



UNIVERSIDAD NACIONAL DE CHIMBORAZO

**VICERRECTORADO DE INVESTIGACIÓN,
VINCULACIÓN Y POSGRADO**

DIRECCIÓN DE POSGRADO

**“ALGORITMOS NUMÉRICOS Y FILTROS BAYESIANOS PARA
ESTIMAR MODELOS DE EFECTOS MIXTOS”**

**TRABAJO DE GRADUACIÓN PARA OPTAR AL TÍTULO DE
MAGÍSTER EN MATEMÁTICA APLICADA CON MENCIÓN EN
MATEMÁTICA COMPUTACIONAL**

AUTOR:

Gabriel Alberto Barragan Ramirez

TUTOR:

Dr. Saba Rafael Infante Quirpa, PhD.

Riobamba – Ecuador, 2024

Certificación del Tutor

Certifico que el presente trabajo de titulación denominado: “**Algoritmos Numéricos y Filtros Bayesianos para Estimar Modelos de Efectos Mixtos**”, ha sido elaborado por Gabriel Alberto Barragan Ramirez, el mismo que ha sido orientado y revisado con el asesoramiento permanente de mi persona en calidad de Tutor. Así mismo, refrendo que dicho trabajo de titulación ha sido revisado por la herramienta antiplagio institucional; por lo que certifico que se encuentra apto para su presentación y defensa respectiva.

Es todo cuanto puedo informar en honor a la verdad.

Riobamba, 07 de marzo, de 2024



Saba Rafael Infante Quirpa, PhD.

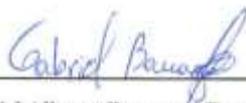
TUTOR

Declaración de Autoría y Cesión de Derechos

Yo, **Gabriel Alberto Barragan Ramirez**, con número único de identificación **092658379-0**, declaro y acepto ser responsable de las ideas, doctrinas, resultados y lineamientos alternativos realizados en el presente trabajo de titulación denominado: "Algoritmos Numéricos y Filtros Bayesianos para Estimar Modelos de Efectos Mixtos" previo a la obtención del grado de Magister en Matemática Aplicada con mención en Matemática Computacional.

- Declaro que mi trabajo investigativo pertenece al patrimonio de la Universidad Nacional de Chimborazo de conformidad con lo establecido en el artículo 20 literal j) de la Ley Orgánica de Educación Superior LOES.
- Autorizo a la Universidad Nacional de Chimborazo que pueda hacer uso del referido trabajo de titulación y a difundirlo como estime conveniente por cualquier medio conocido, y para que sea integrado en formato digital al Sistema de Información de la Educación Superior del Ecuador para su difusión pública respetando los derechos de autor, dando cumplimiento de esta manera a lo estipulado en el artículo 144 de la Ley Orgánica de Educación Superior LOES.

Riobamba, 07 de marzo de 2024



Gabriel Alberto Barragan Ramirez

N.U.I. 092658379-0



Dirección de
Posgrado
VICERRECTORADO DE INVESTIGACIÓN,
VINCULACIÓN Y POSGRADO



Riobamba, 27 de febrero de 2024

ACTA DE SUPERACIÓN DE OBSERVACIONES

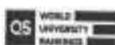
En calidad de miembro del Tribunal designado por la Comisión de Posgrado, CERTIFICO que una vez revisado el Proyecto de Investigación y/o desarrollo denominado "**ALGORITMOS NUMÉRICOS Y FILTROS BAYESIANOS PARA ESTIMAR MODELOS DE EFECTOS MIXTOS**", dentro de la línea de investigación de **Ingeniería Informática**, presentado por el maestrante **Barragan Ramirez Gabriel Alberto**, portador de la CI. 0926583790, del programa de **Maestría en Matemática Aplicada con mención en Matemática Computacional**, cumple al 100% con los parámetros establecidos por la Dirección de Posgrado de la Universidad Nacional de Chimborazo.

Es todo lo que podemos certificar en honor a la verdad.

Atentamente,

Saba Rafael Infante Quirpa

MIEMBRO DEL TRIBUNAL



Campus La Dolores
Av. Eloy Alfaro y 10 de Agosto
Teléfono (593-3) 373-0880 ext. 2002
Riobamba - Ecuador



Dirección de
Posgrado
VICERRECTORADO DE INVESTIGACIÓN,
VINCULACIÓN Y POSGRADO



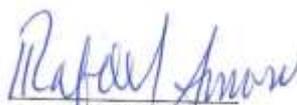
Riobamba, 27 de febrero de 2024

ACTA DE SUPERACIÓN DE OBSERVACIONES

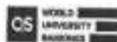
En calidad de miembro del Tribunal designado por la Comisión de Posgrado, CERTIFICO que una vez revisado el Proyecto de Investigación y/o desarrollo denominado "**ALGORITMOS NUMÉRICOS Y FILTROS BAYESIANOS PARA ESTIMAR MODELOS DE EFECTOS MIXTOS**", dentro de la línea de investigación de **Ingeniería Informática**, presentado por el maestrante **Barragan Ramirez Gabriel Alberto**, portador de la CI. 0926583790, del programa de **Maestría en Matemática Aplicada con mención en Matemática Computacional**, cumple al 100% con los parámetros establecidos por la Dirección de Posgrado de la Universidad Nacional de Chimborazo.

Es todo lo que podemos certificar en honor a la verdad.

Atentamente,


Isidro Rafael Amaro

MIEMBRO DEL TRIBUNAL



Campus La Dolorosa
Av. Eloy Alfaro y 10 de Agosto
Teléfono: (593-3) 373-0880 ext. 2002
Riobamba - Ecuador



Dirección de
Posgrado
VICERRECTORADO DE INVESTIGACIÓN,
VINCULACIÓN Y POSGRADO



Riobamba, 27 de febrero de 2024

ACTA DE SUPERACIÓN DE OBSERVACIONES

En calidad de miembro del Tribunal designado por la Comisión de Posgrado, CERTIFICO que una vez revisado el Proyecto de Investigación y/o desarrollo denominado "**ALGORITMOS NUMÉRICOS Y FILTROS BAYESIANOS PARA ESTIMAR MODELOS DE EFECTOS MIXTOS**", dentro de la línea de investigación de **Ingeniería Informática**, presentado por el maestrante **Barragan Ramirez Gabriel Alberto**, portador de la CI. 0926583790, del programa de **Maestría en Matemática Aplicada con mención en Matemática Computacional**, cumple al 100% con los parámetros establecidos por la Dirección de Posgrado de la Universidad Nacional de Chimborazo.

Es todo lo que podemos certificar en honor a la verdad.

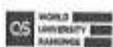
Atentamente,



VERIFICACIÓN DEL
DOCUMENTO
GUILLERMO EDVIN
MACHADO SOTOMAYOR

Guillermo Edvin Machado Sofomayor

MIEMBRO DEL TRIBUNAL



Campus La Dolorosa
Av. Elty Alfaro y 10 de Agosto
Teléfono (593-3) 373-0880, ext. 3003
Riobamba - Ecuador

Unach.edu.ec
en asociación



Dirección de Posgrado
VICERRECTORADO DE INVESTIGACIÓN,
VINCULACIÓN Y POSGRADO

en movimiento

Riobamba, 27 de febrero de 2024

CERTIFICADO

De mi consideración:

Yo Saba Rafael Infante Quirpa, certifico que Gabriel Alberto Barragán Ramírez con cédula de identidad No. 0926583790 estudiante del programa de maestría en Matemática Aplicada con mención en Matemática Computacional, cohorte Primera (2021-2022), presentó su trabajo de titulación bajo la modalidad de Proyecto de titulación con componente de investigación aplicada/desarrollo denominado: ALGORITMOS NUMÉRICOS Y FILTROS BAYESIANOS PARA ESTIMAR MODELOS DE EFECTOS MIXTOS, el mismo que fue sometido al sistema de verificación de similitud de contenido Turnitin identificando el 6% en el texto.

Es todo en cuanto puedo certificar en honor a la verdad.

Atentamente,

Saba Rafael Infante Quirpa

CI: 1757134943

Adj.-

- Resultado del análisis de similitud



Dirección de Posgrado
VICERRECTORADO DE INVESTIGACIÓN,
VINCULACIÓN Y POSGRADO

en movimiento

Riobamba, 5 de marzo de 2024

CERTIFICADO

De mi consideración:

Yo Kléber Augusto Jaramillo Galarza coordinador académico, certifico que Gabriel Alberto Barragan Ramirez con cédula de identidad No. 0926583790 estudiante del programa de maestría en Matemática Aplicada con mención en Matemática Computacional, cohorte Primera (2021-2022), presentó su trabajo de titulación bajo la modalidad de Proyecto de titulación con componente de investigación aplicada/desarrollo denominado: **ALGORITMOS NUMÉRICOS Y FILTROS BAYESIANOS PARA ESTIMAR MODELOS DE EFECTOS MIXTOS**, el mismo que fue sometido al sistema de verificación de similitud de contenido TURNITIN identificando el 6% de similitud en el texto, el cual fue realizado por su tutor.

Es todo en cuanto puedo certificar en honor a la verdad.

Atentamente,



Ing. Kléber Augusto Jaramillo Galarza, MSc
CI: 0703748939

Adj.-

- Informe tutor académico
- Resultado del análisis de similitud

Agradecimiento

Quiero agradecer a todos mis profesores de la maestría que me compartieron sus conocimientos y me inculcaron el propósito de investigar. También agradezco al profesor Saba Infante por proporcionarme un tema de investigación que dio origen al desarrollo del trabajo de titulación, asimismo por compartir sus conocimientos y sus fuentes de información junto con las bases que adquirí en las asignaturas de la maestría ayudaron a estructurar y alcanzar los objetivos planteados en la tesis.

También agradezco a mis compañeros de maestría que me compartieron sus experiencias y sus conocimientos en diferentes áreas del conocimiento, esto me permitió ampliar mi campo de visión sobre temas que me eran ajenos en un inicio y también me proporcionó las herramientas para poder resolver estos temas.

Dedicatoria

Dedicado a mis seres queridos, en especial a mi madre, por haberme apoyado no solo en la maestría, sino también desde que empecé mi etapa universitaria con sus consejos, muchas gracias.

Índice General

Certificación del Tutor	ii
Declaración de Autoría y Cesión de Derechos	iii
Actas de superación de observaciones	iv
Certificados de similitudes	vii
Agradecimiento	ix
Dedicatoria	x
Índice General.....	xi
Índice de Figuras	xiv
Índice de Algoritmos	xxiii
Índice de Tablas	xxiv
Resumen	xxix
Abstract	xxx
1 Introducción	1
1.1 Planteamiento del problema	3
1.2 Objetivos.....	4
1.2.1 Objetivo general	4
1.2.2 Objetivos específicos.....	4
2 Marco teórico	5
2.1 Proceso estocástico	5
2.2 Transformación de variables aleatorias	5
2.3 Modelo estado – espacio.....	6
2.4 Estimación Bayesiana de los parámetros en modelos espacio estado	10
2.5 Ecuaciones diferenciales estocásticas (EDEs).....	14
2.6 Modelos de EDEs con efectos mixtos	15
2.7 Filtros Bayesianos	18

2.8	Métodos de Monte Carlo	19
2.9	Muestreo por rechazo	21
2.9.1	Muestreo por importancia (MI).....	23
2.9.2	Evaluación numérica estable de los pesos.....	26
2.9.3	Método de muestreo por importancia secuencial para un modelo espacio estado 27	
2.9.4	Proceso de remuestreo.....	29
2.10	Filtros de partículas.....	31
2.11	Muestreo de Gibbs	34
2.12	Algoritmo Metropolis – Hastings	35
2.12.1	Algoritmo Metropolis	37
2.12.2	Algoritmo MALA (Metropolis adjusted Langevin algorithm).....	43
2.12.3	Algoritmo de Monte Carlo Hamiltoniano	47
2.13	Algoritmo Metropolis – Hastings de partículas	51
3	Metodología.....	56
3.1	Algoritmos numéricos de simulación de EDEs.....	56
3.1.1	Esquema Euler – Maruyama	57
3.1.2	Puente de difusión modificado.....	57
3.1.3	Puente de difusión residual	58
3.2	Métodos de Monte Carlo por cadena de Markov para la estimación de parámetros 59	
3.2.1	Métodos MCMC pseudo – marginales	59
3.2.2	Enfoque pseudo – marginal correlacionado	61
3.3	Métodos pseudo – marginales de partículas para la estimación de EDEs con efectos mixtos.....	61
3.3.1	Método pseudo – marginal de aumento individual	62

3.3.2	Método pseudo – marginal por componentes	65
4	Aplicaciones.....	67
4.1	Datos reales.....	67
4.1.1	Modelo de crecimiento del árbol de naranja	67
4.2	Datos simulados.....	68
4.2.1	Proceso de Ornstein – Uhlenbeck bidimensional.....	68
5	Resultados	72
5.1	Modelo de crecimiento del árbol de naranja	72
5.1.1	Método pseudo – marginal de aumento individual	72
5.1.2	Método pseudo – marginal por componentes (MALA)	81
5.1.3	Método pseudo – marginal por componentes (Hamiltoniano).....	94
5.1.4	Proceso de Ornstein-Uhlenbeck bidimensional	110
6	Conclusiones.....	144
	Referencias Bibliográficas	145

Índice de Figuras

Figura 1. Representación gráfica de la estructura de un modelo espacio estado (Dahlin y Schön, 2019).....	7
Figura 2. Simulación de los estados y observaciones de una caminata aleatoria Gaussiana.	9
Figura 3. Representación gráfica de un modelo espacio estado (Schön y Lindsten, 2017).	12
Figura 4. Simulación de un proceso OU unidimensional.....	15
Figura 5. Proceso OU modelado por una EDE con efectos mixtos con $M = 40$	17
Figura 6. Datos simulados de la caminata aleatoria Gaussiana.....	33
Figura 7. Caminata aleatoria Gaussiana y trayectoria de filtrado.....	34
Figura 8. Histograma de las observaciones simuladas.	38
Figura 9. Descripción general del algoritmo Metropolis para estimar parámetros (Rampinelli, 2019).....	39
Figura 10. Histograma y diagrama de caja y bigotes de la media μ (Metropolis).....	41
Figura 11. Histograma y diagrama de caja y bigotes de la desviación estándar σ (Metropolis).	42
Figura 12. Diagrama de traza de los parámetros μ y σ (Metropolis).....	42
Figura 13. Histograma y diagrama de caja y bigotes de la media μ (MALA).....	45
Figura 14. Histograma y diagrama de caja y bigotes de la desviación estándar σ (MALA)	46
Figura 15. Diagrama de traza de los parámetros μ y σ (MALA).	46
Figura 16. Histograma y diagrama de caja y bigotes de la media μ (Hamiltoniano)	50

Figura 17. Histograma y diagrama de caja y bigotes de la desviación estándar σ (Hamiltoniano)	50
Figura 18. Diagrama de traza de los parámetros μ y σ (Hamiltoniano).	51
Figura 19. Histograma y diagrama de caja y bigotes del parámetro s_x	53
Figura 20. Histograma y diagrama de caja y bigotes del parámetro s_y	54
Figura 21. Diagrama de traza de los parámetros s_x y s_y	54
Figura 22. Gráfico de las observaciones (datos reales) de los $M = 5$ árboles del modelo de crecimiento del árbol de naranja.....	68
Figura 23. Simulación de un modelo bidimensional OU.	71
Figura 24. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización EM).....	73
Figura 25. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización EM).....	74
Figura 26. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización EM).....	74
Figura 27. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión modificado).....	76

Figura 28. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión modificado).....	77
Figura 29. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión modificado).....	77
Figura 30. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión residual).	79
Figura 31. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión residual).	80
Figura 32. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión residual).	80
Figura 33. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización EM).	82
Figura 34. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización EM).....	83

Figura 35. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización EM).....	84
Figura 36. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización EM).....	85
Figura 37. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado).	87
Figura 38. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado).....	87
Figura 39. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado).....	88
Figura 40. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado).....	89
Figura 41. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual).....	91

Figura 42. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual).	92
Figura 43. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual).	93
Figura 44. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual).	93
Figura 45. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM).	95
Figura 46. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM).	96
Figura 47. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM).	97
Figura 48. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM).	98

Figura 49. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).	100
Figura 50. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).	101
Figura 51. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).	102
Figura 52. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).	103
Figura 53. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión residual).	105
Figura 54. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión residual).	106
Figura 55. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión residual).	107

- Figura 56. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado). 108
- Figura 57. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios $\mathbf{b}(\mathbf{m})$ (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs). 112
- Figura 58. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro α_2 (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs)..... 113
- Figura 59. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro ξ_2 (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs)..... 113
- Figura 60. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros ψ (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs). 115
- Figura 61. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios $\mathbf{b}(\mathbf{m})$ (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs). 117
- Figura 62. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro α_2 (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs). 118
- Figura 63. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro ξ_2 (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs). 119

Figura 64. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros $\boldsymbol{\psi}$ (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).	121
Figura 65. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios $\mathbf{b}(\mathbf{m})$ (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).	123
Figura 66. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro α^2 (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).	124
Figura 67. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro ξ^2 (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).	124
Figura 68. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros $\boldsymbol{\psi}$ (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).	126
Figura 69. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios $\mathbf{b}(\mathbf{m})$ (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).	128
Figura 70. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro α^2 (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).	129
Figura 71. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro ξ^2 (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).	129
Figura 72. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros $\boldsymbol{\psi}$ (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).	131
Figura 73. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios $\mathbf{b}(\mathbf{m})$ (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).	133

Figura 74. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro α^2 (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).	134
Figura 75. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro ξ^2 (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).	134
Figura 76. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros ψ (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).	136
Figura 77. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios $b(m)$ (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).	138
Figura 78. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro α^2 (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).	139
Figura 79. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro ξ^2 (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).	139
Figura 80. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros ψ (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).	141

Índice de Algoritmos

Algoritmo 1. Algoritmo Monte Carlo.....	21
Algoritmo 2. Muestreo por rechazo.....	22
Algoritmo 3. Muestreo por importancia.....	25
Algoritmo 4. Muestreo por importancia secuencial	29
Algoritmo 5. Remuestreo	30
Algoritmo 6. Filtro de partículas	31
Algoritmo 7. Muestreo de Gibbs	35
Algoritmo 8. Algoritmo Metropolis – Hastings (MH)	36
Algoritmo 9. Metropolis – Hastings Langevin Algorithm (MALA).....	44
Algoritmo 10. Algoritmo de Monte Carlo Hamiltoniano.....	48
Algoritmo 11. Algoritmo Metropolis - Hastings de partículas.....	52
Algoritmo 12. Algoritmo Metropolis - Hastings marginal de partículas	60
Algoritmo 13. Método Pseudo - Marginal de aumento individual.....	63
Algoritmo 14. Algoritmo de estimación de verosimilitud individual	64
Algoritmo 15. Método pseudo – marginal por componentes	65

Índice de Tablas

Tabla 1. Resumen estadístico de la cadena de Markov (algoritmo Metropolis) correspondiente a los parámetros μ y σ	43
Tabla 2. Resumen estadístico de las cadenas de Markov (algoritmo MALA) correspondiente a los parámetros μ y σ	47
Tabla 3. Resumen estadístico de las cadenas de Markov (Hamiltoniano) correspondiente a los parámetros μ y σ	51
Tabla 4. Resumen estadístico de las cadenas de Markov (algoritmo Metropolis – Hastings de partículas) correspondiente a los parámetros sx y sy	55
Tabla 5. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización EM).	75
Tabla 6. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión modificado).	78
Tabla 7. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión residual).	81
Tabla 8. Tasa de aceptación de los efectos aleatorios (Método pseudo – marginal por componentes – MALA con esquema de EM).	82

Tabla 9. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización EM).....	85
Tabla 10. Tasa de aceptación de los efectos aleatorios (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado). ..	86
Tabla 11. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado).....	89
Tabla 12. Tasa de aceptación de los efectos aleatorios (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual).....	90
Tabla 13. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual). ..	94
Tabla 14. Tasa de aceptación de los efectos aleatorios (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM).....	95
Tabla 15. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM).....	98
Tabla 16. Tasa de aceptación de los efectos aleatorios (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).....	99

Tabla 17. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).....	103
Tabla 18. Tasa de aceptación de los efectos aleatorios (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión residual).	104
Tabla 19. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión residual).	108
Tabla 20. Indicadores de desempeño de los métodos pseudo – marginales.....	110
Tabla 21. Tasa de aceptación de los efectos aleatorios $\mathbf{b}(\mathbf{m})$ (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs).	111
Tabla 22. Resumen estadístico de los parámetros comunes θ (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs).	114
Tabla 23. Tasa de aceptación de los hiperparámetros ψ (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs).	114
Tabla 24. Resumen estadístico de los hiperparámetros ψ (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs).	115
Tabla 25. Tasa de aceptación de los efectos aleatorios $\mathbf{b}(\mathbf{m})$ (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).	117
Tabla 26. Resumen estadístico de los parámetros comunes θ (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).	119

Tabla 27. Tasa de aceptación de los hiperparámetros ψ (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).	120
Tabla 28. Resumen estadístico de los hiperparámetros ψ (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).	121
Tabla 29. Tasa de aceptación de los efectos aleatorios $b(m)$ (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).	122
Tabla 30. Resumen estadístico de los parámetros comunes θ (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).	125
Tabla 31. Tasa de aceptación de los hiperparámetros ψ (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).	125
Tabla 32. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros ψ (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).	126
Tabla 33. Tasa de aceptación de los efectos aleatorios $b(m)$ (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).	127
Tabla 34. Resumen estadístico de los parámetros comunes θ (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).	130
Tabla 35. Tasa de aceptación de los hiperparámetros ψ (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).	130
Tabla 36. Resumen estadístico de los hiperparámetros ψ (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).	131
Tabla 37. Tasa de aceptación de los efectos aleatorios $b(m)$ (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).	132

Tabla 38. Resumen estadístico de los parámetros comunes θ (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).	135
Tabla 39. Tasa de aceptación de los hiperparámetros ψ (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).	135
Tabla 40. Resumen estadístico de los hiperparámetros ψ (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).	136
Tabla 41. Tasa de aceptación de los efectos aleatorios $b(m)$ (Algoritmo Hamiltoniano correlacionado dentro de un esquema de Muestreo de Gibbs).	138
Tabla 42. Resumen estadístico de los parámetros comunes θ (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).	140
Tabla 43. Tasa de aceptación de los hiperparámetros ψ (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).	140
Tabla 44. Resumen estadístico de los hiperparámetros ψ (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).	141
Tabla 45. Indicadores de desempeño de los algoritmos Metropolis – Hastings dentro de un esquema de muestreo de Gibbs.	142

Resumen

La estimación de parámetros y estados en modelos estado – espacio definidos en la forma de ecuaciones diferenciales estocásticas con efectos mixtos es un problema desafiante en la comunidad estadística desde el punto de vista frecuentista y Bayesiana. Dado que la verosimilitud de los datos observados a partir de modelos de espacio – estado es intratable en problemas prácticos, los métodos pseudo – marginales representan una metodología para estimar parámetros desconocidos y estados latentes en una ecuación diferencial estocástica (EDE) con efectos mixtos. Sin embargo, la implementación de métodos pseudo – marginales puede requerir simular trayectorias de EDEs, esto significa resolver una EDE analíticamente para obtener la densidad de transición. En la mayoría de los casos, las EDEs no tienen una solución explícita, por lo que es necesario utilizar métodos aproximados para la discretización de las EDEs, como el esquema de Euler – Maruyama, los esquemas de puentes de difusión y esquemas Monte Carlo. Comparamos la eficiencia de los métodos pseudo – marginales en escenarios simulados y reales modelados por modelos de efectos mixtos definidos por una ecuación diferencial estocástica (stochastic differential equation mixed – effects models, SDEMEMs). Para la simulación aproximada de las trayectorias del EDE, consideramos tres esquemas de discretización: Euler – Maruyama, Puente de difusión modificado y Puente de difusión residual.

Palabras claves: *Algoritmo Metropolis – Hastings, Filtro de Partículas, modelo estado – espacio, inferencia Bayesiana, efectos aleatorios.*

Abstract

Parameter and state estimation in state–space models, defined as stochastic differential equations with mixed–effects, is a challenging problem in the statistical community from a frequentist and Bayesian point of view. Since the observed data likelihood from the state–space models is intractable in practical issues, pseudo–marginal methods represent a methodology to estimate unknown parameters and latent states in stochastic differential equations (SDE) with mixed – effects. However, the implementation of pseudo–marginal methods may require the simulation of trajectories of SDE; this means solving an SDE analytically to obtain the transition density. In most cases, SDEs do not have an explicit solution. Therefore, it is necessary to use approximate methods for the discretization of SDEs, such as Euler – Maruyama, diffusion bridge schemes, and Monte Carlo schemes. We compare the efficiency of the pseudo–marginal methods in simulated and real scenarios modeled by stochastic differential equation mixed–effect models (SDEMEmS). For the approximate simulation of the paths of the SDE, we consider three discretization schemes: Euler – Maruyama, Modified diffusion bridge, and Residual diffusion bridge.

Keywords: *Metropolis – Hastings Algorithm, Particle Filter, state – space model, Bayesian inference, random effects.*



Revised by
Mario N. Salazar

1 INTRODUCCIÓN

Las Ecuaciones Diferenciales Estocásticas (EDEs) son usadas para modelar sistemas dinámicos complejos con respecto al tiempo y espacio. Las EDEs pueden ser consideradas como una extensión estocástica de las Ecuaciones Diferenciales Ordinarias (EDOs) que permiten considerar las variaciones aleatorias alrededor de la dinámica promedio especificada por las EDOs (Botha et al., 2021). Si se consideran M procesos estocásticos correspondientes a un conjunto de $m = 1, 2, \dots, M$ elementos o individuos, se puede definir un modelo de efectos mixtos con M ecuaciones diferenciales estocásticas con funciones drift y difusión comunes pero con parámetros aleatorios que permiten diferencias entre los procesos (Delattre, 2021). Algunos ejemplos de fenómenos dinámicos que pueden ser modelados mediante EDEs con efectos mixtos son crecimiento de células cancerígenas (Picchini y Forman, 2019), conteo del número de insectos (pulgón del algodón) (Whitaker et al., 2017), procesos Ornstein-Uhlenbeck (Fadwa et al., 2020), entre otros. Desde el enfoque de la Estadística Bayesiana, el objetivo principal es estimar la distribución de los efectos aleatorios y los efectos fijos que son comunes a los M procesos. (Delattre, 2021). Por lo general, los modelos de EDEs con efectos mixtos no tienen soluciones analíticas dado que la verosimilitud suele ser intratable (Botha et al., 2021), lo que conlleva a plantear un esquema de discretización de las EDEs y métodos de Monte Carlo por cadena de Markov para obtener una estimación de los parámetros y estados de los modelos de efectos mixtos (Wiqvist et al., 2021).

En el capítulo 1 se describirán el planteamiento del problema, el objetivo general y los objetivos específicos. En el capítulo 2 se presentará el marco teórico donde se resaltan los modelos estado – espacio, modelos de ecuaciones diferenciales estocásticas con efectos mixtos y filtros bayesianos. En el capítulo 3 se mostrará la metodología empleada para

estimar los parámetros desconocidos o los estados soluciones de una ecuación diferencial estocástica con efectos mixtos, entre los métodos empleados se encontrarán los esquemas de discretización para simular las trayectorias de una ecuación diferencial estocástica, también se mostrarán métodos de Monte Carlo por cadena de Markov (muestreo de Gibbs, algoritmo Metropolis – Hastings, MALA, algoritmo Hamiltoniano) y métodos pseudo – marginales correlacionados. Mientras que en el capítulo 4 se mostrarán algunos ejemplos de ecuaciones diferenciales estocásticas con efectos mixtos (modelo de crecimiento del árbol de naranja y modelo bidimensional de Ornstein – Uhlenbeck). En el capítulo 5 se mostrarán los resultados de la aplicación de los métodos pseudo – marginales para la estimación Bayesiana de los modelos de efectos mixtos. Finalmente, en el capítulo 6 se resumirán las conclusiones.

1.1 Planteamiento del problema

Un proceso espacio estado consta de dos procesos: un proceso latente Markoviano $\{X_t\}_{t \geq 0} \subset \mathbb{R}^n$ donde X_t es observado parcialmente y un proceso de observación $\{Y_t\}_{t \geq 0} \subset \mathbb{R}^m$ donde Y_t es una observación del estado X_t en el instante t (Barragán et al., 2020), (Botha et al., 2021). A partir de las observaciones Y_t se pueden reconstruir los estados ocultos X_t , en particular, consideraremos el problema de estimar los estados y los parámetros de un modelo de efectos mixtos definido por una ecuación diferencial estocástica (EDE). Sea M el número de individuos que forman parte del modelo de efectos mixtos, se puede definir la siguiente ecuación diferencial estocástica

$$\begin{aligned} dX_t^{(m)} &= \mu\left(X_t^{(m)}, \phi_X, b^{(m)}\right)dt + \sqrt{v}\left(X_t^{(m)}, \phi_X, b^{(m)}\right)dW_t^{(m)} & m = 1, \dots, M \\ X_0^{(m)} &= X_0^{(m)}(\phi_X, b^{(m)}) \end{aligned}$$

donde $\mu(\cdot)$ es la función de *drift* y $\sqrt{v}(\cdot)$ es la función de *difusión*, $\{W_t^{(m)}\}_{t \geq 0}$ es un proceso estándar de Wiener, también llamado proceso de movimiento Browniano. El vector ϕ_X contiene los parámetros comunes (o efectos fijos) de la ecuación diferencial estocástica, mientras que $b^{(m)}$ es el vector de parámetros específicos de cada individuo (o efectos aleatorios), donde $b^{(m)} \sim p(b^{(m)}|\psi)$ (Soto et al., 2019).

Si consideramos la siguiente ecuación de observación

$$y_t^{(m)} = g\left(X_t^{(m)}\right) + \epsilon_t^{(m)}$$

obtenemos un proceso espacio estado, donde $y_t^{(m)} \in \{Y_t^{(m)}\}$ es un observación con ruido del individuo m en el instante t . Asumimos que el ruido de las observaciones es originado por un proceso Gaussiano, $\epsilon_t^{(m)} \sim N(0, \sigma^2)$ (Soto y Infante, 2019).

Sea $\theta = (\sigma, \phi_X, \psi)$ el vector de todos los parámetros desconocidos en el modelo de efectos mixtos definido por una ecuación diferencial estocástica, se plantea la siguiente pregunta de investigación:

¿Cómo estimar los estados ocultos y los parámetros desconocidos de un modelo de efectos mixtos definido por una ecuación diferencial estocástica?

1.2 Objetivos

1.2.1 *Objetivo general*

- Estimar los estados y los parámetros desconocidos de un modelo de efectos mixtos definido por una ecuación diferencial estocástica.

1.2.2 *Objetivos específicos*

- Implementar el algoritmo numérico Euler – Maruyama en un modelo de efectos mixtos.
- Implementar el algoritmo numérico de Puente de difusión en un modelo de efectos mixtos.
- Implementar el algoritmo de partículas Monte Carlo por cadena de Markov en un modelo de efectos mixtos.
- Realizar simulaciones de los modelos de efectos mixtos definidos por EDE.
- Ilustrar la metodología propuesta usando datos reales.
- Calcular medidas de bondad de ajuste para validar los modelos de efectos mixtos.

2 MARCO TEÓRICO

2.1 Proceso estocástico

Definición

Un proceso estocástico X_t es una colección de variables aleatorias

$$\{X_t\}_{t \in T} = \{X_t(\omega) : t \in T, \omega \in \Omega\}$$

definido en un espacio probabilístico Ω .

Definición

Un proceso estocástico W_t es llamado proceso de Wiener (o movimiento browniano) si se cumplen las siguientes condiciones:

- El proceso empieza en el origen: $W_0 = 0$.
- W_t tiene incrementos independientes y estacionarios.
- $\forall t > s > 0$, los incrementos $\Delta W = W_t - W_s$ son variables aleatorias con densidad $\Delta W \sim N(\Delta W; 0, \sigma^2 \Delta)$ con $\Delta = t - s$.
- W_t tiene una trayectoria continua pero no diferenciable.

2.2 Transformación de variables aleatorias

Sea X una variable aleatoria restringida en el dominio I con densidad $X \sim p_X(x)$ y $g: \mathbb{R} \rightarrow I$ una función invertible, esto es $h(\cdot) = g^{-1}(\cdot)$ existe, tal que $Y = g(X)$, entonces Y es una nueva variable aleatoria con densidad $p_Y(y) = p_X(h(y)) \left| \frac{dh(y)}{dy} \right|$ (Gelman et al., 2014).

Ejemplo. Transformación de variables aleatorias

Sea X una variable aleatoria con distribución semi-normal con parámetro σ

$$HN(x; \sigma) = \frac{1}{\sigma} \sqrt{\frac{2}{\pi}} \exp\left(-\frac{1}{2} \frac{x^2}{\sigma^2}\right)$$

Consideramos la siguiente transformación $y = g(x) = \log(x)$, cuya inversa es

$g^{-1}(y) = h(y) = \exp(y)$ y derivada $\frac{dh(y)}{dy} = \exp(y)$. Por lo tanto,

$$\begin{aligned} HN(y; \sigma) &= \frac{1}{\sigma} \sqrt{\frac{2}{\pi}} \exp\left(-\frac{1}{2} \frac{(\exp y)^2}{\sigma^2}\right) \exp(y) \\ &= \frac{1}{\sigma} \sqrt{\frac{2}{\pi}} \exp\left(-\frac{1}{2} \frac{\exp(2y)}{\sigma^2} + y\right) \end{aligned}$$

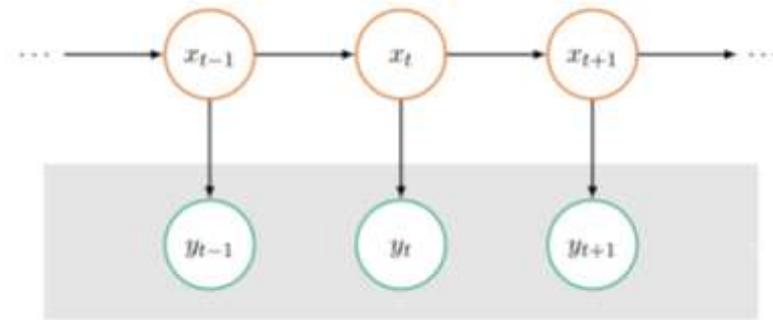
2.3 Modelo estado – espacio

Un modelo estado – espacio, también conocido como modelo de Markov oculto, es un tipo de modelos probabilísticos que consiste de dos procesos: estados latentes u ocultos $\{x_t: t = 0, 1, 2, \dots\}$ y observaciones $\{y_t: t = 1, 2, \dots\}$. Un modelo estado – espacio se caracteriza por cumplir dos propiedad: 1) las variables latentes satisfacen la propiedad de Markov, esto es, el estado latente x_t sólo depende del estado previo x_{t-1} del proceso Markoviano de primer orden, así $p(x_t | x_{1:t-1}, y_{1:t-1}) = p(x_t | x_{t-1})$, y 2) las observaciones son condicionalmente independientes, esto es, la observación actual y_t dado el estado actual x_t es condicionalmente del registro de observaciones y estados previos, así $p(y_t | x_{1:t}, y_{1:t-1}) = p(y_t, x_t)$ (Särkkä, 2013).

En la Figura 1, se muestra una representación gráfica de un modelo espacio estado con los estados latentes o desconocidos (arriba) y las observaciones de estos estados (abajo), sin un control exógeno. También se observa que flechas horizontales representan la probabilidad $p(x_t | x_{t-1})$ de transición de un estado previo x_{t-1} a un estado actual x_t ,

también se pueden apreciar las flechas verticales que indican que las observaciones son condicionalmente independientes dado el estado actual, esto es, $p(y_t|x_t)$.

Figura 1. Representación gráfica de la estructura de un modelo espacio estado (Dahlin y Schön, 2019).



Asumimos que los estados y las observaciones son reales, esto es, $x_t \in \mathbb{R}^{n_x}$ y $y_t \in \mathbb{R}^{n_y}$. El estado inicial x_0 sigue una distribución de acuerdo a la densidad $\mu(x_0|\theta)$ parametrizada por algunos parámetros estáticos desconocidos $\theta \in \Theta \subset \mathbb{R}^p$ (Dahlin y Schön, 2019). Los estados x_t y las observaciones y_t son descritas por las densidades $f(x_t|x_{t-1}, \theta)$ y $g(y_t|x_t, \theta)$, respectivamente. Esto quiere decir que la densidad $f(x_t|x_{t-1}, \theta)$, que describe el proceso de los estados, proporciona la probabilidad del próximo estado x_t dado el anterior estado x_{t-1} . De forma análoga se puede interpretar la densidad de las observaciones $g(y_t|x_t, \theta)$ (Dahlin y Schön, 2019). En un enfoque Bayesiano, la forma de estimar los parámetros θ es considerando una densidad a priori de los parámetros $p(\theta)$ y tratarlos como variables aleatorias adicionales en el modelo (Särkkä, 2013). Con esta notación, se puede definir un modelo probabilístico no lineal espacio estado como sigue

$$\begin{aligned}
 \theta &\sim p(\theta) \\
 x_0|\theta &\sim \mu(x_0|\theta) \\
 x_t|x_{t-1} &\sim f(x_t|x_{t-1}, \theta) \\
 y_t|x_t &\sim g(y_t|x_t, \theta)
 \end{aligned}$$

donde $\mu(x_0|\theta)$ es una densidad de inicialización, $f(x_t|x_{t-1}, \theta)$ es la densidad de transición de los estados latentes y $g(y_t|x_t, \theta)$ es la densidad de observación para.

Ejemplo. Caminata aleatoria Gaussiana.

Un ejemplo de un proceso Markoviano es una caminata aleatoria Gaussiana que está definida por

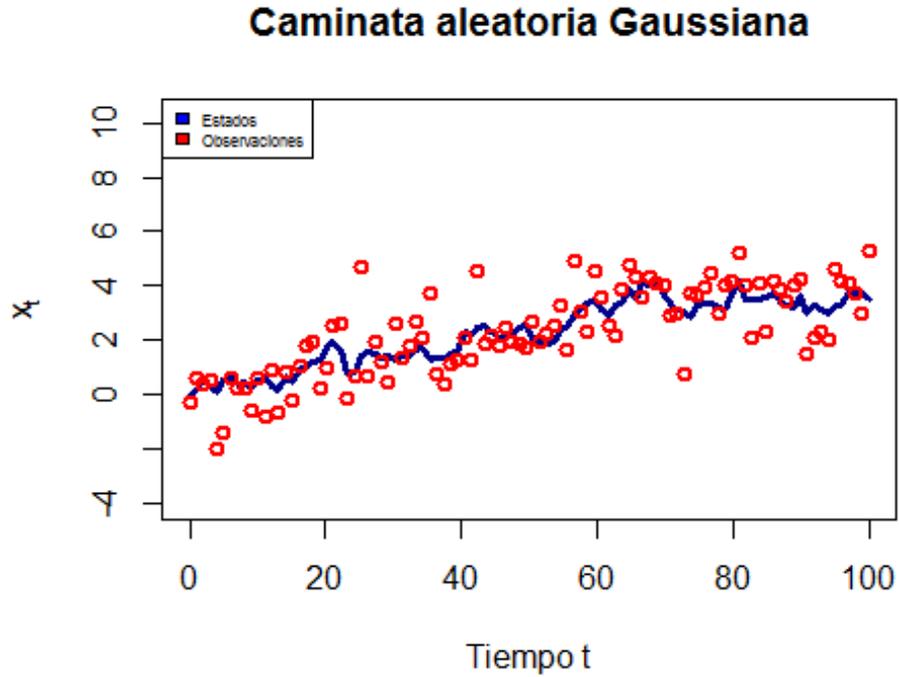
$$\begin{aligned} x_t &= x_{t-1} + q_{t-1} \quad , \quad q_{t-1} \sim N(q_{t-1}; 0, s_x^2) \\ y_t &= x_t + r_t \quad , \quad r_t \sim N(r_t; 0, s_y^2) \end{aligned}$$

En términos de las densidades de probabilidad, el modelo puede ser escrito como

$$\begin{aligned} p(x_t|x_{t-1}) &= N(x_t; x_{t-1}, s_x^2) \\ &= \frac{1}{\sqrt{2\pi s_x^2}} \exp\left(-\frac{1}{2Q} (x_t - x_{t-1})^2\right) \\ p(y_t|x_t) &= N(y_t; x_t, s_y^2) \\ &= \frac{1}{\sqrt{2\pi s_y^2}} \exp\left(-\frac{1}{2R} (y_t - x_t)^2\right) \end{aligned}$$

En la Figura 2, se muestra una simulación de los estados y observaciones de una caminata aleatoria Gaussiana unidimensional con parámetros $s_x = 0.1$ y $s_y = 1$ y estado inicial $x_0 = 0$.

Figura 2. Simulación de los estados y observaciones de una caminata aleatoria Gaussiana.



La inferencia Bayesiana de parámetros en modelos espacio estado tiene como principal objetivo obtener una estimación de los parámetros θ dadas las observaciones $y_{1:T} = \{y_1, y_2, \dots, y_T\}$ mediante el cálculo de la densidad a posteriori de los parámetros, esto es posible mediante la regla de Bayes

$$p(\theta|y_{1:T}) = \frac{p(y_{1:T}|\theta)p(\theta)}{\int_{\theta} p(y_{1:T}|\theta')p(\theta')d\theta'} = \frac{\gamma(\theta)}{Z_{\theta}}$$

donde $p(\theta)$ y $p(y_{1:T}|\theta)$ representan la densidad a priori y la verosimilitud (o densidad de los datos), respectivamente. La densidad a posteriori no normalizada denotada por $\gamma(\theta) = p(y_{1:T}|\theta)p(\theta)$ es un componente importante de los métodos de partículas Metropolis-Hastings. El denominador $Z_{\theta} = \int_{\theta} p(y_{1:T}|\theta')p(\theta')d\theta'$ es conocido como la verosimilitud marginal o evidencia del modelo (Dahlin y Schön, 2019). Dado que la verosimilitud marginal suele ser intratable, los métodos de inferencia Bayesianos se concentran en la

densidad a posteriori no normalizada $\gamma(\theta)$. Entre los métodos Bayesianos de estimación de parámetros, se encuentran los métodos de Monte Carlo por cadena de Markov, que se caracterizan por el muestreo de la densidad a posteriori no normalizada, $p(\theta|y_{1:T}) \propto p(y_{1:T}|\theta)p(\theta)$ (Botha et al., 2021).

Definición. Logaritmo de la verosimilitud

Sea $p(y_{1:T}|\theta)$ la función verosimilitud, entonces definimos el logaritmo de la verosimilitud como $\ell(\theta) = \log p(y_{1:T}|\theta)$.

El logaritmo de la verosimilitud es muy usado en problemas de optimización, por ejemplo, en problemas de verosimilitud máxima donde se usa el hecho de que la función logaritmo es una función estrictamente creciente, otra ventaja para emplear el logaritmo de la verosimilitud es que permite simplificar los cálculos cuando se usan modelos exponenciales, por ejemplo, modelos Gaussianos (ya que $\log \exp(b) = b$). Otra razón importante para definir el logaritmo de la verosimilitud es que permite diseñar algoritmos computacionalmente estables (Schön y Lindsten, 2017).

2.4 Estimación Bayesiana de los parámetros en modelos espacio estado

Dado que $\theta \in \Theta$ es el vector de parámetros desconocidos en un modelo espacio estado, la Estadística Bayesiana procede a considerar una densidad a priori de θ que describa la naturaleza estocástica de los parámetros, así, $\theta \sim p(\theta)$.

