



UNIVERSIDAD
DE PIURA

FACULTAD DE INGENIERÍA

**Modelos de predicción para el caudal del río Chira en la
estación Ardilla**

Tesis para optar el Título de
Ingeniero Industrial y de Sistemas

Julio German Seminario Gastelo

Asesor:
Dr. Ing. Ismael Sánchez Rodríguez-Morcillo

Piura, mayo de 2021

A mi familia, por su constante soporte y aliento.



A Dios, por permitirme prosperar en la vida.
A mis padres y hermano por motivarme a lo largo del proceso.
Mi más profundo agradecimiento al Dr. Ing. Ismael Sánchez,
por su orientación y consejo sin el cual no hubiese sido
posible el desarrollo de la presente investigación.



Resumen

La región Piura hace uso de los recursos hídricos provistos por la cuenca Catamayo – Chira. Las características climatológicas propias de la región, la condición irregular de los cursos de agua superficial y la presencia del fenómeno “El Niño” convierten la gestión de los recursos en una tarea crucial para la seguridad hídrica de Piura. La presente investigación busca encontrar un modelo univariante que permita predecir el caudal del río Chira en la Estación Ardilla, a puertas del reservorio de Poechos, principal infraestructura para la gestión de los recursos hídricos.

Se realiza un análisis estocástico de las series de tiempo de caudales promedio con frecuencia diaria y semanal, utilizando datos recolectados en la Estación Ardilla, entre 1951 a 2017. Los modelos se desarrollan utilizando métodos de suavizamiento exponencial y métodos ARIMA. Para evaluar la capacidad predictiva de los modelos, se utilizan métricas como el error cuadrático medio, el error absoluto medio y el coeficiente de determinación. La comparativa incluye la evaluación de los indicadores en diez horizontes de predicción, una segunda evaluación diferenciando las predicciones por estaciones y la última evalúa el rendimiento de las predicciones en función del valor del caudal.

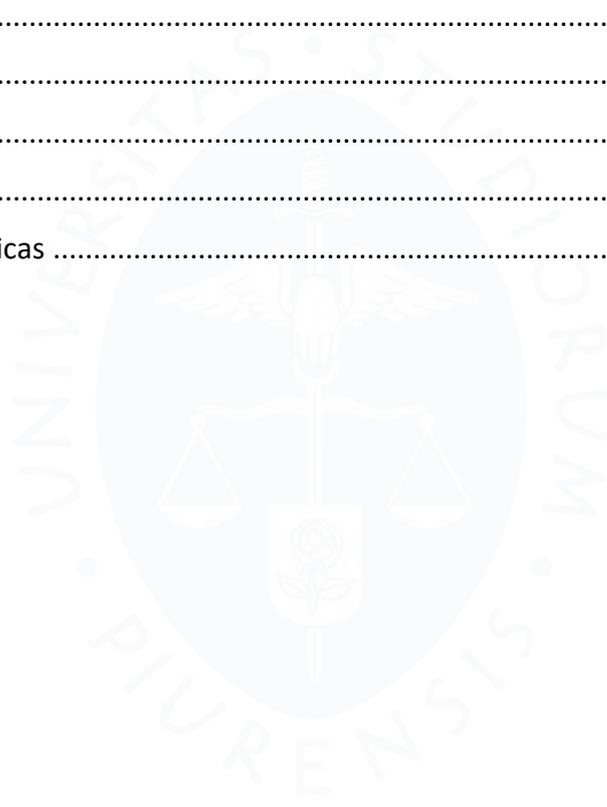
De la comparación entre los modelos, se obtiene el mejor rendimiento para la serie temporal de frecuencia diaria con el Modelo ARIMA (1,1,1). En el caso de la serie temporal con frecuencia semanal, se selecciona el Modelo ARIMA (1,1,1)(0,1,1)₅₂.

La metodología ARIMA, constituye una herramienta útil para la predicción del caudal del río Chira, sin considerar variables hidrológicas distintas al caudal. Los modelos resultantes de la presente investigación permiten la anticipación al caudal del río de una manera rápida y sencilla al prescindir de variables independientes.

Tabla de contenido

Introducción	17
Capítulo 1 Descripción del área de estudio.....	19
1.1 Cuenca binacional Catamayo – Chira	19
1.1.1 Ubicación y extensión	19
1.1.2 Hidrografía del río Chira.....	19
1.1.3 Climatología	24
1.2 Proyecto Especial Chira – Piura	25
1.2.1 Origen.....	25
1.2.2 Ejecución	26
1.2.3 Infraestructura	29
1.3 Reservoirio de Poechos	31
1.3.1 Ubicación.....	32
1.3.2 Infraestructura	32
1.3.3 Operación.....	33
1.3.4 Problema de colmatación	34
Capítulo 2 Fundamentación teórica de series temporales	37
2.1 Métodos de suavizamiento exponencial.....	38
2.1.1 Suavizamiento exponencial simple	38
2.1.2 Suavizamiento exponencial doble (Holt).....	38
2.1.3 Suavizamiento exponencial triple (Holt-Winters)	39
2.2 Métodos ARIMA.....	41
2.2.1 Procesos autorregresivos (AR)	41
2.2.2 Procesos de media móvil (MA).....	42
2.2.3 Funciones de autocorrelación simple y parcial	43
2.2.4 Procesos ARIMA	45

Capítulo 3 Propuesta de modelos de predicción	51
3.1 Datos.....	51
3.2 Modelización diaria	57
3.2.1 <i>Métodos de suavizamiento exponencial</i>	57
3.2.2 <i>Métodos ARIMA</i>	66
3.3 Modelización semanal	73
3.3.1 <i>Métodos de suavizamiento exponencial</i>	73
3.3.2 <i>Métodos ARIMA</i>	82
Capítulo 4 Comparación de modelos	93
4.1 Diario.....	94
4.2 Semanal	110
Conclusiones.....	127
Recomendaciones	129
Referencias bibliográficas	131



Lista de tablas

Tabla 1. División de la cuenca del río Chira.....	23
Tabla 2. Usuarios agrícolas y áreas totales de la cuenca Catamayo - Chira.....	24
Tabla 3. Tipos de clima en la cuenca Catamayo - Chira	25
Tabla 4. Obras ejecutadas en 1° Etapa Proyecto Especial Chira - Piura.....	27
Tabla 5. Obras ejecutadas en 2° Etapa Proyecto Especial Chira - Piura.....	28
Tabla 6. Obras ejecutadas en 3° Etapa Proyecto Especial Chira - Piura.....	29
Tabla 7. Estaciones meteorológicas en la cuenca del río Chira.	31
Tabla 8. Procedimiento para generación de predicciones diarias.	57
Tabla 9. Parámetros promedio estimados - Modelo Holt diario.	58
Tabla 10. Parámetros promedio estimados - Modelo Holt Damped diario.....	59
Tabla 11. Parámetros promedio estimados - Modelo Holt-Winters diario.	61
Tabla 12. Parámetros promedio estimados - Modelo Holt-Winters Damped diario.	62
Tabla 13. Parámetros promedio estimados - Modelo Holt-Winters Multiplicativo diario.....	63
Tabla 14. Parámetros promedio estimados - Modelo Holt-Winters Multiplicativo Damped diario.....	65
Tabla 15. Procedimiento para generación de predicciones semanales.....	73
Tabla 16. Parámetros promedio estimados - Modelo Holt semanal.	74
Tabla 17. Parámetros promedio estimados - Modelo Holt Damped semanal.	75
Tabla 18. Parámetros promedio estimados - Modelo Holt-Winters semanal.	77
Tabla 19. Parámetros promedio estimados - Modelo Holt-Winters Damped semanal.	78
Tabla 20. Parámetros promedio estimados - Modelo Holt-Winters Multiplicativo semanal..	79
Tabla 21. Parámetros promedio estimados - Modelo Holt-Winters Multiplicativo Damped semanal.	81
Tabla 22. Resumen ECMP vs h - Comparativa frecuencia diaria.....	96
Tabla 23. Resumen MAE vs h - Comparativa frecuencia diaria.	97
Tabla 24. Resumen R ² vs h - Comparativa frecuencia diaria.....	97
Tabla 25. Resumen ECMP vs h - Comparativa frecuencia semanal.	112
Tabla 26. Resumen MAE vs h - Comparativa frecuencia semanal.	113
Tabla 27. Resumen R ² vs h - Comparativa frecuencia semanal.	113

Lista de figuras

Figura 1. Mapa de cuenca hidrográfica binacional Catamayo – Chira.....	21
Figura 2. Mapa de subcuencas de la cuenca hidrográfica binacional Catamayo – Chira.	22
Figura 3. Etapas de ejecución Proyecto Especial Chira – Piura.....	26
Figura 4. Primera etapa Proyecto Especial Chira - Piura.....	27
Figura 5. Tercera etapa Proyecto Especial Chira - Piura.	28
Figura 6. Volúmenes aportados por los principales ríos al esquema hidráulico de la cuenca Catamayo - Chira.	34
Figura 7. Acumulación de sedimentos en reservorio de Poechos.	34
Figura 8. Comparativa del volumen del reservorio, según batimetría de 2012.	35
Figura 9. Retención de sedimentos en el reservorio de Poechos.	35
Figura 10. Funciones de autocorrelación simple y parcial de modelos autorregresivos.....	44
Figura 11. Funciones de autocorrelación simple y parcial de procesos de media móvil.....	45
Figura 12. Funciones de autocorrelación simple y parcial de procesos ARMA (1,1).	48
Figura 13. Caudal promedio diario - Estación Ardilla.....	52
Figura 14. Histograma caudal promedio diario -Estación Ardilla.....	52
Figura 15. Caudal promedio semanal - Estación Ardilla.....	53
Figura 16. Histograma caudal promedio semanal - Estación Ardilla.	53
Figura 17. Diagrama de caja del caudal promedio diario por trimestre - Estación Ardilla.....	54
Figura 18. Caudal promedio anual - Estación Ardilla.	54
Figura 19. Varianza anual del caudal - Estación Ardilla.....	55
Figura 20. Transformación Logarítmica del caudal promedio diario.	56
Figura 21. Transformación Logarítmica del caudal promedio semanal.....	56
Figura 22. ECMP vs h - Modelo Holt diario.	58
Figura 23. ECMP vs h - Modelo Holt Damped diario.....	60
Figura 24. ECMP vs h - Modelo Holt-Winters diario.	61
Figura 25. ECMP vs h - Modelo Holt-Winters Damped diario.....	63
Figura 26. ECMP vs h - Modelo Holt-Winters Multiplicativo diario.	64
Figura 27. ECMP vs h - Modelo Holt-Winters Multiplicativo Damped diario.	66
Figura 28. Diferencia regular de logaritmos del caudal promedio diario.	67
Figura 29. Figura 29. FAS para la serie DLC.	68
Figura 30. FAP para la serie DLC.....	68

Figura 31. FAS para la serie DLC (Aumento).....	69
Figura 32. FAP para la serie DLC (Aumento).	69
Figura 33. Parámetros promedio estimados – Modelo ARIMA (1,1,1) diario.	70
Figura 34. FAS para errores del modelo ARIMA (1,1,1) diario.	70
Figura 35. FAP para errores del modelo ARIMA (1,1,1) diario.	71
Figura 36. FAS para errores del modelo ARIMA (1,1,1) diario (Aumento).	71
Figura 37. FAP para errores del modelo ARIMA (1,1,1) diario (Aumento).	72
Figura 38. ECMP vs h - Modelo ARIMA (1,1,1) diario.....	72
Figura 39. ECMP vs h - Modelo Holt semanal.	74
Figura 40. ECMP vs h - Modelo Holt Damped semanal.	76
Figura 41. ECMP vs h – Modelo Holt-Winters semanal.	77
Figura 42. ECMP vs h - Modelo Holt-Winters Damped semanal.	79
Figura 43. ECMP vs h - Modelo Holt-Winters Multiplicativo semanal.....	80
Figura 44. ECMP vs h - Modelo Holt-Winters Multiplicativo Damped semanal.	82
Figura 45. Diferencia regular de logaritmos de caudal promedio semanal.....	83
Figura 46. FAS para la serie DLCS.	83
Figura 47. FAP para la serie DLCS.	84
Figura 48. Diferencia estacional y diferencia regular de logaritmos de caudal promedio semanal.	84
Figura 49. FAS para la serie D52DLCS.....	85
Figura 50. FAP para la serie D52DLCS.....	85
Figura 51. Parámetros promedio estimados - Modelo ARIMA (1,1,1) Semanal.....	86
Figura 52. Parámetros promedio estimados - Modelo ARIMA (2,1,0) Semanal.....	87
Figura 53. Parámetros promedio estimados - Modelo ARIMA (1,1,1)(0,1,1) ₅₂ Semanal.	87
Figura 54. FAS para residuos del modelo ARIMA (1,1,1) semanal.....	88
Figura 55. FAP para residuos del modelo ARIMA (1,1,1) semanal.....	88
Figura 56. FAS para residuos del modelo ARIMA (2,1,0) semanal.....	89
Figura 57. FAP para residuos del modelo ARIMA (2,1,0) semanal.....	89
Figura 58. FAS para residuos del modelo ARIMA (1,1,1)(0,1,1) ₅₂ semanal.....	90
Figura 59. FAP para residuos del modelo ARIMA (1,1,1)(0,1,1) ₅₂ semanal.	90
Figura 60. ECMP vs h - Modelo ARIMA (1,1,1) semanal.	91
Figura 61. ECMP vs h - Modelo ARIMA (2,1,0) semanal.	92
Figura 62. ECMP vs h - Modelo ARIMA (1,1,1)(0,1,1) ₅₂ semanal.	92
Figura 63. ECMP vs h - Comparativa frecuencia diaria.	95
Figura 64. MAE vs h - Comparativa frecuencia diaria.	95
Figura 65. R ² vs h - Comparativa frecuencia diaria.	96
Figura 66. ECMP vs h (1° Trimestre) - Comparativa frecuencia diaria.	99
Figura 67. MAE vs h (1° Trimestre) - Comparativa frecuencia diaria.....	100
Figura 68. R ² vs h (1° Trimestre) - Comparativa frecuencia diaria.	100
Figura 69. ECMP vs h (2° Trimestre) - Comparativa frecuencia diaria.....	101

Figura 70. MAE vs h (2° Trimestre) - Comparativa frecuencia diaria.....	101
Figura 71. R ² vs h (2° Trimestre) - Comparativa frecuencia diaria.....	102
Figura 72. ECMP vs h (3° Trimestre) - Comparativa frecuencia diaria.....	102
Figura 73. MAE vs h (3° Trimestre) - Comparativa frecuencia diaria.....	103
Figura 74. R ² vs h (3° Trimestre) - Comparativa frecuencia diaria.....	103
Figura 75. ECMP vs h (4° Trimestre) - Comparativa frecuencia diaria.....	104
Figura 76. MAE vs h (4° Trimestre) - Comparativa frecuencia diaria.....	104
Figura 77. R ² vs h (4° Trimestre) - Comparativa frecuencia diaria.....	105
Figura 78. Varianza Caudal (4° Trimestre) - Estación Ardilla.....	105
Figura 79. ECMP vs Caudal Medido - Primer horizonte de predicción. Frecuencia diaria.....	107
Figura 80. MAE vs Caudal Medido - Primer horizonte de predicción. Frecuencia diaria.	107
Figura 81. ECMP vs Caudal Medido - Quinto horizonte de predicción. Frecuencia diaria.	108
Figura 82. MAE vs Caudal Medido - Quinto horizonte de predicción. Frecuencia diaria.	108
Figura 83. ECMP vs Caudal Medido - Décimo horizonte de predicción. Frecuencia diaria. ..	109
Figura 84. MAE vs Caudal Medido - Décimo horizonte de predicción. Frecuencia diaria.	109
Figura 85. ECMP vs h - Comparativa frecuencia semanal.....	111
Figura 86. MAE vs h - Comparativa frecuencia semanal.....	111
Figura 87. R ² vs h- Comparativa frecuencia semanal.....	112
Figura 88. ECMP vs h (1° Trimestre) - Comparativa frecuencia semanal.....	114
Figura 89. MAE vs h (1° Trimestre) - Comparativa frecuencia semanal.	115
Figura 90. R ² vs h (1° Trimestre) - Comparativa frecuencia semanal.....	115
Figura 91. ECMP vs h (2° Trimestre) - Comparativa frecuencia semanal.....	116
Figura 92. MAE vs h (2° Trimestre) - Comparativa frecuencia semanal.	116
Figura 93. R ² vs h (2° Trimestre) - Comparativa frecuencia semanal.....	117
Figura 94. ECMP vs h (3° Trimestre) - Comparativa frecuencia semanal.....	118
Figura 95. MAE vs h (3° Trimestre) - Comparativa frecuencia semanal.	119
Figura 96. R ² vs h (3° Trimestre) - Comparativa frecuencia semanal.....	119
Figura 97. ECMP vs h (4° Trimestre) - Comparativa frecuencia semanal.....	120
Figura 98. MAE vs h (4° Trimestre) - Comparativa frecuencia semanal.	120
Figura 99. R ² vs h (4° Trimestre) - Comparativa frecuencia semanal.....	121
Figura 100. ECMP vs Caudal Medido - Primer horizonte de predicción. Frecuencia semanal.	122
Figura 101. MAE vs Caudal Medido - Primer horizonte de predicción. Frecuencia semanal.	123
Figura 102. ECMP vs Caudal Medido - Quinto horizonte de predicción. Frecuencia semanal.	123
Figura 103. MAE vs Caudal Medido - Quinto horizonte de predicción. Frecuencia semanal.	124
Figura 104. ECMP vs Caudal Medido - Décimo horizonte de predicción. Frecuencia semanal.	124

Figura 105. MAE vs Caudal Medido - Décimo horizonte de predicción. Frecuencia semanal.

..... 125



Introducción

Los recursos hídricos son recursos naturales renovables imprescindibles para las actividades primordiales del hombre. Con el crecimiento de la sociedad, se ha tomado especial conciencia en la gestión de los mismos, ya que todo desarrollo depende básicamente de ellos. El aumento de las actividades productivas, el crecimiento de las ciudades, el mejoramiento de la calidad de vida, entre otros, aumentan la demanda de agua de las poblaciones.

La región Piura cuenta con gran cantidad de cuencas hidrográficas de la vertiente del Pacífico y los principales cursos de agua superficial existentes son el río Chira y Piura. Sin embargo, los ríos de la región son irregulares. En este contexto se desarrolló el Proyecto Especial Chira – Piura, creando un sistema hidráulico capaz de asegurar la continuidad de los recursos hídricos.

El Proyecto Especial Chira – Piura se compone de una serie de obras hidráulicas, las cuales se operan para lograr el aprovechamiento e incremento de la producción y la productividad agrícola. Además, se constituye como un sistema de protección, ya que permite controlar las avenidas de los ríos y reserva grandes volúmenes de agua, garantizando la disponibilidad presente y futura del recurso hídrico.

Sin embargo, este proyecto evidencia dificultades para cumplir con sus objetivos debido a la creciente colmatación del reservorio de Poechos, embalse y componente más importante de este conjunto de obras hidráulicas, y la recurrencia del fenómeno El Niño.

Ante la necesidad de conocer con cierta exactitud y facilidad la disponibilidad presente y futura del recurso hídrico, los modelos de predicción constituyen una herramienta importante y permiten una administración racional del agua para sus diferentes fines.

La presente investigación busca determinar el modelo de predicción univariante adecuado dentro de un abanico amplio de opciones, que facilite la gestión de los recursos hídricos.

El estudio busca estimar el caudal a la entrada del embalse de Poechos. Se realiza un análisis estocástico de las series de tiempo de caudales promedio con frecuencia diaria y semanal en la estación Ardilla para el periodo 1951 a 2017. Se compara modelos ARIMA y de suavización exponencial.

Capítulo 1

Descripción del área de estudio

1.1 Cuenca binacional Catamayo – Chira

1.1.1 Ubicación y extensión

Se ubica geográficamente entre las coordenadas 03°30' a 05°08' latitud sur y 79°10' a 81°11' longitud oeste, localizándose en el área fronteriza del Perú y Ecuador. El rango altitudinal va desde la cota 4000 msnm hasta el nivel del mar en la desembocadura del río en el Océano Pacífico.

La cuenca transfronteriza del río Catamayo-Chira ocupa 17199.18 km², de los cuales 41.93% se encuentra en territorio ecuatoriano y 58.07% en territorio peruano. La cuenca se ubica entre la provincia de Loja en Ecuador, abarca los cantones Célica, Pindal, Macará, Sozoranga, Calvas, Espíndola, Gonzanamá, Quilanga, y atraviesa parte de los cantones de Loja, Catamayo, Paltas, Olmedo, Puyango y Zapotillo, y la región Piura en Perú, donde recorre parte de las provincias de Ayabaca, Huancabamba, Morropón, Paita, Piura, Sullana y Talara. (CONSORCIO ATA-UNL-UNP, 2002)

La longitud total de la cuenca hasta su desembocadura en el Océano Pacífico es de 315 km. En el territorio peruano se ubican 119km, donde finalmente desemboca. (Figura 1)

Sus límites hidrográficos son:

- Por el norte, limita con la cuenca del río Puyango-Tumbes.
- Por el sur, limita con la cuenca del río Piura.
- Por el este, limita con las cuencas de Santiago y Chinchipe en Ecuador.
- Por el oeste, limita con las cuencas Pariñas y Fernández y con el Océano Pacífico. (Chinchay Alza, Santana Vera, Panta Samillán, Portocarrero Lau, & Vásquez Arca, 2007)

1.1.2 Hidrografía del río Chira

El Consorcio ATA-UNL-UNP, en su caracterización hídrica de la cuenca binacional Catamayo Chira menciona que la cuenca está integrada por seis subcuencas que pertenecen a cada uno de los ríos que la conforman.

El río Catamayo nace en la Cordillera Occidental de los Andes, pertenece a la vertiente hidrográfica del Pacífico y recibe diferentes denominaciones como: Palmira, Piscobamba, Solanda, Chinguilamanca y El Arenal. El río nace de la unión de dos grandes afluentes: De norte a sur fluye el río llamado Guayabal, conformado por los ríos Gualal y el Ari por un lado y por el otro el río Trapichillo que nacen en la parte norte de la cuenca en la cordillera de Fierrouroco, con altitud máxima de 3754 m.s.n.m., y que recorre 45 km hasta unirse con El Arenal, adoptando finalmente el nombre de río Catamayo.

El río Catamayo recibe pequeños ríos hasta su confluencia con el río Macará, denominado aguas arriba río Calvas, formado por la unión de los ríos Chiriyacu y Espíndola. Por la margen derecha recibe la aportación del río Playas y quebradas como Quiara, Artones, entre otras. En la margen izquierda, Las Quebradas, El Salado, Chinchanga, el río Tagula y la quebrada Matadero. (Figura 2)

La subcuenca del río Catamayo representa el 24.33% del área total de la Cuenca del río Catamayo-Chira (4184 km²) y se ubica íntegramente en Ecuador. (CONSORCIO ATA-UNL-UNP, 2002)

Desde el encuentro del río Catamayo con el río Macará, a una altura de 284 m.s.n.m., el curso principal de la cuenca pasa a denominarse Chira, recorriendo 50 km y sirve del límite entre los países de Perú y Ecuador hasta unirse con el río Alamor, continuando sobre el territorio peruano en dirección suroeste hasta la ciudad de Sullana. A partir de este punto cambia a una dirección este-oeste hasta desembocar en el Océano Pacífico.

Dentro del territorio peruano, los principales afluentes del río Chira son:

- Por su margen derecha, la quebrada Honda, Peroles, Samán, La Tina, Poechos y Cóndor, y el río Pilares.
- Por su margen izquierda, ríos Quiroz y Chipillico y el río Macará.

Recibe también los aportes de otras quebradas que se activan en épocas de lluvia hasta su desembocadura en el océano. (Tabla 1)

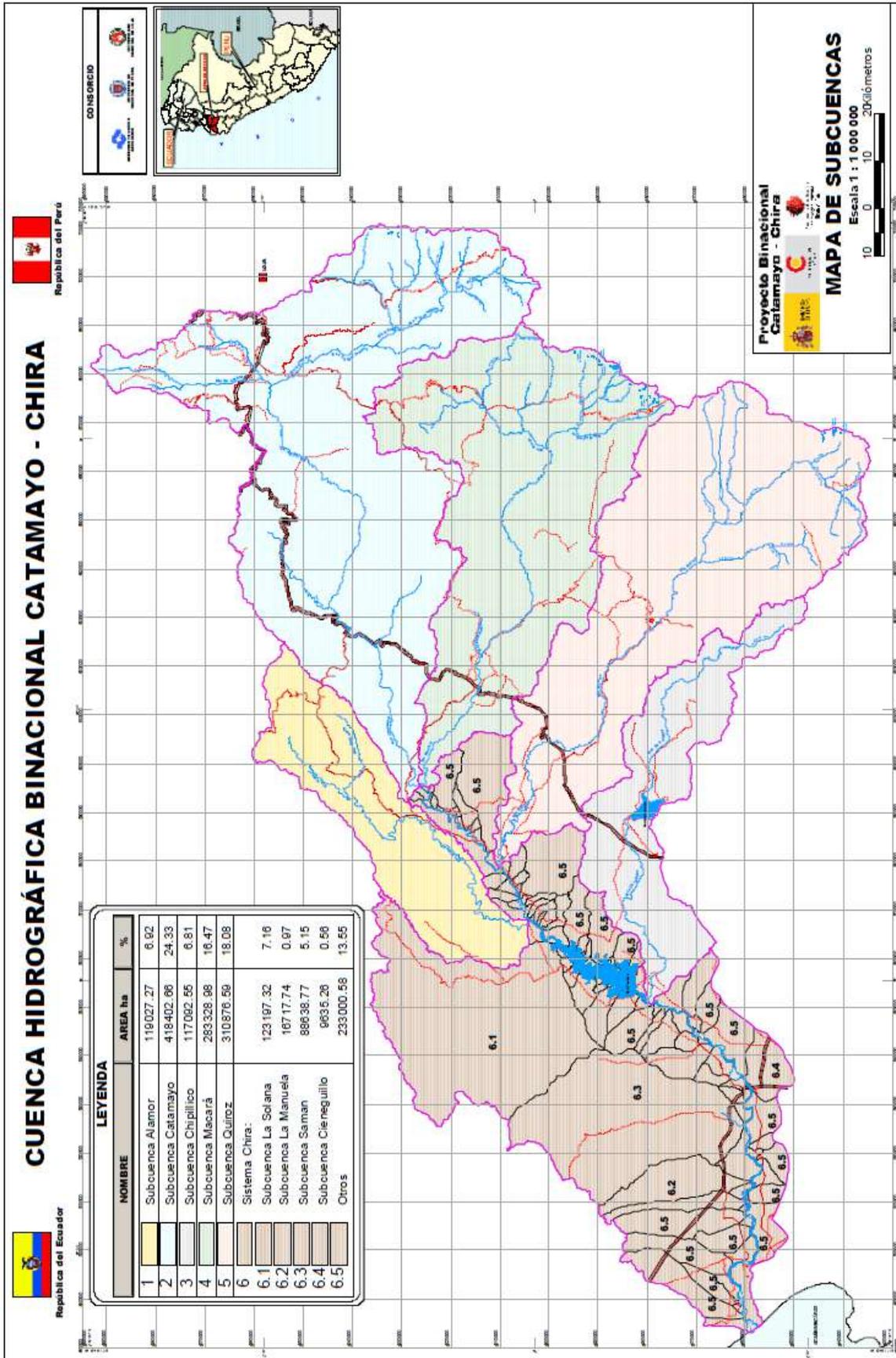


Figura 2. Mapa de subcuencas de la cuenca hidrográfica binacional Catamayo – Chira.

Fuente: CONSORCIO ATA-UNL-UNP (2002)

Tabla 1. División de la cuenca del río Chira

CUENCA	SUBCUENCAS	MICROCUENCAS
Río Chira	Quebrada Samán	Quebrada Sajinos
		Quebrada Caña Dulce
	Río Chipillico	Quebrada Yangas
		Quebrada Totoral
		Quebrada de Pichones
	Quebrada La Solana	Quebrada de Tamarindo
		Quebrada La Encantada
		Quebrada Jaguar
		Quebrada Encuentros
	Río Quiroz	Quebrada Suyo
		Quebrada Guir Guir
		Quebrada Tondopa
		Quebrada Huanta
		Quebrada Olleros
		Quebrada Mangas
		Quebrada Sicacate
		Quebrada Tulman
		Quebrada Santa Rosa
	Río Macará-Calvas	Quebrada Arogoto
Quebrada Chocán		

Fuente: Chinchay Alza, Santana Vera, Panta Samillán, Portocarrero Lau, & Vásquez Arca (2007)

La subcuenca del río Chipillico y parcialmente las aguas de Quiroz que van al reservorio San Lorenzo, atienden a las zonas agrícolas ubicadas en las áreas de la cuenca del Chira y Piura, a través del canal del canal Yuscay y los canales que nacen desde la estructura llamada "Partidor".

El sistema del Chira abarca una superficie de 471189.79 hectáreas, implicando el 27.39% del total de la cuenca Catamayo – Chira.

El río Chira es el segundo con mayor caudal de la costa peruano, solo después del río Santa. Provee de agua a la represa de Poechos, la más grande del Perú, que regula el agua de riego para más de 108 000 ha ubicadas en los valles de Chira y Piura. Por esta razón, el río Chira tiene una importancia estratégica para el desarrollo de la agricultura en la zona.

Los usuarios agrícolas de la cuenca Catamayo – Chira son 35222, según el estudio "Caracterización hídrica y adecuación entre la oferta y demanda en el ámbito de la cuenca binacional Catamayo – Chira", elaborado por el Consorcio ATA-UNL-UNP en 2002. El 30.8% pertenecen al Ecuador y 69.2% al Perú. Dentro de la cuenca existen 51518.22 Ha agrícolas en el territorio peruano. (Tabla 2)

La cuenca del río Chira se relaciona con la zona media y baja de la cuenca del río Piura, a la cual aporta importante caudal.

El canal de derivación del río Chira al río Piura, denominado canal Daniel Escobar, ha permitido incorporar la zona agrícola de Cieneguillo, que por ubicación se encuentra compartiendo las dos cuencas.

Tabla 2. Usuarios agrícolas y áreas totales de la cuenca Catamayo - Chira

PAÍS	CUENCA / SUBCUENCA	USUARIOS	ÁREA TOTAL
Ecuador	Macará	3371	6722.48
	Catamayo	6955	12296.44
	Alamor	507	1501.8
	Subtotal	10833	20520.72
Perú	Chira	16144	34761.15
	Chipillico	2058	9743.77
	Quiroz	2212	2441.43
	Macará (Perú)	3975	4571.87
	Subtotal	24389	51518.22
	Total	35222	72038.94

Fuente: CONSORCIO ATA-U NL-UNP (2002)

1.1.3 Climatología

Se definen dos periodos con diferencias marcadas, uno cálido y húmedo, y otro seco; lo que determina el uso de los suelos. Una gran parte se utiliza en cultivos temporales, el uso agrícola es más intensivo en el área de los valles por su proximidad a las corrientes fluviales.

La temperatura de la cuenca varía según la ubicación geográfica. En la parte alta, sobre los 3200 m.s.n.m. la temperatura oscila alrededor de 7°C. En la cuenca media la temperatura promedio es de 20°C y en la parte más baja varía desde los 24°C.

En la cuenca Catamayo – Chira, las precipitaciones no siguen un ritmo constante, sino que varían en función del espacio y el tiempo. En lo alto de la cuenca, el periodo lluvioso se da entre los meses de octubre a mayo, con una media anual que supera los 1000 mm. En la zona media, las lluvias se desarrollan de diciembre a mayo y la media anual oscila entre 500 y 1000 mm. Por el contrario, en la cuenca baja, los periodos lluviosos son cortos y escasos, principalmente en los meses de enero a abril. Esto a excepción de los años en los que se desarrolla el fenómeno del Niño. (Tabla 3)

La evaporación es 3mm por día en cuenca alta, llegando a 6mm por día en la cuenca baja. (CONSORCIO ATA-U NL-UNP, 2002)

Tabla 3. Tipos de clima en la cuenca Catamayo - Chira

Clima	Superficie (%)	Localidad	Altitud
Cálido	44.6	Esperanza, Mallares, Lancones, Zapotillo, Pindal, Las Lomas, Suyo, Paimas, Saudal de Culucán, Macará, Sabiango, Catamayo, Las Canoas, El Huayco.	< 1000 m.s.n.m.
Semicálido	23.6	Nazaino, Sapillica, Montero, Sícchez, Ciruelo, Amaluza, Vizancio, Zozoranga, Suquianda, Sacapalca, Las Conchas, Malacatos, Vilcabamba.	1000 a 1700 m.s.n.m.
Templado cálido	20.4	Palo Blanco, Lagunas, Pacaipampa, Jimbura, Quilanga, Quinara, Cariamanga, Colaisaca, Mambazo, Changaimina, Bellavista, El Cisne, Catacocha, Lauro Guerrero.	1700 a 2300 m.s.n.m.
Templado frío	7.28	Los Alisos, Minas, Seise, Ayabaca, Yangana, Pacaipampa, Gualel, Chuquiribamba.	2300 a 3000 m.s.n.m.
Semifrío	3.54	Arrendamientos, Las Pircas, Talaneo.	3000 a 3500 m.s.n.m.
Frío moderado	0.66	Las cumbres más altas de la cuenca.	> 3500 m.s.n.m.

Fuente: CONSORCIO ATA-UNL-UNP (2002)

1.2 Proyecto Especial Chira – Piura

1.2.1 Origen

El Proyecto Especial Chira – Piura, en siglas PECHP, es un sistema de riego integrado mediante la unión hidráulica de las cuencas de los ríos Chira y Piura. Se encuentra a cargo del órgano desconcentrado de ejecución del Gobierno Regional de Piura que recibe el mismo nombre. Tiene a su cargo la mayor inversión en materia de irrigación en el Perú, con el objetivo de maximizar la explotación de los valles Chira y Piura. La infraestructura permite una dotación de agua por gravedad para el riego oportuno y en cantidad suficiente, cumpliendo la labor de protección contra inundaciones. De esta manera logra incrementar la producción agrícola, mejorando la rentabilidad y competitividad agropecuaria y agroindustrial. (Proyecto Especial Chira Piura, 2020)

El proyecto busca solucionar situaciones anteriores de dependencia en base a las descargas naturales de ambos ríos, de régimen irregular, y que abastecían con agua suficiente para satisfacer la demanda hídrica de los usuarios agrícolas de la región.

Durante los primeros cincuenta años del siglo XX, la agricultura en el valle del río Piura disminuía gradualmente debido a la falta de agua. Por otro lado, a pesar de la gran cantidad de agua proveniente del río Chira, las tierras de cultivo del bajo Chira eran limitadas debido al alto costo de bombeo.

Después de grandes esfuerzos técnicos se hace realidad la iniciativa de aprovechar los excedentes de agua del río Chira para trasvasarlos a través de un canal al río Piura e irrigar sus valles.

