
PROCESOS DE OPTIMIZACIÓN EN LA DISTRIBUCIÓN



TESIS DOCTORAL

África Ruiz Gándara

Dpto. Estadística, Econometría, I.O., Org. de Empresas y Ec. Aplicada
Universidad de Córdoba

Córdoba 2013

TITULO: *PROCESOS DE OPTIMIZACIÓN EN LA DISTRIBUCIÓN*

AUTOR: *ÁFRICA RUIZ GÁNDARA*

© Edita: Servicio de Publicaciones de la Universidad de Córdoba. 2013
Campus de Rabanales
Ctra. Nacional IV, Km. 396 A
14071 Córdoba

www.uco.es/publicaciones
publicaciones@uco.es

PROCESOS DE OPTIMIZACIÓN EN LA DISTRIBUCIÓN

Memoria que presenta para optar al título de Doctor

África Ruiz Gándara

Dirigida por

Dr. José M^a Caridad y Ocerin (Univ. de Córdoba)

Dr. Jose Luis Pino Mejías (Univ. de Sevilla)

**Dpto. Estadística, Econometría, I.O., Org. de Empresas y Ec. Aplicada
Universidad de Córdoba**

Córdoba 2013



TÍTULO DE LA TESIS:

DOCTORANDO/A: África Ruíz Gándara

INFORME RAZONADO DEL/DE LOS DIRECTOR/ES DE LA TESIS

(se hará mención a la evolución y desarrollo de la tesis, así como a trabajos y publicaciones derivados de la misma).

El trabajo de tesis que presenta la becaria de FPD I A. Ruíz Gándara es el resultado de varios años de desarrollo en los que ha combinado esta investigación con la docencia, y que ahora culmina con esta memoria.

Se abordan dos cuestiones: una de nuevos métodos teóricos de selección de modelos ARIMA, utilizando unos criterios predictivos, como alternativa a los procedimientos basados en medidas de ajuste obtenidas sobre el período muestral; la segunda consiste en la aplicación práctica de las nuevas metodologías a la resolución de problemas reales de predicción asociados a la estimación del consumo de agua en zonas urbanas, consiguiéndose mejoras respecto a los métodos habituales que aparecen en la bibliografía.

El último objetivo ha originado un artículo recientemente aceptado tras el proceso de evaluación en SMSIS (2013) titulado *Automatic identification of Arima models: the case of water consumption forecasting*, y que aparecerá publicado el próximo mes de septiembre, comparándose los resultados con datos empresariales reales usando las metodologías existentes y la basada en los procesos propuestos de especificación.

Otro artículo es el sometido a Investigación Económica (2013), revista JCR, *Modelos de predicción para el consumo de agua en áreas urbanas* que ha pasado el primer filtro de dicha publicación, y está en la fase final de evaluación.

Otras publicaciones de la doctoranda han sido en los siguientes congresos

- XXIX Congreso Nacional de Estadística e IO
- XXX Congreso Nacional de Estadística e IO
- XXXI Congreso Nacional de Estadística e IO
- XXXII Congreso Nacional de Estadística e IO
- XXXIII Congreso Nacional de Estadística e IO

así como la participación en un libro de Estadística en la Universidad de Sevilla.

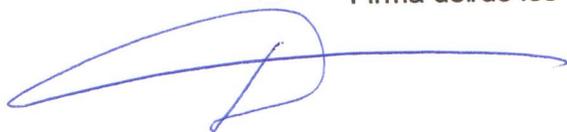
En la actualidad está a punto de someterse a evaluación un nuevo artículo, *Nuevos métodos de especificación de modelos de consumo de agua* con los resultados teóricos de técnicas de simulación en la selección de modelos.

También fueron presentados los resultados en el III Congreso Científico de Investigadores en Formación de la Universidad de Córdoba (2013).

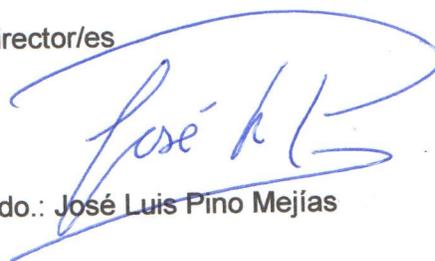
Por todo ello, se autoriza la presentación de la tesis doctoral.

Córdoba, 26 de Junio de 2013

Firma del/de los director/es



Fdo.: José Mª Caridad y Ocerín



Fdo.: José Luis Pino Mejías

A José M^a Caridad y Jose Luis Pino

que han hecho posible esta Tesis

*A todos los que directa o indirectamente
me habéis ayudado, apoyado y animado
en este duro camino.*

Agradecimientos

A todos los que la presente vieron y entendieron.

Inicio de las Leyes Orgánicas. Juan
Carlos I

La presente Tesis es un esfuerzo en el cual, directa o indirectamente, participaron varias personas, leyendo, opinando, corrigiendo, teniéndome paciencia, dándome ánimo, acompañándome en los momentos de crisis y en los momentos de felicidad. Me gustaría que estas líneas sirvieran para expresar mi más profundo y sincero agradecimiento a todas estas personas que con su ayuda han colaborado en la realización de este duro trabajo.

En especial al Dr. Jose M^a Caridad y Ocerín director de esta investigación, por la orientación, el seguimiento y la supervisión de la misma, pero sobre todo por haber confiado en mi persona, por su tenacidad, su capacidad de superación y su infinito valor humano. Al Dr. Jose Luis Pino Mejías, codirector de esta investigación, por el interés mostrado, por la motivación, el ánimo y el apoyo recibido a lo largo de estos años, por su generosidad y paciencia. A ambos por su brillante inteligencia.

Tengo que nombrar también de una manera muy especial al Dr. Rafael Pino Mejías para el que no tengo palabras de agradecimiento y con el que siempre me sentiré en deuda.

Agradezco a mi compañero el Dr. Jose Angel Roldán por suministrarme material importante e interesante para mi investigación y por sus tan necesarias explicaciones, por dedicarme su tiempo y compartir sus conocimientos conmigo, a mi también compañero Dr. Jose Luis Cabello por la ayuda brindada sobre todo en el plano informático, y a mis compañeras Dra. M^a Isabel Sánchez, y Dra. M^a Isabel Tocón, con las que me encuentro en deuda por el ánimo infundido, por los consejos, por su paciencia, su confianza y su siempre compañía.

En especial al Dr. Manuel Vaquero, gran compañero y mejor persona, por su confianza, por su fe en mí, su constante apoyo y sus continuas palabras de ánimo, admiración y cariño, y por todo lo que he aprendido a su lado, y lo que hemos compartido.

Quisiera hacer extensiva mi gratitud a mis compañeros del Departamento tanto de Sevilla como de Córdoba que me han permitido entrar en sus vidas durante estos años, en mayor o menor medida.

Gracias también a la Junta de Andalucía, por concederme esta beca que ha hecho posible la realización de este sueño.

A mis queridos amigos, a todos y cada uno de ellos, que siempre han estado a mi lado, confiando y apostando por mí, por su generosidad, su paciencia, el empeño e interés mostrado, por concederme su tiempo y cariño, porque sin tenerlos a mi lado todo hubiera sido más complicado.

A mis padres, África y Ángel, por darme todo sin esperar nada a cambio, por disfrutar y sufrir conmigo, por enseñarme a vivir, y a mi hermano, Ángel, porque siempre ha estado ahí, a ellos tres en especial, les quiero agradecer su compañía en esta aventura, de forma incondicional, entendiendo mis ausencias y mis malos momentos, a ellos que a pesar de la distancia siempre han estado atentos a mis progresos, a ellos que a pesar, de todo, siempre han estado dispuestos a sacrificarse por hacer que yo consiguiera esto. A ellos que nunca les ha faltado una palabra de ánimo, de preocupación, de apoyo, de comprensión, de valentía, de lucha... Por ellos y para ellos.

Y para terminar, a John, que ha estado a mi lado acompañándome a lo largo de esta intensa aventura, en silencio, con respeto, amor y admiración.

A todos, gracias, muchas gracias, sin vuestra aportación este trabajo no hubiera salido adelante.

Resumen

*A veces gris, a veces blanco,
todo depende del lugar.*

RM.

Los procesos donde interviene la aleatoriedad, son y han sido motivo de estudio debido a su frecuente presencia. En muchos de los casos que se pueden citar, esta aleatoriedad se manifiesta en la demanda de un cierto producto: por ejemplo, las compañías tienen que decidir cuánta ropa deben fabricar, cuántos coches, etc. Si la producción supera a la demanda, se generan pérdidas por almacenaje, materias primas utilizadas, mano de obra, . . . Si la producción es bastante menor que la demanda, se crean pérdidas por venta no realizada. Por tanto, en estos procesos, una buena predicción de la demanda es fundamental para optimizar el beneficio de las empresas. La producción y distribución de un bien perecedero cuya demanda es aleatoria, es lo que se va a estudiar en este trabajo. Por tanto, pueden aplicarse a diversos tipos de procesos, resultando útiles para la toma de las decisiones y medidas pertinentes (fabricación de los bienes, venta, etc.). Asimismo, estos sistemas también pueden usarse para prevención, desde el momento que lo que se estudia es el comportamiento en el futuro.

La aplicación que se llevará a cabo en este trabajo, será sobre la estimación de la demanda de agua. El agua es un recurso natural incomparable, escaso y estratégico, elemento clave para la vida y el desarrollo social, difícilmente se puede exagerar la importancia del agua, es seguramente el más valioso e imprescindible de los existentes en el planeta Tierra. El agua debería ser un derecho humano y, sin embargo, 1100 millones de personas carecen de acceso al agua potable.

La aparente abundancia del agua en el mundo ha dado la impresión, en el pasado, de que se trataba de un bien inagotable. Era también el más barato. En muchos lugares el agua era gratuita. Una política subsidiada que ha fomentado su derroche, su explotación excesiva, originando problemas de escasez, pérdida de calidad o deterioro medioambiental.

En la medida en que el agua se va convirtiendo en un recurso cada vez más escaso y deteriorado por la intervención humana, los poderes públicos han ido estableciendo nuevas medidas para mejorar su gestión. En este contexto, la Directiva Marco del Agua (DMA)¹ establece el objetivo claro de que en el año 2015 se consiga un buen estado ecológico y un uso sostenible para todas las aguas europeas.

Este nuevo marco institucional pone de relieve la necesidad de caracterizar y analizar en profundidad la demanda de agua, lo que motiva el interés por utilizar instrumentos económicos para mejorar la eficiencia en el uso del agua. En este trabajo se realiza un análisis de la demanda de agua urbana en Andalucía, partiendo de los datos de consumo de agua en una zona urbana, correspondientes al periodo 1984 – 2011. Así mismo se desarrolla una nueva metodología que resulta fundamental para obtener predicciones de demanda con mayor precisión.

Para el estudio del comportamiento de este tipo de situaciones se utilizarán tanto técnicas de tipo estadístico como modelos ARIMA. Estos últimos se han empleado con buenos resultados para la predicción en econometría, la elección de estos modelos está basada en el modelo de mejor ajuste, se pretende incorporar técnicas más novedosas para realizar esta elección, de manera que cumpla mejor el objetivo del mismo, que es predecir la demanda correspondiente, por lo que se pretende, en este trabajo, modelar que la elección del modelo se base en la mejor capacidad predictiva del mismo, para ello se utilizaran modelos heurísticos.

En este trabajo se hará uso de las series temporales, éstas no son otra cosa que, datos numéricos de interés estadístico, relacionados con una temática en concreto y recopilados de manera periódica (anualmente, semestralmente, cuatrimestralmente, trimestralmente, bimensualmente, mensualmente, semanalmente o diariamente) e histórica a lo largo de los sucesivos años, cuya principal necesidad es observar la evolución de estos datos y obtener conclusiones estadísticas y/o económicas. La demanda de productos en cada punto se representa mediante un proceso estocástico. Una vez introducido el concepto genérico de proceso estocástico puede decirse que una serie temporal cualquiera es, en realidad, una muestra, una realización concreta con unos valores concretos de un proceso estocástico teórico, real. El análisis de series que vamos a estudiar tratará, a partir de los datos de una serie temporal, inferir las características de la estructura probabilística subyacente, del verdadero proceso estocástico. Se utilizaran los modelos SARIMA (Modelos Autorregresivos Integrados de Medias Móviles Estacionarios)

¹DIRECTIVA 2000/60/CE, del 23 de octubre del 2000

La organización de esta tesis se presenta de la siguiente manera:

- En el capítulo 1 se llevará a cabo una revisión teórica de los procesos estocásticos, desde los modelos más simples $AR(p)$ hasta los modelos más complejos, $SARIMA(p, d, q)(P, D, Q)_7$, con el objetivo de asentar las bases teóricas donde se apoya este trabajo.
- En el capítulo 2 se desarrollará la teoría básica sobre los criterios clásicos actuales que se encuentran en uso para la determinación del modelo seguido por una serie temporal cualquiera de la que no se conoce su procedencia a priori.
- En el capítulo 3 es el capítulo principal de esta tesis, ya que en él se desarrolla la parte innovadora del estudio, se realizará el desarrollo que se ha seguido hasta llegar a la propuesta realizada en este trabajo sobre unos nuevos criterios tanto de ajuste como de predicción, que contemplen y recojan en su estructura algunos aspectos de los que actualmente no se consideran en los criterios existentes en la literatura.
- En el capítulo 4, es más extenso del trabajo de investigación realizado, es precisamente toda la parte de simulación llevada a cabo para la comprobación del buen funcionamiento de los criterios propuestos. Para ello se han utilizado tanto modelos ARIMA como SARIMA.
- En el capítulo 5 se desarrolla la aplicación mediante la cuál se lleva a la práctica el funcionamiento de estos criterios innovadores en datos reales, ya que en muchas ocasiones y sobre todo en este campo, la teoría es un aspecto importante pero realmente el paso de lo teórico a lo práctico hay que considerarlo, ya que en ocasiones teorías perfectamente desarrolladas al llevarlas al campo práctico de datos reales fracasan.
- En el capítulo 6 se expondrán las conclusiones obtenidas a lo largo del estudio y las futuras líneas de investigación que se pueden desarrollar a partir de los alcances realizados en este trabajo.

Tras estos capítulos se han añadido unos apéndices que hemos creído de interés:

- En el apéndice A, se desarrollan algunas nociones básicas sobre el software utilizado para la implementación del programa utilizado en el capítulo de simulación.
- En el apéndice B, se muestra el código fuente del programa implementado.
- En el apéndice C, se recogen todos los modelos que han sido considerados en el estudio realizado.

Índice

Agradecimientos	IX
Resumen	XI
1. Introducción	1
1.1. Introducción	1
1.2. Reseña Histórica	2
1.3. Series Temporales	3
1.4. Procesos Estocásticos	5
1.5. Procesos Estacionarios y Ergódicos	5
1.6. Procesos Lineales	7
1.7. Modelos Lineales	10
1.7.1. Modelos Autorregresivos, AR(p)	10
1.7.2. Modelos de Medias Móviles, MA(q)	12
1.7.3. Modelos Mixtos Autorregresivos-Medias móviles, AR- MA(p,q)	14
1.8. Procesos no Estacionarios	18
1.9. Procesos Estacionarios	21
1.10. Otras Cuestiones sobre Modelos ARIMA	23
2. Criterios de Selección de Modelos	27
2.1. Introducción	27
2.2. Selección de Modelos ARIMA	28
2.3. Criterios de Bondad de Ajuste	28
2.3.1. Criterios basados en el error cuadrático medio	28
2.3.2. Criterios basados en la información de Kullback-Leibler	29
2.3.3. Criterios basados en la teoría Bayesiana	32
2.4. Criterios de Predicción	34
2.4.1. Índice de Desigualdad de Theil	35
3. Una Nueva Metodología para Selección de Modelos	37
3.1. Introducción	37

3.2.	Selección Automática de Modelos ARIMA para Predicción . . .	39
3.3.	Funciones de Proximidad para Evaluar las Predicciones	40
3.4.	Funciones de Proximidad Exponenciales	42
3.5.	Validación	43
4.	Simulación de Modelos	45
4.1.	Introducción	45
4.2.	Procesos no estacionarios, ARIMA(p, d, q)	47
4.3.	Procesos Estacionarios, SARIMA(p, d, q)(P, D, Q) _s	125
5.	Aplicación: La Demanda de Agua	261
5.1.	Introducción	261
5.2.	Planteamiento del problema	262
5.3.	Objetivos del problema	263
5.4.	Metodología	264
5.4.1.	Criterios de Selección de modelos	265
5.4.2.	Metodología Box-Jenkins	269
5.4.3.	Redes neuronales	271
5.5.	Comparación de metodologías	274
6.	Conclusiones y futuras líneas de investigación	277
A.	Algunas nociones sobre R-project	281
A.1.	¿Qué es R?	281
A.2.	El Entorno R	282
A.3.	Fundación R	283
B.	Código Fuente en R	285
B.1.	El Programa de Simulación	285
B.2.	Programa para la Aplicación	296
B.2.1.	Predicciones a partir de los criterios	296
B.2.2.	Red neuronal	299
C.	Espacio de Modelos	301
C.1.	Conjunto de los M Modelos	301
	Bibliografía	307

Índice de figuras

1.1. Ruido Blanco	8
1.2. Camino Aleatorio	10
1.3. Modelo AR(1)	12
1.4. Modelo MA(1)	13
1.5. Modelo ARMA(1,1)	17
1.6. Modelo ARIMA(1)	21
1.7. Modelo SARIMA(1)	23
1.8. Esquema de Identificación	26
4.1. Residuos obtenidos con los modelos de los criterios de ajuste de ARIMA(1,1,0).	48
4.2. Residuos obtenidos con los modelos de los criterios de predicción de ARIMA(1,1,0).	48
4.3. Histograma de los residuos en el conjunto de comprobación por los criterios de ajuste de ARIMA(1,1,0).	50
4.4. Histograma de los residuos en el conjunto de comprobación por los criterios de predicción de ARIMA(1,1,0).	50
4.5. ECM de ajuste con los modelos obtenidos con los criterios de ajuste en ARIMA(1,1,0).	51
4.6. ECM de ajuste con los modelos obtenidos con los criterios de predicción en ARIMA(1,1,0).	52
4.7. Histograma de los ECM de ajuste con los modelos obtenidos con los criterios de ajuste en ARIMA(1,1,0).	52
4.8. Histograma de los ECM de ajuste con los modelos obtenidos con los criterios de predicción en ARIMA(1,1,0).	53
4.9. ECM de predicción con los modelos obtenidos con los criterios de ajuste en ARIMA(1,1,0).	54
4.10. ECM de predicción con los modelos obtenidos con los criterios de predicción en ARIMA(1,1,0).	54
4.11. Histograma de los ECM de predicción con los modelos obtenidos con los criterios de ajuste en ARIMA(1,1,0).	55

4.12. Histograma de los ECM de predicción con los modelos obtenidos con los criterios de predicción en ARIMA(1,1,0).	55
4.13. Residuos de las series correctas e incorrectas con el criterio AIC2 en ARIMA(1,1,0).	56
4.14. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en ARIMA(1,1,0).	57
4.15. Residuos de las series correctas e incorrectas con el criterio BIC en ARIMA(1,1,0).	58
4.16. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC en ARIMA(1,1,0).	59
4.17. Residuos de las series correctas e incorrectas con el criterio BIC2 en ARIMA(1,1,0)	60
4.18. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en ARIMA(1,1,0).	60
4.19. Residuos de las series correctas e incorrectas con el criterio WFMSE en ARIMA(1,1,0).	62
4.20. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMSE en ARIMA(1,1,0).	63
4.21. Residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).	64
4.22. Residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).	64
4.23. Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).	66
4.24. Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).	66
4.25. ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).	68
4.26. ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).	68
4.27. Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).	69
4.28. Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).	69
4.29. ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).	70
4.30. ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).	71
4.31. Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).	71
4.32. Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).	72

4.33. Residuos obtenidos con los modelos de los criterios de ajuste de ARIMA(0,1,1).	74
4.34. Residuos obtenidos con los modelos de los criterios de predicción de ARIMA(0,1,1).	74
4.35. Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de ARIMA(0,1,1).	76
4.36. Histograma de los residuos obtenidos con los modelos de los criterios de predicción de ARIMA(0,1,1).	76
4.37. ECM de ajuste de los modelos obtenidos con los criterios de ajuste en ARIMA(0,1,1).	77
4.38. ECM de ajuste de los modelos obtenidos con los criterios de predicción en ARIMA(0,1,1).	78
4.39. Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en ARIMA(0,1,1).	78
4.40. Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicción en ARIMA(0,1,1).	79
4.41. ECM de predicción de los modelos obtenidos con los criterios de ajuste en ARIMA(0,1,1).	80
4.42. ECM de predicción de los modelos obtenidos con los criterios de predicción en ARIMA(0,1,1).	80
4.43. Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en ARIMA(0,1,1)	81
4.44. Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en ARIMA(0,1,1)	81
4.45. Residuos de las series correctas e incorrectas con el criterio AIC2 en el ARIMA(0,1,1).	82
4.46. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en ARIMA(0,1,1).	83
4.47. Residuos de las series correctas e incorrectas con el criterio BIC en el ARIMA(0,1,1).	84
4.48. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC en ARIMA(0,1,1)	85
4.49. Residuos de las series correctas e incorrectas con el criterio BIC2 en el ARIMA(0,1,1).	86
4.50. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en ARIMA(0,1,1).	87
4.51. Residuos de las series correctas e incorrectas con el criterio WFMAE en el ARIMA(0,1,1).	88
4.52. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en ARIMA(0,1,1).	89
4.53. Residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).	90

4.54. Residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).	90
4.55. Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1)	91
4.56. Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1)	92
4.57. ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).	93
4.58. ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).	93
4.59. Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).	94
4.60. Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).	94
4.61. ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).	95
4.62. ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).	96
4.63. Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).	96
4.64. Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).	97
4.65. Residuos obtenidos con los modelos de los criterios de ajuste de ARIMA(1,1,1)	99
4.66. Residuos obtenidos con los modelos de los criterios de predicción de ARIMA(1,1,1)	99
4.67. Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de ARIMA(1,1,1)	101
4.68. Histograma de los residuos obtenidos con los modelos de los criterios de predicción de ARIMA(1,1,1)	101
4.69. ECM de ajuste de los modelos obtenidos con los criterios de ajuste en ARIMA(1,1,1).	102
4.70. ECM de ajuste de los modelos obtenidos con los criterios de predicción en ARIMA(1,1,1).	103
4.71. Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en ARIMA(1,1,1).	103
4.72. Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicción en ARIMA(1,1,1).	104
4.73. ECM de predicción de los modelos obtenidos con los criterios de ajuste en ARIMA(1,1,1).	105
4.74. ECM de predicción de los modelos obtenidos con los criterios de predicción en ARIMA(1,1,1).	105

4.75. Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en ARIMA(1,1,1).	106
4.76. Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en ARIMA(1,1,1).	106
4.77. Residuos de las series correctas e incorrectas con el criterio AIC2 en el ARIMA(1,1,1).	107
4.78. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en ARIMA(1,1,1.)	108
4.79. Residuos de las series correctas e incorrectas con el criterio BIC en el ARIMA(1,1,1).	109
4.80. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC en ARIMA(1,1,1).	110
4.81. Residuos de las series correctas e incorrectas con el criterio BIC2 en el ARIMA(1,1,1).	111
4.82. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en ARIMA(1,1,1).	111
4.83. Residuos de las series correctas e incorrectas con el criterio WFMSE en el ARIMA(1,1,1).	112
4.84. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMSE en ARIMA(1,1,1).	113
4.85. Residuos de las series correctas e incorrectas con el criterio WFMAE en el ARIMA(1,1,1).	114
4.86. ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en ARIMA(1,1,1).	114
4.87. Residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).	116
4.88. Residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).	116
4.89. Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).	118
4.90. Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).	118
4.91. ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).	119
4.92. ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).	120
4.93. Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).	120
4.94. Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).	121
4.95. ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1)	122

4.96. ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1)	122
4.97. Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1) .	123
4.98. Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1) .	123
4.99. Residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(0, 1, 1)(0, 1, 1) ₇	126
4.100 Residuos obtenidos con los modelos de los criterios de predicción de SARIMA(0, 1, 1)(0, 1, 1) ₇	126
4.101 Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(0, 1, 1)(0, 1, 1) ₇	127
4.102 Histograma de los residuos obtenidos con los modelos de los criterios de predicción de SARIMA(0, 1, 1)(0, 1, 1) ₇	128
4.103 ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(0, 1, 1)(0, 1, 1) ₇	129
4.104 ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(0, 1, 1)(0, 1, 1) ₇	130
4.105 Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(0, 1, 1)(0, 1, 1) ₇	130
4.106 Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(0, 1, 1)(0, 1, 1) ₇	131
4.107 ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(0, 1, 1)(0, 1, 1) ₇	131
4.108 ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(0, 1, 1)(0, 1, 1) ₇	132
4.109 Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(0, 1, 1)(0, 1, 1) ₇	132
4.110 Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(0, 1, 1)(0, 1, 1) ₇	133
4.111 Residuos de las series correctas e incorrectas con el criterio AIC2 en el SARIMA(0, 1, 1)(0, 1, 1) ₇	134
4.112 ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en SARIMA(0, 1, 1)(0, 1, 1) ₇	134
4.113 Residuos de las series correctas e incorrectas con el criterio BIC en el SARIMA(0, 1, 1)(0, 1, 1) ₇	136
4.114 ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC en SARIMA(0, 1, 1)(0, 1, 1) ₇	136
4.115 Residuos de las series correctas e incorrectas con el criterio BIC2 en el SARIMA(0, 1, 1)(0, 1, 1) ₇	138
4.116 ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en SARIMA(0, 1, 1)(0, 1, 1) ₇	139

4.117 Residuos de las series correctas e incorrectas con el criterio WFMAE en el SARIMA(0, 1, 1)(0, 1, 1)₇. 139

4.118 ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en SARIMA(0, 1, 1)(0, 1, 1)₇. 140

4.119 Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇. 141

4.120 Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇. 141

4.121 Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇. 142

4.122 Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇. 143

4.123 ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇. 144

4.124 ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇. 144

4.125 Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇. 145

4.126 Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇. 145

4.127 ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇. 146

4.128 ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇. 147

4.129 Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇. 147

4.130 Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇. 148

4.131 Residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 0)(0, 1, 1)₇. 150

4.132 Residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 1, 0)(0, 1, 1)₇. 150

4.133 Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 0)(0, 1, 1)₇. 151

4.134 Histograma de los residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 1, 0)(0, 1, 1)₇. 152

4.135 ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(0, 1, 1)₇. 153

4.136 ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(0, 1, 1)₇. 153

4.137 Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(0, 1, 1)₇. 154

4.138	Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(0, 1, 1) ₇ . . .	154
4.139	ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(0, 1, 1) ₇	155
4.140	ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(0, 1, 1) ₇	156
4.141	Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(0, 1, 1) ₇ . . .	156
4.142	Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(0, 1, 1) ₇ . . .	157
4.143	Residuos de las series correctas e incorrectas con el criterio AIC2 en el SARIMA(1, 1, 0)(0, 1, 1) ₇	158
4.144	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en SARIMA(1, 1, 0)(0, 1, 1) ₇	159
4.145	Residuos de las series correctas e incorrectas con el criterio BIC en el SARIMA(1, 1, 0)(0, 1, 1) ₇	160
4.146	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC en SARIMA(1, 1, 0)(0, 1, 1) ₇	161
4.147	Residuos de las series correctas e incorrectas con el criterio BIC2 en el SARIMA(1, 1, 0)(0, 1, 1) ₇	161
4.148	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en SARIMA(1, 1, 0)(0, 1, 1) ₇	162
4.149	Residuos de las series correctas e incorrectas con el criterio WFMSE en el SARIMA(1, 1, 0)(0, 1, 1) ₇	163
4.150	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMSE en SARIMA(1, 1, 0)(0, 1, 1) ₇	164
4.151	Residuos de las series correctas e incorrectas con el criterio WFMAE en el SARIMA(1, 1, 0)(0, 1, 1) ₇	164
4.152	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en SARIMA(1, 1, 0)(0, 1, 1) ₇	165
4.153	Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) ₇	166
4.154	Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) ₇	166
4.155	Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) ₇ . . .	167
4.156	Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) ₇ . . .	168
4.157	ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) ₇	169
4.158	ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) ₇	169

4.159	Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) ₇ .	170
4.160	Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) ₇ .	170
4.161	ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) ₇ .	171
4.162	ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) ₇ .	172
4.163	Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) ₇ .	172
4.164	Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) ₇ .	173
4.165	Residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 0)(1, 1, 0) ₇ .	175
4.166	Residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 1, 0)(1, 1, 0) ₇ .	175
4.167	Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 0)(1, 1, 0) ₇ .	176
4.168	Histograma de los residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 1, 0)(1, 1, 0) ₇ .	177
4.169	ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 1, 0) ₇ .	178
4.170	ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(1, 1, 0) ₇ .	178
4.171	Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 1, 0) ₇ .	179
4.172	Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(1, 1, 0) ₇ .	179
4.173	ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 1, 0) ₇ .	180
4.174	ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(1, 1, 0) ₇ .	181
4.175	Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 1, 0) ₇ .	181
4.176	Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(1, 1, 0) ₇ .	182
4.177	Residuos de las series correctas e incorrectas con el criterio AIC2 en el SARIMA(1, 1, 0)(1, 1, 0) ₇ .	183
4.178	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en SARIMA(1, 1, 0)(1, 1, 0) ₇ .	184
4.179	Residuos de las series correctas e incorrectas con el criterio BIC en el SARIMA(1, 1, 0)(1, 1, 0) ₇ .	185

4.180	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC en SARIMA(1, 1, 0)(1, 1, 0) ₇	185
4.181	Residuos de las series correctas e incorrectas con el criterio BIC2 en el SARIMA(1, 1, 0)(1, 1, 0) ₇	187
4.182	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en SARIMA(1, 1, 0)(1, 1, 0) ₇	188
4.183	Residuos de las series correctas e incorrectas con el criterio WFMAE en el SARIMA(1, 1, 0)(1, 1, 0) ₇	188
4.184	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en SARIMA(1, 1, 0)(1, 1, 0) ₇	189
4.185	Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) ₇	190
4.186	Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) ₇	190
4.187	Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) ₇	191
4.188	Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) ₇	192
4.189	ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) ₇	193
4.190	ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) ₇	193
4.191	Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) ₇	194
4.192	Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) ₇	194
4.193	ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) ₇	195
4.194	ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) ₇	196
4.195	Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) ₇	196
4.196	Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) ₇	197
4.197	Residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 0)(1, 0, 1) ₇	199
4.198	Residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 1, 0)(1, 0, 1) ₇	199
4.199	Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 0)(1, 0, 1) ₇	200
4.200	Histograma de los residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 1, 0)(1, 0, 1) ₇	201

4.201	ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 0, 1) ₇	202
4.202	ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(1, 0, 1) ₇	202
4.203	Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 0, 1) ₇	203
4.204	Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(1, 0, 1) ₇	203
4.205	ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 0, 1) ₇	204
4.206	ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(1, 0, 1) ₇	205
4.207	Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 0, 1) ₇	205
4.208	Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(1, 0, 1) ₇	206
4.209	Residuos de las series correctas e incorrectas con el criterio AIC2 en el SARIMA(1, 1, 0)(1, 0, 1) ₇	207
4.210	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en SARIMA(1, 1, 0)(1, 0, 1) ₇	207
4.211	Residuos de las series correctas e incorrectas con el criterio BIC2 en el SARIMA(1, 1, 0)(1, 0, 1) ₇	208
4.212	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en SARIMA(1, 1, 0)(1, 0, 1) ₇	209
4.213	Residuos de las series correctas e incorrectas con el criterio WFMAE en el SARIMA(1, 1, 0)(1, 0, 1) ₇	210
4.214	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en SARIMA(1, 1, 0)(1, 0, 1) ₇	211
4.215	Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) ₇	212
4.216	Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) ₇	212
4.217	Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) ₇	213
4.218	Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) ₇	214
4.219	ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) ₇	215
4.220	ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) ₇	215
4.221	Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) ₇	216

4.222	Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) ₇ .	216
4.223	ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) ₇ .	217
4.224	ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) ₇ .	218
4.225	Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) ₇ .	218
4.226	Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) ₇ .	219
4.227	Residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 0, 1)(0, 0, 1) ₇ .	221
4.228	Residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 0, 1)(0, 0, 1) ₇ .	221
4.229	Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 0, 1)(0, 0, 1) ₇ .	222
4.230	Histograma de los residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 0, 1)(0, 0, 1) ₇ .	223
4.231	ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 0, 1)(0, 0, 1) ₇ .	224
4.232	ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 0, 1)(0, 0, 1) ₇ .	224
4.233	Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 0, 1)(0, 0, 1) ₇ .	225
4.234	Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 0, 1)(0, 0, 1) ₇ .	225
4.235	ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 0, 1)(0, 0, 1) ₇ .	226
4.236	ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 0, 1)(0, 0, 1) ₇ .	227
4.237	Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 0, 1)(0, 0, 1) ₇ .	227
4.238	Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 0, 1)(0, 0, 1) ₇ .	228
4.239	Residuos de las series correctas e incorrectas con el criterio AIC2 en el SARIMA(1, 0, 1)(0, 0, 1) ₇ .	229
4.240	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en SARIMA(1, 0, 1)(0, 0, 1) ₇ .	230
4.241	Residuos de las series correctas e incorrectas con el criterio BIC2 en el SARIMA(1, 0, 1)(0, 0, 1) ₇ .	230
4.242	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en SARIMA(1, 0, 1)(0, 0, 1) ₇ .	231

4.243 Residuos de las series correctas e incorrectas con el criterio WMSE en el SARIMA(1, 0, 1)(0, 0, 1)₇. 232

4.244 ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WMSE en SARIMA(1, 0, 1)(0, 0, 1)₇. 233

4.245 Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇. 233

4.246 Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇. 234

4.247 Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇. 235

4.248 Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇. 235

4.249 ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇. 236

4.250 ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇. 237

4.251 Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇. 237

4.252 Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇. 238

4.253 ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇. 239

4.254 ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇. 239

4.255 Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇. 240

4.256 Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇. 240

4.257 Residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 1)(1, 1, 1)₇. 242

4.258 Residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 1, 1)(1, 1, 1)₇. 242

4.259 Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 1)(1, 1, 1)₇. 243

4.260 Histograma de los residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 1, 1)(1, 1, 1)₇. 244

4.261 ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 1)(1, 1, 1)₇. 245

4.262 ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 1)(1, 1, 1)₇. 245

4.263 Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 1)(1, 1, 1)₇. 246

4.264	Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 1)(1, 1, 1) ₇ . . .	246
4.265	ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 1)(1, 1, 1) ₇	247
4.266	ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 1)(1, 1, 1) ₇	248
4.267	Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 1)(1, 1, 1) ₇ . . .	248
4.268	Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 1)(1, 1, 1) ₇ .	249
4.269	Residuos de las series correctas e incorrectas con el criterio WFMAE en el SARIMA(1, 1, 1)(1, 1, 1) ₇	250
4.270	ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en SARIMA(1, 1, 1)(1, 1, 1) ₇ . . .	250
4.271	Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) ₇	251
4.272	Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) ₇	252
4.273	Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) ₇ . .	253
4.274	Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) ₇ . .	253
4.275	ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) ₇	254
4.276	ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) ₇	255
4.277	Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) ₇ .	255
4.278	Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) ₇ .	256
4.279	ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) ₇	257
4.280	ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) ₇	257
4.281	Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) ₇ .	258
4.282	Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) ₇ .	258
5.1.	Representación gráfica de la serie consumo de agua	265
5.2.	Representación de los residuos obtenidos con cada uno de los modelos estudiados	267

5.3.	Diagrama de cajas y bigotes de los residuos de predicción . . .	267
5.4.	Representación de las predicciones obtenidas frente a los valores reales de la serie	268
5.5.	Representación de la serie de consumo de agua con una diferencia estacional, gráfico de autocorrelación simple y gráfico de autocorrelación parcial	270
5.6.	Representación de la serie de consumo de agua con la primera diferencia y una diferencia estacional, gráfico de autocorrelación simple y gráfico de autocorrelación parcial	271
5.7.	Esquema de una red neuronal artificial	272
5.8.	Representación de las predicciones con los distintos métodos .	275

Índice de Tablas

4.1. Criterios Clásicos de ajuste para ARIMA(1,1,0).	47
4.2. Criterios Clásicos de predicción para ARIMA(1,1,0).	47
4.3. Criterios propuestos de ajuste y predicción para ARIMA(1,1,0).	47
4.4. Medidas descriptivas de los residuos en el conjunto de comprobación para ARIMA(1,1,0).	49
4.5. ECMpre global para cada predicción realizada en el conjunto de comprobación en ARIMA(1,1,0).	51
4.6. El ECM global de ajuste, ECMG, para cada criterio en modelos ARIMA(1,1,0).	53
4.7. ECM global de predicción, ECMGpre, para cada criterio en modelos ARIMA(1,1,0).	56
4.8. Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,0) con el criterio AIC2.	57
4.9. ECM global de predicción en ARIMA(1,1,0) con el criterio AIC2 en el conjunto de comprobación.	57
4.10. Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para ARIMA(1,1,0) con el criterio BIC.	58
4.11. ECM global de predicción en ARIMA(1,1,0) con el criterio BIC en el conjunto de comprobación.	59
4.12. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,0) con el criterio BIC2.	61
4.13. ECM global de predicción en ARIMA(1,1,0) con el criterio BIC2 en el conjunto de comprobación.	61
4.14. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,0) con el criterio WFMSE.	62
4.15. ECM global de predicción en ARIMA(1,1,0) con el criterio WFMSE en el conjunto de comprobación.	62
4.16. Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para ARIMA(1,1,0).	65

4.17. ECM globales para cada predicción realizada en el conjunto de comprobación en ARIMA(1,1,0).	67
4.18. ECM global de ajuste, ECMG, para cada criterio en modelos ARIMA(1,1,0).	70
4.19. ECM global de predicción, ECMGpre, para cada criterio en modelos ARIMA(1,1,0).	72
4.20. Criterios Clásicos de ajuste para ARIMA(0,1,1).	73
4.21. Criterios Clásicos de predicción para ARIMA(0,1,1).	73
4.22. Criterios propuestos para ARIMA(0,1,1).	73
4.23. Medidas descriptivas de los residuos en el conjunto de comprobación para ARIMA(0,1,1).	75
4.24. ECMpre global para cada predicción realizada en el conjunto de comprobación en ARIMA(0,1,1).	77
4.25. El ECM global de ajuste, ECMG, para cada criterio en modelos ARIMA(0,1,1).	79
4.26. ECM global de predicción, ECMGpre, para cada criterio en modelos ARIMA(0,1,1).	82
4.27. Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(0,1,1) con el criterio AIC2.	83
4.28. ECM global de predicción en ARIMA(0,1,1) con el criterio AIC2 en el conjunto de comprobación.	84
4.29. Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para ARIMA(0,1,1) con el criterio BIC.	84
4.30. ECM global de predicción en ARIMA(0,1,1) con el criterio BIC en el conjunto de comprobación.	85
4.31. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(0,1,1) con el criterio BIC2.	86
4.32. ECM global de predicción en ARIMA(0,1,1) con el criterio BIC2 en el conjunto de comprobación.	87
4.33. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(0,1,1) con el criterio WFMAE.	88
4.34. ECM global de predicción en ARIMA(0,1,1) con el criterio WFMAE en el conjunto de comprobación.	89
4.35. Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para ARIMA(0,1,1).	91
4.36. ECM globales para cada predicción realizada en el conjunto de comprobación en ARIMA(0,1,1).	92
4.37. ECM global de ajuste, ECMG, para cada criterio en modelos ARIMA(0,1,1).	95

4.38. ECM global de predicción, ECMGpre, para cada criterio en modelos ARIMA(0,1,1).	97
4.39. Criterios Clásicos de ajuste para ARIMA(1,1,1).	98
4.40. Criterios Clásicos de predicción para ARIMA(1,1,1).	98
4.41. Criterios propuestos para ARIMA(1,1,1).	98
4.42. Medidas descriptivas de los residuos en el conjunto de comprobación para ARIMA(1,1,1).	100
4.43. ECMpre global para cada predicción realizada en el conjunto de comprobación en ARIMA(1,1,1).	102
4.44. El ECM global de ajuste, ECMG, para cada criterio en modelos ARIMA(1,1,1).	104
4.45. ECM global de predicción, ECMGpre, para cada criterio en modelos ARIMA(1,1,1).	107
4.46. Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,1) con el criterio AIC2.	108
4.47. ECM global de predicción en ARIMA(1,1,1) con el criterio AIC2 en el conjunto de comprobación.	108
4.48. Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para ARIMA(1,1,1) con el criterio BIC.	109
4.49. ECM global de predicción en ARIMA(1,1,1) con el criterio BIC en el conjunto de comprobación.	110
4.50. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,1) con el criterio BIC2.	112
4.51. ECM global de predicción en ARIMA(1,1,1) con el criterio BIC2 en el conjunto de comprobación.	112
4.52. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,1) con el criterio WFMSE.	113
4.53. ECM global de predicción en ARIMA(1,1,0) con el criterio WFMSE en el conjunto de comprobación.	113
4.54. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,1) con el criterio WFMAE.	115
4.55. ECM global de predicción en ARIMA(1,1,0) con el criterio WFMAE en el conjunto de comprobación.	115
4.56. Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para ARIMA(1,1,1).	117
4.57. ECM globales para cada predicción realizada en el conjunto de comprobación en ARIMA(1,1,1).	119

4.58. ECM global de ajuste, ECMG, para cada criterio en modelos ARIMA(1,1,1).	121
4.59. ECM global de predicción, ECMGpre, para cada criterio en modelos ARIMA(1,1,1).	124
4.60. Criterios Clásicos de ajuste para SARIMA(0, 1, 1)(0, 1, 1) ₇ . . .	125
4.61. Criterios Clásicos de predicción para SARIMA(0, 1, 1)(0, 1, 1) ₇ . . .	125
4.62. Criterios propuestos para SARIMA(0, 1, 1)(0, 1, 1) ₇	125
4.63. Medidas descriptivas de los residuos en el conjunto de comprobación para SARIMA(0, 1, 1)(0, 1, 1) ₇	127
4.64. ECMpre global para cada predicción realizada en el conjunto de comprobación en SARIMA(0, 1, 1)(0, 1, 1) ₇	128
4.65. El ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(0, 1, 1)(0, 1, 1) ₇	129
4.66. ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(0, 1, 1)(0, 1, 1) ₇	133
4.67. Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(0, 1, 1)(0, 1, 1) ₇ con el criterio AIC2.	135
4.68. ECM global de predicción en SARIMA(0, 1, 1)(0, 1, 1) ₇ con el criterio AIC2 en el conjunto de comprobación.	135
4.69. Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para SARIMA(0, 1, 1)(0, 1, 1) ₇ con el criterio BIC.	137
4.70. ECM global de predicción en SARIMA(0, 1, 1)(0, 1, 1) ₇ con el criterio BIC en el conjunto de comprobación.	137
4.71. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(0, 1, 1)(0, 1, 1) ₇ con el criterio BIC2.	138
4.72. ECM global de predicción en SARIMA(0, 1, 1)(0, 1, 1) ₇ con el criterio BIC2 en el conjunto de comprobación.	138
4.73. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(0, 1, 1)(0, 1, 1) ₇ con el criterio WFMAE.	140
4.74. ECM global de predicción en SARIMA(0, 1, 1)(0, 1, 1) ₇ con el criterio WFMAE en el conjunto de comprobación.	140
4.75. Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para SARIMA(0, 1, 1)(0, 1, 1) ₇	142
4.76. ECM globales para cada predicción realizada en el conjunto de comprobación en SARIMA(0, 1, 1)(0, 1, 1) ₇	143
4.77. ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(0, 1, 1)(0, 1, 1) ₇	146

4.78. ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(0, 1, 1)(0, 1, 1)₇. 148

4.79. Criterios Clásicos de ajuste para SARIMA(1, 1, 0)(0, 1, 1)₇. . . 149

4.80. Criterios Clásicos de predicción para SARIMA(1, 1, 0)(0, 1, 1)₇. 149

4.81. Criterios propuestos para SARIMA(1, 1, 0)(0, 1, 1)₇. 149

4.82. Medidas descriptivas de los residuos en el conjunto de comprobación para SARIMA(1, 1, 0)(0, 1, 1)₇. 151

4.83. ECMpre global para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 0)(0, 1, 1)₇. 152

4.84. El ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 0)(0, 1, 1)₇. 155

4.85. ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 0)(0, 1, 1)₇. 157

4.86. Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio AIC2. 158

4.87. ECM global de predicción en SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio AIC2 en el conjunto de comprobación. 159

4.88. Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio BIC. 160

4.89. ECM global de predicción en SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio BIC en el conjunto de comprobación. 160

4.90. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio BIC2. 162

4.91. ECM global de predicción en SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio BIC2 en el conjunto de comprobación. 162

4.92. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio WFMSE. 163

4.93. ECM global de predicción en SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio WFMSE en el conjunto de comprobación. 163

4.94. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio WFMAE. 165

4.95. ECM global de predicción en SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio WFMAE en el conjunto de comprobación. 165

4.96. Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para SARIMA(1, 1, 0)(0, 1, 1)₇. 167

4.97. ECM globales para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 0)(0, 1, 1)₇. 168

4.98. ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 0)(0, 1, 1) ₇	171
4.99. ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 0)(0, 1, 1) ₇	173
4.100. Criterios Clásicos de ajuste para SARIMA(1, 1, 0)(1, 1, 0) ₇	174
4.101. Criterios Clásicos de predicción para SARIMA(1, 1, 0)(1, 1, 0) ₇	174
4.102. Criterios propuestos para SARIMA(1, 1, 0)(1, 1, 0) ₇	174
4.103. Medidas descriptivas de los residuos en el conjunto de comprobación para SARIMA(1, 1, 0)(1, 1, 0) ₇	176
4.104. ECMpre global para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 0)(1, 1, 0) ₇	177
4.105. El ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 0)(1, 1, 0) ₇	180
4.106. ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 0)(1, 1, 0) ₇	182
4.107. Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(1, 1, 0) ₇ con el criterio AIC2.	183
4.108. ECM global de predicción en SARIMA(1, 1, 0)(1, 1, 0) ₇ con el criterio AIC2 en el conjunto de comprobación.	184
4.109. Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para SARIMA(1, 1, 0)(1, 1, 0) ₇ con el criterio BIC.	186
4.110. ECM global de predicción en SARIMA(1, 1, 0)(1, 1, 0) ₇ con el criterio BIC en el conjunto de comprobación.	186
4.111. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(1, 1, 0) ₇ con el criterio BIC2.	187
4.112. ECM global de predicción en SARIMA(1, 1, 0)(1, 1, 0) ₇ con el criterio BIC2 en el conjunto de comprobación.	187
4.113. Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(1, 1, 0) ₇ con el criterio WFMAE.	189
4.114. ECM global de predicción en SARIMA(1, 1, 0)(1, 1, 0) ₇ con el criterio WFMAE en el conjunto de comprobación.	189
4.115. Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para SARIMA(1, 1, 0)(1, 1, 0) ₇	191
4.116. ECM globales para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 0)(1, 1, 0) ₇	192
4.117. ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 0)(1, 1, 0) ₇	195

4.118	ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 0)(1, 1, 0) ₇	197
4.119	Criterios Clásicos de ajuste para SARIMA(1, 1, 0)(1, 0, 1) ₇	198
4.120	Criterios Clásicos de predicción para SARIMA(1, 1, 0)(1, 0, 1) ₇	198
4.121	Criterios propuestos para SARIMA(1, 1, 0)(1, 0, 1) ₇	198
4.122	Medidas descriptivas de los residuos en el conjunto de comprobación para SARIMA(1, 1, 0)(1, 0, 1) ₇	200
4.123	ECMpre global para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 0)(1, 0, 1) ₇	201
4.124	El ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 0)(1, 0, 1) ₇	204
4.125	ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 0)(1, 0, 1) ₇	206
4.126	Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(1, 0, 1) ₇ con el criterio AIC2.	208
4.127	ECM global de predicción en SARIMA(1, 1, 0)(1, 0, 1) ₇ con el criterio AIC2 en el conjunto de comprobación.	208
4.128	Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(1, 0, 1) ₇ con el criterio BIC2.	209
4.129	ECM global de predicción en SARIMA(1, 1, 0)(1, 0, 1) ₇ con el criterio BIC2 en el conjunto de comprobación.	209
4.130	Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(1, 0, 1) ₇ con el criterio WFMSE.	211
4.131	ECM global de predicción en SARIMA(1, 1, 0)(1, 0, 1) ₇ con el criterio WFMAE en el conjunto de comprobación.	211
4.132	Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para SARIMA(1, 1, 0)(1, 0, 1) ₇	213
4.133	ECM globales para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 0)(1, 0, 1) ₇	214
4.134	ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 0)(1, 0, 1) ₇	217
4.135	ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 0)(1, 0, 1) ₇	219
4.136	Criterios Clásicos de ajuste para SARIMA(1, 0, 1)(0, 0, 1) ₇	220
4.137	Criterios Clásicos de predicción para SARIMA(1, 0, 1)(0, 0, 1) ₇	220
4.138	Criterios propuestos para SARIMA(1, 0, 1)(0, 0, 1) ₇	220
4.139	Medidas descriptivas de los residuos en el conjunto de comprobación para SARIMA(1, 0, 1)(0, 0, 1) ₇	222

4.140	ECMpre global para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 0, 1)(0, 0, 1) ₇	223
4.141	El ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 0, 1)(0, 0, 1) ₇	226
4.142	ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 0, 1)(0, 0, 1) ₇	228
4.143	Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 0, 1)(0, 0, 1) ₇ con el criterio AIC2.	229
4.144	ECM global de predicción en SARIMA(1, 0, 1)(0, 0, 1) ₇ con el criterio AIC2 en el conjunto de comprobación.	229
4.145	Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para SARIMA(1, 0, 1)(0, 0, 1) ₇ con el criterio BIC2.	231
4.146	ECM global de predicción en SARIMA(1, 0, 1)(0, 0, 1) ₇ con el criterio BIC2 en el conjunto de comprobación.	231
4.147	Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 0, 1)(0, 0, 1) ₇ con el criterio WMSE.	232
4.148	ECM global de predicción en SARIMA(1, 0, 1)(0, 0, 1) ₇ con el criterio WMSE en el conjunto de comprobación.	232
4.149	Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para SARIMA(1, 0, 1)(0, 0, 1) ₇	234
4.150	ECM globales para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 0, 1)(0, 0, 1) ₇	236
4.151	ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 0, 1)(0, 0, 1) ₇	238
4.152	ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 0, 1)(0, 0, 1) ₇	241
4.153	Criterios Clásicos de ajuste para SARIMA(1, 1, 1)(1, 1, 1) ₇	241
4.154	Criterios Clásicos de predicción para SARIMA(1, 1, 1)(1, 1, 1) ₇	241
4.155	Criterios propuestos para SARIMA(1, 1, 1)(1, 1, 1) ₇	241
4.156	Medidas descriptivas de los residuos en el conjunto de comprobación para SARIMA(1, 1, 1)(1, 1, 1) ₇	243
4.157	ECMpre global para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 1)(1, 1, 1) ₇	244
4.158	El ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 1)(1, 1, 1) ₇	247
4.159	ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 1)(1, 1, 1) ₇	249

4.160	Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 1)(1, 1, 1) ₇ con el criterio WFMAE.	251
4.161	ECM global de predicción en SARIMA(1, 1, 1)(1, 1, 1) ₇ con el criterio WFMAE en el conjunto de comprobación.	251
4.162	Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para SARIMA(1, 1, 1)(1, 1, 1) ₇	252
4.163	ECM globales para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 1)(1, 1, 1) ₇	254
4.164	ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 1)(1, 1, 1) ₇	256
4.165	ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 1)(1, 1, 1) ₇	259
5.1.	Modelos seleccionados con cada uno de los criterios.	266
5.2.	ECM de ajuste y predicción en cada modelo.	268
5.3.	Predicciones anuales expresadas en 1000001.	269
5.4.	ECM de ajuste y predicción con la metodología Box-Jenkins.	271
5.5.	ECM de ajuste y predicción en Redes Neuronales con distintas capas ocultas.	273
5.6.	ECM de ajuste y predicción con la metodología Box-Jenkins, Sarima y Redes Neuronales.	275
C.1.	Espacio de Modelos	305

Capítulo 1

Introducción

Estudia el pasado si quieres pronosticar el futuro.

Confucio (Pensador y filósofo chino)

Según la concepción lineal del tiempo que tienen la mayoría de las civilizaciones humanas, el futuro es la porción de la línea temporal que todavía no ha sucedido; en otras palabras, es una conjetura que bien puede ser anticipada, predicha, especulada, postulada, teorizada o calculada a partir de datos en un instante de tiempo concreto.

1.1. Introducción

El análisis de series temporales, y la incorporación de técnicas prospectivas de tipo econométrico se basan, en última instancia en considerar que el futuro se parecerá al pasado, o que las relaciones causales actuales se mantendrán durante un cierto tiempo. Sin embargo, aunque el presente y el pasado son de naturaleza determinista, nuestro conocimiento de ambos es incompleto, y la interpretación que se atribuye a sucesos pasados está afectado de un alto grado de subjetividad.

Los procesos de predicción se apoyan en diversas técnicas estadísticas, pero su finalidad está ligada al proceso diario de toma de decisiones.

En definitiva, la necesidad de realizar predicciones es inherente a bastantes actividades humanas. No obstante, sólo en pocas ocasiones se utilizan métodos cuantitativos en los procesos de predicción.

Al analizar una serie temporal es preciso investigar su estructura interna, ya que la desagregación de una serie de componentes permite acceder a la información específica de cada una de éstas. La mayor parte de los métodos univariantes someten a una serie a una sucesión de transformaciones mediante filtros que ponen de manifiesto alguna característica útil.

Los filtros de uso más frecuente son los que corresponden a bandas de frecuencias bajas (para estudiar la evolución a largo plazo), a bandas altas (orientadas a las oscilaciones a muy corto plazo), y a bandas de frecuencias asociadas a componentes cíclicas como la estacionalidad.

Así, al filtrar una serie se persiguen alguno de los objetivos siguientes:

- eliminar componentes o ruidos no deseados,
- resaltar características importantes, y
- descomponer la serie en partes que están asociadas a elementos informativos distintos.

Existen numerosos filtros que se usan con frecuencia: medias móviles, regresión, alisados, diferencias, etc. Desgraciadamente, el filtrado de una serie, aunque permite resaltar algunas de sus características, no es sin inconvenientes, pues, distorsiona otras propiedades de los datos. Caridad y Ocerin y Caridad y López del Río (2003)

1.2. Reseña Histórica

Graunt (1662), hace una interpretación de los movimientos demográficos y de la conducta social, a partir de las cantidades brutas de nacimientos y muertes, por causas, en Londres. La fuente de datos utilizada fue de las Tablas Semanales de la Compañía de Sacristanes Parroquiales. Estas tablas contenían la cantidad de muertes, por causas, y la cantidad de nacimientos, ocurrida en cada parroquia londinense durante el periodo de 1604 a 1661. Se cree que éste haya sido el primer análisis realizado sobre series temporales. Halley (1693), a partir de series temporales, de más de cien años, sobre nacimientos y defunciones ocurridos en la ciudad de Breslaw, Silesia, construye las primeras tablas de esperanza de vida. La metodología utilizada es la vigente hasta hoy.

Farr (1840), siguiendo la filosofía de trabajo de Graunt, pero aplicando el enfoque de S.M. Poisson al tratamiento de series temporales, realiza varios pronósticos de incidencia por viruelas. Para ello ajusta curvas a series temporales de notificaciones de casos en diferentes localidades de Inglaterra y Gales, y procede a su extrapolación. A Farr se debe el concepto de “exceso de mortalidad” utilizado en los análisis de series temporales de mortalidad. Lo definió como “la cantidad de muertes que excede a la esperada por determinada causa para un momento de tiempo y localidad dadas”.

En 1920, W. M. Pearson propone las ideas básicas del enfoque clásico en el tratamiento de series temporales. Este método se desarrolló para el análisis de las series económicas. En la Universidad de Harvard el método propuesto por Pearson fue perfeccionado, aplicándose el tratamiento de series temporales en diversos ámbitos.

Entre las primeras aplicaciones del análisis de series temporales propuesto por Pearson, a un problema de salud, están los trabajos de S. D. Collins para el estudio de la mortalidad excesiva, según el criterio de Farr.

Collins (1932) Collins estima la tasa semanal de muertes esperadas como un promedio de las tasas de muerte observadas para esa semana durante los años epidémicos anteriores. Alrededor de este valor central construye un intervalo de confianza, cuyo límite superior considera como el nivel de mortalidad máxima usual. A partir del trabajo de Collins se ponen en boga los "canales endémicos" sobre mortalidad o morbilidad. Estos canales diferían en la forma de obtener el valor central esperado, a partir de las observaciones de la serie consideradas como "no epidémicas", y en la forma de construir el intervalo de confianza para el pronóstico. Para el valor central se solía utilizar la media aritmética, la mediana, la moda o los promedios tricentrales. Para el intervalo de confianza la desviación típica, la varianza, los recorridos, las amplitudes, el recorrido medio, los percentiles, los mínimos-maximorum (valor mínimo entre los máximos) y los máximos-minimorum (valor máximo entre los mínimos). Cayeron en desuso ante la aparición de los métodos autorregresivos.

Eickhoff et al. (1962) utilizan una ecuación de autorregresión, sobre una serie de tasas semanales de muertes, con la omisión de las tasas correspondientes a las semanas epidémicas, para determinar la tasa de muertes usual esperada para cada semana del año. Serfling (1963), aplica el método de análisis armónico de Fourier a series de tasas de incidencia acumulativa semanal. Define un intervalo de confianza de 1,64 desviaciones típicas para los valores de extrapolación obtenidos. Una de las novedades del método era la autoestimación de valores epidémicos. Serfling aplica este método al análisis de la mortalidad por influenza y neumonía, con fines de vigilancia epidemiológica. Para ello establece el criterio de considerar la situación como de alerta epidémica, cuando dos tasas consecutivas sobrepasen el límite superior del intervalo de confianza, ante evidencias de laboratorio.

Una modificación del método de Serfling es adoptada por los Centros de Control de Enfermedades de los Estados Unidos, como instrumento de vigilancia epidemiológica. Es abandonado en 1982, cuando surge la primera aplicación del método ARIMA con los mismos fines Choi y Thacker (1981a), Choi y Thacker (1981b).

1.3. Series Temporales

Una serie temporal es un conjunto de observaciones tomadas secuencialmente en el tiempo. Muchos conjuntos de datos se presentan como series temporales: serie semanal de accidentes de carretera, consumo anual de aceite en los hogares españoles, etc. Si el conjunto es continuo, se dice que la serie es continua, mientras que si el conjunto es discreto se dice que es discreta. En

este trabajo consideramos solamente series temporales discretas, dado que las variables económicas suelen observarse en tiempo discreto, por lo que, matizamos la definición de serie temporal, enunciando que una serie temporal está constituida por unas observaciones de una variable a intervalos regulares de tiempo, Uriel Jiménez (1985). Así, cuando tenemos N valores sucesivos de una serie discreta escribiremos x_1, x_2, \dots, x_N para denotar observaciones realizadas a intervalos regulares de tiempo $t_0 + h, t_0 + 2h, \dots, t_0 + th, \dots, t_0 + Nh$. Si tomamos t_0 como el origen y h como la unidad de tiempo, podemos considerar x_t como la observación en el período t .

Por otro lado, debemos también hacer la distinción entre serie temporal determinista y estadística. Si los valores futuros de una serie temporal vienen exactamente determinados por alguna función matemática, se dice que la serie temporal es determinista. En cambio, si los valores futuros solo pueden ser descritos en términos de una distribución de probabilidad, la serie temporal se dice que es no determinista o simplemente que es una serie temporal estadística (Box, Jenkins y Reinsel, 1994). En este segundo caso, aunque el comportamiento de la serie hasta el momento actual esté bien definido, es imposible pronosticar el valor exacto que tomará en un nuevo período. Dado que la gran mayoría de series temporales con las que nos encontramos en la práctica son de naturaleza estadística, pues hay una gran cantidad de factores desconocidos que influyen de manera decisiva en el comportamiento de una serie y que el investigador no puede controlar, nuestro estudio se centra precisamente en este tipo de series temporales. La información que se maneja en una serie temporal es generalmente de carácter agregado, aunque el grado de agregación dependerá del sistema que se esté tratando. Así, en una empresa, mediante agregación de las ventas efectuadas a cada cliente, se obtiene la serie temporal de ventas. A nivel macroeconómico la agregación suele hacerse por métodos estadísticos, es decir, mediante inferencia al conjunto de la población de lo observado en un subconjunto de la misma. En cualquier caso, cuando se utiliza una serie temporal no parece razonable adoptar el supuesto de que el valor de la variable estudiada en un período de tiempo es independiente de los valores que ha tomado dicha variable en los períodos que le han precedido. La naturaleza de esta dependencia entre las observaciones de una serie temporal tiene un considerable interés práctico y, precisamente, el análisis de series temporales se centra en las técnicas para estudiar esta dependencia. Se hace necesaria, por tanto, la utilización de la teoría de procesos estocásticos. Puesto que el trabajo se enmarca dentro de las series temporales estadísticas, se examinarán las nociones básicas de los procesos estocásticos en los epígrafes siguientes.

1.4. Procesos Estocásticos

Un proceso estocástico se define como una familia de variables aleatorias que corresponden a momentos sucesivos del tiempo. Será designado por $Y(t, u)$, donde t es el tiempo y u es la variable aleatoria.

La determinación de las características de un proceso estocástico puede hacerse mediante dos formas alternativas, bien a partir de funciones de distribución conjunta o bien a partir de los momentos.

En general, para un conjunto finito de valores del tiempo, se obtiene una función de distribución conjunta.

$$F[Y(t_1), Y(t_2), \dots, Y(t_n)] \quad (1.1)$$

Se dice que un proceso estocástico está perfectamente caracterizado cuando se determinan las funciones de distribución conjunta para cada conjunto finito de variables del proceso, es decir, para cada valor finito de n en (1.1). La determinación de las características del proceso a partir de las funciones de distribución es en general un procedimiento complicado, por lo que se acostumbra a utilizar preferentemente el método de los momentos. En una distribución de probabilidad se pueden calcular momentos de diverso orden. En un proceso estocástico al que se designará, para simplificar la notación, por Y_t , la media o momento de primer orden se define de la siguiente forma:

$$\mu_t = E(Y_t) \quad (1.2)$$

Como momento de segundo orden respecto a la media es preciso considerar, además de la varianza, las covarianzas entre variables referidas a distintos momentos de tiempo o autocovarianzas a las que se designarán por

$$\gamma_{t,s} = Cov(Y_t, Y_s) = E(Y_t - \mu_t)(Y_s - \mu_s) \quad (1.3)$$

Cuando $s = t$ se tiene definida la varianza.

Como forma alternativa de caracterización de un proceso estocástico se utilizan los coeficientes de autocorrelación.

$$\rho_{t,s} = \frac{Con(Y_t, Y_s)}{\sqrt{Var(Y_t)Var(Y_s)}} \quad (1.4)$$

1.5. Procesos Estacionarios y Ergódicos

Para definir la estacionariedad de un proceso se puede utilizar, como se hizo anteriormente para su caracterización, bien las funciones de distribución o los momentos.

Se dice que un proceso estocástico es estacionario en sentido estricto cuando al realizar un mismo desplazamiento en el tiempo de todas las variables de

cualquier distribución conjunta finita, resulta que esta distribución no varía. Considerando la función de distribución conjunta

$$F(T_{t_1}, Y_{t_2}, \dots, Y_{T_k}) \quad (1.5)$$

si se adopta el supuesto de que todos los elementos de la anterior distribución se desplazan m períodos, la nueva función de distribución conjunta será

$$F(T_{t_1+m}, Y_{t_2+m}, \dots, Y_{t_k+m}) \quad (1.6)$$

Si el proceso es estacionario en sentido estricto se deberá verificar que

$$F(T_{t_1}, Y_{t_2}, \dots, Y_{T_k}) = F(T_{t_1+m}, Y_{t_2+m}, \dots, Y_{t_k+m}) \quad (1.7)$$

También es este caso es más complejo el análisis de la estacionariedad a partir de las funciones de distribución que si se efectúa a partir de los momentos. Se dice que un proceso es estacionario de primer orden, o en media, si se verifica que:

$$E(Y_t) = \mu, \quad \forall t \quad (1.8)$$

Se dice que un proceso es estacionario de segundo orden (o en sentido amplio) cuando se verifican las dos condiciones siguientes:

1. La varianza es finita y permanece constante a lo largo del tiempo, es decir:

$$E(Y_t - \mu)^2 = \sigma^2 < \infty \quad \forall t \quad (1.9)$$

2. La autocovarianza entre dos períodos distintos de tiempo únicamente viene afectada por el lapso de tiempo transcurrido entre esos dos períodos. Así:

$$E(Y_{t+k} - \mu)(Y_t - \mu) = \gamma_k \quad \forall t \quad (1.10)$$

Si un proceso es estacionario en sentido estricto también será estacionario en sentido amplio, pero no ocurrirá necesariamente el caso recíproco, pues el proceso puede no ser estacionario para momentos superiores al segundo. También aquí, si el proceso es estacionario en sentido amplio, y además es *normal*, se verifica que el proceso es estacionario en sentido estricto.

En un proceso estacionario las autocorrelaciones quedan definidas de la siguiente forma:

$$\rho_k = \frac{\gamma_k}{\gamma_0} \quad k \geq 0 \quad (1.11)$$

De acuerdo con (1.10), se verificará que $\gamma_k = \gamma_{-k}$, y por consiguiente $\rho_k = \rho_{-k}$. La representación gráfica de ρ_k para $k = 0, 1, 2, 3, \dots$, recibe el nombre de correlograma.

Cuando el proceso es estacionario, en principio pueden estimarse los parámetros $\mu, \gamma_0, \gamma_1, \gamma_2, \dots$ a partir de una sola realización. Disponiendo de la muestra $Y_1, Y_2, Y_3, \dots, Y_T$ se pueden utilizar los siguientes estimadores:

$$\begin{aligned}\hat{\mu} &= \frac{1}{T} \sum_{t=1}^T Y_t \\ \hat{\gamma}_0 &= \frac{1}{T} \sum_{t=1}^T (Y_t - \hat{\mu})^2 \\ \hat{\gamma}_k &= \frac{1}{T} \sum_{t=1}^{T-k} (Y_{t+k} - \hat{\mu})(Y_t - \hat{\mu})\end{aligned}\tag{1.12}$$

Además de la estacionariedad, es necesario que el proceso goce de la propiedad de ergodicidad, con objeto de que el proceso de inferencia pueda realizarse de una forma adecuada.

Cuando valores de la serie temporal alejados en el tiempo están muy correlacionados, es decir, cuando ρ_k se mantiene en unas cotas elevadas para un k grande, sucederá que al aumentar el tamaño de la muestra se añade poca información nueva. La consecuencia de este hecho en el plano estadístico será que los estimadores obtenidos no son consistentes, ya que el aumento del tamaño de la muestra no tendrá una especial utilidad, puesto que se tendrá que calcular un mayor número de autocovarianzas para caracterizar adecuadamente el proceso.

Cuando se verifica la propiedad de ergodicidad las fórmulas (1.13) permiten obtener estimadores consistentes. Una condición necesaria, aunque no suficiente, de la ergodicidad es que $\lim_{K \rightarrow \infty} \rho_K = 0$. En lo sucesivo, se supondrá siempre que el proceso es ergódico.

1.6. Procesos Lineales

Estos procesos son una clase especial de los procesos estacionarios y ergódicos, se caracterizan porque pueden representar como una combinación lineal de variables aleatorias. Se analizarán algunos de los procesos lineales más importantes.

El proceso puramente aleatorio es el más simple de todos. Podría expresarse de la siguiente forma:

$$Y_t = \epsilon_t\tag{1.13}$$

$$\begin{aligned}E[\epsilon_t] &= 0 \quad \forall t \\ E[\epsilon_t^2] &= \sigma^2 \quad \forall t \\ E[\epsilon_t, \epsilon_{t'}] &= 0 \quad t \neq t'\end{aligned}\tag{1.14}$$

En el tratamiento de series temporales, se suele designar a un proceso puramente aleatorio con la denominación de “ruido blanco”.

De ahora en adelante, por ϵ_t designaremos únicamente a una variable aleatoria que tiene las propiedades enunciadas en (1.15). El “ruido blanco” es una pieza clave en la construcción de procesos estocásticos complicados, a pesar, de estar alejado de la concepción intuitiva que hemos dado de proceso estocástico, como pueden ser los procesos autorregresivos y los procesos de medias móviles, que se examinarán a continuación.

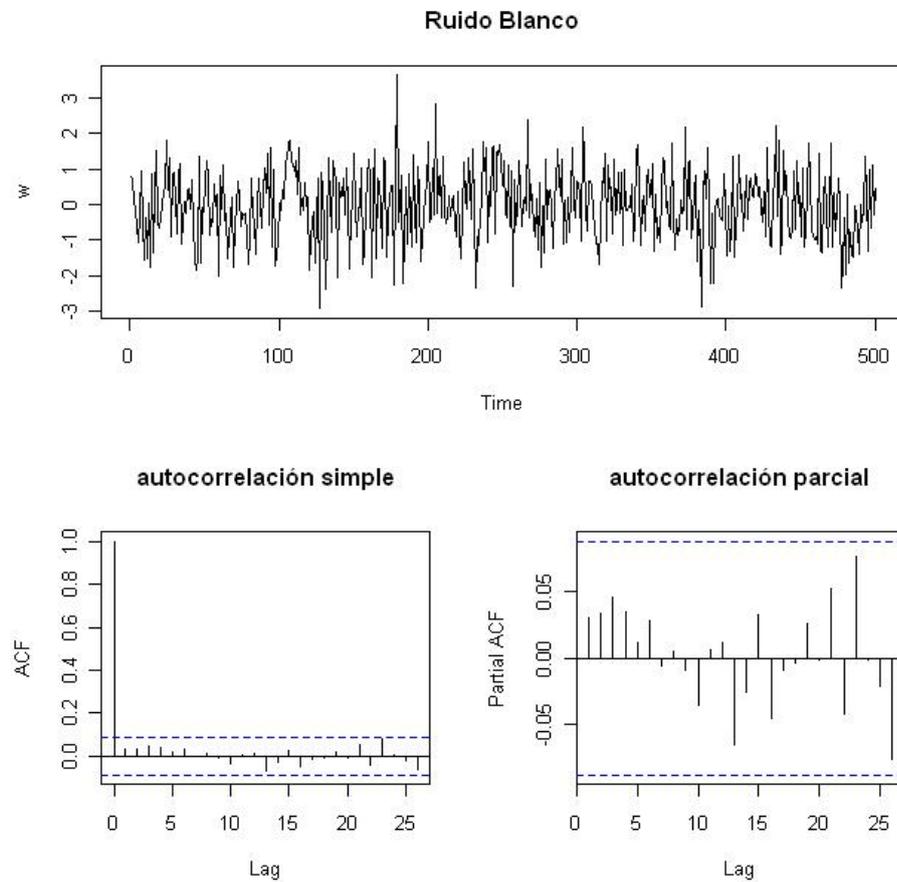


Figura 1.1: Representación gráfica de un ruido blanco

El proceso autorregresivo de orden p , $AR(p)$, se expresa de la siguiente manera:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t \quad (1.15)$$

Como puede verse en un proceso $AR(p)$, aparece un “ruido blanco” referido al momento actual y la variable desfasada para distintos períodos, siendo p el

retardo máximo que aparece en el proceso. La denominación de autorregresivo procede de Y_t se obtiene mediante regresión sobre los valores desfasados de la propia variable. Los procesos autorregresivos fueron introducidos por primera vez por Yule (1927), Yule (1927).

Un proceso de medias móviles de orden q , $MA(q)$ ² viene dado por:

$$Y_t = \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} - \dots - \theta_q\epsilon_{t-q} \quad (1.16)$$

La expresión de medias móviles hace referencia a que la variable Y_t se obtiene como un promedio de variables de “ruido blanco” (en este caso $q+1$) siendo las θ_i los coeficientes de ponderación. La introducción de estos procesos también se debe a Yule (1921) Yule (1926).

Finalmente, mediante una combinación de un proceso autorregresivo y un proceso de medias móviles se obtiene un proceso $ARMA(p, q)$, donde p indica el retardo máximo de la parte autorregresiva y q señala el correspondiente a la parte de las medias móviles. La divulgación y popularización de este tipo de procesos se debe fundamentalmente a la obra de Box (1976) aunque fueron estudiados anteriormente por Wold (1938) y Bartlett (1946). La expresión de un proceso $ARMA(p, q)$ es la siguiente:

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t - \theta_1 \epsilon_{t-1} - \theta_2 \epsilon_{t-2} - \dots - \theta_q \epsilon_{t-q} \quad (1.17)$$

Procede ahora justificar el porqué se ha centrado el estudio en ciertos tipos de procesos. En el año 1938 el profesor Wold enunció y demostró el siguiente teorema:

Cualquier proceso estacionario Y_t puede representarse unívocamente como la suma de dos procesos mutuamente incorrelados.

$$Y_t = D_t + X_t \quad (1.18)$$

donde D_t es linealmente determinista y X_t es un proceso $MA(\infty)$.

Conviene recalcar que la representación (1.18) se aplica a cualquier proceso estacionario, sea o no lineal.

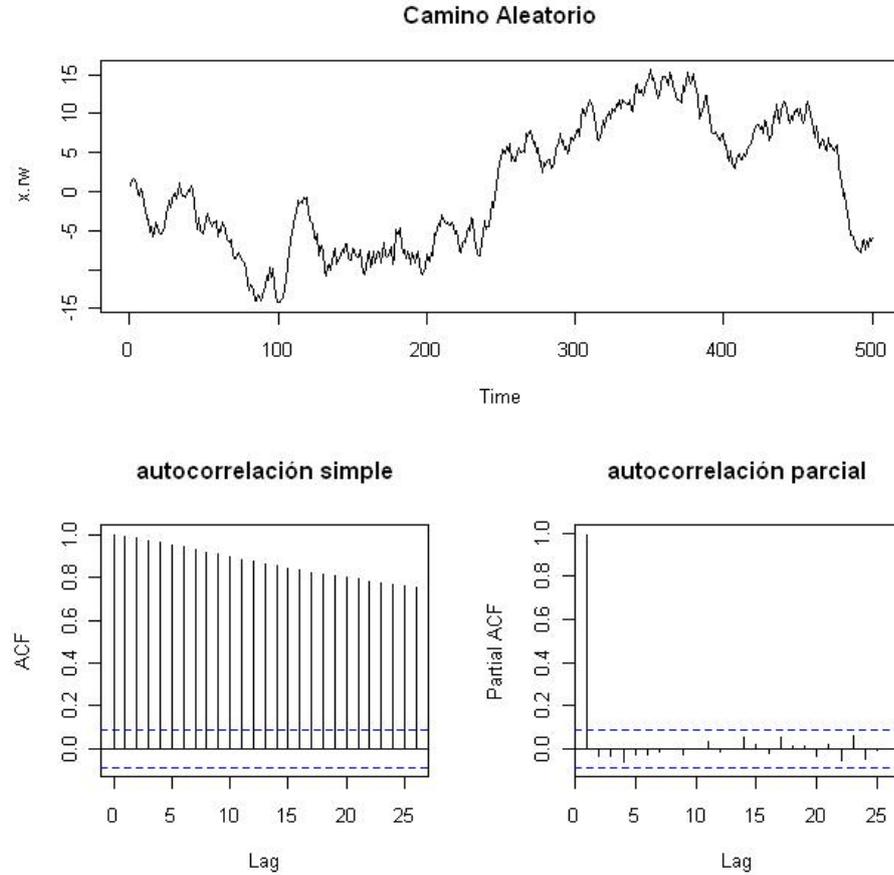


Figura 1.2: Representación gráfica de un camino aleatorio

1.7. Modelos Lineales

1.7.1. Modelos Autorregresivos, AR(p)

Un modelo $AR(p)$ viene definido por

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \epsilon_t \quad (1.19)$$

Utilizando el operador polinomial de retardos

$$\phi(L) = 1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p \quad (1.20)$$

el modelo puede expresarse de forma compacta $\phi(L)Y_t = \epsilon_t$. Para que el proceso sea estacionario se requiere que las raíces de la ecuación polinomial

$$1 - \phi_1 L - \phi_2 L^2 - \dots - \phi_p L^p = 0 \quad (1.21)$$

estén fuera del círculo unidad.

Si en el modelo (1.19) se incluye un término independiente

$$Y_t = \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \dots + \phi_p Y_{t-p} + \delta + \epsilon_t \quad (1.22)$$

entonces bajo el supuesto de estacionariedad, tomando esperanzas en (1.22) y denominado $\mu = E(Y_t)$, $\forall t$, se tiene que

$$\mu = \phi_1 \mu + \phi_2 \mu + \dots + \phi_p \mu + \delta \quad (1.23)$$

por tanto,

$$\mu = \frac{\delta}{1 - \phi_1 - \phi_2 - \dots - \phi_p} \quad (1.24)$$

En lo sucesivo se hará el supuesto, sin pérdida de generalidad, de que $\delta = 0$. Multiplicando ambos miembros de (1.19) por $Y_{t-\tau}$ y tomando esperanzas se tiene que,

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2} + \dots + \phi_p \gamma_{\tau-p} + E(\epsilon_t Y_{t-\tau}) \quad (1.25)$$

Para $\tau = 0$, se obtiene

$$\gamma_0 = \phi_1 \gamma_1 + \phi_2 \gamma_2 + \dots + \phi_p \gamma_p + \sigma_\epsilon^2 \quad (1.26)$$

Para valores de $\tau > 0$ el resultado obtenido es el siguiente:

$$\gamma_\tau = \phi_1 \gamma_{\tau-1} + \phi_2 \gamma_{\tau-2} + \dots + \phi_p \gamma_{\tau-p}; \quad \tau > 0 \quad (1.27)$$

Dividiendo ambos miembros de (1.27) por γ_0 se obtiene la ecuación en diferencias de orden p relativa a las autocorrelaciones

$$\rho_\tau = \phi_1 \rho_{\tau-1} + \phi_2 \rho_{\tau-2} + \dots + \phi_p \rho_{\tau-p} \quad (1.28)$$

Tomando $\rho_0, \rho_1, \rho_2, \dots, \rho_{p-1}$ como condiciones iniciales determinadas a partir de los coeficientes $\phi_1, \phi_2, \dots, \phi_p$, la solución a la ecuación (1.28) permite calcular los valores de ρ_τ para $\tau \geq p$.

Particularizando (1.28) para $\tau = 1, 2, \dots, p$ se obtiene el sistema de ecuaciones de Yule-Walker:

$$\begin{cases} \rho_1 = \phi_1 + \phi_2 \rho_1 + \dots + \phi_p \rho_{p-1} \\ \rho_2 = \phi_1 \rho_1 + \phi_2 + \dots + \phi_p \rho_{p-2} \\ \dots \\ \rho_p = \phi_1 \rho_{p-1} + \phi_2 \rho_{p-2} + \dots + \phi_p \end{cases} \quad (1.29)$$

Resolviendo el sistema anterior resulta que

$$\begin{bmatrix} \phi_1 \\ \phi_2 \\ \vdots \\ \phi_p \end{bmatrix} = \begin{bmatrix} 1 & \rho_1 & \dots & \rho_{p-1} \\ \rho_1 & 1 & \dots & \rho_{p-2} \\ \vdots & \vdots & \ddots & \vdots \\ \rho_{p-1} & \rho_{p-2} & \dots & 1 \end{bmatrix}^{-1} \begin{bmatrix} \rho_1 \\ \rho_2 \\ \vdots \\ \rho_p \end{bmatrix} \quad (1.30)$$

Si en el proceso $\phi(L)Y_t = \epsilon_t$ multiplicamos ambos miembros por $\phi^{-1}(L)$ se obtiene

$$Y_t = \phi^{-1}(L)\epsilon_t = \frac{1}{\phi(L)}\epsilon_t \quad (1.31)$$

De esta forma se ha pasado de un modelo $AR(p)$ a un modelo $AR(\infty)$ según se puede ver al desarrollar el segundo miembro de (1.31).

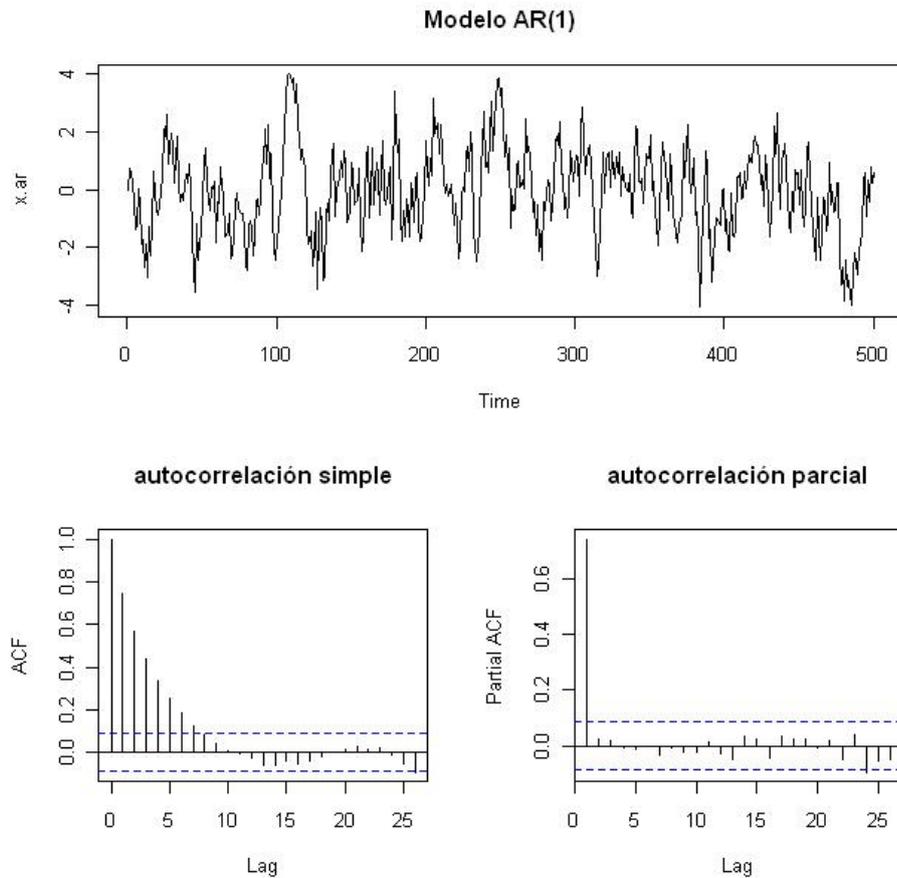


Figura 1.3: Representación gráfica de un AR(1)

1.7.2. Modelos de Medias Móviles, MA(q)

Un modelo de medias móviles de orden q , se define de la siguiente forma:

$$Y_t = \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} - \dots - \theta_q\epsilon_{t-q} \quad (1.32)$$

Conviene señalar que los coeficientes $\theta_1, \dots, \theta_q$ van precedidos por el signo negativo, por cuestión meramente de conveniencia en la notación.

Utilizando el operador polinomial de retardos,

$$\theta(L) = 1 - \theta_1L - \theta_2L^2 - \dots - \theta_qL^q \quad (1.33)$$

El modelo de medias móviles se puede expresar de forma compacta

$$Y_t = \theta(L)\epsilon_t \quad (1.34)$$

En el modelo (1.34) la media es cero, cualesquiera que sean los valores de θ_i . En efecto,

$$E(Y_t) = \theta(L)E(\epsilon_t) = 0 \quad (1.35)$$

Si en el modelo (1.32) se incluye un término constante

$$Y_t = \delta + \epsilon_t - \theta_1\epsilon_{t-1} - \theta_2\epsilon_{t-2} - \dots - \theta_q\epsilon_{t-q} \quad (1.36)$$

Entonces al tomar esperanzas en la expresión anterior resulta

$$E(Y_t) = \delta \quad (1.37)$$

Así pues, en los modelos de medias móviles, la media del proceso coincide con el término independiente, que aparece en el segundo miembro. Sin pérdida de generalidad se supondrá en lo sucesivo que $\delta = 0$.

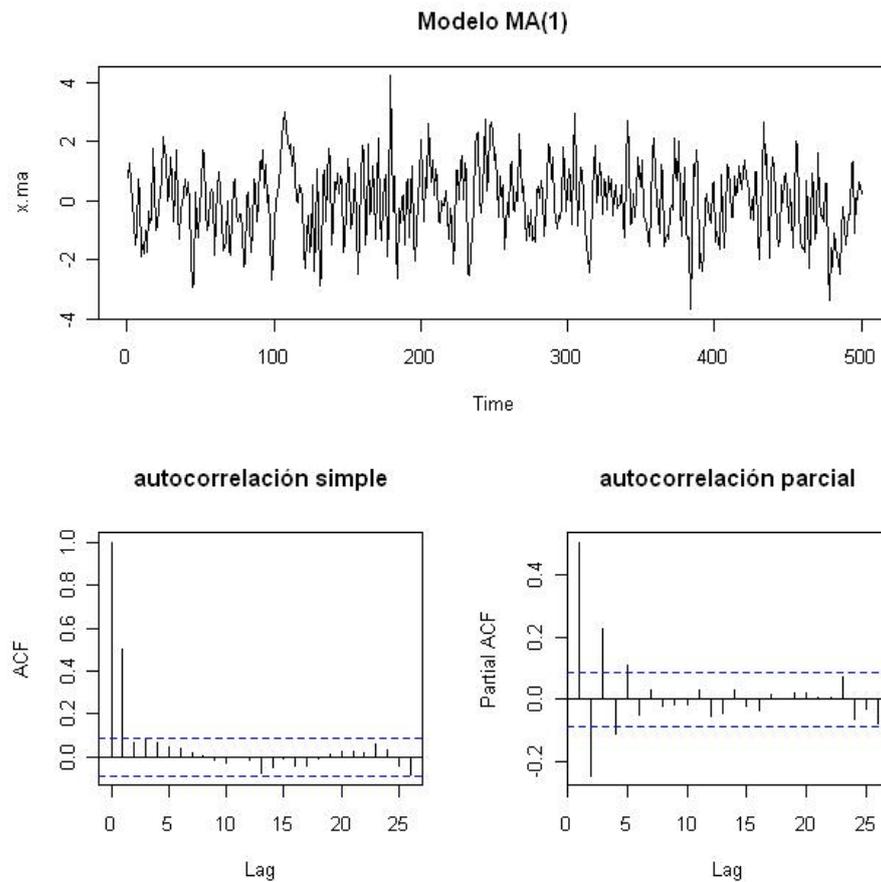


Figura 1.4: Representación gráfica de un MA(1)

A continuación se estudiarán las propiedades de un modelo $MA(q)$. Si en (1.32) se multiplican ambos miembros por $Y_{t-\tau}$ y se toman esperanzas, se obtienen los siguientes resultados

$$\gamma_0 = (1 + \theta_1^2 + \dots + \theta_q^2)\sigma_\epsilon^2 \quad (1.38)$$

$$\gamma_\tau = \begin{cases} (-\theta_\tau + \theta_1\theta_{\tau+1} + \dots + \theta_{q-\tau}\theta_q)\sigma_\epsilon^2 & \tau = 1, 2, \dots, q \\ 0 & \tau > q \end{cases} \quad (1.39)$$

Por tanto, en las autocovarianzas existe un corte brusco en q , ya que después de ese retardo su valor es igual a 0. El mismo fenómeno se presenta con los coeficientes de autocorrelación que vendrán dados por

$$\rho_\tau = \begin{cases} \frac{-\theta_\tau + \theta_1\theta_{\tau+1} + \dots + \theta_{q-\tau}\theta_q}{1 + \theta_1^2 + \dots + \theta_q^2} & \tau = 1, 2, \dots, q \\ 0 & \tau > q \end{cases} \quad (1.40)$$

En un $MA(1)$ se podría ver que $|\rho_1| \leq \frac{1}{2}$. En general, los coeficientes de autocorrelación ρ_τ de un $MA(q)$ están acotados.

Para que un modelo $MA(q)$ sea invertible se requiere que las raíces de la ecuación polinomial

$$1 - \theta_1 L - \dots - \theta_q L^q = 0 \quad (1.41)$$

caigan fuera del círculo unidad.

1.7.3. Modelos Mixtos Autorregresivos-Medias móviles, ARMA(p,q)

Un modelo $ARMA(p, q)$ viene definido de la siguiente forma:

$$Y_t - \phi_1 Y_{t-1} - \dots - \phi_p Y_{t-p} = \epsilon_t - \theta_1 \epsilon_{t-1} - \dots - \theta_q \epsilon_{t-q} \quad (1.42)$$

Utilizando los operadores polinomiales de retardo, el modelo queda expresado en forma compacta del siguiente modo:

$$\phi(L)Y_t = \theta(L)\epsilon_t \quad (1.43)$$

Para que el modelo sea estacionario se requiere que las raíces de la ecuación polinomial

$$\phi(L) = 1 - \phi_1 L - \dots - \phi_p L^p = 0 \quad (1.44)$$

caigan fuera del círculo unidad.

Si se cumplen las condiciones de estacionariedad, el modelo $ARMA(p, q)$ se puede expresar como un $MA(\infty)$, pudiendo representarse de la siguiente forma:

$$Y_t = \frac{\theta(L)}{\phi(L)}\epsilon_t = \psi(L)\epsilon_t \quad (1.45)$$

Por tanto los coeficientes del operador polinomial $\psi(L)$, que tiene infinitos elementos, deben cumplir la siguiente identidad

$$\phi(L)\psi(L) \equiv \theta(L) \quad (1.46)$$

o, en notación más detallada,

$$(1 - \phi_1 L - \dots - \phi_p L^p)(1 + \psi_1 L + \psi_2 L^2 + \dots) \equiv (1 - \theta_1 L - \dots - \theta_q L^q) \quad (1.47)$$

A partir de la identidad (1.47), se pueden deducir un conjunto de ecuaciones que nos permiten obtener los ψ_i , en función de los coeficientes autorregresivos y de medias móviles. Así de un modelo $ARMA(1, 1)$, la identidad anterior será

$$(1 - \phi_1 L)(1 + \psi_1 L + \psi_2 L^2 + \dots) = (1 - \theta_1 L) \quad (1.48)$$

Realizando la multiplicación del primer miembro y agrupando según el grado del operador de retardos, se tiene que:

$$[1 + (\psi_1 - \phi_1)L + (\psi_2 - \phi_1\psi_1)L^2 + (\psi_3 - \phi_1\psi_2)L^3 + \dots] = (1 - \theta_1 L) \quad (1.49)$$

Igualando término a término los coeficientes de las distintas potencias de L en ambos miembros, se obtiene

$$\begin{aligned} \psi_1 - \phi_1 &= -\theta_1 & \text{de donde } \psi_1 &= \phi_1 - \theta_1 \\ \psi_2 - \phi_1\psi_1 &= 0 & \text{de donde } \psi_2 &= \phi_1\psi_1 \\ \psi_3 - \phi_1\psi_2 &= 0 & \text{de donde } \psi_3 &= \phi_1\psi_2 \end{aligned} \quad (1.50)$$

Como puede verse, para $\tau > 1$ los coeficientes ψ_τ se pueden obtener de forma recursiva a partir de

$$\psi_\tau = \phi_1\psi_{\tau-1} \quad \tau > 1 \quad (1.51)$$

De forma análoga en un modelo $ARMA(p, q)$, los coeficientes $\psi_1, \psi_2, \dots, \psi_q$ se pueden obtener de ecuaciones derivadas del sistema (1.46). Los valores de ψ_τ para $\tau > q$ se pueden deducir de la siguiente ecuación

$$\psi_\tau = \phi_1\psi_{\tau-1} + \dots + \phi_p\psi_{\tau-p}; \quad \tau > q \quad (1.52)$$

Por otra parte, para el modelo (1.42) sea invertible, se requiere que las raíces de la ecuación polinomial caiga fuera del círculo unidad.

Si se cumplen las condiciones de invertibilidad, el modelo $ARMA(p, q)$ se puede expresar mediante un $AR(\infty)$

$$\epsilon_t = \frac{\phi(L)}{\theta(L)} Y_t = \pi(L) Y_t \quad (1.53)$$

donde

$$\pi(L) = 1 - \pi_1 L - \pi_2 L^2 - \dots$$

Los coeficientes del operador polinomial $\pi(L)$ deben cumplir la siguiente igualdad

$$\pi(L)\theta(L) = \phi(L) \quad (1.54)$$

Siguiendo un idéntico razonamiento al que se hizo pasar a un $MA(\infty)$, se obtiene que para $\tau > p$ los coeficientes π_τ siguen la siguiente ecuación

$$\pi_\tau = \theta_1\pi_{\tau-1} + \dots + \theta_q\pi_{\tau-q}; \quad \tau > \rho \quad (1.55)$$

En el modelo (1.42) la media es 0. Si se añade al segundo miembro un término constante δ , la media del proceso se deduce a partir de la siguiente expresión:

$$E[\phi(L)Y_t] = \delta + E[\theta(L)\epsilon_t] \quad (1.56)$$

Por tanto,

$$\phi(L)E(Y_t) = \delta \quad (1.57)$$

Si el proceso es estacionario entonces $E(Y_t) = \mu, \forall t$, con lo que

$$\mu = \frac{\delta}{1 - \phi_1L - \phi_pL^p} = \frac{\delta}{1 - \phi_1 - \dots - \phi_p} \quad (1.58)$$

ya que al aplicar el operador L a una constante (en este caso δ) se obtiene el valor constante.

Un modelo $ARMA(p, q)$ viene definido según (1.42). Si en esta expresión se multiplican ambos miembros por $Y_{t-\tau}$ y se toman esperanzas, se obtiene:

$$\gamma_\tau - \phi_1\gamma_{\tau-1} - \dots - \phi_p\gamma_{\tau-p} = E(\epsilon_t Y_{t-\tau}) - \theta_1 E(\epsilon_{t-1} Y_{t-\tau}) - \dots - \theta_q E(\epsilon_{t-q} Y_{t-\tau}) \quad (1.59)$$

En la evaluación de las esperanzas del segundo miembro es preciso tener en cuenta que

$$E(\epsilon_t Y_{t'}) = 0, \quad t' < t \quad (1.60)$$

Por tanto, se verifica que

$$\gamma_\tau - \phi_1\gamma_{\tau-1} - \dots - \phi_p\gamma_{\tau-p} = 0, \quad \tau > q \quad (1.61)$$

Esta ecuación en diferencias permite obtener las autocovarianzas a partir de $\tau > q$. Análogamente, dividiendo en (1.61) por γ_0 se obtiene

$$\rho_\tau - \phi_1\rho_{\tau-1} - \dots - \phi_p\rho_{\tau-p} = 0, \quad \tau > q \quad (1.62)$$

que es la ecuación en diferencias que permite generar los coeficientes de autocorrelación para $\tau > q$. En la determinación de los q primeros valores de ρ_τ interviene la parte de medias móviles del modelo.

En los modelos $ARMA(p, q)$ es conveniente factorizar la parte autorregresiva y la parte de medias móviles, con objeto de analizar si existen raíces comunes. Sean $\lambda_1, \lambda_2, \dots, \lambda_p$ las raíces de la ecuación en diferencias

$$\lambda^p - \phi_1\lambda^{p-1} - \dots - \phi_p = 0$$

y sean $\delta_1, \delta_2, \dots, \delta_q$ las raíces de la ecuación en diferencias

$$\delta^q - \theta_1 \delta^{q-1} - \dots - \theta_q = 0$$

Una forma alternativa, por tanto, de expresar (1.42) sería la siguiente:

$$(1 - \lambda_1 L) \dots (1 - \lambda_p L) Y_t = (1 - \delta_1 L) \dots (1 - \delta_q L) \epsilon_t \quad (1.63)$$

Si existiera alguna raíz idéntica en los dos miembros (por ejemplo, si $\lambda_1 = \delta_q$ entonces en el modelo (1.42) estaría sobrep parametrizado innecesariamente, ya que un modelo con las mismas propiedades sería un $ARMA(p-1, q-1)$. En este caso, se verificaría que

$$(1 - \lambda_2 L) \dots (1 - \lambda_p L) Y_t = (1 - \delta_1 L) \dots (1 - \delta_{q-1} L) \epsilon_t \quad (1.64)$$

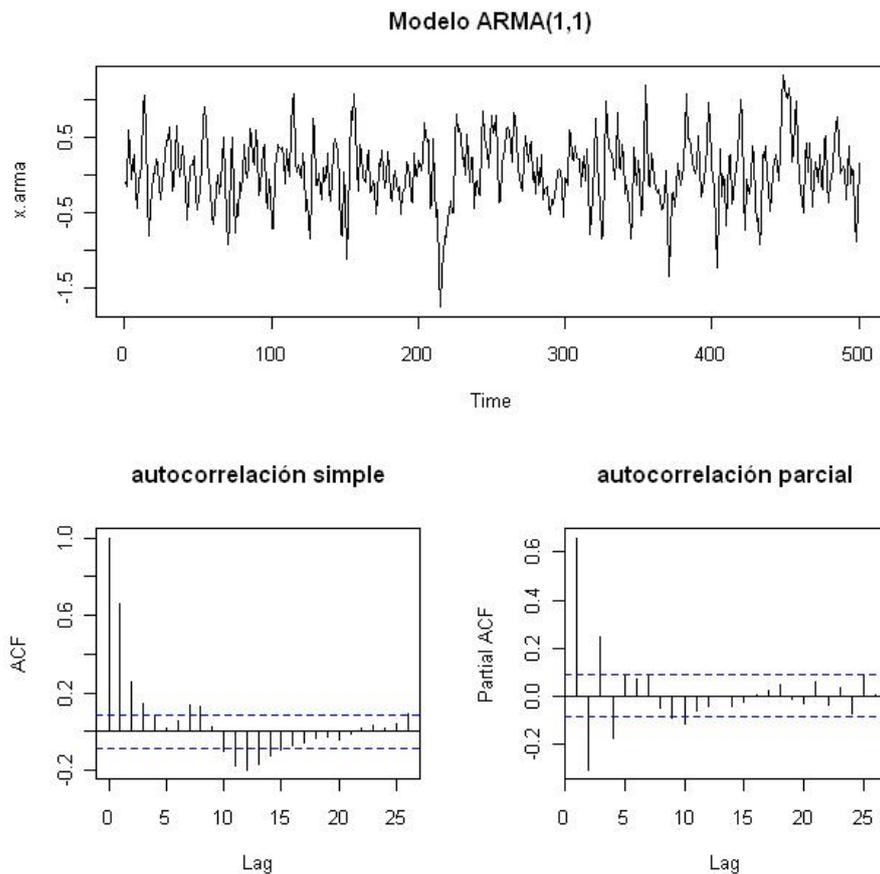


Figura 1.5: Representación gráfica de un $ARMA(1,1)$

1.8. Procesos no Estacionarios

En los procesos tratados en los epígrafes anteriores se ha impuesto las condiciones de estacionariedad y/o invertibilidad. En buena parte de las series económicas se deben considerar generadas por procesos no estacionarios. Así pues, si se desea obtener un tratamiento de las series basado en la teoría de los procesos estocásticos es necesario ampliar nuestro campo para incluir los procesos no estacionarios.

En principio cabe imaginar infinitas formas por las cuales se puede introducir la no estacionariedad en un proceso. Sin embargo, interesa considerar solamente ciertos tipos de procesos no estacionarios que sean adecuados para describir el comportamiento de series económicas, y, al mismo tiempo, sean fácilmente transformables en procesos estacionarios con objeto de utilizar las ventajas que ofrecen estos últimos. Desde esta perspectiva, dentro de los modelos no estacionarios se considerará en primer lugar y a modo de introducción el modelo:

$$Y_t = Y_{t-1} + \epsilon_t \quad (1.65)$$

El modelo anterior es un $AR(1)$ con el coeficiente $\phi_1 = 1$. A este modelo se le denomina también “paseo aleatorio”. Cuando el proceso se inicia en un pasado remoto, mediante sustituciones sucesivas se puede expresar así:

$$Y_t = \sum_{j=0}^{\infty} \epsilon_{t-j} \quad (1.66)$$

La varianza del proceso es infinita, y el proceso, en consecuencia, es no estacionario.

Haciendo de nuevo el supuesto de que el proceso se inicia en $-N + 1$, la varianza para el período t viene dada por la siguiente expresión:

$$\gamma_{0t} = (t + N)\sigma_\epsilon^2 \quad (1.67)$$

El proceso tampoco es estacionario, ya que la varianza toma un valor distinto para cada t . De manera análoga, se obtiene la autocovarianza de orden τ para el momento t .

$$\gamma_{\tau t} = (t + N - \tau)\sigma_\epsilon^2 \quad (1.68)$$

El coeficiente de autocorrelación de orden τ para el momento t será

$$\rho_{\tau t} = \frac{\gamma_{\tau t}}{\sqrt{\gamma_{0t}}\sqrt{\gamma_{0,t-\tau}}} = \frac{t - N - \tau}{\sqrt{t + N}\sqrt{t + N - \tau}} = \sqrt{\frac{t + N - \tau}{t + N}} \quad (1.69)$$

El valor que toma ρ_t depende claramente del período t de referencia. Sin embargo, para t suficientemente grande su valor se situaría muy cerca de 1. Cuando t tiende hacia infinito (o bien el período inicial se remonta hacia $-\infty$), entonces el coeficiente de autocorrelación no está definido.

Sin embargo, el proceso (1.65) puede transformarse fácilmente en estacionario. En efecto, el proceso

$$\omega_t = Y_t - Y_{t-1} = \Delta Y_t = \epsilon_t \quad (1.70)$$

que se obtiene al tomar primeras diferencias en el proceso original es ruido blanco.

Tomando diferencias de primer orden se pasa de Y_t al proceso ω_t . Consideremos ahora el caso inverso, es decir, cómo obtener Y_t a partir del proceso ω_t . Por sustituciones sucesivas se tiene que

$$Y_t = \omega_t + Y_{t-1} = \omega_t + \omega_{t-1} + Y_{t-2} = \dots = \omega_t + \omega_{t-1} + \omega_{t-2} + \omega_{t-3} + \dots \quad (1.71)$$

Por tanto, el proceso Y_t se obtiene sumando, o lo que es lo mismo, integrando el proceso ω_t .

Por esta razón, se dice que el “paseo aleatorio” pertenece a la clase de modelos integrados. Esta clase está constituida por todos aquellos modelos que se pueden transformar en estacionarios mediante la toma de diferencias, o dicho de otra forma, modelos integrados son aquellos que se pueden obtener mediante suma o integración de un proceso estacionario. A estos modelos se les denomina también modelos no estacionarios homogéneos y han sido estudiados, por Tintner et al. (1940), Rao y Tintner (1963), Yaglom (1955) y, más recientemente, por Box (1976), siendo éstos últimos autores los que más han contribuido a su divulgación.

A un proceso integrado Y_t se le denomina proceso $ARIMA(p, d, q)$ si tomando primeras diferencias d veces se obtiene un proceso estacionario ω_t del tipo $ARMA(p, q)$. La I central del término $ARIMA$ indica integrado.

Así pues, se tendría que

$$\omega_t = \Delta^d Y_t = (1 - L)^d Y_t (1 - \phi_1 L - \dots - \phi_p L^p) \omega_t = (1 - \theta_1 L - \dots - \theta_q L^q) \epsilon_t \quad (1.72)$$

o sustituyendo ω_t en la segunda expresión

$$(1 - \phi_1 L - \dots - \phi_p L^p) (1 - L)^d Y_t = (1 - \theta_1 L - \dots - \theta_q L^q) \epsilon_t \quad (1.73)$$

En forma más compacta, un proceso $ARIMA(p, d, q)$ se define así

$$\phi(L) (1 - L)^d Y_t = \theta(L) \epsilon_t \quad (1.74)$$

Los modelos $ARIMA$ constituyen una clase particular de procesos no estacionarios. Sin embargo, en muchos casos son suficientes para representar el comportamiento de las series económicas, es en éstos modelos donde vamos a centrar la investigación de esta Tesis.

Cuando una serie económica se observa a lo largo de un período dilatado de tiempo ocurre con frecuencia que también la varianza viene afectada por una tendencia, y esta tendencia no desaparece al tomar diferencias. Cuando

se presenta esta circunstancia la transformación adecuada puede consistir en tomar logaritmos. Los estadísticos Box y Cox (1964), definieron una transformación instantánea de carácter más general. Esta transformación se define por

$$Y_t^{(\lambda)} = \begin{cases} \frac{Y_t^\lambda - 1}{\lambda} & \lambda \neq 0 \\ \ln Y_t & \lambda = 0 \end{cases} \quad (1.75)$$

Cuando el parámetro es $\lambda = 1$, la transformación de Box-Cox consiste prácticamente en tomar valores originales.

Cuando el parámetro es $\lambda = 0$, sustituyendo en la primera igualdad se ve que da lugar a una indeterminación del tipo $\frac{0}{0}$. Sin embargo aplicando la regla de l'Hôpital se tiene que

$$\lim_{\lambda \rightarrow 0} Y_t^{(\lambda)} = \lim_{\lambda \rightarrow 0} \frac{Y_t^\lambda - 1}{\lambda} = \lim_{\lambda \rightarrow 0} \frac{Y_t^\lambda \ln Y_t}{1} = \ln Y_t, \quad \text{para } Y_t > 0 \quad (1.76)$$

que es, justamente, el valor propuesto para $Y_t^{(\lambda)}$ en (1.75) cuando $\lambda = 0$. Combinando una diferencia con la toma de logaritmos se obtiene la transformación $\Delta \ln Y_t$, que reviste un gran interés, ya que es adecuada para el tratamiento de muchas series económicas, si no se tiene en cuenta el problema de la estacionariedad. Por otra parte a la expresión $\Delta \ln Y_t$ se le denomina tasa de variación natural, siendo una aproximación a la tasa de variación relativa.

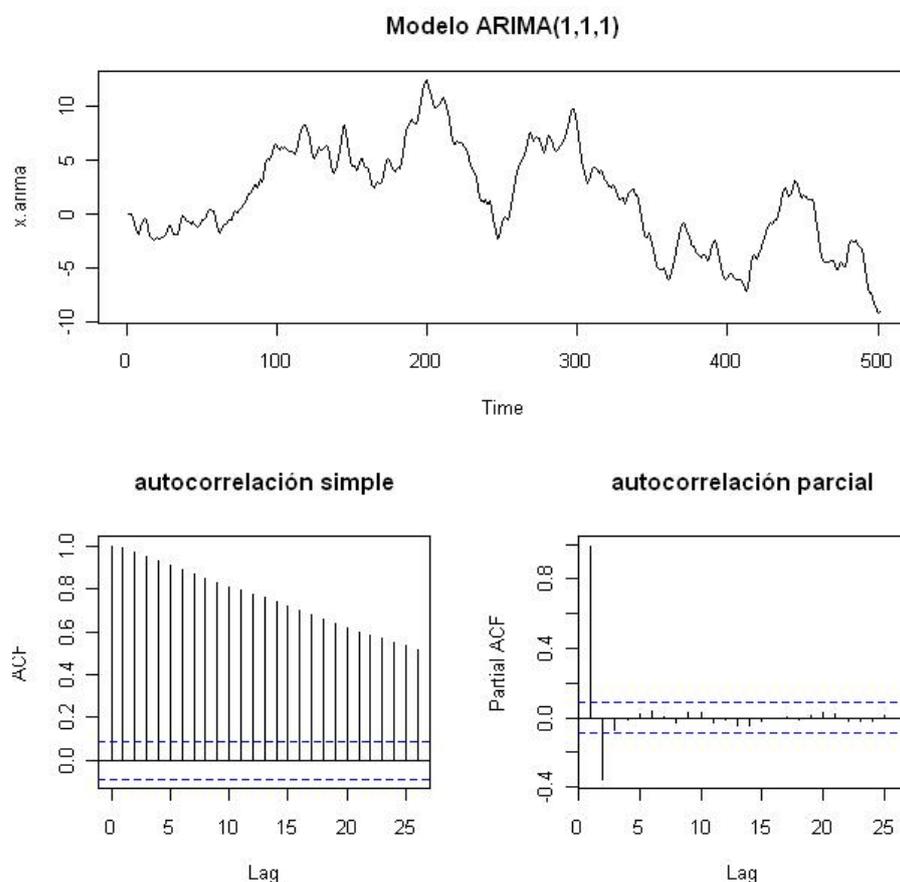


Figura 1.6: Representación gráfica de un ARIMA(1,1,1)

1.9. Procesos Estacionarios

En muchas ocasiones, los datos que se manejan en la elaboración de modelos ARIMA son datos estacionales, es decir, son datos que tienen oscilaciones estrictamente periódicas, donde el período s es igual o inferior al año. En algunas magnitudes, estas oscilaciones obedecen a causas físicas y coinciden con las estaciones del año. En otras magnitudes las oscilaciones se deben a causas sociales, económicas, etc., es decir, a factores humanos que pueden tener a su vez una motivación de orden físico.

En la elaboración de los modelos de series temporales, se debe tener en cuenta el factor estacional cuando los datos utilizados son de esta naturaleza.

Los modelos ARIMA pretenden extraer de una serie la estructura de interrelación temporal entre las variables. Por esta razón surge el operador de retardo, B . Los polinomios autorregresivos y de medias móviles pretenden recoger esa interrelación en el modelo.

En general, si existe estacionalidad, observaciones separadas entre sí s períodos están correlacionadas. En muchos casos se observa que el ciclo estacional evoluciona a lo largo del tiempo, cambiando el nivel. En este caso, tomar diferencias de orden s permitiría estabilizar el proceso alrededor de una media común, obteniéndose pues el modelo.

$$\phi_p(B^s)(1 - B^s)Y_t = \theta_Q(B^s)u_t$$

que se denomina $ARIMA(P, D, Q)_s$, donde P denota el orden del polinomio autorregresivo, D el número de diferencias estacionales tomadas y Q el orden del polinomio de media móvil.

Si la serie original no contiene más estructura de autocorrelación temporal que la estacional, u_t debería ser ruido blanco para que el modelo fuera adecuado.

Sin embargo, en muchas series existe una fuerte inercia: lo que haya pasado con la variable en momentos inmediatamente precedentes tiene un gran peso en lo que sucede actualmente. Esto significa que es de esperar que u_t, u_{t-1}, \dots , etc. estén correlacionados. Se puede proponer un modelo $ARIMA$ para recoger esta situación, obteniéndose

$$\phi_p(B)(1 - B)^d u_t = \theta_q(B)\epsilon_t$$

donde ϵ_t es ahora un ruido blanco.

Para la serie inicial Y_t , que contiene estructura regular y estacional, el modelo construido en estas dos etapas resulta ser

$$\phi_P(B^s)\phi_p(B)(1 - B^s)^D(1 - B)^d Y_t = \theta_Q(B^s)\theta_q(B)\epsilon_t$$

que se denomina $SARIMA(p, d, q)(P, D, Q)_s$.

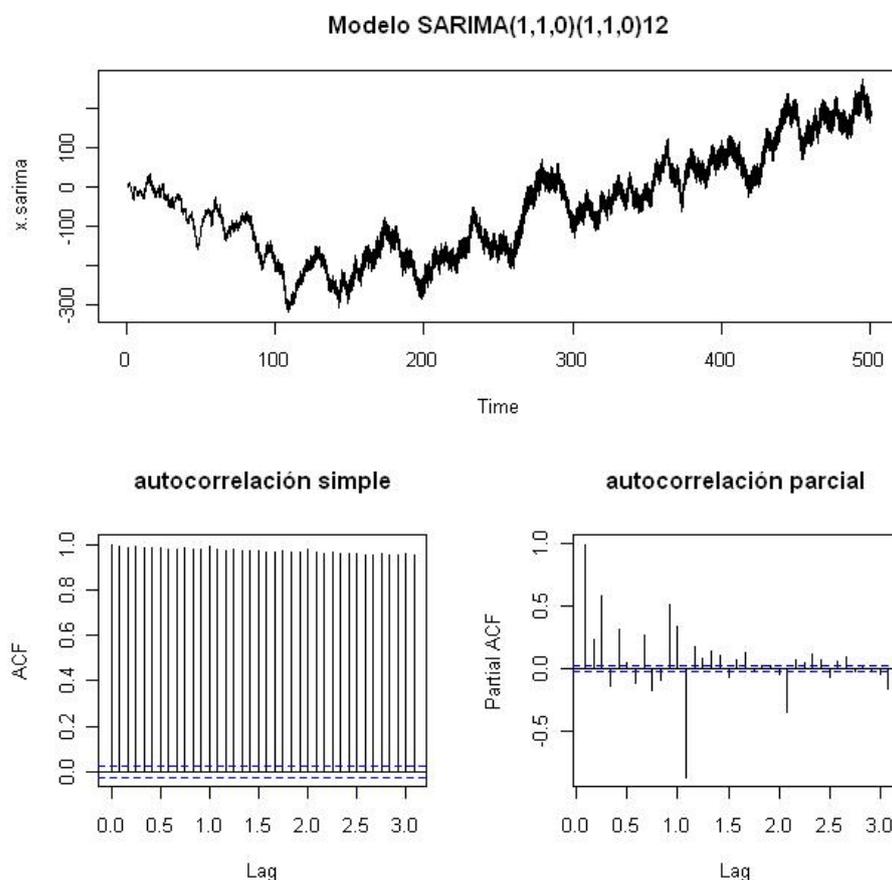


Figura 1.7: Representación gráfica de un SARIMA(1, 1, 0)(1, 1, 0)₁₂

Una vez identificado el modelo que rige una determinada serie cronológica estacional, son de aplicación los diversos métodos generales de estimación de parámetros, así como de validación y predicción.

1.10. Otras Cuestiones sobre Modelos ARIMA

Los modelos ARIMA se construyen a partir de los modelos ARMA, pero considerando que la serie en estudio para que sea estacionaria en media tendrá que diferenciarse una serie de veces. Un modelo ARIMA (p, d, q) es un modelo ARMA (p, q) sobre la serie diferenciada d veces. Habitualmente el orden de diferenciación d , entero, oscila entre 0 y 2.

Una vez hemos visto el comportamiento de un modelo ARIMA, podemos afirmar que éste se puede definir como un modelo de regresión lineal múltiple, donde la variable dependiente es la propia serie (diferenciada o no) y las variables independientes son valores de la serie y valores de los errores de

ajuste retrasados hasta unos órdenes p y q , respectivamente.

De hecho, una vez hemos identificado el modelo, que equivale a identificar los órdenes p , q y el orden de diferenciación (si es requerido), la determinación de los $p+q$ parámetros usados en el modelo se realiza de igual forma que en el caso de la regresión múltiple, es decir, mediante minimización del error cuadrático.

La gran ventaja de los modelos ARIMA con respecto a los ARMA es la incorporación de esta diferenciación dentro del modelo, dentro de la parte de integración.

A continuación se verán las características generales que ha de tener un buen modelo ARIMA.

- **Parsimonia**

Box y Jenkins ponen énfasis en que la clave de un buen modelo ARIMA es que cumpla el principio de parsimonia, que significa sencillez. Así, un modelo se dice que es parsimonioso si se ajusta a la serie de forma adecuada sin usar coeficientes innecesarios. Por ejemplo, si un modelo AR(1) y un modelo AR(2) se comportan de forma prácticamente idéntica, elegiremos el modelo AR(1) ya que así tendremos que estimar un coeficiente menos.

El principio de parsimonia es importante porque, en la práctica, un modelo parsimonioso suele generar mejores predicciones. La idea de la parsimonia nos da una fuerte orientación práctica a la hora de modelar e identificar una modelo ARIMA. Así, no tendremos que buscar el proceso ARIMA que realmente genera la serie temporal, sino que nos conformaremos con encontrar un modelo que se aproxime correctamente, tanto práctica como estadísticamente, al comportamiento de la serie temporal que estudiamos.

- **Estacionariedad**

Asumir que una serie sea estacionaria nos permite desarrollar un marco de trabajo bastante simple y usar herramientas estadísticas de muestreo de gran potencia.

Las técnicas más comunes para conocer si una serie temporal es estacionaria son los contrastes de existencia de raíces unidad.

Si una serie no es estacionaria podemos modificar dicha serie para convertirla en estacionaria. Dado que las modificaciones son conocidas, podemos posteriormente invertirlas para obtener las predicciones en la misma métrica que la serie original. Existen distintas técnicas para estabilizar la media y la varianza.

- **Buenos coeficientes estimados**

Que un modelo tenga unos buenos coeficientes estimados, está relacionado con dos vertientes distintas.

La primera es que los coeficientes, tanto los de la componente au-

torregresiva como los de la componente de media móvil sean significativamente distintos de cero. Esto se realiza mediante contrastes de hipótesis.

La segunda es que las estimaciones de los coeficientes no deben estar altamente correladas entre sí.

- **Los residuos son ruido blanco**

Esta proposición es muy importante a la hora de verificar un modelo ARIMA.

- **Debe ajustarse bien a los datos**

Que un modelo se ajuste todo lo bien posible a los datos de los que es generado, es una hipótesis asumible y lógica.

Los márgenes asumibles del valor de estos errores de ajuste depende, ciertamente, de la naturaleza de la serie, por lo que no hay un criterio unívoco de comprobación de la adecuación del ajuste.

- **Debe dar buenas predicciones**

Aunque el modelo haya sido ajustado y prediga el pasado de una forma suficientemente correcta, lo que realmente se requiere de cualquier modelo de predicción es que realice predicciones satisfactorias.

Nuestro fin a la hora de emplear un modelo ARIMA es obtener un modelo que cumpla las hipótesis que acabamos de formular.

Box y Jenkins proponen un procedimiento práctico en tres etapas para hallar el modelo ARIMA óptimo.

- **Etapas. Identificación**

En esta etapa analizamos mediante distintas técnicas cual es el modelo ARIMA que, a priori, mejor se puede ajustar a la serie. Según la metodología clásica, las herramientas que permiten identificar el patrón que sigue la serie son las funciones de autocorrelación.

- **Etapas. Estimación**

En esta etapa obtenemos las estimaciones de los parámetros del modelo ARIMA, una vez hemos fijados en la etapa de identificación los órdenes autorregresivo y de media móvil. Esta estimación se realiza mediante minimización cuadrática del error de ajuste. Esta etapa nos proporciona señales de aviso sobre si el modelo es adecuado o no. En particular, si los coeficientes no cumplen ciertas inecuaciones derivadas de la invertibilidad y la estacionariedad, el modelo ajustado debe ser rechazado.

- **Etapas. Comprobación**

Box y Jenkins proponen algunas comprobaciones de hipótesis que de-

ben ser realizadas para comprobar que el modelo estimado es estadísticamente adecuado.

Su esquema simplificado es el siguiente:

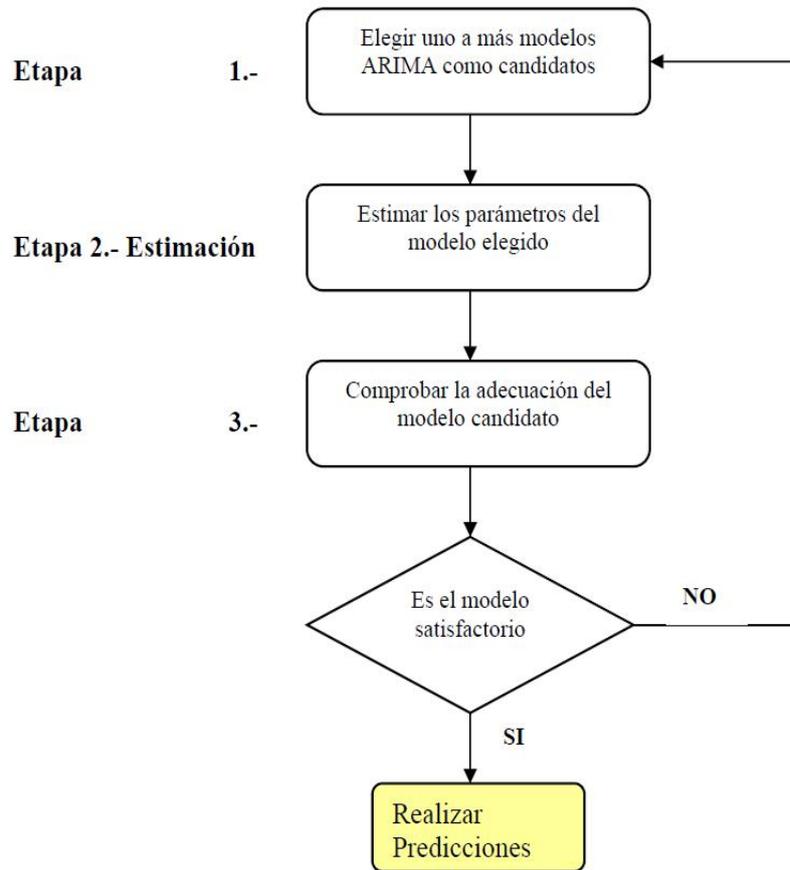


Figura 1.8: Esquema de Identificación Box-Jenkins

Capítulo 2

Criterios de Selección de Modelos

*Todo el mundo se queja de no tener
memoria y nadie se queja de no tener
criterio.*

La Rochefoucauld (Autor de epigramas y
moralista francés)

2.1. Introducción

La clase de modelos ARIMA con estacionalidad multiplicativa, ocupa un lugar central en la modelización de series temporales propuestas por Box-Jenkins Box (1976).

En la práctica, los órdenes del modelo ARIMA son desconocidos y solo se dispone de una realización $y_1, y_2 \dots, y_n$ del proceso generador de la serie para identificarlos, lo que implica que con frecuencia se encuentren varios modelos posibles. En estos casos, suelen estimarse los modelos posibles y seleccionar uno de ellos de acuerdo con algún criterio. Sin embargo, la elección de un modelo atendiendo a un único criterio puede ser a veces, de dudoso valor. Éste es un problema que ha suscitado considerable atención en la literatura especializada, dada la falta de un criterio objetivo último de ajuste, distintos analistas pueden, ante una misma serie, obtener modelos distintos. El problema es importante, porque dos modelos pueden ajustarse prácticamente igual a un conjunto de datos y conducir, sin embargo, a funciones de previsión esencialmente distintas.

Por ello, uno de los objetivos de esta tesis consiste en realizar simulaciones de distintos procesos generadores, para los cuales se estiman distintos modelos ARIMA y se calculan varios estadísticos correspondientes a diferentes criterios de selección, con la idea de comprobar qué criterio o criterios funcionan mejor. Además, como el principal objetivo del análisis de una serie temporal

es la predicción, es decir, la extrapolación de valores futuros a partir de la información disponible, que en este caso se trata de los valores pasados de la serie, otro de los objetivos que se persigue en este trabajo es buscar un criterio de selección que se base, no en el mejor ajuste, sino en obtener la mejor predicción.

