pepita. Las variedades de peral sumaban 24, siendo la principal la William's Bon Chretien, en tanto que se aumentó a 27 la de manzanos, con Jonathan y Delicious como las más importantes.

IMPORTANCIA COMERCIAL

En la actualidad, el Alto Valle de Río Negro y Neuquén es la principal zona exportadora de pera a nivel mundial, con una superficie de 46.104 hectáreas netas con cultivos de fruta de pepita. En ellas, el cultivo de pera William's ocupa 9.545 hectáreas, siendo la principal variedad producida (20,7% de la superficie).

La mayor proporción de la superficie de pera William's se encuentra en el rango de edad productiva y sólo el 15,85% tiene más de 40 años. El ingreso paulatino de nuevas áreas a la etapa productiva explica el crecimiento de la producción de la variedad. En particular, la temporada 2008-2009 significó un incremento de algo más de 70.000 tn, totalizando 344.000 tn de pera William's, equivalentes al 45,5% de la producción total regional de la especie[2].

La actividad frutícola es la mayor fuente de ingresos de la población de la zona, conteniendo unos 2.600 productores, cerca de 50.000 empleados rurales, 250 plantas de empaque y 196 frigoríficos de frutas .

El principal destino de la producción es la exportación. El 61% de la producción ingresa a los mercados del mundo, el 17% se comercializa en el mercado doméstico y el 22% se industrializa para la producción de jugo concentrado como producto central. Los principales mercados de exportación son Brasil, Rusia, Italia y Estados Unidos [2].

Parte II

REVISIÓN DE LITERATURA

Capítulo 3

ASPECTOS AGRÓNOMICOS DEL CRECIMIENTO

3.1

INTRODUCCIÓN

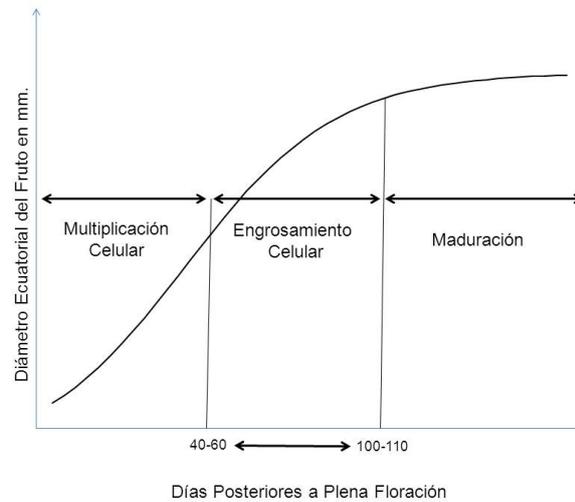
El crecimiento del fruto que resulta de la división y alargamiento celular sigue un modelo representado por una curva sigmoidea cuando se grafica el diámetro ecuatorial del fruto en función del tiempo (Figura 3.1).

Ésta presenta tres fases bien definidas:

- Fase I o período de crecimiento exponencial (multiplicación celular).
- Fase II o período de crecimiento lineal (agrandamiento y expansión celular).
- Fase III que incluye la etapa final del crecimiento y el período de maduración.

El fruto a lo largo de su ciclo incrementa su tamaño de acuerdo a patrones específicos de crecimiento. El tamaño final del fruto en perales depende del número de células al momento del cuaje del fruto, del número posterior de divisiones celulares y de la expansión celular. La expansión celular a lo largo de ciclo de crecimiento se incrementa de manera

Figura 3.1: Curva de Crecimiento del Fruto.



exponencial con el tiempo en peras cultivar William's [6].

En el caso de frutos de pepita en general, las fases de crecimiento se caracterizan por 2 períodos diferentes: un primer período de división celular que típicamente alcanza hasta los 45 días luego de la antesis, seguido de un período de expansión celular que abarca hasta la cosecha. La división celular en peras se establece para los cultivares tempranos entre 25 y 30 días de la floración y en cultivares tardíos hasta 45 días [15]. En algunos ensayos en peras cultivar William's, la división celular se registró hasta 7 u 8 semanas de la plena floración, a partir del cual se identificó la fase de expansión celular donde el fruto experimentó el mayor aumento de volumen [48].

A partir de ese momento se desencadenan en el fruto una serie de transformaciones bioquímicas que conducen al desarrollo de las características organolépticas de la variedad, esta última suele durar de 10 a 30 días.

En la zona del Alto Valle, la edad media del fruto (desde plena floración hasta fecha media de cosecha de los últimos 10 años) es de 106 días, es decir que a fines comerciales los frutos son cosechados antes de la fase final para preservar la calidad de conservación [21].

El crecimiento de los frutos está en gran medida condicionado por factores propios de la planta (internos), ambientales y de manejo (externos).

3.2

FACTORES INTERNOS QUE AFECTAN EL CRECIMIENTO

Tamaño inicial del fruto

El número de células definido en la primera fase de crecimiento determinará el tamaño potencial del fruto. La cantidad de células en el pequeño fruto está regulada por sustancias hormonales y es afectada por la disponibilidad de agua y la provisión de hidratos de carbono y nutrientes. En general, los frutos chicos contienen menos células y más pequeñas que los de tamaño grande. Se ha comprobado que los frutos que durante la primera fase de la curva de crecimiento son grandes llegan a cosecha con tamaños mayores que los que partieron con dimensiones pequeñas.

Número de semillas

El tamaño del fruto es proporcional al número de semillas maduras viables y se debe a la cantidad de hormonas (citocininas, auxinas, giberelinas) producidas por cada semilla. Además de su particularidad de desarrollar frutos sin semilla, la pera William's posee naturalmente un contenido de hormonas más elevado. Por lo tanto, la cantidad de semillas no condiciona significativamente el tamaño final de los frutos.

Estado nutricional de las yemas

La energía inicial para el crecimiento vegetativo y la floración proviene de las reservas alimenticias de la planta. Cuando éstas son escasas, la división celular en los pequeños frutos también cesa en forma temprana limitando el crecimiento potencial posterior de los frutos.

Relación hoja-fruto

Esta relación indica la cantidad de frutos de una planta en función de la superficie foliar fotosintéticamente activa. La disponibilidad de sustancias hidrocarbonadas depende de la superficie foliar y de la intensidad de fotosíntesis. Los frutos actúan como destino para estas sustancias y con frecuencia se produce competencia entre ellos. La relación normal de hojas-fruto que determinan cantidad y calidad de cosecha es de aproximadamente 50 hojas por fruto.

Carga frutal

La cantidad de frutos y el tamaño que alcanzarán están en estrecha relación con el vigor y estructura de la planta. Altos rendimientos asociados a alta carga de fruta por planta conllevan a la formación de frutos más pequeños y el inicio de una serie de procesos fisiológicos que conducen al añerismo (una importante merma en los rendimientos) en la siguiente temporada.

Portainjerto

El portainjerto es el proveedor de agua y nutrientes y actúa como condicionante del vigor de la planta, lo que afecta el tamaño y la calidad de los frutos.

Ubicación de la estructura fructífera

El tamaño potencial de los frutos está condicionado por su ubicación dentro de la planta, en función de la disponibilidad de luz y por lo tanto de los fotoasimilados. Los frutos de las partes bajas y medias del árbol (más sombreadas) en general son más chicos. Se debe considerar, además, que la inflorescencia del peral es un racimo donde la flor central es la última en formarse y por lo tanto la más pequeña. Las flores laterales tienden a producir frutos más grandes que las centrales.

3.3

FACTORES EXTERNOS QUE AFECTAN EL CRECIMIENTO

Los factores del medio ambiente regulan el crecimiento del fruto afectando los procesos metabólicos fundamentales de los árboles frutales como la fotosíntesis y la respiración. El crecimiento de los frutos puede resultar limitado por el suministro de agua y nutrientes y por otros factores que afectan la fotosíntesis, como la baja temperatura y la nubosidad.

Temperatura

La temperatura ambiental es un regulador de los procesos fisiológicos a nivel de las reacciones bioquímicas que intervienen en el crecimiento directa o indirectamente. Si bien tiene un efecto continuo durante todo el ciclo del crecimiento, se destacan dos momentos clave que afectan el tamaño final del fruto.

El primero es el período de división celular, ya que durante éste se establece el número total de células que determinan el tamaño potencial de los frutos. El segundo es la etapa de elongación celular, dos a tres semanas antes de cosecha.

El rango de temperaturas óptimas para el crecimiento vegetativo es de aproximadamente

25 grados centígrados. Dentro de este rango y bajo condiciones adecuadas de intensidad lumínica e hídrica el proceso fotosintético se ve favorecido, lo que contribuye al crecimiento del fruto.

No obstante, cuando las temperaturas ambientales se alejan del óptimo, los árboles frutales han demostrado tener una cierta capacidad de aclimatación basada en una compensación y reajuste de procesos fisiológicos tales como respiración, transpiración, transporte de nutrientes, etc. Estas rectificaciones presentan un costo de energía que será restada a la producción de tejidos en el fruto y a características como color, sabor, etc. Cuando las temperaturas se alejan del óptimo, la influencia de otros factores limitantes se incrementa. Altas temperaturas asociadas a bajos tenores de agua en el suelo reducen la capacidad de refrigeración de la planta a través de la transpiración y por lo tanto la posibilidad de aclimatación.

La temperatura es determinante en la tasa de crecimiento de los frutos [42]. En base a ello se han creado unidades de desarrollo de los frutos o unidades de grados día, que se calculan promediando las temperaturas máximas y mínimas de cada día y restando la temperatura base para el desarrollo que para el caso de peras es de 10 grados centígrados, posteriormente se suman diariamente los resultados desde la floración hasta la cosecha [91].

En zonas templadas, las temperaturas medias de 15 a 16 grados son limitantes, fundamentalmente a comienzos de ciclo. Bepete et al. [72], confirma un aumento de la tasa de crecimiento al incrementar la temperatura mínima en la fase temprana, aunque observaron una importante abscisión de los frutos. Otros ensayos [47] mostraron que la tasa de crecimiento del fruto se incremento hasta 8 veces cuando las temperaturas medias aumentaron de 6 a 18 grados, durante los días posteriores al cuaje, por otro lado cuando se presentan temperaturas bajas inferiores a 10 grados, la división celular se prolonga mas alla de los 40 días posteriores a la antesis.

Luz

La productividad de un monte frutal depende de la eficiencia de la absorción de la luz y la distribución de los carbohidratos. La intercepción de la luz está determinada por la cantidad y disposición de las hojas, el modo de conducción, el marco de plantación, la orientación de las filas y la latitud del lugar. El sombreado limita la capacidad fotosintética de las hojas. Se ha observado que los frutos cercanos a las hojas bien iluminadas son más grandes y tienen mayores probabilidades de alcanzar su potencial de crecimiento. Las hojas de los dardos, localizadas cerca del fruto, cumplen un rol relevante para las etapas tempranas del desarrollo del fruto en la estación de crecimiento.

El sombreado afecta también el desarrollo reproductivo del año, la formación de flores y el cuaje de los frutos del año siguiente. La diferenciación de las yemas de flor no se produce con niveles de luz inferiores al 10 ó 30 por ciento de la luz disponible. Aunque ocurra la diferenciación, la calidad de la yema, el tamaño de la bolsa o del dardo y el cuaje del fruto dependen en forma proporcional de la cantidad de luz directa recibida. Para incrementar la producción es importante maximizar la intercepción lumínica temprana en la estación de crecimiento.

Viento

El viento influye negativamente en el crecimiento del fruto porque afecta la temperatura de la planta, y la pérdida de agua daña tejidos a través de la deshidratación y la fricción mecánica. Vientos moderados pero continuos y altas temperaturas aumentan la tasa de transpiración de las hojas, lo que ocasiona un déficit hídrico a nivel celular. En estas condiciones, el movimiento de agua a los frutos disminuye y provoca una merma en la tasa de crecimiento. El viento también provoca plegado de hojas y daños a nivel epidérmico que alteran la capacidad fotosintética.

Poda

A través de esta práctica se busca aumentar la intercepción y distribución lumínica en el árbol y por lo tanto afectar el tamaño del fruto y maximizar el potencial productivo de la planta. La poda invernal permite controlar el vigor del árbol y mejorar la distribución y disponibilidad de reservas nutricionales. La reducción parcial de las ramas del año durante el verano no afecta el rendimiento general, pero si es muy severa disminuye el tamaño de los frutos al modificar la superficie foliar fotosintéticamente activa.

Fertilización

El crecimiento del fruto no puede progresar sin los nutrientes necesarios para producir los componentes químicos de la planta. La deficiencia de un elemento esencial afecta el crecimiento de la planta y por lo tanto del fruto.

Riego

La fase de elongación celular se caracteriza por una alta tasa de crecimiento debida a la acumulación de agua y sustancias hidrocarbonadas en las células. Por estas razones, durante esa fase es notable el aumento de volumen y peso del fruto y es fundamental la disponibilidad de agua y la capacidad fotosintética de la planta. En ensayos de riego se ha observado una disminución del tamaño del fruto cuando se realiza un estrés hídrico entre los 42 y 188 días después de plena floración (DDPF). Evaluaciones regionales indican un crecimiento compensatorio del fruto en los tratamientos en los que son regados después de un déficit hídrico realizado durante los 60 a 70 DDPF, alcanzando el fruto tamaños finales similares a aquellos que no han recibido restricciones hídricas.

Raleo

Los cultivares de peras Europeas y en particular el cultivar William's poseen, al comienzo de la estación de crecimiento, un excesivo número de flores que determinan tamaños de frutos pequeños a cosecha y una importante reducción en el retorno a floración del próximo año [121]. Para reducir la carga frutal las especies frutales desarrollan mecanismos de autorregulación de caída de frutos en un período temprano de ciclo productivo [34].

El calibre de los frutos puede ser optimizado por una apropiada nutrición, poda y raleo de los frutos aunque los rendimientos sean reducidos por esta práctica. El raleo químico es una de las prácticas culturales más importantes en la producción de frutos [34].

En general, un raleo temprano en la estación de crecimiento incrementa el tamaño de los frutos no raleados debido principalmente a que se estimula una mayor cantidad de células en el fruto. Esta práctica también es importante en etapas posteriores ya que mejora la relación hoja-fruto al disminuir la carga frutal, optimizando la distribución de nutrientes y atenuando la competencia.

3.4

BASES FISIOLÓGICAS DEL CRECIMIENTO DE FRUTOS

Los factores internos y externos que afectan el crecimiento de los frutos pueden ser explicados en función del balance de carbono, una medida de la capacidad potencial que tienen los frutos para desarrollarse.

El suministro de asimilados en el fruto está determinado en la mayoría del período de división celular, por la disponibilidad de reservas almacenadas en la planta y por los fotoasimilados de los dardos y las hojas próximas al fruto [34], en este momento se

genera una competencia entre los frutos y los brotes en crecimiento. En la primavera, cuando las plantas frutales poseen reservas de carbohidratos y nutrientes, los frutos ven restringido su tamaño por la temperatura, durante el verano las reservas de carbohidratos son limitantes y las temperaturas superan ampliamente los umbrales de desarrollo del fruto [72]. La disponibilidad de fotosintatos y el suministro de carbono al comienzo de la estación durante la división celular es crucial para el crecimiento del fruto en las etapas restantes, y está directamente relacionado al tamaño final de los frutos.

El crecimiento de un fruto individual es dependiente del suministro y la adquisición de carbono. El desarrollo del fruto es un proceso no autotrófico e importa la mayoría del carbono de las hojas adyacentes. Las hojas adyacentes asimilan dióxido de carbono y lo convierten en sorbitol para ser transportados a nuevas hojas, a los frutos u otros destinos [81]. Las fuentes de carbono importadas por el fruto dependen fundamentalmente de la fuerza que el fruto posee como destino. Para comprender la regulación de la partición de carbono por los frutos, se debe considerar la capacidad de los destinos llamada la fuerza del destino que determina dicha regulación. Por otro lado se ha observado que la fotosíntesis del fruto puede contribuir con cantidades importantes de carbohidratos particularmente en los comienzos de ciclo de crecimiento [81].

3.5

PREDICCIÓN DEL TAMAÑO DEL FRUTO

Entre los factores que definen la calidad de los frutos, el tamaño reviste una característica de gran importancia económica, diferencias en pocos gramos en el peso medio de los frutos representan importantes pérdidas en los retornos por exportación. Las diferencias en tamaños entre los frutos, se debe principalmente al número y tamaño de las células de los frutos.

En el cultivar William's es importante disponer, al inicio de la cosecha, de una elevada

proporción de frutos con tamaños apropiados a los requerimientos comerciales. El manejo tradicional de la cosecha que se realiza en la región del Alto Valle, es efectuando varias “pasadas” seguidas de un riego para incrementar el tamaño de los frutos, esto incrementa los costos. Además, se deteriora la calidad de la producción por el avance de la madurez de los frutos recolectados en las últimas “pasadas”. Apuntar a realizar la cosecha en dos “pasadas” debería ser considerado como una meta de máxima. Al escenario descrito se suma la concurrencia en las plantas de empaque, a fines del mes de enero, de importantes volúmenes de manzanas Royal Gala y otros clones mejorados, lo que complica sustancialmente su manejo.

En la fruticultura actual, los programas de raleo de frutos son de vital importancia para la obtención de fruta de alta calidad que permita al productor competir en el mercado internacional [34]. El raleo remueve parcialmente las fuentes de giberelinas (semillas) que previenen la formación de yemas de flor. Eliminando una parte de las flores y los frutos se logra incrementar el tamaño de los frutos restantes y aumentar la relación de hojas por fruto, que debería ser de 30 a 40. También se mejora la exposición de los frutos a la luz y se logra una nutrición adecuada, se reduce el daño por enfermedades y plagas y se mejora la distribución del peso en la estructura del árbol. El raleo puede ser efectuado en forma manual o química. Utilizar productos químicos es el método más simple para hacer el trabajo en un corto período de tiempo y ahorrar mano de obra. Estos productos provocan abscisión, lo que altera la fisiología de la planta y el resultado productivo.

Los mayores efectos se logran con un raleo temprano en la estación de crecimiento, ya que se estimula la división y alargamiento celular y se obtienen frutos más grandes. Cuanto más tardíamente se realiza el tratamiento, menor es su efecto. La eliminación de frutos más chicos mejora el calibre porque reduce la competencia y aumenta el desarrollo de los frutos remanentes, además de evitar la presencia de frutos pequeños en la cosecha. La práctica de raleo manual de frutos se efectúa apuntando principalmente a dos fines:

- Equilibrar la relación hojas/frutos cuando el “cuaje” es muy alto y se observa una exagerada cantidad de frutos en relación a la cantidad de hojas. En este caso el raleo manual permitirá que las hojas alimenten en forma apropiada los frutos remanentes y estos puedan crecer con normalidad, de acuerdo con la tasa de crecimiento de esta variedad.
- Racionalizar la cosecha al eliminar temprano en la primavera los frutos deformados, manchados y los de tamaño pequeño que no van a llegar a alcanzar el tamaño mínimo requerido a nivel comercial. De este modo se facilita la cosecha y se reducen los costos.

En la programación del raleo manual de William's es posible utilizar la información generada por el seguimiento sistemático de las curvas de crecimiento de los frutos. Considerando los factores internos y externos que condicionan el patrón de crecimiento, y recordando que en general los frutos que desde el inicio de la curva son grandes llegan a cosecha con tamaños mayores, es posible seleccionar en una etapa temprana de ciclo productivo los frutos que alcanzarán los tamaños de interés comercial para el productor. Por otra parte el conocimiento anticipado de la distribución de tamaños posibilita la confección de estrategias de manejo en post-cosecha de empaque, conservación y la logística de comercialización ya que el precio diferencial por frutos de distintos tamaños comerciales es muy marcada. La organización de una logística de comercialización acorde a los tamaños predichos crea ventajas competitivas importantes posibilitando reorientar acuerdos comerciales previos a la recolección de los frutos [22].

Capítulo 4

ASPECTOS ESTADÍSTICOS

4.1

MODELO DE REGRESIÓN LINEAL CLÁSICO

En el marco de una regresión generalmente consideramos una respuesta y (escalar), y un vector asociado de covariables X , pudiendo formalmente definirlo como:

- y : variable de respuesta dependiente (escalar).
- X : covariable o predictor, una variable independiente de dimensión $(p \times 1)$, con x pudiendo incluir la constante "1" para contemplar una ordenada al origen.

Donde los valores de X pueden ser establecidos por un experimentador o pueden ser producto de una observación, en ambas situaciones se considera que los valores son obtenidos sin error, mientras que Y es una variable aleatoria.

Es decir que dado X , diferentes valores de y puede ser observados, lo cual puede deberse a varias razones:

- El error de medición del instrumento utilizado.
- La variación propia del muestreo aleatorio de poblaciones.
- Las fluctuaciones propias del proceso biológico o físico en estudio.

Siendo posible que de hecho, en muchas aplicaciones, estas razones se produzcan de manera simultánea pero con diferentes magnitudes en cada una de ellas.

Esta situación es susceptible de ser conceptualizada suponiendo que para cada valor de X poseemos una función de distribución de probabilidades que caracteriza los posibles valores observados de la variable y . Sobre dicha base conceptual, el modelo de regresión usual se aboca en obtener un modelo estadístico que considera los valores de y condicionados a los valores de X .

Si expresamos los pares ordenados observados como (y_j, X_j) , con $j = 1, \dots, n$, podemos pensar a estos como realizaciones de la función de distribución conjunta (y, X) en el que el evento X es observado o bien, suponiendo a y_j como realizaciones de la función de distribución y para el valor fijo X_j .

El modelo de regresión lineal expresado para los pares (y_j, X_j) en notación matricial es:

$$Y = X\beta + e \quad (4.1)$$

con $e = (e_1, e_2, \dots, e_n)^T$. Los desvíos $e_j = y_j - X_j^T \beta$ representan el hecho que la relación entre los valores observados para Y y X no siguen exactamente la relación dictada por $X\beta$. Mas bien los valores observados se desvían de esta relación presumiblemente por una o mas de las razones antes citadas. Además del modelo propuesto es necesario establecer ciertos supuestos para poder desarrollar las correspondientes inferencias, pudiendo estos enumerarse de la siguiente manera:

1. La esperanza de e_j es igual a cero: $E(e_j | X_j) = 0$.
2. El modelo $X\beta$ es correcto, es decir pasa por los puntos medios de la distribución de Y para cada uno de los valores de x_j .
3. Los e_j se encuentran idénticamente distribuidos para todos los j , independientemente de los x_j , con $var(e_j) = \sigma^2$.

4. Los e_j son independientes.

5. los e_j se encuentran normalmente distribuidos.

Bajo estos supuestos, los e_j son variables aleatorias normales independientes, lo que implica que los pares (y_j, X_j) son independientes. De lo dicho anteriormente podemos deducir que:

$$E(y_j | X_j) = X_j^T \beta \quad (4.2)$$

$$\text{VAR}(y_j | X_j) = \sigma^2 \quad (4.3)$$

$$Y \sim N(X\beta, \sigma^2 I_n) \quad (4.4)$$

Si β es conocido, entonces el modelo provee una caracterización completa de la respuesta en promedio, pero si β es desconocido debe ser estimado. El método de estimación usual bajo los supuestos antes citados es el de mínimos cuadrados ordinarios(OLS).

El estimador $\hat{\beta}_{OLS}$ se define como:

$$\hat{\beta}_{OLS} = (X^T X)^{-1} X^T Y = \left(\sum_{j=1}^n X_j X_j^T \right)^{-1} \sum_{j=1}^n X_j y_j \quad (4.5)$$

Que es aquel estimador que hace mínimo la función objetivo:

$$\sum_{j=1}^n (y_j - X_j^T \beta)^2 \quad (4.6)$$

Bajo todos los supuestos antes citados $\hat{\beta}_{OLS}$ es también el estimador máximo verosimil, por lo que la siguiente verosimilitud resulta maximizada por:

$$\log L = -(n/2) \log 2\pi - (n/2) \log \sigma^2 - (1/2) \sum_{j=1}^n (y_j - X_j^T \beta)^2 / \sigma^2 \quad (4.7)$$

4.2

MODELOS DE CRECIMIENTO DE FRUTOS

4.2.1

INTRODUCCIÓN

En condiciones de producción y de temperaturas normales, el crecimiento de órganos vegetales, y en particular el de los frutos, se representa mediante modelos sigmoideos. Los modelos sigmoideos son funciones en forma de S, caracterizadas por una asíntota superior, un punto de inflexión y una asíntota inferior. La asíntota superior indica el tamaño máximo del fruto, por otro lado el punto de inflexión representa el punto donde la aceleración del proceso cambia de una velocidad creciente a una decreciente, asimismo la duración y las características del proceso varían según la especie pero también según la variedad y los factores agronómicos y ambientales [87]. Las curvas de crecimiento de frutos pretenden modelar el comportamiento del organismo desde el momento en que se ha producido la fructificación y se inicia el proceso de división celular hasta alcanzar su plena madurez.

4.2.2

MODELO DE REGRESIÓN NO LINEAL

En la elección de un modelo de regresión no lineal para describir como el diámetro ecuatorial varia en función de los DDPF, es posible recurrir a modelos lineales en los parámetros, como los polinómicos. A través del incremento en el orden del modelo polinómico es posible incrementar el grado de certeza en la aproximación a la verdadera función de regresión no lineal, dentro del rango de datos observado. Sin embargo, estos modelos empíricos solo se basan en la relación observada y no incluyen consideraciones teóricas sobre el mecanismo subyacente que produce los datos [58].

Los modelos no lineales, por otra parte, se basan con frecuencia en un modelo que expli-

ca el mecanismo por el cual los datos se producen, consecuencia de ello los parámetros del modelo generalmente tienen una interpretación en términos físicos. Un modelo no lineal generalmente también usa menos parámetros en comparación a un modelo lineal competidor, brindando una descripción de los datos mas parsimoniosa. Estos, además, proveen predicciones mas creíbles para la variable de respuesta por fuera del rango de datos ajustado [80].

Si en vez de postular el modelo de la forma: RESPUESTA=MODELO+DESVIACIÓN, junto con una serie de supuestos sobre los desvíos, lo hacemos en términos de lo que esperamos decir en relación a los dos primeros momentos de la distribución de los y_j dado X_j podríamos proponer un modelo de la forma:

$$E(y_j | X_j) = f(X_j, \beta) \quad (4.8)$$

$$VAR(y_j | X_j) = \sigma_j^2 \quad (4.9)$$

Donde:

- $f(X_j, \beta)$ es una función de valores reales del vector de covariables $X_{r \times 1}$ y el vector de parámetros de regresión $\beta_{p \times 1}$. Esta función no necesariamente debe ser lineal y puede depender del vector de parámetros de manera no lineal.
- La expresión $VAR(y_j | X_j) = \sigma_j^2$ implica que las varianzas de la distribución de y condicionada a los valores de X_j puede variar a lo largo de los j .

El modelo postulado permite responder a cuestionamientos concernientes a valores del vector de parámetros β , por lo que el interés se centra en la respuesta media como una función de X_j , liberándonos del condicionamiento impuesto por los supuestos (2) y (3) del modelo de regresión lineal clásico.

La posibilidad de tener una varianza no necesariamente constante en el modelo habilita el camino para que esta sea susceptible de ser modelada directamente a través de una

nueva función de la forma:

$$\text{VAR}(y_j | X_j) = \sigma^2 g^2(\beta, \theta, X_j) \quad (4.10)$$

Donde g es una función de varianza que caracteriza la varianza en función de covariables contenidas en X_j y parámetros β que caracterizan a $E(y_j | X_j)$ y parámetros adicionales θ . Esta función g se convierte en una manera flexible de acomodar características propias de la variabilidad los datos a través de distintas elecciones de g .

La función objetivo a ser minimizada es:

$$\sum_{j=1}^n w_j (y_j - f(X_j^T, \beta))^2 \quad (4.11)$$

El estimador de β resultante es llamado estimador mínimo cuadrático generalizado ($\hat{\beta}_{GLS}$) y es producto de resolver la ecuación de estimación p-dimensional:

$$\sum_{j=1}^n w_j (y_j - f(X_j, \beta)) f_{\beta}(X_j, \beta) = 0 \quad (4.12)$$

Siendo $f_{\beta}(X_j, \beta)$ un vector de dimension $p \times 1$ cuyos elementos son las derivadas parciales de f con respecto a cada componente de β y $w_j = g^{-2}(\beta, \theta, X_j)$. Por otra parte, es posible obtener estimadores máximo verosímiles del modelo propuesto $\hat{\beta}_{ML}, \hat{\sigma}_{ML}, \hat{\theta}_{ML}$ suponiendo normalidad y como resultado de maximizar la siguiente función de verosimilitud:

$$\log L = -(n/2) \log(2\pi) - n \log \sigma - \sum_{j=1}^n \log g(\beta, \theta, X_j) - (1/2) \sum_{j=1}^n g^{-2}(\beta, \theta, X_j) (y_j - X_j^T \beta)^2 / \sigma^2 \quad (4.13)$$

MODELOS DE CRECIMIENTO SIGMOIDAL

Los parámetros de la familia de curvas sigmoideas y reparametrizaciones poseen interpretaciones relacionadas al crecimiento. Se han descrito para las diferentes especies frutales una ecuación o familia de ecuaciones que caracterizan la evolución del fruto a lo largo de ciclo. Bramardi et al. [114], ha encontrado que la curva logística describe el crecimiento en peras William's y Packam's Triumph. El modelo ajustado que mejor describe el patrón de crecimiento fue el modelo logístico en la tercer parametrización según Ratkowski [103] que, además de poseer una adecuada capacidad predictiva, cumple con las mejores propiedades en cuanto a medidas de no linealidad de los parámetros. Este modelo puede representarse como:

$$\mu_Y = \frac{1}{\beta_1 + \beta_2 * \beta_3^{DDF}} \quad (4.14)$$

donde:

- β_1 es el parámetro relacionado inversamente con la asíntota superior (AS).
- β_2 es el parámetro que relaciona la asíntota superior (AS) con la asíntota inferior (AI), mediante la función: e^{AI} / AS .
- β_3 es el parámetro relacionado con la tasa de crecimiento desde los valores 'iniciales' (determinados por la magnitud de β_2 y los valores 'finales' (determinados por la magnitud de β_1).

Dicha ecuación expresa el diámetro ecuatorial en función de Días Posteriores a Plena Floración (DDPF). Estos diámetros se referencian luego a tamaños comerciales de frutos recolectados en la fecha comercial de cosecha. Esto posibilita el seguimiento de los frutos y la identificación de aquellos que no alcanzarán el tamaño deseado, para su eliminación

en un raleo manual.

Los estudios de análisis de la variancia sobre los parámetros del modelo estimados para cada fruto para el caso de William's y Packam's Triumph indicaron que las características particulares de las plantaciones donde se realizaron las mediciones, así como las condiciones climáticas imperantes en los ciclos productivos analizados, influyen sobre el ciclo de crecimiento de los frutos [114].

— 4.3 —

MODELOS DE RESPUESTA MULTIVARIADA

— 4.3.1 —

INTRODUCCIÓN

Existen circunstancias en las que el supuesto de independencia en los valores observados no se cumple, tal es el caso de medidas tomadas de manera repetida sobre un mismo sujeto o situaciones en las que los datos se constituyen en grupos naturales debido a la forma en la que las observaciones surgen. Un caso particular, lo constituyen los datos longitudinales, medidas repetidas realizadas sobre un mismo sujeto a lo largo del tiempo.

Es natural pensar que en estas situaciones las mediciones se encuentran naturalmente correlacionadas dentro del mismo sujeto, debido a la naturaleza jerárquica de los datos, una fragante violación al supuesto (4) del modelo de regresión clásico.

Si denotamos $Y_i = (y_{i1}, \dots, y_{in_i})^T$ al vector que contiene a las medidas tomadas en forma repetida para un mismo sujeto o grupo i , es posible obtener los pares ordenados (Y_i, X_i) , con $i = 1, \dots, m$. Los Y_i dados los X_i pueden razonablemente ser vistos como independientes entre sí, a pesar que las observaciones dentro de Y_i se encuentran correlacionadas.

El tratamiento de las observaciones como independientes cuando en realidad no lo son nos puede acarrear algunos problemas:

- I) Las inferencias se pueden tornar ineficientes.
- II) Los errores estándar deducidos bajo supuestos incorrectos en la variación pueden no representar apropiadamente la verdadera variación en el muestreo.
- III) El modelo de regresión usual puede no ser el marco adecuado en el cual exponer y analizar los objetivos científicos.

Es posible además asociar a los Y_i , información recolectada a través de covariables de dos tipos:

- *Covariables a nivel intra-sujeto*: Describen las condiciones bajo las cuales las Y_{ij} son registradas en el sujeto i y se denotan a través de μ_i . Para el caso de datos de medidas repetidas u otra condición que provoque cambios en la variable de respuesta para i sobre $j=1, \dots, n_i$, las observaciones pueden ser vistas como realizaciones de un proceso estocástico para el sujeto i . A los efectos de contemplar esta situación, escribimos z_{ij} de dimensión $(r \times 1)$ para denotar todas las ‘situaciones’ asociadas a la recolección de Y_{ij} en i .
- *Covariables a nivel inter-sujeto*: Comunmente no cambian su valor sobre j y frecuentemente pueden ser vistas como características de i o como i es tratado. Estas covariables son consideradas solo cuando el interés se centra en la población de la cual las unidades i son obtenidas y son características de la unidad en su totalidad y no de lo que pasa dentro de ella. Estas covariables se denotan como a_i .

Luego los datos disponibles para i consisten en los pares $(Y_{i1}, Z_{i1}), \dots, (Y_{in_i}, Z_{in_i})$ con su covariable asociada a nivel inter-sujeto a_i . Escribiendo $Z_i = (Z_{i1}^T, \dots, Z_{in_i}^T)^T$ para denotar la colección de situaciones intra-sujeto sobre j , es posible pensar en los datos como triplas (Y_i, z_i, a_i) , con $i = 1, \dots, m$ [76].

También es posible definir $X_i = (z_i^T, a_i^T)^T$ para denotar el conjunto total de covariables asociadas con Y_i , permitiendo representar los datos más suscitantemente como (Y_i, X_i) con

$i = 1, \dots, m.$

Es importante, además, poder discernir las fuentes que originan correlaciones entre observaciones dentro del vector (Y_i, X_i) para cada uno de los i . Ellas pueden surgir a dos niveles:

- I) *Fuentes a nivel de población*: Desde la perspectiva de la población de individuos, es natural pensar que mediciones realizadas sobre un mismo individuo son ‘más parecidas’ entre si, que mediciones realizadas entre diferentes individuos.
- II) *Fuentes a nivel de individuo*: El orden natural de las observaciones para el caso de las mediciones repetidas a lo largo del tiempo, nos pone en sobreaviso sobre la posible presencia de correlaciones seriales. Si en cambio las medidas repetidas no consideran datos longitudinales, la magnitud de la correlación persiste entre las mediciones realizadas dentro de un mismo agrupamiento.

4.3.2

MODELO LINEAL GENERAL MULTIVARIADO

Tradicionalmente se ha reconocido el problema de ajustar curvas cuyos valores están correlacionados. El enfoque de Pothoff y Roy [99] consiste en ajustar un *MANOVA* donde las mediciones en los distintos tiempos se definen como p variables y se modela la curva de crecimiento mediante un polinomio de cierto grado. Los estadísticos de prueba multivariada como Lambda de Wilkins, Traza de Pillai, Traza de Hotelling-Lawley y máxima raíz de Roy pueden entonces calcularse a partir de matrices de sumas de cuadrados, productos cruzados de error e hipótesis análogas a las sumas de cuadrados de error y tratamiento.

Si bien este enfoque no tiene restricciones en cuanto a la covarianza de las mediciones longitudinales para un mismo individuo, posee una serie de restricciones que limitan su uso. La más importante de ellas, es requerir el mismo número de variables para cada

uno de los individuos no permitiendo ningún tipo de desbalance. Si un individuo posee alguna observación (variable) faltante como producto de un proceso aleatorio (missing at random, MAR), una de las opciones consiste en eliminar a este en su totalidad a los efectos del análisis, otra alternativa consiste en rellenar los espacios vacíos mediante estimadores razonables de los datos faltantes [18]. Esta última alternativa solo es posible cuando el mecanismo generador de los datos faltantes es completamente al azar (missing completely at random, MCAR) de acuerdo a Little y Rubin [102].

Enfoques basados en ANOVA, más potentes al MANOVA, han sido desarrollados a partir de una estructura de dependencia marginal en las medidas repetidas parecida a las subparcelas en un diseño de parcelas divididas, donde la matriz de varianza-covarianza para las medidas repetidas en el tiempo tiene una estructura de simetría compuesta. Así es como el análisis de medidas repetidas basado en ANOVA es válido si se cumple esta condición, que en términos prácticos implica que las varianzas son iguales a lo largo del tiempo y las covarianzas entre dos observaciones son siempre las mismas, esta situación es muy poco probable de encontrar en la realidad. El desarrollo de este enfoque ha sido propuesto para análisis de datos provenientes de experimentos diseñados, con medidas repetidas obtenidas en conjuntos de situaciones comunes a todos los individuos, con covariables presentes como factores discretos y datos completos. Esto limita su uso en medidas longitudinales que se encuentran irregularmente espaciadas o incompletas o presentan el requerimiento de contemplar variables cuantitativas en el análisis [18].

Los enfoques actuales al problema simplifican la estructura de correlación y usan modelos no necesariamente lineales. La simplificación de la estructura de correlación se logra a través de la introducción de efectos aleatorios o mediante la introducción de una estructura explícita de covarianza como veremos a continuación.

MODELOS SUJETO-ESPECÍFICOS

Los modelos sujeto-específicos (SS) son utilizados cuando el objetivo del estudio es comprender el comportamiento de la población de sujetos a través de la variabilidad entre el perfil “inherente” de cada individuo y el perfil “típico” o “promedio” de todos los individuos. Para ello se requiere tratar las variaciones originadas a nivel inter-sujeto e intra-sujeto a través de un marco teórico adecuado [74].

VARIACIÓN INTER-INDIVIDUO

El modelo propuesto para describir el perfil “inherente” de cada individuo surge a partir de consideraciones teóricas acerca del proceso que tiene lugar dentro de cada uno de estos. Por consiguiente, tiene sentido sólo como un modelo para el comportamiento individual.

Cuando ciertos ajustes individuales se superponen a los datos es posible apreciar que, aunque la forma general del modelo parece describir bien el patrón para todos los sujetos, la forma no es la misma para todos ellos. Estas diferentes manifestaciones de la pauta común se puede pensar que es debido a diferencias en los valores de los parámetros de la función propuesta, reguladas por procesos subyacentes individuales.

Por lo tanto, los valores medios de la población de cada uno de los parámetros caracterizan el comportamiento "típico", y el patrón de variación (y co-variación) de los valores de los parámetros caracteriza la variación entre sujetos en su perfil “inherente”. Un modelo estadístico adecuado por lo tanto debe:

- Preservar la noción de lo que ocurre a nivel de individuo, es decir, el modelo propuesto debe describir los comportamientos dentro de los individuos.
- Permitir que este comportamiento varíe a lo largo de los individuos, con la posibi-

lidad de variar los parámetros para todos ellos.

- Considerar la correlación originada en los niveles de individuo y población.

Un modelo que haga esto posee β_i donde i denota que los parámetros son específicos para el sujeto i tal lo propuesto por Laird et al. [86], por lo que:

$$E(Y_{ij} | z_{ij}, \beta_i) = f(z_{ij}, \beta_i) \quad (4.15)$$

La media está condicionada en β_i , ya que el modelo describe el comportamiento a nivel individual, donde β_i puede considerarse como un parámetro fijo desde la perspectiva del individuo i . Como cada sujeto tiene sus propios parámetros, una forma natural para formalizar esto es pensar en β_i como vectores aleatorios que surgen de alguna distribución multivariada. Por supuesto, los β_i no son observables, por lo que no podemos usar lo que vemos para modelarlos, pero se pueden hacer algunas suposiciones posibles. Por ejemplo, podríamos asumir que:

$$\beta_i \sim N(\beta, D) \quad (4.16)$$

Un modelo que caracteriza la población de sujetos en términos de la distribución de sus parámetros inherentes, donde β representa los valores típicos en la población, mientras que la matriz de covarianza D describe la variación (o la covariación) de los parámetros en la población. El supuesto de normalidad se hace comúnmente, pero no es necesario. Es posible, además discriminar dos componentes dentro de β_i :

$$\beta_i = \beta + b_i \quad (4.17)$$

donde b_i son vectores aleatorios tomados generalmente como independientes de x_i , tal que $E(b_i) = 0, \text{var}(b_i) = D$. Si incorporamos el supuesto de normalidad obtenemos: $b_i \sim N(0, D)$. Los b_i se refieren generalmente como efectos aleatorios; ellos describen

cómo el vector de parámetros para un sujeto i (elegido al azar) se desvía de la población media β . De hecho, esta ecuación tiene el aspecto de un modelo de regresión simple para el vector de parámetros β y se puede extender fácilmente para incorporar covariables a los efectos fijos. Tal necesidad surge generalmente cuando un número de características basales de los individuos pueden influir en la respuesta observada y, por lo tanto, representan algunas de las variaciones presentes a nivel de población. Estas características están en el nivel del individuo, por lo que son componentes de a_i .

Es por ello posible considerar:

$$\beta_i = A_i\beta + B_i b_i \quad (4.18)$$

- A_i es una matriz de diseño que es función de los componentes de a_i .
- B_i es una matriz de diseño con elementos 0 o 1, que sirve para especificar qué elementos de β_i tienen efectos aleatorios asociados.

Si pretendemos flexibilizar el modelo a nivel de variación poblacional, una propuesta más general sería plantear:

$$\beta_i = d(a_i, \beta, b_i) \quad (4.19)$$

donde d es un vector b -dimensional de funciones de a_i, β y b_i y el modelo (4.18) es un caso especial de este donde $d = I$.

VARIACIÓN INTRA-INDIVIDUO

El modelo, sin embargo, no se encuentra completo. Todavía tenemos que considerar la variación a nivel individual sobre $f(z_{ij}, \beta_i)$. De acuerdo a lo ya considerado escribiendo $e_{ij} = Y_{ij} - f(z_{ij}, \beta_i)$, tenemos $E(e_{ij} | z_{ij}, \beta_i) = 0$. Una perspectiva conceptual para la variación intra-individuo es discutida por diversos autores [93] [40] [76] y presentada en la figura 4.1).

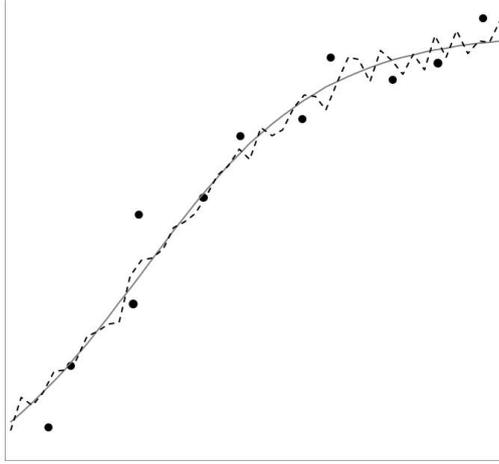


Figura 4.1: Fuentes de Variación Intra-sujeto. La línea sólida es el perfil “inherente”, la punteada es la realización del proceso que actualmente tiene lugar como producto de las llamadas “fluctuaciones” biológicas, y los puntos son medidas en un particular momento en el tiempo de las realizaciones sujetas a error.

De acuerdo a esta perspectiva, podemos pensar que e_{ij} se compone de:

$$e_{ij} = e_{p_{ij}} + e_{M_{ij}} \quad (4.20)$$

por lo que podemos escribir:

$$Y_i = f_i(z_i, \beta_i) + e_i = f_i(z_i, \beta_i) + e_{p_i} + e_{M_i} \quad (4.21)$$

donde $e_{p_{ij}}$ representa la parte de la desviación asociada con las “fluctuaciones” biológicas de las realizaciones del proceso y $e_{M_{ij}}$ se asocia al error de medición.

En la aplicación particular de las curvas de crecimiento de frutos, los $e_{M_{ij}}$ se suponen razonablemente independientes a través de j , debido a que cada valor surge de una medición por separado. También es razonable suponer que $e_{p_{ij}}$ y $e_{M_{ij}}$ son independientes, por lo que:

$$\text{var}(Y_{ij} | z_{ij}, \beta_i) = \text{var}(e_{p_{ij}} | z_{ij}, \beta_i) + \text{var}(e_{M_{ij}} | z_{ij}, \beta_i) \quad (4.22)$$

Esto demuestra que, incluso para el modelado a nivel individual, la varianza asumida por el modelo es un intento de caracterizar la suma de las variaciones debidas a las dos fuentes de variación. Por lo tanto, el modelo más general tendría dos componentes

susceptibles de ser modelados que, dependiendo del aporte relativo de cada una de las fuentes de variación, tendrá un enfoque particular.

Por ejemplo, en el contexto de farmacocinética, a menudo se piensa que el error de medición debido al análisis es la principal fuente de variación alrededor de $f(z_i, \beta_i)$, mientras que las fluctuaciones “biológicas” sobre la trayectoria $f(z_i, \beta_i)$ son, en comparación muy pequeñas [75]. Esta situación no es la concebida en el contexto del crecimiento de frutos donde el error de medición permanece constante, mientras que las fluctuaciones del proceso de crecimiento se incrementan a medida que el desarrollo del fruto avanza.

La variación presente en e_{p_i} es susceptible de ser caracterizada a través de funciones de autocorrelación y varianza que describan el patrón de correlación y variación de las realizaciones. La correlación entre $e_{p_{ij}}$ y $e_{p_{ij'}}$ con $j \neq j'$, puede ser no despreciable si las mediciones están cerca en el tiempo, si ello ocurre, es posible considerar un modelo para describir su patrón con:

$$\Gamma_i(\rho, z_i) \tag{4.23}$$

que expresa la función de correlación de e_{p_i} , donde ρ es un parámetro que caracteriza el modelo y generalmente es el mismo a lo largo de los i , además se habilita al modelo a depender de z_i por lo que algunos modelos de correlación pueden depender del tiempo actual de la observación.

Es natural pensar, además, que las desviaciones asociadas a las realizaciones del proceso biológico no permanezcan constantes por lo que este comportamiento es susceptible de ser modelado a través de una función de varianza (T_i), definida para ser una matriz diagonal dependiendo de un parámetro δ con elementos obtenidos a partir de:

$$T_i(z_{ij}, \beta_i, \delta) \tag{4.24}$$

Por lo que la matriz de covarianza para e_{pi} puede obtenerse como:

$$\text{VAR}(e_{pi} | z_{ij}, \beta_i) = T_i^{1/2}(z_{ij}, \beta_i, \delta) \Gamma_i(\rho, z_i) T_i^{1/2}(z_{ij}, \beta_i, \delta) \quad (4.25)$$

En general la correlación entre los errores de medición e_{Mij} y $e_{Mij'}$ con $j \neq j'$ no existe por lo que es razonable suponer independencia entre las realizaciones en el proceso de medición. La matriz de covarianza de e_{Mi} puede ser definida como una matriz diagonal con elementos obtenidos dependiendo de un parámetro θ a partir de:

$$\text{VAR}(e_{Mi} | z_{ij}, \beta_i) = \Lambda_i(z_{ij}, \beta_i, \theta) \quad (4.26)$$

Si combinamos las consideraciones precedentes 4.25 y 4.26, adoptando 4.22 es posible obtener una representación general de la variación intra-individuo que expresa la función de varianza propuesta en 4.10 como:

$$\text{VAR}(Y_i | z_{ij}, \beta_i) = T_i^{1/2}(z_{ij}, \beta_i, \delta) \Gamma_i(\rho, z_i) T_i^{1/2}(z_{ij}, \beta_i, \delta) + \Lambda_i(z_{ij}, \beta_i, \theta) = R_i(z_i, \beta_i, \xi) \quad (4.27)$$

donde $\xi = (\delta^T, \rho^T, \theta^T)$.

Es una práctica común, adoptar modelos que son simplificaciones de 4.27, cuyas especificaciones se encuentran determinadas por consideraciones prácticas y conocimientos de la materia en estudio.

Otra perspectiva, menos precisa considera que las especificaciones de varianza adoptadas a nivel individual no se encuentran en la forma de una suma de dos expresiones para cada fuente. Este enfoque no trata de distinguir entre las dos fuentes, sino más bien intenta sólo aproximar el patrón global de la variación entre individuos, algo que es sin duda más complicado y genera menos información.

MODELO GENERAL SUJETO-ESPECÍFICO

A partir de las consideraciones anteriores, es posible plantear un modelo completo general a partir de un modelo en dos etapas:

- Etapa I: Modelo Individual.

$$E(Y_i | z_i, a_i, b_i) = E(Y_i | X_i, b_i) = f_i(z_i, \beta_i) = f_i(z_i, a_i, \beta, b_i) = f_i(X_i, \beta, b_i) \quad (4.28)$$

$$\text{var}(Y_i | z_i, a_i, b_i) = R_i(\beta_i, \xi, z_i) = R_i(\beta, \xi, X_i, b_i) \quad (4.29)$$

- Etapa II: Modelo poblacional.

$$\beta_i = d(a_i, \beta, b_i) \quad (4.30)$$

Donde denotamos como a_i a las covariables para el sujeto i que están involucradas en la construcción de la matriz A_i , b_i generalmente se asume i.i.d. con $E(b_i) = 0$, $\text{var}(b_i) = D$. Además, $R_i(\beta, \xi, X_i, b_i)$ es la matriz de covarianza que tiene, implícitas, las hipótesis relacionadas a las fuentes de variación y correlación a nivel intra-individuo y donde ξ representa los parámetros que caracterizan a estas fuentes.

Independientemente de lo que se cree acerca de la variación dentro de los individuos, es necesario tener en cuenta que las observaciones del mismo vector de datos Y_i comparten el mismo efecto aleatorio b_i por lo que el modelo naturalmente tiene en cuenta el hecho que a nivel de población, las observaciones realizadas sobre el mismo individuo tienden a ser “más parecidas” entre sí en relación a las observaciones realizadas entre distintos individuos.

El modelo en (4.28) y (4.30) es un modelo sujeto-específico, donde se modela en dos etapas de manera que:

- Es adecuado para los objetivos científicos planteados para el caso de la curvas de crecimiento: β y D caracterizan la media y la variación en la población de parámetros, los cuales son de interés científico directo.
- Reconoce que el comportamiento a nivel individual es bien entendido en el sentido que un modelo teórico puede resultar incorporado directamente.

Modelos de la forma dada en (4.28) y (4.30) se conocen como modelos no lineales de efectos mixtos (NLME). Modelos-no lineales debido a que algunos de los parámetros ingresan de forma no lineal en la función del modelo y de efectos mixtos por reconocer la presencia tanto de parámetros fijos (β, D, ξ) como efectos aleatorios (b_i) en el modelo.

MODELO MARGINAL INDUCIDO

En los modelos univariados de regresión, “se modela lo que se ve” en el sentido que el modelo de media-varianza es un modelo para los momentos marginales de la respuesta, condicionada en las covariables donde nosotros observamos (Y_i, X_i) , $i = 1, \dots, m$.

El modelo en (4.28) y (4.30) no proporciona directamente un modelo análogo para el caso multivariado, esto es un modelo para $E(Y | X_i)$ y $VAR(Y | X_i)$.

Consideremos ahora lo que el modelo implica sobre estos momentos marginales de Y_i , condicionado sólo por las covariables.

En el caso de la esperanza marginal tenemos que, a partir de (4.28), es posible plantear:

$$E(Y_i | X_i) = E(E(Y_i | X_i, b_i) | X_i) = E(f(X_i, \beta, b_i) | X_i) \quad (4.31)$$

donde la esperanza es sobre la distribución de los b_i dado X_i . Es decir, para obtener el promedio de todos los vectores respuesta con covariables $X_i = (z_i^T, a_i^T)^T$, debemos promediar sobre la población (la distribución de los b_i). Escribiéndolo de otra forma

y denotando como F_b a la función de distribución acumulada de b_i y recurriendo a la independencia asumida entre b_i y X_i podemos expresar esta esperanza como:

$$E(Y_i | X_i) = \int f_i(X_i, \beta, b_i) dF_b(b_i) \quad (4.32)$$

Cuando f es una función no lineal en b_i , es probable que esta integral no se puede evaluar analíticamente debido a la imposibilidad de obtener una forma cerrada de esta. Por ello, esta estrategia de modelado no tiene una analogía directa con el modelo de medias del modelo de regresión univariado en la respuesta.