1.2.2 Ejecución

La ejecución del proyecto se divide en tres grandes etapas, debido a la magnitud de las obras y la gran inversión a realizar. (Figura 3)

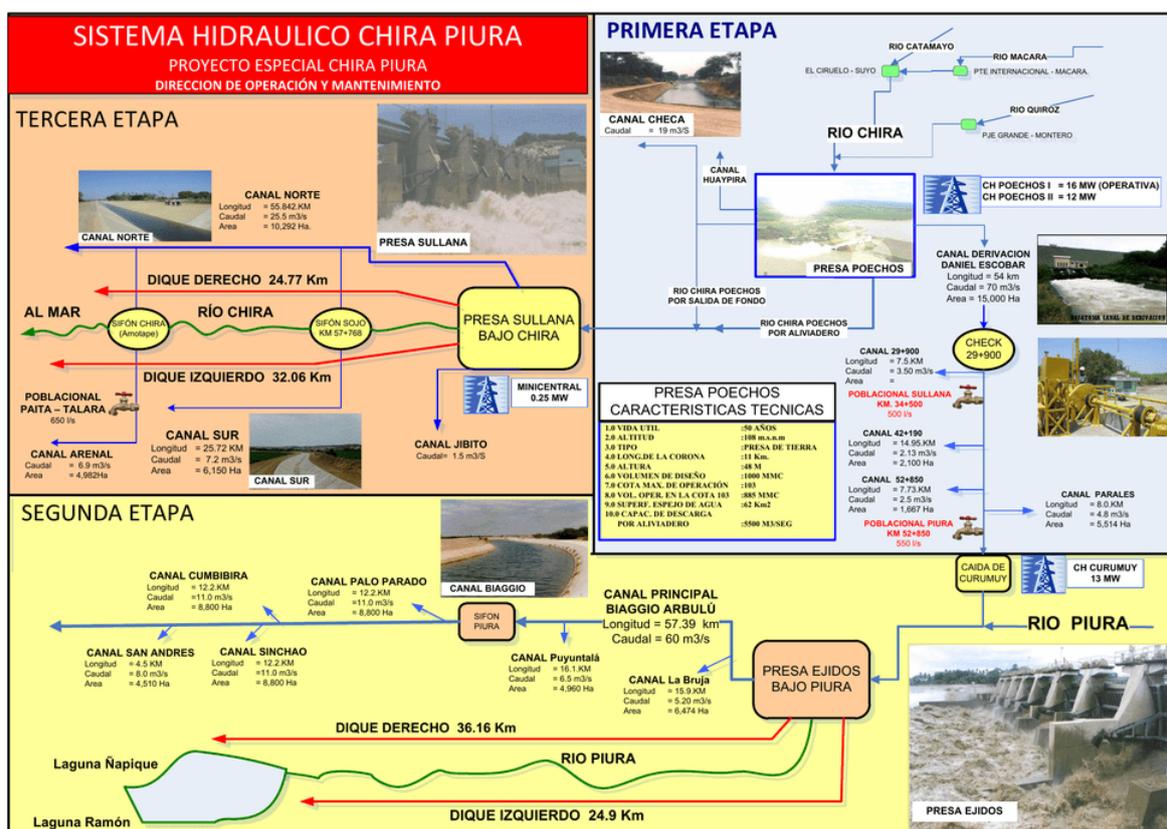


Figura 3. Etapas de ejecución Proyecto Especial Chira – Piura.

Fuente: Enríquez Beck (2016)

1º Etapa

Se ejecutó con el propósito de asegurar el abastecimiento de agua de riego regulada en los valles del Medio y Bajo Piura, además de realizar obras de drenaje para rehabilitar terrenos de cultivo. La obra fue autorizada con el D.S. N°290-70 AG del 4 de octubre de 1970. (Figura 4) Los trabajos consistieron en trasvasar el agua del río Chira a través del canal de derivación Daniel Escobar y se desarrollaron desde el año 1972 hasta su finalización en el año 1979. (Tabla 4)

Tabla 4. Obras ejecutadas en 1° Etapa Proyecto Especial Chira - Piura.

Obras	Descripción
Represa de Poechos	Con una capacidad de diseño para 1000 MMC y capacidad desde la cota de operación de 885 MMC.
Canal de derivación "Daniel Escobar"	Trasvasa agua del río Chira al río Piura. Tiene una longitud de 54 km y una capacidad de 70 m ³ /s.
Canal Parales	Tiene una longitud de 8 km y una capacidad de 4.8 m ³ /s para irrigar 5514 ha.
Canal Paralelo Cieneguillo	Tiene una longitud de 7.8 km y una capacidad de 6.2 m ³ /s para irrigar 5422 ha.
Drenes Troncales	Construcción de 452 km de drenes troncales en el Bajo Piura.
Ampliación Agrícola	Incremento de 5422 ha de frontera agrícola. Establecimiento agrícola de Cieneguillo.
Defensas	Construcción de 18 km de defensas contra inundaciones en puntos críticos del valle del Bajo Piura.

Fuente: Elaboración propia. Proyecto Especial Chira Piura I (2020)

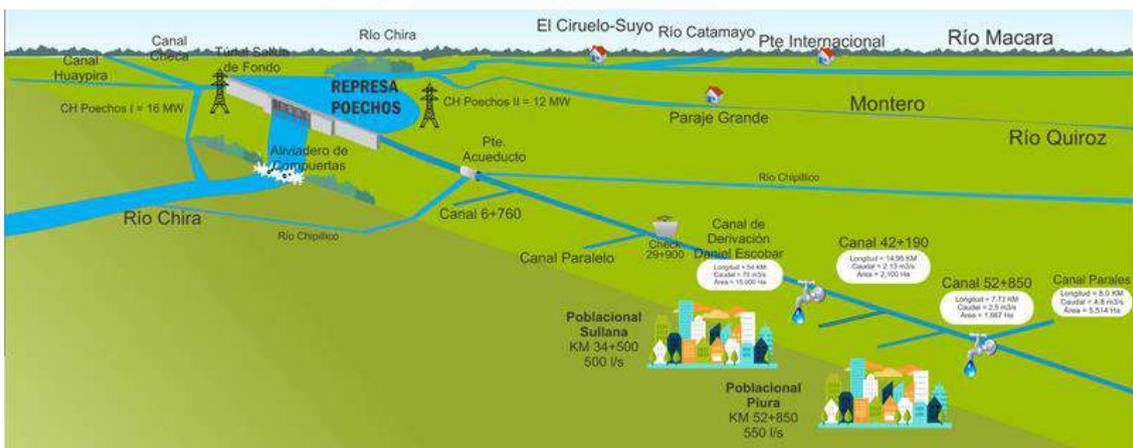


Figura 4. Primera etapa Proyecto Especial Chira - Piura.

Fuente: Proyecto Especial Chira Piura I (2020)

2° Etapa

Se desarrolló con el objetivo de aumentar la producción y productividad agrícola del valle del Bajo Piura e incorporar 5615 ha bajo riego. La obra fue autorizada con el D.S. N°101-79 EF del 25 de julio de 1970. Las obras comenzaron en 1980, durante el mes de enero y en 1989 se concluyó con la ejecución de los trabajos de reconstrucción de las obras dañadas por el fenómeno El Niño. (Tabla 5)

Tabla 5. Obras ejecutadas en 2° Etapa Proyecto Especial Chira - Piura.

Obras	Descripción
Presa Derivadora de los Ejidos	Capta las aguas provenientes de Poechos y del río Piura, derivándolas por el canal principal Biaggio Arbulú para irrigar el Valle del Bajo Piura y tiene una capacidad de 4.8 MMC.
Canal Principal Biaggio Arbulú	Tiene una longitud de 56 km desde Ejidos hasta Sechura (Chusis) y 60 m ³ /s de bocatoma.
Construcción de diques	Construcción de 63 km de diques de encauzamiento del río Piura, desde el puente Bolognesi en la ciudad de Piura hasta la Laguna Ramón.
Rehabilitación de tierras	Rehabilitación de 7980 ha de tierras afectadas con problemas de salinidad y drenaje.
Riego	Obras de riego y drenaje a nivel parcelario
Canales secundarios y terciarios	Construcción de 86 km de canales secundarios y terciarios revestidos de concreto.
Obras de reconstrucción FEN	Obras de reconstrucción Fenómenos El Niño 1983: Canal de derivación Chira - Piura.
Canal Principal Bajo Piura	Canal Principal del Bajo Piura, drenaje troncal del Bajo Piura, diques de encauzamiento del río Piura.
Extensión Agrícola	Servicios de extensión agrícola y riego tecnificado.
Estudios	Estudios definitivos de remodelación del Valle del Chira y estudio de factibilidad del Alto Piura.

Fuente: Elaboración propia. Proyecto Especial Chira Piura II (2020)

3° Etapa

Se realizó con la finalidad de irrigar por gravedad 37277.6 ha e incorporar a la agricultura 4908.4 ha ubicadas en el Valle del Chira. (Figura 5) De esta manera se eliminó el costoso y anticuado sistema de riego por bombeo. La ejecución de las obras se aprobó con el D.S. N° 072-89-MIRE el 19 de diciembre de 1989. (Tabla 6)



Figura 5. Tercera etapa Proyecto Especial Chira - Piura.

Fuente: Proyecto Especial Chira Piura III (2020)

Tabla 6. Obras ejecutadas en 3° Etapa Proyecto Especial Chira - Piura.

Obras	Descripción
Canal Miguel Checa	El canal permite el desarrollo de aproximadamente 14481 ha.
Presa derivadora Sullana	Tiene una capacidad de 6 MMC.
Canal Norte	Capacidad variable de 25.5 m ³ /s a 3.8 m ³ /s y con una longitud de 39.2 km. Tiene como estructura más importante Sifón Chira de 687 metros para trasvasar 6.9 m ³ /s.
Canal Sur	Capacidad variable de 7 m ³ /s a 0.55 m ³ /s y con una longitud de 25.75 km. Tiene como estructura más importante Sifón Sojo de 1515 metros de longitud.
Sistema de Drenaje	Comprende una red de drenes principales con una longitud de 52.69 km para drenar las áreas agrícolas afectadas por salinidad en el Valle del Chira.
Diques de Encauzamiento	Tiene una longitud de 57.03 km de diques de defensa y encauzamiento con sus respectivos espigones en ambas márgenes del río Chira.

Fuente: Elaboración propia. Proyecto Especial Chira Piura III (2020)

1.2.3 Infraestructura

Presas

Represa de Poechos

Tiene una capacidad de 885 Hm³ en su cota de operación normal a 103 m.s.n.m. Se dará un mayor alcance en el apartado 1.3.

Presa derivadora de los Ejidos

Ubicada en el cauce del río Piura, 7 km arriba de la ciudad. Su función es captar las aguas reguladas de la presa Poechos, así como del río Piura para suministrar agua al valle del Bajo Piura. Cuenta con un volumen de 4.8 Hm³ y su nivel normal de operación es a 30.5 m.s.n.m. Principales estructuras de la presa:

- Vertedero libre: Tiene una longitud de 120 m y una capacidad de evacuación de 950 m³/s.
- Aliviadero de compuertas: Está formado por 7 compuertas radiales, con una capacidad de evacuación de 1550 m³/s.
- Canal de limpia: Conformado por 2 compuertas radiales, con una capacidad de evacuación de 700 m³/s.
- Bocatoma canal principal: Con 3 compuertas radiales y capacidad de captación de 64 m³/s. (Aguilar Villena, 2016)

Presa derivadora de Sullana

Ubicada en el cauce del río Chira, a medio kilómetro aguas debajo de la ciudad de Sullana. Es una de las obras ejecutadas para la remodelación del Valle del Chira y tuvo un costo de 24.933 millones de dólares. Tiene una longitud de coronamiento de 362 m que permite aprovechar 250 millones de metros cúbicos de agua que anualmente se perdían en el mar. El nivel normal de operación es a 36.5 m.s.n.m. con un volumen de 6 Hm³. La estructura está compuesta de:

- Tomas de derivación hacia los canales norte, la minicentral hidroeléctrica y canal Capilla – Jíbito.
- Aliviadero de compuertas de 76 m de longitud con 8 compuertas radiales diseñadas para evacuar 2300 m³/s y rediseñadas.
- Aliviadero fijo a 290 m diseñado para evacuar 2700 m³/s y rediseñado para 4400 m³/s. (Aguilar Villena, 2016)

Canales Principales

Canal de derivación “Daniel Escobar”

El canal cuenta con 54 km de longitud, es de sección trapezoidal y está totalmente revestido en concreto. Tiene una capacidad máxima de 75 m³/s.

Transporta las aguas reguladas del río Chira desde Poechos hasta la cuenca del río Piura. Estas aguas abastecen demanda agrícola, industrial, energética y poblacional, con un total de 27 tomas de diferentes tamaños a lo largo de su longitud, con anchos de entre 2 m a 5 m para captar caudales de 10 l/s, 100 l/s, etc. (Aguilar Villena, 2016)

Canal Miguel Checa

El canal más importante del distrito de riego del valle del Chira. Es un canal de sección trapezoidal de 79 km de longitud. Diseñado para un caudal inicial de 19 m³/s en bocatoma y en el extremo final de 1 m³/s. Transporta las aguas del río Chira reguladas en Poechos hacia el valle del Chira. (Aguilar Villena, 2016)

Canal Norte

Canal que conduce 25.5 m³/s y tiene un recorrido de 39 km hasta la altura de la ciudad de Amotape. Recorre la margen derecha del río Chira y cruza hacia la margen izquierda a través del Sifón Chira. Trasvasa 6.9 m³/s para irrigar 4550 ha en el sector “El Arenal”. Conduce las aguas reguladas en la Presa de Sullana hacia la parte baja del valle. (Aguilar Villena, 2016)

Estaciones Hidrometeorológicas

Desde el año 1972 se han construido estructuras de medición para la evaluación del recurso hídrico, que generan información utilizada para diversos estudios y constituye la herramienta principal del Reservorio de Poechos.

Cada estación cuenta con:

- Infraestructura limnigráfica constituida por la caseta y tubo de fierro.
- Equipo registrador (limnigrafo Steveen) de bandas de papel especial.
- Batería de miras de fierro o pintadas en la estructura para lecturas de nivel de agua.
- Sistema de comunicación con radios HF5 antenas y batería para generar energía.
- Módulo de material noble que sirve como vivienda – oficina.
- Un equipo (pluviómetro) para medir las precipitaciones pluviales.

Las estaciones establecidas en la cuenca del Chira se detallan en la Tabla 7.

Tabla 7. Estaciones meteorológicas en la cuenca del río Chira.

N°	Estación	Ubicación Geográfica		Cuenca	Coordenadas Geográficas		Altitud (m.s.n.m.)	Categoría	Entidad responsable	Estado
		Provincia	Distrito		Latitud	Longitud				
1	Ardilla	Sullana	Lancones	Chira	04°31'	80°26'	150	H-PLU-PG	PECHP	Operativa
2	Ciruelo	Ayabaca	Suyo	Chira	04°16'06"	80°09'11"	250	H-PLU-PG	PECHP	Operativa
3	Paraje Grande	Ayabaca	Paimas	Chira	04°37'	79°54'	555	H-PLU	PECHP	Operativa
4	Puente Internacional	Ayabaca	Macará - Suyo	Chira	04°23'19"	79°57'43"	408	H-PLU	PECHP	Operativa
5	Alamor	Sullana	Lancones	Chira	04°28'48,41"	80°23'56,9"	133	PLU	SENAMHI	Operativa
6	Los Encuentros	Sullana	Lancones	Chira	04°26'1"	80°17'1"	150	H	SENAMHI	Cerrada
7	Jilili	Ayabaca	Jilili	Chira	04°35'1'	79°48'1"	1236	PLU	SENAMHI	Cerrada
8	Sicchez	Ayabaca	Sicchez	Chira	04°34'	79°46'	1292	PLU	SENAMHI	Cerrada
9	Aul (C. Membrillo)	Ayabaca	Ayabaca	Chira	04°33'	79°42'	2450	PLU	SENAMHI	Cerrada
10	Vado Grande	Ayabaca	Ayabaca	Chira	04°27'37"	79°36'23"	1006	PLU	SENAMHI	Cerrada
11	Huara de Veras	Ayabaca	Ayabaca	Chira	04°35'1"	79°34'1"	1243	PLU	SENAMHI	Cerrada
12	Espíndola	Ayabaca	Ayabaca	Chira	04°38'1"	79°30'1"	1900	PLU	SENAMHI	Cerrada

Fuente: Enríquez Beck (2016)

1.3 Reservoirio de Poechos

El Reservoirio de Poechos constituye la obra más importante dentro de la primera etapa del Proyecto Especial Chira – Piura y es la estructura principal en el sistema de riego Chira – Piura. Su objetivo principal es almacenar los excedentes de agua del río Chira durante el periodo lluvioso y con ella asegurar la satisfacción de la demanda hídrica de los valles Piura y Chira.

Su construcción se inició en el año de 1972 y culminó en 1976. La obra fue construida por Energoprojekt y supervisada por Binnie & Partners Corporación Peruana de Ingeniería S.A. y fue puesta en operación el 4 de junio de 1976.

Está conformado por el dique principal que cierra el lecho del río Chira, con 13 km de largo y 48 m de altura, los diques laterales izquierdo y derecho; dando lugar a un embalse con capacidad de diseño de 1000 MMC.

El reservorio permite irrigar más de 108 mil hectáreas, en los valles de Chira y Piura. Permite la generación de 42 MW de energía, lograda en las centrales hidroeléctricas de Curumuy, Poechos I y II. Además, abastece de agua para uso poblacional a las ciudades de Piura, Sullana y Paita. (Varona Alama, 2018)

1.3.1 Ubicación

La represa de Poechos está ubicado sobre el río Chira, al noroeste del Perú, en una zona costera y desértica, a 40 km al noreste de la ciudad de Sullana y a unos 60 km al norte de la ciudad de Piura. La zona rodea el reservorio pertenece al sistema climático de bosque seco subtropical. Se encuentra dentro de la Cuenca Binacional Catamayo – Chira.

1.3.2 Infraestructura

Aliviadero de compuertas

Busca controlar las descargas de agua hacia el río Chira en época de avenidas. Está compuesto por tres compuertas radiales, de 12 m de altura por 10m de ancho y 210 toneladas de peso cada una. Cuenta con una capacidad máxima de descarga de 5400 m³/s. La operación se realiza basado en el análisis de los datos hidrológicos obtenidos en las estaciones Ciruelo, Paraje Grande y Ardilla. (Varona Alama, 2018)

Aliviadero de emergencia

Ubicado en el dique izquierdo de la presa. Es un solado de concreto de 400 m de longitud con muros laterales. Está compuesto por 4 cuerpos individuales de 100 m cada uno, sobre el cual existe un relleno fusible de tierra provisto para ser erosionado cuando las aguas excedan al nivel máximo de seguridad de la presa. La capacidad máxima de descarga es de 10000 m³/s. (Varona Alama, 2018)

Salida al canal de derivación “Daniel Escobar”

Se encuentra localizado en el dique izquierdo de la Presa. El canal es de sección trapezoidal, revestido de concreto a lo largo de sus 54 km de longitud, con una capacidad máxima de 70 m³/s. Posee dos compuertas radiales al final de dos conductos de 2.4 m de diámetro. (Varona Alama, 2018)

Salida a fondo

Es una estructura de concreto armado con blindaje. Tiene una longitud de 415 m de longitud y 300 m³/s de capacidad. Está constituida de compuertas de rueda de 4.5 m de diámetro, válvula de mariposa y compuerta radial. Permite la irrigación del valle del Chira, siendo la salida hacia los canales laterales Miguel Checa y Huaypirá. (Varona Alama, 2018)

1.3.3 Operación

La operación del reservorio se realiza en dos modalidades:

En el periodo de operación normal, los aportes hídricos son siempre menores a la demanda. Para este caso, el vaciado del reservorio se realiza de acuerdo a los PCR, Planes de Cultivo y Riego (m³/ha/campaña), y a las necesidades de los usuarios; realizando un permanente control de las descargas y los tiempos de duración.

En el periodo de operación de emergencia, el cual se da cuando los aportes hídricos superan a la demanda, se descarga al río Chira, por la necesidad de mantener el almacenamiento al nivel normal de operación, de acuerdo con las reglas de operación vigente. (Enríquez Beck, 2016)

La operación del reservorio se rige sobre las siguientes consideraciones:

- El embalse no tiene un volumen específico de diseño para la mitigación de avenidas extraordinarias.
- Al final de cada periodo lluvioso es preciso asegurar un embalse completamente lleno.
- Coordinar la limpieza del cauce del río Chira aguas abajo de la presa, para reducir el riesgo de inundación.

La operación del Reservorio de Poechos se torna compleja pues el volumen de agua acumulado dentro del reservorio solamente asegura la regulación parcial del Balance Hídrico en años lluviosos y porque ante avenidas extraordinarias, la capacidad física del cauce del río aguas debajo de la represa es limitada y no admite desembalses incontrolados. (Figura 6)

Los protocolos de operación para el reservorio de Poechos han variado a lo largo de los años:

- 1976: Manual de operación preparado por Energoprojekt.
- 1977: Reglas modificadas – Energoprojekt.
- 1978: Acuerdos sobre la operación del Embalse de Poechos (“Reglas Hudson”)
- 1984: Reglas de Operación Presa Poechos (“Reglas Araujo”)
- 1999: Reglas formuladas por la Empresa Consultora Colpex Project (Vigentes actualmente) (Enríquez Beck, 2016)

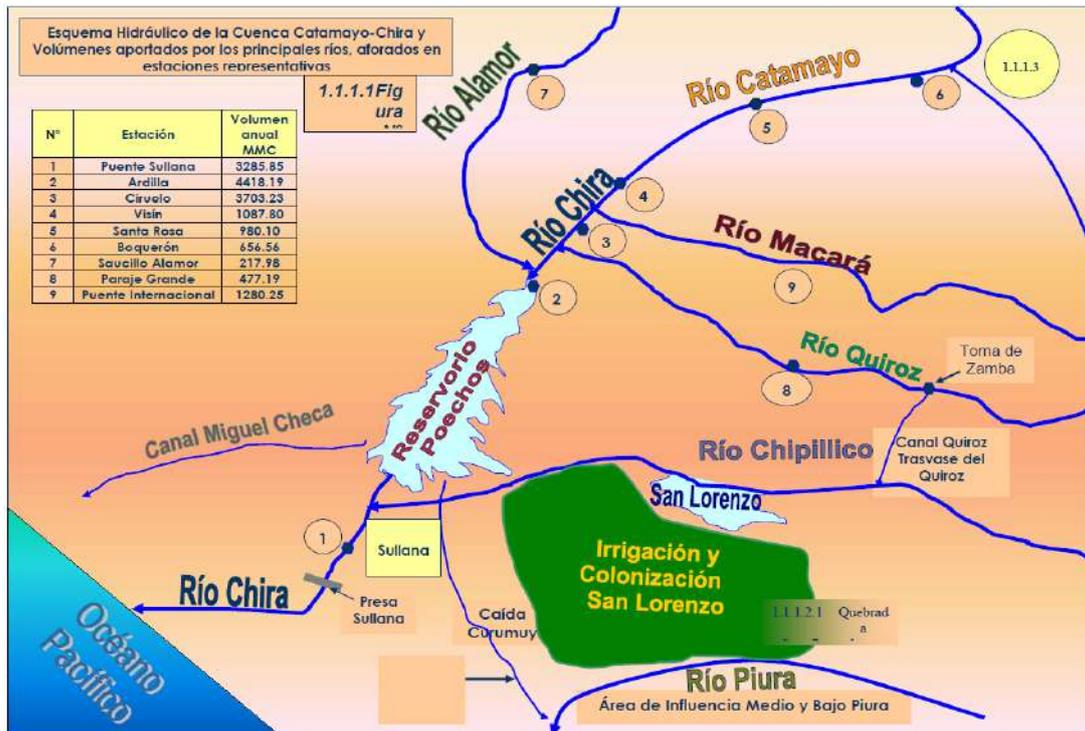


Figura 6. Volúmenes aportados por los principales ríos al esquema hidráulico de la cuenca Catamayo - Chira.

Fuente: CONSORCIO ATA-UNL-UNP (2002)

1.3.4 Problema de colmatación

El reservorio de Poechos fue diseñado para almacenar una capacidad de 885 MMC, a la cota de operación normal 103, incluido el volumen muerto (180 MMC). Sin embargo, debido a la acumulación de sedimentos, el reservorio actualmente almacena 397 MMC. (Figura 7)

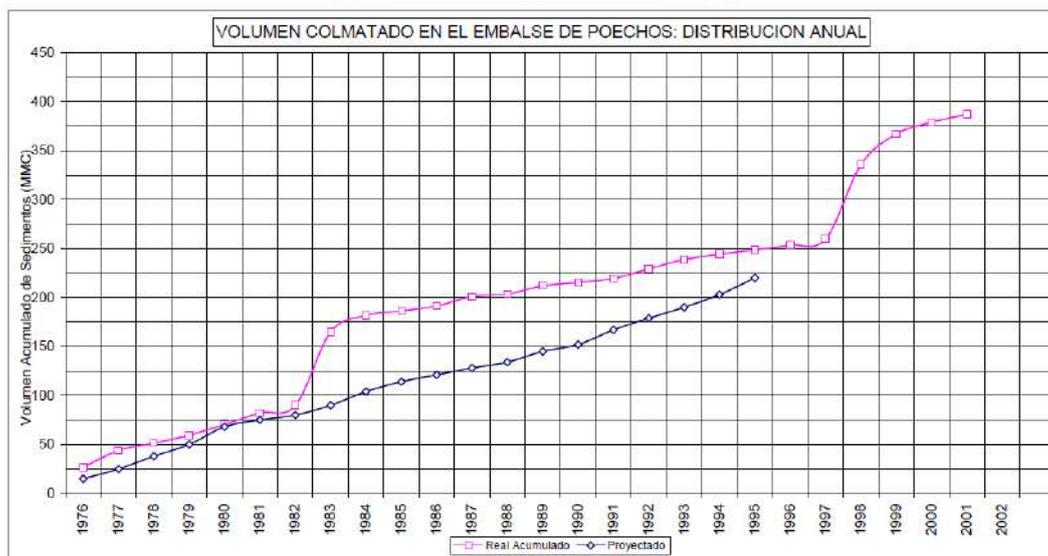


Figura 7. Acumulación de sedimentos en reservorio de Poechos.

Fuente: CONSORCIO ATA-UNL-UNP (2002)

El proceso de colmatación corresponde a la erosión de los suelos y de las riveras de los cursos de agua de la cuenca, por acción de la lluvia y actividades antropogénicas no controladas. (Figura 8) El volumen acumulado de sedimentos representa el 55.14% del volumen inicial de operación. (Reyes Salazar, Farías de Reyes, Guerrero Asmad, & Tesén Arámbulo, 2018)

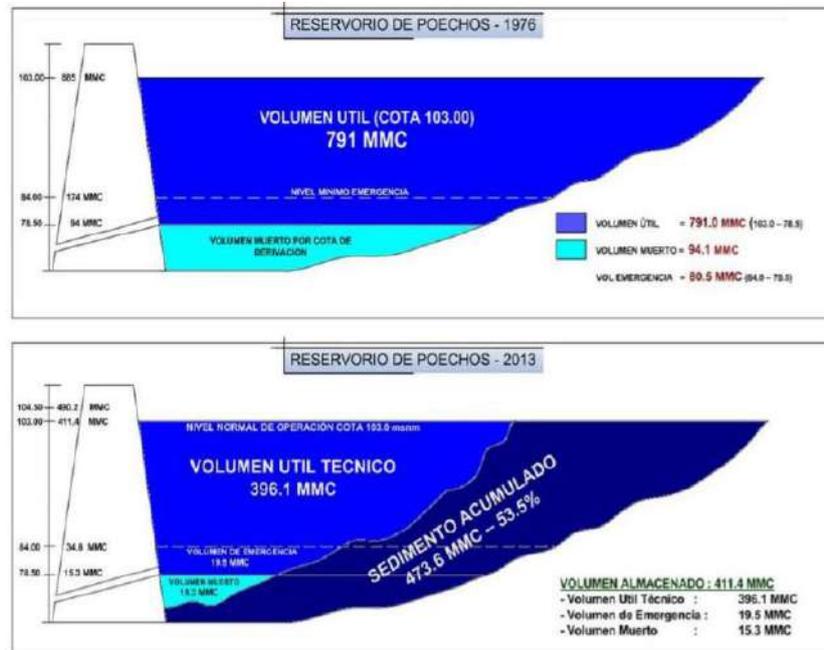


Figura 8. Comparativa del volumen del reservorio, según batimetría de 2012.

Fuente: Aguilar Villena (2016)

Desde el comienzo de su operación, se han realizado 26 campañas de medición. Los resultados indican que en años hidrológicos normales, la acumulación de sedimentos corresponde a lo previsto en los parámetros de diseño (8 a 9 MMC por año). Por el contrario, después de los fenómenos El Niño sucedidos en 1983 y 1998, el depósito de sedimentos se elevó hasta 75 MMC por año. (Figura 9)

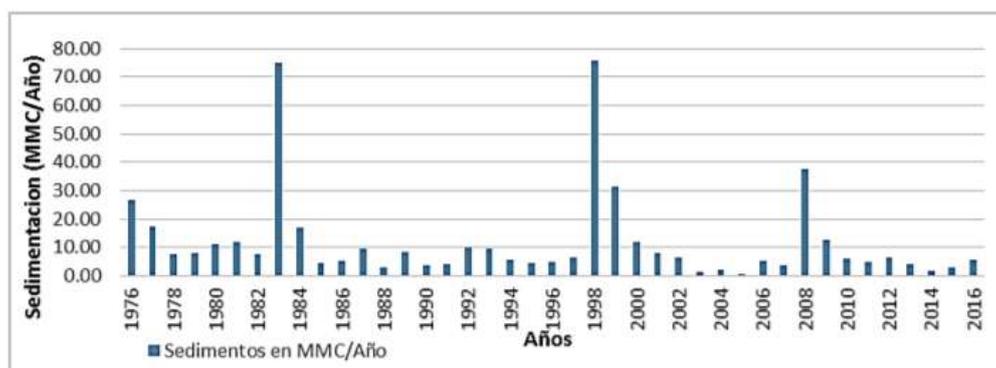


Figura 9. Retención de sedimentos en el reservorio de Poechos.

Fuente: Reyes Salazar, Farías de Reyes, Guerrero Asmad, & Tesén Arámbulo (2018)

Capítulo 2

Fundamentación teórica de series temporales

El capítulo presenta la fundamentación teórica que constituye la base para la construcción de modelo de series temporales. La información para el desarrollo de este capítulo es extraída de los libros: Análisis de series temporales (Peña, 2005), Time series analysis: forecasting and control (Box & Jenkins, 1976) y Forecasting: principles and practice (Hyndman & Athanasopoulos, 2018)

Una serie temporal se define como una colección de observaciones de una variable a lo largo del tiempo en intervalos regulares, esto es, en instantes de tiempo equiespaciados. Si se puede predecir los valores exactos de la serie se considera determinística. En caso los valores de la serie solo se puedan determinar de manera parcial, la serie se denomina estocástica. (Peña, 2005)

Los modelos univariantes buscan predecir valores futuros de la serie a través de la explicación histórica de una variable a lo largo del tiempo, es decir, se basan únicamente en la historia de la propia serie.

Definiendo una serie temporal compuesta por una secuencia de observaciones de variables aleatorias (z_t) donde la serie se considera como un proceso estocástico. El modelo univariante para la predicción tiene la forma:

$$\hat{z}_{t+h|t} = f(z_t, z_{t-1}, z_{t-2}, z_{t-3}, \dots)$$

El subíndice t representa el tiempo. La variable $\hat{z}_{t+h|t}$ representa la predicción en el instante $t + h$ respecto del instante t , donde h representa el horizonte de predicción. Esta predicción es una función de los valores pasados de la serie, sin considerar variables externas que puedan afectar el proceso.

La observación real de z en el instante $t + h$ difiere de la predicción en un valor al que se denomina $e_{t+h|t}$ que representa el error debido a las innovaciones en los instantes $t + 1, \dots, t + h$. El error de predicción es producto de las “innovaciones”, que son la variación aleatoria de la serie y los efectos de las variables externas relevantes que no están incluidas en el modelo. (Hyndman & Athanasopoulos, 2018)

$$z_{t+h} = \hat{z}_{t+h|t} + e_{t+h|t}$$

2.1 Métodos de suavizamiento exponencial

Los métodos de suavización exponencial datan de la década de 1950 y en base a ellos se ha desarrollado una gran cantidad de modelos de pronóstico exitosos. La predicción se realiza utilizando promedios ponderados de observaciones anteriores dentro de una serie. La importancia de las observaciones para el modelo va decayendo exponencialmente a medida que avanza la serie. De esta manera, las observaciones recientes tienen mayor influencia en los modelos. (Hyndman & Athanasopoulos, 2018)

2.1.1 Suavizamiento exponencial simple

El método de suavizamiento exponencial simple consiste en explicar una observación dentro de una serie de tiempo a través de una media ponderada de los valores previos de la serie con pesos que decrecen geométricamente. Así, las observaciones más recientes de la serie son más relevantes para el modelo que las observaciones anteriores.

$$\hat{z}_{T+1|T} = w_0 z_T + w_1 z_{T-1} + w_2 z_{T-2} + \dots$$

$$\hat{z}_{T+1|T} = \sum_{t=0}^{\infty} w_t z_{T-t}$$

Siendo w_t los pesos asignados a los valores pasados de la serie, los cuales suman uno. Para imponer la condición de que las observaciones próximas son más importantes y que decrecen geométricamente, utilizamos la constante de suavizamiento α , donde $0 < \alpha < 1$, tal que:

$$w_t = (1 - \alpha)\alpha^t$$

La ecuación de predicción queda expresada de la siguiente manera:

$$\hat{z}_{T+1} = \alpha z_T + (1 - \alpha)\hat{z}_T$$

Los modelos de suavizamiento exponencial simple pueden ser usados solamente con series de tiempo sin tendencia sistemática o componentes estacionales.

2.1.2 Suavizamiento exponencial doble (Holt)

El método de suavizamiento exponencial doble, también conocido como método de Holt en honor a su desarrollador, permite analizar series temporales que muestran tendencia, sea esta constante o no. Basada en el modelo anterior, intenta explicar una observación a través de las observaciones previas de la serie. En esta ocasión, se utilizan dos ecuaciones de suavizamiento: una para el nivel de la serie y otra para la tendencia de la misma. (Hyndman & Athanasopoulos, 2018)

$$\hat{z}_{t+h|t} = L_t + hT_t$$

Donde:

L_t representa la ecuación de nivel:

$$L_t = \alpha z_t + (1 - \alpha)(L_{t-1} + T_{t-1})$$

T_t representa la ecuación de tendencia:

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}$$

Para construir las predicciones, el método de Holt generaliza la idea del suavizamiento exponencial simple introduciendo dos factores de descuento, donde α es el parámetro de suavizamiento para el nivel y β es el parámetro de suavizamiento para la tendencia. Se cumple que $0 < \alpha < 1$ y $0 < \beta < 1$.

Las predicciones a partir del método de Holt evidencian una tendencia que aumenta o disminuye en el tiempo de forma indefinida. Los estudios empíricos demuestran que esta tendencia constante disminuye la efectividad del método para horizontes de pronóstico (h) más largos. En respuesta a este problema, Gardner y Mc Kenzie desarrollaron un modelo que introduce un parámetro adicional con el objetivo de amortiguar la tendencia para futuras observaciones. Este modelo más flexible se conoce como modelo de suavizamiento exponencial con tendencia amortiguada y cuenta con gran aceptación por su mayor precisión de pronóstico. (Hyndman & Athanasopoulos, 2018)

Introduciendo este nuevo parámetro, la ecuación de predicción se expresa como:

$$\hat{z}_{t+h|t} = L_t + (\varphi + \varphi^2 + \dots + \varphi^h)T_t$$

Donde:

φ es el parámetro de amortización y $0 < \varphi < 1$

L_t representa la ecuación de nivel:

$$L_t = \alpha z_t + (1 - \alpha)(L_{t-1} + \varphi T_{t-1})$$

T_t representa la ecuación de tendencia:

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)\varphi T_{t-1}$$

2.1.3 Suavizamiento exponencial triple (Holt-Winters)

El modelo de suavizamiento exponencial triple, introducido por Holt y Winters, extiende el método de Holt, capturando la estacionalidad de la serie. Se compone de tres ecuaciones de suavizamiento: la primera para el nivel de la serie, la segunda para la tendencia y se adiciona una tercera, para la componente estacional de la serie. Así, el método extrapola el nivel y la tendencia de la serie y utiliza la información de ciclos anteriores para extrapolar la estacionalidad.