2.2. Selección de Modelos ARIMA

Un punto clave en el análisis de una serie temporal a la que se quiere ajustar un modelo ARIMA es determinar sus órdenes. Generalmente, cuantas más variables se añaden a un modelo, mejor es el ajuste aparente a los datos observados. Uno de los objetivos de la selección de modelos es hallar un equilibrio entre la bondad de ajuste y la complejidad del modelo. No obstante, si el objetivo es la predicción, como es el caso del análisis de una serie temporal, problema que nos aborda, la calidad de un modelo debe juzgarse más por su capacidad predictiva que por su ajuste a los datos reales.

Se pueden clasificar los criterios de selección en dos grupos: los criterios basados en bondad de ajuste; son aquellos que seleccionan el modelo que más se aproxime al modelo verdadero y los criterios basados en medidas de predicción; son aquellos que seleccionan el modelo que predice mejor fuera de la muestra.

Supongamos que para una serie temporal $S = \{y_t; t = 1, 2, \dots, n + h\}$, se desea realizar predicciones para el intervalo $t = n + 1, n + 2, \dots, n + h$. La estimación del modelo ARIMA se lleva a cabo con el subconjunto de datos $S_n = \{y_1, y_2, \dots, y_n\}$, lo que proporciona un subconjunto de n valores estimados $P_n = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$ y otro de h predicciones $P_h = \{\hat{y}_{n+1}, \hat{y}_{n+2}, \dots, \hat{y}_{n+h}\}$ quedando el conjunto $S_h = \{y_{n+1}, y_{n+2}, \dots, y_{n+h}\}$ de datos futuros.

2.3. Criterios de Bondad de Ajuste

En este caso, los criterios son funciones de los conjuntos S_n y P_n , ya que el objetivo de estos criterios como se ha descrito anteriormente es buscar el modelo de un conjunto, M , de M modelos $\text{ARIMA}(p, d, q) \times (P, D, Q)_s$, entre los que se supone que existe un modelo “óptimo”, o un número reducido de “mejores” modelos, en cuánto a “óptimo ajuste”, basándonos en la función de verosimilitud de la serie.

2.3.1. Criterios basados en el error cuadrático medio

El primer criterio de selección que se conoce fue introducido por Akaike (1969) y se denomina error de predicción final (FPE). Dada una serie $S_n = \{y_1, y_2, \dots, y_n\}$, se quiere ajustar modelos autorregresivos, utilizando

un conjunto, M , buscando aquel modelo que minimize el error de predicción.

$$(FPE)_M = \left(\frac{n + M + 1}{n - M - 1} \right) R_M \quad m = 1, 2, \dots, L \quad (2.1)$$

donde R_M es la solución obtenida al aplicar el método de mínimos cuadrados sobre la media de cuadrados residuales. Este criterio, por su definición, tiende a seleccionar modelos con muchos parámetros. Soderstrom (1977) demuestra que este criterio se puede aplicar a otros modelos de series temporales como los modelos ARMA reemplazando $M + 1$ por el número de parámetros del modelo. Al igual que para los modelos ARMA. Bhansali y Downham (1977) sugieren el uso del criterio (FPE) generalizado, dado por:

$$(FPE)_{M\alpha} = \left(1 + \frac{\alpha(M + 1)}{n} \right) R_M \quad (2.2)$$

donde α es una constante positiva (generalmente $\alpha > 2$). Akaike (1979) analiza esta generalización detallando algunos inconvenientes de la misma. En el caso de los modelos ARIMA presenta un grave inconveniente, y es que este estadístico únicamente tiene validez cuando se comparan modelos en que se haya tomado idéntico número de diferencias.

2.3.2. Criterios basados en la información de Kullback-Leibler

Un criterio diferente de gran utilización en la selección de modelos es el propuesto por Akaike (1973) basado en la teoría de la información. Sea una serie temporal $S_n = \{y_1, y_2, \dots, y_n\}$ que ha sido generada por un función de densidad, $f(y|\alpha_0)$, donde α_0 es el vector de parámetros verdaderos. Supongamos un conjunto de modelos ARMA(p, q) que denotamos por $M_{p,q}$ con parámetros $\alpha_{p,q}$ donde $p = 0, 1, \dots, L_p$ y $q = 0, 1, \dots, l_q$, siendo L_p, L_q dos cotas positivas. La información de Kullback-Leibler para dos funciones de densidad $f(\cdot|\alpha_0)$ y $f(\cdot|\alpha_{p,q})$ viene dada por

$$I(f(\cdot|\alpha_0), f(\cdot|\alpha_{p,q})) = \int f(x|\alpha_0) \log \left(\frac{f(x|\alpha_0)}{f(x|\alpha_{p,q})} \right) dx \geq 0 \quad (2.3)$$

de tal manera que dos densidades estarán cerca cuanto menor sea la información de Kullback-Leibler. Akaike propone seleccionar el modelo que proporcione el valor mínimo de (2.3). Puesto que la distribución $f(x|\alpha_0)$ es desconocida, desarrollando en (2.3) se tiene que,

$$\begin{aligned} I(f(\cdot|\alpha_0), f(\cdot|\alpha_{p,q})) &= \int f(x|\alpha_0) \log f(x|\alpha_0) dx - \int f(x|\alpha_0) \log f(x|\alpha_{p,q}) dx = \\ &= E_x(\log f(x|\alpha_0)) - E_x(\log f(x|\alpha_{p,q})) \end{aligned} \quad (2.4)$$

donde las esperanzas se toman con respecto a la verdadera distribución $f(\cdot|\alpha_0)$. Puesto que $E_x(\log f(x|\alpha_0))$ es una constante desconocida, minimizar (2.3) es equivalente a maximizar $E_x(\log f(x|\alpha_{p,q}))$, es decir, seleccionar

el modelo que da el máximo del valor promedio de $\log f(x|\alpha_{p,q})$ para todos los valores de x de la densidad verdadera $f(x|\alpha_0)$.

En la práctica, dada la serie $y = (y_1, y_2, \dots, y_n)'$, la idea más natural es estimar (2.3) mediante:

$$I(f(\cdot|\alpha_0), f(\cdot|\hat{\alpha}_{p,q})) = \int f(x|\alpha_0) \log \left(\frac{f(x|\alpha_0)}{f(x|\hat{\alpha}_{p,q})} \right) dx \quad (2.5)$$

donde $\hat{\alpha}_{p,q}$ es el estimador de máxima verosimilitud basado en la serie y . En lugar de esto, Akaike propone obtener el estimador $\hat{\alpha}_{p,q}$ que minimice el valor esperado:

$$E_y(I(f(\cdot|\alpha_0), f(\cdot|\hat{\alpha}_{p,q}))) = E_y \left(E_x \left(\log \left(\frac{f(x|\alpha_0)}{f(x|\hat{\alpha}_{p,q})} \right) \right) \right) \quad (2.6)$$

Desarrollando (2.6), se tiene que:

$$E_y(I(f(\cdot|\alpha_0), f(\cdot|\hat{\alpha}_{p,q}))) = E_y(E_x(\log f(x|\alpha_0))) - E_y(E_x(\log f(x|\hat{\alpha}_{p,q}))) \quad (2.7)$$

Siendo el primer término de (2.7) un valor constante. Por lo tanto, minimizar (2.6) es equivalente a maximizar $E_y(E_x(\log f(x|\hat{\alpha}_{p,q})))$. Akaike deriva la siguiente aproximación (ver Akaike (1973)) y proporciona el conocido criterio de Akaike:

$$E_y(E_x(\log f(x|\hat{\alpha}_{p,q}))) \simeq -\frac{n}{2} \log \hat{\sigma}_{p,q}^2 - (p + q + 1) \quad (2.8)$$

$$AIC(p, q) = n \log \hat{\sigma}_{p,q}^2 + 2(p + q + 1) \quad (2.9)$$

El AIC establece como mejor modelo aquel para el que el estadístico tome el valor más bajo, que es el que encuentra el mejor equilibrio entre el ajuste del modelo y su complejidad (medida por el número de parámetros).

El problema del AIC es que tiende a sobreestimar el número de parámetros en el modelo y este efecto puede ser muy acusado en muestras pequeñas.

Son muchos los autores que han estudiado y propuesto derivaciones del criterio AIC bajo hipótesis alternativas, como por ejemplo, Sugiura (1978), Ogata (1980), Chow (1981), Findley (1985), Bozdogan (1987), Shibata (1989) o Cavanaugh y Shumway (1997) entre otros.

Akaike (1974) aplica el criterio AIC para determinar el orden de ajuste de modelos autorregresivos, que constituye una generalización del criterio (FPE).

$$(FPE)_p = \left(\frac{n + p + 1}{n - p + 1} \right) \hat{\sigma}_p^2 = \left(1 + \frac{2(p + 1)}{n} \right) \hat{\sigma}_p^2 + O(n^{-2}) \quad (2.10)$$

y tomando logaritmos tenemos que asintóticamente, el (FPE) y el AIC son equivalentes.

Shibata (1976) obtiene la distribución asintótica del orden seleccionado por

el criterio de Akaike para modelos autorregresivos y demuestra que dicha distribución depende de los órdenes máximo y mínimo ajustados. A partir de dicha distribución se puede concluir que si el modelo real de la serie se encuentra entre los modelos candidatos, incluso asintóticamente el criterio de Akaike no tiene porque seleccionar el orden verdadero. Hannan (1980) generaliza este resultado para modelos ARMA. Shibata (1980) demuestra que los criterios AIC y (FPE) son asintóticamente eficientes para procesos autorregresivos infinitos, en el sentido en que los órdenes seleccionados alcanzan asintóticamente el menor error de predicción. Tsay (1984) generaliza estos resultados para el caso de modelos autorregresivos no estacionarios. Sugiura (1978) obtiene una mejor aproximación para muestras finitas de la distancia de Kullback-Leibler que se deriva en un criterio AICc y que para modelos ARMA(p,q) viene dado por,

$$AICc(p, q) = \log \hat{\sigma}_{p,q}^2 + \frac{2n(p+q+1)}{n-p-q-2} \quad (2.11)$$

Hurvich y Tsai (1989) derivan el criterio de Sugiura para el caso particular de regresión lineal y modelos autorregresivos, demuestra que mejora el comportamiento del AIC en muestras finitas y que es un criterio asintóticamente eficiente. Hurvich et al. (1990) y Hurvich y Tsai (1991) estudian el sesgo de los criterios AIC y AICc para estimar la información de Kullback-Leibler y proponen un refinamiento basado en un método de Monte Carlo.

Konishi y Kitagawa (1996) derivan un criterio similar al de Akaike y que denotan por GIC, en el sentido de que es un estimador de la distancia de Kullback-Liebler pero suavizando la hipótesis de partida de Akaike. Es decir, no es necesaria la estimación máximo verosímil y la familia de distribuciones paramétricas puede o no contener al verdadero modelo. El criterio obtenido es bastante complicado y depende de funcionales, pero es válido para estimadores de la densidad robustos, Bayesianos y Bootstrap.

McQuarrie et al. (1997) introducen una corrección para modelos de regresión lineal al criterio AICc que denominan AICu que aparentemente funciona mejor en pequeñas muestras en el sentido de que selecciona el verdadero modelo más veces que otros criterios y que está dado por:

$$AICu(p, q) = \log \hat{\sigma}_{p,q}^2 + \frac{2n(p+q+1)}{n-p-q-2} + \log \frac{n}{n-p-q-2} \quad (2.12)$$

El inconveniente es que el criterio resultante no es ni consistente, ni eficiente, por lo que no posee propiedades asintóticas conocidas.

Shi y Tsai (1998), Shi y Tsai (2002) derivan versiones modificadas robustas de los criterios AIC y AICc y proponen un criterio que denominan RIC basada en la verosimilitud de los residuos y que estudia el comportamiento residual después del ajuste de cada modelo candidato.

Peña Sánchez de Rivera y Arnáiz Tovar (1981) realiza una adaptación del

AIC para modelos ARIMA, donde se prueba además, que este criterio equivale a un contraste F para probar si la reducción de varianza aportada por la introducción de nuevos parámetros es significativa.

$$AIC(p, d, q) = \ln \hat{\sigma}_a^2 + n + 2K \quad (2.13)$$

Donde K representa el número de parámetros estimados que será, para un modelo estacional, igual a la suma de los órdenes de los operadores más uno (debido a la estimación de la varianza residual $\hat{\sigma}_a^2$), más uno (únicamente cuando la media de la serie estacionaria sea distinta de cero).

El enfoque de Akaike es atractivo porque amplía la teoría de estimación clásica de Fisher-Neyman y Pearson, donde la forma del modelo se supone conocida a priori, a situaciones más generales donde la identificación de la estructura del modelo puede resolverse como un problema de estimación. Akaike (1976), ha utilizado este criterio para investigar representaciones canónicas multivariantes para series temporales y ha presentado sus implicaciones en el análisis bayesiano, en especial en el establecimiento de distribuciones a priori Akaike (1978a), y en la selección bayesiana del orden de un proceso autorregresivo Akaike (1978b).

Dado que el establecimiento de este criterio se basa en un procedimiento heurístico su “racionalidad” para la selección de modelos depende de su eficacia empírica. Ozaki (1977) ha aplicado este criterio a modelos ARIMA(p,d,q), utilizando la siguiente expresión.

$$AIC = n \log \hat{\sigma}_a^2 + \frac{n}{n-d} 2(p+q+1 + \delta_{d0}) + n \log(2\pi) + n \quad (2.14)$$

donde

$$\delta_{d0} = \begin{cases} 1 & \text{si } d = 0, \\ 0 & \text{si } d \neq 0. \end{cases}$$

2.3.3. Criterios basados en la teoría Bayesiana

La hipótesis de partida de estos criterios, supone que el verdadero generador del conjunto de datos es uno de los modelos candidatos, lo que lleva de manera natural a considerar un punto de vista Bayesiano y seleccionar el modelo con la mayor probabilidad a posteriori dados los datos.

Supongamos una serie temporal $S_n = \{y_1, y_2, \dots, y_n\}$ y una serie de modelos $M_{p,q}$ de parámetros $\alpha_{p,q}$ donde $p = 0, 1, \dots, L_p$ y $q = 0, 1, \dots, l_q$, siendo L_p, L_q dos cotas positivas.

Schwarz (1978) con el objetivo de determinar el modelo con mayor probabilidad a posteriori, deriva el criterio Bayesiano (BIC) para el caso de observaciones independientes e idénticamente distribuidas y modelos lineales bajo la hipótesis de que la verosimilitud pertenece a la familia exponencial regular. Para el caso de modelos ARMA(p,q) el criterio BIC se define por:

$$BIC(p, q) = n \log \hat{\sigma}_{p,q}^2 + \log n(p+q+1) \quad (2.15)$$

El BIC es un criterio que trata de maximizar la probabilidad a posteriori del modelo dada la serie a estudiar, suponiendo que las probabilidad a priori son las mismas para todos ellos.

Es similar y alternativo al de Akaike, con la diferencia de que el BIC penaliza aún más la inclusión de nuevos términos en el modelo, por lo que funciona mejor en la práctica. Esto se debe a que éste criterio supone que los datos se han generado por un modelo con un número de parámetros finito, mientras que el AIC permite que el modelo correcto pueda no ser de orden finito. El criterio de Schwarz nunca seleccionará modelos con mayor número de parámetros que el criterio de Akaike.

La ventaja del BIC frente al AIC radica en su consistencia, bajo la hipótesis de que existe un verdadero proceso generador de los datos. No obstante, su rendimiento en muestras finitas no es necesariamente superior al AIC, véase Lütkepohl y NetLibrary (1991), Lütkepohl y NetLibrary (2005) para una discusión detallada de estos criterios.

Se han propuesto diferentes derivaciones del criterio BIC que se pueden encontrar en Chow (1981), Haughton (1988), Cavanaugh y Neath (1999). En Chow (1981) bajo el modelo $M_{p,q}$, x tiene densidad $p(x|\alpha_{p,q}, M_{p,q})$ donde $\alpha_{p,q}$ es el vector de parámetros desconocidos. El procedimiento Bayesiano asigna una distribución de probabilidad a priori a los parámetros del modelo, $p(\alpha_{p,q}|M_{p,q})$ y una probabilidad a priori para cada modelo, $p(M_{p,q})$. La probabilidad de que $M_{p,q}$ sea el modelo generador de los datos, condicionado a la serie x , es la probabilidad a posteriori del modelo. Tras eliminar constantes e ignorar las probabilidades a priori se deriva el criterio Bayesiano siguiente:

$$BIC(p, q) = n \log \hat{\sigma}_{p,q}^2 + \log n(p + q + 1) \quad (2.16)$$

Si estudiamos Haughton (1988), podemos ver un complejo desarrollo teórico que aborda el problema de, considerando una serie temporal procedente de una familia exponencial, dar un número finito de modelos M , y escoger el que mejor se ajuste a la serie dada, con algunas penalizaciones al escoger modelos con dimensiones demasiado grandes. El criterio de Schwarz se hace más específico en esta ocasión sugiriendo para ello una expansión asintótica, cuya formulación es más compleja. En Cavanaugh y Neath (1999) se puede ver también otra propuesta para el criterio BIC, donde las hipótesis de partida se relajan, es decir, no se asume ninguna especificación sobre la función de verosimilitud y lo único que se requiere es cumplir ciertas condiciones no restrictivas de regularidad.

$$SIC = -2 \ln(\hat{\theta}_n^k | S_n) + D_k \ln n \quad (2.17)$$

Donde θ_n^k es el vector de parámetros a priori del modelo M_k , D_k es la dimensión del modelo, M_k y k^{th} el modelo candidato.

Akaike (1981) realiza un análisis crítico de este criterio, argumentando que el uso de (2.15) sólo puede ser justificado tras definir las probabilidades a

priori y no omitiéndolas y Chow (1981) demuestra que la aproximación del logaritmo de la probabilidad a posteriori que conlleva el criterio BIC no es buena.

Simultáneamente al trabajo de Schwarz, H. (1977) propone maximizar la esperanza del logaritmo de la probabilidad a posteriori del modelo, y que está dado por

$$\begin{aligned} BAIC(p, q) = n \log \hat{\sigma}_{p,q}^2 - (n - p - q) \log \left(1 + \frac{p + q}{n} \right) + \\ + (p + q) \log n + (p + q) \log \left(\frac{\frac{\hat{\sigma}_{p,q}^2}{\hat{\sigma}_y^2} - 1}{p + q} \right) \end{aligned} \quad (2.18)$$

y que, asintóticamente, es similar al criterio propuesto por Schwarz. Además, Akaike justifica su uso porque no requiere probabilidades a priori para cada modelo.

También, mediante un procedimiento no Bayesiano, Rissanen (1978) deriva el criterio (2.15) basándose en el principio de minimizar el número de dígitos binarios requeridos para describir los datos observados. Hannan y Quinn (1979) buscan un criterio de selección de un modelo autorregresivo de la forma:

$$\log \hat{\sigma}_p^2 + C_n \times p \quad (2.19)$$

de manera que el criterio sea consistente y C_n decrezca tan rápido a 0 como sea posible. El criterio obtenido está dado por:

$$HQ(p) = \log \hat{\sigma}_p^2 + 2cp \frac{\log(\log n)}{n} \quad c > 1 \quad (2.20)$$

pero este criterio tiene un peor comportamiento en muestras finitas que el criterio BIC.

Hannan (1980) demuestra que el orden de un proceso ARMA es estimado consistentemente por los criterios BIC y HQ, pero no por el AIC y determina la distribución asintótica del orden seleccionado por el AIC. Estos resultados son generalizados para procesos autorregresivos no estacionarios por Tsay (1984).

2.4. Criterios de Predicción

Estos criterios se basan en medidas que calculan la capacidad predictiva del modelo en cuestión. En este caso se utilizan los conjuntos de datos actuales S_n y P_n , se reserva el conjunto S_h para evaluar la capacidad predictiva estudiando el conjunto P_h , que se ha obtenido basándose en el modelo seleccionado por el criterio en cuestión, éste modelo, no tiene que ser el modelo generador de la serie, sino el que asegure una mejor predicción en los h valores seleccionados.

2.4.1. Índice de Desigualdad de Theil

En el contexto de la teoría de la información, Theil y Scholes (1967), Theil (1972) define el concepto de entropía o información esperada de una distribución de probabilidades como el valor esperado del logaritmo de las probabilidades con signo negativo.

Este concepto, que corresponde básicamente a una medida de incertidumbre o desorden, tuvo su origen en la ciencia física, pero ha tenido aplicaciones en el campo de la economía y política. Dos ejemplos de estas aplicaciones son el Índice de Concentración Industrial de Hirschman (1945) y de Herfindahl (1950) y el Índice de Cohesión de Rice (1928).

A partir de este concepto de entropía Theil (1958) propone como índice de Theil, U_1 la siguiente expresión:

$$U_1 = U_1(S_h, P_h) = \frac{\sqrt{\frac{1}{h} \sum_{t=1}^h (Y_{n+t} - \hat{Y}_{n+t})^2}}{\sqrt{\frac{1}{h} \sum_{t=1}^h Y_{n+t}^2} + \sqrt{\frac{1}{h} \sum_{t=1}^h \hat{Y}_{n+t}^2}} \quad (2.21)$$

Este índice varía entre 0 y 1 correspondiendo 0 a la previsión perfecta y 1 a la predicción del modelo sin cambios (todas las predicciones son cero). En este contexto, el valor 1 debería corresponder al peor caso, de manera que valores inferiores a 1 siempre deberían ir asociados a modelos estimados con mejor capacidad predictiva que el modelo sin cambios. No obstante, como se puede comprobar fácilmente, el índice tomará valor inferior a 1 incluso con modelos que predicen peor que el citado modelo sin cambios. Para solventar este problema Theil (1966) propone una nueva expresión que supera el inconveniente anterior, continuando con la misma idea de medir la capacidad predictiva del modelo estimado.

$$U_2 = U_2(S_h, P_h) = \sqrt{\frac{\frac{1}{h} \sum_{t=1}^h (Y_{n+t} - \hat{Y}_{n+t})^2}{\frac{1}{h} \sum_{t=1}^h Y_{n+t}^2}} \quad (2.22)$$

de tal forma que está acotado inferiormente por cero (caso de predicción perfecta). De nuevo, el valor 1 se corresponde con el caso del modelo sin cambios, pero a diferencia de U_1 , modelos estimados con mejor capacidad predictiva que el modelo sin cambios presentará siempre un valor de U_2 inferior a 1, mientras que los de peor capacidad predictiva siempre llevarán asociado un valor de U_2 superior a 1. En su origen este índice asume que las predicciones son realizadas año a año, por lo cual los mayores errores pueden esperarse cuando las diferencias entre los valores reales de años sucesivos son mayores.

El coeficiente U se interpreta como la relación entre la raíz cuadrada del error cuadrático medio de las predicciones y la raíz cuadrada del error cuadrático medio correspondiente al modelo "naive" que supone que no habrá cambios

en el futuro. Para este modelo, $U_2 = 1$, lo que representa un límite superior, ya que si lo sobrepasa, implica que el modelo predice peor que el modelo “de paseo aleatorio”.

Este índice de Theil tiene ciertas propiedades de descomposición que la hacen particularmente atractiva, Estas propiedades se deben a la aditividad de la función de probabilidad de la información propuesta por Shannon (2001). La descomposición del Índice de Theil permite analizar la asociación existente entre el grado de desigualdad de una determinada variable y el grado de desigualdad de una serie de variables explicativas.

El numerador del índice de Theil, ECM, se puede descomponer en sesgo, varianza y covarianza de la siguiente manera,

$$ECM = \frac{1}{h} \sum_{t=1}^h (Y_{n+t} - \hat{Y}_{n+t})^2 = (\bar{\hat{y}} - y)^2 + (S_{\hat{y}} - S_y)^2 + 2(1 - r_{\hat{y}y})S_{\hat{y}}S_y$$

Y la descomposición del componente de Theil se puede utilizar como una medida de lo que puede esperarse en el poder predictivo.

$$U = U_S + U_V + U_C$$

A la componente U_S se le conoce como componente de sesgo, ya que recoge las diferencias entre las medias, la componente U_V es la componente varianza, esta componente es mayor cuanto más difieran entre si las desviaciones estándar de las predicciones y de los datos reales, y por último, la componente U_C , que es la de correlación, que recoge las diferencias entre la variación conjunta entre las predicciones y realizaciones. Si $U_C = 0$ implica que el coeficiente de correlación es 1. A medida que U_C se va acercando a 1, la predicción será más imperfecta.

El método de Theil, ha sido estudiado por muchos autores en el campo de la distribución personal del ingreso, al permitir abordar el problema sin necesidad de supuestos previos, ya que presenta claras ventajas con respecto a aquellos métodos que se ciñen a un modelo basado en el conocimiento preexistente, limitándose a la estimación del valor de ciertos parámetros de relaciones funcionales previamente determinadas, Malinvaud (1970). La aplicación del Índice de Theil, como método de descomposición, ha sido ya utilizado en algunos análisis empíricos de distribución de ingresos. (Fishlow (1972), Van Ginneken (1975), Chiswick (1976), Altimir y Pinera (1977)). Sin embargo, estas aplicaciones han sido condicionadas por las limitaciones y características de las bases de datos utilizados.

Altimir y Piñera (1979) presenta un detallado trabajo sobre una generalización formal del método de descomposición de Theil al caso de N variables.

Capítulo 3

Una Nueva Metodología para Selección de Modelos

*Aquel que duda y no investiga, se torna
no solo infeliz, sino también injusto*

Blaise Pascal (Científico, filósofo y
escritor francés).

3.1. Introducción

En este capítulo se propone un nuevo enfoque para la selección del modelo asociado a una serie temporal, la tarea principal es encontrar el mejor modelo capaz de explicar los datos reales. ¿Cuál es el mejor modelo para asociar a una serie de tiempo? Normalmente, los criterios utilizados se basan en medidas de ajuste como AIC, BIC ... pero tiene un problema: en la estimación de la varianza residual todas las observaciones se ponderan por igual y por eso sería mejor utilizar un sistema más completo de medidas para seleccionar el mejor modelo de ajuste. Por otro lado, se utilizan también criterios basados en buscar la mejor predicción como la U de Theil ... pero cometen el mismo problema ya señalado, todas las observaciones tienen la misma ponderación.

Actualmente, es posible utilizar nuevas metodologías que eliminan el problema anterior y ofrecen una información más completa. Una alternativa, que ofrecemos a esto, podría ser el uso de medidas que dan más importancia a los residuos presentes para obtener predicciones más exactas en las series de tiempo.

Nuestro objetivo en este capítulo, es mostrar unos criterios de ajuste y de predicción, donde se ponderen las observaciones de la serie temporal, de manera que, las observaciones más recientes tengan un peso superior a las observaciones más lejanas en el tiempo, es de lógica pensar que el comportamiento

de la serie futura será más semejante al comportamiento de las últimas observaciones conocidas que al comportamiento de las primeras observaciones tomadas de la misma.

En nuestro caso además, nos interesa especialmente que el modelo a seleccionar, no sea obligatoriamente el que mejor se ajuste a los datos, si no el que mejor prediga en un corto espacio de tiempo, que puede ser o no el mismo modelo de ajuste. La selección de un modelo de predicción es un proceso que lleva tiempo. En un entorno de empresa, donde los pronósticos tienen que ser actualizados de forma mensual o trimestral, para muchas series de tiempo, es una tarea importante. Es necesario disponer de un método automático que sea capaz de identificar el mejor modelo para cada serie, y por lo tanto, sea capaz de generar automáticamente las previsiones. Existen varios paquetes de software (Forecast Pro, Demetra, Autobox,...) para proporcionar una solución a la especificación de modelos de predicción, por lo general ARIMA o de suavizado exponencial, que están basados en alguna medida de bondad de ajuste, como AIC, BIC o HQ, es decir, utilizando la probabilidad asociada a cada modelo.

Un nuevo enfoque para la selección de modelos se presenta en este trabajo. Puede ser implementado en la mayoría de los paquetes estadísticos o económicos, así como en R-project.

Reschenhofer (2006), estudia muchos de los criterios de selección de modelos, los cuales se basan en el error de predicción y en penalizaciones sobre el número de parámetros. También Galeano y Peña (2004), afirma que la mayor parte de los criterios de selección de modelos para series temporales de tipo lineal, se escriben, asintóticamente de la siguiente manera:

$$\min_k n \ln \bar{s}_a^2 + kC(n, k)$$

siendo \bar{s}_a^2 la estimación de máxima verosimilitud de la varianza residual, n el tamaño de la muestra y k el número de parámetros del modelo. El último término, es una penalización asociada a modelos con mayor k . Este es el caso del AIC, BIC o HQ, estudiados en el capítulo anterior. Sin embargo, en casi todos los métodos utilizados para la selección de modelos, García Hiernaux (2006), con un tamaño de muestra pequeño, todos estos criterios producen resultados diferentes y no unificados, como se podía esperar de la dependencia de la función de verosimilitud del modelo y del número real de datos, n , que depende del número de retardos autorregresivos.

Las medidas usuales, vistas en el capítulo anterior, tienden a escoger modelos que no usan la importancia temporal de los errores. Estos no tienen en consideración que partes de la serie temporal son más o menos importantes, ya que a la hora de predecir, es de esperar que las observaciones más recientes tengan un mayor peso en la predicción de las futuras observaciones. Además, $C(n, k)$ converge a 0 cuando n aumenta, en este sentido, con series de tiempo de gran tamaño, la selección se reduce a minimizar la varianza residual.

En este capítulo, se introduce un esquema de ponderación, atribuyendo más importancia a los últimos errores y menos a los del pasado más lejano, con un parámetro de suavización.

3.2. Selección Automática de Modelos ARIMA para Predicción

El problema que abordamos es el siguiente:

Dada una serie de datos, $S = \{y_t : t = 1, 2, \dots, n\}$ se desean realizar predicciones para el intervalo $t = n + 1, n + 2, \dots, n + h$, utilizando un conjunto, M , de M modelos $ARIMA(p, d, q)x(P, D, Q)_s$, entre los que se supone que existe un modelo “óptimo”, o un número reducido de “mejores” modelos.

Se considerarán entonces los conjuntos $S_n = \{y_1, y_2, \dots, y_n\}$ de datos reales, y $P_n = \{\hat{y}_1, \hat{y}_2, \dots, \hat{y}_n\}$ de datos estimados, se desean realizar h predicciones, $P_h = \{\hat{y}_{n+1}, \hat{y}_{n+2}, \dots, \hat{y}_{n+h}\}$, quedando el conjunto $S_h = \{y_{n+1}, y_{n+2}, \dots, y_{n+h}\}$ de datos futuros.

Para cada ARIMA se consideran los errores de predicción

$$a_t = y_t - \hat{y}_t$$

siendo $t = m + 1, m + 2, \dots, n + h$, donde $m = p + d + Ps + Ds$ el número de observaciones perdidas debido a los términos autorregresivos, $m = m_{M_s, k}$. El problema no es uno de “óptimo ajuste”, para lo cual bastaría minimizar alguna medida como los criterios de ajuste estudiados en el capítulo anterior, tampoco se trata sólo de un problema de predicción para lo cual se podrían utilizar cualquiera de los criterios de predicción del capítulo 2. Se trata de obtener las “mejores” predicciones pero considerando pesos en los errores.

Es necesario especificar que se entiende por “mejores” predicciones, para lo cual se pueden obtener distintas medidas, que pueden variar en función de los objetivos, pero que todas ellas se basarán en comparar los valores futuros reales, $S_h = \{y_{n+1}, y_{n+2}, \dots, y_{n+h}\}$ con los obtenidos con el modelo i -ésimo (del espacio de modelos M), $P_{h,i} = \{\hat{y}_{n+1}, \hat{y}_{n+2}, \dots, \hat{y}_{n+h}\}$

Así pues, se define una medida, D_F , de proximidad predictiva entre ambos conjuntos,

$$D_{F,i} = d_F(S_{h,i}, P_{h,i})$$

y el problema se reduciría a minimizar $D_{F,i}$ para $i = 1, 2, \dots, M$.

Otra alternativa, sería definir estas funciones de proximidad, D , en el conjunto original de datos, S , comparándolas con los datos estimados $P_{\hat{y},i} = \{\hat{y}_{m+1}, \hat{y}_{m+2}, \dots, \hat{y}_{n+h}\}$, con el modelo i -ésimo y esperar que actuasen de forma similar en el futuro.

Los $m = m_i = p + d + Ps + Ds$ primeros datos correspondientes a retardos autorregresivos del modelo ARIMA se pierden. Este valor m depende el modelo M_i elegido. La medida de proximidad dependerá del conjunto de datos

originales, S , y del conjunto de datos estimados con el modelo M_i

$$D_i = d(S, P_{\hat{y},i})$$

y de nuevo, el problema de selección del modelo óptimo se reduciría a minimizar D_i para $i = 1, 2, \dots, M$.

Ambas funciones de proximidad deben verificar que $d_F(S_h, S_h) = 0 = d(S, S)$, y positivas, en presencia de errores de predicción, y “crecientes” con éstos.

Lógicamente en la práctica se dispone sólo del conjunto S , $S = \{y_t : t = 1, 2, \dots, n\}$, el cual se puede dividir en dos subconjuntos: uno de entrenamiento hasta la observación T , $S_T = \{y_1, y_2, \dots, y_T\}$ y los últimos h datos de la serie formarían el conjunto de comprobación, $S_h = \{y_{T+1}, y_{T+2}, \dots, y_{T+h}\}$. Los valores de T serán definidos de antemano, seleccionando razonadamente el tamaño de la serie de entrenamiento, que será la mayor parte de la serie original. El valor de h ha de ser elegido atendiendo al número de predicciones que se desean realizar, en nuestro caso el horizonte de predicción será a corto plazo, por lo que tomará valores pequeños. Es obvio, que la suma de los valores seleccionados para T y para h , no debe superar el tamaño de la serie original, es decir, $T + h = n$. Los distintos modelos Arima ajustados a la serie, S_h (uno será el correcto y los otros, con error de especificación) producirán las h predicciones, dando lugar al conjunto P_h , y usando las medidas de proximidad y de ajuste introducidas, se podrá seleccionar el modelo que produzca mejores predicciones. En el caso de las medidas de ajuste se estudiarán los conjuntos S_T con P_T , y en el caso de medidas de predicción serán los conjuntos S_h frente a P_h . Convendría comparar siempre las medidas de proximidad en el modelo correcto y en el seleccionado.

3.3. Funciones de Proximidad para Evaluar las Predicciones

En primer lugar es necesario determinar que tipo de función $d()$ y $d_F()$ van a utilizarse. Unas clásicas son los errores cuadráticos medios o con los errores absolutos medios.

$$d(S, S_{\hat{y}}) = MSE_{M_i} = \frac{1}{T - m_i} \sum_{t=m_i+1}^T a_{t,M_i}^2$$

$$d(S, S_{\hat{y}}) = MAE_{M_i} = \frac{1}{T - m_i} \sum_{t=m_i+1}^T |a_{t,M_i}|$$

o bien, si se trata de evaluar las predicciones con horizonte h :

$$d(S, S_{\hat{y}}) = MSEF_{M_i} = \frac{1}{h} \sum_{t=0}^{h-1} a_{t,M_i}^2$$

$$d(S, S_{\hat{y}}) = MAEF_{M_i} = \frac{1}{h} \sum_{t=0}^{h-1} |a_{t, M_i}|$$

siendo los errores $a_{t, M_i} = y_t - \hat{y}_{t, i}$ en cada caso.

Una forma más general de evaluar estas proximidades sería introducir una función de ponderación, w , para valorar de forma adecuada los errores de ajuste o de predicción. Este parámetro de ponderación se puede cambiar para dar más o menos importancia a los errores más recientes.

Así se tendrían las medidas de proximidad w -ponderadas (de tipo error cuadrático medio)

$$d(S, S_{\hat{y}}) = WMSE_{M_i} = \frac{1}{T - m_i} \sum_{t=m_i+1}^T w(T, t) a_{t, M_i}^2$$

$$d(S, S_{\hat{y}}) = WMSEF_{M_i} = \frac{1}{h} \sum_{t=0}^{h-1} w_F(T, t) a_{t, M_i}^2$$

De esta forma cabría atribuir mayor importancia a los errores cercanos al tiempo presente con respecto a los más alejados hacia el pasado, o valor de forma asimétrica los errores positivos o negativos, etc. En el caso de medidas predictivas, podría interesar valorar como más importantes algunos errores en el rango de predicción $\{n + 1, \dots, n + h\}$ o según el signo del error, o incluso el considerar un error agregado sobre el horizonte de predicción.

Unas funciones simples de ponderación pueden definirse en función de un parámetro, α , de valoración exponencial de errores a lo largo del tiempo.

$$d(S, S_{\hat{y}}) = WMSE_{\alpha, M_i} = \frac{1}{T - m_i} \sum_{t=m_i+1}^T \alpha^{T-t} a_{t, M_i}^2 = \frac{1}{T - m} \sum_{t=m+1}^T e_{\alpha, M_i}^2$$

$$d_{Fw}(S, S_{\hat{y}}) = WMSEF_{\alpha, M_i} = \frac{1}{h} \sum_{t=0}^{h-1} \alpha^t a_{t, M_i}^2$$

En la primera expresión, los errores se irían ponderando de forma creciente, al ir retrocediendo en el tiempo, atribuyendo errores ponderados menores a los alejados hacia el pasado; la idea es similar a la de los operadores de alisado exponencial. En la medida de proximidad hacia el futuro, se consideran menos importantes los errores cometidos en el futuro más lejano que en el más reciente.

En lugar de ponderaciones exponenciales, cabría parametrizar el problema de forma más general, usando ponderaciones polinomiales, con las que, con un número reducido de parámetros, se podrían obtener funciones de ponderación muy diversas.

3.4. Funciones de Proximidad Exponenciales

Las funciones de proximidad antes definidas

$$WMSE_{\alpha, M_i} = \frac{1}{T - m_i} \sum_{t=m_i+1}^T \alpha^{T-t} a_{t, M_i}^2 = \frac{1}{T - m} \sum_{t=m+1}^T e_{\alpha, M_i}^2$$

$$WMSEF_{\alpha, M_i} = \frac{1}{h} \sum_{t=0}^{h-1} \alpha^t a_{t, M_i}^2$$

dependen de la constante α . Este parámetro se puede cambiar para dar más o menos importancia a los errores más recientes, obviamente cuando $\alpha = 1$, $WMSE_{1, M_i} = s_{\alpha, M_i}^2$, que es la varianza residual del modelo, y, si el tamaño de la muestra tiende a infinito, la medida es similar a los criterios de información ya estudiados. Muchas series temporales son modeladas, usando $\alpha \in [0, 7, 1]$ con buenos resultados, como se podrá ver en el capítulo posterior. Dado un valor para α , el modelo óptimo se obtendría minimizando una de las medidas de proximidad anteriores

$$\min_{i=1, 2, \dots, M} WMSE_{\alpha, M_i}$$

$$\min_{i=1, 2, \dots, M} WMSEF_{\alpha, M_i}$$

Surge el problema de querer comparar modelos pero con ponderaciones distintas de α para evitar que al disminuir α las funciones de proximidad tiendan a disminuir también (no necesariamente de forma monótona).

Si se modifican las funciones anteriores introduciendo una constante que dependa de α , se tiene

$$\begin{aligned} F_{MSE, M_i}(\alpha) &= WMSE_{\alpha, M_i} = \frac{T - m}{\sum_{t=m+1}^T \alpha^{T-t}} \frac{1}{T - m} \sum_{t=m+1}^T \alpha^{T-t} a_{t, M_i}^2 = \\ &= \frac{1 - \alpha}{1 - \alpha^{T-m}} \sum_{t=m+1}^T \alpha^{T-t} a_{t, M_i}^2 \end{aligned}$$

$$F_{MSE, M_i}^F(\alpha) = WMSEF_{\alpha, M_i} = \frac{h}{\sum_{t=0}^{h-1} \alpha^t} \frac{1}{h} \sum_{t=0}^{h-1} \alpha^t a_{t, M_i}^2 = \alpha^{T+1} \frac{1 - \alpha^h}{1 - \alpha}$$

De igual forma cabe definir unas funciones análogas basadas en el error absoluto medio

$$F_{MAE, M_i}(\alpha) = WMAE_{\alpha, M_i} = \frac{1 - \alpha}{1 - \alpha^{T-m}} \sum_{t=m+1}^T \alpha^{T-t} |a_{t, M_i}|$$

$$F_{MAE, M_i}^F(\alpha) = WMAEF_{\alpha, M_i} = \alpha^{T+1} \frac{1 - \alpha^h}{1 - \alpha} \sum_{t=0}^{h-1} \alpha^t |a_{t, M_i}|$$

El obtener un modelo óptimo minimizando las funciones anteriores produce los mismos resultados que minimizando las funciones del apartado anterior.

$$\min_{i=1,2,\dots,M} F_{MSE, M_i}(\alpha) = F_{MSE}(\alpha)$$

$$\min_{i=1,2,\dots,M} F_{MSE, M_i}^F(\alpha) = F_{MSE}^F(\alpha)$$

$$\min_{i=1,2,\dots,M} WMAE_{\alpha, M_i} = F_{MAE}(\alpha)$$

$$\min_{i=1,2,\dots,M} WMAEF_{\alpha, M_i} = F_{MAE}^F(\alpha)$$

El modelo óptimo se obtendrá seleccionando el mejor valor de α ,

$$WMAE_{\alpha^*, M_i^*} = \min_{\alpha \in [0,7,1]} \min_{i=1,2,\dots,M} WMAE_{\alpha, M_i}$$

3.5. Validación

Para una serie dada y_t , el objetivo es identificar un modelo SARIMA (o varios) usando algún criterio de optimización, C , dirigido a obtener las mejores predicciones posibles. El criterio C deberá inducir un orden en el conjunto de todos los posibles modelos, por lo que sería posible seleccionar el mejor según el criterio C .

El criterio C podría ser de dos tipos:

$C = G$ criterios de bondad de ajuste

$C = U$ criterios basados en la capacidad predictiva, teniendo en cuenta los errores de predicción

Una vez que el modelo ha sido seleccionado con C , se obtienen las predicciones sobre un horizonte h .

La evaluación de los criterios empleados, C , se realiza a través de simulaciones, por ejemplo,

$$A(1, 1, 1) = \{ARIMA(1, 1, 1)/f = f_1, f_2, \dots, \theta = \theta_1, \theta_2, \dots\}$$

y simulamos N series de longitud n . Aplicando los criterios seleccionados (G y/o U), seleccionamos un modelo ARIMA(p,d,q), y obtenemos las h predicciones para compararlas con los datos reales. En algunos casos, el modelo correcto puede ser seleccionado por N_c series de N y el modelo incorrecto por $N - N_c$ series. En cada serie se calculan las predicciones, y se evalúa la capacidad predictiva.

- Usando $C = G$: para las N_c series con el modelo correcto, se calcularán las correspondientes medidas de bondad de predicción y lo mismo se hará con las $N - N_c$ series con la especificación incorrecta. Estudiando en cada caso los residuos de las mismas.
- Usando $C = U$: se procederá de igual manera.

Puede ser estudiado por lo tanto cuál es el mejor procedimiento para seleccionar un modelo adecuado que garantice unas buenas predicciones: G o U. Esto se hará en nuestro caso con diferentes medidas que definen, y proponen la mejor línea de acción (se utilizarán algunas de las medidas de ajuste y predicción estudiadas en el capítulo anterior y las medidas de ajuste y predicción arriba desarrolladas).

Capítulo 4

Simulación de Modelos

*Largo es el camino de la enseñanza por
medio de teorías; breve y eficaz por
medio de ejemplos.*

Séneca (Filósofo latino).

4.1. Introducción

Para acotar el problema vamos a tomar $p, q = 0, 1, 2$ y $d = 0, 1$, con respecto a la parte estacionaria vamos a seleccionar los mismos valores; $P, Q = 0, 1, 2$ y $D = 0, 1$. Esta elección de valores se debe a que en la práctica estos modelos suelen ser simples, pues en la mayor parte e las ocasiones los órdenes de estos polinomios son valores muy pequeños.

$$p + q \leq 5$$

$$P + Q \leq 3$$

$$d \leq 3$$

$$D \leq 1$$

El número M de modelos en el espacio de modelos sería $M = 324$ modelos, desde un ruido blanco hasta el modelo $SARIMA(2, 1, 2)(2, 1, 2)_s$ con todos sus términos. Modelos con el polinomio autorregresivo de orden 2, pueden tener dos coeficientes autorregresivos o sólo el coeficiente de mayor retardo. Es decir, que en un polinomio $\phi(B)$ o en $\theta(B)$, pueden faltar términos intermedios.

Los modelos simulados deben cumplir las condiciones de estacionariedad e invertibilidad. Con respecto a los modelos $ARIMA(p, d, q)$, esas condiciones son las siguientes:

$$\begin{array}{ll}
\phi(B) = 1 - \phi B & |\phi| < 1 \\
\theta(B) = 1 + \theta B & |\theta| < 1 \\
\phi(B) = 1 - \phi_1 B - \phi_2 B^2 & |\phi_2| < 1, \phi_2 \pm \phi_1 < 1 \\
\phi(B) = 1 - \phi_2 B^2 & |\phi_2| < 1 \\
\theta(B) = 1 - \theta_1 B - \theta_2 B^2 & |\theta_2| < 1, \theta_2 \pm \theta_1 < 1 \\
\theta(B) = 1 - \theta_2 B^2 & |\theta_2| < 1
\end{array}$$

En cuanto a los modelos con parte estacionaria, $SARIMA(p, d, q)(P, D, Q)_s$, las condiciones son similares; las raíces unitarias de los polinomios, en módulo, deben ser mayor que 1.

Para llevar a cabo esta simulación, se hace uso del software libre R-project (ver apéndice A). Se han simulado series ARIMA y SARIMA de tamaño 180 de manera aleatoria, y para cada modelo estudiado se han estudiado 200 series. Las 200 series obtenidas han sido evaluadas en cada uno de los 324 modelos considerados y se ha evaluado en cada uno de ellos algunos de los criterios estudiados más usuales; el criterio de Akaike, en dos de sus versiones (AIC, AIC2), el criterio de Schwartz, en dos versiones también (BIC, BIC2), el error cuadrático medio tanto de ajuste como de predicción (ECM, ECMpre), el criterio de la U de Theil en sus dos modalidades (U_1 , U_2) y los criterios de ajuste y predicción estudiados en el capítulo anterior (WMSE, WFMSE, WMAE, WFMAE). El procedimiento descrito para la selección del modelo “óptimo” ha sido en un primer lugar seleccionar el modelo que minimiza cada criterio en cada una de las 200 series y el mínimo de estos a su vez, en un segundo lugar, se han separado las series donde el modelo mínimo coincidía con el modelo generador y se ha estudiado el comportamiento de las predicciones en ambos grupos, y en último lugar se han estudiado las predicciones seleccionando en cada serie el modelo que minimiza el criterio aplicado, sin tener en cuenta si era o no el generador y sin establecer un modelo único para las series procedentes del mismo modelo generador, es decir, se han realizado las predicciones en las 200 series de un mismo modelo utilizando en cada caso el modelo seleccionado por el criterio empleado.

Se tiene la serie $S = \{y_1, y_2, \dots, y_{180}\}$, se han definido los conjuntos siguientes:

- Conjunto de entrenamiento: $S_T = \{y_1, y_2, \dots, y_T\}$
- Conjunto de comprobación: $S_h = \{y_{T+1}, y_{T+2}, \dots, y_{T+h}\}$

En la simulación realizada se han considerado que las observaciones de la serie se han tomado diariamente, por tanto, en el caso de la simulación de los modelos SARIMA se ha establecido $s = 7$. Para todos los modelos ARIMA y SARIMA se ha establecido el horizonte de predicción semanal, es decir, $h = 7$, de manera que el conjunto de entrenamiento esta compuesto por las observaciones y_1, y_2, \dots, y_{173} y el conjunto de comprobación quedaría, por tanto, formado por las observaciones $y_{174}, y_{175}, \dots, y_{180}$.

Se ha implementado todo el procedimiento realizado utilizando el programa ya mencionado R-project (ver apéndice B), de manera que el proceso se automatiza, de manera fácil de aplicar y sin necesidad de tener grandes conocimientos estadísticos, con el objetivo de que pueda estar al alcance de cualquier persona interesada en ello.

4.2. Procesos no estacionarios, ARIMA(p, d, q)

Para determinar el modelo en los criterios propuestos se han considerado valores de $\alpha \in [0,8,0,99]$ y en cada caso se ha determinado el valor de α que minimizaba el criterio.

Empezamos con el **modelo ARIMA(1, 1, 0)**, cuyos coeficientes son $\phi_1 = 0,7$. (**Modelo 163**)

	AIC	AIC2	ECM	BIC	BIC2
Valor	2.39642	-0.3530208	0.592735	2.475114	-0.2983395
Modelo	320	289	321	176	289

Tabla 4.1: Criterios Clásicos de ajuste para ARIMA(1,1,0).

	U_1	U_2	ECMpre
Valor	0.01007693	0.01007745	0.043910
Modelo	321	321	309

Tabla 4.2: Criterios Clásicos de predicción para ARIMA(1,1,0).

	Ajuste		Predicción	
	WMSE	WMAE	WFMSE	WFMAE
Valor	0.2180055	0.2960323	0.03008745	0.1275013
Modelo	250	165	309	309

Tabla 4.3: Criterios propuestos de ajuste y predicción para ARIMA(1,1,0).

A la vista de las tablas 4.1, 4.2 y 4.3, se observa como hay criterios en los que los modelos elegidos coinciden, esto sucede con los criterios AIC2 y BIC2, que se decantan por el modelo 289 (ver anexo C), ocurre también en el caso de los criterios ECM, U_1 y U_2 , cuyo modelo seleccionado es el 321 (ver anexo C), y en el caso del ECMpre con los criterios de predicción propuestos cuyo modelo seleccionado es el 309 (ver anexo C).

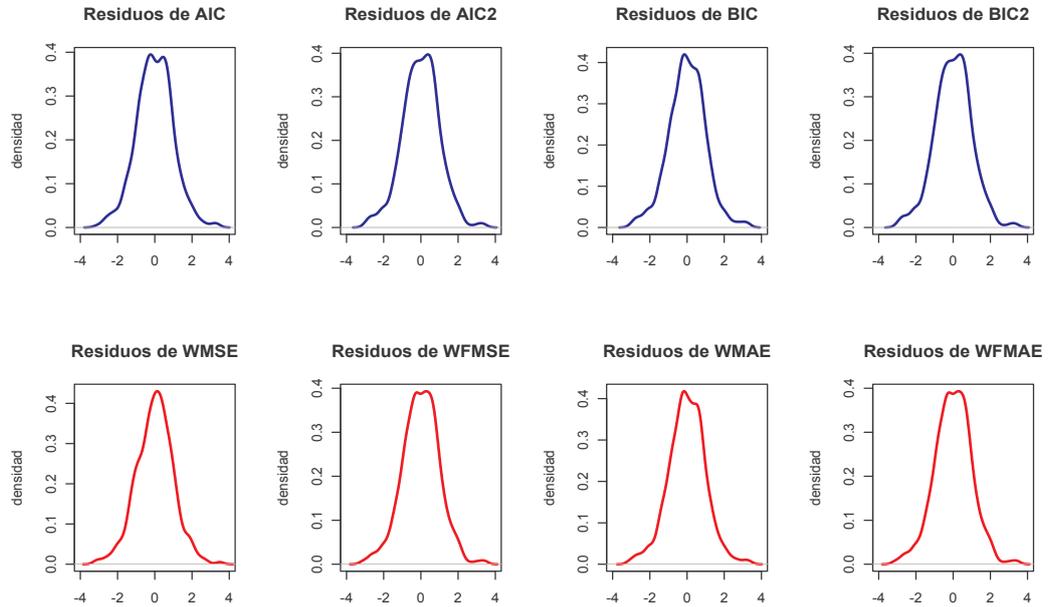


Figura 4.1: Residuos obtenidos con los modelos de los criterios de ajuste de ARIMA(1,1,0).

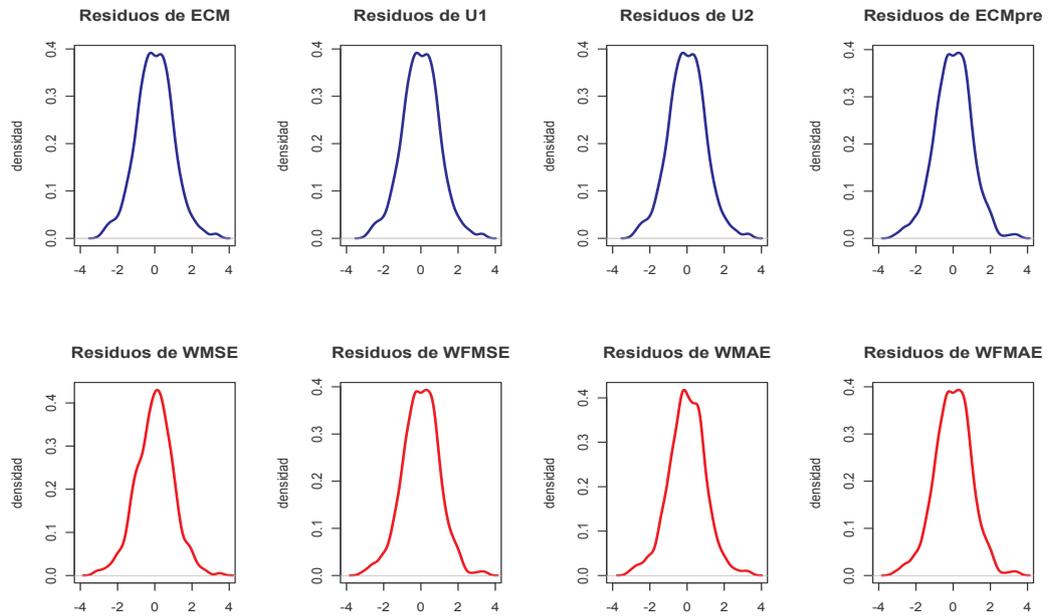


Figura 4.2: Residuos obtenidos con los modelos de los criterios de predicción de ARIMA(1,1,0).

En los gráficos 4.1 y 4.2 se puede ver la representación de los residuos obtenidos ante el empleo del modelo seleccionado con cada criterio, para todas las representaciones que serán mostradas a lo largo de este capítulo, se han representado los criterios clásicos en azul y bajo ellos siempre se muestran los criterios que se han propuestos en el capítulo 3 de este trabajo, con el objetivo de que puedan ser fácilmente comparados, además y para que la comparación sea factible en todo momento se ha establecido la misma escala del eje X, tipificando en cada caso las distribuciones representadas. Observando las gráficas citadas, se ve que la distribución de los residuos es bastante similar con cada criterio, a pesar de que el modelo utilizado en cada caso ha sido diferente salvo en las ocasiones en que los criterios han empatado su elección, veamos en la tabla 4.4, donde se muestra la media y varianza de cada distribución si podemos diferenciar entre los criterios.

Criterio	Media	Varianza
AIC	-0.2083	17.4456
AIC2	-0.2084	16.6700
ECM	-0.2584	17.6268
BIC	-0.1976	16.7512
BIC2	-0.2084	16.6700
U1	-0.2584	17.6268
U2	-0.2584	17.6268
ECMpre	-0.2227	16.8755
WMSE	-0.2609	20.9829
WFMSE	-0.2227	16.8755
WMAE	-0.2128	16.8292
WFMAE	-0.2227	16.8755

Tabla 4.4: Medidas descriptivas de los residuos en el conjunto de comprobación para ARIMA(1,1,0).

Las medias están muy cercas a 0, por lo que no se considera que haya grandes diferencias, se podría determinar que el mejor modelo es aquel que tenga mínima varianza, en ese caso, serían los modelos elegidos por los criterios AIC2 y BIC2, que se corresponde con un ARIMA(2, 1, 1), no obstante, las diferencias entre los siguientes valores son mínimas, lo que nos lleva a realizar un estudio más detallado de la situación con el objetivo de poder tomar una decisión más solida.

En los gráficos 4.3 y 4.4 se representara el histograma de los residuos, el objetivo ahora es estudiar las predicciones realizadas, una a una, debido a que en la construcción de nuestros criterios se han penalizado los residuos más próximos a la última observación conocida de la serie, con lo cuál, la primera de las predicciones realizadas sería la que mayor interés suscita, por lo que se busca una mayor precisión en la misma, es decir, un menor ECM.

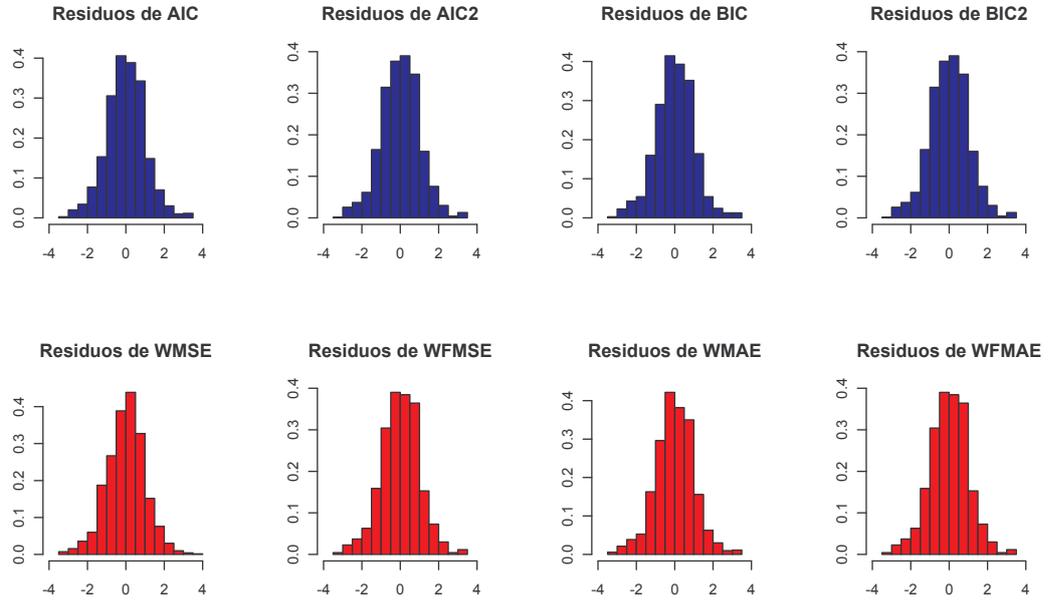


Figura 4.3: Histograma de los residuos en el conjunto de comprobación por los criterios de ajuste de ARIMA(1,1,0).

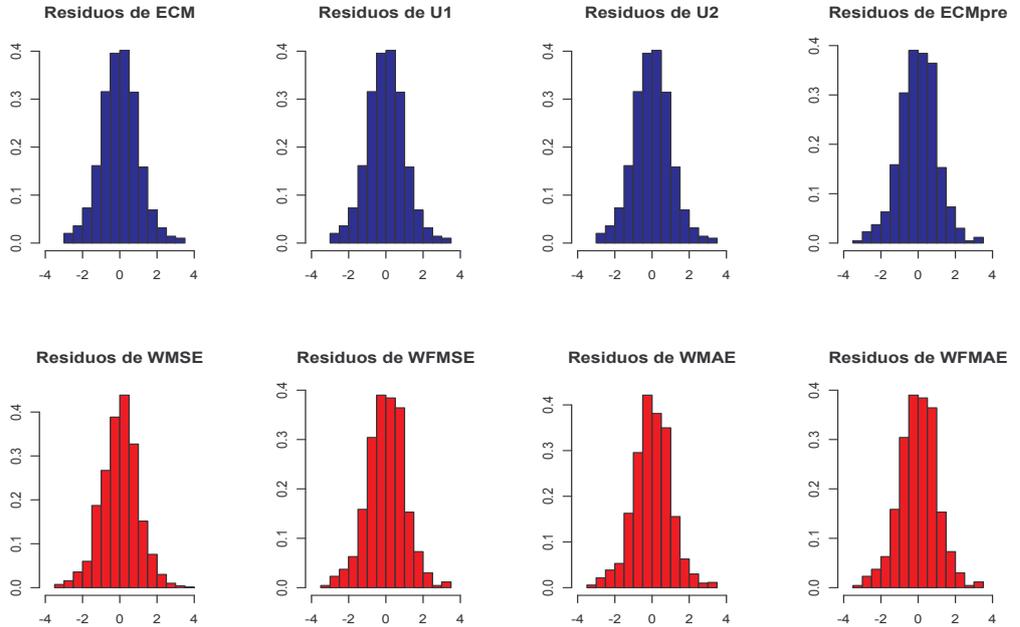


Figura 4.4: Histograma de los residuos en el conjunto de comprobación por los criterios de predicción de ARIMA(1,1,0).

Criterio	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
AIC	0.9040	3.7031	8.3795	14.5209	22.6414	31.7268	40.5473
AIC2	0.8803	3.6081	8.1614	14	21.6344	30.2484	38.4615
ECM	0.9642	3.8857	8.5733	14.6408	22.8776	32.0078	40.9058
BIC	0.8843	3.6328	8.1742	14.1033	21.8244	30.4052	38.5074
BIC2	0.8803	3.6081	8.1614	14	21.6344	30.2484	38.4615
U1	0.9642	3.8857	8.5733	14.6408	22.8776	32.0078	40.9058
U2	0.9642	3.8857	8.5733	14.6408	22.8776	32.0078	40.9058
ECMpre	0.9291	3.7572	8.3481	14.1983	21.9859	30.5547	38.7024
WMSE	1.2510	4.8191	10.7170	18.3898	28.7105	38.1861	45.2834
WFMSE	0.9291	3.7572	8.3481	14.1983	21.9859	30.5547	38.7024
WMAE	0.8664	3.6393	8.2231	14.1130	21.8804	30.5706	38.8086
WFMAE	0.9291	3.7572	8.3481	14.1983	21.9859	30.5547	38.7024

Tabla 4.5: ECMpre global para cada predicción realizada en el conjunto de comprobación en ARIMA(1,1,0).

Observando la tabla podemos ver como el menor ECMpre1 se consigue con el modelo 165, que corresponde con un SARIMA(1, 1, 0)(0, 0, 2)₇, sin embargo, el ECMGpre es 16,8745 que difiere muy poco del 16,7135 que se correspondía con el valor mínimo (ver tabla 4.7). Sin embargo, si miramos la segunda observación es el modelo 289 el que consigue el menor error igual ocurriría con las predicciones sucesivas.

Podemos observar de manera gráfica como se comportan los ECM de ajuste y predicción en cada una de las series generadas.

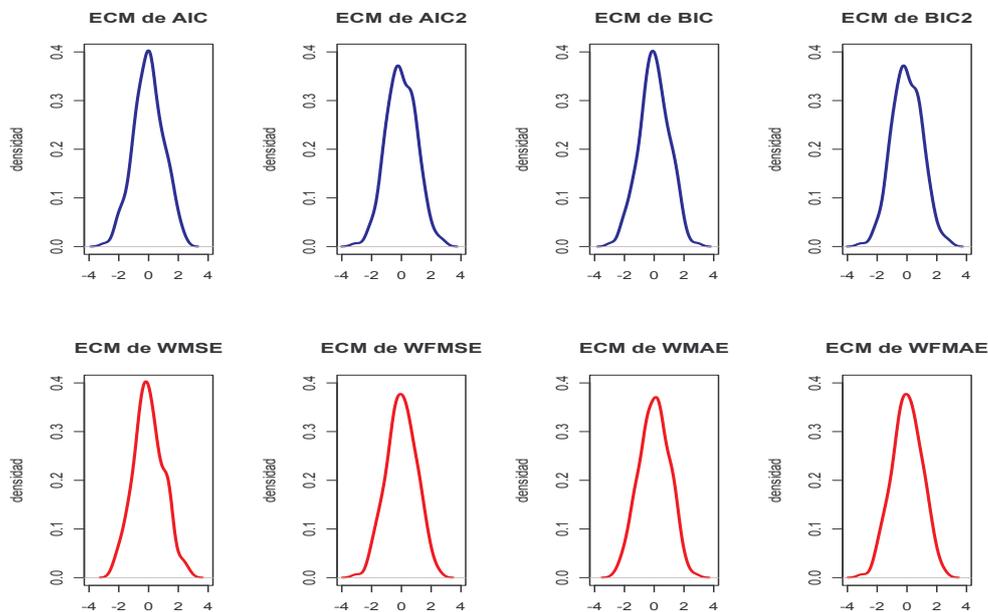


Figura 4.5: ECM de ajuste con los modelos obtenidos por los criterios de ajuste en ARIMA(1,1,0).

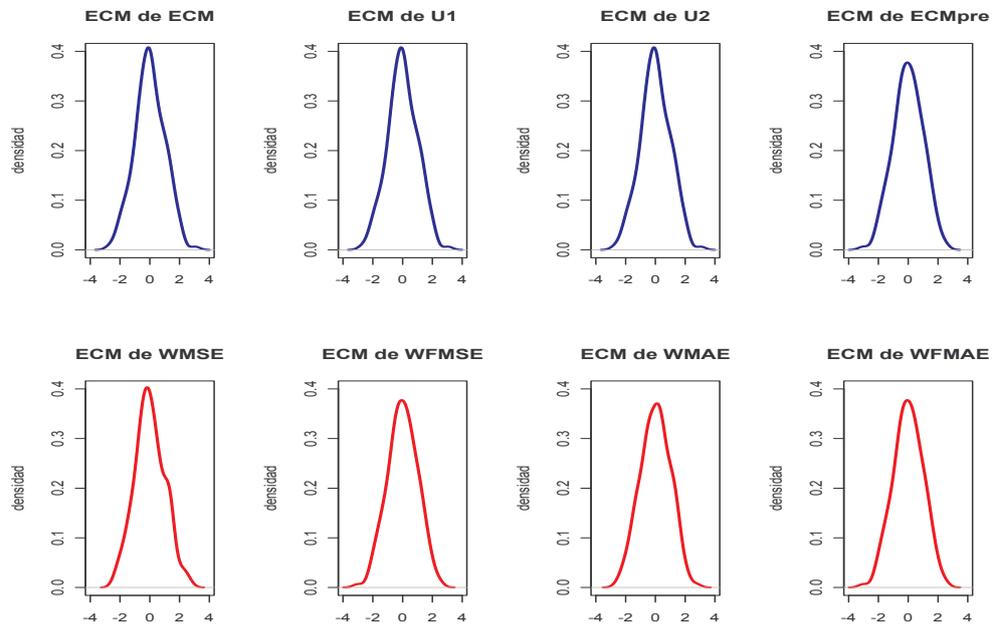


Figura 4.6: ECM de ajuste con los modelos obtenidos con los criterios de predicción en ARIMA(1,1,0).

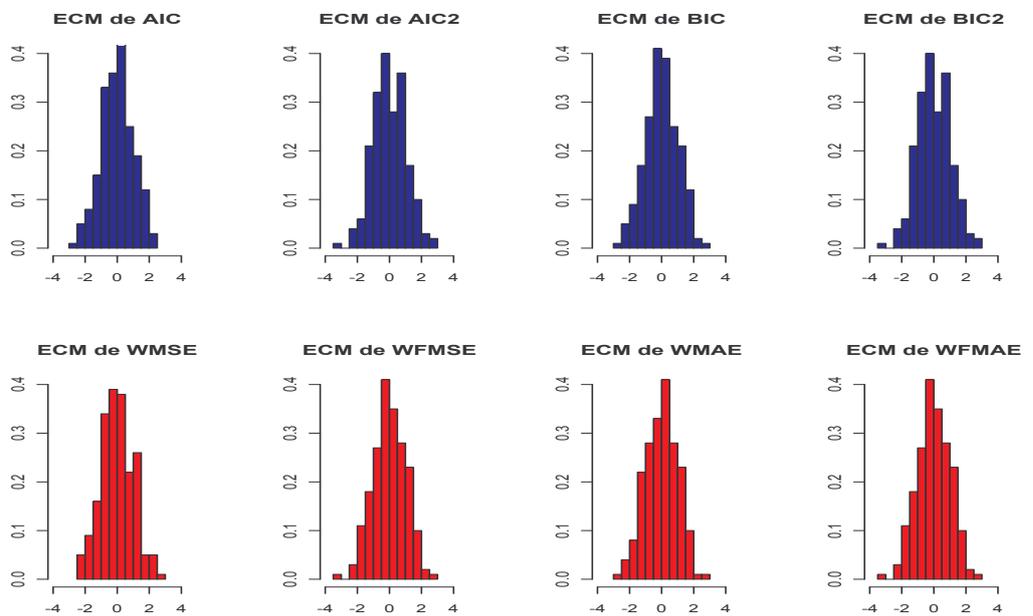


Figura 4.7: Histograma de los ECM de ajuste con los modelos obtenidos por los criterios de ajuste en ARIMA(1,1,0).

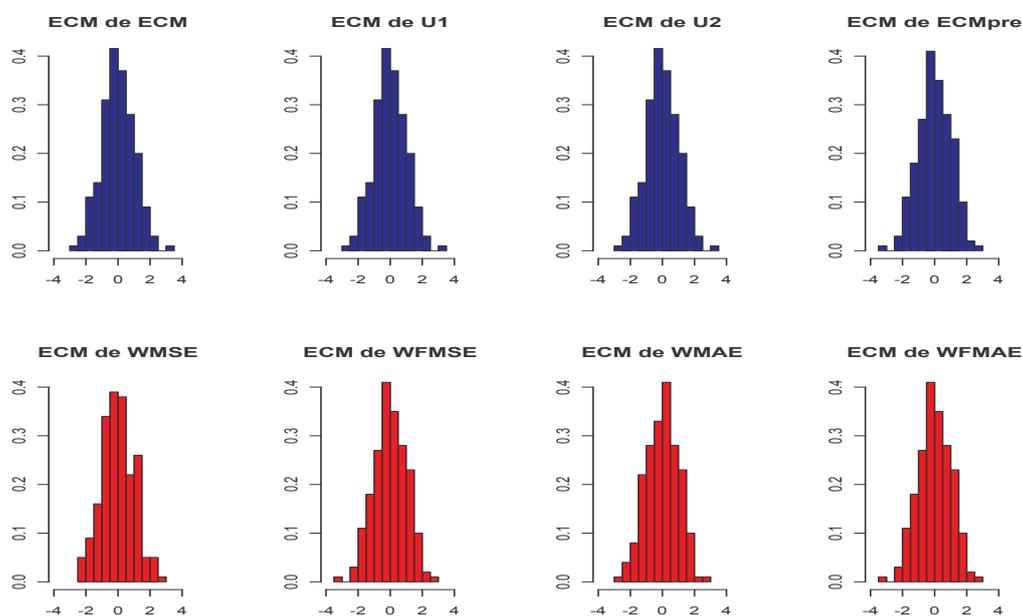


Figura 4.8: Histograma de los ECM de ajuste con los modelos obtenidos con los criterios de predicción en ARIMA(1,1,0).

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8605	0.9703	0.8480	0.8906	0.9703	0.8480
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8480	0.9506	1.0803	0.9506	0.9801	0.9506

Tabla 4.6: El ECM global de ajuste, ECMG, para cada criterio en modelos ARIMA(1,1,0).

En cuanto al ajuste, los mejores resultados se obtienen con el criterio del ECM, que se corresponde con el modelo 321, que es un SARIMA(2, 1, 2)(2, 0, 2)₇ lo que complica mucho al modelo de partida, no obstante, el objetivo de este estudio se basa en buscar el modelo con mayor capacidad predictiva para la serie en estudio, con lo que conseguir o no un buen ajuste no será problema para conseguir la meta fijada, es por ello, que el interés mayor se encuentra en los valores del ECMpre, que se muestra en las gráficas siguientes.

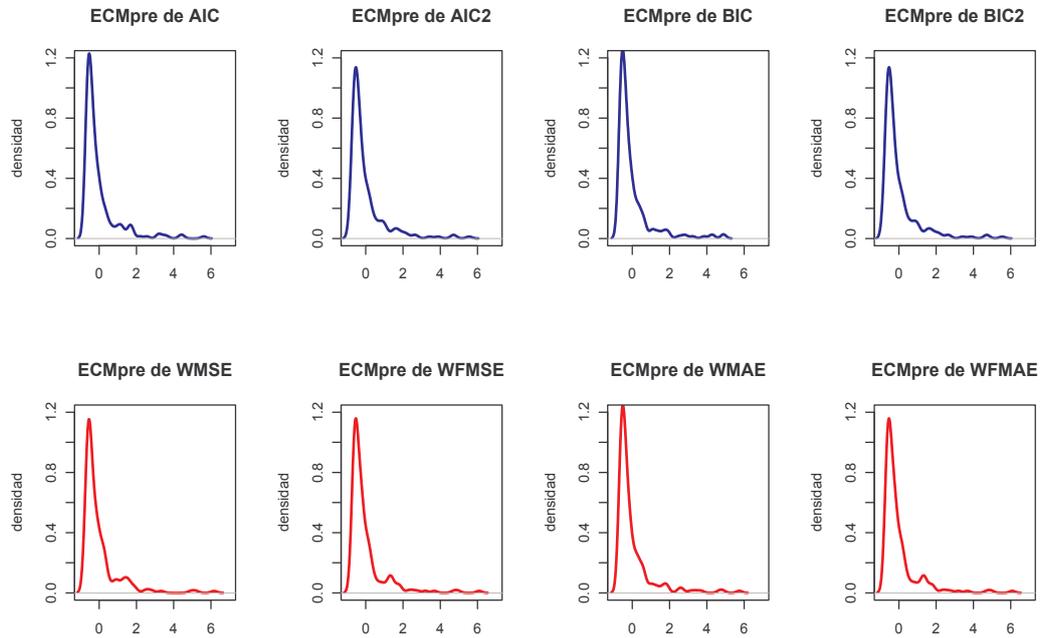


Figura 4.9: ECM de predicción de ajuste con los modelos obtenidos por los criterios de ajuste en ARIMA(1,1,0).

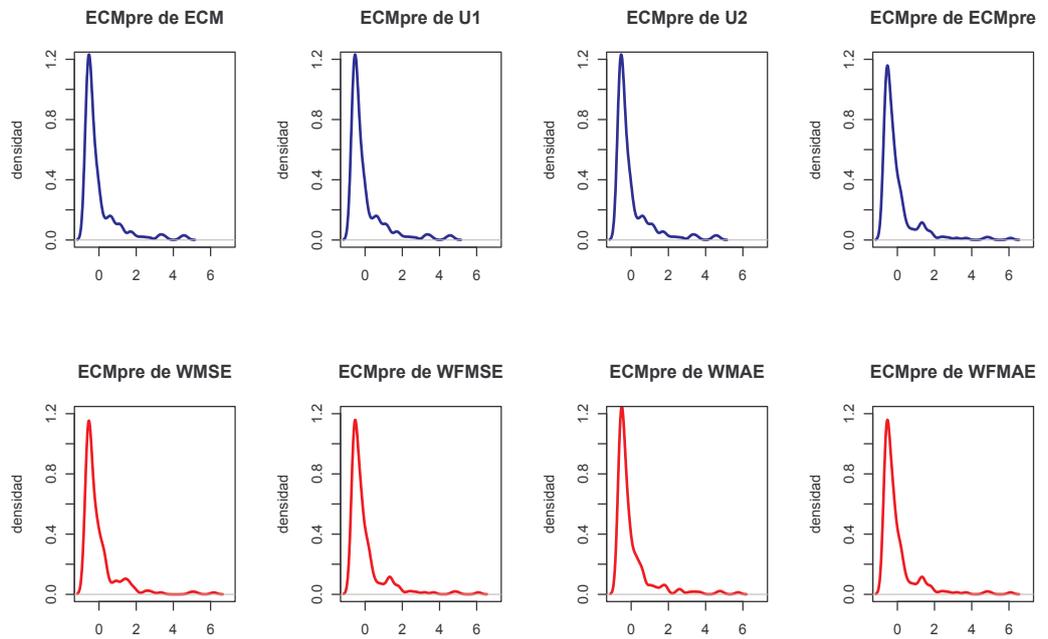


Figura 4.10: ECM de predicción de ajuste con los modelos obtenidos por los criterios de predicción en ARIMA(1,1,0).

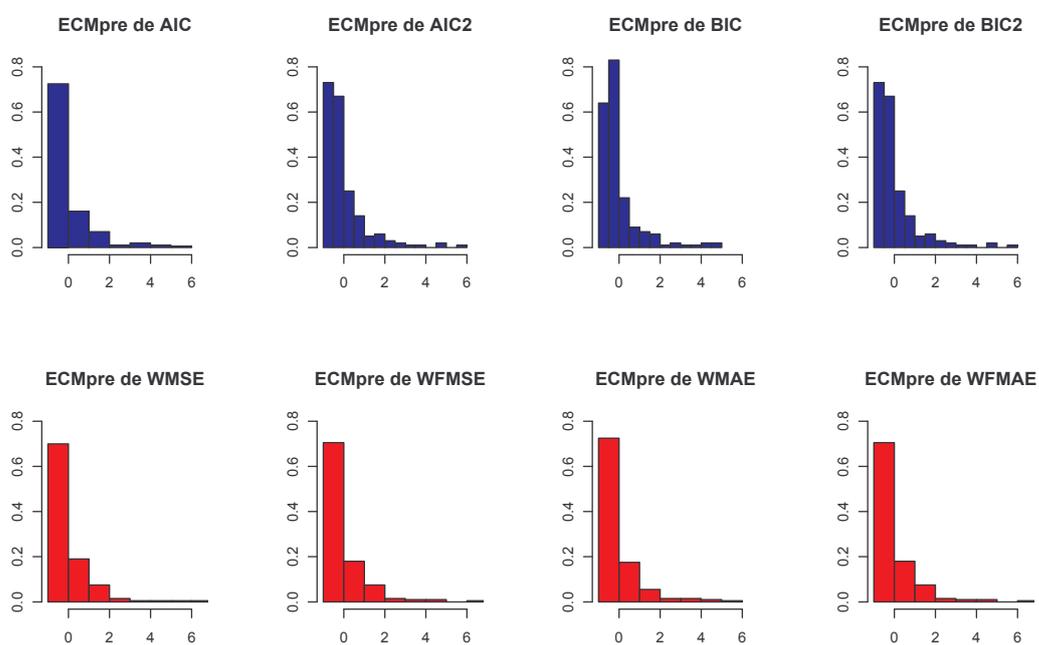


Figura 4.11: Histograma de los ECM de predicción de ajuste con los modelos obtenidos por los criterios de ajuste en ARIMA(1,1,0).

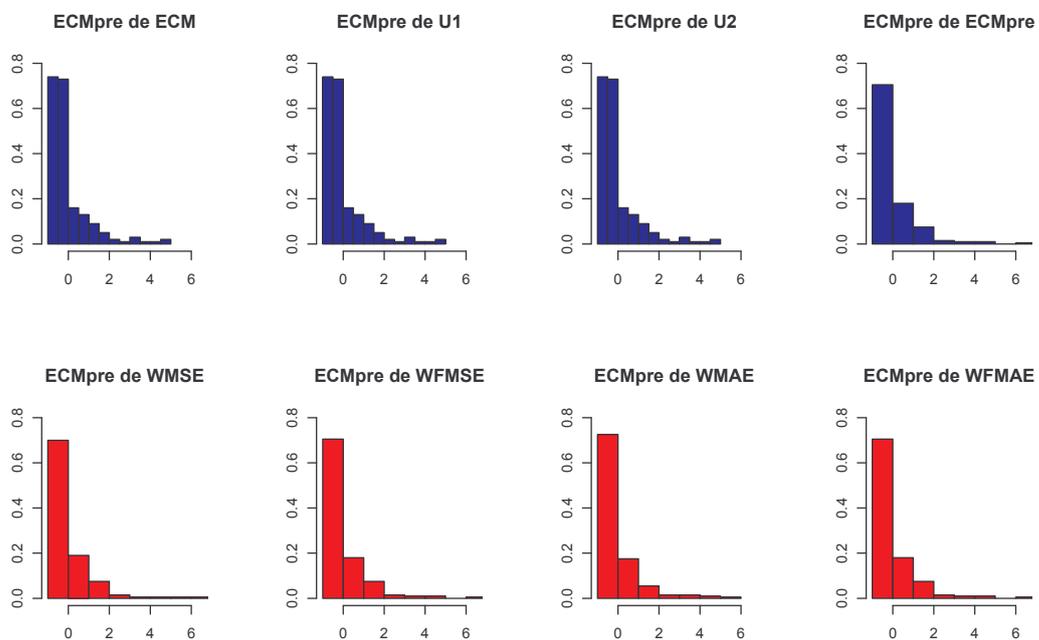


Figura 4.12: Histograma de los ECM de predicción de ajuste con los modelos obtenidos por los criterios de predicción en ARIMA(1,1,0).

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG _{pre}	17.4890	16.7135	17.6936	16.7902	16.7135	17.6936
	U2	ECM _{pre}	WMSE	WFMSE	WMAE	WFMAE
ECMG _{pre}	17.6936	16.9251	21.0510	16.9251	16.8745	16.9251

Tabla 4.7: ECM global de predicción, ECMG_{pre}, para cada criterio en modelos ARIMA(1,1,0).

En cuanto a predicción, las diferencias entre los ECM_{pre} no son muy elevadas, con lo que tampoco se podría concluir determinando un único modelo.

En general, se ha comprobado que considerando todas las series y realizando las predicciones con el modelo seleccionado no se alcanzan buenos resultados, por tanto, se considerarán dos grupos de series; uno que llamaremos el grupo de series correctas, que serán aquellas donde el criterio empleado ha seleccionado como modelo el mismo que ha generado la serie, y el grupo de series incorrectas, que serán aquellas donde el criterio ha elegido un modelo distinto. Se probarán todos los criterios, aunque solo se estudiarán aquellos que han sido capaces de seleccionar el modelo “correcto” en alguna de las 200 series (al menos 2 de las 200), en el caso del ARIMA(1, 1, 0) que estamos estudiando, los criterios que han elegido en alguna ocasión el modelo mencionado han sido AIC2, BIC, BIC2 y WFMSE.

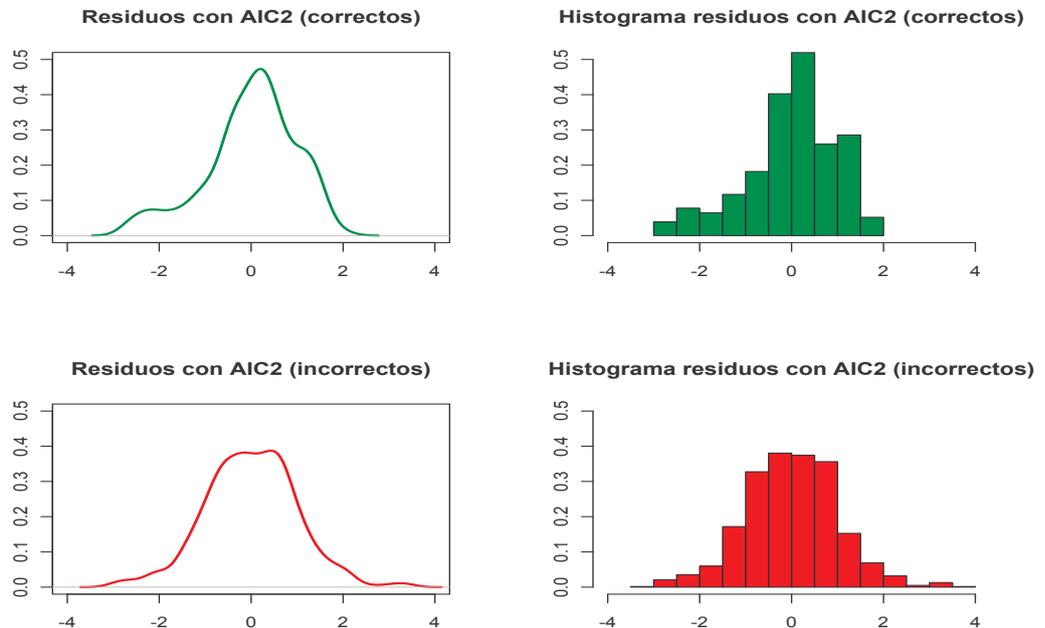


Figura 4.13: Residuos de las series correctas e incorrectas con el criterio AIC2 en ARIMA(1,1,0).

	Media	Varianza	ECMG	ECMGpre
Correctos	-0.0118	18.9658	1.0182	18.9659
Incorrectos	-0.2327	16.3809	0.8658	16.4351

Tabla 4.8: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,0) con el criterio AIC2.

En las series correctas se ha obtenido una media menor, sin embargo, la varianza es mayor que en las series incorrectas, con lo cuál, en este caso funcionan mejor en predicción, las series incorrectas, es decir, aquellas donde la predicción se ha realizado con un modelo distinto del modelo que las generó, se puede observar claramente como el ECMGpre de las series incorrectas es menor que en las series correctas, lo mismo pasa en cuanto a mejor ajuste, es en las series donde el modelo ajustado no coincide con el generador donde se consigue un mejor ajuste de la misma.

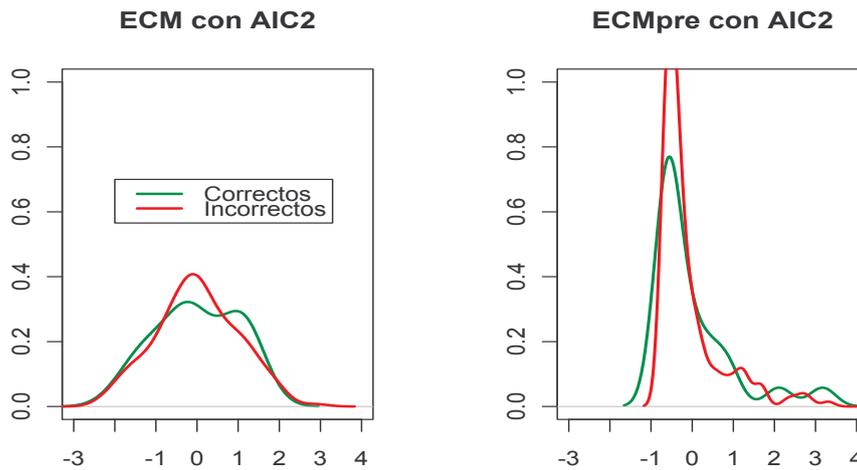


Figura 4.14: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en ARIMA(1,1,0).

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.9960	3.8697	7.8204	15.6103	26.1471	35.7813	42.5365
Incorrectos	0.8660	3.5758	7.2036	13.8010	21.0766	29.5646	37.9578

Tabla 4.9: ECM global de predicción en ARIMA(1,1,0) con el criterio AIC2 en el conjunto de comprobación.

Como podemos comprobar en la tabla 4.9, las predicciones, estudiadas en

este caso una a una, son mejores en el grupo de las series incorrectas que en el grupo de las series correctas, por tanto, en la tabla 4.8 determinamos que las mejores predicciones globales se conseguían con las series incorrectas, hecho que se verifica en la tabla 4.9, donde se ve que las predicciones consideradas “individualmente” son más precisas cuando no se utiliza para su cálculo el modelo generador de la serie.

Veamos que ocurre en el caso, donde el criterio de selección sea BIC.

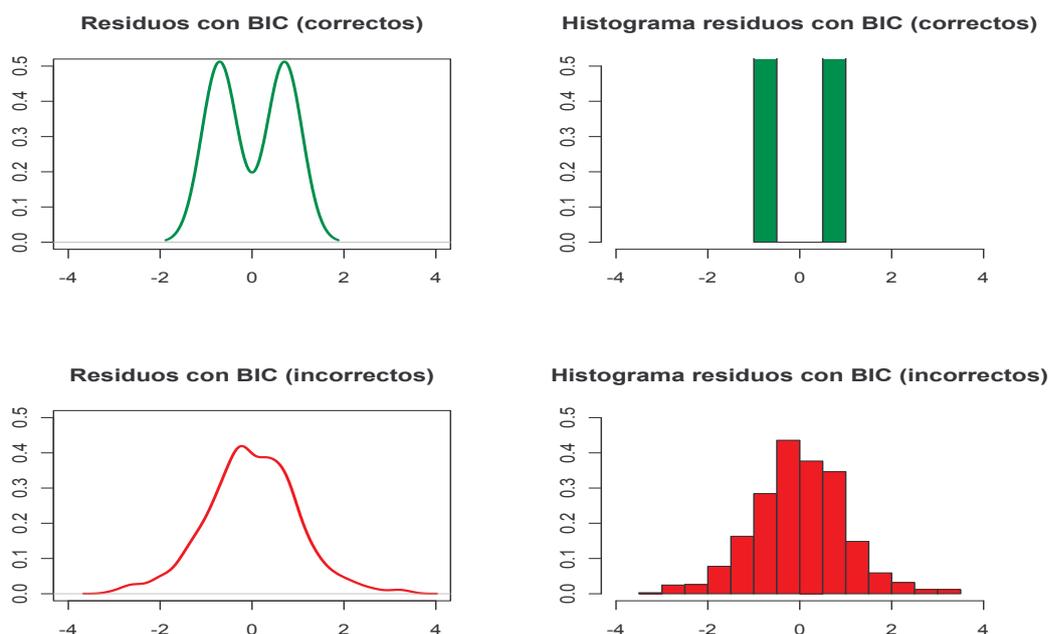


Figura 4.15: Residuos de las series correctas e incorrectas con el criterio BIC en ARIMA(1,1,0).

	Media	Varianza	ECMG	ECMGpre
Correctos	0.4686	4.3718	0.8971	4.5915
Incorrectos	-0.2406	17.4717	0.8757	17.5297

Tabla 4.10: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para ARIMA(1,1,0) con el criterio BIC.

En cuanto a la predicción global se ve una importante diferencia entre las series correctas e incorrectas, siendo en este caso las predicciones mejores en el primer grupo (tabla 4.10), esto se corrobora en la tabla 4.11 donde se consideran las predicciones una a una y se observa la mejoría que se obtienen al utilizar el modelo ARIMA(1, 1, 0).

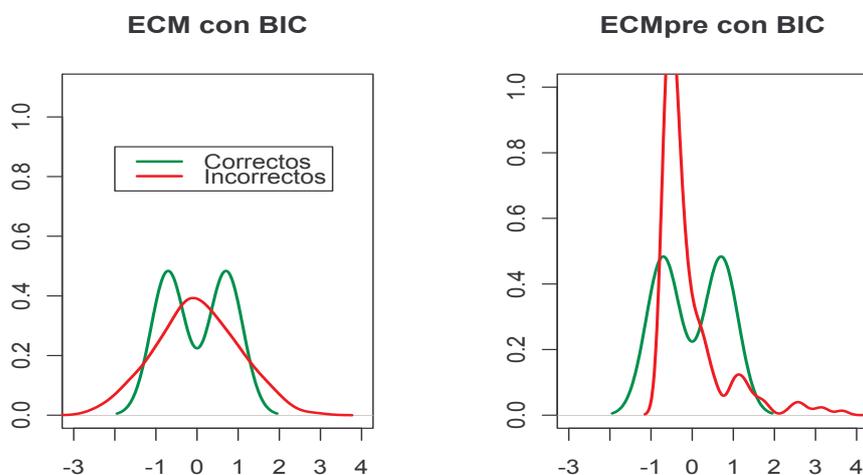


Figura 4.16: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC en ARIMA(1,1,0).

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.4360	1.8819	0.9613	1.2647	1.8075	7.4466	18.3423
Incorrectos	0.9151	3.7356	8.3799	14.5544	22.7782	31.9149	40.4296

Tabla 4.11: ECM global de predicción en ARIMA(1,1,0) con el criterio BIC en el conjunto de comprobación.

Con el criterio BIC si observamos los ECM de ajuste y predicción, los mejores resultados se consiguen en las series correctas, como puede verse en la tabla 4.10, sin embargo si observamos la representación de los residuos gráfica 4.15, la distribución en las correctas se ajusta más al comportamiento de una distribución normal, y la media que se consigue es inferior a la obtenida en las series correctas, aunque es fácil comprobar que la varianza en las series incorrectas es mucho superior a la varianza en las series correctas, y en cuanto a la predicción individual (tabla 4.11) en las series correctas se consiguen mejores predicciones, aunque se aprecia un aumento en el error cometido a medida que nos alejamos del instante último del conjunto de entrenamiento, no obstante, el incremento de ese error es significativamente mayor en las series incorrectas. No obstante, el criterio BIC tan sólo ha conseguido identificar al 1% de las series estudiadas, es decir, tan sólo dos de las series generadas con el modelo ARIMA(1, 1, 0) han sido detectadas con el criterio, frente al 11% que se conseguía con el criterio anterior, AIC2.

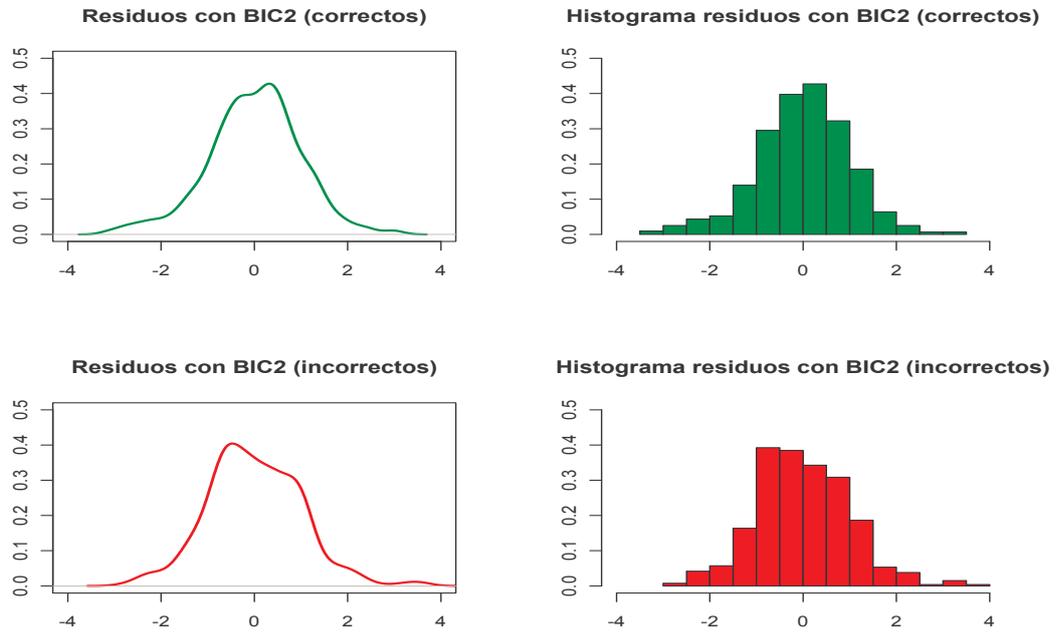


Figura 4.17: Residuos de las series correctas e incorrectas con el criterio BIC2 en ARIMA(1,1,0).

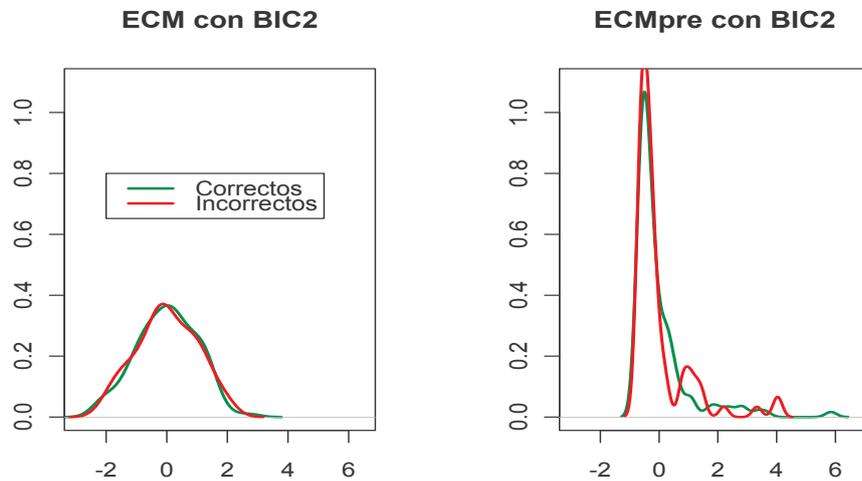


Figura 4.18: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en ARIMA(1,1,0).

	Media	Varianza	ECMG	ECMGpre
Correctos	0.0491	15.7781	0.9922	15.7806
Incorrectos	-0.7254	18.2197	0.8834	18.7459

Tabla 4.12: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,0) con el criterio BIC2.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.9207	3.8458	8.0465	13.8140	20.9711	28.2070	34.6587
Incorrectos	0.7890	3.0585	7.9520	13.9547	23.1460	35.1864	47.1346

Tabla 4.13: ECM global de predicción en ARIMA(1,1,0) con el criterio BIC2 en el conjunto de comprobación.

Con el criterio BIC2, la distribución de los residuos en la gráfica 4.17 se comporta mejor en las series correctas, lo que puede comprobarse en la tabla 4.12 donde vemos que el ECMGpre es menor en este grupo, además, la media y la varianza también consiguen ser menor en el grupo de las series bien identificadas, en cuanto a la predicción es en las series incorrectas donde se consiguen mejores resultados, sin ser éstos significativos como se puede observar en la gráfica 4.18, donde las curvas casi están superpuestas. En la predicción, a pesar, de que el menor valor del ECMGpre se consigue en las series correctas, se comprueba en la tabla 4.13 que la primera predicción realizada es mejor en las series mal identificadas, es decir, en aquellas en la predicción ha sido obtenida empleando un modelo distinto al generador y ese comportamiento se extiende hasta la tercera de las predicciones realizadas, cambiando el comportamiento radicalmente a partir de la quinta de las predicciones donde los mejores resultados son obtenidos en las series correctas, no obstante, éstos resultados no son buenos en ninguno de los casos, dado que el error predictivo aumenta mucho en los últimos casos. Con este criterio ha sido posible detectar al 62,5% de las series dentro del grupo de series correctas, es decir, bien identificadas.

En último lugar estudiaremos el comportamiento de estos dos grupos de series que se han considerando cuando el criterio utilizado es el criterio de predicción propuesto, WFMSE, aunque al igual que pasaba con el criterio BIC, tan sólo ha sido posible identificar al 1% de las 200 series estudiadas, lo que se puede ver claramente en el histograma del gráfico 4.19.

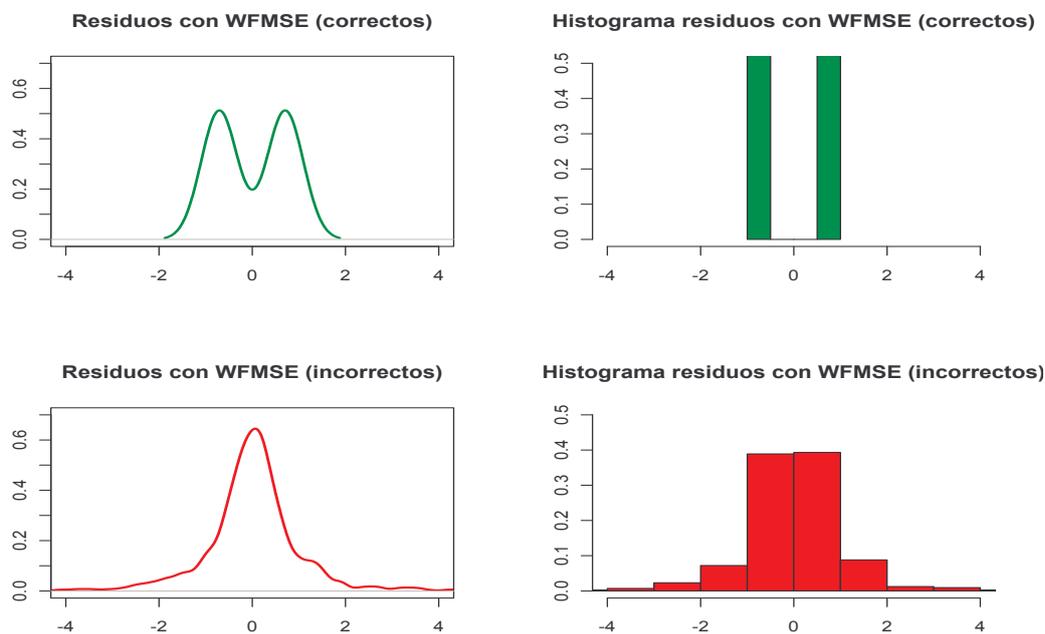


Figura 4.19: Residuos de las series correctas e incorrectas con el criterio WFMSE en ARIMA(1,1,0).

	Media	Varianza	ECMG	ECMGpre
Correctos	-0.2485	2.2006	1.0403	2.2624
Incorrectos	-0.0510	4.3405	4.9607	4.3431

Tabla 4.14: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,0) con el criterio WFMSE.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	1.7222	2.5125	1.0521	0.3043	1.8789	3.8273	4.5395
Incorrectos	0.6284	1.4493	2.3366	2.6503	4.3648	7.4506	11.5219

Tabla 4.15: ECM global de predicción en ARIMA(1,1,0) con el criterio WFMSE en el conjunto de comprobación.

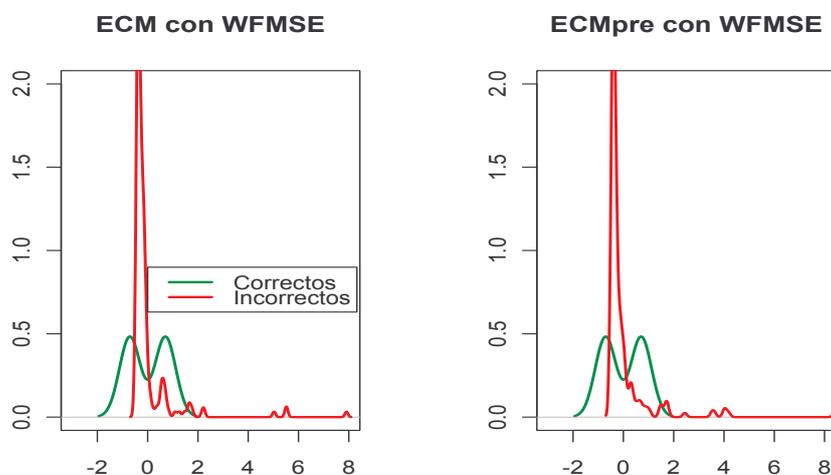


Figura 4.20: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMSE en ARIMA(1,1,0).

En esta ocasión, también la mejor predicción global se consigue en las series correctas como se puede comprobar en la tabla 4.14, es también en este grupo donde se alcanza el mejor ajuste de las series, al igual que ocurría con el criterio BIC2, en las primeras predicciones individuales se consiguen mejores resultados con los modelos utilizados en las series que no han sido identificadas, en este caso en las dos primeras, aumentando mucho el error en las siguientes, y siendo con este criterio con el que se consiguen los mejores resultados globales con respecto a los criterios AIC2, BIC y BIC2, estudiados anteriormente, incluso las mejores predicciones individuales.

A la vista de los resultados anteriores no se puede determinar un procedimiento de actuación efectivo cuando el objetivo perseguido es obtener el modelo cuya capacidad predictiva sea mayor, y con el que se consigan las mejores predicciones, teniendo en cuenta una penalización diferente sobre las mismas.

En último lugar, se llevará a cabo el estudio ya citado sobre las predicciones de las series con los distintos criterios utilizando en cada caso y para cada serie estudiada el modelo seleccionado, no teniendo que ser el mismo modelo para todas las series, sino que cada una se asociará a la serie que determine cada criterio en el conjunto de entrenamiento, con ese modelo se realizarán las predicciones pertinentes en el conjunto de comprobación. Se puede observar en las gráficas 4.21 y 4.22 el comportamiento de los residuos obtenidos.

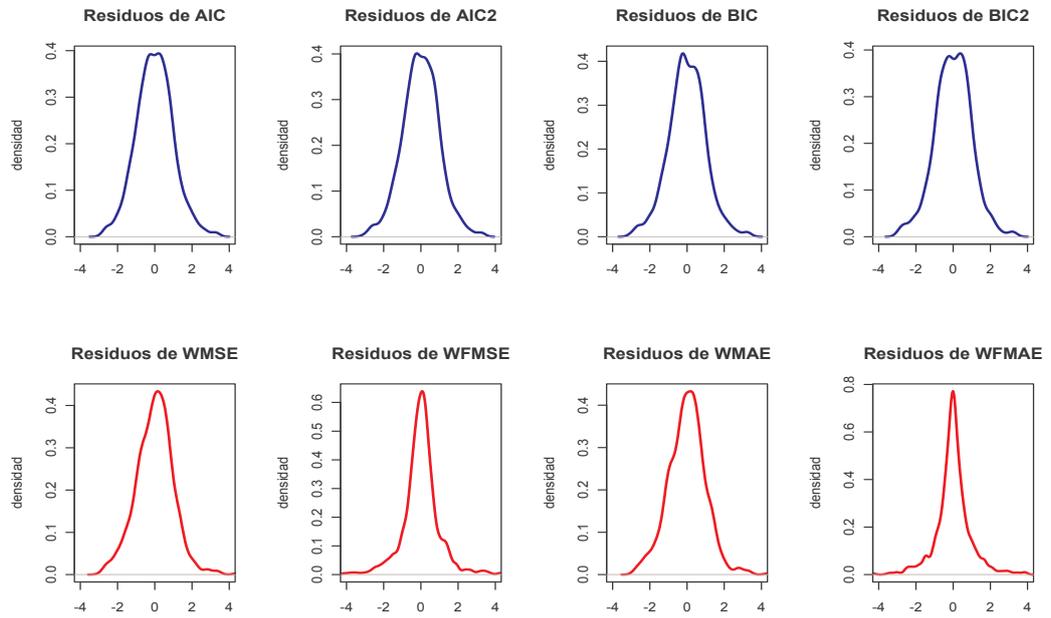


Figura 4.21: Residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).

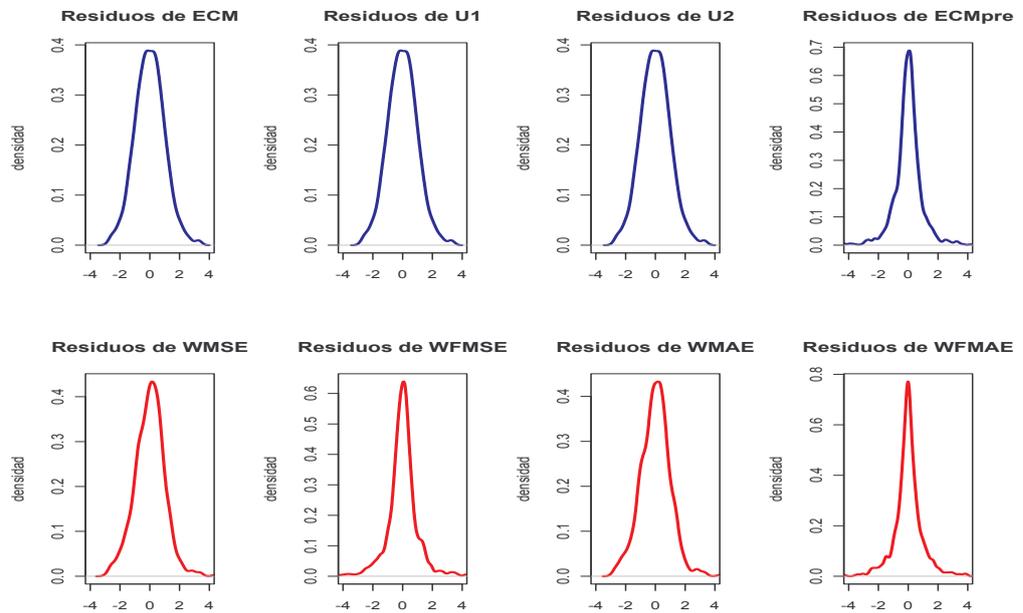


Figura 4.22: Residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).

Criterio	Media	Varianza
AIC	-0.1410	17.3084
AIC2	-0.1366	17.2881
ECM	-0.2195	17.8693
BIC	-0.2336	17.3458
BIC2	-0.2414	16.8343
U1	-0.2195	17.8693
U2	-0.2195	17.8693
ECMpre	-0.0675	4.9832
WMSE	-0.2243	20.9815
WFMSE	-0.0530	4.3195
WMAE	-0.0745	21.0306
WFMAE	-0.0377	4.8441

Tabla 4.16: Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para ARIMA(1,1,0).

Observando los resultados recogidos en la tabla 4.16, vemos que las medias son muy similares entre si, encontrándose una mejoría en las mismas en los criterios ECMpre, WFMSE y WFMAE, que además toma más fuerza cuando se comprueba claramente una menor varianza en relación al resto de criterios estudiados, esta mejoría se mantiene si se comparan estos resultados con los obtenidos en la tabla 4.4. En las gráficas 4.21 y 4.22 se observa como las curvas de los criterios citados son más “estrechas”, lo que muestra la menor varianza que se observa en la tabla 4.16.

Entre los tres criterios citados los resultados son mejores en el caso de los criterios propuestos, aunque en cuanto al error de predicción global no se aprecien grandes diferencias, no obstante, veremos en la tabla 4.17 que ocurre con los errores individuales donde se apreciará una clara mejoría en las primeras predicciones con los criterios de predicción propuestos, resultados que también son mejores que los que se exponían en la tabla 4.5.

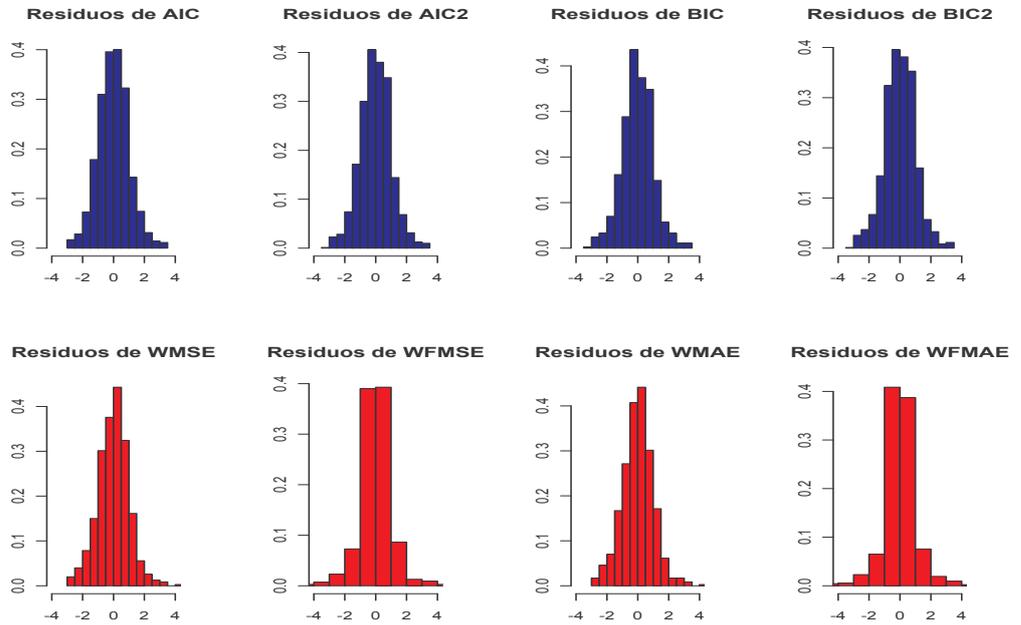


Figura 4.23: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).

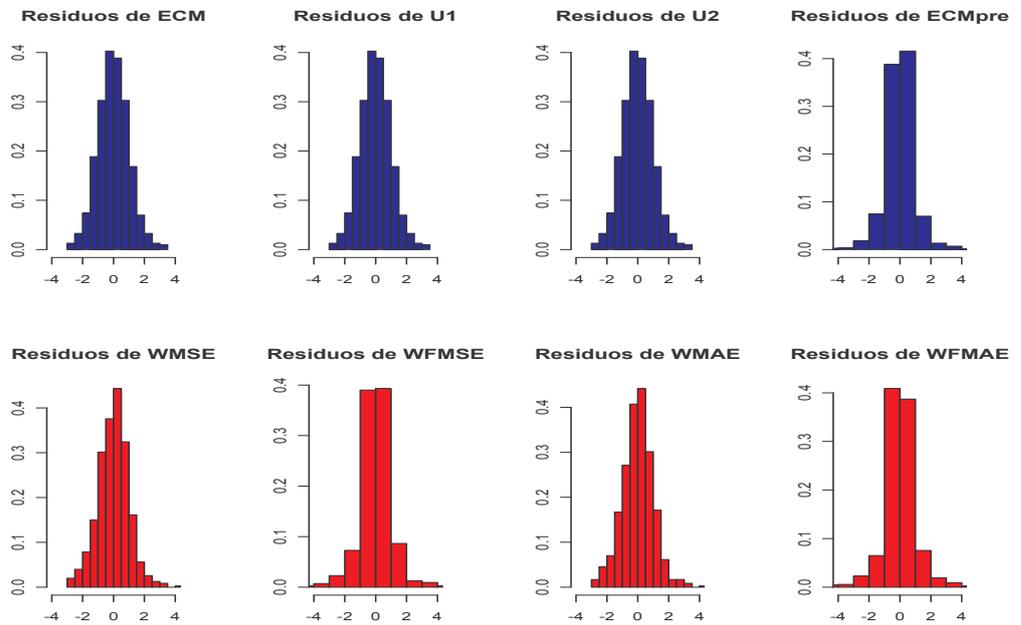


Figura 4.24: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).

Criterio	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
AIC	0.9550	3.7276	8.3420	14.1766	22.1465	31.3194	40.6311
AIC2	0.9491	3.6662	8.1205	13.8662	21.8523	31.4876	41.1494
ECM	0.9940	3.8911	8.5792	14.6001	22.9503	32.4591	41.9487
BIC	0.9103	3.7171	8.3058	14.4215	22.5686	31.6702	40.2088
BIC2	0.8713	3.5506	8.0110	13.8668	21.7867	30.8243	39.3372
U1	0.9940	3.8911	8.5792	14.6001	22.9503	32.4591	41.9487
U2	0.9940	3.8911	8.5792	14.6001	22.9503	32.4591	41.9487
ECMpre	1.2411	2.5358	3.1573	2.9811	3.8889	6.1950	9.2991
WMSE	1.2494	4.7931	9.9349	17.2027	27.5901	38.3627	48.0900
WFMSE	0.6394	1.4599	2.3237	2.6268	4.3399	7.4144	11.4521
WMAE	1.2205	4.7715	10.0703	17.3178	27.7377	38.4084	47.7269
WFMAE	0.4821	1.1948	2.1692	2.6160	4.8813	8.9370	13.6381

Tabla 4.17: ECM globales para cada predicción realizada en el conjunto de comprobación en ARIMA(1,1,0).

En la tabla 4.17 podemos comprobar como los modelos seleccionados por los criterios WFMSE y WFMAE son los que consiguen las mejores predicciones no sólo globales sino también en los cuatro primeros momentos, siendo el ECMpre el que consigue que las 3 últimas de las predicciones realizadas se ajusten mejor a los datos observados. Por tanto, y considerando que es de mayor interés conseguir mayor precisión en las primeras predicciones, podemos considerar que el plan de actuación para resolver el problema que venimos estudiando en este caso de los ARIMA(1, 1, 0) sería considerar para cada serie el modelo que minimiza el criterio de predicción que ha sido propuesto, ya sea WFMSE o WFMAE, considerando además un modelo que no tiene porque ser el mismo en cada una de las 200 series evaluadas, es decir, no existe un modelo “óptimo” de predicción para el caso de los modelos ARIMA(1, 1, 0), sino un conjunto de modelos que será en cada caso el que minimice el criterio empleado.

Como el programa implementado para realizar dichas predicciones está bastante automatizado siempre se podrá actualizar las últimas predicciones donde los resultados son peores, conforme se acerque el instante.

Para completar este estudio se presentan a continuación, al igual que anteriormente las representaciones y resultados referentes a los ECM de ajuste y predicción de las series objeto de estudio.

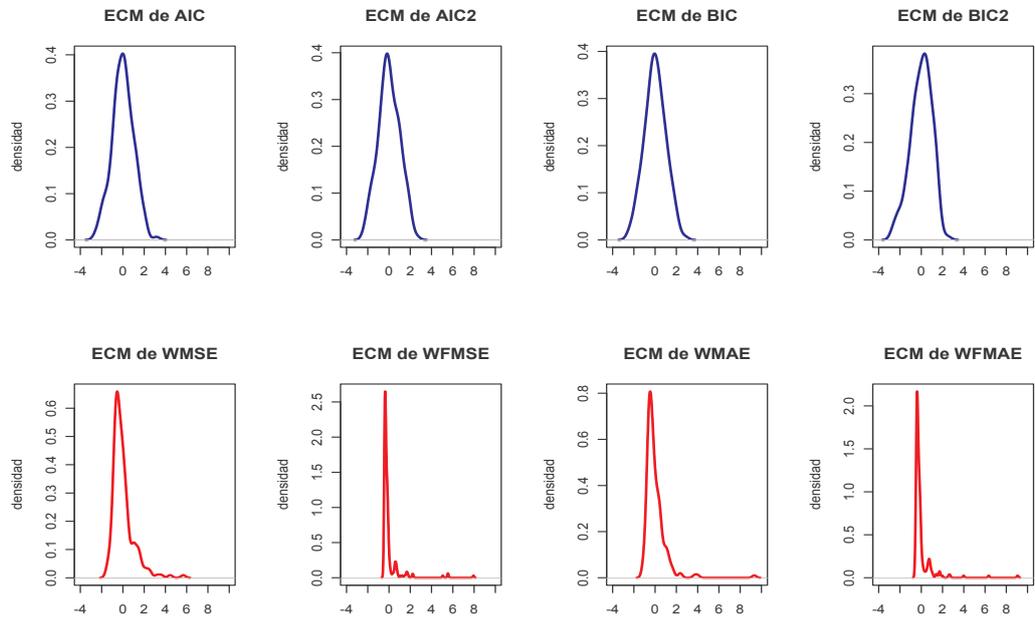


Figura 4.25: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).

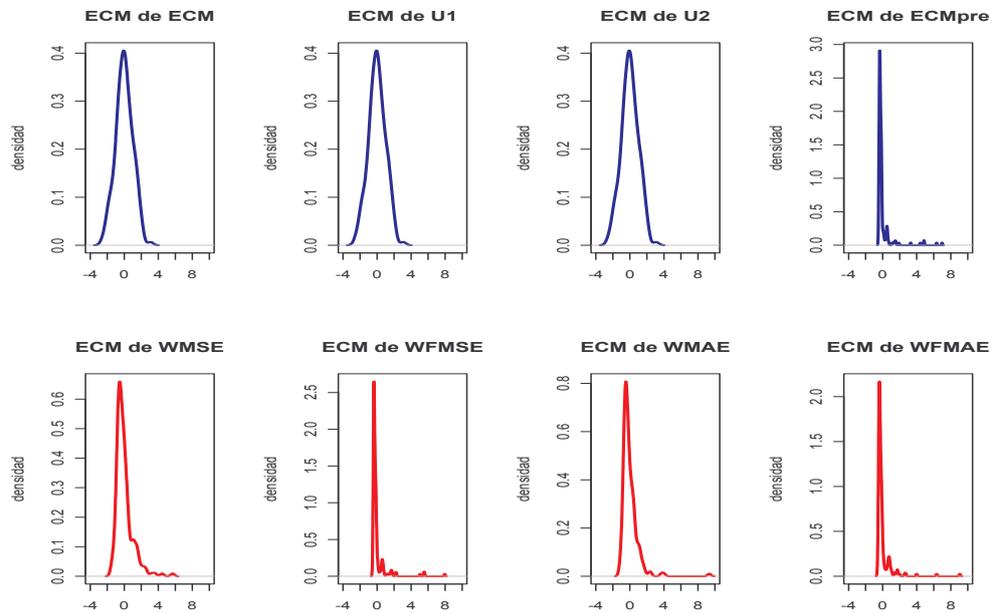


Figura 4.26: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).

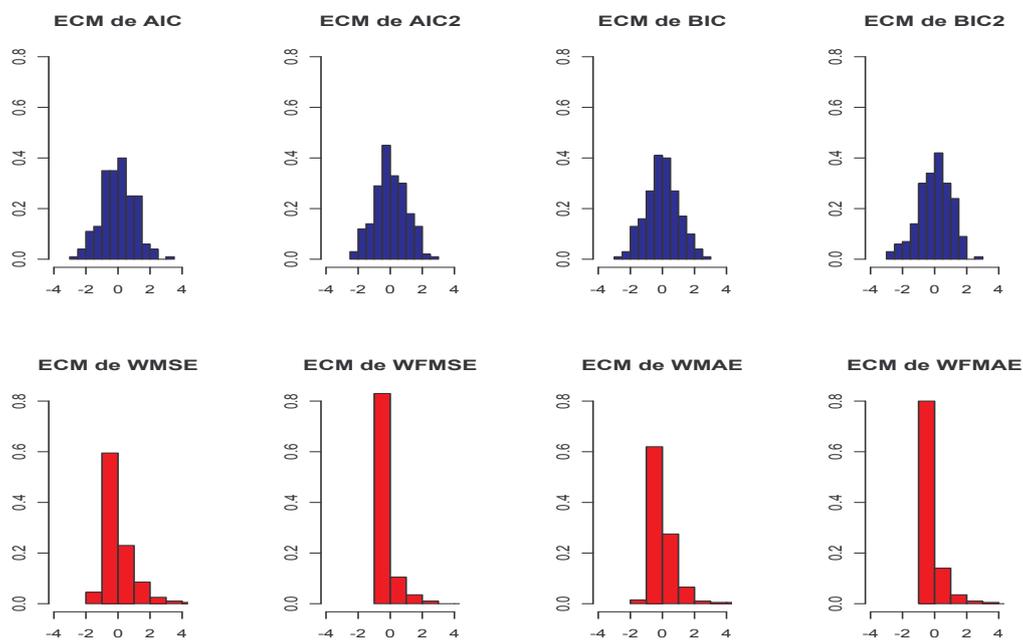


Figura 4.27: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).