Para el caso de la varianza marginal, el análisis se puede realizar a partir de la relación que dos vectores aleatorios cualesquiera Z y V poseen:

$$\text{VAR}(Z) = E(\text{VAR}(Z | V)) + \text{VAR}(E(Z | V)) \quad (4.33)$$

por lo que a partir de 4.28 es posible obtener:

$$\text{VAR}(Y_i | X_i) = E(R_i(\beta, \xi, X_i, b_i) | X_i) + \text{VAR}(f_i(X_i, \beta, b_i) | X_i) \quad (4.34)$$

donde la esperanza y la varianza nuevamente aquí lo son con respecto a la distribución de los b_i dado X_i .

De lo expuesto al momento es posible concluir que:

- Al igual que la media, en la varianza marginal los componentes de R_i y f son no lineales en b_i , por lo que no necesariamente existen formas cerradas de expresar $\text{VAR}(Y_i | x_i)$, ya que ambos términos implican integrales sobre la distribución de los b_i .
- Debido a que b_i aparece en cada elemento de $f_i(X_i, \beta, b_i)$, el segundo término en

(4.34) muy posiblemente no es una matriz diagonal.

- Incluso si la covarianza intra-individuo R_i es diagonal, basándonos en la creencia que la correlación entre mediciones de un mismo individuo es despreciable, el segundo término no necesariamente debe ser diagonal. Por ello, es también poco probable que $VAR(Y_i | X_i)$ sea una matriz diagonal en general, lo que demuestra como el modelo habilita la posibilidad de una mayor semejanza entre las respuestas de un mismo individuo.

Es frecuente encontrar estudios que asumen a R_i como una matriz diagonal, a partir de la idea que durante el proceso de estimación del segundo término se captura cualquier correlación entre los elementos de Y_i . Una aproximación motivada en limitaciones en la cantidad de datos sobre cada individuo.

4.3.4

**MODELOS SUJETOS-ESPECIFICOS VERSUS MODELOS PROMEDIO
POBLACIONALES**

Un enfoque alternativo a modelar los pares (Y_i, X_i) , $i = 1, \dots, m$ como lo planteado por los modelos (SS), es de hecho construir directamente un modelo para la media y la matriz de covarianza marginal condicional $E(Y_i | X_i)$ y $VAR(Y_i | X_i)$ [117].

Este último modelo se denomina Promedio Poblacional (PP), se obtiene promediado directamente a través de la población, no reconociendo explícitamente comportamientos individuales a través de parámetros específicos de individuo tales como β_i . Aquí, el interés no se centra directamente en los momentos de los vectores respuesta condicionales a las covariables per se, sino que estos vectores son de interés sólo como fuente disponible de información sobre los β_i , que no son observables directamente.

La forma general del modelo postulada tiene una analogía directa con el caso univariado:

$$E(Y_i | X_i) = f_i(X_i, \beta) \tag{4.35}$$

$$\text{VAR}(Y_i | X_i) = V_i(\beta, \varepsilon, x_i) \quad (4.36)$$

Donde β contiene a los parámetros de la regresión, y ε es un vector de parámetros de varianza y correlación. Una comparación entre ambas estrategias de modelación nos permite apreciar que:

■ Los modelos SS:

- Son preferidos en situaciones donde el interés se focaliza en la distribución de aspectos particulares del comportamiento individual, basado quizás en consideraciones teóricas, si estas se encuentran disponibles.
- En caso que el interés está centrado en la población de respuestas y la aproximación SS sólo se utiliza como un mecanismo para modelar la correlación, los efectos aleatorios inducen una estructura de correlación.
- Si f es un modelo postulado en la etapa I para el comportamiento individual, entonces :

$$E(Y_i | X_i) = E(f_i(X_i, \beta, b_i) | X_i) \quad (4.37)$$

■ Los modelos PP:

- Esta aproximación es utilizada cuando el interés se centra en la población de respuestas. El promedio poblacional de las respuestas es modelado directamente, a través de postular el modelo de la forma:

$$E(Y_i | X_i) = f_i(X_i, \beta) \quad (4.38)$$

- El patrón total de varianza y correlación tanto a nivel individual como poblacional es modelado explícitamente.

Si f es una función lineal, las dos estrategias llevan al mismo modelo para la media mar-

ginal $E(Y_i | X_i)$, no así para $VAR(Y_i | X_i)$. Sin embargo la interpretación de los parámetros fijos β es dual:

- Desde la perspectiva SS, β se interpreta como la media de la población de parámetros de regresión de los individuos β_i , es decir β representa el ‘valor del parámetro típico’.
- Desde la perspectiva PP, β puede ser interpretado como el parámetro que produce el ‘vector de respuesta típico’.

En los modelos lineales, la interpretación de la media marginal, obtenida de cualquiera de las perspectivas es la misma por lo que allí no necesario realizar las disquisiciones realizadas precedentemente: la interpretación del analista de β es correcta de cualquier manera. Está claro que esta propiedad no es propia de funciones no lineales f .

4.4

MODELOS NO LINEALES DE EFECTOS MIXTOS

4.4.1

INTRODUCCIÓN

Los NLME constituyen un enfoque ampliamente utilizado en la actualidad para el análisis de medidas repetidas cuando el interés se centra en características específicas de los individuos. Históricamente estos modelos han recibido una gran atención en la literatura estadística desde finales de 1980, desarrollándose una serie de nuevos métodos computacionales para su tratamiento en la década de 1990. En la actualidad, los NLME son ampliamente utilizados en numerosos campos, como biología, agricultura, medio ambiente, medicina y economía, y están disponibles a través de varios paquetes de software, pudiendo ser pensados como una extensión de:

- Modelos de regresión no lineales para ajustar datos de medidas repetidas.
- Modelos mixtos lineales para el caso de funciones no lineales en la respuesta.

Los modelos NLME, tal como se ha presentado en la sección 4.3.3, consisten en modelos jerárquicos en dos etapas, los cuales, en su forma más general y de acuerdo a 4.28 y 4.30 se plantean como:

$$Y_{ij} = f(z_{ij}, \beta_i) + e_{ij}, \quad j = 1, \dots, n_i \quad (4.39)$$

$$\beta_i = d(a_i, \beta, b_i) \quad (4.40)$$

Con e_{ij} y b_i variables aleatorias que poseen una densidad con $e_{ij} \sim (0, R_i(\beta_i, \xi, z_i))$ y $b_i \sim (0, D)$.

En alguna literatura de modelación multinivel, especialmente en Goldstein [44], el modelo constituido por 4.40 es denominado a dos niveles debido a la presencia de dos niveles de variación aleatoria: b_i y e_{ij} . En otras referencias tal como Pinheiro et al. [58], este modelo es descrito como teniendo un nivel de efectos aleatorios en el modelo. En la presente tesis, esta última convención es la que se adopta y se contabilizan los niveles de un modelo multinivel como el número de niveles de efectos aleatorios anidados o cruzados. Es así como un modelo multinivel con dos niveles de efectos aleatorios con respuestas para las unidades del j -ésimo segundo nivel anidadas dentro de las unidades del i -ésimo primer nivel puede expresarse como:

$$Y_{ijk} = f(z_{ijk}, \beta_{ij}) + e_{ijk}, \quad k = 1, \dots, n_i$$

$$\beta_{ij} = d(a_i, a_{ij}, \beta, b_i, b_{ij})$$

Con e_{ijk} , b_i y b_{ij} variables aleatorias que poseen una densidad con $e_{ijk} \sim (0, R_i(\beta_i, \xi, z_i))$, $b_i \sim (0, D_1)$ y $b_{ij} \sim (0, D_2)$.

Los primeros desarrollos de NLME se concretan a partir de los trabajos de Sheiner and Beal [71]. Luego de ellos, una serie de enfoques estadísticos se ha discutido en la literatura y ha sido aplicado en el análisis de datos. Básicamente, los NLME pueden ser clasificados en cuatro categorías: modelado paramétrico, modelado no paramétrico, modelado semi-paramétrico y modelado bayesiano.

4.4.2

NLME PARAMÉTRICOS

Para un NLME paramétrico se tiene un modelo similar a 4.28 y 4.30 con el supuesto adicional que los efectos aleatorios b_i tienen una distribución normal. La mayor complicación de la estimación de parámetros en los NLME paramétricos es la función de verosimilitud basada en la distribución marginal de la variable de respuesta 4.31, la cual generalmente no tiene una expresión cerrada. Se han presentado diferentes métodos para estimar la función de verosimilitud, estos puede clasificarse en cuatro grandes categorías: métodos basados en estimaciones individuales [76], métodos de linealización ([71] [80], [104]), métodos de aproximación de la integral ([100],[57],[27]), y algoritmos EM ([113],[24],[69]).

Los métodos de linealización aproximan los NLME a través de una expansión de series de Taylor de primer orden para arribar a un pseudo-modelo que es de la forma de un modelo lineal mixto. Los métodos de aproximación de la integral utilizan la aproximación de Laplace, la cuadratura gaussiana o el muestreo de importancia para aproximar la distribución marginal de la variable de respuesta Y para luego maximizar la verosimilitud directamente. Como una alternativa a la aproximación directa de la verosimilitud marginal, los algoritmos EM aproximan la esperanza condicional del logaritmo de la verosimilitud en el paso E y luego maximizan esta esperanza en el paso M, para obtener estimadores en el siguiente paso E.

De todos los métodos mencionados, los métodos de linealización son, por lejos, los enfoques más populares debido a su simplicidad numérica.

4.4.3

NLME NO PARAMÉTRICOS

Para los NLME no paramétricos, un método de máxima verosimilitud ha sido propuesto por Mallet et al.[5]. La diferencia entre este modelo y uno paramétrico es que aquí no se realizan presunciones sobre la distribución de los efectos aleatorios, excepto que estos son una medida de probabilidad. La distribución condicional de la respuesta Y dados los efectos aleatorios se asume conocida y el objetivo del procedimiento de estimación es obtener la distribución de probabilidad de los efectos sujeto-específicos b_i que maximicen la verosimilitud. Mallet [5] probó que la solución máximo verosímil es una distribución discreta con un número de puntos de discontinuidad menor o igual al número de agrupamientos en la muestra.

4.4.4

NLME SEMI-PARAMÉTRICOS

Para un NLME semi-paramétrico, un método de estimación máximo verosímil ha sido propuesto por Davidian et al. [76]. La diferencia entre este modelo y el paramétrico es que aquí se permite que la distribución de los efectos aleatorios sea libre dentro de una clase de densidades H , definida en Gallant et al. [10]. Una densidad de H puede ser expresada como una combinación lineal infinita de densidades normales, en los cálculos de la verosimilitud esta sumatoria es truncada a un número finito de términos y la integración numérica es llevada a cabo utilizando la cuadratura de Gauss-Hermite.

ENFOQUE BAYESIANO A LOS NLME

Un enfoque bayesiano a NLME es propuesto por diversos investigadores utilizando un modelo en tres etapas y MCMC [13], [59], [60]. En la primer etapa se especifican la densidad condicional de la observación Y dados los efectos aleatorios b_i y los efectos fijos β . En la segunda etapa, se especifica la densidad de los efectos aleatorios b_i dados los β y en la tercer etapa una densidad previa para β . La densidad posterior de los efectos aleatorios puede ser obtenida utilizando métodos MCMC y luego cualquier característica deseada de la densidad posterior, como la moda, los momentos, probabilidades e intervalos de confianza pueden ser aproximados mediante técnicas MCMC [90]. Las técnicas MCMC tienen gran potencia por su capacidad de manejar modelos de efectos mixtos con efectos aleatorios con alta dimensionalidad y se fundamenta en la generación de muestras aleatorias para obtener la distribución de interés. En otros enfoques tales como INLA (Rue et al. [46]) o métodos basados en variaciones (optimización sobre el espacio de funciones) (Lee et al. [127]) se utilizan métodos para aproximar dicha distribución sin necesidad de generar las muestras iterativamente.

IMPLEMENTACIÓN DE LOS NLME

INTRODUCCIÓN

A partir de las diferentes formulaciones de NLME disponibles en la literatura es posible implementar un sinnúmero de métodos inferenciales. En la presente tesis nos enfocaremos en los NLME paramétricos con efectos aleatorios y términos de error distribuidos normalmente, por lo que el modelo final propuesto está dado por:

- Nivel I: Modelo a nivel de Individuo

$$Y_{ij} = f(z_{ij}, \beta_i) + e_{ij}, \quad i = 1 \dots m \quad j = 1, \dots, n_i \quad (4.41)$$

- Nivel II: Modelo a nivel de Población

$$\beta_i = d(a_i, \beta, b_i) \quad (4.42)$$

Con e_{ij} y b_i variables aleatorias con;

$$E(e_{ij}) = E(b_i) = 0 \quad (4.43)$$

$$\text{VAR}(e_{ij}) \sim N(0, R_i(\beta_i, \xi, z_i)) \quad (4.44)$$

$$\text{VAR}(b_i) \sim N(0, D) \quad (4.45)$$

INFERENCIA BASADA EN VEROSIMILITUD

El modelo a nivel de individuo 4.41 junto con 4.44 produce una densidad normal condicional $p(z_i, b_i, \beta, \xi)$. Bajo el supuesto de independencia de los m individuos es posible obtener la densidad conjunta como producto de las m contribuciones $p(z_i, b_i, \beta, \xi)$. El modelo a nivel de población 4.42 junto a 4.45 hacen posible asumir una densidad k -variada $p(b_i | a_i; D)$ para b_i .

A partir de las especificaciones precedentes es posible postular la densidad marginal de Y_i dado z_i como:

$$p(Y_i | z_i; \beta, \xi, D) = \int p(Y_i | z_i, b_i; \beta, \xi) p(b_i | a_i; D) db_i \quad (4.46)$$

Por lo que la verosimilitud para β, ξ, D puede basarse en la densidad conjunta de los

valores observados Y_i dados z_i :

$$\prod_{i=1}^m p(Y_i | z_i; \beta, \xi, D) = \prod_{i=1}^m \int p(Y_i | z_i, b_i; \beta, \xi) p(b_i | a_i; D) db_i \quad (4.47)$$

Debido a la no linealidad existente en las m integraciones k -dimensionales en 4.47, estas no pueden realizarse de una forma cerrada, por lo que los algoritmos iterativos que se implementen para maximizar β, ξ, D requieren una manera de obtener estas integrales.

4.5.2

INFERENCIA A NIVEL DE SUJETO

En un modelo NLME general, b_i representa el efecto que corresponde a individuos particulares, y el modelo a nivel de individuo $f_i(z_i, \beta, b_i)$ describe la respuesta para individuos particulares. Por lo que, desde un punto de vista sujeto-específico, para caracterizar el comportamiento de los individuos el interés se centra tanto en β como en b_i . La estimación de parámetros fijos como β puede llevarse a cabo en principio a través de máxima verosimilitud.

Debido a que los b_i son vectores aleatorios, la inferencia de ellos es un tanto más complicada. Desde el punto de vista de la población de individuos, los b_i representan muestras aleatorias de la población, por lo que caracterizar b_i es equivalente a predecir el valor que toma una variable aleatoria correspondiente a un individuo seleccionado aleatoriamente de la población. La inferencia en b_i es, por ello, un problema de predicción.

Si se desea predecir el valor del vector aleatorio b_i , es posible obtener información sobre este en la variable de respuesta Y_i . Por lo que es natural caracterizar este problema de predicción como una caracterización de b_i dado Y_i .

El enfoque usual a este problema es utilizar como predictor de b_i , el valor más probable de este dada la respuesta que se ha observado, es decir el valor de b_i que maximiza la

densidad posterior, condicionada a las covariables:

$$p(b_i | Y_i, z_i; \beta, \xi, D) = \frac{p(Y_i | b_i, z_i; \beta, \xi)p(b_i | z_i; D)}{p(Y_i | z_i; \beta, \xi, D)} \quad (4.48)$$

Si β, ξ y D son conocidas, 4.48 puede ser maximizado como una función de b_i , a través de búsquedas numéricas, esto último debido a que no existe una forma cerrada para la expresión.

Como los parámetros fijos no son conocidos pero se encuentran disponibles estimadores de ellos, una sugerencia natural es sustituir los valores estimados en 4.48 y luego maximizar en b_i . Esto es conocido como inferencia Empírica de Bayes [11]. El máximo o la moda posterior b_i así obtenida es llamada Estimador Empírico de Bayes (EBEs).

Desde un punto de vista bayesiano estándar, no existen diferencias en el tratamiento de los parámetros fijos β, ξ y D y los efectos aleatorios b_i . Por ello una aproximación bayesiana estándar para estimar todos estos parámetros podría ser encontrar la moda de la densidad posterior para los parámetros particulares, condicionados sólo en los datos. En particular para los b_i , condicionados en las covariables, la aproximación bayesiana encontraría la moda a partir de:

$$p(b_i | Y_i, z_i) = \int p(b_i | z_i; \beta, \xi, D)p(\beta, \xi, D)d\beta d\xi dD \quad (4.49)$$

Este enfoque requiere la resolución de complicadas integraciones y la especificación de las distribuciones conjuntas previas de (β, ξ, D) . La utilización de implementaciones MCMC hacen posible en este punto la inferencia Bayesiana [74].

MÉTODOS DE ESTIMACIÓN EN NLME

MÉTODOS DE LINEALIZACIÓN

Como se ha mencionada anteriormente, una estrategia atractiva para obtener la verosimilitud 4.47 consiste en aproximarla una manera que sea posible tratar la integración. En particular para cada i , una aproximación de

$$p(Y_i | z_i; \beta, \xi, D) = \int p(Y_i | z_i, b_i; \beta, \xi) p(b_i | a_i; D) db_i \quad (4.50)$$

es propuesta por Beal et al. [115] a través de un método de primer orden, el cual aproxima la verosimilitud tomando una expansión de series de Taylor. Si expresamos $R_i^{1/2}$ a través de la descomposición de Cholesky de R_i y con $\epsilon_i = R_i^{1/2}(z_i, \beta, b_i, \xi)(Y_i - f_i(z_i, \beta, b_i))$ y escribiendo 4.41 y 4.42 como:

$$Y_i = f_i(z_i, \beta, b_i) + R_i^{1/2}(z_i, \beta, b_i, \xi)\epsilon_i, \quad \epsilon_i \sim (0, I_{m_i}), b_i \sim (0, I_{n_i}) \quad (4.51)$$

Como la no-linealidad en b_i ocasiona la dificultad para la integración de 4.47, una propuesta es considerar una aproximación lineal. Una serie de Taylor de 4.51 sobre $b_i = 0$ en los términos lineales, despreciando los términos que involucran a b_i y ϵ_i como pequeños y denotando $Z_i(z_i, \beta, b^*) = \frac{\partial f_i(z_i, \beta, b_i)}{\partial b_i} |_{b_i=b^*}$ nos conduce a:

$$Y_i \approx f_i(z_i, \beta, 0) + Z_i(z_i, \beta, 0)b_i + R_i^{1/2}(z_i, \beta, \xi)\epsilon_i, \quad (4.52)$$

$$E(Y_i | z_i) \approx f_i(z_i, \beta, 0), \quad \text{VAR}(Y_i | z_i) \approx Z_i(z_i, \beta, 0)DZ_i^T(z_i, \beta, 0) + R_i(z_i, \beta, 0, \xi) \quad (4.53)$$

Cuando $p(Y_i | z_i, b_i; \beta, \xi)$ en 4.47 es una densidad normal, 4.52 equivale a otra densidad normal con media lineal en b_i y matriz de varianza independiente de b_i . Si $p(b_i | a_i; D)$

es también normal, la integral es calculable analíticamente de manera análoga a un modelo mixto lineal y produce una densidad normal n_i -variada $p(Y_i | z_i; \beta, \xi, D)$ para cada i con media y matriz de covarianza 4.53. La estimación de β, ξ, D se obtiene a través de la maximización simultánea de: $\prod_{i=1}^m p(Y_i | z_i; \beta, \xi, D)$, lo que es equivalente a maximizar la verosimilitud bajo el supuesto que la distribución marginal $Y_i | z_i$ es normal con momentos dados por 4.53.

Un inconveniente obvio de la presente aproximación es que generalmente resulta pobre, ya que esencialmente reemplaza: $\int f(z_i, \beta, b_i)p(b_i | a_i; D)$ por $f(z_i, \beta, 0)$, por lo que remueve toda consideración de los aspectos individuales al sustituir todos los b_i por 0, para todos los i .

El método de linealización propuesto por Beal et al. [115] es conocido como de primer orden (FO) y es implementado por sus autores en el software NONMEM [116], el cual es muy utilizado en farmacocinética. También se encuentra disponible en SAS proc nlmixed [3] a través del método firo.

Lindstrom and Bates [80] sugieren una mejor aproximación en comparación al método FO. En ella la expansión de Taylor sobre los términos lineales alrededor de b_i en 4.51 se realiza sobre un valor cercano a b_i denominado b^* .

Despreciando los términos que involucran b_i y ϵ_i como pequeños, tratando a b^* como una constante fija y denotando a Z_i de la misma manera que en 4.52, los autores sugieren la siguiente aproximación:

$$Y_i \approx f_i(z_i, \beta, b_i^*) - Z_i(z_i, \beta, b_i^*)b_i^* + Z_i(z_i, \beta, b_i^*)b_i + R^{1/2}(z_i, \beta, \xi)\epsilon_i \quad (4.54)$$

Con los siguientes momentos marginales:

$$E(Y_i | z_i) \approx f_i(z_i, \beta, b_i^*) - Z_i(z_i, \beta, b_i^*)b_i^* \quad (4.55)$$

$$\text{VAR}(Y_i | z_i) \approx Z_i(z_i, \beta, b_i^*) D Z_i^T(z_i, \beta, b_i^*) + R_i(z_i, \beta, b_i^*, \xi) \quad (4.56)$$

Aquí es necesario considerar alternativas válidas para sustituir a b^* . Bajo el supuesto de normalidad de $p(Y | b_i, z_i, \beta, \xi)$ y b_i , Lindstrom and Bates [80] proponen la moda posterior de b_i , obtenida a través de los mejores predictores lineales insesgados (BLUP) actuales de los efectos aleatorios:

$$\hat{b}_i = D Z_i^T(z_i, \beta, b_i) R_i(z_i, \beta, \xi) (Y_i - f_i(z_i, \beta, \hat{b}_i)) \quad (4.57)$$

En general, la densidad posterior $p(b_i | Y_i, z_i; \beta, \xi, D)$ es proporcional a $p(Y_i | b_i, z_i; \beta, \xi) p(b_i | a_i; D)$ de acuerdo a 4.48. Bajo el supuesto de normalidad de los dos términos mencionados, la moda posterior maximiza el logaritmo del producto, el cual, ignorando constantes toma la forma:

$$\begin{aligned} -(1/2) \log |R_i(\beta, \xi, z_i, b_i)| - (1/2) (Y_i - f_i(z_i, \beta, b_i))^T R_i^{-1}(\beta, \xi, z_i, b_i) (Y_i - f_i(z_i, \beta, b_i)) \\ - (1/2) b_i^T D^{-1} b_i \end{aligned} \quad (4.58)$$

Las estimaciones finalmente se obtienen a través de un proceso iterativo que, en su esencia es el siguiente:

- I) Dadas las estimaciones actuales de $\hat{\beta}_i, \hat{\xi}, \hat{D}$ y \hat{b}_i , actualizar el valor de \hat{b}_i en 4.57.
- II) Manteniendo \hat{b}_i fijo, actualizar las estimaciones de β, ξ y D basados en los momentos de 4.55 y 4.56.
- III) Repetir el proceso hasta el cumplimiento de algún criterio de convergencia.

El método de Lindstrom and Bates [80] es referenciado comunmente como el método de linealización EBLUP y es implementado en la función NLME [58], disponible tanto en

S-PLUS [1] como en R [120]. La función NLME solo puede utilizarse para resolver NLME con efectos anidados y una estructura de efectos cruzados muy simple. Wolfinger et al. [104] utilizan un algoritmo diferente para implementar los dos métodos de linealización antes mencionados en la macro de SAS NLINMIX, el cual puede ser utilizada para resolver efectos aleatorios anidados o cruzados. El comportamiento de esta aproximación se encuentra ampliamente documentado y se ha demostrado que llega a valores adecuados en la mayoría de las situaciones, Hartford et al. [4] demuestran que esto es así incluso cuando n_i no es grande o los supuestos de normalidad son violados.

Los métodos de linealización son una extensión natural de los algoritmos de linealización para la regresión no lineal clásica. Las ventajas de estos pueden resumirse en que:

- Son computacionalmente simples.
- Pueden acomodarse múltiples niveles de anidamiento o cruces para los efectos aleatorios.
- Son implementados en varios paquetes como la macro NLINMIX en SAS y NLME en S-PLUS y R.

Sin embargo, a pesar de su popularidad, los inconvenientes de los métodos de linealización son variados:

- Pueden producir un sesgo sustancial en la estimación de parámetros cuando existe un número limitado de observaciones por unidad de agrupamiento y una gran variabilidad en los efectos aleatorios [132].
- La maximización de la verosimilitud se realiza en los llamados pseudo-datos, es decir en la linealización, y no en los datos originales.

Debido a que 4.53, 4.55 y 4.56 definen aproximaciones para las medias y matrices de covarianza marginales para Y_i dados z_i , una alternativa para la estimación de β, ξ, D es

a través de la resolución de ecuaciones de estimación generalizadas (GGE) [93]. Esta aproximación es aplicable a cualquier modelo con momentos marginales. Davidian et al. [75] y Vonesh et al. [30] describen las relaciones existentes entre los GGEs y los métodos aproximados descritos tanto para FO como para EBLUP.

MÉTODOS DE APROXIMACIÓN A LA INTEGRAL
--

La idea básica de la aproximación a la integral es la siguiente:

- Aproximar la verosimilitud marginal de la variable de respuesta utilizando una rutina de integración numérica disponible.
- Maximizar la verosimilitud aproximada numéricamente.

Las aproximaciones a la integral generalmente tienen mayores demandas computacionales que los métodos de linealización. Sin embargo las aproximaciones a la integral maximizan la verosimilitud de los datos originales y pueden generar mejores y más consistentes estimaciones comparados con los métodos de linealización [89]. Sin embargo, es recomendable utilizar los métodos de linealización para proveer valores iniciales adecuados para los métodos de aproximación.

Diferentes métodos basados en la aproximación de la integral han sido propuestos para ajustar los NLME, entre ellos se incluyen la aproximación de Laplace, el Muestreo de Importancia y los Métodos de Cuadratura Gaussiana.

La aproximación de Laplace es un método para aproximar integrales utilizando información local acerca del integrando en su máximo. Por lo tanto es más útil cuando el integrando se encuentra concentrado en las cercanías de su valor máximo. La aproximación de Laplace ha sido ampliamente utilizada en análisis bayesiano para obtener las medias y varianzas posteriores de funciones paramétricas y es también útil para aproxi-

mar la verosimilitud en NLME cuando las integrales en la verosimilitud no tiene solución de una forma cerrada [55],[67],[118],[68]. En ambos casos, la aproximación de Laplace converge al valor correcto del integral a medida que el tamaño de la muestra tiende a infinito. Dada un integral uni-dimensional de una función positiva $p(b)$ y denotando el logaritmo de esta como $l(b)$, la aproximación de Laplace en su forma más simple es:

$$\int p(b)db \approx \sqrt{2\pi}e^{l(\hat{b})}(-l''(\hat{b}))^{-1/2} \quad (4.59)$$

donde \hat{b} maximiza $l(b)$. La aproximación de Laplace para los NLME ha tenido diferentes variaciones. La versión máximo verosímil (ML) propuesta por Pinheiro and Bates [57] trata la verosimilitud marginal de los NLME como una integral con respecto sólo de los efectos aleatorios mientras que la versión máximo verosímil restringida (REML) propuesta por Wolfinger [100] integra sobre los efectos aleatorios y los fijos asumiendo para estos una densidad no informativa previa para la verosimilitud marginal de los NLME. La aproximación de Laplace provee mejores estimaciones que los métodos de linealización y es computacionalmente eficiente. Sin embargo la versión ML puede introducir un sesgo para los modelos de efectos mixtos cuando el número de individuos es pequeño [83],[134],[36]. El desempeño de la versión REML todavía necesita ser investigado.

La versión ML para la aproximación de Laplace es implementada en el procedimiento `nlmixed` de SAS [3] para ajustar NLME para un nivel de efectos aleatorios. Bates et al. [12] extienden la versión ML de la aproximación de Laplace a NLME multinivel con efectos tanto anidados como cruzados y lo implementan en el paquete LMER 4.1-7 del software estadístico R. Sin embargo, esta última no habilita la modelación de efectos fijos para los parámetros del modelo de una forma directa.

Otro método para aproximar una integral numéricamente se denomina Muestreo de Importancia, este toma ventaja del hecho que cualquier integral puede ser pensada como función de esperanzas. Por ello una integral arbitraria múltiple puede ser expresada

como:

$$I = \int f(x)dx = \int \frac{f(x)}{P_x(x)}dx = E\left(\frac{f(x)}{P_x(x)}\right) \quad (4.60)$$

donde x es cualquier variable aleatoria con función de distribución de probabilidad $P_x(x)$ y esperanza $E(\cdot)$. Luego, es posible tomar una muestra de $P_x(x)$ y aproximar la integral a través de la media muestral:

$$\hat{I} = \frac{1}{n} \sum_{i \in n} \frac{f(x_i)}{P_x(x)} \quad (4.61)$$

De esta manera el Muestreo de Importancia provee una manera simple y eficiente de realizar integración Monte Carlo [57].

La Cuadratura Gaussiana es un método que utiliza la regla de cuadratura de Gauss-Hermite para aproximar una integral, esta esencialmente reemplaza la integral por un conjunto fijo de intervalos espaciados con pesos asociados. En el caso univariado esta puede expresarse como:

$$\int h(x)d(x) = \int f(x)w(x)d(x) \approx \sum_{i \in N_{GQ}} w_i f(x_i) \quad (4.62)$$

donde $h(x) = f(x)w(x)$, con $w(x)$ función de ponderación, w_i son ponderaciones, x_i son abscisas y N_{GQ} es el número de puntos de cuadratura. Si los intervalos se seleccionan en las cercanías de la moda condicional de los efectos aleatorios, la aproximación de cuadratura resultante es llamada aproximación de cuadratura gaussiana adaptativa [57].

Los métodos de integración numérica incluyendo la Cuadratura Gaussiana, el Muestreo de Importancia y otros métodos de integración Monte Carlo funcionan extremadamente bien para NLME con uno o dos efectos aleatorios, pero se vuelven computacionalmente intensivos a medida que el número de efectos aleatorios se incrementa [28].

ALGORITMOS EM

El algoritmo EM introducido por Dempster et al. [9] es un método iterativo que alterna entre realizar un paso de esperanza (paso E) y uno de maximización (paso M). Su simplicidad y estabilidad lo han convertido en un enfoque muy utilizado para obtener estimaciones máximo verosímiles de parámetros en modelos estadísticos que contienen datos faltantes o variables no observadas. Si y_o , y_f y θ representan los datos observados, los datos faltantes y el vector de parámetros a ser estimados respectivamente. En la iteración $k+1$, el paso E calcula la esperanza del logaritmo de la verosimilitud de los datos completos $l(y_o, y_f, \theta)$ con respecto a la distribución condicional de los datos faltantes y_f dados los datos observados y_o bajo las estimaciones actuales de los parámetros $\theta^{(k)}$ de la forma:

$$Q(\theta | \theta^{(k)}) = E[l(y_o, y_f; \theta) | y_o; \theta^{(k)}] \quad (4.63)$$

luego el paso M consiste en encontrar $\theta^{(k+1)}$ que maximice $Q(\theta | \theta^{(k)})$ para todo θ en el espacio de parámetros Ω . Estos pasos se repiten hasta el cumplimiento de un criterio de convergencia. Para los NLME, al igual que la evaluación del logaritmo de la función de verosimilitud, la evaluación de la esperanza $Q(\theta | \theta^{(k)})$ involucra la evaluación de integrales múltiples que, en muchos casos, no tienen una expresión con forma cerrada. Varios algoritmos EM simulados han sido presentados para aproximar el paso E. Walker [113] sugiere utilizar integración Monte Carlo para obtener las esperanzas condicionales en el paso E y demuestra que el paso M es tratable analíticamente. Wang [69] presenta un algoritmo Monte Carlo para el procedimiento EM (MCEM) que utiliza muestras obtenidas de un muestreo de importancia. Otra propuesta es la aproximación estocástica del procedimiento EM (SAEM) propuesto por Kuhn et al. [24], el cual descompone el paso E en una etapa de simulación y otra de aproximación estocástica y donde las muestras de los datos faltantes son simuladas bajo la distribución condicional de $p(y_f, y_o; \theta^{(k)})$ u

obtenida por procedimientos MCMC.

4.6

ESTRATEGIAS PARA LA CONSTRUCCIÓN DE MODELOS

4.6.1

INTRODUCCIÓN

La construcción de un modelo mixto involucra cuestiones que no tienen un paralelo en modelos lineales y no lineales [56]. Algunas de estas cuestiones son:

- Determinar cuáles efectos deben tener asociado a un componente aleatorio y cuales deben ser puramente fijos.
- El uso de covariables que expliquen la variabilidad modelada en los efectos aleatorios.
- El uso de matrices de varianza-covarianza estructuradas para reducir el número de parámetros en el modelo.

Por otra parte, para obtener un modelo candidato a partir de los datos es necesario evaluar la precisión de las estimaciones, la significancia de los distintos términos que componen el modelo o comparar que tan bien ajusta un modelo a los datos en relación a otro modelo a partir de pruebas de hipótesis sobre efectos fijos y estructuras de varianza-covarianza en los efectos aleatorios fundamentadas en las propiedades asintóticas de los NLME.

4.6.2

PROPIEDADES ASINTÓTICAS DE LOS NLME

En el modelo lineal univariado, los efectos fijos aparecen en forma lineal y existen pruebas óptimas cuyos estadísticos tienen una distribución nula conocida para cualquier

tamaño muestral. Para el caso multivariado, o para modelos mixtos, no existe una única prueba óptima, como las distribuciones exactas son complejas o intratables, se usan aproximaciones relacionadas con las distribuciones asintóticas.

Para modelos mixtos lineales, cuando el número de individuos m , tiende a infinito y el número de observaciones por individuo n_i , permanece finita, las estimaciones ML Y REML son consistentes, asintóticamente normales y eficientes [56], [119].

Poco se conoce sobre las propiedades de las estimaciones para NLME incluso en grandes muestras [75], [119]. El único camino para realizar afirmaciones generales en la estimación de NLME es considerar propiedades asintóticas de los estimadores y estudiar su comportamiento para tamaños muestrales pequeños para mejorar las estimaciones. Vonesh et al.[29] distingue dos situaciones:

- Cuando el número de individuos tiende a infinito, m tiende a infinito y el número de observaciones por individuo, n_i permanece finito, es decir el caso acotado de manera uniforme.
- El número de individuos m tiende a infinito junto con n_i .

Demidenko [23] ha establecido algunas propiedades asintóticas para cuatro clases de estimadores: El estimador máximo verosímil (MLE), un estimador basado en la aproximación de primer orden de Vonesh [29], y el estimador en dos etapas de Lindstrom and Bates [80].

4.6.3

PRUEBAS DE HIPÓTESIS SOBRE EFECTOS FIJOS

Dos pruebas asintóticas basadas en el principio de verosimilitud son el cociente de verosimilitud (LRT) y la prueba de Wald.

El test de razón de verosimilitud (LRT) [31] es un método general para comparar

modelos anidados obtenidos por máxima verosimilitud. Un modelo estadístico se dice anidado dentro de otro si este representa un caso especial del otro modelo. Si L_2 es la verosimilitud para el modelo más general y L_1 lo es para el modelo con restricciones tendríamos que $L_2 > L_1$ y $\log L_2 > \log L_1$. El test de razón de verosimilitud:

$$2\log(L_2/L_1) = 2[\log(L_2) - \log(L_1)] \quad (4.64)$$

debe ser positivo. Si k_i es el número de parámetros a ser estimados en el modelo i , entonces la distribución asintótica para el estadístico LRT, bajo la hipótesis nula (el modelo restringido es adecuado), es una distribución χ^2 con $k_2 - k_1$ grados de libertad.

La prueba de Wald es la extensión del estadístico t : estimador/error estándar asintótico (*ASE*). El *ASE* tiene que ver con la curvatura de la función de verosimilitud en el máximo: cuanto más cerrada sea la curva, menor va a ser el *ASE*. Para modelos no lineales Wald no es invariante, y por lo tanto depende de la parametrización particular, por lo que se debe considerar la no linealidad en los parámetros del modelo propuesto. Para más de un parámetro, la prueba es asintóticamente chi-cuadrado [53].

Para el test LRT se debe trabajar con ML, debido a que si se estima mediante REML no se considera la contribución de los efectos fijos a la función de verosimilitud.

Para realizar la prueba de Wald es posible utilizar tanto ML como REML. No existe un consenso generalizado sobre cual de ellos es mejor, Kenward et al [78] afirman que el comportamiento de Wald para tamaños de muestra pequeños es mejor con REML .

Los test t basados en el estadístico de Wald evalúan la significancia marginal de cada coeficiente de efectos fijos cuando los otros efectos fijos se encuentran presentes en el modelo, las pruebas F condicionales no sólo son posibles, sino que son preferibles cuando el número de individuos es escaso.

CORRECCIONES A LAS PRUEBAS ASINTÓTICAS

Existen varios métodos para corregir los estadísticos de Wald o LRT con el fin que su distribución se aproxime, en algún sentido, a la distribución teórica de la cual se obtienen los valores críticos.

Un enfoque consiste en corregir los grados de libertad de la distribución teórica de la cual se obtienen los valores críticos, tal es el caso de: los grados de libertad de la varianza residual, el Método de Satterthwaite, el método Containment, el método Between-Within o el método de Kenward-Rogger.

Otros métodos que suelen emplearse para que las pruebas de hipótesis mantengan su nivel de significancia es el uso de otros estimadores de la estructura de covarianza. El método más difundido es el Estimador Sándwich Empírico de la estructura de covarianza, que tiene la ventaja de ser consistente aunque la estructura de covarianza no este correctamente especificada y la desventaja que sus propiedades no son buenas en muestras pequeñas [35].

Por otra parte para el caso de los estadísticos de Wald, Douglas Bates sostiene que no está claro que la distribución de la razón de sumas de cuadrados calculadas en la hipótesis nula sea realmente una distribución teórica, para cualquier elección en los grados de libertad del denominador. Esto motiva la inexistencia de los grados de libertad y de significancias en el paquete LMER 4.1-7 del software estadístico R. Esta falta de implementación surge del convencimiento de las grandes limitaciones del test frente a un estadístico con referencia a una distribución nula aproximada [17], lo que representa un claro cambio de enfoque en comparación al procedimiento NLME del paquete LME, el cual utiliza el método Between-Within.

Una alternativa final a considerar a las planteadas anteriormente consiste en asumir infinitos grados de libertad en el denominador, si el número de grupos en los efectos alea-

torios es lo suficientemente grande (mayor a 50) [8] .

4.6.4

DETERMINACIÓN DE ESTRUCTURAS EN LOS EFECTOS ALEATORIOS

PRUEBA LRT

Cuando se estiman los efectos aleatorios a partir de ML o REML, si la hipótesis nula está anidada dentro de la alternativa, las pruebas LRT tienen bajo la hipótesis nula, una distribución asintótica χ^2 . Los grados de libertad que le corresponden a dicha distribución se obtienen a partir de la diferencia entre la cantidad de parámetros del modelo bajo la hipótesis alternativa (k_2) y la cantidad de parámetros del modelo bajo la hipótesis nula (k_1).

Sin embargo, Stram et al.[19], utilizando los resultados de Self et al.[112], argumentan que esta prueba resulta ser conservadora. Es decir las significancias calculadas a partir de la distribución $\chi^2_{k_2-k_1}$ resultan mayores a las reales, motivado esto en la restricción de nulidad impuesta en los componentes de varianza para el modelo restringido, que la sitúa en el borde del espacio de parámetros. Por ello, Stram [19] sugiere utilizar como distribución de referencia la mezcla de $0,5\chi^2_{k_2-k_1-1} + 0,5\chi^2_{k_2-k_1}$.

Cuando nos encontramos frente a tamaños de muestra pequeños, la aproximación asintótica tampoco es adecuada. Existen algunas correcciones de los grados de libertad para mejorar este problema, como Satterthwaite, Lawley o Bartlett. Estas correcciones tratan, en esencia, de lograr la coincidencia entre el valor esperado del estadístico y los grados de libertad.

CRITERIOS DE INFORMACIÓN

Los Criterios de información (CI) constituyen los también llamados criterios de verosimilitud penalizada, entre los cuales el Criterio de Información de Akaike (AIC) [128] y el Criterio de Información Bayesiano (BIC) [38] presentan un amplio uso en la actualidad. Estos, se obtienen a partir de la formulación marginal de los modelos como:

$$AIC = -2\log L + 2n_{par} \quad (4.65)$$

$$BIC = -2\log L + 2n_{par}\log(N) \quad (4.66)$$

donde $\log L$ denota el log de la verosimilitud, n_{par} el número de parámetros y N el número total de observaciones.

Los CI permiten comparar dos o más modelos, para un mismo conjunto de datos, a partir del criterio “mientras más pequeño mejor”. Si en cambio consideramos las diferencias para un mismo CI de modelos candidatos, es posible afirmar que un valor de dos a cinco unidades implica que aquel modelo con menores valores de CI tiene altas chances de resultar el más adecuado [110].

Mas allá de sus similitudes, cada uno de los CI propuestos cumplen objetivos distintos. AIC constituye una métrica predictiva ya que allí el modelo no pretende inferir con precisión la verdadera distribución, sino predecir datos futuros con la mayor precisión posible [14],[110], así es como AIC favorece al modelo con mayor verosimilitud. BIC, en cambio, se constituye como un criterio más conservador que puede terminar favoreciendo al modelo con menores parámetros.

Algunos autores, entre ellos Sober [26], sugieren que AIC mide la capacidad predictiva mientras que BIC lo hace con la bondad de ajuste, por lo que, al proveer estimaciones de

métricas diferentes no deberían considerarse como criterios en competencia.

En cuanto al método de ajuste más convenientes para trabajar con CI, la bibliografía estadística especializada [37],[98] recomienda el uso de ML, sin embargo hallazgos recientes ponen nuevamente en valor el uso de REML [77],[39].

OTROS ENFOQUES

Otra estrategia propuesta por Pinheiro [56] para la elección de los efectos aleatorios a ser incluidos en el modelo es comenzar la modelación con todos los parámetros como efectos mixtos cuando no existe información previa disponible sobre la estructura de varianza de los efectos aleatorios y la convergencia es posible. Luego examinar los autovalores de las matrices D , chequeando la presencia de magnitudes cercanas al cero. Los autovectores asociados a estos darían una estimación de la combinación lineal de los parámetros que pueden ser considerados como fijos. Si esto sucede, es posible proponer un modelo reducido en el cual la combinación correspondiente de efectos aleatorios es eliminada, ajustándolo y comparándolo con el modelo original por medio de las pruebas LRT, AIC o BIC.

4.6.5

INCORPORACIÓN DE COVARIABLES

Las estimaciones de los b_i representan desviaciones de los β_i , los parámetros individuales de los parámetros fijos. Estas desviaciones provienen de variaciones no explicadas entre los individuos, que en ocasiones pueden contemplarse en forma parcial a través de diferencias en las covariables a_i .

La inclusión de covariables en el modelo para explicar varianzas de los efectos aleatorios generalmente produce una reducción en el número de efectos aleatorios y conduce a una mejor comprensión de los mecanismos que producen la respuesta. Algunas de las

cuestiones que deben ser abordadas en el proceso de modelación de las covariables a_i son:

- Cúales de las a_i candidatas son potencialmente útiles en explicar la variación de b_i ?
- Cúales b_i tienen mejor explicada su variación por las a_i ?
- Como se debe testear las potenciales a_i para su inclusión en el modelo?
- Deben incluirse o eliminarse los b_i del modelo modificado?

Para poder contestar estas preguntas es necesario desarrollar un adecuado método de cribado de a_i candidatas, junto a pruebas que permitan evaluar la significancia de los términos que se incorporan al modelo.

Wu et al.[70] comparan tres métodos de cribado de covariables:

- Mínimos cuadrados no lineales(NL)
- Basados en EBEs
- Cribado directo de las covariables a través de la inclusión de estas y el test LRT

Ellos concluyen que el test LRT es demasiado liberal para datos ralos, los métodos basados en NL son demasiado conservadores, mientras que los métodos basados en EBEs resultan ser los más confiables.

El método basado en EBEs consiste en ajustar un modelo inicial en el cual no se incluyen covariables a_i , obtener los estimadores bayesianos o estimadores bayesianos empíricos (EBEs) y graficar los componentes de \hat{b}_i versus cada elemento de los a_i candidatos[56]. La presencia de patrones sistemáticos aparentes se considera como un indicador que habilita la inclusión de a_i en el modelo a través de una forma funcional sugerida para esa

dependencia. Wakefield [54] demuestra este enfoque en una aplicación particular. Mandema et al. [50] utilizan modelos aditivos generalizados para ayudar a la interpretación. Luego que una a_i ha sido incluida, repetimos el procedimiento con los b_i actualizados y las a_i candidatas remanentes. El número de parámetros adicionales a ser estimados tiende a crecer considerablemente con la inclusión de las a_i y sus b_i asociadas.

Si el número de combinaciones a_i - b_i es grande, se sugiere utilizar una aproximación forward-stepwise en la que los a_i son introducidos uno a la vez y evaluados a partir de criterios gráficos, CI o LRT.

Bonate et al.[94] han demostrado que a medida que la correlación entre los a_i se incrementa, a pesar de que estos no se encuentran en los mismos parámetros, las estimaciones obtenidas resultan menos precisas. Esto puede conllevar la falta de significancia en las pruebas t asociadas en los casos que dicha correlación excede un valor de 0.5.

La inclusión de nuevos efectos aleatorios en el modelo cuando una covariable es incluida es una situación atípica, pero debe ser investigado. El común de las situaciones es que efectos aleatorios pueden ser eliminados a partir de la inclusión de a_i antes inexplicadas. En ambos casos se procede a través de la comparación de modelos por LRT, AIC o BIC. Finalmente, luego que las a_i son incluidas en el modelo, una pregunta natural es qué efectos aleatorios deben permanecer. La razón entre el desvío estándar del efecto aleatorio y el valor absoluto del correspondiente efecto fijo da una idea de la variabilidad relativa para el coeficiente, el cual es una herramienta útil para decidir qué efecto aleatorio debe testearse para ser eliminado.

4.6.6

EXAMINANDO EL MODELO AJUSTADO

Para realizar inferencias acerca del NLME es necesario verificar el cumplimiento de los supuestos para los datos en los cuales se ha obtenido el modelo, estos pueden ser:

- Los errores residuales son independientes, normales e idénticamente distribuidos,

con media cero y varianza σ^2 e independientes, además, de los b_i .

- los b_i se encuentran normalmente distribuidos con media cero, estructura de varianza D e independientes entre los diferentes individuos.

El primer supuesto implica el estudio de las dependencias entre los errores, susceptibles de ser modeladas a través de estructuras de correlación. A efectos de su verificación, es factible evaluar los residuales para cada uno de los individuos, además de los valores ajustados y las a_i de interés. Para ello se confeccionan gráficas de residuales versus la variable independiente y las a_i candidatas, valores observados versus predichos, gráficas de probabilidad normal para los residuales considerados junto a correlogramas y qq plots de residuales para evaluar correlaciones.

El segundo supuesto se evalúa a través de los EBEs a partir de gráficas de probabilidad normal para evaluar normalidad e identificar valores extremos “outliers” y de EBEs versus factores de diseño y variable independiente para el chequeo de supuestos de homogeneidad en la estructura de varianza de los b_i .

4.7

MÉTRICAS DE COMPARACIÓN DE MODELOS

Las distribuciones aproximadas de los estimadores se utilizan para producir pruebas de hipótesis e intervalos de confianza para parámetros y predicciones del modelo. La selección del modelo óptimo entre múltiples modelos candidatos representa un importante desafío, esta suele realizarse a partir de los tradicionales CI y la prueba LRT, los cuales se obtienen a partir de funciones de las verosimilitudes marginales.

Por otra parte, se dispone también de los llamados Criterios Predictivos (CP). Estos hacen uso de valores predichos que permiten construir métricas de capacidad predictiva.

Los CP más utilizados son: Coeficiente R^2 , Coeficiente de Concordancia (CCC), Suma de Cuadrados de Errores de Predicción (PRESS), Raíz del Cuadrado Medio del Error

(RMSE), Sesgo Medio (\bar{e}), Sesgo Medio Relativo ($\bar{e} \%$), Sesgo Medio Absoluto ($|\bar{e}|$) y Sesgo Medio Relativo Absoluto ($|\bar{e} \%$).

CRITERIOS PREDICTIVOS

En el análisis de regresión por mínimos cuadrados clásico, el coeficiente de determinación (R^2 , o su forma ajustada R_{aj}^2) es extensamente utilizado a pesar de sus limitaciones como las descritas por Lvalseth [122]. R^2 , RMSE, y los gráficos de residuales son probablemente las medidas de bondad de ajuste más comúnmente utilizadas en modelación. Aquí, R^2 es muy popular ya que representa la proporción de la variación total en la variable dependiente que es explicada por un modelo ajustado. Es una medida directa y fácilmente comprensible del éxito en la predicción de la variable dependiente por las variables independientes.

Sin embargo, para los modelos mixtos en general, y los NLME en particular, el concepto de variación total puede ser definida de diferentes maneras dependiendo del criterio y el método utilizado en la estimación [85]. Los modelos mixtos pueden tener varios componentes de varianza, además el modelo ajustado puede promediar un único modelo para la población, o un modelo por cada individuo muestreado dentro de la población. No hay una simple definición de R^2 para modelos mixtos, y varios R^2 han sido propuestos [66] [85] [133] [45]. Estos R^2 , sin embargo, pueden indicar cosas muy diferentes, con interpretaciones también diferentes [79] [61]. Mientras que los estadísticos clásicos como LRT, AIC y BIC son utilizados generalmente para comparar NLME alternativos, los valores actuales de estos estadísticos poco dicen acerca de lo bueno o malo que es un modelo. Ellos son imposibles de interpretar sin un marco de referencia [95].

He aquí una reseña de los CP propuestos para su utilización en la presente tesis:

- **COEFICIENTE R^2 CONDICIONAL:** Este coeficiente, también referenciado como efi-

ciencia de modelado o índice de ajuste puede ser obtenido como:

$$R_C^2 = 1 - \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \hat{Y}_{ij})^2}{\sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y})^2} \quad (4.67)$$

donde Y_{ij} y \hat{Y}_{ij} se corresponden al j ésimo valor observado y predicho respectivamente para el sujeto i y \bar{Y} la media general para los valores observados. R_C^2 toma en cuenta de esta manera tanto los efectos fijos como los aleatorios a través de las predicciones \hat{Y}_{ij} .

- CCC CONDICIONAL: El CCC es una medida que combina una medida del grado de sesgo y precisión en las que pares de valores de Y_i y \hat{Y}_i caen en la línea identidad, también conocida como la línea de concordancia, que parte del origen con un ángulo de 45 grados. Vonesh et al. [27] [30] proponen el uso de CCC denotado aquí como CC_1 :

$$CC_1 = 1 - \frac{\sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \hat{Y}_{ij})^2}{\sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \bar{Y})^2 + \sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \bar{\hat{Y}})^2 + N(\bar{Y} - \bar{\hat{Y}})^2} \quad (4.68)$$

Los valores de CC_1 satisfacen: $-1 \leq CC_1 \leq 1$. Un valor de $CC_1 = 1$ o $CC_1 = -1$ corresponde a una perfecta concordancia entre los valores observados y predichos directa e inversa respectivamente [27] [30].

Recientemente, un CCC mejorado es propuesto por Liao [62]:

$$CC_2 = \rho \frac{4S_1S_2 - \rho(S_1^2 + S_2^2)}{(2 - \rho)(S_1^2 + S_2^2) + (\bar{Y} - \bar{\hat{Y}})^2} \quad (4.69)$$

donde $\rho = S_{12}/(S_1S_2)$, con S_1^2, S_2^2 y S_{12}^2 varianzas y covarianzas asociadas con Y_{ij} y \hat{Y}_{ij} .