Existen dos variantes de este modelo. Cuando las variaciones del componente estacional son aproximadamente constantes se utiliza el método aditivo. Por el contrario,

cuando las variaciones estacionales cambian de manera proporcional al nivel de la serie, se utiliza el método multiplicativo. (Hyndman & Athanasopoulos, 2018)

Método Aditivo

$$\hat{z}_{t+h|t} = L_t + hT_t + S_{t+h-s(k+1)}$$

Donde:

L_t representa la ecuación de nivel:

$$L_t = \alpha(z_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + T_{t-1})$$

T_t representa la ecuación de tendencia:

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}$$

S_t representa la ecuación de estacionalidad:

$$S_t = \gamma(z_t - L_{t-1} - T_{t-1}) + (1 - \gamma)S_{t-s}$$

El valor s representa la frecuencia de la estacionalidad y k es la parte entera de $\frac{(h-1)}{s}$, con el objetivo las estimaciones de los índices estacionales utilizados para la revisión provengan del último año de la muestra. Los parámetros de suavizamiento $\alpha, \beta, \gamma \in [0,1]$.

Método Multiplicativo

$$\hat{z}_{t+h|t} = (L_t + hT_t)S_{t+h-s(k+1)}$$

Donde:

L_t representa la ecuación de nivel:

$$L_t = \alpha \frac{z_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + T_{t-1})$$

T_t representa la ecuación de tendencia:

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)T_{t-1}$$

S_t representa la ecuación de estacionalidad:

$$S_t = \gamma \left(\frac{z_t}{L_{t-1} + T_{t-1}} \right) + (1 - \gamma)S_{t-s}$$

El valor s representa la frecuencia de la estacionalidad y k es la parte entera de $\frac{(h-1)}{s}$, con el objetivo las estimaciones de los índices estacionales utilizados para la revisión provengan del último año de la muestra. Los parámetros de suavizamiento $\alpha, \beta, \gamma \in [0,1]$.

Utilizando el método multiplicativo, el componente estacional se expresa en términos relativos y la serie se ajusta estacionalmente dividiendo por el componente estacional.

Se puede mejorar las estimaciones del método de Holt-Winters utilizando el método de tendencia amortiguada. De esta manera, se podrá ampliar el horizonte de predicción del método. Las ecuaciones afectadas por el parámetro de amortización (φ) resultan:

Método Aditivo con tendencia amortiguada

$$\hat{z}_{t+h|t} = L_t + (\varphi + \varphi^2 + \dots + \varphi^h)T_t + S_{t+h-s(k+1)}$$

Donde:

L_t representa la ecuación de nivel:

$$L_t = \alpha(z_t - S_{t-s}) + (1 - \alpha)(L_{t-1} + \varphi T_{t-1})$$

T_t representa la ecuación de tendencia:

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)\varphi T_{t-1}$$

S_t representa la ecuación de estacionalidad:

$$S_t = \gamma(z_t - L_{t-1} - \varphi T_{t-1}) + (1 - \gamma)S_{t-s}$$

Los parámetros $\alpha, \beta, \gamma, \varphi \in [0,1]$

Método Multiplicativo con tendencia amortiguada

$$\hat{z}_{t+h|t} = (L_t + (\varphi + \varphi^2 + \dots + \varphi^h)T_t)S_{t+h-s(k+1)}$$

Donde:

L_t representa la ecuación de nivel:

$$L_t = \alpha \frac{z_t}{S_{t-s}} + (1 - \alpha)(L_{t-1} + \varphi T_{t-1})$$

T_t representa la ecuación de tendencia:

$$T_t = \beta(L_t - L_{t-1}) + (1 - \beta)\varphi T_{t-1}$$

S_t representa la ecuación de estacionalidad:

$$S_t = \gamma \left(\frac{z_t}{L_{t-1} + \varphi T_{t-1}} \right) + (1 - \gamma)S_{t-s}$$

Los parámetros $\alpha, \beta, \gamma, \varphi \in [0,1]$

2.2 Métodos ARIMA

2.2.1 Procesos autorregresivos (AR)

Los procesos autorregresivos son modelos estocásticos útiles para una representación práctica de una serie. Son modelos sencillos que buscan explicar el valor actual de la variable de interés como una combinación lineal de los valores previos del proceso y las innovaciones del mismo (ruido blanco). Se consideran como procesos de memoria larga porque el valor

actual está relacionado con todos los anteriores, aunque en diferente proporción. (Box & Jenkins, 1976)

Estos modelos se caracterizan por utilizar gran cantidad de coeficientes de autocorrelación, distintos de cero, y que decrecen de forma exponencial con el retardo. Además, pueden ser procesos estacionarios o no.

Una serie temporal z_t , con valores del proceso espaciados de igual manera en el tiempo, sigue un proceso autorregresivo de orden p , si:

$$z_t = \phi_1 z_{t-1} + \phi_2 z_{t-2} + \dots + \phi_p z_{t-p} + a_t$$

Este es un proceso de media μ , donde ϕ es conocido como el operador de autorregresión, cumpliendo que $|\phi_p| < 1$ para procesos estacionarios y a_t representa el término residual que no puede ser explicado por el modelo. Es el error de predicción para el primer horizonte, también conocido como innovación ($e_{t|t-1} = a_t$), el cual es un proceso de ruido blanco. La serie temporal también puede ser representada como un proceso al cual se le ha restado la media:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} + a_t$$

$$\tilde{z}_t = z_t - \mu$$

Una serie autorregresiva se puede expresar también a través del operador de retardos (B). Este es un operador tal que al aplicarlo a un elemento de una serie de tiempo produce el elemento anterior: $B(z_t) = z_{t-1}$. (Box & Jenkins, 1976) De esta forma, los procesos autorregresivos se definen como:

$$\phi_p(B)\tilde{z}_t = a_t$$

Los procesos autorregresivos se representan por las siglas AR(p), donde p indica el orden de dependencia de las observaciones pasadas de la serie. En estos modelos, ϕ es una constante a determinar y a_t es un proceso de ruido blanco con varianza σ^2 .

2.2.2 Procesos de media móvil (MA)

A diferencia de los procesos autorregresivos, los procesos de media móvil buscan explicar el valor actual de la variable de interés a través de las innovaciones pasadas del proceso. Es decir, es un modelo que usa una regresión de los términos residuales. Son procesos memoria más corta y se representan por las siglas MA(q), donde q indica un número finito de innovaciones anteriores. (Peña, 2005)

La representación de un modelo MA(q) es:

$$z_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}$$

Siendo z_t una serie temporal de media μ , θ son los pesos que cumplen $|\theta_q| < 1$ para que las innovaciones pasadas influyan menos y a_t un proceso de ruido blanco con varianza

σ^2 . Este proceso se puede interpretar como una aportación nueva de información de la innovación a_t que tiene validez en el instante t hasta el instante $t-q$.

El proceso también puede ser representado restando la media:

$$\tilde{z}_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}$$

$$\tilde{z}_t = z_t - \mu$$

De igual manera, se puede escribir de forma compacta, utilizando el operador de retardos:

$$\tilde{z}_t = \theta_q(B)a_t$$

Un proceso MA(q) es un proceso estacionario, por ser una combinación lineal de procesos de ruido blanco (a_t).

La función de autocorrelación queda descrita como:

$$\rho_k = \frac{\sum_{i=0}^{i=q-k} \theta_i \theta_{k+i}}{\sum_{i=0}^{i=q} \theta_i^2}, \quad k = 1, \dots, q$$

$$\rho_k = 0, \quad k > q$$

Donde $\theta_0 = -1$ y $\theta_k = 0$ para $k \geq q + 1$

2.2.3 Funciones de autocorrelación simple y parcial

La función de autocorrelación se define como la correlación de la serie de tiempo consigo mismo. La autocorrelación entre el valor z_t y z_{t+k} , separados por k intervalos de tiempo sería:

$$\rho_k = \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sqrt{E(z_t - \mu)^2 E(z_{t+k} - \mu)^2}} = \frac{E[(z_t - \mu)(z_{t+k} - \mu)]}{\sigma_z^2}$$

En un proceso estacionario, la varianza de la serie es igual a la autocovarianza ($\gamma_k = cov[z_t, z_{t+k}] = E[(z_t - \mu)(z_{t+k} - \mu)]$) en el instante cero, $\sigma_z^2 = \gamma_0$. Así, la autocorrelación para el retardo k , que es la correlación entre z_t y z_{t+k} se puede expresar:

$$\rho_k = \frac{\gamma_k}{\gamma_0}$$

La función de autocorrelación simple (FAS) proporciona el coeficiente de correlación entre las observaciones de una misma serie, comprendidas a lo largo de un número determinado de periodos.

La función de autocorrelación parcial (FAP) muestra la relación directa que existe entre dos observaciones, separadas por k retardos. (Box & Jenkins, 1976)

Ambas funciones sirven como patrones de referencia para identificar las funciones de autocorrelación muestral y de autocorrelación parcial muestral de una serie temporal en

estudio. Los procesos autorregresivos AR(p) en la FAS presentan gran cantidad de coeficientes no nulos, mientras que, en la FAP, los primeros p reciben coeficientes no nulos y los siguientes cero. En el caso de los procesos de media móvil MA(q), en la FAS los primeros q reciben coeficientes no nulos y los siguientes cero, pero en la FAP muchos coeficientes son no nulos. En la Figura 10 se aprecia las funciones de autocorrelación para los procesos AR (1) y AR (2). En la Figura 11 se presenta las funciones de autocorrelación para los procesos MA (1) y MA (2). Los diagramas de la función de autocorrelación son también conocidos como correlograma.

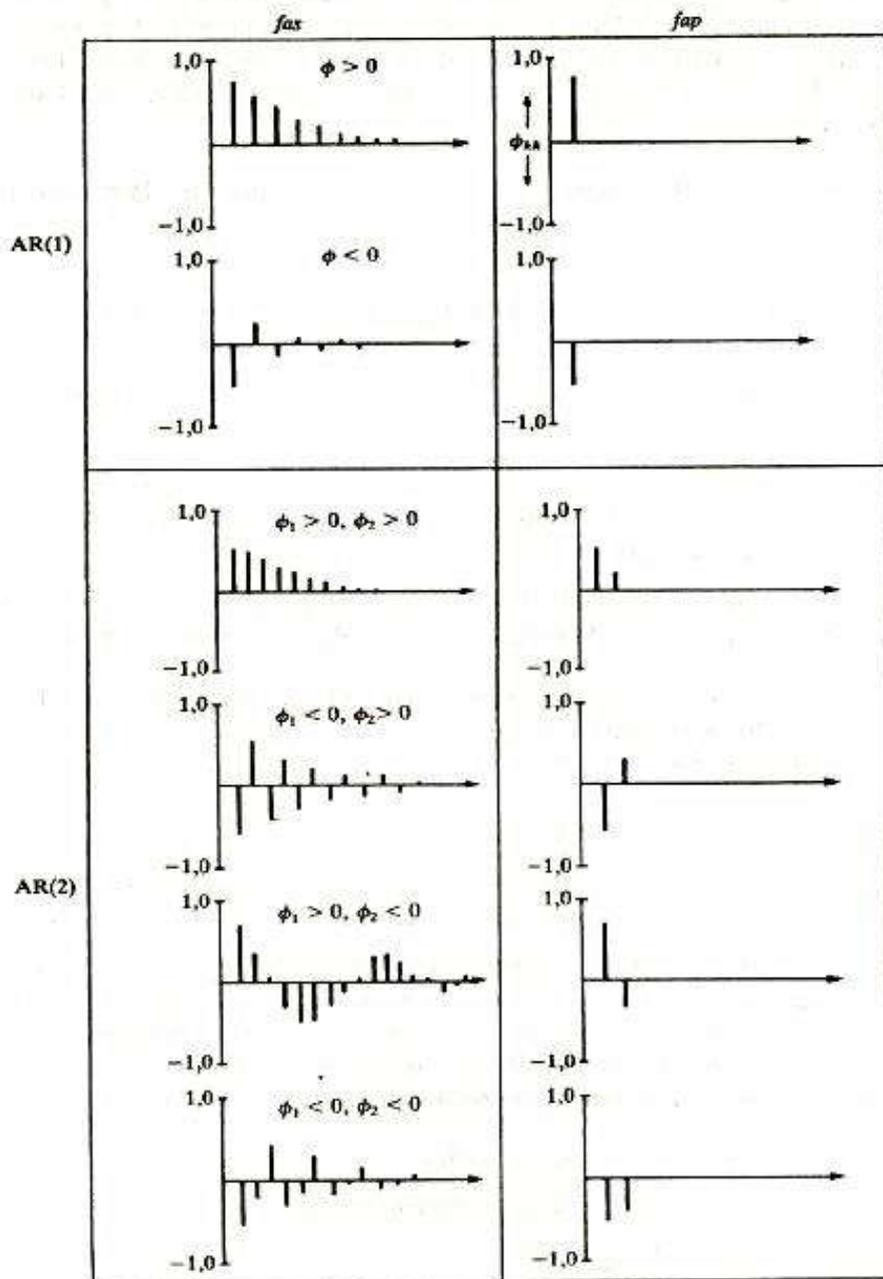


Figura 10. Funciones de autocorrelación simple y parcial de modelos autorregresivos.

Fuente: Peña (2005)

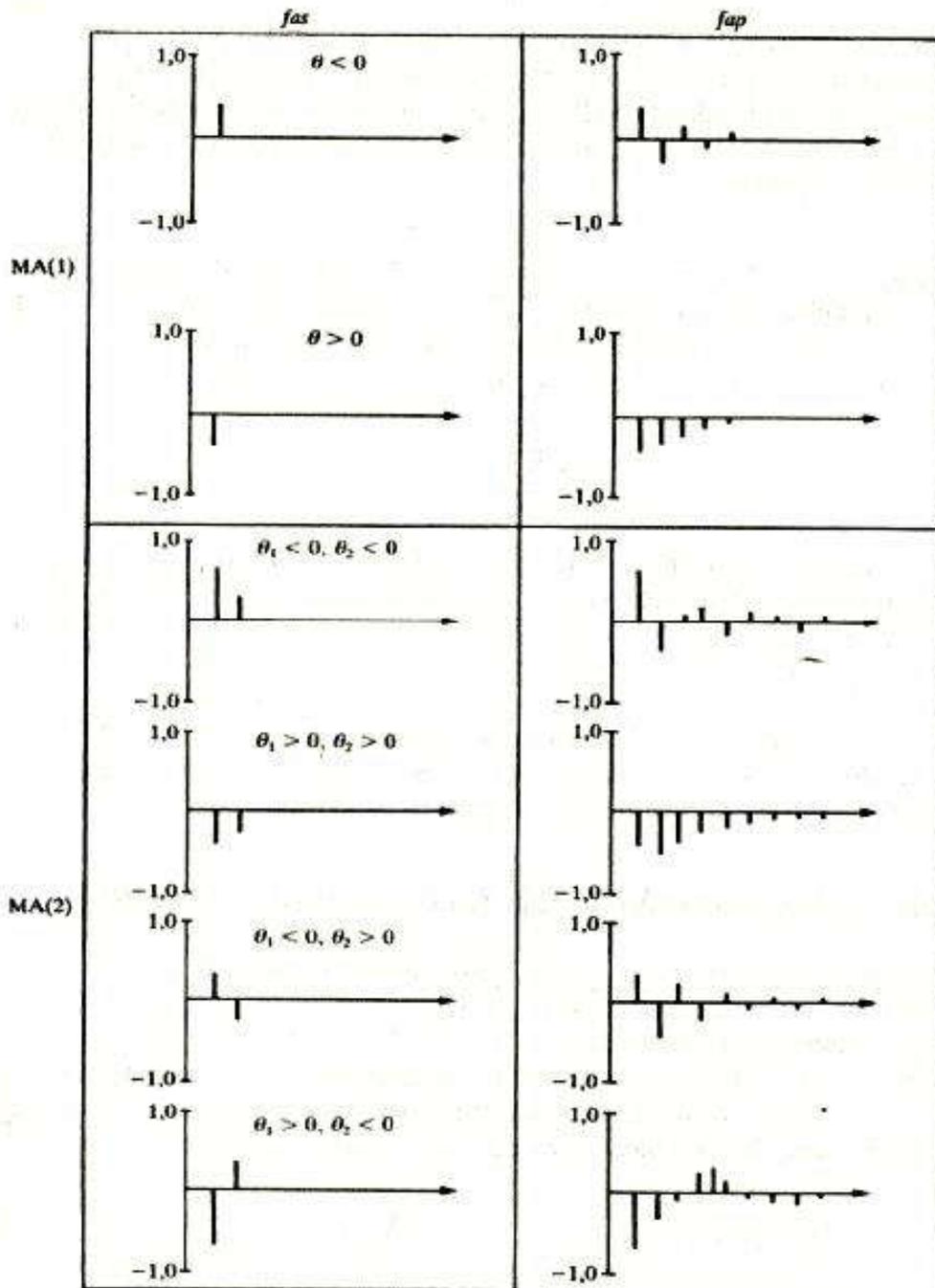


Figura 11. Funciones de autocorrelación simple y parcial de procesos de media móvil.

Fuente: Peña (2005)

2.2.4 Procesos ARIMA

Para mejorar la flexibilidad y el ajuste en las series temporales, nacen los procesos ARMA (p,q), una combinación de las dos estructuras (autorregresiva y de media móvil) donde la observación actual de la variable de interés se explica a través de las observaciones anteriores, las innovaciones anteriores y la innovación en ese instante.

Un proceso ARMA (p,q) se expresa:

$$\tilde{z}_t = \phi_1 \tilde{z}_{t-1} + \phi_2 \tilde{z}_{t-2} + \dots + \phi_p \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \theta_2 a_{t-2} - \dots - \theta_q a_{t-q}$$

La expresión se puede reducir utilizando el operador de retardos:

$$(1 - \phi_1 B - \dots - \phi_p B^p) \tilde{z}_t = (1 - \theta_1 B - \dots - \theta_q B^q) a_t$$

$$\phi_p(B) \tilde{z}_t = \theta_q(B) a_t$$

Donde las raíces de $\phi_p(B) = 0$ están fuera del círculo unidad para que el proceso sea estacionario y el proceso será invertible si lo están las de $\theta_q(B) = 0$. Además, se debe cumplir que $\phi_p(B)$ y $\theta_q(B)$ no tengan raíces comunes. Si no se cumple esta condición resulta que $z_t = a_t$ y el proceso sería ruido blanco.

Los procesos ARIMA son una combinación de la diferenciación con los procesos autorregresivos y los de media móvil estudiados anteriormente.

La diferenciación es un método realizado para combatir el comportamiento no estacionario de algunas series temporales. Consiste en suponer que la tendencia evoluciona con el tiempo y si restamos cada valor de la serie con el valor anterior se librerá a esta de la tendencia. (Box & Jenkins, 1976)

$$B(z_t) = z_{t-1}$$

$$\nabla z_t = z_t - z_{t-1} = z_t - B(z_t) = (1 - B)z_t$$

$$\therefore \nabla^d z_t = (1 - B)^d z_t$$

Donde d representa la cantidad de diferencias aplicadas a la serie.

La operación inversa a la diferenciación se conoce como integración ($z_t = z_{t-1} + \nabla z_t$) y gracias a ella revisen su nombre los procesos ARIMA.

Un proceso ARIMA (p,d,q) se expresa:

$$\nabla^d \tilde{z}_t = \phi_1 \nabla^d \tilde{z}_{t-1} + \dots + \phi_p \nabla^d \tilde{z}_{t-p} + a_t - \theta_1 a_{t-1} - \dots - \theta_q a_{t-q}$$

Utilizando el operador de retardos, la expresión se simplifica a:

$$\phi_p(B) \nabla^d \tilde{z}_t = \theta_q(B) a_t$$

En la ecuación, p especifica los periodos anteriores relevantes para la componente autorregresiva, la letra q denota las innovaciones anteriores consideradas para estimar el modelo y del grado de diferenciación aplicada a la serie original.

Las condiciones de estacionariedad e invertibilidad son las mismas utilizadas para los procesos ARMA.

2.2.4.1 Identificación de modelos ARIMA. La identificación de los órdenes p , q y d del modelo se determina a partir de la observación de las funciones de autocorrelación simple y parcial con las funciones teóricas.

El decrecimiento de las barras en el correlograma (coeficientes de función) debe ser exponencial. En ocasiones existe alternancia de signos o forma sinusoidal como resultado de los signos de los parámetros o la existencia de raíces complejas en la ecuación característica del proceso. Si el decrecimiento apreciado es lineal, se presume que la serie es integrada. En este caso se debe diferenciar y volver a realizar la inspección. El número de diferencias indicará el grado de diferenciación del modelo (d).

Comenzar con la identificación de modelos simples AR o MA. Si la FAS presenta el decrecimiento exponencial podemos presumir un modelo AR y la FAP nos indicará el orden (p) dependiendo de la cantidad de coeficientes que superen las líneas de significatividad (Figura 10). Si, por el contrario, el decrecimiento es detectado en la FAP se presume un modelo MA cuyo orden (q) dependerá del número de coeficientes que superen las líneas de significatividad en la FAS (Figura 11).

Es posible que el correlograma no se muestre claro, lo que indica un modelo mixto compuesto por una parte AR y otra MA. En la Figura 12 se muestra las funciones teóricas para procesos mixtos ARMA (1,1). (Peña, 2005)

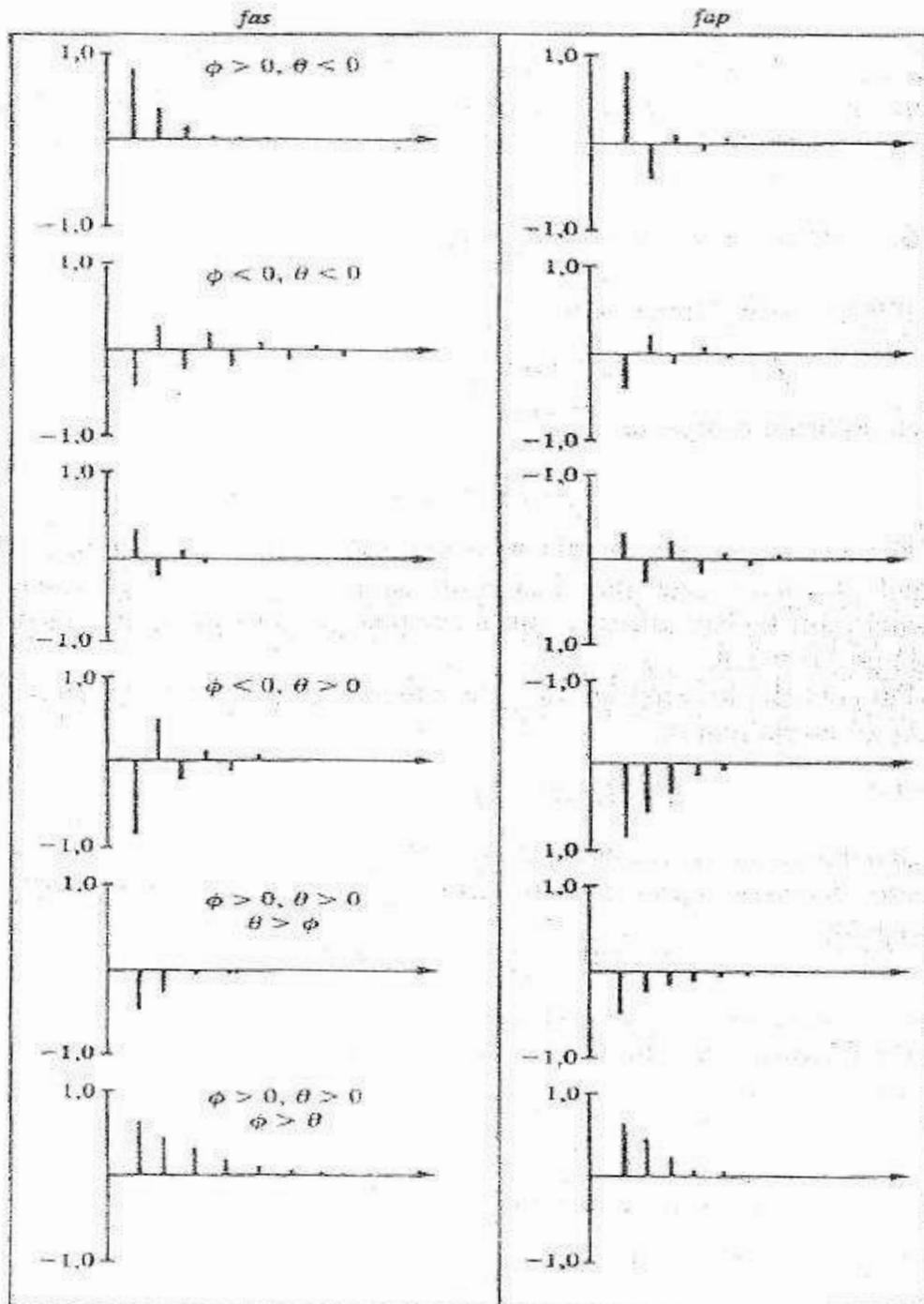


Figura 12. Funciones de autocorrelación simple y parcial de procesos ARMA (1,1).

Fuente: Peña (2005)

2.2.4.2 ARIMA estacional. En una serie estacional, el valor esperado no es constante, sino que varía con una pauta cíclica:

$$E(z_t) = E(z_{t+k})$$

Donde la serie tiene una estacionalidad de periodo k.

Se puede convertir series no estacionarias, a través de diferencias regulares, en series estacionarias. De la misma manera se puede eliminar la estacionalidad mediante diferencias estacionales. La componente estacional de una serie de tiempo también puede explicarse mediante procesos autorregresivos y de media móvil.

Definiremos entonces un proceso ARIMA estacional o ARIMA (P,Q,D)_s como:

$$\Phi_P(B)\nabla_S^D \tilde{z}_t = \Theta_Q(B)a_t$$

Donde D es el grado de diferenciación estacional, P el orden de la componente autorregresiva estacional y Q el orden de la componente de media móvil estacional.

Los modelos ARIMA se pueden completar incorporando además de la dependencia regular la dependencia estacional, asociada para observaciones separadas k periodos.

El modelo se construye incorporando la dependencia estacional de forma multiplicativa ARIMA (p,d,q)(P,Q,D)_s y se expresa:

$$\phi_p(B)\Phi_P(B)\nabla^d\nabla_S^D \tilde{z}_t = \theta_q(B)\Theta_Q(B)a_t$$

Donde:

p es el orden del componente autorregresivo regular

q es el orden del componente de media móvil regular

d es el grado de diferenciación regular de la serie

P es el orden del componente autorregresivo estacional

Q es el orden del componente de media móvil estacional

D es el grado de diferenciación estacional de la serie (Peña, 2005)

Capítulo 3

Propuesta de modelos de predicción

La predicción del caudal del río Chira es fundamental para permitir la operación óptima del embalse de Poechos y de esta manera asegurar la disponibilidad del recurso hídrico en la región. En el desarrollo de este capítulo se propone modelos de predicción univariante que permitan satisfacer esta necesidad.

3.1 Datos

En el desarrollo de esta investigación se utiliza los registros de caudal promedio diario recolectados en la estación Ardilla.

La estación Ardilla se encuentra ubicada en el distrito de Lancones, en la provincia de Sullana y mide el caudal aguas arriba del reservorio de Poechos. Permite la evaluación de los recursos hídricos provenientes de la corriente del río Chira.

La estación pertenece a la red hidrográfica del río Chira y es controlada por el Proyecto Especial Chira - Piura (PECHP) desde el 01 de enero del año 1976. Facilita los registros hidrográficos necesarios para la operación del embalse. (Aguilar Villena, 2016)

El análisis se realiza en el periodo del 01 de enero de 1951 al 14 de agosto de 2017.

En la Figura 13 se describe la situación hidrológica del río Chira a partir de los caudales promedio diarios recolectados en la estación Ardilla. El caudal medio diario máximo fue alcanzado el 07/04/1998, con un registro de 3620.8 m³/s. En la Figura 14 se muestra la distribución de los caudales medios diario, observando que la mayoría de los valores se distribuyen entre 0 y 500 m³/s.

En la Figura 15 y Figura 16 se presenta la situación hidrológica expresada en caudales promedio semanales. El máximo caudal medio semanal fue de 2786.6 m³/s, registrado en abril de 1998.

Es importante notar que el comportamiento en la serie para cada trimestre del año varía, producto de las estaciones como se aprecia en la Figura 17.

En la Figura 18 se aprecia el comportamiento del caudal durante el periodo de estudio. El máximo caudal medio anual se alcanzó en el año 1998, con un valor de 556.7 m³/s. La serie

temporal del caudal presenta gran variabilidad, que se evidencia con mayor detalle en la Figura 19.

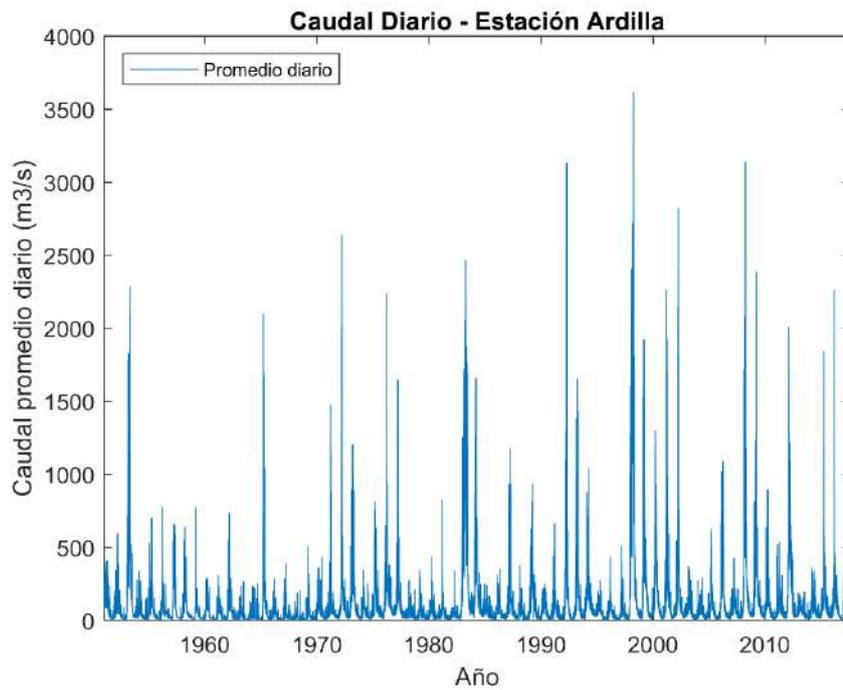


Figura 13. Caudal promedio diario - Estación Ardilla.

Fuente: Elaboración propia.

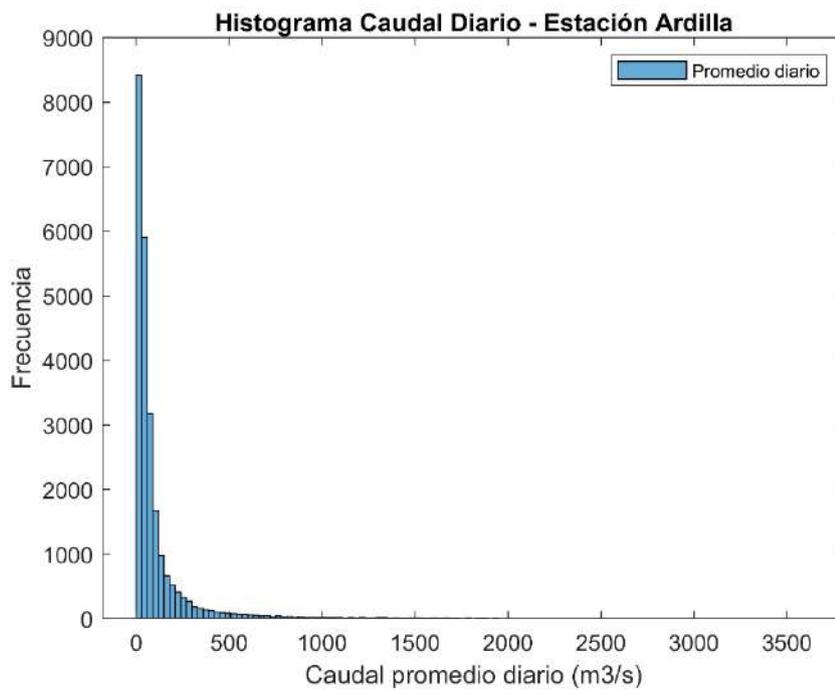


Figura 14. Histograma caudal promedio diario -Estación Ardilla.

Fuente: Elaboración propia.

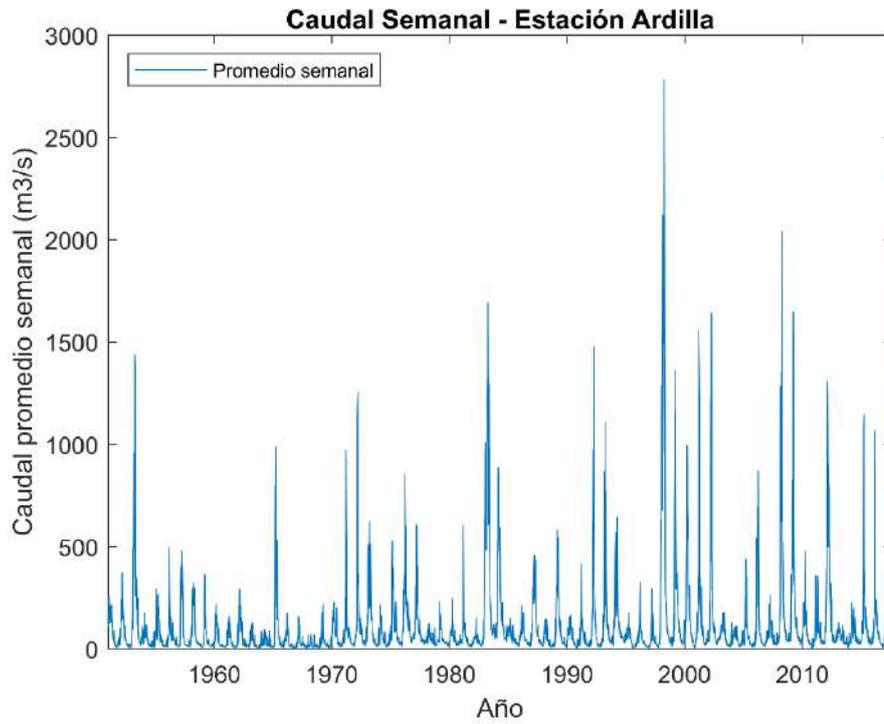


Figura 15. Caudal promedio semanal - Estación Ardilla.

Fuente: Elaboración propia.

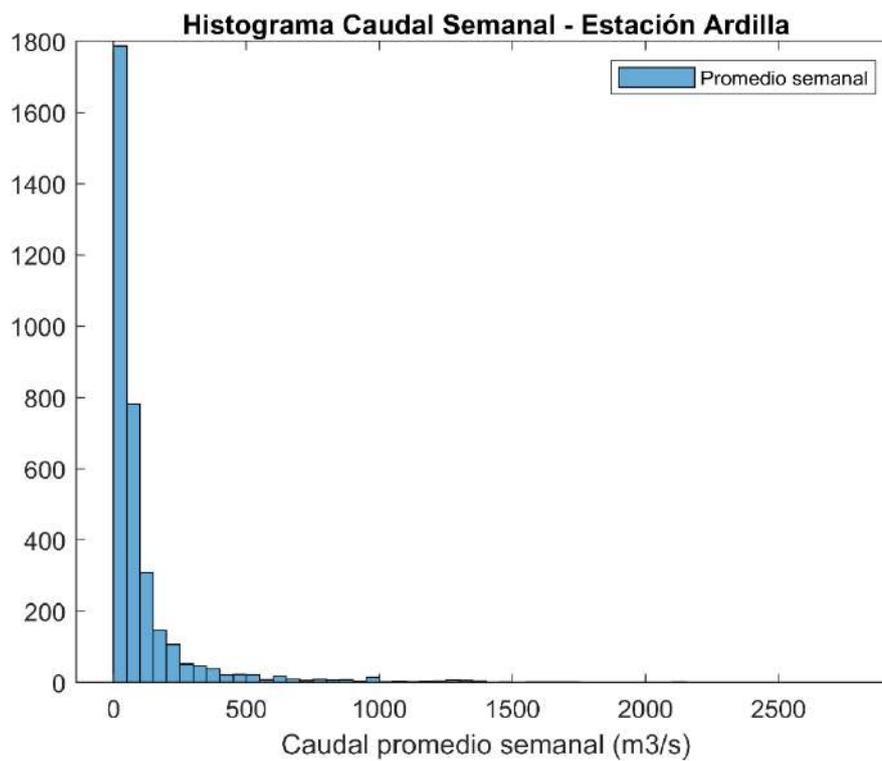


Figura 16. Histograma caudal promedio semanal - Estación Ardilla.