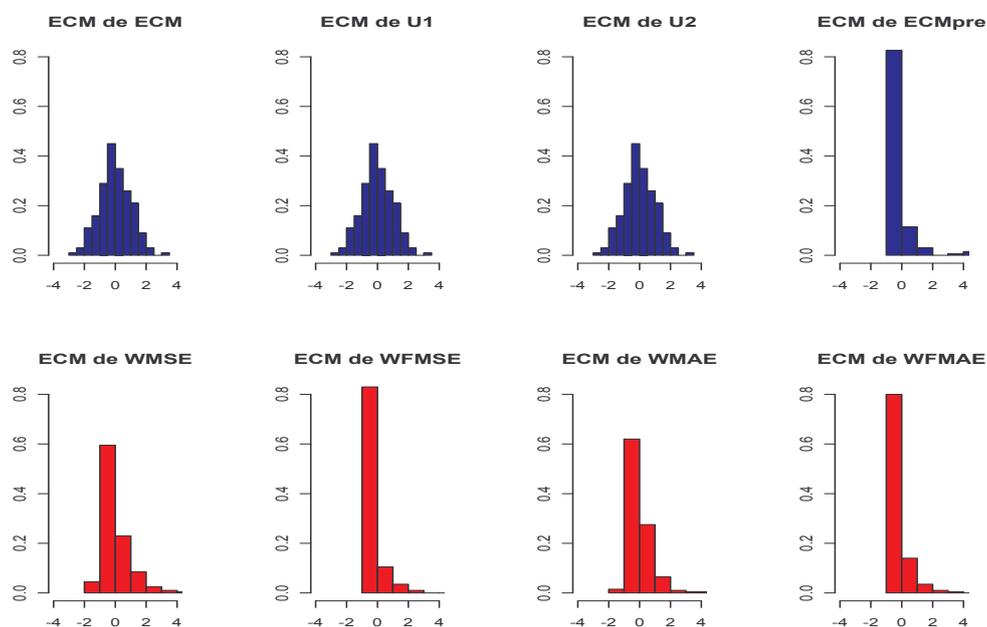


Figura 4.28: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8485	0.8825	0.8391	0.8759	0.9514	0.8391
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8391	5.7503	1.0973	4.9215	1.1894	4.2657

Tabla 4.18: ECM global de ajuste, ECMG, para cada criterio en modelos ARIMA(1,1,0).

En las gráficas 4.25, 4.26, 4.27 y 4.28 se observa que el comportamiento del ECM de ajuste en los criterios propuestos, tanto los de ajuste como los de predicción, es bastante peor que en los criterios clásicos. En la tabla 4.18 se puede comprobar como el ajuste en los criterios citados no obtiene buenos resultados globales es más, empeora en cuanto a los resultados mostrados en la tabla 4.6, el ECMpre es otro de los criterios que consigue malos resultados de ajuste aunque en este caso este criterio es de predicción.

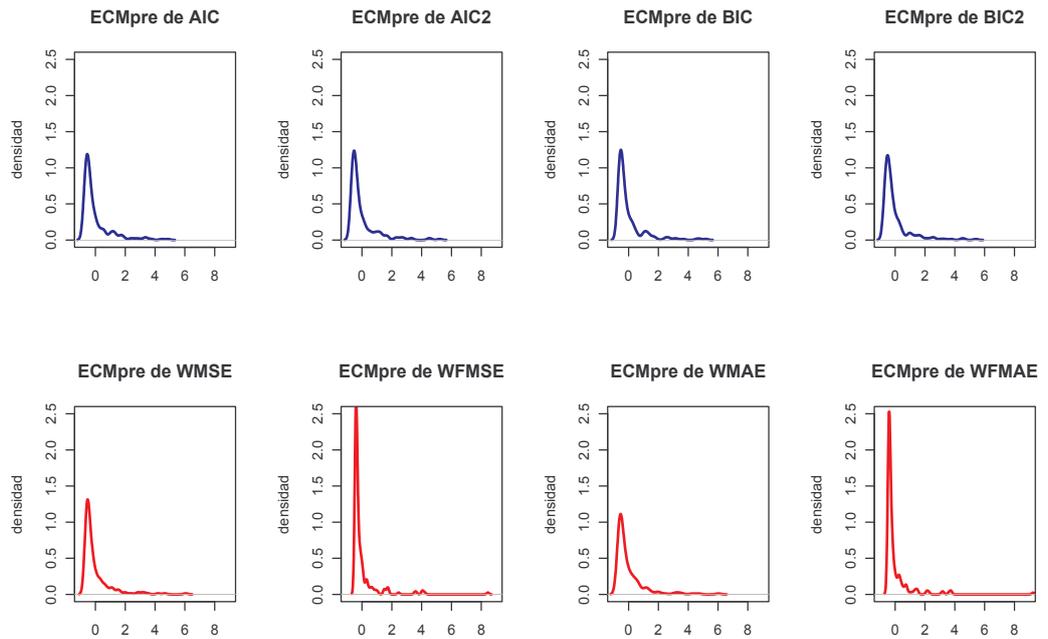


Figura 4.29: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).

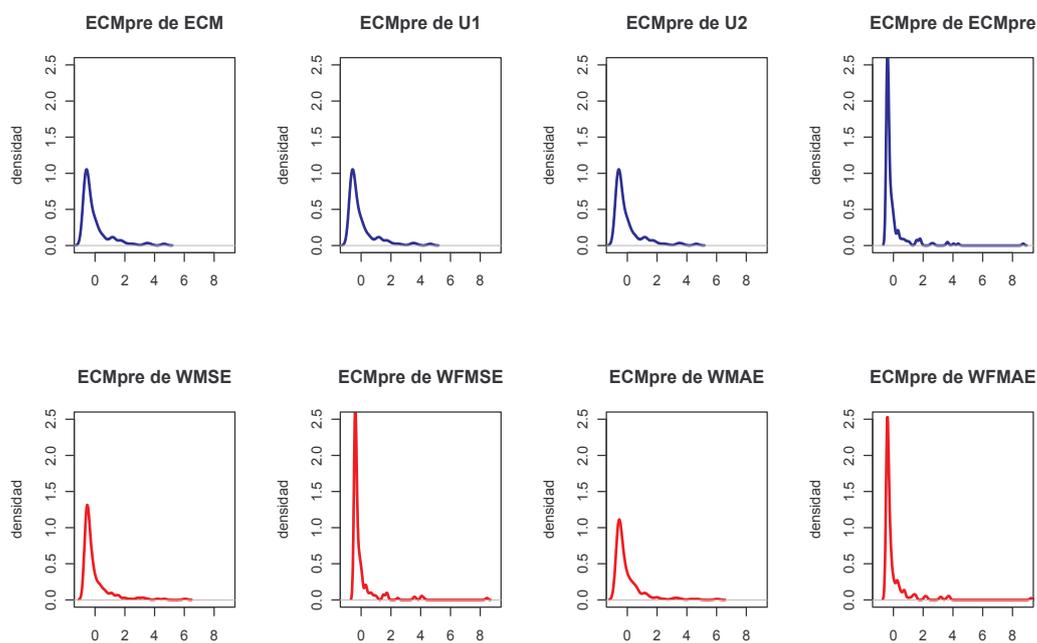


Figura 4.30: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).

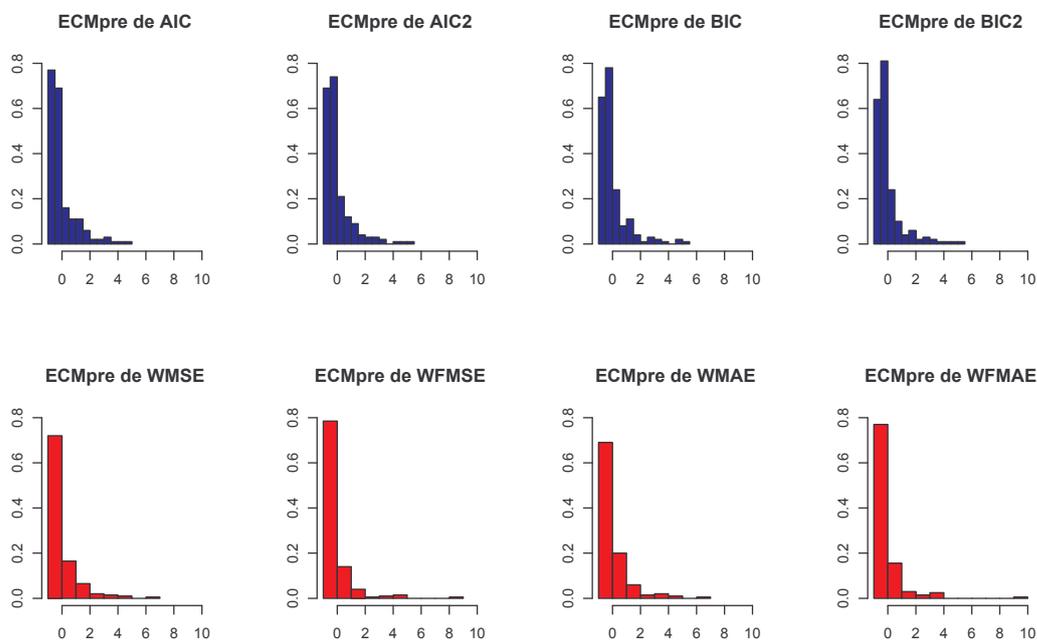


Figura 4.31: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,0).

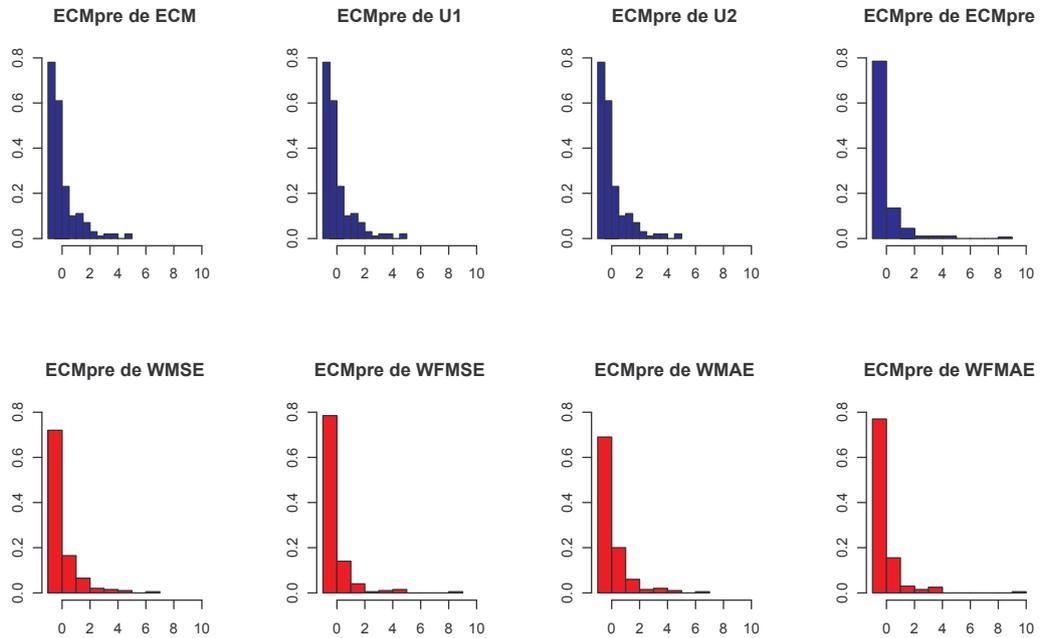


Figura 4.32: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie $ARIMA(1,1,0)$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	17.3283	17.2988	17.9175	17.4003	16.8926	17.9175
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	17.9175	4.1855	21.0318	4.3223	21.0362	4.8455

Tabla 4.19: ECM global de predicción, ECMGpre, para cada criterio en modelos $ARIMA(1,1,0)$.

A la vista de la tabla 4.19 se observa claramente, la diferencia de los criterios propuestos para predicción con los criterios clásicos, por lo que se puede decir que los criterios propuestos para predicción alcanzan mejores resultados al menos en los modelos $ARIMA(1,1,0)$, aunque veamos que el criterio del ECMpre mejora un poco estos criterios propuestos, observamos en la tabla 4.17, que las primeras predicciones se consiguen con menor error con los criterios de predicción propuestos, es por ello que sean preferibles. Además, si comparamos los resultados obtenidos por los criterios propuestos en la tabla 4.7, podemos ver la gran mejoría conseguida con este procedimiento de actuación, lo que nos lleva a determinar que para conseguir unas “buenas” predicciones hemos de considerar el modelo seleccionado por los criterios de predicción propuestos en la serie de interés sin preocuparnos por

el modelo de ajuste de la misma.

Este estudio se ha repetido cambiando los coeficientes del modelo y los parámetros siempre considerando que el modelo tenga tan sólo parte autorregresiva, y en todos los casos se ha llegado a un comportamiento similar, por tanto, en el siguiente estudio que se expone se ha cambiado a un modelo integrado de medias móviles, con el objeto de poder establecer el comportamiento de los criterios en este tipo de series y poder determinar la manera de proceder que mejor se ajuste para alcanzar el objetivo que se persigue a lo largo de todo este trabajo.

Modelo ARIMA(0, 1, 1) con los siguientes coeficientes, $\theta_1 = 0,8$ (Modelo 73)

	AIC	AIC2	ECM	BIC	BIC2
Valor	2.32566	-0.4915852	0.548911	2.42183	-0.4029472
Modelo	320	309	267	307	73

Tabla 4.20: Criterios Clásicos de ajuste para ARIMA(0,1,1).

	U_1	U_2	ECMpre
Valor	0.02050853	0.02051285	0.081778
Modelo	162	162	46

Tabla 4.21: Criterios Clásicos de predicción para ARIMA(0,1,1).

	Ajuste		Predicción	
	WMSE	WMAE	WFMSE	WFMAE
Valor	0.1944577	0.3061104	0.04636347	0.1459059
Modelo	115	248	46	46

Tabla 4.22: Criterios propuestos para ARIMA(0, 1, 1).

Para comprobar con que modelo exactamente se corresponde cada uno de los valores mostrados en las tablas 4.20, 4.21 y 4.22 podemos consultar el apéndice C. Cada criterio ha seleccionado dentro del modelo mínimo para cada serie el mínimo de éstos. Sólo en el caso del ECMpre y los criterios de predicción propuestos se ha seleccionado el mismo modelo, en el caso del criterio de Theil ambos han coincidido también en la elección.

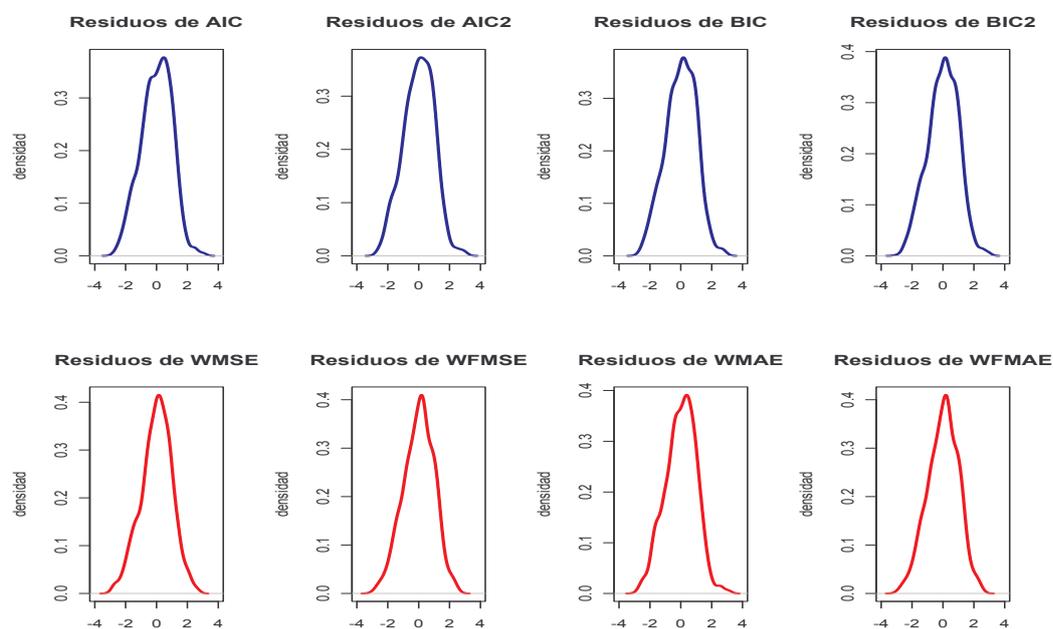


Figura 4.33: Residuos obtenidos con los modelos de los criterios de ajuste de ARIMA(0,1,1).

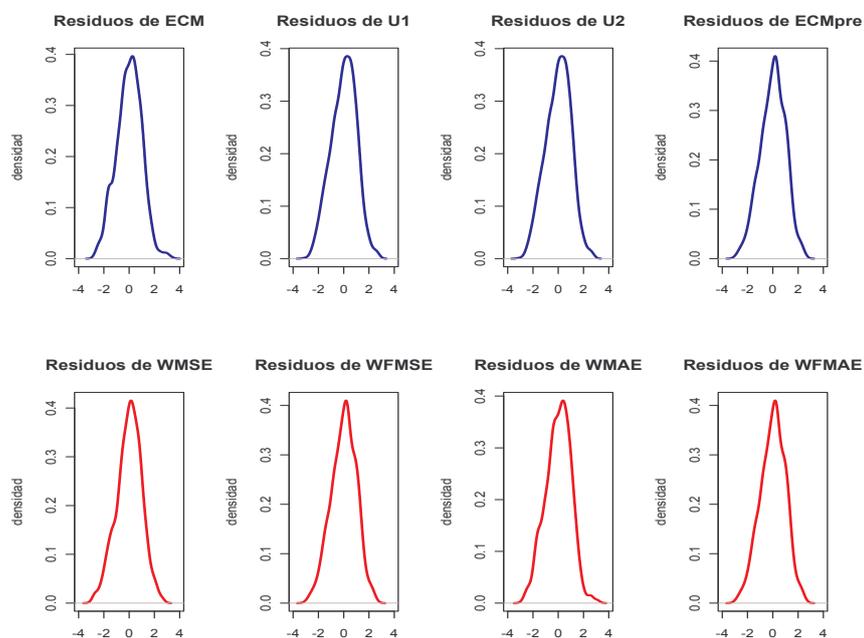


Figura 4.34: Residuos obtenidos con los modelos de los criterios de predicción de ARIMA(0,1,1).

Criterio	Media	Varianza
AIC	-0.4313	13.1613
AIC2	-0.4023	13.1340
ECM	-0.3515	14.42
BIC	-0.4192	12.8566
BIC2	-0.4230	12.4956
U1	-0.3856	13.1008
U2	-0.3856	13.1008
ECMpre	-0.3027	21.0679
WMSE	-0.4980	13.4967
WFMSE	-0.3027	21.0679
WMAE	-0.4127	13.3953
WFMAE	-0.3027	21.0679

Tabla 4.23: Medidas descriptivas de los residuos en el conjunto de comprobación para ARIMA(0,1,1).

Observando los valores de las medias de los residuos obtenidos en el conjunto de comprobación, no se aprecian grandes diferencias, nos inclinaríamos por tanto, a buscar el criterio donde se alcanza la varianza mínima, en ese caso el criterio sería al BIC2 o BIC, no obstante, el resto de criterios a excepción del ECMpre, WFMSE y WFMAE, no difieren significativamente con el resto, por lo que no se aprecian grandes mejorías. Esto puede verse también en los gráficos 4.35 y 4.36, donde se puede ver que los histogramas son muy similares entre sí. Veamos en la tabla 4.24 que ocurre con las predicciones si se consideran de manera individual, en la que se puede apreciar a grandes rasgos como hay una diferencia significativa en cuanto al aumento del error conforme la predicción se aleja del instante inicial.

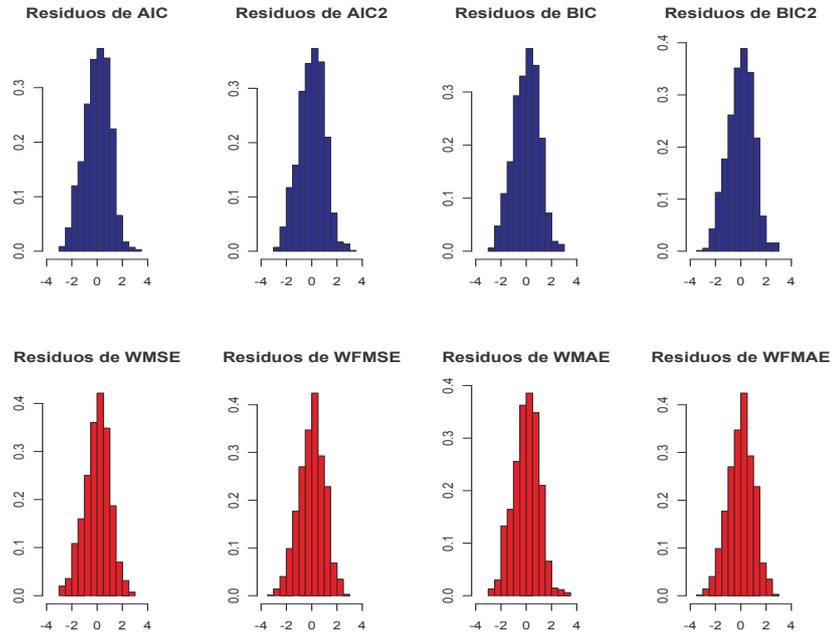


Figura 4.35: Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de ARIMA(0,1,1).

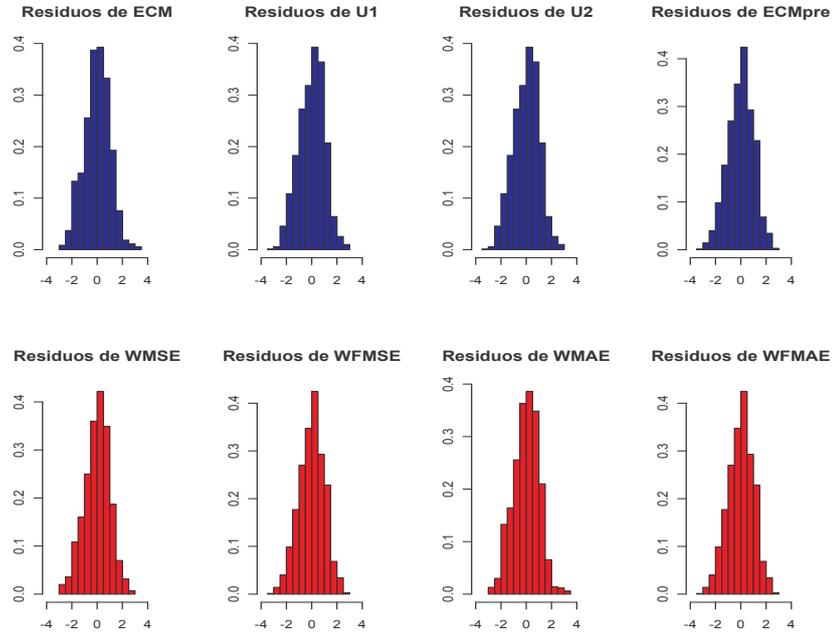


Figura 4.36: Histograma de los residuos obtenidos con los modelos de los criterios de predicción de ARIMA(0,1,1).

Las semejanzas entre los resultados obtenidos con unos criterios y otros son evidentes, este comportamiento continua a la vista de los resultados expuestos en la tabla 4.24 donde se recogen los valores medios del ECMpre, ECMGpre, en función de cada predicción individualmente, donde se puede comprobar que los mejores resultados para las primeras predicciones se obtienen con el modelo seleccionado por el criterio BIC2, seguido del BIC, AIC, AIC2 y ECM, sin que existan diferencias significativas entre los mismos.

Criterio	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
AIC	1.3758	5.5820	8.8530	12.7628	17.4317	21.7944	25.6318
AIC2	1.3896	5.6717	8.8809	12.5947	17.2244	21.6389	25.6706
ECM	1.4122	5.7424	9.3891	13.8276	18.8708	23.8825	28.6775
BIC	1.3536	5.4744	8.6805	12.4238	16.8519	21.1978	25.2447
BIC2	1.3154	5.2753	8.3775	12.1026	16.4115	20.6379	24.6014
U1	1.4759	5.7617	8.9322	12.7270	16.8311	21.2760	25.7423
U2	1.4759	5.7617	8.9322	12.7270	16.8311	21.2760	25.7423
ECMpre	3.1291	16.5222	25.0989	25.4413	25.3479	25.6716	26.9056
WMSE	2.0071	6.0554	9.2019	13.2964	17.8102	21.8840	25.9578
WFMSE	3.1291	16.5222	25.0989	25.4413	25.3479	25.6716	26.9056
WMAE	1.3676	5.6005	8.9855	12.9901	17.6280	22.0160	26.3717
WFMAE	3.1291	16.5222	25.0989	25.4413	25.3479	25.6716	26.9056

Tabla 4.24: ECMpre global para cada predicción realizada en el conjunto de comprobación en ARIMA(0,1,1).

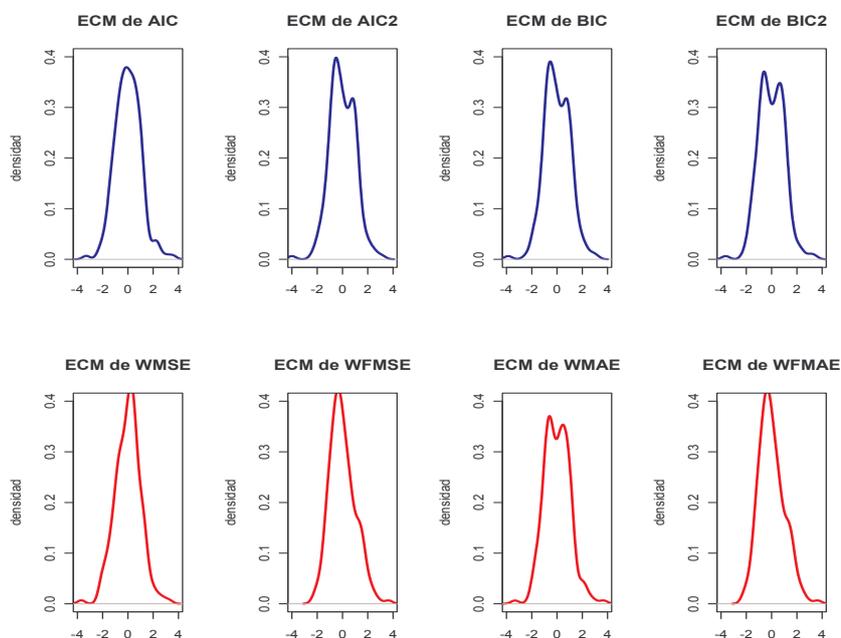


Figura 4.37: ECM de ajuste de los modelos obtenidos con los criterios de ajuste en ARIMA(0,1,1).

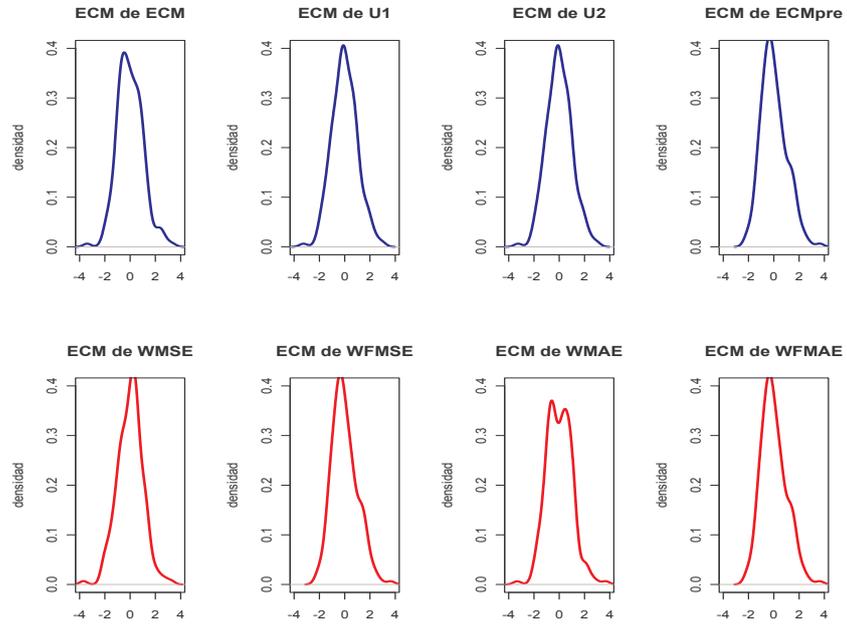


Figura 4.38: ECM de ajuste de los modelos obtenidos con los criterios de predicción en ARIMA(0,1,1).

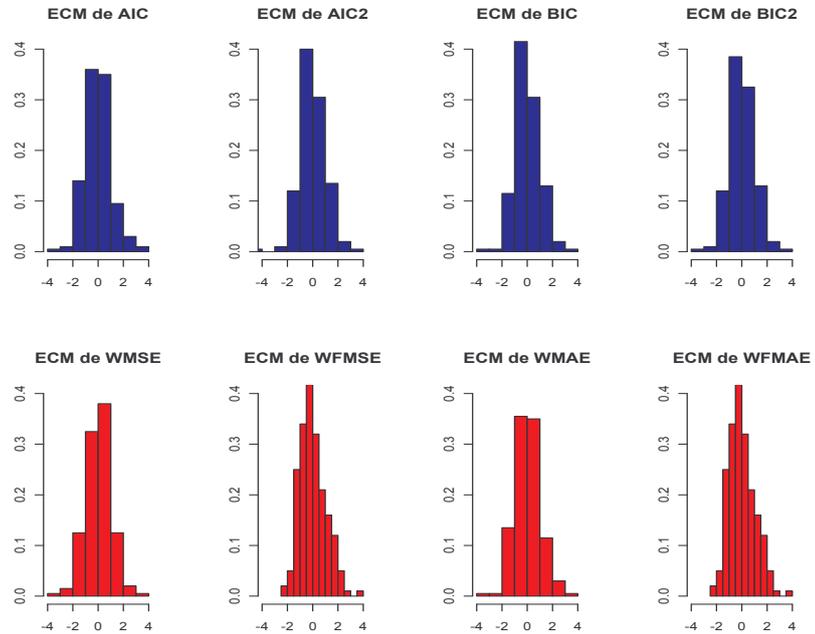


Figura 4.39: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en ARIMA(0,1,1).

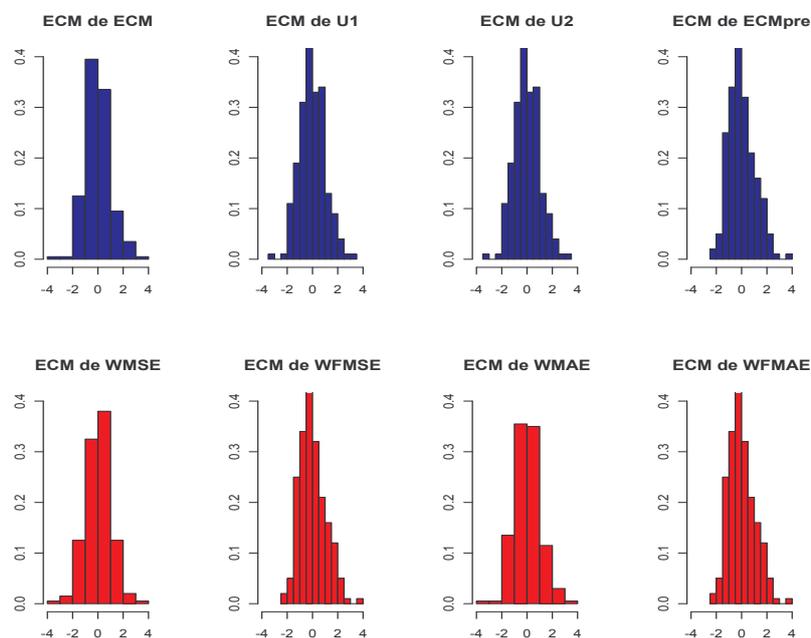


Figura 4.40: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicci3n en ARIMA(0,1,1).

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8536	0.9394	0.8469	0.9510	0.9906	0.9525
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.9525	2.7422	1.5383	2.7422	0.8749	2.7422

Tabla 4.25: El ECM global de ajuste, ECMG, para cada criterio en modelos ARIMA(0,1,1).

En la tabla 4.25 se muestran los ECM de los residuos obtenidos en el conjunto de entrenamiento, es decir, los residuos de ajuste del modelo, donde se puede observar que el mejor ajuste se ha obtenido con el modelo seleccionado por el criterio AIC, ECM y por el criterio WMAE. En los gr1ficos 4.41, 4.42, 4.43 y 4.44 se puede ver de manera gr1fica el comportamiento de los valores del ECM en cada una de las series evaluadas en los criterios.

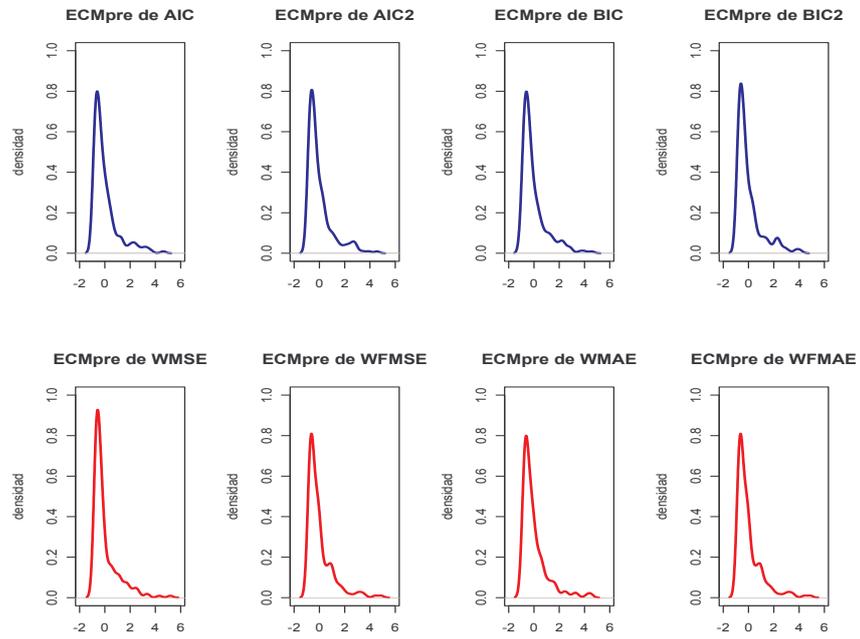


Figura 4.41: ECM de predicción de los modelos obtenidos con los criterios de ajuste en ARIMA(0,1,1).

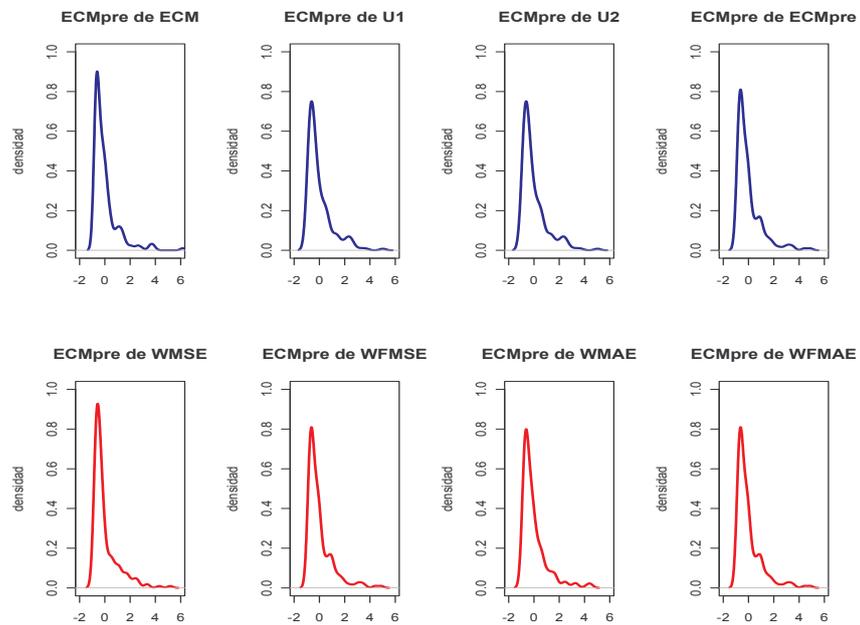


Figura 4.42: ECM de predicción de los modelos obtenidos con los criterios de predicción en ARIMA(0,1,1).

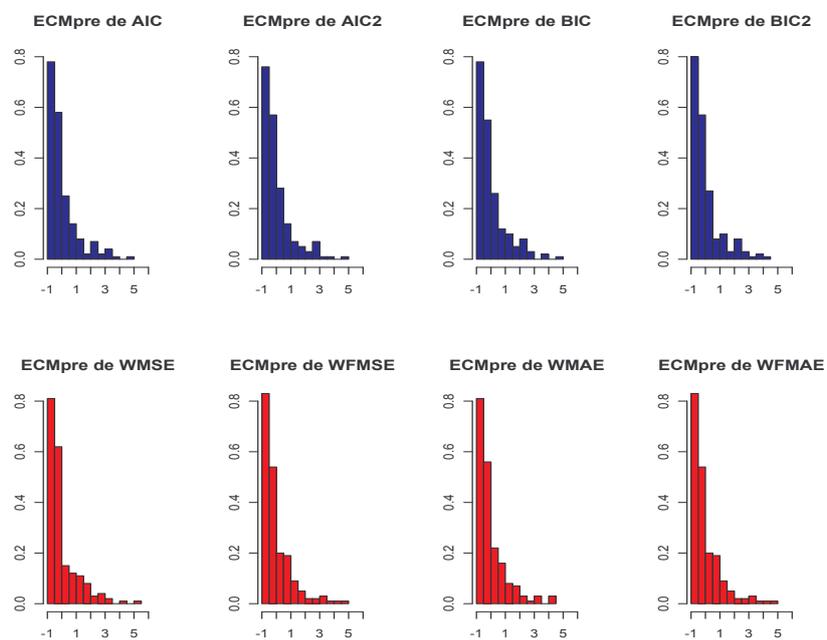


Figura 4.43: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en ARIMA(0,1,1).

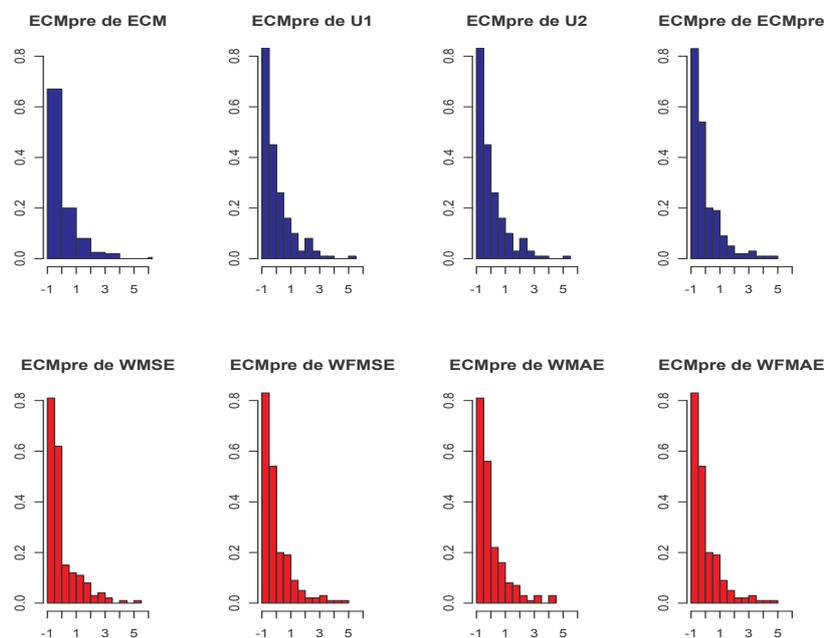


Figura 4.44: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en ARIMA(0,1,1).

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG _{pre}	13.3474	13.2958	14.5436	13.0324	12.6745	13.2494
	U2	ECM _{pre}	WMSE	WFMSE	WMAE	WFMAE
ECMG _{pre}	13.2494	21.1595	13.7447	21.1595	13.5656	21.1595

Tabla 4.26: ECM global de predicción, ECMG_{pre}, para cada criterio en modelos ARIMA(0,1,1).

Se puede determinar que considerando todas las series y realizando las predicciones con el modelo seleccionado no se obtienen resultados significativamente mejores, por tanto, al igual que se hizo en el caso anterior, se considerarán dos grupos de series aquellas a las que hemos llamado series correctas y las incorrectas. Los criterios que han sido capaces de detectar en alguna serie (al menos 2 de las 200) el modelo generador de los datos han sido AIC2 (9,5%), BIC (4%), BIC2 (55,5%) y WFMAE (1%).

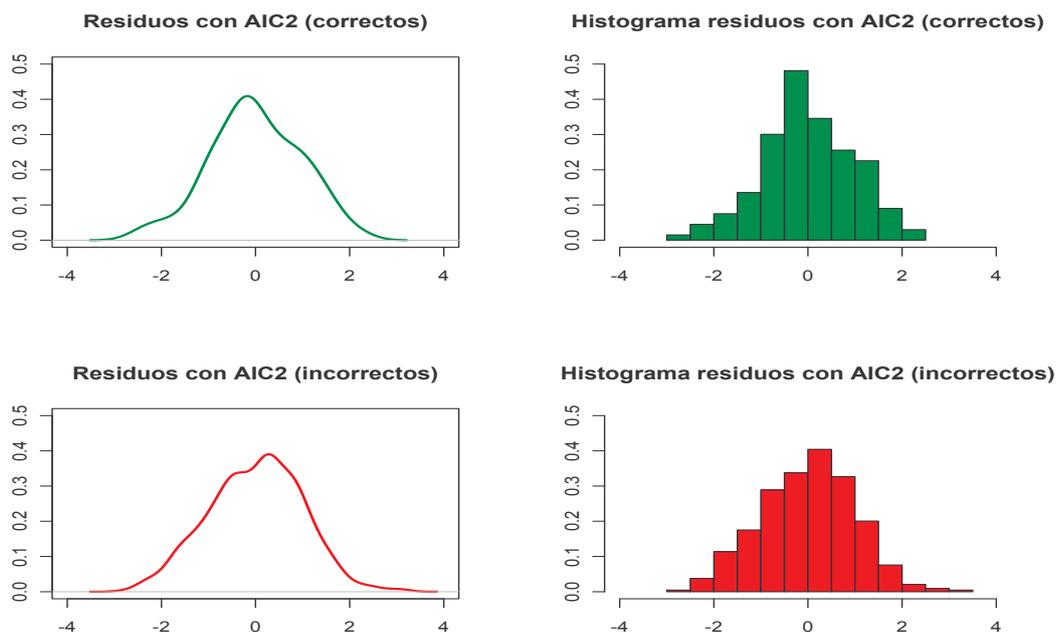


Figura 4.45: Residuos de las series correctas e incorrectas con el criterio AIC2 en el ARIMA(0,1,1).

	Media	Varianza	ECMG	ECMGpre
Correctos	0.3275	6.4995	0.9704	6.6068
Incorrectos	-0.5364	14.3319	0.8688	14.6197

Tabla 4.27: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(0,1,1) con el criterio AIC2.

En la tabla 4.7, se aprecia una mejor predicción en el caso de las series correctas, sin embargo, el mejor ajuste se obtiene en las series incorrectas (gráfico 4.46), nuestro objetivo sería conseguir la mejor predicción posible, sobre todo en las primeras observaciones a las que hemos dotado de una mayor importancia, para ello analizamos los resultados obtenidos en la tabla 4.28, donde se aprecia una leve mejoría en la primera de las predicciones en las incorrectas no obstante los errores aumentan mucho más en las siguientes predicciones lo que lleva a una mayor varianza como puede observarse en el gráfico 4.46 que en las series correctas, donde el aumento del error es de un crecimiento más moderado.

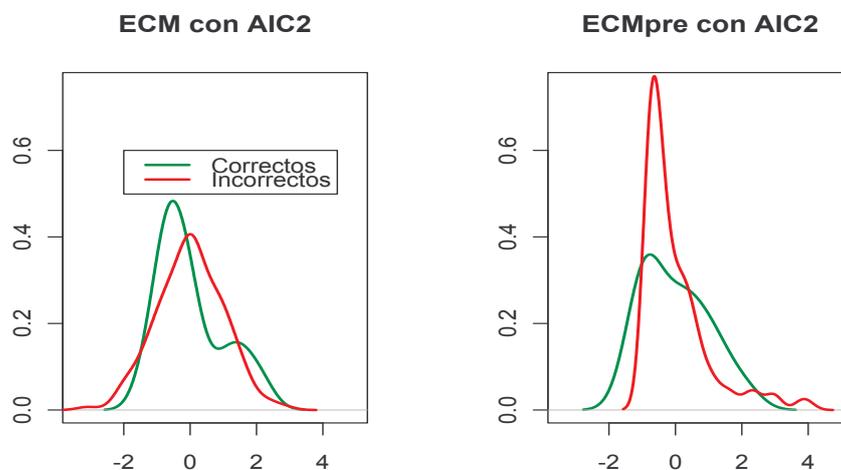


Figura 4.46: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en ARIMA(0,1,1).

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	1.4115	5.2083	5.6573	5.9785	6.3083	8.7046	12.9789
Incorrectos	1.3847	5.7623	9.5188	13.7767	18.9186	24.0605	28.9160

Tabla 4.28: ECM global de predicción en ARIMA(0,1,1) con el criterio AIC2 en el conjunto de comprobación.

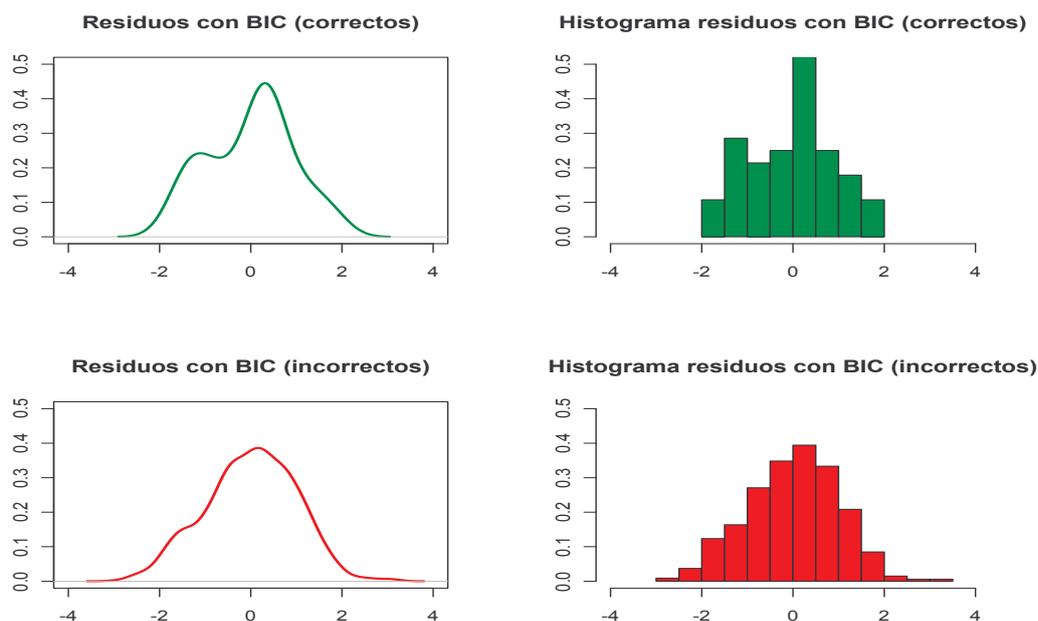


Figura 4.47: Residuos de las series correctas e incorrectas con el criterio BIC en el ARIMA(0,1,1).

	Media	Varianza	ECMG	ECMGpre
Correctos	0.5344	6.5128	0.9567	6.7984
Incorrectos	-0.4238	13.0234	0.8711	13.2239

Tabla 4.29: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para ARIMA(0,1,1) con el criterio BIC.

En el caso del criterio BIC, vemos en la tabla 4.29 como la media es menor en el caso de las series incorrectas pero a su vez éstas tienen también mayor varianza, lo que nos lleva a obtener un ECMGpre mayor, esto quiere decir que dentro de este grupo hay series no son homogéneas lo que hace que

la varianza sea mayor y el resultado global de la predicción tampoco sea el deseado, puede comprobarse en la gráfica 4.48, ese aumento de la varianza a pesar de conseguir una mejor media.

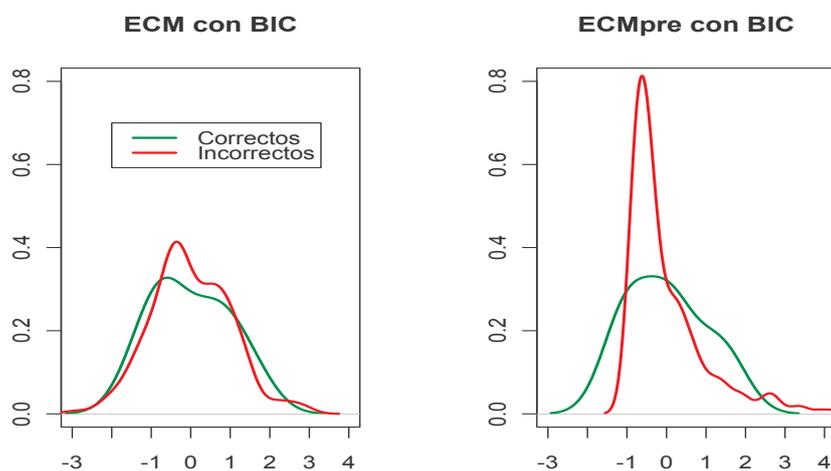


Figura 4.48: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC en ARIMA(0,1,1).

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	1.7744	6.7553	5.2302	4.2318	7.7301	10.1056	11.7616
Incorrectos	1.3372	5.4232	8.7583	12.5662	17.0469	21.6286	25.8071

Tabla 4.30: ECM global de predicción en ARIMA(0,1,1) con el criterio BIC en el conjunto de comprobación.

A la vista de estos ECMGpre individual (tabla 4.30), se puede ver como en las series incorrectas las dos primeras predicciones, que son a las que se les ha dado un mayor peso, se mejoran con respecto a las series correctas, sin embargo el resto de las mismas, se obtienen mejores en el grupo de series correctas. También es destacable el hecho ya comentado en alguna ocasión del aumento mucho mayor en el error cometido en las predicciones entre unas y otras en el grupo de las series que no han sido catalogadas con el modelo con el que han sido simuladas.

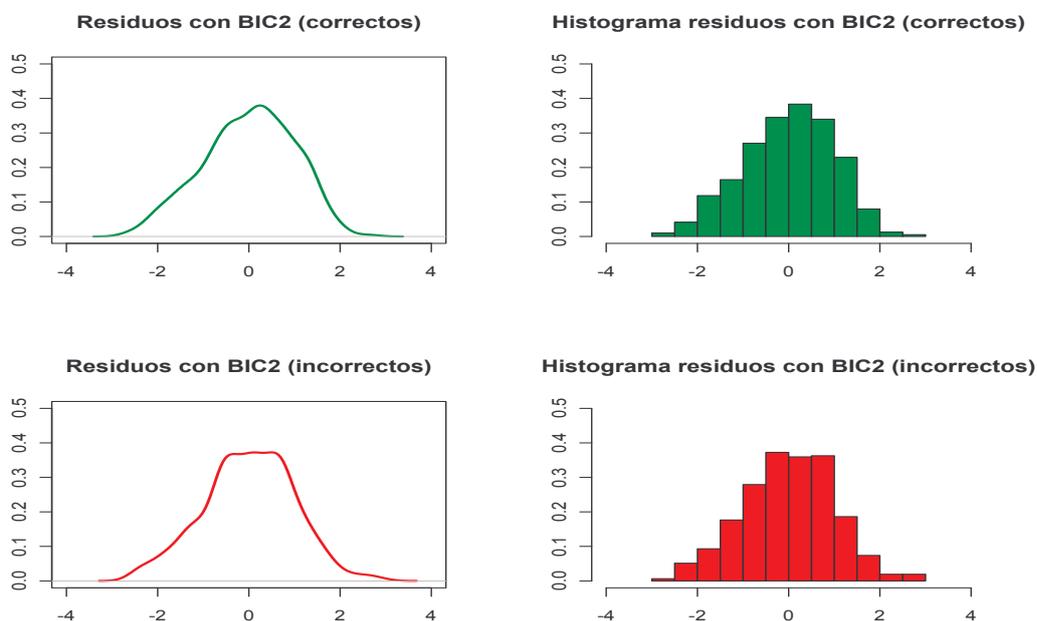


Figura 4.49: Residuos de las series correctas e incorrectas con el criterio BIC2 en el ARIMA(0,1,1).

	Media	Varianza	ECMG	ECMGpre
Correctos	-0.7627	11.1503	0.9859	11.7320
Incorrectos	0.0262	13.4735	0.8924	13.4742

Tabla 4.31: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(0,1,1) con el criterio BIC2.

Como viene siendo habitual en este modelo objeto de estudio, las mejores predicciones se consiguen en el conjunto de series correctas, y sin embargo, el mejor ajuste en el de las series incorrectas, en el gráfico 4.49 se puede apreciar la leve mejoría en las series correctas ante las incorrectas. En la gráfica 4.50 se observa como la diferencia en cuanto al comportamiento en un grupo o en otro en función de los valores obtenidos en el ECM de ajuste y predicción por cada una de las series no es significativo, ya que ambas curvas tienen un comportamiento muy similar.

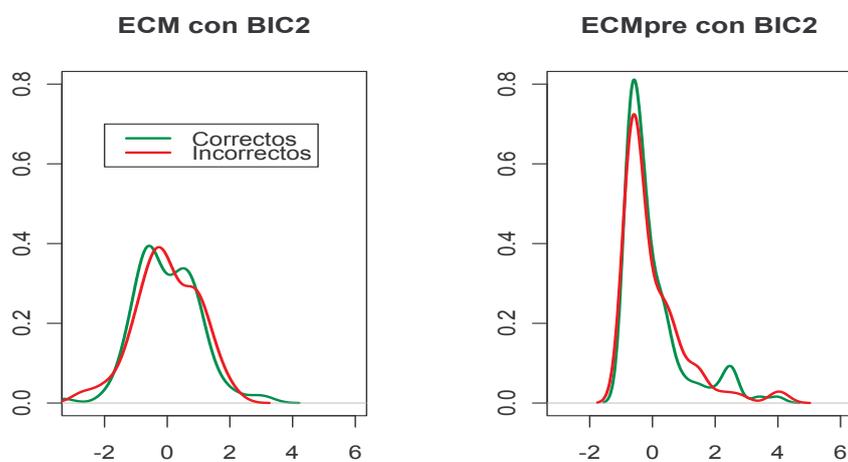


Figura 4.50: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en ARIMA(0,1,1).

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	1.5402	5.4922	7.9418	10.9033	14.0427	18.8382	23.3658
Incorrectos	1.0323	5.0897	8.8088	12.9387	18.6481	22.2795	25.5220

Tabla 4.32: ECM global de predicción en ARIMA(0,1,1) con el criterio BIC2 en el conjunto de comprobación.

Ante el criterio BIC2, en esta ocasión el comportamiento entre un grupo u otro no tiene diferencias ni en función del ajuste ni de predicción. Tampoco hay diferencias importantes con el resto de los criterios estudiados, se mantienen más o menos los mismos resultados. Estudiamos por último el criterio WFMAE donde el porcentaje de series que han sido determinadas correctamente es tan sólo del 1% como puede apreciarse en la gráfica 4.51.

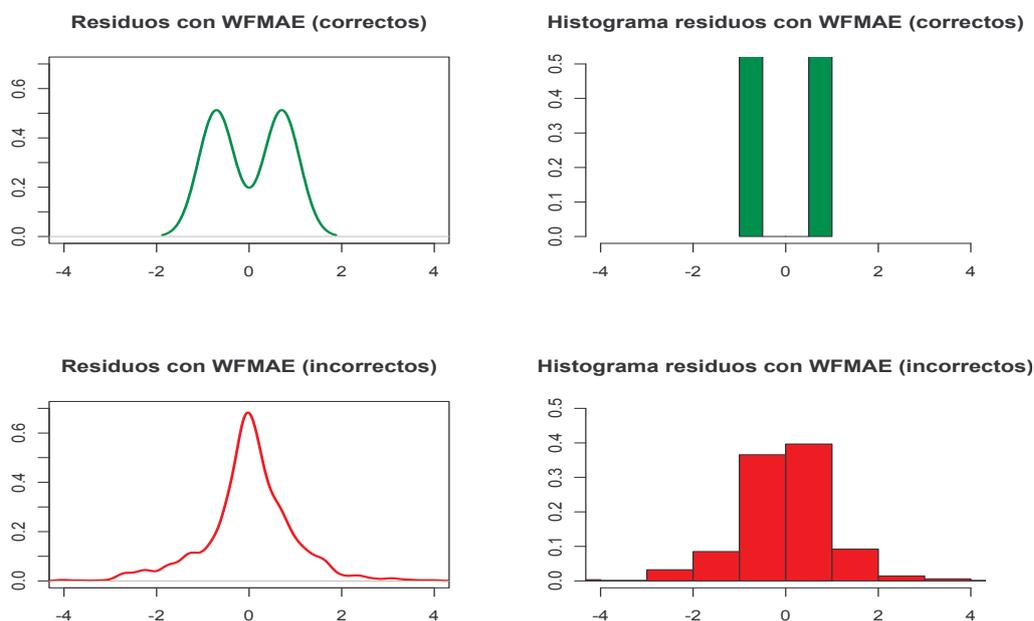


Figura 4.51: Residuos de las series correctas e incorrectas con el criterio WFMAE en el ARIMA(0,1,1).

	Media	Varianza	ECMG	ECMG _{pre}
Correctos	-0.4170	1.8686	1.0323	2.0425
Incorrectos	-0.1662	4.2375	3.6980	4.2651

Tabla 4.33: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(0,1,1) con el criterio WFMAE.

En la tabla 4.33 se observa una clara mejoría en el grupo de las series catalogadas como correctas, tanto en predicción como en ajuste, sin embargo, hay que tener en cuenta que el grupo de las series correctas se compone únicamente por dos de las series, lo que explica que la varianza sea mucho inferior al grupo de las series incorrectas, pero ante este criterio no se casar ninguna conclusión de validez estadística.

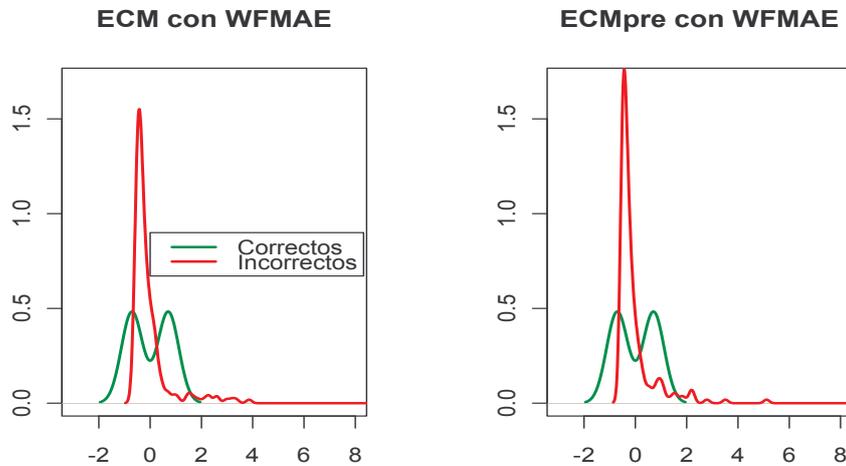


Figura 4.52: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en ARIMA(0,1,1).

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.0702	0.1059	0.0340	0.7212	0.4292	4.0757	8.8612
Incorrectos	0.6419	1.6530	1.9967	3.2675	4.7811	7.1659	10.3496

Tabla 4.34: ECM global de predicción en ARIMA(0,1,1) con el criterio WFMAE en el conjunto de comprobación.

En la tabla 4.34 donde se muestran los ECMGpre individual, se consiguen predicciones mejores en el grupo de las series bien catalogadas, pero insistimos en que no son conclusiones determinantes ante el tamaño muestral del grupo de las series correctas.

Tras analizar las representaciones gráficas y resultados expuestos, no es posible determinar una buena política de actuación, donde se garanticen la obtención de buenos resultados en todos los casos, es decir, no se pudo determinar si el mejor funcionamiento en cuanto a predicción se refiere se obtendría cuando un criterio es capaz de determinar el modelo generador de una serie o por el contrario asociar un modelo distinto. Continuamos con el estudio sobre el conjunto de las 200 series considerando para cada una de ellas el modelo que determine cada criterio en cada caso, es decir, sin establecer un modelo fijo para todas.

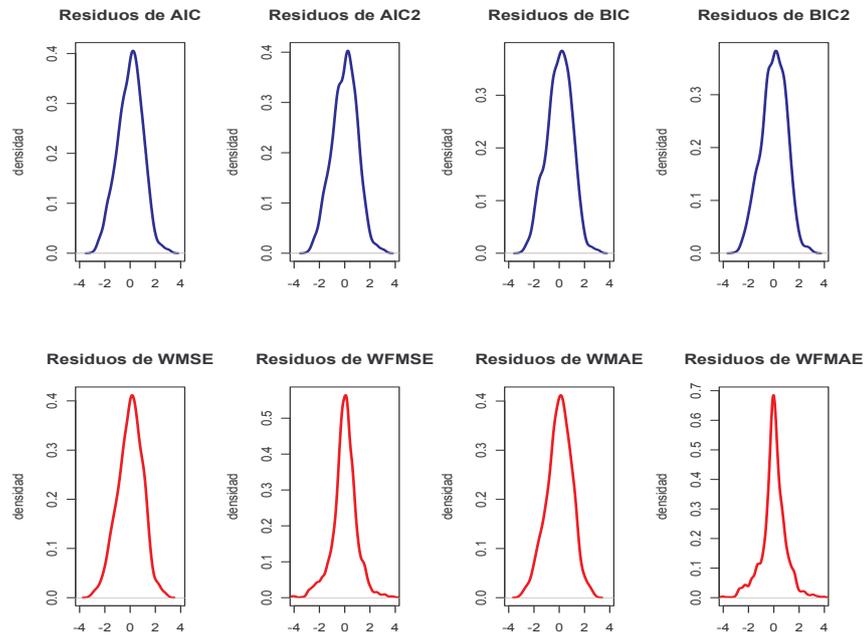


Figura 4.53: Residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).

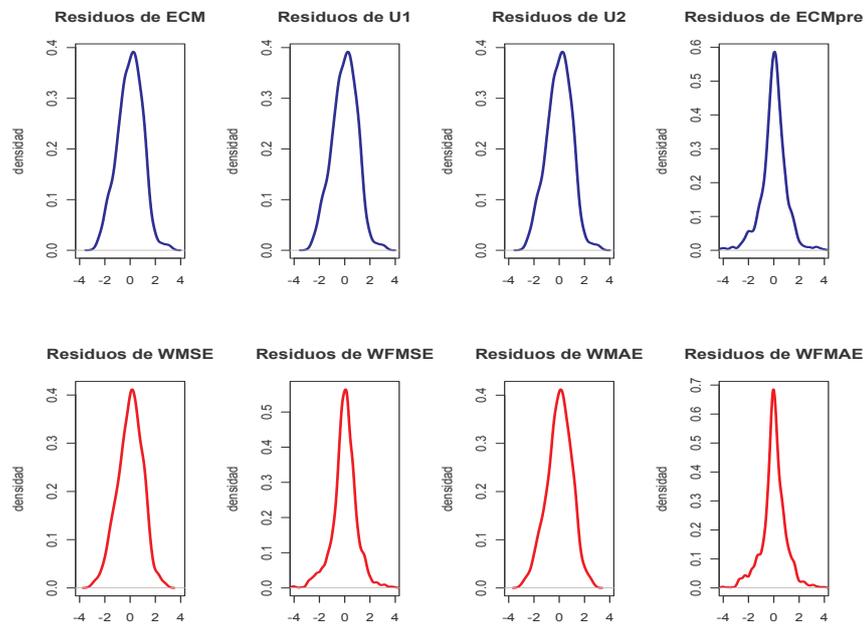


Figura 4.54: Residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).

Criterio	Media	Varianza
AIC	-0.421	14.1849
AIC2	-0.4543	13.6520
ECM	-0.4482	14.6695
BIC	-0.4085	12.8
BIC2	-0.4116	12.3378
U1	-0.4482	14.6695
U2	-0.4482	14.6695
ECMpre	-0.1430	3.7704
WMSE	-0.4795	14.7531
WFMSE	-0.1526	3.9478
WMAE	-0.3271	15.3974
WFMAE	-0.1687	4.2144

Tabla 4.35: Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para ARIMA(0,1,1).

Tan sólo con las medidas descriptivas se puede observar una clara mejoría utilizando los criterios ECMpre, WFMSE y WFMAE, donde no sólo la media es significativamente menor, sino también la varianza.

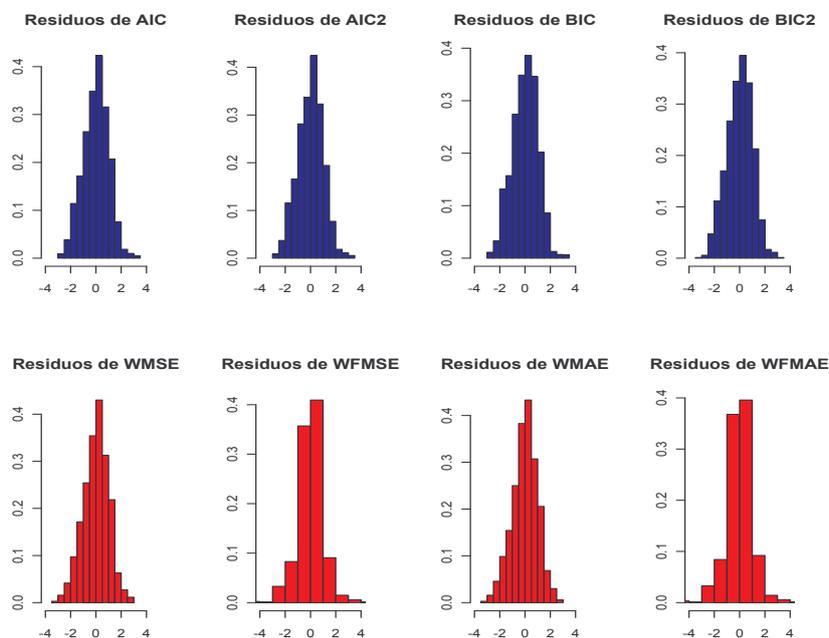


Figura 4.55: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).

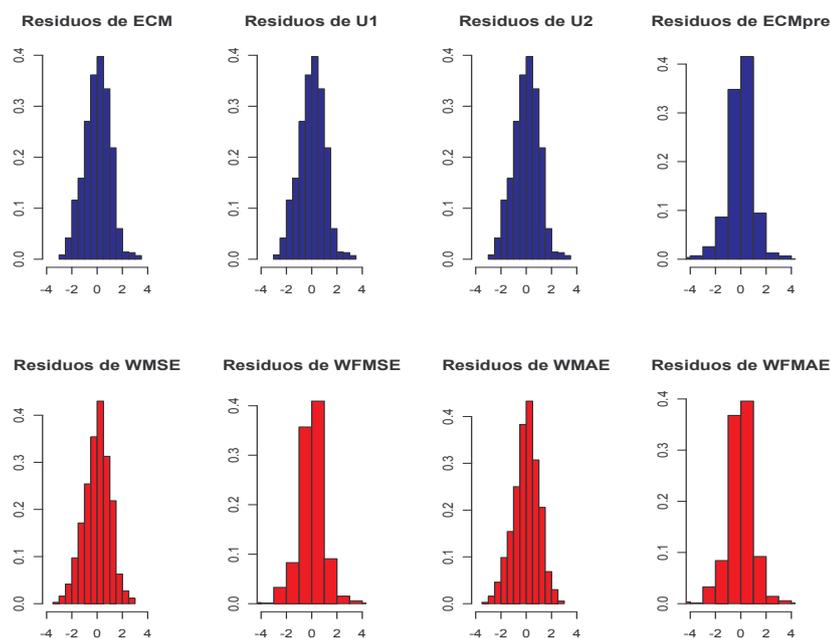


Figura 4.56: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).

Criterio	ECMpre1	ECMpre2	ECMpre3	ECMpre4	ECMpre5	ECMpre6	ECMpre7
AIC	1.4013	5.6776	9.1331	13.2452	18.3368	23.8609	28.88
AIC2	1.3872	5.7097	9.1520	13.0359	17.7207	22.6017	27.4020
ECM	1.3799	5.6123	9.1747	13.6046	19.1342	25.0395	30.1477
BIC	1.3547	5.4765	8.6172	12.2328	16.6742	21.1677	25.2453
BIC2	1.3142	5.3191	8.3276	11.8091	16.0921	20.3695	24.3253
U1	1.3799	5.6123	9.1747	13.6046	19.1342	25.0395	30.1477
U2	1.3799	5.6123	9.1747	13.6046	19.1342	25.0395	30.1477
ECMpre	1.0561	2.3590	2.8732	3.7797	4.1430	5.2028	7.1221
WMSE	0.7254	6.4926	10.0355	14.3595	19.4824	24.3767	28.4087
WFMSE	0.7294	1.7376	2.0464	3.1258	4.4425	6.5371	9.1786
WMAE	1.6783	6.6719	10.8561	15.6651	20.4322	24.8791	28.3482
WFMAE	0.6361	1.6375	1.9770	3.2420	4.7376	7.1350	10.3347

Tabla 4.36: ECM globales para cada predicción realizada en el conjunto de comprobación en ARIMA(0,1,1).

Como ocurría con el modelo anterior los criterios ECMpre, y los propuestos de predicción, son los que consiguen mejores resultados (tabla 4.36), siendo los criterios propuestos de predicción donde se consiguen las mejores predicciones en los 4 primeros casos, con respecto al ECMpre donde es en las últimas predicciones donde mejoraría a los propuestos.

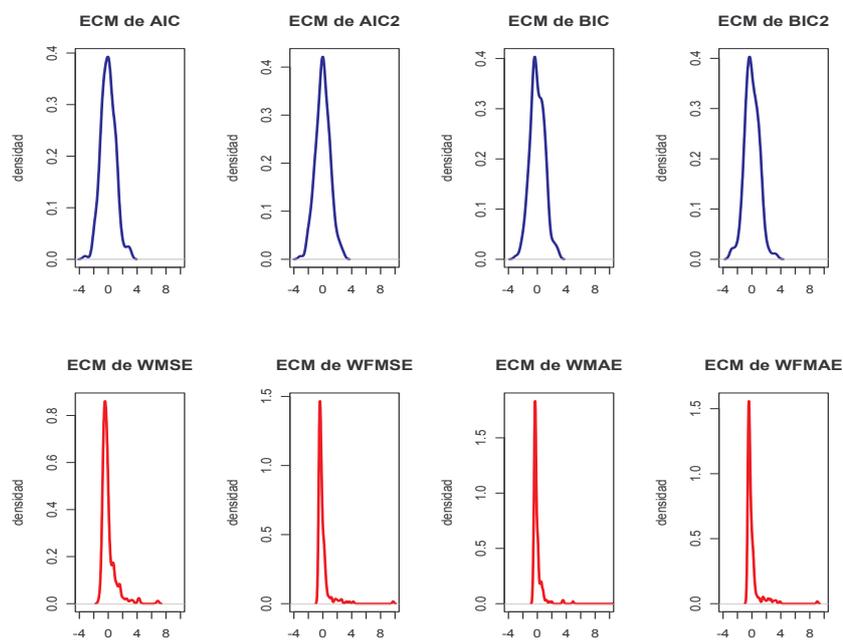


Figura 4.57: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).

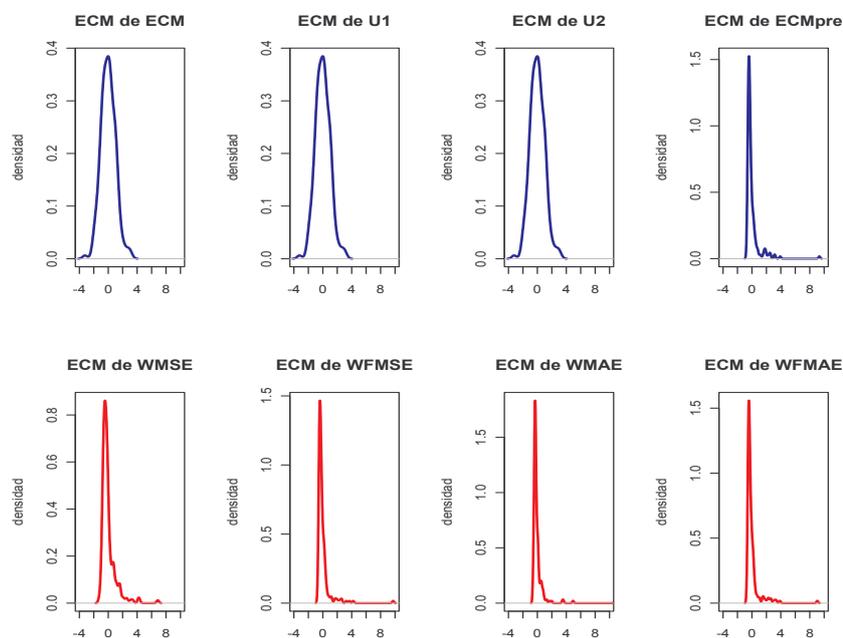


Figura 4.58: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).

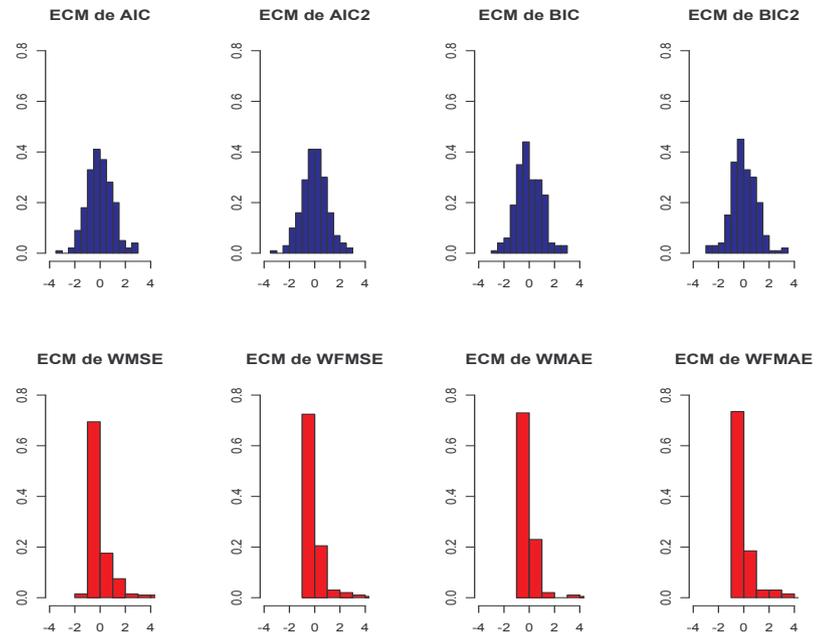


Figura 4.59: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $ARIMA(0,1,1)$.

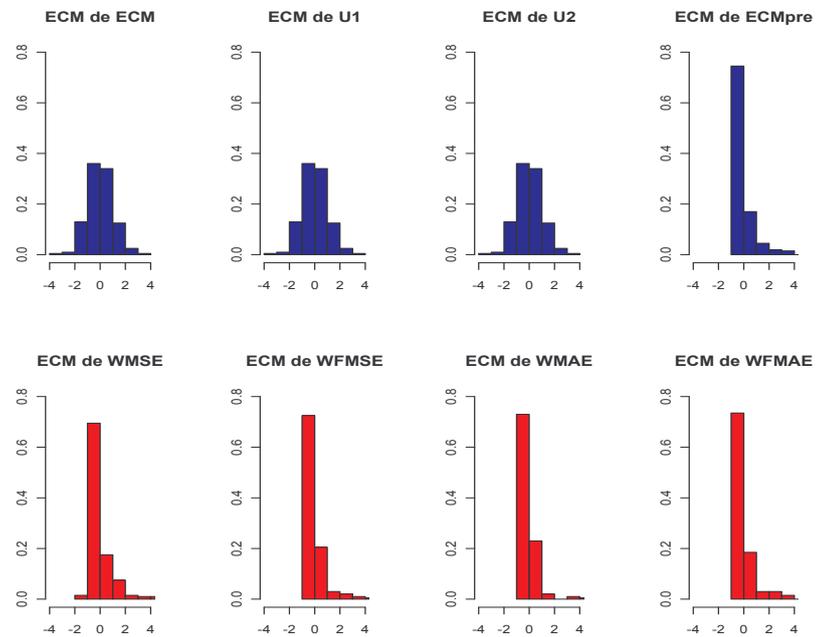


Figura 4.60: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $ARIMA(0,1,1)$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8376	0.8785	0.8289	0.8745	0.9443	0.8289
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8289	3.5470	1.1536	3.3068	1.5411	3.6714

Tabla 4.37: ECM global de ajuste, ECMG, para cada criterio en modelos ARIMA(0,1,1).

Atendiendo al ajuste (tabla 4.37) los criterios propuestos son los que peor funcionan, con una clara desventaja con el resto de los criterios estudiados, sin embargo, no se pretende conseguir un buen ajuste sino una buena capacidad predictiva, lo que puede verse en la tabla 4.38 donde son los criterios propuestos junto con el ECMpre donde se consiguen los resultados más favorables.

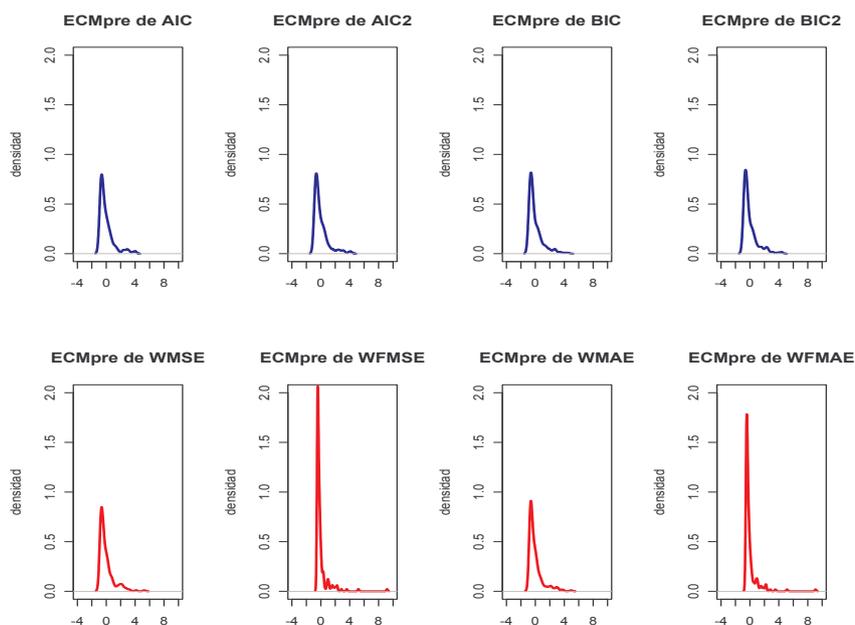


Figura 4.61: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).

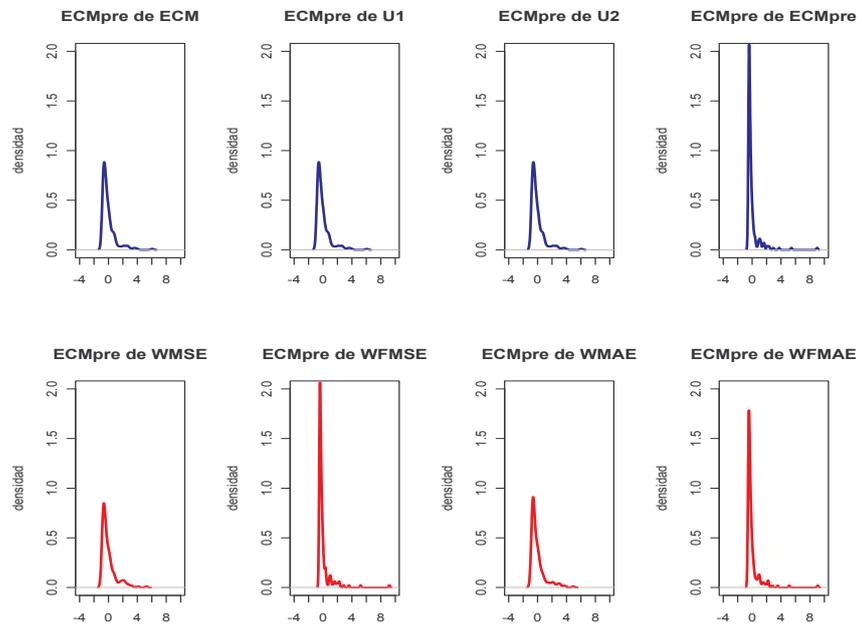


Figura 4.62: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie $ARIMA(0,1,1)$.

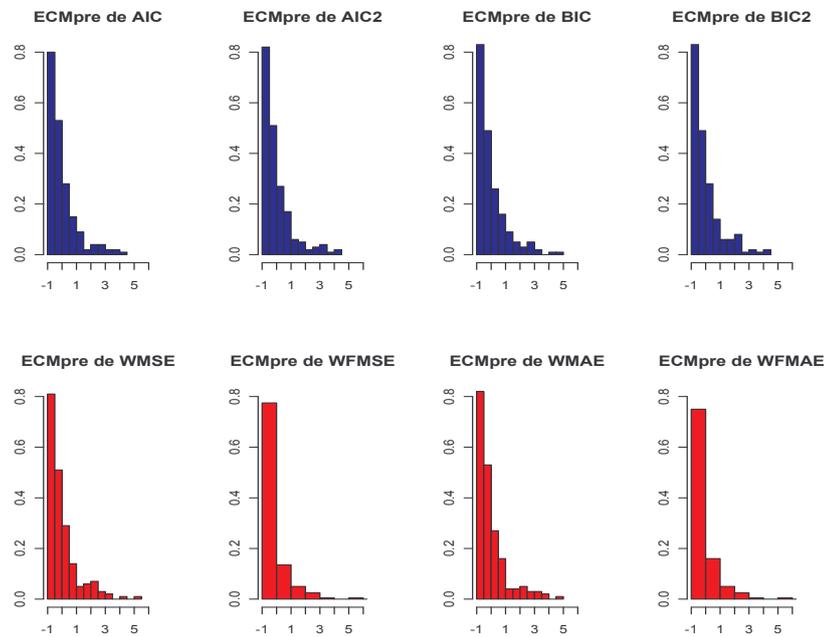


Figura 4.63: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie $ARIMA(0,1,1)$.

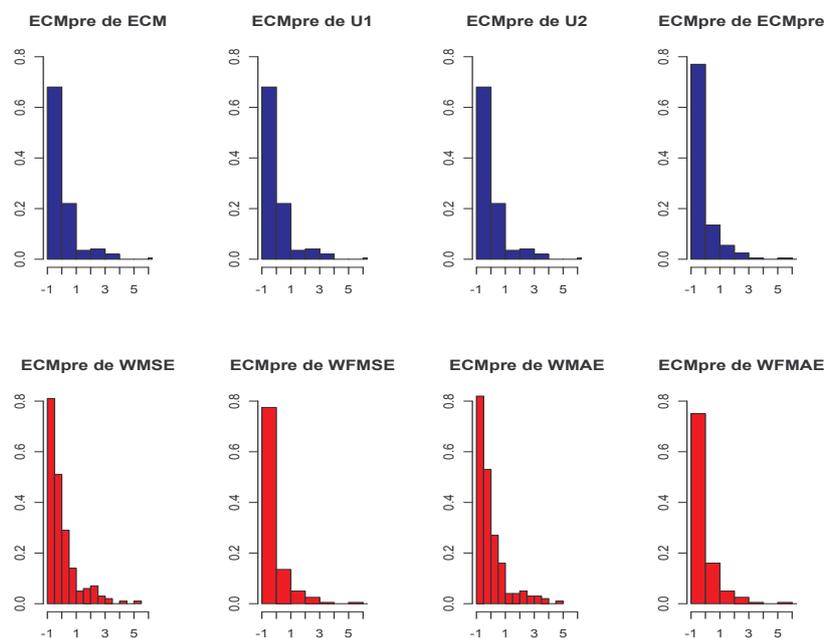


Figura 4.64: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(0,1,1).

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	14.3621	13.8584	14.8704	12.9669	12.5073	14.8704
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	14.8704	3.7909	14.9830	3.9711	15.5044	4.2429

Tabla 4.38: ECM global de predicción, ECMGpre, para cada criterio en modelos ARIMA(0,1,1).

Por último, y para concluir con este modelo, se observa en la tabla 4.28 que el menor valor del ECMGpre se obtiene con el criterio ECMpre, sin embargo, es preferible y se consiguen mejores predicciones con los criterios propuestos, ya que, éstos no sólo obtienen una ventaja en cuanto a predicción global con el resto de criterios, sino que además lo consiguen sobre todo en las primeras observaciones estimadas como se veía en la tabla 4.36. Y si comparamos con los resultados obtenidos en las tablas 4.24 y 4.26, veremos la gran diferencia en cuanto al ECMGpre en los criterios de predicción propuestos mejorando significativamente con relación a las citadas tablas. Por tanto, y al igual que ocurría con los modelos ARIMA(1, 1, 0) para conseguir los mejores resultados predictivos se debe considerar alguno de los dos criterios propuestos de predicción y realizar las predicciones correspondientes

con el modelo seleccionado sin cuestionarse el modelo real de la serie y sin determinar un modelo único para tal fin.

Modelo ARIMA(1, 1, 1) cuyos coeficientes son, $\phi_1 = 0,7$ y $\theta_1 = 0,8$ (Modelo 181)

	AIC	AIC2	ECM	BIC	BIC2
Valor	2.438309	-0.3003134	0.619846	2.529445	-0.2372095
Modelo	301	249	249	301	181

Tabla 4.39: Criterios Clásicos de ajuste para ARIMA(1,1,1).

	U_1	U_2	ECMpre
Valor	4.314270e-3	4.304311e-3	0.064670
Modelo	321	321	228

Tabla 4.40: Criterios Clásicos de predicción para ARIMA(1,1,1).

	Ajuste		PREDICCIÓN	
	WMSE	WMAE	WFMSE	WFMAE
Valor	0.2323977	0.3416546	0.04608275	0.1346496
Modelo	246	162	228	228

Tabla 4.41: Criterios propuestos para ARIMA(1,1,1).

En esta ocasión, hay muchos criterios que han seleccionado el mismo modelo, por ejemplo, el criterio AIC y BIC han elegido el modelo 301 que se corresponde con un SARIMA(2, 1, 1)(2, 0, 0)₇, el criterio AIC2 y ECM también han empatado en el modelo, siendo en este caso el modelo 249 que es un SARIMA(2, 0, 1)(2, 0, 2)₇, los criterios de Theil, U_1 y U_2 , de nuevo han coincidido, decantándose por el 321 que es un SARIMA(2, 1, 2)(2, 0, 2)₇ y por último, los criterios ECMpre, WFMSE y WFMAE han seleccionado un SARIMA(2, 0, 0)(1, 1, 2)₇, modelo 228.

Veamos que ocurre con el comportamiento de las predicciones en este caso, para ello, como hemos hecho en ocasiones anteriores, representaremos la distribución de los residuos obtenidos en el conjunto de comprobación al aplicar para l predicción los modelos indicados en las tablas anteriores.

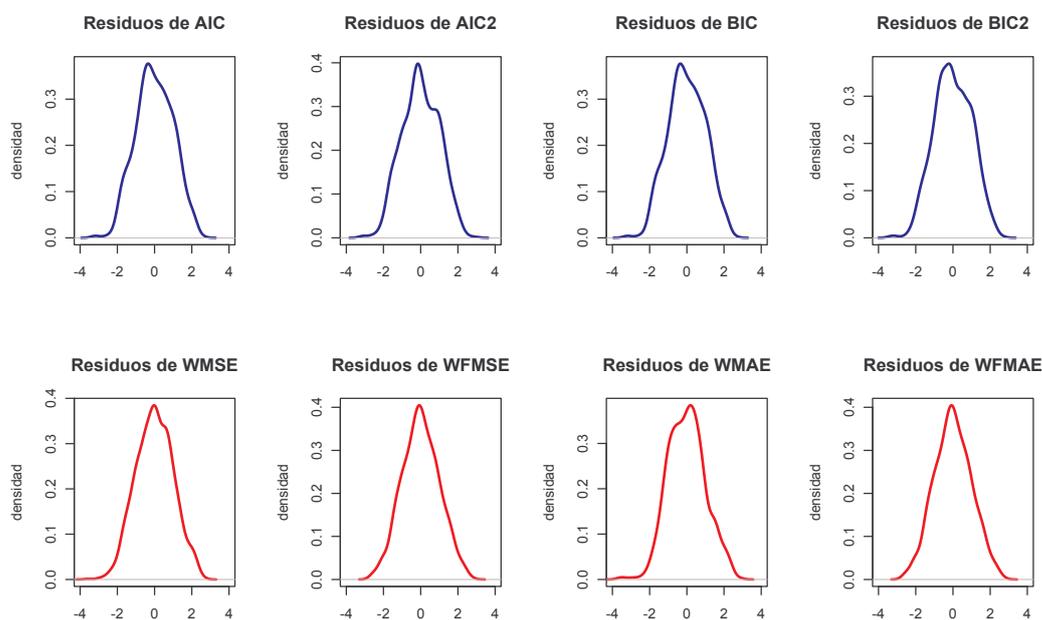


Figura 4.65: Residuos obtenidos con los modelos de los criterios de ajuste de ARIMA(1,1,1).

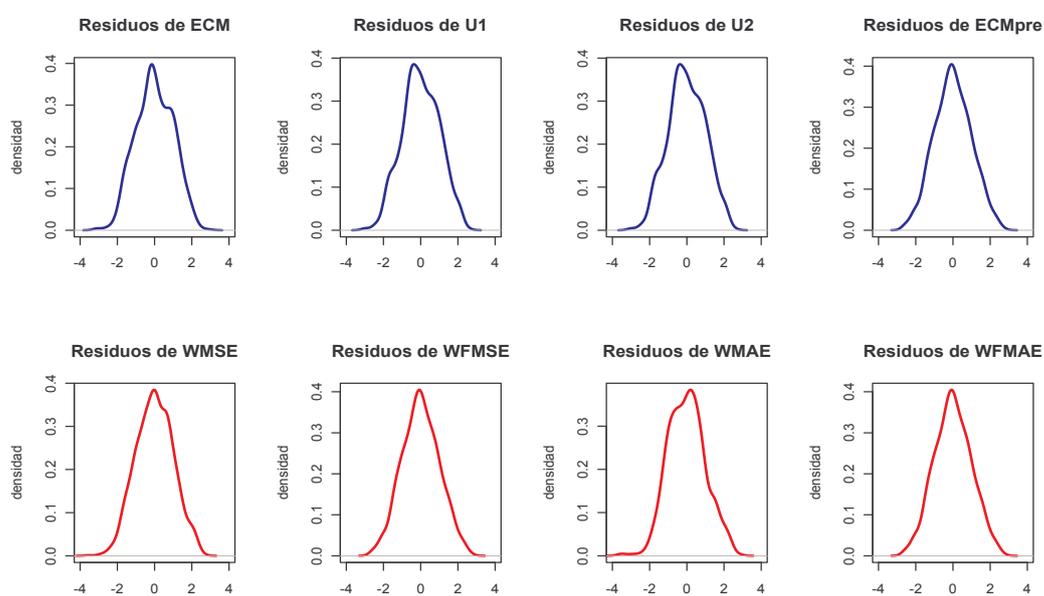


Figura 4.66: Residuos obtenidos con los modelos de los criterios de predicci3n de ARIMA(1,1,1).

Criterio	Media	Varianza
AIC	-0.1540	70.1440
AIC2	-0.2807	80.0390
ECM	-0.2807	80.0390
BIC	-0.1540	70.1440
BIC2	-0.1289	66.2084
U1	-0.3385	71.4974
U2	-0.3385	71.4974
ECM _{pre}	-0.3882	87.7777
WMSE	-0.1255	76.1692
WFMSE	-0.3882	87.7777
WMAE	1.3859	91.5392
WFMAE	-0.3882	87.7777

Tabla 4.42: Medidas descriptivas de los residuos en el conjunto de comprobación para ARIMA(1,1,1).

El comportamiento medio de los residuos es poco significativo, a excepción de los residuos obtenidos con el modelo seleccionado por el criterio WMAE, donde la media es superior que en el resto de los criterios al igual que ocurre con la varianza. Podemos ver estos comportamientos en las gráficas 4.67 y 4.68 donde se representa el histograma correspondiente a la distribución de los residuos de todas las series. En la tabla 4.43 se estudiarán los ECM_{Gpre} para cada estimación realizada, donde se podrá apreciar el comportamiento homogéneo en la primera predicción a lo largo de cada uno de los criterios considerados. En las gráficas 4.69 a la gráfica 4.72 se pueden ver diferencias en cuanto al ajuste para cada uno de los criterios, algo que se estudiará mejor con la tabla 4.44. En cuanto a las predicciones se analizarán en los gráficos 4.73 al 4.76 y se determinará con la tabla 4.45 que los resultados no son los deseados en este caso, pero al igual que en los modelos anteriores se depurará el procedimiento para poder alcanzar resultados más satisfactorios.

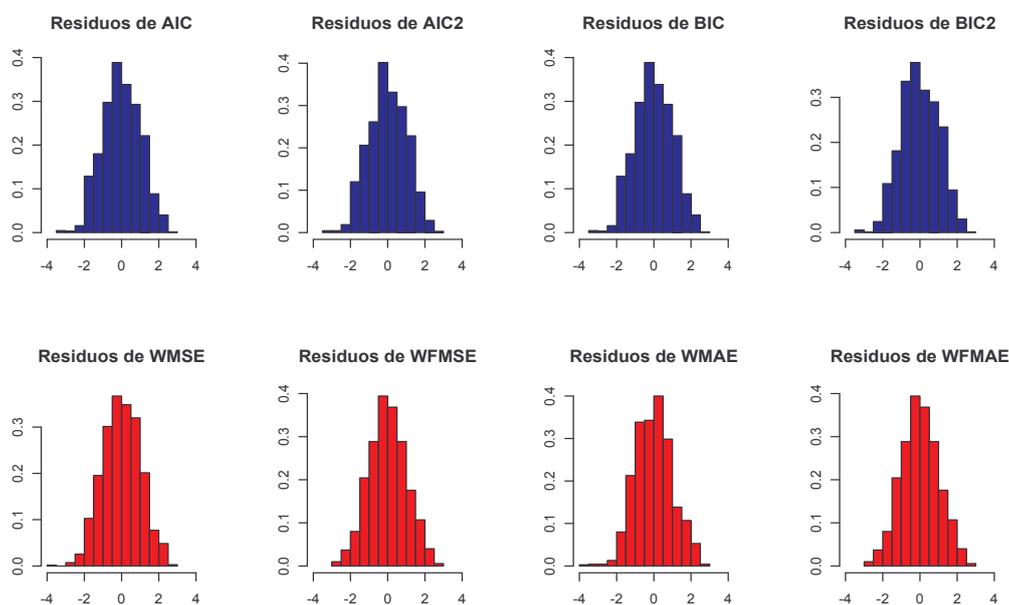


Figura 4.67: Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de ARIMA(1,1,1).

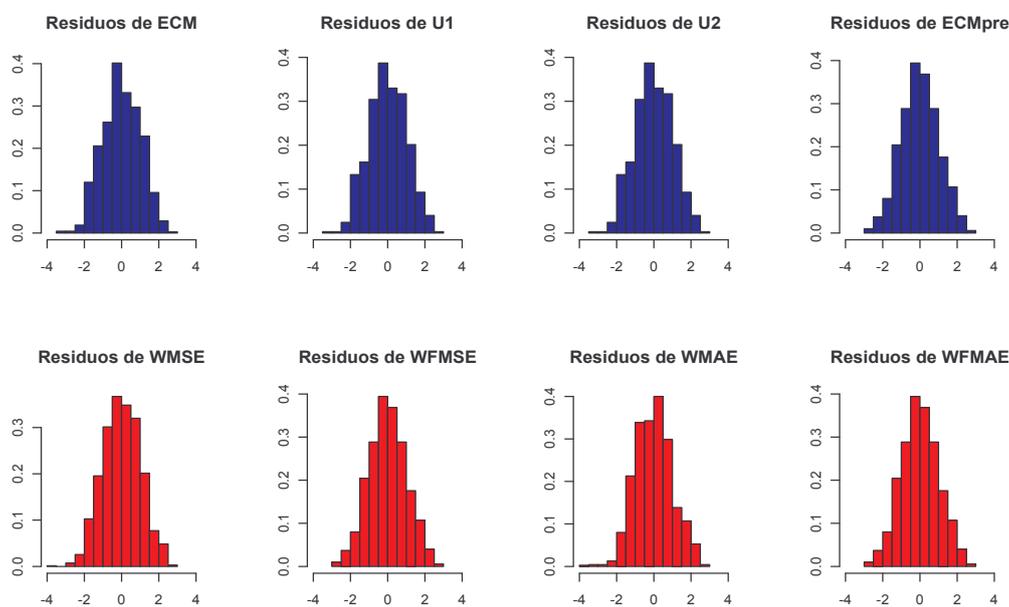


Figura 4.68: Histograma de los residuos obtenidos con los modelos de los criterios de predicción de ARIMA(1,1,1).

Criterio	ECMpre1	ECMpre2	ECMpre3	ECMpre4	ECMpre5	ECMpre6	ECMpre7
AIC	1.0411	8.6698	26.3974	53.9626	89.3966	132.0398	170.6674
AIC2	1.1226	9.5825	29.3711	60.3741	100.6886	150.3359	209.3465
ECM	1.1226	9.5825	29.3711	60.3741	100.6886	150.3359	209.3465
BIC	1.0411	8.6698	26.3974	53.9626	89.3966	132.0398	170.6674
BIC2	0.9841	8.1241	26.6224	50.5552	84.2758	124.8226	170.1912
U1	1.0243	8.666	26.4376	54.2401	90.5778	134.8758	185.4624
U2	1.0243	8.666	26.4376	54.2401	90.5778	134.8758	185.4624
ECMpre	1.8355	11.9656	33.5846	67.6016	111.7816	164.6410	224.0895
WMSE	1.2969	10.2314	29.7132	59.5776	97.7330	142.6411	192.1012
WFMSE	1.8355	11.9656	33.5846	67.6016	111.7816	164.6410	224.0895
WMAE	1.7089	14.0989	40.2108	76.5048	121.5157	172.9528	227.2271
WFMAE	1.8355	11.9656	33.5846	67.6016	111.7816	164.6410	224.0895

Tabla 4.43: ECMpre global para cada predicción realizada en el conjunto de comprobación en ARIMA(1,1,1).

En la tabla 4.43 podemos observar como el error que se comete a medida que avanzamos en el dato a predecir aumenta escandalosamente, de manera que las últimas predicciones realizadas tienen un error muy alto. Mostramos a continuación los gráficos que recogen el comportamiento de la distribución de los ECM de cada una de las series estudiadas con cada uno de los modelos seleccionados por cada criterio.

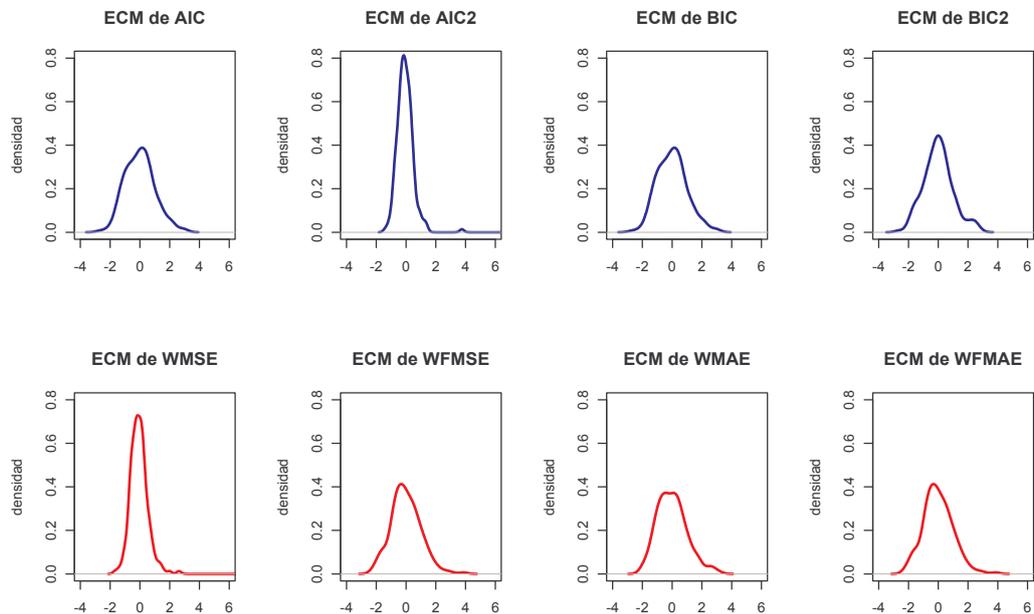


Figura 4.69: ECM de ajuste de los modelos obtenidos con los criterios de ajuste en ARIMA(1,1,1).

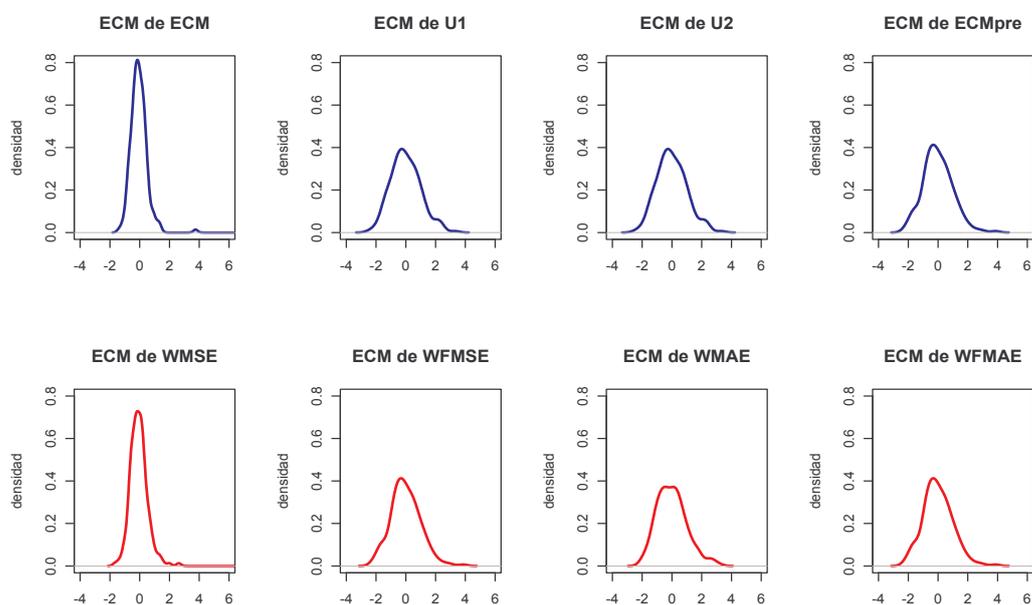


Figura 4.70: ECM de ajuste de los modelos obtenidos con los criterios de predicci3n en ARIMA(1,1,1).

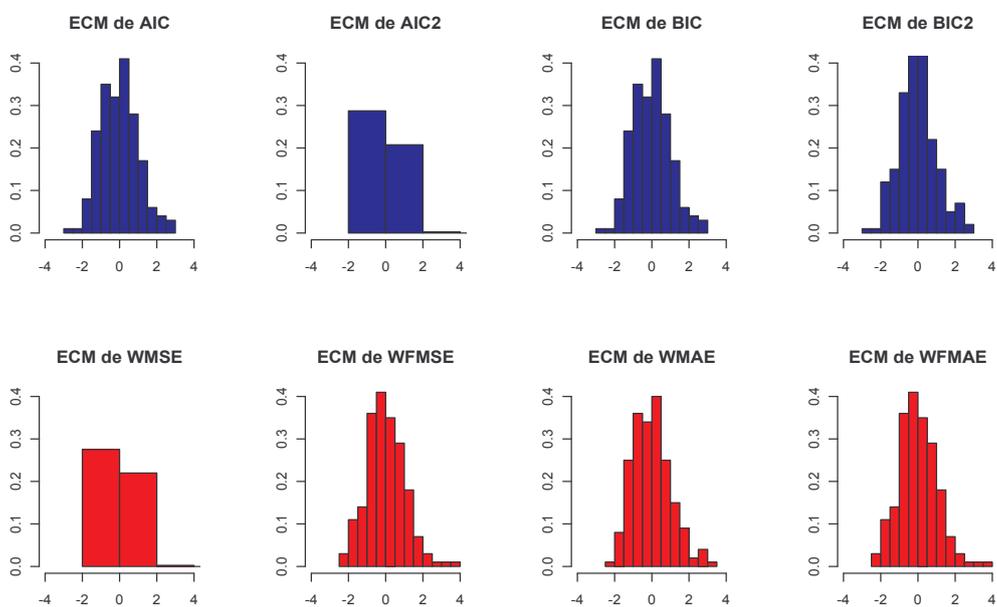


Figura 4.71: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en ARIMA(1,1,1).

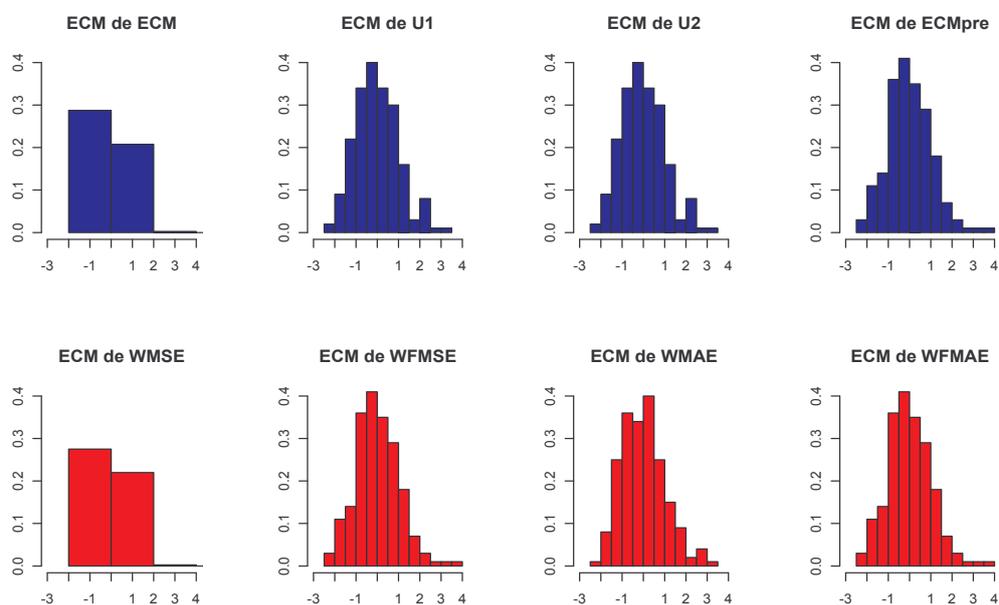


Figura 4.72: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicción en ARIMA(1,1,1).

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8802	0.8876	0.8876	0.8802	1.0220	0.8437
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8437	1.5298	1.0339	1.5298	1.3869	1.5298

Tabla 4.44: El ECM global de ajuste, ECMG, para cada criterio en modelos ARIMA(1,1,1).

Se repite de nuevo, como en los modelos anteriores, que en base al ajuste los criterios clásicos son los que mejor se comportan, y casualmente sin distinción entre criterios de ajuste o predicción.

Veamos que ocurre cuando lo que se estudia es la capacidad predictiva del modelo.

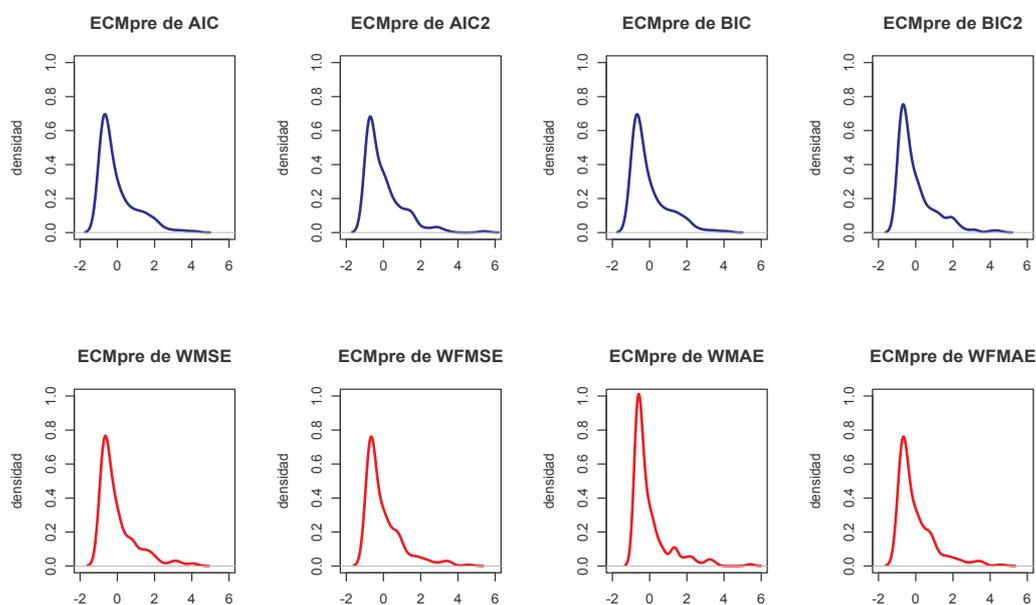


Figura 4.73: ECM de predicción de los modelos obtenidos con los criterios de ajuste en ARIMA(1,1,1).

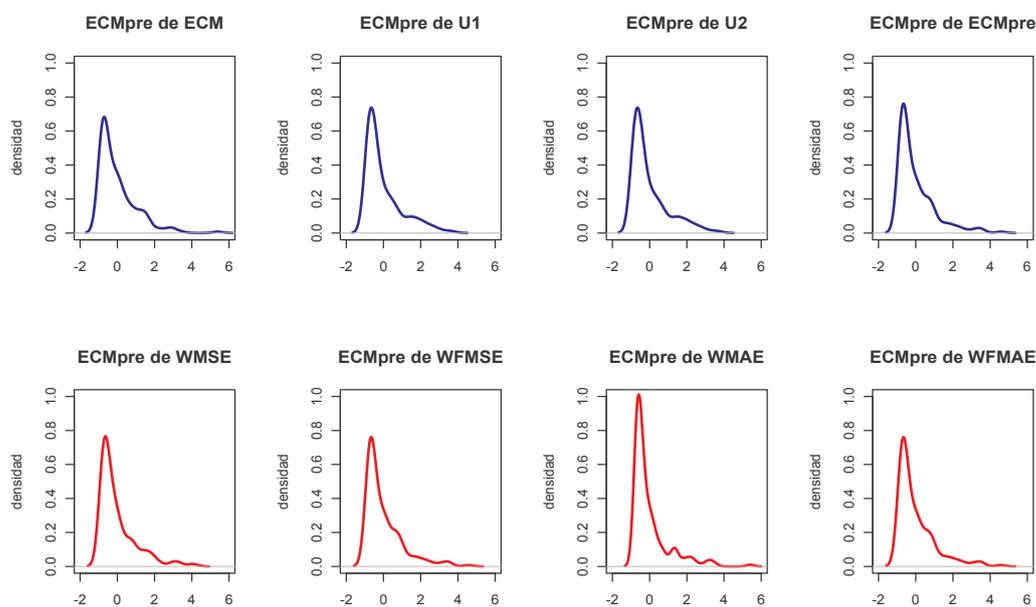


Figura 4.74: ECM de predicción de los modelos obtenidos con los criterios de predicción en ARIMA(1,1,1).

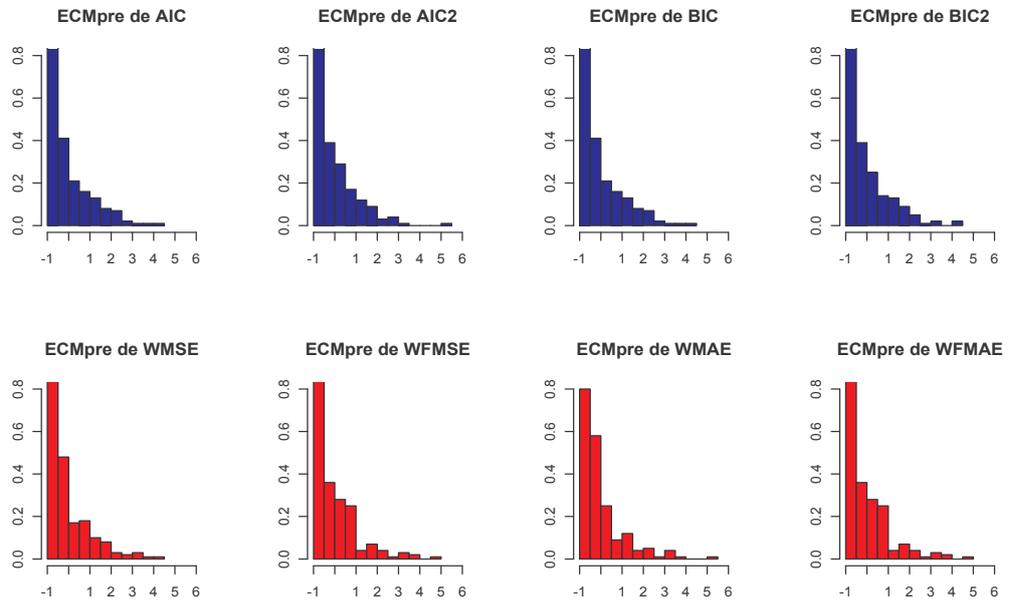


Figura 4.75: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en ARIMA(1,1,1).

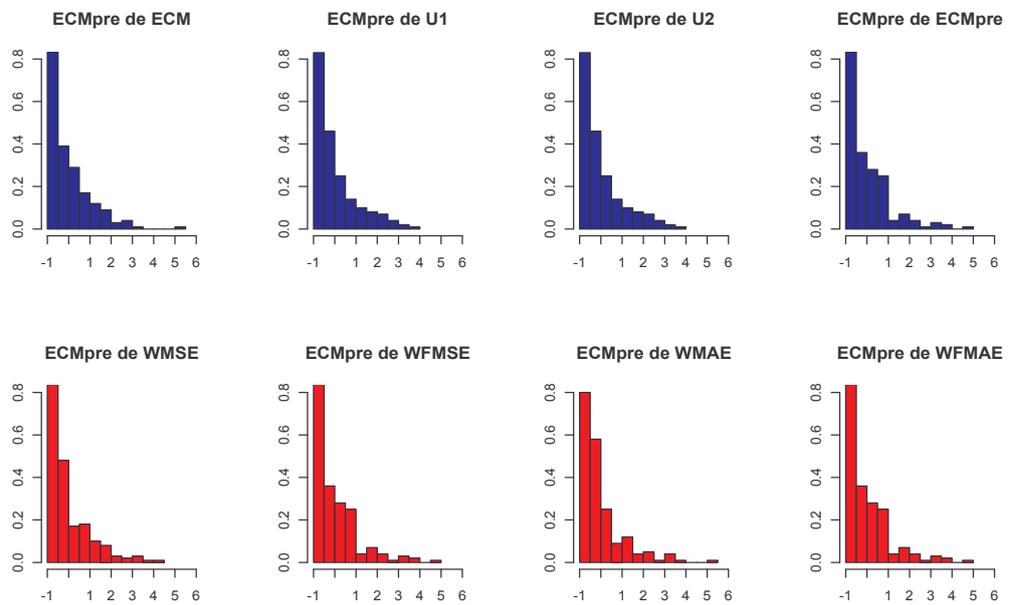


Figura 4.76: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en ARIMA(1,1,1).

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	70.1678	80.1179	80.1179	70.1678	66.2251	71.6120
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	71.6120	87.9285	76.1849	87.9285	93.4599	87.9285

Tabla 4.45: ECM global de predicción, ECMGpre, para cada criterio en modelos ARIMA(1,1,1).

Con respecto a la tabla 4.45, los ECMGpre alcanzan valores elevados, por lo que no se consiguen buenos resultados al respecto mejorando en los criterios clásicos, con una diferencia algo mayor en el criterio BIC2 con respecto al resto.

Una vez más el procedimiento de buscar un mismo modelo para todas las series en función de algún criterio establecido, no es la manera de proceder que asegure conseguir unos buenos resultados predictivos.

Veamos que ocurre en esta ocasión, con las series correctas e incorrectas, en cuanto a predicción y ajuste con cada uno de los criterios posibles, éstos serán el AIC2, BIC, BIC2, WFMSE y WFMAE.

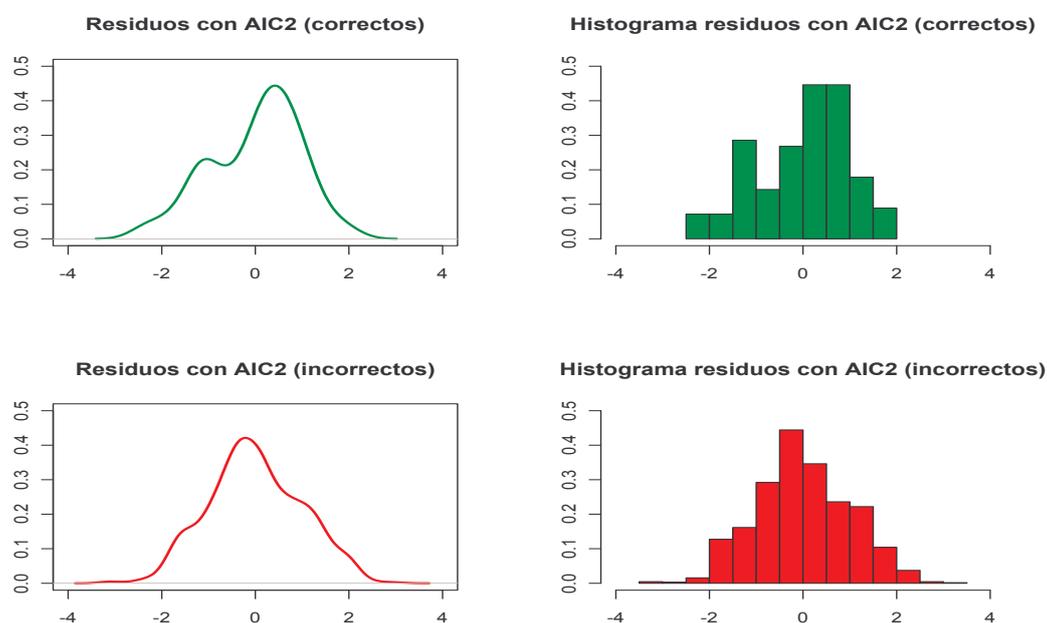


Figura 4.77: Residuos de las series correctas e incorrectas con el criterio AIC2 en el ARIMA(1,1,1).

	Media	Varianza	ECMG	ECMGpre
Correctos	2.7259	59.9963	0.9635	67.4267
Incorrectos	-0.4036	74.3062	0.8733	74.4691

Tabla 4.46: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,1) con el criterio AIC2.

La mejor media se obtiene en el grupo de series incorrectas, no obstante éste tiene una varianza superior al grupo de los correctos, y el error de predicción también es más elevado, en cuanto al ajuste el comportamiento es mejor. En esta ocasión se han conseguido detectar correctamente 16 de las 200 series, por lo que el porcentaje es bastante bajo.

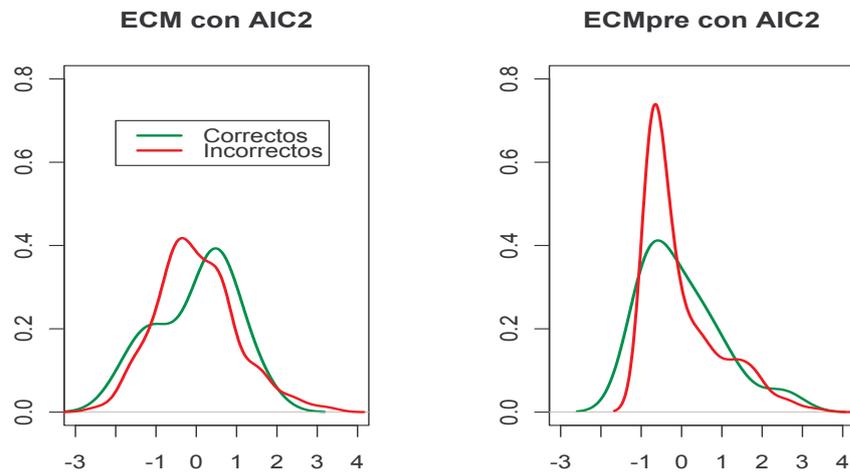


Figura 4.78: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en ARIMA(1,1,1).

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.0027	0.0413	0.1250	0.2078	0.2632	0.2847	0.2679
Incorrectos	1.0595	9.1164	27.8770	56.5298	93.5915	139.4517	193.6580

Tabla 4.47: ECM global de predicción en ARIMA(1,1,1) con el criterio AIC2 en el conjunto de comprobación.

Se observa claramente, como la predicción en las series incorrectas ha fracasado completamente con este criterio.

A continuación podemos ver la representación de los residuos en el conjunto

de comprobación en función de si el criterio ha sido capaz de determinar el modelo generador (correctas) o no (incorrectas).

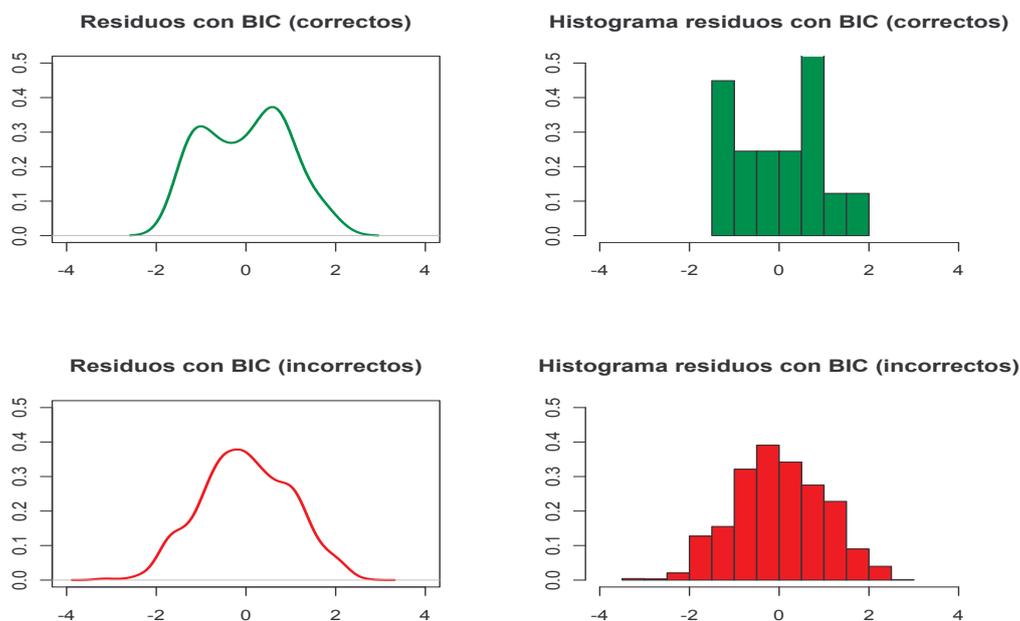


Figura 4.79: Residuos de las series correctas e incorrectas con el criterio BIC en el ARIMA(1,1,1).