Mediante la regla de Bayes, la densidad a posteriori completa de los estados latentes y los parámetros puede ser expresada como sigue

$$\begin{aligned}
p(x_{0:T}, \theta | y_{1:T}) &= \frac{p(x_{0:T}, \theta, y_{1:T})}{p(y_{1:T})} \\
&= \frac{p(y_{1:T} | x_{0:T}, \theta) p(x_{0:T}, \theta)}{p(y_{1:T})} \\
&= \frac{p(y_{1:T} | x_{0:T}, \theta) p(x_{0:T} | \theta) p(\theta)}{p(y_{1:T})} \\
&\propto p(y_{1:T} | x_{0:T}, \theta) p(x_{0:T} | \theta) p(\theta) \\
&= \gamma(x_{0:T}, \theta | y_{1:T})
\end{aligned}$$

donde $p(x_{0:T}, \theta, y_{1:T})$ es conocida como el modelo probabilístico completo y $\gamma(x_{0:T}, \theta | y_{1:T})$ es llamada la densidad a posteriori conjunta no normalizada. Esta densidad se puede factorizar en la densidad de los datos $p(y_{1:T} | x_{0:T}, \theta)$ y en la densidad a priori de los estados latentes y los parámetros $p(x_{0:T}, \theta) = p(x_{0:T} | \theta) p(\theta)$ (Schön y Lindsten, 2017).

Mediante probabilidades condicionadas y la propiedad de observaciones condicionalmente independientes podemos factorizar la densidad de los datos como sigue

$$\begin{aligned}
p(y_{1:T} | x_{0:T}, \theta) &= p(y_T | y_{1:T-1}, x_{0:T}, \theta) p(y_{1:T-1} | x_{0:T}, \theta) \\
&= p(y_T | x_T, \theta) p(y_{1:T-1} | x_{0:T-1}, \theta) \\
&\vdots \\
&= \prod_{t=1}^T g(y_t | x_t, \theta)
\end{aligned}$$

donde $g(y_t | x_t, \theta)$ es la densidad de observación asociada al modelo espacio estado.

Ahora usando la densidad de transición de los estados latentes $f(x_t | x_{t-1}, \theta)$ y la propiedad Markov podemos factorizar $p(x_{0:T} | \theta)$

$$\begin{aligned}
p(x_{0:T} | \theta) &= p(x_T | x_{0:T-1}) p(x_{0:T-1} | \theta) \\
&= p(x_T | x_{T-1}) p(x_{0:T-1} | \theta) \\
&\vdots \\
&= \mu(x_0 | \theta) \prod_{t=1}^T f(x_t | x_{t-1}, \theta)
\end{aligned}$$

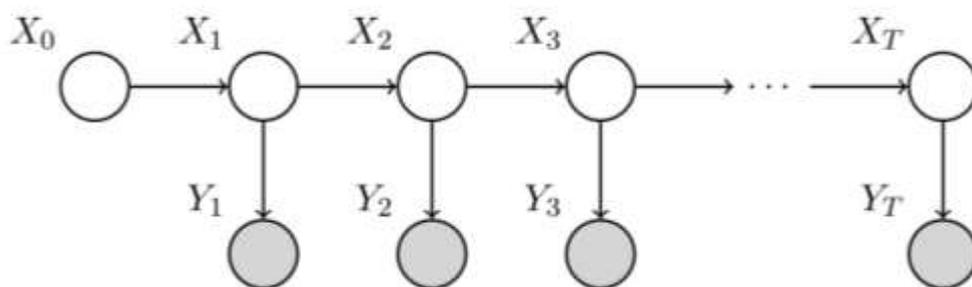
Ejemplo. Representación gráfica de un modelo espacio estado

A continuación, consideramos la densidad conjunta de los estados latentes y las observaciones de un modelo espacio estado

$$\begin{aligned}
 p(x_{0:T}, y_{1:T} | \theta) &= p(y_{1:T} | x_{0:T}, \theta) p(x_{0:T} | \theta) \\
 &= \left(\prod_{t=1}^T g(y_t | x_t, \theta) \right) \left(\mu(x_0 | \theta) \prod_{t=1}^T f(x_t | x_{t-1}, \theta) \right) \\
 &= \mu(x_0 | \theta) \prod_{t=1}^T f(x_t | x_{t-1}, \theta) g(y_t | x_t, \theta)
 \end{aligned}$$

En la Figura 3, se muestra la representación gráfica de un modelo espacio estado en donde cada variable aleatoria (X_t o Y_t) está representada como un nodo. Los nodos en color blanco corresponden a los estados latentes, esto es, no observados, y los nodos en color gris indican que las variables son observadas. Las relaciones probabilísticas entre las variables aleatorias están representadas mediante flechas. Cada flecha corresponde a uno de los factores de la descomposición de la densidad conjunta $p(x_{0:t}, y_{1:t} | \theta)$ (Schön y Lindsten, 2017).

Figura 3. Representación gráfica de un modelo espacio estado (Schön y Lindsten, 2017).



A partir de la factorizaciones anteriores y acomodando términos, la densidad a posteriori no normalizada de los estados latentes y parámetros $\gamma(x_{0:t}, \theta | y_{1:t})$ se puede escribir como

$$\begin{aligned}
\gamma(x_{0:T}, \theta | y_{1:T}) &= p(y_{1:T} | x_{0:T}, \theta) p(x_{0:T} | \theta) p(\theta) \\
&= p(\theta) \mu(x_0 | \theta) \prod_{t=1}^T f(x_t | x_{t-1}, \theta) g(y_t | x_t, \theta)
\end{aligned}$$

Continuando con la Inferencia Bayesiana de los parámetros θ , se tiene que la densidad a posteriori no normalizada de los parámetros se escribe como $\gamma(\theta) = p(y_{1:T} | \theta) p(\theta)$. Usando marginalización en la densidad conjunta $p(x_{0:T}, y_{1:T} | \theta)$ se sigue que

$$p(y_{1:T} | \theta) = \int_X p(x_{0:T}, y_{1:T} | \theta) dx_{0:T}$$

Así, la verosimilitud $p(y_{1:T} | \theta)$ se puede interpretar como un promedio de la densidad conjunta de los estados y observaciones $p(x_{0:T}, y_{1:T} | \theta)$ sobre todas las posibles trayectorias de los estados $x_{0:T} = \{x_0, x_1, \dots, x_T\}$ (Schön y Lindsten, 2017).

Luego, usando las factorizaciones anteriores y acomodando los términos, se tiene que la densidad no normalizada $\gamma(\theta)$ se puede escribir como

$$\begin{aligned}
\gamma(\theta) &= p(y_{1:T} | \theta) p(\theta) \\
&= p(\theta) \int_X p(x_{0:T}, y_{1:T} | \theta) dx_{0:T} \\
&= p(\theta) \int_X \mu(x_0 | \theta) \prod_{t=1}^T f(x_t | x_{t-1}, \theta) g(y_t | x_t, \theta) dx_{0:T}
\end{aligned}$$

Sin embargo, la integral anterior que surge de la verosimilitud no suele tener una solución analítica, salvo en algunos modelos lineales y con densidades Gaussianas (Doucet y Johansen, 2009). Adicionalmente, en algunos modelos las EDEs no suelen tener soluciones analíticas, esto es, densidades de transición intratables (Botha et al., 2021).

2.5 Ecuaciones diferenciales estocásticas (EDEs)

Un proceso de Itô unidimensional es un proceso estocástico $\{X_t\}_{t \geq 0}$ que satisface la siguiente ecuación

$$X_t = X_0 + \int_0^t \mu(X_u, u) du + \int_0^t \sqrt{v}(X_u, u) dW_u$$

donde $W_t = \int_0^t dW_u$ es un proceso de Wiener unidimensional estándar y $X_0 = x_0$ es la condición inicial en $t = 0$. Mientras que funciones μ y \sqrt{v} son la función drift y función de difusión, respectivamente (Øksendal, 2000).

Dado un proceso de Itô $\{X_t\}_{t \geq 0}$, la forma general para una ecuación diferencial estocástica (EDE) que describe el proceso $\{X_t\}_{t \geq 0}$ es

$$\begin{aligned} dX_t &= \mu(X_t, t)dt + \sqrt{v}(X_t, t)dW_t \\ X_0 &= x_0 \end{aligned}$$

Ejemplo. Proceso Ornstein-Uhlenbeck unidimensional

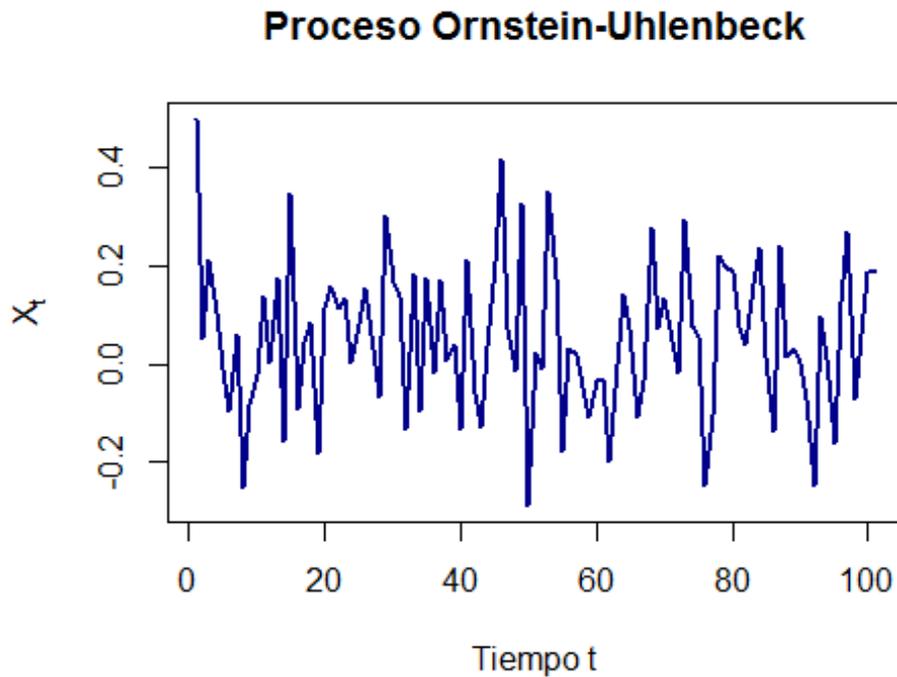
A continuación, un proceso de Ornstein-Uhlenbeck (OU) unidimensional satisface la siguiente EDE

$$dX_t = \beta(\alpha - X_t)dt + \sigma dW_t$$

donde $\beta > 0$ es la tasa de crecimiento, esto es, un indicador de la reacción del sistema frente a perturbaciones, $\alpha \in \mathbb{R}$ es el promedio estacionario para el proceso X_t y σ es el coeficiente de difusión del proceso OU (Wiqvist et al., 2021).

En la Figura 4 se muestra una simulación de un proceso OU unidimensional con parámetros $\beta = 3$, $\alpha = 1$, $\sigma = 1.5$ y condición inicial $x_0 = 0.5$.

Figura 4. Simulación de un proceso OU unidimensional.



2.6 Modelos de EDEs con efectos mixtos

Las EDEs pueden ser extendidas para considerar la variabilidad intersubjetiva e intrasubjetiva de $m = 1, \dots, M$ elementos o individuos, esto permite definir una EDE con efectos mixtos como sigue

$$dX_t^{(m)} = \mu(X_t^{(m)}, \phi_X, b^{(m)}, t)dt + \sqrt{v}(X_t^{(m)}, \phi_X, b^{(m)}, t)dW_t^{(m)}$$

$$X_0^m(\phi_X, b^{(m)}) = x_0$$

donde $W_t^{(m)}$ es un proceso de Wiener estándar e independiente, ϕ_X son los parámetros fijos o comunes para los M individuos del modelo, también llamados efectos fijos y mientras que $b^{(m)}$ son los parámetros individuales o específicos a cada individuo $m = 1, \dots, M$. Desde el enfoque Bayesiano de la Estadística Inferencial, los parámetros $b^{(m)}$ son considerados como variables aleatorias (efectos aleatorios) y además se asume que la densidad de los parámetros

$b^{(m)}$ está parametrizada por los hiperparámetros o efectos poblacionales ψ , esto es, $b^{(m)} \sim p(b^{(m)}|\psi)$ (Wiqvist et al., 2021).

Ejemplo. Proceso OU modelado por una EDE con efectos mixtos

A continuación, consideramos una EDE con efectos mixtos que modela un proceso OU unidimensional.

$$dX_t^{(m)} = \beta^{(m)}(\alpha^{(m)} - X_t^{(m)})dt + \sigma^{(m)}dW_t^{(m)}, \quad m = 1, \dots, M$$

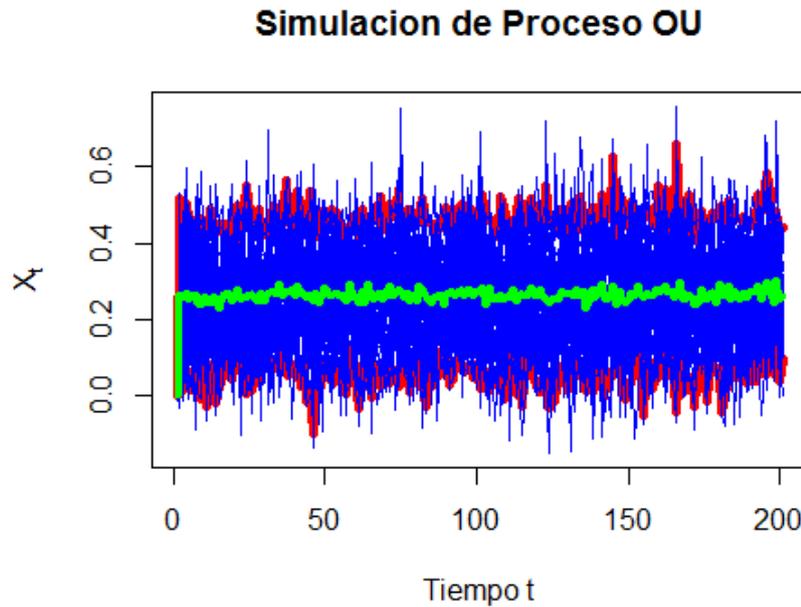
donde $\beta^{(m)}$, $\alpha^{(m)}$ y $\sigma^{(m)}$ son los efectos aleatorios. En la literatura se asume que los efectos aleatorios son estrictamente positivos por lo que se consideran sus versiones logarítmica, esto es, $\phi_1^{(m)} = \log \beta^{(m)}$, $\phi_2^{(m)} = \log \alpha^{(m)}$ y $\phi_3^{(m)} = \log \sigma^{(m)}$, así se tiene $b^{(m)} = (\phi_1^{(m)}, \phi_2^{(m)}, \phi_3^{(m)})$ (Wiqvist et al., 2021). Estos parámetros se distribuyen de acuerdo a las siguientes densidades

$$\phi_j^{(m)} \sim N(\phi_j^{(m)}; \mu_j, \tau_j^{-1}), \quad j = 1, 2, 3$$

donde $\psi = (\mu_1, \tau_1, \mu_2, \tau_2, \mu_3, \tau_3)$ son los hiperparámetros del modelo de efectos mixtos con τ_j como la precisión de $\phi_j^{(m)}$.

En la Figura 5 se muestra una simulación de un proceso OU con efectos mixtos ($M = 40$), las trayectorias de los procesos estocásticos $X_t^{(m)}$ están en color azul, la trayectoria promedio de los procesos estocásticos aparece en color verde y los cuantiles están en color rojo. Los parámetros usados para la simulación son $\psi = (-0.7, -0.9, 10, 2.3, 4, 4)$ y condición inicial $x_0^{(m)} = 0$ para $m = 1, \dots, 40$.

Figura 5. Proceso OU modelado por una EDE con efectos mixtos con $M = 40$.



Se asume que cada uno de los estados $\{X_t^{(m)}\}_{t \geq 0}$ no puede ser observado directamente, sino mediante observaciones con ruido $Y_t^{(m)} = (Y_1^{(m)}, Y_2^{(m)}, \dots, Y_{t_m}^{(m)})$, donde $y_{t_m}^{(m)}$ representa las observaciones en el tiempo t_m es el número de observaciones para el individuo m . También, se asume que las observaciones son condicionalmente independiente dado los estados latentes. Sea F una matriz constante, definimos un modelo espacio estado como sigue

$$\begin{aligned} dX_t^{(m)} &= \mu(X_t^{(m)}, \phi_X, b^{(m)}, t)dt + \sqrt{v}(X_t^{(m)}, \phi_X, b^{(m)}, t)dW_t^{(m)} \\ Y_t^{(m)} &= FX_t^{(m)} + \epsilon_t, \quad \epsilon_t | \Sigma \sim N(\epsilon_t; 0, \Sigma) \\ &\quad b^{(m)} \sim p(b^{(m)} | \psi) \end{aligned}$$

De este modo es posible definir el vector de parámetros desconocidos a estimar $\theta = (\phi_X, \psi, \Sigma)$.

2.7 Filtros Bayesianos

Volvamos a considerar el problema de aproximar la densidad a posteriori $p(\theta|y_{1:T}) \propto p(y_{1:T}|\theta)p(\theta)$. Dado que la densidad a priori puede ser seleccionada tal que sea fácil de evaluar, usualmente la dificultad se produce en evaluar la verosimilitud de los datos $p(y_{1:T}|\theta)$. A continuación, factorizamos la verosimilitud mediante el uso de las probabilidades condicionales varias veces como sigue

$$\begin{aligned} p(y_{1:T}|\theta) &= p(y_{1:T-1}, y_T|\theta) \\ &= p(y_T|y_{1:T-1}, \theta)p(y_{1:T-1}|\theta) \\ &\vdots \\ &= \prod_{t=1}^T p(y_t|y_{1:t-1}, \theta) \end{aligned}$$

donde $p(y_1|y_{1:0}, \theta) = p(y_1|\theta)$ es una inicialización. Esta factorización de $p(y_{1:T}|\theta)$ es conocida como descomposición del error de predicción (Särkkä, 2013). Cada uno de los factores $p(y_t|y_{1:t-1}, \theta)$ se llama verosimilitud predictiva y se pueden calcular recursivamente como

$$p(y_t|y_{1:t-1}, \theta) = \int_{\mathcal{X}} p(y_t|x_t, \theta)p(x_t|y_{1:t-1}, \theta)dx_t$$

donde $p(y_t|x_t, \theta) = g(y_t|x_t, \theta)$ es la densidad de observación del modelo espacio estado y $p(x_t|y_{1:t-1}, \theta)$ es la densidad predictiva de los estados, esta última puede ser calculada usando las ecuaciones recursivas de filtrado Bayesianas como sigue

- Primero, inicializar con la densidad inicial $\mu(x_0)$
- Para cada paso $t = 1, 2, \dots, T$ hacer
 - Paso de predicción. La densidad predictiva de los estados x_t puede ser calculada con la ecuación de Chapman – Kolmogorov

$$p(x_t|y_{1:t-1}, \theta) = \int_X p(x_t|x_{t-1}, \theta)p(x_{t-1}|y_{1:t-1}, \theta)dx_{t-1}$$

donde $p(x_t|x_{t-1}, \theta) = f(x_t|x_{t-1}, \theta)$ es la densidad de transición de los estados.

- Paso de actualización de las observaciones. Dada la observación y_t en el tiempo t , calcular la densidad a posteriori de los estados $p(x_t|y_{1:t}, \theta)$ usando la regla de Bayes como sigue

$$p(x_t|y_{1:t}, \theta) = \frac{p(y_t|x_t, \theta)p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1}, \theta)} = \frac{g(y_t|x_t, \theta)p(x_t|y_{1:t-1})}{p(y_t|y_{1:t-1}, \theta)}$$

2.8 Métodos de Monte Carlo

Para motivar los métodos de Monte Carlo, vamos a considerar dos problemas muy comunes en Estadística Bayesiana: Valor esperado y marginalización.

1. Valor esperado

Sea X una variable aleatoria que se distribuye de acuerdo a $p_X(x)$, esto es, $X \sim p_X(x)$, y g una función de interés, también llamada función de prueba, entonces el valor esperado de $g(X)$ es

$$\mathbb{E}[g(X)] = \int g(x) p_X(x)dx$$

El valor esperado permite calcular estimadores puntuales, por ejemplo, un estimador puntual de los estados X_t (o parámetros θ) dadas T observaciones $y_{1:T} = \{y_1, y_2, \dots, y_T\}$, esto es,

$$\mathbb{E}[X_t|y_{1:T}] = \int x_t p(x_t|y_{1:T})dx_t \quad \left(\text{o } \mathbb{E}[\theta|y_{1:T}] = \int \theta p(\theta|y_{1:T})d\theta \right)$$

2. Marginalización

Sean X y Y dos variables aleatorias con densidad conjunta $p(x, y)$, entonces la densidad marginal de X (o Y) se puede calcular como

$$p_X(x) = \int p(x, y) dy \quad \left(\text{o } p_Y(y) = \int p(x, y) dx \right)$$

La marginalización puede interpretarse como un promedio de la densidad conjunta sobre X (o Y). Un ejemplo de marginalización surge en los modelos probabilísticos espacio estado cuando se calcula la verosimilitud marginal $p(y_{1:T})$

$$p(y_{1:T}) = \int p(y_{1:T}|\theta)p(\theta)d\theta$$

La idea principal del método de Monte Carlo se basa en considerar una variable aleatoria X junto con su respectiva densidad $p_X(x)$, que suele ser llamada densidad objetivo, así $x \sim p_X(x)$, y aproximar la integral del valor esperado de alguna función de prueba g , que será denotada como I

$$I(g) = \mathbb{E}_p[g(X)] = \int g(x) p_X(x) dx$$

Supongamos que podemos generar muestras independientes $\{x^i\}_{i=1}^N$ distribuidas de acuerdo a $p_X(x)$. A partir de estas muestras, aproximamos la integral I por el promedio muestral $\bar{I}^N(g)$, que es conocido como estimador de Monte Carlo (Schön y Lindsten, 2017).

$$I(g) \approx \frac{1}{N} \sum_{i=1}^N g(x^i) = \bar{I}^N(g)$$

Entre las principales propiedades del estimador de Monte Carlo se tiene que es un estimador insesgado, esto es, $\mathbb{E}[\bar{I}^N(g)] = I(g)$ (Luengo et al., 2020). Segundo, la ley de los grandes números implica convergencia casi segura al valor esperado verdadero (Schön y Lindsten, 2017).

$$\bar{I}^N(g) \rightarrow I(g), \quad N \rightarrow \infty$$

En el Algoritmo 1, se muestra el pseudocódigo del algoritmo de Monte Carlo.

Algoritmo 1. Algoritmo Monte Carlo

Entrada: Número de muestras N y densidad $p_X(x)$.

Salida: Aproximación $\bar{I}^N(g)$.

1. Generar muestras $x^i \sim p_X(x)$
2. Aproximar la integral $I(g)$ por

$$\bar{I}^N(g) = \frac{1}{N} \sum_{i=1}^N g(x^i)$$

2.9 Muestreo por rechazo

Un inconveniente con los métodos de Monte Carlo que suele ocurrir en situaciones prácticas es que la densidad objetivo $p_X(x)$ no es conocida, sino hasta una proporcionalidad, esto es, $p_X(x) = \tilde{p}(x)/Z$, donde $\tilde{p}(x)$ es una densidad simple de evaluar y Z es una constante de normalización independiente de x que es desconocida. El método de muestreo por rechazo, también llamado muestreo por aceptación – rechazo, permite generar muestras de $p_X(x)$, a partir de una densidad de propuesta simple $q(x)$ (Luengo et al., 2020). La idea básica detrás del muestreo por rechazo consiste en la generación de muestras a partir de la densidad de propuesta, esto es, $x^i \sim q(x)$ y aceptarlas con probabilidad

$$p_A(x^i) = \frac{\tilde{p}(x^i)}{Cq(x^i)} \leq 1$$

donde C es una constante tal que $Cq(x)$ es una función de envoltura para $\tilde{p}(x)$, esto significa que $p(x) \leq Cq(x)$, para $\forall x$. En el Algoritmo 2, se resume la generación de N muestras de la densidad objetivo usando el método de muestreo por rechazo.

Ejemplo. Muestreo de la densidad a posteriori no normalizada

En muchos casos prácticos no es posible generar muestras de la densidad a posteriori $p(\theta|y_{1:t}) = \gamma(\theta)/Z_\theta$ directamente. En estos casos, es posible evaluar la densidad no normalizada $\gamma(\theta)$ (densidad objetivo), por lo que el método de muestreo por rechazo es aplicable si consideramos una densidad de propuesta, $q(\theta)$, de donde generar muestras, esto es, $\tilde{\theta} \sim q(\theta)$.

Algoritmo 2. Muestreo por rechazo

Entrada: Densidad de propuesta $q(x)$, densidad no normalizada $\tilde{p}(x)$, número de muestras N y una constante C tal que $\tilde{p}(x) \leq Cq(x)$, $\forall x$.

Salida: Muestras $\{x^1, x^2, \dots, x^N\}$ de la densidad objetivo $p_X(x)$

1. Para $n = 1$ hasta N hacer
2. No_encontrado \leftarrow Verdadero
3. Mientras no_encontrado hacer
4. Generar una muestra $\tilde{x} \sim q(x)$
5. Generar una muestra $u \sim \mathcal{U}[0,1]$ # Distribución uniforme
6. Si $u \leq \tilde{p}(x)/Cq(x)$ entonces
7. $x^i = \tilde{x}$
8. No_encontrado \leftarrow Falso

Si bien el método de muestreo por rechazo presenta una estructura fácil de programar, no obstante, presenta algunos inconvenientes al momento de implementarlo en problemas prácticos (Luengo et al., 2020).

- 1) Dificultad para encontrar la constante C , especialmente para problemas con alta dimensionalidad.

- 2) El método de muestreo por rechazo puede ser ineficiente si la razón $\tilde{p}(x)/Cq(x)$ es muy pequeña.
- 3) El número de iteraciones en el lazo *Mientras ... hacer* necesario para generar N muestras, T , es una variable aleatoria con valor esperado $\mathbb{E}[T] = N/p_A(\cdot)$ donde $p_A(\cdot)$ es la probabilidad de aceptación. Por lo tanto, el número de iteraciones T requerido para generar las N muestras no es conocido, esto suele ser un inconveniente en muchas aplicaciones.

2.9.1 Muestreo por importancia (MI)

Además del método de muestreo por rechazo, existe otro método de muestreo llamado método de muestreo por importancia (MI). Este se caracteriza por generar muestras de una densidad de propuesta $x \sim q(x)$ (también llamada densidad de importancia), pero a diferencia del método de muestreo de rechazo que descarta algunos valores simulados, todas las muestras son usadas. Las muestras generadas por el método de muestreo por importancia son asignadas con pesos individuales dependiendo de que tan bien se ajustan a la densidad objetivo (Schön y Lindsten, 2017).

A continuación, veremos la derivación del método de muestreo por importancia. Sea $X' \sim q(x)$ una variable aleatoria distribuida de acuerdo a la densidad de importancia $q(x)$. Entonces, la integral $I(g) = \mathbb{E}_p[g(X)]$ se puede escribir como

$$\begin{aligned}
 \mathbb{E}_p[g(X)] &= \int_{\mathcal{X}} g(x)p_X(x)dx \\
 &= \int_{\mathcal{X}} g(x)\frac{p_X(x)}{q(x)}q(x)dx \\
 &= \int_{\mathcal{X}} g(x)\omega(x)q(x)dx \\
 &= \mathbb{E}_q[g(X')\omega(X')]
 \end{aligned}$$

donde $\omega(x) = p_X(x)/q(x)$ es conocida como la función de peso y donde hemos asumido que $q(x) > 0$ para $\forall x$ donde $p_X(x) > 0$. Así hemos escrito el valor esperado de $g(x)$ como el valor esperado de la variable X' mapeada a través de otra función. Esto permite un camino factible para aproximar $\mathbb{E}_p[g(X)]$ (Schön y Lindsten, 2017). Por construcción, se puede generar muestras de $q(x)$ por lo tanto un estimador Monte Carlo para $\mathbb{E}_q[g(X')\omega(X')]$ se consigue obteniendo N muestras de $q(x)$, esto es, $x^i \sim q(x)$, $i = 1, \dots, N$ y

$$\bar{I}_{\text{MI}}^N(g) = \frac{1}{N} \sum_{i=1}^N \omega(x^i) g(x^i)$$

Sin embargo, la densidad objetivo $p_X(x)$ no es conocida sino hasta una constante de normalización Z , usualmente desconocida. Esto nos lleva a considerar la siguiente derivación

$$\begin{aligned} I(g) &= \mathbb{E}_p[g(x)] \\ &= \int_X g(x) \frac{\tilde{p}(x)}{Z} dx \\ &= \frac{1}{Z} \int_X g(x) \frac{\tilde{p}(x)}{q(x)} q(x) dx \\ &= \frac{1}{Z} \int_X g(x) \omega(x) q(x) dx \\ &= \frac{1}{Z} \mathbb{E}_q[g(x) \omega(x)] \end{aligned}$$

donde $\omega(x) = \tilde{p}(x)/q(x)$. Si consideramos las muestras de importancia $x^i \sim q(x)$, $i = 1, \dots, N$, y definimos $w^i = \omega(x^i)$, podemos construir un estimador del método de muestreo por importancia como

$$\bar{I}_{\text{MI}}^N(g) = \frac{1}{NZ} \sum_{i=1}^N w^i g(x^i)$$

Notamos que w^i es una variable aleatoria ya que es el resultado de la transformación de otra variable aleatoria.

A continuación, vamos a considerar una aproximación de la constante Z para lo cual usamos la densidad de importancia como sigue

$$Z = \int_X \tilde{p}(x) dx = \int_X \frac{\tilde{p}(x)}{q(x)} q(x) dx = \int_X w^i q(x) dx \approx \frac{1}{N} \sum_{i=1}^N w^i$$

Por lo tanto, el estimador del método de muestreo por importancia se puede escribir como

$$\bar{I}_{\text{MI}}^N(g) = \frac{\frac{1}{N} \sum_{i=1}^N w^i g(x^i)}{\frac{1}{N} \sum_{i=1}^N w^i} = \sum_{i=1}^N \bar{w}^i g(x^i)$$

donde $\bar{w}^i = w^i / \sum_{j=1}^N w^j$ representa un peso normalizado.

En el estimador del método de muestreo por importancia, todas las muestras generadas x^i son ponderadas por los pesos w^i . Estos pesos permiten cuantificar la discrepancia entre la densidad de importancia y la densidad objetivo (Schön y Lindsten, 2017).

En el Algoritmo 3, se muestra el pseudocódigo del método de muestreo por importancia.

Algoritmo 3. Muestreo por importancia.

Entrada: Densidad de propuesta $q(x)$, densidad no normalizada $\tilde{p}(x)$, número de muestras N

Salida: Muestras $\{x^i\}_{i=1}^N$ junto con los pesos normalizados $\{\bar{w}^i\}_{i=1}^N$.

1. Generar muestras de $x^i \sim q(x)$, $\forall i = 1, \dots, N$.

2. Calcular los pesos de x^i como

$$w^i = \frac{\tilde{p}(x^i)}{q(x^i)}, \quad \forall i = 1, \dots, N$$

3. Normalizar los pesos

$$\bar{w}^i = \frac{w^i}{\sum_{j=1}^N w^j}, \quad \forall i = 1, \dots, N$$

2.9.2 Evaluación numérica estable de los pesos

Para evitar imprecisiones numéricas o errores de ejecución debidos a divisiones para cero durante la normalización de los pesos, debidas a pesos muy pequeños, conviene considerar el logaritmo de los pesos $\log w^i$ (Schön y Lindsten, 2017).

Asumamos que $\log w^i$ es el logaritmo del peso w^i para $i = 1, \dots, N$. Luego, se calcula el máximo de los logaritmos de los pesos

$$m = \max_{i=1, \dots, N} \{\log w^i\}$$

y luego sustraer m de todos los logaritmos de los pesos $\widetilde{\log w}^i = \log w^i - m$ para $i = 1, \dots, N$. A continuación, normalizamos los logaritmos de los pesos como sigue

$$\begin{aligned} v^i &= \frac{\exp(\widetilde{\log w}^i)}{\sum_{j=1}^N \exp(\widetilde{\log w}^j)} \\ &= \frac{\exp(\log w^i - m)}{\sum_{j=1}^N \exp(\log w^j - m)} \\ &= \frac{\exp(\log w^i) \exp(-m)}{\sum_{j=1}^N \exp(\log w^j) \exp(-m)} \\ &= \frac{\exp(-m) \exp(\log w^i)}{\exp(-m) \sum_{j=1}^N \exp(\log w^j)} \\ &= \frac{\exp(\log w^i)}{\sum_{j=1}^N \exp(\log w^j)} \\ &= \bar{w}^i \end{aligned}$$

Observamos que la substracción del máximo de los logaritmos de los pesos no afecta el peso relativo de los pesos normalizados \bar{w}^i .

Adicionalmente, los logaritmos de los pesos normalizados se pueden escribir como

$$\log \bar{w}^i = \widetilde{\log w}^i - \log \sum_{j=1}^N \exp(\widetilde{\log w}^j)$$

donde el logaritmo del segundo término está bien definido ya que la suma tiene una cota inferior en 1.

Además, el cálculo del logaritmo de la constante de normalización para $\tilde{p}(x)$, es decir, $\log Z$ se puede calcular usando el logaritmos como sigue

$$\log Z = \log \left(\frac{1}{N} \sum_{i=1}^N w^i \right) = \log \sum_{i=1}^N \exp(\widetilde{\log w}^i) + m - \log N$$

2.9.3 Método de muestreo por importancia secuencial para un modelo espacio estado

Ahora, consideramos el problema de aproximar las ecuaciones de filtrado de un modelo espacio estado mediante el método de muestreo por importancia. La solución a este problema es el algoritmo de muestreo por importancia secuencial. Para derivar este método, fijamos los parámetros θ y consideramos la densidad de filtrado, para $t = 1, \dots, T$

$$\begin{aligned} p(x_{0:t}|y_{1:t}) &\propto p(y_{1:t}|x_{0:t})p(x_{0:t}) \\ &= p(y_t, y_{1:t-1}|x_{0:t})p(x_{0:t}) \\ &= p(y_t|y_{1:t}, x_{0:t})p(y_{1:t-1}|x_{0:t})p(x_{0:t}) \\ &\propto p(y_t|x_t)p(x_{0:t}|y_{1:t-1}) \\ &= p(y_t|x_t)p(x_t, x_{0:t-1}|y_{1:t-1}) \\ &= p(y_t|x_t)p(x_t|x_{0:t-1}, y_{1:t-1})p(x_{0:t-1}|y_{1:t-1}) \\ &= p(y_t|x_t)p(x_t|x_{t-1})p(x_{0:t-1}|y_{1:t-1}) \end{aligned}$$

Si consideramos una densidad de importancia tal que $x_{0:t}^i \sim q(x_{0:t}|y_{1:t})$, para $\forall i = 1, \dots, N$ podemos escribir los pesos de importancia como

$$w_t^i \propto \frac{p(y_t|x_t^i)p(x_t^i|x_{t-1}^i)p(x_{0:t-1}^i|y_{1:t-1})}{q(x_{0:t}^i|y_{1:t})}, \quad \forall i = 1, \dots, N$$

Tomando una densidad de importancia para los estados x_t que se pueda factorizar como sigue

$$q(x_{0:t}|y_{1:t}) = q(x_t|x_{0:t-1}, y_{1:t})q(x_{0:t-1}|y_{1:t-1})$$

Entonces, la expresión de los pesos se puede reescribir como

$$\begin{aligned} w_t^i &\propto \frac{p(y_t|x_t^i)p(x_t^i|x_{t-1}^i)p(x_{0:t-1}^i|y_{1:t-1})}{q(x_{0:t}^i|y_{1:t})} \\ &= \frac{p(y_t|x_t^i)p(x_t^i|x_{t-1}^i)}{q(x_t^i|x_{0:t-1}^i, y_{1:t})} \cdot \frac{p(x_{0:t-1}^i|y_{1:t-1})}{q(x_{0:t-1}^i|y_{1:t-1})}, \quad \forall i = 1, \dots, N \end{aligned}$$

Si asumimos que hemos generados una muestra $x_{0:t-1}^i$ de la densidad de importancia $q(x_{0:t-1}^i|y_{1:t-1})$ y calculado los correspondientes pesos w_{t-1}^i , entonces podemos obtener las muestras $x_{0:t}^i$ generadas de la densidad de importancia $q(x_{0:t}^i|y_{1:t})$ mediante la actualización de nuevas muestras para el paso t , esto es, $x_t^i \sim q(x_t|x_{0:t-1}, y_{1:t})$. Los pesos del paso previo $t - 1$, se pueden expresar como

$$w_{t-1}^i \propto \frac{p(x_{0:t-1}^i|y_{1:t-1})}{q(x_{0:t-1}^i|y_{1:t-1})}$$

Por lo que los pesos satisfacen la recursión

$$w_t^i \propto \frac{p(y_t|x_t^i)p(x_t^i|x_{t-1}^i)}{q(x_t^i|x_{0:t-1}^i, y_{1:t})} w_{t-1}^i$$

Algoritmo 4. Muestreo por importancia secuencial

Entrada: Observaciones $y_{1:T} = \{y_1, y_2, \dots, y_T\}$, densidad de propuesta $q(x_t|x_{t-1}, y_{1:t})$, densidad de observación $p(y_t|x_t)$, densidad de transición $p(x_t|x_{t-1})$, número de partículas N , estados iniciales $\{x_0^i\}_{i=1}^N$ (estos estados pueden ser inicializados a partir de una densidad inicial $x_0^i \sim p(x_0)$)

Salida: Muestras $\{x_t^i\}_{i=1}^N$ junto con los pesos normalizados $\{\bar{w}_t^i\}_{i=1}^N$.

1. Inicializar $x_1^i = x_0^i$, $w_1^i = p(y_1|x_1^i)$ y $\bar{w}_1^i = 1/N$, $\forall i = 1, \dots, N$
2. Para $t = 2$ hasta T hacer
3. Generar muestras de la densidad de importancia

$$x_t^i \sim q(x_t^i|x_{t-1}^i, y_{1:t}), \quad \forall i = 1, \dots, N$$

4. Calcular los pesos de x_t^i como

$$w_t^i = w_{t-1}^i \frac{p(y_t|x_t^i)p(x_t^i|x_{t-1}^i)}{q(x_t^i|x_{t-1}^i, y_{1:t})}, \quad \forall i = 1, \dots, N$$

5. Normalizar los pesos

$$\bar{w}_t^i = \frac{w_t^i}{\sum_{j=1}^N w_t^j}, \quad \forall i = 1, \dots, N$$

2.9.4 Proceso de remuestreo

Un problema en los algoritmos de muestreo por importancia secuencial es que puede suceder que las partículas tengan un peso de cero o cercano a cero. Este problema es conocido en la literatura de los filtros de partículas como problema de degeneración. Una solución para este problema de degeneración es el proceso de remuestreo. El remuestreo consiste en generar N muestras de la distribución discreta definida por los pesos y

reemplazar el conjunto antiguo de N muestras por el conjunto de muestras nuevas. En resumen, el proceso de remuestreo remueve partículas con pesos bajos y multiplica las partículas con pesos altos (Särkkä, 2013).

Algoritmo 5. Remuestreo

Entrada: Conjunto de partículas antiguas $\{x_t^{i-\text{antiguas}}\}_{i=1}^N$ y sus respectivos pesos normalizados $\{\bar{w}_t^{i-\text{antiguos}}\}_{i=1}^N$.

Salida: Conjunto de partículas nuevas (remuestreadas) $\{x_t^{i-\text{nuevas}}\}_{i=1}^N$ y sus respectivos pesos normalizados $\{\bar{w}_t^{i-\text{nuevos}}\}_{i=1}^N$.

1. Interpretar cada uno de los pesos $\bar{w}_t^{i-\text{antiguos}}$ como la probabilidad de obtener el índice i en el conjunto de partículas $\{x_t^{i-\text{antiguas}}\}_{i=1}^N$.
2. Generar un nuevo conjunto de N partículas de la distribución discreta definida por los pesos, esto es, $\{x_t^{i-\text{nuevas}}\}_{i=1}^N$ y reemplazar el conjunto de muestras antiguas con el conjunto de muestras nuevas.
3. Igualar todos los pesos al valor constante $\bar{w}_t^{i-\text{nuevos}} = 1/N$.

En general, el proceso de remuestreo puede ser aplicado dentro de un muestreo por importancia secuencial. Sin embargo, el paso de remuestreo puede ser una fuente adicional de varianza. Si las partículas tienen pesos no normalizados con baja varianza, entonces el paso de remuestreo puede ser innecesario. Por lo tanto, en casos prácticos es conveniente aplicar el paso de remuestreo sólo cuando la varianza de los pesos no normalizados es superior a un umbral pre establecido (Doucet y Johansen, 2009). Esto último es posible de

evaluar observando la variabilidad de los pesos no normalizados mediante el número de muestra efectivo dado por

$$n_{\text{efectivo}} = \frac{1}{\sum_{i=1}^N (\bar{w}_t^i)^2}$$

donde \bar{w}_t^i es el peso normalizado de la partícula i en el tiempo t . El paso de remuestreo se aplica cuando el número de muestra efectivo es significativamente menor al número total de partículas N , por ejemplo, $n_{\text{efectivo}} < N/2$ (Doucet y Johansen, 2009) o $n_{\text{efectivo}} < N/10$ (Särkkä, 2013).

2.10 Filtros de partículas

Los filtros de partículas, también llamados métodos de Monte Carlo secuencial, son métodos de Monte Carlo que permiten aproximar la solución del problema de filtrado no lineal, donde la exactitud de la aproximación está sólo limitada por la capacidad computacional. Adicionalmente, los filtros de partículas permiten calcular una estimación insesgada de la verosimilitud marginal $p(y_{1:T}|\theta)$, a partir de un sistema de partículas ponderadas.

$$p(y_{1:T}|\theta) \approx p^N(y_{1:T}|\theta) = \prod_{t=1}^T p^N(y_t|y_{1:t-1}, \theta) = \prod_{t=1}^T \left[\frac{1}{N} \sum_{i=1}^N w_t^i \right]$$

El algoritmo de filtro de partícula se puede diseñar combinando el muestreo por importancia secuencial con un paso de remuestreo. En el Algoritmo 6 se resume la estructura del filtro de partículas.

Algoritmo 6. Filtro de partículas

Entrada: Observaciones $y_{1:T} = \{y_1, y_2, \dots, y_T\}$, densidad de propuesta $q(x_t|x_{t-1}, y_{1:t}, \theta)$, densidad de observación $p(y_t|x_t, \theta)$, densidad de transición $p(x_t|x_{t-1}, \theta)$, número de partículas N , estados iniciales $\{x_0^i\}_{i=1}^N$ (estos estados pueden ser inicializados a partir de una densidad inicial $x_0^i \sim p(x_0)$)

Salida: Muestras $\{x_t^i\}_{i=1}^N$ junto con los pesos normalizados $\{\bar{w}_t^i\}_{i=1}^N$, estimación de la verosimilitud \hat{Z}

1. Inicializar $x_1^i = x_0^i$, $w_1^i = p(y_1|x_1^i, \theta)$, $\bar{w}_1^i = 1/N$, $\forall i = 1, \dots, N$ y $\hat{Z} = \frac{1}{N} \sum_{i=1}^N w_1^i$
2. Para $t = 2$ hasta T hacer
3. Remuestrear N muestras de $\{x_t^i\}_{i=1}^N$ de acuerdo a $\{\bar{w}_t^i\}_{i=1}^N$
4. Generar muestras de la densidad de importancia

$$x_t^i \sim q(x_t^i|x_{t-1}^i, y_{1:t}, \theta), \quad \forall i = 1, \dots, N$$

5. Calcular los pesos de x_t^i como

$$w_t^i = w_{t-1}^i \frac{g(y_t|x_t^i, \theta)f(x_t^i|x_{t-1}^i, \theta)}{q(x_t^i|x_{t-1}^i, y_{1:t}, \theta)}, \quad \forall i = 1, \dots, N$$

6. Normalizar los pesos

$$\bar{w}_t^i = \frac{w_t^i}{\sum_{j=1}^N w_t^j}, \quad \forall i = 1, \dots, N$$

7. Actualizar la estimación de la verosimilitud

$$\hat{Z} = \hat{Z} \times \frac{1}{N} \sum_{i=1}^N w_t^i$$

El filtro de partículas Bootstrap es caso particular del filtro de partículas del Algoritmo 6, donde la densidad de transición de los estados es tomada como la densidad de propuesta, esto es, $q(x_t^i|x_{t-1}^i, y_{1:t}, \theta) = f(x_t^i|x_{t-1}^i, \theta)$, por lo tanto, los pesos no normalizados se simplifican a $w_t^i = \bar{w}_{t-1}^i g(y_t|x_t^i, \theta)$ (Botha et al., 2021).