Fuente: Elaboración propia.

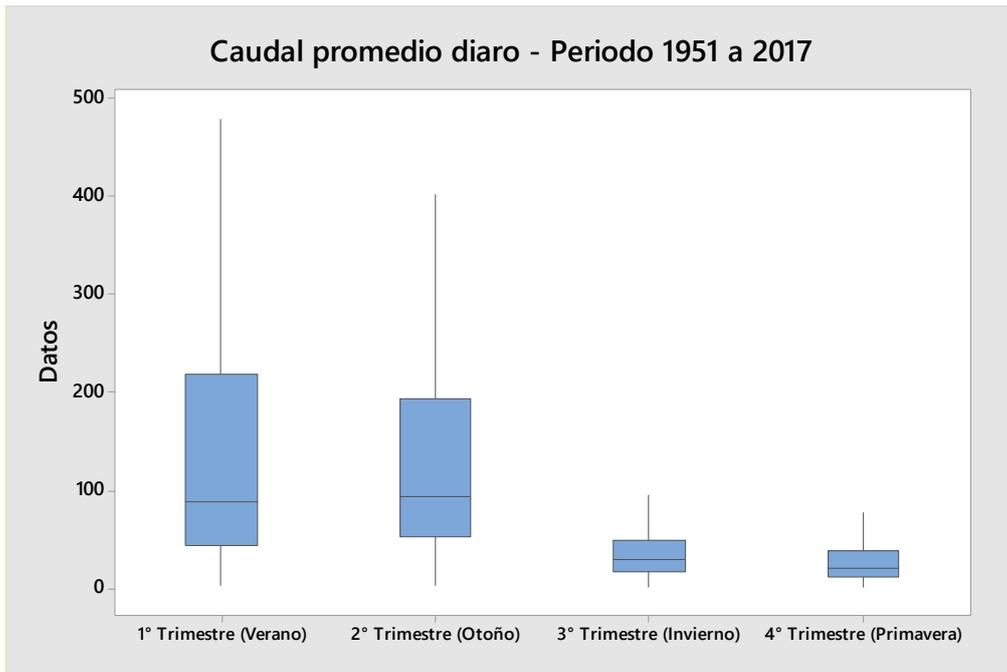


Figura 17. Diagrama de caja del caudal promedio diario por trimestre - Estación Ardilla.

Fuente: Elaboración propia.

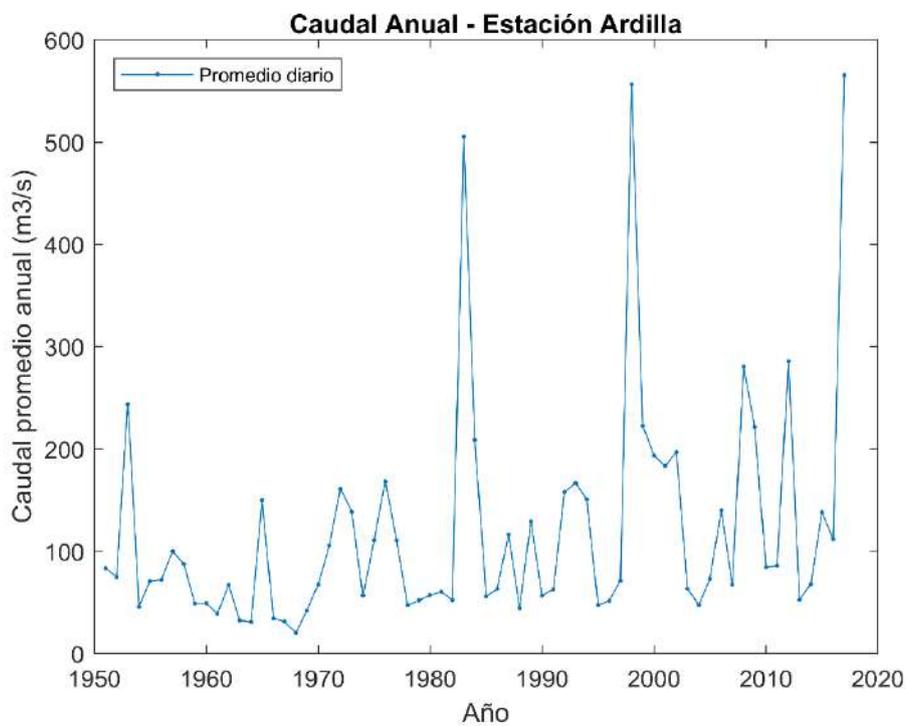


Figura 18. Caudal promedio anual - Estación Ardilla.

Fuente: Elaboración propia.

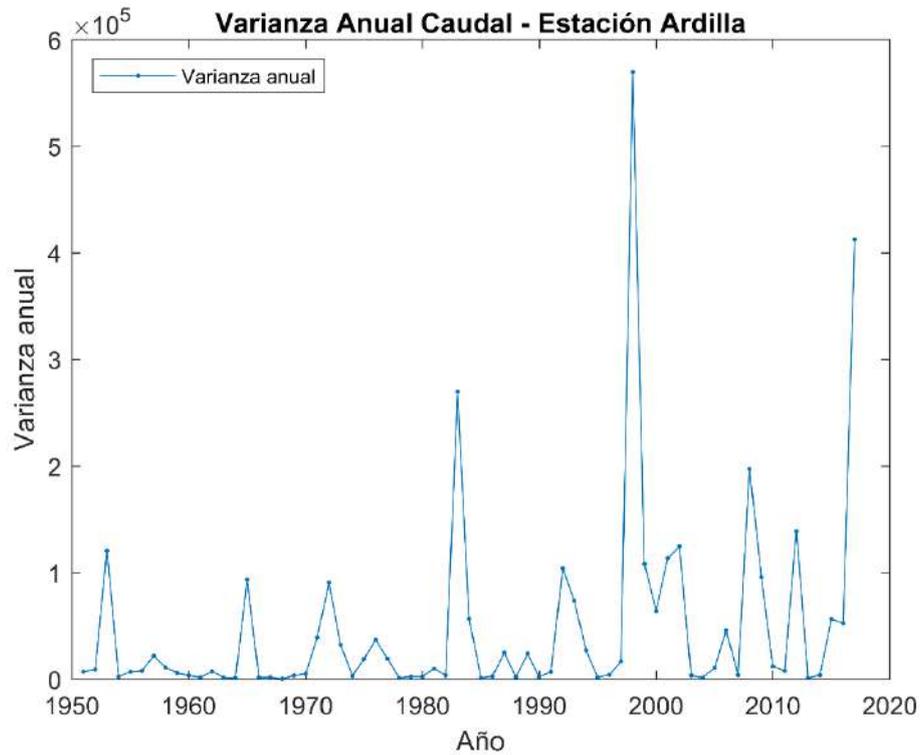


Figura 19. Varianza anual del caudal - Estación Ardilla.

Fuente: Elaboración propia.

El comportamiento heterocedástico de la serie temporal de caudales dificulta la modelización estadística, que requiere de una varianza constante a lo largo del tiempo. Por tanto, es necesario realizar una transformación a la serie temporal original.

Para conseguir una serie estacionaria se utiliza la transformación logarítmica, la más usada de la familia de transformaciones Box - Cox.

$$z'_t = \ln(z_t)$$

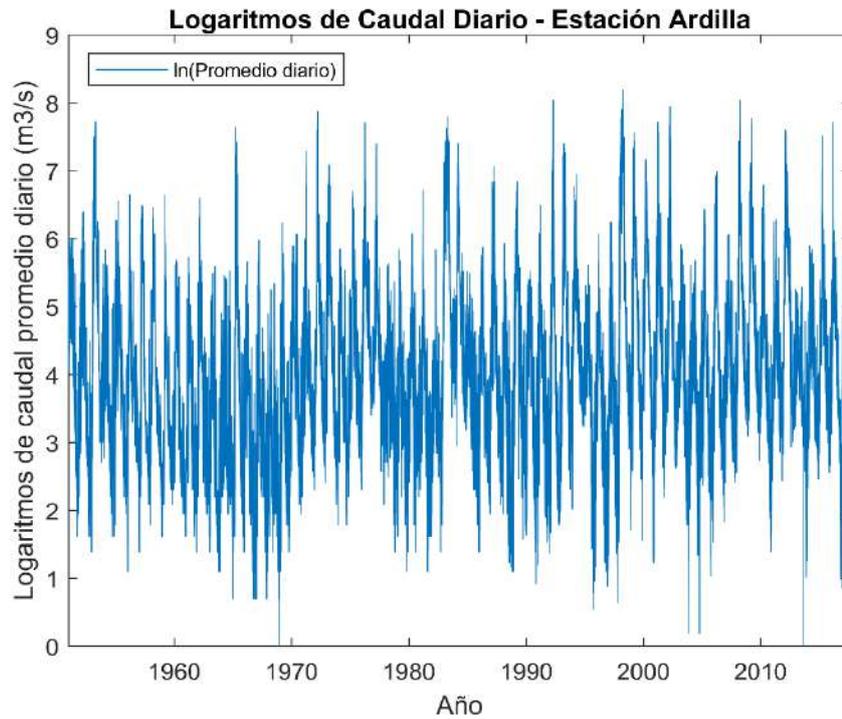


Figura 20. Transformación Logarítmica del caudal promedio diario.

Fuente: Elaboración propia.

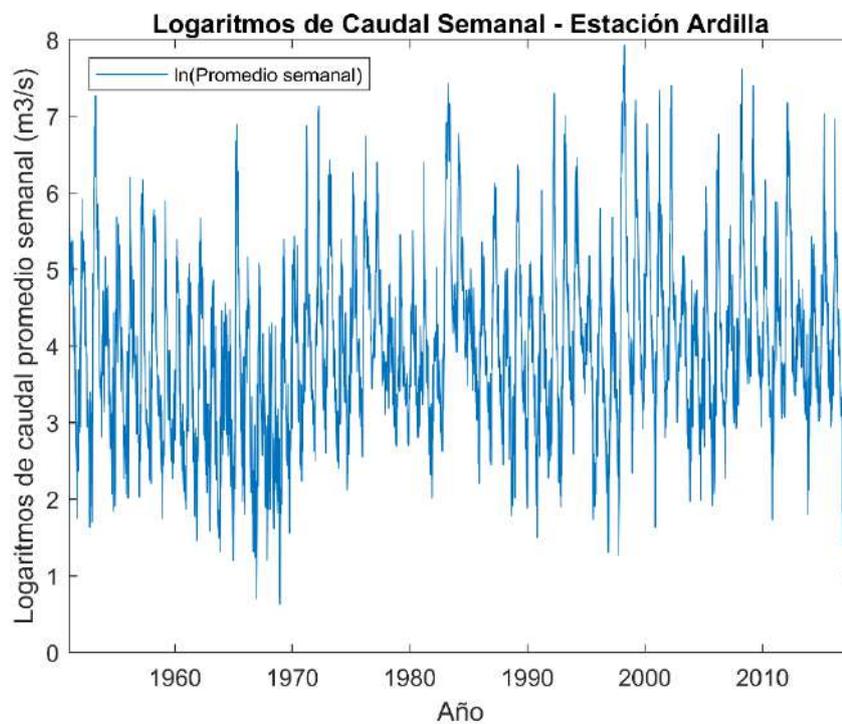


Figura 21. Transformación Logarítmica del caudal promedio semanal.

Fuente: Elaboración propia.

3.2 Modelización diaria

La presente investigación evalúa el periodo comprendido entre el 01 de enero de 1951 hasta el 14 de agosto de 2017.

Los modelos realizan predicciones en un horizonte de 10 días ($h = 10$). Para inicializar los modelos se utiliza un periodo de muestra que comprende el caudal medio diario entre el 01 de enero de 1951 hasta el 31 de diciembre de 1979. Las primeras predicciones corresponden al caudal medio diario del 01 de enero de 1980 al 10 de enero de 1980. El ejercicio se repite incorporando en cada iteración un día al periodo de muestra y prediciendo los 10 días siguientes hasta agotar los datos del periodo en evaluación (13731 iteraciones).

Tabla 8. Procedimiento para generación de predicciones diarias.

	Fin del periodo de muestra	Periodo de muestra	Horizonte de predicción
Primera iteración ($i = 1$)	Fecha ₁ = 31/12/1979	Caudal (01/01/1951 hasta 31/12/1979)	Caudal (01/01/1980 hasta 10/01/1980)
Siguientes iteraciones (i)	Fecha _i = Fecha _{i-1} + 1	Caudal (01/01/1951 hasta Fecha _i)	Caudal (Fecha _i + 1 hasta Fecha _i + 10)

Fuente: Elaboración propia.

3.2.1 Métodos de suavizamiento exponencial

Se evalúa los diferentes métodos de suavización exponencial considerados en el apartado 2.1 incluyendo las versiones con tendencia amortiguada de los mismos.

3.2.1.1 Holt. El método de suavizamiento exponencial doble considera la variación en el nivel y la tendencia de la serie para efectuar las predicciones. Se emplean dos parámetros:

- Alfa (α): Parámetro de la componente de nivel
- Beta (β): Parámetro de la componente de tendencia.

Para el desarrollo del presente modelo, en adelante Modelo Holt, se utiliza una función de optimización, que permite encontrar un parámetro óptimo que minimice el error cuadrático medio de las predicciones. Además, se hace uso de la transformación logarítmica, propia de la familia Box - Cox, con el objetivo de disminuir el impacto de la variabilidad propia de la serie.

En cada ejercicio de predicción se inicializan los parámetros del modelo con los parámetros óptimos de la iteración anterior. De esta manera, los parámetros promedio obtenidos para este modelo, siguiendo el procedimiento de evaluación descrito en el apartado 3.2 son:

Tabla 9. Parámetros promedio estimados - Modelo Holt diario.

Modelo	Parámetros	
	α	β
Holt	0.8144	0.0013

Fuente: Elaboración propia.

El modelo queda expresado como:

$$\hat{z}_{t+h|t} = L_t + hT_t$$

Donde:

$$L_t = (0.8144)z_t + (1 - 0.8144)(L_{t-1} + T_{t-1})$$

$$T_t = (0.0013)(L_t - L_{t-1}) + (1 - 0.0013)T_{t-1}$$

$$h = 1, 2, \dots, 10$$

El error cuadrático medio obtenido con este modelo para cada horizonte de predicción se encuentra en la Figura 22. Se puede resaltar que el error cuadrático medio incrementa a media que aumenta el horizonte de predicción. El valor mínimo de ECMP alcanzado por el Modelo Holt es 1.0901×10^4 mientras que el valor máximo es 4.1981×10^4 .

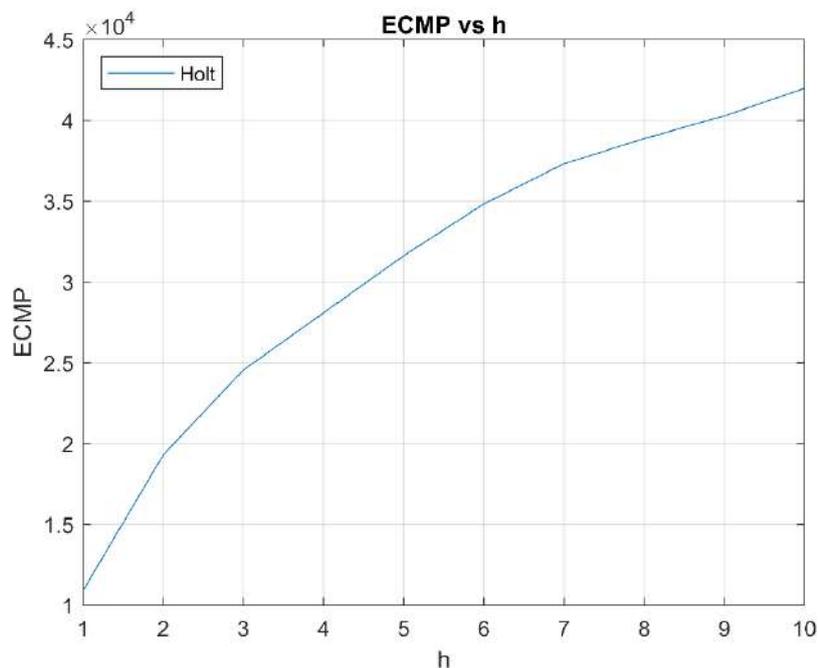


Figura 22. ECMP vs h - Modelo Holt diario.

Fuente: Elaboración propia.

3.2.1.2 Holt con tendencia amortiguada. Este método posee los mismos parámetros mencionados en el método de suavización doble e incluye un nuevo parámetro que amortigua la tendencia en observaciones futuras, al que se denomina como ϕ .

Para el desarrollo de este modelo, en adelante Modelo Holt Damped, se utiliza la función de optimización para la búsqueda de los parámetros óptimos. Del mismo modo, se emplea también la transformación logarítmica de la serie para reducir la variabilidad. En cada iteración, los parámetros se inicializan con los parámetros óptimos de la iteración anterior.

Tabla 10. Parámetros promedio estimados - Modelo Holt Damped diario.

Modelo	Parámetros		
	α	β	ϕ
Holt Damped	0.7701	0.3448	2.9103×10^{-5}

Fuente: Elaboración propia.

Como resultado, el modelo queda expresado como:

$$\hat{z}_{t+h|t} = L_t + ((2.903 \times 10^{-5}) + \dots + (2.903 \times 10^{-5})^h)T_t$$

Donde:

$$L_t = (0.7701)z_t + (1 - 0.7701)(L_{t-1} + (2.903 \times 10^{-5})T_{t-1})$$

$$T_t = (0.3448)(L_t - L_{t-1}) + (1 - 0.3448)(2.903 \times 10^{-5})T_{t-1}$$

$$h = 1, 2, \dots, 10$$

El error cuadrático medio resultante para cada horizonte de predicción está representado en la Figura 23. El valor máximo del ECMP se alcanza en el último horizonte de predicción y tiene un valor de 4.1543×10^4 , mientras que el valor mínimo se produce en el primer horizonte de predicción con un valor de 1.0978×10^4 .

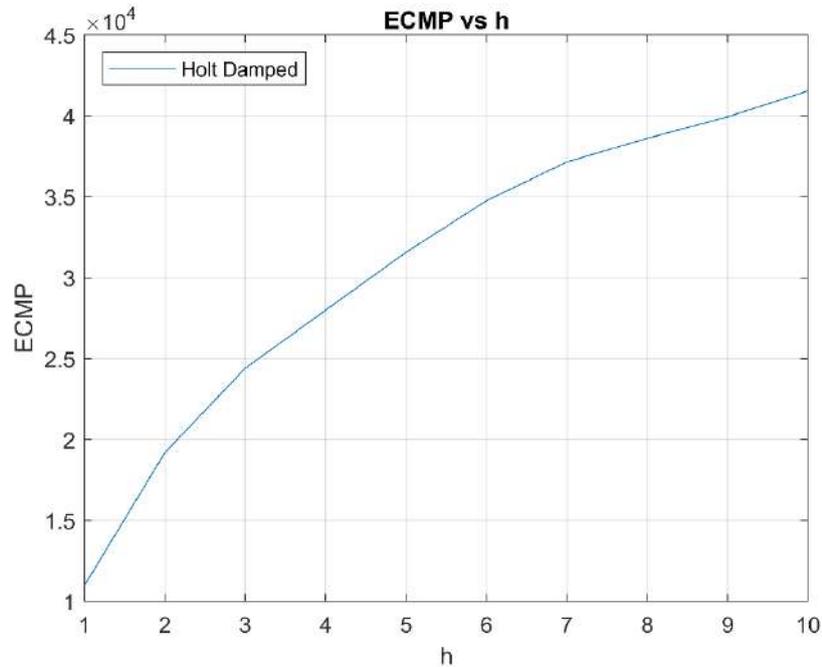


Figura 23. ECMP vs h - Modelo Holt Damped diario.

Fuente: Elaboración propia.

3.2.1.3 Holt-Winters aditivo. El método de suavizamiento exponencial triple es una extensión del método de suavizamiento doble y considera tres exponentes suavizantes, los dos primeros correspondientes al nivel y la tendencia, acompañados de un tercero que corresponde a la estacionalidad. Los parámetros considerados para este método son:

- Alfa (α): Parámetro de la componente de nivel.
- Beta (β): Parámetro de la componente de tendencia.
- Gamma (γ): Parámetro de la componente estacional.

En el desarrollo de este modelo, en adelante denominado Modelo Holt-Winters, se hace uso de la misma función de optimización utilizada en el apartado 3.2.1.1 que permite encontrar el parámetro óptimo el cual reduce el error cuadrático medio de las predicciones. La serie está compuesta de caudales promedio diarios y para este modelo se presupone que existe variación periódica y predecible cada año, considerando una estacionalidad de 365 días ($s = 365$). Además, se realiza la transformación logarítmica de la serie, para reducir el impacto de la variabilidad de la misma. En cada ejercicio de predicción, los parámetros se inicializan utilizando el parámetro óptimo de la iteración anterior.

En consecuencia, los parámetros obtenidos para este modelo son:

Tabla 11. Parámetros promedio estimados -
Modelo Holt-Winters diario.

Modelo	Parámetros		
	α	β	γ
Holt-Winters	0.6782	2.8565×10^{-11}	0.1014

Fuente: Elaboración propia.

El modelo se expresa como:

$$\hat{z}_{t+h|t} = L_t + hT_t + S_{t+h-365(k+1)}$$

Donde:

$$L_t = (0.6782)(z_t - S_{t-365}) + (1 - 0.6782)(L_{t-1} + T_{t-1})$$

$$T_t = (2.8565 \times 10^{-11})(L_t - L_{t-1}) + (1 - 2.8565 \times 10^{-11})T_{t-1}$$

$$S_t = (0.1014)(z_t - L_{t-1} - T_{t-1}) + (1 - 0.1014)S_{t-365}$$

$$h = 1, 2, \dots, 10$$

$$k = \frac{(h-1)}{365}$$

En la Figura 24 se puede comprobar el error cuadrático medio obtenido para cada horizonte de predicción. El error cuadrático medio para cada horizonte de predicción obtenido con el modelo Holt-Winters se muestra como un trazo un trazo cóncavo. El mínimo valor de ECMP se obtuvo en el primer horizonte de predicción con un valor de 1.2247×10^4 y el valor máximo fue 4.3197×10^4 en el último horizonte de predicción.

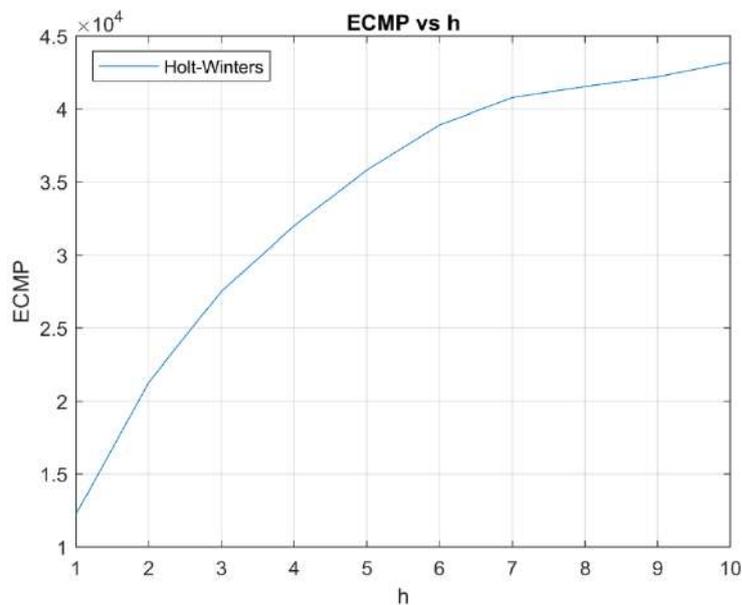


Figura 24. ECMP vs h - Modelo Holt-Winters diario.

Fuente: Elaboración propia.

3.2.1.4 Holt-Winters aditivo con tendencia amortiguada. Este método incluye los parámetros mencionados en el apartado 3.2.1.3 e incorpora el parámetro ϕ que amortigua la tendencia en observaciones futuras.

Este modelo, en adelante Modelo Holt-Winters Damped, emplea la función de optimización para encontrar parámetros óptimos. Así mismo, se utiliza la serie transformada en logaritmos para reducir la variabilidad. Se supone cambios regulares de la serie en periodos de 365 días ($s = 365$). En cada iteración, los parámetros se inicializan con el parámetro óptimo de la iteración anterior.

Tabla 12. Parámetros promedio estimados - Modelo Holt-Winters Damped diario.

Modelo	Parámetros			
	α	β	γ	ϕ
Holt-Winters Damped	0.6777	0.3313	0.1014	2.9136×10^{-5}

Fuente: Elaboración propia.

Por tanto, el modelo resultante se expresa como:

$$\hat{z}_{t+h|t} = L_t + ((2.9136 \times 10^{-5}) + \dots + (2.9136 \times 10^{-5})^h)T_t + S_{t+h-365(k+1)}$$

Donde:

$$L_t = (0.6777)(z_t - S_{t-365}) + (1 - 0.6777)(L_{t-1} + (2.9136 \times 10^{-5})T_{t-1})$$

$$T_t = (0.3313)(L_t - L_{t-1}) + (1 - 0.3313)(2.9136 \times 10^{-5})T_{t-1}$$

$$S_t = (0.1014)(z_t - L_{t-1} - (2.9136 \times 10^{-5})T_{t-1}) + (1 - 0.1014)S_{t-365}$$

$$h = 1, 2, \dots, 10$$

$$k = \frac{(h-1)}{365}$$

El error cuadrático medio para cada horizonte de predicción se observa en la Figura 25. La curva del error cuadrático medio para cada horizonte de predicción de este modelo tiene un comportamiento similar al Modelo Holt-Winters, con un valor máximo de ECMP de 4.3394×10^4 y un valor mínimo de 1.2253×10^4 .

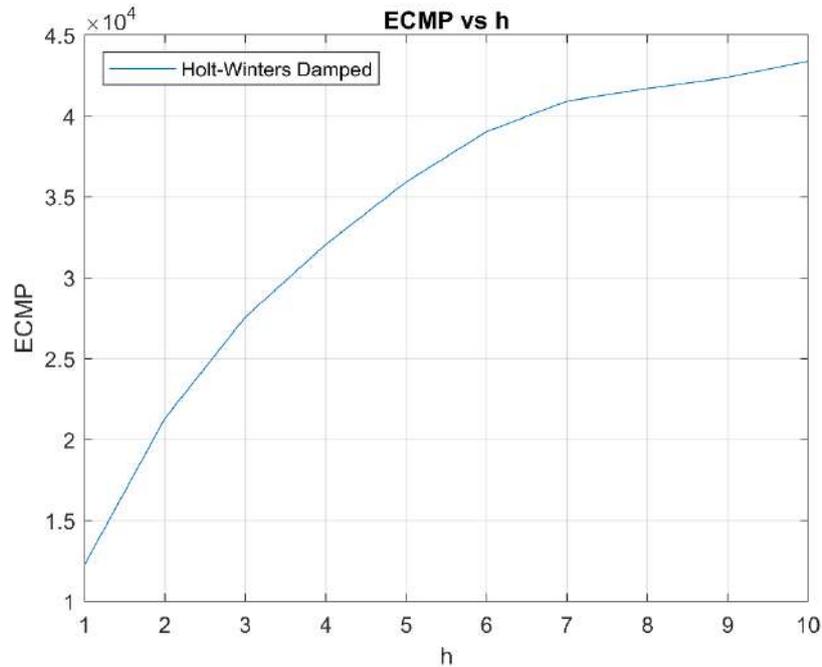


Figura 25. ECMP vs h - Modelo Holt-Winters Damped diario.

Fuente: Elaboración propia.

3.2.1.5 Holt-Winters multiplicativo. El método de suavizamiento exponencial triple en su variación multiplicativo considera los mismos componentes vistos en el apartado 3.2.1.3. Este método es usado en series que presentan una variación estacional que cambia proporcionalmente al nivel de la serie. (Hyndman & Athanasopoulos, 2018)

El modelo, al que en adelante se hará referencia como Modelo Holt-Winters Multiplicativo, se vale de una función de optimización para hallar parámetros óptimos, asegurando la disminución del error cuadrático medio. Se supone que la serie de caudal medio diario experimenta fluctuaciones regulares en periodos anuales ($s = 365$). Debido a la naturaleza del método, que trata las variaciones estacionales que crecen proporcional nivel, no se utiliza las transformaciones de la familia Box-Cox. En cada ejercicio de predicción, los parámetros del modelo se inicializan con el parámetro óptimo de la iteración anterior.

Los parámetros promedio obtenidos para este modelo se ubican en la Tabla 13.

Tabla 13. Parámetros promedio estimados - Modelo Holt-Winters Multiplicativo diario.

Modelo	Parámetros		
	α	β	γ
Holt-Winters Multiplicativo	0.9844	0.0294	0.4631

Fuente: Elaboración propia.

El modelo queda expresado como:

$$\hat{z}_{t+h|t} = (L_t + hT_t)S_{t+h-365(k+1)}$$

Donde:

$$L_t = (0.9844)\left(\frac{z_t}{S_{t-365}}\right) + (1 - 0.9844)(L_{t-1} + T_{t-1})$$

$$T_t = (0.0294)(L_t - L_{t-1}) + (1 - 0.0294)T_{t-1}$$

$$S_t = (0.4631)\left(\frac{z_t}{L_{t-1} + T_{t-1}}\right) + (1 - 0.4631)S_{t-365}$$

$$h = 1, 2, \dots, 10$$

$$k = \frac{(k+1)}{365}$$

En la Figura 26 se observa el error cuadrático medio para cada horizonte de predicción obtenido con este modelo. En la gráfica se puede distinguir el comportamiento errático del error cuadrático medio. Debido a las características propias de la serie de caudal medio diario, el modelo Holt-Winters Multiplicativo presenta errores muy altos y se reduce su capacidad de predicción.

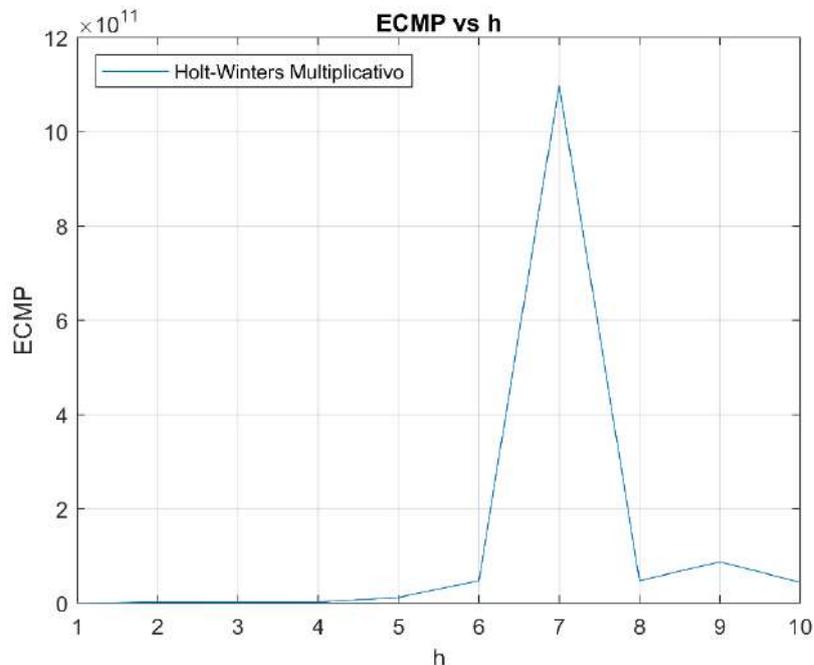


Figura 26. ECMP vs h - Modelo Holt-Winters Multiplicativo diario.

Fuente: Elaboración propia.

3.2.1.6 Holt-Winters multiplicativo con tendencia amortiguada. Este método incluye los parámetros considerados en el apartado 3.2.1.5 e incorpora el parámetro ϕ que amortigua la tendencia en observaciones futuras.

El modelo, en adelante denominado Modelo Holt-Winters Multiplicativo Damped, utiliza una función de optimización para localizar los parámetros óptimos. Se supone una componente estacional de la serie de 365 días ($s = 365$). En cada iteración, los parámetros del modelo se inicializan con los parámetros óptimos de la iteración anterior.

Tabla 14. Parámetros promedio estimados - Modelo Holt-Winters Multiplicativo Damped diario.

Modelo	Parámetros			
	α	β	γ	ϕ
Holt-Winters Multiplicativo Damped	0.4262	0.373	0.9256	0.000846

Fuente: Elaboración propia.

El modelo resultante se expresa como:

$$\hat{Z}_{t+h|t} = (L_t + ((0.000846) + \dots + (0.000846)^h)T_t)S_{t+h-365(k+1)}$$

Donde:

$$L_t = (0.4262)\left(\frac{Z_t}{S_{t-365}}\right) + (1 - 0.4262)(L_{t-1} + (0.000846)T_{t-1})$$

$$T_t = (0.373)(L_t - L_{t-1}) + (1 - 0.373)(0.000846)T_{t-1}$$

$$S_t = (0.9256)\left(\frac{Z_t}{L_{t-1} + (0.000846)T_{t-1}}\right) + (1 - 0.9256)S_{t-365}$$

$$h = 1, 2, \dots, 10$$

$$k = \frac{(k+1)}{365}$$

En la Figura 27 se aprecia el error cuadrático medio para cada horizonte de predicción obtenido con este modelo. Se observa que el comportamiento del error cuadrático medio respecto al horizonte de predicción es errático, aunque en menor medida que el modelo Holt-Winters Multiplicativo. El modelo no reacciona de manera favorable a las características propias de la serie de caudal medio diario, razón por la cual se reduce la capacidad de predicción.

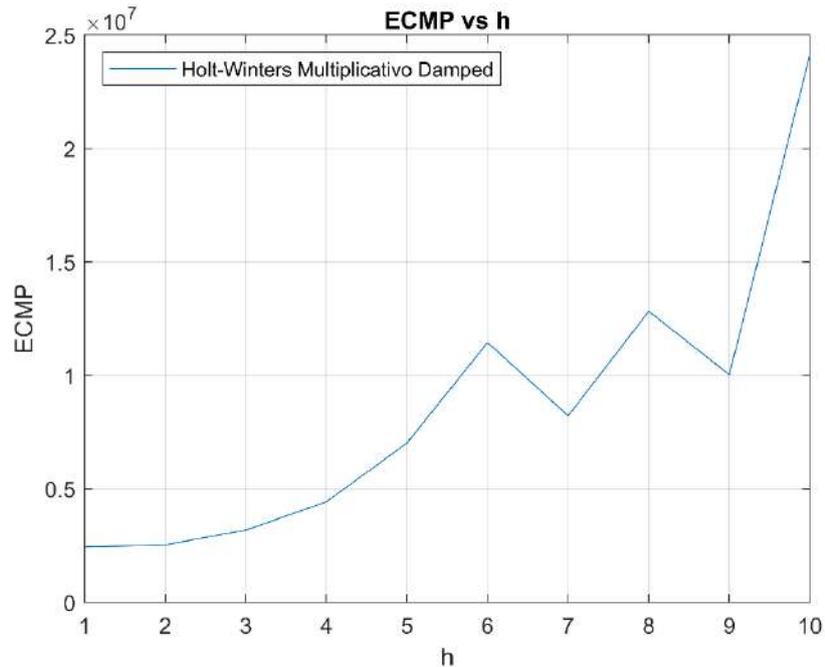


Figura 27. ECMP vs h - Modelo Holt-Winters Multiplicativo Damped diario.

Fuente: Elaboración propia.