	Media	Varianza	ECMG	ECMGpre
Correctos	5.5983	65.6473	0.9704	96.9888
Incorrectos	-0.3549	71.2013	0.8688	71.3273

Tabla 4.48: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para ARIMA(1,1,1) con el criterio BIC.

Con el criterio BIC, las series que tienen mejor media aunque mayor varianza son las incorrectas, pero en esta ocasión son las que consiguen una mejor predicción media. Veamos si observando las predicciones de manera individual este comportamiento se mantiene (ver tabla 4.49).

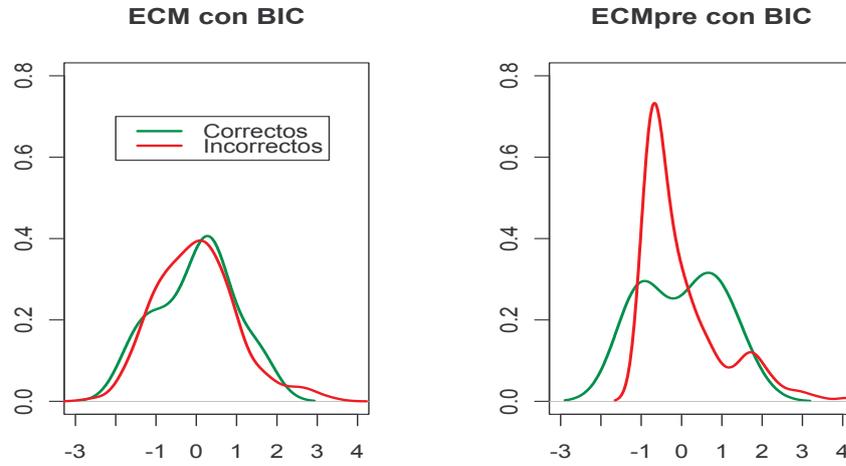


Figura 4.80: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC en ARIMA(1,1,1).

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.9791	9.7075	33.9933	78.9776	131.7884	182.1453	241.33
Incorrectos	1.0520	8.7753	26.6874	54.4736	90.4620	134.1576	183.6831

Tabla 4.49: ECM global de predicción en ARIMA(1,1,1) con el criterio BIC en el conjunto de comprobación.

En relación a las predicciones individuales, las series correctas consiguen un menor ECMG en la primera predicción, sin embargo en las restantes el error es mayor, por lo que habría que plantearse si nos interesa “sacrificar” el resto de predicciones para asegurar que en el instante inmediatamente siguiente a la última observación la predicción es más precisa. Con relación al gráfico 4.80, se puede ver como en cuanto a resultados de ajuste las diferencias entre los grupos son mínimas, sin embargo, se aprecia que en el ECMpre el grupo de series incorrectas es muy heterogéneo, como cabe esperar, ya que las series pertenecientes a este grupo utilizan modelos diferentes entre sí para obtener las predicciones mientras que en el grupo de las series correctas todas predicen bajo el mismo modelo, el modelo generador, ARIMA(1, 1, 1).

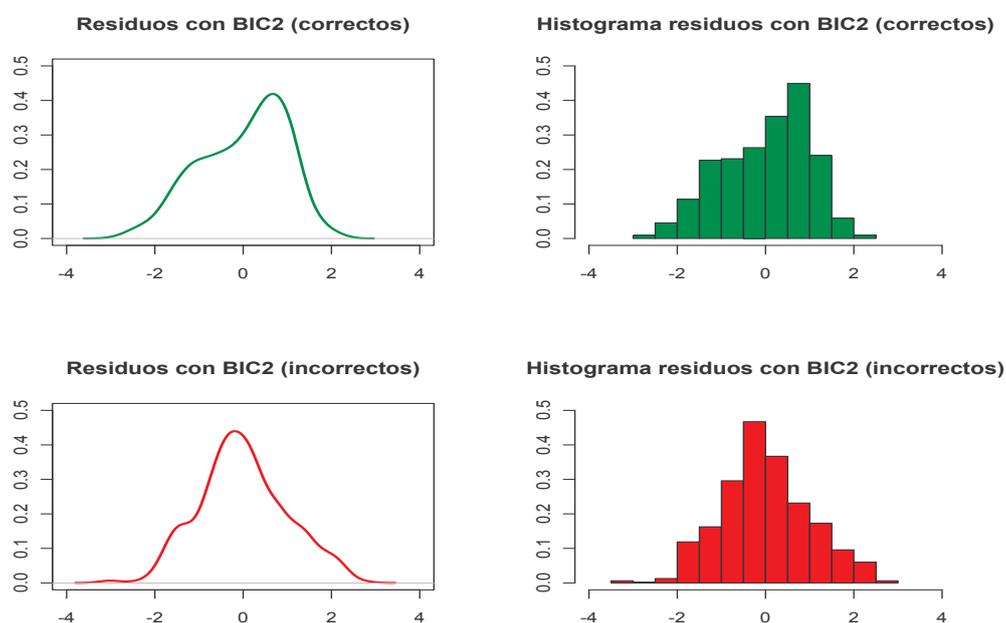


Figura 4.81: Residuos de las series correctas e incorrectas con el criterio BIC2 en el ARIMA(1,1,1).

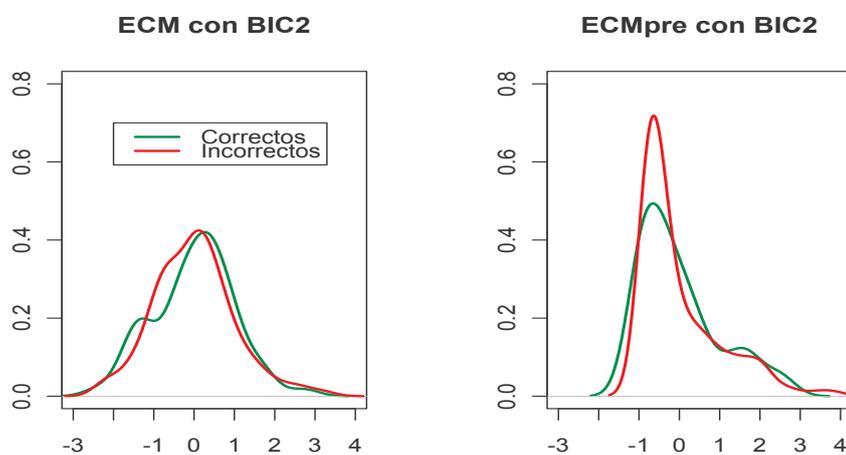


Figura 4.82: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en ARIMA(1,1,1).

A la vista del gráfico 4.81 los residuos de las series incorrectas se asemejan más a una distribución normal, no obstante puede verse en la tabla 4.50, que aunque se consigue una media menor la varianza es mayor en el grupo, lo que

provoca un peor resultado en la predicción. Con este criterio se ha llegado a un porcentaje de acierto, es decir, de detectar en modelo generador en el 31,5 % de las series.

	Media	Varianza	ECMG	ECMGpre
Correctos	2.1083	58.7755	0.9769	63.2206
Incorrectos	-1.1735	70.8572	0.9078	72.2343

Tabla 4.50: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,1) con el criterio BIC2.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.8074	7.2651	22.6149	49.0457	82.7032	119.5105	160.5972
Incorrectos	1.1022	9.0969	27.4361	55.0929	91.1716	135.8450	185.8950

Tabla 4.51: ECM global de predicción en ARIMA(1,1,1) con el criterio BIC2 en el conjunto de comprobación.

Como se puede apreciar en las tablas 4.50 y 4.51, en el caso de utilizar el criterio BIC2, las cosas tanto global como individualmente funcionan mejor en el caso en que las series son identificadas correctamente mediante el modelo simulador de las mismas, en cuanto a mejor ajuste, se puede ver en el gráfico 4.82 que las diferencias no son determinantes entre los grupos.

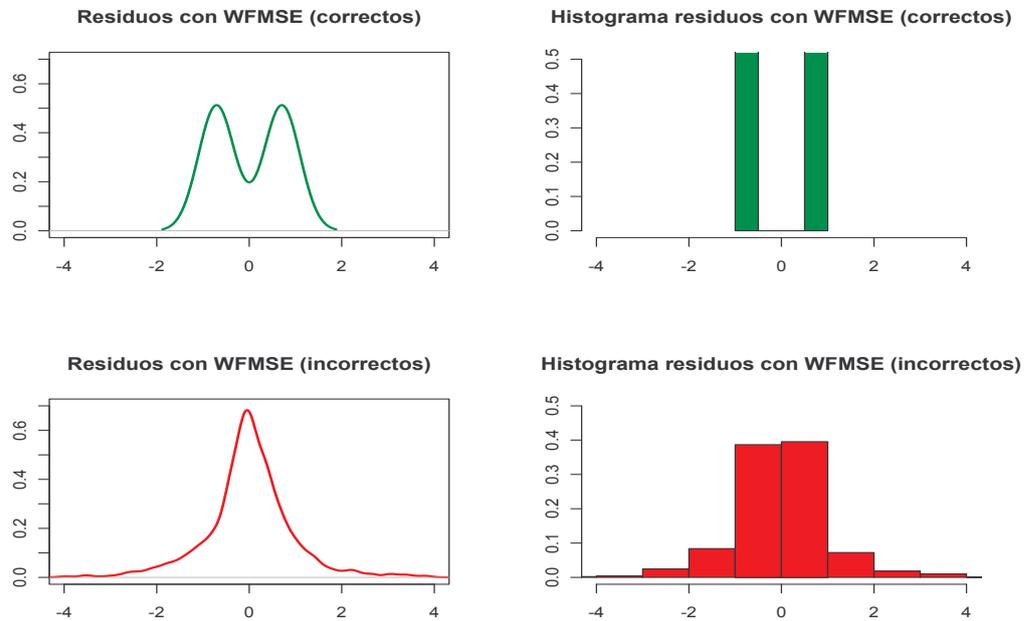


Figura 4.83: Residuos de las series correctas e incorrectas con el criterio WFMSE en el ARIMA(1,1,1).

	Media	Varianza	ECMG	ECMGpre
Correctos	-0.1931	0.7127	1.0302	0.75
Incorrectos	0.0576	10.6160	6.3079	10.6194

Tabla 4.52: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,1) con el criterio WFMSE.

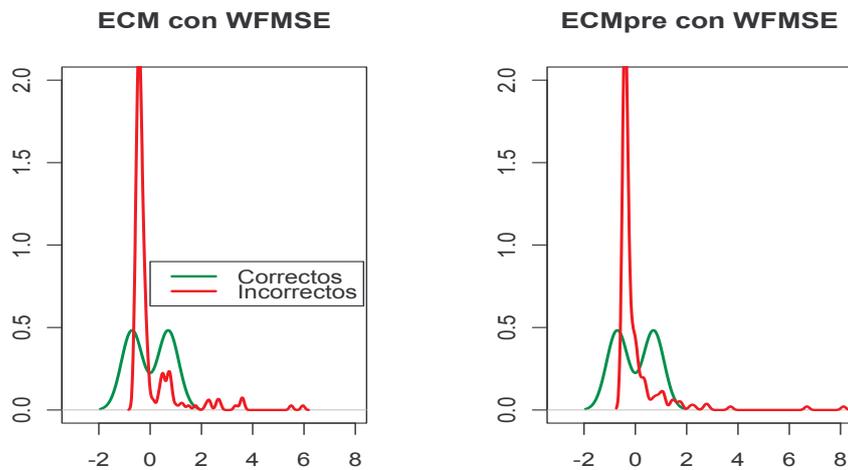


Figura 4.84: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMSE en ARIMA(1,1,1).

Se aprecia una clara mejoría en cuanto a detectar el modelo con el criterio WFMSE tanto en predicción como en ajuste (tabla 4.52), de manera global e individual (tabla 4.53), sin embargo, tan sólo ha sido posible identificar 2 series de las 200 generadas con el modelo ARIMA(1, 1, 1).

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.9367	1.7903	0.5596	0.6327	0.0210	0.7517	0.5580
Incorrectos	1.1463	3.6799	5.7916	6.9934	9.8202	16.9171	29.9870

Tabla 4.53: ECM global de predicción en ARIMA(1,1,0) con el criterio WFMSE en el conjunto de comprobación.

Por último, estudiaremos el criterio WFMAE, en cuál al igual que el WFMSE tan sólo cuenta con dos series en el grupo de series correctas. Por lo que los resultados obtenidos van en la misma línea, como se puede apreciar en los gráficos 4.85 y 4.86 y en las tablas 4.54 y 4.55.

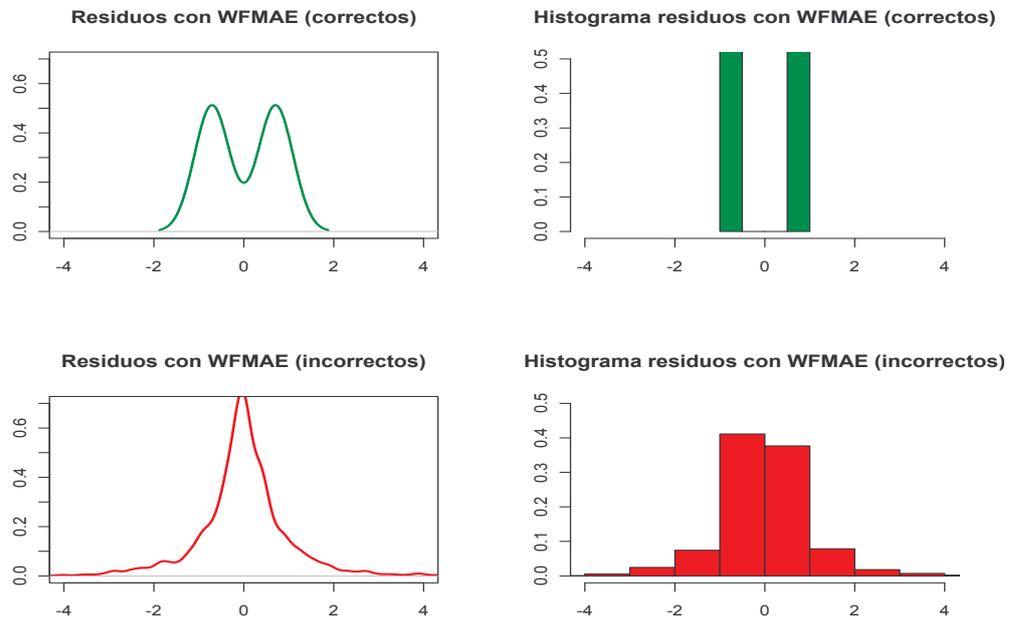


Figura 4.85: Residuos de las series correctas e incorrectas con el criterio WFMAE en el ARIMA(1,1,1).

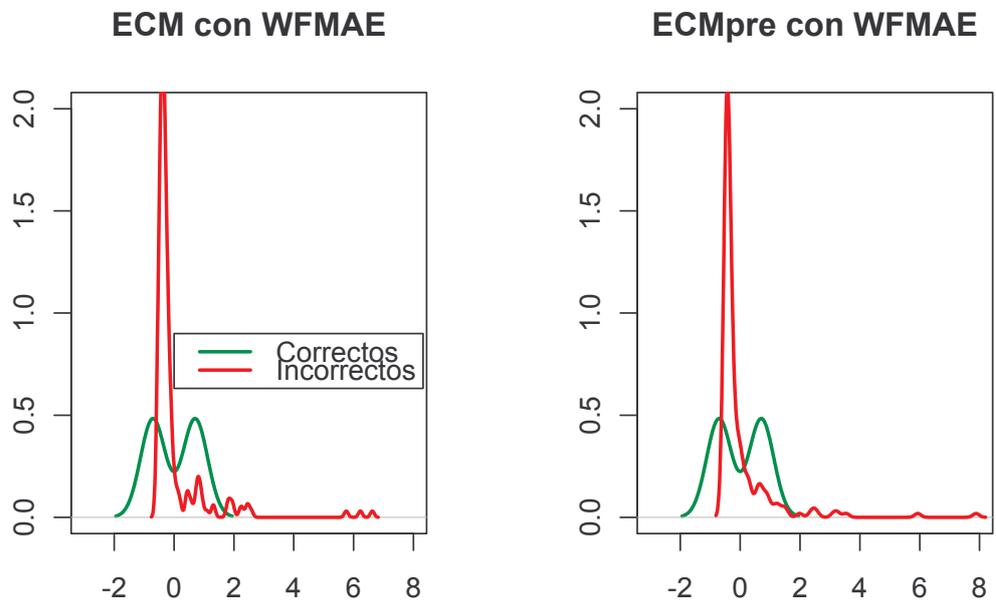


Figura 4.86: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en ARIMA(1,1,1).

	Media	Varianza	ECMG	ECMGpre
Correctos	-0.1931	0.7127	1.0302	0.75
Incorrectos	0.0262	12.6329	5.7176	12.6335

Tabla 4.54: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para ARIMA(1,1,1) con el criterio WFMAE.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.9367	1.7903	0.5596	0.6327	0.0210	0.7517	0.5580
Incorrectos	0.8542	2.3851	4.3589	6.6875	11.8042	22.3604	39.9844

Tabla 4.55: ECM global de predicción en ARIMA(1,1,0) con el criterio WFMAE en el conjunto de comprobación.

Al igual que con el criterio anterior el comportamiento con respecto a la predicción global es el mismo, sin embargo, se aprecia una cierta mejoría en la primera predicción individual, como ya se ha comentado dependerá del “coste” que tenga esa primera predicción, lo que marcará la forma de proceder en cada estudio.

No obstante, no hay un comportamiento homogéneo entre las series correctas e incorrectas, con lo que el ser capaz de identificar las series no tiene gran importancia a la hora de buscar las mejores predicciones de una serie de tiempo. Entre los criterios AIC2, BIC y BIC2 no se aprecian diferencias significativas, sin embargo, en los criterios de predicción propuestos en el capítulo 3, WFMSE y WFMAE los errores conseguidos son menores que los casos anteriores, no obstante, las conclusiones no son aceptadas dado que el grupo de series correctas tiene un tamaño demasiado pequeño.

Veamos que ocurre a continuación con los gráficos y tablas de residuos y ECM de ajuste y predicción, en el caso en que se respeten los criterios asociados a cada serie con cada criterio elegido.

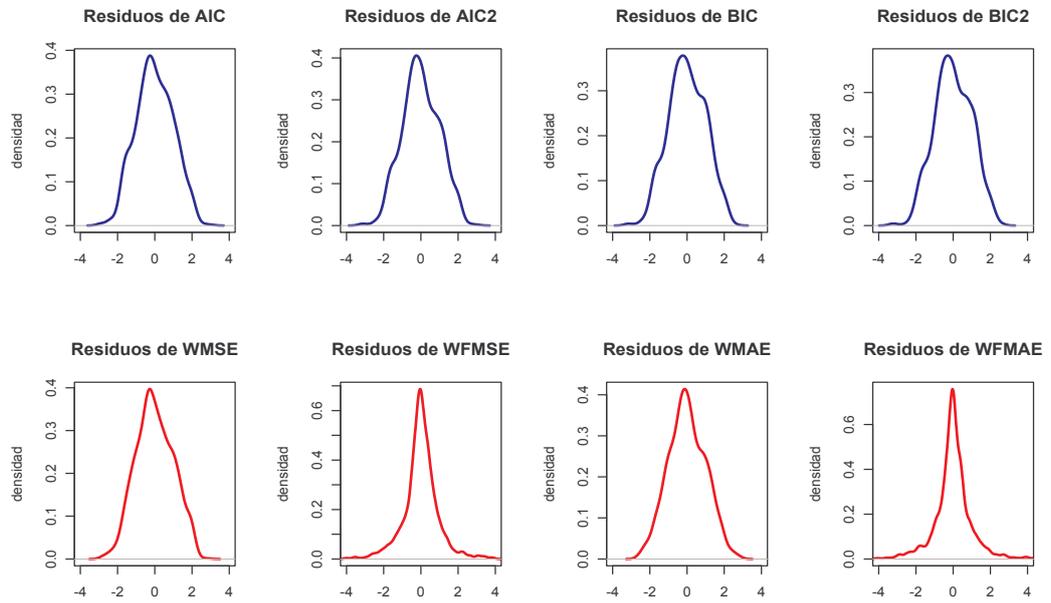


Figura 4.87: Residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).

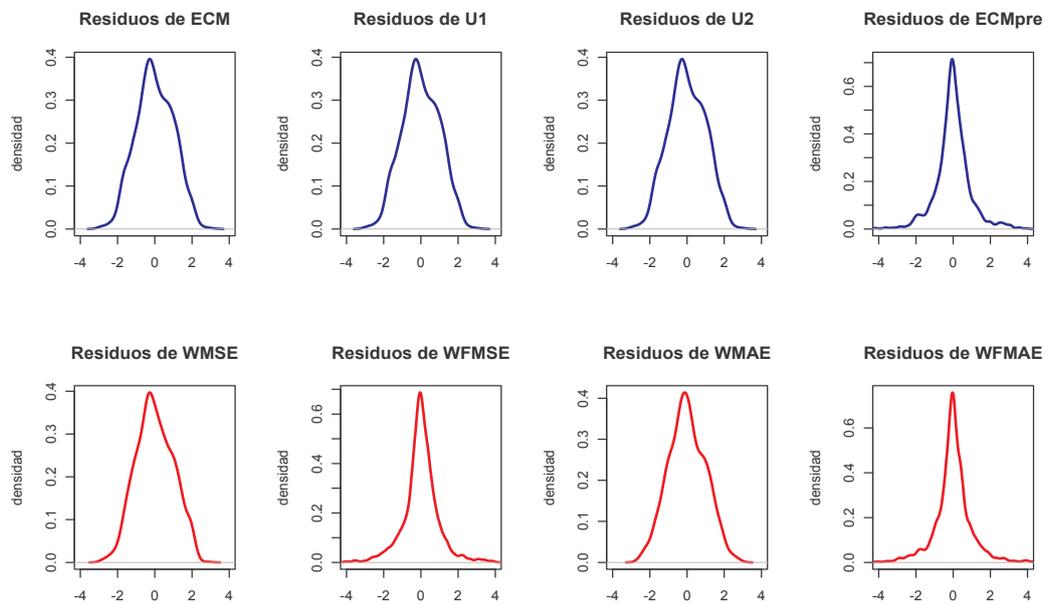


Figura 4.88: Residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).

Criterio	Media	Varianza
AIC	-0.2265	75.9539
AIC2	-0.1533	73.8822
ECM	-0.3768	77.4933
BIC	-0.1466	72.2040
BIC2	-0.1397	69.3754
U1	-0.3768	77.4933
U2	-0.3768	77.4933
ECMpre	0.1189	10.1386
WMSE	0.0746	85.4147
WFMSE	0.0551	10.5176
WMAE	-0.0228	89.9130
WFMAE	0.0240	12.5141

Tabla 4.56: Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para ARIMA(1,1,1).

A la vista de los gráficos 4.87 y 4.88 se aprecia un comportamiento mejor en los residuos obtenidos con el empleo de los modelos seleccionados por los criterios ECMpre, WFMSE y WFMAE, donde la varianza es menor, significativamente que en el resto de las curvas representadas. En la tabla 4.56 se puede comprobar que efectivamente la varianza es mucho menor en los criterios nombrados. Comparando los resultados obtenidos en estos criterios con los que se recogen en la tabla 4.42 se puede ver la disminución producida tanto en la media como en la varianza, no ocurre igual con el resto de criterios. Veamos en la tabla 4.57 que ocurre en cuanto a las predicciones individuales, aunque es de esperar que sean también significativamente mejores en los criterios de predicción propuestos en este trabajo, como viene ocurriendo en los casos estudiados anteriormente.

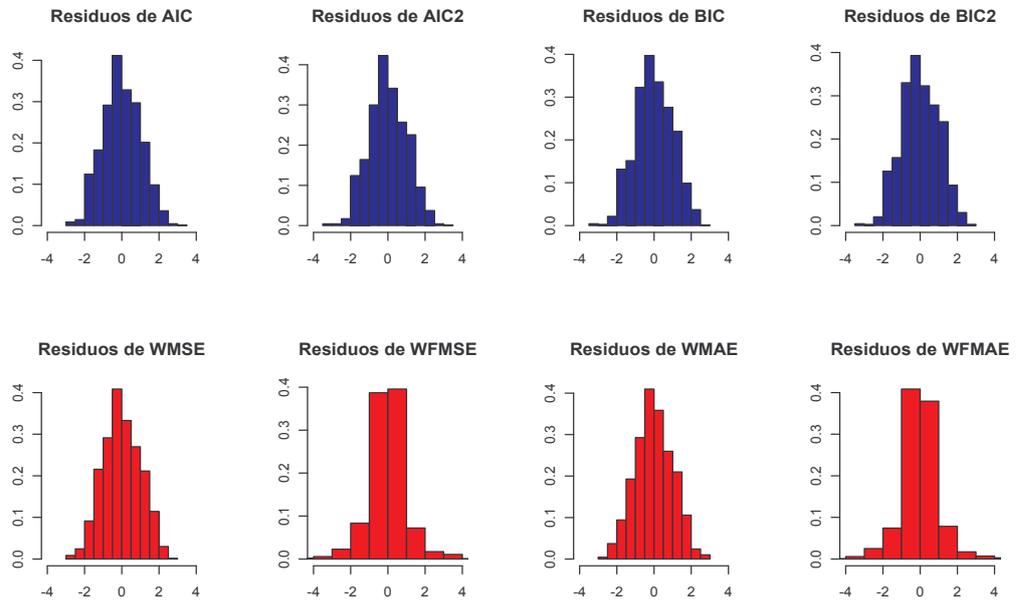


Figura 4.89: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).

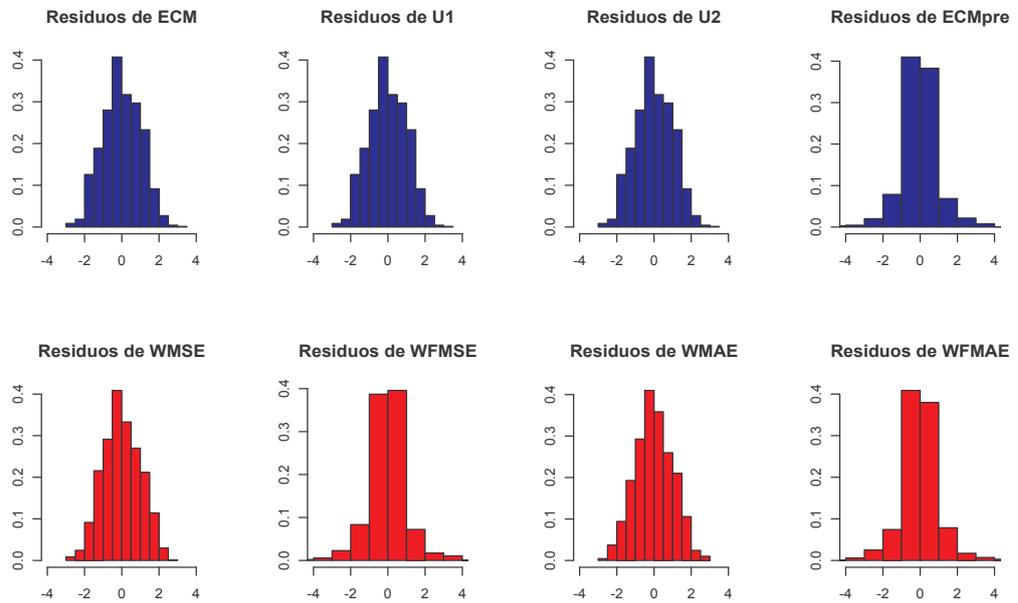


Figura 4.90: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).

Criterio	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
AIC	1.0628	9.0409	27.8392	57.2381	95.4906	142.8898	198.4752
AIC2	1.0526	9.0122	27.8212	56.9548	93.7961	138.3311	190.3723
ECM	1.0620	9.1178	28.0612	57.8237	97.1636	146.2331	203.9860
BIC	1.0495	8.8079	26.9431	55.3313	91.9084	135.8372	185.8372
BIC2	1.0094	8.5199	25.9174	53.1880	88.5041	130.6996	177.9262
U1	1.0620	9.1178	28.0612	57.8237	97.1636	146.2331	203.9860
U2	1.0620	9.1178	28.0612	57.8237	97.1636	146.2331	203.9860
ECMpre	1.8528	5.5180	7.0569	7.7966	9.4978	14.6491	24.6978
WMSE	1.4363	10.9858	32.0239	63.8358	106.1446	160.1374	223.3780
WFMSE	1.1442	3.6610	5.7393	6.9298	9.7222	16.7555	29.6927
WMAE	1.6140	12.1355	34.4259	67.8559	113.0911	168.6719	231.6009
WFMAE	0.8551	2.3791	4.3209	6.6270	11.6864	22.1443	39.5902

Tabla 4.57: ECM globales para cada predicción realizada en el conjunto de comprobación en ARIMA(1,1,1).

Se puede decir que con una gran diferencia con los demás, el criterio que mejor funciona en cuanto a predicciones individuales, es decir, consigue obtener el menor error de predicción en la primera predicción, al igual que en la segunda, y así hasta la séptima, es el de WFMAE e incluso el WFMSE, donde se ve claramente una mejoría observación a observación (ver tabla 4.57).

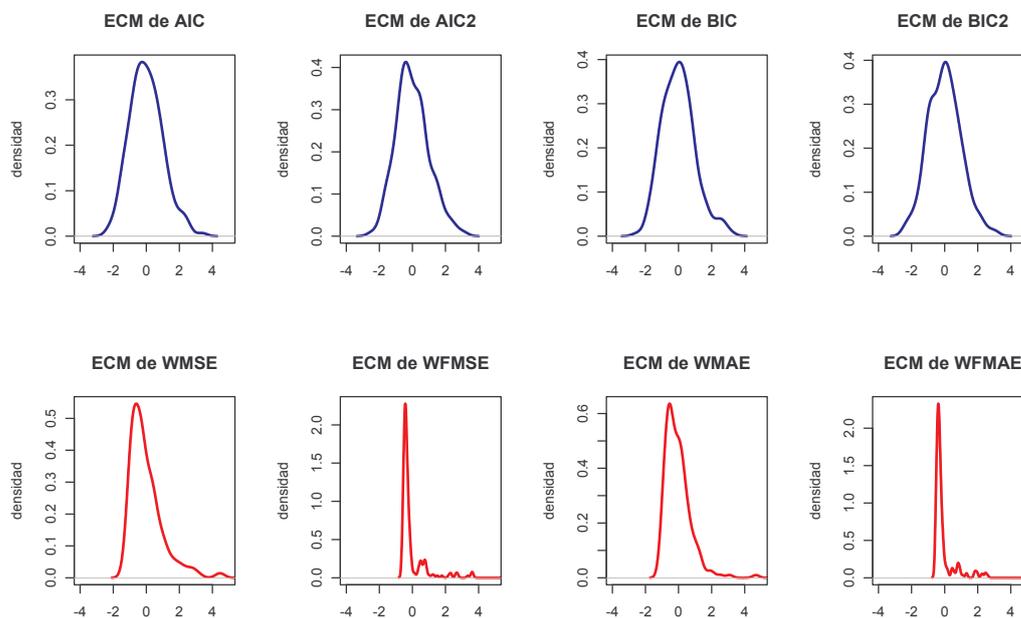


Figura 4.91: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).

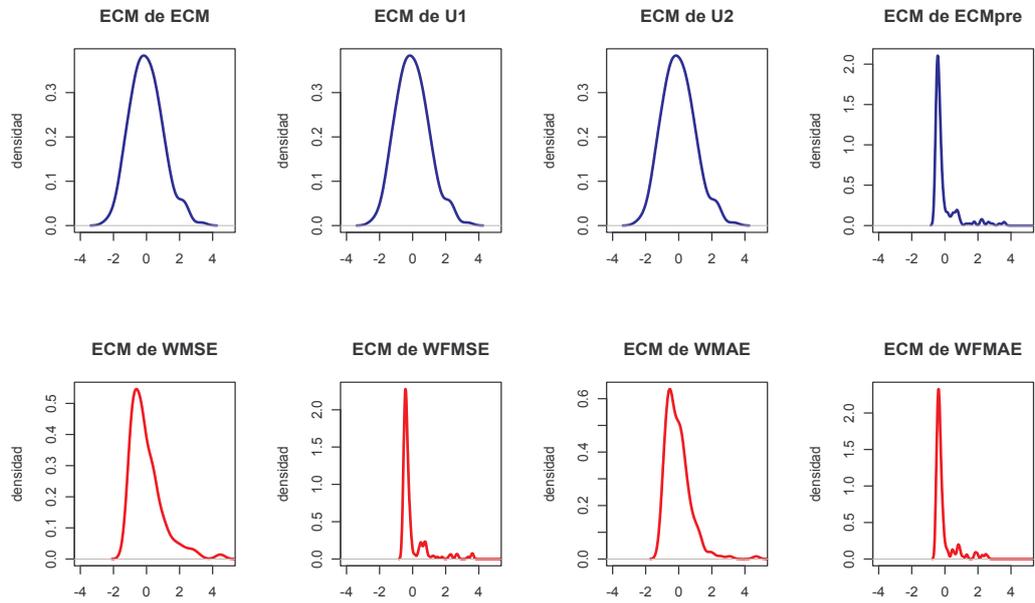


Figura 4.92: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).

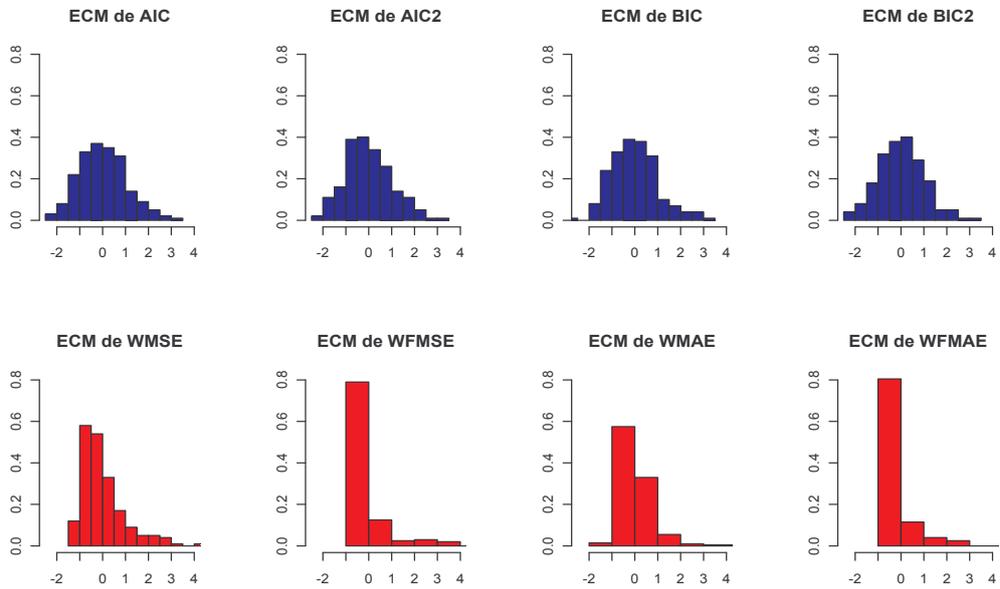


Figura 4.93: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).

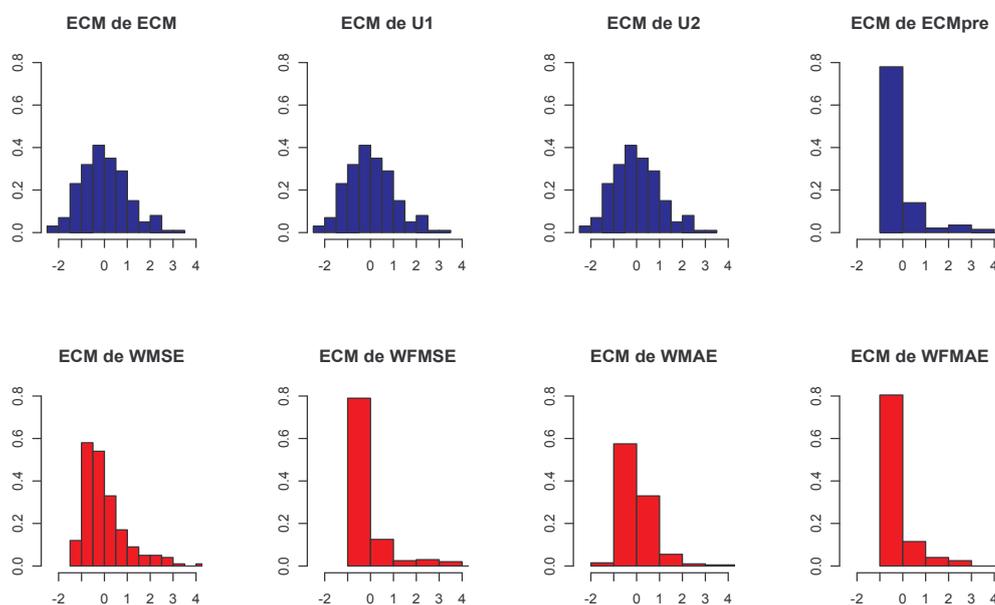


Figura 4.94: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8453	0.8805	0.8376	0.8723	0.9295	0.8376
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8376	6.2649	1.3026	6.2551	1.4790	5.6706

Tabla 4.58: ECM global de ajuste, ECMG, para cada criterio en modelos ARIMA(1,1,1).

A la vista de la tabla 4.58 donde se recogen los errores cuadráticos medios de ajuste, se observa un comportamiento que ya se nos ha dado en alguna ocasión, y es que, los criterios que consiguen las mejores predicciones, en muchos casos, son a su vez los que tienen un peor ajuste. En relación al mejor ajuste, éste se consigue con los criterios clásicos.

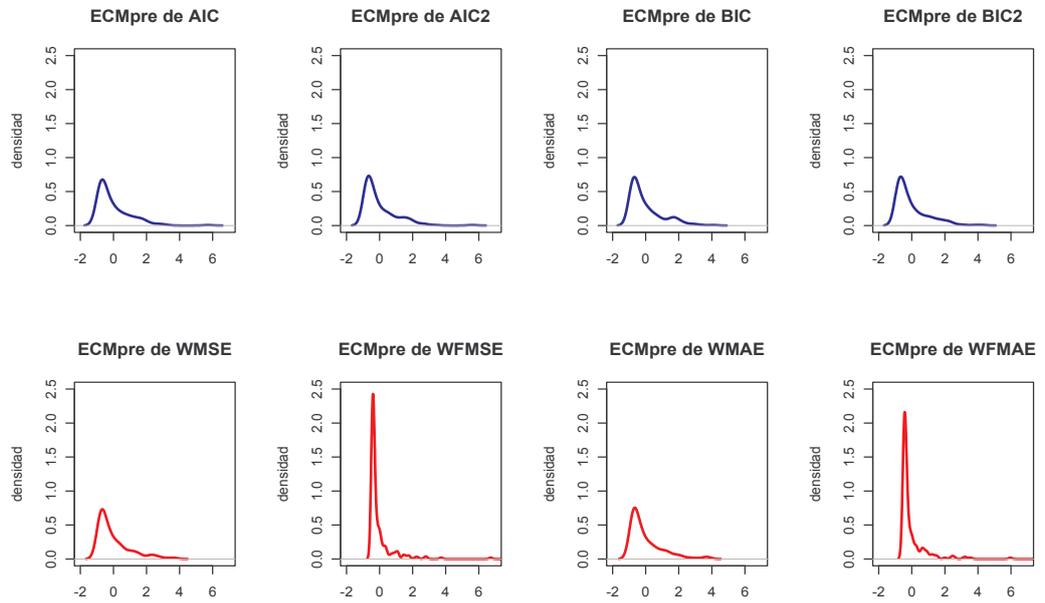


Figura 4.95: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).

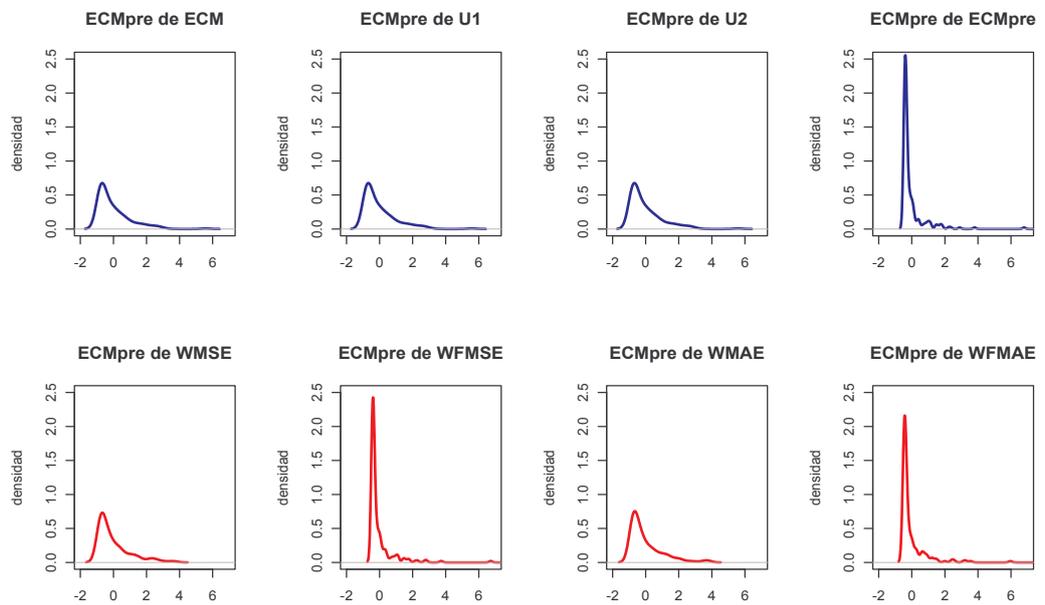


Figura 4.96: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).

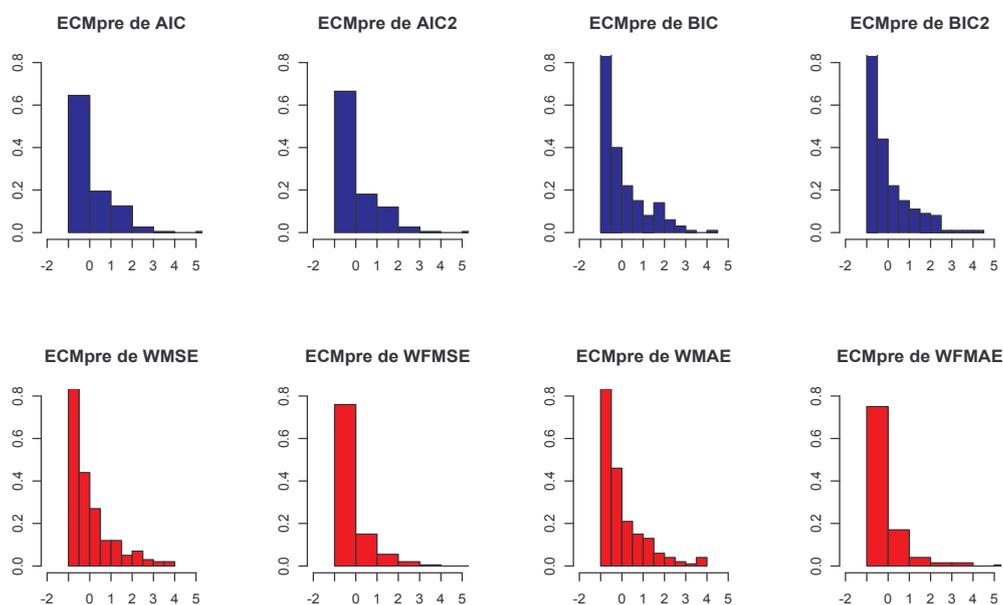


Figura 4.97: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).

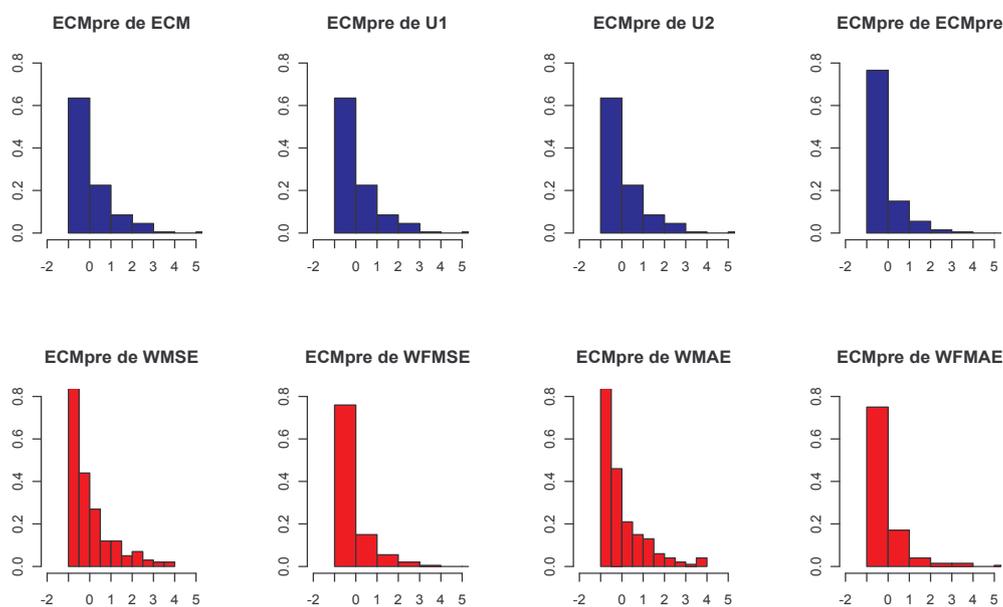


Figura 4.98: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie ARIMA(1,1,1).

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG _{pre}	76.0052	73.9057	77.6353	72.2254	69.3949	77.6353
	U2	ECM _{pre}	WMSE	WFMSE	WMAE	WFMAE
ECMG _{pre}	77.6353	10.1527	85.4202	10.5207	89.9135	12.5147

Tabla 4.59: ECM global de predicción, ECMG_{pre}, para cada criterio en modelos ARIMA(1,1,1).

Se ve claramente la increíble mejoría de los errores de predicción entre los criterios ECM_{pre}, WFMSE y WFMAE, con el resto de criterios estudiados, dentro de éstos tres, el que obtiene un mejor valor en cuanto al error medio de predicción global es el ECM_{pre}, sin embargo si observamos de nuevo la tabla 4.57 podemos comprobar como son los otros dos criterios los que alcanzan mejores valores en las predicciones iniciales, lo que adquiere un mayor interés, aunque ciertamente es en las 2 últimas de éstas 7 predicciones realizadas donde el ECM_{pre} consigue mejores resultados, sin embargo, como ya hemos comentado en algún momento nuestro interés mayor radica en estas primeras, ya que son las que no pueden ser actualizadas, mientras que las demás se podrían ir perfilando. No obstante, determinar la mejor línea de proceder en cada caso dependerá del objetivo fijado en el estudio que se realice en cada momento. En nuestro caso, el objetivo siempre fue penalizar los errores dando una mayor importancia a los más próximos en el tiempo, es por ello, que para solventar esto, la mejor manera de proceder sea utilizando los criterios de predicción propuestos en este trabajo, es decir, los criterios que hemos llamado WFMSE y WFMAE.

En general, en cuanto a los procesos no estacionarios podríamos apuntar que, a la hora de conseguir una buenas predicciones, lo más cercanas al valor real de la serie, la forma de actuar, sería utilizar cualquiera de los dos criterios de predicción propuestos en esta tesis, WFMSE y/o WFMAE, que asegura una mejor predicción global e individual, respetando en cada caso el modelo elegido como “óptimo” para cada serie, ya que no parece existir un modelo único “mejor” para las series generadas con un mismo modelo fijo, esto ha sido probado no sólo en los modelos ARIMA expuestos en este capítulo, sino en otros donde se le han variado tanto los parámetros como los coeficientes y el comportamiento aquí detallado se ha dado en todos ellos. Además se ha probado a seleccionar los modelos atendiendo a otra decisión, que no sea el mínimo sino la moda, ya que al evaluar 200 series de un mismo modelo podemos llegar a pensar que el mejor modelo será el que se elija con mayor frecuencia y no el que minimice, sin embargo, los resultados no han sido mejores en ningún caso, y esto ha presentado además una complejidad añadida que es el caso en que la moda no sea única.

4.3. Procesos Estacionarios, SARIMA(p, d, q)(P, D, Q) $_s$

En esta sección se llevará a cabo el estudio de algunos de los procesos estacionarios, siguiendo la misma la línea que en el apartado anterior, se tomará $s = 7$, suponiendo por tanto, que los datos son diarios y presentan una estacionalidad semanal.

Modelo SARIMA(0, 1, 1)(0, 1, 1) $_7$, cuyos coeficientes son $\theta_1 = 0,7$ y $\Theta_1 = 0,5$. (Modelo 77)

	AIC	AIC2	ECM	BIC	BIC2
Valor	2.303203	-0.4091566	0.548566	2.357884	-0.3727023
Modelo	88	77	324	88	77

Tabla 4.60: Criterios Clásicos de ajuste para SARIMA(0, 1, 1)(0, 1, 1) $_7$.

	U_1	U_2	ECMpre
Valor	0.01307016	0.01307128	0.052905
Modelo	317	317	123

Tabla 4.61: Criterios Clásicos de predicción para SARIMA(0, 1, 1)(0, 1, 1) $_7$.

	Ajuste		Predicción	
	WMSE	WMAE	WFMSE	WFMAE
Valor	0.1671818	0.2732851	0.04187303	0.1648211
Modelo	178	177	123	123

Tabla 4.62: Criterios propuestos para SARIMA(0, 1, 1)(0, 1, 1) $_7$.

Observando las tablas 4.60, 4.61 y 4.62, vemos como el criterio AIC y BIC seleccionan el mismo modelo, el criterio AIC2 y BIC2 también coinciden en su elección, siendo además éste el modelo generador de la serie. En el caso de los criterios de Theil, una vez más, su elección es la misma, los criterios ECMpre, WFMSE y WFMAE también se decantan por el mismo modelo (ver apéndice C).

A la vista de los gráficos 4.99 y 4.100 se observa que la distribución de los residuos en cada uno de los criterios no presenta claras diferencias entre si, esto mismo se puede ver en los gráficos 4.101 y 4.102, no obstante será en las tablas 4.63, 4.64 y 4.66 donde se intentará obtener diferencias entre ellos y establecer con cuál se consiguen mejores resultados.

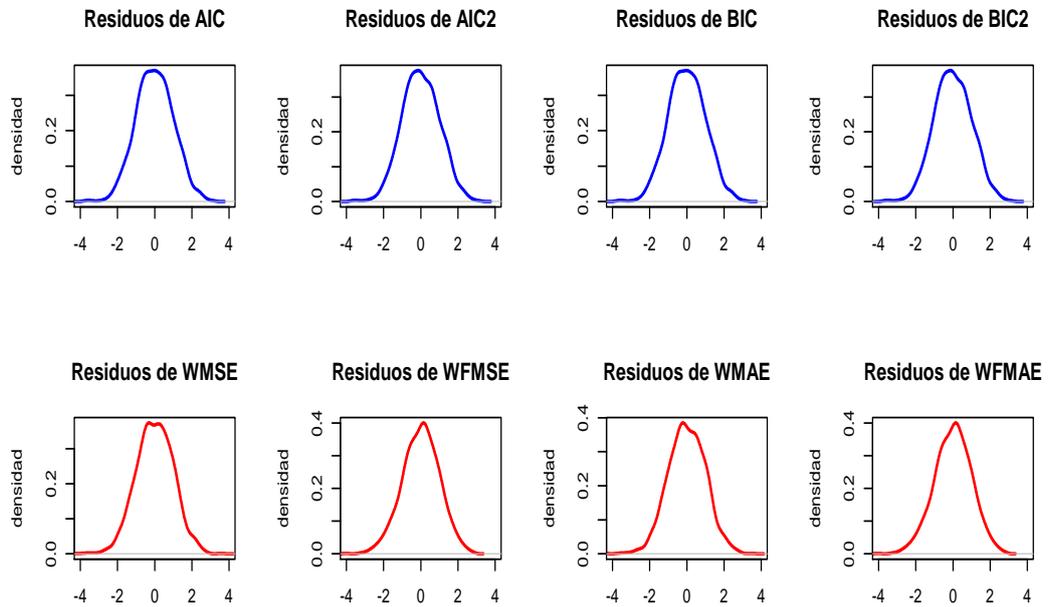


Figura 4.99: Residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(0, 1, 1)(0, 1, 1)₇.

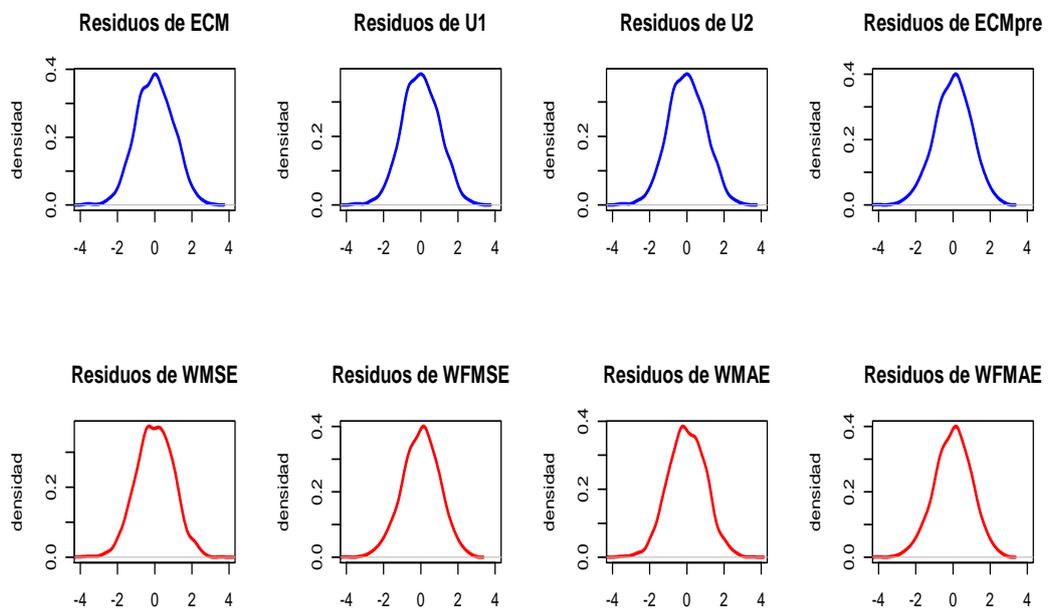


Figura 4.100: Residuos obtenidos con los modelos de los criterios de predicción de SARIMA(0, 1, 1)(0, 1, 1)₇.

Criterio	Media	Varianza
AIC	-0.0276	1.3236
AIC2	-0.0218	1.2886
ECM	-0.0241	1.3499
BIC	-0.0276	1.3236
BIC2	-0.0218	1.2886
U1	-0.0350	1.3301
U2	-0.0350	1.3301
ECMpre	-0.0601	2.4202
WMSE	-0.0203	1.5578
WFMSE	-0.0601	2.4202
WMAE	-0.0073	1.5632
WFMAE	-0.0601	2.4202

Tabla 4.63: Medidas descriptivas de los residuos en el conjunto de comprobación para SARIMA(0, 1, 1)(0, 1, 1)₇.

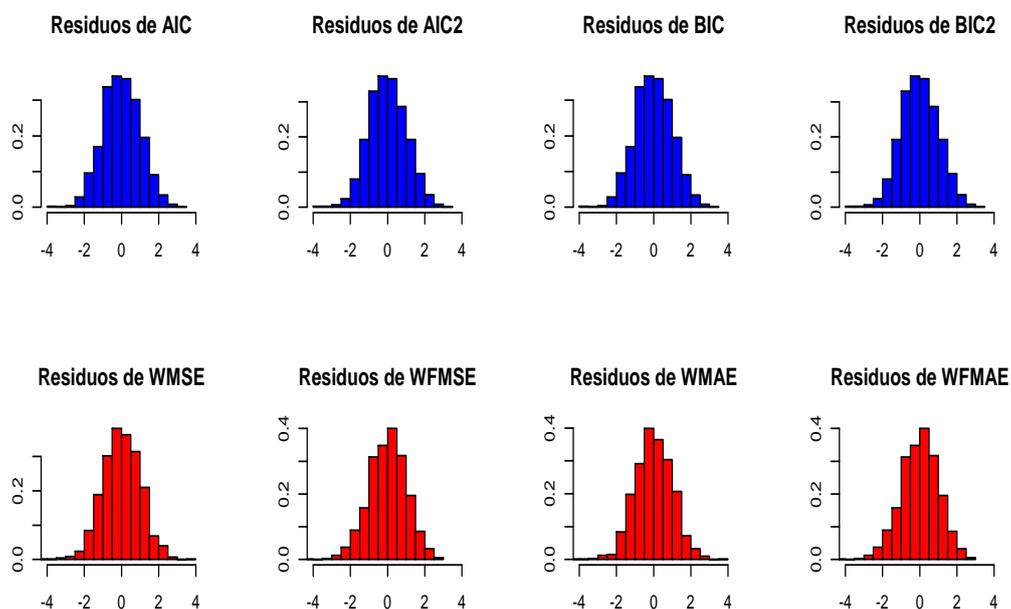


Figura 4.101: Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(0, 1, 1)(0, 1, 1)₇.

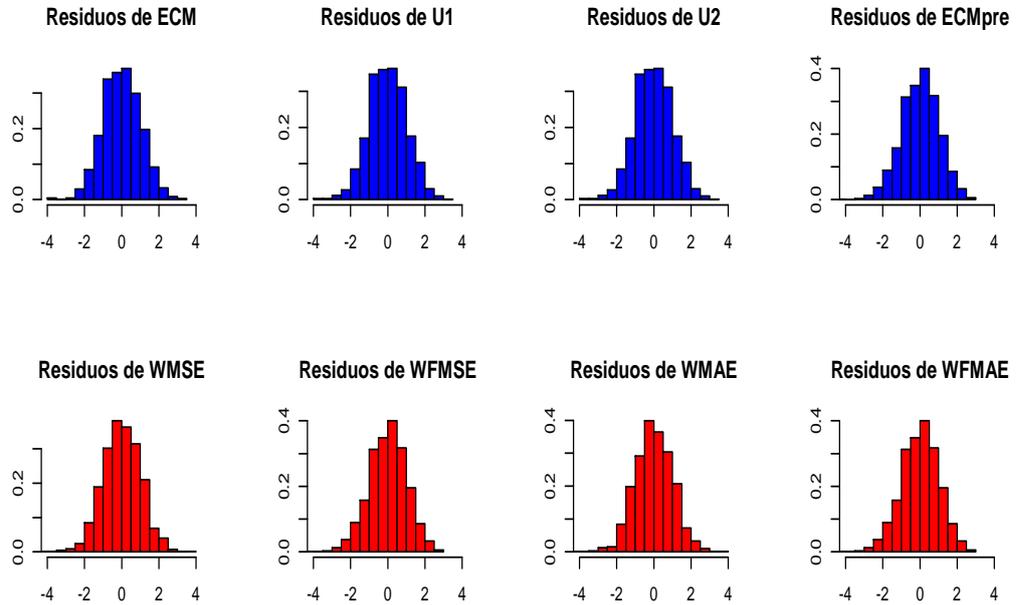


Figura 4.102: Histograma de los residuos obtenidos con los modelos de los criterios de predicción de $\text{SARIMA}(0, 1, 1)(0, 1, 1)_7$.

Criterio	ECMpre1	ECMpre2	ECMpre3	ECMpre4	ECMpre5	ECMpre6	ECMpre7
AIC	1.2021	1.1712	1.2984	1.2613	1.2357	1.5516	1.5505
AIC2	1.1652	1.0984	1.2602	1.2206	1.2074	1.5365	1.5354
ECM	1.2498	1.2094	1.3450	1.2802	1.2509	1.5539	1.5638
BIC	1.2021	1.1712	1.2984	1.2613	1.2357	1.5516	1.5505
BIC2	1.1652	1.0984	1.2602	1.2206	1.2074	1.5365	1.5354
U1	1.2026	1.1327	1.3114	1.2788	1.2261	1.5897	1.5781
U2	1.2026	1.1327	1.3114	1.2788	1.2261	1.5897	1.5781
ECMpre	1.6562	1.9802	2.1458	2.3180	2.5154	2.9032	3.4480
WMSE	1.3705	1.4778	1.5533	1.5610	1.5055	1.6758	1.7633
WFMSE	1.6562	1.9802	2.1458	2.3180	2.5154	2.9032	3.4480
WMAE	1.3947	1.4688	1.5768	1.5847	1.5457	1.6465	1.7257
WFMAE	1.6562	1.9802	2.1458	2.3180	2.5154	2.9032	3.4480

Tabla 4.64: ECMpre global para cada predicción realizada en el conjunto de comprobación en $\text{SARIMA}(0, 1, 1)(0, 1, 1)_7$.

En la tabla 4.63 podemos ver las medidas descriptivas para los residuos obtenidos en el conjunto de comprobación haciendo uso del modelo mínimo seleccionado por cada criterio en cada una de las 200 series estudiadas, se observa unas medias bajas, por lo que la media de errores es bastante buena, y una buena varianza también, más o menos similar en los distintos criterios. Estudiando las predicciones individuales (tabla 4.64), se puede determinar

que en este caso, al contrario que ocurría en los procesos no estacionarios, los errores de predicción se mantienen más o menos homogéneos a medida que avanzamos el instante de predicción, además no se aprecian diferencias importantes en ningún criterio concreto. Veamos a continuación lo que ocurre con el ajuste.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.87	0.9674	0.8092	0.87	0.9674	0.8711
	U2	ECM _{pre}	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8711	1.1764	0.9929	1.1764	1.0181	1.1764

Tabla 4.65: El ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(0, 1, 1)(0, 1, 1) $_7$.

En cuanto al ajuste, en este tipo de modelos, observamos en la tabla 4.65, que al igual que en los procesos estacionarios los criterios propuestos no establecen buenos resultados al respecto y son los clásicos los que consiguen los mejores ajustes.

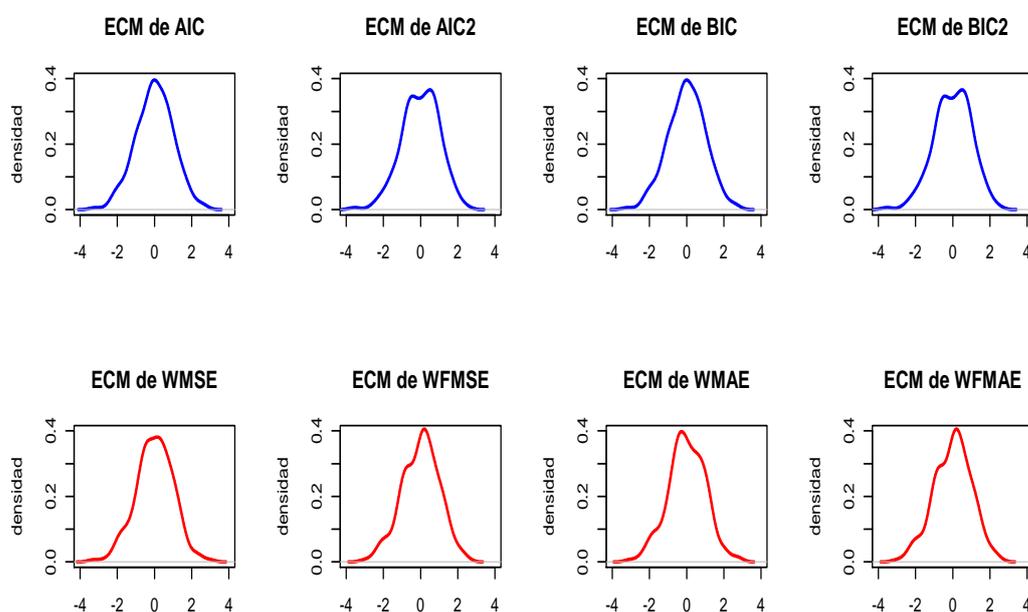


Figura 4.103: ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(0, 1, 1)(0, 1, 1) $_7$.

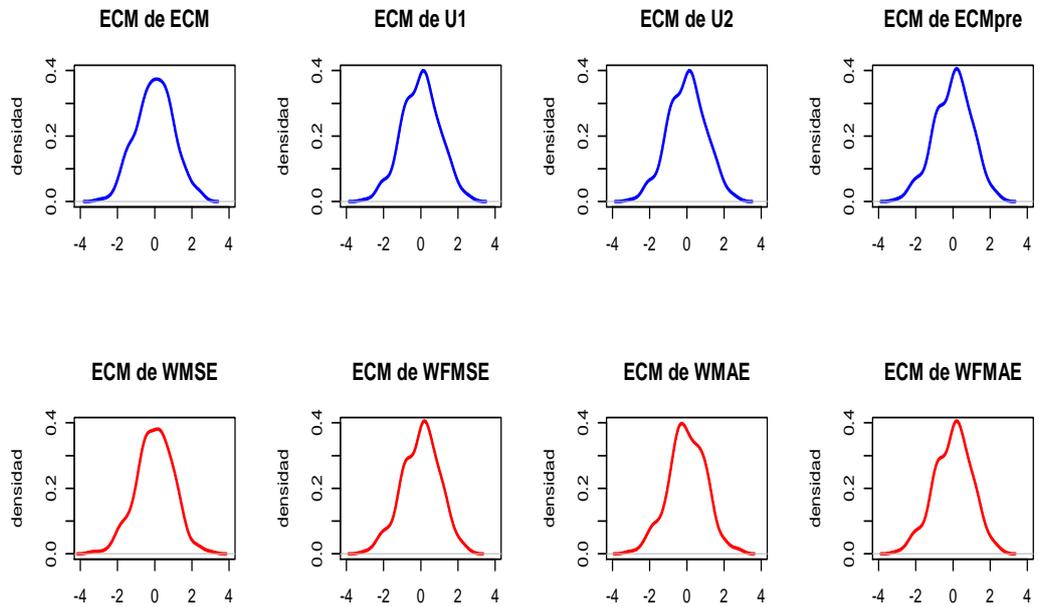


Figura 4.104: ECM de ajuste de los modelos obtenidos con los criterios de predicción en $\text{SARIMA}(0, 1, 1)(0, 1, 1)_7$.

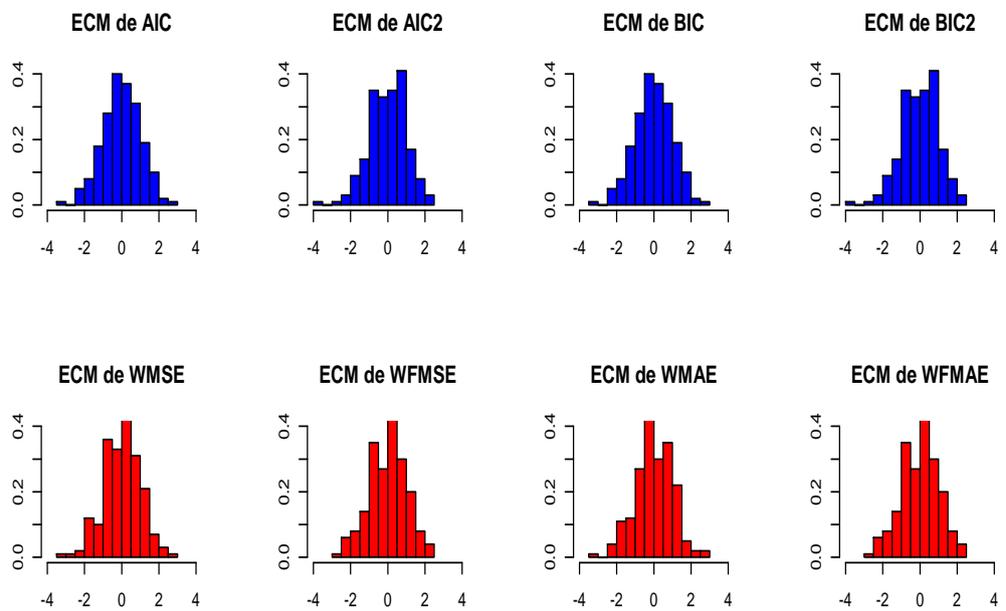


Figura 4.105: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en $\text{SARIMA}(0, 1, 1)(0, 1, 1)_7$.

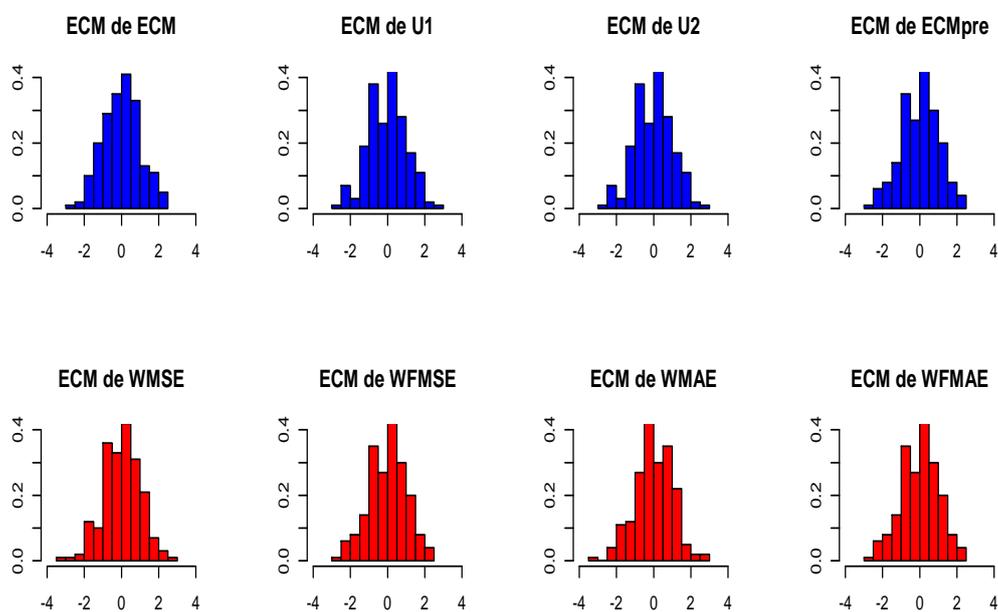


Figura 4.106: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicci3n en $SARIMA(0, 1, 1)(0, 1, 1)_7$.

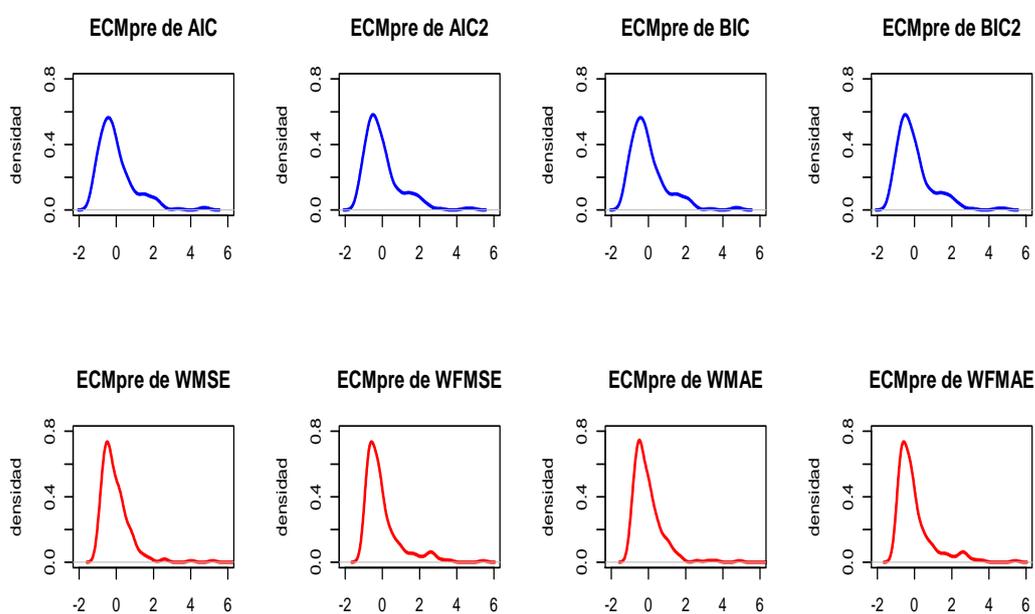


Figura 4.107: ECM de predicci3n de los modelos obtenidos con los criterios de ajuste en $SARIMA(0, 1, 1)(0, 1, 1)_7$.

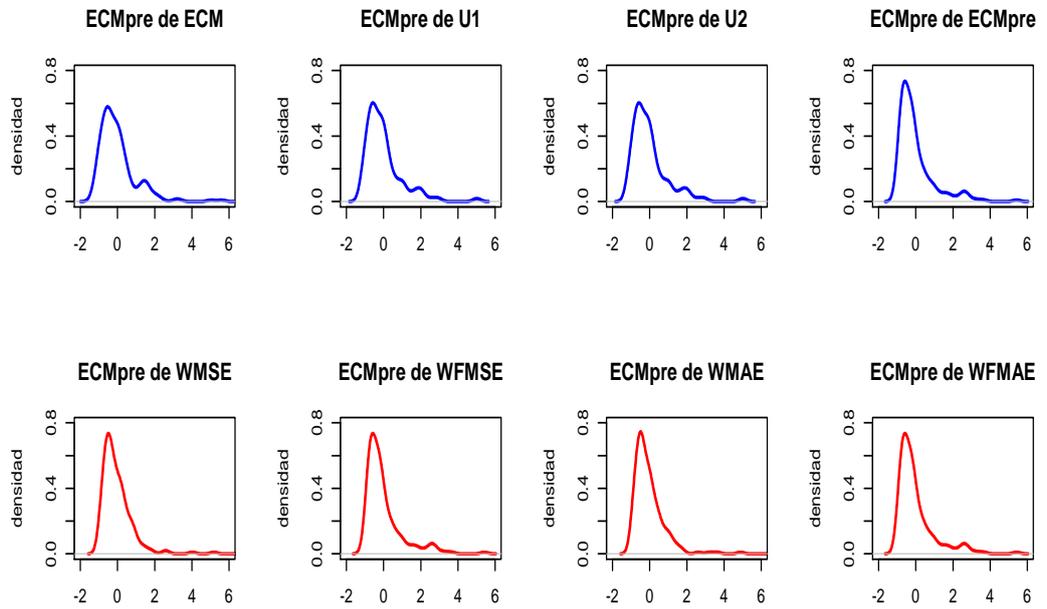


Figura 4.108: ECM de predicción de los modelos obtenidos con los criterios de predicción en $\text{SARIMA}(0, 1, 1)(0, 1, 1)_7$.

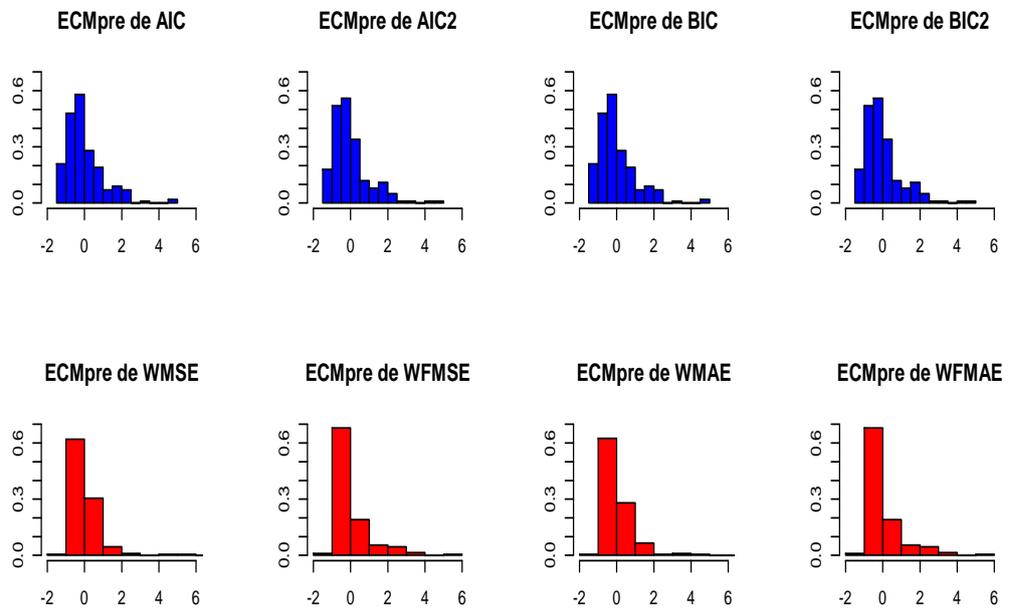


Figura 4.109: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en $\text{SARIMA}(0, 1, 1)(0, 1, 1)_7$.

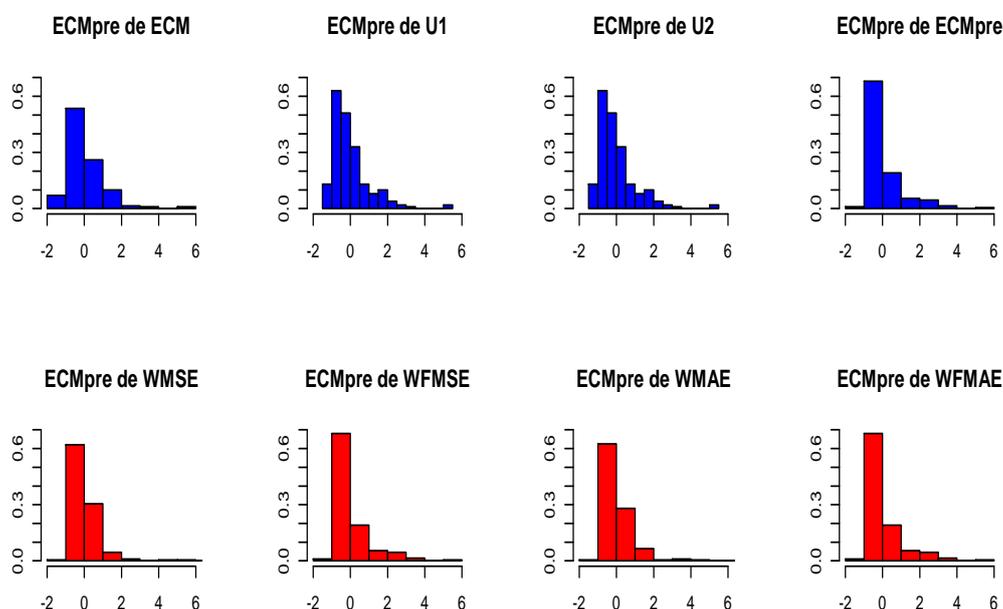


Figura 4.110: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA($0, 1, 1$)($0, 1, 1$) $_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	1.3244	1.2891	1.3504	1.3244	1.2891	1.3313
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	1.3313	2.4238	1.5582	2.4238	1.5633	2.4238

Tabla 4.66: ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA($0, 1, 1$)($0, 1, 1$) $_7$.

En la tabla 4.66 donde se recogen los ECM globales de predicción, podemos observar que los resultados obtenidos con los diversos criterios no son satisfactorios, lo que nos lleva a estudiar el caso de diferenciar entre las series donde se detecta el modelo por el que fueron simuladas de aquellas en que no es posible identificarlas.

Los criterios que han sido capaces de detectar en alguna serie (al menos 2 de las 200) el modelo generador de los datos han sido AIC2, BIC, BIC2 y WFMAE.

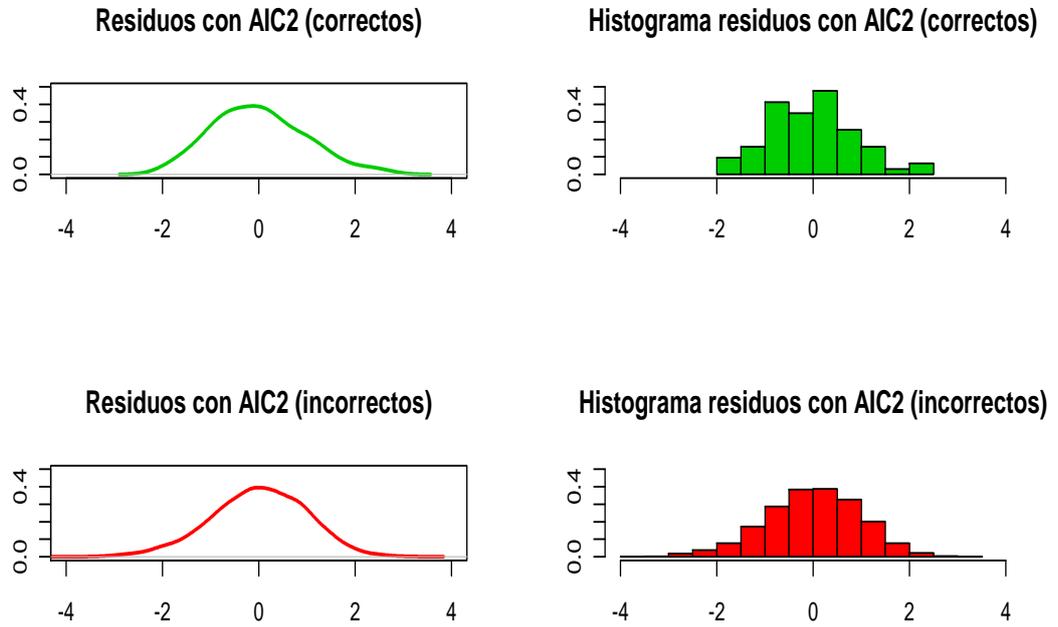


Figura 4.111: Residuos de las series correctas e incorrectas con el criterio AIC2 en el SARIMA(0, 1, 1)(0, 1, 1)₇.

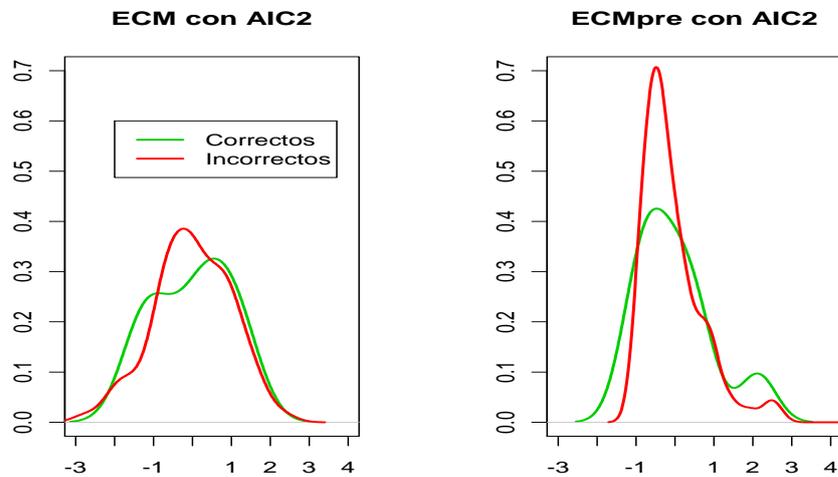


Figura 4.112: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en SARIMA(0, 1, 1)(0, 1, 1)₇.

	Media	Varianza	ECMG	ECMGpre
Correctos	0.1823	1.3225	0.7746	1.3557
Incorrectos	-0.0614	1.4110	0.8316	1.4148

Tabla 4.67: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(0, 1, 1)(0, 1, 1) $_7$ con el criterio AIC2.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	1.0259	0.6201	0.6411	1.5699	2.1886	1.3848	2.0597
Incorrectos	1.1810	1.2162	1.2064	1.4520	1.4768	1.7430	1.6283

Tabla 4.68: ECM global de predicción en SARIMA(0, 1, 1)(0, 1, 1) $_7$ con el criterio AIC2 en el conjunto de comprobación.

A la vista del gráfico 4.111 las diferencias que se aprecian entre los dos grupos de series son mínimas, si observamos los resultados recogidos en la tabla 4.67 se puede comprobar como la varianza en ambas series, el ECMG y el ECMGpre toman valores muy similares, aunque en las series incorrectas se consigue una media de error inferior que en las correctas.

En el gráfico 4.112 se puede apreciar con un poco más de claridad las diferencias establecidas entre ambos grupos sobre todo con respecto al ECMpre. Sin embargo, ante la tabla 4.68 se consiguen menores errores de predicción en los primeros instantes de predicción en las series correctas aumentando en las últimas estimaciones realizadas, en cambio en las series incorrectas los errores cometidos se mantienen homogéneos a medida que aumenta el instante de predicción.

Con el criterio que acabamos de estudiar, AIC2, se consigue que el grupo de las series bien identificadas esté formado por 9 de las series estudiadas quedando por tanto, en el grupo de las incorrectas un tamaño de 191 series. Continuamos con el estudio del criterio BIC, donde se establece un menor porcentaje de identificación, siendo éste tan sólo del 1,5% de series correctamente identificadas.

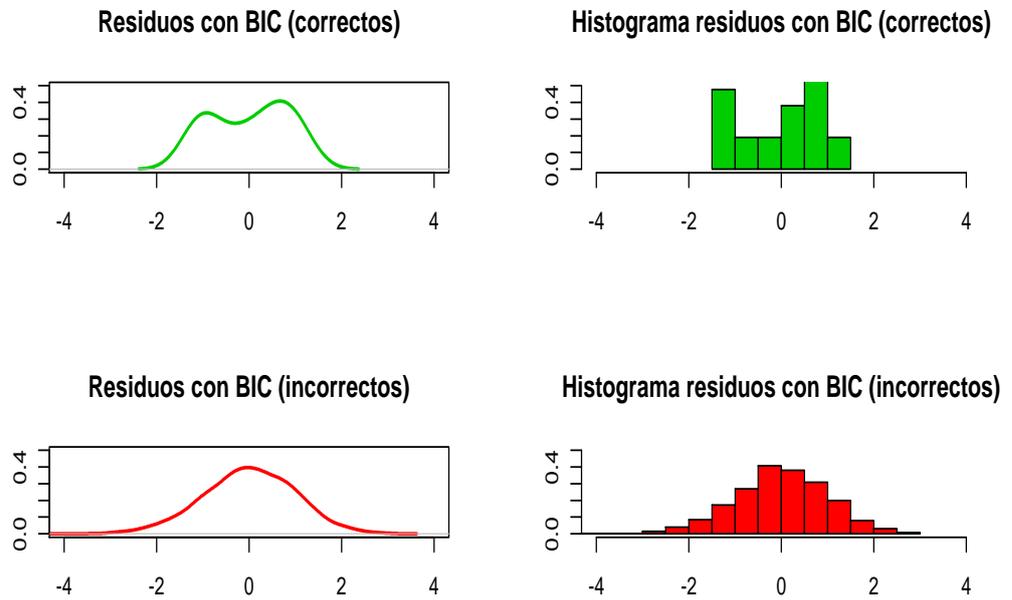


Figura 4.113: Residuos de las series correctas e incorrectas con el criterio BIC en el SARIMA(0, 1, 1)(0, 1, 1)₇.

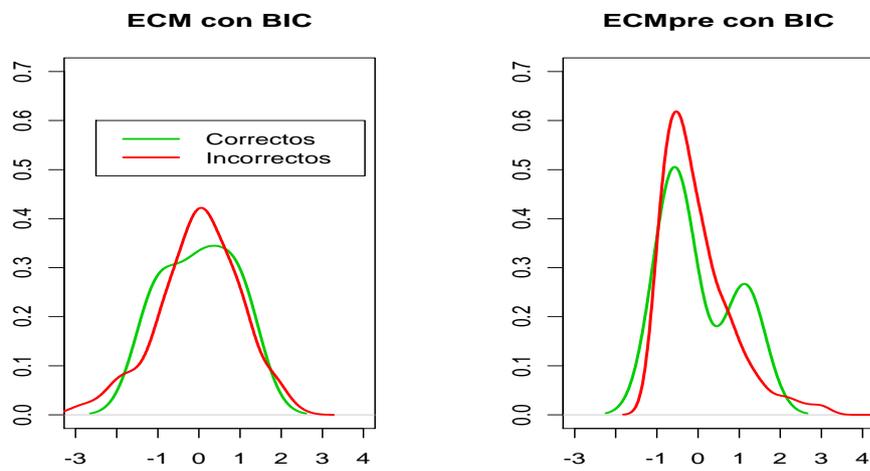


Figura 4.114: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC en SARIMA(0, 1, 1)(0, 1, 1)₇.

	Media	Varianza	ECMG	ECMGpre
Correctos	0.8490	0.5715	0.8258	1.2924
Incorrectos	-0.0405	1.3776	0.8347	1.3792

Tabla 4.69: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para SARIMA(0, 1, 1)(0, 1, 1) $_7$ con el criterio BIC.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.5035	0.4215	1.4810	0.3666	2.7579	1.5573	1.9588
Incorrectos	1.1844	1.2040	1.1970	1.4186	1.4411	1.6454	1.5638

Tabla 4.70: ECM global de predicción en SARIMA(0, 1, 1)(0, 1, 1) $_7$ con el criterio BIC en el conjunto de comprobación.

En los gráficos 4.113 se comprueba como la curva que más se asemeja a una normal es la obtenida en el caso de las series incorrectas, sin embargo y a la vista de los datos recogidos en la tabla 4.69 aunque la menor media se consiga en este grupo, en el grupo de series correctas se alcanza una varianza menor, y por tanto, una leve mejoría en cuanto al ECMGpre, comportando el ajuste de manera muy similar entre los dos grupos donde no se pueden establecer diferencias significativas al respecto, lo que puede comprobarse de manera gráfica en el gráfico 4.114, donde además tampoco se observa diferencias importantes en cuanto al ECMpre, aunque cabe destacar que según la tabla 4.70, es en las series correctas donde se consigue una mejor precisión es la estimación de los dos primeros instantes de predicción, comportándose el aumento de los errores a lo largo del horizonte de predicción de manera similar a lo que ocurría con el criterio AIC2.

Con el criterio BIC2 que se mostrará a continuación, el comportamiento global de los residuos y de los ECMG de ajuste y predicción se distribuyen de manera similar a los criterios ya estudiados (tabla 4.71), cabe mencionar que en cuanto a los ECMGpre individual (tabla 4.72), en esta ocasión no se establecen diferencias entre los grupos, y el error cometido a lo largo de las 7 predicciones realizadas es homogéneo y sin diferencias entre grupos.

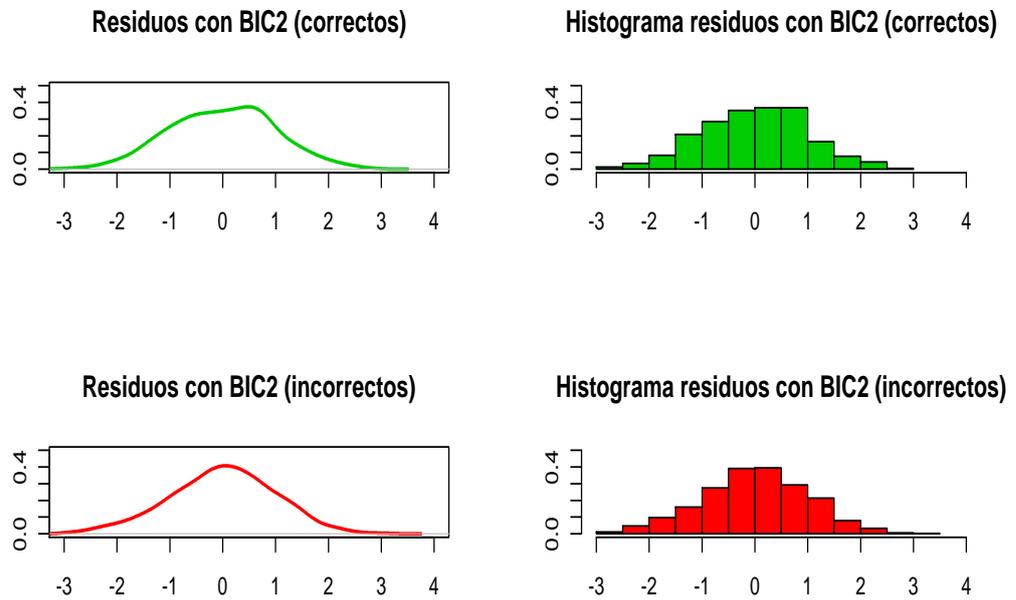


Figura 4.115: Residuos de las series correctas e incorrectas con el criterio BIC2 en el SARIMA(0, 1, 1)(0, 1, 1)₇.

	Media	Varianza	ECMG	ECMGpre
Correctos	0.1289	1.4871	0.8655	1.5038
Incorrectos	-0.0828	1.2496	0.8837	1.2565

Tabla 4.71: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(0, 1, 1)(0, 1, 1)₇ con el criterio BIC2.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	1.2306	1.2628	1.2451	1.5824	1.6945	1.8784	1.6325
Incorrectos	1.1554	1.0716	1.1307	1.3071	1.2989	1.4269	1.4049

Tabla 4.72: ECM global de predicción en SARIMA(0, 1, 1)(0, 1, 1)₇ con el criterio BIC2 en el conjunto de comprobación.

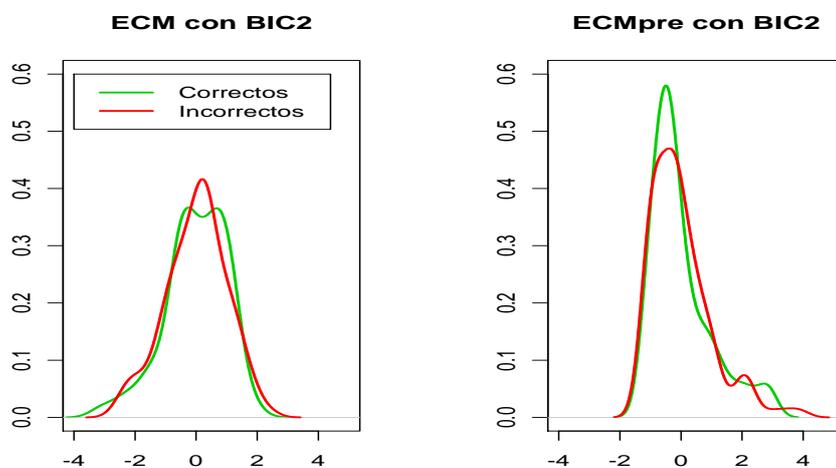


Figura 4.116: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en $SARIMA(0, 1, 1)(0, 1, 1)_7$.

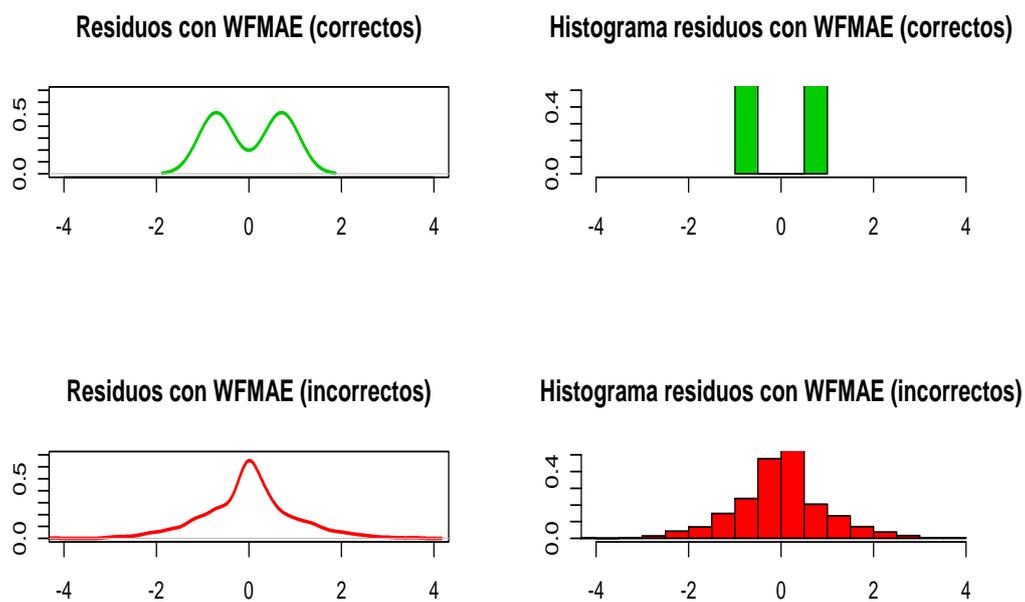


Figura 4.117: Residuos de las series correctas e incorrectas con el criterio WFMAE en el $SARIMA(0, 1, 1)(0, 1, 1)_7$.

	Media	Varianza	ECMG	ECMGpre
Correctos	-0.0033	0.6759	1.0596	0.6759
Incorrectos	-0.0147	0.8388	1.3764	0.8390

Tabla 4.73: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(0, 1, 1)(0, 1, 1)₇ con el criterio WFMAE.

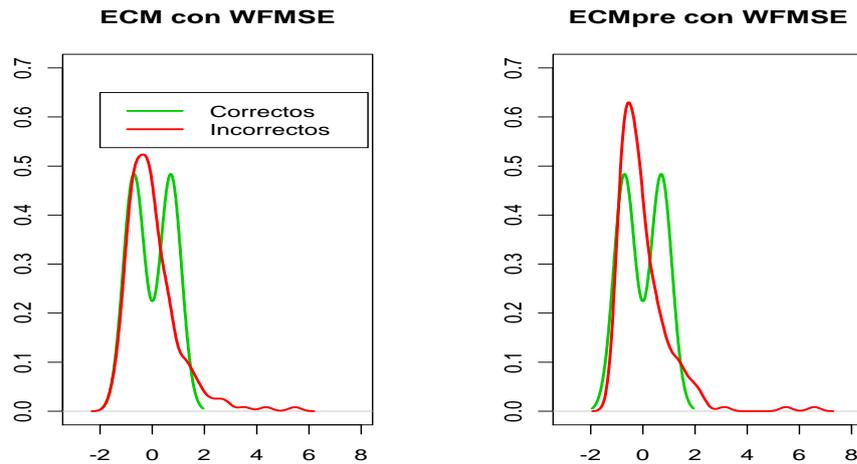


Figura 4.118: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en SARIMA(0, 1, 1)(0, 1, 1)₇.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.3327	0.1381	0.0311	2.7731	0.2835	0.8764	0.2962
Incorrectos	0.3882	0.5871	0.7715	0.8154	0.9265	1.0181	1.3663

Tabla 4.74: ECM global de predicción en SARIMA(0, 1, 1)(0, 1, 1)₇ con el criterio WFMAE en el conjunto de comprobación.

A la vista de la gráfica 4.118 no se aprecian diferencias significativas entre conseguir detectar el modelo generador de no conseguirlo, sin embargo en la gráfica 4.117 se muestra como el tamaño del grupo de las series correctas es tan sólo de dos series, con lo que no podemos establecer conclusiones fiables al respecto del comportamiento de este criterio en los grupos establecidos. En la tabla 4.73 se establecen una ciertas mejoras en los resultados obtenidos en las series correctas contra las incorrectas sin ser éstos decisivos, en cuanto a las predicciones individuales (tabla 4.74) el comportamiento no presenta diferencias notables. En relación a los otros criterios AIC2, BIC, y BIC2 se consigue con este criterio, WFMAE unos resultados más esperanzadores.

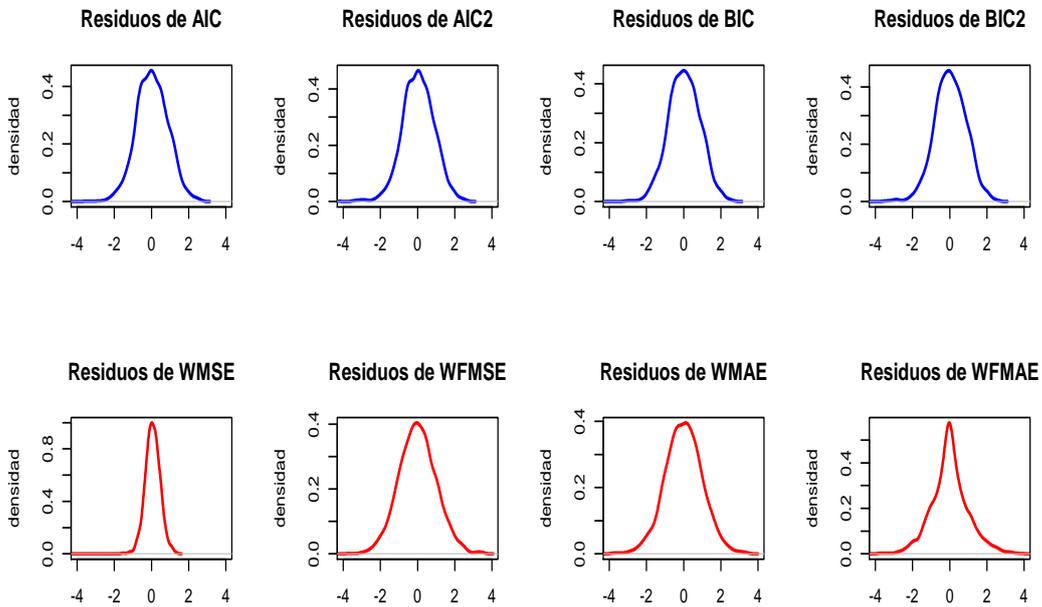


Figura 4.119: Residuos obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(0, 1, 1)(0, 1, 1)_7$.

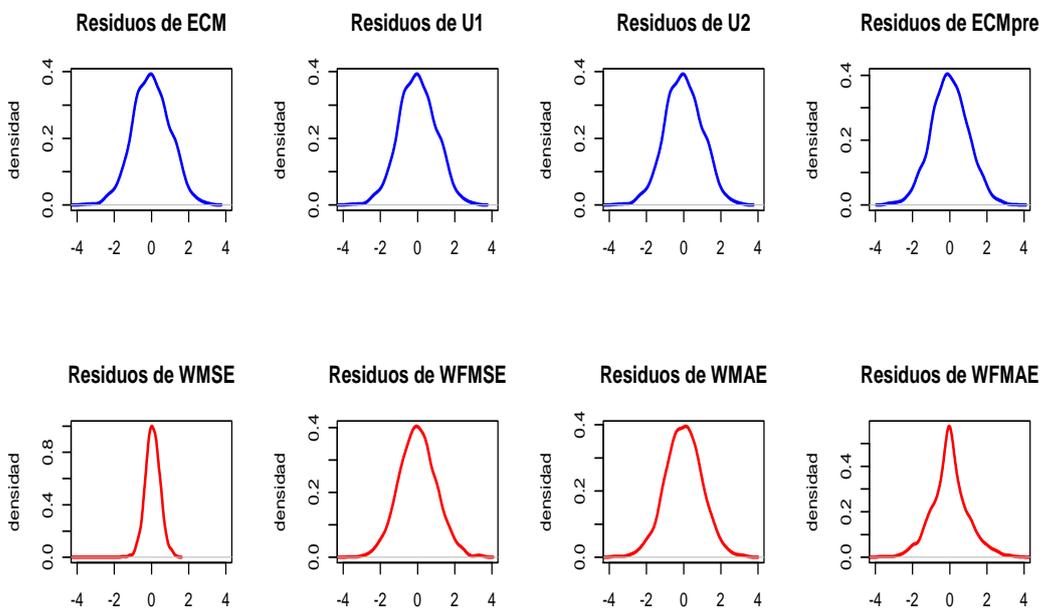


Figura 4.120: Residuos obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(0, 1, 1)(0, 1, 1)_7$.

Criterio	Media	Varianza
AIC	-0.0311	1.44
AIC2	-0.0505	1.4096
ECM	-0.0254	1.4275
BIC	-0.0272	1.3772
BIC2	-0.0130	1.3379
U1	-0.0254	1.4275
U2	-0.0254	1.4275
ECMpre	0.0080	0.6807
WMSE	0.1454	6.6399
WFMSE	-0.0160	0.7182
WMAE	0.0107	1.76
WFMAE	-0.0146	0.8372

Tabla 4.75: Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para SARIMA(0, 1, 1)(0, 1, 1)₇.

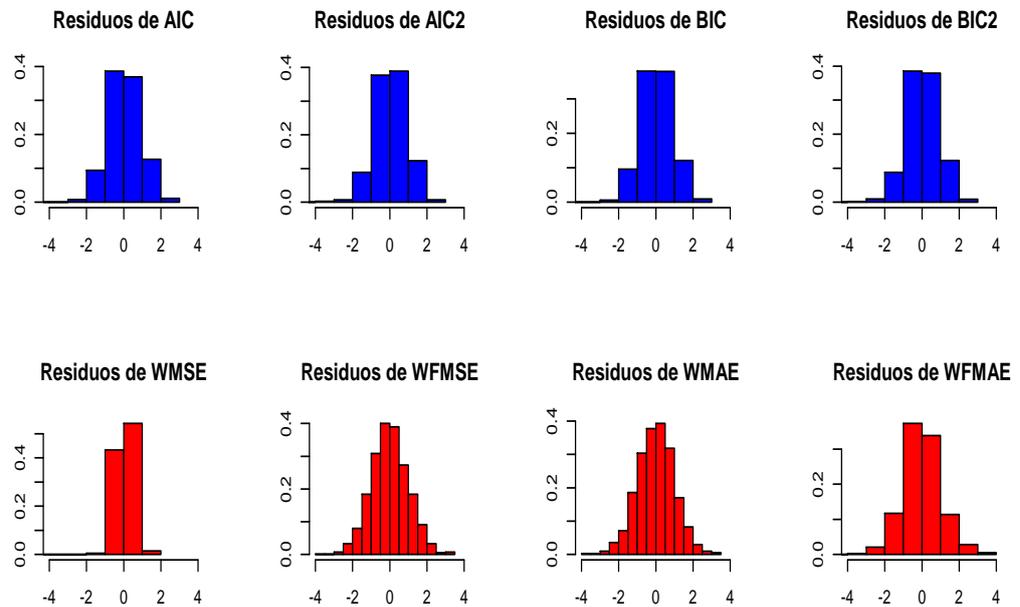


Figura 4.121: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇.

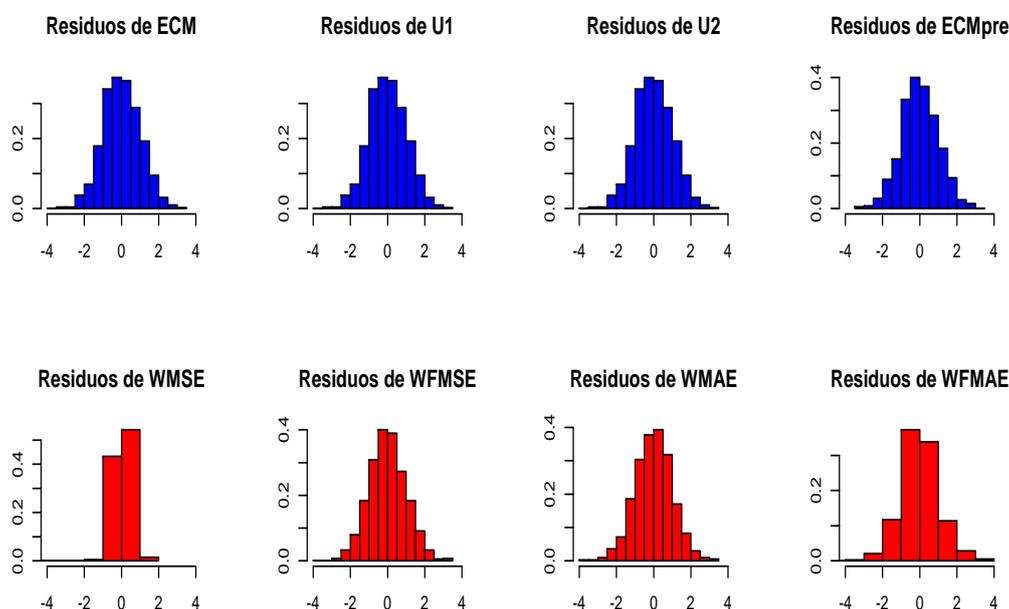


Figura 4.122: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1) $_7$.

Criterio	ECMpre1	ECMpre2	ECMpre3	ECMpre4	ECMpre5	ECMpre6	ECMpre7
AIC	1.2457	1.2527	1.1712	1.4835	1.4942	1.7336	1.7060
AIC2	1.1741	1.1894	1.1809	1.4573	1.5088	1.7269	1.6478
ECM	1.2505	1.2237	1.1740	1.4562	1.5123	1.6860	1.6940
BIC	1.1742	1.1923	1.2012	1.4028	1.4609	1.6440	1.5697
BIC2	1.1802	1.1347	1.1685	1.3979	1.4295	1.5759	1.48
U1	1.2505	1.2237	1.1740	1.4562	1.5123	1.6860	1.6940
U2	1.2505	1.2237	1.1740	1.4562	1.5123	1.6860	1.6940
ECMpre	0.6322	0.6623	0.7470	0.7056	0.7153	0.6757	0.6273
WMSE	4.7641	6.3009	6.3996	6.6371	7.2854	7.8307	7.4693
WFMSE	0.4371	0.5443	0.7040	0.6784	0.8457	0.8423	0.9770
WMAE	1.5392	1.5381	1.7041	1.5880	2.0374	2.0906	1.8237
WFMAE	0.3877	0.5826	0.7641	0.8350	0.9201	1.0167	1.3556

Tabla 4.76: ECM globales para cada predicción realizada en el conjunto de comprobación en SARIMA(0, 1, 1)(0, 1, 1) $_7$.

En la tabla 4.75 se observa una disminución de la varianza en los criterios ECMpre, WFMSE y WFMAE con respecto al resto de criterios y con respecto a los resultados obtenidos en estos mismos criterios en la tabla 4.63, donde la mejoría es más acusada. Atendiendo a la tabla 4.76 las mejores predicciones se consiguen en los criterios mencionados manteniéndose además a lo largo de todo el horizonte de predicción y mejorando también los resultados analizados en la tabla 4.64.

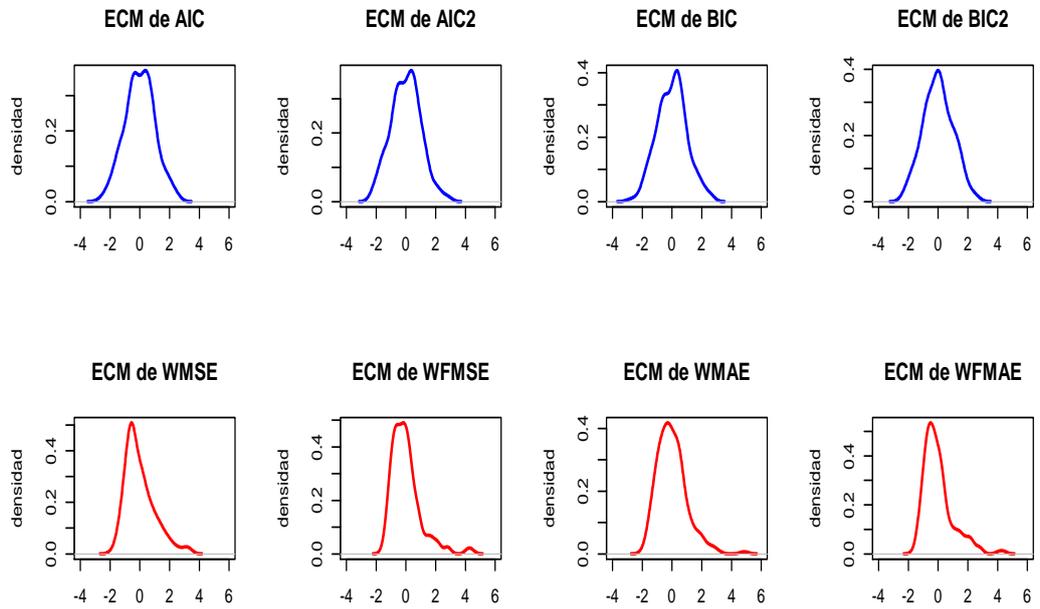


Figura 4.123: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(0, 1, 1)(0, 1, 1)_7$.

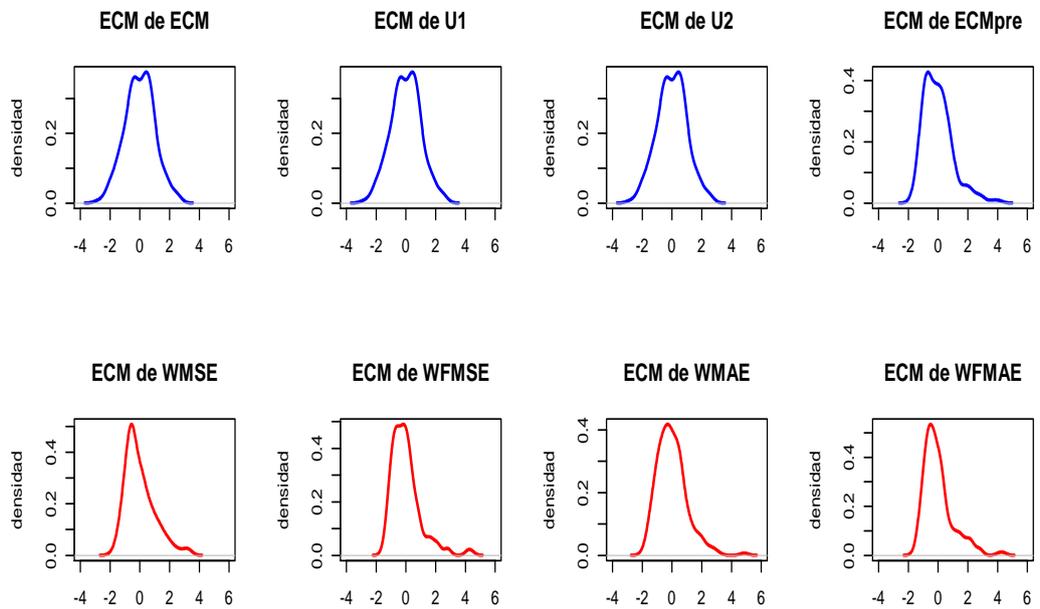


Figura 4.124: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(0, 1, 1)(0, 1, 1)_7$.

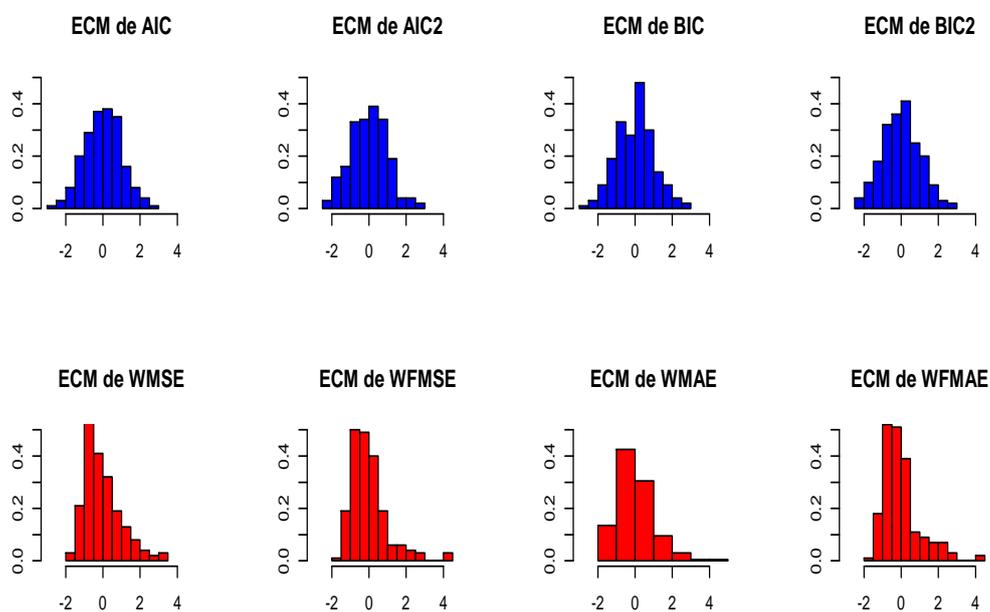


Figura 4.125: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(0, 1, 1)(0, 1, 1)_7$.

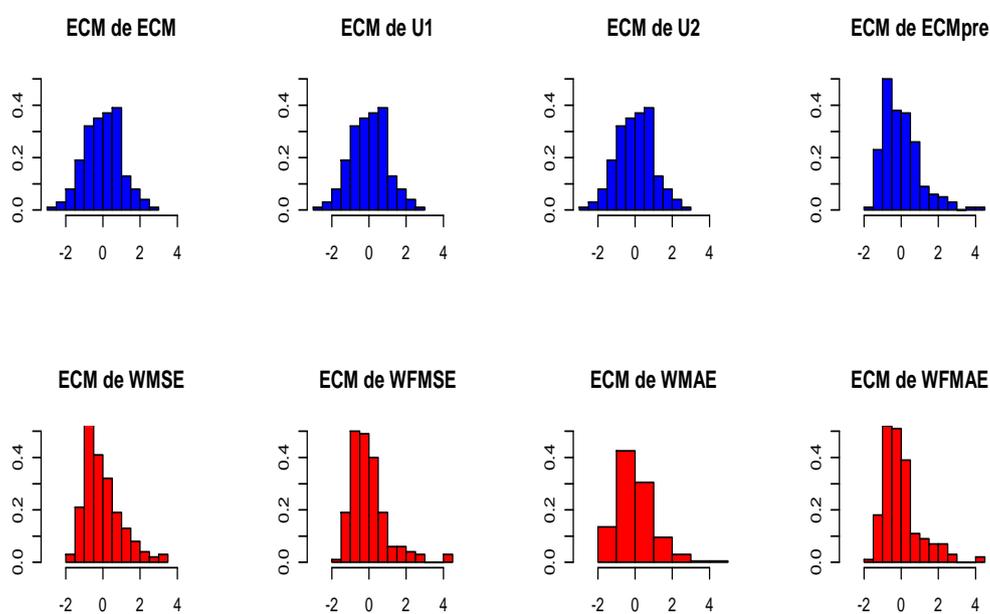


Figura 4.126: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(0, 1, 1)(0, 1, 1)_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8053	0.8290	0.7978	0.8345	0.8770	0.7978
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.7978	1.2762	1.0343	1.3325	1.0895	1.3732

Tabla 4.77: ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(0, 1, 1)(0, 1, 1)₇.

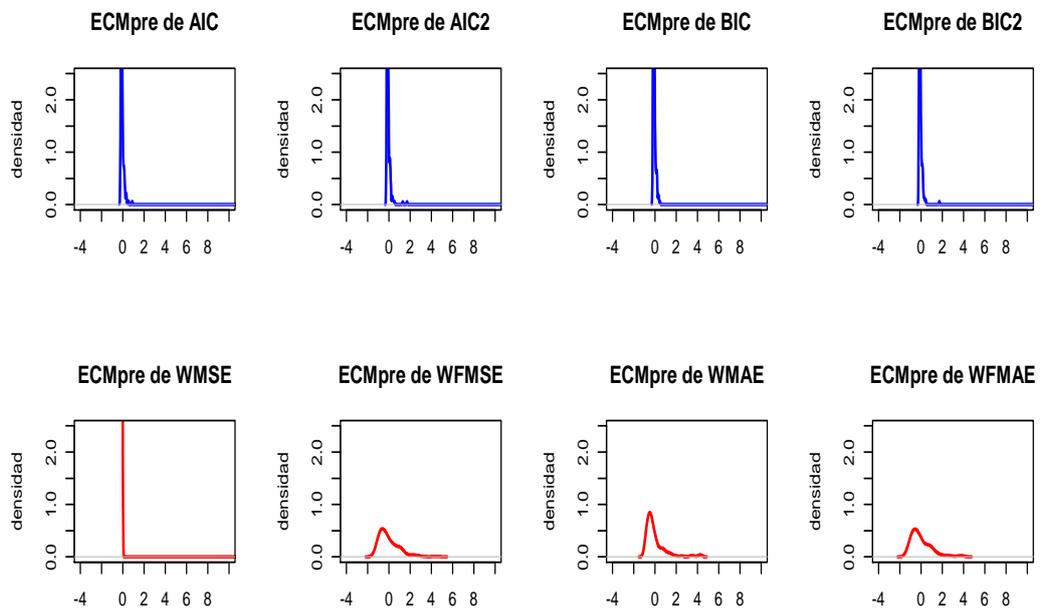


Figura 4.127: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇.

En la tabla 4.77 obtenemos lo que viene siendo habitual a lo largo de los estudios realizados, y es que el mejor ajuste se alcanza con los criterios clásicos, siendo el peor de los ajustes coincidentes en los criterios donde se consiguen los mejores resultados de predicción. Este comportamiento se observa de manera visual en las gráficas 4.123 a 4.126.