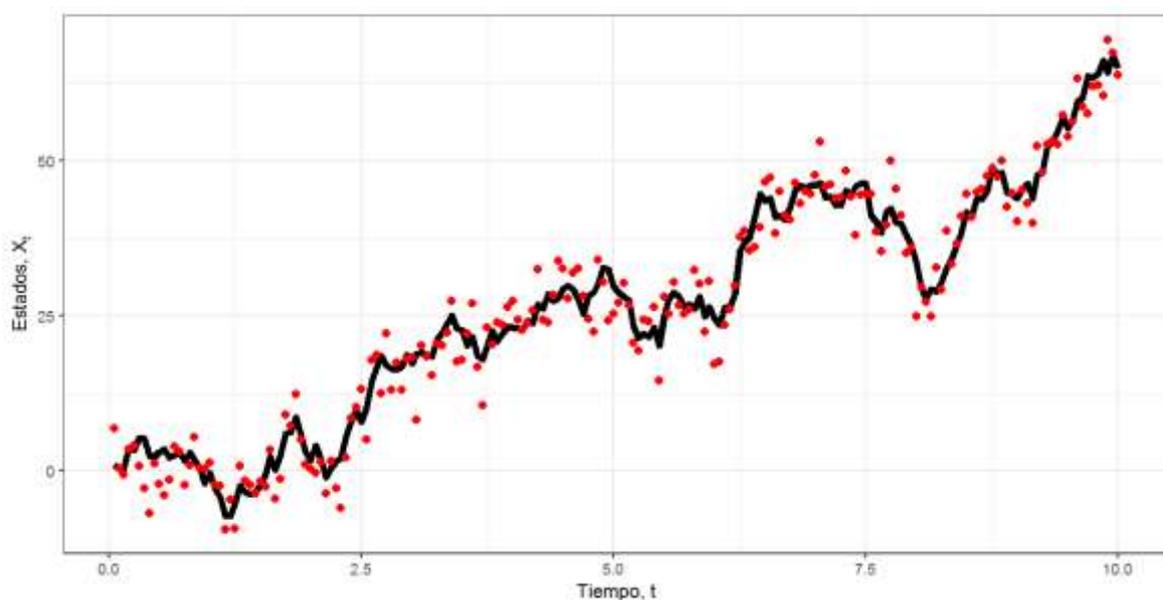
Ejemplo. Caminata aleatoria Gaussiana con un filtro de partícula

Para observar el desempeño del filtro de partículas para estimar los estados latentes de una caminata aleatoria Gaussiana unidimensional dadas las observaciones, creamos datos sintéticos del modelo espacio estado

$$\begin{aligned} x_t &= x_{t-1} + q_{t-1} & , & & q_{t-1} &\sim N(q_{t-1}; 0, s_x^2) \\ y_t &= x_t + r_t & , & & r_t &\sim N(r_t; 0, s_y^2) \end{aligned}$$

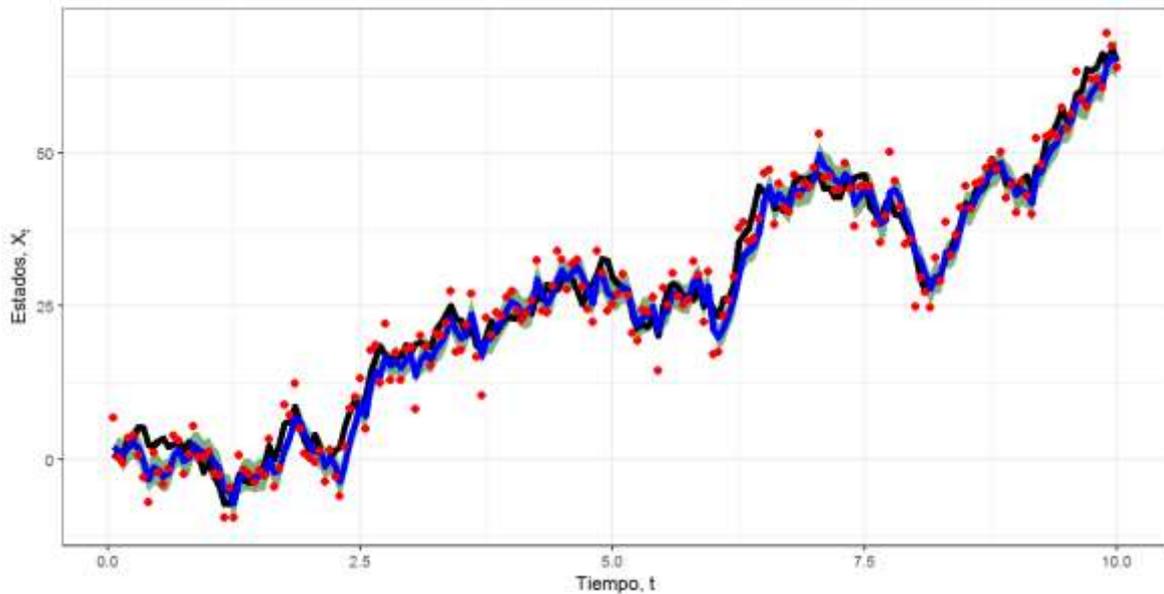
con los parámetros verdaderos $s_x = 2$, $s_y = 3.5$, $x_0 = 0$ con un tiempo de simulación $T = 10$ y número de tiempos de observación $n = 200$, esto significa que $\Delta t = T/n = 0.05$. En la Figura 6, se muestra la trayectoria de una caminata aleatoria Gaussiana (color negro) y las observaciones (puntos en color rojo) simulados con los parámetros verdaderos.

Figura 6. Datos simulados de la caminata aleatoria Gaussiana



Usando las observaciones de los estados y_t , ejecutamos el filtro de partículas Bootstrap con $N = 400$ partículas y parámetros $s'_x = 1.25$, $s'_y = 2$, $x'_0 = 2$. En la Figura 7, se observa la trayectoria de filtrado (color azul) y los intervalos de confianza (color verde).

Figura 7. Caminata aleatoria Gaussiana y trayectoria de filtrado



2.11 Muestreo de Gibbs

El muestreo de Gibbs es un método de Monte Carlo por cadena de Markov que se caracteriza por generar muestras de una distribución conjunta de los parámetros a partir de las distribuciones condicionales dados los otros parámetros. El muestreo de Gibbs se puede definir en términos de los subvectores de θ . Dado el vector de parámetros θ , se procede a dividirlo en D componentes, esto es, $\theta = \{\theta_1, \theta_2, \dots, \theta_D\}$. Cada iteración del método de muestreo de Gibbs ingresa a un ciclo que recorre las componentes de θ , generando componentes de θ condicionados en los valores de las demás componentes. Así, en cada iteración i , el método de muestreo de Gibbs realiza d pasos (Gelman et al., 2014). A continuación, se muestra el Algoritmo 7, donde se indican los pasos del muestreo de Gibbs.

Algoritmo 7. Muestreo de Gibbs

Entrada: Valores iniciales θ^0 , número de iteraciones I , periodo de quemado I_b , observaciones $y_{1:T} = \{y_1, y_2, \dots, y_T\}$,

Salida: Cadena de Markov después del periodo de quemado, esto es, $\{\theta^{I_b+1}, \dots, \theta^I\}$

1. Inicializar $\theta^1 = \theta^0$
2. Para $i = 1$ hasta I hacer

Actualizar θ_j^{i+1} condicionado en $\theta_{-j}^i = (\theta_1^{i+1}, \theta_2^{i+1}, \dots, \theta_{j-1}^{i+1}, \theta_{j+1}^i, \dots, \theta_D^i)$ y las observaciones $y_{1:T}$ (para $j = 1, \dots, D$)

2.12 Algoritmo Metropolis – Hastings

Los métodos de Monte Carlo por cadena de Markov, en inglés Markov chain Monte Carlo (MCMC), son una familia de algoritmos para generar variables aleatorias de una densidad dada mediante la simulación de una cadena de Markov ergódica que tiene la densidad requerida como su densidad estacionaria. Estos métodos son adecuados para simular muestras de una densidad a posteriori $p(\theta|y_{1:T})$, para implementar estos métodos sólo es necesario conocer la densidad a posteriori no normalizada $\gamma(\theta) = p(y_{1:T}|\theta)p(\theta)$ (o equivalentemente $\varphi_T(\theta) = \log \gamma(\theta) = \log p(y_{1:T}|\theta) + \log p(\theta)$, donde $\varphi_T(\theta)$ es el logaritmo de la densidad a posteriori no normalizada) (Särkkä, 2013). El algoritmo Metropolis – Hastings (MH) es un método MCMC estándar que acepta valores candidatos $\theta' \sim q(\theta'|\theta)$ con probabilidad

$$\begin{aligned}
\alpha(\theta'|\theta) &= \min \left[1, \frac{\gamma(\theta') q(\theta|\theta')}{\gamma(\theta) q(\theta'|\theta)} \right] \\
&= \min \left[1, \frac{p(y_{1:T}|\theta') p(\theta') q(\theta|\theta')}{p(y_{1:T}|\theta) p(\theta) q(\theta'|\theta)} \right] \\
&= \min \left[1, (\varphi_T(\theta') - \varphi_T(\theta)) \frac{q(\theta|\theta')}{q(\theta'|\theta)} \right]
\end{aligned}$$

En el Algoritmo 8 se resume el algoritmo Metropolis – Hastings con probabilidad de aceptación $\alpha(\theta'|\theta)$. El periodo de quemado, en inglés burn – in period, I_b , es el número de muestras iniciales removidas para garantizar que la cadena ha convergido aproximadamente a su densidad estacionaria (Luengo et al., 2020).

Algoritmo 8. Algoritmo Metropolis – Hastings (MH)

Entrada: Densidad de propuesta $q(\theta'|\theta)$, iteración inicial θ^0 (generada de una densidad inicial arbitraria o inicializada con un valor arbitrario), número de iteraciones I y periodo de quemado I_b , observaciones $y_{1:T} = \{y_1, y_2, \dots, y_T\}$.

Salida: Cadena de Markov después del periodo de quemado, esto es, $\{\theta^{I_b+1}, \dots, \theta^I\}$

1. Inicializar $\theta^1 = \theta^0$
2. Calcular una estimación de $p(\theta^1|y_{1:T}) \propto p(y_{1:T}|\theta^1)p(\theta^1)$
3. Para $i = 2$ hasta I hacer
4. Generar una muestra de valores candidatos θ' de la densidad de propuesta $q(\theta'|\theta^{i-1})$, esto es, $\theta' \sim q(\theta'|\theta^{i-1})$
5. Calcular la probabilidad de aceptación:

$$\alpha_i = \alpha(\theta', \theta^{i-1}) = \min \left[1, \frac{p(y_{1:T}|\theta') p(\theta') q(\theta^{i-1}|\theta')}{p(y_{1:T}|\theta^{i-1}) p(\theta^{i-1}) q(\theta'|\theta^{i-1})} \right]$$

6. Generar una variable aleatoria uniforme $u \sim \mathcal{U}[0,1]$ y establecer:

$$\theta^i = \begin{cases} \theta', & \text{si } u \leq \alpha_i \\ \theta^{i-1}, & \text{en otro caso} \end{cases}$$

7. Descartar las primeras I_b iteraciones de quemado.

Una de las ventajas del algoritmo MH es que es un método genérico que se puede aplicar casi cualquier densidad de propuesta y densidad objetivo. Sin embargo, la velocidad de convergencia y la calidad de los estimadores del algoritmo MH dependen de la calidad de la densidad de propuesta (Luengo et al., 2020). Algunas elecciones de densidades de propuesta pueden conllevar que las muestras estén altamente correlacionadas o que la tasa de rechazo sea muy alta (Särkkä, 2013).

2.12.1 Algoritmo Metropolis

El algoritmo Metropolis es un caso especial del método Metropolis – Hastings, donde la densidad de propuesta es simétrica, esto es, $q(\theta'|\theta) = q(\theta|\theta')$ (Särkkä, 2013). En este caso la probabilidad de aceptación se reduce a

$$\alpha(\theta'|\theta) = \min \left[1, \frac{\gamma(\theta')}{\gamma(\theta)} \right] = \min \left[1, \frac{p(y_{1:T}|\theta')p(\theta')}{p(y_{1:T}|\theta)p(\theta)} \right] = \min [1, (\varphi_T(\theta') - \varphi_T(\theta))]$$

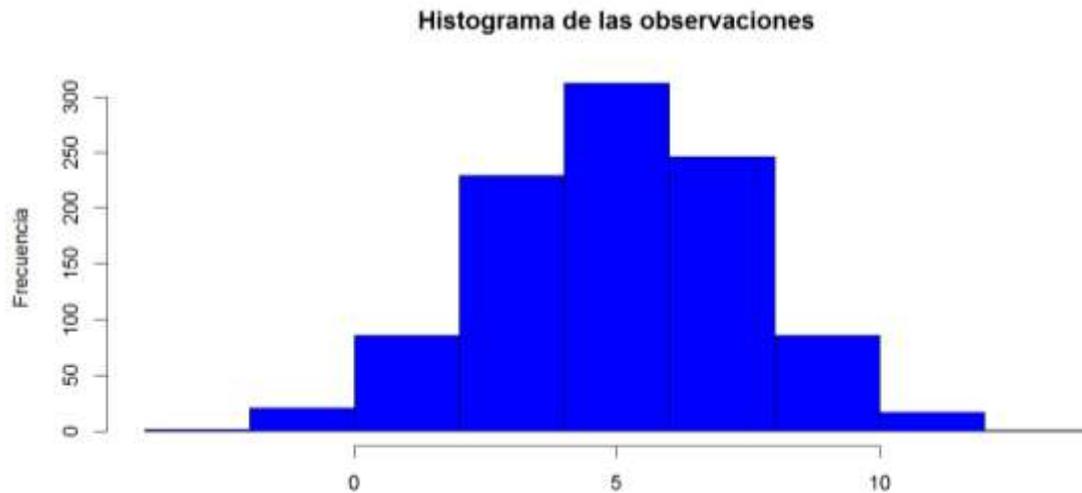
Si consideramos una densidad de propuesta Gaussiana (simétrica) $q(\theta'|\theta) = N(\theta'; \theta, \Sigma)$ con matrix de covarianza Σ , obtenemos el algoritmo Metropolis caminata aleatoria, o en inglés random walk Metropolis algorithm. La densidad de transición de este algoritmo se caracteriza por definir una caminata aleatoria Gaussiana en el espacio de los parámetros (Särkkä, 2013). Así, el valor candidato propuesto en la i -ésima iteración se puede escribir como $\theta' \sim \theta^i + \vartheta'$, donde $\vartheta' \sim N(\vartheta'; 0, \Sigma_i)$.

Ejemplo. Estimación de parámetros usando el algoritmo Metropolis.

Para mostrar el desempeño del algoritmo Metropolis para estimar parámetros dentro de un enfoque Bayesiano, vamos a simular observaciones $y_{1:T} = \{y_1, y_2, \dots, y_T\}$ de una

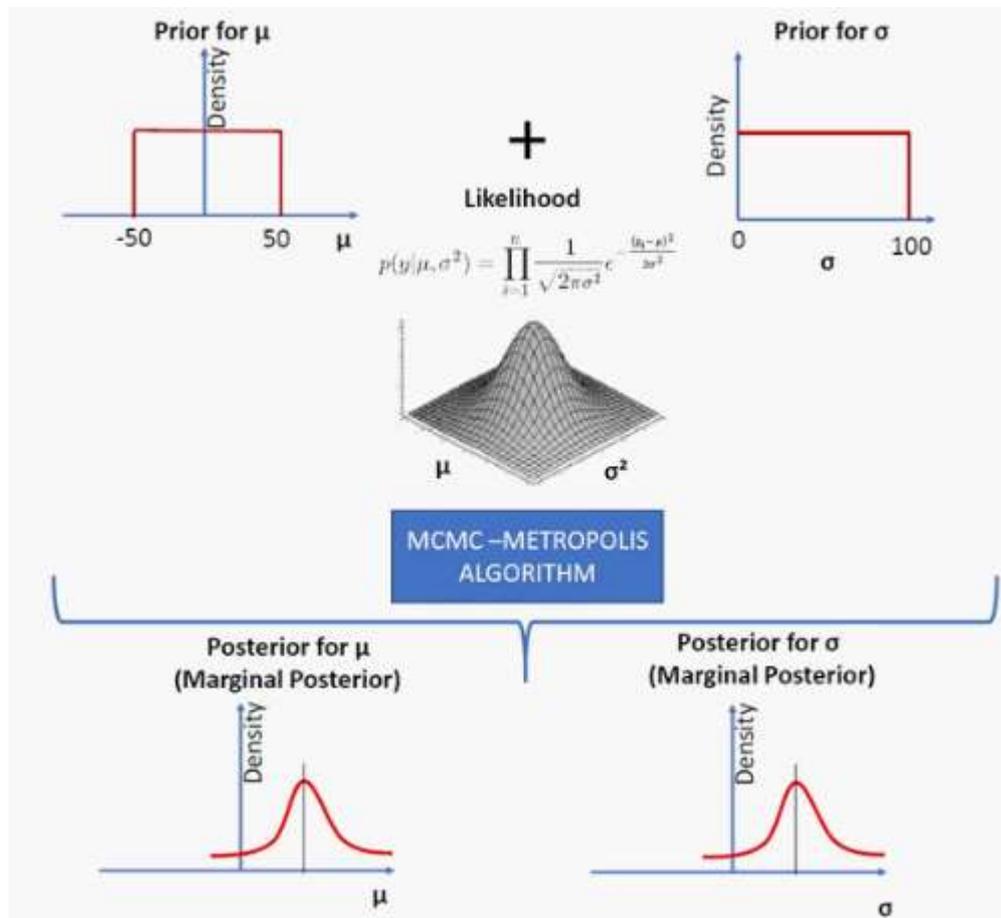
distribución Gaussiana con media $\mu = 5$ y desviación estándar $\sigma = 2.5$, así $y_t \sim N(y_t; \mu = 5, \sigma^2 = 2.5^2)$. En la Figura 8, se muestra el histograma de los datos simulados.

Figura 8. Histograma de las observaciones simuladas.



En la Figura 9, se muestra la estructura para estimar los parámetros θ del ejemplo bajo el enfoque Bayesiano. Además, se observan las densidades asignadas a las observaciones $y_t \sim N(y_t; \mu, \sigma)$ y las densidades a priori asignadas a los parámetros son $\mu \sim \mathcal{U}[\mu; -50, 50]$ y $\sigma \sim \mathcal{U}[\sigma; 0, 100]$.

Figura 9. Descripción general del algoritmo Metropolis para estimar parámetros (Rampinelli, 2019).



Para este ejemplo, consideramos que los parámetros $\theta = (\mu, \sigma)$ son desconocidos. Siguiendo el enfoque Bayesiano, vamos a generar muestras de la densidad a posteriori no normalizada, $\gamma(\theta) = p(y_{1:T}|\theta)p(\theta)$, ya que $p(\theta|y_{1:T}) \propto p(y_{1:T}|\theta)p(\theta)$. Así,

$$\begin{aligned} p(\theta|y_{1:T}) &= p(\mu, \sigma|y_{1:T}) \\ p(y_{1:T}|\theta) &= p(y_{1:T}|\mu, \sigma) \\ p(\theta) &= p(\mu, \sigma) \end{aligned}$$

Primero, consideramos la verosimilitud de los datos como sigue

$$p(y_{1:T}|\theta) = \prod_{t=1}^T N(y_t; \mu, \sigma) = \prod_{t=1}^T \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(y_t - \mu)^2}{\sigma^2}\right)$$

A efectos de conseguir estabilidad numérica, consideramos el logaritmo natural de la verosimilitud, esto es, $\ell(\theta) = \log p(y_{1:T}|\theta)$. Usando las propiedades de los logaritmos y acomodando los términos se obtiene

$$\begin{aligned}
 \ell(\theta) &= \log p(y_{1:T}|\theta) \\
 &= \log p(y_{1:T}|\mu, \sigma) \\
 &= \log \prod_{t=1}^T \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(y_t - \mu)^2}{\sigma^2}\right) \\
 &= \sum_{t=1}^T \log\left(\frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{1}{2} \frac{(y_t - \mu)^2}{\sigma^2}\right)\right) \\
 &= -\frac{T}{2} \cdot \log 2\pi\sigma^2 - \frac{1}{2\sigma^2} \sum_{t=1}^T (y_t - \mu)^2
 \end{aligned}$$

Asumiendo que los parámetros μ y σ son independientes, podemos escribir la densidad a priori como

$$p(\theta) = p(\mu, \sigma) \stackrel{\text{ind}}{=} p(\mu) \cdot p(\sigma)$$

Dado que la desviación estándar es estrictamente positiva $\sigma > 0$, consideramos la transformación $\kappa = \log \sigma$, por lo que la densidad a priori de los parámetros sería

$$\begin{aligned}
 p(\theta) &= p(\mu) \cdot p(\kappa) \\
 &= \mathcal{U}[\mu; -50, 50] \cdot (\mathcal{U}[\kappa; 0, 100] \cdot \exp(\kappa)) \\
 &= \frac{1}{50 - (-50)} \cdot \left(\frac{1}{100 - 0} \cdot \exp(\kappa)\right) \\
 &= \frac{1}{100} \cdot \frac{\exp(\kappa)}{100} \\
 &= \frac{\exp(\kappa)}{10000} \\
 &= \exp(\kappa) \cdot 10^{-4}
 \end{aligned}$$

Si consideramos el logaritmo natural $\log(\cdot)$ de la densidad a priori $p(\theta)$, se tiene

$$\log p(\theta) = \log(\exp(\kappa) \cdot 10^{-4}) = \kappa - 4 \log 10$$

Mientras que el logaritmo de la verosimilitud con la transformación $\kappa = \log \sigma$ (se tiene que $\sigma = \exp \kappa$)

$$\begin{aligned} \log p(y_{1:T}|\theta) &= \log p(y_{1:T}|\mu, \kappa) \\ &= -\frac{T}{2} \cdot \log(2\pi(\exp(\kappa))^2) - \frac{1}{2(\exp(\kappa))^2} \sum_{t=1}^T (y_t - \mu)^2 \\ &= -\frac{T}{2} \cdot \log 2\pi - \kappa T - \frac{1}{2 \exp(2 \cdot \kappa)} \sum_{t=1}^T (y_t - \mu)^2 \end{aligned}$$

Después de ejecutar el algoritmo Metropolis con $I = 20000$ iteraciones, la tasa de aceptación del método fue 28% aproximadamente. En las Figura 10 y Figura 11, se muestran los histogramas con las densidades empíricas de las cadenas de Markov de los parámetros μ y σ , las líneas continuas rojas representan los valores promedios de la cadena de Markov: $\mu_{\text{promedio}} = 5.03$ y $\sigma_{\text{promedio}} = 2.45$, y las líneas discontinuas azules indican el valor verdadero de los parámetros: $\mu_{\text{verdadero}} = 5$ y $\sigma_{\text{verdadero}} = 2.5$.

Figura 10. Histograma y diagrama de caja y bigotes de la media μ (Metropolis).

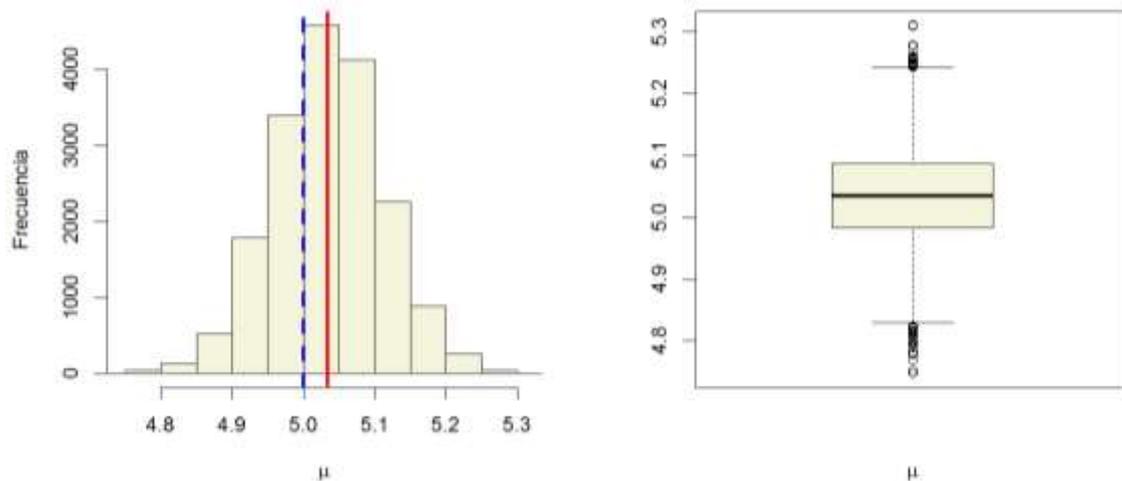
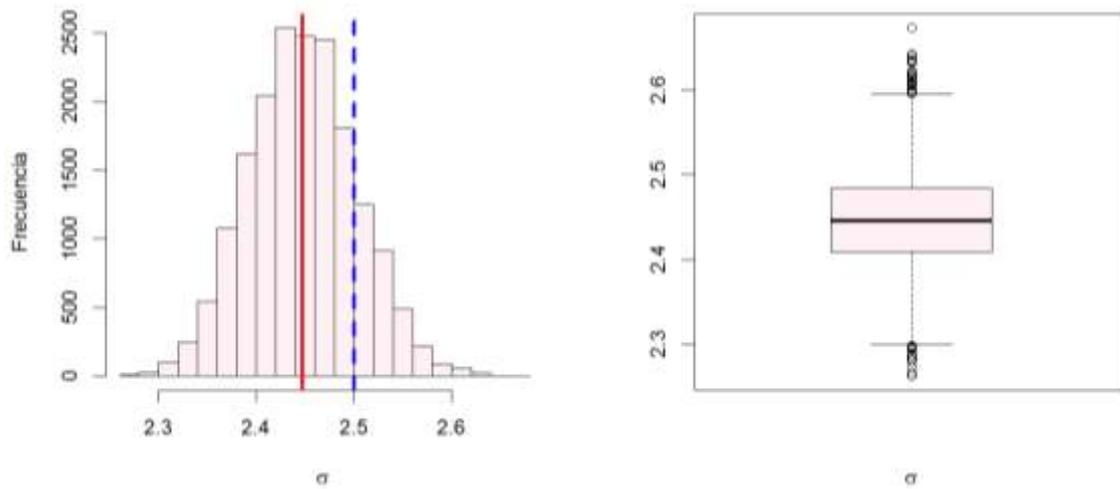
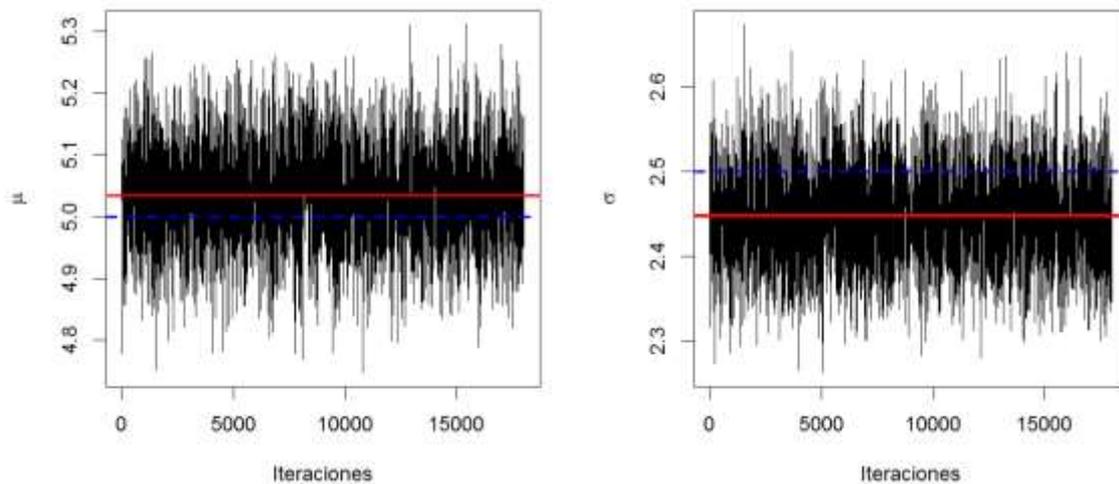


Figura 11. Histograma y diagrama de caja y bigotes de la desviación estándar σ (Metropolis).



En la Figura 12 se observan los diagramas de traza (líneas continuas negras) de las cadenas de Markov después del periodo de quemado, junto con el valor promedio de la cadena de Markov (líneas discontinuas azules) y el valor verdadero (líneas continuas rojas).

Figura 12. Diagrama de traza de los parámetros μ y σ (Metropolis).



En la Tabla 1, se muestra el resumen estadístico de las cadenas de Markov resultante del algoritmo Metropolis junto con los valores verdaderos.

Tabla 1. Resumen estadístico de la cadena de Markov (algoritmo Metropolis) correspondiente a los parámetros μ y σ .

(Valor verdadero)	$\mu = 5$	$\sigma = 2.5$
Media	5.03	2.45
Desviación estándar	0.08	0.06

2.12.2 Algoritmo MALA (*Metropolis adjusted Langevin algorithm*)

El algoritmo MALA es un método MCMC de muestreo que usa una EDE de difusión de Langevin para formar la densidad de propuesta. Este algoritmo forma parte de la familia de métodos basados en el gradiente del logaritmo de la densidad a posteriori, $\nabla \log p(\theta)$, para aumentar la eficiencia del proceso de muestreo. La EDE de tipo de difusión usada como densidad de propuesta es aproximada por medio del método de Euler-Maruyama (Luengo et al., 2020).

$$\theta' = \theta + \frac{\epsilon^2}{2} M \nabla_{\theta} \log p(\theta | y_{1:T}) + \epsilon \sqrt{M} z$$

donde $z \sim N(z; 0, I)$, ϵ es el tamaño del paso de discretización y M es una matriz de preconditionamiento que debe ser diagonalizable o tener una descomposición de Cholesky tal que $M = AA^T$ y $A = \sqrt{M}$. Esta matriz de preconditionamiento es usada para tomar en cuenta variables correlacionadas θ con varianzas muy diferentes (Girolami y Calderhead, 2011). Por lo tanto, la densidad de propuesta (no simétrica) del algoritmo MALA es

$$q(\theta' | \theta) = N\left(\theta'; \theta + \frac{\epsilon^2}{2} M \nabla_{\theta} \log p(\theta | y_{1:T}), \epsilon^2 M\right)$$

En el Algoritmo 9, se muestra del pseudocódigo del método MALA.

Algoritmo 9. Metropolis – Hastings Langevin Algorithm (MALA)

Entrada: Gradiente del logaritmo de la densidad a posteriori $\nabla_{\theta} \log p(\theta|y_{1:T})$, iteración inicial θ^0 (generada de una densidad inicial arbitraria o inicializada con un valor arbitrario), Matriz preconditionamiento M , paso de discretización ϵ , número de iteraciones I , periodo de quemado I_b , observaciones $y_{1:T} = \{y_1, y_2, \dots, y_T\}$.

Salida: Cadena de Markov después del periodo de quemado, esto es, $\{\theta^{I_b+1}, \dots, \theta^I\}$

1. Inicializar $\theta^1 = \theta^0$
2. Calcular una estimación de $p(\theta^1|y_{1:T}) \propto p(y_{1:T}|\theta^1)p(\theta^1)$
3. Para $i = 2$ hasta I hacer
4. Generar $z \sim N(z; 0, I)$ y simular una nueva muestra de la difusión de Langevin

$$\theta' = \theta^{i-1} + \frac{\epsilon^2}{2} M \nabla_{\theta} \log p(\theta^{i-1}|y_{1:T}) + \epsilon \sqrt{M} z$$

5. Calcular la probabilidad de aceptación:

$$\alpha_i = \alpha(\theta', \theta^{i-1})$$

$$= \min \left[1, \frac{p(y_{1:T}|\theta')p(\theta')}{p(y_{1:T}|\theta^{i-1})p(\theta^{i-1})} \frac{N\left(\theta^{i-1}; \theta' + \frac{\epsilon^2}{2} M \nabla_{\theta} \log p(\theta'|y_{1:T}), \epsilon^2 M\right)}{N\left(\theta'; \theta^{i-1} + \frac{\epsilon^2}{2} M \nabla_{\theta} \log p(\theta^{i-1}|y_{1:T}), \epsilon^2 M\right)} \right]$$

6. Generar una variable aleatoria uniforme $u \sim \mathcal{U}[0,1]$ y establecer:

$$\theta^i = \begin{cases} \theta', & \text{si } u \leq \alpha_i \\ \theta^{i-1}, & \text{en otro caso} \end{cases}$$

7. Descartar las primeras I_b iteraciones de quemado.

Ejemplo. Estimación de parámetros usando el algoritmo MALA.

Utilizando los datos del ejercicio anterior, se implementó el algoritmo MALA con $I = 20\,000$ iteraciones y tamaño de paso $\epsilon = 0.0383$, este último parámetro del algoritmo fue obtenido a través de experimentación. El gradiente del logaritmo de la densidad a posteriori $\nabla_{\theta} \log p(\theta|y_{1:T})$ donde $\theta = (\mu, \kappa)$ está dado por

$$\nabla_{\theta} \log p(\theta|y_{1:T}) \propto \nabla_{\theta} (\log p(y_{1:T}|\theta) + \log p(\theta)) = \begin{pmatrix} \frac{1}{\exp(2 \cdot \kappa)} \sum_{t=1}^T (y_t - \mu) \\ 1 - \kappa + \frac{1}{\exp(2 \cdot \kappa)} \sum_{t=1}^T (y_t - \mu)^2 \end{pmatrix}$$

Después de ejecutar el algoritmo MALA con $I = 20000$ iteraciones, la tasa de aceptación del método fue 59% aproximadamente. En las Figura 13 y Figura 14, se muestran los histogramas con las densidades empíricas de las cadenas de Markov de los parámetros μ y σ , las líneas continuas rojas representan los valores promedios de la cadena de Markov: $\mu_{\text{promedio}} = 5.04$ y $\sigma_{\text{promedio}} = 2.45$, y las líneas discontinuas azules indican el valor verdadero de los parámetros: $\mu_{\text{verdadero}} = 5$ y $\sigma_{\text{verdadero}} = 2.5$.

Figura 13. Histograma y diagrama de caja y bigotes de la media μ (MALA).

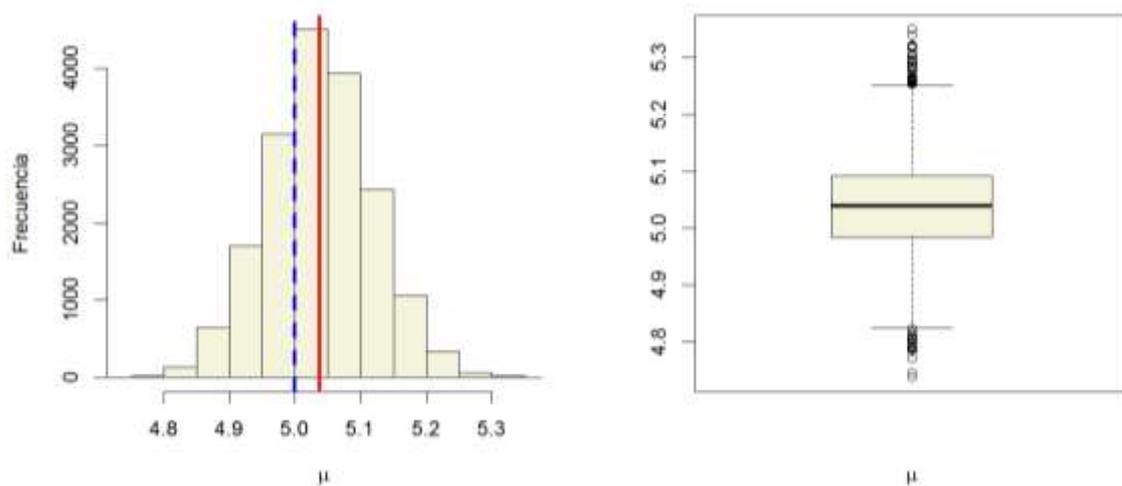
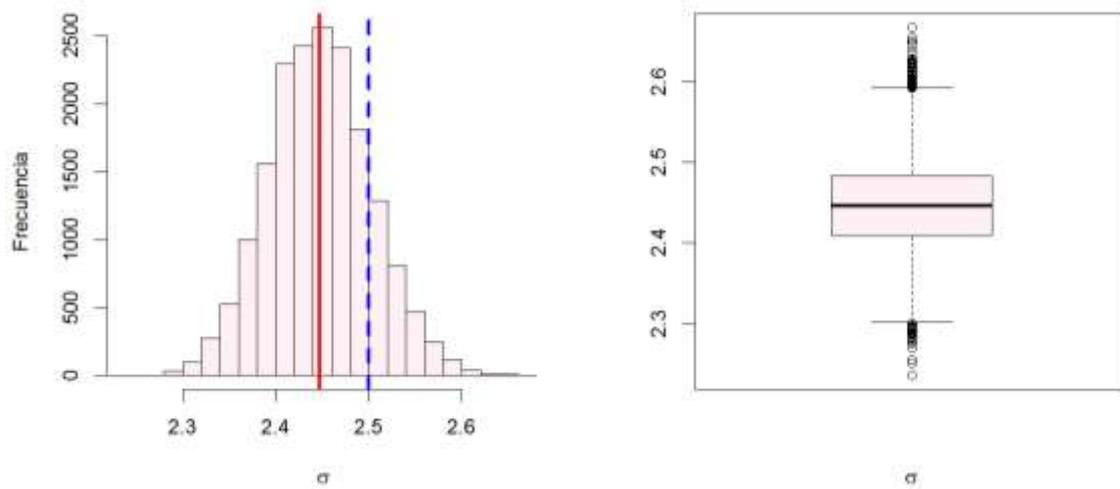
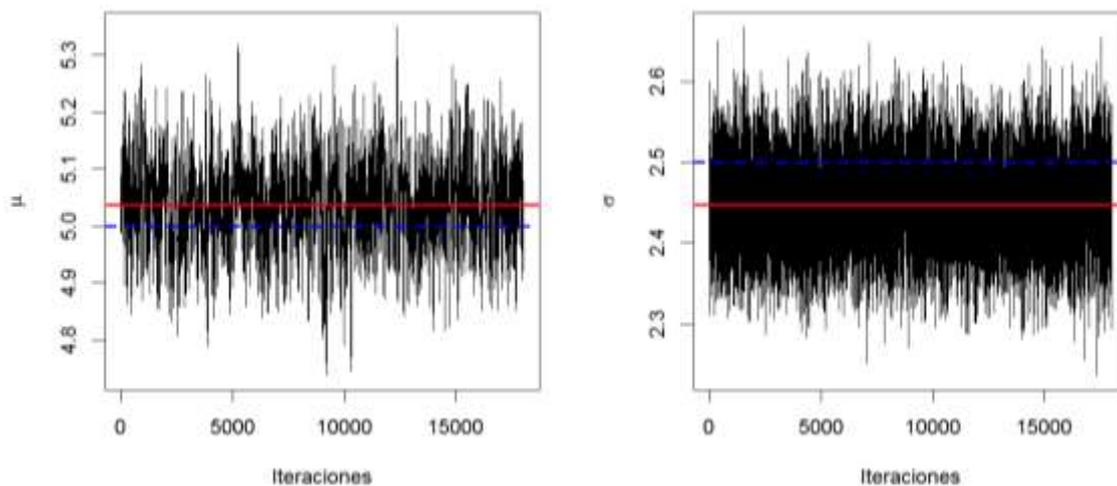


Figura 14. Histograma y diagrama de caja y bigotes de la desviación estándar σ (MALA)



En la Figura 15 se observan los diagramas de traza de MALA (líneas continuas negras) de la cadenas de Markov después del periodo de quemado, junto con el valor promedio de la cadena de Markov (líneas discontinuas azules) y el valor verdadero (líneas continuas rojas).

Figura 15. Diagrama de traza de los parámetros μ y σ (MALA).



En la Tabla 2, se muestra el resumen estadístico de la cadena de Markov resultante del algoritmo MALA junto con el valor verdadero.

Tabla 2. Resumen estadístico de las cadenas de Markov (algoritmo MALA) correspondiente a los parámetros μ y σ .

(Valor verdadero)	$\mu = 5$	$\sigma = 2.5$
Media	5.04	2.45
Desviación estándar	0.08	0.05

2.12.3 Algoritmo de Monte Carlo Hamiltoniano

El algoritmo de Monte Carlo Hamiltoniano, Monte Carlo híbrido o Hamiltoniano, es un método MCMC de muestreo basado en gradiente, caracterizado por realizar múltiples pasos en el espacio de parámetros antes de aceptar o rechazar una cadena de propuesta. Para cada parámetro θ (llamado posición) en el espacio de parámetros, el algoritmo Hamiltoniano adiciona una variable de momento p . Este algoritmo se basa en considerar un sistema de partículas con el siguiente Hamiltoniano: (Luengo et al., 2020).

$$H(\theta, p) = -\log(\theta|y_{1:T}) + \frac{1}{2}p^T p$$

Las ecuaciones Hamiltonianas para la dinámica de las partículas en un tiempo ficticio τ son dadas por

$$\begin{aligned} \frac{d\theta}{d\tau} &= \nabla_p H(\theta, p) = p \\ \frac{dp}{d\tau} &= -\nabla_\theta H(\theta, p) = \nabla \log(\theta|y_{1:T}) \end{aligned}$$

El algoritmo Hamiltoniano construye la distribución de propuesta mediante simulación de trayectorias de las ecuaciones Hamiltonianas. Dado que una simulación exacta no es posible, se deben utilizar métodos numéricos para simular las trayectorias, por

ejemplo, un integrador simplético como el método de salto de rana (Leapfrog method). Un paso del método de salto de rana para las ecuaciones Hamiltonianas desde τ con tamaño de paso $\Delta\tau$ está dado por

$$\begin{aligned}\tilde{p}^{(\tau+\Delta\tau/2)} &= \tilde{p}^{(\tau)} + \frac{\Delta\tau}{2} \nabla \log(\tilde{\theta}^{(\tau)} | y_{1:T}) \\ \tilde{\theta}^{(\tau+\Delta\tau)} &= \tilde{\theta}^{(\tau)} + \Delta\tau \tilde{p}^{(\tau+\Delta\tau/2)} \\ \tilde{p}^{(\tau+\Delta\tau)} &= \tilde{p}^{(\tau+\Delta\tau/2)} + \frac{\Delta\tau}{2} \nabla \log(\tilde{\theta}^{(\tau+\Delta\tau)} | y_{1:T})\end{aligned}$$

En el Algoritmo 10, se muestra del pseudocódigo del algoritmo de Monte Carlo Hamiltoniano.

Algoritmo 10. Algoritmo de Monte Carlo Hamiltoniano

Entrada: Gradiente del logaritmo de la densidad a posteriori $\nabla_{\theta} \log p(\theta | y_{1:T})$, iteración inicial θ^0 (generada de una densidad inicial arbitraria o inicializada con un valor arbitrario), paso de discretización $\Delta\tau$, número de pasos de integración L , número de iteraciones I , periodo de quemado I_b , observaciones $y_{1:T} = \{y_1, y_2, \dots, y_T\}$.

Salida: Cadena de Markov después del periodo de quemado, esto es, $\{\theta^{I_b+1}, \dots, \theta^I\}$

1. Inicializar $\theta^1 = \theta^0$
2. Calcular una estimación de $p(\theta^1 | y_{1:T}) \propto p(y_{1:T} | \theta^1) p(\theta^1)$
3. Para $i = 2$ hasta I hacer
4. Resolver numéricamente las ecuaciones Hamiltonianas usando L pasos de un método de salto de rana (Leapfrog method) empezando con $\tilde{\theta}^{(0)} = \theta^{(i-1)}$ y $\tilde{p} \sim N(0, I)$. Establecer $\theta' = \tilde{\theta}^{(L\Delta\tau)}$ y $p' = \tilde{p}^{(L\Delta\tau)}$.
5. Calcular la probabilidad de aceptación:

$$\alpha_i = \alpha(\theta', \theta^{i-1}) = \min \left[1, \exp \left(-H(\theta', p') + H(\theta^{(i-1)}, p^{(i-1)}) \right) \right]$$

6. Generar una variable aleatoria uniforme $u \sim \mathcal{U}[0,1]$ y establecer:

$$\theta^i = \begin{cases} \theta', & \text{si } u \leq \alpha_i \\ \theta^{i-1}, & \text{en otro caso} \end{cases}$$

7. Descartar las primeras I_b iteraciones de quemado.

Ejemplo. Estimación de parámetros usando el algoritmo de Monte Carlo Hamiltoniano.