- RMSE:

$$RSME = \sqrt{\frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \hat{Y}_{ij})^2} \quad (4.70)$$

- Sesgo Medio (\bar{e}):

$$\bar{e} = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} (Y_{ij} - \hat{Y}_{ij}) \quad (4.71)$$

- Precisión de los Errores(SD):

$$SD = \sqrt{\frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} (e_{ij} - \bar{e})^2} \quad (4.72)$$

- Precisión Global de los Errores (δ):

$$\delta = \bar{e}^2 + SD^2 \quad (4.73)$$

- Sesgo Medio Relativo ($\bar{e} \%$) :

$$\bar{e} \% = \frac{\bar{Y} - \hat{\bar{Y}}}{\bar{Y}} 100 \quad (4.74)$$

- Sesgo Medio Absoluto ($|\bar{e}|$):

$$|\bar{e}| = \frac{1}{N} \sum_{i=1}^m \sum_{j=1}^{n_i} |Y_{ij} - \hat{Y}_{ij}| \quad (4.75)$$

- Sesgo Medio Relativo Absoluto ($|\bar{e} \%|$) :

$$|\bar{e} \%| = \frac{|\bar{Y} - \hat{\bar{Y}}|}{\bar{Y}} 100 \quad (4.76)$$

MÉTODOS PARA VALIDAR MODELOS

Para realizar predicciones en un NLME es necesario realizar un procedimiento en dos etapas [82]:

- La primer etapa consiste en calibrar el modelo que se ha ajustado, lo que involucra la predicción de los efectos aleatorios (EBEs) para los sujetos de interés. Estos sujetos preferentemente deben ser nuevos individuos, es decir individuos no utilizados en el procedimiento de ajuste del modelo.

Para ello se requieren algunos datos de estos individuos para poder calibrar el modelo, pero generalmente la cantidad disponible de estos es menor a la presente en el ajuste del modelo [108].

- La segunda etapa consiste en predecir el valor de la variable de respuesta con las estimaciones de los parámetros de los efectos fijos obtenidos del modelo ajustado, las predicciones de los efectos aleatorios resultantes de la calibración del modelo y el modelo en cuestión.

PREDICCIÓN DE NUEVOS EFECTOS ALEATORIOS

Para predecir las respuestas de un nuevo individuo no contemplado en el ajuste del modelo, podemos descomponer el vector Y_i como: $Y_i = \begin{pmatrix} Y_{i,1} \\ Y_{i,2} \end{pmatrix}$ donde $Y_{i,1}$ son las respuestas que han sido observadas y $Y_{i,2}$ son las respuestas a ser predichas. Utilizando β, D y ξ estimadas de la etapa de modelado, $Y_{i,2}$ puede ser predicha utilizando sólo los parámetros fijos estimados ($\hat{\beta}$) comunes a todos los individuos o a través de tanto $\hat{\beta}$ como \hat{b}_i para realizar predicciones sujeto-específicas. La predicción de los EBEs de b_i obtenida a través de los enfoques descritos en la sección 4.5.2 suelen implementarse a través de métodos de linealización o bien a través de MCMC [63]. En el primero de los casos, la linealización

se realiza a través de una expansión de series de Taylor de primer orden, ya sea sobre el valor esperado de b_i (FO) o sobre su valor predicho (BLUP).

Si aplicamos la linealización propuesta en el enfoque FO en la ecuación 4.41 para el conjunto de respuestas observadas obtenemos:

$$Y_{i,1} = f(X_{i,1}, \beta, A_i) + Z_{i,1}b_i + e_{i,1} \quad (4.77)$$

donde $Z_i = \frac{\partial f_i(X_i, A_i \beta + B_i b_i)}{\partial b_i} \Big|_{b_i=0}$, la matriz de derivadas parciales con respecto a b_i evaluadas en $b_i = 0$. Como $e_{i,1}$ se asume independiente a b_i , las variabilidades de $Y_{i,1}$ se expresan como:

$$VAR(Y_{i,1}) = Z_{i,1}DZ_{i,1}^T + R_{i,1} \quad (4.78)$$

$$COV(b_i, Y_i) = DZ_{i,1}^T \quad (4.79)$$

Si asumimos β, D y R_i como conocidos, apilando b_i y $Y_{i,1}$ es posible obtener la distribución normal multivariada conjunta:

$$\begin{pmatrix} b_i \\ Y_{i,1} \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ f(A_i \beta, X_{i,1}) \end{pmatrix}, \begin{pmatrix} D & DZ_{i,1}^T \\ Z_{i,1}D^T & Z_{i,1}DZ_{i,1}^T + R_{i,1} \end{pmatrix} \right) \quad (4.80)$$

Es posible derivar una predicción de b_i a partir de 4.80, si suponemos que z_1 y z_2 son vectores aleatorios normales, y su distribución multivariada conjunta es:

$$\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim N \left(\begin{pmatrix} E(z_1) \\ E(z_2) \end{pmatrix}, \begin{pmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{pmatrix} \right) \quad (4.81)$$

donde $W_{11} = VAR(z_1), W_{22} = VAR(z_2)$ y $W_{12} = COV(z_1, z_2)$. La esperanza de z_1 condicionada en z_2 o BLUP para z_1 está dado por [101]:

$$\hat{z}_1 = E(z_1) + W_{12}W_{22}^{-1}(z_2 - E(z_2)) \quad (4.82)$$

Aplicando 4.82 en 4.80, la esperanza de b_i condicionada en $Y_{i,1}$ puede expresarse como:

$$\hat{b}_i = DZ_{i,1}^T (Z_{i,1} DZ_{i,1}^T + R_{i,1})^{-1} (Y_{i,1} - f(A_i \beta, X_{i,1})) \quad (4.83)$$

Cuando sustituimos todos los parámetros por sus estimaciones, el BLUP para b_i a partir del enfoque FO, también llamado ZEP_b puede expresarse como:

$$\hat{b}_i = \hat{D} \hat{Z}_{i,1}^T (\hat{Z}_{i,1} \hat{D} \hat{Z}_{i,1}^T + \hat{R}_{i,1})^{-1} (Y_{i,1} - f(A_i \hat{\beta}, X_{i,1})) \quad (4.84)$$

Si aplicamos la linealización propuesta en el enfoque EBLUP en la ecuación 4.41 obtenemos:

$$Y_{i,1} = f(X_{i,1}, \beta, A_i) - Z_{i,1} \hat{b}_i + Z_{i,1} b_i + e_{i,1} \quad (4.85)$$

donde $Z_i = \frac{\partial f_i(X_i, A_i \beta + B_i b_i)}{\partial b_i} \Big|_{b_i = \hat{b}_i}$, la matriz de derivadas parciales con respecto a b_i evaluadas en $b_i = \hat{b}_i$. Además, si construimos la distribución normal multivariada de $Y_{i,1}$ y b_i y aplicamos 4.82 a la distribución, obtenemos el BLUP para b_i a partir del enfoque EBLUP denominado $EBLUP_b$ obtenemos:

$$\hat{b}_i = \hat{D} \hat{Z}_{i,1}^T (\hat{Z}_{i,1} \hat{D} \hat{Z}_{i,1}^T + \hat{R}_{i,1})^{-1} (Y_{i,1} - f(A_i \hat{\beta}, z_{i,1}) + Z_{i,1} \hat{b}_i) \quad (4.86)$$

Como b_i se encuentra en ambos lados de la ecuación 4.86 y esta es no lineal en b_i , debe ser resuelta de manera iterativa hasta el cumplimiento de algún criterio de convergencia [108].

— 4.8.2 —

PREDICCIÓN DE FUTURAS OBSERVACIONES

Con b_i predicha, se proponen dos métodos para predecir $Y_{i,2}$, los cuales se derivan de las ecuaciones 4.77 y 4.85 respectivamente. El método de derivación es el mismo que el utilizado para b_i , es decir construir la distribución multivariada de $Y_{i,1}$ y $Y_{i,2}$ para luego

aplicar la ecuación 4.82 a la distribución. Esto da como resultado la esperanza de $Y_{i,2}$ condicionada a $Y_{i,1}$.

Basados en la ecuación 4.77, la distribución conjunta de $Y_{i,1}$ y $Y_{i,2}$ es:

$$\begin{pmatrix} Y_{i,2} \\ Y_{i,1} \end{pmatrix} \sim N \left(\begin{pmatrix} f(A_i \hat{\beta}, X_{i,2}) \\ f(A_i \hat{\beta}, X_{i,1}) \end{pmatrix}, \begin{pmatrix} W_{11} & W_{21} \\ W_{12} & W_{22} \end{pmatrix} \right) \quad (4.87)$$

donde:

- $W_{12} = COV(Y_{i,2}, Y_{i,1}) = Z_{i,2} D Z_{i,1}^T + COV(e_{i,1}^T, e_{i,2})$
- $W_{11} = VAR(Y_{i,1}) = Z_{i,1} D Z_{i,1}^T + R_{i,1}$
- $Z_{i,j} = \frac{\partial \hat{f}_i(X_i, A_i \beta + B_i b_i)}{\partial b_i} \Big|_{b_i=0}$

Luego, aplicando la ecuación 4.82 a la distribución conjunta de $Y_{i,1}$ y $Y_{i,2}$ y substituyendo los parámetros desconocidos por sus estimaciones obtenemos el predictor de $Y_{i,2}$:

$$\hat{Y}_{i,2} = f(A_i \hat{\beta}, X_{i,2}) + \hat{W}_{21} \hat{W}_{11}^{-1} (Y_{i,1} - f(A_i \hat{\beta}, X_{i,1})) \quad (4.88)$$

Esta ecuación tiene en cuenta la correlación existente entre observaciones, lo que permite mejorar la calidad de las predicciones [20]. Cuando $e_{i,1}$ es independiente de $e_{i,2}$, 4.88 es equivalente a:

$$\hat{Y}_{i,2} = f(A_i \hat{\beta}, X_{i,2}) + \hat{Z}_{i,2} \hat{b}_i \quad (4.89)$$

Hall et al. [20] proponen otro método para predecir $Y_{i,2}$ que se deriva del enfoque EBLUP a través de la ecuación 4.85, con b_i tomados como fijos, la distribución conjunta de $Y_{i,1}$ y $Y_{i,2}$ derivadas de la ecuación 4.85 es:

$$\begin{pmatrix} Y_{i,2} \\ Y_{i,1} \end{pmatrix} \sim N \left(\begin{pmatrix} f(A_i \hat{\beta}, X_{i,2}) - Z_{i,2} \hat{b}_i \\ f(A_i \hat{\beta}, X_{i,1}) - Z_{i,1} \hat{b}_i \end{pmatrix}, \begin{pmatrix} W_{11} & W_{21} \\ W_{12} & W_{22} \end{pmatrix} \right) \quad (4.90)$$

Por lo que aplicando la ecuación 4.82 a la distribución previa y sustituyendo los valores desconocidos por sus predichos, tenemos:

$$\hat{Y}_{i,2} = f(A_i \hat{\beta}, X_{i,2}) - Z_{i,2} \hat{b}_i + \hat{W}_{21} \hat{W}_{11}^{-1} (Y_{i,1} - f(A_i \hat{\beta}, X_{i,1}) + Z_{i,1} \hat{b}_i) \quad (4.91)$$

donde W_{21} y W_{11}^{-1} son equivalentes a los definidos anteriormente. Cuando $e_{i,1}$ es independiente de $e_{i,2}$, 4.91 es equivalente a:

$$\hat{Y}_{i,2} = f(A_i \hat{\beta} + B_i \hat{b}_i, X_{i,2}) \quad (4.92)$$

Las ecuaciones 4.89 y 4.92 difieren en los predictores para b_i , y, en como los efectos aleatorios predichos son aplicados. La ecuación 4.92 es no lineal en \hat{b}_i , mientras que la ecuación 4.89 es lineal en \hat{b}_i , por lo que a partir aquí nos referiremos a 4.92 como el predictor no lineal (PNL_y) y a 4.89 como el predictor lineal (PL_y).

Cuando se obtienen modelos con estructuras que dependen de un número reducido de parámetros es factible incorporarlos a efectos de obtener predicciones. Así es como diversos autores [107] [41], proponen ecuaciones de predicción más explícitas. Si la estructura propuesta se corresponde a una autocorrelación continua de primer orden CAR(1), es posible plantear la ecuación 4.88 como:

$$\hat{Y}_{i,2} = f(A_i \hat{\beta}, X_{i,2}) + \hat{\rho}^d (Y_{i,1} - f(A_i \hat{\beta}, X_{i,1})) \quad (4.93)$$

De la misma manera es posible plantear una alternativa para 4.91 para el enfoque EBLUP:

$$\hat{Y}_{i,2} = f(A_i \hat{\beta}, X_{i,2}) - Z_{i,2} \hat{b}_i + \hat{\rho}^d (Y_{i,1} - f(A_i \hat{\beta}, X_{i,1}) + Z_{i,1} \hat{b}_i) \quad (4.94)$$

donde d es el número de días proyectados para la predicción a partir de la fecha correspondiente a la calibración del modelo.

Parte III

EXPERIMENTACIÓN

Capítulo 5

MATERIALES Y MÉTODOS

La recopilación de medidas de diámetro ecuatorial de frutos se corresponde a esquemas de muestreo de tipo LONGITUDINAL y TRANSVERSAL para los diámetros ecuatoriales de frutos.

5.1

OBSERVACIONES LONGITUDINALES

Durante 17 temporadas se recolecta información correspondiente a 1.599 frutos individualizados a partir de los cuales se obtuvieron 17.667 registros de diámetro ecuatorial. Los establecimientos a partir de donde se generan las mediciones se corresponden a 7 parcelas experimentales de establecimientos representativos de los montes comerciales de la región del Alto Valle de Río Negro y Neuquén.

Para cada una de las temporadas y en cada parcela experimental, se seleccionan en forma aleatoria árboles de Peral (*Pirus communis*) correspondientes al cultivar William's sujetos a prácticas habituales en la región en cuanto a poda, tratamientos sanitarios, raleo, fertilización y riego.

Sobre cada árbol del sitio elegido se señalan temprano, en la primavera, frutos elegidos al azar en un número igual para tres estratos de tamaño para ese momento: pequeños,

medianos y grandes. A estos frutos, con intervalos de una semana y hasta, por lo menos, la autorización oficial de la cosecha, se les mide el diámetro ecuatorial como resultante del promedio de dos mediciones perpendiculares.

Ante la caída eventual de alguno de estos frutos se procede a su reemplazo, seleccionando en oportunidad de una nueva medición otro fruto del mismo árbol y de tamaño similar al esperado para el fruto caído. Los momentos en que se efectúan las mediciones en cada temporada se referencian como días transcurridos desde la fecha de plena floración (80% de flores abiertas) en la primavera correspondiente. De este modo la variable predictora para los modelos ajustados son los Días Posteriores a Plena Floración (DDPF), con lo que es posible hacer comparaciones en base a estados fenológicos y no a fechas cronológicas.

Cuando es factible, se realizan aplicaciones de productos para retención de frutos, a fin de controlar su caída natural por abscisión y permitir proseguir con las mediciones hasta unos 30 días después de autorizada la cosecha.

La fecha de autorización de cosecha se determina a través del Programa Regional de Madurez en el que participan múltiples instituciones con la coordinación del Área de Postcosecha del Instituto Nacional de Tecnología Agropecuaria Alto Valle (INTA Alto Valle). Con el propósito de obtener métricas de validación externa, la información histórica se desagrega en una base de modelación, consistente en 1.374 frutos (Cuadro 5.1) y otra base de validación externa consistente en 225 frutos (Cuadro 5.2).

5.2

OBSERVACIONES TRANSVERSALES

A lo largo de 8 sitios, procedentes de 4 temporadas y 2 establecimientos, se identifican plantas de Pera variedad William's y durante el mes de diciembre se registran los diámetros de la totalidad de los frutos de cada una de las ellas. Luego en fecha de cosecha se recolecta la totalidad de los frutos de los árboles identificados y registra el diámetro y

	Allen	Cip.I	C.Saltos	Fdez.Oro	Guerr.I	Guerr.II	J.J.Gomez
A75-76	370(74)	.	330(66)
A76-77	483(69)	.	511(73)
A77-78	438(73)	.	432(72)
A91-92	686(73)	.	.
A92-93	.	.	1,146(75)	.	1,217(74)	.	.
A96-97	248(31)	.	.	1,256(75)	.	.	.
A98-99	.	.	.	1,058(74)	.	.	.
A99-00	.	.	.	1,149(88)	.	.	.
A00-01	.	.	.	1,118(83)	.	.	.
A01-02	.	.	.	910(81)	.	.	.
A02-03	828(69)	.
A08-09	.	1,076(75)
A10-11	.	900(75)
A11-12	.	976(74)

Cuadro 5.1: Número de mediciones y frutos en Base de Modelación LONGITUDINALES

	Cip.I	Cip.II
A09-10	960(75)	.
A13-14	.	750(75)
A14-15	.	825(75)

Cuadro 5.2: Número de mediciones y frutos en Base de Validación LONGITUDINALES

el peso de cada uno de ellos.

A efectos de obtener métricas de validación externa, la información histórica se desagrega en una base de modelación, consistente en 6.540 frutos (cuadro 5.3) y otra base de validación externa consistente en 1.021 frutos (cuadro 5.4). La figura 5.1 representa la información disponible para la elaboración del modelo, mientras que en 5.2 se presenta las mediciones utilizadas para validar los modelos obtenidos.

	Cip.II	Cte.Cordero
A10-11	1,155	881
A11-12	990	1,277
A12-13	829	1,408

Cuadro 5.3: Número de frutos en Base de Modelación TRANSVERSALES

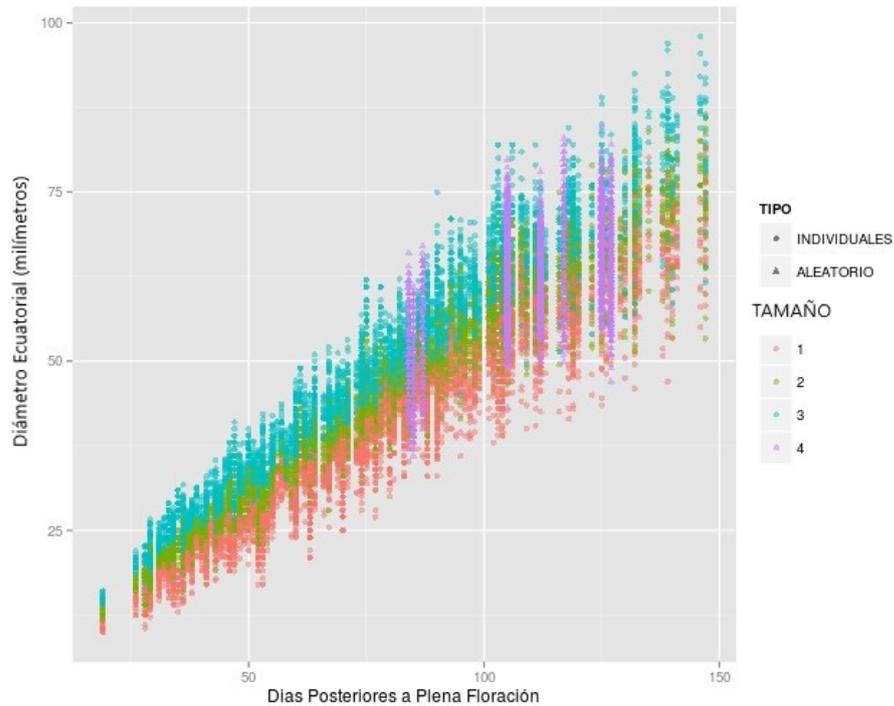


Figura 5.1: Mediciones de las Bases de Modelación LONGITUDINALES y TRANSVERSALES

	Cip.II	Cte.Cordero
A13-14	510	511

Cuadro 5.4: Número de frutos en Base de Validación TRANSVERSALES

5.3

INFORMACIÓN AGROCLIMÁTICA

A efectos de considerar posibles covariables ambientales, se procede al relevamiento de la información concerniente a la temperatura media diaria correspondiente a las 18 temporadas consideradas en los relevamientos de diámetro ecuatorial de frutos. Los registros disponibles se corresponden al promedio diario de la temperatura máxima y mínima registrada desde el primero de agosto hasta el 31 de enero para cada una de las temporadas en consideración obtenidos en la estación agroclimática del INTA Alto Valle. A partir de la temperatura diaria se generan 976 índices de termoacumulación (ITAC)

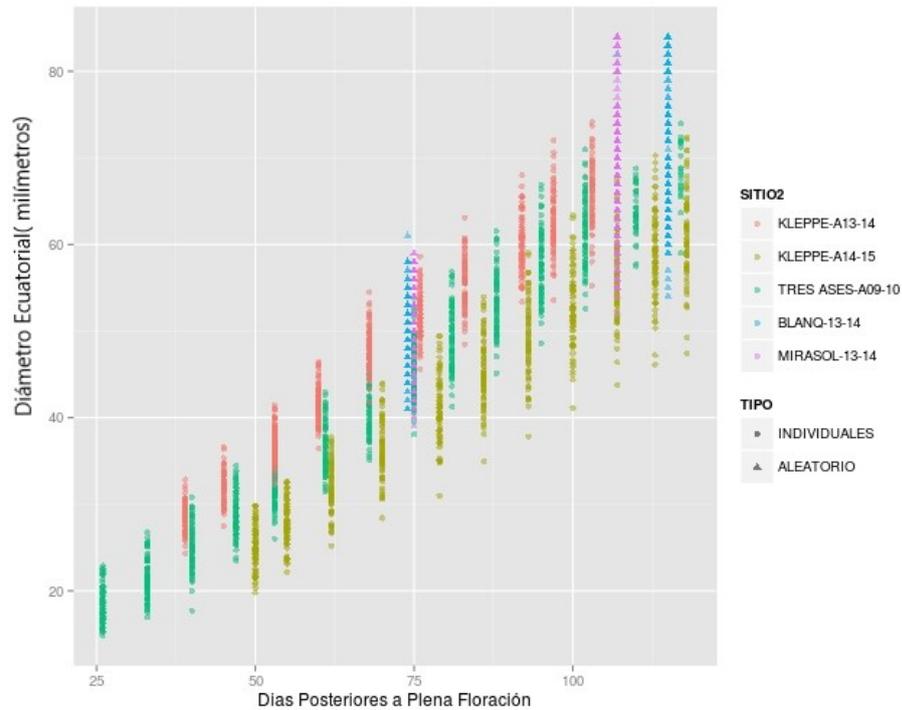


Figura 5.2: Mediciones de las Bases de Validación LONGITUDINALES y TRANSVERSALES

para distintos momentos de inicio de temporada y 594 ITAC para el mes de diciembre, en cada una de las temporadas. Estos índices se obtienen a partir de criterios que varían en función de: momento de inicio de la acumulación, momento de finalización de la acumulación y la temperatura base de acumulación. El rango de los criterios considerados es el siguiente:

Para los ITAC de inicio de Temporada:

- Inicio de la Acumulación: 44 a 110 días posteriores al primero de agosto de cada temporada (DDA).
- Fin de la Acumulación: 50 a 121 días DDA de cada temporada.
- Temperatura diaria Base: 0 a 20 grados centígrados.

Para los ITAC del mes de Diciembre:

- Inicio de la Acumulación: 122 a 145 días posteriores al primero de agosto de cada temporada (DDA).
- Fin de la Acumulación: 130 a 151 días DDA de cada temporada.
- Temperatura diaria Base: 0 a 20 grados centígrados.

5.4

MODELACIÓN

Las estimaciones correspondientes a los modelos candidatos se realizan a partir de las funciones NLME y LMER del software estadístico R.

El primero de ellos: NLME es una función del paquete NLME en su versión 3.1-117 que obtiene la verosimilitud a partir de una expansión de Taylor con el método de linealización BLUP propuesto por Linsdtrom et al. [80]. Los modelos se corresponden a NLME con efectos anidados o una estructura de efectos cruzados muy simples, con la posibilidad de modelar directamente la heterocedasticidad y la correlación de los residuales.

Una vez ajustado el modelo, las componentes de varianza se evalúan a partir de CI o LRT en una primera instancia para luego confirmarse, en lo posible, con intervalos de confianza obtenidos a partir del estadístico de Wald.

La significancia de los efectos fijos se evalúa a partir de pruebas t basadas en el estadístico Wald con grados de libertad obtenidos a partir del método Between-Within.

No se recomiendan, aquí, las pruebas LRT debido a que la maximización de la verosimilitud se realiza en los llamados pseudo-datos, es decir en la linealización, y no en los datos originales.

Otro paquete utilizado para obtener estimaciones es lme4 en su versión 1.1-7, a partir de la función NLMER. Esta calcula la función de verosimilitud ML a partir de la Aproximación de LAPLACE para NLME multinivel, tanto anidados como cruzados.

LMER recurre el paquete NLOPTR para realizar optimizaciones no lineales a partir de va-

rios algoritmos de optimización. Para la presente implementación el algoritmo PRAXIS, un algoritmo de optimización local libre de derivadas segundas, es el que se utiliza por presentar las mayores tasas de convergencia exitosas.

Además, para los casos en los que fue necesario, se generan variables dummy a los efectos de modelar los efectos fijos de los parámetros, en la imposibilidad de obtenerlos de forma directa con la función NLMER. Ello ha requerido el desarrollo de funciones que cumplan tal cometido que, junto a sus derivadas primeras, ha permitido un adecuado comportamiento del algoritmo de optimización.

Una vez finalizado el proceso de estimación con NLMER, se evalúa la estructura de las componentes de varianza a partir de los CI, mientras que la inferencia de los efectos fijos se realiza a partir de pruebas LRT.

Eventualmente se realiza para NLMER, un test de hipótesis para cada término fijo a partir de la prueba de Wald en una prueba t con distribución asintótica normal. La presencia de un gran número de sujetos en alguno de los niveles de efectos aleatorios considerados fundamenta esta estrategia de análisis.

Como los propósitos finales de los modelos obtenidos tienen objetivos predictivos, además de las pruebas usuales utilizadas para el proceso de modelación, la elección final está condicionada a la performance de los criterios predictivos (CP) obtenidos a partir de la Base de Modelación.

Los CP, para cada uno de los modelos candidatos, se obtienen a partir de funciones creadas a tal fin y obteniendo información requerida a partir de los objetos del tipo NLME o NLMER del paquete estadístico R.

También se han creado funciones en el paquete estadístico R que obtienen datos sintéticos a partir de simulaciones Monte Carlo de los modelos candidatos. Para la simulación multivariada en particular, además, se ha desarrollado una función denominada *MULTIRANDON* cuya implementación se fundamenta en el principio de descomposi-

ción espectral.

La base de covariables ambientales general se obtiene a partir de datos agrometeorológicos históricos en cada temporada, para ello se construyen funciones que obtienen valores representativos de cada una de las temporadas a partir de criterios de termoacumulación. El proceso de selección de covariables ambientales candidatas se desarrolla a partir las correlaciones entre los EBEs de efectos aleatorios a nivel de SITIO en los modelos candidatos y las covariables ambientales generales. La elección final se realiza a partir de apreciaciones visuales de la distribución de las covariables ambientales candidatas y su relación con los EBEs en consideración.

Las covariables ambientales seleccionadas se consideran en la construcción de nuevos modelos NLME en la forma de a_i y considerando presencias y/o ausencias de términos en la estructura de varianza de los efectos aleatorios.

— 5.5 —

CALIBRACIÓN

La calibración se realiza a partir de la función construida en R: BLUP_OF. La misma permite obtener la esperanza de b_i condicionada en $Y_{i,1}$, para el enfoque ZEP_b a partir de la ecuación:

$$\hat{b}_i = \hat{D}\hat{Z}_{i,1}^T(\hat{Z}_{i,1}\hat{D}\hat{Z}_{i,1}^T + \hat{R}_{i,1})^{-1}(Y_{i,1} - f(A_i\hat{\beta}, z_{i,1})) \quad (5.1)$$

mientras que el enfoque $EBLUP_b$ se implementa a partir de la ecuación:

$$\hat{b}_i = \hat{D}\hat{Z}_{i,1}^T(\hat{Z}_{i,1}\hat{D}\hat{Z}_{i,1}^T + \hat{R}_{i,1})^{-1}(Y_{i,1} - f(A_i\hat{\beta}, z_{i,1})) + Z_{i,1}(\hat{b}_i) \quad (5.2)$$

Donde \hat{b}_i es un vector de los parámetros aleatorios predichos en los niveles deseados. \hat{D} es la matriz de varianza-covarianza para \hat{b}_i . Cuando los \hat{b}_i poseen dos niveles, la matriz \hat{D} es una matriz diagonal en bloques que contiene a \hat{D}_1 y \hat{D}_2 con \hat{D}_2 repetido m_i veces

y donde \hat{D}_1 y \hat{D}_2 son las matrices de varianza-covarianza estimadas para los parámetros a nivel de SITIO y FRUTO respectivamente, para los m_i FRUTOS del SITIO i . $\hat{Z}_{i,1}$ es la matriz de derivadas parciales de la función no lineal con respecto a b_i . Las primeras g_i columnas de $\hat{Z}_{i,1}$ contienen las derivadas parciales de la función no lineal con respecto al nivel de SITIO, mientras que las restantes columnas contienen las derivadas parciales con respecto al nivel de FRUTO. Las matrices \hat{D} y $\hat{Z}_{i,1}$ pueden expresarse como:

$$\hat{D} = \begin{pmatrix} \hat{D}_1 & 0 & \cdots & 0 \\ 0 & \hat{D}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ 0 & 0 & 0 & \hat{D}_2 \end{pmatrix} \quad (5.3)$$

$$\hat{Z}_{i,1} = \begin{pmatrix} \hat{Z}_{i1} & \hat{Z}_{i1} & 0 & \cdots & 0 \\ \hat{Z}_{i2} & 0 & \hat{Z}_{i2} & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & 0 \\ \hat{Z}_{in_i} & 0 & 0 & & \hat{Z}_{in_i} \end{pmatrix} \quad (5.4)$$

Por otra parte, la matriz $\hat{R}_{i,1}$ representa la matriz de varianza-covarianza para el término del error, que puede expresarse a partir de una formulación general propuesta en la ecuación 4.25 a la cual se extrae el factor de escala σ^2 para la dispersión del error obteniendo:

$$\hat{R}_{i,1} = \sigma^2 G_{i,1}^{0,5} \Gamma_{i,1} G_{i,1}^{0,5} \quad (5.5)$$

donde $G_{i,1}$ es una matriz diagonal que describe la estructura de varianza, mientras que $\Gamma_{i,1}$ es una matriz de dimensión $n_i \times n_i$ que describe los patrones de correlación existentes.

PREDICCIONES

Para el caso de la predicción de las futuras observaciones, la función construida en R denominada PREDICHO posibilita obtenerlas utilizando el método PL_y a partir de la expresión:

$$\hat{Y}_{i,2} = f(A_i \hat{\beta}, z_{i,2}) + \hat{Z}_{i,2} \hat{b}_i \quad (5.6)$$

cuando las observaciones son independientes, mientras que cuando existe una estructura de correlación en los errores de tipo CAR(1) , esta se obtiene como:

$$\hat{Y}_{i,2} = f(A_i \hat{\beta}, z_{i,2}) + \hat{\rho}^d (Y_{i,1} - f(A_i \hat{\beta}, z_{i,1})) \quad (5.7)$$

Por otra parte, cuando el método utilizado para la predicción de futuras observaciones se corresponde a PNL_y , se propone la siguiente expresión para el caso de errores independientes:

$$\hat{Y}_{i,2} = f(A_i \hat{\beta} + B_i \hat{b}_i, z_{i,2}) \quad (5.8)$$

mientras que para errores con la estructura tipo CAR(1) tenemos:

$$\hat{Y}_{i,2} = f(A_i \hat{\beta}, z_{i,2}) - Z_{i,2} \hat{b}_i + \hat{\rho}^d (Y_{i,1} - f(A_i \hat{\beta}, z_{i,1}) + Z_{i,1} \hat{b}_i) \quad (5.9)$$

donde d es el número de días entre la fecha de predicción y la fecha correspondiente a la calibración del modelo, mientras que $\hat{\rho}$ expresa la correlación entre medidas de un mismo FRUTO por día transcurrido.

La totalidad de las funciones escritas e implementadas en este y en los apartados anteriores se transcriben en los Anexos de la presente Tesis por motivos de espacio.

Capítulo 6

RESULTADOS Y DISCUSIÓN

6.1

MODELACIÓN

6.1.1

INTRODUCCIÓN

El primer paso para desarrollar un modelo NLME consiste en proponer la componente estructural más adecuada para el conjunto de datos bajo estudio. Dicha ecuación debe expresar el diámetro ecuatorial en función de los DDPF, describiendo el período comprendido desde los 40 DDPF hasta 20 días posteriores a la cosecha. De acuerdo a Bramardi et al.[114],tal cual lo presentado en la ecuación 4.14, la función propuesta para los datos obtenidos puede expresarse como:

$$Y = \frac{1}{\beta_1 + \beta_2 * \beta_3^{DDPF}} \quad (6.1)$$

con:

- β_1 : parámetro relacionado inversamente con la asíntota superior.
- β_2 : parámetro que relaciona la asíntota superior (AS) con la inferior (AI), mediante

la función: e^{At} / AS .

- β_3 : parámetro relacionado con la tasa de crecimiento desde los valores ‘iniciales’ (determinados por la magnitud de β_2 y los valores ‘finales’ (determinados por la magnitud de β_1).

Con el propósito de mejorar las condiciones de convergencia de los algoritmos de optimización se reparametriza el modelo a través de una propuesta que posibilita obtener estimaciones con magnitudes equivalentes ya que la relación de valores absolutos con respecto a β_3 es de 1 en 10 para β_2 y de 1 en 100 para β_1 . El modelo propuesto es:

$$Y = \frac{1}{\beta_1 * 0,01 + (\beta_2 * 0,01)(\beta_3 * 0,1)^{DDPF}} \quad (6.2)$$

6.1.2

MODELACIÓN EN LAS MEDICIONES LONGITUDINALES

El diseño de la experiencias asociadas con las mediciones LONGITUDINALES posibilita la evaluación de diferentes factores implicados en el proceso de muestreo y que deben ser considerados al momento de proponer un modelo candidato.

En el presente trabajo la información combinada de varias temporadas y establecimientos se denota SITIO. Bajo esta premisa, es posible describir el proceso de selección de las unidades experimentales para cada uno de los SITIOS como sigue: se eligen al azar ÁRBOLES y de estos se toman FRUTOS al azar de 3 estratos de TAMAÑOS determinados en función del desarrollo que presentan los frutos en el correspondiente SITIO. A los FRUTOS seleccionados se los identifica para luego medirles, regularmente en el tiempo, su diámetro ecuatorial.

En función de ello es factible plantear para el modelo 4.14, efectos fijos de TAMAÑO para los parámetros de la componente estructural. El modelo propuesto completo puede expresarse como:

$$Y_{ijklmn} = \frac{1}{(\beta_{1,TAMA_k} + b_{1,jklm}) *_{0,01} + ((\beta_{2,TAMA_k} + b_{2,jklm}) *_{0,01})((\beta_{3,TAMA_k} + b_{3,jklm}) *_{0,1})^{DDPF}} + \varepsilon_{ijklmn} \quad (6.3)$$

con

$$b_{i,jklm} = b_{i,SITIO_j} + b_{i,ARB(SITIO)_{l(j)}} + b_{i,TAMA*ARB(SITIO)_{kl(j)}} + b_{i,FRUTO_{m(jkl)}} \quad (6.4)$$

dónde:

- $\beta_{i,TAMA_k}$ efecto del k-ésimo TAMAÑO para el i-ésimo parámetro del efecto fijo.
- $b_{i,SITIO_j}$ efecto del j-ésimo SITIO del i-ésimo parámetro del efecto aleatorio con $b_{i,SITIO_j} \sim N([0, 0, 0], D_{SITIO})$.
- $b_{i,ARB(SITIO)_{l(j)}}$ efecto del l-ésimo ARBOL dentro del j-ésimo SITIO del i-ésimo parámetro del efecto aleatorio con $b_{i,ARB(SITIO)_{l(j)}} \sim N([0, 0, 0], D_{ARB})$.
- $b_{i,TAMA*ARB(SITIO)_{kl(j)}}$ Interacción entre el k-ésimo TAMAÑO y el l-ésimo ARBOL correspondientes al j-ésimo SITIO del i-ésimo parámetro del efecto aleatorio con $b_{i,TAMA*ARB(SITIO)_{kl(j)}} \sim N([0, 0, 0], D_{TAMA*ARB})$.
- $b_{i,FRUTO_{m(jkl)}}$ efecto del m-ésimo FRUTO del j-ésimo SITIO, el k-ésimo TAMAÑO y el l-ésimo ARBOL del efecto aleatorio con $b_{i,FRUTO_{m(jkl)}} \sim N([0, 0, 0], D_{FRUTO})$.
- ε_{ijkl} error aleatorio con $\varepsilon_{ijkl} \sim N(0, \sigma_\varepsilon^2)$ y $\sigma_\varepsilon^2 = \sigma^2 I_{n_i \times n_i}$, donde $I_{n_i \times n_i}$ es una matriz Identidad y n_i es el número de medidas repetidas por FRUTO.

con D_{SITIO} , D_{ARB} , $D_{TAMA*ARB}$, D_{FRUTO} matrices de varianza-covarianza de los efectos aleatorios en sus correspondientes niveles.

Se generan efectos aleatorios en los factores ARBOL, FRUTO y TAMAÑO*ARBOL debido a que media en ellos un proceso de muestreo aleatorio, el factor SITIO se considera efecto aleatorio con el propósito de poder ampliar el espacio de inferencia más allá de la

combinación de temporadas y establecimientos considerados.

MODELADO INICIAL

A partir del diseño original, se plantean modelos alternativos que consideran $\beta_{1,TAMA}$, $\beta_{2,TAMA}$, $\beta_{3,TAMA}$ y los efectos aleatorios para cada uno de los parámetros a saber:

- MODELO 1: Considerando solo efecto aleatorio a nivel de FRUTO: $b_{i,FRUTO_{m(jkl)}}$.
- MODELO 2: Considerando efecto aleatorio a nivel de SITIO: $b_{i,SITIO_j}$ y a nivel de FRUTO: $b_{i,FRUTO_{m(jkl)}}$.
- MODELO 3: Considerando efecto aleatorio a nivel de SITIO: $b_{i,SITIO_j}$, a nivel de TAMAÑO*ARBOL: $b_{i,TAMA*ARB(SITIO)_{kl(j)}}$ y a nivel de FRUTO: $b_{i,FRUTO_{m(jkl)}}$.
- MODELO 4: Considerando efecto aleatorio a nivel de SITIO: $b_{i,SITIO_j}$, a nivel de ARBOL: $b_{i,ARB(SITIO)_{l(j)}}$ y a nivel de FRUTO: $b_{i,FRUTO_{m(jkl)}}$.

Los modelos propuestos se estiman a partir de la base de modelación de las medidas LONGITUDINALES a través del procedimiento LMER 4.1-7 del paquete estadístico R y utilizando estimación máximo verosímil (ML), los CI obtenidos se presentan en el cuadro 6.1. Los valores de AIC y BIC obtenidos sugieren a MODELO 2 como un primer candida-

MODELO	Df	AIC	BIC	logLik
1	16	53.934	54.056	-26.951
2	22	51.986	52.154	-25.971
4	28	54.364	54.577	-27.154
3	34	60.423	60.682	-30.178

Cuadro 6.1: CI de Modelos Propuestos estimados por ML en Base de Modelación LONGITUDINALES

to a ser contemplado, descartando por ello a los efectos aleatorios posibles en ARBOL Y TAMAÑO*ARBOL. Las estimaciones de los parámetros se aprecian en el Cuadro 6.6.

INCIDENCIA DE LOS EFECTOS ALEATORIOS SOBRE EL DIÁMETRO FINAL

Amén de lo expresado anteriormente en base a los CI, es importante considerar otros aspectos de los modelos propuestos debido a los propósitos predictivos que se requiere en el modelo final obtenido. Uno de ellos es la incidencia que los efectos aleatorios poseen en cada uno de los parámetros sobre la distribución marginal del modelo inducido a un tiempo dado.

Para evaluar dicha incidencia, se generan datos sintéticos con diámetros ecuatoriales de frutos en base a un modelo que considera b_i a nivel de FRUTO en función de la combinación de varios escenarios contemplados:

- Dos fechas de cosecha (109 y 116 DDPF).
- Presencia y Ausencia de efectos aleatorios.

Para cada escenario posible, se realizan 1.000 simulaciones Monte Carlo, los resultados obtenidos permiten obtener estimaciones de la variabilidad potencial del diámetro ecuatorial para cada uno de ellos. Los resultados se expresan como Coeficientes de Variación (CV) obtenidos como la razón entre el desvío estandar y el valor medio del Diámetro Ecuatorial a cosecha (cuadros 6.2 y 6.3).

	CV _{b_i} Simulados							
	Parcial						Total	
	Presencia			Ausencia			Presencia	Ausencia
MOMENTO (DDPF)	$b_1^{(1)}$	$b_2^{(2)}$	$b_3^{(3)}$	$b_1^{(4)}$	$b_2^{(5)}$	$b_3^{(6)}$	$(b_1 + b_2 + b_3)^{(7)}$	$(e)^{(8)}$
109	19.74%	5.44%	13.18%	13.86%	12.65%	19.94%	12.94%	1.54%
116	19.90%	4.80%	12.33%	12.84%	12.84%	20.01%	12.97%	1.46%

Cuadro 6.2: CV del Diámetro Ecuatorial de Frutos

Los resultados evidencian elevados valores de cambio potencial para b_1 y b_3 cuando estos se incorporan como únicos efectos aleatorios a nivel de FRUTO, con magnitudes

MOMENTO (DDPF)	Presencia vs. e			Ausencia vs. $b_1 + b_2 + b_3$		
	(1)-(8)	(2)-(8)	(3)-(8)	(4)-(7)	(5)-(7)	(6)-(7)
109	18.19%	3.89%	11.63%	0.92%	-0.29%	7.00%
116	18.44%	3.34%	10.87%	-0.13%	-0.12%	7.05%

Cuadro 6.3: Cambios de CV de Diámetros Ecuatoriales en Frutos Simulados

máximas de 18.44% y 10.87% respectivamente.

Otro aspecto a considerar son los cambios, expresados en CV, que se obtienen como producto de descartar efectos aleatorios en alguno de los parámetros. Un análisis del Cuadro 6.3 revela que las mayores magnitudes se obtienen en b_3 con un valor máximo de 7.05%. Los efectos de b_1 y b_2 le siguen con valores máximos de 0.92% y -0.29% respectivamente.

INCIDENCIA DE LOS EFECTOS ALEATORIOS EN CADA NIVEL CONSIDERADO

Es necesario, además, evaluar la importancia relativa que cada efecto aleatorio posee en cada uno de los niveles considerados. Los resultados se presentan como Coeficientes de Variación de b_i (CV_{b_i}) obtenidos como razón entre el desvío estandar de b_i y el valor absoluto del correspondiente efecto fijo de la componente estructural. Los valores obtenidos a partir del MODELO 4 se presentan en el Cuadro 6.4.

NIVEL	CV_{b_i}		
	b_1	b_2	b_3
SITIO	3.92%	10.36%	0.19%
ARBOL:SITIO	5.71%	2.81%	0.13%
FRUTO:TAMAÑO:SITIO	10.88%	5.25%	0.26%

Cuadro 6.4: Valores de CV_{b_i} de MODELO 4

Del Cuadro 6.4 es posible apreciar para cada coeficiente que:

- b_1 predomina en el nivel FRUTO (10,88%) y en ARBOL (5,71%).

- b_2 predomina en el nivel SITIO (10,36%) y en FRUTO (5,25%).
- b_3 predomina en el nivel FRUTO (0,26%) y en SITIO (0,19%).

Aquí, surge que el factor FRUTO se encuentra presente en todos los coeficientes mientras que SITIO prevalece en b_2 y b_3 .

A raíz de los resultados obtenidos en los cuadros 6.2, 6.3 y 6.4, los efectos aleatorios b_3 y b_1 en los factores SITIO y FRUTO son aquellos que mayor incidencia tienen en la variable de respuesta media a tiempo de cosecha. Estos hallazgos corroboran la elección inicial de MODELO 2 en 6.1.2 como el mejor modelo candidato al momento.

ESTRUCTURA DE CORRELACIÓN EN LOS ERRORES

Otro aspecto a considerar en el proceso de construcción del modelo reside en la modelización directa de la correlación existente en las medidas repetidas. La acción de los efectos aleatorios considerados genera de manera indirecta correlaciones en la variable de respuesta, pero el gran número de mediciones efectuadas durante el ciclo de crecimiento para un mismo fruto (un promedio de 11) conlleva al planteo de la eventual necesidad de un tratamiento específico en la estructura del error residual. Actualmente no se han desarrollado rutinas que posibiliten una modelización directa a partir del paquete LMER, por lo que se propone como estrategia obtener nuevas estimaciones de MODELO 2 en el paquete NLME del mismo software estadístico, denominado MODELO 2b. Este se expresa como:

$$Y_{jmn} = \frac{1}{(\beta_{1,TAMA_k} + b_{1,jm}) * 0,01 + ((\beta_{2,TAMA_k} + b_{2,jm}) * 0,01)((\beta_{3,TAMA_k} + b_{3,jm}) * 0,1)^{DDPF}} + \varepsilon_{jmn} \quad (6.5)$$

con

$$b_{i,jm} = b_{i,SITIO_j} + b_{i,FRUTO_{m(j)}} \quad (6.6)$$

dónde:

- $\beta_{i,TAMA_k}$ efecto del k-ésimo TAMAÑO para el i-ésimo parámetro del efecto fijo.
- $b_{i,SITIO_j}$ efecto del j-ésimo SITIO del i-ésimo parámetro del efecto aleatorio con $b_{i,SITIO_j} \sim N([0, 0, 0], D_{SITIO})$.
- $b_{i,FRUTO_{m(j)}}$ efecto del m-ésimo FRUTO del j-ésimo SITIO del efecto aleatorio con $b_{i,FRUTO_{m(j)}} \sim N([0, 0, 0], D_{FRUTO})$.
- ε_{jmn} error aleatorio con $\varepsilon_{jm} \sim N(0, \sigma_\varepsilon^2)$ y $\sigma_\varepsilon^2 = \sigma^2 I_{n_i \times n_i}$, donde $I_{n_i \times n_i}$ es una matriz Identidad y n_i es el número de medidas repetidas por FRUTO.

con D_{SITIO} y D_{FRUTO} matrices de varianza-covarianza de los efectos aleatorios en sus correspondientes niveles. En el Cuadro 6.6 se aprecian los valores estimados y su correlograma correspondiente (Figura 6.1).

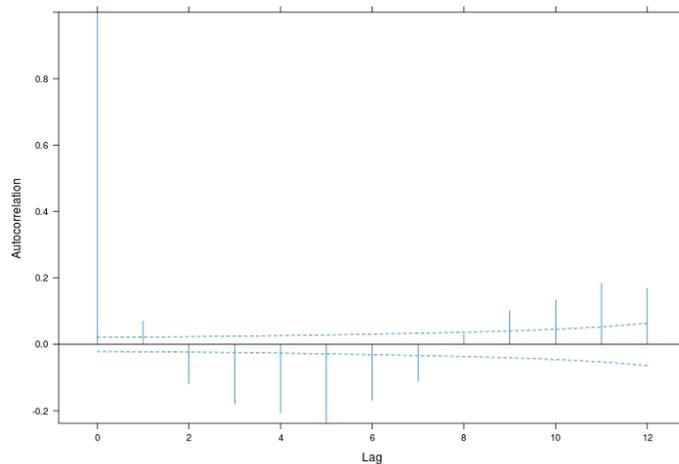


Figura 6.1: Correlograma del MODELO 2b

Los modelos propuestos son:

- MODELO 8: Ídem al MODELO 2b, pero considerando una Autocorrelación de Primer Orden (AR(1)) en el término del error.
- MODELO 9: Ídem al MODELO 2b, pero contemplando una Autocorrelación Continua de Primer Orden en el término del error (CAR(1)) con DDPF como covariable de tiempo.
- MODELO 10: Ídem al MODELO 2b, pero contemplando una Autocorrelación Continua de Primer Orden en el término del error (CAR(1)) con el logaritmo natural de DDPF como covariable de tiempo.

Los CI obtenidos a través de máxima verosimilitud se presentan en términos comparativos al MODELO 2b en el Cuadro 6.5. Los valores de CI indican al MODELO 9 como el mejor

MODELO	Df	AIC	BIC	logLik
2b	22	51.649	51.817	-25.802
8	23	50.106	50.281	-25.030
9	23	49.083	49.259	-24.518
10	23	49.825	50.001	-24.889

Cuadro 6.5: CI de modelos estimados por ML considerando Correlación en los Errores

modelo candidato, sin embargo este resultado debe tomarse como precaución ya que la verosimilitud marginal sobre la cual se obtienen estos criterios tiene su origen en pseudo-datos, y no en los datos originales, un planteo que también se extiende a la prueba LRT.

La distribución normal asintótica del criterio de Wald se constituye aquí, como una alternativa válida a considerar. Al fundamentarse el criterio de Wald en la curvatura de verosimilitud y no en los valores absolutos de esta, es posible conservar propiedades asintóticas deseables. El Cuadro 6.6 contiene los parámetros ajustados junto a sus correspondientes intervalos de confianza obtenidos a partir del criterio de Wald. El correlograma correspondiente al MODELO 9 indica la persistencia de correlaciones residuales en el término del error (Figura 6.2). Una verificación gráfica de supuestos se aprecia en la Figura B.1 del Anexo para el MODELO 9.

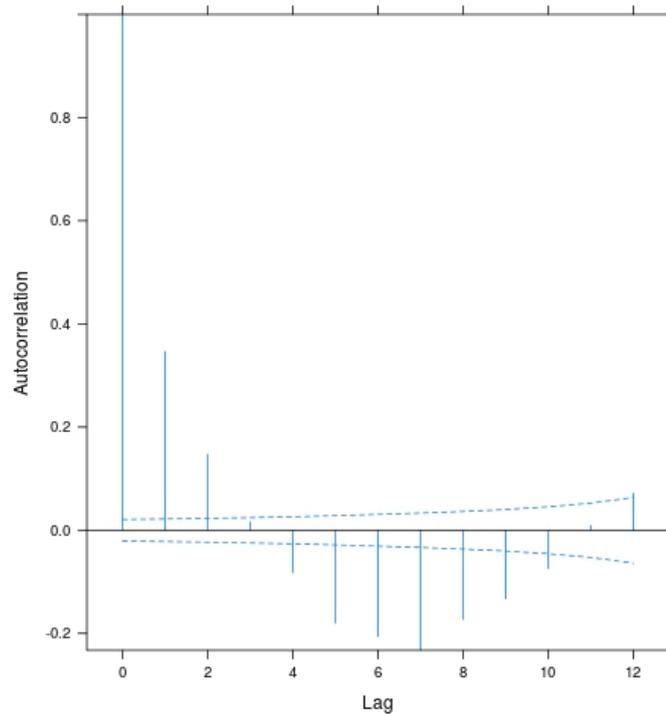


Figura 6.2: Correlograma de MODELO 9

6.1.3

INCORPORACIÓN DE MEDIDAS TRANSVERSALES

La incorporación de medidas TRANSVERSALES implica la consideración de un nuevo esquema de muestreo, donde la información relevada proviene de diámetros ecuatoriales de la totalidad de frutos presentes en árboles representativos durante el mes de diciembre y en fecha de recolección comercial. El nuevo esquema de muestreo se agrega al antes contemplado por las medidas LONGITUDINALES que caracterizaban una selección aleatoria de frutos inicial condicionada por categorías de tamaño prefijadas, su identificación y posterior seguimiento a lo largo de ciclo de crecimiento a intervalos regulares.