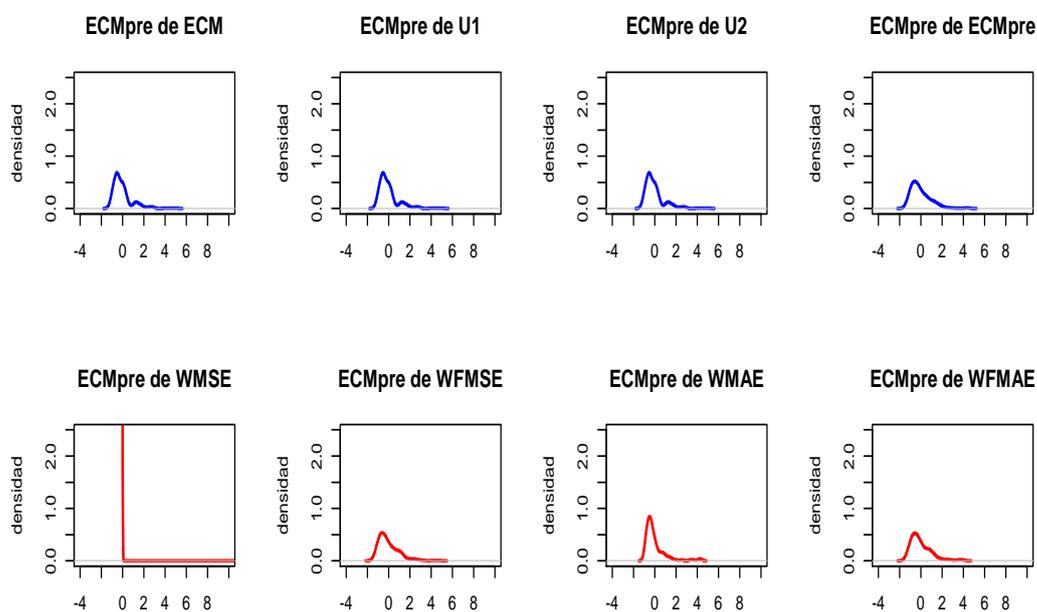


Figura 4.128: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇.

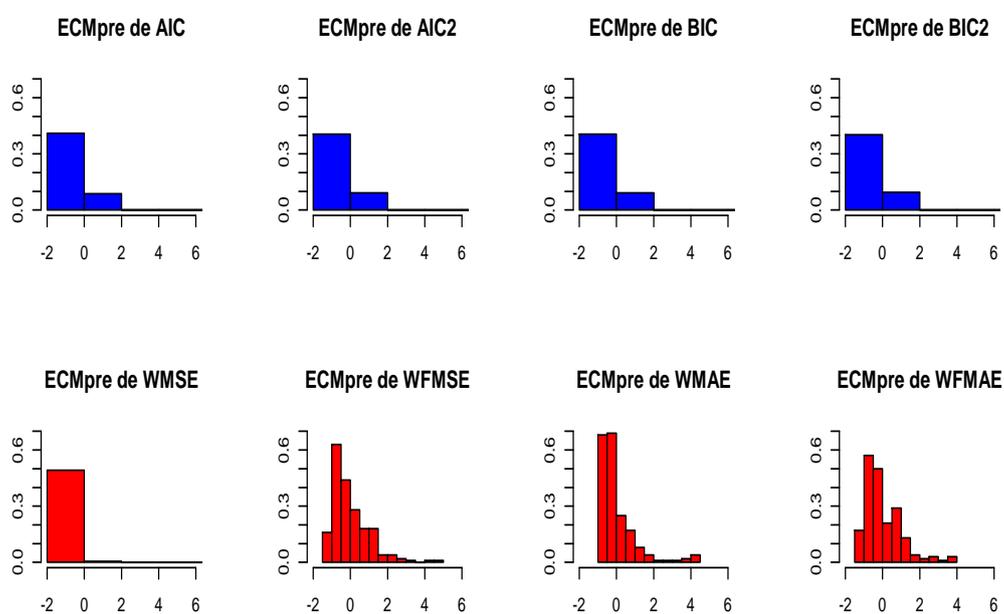


Figura 4.129: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇.

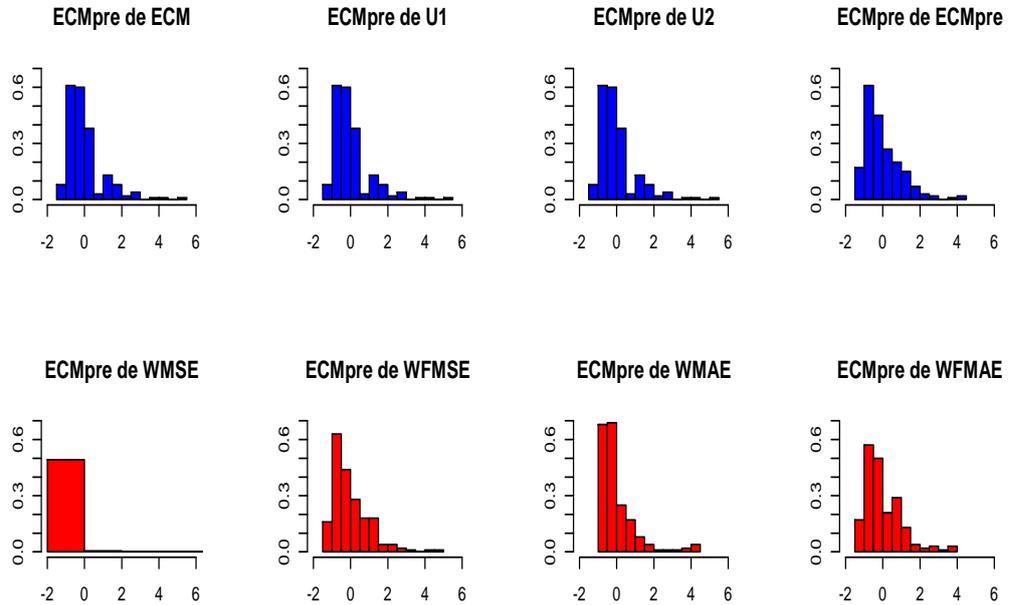


Figura 4.130: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(0, 1, 1)(0, 1, 1)₇.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	1.4410	1.4122	1.4281	1.3779	1.3381	1.4281
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	1.4281	0.6808	6.6610	0.7184	1.7602	0.8374

Tabla 4.78: ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(0, 1, 1)(0, 1, 1)₇.

En esta tabla, 4.78 se aprecian diferencias significativas en cuanto a los resultados globales de predicción, donde se alcanzan los mejores resultados en los criterios ECMpre, WFMSE y WFMAE, no obstante, es preferible el empleo de los criterios de predicción propuestos ante el ECMpre, ya que no sólo en los resultados globales se mejora con respecto al resto sino también en los resultados individuales como veíamos en la tabla 4.76. Los resultados de la tabla 4.66 también son mejorados en los criterios de predicción nombrados por los expuestos en la tabla superior (4.78), donde se aprecia la importante disminución de los errores de predicción, de forma inversa se observa también como el ajuste empeora de la tabla 4.65 a los resultados de la tabla 4.77 atendiendo a los criterios que se han citado como los mejores para la predicción, ECMpre, WFMSE y WFMAE.

Modelo SARIMA(1, 1, 0)(0, 1, 1) $_7$, cuyos coeficientes son $\phi_1 = 0,7$ y $\Theta_1 = 0,5$. (Modelo 167)

	AIC	AIC2	ECM	BIC	BIC2
Valor	2.304187	-0.4141506	0.551709	2.358869	-0.3635506
Modelo	178	293	324	178	167

Tabla 4.79: Criterios Clásicos de ajuste para SARIMA(1, 1, 0)(0, 1, 1) $_7$.

	U_1	U_2	ECMpre
Valor	1.300835e-3	1.300837e-3	0.06946
Modelo	303	303	133

Tabla 4.80: Criterios Clásicos de predicción para SARIMA(1, 1, 0)(0, 1, 1) $_7$.

	Ajuste		Predicción	
	WMSE	WMAE	WFMSE	WFMAE
Valor	0.1203786	0.2395678	0.070043	0.2181408
Modelo	88	160	133	133

Tabla 4.81: Criterios propuestos para SARIMA(1, 1, 0)(0, 1, 1) $_7$.

En la tabla 4.79, se puede comprobar como el criterio AIC y BIC seleccionan el mismo modelo para predecir, en esta ocasión ese modelo es SARIMA(1, 1, 0)(2, 1, 0) $_7$, es el criterio BIC2 el único que seleccionan el modelo coincidente con el modelo SARIMA utilizado para generar las series. En la tabla 4.80, se tiene una vez más que los criterios de Theil han elegido el mismo modelo (SARIMA(2, 1, 1)(2, 0, 2) $_7$) y observando también la tabla 4.81, veremos, como los criterios ECMpre, WFMSE y WFMAE, como ya ha ocurrido también en alguna otra ocasión, han seleccionado el mismo modelo, en esta ocasión un SARIMA(1, 0, 1)(1, 0, 0) $_7$.

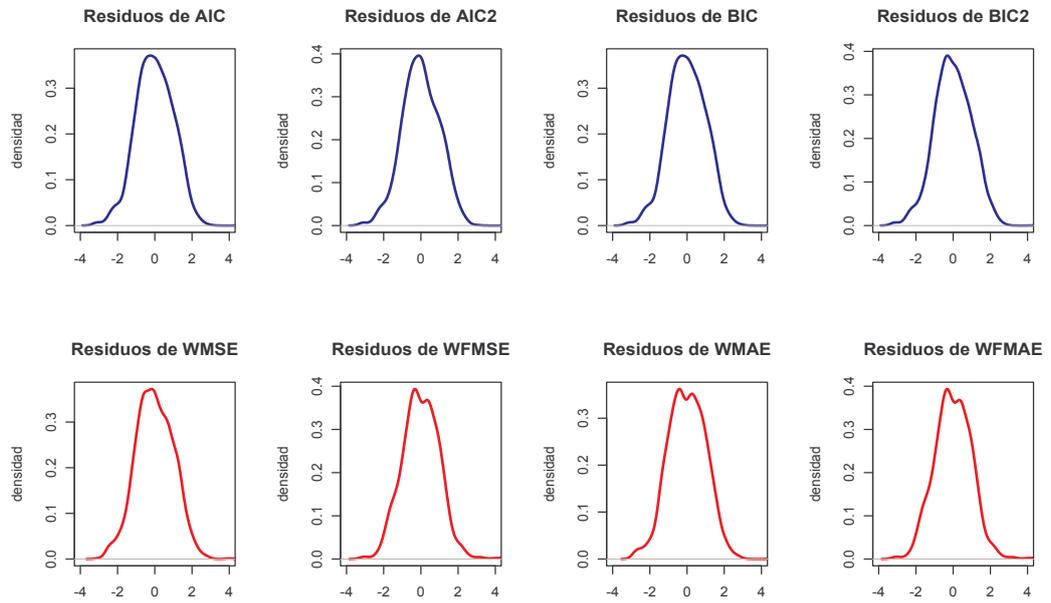


Figura 4.131: Residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 0)(0, 1, 1)₇.

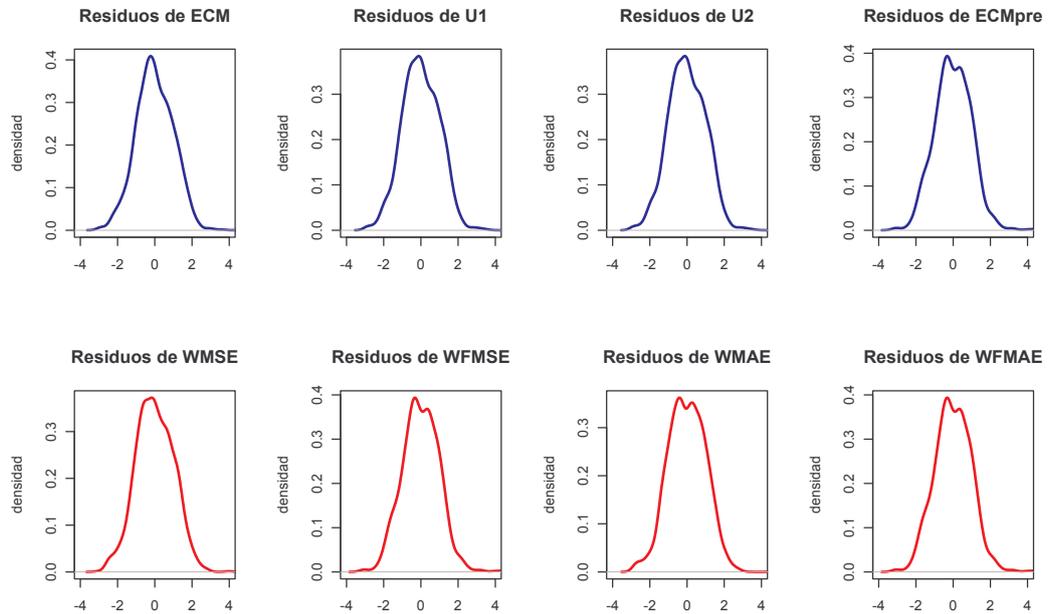


Figura 4.132: Residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 1, 0)(0, 1, 1)₇.

Criterio	Media	Varianza
AIC	-0.1350	16.7193
AIC2	-0.1284	16.6531
ECM	-0.1117	18.5176
BIC	-0.1350	16.7193
BIC2	-0.1069	16.5249
U1	-0.0904	19.0565
U2	-0.0904	19.0565
ECMpre	0.0424	24.1449
WMSE	-0.3263	22.1796
WFMSE	0.0424	24.1449
WMAE	-0.3289	20.1048
WFMAE	0.0424	24.1449

Tabla 4.82: Medidas descriptivas de los residuos en el conjunto de comprobación para SARIMA(1, 1, 0)(0, 1, 1)₇.

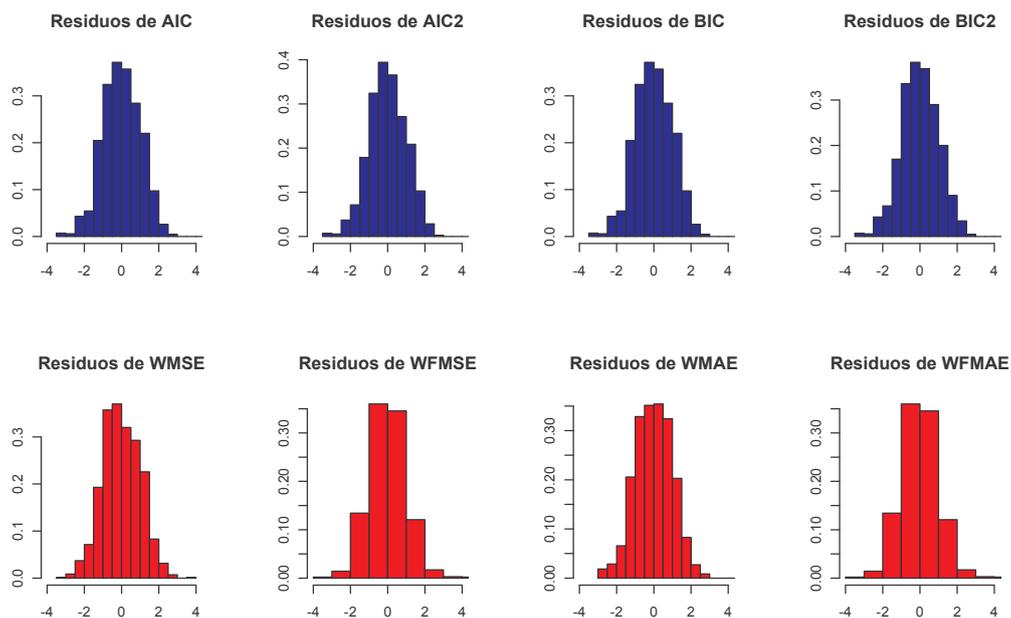


Figura 4.133: Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 0)(0, 1, 1)₇.

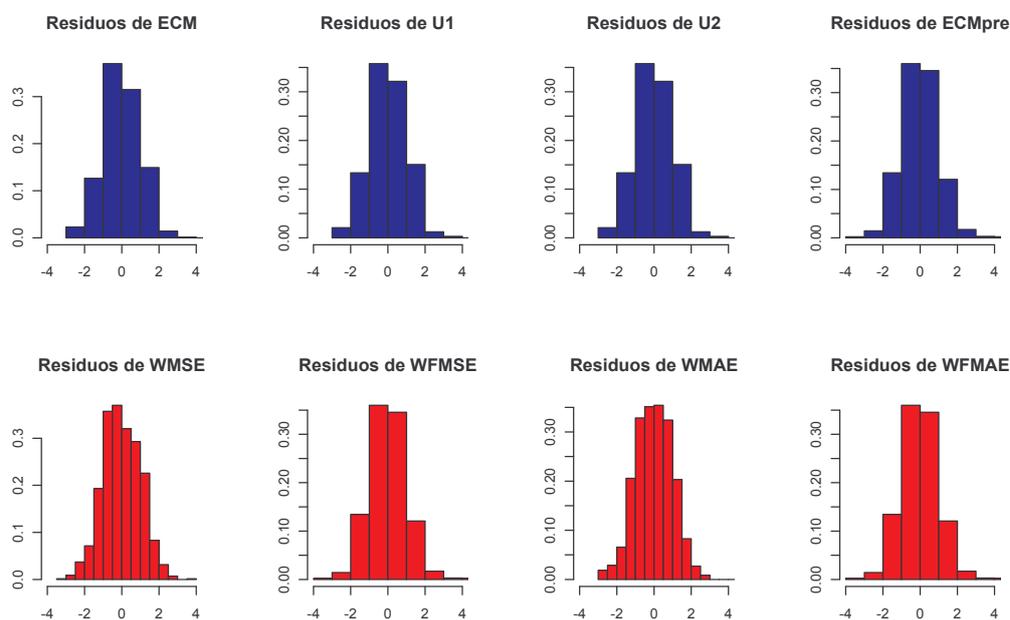


Figura 4.134: Histograma de los residuos obtenidos con los modelos de los criterios de predicción de $\text{SARIMA}(1, 1, 0)(0, 1, 1)_7$.

Criterio	ECMpre1	ECMpre2	ECMpre3	ECMpre4	ECMpre5	ECMpre6	ECMpre7
AIC	1.2973	3.7218	8.0328	13.5712	21.5255	29.7147	39.2994
AIC2	1.3020	3.7079	7.9392	13.6194	21.6108	29.5766	38.9309
ECM	1.5527	4.2678	9.0046	14.9441	23.2271	32.8749	43.8394
BIC	1.2973	3.7218	8.0328	13.5712	21.5255	29.7147	39.2994
BIC2	1.2295	3.6236	7.7989	13.5308	21.5060	29.4007	38.6650
U1	1.5058	4.1554	9.2118	15.3338	23.7623	33.6421	45.8417
U2	1.5058	4.1554	9.2118	15.3338	23.7623	33.6421	45.8417
ECMpre	2.1475	6.3152	12.9828	19.9953	29.2966	42.7398	55.5494
WMSE	1.7528	5.9936	12.4138	19.3926	28.6365	38.6923	49.1212
WFMSE	2.1475	6.3152	12.9828	19.9953	29.2966	42.7398	55.5494
WMAE	1.5066	4.6767	10.3881	17.2923	26.3291	35.5723	45.7253
WFMAE	2.1475	6.3152	12.9828	19.9953	29.2966	42.7398	55.5494

Tabla 4.83: ECMpre global para cada predicción realizada en el conjunto de comprobación en $\text{SARIMA}(1, 1, 0)(0, 1, 1)_7$.

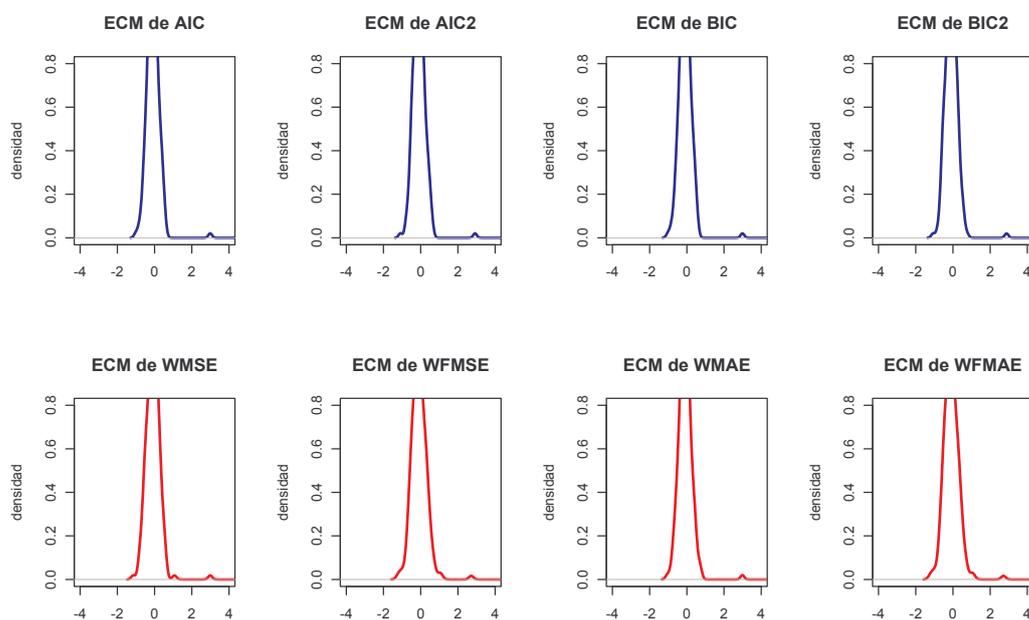


Figura 4.135: ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(0, 1, 1)₇.

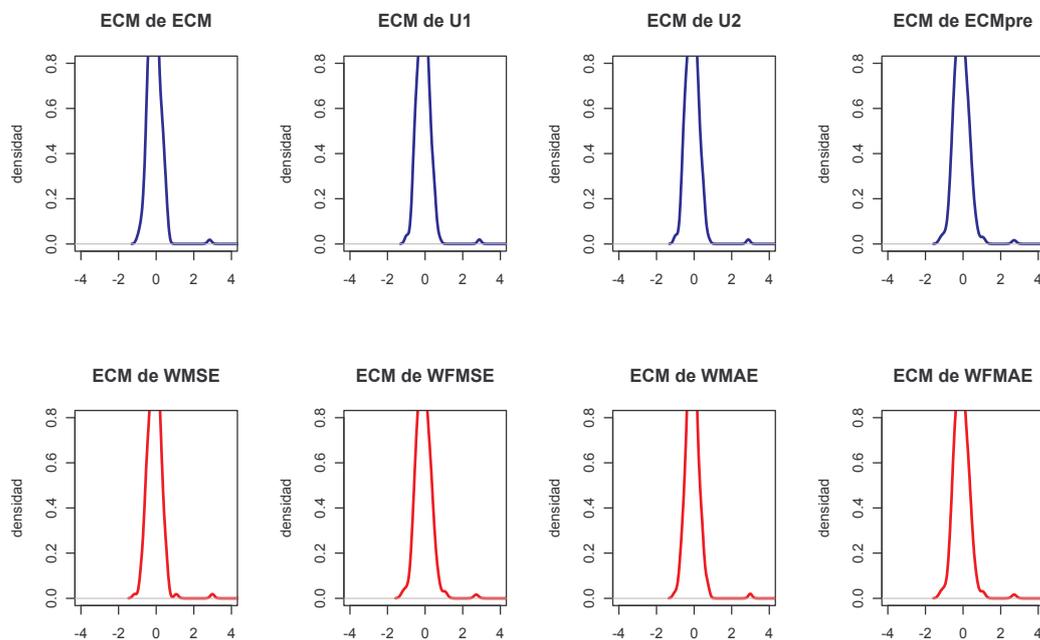


Figura 4.136: ECM de ajuste de los modelos obtenidos con los criterios de predicci3n en SARIMA(1, 1, 0)(0, 1, 1)₇.

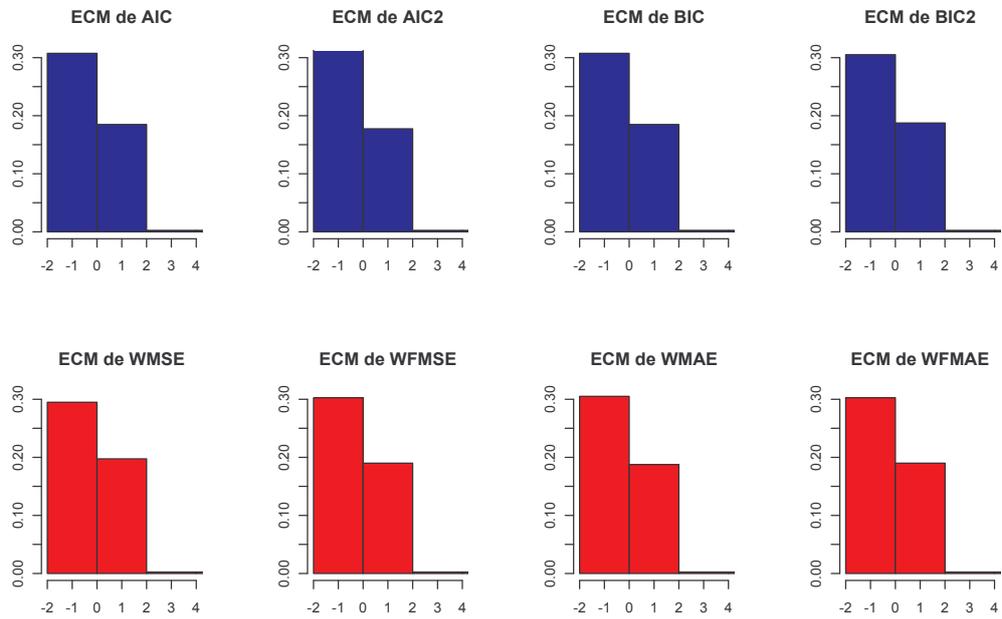


Figura 4.137: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en $SARIMA(1, 1, 0)(0, 1, 1)_7$.

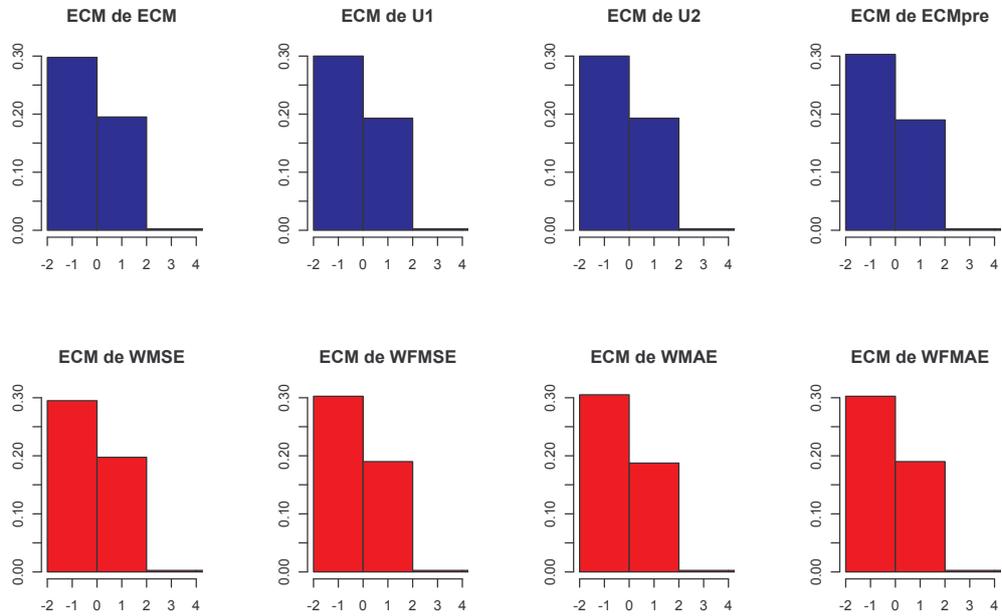


Figura 4.138: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicción en $SARIMA(1, 1, 0)(0, 1, 1)_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.9210	0.9879	0.8621	0.9210	1.0109	0.9141
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.9141	1.3912	1.1492	1.3912	0.9994	1.3912

Tabla 4.84: El ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 0)(0, 1, 1) $_7$.

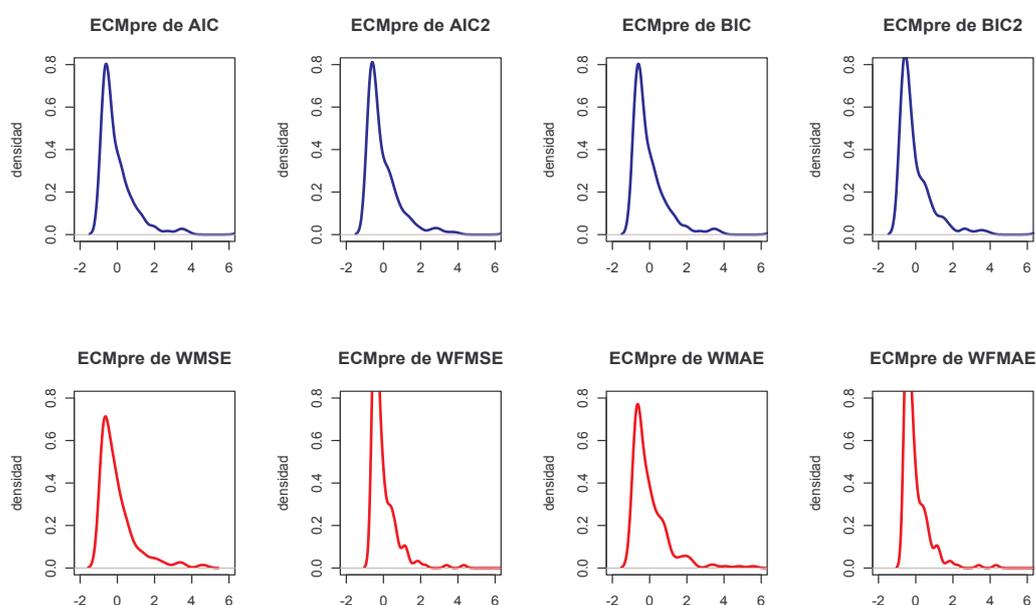


Figura 4.139: ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(0, 1, 1) $_7$.

Existen diferencias mínimas entre los criterios, aunque cabe mencionar que la varianza es peor en los últimos 5 criterios mostrados en la tabla 4.82, también en éstos últimos criterios se alcanzan peores resultados atendiendo a los ECMGpre individual (ver tabla 4.83). En las predicciones realizadas a lo largo del horizonte establecido el aumento en cuanto a los errores cometidos va creciendo en mayor medida cada vez, aunque es un comportamiento que tienen todos los criterios en una forma más acusada en los criterios mencionados. En cuanto al ajuste también en estos criterios es peor aunque en este caso las diferencias con el resto son menos acusadas (tabla 4.84).

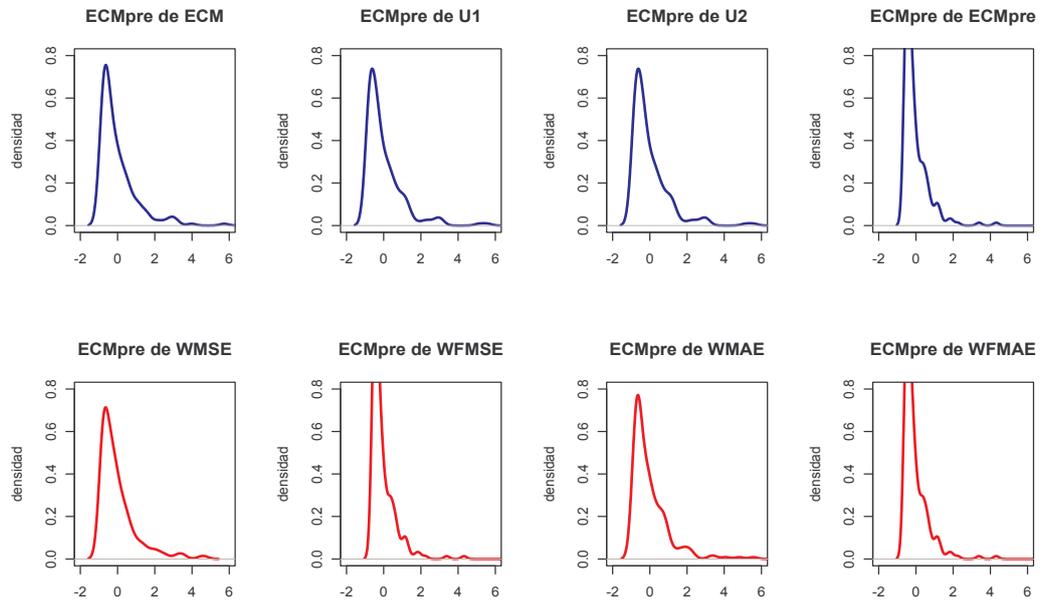


Figura 4.140: ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(0, 1, 1)₇.

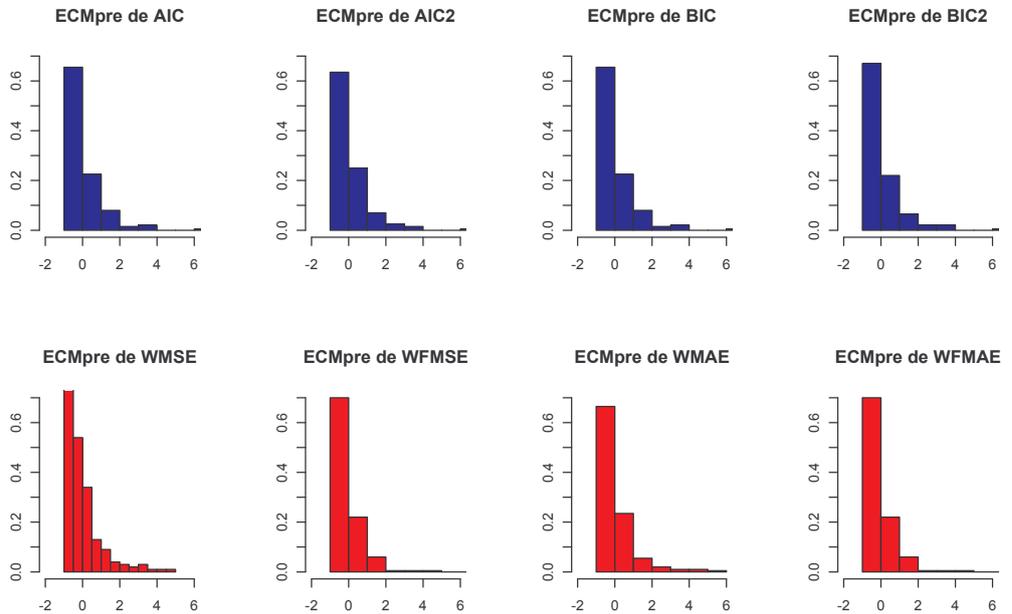


Figura 4.141: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(0, 1, 1)₇.

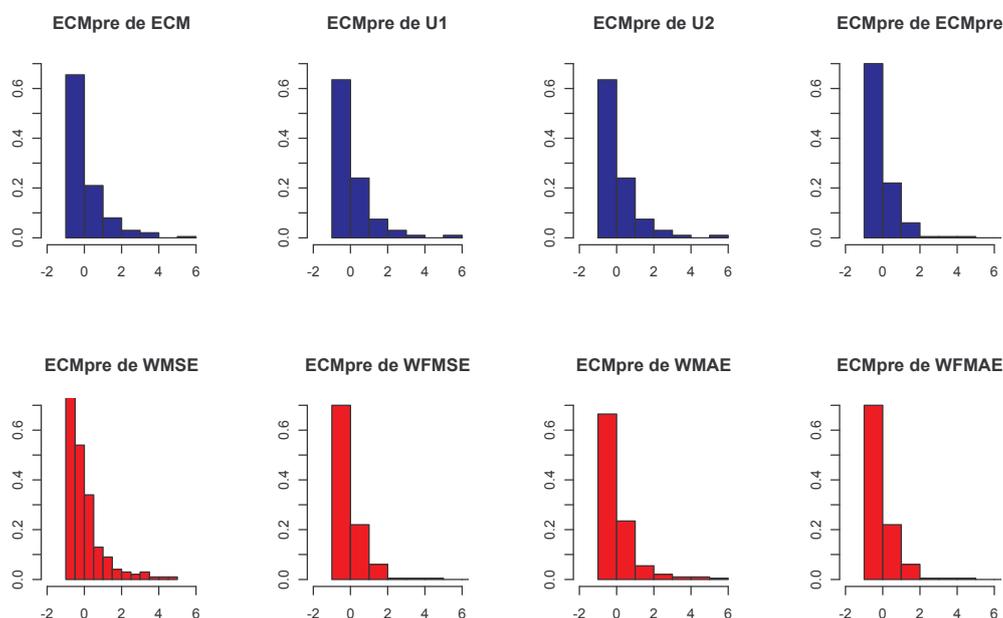


Figura 4.142: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(0, 1, 1) $_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	16.7375	16.6695	18.5301	16.7375	16.5364	19.0647
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	19.0647	24.1467	22.2861	24.1467	20.2129	24.1467

Tabla 4.85: ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 0)(0, 1, 1) $_7$.

En la tabla 4.85 se recoge de nuevo el mal comportamiento de los últimos 5 criterios, donde se puede ver como los resultados globales de predicción son superiores al resto de criterios evaluados.

Continuamos, con el análisis de los series diferenciando entre correctas y no correctas, como ya se ha comentado no todos los criterios son evaluados de esta manera ya que existen algunos de ellos que no son capaces de detectar en ningún caso el modelo real de las series. Los criterios que han sido capaces de detectar el modelo generador de los datos han sido AIC2 (7,5%), BIC (3%), BIC2 (43%), WFMSE (2%) y WFMAE (2%).

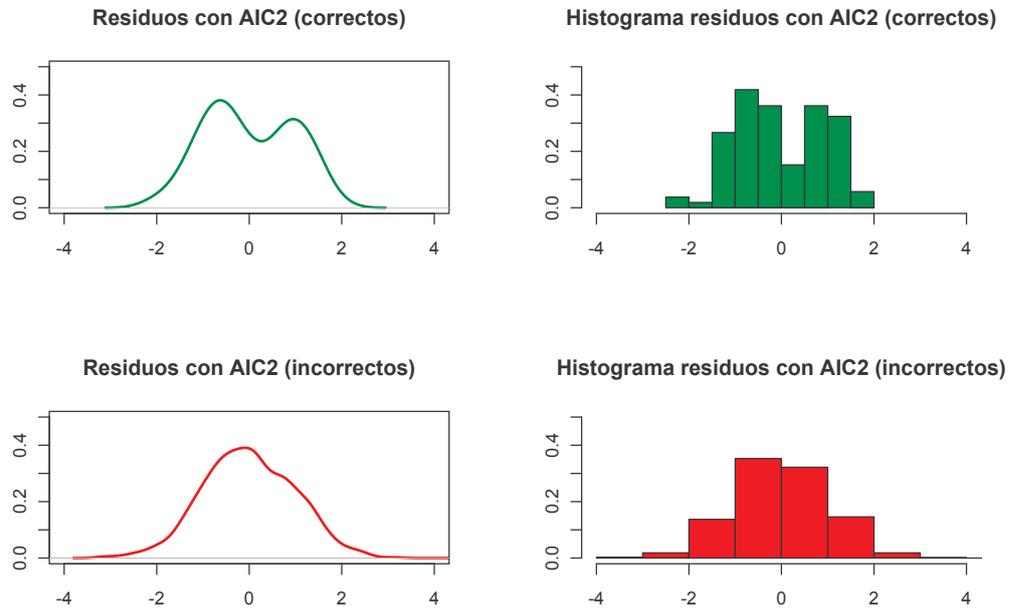


Figura 4.143: Residuos de las series correctas e incorrectas con el criterio AIC2 en el SARIMA(1, 1, 0)(0, 1, 1)₇.

	Media	Varianza	ECMG	ECMGpre
Correctos	0.8209	20.2739	0.9321	20.9478
Incorrectos	-0.2672	17.7095	0.8933	17.7809

Tabla 4.86: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio AIC2.

En el gráfico 4.143 se observa como el comportamiento de los residuos en el conjunto de comprobación de las series incorrectas tiene una mejor distribución que en el conjunto de las series correctas, esto se refleja también en los resultados recogidos en la tabla 4.86, donde la media es inferior, también lo es la varianza, y atendiendo a los resultados globales del ECM y del ECMpre también se obtiene una mejoría con respecto al grupo de las series bien determinadas. Sin embargo, al analizar los resultados individuales (tabla 4.87) la primera estimación es mejor en el conjunto de series correctas, pero en el resto a excepción de la última se comportan mejor en el conjunto de series incorrectas, cabría destacar también en esta ocasión que el aumento de los errores a lo largo del horizonte de predicción se produce de manera mas acusada, aunque de igual manera en un conjunto que en otro.

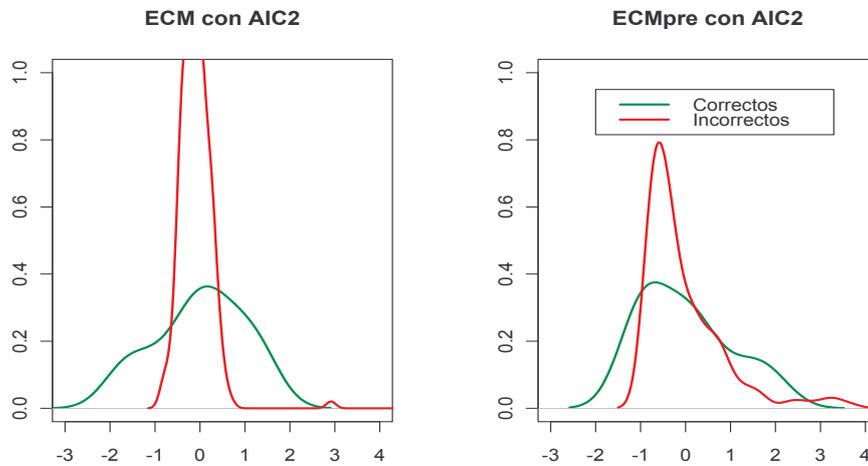


Figura 4.144: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en SARIMA(1, 1, 0)(0, 1, 1) $_7$.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	1.2045	4.5404	9.7374	18.3655	31.9843	39.3747	41.4281
Incorrectos	1.3819	3.8762	8.2907	14.2014	22.0123	31.4073	43.2963

Tabla 4.87: ECM global de predicción en SARIMA(1, 1, 0)(0, 1, 1) $_7$ con el criterio AIC2 en el conjunto de comprobación.

En el criterio BIC se puede comprobar que el comportamiento de los residuos es igual que el descrito en el caso del criterio AIC2, esto se puede ver en el gráfico 1.145 y se puede corroborar con los resultados expuestos en la tabla 4.88.

En cuanto a los resultados de predicción global individual, tabla 4.89, hay algunas diferencias con respecto al anterior criterio ya que en esta ocasión la primera de las predicciones es mejor en las series correctas, pero la segunda también lo es aunque con una diferencia menor, sin embargo en el resto de las estimaciones es en el conjunto de las series incorrectas donde se llegan a cometer errores menores, aumentando éstos en ambos grupos a medida que la predicción se aleja en el tiempo.

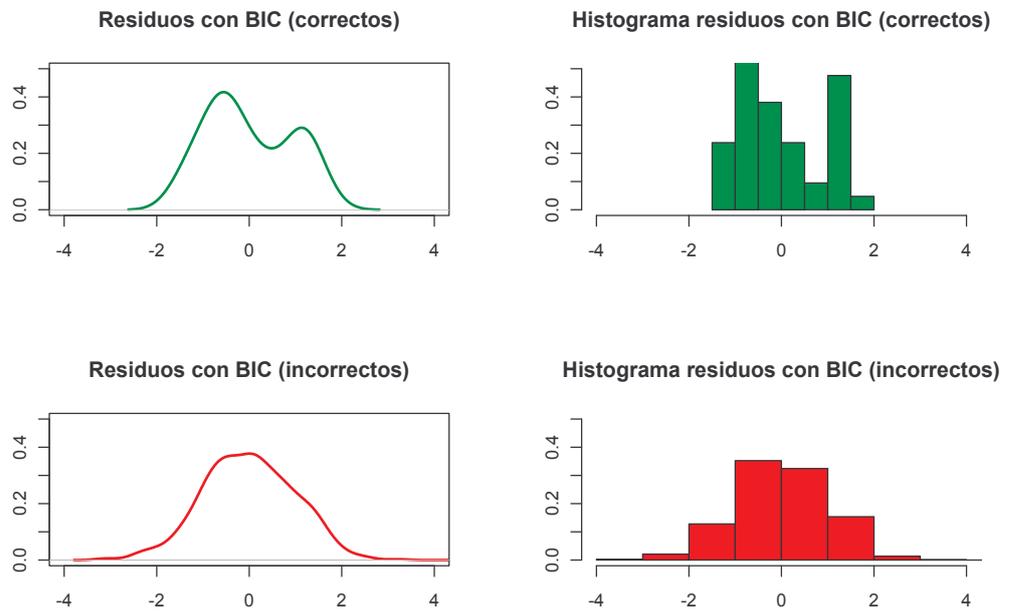


Figura 4.145: Residuos de las series correctas e incorrectas con el criterio BIC en el SARIMA(1, 1, 0)(0, 1, 1)₇.

	Media	Varianza	ECMG	ECMGpre
Correctos	0.2145	29.4149	0.9494	29.4610
Incorrectos	-0.0821	17.5050	0.8909	17.5117

Tabla 4.88: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio BIC.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.8108	3.9279	9.9781	22.3213	40.7690	58.2274	70.1924
Incorrectos	1.4311	4.0409	8.5019	14.1221	21.7307	30.8720	41.8834

Tabla 4.89: ECM global de predicción en SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio BIC en el conjunto de comprobación.

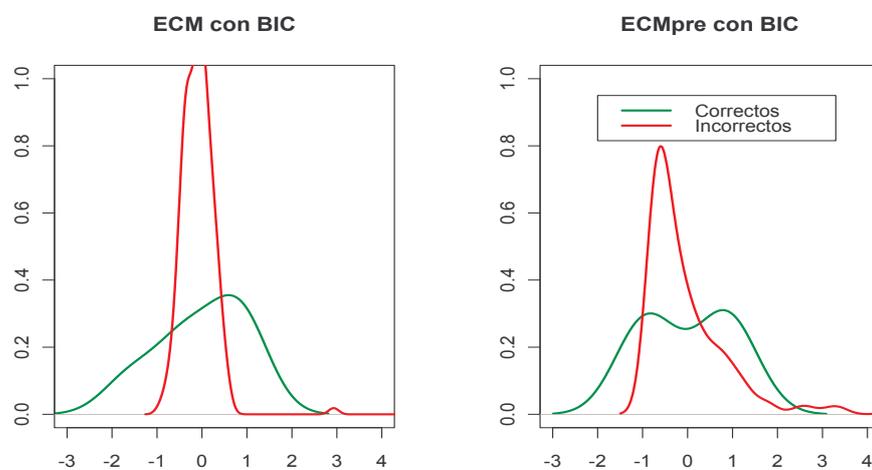


Figura 4.146: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC en SARIMA(1, 1, 0)(0, 1, 1) $_7$.

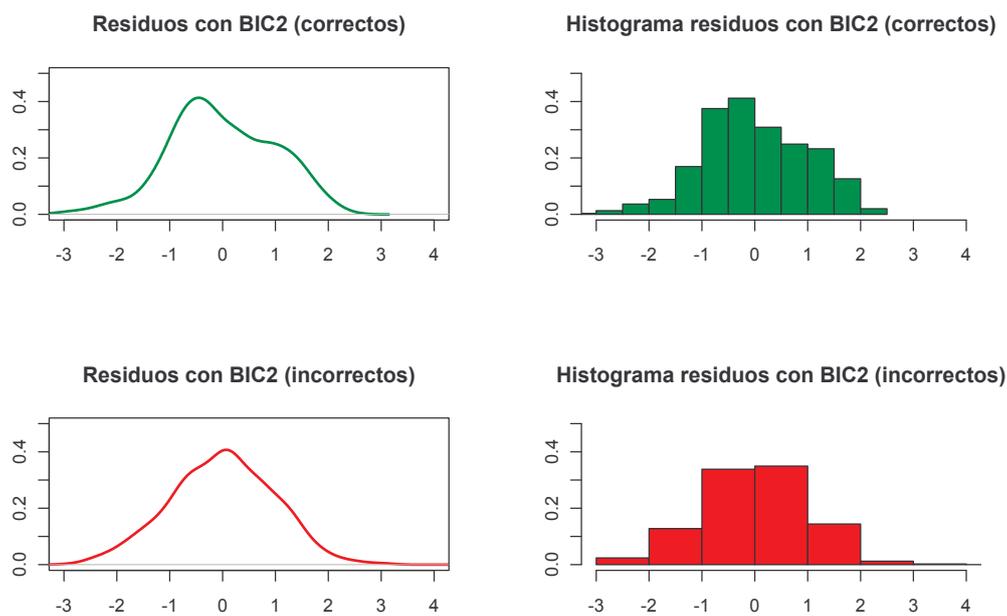


Figura 4.147: Residuos de las series correctas e incorrectas con el criterio BIC2 en el SARIMA(1, 1, 0)(0, 1, 1) $_7$.

	Media	Varianza	ECMG	ECMGpre
Correctos	-0.1484	19.8774	0.9368	19.8994
Incorrectos	0.0265	16.1628	0.9489	16.1635

Tabla 4.90: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio BIC2.

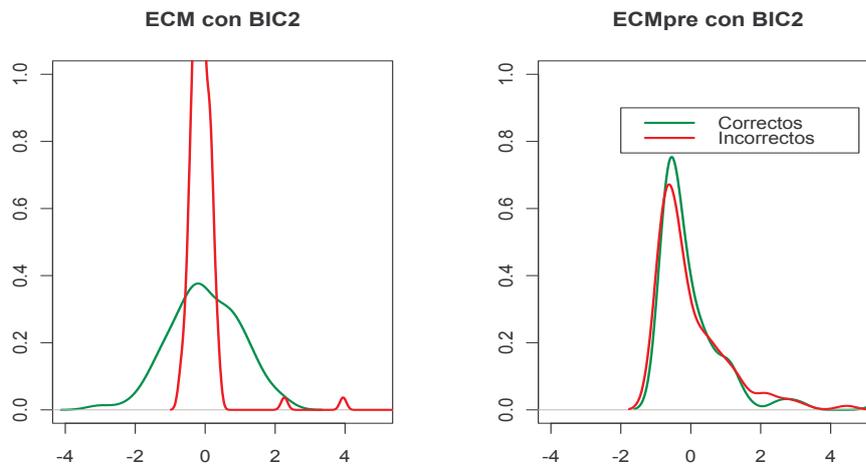


Figura 4.148: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en SARIMA(1, 1, 0)(0, 1, 1)₇.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	1.1757	3.8277	8.6656	16.5814	26.9242	36.0894	46.0316
Incorrectos	1.4621	3.7149	7.5768	12.0932	18.5942	28.4926	41.2105

Tabla 4.91: ECM global de predicción en SARIMA(1, 1, 0)(0, 1, 1)₇ con el criterio BIC2 en el conjunto de comprobación.

El criterio BIC2 que se representa en los gráficos 4.147 y 4.148 tiene un comportamiento prácticamente igual que el que se ha analizado para el criterio BIC, por lo que no hay ninguna característica especial que nombrar en este estudio.

Veamos por último que ocurre con los criterios WFMSE y WFMAE, teniendo en cuenta que los porcentajes de acierto han sido muy bajos.

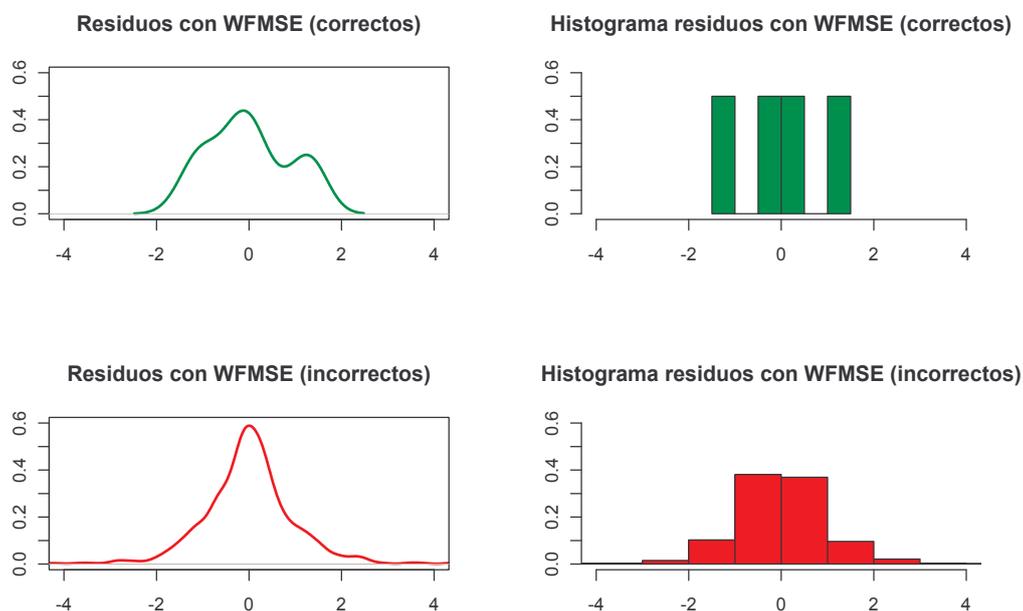


Figura 4.149: Residuos de las series correctas e incorrectas con el criterio WFMSE en el SARIMA(1, 1, 0)(0, 1, 1) $_7$.

	Media	Varianza	ECMG	ECMGpre
Correctos	1.6646	7.64	4.3721	10.4110
Incorrectos	-0.0035	3.1651	4.3120	3.1651

Tabla 4.92: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(0, 1, 1) $_7$ con el criterio WFMSE.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	10.9112	2.9937	6.1459	0.5709	7.9553	7.6191	36.6809
Incorrectos	0.8070	1.7877	2.1494	2.6948	3.4576	4.2876	6.9718

Tabla 4.93: ECM global de predicción en SARIMA(1, 1, 0)(0, 1, 1) $_7$ con el criterio WFMSE en el conjunto de comprobación.

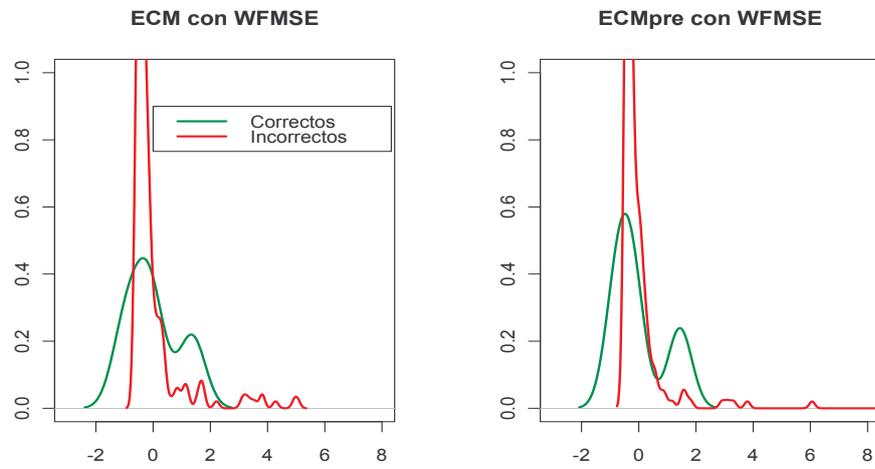


Figura 4.150: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMSE en $\text{SARIMA}(1, 1, 0)(0, 1, 1)_7$.

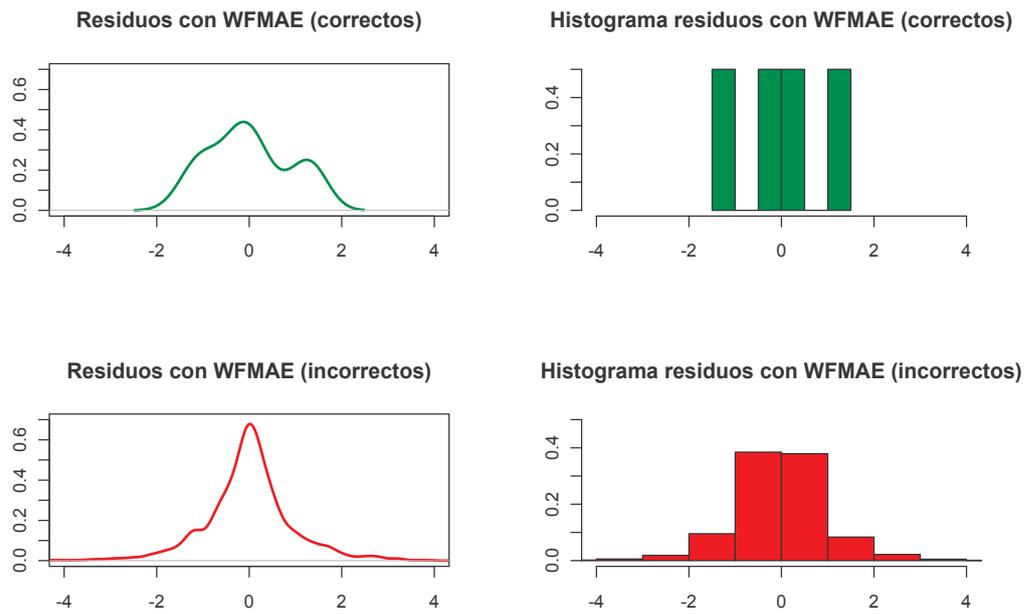


Figura 4.151: Residuos de las series correctas e incorrectas con el criterio WFMAE en el $\text{SARIMA}(1, 1, 0)(0, 1, 1)_7$.

	Media	Varianza	ECMG	ECMGpre
Correctos	1.6646	7.64	4.3721	10.4110
Incorrectos	-0.0239	3.4030	3.9744	3.4036

Tabla 4.94: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(0, 1, 1) $_7$ con el criterio WFMAE.

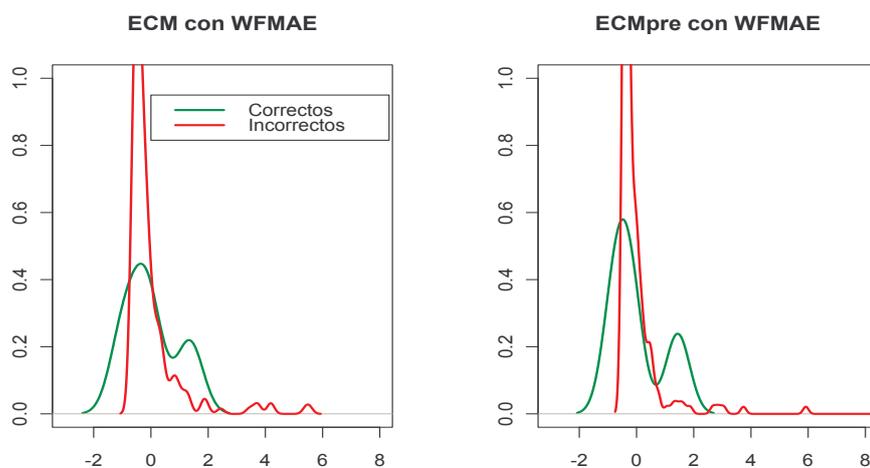


Figura 4.152: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en SARIMA(1, 1, 0)(0, 1, 1) $_7$.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	10.9112	2.9937	6.1459	0.5709	7.9553	7.6191	36.6809
Incorrectos	0.7164	1.6375	1.8259	2.2742	3.2667	5.0041	9.1001

Tabla 4.95: ECM global de predicción en SARIMA(1, 1, 0)(0, 1, 1) $_7$ con el criterio WFMAE en el conjunto de comprobación.

En la tabla 4.92 se aprecia una acusada mejoría en las series incorrectas ante las correctas, gráficamente se ve en gráfica 4.149, esta mejoría se da incluso si se compara con los resultados obtenidos en los criterios analizados anteriormente, también en las predicciones individuales (tabla 4.93), el comportamiento es mejor en las incorrectas y además el aumento en los errores cometidos a lo largo del horizonte de predicción crece de una manera más lenta que en las correctas, este mismo comportamiento se repite con el criterio WFMAE (ver tabla 4.94 y 4.95).

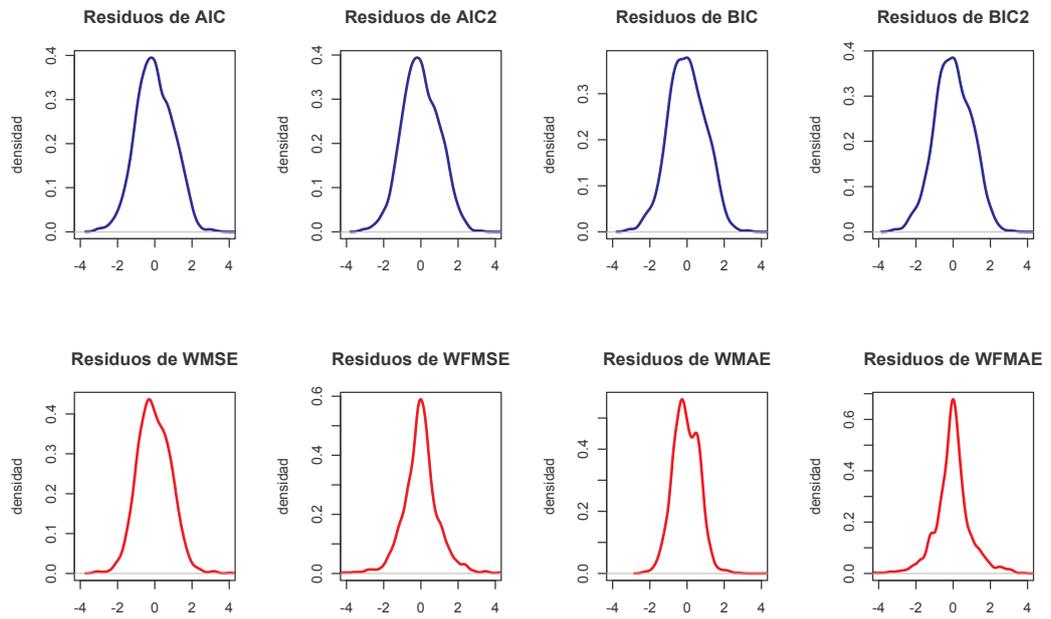


Figura 4.153: Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1)₇.

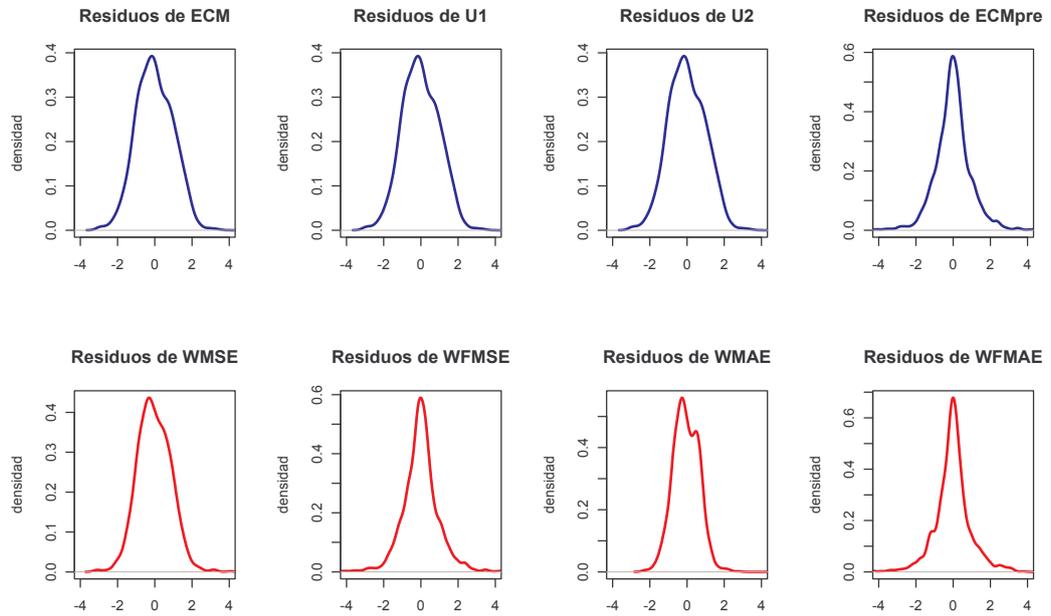


Figura 4.154: Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1)₇.

Criterio	Media	Varianza
AIC	-0.2098	17.8996
AIC2	-0.1855	17.9840
ECM	-0.1882	18.7299
BIC	-0.0732	17.8648
BIC2	-0.0487	17.7675
U1	-0.1882	18.7299
U2	-0.1882	18.7299
ECM _{pre}	0.0313	3.3090
WMSE	0.2742	26.6889
WFMSE	0.0298	3.3092
WMAE	0.3112	51.4154
WFMAE	0.0099	3.5436

Tabla 4.96: Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para SARIMA(1, 1, 0)(0, 1, 1) $_7$.

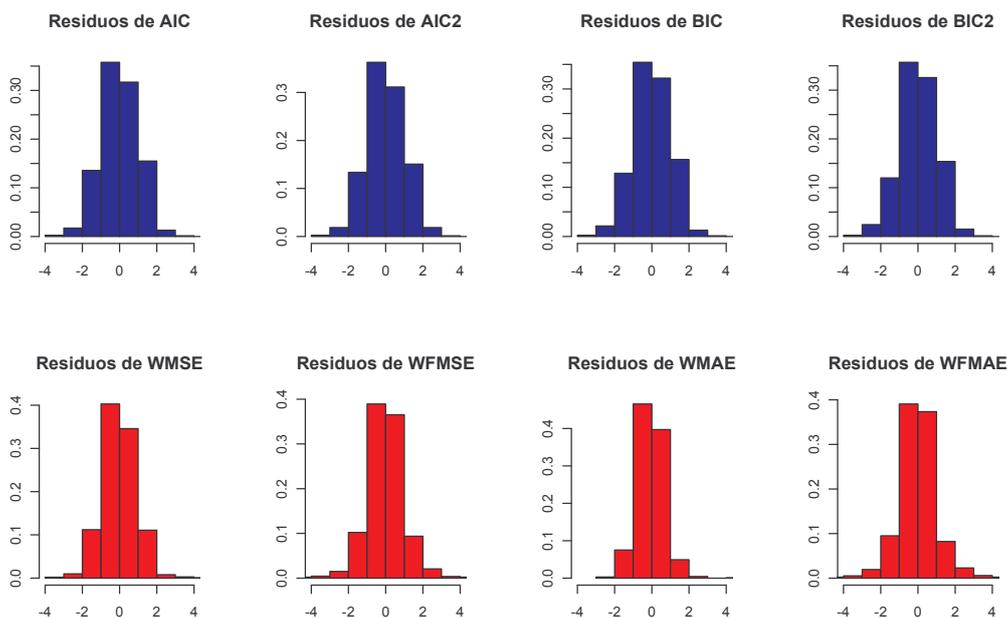


Figura 4.155: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) $_7$.

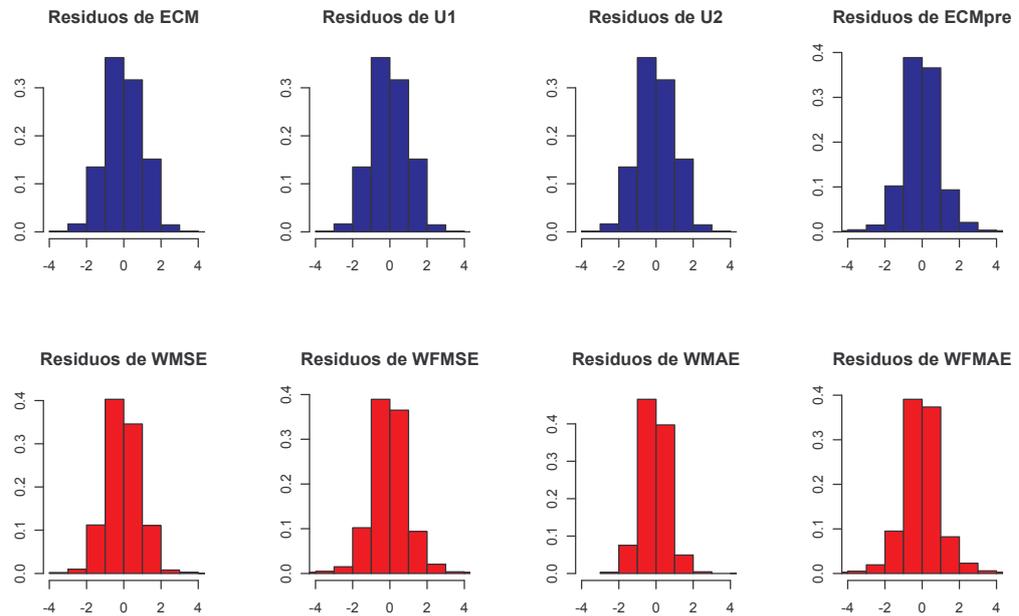


Figura 4.156: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1)₇.

Criterio	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
AIC	1.4761	4.0640	8.7250	14.6755	22.6707	31.6804	42.3137
AIC2	1.3686	3.9260	8.3992	14.5137	22.7602	32.0049	43.1562
ECM	1.5317	4.2107	9.0141	15.0821	23.4273	33.2255	44.8657
BIC	1.4124	4.0375	8.5462	14.3680	22.3019	31.6926	42.7327
BIC2	1.3390	3.7634	8.0450	14.0232	22.1761	31.7592	43.2836
U1	1.5317	4.2107	9.0141	15.0821	23.4273	33.2255	44.8657
U2	1.5317	4.2107	9.0141	15.0821	23.4273	33.2255	44.8657
ECMpre	1.0093	1.8132	2.2337	2.6549	3.5475	4.3512	7.5604
WMSE	2.0103	5.7044	12.5563	20.7007	32.5525	48.7108	65.1136
WFMSE	1.0091	1.8119	2.2293	2.6524	3.5476	4.3542	7.5660
WMAE	2.7141	9.8020	22.4624	40.4550	64.6540	94.3537	126.1414
WFMAE	0.9203	1.6646	1.9123	2.2402	3.3605	5.0564	9.6517

Tabla 4.97: ECM globales para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 0)(0, 1, 1)₇.

En los gráficos 4.153 y 4.154 podemos comprobar como las curvas de los residuos obtenidos en los criterios ECMpre, WFMSE y WFMAE tienen una menor varianza con respecto al resto, algo que se verifica en la tabla 4.96 donde además se muestra un menor valor en cuanto a la media de los residuos se refiere, también en la tabla 4.97 son estos tres criterios los que alcanzan los mejores resultados en el ECMGpre individual, siendo en esta ocasión el criterio WFMAE el que mejor funciona y además en los tres criterios el aumento del error a lo largo de la predicción tiene un crecimiento inferior que en el resto de criterios. Estos resultados a su vez mejoran a los obtenidos en los mismos criterios en las tablas 4.82 y 4.83.

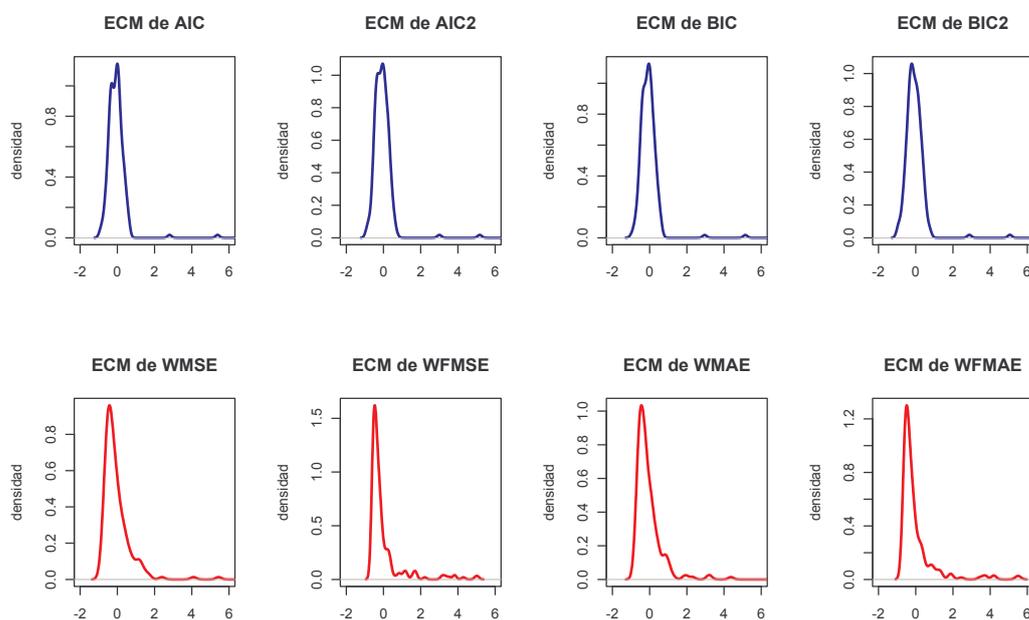


Figura 4.157: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1)₇.

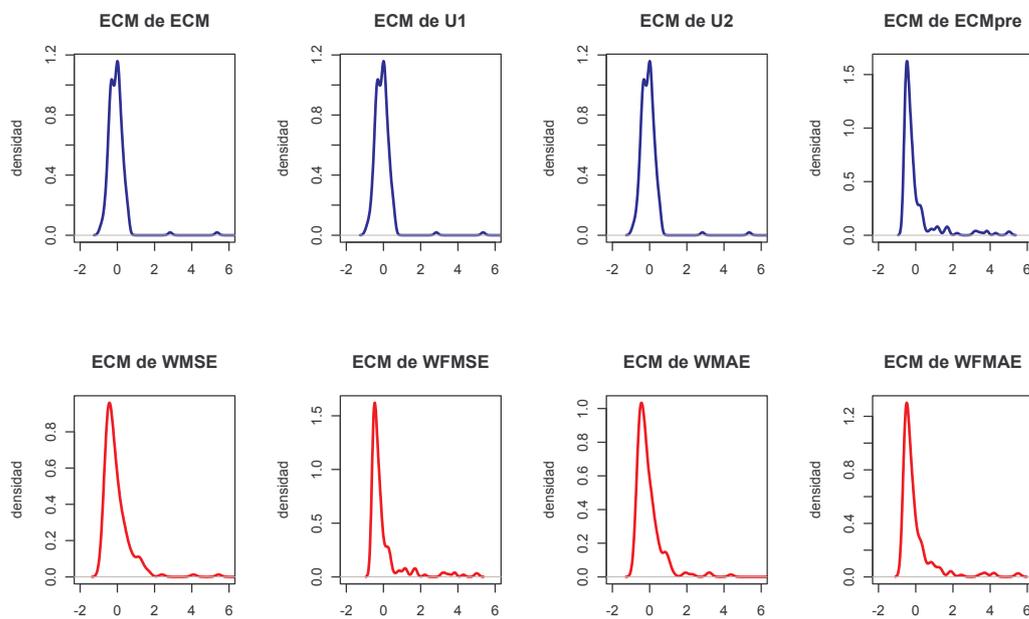


Figura 4.158: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1)₇.

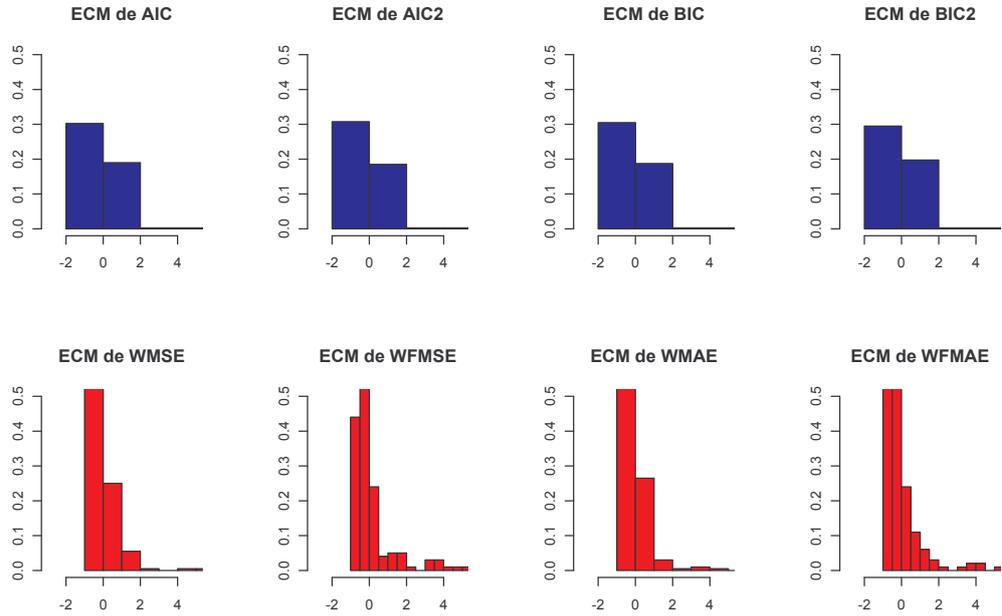


Figura 4.159: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 0)(0, 1, 1)_7$.

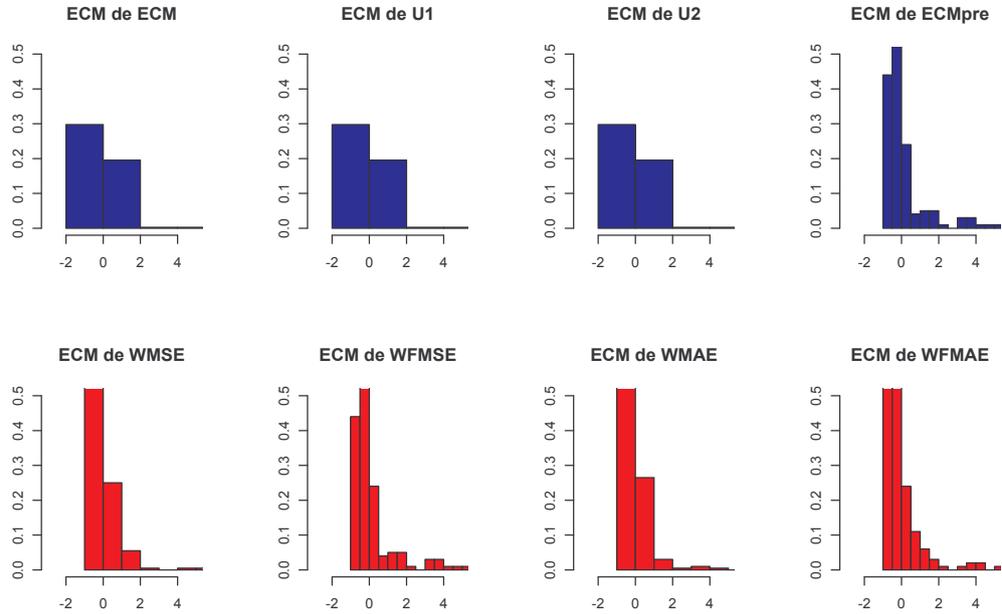


Figura 4.160: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 0)(0, 1, 1)_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8642	0.8962	0.8546	0.8927	0.9437	0.8546
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8546	4.3137	1.1694	4.3132	1.2913	3.9824

Tabla 4.98: ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 0)(0, 1, 1)₇.

En los gráficos 4.159 y 4.160 se aprecia de manera visual que los ECM de los criterios ECMpre, WFMSE y WFMAE tienen una mayor dispersión con respecto al resto, y rectángulos muy alejados del grueso de la distribución, por lo en cuanto a ajuste será en estos tres criterios donde las cosas funcionen peor, como puede verse en la tabla 4.98 hay una gran diferencia en los valores obtenidos por éstos. Ya viene siendo habitual este hecho, de manera que si se mejora la predicción en contrapartida se empeora el ajuste, en nuestro trabajo es un tema en el que no vamos a profundizar, dado que nuestro objetivo es conseguir la mejor predicción posible independientemente del ajuste que se alcance, ya que es sabido que ambas cosas no tienen porque darse en un modelo, existen en muchos estudios donde se ha demostrado que una buena capacidad predictiva no implica un buen ajuste y viceversa, nuestro estudio podría ser un buen ejemplo de ello.

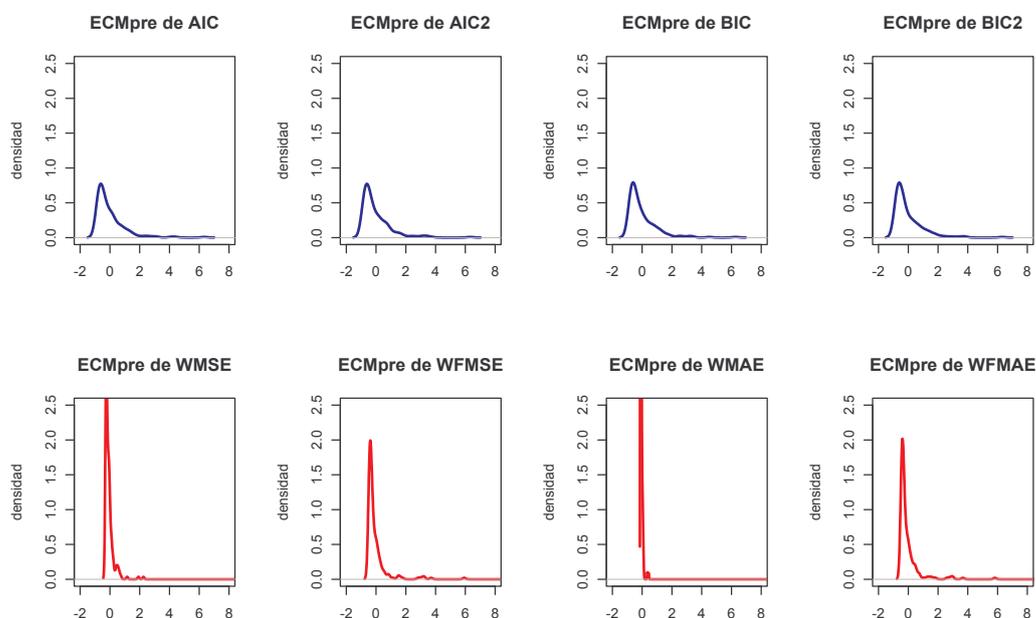


Figura 4.161: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1)₇.

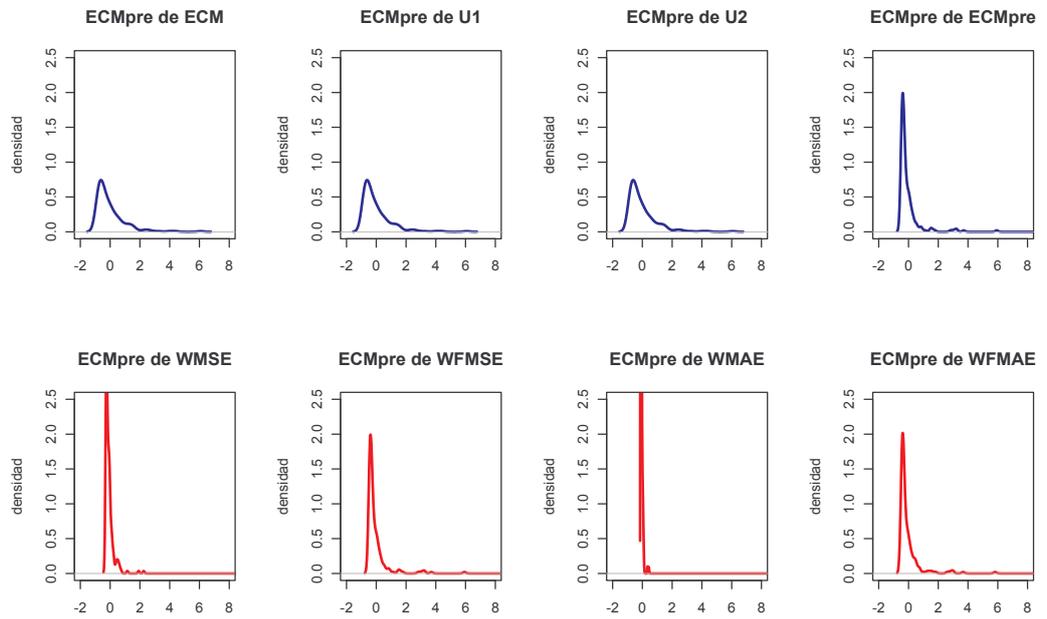


Figura 4.162: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 0)(0, 1, 1)_7$.

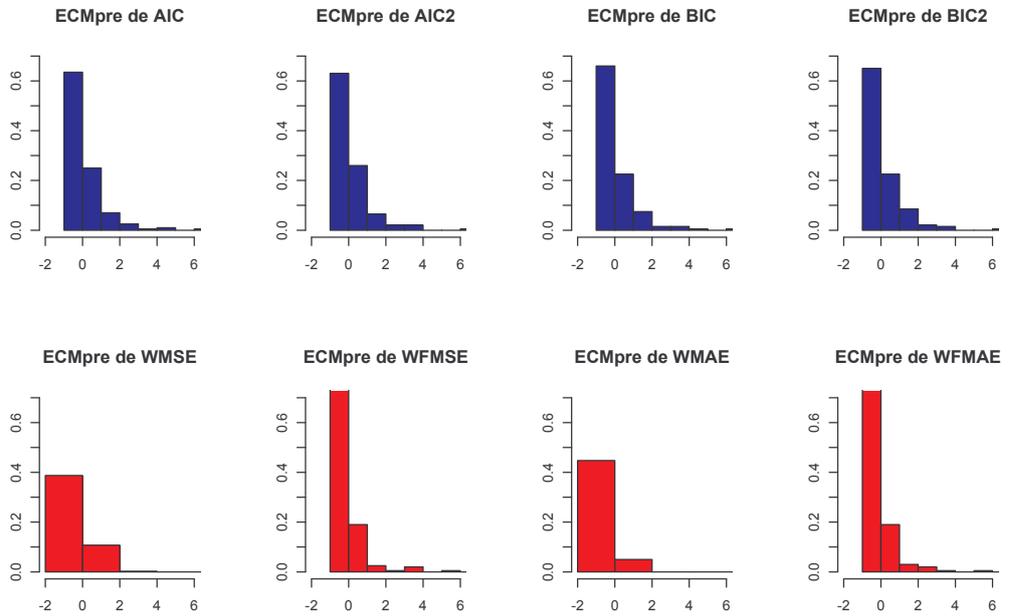


Figura 4.163: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 0)(0, 1, 1)_7$.

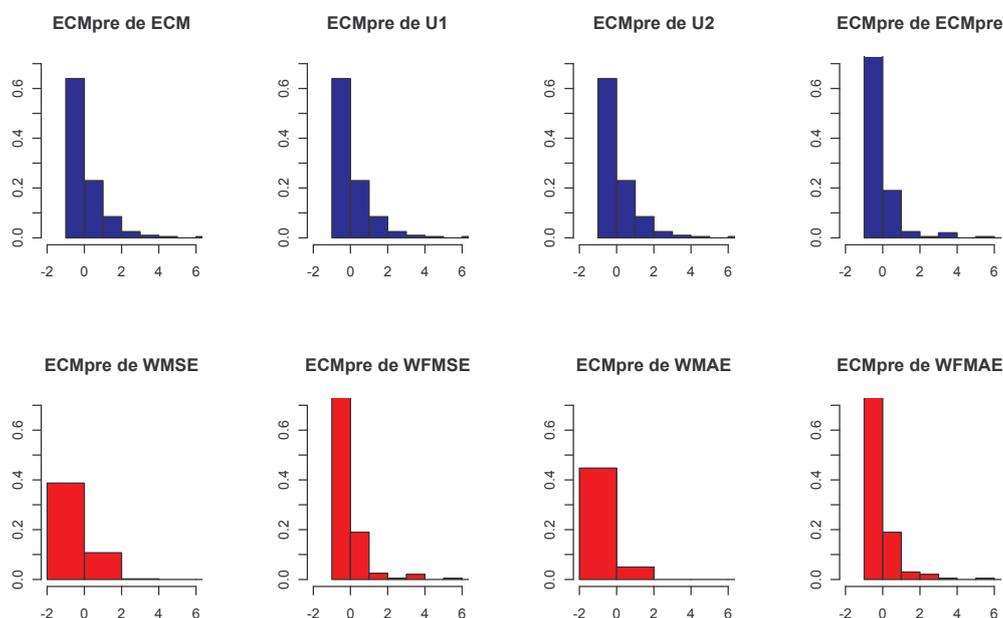


Figura 4.164: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(0, 1, 1) $_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	17.9436	18.0184	18.7653	17.8702	17.7699	18.7653
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	18.7653	3.31	26.7641	3.3101	51.5123	3.5437

Tabla 4.99: ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 0)(0, 1, 1) $_7$.

En la tabla 4.99, se ve la gran disminución sufrida en el ECMGpre que se logra con los criterios ECMpre, WFMSE y WFMAE, siendo esta disminución también importante si se compara con los resultados obtenidos en la tabla 4.85. Dentro de estos tres criterios el preferible es WFMAE, ya que sólo consigue unos resultados globales mejores, aunque sea algo superior que en los otros dos, sino que consigue unas predicciones más precisas en los primeros momentos, lo que nos lleva a un mayor porcentaje de acierto en las estimaciones inmediatamente siguientes al último instante del conjunto de entrenamiento, de manera que para asegurar estas últimas se podrán actualizar las predicciones más tarde.

Modelo SARIMA(1, 1, 0)(1, 1, 0)₇, cuyos coeficientes son $\phi_1 = 0,7$ y $\Phi_1 = 0,6$. (Modelo 172)

	AIC	AIC2	ECM	BIC	BIC2
Valor	2.333354	-0.3701187	0.561796	2.422665	-0.2986159
Modelo	304	297	324	178	172

Tabla 4.100: Criterios Clásicos de ajuste para SARIMA(1, 1, 0)(1, 1, 0)₇

	U_1	U_2	ECMpre
Valor	4.894569e-4	4.89457e-4	0.049963
Modelo	306	306	257

Tabla 4.101: Criterios Clásicos de predicción para SARIMA(1, 1, 0)(1, 1, 0)₇

	Ajuste		Predicción	
	WMSE	WMAE	WFMSE	WFMAE
Valor	0.1879612	0.3096225	0.03770242	0.1377073
Modelo	234	167	257	268

Tabla 4.102: Criterios propuestos para SARIMA(1, 1, 0)(1, 1, 0)₇

Analizando las tablas 4.100 a la 4.102, tenemos que el criterio BIC2 selecciona el modelo generador de las series, los criterios de Theil coinciden una vez más en la elección de su modelo, optando en este caso por un SARIMA(2, 1, 1)(2, 1, 2)₇, y como se puede ver tienden a seleccionar modelos de mayor complejidad que el modelo real, aunque este es un tema que tampoco abordamos en este trabajo, ya que, al estar automatizado el procedimiento la complejidad sería tan sólo a nivel computacional, y se ha comprobado que el tiempo de ejecución entre modelos más simples a modelos más complejos no difiere significativamente para que esto sea un problema. Por último, el criterio ECMpre coincide con el criterio de predicción propuesto en el estudio, WFMSE.

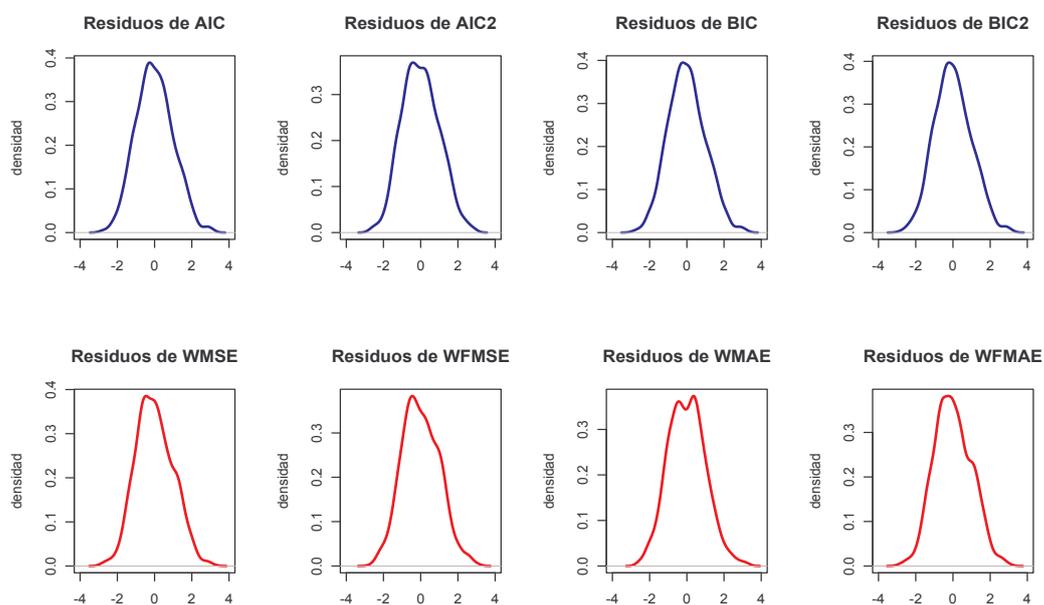


Figura 4.165: Residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 0)(1, 1, 0)₇.

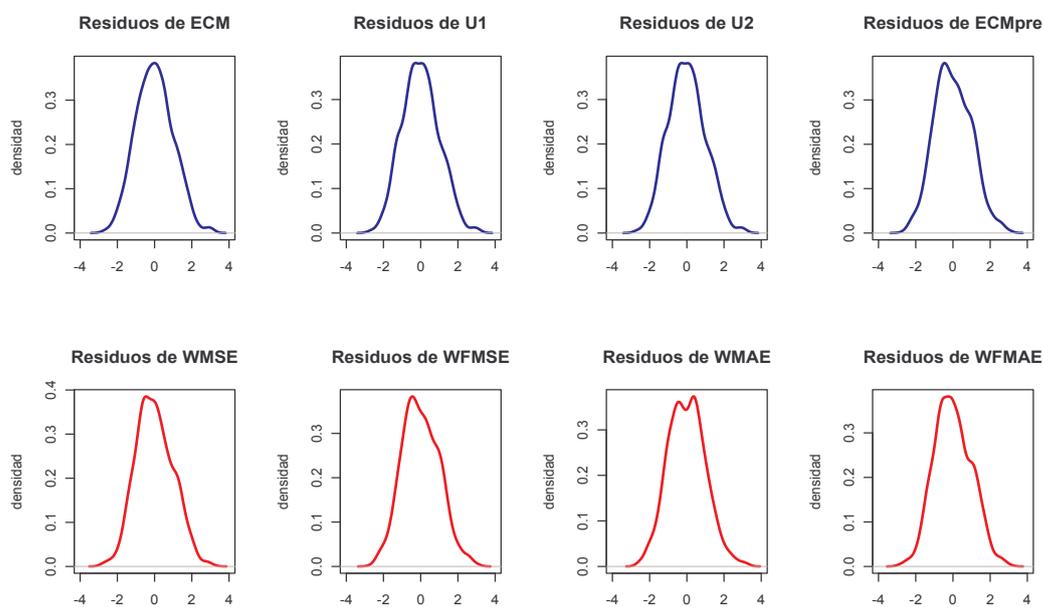


Figura 4.166: Residuos obtenidos con los modelos de los criterios de predicci3n de SARIMA(1, 1, 0)(1, 1, 0)₇.

Criterio	Media	Varianza
AIC	0.2846	15.7827
AIC2	0.2015	17.0990
ECM	0.3448	16.0791
BIC	0.32	15.5894
BIC2	0.3085	15.5159
U1	0.3566	16.1758
U2	0.3566	16.1758
ECMpre	0.2199	21.1807
WMSE	0.3020	18.1447
WFMSE	0.2199	21.1807
WMAE	0.2906	19.2335
WFMAE	0.2212	18.2409

Tabla 4.103: Medidas descriptivas de los residuos en el conjunto de comprobación para SARIMA(1, 1, 0)(1, 1, 0)₇.

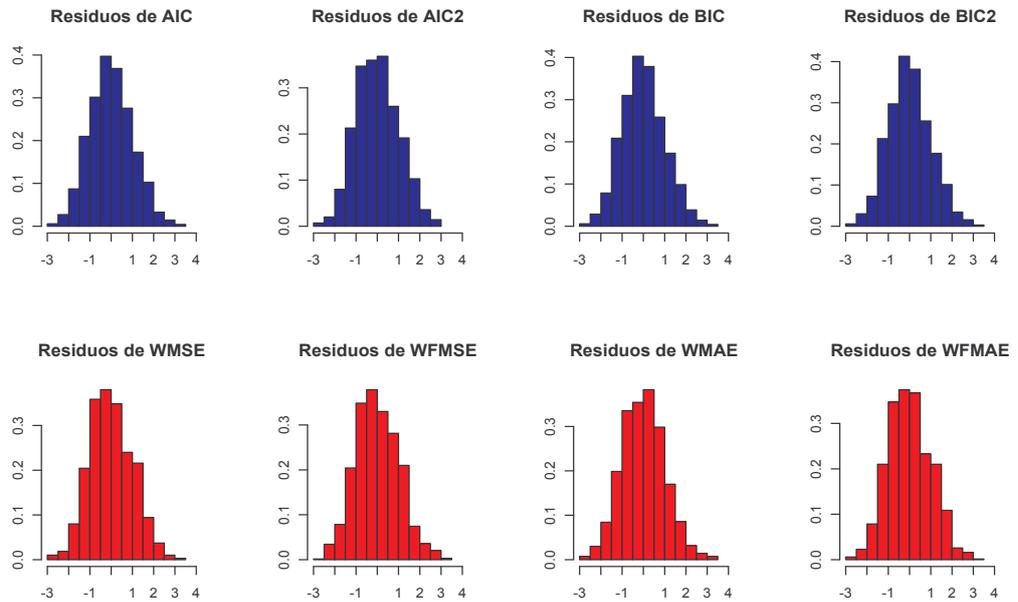


Figura 4.167: Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 0)(1, 1, 0)₇.

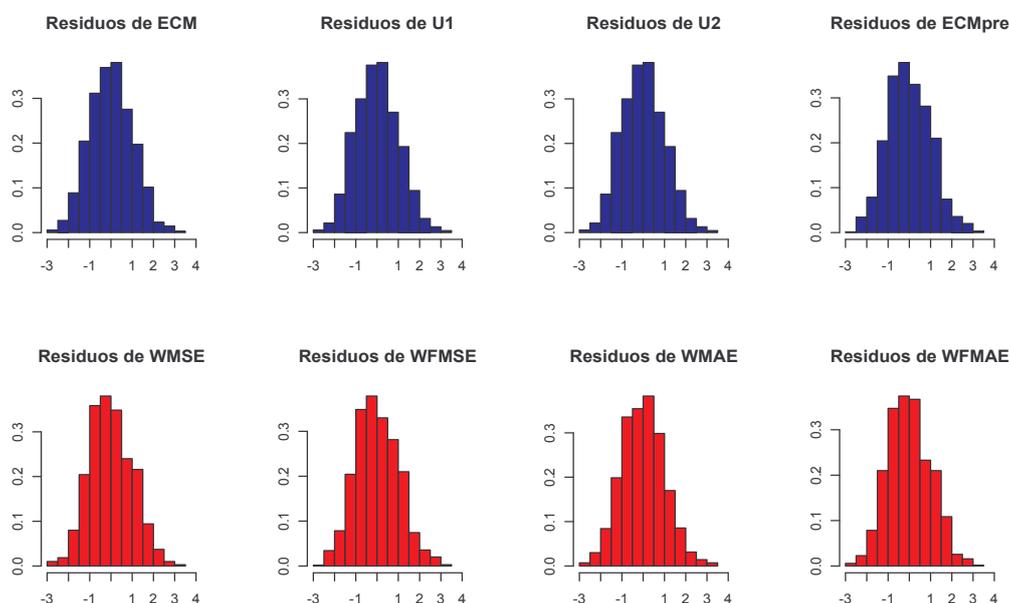


Figura 4.168: Histograma de los residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 1, 0)(1, 1, 0) $_7$.

Criterio	ECMpre1	ECMpre2	ECMpre3	ECMpre4	ECMpre5	ECMpre6	ECMpre7
AIC	1.0366	3.7035	7.8934	13.5385	20.4477	27.7751	36.6509
AIC2	1.0831	3.8866	8.4995	14.5230	21.9322	30.1398	39.9128
ECM	1.0507	3.8137	7.9722	13.6883	20.8197	28.5081	37.5329
BIC	1.0174	3.6378	7.7891	13.4648	20.2642	27.4731	36.1958
BIC2	1.0134	3.65	7.7966	13.3942	20.0998	27.4731	36.1958
U1	1.0651	3.8252	8.0618	13.8130	20.9154	28.6368	37.8036
U2	1.0651	3.8252	8.0618	13.8130	20.9154	28.6368	37.8036
ECMpre	1.1931	4.5037	10.0370	17.7646	27.6786	37.7179	49.7086
WMSE	1.0466	3.8678	8.3761	15.0931	23.4347	32.3710	43.4621
WFMSE	1.1931	4.5037	10.0370	17.7646	27.6786	37.7179	49.7086
WMAE	1.1652	4.1975	9.3271	16.3255	25.3185	34.2891	44.6026
WFMAE	1.1027	4.0618	8.7404	15.4499	23.6832	32.1933	42.7980

Tabla 4.104: ECMpre global para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 0)(1, 1, 0) $_7$.

Ante la tabla 4.103 se tienen que las menores varianzas se obtienen en los criterios AIC, BIC y BIC2 (ver gráfico 4.165), en cuanto a las medias no existen diferencias importantes entre los criterios expuestos en la tabla. En la tabla 4.104 donde se muestran los ECMGpre individuales no hay nada que destacar entre los criterios pues el comportamiento entre ellos es bastante similar.

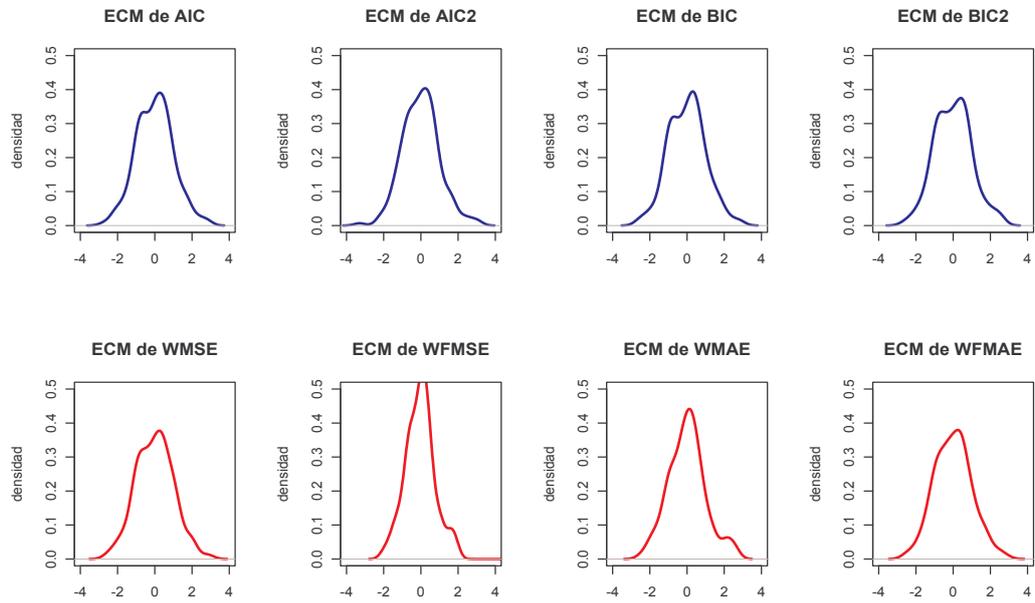


Figura 4.169: ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 1, 0)₇.

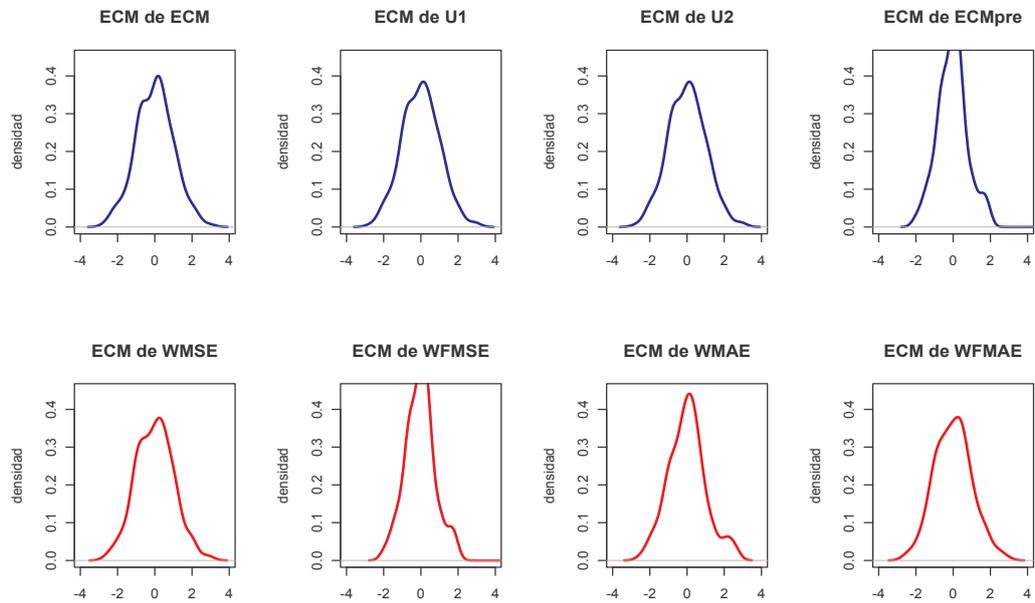


Figura 4.170: ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(1, 1, 0)₇.

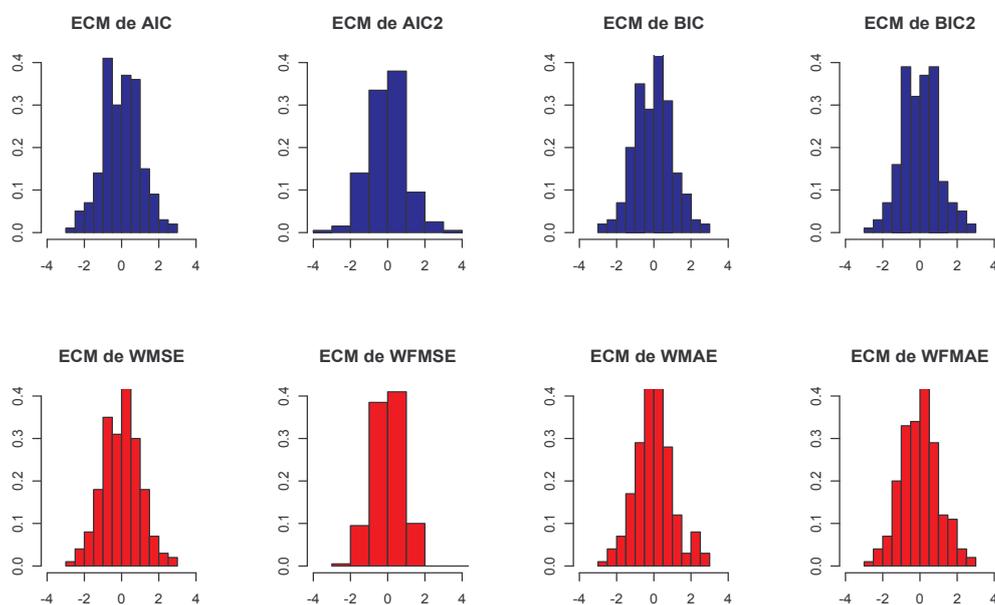


Figura 4.171: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 1, 0) $_7$.

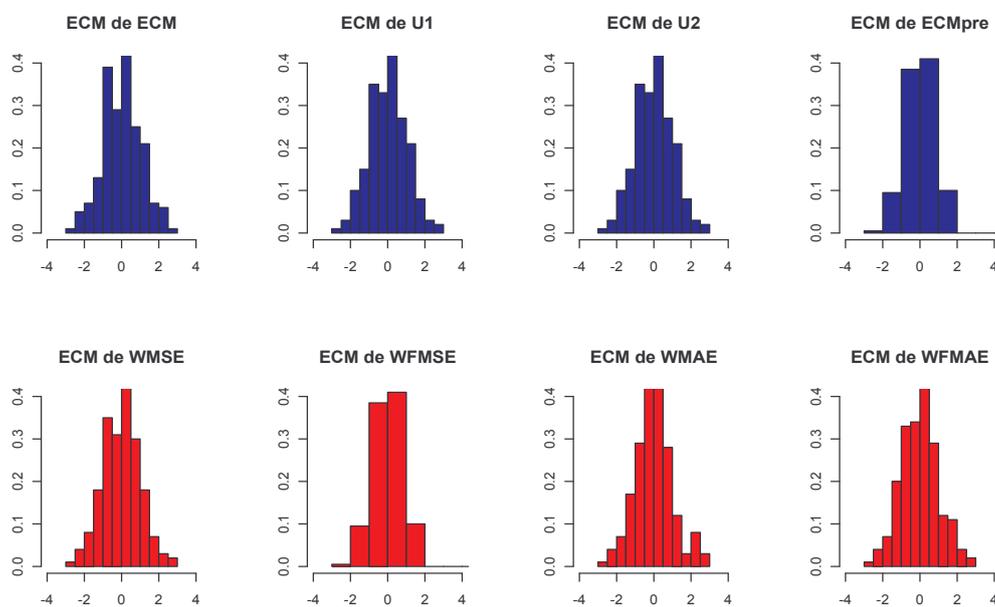


Figura 4.172: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicci3n en SARIMA(1, 1, 0)(1, 1, 0) $_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8244	0.9364	0.8036	0.8479	0.8930	0.8069
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8069	1.0394	0.8267	1.0394	1.0549	0.83

Tabla 4.105: El ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 0)(1, 1, 0)₇.

En los gráficos 4.169 al 4.172 se representan los ECM de cada serie obtenido con cada uno de los criterios y el comportamiento es bastante homogéneo entre criterios a excepción del ECMpre y WFMSE donde se ve claramente una distribución peor, estos datos quedan reflejados en la tabla 4.105, donde destacan por buen ajuste los criterio de predicción de Theil, U_1 , U_2 junto con el ECM.

Continuamos con el estudio de los ECM de predicción que se muestran en los gráficos 4.173 al 4.176 donde se intuye que los resultados no son los deseados.

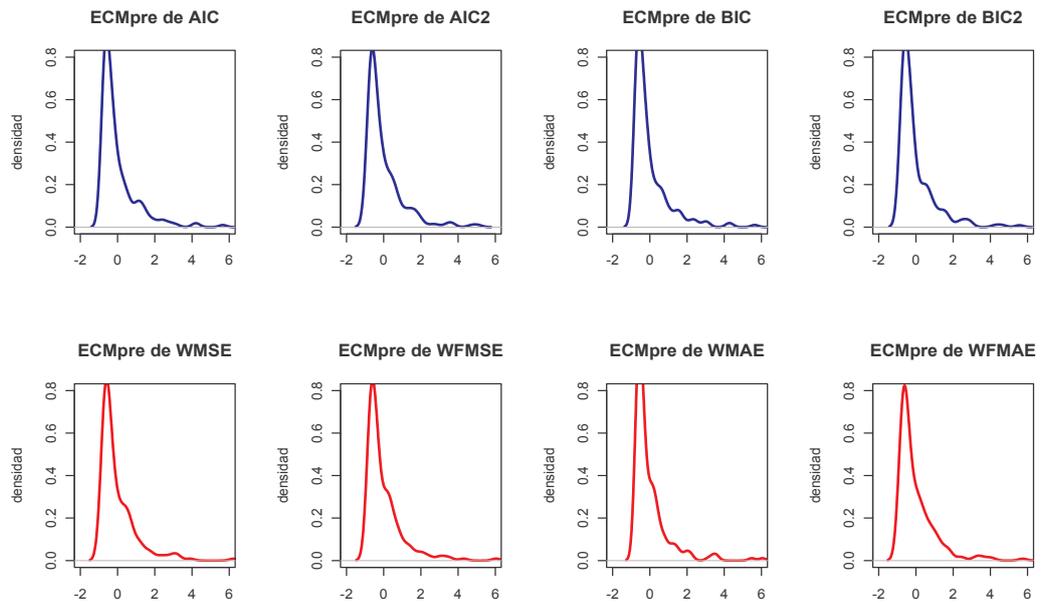


Figura 4.173: ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 1, 0)₇.

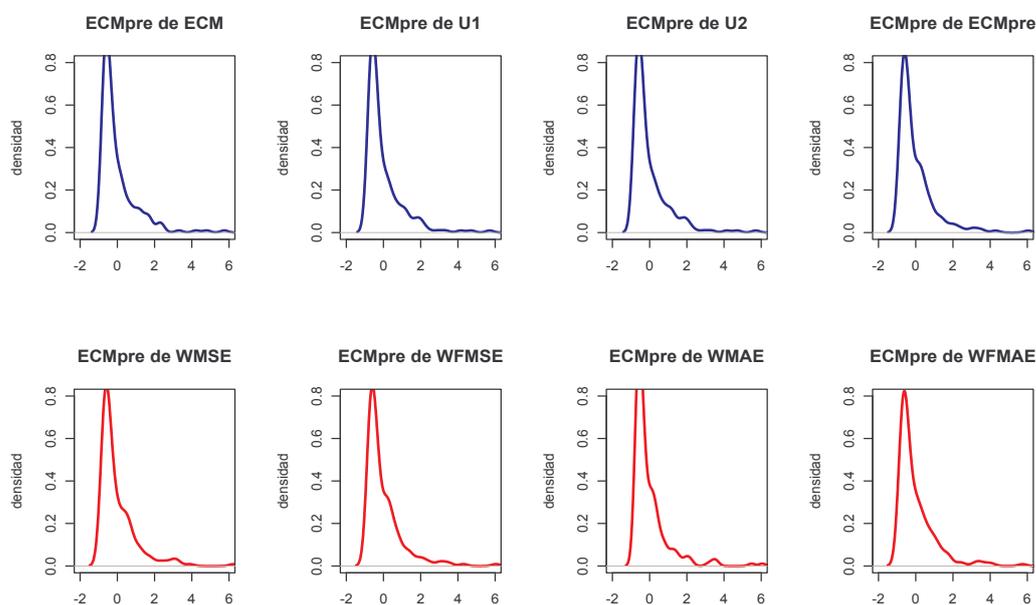


Figura 4.174: ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(1, 1, 0)₇.

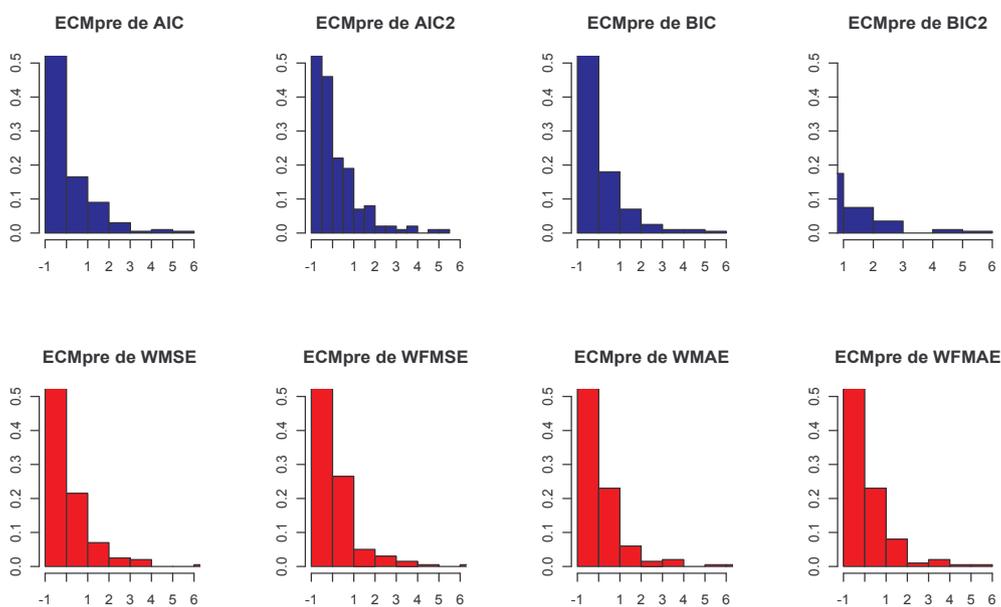


Figura 4.175: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 1, 0)₇.

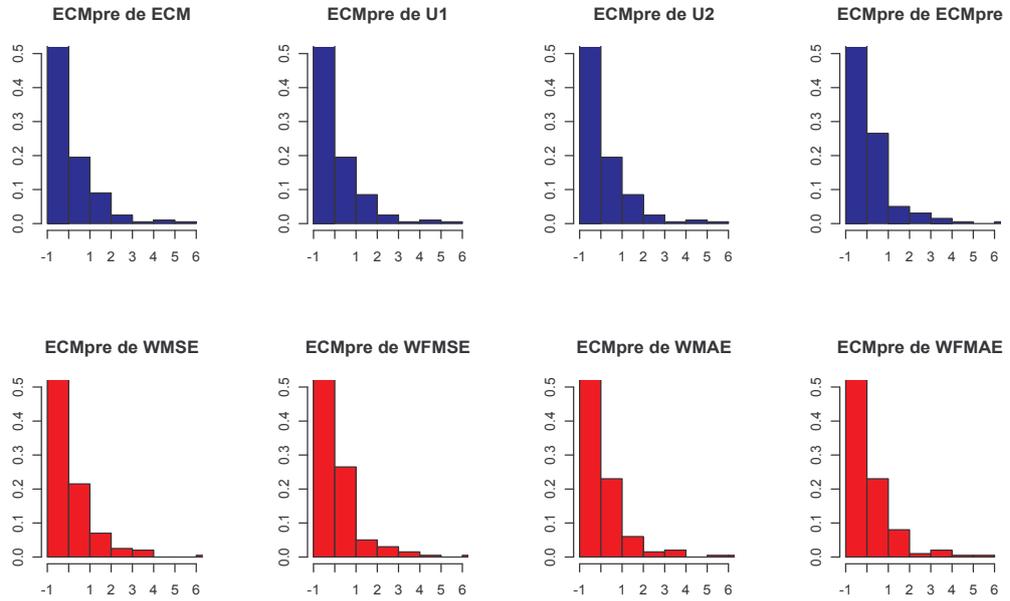


Figura 4.176: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en $SARIMA(1, 1, 0)(1, 1, 0)_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	15.8637	17.1396	16.1979	15.6917	15.6110	16.3030
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	16.3030	21.2291	18.2359	21.2291	19.3179	18.2899

Tabla 4.106: ECM global de predicción, ECMGpre, para cada criterio en modelos $SARIMA(1, 1, 0)(1, 1, 0)_7$.

Comprobamos con la tabla 4.106 que las predicciones no gozan de una gran fiabilidad, por lo que una vez más se depurará el procedimiento con objeto de mejorar los resultados mostrados aquí.

Diferenciamos ahora las series donde se ha logrado identificar el criterio generador de las mismas de aquellas en las que no ha sido posible esa identificación. Los criterios que han sido capaces de detectar en alguna serie (al menos 2 de las 200) el modelo real han sido AIC2 (9%), BIC (10%), BIC2 (62%) y WFMAE (1%).

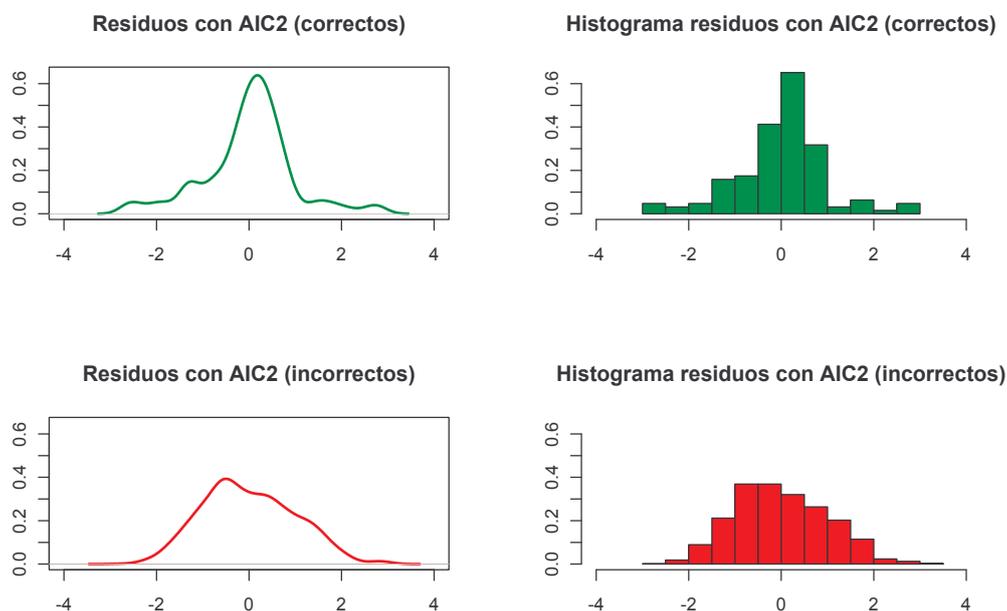


Figura 4.177: Residuos de las series correctas e incorrectas con el criterio AIC2 en el SARIMA(1, 1, 0)(1, 1, 0)₇.

En el gráfico 4.177 se ve como la distribución de residuos en el grupo de series correctas se comporta de mejor manera que en las incorrectas, dato que se puede verificar con la tabla 4.107, en la que se recoge una media y una varianza inferior, lo que nos lleva a un menor valor en el ECMGpre.

	Media	Varianza	ECMG	ECMGpre
Correctos	-0.0705	14.3985	0.9016	14.4034
Incorrectos	0.2863	17.8491	0.8106	17.9311

Tabla 4.107: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(1, 1, 0)₇ con el criterio AIC2.

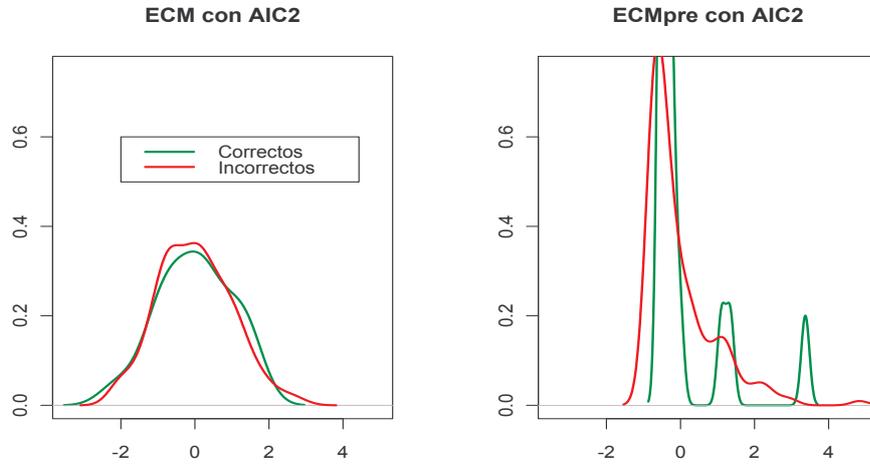


Figura 4.178: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en $SARIMA(1, 1, 0)(1, 1, 0)_7$.