Utilizando los datos del ejercicio anterior, se implementó el algoritmo Monte Carlo Hamiltoniano con $I = 20\,000$ iteraciones, número de pasos de integración $L = 10$ y tamaño de paso $\Delta\tau = 0.015$, este último parámetro del algoritmo fue obtenido a través de experimentación. El gradiente del logaritmo de la densidad a posteriori $\nabla_{\theta} \log p(\theta|y_{1:T})$ donde $\theta = (\mu, \kappa)$ está dado por

$$\nabla_{\theta} \log p(\theta|y_{1:T}) \propto \nabla_{\theta} (\log p(y_{1:T}|\theta) + \log p(\theta)) = \begin{pmatrix} \frac{1}{\exp(2 \cdot \kappa)} \sum_{t=1}^T (y_t - \mu) \\ 1 - \kappa + \frac{1}{\exp(2 \cdot \kappa)} \sum_{t=1}^T (y_t - \mu)^2 \end{pmatrix}$$

Después de ejecutar el algoritmo de Monte Carlo Hamiltoniano con $I = 20000$ iteraciones, la tasa de aceptación del método fue 50.59% aproximadamente. En la Figura 16 y

Figura 17, se muestran los histogramas con las densidades empíricas de las cadenas de Markov de los parámetros μ y σ , las líneas continuas rojas representan los valores promedios de la cadena de Markov: $\mu_{\text{promedio}} = 5.03$ y $\sigma_{\text{promedio}} = 2.45$, y las líneas discontinuas azules indican el valor verdadero de los parámetros: $\mu_{\text{verdadero}} = 5$ y $\sigma_{\text{verdadero}} = 2.5$.

Figura 16. Histograma y diagrama de caja y bigotes de la media μ (Hamiltoniano)

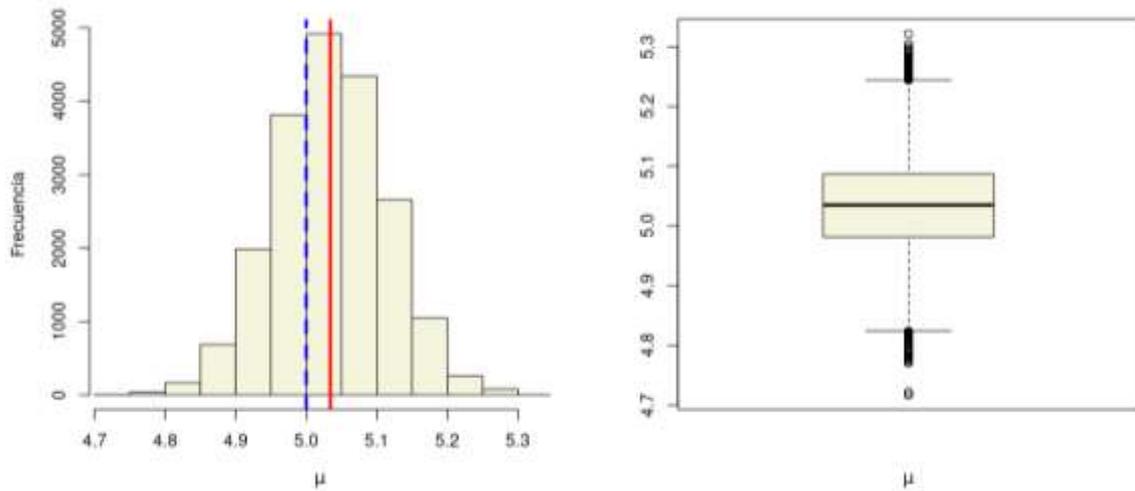
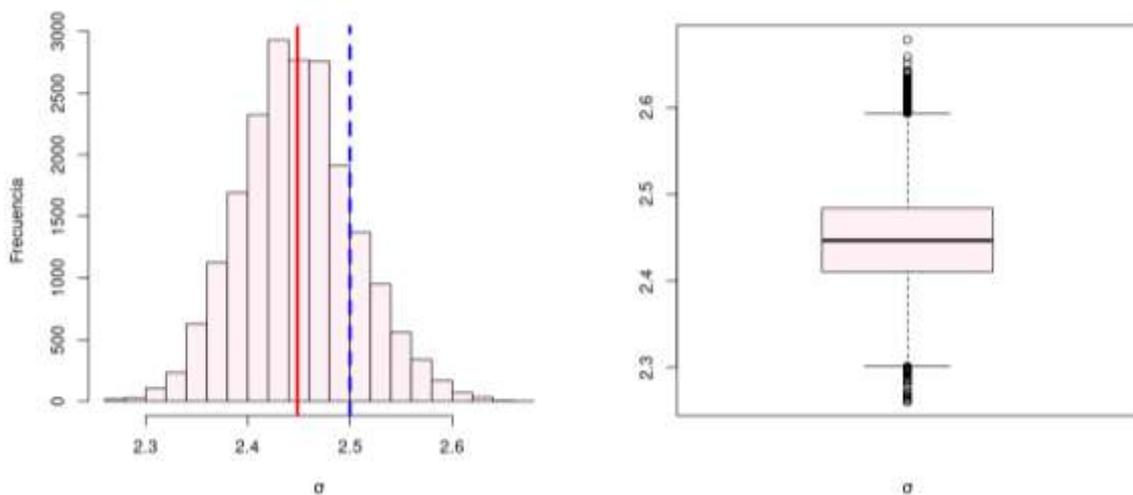
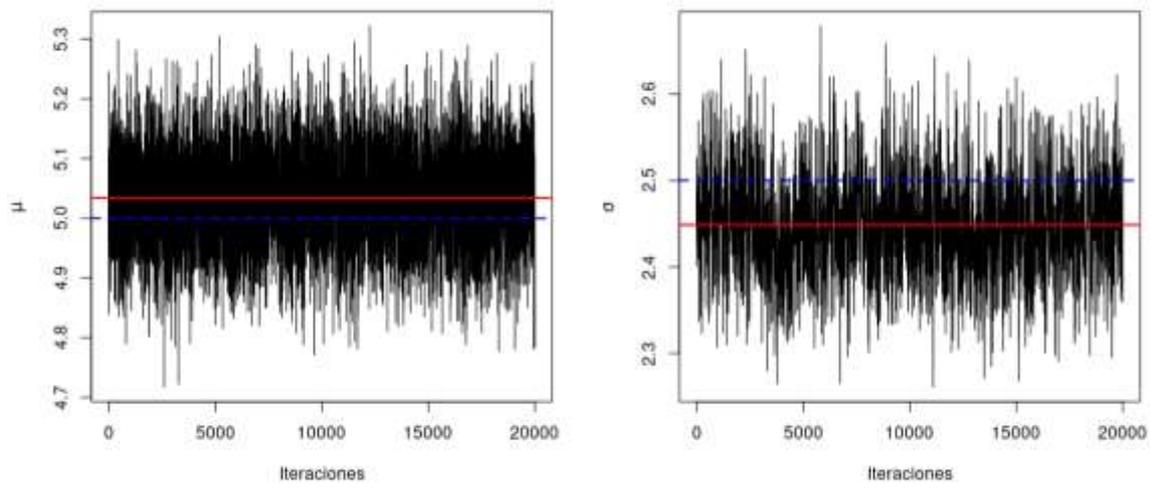


Figura 17. Histograma y diagrama de caja y bigotes de la desviación estándar σ (Hamiltoniano)



En la Figura 18 se observan los diagramas de traza del Algoritmo Monte Carlo Hamiltoniano (líneas continuas negras) de la cadenas de Markov después del periodo de quemado, junto con el valor promedio de la cadena de Markov (líneas discontinuas azules) y el valor verdadero (líneas continuas rojas).

Figura 18. Diagrama de traza de los parámetros μ y σ (Hamiltoniano).



En la Tabla 3, se muestra el resumen estadístico de la cadena de Markov resultante del algoritmo Monte Carlo Hamiltoniano junto con el valor verdadero.

Tabla 3. Resumen estadístico de las cadenas de Markov (Hamiltoniano) correspondiente a los parámetros μ y σ .

(Valor verdadero)	$\mu = 5$	$\sigma = 2.5$
Media	5.03	2.45
Desviación estándar	0.08	0.06

2.13 Algoritmo Metropolis – Hastings de partículas

El algoritmo Metropolis – Hastings de partículas es un método de Monte Carlo que combina el algoritmo Metropolis – Hastings con el filtro de partículas. Este método reemplaza la verosimilitud intratable del algoritmo Metropolis – Hastings por la estimación insesgada del filtro de partículas (Botha et al., 2021). En el Algoritmo 11 se resume la estructura del algoritmo Metropolis – Hastings de partículas.

Algoritmo 11. Algoritmo Metropolis - Hastings de partículas

Entrada: Densidad de propuesta $q(\theta'|\theta)$, iteración inicial θ^0 (generada de una densidad inicial arbitraria o inicializada con un valor arbitrario), número de iteraciones I y periodo de quemado I_b , observaciones $y_{1:T} = \{y_1, y_2, \dots, y_T\}$, número de partículas N .

Salida: Cadena de Markov después del periodo de quemado, esto es, $\{\theta^{I_b+1}, \dots, \theta^I\}$

1. Inicializar $\theta^1 = \theta^0$
2. Ejecutar el algoritmo de filtro de partículas para obtener una estimación insesgada de $p^N(y_{1:T}|\theta^1)$
3. Para $i = 2$ hasta I hacer
4. Generar una muestra de valores candidatos θ' de la densidad de propuesta $q(\theta'|\theta^{i-1})$, esto es, $\theta' \sim q(\theta'|\theta^{i-1})$
5. Ejecutar el algoritmo de filtro de partículas para obtener una estimación insesgada de $p^N(y_{1:T}|\theta')$
6. Calcular la probabilidad de aceptación:

$$\alpha_i = \alpha(\theta', \theta^{i-1}) = \min \left[1, \frac{p^N(y_{1:T}|\theta')p(\theta')}{p^N(y_{1:T}|\theta^{i-1})p(\theta^{i-1})} \frac{q(\theta^{i-1}|\theta')}{q(\theta'|\theta^{i-1})} \right]$$

7. Generar una variable aleatoria uniforme $u \sim \mathcal{U}[0,1]$ y establecer:

$$\theta^i = \begin{cases} \theta', & \text{si } u \leq \alpha_i \\ \theta^{i-1}, & \text{en otro caso} \end{cases}$$

8. Descartar las primeras I_b iteraciones de quemado.

Ejemplo. Implementación del algoritmo Metropolis – Hastings de partículas

Consideramos el problema de estimar los parámetros s_x y s_y del modelo estado espacio (caminata aleatoria Gaussiana unidimensional) a partir de un conjunto de observaciones $y_{1:T}$.

$$\begin{aligned}x_t &= x_{t-1} + q_{t-1} \quad , \quad q_{t-1} \sim N(q_{t-1}; 0, s_x^2) \\y_t &= x_t + r_t \quad , \quad r_t \sim N(r_t; 0, s_y^2)\end{aligned}$$

Simulando datos con los parámetros verdaderos $s_x = 2$ y $s_y = 3.5$ y aplicando el algoritmo Metropolis – Hastings de partículas con 10 000 iteraciones (1 000 iteraciones del periodo de quemado) y $N = 200$ partículas se estimaron los parámetros verdaderos usando los valores iniciales $s'_x = 1.25$ y $s'_y = 2$. En la Figura 19 y Figura 20 se muestran los histogramas con las densidades empíricas de las cadenas de Markov de los parámetros s_x y s_y , las líneas discontinuas amarillas indican los valores iniciales del algoritmo, líneas continuas rojas representan los valores promedios de las cadenas de Markov: $s_{x\text{promedio}} = 2.62$ y $s_{y\text{promedio}} = 3.22$, y las líneas discontinuas azules indican los valores verdaderos de los parámetros: $s_{x\text{verdadero}} = 2$ y $s_{y\text{verdadero}} = 3.5$.

Figura 19. Histograma y diagrama de caja y bigotes del parámetro s_x

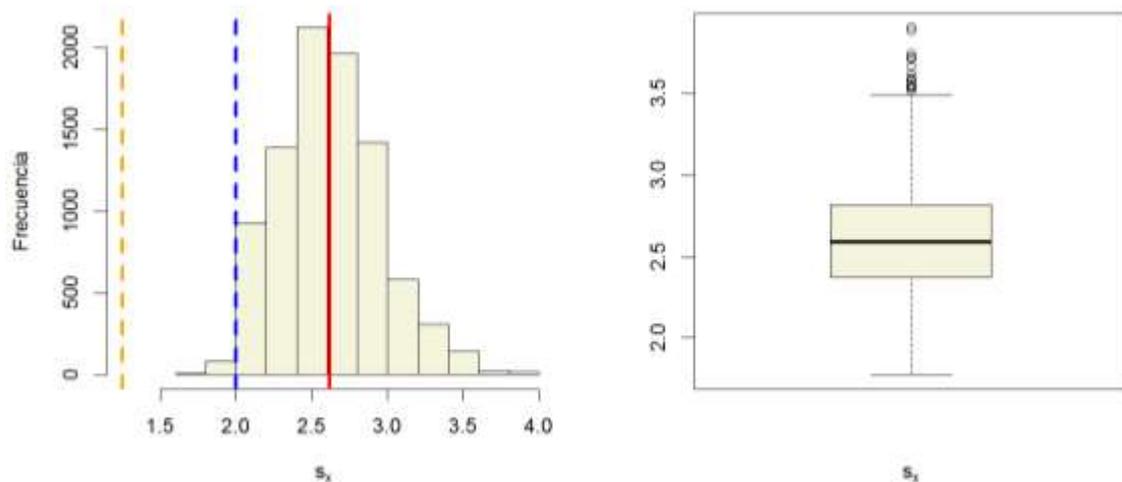
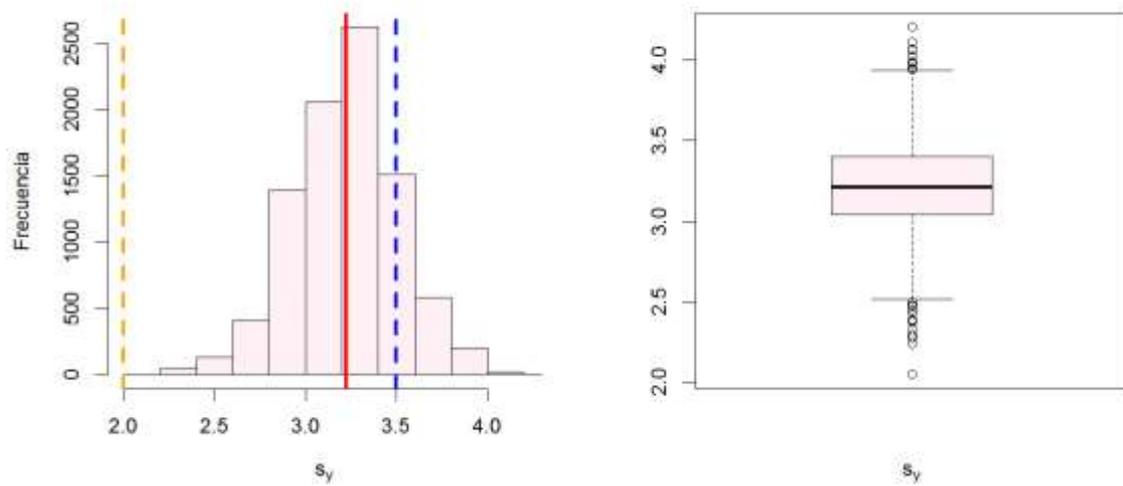
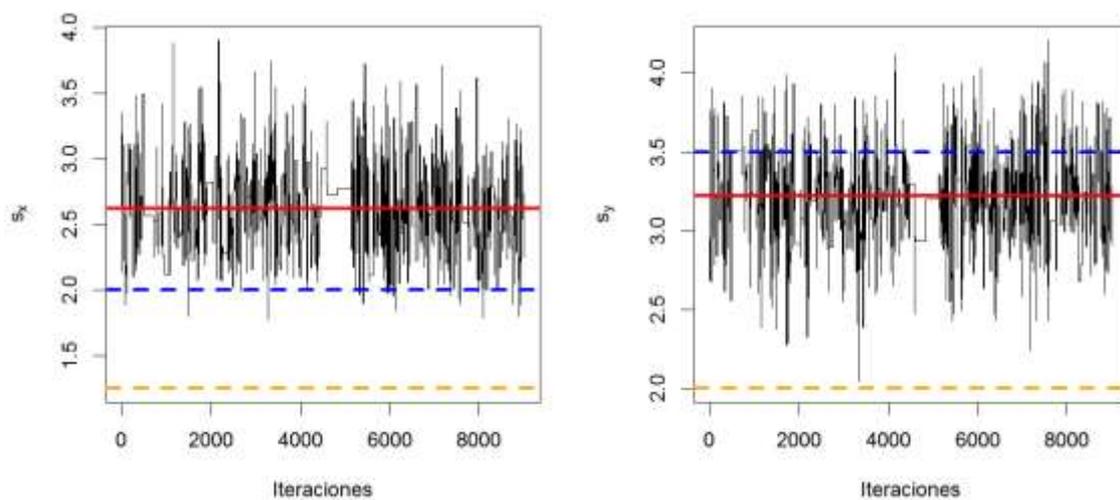


Figura 20. Histograma y diagrama de caja y bigotes del parámetro s_y



En la Figura 21 se observan los diagramas de traza del algoritmo Metropolis – Hastings de partículas (líneas continuas negras) de las cadenas de Markov después del periodo de quemado, junto con los valores iniciales (líneas discontinuas amarillas), los valores promedios de las cadenas de Markov (líneas discontinuas azules) y los valores verdaderos (líneas continuas rojas).

Figura 21. Diagrama de traza de los parámetros s_x y s_y



En la Tabla 4, se muestra el resumen estadístico de las cadenas de Markov resultante del algoritmo Metropolis – Hastings de partículas junto con los valores verdaderos.

Tabla 4. Resumen estadístico de las cadenas de Markov (algoritmo Metropolis – Hastings de partículas) correspondiente a los parámetros s_x y s_y .

(Valor verdadero)	$s_x = 2$	$s_y = 3.5$
Media	2.62	3.22
Desviación estándar	0.34	0.28

3 METODOLOGÍA

3.1 Algoritmos numéricos de simulación de EDEs

Consideramos una EDE con efectos mixtos para un individuo

$$dX_t = \mu(X_t, \phi_X, b)dt + \sqrt{v}(X_t, \phi_X, b)dW_t$$

Para aproximar esta EDE seguiremos un enfoque de discretización de Euler – Maruyama (EM) y de puente de difusión. Ambos enfoques simulan la EDE entre puntos de tiempos discretos, generalmente correspondientes a los tiempos de observación, a lo largo de la trayectoria completa de difusión, esto es, de $t = 0, \dots, T$.

Primero, consideramos la partición del intervalo $[0, T]$ como sigue

$$0 = t_0 < t_1 < t_2 < \dots < t_{m-1} < t_m = T$$

donde m es el número de tiempos de observación. Luego, consideramos el intervalo entre dos observaciones consecutivas $[t_{k-1}, t_k]$ con $k = 1, \dots, m$, y lo dividimos en D subintervalos, donde D denota el nivel de discretización, esto es,

$$t_{k-1} = \tau_0 < \tau_1 < \dots < \tau_D = t_k$$

con longitud entre cada subintervalo $\Delta\tau = (t_k - t_{k-1})/D$.

El esquema de EM y puente de difusión simula sobre cada subintervalos como sigue

$$X_{\tau_{j+1}} = X_{\tau_j} + \mu_{ED}(\cdot)\Delta\tau + \sqrt{v_{ED}(\cdot)}\Delta W_{\tau_j}$$

donde $\mu_{ED}(\cdot)$ y $\sqrt{v_{ED}(\cdot)}$ son determinados por el esquema de discretización (ED) utilizado, y donde $\Delta W_{\tau_j} = W_{\tau_{j+1}} - W_{\tau_j}$. Dado que $\Delta W_{\tau_j} \sim N(\Delta W_{\tau_j}; 0, \Delta\tau I_{n_x})$ por definición, la trayectoria es simulada recursivamente usando la siguiente densidad de transición

$$x_{\tau_{j+1}} | x_{\tau_j} \sim N \left(x_{\tau_{j+1}}; x_{\tau_j} + \mu_{\text{ED}}(\cdot) \Delta\tau, v_{\text{ED}}(\cdot) \Delta\tau \right)$$

La densidad conjunta de este esquema de discretización es

$$q(x_{\tau_1:\tau_D} | x_{\tau_0}, \phi_X, b) \propto \prod_{j=0}^{D-1} N \left(x_{\tau_{j+1}}; x_{\tau_j} + \mu_{\text{ED}}(\cdot) \Delta\tau, v_{\text{ED}}(\cdot) \Delta\tau \right)$$

3.1.1 Esquema Euler – Maruyama

El esquema de discretización Euler – Maruyama es el método más simple para simular una trayectoria aproximada de una EDE. Asumiendo que los coeficientes de drift y difusión son localmente constantes

$$\begin{aligned} \mu(X_{\tau_j}, \phi_X, b) &= \mu_j \\ \sqrt{v}(X_{\tau_j}, \phi_X, b) &= \sqrt{v_j} \end{aligned}$$

La discretización EM usa la propuesta

$$x_{\tau_{j+1}} | x_{\tau_j} \sim N \left(x_{\tau_{j+1}}; x_{\tau_j} + \mu_j, v_j \Delta\tau \right)$$

Esta propuesta aproxima la densidad de transición $f(x_{\tau_{j+1}} | x_{\tau_j}, \phi_X, b)$. Si la EDE tiene coeficientes de drift y de difusión constantes, entonces la discretización EM proporciona la solución exacta, esto es, la densidad de transición.

3.1.2 Puente de difusión modificado

Simular desde la densidad de transición aproximada puede ser subóptimo si algunas observaciones son muy informativas o hay poco ruido de observación. Trayectorias más efectivas pueden ser obtenidas si la propuesta para x_t puede ser dirigida hacia y_t . Esto es posible usando un puente de difusión.

El puente de difusión modificado (PDM) dirige x_t linealmente hacia y_t . El puente de difusión modificado es derivado a partir de aproximar la distribución conjunta de $X_{\tau_{j+1}}, Y_{t_k} | x_{\tau_j}$ usando una distribución Gaussiana multivariante, y luego condicionar en $Y_{t_k} = y_{t_k}$. La distribución de $X_{\tau_{j+1}}, Y_{t_k} | x_{\tau_j}$ se obtiene de la densidad de observación y la discretización EM de $X_{\tau_{j+1}} | x_{\tau_j}$. El puente de difusión es una propuesta de discretización de la forma

$$x_{\tau_{j+1}} | x_{\tau_j}, y_{t_k} \sim N \left(x_{\tau_{j+1}}; x_{\tau_j} + \mu_{\text{PDM}}(x_{\tau_j}, y_{t_k}), \psi_{\text{PDM}}(x_{\tau_j}) \Delta \tau \right)$$

donde

$$\begin{aligned} \mu_{\text{PDM}}(x_{\tau_j}, y_{t_k}) &= \mu_j + v_j F' (F v_j F' \Delta_j + \Sigma)^{-1} \left(y_{t_k} - F(x_{\tau_j} + \mu_j \Delta_j) \right) \\ \psi_{\text{PDM}}(x_{\tau_j}) &= v_j - v_j F' (F v_j F' \Delta_j + \Sigma)^{-1} F v_j \Delta \tau \end{aligned}$$

y $\Delta_j = t_k - \tau_j$.

3.1.3 Puente de difusión residual

Supongamos que X_t es dividido como $X_t = \zeta_t + R_t$ donde $\{\zeta_t\}_{t \geq 0}$ es un proceso determinístico y $\{R_t\}_{t \geq 0}$ es un proceso estocástico residual, además los procesos deben satisfacer las siguientes ecuaciones

$$\begin{aligned} d\zeta_t &= f(\zeta_t) dt, \quad \zeta_0 = x_0 \\ dR_t &= \{\mu_t(X_t, \phi_X, b) - f(\zeta_t)\} dt + \sqrt{v_t(X_t, \phi_X, b)} dW_t, \quad R_0 = 0 \end{aligned}$$

La elección de ζ_t y $f(\cdot)$ debe tomar en cuenta las dinámicas no lineales de la EDE (esta elección debe proponer que el coeficiente de drift sea aproximadamente constante). El enfoque más simple es tomar $\zeta_t = \eta_t$ y $f(\cdot) = \mu(\cdot)$ por lo tanto $X_t = \eta_t + R_t$ y

$$\begin{aligned} d\eta_t &= \mu(\eta_t, \phi_X, b) dt, \quad \eta_0 = x_0 \\ dR_t &= \{\mu_t(X_t, \phi_X, b) - \mu(\eta_t, \phi_X, b)\} dt + \sqrt{v_t(X_t, \phi_X, b)} dW_t, \quad R_0 = 0 \end{aligned}$$

notando que $Y_{t_k} - \eta_{t_k} = R_{t_k} + \epsilon_{t_k}$

El puente de difusión residual (PDR) se obtiene construyendo el PDM para el proceso estocástico residual en lugar del proceso objetivo. Este puente de difusión residual es

$$x_{\tau_{j+1}} | x_{\tau_j}, y_{t_k} \sim N \left(x_{\tau_{j+1}}; x_{\tau_j} + \mu_{\text{PDR}}(x_{\tau_j}, y_{t_k}), \psi_{\text{PDR}}(x_{\tau_j}) \Delta \tau \right)$$

donde

$$\begin{aligned} \mu_{\text{PDR}}(x_{\tau_j}, y_{t_k}) &= \mu_j + v_j F' (F v_j F' \Delta_j + \Sigma)^{-1} \left(y_{t_k} - F \left(\eta_{t_k} + r_{\tau_j} + (\mu_j - \delta_j^\eta) \Delta_j \right) \right) \\ \psi_{\text{PDR}}(x_{\tau_j}) &= \psi_{\text{PDM}}(x_{\tau_j}) \end{aligned}$$

y

$$\delta_j^\eta = \frac{\eta_{\tau_{j+1}} - \eta_{\tau_j}}{\Delta \tau}$$

3.2 Métodos de Monte Carlo por cadena de Markov para la estimación de parámetros

3.2.1 Métodos MCMC pseudo – marginales

Los métodos MCMC pseudo – marginales son algoritmos Metropolis – Hastings estándares que permiten la inferencia exacta en modelos con verosimilitudes intratables. En este enfoque, la verosimilitud intratable es reemplazada con una estimación insegada no negativa $\hat{p}(y_{1:T} | \theta) = p(y_{1:T} | \theta, u)$ donde $u \sim g(u)$ son variables auxiliares usadas para construir la estimación de la verosimilitud. Un método MCMC pseudo – marginal puede ser definido como un algoritmo Metropolis – Hastings en un espacio aumentado, esto es, el espacio de θ aumentado con las variables auxiliares u . En el contexto de inferencia para EDEs, las variables auxiliares u pueden ser una colección de números pseudo – aleatorios

necesarios para simular los incrementos de las trayectorias de un movimiento Browniano o realizar el paso de remuestreo en los filtros de partículas. El algoritmo Metropolis – Hastings marginal de partículas es un método pseudo – marginal donde la verosimilitud es reemplazada por una estimación insesgada del filtro de partículas (Wiqvist et al., 2021).

Algoritmo 12. Algoritmo Metropolis - Hastings marginal de partículas

Entrada: Densidad de propuesta $q(\theta'|\theta)$, iteración inicial θ^0 (generada de una densidad inicial arbitraria o inicializada con un valor arbitrario), número de iteraciones I y periodo de quemado I_b , observaciones $y_{1:T} = \{y_1, y_2, \dots, y_T\}$, número de partículas N .

Salida: Cadena de Markov después del periodo de quemado, esto es, $\{\theta^{I_b+1}, \dots, \theta^I\}$

1. Inicializar $\theta^1 = \theta^0$
2. Generar $u \sim g(\cdot)$
3. Ejecutar el algoritmo de filtro de partículas para obtener una estimación insesgada de $p^N(y_{1:T}|\theta^1, u)$
4. Para $i = 2$ hasta I hacer
 5. Generar una muestra de valores candidatos θ' de la densidad de propuesta $q(\theta'|\theta^{i-1})$, esto es, $\theta' \sim q(\theta'|\theta^{i-1})$ y variables auxiliares $u' \sim g(\cdot)$
 6. Ejecutar el algoritmo de filtro de partículas para obtener una estimación insesgada de $p^N(y_{1:T}|\theta', u')$
 7. Calcular la probabilidad de aceptación:

$$\alpha_i = \alpha(\theta', \theta^{i-1}) = \min \left[1, \frac{p^N(y_{1:T}|\theta', u')p(\theta')}{p^N(y_{1:T}|\theta^{i-1}, u)p(\theta^{i-1})} \frac{q(\theta^{i-1}|\theta')}{q(\theta'|\theta^{i-1})} \right]$$

8. Generar una variable aleatoria uniforme $\tilde{u} \sim \mathcal{U}[0,1]$ y establecer:

$$(\theta^i, u) = \begin{cases} (\theta', u'), & \text{si } \tilde{u} \leq \alpha_i \\ (\theta^{i-1}, u), & \text{en otro caso} \end{cases}$$

9. Descartar las primeras I_b iteraciones de quemado.

3.2.2 *Enfoque pseudo – marginal correlacionado*

Los métodos MCMC pseudo – marginales correlacionados son una generalización de los métodos pseudo – marginales. Una alta correlación entre las variables auxiliares u y u' conlleva una mayor correlación entre las verosimilitudes $p(y_{1:T}|\theta', u')$ y $p(y_{1:T}|\theta, u)$. Esto último permite reducir la varianza del logaritmo del ratio de las verosimilitudes, $\log(p(y_{1:T}|\theta', u')/p(y_{1:T}|\theta, u))$, aumentando la mezcla de la cadena. Los métodos pseudo – marginales correlacionados se caracterizan por generar una nueva variable u' del kernel $K(u'|u)$, donde $K(\cdot | \cdot)$ satisface la ecuación de balance $g(u)K(u'|u) = g(u')K(u|u')$ (Wiqvist et al., 2021).

Asumiendo que los números aleatorios están normalmente distribuidos, esto es, $u \sim g(u) = N(u|0, I)$ y tomando $K(\cdot | \cdot)$ como el kernel asociado a la densidad de propuesta de Crank – Nicolson ($K(u'|u) = N(u'|\rho u, (1 - \rho^2)I)$), se puede inducir una correlación ρ entre las variables auxiliares u y u' . Si el parámetro ρ es elegido cercano a 1, entonces se induce una correlación positiva entre las verosimilitudes $p(y_{1:T}|\theta', u')$ y $p(y_{1:T}|\theta, u)$, reduciendo la varianza de la probabilidad de aceptación del algoritmo Metropolis – Hastings, lo cual es beneficioso porque reduce la posibilidad de aceptar una sobreestimación de la función de verosimilitud. Si se toma $\rho = 0$, entonces se tiene $K(u'|u) = g(u')$, lo cual corresponde al método pseudo – marginal estándar (Wiqvist et al., 2021).

3.3 **Métodos pseudo – marginales de partículas para la estimación de EDEs con efectos mixtos**

A continuación, consideramos algoritmos pseudo – marginales de partículas para estimar los parámetros desconocidos de una EDE con efectos mixtos. Los filtros de

partículas de estos algoritmos permiten calcular una estimación de la verosimilitud individual $p(y_{1:T_m}^{(m)} | b^{(m)}, \Sigma, \phi_X)$ para $m = 1, \dots, M$, la cual puede ser usada para estimar la verosimilitud condicional de todos los individuos.

$$\hat{p}(y_{1:T_m}^{(1:M)} | b^{(1:M)}, \Sigma, \phi_X) = \prod_{m=1}^M \hat{p}(y_{1:T_m}^{(m)} | b^{(m)}, \Sigma, \phi_X)$$

Si la solución de la EDE no se puede obtener de forma cerrada, la densidad de transición puede ser aproximada usando la discretización Euler – Maruyama, por lo tanto, la distribución objetivo es exacta hasta el error de discretización de la EDE. Este error de discretización puede ser disminuido si se aumenta el nivel de discretización, a expensas de aumentar el costo computacional (Botha et al., 2021).

3.3.1 Método pseudo – marginal de aumento individual

Este método se caracteriza por utilizar variables auxiliares adicionales para estimar la verosimilitud de cada individuo. A continuación, la estimación de la verosimilitud para el individuo m

$$\begin{aligned} \hat{p}(y_{1:T_m}^{(m)} | \theta) &= \int \hat{p}(y_{1:T_m}^{(m)} | b^{(m)}, \Sigma, \phi_X) p(b^{(m)} | \psi) db^{(m)} \\ &\approx \frac{1}{N} \sum_{l=1}^L \frac{\hat{p}(y_{1:T_m}^{(m)} | b_l^{(m)}, \Sigma, \phi_X) p(b_l^{(m)} | \psi)}{p(b_l^{(m)} | \theta)} \end{aligned}$$

donde $\theta = (\Sigma, \phi_X, \psi)$ y $b_l^{(m)} \sim p(b^{(m)} | \theta)$ es la distribución de importancia dentro del algoritmo Metropolis – Hastings de partículas.

La variabilidad de la verosimilitud $\hat{p}(y^{(m)} | \theta)$ para un distribución de importancia dada $p(b^{(m)} | \theta)$ puede ser controlada tanto por el número de partículas N como por las generaciones de efectos aleatorios L . Si se elige $p(b^{(m)} | \theta) = p(b^{(m)} | \psi)$, entonces se

simplifica el cálculo de la verosimilitud a un promedio de las verosimilitudes generadas por L efectos aleatorios, esto es, $\hat{p}(y^{(m)}|\theta) \approx \sum_{l=1}^L \hat{p}(y^{(m)}|b_l^{(m)}, \Sigma, \phi_X) / L$ (Botha et al., 2021).

Algoritmo 13. Método Pseudo - Marginal de aumento individual

Entrada: Densidad de propuesta $q(\theta'|\theta)$, iteración inicial θ^0 (generada de una densidad inicial arbitraria o inicializada con un valor arbitrario), número de iteraciones I y periodo de quemado I_b , observaciones de los M individuos $y_{1:T_m}^{(1:M)} = \{y_1^{(m)}, y_2^{(m)}, \dots, y_{T_m}^{(m)}\}_{m=1}^M$, número de partículas N , número de efectos aleatorios L .

Salida: Cadena de Markov después del periodo de quemado, esto es, $\{\theta^{I_b+1}, \dots, \theta^I\}$

1. Inicializar $\theta^1 = \theta^0$
2. Generar $u \sim g(\cdot)$
3. Ejecutar el algoritmo de estimación de verosimilitud individual (filtro de partículas) con L efectos aleatorios para obtener una estimación de la verosimilitud $p^N(y_{1:T_m}^{(1:M)}|\theta^1, u)$
4. Para $i = 2$ hasta I hacer
5. Generar una muestra de valores candidatos θ' de la densidad de propuesta $q(\theta'|\theta^{i-1})$, esto es, $\theta' \sim q(\theta'|\theta^{i-1})$, y variables auxiliares $u' \sim g(\cdot)$
6. Ejecutar el algoritmo de estimación de verosimilitud individual (filtro de partículas) con L efectos aleatorios para obtener una estimación de la verosimilitud $p^N(y_{1:T_m}^{(1:M)}|\theta', u')$
7. Calcular la probabilidad de aceptación:

$$\alpha_i = \alpha(\theta', \theta^{i-1}) = \min \left[1, \frac{p^N(y_{1:T_m}^{(1:M)}|\theta', u') p(\theta')}{p^N(y_{1:T_m}^{(1:M)}|\theta^{i-1}, u) p(\theta^{i-1})} \frac{q(\theta^{i-1}|\theta')}{q(\theta'|\theta^{i-1})} \right]$$

8. Generar una variable aleatoria uniforme $\tilde{u} \sim \mathcal{U}[0,1]$ y establecer:

$$(\theta^i, u) = \begin{cases} (\theta', u'), & \text{si } \tilde{u} \leq \alpha_i \\ (\theta^{i-1}, u), & \text{en otro caso} \end{cases}$$

9. Descartar las primeras I_b iteraciones de quemado.

Algoritmo 14. Algoritmo de estimación de verosimilitud individual

Entrada: Observaciones de los M individuos $y_{1:T_m}^{(1:M)} = \{y_1^{(m)}, y_2^{(m)}, \dots, y_{T_m}^{(m)}\}_{m=1}^M$, número de efectos aleatorios L y número de partículas N .

Salida: Estimación de la verosimilitud $p^N(y_{1:T_m}^{(1:M)} | \theta)$.

1. Para $m = 1$ hasta M hacer
 2. Para $l = 1$ hasta L hacer
 3. Extraer $b_l^{(m)} \sim p(\cdot | \theta)$
 4. Ejecutar el algoritmo de partículas (N partículas) con $b_l^{(m)}$ para obtener la estimación de la verosimilitud $Z_l^{(m)}$
 5. Corregir la distribución de importancia $Z_l^{(m)} = Z_l^{(m)} / p(b_l^{(m)} | \theta)$
 6. Calcular

$$p^N(y_{1:T_m}^{(m)} | \theta) = \frac{1}{L} \sum_{i=1}^L Z_i^{(m)}$$

7. Calcular

$$p^N(y_{1:T_m}^{(1:M)} | \theta) = \prod_{m=1}^M p^N(y_{1:T_m}^{(m)} | \theta)$$

3.3.2 Método pseudo – marginal por componentes

El método pseudo – marginal por componentes es un método de muestro de Gibbs que se caracteriza por actualizar los efectos aleatorios $b^{(1:M)}$ junto con θ , esto lleva a considerar los bloques de parámetros $b^{(1:M)}$, $\{\sigma, \phi_X\}$ y ψ . Si $\theta_X = \{\sigma, \phi_X\}$, entonces la densidad a posteriori (conjunta) completa tiene la forma

$$p(\theta_X, \psi, b^{(1:M)} | y_{1:T_m}^{(1:M)}) \propto p(y_{1:T_m}^{(1:M)} | b^{(1:M)}, \theta_X) p(b^{(1:M)} | \psi) p(\theta_X) p(\psi)$$

y las densidades a posteriori condicionales completas para cada uno de los bloques de parámetros son

$$\begin{aligned} p(b^{(m)} | y_{1:T_m}^{(1:M)}, \theta_X, \psi) &\propto p(y_{1:T_m}^{(1:M)} | b^{(1:M)}, \theta_X) p(b^{(1:M)} | \psi) \\ p(\theta_X | y_{1:T_m}^{(1:M)}, b^{(1:M)}) &\propto p(y_{1:T_m}^{(1:M)} | b^{(1:M)}, \theta_X) p(\theta_X) \\ p(\psi | b^{(1:M)}) &\propto p(b^{(1:M)} | \psi) p(\psi) \end{aligned}$$

Una estimación dada por un filtro de partículas de la verosimilitud $p(y_{1:T_m}^{(1:M)} | b^{(1:M)}, \theta_X)$ es usada cuando se actualizan los efectos aleatorios $b^{(1:M)}$ y los parámetros θ_X . El parámetro de los efectos poblacionales ψ es actualizado directamente dado que la densidad $p(\psi | b^{(1:M)})$ es tratable (Botha et al., 2021).

Algoritmo 15. Método pseudo – marginal por componentes

Entrada: Densidad de propuesta $q(\theta' | \theta)$, iteración inicial θ^0 (generada de una densidad inicial arbitraria o inicializada con un valor arbitrario), número de iteraciones I y periodo de quemado I_b , observaciones de los M individuos $y_{1:T_m}^{(1:M)} = \{y_1^{(m)}, y_2^{(m)}, \dots, y_{T_m}^{(m)}\}_{m=1}^M$, número de partículas N , valores iniciales $b^{(1)} = b^{(1,1:M)}$, $\theta_X^{(1)}$ y $\psi^{(1)}$.

Salida: Cadena de Markov después del periodo de quemado, esto es, $\{\theta^{I_b+1}, \dots, \theta^I\}$

1. Inicializar $\theta^1 = \theta^0$
2. Generar $u \sim g(\cdot)$
3. Ejecutar el filtro de partículas para obtener una estimación de la verosimilitud

$$p^N(y_{1:T_m}^{(1:M)} | b^{(1)}, \theta_X^{(1)}, u)$$

4. Para $i = 2$ hasta I hacer
5. Generar efectos aleatorios $b' \sim q(\cdot | b^{(i-1)})$ y variables auxiliares $u' \sim g(\cdot)$
6. Ejecutar el filtro de partículas para obtener una estimación de la verosimilitud

$$p^N(y_{1:T_m}^{(1:M)} | b', \theta_X^{(i-1)}, u')$$

7. Aceptar b' y u' con probabilidad de aceptación:

$$\alpha = \min \left[1, \frac{p^N(y_{1:T_m}^{(1:M)} | b', \theta_X^{(i-1)}, u') p(b' | \psi^{(i-1)}) q(b^{(i-1)} | b')}{p^N(y_{1:T_m}^{(1:M)} | b^{(i-1)}, \theta_X^{(i-1)}, u) p(b^{(i-1)} | \psi^{(i-1)}) q(b' | b^{(i-1)})} \right]$$

8. Generar $\theta'_X \sim q(\cdot | \theta_X^{(i-1)})$ y variables auxiliares $u' \sim g(\cdot)$
9. Ejecutar el filtro de partículas para obtener una estimación de la verosimilitud

$$p^N(y_{1:T_m}^{(1:M)} | b^{(i)}, \theta'_X, u')$$

10. Aceptar θ'_X y u' con probabilidad de aceptación:

$$\alpha = \min \left[1, \frac{p^N(y_{1:T_m}^{(1:M)} | b^{(i)}, \theta'_X, u') p(\theta'_X) q(\theta_X^{(i-1)} | \theta'_X)}{p^N(y_{1:T_m}^{(1:M)} | b^{(i)}, \theta_X^{(i-1)}, u) p(\theta_X^{(i-1)}) q(\theta'_X | \theta_X^{(i-1)})} \right]$$

11. Generar $\psi' \sim q(\cdot | \psi^{(i-1)})$
12. Aceptar ψ' con probabilidad de aceptación:

$$\alpha = \min \left[1, \frac{p(b^{(i)} | \psi') p(\psi') q(\psi^{(i-1)} | \psi')}{p^N(b^{(i)} | \psi^{(i-1)}) p(\psi^{(i-1)}) q(\psi' | \psi^{(i-1)})} \right]$$

13. Descartar las primeras I_b iteraciones de quemado.

4 APLICACIONES

En este capítulo, se presentan ejemplos específicos de aplicaciones en el campo de la biología (modelo de crecimiento del árbol de naranja) y procesos bivariados usados en finanzas, ecología, ingeniería (proceso de Ornstein-Uhlenbeck bidimensional).