3.2.2 Métodos ARIMA

3.2.2.1 Identificación. La serie consta de 24333 observaciones del caudal medio diario, que siguen la distribución que se aprecia en la Figura 13. A causa de la variabilidad presente en la serie, se requiere aplicar una transformación de la familia Box-Cox que revierta el comportamiento heterocedástico del caudal. La transformación utilizada es la transformación logarítmica, resultando en la nueva serie descrita en la Figura 20, a la que se denomina LC. Por último, la modelización estadística requiere que la serie sea estacionaria. La serie LC tiene una media que no es constante, por lo que se toma una diferencia regular de la serie. Esto consiste en restar al dato LC_t con su inmediato anterior LC_{t-1} . Esta nueva serie, denominada DLC, es estacionaria como se muestra en la Figura 28.

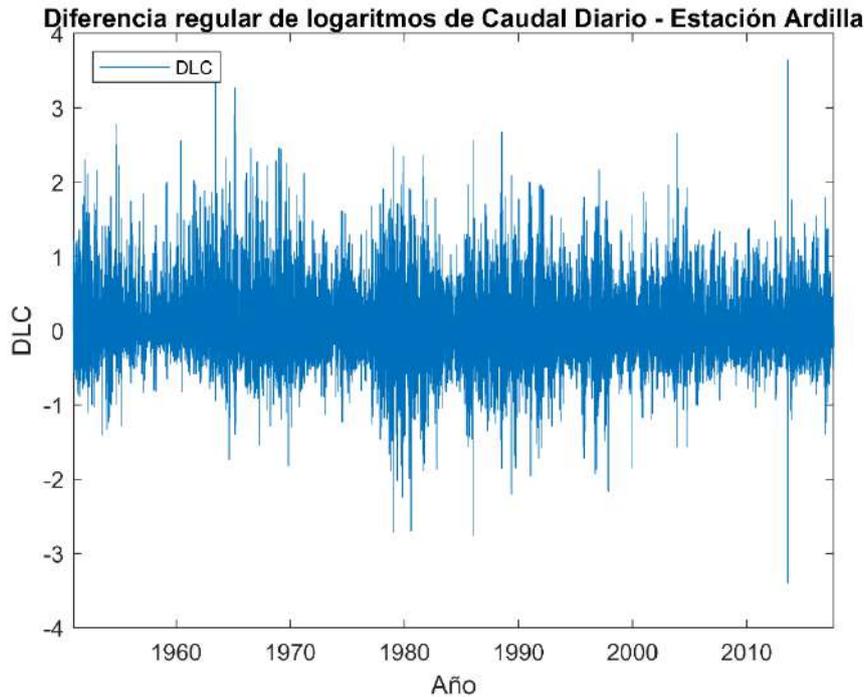


Figura 28. Diferencia regular de logaritmos del caudal promedio diario.

Fuente: Elaboración propia.

Para definir el modelo ARIMA óptimo para predecir la serie, es necesario comprender la correlación existente entre los diferentes instantes en la serie temporal, con un desfase de uno o más periodos. El análisis de la FAS y la FAP permite conocer la correlación entre las observaciones en el instante t hasta el instante $t-k$.

En la Figura 29 y Figura 30 se cumple que los coeficientes de la función decrecen de manera exponencial. Tanto la FAS como la FAP muestran valores fuera de las bandas de significatividad en la componente regular, lo que sugiere un modelo ARMA. No se identifica valores significativos en los coeficientes estacionales y se infiere que la estacionalidad es lo suficientemente débil para despreciarla.

Comparando la FAS y FAP obtenidas de la serie (Figura 31 y Figura 32) contra las gráficas FAS y FAP teóricas, se percibe una gran similitud con el quinto ejemplo de la Figura 12, lo que sugiere un modelo ARMA(1,1).

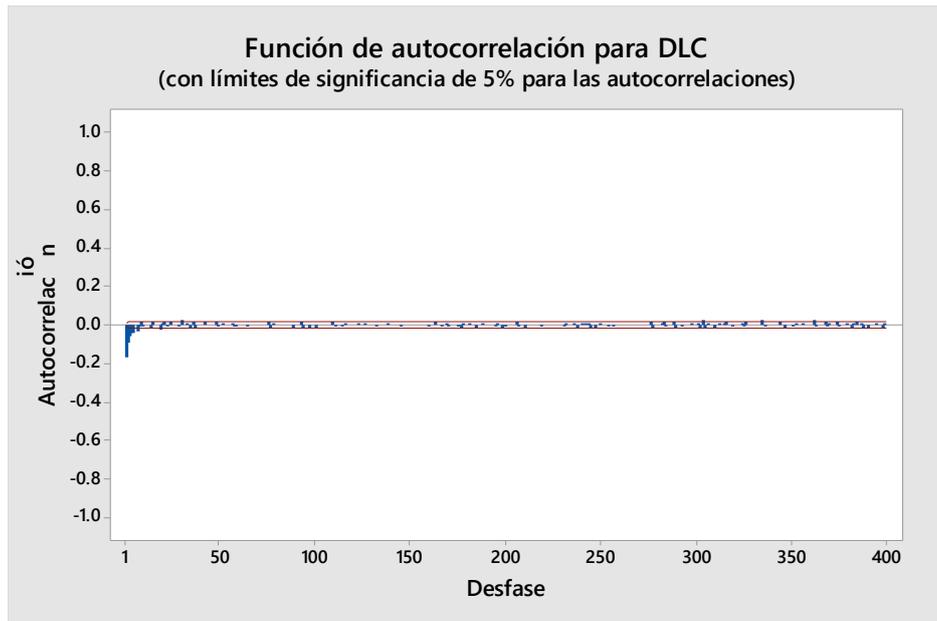


Figura 29. Figura 29. FAS para la serie DLC.

Fuente: Elaboración propia.

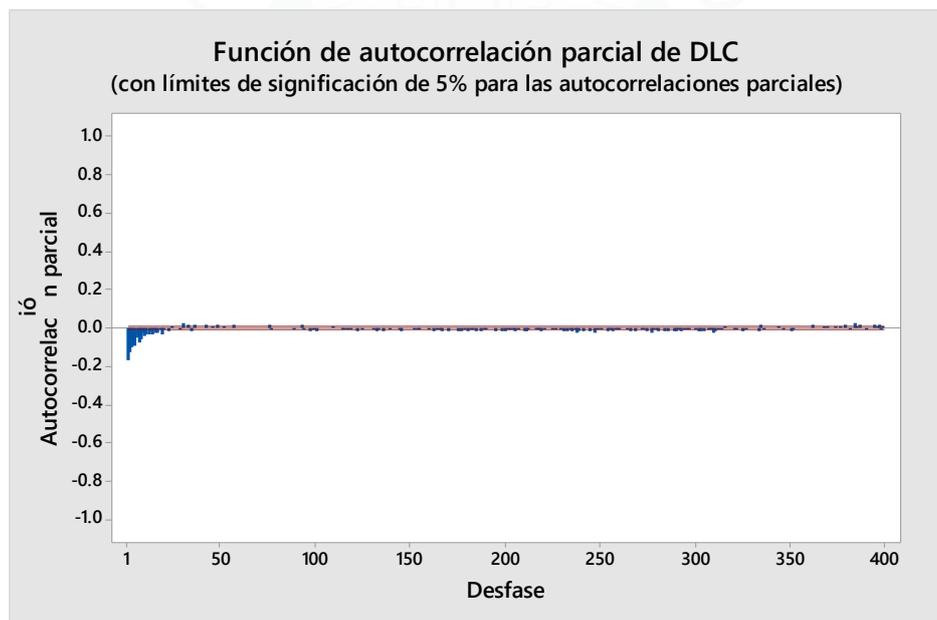


Figura 30. FAP para la serie DLC.

Fuente: Elaboración propia.

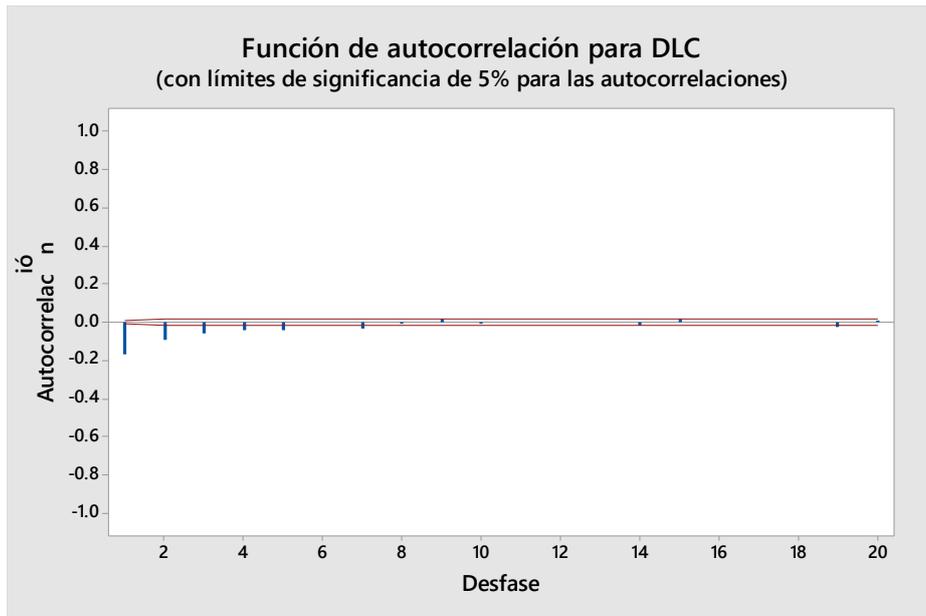


Figura 31. FAS para la serie DLC (Aumento).

Fuente: Elaboración propia.

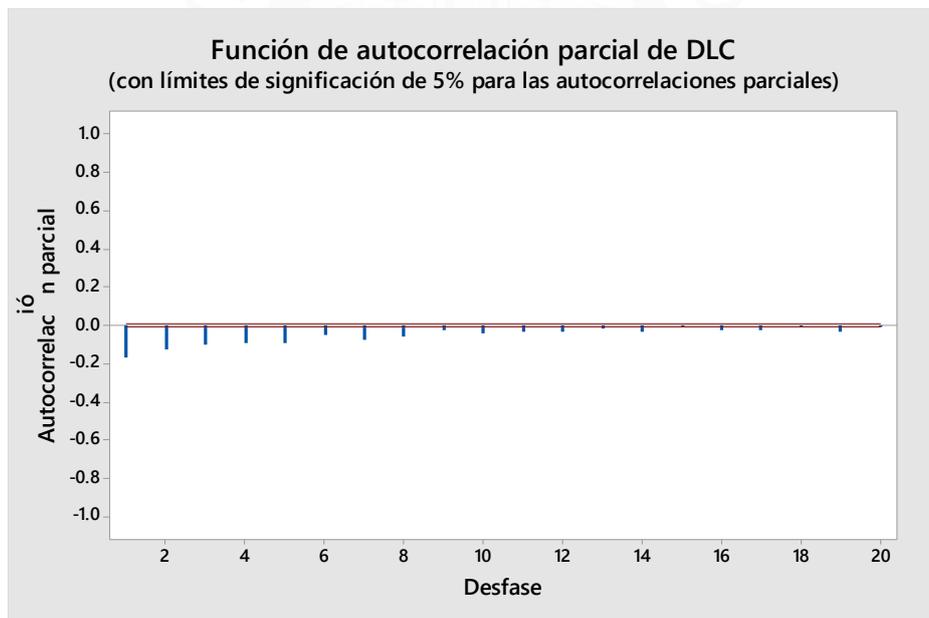


Figura 32. FAP para la serie DLC (Aumento).

Fuente: Elaboración propia.

En resumen, para eliminar la tendencia de la serie y volverla estacionaria se requiere de transformación logarítmica ($z_t \rightarrow z'_t = \ln(z_t)$) y una diferencia regular ($d = 1$). Además, la FAS y FAP sugieren un patrón autorregresivo de orden 1 ($p = 1$) y un patrón de media móvil de orden 1 ($d = 1$). El modelo ARIMA que subyace de este análisis es ARIMA (1,1,1).

$$(1 - \phi_1 B)(1 - B)^1 z'_t = (1 - \theta_1 B) a_t$$

Al finalizar el ejercicio de predicción, se revierte la transformación logarítmica, de modo que:

$$\hat{z}_{t+h|t} = \exp(\hat{z}'_{t+h|t})$$

3.2.2.2 Estimación. Se estimaron los parámetros, siguiendo el método detallado en el apartado 3.2, obteniendo un parámetro promedio:

	Value	StandardError	TStatistic	PValue
Constant	0	0	NaN	NaN
AR{1}	0.63411	0.0080401	78.869	0
MA{1}	-0.87304	0.005589	-156.21	0
Variance	0.1521	0.00077243	196.91	0

Figura 33. Parámetros promedio estimados – Modelo ARIMA (1,1,1) diario.

Fuente: Elaboración propia.

El modelo ARIMA (1,1,1) estimado es:

$$(1 - 0.6341B)(1 - B)z'_t = (1 + 0.8730B)a_t$$

Se confirma la validez del modelo al revisar el p-valor del contraste es menor al valor crítico de 0.05, confirmando que los coeficientes son necesarios para el modelo. Además, se analiza los residuos del modelo. En la Figura 34 a la Figura 37 aprecia que la FAS y FAP corresponden a ruido blanco.

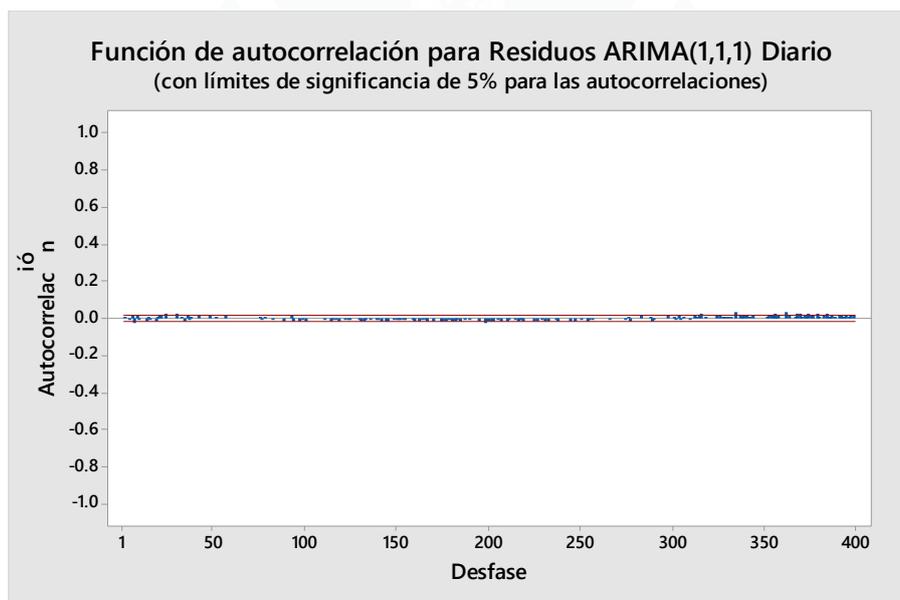


Figura 34. FAS para errores del modelo ARIMA (1,1,1) diario.

Fuente: Elaboración propia.

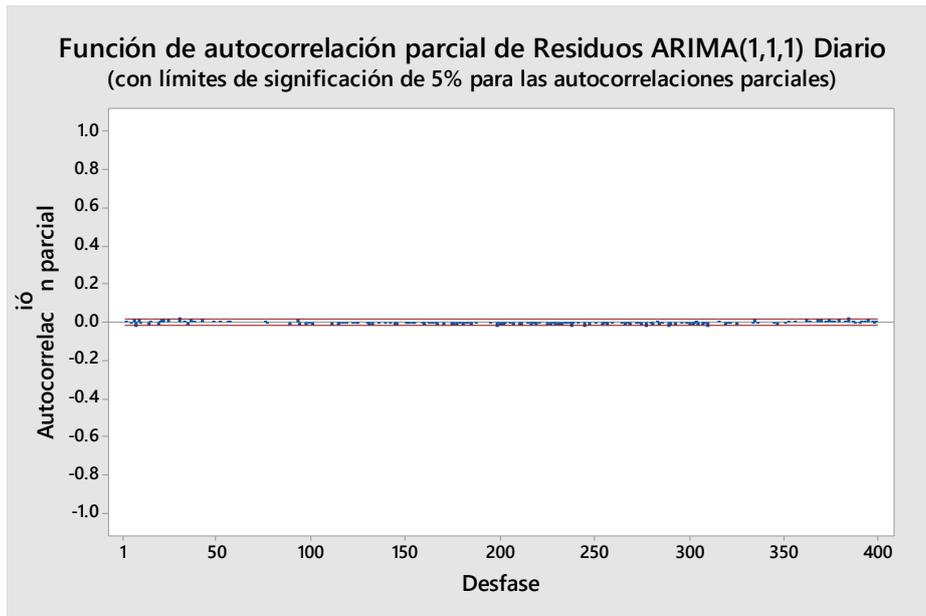


Figura 35. FAP para errores del modelo ARIMA (1,1,1) diario.

Fuente: Elaboración propia.

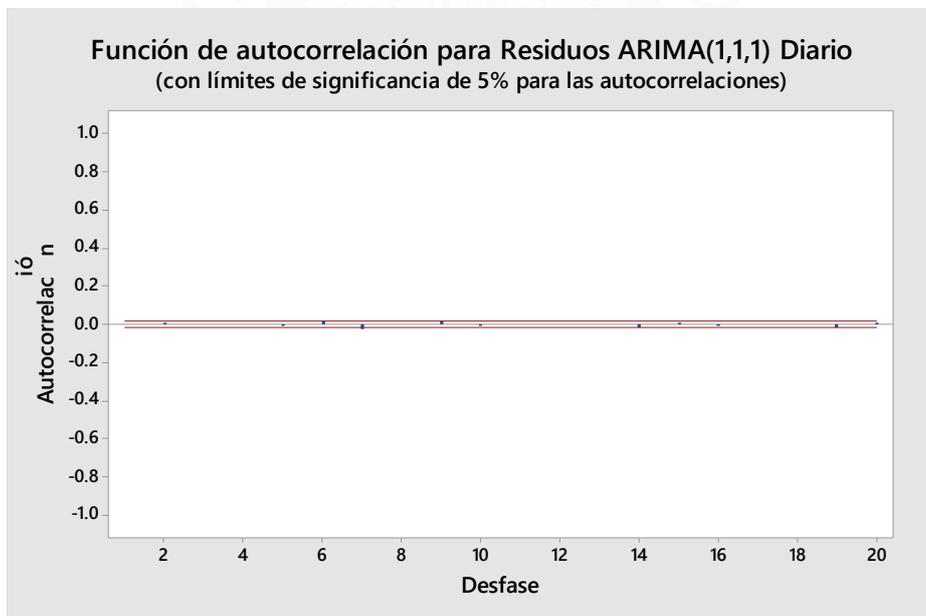


Figura 36. FAS para errores del modelo ARIMA (1,1,1) diario (Aumento).

Fuente: Elaboración propia.

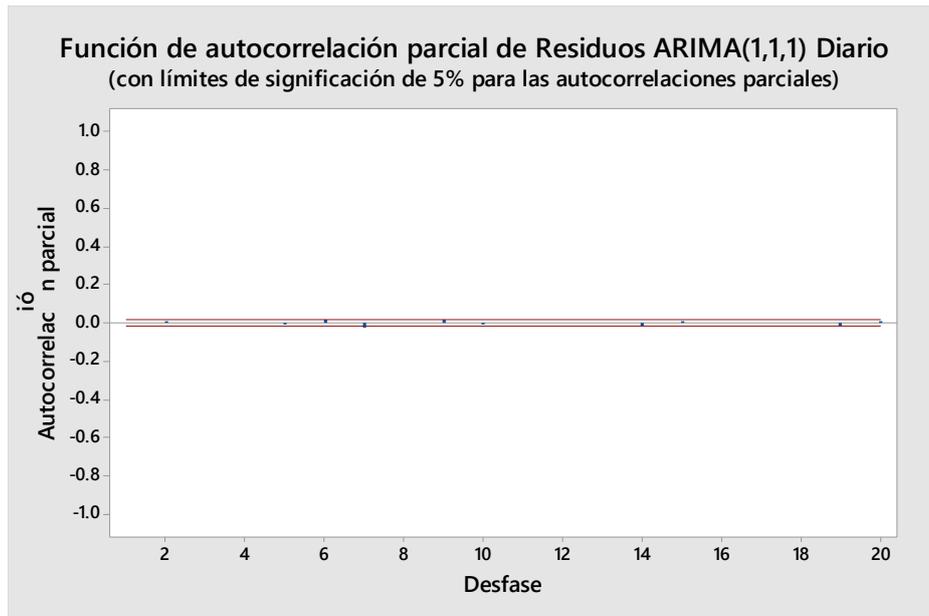


Figura 37. FAP para errores del modelo ARIMA (1,1,1) diario (Aumento).

Fuente: Elaboración propia.

El error cuadrático medio obtenido para horizonte de predicción con este modelo se observa en la Figura 38. Es posible notar que el error cuadrático medio aumenta en la medida que el horizonte de predicción se aleja de la data de muestra del modelo. La curva representada es ligeramente cóncava y el valor máximo de ECMP obtenido con este modelo es de 3.6285×10^4 .

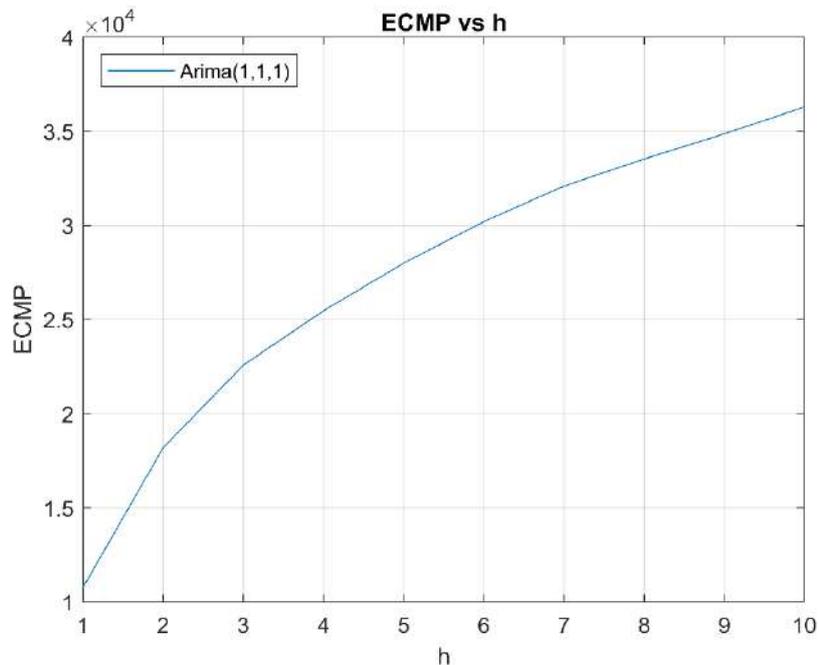


Figura 38. ECMP vs h - Modelo ARIMA (1,1,1) diario.

Fuente: Elaboración propia.

3.3 Modelización semanal

La modelización utiliza los datos del caudal promedio semanal recolectados en la Estación Ardilla. El periodo de evaluación está comprendido entre el 01 de enero de 1951 al 14 de agosto de 2017.

Los modelos realizan predicciones en un horizonte de 10 semanas ($h = 10$). Para inicializar los modelos, se utiliza un periodo de muestra que comprende el caudal medio semanal entre la primera semana de 1951 hasta la última semana de 1979. Las primeras predicciones corresponden al caudal medio semanal de la primera semana de 1980 hasta la décima semana de 1980. El ejercicio de predicción se repite incorporando en cada iteración una semana al periodo de muestra y prediciendo las 10 semanas siguientes hasta agotar los datos del periodo de evaluación (1953 iteraciones).

Tabla 15. Procedimiento para generación de predicciones semanales.

	Fin del periodo de muestra	Periodo de muestra	Horizonte de predicción
Primera iteración ($i = 1$)	Semana ₁ = Última semana 1979	Caudal (Primera semana 1951 hasta Última semana 1979)	Caudal (Primera semana 1980 hasta Décima semana 1980)
Siguientes iteraciones (i)	Semana _i = Semana _{i-1} + 1	Caudal (Primera semana 1951 hasta Semana _i)	Caudal (Semana _i + 1 hasta Semana _i + 10)

Fuente: Elaboración propia.

3.3.1 Métodos de suavizamiento exponencial

Se evalúa los diferentes métodos de suavización exponencial considerados en el apartado 2.1, incluyendo las versiones con tendencia amortiguada de los mismos.

3.3.1.1 Holt. El método considera dos ecuaciones de suavizamiento. La primera para el nivel y la segunda para la tendencia. Los parámetros utilizados son:

- Alfa (α): Parámetro de la componente de nivel
- Beta (β): Parámetro de la componente de tendencia.

Para el desarrollo del presente modelo, en adelante Modelo Holt, se utiliza una función de optimización, que permite encontrar un parámetro óptimo que minimice el error cuadrático medio de las predicciones. Se hace uso de la transformación logarítmica para disminuir el impacto de la variabilidad propia de la serie. En cada ejercicio de predicción los parámetros del modelo se inicializan con los parámetros óptimos de la iteración anterior.

Los parámetros promedio resultantes de aplicar el método descrito en el apartado 3.3 son:

Tabla 16. Parámetros promedio estimados - Modelo Holt semanal.

Modelo	Parámetros	
	α	β
Holt	0.8810	0.0114

Fuente: Elaboración propia.

El modelo queda expresado como:

$$\hat{z}_{t+h|t} = L_t + hT_t$$

Donde:

$$L_t = (0.8810)z_t + (1 - 0.8810)(L_{t-1} + T_{t-1})$$

$$T_t = (0.0114)(L_t - L_{t-1}) + (1 - 0.0114)T_{t-1}$$

$$h = 1, 2, \dots, 10$$

El error cuadrático medio obtenido con este modelo para cada horizonte de predicción se encuentra en la Figura 39. La curva del error cuadrático medio es convexa y aumenta con el incremento del horizonte de predicción. El valor mínimo de ECMP es 1.9688×10^4 y el valor máximo es 16.919×10^4 .

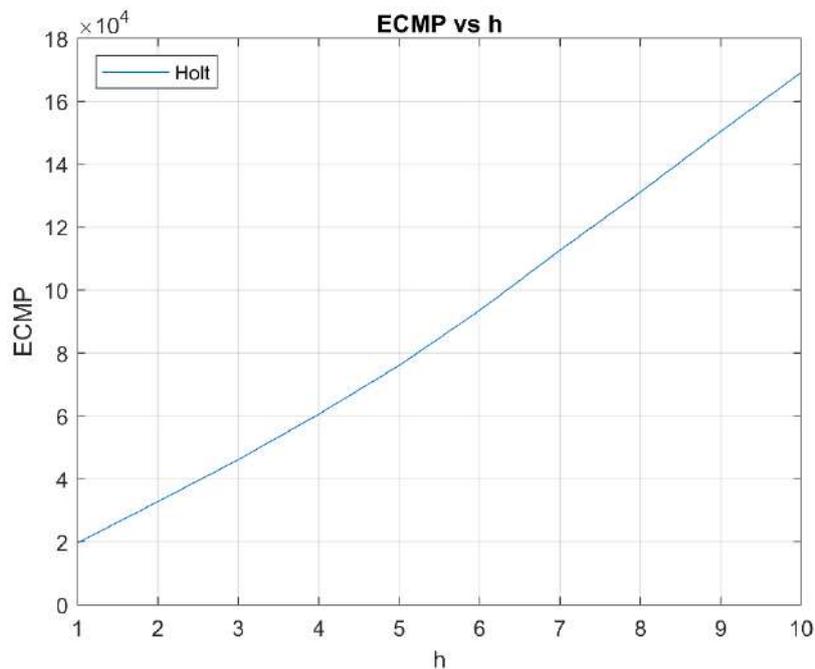


Figura 39. ECMP vs h - Modelo Holt semanal.

Fuente: Elaboración propia.

3.3.1.2 Holt con tendencia amortiguada. Este método considera los mismos parámetros que el método de Holt y agrega un nuevo parámetro, ϕ (Φ), con la función de amortiguar la tendencia en observaciones futuras.

El modelo desarrollado, en adelante Modelo Holt Damped, utiliza una función de optimización para la búsqueda de parámetros óptimos. Así mismo, se emplea la serie transformada en logaritmos, para reducir el impacto de la variabilidad. En cada iteración, los parámetros se inicializan con los parámetros óptimos de la iteración anterior.

Tabla 17. Parámetros promedio estimados - Modelo Holt Damped semanal.

Modelo	Parámetros		
	α	β	ϕ
Holt Damped	0.8432	0.4587	2.346×10^{-5}

Fuente: Elaboración propia.

El modelo resultante se expresa:

$$\hat{z}_{t+h|t} = L_t + ((2.346 \times 10^{-5}) + \dots + (2.346 \times 10^{-5})^h)T_t$$

Donde:

$$L_t = (0.8432)z_t + (1 - 0.8432)(L_{t-1} + (2.346 \times 10^{-5})T_{t-1})$$

$$T_t = (0.4587)(L_t - L_{t-1}) + (1 - 0.4587)(2.346 \times 10^{-5})T_{t-1}$$

$$h = 1, 2, \dots, 10$$

El error cuadrático medio resultante para cada horizonte de predicción está representado en la Figura 40. El incremento del error cuadrático medio respecto del horizonte de predicción es aproximadamente lineal. El valor mínimo de ECMP es 1.8852×10^4 en el primer horizonte de predicción y el valor máximo es 10.789×10^4 en el último horizonte de predicción.

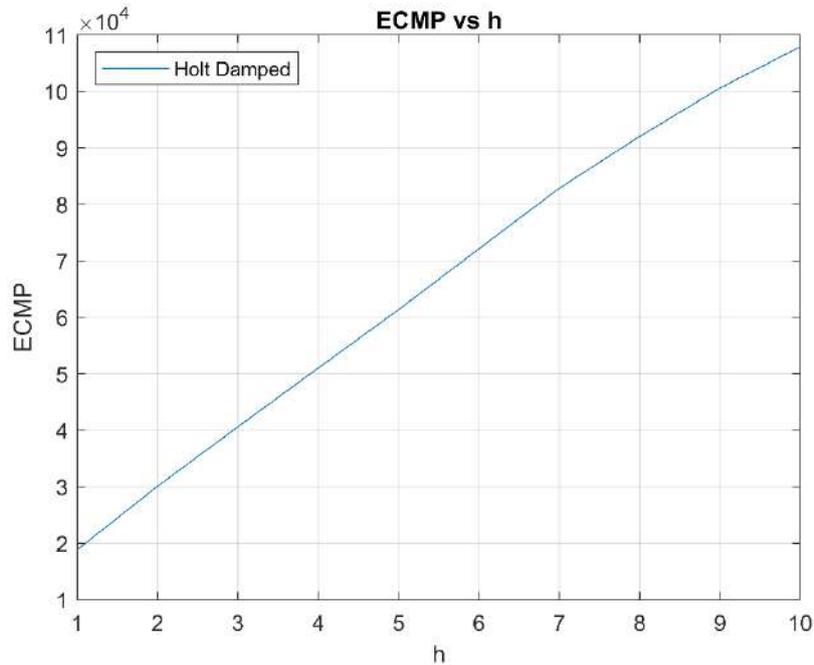


Figura 40. ECMP vs h - Modelo Holt Damped semanal.

Fuente: Elaboración diaria.

3.3.1.3 Holt-Winters aditivo. El método considera tres ecuaciones de suavizamiento. Las dos primeras correspondientes al nivel y la tendencia, igual al método de Holt, y se añade una tercera ecuación que corresponde a la estacionalidad. Los parámetros considerados por este método son:

- Alfa (α): Parámetro de la componente de nivel.
- Beta (β): Parámetro de la componente de tendencia.
- Gamma (γ): Parámetro de la componente estacional.

Para el desarrollo de este modelo, en adelante denominado Modelo Holt-Winters, se emplea la misma función de optimización utilizada en el apartado 3.3.1.1, que permite encontrar los parámetros óptimos que reducen el error cuadrático medio de las predicciones. La serie está compuesta de caudales promedio semanales y para este modelo se presupone que existe variación periódica y predecible cada año, considerando una estacionalidad de 52 semanas ($s = 52$). Además, se realiza la transformación logarítmica de la serie, para reducir el impacto de la variabilidad. En cada ejercicio de predicción, los parámetros se inicializan utilizando el parámetro óptimo de la iteración anterior. Los parámetros promedio obtenidos son:

Tabla 18. Parámetros promedio estimados - Modelo Holt-Winters semanal.

Modelo	Parámetros		
	α	β	γ
Holt-Winters	0.8292	5.6906×10^{-4}	0.0451

Fuente: Elaboración propia.

El modelo se expresa como:

$$\hat{z}_{t+h|t} = L_t + hT_t + S_{t+h-52(k+1)}$$

Donde:

$$L_t = (0.8292)(z_t - S_{t-52}) + (1 - 0.8292)(L_{t-1} + T_{t-1})$$

$$T_t = (5.6906 \times 10^{-4})(L_t - L_{t-1}) + (1 - 5.6906 \times 10^{-4})T_{t-1}$$

$$S_t = (0.0451)(z_t - L_{t-1} - T_{t-1}) + (1 - 0.0451)S_{t-52}$$

$$h = 1, 2, \dots, 10$$

$$k = \frac{(h-1)}{52}$$

En la Figura 41 se puede comprobar el error cuadrático medio obtenido para cada horizonte de predicción. La curva del error cuadrático medio es ligeramente cóncava y es directamente proporcional al horizonte de predicción. El valor máximo de ECMP es 9.1453×10^4 y el valor mínimo es 2.076×10^4 .

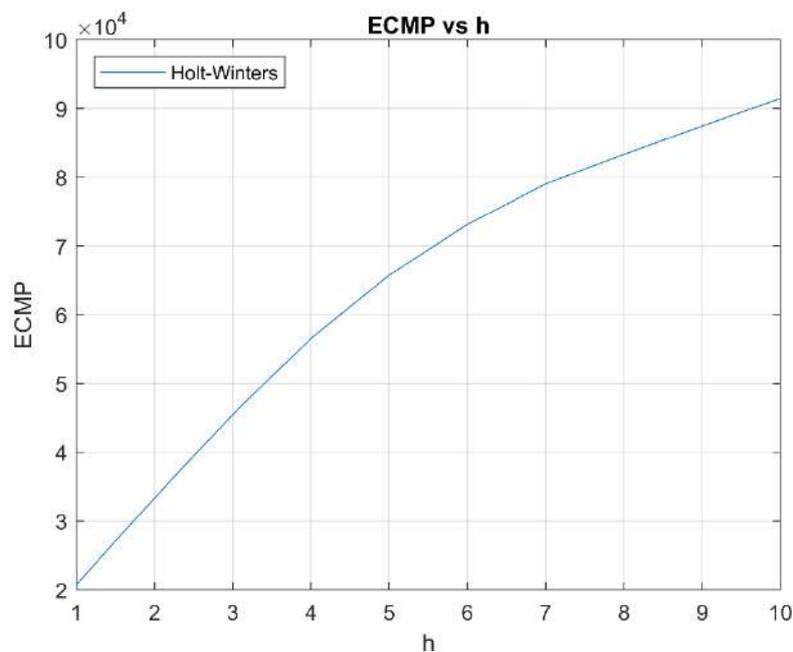


Figura 41. ECMP vs h – Modelo Holt-Winters semanal.

Fuente: Elaboración propia.

3.3.1.4 Holt-Winters aditivo con tendencia amortiguada. Este método emplea los mismos parámetros detallados en el apartado 3.3.1.3 e incorpora un nuevo parámetro con el fin de amortiguar la tendencia en observaciones futuras, denominado ϕ (ϕ).

Este modelo, en adelante Modelo Holt-Winters Damped, emplea una función de optimización para la búsqueda de parámetros óptimos. De igual manera, se reduce la variabilidad de la serie utilizando la transformación logarítmica. Se supone cambios regulares de la serie en periodos de 52 semanas ($s = 52$). En cada iteración, los parámetros se inicializan con los parámetros óptimos de la iteración anterior.

Tabla 19. Parámetros promedio estimados - Modelo Holt-Winters Damped semanal.

Modelo	Parámetros			
	α	β	γ	ϕ
Holt-Winters Damped	0.8277	0.3448	0.0451	8.6562×10^{-5}

Fuente: Elaboración propia.