En el gráfico 4.178 las curvas de los valores de los ECM en las series correctas e incorrectas casi se superponen, por lo que las diferencias en cuanto al ajuste en estos grupos no cabe mención alguna.

En la tabla 4.108 se observa como en el primero de los valores estimados el resultado mejora en las series incorrectas, no obstante en los demás es en las series correctas donde se alcanzan las mejores predicciones, es por ello, que el $ECMG_{pre}$ sea inferior en este grupo que en las incorrectas, aunque a la vista del gráfico 4.178 las diferencias no son notables.

	$ECMG_{pre1}$	$ECMG_{pre2}$	$ECMG_{pre3}$	$ECMG_{pre4}$	$ECMG_{pre5}$	$ECMG_{pre6}$	$ECMG_{pre7}$
Correctos	1.1444	3.2863	6.4880	11.6841	18.0985	26.7875	33.3353
Incorrectos	1.0680	4.0201	8.6192	15.0177	22.7987	31.3466	42.6470

Tabla 4.108: ECM global de predicción en $SARIMA(1, 1, 0)(1, 1, 0)_7$ con el criterio AIC2 en el conjunto de comprobación.

Continuamos el estudio con el criterio BIC donde han sido 20 las series que han podido identificarse correctamente.

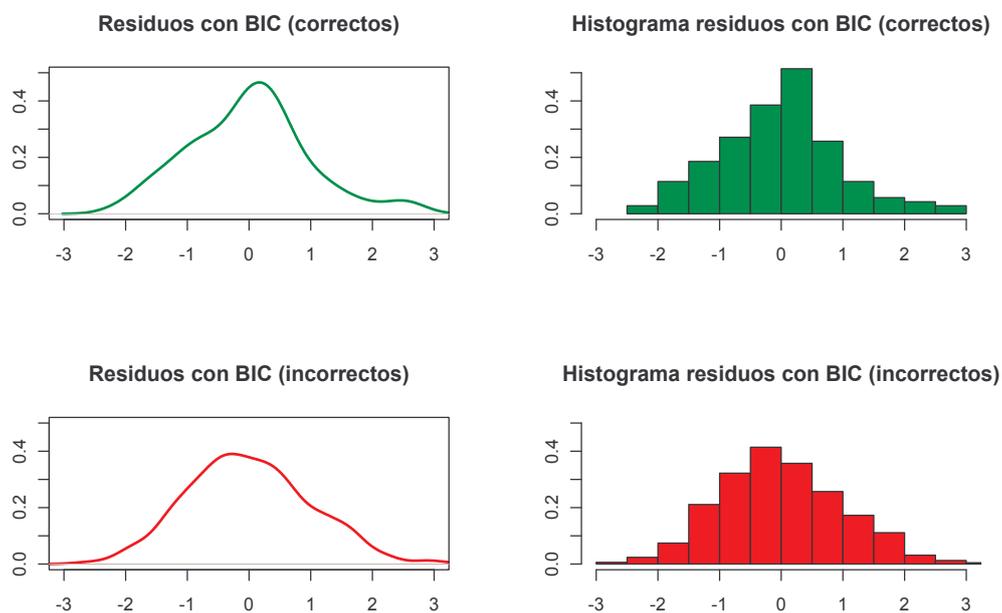


Figura 4.179: Residuos de las series correctas e incorrectas con el criterio BIC en el $SARIMA(1, 1, 0)(1, 1, 0)_7$.

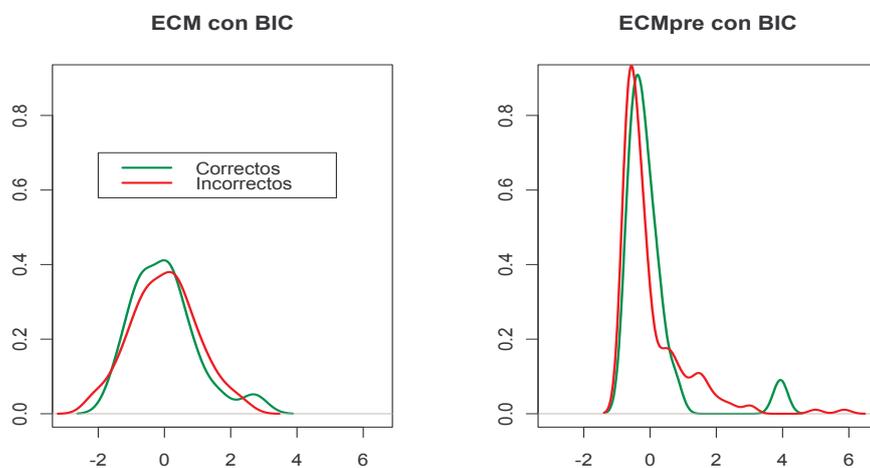


Figura 4.180: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC en $SARIMA(1, 1, 0)(1, 1, 0)_7$.

	Media	Varianza	ECMG	ECMGpre
Correctos	1.7202	12.2015	0.9030	15.1605
Incorrectos	0.1560	16.0972	0.8217	16.1216

Tabla 4.109: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para SARIMA(1, 1, 0)(1, 1, 0)₇ con el criterio BIC.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.7680	2.7464	6.7932	11.9251	19.2864	28.2002	36.4044
Incorrectos	1.0542	3.7757	8.0192	13.9129	20.8224	27.9656	37.3008

Tabla 4.110: ECM global de predicción en SARIMA(1, 1, 0)(1, 1, 0)₇ con el criterio BIC en el conjunto de comprobación.

En el gráfico 4.179 la curva que más se ajusta a la normal es la de los residuos obtenidos con las series incorrectas, sin embargo, a pesar de tener una media significativamente inferior que las correctas, también tienen una varianza mayor, tabla 4.109, lo que lleva a una leve diferencia en los resultados globales tanto de ajuste como de predicción.

En el gráfico 4.180 queda latente que las diferencias entre los grupos son mínimas, y analizando la tabla 4.110 las cinco primeras predicciones se presentan mejores en las series correctas, tan sólo en la predicción 6 las series incorrectas mejoran el dato, aunque esta apreciación carece de relevancia estadística.

Analizando el criterio BIC2, vemos que el comportamiento de los residuos es muy similar a lo que ocurría con el criterio BIC (gráfico 4.181), a la vista de la tabla 4.111 se comprueba la similitud en el comportamiento de ambos criterios, tan sólo en la tabla 4.112 se puede apreciar una leve diferencia, y es que en esta ocasión al contrario de como ocurría anteriormente la primera de las predicciones realizadas alcanza un valor más preciso o lo que es lo mismo presenta un menor ECMpre en las series incorrectas, aunque una vez más las diferencias son mínimas como se representa gráficamente en la figura 4.182.

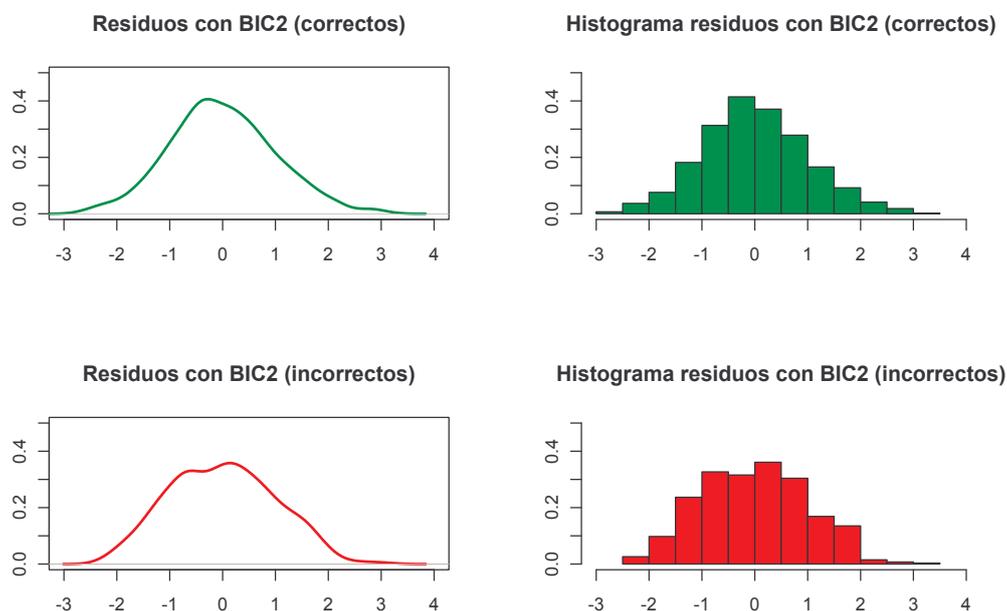


Figura 4.181: Residuos de las series correctas e incorrectas con el criterio BIC2 en el SARIMA(1, 1, 0)(1, 1, 0) $_7$.

	Media	Varianza	ECMG	ECMGpre
Correctos	0.1908	15.2737	0.8844	15.3101
Incorrectos	0.5406	16.6776	0.8267	16.9698

Tabla 4.111: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(1, 1, 0) $_7$ con el criterio BIC2.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	1.1369	3.9163	8.0367	13.6422	19.8609	26.1033	34.4744
Incorrectos	0.8459	3.3696	7.8816	13.8057	21.7027	30.7573	40.4261

Tabla 4.112: ECM global de predicción en SARIMA(1, 1, 0)(1, 1, 0) $_7$ con el criterio BIC2 en el conjunto de comprobación.

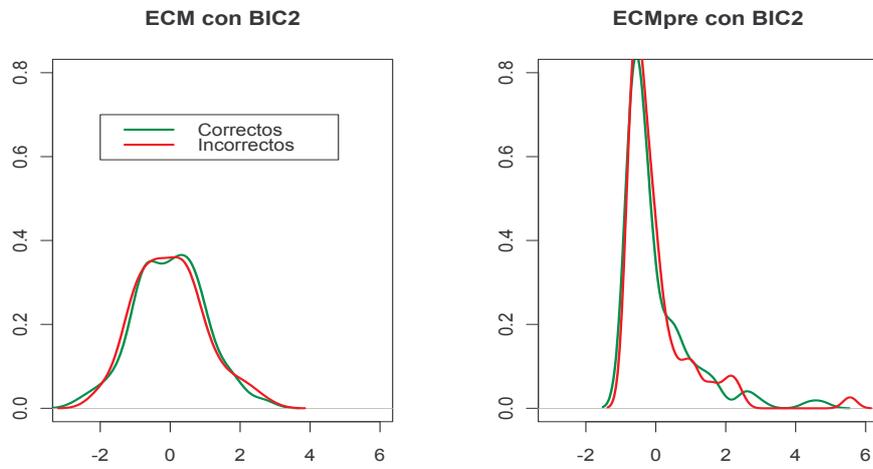


Figura 4.182: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en $\text{SARIMA}(1, 1, 0)(1, 1, 0)_7$.

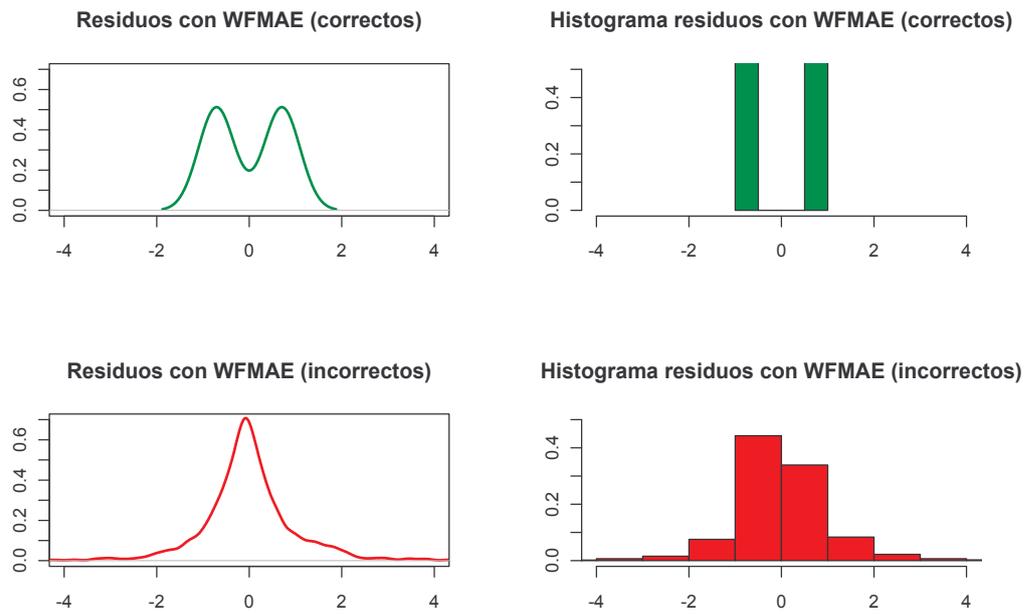


Figura 4.183: Residuos de las series correctas e incorrectas con el criterio WFMAE en el $\text{SARIMA}(1, 1, 0)(1, 1, 0)_7$.

	Media	Varianza	ECMG	ECMGpre
Correctos	2.2759	3.8680	0.7229	9.0475
Incorrectos	0.1216	3.9474	5.7613	3.9622

Tabla 4.113: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(1, 1, 0) $_7$ con el criterio WFMAE.

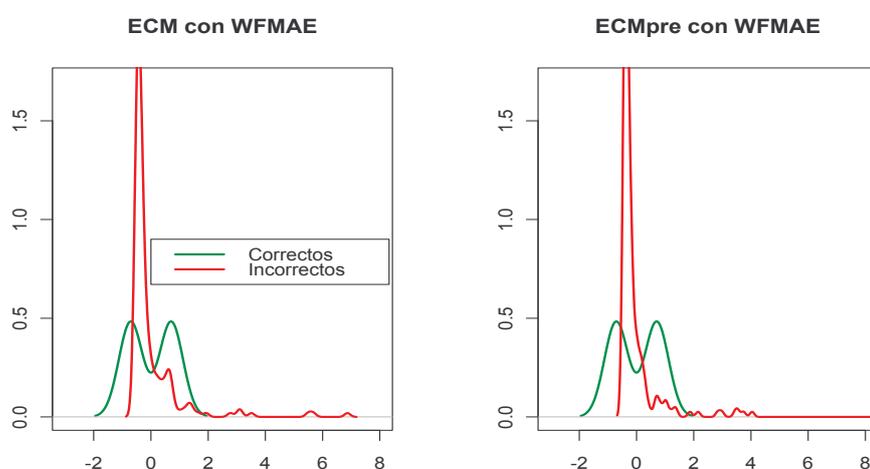


Figura 4.184: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en SARIMA(1, 1, 0)(1, 1, 0) $_7$.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.0024	0.1295	0.0370	4.5341	15.4126	18.6290	24.5880
Incorrectos	0.5491	0.9744	1.8704	2.8791	4.35	6.5704	10.5423

Tabla 4.114: ECM global de predicción en SARIMA(1, 1, 0)(1, 1, 0) $_7$ con el criterio WFMAE en el conjunto de comprobación.

En cuanto al criterio WFMAE se puede observar en el gráfico 4.183 que a pesar de que en el grupo de las series correctas tan sólo se recogen dos series, el grupo de las incorrectas tiene un mejor comportamiento en los residuos lo que se verifica con la tabla 4.113, sin embargo, en la tabla 4.114 se puede ver como las mejores predicciones se consiguen con las series correctas hasta la cuarta de las mismas.

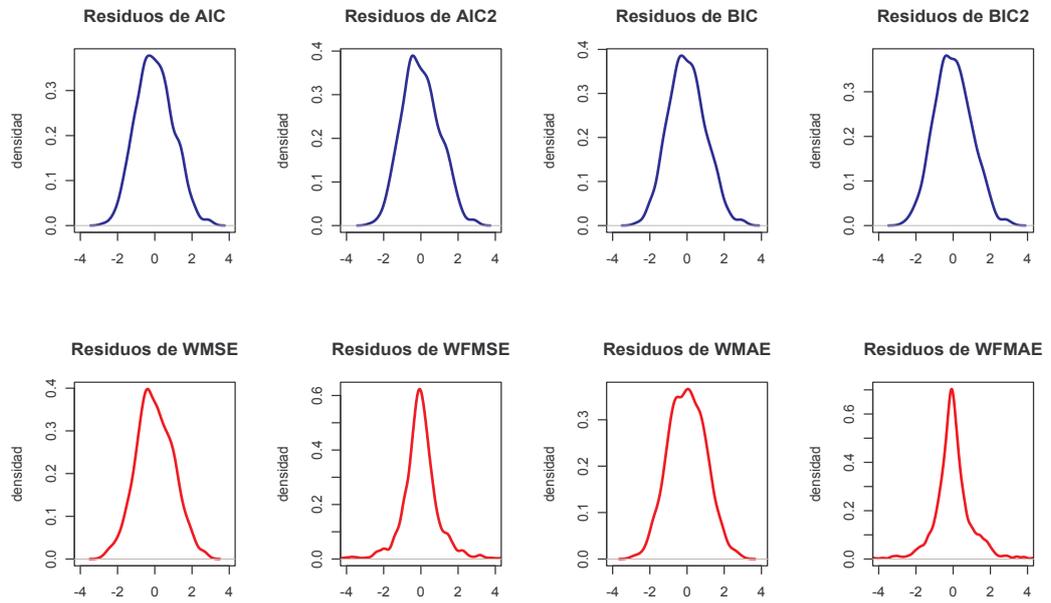


Figura 4.185: Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0)₇.

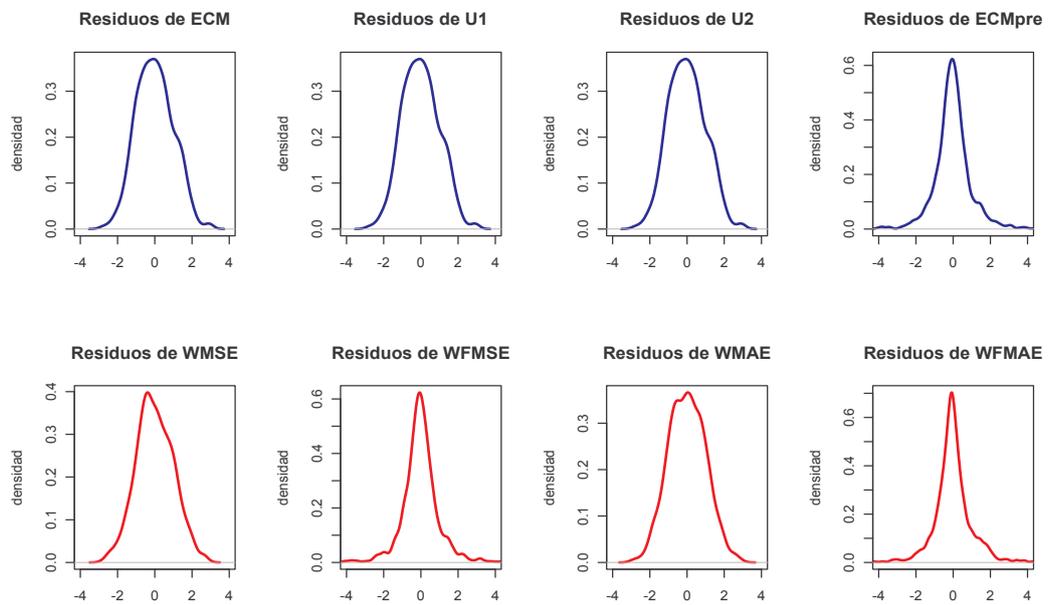


Figura 4.186: Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0)₇.

Criterio	Media	Varianza
AIC	0.2248	17.2449
AIC2	0.2542	17.5490
ECM	0.1918	17.5718
BIC	0.3124	15.9279
BIC2	0.3237	15.8360
U1	0.1918	17.5718
U2	0.1918	17.5718
ECM _{pre}	0.0934	3.4948
WMSE	0.11	19.5627
WFMSE	0.1276	3.5496
WMAE	0.2243	20.1882
WFMAE	0.1431	3.9924

Tabla 4.115: Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para SARIMA(1, 1, 0)(1, 1, 0)₇.

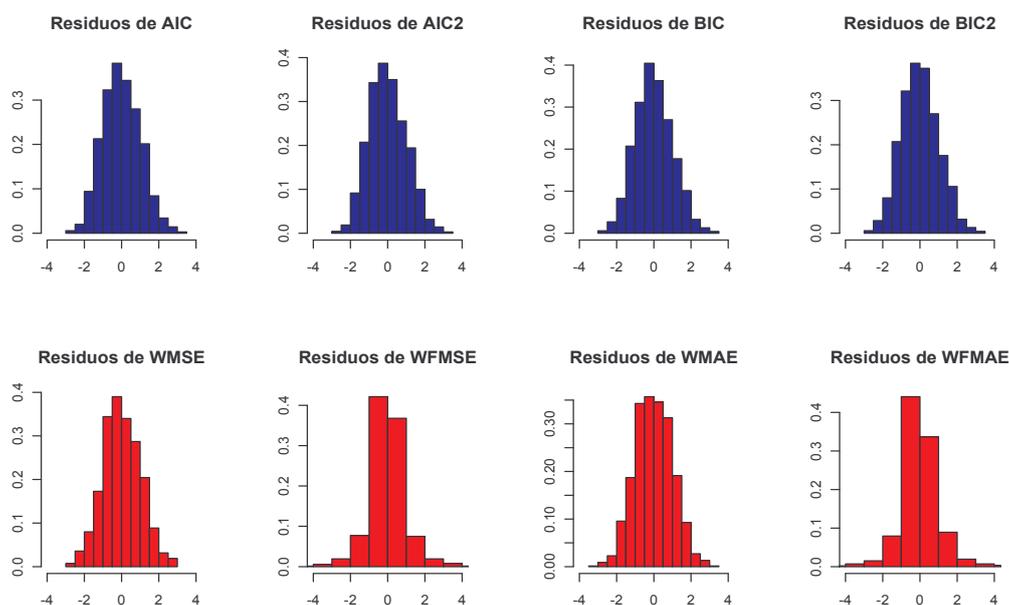


Figura 4.187: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0)₇.

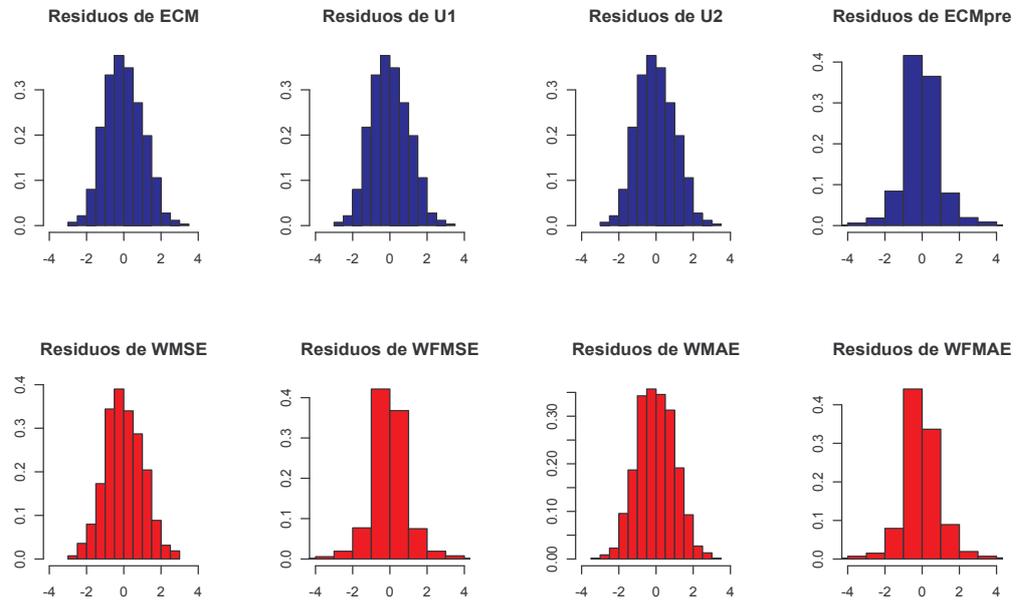


Figura 4.188: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0)₇.

Criterio	ECMpre1	ECMpre2	ECMpre3	ECMpre4	ECMpre5	ECMpre6	ECMpre7
AIC	1.0689	3.9092	8.2516	14.4425	22.1073	30.3758	40.9129
AIC2	1.0749	3.9540	8.4274	14.7177	22.3757	30.9363	41.8090
ECM	1.07	3.9606	8.4051	14.7917	22.6573	31.0370	41.3383
BIC	1.0256	3.6728	7.8966	13.7142	20.6688	27.9891	37.2112
BIC2	1.0263	3.7086	7.9778	13.7043	20.5608	27.8718	36.7361
U1	1.07	3.9606	8.4051	14.7917	22.6573	31.0370	41.3383
U2	1.07	3.9606	8.4051	14.7917	22.6573	31.0370	41.3383
ECMpre	0.9292	1.6192	2.3537	2.8812	3.8984	5.0966	7.7462
WMSE	1.1275	4.4172	9.6566	16.8893	25.3880	34.5195	45.0259
WFMSE	0.8104	1.3997	2.0956	2.7449	3.9427	5.4343	8.5335
WMAE	1.2028	4.7014	10.0839	17.4665	26.3584	35.3265	46.5304
WFMAE	0.5436	0.9659	1.8521	2.8954	4.4602	6.6909	10.6823

Tabla 4.116: ECM globales para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 0)(1, 1, 0)₇.

Sin necesidad de mirar nada más que la s gráficas 4.185 y 4.186 se puede detectar una clara mejoría en los criterios ECMpre, WFMSE y WFMAE, no obstante esto es apoyado por la tabla 4.115 donde se aprecia la disminución en la varianza, aunque es en la tabla 4.188 donde queda reflejado detalladamente la mejoría en cuanto a predicción se refiere, ya que podemos ver como las predicciones son mejores una a una en estos criterios, mejorando además el aumento del error a lo largo del horizonte de predicción, donde es más lento que en el resto de criterios.

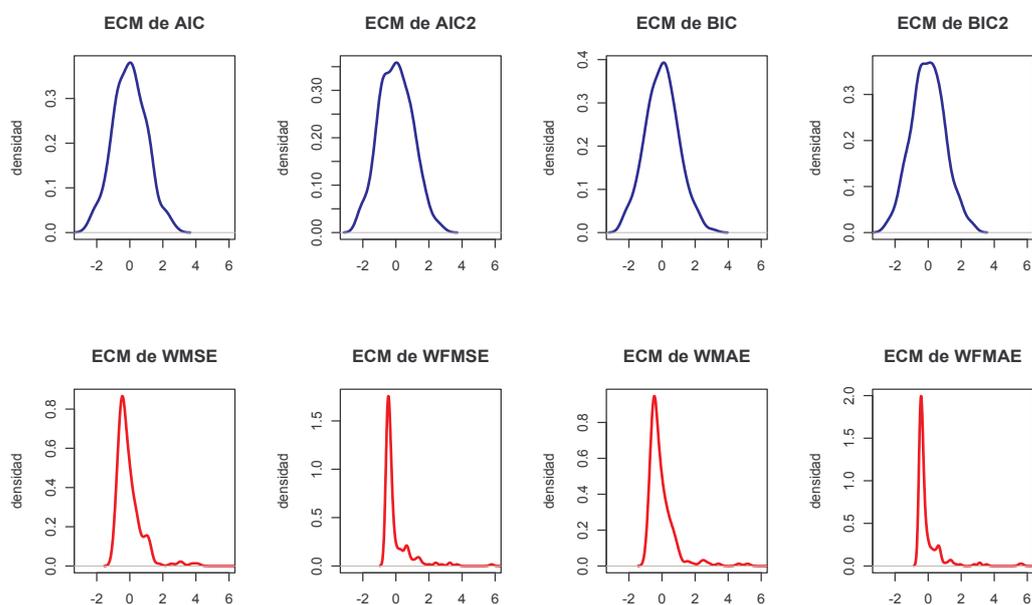


Figura 4.189: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0)₇.

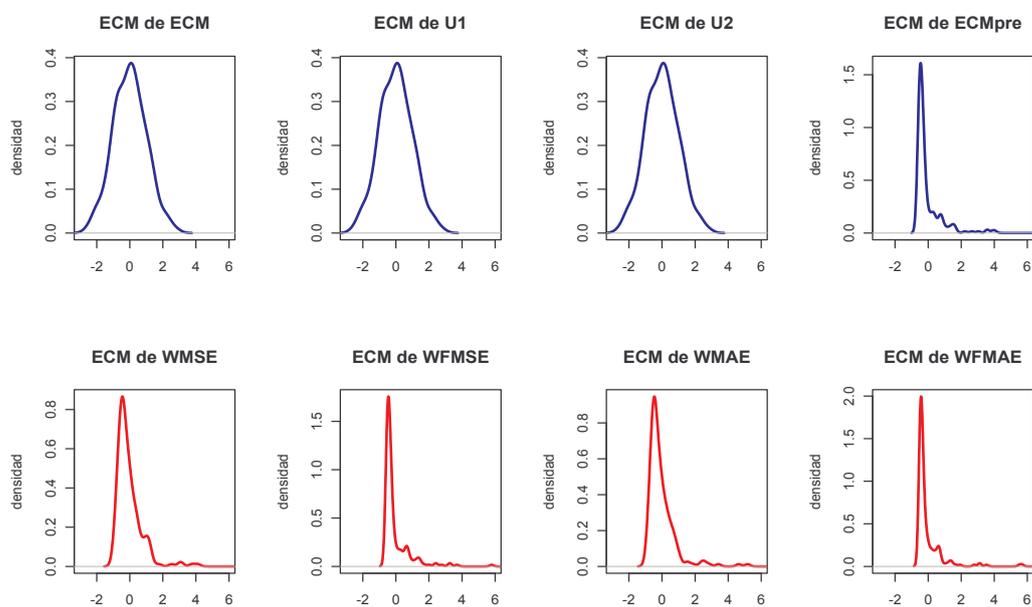


Figura 4.190: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0)₇.

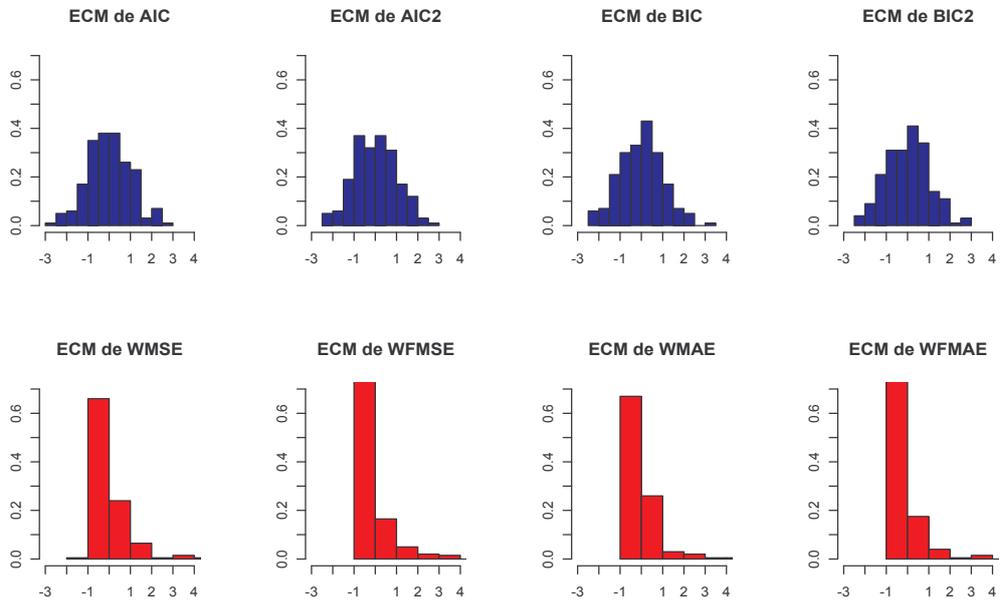


Figura 4.191: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 0)(1, 1, 0)_7$.

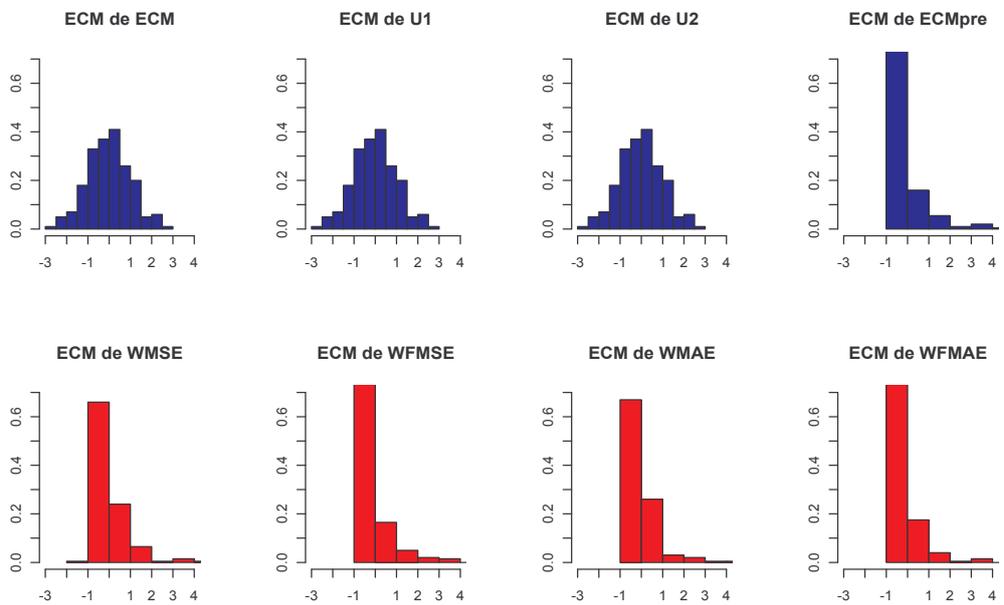


Figura 4.192: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 0)(1, 1, 0)_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8050	0.8188	0.7950	0.8298	0.8625	0.7950
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.7950	5.0185	1.1984	5.5606	1.2918	5.7117

Tabla 4.117: ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 0)(1, 1, 0) $_7$.

En los gráficos 4.189 y 4.190 y con mayor claridad en los gráficos 4.191 y 4.192 se puede ver como en los criterios ECMpre, WMSE, WFMSE, WMAE y WFMAE el ajuste es inexistente, como viene ocurriendo en los demás casos en cuanto a ajuste el fracaso de estos criterios es más que evidente, y una vez más a la vista de la tabla 4.117 el criterio de Theil funciona bien al respecto, a pesar de ser un criterio de predicción.

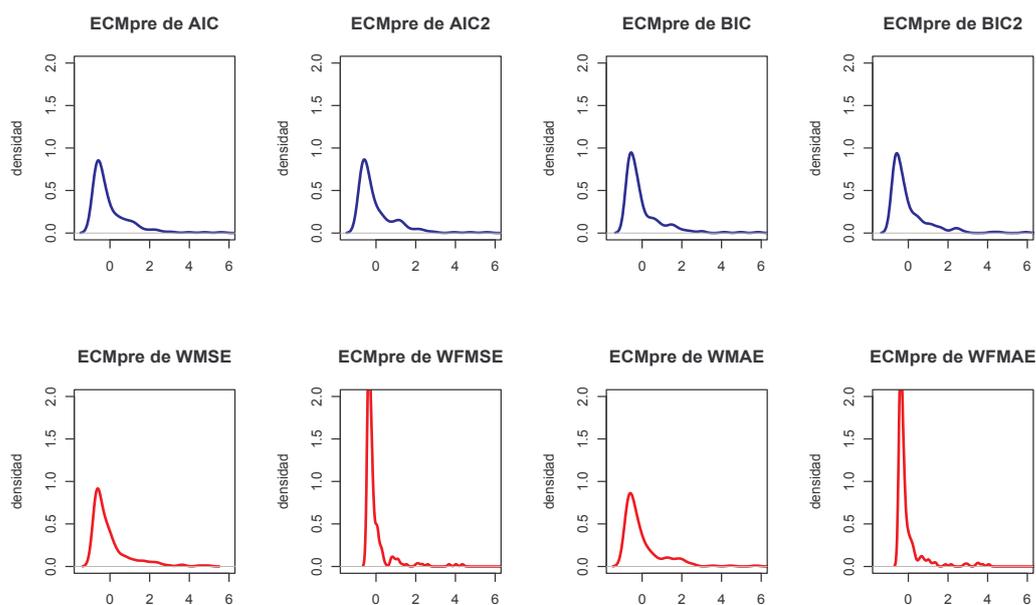


Figura 4.193: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) $_7$.

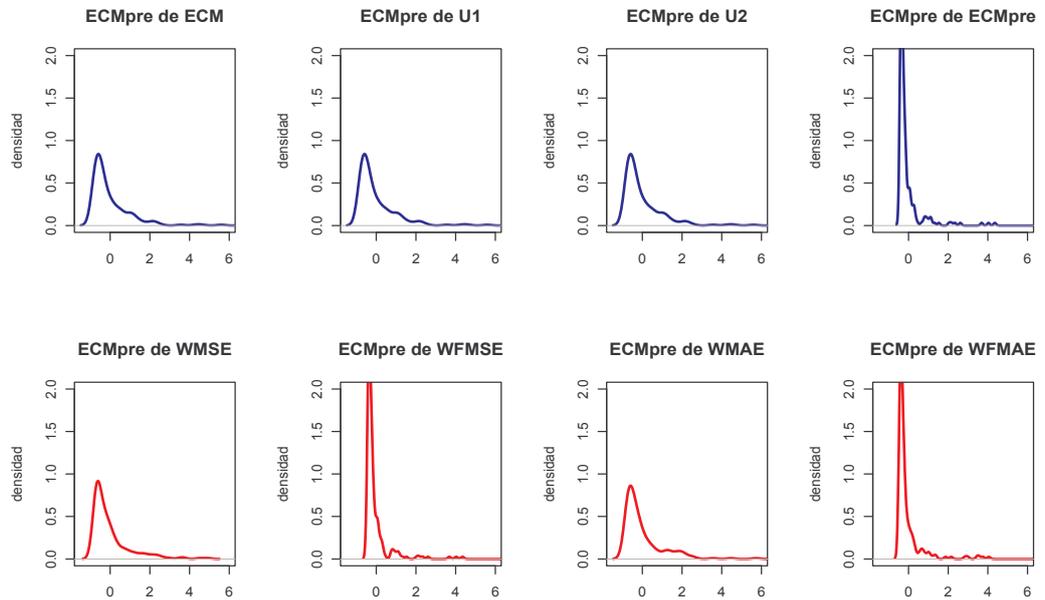


Figura 4.194: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 0)(1, 1, 0)_7$.

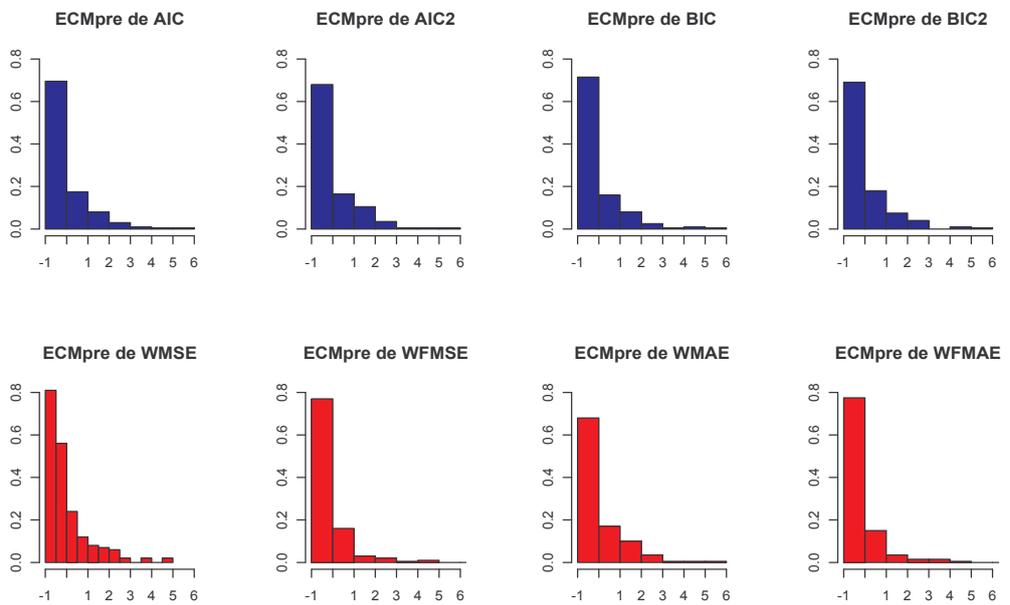


Figura 4.195: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 0)(1, 1, 0)_7$.

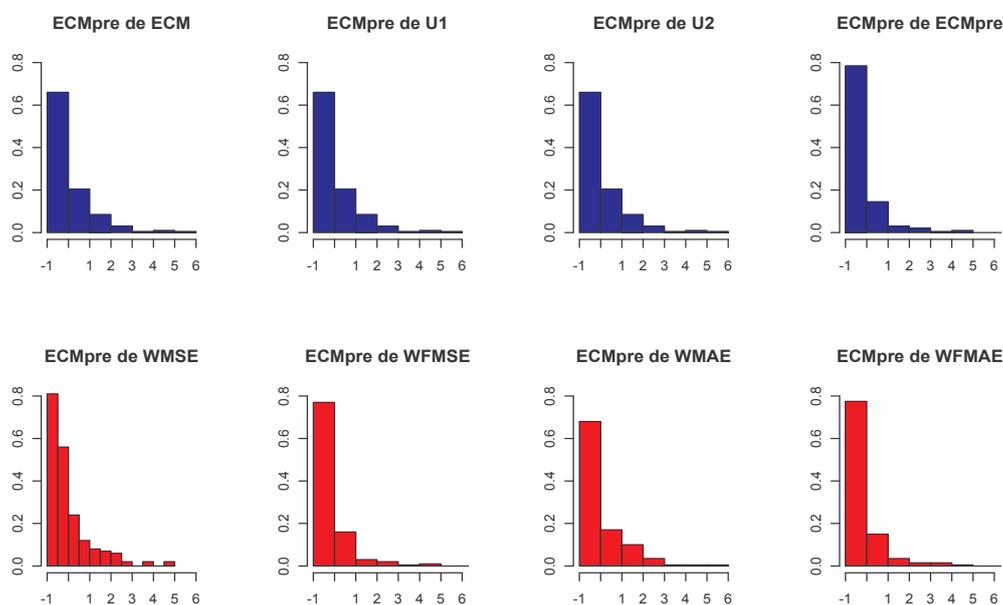


Figura 4.196: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 1, 0) $_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	17.2955	17.6136	17.6086	16.0255	15.9408	17.6086
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	17.6086	3.5035	19.5748	3.5659	20.2385	4.0129

Tabla 4.118: ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 0)(1, 1, 0) $_7$.

Es en la tabla 4.194 donde queda más que probado el éxito obtenido en la predicción de los criterios ECMpre, WFMSE y WFMAE, esto ya viene ocurriendo en los casos anteriores y ante modelos de todo tipo, mejorando también en este caso los resultados obtenidos en las tablas 4.104 y 4.106. El resto de criterios mantienen su comportamiento de manera similar.

Modelo SARIMA(1, 1, 0)(1, 0, 1)₇, cuyos coeficientes son $\phi_1 = 0,7$, $\Phi_1 = 0,5$ y $\Theta = 0,3$. (**Modelo 170**)

	AIC	AIC2	ECM	BIC	BIC2
Valor	-0.6357156	-3.376547	0.0291725	-0.5750404	-3.315872
Modelo	193	193	321	175	175

Tabla 4.119: Criterios Clásicos de ajuste para SARIMA(1, 1, 0)(1, 0, 1)₇.

	U_1	U_2	ECMpre
Valor	8.534454e-3	8.534765e-3	0.037494
Modelo	296	296	296

Tabla 4.120: Criterios Clásicos de predicción para SARIMA(1, 1, 0)(1, 0, 1)₇.

	Ajuste		Predicción	
	WMSE	WMAE	WFMSE	WFMAE
Valor	0.02451431	0.1393009	0.02860281	0.1247465
Modelo	155	155	216	216

Tabla 4.121: Criterios propuestos para SARIMA(1, 1, 0)(1, 0, 1)₇.

En este modelo, la selección del modelo “óptimo” es más afinada que en el resto de casos, todos los modelos seleccionados a excepción del 321, han sido elegidos al menos por dos de los criterios (tabla 4.119-4.121), los criterios AIC y AIC2 no presentan diferencias en esta ocasión ante la determinación del modelo, los criterios BIC y BIC2 también han seleccionado el mismo modelo, que equivale a un SARIMA(1, 1, 0)(2, 0, 0)₇, los nuevos criterios de ajuste que se han introducido también se decantan por el mismo modelo, todos los criterios de predicción clásicos estudiados han seleccionado como mejor modelo el mismo, siendo éste el 296 (ver apéndice C), y por último, los criterios de predicción desarrollados en este trabajo también determinan el mismo modelo en su elección.

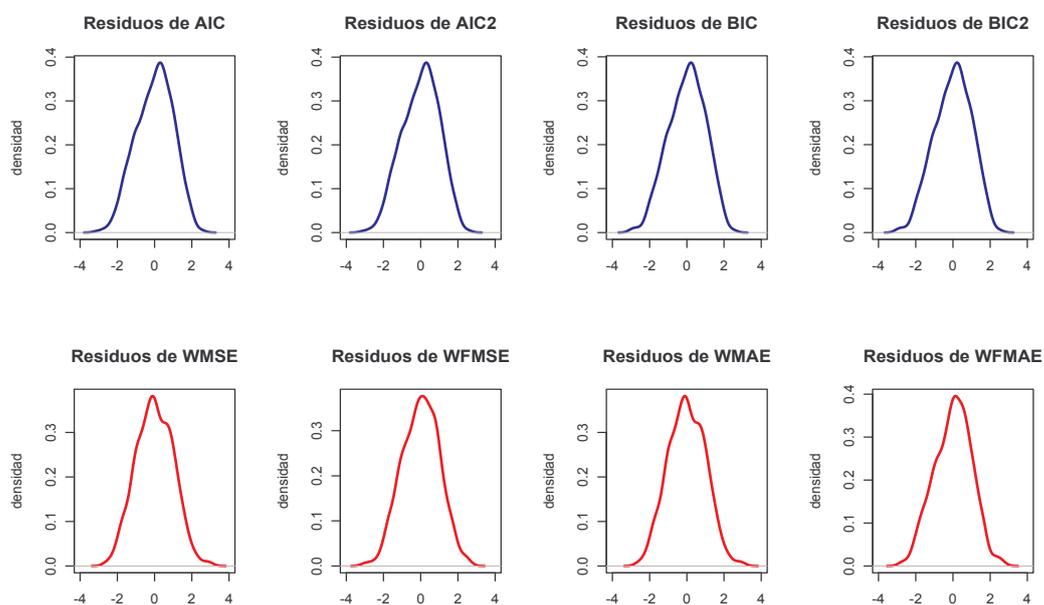


Figura 4.197: Residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 0)(1, 0, 1)₇.

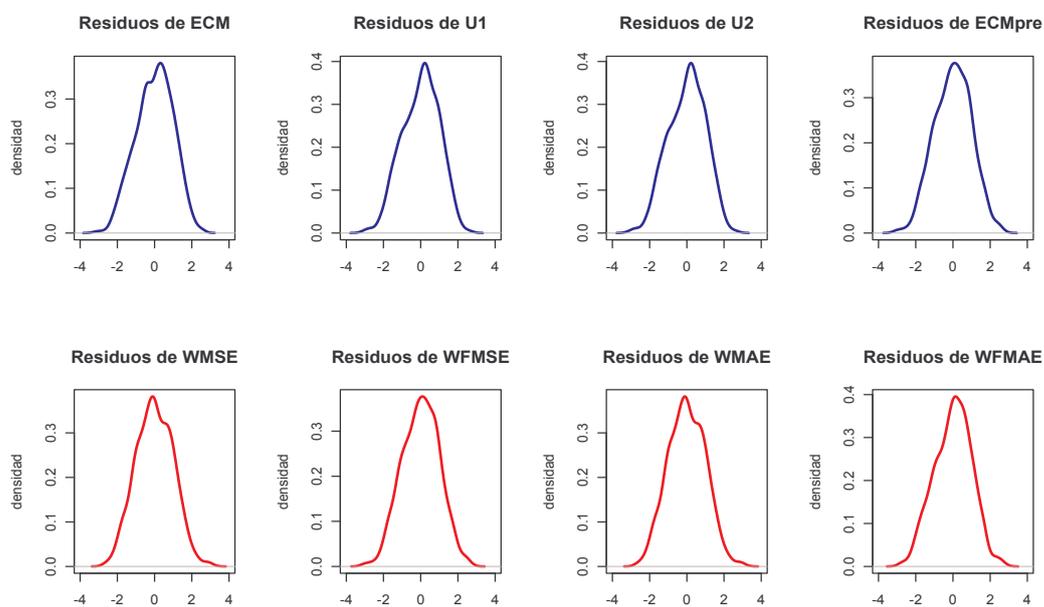


Figura 4.198: Residuos obtenidos con los modelos de los criterios de predicci3n de SARIMA(1, 1, 0)(1, 0, 1)₇.

Criterio	Media	Varianza
AIC	-0.2946	17.8207
AIC2	-0.2946	17.8207
ECM	-0.3277	18.0168
BIC	-0.2850	18.1425
BIC2	-0.2850	18.1425
U1	-0.3276	17.4546
U2	-0.3276	17.4546
ECMpre	-0.2427	21.8069
WMSE	-0.3263	21.3674
WFMSE	-0.2427	21.8069
WMAE	-0.3263	21.3674
WFMAE	-0.2525	20.9816

Tabla 4.122: Medidas descriptivas de los residuos en el conjunto de comprobación para SARIMA(1, 1, 0)(1, 0, 1)₇.

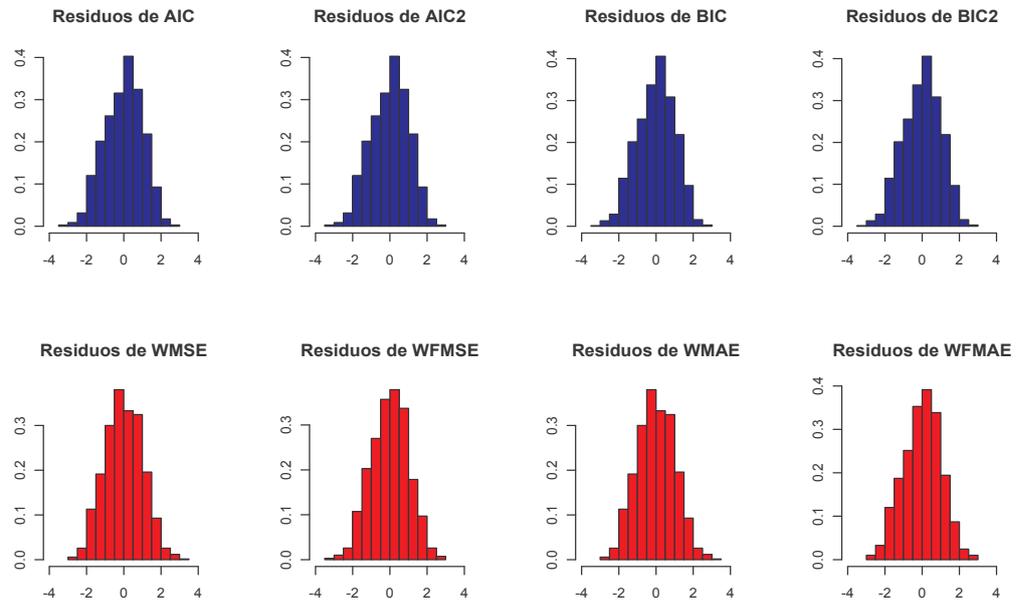


Figura 4.199: Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 0)(1, 0, 1)₇.

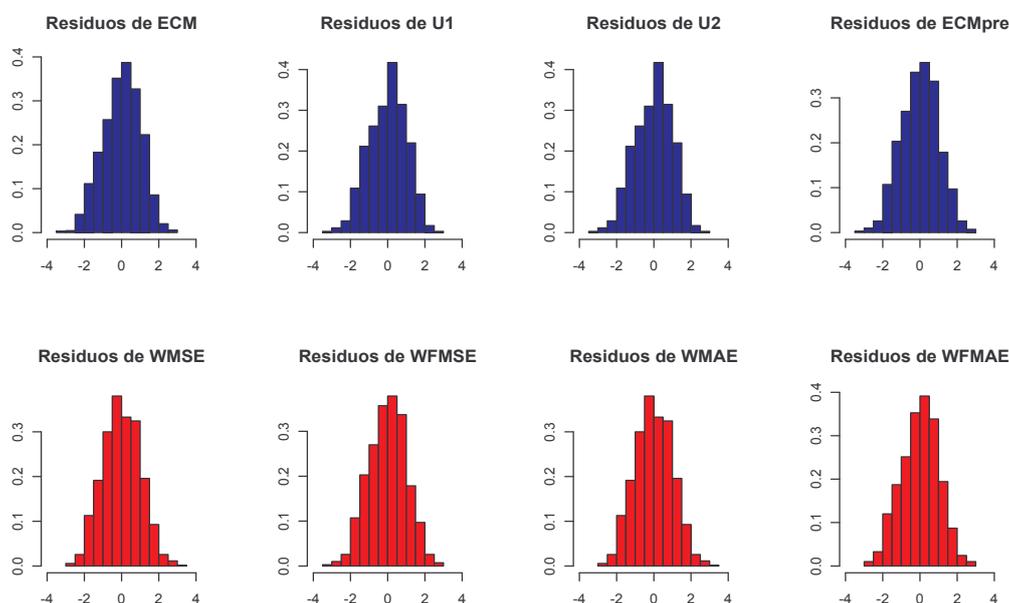


Figura 4.200: Histograma de los residuos obtenidos con los modelos de los criterios de predicción de SARIMA(1, 1, 0)(1, 0, 1) $_7$.

Criterio	ECMpre1	ECMpre2	ECMpre3	ECMpre4	ECMpre5	ECMpre6	ECMpre7
AIC	1.1558	3.7847	7.8544	14.4660	22.6972	32.1920	43.2024
AIC2	1.1558	3.7847	7.8544	14.4660	22.6972	32.1920	43.2024
ECM	1.1535	3.7860	7.9604	14.8044	23.1623	32.5950	43.4076
BIC	1.1611	3.8185	7.9838	14.7628	23.1723	32.8083	43.8590
BIC2	1.1611	3.8185	7.9838	14.7628	23.1723	32.8083	43.8590
U1	1.1469	3.7102	7.6536	14.1870	22.3970	31.5957	42.2429
U2	1.1469	3.7102	7.6536	14.1870	22.3970	31.5957	42.2429
ECMpre	1.3640	4.6192	9.7325	17.7961	28.2325	39.7673	51.5489
WMSE	1.3572	4.5601	10.1456	18.3193	28.0086	38.4865	49.4396
WFMSE	1.3640	4.6192	9.7325	17.7961	28.2325	39.7673	51.5489
WMAE	1.3572	4.5601	10.1456	18.3193	28.0086	38.4865	49.4396
WFMAE	1.3025	4.4647	9.5042	17.2384	27.0924	38.1588	49.5566

Tabla 4.123: ECMpre global para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 0)(1, 0, 1) $_7$.

En esta ocasión ni los gráficos 4.197-4.200 ni las tablas 4.122 y 4.123 merecen mención alguna dado que el comportamiento no presenta importantes diferencias a lo largo de los criterios. Por tanto, pasamos directamente al estudio de los ECM de ajuste y predicción.

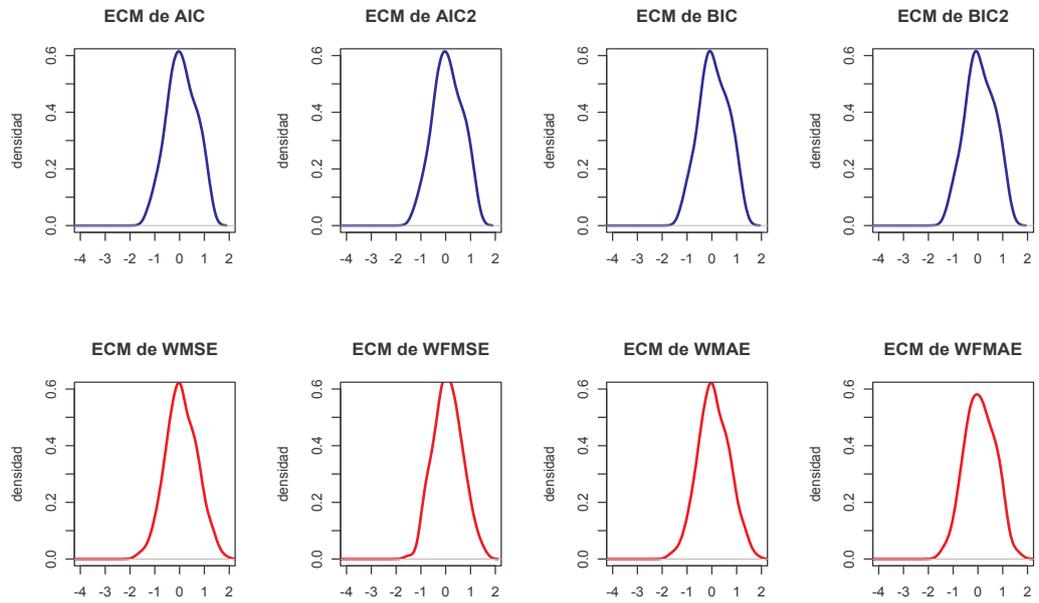


Figura 4.201: ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 0, 1)₇.

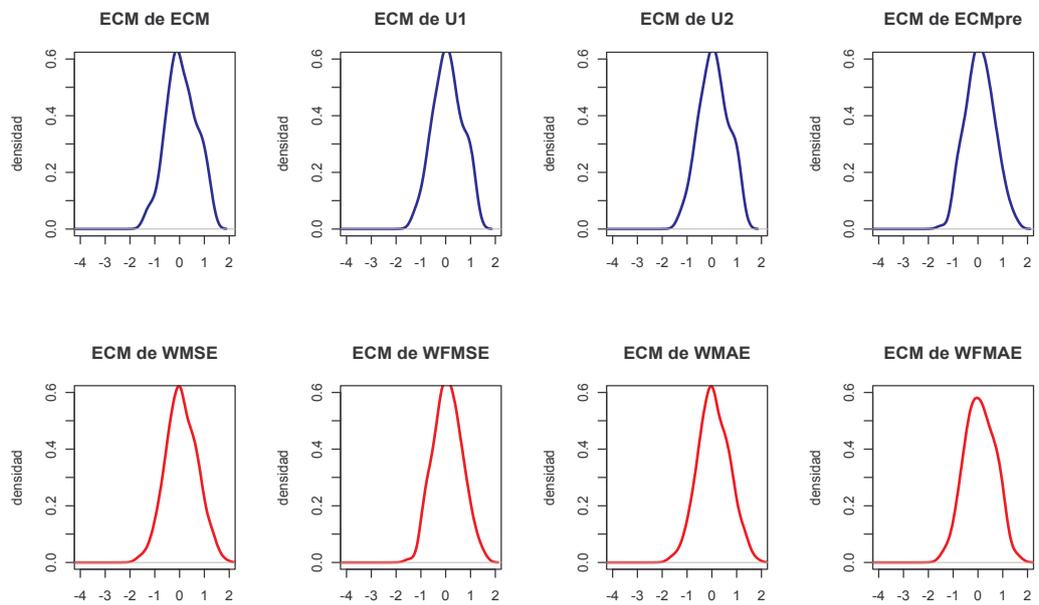


Figura 4.202: ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(1, 0, 1)₇.

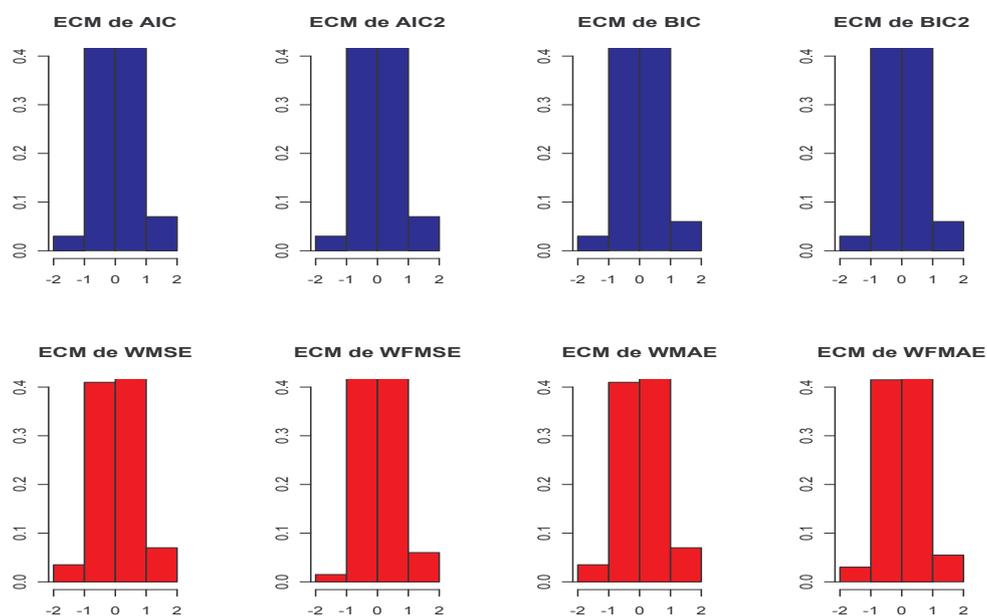


Figura 4.203: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 0, 1)₇.

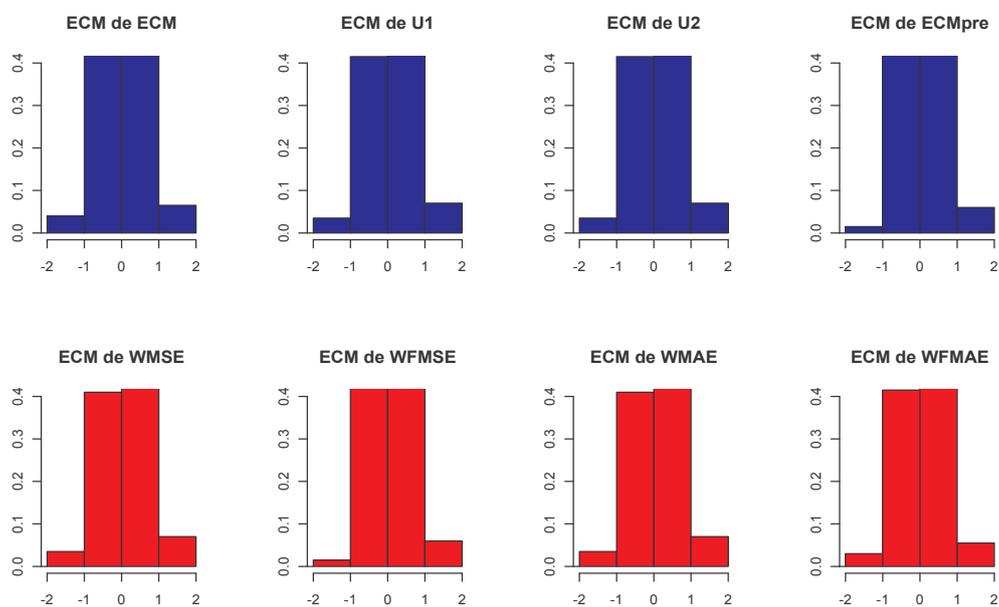


Figura 4.204: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicci3n en SARIMA(1, 1, 0)(1, 0, 1)₇.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8745	0.8745	0.8341	0.8812	0.8812	0.8966
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8966	1.0716	1.0978	1.0716	1.0978	0.9197

Tabla 4.124: El ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 0)(1, 0, 1)₇.

En la tabla superior (4.124) los resultados referentes al ajuste una vez más se comportan de manera similar, siendo el mejor ajuste el determinado por el modelo seleccionado por el criterio ECM, que es el único que ha coincidido en su elección con ningún otro, no obstante, este resultado tiene una validez dudosa ya que el modelo seleccionado es un SARIMA(2, 1, 2)(2, 0, 2)₇ lo que complica al modelo real, y por tanto, no se puede determinar si es porque el comportamiento de las series se ajusta mejor a éste, o porque, cuando más parámetros seleccionamos más se perfilan los resultados. No obstante, este es un problema que podría abordarse en estudios futuros pero en este no es un tema en que vayamos a reparar.

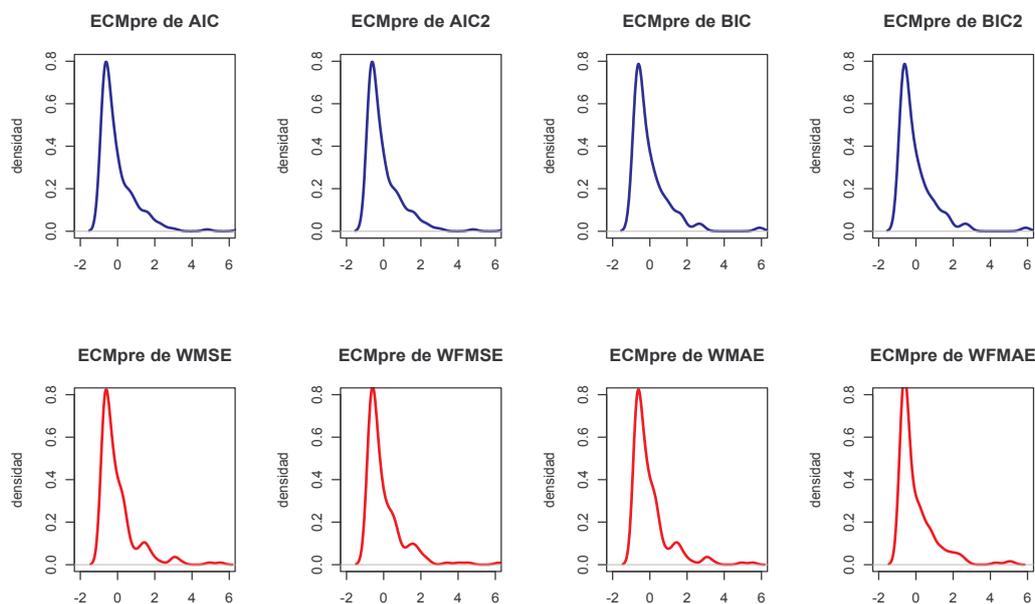


Figura 4.205: ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 0, 1)₇.

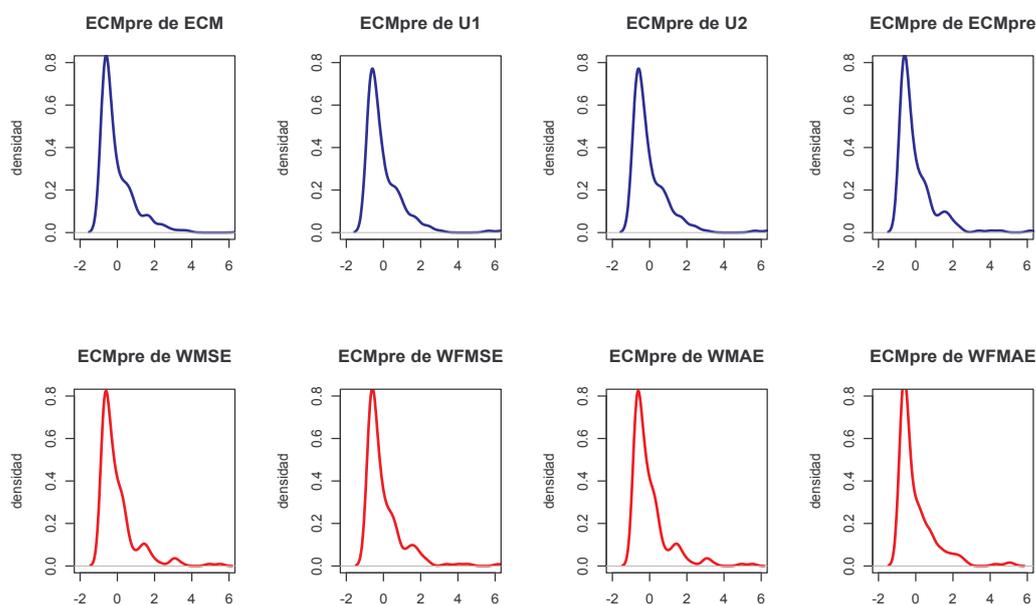


Figura 4.206: ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 0)(1, 0, 1)₇.

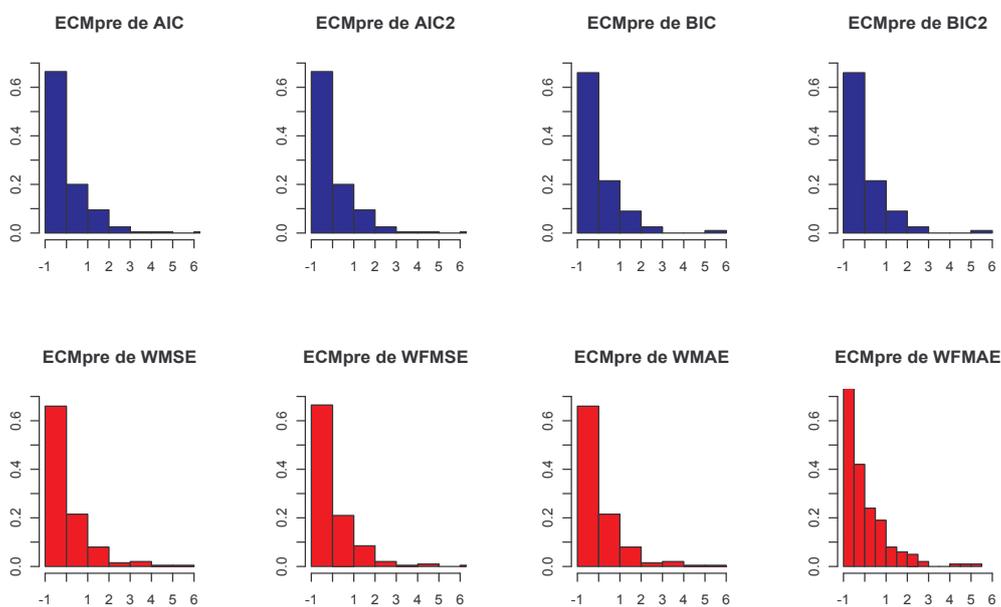


Figura 4.207: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 0)(1, 0, 1)₇.

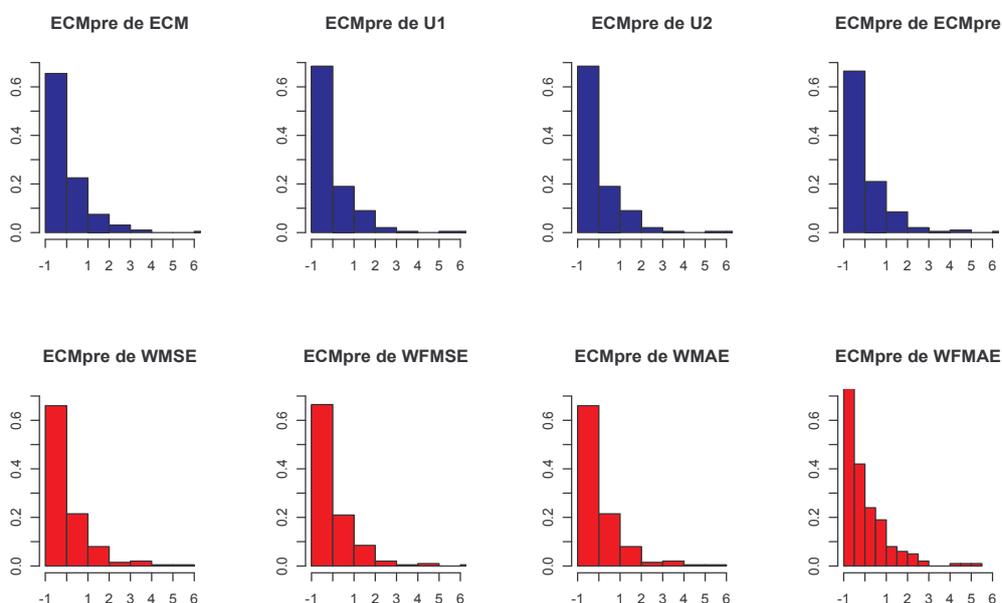


Figura 4.208: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en $SARIMA(1, 1, 0)(1, 0, 1)_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	17.9075	17.9075	18.1242	18.2237	18.2237	17.5619
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	17.5619	21.8658	21.4739	21.8658	21.4739	21.0454

Tabla 4.125: ECM global de predicción, ECMGpre, para cada criterio en modelos $SARIMA(1, 1, 0)(1, 0, 1)_7$.

La tabla 4.125 es muy similar a la obtenida en casos anteriores lo que sería válidas las explicaciones desarrolladas entonces, por ejemplo, para el modelo $SARIMA(1, 1, 0)(1, 1, 0)_7$ que se vió anteriormente a este que se expone en estos momentos (pág. 174).

Continuamos por tanto, con el estudio de validación, donde se diferencian dos grupos de series. Los criterios utilizados como discriminantes para los grupos han sido en esta ocasión el AIC2 (3%), BIC2 (2%) y WFMAE (1%). No vamos a extender los comentarios al respecto dado que los porcentajes son bajísimos lo que nos complica poder obtener alguna conclusión estable.

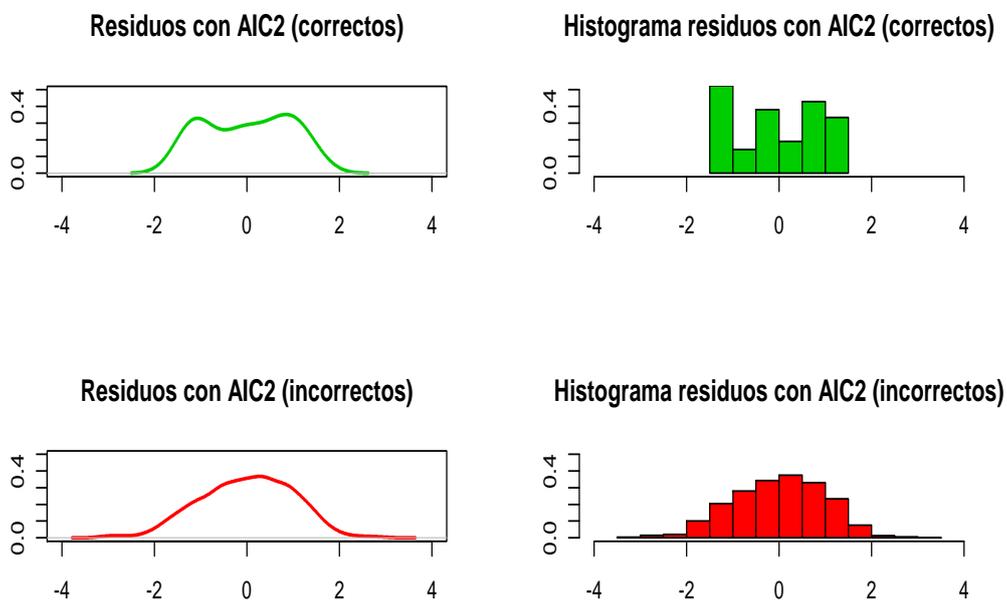


Figura 4.209: Residuos de las series correctas e incorrectas con el criterio AIC2 en el SARIMA(1, 1, 0)(1, 0, 1)₇.

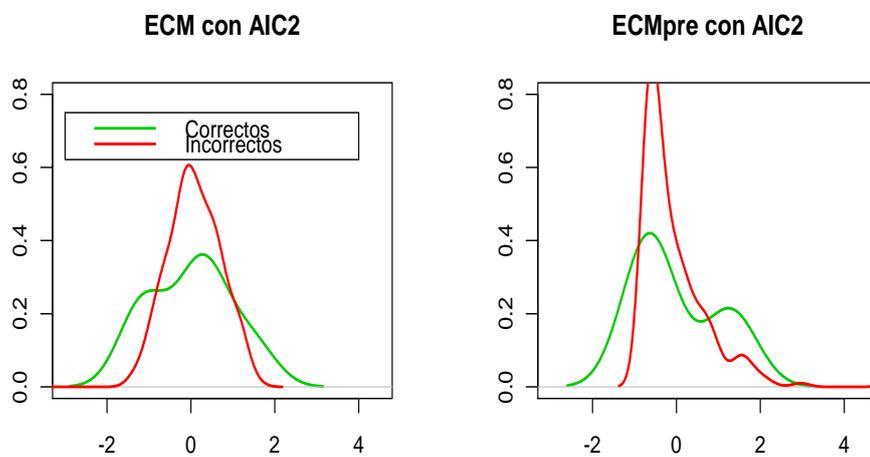


Figura 4.210: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en SARIMA(1, 1, 0)(1, 0, 1)₇.

	Media	Varianza	ECMG	ECMGpre
Correctos	-1.7625	11.7303	0.9145	14.8368
Incorrectos	1.2142	19.1872	0.865	19.2510

Tabla 4.126: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(1, 0, 1)₇ con el criterio AIC2.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.5138	1.9566	6.0223	16.0672	23.7908	29.7313	25.7759
Incorrectos	1.2142	3.9229	8.2636	15.2982	24.1957	34.6644	47.1982

Tabla 4.127: ECM global de predicción en SARIMA(1, 1, 0)(1, 0, 1)₇ con el criterio AIC2 en el conjunto de comprobación.

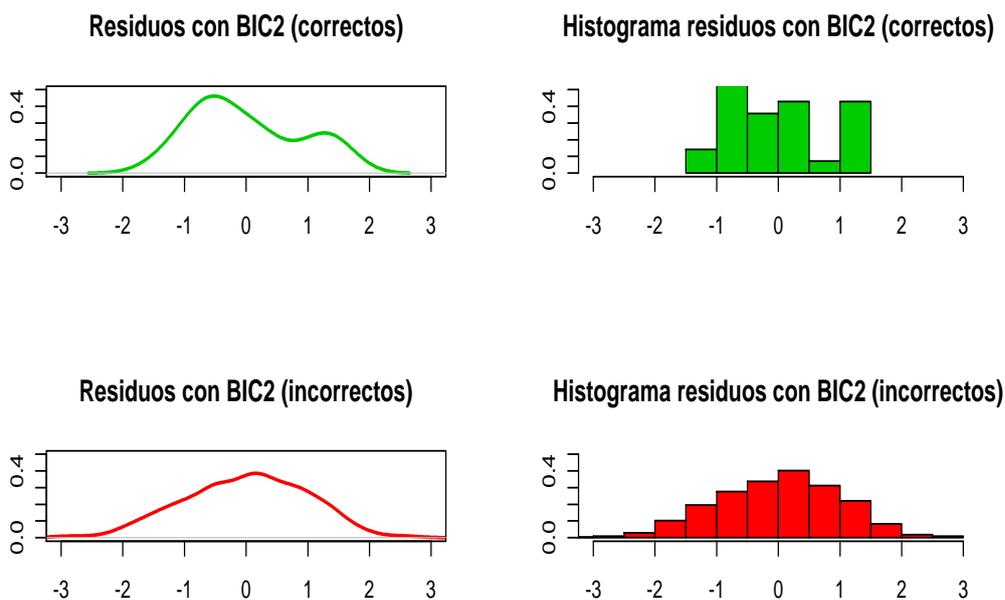


Figura 4.211: Residuos de las series correctas e incorrectas con el criterio BIC2 en el SARIMA(1, 1, 0)(1, 0, 1)₇.

	Media	Varianza	ECMG	ECMGpre
Correctos	-3.9376	13.2463	0.9441	28.7512
Incorrectos	-0.2063	17.9625	0.9132	18.0051

Tabla 4.128: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(1, 0, 1)₇ con el criterio BIC2.

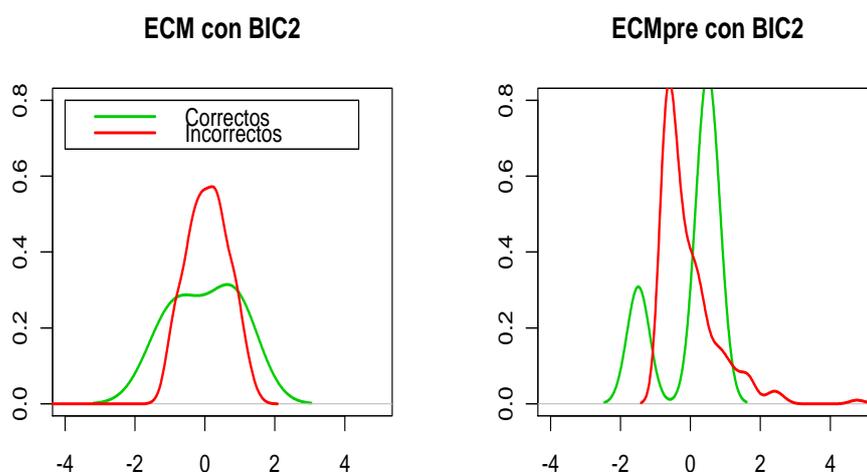


Figura 4.212: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en SARIMA(1, 1, 0)(1, 0, 1)₇.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.2867	1.3443	7.7696	24.4852	40.9735	63.3921	63.0070
Incorrectos	1.1737	3.8401	7.9105	14.5840	22.8460	32.0929	43.5883

Tabla 4.129: ECM global de predicción en SARIMA(1, 1, 0)(1, 0, 1)₇ con el criterio BIC2 en el conjunto de comprobación.

Con respecto al criterio AIC2, en la gráfica 4.209 se puede apreciar la similitud de la curva de los residuos de las series incorrectas con la curva de una distribución normal, sin embargo, en la tabla 4.126 se aprecia una mayor varianza en este grupo frente al de las correctas, y se puede ver como el ECMGpre es más elevado también, por tanto, las predicciones serían mejores en las series correctas, no así el ajuste donde es en las series incorrectas donde se alcanza un valor mejor. A la vista de la tabla 4.127 donde se presentan los ECMGpre individuales se observan mejores resultados en el conjunto de series correctas.

En cuanto al criterio BIC2, se tiene en la gráfica 4.211 el mismo resultado que para el criterio AIC2, también se comprueba en la tabla 4.128 una mayor varianza en el grupo de series incorrectas, no obstante, en este caso el mejor valor del ECMGpre es también en este grupo donde se alcanza al igual que ocurre con el valor del ECMG, sin embargo, analizando la tabla 4.129 tenemos que las mejores predicciones en las tres primeras observaciones se consiguen con mayor precisión en las series correctas, siendo el resto de las predicciones mejor en las incorrectas, de ahí, que el resultado global sea mejor en las series incorrectas.

Por último, con respecto al criterio WFMAE, en la gráfica 4.213 obtenemos el mismo resultado que en los casos anteriores, en la tabla 4.130 se comprueba también la mayor varianza dentro del grupo de las series incorrectas, y en esta ocasión también los resultados son peores en cuanto al ECMG y al ECMGPre, en la tabla 4.131 se puede ver que las mejores predicciones se dan en el grupo de series correctas, es decir, utilizando el modelo generador de las series, y además se observa que el crecimiento del error cometido a lo largo del periodo de predicción aumenta de manera más lineal que el aumento que se baraja en las series incorrectas.

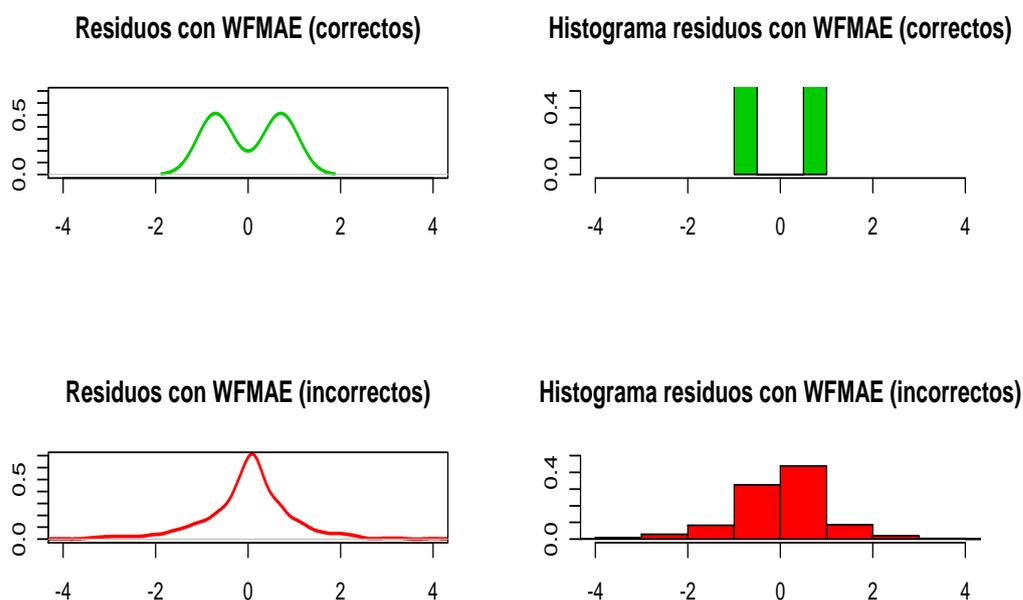


Figura 4.213: Residuos de las series correctas e incorrectas con el criterio WFMAE en el SARIMA(1, 1, 0)(1, 0, 1)₇.

	Media	Varianza	ECMG	ECMGpre
Correctos	-0.26	0.7403	0.4195	0.8078
Incorrectos	-0.1662	5.3817	4.6181	5.4093

Tabla 4.130: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 0)(1, 0, 1) $_7$ con el criterio WFMSE.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.0911	0.4059	0.1347	0.1576	0.5482	0.0034	4.3140
Incorrectos	0.9297	2.0399	2.3976	3.1711	5.8416	9.5625	13.9227

Tabla 4.131: ECM global de predicción en SARIMA(1, 1, 0)(1, 0, 1) $_7$ con el criterio WFMAE en el conjunto de comprobación.

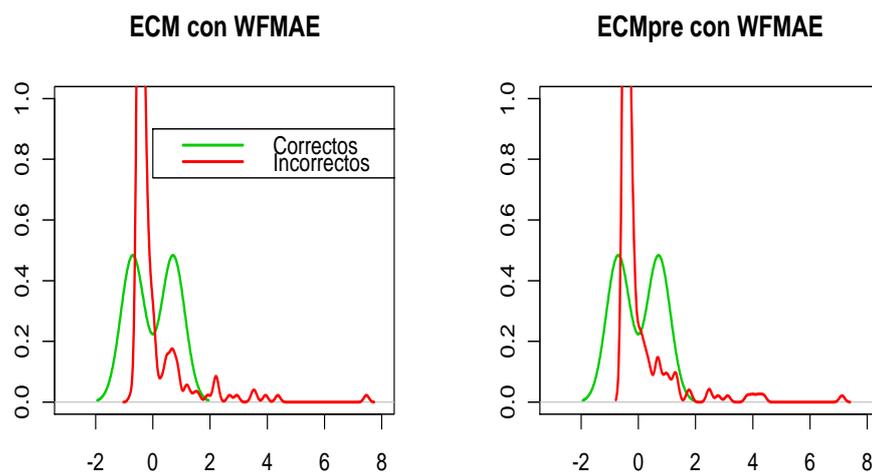


Figura 4.214: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en SARIMA(1, 1, 0)(1, 0, 1) $_7$.

Tampoco en esta ocasión se pueden obtener conclusiones al respecto de este procedimiento empleado, dado que no se detecta un comportamiento similar entre unos y otros y además el tamaño de muestra de los grupos de series bien identificadas son demasiado pequeños.

Continuamos con la última de las líneas de actuación determinadas a lo largo de este trabajo de investigación.

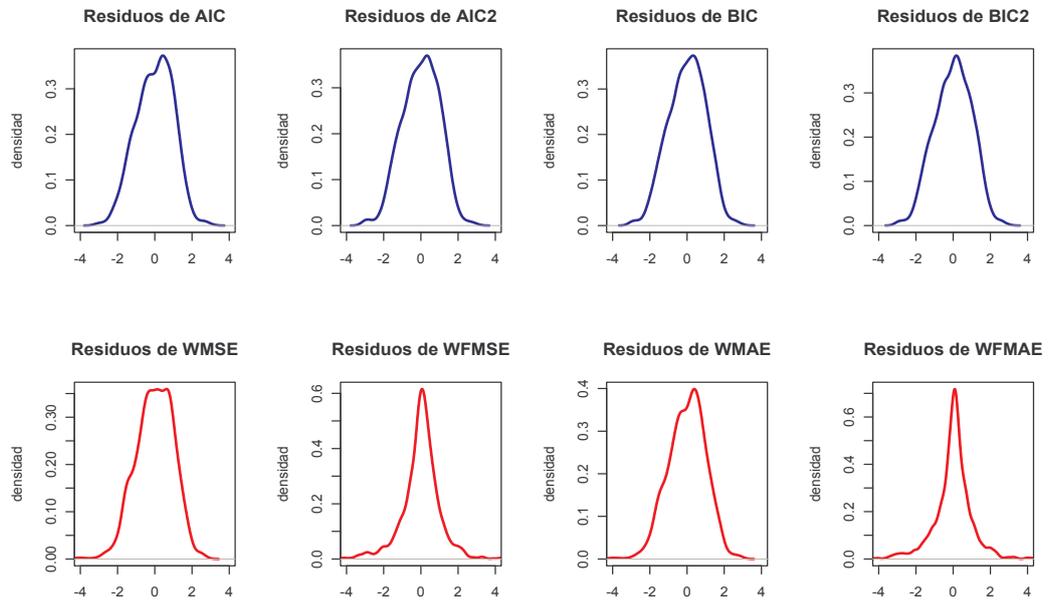


Figura 4.215: Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1)₇.

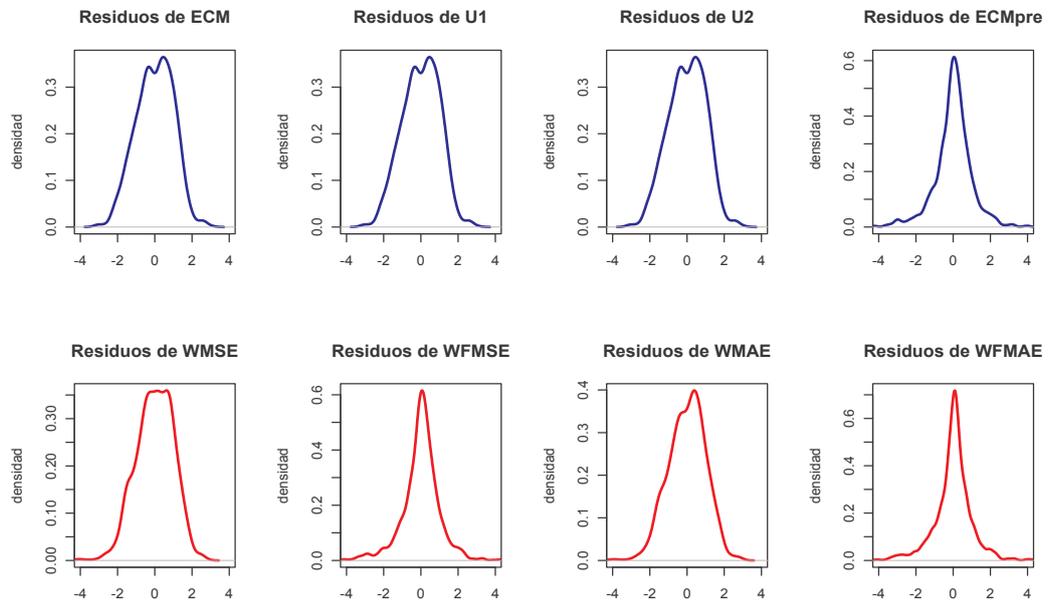


Figura 4.216: Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1)₇.

Criterio	Media	Varianza
AIC	-0.2605	18.9446
AIC2	-0.2979	19.0299
ECM	-0.3311	18.7215
BIC	-0.2554	18.3464
BIC2	-0.2809	18.1411
U1	-0.3311	18.7215
U2	-0.3311	18.7215
ECMpre	-0.1335	5.1728
WMSE	-0.2063	22.6872
WFMSE	-0.1271	5.1842
WMAE	-0.2702	22.3407
WFMAE	-0.1671	5.3354

Tabla 4.132: Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para SARIMA(1, 1, 0)(1, 0, 1) $_7$.

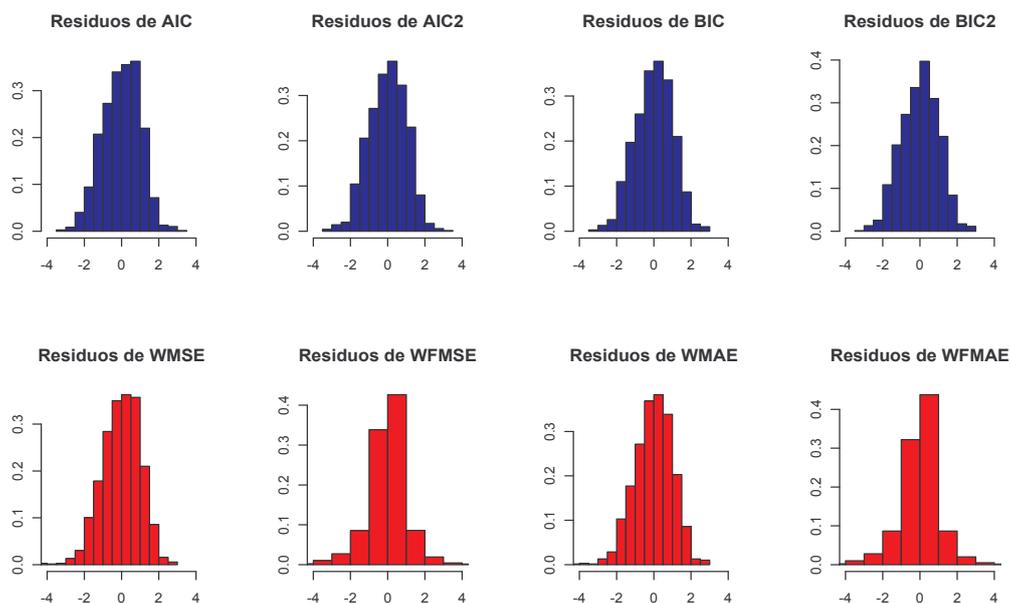


Figura 4.217: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1) $_7$.

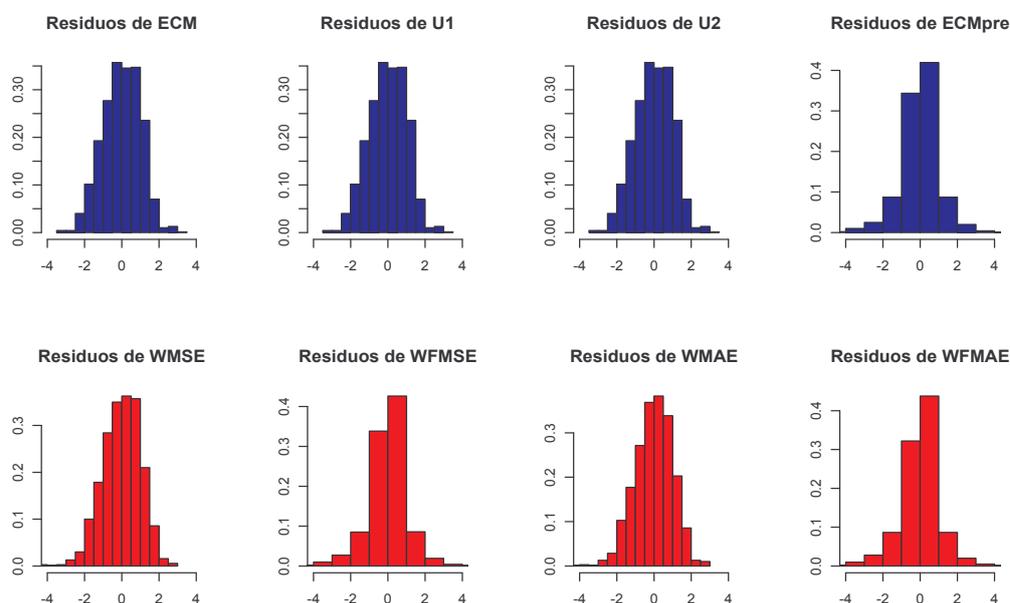


Figura 4.218: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1)₇.

Criterio	ECMpre1	ECMpre2	ECMpre3	ECMpre4	ECMpre5	ECMpre6	ECMpre7
AIC	1.1967	3.9583	8.2883	15.3499	24.0566	34.2516	45.9859
AIC2	1.1932	3.8639	8.1963	15.3213	24.1836	34.5164	46.5555
ECM	1.1770	3.9319	8.1812	15.1657	23.8417	33.9724	45.5480
BIC	1.1703	3.8585	8.0707	14.8884	23.3012	33.1702	44.4218
BIC2	1.1560	3.7901	7.9077	14.7820	23.2085	32.7189	43.9767
U1	1.1770	3.9319	8.1812	15.1657	23.8417	33.9724	45.5480
U2	1.1770	3.9319	8.1812	15.1657	23.8417	33.9724	45.5480
ECMpre	1.0435	2.1901	2.6709	3.2992	5.5735	8.7973	12.7598
WMSE	1.4474	5.1268	10.8009	18.8327	29.0453	40.6666	53.1885
WFMSE	0.9768	1.8798	2.3984	3.2123	5.7456	9.1133	13.0766
WMAE	1.4757	5.3998	10.9039	18.9488	28.7984	39.8136	51.5660
WFMAE	0.9213	2.0236	2.3749	3.1409	5.7887	9.4669	13.8266

Tabla 4.133: ECM globales para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 0)(1, 0, 1)₇.

En la gráfica 4.215 y 4.216 ya se ve en un primer golpe de vista la mejoría alcanzada en los criterios ECMpre, WFMSE y WFMAE, esto lo verificamos con los resultados de la tabla 4.132 donde se aprecia la disminución de la varianza en relación al resto de criterios y en relación a la tabla 4.122. Ante la tabla 4.133 se disipan todas las posibles dudas ante el éxito conseguido en esta ocasión por los criterios citados, se comprueba además que el incremento del error se da de manera más paulatina que en el resto de los criterios y que los resultados han mejorado también a los obtenidos en la tabla 4.123.

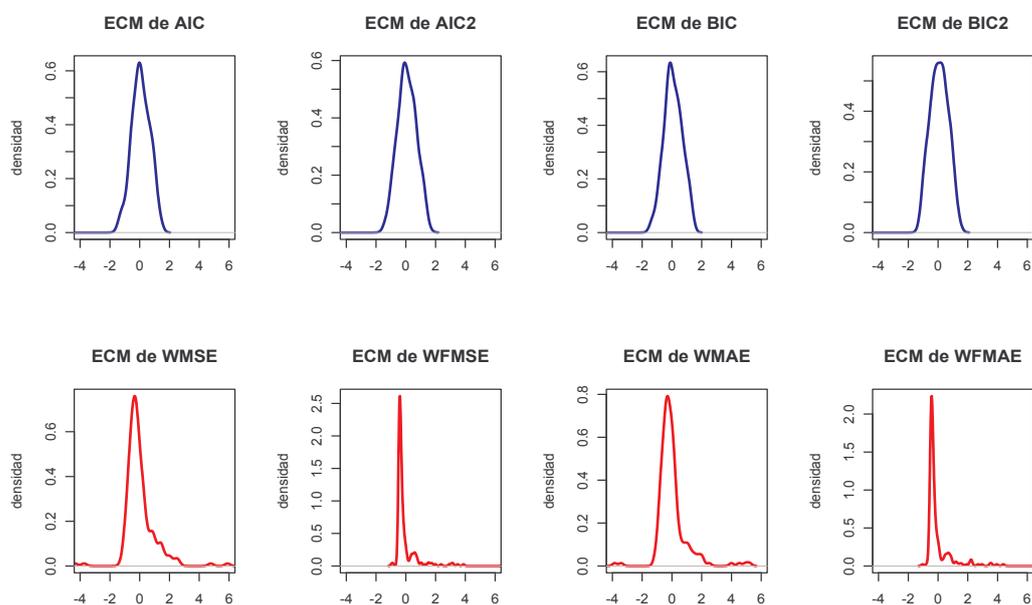


Figura 4.219: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1)₇.

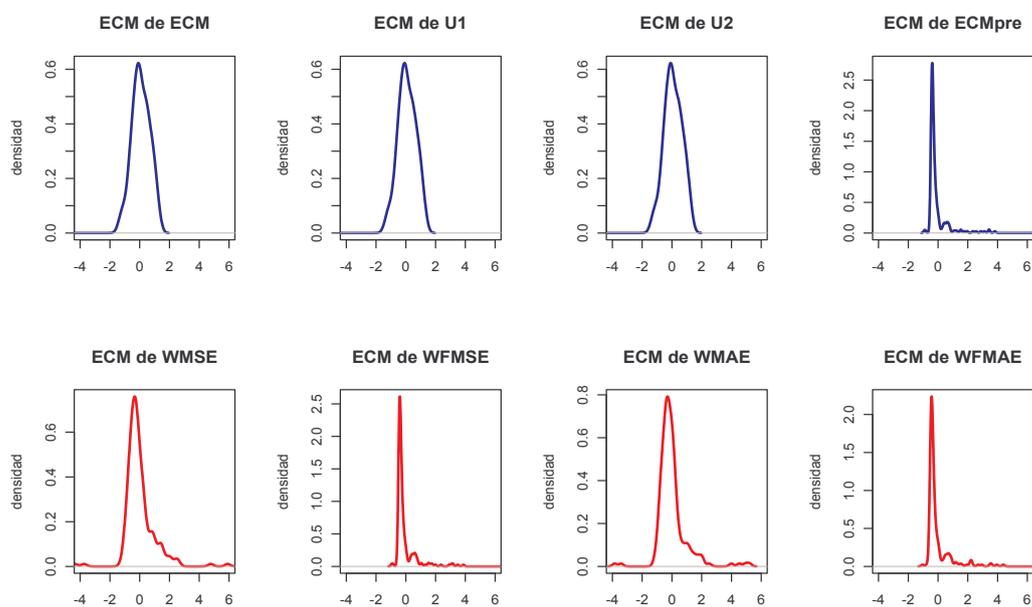


Figura 4.220: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1)₇.

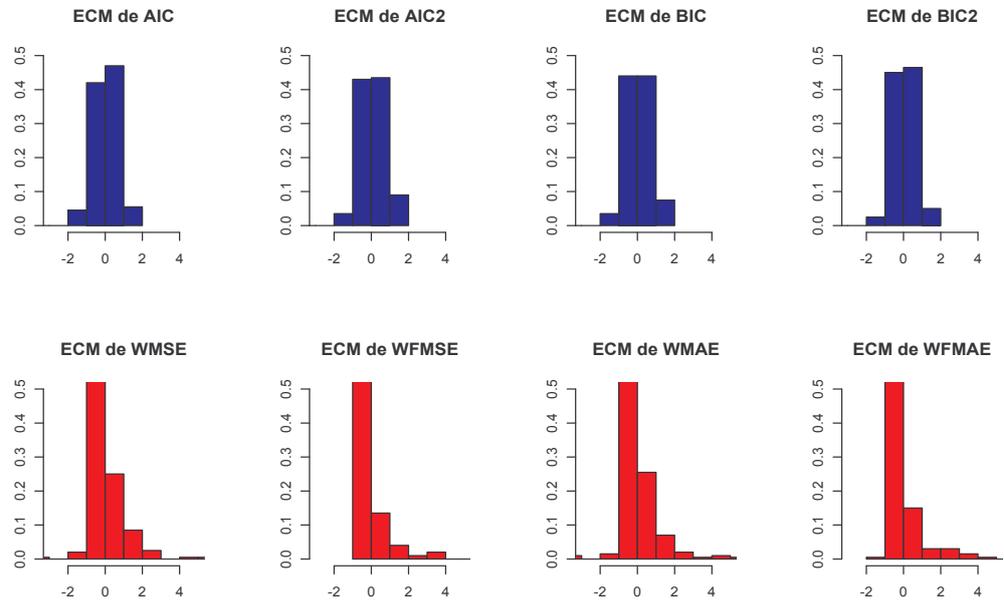


Figura 4.221: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 0)(1, 0, 1)_7$.

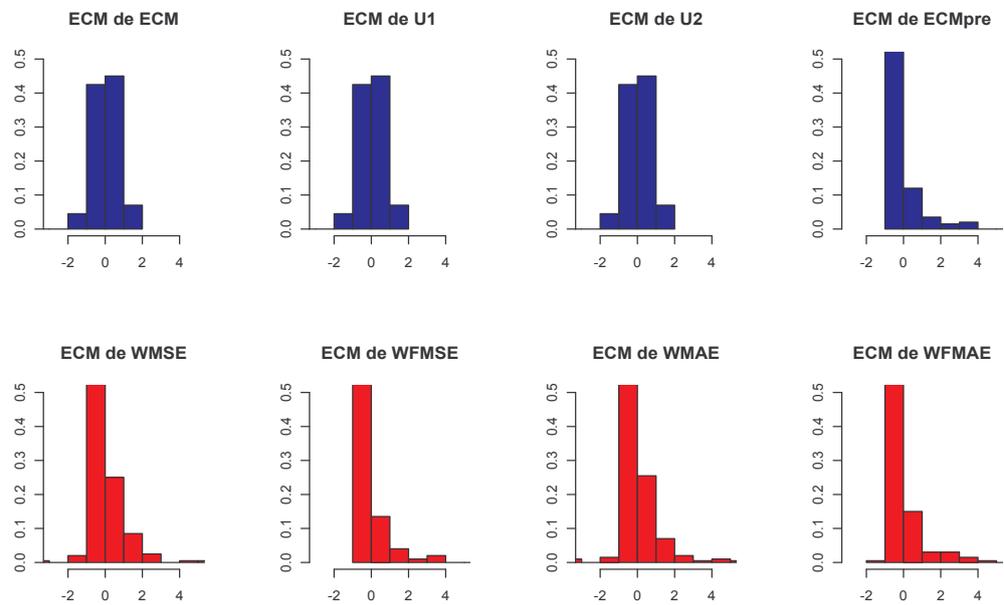


Figura 4.222: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 0)(1, 0, 1)_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8391	0.8667	0.8276	0.8662	0.9138	0.8276
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8276	5.2294	1.1220	5.1728	1.1861	4.5760

Tabla 4.134: ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 0)(1, 0, 1)₇.

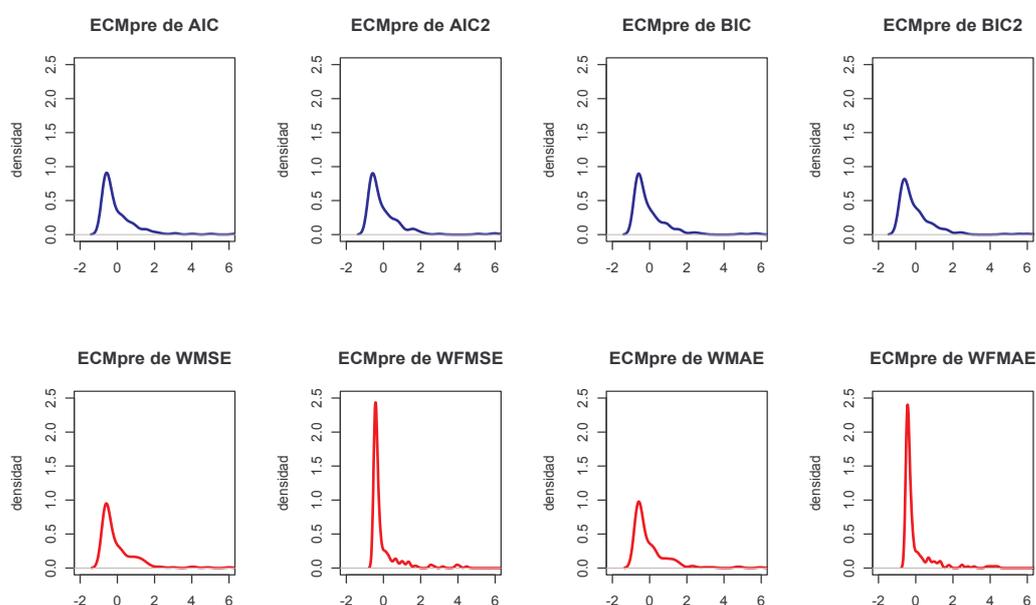


Figura 4.223: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1)₇.

En la tabla 4.134 se comprueba, además de que se puede ver también de forma evidente en los gráficos 4.129 y 4.220 como una vez más los resultados de ajuste en los criterios ECMpre, WFMSE y WFMAE, sobre todos, se elevan significativamente. Estos resultados también se ven incrementados al compararlos con los mostrados en la tabla 4.124, lo que nos lleva una vez más a ver que existe un comportamiento inverso entre el ajuste y la predicción.

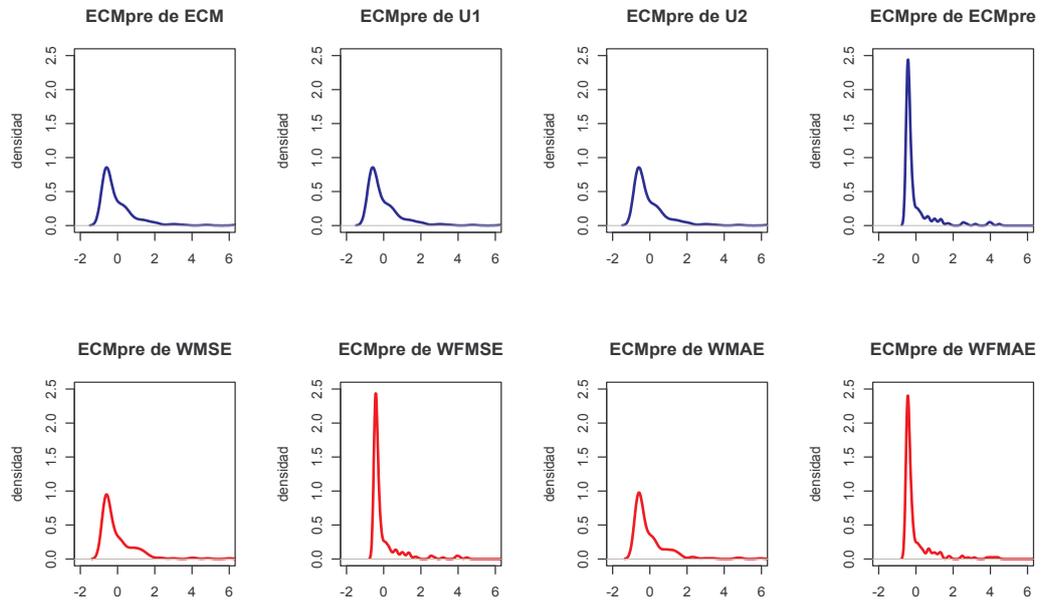


Figura 4.224: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 0)(1, 0, 1)_7$.

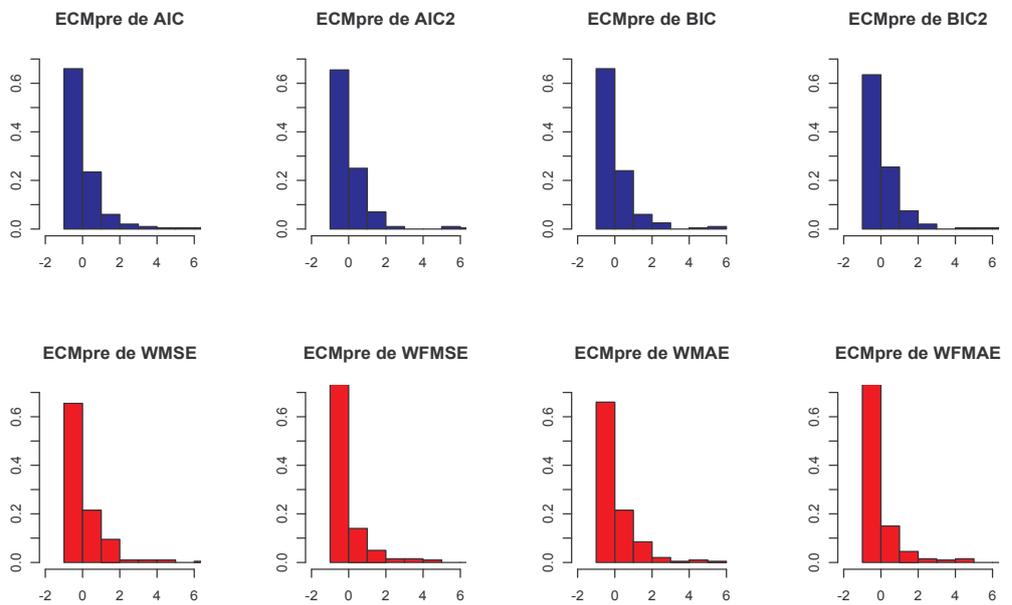


Figura 4.225: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 0)(1, 0, 1)_7$.

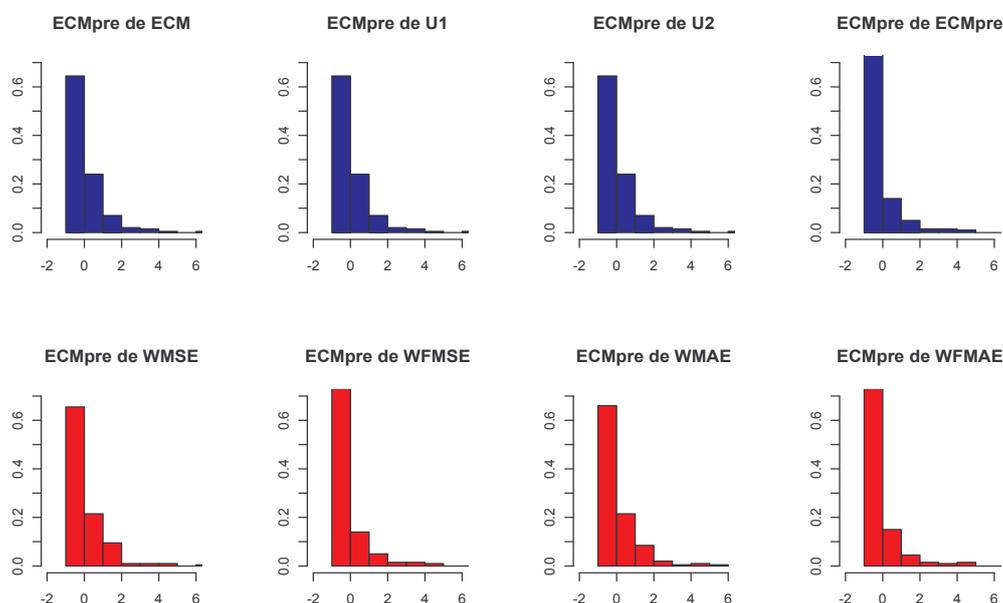


Figura 4.226: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 0)(1, 0, 1)₇.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	19.0125	19.1186	18.8311	18.4116	18.22	18.8311
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	18.8311	5.1906	22.7297	5.2004	22.4137	5.3633

Tabla 4.135: ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 0)(1, 0, 1)₇.

Para terminar con este modelo, vemos en la tabla 4.135 la mejoría experimentada en los criterios citados, una vez más aun que el menor ECMGpre se consigue con el criterio del ECMpre, los criterios de predicción, WFMSE y WFMAE, son preferibles, ya que complementan la predicción global con las predicciones individuales (tabla 4.133) donde se ve como es en estos donde las primeras de las predicciones son más precisas en sus valores. En cuanto a la comparación con la tabla 4.125, se ve también en ese caso la mejoría que se alcanza ante este procedimiento de actuación.

Modelo SARIMA(1, 0, 1)(0, 0, 1)₇, cuyos coeficientes son $\phi_1 = 0,1$, $\theta_1 = 0,6$ y $\Theta = -0,8$. (Modelo 128)

	AIC	AIC2	ECM	BIC	BIC2
Valor	2.424756	-0.3318941	0.617247	2.534118	-0.2761620
Modelo	195	147	195	195	20

Tabla 4.136: Criterios Clásicos de ajuste para SARIMA(1, 0, 1)(0, 0, 1)₇.

	U_1	U_2	ECMpre
Valor	0.4799237	0.5470398	0.0637416
Modelo	267	267	303

Tabla 4.137: Criterios Clásicos de predicción para SARIMA(1, 0, 1)(0, 0, 1)₇.

	Ajuste		Predicción	
	WMSE	WMAE	WFMSE	WFMSE
Valor	0.1351331	0.2512618	0.06116104	0.1743554
Modelo	35	240	303	11

Tabla 4.138: Criterios propuestos para SARIMA(1, 0, 1)(0, 0, 1)₇.

En la tabla 4.136 se puede comprobar como los criterios AIC, ECM y BIC se decantan por el mismo modelo, el 195 (ver apéndice C), En la tabla 4.137 se observa como los criterios de Theil seleccionan el modelo 267 (ver apéndice C) y junto con la tabla 4.138 la elección se empata también ante el modelo 303 (ver apéndice C) para los criterios ECMpre y WFMSE.

A continuación y siguiendo la misma línea que en el resto de modelos, estudiamos gráficamente la distribución de los residuos en el conjunto de comprobación que se han obtenido al aplicar los modelos que se recogen en las tablas 4.136 a 4.138 para realizar las predicciones. A la vista de los gráficos 4.227 a 4.230 no se ven diferencias evidentes entre los criterios, si miramos en la tabla 4.139 se corrobora que la distribución de los residuos en cada uno de los modelos sigue un comportamiento similar.

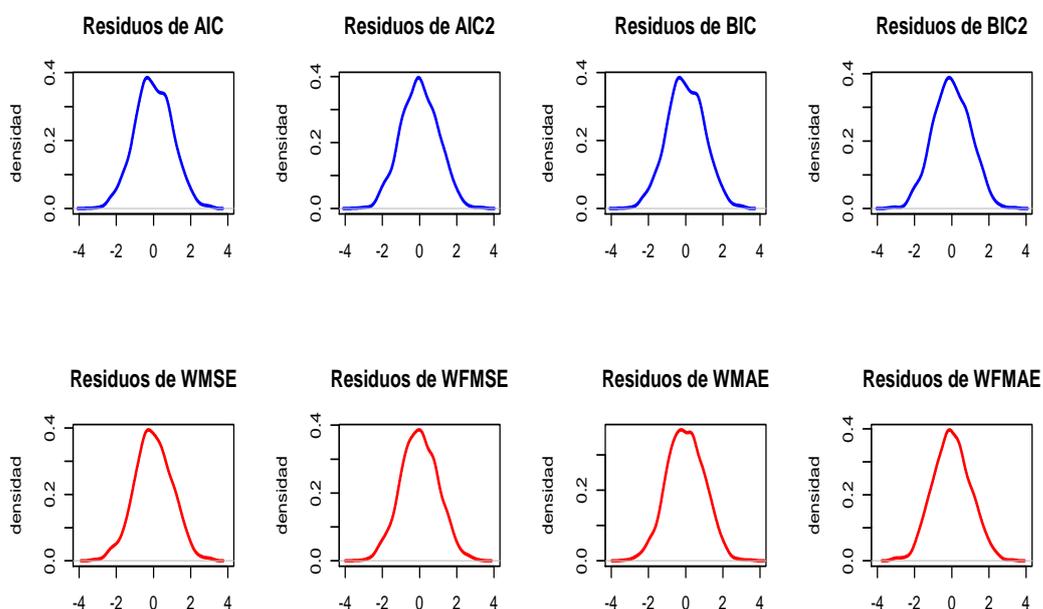


Figura 4.227: Residuos obtenidos con los modelos de los criterios de ajuste de $SARIMA(1, 0, 1)(0, 0, 1)_7$.

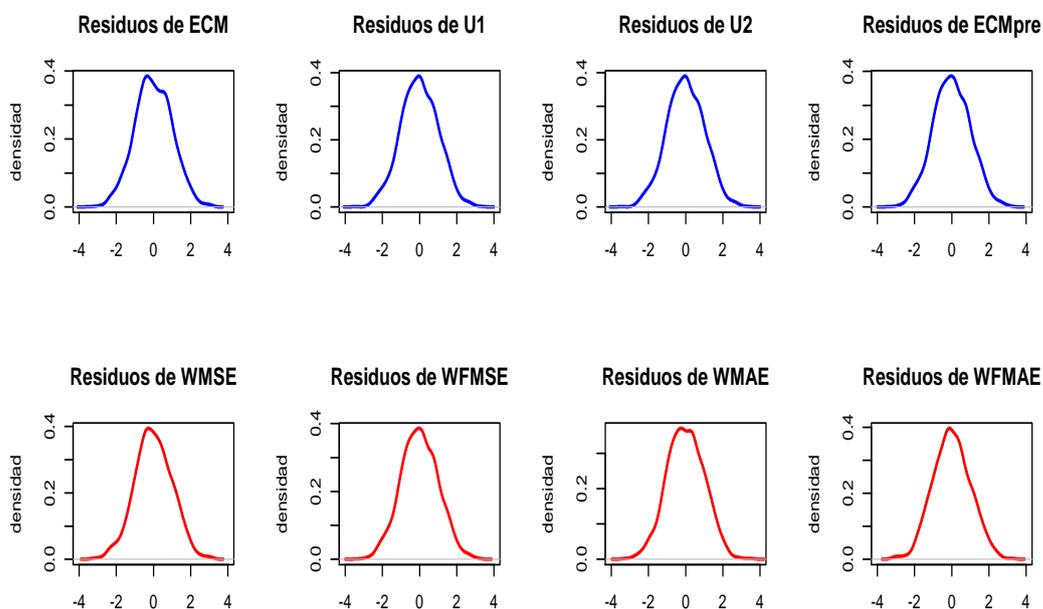


Figura 4.228: Residuos obtenidos con los modelos de los criterios de predicci3n de $SARIMA(1, 0, 1)(0, 0, 1)_7$.

Criterio	Media	Varianza
AIC	-0.0025	1.2654
AIC2	-0.0014	1.2337
ECM	-0.0025	1.2654
BIC	-0.0025	1.2654
BIC2	-0.0022	1.2305
U1	0.0026	1.2655
U2	0.0026	1.2655
ECMpre	-0.0054	1.2707
WMSE	-0.0145	1.4432
WFMSE	-0.0054	1.2707
WMAE	-0.0050	1.2911
WFMAE	-0.0126	1.8071

Tabla 4.139: Medidas descriptivas de los residuos en el conjunto de comprobación para SARIMA(1, 0, 1)(0, 0, 1)₇.