El propósito de mancomunar la información de ambos esquemas consiste en potenciar las fortalezas disponibles en cada una de las fuentes de datos: las medidas LONGITUDINALES contribuyen a estimar los parámetros de forma, mientras que las medidas TRANS-

PARÁMETROS	ESTIMACIÓN MODELO		
	2	2b	9
$\beta_{1_INTERCEPT}$	1.284 (**)	1.274 (**)	1.275 (**) (1.236,1.314)
$\beta_{1_TAMA\ 2}$	-0.114 (**)	-0.110 (**)	-0.101 (**) (-0.127,-0.074)
$\beta_{1_TAMA\ 3}$	-0.213 (**)	-0.020 (**)	-0.197 (**) (-0.223,-0.171)
$\beta_{2_INTERCEPT}$	8.258 (**)	10.388 (**)	8.305 (**) (7.858,8.753)
$\beta_{2_TAMA\ 2}$	-0.669 (**)	-1.633 (**)	-0.715 (**) (-0.847,-0.582)
$\beta_{2_TAMA\ 3}$	-1.396 (**)	-3.116 (**)	-1.426 (**) (-1.553,-1.300)
$\beta_{3_INTERCEPT}$	9.731 (**)	9.732 (**)	9.732 (**) (9.721,9.743)
$\beta_{3_TAMA\ 2}$	0.004 (NS)	0.004 (NS)	0.003 (NS) (-0.002,0.008)
$\beta_{3_TAMA\ 3}$	0.008 (NS)	0.005 (NS)	0.005 (NS) (0.001,0.0108)
$\sigma_{b_{1,SITIO_j}}$	0.072	0.195	0.074 (0.051,0.106)
$\sigma_{b_{2,SITIO_j}}$	1.020	1.075	0.967 (0.802,1.166)
$\sigma_{b_{3,SITIO_j}}$	0.022	0.037	0.023 (0.167,0.033)
$r_{b_{12,SITIO_j}}$	-0.06	0.80	-0.08 (-0.048,0.341)
$r_{b_{13,SITIO_j}}$	-0.66	-0.87	-0.69 (-0.879,-0.329)
$r_{b_{23,SITIO_j}}$	-0.26	-0.85	-0.23 (-0.615,0.246)
$\sigma_{b_{1,FRUTO_{m(j)}}$	0.175	0.115	0.167 (0.159,0.176)
$\sigma_{b_{2,FRUTO_{m(j)}}$	0.734	1.119	0.551 (0.504,0.603)
$\sigma_{b_{3,FRUTO_{m(j)}}$	0.033	0.537	0.028 (0.026,0.030)
$r_{b_{12,FRUTO_{m(j)}}$	-0.66	-0.24	-0.99 (-0.999,-0.971)
$r_{b_{13,FRUTO_{m(j)}}$	-0.82	-0.04	-0.85 (-0.872,-0.821)
$r_{b_{23,FRUTO_{m(j)}}$	0.31	-0.07	0.81 (0.702,0.881)
ρ	.	.	0.94 (0.935,0.946)
σ_{RES}	0.936	0.930	1.248 (1.213,1.284)

$\sigma_{b_{1,SITIO_j}}$, $\sigma_{b_{2,SITIO_j}}$ y $\sigma_{b_{3,SITIO_j}}$ son efectos aleatorios a nivel de SITIO.

$\sigma_{b_{1,FRUTO_{m(j)}}$, $\sigma_{b_{2,FRUTO_{m(j)}}$ y $\sigma_{b_{3,FRUTO_{m(j)}}$ son efectos aleatorios a nivel de FRUTO

(NS): No Significativo, (**): Altamente Significativo (p<0.01).

Test t e intervalos basados en el estadístico de Wald

Cuadro 6.6: Parámetros estimados en Modelos estimados por REML en Base de Modelación LONGITUDINALES

VERSALES contribuyen a mejorar la estimación de la variabilidad de los efectos aleatorios generalizando la inferencia a una población mas amplia [105], [126].

Los nuevos modelos candidatos incorporan una categoría adicional en la variable TAMAÑO y una variable denominada TIPO que discrimina los esquemas de muestreo LONGITUDINAL y TRANSVERSAL. Las estimaciones se realizan a partir del procedimiento NLME con las siguientes propuestas:

- MODELO 11: Considerando como efectos Fijos: $\beta_{1_{TAMA_k}}$, $\beta_{3_{TAMA_k}}$. Efectos Aleatorios a nivel de SITIO: $b_{3,SITIO_j}$ y a nivel de FRUTO: $b_{1,FRUTO_{m(j)}}$, $b_{2,FRUTO_{m(j)}}$ y $b_{3,FRUTO_{m(j)}}$.
- MODELO 12: Ídem al MODELO 11, pero considerando varianzas que cambian para cada TIPO en el término del error.
- MODELO 13: Ídem al MODELO 12, pero considerando una estructura CAR(1) sobre $\log(DDPF)$ en el término del error.
- MODELO 14: Considerando como efectos Fijos: $\beta_{1_{TAMA_k}}$, $\beta_{3_{TAMA_k}}$. Efectos Aleatorios a nivel de SITIO: $b_{TRANS,3,SITIO_j}$, $b_{LONGI,3,SITIO_j}$ y a nivel de FRUTO: $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,2,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.
- MODELO 15: Considerando como efectos Fijos: $\beta_{1_{TAMA_k}}$, $\beta_{3_{TIPO_r}}$. Efectos Aleatorios a nivel de SITIO: $b_{TRANS,3,SITIO_j}$, $b_{LONGI,3,SITIO_j}$ y a nivel de FRUTO: $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,2,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.
- MODELO 16: Ídem al MODELO 15, pero considerando varianzas que cambian para cada TIPO en el término del error.

Los CI obtenidos a partir de los modelos estimados con el procedimiento NLME se aprecian en el Cuadro 6.7.

MODELO	Df	AIC	BIC	logLik	Coc. Verom.	p-valor
11	20	92.829	92.989	-46.394		
12	18	92.238	92.382	-46.101		
13	19	90.584	90.735	-45.273		
14	20	92.559	92.719	-46.259		
15	16	92.612	92.685	-46.262	5.59	0.21
16	21	93.974	94.085	-46.973		

Cuadro 6.7: CI en Modelos con Información Combinada estimados por ML

Aquí el MODELO 13, el cual considera la correlación serial, se propone como mejor candidato. Sin embargo, un estudio de CP pone en evidencia la pobre performance pre-

dictiva del MODELO 13 como se puede apreciar en el Cuadro 6.8.

Del Cuadro 6.8, antes citado, surgen los MODELOS 14,15 y 16 como aquellos con mejores métricas predictivas en la validación interna. Una prueba de Wald sobre los parámetros estimados en los efectos fijos del MODELO 14 pone en evidencia la escasa significancia de algunos de los términos de TAMAÑO en $\beta_{3_{TAMA_k}}$ tal como se aprecia en el Cuadro 6.11. Los mejores modelos candidatos, se constituyen de esta forma, como los MODELOS 15 y 16 al modelar β_3 a partir de la variable TIPO. Una verificación gráfica de supuestos del MODELO 15, uno de los modelos propuestos, se aprecia en la Figura B.2 del Anexo.

MODELO	R^2 Cond	CC_1	CC_2	RMSE	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD
11	0.9979	0.9989	0.9989	0.7060	0.0331	-0.070	0.4865	1.250	0.705
12	0.9954	0.9977	0.9976	1.0474	-0.0579	-0.283	0.7587	1.764	1.045
13	0.9955	0.9977	0.9977	1.0274	-0.0100	-0.132	0.7811	1.884	1.027
14	0.9980	0.9989	0.9989	0.6914	-0.0121	-0.147	0.4615	1.216	0.691
15	0.9980	0.9989	0.9989	0.6914	-0.0121	-0.147	0.4615	1.216	0.691
16	0.9980	0.9989	0.9989	0.6899	-0.0115	-0.137	0.4605	1.208	0.690

Cuadro 6.8: Capacidad Predictiva en Modelos con Información Combinada

Las inconsistencias encontradas entre lo sugerido por los CI y los CP ya han sido encontrados en trabajos donde se modela la estructura de varianzas-covarianzas del error. Así es como Huang S. et al. [107] encontraron estas discrepancias en modelos NLME considerando y sin considerar la correlación serial de las medias repetidas en pinos contorta (*Pinus contorta var. latifolia*). Resultados similares han sido hallados recientemente en abetos negros por Meng et al. [111] para los mismos modelos. Este estudio se amplía para NLME con una estructura generalizada en la matriz del error residual en otro estudio de Huang et al. [106]. En los trabajos citados se enfatiza el propósito bajo el cual se debe juzgar a los modelos candidatos porque ello determina la manera en la cual estos son seleccionados. De manera tal que si el propósito es obtener modelos descriptivos,

los CI se constituyen como una métrica adecuada de selección, mientras que si modelos tienen propósitos predictivos el proceso de selección debe ser conducido por CP obtenidos a partir de validaciones internas y externas. La totalidad de los estudios citados coinciden en sus conclusiones con el hecho que los NLME con la estructura más simple en la varianza residual (una estructura identidad) obtienen los peores CI, pero las mejores performances predictivas. Este comportamiento no ha ocurrido en modelos que solo consideran efectos fijos y han sido sujetos a modelación en la estructura del error, tal como lo obtenido por Meng.et al. [111].

En la presente tesis el efecto aleatorio FRUTO compite con la modelación de la estructura residual del error cuando se considera la correlación serial en el tiempo para un mismo FRUTO en la varianza del modelo marginal inducido tal lo expresado en las conclusiones obtenidas a partir de la ecuación 4.34, una situación que necesariamente genera condiciones de sobreajuste que termina explicando los descalces obtenidos entre los CI y los CP.

Una situación similar se presenta al momento de estimar más de un efecto aleatorio a partir de frutos que poseen una única medición, tal como ocurre con las medidas TRANSVERSALES. Esto se aprecia claramente en los estudios de residuales tal como el presentado en la Figura B.3 para el caso del MODELO 11. La estrategia propuesta para salvar el citado condicionamiento consiste en generar efectos aleatorios diferenciales (b_i) que consideran la naturaleza LONGITUDINAL y TRANSVERSAL de las fuentes de muestreo tal como los MODELOS 14,15 y 16 generados. En todos ellos no se modela la estructura de la varianza residual por lo que es factible realizar propuestas de modelos candidatos y estimaciones que consideren al paquete LMER 4.1-7 del software estadístico R. En los siguientes modelos candidatos se modelan los efectos fijos de la componente estructural considerando las variantes involucradas en la variable TIPO:

- MODELO 17: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$. Efectos Aleatorios

a nivel de SITIO: $b_{TRANS,3,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.

■ MODELO 18: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$. Efectos aleatorios a nivel de SITIO: $b_{LONGI,3,SITIO_j}$ y a nivel de FRUTO: $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.

■ MODELO 19: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$. Efectos aleatorios a nivel de SITIO: $b_{LONGI,3,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.

■ MODELO 20: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$. Efectos aleatorios a nivel de SITIO: $b_{LONGI,3,SITIO_j}$, $b_{TRANS,3,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.

■ MODELO 21: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$. Efectos aleatorios a nivel de SITIO: $b_{2,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.

En el Cuadro 6.9 se aprecian los CI, a partir de los cuales el MODELO 20 se propone como el más parsimonioso, seguido por MODELO 21. La apreciación de los CP (Cuadro 6.10) evidencia que los CP sugieren también a estos mismos modelos, con valores muy similares. Los parámetros estimados se presentan en el Cuadro 6.11 .

MODELO	Df	AIC	BIC	logLik
17	11	167.340	167.428	-83.659
18	10	179.950	180.030	-89.965
19	11	171.105	171.193	-85.541
20	12	108.509	108.604	-54.242
21	11	116.460	116.548	-58.219

Cuadro 6.9: CI en Modelos con Información Combinada II estimados por ML

MODELO	R ² Cond	CC ₁	CC ₂	RMSE	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD
17	0.9981	0.9990	0.9990	0.6695	-0.0073	-0.1260	0.4070	1.102	0.667
18	0.9963	0.9981	0.9881	0.9366	-0.0121	-0.2415	0.5615	1.641	0.937
19	0.9981	0.9990	0.9990	0.6707	-0.0073	-0.1278	0.4073	1.105	0.671
20	0.9979	0.9999	0.9999	0.6939	-0.0077	-0.1359	0.4241	1.146	0.694
21	0.9979	0.9999	0.9999	0.6931	-0.0113	-0.1614	0.4262	1.156	0.694

Cuadro 6.10: Capacidad Predictiva en Modelos con Información Combinada II

PARÁMETROS	ESTIMACIÓN MODELO			
	14	15	20	21
$\beta_{1_INTERCEPT}$	1.298 (**)	1.296 (**)	1.334 (**)	-5.578 (**)
$\beta_{1_LONGITUDINAL}$.	.	0.067 (**)	7.749 (**)
$\beta_{1_TAMA\ 2}$	-0.152 (**)	-0.148 (**)	.	.
$\beta_{1_TAMA\ 3}$	-0.292 (**)	-0.292 (**)	.	.
$\beta_{1_TAMA\ 4}$	-0.866 (**)	-0.864 (**)	.	.
$\beta_{2_INTERCEPT}$	7.958 (**)	7.957 (**)	5.81 (**)	9.012 (**)
$\beta_{3_INTERCEPT}$	9.733 (**)	9.905 (**)	9.739 (**)	9.984 (**)
$\beta_{3_LONGITUDINAL}$.	-0.172 (**)	-0.004 (NS)	-0.248 (**)
$\beta_{1_TAMA\ 2}$	0.001 (NS)	.	.	.
$\beta_{1_TAMA\ 3}$	0.0004 (NS)	.	.	.
$\beta_{1_TAMA\ 4}$	0.172 (**)	.	.	.
$\sigma_{b_{2_SITIO_j}}$.	.	.	1.276
$\sigma_{b_{3_SITIO_j}}$	0.0189	0.017	.	.
$\sigma_{b_{TRANS,3_SITIO_j}}$.	.	0.035	.
$\sigma_{b_{LONGI,3_SITIO_j}}$.	.	0.562	.
$\sigma_{b_{LONGI,1_FRUTO_{m(j)}}$	0.197	0.197	0.0352	0.004
$\sigma_{b_{LONGI,2_FRUTO_{m(j)}}$	1.403	1.403	.	1.019
$\sigma_{b_{2_FRUTO_{m(j)}}$.	.	5.818	1.019
$\sigma_{b_{LONGI,3_FRUTO_{m(j)}}$	0.041	0.041	0.001	0.003
$\sigma_{b_{TRANS,3_FRUTO_{m(j)}}$.	0.011	0.005	0.001
$r_{b_{12_FRUTO_{m(j)}}$	-0.448	-0.448	.	.
$r_{b_{13_FRUTO_{m(j)}}$	-0.732	-0.732	.	.
$r_{b_{23_FRUTO_{m(j)}}$	-0.183	-0.182	.	.
σ_{RES}	0.934	0.934	0.562	0.568

$\sigma_{b_{X,1,i}}, \sigma_{b_{X,2,i}}$ y $\sigma_{b_{X,3,i}}$ son efectos aleatorios a nivel de SITIO.

$\sigma_{b_{X,1,ij}}, \sigma_{b_{X,2,ij}}$ y $\sigma_{b_{X,3,ij}}$ son efectos aleatorios a nivel de FRUTO

(NS): No Significativo, (**): Altamente Significativo (p<0.01). Test t basado en el estadístico de Wald con distribución asintótica normal

Cuadro 6.11: Parámetros Estimados en Modelos con Información Combinada estimados por REML

6.1.4

INCORPORACIÓN DE COVARIABLES AMBIENTALES

A efectos de incrementar la capacidad predictiva, se evalúa la factibilidad de incorporar covariables ambientales que posean una cierta capacidad explicativa por sobre b_i

a nivel de SITIO. Para ello se consideran los índices de termoacumulación (ITAC) en dos momentos de ciclo de crecimiento de los FRUTOS: inmediatamente posterior a plena floración y durante el transcurso del mes de diciembre. Motiva la generación diferencial de bases de ITAC, estudios previos realizados por Tassile et al. [124] que demuestran el efecto diferencial que generan los ITAC en función del momento de ciclo en el cual se encuentran los frutos. Altas temperaturas dan lugar a frutos con calibre mayor cuando estas se producen en etapas tempranas del crecimiento de los mismos, mientras que su ocurrencia durante el transcurso del mes de diciembre dan lugar a frutos con menores calibres comerciales.

La base que contiene las condiciones térmicas en las etapas previas del crecimiento contiene 976 ITAC, mientras que la base que expresa las condiciones transcurridas durante el mes de diciembre contiene 594 ITAC.

A partir de los ITAC obtenidos, se propone un procedimiento de selección de variables en base a su relación con los EBEs de $b_{LONGI,3,SITIO_j}$ y $b_{TRANS,3,SITIO_j}$ del MODELO 20 y $b_{2,SITIO_j}$ del MODELO 21 en ambos momentos considerados. La magnitud en la correlación lineal, los patrones sistemáticos aparentes y la distribución de los ITAC dentro del rango de estudio puede apreciarse gráficamente en las Figuras 6.3, 6.4, 6.5 para las ITAC candidatas resultantes.

A partir de la información recopilada, se consideran las siguientes ITAC:

- B.46.T.100.MIN.16 para β_3 durante las Etapas Tempranas del Ciclo de Crecimiento, obtenida por la acumulación diaria de temperaturas que superan los 16 grados centígrados desde los 46 DDA hasta los 100 DDA. Si lo expresamos en edad fisiológica del fruto, el criterio se corresponde a una acumulación de temperaturas diarias que superan los 16 Grados centígrados desde 12 días anteriores a plena floración hasta 42 DDPF.

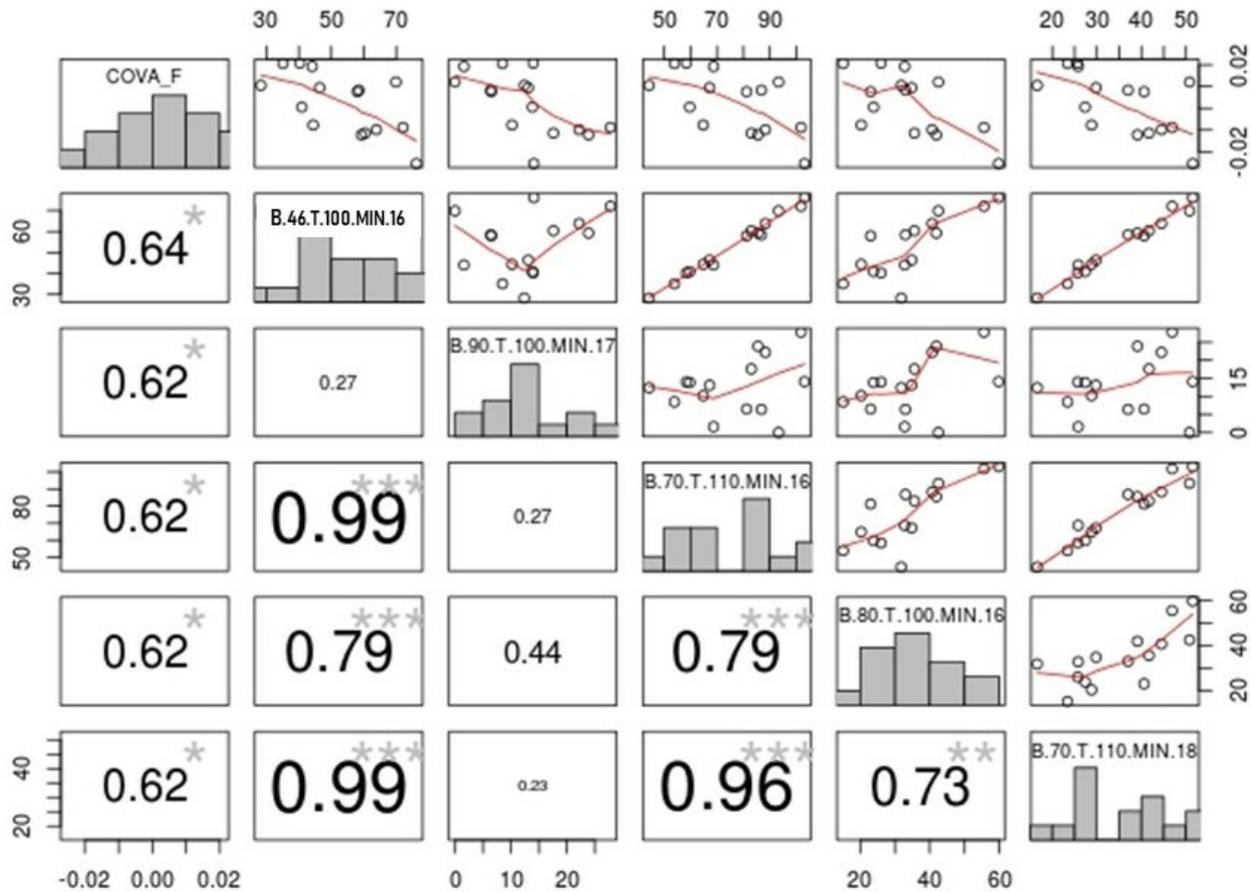


Figura 6.3: $b_{LONGI,3,SITIO_i}$ y $b_{TRANS,3,SITIO_i}$ de MODELO 20 vs ITAC propuestos en Etapas Tempranas del Ciclo de Crecimiento

- B.122.T.150.MIN.22 para β_3 en el Mes de Diciembre, obtenida por la acumulación diaria de temperaturas que superan los 22 grados centígrados desde los 122 DDA hasta los 150 DDA. Si lo expresamos en edad fisiológica del fruto, el criterio se corresponde a una acumulación de temperaturas diarias que superan los 22 Grados centígrados desde 64 hasta 92 DDPF.
- B.60.T.100.MIN.17 para β_2 durante las Etapas Tempranas del Ciclo de Crecimiento, obtenida por la acumulación diaria de temperaturas que superan los 17 grados centígrados desde los 60 DDA hasta los 100 DDA. Si lo expresamos en edad fisiológica del fruto, el criterio se corresponde a una acumulación de temperaturas diarias que

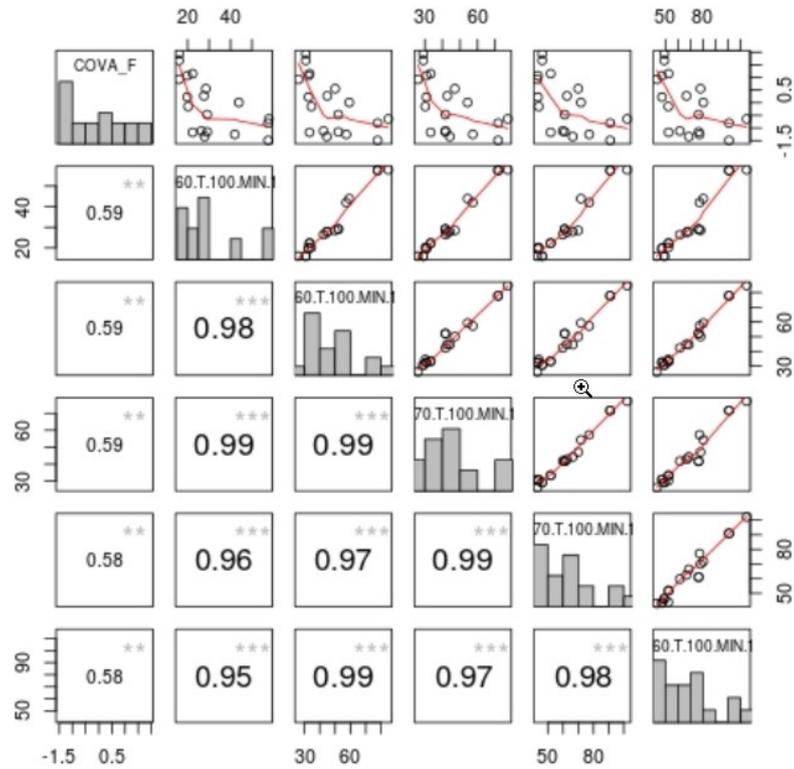


Figura 6.4: $b_{2,SITIO_j}$ de MODELO 21 vs ITAC propuestos en Etapas Tempranas del Ciclo de Crecimiento

superan los 17 Grados centígrados desde 2 hasta 42 DDPF.

A partir de las cuales se proponen los siguientes modelos, con las medidas TRANS-VERSALES como Baseline:

- MODELO 22: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$ y la covariable ambiental B.46.T.100.MIN.16 para β_3 ($a_{3_{TEMP}}$). Efectos aleatorios a nivel de SITIO: $b_{LONGI,3,SITIO_j}$, $b_{TRANS,3,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.
- MODELO 23: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$ y la covariable ambiental B.46.T.100.MIN.16 para β_3 ($a_{3_{TEMP}}$). Efectos aleatorios a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.
- MODELO 24: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$ y la covariable am-

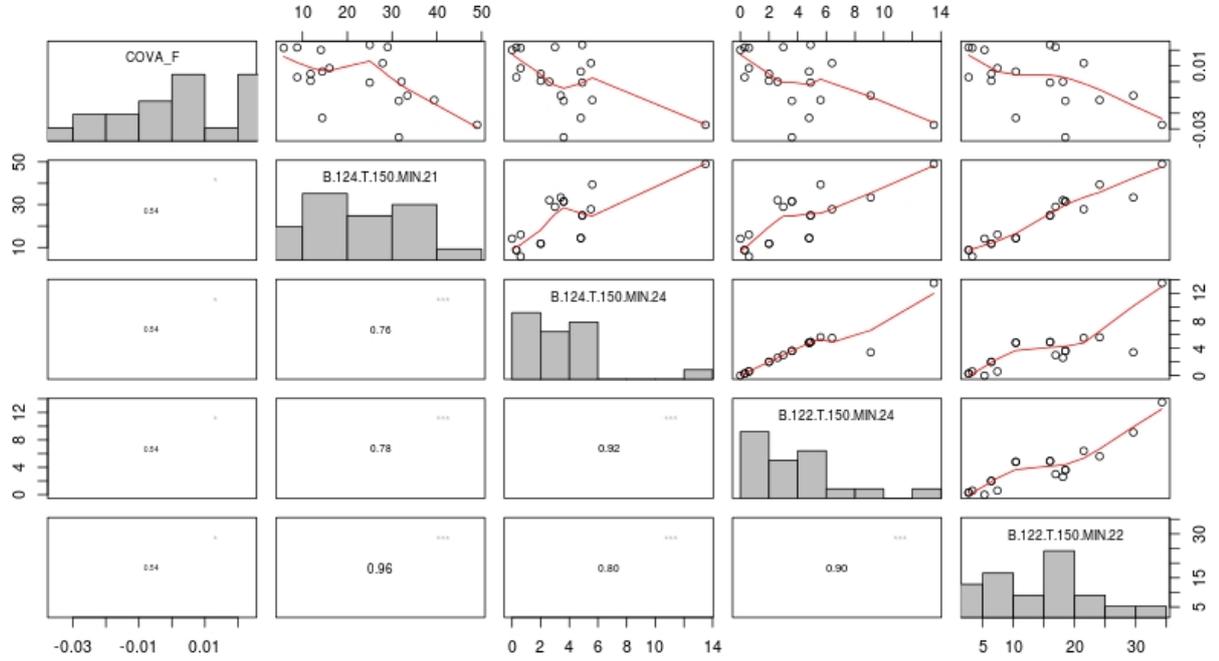


Figura 6.5: $b_{LONGI,3,SITIO_j}$, $b_{TRANS,3,SITIO_j}$ de MODELO 20 vs ITAC propuestos para el transcurso del mes de Diciembre

biental B.122.T.150.MIN.22 para β_3 ($a_{3_{TEMP}}$). Efectos aleatorios a nivel de SITIO:

$b_{LONGI,3,SITIO_j}$, $b_{TRANS,3,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.

- MODELO 25: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$ y la covariable ambiental B.122.T.150.MIN.22 para β_3 ($a_{3_{TEMP}}$). Efectos aleatorios a nivel de SITIO:

$b_{2,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.

- MODELO 26: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$ y la covariable ambiental B.122.T.150.MIN.22 para β_3 ($a_{3_{TEMP}}$) y B.60.T.100.MIN.17 para β_2 ($a_{2_{TEMP}}$).

Efectos aleatorios a nivel de SITIO: $b_{LONGI,3,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.

- MODELO 27: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$ y la covariable ambiental

biental B.122.T.150.MIN.22 para β_3 ($a_{3_{TEMP}}$) y B.60.T.100.MIN.17 para β_2 ($a_{2_{TEMP}}$).
Efectos aleatorios a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$,
 $b_{TRANS,3,FRUTO_{m(j)}}$.

- MODELO 28: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$ y la covariable ambiental B.122.T.150.MIN.22 para β_3 ($a_{3_{TEMP}}$) y B.60.T.100.MIN.17 para β_2 ($a_{2_{TEMP}}$).
. Efectos aleatorios a nivel de SITIO: $b_{LONGI,3,SITIO_j}$, $b_{TRANS,3,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.
- MODELO 29: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$ y la covariable ambiental B.122.T.150.MIN.22 para β_3 ($a_{3_{TEMP}}$) y B.60.T.100.MIN.17 para β_2 ($a_{2_{TEMP}}$).
. Efectos aleatorios a nivel de SITIO: $b_{LONGI,3,SITIO_j}$, $b_{TRANS,3,SITIO_j}$, $b_{2,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.
- MODELO 30: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$ y la covariable ambiental B.122.T.150.MIN.22 para β_3 ($a_{3_{TEMP}}$) y B.60.T.100.MIN.17 para β_2 ($a_{2_{TEMP}}$).
. Efectos aleatorios a nivel de SITIO: $b_{2,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.
- MODELO 31: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$ y la covariable ambiental B.122.T.150.MIN.22 para β_3 ($a_{3_{TEMP}}$) y B.60.T.100.MIN.17 para β_2 ($a_{2_{TEMP}}$).
. Efectos aleatorios a nivel de SITIO: $b_{LONGI,3,SITIO_j}$, $b_{2,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.
- MODELO 32: Considerando como efectos Fijos: $\beta_{1_{TIPO_r}}$, $\beta_{3_{TIPO_r}}$ y la covariable ambiental B.122.T.150.MIN.22 para β_3 ($a_{3_{TEMP}}$) y B.60.T.100.MIN.17 para β_2 ($a_{2_{TEMP}}$).
Efectos aleatorios a nivel de SITIO: $b_{LONGI,3,SITIO_j}$, $b_{2,SITIO_j}$ y a nivel de FRUTO: $b_{2,FRUTO_{m(j)}}$, $b_{LONGI,1,FRUTO_{m(j)}}$, $b_{LONGI,3,FRUTO_{m(j)}}$, $b_{TRANS,3,FRUTO_{m(j)}}$.

La evaluación de performance predictiva para los modelos estimados que incorporan las ITAC se realiza a partir de los CP en el Cuadro 6.13. Las verificaciones gráficas de

supuestos se obtienen en la Figura B.5 del Anexo. Los MODELOS 22,24,25,26,27,30 y 32 se presentan con los mejores CP, en solo uno de ellos las temperaturas tempranas actúan sobre β_3 (MODELO 22), en dos de ellos las temperaturas tardías del mes de diciembre lo hacen sobre β_3 (MODELO 24 y 25), mientras que en los restantes modelos existe una acción simultanea de las temperaturas tempranas y tardías sobre β_2 y β_3 respectivamente.

En general la inclusión de variables que expresan fluctuaciones temporales o espaciales locales incrementa la precisión y reduce el sesgo de las predicciones en los NLME, lo obtenido en esta tesis concuerda con los hallazgos obtenidos en distintas áreas de estudio. En Forestaría se puede citar el trabajo de Crecente-Campo et al. [32] en *Eucaliptos*, Calama et al. [96] en pino piñonero, Newton et al.[92] en abeto negro. En farmacocinética y farmacodinámica, el estudio de covariables de NLME es un tópico de fundamental importancia tal como lo demuestran extensos estudios realizados al respecto con el objetivo último de ampliar la población objeto de estudio sin resignar capacidad predictiva [49],[54], [50], [95], [52], [123], [84]. En estudios epidemiológicos, podemos citar a Distiller et al. [33] sobre el parásito de la malaria y en producción animal a Furtado et al. [43] en crianza de cerdos.

MODELO	R^2 Cond	CC_1	CC_2	RMSE	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD
20	0.9979	0.9999	0.9999	0.6939	-0.0077	-0.1359	0.4241	1.146	0.694
22	0.9980	0.9990	0.9990	0.6852	-0.0109	-0.1579	0.4161	1.141	0.685
23	0.9978	0.9988	0.9988	0.7254	-0.0134	-0.1690	0.4432	1.207	0.725
24	0.9980	0.9990	0.9990	0.6860	-0.0108	-0.1559	0.4168	1.141	0.686
25	0.9980	0.9990	0.9990	0.6818	-0.0096	-0.1474	0.4142	1.129	0.681
26	0.9980	0.9990	0.9990	0.6840	-0.0102	-0.1511	0.4155	1.135	0.684
27	0.9980	0.9990	0.9990	0.6787	-0.0092	-0.1425	0.4123	1.123	0.698
28	0.9980	0.9990	0.9990	0.6814	-0.0097	-0.1465	0.4141	1.128	0.6813
29	0.9980	0.9990	0.9990	0.6851	-0.0099	-0.1505	0.4164	1.135	0.6850
30	0.9980	0.9990	0.9990	0.6810	-0.0091	-0.1433	0.4137	1.126	0.6810
31	0.9980	0.9990	0.9990	0.6865	-0.0101	-0.152	0.4172	1.138	0.6865
32	0.9980	0.9990	0.9990	0.6792	-0.0091	-0.1427	0.4126	1.123	0.679

Cuadro 6.12: Capacidad Predictiva en Modelos con Información Combinada e ITAC

PARÁMETROS	ESTIMACIÓN MODELO						
	22	24	25	26	27	30	32
$\beta_{1INTERCEPT}$	-5.501 (**)	-5.508 (**)	-5.858 (**)	-5.575 (**)	-5.553 (**)	-5.534 (**)	-5.558 (**)
$\beta_{1LONGITUDINAL}$	6.671 (**)	6.678 (**)	7.030 (**)	6.748 (**)	6.729 (**)	0.565 (**)	6.731 (**)
$\beta_{2INTERCEPT}$	8.752 (**)	8.757 (**)	8.875 (**)	9.305 (**)	9.372 (**)	9.401 (**)	9.418 (**)
$\beta_{3INTERCEPT}$	10.001 (**)	9.982 (**)	9.984 (**)	9.982 (**)	9.982 (**)	9.989 (**)	9.984 (**)
$\beta_{3LONGITUDINAL}$	-0.242 (**)	-0.245 (**)	-0.248 (**)	-0.246 (**)	-0.246 (**)	-0.247 (**)	-0.246 (**)
a_{2TEMP}				-0.00168 (3) (**)	-0.01729 (3) (**)	-0.01845 (3) (NS)	-0.01934 (NS)
a_{3TEMP}	-0.00044 (1) (**)	-0.00004 (2) (NS)	-0.00058 (2) (NS)	-0.00001 (2) (NS)	-0.00006 (2) (NS)	-0.00616 (2) (**)	-0.00018 (2) (NS)
$\sigma_{b_{2,SITIO}_j}$	0.033	0.002	8.263	.	.	0.531	0.524
$\sigma_{b_{TRANS,3,SITIO}_j}$	0.033	0.002	.	0.040	.	.	.
$\sigma_{b_{LONGI,3,SITIO}_j}$	0.021	0.495	.	.	0.041	.	0.001
$\sigma_{b_{LONGI,1,FRUTO}_{m(j)}}$	0.051	0.035	0.042	0.038	2.954	0.039	0.038
$\sigma_{b_{2,FRUTO}_{m(j)}}$	1.672	1.805	1.646	2.113	.	2.052	2.087
$\sigma_{b_{LONGI,3,FRUTO}_{m(j)}}$	0.001	0.045	0.297	0.001	0.002	0.001	0.001
$\sigma_{b_{TRANS,3,FRUTO}_{m(j)}}$	0.079	0.007	0.021	0.004	0.003	0.004	0.033
σ_{RES}	0.545	0.554	0.529	0.549	0.533	0.535	0.525

$\sigma_{b_{X,1,i}}, \sigma_{b_{X,2,i}}$ y $\sigma_{b_{X,3,i}}$ son efectos aleatorios a nivel de SITIO. $\sigma_{b_{X,1,ij}}, \sigma_{b_{X,2,ij}}$ y $\sigma_{b_{X,3,ij}}$ son efectos aleatorios a nivel de FRUTO
 (1) Covariable Ambiental B.46.T.100.MIN.16, (2) Covariable Ambiental B.122.T.150.MIN.22, (3) Covariable Ambiental B.60.T.100.MIN.17
 (NS): No Significativo, (**): Altamente Significativo (p<0.01). Test t basado en el estadístico de Wald con distribución asintótica normal

Cuadro 6.13: Parámetros Estimados en Modelos con Información Combinada e ITAC estimados por REML

6.2

CALIBRACIÓN

Una vez concluida la etapa de modelado, es necesario validar los modelos propuestos, lo recomendable aquí es realizarlo a partir de datos no contenidos en los procedimientos de ajuste ya que los resultados obtenidos a partir de los CP tienden a ser demasiados optimistas en relación a los obtenidos con datos externos. La validación externa es un procedimiento necesario a efectos de garantizar métricas predictivas más realistas y es una práctica habitual en farmacocinética, farmacodinámica y forestaría para NLME predictivos [95], [130].

Una primer etapa de la validación externa consiste en someter los modelos propuestos

a una calibración, proceso que consiste básicamente en predecir los nuevos efectos aleatorios (b_i) a partir de datos no contenidos en los procedimientos de ajuste. Los modelos sometidos a calibración no se restringen a los candidatos finales, la evaluación se amplía a modelos que han tenido un adecuado comportamiento en alguna de las etapas del proceso de modelado (MODELOS 2,2b,9,15,17,18,20,19,21,22,23,24,25,26,27,28,29,30,31 y 32) ya que es común obtener discrepancias entre las performances obtenidas en las validaciones internas y externas [95].

La función desarrollada en esta tesis para realizar la calibración se denomina BLUP_OF y realiza esta tarea a través de linealizaciones basadas en los enfoques FO y BLUP, denominados ZEP_b y $EBLUP_b$ respectivamente. La información utilizada para realizar el procedimiento se corresponde a las bases de Validación de Datos LONGITUDINALES y TRANSVERSALES detallados en los Cuadros 5.2 y 5.4 de observaciones con, por lo menos, un mes de anticipación a la fecha de recolección. En los Cuadros 6.14,6.15 se detallan las combinaciones consideradas para la Base de Validación, a partir de las cuales se estiman los b_i para los efectos aleatorios y niveles propios de cada modelo candidato ajustado en base a sus parámetros estimados.

6.3

PREDICCIONES

A partir de los modelos candidatos calibrados se procede a predecir futuras observaciones para cada uno de los frutos considerados en las bases de Validación de Datos LONGITUDINALES y TRANSVERSALES. La función PREDICHO desarrollada en la presente tesis para tal fin, realiza esta tarea a través de linealizaciones obtenidas a partir de los enfoques FO y BLUP, denominadas PL_y y PNL_y respectivamente.

Una vez obtenidos los valores predichos, se procede a evaluar la performance predictiva de los modelos candidatos a través de las métricas de validación externa propuestas.

Sitio	Condición	Nro de Mediciones	DDPF Considerados
TRES ASES-A09-10	1	1	61
TRES ASES-A09-10	2	1	68
TRES ASES-A09-10	3	1	75
TRES ASES-A09-10	4	2	61-68
TRES ASES-A09-10	5	2	61-75
TRES ASES-A09-10	6	2	68-75
TRES ASES-A09-10	7	3	61-68-75
KLEPPE-A13-14	1	1	60
KLEPPE-A13-14	2	1	68
KLEPPE-A13-14	3	1	76
KLEPPE-A13-14	4	2	60-76
KLEPPE-A13-14	5	2	68-76
KLEPPE-A13-14	6	2	60-68-76
KLEPPE-A14-15	1	1	62
KLEPPE-A14-15	2	1	70
KLEPPE-A14-15	3	1	79
KLEPPE-A14-15	4	2	62-70
KLEPPE-A14-15	5	2	62-79
KLEPPE-A14-15	6	2	70-79
KLEPPE-A14-15	7	3	62-70-79

Cuadro 6.14: Combinaciones Consideradas para Base de Validación LONGITUDINALES

Temporada	Condición	Nro de Mediciones	DDPF Considerados
BLANQ-13-14	1	1	74
MIRASOL-13-14	1	1	75

Cuadro 6.15: Combinaciones Consideradas para Base de Validación TRANSVERSALES

Idealmente un mismo enfoque en la linealización debe utilizarse para ajustar un modelo NLME, para predecir parámetros aleatorios y para realizar predicciones sujeto específicas [82].

Es importante, además, poder considerar la extensión y la forma en que las combinaciones de los métodos de expansión afectan la habilidad predictiva de los modelos. Estos descalces pueden ocurrir entre la estimación y la predicción de los b_i , o entre la predicción de los b_i y la predicción de $Y_{i,2}$. Esta investigación tiene sentido, no sólo para modelos que utilizan las expansiones de Taylor en la etapa de modelado (el procedimiento NLME de R en nuestro caso), sino también para modelos estimados por otros enfoques pero que también necesitan predecir b_i y $Y_{i,2}$. Este último caso es un escenario también presente en esta tesis, ya que los estimaciones obtenidas a través del al paquete LMER

4.1-7 del software estadístico R, son obtenidas a partir de métodos de aproximación a la integral.

Por todo lo enunciado precedentemente, se decide evaluar las métricas de validación para la base de validación de los datos LONGITUDINALES para combinaciones de las predicción de los b_i surgidas de los métodos ZEP_b y $EBLUP_b$ y predicciones de los $Y_{i,2}$ procedentes de los métodos PL_y y PNL_y . Los resultados obtenidos se transcriben en el Cuadro A.1 del Anexo.

Allí se aprecia que la combinación $EBLUP_b$ - PNL_y surge como la que mejor performance logra tanto en los modelos estimados a partir de linealizaciones como aquellos estimados a partir de aproximaciones a la integral.

6.3.1

BASE DE VALIDACIÓN DE DATOS LONGITUDINALES

Una vez seleccionada la combinación de enfoques con mayor capacidad predictiva, se torna necesario evaluar la performance de los modelos candidatos bajo la situación de una observación por fruto, el común de las situaciones en el programa de Pronóstico de Producción .

Los resultados obtenidos se detallan ordenados en orden decreciente de Precisión Global de los Errores (δ) resumidos por modelo en el Cuadro 6.16 para los Datos LONGITUDINALES, mientras que una desagregación por Condición se presenta en el Anexo a partir de los Cuadros A.2,A.3,A.4 junto a Histogramas comparativos de las distribuciones de diámetros observada y esperada para dos de los modelos con menores valores de δ (Figura A.1).

Es así como los MODELOS 32, 21, 25, 30, 27, 26, 24 y 17 son los que de acuerdo a δ obtienen la mayor performance predictiva. Llama la atención el hecho que la totalidad de los modelos consignados anteriormente se han obtenido a partir de la información

Modelo	Diám.	Diám. Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	P_{50} $ \bar{e}\% $	P_{75} $ \bar{e}\% $	P_{90} $ \bar{e}\% $	SD	δ
32	64.33	63.55	0.78	1.12	1.91	2.95	2.44	3.97	6.32	2.46	6.68
21	64.40	63.59	0.81	1.15	1.95	3.02	2.41	4.05	6.47	2.47	6.75
25	63.89	63.60	0.29	0.40	2.18	3.43	3.12	4.59	6.65	2.67	7.20
30	63.89	63.33	0.56	0.78	1.99	3.15	2.37	4.67	7.32	2.65	7.34
27	64.42	64.29	0.13	0.14	2.19	3.40	3.02	4.45	6.24	2.72	7.40
26	64.13	64.00	0.13	0.16	2.15	3.37	2.75	4.49	7.12	2.78	7.74
24	64.31	64.02	0.29	0.37	2.27	3.56	3.22	4.70	6.63	2.77	7.78
17	64.00	63.66	0.35	0.42	2.35	3.68	3.22	4.95	7.05	2.87	8.34
20	64.41	63.31	1.09	1.59	2.35	3.62	3.27	5.18	6.75	2.69	8.46
22	64.32	64.40	-0.08	-0.11	2.40	3.75	3.40	5.09	6.85	2.91	8.50
23	63.85	63.53	0.32	0.42	2.37	3.74	3.25	5.17	7.24	2.93	8.67
19	63.96	63.11	0.85	1.20	2.50	3.92	3.61	5.52	7.27	3.01	9.79
2	64.10	63.78	0.32	0.44	2.66	4.13	3.79	5.90	8.53	3.30	11.02
29	64.23	64.75	-0.52	-0.77	3.41	5.36	4.82	7.03	10.32	4.14	17.44
18	63.76	63.49	0.26	0.57	3.31	5.21	4.49	7.29	10.64	4.18	17.52
15	63.92	63.21	0.71	1.07	3.32	5.18	4.19	7.49	10.71	4.15	17.69
28	64.26	64.11	0.15	0.31	3.54	5.58	4.50	8.06	12.17	4.43	19.68
31	64.07	67.07	-3.00	-4.62	3.88	6.02	5.11	9.03	12.40	3.69	22.63
9	63.66	62.86	0.81	0.73	5.29	8.49	7.84	11.75	16.00	6.20	39.11
2b	63.82	66.66	-2.84	-4.91	8.52	13.46	12.84	16.24	19.12	8.67	83.26

Cuadro 6.16: Capacidad Predictiva para una Obs. en Base de Validación LONGITUDINALES

mancomunada de las mediciones LONGITUDINALES y TRANSVERSALES.

Esto confirma el aporte valioso que realizan las mediciones TRANSVERSALES en la estimación de la verdadera variabilidad de los b_i a la luz de la capacidad predictiva final para conjuntos de datos independientes obtenidos a partir de un esquema LONGITUDINAL en concordancia con los resultados obtenidos por Jerez et al.[97] y Macchiavelli [105] para mediciones de parcelas permanentes (experimentales de aclareo rendimiento) y parcelas temporales (de inventario con muestreo sistemático) de Tecas en Venezuela.

Sólo dos de los modelos propuestos no incorporan las ITAC: 17 y 21. Ambos contemplan efectos aleatorios a nivel de Sitio: en β_3 para el MODELO 17 y en β_2 para el MODELO 21.

En cuanto a los modelos que incorporan las ITAC tenemos aquellos que:

- Las incluyen para eventos incurridos en etapas tardías del crecimiento para β_3 :

MODELOS 24 y 25.

- Las incluyen para eventos incurridos en etapas tardías del crecimiento para β_3 y para etapas tempranas del crecimiento en β_2 : MODELOS 26,27,30 y 32.

La totalidad de los modelos que consideran ITAC se nutren de información posterior al momento de medición de diámetros. El momento de medición suele ocurrir durante la primer quincena de diciembre, mientras que algunas ITAC contempladas en los modelos propuestos sólo pueden obtenerse una vez concluida la segunda quincena de diciembre. Por ello, si la predicción se realiza durante la segunda quincena de diciembre sólo pueden utilizarse los modelos 17 y 21, mientras que si la predicción se realiza los primeros días de enero es posible utilizar cualquiera de los modelos propuestos.

Cual sería la ganancia de incorporar las covariables ambientales en los modelos candidatos?. La respuesta la podremos encontrar a partir la información suministrada por los Cuadros A.2,A.3,A.4. A partir de ellos podemos concluir que las performances predictivas de los MODELOS 17 y 21 dependen en gran manera del momento en el cual se efectiviza la predicción. No ocurre lo mismo con los modelos que incorporan las covariables ambientales, los cuales presentan altas performances predictivas para la totalidad de las situaciones contempladas.

El comportamiento adecuado de los modelos que contemplan ITAC puede explicarse a partir de la acción propia que la temperatura tendría sobre los procesos fisiológicos involucrados en la expansión celular a lo largo del mes de diciembre y, por ende, en el tamaño final obtenido por los frutos, tal lo expresado en estudios previos realizados por Tassile et al. [124] y Giménez en su tesis de maestría [34].

Por otra parte, como la estimación de los efectos b_i se nutre de información proveniente de la temporada en curso para actualizar la esperanza condicional, es importante

evaluar la ganancia en capacidad predictiva obtenida como consecuencia de incorporar observaciones para un mismo fruto. Una forma de lograrlo es a partir de los δ para un número variable de observaciones previas por fruto. Los resultados se aprecian en el Cuadro 6.17, el cual expresa los δ promedio obtenidos como función del número de observaciones previas para cada una de las combinaciones consideradas.

	Obs.	Modelo	Diám.	Diám. Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
1	1	17	63.43	63.13	0.30	0.37	2.08	3.26	2.51	6.41
2	2	17	63.58	62.75	0.83	1.18	2.00	3.13	2.54	7.12
3	3	17	62.60	62.58	0.02	-0.10	2.07	3.32	2.86	8.19
4	3	21	62.11	62.72	-0.61	-1.05	1.38	2.25	1.81	3.66
5	1	21	63.55	63.07	0.48	0.64	1.85	2.91	2.44	6.18
6	2	21	64.01	62.85	1.16	1.63	2.22	3.40	2.64	8.31
7	3	24	63.56	63.78	-0.22	-0.49	1.40	2.27	1.99	4.00
8	1	24	64.34	64.09	0.24	0.33	2.25	3.49	2.72	7.45
9	2	24	63.73	62.60	1.14	1.66	2.26	3.56	2.52	7.65
10	3	25	65.02	63.83	1.19	1.75	1.59	2.35	2.05	5.60
11	1	25	64.62	64.23	0.39	0.50	2.35	3.69	2.92	8.69
12	2	25	63.91	63.01	0.90	1.30	2.12	3.33	2.94	9.43
13	1	26	64.13	63.66	0.47	0.66	1.83	2.88	2.24	5.23
14	2	26	64.47	63.84	0.63	0.91	1.79	2.77	2.25	5.47
15	3	26	63.80	63.94	-0.15	-0.30	1.45	2.29	2.36	5.61
16	3	27	61.94	61.86	0.08	0.13	1.43	2.29	1.85	3.42
17	2	27	64.32	63.78	0.54	0.79	1.97	3.07	2.84	8.37
18	1	27	63.82	64.27	-0.45	-0.74	2.37	3.73	2.94	8.82
19	1	30	64.62	64.05	0.57	0.79	1.76	2.77	2.25	5.38
20	2	30	64.63	63.26	1.36	2.00	2.15	3.30	2.26	6.98
21	3	30	63.02	63.60	-0.58	-0.94	1.71	2.71	3.29	11.18
22	3	32	63.36	63.54	-0.18	-0.29	1.17	1.84	1.54	2.40
23	2	32	64.73	63.54	1.19	1.77	1.78	2.71	2.08	5.74
24	1	32	63.86	62.93	0.93	1.35	2.08	3.23	2.55	7.39

Cuadro 6.17: Capacidad Predictiva en Base de Validación TRANSVERSALES según nro. de Obs.

Del Cuadro 6.17 se desprende que los MODELOS 21, 24, 25, 27 y 32 obtienen las mejores performances predictivas a partir de tres observaciones por fruto, con una caída de gran magnitud en valores de δ cuando se los compara con sus respectivas calibraciones resultantes con una o dos observaciones.

Por otra parte, los MODELOS 17, 26 y 30 logran una mejor performance predicti-

va cuando sus b_i se calibran a partir de una observación, pero con valores de δ que no presentan grandes mejoras en relación a los modelos calibrados a partir de dos observaciones por fruto.

En función de lo expresado anteriormente, podemos concluir que una ganancia en capacidad predictiva se logra en la mayoría de los modelos a partir de su calibración con tres observaciones consecutivas. Por otra parte, no se aprecian diferencias prácticas entre una o dos observaciones consecutivas por fruto para el proceso de calibración.

6.3.2

BASE DE VALIDACIÓN DE DATOS TRANSVERSALES

Los resultados obtenidos para la Base de Validación de Datos TRANSVERSALES se presentan en el Cuadro 6.18. A las usuales métricas predictivas propuestas se incorporan las correspondientes a la Mediana debido a la presencia de distribuciones alejadas de la simetría, tal como se aprecia en la Figura A.2 del Anexo.

A partir de estos resultados es posible concluir que los menores Errores Medianos y de Desvío se obtienen a partir de los MODELOS 21, 22, 24 y 27.

6.3.3

ESPERANZAS Y CORRELACIONES MARGINALES INDUCIDAS

Los efectos fijos estimados como parte de un NLME no caracterizan totalmente la respuesta media de la población, especialmente cuando la variabilidad inter-sujeto es grande [74], por ello se torna necesario estimar las esperanzas marginales inducidas a partir de los términos del modelo. Una propuesta para su cálculo es posible a partir de simulaciones de los b_i que posibilitan generar curvas medianas y curvas percentiles 05 y 95.