4.1 Datos reales

4.1.1 Modelo de crecimiento del árbol de naranja

La dinámica del crecimiento la circunferencia del tronco de los árboles de naranja puede ser modelada mediante EDEs con efectos mixtos (Whitaker et al., 2017). Este enfoque combina los conceptos de EDEs y efectos mixtos para capturar tanto la dinámica determinística del crecimiento como la variabilidad aleatoria asociada. La aplicación de un modelo de EDE con efectos mixtos proporciona un enfoque para el estudio de crecimiento de otros organismos y cultivos, brindando información del desarrollo de una especie necesaria para la toma de decisiones en el campo de la agricultura. Sea $X_t^{(m)}$ un proceso estocástico que describe la circunferencia del árbol $m = 1, \dots, M$ que está definido por la siguiente EDE con condición inicial $X_0^{(m)}$ y efectos aleatorios $\phi_1^{(m)}$ y $\phi_2^{(m)}$.

$$dX_t^{(m)} = \frac{1}{\phi_1^{(m)} \phi_2^{(m)}} X_t^{(m)} (\phi_1^{(m)} - X_t^{(m)}) dt + \sigma \sqrt{X_t^{(m)}} dW_t^{(m)}$$

$$\begin{aligned} X_0^{(m)} &= x_0^{(m)} \\ \phi_1^{(m)} &\sim N(\phi_1^{(m)}; \mu_{\phi_1}, \sigma_{\phi_1}^2) \\ \phi_2^{(m)} &\sim N(\phi_2^{(m)}; \mu_{\phi_2}, \sigma_{\phi_2}^2) \end{aligned}$$

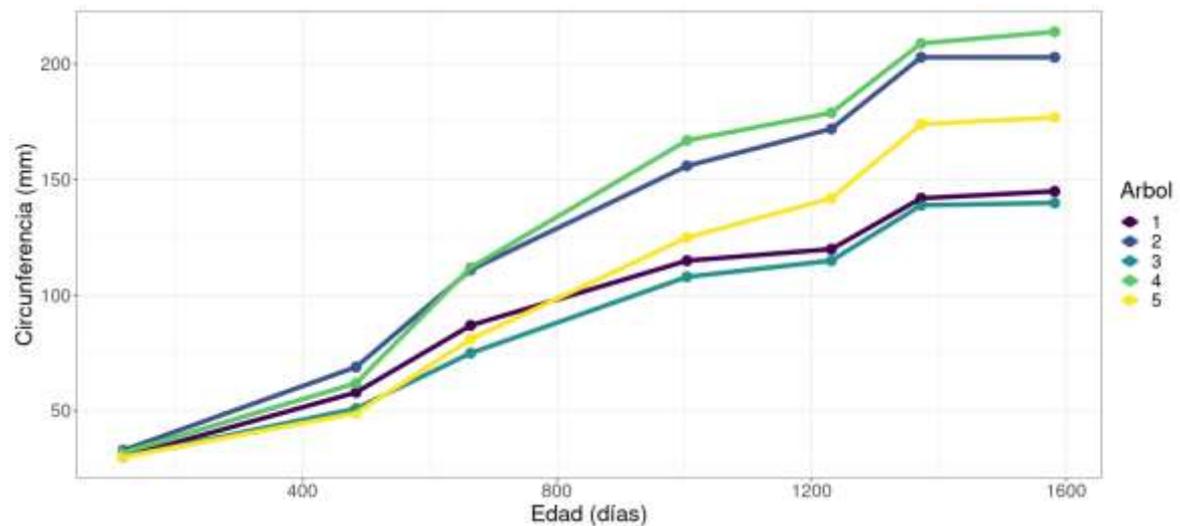
Se considera que el proceso es parcialmente observado, así que las observaciones son modeladas por

$$Y_t^{(m)} = X_t^{(m)} + \epsilon_t^{(m)}$$

donde $\epsilon_t^{(m)} \sim N(\epsilon_t^{(m)}; 0, \xi^2)$ es el ruido de observación con parámetro fijo ξ .

Aquí, $\theta = (\sigma, \xi)$ son comunes para todos los árboles (efecto fijo), los efectos aleatorios son $b^{(m)} = (\phi_1^{(m)}, \phi_2^{(m)})^T$ y el vector de los parámetros que gobiernan las distribuciones de efectos aleatorios es $\psi = (\mu_{\phi_1}, \sigma_{\phi_1}, \mu_{\phi_2}, \sigma_{\phi_2})^T$ donde μ_{ϕ_1} y μ_{ϕ_2} con densidades a priori $N(0, 100^2)$ y $1/\sigma_{\phi_1}^2$, $1/\sigma_{\phi_2}^2$, $1/\sigma^2$ y $1/\xi^2$ con densidades a priori $Ga(1, 0.01)$ (Whitaker et al., 2017).

Figura 22. Gráfico de las observaciones (datos reales) de los $M = 5$ árboles del modelo de crecimiento del árbol de naranja.



4.2 Datos simulados

4.2.1 Proceso de Ornstein – Uhlenbeck bidimensional

Un modelo de Ornstein – Uhlenbeck (OU) bidimensional con efectos mixtos es un modelo estadístico que combina un proceso OU clásico con un modelo de efectos mixtos que incorpora tanto efectos fijos como efectos aleatorios en un contexto de dos variables. El

proceso OU es una ecuación diferencial estocástica que es usada para modelar la dinámica de un sistema que es sujeto de fuerzas aleatorias y determinísticas. Algunas aplicaciones del proceso OU se encuentran en estudios dinámicos fármaco-dinámicos/fármaco-cinéticos (Fadwa et al., 2020) y en Neurociencias (Dion, 2016), dado que el proceso OU puede describir la dinámica estocástica de variables continuas a lo largo del tiempo, donde las variables tienden a converger hacia una media o punto de equilibrio. Sean $m = 1, \dots, M$ procesos estocásticos, consideramos un proceso de difusión bidimensional $X_t^{(m)} =$

$(X_t^{(1,m)}, X_t^{(2,m)})^T \in \mathbb{R}^2$ que satisface las siguientes EDEs

$$\begin{aligned} dX_t^{(1,m)} &= -\left(\beta_{11}b_{11}^{(m)}(X_t^{(1,m)} - \alpha_1) + \beta_{12}b_{12}^{(m)}(X_t^{(2,m)} - \alpha_2)\right)dt + \sigma_1dW_t^{(1,m)} \\ dX_t^{(2,m)} &= -\left(\beta_{21}b_{21}^{(m)}(X_t^{(1,m)} - \alpha_1) + \beta_{22}b_{22}^{(m)}(X_t^{(2,m)} - \alpha_2)\right)dt + \sigma_2dW_t^{(2,m)} \end{aligned}$$

con condiciones iniciales

$$\begin{aligned} X_0^{(1,m)} &= x_0^{(1,m)} \\ X_0^{(2,m)} &= x_0^{(2,m)} \end{aligned}$$

Además consideramos los efectos aleatorios como $b_{l,k} \sim Ga(v_{l,k}, v_{l,k})$ para $l, k = 1, 2$. Se tiene que $Ga(\alpha, \beta)$, es una distribución Gamma con parámetros α y β , y $\psi = \{v_{l,k}\}_{l,k=1,2}$ es el vector de hiperparámetros.

Si se considera $X_t^{(m)} = \begin{pmatrix} X_t^{(1)(m)} \\ X_t^{(2)(m)} \end{pmatrix}$, $\beta = \begin{pmatrix} \beta_{11} & \beta_{12} \\ \beta_{21} & \beta_{22} \end{pmatrix}$, $\alpha = \begin{pmatrix} \alpha_1 \\ \alpha_2 \end{pmatrix}$, $\sigma = \begin{pmatrix} \sigma_1 & 0 \\ 0 & \sigma_2 \end{pmatrix}$,

$$W_t^{(m)} = \begin{pmatrix} W_t^{(1)(m)} \\ W_t^{(2)(m)} \end{pmatrix}, X_0^{(m)} = \begin{pmatrix} X_0^{(1)(m)} \\ X_0^{(2)(m)} \end{pmatrix} \text{ y } b^{(m)} = \begin{pmatrix} b_{11}^{(m)} & \beta_{12}^{(m)} \\ \beta_{21}^{(m)} & \beta_{22}^{(m)} \end{pmatrix}$$

Sea (\star) la operación que denota la multiplicación elemento por elemento, se puede reescribir el sistema de ecuaciones en notación matricial como sigue

$$dX_t^{(m)} = \beta \star b^{(m)}(\alpha - X_t^{(m)})dt + \sigma dW_t^{(m)}$$

Consideraremos que el proceso es parcialmente observado, así que las observaciones son modeladas por

$$Y_t^{(m)} = X_t^{(m)} + \epsilon_t^{(m)}$$

Donde $Y_t^{(m)} = (Y_1^{(m)}, Y_2^{(m)})^T$ y $\epsilon_t^{(m)} \sim N(\epsilon_t^{(m)}; 0, \Sigma)$ es el ruido de observación con

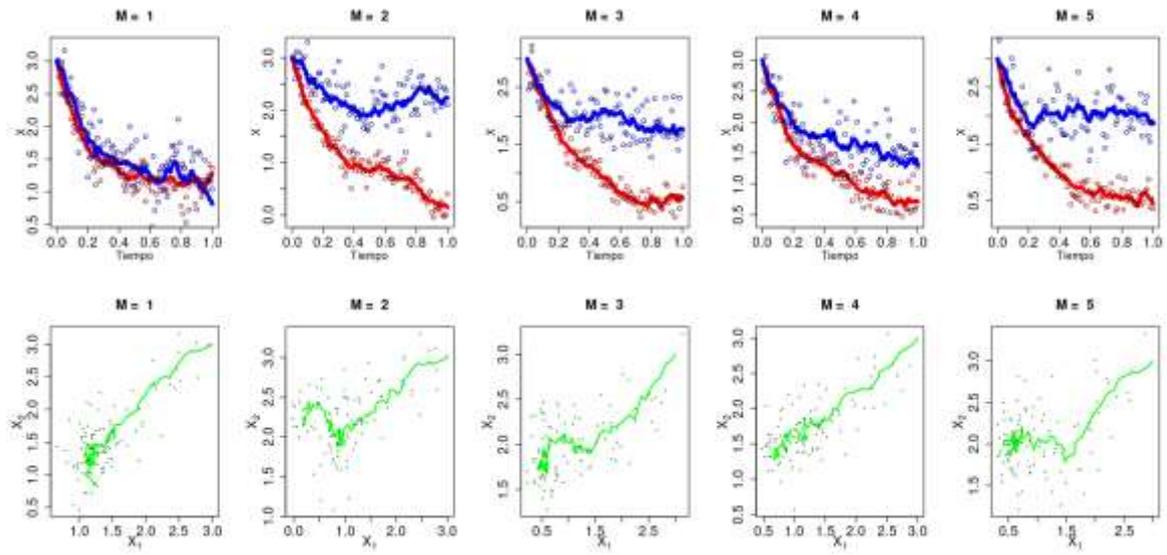
$$\Sigma = \begin{pmatrix} \xi_1^2 & 0 \\ 0 & \xi_2^2 \end{pmatrix}.$$

El vector de parámetros $\theta = (\alpha_1, \alpha_2, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \sigma_1, \sigma_2, \xi_1, \xi_2)$ es común para todas las unidades experimentales (efecto fijo), los efectos aleatorios son $b^{(m)} =$

$\begin{pmatrix} b_{11}^{(m)} & \beta_{12}^{(m)} \\ \beta_{21}^{(m)} & \beta_{22}^{(m)} \end{pmatrix}$ y el vector de los parámetros poblacionales o hiperparámetros que gobiernan

las distribuciones de efectos aleatorios es $\psi = (v_{11}, v_{12}, v_{21}, v_{22})$ (Fadwa et al., 2020).

En la Figura 23, en la primera fila se muestran las trayectorias de los estados X_1 (línea de color rojo) y X_2 (línea de color azul) con respecto al tiempo de $M = 5$ unidades experimentales, se observa que las trayectorias tienden hacia las medias asintóticas α_1 y α_2 . En la segunda fila se observan los gráficos del estado X_1 frente al estado X_2 (línea de color verde). En estos gráficos las observaciones se representan con círculos.

Figura 23. Simulación de un modelo bidimensional OU.

5 RESULTADOS

5.1 Modelo de crecimiento del árbol de naranja

5.1.1 Método pseudo – marginal de aumento individual

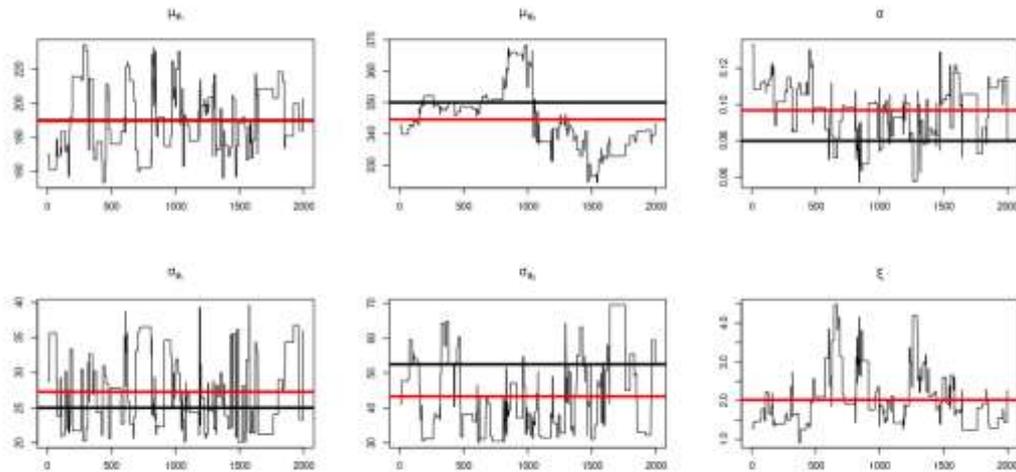
5.1.1.1 Esquema de discretización Euler – Maruyama (Método pseudo – marginal de aumento individual)

A continuación, se muestran los resultados a posteriori obtenidos del método pseudo – marginal de aumento individual, donde se utilizó un esquema de discretización Euler – Maruyama con un número de discretización $D = 50$ para simular las trayectorias de la EDE, mientras que la verosimilitud individual fue estimada usando $N = 20$ partículas, y la verosimilitud de los $M = 5$ árboles (unidades experimentales) fue estimada con $L = 20$ muestras de efectos aleatorios. El algoritmo fue ejecutado con $I = 2000$ iteraciones con $B = 200$ iteraciones de periodo de quemado (burn-in). Después de ejecutar el método pseudo – marginal de aumento individual se alcanzó una tasa de aceptación de 7.72% con un tiempo de ejecución de 5.59 minutos.

En la

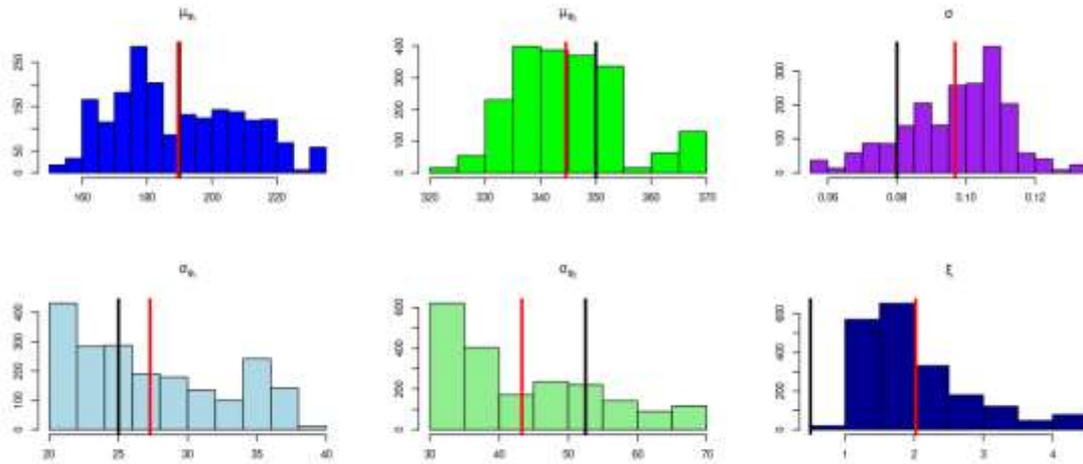
Figura 24, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal de aumento individual con esquema de discretización EM y la línea negra indica el valor inicial del parámetro. Se observa que el parámetro μ_{ϕ_2} no presenta una convergencia fuerte.

Figura 24. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización EM).



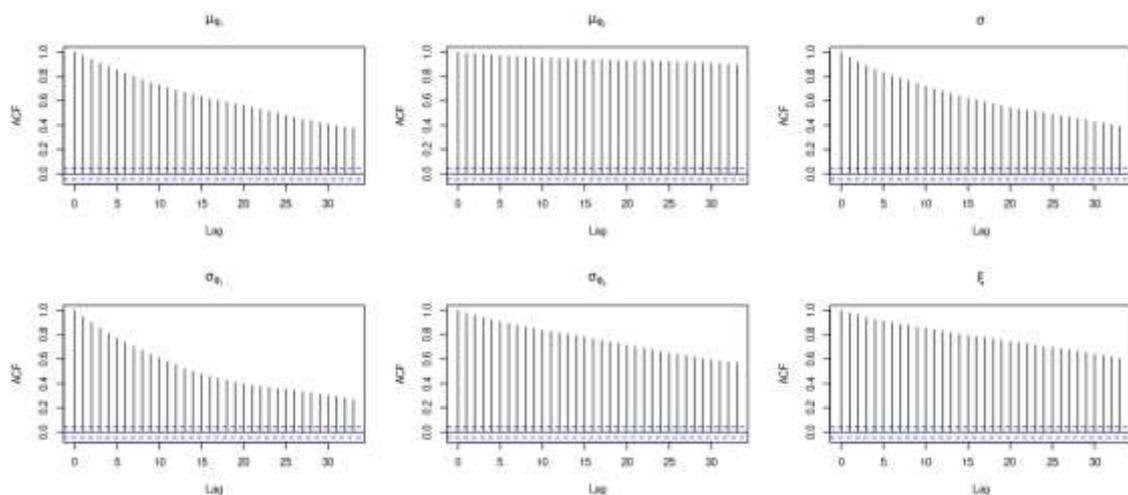
En la Figura 25, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal de aumento individual (EM) y la línea negra indica el valor inicial del parámetro. Se observa que los histogramas empíricos de los parámetros σ_{ϕ_1} , σ_{ϕ_2} y ξ están sesgados hacia la derecha.

Figura 25. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización EM).



En la Figura 26, se observa que las muestras a posteriori del parámetro μ_{ϕ_2} están altamente correlacionadas, mientras que las muestras a posteriori del parámetro ξ están medianamente correlacionadas.

Figura 26. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización EM).



En la Tabla 5, se muestran las estimaciones de los parámetros comunes θ obtenidas del Método pseudo – marginal de aumento individual con esquema de discretización EM.

Tabla 5. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización EM).

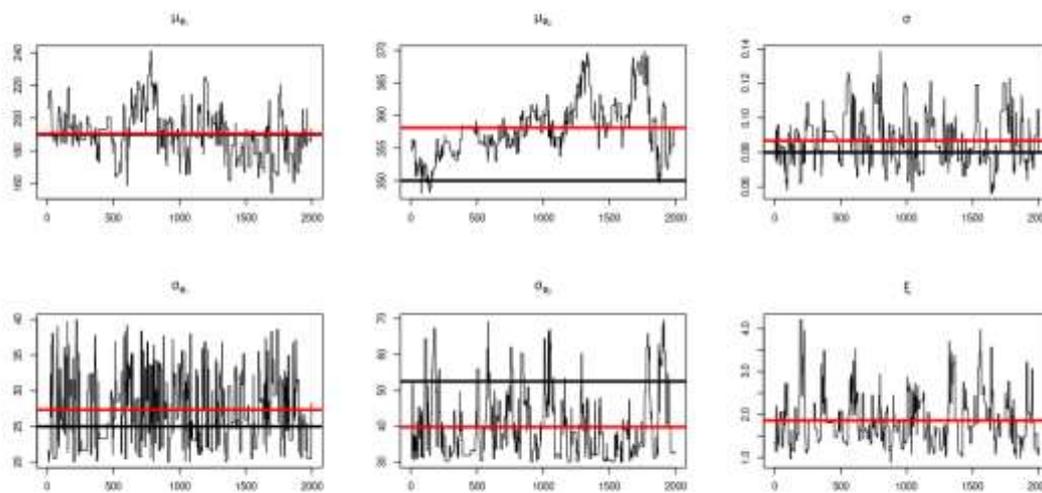
	μ_{ϕ_1}	σ_{ϕ_1}	μ_{ϕ_2}	σ_{ϕ_2}	σ	ξ
Media	189.53	27.28	344.55	43.31	0.1	2.02
Desviación estándar	19.71	5.39	9.58	11.56	0.01	0.74

5.1.1.2 Esquema de discretización Puente de difusión modificado (Método pseudo – marginal de aumento individual)

En esta sección se indican los resultados a posteriori obtenidos del método pseudo – marginal de aumento individual, donde se utilizó un esquema de discretización Puente de difusión modificado con un número de discretización $D = 50$ para simular las trayectorias de la EDE, mientras que la verosimilitud individual fue estimada usando $N = 20$ partículas, y la verosimilitud de los $M = 5$ árboles (unidades experimentales) fue estimada con $L = 20$ muestras de efectos aleatorios. El algoritmo fue ejecutado con $I = 2000$ iteraciones con $B = 200$ iteraciones de periodo de quemado (burn-in). Después de ejecutar el método pseudo – marginal de aumento individual se alcanzó una tasa de aceptación de 19.23% con un tiempo de ejecución de 9.62 minutos.

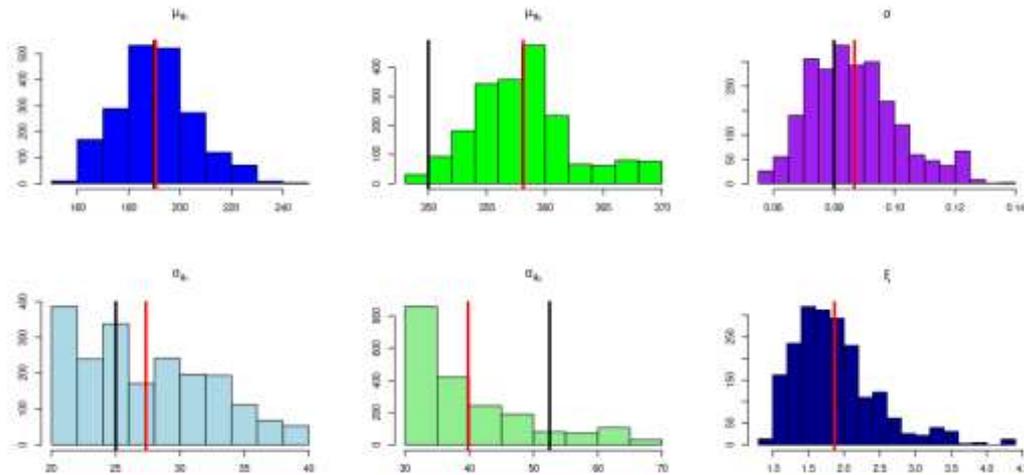
En la Figura 27, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión modificado y la línea negra indica el valor inicial del parámetro. Se observa que el parámetro μ_{ϕ_2} presenta una mejor convergencia hacia la media a posteriori.

Figura 27. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión modificado).



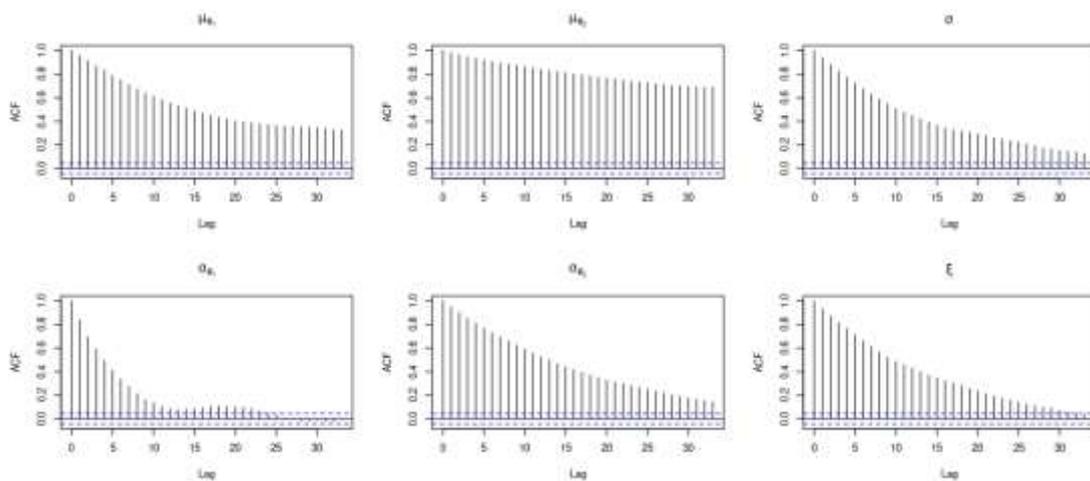
En la Figura 28, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión modificado y la línea negra indica el valor inicial del parámetro. Se observa que los histogramas empíricos de los parámetros σ_{ϕ_1} , σ_{ϕ_2} y ξ están sesgados hacia la derecha.

Figura 28. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión modificado).



En la Figura 29, se observa que las muestras a posteriori del parámetro $\mu\phi_2$ están medianamente correlacionadas.

Figura 29. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión modificado).



En la Tabla 6, se muestran las estimaciones de los parámetros comunes θ obtenidas del Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión modificado.

Tabla 6. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión modificado).

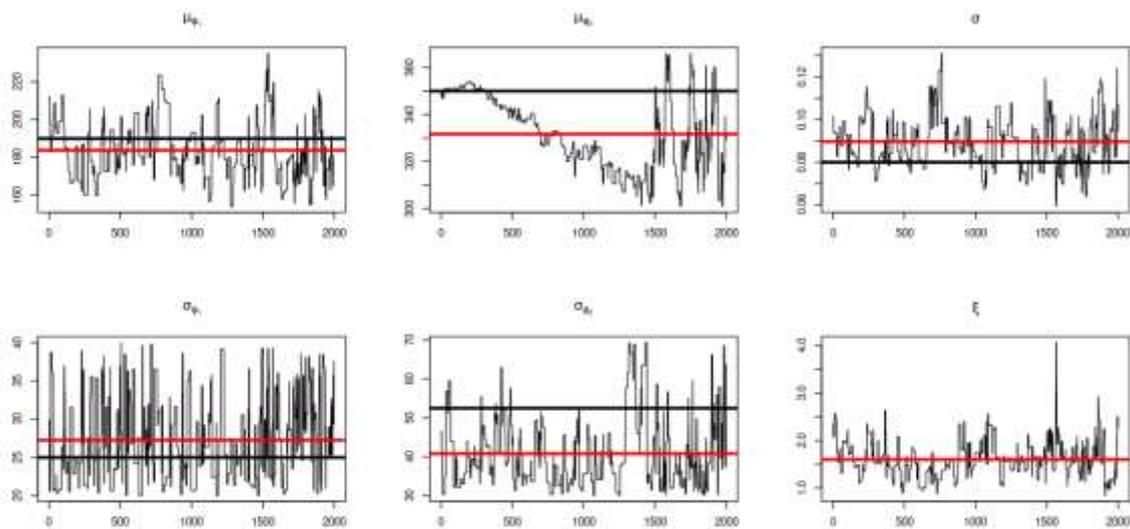
	μ_{ϕ_1}	σ_{ϕ_1}	μ_{ϕ_2}	σ_{ϕ_2}	σ	ξ
Media	190.56	27.34	358.12	39.8	0.09	1.86
Desviación estándar	14.83	5.05	4.30	9.46	0.01	0.58

5.1.1.3 Esquema de discretización Puente de difusión residual (Método pseudo – marginal de aumento individual)

En este apartado se muestran los resultados a posteriori obtenidos del método pseudo – marginal de aumento individual, donde se utilizó un esquema de discretización Puente de difusión residual con un número de discretización $D = 50$ para simular las trayectorias de la EDE, mientras que la verosimilitud individual fue estimada usando $N = 20$ partículas, y la verosimilitud de los $M = 5$ árboles (unidades experimentales) fue estimada con $L = 20$ muestras de efectos aleatorios. El algoritmo fue ejecutado con $I = 2000$ iteraciones con $B = 200$ iteraciones de periodo de quemado (burn-in). Después de ejecutar el método pseudo – marginal de aumento individual se alcanzó una tasa de aceptación de 15.64% con un tiempo de ejecución de 23.36 minutos.

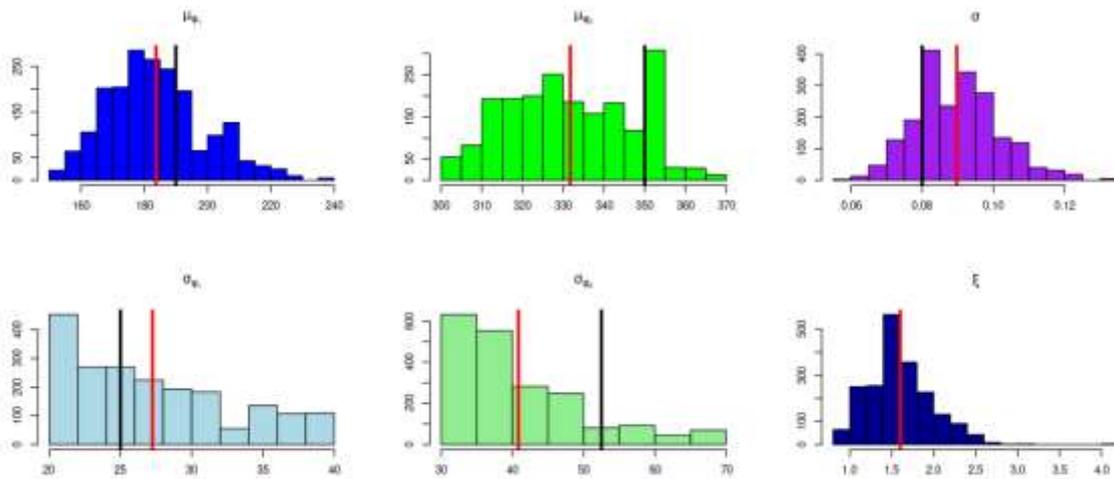
En la Figura 30, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal de aumento individual con esquema de Puente de difusión residual y la línea negra indica el valor inicial del parámetro. Se observa que el parámetro μ_{ϕ_2} empieza a converger en las últimas 100 iteraciones.

Figura 30. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión residual).



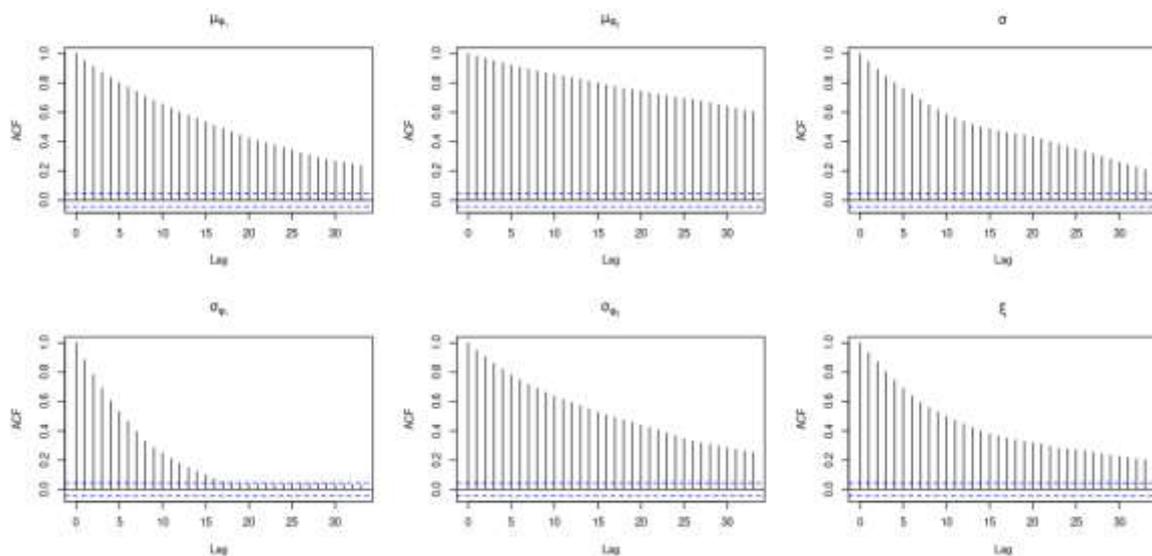
En la Figura 31, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión residual y la línea negra indica el valor inicial del parámetro. Se observa que los histogramas empíricos de los parámetros σ_{ϕ_1} y σ_{ϕ_2} están sesgados hacia la derecha.

Figura 31. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión residual).



En la Figura 32, se observa que las muestras a posteriori del parámetro μ_{ϕ_2} están medianamente correlacionadas.

Figura 32. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión residual).



En la Tabla 7, se muestran las estimaciones obtenidas del Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión residual.

Tabla 7. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal de aumento individual con esquema de discretización Puente de difusión residual).

	μ_{ϕ_1}	σ_{ϕ_1}	μ_{ϕ_2}	σ_{ϕ_2}	σ	ξ
Media	183.76	27.25	331.72	40.85	0.09	1.6
Desviación estándar	15.63	5.64	15.38	9.39	0.01	0.38

5.1.2 Método pseudo – marginal por componentes (MALA)

5.1.2.1 Esquema de discretización Puente de difusión modificado (MALA)

En esta sección se muestran los resultados a posteriori obtenidos del método pseudo – marginal por componentes con actualización de los hiperparámetros mediante el algoritmo MALA (tamaño de paso $\epsilon = 0.2$), donde se utilizó un esquema de discretización Euler – Maruyama con un número de discretización $D = 50$ para simular las trayectorias de la EDE, mientras que la verosimilitud individual fue estimada usando $N = 20$ partículas, y la verosimilitud de los $M = 5$ árboles (unidades experimentales) fue estimada con $L = 20$ muestras de efectos aleatorios. El algoritmo fue ejecutado con $I = 2000$ iteraciones con $B = 200$ iteraciones de periodo de quemado (burn-in). Después de ejecutar el método pseudo – por componentes (MALA) se alcanzó una tasa de aceptación de los hiperparámetros de

61.73% y una tasa de aceptación de los parámetros comunes de 7.14% y un tiempo de ejecución de 2.36 minutos.

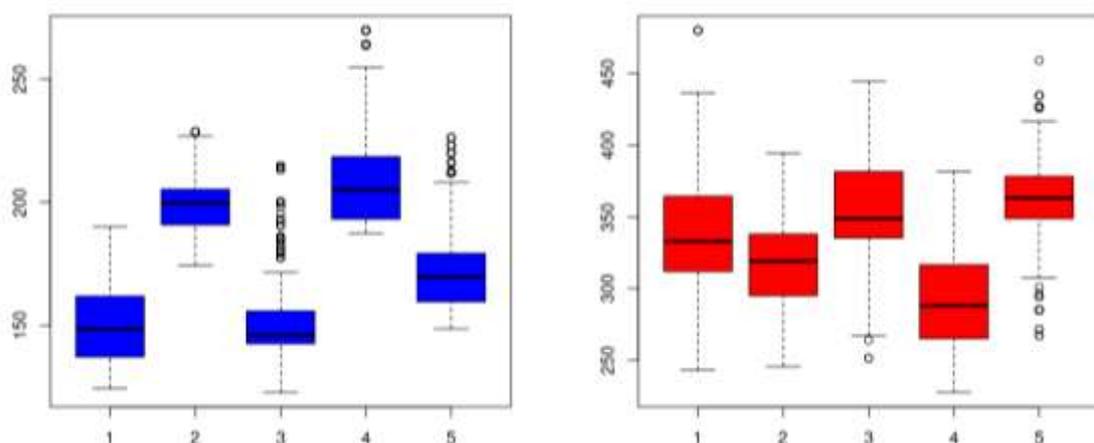
En la Tabla 8, se muestran las tasas de aceptación de los efectos aleatorios $b^{(m)}$ para $m = 1, \dots, 5$ obtenidas del Método pseudo – marginal por componentes – MALA con esquema de EM.

Tabla 8. Tasa de aceptación de los efectos aleatorios (Método pseudo – marginal por componentes – MALA con esquema de EM).

$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
7.73%	8.14%	9.50%	5.86%	6.77%

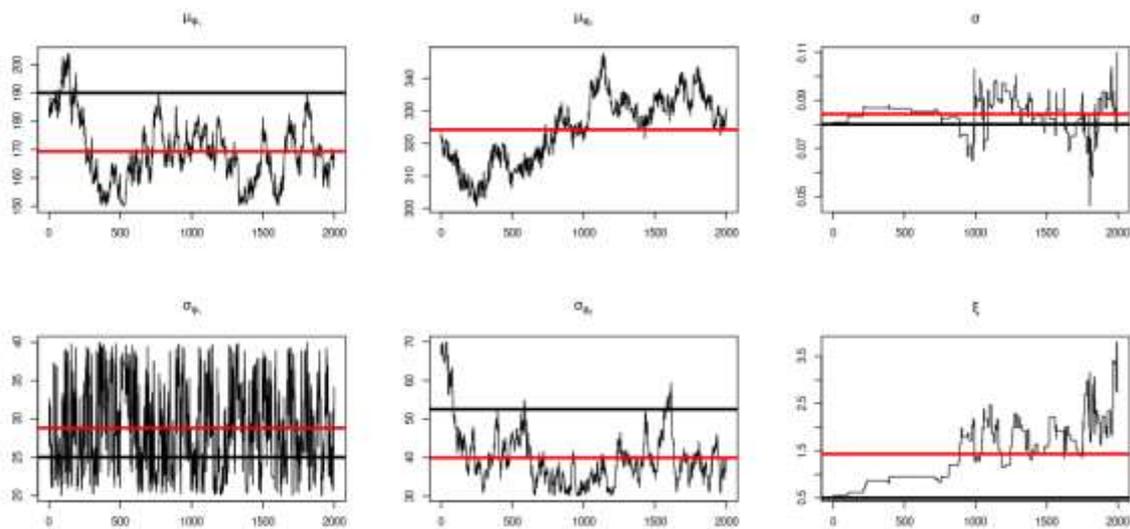
En la Figura 33, se observan los diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del parámetro ϕ_1 (color azul) y ϕ_2 (color rojo).

Figura 33. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización EM).



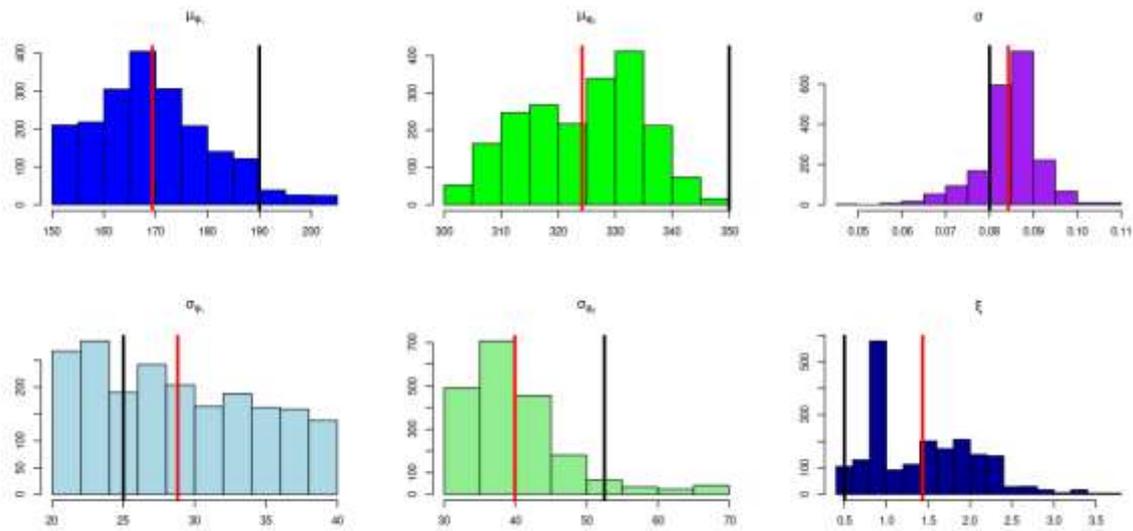
En la Figura 34, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal por componentes (MALA) con esquema de discretización EM y la línea negra indica el valor inicial del parámetro. Se observa que las muestras a posteriori del parámetro σ_{ϕ_1} tiende a converger hacia el valor de la media. Las muestras a posteriori del parámetro σ convergen a partir de las últimas 1000 iteraciones.

Figura 34. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización EM).



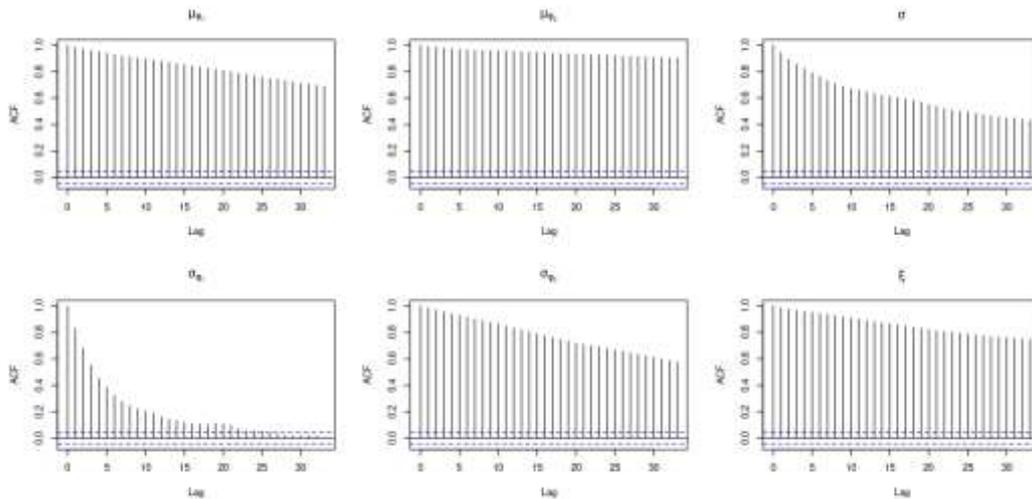
En la Figura 35, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal por componentes (MALA) con esquema de discretización EM y la línea negra indica el valor inicial del parámetro. Se muestra que el histograma de las muestras a posteriori del parámetro σ_{ϕ_2} está sesgado a la derecha.

Figura 35. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización EM).



En la Figura 36, se observa que las muestras a posteriori de los parámetros μ_{ϕ_1} , μ_{ϕ_2} y ξ están altamente correlacionadas, mientras que las muestras a posteriori del parámetro σ están medianamente correlacionadas.

Figura 36. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización EM).



En la Tabla 9, se muestran las estimaciones de los parámetros comunes θ obtenidas del Método pseudo – marginal por componentes – MALA con esquema de discretización EM.

Tabla 9. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización EM).

	μ_{ϕ_1}	σ_{ϕ_1}	μ_{ϕ_2}	σ_{ϕ_2}	σ	ξ
Media	169.34	28.81	324.2	39.94	0.08	1.43
Desviación estándar	11.21	5.73	10.40	7.43	0.01	0.63

5.1.2.2 Esquema de discretización Puente de difusión modificado (MALA)

A continuación, se muestran los resultados a posteriori obtenidos del método pseudo – marginal por componentes con actualización de los hiperparámetros mediante el algoritmo MALA (tamaño de paso $\epsilon = 0.75$), donde se utilizó un esquema de discretización Puente de difusión modificado con un número de discretización $D = 50$ para simular las trayectorias de la EDE, mientras que la verosimilitud individual fue estimada usando $N = 20$ partículas, y la verosimilitud de los $M = 5$ árboles (unidades experimentales) fue estimada con $L = 20$ muestras de efectos aleatorios. El algoritmo fue ejecutado con $I = 2000$ iteraciones con $B = 200$ iteraciones de periodo de quemado (burn-in). Después de ejecutar el método pseudo – por componentes (MALA) se alcanzó una tasa de aceptación de los hiperparámetros de 57.36% y una tasa de aceptación de los parámetros comunes de 52.50% y un tiempo de ejecución de 2.66 minutos.