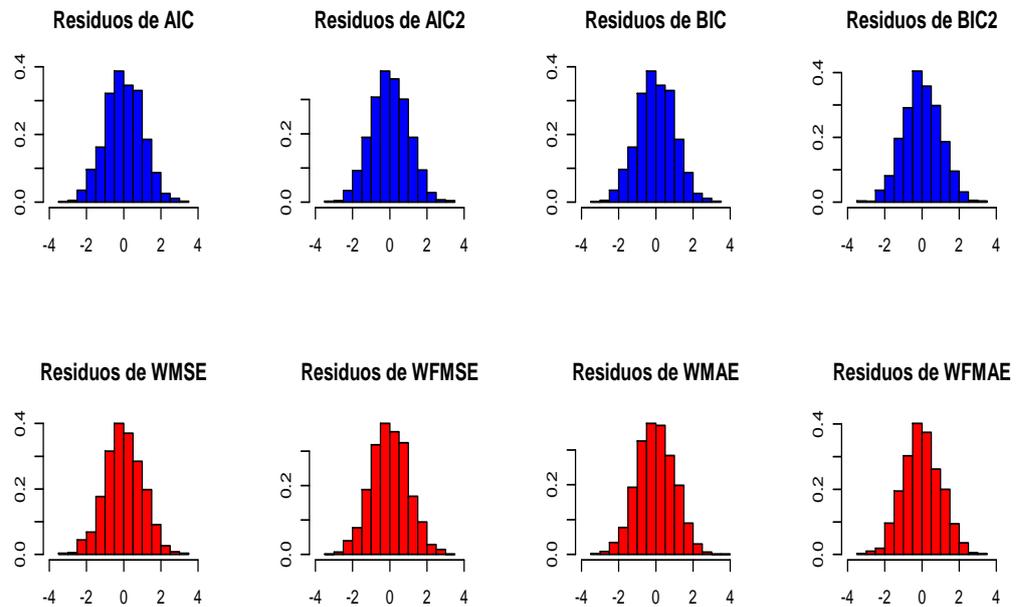


Figura 4.229: Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 0, 1)(0, 0, 1)₇.

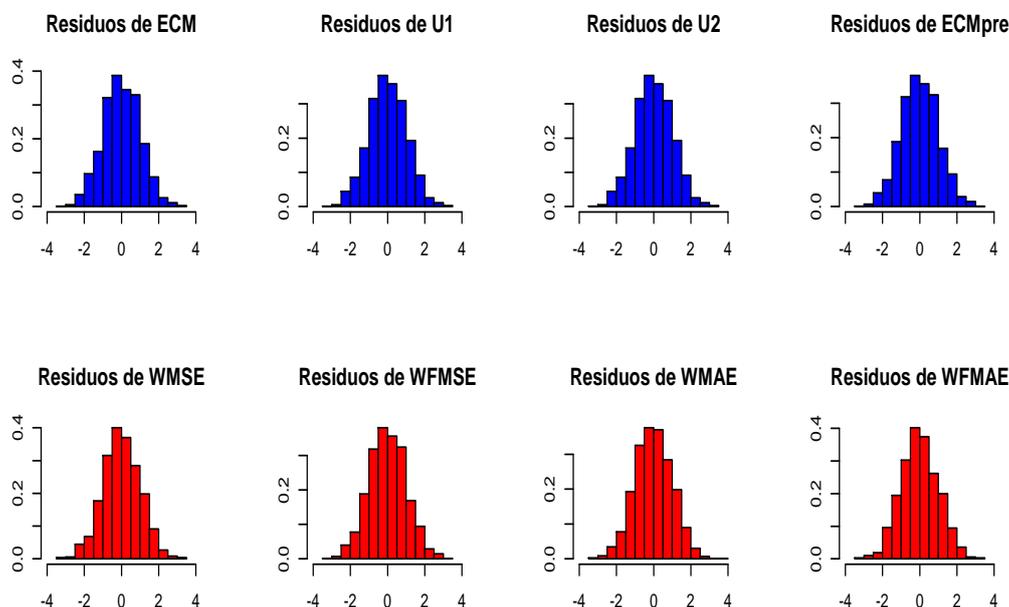


Figura 4.230: Histograma de los residuos obtenidos con los modelos de los criterios de predicción de SARIMA($1, 0, 1$)($0, 0, 1$) $_7$.

Criterio	ECMpre1	ECMpre2	ECMpre3	ECMpre4	ECMpre5	ECMpre6	ECMpre7
AIC	1.1418	1.2702	1.1041	1.2508	1.4730	1.3562	1.2619
AIC2	1.0326	1.1737	1.1224	1.3036	1.5081	1.3146	1.1812
ECM	1.1418	1.2702	1.1041	1.2508	1.4730	1.3562	1.2619
BIC	1.1418	1.2702	1.1041	1.2508	1.4730	1.3562	1.2619
BIC2	1.0378	1.1372	1.1074	1.2827	1.5109	1.3238	1.2137
U1	1.0605	1.2673	1.1555	1.2795	1.4970	1.3392	1.2596
U2	1.0605	1.2673	1.1555	1.2795	1.4970	1.3392	1.2596
ECMpre	1.0910	1.2768	1.1819	1.2501	1.5319	1.3299	1.2336
WMSE	1.3037	1.4086	1.3133	1.3714	1.7450	1.5577	1.4041
WFMSE	1.0910	1.2768	1.1819	1.2501	1.5319	1.3299	1.2336
WMAE	1.1822	1.2213	1.1165	1.3096	1.5457	1.4179	1.2449
WFMAE	1.8622	1.5669	1.8103	1.7331	2.0056	2.0191	1.6536

Tabla 4.140: ECMpre global para cada predicción realizada en el conjunto de comprobación en SARIMA($1, 0, 1$)($0, 0, 1$) $_7$.

En la tabla superior, 4.140, se comprueba de igual manera el comportamiento seguido por los ECMGpre en cada una de las predicciones realizadas y se ve una vez más la falta de diferenciación entre ellos, no obstante en este caso se observa como el aumento de los errores a lo largo del horizonte de predicción se mantiene constante, al contrario que ocurría en casos anteriores donde estos errores aumentaban rápidamente.

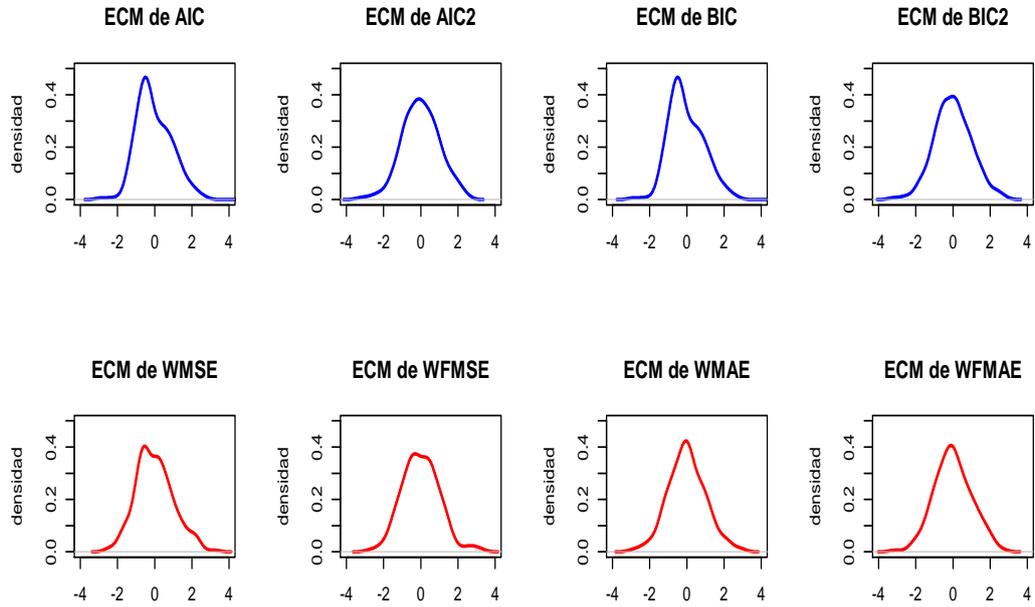


Figura 4.231: ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 0, 1)(0, 0, 1)₇.

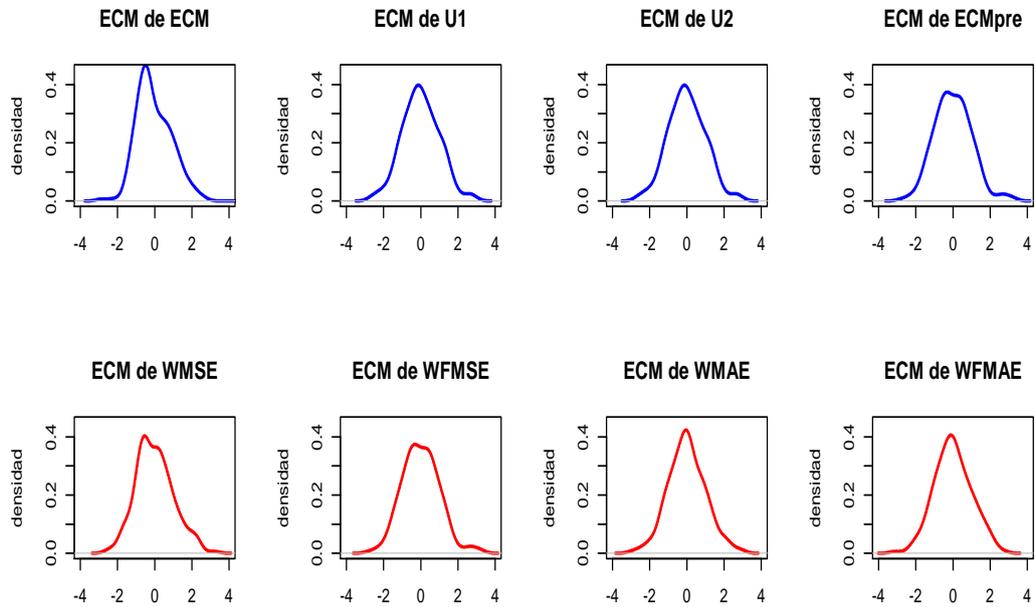


Figura 4.232: ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 0, 1)(0, 0, 1)₇.

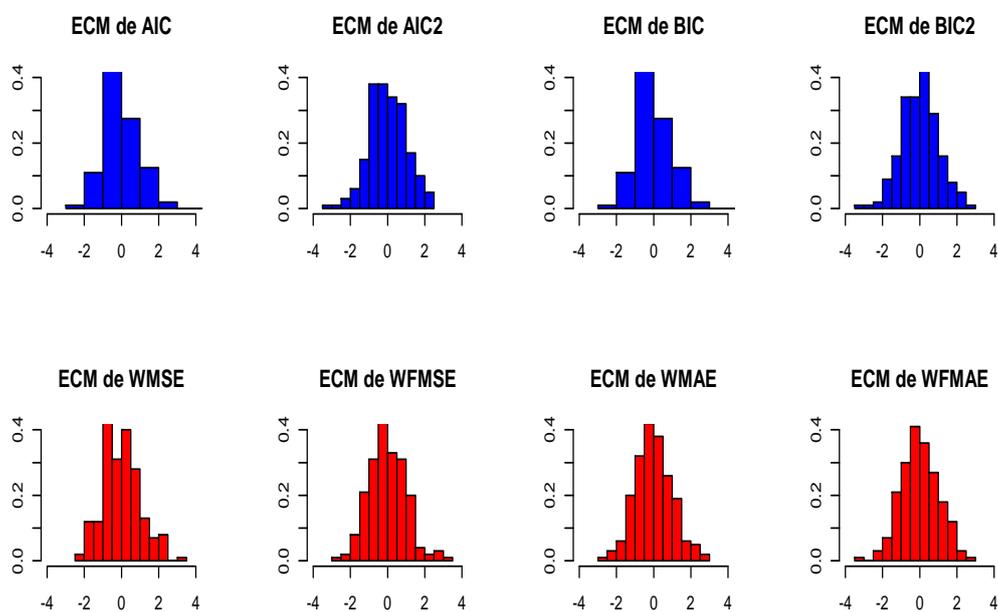


Figura 4.233: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 0, 1)(0, 0, 1)₇.

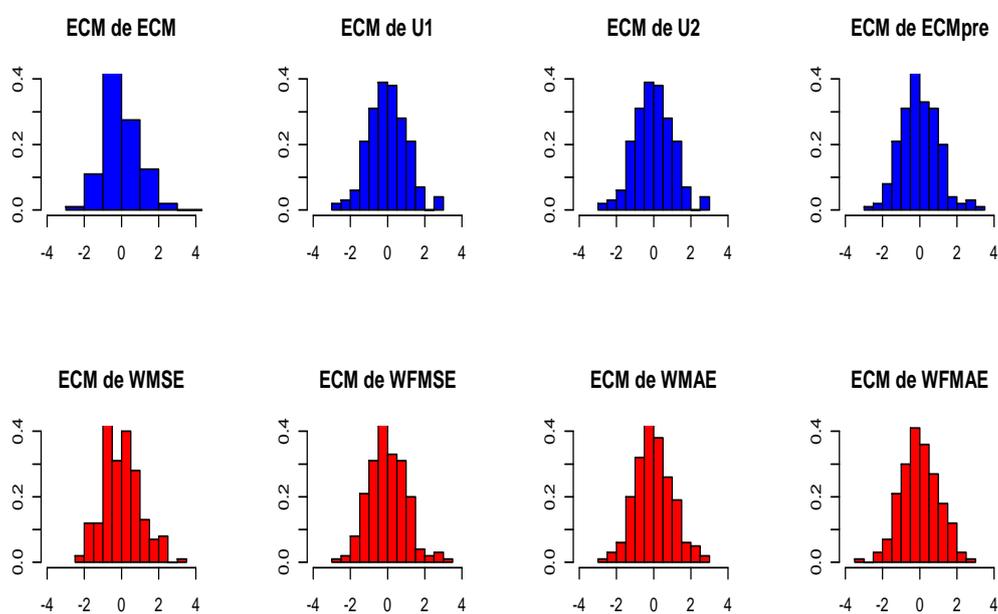


Figura 4.234: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicci3n en SARIMA(1, 0, 1)(0, 0, 1)₇.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.9987	1.0179	0.9987	0.9987	1.0597	0.8730
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8730	0.9409	1.0596	0.9409	1.0750	1.7129

Tabla 4.141: El ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 0, 1)(0, 0, 1)₇.

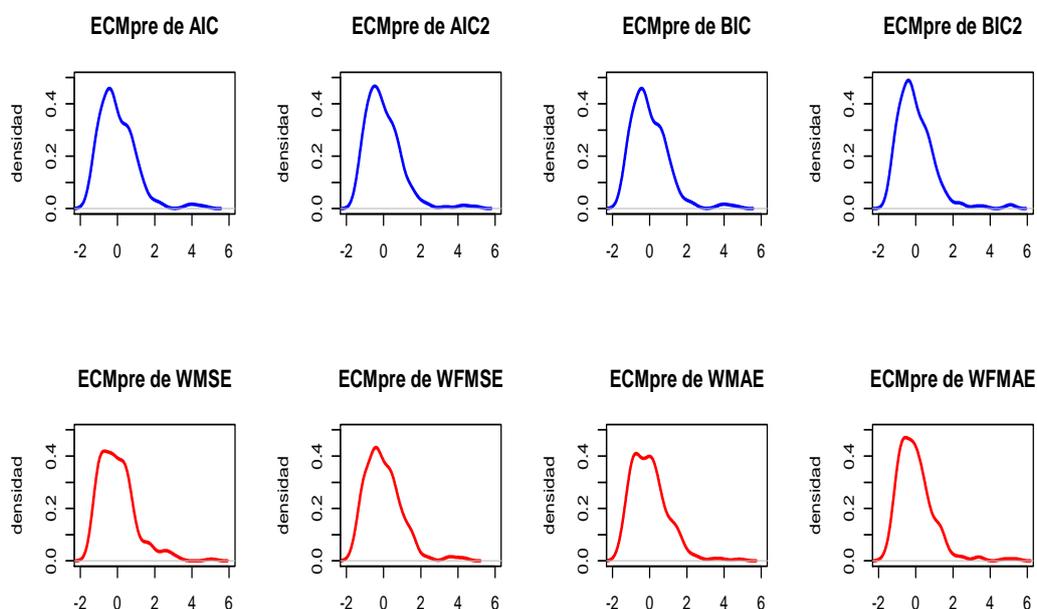


Figura 4.235: ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 0, 1)(0, 0, 1)₇.

En la tabla 4.141 se comprueba también que en esta ocasión los errores de ajuste tampoco experimentan grandes cambios entre los criterios, incluyendo en estos los criterios ECMpre, WMSE, WFMSE, WMAE y WFMAE, donde se alcanzaban peores resultados en los modelos anteriores, en esta ocasión el comportamiento es más similar en todos los casos.

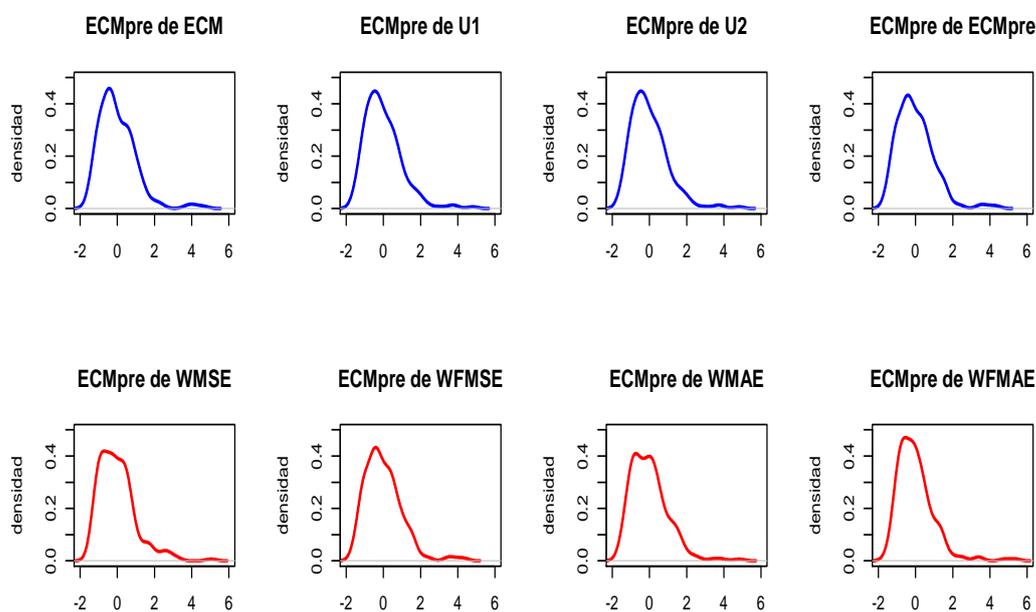


Figura 4.236: ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 0, 1)(0, 0, 1)₇.

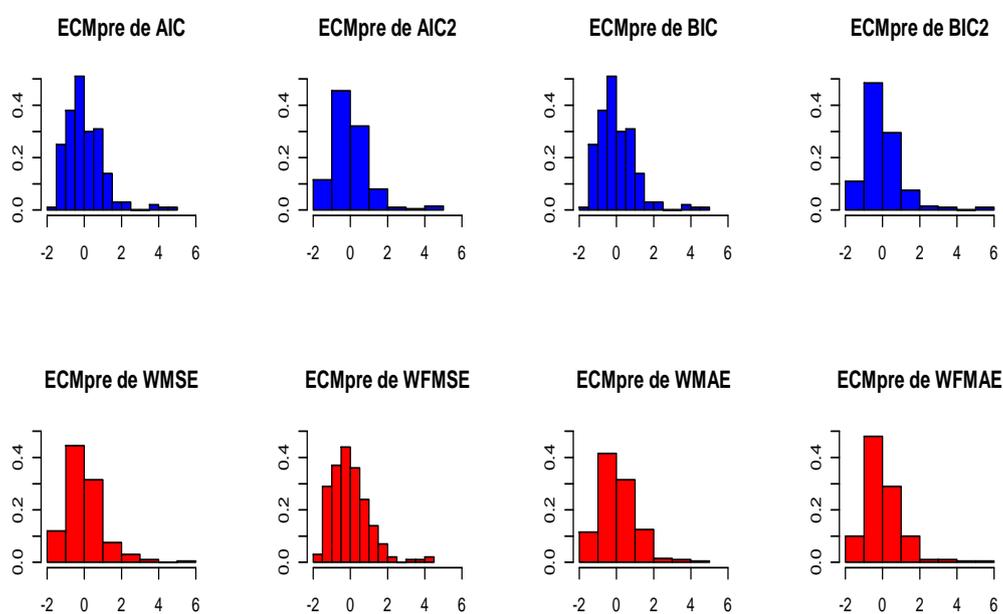


Figura 4.237: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 0, 1)(0, 0, 1)₇.

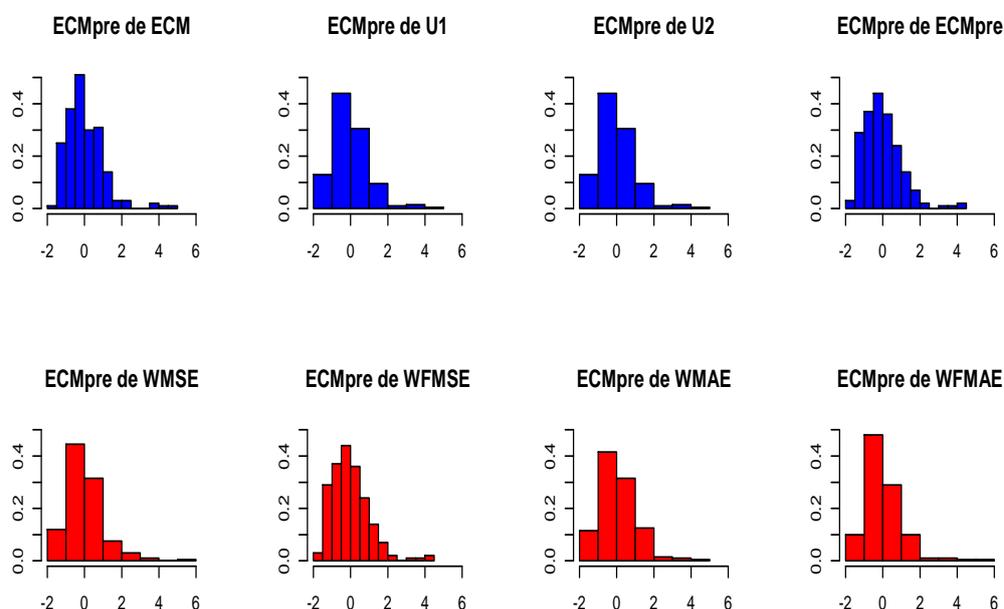


Figura 4.238: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en $SARIMA(1, 0, 1)(0, 0, 1)_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	1.2654	1.2337	1.2654	1.2654	1.2305	1.2655
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	1.2655	1.2707	1.4434	1.2707	1.2911	1.8072

Tabla 4.142: ECM global de predicción, ECMGpre, para cada criterio en modelos $SARIMA(1, 0, 1)(0, 0, 1)_7$.

En la tabla 4.142 los resultados sobre predicción también se mantienen constantes a lo largo de los criterios, por lo que en cuestión de este modelo, el procedimiento no discrimina en ningún sentido entre los criterios ni en predicción, ni en ajuste.

En esta ocasión para llevar a cabo el estudio de las series bien determinadas y las que no, se han estudiado los criterios AIC2, BIC2 y WMSE, al ser éstos los únicos capaces de identificar a algunas de las 200 series estudiadas, en el caso del AIC2 han sido 4 las series correctas, en el caso del criterio BIC2 son 5 las series detectadas y 3 con el criterio WMSE, una vez más se mostrarán los resultados obtenidos al respecto, con el cuidado del bajo porcentaje de acierto que se ha alcanzado.

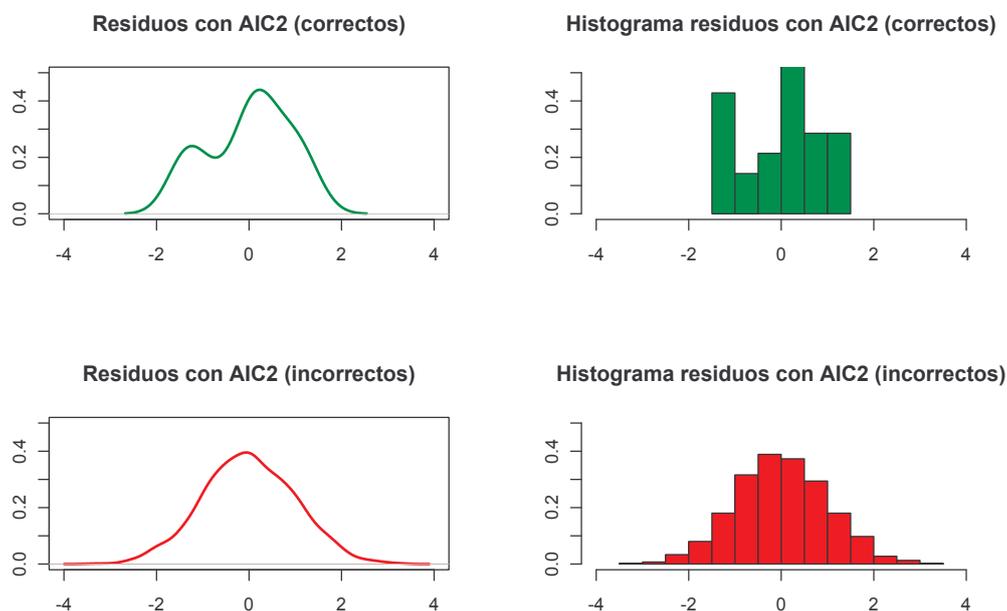


Figura 4.239: Residuos de las series correctas e incorrectas con el criterio AIC2 en el SARIMA(1, 0, 1)(0, 0, 1) $_7$.

	Media	Varianza	ECMG	ECMGpre
Correctos	0.0103	0.8912	0.9399	0.8913
Incorrectos	0.0052	1.2665	0.8942	1.2665

Tabla 4.143: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 0, 1)(0, 0, 1) $_7$ con el criterio AIC2.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.6258	0.5151	0.4941	1.3930	0.8050	0.9091	1.4970
Incorrectos	1.0839	1.2172	1.1950	1.2842	1.4978	1.3768	1.2108

Tabla 4.144: ECM global de predicción en SARIMA(1, 0, 1)(0, 0, 1) $_7$ con el criterio AIC2 en el conjunto de comprobación.

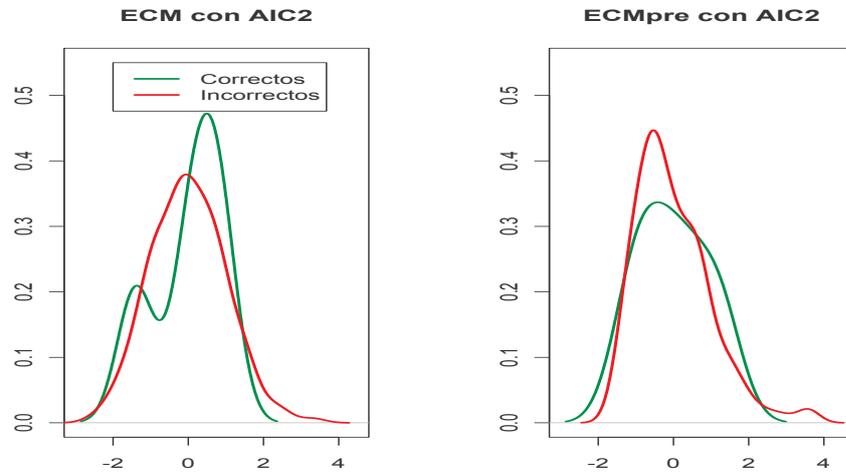


Figura 4.240: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio AIC2 en $\text{SARIMA}(1, 0, 1)(0, 0, 1)_7$.

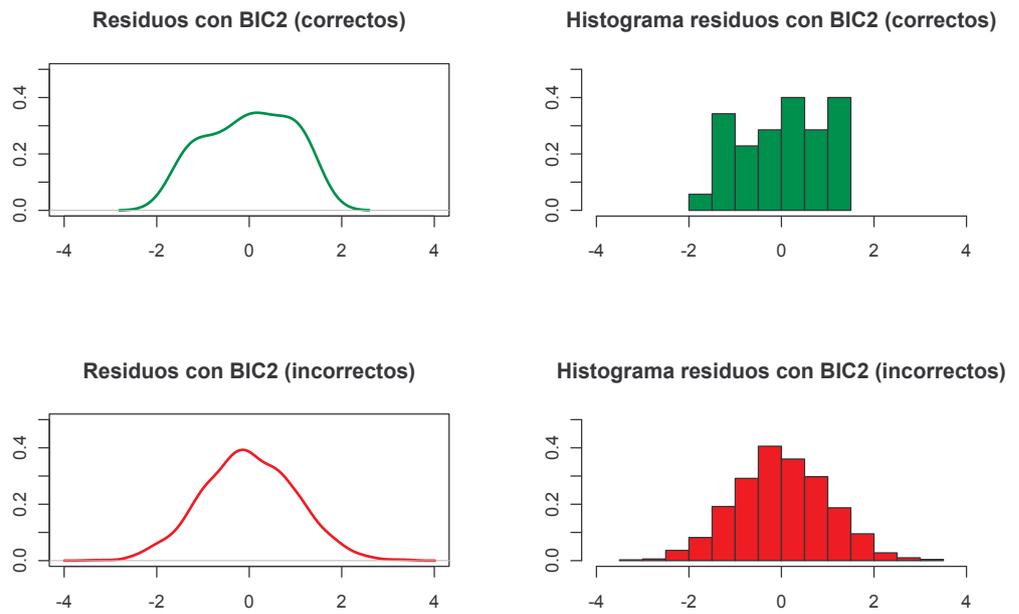


Figura 4.241: Residuos de las series correctas e incorrectas con el criterio BIC2 en el $\text{SARIMA}(1, 0, 1)(0, 0, 1)_7$.

	Media	Varianza	ECMG	ECMGpre
Correctos	0.0679	0.8907	0.9004	0.8953
Incorrectos	0.0029	1.2525	0.9492	1.2525

Tabla 4.145: Medidas descriptivas de los residuos del conjunto de comprobación y ECMG de ajuste y predicción global para SARIMA(1, 0, 1)(0, 0, 1)₇ con el criterio BIC2.

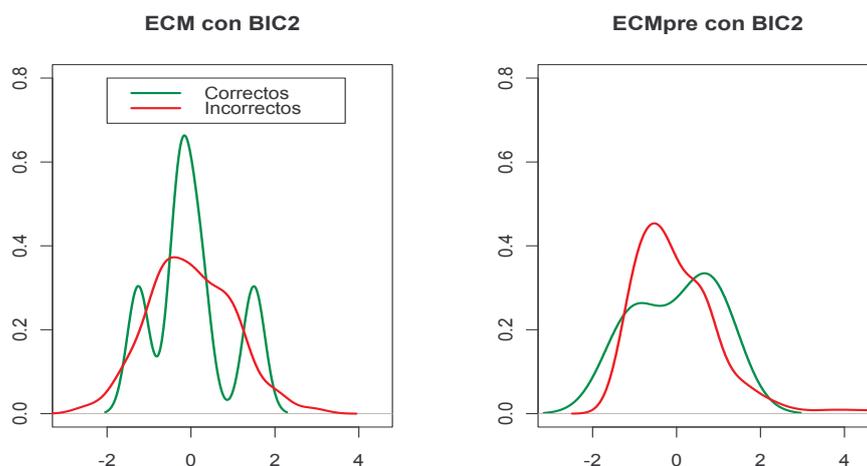


Figura 4.242: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio BIC2 en SARIMA(1, 0, 1)(0, 0, 1)₇.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.7184	0.9902	0.4770	0.7757	0.9765	0.9857	1.3238
Incorrectos	1.0579	1.1807	1.1580	1.2925	1.4920	1.3780	1.2087

Tabla 4.146: ECM global de predicción en SARIMA(1, 0, 1)(0, 0, 1)₇ con el criterio BIC2 en el conjunto de comprobación.

En la gráfica 4.239 se observa que la mejor distribución de residuos se contempla en el grupo de los incorrectos, hecho que se puede verificar con la tabla 4.143 donde se contempla que la media es prácticamente cero, no obstante la varianza es mayor, aunque se consigue un mejor ajuste (gráfico 4.240) y peores resultados en predicción, se ve en la tabla 4.144 como es en las series correctas donde se determinan los menores errores.

En el caso del criterio BIC2 el comportamiento es muy parecido al anterior como puede verse en el gráfico 4.241, no obstante en la tabla 4.145 se aprecia en este caso un peor valor también en cuanto al ajuste, sin embargo a la vista de la gráfica 4.242 la curva “mejor” es la de las series incorrectas. En la tabla 4.146 el comportamiento es el mismo que el descrito en el caso del

criterio AIC2.

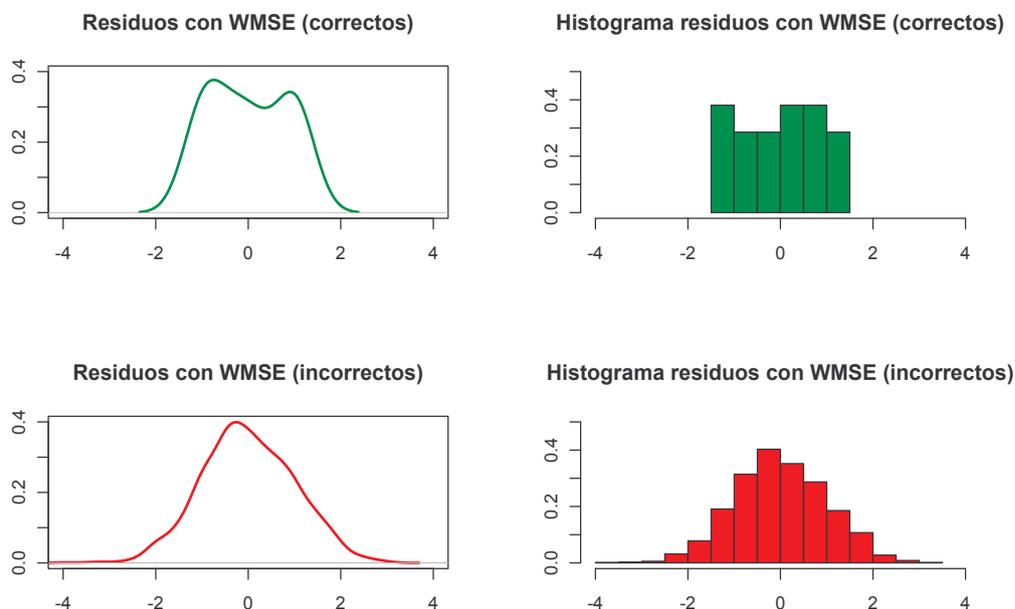


Figura 4.243: Residuos de las series correctas e incorrectas con el criterio WMSE en el SARIMA(1,0,1)(0,0,1)₇.

	Media	Varianza	ECMG	ECMGpre
Correctos	0.1586	0.7361	0.9465	0.7613
Incorrectos	-0.0071	1.3381	0.9291	1.3381

Tabla 4.147: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1,0,1)(0,0,1)₇ con el criterio WMSE.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	0.3229	0.2523	0.7698	1.1935	1.2506	0.8070	0.7328
Incorrectos	1.1775	1.3010	1.2955	1.3162	1.6062	1.4218	1.2489

Tabla 4.148: ECM global de predicción en SARIMA(1,0,1)(0,0,1)₇ con el criterio WMSE en el conjunto de comprobación.

En referencia al criterio WMSE se observa en el gráfico 4.243 el mismo comportamiento descrito para los criterios anteriores, al igual que en la tabla 4.147 se aprecian resultados similares a los obtenidos con el criterio AIC2,

sin muchas diferencias en cuanto a ajuste como se puede ver gráficamente (gráfico 4.244), con respecto a la tabla 4.148 las mejores predicciones se alcanzan en el grupo de series correctas.

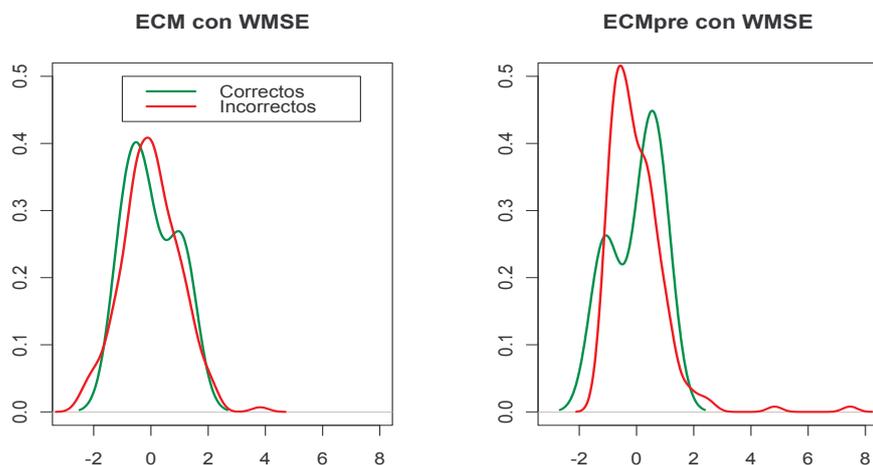


Figura 4.244: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WMSE en SARIMA(1, 0, 1)(0, 0, 1) $_7$.

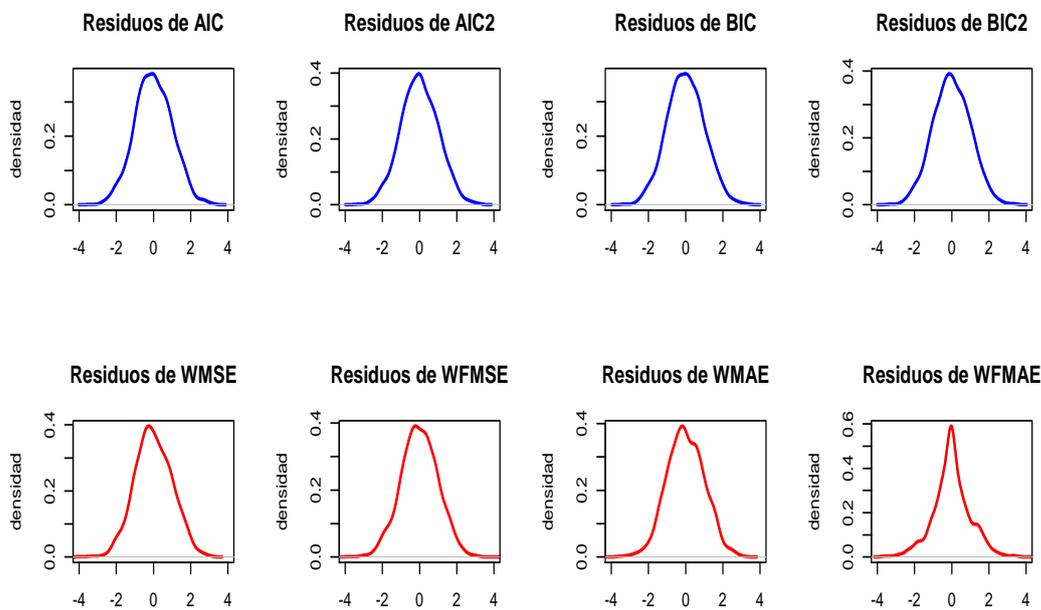


Figura 4.245: Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1) $_7$.

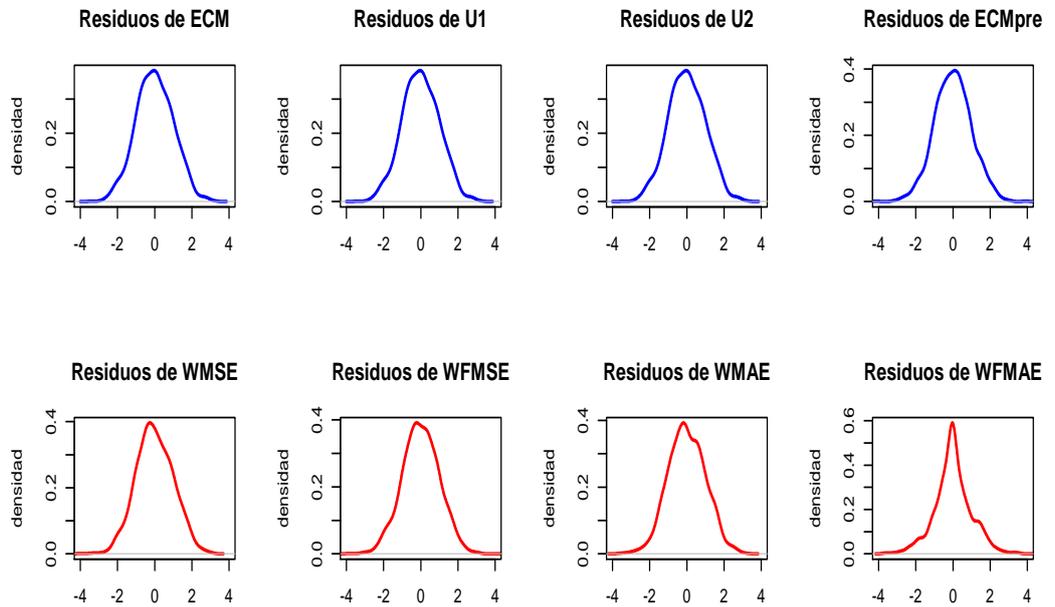


Figura 4.246: Residuos obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 0, 1)(0, 0, 1)_7$.

Criterio	Media	Varianza
AIC	0.0031	1.2802
AIC2	0.0053	1.2590
ECM	0.0036	1.2782
BIC	0.0029	1.2625
BIC2	0.0045	1.2436
U1	0.0036	1.2782
U2	0.0036	1.2782
ECMpre	0.0162	0.7696
WMSE	-0.0046	1.3295
WFMSE	-0.0019	0.7949
WMAE	-0.0361	1.5713
WFMAE	-0.0031	0.8795

Tabla 4.149: Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para $SARIMA(1, 0, 1)(0, 0, 1)_7$.

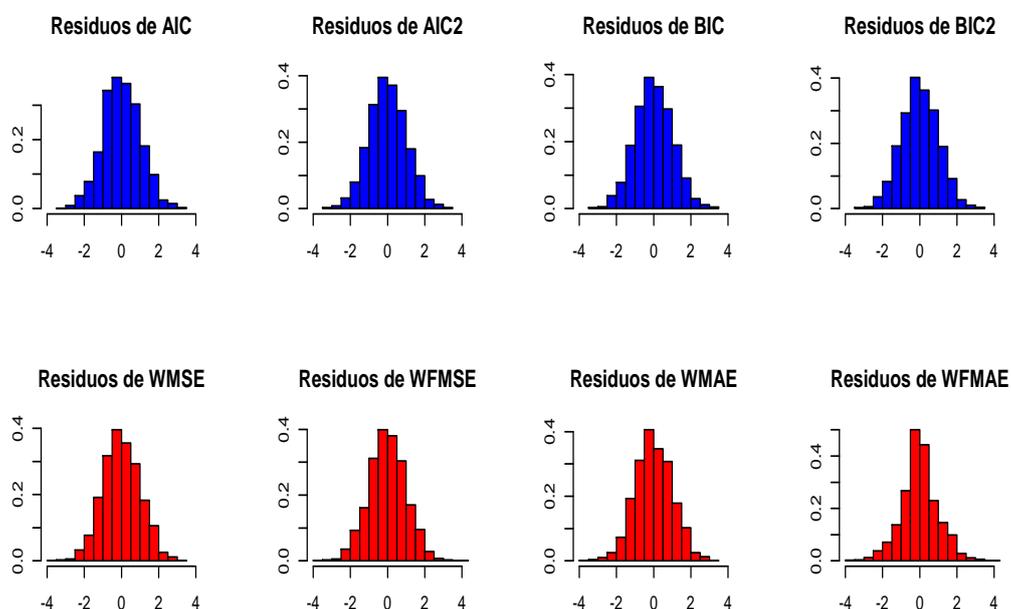


Figura 4.247: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1,0,1)(0,0,1)₇.

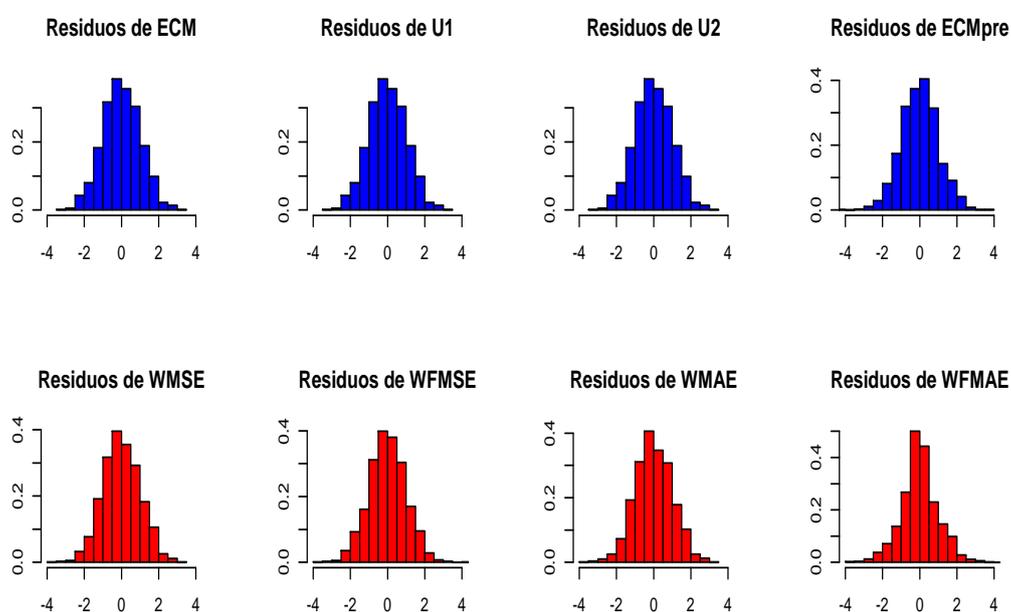


Figura 4.248: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1,0,1)(0,0,1)₇.

Criterio	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
AIC	1.1206	1.2380	1.1741	1.2916	1.4965	1.4206	1.2201
AIC2	1.0747	1.2031	1.1810	1.2864	1.4840	1.3674	1.2166
ECM	1.0923	1.2761	1.1664	1.2971	1.4808	1.4069	1.2277
BIC	1.0393	1.2332	1.1676	1.2481	1.5126	1.4064	1.2304
BIC2	1.0494	1.1760	1.1410	1.2796	1.4796	1.3682	1.2115
U1	1.0923	1.2761	1.1664	1.2971	1.4808	1.4069	1.2277
U2	1.0923	1.2761	1.1664	1.2971	1.4808	1.4069	1.2277
ECMpre	0.5454	0.7401	0.6957	0.7163	0.9606	0.8759	0.8554
WMSE	1.1647	1.2852	1.2876	1.3143	1.6009	1.4126	1.2412
WFMSE	0.4202	0.5648	0.6588	0.7373	1.1149	1.1028	0.9697
WMAE	1.4642	1.3858	1.3176	1.4487	1.9650	1.8795	1.5476
WFMAE	0.3678	0.5518	0.8015	0.8974	1.2677	1.1943	1.0758

Tabla 4.150: ECM globales para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 0, 1)(0, 0, 1)₇.

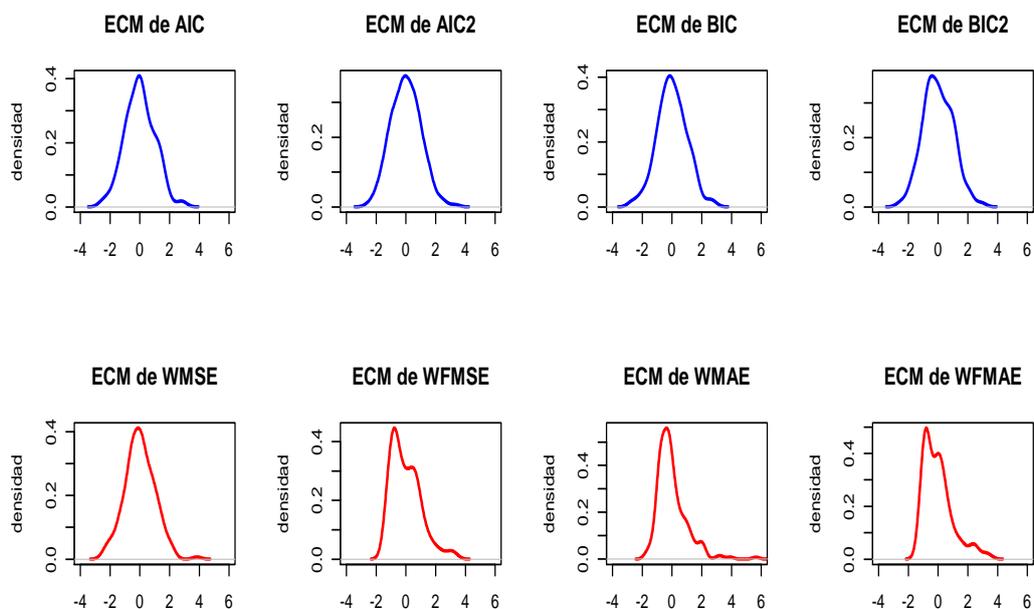


Figura 4.249: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇.

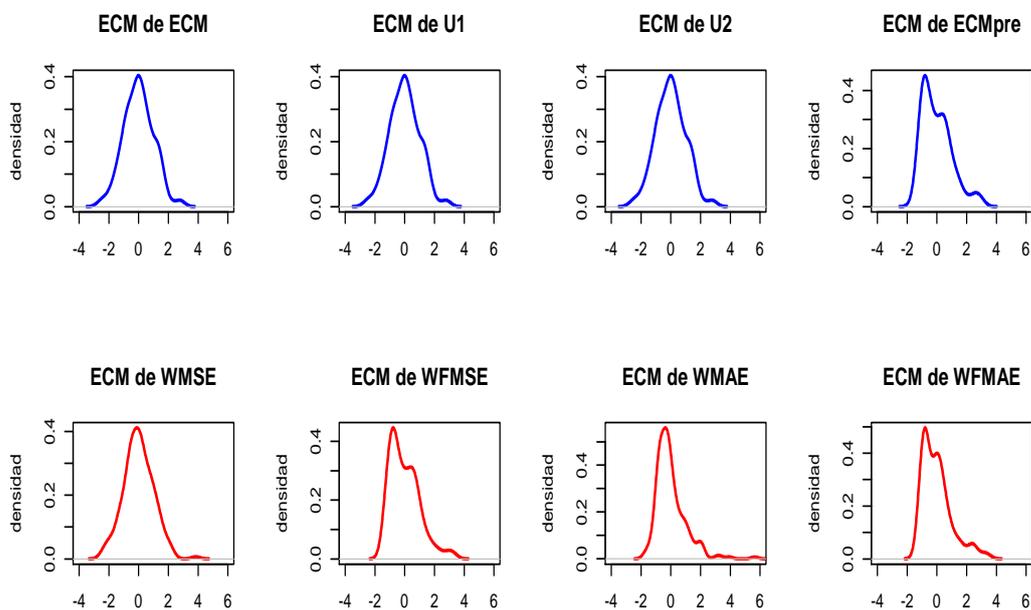


Figura 4.250: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇.

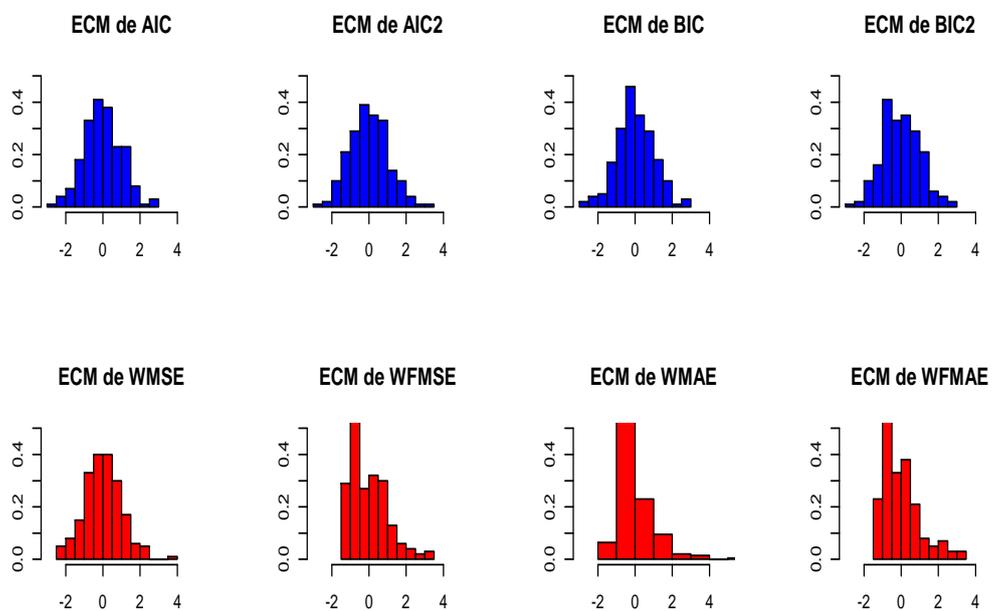


Figura 4.251: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇.

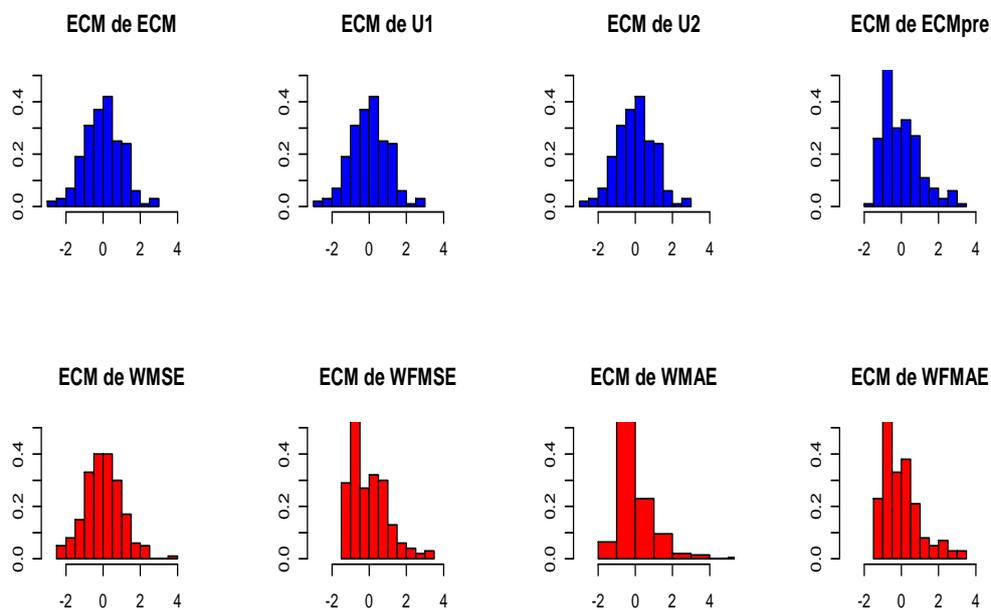


Figura 4.252: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 0, 1)(0, 0, 1)_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8668	0.8951	0.8611	0.8953	0.9480	0.8611
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8611	1.4870	0.9293	1.4642	1.1880	1.5740

Tabla 4.151: ECM global de ajuste, ECMG, para cada criterio en modelos $SARIMA(1, 0, 1)(0, 0, 1)_7$.

En la tabla 4.149 se contempla una mejoría aunque más moderada en esta ocasión en los criterios ECMpre, WFMSE y WFMAE, lo que puede comprobarse en los gráficos 4.245 y 4.246, sin embargo en la tabla 4.150 se observa una mejoría más destacada sobre todo en el criterio WFMAE, y en la tabla 4.151 también existe un empeoramiento en los resultados pero de una manera más moderada también en este caso.

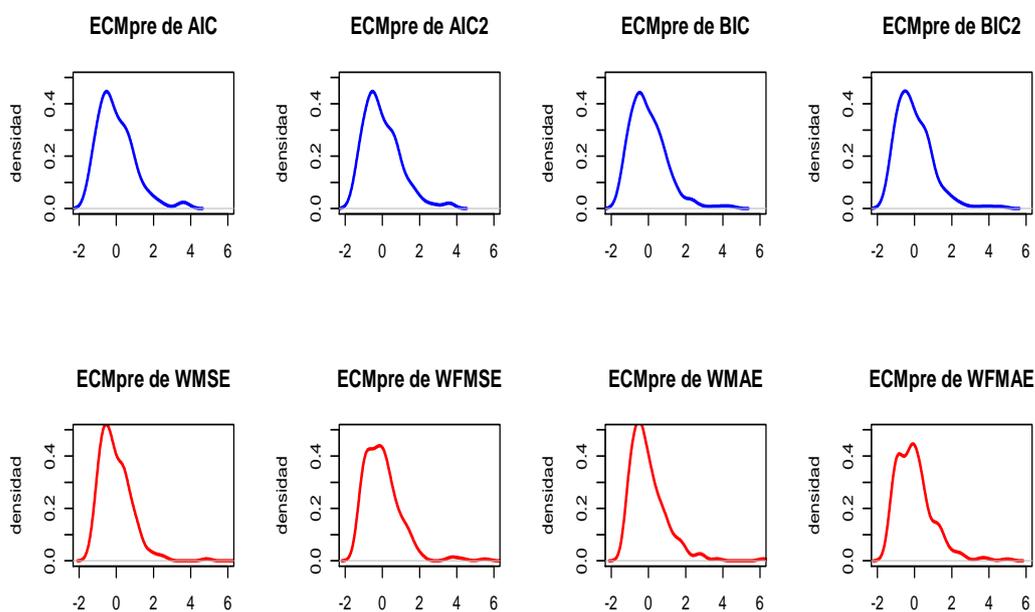


Figura 4.253: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1,0,1)(0,0,1)₇.

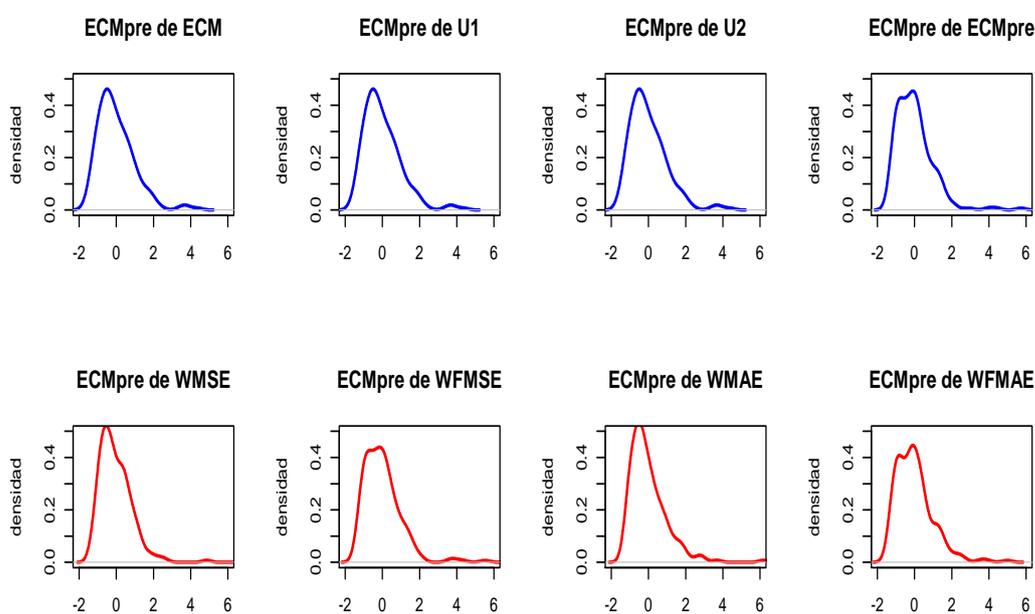


Figura 4.254: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1,0,1)(0,0,1)₇.

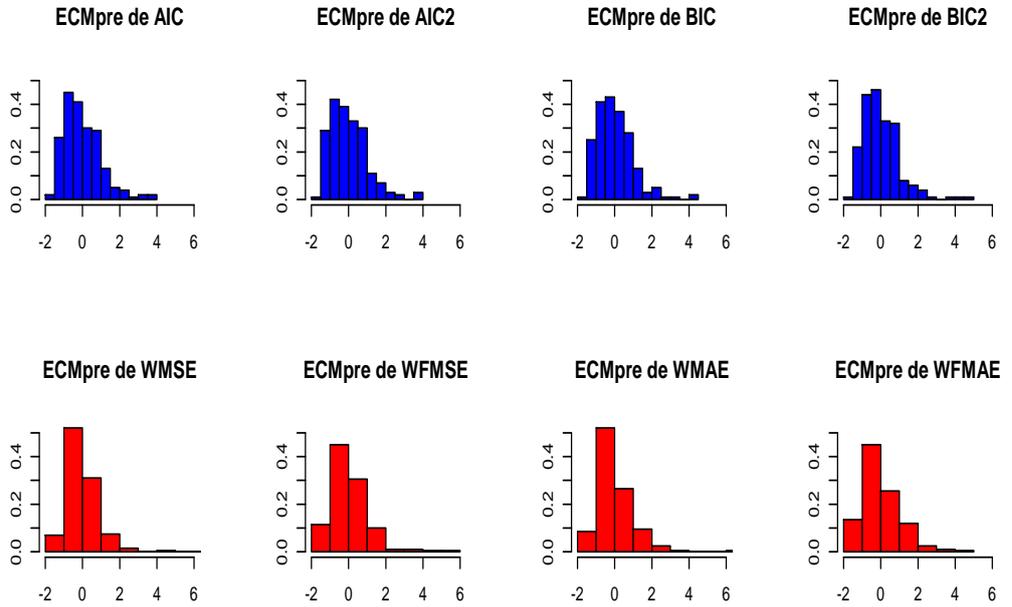


Figura 4.255: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇.

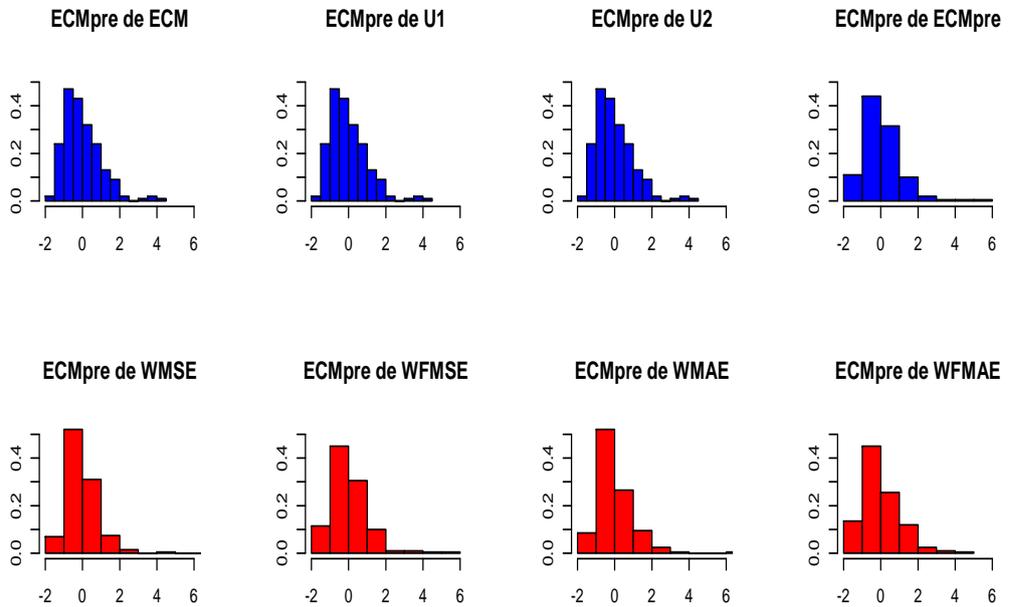


Figura 4.256: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 0, 1)(0, 0, 1)₇.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	1.2802	1.2590	1.2782	1.2625	1.2436	1.2782
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	1.2782	0.7699	1.3295	0.7949	1.5726	0.8795

Tabla 4.152: ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 0, 1)(0, 0, 1)₇.

En la tabla 4.152 se recogen los mejores resultados en el caso de los criterios WFMSE, WFMAE a pesar de que su valor en el ECMGpre es algo superior que en el criterio ECMpre pero en la tabla 4.151, de las individuales se comprobaban que funcionaban mejor éstos frente al ECMpre.

Para terminar con este capítulo, y no alargarlo demasiado analizaremos el último modelo que se mostrará.

Modelo SARIMA(1, 1, 1)(1, 1, 1)₇, cuyos coeficientes son $\phi_1 = -0,7$, $\theta_1 = 0,5$, $\Phi_1 = -0,3$ y $\Theta = 0,3$. (Modelo 191)

	AIC	AIC2	ECM	BIC	BIC2
Valor	2.372309	-0.3229122	0.5723217	2.506997	-0.206271
Modelo	270	270	270	215	190

Tabla 4.153: Criterios Clásicos de ajuste para SARIMA(1, 1, 1)(1, 1, 1)₇.

	U_1	U_2	ECMpre
Valor	0.02071298	0.02071743	0.037946
Modelo	270	270	132

Tabla 4.154: Criterios Clásicos de predicción para SARIMA(1, 1, 1)(1, 1, 1)₇.

	Ajuste		Predicción	
	WMSE	WMAE	WFMSE	WFMAE
Valor	0.1949824	0.3102977	0.03438867	0.1382251
Modelo	42	42	132	166

Tabla 4.155: Criterios propuestos para SARIMA(1, 1, 1)(1, 1, 1)₇.

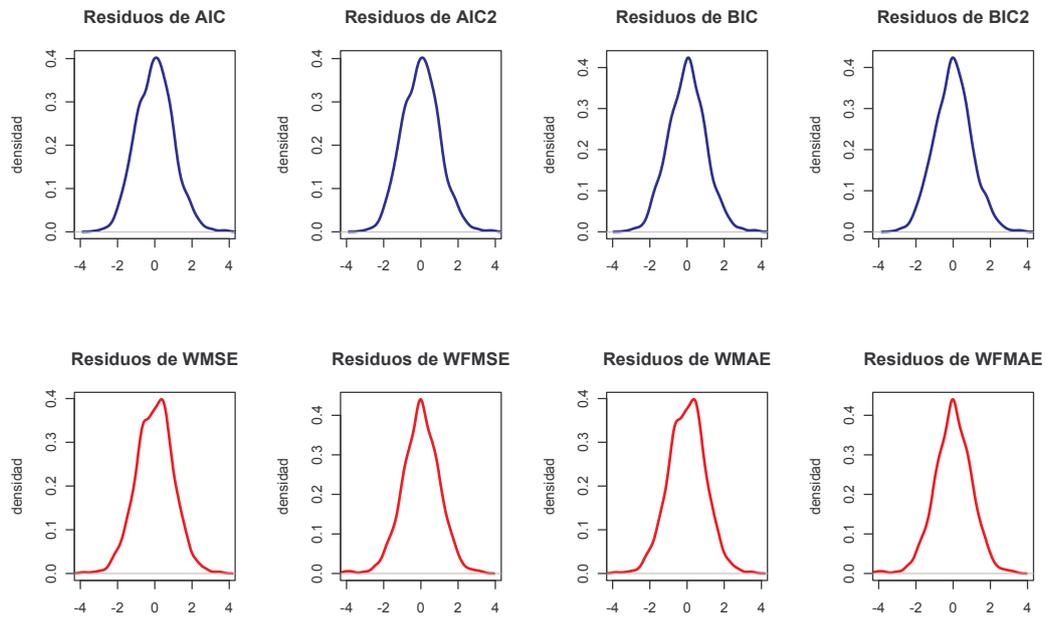


Figura 4.257: Residuos obtenidos con los modelos de los criterios de ajuste de $SARIMA(1, 1, 1)(1, 1, 1)_7$.

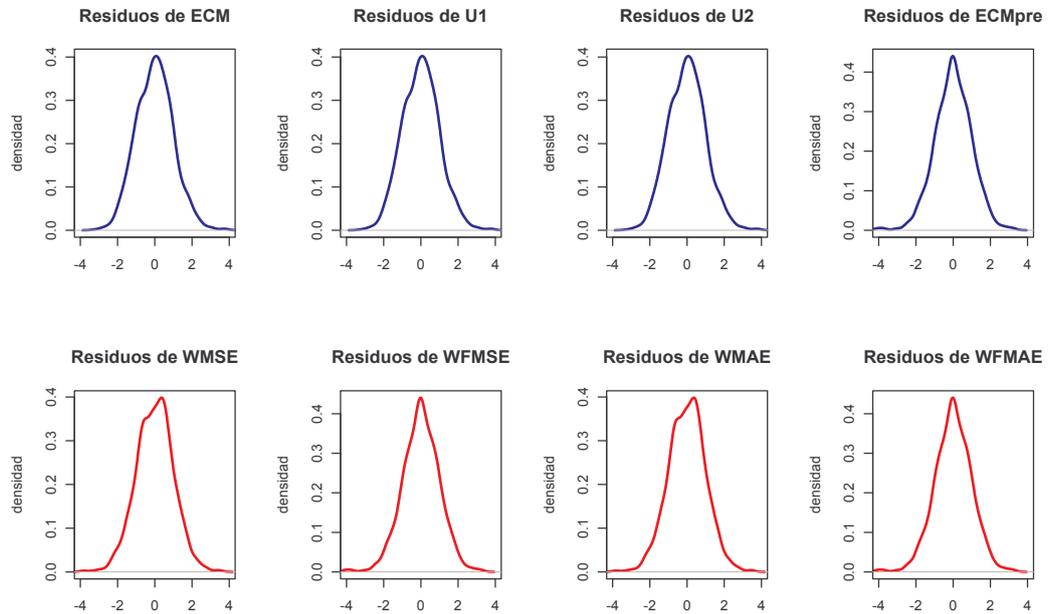


Figura 4.258: Residuos obtenidos con los modelos de los criterios de predicción de $SARIMA(1, 1, 1)(1, 1, 1)_7$.

Criterio	Media	Varianza
AIC	0.0409	1.6153
AIC2	0.0409	1.6153
ECM	0.0409	1.6153
BIC	0.0222	1.5188
BIC2	0.0343	1.5798
U1	0.0409	1.6153
U2	0.0409	1.6153
ECMpre	0.0425	2.5327
WMSE	0.1753	3.6566
WFMSE	0.0425	2.5327
WMAE	0.1753	3.6566
WFMAE	0.0425	2.5327

Tabla 4.156: Medidas descriptivas de los residuos en el conjunto de comprobación para SARIMA(1, 1, 1)(1, 1, 1) $_7$.

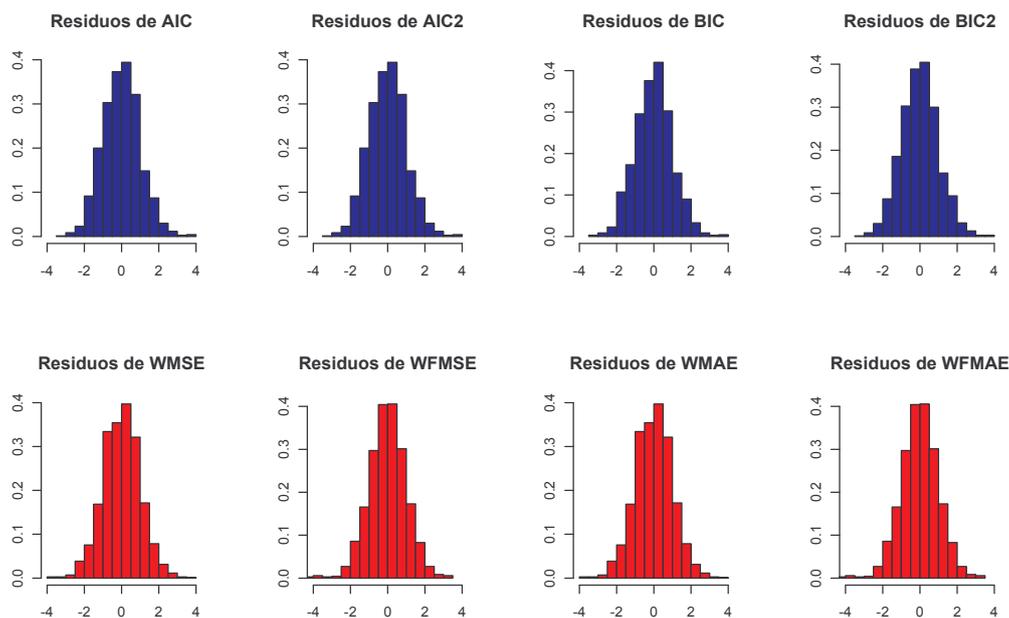


Figura 4.259: Histograma de los residuos obtenidos con los modelos de los criterios de ajuste de SARIMA(1, 1, 1)(1, 1, 1) $_7$.

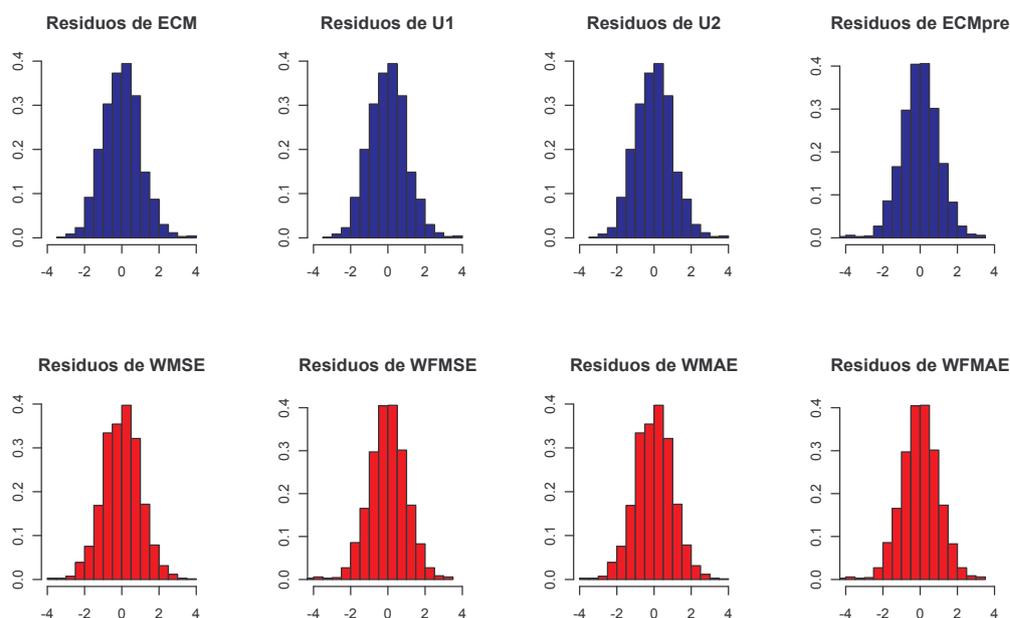


Figura 4.260: Histograma de los residuos obtenidos con los modelos de los criterios de predicción de $SARIMA(1, 1, 1)(1, 1, 1)_7$.

Criterio	ECMpre1	ECMpre2	ECMpre3	ECMpre4	ECMpre5	ECMpre6	ECMpre7
AIC	1.1527	1.3506	1.6689	1.4536	2.0065	1.7592	1.9275
AIC2	1.1527	1.3506	1.6689	1.4536	2.0065	1.7592	1.9275
ECM	1.1527	1.3506	1.6689	1.4536	2.0065	1.7592	1.9275
BIC	1.0426	1.3298	1.4967	1.4072	1.8731	1.6867	1.7989
BIC2	1.1998	1.3481	1.5580	1.4939	1.9548	1.7494	1.8427
U1	1.1527	1.3506	1.6689	1.4536	2.0065	1.7592	1.9275
U2	1.1527	1.3506	1.6689	1.4536	2.0065	1.7592	1.9275
ECMpre	2.1890	2.1264	2.5524	2.2151	3.1822	2.6728	2.8039
WMSE	2.0388	2.4661	3.9026	4.2189	4.6109	4.0888	4.4849
WFMSE	2.1890	2.1264	2.5524	2.2151	3.1822	2.6728	2.8039
WMAE	2.0388	2.4661	3.9026	4.2189	4.6109	4.0888	4.4849
WFMAE	2.1890	2.1264	2.5524	2.2151	3.1822	2.6728	2.8039

Tabla 4.157: ECMpre global para cada predicción realizada en el conjunto de comprobación en $SARIMA(1, 1, 1)(1, 1, 1)_7$.

A la vista de la tabla 4.157 y junto con la tabla 4.156, lo único más destacable es el peor comportamiento de los 5 últimos criterios mostrados en la tabla, los criterios ECMpre, WFMSE y WFMAE han seleccionado el mismo modelo, lo mismo ha ocurrido con los criterios WMSE y WMAE (tabla 4.154 y 4.155).

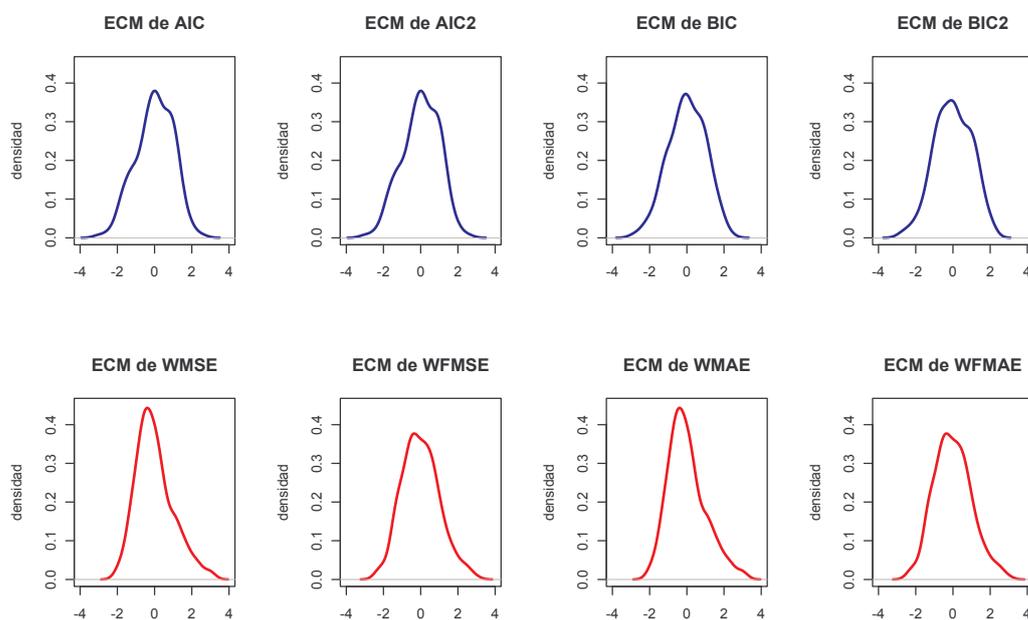


Figura 4.261: ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 1)(1, 1, 1) $_7$.

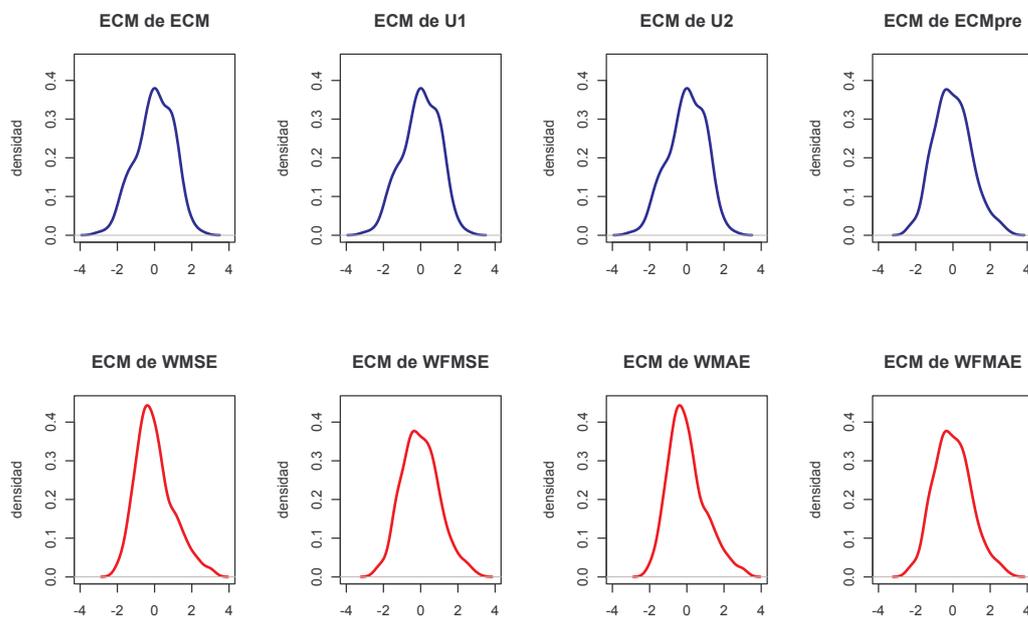


Figura 4.262: ECM de ajuste de los modelos obtenidos con los criterios de predicci3n en SARIMA(1, 1, 1)(1, 1, 1) $_7$.

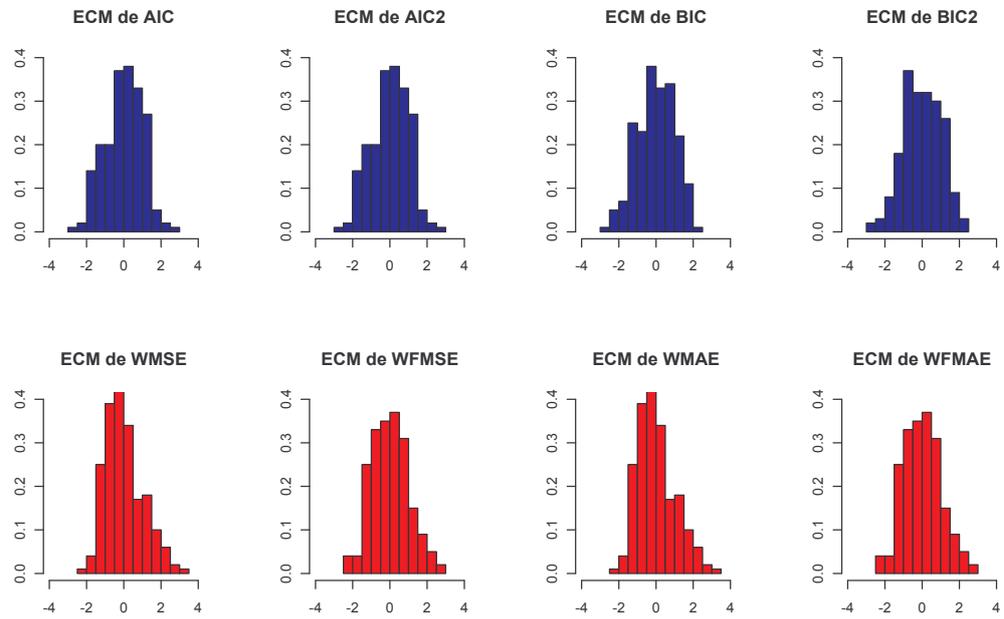


Figura 4.263: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 1)(1, 1, 1)₇.

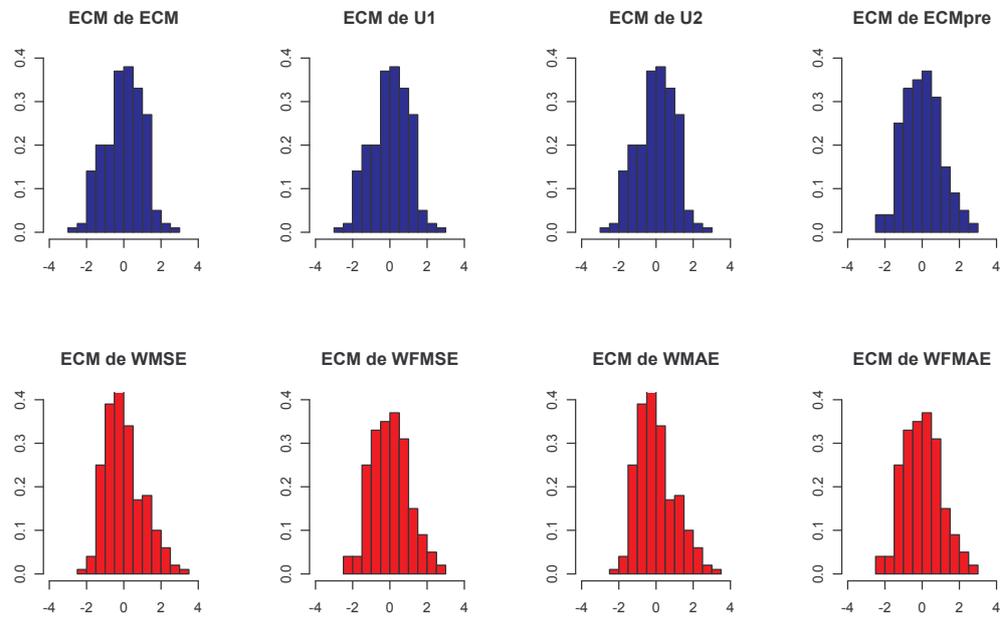


Figura 4.264: Histograma de los ECM de ajuste de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 1)(1, 1, 1)₇.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8863	0.8863	0.8863	0.9061	1.0115	0.8863
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8863	1.6698	1.6558	1.6698	1.6558	1.6698

Tabla 4.158: El ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 1)(1, 1, 1)₇.

En la tabla 4.158 se comprueba que los peores resultados una vez más, se alcanzan en los últimos 5 criterios aunque en esta ocasión junto con el BIC2, que tampoco consigue un buen ajuste, aunque en general no hay ningún criterio que destaque en este aspecto significativamente con el resto, puede verse en las gráficas 4.261 y 4.262.

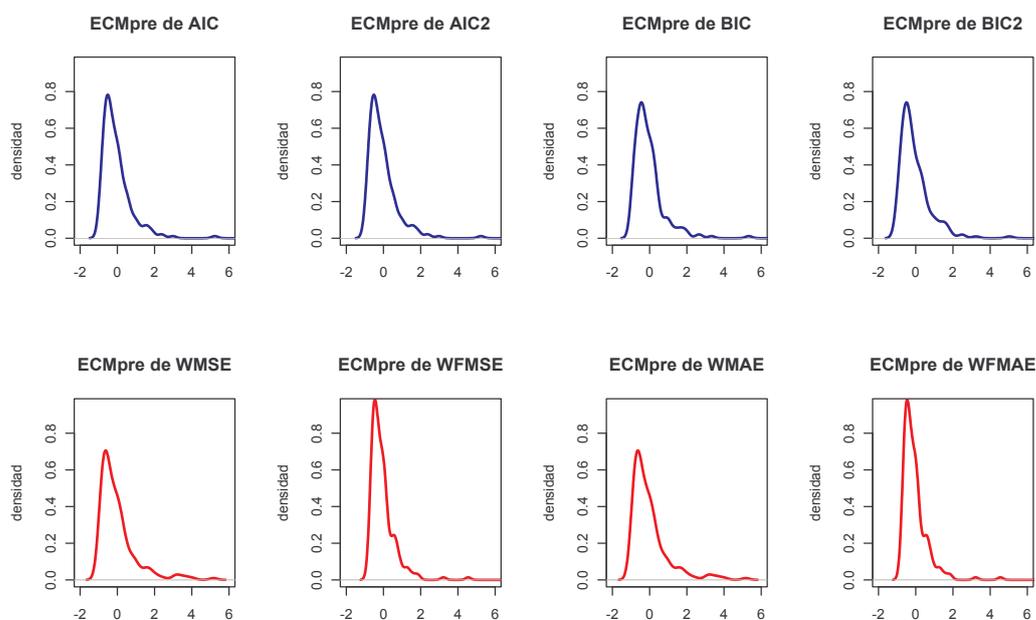


Figura 4.265: ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 1)(1, 1, 1)₇.

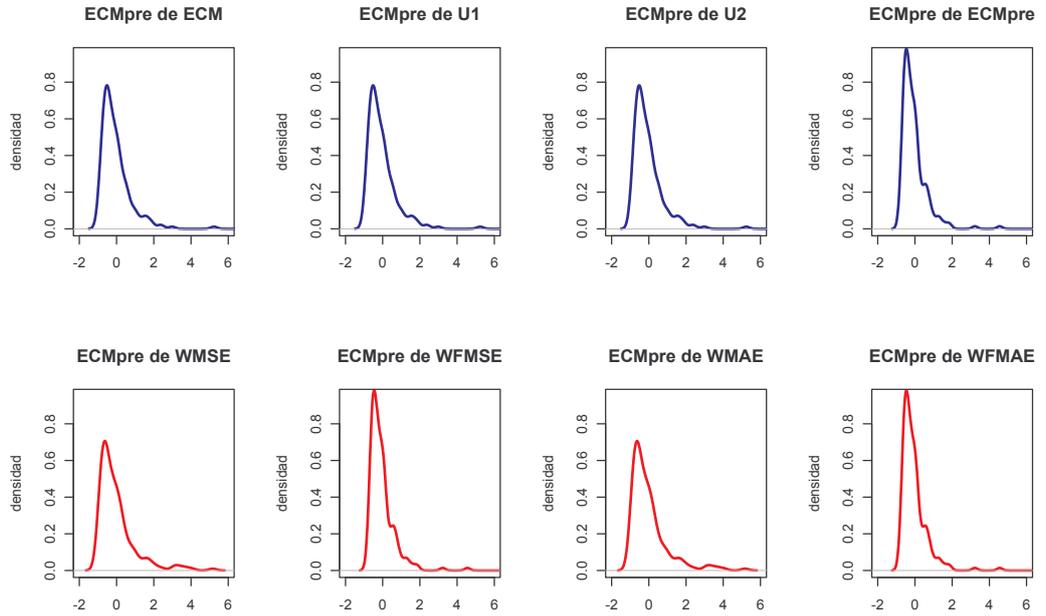


Figura 4.266: ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 1)(1, 1, 1)₇.

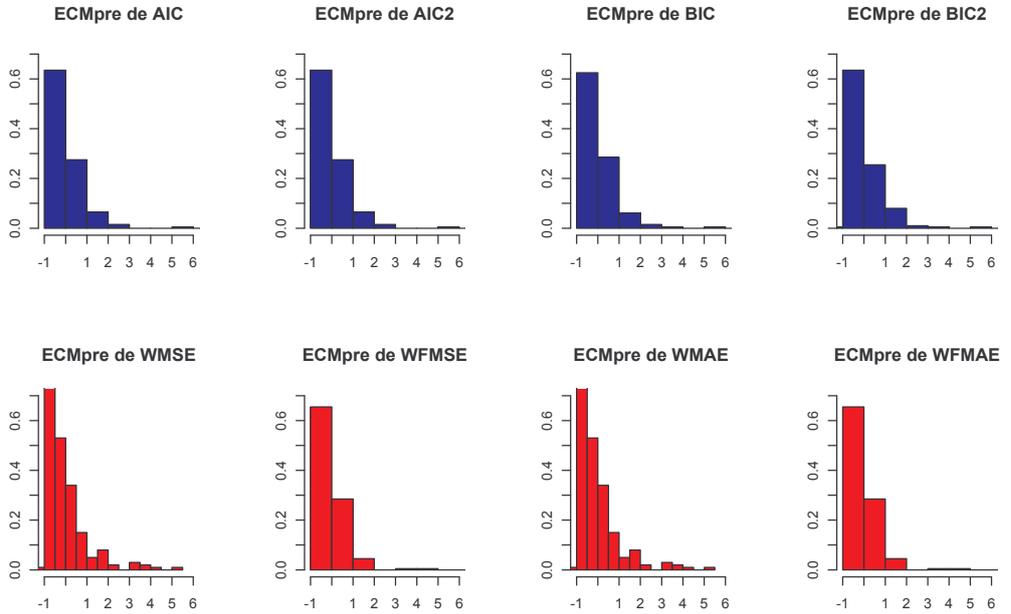


Figura 4.267: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de ajuste en SARIMA(1, 1, 1)(1, 1, 1)₇.

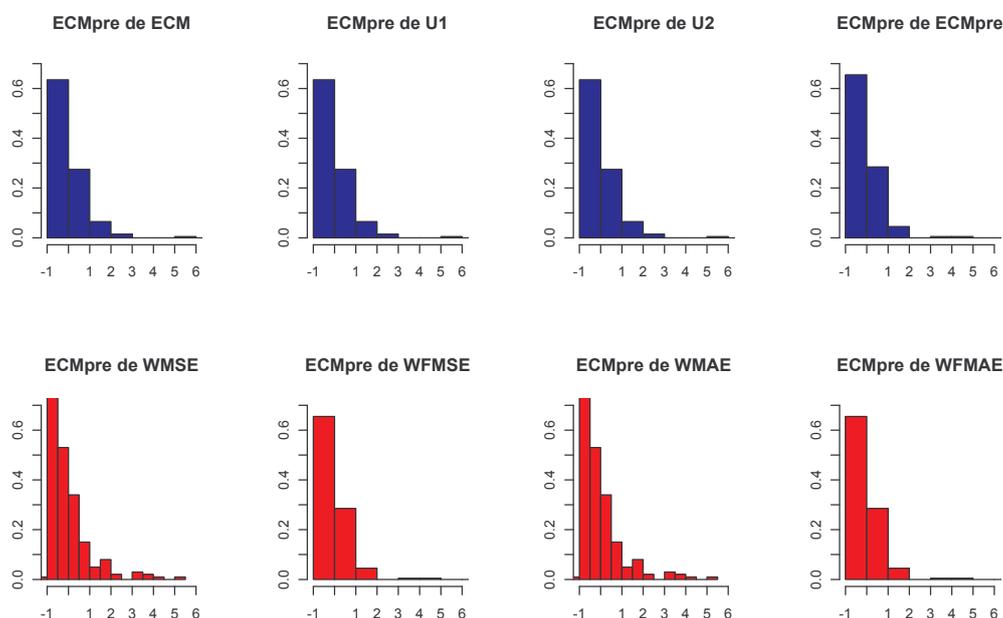


Figura 4.268: Histograma de los ECM de predicción de los modelos obtenidos con los criterios de predicción en SARIMA(1, 1, 1)(1, 1, 1)₇.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	1.6170	1.6170	1.6170	1.5193	1.5810	1.6170
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	1.6170	2.5345	3.6873	2.5345	3.6873	2.5345

Tabla 4.159: ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 1)(1, 1, 1)₇.

Con respecto a los errores cuadráticos medios de predicción global, los criterios citados alcanzan también en esta tabla 4.159 los peores resultados al respecto, sin que existe ningún comportamiento que sea digno de resaltar en este campo.

En esta ocasión tan sólo ha sido posible detectar el modelo generador de la serie con el criterio WFMAE, aunque el porcentaje de acierto ha sido de tan sólo un 1%.

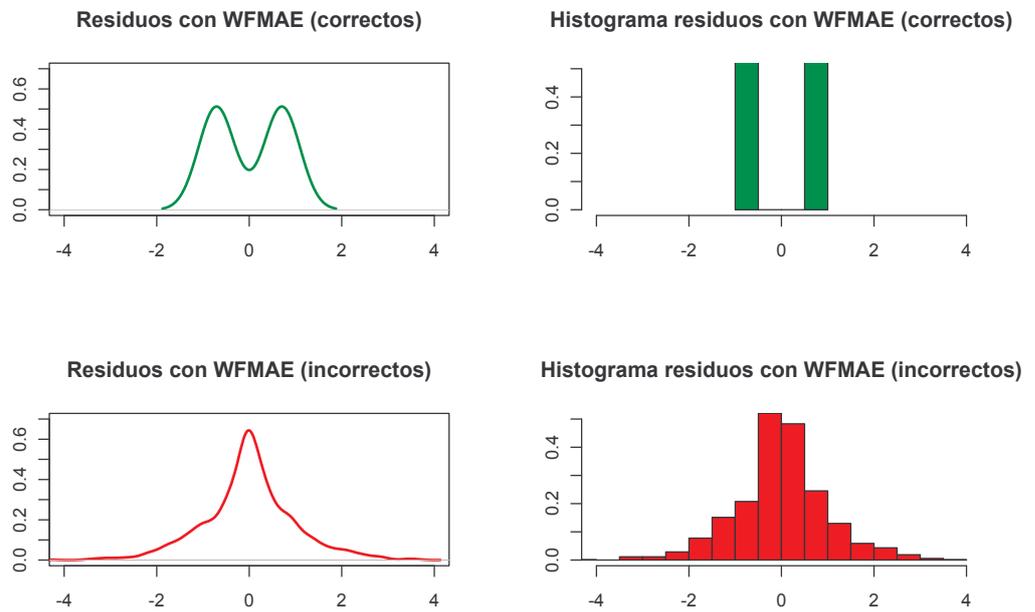


Figura 4.269: Residuos de las series correctas e incorrectas con el criterio WFMAE en el SARIMA(1, 1, 1)(1, 1, 1)₇.

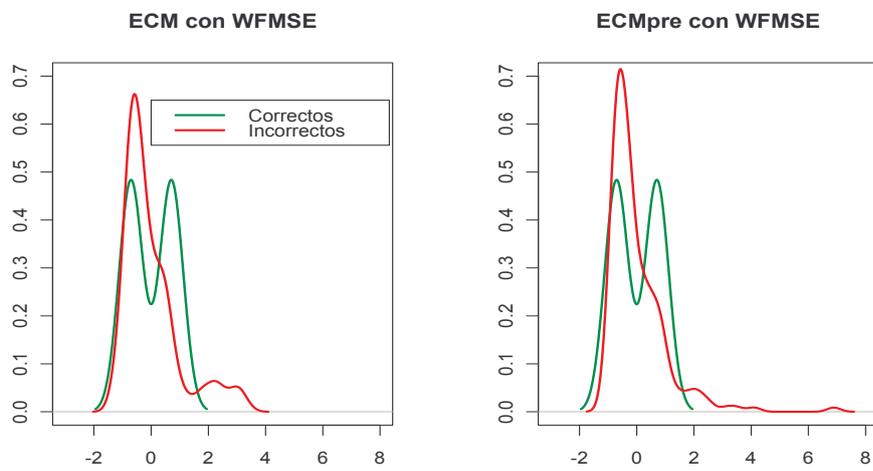


Figura 4.270: ECM de ajuste y predicción en las series correctas e incorrectas con el criterio WFMAE en SARIMA(1, 1, 1)(1, 1, 1)₇.

	Media	Varianza	ECMG	ECMGpre
Correctos	0.0569	1.0335	0.9673	1.0367
Incorrectos	0.0058	0.6675	1.6934	0.8676

Tabla 4.160: Medidas descriptivas en el conjunto de comprobación y ECMG de ajuste y predicción para SARIMA(1, 1, 1)(1, 1, 1) $_7$ con el criterio WFMAE.

	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
Correctos	1.0357	0.0841	0.7737	0.4699	2.9526	0.5674	1.3738
Incorrectos	0.4494	0.7509	0.6368	0.9511	0.8656	1.3355	1.0835

Tabla 4.161: ECM global de predicción en SARIMA(1, 1, 1)(1, 1, 1) $_7$ con el criterio WFMAE en el conjunto de comprobación.

Ante el criterio WFMAE, el mejor comportamiento se consigue en el grupo de las series incorrectas (tabla 4.160) en las que se consigue una media y una varianza inferior a las correctas, y un ECMGpre que también mejora los resultados con respecto a las correctas, es variante el comportamiento sufrido en relación a las predicciones individuales, aunque la primera de ellas es más precisa en el caso de las series incorrectas (tabla 4.161)

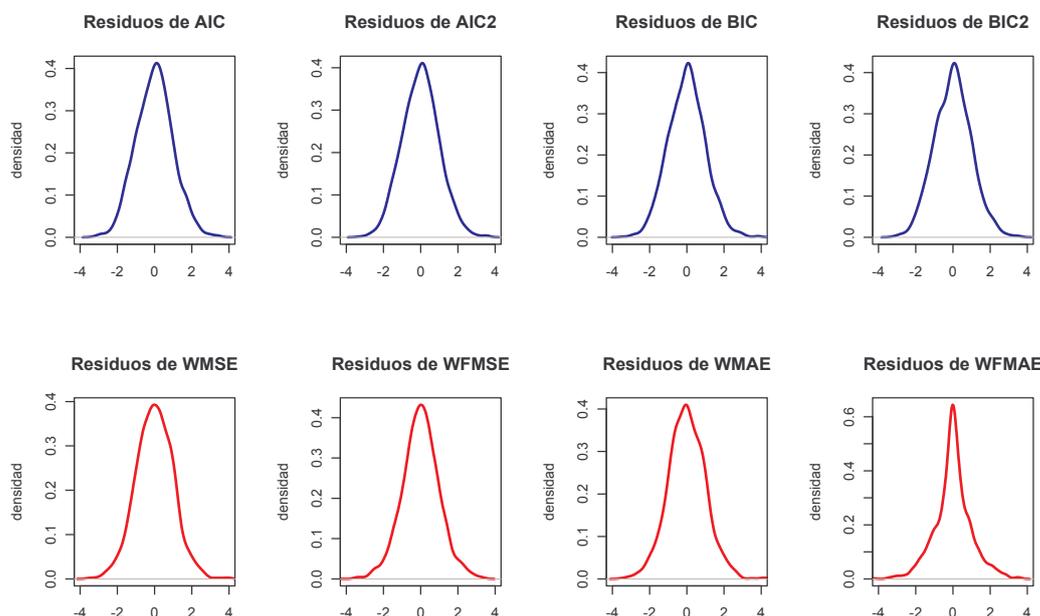


Figura 4.271: Residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1) $_7$.

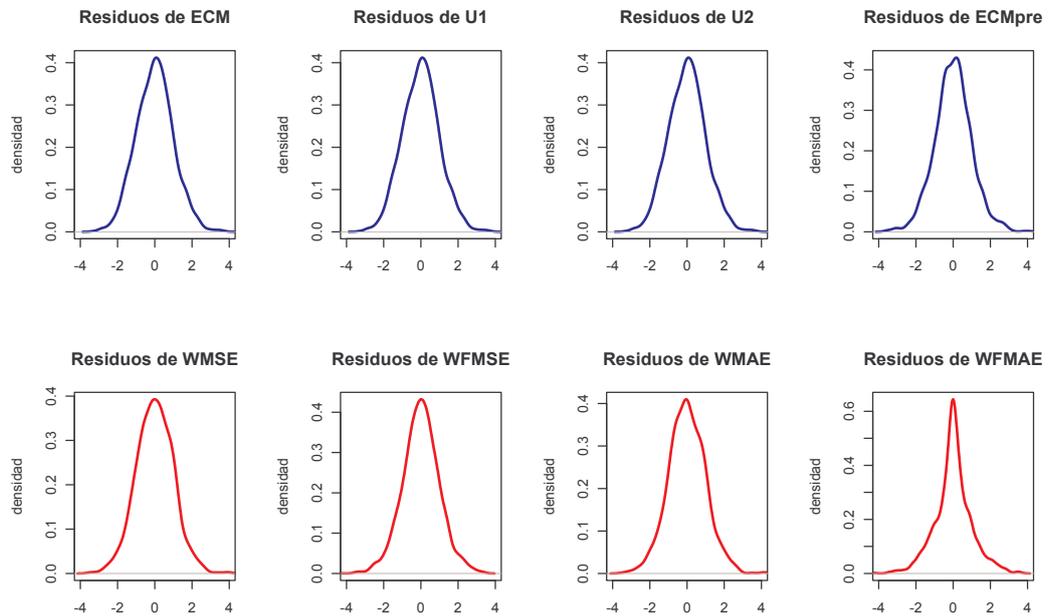


Figura 4.272: Residuos obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 1)(1, 1, 1)_7$.

Criterio	Media	Varianza
AIC	0.0352	1.5110
AIC2	0.0292	1.4673
ECM	0.0387	1.5248
BIC	0.0396	1.5276
BIC2	0.0352	1.4837
U1	0.0387	1.5248
U2	0.0387	1.5248
ECMpre	-0.0041	0.8040
WMSE	0.0291	1.7768
WFMSE	-0.0009	0.8441
WMAE	0.0571	1.8114
WFMAE	0.0063	0.8692

Tabla 4.162: Medidas descriptivas de los residuos de predicción (conjunto de comprobación) para $SARIMA(1, 1, 1)(1, 1, 1)_7$.

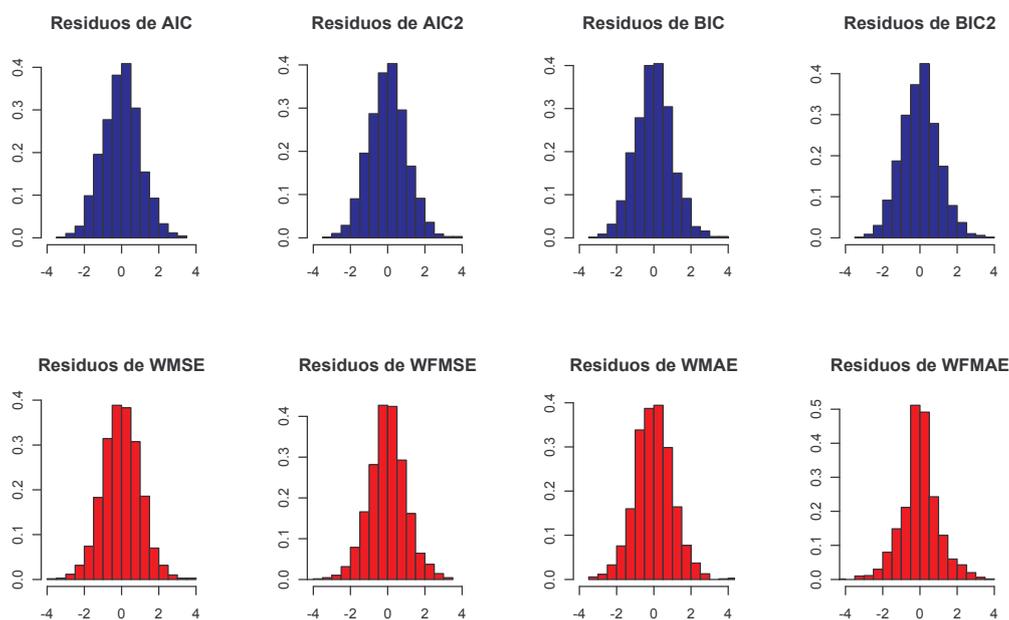


Figura 4.273: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1)₇.

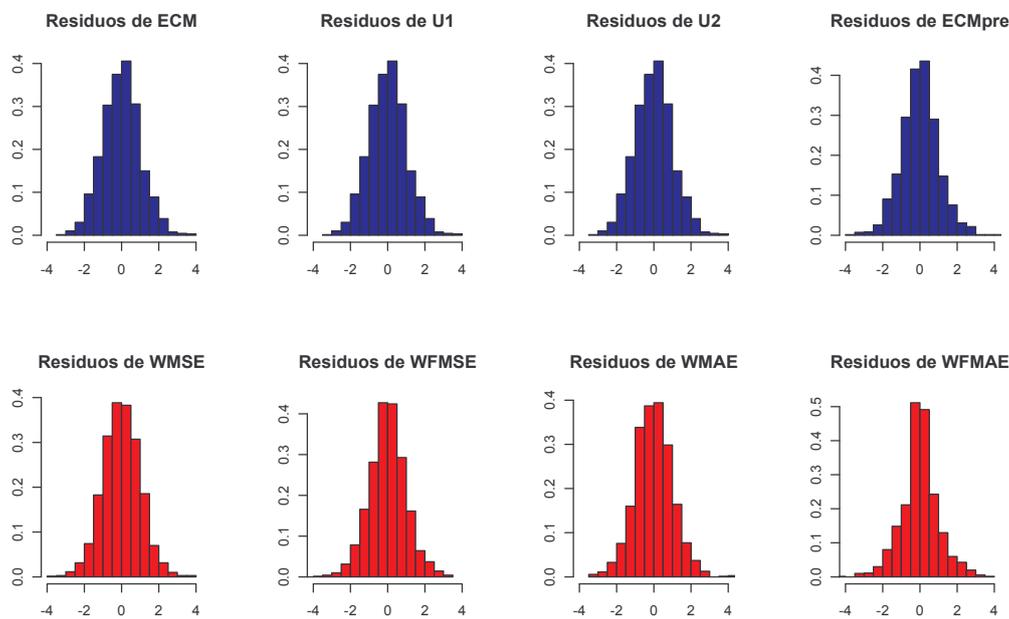


Figura 4.274: Histograma de los residuos obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1)₇.

Criterio	ECMGpre1	ECMGpre2	ECMGpre3	ECMGpre4	ECMGpre5	ECMGpre6	ECMGpre7
AIC	1.0269	1.3246	1.5181	1.3965	1.8181	1.6442	1.8574
AIC2	1.0515	1.2789	1.4689	1.3541	1.7653	1.5502	1.8080
ECM	1.0780	1.3270	1.5379	1.4155	1.8456	1.6263	1.8540
BIC	1.0491	1.3112	1.5580	1.4317	1.8840	1.6179	1.8521
BIC2	1.0257	1.2464	1.4690	1.3738	1.8094	1.6176	1.8514
U1	1.0780	1.3270	1.5379	1.4155	1.8456	1.6263	1.8540
U2	1.0780	1.3270	1.5379	1.4155	1.8456	1.6263	1.8540
ECMpre	0.4467	0.7460	0.6716	0.8778	0.8686	1.0713	0.9459
WMSE	1.1845	1.4174	1.8254	1.6514	2.1355	2.0463	2.1833
WFMSE	0.3062	0.5859	0.5446	0.8662	1	1.3865	1.2193
WMAE	1.1621	1.5412	1.7827	1.7417	2.1086	2.1415	2.2249
WFMAE	0.4553	0.7443	0.6382	0.9463	0.8865	1.3278	1.0864

Tabla 4.163: ECM globales para cada predicción realizada en el conjunto de comprobación en SARIMA(1, 1, 1)(1, 1, 1)₇.

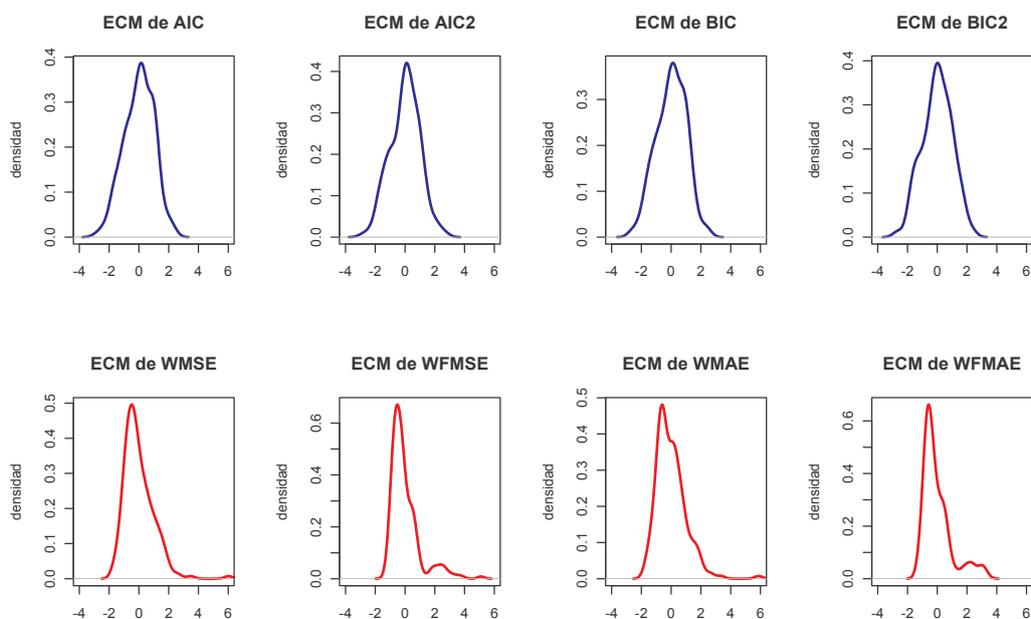


Figura 4.275: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1)₇.

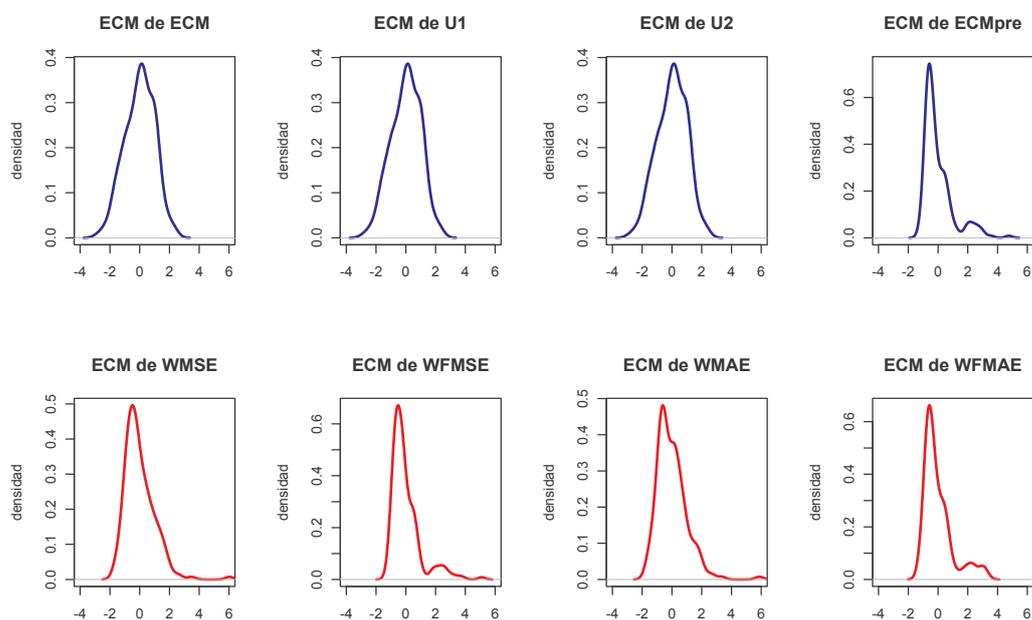


Figura 4.276: ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1)₇.

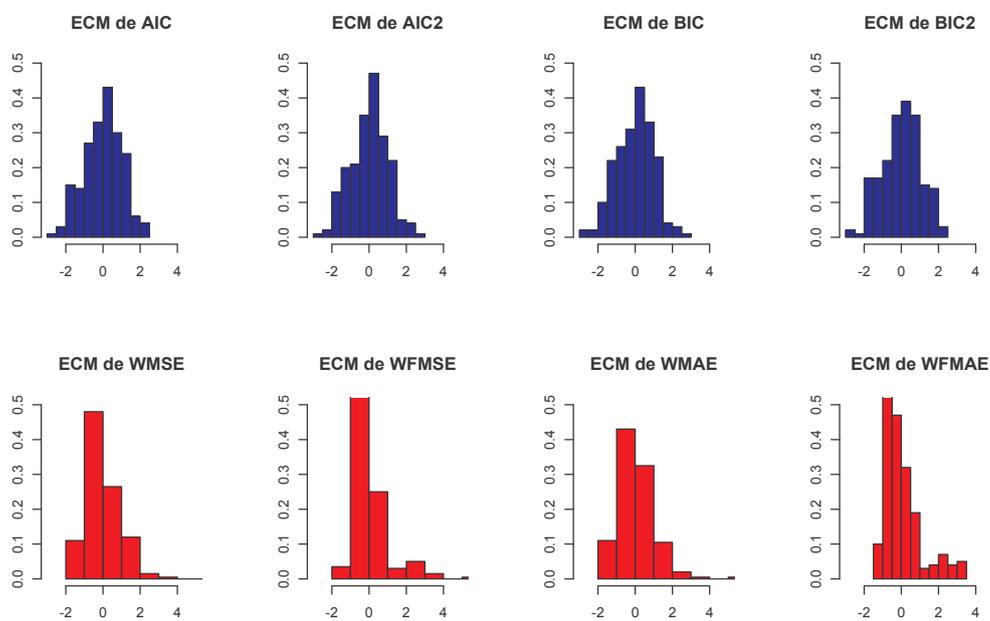


Figura 4.277: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1)₇.

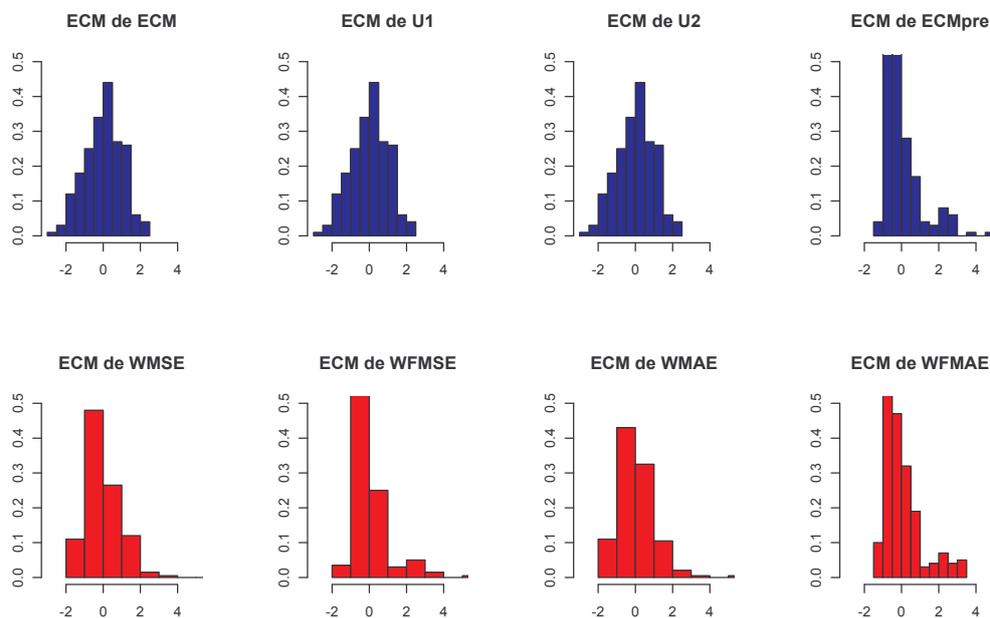


Figura 4.278: Histograma de los ECM de ajuste obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1, 1, 1)(1, 1, 1)₇.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMG	0.8673	0.8934	0.8605	0.8838	0.93	0.8605
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMG	0.8605	1.6975	1.2505	1.6522	1.2931	1.6861

Tabla 4.164: ECM global de ajuste, ECMG, para cada criterio en modelos SARIMA(1, 1, 1)(1, 1, 1)₇.

En los gráficos 4.271 y 4.272 ya se puede apreciar la mejoría sufrida por los criterios ECMpre, WFMSE y WFMAE, en la tabla 4.162 podemos verificar esta apreciación, siendo en esta ocasión el criterio WFMSE el que mejor funciona, como se puede ver en la tabla 4.163, y en esta última tabla (4.164) podemos ver que en cuanto al ajuste una vez más en estos criterios es donde peor funciona aunque en este caso no se consiguen buenos resultados en general.

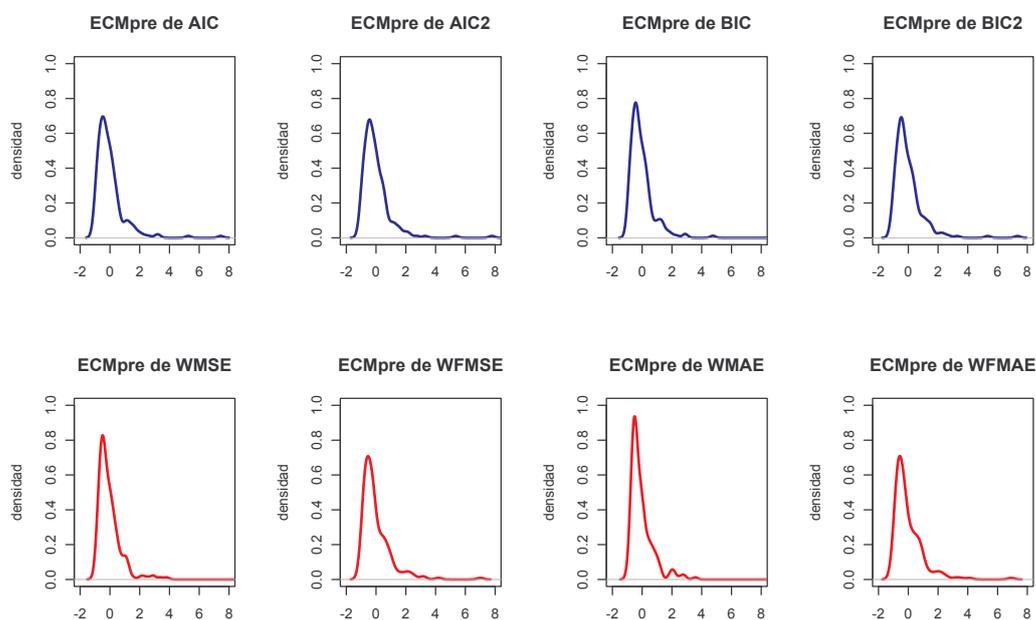


Figura 4.279: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1,1,1)(1,1,1)₇.