Además, en los modelos finales propuestos se asume que los errores son independien-

MODELO	Media			Mediana			Desvio		CV	
	Pred.	Error	Error Rel. %	Pred	Error	Error Rel. %	Pred.	Error	Pred.	Error
19	65.65	1.95	2.89	65.78	0.21	0.33	9.69	-3.15	14.77	-5.09
21	65.28	2.32	3.44	65.04	0.95	1.45	7.00	-0.45	10.73	-1.04
24	67.64	-0.03	-0.04	67.36	-1.36	-2.07	7.62	-1.08	11.27	-1.59
22	67.66	-0.04	-0.07	67.39	-1.39	-2.11	7.61	-1.07	11.25	-1.57
18	67.57	0.03	0.05	67.40	-1.40	-2.13	10.24	-3.69	15.15	-5.47
27	67.80	-0.19	-0.28	67.48	-1.48	-2.24	7.64	-1.10	11.27	-1.59
17	68.37	-0.76	-1.12	68.14	-2.14	-3.25	7.80	-1.26	11.41	-1.73
25	63.64	3.97	5.87	63.58	2.41	3.66	6.61	-0.071	10.39	-0.71
23	71.14	-3.53	-5.22	71.10	-5.10	-7.73	8.55	-2.00	12.02	-2.34
20	70.93	-3.32	-4.91	71.27	-5.27	-7.98	8.11	-1.56	11.43	-1.75
32	60.03	7.57	11.20	60.00	5.99	9.08	8.12	-1.57	13.53	-3.84
26	59.99	7.61	11.26	59.96	6.03	9.14	8.11	-1.56	13.52	-3.84
31	76.19	-8.58	-12.69	75.97	-9.97	-15.10	9.21	-2.66	12.09	-2.41
15	86.97	-19.36	-28.63	89.82	-23.82	-36.10	5.29	1.24	6.09	3.59
30	116.16	-48.55	-71.81	114.38	-48.38	-73.305	27.26	-20.71	23.46	-13.78
29	-5.05	72.66	107.47	-0.23	66.23	100.355	4.83	1.71	-95.65	105.34
28	-7.68	75.30	111.36	-11.33	77.33	117.175	3.69	2.84	-48.09	57.77

Cuadro 6.18: Capacidad Predictiva en Base de Validación TRANSVERSALES

tes normales y aleatorios con varianza constante, lo cual implica que las respuestas son condicionalmente independientes, dados los b_i . Sin embargo, marginalmente, las observaciones que comparten un b_i se encuentran correlacionadas lo cual implica una estructura de correlación de que no posee una forma cerrada [20]. Por esa razón, la aproximación propuesta es un ejemplo de modelado indirecto de la estructura de varianza-covarianza [74]. Una estrategia alternativa para determinar dicha estructura, consiste en obtener, por medio de simulaciones, distribuciones marginales bivariadas de dos observaciones del mismo FRUTO en distintos tiempos. A partir de estas distribuciones marginales, es posible calcular las correlaciones entre observaciones del mismo FRUTO [73], [97]. Estas correlaciones se presentan para los modelos que tienen un adecuado comportamiento en las bases de validación de datos longitudinales y transversales: MODELOS 21, 24 y 27 (Figuras del Anexo C.1, C.2, C.3). En ellas se ven las importantes magnitudes alcanzadas por la correlación marginal junto a la dependencia de esta con respecto a los modelos elegidos y los momentos de ciclo de crecimiento contemplados.

Capítulo 7

CONCLUSIONES

Una primer cuestión a considerar reside en la escasa performance predictiva de los modelos que consideran la correlación serial de las medidas repetidas a partir de la modelación directa en la matriz de varianza-covarianza de los errores para modelos NLME. Si bien los CI obtenidos perfilan a estos como los modelos candidatos finales, los CP alcanzados por estos no indican lo mismo. Por el contrario, los modelos con estructura del error sin correlación y/o heterocedasticidad logran consistentemente mejores métricas predictivas a partir de los CP en todos los escenarios contemplados, un resultado también hallado en estudios de forestería [107], [106].

Cuando un NLME es utilizado para realizar predicciones de frutos no disponibles en el proceso de modelado, una o más observaciones de los FRUTOS deben estar disponibles en el proceso de calibración. Deben abordarse, además, cuestiones en relación a los enfoques implementados para la calibración de los efectos aleatorios y las predicciones de observaciones futuras de estos nuevos FRUTOS.

Cuando un subconjunto de observaciones por FRUTO es utilizada para calibrar los modelos, las predicciones resultantes son más precisas y exactas a medida que más observaciones se utilizan [129]. En esta tesis esto es verdad a partir de tres observaciones

consecutivas tomadas una vez por semana para el proceso de calibración, no apreciándose diferencias prácticas entre una o dos observaciones.

Existe un consenso respecto a que los mejores CP en modelos estimados en la modelación a partir de FO, y EBLUP, se obtienen a partir del uso de un mismo enfoque para la calibración de los efectos aleatorios (b_i) y las predicciones sujeto específicas [82] [129] [130]. Cuando el ajuste del modelo se realiza a partir de enfoques que no implican la obtención de pseudo-datos, Chengcai et al. [82] afirman que la calibración de los b_i y la predicciones de futuras observaciones pueden obtenerse de manera indistinta a partir de métodos de expansión FO o EBLUP [82].

Esta afirmación no concuerda con los resultados obtenidos en la presente tesis, donde los mejores CP se logran a través de la dupla $EBLUP_b-PNL_y$ para la totalidad de los enfoques incurridos para la estimación en la etapa de modelado (EBLUP y aproximación de Laplace).

Los NLME multinivel tienen una ventaja al discernir y cuantificar fuentes de variabilidad estocástica a distintos niveles, esto se traduce en mejores CI y CP en relación a modelos que consideran solo un nivel de variabilidad en sus efectos aleatorios. En esta tesis, los MODELOS candidatos obtenidos tanto a partir de la Base de Validación LONGITUDINALES como de la Base de Validación TRANSVERSALES se corresponden con modelos multinivel que contienen efectos aleatorios a nivel de SITIO y de FRUTO, estos resultados concuerdan con la bibliografía consultada [20] [64] [129] [65] [131] [109]. Es llamativa, además, la capacidad que poseen algunos NLME multinivel en lograr altos CI a pesar de prescindir de covariables ambientales. Esto se debe a la habilidad que tienen los efectos aleatorios a nivel de SITIO en contemplar variabilidad estocástica sin tener que exigir una identificación o medición de la fuente de variación, actuando por ello como una aproximación de factores ambientales no considerados [129].

Por otro lado, generalmente, se toma por dada la existencia de una relación directa entre capacidad predictiva y número de covariables que posee un modelo. Esto es verda-

dero en regresión mínimo cuadrática tradicional y también puede serlo, en predicciones obtenidas a partir de modelos con efectos fijos. Sin embargo, en términos de NLME la incorporación de covariables debe realizarse de manera muy cuidadosa ya que existen diversos motivos potenciales que conllevan un sobreajuste con la consecuente reducción en la capacidad predictiva, existen algunos ejemplos de ello en la bibliografía [109]. En ese sentido, los resultados obtenidos en esta tesis son claros en favor de la incorporación de covariables que caracterizan momentos variables del crecimiento de los frutos, lo que concuerda con numerosa bibliografía [32] [96] [92] [33]. De los MODELOS calibrados a partir de la Base de Validación LONGITUDINALES, seis de los ocho MODELOS candidatos tienen ITAC incorporadas, mientras que para el caso de la Base de Validación TRANSVERSALES, tres de los cuatro MODELOS candidatos poseen covariables ambientales.

Cuando consideramos las implicancias biológicas de las ITAC en el marco de los NLME, es útil evaluar la significancia de estas al momento de la incorporación al modelo. Los resultados obtenidos han permitido corroborar lo obtenido en estudios anteriores sobre el efecto diferencial de la termoacumulación sobre los distintos momentos fisiológicos del fruto. Es así como, en etapas tempranas del crecimiento, altas temperaturas favorecen el crecimiento celular lo que incrementa el tamaño potencial de los frutos, mientras que en momentos finales de ciclo, elevadas temperaturas inciden negativamente sobre las tasas de crecimiento [124] [125] [6] [34].

Un argumento adicional a favor de modelos que contemplan ITAC, surge en la robustez obtenida en cuanto a los CP. Esta es una propiedad de la cual adolecen los modelos que carecen de los ITAC (MODELOS 17 y 21) con performances predictivas que dependen del momento de ciclo de crecimiento en el cual se efectivizan las mediciones utilizadas para calibrar los efectos aleatorios. Si bien el uso de ITAC está limitado a predecir observaciones tardías, este aspecto de los NLME multinivel con covariables ambientales es un

tópico no considerado en la bibliografía consultada.

Apéndice A

MÉTRICAS DE VALIDACIÓN.

RESULTADOS OBTENIDOS

DETALLADOS

Cuadro A.1: Métricas de Validación para Modelos Propuestos

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
1	<i>EBLUP_b</i>	<i>PNL_y</i>	17	64.09	63.35	0.75	1.00	2.17	3.38	2.66	7.64
1	<i>ZEP_b</i>	<i>PNL_y</i>	17	64.09	64.11	-0.02	-0.17	2.38	3.72	2.94	8.64
1	<i>EBLUP_b</i>	<i>PL_y</i>	17	64.09	66.28	-2.19	-3.16	5.70	8.84	6.56	47.88
1	<i>ZEP_b</i>	<i>PL_y</i>	17	64.09	67.85	-3.76	-5.55	6.73	10.30	8.02	78.44
1	<i>EBLUP_b</i>	<i>PNL_y</i>	18	64.09	63.61	0.48	0.92	3.74	5.85	4.64	21.74
1	<i>ZEP_b</i>	<i>PNL_y</i>	18	64.09	63.91	0.18	0.47	4.04	6.24	5.25	27.64
1	<i>EBLUP_b</i>	<i>PL_y</i>	18	64.09	65.35	-1.26	-1.42	7.56	11.79	8.92	81.15
1	<i>ZEP_b</i>	<i>PL_y</i>	18	64.09	65.96	-1.86	-2.32	7.98	12.31	10.08	105.06
1	<i>ZEP_b</i>	<i>PNL_y</i>	19	64.09	62.91	1.19	1.78	2.49	3.90	2.84	9.51
1	<i>EBLUP_b</i>	<i>PNL_y</i>	19	64.09	62.75	1.34	2.02	2.53	3.97	2.82	9.76
1	<i>EBLUP_b</i>	<i>PL_y</i>	19	64.09	65.56	-1.47	-1.96	5.97	9.40	6.84	48.92

Continúa en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
1	ZEP _b	PL _y	19	64.09	65.84	-1.74	-2.40	6.08	9.54	6.93	51.01
1	ZEP _b	PNL _y	20	64.09	63.15	0.94	1.33	2.43	3.80	2.86	9.04
1	EBLUP _b	PNL _y	20	64.09	62.74	1.36	1.99	2.60	4.06	2.87	10.11
1	EBLUP _b	PL _y	20	64.09	65.36	-1.27	-1.70	5.87	9.18	6.70	46.51
1	ZEP _b	PL _y	20	64.09	66.04	-1.95	-2.77	5.90	9.18	6.73	49.04
1	ZEP _b	PNL _y	21	64.09	63.47	0.62	0.84	2.06	3.24	2.63	7.31
1	EBLUP _b	PNL _y	21	64.09	63.04	1.05	1.51	2.10	3.27	2.52	7.43
1	EBLUP _b	PL _y	21	64.09	66.04	-1.94	-2.75	5.38	8.42	6.16	41.74
1	ZEP _b	PL _y	21	64.09	66.82	-2.73	-3.97	5.90	9.18	6.69	52.17
1	EBLUP _b	PNL _y	22	64.09	63.72	0.37	0.58	2.54	3.96	3.11	9.79
1	ZEP _b	PNL _y	22	64.09	64.07	0.02	0.03	2.59	4.04	3.27	10.71
1	EBLUP _b	PL _y	22	64.09	66.56	-2.47	-3.47	5.93	9.24	6.75	51.65
1	ZEP _b	PL _y	22	64.09	67.23	-3.14	-4.50	6.29	9.70	7.27	62.63
1	ZEP _b	PNL _y	23	64.09	63.51	0.58	0.84	2.41	3.79	2.92	8.84
1	EBLUP _b	PNL _y	23	64.09	63.36	0.74	1.10	2.43	3.83	2.92	9.05
1	EBLUP _b	PL _y	23	64.09	65.50	-1.41	-1.85	5.92	9.32	6.78	47.98
1	ZEP _b	PL _y	23	64.09	65.77	-1.68	-2.28	5.98	9.39	6.80	49.01
1	ZEP _b	PNL _y	24	64.09	63.57	0.52	0.75	2.41	3.78	2.92	8.79
1	EBLUP _b	PNL _y	24	64.09	63.41	0.69	1.02	2.41	3.80	2.91	8.93
1	EBLUP _b	PL _y	24	64.09	65.63	-1.53	-2.05	6.06	9.54	6.96	50.76
1	ZEP _b	PL _y	24	64.09	65.91	-1.82	-2.50	6.14	9.62	7.00	52.32
1	ZEP _b	PNL _y	25	64.09	63.67	0.42	0.59	2.36	3.71	2.89	8.54
1	EBLUP _b	PNL _y	25	64.09	63.52	0.58	0.84	2.36	3.73	2.89	8.67
1	EBLUP _b	PL _y	25	64.09	65.75	-1.66	-2.25	5.88	9.25	6.73	48.02
1	ZEP _b	PL _y	25	64.09	66.02	-1.92	-2.68	5.95	9.33	6.76	49.36
1	ZEP _b	PNL _y	26	64.09	63.94	0.15	0.16	2.15	3.37	2.69	7.25
1	EBLUP _b	PNL _y	26	64.09	63.64	0.46	0.65	2.15	3.38	2.65	7.26
1	EBLUP _b	PL _y	26	64.09	66.06	-1.97	-2.74	5.65	8.83	6.41	45.00

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
1	ZEP _b	PL _y	26	64.09	66.62	-2.53	-3.62	5.79	8.99	6.58	49.71
1	EBLUP _b	PNL _y	27	64.09	63.83	0.27	0.35	2.14	3.37	2.67	7.19
1	ZEP _b	PNL _y	27	64.09	64.08	0.01	-0.06	2.20	3.45	2.73	7.45
1	EBLUP _b	PL _y	27	64.09	66.24	-2.15	-3.03	5.64	8.81	6.36	45.01
1	ZEP _b	PL _y	27	64.09	66.79	-2.69	-3.88	5.90	9.16	6.61	50.95
1	EBLUP _b	PNL _y	28	64.09	63.48	0.61	1.05	4.22	6.70	5.18	27.20
1	ZEP _b	PNL _y	28	64.09	63.93	0.17	0.32	4.29	6.74	5.32	28.34
1	EBLUP _b	PL _y	28	64.09	63.32	0.78	1.60	6.55	10.43	8.13	66.63
1	ZEP _b	PL _y	28	64.09	64.05	0.04	0.41	6.67	10.52	8.53	72.69
1	EBLUP _b	PNL _y	29	64.09	64.42	-0.33	-0.44	4.01	6.31	4.87	23.85
1	ZEP _b	PNL _y	29	64.09	63.43	0.67	1.09	4.22	6.61	5.22	27.68
1	EBLUP _b	PL _y	29	64.09	63.86	0.24	0.73	6.49	10.30	8.05	64.79
1	ZEP _b	PL _y	29	64.09	63.17	0.93	1.76	6.83	10.78	8.53	73.64
1	ZEP _b	PNL _y	30	64.09	63.53	0.57	0.73	2.19	3.43	2.90	8.72
1	EBLUP _b	PNL _y	30	64.09	63.21	0.88	1.23	2.23	3.49	2.87	9.00
1	EBLUP _b	PL _y	30	64.09	64.84	-0.75	-0.97	3.64	5.73	4.49	20.69
1	ZEP _b	PL _y	30	64.09	65.58	-1.48	-2.13	3.95	6.13	4.88	26.02
1	EBLUP _b	PNL _y	31	64.09	67.02	-2.93	-4.51	3.81	5.92	3.61	21.59
1	ZEP _b	PNL _y	31	64.09	66.32	-2.22	-3.35	4.48	6.92	4.89	28.88
1	EBLUP _b	PL _y	31	64.09	69.24	-5.14	-7.59	7.60	11.76	7.53	83.10
1	ZEP _b	PL _y	31	64.09	69.22	-5.13	-7.49	8.81	13.57	9.48	116.11
1	ZEP _b	PNL _y	32	64.09	63.53	0.57	0.75	1.96	3.08	2.54	6.78
1	EBLUP _b	PNL _y	32	64.09	62.99	1.10	1.60	2.09	3.27	2.48	7.36
1	EBLUP _b	PL _y	32	64.09	65.26	-1.16	-1.55	4.53	7.11	5.31	29.53
1	ZEP _b	PL _y	32	64.09	66.09	-1.99	-2.85	4.82	7.49	5.59	35.23
1	ZEP _b	PNL _y	2	64.09	65.33	-1.24	-2.04	2.47	3.90	2.96	10.27
1	ZEP _b	PL _y	2	64.09	68.63	-4.53	-6.98	5.59	8.64	5.88	55.19
1	EBLUP _b	PL _y	2	64.09	67.91	-3.82	-6.30	8.21	12.82	8.12	80.44

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

MÉTODO		Diámetro									
Cond. 1	2	MODELO	Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ	
1	<i>EBLUP_b</i>	<i>PNL_y</i>	2	64.09	67.03	-2.94	-5.05	9.17	14.42	9.34	95.79
1	<i>EBLUP_b</i>	<i>PNL_y</i>	15	64.09	62.93	1.16	1.83	3.42	5.34	4.08	18.01
1	<i>ZEP_b</i>	<i>PNL_y</i>	15	64.09	62.77	1.32	2.08	3.49	5.45	4.10	18.56
1	<i>EBLUP_b</i>	<i>PL_y</i>	15	64.09	62.34	1.75	3.07	6.43	10.15	7.21	55.08
1	<i>ZEP_b</i>	<i>PL_y</i>	15	64.09	62.18	1.91	3.33	6.42	10.16	7.22	55.82
1	<i>EBLUP_b</i>	<i>PNL_y</i>	2	64.09	62.79	1.30	1.93	2.71	4.22	3.11	11.36
1	<i>ZEP_b</i>	<i>PNL_y</i>	2	64.09	62.00	2.10	3.13	3.07	4.75	3.13	14.18
1	<i>ZEP_b</i>	<i>PL_y</i>	2	64.09	60.68	3.42	5.47	4.86	7.71	4.64	33.21
1	<i>EBLUP_b</i>	<i>PL_y</i>	2	64.09	61.61	2.48	4.07	5.51	8.70	5.65	38.08
1	<i>EBLUP_b</i>	<i>PL_y</i>	9	64.09	61.44	2.66	4.01	3.40	5.27	3.08	16.53
1	<i>ZEP_b</i>	<i>PL_y</i>	9	64.09	60.84	3.26	4.91	3.84	5.91	3.17	20.65
1	<i>EBLUP_b</i>	<i>PNL_y</i>	9	64.09	62.71	1.38	1.75	5.13	8.08	5.89	36.63
1	<i>ZEP_b</i>	<i>PNL_y</i>	9	64.09	62.21	1.89	2.52	5.44	8.55	6.12	41.04
2	<i>EBLUP_b</i>	<i>PNL_y</i>	17	64.09	64.11	-0.01	-0.14	2.59	4.03	3.11	9.68
2	<i>ZEP_b</i>	<i>PNL_y</i>	17	64.09	64.90	-0.81	-1.35	3.03	4.70	3.57	13.38
2	<i>EBLUP_b</i>	<i>PL_y</i>	17	64.09	67.88	-3.79	-5.55	7.44	11.50	8.51	86.85
2	<i>ZEP_b</i>	<i>PL_y</i>	17	64.09	69.63	-5.53	-8.18	8.69	13.26	10.28	136.23
2	<i>EBLUP_b</i>	<i>PNL_y</i>	18	64.09	64.17	-0.07	0.02	3.88	6.04	4.76	22.64
2	<i>ZEP_b</i>	<i>PNL_y</i>	18	64.09	64.46	-0.37	-0.42	4.12	6.31	5.33	28.60
2	<i>EBLUP_b</i>	<i>PL_y</i>	18	64.09	66.64	-2.55	-3.40	8.81	13.68	10.06	107.68
2	<i>ZEP_b</i>	<i>PL_y</i>	18	64.09	67.32	-3.23	-4.39	9.23	14.16	11.39	140.05
2	<i>ZEP_b</i>	<i>PNL_y</i>	19	64.09	63.31	0.78	1.12	2.54	3.97	3.01	9.66
2	<i>EBLUP_b</i>	<i>PNL_y</i>	19	64.09	63.27	0.82	1.19	2.55	3.99	3.01	9.72
2	<i>ZEP_b</i>	<i>PL_y</i>	19	64.09	66.83	-2.74	-3.93	6.86	10.72	7.70	66.85
2	<i>EBLUP_b</i>	<i>PL_y</i>	19	64.09	66.76	-2.67	-3.81	6.88	10.76	7.75	67.15
2	<i>ZEP_b</i>	<i>PNL_y</i>	20	64.09	63.63	0.47	0.60	2.50	3.91	3.01	9.30
2	<i>EBLUP_b</i>	<i>PNL_y</i>	20	64.09	63.36	0.74	1.04	2.59	4.03	3.04	9.81
2	<i>EBLUP_b</i>	<i>PL_y</i>	20	64.09	66.71	-2.61	-3.73	6.91	10.74	7.84	68.25

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
2	ZEP _b	PL _y	20	64.09	67.18	-3.09	-4.49	6.90	10.69	7.77	69.90
2	EBLUP _b	PNL _y	21	64.09	63.62	0.47	0.61	2.13	3.34	2.70	7.51
2	ZEP _b	PNL _y	21	64.09	63.96	0.13	0.08	2.28	3.57	2.87	8.26
2	EBLUP _b	PL _y	21	64.09	67.30	-3.20	-4.67	6.66	10.36	7.45	65.81
2	ZEP _b	PL _y	21	64.09	67.89	-3.79	-5.58	7.07	10.96	7.88	76.50
2	EBLUP _b	PNL _y	22	64.09	64.26	-0.16	-0.25	2.77	4.31	3.43	11.77
2	ZEP _b	PNL _y	22	64.09	64.55	-0.45	-0.69	2.95	4.56	3.72	14.04
2	EBLUP _b	PL _y	22	64.09	67.85	-3.76	-5.43	7.12	11.03	7.75	74.13
2	ZEP _b	PL _y	22	64.09	68.47	-4.37	-6.36	7.55	11.59	8.47	90.88
2	ZEP _b	PNL _y	23	64.09	63.96	0.14	0.13	2.56	4.00	3.09	9.57
2	EBLUP _b	PNL _y	23	64.09	63.90	0.19	0.23	2.58	4.03	3.10	9.62
2	ZEP _b	PL _y	23	64.09	66.77	-2.68	-3.82	6.80	10.63	7.60	64.89
2	EBLUP _b	PL _y	23	64.09	66.67	-2.57	-3.65	6.82	10.67	7.65	65.17
2	ZEP _b	PNL _y	24	64.09	64.00	0.10	0.07	2.55	3.99	3.09	9.54
2	EBLUP _b	PNL _y	24	64.09	63.93	0.16	0.18	2.56	4.01	3.09	9.56
2	ZEP _b	PL _y	24	64.09	66.88	-2.79	-3.99	6.96	10.87	7.83	69.01
2	EBLUP _b	PL _y	24	64.09	66.77	-2.68	-3.81	6.98	10.91	7.87	69.06
2	ZEP _b	PNL _y	25	64.09	64.06	0.04	-0.03	2.52	3.94	3.06	9.38
2	EBLUP _b	PNL _y	25	64.09	64.00	0.09	0.07	2.53	3.96	3.07	9.41
2	ZEP _b	PL _y	25	64.09	66.92	-2.83	-4.07	6.76	10.55	7.55	64.97
2	EBLUP _b	PL _y	25	64.09	66.83	-2.73	-3.90	6.78	10.59	7.60	65.19
2	EBLUP _b	PNL _y	26	64.09	64.19	-0.09	-0.19	2.41	3.76	2.98	8.91
2	ZEP _b	PNL _y	26	64.09	64.40	-0.31	-0.53	2.44	3.80	3.06	9.49
2	EBLUP _b	PL _y	26	64.09	67.30	-3.20	-4.60	6.88	10.69	7.61	68.13
2	ZEP _b	PL _y	26	64.09	67.74	-3.65	-5.29	7.05	10.88	7.82	74.38
2	EBLUP _b	PNL _y	27	64.09	64.53	-0.44	-0.74	2.63	4.10	3.16	10.18
2	ZEP _b	PNL _y	27	64.09	64.75	-0.66	-1.08	2.72	4.24	3.25	10.98
2	EBLUP _b	PL _y	27	64.09	67.72	-3.62	-5.25	7.13	11.08	7.87	75.04

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

MÉTODO		Diámetro									
Cond. 1	2	MODELO	Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ	
2	ZEP _b	PL _y	27	64.09	68.25	-4.16	-6.09	7.43	11.50	8.18	84.27
2	EBLUP _b	PNL _y	28	64.09	64.38	-0.29	-0.37	3.85	6.07	4.47	20.08
2	ZEP _b	PNL _y	28	64.09	64.73	-0.63	-0.94	3.92	6.14	4.55	21.12
2	EBLUP _b	PL _y	28	64.09	64.57	-0.48	-0.32	6.53	10.33	7.73	59.97
2	ZEP _b	PL _y	28	64.09	65.23	-1.13	-1.39	6.67	10.45	8.07	66.41
2	EBLUP _b	PNL _y	29	64.09	65.02	-0.92	-1.38	3.54	5.57	4.00	16.85
2	ZEP _b	PNL _y	29	64.09	64.30	-0.21	-0.29	3.54	5.57	4.37	19.15
2	EBLUP _b	PL _y	29	64.09	64.75	-0.66	-0.63	6.03	9.55	7.42	55.45
2	ZEP _b	PL _y	29	64.09	64.35	-0.25	-0.05	6.16	9.70	7.92	62.74
2	EBLUP _b	PNL _y	30	64.09	63.70	0.39	0.48	2.32	3.63	3.03	9.33
2	ZEP _b	PNL _y	30	64.09	64.06	0.04	-0.07	2.34	3.67	3.08	9.49
2	EBLUP _b	PL _y	30	64.09	65.99	-1.89	-2.68	4.89	7.61	5.55	34.34
2	ZEP _b	PL _y	30	64.09	66.86	-2.77	-4.05	5.36	8.26	6.04	44.11
2	EBLUP _b	PNL _y	31	64.09	67.59	-3.50	-5.37	4.35	6.74	4.22	30.01
2	ZEP _b	PNL _y	31	64.09	67.39	-3.30	-4.98	5.30	8.14	5.81	44.67
2	EBLUP _b	PL _y	31	64.09	70.61	-6.51	-9.65	9.11	14.05	9.02	123.78
2	ZEP _b	PL _y	31	64.09	71.40	-7.30	-10.75	10.84	16.61	11.66	189.21
2	EBLUP _b	PNL _y	32	64.09	63.65	0.44	0.59	2.15	3.37	2.75	7.74
2	ZEP _b	PNL _y	32	64.09	64.07	0.02	-0.07	2.20	3.44	2.84	8.04
2	EBLUP _b	PL _y	32	64.09	66.61	-2.52	-3.59	5.95	9.26	6.60	49.84
2	ZEP _b	PL _y	32	64.09	67.31	-3.21	-4.68	6.24	9.65	6.90	57.99
2	ZEP _b	PNL _y	2	64.09	63.79	0.30	0.34	2.13	3.34	2.82	8.05
2	ZEP _b	PL _y	2	64.09	66.64	-2.55	-3.90	4.59	7.12	5.23	33.82
2	EBLUP _b	PL _y	2	64.09	67.23	-3.14	-5.21	8.35	13.00	8.67	84.96
2	EBLUP _b	PNL _y	2	64.09	67.14	-3.04	-5.18	8.68	13.66	8.77	86.19
2	EBLUP _b	PNL _y	15	64.09	63.90	0.19	0.31	3.37	5.23	4.27	18.29
2	ZEP _b	PNL _y	15	64.09	63.42	0.68	1.09	3.47	5.40	4.29	18.86
2	EBLUP _b	PL _y	15	64.09	63.73	0.36	0.94	7.22	11.33	8.27	68.49

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
2	ZEP _b	PL _y	15	64.09	63.11	0.99	1.95	7.14	11.26	8.22	68.56
2	ZEP _b	PNL _y	2	64.09	62.24	1.85	2.76	2.95	4.56	3.08	12.94
2	EBLUP _b	PNL _y	2	64.09	64.35	-0.26	-0.46	2.86	4.44	3.63	13.25
2	ZEP _b	PL _y	2	64.09	61.02	3.07	5.01	5.44	8.65	5.57	40.52
2	EBLUP _b	PL _y	2	64.09	63.77	0.32	0.81	7.23	11.33	7.99	63.99
2	EBLUP _b	PL _y	9	64.09	62.35	1.74	2.62	3.09	4.79	3.33	14.12
2	ZEP _b	PL _y	9	64.09	60.69	3.40	5.16	3.90	6.00	3.04	20.81
2	EBLUP _b	PNL _y	9	64.09	63.16	0.93	1.03	5.18	8.20	6.03	37.23
2	ZEP _b	PNL _y	9	64.09	62.00	2.09	2.82	5.66	8.89	6.26	43.58
3	EBLUP _b	PNL _y	17	64.09	63.96	0.13	0.13	1.90	2.96	2.36	5.57
3	ZEP _b	PNL _y	17	64.09	64.60	-0.50	-0.85	2.19	3.39	2.67	7.38
3	EBLUP _b	PL _y	17	64.09	67.54	-3.44	-4.93	6.88	10.68	7.62	69.92
3	ZEP _b	PL _y	17	64.09	68.93	-4.83	-7.05	7.82	11.98	8.88	102.25
3	EBLUP _b	PNL _y	18	64.09	63.79	0.31	0.58	2.71	4.25	3.46	12.08
3	ZEP _b	PNL _y	18	64.09	64.04	0.05	0.19	2.97	4.57	3.91	15.30
3	EBLUP _b	PL _y	18	64.09	66.14	-2.04	-2.60	7.56	11.79	8.64	78.76
3	ZEP _b	PL _y	18	64.09	66.75	-2.65	-3.51	7.95	12.24	9.74	101.83
3	EBLUP _b	PNL _y	19	64.09	63.12	0.97	1.42	1.96	3.06	2.25	5.99
3	ZEP _b	PNL _y	19	64.09	63.05	1.05	1.53	1.97	3.07	2.23	6.05
3	ZEP _b	PL _y	19	64.09	66.37	-2.28	-3.18	5.90	9.28	6.51	47.57
3	EBLUP _b	PL _y	19	64.09	66.52	-2.43	-3.40	6.06	9.52	6.69	50.68
3	ZEP _b	PNL _y	20	64.09	63.41	0.68	0.96	1.81	2.82	2.22	5.39
3	EBLUP _b	PNL _y	20	64.09	63.26	0.83	1.21	1.88	2.93	2.24	5.72
3	ZEP _b	PL _y	20	64.09	66.84	-2.75	-3.91	6.12	9.52	6.73	52.89
3	EBLUP _b	PL _y	20	64.09	66.56	-2.47	-3.45	6.20	9.68	6.89	53.56
3	EBLUP _b	PNL _y	21	64.09	63.54	0.55	0.75	1.66	2.59	2.14	4.87
3	ZEP _b	PNL _y	21	64.09	63.69	0.41	0.52	1.68	2.63	2.19	4.97
3	EBLUP _b	PL _y	21	64.09	66.94	-2.85	-4.06	6.02	9.42	6.59	51.52

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
3	ZEP _b	PL _y	21	64.09	67.16	-3.07	-4.41	6.13	9.58	6.65	53.58
3	EBLUP _b	PNL _y	21b	64.09	64.77	-0.68	-1.07	2.22	3.47	2.67	7.56
3	ZEP _b	PNL _y	21b	64.09	65.11	-1.01	-1.56	3.07	4.72	3.61	14.04
3	EBLUP _b	PL _y	21b	64.09	67.57	-3.48	-4.91	7.53	11.71	8.41	82.87
3	ZEP _b	PL _y	21b	64.09	68.94	-4.85	-6.99	9.03	13.90	10.44	132.50
3	ZEP _b	PNL _y	21c	64.09	64.28	-0.18	-0.25	2.78	4.29	3.31	10.98
3	ZEP _b	PL _y	21c	64.09	69.49	-5.40	-7.83	8.78	13.48	9.82	125.47
3	EBLUP _b	PNL _y	21c	64.09	79.39	-15.30	-23.70	15.30	23.70	4.21	251.76
3	EBLUP _b	PL _y	21c	64.09	76.00	-11.90	-17.46	15.91	24.49	14.67	356.77
3	EBLUP _b	PNL _y	22	64.09	64.02	0.08	0.13	2.00	3.13	2.64	6.99
3	ZEP _b	PNL _y	22	64.09	64.24	-0.15	-0.21	2.25	3.47	3.01	9.08
3	EBLUP _b	PL _y	22	64.09	67.55	-3.46	-4.92	6.45	10.01	6.97	60.56
3	ZEP _b	PL _y	22	64.09	68.07	-3.98	-5.70	6.82	10.49	7.85	77.48
3	ZEP _b	PNL _y	23	64.09	63.65	0.44	0.60	1.79	2.81	2.23	5.18
3	EBLUP _b	PNL _y	23	64.09	63.69	0.40	0.54	1.81	2.83	2.25	5.23
3	ZEP _b	PL _y	23	64.09	66.21	-2.11	-2.90	5.83	9.19	6.44	45.97
3	EBLUP _b	PL _y	23	64.09	66.29	-2.19	-3.02	5.94	9.36	6.58	48.13
3	ZEP _b	PNL _y	24	64.09	63.67	0.42	0.57	1.78	2.78	2.23	5.13
3	EBLUP _b	PNL _y	24	64.09	63.72	0.38	0.51	1.80	2.81	2.25	5.19
3	ZEP _b	PL _y	24	64.09	66.29	-2.19	-3.03	5.98	9.41	6.62	48.70
3	EBLUP _b	PL _y	24	64.09	66.38	-2.28	-3.15	6.10	9.59	6.78	51.10
3	ZEP _b	PNL _y	25	64.09	63.71	0.38	0.50	1.76	2.76	2.22	5.08
3	EBLUP _b	PNL _y	25	64.09	63.76	0.34	0.44	1.78	2.79	2.24	5.13
3	ZEP _b	PL _y	25	64.09	66.29	-2.19	-3.04	5.79	9.11	6.38	45.50
3	EBLUP _b	PL _y	25	64.09	66.38	-2.28	-3.17	5.90	9.28	6.52	47.74
3	EBLUP _b	PNL _y	26	64.09	63.96	0.14	0.18	1.70	2.67	2.28	5.22
3	ZEP _b	PNL _y	26	64.09	64.09	0.00	-0.04	1.76	2.74	2.40	5.77
3	EBLUP _b	PL _y	26	64.09	66.95	-2.85	-4.00	6.25	9.74	6.86	55.19

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
3	ZEP _b	PL _y	26	64.09	67.27	-3.18	-4.51	6.34	9.81	7.10	60.47
3	EBLUP _b	PNL _y	27	64.09	64.22	-0.13	-0.25	1.81	2.83	2.29	5.25
3	ZEP _b	PNL _y	27	64.09	64.39	-0.29	-0.51	1.86	2.91	2.34	5.55
3	EBLUP _b	PL _y	27	64.09	67.25	-3.16	-4.48	6.42	10.03	6.96	58.40
3	ZEP _b	PL _y	27	64.09	67.67	-3.57	-5.13	6.62	10.29	7.12	63.46
3	EBLUP _b	PNL _y	28	64.09	63.87	0.22	0.38	2.66	4.21	3.32	11.09
3	ZEP _b	PNL _y	28	64.09	64.24	-0.15	-0.23	2.83	4.43	3.63	13.21
3	EBLUP _b	PL _y	28	64.09	63.72	0.38	1.01	5.76	9.21	7.23	52.44
3	ZEP _b	PL _y	28	64.09	64.44	-0.35	-0.18	5.98	9.45	7.89	62.31
3	EBLUP _b	PNL _y	29	64.09	64.32	-0.22	-0.32	2.55	4.01	3.18	10.14
3	ZEP _b	PNL _y	29	64.09	64.00	0.10	0.15	2.83	4.42	3.63	13.18
3	EBLUP _b	PL _y	29	64.09	63.69	0.40	1.02	5.89	9.41	7.33	53.94
3	ZEP _b	PL _y	29	64.09	63.76	0.33	0.86	6.23	9.86	8.00	64.16
3	ZEP _b	PNL _y	30	64.09	63.93	0.16	0.15	1.73	2.71	2.45	6.01
3	EBLUP _b	PNL _y	30	64.09	63.61	0.49	0.66	1.78	2.77	2.42	6.08
3	EBLUP _b	PL _y	30	64.09	65.66	-1.57	-2.12	4.37	6.84	5.23	29.77
3	ZEP _b	PL _y	30	64.09	66.46	-2.36	-3.36	4.74	7.35	5.68	37.80
3	EBLUP _b	PNL _y	31	64.09	66.47	-2.37	-3.64	3.06	4.75	2.84	13.71
3	ZEP _b	PNL _y	31	64.09	66.31	-2.21	-3.34	3.75	5.78	3.92	20.29
3	EBLUP _b	PL _y	31	64.09	69.28	-5.18	-7.56	7.83	12.14	7.70	86.17
3	ZEP _b	PL _y	31	64.09	69.96	-5.86	-8.54	9.07	13.96	9.62	127.00
3	EBLUP _b	PNL _y	32	64.09	63.61	0.48	0.67	1.61	2.52	2.18	4.96
3	ZEP _b	PNL _y	32	64.09	63.88	0.22	0.26	1.60	2.51	2.23	5.02
3	EBLUP _b	PL _y	32	64.09	66.36	-2.27	-3.14	5.43	8.49	6.05	41.77
3	ZEP _b	PL _y	32	64.09	66.81	-2.72	-3.85	5.58	8.66	6.20	45.84
3	ZEP _b	PNL _y	2	64.09	62.30	1.79	2.66	2.52	3.89	2.61	10.03
3	ZEP _b	PL _y	2	64.09	63.99	0.11	0.23	2.97	4.69	3.61	13.06
3	EBLUP _b	PNL _y	2	64.09	67.03	-2.94	-4.87	7.76	12.14	7.59	66.29

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
3	<i>EBLUP_b</i>	<i>PL_y</i>	2	64.09	66.59	-2.49	-4.01	8.03	12.35	9.02	87.64
3	<i>EBLUP_b</i>	<i>PNL_y</i>	15	64.09	63.88	0.22	0.33	2.72	4.24	3.42	11.73
3	<i>ZEP_b</i>	<i>PNL_y</i>	15	64.09	63.22	0.87	1.39	2.98	4.67	3.57	13.49
3	<i>EBLUP_b</i>	<i>PL_y</i>	15	64.09	63.36	0.74	1.54	6.18	9.80	6.97	49.07
3	<i>ZEP_b</i>	<i>PL_y</i>	15	64.09	62.43	1.66	3.04	6.18	9.89	7.15	53.87
3	<i>EBLUP_b</i>	<i>PNL_y</i>	2	64.09	64.78	-0.69	-1.12	2.36	3.67	2.97	9.27
3	<i>ZEP_b</i>	<i>PNL_y</i>	2	64.09	62.11	1.98	3.02	2.77	4.31	2.73	11.39
3	<i>ZEP_b</i>	<i>PL_y</i>	2	64.09	60.46	3.63	5.98	5.31	8.59	5.61	44.65
3	<i>EBLUP_b</i>	<i>PL_y</i>	2	64.09	63.87	0.22	0.71	6.46	10.21	7.26	52.81
3	<i>EBLUP_b</i>	<i>PL_y</i>	9	64.09	62.17	1.92	2.92	2.72	4.23	2.67	10.80
3	<i>ZEP_b</i>	<i>PL_y</i>	9	64.09	60.14	3.95	6.06	4.18	6.45	2.60	22.37
3	<i>EBLUP_b</i>	<i>PNL_y</i>	9	64.09	63.27	0.83	0.86	5.19	8.23	6.07	37.48
3	<i>ZEP_b</i>	<i>PNL_y</i>	9	64.09	61.82	2.27	3.10	5.63	8.83	6.08	42.17
4	<i>EBLUP_b</i>	<i>PNL_y</i>	17	63.25	63.67	-0.42	-0.76	2.54	4.03	3.80	14.60
4	<i>ZEP_b</i>	<i>PNL_y</i>	17	63.25	64.30	-1.04	-1.73	2.95	4.67	3.91	16.38
4	<i>EBLUP_b</i>	<i>PL_y</i>	17	63.25	66.69	-3.43	-4.95	7.37	11.47	8.26	80.03
4	<i>ZEP_b</i>	<i>PL_y</i>	17	63.25	68.05	-4.79	-7.03	8.14	12.51	9.34	110.25
4	<i>ZEP_b</i>	<i>PNL_y</i>	18	63.25	63.18	0.08	0.29	3.48	5.48	4.36	19.02
4	<i>EBLUP_b</i>	<i>PNL_y</i>	18	63.25	62.92	0.33	0.70	3.46	5.56	4.38	19.26
4	<i>EBLUP_b</i>	<i>PL_y</i>	18	63.25	64.71	-1.45	-1.57	8.35	13.21	9.39	90.27
4	<i>ZEP_b</i>	<i>PL_y</i>	18	63.25	65.27	-2.02	-2.44	8.39	13.07	10.08	105.70
4	<i>ZEP_b</i>	<i>PNL_y</i>	19	63.25	62.36	0.89	1.33	2.30	3.66	2.84	8.85
4	<i>EBLUP_b</i>	<i>PNL_y</i>	19	63.25	62.39	0.86	1.30	2.34	3.73	2.89	9.07
4	<i>ZEP_b</i>	<i>PL_y</i>	19	63.25	65.05	-1.80	-2.40	6.58	10.42	7.18	54.77
4	<i>EBLUP_b</i>	<i>PL_y</i>	19	63.25	65.14	-1.89	-2.51	6.76	10.71	7.41	58.41
4	<i>ZEP_b</i>	<i>PNL_y</i>	20	63.25	62.94	0.32	0.37	2.22	3.54	3.16	10.08
4	<i>EBLUP_b</i>	<i>PNL_y</i>	20	63.25	62.82	0.43	0.58	2.25	3.58	3.23	10.60
4	<i>ZEP_b</i>	<i>PL_y</i>	20	63.25	65.63	-2.38	-3.36	6.44	10.05	7.11	56.28

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
4	<i>EBLUP_b</i>	<i>PL_y</i>	20	63.25	65.43	-2.17	-2.99	6.69	10.49	7.42	59.81
4	<i>ZEP_b</i>	<i>PNL_y</i>	21	63.25	63.51	-0.25	-0.55	2.03	3.26	2.89	8.43
4	<i>EBLUP_b</i>	<i>PNL_y</i>	21	63.25	63.34	-0.09	-0.28	1.96	3.14	2.92	8.56
4	<i>EBLUP_b</i>	<i>PL_y</i>	21	63.25	66.25	-3.00	-4.33	6.28	9.86	6.81	55.36
4	<i>ZEP_b</i>	<i>PL_y</i>	21	63.25	66.55	-3.30	-4.83	6.35	9.94	6.80	57.17
4	<i>EBLUP_b</i>	<i>PNL_y</i>	21b	63.25	64.23	-0.97	-1.57	2.81	4.42	3.32	12.00
4	<i>ZEP_b</i>	<i>PNL_y</i>	21b	63.25	64.99	-1.73	-2.73	3.53	5.50	3.89	18.13
4	<i>EBLUP_b</i>	<i>PL_y</i>	21b	63.25	66.55	-3.30	-4.66	7.85	12.25	8.73	87.16
4	<i>ZEP_b</i>	<i>PL_y</i>	21b	63.25	68.42	-5.16	-7.50	9.25	14.25	10.65	140.04
4	<i>ZEP_b</i>	<i>PNL_y</i>	21c	63.25	64.94	-1.68	-2.71	3.43	5.39	4.20	20.48
4	<i>ZEP_b</i>	<i>PL_y</i>	21c	63.25	68.46	-5.21	-7.56	9.00	13.88	10.09	128.91
4	<i>EBLUP_b</i>	<i>PNL_y</i>	21c	63.25	77.34	-14.09	-22.38	14.11	22.42	7.80	259.28
4	<i>EBLUP_b</i>	<i>PL_y</i>	21c	63.25	78.35	-15.10	-23.16	15.30	23.53	11.95	370.80
4	<i>ZEP_b</i>	<i>PNL_y</i>	22	63.25	63.62	-0.37	-0.58	2.50	3.93	3.67	13.58
4	<i>EBLUP_b</i>	<i>PNL_y</i>	22	63.25	63.42	-0.16	-0.25	2.43	3.86	3.68	13.60
4	<i>EBLUP_b</i>	<i>PL_y</i>	22	63.25	65.87	-2.62	-3.63	6.40	10.06	7.26	59.51
4	<i>ZEP_b</i>	<i>PL_y</i>	22	63.25	66.27	-3.02	-4.27	6.47	10.06	7.54	65.99
4	<i>ZEP_b</i>	<i>PNL_y</i>	23	63.25	62.92	0.33	0.45	2.11	3.37	2.61	6.94
4	<i>EBLUP_b</i>	<i>PNL_y</i>	23	63.25	62.89	0.37	0.53	2.16	3.46	2.66	7.20
4	<i>ZEP_b</i>	<i>PL_y</i>	23	63.25	64.88	-1.62	-2.12	6.25	9.92	6.82	49.11
4	<i>EBLUP_b</i>	<i>PL_y</i>	23	63.25	64.84	-1.58	-2.03	6.44	10.22	7.07	52.47
4	<i>ZEP_b</i>	<i>PNL_y</i>	24	63.25	63.07	0.19	0.22	2.17	3.46	2.93	8.63
4	<i>EBLUP_b</i>	<i>PNL_y</i>	24	63.25	63.05	0.20	0.26	2.24	3.57	3.01	9.11
4	<i>ZEP_b</i>	<i>PL_y</i>	24	63.25	65.15	-1.90	-2.55	6.54	10.35	7.15	54.77
4	<i>EBLUP_b</i>	<i>PL_y</i>	24	63.25	65.15	-1.89	-2.51	6.73	10.65	7.39	58.24
4	<i>ZEP_b</i>	<i>PNL_y</i>	25	63.25	63.14	0.11	0.10	2.18	3.48	3.02	9.12
4	<i>EBLUP_b</i>	<i>PNL_y</i>	25	63.25	63.14	0.11	0.12	2.24	3.57	3.11	9.69
4	<i>ZEP_b</i>	<i>PL_y</i>	25	63.25	65.15	-1.90	-2.57	6.32	10.00	6.92	51.54

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
4	<i>EBLUP_b</i>	<i>PL_y</i>	25	63.25	65.17	-1.91	-2.56	6.49	10.28	7.15	54.77
4	<i>ZEP_b</i>	<i>PNL_y</i>	26	63.25	63.47	-0.21	-0.40	2.13	3.39	3.13	9.87
4	<i>EBLUP_b</i>	<i>PNL_y</i>	26	63.25	63.32	-0.07	-0.15	2.15	3.41	3.19	10.20
4	<i>EBLUP_b</i>	<i>PL_y</i>	26	63.25	65.56	-2.30	-3.16	6.43	10.12	7.18	56.89
4	<i>ZEP_b</i>	<i>PL_y</i>	26	63.25	65.83	-2.58	-3.61	6.34	9.90	7.14	57.59
4	<i>EBLUP_b</i>	<i>PNL_y</i>	27	63.25	63.62	-0.37	-0.63	2.36	3.75	3.28	10.87
4	<i>ZEP_b</i>	<i>PNL_y</i>	27	63.25	63.84	-0.59	-0.99	2.44	3.87	3.27	11.05
4	<i>EBLUP_b</i>	<i>PL_y</i>	27	63.25	65.93	-2.68	-3.74	6.68	10.50	7.40	61.99
4	<i>ZEP_b</i>	<i>PL_y</i>	27	63.25	66.42	-3.16	-4.52	6.78	10.59	7.49	66.11
4	<i>ZEP_b</i>	<i>PNL_y</i>	28	63.25	62.90	0.35	0.48	1.96	3.13	2.76	7.76
4	<i>EBLUP_b</i>	<i>PNL_y</i>	28	63.25	62.55	0.70	1.09	2.11	3.39	2.87	8.73
4	<i>ZEP_b</i>	<i>PL_y</i>	28	63.25	61.81	1.44	2.53	4.08	6.61	4.98	26.84
4	<i>EBLUP_b</i>	<i>PL_y</i>	28	63.25	61.25	2.00	3.49	4.41	7.22	5.27	31.75
4	<i>ZEP_b</i>	<i>PNL_y</i>	29	63.25	63.10	0.16	0.10	1.85	2.97	2.70	7.32
4	<i>EBLUP_b</i>	<i>PNL_y</i>	29	63.25	63.08	0.18	0.18	1.80	2.88	2.70	7.34
4	<i>ZEP_b</i>	<i>PL_y</i>	29	63.25	61.56	1.70	2.83	3.83	6.19	4.44	22.59
4	<i>EBLUP_b</i>	<i>PL_y</i>	29	63.25	61.27	1.98	3.36	4.05	6.60	4.65	25.50
4	<i>EBLUP_b</i>	<i>PNL_y</i>	30	63.25	63.41	-0.16	-0.37	2.06	3.29	3.29	10.88
4	<i>ZEP_b</i>	<i>PNL_y</i>	30	63.25	63.66	-0.41	-0.78	2.13	3.43	3.30	11.09
4	<i>EBLUP_b</i>	<i>PL_y</i>	30	63.25	64.38	-1.12	-1.47	4.25	6.74	5.30	29.35
4	<i>ZEP_b</i>	<i>PL_y</i>	30	63.25	64.95	-1.70	-2.40	4.34	6.81	5.41	32.15
4	<i>EBLUP_b</i>	<i>PNL_y</i>	31	63.25	65.28	-2.02	-3.19	3.18	4.99	3.64	17.35
4	<i>ZEP_b</i>	<i>PNL_y</i>	31	63.25	65.74	-2.49	-3.89	3.61	5.63	3.88	21.24
4	<i>EBLUP_b</i>	<i>PL_y</i>	31	63.25	67.14	-3.89	-5.65	6.86	10.69	7.42	70.22
4	<i>ZEP_b</i>	<i>PL_y</i>	31	63.25	68.28	-5.02	-7.38	7.63	11.74	8.47	96.95
4	<i>EBLUP_b</i>	<i>PNL_y</i>	32	63.25	63.38	-0.13	-0.31	1.99	3.19	3.23	10.48
4	<i>ZEP_b</i>	<i>PNL_y</i>	32	63.25	63.63	-0.37	-0.72	2.07	3.32	3.24	10.65
4	<i>EBLUP_b</i>	<i>PL_y</i>	32	63.25	65.34	-2.09	-2.91	5.41	8.52	6.17	42.38