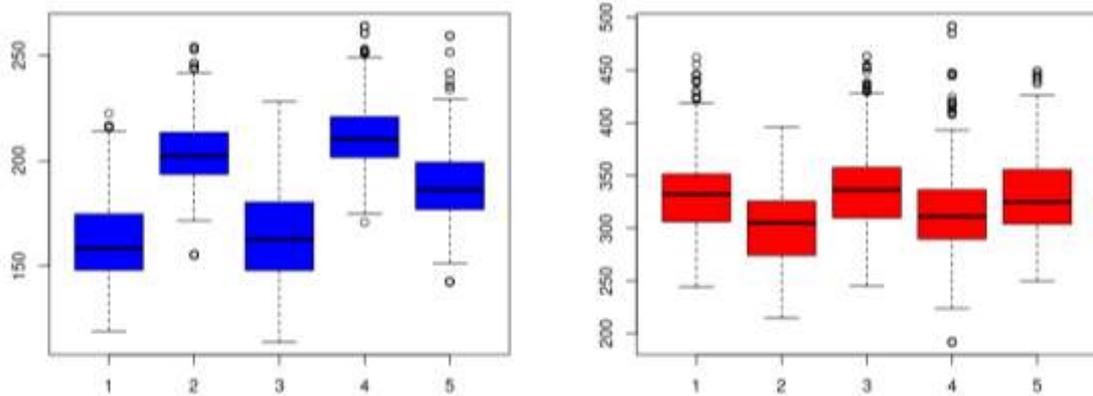
En la Tabla 8, se muestran las tasas de aceptación de los efectos aleatorios $b^{(m)}$ para $m = 1, \dots, 5$ obtenidas del Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado.

Tabla 10. Tasa de aceptación de los efectos aleatorios (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado).

$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
14.27%	12.77%	13.82%	12.91%	13.64%

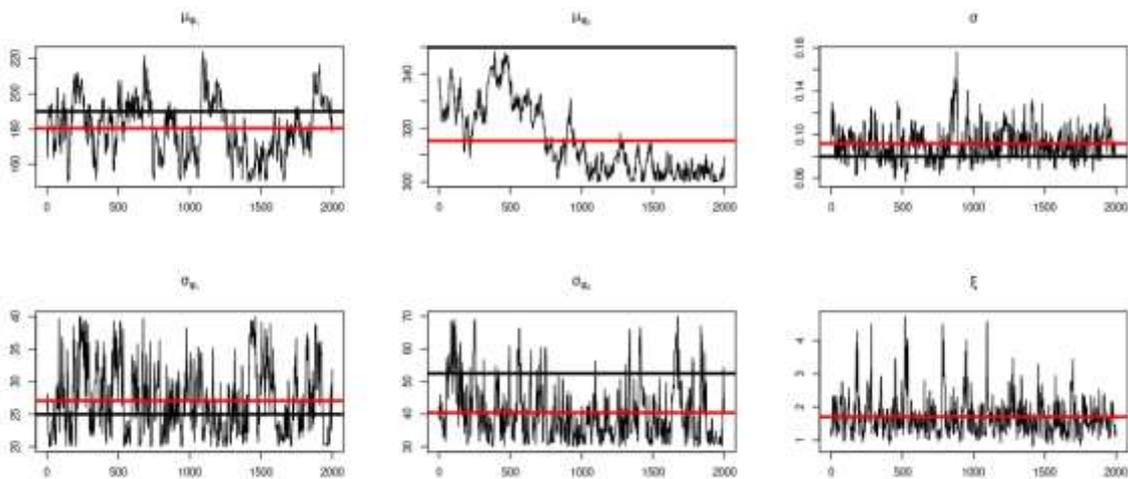
En la Figura 37, se observan los diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del parámetro ϕ_1 (color azul) y ϕ_2 (color rojo).

Figura 37. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado).



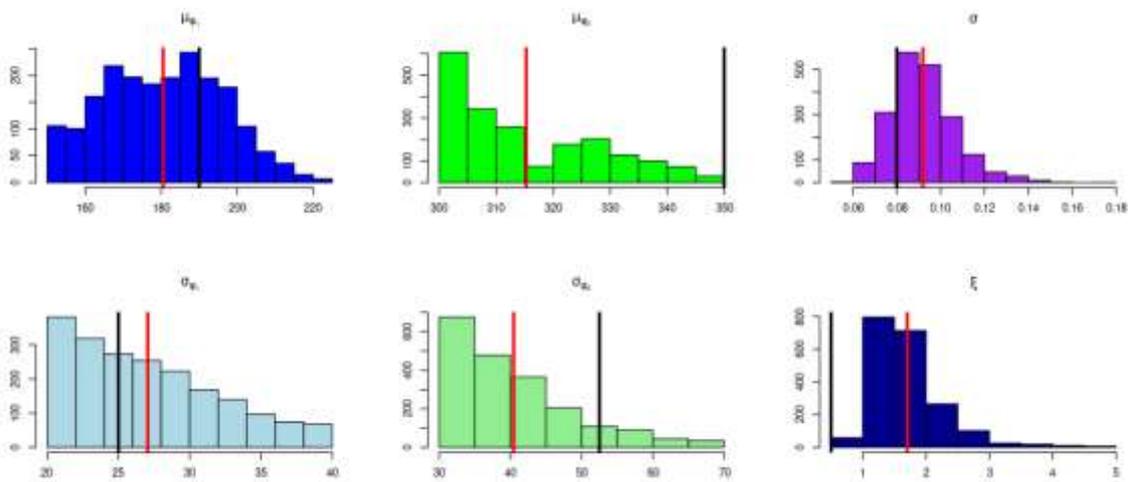
En la Figura 38, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal por componentes (MALA) con esquema de discretización Puente de difusión modificado y la línea negra indica el valor inicial del parámetro. Se observa que existe una convergencia en el parámetro σ_{ϕ_1} , σ_{ϕ_2} , σ y ξ .

Figura 38. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado).



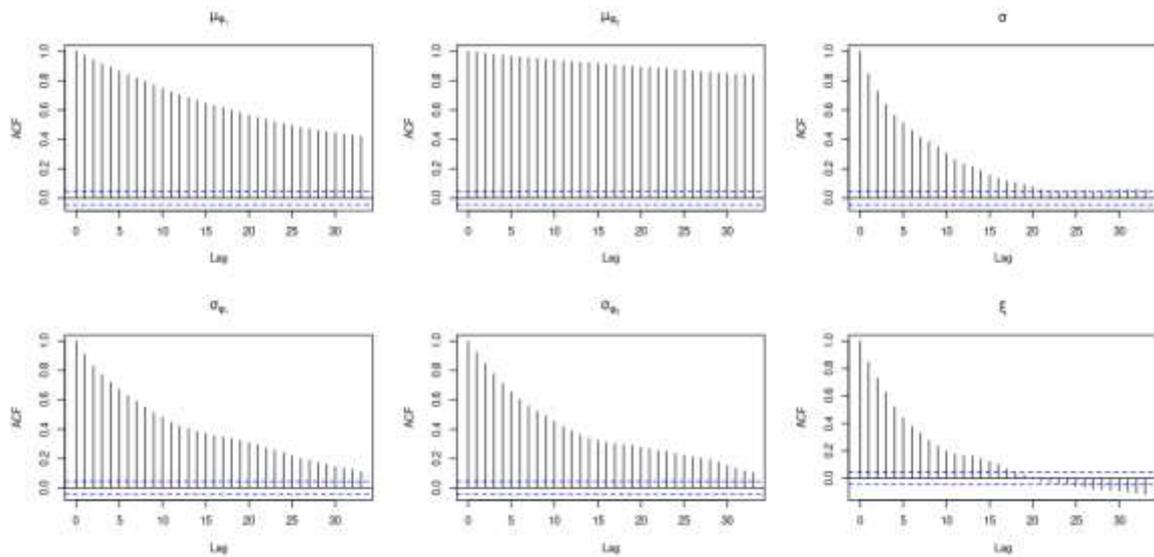
En la Figura 39, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal por componentes (MALA) con esquema de discretización Puente de difusión modificado y la línea negra indica el valor inicial del parámetro. Se observan que los histogramas de las muestras a posteriori de los parámetros μ_{ϕ_2} , σ_{ϕ_2} y σ_{ϕ_1} están sesgados a la derecha.

Figura 39. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado).



En la Figura 40, se observa que las muestras a posteriori del parámetro μ_{ϕ_2} están altamente correlacionadas de acuerdo al diagrama de autocorrelación, mientras que las muestras a posteriori de los demás parámetros presentan una baja correlación.

Figura 40. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado).



En la

Tabla 11, se muestran las estimaciones de los parámetros comunes θ obtenidas del Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado.

Tabla 11. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión modificado).

	μ_{ϕ_1}	σ_{ϕ_1}	μ_{ϕ_2}	σ_{ϕ_2}	σ	ξ
Media	180.51	27.06	315.3	40.46	0.09	1.7
Desviación estándar	15.72	5.12	13.05	8.62	0.01	0.57

5.1.2.3 Esquema de discretización Puente de difusión residual (MALA)

En esta sección se resumen los resultados a posteriori obtenidos del método pseudo – marginal por componentes con actualización de los hiperparámetros mediante el algoritmo MALA (tamaño de paso $\epsilon = 0.5$), donde se utilizó un esquema de discretización Puente de difusión residual con un número de discretización $D = 50$ para simular las trayectorias de la EDE, mientras que la verosimilitud individual fue estimada usando $N = 20$ partículas, y la verosimilitud de los $M = 5$ árboles (unidades experimentales) fue estimada con $L = 20$ muestras de efectos aleatorios. El algoritmo fue ejecutado con $I = 2000$ iteraciones con $B = 200$ iteraciones de periodo de quemado (burn-in). Después de ejecutar el método pseudo – por componentes (MALA) se alcanzó una tasa de aceptación de los hiperparámetros de 43.86% y una tasa de aceptación de los parámetros comunes de 56% y un tiempo de ejecución de 8.11 minutos.

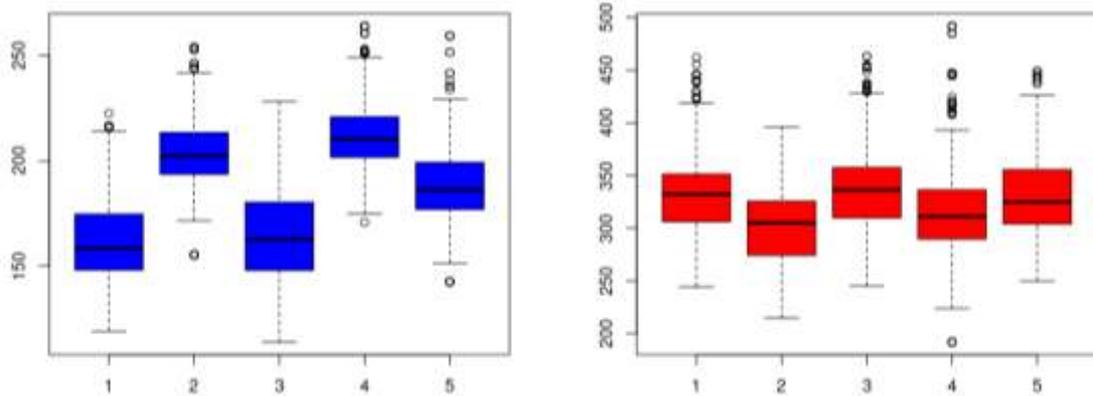
En la Tabla 12, se muestran las tasas de aceptación de los efectos aleatorios $b^{(m)}$ para $m = 1, \dots, 5$ obtenidas del Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual.

Tabla 12. Tasa de aceptación de los efectos aleatorios (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual).

$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
19.50%	19.18%	19.05%	19.82%	19.64%

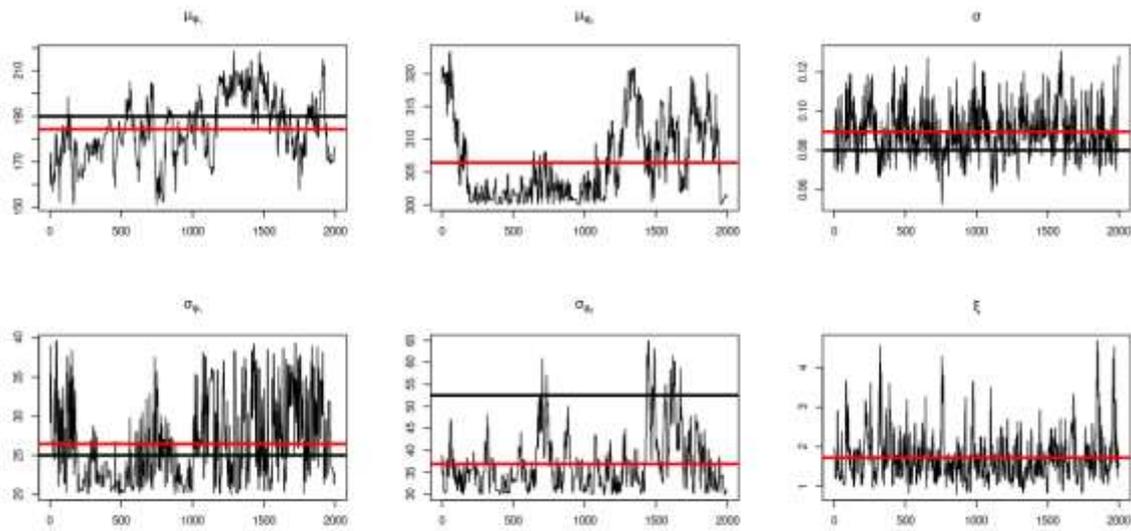
En la Figura 41, se observan los diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del parámetro ϕ_1 (color azul) y ϕ_2 (color rojo).

Figura 41. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual).



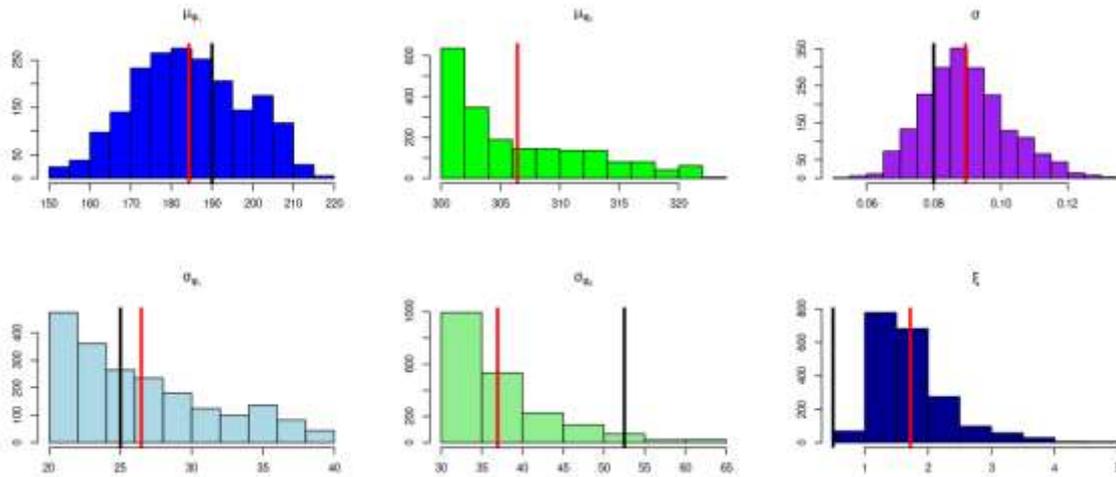
En la Figura 42, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal por componentes (MALA) con esquema de discretización Puente de difusión residual y la línea negra indica el valor inicial del parámetro. Se observa que las muestras a posteriori de los parámetros presentan una mejor convergencia, en particular con los parámetros σ y ξ .

Figura 42. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual).



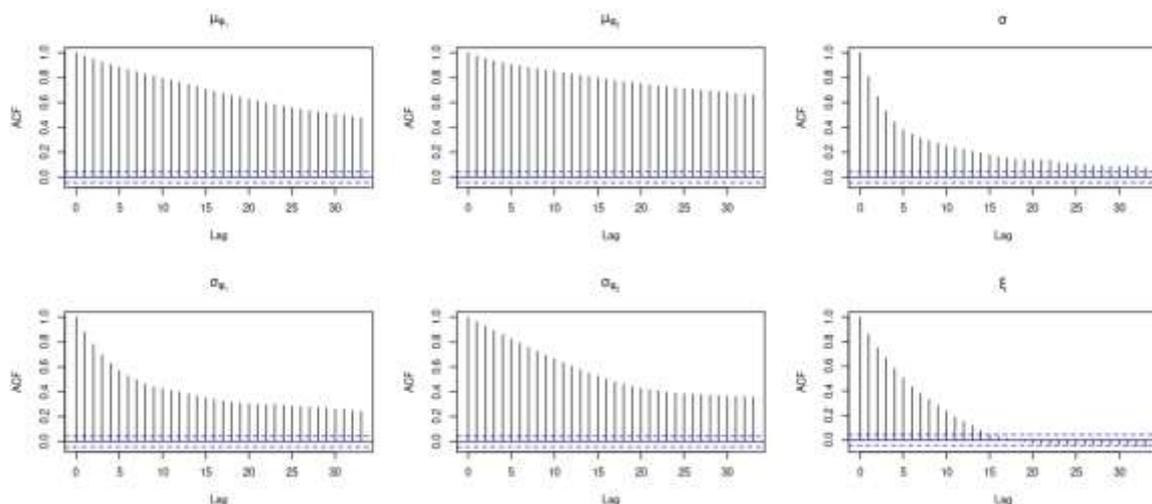
En la Figura 43, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal por componentes (MALA) con esquema de discretización Puente difusión residual y la línea negra indica el valor inicial del parámetro. Se observan que los histogramas de las muestras a posteriori de los parámetros μ_{ϕ_2} , σ_{ϕ_2} y σ_{ϕ_1} están sesgados a la derecha, mientras que los histogramas de las muestras a posteriori de los parámetros μ_{ϕ_1} y σ tienen una forma simétrica.

Figura 43. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual).



En la Figura 44, se observa que las muestras a posteriori de los parámetros μ_{ϕ_1} y μ_{ϕ_2} están medianamente correlacionadas.

Figura 44. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual).



En la Tabla 13, se muestran las estimaciones de los parámetros comunes θ obtenidas del Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual.

Tabla 13. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – MALA con esquema de discretización Puente de difusión residual).

	μ_{ϕ_1}	σ_{ϕ_1}	μ_{ϕ_2}	σ_{ϕ_2}	σ	ξ
Media	184.31	26.47	306.42	36.94	0.09	1.72
Desviación estándar	13.61	5.17	5.89	6.56	0.01	0.60

5.1.3 Método pseudo – marginal por componentes (Hamiltoniano)

5.1.3.1 Esquema de discretización Euler – Maruyama (Hamiltoniano)

En esta sección se muestran los resultados a posteriori obtenidos del método pseudo – marginal por componentes con actualización de los hiperparámetros mediante el algoritmo Hamiltoniano (tamaño de paso $\epsilon = 1/15$ y número de paso $L = 15$), donde se utilizó un esquema de discretización Euler – Maruyama con un número de discretización $D = 50$ para simular las trayectorias de la EDE, mientras que la verosimilitud individual fue estimada usando $N = 20$ partículas, y la verosimilitud de los $M = 5$ árboles (unidades experimentales) fue estimada con $L = 20$ muestras de efectos aleatorios. El algoritmo fue ejecutado con $I = 2000$ iteraciones con $B = 200$ iteraciones de periodo de quemado (burn-in). Después de ejecutar el método pseudo – por componentes (Hamiltoniano) se alcanzó

una tasa de aceptación de los hiperparámetros de 52.14% y una tasa de aceptación de los parámetros comunes de 12.64% y un tiempo de ejecución de 2.25 minutos.

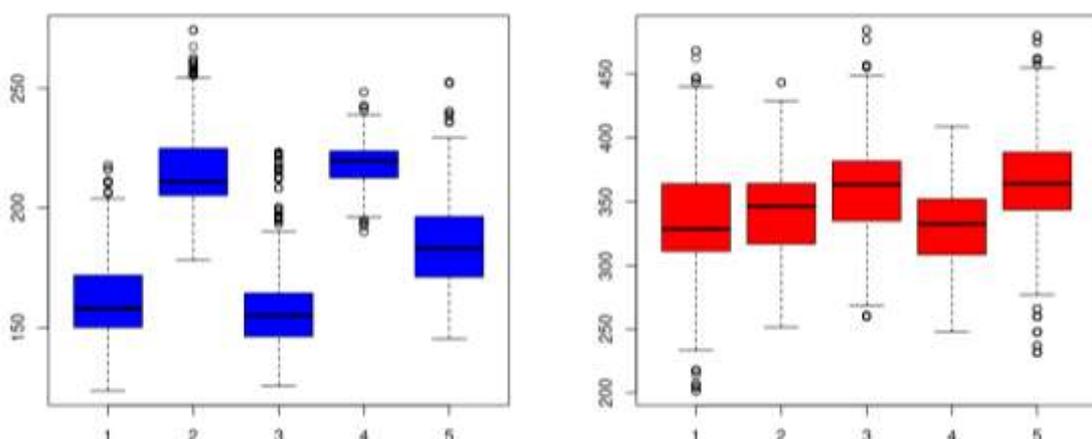
En la Tabla 14, se muestran las tasas de aceptación de los efectos aleatorios $b^{(m)}$ para $m = 1, \dots, 5$ obtenidas del Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM.

Tabla 14. Tasa de aceptación de los efectos aleatorios (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM).

$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
14.46%	12.41%	13.36%	10.14%	11.91%

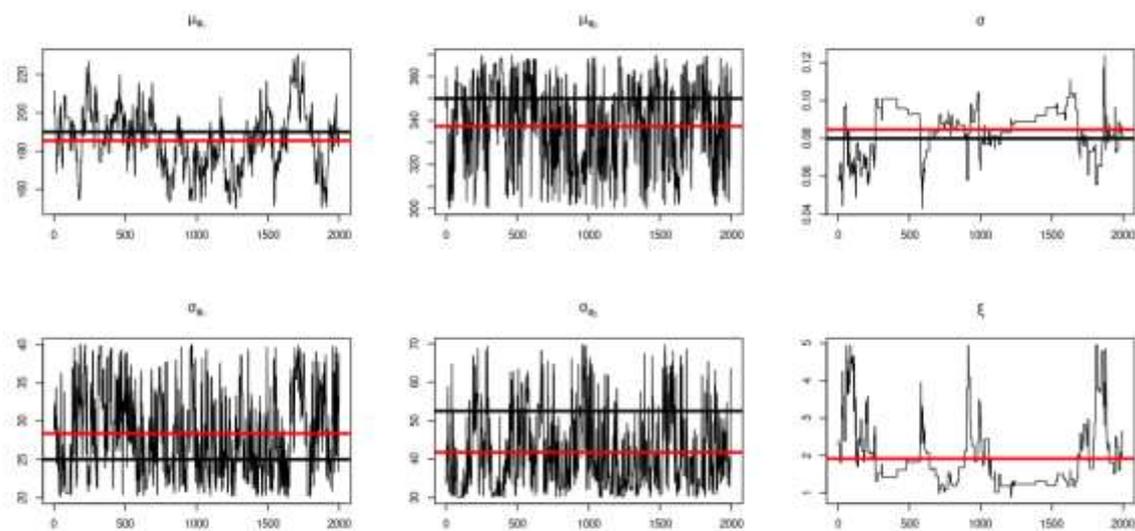
En la Figura 45, se observan los diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del parámetro ϕ_1 (color azul) y ϕ_2 (color rojo).

Figura 45. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM).



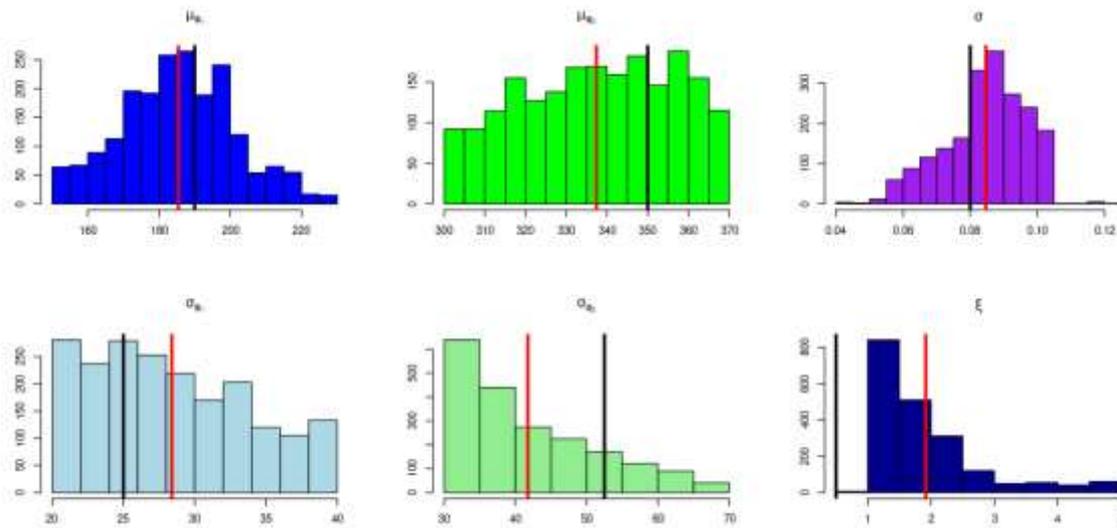
En la Figura 46, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal por componentes (Hamiltoniano) con esquema de discretización EM y la línea negra indica el valor inicial del parámetro. Se observa que las muestras a posteriori presenta una mayor convergencia en los parámetros μ_{ϕ_2} , σ_{ϕ_2} y σ_{ϕ_1} .

Figura 46. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM).



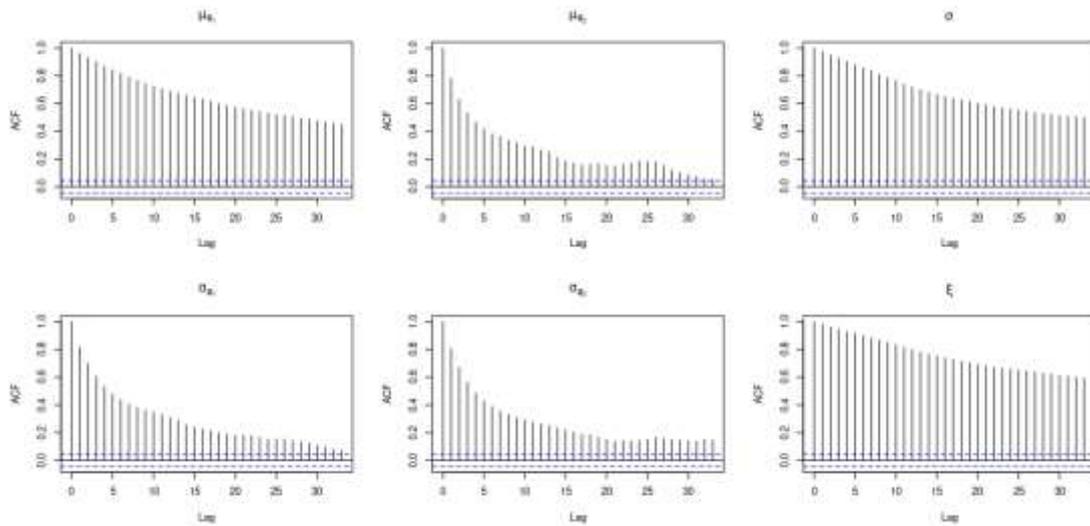
En la Figura 47, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal por componentes (Hamiltoniano) con esquema de discretización EM y la línea negra indica el valor inicial del parámetro. Se muestra que el histograma de las muestras a posteriori del parámetro σ_{ϕ_2} está sesgado hacia la derecha, mientras que el histograma de las muestras a posteriori de los parámetros μ_{ϕ_2} y σ_{ϕ_1} presentan una forma más uniforme.

Figura 47. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM).



En la Figura 48, se observa que las muestras a posteriori de los parámetros μ_{ϕ_1} , σ y ξ están medianamente correlacionadas.

Figura 48. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM).



En la Tabla 15, se muestran las estimaciones de los parámetros comunes θ obtenidas del Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM.

Tabla 15. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización EM).

	μ_{ϕ_1}	σ_{ϕ_1}	μ_{ϕ_2}	σ_{ϕ_2}	σ	ξ
Media	185.43	28.39	337.42	41.76	0.08	1.92
Desviación estándar	15.83	5.5	18.82	9.88	0.01	0.88

5.1.3.2 Esquema de discretización Puente de difusión modificado (Hamiltoniano)

A continuación, se muestran los resultados a posteriori obtenidos del método pseudo – marginal por componentes con actualización de los hiperparámetros mediante el algoritmo Hamiltoniano (tamaño de paso $\epsilon = 1/15$ y número de paso $L = 15$), donde se utilizó un esquema de discretización Puente de difusión modificado con un número de discretización $D = 50$ para simular las trayectorias de la EDE, mientras que la verosimilitud individual fue estimada usando $N = 20$ partículas, y la verosimilitud de los $M = 5$ árboles (unidades experimentales) fue estimada con $L = 20$ muestras de efectos aleatorios. El algoritmo fue ejecutado con $I = 2000$ iteraciones con $B = 200$ iteraciones de periodo de quemado (burn-in). Después de ejecutar el método pseudo – por componentes (Hamiltoniano) se alcanzó una tasa de aceptación de los hiperparámetros de 53.14% y una tasa de aceptación de los parámetros comunes de 49.96% y un tiempo de ejecución de 2.69 minutos.

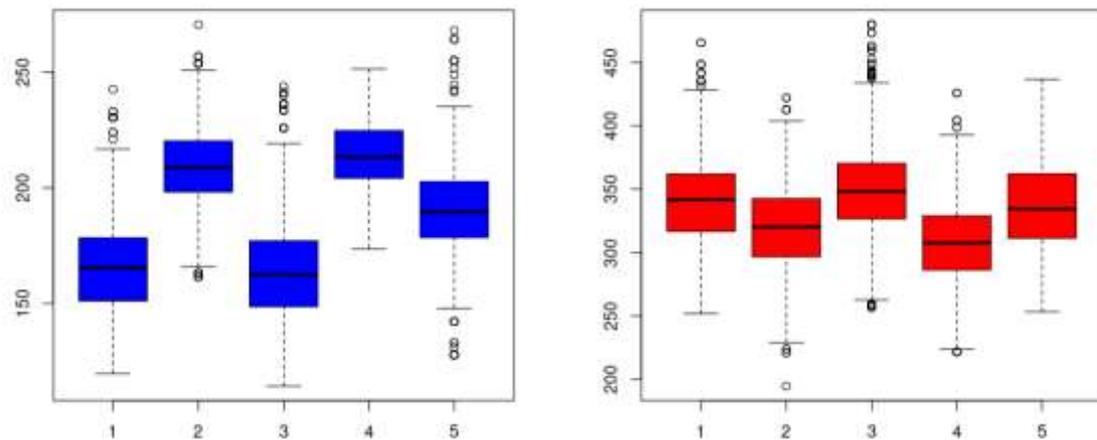
En la Tabla 16, se muestran las tasas de aceptación de los efectos aleatorios $b^{(m)}$ para $m = 1, \dots, 5$ obtenidas del Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado.

Tabla 16. Tasa de aceptación de los efectos aleatorios (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).

$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
18.64%	19.59%	17.32%	17.96%	18.55%

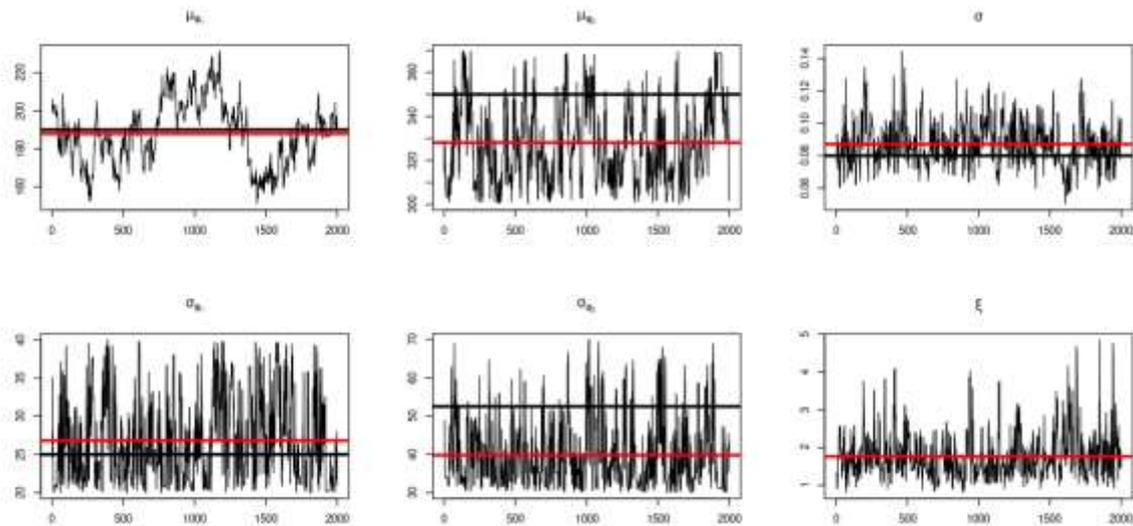
En la Figura 49, se observan los diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del parámetro ϕ_1 (color azul) y ϕ_2 (color rojo).

Figura 49. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).



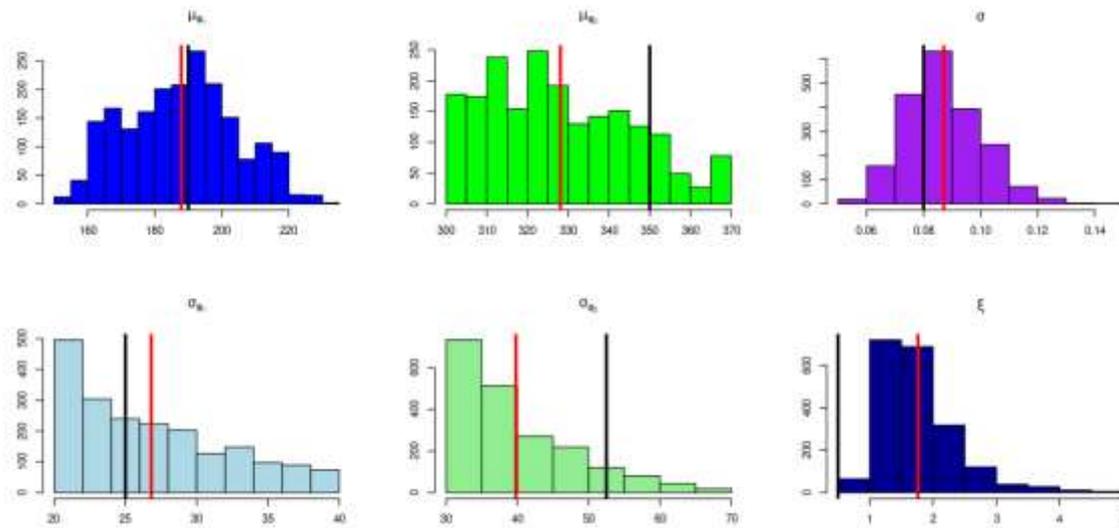
En la Figura 50, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal por componentes (Hamiltoniano) con esquema de discretización Puente de difusión modificado y la línea negra indica el valor inicial del parámetro. Se observa que las muestras a posteriori de los parámetros convergen, en el caso del parámetro μ_{ϕ_1} presenta una convergencia más lenta a comparación de los parámetros restantes.

Figura 50. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).



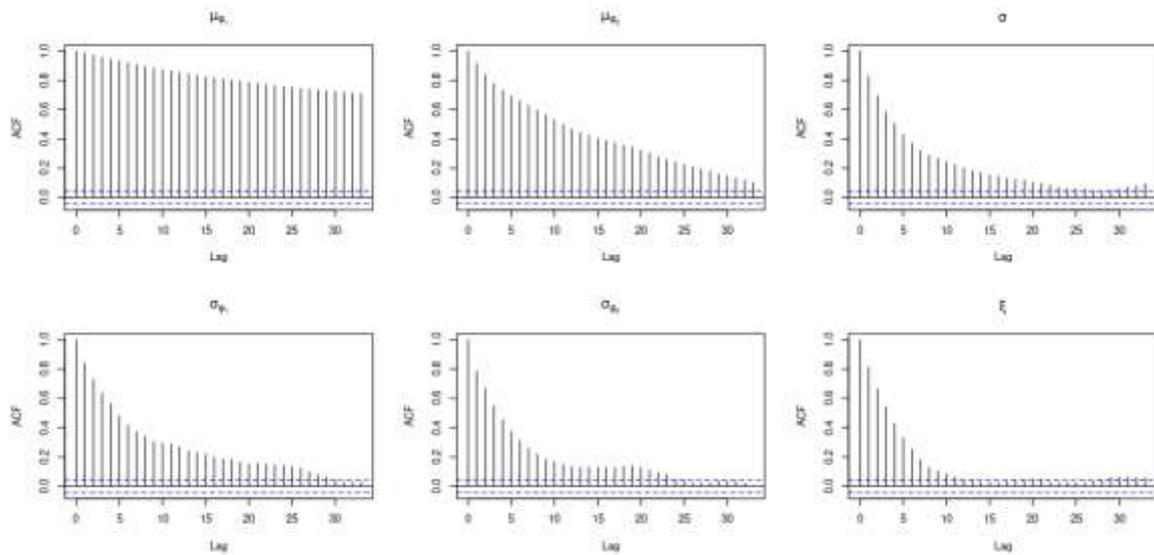
En la Figura 51, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal por componentes (Hamiltoniano) con esquema de discretización Puente de difusión modificado y la línea negra indica el valor inicial del parámetro. Se observa que los histogramas de las muestras a posteriori de los parámetros σ_{ϕ_1} y σ_{ϕ_2} están sesgados a la derecha, mientras que el histograma de las muestras a posteriori del parámetro μ_{ϕ_2} tiene una forma uniforme.

Figura 51. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).



En la Figura 52, se observa que las muestras a posteriori del parámetro μ_{ϕ_1} están medianamente correlacionadas, mientras que las muestras a posteriori de los demás parámetros presentan una baja correlación.

Figura 52. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).



En la Tabla 17, se muestran las estimaciones de los parámetros comunes θ obtenidas del Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado.

Tabla 17. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).

	μ_{ϕ_1}	σ_{ϕ_1}	μ_{ϕ_2}	σ_{ϕ_2}	σ	ξ
Media	187.9	26.82	328.06	39.79	0.09	1.76
Desviación estándar	16.51	5.46	17.79	8.51	0.01	0.6

5.1.3.3 Esquema de discretización Puente de difusión residual (Hamiltoniano)

En este apartado se resumen los resultados a posteriori obtenidos del método pseudo – marginal por componentes con actualización de los hiperparámetros mediante el algoritmo Hamiltoniano (tamaño de paso $\epsilon = 1/15$ y número de paso $L = 15$), donde se utilizó un esquema de discretización Puente de difusión residual con un número de discretización $D = 50$ para simular las trayectorias de la EDE, mientras que la verosimilitud individual fue estimada usando $N = 20$ partículas, y la verosimilitud de los $M = 5$ árboles (unidades experimentales) fue estimada con $L = 20$ muestras de efectos aleatorios. El algoritmo fue ejecutado con $I = 2000$ iteraciones con $B = 200$ iteraciones de periodo de quemado (burn-in). Después de ejecutar el método pseudo – por componentes (Hamiltoniano) se alcanzó una tasa de aceptación de los hiperparámetros de 59.73% y una tasa de aceptación de los parámetros comunes de 50.77% y un tiempo de ejecución de 8.75 minutos.

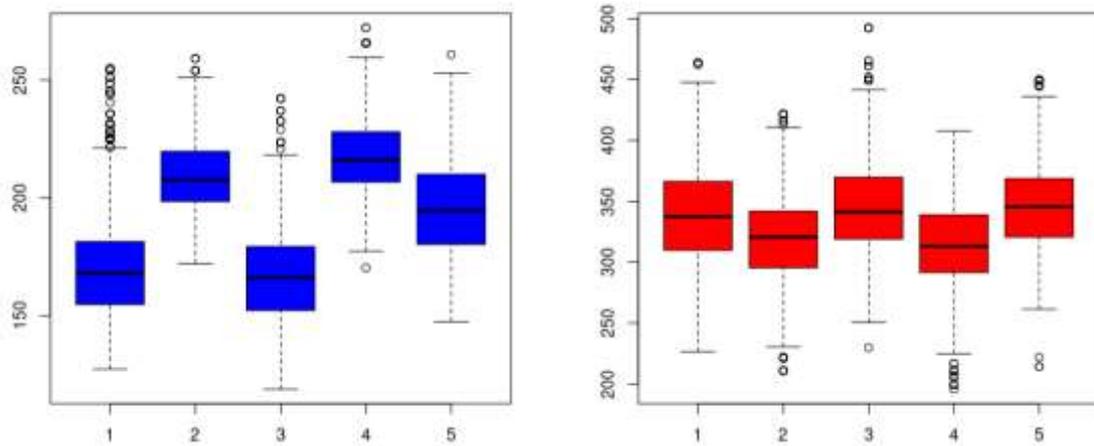
En la Tabla 18, se muestran las tasas de aceptación de los efectos aleatorios $b^{(m)}$ para $m = 1, \dots, 5$ obtenidas del Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión residual.

Tabla 18. Tasa de aceptación de los efectos aleatorios (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión residual).

$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
24.59%	26.36%	23.50%	24.96%	27.27%

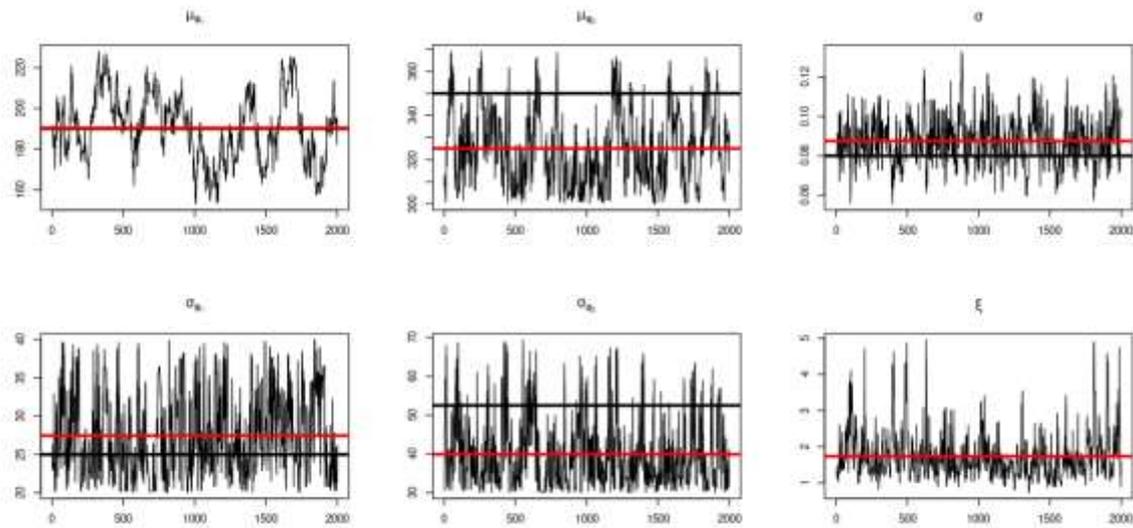
En la Figura 53, se observan los diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del parámetro ϕ_1 (color azul) y ϕ_2 (color rojo).

Figura 53. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión residual).



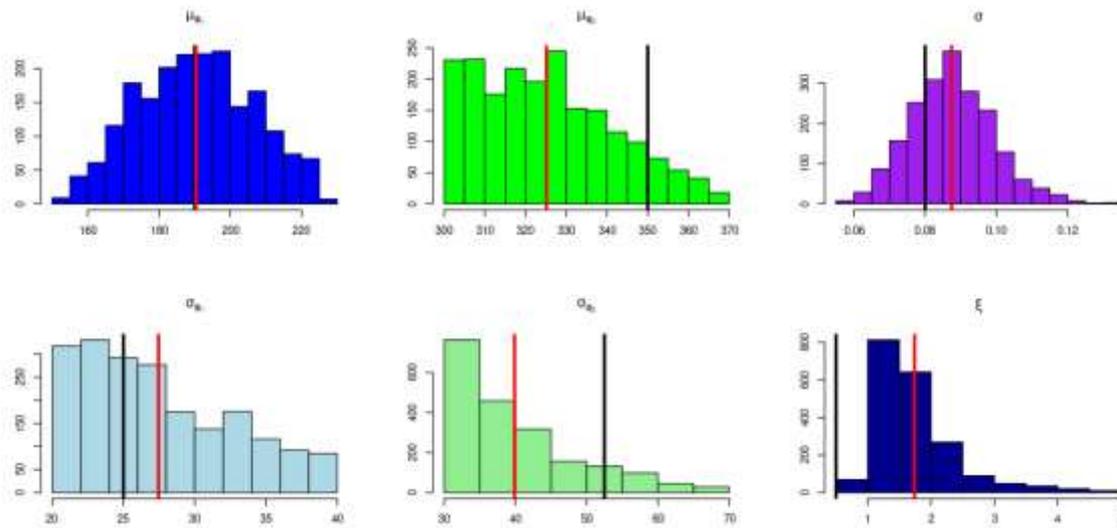
En la Figura 54, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal por componentes (Hamiltoniano) con esquema de discretización Puente de difusión residual y la línea negra indica el valor inicial del parámetro. Se observa las muestras a posteriori de los parámetros convergen, en el caso del parámetro μ_{ϕ_1} se tiene que hay una convergencia más lenta a comparación de los restantes parámetros.