En consecuencia, el modelo queda expresado como:

$$\hat{z}_{t+h|t} = L_t + ((8.6562 \times 10^{-5}) + \dots + (8.6562 \times 10^{-5})^h)T_t + S_{t+h-52(k+1)}$$

Donde:

$$L_t = (0.8277)(z_t - S_{t-52}) + (1 - 0.8277)(L_{t-1} + (8.6562 \times 10^{-5})T_{t-1})$$

$$T_t = (0.3448)(L_t - L_{t-1}) + (1 - 0.3448)(8.6562 \times 10^{-5})T_{t-1}$$

$$S_t = (0.0451)(z_t - L_{t-1} - (8.6562 \times 10^{-5})T_{t-1}) + (1 - 0.0451)S_{t-52}$$

$$h = 1, 2, \dots, 10$$

$$k = \frac{(h-1)}{52}$$

El error cuadrático medio para cada horizonte de predicción se observa en la Figura 42. El comportamiento del error cuadrático medio con este método es similar al Modelo Holt-Winters. En el primer horizonte de predicción se obtiene el valor mínimo de ECMP de 2.0708×10^4 y el valor máximo de 9.0039×10^4 se encuentra en el último horizonte de predicción.

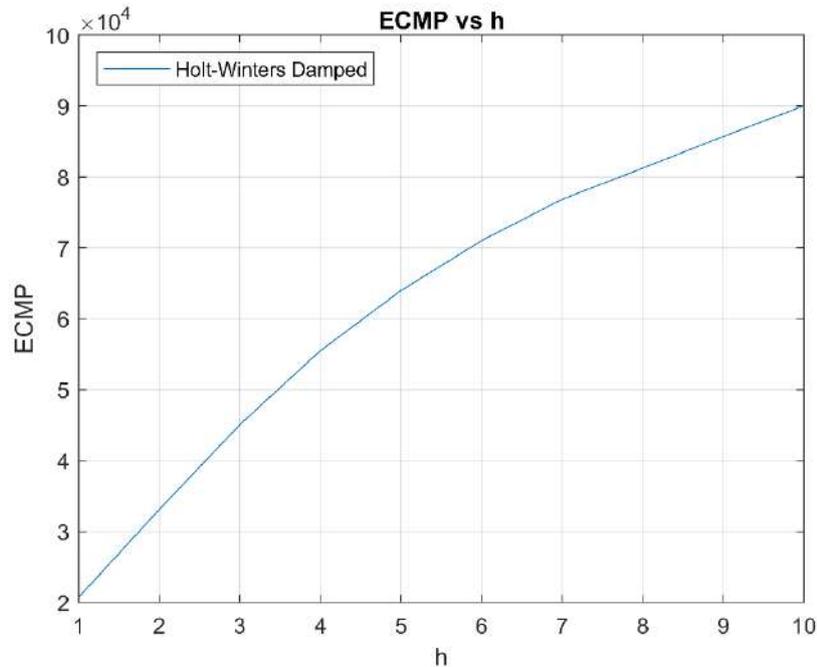


Figura 42. ECMP vs h - Modelo Holt-Winters Damped semanal.

Fuente: Elaboración propia.

3.3.1.5 Holt-Winters multiplicativo. Este método es usado en series que presentan una variación estacional que cambia proporcionalmente al nivel de la serie. (Hyndman & Athanasopoulos, 2018)

Utiliza los mismos parámetros detallados en el apartado 3.3.1.3. El modelo, que en adelante se denomina Modelo Holt-Winters Multiplicativo, se vale de una función de optimización para encontrar los parámetros óptimos, asegurando el menor error cuadrático medio. Se supone que la serie de caudal medio semanal experimenta fluctuaciones regulares en periodos anuales ($s = 52$). Debido a la naturaleza del método, no se emplean transformaciones en la serie. En cada ejercicio de predicción, los parámetros del modelo se inicializan con el parámetro óptimo de la iteración anterior.

Los parámetros promedio obtenidos son:

Tabla 20. Parámetros promedio estimados - Modelo Holt-Winters Multiplicativo semanal.

Modelo	Parámetros		
	α	β	γ
Holt-Winters Multiplicativo	0.2673	0.0992	0.3085

Fuente: Elaboración propia.

Por tanto, el modelo que resulta es:

$$\hat{z}_{t+h|t} = (L_t + hT_t)S_{t+h-52(k+1)}$$

Donde:

$$L_t = (0.2673)\left(\frac{z_t}{S_{t-52}}\right) + (1 - 0.2673)(L_{t-1} + T_{t-1})$$

$$T_t = (0.0992)(L_t - L_{t-1}) + (1 - 0.0992)T_{t-1}$$

$$S_t = (0.3085)\left(\frac{z_t}{L_{t-1} + T_{t-1}}\right) + (1 - 0.3085)S_{t-52}$$

$$h = 1, 2, \dots, 10$$

$$k = \frac{(k+1)}{52}$$

En la Figura 43 se observa el error cuadrático medio para cada horizonte de predicción obtenido con este modelo. El error cuadrático medio tiene un crecimiento exponencial a media que se incrementa el horizonte de predicción. Alcanza grandes valores de ECMP, siendo el máximo de 2.7974×10^9 y el mínimo de 2.7974×10^7 .

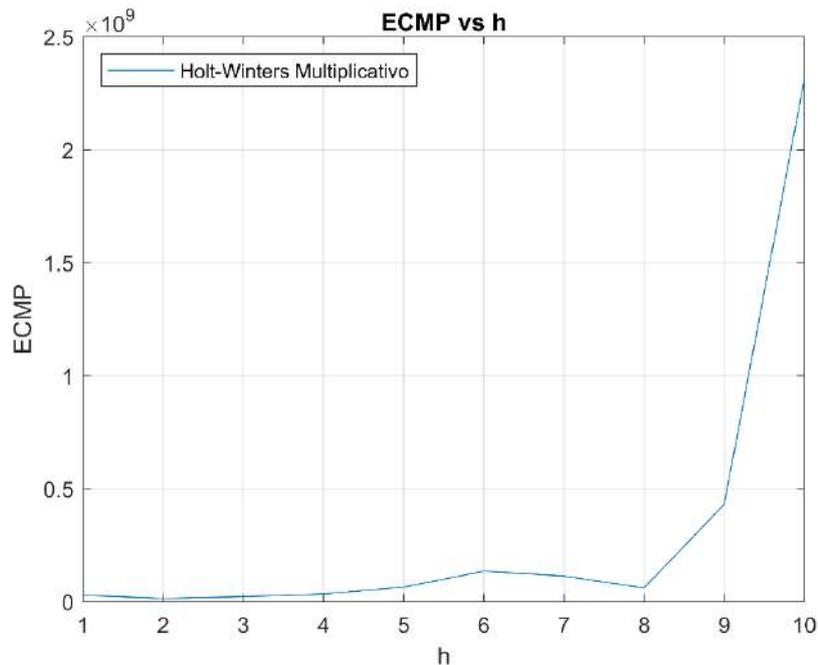


Figura 43. ECMP vs h - Modelo Holt-Winters Multiplicativo semanal.

Fuente: Elaboración propia.

3.3.1.6 Holt-Winters multiplicativo con tendencia amortiguada. El método considera los mismos parámetros detallados en el método Holt-Winters Multiplicativo, agregando el parámetro F_i (ϕ) que amortigua la tendencia en observaciones futuras.

El modelo a construir, en adelante denominado Holt-Winters Multiplicativo Damped, emplea una función de optimización para hallar los parámetros óptimos que reduzcan el error cuadrático medio de las predicciones. Se supone una componente estacional en la serie de 52 semanas ($s = 52$). En cada iteración, los parámetros del modelo se inicializan con los parámetros óptimos de la iteración anterior.

Tabla 21. Parámetros promedio estimados - Modelo Holt-Winters Multiplicativo Damped semanal.

Modelo	Parámetros			
	α	β	γ	ϕ
Holt-Winters Multiplicativo Damped	0.5490	0.3487	0.0127	0.4980

Fuente: Elaboración propia.

El modelo resultante se expresa como:

$$\hat{z}_{t+h|t} = (L_t + ((0.4980) + \dots + (0.4980)^h)T_t)S_{t+h-52(k+1)}$$

Donde:

$$L_t = (0.5490)\left(\frac{z_t}{S_{t-52}}\right) + (1 - 0.5490)(L_{t-1} + (0.4980)T_{t-1})$$

$$T_t = (0.3487)(L_t - L_{t-1}) + (1 - 0.3487)(0.4980)T_{t-1}$$

$$S_t = (0.0127)\left(\frac{z_t}{L_{t-1} + (0.4980)T_{t-1}}\right) + (1 - 0.0127)S_{t-52}$$

$$h = 1, 2, \dots, 10$$

$$k = \frac{(k+1)}{52}$$

En la Figura 44 se aprecia el error cuadrático medio para cada horizonte de predicción obtenido con este modelo. En la gráfica se observa un comportamiento errático producto de las características propias de la serie de caudal semanal. El modelo presenta un error cuadrático medio de gran magnitud, reduciendo la capacidad de predicción.

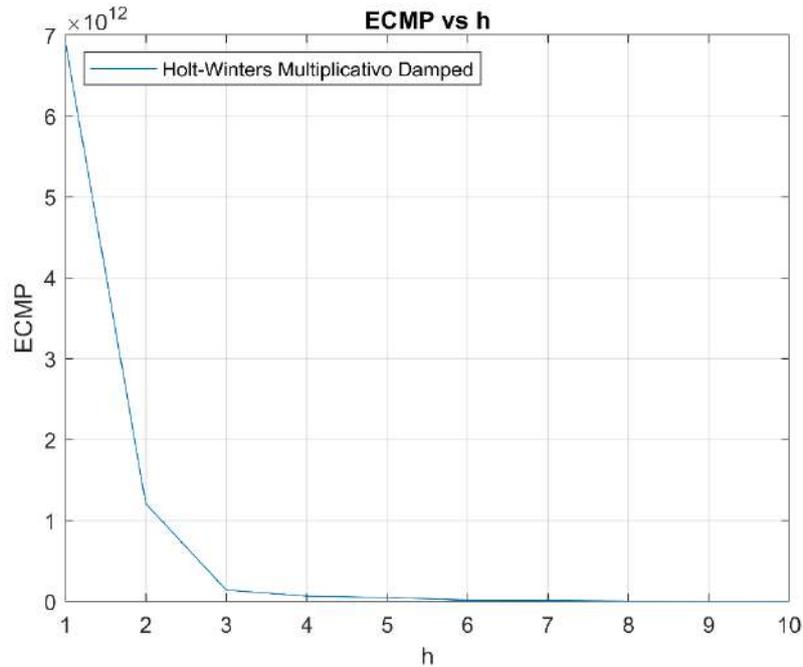


Figura 44. ECMP vs h - Modelo Holt-Winters Multiplicativo Damped semanal.

Fuente: Elaboración propia.

3.3.2 Métodos ARIMA

3.3.2.1 Identificación. La serie de caudal medio semanal cuenta con 3477 observaciones y una distribución presentada en la Figura 15. Con el objetivo de reducir la variabilidad de la serie, se requiere la utilización de una transformación de la familia Box-Cox. La transformación aplicada para revertir el comportamiento heterocedástico es la transformación logarítmica. La distribución de la serie transformada se observa en la Figura 21, en adelante se denomina LCS. Para completar los criterios de estacionariedad demandados por el método ARIMA, se busca que la serie transformada tenga una media constante. Es necesario la toma de diferencias regulares de la serie. La serie resultante, nombrada DLCS ($DLCS_t = LCS_t - LCS_{t-1}$), es estacionaria como se aprecia en la Figura 45.

Se analiza la correlación entre las observaciones en diferentes instantes de tiempo para definir el modelo ARIMA óptimo. Se emplean la FAS y FAP para evaluar la correlación entre las observaciones t hasta el instante $t-k$.

Analizando la Figura 46 y Figura 47, se observa el cumplimiento del decrecimiento exponencial de las funciones. Además, se observa alternancia de signos y una forma sinusoidal. Esta alternancia depende de los signos de los parámetros y la forma sinusoidal se explica con la existencia de raíces complejas en la ecuación característica del proceso. Ambas funciones muestran valores fuera de las bandas de significatividad en la componente regular, además, existen algunos coeficientes ligeramente fuera de las bandas en la componente estacional, por lo que se debe evaluar la inclusión en el modelo.

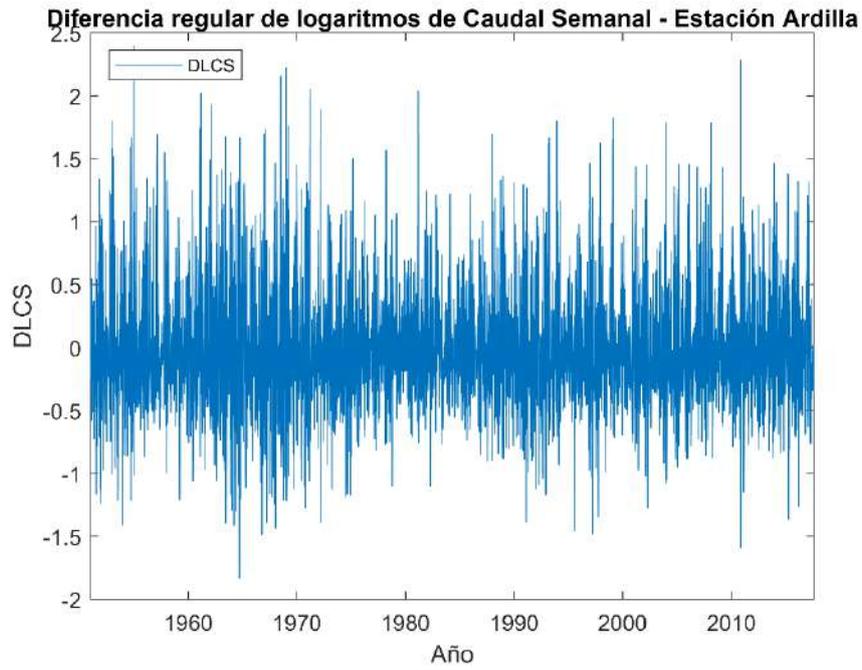


Figura 45. Diferencia regular de logaritmos de caudal promedio semanal.

Fuente: Elaboración propia.

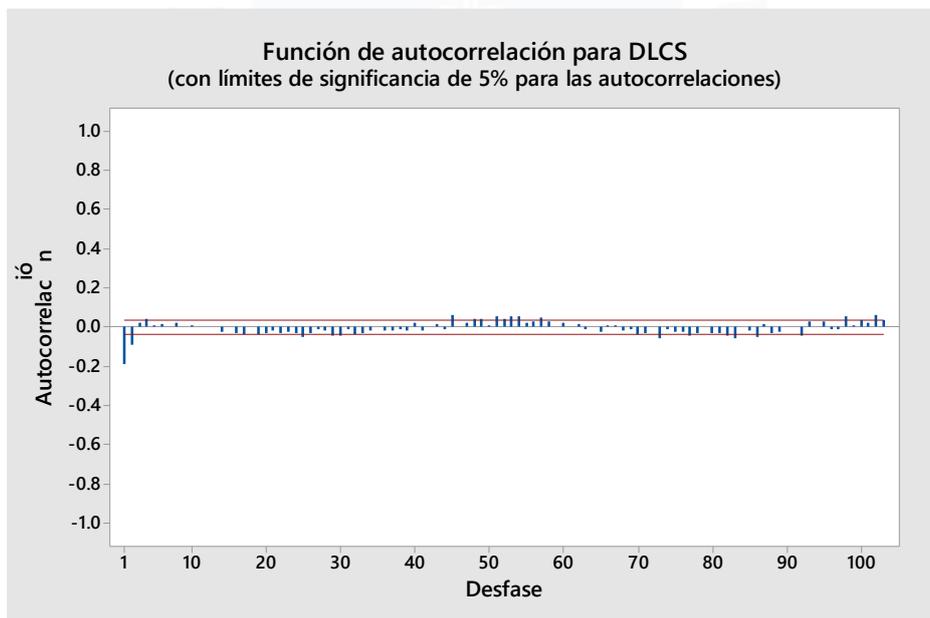


Figura 46. FAS para la serie DLCS.

Fuente: Elaboración propia.

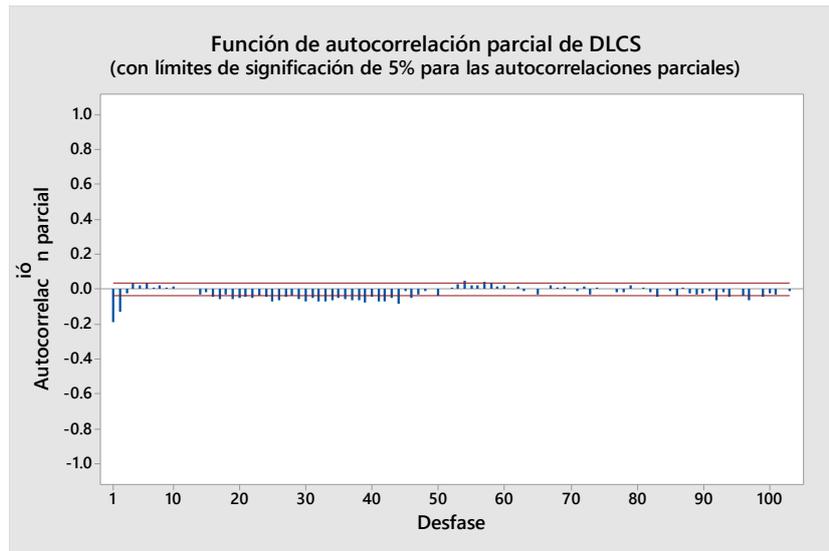


Figura 47. FAP para la serie DLCS.

Fuente: Elaboración propia.

Comparando la FAS y FAP obtenidas de la serie con las gráficas FAS y FAP teóricas, existe una gran similitud con el último ejemplo de la Figura 10, correspondiente a un modelo AR (2), aunque también existe similitud con el quinto ejemplo de la Figura 12, lo que sugiere un modelo ARMA (1,1). Se conviene considerar los modelos para la comparación.

A fin de evaluar la relevancia de la componente estacional de la serie, se toma una diferencia estacional, originando la serie denominada como D52DLCS, donde $D52DLCS_t = DLCS_t - DLCS_{t-52}$.

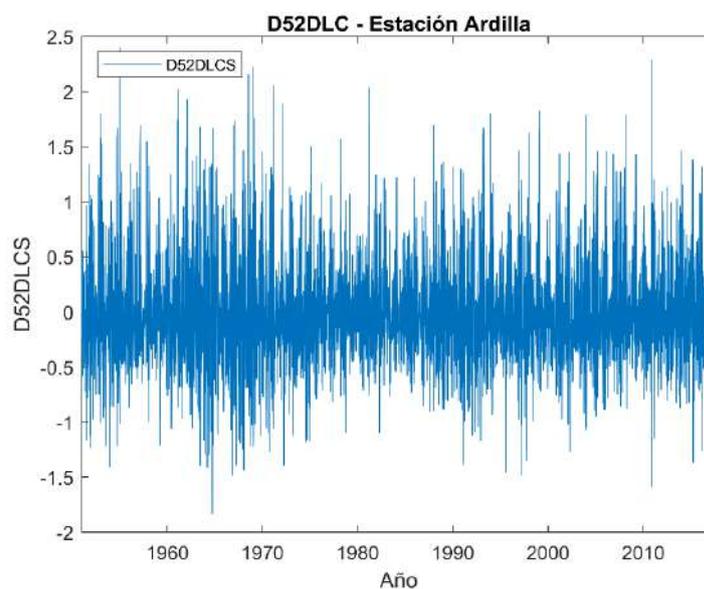


Figura 48. Diferencia estacional y diferencia regular de logaritmos de caudal promedio semanal.

Fuente: Elaboración propia.

Los coeficientes de la FAS y FAP de la nueva serie D52DLCS decrecen de manera exponencial. Se muestran valores fuera de las bandas de significatividad y se revela una fuerte correlación en la componente estacional.

Al comparar la Figura 49 y Figura 50 contra las gráficas FAS y FAP teóricas, en la componente regular se aprecia un comportamiento similar al quinto ejemplo de la Figura 12, lo que sugiere un modelo ARMA (1,1). En el componente estacional se observa el comportamiento marcado de un modelo MA (1), similar al presentado en el segundo ejemplo de la Figura 11.

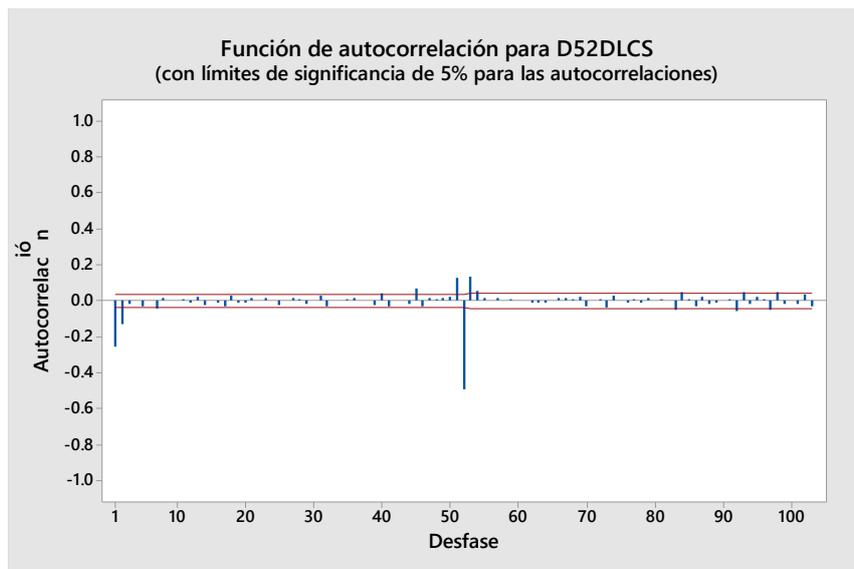


Figura 49. FAS para la serie D52DLCS.

Fuente: Elaboración propia.

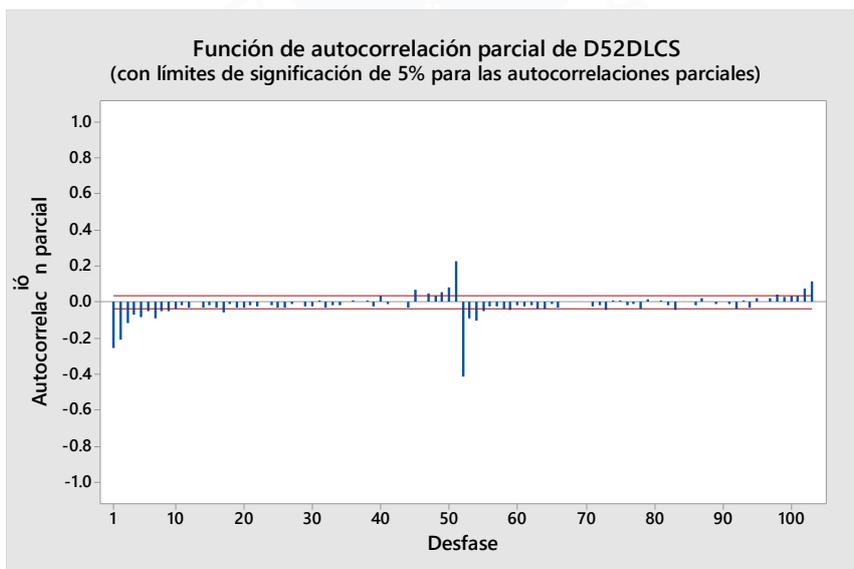


Figura 50. FAP para la serie D52DLCS.

Fuente: Elaboración propia.

En resumen, para eliminar la tendencia de la serie y conseguir la estacionariedad se utiliza la transformación logarítmica ($z_t \rightarrow z'_t = \ln(z_t)$) y también una diferencia regular ($d = 1$). El análisis realizado de la FAS y FAP sugiere tres modelos distintos.

El primer modelo con un patrón autorregresivo de orden 1 ($p = 1$) y un patrón de media móvil de orden 1 ($q = 1$), decantando en un modelo ARIMA (1,1,1).

$$(1 - \phi_1 B)(1 - B)^1 z'_t = (1 - \theta_1 B) a_t$$

El segundo modelo con un patrón autorregresivo de orden 2 ($p = 2$) y sin considerar un patrón de media móvil. El modelo que subyace es ARIMA (2,1,0).

$$(1 - \phi_1 B - \phi_2 B^2)(1 - B)^1 z'_t = a_t$$

El tercer modelo considera un patrón autorregresivo de orden 1 ($p = 1$) y un patrón de media móvil de orden 1 ($q = 1$) para la componente regular. Además, se requiere de una diferencia en la componente estacional ($D = 1$) y se considera un patrón de media móvil de orden 1 ($Q = 1$). El modelo resultante es ARIMA (1,1,1)(0,1,1)₅₂.

$$(1 - \phi_1 B)(1 - B)(1 - B^{52}) z'_t = (1 - \theta_1 B)(1 - \Theta_{52} B^{52}) a_t$$

Al finalizar el ejercicio de predicción, es necesario revertir la transformación logarítmica, de modo que:

$$\hat{z}_{t+h|t} = \exp(\hat{z}'_{t+h|t})$$

3.3.2.2 Estimación. Se estimaron los parámetros, siguiendo el método detallado en el apartado 3.3, obteniendo como parámetro promedio para los modelos:

```
ARIMA(1,1,1) Model (Gaussian Distribution):
```

	Value	StandardError	TStatistic	PValue
Constant	0	0	NaN	NaN
AR{1}	0.21645	0.073179	2.9578	0.0030981
MA{1}	-0.4519	0.068191	-6.6271	3.4246e-11
Variance	0.26415	0.0062899	41.997	0

Figura 51. Parámetros promedio estimados - Modelo ARIMA (1,1,1) semanal.

Fuente: Elaboración propia.

ARIMA(2,1,0) Model (Gaussian Distribution):

	Value	StandardError	TStatistic	PValue
Constant	0	0	NaN	NaN
AR{1}	-0.21694	0.018176	-11.935	7.7496e-33
AR{2}	-0.14325	0.018259	-7.8455	4.3119e-15
Variance	0.25642	0.005631	45.537	0

Figura 52. Parámetros promedio estimados - Modelo ARIMA (2,1,0) semanal.

Fuente: Elaboración propia.

ARIMA(1,1,1) Model with Seasonal MA(52) (Gaussian Distribution):

	Value	StandardError	TStatistic	PValue
Constant	0	0	NaN	NaN
AR{1}	0.22241	0.067325	3.3036	0.00095461
MA{1}	-0.45416	0.06291	-7.2192	5.2285e-13
SMA{52}	0.071732	0.017837	4.0216	5.7793e-05
Variance	0.25559	0.0054888	46.565	0

Figura 53. Parámetros promedio estimados - Modelo ARIMA (1,1,1)(0,1,1)₅₂ semanal.

Fuente: Elaboración diaria.

Los modelos ARIMA estimados son:

ARIMA (1,1,1)

$$(1 - 0.21645B)(1 - B)z'_t = (1 + 0.4519B)a_t$$

ARIMA (2,1,0)

$$(1 + 0.21694B + 0.14325B^2)(1 - B)z'_t = a_t$$

ARIMA (1,1,1)(0,1,1)₅₂

$$(1 - 0.22241B)(1 - B)(1 - B^{52})z'_t = (1 + 0.45416B)(1 - 0.071732B^{52})a_t$$

Se confirma la validez de los modelos al revisar el p-valor del contraste es menor al valor crítico de 0.05 en todos los casos, confirmando que los coeficientes son necesarios para el modelado.

También se realiza el análisis de los residuos propios de cada modelo. De la Figura 54 hasta la Figura 59 se aprecia que la FAS y FAP de los modelos ARIMA (1,1,1) Semanal, ARIMA (2,1,0) Semanal y ARIMA (1,1,1)(0,1,1)₅₂.

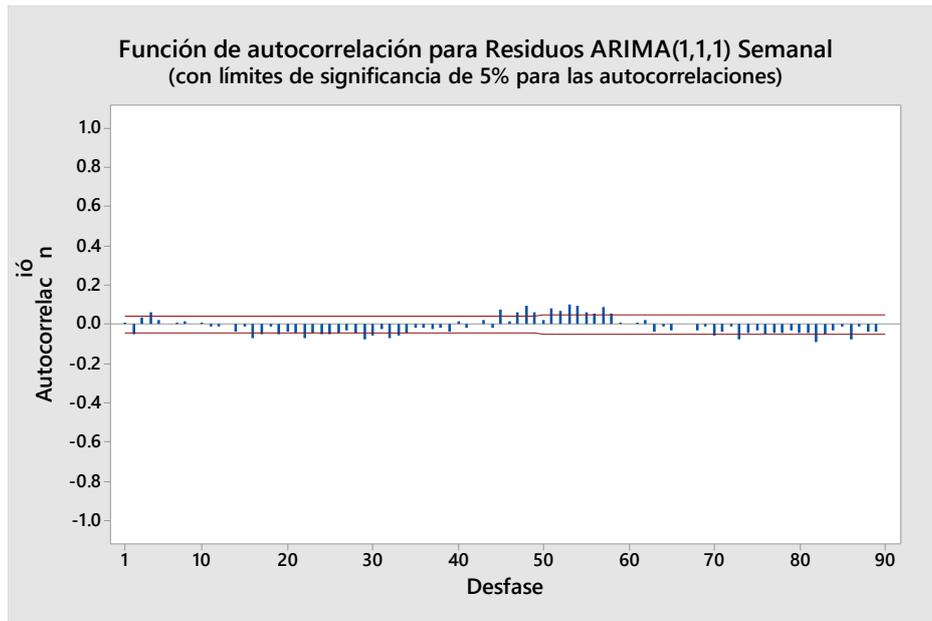


Figura 54. FAS para residuos del modelo ARIMA (1,1,1) semanal.

Fuente: Elaboración propia.

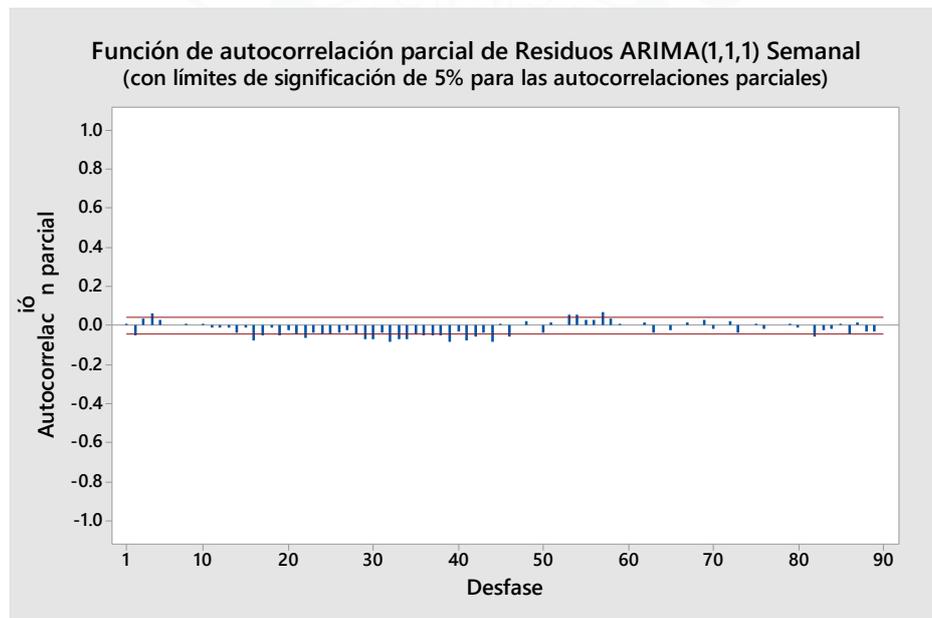


Figura 55. FAP para residuos del modelo ARIMA (1,1,1) semanal.

Fuente: Elaboración propia.

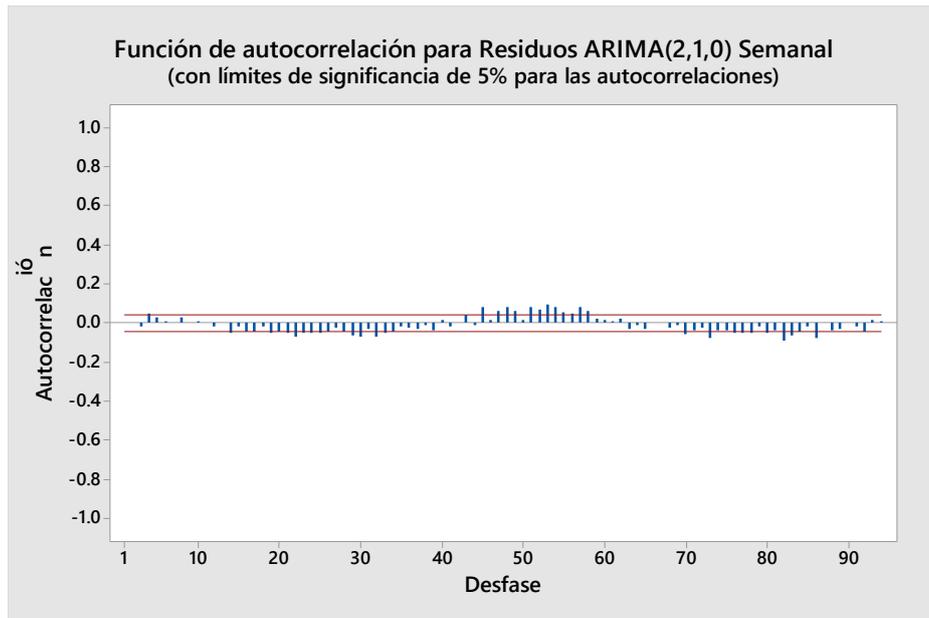


Figura 56. FAS para residuos del modelo ARIMA (2,1,0) semanal.

Fuente: Elaboración propia.

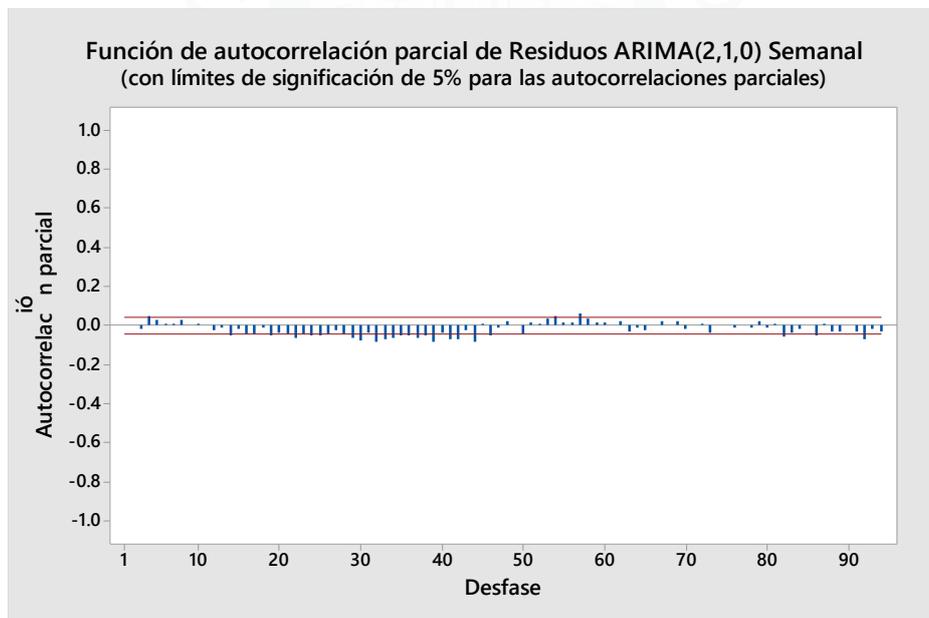


Figura 57. FAP para residuos del modelo ARIMA (2,1,0) semanal.

Fuente: Elaboración propia.