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
4	ZEP _b	PL _y	32	63.25	65.81	-2.55	-3.67	5.43	8.47	6.19	44.85
4	ZEP _b	PNL _y	2	63.25	61.83	1.42	2.02	2.87	4.52	3.20	12.26
4	ZEP _b	PL _y	2	63.25	63.13	0.12	0.20	2.98	4.73	3.53	12.49
4	EBLUP _b	PNL _y	2	63.25	62.28	0.97	1.06	5.53	8.74	6.15	38.75
4	EBLUP _b	PL _y	2	63.25	60.90	2.35	3.37	4.66	7.21	5.95	40.92
4	EBLUP _b	PNL _y	15	63.25	63.02	0.24	0.39	2.92	4.61	3.65	13.39
4	ZEP _b	PNL _y	15	63.25	62.43	0.82	1.34	3.19	5.06	3.78	14.99
4	EBLUP _b	PL _y	15	63.25	62.32	0.93	1.95	6.81	10.86	7.44	56.15
4	ZEP _b	PL _y	15	63.25	61.53	1.72	3.23	6.71	10.79	7.44	58.37
4	ZEP _b	PNL _y	2	63.25	62.25	1.01	1.41	2.77	4.36	3.26	11.64
4	EBLUP _b	PNL _y	2	63.25	64.77	-1.51	-2.51	3.00	4.78	3.39	13.79
4	ZEP _b	PL _y	2	63.25	60.74	2.52	4.19	4.49	7.27	4.64	27.84
4	EBLUP _b	PL _y	2	63.25	63.95	-0.69	-0.76	6.29	9.90	6.84	47.27
4	EBLUP _b	PL _y	9	63.25	62.51	0.75	1.00	2.62	4.18	3.20	10.78
4	ZEP _b	PL _y	9	63.25	60.56	2.69	4.03	3.63	5.65	3.33	18.34
4	EBLUP _b	PNL _y	9	63.25	63.68	-0.43	-1.26	5.83	9.44	6.89	47.62
4	ZEP _b	PNL _y	9	63.25	62.30	0.95	0.92	5.98	9.56	6.84	47.71
5	ZEP _b	PNL _y	17	64.56	64.43	0.13	0.10	1.77	2.77	2.34	5.51
5	EBLUP _b	PNL _y	17	64.56	63.59	0.97	1.38	1.86	2.89	2.20	5.76
5	EBLUP _b	PL _y	17	64.56	69.08	-4.52	-6.65	7.05	10.80	6.93	68.47
5	ZEP _b	PL _y	17	64.56	70.98	-6.42	-9.48	8.58	12.98	8.71	117.09
5	EBLUP _b	PNL _y	18	64.56	64.20	0.35	0.65	2.53	3.96	3.34	11.31
5	ZEP _b	PNL _y	18	64.56	64.59	-0.04	0.10	3.13	4.80	4.09	16.76
5	EBLUP _b	PL _y	18	64.56	67.90	-3.34	-4.64	7.69	11.84	8.28	79.76
5	ZEP _b	PL _y	18	64.56	68.62	-4.06	-5.64	8.34	12.66	9.98	116.09
5	ZEP _b	PNL _y	19	64.56	62.33	2.23	3.30	2.74	4.18	2.44	10.90
5	EBLUP _b	PNL _y	19	64.56	62.35	2.20	3.29	2.72	4.17	2.48	11.02
5	ZEP _b	PL _y	19	64.56	66.04	-1.48	-2.10	4.54	7.17	5.05	27.66

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
 RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
5	<i>EBLUP_b</i>	<i>PL_y</i>	19	64.56	66.01	-1.45	-2.00	5.05	7.95	5.79	35.64
5	<i>ZEP_b</i>	<i>PNL_y</i>	20	64.56	63.14	1.42	2.08	2.08	3.22	2.12	6.50
5	<i>EBLUP_b</i>	<i>PNL_y</i>	20	64.56	62.92	1.63	2.41	2.20	3.39	2.11	7.10
5	<i>EBLUP_b</i>	<i>PL_y</i>	20	64.56	67.25	-2.70	-3.87	5.70	8.82	5.95	42.71
5	<i>ZEP_b</i>	<i>PL_y</i>	20	64.56	67.87	-3.32	-4.81	5.97	9.16	6.19	49.32
5	<i>EBLUP_b</i>	<i>PNL_y</i>	21	64.56	62.83	1.72	2.54	2.29	3.51	2.13	7.52
5	<i>ZEP_b</i>	<i>PNL_y</i>	21	64.56	65.59	-1.03	-1.69	2.69	4.21	3.11	10.73
5	<i>EBLUP_b</i>	<i>PL_y</i>	21	64.56	70.50	-5.95	-8.88	8.39	12.95	8.12	101.32
5	<i>ZEP_b</i>	<i>PL_y</i>	21	64.56	81.21	-16.65	-25.23	18.94	29.04	17.79	593.83
5	<i>EBLUP_b</i>	<i>PNL_y</i>	21b	64.56	63.09	1.47	2.15	2.18	3.35	2.20	7.01
5	<i>ZEP_b</i>	<i>PNL_y</i>	21b	64.56	62.99	1.57	2.30	2.23	3.43	2.16	7.13
5	<i>EBLUP_b</i>	<i>PL_y</i>	21b	64.56	67.90	-3.35	-4.84	6.07	9.38	6.13	48.72
5	<i>ZEP_b</i>	<i>PL_y</i>	21b	64.56	69.20	-4.64	-6.83	7.45	11.44	7.18	73.12
5	<i>EBLUP_b</i>	<i>PL_y</i>	21c	64.56	70.31	-5.76	-8.55	7.31	11.25	7.85	94.75
5	<i>EBLUP_b</i>	<i>PNL_y</i>	21c	64.56	70.11	-5.55	-8.88	9.91	15.44	10.52	141.56
5	<i>ZEP_b</i>	<i>PNL_y</i>	21c	64.56	56.38	8.18	12.25	13.41	20.41	13.75	256.05
5	<i>ZEP_b</i>	<i>PL_y</i>	21c	64.56	52.73	11.83	18.10	18.65	28.48	18.52	482.91
5	<i>EBLUP_b</i>	<i>PNL_y</i>	22	64.56	63.13	1.42	2.16	2.12	3.30	2.22	6.94
5	<i>ZEP_b</i>	<i>PNL_y</i>	22	64.56	63.27	1.28	1.95	2.32	3.60	2.60	8.38
5	<i>EBLUP_b</i>	<i>PL_y</i>	22	64.56	66.71	-2.15	-3.00	4.65	7.20	5.45	34.27
5	<i>ZEP_b</i>	<i>PL_y</i>	22	64.56	66.99	-2.43	-3.41	5.01	7.68	6.28	45.41
5	<i>ZEP_b</i>	<i>PNL_y</i>	23	64.56	64.33	0.23	0.27	1.90	2.99	2.42	5.92
5	<i>EBLUP_b</i>	<i>PNL_y</i>	23	64.56	64.29	0.27	0.34	1.91	3.02	2.44	6.04
5	<i>ZEP_b</i>	<i>PL_y</i>	23	64.56	68.29	-3.74	-5.47	6.69	10.40	6.53	56.60
5	<i>EBLUP_b</i>	<i>PL_y</i>	23	64.56	68.25	-3.69	-5.38	6.75	10.49	6.64	57.69
5	<i>ZEP_b</i>	<i>PNL_y</i>	24	64.56	63.33	1.22	1.77	2.01	3.10	2.25	6.55
5	<i>EBLUP_b</i>	<i>PNL_y</i>	24	64.56	63.23	1.33	1.95	2.10	3.25	2.32	7.16
5	<i>ZEP_b</i>	<i>PL_y</i>	24	64.56	66.83	-2.27	-3.26	5.26	8.25	5.48	35.19

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
5	<i>EBLUP_b</i>	<i>PL_y</i>	24	64.56	66.58	-2.02	-2.85	5.32	8.35	5.72	36.77
5	<i>ZEP_b</i>	<i>PNL_y</i>	25	64.56	63.03	1.52	2.22	2.23	3.42	2.36	7.87
5	<i>EBLUP_b</i>	<i>PNL_y</i>	25	64.56	63.00	1.56	2.29	2.26	3.49	2.44	8.39
5	<i>ZEP_b</i>	<i>PL_y</i>	25	64.56	66.16	-1.61	-2.28	4.50	7.09	4.95	27.05
5	<i>EBLUP_b</i>	<i>PL_y</i>	25	64.56	66.00	-1.45	-1.99	4.84	7.64	5.51	32.43
5	<i>EBLUP_b</i>	<i>PNL_y</i>	26	64.56	63.52	1.04	1.53	1.78	2.77	2.04	5.22
5	<i>ZEP_b</i>	<i>PNL_y</i>	26	64.56	63.69	0.87	1.27	1.82	2.84	2.19	5.53
5	<i>EBLUP_b</i>	<i>PL_y</i>	26	64.56	67.19	-2.63	-3.75	5.24	8.10	5.59	38.12
5	<i>ZEP_b</i>	<i>PL_y</i>	26	64.56	67.53	-2.97	-4.26	5.41	8.30	5.91	43.74
5	<i>ZEP_b</i>	<i>PNL_y</i>	27	64.56	64.17	0.39	0.51	1.61	2.53	2.12	4.65
5	<i>EBLUP_b</i>	<i>PNL_y</i>	27	64.56	64.01	0.55	0.76	1.64	2.56	2.11	4.77
5	<i>EBLUP_b</i>	<i>PL_y</i>	27	64.56	68.21	-3.65	-5.32	6.14	9.50	6.04	49.77
5	<i>ZEP_b</i>	<i>PL_y</i>	27	64.56	68.51	-3.95	-5.80	6.26	9.64	6.03	51.99
5	<i>EBLUP_b</i>	<i>PNL_y</i>	28	64.56	62.77	1.79	2.75	3.48	5.41	3.60	16.18
5	<i>ZEP_b</i>	<i>PNL_y</i>	28	64.56	62.62	1.94	2.92	4.06	6.24	4.18	21.25
5	<i>EBLUP_b</i>	<i>PL_y</i>	28	64.56	62.56	2.00	3.37	5.68	8.99	6.66	48.34
5	<i>ZEP_b</i>	<i>PL_y</i>	28	64.56	62.38	2.18	3.54	6.54	10.20	7.50	60.97
5	<i>EBLUP_b</i>	<i>PNL_y</i>	29	64.56	62.83	1.73	2.61	3.89	5.99	3.92	18.35
5	<i>ZEP_b</i>	<i>PNL_y</i>	29	64.56	62.61	1.95	2.93	4.16	6.37	4.25	21.87
5	<i>EBLUP_b</i>	<i>PL_y</i>	29	64.56	62.17	2.39	3.91	6.48	10.19	7.03	55.12
5	<i>ZEP_b</i>	<i>PL_y</i>	29	64.56	62.12	2.43	3.92	6.88	10.71	7.67	64.78
5	<i>ZEP_b</i>	<i>PNL_y</i>	30	64.56	63.52	1.03	1.46	2.04	3.17	2.42	6.92
5	<i>EBLUP_b</i>	<i>PNL_y</i>	30	64.56	63.14	1.42	2.06	2.23	3.45	2.39	7.73
5	<i>EBLUP_b</i>	<i>PL_y</i>	30	64.56	65.98	-1.42	-1.97	3.64	5.65	4.42	21.56
5	<i>ZEP_b</i>	<i>PL_y</i>	30	64.56	67.05	-2.49	-3.62	4.27	6.56	4.90	30.23
5	<i>EBLUP_b</i>	<i>PNL_y</i>	31	64.56	64.08	0.48	0.63	2.88	4.45	3.41	11.86
5	<i>ZEP_b</i>	<i>PNL_y</i>	31	64.56	63.64	0.92	1.33	3.22	4.95	3.76	14.97
5	<i>EBLUP_b</i>	<i>PL_y</i>	31	64.56	68.00	-3.45	-5.04	5.03	7.75	5.49	42.06

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
5	ZEP _b	PL _y	31	64.56	68.48	-3.92	-5.73	5.77	8.84	6.80	61.67
5	ZEP _b	PNL _y	32	64.56	64.27	0.29	0.36	1.53	2.40	2.13	4.62
5	EBLUP _b	PNL _y	32	64.56	63.10	1.45	2.13	2.14	3.30	2.17	6.81
5	EBLUP _b	PL _y	32	64.56	68.13	-3.58	-5.20	5.90	9.09	5.88	47.40
5	ZEP _b	PL _y	32	64.56	71.54	-6.98	-10.36	9.00	13.71	8.38	119.00
5	EBLUP _b	PL _y	2	64.56	64.94	-0.38	-0.84	4.96	7.66	5.92	35.23
5	EBLUP _b	PNL _y	2	64.56	64.05	0.50	0.43	5.65	8.81	6.27	39.56
5	ZEP _b	PNL _y	2	64.56	55.95	8.61	12.91	9.66	14.70	8.68	149.46
5	ZEP _b	PL _y	2	64.56	53.58	10.97	16.54	11.89	18.03	11.18	245.33
5	EBLUP _b	PNL _y	15	64.56	63.66	0.90	1.37	2.58	4.03	3.09	10.34
5	ZEP _b	PNL _y	15	64.56	62.96	1.60	2.45	3.05	4.75	3.32	13.60
5	ZEP _b	PL _y	15	64.56	63.53	1.03	1.87	4.97	7.88	6.06	37.73
5	EBLUP _b	PL _y	15	64.56	64.60	-0.04	0.22	5.47	8.62	6.27	39.28
5	EBLUP _b	PNL _y	2	64.56	63.99	0.56	0.77	2.00	3.14	2.50	6.56
5	ZEP _b	PNL _y	2	64.56	61.45	3.11	4.65	3.82	5.86	3.59	22.55
5	ZEP _b	PL _y	2	64.56	61.43	3.12	4.97	4.15	6.58	4.04	26.10
5	EBLUP _b	PL _y	2	64.56	67.00	-2.44	-3.48	6.79	10.52	7.16	57.24
5	EBLUP _b	PL _y	9	64.56	62.00	2.56	3.83	3.16	4.88	2.70	13.85
5	ZEP _b	PL _y	9	64.56	59.83	4.72	7.13	5.11	7.81	3.56	34.95
5	ZEP _b	PNL _y	9	64.56	60.25	4.31	6.23	7.32	11.33	7.03	68.02
5	EBLUP _b	PNL _y	9	64.56	60.39	4.16	5.99	7.68	11.91	7.67	76.18
6	EBLUP _b	PNL _y	17	64.61	64.55	0.05	0.03	1.88	2.93	2.40	5.77
6	ZEP _b	PNL _y	17	64.61	65.17	-0.56	-0.91	2.20	3.40	2.79	8.12
6	EBLUP _b	PL _y	17	64.61	68.33	-3.72	-5.47	6.55	10.12	7.08	63.94
6	ZEP _b	PL _y	17	64.61	69.99	-5.38	-7.99	7.87	12.04	8.71	104.75
6	EBLUP _b	PNL _y	18	64.61	64.10	0.51	0.87	3.26	5.05	3.94	15.80
6	ZEP _b	PNL _y	18	64.61	64.61	-0.01	0.08	3.67	5.61	4.58	20.93
6	EBLUP _b	PL _y	18	64.61	66.55	-1.94	-2.60	7.53	11.68	8.56	77.00

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
6	ZEP _b	PL _y	18	64.61	67.50	-2.90	-4.05	8.25	12.67	9.78	104.00
6	EBLUP _b	PNL _y	19	64.61	63.34	1.27	1.88	1.92	2.96	2.07	5.88
6	ZEP _b	PNL _y	19	64.61	62.75	1.86	2.75	2.32	3.56	2.11	7.91
6	ZEP _b	PL _y	19	64.61	65.78	-1.17	-1.63	4.30	6.75	4.96	26
6	EBLUP _b	PL _y	19	64.61	66.74	-2.14	-3.08	5.20	8.11	5.76	37.75
6	EBLUP _b	PNL _y	20	64.61	63.68	0.93	1.36	1.75	2.70	2.12	5.34
6	ZEP _b	PNL _y	20	64.61	63.60	1.01	1.47	1.82	2.80	2.09	5.37
6	ZEP _b	PL _y	20	64.61	67.07	-2.46	-3.58	5.17	7.99	5.73	38.94
6	EBLUP _b	PL _y	20	64.61	66.97	-2.37	-3.42	5.37	8.34	5.91	40.55
6	EBLUP _b	PNL _y	21	64.61	64.37	0.24	0.28	1.46	2.27	1.97	3.95
6	ZEP _b	PNL _y	21	64.61	64.19	0.41	0.53	1.51	2.34	1.98	4.11
6	EBLUP _b	PL _y	21	64.61	67.76	-3.15	-4.65	5.49	8.52	5.81	43.68
6	ZEP _b	PL _y	21	64.61	67.91	-3.30	-4.90	5.65	8.78	5.97	46.60
6	EBLUP _b	PNL _y	21b	64.61	65.06	-0.45	-0.74	1.67	2.61	2.13	4.75
6	ZEP _b	PNL _y	21b	64.61	65.48	-0.87	-1.39	2.10	3.25	2.48	6.90
6	EBLUP _b	PL _y	21b	64.61	67.84	-3.23	-4.71	6.44	9.99	6.97	59.11
6	ZEP _b	PL _y	21b	64.61	68.99	-4.38	-6.47	7.42	11.43	8.01	83.37
6	EBLUP _b	PNL _y	21c	64.61	69.34	-4.74	-7.42	5.77	8.98	4.80	45.45
6	ZEP _b	PNL _y	21c	64.61	61.51	3.10	4.71	6.70	10.25	8.73	85.84
6	EBLUP _b	PL _y	21c	64.61	71.36	-6.75	-10.07	7.74	11.80	6.49	87.62
6	ZEP _b	PL _y	21c	64.61	62.84	1.77	2.94	10.60	16.26	13.05	173.41
6	EBLUP _b	PNL _y	22	64.61	64.05	0.55	0.84	1.77	2.74	2.42	6.18
6	ZEP _b	PNL _y	22	64.61	63.96	0.65	0.99	2.16	3.32	2.82	8.37
6	EBLUP _b	PL _y	22	64.61	67.79	-3.18	-4.62	5.70	8.80	6.08	47.02
6	ZEP _b	PL _y	22	64.61	67.82	-3.22	-4.67	5.66	8.68	6.85	57.30
6	ZEP _b	PNL _y	23	64.61	63.91	0.70	0.98	1.89	2.93	2.29	5.72
6	EBLUP _b	PNL _y	23	64.61	63.95	0.66	0.92	1.92	2.98	2.31	5.79
6	ZEP _b	PL _y	23	64.61	66.76	-2.16	-3.12	5.38	8.40	5.93	39.84

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
6	<i>EBLUP_b</i>	<i>PL_y</i>	23	64.61	66.79	-2.18	-3.14	5.45	8.52	6.03	41.11
6	<i>EBLUP_b</i>	<i>PNL_y</i>	24	64.61	63.88	0.72	1.04	1.67	2.58	2.06	4.79
6	<i>ZEP_b</i>	<i>PNL_y</i>	24	64.61	63.61	0.99	1.43	1.74	2.68	2.02	5.08
6	<i>ZEP_b</i>	<i>PL_y</i>	24	64.61	66.23	-1.62	-2.30	4.77	7.47	5.31	30.83
6	<i>EBLUP_b</i>	<i>PL_y</i>	24	64.61	66.60	-1.99	-2.85	5.15	8.05	5.73	36.75
6	<i>EBLUP_b</i>	<i>PNL_y</i>	25	64.61	63.88	0.72	1.03	1.64	2.54	2.04	4.68
6	<i>ZEP_b</i>	<i>PNL_y</i>	25	64.61	63.51	1.09	1.58	1.82	2.80	2.04	5.37
6	<i>ZEP_b</i>	<i>PL_y</i>	25	64.61	66.04	-1.44	-2.02	4.38	6.87	4.96	26.62
6	<i>EBLUP_b</i>	<i>PL_y</i>	25	64.61	66.58	-1.98	-2.83	4.92	7.68	5.45	33.67
6	<i>EBLUP_b</i>	<i>PNL_y</i>	26	64.61	64.17	0.44	0.63	1.60	2.50	2.20	5.03
6	<i>ZEP_b</i>	<i>PNL_y</i>	26	64.61	64.11	0.50	0.71	1.71	2.64	2.31	5.56
6	<i>EBLUP_b</i>	<i>PL_y</i>	26	64.61	67.30	-2.69	-3.89	5.52	8.56	5.97	42.90
6	<i>ZEP_b</i>	<i>PL_y</i>	26	64.61	67.36	-2.75	-3.99	5.43	8.37	6.11	44.87
6	<i>EBLUP_b</i>	<i>PNL_y</i>	27	64.61	64.52	0.08	0.07	1.67	2.61	2.18	4.76
6	<i>ZEP_b</i>	<i>PNL_y</i>	27	64.61	64.56	0.04	0.00	1.65	2.57	2.19	4.82
6	<i>EBLUP_b</i>	<i>PL_y</i>	27	64.61	67.75	-3.15	-4.60	5.75	8.93	6.06	46.64
6	<i>ZEP_b</i>	<i>PL_y</i>	27	64.61	68.00	-3.40	-4.99	5.85	9.05	6.12	49.02
6	<i>EBLUP_b</i>	<i>PNL_y</i>	28	64.61	64.94	-0.33	-0.54	1.89	2.96	2.44	6.07
6	<i>ZEP_b</i>	<i>PNL_y</i>	28	64.61	65.51	-0.90	-1.42	2.41	3.73	3.06	10.20
6	<i>EBLUP_b</i>	<i>PL_y</i>	28	64.61	65.56	-0.95	-1.19	4.77	7.51	6.31	40.76
6	<i>ZEP_b</i>	<i>PL_y</i>	28	64.61	66.51	-1.90	-2.67	5.43	8.45	7.54	60.46
6	<i>EBLUP_b</i>	<i>PNL_y</i>	29	64.61	64.94	-0.34	-0.55	1.66	2.59	2.21	4.98
6	<i>ZEP_b</i>	<i>PNL_y</i>	29	64.61	65.50	-0.89	-1.41	2.15	3.33	2.81	8.71
6	<i>EBLUP_b</i>	<i>PL_y</i>	29	64.61	65.17	-0.57	-0.61	4.73	7.45	6.20	38.79
6	<i>ZEP_b</i>	<i>PL_y</i>	29	64.61	66.14	-1.53	-2.11	5.33	8.29	7.37	56.70
6	<i>EBLUP_b</i>	<i>PNL_y</i>	30	64.61	64.27	0.33	0.45	1.68	2.62	2.23	5.10
6	<i>ZEP_b</i>	<i>PNL_y</i>	30	64.61	64.44	0.16	0.18	1.71	2.65	2.35	5.55
6	<i>EBLUP_b</i>	<i>PL_y</i>	30	64.61	66.82	-2.21	-3.17	4.57	7.09	5.21	32.04

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
6	ZEP _b	PL _y	30	64.61	67.40	-2.80	-4.08	4.95	7.63	5.74	40.82
6	EBLUP _b	PNL _y	31	64.61	65.70	-1.09	-1.69	1.86	2.90	2.17	5.89
6	ZEP _b	PNL _y	31	64.61	66.61	-2.00	-3.08	2.65	4.09	2.93	12.61
6	EBLUP _b	PL _y	31	64.61	68.84	-4.23	-6.23	6.35	9.82	6.35	58.20
6	ZEP _b	PL _y	31	64.61	70.57	-5.96	-8.86	7.79	11.95	8.05	100.34
6	ZEP _b	PNL _y	32	64.61	64.45	0.16	0.18	1.56	2.43	2.18	4.79
6	EBLUP _b	PL _y	32	64.61	67.48	-2.88	-4.19	5.32	8.25	5.64	40.15
6	ZEP _b	PL _y	32	64.61	67.69	-3.08	-4.52	5.37	8.30	5.74	42.45
6	EBLUP _b	PNL _y	2	64.61	64.46	0.15	0.10	1.94	3.01	2.48	6.18
6	EBLUP _b	PL _y	2	64.61	66.47	-1.86	-2.89	3.52	5.49	4.00	19.49
6	ZEP _b	PL _y	2	64.61	61.72	2.88	4.37	4.05	6.20	4.17	25.68
6	ZEP _b	PNL _y	2	64.61	60.86	3.75	5.63	4.30	6.55	4.27	32.28
6	EBLUP _b	PNL _y	15	64.61	64.16	0.45	0.66	2.30	3.58	2.92	8.73
6	ZEP _b	PNL _y	15	64.61	63.46	1.14	1.76	2.55	3.97	2.99	10.27
6	EBLUP _b	PL _y	15	64.61	63.90	0.71	1.30	5.24	8.25	5.93	35.67
6	ZEP _b	PL _y	15	64.61	62.97	1.64	2.79	5.07	8.04	5.91	37.63
6	EBLUP _b	PNL _y	2	64.61	65.05	-0.44	-0.78	1.81	2.83	2.40	5.95
6	ZEP _b	PNL _y	2	64.61	63.41	1.19	1.78	2.29	3.56	2.59	8.12
6	EBLUP _b	PL _y	2	64.61	64.41	0.20	0.44	4.39	6.90	5.07	25.78
6	ZEP _b	PL _y	2	64.61	62.47	2.14	3.48	3.95	6.30	4.74	27.04
6	EBLUP _b	PL _y	9	64.61	62.69	1.91	2.86	2.66	4.10	2.51	9.95
6	ZEP _b	PL _y	9	64.61	61.40	3.21	4.89	3.56	5.48	2.56	16.85
6	EBLUP _b	PNL _y	9	64.61	63.53	1.08	1.37	5.16	8.09	5.99	37.00
6	ZEP _b	PNL _y	9	64.61	62.45	2.15	3.05	5.25	8.17	5.75	37.72
7	EBLUP _b	PNL _y	17	62.20	61.87	0.32	0.45	2.11	3.34	3.70	13.80
7	ZEP _b	PNL _y	17	62.20	62.11	0.08	0.04	2.09	3.29	3.88	15.07
7	ZEP _b	PL _y	17	62.20	60.82	1.38	2.33	3.68	6.00	5.23	29.27
7	EBLUP _b	PL _y	17	62.20	60.38	1.81	3.08	3.94	6.46	5.20	30.35

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
7	ZEP _b	PNL _y	18	62.20	60.58	1.62	2.58	3.19	5.11	3.58	15.40
7	EBLUP _b	PNL _y	18	62.20	60.37	1.82	2.98	3.63	5.88	3.99	19.24
7	ZEP _b	PL _y	18	62.20	58.26	3.93	6.51	5.55	9.10	5.02	40.72
7	EBLUP _b	PL _y	18	62.20	57.95	4.25	7.11	6.00	9.89	5.71	50.68
7	EBLUP _b	PNL _y	19	62.20	61.20	1.00	1.49	2.18	3.42	3.17	11.02
7	ZEP _b	PNL _y	19	62.20	61.18	1.01	1.52	2.23	3.50	3.26	11.65
7	EBLUP _b	PL _y	19	62.20	60.02	2.17	3.64	3.93	6.46	4.69	26.70
7	ZEP _b	PL _y	19	62.20	59.99	2.20	3.68	4.04	6.62	4.83	28.16
7	EBLUP _b	PNL _y	20	62.20	61.63	0.56	0.80	2.05	3.22	3.50	12.58
7	ZEP _b	PNL _y	20	62.20	61.73	0.46	0.62	2.07	3.24	3.64	13.43
7	ZEP _b	PL _y	20	62.20	60.49	1.70	2.85	3.77	6.15	5.00	27.86
7	EBLUP _b	PL _y	20	62.20	60.24	1.96	3.29	3.88	6.36	4.92	28.08
7	EBLUP _b	PNL _y	21	62.20	62.05	0.15	0.11	1.74	2.71	3.05	9.30
7	ZEP _b	PNL _y	21	62.20	62.05	0.14	0.10	1.74	2.72	3.06	9.37
7	EBLUP _b	PL _y	21	62.20	60.76	1.44	2.43	3.43	5.64	4.51	22.43
7	ZEP _b	PL _y	21	62.20	60.87	1.32	2.25	3.43	5.62	4.61	22.98
7	EBLUP _b	PNL _y	21b	62.20	61.73	0.47	0.66	1.85	2.91	2.94	8.85
7	ZEP _b	PNL _y	21b	62.20	61.88	0.32	0.40	1.80	2.83	2.98	9.00
7	ZEP _b	PL _y	21b	62.20	59.82	2.38	3.95	4.07	6.65	4.48	25.77
7	EBLUP _b	PL _y	21b	62.20	59.43	2.76	4.60	4.35	7.14	4.61	28.85
7	ZEP _b	PNL _y	21c	62.20	63.32	-1.13	-1.79	2.22	3.52	3.73	15.21
7	EBLUP _b	PNL _y	21c	62.20	65.69	-3.49	-5.62	3.60	5.78	3.62	25.27
7	ZEP _b	PL _y	21c	62.20	62.37	-0.18	-0.05	4.06	6.61	5.79	33.51
7	EBLUP _b	PL _y	21c	62.20	64.19	-2.00	-2.95	4.26	6.87	5.72	36.68
7	EBLUP _b	PNL _y	22	62.20	62.12	0.08	0.09	2.29	3.63	3.92	15.35
7	ZEP _b	PNL _y	22	62.20	62.24	-0.05	-0.13	2.23	3.51	4.00	16.04
7	EBLUP _b	PL _y	22	62.20	61.83	0.37	0.82	4.11	6.75	5.79	33.64
7	ZEP _b	PL _y	22	62.20	62.10	0.09	0.35	4.10	6.68	5.84	34.15

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
 RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
7	ZEP _b	PNL _y	23	62.20	61.21	0.99	1.46	1.93	3.01	2.62	7.86
7	EBLUP _b	PNL _y	23	62.20	61.18	1.02	1.50	1.95	3.04	2.62	7.88
7	ZEP _b	PL _y	23	62.20	59.63	2.56	4.26	4.00	6.56	4.17	23.95
7	EBLUP _b	PL _y	23	62.20	59.55	2.64	4.39	4.03	6.62	4.13	24.07
7	EBLUP _b	PNL _y	24	62.20	61.69	0.51	0.70	1.87	2.93	3.10	9.89
7	ZEP _b	PNL _y	24	62.20	61.67	0.52	0.73	1.89	2.95	3.14	10.15
7	EBLUP _b	PL _y	24	62.20	59.93	2.26	3.79	3.96	6.51	4.67	26.92
7	ZEP _b	PL _y	24	62.20	59.97	2.23	3.73	4.02	6.60	4.82	28.23
7	EBLUP _b	PNL _y	25	62.20	61.80	0.40	0.53	1.89	2.95	3.21	10.46
7	ZEP _b	PNL _y	25	62.20	61.77	0.43	0.58	1.91	2.99	3.27	10.90
7	EBLUP _b	PL _y	25	62.20	60.21	1.99	3.35	3.84	6.32	4.77	26.72
7	ZEP _b	PL _y	25	62.20	60.23	1.97	3.31	3.93	6.45	4.95	28.37
7	EBLUP _b	PNL _y	26	62.20	62.02	0.18	0.21	2.04	3.21	3.57	12.75
7	ZEP _b	PNL _y	26	62.20	62.10	0.09	0.06	2.00	3.15	3.63	13.17
7	ZEP _b	PL _y	26	62.20	60.98	1.21	2.11	3.80	6.23	5.20	28.49
7	EBLUP _b	PL _y	26	62.20	60.78	1.41	2.46	3.91	6.43	5.23	29.33
7	EBLUP _b	PNL _y	27	62.20	61.87	0.33	0.43	1.89	2.97	3.25	10.68
7	ZEP _b	PNL _y	27	62.20	61.97	0.23	0.27	1.91	3.00	3.41	11.66
7	EBLUP _b	PL _y	27	62.20	60.68	1.51	2.60	3.73	6.15	4.98	27.14
7	ZEP _b	PL _y	27	62.20	60.93	1.27	2.19	3.70	6.07	5.18	28.46
7	EBLUP _b	PNL _y	28	62.20	62.15	0.04	0.02	1.98	3.16	3.06	9.37
7	ZEP _b	PNL _y	28	62.20	62.30	-0.10	-0.22	1.90	3.01	3.11	9.66
7	ZEP _b	PL _y	28	62.20	60.64	1.55	2.70	4.50	7.37	5.43	31.88
7	EBLUP _b	PL _y	28	62.20	60.29	1.91	3.30	4.65	7.66	5.44	33.28
7	EBLUP _b	PNL _y	29	62.20	62.17	0.03	-0.01	1.97	3.14	3.05	9.32
7	ZEP _b	PNL _y	29	62.20	62.40	-0.20	-0.39	1.86	2.95	3.08	9.52
7	ZEP _b	PL _y	29	62.20	60.72	1.47	2.56	4.41	7.21	5.31	30.37
7	EBLUP _b	PL _y	29	62.20	60.22	1.97	3.40	4.69	7.72	5.43	33.33

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

Cond. 1	MÉTODO		MODELO	Diámetro							
	2			Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ
7	<i>EBLUP_b</i>	<i>PNL_y</i>	30	62.20	62.08	0.12	0.09	1.85	2.90	3.27	10.71
7	<i>ZEP_b</i>	<i>PNL_y</i>	30	62.20	62.23	-0.03	-0.16	1.89	2.98	3.46	11.95
7	<i>EBLUP_b</i>	<i>PL_y</i>	30	62.20	61.24	0.95	1.71	3.63	5.98	5.07	26.57
7	<i>ZEP_b</i>	<i>PL_y</i>	30	62.20	61.60	0.60	1.12	3.56	5.82	5.24	27.83
7	<i>EBLUP_b</i>	<i>PNL_y</i>	31	62.20	62.40	-0.20	-0.36	1.88	3.00	3.21	10.32
7	<i>ZEP_b</i>	<i>PNL_y</i>	31	62.20	62.74	-0.54	-0.92	1.81	2.87	3.32	11.28
7	<i>EBLUP_b</i>	<i>PL_y</i>	31	62.20	61.41	0.79	1.49	4.06	6.68	5.40	29.79
7	<i>ZEP_b</i>	<i>PL_y</i>	31	62.20	61.98	0.22	0.55	3.91	6.38	5.46	29.90
7	<i>EBLUP_b</i>	<i>PNL_y</i>	32	62.20	62.14	0.05	-0.01	1.86	2.92	3.31	10.96
7	<i>ZEP_b</i>	<i>PNL_y</i>	32	62.20	62.24	-0.05	-0.18	1.90	2.99	3.49	12.15
7	<i>EBLUP_b</i>	<i>PL_y</i>	32	62.20	61.09	1.11	1.95	3.63	5.98	5.02	26.44
7	<i>ZEP_b</i>	<i>PL_y</i>	32	62.20	61.35	0.84	1.50	3.56	5.84	5.18	27.58
7	<i>EBLUP_b</i>	<i>PNL_y</i>	2	62.20	62.83	-0.64	-1.18	1.82	2.90	2.70	7.69
7	<i>ZEP_b</i>	<i>PNL_y</i>	2	62.20	61.19	1.01	1.44	2.13	3.31	3.00	9.99
7	<i>ZEP_b</i>	<i>PL_y</i>	2	62.20	60.27	1.93	3.02	3.07	4.92	3.21	14.05
7	<i>EBLUP_b</i>	<i>PL_y</i>	2	62.20	61.40	0.79	1.24	3.17	5.13	4.03	16.86
7	<i>EBLUP_b</i>	<i>PNL_y</i>	15	62.20	61.88	0.32	0.42	2.52	3.95	3.60	13.07
7	<i>ZEP_b</i>	<i>PNL_y</i>	15	62.20	61.00	1.19	1.87	2.82	4.47	3.57	14.13
7	<i>EBLUP_b</i>	<i>PL_y</i>	15	62.20	57.27	4.93	8.05	5.79	9.43	4.05	40.67
7	<i>ZEP_b</i>	<i>PL_y</i>	15	62.20	56.25	5.94	9.75	6.74	11.04	4.59	56.40
7	<i>ZEP_b</i>	<i>PNL_y</i>	2	62.20	61.63	0.56	0.78	2.05	3.19	3.10	9.92
7	<i>EBLUP_b</i>	<i>PNL_y</i>	2	62.20	62.90	-0.71	-1.29	2.30	3.65	3.22	10.87
7	<i>EBLUP_b</i>	<i>PL_y</i>	2	62.20	58.15	4.04	6.55	4.76	7.72	3.44	28.20
7	<i>ZEP_b</i>	<i>PL_y</i>	2	62.20	57.12	5.07	8.24	5.70	9.26	3.90	40.97
7	<i>EBLUP_b</i>	<i>PL_y</i>	9	62.20	61.07	1.13	1.65	2.54	3.98	3.21	11.59
7	<i>ZEP_b</i>	<i>PL_y</i>	9	62.20	60.05	2.15	3.31	2.91	4.55	2.99	13.58
7	<i>ZEP_b</i>	<i>PNL_y</i>	9	62.20	64.45	-2.26	-3.95	4.82	7.79	5.22	32.38

Continua en la página siguiente —

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Cuadro A.1 – Continua de la pagina previa

MÉTODO		Diámetro									
Cond. 1	2	MODELO	Obs.	Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	SD	δ	
7	<i>EBLUP_b</i>	<i>PNL_y</i>	9	62.20	65.62	-3.43	-5.85	5.59	9.10	5.63	43.49

Modelo	Diám.	Diám. Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	P_{50} $ \bar{e}\% $	P_{75} $ \bar{e}\% $	P_{90} $ \bar{e}\% $	SD	δ
32	64.63	63.55	1.08	1.57	2.04	3.14	2.54	3.97	6.55	2.56	7.72
27	64.50	64.03	0.48	0.65	2.21	3.44	3.16	4.49	6.46	2.76	7.84
30	63.22	62.40	0.82	1.14	2.17	3.52	2.67	5.30	7.21	2.72	8.06
26	64.09	63.43	0.67	1.02	2.26	3.55	2.92	4.53	7.89	2.78	8.16
22	64.38	64.24	0.14	0.26	2.44	3.83	3.40	5.28	6.87	2.88	8.33
24	64.07	63.39	0.68	0.98	2.41	3.81	3.82	4.93	6.95	2.89	8.84
25	64.64	63.94	0.71	1.01	2.54	3.93	4.05	5.25	7.53	2.96	9.24
17	64.30	63.27	1.03	1.43	2.47	3.87	3.44	4.86	7.50	2.89	9.39
21	64.70	63.30	1.40	2.04	2.39	3.69	3.10	4.92	7.13	2.76	9.56
23	64.22	63.19	1.04	1.55	2.58	4.06	3.90	5.56	7.10	3.01	10.12
20	64.35	62.95	1.39	2.03	2.58	3.99	3.34	5.69	7.96	2.97	10.77
2	63.97	62.53	1.44	2.17	2.80	4.41	4.05	6.34	8.36	3.04	11.33
19	64.24	63.05	1.19	1.71	2.75	4.34	3.89	5.85	8.23	3.34	12.57
15	63.92	62.64	1.28	1.98	3.69	5.76	5.45	8.62	11.34	4.41	21.11
18	63.80	63.24	0.56	1.08	3.89	6.10	5.52	9.07	12.98	4.80	23.33
31	64.44	67.42	-2.98	-4.56	4.12	6.39	6.09	9.04	12.02	3.88	23.95
29	64.92	65.66	-0.74	-1.08	4.01	6.25	5.51	9.68	11.89	4.90	24.56
28	63.88	63.31	0.57	1.03	4.38	6.96	5.79	11.33	13.96	5.41	29.62
29	63.65	62.42	1.23	1.39	5.44	8.72	7.89	12.12	15.94	6.35	41.85
2b	63.27	66.63	-3.36	-5.79	8.96	14.34	14.62	16.65	22.85	9.11	94.35

Cuadro A.2: Condición I en Base de Validación LONGITUDINALES

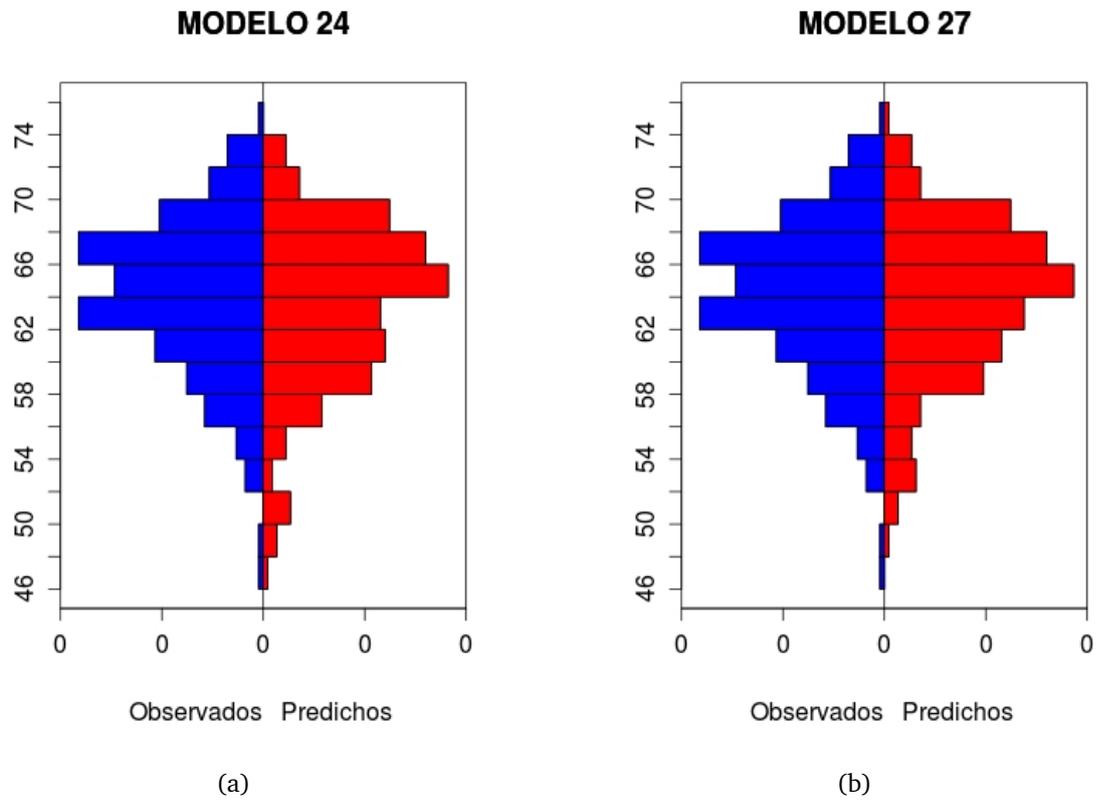


Figura A.1: Distribución de Diámetros Obs. para Base de Validación LONGITUDINALES

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

Modelo	Diám.	Diám. Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	P_{50} $ \bar{e}\% $	P_{75} $ \bar{e}\% $	P_{90} $ \bar{e}\% $	SD	δ
21	64.78	64.12	0.66	0.94	2.01	3.11	3.07	4.04	5.71	2.41	6.25
32	64.18	63.72	0.46	0.64	2.05	3.20	2.69	4.41	6.68	2.54	6.67
27	64.56	64.64	-0.08	-0.16	2.45	3.78	3.65	4.64	6.83	2.91	8.47
17	63.81	63.89	-0.09	-0.19	2.43	3.76	3.36	5.65	6.96	2.91	8.50
25	63.90	63.88	0.01	-0.02	2.43	3.81	3.30	5.48	7.84	2.92	8.55
26	64.56	64.83	-0.27	-0.46	2.34	3.62	3.25	5.02	7.42	2.92	8.62
24	64.24	64.29	-0.05	-0.15	2.54	3.98	3.50	5.44	7.95	2.98	8.86
30	64.30	63.84	0.46	0.65	2.18	3.36	2.42	5.02	7.75	2.97	9.04
20	64.23	63.49	0.74	1.08	2.56	3.95	3.95	5.52	7.05	2.94	9.21
19	63.49	63.09	0.39	0.55	2.54	4.01	3.87	5.60	7.35	3.09	9.73
23	63.05	63.29	-0.24	-0.43	2.60	4.11	3.38	5.76	7.97	3.21	10.34
22	64.47	64.90	-0.44	-0.66	2.76	4.27	3.78	6.10	7.07	3.32	11.21
2	63.96	64.10	-0.14	-0.27	2.74	4.26	3.70	6.46	9.16	3.38	11.44
29	64.10	64.42	-0.32	-0.43	3.64	5.71	5.27	7.57	9.28	4.20	17.73
15	64.00	63.52	0.48	0.72	3.51	5.47	4.97	7.78	10.35	4.32	18.88
18	63.72	63.92	-0.20	-0.19	3.54	5.56	5.40	7.48	10.25	4.48	20.15
28	64.63	65.07	-0.45	-0.67	3.84	6.00	5.13	8.09	11.81	4.48	20.25
31	64.55	68.18	-3.63	-5.60	4.45	6.86	5.01	11.13	14.13	4.13	30.28
9	63.56	62.76	0.80	0.85	4.92	7.86	7.44	11.41	13.69	5.78	34.01
2b	64.07	67.17	-3.10	-5.28	8.89	13.96	14.16	17.00	20.20	9.02	91.02

Cuadro A.3: Condición II en Base de Validación LONGITUDINALES

Modelo	Diám.	Diám. Pred.	\bar{e}	$\bar{e}\%$	$ \bar{e} $	$ \bar{e}\% $	P_{50} $ \bar{e}\% $	P_{75} $ \bar{e}\% $	P_{90} $ \bar{e}\% $	SD	δ
25	63.21	63.01	0.20	0.26	1.61	2.58	2.41	3.51	5.02	2.03	4.17
21	63.73	63.36	0.37	0.44	1.46	2.27	1.49	2.90	5.59	2.10	4.55
30	64.16	63.76	0.40	0.56	1.62	2.53	2.05	3.55	5.64	2.20	5.02
20	64.65	63.46	1.19	1.73	1.93	2.94	2.73	4.19	5.87	2.07	5.71
28	64.19	63.38	0.81	1.17	1.62	2.50	1.57	3.47	6.20	2.26	5.77
27	64.17	64.17	0.00	-0.06	1.88	2.93	2.57	4.20	5.98	2.43	5.92
23	64.20	64.08	0.12	0.06	1.96	3.10	2.51	4.50	6.37	2.44	5.97
24	64.60	64.36	0.24	0.27	1.88	2.95	2.52	3.89	5.82	2.44	6.02
22	64.10	64.06	0.04	0.05	2.01	3.15	2.95	4.40	6.01	2.48	6.17
26	63.73	63.67	0.07	0.03	1.87	2.95	1.92	3.94	6.80	2.58	6.69
19	64.23	63.18	1.05	1.47	2.19	3.40	3.09	5.03	6.24	2.50	7.34
17	63.93	63.78	0.15	0.11	2.18	3.42	2.82	4.68	6.52	2.73	7.45
18	63.75	63.34	0.41	0.80	2.52	4.00	3.57	6.00	8.93	3.09	9.71
28	64.29	63.98	0.30	0.53	2.47	3.91	3.18	6.22	7.76	3.20	10.36
2	64.43	64.98	-0.55	-0.92	2.39	3.65	2.70	5.33	8.94	3.19	10.47
29	63.74	64.25	-0.51	-0.79	2.70	4.32	3.89	5.45	6.58	3.38	11.69
31	63.10	65.43	-2.33	-3.59	2.97	4.68	4.38	6.50	8.56	2.75	13.01
15	63.83	63.54	0.29	0.41	2.72	4.25	3.59	5.99	8.61	3.62	13.17
9	63.80	63.51	0.29	-0.26	5.54	8.91	8.28	11.46	17.21	6.52	42.65
2b	64.05	66.18	-2.13	-3.75	7.77	12.19	11.69	13.52	18.09	7.94	67.51