Figura 54. Diagrama de traza de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión residual).



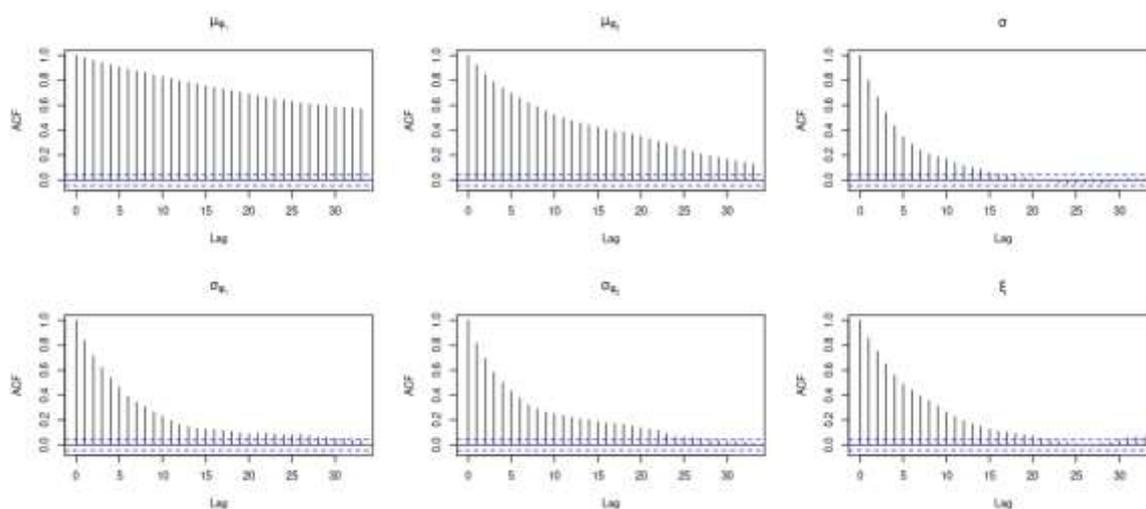
En la Figura 55, la línea roja indica la media a posteriori de las muestras obtenidas del método pseudo – marginal por componentes (Hamiltoniano) con esquema de discretización Puente de difusión residual y la línea negra indica el valor inicial del parámetro. Se observa que

Figura 55. Histogramas de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión residual).



En la Figura 56, se observa que las muestras a posteriori del parámetro μ_{ϕ_1} están medianamente correlacionadas, mientras que las muestras a posteriori de los demás parámetros presentan una baja correlación.

Figura 56. Diagramas de autocorrelación de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión modificado).



En la Tabla 19, se muestran las estimaciones de los parámetros comunes θ obtenidas del Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión residual.

Tabla 19. Resumen estadístico de las muestras a posteriori del modelo de crecimiento del árbol de naranja (Método pseudo – marginal por componentes – Hamiltoniano con esquema de discretización Puente de difusión residual).

	μ_{ϕ_1}	σ_{ϕ_1}	μ_{ϕ_2}	σ_{ϕ_2}	σ	ξ
Media	190.41	27.46	325.12	39.88	0.09	1.74
Desviación estándar	16.27	5.33	16.63	8.85	0.01	0.67

En la Tabla 20, se muestra el desempeño de los tres métodos pseudo – marginales: Aumento Individual (AI), MALA y Hamiltoniano con cada uno de los esquemas de discretización de la EDE, esto es, Euler – Maruyama (EM), Puente de difusión modificado (PDM) y Puente de difusión residual (PDR). Los indicadores de desempeño son tiempo de ejecución (segundos), tamaño de muestra efectivo (ESS: effective sample size) promedio, tasa de tamaño de muestra efectivo (ESS/s) y tasa de aceptación de los parámetros. El tamaño de muestra efectivo es un indicador de correlación entre las muestras a posteriori obtenidas de los métodos MCMC. Si el tamaño de muestra es menor, entonces las muestras presentan una baja correlación. Se observa que los métodos MALA y Hamiltoniano presentan un menor tiempo de ejecución con respecto al método de Aumento Individual (IA), esto se debe a que el método de Aumento Individual requiere ejecutar LN veces el filtro de partículas en comparación a los otros métodos pseudo – marginales que sólo requieren invocar $2N$ veces el filtro de partículas. Mientras que los esquemas de discretización de EM y Puente de difusión modificado presentan un menor tiempo de ejecución con respecto al esquema de Puente de difusión residual, esto se debe a que el esquema de discretización de Puente de difusión residual requiere la ejecución del método de Euler para aproximar el término drift de la EDE, mientras que los otros dos esquemas restantes no lo tienen que hacer. También se observa que el método Hamiltoniano con los esquemas de EM y Puente de difusión modificado presentan una mayor tasa de tamaño de muestra a comparación de los otros métodos pseudo – marginales, lo que indica que las muestras a posteriori obtenidas con el método Hamiltoniano (EM y PDM) presentan una menor correlación y su tiempo de ejecución es menor. Finalmente, se observa que los métodos MALA y Hamiltoniano (ambos con el esquema PDM) presentan altas tasas de aceptación de los parámetros comunes θ .

Tabla 20. Indicadores de desempeño de los métodos pseudo – marginales.

	Tiempo de ejecución (segundo)	Tamaño de muestra efectivo promedio (ESS)	Tasa de tamaño de muestra efectivo (ESS/segundo)	Tasa de aceptación de parámetros (%)
IA – EM	335.52	24.28	0.07	7.40
IA – PDM	576.93	70.96	0.12	18.21
IA – PDR	1401.30	62.11	0.04	15.41
MALA – EM	141.75	42.40	0.30	43.22
MALA – PDM	159.70	81.51	0.51	55.31
MALA – PDR	486.77	71.58	0.15	46.99
Hamiltoniano – EM	134.86	75.48	0.56	38.19
Hamiltoniano – PDM	161.58	126.75	0.78	51.93
Hamiltoniano – PDR	524.77	122.25	0.23	56.34

5.1.4 *Proceso de Ornstein-Uhlenbeck bidimensional*

Las observaciones del proceso de Ornstein – Uhlenbeck bidimensional fueron obtenidas a partir de la simulación de los estados no observables $X_t^{(m)} = (X_t^{(1,m)}, X_t^{(2,m)})^T$ mediante un esquema de Euler – Maruyama con la misma condición inicial $x_0^{(m)} = (1,1)^T$

para $m = 1, \dots, 5$ (unidades experimentales) y paso de discretización $\Delta = 0.1$ en el intervalo $[0,1]$ con los parámetros comunes verdaderos $\theta = (\alpha_1, \alpha_2, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}, \sigma_1, \sigma_2, \xi_1, \xi_2) = (3, 2.5, 1.8, 2, 1, 1.5, 0.3, 0.5, 0.1, 0.5)$ y parámetros poblacionales $\psi = (45, 100, 100, 25)$.

5.1.4.1 Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs

En esta sección se resumen los resultados a posteriori del algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs. El esquema de muestreo de Gibbs empleó tres bloques de actualización: efectos aleatorios, parámetros comunes e hiperparámetros. Se utilizó un esquema de discretización Euler – Maruyama con un número de discretización $D = 20$ para simular las trayectorias de la EDE, y un filtro de partículas con $N = 20$ partículas para estimar la verosimilitud de $M = 5$ unidades experimentales. El algoritmo fue ejecutado con $I = 5000$ iteraciones, donde las primeras $B = 2500$ iteraciones son del periodo de quemado. Después de ejecutar el algoritmo se obtuvo una tasa de aceptación de los parámetros comunes de 2.40% y un tiempo de ejecución de 69.85 minutos.

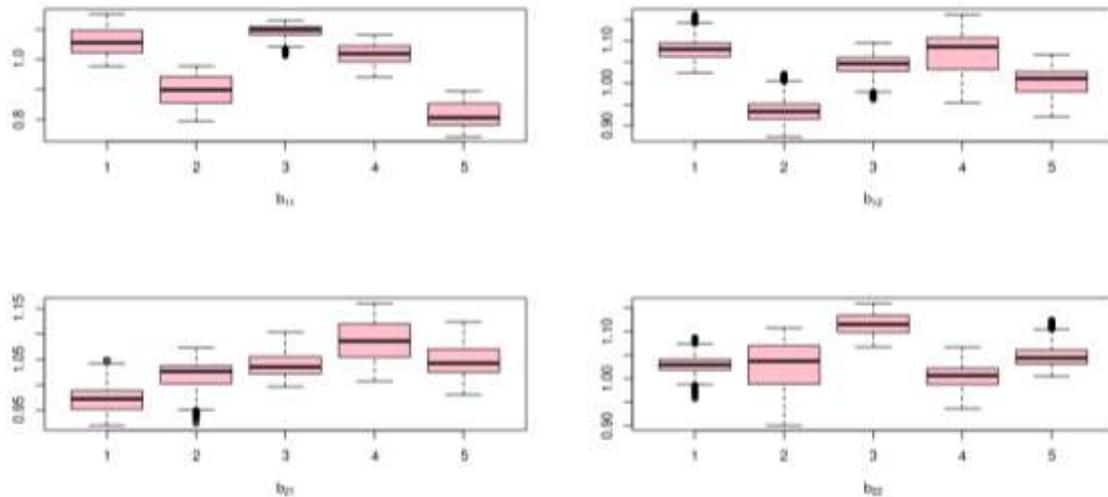
En la Tabla 21, se muestran las tasas de aceptación de los efectos aleatorios $b^{(m)}$ para $m = 1, \dots, 5$ obtenidas del Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs.

Tabla 21. Tasa de aceptación de los efectos aleatorios $b^{(m)}$ (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs).

$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
38.32%	30.12%	24.60%	29.04%	34.36%

En la Figura 57, se muestran los diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo $(b_{11}^{(m)}, b_{12}^{(m)}, b_{21}^{(m)}, b_{22}^{(m)})$ para $m = 1, \dots, 5$ obtenidas del Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs.

Figura 57. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios $b^{(m)}$ (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs).



En la Figura 58 y la Figura 59, se observan los histogramas, diagrama de traza y diagramas de autocorrelación de los parámetros α_2 y ξ_2 , respectivamente. La línea negra indica la media a posteriori de las muestras obtenidas del Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs y la línea roja indica el valor verdadero del parámetro.

Figura 58. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro α_2 (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs).

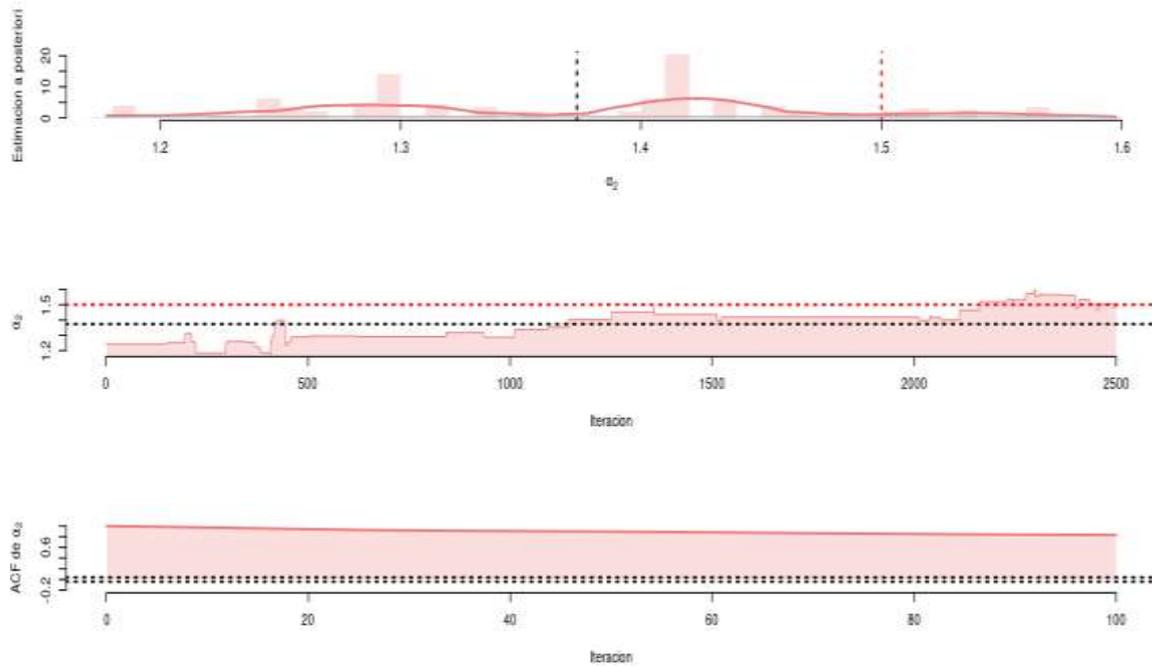
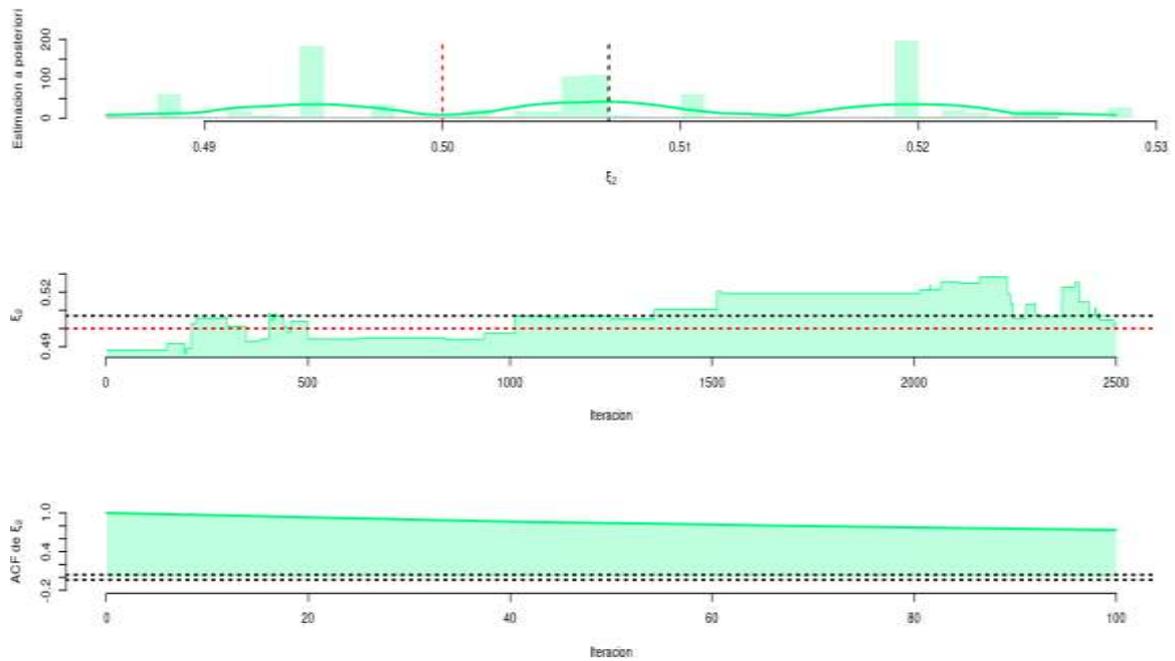


Figura 59. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro ξ_2 (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs).



En la Tabla 22, se muestran las estimaciones de los parámetros comunes θ obtenidas del Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs. Como parte del resumen estadístico se contempla el IACT (integrated autocorrelation time), que corresponde al área bajo la curva del diagrama de autocorrelación. Si el valor bajo de IACT, entonces la correlación entre las muestras a posteriori es baja.

Tabla 22. Resumen estadístico de los parámetros comunes θ (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs).

	α_1	α_2	β_{11}	β_{12}	β_{21}	β_{22}	σ_1	σ_2	ξ_1	ξ_2
Valor verdadero	3	2.5	1.8	2	1	1.5	0.3	0.5	0.1	0.5
Media	1.86	1.30	1.26	1.13	1.22	1.37	0.24	0.17	0.04	0.51
Desviación estándar	0.05	0.10	0.06	0.05	0.06	0.10	0.01	0.01	0.01	0.01
IACT	152.24	168.91	165.80	151.18	171.14	182.42	168.32	158.73	165.27	173.20

En la Tabla 23, se muestran las tasas de aceptación de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$ obtenidas del Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs.

Tabla 23. Tasa de aceptación de los hiperparámetros ψ (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs).

v_{11}	v_{12}	v_{21}	v_{22}
----------	----------	----------	----------

41.52%

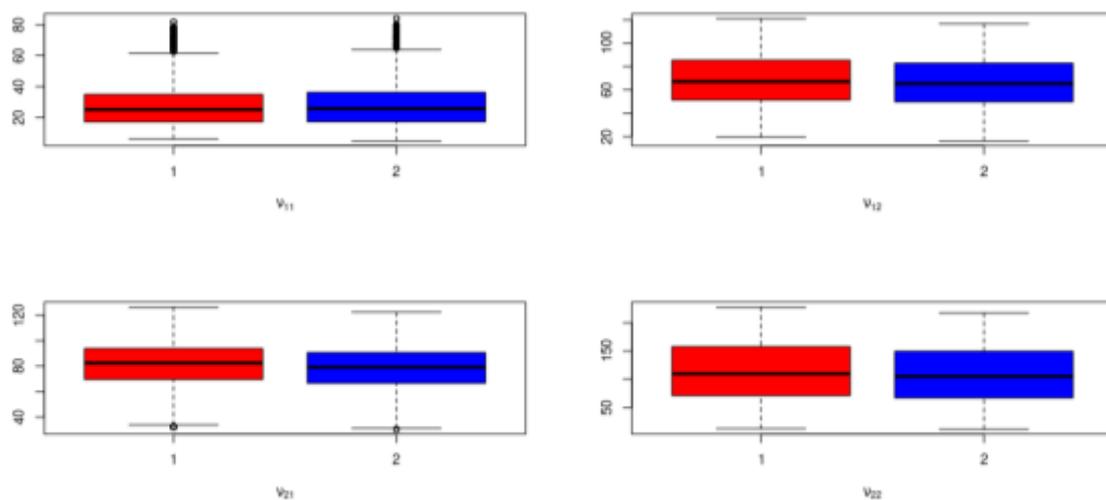
58.92%

60.36%

63.12%

En la Figura 60, se observan los diagramas de caja y bigotes de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$, obtenidos del Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs.

Figura 60. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros ψ (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs).



En la Tabla 24, se reportan las estimaciones de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$ del Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs.

Tabla 24. Resumen estadístico de los hiperparámetros ψ (Algoritmo Metropolis – Hastings dentro de un esquema de muestreo de Gibbs).

Valor	$v_{11} = 25$	$v_{12} = 100$	$v_{21} = 100$	$v_{22} = 45$
verdadero				

$Ga(\mathbf{v}_{l,k}, \mathbf{v}_{l,k})$								
$Ga(\mathbf{v}_{l,k}^{(\alpha)}, \mathbf{v}_{l,k}^{(\beta)})$	$v_{11}^{(\alpha)}$	$v_{11}^{(\beta)}$	$v_{12}^{(\alpha)}$	$v_{12}^{(\beta)}$	$v_{21}^{(\alpha)}$	$v_{21}^{(\beta)}$	$v_{22}^{(\alpha)}$	$v_{22}^{(\beta)}$
Media	27.89	28.47	68.85	67.00	81.62	78.61	113.99	109.36
Desviación estándar	14.98	15.53	21.99	21.43	18.69	18.21	53.02	51.59

5.1.4.2 Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs

En esta sección se resumen los resultados a posteriori del algoritmo Metropolis – Hastings correlacionado ($\rho = 0.999$) dentro de un esquema de muestreo de Gibbs. El esquema de muestreo de Gibbs empleó tres bloques de actualización: efectos aleatorios, parámetros comunes e hiperparámetros. Se utilizó un esquema de discretización Euler – Maruyama con un número de discretización $D = 20$ para simular las trayectorias de la EDE, y un filtro de partículas con $N = 20$ partículas para estimar la verosimilitud de $M = 5$ unidades experimentales. El algoritmo fue ejecutado con $I = 5000$ iteraciones, donde las primeras $B = 2500$ iteraciones son del periodo de quemado. Después de ejecutar el algoritmo se obtuvo una tasa de aceptación de los parámetros comunes de 18.28% y un tiempo de ejecución de 65.03 minutos.

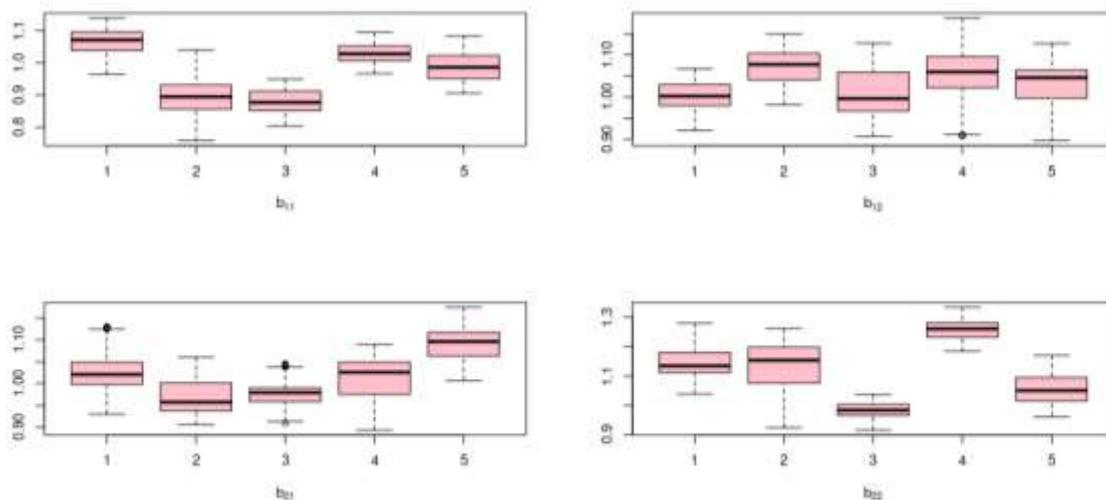
En la Tabla 25, se muestran las tasas de aceptación de los efectos aleatorios $b^{(m)}$ para $m = 1, \dots, 5$ obtenidas del Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs.

Tabla 25. Tasa de aceptación de los efectos aleatorios $\mathbf{b}^{(m)}$ (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).

$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
72.00%	68.44%	57.52%	71.16%	71.04%

En la Figura 61, se muestran los diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo $(b_{11}^{(m)}, b_{12}^{(m)}, b_{21}^{(m)}, b_{22}^{(m)})$ para $m = 1, \dots, 5$ obtenidas del Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs.

Figura 61. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios $\mathbf{b}^{(m)}$ (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).



En la Figura 62 y Figura 63, se observan los histogramas, diagrama de traza y diagramas de autocorrelación de los parámetros α_2 y ξ_2 , respectivamente. La línea negra indica la media a posteriori de las muestras obtenidas del Algoritmo Metropolis – Hastings

correlacionado dentro de un esquema de muestreo de Gibbs y la línea roja indica el valor verdadero del parámetro.

Figura 62. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro α_2 (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).

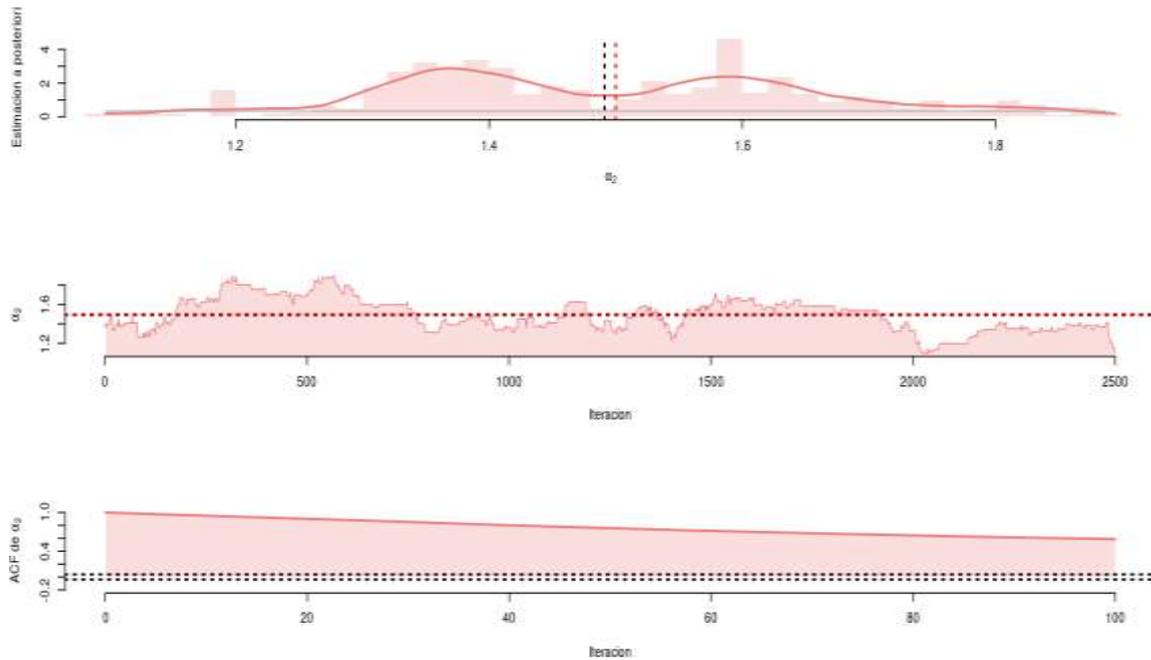
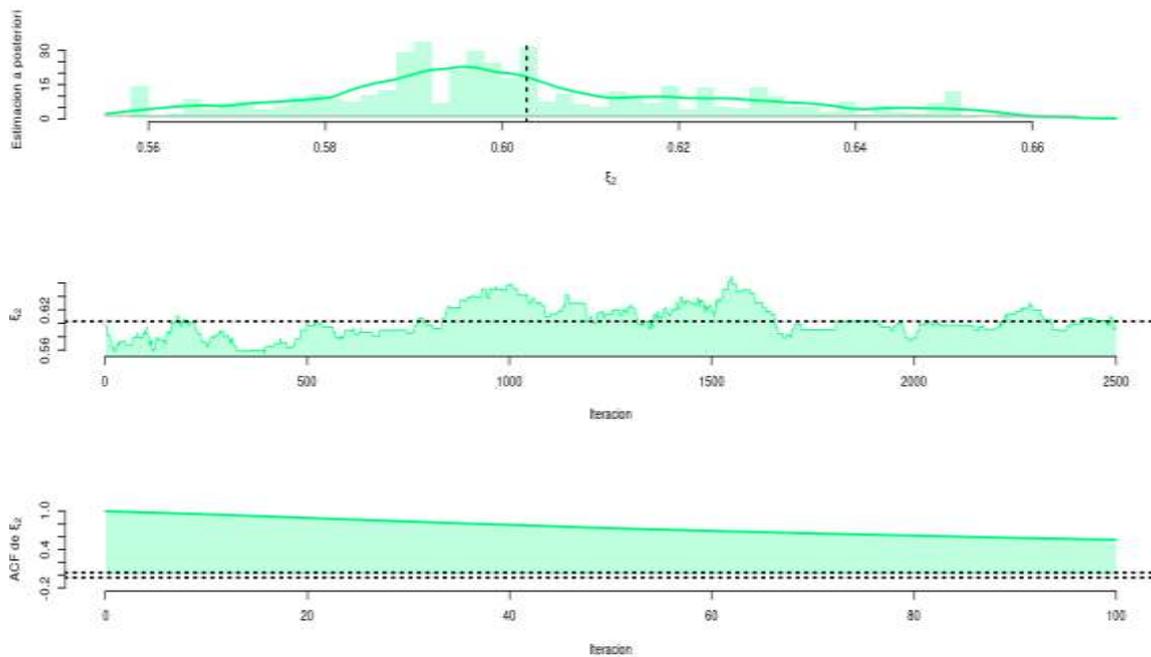


Figura 63. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro ξ_2 (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).



En la Tabla 26, se muestran las estimaciones de los parámetros comunes θ obtenidas del Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs.

Tabla 26. Resumen estadístico de los parámetros comunes θ (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).

	α_1	α_2	β_{11}	β_{12}	β_{21}	β_{22}	σ_1	σ_2	ξ_1	ξ_2
Valor verdadero	3	2.5	1.8	2	1	1.5	0.3	0.5	0.1	0.5
Media	2.52	1.24	1.50	0.97	1.13	1.49	0.17	0.26	0.05	0.60

Desviación estándar	0.21	0.25	0.52	0.28	0.11	0.17	0.03	0.03	0.01	0.02
IACT	173.07	180.45	190.90	182.64	170.53	156.67	180.25	171.37	138.31	153.01

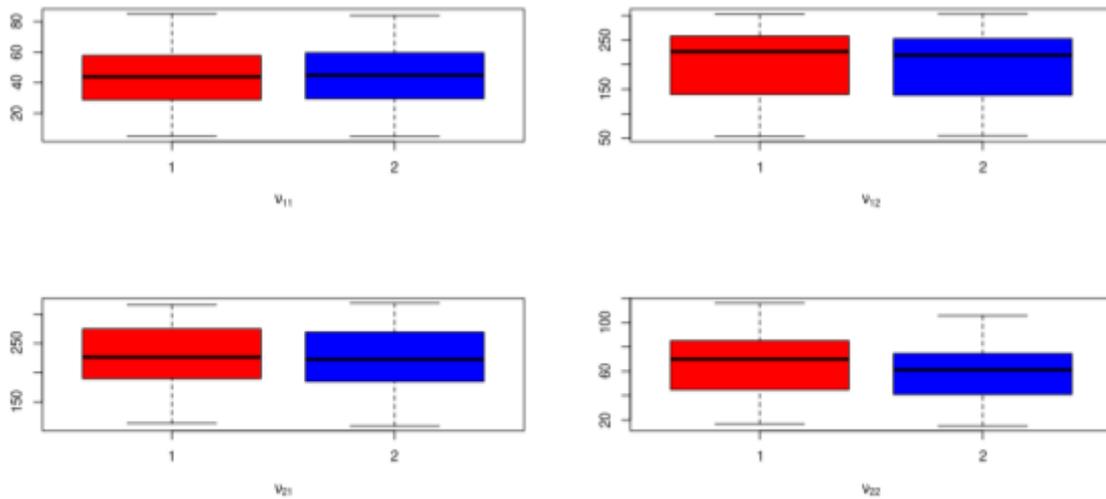
En la Tabla 27, se muestran las tasas de aceptación de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$ obtenidas del Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs

Tabla 27. Tasa de aceptación de los hiperparámetros ψ (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).

v_{11}	v_{12}	v_{21}	v_{22}
50.12%	73.60%	74.52%	56.48%

En la Figura 64, se observan los diagramas de caja y bigotes de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$, obtenidos del Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs.

Figura 64. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros ψ (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).



En la Tabla 28, se reportan las estimaciones de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$ del Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs.

Tabla 28. Resumen estadístico de los hiperparámetros ψ (Algoritmo Metropolis – Hastings correlacionado dentro de un esquema de muestreo de Gibbs).

Valor								
verdadero	$v_{11} = 25$	$v_{12} = 100$	$v_{21} = 100$	$v_{22} = 45$				
$Ga(v_{l,k}, v_{l,k})$								
$Ga(v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$	$v_{11}^{(\alpha)}$	$v_{11}^{(\beta)}$	$v_{12}^{(\alpha)}$	$v_{12}^{(\beta)}$	$v_{21}^{(\alpha)}$	$v_{21}^{(\beta)}$	$v_{22}^{(\alpha)}$	$v_{22}^{(\beta)}$
Media	43.55	44.72	200.98	194.92	228.25	225.50	65.47	58.42

Desviación estándar	17.47	18.12	72.58	71.88	47.95	48.61	22.88	19.71
--------------------------------	-------	-------	-------	-------	-------	-------	-------	-------

5.1.4.3 Algoritmo MALA dentro de un esquema de muestreo de Gibbs

En esta sección se resumen los resultados a posteriori del algoritmo MALA dentro de un esquema de muestreo de Gibbs. El algoritmo MALA utilizó el parámetro $\epsilon = 0.25$. El esquema de muestreo de Gibbs empleó tres bloques de actualización: efectos aleatorios, parámetros comunes e hiperparámetros. Se utilizó un esquema de discretización Euler – Maruyama con un número de discretización $D = 20$ para simular las trayectorias de la EDE, y un filtro de partículas con $N = 20$ partículas para estimar la verosimilitud de $M = 5$ unidades experimentales. El algoritmo fue ejecutado con $I = 5000$ iteraciones, donde las primeras $B = 2500$ iteraciones son del periodo de quemado. Después de ejecutar el algoritmo se obtuvo una tasa de aceptación de los parámetros comunes de 0.4% y un tiempo de ejecución de 71.20 minutos.

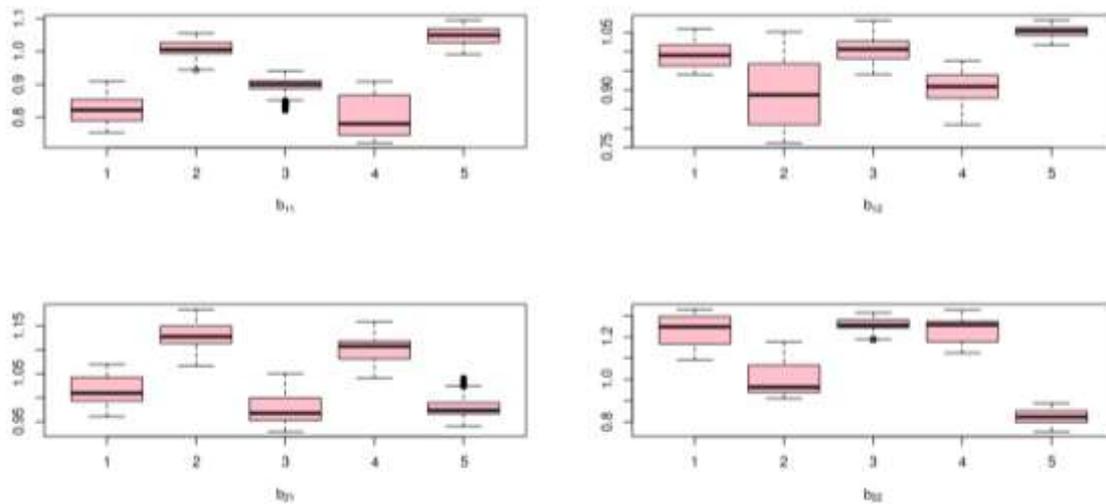
En la Tabla 29, se muestran las tasas de aceptación de los efectos aleatorios $b^{(m)}$ para $m = 1, \dots, 5$ obtenidas del Algoritmo MALA dentro de un esquema de muestreo de Gibbs.

Tabla 29. Tasa de aceptación de los efectos aleatorios $b^{(m)}$ (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).

$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
56.32%	60.20%	58.04%	57.96%	59.60%

En la Figura 65, se muestran los diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo $(b_{11}^{(m)}, b_{12}^{(m)}, b_{21}^{(m)}, b_{22}^{(m)})$ para $m = 1, \dots, 5$ obtenidas del Algoritmo MALA dentro de un esquema de muestreo de Gibbs.

Figura 65. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios $b^{(m)}$ (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).



En la Figura 66 y Figura 67, se observan los histogramas, diagrama de traza y diagramas de autocorrelación de los parámetros α_2 y ξ_2 , respectivamente. La línea negra indica la media a posteriori de las muestras obtenidas del Algoritmo MALA dentro de un esquema de muestreo de Gibbs y la línea roja indica el valor verdadero del parámetro.

Figura 66. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro α_2 (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).

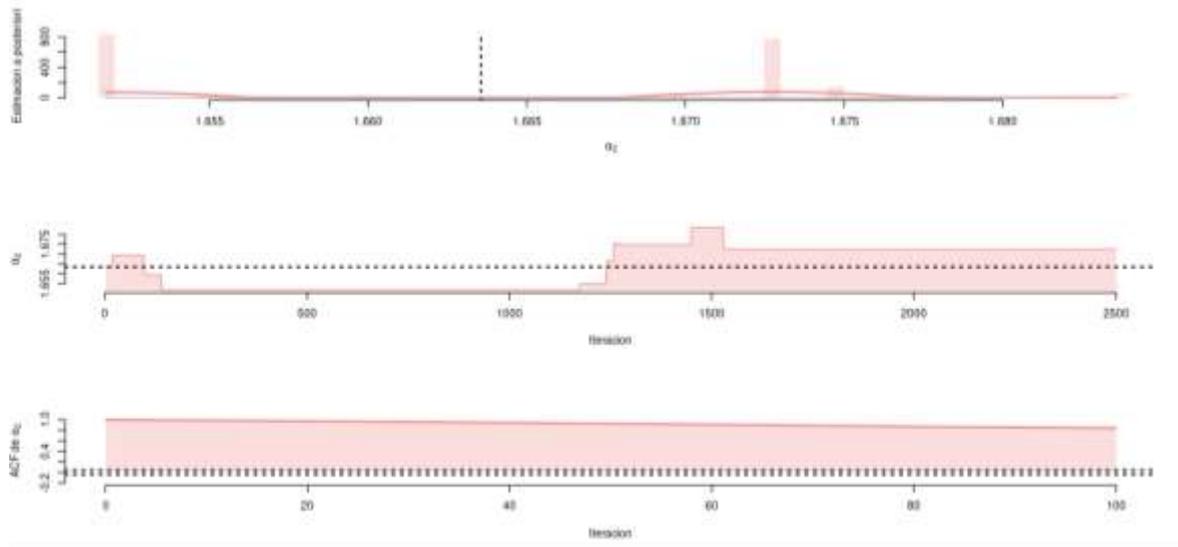
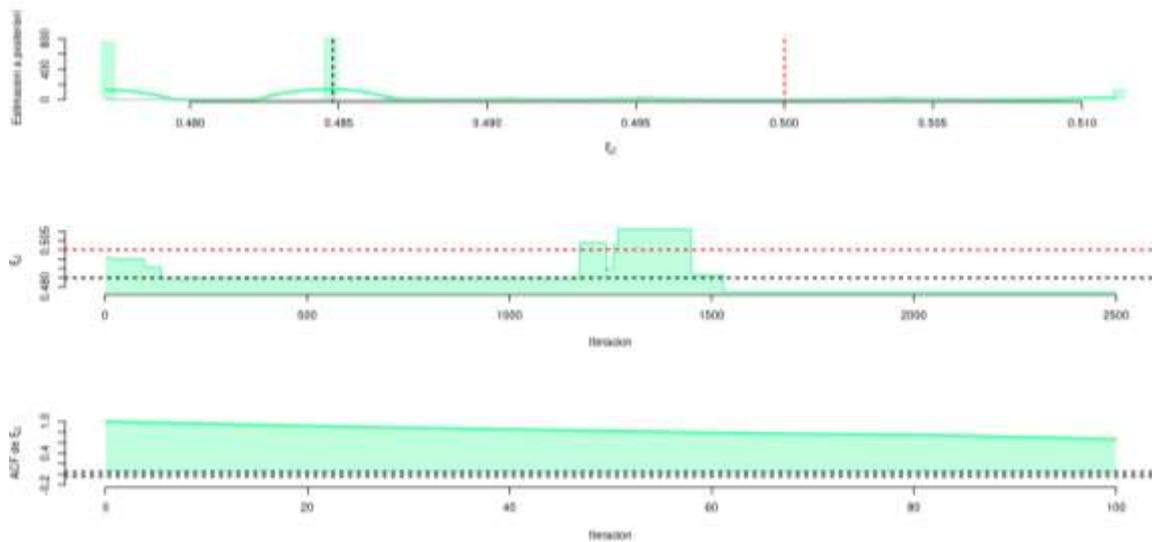


Figura 67. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro ξ_2 (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).



En la Tabla 30, se muestran las estimaciones de los parámetros comunes θ obtenidas del Algoritmo MALA dentro de un esquema de muestreo de Gibbs.

Tabla 30. Resumen estadístico de los parámetros comunes θ (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).

	α_1	α_2	β_{11}	β_{12}	β_{21}	β_{22}	σ_1	σ_2	ξ_1	ξ_2
Valor verdadero	3	2.5	1.8	2	1	1.5	0.3	0.5	0.1	0.5
Media	1.89	2.11	1.40	1.71	0.83	1.66	0.23	0.13	0.04	0.48
Desviación estándar	0.005	0.01	0.03	0.04	0.01	0.01	0.01	0.01	0.01	0.01
IACT	114.92	171.28	188.45	196.00	187.85	187.70	178.49	101.62	179.15	168.13

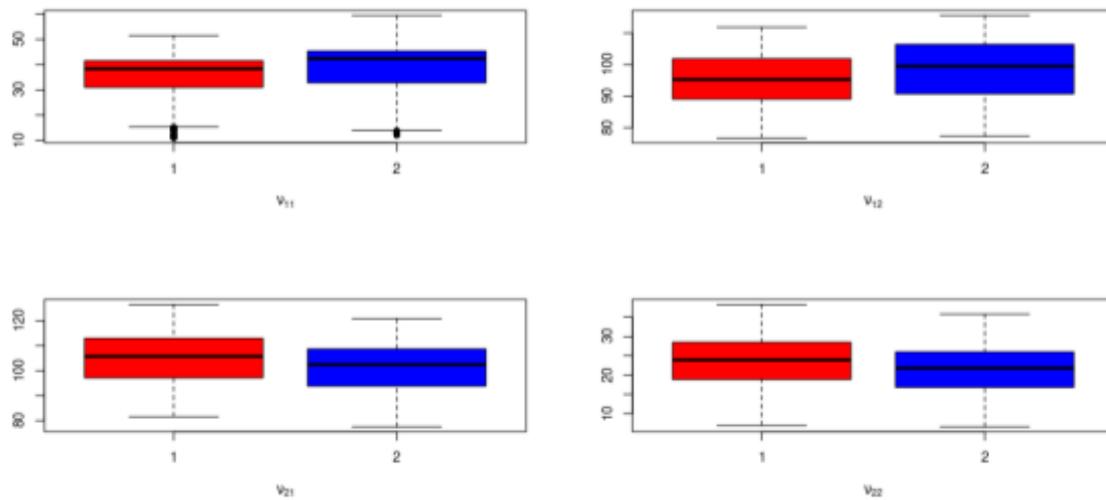
En la Tabla 31, se muestran las tasas de aceptación de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$ obtenidas del Algoritmo MALA dentro de un esquema de muestreo de Gibbs

Tabla 31. Tasa de aceptación de los hiperparámetros ψ (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).

v_{11}	v_{12}	v_{21}	v_{22}
91.36%	93.76%	93.48%	87.24%

En la Figura 68, se observan los diagramas de caja y bigotes de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$, obtenidos del Algoritmo MALA dentro de un esquema de muestreo de Gibbs.

Figura 68. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros ψ (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).



En la Tabla 32, se reportan las estimaciones de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$ del Algoritmo MALA dentro de un esquema de muestreo de Gibbs.

Tabla 32. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros ψ (Algoritmo MALA dentro de un esquema de muestreo de Gibbs).