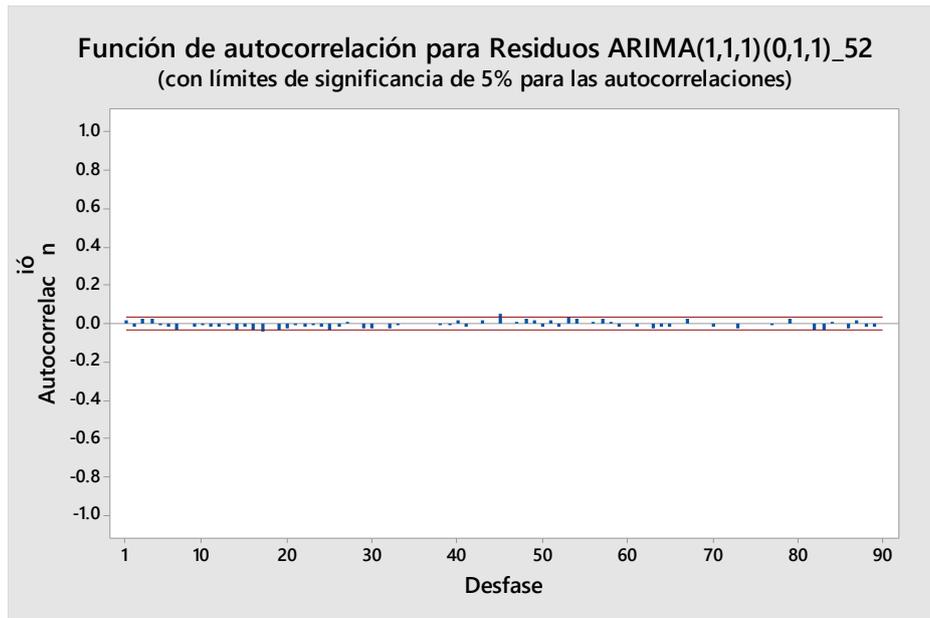


Figura 58. FAS para residuos del modelo ARIMA (1,1,1)(0,1,1)₅₂ semanal.

Fuente: Elaboración propia.

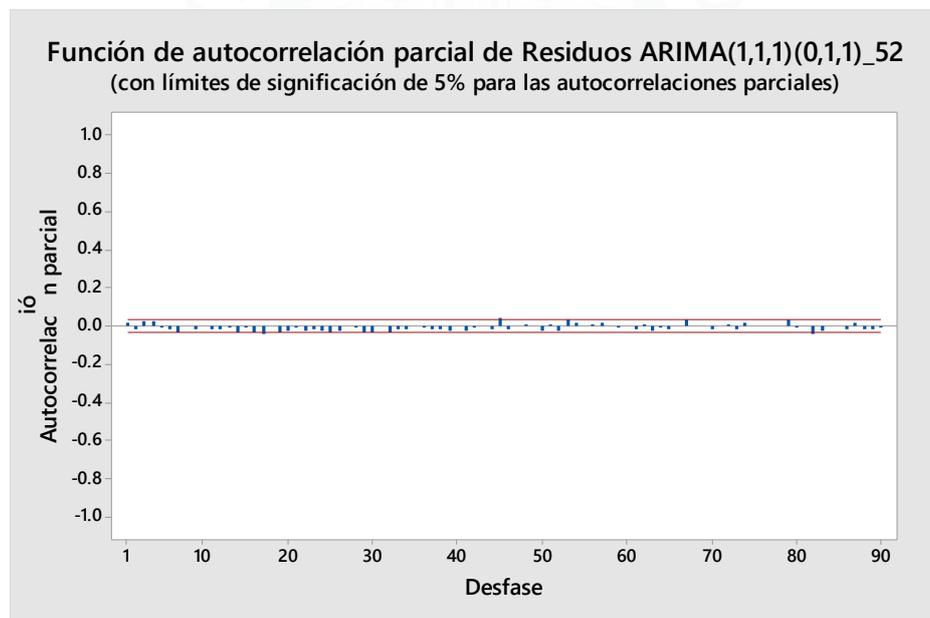


Figura 59. FAP para residuos del modelo ARIMA (1,1,1)(0,1,1)₅₂ semanal.

Fuente: Elaboración propia.

Se observa que algunos valores superan ligeramente las bandas de significatividad, denotando una posible correlación en la componente estacional en los dos primeros modelos. Este comportamiento se reduce en el último modelo que considera primeras diferencias en la componente estacional, donde la FAS y FAP corresponden a ruido blanco.

El error cuadrático medio obtenido para cada horizonte es la métrica que nos permite observar el rendimiento de los modelos y están representados desde la Figura 60 a la Figura 62. En todas las gráficas se aprecia que el error cuadrático medio aumenta a medida que se incrementa el horizonte de predicción. Los modelos ARIMA (1,1,1) y ARIMA (2,1,0) presentan un comportamiento más lineal, mientras que la curva del modelo ARIMA (1,1,1)(0,1,1)₅₂ es cóncava. El modelo ARIMA (1,1,1) alcanzó un valor mínimo de 1.8731×10^4 y un valor máximo de 10.504×10^4 , el modelo ARIMA (2,1,0) alcanzó un valor mínimo de 1.8866×10^4 y un valor máximo de 10.630×10^4 , y el modelo ARIMA (1,1,1)(0,1,1)₅₂ alcanzó un valor mínimo de 1.7875×10^4 y un valor máximo de 4.4557×10^4 .

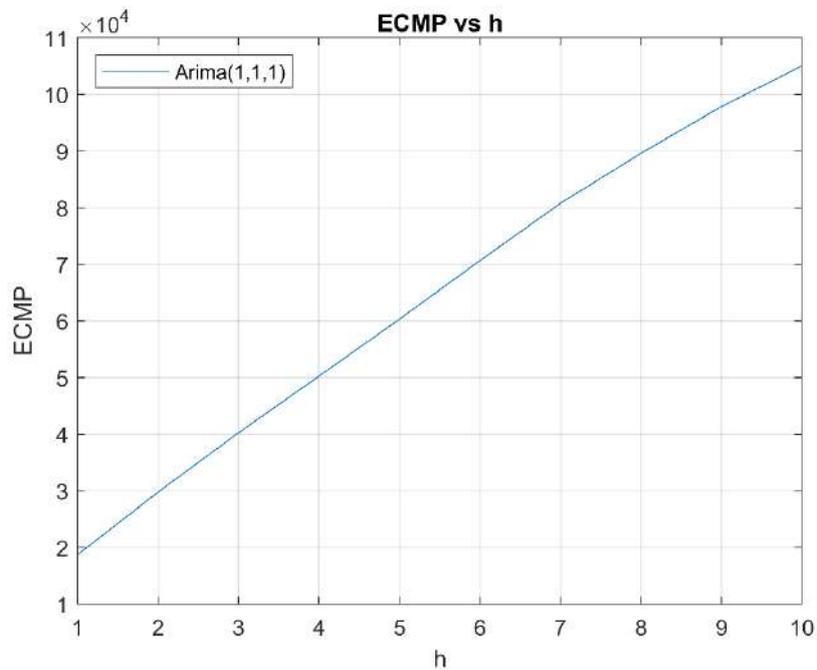


Figura 60. ECMP vs h - Modelo ARIMA (1,1,1) semanal.

Fuente: Elaboración propia.

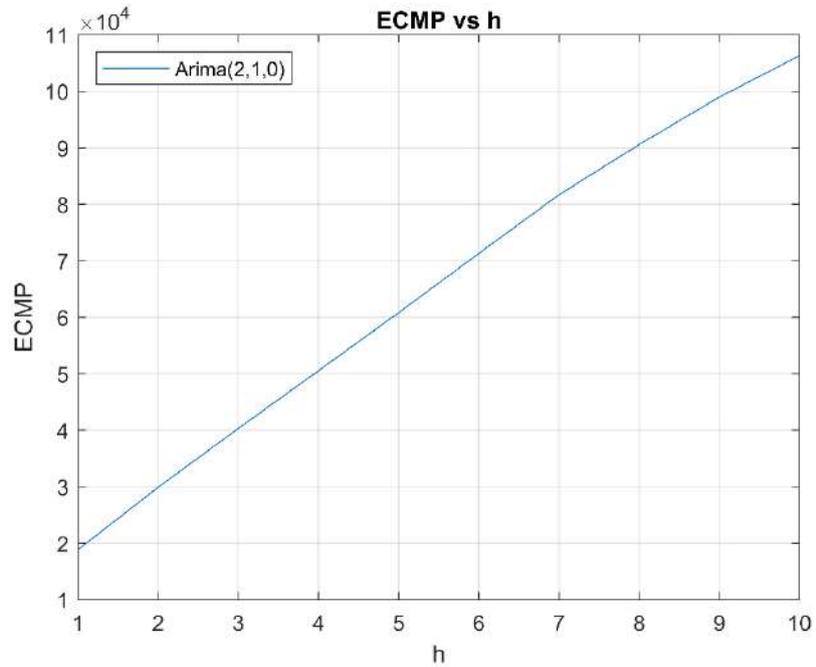


Figura 61. ECMP vs h - Modelo ARIMA (2,1,0) semanal.

Fuente: Elaboración propia.

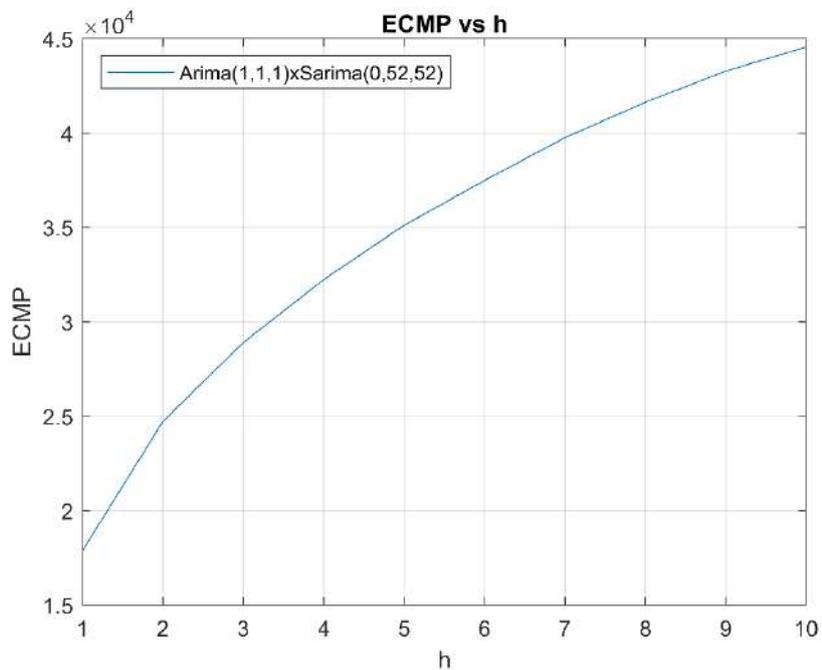


Figura 62. ECMP vs h - Modelo ARIMA (1,1,1)(0,1,1)₅₂ semanal.

Fuente: Elaboración propia.

Capítulo 4

Comparación de modelos

En este capítulo se compara los diferentes modelos de predicción univariante desarrollados en el Capítulo 3. Todas las predicciones se han generado en el periodo del 01 de enero de 1980 hasta el final de la data recaudada en la Estación Ardilla, siguiendo el procedimiento detallado en el apartado 3.2 para el análisis del caudal con frecuencia diaria y en el apartado 3.3 para el caudal con frecuencia semanal. Para realizar la comparativa se utiliza diferentes métricas que permitan evaluar el rendimiento de las predicciones.

El primer criterio de evaluación utilizado es el error cuadrático medio:

$$ECMP = \frac{1}{n} \sum_{i=1}^n (z_i - \hat{z}_i)^2$$

Donde el menor valor de error cuadrático medio indica que la predicción realizada por el modelo se acerca más al valor real. Esta medida se caracteriza por penalizar en mayor medida los errores de predicción muy altos. El ECMP no incrementa necesariamente con la variación de los errores sino con la varianza de la distribución de frecuencia de la magnitud de los errores.

El segundo criterio de evaluación utilizado es el error absoluto medio:

$$MAE = \frac{1}{n} \sum_{i=1}^n |z_i - \hat{z}_i|$$

Igual que el criterio anterior, valores cercanos a cero de error absoluto medio indica una mejor aproximación de la predicción al valor real. A diferencia del error cuadrático medio, todas las diferencias individuales se ponderan por igual y no es tan sensible a valores atípicos.

El último criterio a utilizar es el coeficiente de determinación o R^2 :

$$R^2 = 1 - \frac{\sum_{i=1}^n (z_i - \hat{z}_i)^2}{\sum_{i=1}^n (z_i - \bar{z})^2} = 1 - \frac{ECMP}{\sigma^2}$$

El coeficiente de determinación representa la proporción de varianza explicada por el modelo. Puede tomar valores entre 1 y ∞ . Valores menores o iguales a cero indican que el modelo no explica ninguna porción de la variabilidad de los datos.

De esta manera, se puede cuantificar la capacidad predictiva de los modelos y optar por el de mejor rendimiento.

4.1 Diario

A partir de los modelos desarrollados en el apartado 3.2 que incluyen métodos de suavizamiento exponencial y métodos ARIMA, se seleccionan cinco modelos que se comparan entre sí:

- Modelo Holt
- Modelo Holt Damped
- Modelo Holt-Winters
- Modelo Holt-Winters Damped
- Modelo ARIMA (1,1,1)

Se excluyen los modelos Holt-Winters multiplicativos al presentar un ajuste pobre.

La evaluación se realiza comparando el error cuadrático medio de los modelos, el error absoluto medio y el coeficiente de determinación para cada horizonte de predicción ($h = 10$).

En la Figura 63 se aprecia la comparativa del error cuadrático medio producido por los modelos en cada horizonte de predicción. El comportamiento de los modelos que utilizan el método Holt-Winters aditivo es similar. De igual forma, el comportamiento de los modelos que utilizan los métodos de suavizamiento exponencial doble es similar, siendo ligeramente mejor el modelo con tendencia amortiguada. Estos modelos de suavizamiento exponencial doble presentan un menor error cuadrático medio que los modelos de suavizamiento exponencial triple, lo que revela una baja relevancia del componente estacional para la serie de caudal medio diario. Esto se refuerza al observar la Figura 64 correspondiente al error absoluto medio en cada horizonte de predicción y la Figura 65 que muestra el R^2 en cada horizonte de predicción. Los modelos de suavizamiento exponencial doble presentan un menor valor MAE y un mayor valor R^2 respecto a los modelos de suavizamiento exponencial triple, indicando un mejor ajuste.

Además, es posible reconocer la superioridad del Modelo ARIMA (1,1,1) frente a los modelos de suavizamiento exponencial, alcanzando el menor valor de ECMP y MAE, así como el más alto R^2 en todos los horizontes de predicción. El modelo obtiene un valor mínimo de 1.0766×10^4 en el primero horizonte y el valor máximo de 3.6285×10^4 en el último horizonte de predicción para el ECMP. Con la métrica MAE obtiene un valor mínimo de 36.74 en el primer horizonte de predicción y el máximo de 74.35 en el último horizonte y con la evaluación del R^2 obtiene el valor máximo en el primer horizonte con 0.8697 y el mínimo en el último horizonte con 0.5608.

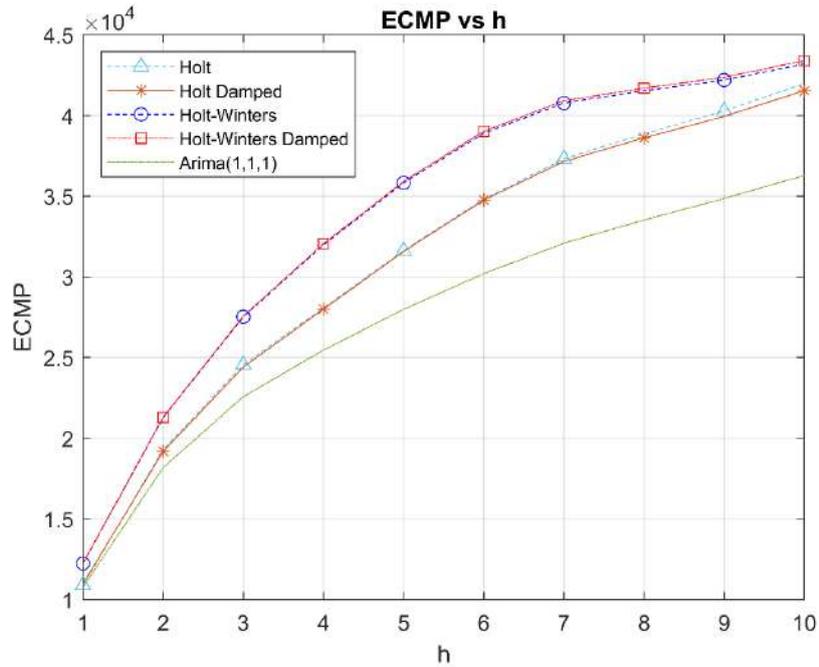


Figura 63. ECMP vs h - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

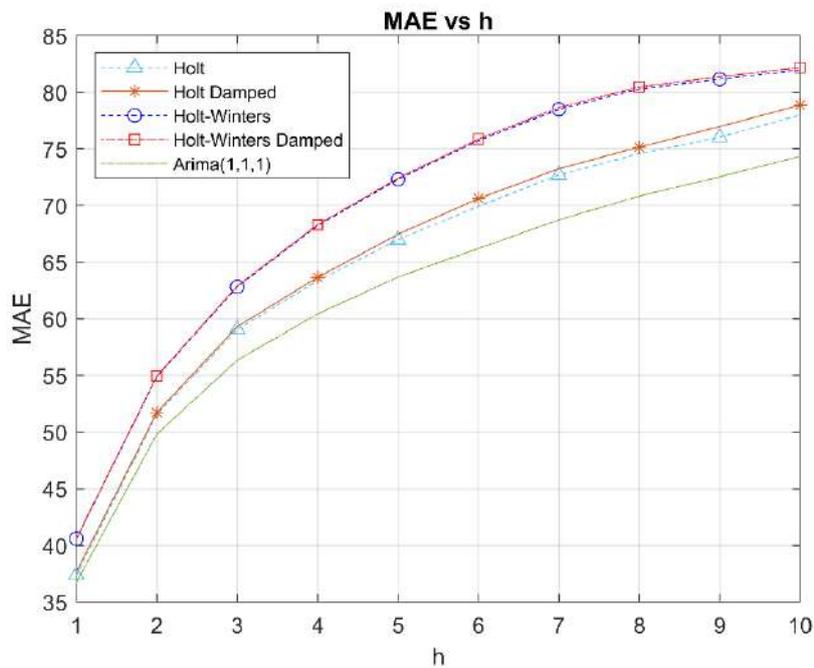


Figura 64. MAE vs h - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

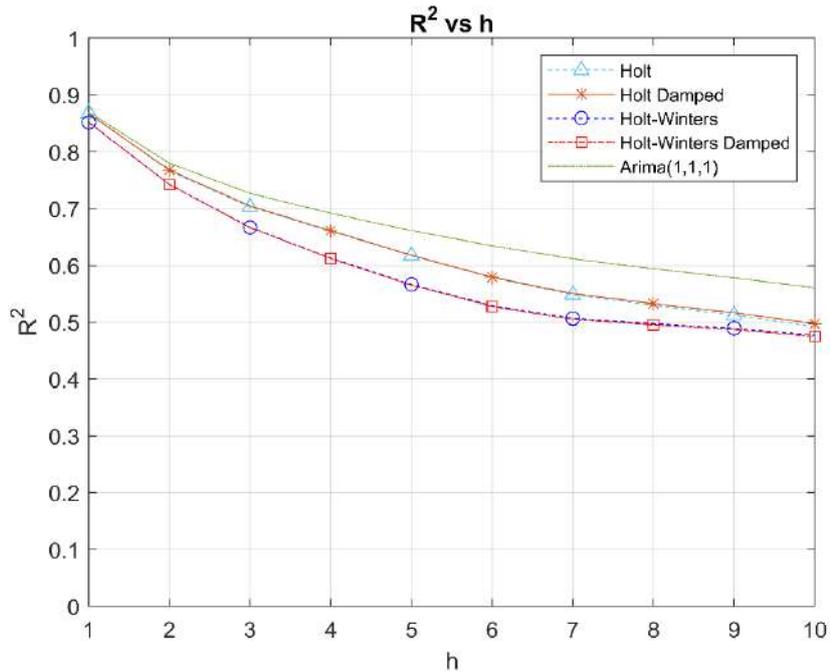


Figura 65. R^2 vs h - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

Tabla 22. Resumen ECMP vs h - Comparativa frecuencia diaria.

h	Holt	Holt Damped	Holt-Winters	Holt-Winters Damped	ARIMA (1,1,1)
1	10900.55	10977.95	12247.33	12253.44	10766.47
2	19293.72	19187.45	21281.22	21300.96	18191.64
3	24551.49	24408.56	27526.60	27567.12	22576.90
4	28082.60	28000.63	31987.00	32050.95	25462.18
5	31620.41	31576.37	35827.61	35916.31	27992.42
6	34826.06	34758.82	38908.16	39023.38	30200.49
7	37313.80	37155.93	40774.68	40913.97	32086.43
8	38865.59	38613.29	41535.11	41694.96	33525.79
9	40277.54	39946.28	42197.82	42375.59	34863.81
10	41981.31	41543.32	43197.19	43393.60	36285.38

Fuente: Elaboración propia.

Tabla 23. Resumen MAE vs h - Comparativa frecuencia diaria.

h	Holt	Holt Damped	Holt-Winters	Holt-Winters Damped	ARIMA (1,1,1)
1	37.3492	37.6757	40.5901	40.6151	36.7391
2	51.5637	51.7298	54.8944	54.9417	49.8228
3	59.0743	59.3121	62.8317	62.8964	56.3440
4	63.3644	63.6600	68.2291	68.3127	60.4354
5	66.9861	67.4679	72.3141	72.4231	63.6939
6	69.9365	70.6144	75.7339	75.8637	66.2135
7	72.6999	73.2520	78.4964	78.6528	68.7355
8	74.5842	75.1556	80.2865	80.4667	70.8350
9	76.0274	76.9688	81.1592	81.3662	72.5348
10	77.9985	78.8619	81.9551	82.1753	74.3484

Fuente: Elaboración propia.

Tabla 24. Resumen R² vs h - Comparativa frecuencia diaria.

h	Holt	Holt Damped	Holt-Winters	Holt-Winters Damped	ARIMA (1,1,1)
1	0.8681	0.8671	0.8518	0.8517	0.8697
2	0.7665	0.7678	0.7424	0.7422	0.7798
3	0.7029	0.7046	0.6669	0.6664	0.7268
4	0.6601	0.6611	0.6129	0.6121	0.6918
5	0.6173	0.6178	0.5664	0.5653	0.6612
6	0.5785	0.5793	0.5291	0.5277	0.6345
7	0.5484	0.5503	0.5065	0.5048	0.6117
8	0.5296	0.5327	0.4973	0.4954	0.5942
9	0.5125	0.5165	0.4893	0.4871	0.5780
10	0.4919	0.4972	0.4772	0.4748	0.5608

Fuente: Elaboración propia.

En el análisis de la serie del caudal medio diario, detallado en el apartado 3.1, se observa que la distribución de la serie varía según las estaciones en cada año (Figura 17). Por esta razón, se considera una evaluación de las predicciones, diferenciando la estación del año en la que se realizan. Se divide el año en trimestres, donde el primer trimestre, de enero a marzo, corresponde a la estación de verano, el segundo trimestre, de abril a junio, corresponde al otoño, el tercer trimestre, de julio a setiembre, corresponde al invierno y el último trimestre, de octubre a diciembre corresponde a la primavera.

Se analiza el error cuadrático medio, el error absoluto medio y el coeficiente de determinación por horizonte de predicción para cada uno de los trimestres antes detallados.

El primer trimestre es el más complicado de predecir, pues la variabilidad de la serie es más acentuada en el verano (Figura 17). En la Figura 66 se aprecia el error cuadrático medio, en la Figura 67 el error absoluto medio y en la Figura 68 el coeficiente de determinación. Los modelos de suavizamiento exponencial doble son superiores a los de suavizamiento exponencial triple, tal como se observa en todas las gráficas, resaltando el Modelo Holt Damped. Sin embargo, el menor rendimiento se obtiene con la utilización del Modelo ARIMA (1,1,1), con los valores de ECMP más bajos, desde 2.8804×10^4 hasta 9.1935×10^4 , los menores valores de MAE desde 76.60 hasta 147.63 y el mejor R^2 desde 0.8333 en el primer horizonte hasta 0.4681 en el último.

El segundo trimestre, que corresponde al periodo de otoño, se presenta desde la Figura 69 a la Figura 71 y muestra valores de error cuadrático medio y error absoluto medio por debajo del trimestre anterior. A diferencia de este último, el comportamiento de los modelos es bastante distinto. Los modelos de suavizamiento exponencial doble presentan menor error cuadrático medio y error absoluto medio que los de suavizamiento exponencial triple, aunque esto varía en el último horizonte debido al crecimiento aproximadamente lineal de los errores. El Modelo ARIMA (1,1,1) mantiene el mejor rendimiento de todos los modelos. El error cuadrático medio de este modelo alcanza un máximo de 5.0483×10^4 en el último horizonte de predicción, siendo el más bajo en todos los horizontes de predicción. El error absoluto medio es el más bajo desde el segundo al noveno horizonte de predicción. Es superado ligeramente por el Modelo Holt en el primer horizonte y el Modelo Holt-Winters en el último. El modelo ARIMA presenta el mayor R^2 en todos los horizontes de predicción, con un máximo de 0.8784 en el primer horizonte y un mínimo de 0.5324 en el último.

El tercer trimestre, que responde al periodo de invierno, se muestra en la Figura 72 a la Figura 74 y presenta los valores más bajos de error cuadrático medio y error absoluto medio. Los modelos que utilizan el suavizamiento exponencial triple mantienen los valores más altos de ECMP y MAE. Dentro de los modelos de suavizamiento exponencial destaca el Modelo Holt, no obstante, el modelo que utiliza la metodología ARIMA supera con amplitud su rendimiento. El Modelo ARIMA (1,1,1) muestra el valor más bajo de ECMP en el primer horizonte de predicción: 337, y el valor máximo de 809.38, los menores valores de MAE con 11.38 en el primer horizonte y 18.21 en el último y el mayor R^2 con un valor máximo de 0.7134 en el primer horizonte de predicción y un mínimo de 0.3117 en el último.

El último trimestre, correspondiente a la estación primavera, se presenta en la Figura 75 a Figura 77. El error cuadrático medio y el error absoluto medio se incrementa en gran medida para los modelos de suavizamiento exponencial triple. Se obtiene los mejores rendimientos con los modelos de suavizamiento exponencial doble y el modelo ARIMA. En la gráfica se aprecia una disminución del ECMP en los últimos horizontes de predicción para los

modelos de suavizamiento exponencial responde a la variabilidad muestral. La temporada de primavera presenta valores atípicos producto de los años en los que sucede el fenómeno del niño y esto afecta la variabilidad de la muestra (Figura 78). Indicadores como el error cuadrático medio y el coeficiente de determinación se ven afectados por esta variabilidad, lo cual se refleja en la Figura 75 y Figura 77 respectivamente. El error absoluto medio, al ponderar por igual todas las diferencias individuales no se afecta gran medida. En la Figura 76 se observa que el Modelo ARIMA (1,1,1) presenta el menor error en todos los horizontes de predicción, con un mínimo de 10.53 en el primer horizonte y un máximo de 18.32 en el último horizonte.

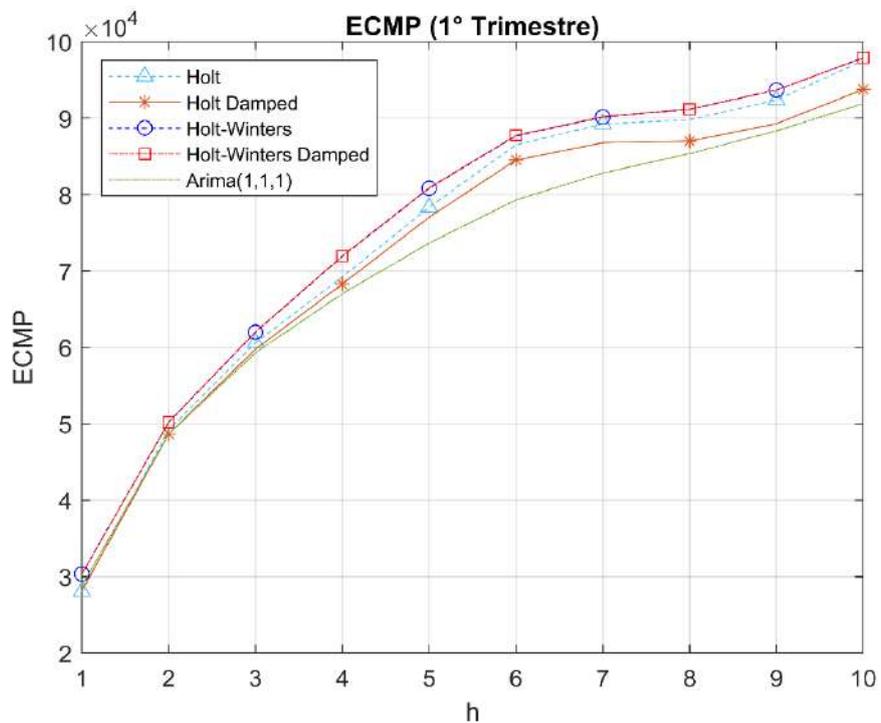


Figura 66. ECMP vs h (1° Trimestre) - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

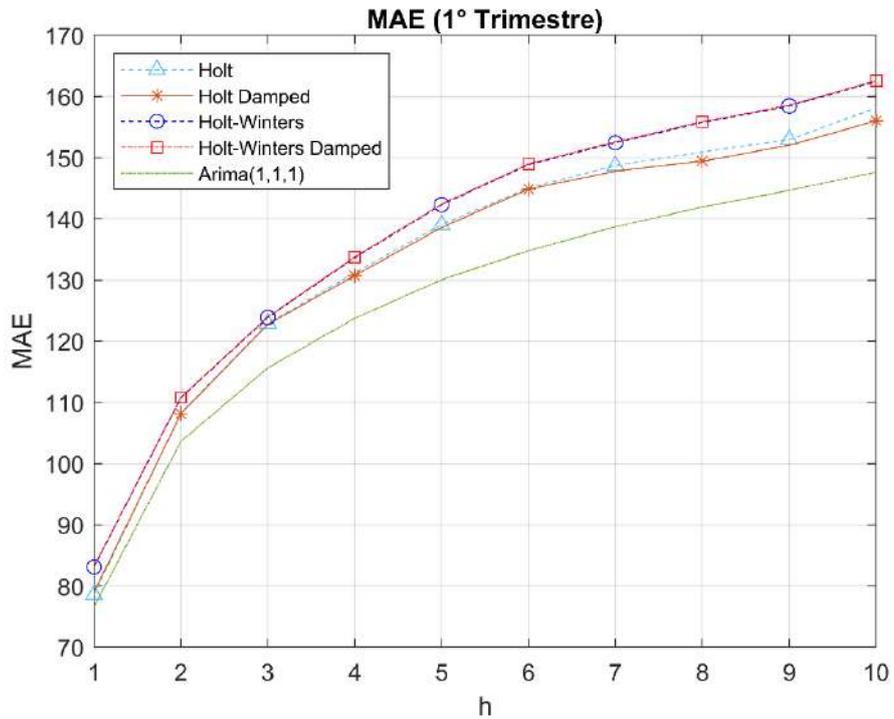


Figura 67. MAE vs h (1° Trimestre) - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

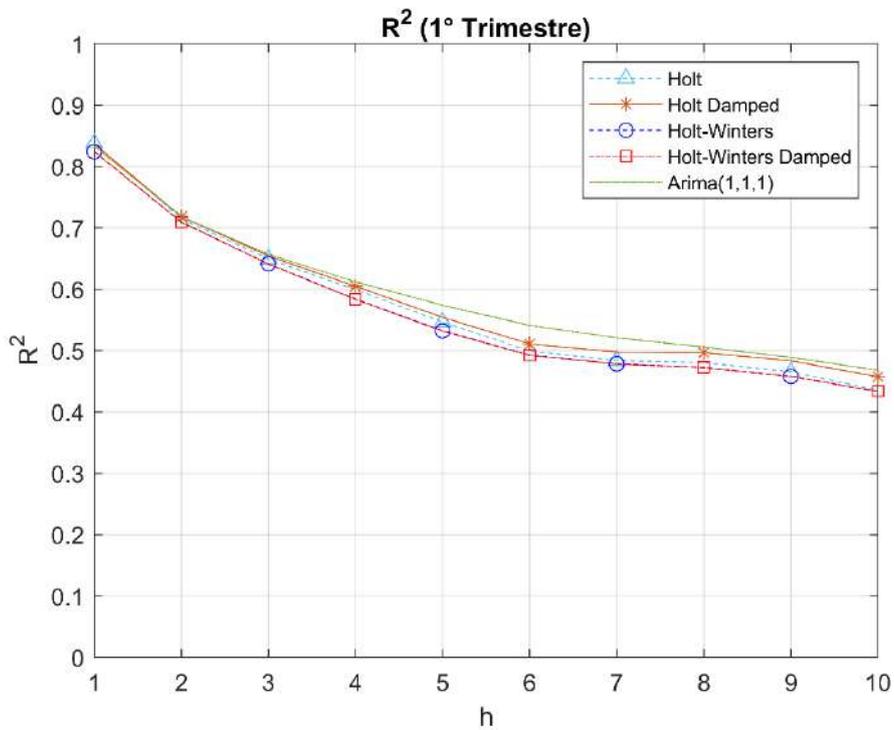


Figura 68. R² vs h (1° Trimestre) - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

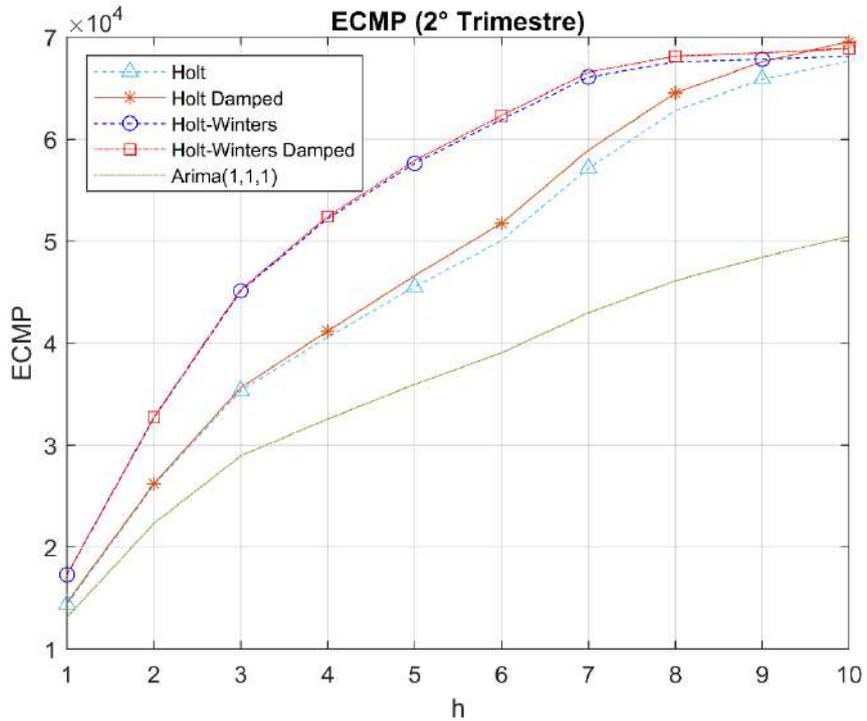


Figura 69. ECMP vs h (2° Trimestre) - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