Cuadro A.4: Condición III en Base de Validación LONGITUDINALES

Apéndice A. MÉTRICAS DE VALIDACIÓN.
RESULTADOS OBTENIDOS DETALLADOS

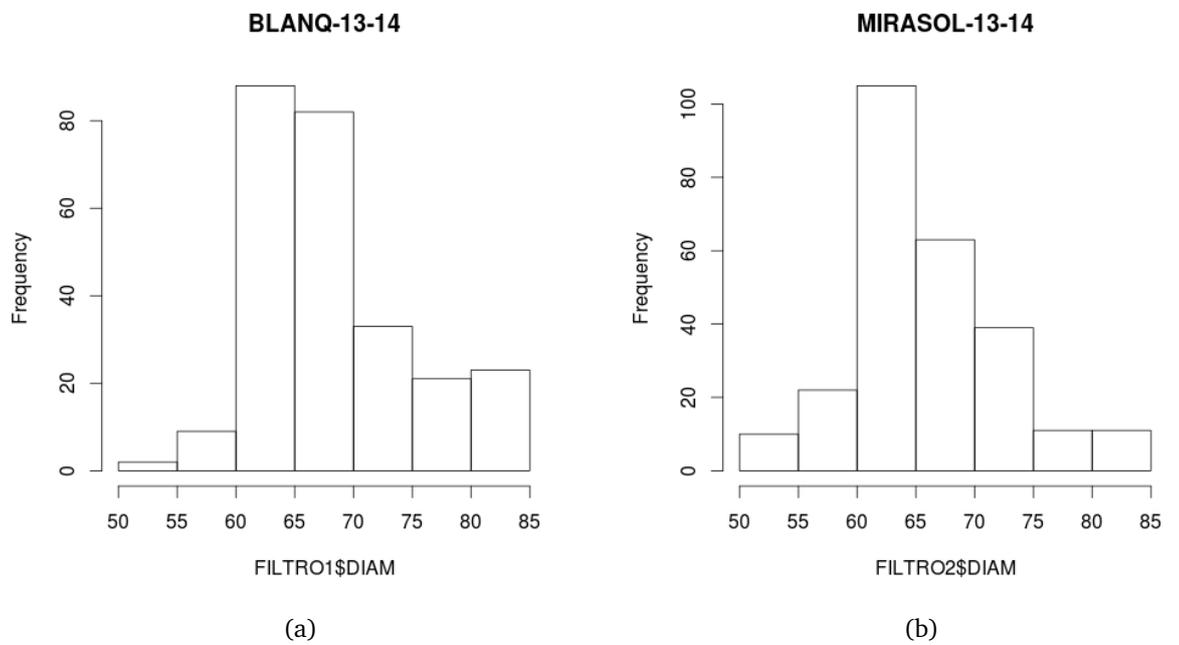


Figura A.2: Distribución de Diámetros Obs.para Base de Validación TRANSVERSALES

Apéndice B

VERIFICACIÓN GRÁFICA DE SUPUESTOS

Apéndice B. VERIFICACIÓN GRÁFICA DE SUPUESTOS

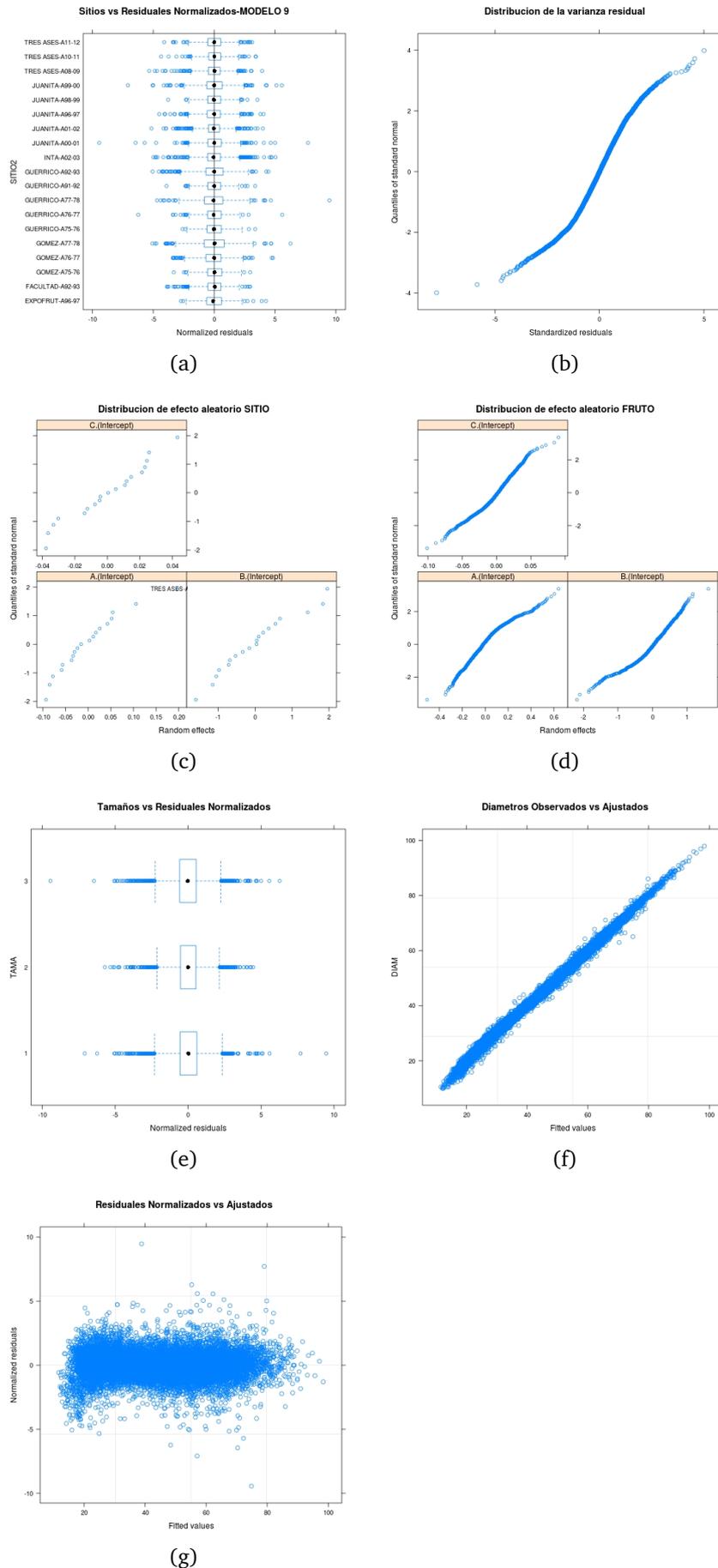


Figura B.1: Verificación Gráfica de Supuestos de MODELO 9

Apéndice B. VERIFICACIÓN GRÁFICA DE SUPUESTOS

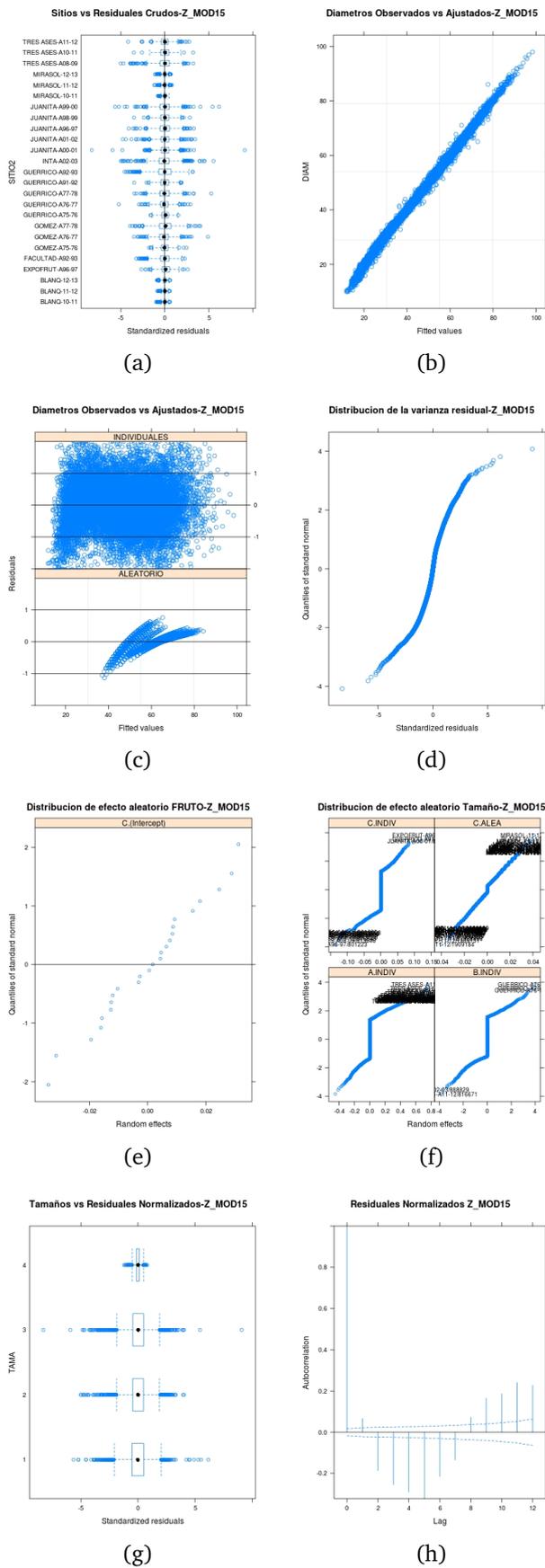


Figura B.2: Verificación Gráfica de Supuestos de MODELO 15

Diametros Observados vs Ajustados-Z_MOD11

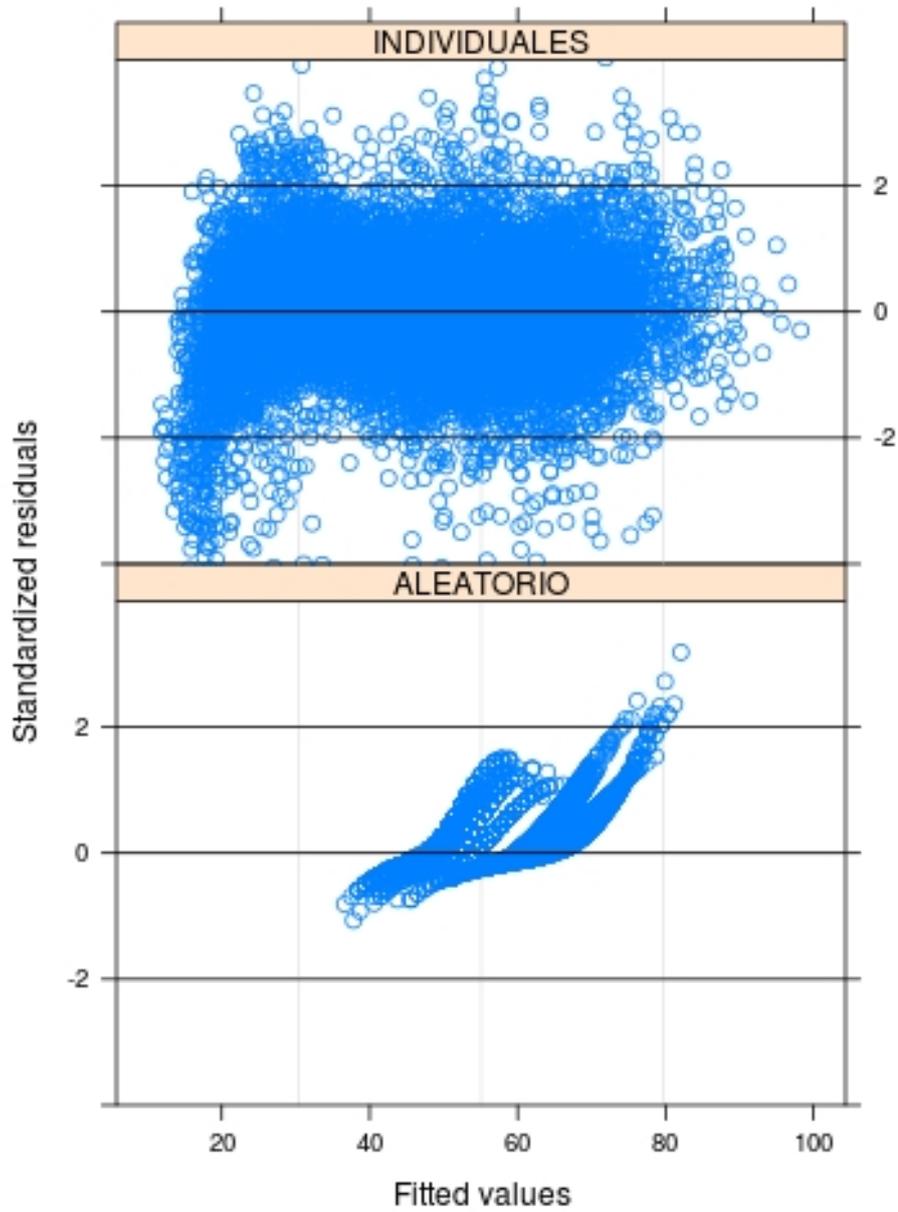


Figura B.3: Residuales de MODELO 11

Apéndice B. VERIFICACIÓN GRÁFICA DE SUPUESTOS

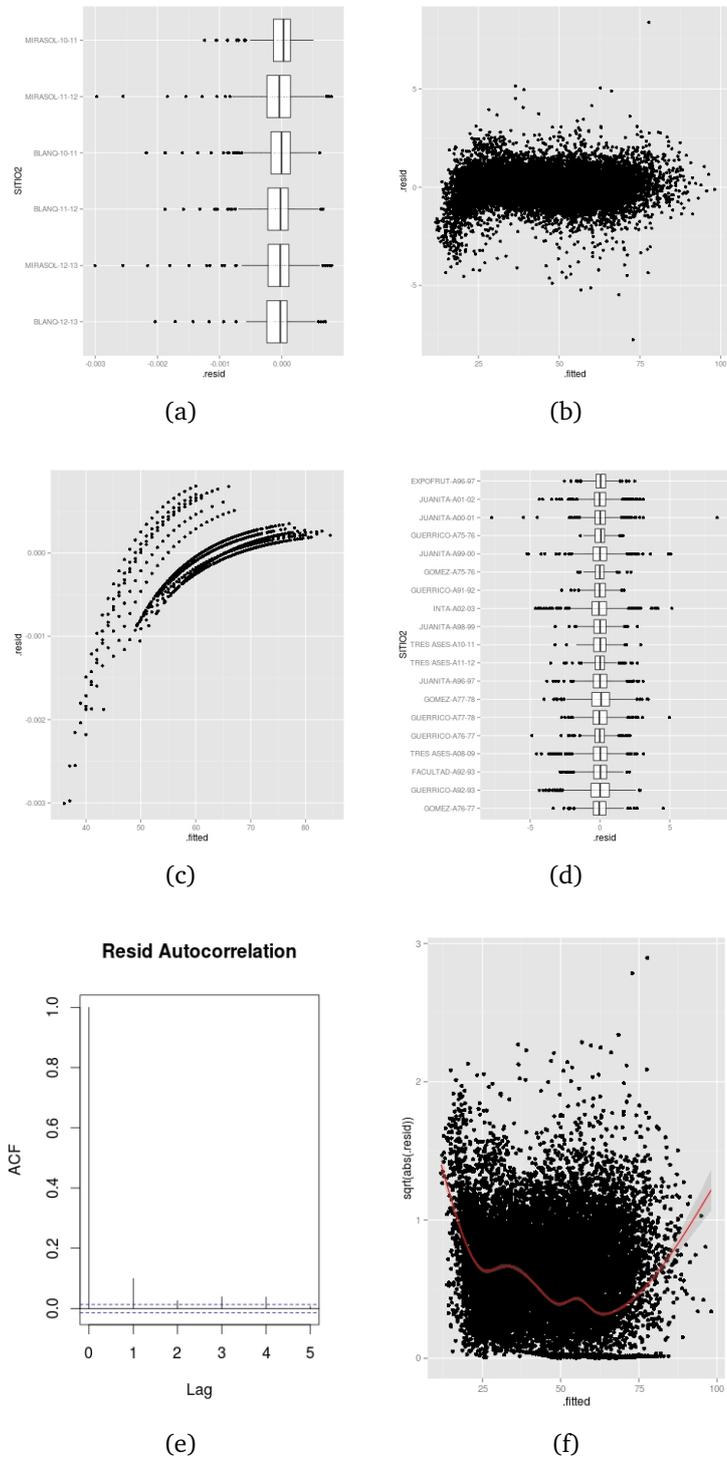


Figura B.4: Verificación Gráfica de Supuestos de MODELO 17

Apéndice B. VERIFICACIÓN GRÁFICA DE SUPUESTOS

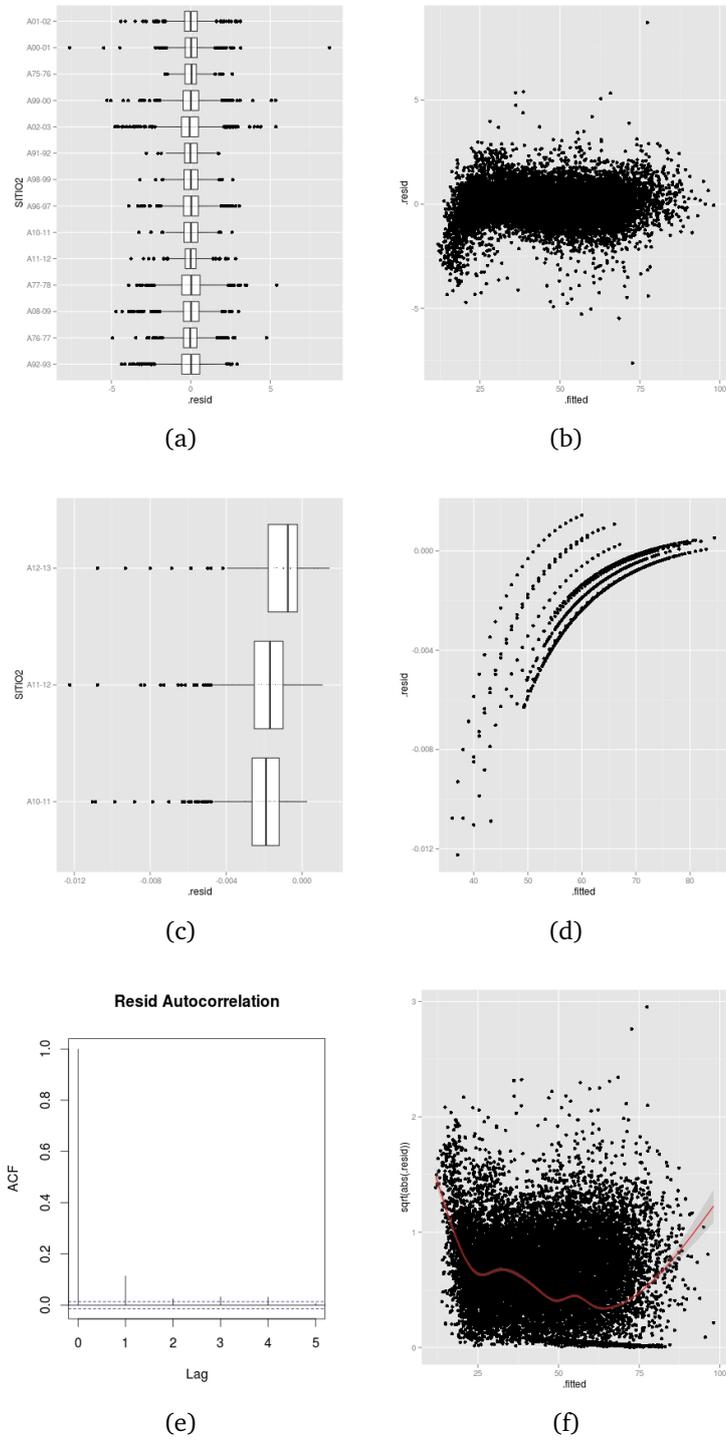


Figura B.5: Verificación Gráfica de Supuestos de MODELO 22

Apéndice C

CORRELACIONES MARGINALES

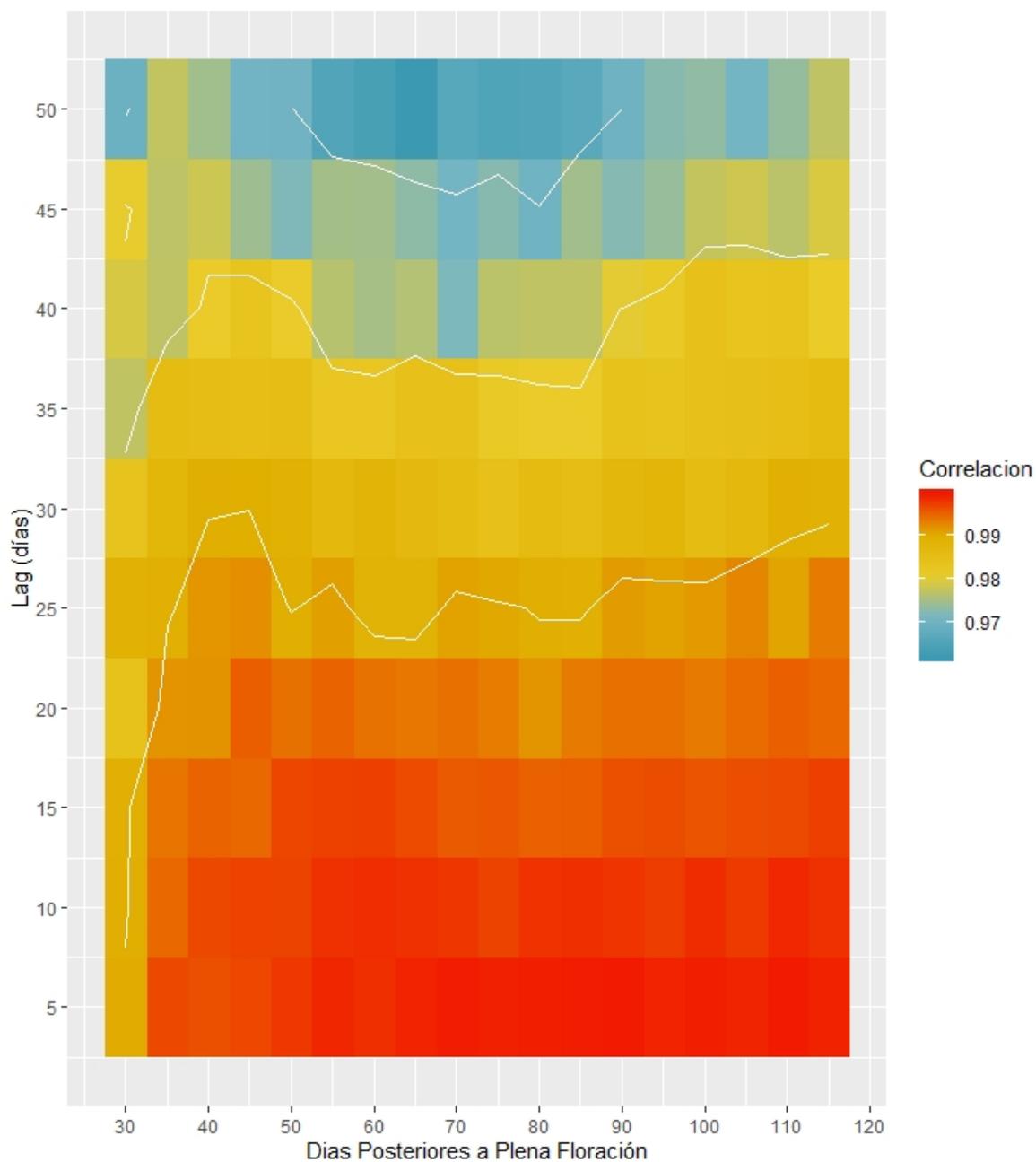


Figura C.1: Correlación Marginal para obs. del mismo Fruto en MODELO 21

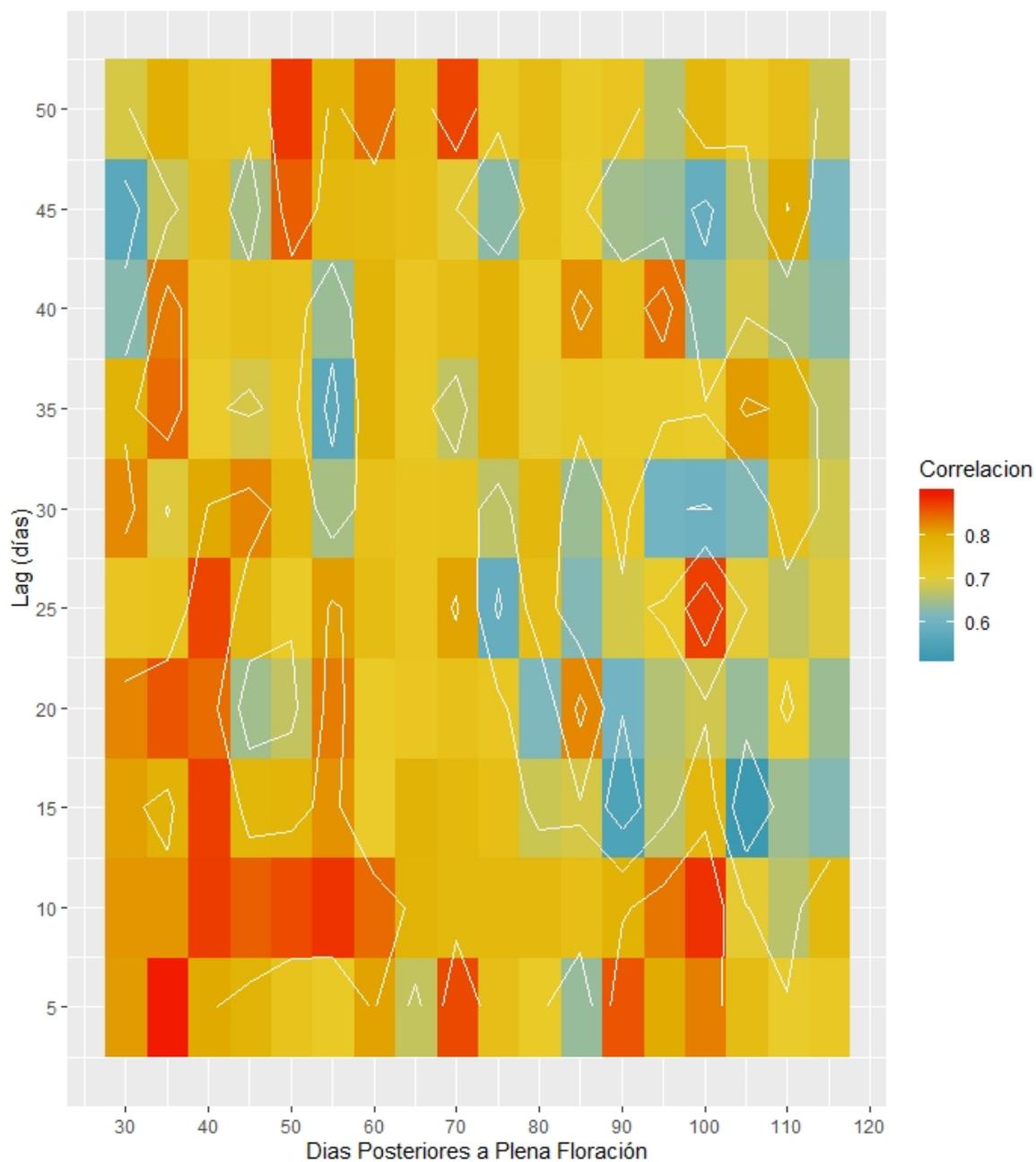


Figura C.2: Correlación Marginal para obs. del mismo Fruto en MODELO 24

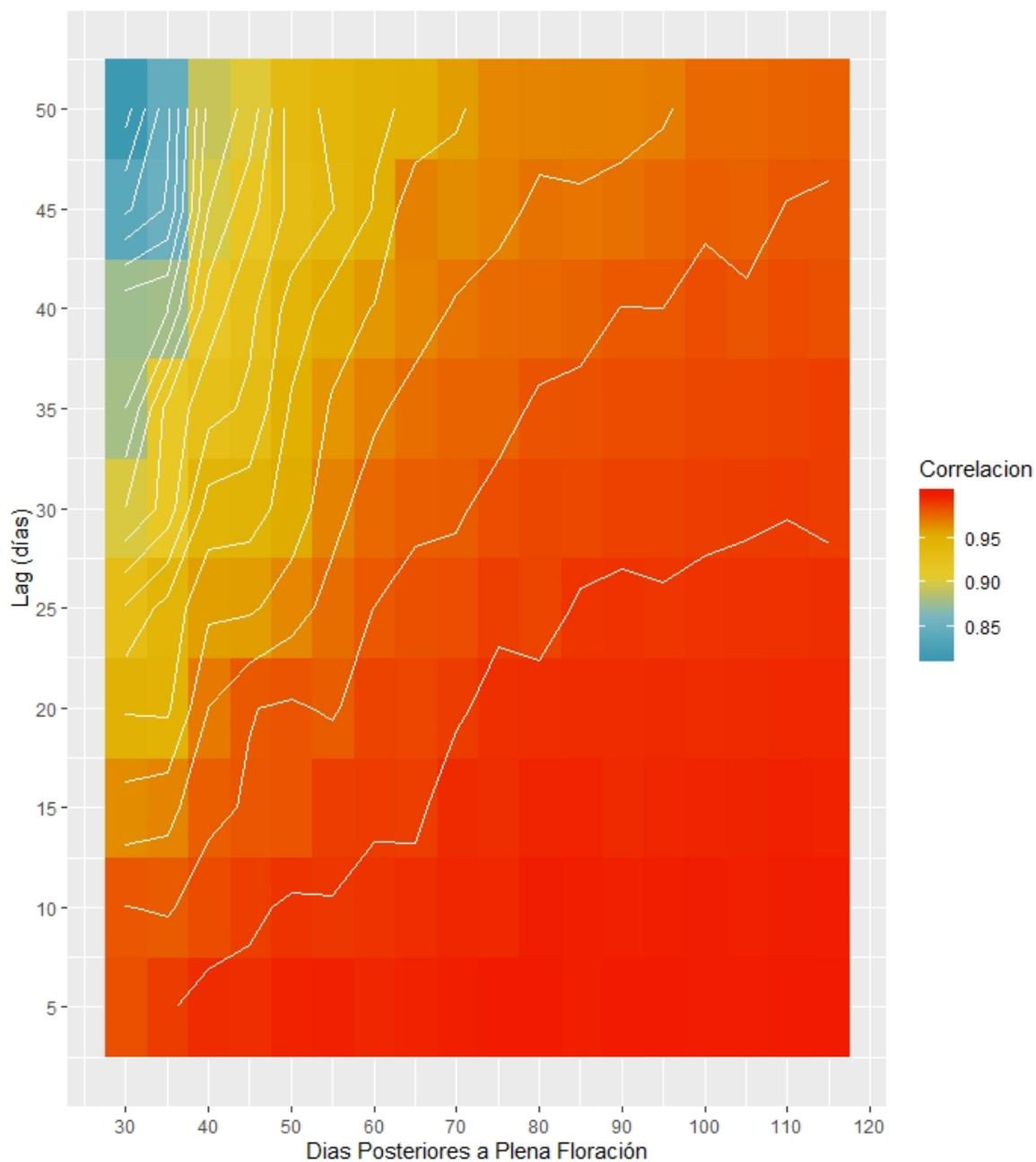


Figura C.3: Correlación Marginal para obs. del mismo Fruto en MODELO 27

Apéndice D

SCRIPTS: PREDICCIÓN DE EFECTOS ALEATORIOS

```
##### VERSION 5.0 #####

# LA MACRO FUNCIONA CON LME Y LMER
# DEFINICION DE LAS FUNCIONES A UTILIZAR EN LA ESTIMACION
# DE LOS BLUPS
# FUNCIONA CON DOS NIVELES DE EFECTOS ALEATORIOS
# LA FUNCION OBTIENE_FIJOS SE DEBE CONFIGURAR MANUALMENTE EN CADA
# UNO DE LOS MODELOS

BLUP_OF<-function(MODELO,VALIDA,NOMBRE_FIJOS,VAR_DEP_CORR="DDF",METODO="BLUP",
                  DELTA_MINIMO=0.20,MAXIMO_CICLO=1,TIPO="LONGI",MODELO_N){

  if(class(MODELO)[1]=="nlme")
  {
    VAR_DEP<-formula(MODELO)[[2]]
  }
  else
  {
    VAR_DEP<-as.character(formula(MODELO)[[2]][[2]])
  }

  # DEFINICION DE EFECTOS ALEATORIOS
  # PARA A_MOD2
  if(MODELO_N=="A_MOD2")
  {
    n_aleat<-length(names(MODELO@cnms))
    NOMBRES<- names(MODELO@cnms)
    nivel_1<-as.matrix(VarCorr(MODELO)[[2]])
    nivel_2<-as.matrix(VarCorr(MODELO)[[1]])
    LISTA<-list(NOMBRE_1=nivel_1,NOMBRE_2=nivel_2)
    names(LISTA)[1]<-"SITIO2"
```

Apéndice D. SCRIPTS: PREDICCIÓN DE EFECTOS ALEATORIOS

```
names(LISTA)[2]<-"FRUTO4"
}

# DEFINICION DE EFECTOS ALEATORIOS
# PARA Z_MOD2
if(MODELO_N=="Z_MOD2")
{
  coef_f<-summary(MODELO)$coefficients$fixed
  aa<-summary(MODELO)$modelStruct
  # DEFINICION DE LA LISTA DEL NRO DE NIVELES, DE EFECTOS ALEATORIOS PARA CADA NIVEL
  # Y LAS MATRICES DE VARIANZA/COVARIANZA DE CADA UNO DE ELLOS
  # NRO Y COEFICIENTES DE LAS MATRICES DE EFECTOS ALEATORIOS
  n_aleat<-length(summary(aa)$reStruct)
  NOMBRES<-names(summary(aa)$reStruct)
  if(n_aleat==1)
  {
    nivel_1<-as.matrix(summary(aa)$reStruct[[1]])
    LISTA<-list(NOMBRE_1=nivel_1)
    names(LISTA)[1]<-NOMBRES[1]
  }
  if(n_aleat==2)
  {
    nivel_1<-as.matrix(summary(aa)$reStruct[[2]])
    nivel_2<-as.matrix(summary(aa)$reStruct[[1]])
    LISTA<-list(NOMBRE_1=nivel_1,NOMBRE_2=nivel_2)
    names(LISTA)[1]<-NOMBRES[2]
    names(LISTA)[2]<-NOMBRES[1]
  }
  if(n_aleat==3)
  {
    nivel_1<-as.matrix(summary(aa)$reStruct[[3]])
    nivel_2<-as.matrix(summary(aa)$reStruct[[2]])
    nivel_3<-as.matrix(summary(aa)$reStruct[[1]])
    LISTA<-list(NOMBRE_1=nivel_1,NOMBRE_2=nivel_2,NOMBRE_3=nivel_3)
    names(LISTA)[1]<-NOMBRES[3]
    names(LISTA)[2]<-NOMBRES[2]
    names(LISTA)[3]<-NOMBRES[1]
  }
}

# DEFINICION DE EFECTOS ALEATORIOS
# PARA Z_MOD9
if(MODELO_N=="Z_MOD9")
{
  coef_f<-summary(MODELO)$coefficients$fixed
  aa<-summary(MODELO)$modelStruct
  # DEFINICION DE LA LISTA DEL NRO DE NIVELES, DE EFECTOS ALEATORIOS PARA CADA NIVEL
  # Y LAS MATRICES DE VARIANZA/COVARIANZA DE CADA UNO DE ELLOS
  # NRO Y COEFICIENTES DE LAS MATRICES DE EFECTOS ALEATORIOS
  n_aleat<-length(summary(aa)$reStruct)
  NOMBRES<-names(summary(aa)$reStruct)
  if(n_aleat==1)
  {
    nivel_1<-as.matrix(summary(aa)$reStruct[[1]])
```

```

LISTA<-list(NOMBRE_1=nivel_1)
names(LISTA)[1]<-NOMBRES[1]
}
if(n_aleat==2)
{
  nivel_1<-as.matrix(summary(aa)$reStruct[[2]])
  nivel_2<-as.matrix(summary(aa)$reStruct[[1]])
  LISTA<-list(NOMBRE_1=nivel_1,NOMBRE_2=nivel_2)
  names(LISTA)[1]<-NOMBRES[2]
  names(LISTA)[2]<-NOMBRES[1]
}
if(n_aleat==3)
{
  nivel_1<-as.matrix(summary(aa)$reStruct[[3]])
  nivel_2<-as.matrix(summary(aa)$reStruct[[2]])
  nivel_3<-as.matrix(summary(aa)$reStruct[[1]])
  LISTA<-list(NOMBRE_1=nivel_1,NOMBRE_2=nivel_2,NOMBRE_3=nivel_3)
  names(LISTA)[1]<-NOMBRES[3]
  names(LISTA)[2]<-NOMBRES[2]
  names(LISTA)[3]<-NOMBRES[1]
}
}

# DEFINICION DE EFECTOS ALEATORIOS
# PARA AB_MOD17
if(MODELO_N=="AB_MOD17")
{
  n_aleat<-2
  NOMBRES<- c("FRUTO4","SITIO2")
  nivel_1<-as.matrix(VarCorr(MODELO)[[5]])
  if(TIPO=="LONGI")
  {
    aa<-as.numeric(VarCorr(MODELO)[[1]])
    bb<-as.numeric(VarCorr(MODELO)[[3]])
    cc<-as.numeric(VarCorr(MODELO)[[2]])
    XX<-matrix(data=0,nrow=3,ncol=3)
    diag(XX)<-c(aa,bb,cc)
  }
  else
  {
    bb<-as.numeric(VarCorr(MODELO)[[3]])
    cc<-as.numeric(VarCorr(MODELO)[[4]])
    XX<-matrix(data=0,nrow=2,ncol=2)
    diag(XX)<-c(bb,cc)
  }
  nivel_2<-XX
  LISTA<-list(NOMBRE_1=nivel_1,NOMBRE_2=nivel_2)
  names(LISTA)[1]<-"SITIO2"
  names(LISTA)[2]<-"FRUTO4"
}

# DEFINICION DE EFECTOS ALEATORIOS
# PARA Z_MOD15
if(MODELO_N=="Z_MOD15")

```

Apéndice D. SCRIPTS: PREDICCIÓN DE EFECTOS ALEATORIOS

```
{
  aa<-summary(MODELO)$modelStruct
  #PARA MEDIDAS A NIVEL DE SITIO
  bb2<-summary(aa)$reStruct[2]
  #PARA MEDIDAS A NIVEL DE FRUTO
  bb<-summary(aa)$reStruct[1]
  # DEFINICION DE LA LISTA DEL NRO DE NIVELES, DE EFECTOS ALEATORIOS PARA CADA NIVEL
  # Y LAS MATRICES DE VARIANZA/COVARIANZA DE CADA UNO DE ELLOS
  # NRO Y COEFICIENTES DE LAS MATRICES DE EFECTOS ALEATORIOS
  n_aleat<-length(summary(aa)$reStruct)
  NOMBRES<-names(summary(aa)$reStruct)
  nivel_1<-as.matrix(bb2[[1]])
  if(TIPO=="LONGI")
  {
    nivel_2<-as.matrix(bb$FRUTO4[[1]])
  }
  else
  {
    nivel_2<-as.matrix(bb$FRUTO4[[2]])
  }
  LISTA<-list(NOMBRE_1=nivel_1,NOMBRE_2=nivel_2)
  names(LISTA)[1]<-NOMBRES[2]
  names(LISTA)[2]<-NOMBRES[1]
}

# DEFINICION DE EFECTOS ALEATORIOS
# PARA AB_MOD22
if(MODELO_N=="AB_MOD22")
{
  n_aleat<-1
  NOMBRES<- c("FRUTO4")
  if(TIPO=="LONGI")
  {
    aa<-as.numeric(VarCorr(MODELO)[[1]])
    bb<-as.numeric(VarCorr(MODELO)[[3]])
    cc<-as.numeric(VarCorr(MODELO)[[2]])
    XX<-matrix(data=0,nrow=3,ncol=3)
    diag(XX)<-c(aa,bb,cc)
  }
  nivel_1<-XX
  LISTA<-list(NOMBRE_1=nivel_1)
  names(LISTA)[1]<-"FRUTO4"
}

# PREPARAMOS LAS MATRICES PARA PODER OBTENER LOS BLUPs
# NRO DE ELEMENTOS EN
MAX_NIVEL<-length(unique(VALIDA[,as.character(names(LISTA)[1])]))
MANTIENE<-1
CICLO<-1
# OBTIENE LA LONGITUD TOTAL DE LOS EFECTOS ALEATORIOS PARA EL PRESENTE SITIO
LONGITUD<-LARGO_UU(VALIDA,LISTA,1)
UU<-diag(x=0,nrow=LONGITUD,ncol=MAX_NIVEL)
DELTA1<-diag(x=0,nrow=LONGITUD,ncol=MAXIMO_CICLO)
DELTA2<-diag(x=0,nrow=LONGITUD,ncol=MAXIMO_CICLO)
```

Apéndice D. SCRIPTS: PREDICCIÓN DE EFECTOS ALEATORIOS

```
# INICIA EL CICLO PARA OBTENER LOS BLUPS
cat("Macro_Blups_Ver_4.0","\n")
cat("Metodo:...",METODO,"\n")
while(MANTIENE==1){
  for(i in 1:MAX_NIVEL) {
    # OBTENEMOS LA MATRIZ D PARA EL SITIO2 i
    # ESTE CORRE
    D<-OBTIENE_D(LISTA,VALIDA,i)
    # OBTENEMOS LA MATRIZ Z PARA EL SITIO2 i
    ZQ<-OBTIENE_Z(VALIDA,LISTA,DERIVADAS,UU[,i],i,MODELO,NOMBRE_FIJOS)
    # OBTENEMOS LOS VALORES AJUSTADOS PARA EL SITIO2 i
    PP<-AJUSTADO(VALIDA,LISTA,EXPRESION,DERIVADAS,UU[,i],i,MODELO,NOMBRE_FIJOS)
    DIF<-PP[,as.character(VAR_DEP)]-PP[,"Pred"]
    # OBTENEMOS LA MATRIZ R PARA EL SITIO2 i
    RR<-OBTIENE_R(LISTA,MODELO,PP,VAR_DEP_CORR="DDF",i)
    # OPERACIONES MATRICIALES
    TEMPO1 <- D%*%t(ZQ)
    TEMPO2 <- ZQ%*%D
    TEMPO3<- TEMPO2%*%t(ZQ)
    TEMPO4<-TEMPO3+RR
    TEMPO5<-solve(TEMPO4)
    TEMPO6<-TEMPO1%*%TEMPO5
    TEMPO7<-ZQ%*%UU[,i]
    DIF2<-DIF+TEMPO7
    if (METODO=="BLUP") UV<-TEMPO6%*%DIF2
    if (METODO=="ZERO") UV<-TEMPO6%*%DIF
    DELTA1[,CICLO]<-UV-UU[,i]
    DELTA2[,CICLO]<-(UV-UU[,i])/UV*100
    UU[,i]<-UV
  }
  ### ACLARACIONES PARA SALIDA POR CICLOS
  U_FINAL<-UU

  for(i in 1:MAX_NIVEL) {
    PP_F<-AJUSTADO(VALIDA,LISTA,EXPRESION,DERIVADAS,U_FINAL[,i],i,MODELO,NOMBRE_FIJOS)
    if(i==1) PP_FINAL<-PP_F
    if(i>1) PP_FINAL<-rbind(PP_FINAL,PP_F)
  }
  PP_FINAL$error<-PP_FINAL[,as.character(VAR_DEP)]-PP_FINAL[,"Pred"]
  MM<-mean(PP_FINAL$error)
  VV<-var(PP_FINAL$error)
  cat("Ciclo:...",CICLO,"Media del Error:...",MM,"Varianza del Error:...",VV,"\n")
  if(CICLO==MAXIMO_CICLO) break
  if(max(abs(DELTA2[,CICLO]))<DELTA_MINIMO) break
  if(METODO=="ZERO") break
  CICLO<-CICLO+1
}
PP_FINAL$METODO1<-METODO
PP_FINAL
}

Z_2<-function(DERIVADAS,EVALUA,DATOS,MODELO,LISTA,NOMBRE_FIJOS,ALEATORIOS,NRO)
{
  M_FIJOS<-OBTIENE_FIJOS(MODELO,NOMBRE_FIJOS,DATOS)
```

Apéndice D. SCRIPTS: PREDICCIÓN DE EFECTOS ALEATORIOS

```
EXPRESION<-DERIVADAS[[EVALUA]]
UU_LISTOS<-OBTIENE_UU(VVALIDA ,DERIVADAS ,DATOS ,LISTA ,ALEATORIOS ,NRO)
d <- data.frame((cbind(DATOS ,M_FIJOS ,UU_LISTOS)))
Z2<-eval(EXPRESION ,d)
Z2
}

AJUSTADO<-function(VVALIDA ,LISTA ,EXPRESION ,DERIVADAS ,ALEATORIOS ,NRO ,MODELO ,NOMBRE_FIJOS)
{
  NOM_UU<-names(DERIVADAS)
  ANCHO_UU<-length(DERIVADAS)
  NOMBRE_1<-as.character(names(LISTA)[1])
  V_N1<-unique(VVALIDA[,NOMBRE_1])
  TAMIZ1<-VVALIDA[VVALIDA[,NOMBRE_1]==V_N1[NRO],]
  LARGO_UU<-dim(TAMIZ1)[1]
  UU_LISTOS<-matrix(0,LARGO_UU,ANCHO_UU)
  M_FIJOS<-matrix(0,LARGO_UU,length(NOMBRE_FIJOS))
  colnames(UU_LISTOS)<-NOM_UU
  # CREA UN VECTOR LLAMADO UU_LISTOS
  for(i in 1:LARGO_UU) {
    # ESTO SOLO FUNCIONA CON UNO O DOS NIVELES
    M_FIJOS[i,]<-as.numeric(OBTIENE_FIJOS(MODELO ,NOMBRE_FIJOS ,TAMIZ1[i,]))
    UU_LISTOS[i,]<-OBTIENE_UU(VVALIDA ,DERIVADAS ,TAMIZ1[i,] ,LISTA ,ALEATORIOS ,NRO)
  }
  M_FIJOS<-as.data.frame(M_FIJOS)
  names(M_FIJOS)<-NOMBRE_FIJOS
  d <- data.frame((cbind(TAMIZ1 ,M_FIJOS ,UU_LISTOS)))
  d$Pred<-eval(EXPRESION ,d)
  d
}

OBTIENE_UU<-function(VVALIDA ,DERIVADAS ,DATOS ,LISTA ,ALEATORIOS ,NRO ,TOTAL=FALSE){
  NIVELES=length(LISTA)
  ### DEFINE VECTOR DE EFECTOS ALEATORIOS
  NOM_UU<-names(DERIVADAS)
  LARGO_UU<-length(DERIVADAS)
  UU_LIST<-matrix(0,1,LARGO_UU)
  colnames(UU_LIST)<-NOM_UU
  ##### DEFINICIONES DEL NIVEL 1
  NOMBRE_1<-as.character(names(LISTA)[1])
  ANCHO_1<-dim(LISTA[[1]])[1]
  V_N1<-unique(VVALIDA[,NOMBRE_1])
  TAMIZ1<-VVALIDA[VVALIDA[,NOMBRE_1]==V_N1[NRO],]
  ##### ASIGNA LOS EFECTOS ALEATORIOS DEL NIVEL 1 AL VECTOR
  for(i in 1:ANCHO_1) {
    UU_LIST[i]<-ALEATORIOS[i]
  }
  if(NIVELES>1) {
    ##### DEFINICIONES DEL NIVEL 2
    NOMBRE_2<-as.character(names(LISTA)[2])
    ANCHO_2<-dim(LISTA[[2]])[1]
    V_N2<-as.matrix(unique(TAMIZ1[,NOMBRE_2]))
    VALOR_N2<-as.character(DATOS[,NOMBRE_2])
  }
}
```

Apéndice D. SCRIPTS: PREDICCIÓN DE EFECTOS ALEATORIOS

```
N2<-length(unique(TAMIZ1[,NOMBRE_2]))
##### BUSCA LA POSICION DEL EFECTO ALEATORIO ACTUAL EN LOS PRESENTES EN ESE NIVEL
POS_N2<-which(V_N2==VALOR_N2)
for(i in 1:ANCHO_2) {
  POS_LIST<-ANCHO_1+i
  POS_ALEA<-ANCHO_1+ANCHO_2*(POS_N2-1)+i
  UU_LIST[POS_LIST]<-ALEATORIOS[POS_ALEA]
}
if(NIVELES>2) {
  ##### DEFINICIONES DEL NIVEL 3
  NOMBRE_3<-as.character(names(LISTA)[3])
  ANCHO_3<-dim(LISTA[[3]])[1]
  # FRUTOS EN EL TAMANO
  N3<-0
  for(i in 1:N2) {
    TAMIZ2<-TAMIZ1[TAMIZ1[,NOMBRE_2]==V_N2[i],]
    # FRUTOS EN EL TAMANO
    V_N3<-as.matrix(unique(TAMIZ2[,NOMBRE_3]))
    N3_SUB<-length(V_N3)
    VALOR_N3<-as.character(DATOS[,NOMBRE_3])
    POS_N3<-match(VALOR_N3,V_N3)
    ifelse(is.na(POS_N3),N3<-N3+N3_SUB,break)
  }
  N3<-N3+POS_N3
  for(i in 1:ANCHO_3) {
    POS_LIST<-ANCHO_1+ANCHO_2+i
    POS_ALEA<-ANCHO_1+ANCHO_2+ANCHO_3*(N3-1) +i
    UU_LIST[POS_LIST]<-ALEATORIOS[POS_ALEA]
  }
}
}
UU_LIST
}

### ESTIMA D #####
OBTIENE_D<-function(LISTA,VALIDA,NRO) {
  NIVELES=length(LISTA)
  # OBTIENE EL BLOQUE DEL PRIMER NIVEL
  BLOQUE_1<-LISTA[[1]]
  NOMBRE_1<-as.character(names(LISTA)[1])
  ANCHO_1<-dim(LISTA[[1]])[1]
  LONGITUD<-ANCHO_1
  RETORNA<-BLOQUE_1
  if(NIVELES>1) {
    # NIVEL 2 DE EFECTOS ALEATORIOS
    NOMBRE_2<-as.character(names(LISTA)[2])
    ANCHO_2<-dim(LISTA[[2]])[1]
    V_N1<-unique(VALIDA[,NOMBRE_1])
    TAMIZ1<-VALIDA[VALIDA[,NOMBRE_1]==V_N1[NRO],]
    V_N2<-unique(TAMIZ1[,NOMBRE_2])
    N2<-length(V_N2)
    LONGITUD<-ANCHO_1+ANCHO_2*N2
    BLOQUE_2<-REPLICA_M(LISTA[[2]],N2)
```

Apéndice D. SCRIPTS: PREDICCIÓN DE EFECTOS ALEATORIOS

```
RETORNA<-as.matrix(bdiag(BLOQUE_1,BLOQUE_2))
if(NIVELES>2) {
  # NIVEL 3 DE EFECTOS ALEATORIOS
  NOMBRE_3<-as.character(names(LISTA)[3])
  ANCHO_3<-dim(LISTA[[3]])[1]
  N3<-0
  for(i in 1:N2) {
    TAMIZ2<-TAMIZ1[TAMIZ1[,NOMBRE_2]==V_N2[i],]
    # FRUTOS EN EL TAMANO
    V_N3<-unique(TAMIZ2[,NOMBRE_3])
    N3_SUB<-length(V_N3)
    N3<-N3+N3_SUB
  }
  LONGITUD<-ANCHO_1+ANCHO_2*N2+ANCHO_3*N3
  BLOQUE_3<-REPLICA_M(LISTA[[3]],N3)
  RETORNA<-as.matrix(bdiag(BLOQUE_1,BLOQUE_2,BLOQUE_3))
}
}
RETORNA
}

##### ESTIMA Z #####

OBTIENE_Z<-function (VALIDA,LISTA,DERIVADAS,ALEATORIOS,NRO,MODELO,NOMBRE_FIJOS) {
  NIVELES=length(LISTA)
  # OBTIENE EL BLOQUE DEL PRIMER NIVEL
  NOMBRE_1<-as.character(names(LISTA)[1])
  ### ANCHO DEL BLOQUE
  ANCHO_1<-dim(LISTA[[1]])[1]
  V_N1<-unique(VALIDA[,NOMBRE_1])
  TAMIZ1<-VALIDA[VALIDA[,NOMBRE_1]==V_N1[NRO],]
  ### LONGITUD DEL BLOQUE
  LARGO_1<-dim(TAMIZ1)[1]
  ### CONSTRUYE EL PRIMER BLOQUE
  BLOQUE_1<-matrix(0,LARGO_1,ANCHO_1)
  CORRE<-1
  for(i in 1:ANCHO_1) {
    for(ii in 1:LARGO_1) {
      DATOS<-TAMIZ1[ii,]
      BLOQUE_1[ii,i]<-Z_2(DERIVADAS,CORRE,DATOS,MODELO,LISTA,NOMBRE_FIJOS,ALEATORIOS,NRO)
    }
    CORRE<-CORRE+1
  }
  RETORNA<-BLOQUE_1
  if(NIVELES>1) {
    # NIVEL 2 DE EFECTOS ALEATORIOS
    NOMBRE_2<-as.character(names(LISTA)[2])
    ANCHO_2<-dim(LISTA[[2]])[1]
    V_N2<-unique(TAMIZ1[,NOMBRE_2])
    N2<-length(V_N2)
    ### CONSTRUYE EL SEGUNDO BLOQUE
    for(i in 1:N2) {
```

Apéndice D. SCRIPTS: PREDICCIÓN DE EFECTOS ALEATORIOS

```

CORRE<-ANCHO_1+1
TAMIZ2<-TAMIZ1[TAMIZ1[,NOMBRE_2]==V_N2[i],]
### LONGITUD DEL BLOQUE
LARGO_2<-dim(TAMIZ2)[1]
SUB_BLOQUE_2<-matrix(0,LARGO_2,ANCHO_2)
for(ii in 1:ANCHO_2) {
  for(iii in 1:LARGO_2) {
    DATOS<-TAMIZ2[iii,]
    SUB_BLOQUE_2[iii,ii]<-Z_2(DERIVADAS,CORRE,DATOS,MODELO,LISTA,NOMBRE_FIJOS,ALEATORIOS,NRO)
  }
  CORRE<-CORRE+1
}
if(i==1) BLOQUE_2<-SUB_BLOQUE_2
else BLOQUE_2<-as.matrix(bdiag(BLOQUE_2,SUB_BLOQUE_2))
}
RETORNA<-as.matrix(cbind(BLOQUE_1,BLOQUE_2))
if(NIVELES>2) {
  # NIVEL 3 DE EFECTOS ALEATORIOS
  NOMBRE_3<-as.character(names(LISTA)[3])
  ANCHO_3<-dim(LISTA[[3]])[1]
  ### CONSTRUYE EL TERCER BLOQUE
  for(i in 1:N2) {
    NOMBRE_3<-as.character(names(LISTA)[3])
    ANCHO_3<-dim(LISTA[[3]])[1]
    TAMIZ2<-TAMIZ1[TAMIZ1[,NOMBRE_2]==V_N2[i],]
    V_N3<-unique(TAMIZ2[,NOMBRE_3])
    N3<-length(V_N3)
    for(v in 1:N3) {
      TAMIZ3<-TAMIZ2[TAMIZ2[,NOMBRE_3]==V_N3[v],]
      ### LONGITUD DEL BLOQUE
      LARGO_3<-dim(TAMIZ3)[1]
      S_SUB_BLOQUE_3<-matrix(0,LARGO_3,ANCHO_3)
      CORRE<-ANCHO_1+ANCHO_2+1
      for(ii in 1:ANCHO_3) {
        for(iii in 1:LARGO_3) {
          DATOS<-TAMIZ3[iii,]
          S_SUB_BLOQUE_3[iii,ii]<-Z_2(DERIVADAS,CORRE,DATOS,MODELO,LISTA,NOMBRE_FIJOS,ALEATORIOS,NRO)
        }
        CORRE<-CORRE+1
      }
      if(v==1) S_BLOQUE_3<-S_SUB_BLOQUE_3
      else S_BLOQUE_3<-as.matrix(bdiag(S_BLOQUE_3,S_SUB_BLOQUE_3))
    }
    if(i==1) BLOQUE_3<-S_BLOQUE_3
    else BLOQUE_3<-as.matrix(bdiag(BLOQUE_3,S_BLOQUE_3))
  }
  RETORNA<-as.matrix(cbind(BLOQUE_1,BLOQUE_2,BLOQUE_3))
}
}
RETORNA
}

### REPLICA MATRICES DIAGONIZANDOLES
REPLICA_M<-function(MATRIZ,n) {

```

```

ES<-MATRIZ
if(n>1) {
  LIMITE<-n-1
  for(i in 1:LIMITE) {
    ES<-as.matrix(bdiag(MATRIZ,ES))
  }
}
ES
}

#### FUNCION QUE CALCULA LA LONGITUD DEL VECTOR DE BLUPS ESTIMADOS
LARGO_UU<-function(VALIDA,LISTA,NRO) {
  NIVELES=length(LISTA)
  # OBTIENE EL BLOQUE DEL PRIMER NIVEL
  NOMBRE_1<-as.character(names(LISTA)[1])
  ANCHO_1<-dim(LISTA[[1]])[1]
  V_N1<-unique(VALIDA[,NOMBRE_1])
  TAMIZ1<-VALIDA[VALIDA[,NOMBRE_1]==V_N1[NRO],]
  LONGITUD<-ANCHO_1
  if(NIVELES>1) {
    # NIVEL 2 DE EFECTOS ALEATORIOS
    NOMBRE_2<-as.character(names(LISTA)[2])
    ANCHO_2<-dim(LISTA[[2]])[1]
    V_N2<-unique(TAMIZ1[,NOMBRE_2])
    N2<-length(V_N2)
    LONGITUD<-ANCHO_1+ANCHO_2*N2
    if(NIVELES>2) {
      # NIVEL 3 DE EFECTOS ALEATORIOS
      NOMBRE_3<-as.character(names(LISTA)[3])
      ANCHO_3<-dim(LISTA[[3]])[1]
      N3<-0
      for(i in 1:N2) {
        TAMIZ2<-TAMIZ1[TAMIZ1[,NOMBRE_2]==V_N2[i],]
        # FRUTOS EN EL TAMANO
        V_N3<-unique(TAMIZ2[,NOMBRE_3])
        N3_SUB<-length(V_N3)
        N3<-N3+N3_SUB
      }
      LONGITUD<-ANCHO_1+ANCHO_2*N2+ANCHO_3*N3
    }
  }
  LONGITUD
}

OBTIENE_R<-function(LISTA,MODELO,PP,VAR_DEP_CORR="DDF",NRO) {
  NIVELES=length(LISTA)
  # VER SI EXISTE ESTRUCTURA DE: CORRELACION - DETECTAR EL COEF Y COVARIABLE
  # / VARIANZAS CON VARIDENT - DETECTAR LA COVARIABLE
  # / VARIANZAS CON FUNCION - DETECTAR EL COEF Y LA COVARIABLE

  # OBTIENE EL PRIMER NIVEL
  NOMBRE_1<-as.character(names(LISTA)[1])
  V_N1<-unique(PP[,NOMBRE_1])
  TAMIZ1<-PP[PP[,NOMBRE_1]==V_N1[NRO],]