Valor								
verdadero	$v_{11} = 25$	$v_{12} = 100$	$v_{21} = 100$	$v_{22} = 45$				
$Ga(v_{l,k}, v_{l,k})$								
$Ga(v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$	$v_{11}^{(\alpha)}$	$v_{11}^{(\beta)}$	$v_{12}^{(\alpha)}$	$v_{12}^{(\beta)}$	$v_{21}^{(\alpha)}$	$v_{21}^{(\beta)}$	$v_{22}^{(\alpha)}$	$v_{22}^{(\beta)}$
Media	36.01	39.30	95.07	98.27	104.64	101.10	23.56	21.21
Desviación estándar	8.84	10.06	8.28	9.21	10.53	10.51	6.67	6.13

5.1.4.4 Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs

En esta sección se resumen los resultados a posteriori del algoritmo MALA correlacionado ($\rho = 0.999$) dentro de un esquema de muestreo de Gibbs. El algoritmo MALA utilizó el parámetro $\epsilon = 0.5$. El esquema de muestreo de Gibbs empleó tres bloques de actualización: efectos aleatorios, parámetros comunes e hiperparámetros. Se utilizó un esquema de discretización Euler – Maruyama con un número de discretización $D = 20$ para simular las trayectorias de la EDE, y un filtro de partículas con $N = 20$ partículas para estimar la verosimilitud de $M = 5$ unidades experimentales. El algoritmo fue ejecutado con $I = 5000$ iteraciones, donde las primeras $B = 2500$ iteraciones son del periodo de quemado. Después de ejecutar el algoritmo se obtuvo una tasa de aceptación de los parámetros comunes de 16.52% y un tiempo de ejecución de 60.44 minutos.

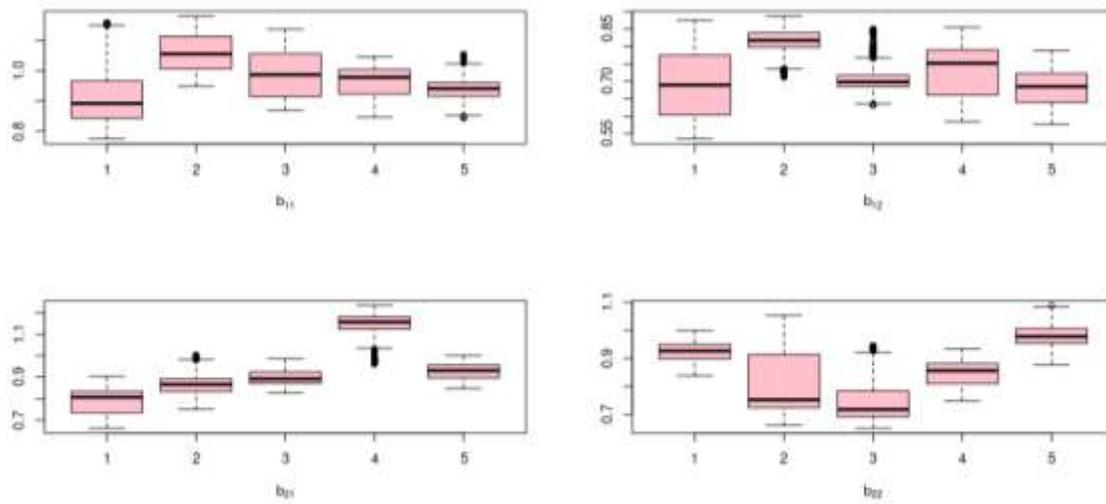
En la Tabla 33, se muestran las tasas de aceptación de los efectos aleatorios $b^{(m)}$ para $m = 1, \dots, 5$ obtenidas del Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs.

Tabla 33. Tasa de aceptación de los efectos aleatorios $b^{(m)}$ (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).

$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
56.32%	60.20%	58.04%	57.96%	59.60%

En la Figura 69, se muestran los diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo $(b_{11}^{(m)}, b_{12}^{(m)}, b_{21}^{(m)}, b_{22}^{(m)})$ para $m = 1, \dots, 5$ obtenidas del Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs.

Figura 69. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios $b^{(m)}$ (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).



En la Figura 70 y Figura 71, se observan los histogramas, diagrama de traza y diagramas de autocorrelación de los parámetros α_2 y ξ_2 , respectivamente. La línea negra indica la media a posteriori de las muestras obtenidas del Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs y la línea roja indica el valor verdadero del parámetro.

Figura 70. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro α_2 (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).

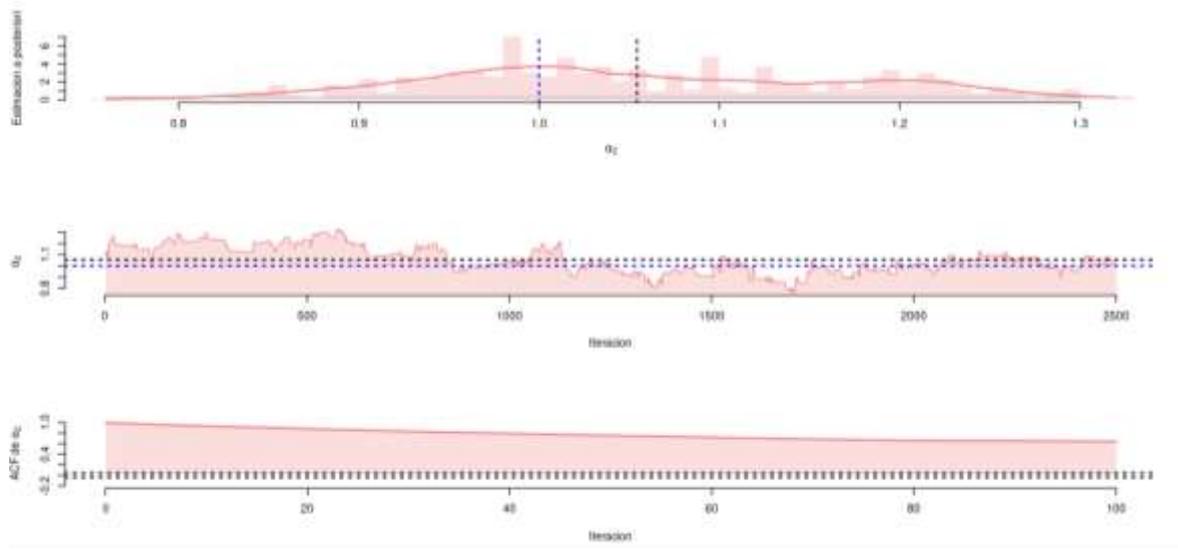
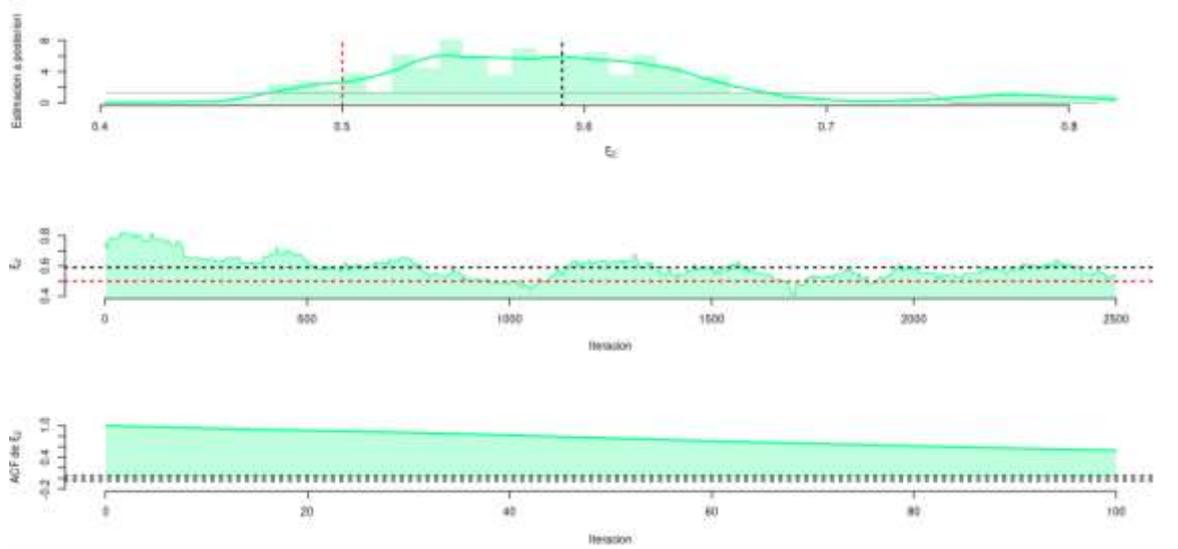


Figura 71. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro ξ_2 (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).



En la Tabla 34, se muestran las estimaciones de los parámetros comunes θ obtenidas del Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs.

Tabla 34. Resumen estadístico de los parámetros comunes θ (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).

	α_1	α_2	β_{11}	β_{12}	β_{21}	β_{22}	σ_1	σ_2	ξ_1	ξ_2
Valor verdadero	3	2.5	1.8	2	1	1.5	0.3	0.5	0.1	0.5
Media	1.45	1.23	1.21	0.85	1.38	1.05	0.06	0.12	0.08	0.59
Desviación estándar	0.16	0.17	0.10	0.18	0.07	0.12	0.04	0.09	0.02	0.08
IACT	151.38	161.60	112.40	163.03	128.85	157.05	102.51	169.10	96.31	155.24

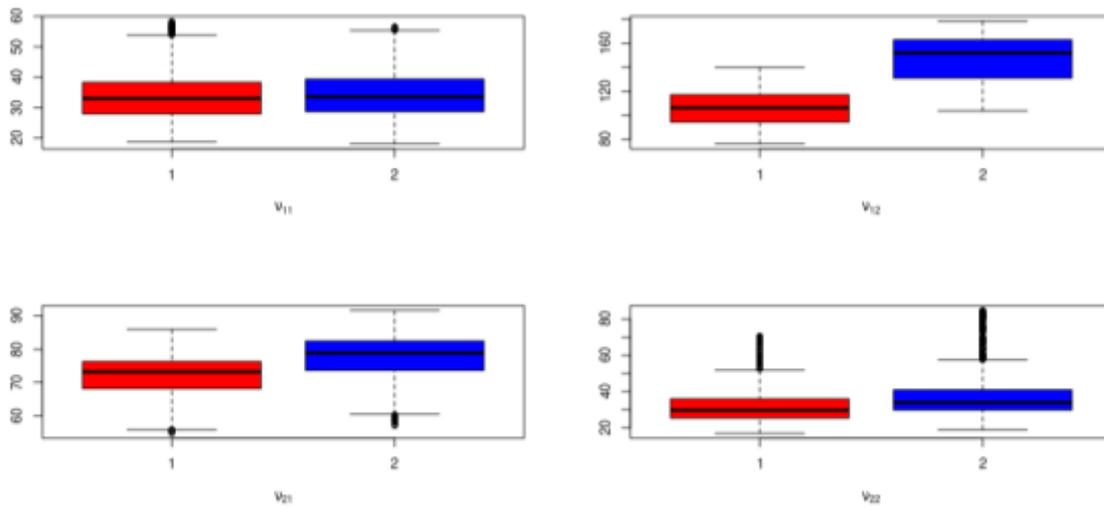
En la Tabla 35, se muestran las tasas de aceptación de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$ obtenidas del Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs

Tabla 35. Tasa de aceptación de los hiperparámetros ψ (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).

v_{11}	v_{12}	v_{21}	v_{22}
89.96%	94.20%	91.72%	87.08%

En la Figura 72, se observan los diagramas de caja y bigotes de los hiperparámetros $v_{l,k}^{(\alpha)}$ y $v_{l,k}^{(\beta)}$ para $l, k = 1, 2$, obtenidos del Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs

Figura 72. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros ψ (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).



En la Tabla 36, se reportan las estimaciones de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$ del Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs.

Tabla 36. Resumen estadístico de los hiperparámetros ψ (Algoritmo MALA correlacionado dentro de un esquema de muestreo de Gibbs).

Valor								
verdadero	$v_{11} = 25$	$v_{12} = 100$	$v_{21} = 100$	$v_{22} = 45$				
$Ga(v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$								
$Ga(v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$	$v_{11}^{(\alpha)}$	$v_{11}^{(\beta)}$	$v_{12}^{(\alpha)}$	$v_{12}^{(\beta)}$	$v_{21}^{(\alpha)}$	$v_{21}^{(\beta)}$	$v_{22}^{(\alpha)}$	$v_{22}^{(\beta)}$
Media	33.65	34.24	106.53	147.13	71.79	77.58	32.15	37.46
Desviación estándar	7.99	7.78	15.33	19.56	6.38	6.66	10.36	12.47

5.1.4.5 Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs

En esta sección se resumen los resultados a posteriori del algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs. El algoritmo Hamiltoniano utilizó un tamaño de paso $\epsilon = 1/10$ y un número de paso $L = 10$. El esquema de muestreo de Gibbs empleó tres bloques de actualización: efectos aleatorios, parámetros comunes e hiperparámetros. Se utilizó un esquema de discretización Euler – Maruyama con un número de discretización $D = 20$ para simular las trayectorias de la EDE, y un filtro de partículas con $N = 20$ partículas para estimar la verosimilitud de $M = 5$ unidades experimentales. El algoritmo fue ejecutado con $I = 5000$ iteraciones, donde las primeras $B = 2500$ iteraciones son del periodo de quemado. Después de ejecutar el algoritmo se obtuvo una tasa de aceptación de los parámetros comunes de 14.72% y un tiempo de ejecución de 68.87 minutos.

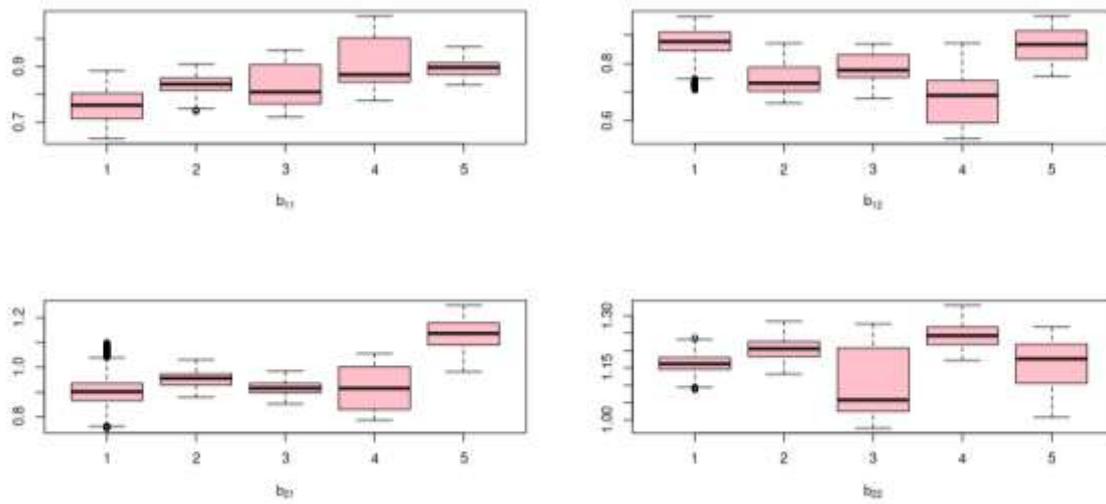
En la Tabla 37, se muestran las tasas de aceptación de los efectos aleatorios $b^{(m)}$ para $m = 1, \dots, 5$ obtenidas del Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs.

Tabla 37. Tasa de aceptación de los efectos aleatorios $b^{(m)}$ (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).

$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
56.32%	60.20%	58.04%	57.96%	59.60%

En la Figura 73, se muestran los diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo $(b_{11}^{(m)}, b_{12}^{(m)}, b_{21}^{(m)}, b_{22}^{(m)})$ para $m = 1, \dots, 5$ obtenidas del Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs.

Figura 73. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios $b^{(m)}$ (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).



En la Figura 74 y Figura 75, se observan los histogramas, diagrama de traza y diagramas de autocorrelación de los parámetros α_2 y ξ_2 , respectivamente. La línea negra indica la media a posteriori de las muestras obtenidas del Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs y la línea roja indica el valor verdadero del parámetro.

Figura 74. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro α_2 (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).

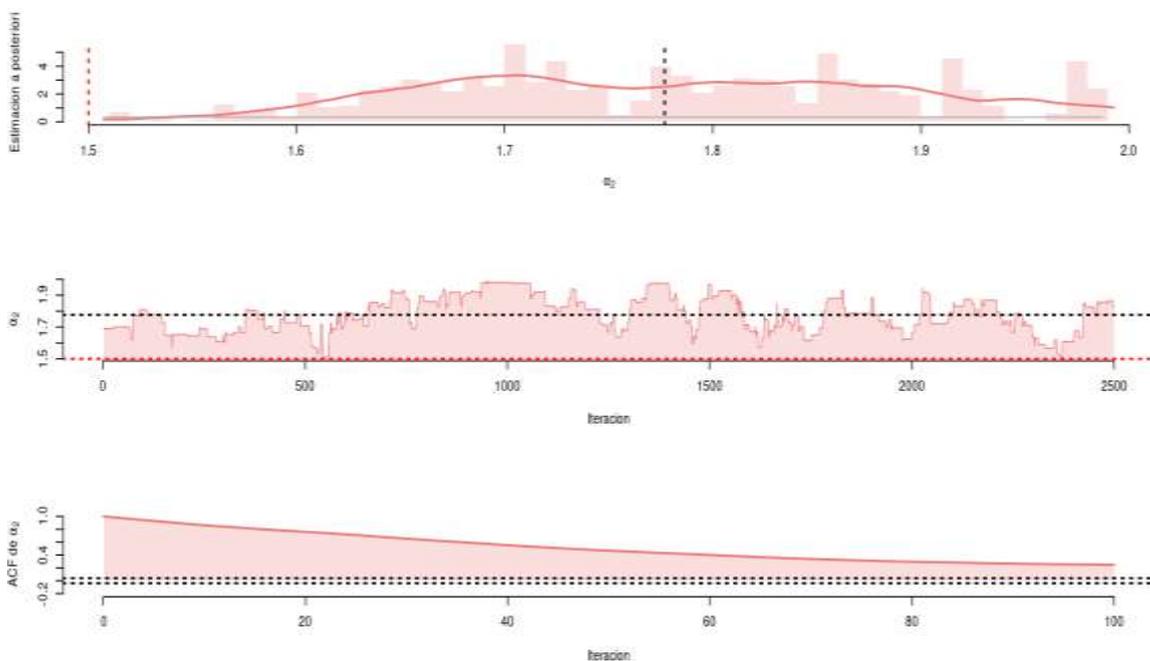
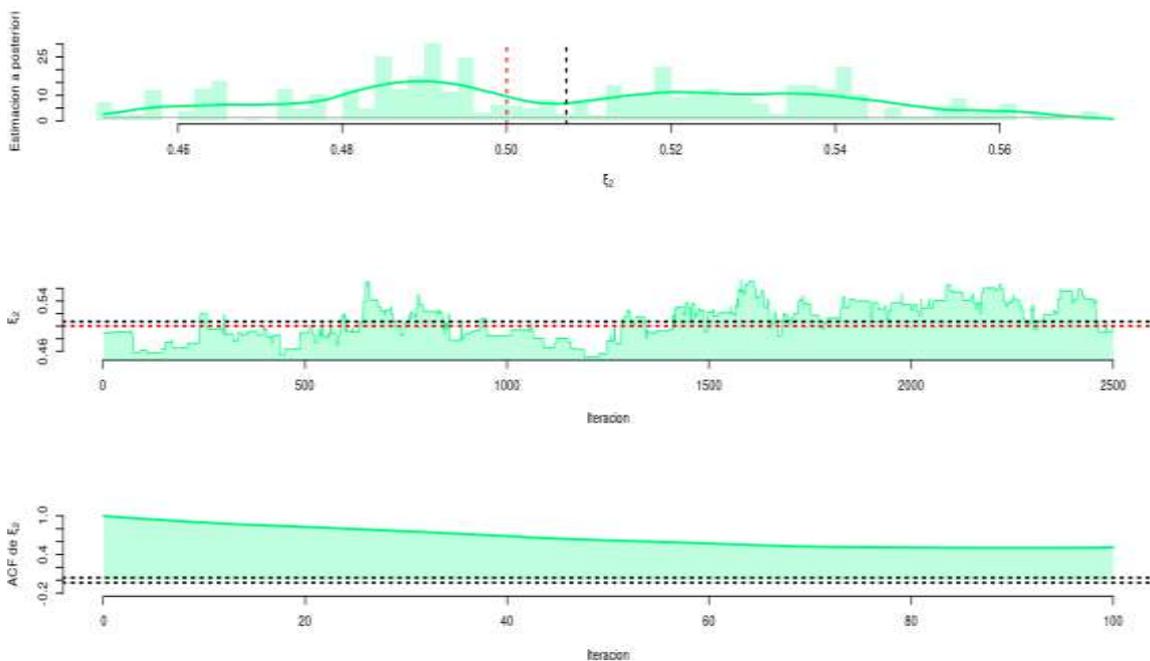


Figura 75. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro ξ_2 (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).



En la Tabla 38, se muestran las estimaciones de los parámetros comunes θ obtenidas del Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs.

Tabla 38. Resumen estadístico de los parámetros comunes θ (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).

	α_1	α_2	β_{11}	β_{12}	β_{21}	β_{22}	σ_1	σ_2	ξ_1	ξ_2
Valor verdadero	3	2.5	1.8	2	1	1.5	0.3	0.5	0.1	0.5
Media	3.55	1.96	3.18	1.62	0.96	1.78	0.13	0.27	0.08	0.51
Desviación estándar	0.25	0.26	0.27	0.12	0.06	0.11	0.04	0.02	0.02	0.03
IACT	186.75	155.90	160.20	137.50	115.52	106.45	179.82	153.06	148.52	135.25

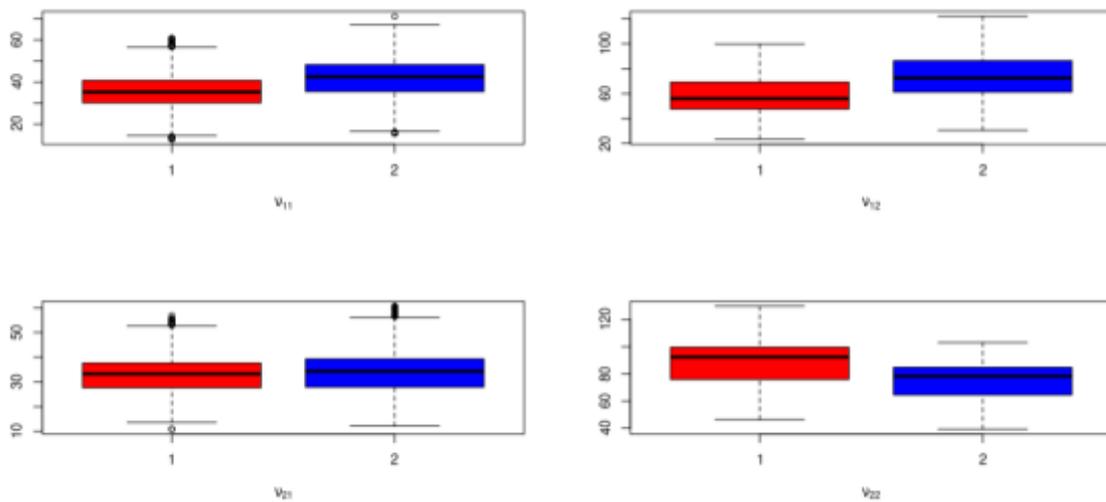
En la Tabla 39, se muestran las tasas de aceptación de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$ obtenidas del Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs

Tabla 39. Tasa de aceptación de los hiperparámetros ψ (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).

v_{11}	v_{12}	v_{21}	v_{22}
73.44%	77.88%	68.88%	78.60%

En la Figura 76, se observan los diagramas de caja y bigotes de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$, obtenidos del Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs

Figura 76. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros ψ (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).



En la Tabla 40, se reportan las estimaciones de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$ del Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs.

Tabla 40. Resumen estadístico de los hiperparámetros ψ (Algoritmo Hamiltoniano dentro de un esquema de muestreo de Gibbs).

Valor								
verdadero	$v_{11} = 25$	$v_{12} = 100$	$v_{21} = 100$	$v_{22} = 45$				
$Ga(v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$								
$Ga(v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$	$v_{11}^{(\alpha)}$	$v_{11}^{(\beta)}$	$v_{12}^{(\alpha)}$	$v_{12}^{(\beta)}$	$v_{21}^{(\alpha)}$	$v_{21}^{(\beta)}$	$v_{22}^{(\alpha)}$	$v_{22}^{(\beta)}$

Media	35.50	41.79	58.89	74.74	32.79	33.84	88.01	74.78
Desviación estándar	8.80	9.76	16.05	19.20	8.25	9.01	16.85	13.95

5.1.4.6 Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs

En esta sección se resumen los resultados a posteriori del algoritmo Hamiltoniano correlacionado ($\rho = 0.999$) dentro de un esquema de muestreo de Gibbs. El algoritmo Hamiltoniano utilizó un tamaño de paso $\epsilon = 1/10$ y un número de paso $L = 10$. El esquema de muestreo de Gibbs empleó tres bloques de actualización: efectos aleatorios, parámetros comunes e hiperparámetros. Se utilizó un esquema de discretización Euler – Maruyama con un número de discretización $D = 20$ para simular las trayectorias de la EDE, y un filtro de partículas con $N = 20$ partículas para estimar la verosimilitud de $M = 5$ unidades experimentales. El algoritmo fue ejecutado con $I = 5000$ iteraciones, donde las primeras $B = 2500$ iteraciones son del periodo de quemado. Después de ejecutar el algoritmo se obtuvo una tasa de aceptación de los parámetros comunes de 25.56% y un tiempo de ejecución de 65.57 minutos.

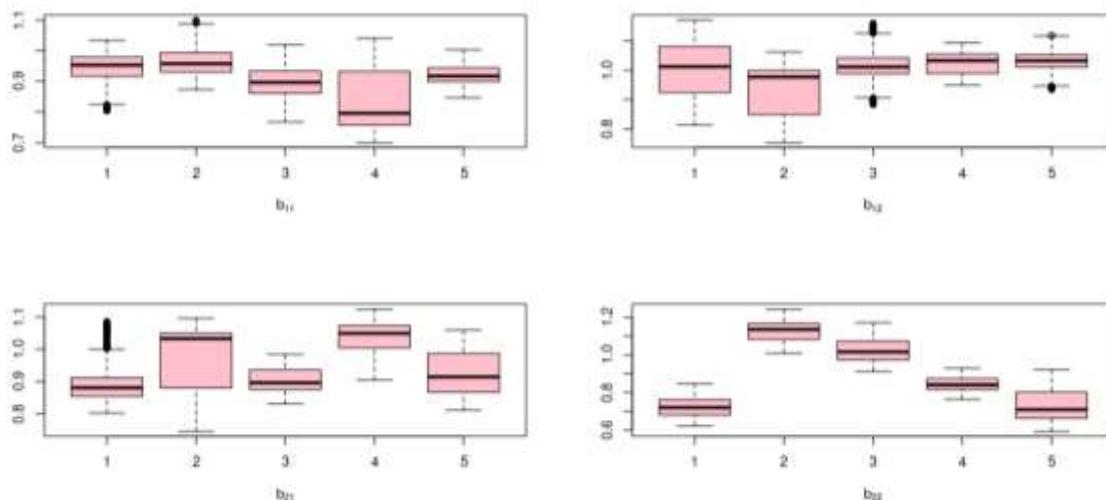
En la Tabla 41, se muestran las tasas de aceptación de los efectos aleatorios $b^{(m)}$ para $m = 1, \dots, 5$ obtenidas del Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs.

Tabla 41. Tasa de aceptación de los efectos aleatorios $b^{(m)}$ (Algoritmo Hamiltoniano correlacionado dentro de un esquema de Muestreo de Gibbs).

$m = 1$	$m = 2$	$m = 3$	$m = 4$	$m = 5$
75.96%	81.12%	80.20%	70.76%	76.68%

En la Figura 77, se muestran los diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios del modelo $(b_{11}^{(m)}, b_{12}^{(m)}, b_{21}^{(m)}, b_{22}^{(m)})$ para $m = 1, \dots, 5$ obtenidas del Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs.

Figura 77. Diagramas de caja y bigotes de las muestras a posteriori de los efectos aleatorios $b^{(m)}$ (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).



En la Figura 78 y Figura 79, se observan los histogramas, diagrama de traza y diagramas de autocorrelación de los parámetros α_2 y ξ_2 , respectivamente. La línea negra indica la media a posteriori de las muestras obtenidas del Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs y la línea roja indica el valor verdadero del parámetro.

Figura 78. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro α_2 (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).

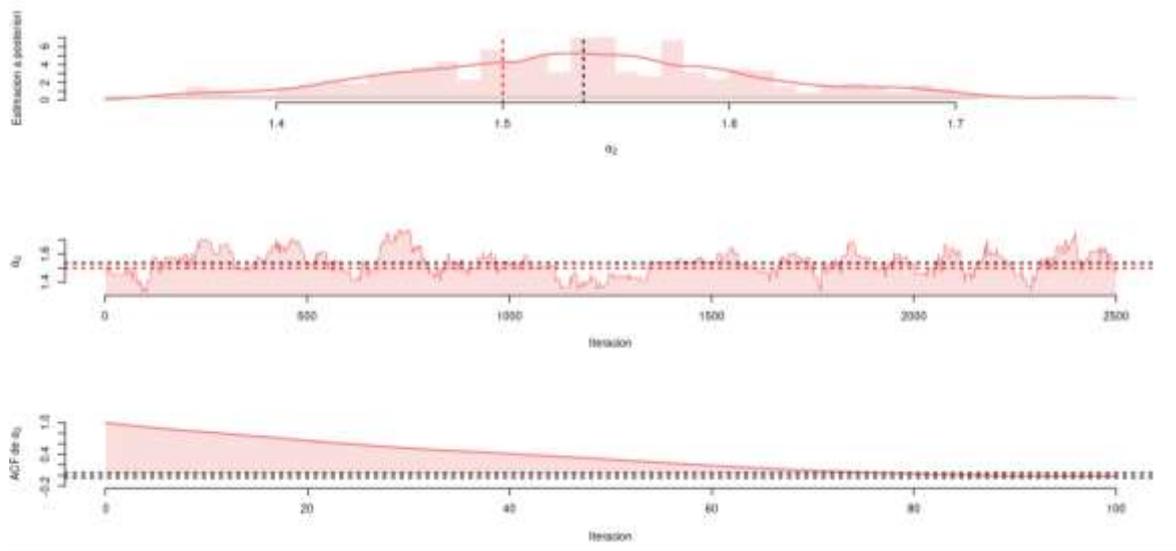
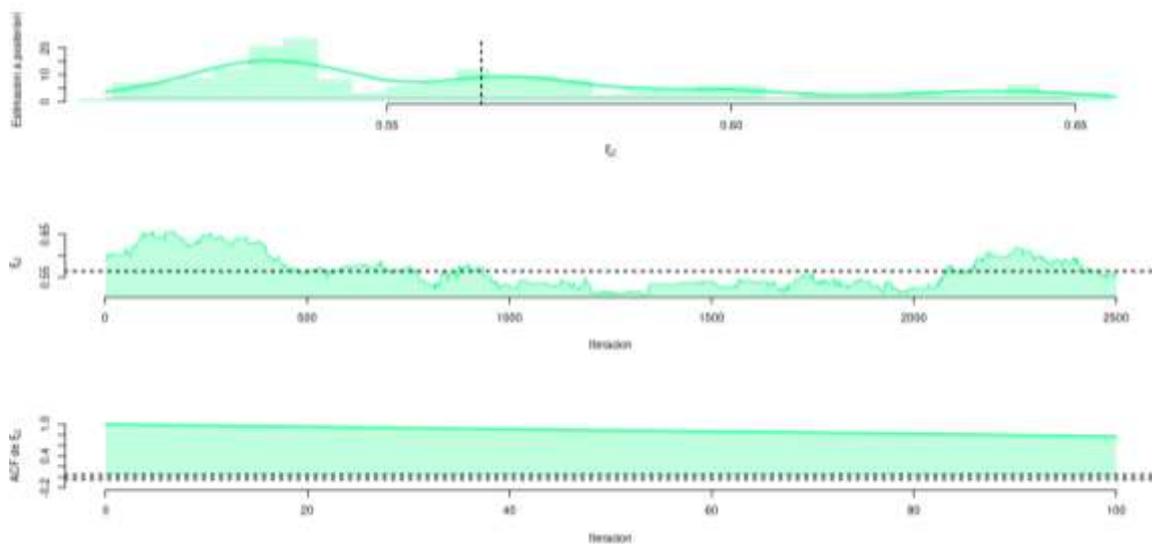


Figura 79. Histograma, diagrama de traza y diagrama de autocorrelación del parámetro ξ_2 (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).



En la Tabla 42, se muestran las estimaciones de los parámetros comunes θ obtenidas del Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs.

Tabla 42. Resumen estadístico de los parámetros comunes θ (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).

	α_1	α_2	β_{11}	β_{12}	β_{21}	β_{22}	σ_1	σ_2	ξ_1	ξ_2
Valor verdadero	3	2.5	1.8	2	1	1.5	0.3	0.5	0.1	0.5
Media	1.81	1.13	1.27	1.14	1.09	1.54	0.16	0.10	0.05	0.56
Desviación estándar	0.20	0.19	0.25	0.16	0.08	0.08	0.03	0.05	0.02	0.04
IACT	155.88	152.62	183.90	125.70	111.65	71.42	168.78	182.71	137.33	180.15

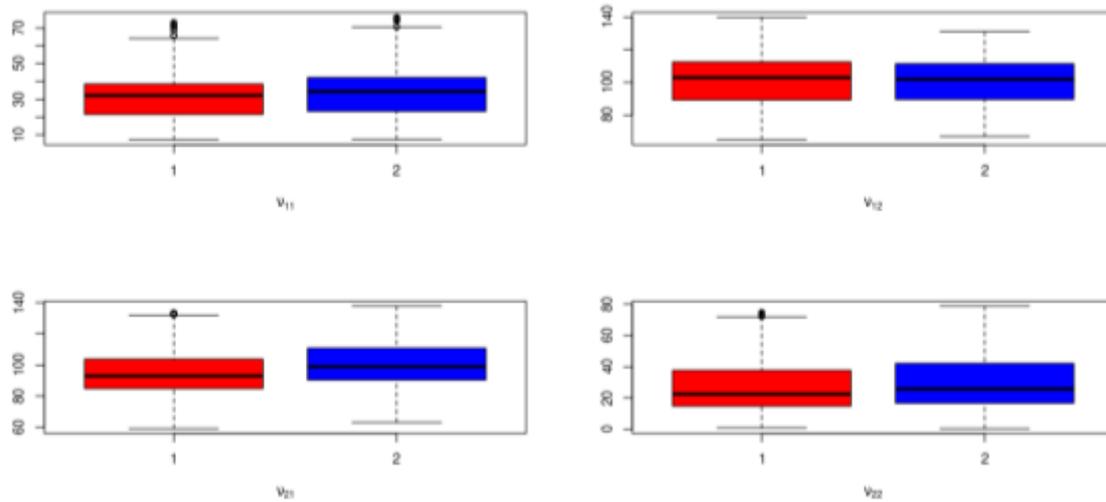
En la Tabla 43, se muestran las tasas de aceptación de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$ obtenidas del Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs.

Tabla 43. Tasa de aceptación de los hiperparámetros ψ (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).

v_{11}	v_{12}	v_{21}	v_{22}
69.52%	80.44%	82.72%	66.92%

En la Figura 80, se observan los diagramas de caja y bigotes de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$, obtenidos del Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs

Figura 80. Diagramas de caja y bigotes de las muestras a posteriori de los hiperparámetros ψ (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).



En la Tabla 44, se reportan las estimaciones de los hiperparámetros $\psi = (v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$ para $l, k = 1, 2$ del Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs.

Tabla 44. Resumen estadístico de los hiperparámetros ψ (Algoritmo Hamiltoniano correlacionado dentro de un esquema de muestreo de Gibbs).

Valor								
verdadero	$v_{11} = 25$	$v_{12} = 100$	$v_{21} = 100$	$v_{22} = 45$				
$Ga(v_{l,k}, v_{l,k})$								
$Ga(v_{l,k}^{(\alpha)}, v_{l,k}^{(\beta)})$	$v_{11}^{(\alpha)}$	$v_{11}^{(\beta)}$	$v_{12}^{(\alpha)}$	$v_{12}^{(\beta)}$	$v_{21}^{(\alpha)}$	$v_{21}^{(\beta)}$	$v_{22}^{(\alpha)}$	$v_{22}^{(\beta)}$
Media	31.37	33.88	101.38	100.60	94.99	100.99	27.00	30.01
Desviación estándar	12.64	13.26	15.39	14.15	13.71	14.27	16.04	17.14

En la Tabla 45, se muestra el desempeño de los algoritmos Metropolis – Hastings, MALA, Hamiltoniano y sus respectivas versiones correlacionadas. Los indicadores de desempeño son tiempo de ejecución (segundos), tamaño de muestra efectivo (ESS: effective sample size) promedio, tasa de tamaño de muestra efectivo (ESS/s), tasa de aceptación de los parámetros y promedio de IACT de los parámetros comunes. Se observa que los métodos pseudo – marginales correlacionados presentan una mayor tasa de tamaño de muestra efectivo, lo que indica que las muestras a posteriori presentan una baja correlación y un tiempo de ejecución menor con respecto a los métodos pseudo – marginales no correlacionados. Además, se tiene que el método Hamiltoniano correlacionado presentan la mayor tasa de aceptación de los parámetros comunes θ . Finalmente, se muestra que los métodos Hamiltoniano (correlacionado y no correlacionado) presentan un bajo promedio de IACT de los parámetros θ , indicando que estos métodos producen muestras a posteriori de los parámetros θ con baja correlación.

Tabla 45. Indicadores de desempeño de los algoritmos Metropolis – Hastings dentro de un esquema de muestreo de Gibbs.

	Tamaño de muestra efectivo promedio (ESS)	Tasa de tamaño de muestra efectivo (ESS/segundo)	Tasa de aceptación de parámetros θ (%)	Promedio IACT θ
--	--	---	---	--

MH – Gibbs	4191.00	5.10	0.0012	2.40	165.72
MH – Gibbs correlacionado	3901.66	4.92	0.0013	18.29	169.72
MALA – Gibbs	4271.93	5.58	0.0013	0.40	167.36
MALA – Gibbs correlacionado	3806.18	14.18	0.0037	16.52	139.75
Hamiltoniano – Gibbs	4131.88	10.70	0.0026	14.73	147.90
Hamitoniano – Gibbs correlacionado	3934.39	10.19	0.0026	25.57	147.01

6 CONCLUSIONES

En esta tesis se usaron métodos pseudo – marginales para estimar los parámetros (mediante el algoritmo Metropolis – Hastings) y estados (mediante el filtro de partículas) desconocidos de una EDE que modelaba el crecimiento del árbol de naranja (Caso unidimensional con datos reales). En este primer caso, se usaron los siguientes métodos pseudo – marginales: Aumento Individual (AI), MALA y Hamiltoniano, cada uno de estos métodos fue implementado con un esquema de discretización para simular las trayectorias de la EDE, estos esquemas fueron el método Euler – Maruyama (EM), Puente de difusión modificado (PDM) y Puente de difusión residual (PDR). Se observó que los métodos MALA y Hamiltoniano con los esquemas de discretización EM y PDM tuvieron un buen desempeño en la estimación de los parámetros y estados desconocidos, ya que alcanzaron las mayores tasas de aceptación de los efectos comunes $b^{(m)}$, parámetros comunes θ e hiperparámetros ψ en comparación a los demás métodos pseudo – marginales. Adicionalmente, las muestras a posteriori del método Hamiltoniano con el esquema de Puente de difusión modificado presentaron menor correlación a comparación de los demás métodos pseudo – marginales con los diferentes esquemas de discretización.

También se usaron métodos pseudo – marginales (Metropolis – Hastings, MALA, Hamiltoniano) correlacionados y no correlacionados con el esquema de discretización Euler – Maruyama para estimar los parámetros desconocidos de un proceso bidimensional Ornstein – Uhlenbeck. Se observó que los métodos correlacionados tuvieron un menor tiempo de ejecución a comparación de los métodos no correlacionados. También se obtuvo que las muestras a posteriori de los métodos MALA y Hamiltoniano presentaron una menor correlación con respecto al método Metropolis – Hastings.

Referencias Bibliográficas

Barragán, G., Infante, S., y Hernández, A. (2020). *Unscented Kalman Filter and Gauss-Hermite Kalman Filter for Range-Bearing Target Tracking*. 537-544.

Botha, I., Kohn, R., y Drovandi, C. (2021). Particle Methods for Stochastic Differential Equation Mixed Effects Models. *Bayesian Analysis*, 16(2). <https://doi.org/10.1214/20-BA1216>

Dahlin, J., y Schön, T. B. (2019). Getting Started with Particle Metropolis-Hastings for Inference in Nonlinear Dynamical Models. *Journal of Statistical Software*, 88(Code Snippet 2). <https://doi.org/10.18637/jss.v088.c02>

Delattre, M. (2021). A review on asymptotic inference in stochastic differential equations with mixed effects. *Japanese Journal of Statistics and Data Science*, 4(1), 543-575. <https://doi.org/10.1007/s42081-021-00105-3>

Dion, C. (2016). Nonparametric estimation in a mixed-effect Ornstein–Uhlenbeck model. *Metrika*, 79(8), 919-951. <https://doi.org/10.1007/s00184-016-0583-y>

Doucet, A., y Johansen, A. M. (2009). A tutorial on particle filtering and smoothing: Fifteen years later. *Handbook of Nonlinear Filtering*, 12(3), 656-704.

Fadwa, B., El Maroufy, H., y Ait Mousse, H. (2020). Simulation and Parametric Inference of a Mixed Effects Model with Stochastic Differential Equations Using the Fokker-Planck Equation Solution. En D. M. Cvetković y G. A. Birajdar (Eds.), *Numerical Modeling and Computer Simulation*. IntechOpen. <https://doi.org/10.5772/intechopen.90751>

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., y Rubin, D. B. (2014). *Bayesian Data Analysis* (3.^a ed.). Chapman and Hall/CRC.

Girolami, M., y Calderhead, B. (2011). Riemann manifold Langevin and Hamiltonian Monte Carlo methods. *Journal of the Royal Statistical Society*, 73(2), 123-214.

Luengo, D., Martino, L., Bugallo, M., Elvira, V., y Särkkä, S. (2020). A survey of Monte Carlo methods for parameter estimation. *EURASIP Journal on Advances in Signal Processing*, 2020(1), 25. <https://doi.org/10.1186/s13634-020-00675-6>

Øksendal, B. (2000). *Stochastic Differential Equations: An Introduction with Applications* (5.^a ed.). Springer.

Picchini, U., y Forman, J. L. (2019). Bayesian inference for stochastic differential equation mixed effects models of a tumour xenography study. *Journal of the Royal Statistical Society: Series C (Applied Statistics)*, 68(4), 887-913. <https://doi.org/10.1111/rssc.12347>

Rampinelli, C. (2019). *Metropolis algorithm applied to a normal distribution assuming unknown mean and variance* [R]. <https://rpubs.com/cassiorampinelli/431636>

Särkkä, S. (2013). *Bayesian Filtering and Smoothing* (1.^a ed.). Cambridge University Press. <https://doi.org/10.1017/CBO9781139344203>

Schön, T. B., y Lindsten, F. (2017). *Learning of dynamical systems: Particle filters and Markov chain methods* (Draft).

Soto, J., y Infante, S. (2019). Ensemble Kalman filter and extended Kalman filter for state-parameter dual estimation in mixed effects models defined by a stochastic differential equation. *Technology, Sustainability and Educational Innovation*, 285-300.

Soto, J., Infante, S., Camacho, F., y Amaro, I. R. (2019). Estimación de un modelo de efectos mixtos usando un proceso de difusión parcialmente observado. *Revista de Matemática: Teoría y aplicaciones*, 26(1), 83-98.

Whitaker, G. A., Golightly, A., Boys, R. J., y Sherlock, C. (2017). Bayesian Inference for Diffusion-Driven Mixed-Effects Models. *Bayesian Analysis*, 12(2). <https://doi.org/10.1214/16-BA1009>

Wiqvist, S., Golightly, A., McLean, A. T., y Picchini, U. (2021). Efficient inference for stochastic differential equation mixed-effects models using correlated particle pseudo-marginal algorithms. *Computational Statistics & Data Analysis*, 157, 107151. <https://doi.org/10.1016/j.csda.2020.107151>