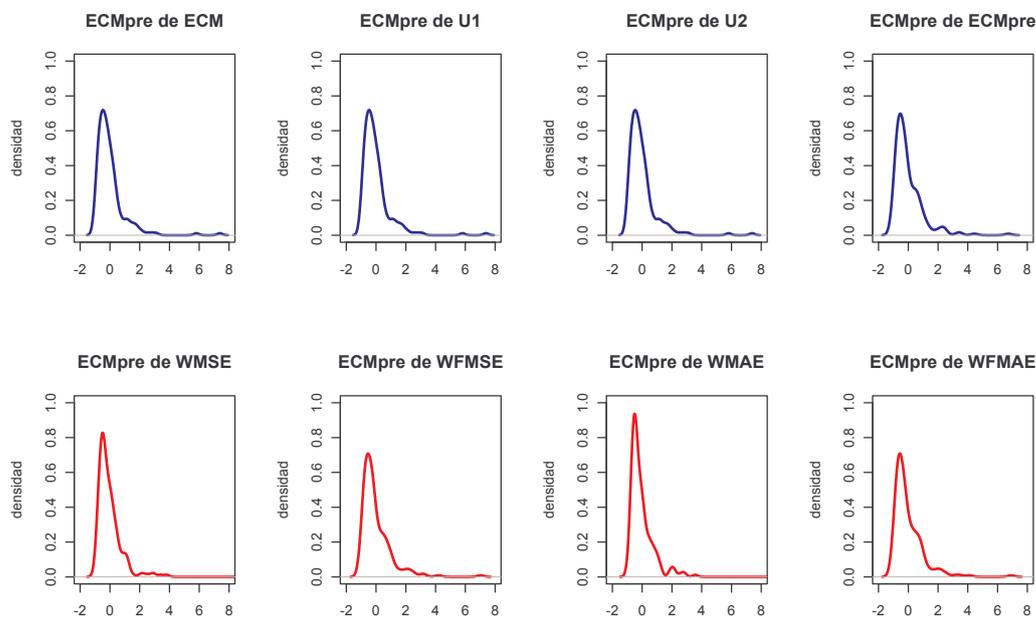


Figura 4.280: ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie SARIMA(1,1,1)(1,1,1)₇.

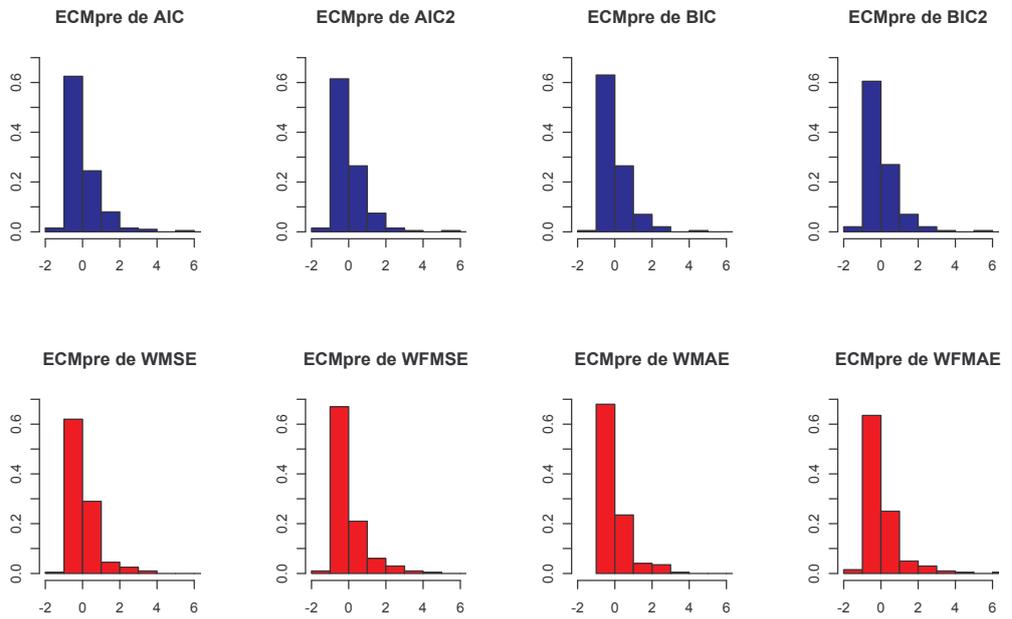


Figura 4.281: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 1)(1, 1, 1)_7$.

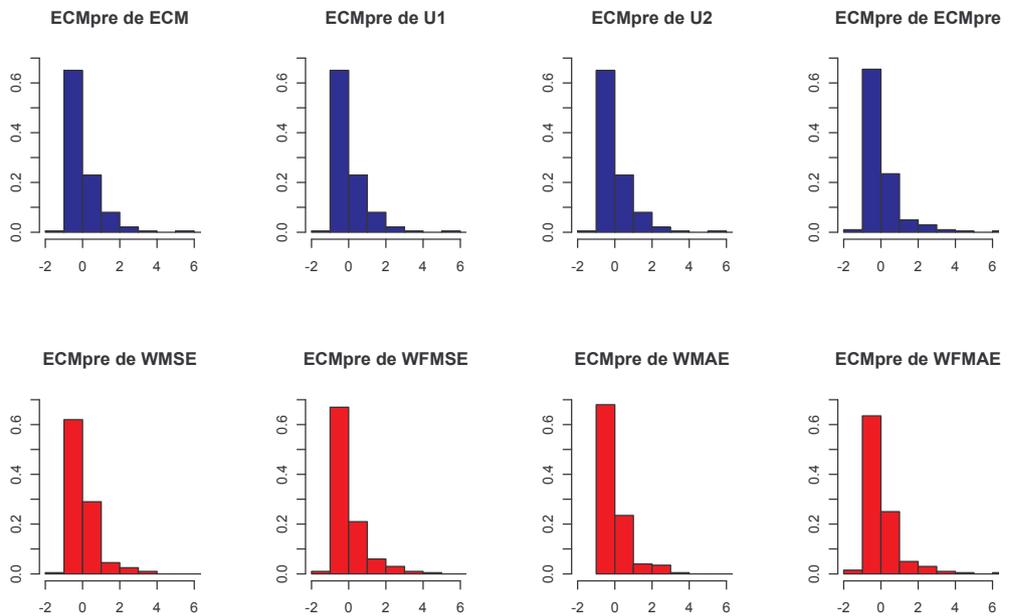


Figura 4.282: Histograma de los ECM de predicción obtenidos con los modelos asignados por cada criterio a cada serie $SARIMA(1, 1, 1)(1, 1, 1)_7$.

	AIC	AIC2	ECM	BIC	BIC2	U1
ECMGpre	1.5122	1.4681	1.5263	1.5291	1.4848	1.5263
	U2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
ECMGpre	1.5263	0.8040	1.7777	0.8441	1.8147	0.8693

Tabla 4.165: ECM global de predicción, ECMGpre, para cada criterio en modelos SARIMA(1, 1, 1)(1, 1, 1)₇.

Por último, y como era de esperar en la tabla 4.165 se puede comprobar como los criterios ya mencionados en diversas ocasiones son los que mejores resultados en predicción obtienen (ECMpre, WFMSE y WFMAE), y además, y siguiendo la misma línea que en los modelos anteriores, los resultados mostrados en las tablas asociadas a este procedimiento de actuación, superan a los resultados mostrados en el procedimiento inicial, donde se determinaba como modelo “óptimo” aquel modelo que minimizaba el criterio empleado en las 200 series estudiadas.

En general, ante los procesos estacionarios se puede decir, que el procedimiento para obtener buenos resultados en la predicción es utilizar el los criterios de predicción aportados en este trabajo y utilizar el modelo que determinen éstos, ya que no sólo se ha conseguido disminuir el ECMGpre sino los ECMGpre de cada una de las predicciones realizadas.

Capítulo 5

Aplicación: La Demanda de Agua

El agua es el elemento y principio de las cosas

Tales de Mileto (Filósofo y matemático griego)

5.1. Introducción

El agua es un recurso natural incomparable, escaso y estratégico, elemento clave para la vida y el desarrollo social, difícilmente se puede exagerar la importancia del agua, es seguramente el más valioso e imprescindible de los existentes en el planeta Tierra. El agua debería ser un derecho humano y, sin embargo, 1100 millones de personas carecen de acceso al agua potable. La aparente abundancia del agua en el mundo ha dado la impresión, en el pasado, de que se trataba de un bien inagotable. Era también el más barato. En muchos lugares el agua era gratuita. La escasez de agua afecta ya a todos los continentes. Cerca de 1200 millones de personas, casi una quinta parte de la población mundial, vive en áreas de escasez física de agua, mientras que 500 millones se aproximan a esta situación. Otros 1600 millones, alrededor de un cuarto de la población mundial, se enfrentan a situaciones de escasez económica de agua, donde los países carecen de la infraestructura necesaria para transportar el agua desde ríos y acuíferos.

La escasez de agua constituye uno de los principales desafíos del siglo XXI al que se están enfrentando ya numerosas sociedades de todo el mundo. A lo largo del último siglo, el uso y consumo de agua creció a un ritmo dos veces superior al de la tasa de crecimiento de la población y, aunque no se puede hablar de escasez hídrica a nivel global, va en aumento el número de regiones con niveles crónicos de carencia de agua.

La escasez de agua es un fenómeno no solo natural sino también causado por la acción del ser humano. Hay suficiente agua potable en el planeta para abastecer a 6000 millones de personas, pero ésta está distribuida de forma irregular, se desperdicia, está contaminada y se gestiona de forma insostenible. Una política subsidiada que ha fomentado su derroche, su explotación excesiva, originando problemas de escasez, pérdida de calidad o deterioro medioambiental.

A día de hoy, cerca de 700 millones de personas procedentes de 43 países diferentes sufren escasez de agua. En 2025, 1800 millones de personas vivirán en países o regiones con escasez absoluta de agua y dos terceras partes de la población mundial podrían hacerlo en condiciones de estrés hídrico. Bajo el contexto actual de cambio climático, en el 2030, casi la mitad de la población mundial vivirá en áreas de estrés hídrico, incluidos entre 75 y 250 millones de personas de África. Además, la escasez de agua en áreas áridas o semi-áridas provocará el desplazamiento de entre 24 y 700 millones de personas. En el África Subsahariana se concentra el mayor número de países con estrés hídrico.

En la medida en que el agua se va convirtiendo en un recurso cada vez más escaso y deteriorado por la intervención humana, los poderes públicos han ido estableciendo nuevas medidas para mejorar su gestión. En este contexto, la Directiva Marco del Agua (DMA)¹ establece el objetivo claro de que en el año 2015 se consiga un buen estado ecológico y un uso sostenible para todas las aguas europeas. Además, recoge que los Estados miembros garantizarán, a más tardar en 2010, que la política de tarifas del agua proporcione incentivos para el uso eficiente y la recuperación de los costes de los servicios relacionados con el agua.

Este nuevo marco institucional pone de relieve la necesidad de caracterizar y analizar en profundidad la demanda de agua, lo que motiva el interés por utilizar instrumentos económicos para mejorar la eficiencia en el uso del agua.

5.2. Planteamiento del problema

La demanda de agua se ha convertido a lo largo de los últimos años y, en especial en el nuevo siglo, en uno de los principales problemas a nivel mundial, a medida que numerosos y variados factores confieren que el análisis del agua sea una importante preocupación para los ciudadanos. El primero de ellos, es el hecho de que hemos asistido a una mejora de la actividad económica y del bienestar general que han provocado de un modo incesante las demandas de servicios de agua para todos los usos, el consumo global de agua en España se ha multiplicado más del doble que el incremento de la población, además, las fuentes, los manantiales, las cuencas o cañadas están en acelerada vía de extinción y se sufren cambios de clima y de suelo, inun-

¹DIRECTIVA 2000/60/CE, del 23 de octubre del 2000

daciones, sequías y desertización, que condicionan gravemente la seguridad de su abastecimiento.

El crecimiento de la demanda de agua y la escasez física son factores que requieren mejorar la eficiencia económica del uso del agua y actuar de acuerdo a los objetivos ambientales y de sostenibilidad en el contexto de la Directiva Marco del Agua y de una nueva política del agua. Muchos de los debates que se plantean en torno al agua suscitan enfrentamientos que, en ocasiones, parten de un mal entendimiento del agua como bien económico.

El funcionamiento de la economía, el crecimiento económico y el bienestar individual y colectivo de las personas depende de los servicios de agua. El agua es indispensable para cualquier actividad: la industrial, la agrícola y la urbana ya que promueve su desarrollo económico y social. Para comprender la importancia del agua, como elemento esencial para la generación de riqueza y bienestar, es necesario analizar profundamente la evolución de la demanda de agua y las actuaciones de intervención pública.

En España la capacidad de agua embalsada es en la actualidad superior a 50000 Hm^3 al año, lo que ofrece una disponibilidad de agua de unos 2800 m^3 al año (2800000 litros) por persona al año. Esta disponibilidad es mayor que la de la media de la Unión Europea pero el problema fundamental es que se distribuye de forma muy desigual, por lo que algunas zonas secas sufren problemas de escasez de agua. Los planes hidrológicos están encaminados al trasvase de agua de zonas que la tienen en abundancia a otras más secas a través de la construcción de grandes presas y embalses, túneles y tuberías. En los últimos años se están utilizando otras tecnologías como la desalinización, la destilación o la ósmosis inversa para obtener agua.

La creciente necesidad de lograr el equilibrio hidrológico que asegure el abastecimiento suficiente de agua a la población, supone un importante desafío que requiere armonizar la disponibilidad natural con las extracciones del recurso mediante el uso eficiente del agua.

En este contexto, dada la incidencia a nivel mundial y las múltiples actividades que directa o indirectamente están afectadas por el déficit de agua, la planificación hidrográfica se plantea como un reto ineludible, para el que la elaboración de predicciones de la demanda de agua urbana con la mayor precisión posible tiene un carácter prioritario.

5.3. Objetivos del problema

El presente estudio debe entenderse, en líneas generales, como una contribución al conocimiento del problema del agua en España, especialmente, en Andalucía, y una aportación al reto de una gestión eficiente de la demanda de agua. El hecho de analizar el problema del agua en un determinado contexto geográfico y social, permite llevar a cabo un análisis más detallado y profundo y adentrarse en las causas concretas y en las consecuencias del

problema, lo que no hubiera sido posible con un estudio mas amplio. Se centra en la predicción del consumo de agua en los núcleos urbanos, que si bien no tienen un gran peso específico desde el punto de vista del volumen consumido en comparación con otros usos, hay que subrayar que el abastecimiento a la población es el “uso” fundamental por su prioridad social, legal y económica y, además, en los últimos años ha experimentado el mayor crecimiento en términos relativos.

Aunque las metas concretas del estudio presentado son esencialmente:

- Un análisis de la evolución del consumo de agua en un área urbana durante el periodo comprendido entre 1984 y 2011.
- El objetivo principal de este trabajo es desarrollar una novedosa metodología de predicción del consumo de agua, que trate de superar las limitaciones de los métodos tradicionales, minimizando el error estimado de predicción, y con la que se complete la información precisa para la toma de decisiones tanto de las administraciones públicas como de las empresas suministradoras.

Para llevar a cabo las metas fijadas, se va desarrollar el estudio de la serie de consumo de agua en áreas urbanas aplicándose los criterios clásicos frente a los criterios propuestos en una primera sección, para comprobar la efectividad de los mismos. Se continuará el estudio desarrollando el ajuste clásico mediante correlogramas, y en un último momento mediante técnicas más sofisticadas como son las redes neuronales. Finalmente se concluirá la aplicación con una sección donde se comparen los mejores resultados obtenidos en cada apartado, con el objetivo de determinar que metodología es capaz de dar una solución más ajustada al problema que se aborda.

5.4. Metodología

Tenemos una serie temporal Y_t donde se recogen los datos mensuales de consumo de agua en una zona urbana de Andalucía, desde el año 1984 hasta 2011, de manera que se tienen $n = 337$ observaciones, de las que $T = 325$ se consideran datos actuales y formaran el grupo de entrenamiento y $n - T = 12$ se definen como datos futuros ($h = 12$), que será el grupo de comprobación. De manera, que para ajustar la serie se utilizara el grupo de entrenamiento, y el grupo de comprobación para comparar las predicciones realizadas con el modelo seleccionado por cada uno de los criterios con los datos reales que se han reservado para la validación. El conjunto de modelos, M , a considerar, serán todos aquellos que cumplan

$$p + q \leq 4$$

$$P + Q \leq 4$$

$$d \leq 1$$

$$D \leq 1$$

lo que hace un total de 324 modelos.

Para ilustrar este estudio vamos a mostrar la representación de la serie a analizar, con objeto de tener una primera visión general sobre el comportamiento de la misma.

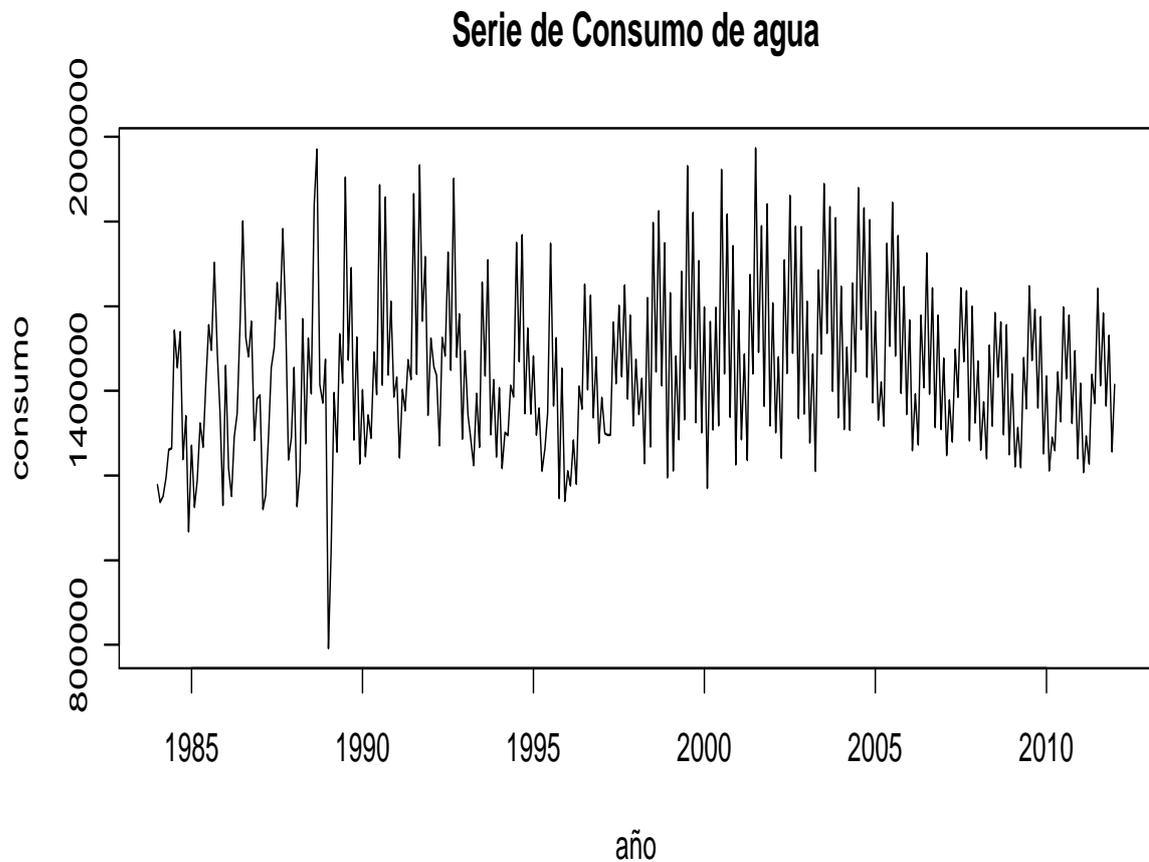


Figura 5.1: Representación gráfica de la serie sobre el consumo de agua

5.4.1. Criterios de Selección de modelos

Para seleccionar un modelo SARIMA adecuado para la serie objeto de estudio, se evalúa la serie en cada uno de los criterios seleccionados considerando cada uno de los 324 modelos, y se determina como modelo el que obtenga el mínimo valor para cada uno de los criterios. Los criterios utilizados serán los mismos que los empleados en el capítulo 4 de simulación.

Criterio	AIC	AIC2	ECM	BIC	BIC2	U_1
Modelo	270	135	270	195	80	270
Criterio	U_2	ECMpre	WMSE	WFMSE	WMAE	WFMAE
Modelo	270	224	269	224	229	155

Tabla 5.1: Modelos seleccionados con cada uno de los criterios.

Se estudiarán por tanto las predicciones obtenidas con los modelos SARIMA(2,0,2)(2,1,2)₁₂, que ha sido seleccionado por los criterios de ajuste AIC y ECM, y por los criterios de predicción U de Theil 1, U de Theil 2, el modelo SARIMA(1,0,1)(1,0,2)₁₂, que ha sido escogido por el criterio de ajuste AIC2, el SARIMA(1,1,1)(2,0,2)₁₂, elegido en este caso por el criterio de ajuste BIC, SARIMA(0,1,1)(1,0,1)₁₂, por el que se ha decantado el criterio de ajuste BIC2, el modelo SARIMA(2,0,0)(1,0,1)₁₂, determinado por los criterios de predicción del ECM y WFMSE propuesto en este trabajo, SARIMA(2,0,2)(2,1,1)₁₂, adoptado en esta ocasión por el criterio de ajuste WMSE, expuesto en este trabajo, el SARIMA(2,0,0)(2,0,0)₁₂ preferido por el criterio de ajuste, WMAE, formulado también en este documento y por último, el modelo SARIMA(1,0,2)(1,1,1)₁₂ al que ha optado el criterio de predicción, WFMAE, enunciado en el capítulo 3.

Para la elección de los modelos se ha definido un α con valores comprendidos entre 0,94 y 0,99, de manera que al calcular sus exponentes el decrecimiento sea lento, ya que esas potencias son los pesos adjudicados a las observaciones, y dado que en la predicción se realizan 12 meses de tener cuidado de elegir un valor de α tal que vaya disminuyendo el peso dado a cada predicción sin llegar a anular a ninguna de ellas.

Se han realizado las predicciones a 12 meses, utilizando cada uno de los modelos anteriores y se han representado las distribuciones de los residuos obtenidos con cada uno de ellos en el conjunto de comprobación, para facilitar la comparación entre las distintas curvas, la serie ha sido tipificada.

Se observa en el gráfico 5.2 que las mejores distribuciones son las cuatro últimas gráficas, sobre todo las correspondientes al modelo SARIMA(2, 0, 0)(1, 0, 1)₁₂ y SARIMA(1,0,2)(1,1,1)₁₂ donde se observa una menor varianza que en el resto con la salvedad de la “cola” de la derecha, que si observamos el gráfico 5.3 comprobaremos que hay un residuo que se dispara, este residuo corresponde con la predicción 9, con lo cuál se considera un valor outlier al resto de la distribución, aunque no es preocupante ya que no se trata de las primeras predicciones que son las que mayor interés nos suscitan.

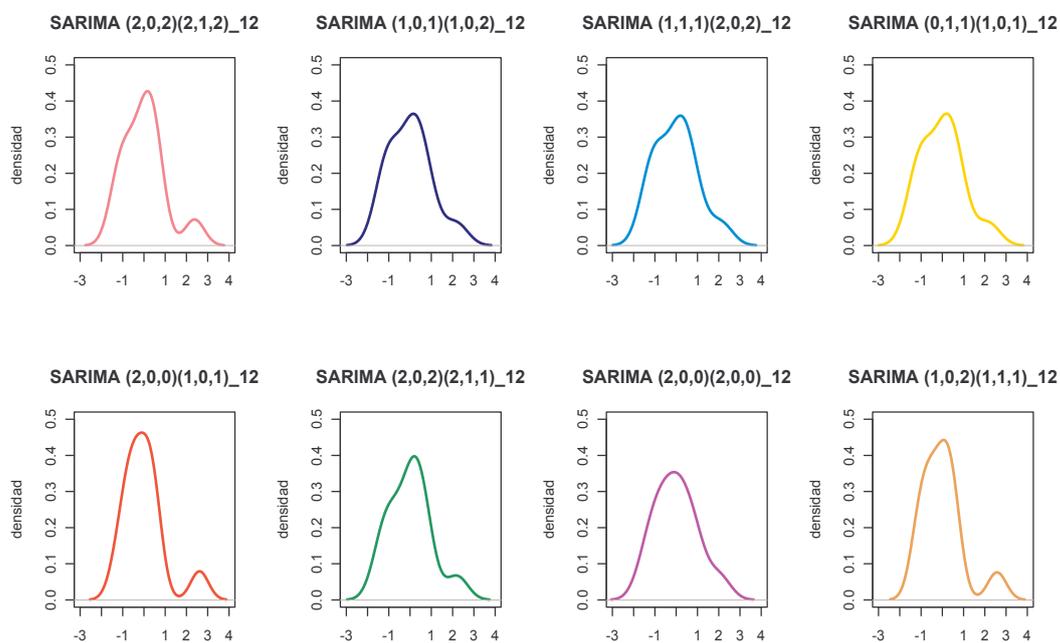


Figura 5.2: Representación de los residuos obtenidos en cada modelo.

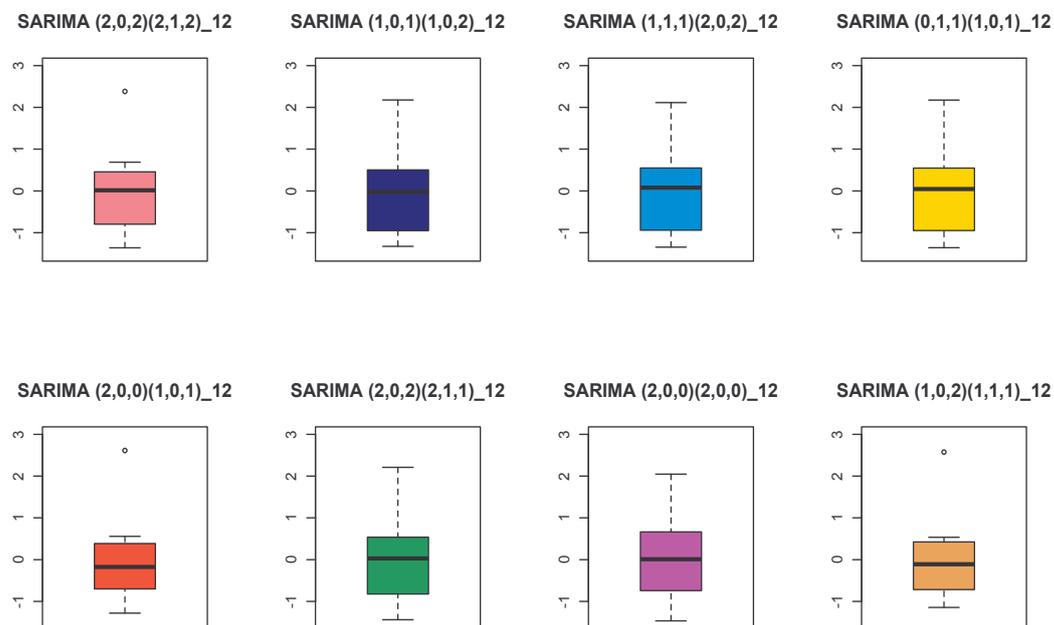


Figura 5.3: Representación del diagrama de cajas y bigotes de los residuos de predicción.

	$(2,0,2)(2,1,2)_{12}$	$(1,0,1)(1,0,2)_{12}$	$(1,1,1)(2,0,2)_{12}$	$(0,1,1)(1,0,1)_{12}$
media	0.313237	0.672788	0.573356	0.504668
ECM	1.630772	1.723538	1.654099	1.771196
ECMpre	0.487196736	1.146151	1.036783	0.888327

	$(2,0,0)(1,0,1)_{12}$	$(2,0,2)(2,1,1)_{12}$	$(2,0,0)(2,0,0)_{12}$	$(1,0,2)(1,1,1)_{12}$
media	-0.031054	0.386064	0.261596	0.180845
ECM	2.003823	1.708121	2.167396	1.778594
ECMpre	0.359088	0.549686	0.458306	0.390951

Tabla 5.2: ECM de ajuste y predicción en cada modelo.

En la tabla 5.2, se puede verificar que los modelos donde se alcanzan las mejores predicciones, son el SARIMA(2, 0, 0)(1, 0, 1)₁₂ que se correspondía con la elección del criterio WFMSE y el SARIMA(1, 0, 2)(1, 1, 1)₁₂ que procede del criterio WFMAE, por tanto, las mejores predicciones se alcanzan en los criterios propuestos en este trabajo ante el caso real de la demanda de agua, siendo además el primero de ellos el que alcanza una menor media, y una expresión del modelo más simple. Podemos observar en el gráfico 5.4, los datos reales del conjunto de comprobación frente a las predicciones realizadas con cada uno de los modelos seleccionados.

Predicciones en el Cjto. de comprobación

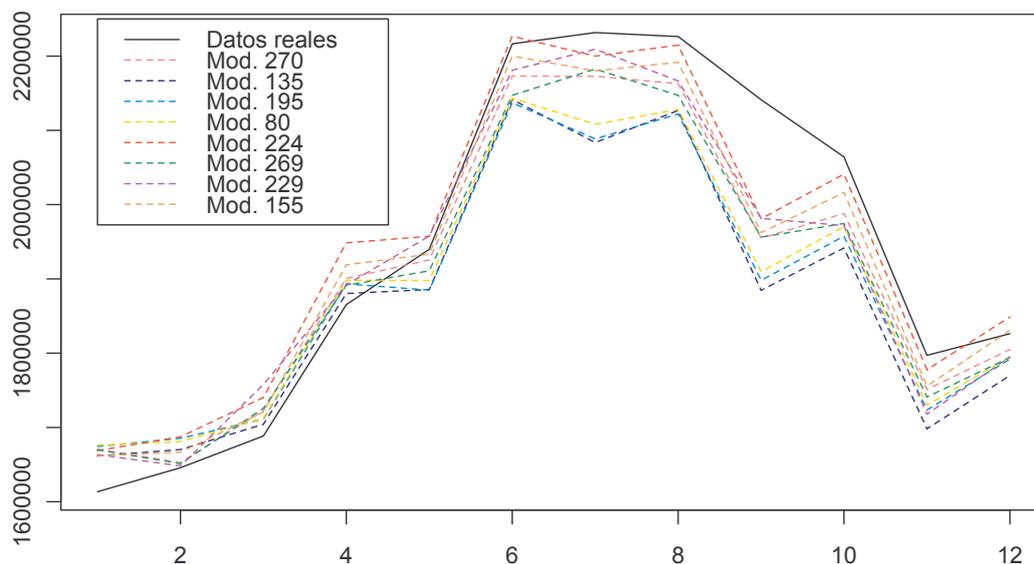


Figura 5.4: Representación de las predicciones obtenidas frente a los valores reales de la serie.

Otra alternativa al caso de la predicción sería comprobar que ocurre con una predicción agregada, de manera que, en vez de realizar las predicciones mes a mes para un año, como se ha hecho, se puede plantear estudiar que ocurre si se realiza una predicción anual, es decir, no diferenciada por meses sino de manera global.

En esta ocasión, lo que se hace es realizar una única predicción que sea el consumo total de agua en el último año, para poder validar la calidad de la predicción se toma el consumo real del último año conocido como conjunto de comprobación, se tiene que el consumo ha sido 232559301.

Realizamos la predicción de ese año con los modelos seleccionados anteriormente y estudiamos los residuos obtenidos.

	$(2,0,2)(2,1,2)_{12}$	$(1,0,1)(1,0,2)_{12}$	$(1,1,1)(2,0,2)_{12}$	$(0,1,1)(1,0,1)_{12}$
predicción	228.8005	224.4859	225.6791	226.5033
residuo	3.758855	8.0734641	6.8802731	6.0560213
	$(2,0,0)(1,0,1)_{12}$	$(2,0,2)(2,1,1)_{12}$	$(2,0,0)(2,0,0)_{12}$	$(1,0,2)(1,1,1)_{12}$
predicción	232.9320	227.9266	229.4202	230.3892
residuo	-0.3726501	4.632779	3.139160	2.170143

Tabla 5.3: Predicciones anuales expresadas en 1000001.

En la tabla 5.3 es fácil comprobar que la predicción más acertada es la realizada con el modelo seleccionado por el criterio WFMSE, no solo porque se consiga alcanzar un residuo menor que en el resto de los casos sino también porque con el resto de modelos la predicción es inferior al consumo real. Con este criterio la predicción supera en 372701 el consumo real, mientras que seleccionar un criterio diferente supondría en el mejor de los casos, 21701431 menos de los reales, por lo que no se hubiera podido satisfacer la demanda de agua. Este hecho puede tener consecuencias graves ante la política de actuación para el aprovechamiento de los recursos naturales que se haya marcado en el año, ya que ha podido generar períodos restricciones innecesarios.

5.4.2. Metodología Box-Jenkins

Para ajustar la serie siguiendo la metodología Box (1976), se han de estimar los parámetros del modelo. Lo primero que se determina son los valores de d y D , que hagan que la serie diferenciada sea estacionaria. En la figura 5.1 se puede observar que la serie presenta estacionalidad a distancia $s = 12$. Por lo que se determina $D = 1$, se necesita diferenciar en la parte estacional. En cuanto a d , podría considerarse como posibles valores el 0 y el 1. Mediante los gráficos de correlación simple (ACF) y parcial (PACF) se determinarían el resto de parámetros.

Se muestra la gráfica para el caso $d = 0$ y $D = 1$ (gráfico 5.5).

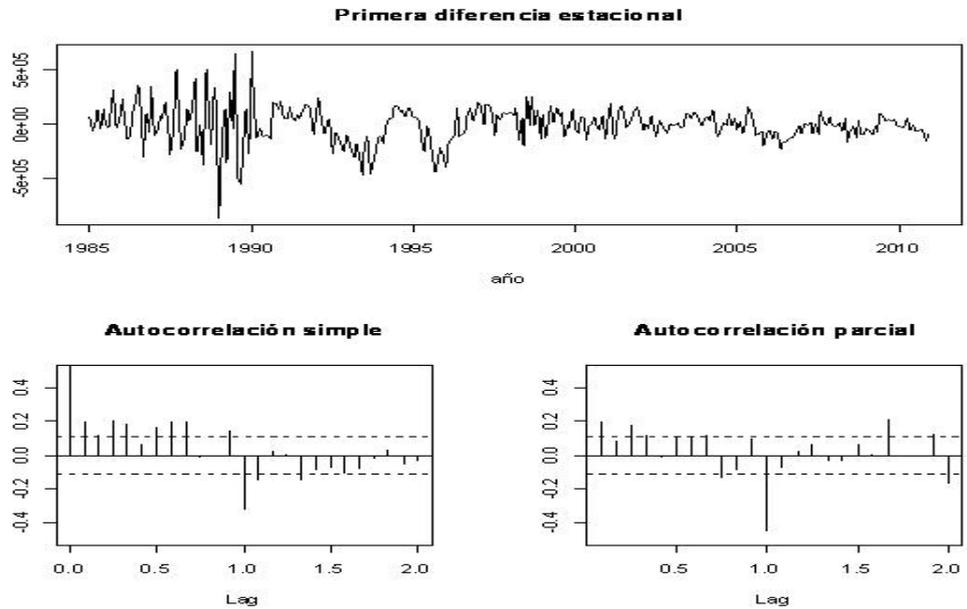


Figura 5.5: Representación de la serie de consumo de agua con una diferencia estacional, gráfico de autocorrelación simple y gráfico de autocorrelación parcial.

A la vista del gráfico 5.5 se determina como modelo un $ARMA(1,1)$ para los datos mensuales, y un $ARMA(1,1)$ para los datos anuales, por lo tanto el modelo ajustado se determina como un $SARIMA(1,0,1) \times (1,1,1)_{12}$. Se analizan ahora las gráficas de las ACF y PACF para el caso $d = 1$ y $D = 1$ (gráfico 5.6).

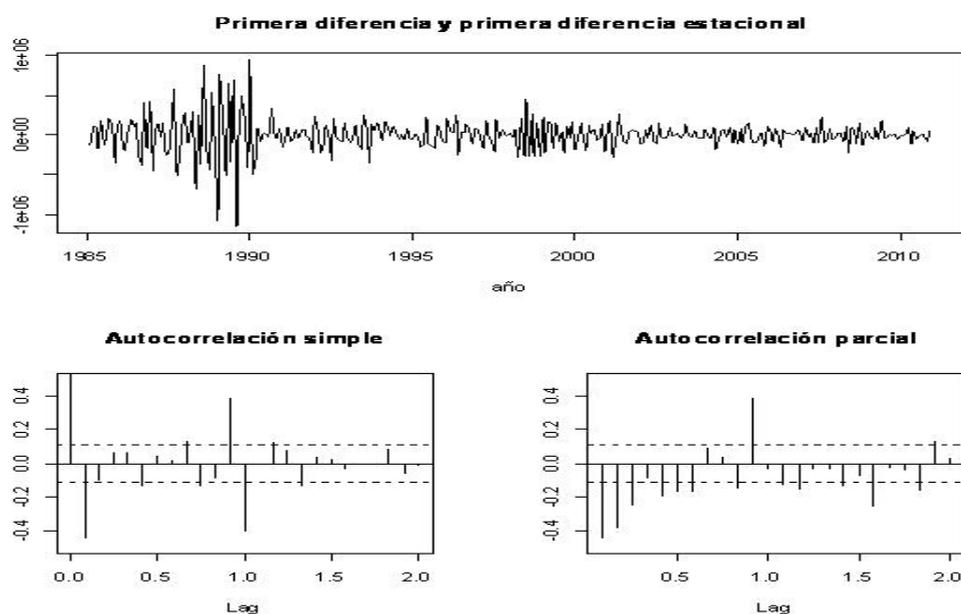


Figura 5.6: Representación de la serie de consumo de agua con la primera diferencia y una diferencia estacional, gráfico de autocorrelación simple y gráfico de autocorrelación parcial.

En este caso, se podría determinar un MA(1) para los datos mensuales y un MA(1) para los datos anuales. El modelo por tanto, en este caso sería un SARIMA(0,1,1)x(0,1,1)₁₂.

Utilizando ambos modelos se realizan las predicciones correspondientes al año 2012, y se estudia el ECM tanto de ajuste como de predicción.

	(1,0,1)(1,1,1) ₁₂	(0,1,1)(0,1,1) ₁₂
ECM	1.663971	1.710733
ECMpre	0.46431	0.466847

Tabla 5.4: ECM de ajuste y predicción con la metodología Box-Jenkins.

En la tabla 5.4 se puede ver como en cuanto a predicción no existen diferencias significativas entre los dos modelos, por tanto, hemos optado por seleccionar el modelo que además tiene el mejor ajuste, es decir, SARIMA(1,0,1)(1,1,1)₁₂.

5.4.3. Redes neuronales

Se aplican en esta ocasión mecanismos de redes neuronales artificiales, éstas buscan reproducir algunas de las capacidades del cerebro imitando su

estructura neuronal mediante un estilo de computación paralelo, distribuido y adaptativo con lo que son capaces de aprender a partir de ejemplos.

El perceptrón multicapa es un aproximador universal de relaciones no lineales entre datos de entrada y salida. La arquitectura típica de un perceptrón multicapa, como se puede ver en la imagen inferior, consiste en tres partes: una capa de entrada, una capa de salida y una o varias capas ocultas, puede verse un esquema en la gráfica 5.7.

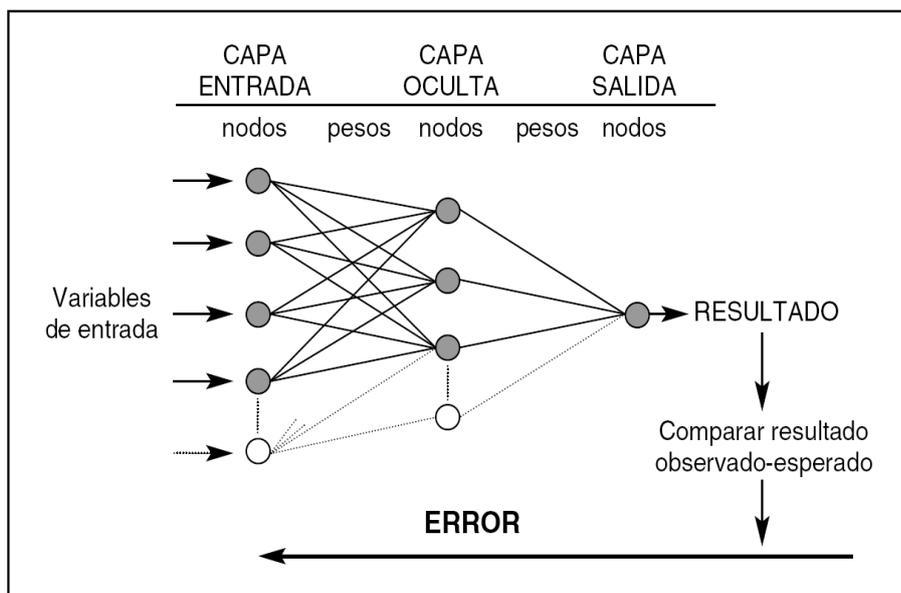


Figura 5.7: Esquema de una red neuronal artificial.

En cuanto al diseño de la arquitectura, el simulador del perceptrón multicapa emplea la función sigmoideal, Chandra y Singh (2004), donde sus parámetros suelen ser determinados por el método de prueba y error, Shepherd (1997). Es necesario remarcar este punto porque no existe una regla exacta que determine los valores óptimos de una red basada en perceptrón multicapa. Esta arquitectura emplea el algoritmo de retropropagación del error (backpropagation), Rumelhart et al. (1986), Rumelhart et al. (2002). Consiste en ajustar los pesos de la red para minimizar el error global, con el objetivo de que la salida de la red sea lo más próxima posible a la salida deseada. Este algoritmo es uno de los puntos característicos del perceptrón multicapa, no existe un proceso automático para determinar el número óptimo de neuronas en esta red neuronal.

La principal dificultad en el entrenamiento de redes de varias capas es encontrar los errores asociados con las capas ocultas; es decir, en las capas que no son la de salida. Esto es debido a que los errores son necesarios para el aprendizaje, para saber cómo modificar los pesos de las neuronas en las ca-

pas ocultas. Rumelhart et al. (1986) formalizaron un método para que una red del tipo perceptrón multicapa aprendiera la asociación que existe entre un conjunto de patrones de entrada y sus salidas correspondientes: método backpropagation error (propagación del error hacia atrás).

Así el funcionamiento del perceptrón multicapa es básicamente: se aplica una entrada cuya salida se conoce, luego se calcula primero la salida de las neuronas de entrada, estas salidas son las entradas de las neuronas de la capa oculta, con estas entradas se calcula la salida de las neuronas ocultas, y con éstas como entrada para las neuronas de salida, se calculan las salidas finales. Esta es la fase hacia delante, por así decirlo, en la red.

Luego se obtiene el error con respecto a la señal deseada y finalmente este error se retro propaga (de atrás hacia delante) modificando los pesos.

No se tratará la parte matemática de cómo se hace esto, pero mencionaremos que las neuronas de la capa oculta usan como regla de propagación, la suma ponderada de las entradas con los pesos sinápticos y sobre esa suma ponderada se aplica una función de transferencia de tipo sigmoide, que es acotada en respuesta. Similarmente en la retropropagación, el error en los nodos de las capas ocultas es proporcional a la sumatoria de los gradientes de los nodos de la capa siguiente pesados por los pesos de conexión.

Se ha llevado a cabo el estudio de la red neuronal perceptrón multicapa, con el software libre R-project, que se viene utilizando a lo largo de toda esta tesis, donde se ha implementado la red utilizando el paquete nnet (apéndice B), y a la que se le ha pedido que seleccione el número de capas ocultas (H) entre 1 y 12, se muestran en la tabla 5.3 los resultados diferentes obtenidos en función del número de capas ocultas utilizadas, para no hacer muy extensa la salida se muestra los valores para 1, 2 ó 3 capas ocultas, que son las que mejores resultados consiguen.

H	ECM	ECMpre
1	1.864918	0.696281
2	1.588207	0.737552
3	1.441918	0.955957

Tabla 5.5: ECM de ajuste y predicción en Redes Neuronales con distintas capas ocultas.

Se observa en la tabla 5.5 como para el caso de buscar un buen ajuste el número de capas ocultas óptimo serían 3, mientras que la mayor capacidad predictiva se consigue con una sola capa oculta. En nuestro problema la red seleccionada sería la que considera una sola capa oculta, al ser esta la que mejor predice.

Las redes neuronales artificiales han ido cobrando una gran importancia en muchas aplicaciones reales, Allende (2002). Los resultados empíricos muestran que bajo ciertas condiciones, incluso la red neuronal más simple es capaz de aproximar cualquier función medible de cualquier grado, Zhang et al. (1998). El primer modelo de red neuronal fue propuesto por McCulloch y Pitts (1943) en términos de un modelo computacional de actividad nerviosa, otros modelos han sido estudiados posteriormente, Elman (1990). Estos modelos generalmente utilizan la red de perceptrón multicapa. Las redes neuronales han encontrado muchas aplicaciones con éxito en la visión artificial, en el procesamiento de señales e imágenes, reconocimiento del habla y de caracteres, sistemas expertos, análisis de imágenes médicas, control remoto, control de robots e inspección industrial, y cada vez más se aplican a nuevos campos, como el de las ciencias económicas, Santana (2006), Patiño (2012). En relación al problema que nos ocupa, son frecuentes el empleo de redes neuronales para predecir series de tiempo no lineal, Kajitani et al. (2005). No obstante, en la realidad, existen series de tiempo que no son absolutamente lineales, ni no lineales, es decir, que tienen componentes lineales y tienen a su vez, componentes no lineales, lo que hacen a menudo, que el modelo se vea afectado por irregularidades y eventos infrecuentes, que hacen que el pronóstico sea un problema difícil y complejo. Para abordar este tipo de series varios investigadores han ido proponiendo no sólo nuevos modelos ARIMA, sino de redes neuronales, Santana (2006), e incluso modelos híbridos que utilizan tanto modelos ARIMA como redes neuronales, como por ejemplo, Caridad y OcerÁN (2011), Zhang (2003). Tseng et al. (2002) combinan series temporales estacionarias con redes neuronales de tipo feedforward. Estos modelos híbridos obtienen buenos resultados en muchos de los campos en los que han sido aplicados, ya que parecen corregir parcialmente los errores de especificación en modelos ARIMA. Este es un estudio que se quiere abordar en una futura línea de investigación.

5.5. Comparación de metodologías

Se ha desarrollado en el capítulo anterior diversas metodologías existentes en la literatura. En esta ocasión lo que se va a desarrollar es el estudio de los resultados obtenidos con cada una de ellas para establecer la metodología más eficaz que de solución a nuestro objetivo, que es predecir de la manera más precisa posible con el objetivo de poder desarrollar mecanismos de actuación eficaces para poder satisfacer la demanda en un área determinada. Se presentan en la tabla 5.6 los valores del ECM de ajuste y predicción obtenidos en las tres metodologías desarrolladas; Box-Jenkins, modelos SARIMA y redes neuronales.

	$(1,0,1)(1,1,1)_{12}$	WFMSE	Red Neuronal
ECM	1.663971	1.415565	1.864918
ECMpre	0.46431	0.359088	0.696281

Tabla 5.6: ECM de ajuste y predicción con la metodología Box-Jenkins, Sarima y Redes Neuronales.

Observando los resultados recogidos en la tabla 5.6, podemos determinar de manera clara la efectividad y el buen funcionamiento de los criterios de predicción que se proponen en este trabajo. Como puede comprobarse, la capacidad predictiva y el error que se comete con nuestros criterios es menor que en el caso de desarrollar una metodología basada en correlogramas, que además de ser obsoleta es más laboriosa y no tiene posibilidad de ser automatizada, y como se vió anteriormente también obtiene mejores resultados que utilizando cualquier otro de los criterios clásicos estudiados. Pero además, en nuestro ejemplo, estos criterios nuevos de predicción consiguen alcanzar mejores resultados ante metodologías más complejas, como es el caso de las redes. Además en esta ocasión los resultados de ajuste también son mejores utilizando el modelo determinado por el criterio WFMSE.

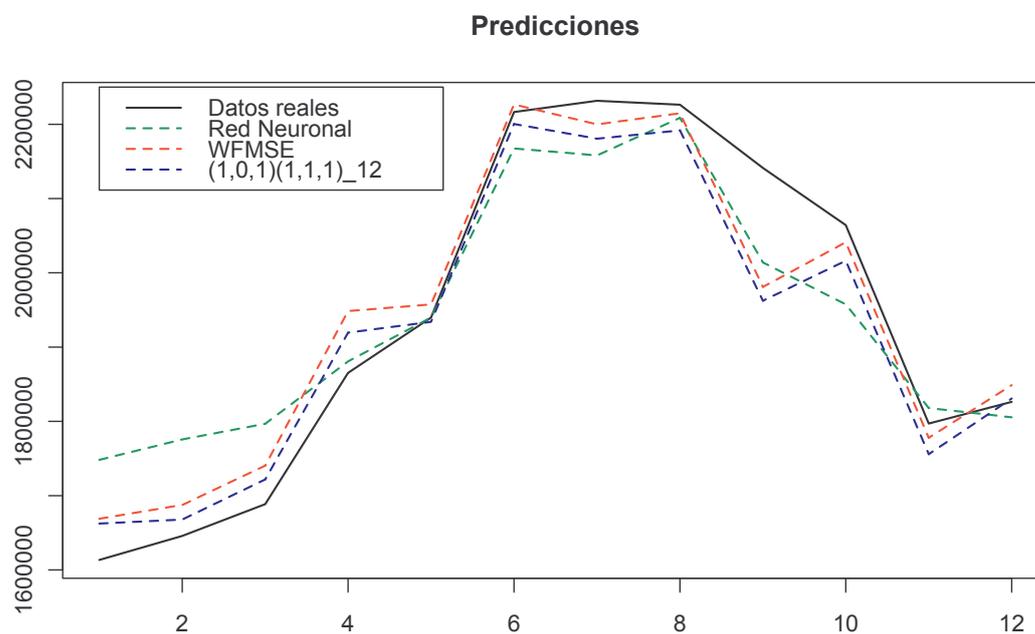


Figura 5.8: Representación de las predicciones con los distintos métodos.

Capítulo 6

Conclusiones y futuras líneas de investigación

Nunca se alcanza la verdad total, ni nunca se está totalmente alejado de ella.

Aristóteles. (Filósofo lógico y científico de la antigua Grecia)

Son muchos los estudios de existentes en la literatura sobre el estudio de las series temporales, y sobre todo el estudio sobre la modelización de las mismas y el ajuste de un modelo adecuado, son muchos los desarrollos teóricos que existen, sin embargo, en el campo de las series temporales, es muy difícil llevar a la práctica todos esos estudios con éxito.

En primer lugar, y a pesar de que los criterios de ajuste y predicción que se han expuesto en este trabajo, son más que conocidos, estudiados y utilizados, hemos de empezar diciendo que su capacidad de detectar el modelo generador de una serie temporal obtenida, es, en general, muy pobre, baraja unos porcentajes de aciertos muy bajos, en el caso de los criterios de ajuste clásicos, llegando en el mejor de los casos a obtener un 60% y tan sólo con uno de los criterios, con el BIC2, en el caso de los criterios de predicción aún esos porcentajes son menores, más aún en la mayoría de los casos nulos, es decir, no son capaces de detectar el modelo generador de las series.

En cuanto al procedimiento a desarrollar para alcanzar de la manera más beneficiosa el objetivo establecido, se ha buscado en un primer lugar un modelo “óptimo” para la predicción en cada uno de los modelos estudiados, tomando la decisión de seleccionar aquel que minimiza el criterio empleado en un grupo de series procedentes de un mismo modelo conocido y determinar a su vez el mínimo de estos modelos, se decir, tomando el mínimo del mínimo de un grupo de series con la misma estructura, sin embargo, mediante esta actuación no se han conseguido discriminar los modelos ni elegir

un criterio que se comporte mejor ante esta situación.

En segundo lugar (aunque no se ha mostrado en este documento), se ha intentado llevar a cabo la decisión de seleccionar el modelo que más se repetía entre el grupo de series evaluadas, de manera que en cada serie se tomaba el modelo que minimizaba al criterio y el modelo final elegido era la moda de éstos, sin embargo, los resultados al respecto han sido peores si cabe, que en el proceso anterior, además se nos presenta el problema de que en un grupo de series puede existir más de una moda, lo que complica el desarrollo y no nos lleva a un buen camino.

En un tercer momento se ha estudiado la importancia de que el criterio sea capaz de determinar correctamente a las series evaluadas y si ese procedimiento nos llevaba a obtener unos resultados mejores, sin embargo, detectar las series no ha sido fácil, no obstante, en los casos en los que se ha llevado a cabo esta discriminación los resultados obtenidos en cuanto a la predicción no se han visto mejorados notablemente en ninguno de los casos.

Por último, se ha desarrollado un mecanismo donde la importancia del modelo que se elegía toma un segundo plano, es decir, no nos interesa tanto el modelo utilizado para la predicción sino que lo que se persigue es obtener unos buenos índices de la misma, por tanto, se ha considerado para cada serie independientemente de su procedencia el modelo que en cada caso ha establecido el criterio empleado, y este procedimiento en todos los procesos estudiados nos ha conducido al éxito de manera determinante.

En cuanto a los modelos ARIMA, los mejores resultados en cuanto a la predicción se han alcanzado siguiendo el procedimiento de elegir en cada caso em modelo determinado por el criterio, y los criterios que han conseguido, con diferencia al resto, al resto las mejores predicciones han sido los criterios de predicción que se han desarrollado en el capítulo 3. No obstante, se ha comprobado como la primera de las predicciones realizadas es bastante exacta pero el incremento en el error cometido a lo largo del horizonte de predicción aumenta muy rápidamente.

En cuanto a los modelos SARIMA, en esta ocasión las mejores predicciones también han sido alcanzadas siguiendo el mismo procedimiento que en el caso anterior y de nuevo los criterios de predicción propuestos en este trabajo han alcanzado los mejores resultados con diferencia del resto de criterios evaluados, además, en estos modelos se ha conseguido en la mayoría de las situaciones que el crecimiento de los errores cometidos a lo largo de la predicción si no se han mantenido estables, el crecimiento entre unos y otros ha sido mucho más lento.

En cuanto a los criterios propuestos, hemos llegado a obtener una expresión más completa que la que se tiene en los criterios clásicos existentes en la literatura, ya que en nuestro caso se contempla el coste que supone

cometer errores de predicción en instantes próximos a la última observación conocida, dada la poca capacidad de reacción que se tiene para poder reparar ese error, de manera que, se ha penalizado con más dureza el error que se comete en el primer instante de la predicción que en el último, ya que éste siempre puede ser actualizado conforme se acerque el momento. Para ello se ha utilizado un valor α determinado en cada caso según sea el objetivo perseguido, en nuestro caso los mejores valores de α han sido los comprendidos entre 0,8 y 0,99, tomando en cada caso aquel que alcance el valor inferior en el criterio con respecto al resto, este valor debe ser elegido con cuidado ya que es el que determina el peso que se le da al error cometido, de manera que debemos tomar un α tal que disminuya lentamente al ser elevado a los distintos exponentes definidos.

En las pruebas realizadas se ha comprobado que en el cambio de proceder de una u otra manera los criterios que han sufrido mas modificaciones han sido precisamente los que se han propuesto en este trabajo, esto es debido a que en la elección del modelo si nos ceñíamos a uno único (primer procedimiento) la frecuencia con la que este modelo era elegido era muy baja, mientras que en el resto de criterios esa frecuencia era mayor, por tanto, al cambiar el procedimiento y dejar que cada serie prediga con el modelo establecido para ella, éstos criterios se ven muy afectados beneficiosamente, mientras que en los demás el cambio es menor.

Futuras líneas de investigación

Tras la realización de este trabajo van surgiendo distintas ideas que no pueden ser atajadas al mismo tiempo, dado a que hay que darle un final en algún momento al estudio, lo que no significa que se le de fin a la investigación.

- Una de las propuestas que tenemos para futuras investigaciones, es estudiar la causa por la cuál a lo largo de todas las simulaciones ha ocurrido que a medida que hemos conseguido precisar en la predicción se ha disparado el ajuste. Por lo que se quiere estudiar la capacidad de conseguir un buen ajuste.
- Por otro lado, nos gustaría ver que ocurre en el caso de predicciones agregadas y desagregadas, para comprobar si el comportamiento se mantiene en esos casos o por el contrario sufre modificaciones.
- Otra posible línea, que nos gustaría desarrollar es el empleo de otro tipo de modelos para predicción, las redes neuronales, como se ha visto en la aplicación, ya que es conocido el buen funcionamiento de las mismas en otro tipo de estudios, con el objetivo de adaptarlas a nuestro problema.
- Combinar ambos métodos de manera que se puedan desarrollar modelos híbridos de predicción, de los que también es sabido su buen funcionamiento en predicción en estudios ya realizados en econometría.

Apéndice A

Algunas nociones sobre R-project

*Las únicas personas que tienen algo que
temer de software libre son aquellos
cuyos productos tienen un valor aún
menor.*

David Emery (Escritor independiente)

La información que se detalla a continuación, ha sido extraída de la página oficial del programa, donde se puede comprobar y ampliar la misma. <http://www.r-project.org/>

A.1. ¿Qué es R?

R es un lenguaje y un entorno para computación y gráficos estadísticos. Es un proyecto GNU ¹, que es similar al lenguaje y entorno S que se ha desarrollado en los Laboratorios Bell (antes AT and T, ahora Lucent Technologies) por John Chambers y otros. R puede ser considerado como una implementación diferente de S. Hay algunas diferencias importantes, pero mucho código escrito para S funciona sin perfeccionar en R.

R ofrece una gran variedad de estadísticas (modelos lineales y no lineales, clásicos contrastes estadísticos, análisis de series de tiempo, clasificación, clustering, ...) y las técnicas gráficas, y es altamente extensible. El lenguaje S es a menudo el vehículo de elección para la investigación en metodología

¹GNU es un sistema operativo similar a Unix que es software libre y respeta su libertad. El Proyecto GNU se inició en 1984 para desarrollar el sistema GNU. El nombre "GNU" (que significa "ñ" en inglés) es un acrónimo recursivo de "¡GNU No es Unix!". Se construyen a partir de un conjunto de aplicaciones, bibliotecas y herramientas de programación, además de un programa para alojar recursos e interactuar con el hardware, denominado núcleo.

estadística, y R proporciona una ruta de código abierto para la participación en esa actividad.

Uno de los puntos fuertes de R es la facilidad con la que pueden ser diseñados los estudios estadísticos, incluidos los símbolos y fórmulas matemáticas cuando sea necesario. R dispone de valores predeterminados para las opciones de diseño de menor importancia en los gráficos, pero el usuario mantiene el control completo.

R está disponible como software libre bajo los términos de Free Software Foundation's GNU General Public License en forma de código fuente. Se compila y se ejecuta en una amplia variedad de plataformas UNIX y sistemas similares (incluyendo FreeBSD y Linux), Windows y MacOS.

A.2. El Entorno R

R es un conjunto integrado de servicios de software para la manipulación de datos, cálculo y representación gráfica. incluye

- un manejo eficaz de datos y almacenamiento,
- un conjunto de operadores para los cálculos de matrices, matrices en particular,
- una gran colección coherente e integrado de herramientas intermedias para el análisis de datos,
- facilidades gráficas para el análisis y visualización de datos ya sea en pantalla o en papel, y
- un bien desarrollado, el lenguaje de programación sencillo y eficaz que incluye condicionales, bucles, funciones definidas por el usuario recursivas y las instalaciones de entrada y de salida.

El término “entorno” se pretende que lo caracterizan como un sistema totalmente planificado y coherente, en lugar de una acumulación incremental de herramientas muy específicas y poco flexibles, como suele ser el caso con otro software de análisis de datos.

R, como S, está diseñado en torno a un lenguaje de programación real, y permite a los usuarios añadir funcionalidad adicional mediante la definición de nuevas funciones. Gran parte del sistema es en sí mismo escrito en el dialecto R de S, lo que hace que sea fácil para los usuarios a seguir las elecciones hechas algorítmicos. Para las tareas computacionalmente intensivas, C, C++ y Fortran se puede vincular y llamar en tiempo de ejecución. Los usuarios avanzados pueden escribir código C para manipular objetos de investigación directamente.

Muchos usuarios piensan de R como un sistema de estadísticas. Preferimos pensar en él de un entorno en el que las técnicas estadísticas se aplican. R puede ser extendido (fácilmente) a través de paquetes. Hay alrededor de ocho paquetes suministrados con la distribución de R y muchos más están disponibles a través de la familia CRAN de sitios de Internet que cubren una gama muy amplia de la estadística moderna.

R tiene su propio formato similar al *l*átex documentación, que se utiliza para suministrar documentación completa, tanto en línea en un número de formatos y en papel.

A.3. Fundación R

La Fundación R es una entidad sin ánimo de lucro que trabaja en el interés público. Ha sido fundada por los miembros del R Development Core Team con el fin de

- Prestar apoyo para el proyecto de I y otras innovaciones en la computación estadística. Creemos que la investigación se ha convertido en una herramienta valiosa y madura y nos gustaría asegurar su desarrollo, y el desarrollo de las futuras innovaciones en software para la investigación estadística y computacional.
- Proporcionar un punto de referencia para los individuos, instituciones o empresas comerciales que quieran apoyar o interactuar con la comunidad R development.
- Mantener y administrar el copyright de los software de R y documentación.

R es una parte oficial del Free Software Foundation's GNU project y la Fundación R tiene objetivos similares a otras fundaciones de software libre como la Fundación Apache o la Fundación GNOME.

Entre los objetivos de la Fundación R son el soporte de desarrollo continuo de R, la exploración de nuevos métodos de enseñanza y capacitación de computación estadística y la organización de reuniones y conferencias orientadas a la computación estadística.

Apéndice B

Código Fuente en R

*La mejor forma de predecir el futuro es
implementarlo*

David Heinemeier Hansson (Creador de
Ruby on Rails)

B.1. El Programa de Simulación

```
library(lattice)
library(CombMSC)
library(tseries)
library(zoo)
library(forecast)
num-iter<-200
alpha<-c(0.8, 0.85, 0.90, 0.95, 0.96, 0.97, 0.98, 0.99)
cont<-matrix(0,num-iter,8+4*length(alpha))
mod-selec<-matrix(0,num-iter,8+4*length(alpha))
minimo<-matrix(999,num-iter,8+4*length(alpha)) pos<-matrix(0,num-iter,8+4*length(alpha))
aux<-0
for(iter in 1:num-iter)
{ cat("iteracion ",iter" \ n")
ar1<- 0
ar2<- 0
d0<- 0
ma1<- 0
ma2<- 0
sar1<-0
sar2<-0
D0<-0
sma1<-0
sma2<-0
```

```

s<-0
n0<-180
p0<-2
q0<-2
P0<-2
Q0<-2
ts <- arima.sim(list(order=c(p0,d0,q0),ar=c(ar1,ar2),ma=c(ma1,ma2), sea-
sonal=list(order=c(P0,D0,Q0),sar=c(sar1,sar2),sma=c(ma1,ma2))), period=s),n=n0)
vp<- 0:2
vd<- 0:1
vq<- 0:2
vP<- 0:2
vD<- 0:1
vQ<- 0:2
numod<- length(vp)*length(vd)*length(vq)*length(vP)*length(vD)*length(vQ)
entrena<-ts[0:173]
mat-mod<-matrix(0,numod,6)
i<-1
for (p in vp)
for (d in vd)
for (q in vq)
for(P in vP)
for(D in vD)
for(Q in vQ)
{mat-mod[i,]<-c(p,d,q,P,D,Q)
i<-i+1}
log-mod<-matrix(FALSE,numod,1)
if (ar2!=0) pp<-2 else
{if (ar1!=0) pp<-1 else pp<-0}
dd<-d0
if (ma2!=0) qq<-2 else
{if (ma1!=0) qq<-1 else qq<-0}
if (sar2!=0) PP<-2 else
{if (sar1!=0) PP<-1 else PP<-0}
DD<-D0
if (sma2!=0) QQ<-2 else
{if (sma1!=0) QQ<-1 else QQ<-0}
ord<-c(pp,dd,qq,PP,DD,QQ)
i<-1
for (p in vp)
for (d in vd)
for (q in vq)
for(P in vP)
for(D in vD)

```

```

for(Q in vQ)
{if (p==pp & d==dd & q==qq & P==PP & D==DD & Q==QQ) log-
mod[i]=TRUE
i<-i+1 }
resul<- matrix(99999999,numod,8+4*length(alpha))
W-MSE<- matrix(NA,1,length(alpha))
Wf-MSE<- matrix(NA,1,length(alpha))
W-MAE<- matrix(NA,1,length(alpha))
Wf-MAE<- matrix(NA,1,length(alpha))
s<-0
h<-7
T<-173
rownames(resul)<- 1:numod
observado<-ts[174:180]
modelos<-list()
i<-0
for (p in vp)
for (d in vd)
for (q in vq)
for(P in vP)
for(D in vD)
for(Q in vQ) {
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")
i<-i+1
modelos[[i]]<-modelo
rownames(resul)[i]<- paste("ARIMA(",p," ",d," ",q," ",P," ",D," ",Q,")",sep="")
if(sum(diag(modelo$var.coef)<0)==0) {
res=modelo$residuals
serienueva<- predict(modelo, n.ahead=7)
pred=serienueva$pred
l<- -(length(entrena)/2)-(length(entrena)/2)*log(2*pi)-(length(entrena)/2)*
log((1/length(entrena))*sum(res2))
AIC<-2*(1/length(entrena))+2*(length(modelo$coef)/length(entrena))
BIC<-2*(1/length(entrena))+(length(modelo$coef)/length(entrena))*log(length(entrena))
AIC2<- log(modelo$sigma2)+length(modelo$coef)*(2/length(entrena))
BIC2<- log(modelo$sigma2)+length(modelo$coef)*(log(length(entrena))/length(entrena))
theil1<- sqrt(((1/length(entrena))*sum(res2)/(((1/length(entrena))*sum(entrena2))+
+((1/length(entrena))*sum(res2))))
theil2<- sqrt(((1/length(entrena))*sum(res2)/((1/length(entrena))*sum(entrena2))))
ECMpre<-(1/length(observado))*sum((observado-pred)2)
ECM<-(1/length(entrena))*sum(res2)
mi<-p+d+P*s+D*s
exp1<-(T-(mi+1)):0

```

```

exp2<- 0:(h-1)
for (j in 1:length(alpha)){
error1 <- (((1-alpha[j])/(1-alpha[j]^(T-mi))))*sum(alpha[j]^(exp1)*res[(mi+1):T]^2)
W-MSE[j] <- error1
error2 <- (((1-alpha[j])/(1-alpha[j]^h))*sum(alpha[j]^(exp2)*(observado-pred)^2)
Wf-MSE[j] <- error2
error3 <- (((1-alpha[j])/(1-alpha[j]^(T-mi))))*sum(alpha[j]^(exp1)*abs(res[(mi+1):T]))
W-MAE[j] <- error3
error4 <- (((1-alpha[j])/(1-alpha[j]^h))*sum(alpha[j]^(exp2)*abs(observado-pred))
Wf-MAE[j] <- error4}
resul[i,1] <- AIC
resul[i,2] <- AIC2
resul[i,3] <- ECM
resul[i,4] <- BIC
resul[i,5] <- BIC2
resul[i,6] <- theil1
resul[i,7] <- theil2
resul[i,8] <- ECMpre
resul[i,9:(8+length(alpha))] <- W-MSE
resul[i,(8+length(alpha)+1):(8+2*length(alpha))] <- Wf-MSE
resul[i,(8+2*length(alpha)+1):(8+3*length(alpha))] <- W-MAE
resul[i,(8+3*length(alpha)+1):(8+4*length(alpha))] <- Wf-MAE } }
aux<-aux+1
for (j in 1:(8+4*length(alpha)))
{if (resul[which(log-mod==TRUE),j]==min(resul[,j])) cont[iter,j] <- 1}
for (j in 1:(8+4*length(alpha)))
{mod-selec[iter,j] <- which.min(resul[,j])}
for (j in 1:(8+4*length(alpha)))
{minimo[iter, j] <- min(resul[,j])}
100*apply(cont,2,mean)
which(log-mod==TRUE)
mat-mod[which(log-mod==TRUE),]
moda=function(x)
{m1=sort(table(x),decreasing=TRUE)
moda=names(m1[m1==m1[1]])
moda=as.numeric(moda)
return(moda)}
x<-matrix(0,1,ncol(cont))
y<-matrix(0,1,ncol(cont))
for (j in 1:(ncol(cont)))
{for (i in 1:(nrow(cont)))
{if (cont[i,j]==1) x[j] <- x[j]+minimo[i,j]
else x[j] <- x[j]} } for (j in 1:(ncol(cont)))
{for (i in 1:(nrow(cont)))

```

```

{if (cont[i,j]==0) y[j]<-y[j]+minimo[i,j]
else y[j]<-y[j]}} h <- list()
for (i in 1:ncol(mod-selec))
{ H<- abs(sort(table(mod-selec[,i])))
h[[i]] <- H} minimo-minimo <- matrix(0, 2, ncol(minimo))
for (j in 1:ncol(minimo))
{ m <- min(minimo[,j])
minimo-minimo[1,j] <- m
a <- which.min(minimo[,j])
minimo-minimo[2,j] <- mod-selec[a,j]}
vec-moda<-matrix(0,1,8+4*length(alpha))
for (j in 1:(8+4*length(alpha)))
{vec-moda[j]<-moda(mod-selec[,j])}
valores-cri<-matrix(99, 1, 8+4*length(alpha))
for (n in 1:(8+4*length(alpha)))
{valores-cri[1, n]<-resul[vec-moda[1,n],n]}
predaic<-matrix(0, num-iter, h)
resiaic<-matrix(0,num-iter,h)
predaic2<-matrix(0, num-iter, 7)
resiaic2<-matrix(0,num-iter,7)
predecn<-matrix(0, num-iter, 7)
resiecn<-matrix(0, num-iter, h)
predbic<-matrix(0, num-iter, h)
resibic<-matrix(0, num-iter, h)
predbic2<-matrix(0, num-iter, h)
resibic2<-matrix(0, num-iter, h)
predu1<-matrix(0, num-iter, h)
resiu1<-matrix(0, num-iter, h)
predu2<-matrix(0, num-iter, h)
resiu2<-matrix(0, num-iter, h)
predecmpre<-matrix(0, num-iter, h)
resiecmpre<-matrix(0, num-iter, h)
predwmse<-matrix(0, num-iter, h)
resiwmse<-matrix(0, num-iter, h)
predwfmse<-matrix(0, num-iter, h)
resiwfmse<-matrix(0, num-iter, h)
predwmae<-matrix(0, num-iter, h)
resiwmae<-matrix(0, num-iter, h)
predwfmae<-matrix(0, num-iter, h)
resiwfmae<-matrix(0, num-iter, h)
Datos<-ts
elegido<-minimo-minimo[2,]
ECM<-matrix(0,24, ncol(Datos))
for(iter in 1:ncol(Datos))

```

```

{ts <- ts(Datos[,iter], star=c(1984,1), frequency = 7)
entrena<-ts[0:173]
s<-7
h<-7
T<-173
observado<-ts[174:180]
p<-mat-mod[mod-selec[iter,1],1]
d<-mat-mod[mod-selec[iter,1],2]
q<-mat-mod[mod-selec[iter,1],3]
P<-mat-mod[mod-selec[iter,1],4]
D<-mat-mod[mod-selec[iter,1],5]
Q<-mat-mod[mod-selec[iter,1],6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")
res=modelo$residuals
serienueva<- predict(modelo, n.ahead=7)
predaic[iter,]=serienueva$pred
resiaic[iter,]<-observado-predaic[iter,]
ecm-pre<-(1/length(observado))*sum(resiaic[iter,]^2)
ecm-aju<-(1/length(entrena))*sum(res^2)
ECM[1,iter]<-ecm-aju
ECM[2,iter]<-ecm-pre}
for(iter in 1:ncol(Datos))
{ts <- ts(Datos[,iter], star=c(1984,1), frequency = 7)
entrena<-ts[0:173]
s<-7
h<-7
T<-173
observado<-ts[174:180]
p<-mat-mod[mod-selec[iter,2],1]
d<-mat-mod[mod-selec[iter,2],2]
q<-mat-mod[mod-selec[iter,2],3]
P<-mat-mod[mod-selec[iter,2],4]
D<-mat-mod[mod-selec[iter,2],5]
Q<-mat-mod[mod-selec[iter,2],6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")
res=modelo$residuals
serienueva<- predict(modelo, n.ahead=7)
predaic2[iter,]=serienueva$pred
resiaic2[iter,]<-observado-predaic2[iter,]
ecm-pre<-(1/length(observado))*sum(resiaic2[iter,]^2)
ecm-aju<-(1/length(entrena))*sum(res^2)
ECM[3,iter]<-ecm-aju

```

```
ECM[4,iter]<-ecm-pre}
for(iter in 1:ncol(Datos))
{ts <- ts(Datos[,iter], star=c(1984,1), frequency = 7)
entrena<-ts[0:173]
s<-7
h<-7
T<-173
observado<-ts[174:180]
p<-mat-mod[mod-selec[iter,3],1]
d<-mat-mod[mod-selec[iter,3],2]
q<-mat-mod[mod-selec[iter,3],3]
P<-mat-mod[mod-selec[iter,3],4]
D<-mat-mod[mod-selec[iter,3],5]
Q<-mat-mod[mod-selec[iter,3],6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")
res=modelo$residuals
serienueva<- predict(modelo, n.ahead=7)
predec[iter,]=serienueva$pred
resiecm[iter,]<-observado-predec[iter,]
ecm-pre<-(1/length(observado))*sum(resiecm[iter,]^2)
ecm-aju<-(1/length(entrena))*sum(res^2)
ECM[5,iter]<-ecm-aju
ECM[6,iter]<-ecm-pre}
for(iter in 1:ncol(Datos))
{ts <- ts(Datos[,iter], star=c(1984,1), frequency = 7)
entrena<-ts[0:173]
s<-7
h<-7
T<-173
observado<-ts[174:180]
p<-mat-mod[mod-selec[iter,4],1]
d<-mat-mod[mod-selec[iter,4],2]
q<-mat-mod[mod-selec[iter,4],3]
P<-mat-mod[mod-selec[iter,4],4]
D<-mat-mod[mod-selec[iter,4],5]
Q<-mat-mod[mod-selec[iter,4],6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")
res=modelo$residuals
serienueva<- predict(modelo, n.ahead=7)
predbic[iter,]=serienueva$pred
resibic[iter,]<-observado-predbic[iter,]
ecm-pre<-(1/length(observado))*sum(resibic[iter,]^2)
```

```

ecm-aju<-(1/length(entrena))*sum(res2)
ECM[7,iter]<-ecm-aju
ECM[8,iter]<-ecm-pre}
for(iter in 1:ncol(Datos))
{ts <- ts(Datos[,iter], star=c(1984,1), frequency = 7)
entrena<-ts[0:173]
s<-7
h<-7
T<-173
observado<-ts[174:180]
p<-mat-mod[mod-selec[iter,5],1]
d<-mat-mod[mod-selec[iter,5],2]
q<-mat-mod[mod-selec[iter,5],3]
P<-mat-mod[mod-selec[iter,5],4]
D<-mat-mod[mod-selec[iter,5],5]
Q<-mat-mod[mod-selec[iter,5],6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS") res=modelo$residuals
serienueva<- predict(modelo, n.ahead=7)
predbic2[iter,]=serienueva$pred
resibic2[iter,]<-observado-predbic2[iter,]
ecm-pre<-(1/length(observado))*sum(resibic2[iter,]2)
ecm-aju<-(1/length(entrena))*sum(res2)
ECM[9,iter]<-ecm-aju
ECM[10,iter]<-ecm-pre}
for(iter in 1:ncol(Datos))
{ts <- ts(Datos[,iter], star=c(1984,1), frequency = 7)
entrena<-ts[0:173]
s<-7
h<-7
T<-173
observado<-ts[174:180]
p<-mat-mod[mod-selec[iter,6],1]
d<-mat-mod[mod-selec[iter,6],2]
q<-mat-mod[mod-selec[iter,6],3]
P<-mat-mod[mod-selec[iter,6],4]
D<-mat-mod[mod-selec[iter,6],5]
Q<-mat-mod[mod-selec[iter,6],6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")
res=modelo$residuals
serienueva<- predict(modelo, n.ahead=7)
predu1[iter,]=serienueva$pred
resiu1[iter,]<-observado-predu1[iter,]

```

```
ecm-pre<-(1/length(observado))*sum(resiu1[iter,]2)
ecm-aju<-(1/length(entrena))*sum(res2)
ECM[11,iter]<-ecm-aju
ECM[12,iter]<-ecm-pre}
for(iter in 1:ncol(Datos))
{ts <- ts(Datos[,iter], star=c(1984,1), frequency = 7)
entrena<-ts[0:173]
s<-7
h<-7
T<-173
observado<-ts[174:180]
p<-mat-mod[mod-selec[iter,7],1]
d<-mat-mod[mod-selec[iter,7],2]
q<-mat-mod[mod-selec[iter,7],3]
P<-mat-mod[mod-selec[iter,7],4]
D<-mat-mod[mod-selec[iter,7],5]
Q<-mat-mod[mod-selec[iter,7],6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")
res=modelo$residuals
serienueva<- predict(modelo, n.ahead=7)
predu2[iter,]=serienueva$pred
resiu2[iter,]<-observado-predu2[iter,] ecm-pre<-(1/length(observado))*sum(resiu2[iter,]2)
ecm-aju<-(1/length(entrena))*sum(res2)
ECM[13,iter]<-ecm-aju
ECM[14,iter]<-ecm-pre}
for(iter in 1:ncol(Datos))
{ts <- ts(Datos[,iter], star=c(1984,1), frequency = 7)
entrena<-ts[0:173]
s<-7
h<-7
T<-173
observado<-ts[174:180]
p<-mat-mod[mod-selec[iter,8],1]
d<-mat-mod[mod-selec[iter,8],2]
q<-mat-mod[mod-selec[iter,8],3]
P<-mat-mod[mod-selec[iter,8],4]
D<-mat-mod[mod-selec[iter,8],5]
Q<-mat-mod[mod-selec[iter,8],6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")
res=modelo$residuals
serienueva<- predict(modelo, n.ahead=7)
predecmpre[iter,]=serienueva$pred
```

```

resiecmpre[iter,]<-observado-predecmpre[iter,]
ecm-pre<-(1/length(observado))*sum(resiecmpre[iter,]2)
ecm-aju<-(1/length(entrena))*sum(res2)
ECM[15,iter]<-ecm-aju
ECM[16,iter]<-ecm-pre}
for(iter in 1:ncol(Datos))
{ts <- ts(Datos[,iter], star=c(1984,1), frequency = 7)
entrena<-ts[0:173]
s<-7
h<-7
T<-173
observado<-ts[174:180]
p<-mat-mod[mod-selec[iter,which.min(minimo[iter,9:16]),1]
d<-mat-mod[mod-selec[iter,which.min(minimo[iter,9:16]),2]
q<-mat-mod[mod-selec[iter,which.min(minimo[iter,9:16]),3]
P<-mat-mod[mod-selec[iter,which.min(minimo[iter,9:16]),4]
D<-mat-mod[mod-selec[iter,which.min(minimo[iter,9:16]),5]
Q<-mat-mod[mod-selec[iter,which.min(minimo[iter,9:16]),6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")
res=modelo$residuals
serienueva<- predict(modelo, n.ahead=7)
predwmse[iter,]=serienueva$pred
resiwmse[iter,]<-observado-predwmse[iter,]
ecm-pre<-(1/length(observado))*sum(resiwmse[iter,]2)
ecm-aju<-(1/length(entrena))*sum(res2)
ECM[17,iter]<-ecm-aju
ECM[18,iter]<-ecm-pre}
for(iter in 1:ncol(Datos))
{ts <- ts(Datos[,iter], star=c(1984,1), frequency = 7)
entrena<-ts[0:173]
s<-7
h<-7
T<-173
observado<-ts[174:180]
p<-mat-mod[mod-selec[iter,which.min(minimo[iter,17:24]),1]
d<-mat-mod[mod-selec[iter,which.min(minimo[iter,17:24]),2]
q<-mat-mod[mod-selec[iter,which.min(minimo[iter,17:24]),3]
P<-mat-mod[mod-selec[iter,which.min(minimo[iter,17:24]),4]
D<-mat-mod[mod-selec[iter,which.min(minimo[iter,17:24]),5]
Q<-mat-mod[mod-selec[iter,which.min(minimo[iter,17:24]),6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")
res=modelo$residuals

```

```

serienueva<- predict(modelo, n.ahead=7)
predwfmse[iter,]=serienueva$pred
resiwmse[iter,]<-observado-predwfmse[iter,]
ecm-pre<-(1/length(observado))*sum(resiwmse[iter,]^2)
ecm-aju<-(1/length(entrena))*sum(res^2)
ECM[19,iter]<-ecm-aju
ECM[20,iter]<-ecm-pre}
for(iter in 1:ncol(Datos))
{ts <- ts(Datos[,iter], star=c(1984,1), frequency = 7)
entrena<-ts[0:173]
s<-7
h<-7
T<-173
observado<-ts[174:180]
p<-mat-mod[mod-selec[iter,which.min(minimo[iter,25:32]),1]
d<-mat-mod[mod-selec[iter,which.min(minimo[iter,25:32]),2]
q<-mat-mod[mod-selec[iter,which.min(minimo[iter,25:32]),3]
P<-mat-mod[mod-selec[iter,which.min(minimo[iter,25:32]),4]
D<-mat-mod[mod-selec[iter,which.min(minimo[iter,25:32]),5]
Q<-mat-mod[mod-selec[iter,which.min(minimo[iter,25:32]),6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")
res=modelo$residuals
serienueva<- predict(modelo, n.ahead=7)
predwmae[iter,]=serienueva$pred
resiwmae[iter,]<-observado-predwmae[iter,]
ecm-pre<-(1/length(observado))*sum(resiwmae[iter,]^2)
ecm-aju<-(1/length(entrena))*sum(res^2)
ECM[21,iter]<-ecm-aju
ECM[22,iter]<-ecm-pre}
for(iter in 1:ncol(Datos))
{ts <- ts(Datos[,iter], star=c(1984,1), frequency = 7)
entrena<-ts[0:173]
s<-7
h<-7
T<-173
observado<-ts[174:180]
p<-mat-mod[mod-selec[iter,which.min(minimo[iter,33:40]),1]
d<-mat-mod[mod-selec[iter,which.min(minimo[iter,33:40]),2]
q<-mat-mod[mod-selec[iter,which.min(minimo[iter,33:40]),3]
P<-mat-mod[mod-selec[iter,which.min(minimo[iter,33:40]),4]
D<-mat-mod[mod-selec[iter,which.min(minimo[iter,33:40]),5]
Q<-mat-mod[mod-selec[iter,which.min(minimo[iter,33:40]),6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =

```

```

s), include.mean=FALSE, method="CSS")
res=modelo$residuals
serienueva<- predict(modelo, n.ahead=7)
predwfmae[iter,]=serienueva$pred
resiwfmae[iter,]<-observado-predwfmae[iter,]
ecm-pre<-(1/length(observado))*sum(resiwfmae[iter,]^2)
ecm-aju<-(1/length(entrena))*sum(res^2)
ECM[23,iter]<-ecm-aju
ECM[24,iter]<-ecm-pre}

```

B.2. Programa para la Aplicación

B.2.1. Predicciones a partir de los criterios

```

library(lattice)
library(CombMSC)
library(tseries)
library(zoo)
library(forecast)
ts<-ts(agua/100000, star=c(1984,1), frequency = 12)
alpha<-c(0.94, 0.95, 0.96, 0.97, 0.98, 0.99)
cont<-matrix(0,1,8+4*length(alpha))
pos<-matrix(0,1,8+4*length(alpha))
aux<-0
vp<- 0:2
vd<- 0:1
vq<- 0:2
vP<- 0:2
vD<- 0:1
vQ<- 0:2
numod<- length(vp)*length(vd)*length(vq)*length(vP)*length(vD)*length(vQ)
resul<- matrix(99999999,numod,8)
resul2<- matrix(99999999,numod,4*length(alpha))
W-MSE<- matrix(NA,1,length(alpha))
Wf-MSE<- matrix(NA,1,length(alpha))
W-MAE<- matrix(NA,1,length(alpha))
Wf-MAE<- matrix(NA,1,length(alpha))
s<-12
h<-12
T<-325
entrena<-ts[0:325]
rownames(resul)<- 1:numod

```

```

observado<-ts[326:337]
modelos<-list()
prediccion <- list()
i<-0
for (p in vp)
for (d in vd)
for (q in vq)
for(P in vP)
for(D in vD)
for(Q in vQ){
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")
i<-i+1
modelos[[i]]<-modelo
rownames(resul)[i] <- paste(."ARIMA(",p," ",d," ",q," ",P," ",D," ",Q,")",sep=
)
if(sum(diag(modelo$var.coef)<0)==0)
{res=modelo$residuals
serienueva<- predict(modelo, n.ahead=12)
prediccion[[i]]<-serienueva
pred=serienueva$pred
l<- -(length(entrena)/2)-(length(entrena)/2)*log(2*pi)-(length(entrena)/2)*
log((1/length(entrena))*sum(res2))
AIC<-2*(1/length(entrena))+2*(length(modelo$coef)/length(entrena))
BIC<-2*(1/length(entrena))+length(modelo$coef)/length(entrena)*log(length(entrena))
AIC2<- log(modelo$sigma2)+length(modelo$coef)*(2/length(entrena))
BIC2<- log(modelo$sigma2)+length(modelo$coef)*(log(length(entrena))/length(entrena))
ecmpre<-(1/length(observado))*sum((observado-pred)2)
theil1<- sqrt((1/length(entrena))*sum(res2)/(((1/length(entrena))*sum(entrena2))+
((1/length(entrena))*sum(res2))))
theil2<- sqrt((1/length(entrena))*sum(res2)/((1/length(entrena))*sum(entrena2)))
ecm<-(1/length(entrena))*sum(res2)
mi<-p+d+P*s+D*s
exp1<-(length(entrena)-1):0
exp2<- 0:(h-1)
for (j in 1:length(alpha))
{error1 <- ((1-alpha[j])/(1-alpha[j]T-mi))*sum(alpha[j]exp1*res2)
W-MSE[j] <- error1
error2 <- ((1-alpha[j])/(1-alpha[j]h))*sum(alpha[j]exp2*(observado-pred)2)
Wf-MSE[j] <- error2
error3 <-((1-alpha[j])/(1-alpha[j]T-mi))*sum(alpha[j]exp1*abs(res))
W-MAE[j] <- error3 error4 <- ((1-alpha[j])/(1-alpha[j]h))*sum(alpha[j]exp2*abs(observado-
pred))
Wf-MAE[j] <- error4}

```

```

resul[i,1]<- AIC
resul[i,2]<- AIC2
resul[i,3]<- ecm
resul[i,4]<-BIC
resul[i,5]<-BIC2
resul[i,6]<- theil1
resul[i,7]<- theil2
resul[i,8]<- ecmpre
resul2[i,1:(length(alpha))] <- W-MSE
resul2[i,(length(alpha)+1):(2*length(alpha))]<-Wf-MSE
resul2[i,(2*length(alpha)+1):(3*length(alpha))] <- W-MAE
resul2[i,(3*length(alpha)+1):(4*length(alpha))] <- Wf-MAE }}
minimo<- matrix (0, 1, 8)
for (h in 1:8)
{minimo[h] <- min(resul[ ,h])}
modelo<- matrix (0, 1, 8)
for (m in 1:8)
{modelo[m] <- which.min(resul[ ,m])}
minimo2<- matrix (0, 1, 4*length(alpha))
for (h in 1: (4*length(alpha)))
{minimo2[h] <- min(resul2[ ,h])}
modelo2<- matrix (0, 1, 4*length(alpha))
for (m in 1: (4*length(alpha)))
{modelo2[m] <- which.min(resul2[ ,m])}
modelo<-c(270,135,195,80,224,269,229,155)
m<-matrix(c(2,0,2,2,1,2,1,0,1,1,0,2,1,1,1,2,0,2,0,1,1,1,0,1,2,0,0,1,
0,1,2,0,2,2,1,1,2,0,0,2,0,0,1,0,2,1,1,1), nrow=8, ncol=6, byrow=T)
res<-matrix(0, 8, 325)
pred<-matrix(0, 8, 12)
residuos<-matrix(0, 8, 12)
ECM<-matrix(0, 2, 8)
for (iter in 1:8)
{entrena<-ts[1:325]
s<-12
h<-12
T<-325
observado<-ts[326:337] p<-m[iter,1]
d<-m[iter,2]
q<-m[iter,3]
P<-m[iter,4]
D<-m[iter,5]
Q<-m[iter,6]
modelo<-arima(entrena, order=c(p,d,q),seas=list(order=c(P,D,Q), period =
s), include.mean=FALSE, method="CSS")

```

```

res[iter,]=modelo$residuals
serienueva<- predict(modelo, n.ahead=12)
pred[iter,]=serienueva$pred
residuos[iter,]<-observado-pred[iter,]
ecm-pre<-(1/length(observado))*sum(residuos[iter,]2)
ecm-aju<-(1/length(entrena))*sum(res[iter,]2)
ECM[1,iter]<-ecm-aju
ECM[2,iter]<-ecm-pre}
apply(residuos,1,mean)

```

B.2.2. Red neuronal

```

library(nnet)
library(rminer)
agua<-ts(agua/100000, start=1984, end=2012, frequency=12)
ts.plot(agua)
matriz<-CasesSeries(agua,1:12)
entre<- 1:313
test<- 314:325
red<- fit(y ,data=matriz[entre,],model="ml",scale="all",search=1:12)
red@object
predicttest<- predict(red,matriz[test,])
predicent<- predict(red,matriz[entre,])
residuostest<- agua[326:337]-predicttest
residuosenentre<- agua[13:325]-predicent
summary(residuostest)
summary(residuostest2)
summary(residuosenentre)
summary(residuosenentre2)
pred1<-predicttest
obs<-agua[326:337]
ecm<-(1/12)*sum((obs-pred1)2)
datosentre<- agua[13:325]
S<- sum(residuosenentre2)
numpar<- red@object$mlp$nunits
n<- length(datosentre)
ecmpre<-(1/n)*sum(residuosenentre2)

```


Apéndice C

Espacio de Modelos

*Son vanas y están cargadas de errores
las ciencias que no han nacido del
experimento*

Leonardo Da Vinci (Notable polímata
del renacimiento italiano)

C.1. Conjunto de los M Modelos

Número	Modelo	Número	Modelo
1	SARIMA(0,0,0)(0,0,0)	2	SARIMA(0,0,0)(0,0,1)
3	SARIMA(0,0,0)(0,0,2)	4	SARIMA(0,0,0)(0,1,0)
5	SARIMA(0,0,0)(0,1,1)	6	SARIMA(0,0,0)(0,1,2)
7	SARIMA(0,0,0)(1,0,0)	8	SARIMA(0,0,0)(1,0,1)
9	SARIMA(0,0,0)(1,0,2)	10	SARIMA(0,0,0)(1,1,0)
11	SARIMA(0,0,0)(1,1,1)	12	SARIMA(0,0,0)(1,1,2)
13	SARIMA(0,0,0)(2,0,0)	14	SARIMA(0,0,0)(2,0,1)
15	SARIMA(0,0,0)(2,0,2)	16	SARIMA(0,0,0)(2,1,0)
17	SARIMA(0,0,0)(2,1,1)	18	SARIMA(0,0,0,2,1,2)
19	SARIMA(0,0,1)(0,0,0)	20	SARIMA(0,0,1)(0,0,1)
21	SARIMA(0,0,1)(0,0,2)	22	SARIMA(0,0,1)(0,1,0)
23	SARIMA(0,0,1)(0,1,1)	24	SARIMA(0,0,1)(0,1,2)
25	SARIMA(0,0,1)(1,0,0)	26	SARIMA(0,0,1)(1,0,1)
27	SARIMA(0,0,1)(1,0,2)	28	SARIMA(0,0,1)(1,1,0)
29	SARIMA(0,0,1)(1,1,1)	30	SARIMA(0,0,1)(1,1,2)
31	SARIMA(0,0,1)(2,0,0)	32	SARIMA(0,0,1)(2,0,1)
33	SARIMA(0,0,1)(2,0,2)	34	SARIMA(0,0,1)(2,1,0)
35	SARIMA(0,0,1)(2,1,1)	36	SARIMA(0,0,1)(2,1,2)
37	SARIMA(0,0,2)(0,0,0)	38	SARIMA(0,0,2)(0,0,1)

Número	Modelo	Número	Modelo
39	SARIMA(0,0,2)(0,0,2)	40	SARIMA(0,0,2)(0,1,0)
41	SARIMA(0,0,2)(0,1,1)	42	SARIMA(0,0,2)(0,1,2)
43	SARIMA(0,0,2)(1,0,0)	44	SARIMA(0,0,2)(1,0,1)
45	SARIMA(0,0,2)(1,0,2)	46	SARIMA(0,0,2)(1,1,0)
47	SARIMA(0,0,2)(1,1,1)	48	SARIMA(0,0,2)(1,1,2)
49	SARIMA(0,0,2)(2,0,0)	50	SARIMA(0,0,2)(2,0,1)
51	SARIMA(0,0,2)(2,0,2)	52	SARIMA(0,0,2)(2,1,0)
53	SARIMA(0,0,2)(2,1,1)	54	SARIMA(0,0,2)(2,1,2)
55	SARIMA(0,1,0)(0,0,0)	56	SARIMA(0,1,0)(0,0,1)
57	SARIMA(0,1,0)(0,0,2)	58	SARIMA(0,1,0)(0,1,0)
59	SARIMA(0,1,0)(0,1,1)	60	SARIMA(0,1,0)(0,1,2)
61	SARIMA(0,1,0)(1,0,0)	62	SARIMA(0,1,0)(1,0,1)
63	SARIMA(0,1,0)(1,0,2)	64	SARIMA(0,1,0)(1,1,0)
65	SARIMA(0,1,0)(1,1,1)	66	SARIMA(0,1,0)(1,1,2)
67	SARIMA(0,1,0)(2,0,0)	68	SARIMA(0,1,0)(2,0,1)
69	SARIMA(0,1,0)(2,0,2)	70	SARIMA(0,1,0)(2,1,0)
71	SARIMA(0,1,0)(2,1,1)	72	SARIMA(0,1,0)(2,1,2)
73	SARIMA(0,1,1)(0,0,0)	74	SARIMA(0,1,1)(0,0,1)
75	SARIMA(0,1,1)(0,0,2)	76	SARIMA(0,1,1)(0,1,0)
77	SARIMA(0,1,1)(0,1,1)	78	SARIMA(0,1,1)(0,1,2)
79	SARIMA(0,1,1)(1,0,0)	80	SARIMA(0,1,1)(1,0,1)
81	SARIMA(0,1,1)(1,0,2)	82	SARIMA(0,1,1)(1,1,0)
83	SARIMA(0,1,1)(1,1,1)	84	SARIMA(0,1,1)(1,1,2)
85	SARIMA(0,1,1)(2,0,0)	86	SARIMA(0,1,1)(2,0,1)
87	SARIMA(0,1,1)(2,0,2)	88	SARIMA(0,1,1)(2,1,0)
89	SARIMA(0,1,1)(2,1,1)	90	SARIMA(0,1,1)(2,1,2)
91	SARIMA(0,1,2)(0,0,0)	92	SARIMA(0,1,2)(0,0,1)
93	SARIMA(0,1,2)(0,0,2)	94	SARIMA(0,1,2)(0,1,0)
95	SARIMA(0,1,2)(0,1,1)	96	SARIMA(0,1,2)(0,1,2)
97	SARIMA(0,1,2)(1,0,0)	98	SARIMA(0,1,2)(1,0,1)
99	SARIMA(0,1,2)(1,0,2)	100	SARIMA(0,1,2)(1,1,0)
101	SARIMA(0,1,2)(1,1,1)	102	SARIMA(0,1,2)(1,1,2)
103	SARIMA(0,1,2)(2,0,0)	104	SARIMA(0,1,2)(2,0,1)
105	SARIMA(0,1,2)(2,0,2)	106	SARIMA(0,1,2)(2,1,0)
107	SARIMA(0,1,2)(2,1,1)	108	SARIMA(0,1,2)(2,1,2)
109	SARIMA(1,0,0)(0,0,0)	110	SARIMA(1,0,0)(0,0,1)
111	SARIMA(1,0,0)(0,0,2)	112	SARIMA(1,0,0)(0,1,0)
113	SARIMA(1,0,0)(0,1,1)	114	SARIMA(1,0,0)(0,1,2)
115	SARIMA(1,0,0)(1,0,0)	116	SARIMA(1,0,0)(1,0,1)
117	SARIMA(1,0,0)(1,0,2)	118	SARIMA(1,0,0,1,1,0)
119	SARIMA(1,0,0)(1,1,1)	120	SARIMA(1,0,0)(1,1,2)

Número	Modelo	Número	Modelo
121	SARIMA(1,0,0)(2,0,0)	122	SARIMA(1,0,0)(2,0,1)
123	SARIMA(1,0,0)(2,0,2)	124	SARIMA(1,0,0)(2,1,0)
125	SARIMA(1,0,0)(2,1,1)	126	SARIMA(1,0,0)(2,1,2)
127	SARIMA(1,0,1)(0,0,0)	128	SARIMA(1,0,1)(0,0,1)
129	SARIMA(1,0,1)(0,0,2)	130	SARIMA(1,0,1)(0,1,0)
131	SARIMA(1,0,1)(0,1,1)	132	SARIMA(1,0,1)(0,1,2)
133	SARIMA(1,0,1)(1,0,0)	134	SARIMA(1,0,1)(1,0,1)
135	SARIMA(1,0,1)(1,0,2)	136	SARIMA(1,0,1)(1,1,0)
137	SARIMA(1,0,1)(1,1,1)	138	SARIMA(1,0,1)(1,1,2)
139	SARIMA(1,0,1)(2,0,0)	140	SARIMA(1,0,1)(2,0,1)
141	SARIMA(1,0,1)(2,0,2)	142	SARIMA(1,0,1)(2,1,0)
143	SARIMA(1,0,1)(2,1,1)	144	SARIMA(1,0,1)(2,1,2)
145	SARIMA(1,0,2)(0,0,0)	146	SARIMA(1,0,2)(0,0,1)
147	SARIMA(1,0,2)(0,0,2)	148	SARIMA(1,0,2)(0,1,0)
149	SARIMA(1,0,2)(0,1,1)	150	SARIMA(1,0,2)(0,1,2)
151	SARIMA(1,0,2)(1,0,0)	152	SARIMA(1,0,2)(1,0,1)
153	SARIMA(1,0,2)(1,0,2)	154	SARIMA(1,0,2)(1,1,0)
155	SARIMA(1,0,2)(1,1,1)	156	SARIMA(1,0,2)(1,1,2)
157	SARIMA(1,0,2)(2,0,0)	158	SARIMA(1,0,2)(2,0,1)
159	SARIMA(1,0,2)(2,0,2)	160	SARIMA(1,0,2)(2,1,0)
161	SARIMA(1,0,2)(2,1,1)	162	SARIMA(1,0,2)(2,1,2)
163	SARIMA(1,1,0)(0,0,0)	164	SARIMA(1,1,0)(0,0,1)
165	SARIMA(1,1,0)(0,0,2)	166	SARIMA(1,1,0)(0,1,0)
167	SARIMA(1,1,0)(0,1,1)	168	SARIMA(1,1,0)(0,1,2)
169	SARIMA(1,1,0)(1,0,0)	170	SARIMA(1,1,0)(1,0,1)
171	SARIMA(1,1,0)(1,0,2)	172	SARIMA(1,1,0)(1,1,0)
173	SARIMA(1,1,0)(1,1,1)	174	SARIMA(1,1,0)(1,1,2)
175	SARIMA(1,1,0)(2,0,0)	176	SARIMA(1,1,0)(2,0,1)
177	SARIMA(1,1,0)(2,0,2)	178	SARIMA(1,1,0)(2,1,0)
179	SARIMA(1,1,0)(2,1,1)	180	SARIMA(1,1,0)(2,1,2)
181	SARIMA(1,1,1)(0,0,0)	182	SARIMA(1,1,1)(0,0,1)
183	SARIMA(1,1,1)(0,0,2)	184	SARIMA(1,1,1)(0,1,0)
185	SARIMA(1,1,1)(0,1,1)	186	SARIMA(1,1,1)(0,1,2)
187	SARIMA(1,1,1)(1,0,0)	188	SARIMA(1,1,1)(1,0,1)
189	SARIMA(1,1,1)(1,0,2)	190	SARIMA(1,1,1)(1,1,0)
191	SARIMA(1,1,1)(1,1,1)	192	SARIMA(1,1,1)(1,1,2)
193	SARIMA(1,1,1)(2,0,0)	194	SARIMA(1,1,1)(2,0,1)
195	SARIMA(1,1,1)(2,0,2)	196	SARIMA(1,1,1)(2,1,0)
197	SARIMA(1,1,1)(2,1,1)	198	SARIMA(1,1,1)(2,1,2)
199	SARIMA(1,1,2)(0,0,0)	200	SARIMA(1,1,2)(0,0,1)
201	SARIMA(1,1,2)(0,0,2)	202	SARIMA(1,1,2)(0,1,0)

Número	Modelo	Número	Modelo
203	SARIMA(1,1,2)(0,1,1)	204	SARIMA(1,1,2)(0,1,2)
205	SARIMA(1,1,2)(1,0,0)	206	SARIMA(1,1,2)(1,0,1)
207	SARIMA(1,1,2)(1,0,2)	208	SARIMA(1,1,2)(1,1,0)
209	SARIMA(1,1,2)(1,1,1)	210	SARIMA(1,1,2)(1,1,2)
211	SARIMA(1,1,2)(2,0,0)	212	SARIMA(1,1,2)(2,0,1)
213	SARIMA(1,1,2)(2,0,2)	214	SARIMA(1,1,2)(2,1,0)
215	SARIMA(1,1,2)(2,1,1)	216	SARIMA(1,1,2)(2,1,2)
217	SARIMA(2,0,0)(0,0,0)	218	SARIMA(2,0,0)(0,0,1)
219	SARIMA(2,0,0)(0,0,2)	220	SARIMA(2,0,0)(0,1,0)
221	SARIMA(2,0,0)(0,1,1)	222	SARIMA(2,0,0)(0,1,2)
223	SARIMA(2,0,0)(1,0,0)	224	SARIMA(2,0,0)(1,0,1)
225	SARIMA(2,0,0)(1,0,2)	226	SARIMA(2,0,0)(1,1,0)
227	SARIMA(2,0,0)(1,1,1)	228	SARIMA(2,0,0)(1,1,2)
229	SARIMA(2,0,0)(2,0,0)	230	SARIMA(2,0,0)(2,0,1)
231	SARIMA(2,0,0)(2,0,2)	232	SARIMA(2,0,0)(2,1,0)
233	SARIMA(2,0,0)(2,1,1)	234	SARIMA(2,0,0)(2,1,2)
235	SARIMA(2,0,1)(0,0,0)	236	SARIMA(2,0,1)(0,0,1)
237	SARIMA(2,0,1)(0,0,2)	238	SARIMA(2,0,1)(0,1,0)
239	SARIMA(2,0,1)(0,1,1)	240	SARIMA(2,0,1)(0,1,2)
241	SARIMA(2,0,1)(1,0,0)	242	SARIMA(2,0,1)(1,0,1)
243	SARIMA(2,0,1)(1,0,2)	244	SARIMA(2,0,1)(1,1,0)
245	SARIMA(2,0,1)(1,1,1)	246	SARIMA(2,0,1)(1,1,2)
247	SARIMA(2,0,1)(2,0,0)	248	SARIMA(2,0,1)(2,0,1)
249	SARIMA(2,0,1)(2,0,2)	250	SARIMA(2,0,1)(2,1,0)
251	SARIMA(2,0,1)(2,1,1)	252	SARIMA(2,0,1)(2,1,2)
253	SARIMA(2,0,2)(0,0,0)	254	SARIMA(2,0,2)(0,0,1)
255	SARIMA(2,0,2)(0,0,2)	256	SARIMA(2,0,2)(0,1,0)
257	SARIMA(2,0,2)(0,1,1)	258	SARIMA(2,0,2)(0,1,2)
259	SARIMA(2,0,2)(1,0,0)	260	SARIMA(2,0,2)(1,0,1)
261	SARIMA(2,0,2)(1,0,2)	262	SARIMA(2,0,2)(1,1,0)
263	SARIMA(2,0,2)(1,1,1)	264	SARIMA(2,0,2)(1,1,2)
265	SARIMA(2,0,2)(2,0,0)	266	SARIMA(2,0,2)(2,0,1)
267	SARIMA(2,0,2)(2,0,2)	268	SARIMA(2,0,2)(2,1,0)
269	SARIMA(2,0,2)(2,1,1)	270	SARIMA(2,0,2)(2,1,2)
271	SARIMA(2,1,0)(0,0,0)	271	SARIMA(2,1,0)(0,0,1)
273	SARIMA(2,1,0)(0,0,2)	274	SARIMA(2,1,0)(0,1,0)
275	SARIMA(2,1,0)(0,1,1)	276	SARIMA(2,1,0)(0,1,2)
277	SARIMA(2,1,0)(1,0,0)	278	SARIMA(2,1,0)(1,0,1)
279	SARIMA(2,1,0)(1,0,2)	280	SARIMA(2,1,0)(1,1,0)
281	SARIMA(2,1,0)(1,1,1)	282	SARIMA(2,1,0)(1,1,2)
283	SARIMA(2,1,0)(2,0,0)	284	SARIMA(2,1,0)(2,0,1)

Número	Modelo	Número	Modelo
285	SARIMA(2,1,0)(2,0,2)	286	SARIMA(2,1,0)(2,1,0)
287	SARIMA(2,1,0)(2,1,1)	288	SARIMA(2,1,0)(2,1,2)
289	SARIMA(2,1,1)(0,0,0)	290	SARIMA(2,1,1)(0,0,1)
291	SARIMA(2,1,1)(0,0,2)	292	SARIMA(2,1,1)(0,1,0)
293	SARIMA(2,1,1)(0,1,1)	294	SARIMA(2,1,1)(0,1,2)
295	SARIMA(2,1,1)(1,0,0)	296	SARIMA(2,1,1)(1,0,1)
297	SARIMA(2,1,1)(1,0,2)	298	SARIMA(2,1,1)(1,1,0)
299	SARIMA(2,1,1)(1,1,1)	300	SARIMA(2,1,1)(1,1,2)
301	SARIMA(2,1,1)(2,0,0)	302	SARIMA(2,1,1)(2,0,1)
303	SARIMA(2,1,1)(2,0,2)	304	SARIMA(2,1,1)(2,1,0)
305	SARIMA(2,1,1)(2,1,1)	306	SARIMA(2,1,1)(2,1,2)
307	SARIMA(2,1,2)(0,0,0)	308	SARIMA(2,1,2)(0,0,1)
309	SARIMA(2,1,2)(0,0,2)	310	SARIMA(2,1,2)(0,1,0)
311	SARIMA(2,1,2)(0,1,1)	312	SARIMA(2,1,2)(0,1,2)
313	SARIMA(2,1,2)(1,0,0)	314	SARIMA(2,1,2)(1,0,1)
315	SARIMA(2,1,2)(1,0,2)	316	SARIMA(2,1,2)(1,1,0)
317	SARIMA(2,1,2)(1,1,1)	318	SARIMA(2,1,2)(1,1,2)
319	SARIMA(2,1,2)(2,0,0)	320	SARIMA(2,1,2)(2,0,1)
321	SARIMA(2,1,2)(2,0,2)	322	SARIMA(2,1,2)(2,1,0)
323	SARIMA(2,1,2)(2,1,1)	324	SARIMA(2,1,2)(2,1,2)

Tabla C.1: Espacio de Modelos

Bibliografía

En Egipto se llamaban las bibliotecas el tesoro de los remedios del alma. En efecto, curábase en ellas de la ignorancia, la más peligrosa de las enfermedades y el origen de todas las demás.

Jackes Benigne Bossuet (Clérigo,
predicador e intelectual francés)

- AKAIKE, H. Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics*, vol. 21, páginas 243–247, 1969. ISSN 0020-3157. 10.1007/BF02532251.
- AKAIKE, H. *Information theory and an extension of the maximum likelihood principle*, vol. 1, páginas 267–281. Akademiai Kiado, 1973.
- AKAIKE, H. A new look at the statistical model identification. *Automatic Control, IEEE Transactions on*, vol. 19(6), páginas 716–723, 1974.
- AKAIKE, H. Canonical correlation analysis of time series and the use of an information criterion. *System identification: advances and case studies*, vol. 126, página 27, 1976.
- AKAIKE, H. A new look at the bayes procedure. *Biometrika*, vol. 65(1), páginas 53–59, 1978a.
- AKAIKE, H. On the likelihood of a time series model. *The Statistician*, páginas 217–235, 1978b.
- AKAIKE, H. A bayesian extension of the minimum aic procedure of autoregressive model fitting. *Biometrika*, vol. 66(2), páginas 237–242, 1979.
- AKAIKE, H. Likelihood of a model and information criteria. *Journal of Econometrics*, vol. 16(1), páginas 3 – 14, 1981. ISSN 0304-4076.
- ALLENDE, M. C. Y. S. R., H. Artificial neural networks in the time series forecasting; a comparative analysis. *Kybemetik*, vol. 38, páginas 685–707, 2002.

- ALTIMIR, O. y PINERA, S. Análisis de descomposición de las desigualdades de los ingresos primarios en países de América Latina. *CEPAL y Banco Internacional de Reconstrucción y Fomento*, enero, 1977.
- ALTIMIR, O. y PIÑERA, S. Análisis de descomposición: Una generalización del método de Theil. *Latin American Journal of Economics-formerly Cuadernos de Economía*, vol. 16(48), páginas 207–236, 1979.
- BARTLETT, M. S. On the theoretical specification and sampling properties of autocorrelated time-series. *Supplement to the Journal of the Royal Statistical Society*, vol. 8(1), páginas pp. 27–41, 1946. ISSN 14666162.
- BHANSALI, R. J. . y DOWNHAM, D. Y. Some properties of the order of an autoregressive model selected by a generalization of akaike's epf criterion. *Biometrika*, vol. 64(3), páginas 547–551, 1977.
- BOX, G. G.. m. Jenkins, time series analysis: forecasting and control. *Rev. Ed. Holdem Day*, 1976.
- BOX, G. E. P. y COX, D. R. An analysis of transformations. *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 26(2), páginas pp. 211–252, 1964. ISSN 00359246.
- BOZDOGAN, H. Model selection and akaike's information criterion (aic): The general theory and its analytical extensions. *Psychometrika*, vol. 52, páginas 345–370, 1987. ISSN 0033-3123. 10.1007/BF02294361.
- CAVANAUGH, J. E. y NEATH, A. A. Generalizing the derivation of the schwarz information criterion. *Communications in Statistics - Theory and Methods*, vol. 28(1), páginas 49–66, 1999.
- CAVANAUGH, J. E. y SHUMWAY, R. H. A bootstrap variant of aic for state space model selection. *Statistica Sinica*, vol. 7(2), páginas 473–496, 1997.
- CHANDRA, P. y SINGH, Y. An activation function adapting training algorithm for sigmoidal feedforward networks. *Neurocomputing*, vol. 61, páginas 429–437, 2004.
- CHISWICK, C. U. *Application of the Theil Index to income inequality*. World Bank, 1976.
- CHOI, K. y THACKER, S. B. An evaluation of influenza mortality surveillance, (1962,1979).i time series forecasts of expected pneumonia and influenza deaths. *American Journal of Epidemiology*, vol. 113(3), páginas 215–226, 1981a.
- CHOI, K. y THACKER, S. B. An evaluation of influenza mortality surveillance, (1962,1979).ii percentage of pneumonia and influenza deaths as an

- indicator of influenza activity. *American Journal of Epidemiology*, vol. 113(3), páginas 227–235, 1981b.
- CHOW, G. C. A comparison of the information and posterior probability criteria for model selection. *Journal of Econometrics*, vol. 16(1), páginas 21 – 33, 1981. ISSN 0304-4076.
- COLLINS, S. Excess mortality from causes other than influenza and pneumonia during influenza epidemics. *Public Health Rep*, vol. 47, páginas 2159–2180, 1932.
- EICKHOFF, T., SHERMAN, I. y SERFLING, R. Observations on excess mortality associated with epidemic influenza. *Maryland state medical journal*, vol. 11, página 115, 1962.
- ELMAN, J. L. Finding structure in time. *Cognitive science*, vol. 14(2), páginas 179–211, 1990.
- FARR, W. Progress of epidemics. *Second report of the Registrar General of England and Wales*, páginas 16–20, 1840.
- FINDLEY, D. F. On the unbiasedness property of aic for exact or approximating linear stochastic time series models. *Journal of Time Series Analysis*, vol. 6(4), páginas 229–252, 1985. ISSN 1467-9892.
- FISHLOW, A. Brazilian size distribution of income. *The American Economic Review*, vol. 62(1/2), páginas 391–402, 1972.
- GALEANO, P. y PEÑA, D. Model selection criteria and quadratic discrimination in arma and setar time series models. *N.: UC3M Working Papers. Statistics and Econometrics 2004-06*, 2004.
- GARCÍA HIERNAUX, A. *Identificación de modelos para series temporales mediante métodos de subespacios*. Tesis Doctoral, Universidad Complutense de Madrid, Servicio de Publicaciones, 2006.
- GRAUNT, J. *Natural and political observations mentioned in a following index, and made upon the bills of mortality*. London. Reprinted Johns Hopkins University Press, Baltimore, 1939, 1662.
- H., A. On entropy maximization principle. *Application of Statistics*, 1977.
- HALLEY, E. *An Estimate of the Degrees of the Mortality of Mankind Drawn from Curious Tables of the Births and Funerals at the City of Breslau; with an Attempt of Ascertain the Prices of Annuities upon Lives*. Phil Trans Roy Soc London; XVII: 6-49, 1693.
- HANNAN, E. J. The estimation of the order of an arma process. *The Annals of Statistics*, vol. 8(5), páginas pp. 1071–1081, 1980. ISSN 00905364.

- HANNAN, E. J. y QUINN, B. G. The determination of the order of an autoregression. *Journal of the Royal Statistical Society. Series B (Methodological)*, vol. 41(2), páginas pp. 190–195, 1979. ISSN 00359246.
- HAUGHTON, D. M. A. On the choice of a model to fit data from an exponential family. *The Annals of Statistics*, vol. 16(1), páginas pp. 342–355, 1988. ISSN 00905364.
- HERFINDAHL, O. C. *Concentration in the steel industry*. Tesis Doctoral, Columbia university., 1950.
- HIRSCHMAN, A. O. *National power and the structure of foreign trade*, vol. 18. University of California Press Berkeley, 1945.
- HURVICH, C. M., SHUMWAY, R. y TSAI, C.-L. Improved estimators of kullback leibler information for autoregressive model selection in small samples. *Biometrika*, vol. 77(4), páginas 709–719, 1990.
- HURVICH, C. M. y TSAI, C.-L. Regression and time series model selection in small samples. *Biometrika*, vol. 76(2), páginas 297–307, 1989.
- HURVICH, C. M. y TSAI, C.-L. Bias of the corrected aic criterion for underfitted regression and time series models. *Biometrika*, vol. 78(3), páginas 499–509, 1991.
- KAJITANI, Y., HIPEL, K. W. y MCLEOD, A. I. Forecasting nonlinear time series with feed-forward neural networks: a case study of canadian lynx data. *Journal of Forecasting*, vol. 24(2), páginas 105–117, 2005.
- KONISHI, S. y KITAGAWA, G. Generalised information criteria in model selection. *Biometrika*, vol. 83(4), páginas 875–890, 1996.
- LÜTKEPOHL, H. y NETLIBRARY, I. Introduction to multiple time series analysis. 1991.
- LÜTKEPOHL, H. y NETLIBRARY, I. New introduction to multiple time series analysis. 2005.
- MALINVAUD, E. *Statistical methods of econometrics*. North-Holland, Amsterdam, 1970.
- MCCULLOCH, W. S. y PITTS, W. A logical calculus of the ideas immanent in nervous activity. *The Bulletin of Mathematical Biophysics*, vol. 5(4), páginas 115–133, 1943.
- MCQUARRIE, A., SHUMWAY, R. y TSAI, C.-L. The model selection criterion aicu. *Statistics; Probability Letters*, vol. 34(3), páginas 285 – 292, 1997. ISSN 0167-7152.

- CARIDAD Y OCERIN, J. y CARIDAD Y LÓPEZ DEL RÍO, D. *Análisis de datos en Ciencias Sociales*. Universidad de Córdoba, 2003.
- CARIDAD Y OCERÁN, J. S. y E. R., J.M. Hybrida-arma-neural network model for prediction of water consumption aggregate. Informe técnico, International Conference on Social Sciences and Society, 2011.
- OGATA, Y. Maximum likelihood estimates of incorrect markov models for time series and the derivation of aic. *Journal of Applied Probability*, vol. 17(1), páginas pp. 59–72, 1980. ISSN 00219002.
- OZAKI, T. On the order determination of arima models. *Applied Statistics*, páginas 290–301, 1977.
- PATIÑO, V. M. O. Redes neuronales artificiales en las ciencias económicas. Informe técnico, UN-RCE-CID, 2012.
- RAO, J. N. K. y TINTNER, G. On the variate difference method. *Australian Journal of Statistics*, vol. 5(3), páginas 106–116, 1963. ISSN 1467-842X.
- RESCHENHOFER, E. Selecting selection methods. *InterStat*, 2006.
- RICE, S. A. Quantitative methods in politics. 1928.
- RISSANEN, J. Modeling by shortest data description. *Automatica*, vol. 14(5), páginas 465 – 471, 1978. ISSN 0005-1098.
- PEÑA SÁNCHEZ DE RIVERA, D. y ARNÁIZ TOVAR, G. Criterios de selección de modelos arima. *Trabajos de estadística e investigación operativa*, vol. 32(1), páginas 70–93, 1981.
- RUMELHART, D., HINTON, G. E. y WILLIAMS, R. J. Learning representations by back-propagating errors. *Nature*, vol. 323(9), páginas 533–535, 1986.
- RUMELHART, D. E., HINTON, G. E. y WILLIAMS, R. J. Learning representations by back-propagating errors. *Cognitive modeling*, vol. 1, página 213, 2002.
- SANTANA, J. C. Predicción de series temporales con redes neuronales: una aplicación a la inflación colombiana. *Revista Colombiana de Estadística*, vol. 29(1), páginas 77–92, 2006.
- SCHWARZ, G. Estimating the dimension of a model. *The Annals of Statistics*, vol. 6(2), páginas pp. 461–464, 1978. ISSN 00905364.
- SERFLING, R. Methods for current statistical analysis of excess pneumonia-influenza deaths. *Public health reports*, vol. 78(6), página 494, 1963.

- SHANNON, C. E. A mathematical theory of communication. *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5(1), páginas 3–55, 2001.
- SHEPHERD, A. J. *Second-order methods for neural networks*. Springer-Verlag New York, Inc., 1997.
- SHI, P. y TSAI, C.-L. A note on the unification of the akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 60(3), páginas 551–558, 1998. ISSN 1467-9868.
- SHI, P. y TSAI, C.-L. Regression model selection: a residual likelihood approach. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 64(2), páginas 237–252, 2002. ISSN 1467-9868.
- SHIBATA, R. Selection of the order of an autoregressive model by akaike's information criterion. *Biometrika*, vol. 63(1), páginas 117–126, 1976.
- SHIBATA, R. Asymptotically efficient selection of the order of the model for estimating parameters of a linear process. *The Annals of Statistics*, vol. 8(1), páginas pp. 147–164, 1980. ISSN 00905364.
- SHIBATA, R. From data to model. capítulo Statistical aspects of model selection, páginas 215–240. Springer-Verlag New York, Inc., New York, NY, USA, 1989. ISBN 0-387-51571-2.
- SODERSTROM, T. On model structure testing in system identification. *International Journal of Control*, vol. 26(1), páginas 1–18, 1977.
- SUGIURA, N. Further analysts of the data by akaike's information criterion and the finite corrections. *Communications in Statistics - Theory and Methods*, vol. 7(1), páginas 13–26, 1978.
- THEIL, H. Economic forecasts and policy. 1958.
- THEIL, H. *Applied economic forecasting*, vol. 4. North-Holland Pub. Co., 1966.
- THEIL, H. Statistical decomposition analysis. 1972.
- THEIL, H. y SCHOLLES, M. Forecast evaluation based on a multiplicative decomposition of mean square errors. *Econometrica: journal of the Econometric Society*, páginas 70–88, 1967.
- TINTNER, G. ET AL. The variate difference method. *The variate difference method.*, 1940.
- TSAY, R. S. Order selection in nonstationary autoregressive models. *The Annals of Statistics*, vol. 12(4), páginas pp. 1425–1433, 1984. ISSN 00905364.

- TSENG, F.-M., YU, H.-C. y TZENG, G.-H. Combining neural network model with seasonal time series arima model. *Technological Forecasting and Social Change*, vol. 69(1), páginas 71–87, 2002.
- URIEL JIMÉNEZ, E. *Análisis de series temporales modelos Arima*. Thomson-Paraninfo, 1985.
- VAN GINNEKEN, W. Análisis de descomposición del índice de theil aplicado a la distribución del ingreso familiar en México. *Demografía y economía*, páginas 93–112, 1975.
- WOLD, H. *A study in the analysis of stationary time series*. Tesis Doctoral, Stockholm, 1938.
- YAGLOM, A. The correlation theory of processes whose nth difference constitutes a stationary process. *Matem Sb*, vol. 141, 1955.
- YULE, G. U. On the time-correlation problem, with especial reference to the variate-difference correlation method. *Journal of the Royal Statistical Society*, vol. 84(4), páginas pp. 497–537, 1921. ISSN 09528385.
- YULE, G. U. Why do we sometimes get nonsense-correlations between time-series?—a study in sampling and the nature of time-series. *Journal of the Royal Statistical Society*, vol. 89(1), páginas pp. 1–63, 1926. ISSN 09528385.
- YULE, G. U. On a method of investigating periodicities in disturbed series, with special reference to wolfer's sunspot numbers. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character*, vol. 226(636-646), páginas 267–298, 1927.
- ZHANG, G., EDDY PATUWO, B. y Y HU, M. Forecasting with artificial neural networks:: The state of the art. *International journal of forecasting*, vol. 14(1), páginas 35–62, 1998.
- ZHANG, G. P. Time series forecasting using a hybrid arima and neural network model. *Neurocomputing*, vol. 50, páginas 159–175, 2003.

Debe evitarse hablar a los jóvenes del éxito como si se tratase del principal objetivo en la vida. La razón más importante para trabajar en la escuela y en la vida es el placer de trabajar, el placer de su resultado y el conocimiento del valor del resultado para la comunidad.

Albert Einstein. (Físico alemán)

*Un ciclo que termina,
una página que llega a su fin;
un peldaño más que subir.
La oportunidad comienza de nuevo...*