```

```

TAMANO_1<-dim(TAMIZ1)[1]
if(NIVELES==1) {
  VAR_POND<-OBTIENE_TVAR(MODELO,TAMIZ1)
  VAR_CORR<-OBTIENE_CORR(MODELO,TAMIZ1,VAR_DEP_CORR)
  RR<-VAR_POND%*%VAR_CORR%*%VAR_POND
}

if(NIVELES==2) {
  # NIVEL 2 DE EFECTOS ALEATORIOS
  NOMBRE_2<-as.character(names(LISTA)[2])
  ANCHO_2<-dim(LISTA[[2]])[1]
  V_N2<-unique(TAMIZ1[,NOMBRE_2])
  N2<-length(V_N2)
  for(v in 1:N2) {
    TAMIZ2<-TAMIZ1[TAMIZ1[,NOMBRE_2]==V_N2[v],]
    # ESTE MATRIZ TIENE ELEMENTOS EN LA DIAGONAL
    # SE OBTIENEN LAS PONDERACIONES POR GRUPOS Y
    # Y POR FUNCIONES DE POTENCIA Y EXPONENTE
    VAR_POND<-OBTIENE_TVAR(MODELO,TAMIZ2)
    VAR_CORR<-OBTIENE_CORR(MODELO,TAMIZ2,VAR_DEP_CORR)
    S_RR<-VAR_POND%*%VAR_CORR%*%VAR_POND
    if(v==1) RR<-S_RR
    else RR<-as.matrix(bdiag(RR,S_RR))
  }
}

if(NIVELES==3) {
  # NIVEL 3 DE EFECTOS ALEATORIOS
  NOMBRE_2<-as.character(names(LISTA)[2])
  ANCHO_2<-dim(LISTA[[2]])[1]
  NOMBRE_3<-as.character(names(LISTA)[3])
  ANCHO_3<-dim(LISTA[[3]])[1]
  V_N2<-unique(TAMIZ1[,NOMBRE_2])
  N2<-length(V_N2)
  for(v in 1:N2) {
    TAMIZ2<-TAMIZ1[TAMIZ1[,NOMBRE_2]==V_N2[v],]
    V_N3<-unique(TAMIZ2[,NOMBRE_3])
    N3<-length(V_N3)
    for(w in 1:N3) {
      TAMIZ3<-TAMIZ2[TAMIZ2[,NOMBRE_3]==V_N3[w],]
      VAR_POND<-OBTIENE_TVAR(MODELO,TAMIZ2)
      VAR_CORR<-OBTIENE_CORR(MODELO,TAMIZ2,VAR_DEP_CORR)
      S_RR<-VAR_POND%*%VAR_CORR%*%VAR_POND
      if(w==1) S_RR<-S_S_RR
      else S_RR<-as.matrix(bdiag(S_RR,S_S_RR))
    }
    if(v==1) RR<-S_RR
    else RR<-as.matrix(bdiag(RR,S_RR))
  }
}
SIGMA<-summary(MODELO)$sigma
RR<-SIGMA*SIGMA*RR
RR
}

```

```

OBTIENE_TVAR<-function(MODELO,DATA){
  # esta es la forma de leer el tipo de estructura de varianza y correlacion del error
  RETORNA<-rep(1,dim(DATA)[1])
  if(class(MODELO)[1]=="nlme"){
    T_VAR<-class(summary(MODELO)$modelStruct$varStruct)[1]
    if(T_VAR=="varIdent"){
      {
        # con esto encuentro la covariable para varId
        TEMPO1<-summary(MODELO)$modelStruct$varStruct
        (TEMPO2<-attr(TEMPO1,"formula"))
        COVA<-as.character(TEMPO2[[2]])[3]
        GRUPOS<-attr(TEMPO1,"groupNames")
        MARCO_VAR<-matrix(0,length(GRUPOS),2)
        MARCO_VAR<-as.data.frame(MARCO_VAR)
        MARCO_VAR[1,2]<-1
        MARCO_VAR[1,1]<-GRUPOS[1]
        for(ww in 2:length(GRUPOS)) {
          MARCO_VAR[ww,1]<-GRUPOS[ww]
          MARCO_VAR[ww,2]<-exp(TEMPO1[ww-1])
        }
        DATA<-as.data.frame(DATA)
        RETORNA<-merge(DATA, MARCO_VAR, by.x = COVA, by.y = "V1", sort=F,all.x=T,all.y=F)[,"V2"]
      }
    }
    if(T_VAR=="varPower"){
      {
        # con esto encuentro la covariable para varPower
        TEMPO1<-summary(MODELO)$modelStruct$varStruct
        TEMPO2<-attr(TEMPO1,"formula")
        COVA<-as.character(TEMPO2[[2]])[1]
        COVA2<-ifelse(COVA=="fitted","Pred",COVA)
        COEF<-as.numeric(aa$varStruct)
        RETORNA<-(DATA[,COVA2]^COEF)
      }
    }
    if(T_VAR=="varExp"){
      {
        # con esto encuentro la covariable para varPower
        TEMPO1<-summary(MODELO)$modelStruct$varStruct
        TEMPO2<-attr(TEMPO1,"formula")
        COVA<-as.character(TEMPO2[[2]])[1]
        COVA2<-ifelse(COVA=="fitted","Pred",COVA)
        COEF<-as.numeric(aa$varStruct)
        RETORNA<-exp(DATA[,COVA2]*COEF)
      }
    }
    RETORNA2<-diag(length(RETORNA))
    diag(RETORNA2)<-RETORNA
    RETORNA2
  }
}

OBTIENE_CORR<-function(MODELO,DATA,VAR_DEP_CORR){
  RETORNA<-rep(1,dim(DATA)[1])

```

```

RETORNA2<-diag(length(RETORNA))
diag(RETORNA2)<-RETORNA
if(class(MODELO)[1]=="nlme"){
  T_CORR<-class(summary(MODELO)$modelStruct$corStruct)[1]
  # OBTIENE LA CORRELACION SI EXISTE Y CONSTRUYE LA MATRIZ
  if (T_CORR=="corAR1") {
    CORR_P<-MODELO$modelStruct$corStruct[1]
    CORR<-exp(CORR_P)
    CORR <- c((CORR - 1)/(CORR + 1))
    COEF_CORR<-CORR
    CORR <- corAR1(COEF_CORR)
    #M_CORR<-corMatrix(CORR, covariate = 1:4)
    # ARMO LA MATRIZ C DE LA CORRELACION
    LARGO<-length(DATA[,VAR_DEP_CORR])
    VECTOR_1<- DATA[1,VAR_DEP_CORR]
    UNOS<-rep(1,LARGO)
    CONSTANTE<-UNOS%*%t(VECTOR_1)
    VECTOR_2<- ((DATA[,VAR_DEP_CORR]-CONSTANTE)/7)+1
    MARCO_CORR<-matrix(0,LARGO,LARGO)
    for(i in 1:LARGO) {
      for(ii in 1:LARGO) {
        MARCO_CORR[i,ii]<-COEF_CORR^(abs(VECTOR_2[i]-VECTOR_2[ii]))
      }
    }
    RETORNA2<-MARCO_CORR
  }

  if (T_CORR=="corCAR1") {
    CORR_P<-MODELO$modelStruct$corStruct[1]
    CORR<-exp(CORR_P)
    CORR <- c((CORR )/(CORR + 1))
    COEF_CORR<-CORR
    CORR <- corCAR1(COEF_CORR)
    #M_CORR<-corMatrix(CORR, covariate = 1:4)
    # ARMO LA MATRIZ C DE LA CORRELACION
    LARGO<-length(DATA[,VAR_DEP_CORR])
    VECTOR_1<- DATA[1,VAR_DEP_CORR]
    UNOS<-rep(1,LARGO)
    CONSTANTE<-UNOS%*%t(VECTOR_1)
    VECTOR_2<- ((DATA[,VAR_DEP_CORR]-CONSTANTE))+1
    MARCO_CORR<-matrix(0,LARGO,LARGO)
    # SE ARMA COVA_R PARA QUE OBTENGA LA VARIABLE DEPENDIENTE TRANSFORMADA
    attach(DATA)
    TEMPO1<-MODELO$modelStruct$corStruct
    TEMPO2<-attr(TEMPO1,"formula")
    TEMPO3<-TEMPO2[[2]][2]
    TEMPO4<-languageEl(TEMPO3, 1)
    COVA_R<-eval(TEMPO4)
    detach(DATA)
    for(i in 1:LARGO) {
      for(ii in 1:LARGO) {
        MARCO_CORR[i,ii]<-COEF_CORR^((abs(VECTOR_2[i]-VECTOR_2[ii])*COVA_R))
      }
    }
  }
}

```

Apéndice D. SCRIPTS: PREDICCIÓN DE EFECTOS ALEATORIOS

```
    RETORNA2<-MARCO_CORR
  }
}
RETORNA2
}

OBT_Z<-function(DATA,DERIVADAS) {
  ANCHO<-length(DERIVADAS)
  LARGO<-dim(DATA)[1]
  MARCO<-matrix(0,LARGO,ANCHO)
  for(i in 1:ANCHO) {
    EXPRESION<-DERIVADAS[[i]]
    MARCO[,i]<-eval(EXPRESION,DATA)
  }
  MARCO
}

OBT_B<-function(DATA,DERIVADAS) {
  ANCHO<-length(DERIVADAS)
  LARGO<-dim(DATA)[1]
  MARCO<-matrix(0,LARGO,ANCHO)
  for(i in 1:ANCHO) {
    NOMBRE<-names(DERIVADAS)[i]
    MARCO[,i]<-DATA[,NOMBRE]
  }
  MARCO
}

OBT_COR<-function(MODELO,DIF_DDF){
  APLICA<-rep(0,length(DIF_DDF))
  if(class(MODELO)[1]=="nlme"){
    T_CORR<-class(summary(MODELO)$modelStruct$corStruct)[1]
    if (T_CORR=="corAR1") {
      CORR_P<-MODELO$modelStruct$corStruct[1]
      CORR<-exp(CORR_P)
      CORR <- c((CORR - 1)/(CORR + 1))
      COEF_CORR<-CORR
      APLICA<-COEF_CORR^(DIF_DDF/7)
    }
    # AHORA HAY QUE CONSIDERAR UNA FUNCION QUE PERMITA OBTENER UNA FUNCION
    if (T_CORR=="corCAR1") {
      CORR_P<-MODELO$modelStruct$corStruct[1]
      CORR<-exp(CORR_P)
      CORR <- c((CORR )/(CORR + 1))
      COEF_CORR<-CORR
      CORR <- corCAR1(COEF_CORR)
      APLICA<-COEF_CORR^(log(DIF_DDF))
    }
  }
  APLICA
}

##### FUNCION NLMER
```

Apéndice D. SCRIPTS: PREDICCIÓN DE EFECTOS ALEATORIOS

```
#### EFECTOS FIJOS DE TAMANO PARA 3 TAMANOS
OBTIENE_FIJOS<-function(MODELO,NOMBRE_FIJOS,DATOS,MODELO_N)
{
  ### ACA LE PONEMOS LA VARIABLE DE EFECTO FIJO MANUALMENTE
  ### DEFINIMOS PARA A_MOD2
  if(MODELO_N=="A_MOD2")
  {
    FIJOS<-fixef(MODELO)
    VAR_CLASS<-"TAMA"
    if(is.null(VAR_CLASS))
    {
      RETORNA<-as.data.frame(t(FIJOS))
    }
    else
    {
      if(DATOS[,VAR_CLASS]==1)
      {
        A<-FIJOS[1]
        B<-FIJOS[4]
        C<-FIJOS[7]
      }
      if(DATOS[,VAR_CLASS]==2)
      {
        A<-FIJOS[1]+FIJOS[2]
        B<-FIJOS[4]+FIJOS[5]
        C<-FIJOS[7]+FIJOS[8]
      }
      if(DATOS[,VAR_CLASS]==3)
      {
        A<-FIJOS[1]+FIJOS[3]
        B<-FIJOS[4]+FIJOS[6]
        C<-FIJOS[7]+FIJOS[9]
      }
      AGRUPA<-c(A,B,C)
      RETORNA<-as.data.frame(t(AGRUPA))
    }
  }
  ### DEFINIMOS PARA Z_MOD2
  if(MODELO_N=="Z_MOD2")
  {
    FIJOS<-summary(MODELO)$coefficients$fixed
    VAR_CLASS<-names((summary(MODELO))$contrasts[1])
    if(is.null(VAR_CLASS))
    {
      RETORNA<-as.data.frame(t(FIJOS))
    }
    else
    {
      if(DATOS[,VAR_CLASS]==1)
      {
        A<-FIJOS[1]
        B<-FIJOS[4]
        C<-FIJOS[7]
      }
    }
  }
}
```

```

if (DATOS[,VAR_CLASS]==2)
{
  A<-FIJOS[1]+FIJOS[2]
  B<-FIJOS[4]+FIJOS[5]
  C<-FIJOS[7]+FIJOS[8]
}
if (DATOS[,VAR_CLASS]==3)
{
  A<-FIJOS[1]+FIJOS[3]
  B<-FIJOS[4]+FIJOS[6]
  C<-FIJOS[7]+FIJOS[9]
}
AGRUPA<-c(A,B,C)
RETORNA<-as.data.frame(t(AGRUPA))
}
}
### DEFINIMOS PARA Z_MOD9
if (MODELO_N=="Z_MOD9")
{
  FIJOS<-summary(MODELO)$coefficients$fixed
  VAR_CLASS<-names((summary(MODELO))$contrasts[1])
  if (is.null(VAR_CLASS))
  {
    RETORNA<-as.data.frame(t(FIJOS))
  }
  else
  {
    if (DATOS[,VAR_CLASS]==1)
    {
      A<-FIJOS[1]
      B<-FIJOS[4]
      C<-FIJOS[7]
    }
    if (DATOS[,VAR_CLASS]==2)
    {
      A<-FIJOS[1]+FIJOS[2]
      B<-FIJOS[4]+FIJOS[5]
      C<-FIJOS[7]+FIJOS[8]
    }
    if (DATOS[,VAR_CLASS]==3)
    {
      A<-FIJOS[1]+FIJOS[3]
      B<-FIJOS[4]+FIJOS[6]
      C<-FIJOS[7]+FIJOS[9]
    }
    AGRUPA<-c(A,B,C)
    RETORNA<-as.data.frame(t(AGRUPA))
  }
}
### DEFINIMOS PARA Z_MOD15
if (MODELO_N=="Z_MOD15")
{
  FIJOS<-summary(MODELO)$coefficients$fixed
  VAR_CLASS_TAMA<-names((summary(MODELO))$contrasts[1])

```

Apéndice D. SCRIPTS: PREDICCIÓN DE EFECTOS ALEATORIOS

```
VAR_CLASS_TIPO<-names((summary(MODELO))$contrasts[2])
if(is.null(VAR_CLASS_TAMA))
{
  RETORNA<-as.data.frame(t(FIJOS))
}
else
{
  if(DATOS[,VAR_CLASS_TAMA]==1) { A<-FIJOS[1] }
  if(DATOS[,VAR_CLASS_TAMA]==2) { A<-FIJOS[1]+FIJOS[2] }
  if(DATOS[,VAR_CLASS_TAMA]==3) { A<-FIJOS[1]+FIJOS[3] }
  if(DATOS[,VAR_CLASS_TAMA]==4) { A<-FIJOS[1]+FIJOS[4] }
  if(DATOS[,VAR_CLASS_TIPO]=="INDIVIDUALES") { C<-FIJOS[6]+FIJOS[7] }
  if(DATOS[,VAR_CLASS_TIPO]=="ALEATORIO") { C<-FIJOS[6] }
  B<-FIJOS[5]
  AGRUPA<-c(A,B,C)
  RETORNA<-as.data.frame(t(AGRUPA))
}
}
### DEFINIMOS PARA AB_MOD17
if(MODELO_N=="AB_MOD17")
{
  FIJOS<-fixef(MODELO)
  # if(DATOS[,VAR_CLASS_TIPO]=="INDIVIDUALES") { A<-FIJOS[1]+FIJOS[2] }
  A<-FIJOS[1]+FIJOS[2]
  # if(DATOS[,VAR_CLASS_TIPO]=="ALEATORIO") { A<-FIJOS[1] }
  # if(DATOS[,VAR_CLASS_TIPO]=="INDIVIDUALES") { C<-FIJOS[4]+FIJOS[5] }
  C<-FIJOS[4]+FIJOS[5]
  # if(DATOS[,VAR_CLASS_TIPO]=="ALEATORIO") { C<-FIJOS[4] }
  B<-FIJOS[3]
  AGRUPA<-c(A,B,C)
  RETORNA<-as.data.frame(t(AGRUPA))
}
### DEFINIMOS PARA AB_MOD22
if(MODELO_N=="AB_MOD22")
{
  FIJOS<-fixef(MODELO)
  if(DATOS[,VAR_CLASS_TIPO]=="INDIVIDUALES") { A<-FIJOS[1]+FIJOS[2] }
  if(DATOS[,VAR_CLASS_TIPO]=="ALEATORIO") { A<-FIJOS[1] }
  if(DATOS[,VAR_CLASS_TIPO]=="INDIVIDUALES") { C<-FIJOS[5]+FIJOS[6] }
  if(DATOS[,VAR_CLASS_TIPO]=="ALEATORIO") { C<-FIJOS[5] }
  C<-C+FIJOS[4]*DATOS$B.46.T.100.MIN.16
  B<-FIJOS[3]
  AGRUPA<-c(A,B,C)
  RETORNA<-as.data.frame(t(AGRUPA))
}
names(RETORNA)<-NOMBRE_FIJOS
RETORNA
}
```

Apéndice E

SCRIPTS: PREDICCIÓN DE VARIABLES DE RESPUESTA

```
PREDICHO<-function(MODELO , EXPRESION , PP_FINAL , VAR_DEP_CORR="DDF" , METODO="BLUP" , DDF_PRED){
  #VAR_DEP<-formula(MODELO) [[2]]
  VAR_DEP<-as.character(formula(MODELO) [[2]] [[2]])

  # EXTRACCION DE PARAMETROS DEL MODELO AJUSTADO
  #aa<-summary(MODELO)$modelStruct
  # DEFINICION DE LA LISTA DEL NRO DE NIVELES, DE EFECTOS ALEATORIOS PARA CADA NIVEL
  # Y LAS MATRICES DE VARIANZA/COVARIANZA DE CADA UNO DE ELLOS
  # NRO Y COEFICIENTES DE LAS MATRICES DE EFECTOS ALEATORIOS
  cat("Macro □ Predicho □ Ver □ 1.0 . . ")
  cat("Metodo : . . . " , METODO , "\n")
  D_PREV<-PP_FINAL
  D_POST<-PP_FINAL
  if (DDF_PRED==0) D_POST$DDF<-D_POST$DDF_COS
  if (DDF_PRED>0) D_POST$DDF<-DDF_PRED

  D_PREV$Pred_1<-eval(EXPRESION ,D_PREV)
  D_PREV$Pred_2<-eval(EXPRESION ,D_POST)
  DIF<-D_PREV[,as.character(VAR_DEP)]-D_PREV[, "Pred_2"]
  DIF_DDF<-D_POST$DDF-D_PREV$DDF
  Z1<-OBT_Z(D_PREV , DERIVADAS)
  Z2<-OBT_Z(D_PREV , DERIVADAS)
  BB<-OBT_B(D_PREV , DERIVADAS)
  BCORR<-OBT_COR(MODELO , DIF_DDF)

  TEMPO1<-rowSums(Z1*BB)
  TEMPO2<-rowSums(Z2*BB)

  if (BCORR[1]==0){
    if (METODO=="ZERO") { D_PREV$Pred_f<-D_PREV[, "Pred_2"]+TEMPO2 }
    if (METODO=="BLUP") { D_PREV$Pred_f<-D_PREV[, "Pred_2" ] }
```

Apéndice E. SCRIPTS: PREDICCIÓN DE VARIABLES DE RESPUESTA

```
}

if (BCORR[1]>0){
  if (METODO=="ZERO"){ D_PREV$Pred_f<-D_PREV[, "Pred_2"]+BCORR*(DIF) }
  if (METODO=="BLUP"){ D_PREV$Pred_f<-D_PREV[, "Pred_2"]-TEMPO2+BCORR*(DIF+TEMPO1) }
}
D_PREV$METODO2<-METODO
if (DDF_PRED==0) {
  D_PREV$ER<-D_PREV[, "DIAM_COS"]-D_PREV[, "Pred_f"]
  D_PREV$ER_ABS<-abs(D_PREV[, "DIAM_COS"]-D_PREV[, "Pred_f"])
  D_PREV$ER_REL<-D_PREV$ER/D_PREV[, "DIAM_COS"]
  D_PREV$ER_REL_ABS<-D_PREV$ER_ABS/D_PREV[, "DIAM_COS"]
}
D_PREV
}
```

Apéndice F

SCRIPTS: MÉTRICAS DE EVALUACIÓN

```
R2<- function(MODELO)
{
  if(class(MODELO)[1]=="nlme"){
    mydata=eval(MODELO$call[[3]])
    y=mydata[,as.character(formula(MODELO)[[2]])]
  }
  if(class(MODELO)[1]=="nlmerMod"){
    y<-model.response(MODELO@frame)
  }
  C<-cor(fitted(MODELO),y)
  C*C
}

R2_COND<- function(MODELO)
{
  if(class(MODELO)[1]=="nlme"){
    mydata=eval(MODELO$call[[3]])
    y=mydata[,as.character(formula(MODELO)[[2]])]
  }
  if(class(MODELO)[1]=="nlmerMod"){
    y<-model.response(MODELO@frame)
  }
  y_hat<-fitted(MODELO)
  y_media<-mean(y)
  y1<-y-y_hat
  y2<-y1*y1
  y_num<-sum(y2)
  y3<-y-y_media
  y4<-y3*y3
  y_den<-sum(y4)
  final<-1-(y_num/y_den)
  final
}

CC1<- function(MODELO)
```

```

{
  if(class(MODELO)[1]=="nlme"){
    mydata=eval(MODELO$call[[3]])
    y=mydata[,as.character(formula(MODELO)[[2]])]
  }
  if(class(MODELO)[1]=="nlmerMod"){
    y<-model.response(MODELO@frame)
  }
  y_hat<-fitted(MODELO)
  y_media<-mean(y)
  y1<-y-y_hat
  y2<-y1*y1
  y_num<-sum(y2)
  y3<-y-y_media
  y4<-y3*y3
  y_med_p<-mean(y_hat)
  y5<-y-y_med_p
  y6<-y5*y5
  den_1<-sum(y4)
  den_2<-sum(y6)
  den_3<-length(y_hat)*(y_media-y_med_p)*(y_media-y_med_p)
  final<-1-(y_num/(den_1+den_2+den_3))
  final
}

CC2<- function(MODELO)
{
  if(class(MODELO)[1]=="nlme"){
    mydata=eval(MODELO$call[[3]])
    y=mydata[,as.character(formula(MODELO)[[2]])]
  }
  if(class(MODELO)[1]=="nlmerMod"){
    y<-model.response(MODELO@frame)
  }
  y_hat<-fitted(MODELO)
  S1<-sd(y)
  S2<-sd(y_hat)
  S12<-cov(y,y_hat)
  ro<-S12/(S1*S2)
  y_media<-mean(y)
  y_med_p<-mean(y_hat)
  retorno<- ro*((4*S1*S2-ro*(S1*S1+S2*S2))/
    ((2-ro)*(S1*S1+S2*S2)+(y_media-y_med_p)*(y_media-y_med_p)))
  retorno
}

RSME<- function(MODELO)
{
  if(class(MODELO)[1]=="nlme"){
    mydata=eval(MODELO$call[[3]])
    y=mydata[,as.character(formula(MODELO)[[2]])]
  }
  if(class(MODELO)[1]=="nlmerMod"){

```

```
    y<-model.response(MODELO@frame)
  }
  y_hat<-fitted(MODELO)
  y_media<-mean(y)
  y1<-y-y_hat
  y2<-y1*y1
  retorno<-sqrt(1/length(y_hat)*sum(y2))
  retorno
}

SESGO_M<- function(MODELO)
{
  if(class(MODELO)[1]=="nlme"){
    mydata=eval(MODELO$call[[3]])
    y=mydata[,as.character(formula(MODELO)[[2]])]
  }
  if(class(MODELO)[1]=="nlmerMod"){
    y<-model.response(MODELO@frame)
  }
  y_hat<-fitted(MODELO)
  y_media<-mean(y)
  y1<-y-y_hat
  retorno<-1/length(y_hat)*sum(y1)
  retorno
}

SESGO_MA<- function(MODELO)
{
  if(class(MODELO)[1]=="nlme"){
    mydata=eval(MODELO$call[[3]])
    y=mydata[,as.character(formula(MODELO)[[2]])]
  }
  if(class(MODELO)[1]=="nlmerMod"){
    y<-model.response(MODELO@frame)
  }
  y_hat<-fitted(MODELO)
  y_media<-mean(y)
  y1<-abs(y-y_hat)
  retorno<-1/length(y_hat)*sum(y1)
  retorno
}

SESGO_MR<- function(MODELO)
{
  if(class(MODELO)[1]=="nlme"){
    mydata=eval(MODELO$call[[3]])
    y=mydata[,as.character(formula(MODELO)[[2]])]
  }
  if(class(MODELO)[1]=="nlmerMod"){
    y<-model.response(MODELO@frame)
  }
  y_hat<-fitted(MODELO)
  y_media<-mean(y)
  y1<-(y-y_hat)/y
}
```

Apéndice F. SCRIPTS: MÉTRICAS DE EVALUACIÓN

```
    retorno<-1/length(y_hat)*sum(y1)*100
  }
  returno
}

SESGO_MRA<- function(MODELO)
{
  if(class(MODELO)[1]=="nlme"){
    mydata=eval(MODELO$call[[3]])
    y=mydata[,as.character(formula(MODELO)[[2]])]
  }
  if(class(MODELO)[1]=="nlmerMod"){
    y<-model.response(MODELO@frame)
  }
  y_hat<-fitted(MODELO)
  y_media<-mean(y)
  y1<-abs(y-y_hat)/y
  retorno<-1/length(y_hat)*sum(y1)*100
  returno
}

# MEDIDA=1 PROMEDIO DE ERRORES, MEDIDA=2 PROMEDIO DE ERRORES ABSOLUTOS
# MEDIDA=3 PROMEDIO DE ERRORES RELATIVOS , MEDIDA=4 PROMEDIO DE ERRORES ABSOLUTOS RELATIVOS
# MEDIDA=5 VARIANZA DEL ERROR MEDIDA=6 RAIZ DEL CUADRADO MEDIO DEL ERROR

CAPACIDAD<-function(PP, OBSERVADO, PREDICHO, MEDIDA) {
  if(MEDIDA==1){
    RETORNA<-mean(PP[,OBSERVADO]-PP[,PREDICHO])
  }
  if(MEDIDA==2){
    RETORNA<-mean(abs(PP[,OBSERVADO]-PP[,PREDICHO]))
  }
  if(MEDIDA==3){
    RETORNA<-mean((PP[,OBSERVADO]-PP[,PREDICHO])/PP[,OBSERVADO])*100
  }
  if(MEDIDA==4){
    RETORNA<-mean(abs(PP[,OBSERVADO]-PP[,PREDICHO])/PP[,OBSERVADO])*100
  }
  if(MEDIDA==5){
    RETORNA<-var(PP[,OBSERVADO]-PP[,PREDICHO])
  }
  if(MEDIDA==6){
    RR1<-mean(PP[,OBSERVADO]-PP[,PREDICHO])
    RR2<-var(PP[,OBSERVADO]-PP[,PREDICHO])
    RETORNA<-(RR1**2+RR1)**0.5
  }
  RETORNA
}
```

Bibliografía

- [1] *S-PLUS 4 Guide to Statistics*. Data Analysis Products Division, MathSoft, Seattle, 1998.
- [2] *Anuario Estadístico 2013. Centro Regional Patagonia Norte*. Servicio Nacional de Sanidad y Calidad Agroalimentaria (SENASA), General Roca. Río Negro. Argentina, 2013.
- [3] *SAS/STAT 14.1 User's Guide*. SAS Institute Inc., New York, 2015.
- [4] Hartford A. and Davidian M. Consequences of misspecifying assumptions in nonlinear mixed effects models. *Computational Statistics and Data Analysis*, 34:139–164, 2000.
- [5] Mallet A., Mentre F., Steimer J.L., and Lokiec F. Nonparametric maximum likelihood estimation for population pharmacokinetics, with applications to cyclosporine. *Journal of Pharmacokinetics and Biopharmaceutics*, (16):311–327, 1988.
- [6] Rodriguez A.B. Influencia de la temperatura del aire y del raleo sobre el tamaño del fruto de peral (*pyrus communis*) cv williams durante sus distintas fases de desarrollo. Master's thesis, Universidad Nacional de Cordoba, 2009.
- [7] Goicochea A.M., Silveira M., Macchi A., Lavallo M., and Radonich M. *La estructura Urbana desde el uso del suelo*. Universidad Nacional del Comahue, Neuquén, 1990.
- [8] Joshua D. Angrist and Jörn-Steffen Pischke. *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press, Princeton, December 2008.
- [9] Dempster A.P., Laird N.M., and Rubin D.B. Maximum likelihood from incomplete data via the em algorithm. *Journal of the Royal Statistical Society*, (39):1–22, 1977.
- [10] Gallant A.R. and Nychka W.D. Semi-nonparametric maximum likelihood estimation. *Econometrica*, (55):363–390, 1987.

- [11] Carlin B. and Louis T. *Bayes and Empirical Bayes Methods for Data Analysis*. Chapman and Hall-CRC Press, New York, 2000.
- [12] Douglas M. Bates. *lme4: Mixed-effects modeling with R*. 2010.
- [13] Jeffrey E. Bennett, Amy Racine-Poon, and Jonathan Wakefield. *Markov chain Monte Carlo for non-linear hierarchical models*. Chapman and Hall, London, 1996.
- [14] R.A. Berk. *Statistical Learning from a Regression Perspective*. Springer, New York, 2008.
- [15] Zhang C., Tanabe K., Tamura F., Itai A., Wang S., Yoshida A., and Matsumoto K. The impact of cell division and cell enlargement on the evolution of fruit size in *pyrus pyrifolia*. *Annals of Botany*, (98):537–543, 2006.
- [16] Thornthwaite C.W. An Approach toward a Rational Classification of Climate. *Geographical Review*, 38(38):55–94, 1948.
- [17] Bates D. lmer, p-values and all that. Technical report, DRAFT r-sig-mixeg-models FAQ, 2006.
- [18] Hedeker D. and Gibbons R.D. *Longitudinal Data Analysis*. John Wiley, New York, 2006.
- [19] Stram D. and Lee J. Variance components testing in the longitudinal mixed-effects models. *Biometrics*, 50:1171–1177, 1994.
- [20] Hall D.B. and Bailey R.L. Modeling and prediction of forest growth variables based on multilevel nonlinear mixed models. *Forest Science*, 47:311–321, 2001.
- [21] Programa de Madurez. Temporada 2014. Technical report, Instituto Nacional de Tecnología Agropecuaria, Estación Experimental Alto Valle, 2014.
- [22] Programa Pronóstico de Producción. Pronostico de producción de frutos de pepita para rio negro y neuquén. Technical report, Facultad de Ciencias Agrarias. Universidad Nacional del Comahue-Secretaría de Estado de Fruticultura de Rio Negro- Ministerio de Desarrollo Territorial de la Provincia del Neuquén, Estación Experimental Alto Valle, 2013.
- [23] Demidenko E. 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, 1995.

- [24] Kuhn E. and Lavielle M. Maximum likelihood estimation in nonlinear mixed effects models. *Computational Statistics & Data Analysis*, 49(4):1020–1038, June 2005.
- [25] Sanchez E., Álvarez J., Colavita G., Dussi M., Fernandez D., Giayetto A., and Reeb P. *Pera Williams: Manual para el Productor y el Empacador*. Instituto Nacional de Tecnología Agropecuaria, General Roca. Río Negro, 2010. ISBN 978-987-25872-0-8.
- [26] Sober E. Instrumentalism, parsimony and the akaike framework. *Philosophy Science*, 69:112–123, 2002.
- [27] Vonesh E.F. A note on the use of laplace’s approximation for nonlinear mixed-effects models. *Biometrika*, (83):447–452, 1996.
- [28] Vonesh E.F., Wang H., Nie L., and Majumdar D. Conditional second-order generalized estimating equations for generalized linear and nonlinear mixed-effects models. *Journal of the American Statistical Association*, (97):271–283, 2002.
- [29] Vonesh E.F. and Carter R.L. Mixed effects nonlinear regression for unbalanced repeated measures. *Biometrics*, (48):1–17, 1992.
- [30] Vonesh E.F. and Chinchilli V.M. *Linear and Nonlinear Models for the Analysis of Repeated Measurements*. Marcel Dekker, New York, 1997.
- [31] Lehmann E.L. *Testing Statistical Hypotheses*. Wiley, New York, 1986.
- [32] Crecente-Campo F., Tomé M., Soares P., and Diéguez-Arand U. A generalized nonlinear mixed-effects height-diameter model for eucalyptus globulus l. in northwestern spain. *Forest Ecology and Management*, 259:494–952, 2010.
- [33] Distiller G., Little F., and Barnes K. Nonlinear mixed effects modeling of gametocyte carriage in patients with uncomplicated malaria. *Malaria Journal*, 60:1–12, 2010.
- [34] Gimenez G. Determinación de los patrones de crecimiento y abscisión de frutos para ajustar el momento oportuno de raleo en perales cv. williams. Master’s thesis, Universidad de Bologna-Universidad Nacional del Comahue-Inta, General Roca. Río Negro, 2013.
- [35] Guillermo Vallejo Seco G., Fernández Hermida J.R., and Secades Villa R. Análisis estadístico y consideraciones de potencia en la evaluación de programas mediante diseños de muestra en dos etapas. *Psicothema*, 15:300–308, 2003.

- [36] Kauermann G., Xu R., and Vaida F. Stacked laplace-em algorithm for duration models with time-varying and random effects. *Computational Statistics and Data Analysis*, (52):2514–2528, 2008.
- [37] Molenberghs G. and Verbeke G. A review on linear mixed models for longitudinal data, possibly subject to dropout. *Statistical Modeling*, page 235–269, 2001.
- [38] Schwarz G. Estimating the Dimension of a Model. *The Annals of Statistics*, 6(2):461–464, 1978.
- [39] Seco G., Grass J., Cabre R., and García P. Selección de modelos anidados para datos longitudinales usando criterios de información y la estrategia de ajuste condicional. *Psicothema*, 22:323–333, 2010.
- [40] Verbeke G. and Molenberghs G. *Linear Mixed Models for Longitudinal Data*. Springer, New York, 2000.
- [41] Judge G.G., Griffiths W.E., and Hill R.C. *The Theory and Practice of Econometrics*. John Wiley, New York, 1865.
- [42] De Silva H. Variation of Fruit Size and Growth within an Apple Tree and its Influence on Sampling Methods for Estimating the Parameters of Mid-season Size Distributions. *Annals of Botany*, 86(3):493–501, September 2000.
- [43] Furtado Campos H., Noblet J., Jaguelin-Peyraud Y., Gilbert H., Mormède P., Miranda F., and Renaudeau D. Thermoregulatory responses during thermal acclimation in pigs divergently selected for residual feed intake. *International Journal of Biometeorology*, 8:1545–1557, 2014.
- [44] Goldstein H. *Multilevel Statistical Models*. Halstead Press., New York, 1995.
- [45] Liu H., Zheng Y., and Shen J. Goodness-of-fit measures or r^2 for repeated measures mixed effect models. *Applied Statistics*, 35(10):1081–1092, 2008.
- [46] Rue H. and Martino S. Approximate bayesian inference for latent gaussian models by using integrated nested laplace approximations. *Journal of the Royal Statistical Society*, pages 319–292, 2009.
- [47] Warrington I., Fulton T.A., Halligan E.A., and Silva H. Apple fruit growth and maturity are affected by early season temperatures. *Journal of the American Society for Horticultural Science*, (124):468–477, 1999.

- [48] Barceló Coll J., Nicolás R.G., Sabater García B., and Sánchez R. *Fisiología Vegetal. Sexta Edición*. Editorial Piramides., Madrid, 1992.
- [49] Bertrand J., Laffont C.M., Mentré F., Chenel M., and Comets M. Development of a complex parent-metabolite joint population pharmacokinetic model. *The AAPS Journal*, 13:390–404, 2011.
- [50] Mandema J., Verotta D., and Sheiner L. Building population pharmacokinetics/pharmacodynamic models. *Journal of Pharmacokinetics and Biopharmaceutics*, 20:511–529, 1992.
- [51] Papadakis J. *Ecología y manejo de cultivos, pasturas y suelos*. Albatros, Ciudad Autónoma de Buenos Aires, 1980.
- [52] Ribbing J and Jonsson EN. Power, selection bias and predictive performance of the population pharmacokinetic covariate model. *Journal of Pharmacokinetics and Pharmacodynamics*, 31:109–134, 2004.
- [53] Serroyen J., Molenberghs G., Verbeke G., and Davidian M. Nonlinear Models for Longitudinal Data. *The American Statistician*, (63):378–388, 2009.
- [54] Wakefield J. The bayesian analysis of population pharmacokinetic models. *Journal of the American Statistical Association*, 91:62–75, 1996.
- [55] Naylor J.C. and Smith A.F.M. Applications of a method for the efficient computation of posterior distributions. *Applied Statistics*, (31):214–225, 1982.
- [56] Pinheiro J.C. *Topics in Mixed Effects Models*. University of Wisconsin, Madison, 1994.
- [57] Pinheiro J.C. and Bates D.M. Approximations to the log-likelihood function in the nonlinear mixed-effects model. *Journal of Computational and Graphical Statistics*, (4):12–35, 1995.
- [58] Pinheiro J.C. and Bates D.M. *Mixed-Effects Models in S and S-Plus*. Springer, 2000. ISBN 0-387-98957-0.
- [59] Wakefield J.C. The bayesian analysis of population pharmacokinetic models. technical report. *Statistics Section, Imperial College.*, (93-11), 1993.
- [60] Wakefield J.C., Smith A.F.M., Racine-Poon A., and Gelfand A.E. Bayesian analysis of linear and nonlinear population models by using the gibbs sampler. *Applied Statistics*, (43):201–221, 1994.

- [61] Orelien J.G. and Edwards L.J. Fixed-effect variable selection in linear mixed models using r^2 statistics. *Computational Statistics Data Analysis*, 52(4):1896–1907, 2008.
- [62] Liao J.J. An improved concordance coefficient. *Pharmaceutical Statistics*, 2(4):253–261, 2003.
- [63] Stirnemann J.J., Samson A, and Thalabard J.C. Individual predictions based on nonlinear mixed modeling: application to prenatal twin growth. *Statistics in Medicine*, (1):1–15, 2011.
- [64] Fu L., Sun H, Sharma C, Yuancai L., Zhang H., and Tang S. Nonlinear mixed-effects crown width models for individual trees of chinese fir (*cunninghamia lanceolata*) in south-central china. *Forest Ecology and Management*, 302:210–220, 2013.
- [65] Jiang L. and Li Y. Application of Nonlinear Mixed-Effects Modeling Approach in Tree Height Prediction. *Journal of Computers*, 5(10):1575–1581, October 2010.
- [66] Magee L. R^2 measures based on wald and likelihood ratio joint significance test. *American Statistician*, 44(3):250–253, 1990.
- [67] Tierney L. and Kadane J.B. Accurate approximations for posterior moments and densities. *Journal of the American Statistical Association*, 393(81):82–86, 1986.
- [68] Tierney L., Kass R.E., and Kadane J.B. Fully exponential laplace approximations to expectations and variances of nonpositive functions. *Journal of the American Statistical Association*, (84):710–716, 1989.
- [69] Wang L. Estimating Nonlinear Mixed-Effects Models by the Generalized Profiling Method and its Application to Pharmacokinetics. *McGill University*, 2007.
- [70] Wu L. and Wu H. Missing time-dependent covariates in HIV viral dynamic models. *J. Roy. Statist. Soc. Ser. C*, (51):297–318, 2002.
- [71] Sheiner L.B. and Beal S.L. Evaluation of methods for estimating population pharmacokinetics. i. michaelis-menton model: Routine clinical pharmacokinetics data. *Journal of Pharmacokinetics and Biopharmaceutics*, (8):673–687, 1980.
- [72] Bepete M. and Lakso A. Differential Effects of Shade on Early-season Fruit and Shoot Growth Rates in Empire Apple. *HortScience*, 5(33):823–825, 1999.

- [73] Conles M., Cragolini V., Balzarini M., and Macchiavelli R. Estimación de curvas de progreso de la incidencia de podredumbre blanca (*sclerotium cepivorum*) en cultivos de ajo mediante un modelo no lineal mixto. *Agriscientia*, XXVIII:61–74, 2011.
- [74] Davidian M. *Nonlinear Models for Univariate and Multivariate Response, Lecture Notes*. Department of Statistics, North Carolina State University. University of Wisconsin, Madison, 2009.
- [75] Davidian M. and Giltinan D.M. *Non-linear Models for Repeated Measurement Data*. Chapman and Hall, London, 1995.
- [76] Davidian M. and Giltinan D.M. Nonlinear models for repeated measurement data: an overview and update. *Journal of Agricultural, Biological, and Environmental Statistics*, 8:1–42, 2003.
- [77] Gurka M. Selecting the best linear mixed model under reml. *The American Statistician*, 60:19–26, 2006.
- [78] Kenward M. and Roger J. An improved approximation to the precision of fixed effects from restricted maximum likelihood. *Computational Statistics and Data Analysis*, 53:2583–2595, 2009.
- [79] Kramer M. R² as a goodness of fit statistics for mixed models. In *Workshop on Spatial Statistics for Agricultural and Environmental Research Applications*, pages 221–225, Beltsville, 2006. Agricultural Research Service, USDA.
- [80] Lindstrom M.L. and Bates D.M. Nonlinear mixed effects models for repeated measures data. *Biometrics*, 46(3):673–687, September 1990.
- [81] Wibbe M.L., Blanke M.M., and Lenz F. Effect of fruiting on carbon budgets of apple tree canopies. *Trees*, 8(1):56–60, 1993.
- [82] Chengcai N. and Gordon D. An analysis and comparison of predictors of random parameters demonstrated on planted loblolly pine diameter growth prediction. *Forestry*, 85(2):271–281, 2012.
- [83] Breslow N.E. and Lin X. Bias correction in generalised linear mixed models with a single component of dispersion. *Biometrika*, (82):81–91, 1995.
- [84] Hagenbuch Niels. A comparison of four methods to analyse a non-linear mixed-effects model using simulated pharmacokinetic data. Master's thesis, Swiss Federal Institute of Technology Zurich, 2011.

- [85] Nagelkerke N.J. A note on a general definition of the coefficient of determination. *Biometrika*, 78(3):691–692, 1991.
- [86] Laird N.M. and Ware J.H. Random-effects models for longitudinal data. *Biometrics*, (38):963–974, 1982.
- [87] Alvarez O. and Boche S. Modelos matematicos para describir crecimientos doble sigmoideos en frutos de nectaria tardio(c.v. Sun Grand). *AgroSur*, (27):21–28, 1999.
- [88] Alvarez O., Bramardi S.J., and Stangaferro S. Modelos de predicción anticipada del peso medio de los frutos en producción de manzanas. *Congreso Latino Americano de Biomatemática X ALAB - V ELAEM. Campinas, Brasil*, 2001.
- [89] Schabenberger O. and Pierce J.F. *Contemporary Statistical Models for the Plant and Soil Sciences*. CRC Press., Florida. United States, 2001.
- [90] Carlin B. P. and Louis T. A. *Bayesian Methods for Data Analysis*. CRC Press, 1 edition., 2009.
- [91] Minchin P and Thorpe M.R. What determines carbon partitioning between competing sinks?. *Journal of Experimental Botany*, 47:1293–1296, 1996.
- [92] Newton P.F. and Amposah I.G. Comparative evaluation of five height-diameter models developed for black spruce and jack pine stand-types in terms of goodness-of-fit, lack-of-fit and predictive ability. *Forest Ecology and Management*, 247:149–166, 2007.
- [93] Diggle P.J., Heagerty P, Liang K., and Zeger S. *Analysis of Longitudinal Data*. Oxford University Press, Oxford, 2001.
- [94] Bonate P.L. The effect of collinearity on parameter estimates in nonlinear mixed effect models. *Pharmaceutical Research*, 16(5):709–717, 1999.
- [95] Bonate P.L. *Pharmacokinetics-pharmacodynamic modeling and simulation*. Springer, New York, 2006.
- [96] Calama R. and Montero G. Multilevel linear mixed model for tree diameter increment in stone pine(*Pinus pinea*). *Silva Fenn.*, (39):37–54, 2005.
- [97] Jerez R., Ana Moret B., Carrero G., Macchiavelli R., and Quevedo R. Site index curves based on mixed models for teak plantations in the venezuelan plains. *Agrociencia*, 45:135–145, 2011.

- [98] Littell R., Milliken G., Stroup W., Wolfinger R., and Schabenber O. *SAS System for Mixed Mixed Models. 2nd edition*. SAS Institute Inc., New York, 2006.
- [99] Pothoff R. and Roy S. A generalized multivariate analysis of variance model useful especially for growth curve problems. *Biometrika*, 51:313–326, 1964.
- [100] Wolfinger R. Laplace's approximation for nonlinear mixed models. *Biometrika*, (80):791–795, 1993.
- [101] Johnson R.A. and Wichern D. *Applied Multivariate Statistical Analysis*. Prentice Hall, Upper Saddle River. New Jersey, 1999.
- [102] Litte R.A and Rubin B.D. *Statistical Analysis with Missing Data*. John Wiley, New York, 1987.
- [103] David A. Ratkowsky. *Nonlinear Regression Modeling : a Unified Practical Approach*. M. Dekker, New York, 1983.
- [104] Wolfinger R.D. and Lin X. Two taylor-series approximation methods for nonlinear mixed models. *Computational Statistics and Data Analysis*, (25):465–490, 1997.
- [105] Macchiavelli R.E. Combinando información longitudinal y transversal en estudios forestales: Construcción de curvas de índice de sitio mediante modelos no lineales mixtos. *XIX Reunión del Grupo Argentino de Biometria*, 2014.
- [106] Huang S., Meng S., and Yang Y. Prediction Implications of Nonlinear Mixed-effects Forest Biometric Models Estimated with a Generalizad Error Structure. *American Statistical Association*, (6):1174–1188, 2009.
- [107] Huang S. and Meng S.X. Estimating a Nonlinear Mixed Volume-age Model with and without Taking into Account Serrially-correlated Errors:Differences and Implications. *Modern Appied Science*, 3(5):3–20, May 2009.
- [108] Huang S., Meng S.X., and Yang Y. Assessing the goodness of fit of forest models estimated by nonlinear mixed-model methods. *Canadian Journal of Forest Research*, 39(12):2418–2436, December 2009.
- [109] Huang S., Shawn X., and Yuqing Y. Using Nonlinear Mixed Model Technique to Determine the Optimal Tree Height Prediction Model for Black Spruce. *Modern Applied Science*, 3(4):3–18, 2009.

- [110] Konishi S. and Kitagawa G. *Information Criteria and Statistical Modeling*. Springer, New York, 2007.
- [111] Meng S. and Huang S. Accounting for serial correlation and its impact on forecasting ability of a fixed and mixed effects basal area model:a case study. *European Forest Research*, (131):541–552, 2012.
- [112] Self S. and Liang K. Asymptotic properties of maximum likelihood estimators and likelihood ratio test under nonstandard conditions. *Journal of the American Statistical Association*, 82:605–610, 1987.
- [113] Walker S. An em algorithm for nonlinear random effects models. *Biometrics*, (52):934–944, 1996.
- [114] Bramardi S.J., Castro H.R., and Zanelli M.R. Fruit growth pattern of pear cv. Bartlett and Packham 's Triumph to improve hand thinning. *Acta Horticulturae*, (475):283–293, 1997.
- [115] Beal S.L. and Sheiner L.B. Estimating population kinetics.crc critical reviews. *Biomedical Engineering*, (8):195–222, 1982.
- [116] Beal S.L. and Sheiner L.B. *NONMEM User's Guides*. NONMEM Project Group. University of California, San Francisco., 1992.
- [117] Zeger S.L., Liang K.Y., and Albert P.S. Models for longitudinal data: a generalized estimating equation approach. *Biometrics*, 44(4):1049–60, 1988.
- [118] Leonard T., Hsu J.S.J., and Tsui K.W. Bayesian marginal inference. *Journal of the American Statistical Association*, (84):1051–1058, 1989.
- [119] Stukel T.A. and Demidenko E. *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, 1997.
- [120] R Development Core Team. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, 2009.
- [121] Chabikwa T.G. *Chemical thinning of European pear cultivars (Pyrus communis L.)*. PhD thesis, Stellenbosch: Stellenbosch University, 2008.
- [122] Kvalseth T.O. Cautionary note about R2. *American Statistician*, 4(39):279–282, 1985.

- [123] Wahlby Ulrika. *Methodological Studies on Covariate Model Building in Population Pharmacokinetic-Pharmacodynamics Analysis*. PhD thesis, Uppsala University, 2002.
- [124] Tassile V, Bramardi S., Reeb P, and De Bernardin F. Modelos de crecimiento de frutos de pepita: mejora en la precisión de las estimaciones a través de la inclusión de covariables ambientales y componentes aleatorias. In *X Reunión Científica del Grupo Argentino de Biometría. Universidad Nacional del Nordeste. Corrientes*, 2005.
- [125] Tassile V, Bramardi S., Reeb P, and De Bernardin F. Validación de modelos no lineales mixtos y su aplicación con fines de pronóstico. In *VII CLATSE. Rosario*, 2006.
- [126] Tassile V, Bramardi S., and Macchiavelli R. Modelos mixtos obtenidos a partir de información longitudinal y transversal. In *I Congreso Argentino de Estadística. Universidad Nacional de Córdoba. Córdoba*, 2015.
- [127] Lee C. Y. Y. and Wand M. P. Variational methods for fitting complex bayesian mixed effects models to health data. *Statistics in medicine*, 35:165–188, 2015.
- [128] Sakamoto Y., Ishiguro M., and Kitagawa G. *Akaike information criterion statistics*. Mathematics and its application. Japanese series. KTK Scientific Publishers, 1986.
- [129] Yang Y. and Huang S. Comparison of different methods for fitting nonlinear mixed forest models and for making predictions. 1686(May):1671–1686, 2011.
- [130] Yang Y and Huang S. On the statistical and biological behaviors of nonlinear mixed forest models. *European Journal Forest*, 132:727–736, 2013.
- [131] Yang Y., Huang S., and Meng S.X. Development of a tree-specific stem profile model for white spruce: a nonlinear mixed model approach with a generalized covariance structure. *Forestry*, 82(5):541–555, 2009.
- [132] Ge Z., Bickel P.J., and Rice J.A. An approximate likelihood approach to nonlinear mixed effects models via spline approximation. *Computational Statistics & Data Analysis*, 46(4):747–776, July 2004.
- [133] Pan Z. and Lin D.Y. Goodness-of-fit methods for generalized linear mixed models. *Biometrics*, 61(4):1000–1009, 2005.

Bibliografía

- [134] Shun Z. and McCullagh P. Laplace approximation of high dimensional integrals. *Journal of the Royal Statistical Society*, (57):749–760, 1995.