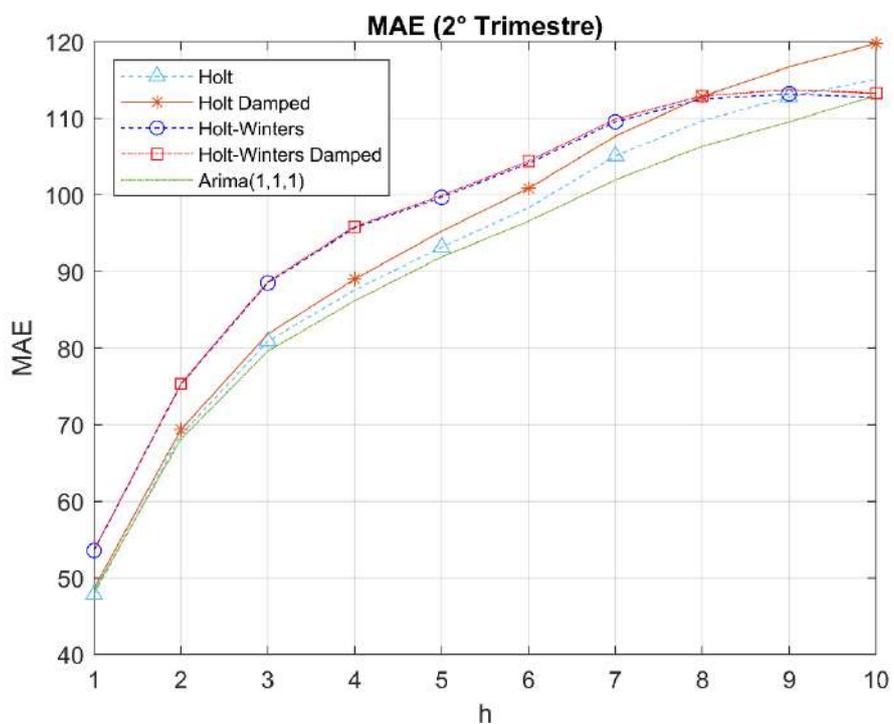


Figura 70. MAE vs h (2° Trimestre) - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

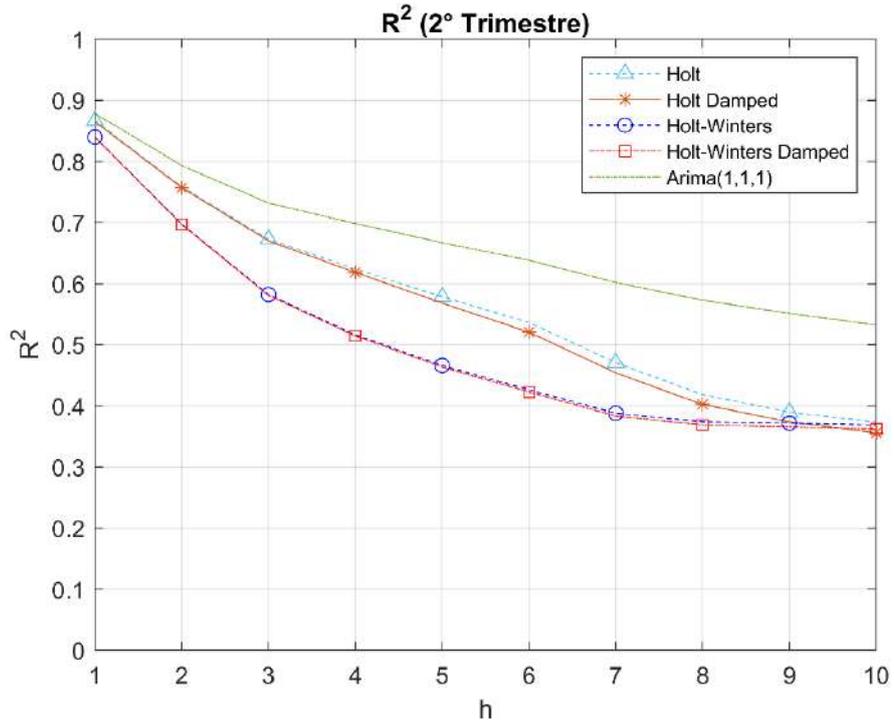


Figura 71. R² vs h (2° Trimestre) - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

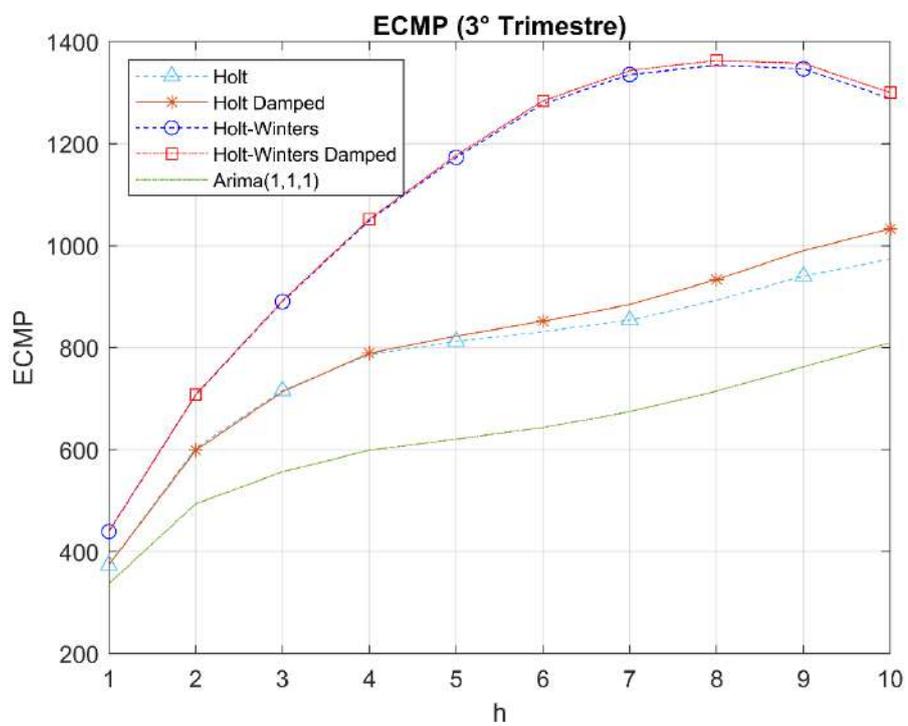


Figura 72. ECMP vs h (3° Trimestre) - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

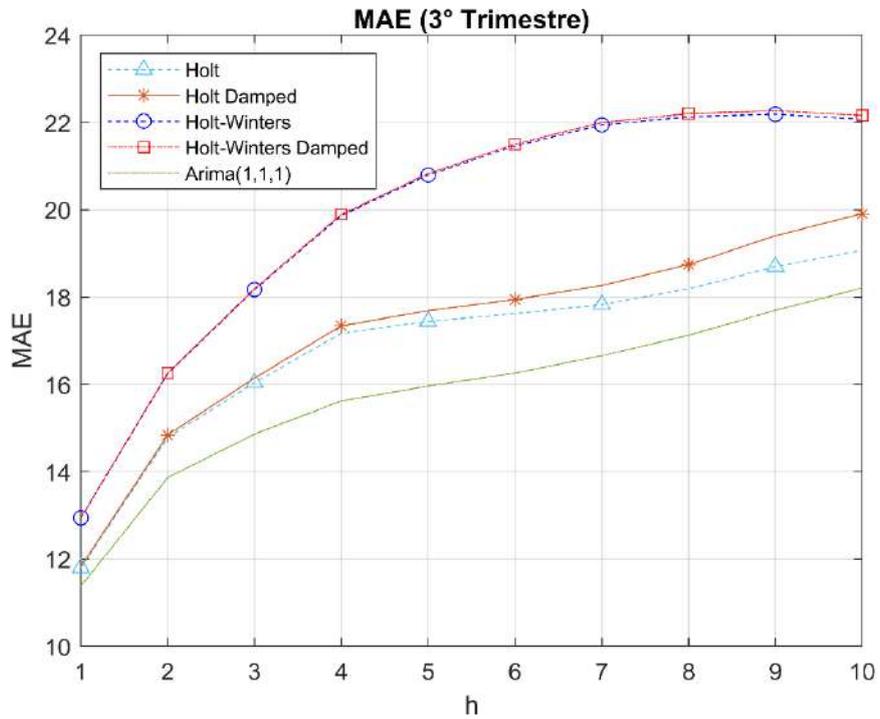


Figura 73. MAE vs h (3° Trimestre) - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

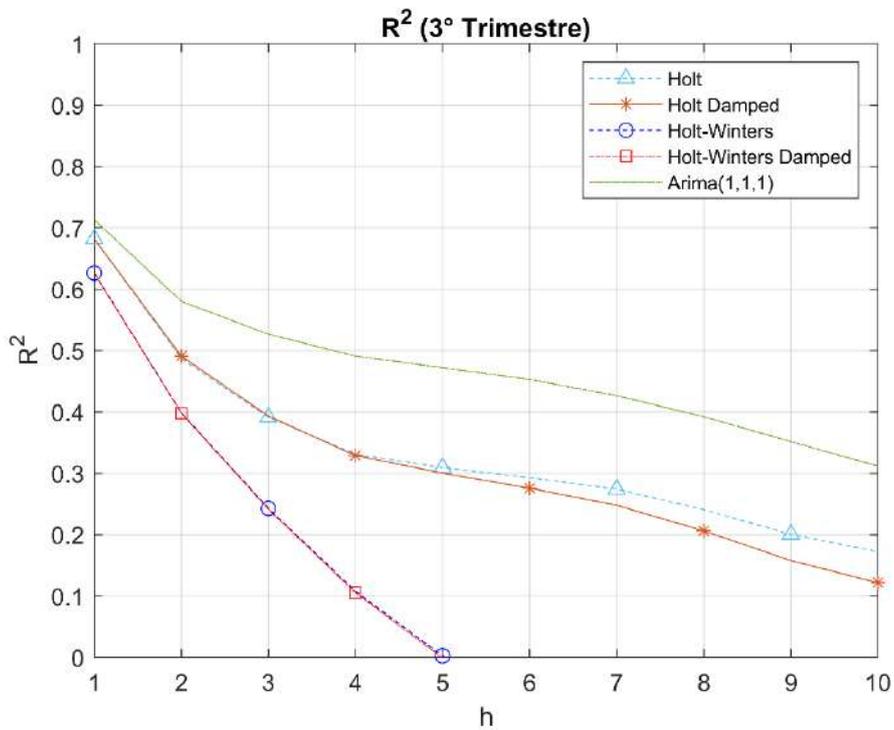


Figura 74. R² vs h (3° Trimestre) - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

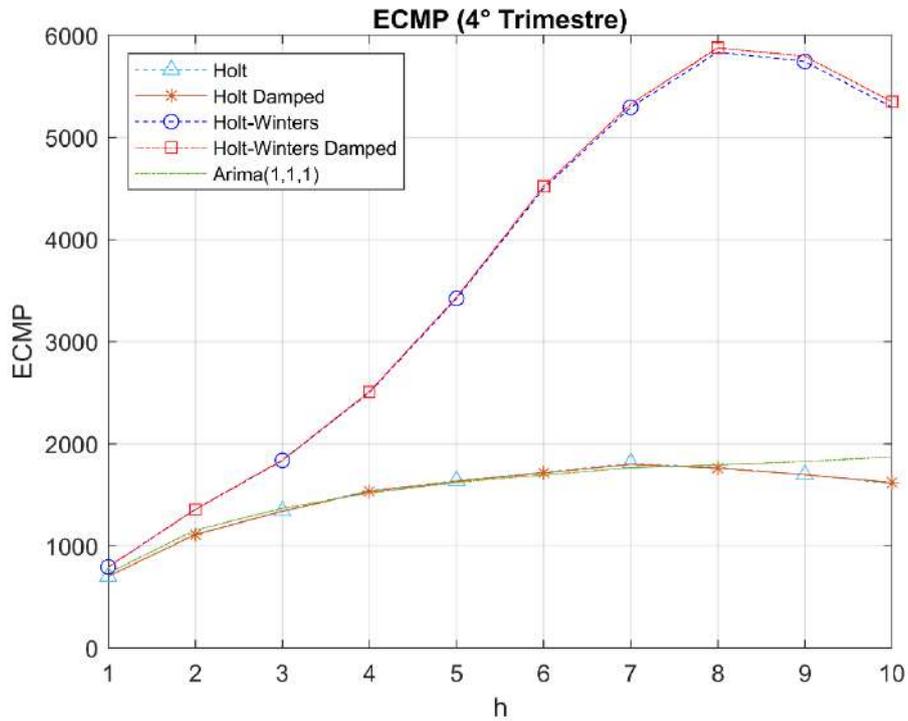


Figura 75. ECMP vs h (4° Trimestre) - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

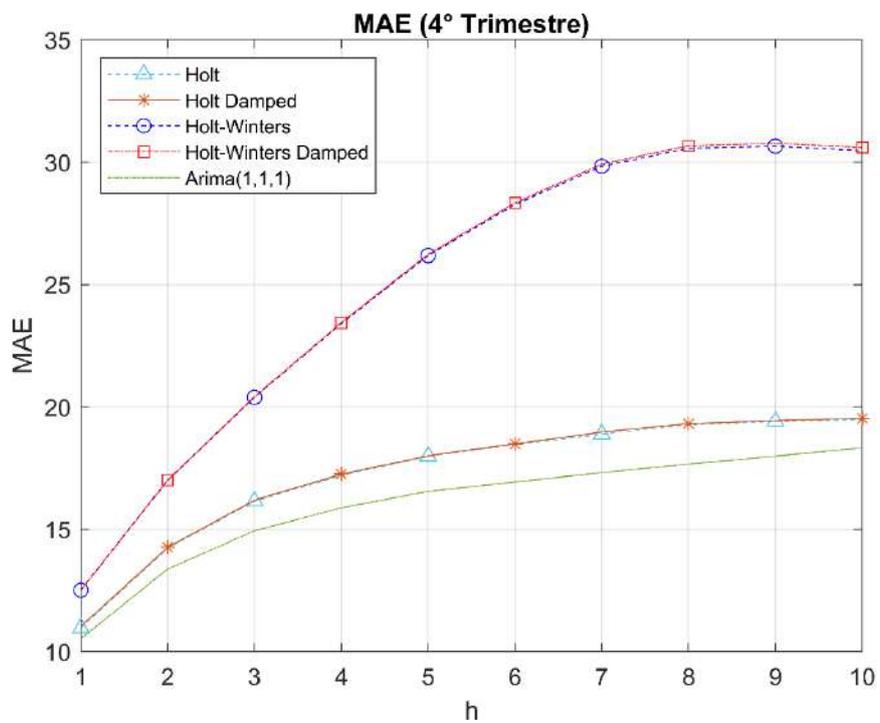


Figura 76. MAE vs h (4° Trimestre) - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

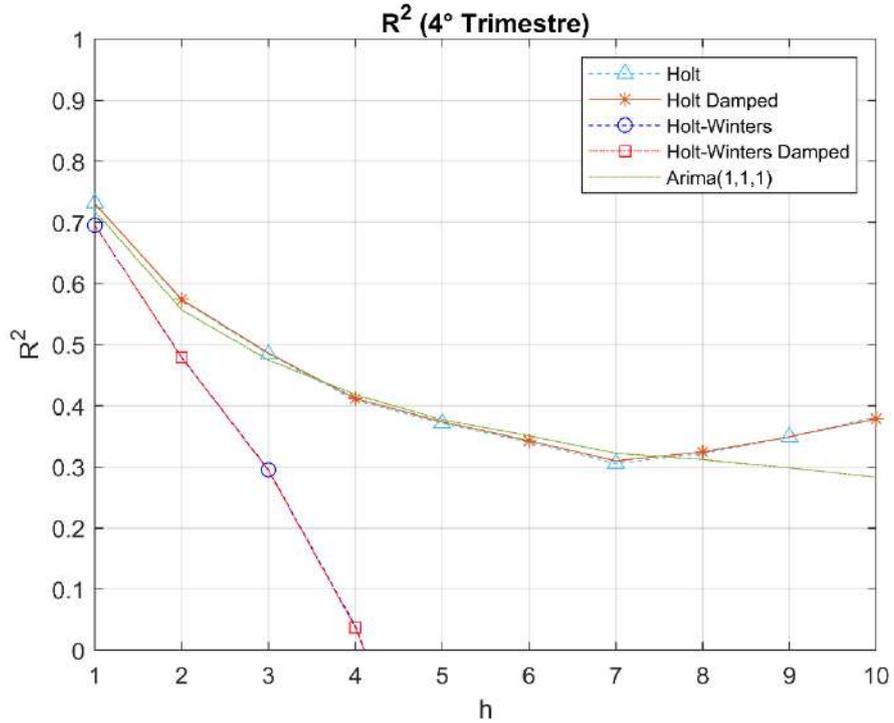


Figura 77. R² vs h (4° Trimestre) - Comparativa frecuencia diaria.

Fuente: Elaboración propia.

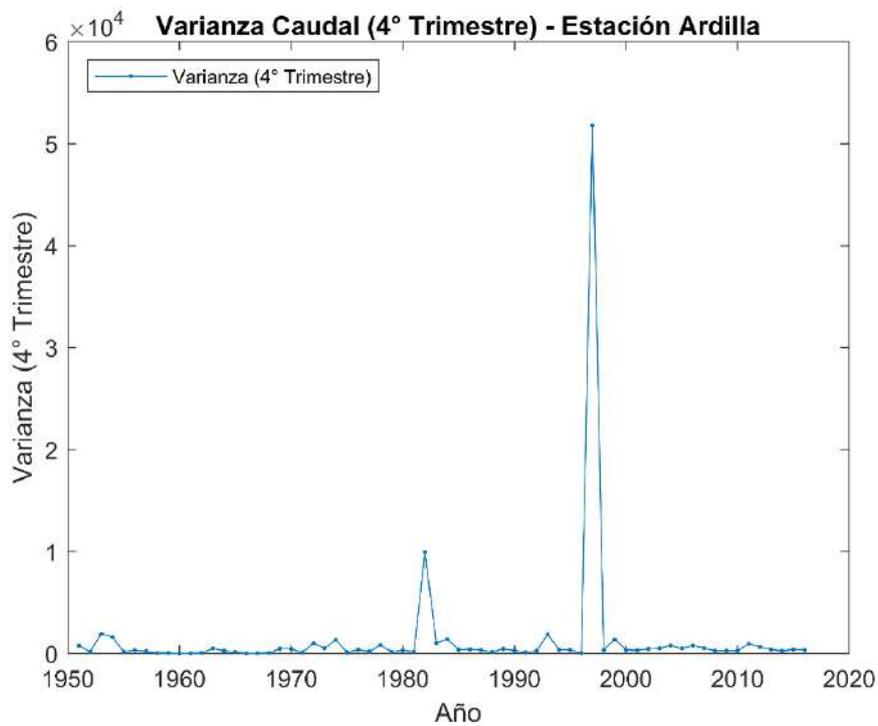


Figura 78. Varianza Caudal (4° Trimestre) - Estación Ardilla.

Fuente: Elaboración propia.

Analizados los resultados por trimestre, es interesante también evaluar la variación en el rendimiento de las predicciones respecto al valor del caudal medido. En la Figura 14, se representa en un histograma la distribución de la serie de caudal medio diario. La mayor frecuencia de caudal se encuentra en el intervalo de 0 a 500 m³/s. Por consiguiente, se evaluará la evolución del error cuadrático medio y el error absoluto medio respecto del caudal real, en el intervalo de mayor frecuencia de la serie. El análisis se realiza diferenciando el rendimiento para cada horizonte de predicción.

En la Figura 79 se presenta el error cuadrático medio respecto del caudal en el primer horizonte de predicción. Es evidente que el rendimiento disminuye a medida que el valor del caudal aumenta. En un primer intervalo de 0 a 100 m³/s, el comportamiento de los modelos es similar, con un ECMP bajo. Las diferencias se acrecientan a mayores valores de caudal, donde los modelos de suavizamiento triple cuentan con mayor error de todos. El Modelo ARIMA (1,1,1) obtiene el mejor rendimiento en el primer horizonte de predicción, seguido del Modelo Holt Damped como el mejor de los modelos basados en suavizamiento exponencial. En la Figura 80 se muestra el error absoluto medio respecto del caudal. Dentro del intervalo de 0 a 100 m³/s los modelos de suavizamiento exponencial doble junto al modelo ARIMA presentan los menores valores de MAE. Sin embargo, con el aumento en los valores de caudal los métodos de suavizamiento exponencial doble superan al modelo ARIMA. Esta diferencia entre las dos métricas empleadas para la evaluación significa que los errores producidos por los modelos de suavizamiento exponencial doble son de mayor magnitud que los producidos por el modelo ARIMA, debido que son penalizados en mayor medida por el error cuadrático medio.

En la Figura 81 se observa la misma gráfica, esta vez en el quinto horizonte de predicción. Al compararlo con la Figura 79, se observa un aumento en los valores del ECMP. Este mismo aumento se detecta en el error absoluto medio, al comparar la Figura 80 y la Figura 82. Se aprecia que el Modelo ARIMA (1,1,1) consigue el menor error cuadrático medio a medida que aumenta el valor del caudal, haciéndose más notoria la diferencia al compararlo con los modelos de suavizamiento exponencial. Al observar el error absoluto medio en la Figura 82 se aprecia un comportamiento similar entre los modelos, obteniendo el menor MAE con los modelos de suavizamiento exponencial doble y el modelo ARIMA en el primer tramo del intervalo en evaluación y el modelo ARIMA en el último tramo.

Al observar la Figura 83, correspondiente al décimo horizonte de predicción, se nota el aumento en la escala del ECMP. El comportamiento se repite para el error absoluto medio, presente en la Figura 84. El error cuadrático medio aumenta en función del valor del caudal y el horizonte de predicción. En esta última gráfica se distingue de mejor manera la diferencia de rendimiento entre los modelos de suavizamiento exponencial y el modelo ARIMA, siendo este último el que consigue el mejor rendimiento ante el aumento del valor del caudal. El error absoluto medio mostrado en la Figura 84 presenta comportamientos similares de todos los

métodos. Al evaluar los dos indicadores se deduce que, a pesar de presentar en promedio errores similares, el modelo ARIMA produce errores de menor magnitud que los otros modelos.

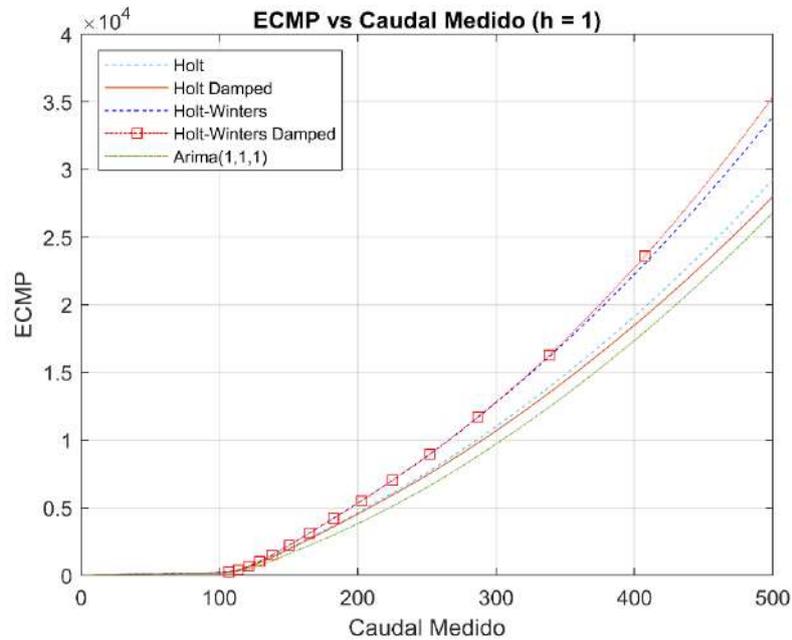


Figura 79. ECMP vs Caudal Medido - Primer horizonte de predicción. Frecuencia diaria.

Fuente: Elaboración propia.

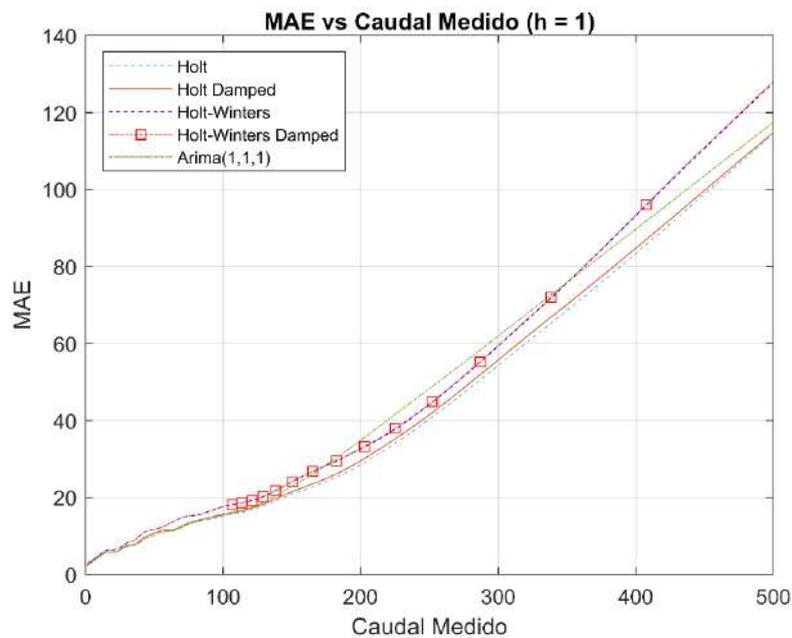


Figura 80. MAE vs Caudal Medido - Primer horizonte de predicción. Frecuencia diaria.

Fuente: Elaboración propia.

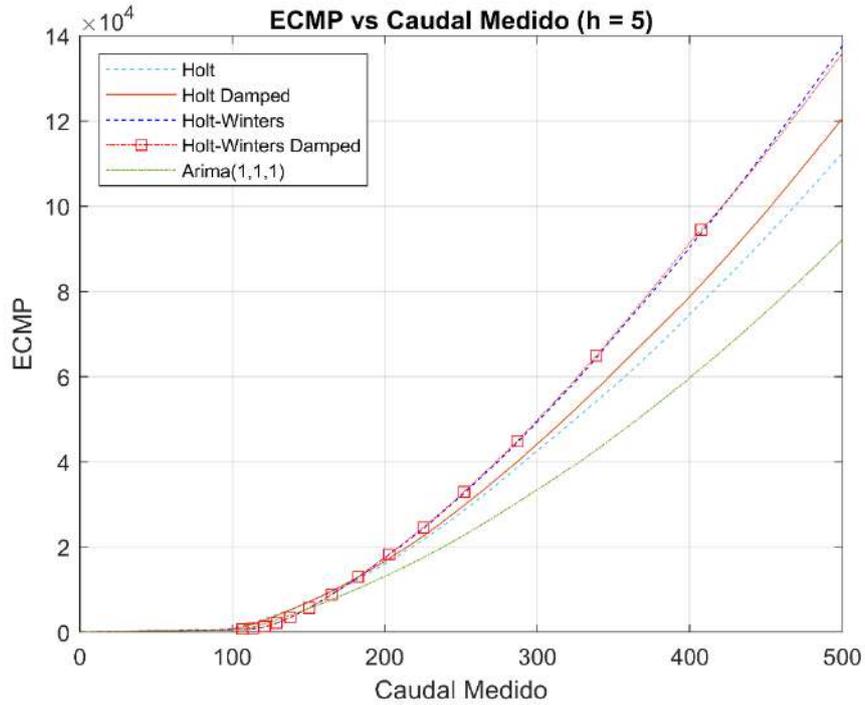


Figura 81. ECMP vs Caudal Medido - Quinto horizonte de predicción. Frecuencia diaria.

Fuente: Elaboración propia.

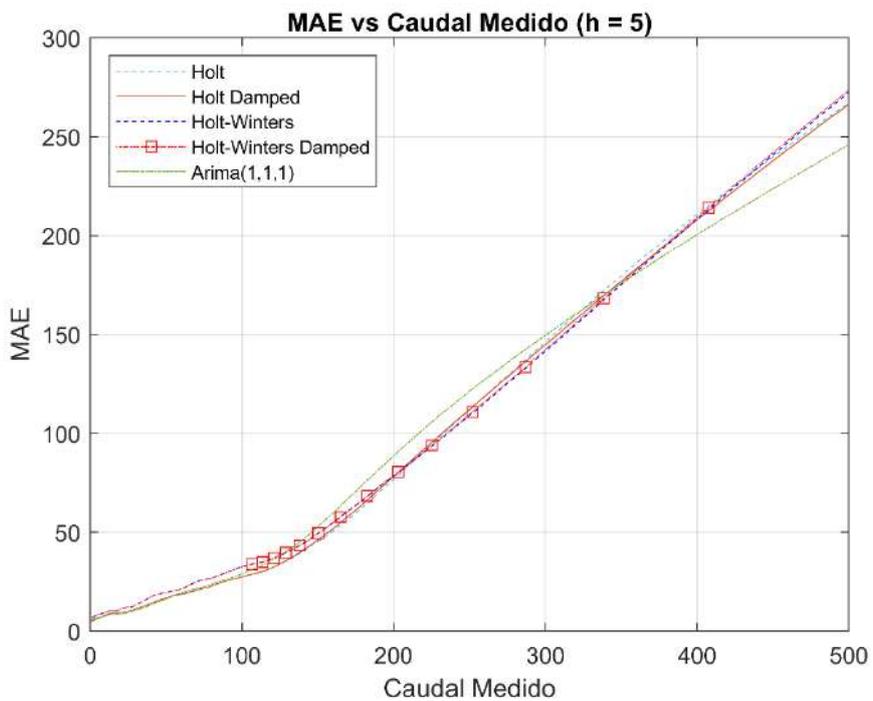


Figura 82. MAE vs Caudal Medido - Quinto horizonte de predicción. Frecuencia diaria.

Fuente: Elaboración propia.

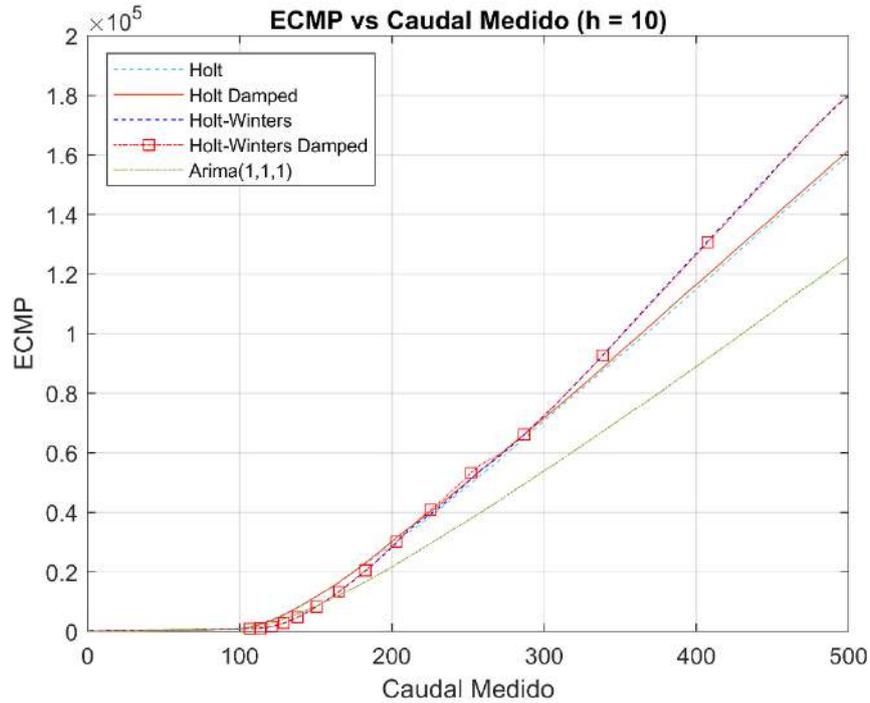


Figura 83. ECMP vs Caudal Medido - Décimo horizonte de predicción. Frecuencia diaria.

Fuente: Elaboración propia.

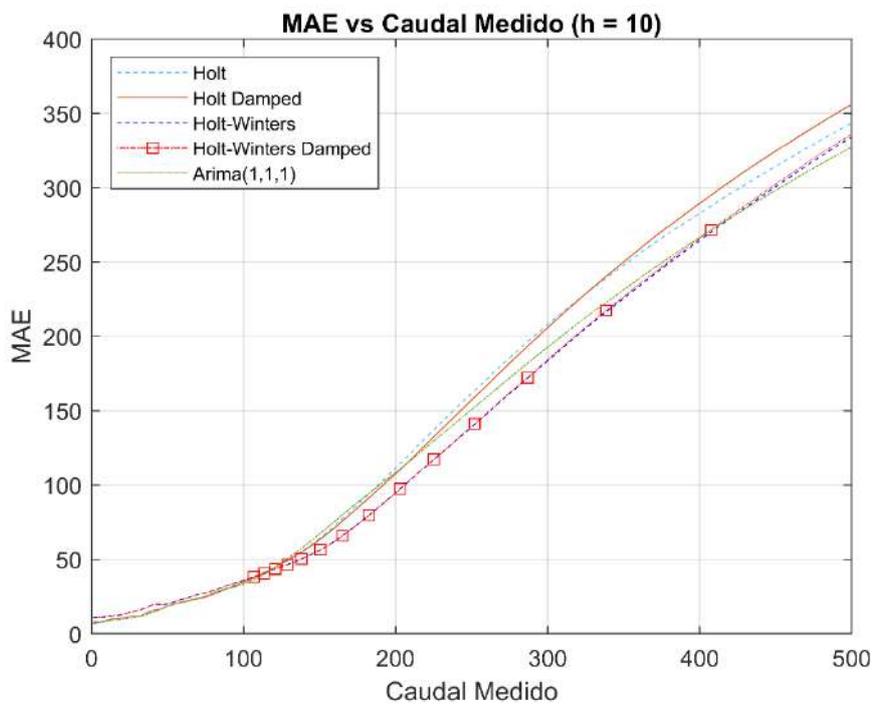


Figura 84. MAE vs Caudal Medido - Décimo horizonte de predicción. Frecuencia diaria.

Fuente: Elaboración propia.

En definitiva, el Modelo ARIMA (1,1,1) presenta los mejores resultados de predicción, dentro del abanico propuesto en el apartado 3.2.

4.2 Semanal

De los modelos presentados en el apartado 3.3, desarrollados en base a la serie temporal de los caudales promedio con frecuencia diaria. Se seleccionan cinco modelos para realizar la comparativa:

- Modelo Holt
- Modelo Holt Damped
- Modelo Holt-Winters
- Modelo Holt-Winters Damped
- Modelo ARIMA (1,1,1)(0,1,1)₅₂

Se excluyen los modelos Holt-Winters multiplicativos por el mal ajuste. Tampoco se incluye el Modelo ARIMA (1,1,1) y el Modelo ARIMA (2,1,0) debido a la existencia de correlación en la FAS y FAP de los residuos (Figura 54 a Figura 57).

La evaluación se realiza comparando el error cuadrático medio, el error absoluto medio y el coeficiente de determinación producido por los modelos seleccionados en cada horizonte de predicción ($h = 10$).

Esta comparativa se observa en la Figura 85 a la Figura 87, donde el Modelo Holt Damped presenta un menor error cuadrático medio que los métodos de suavizamiento triple hasta el quinto horizonte de predicción. Este modelo es superado en los horizontes siguientes, siendo el Modelo Holt-Winters Damped el de mejor rendimiento dentro de los modelos de suavizamiento exponencial, a partir del sexto horizonte de predicción. Al revisar el error absoluto simple se observa que el modelo Holt Damped presenta los valores más bajo en todos los horizontes de predicción. Sin embargo, el Modelo ARIMA (1,1,1)(0,1,1)₅₂ es el método que consigue el menor error cuadrático medio en cada horizonte de predicción, alcanzando un valor mínimo de 1.7875×10^4 en el primer horizonte y un máximo de 4.4557×10^4 en el último horizonte y también el menor error absoluto medio, con un mínimo de 51.65 en el primer horizonte y un máximo de 89.74 en el último horizonte. El coeficiente de dispersión que se aprecia en la Figura 87 refuerza la afirmación, evidenciando que el modelo ARIMA explica en mayor proporción la varianza de la serie. En el primer horizonte obtiene un valor R^2 de 0.7568 y en el último horizonte un valor de 0.3937.

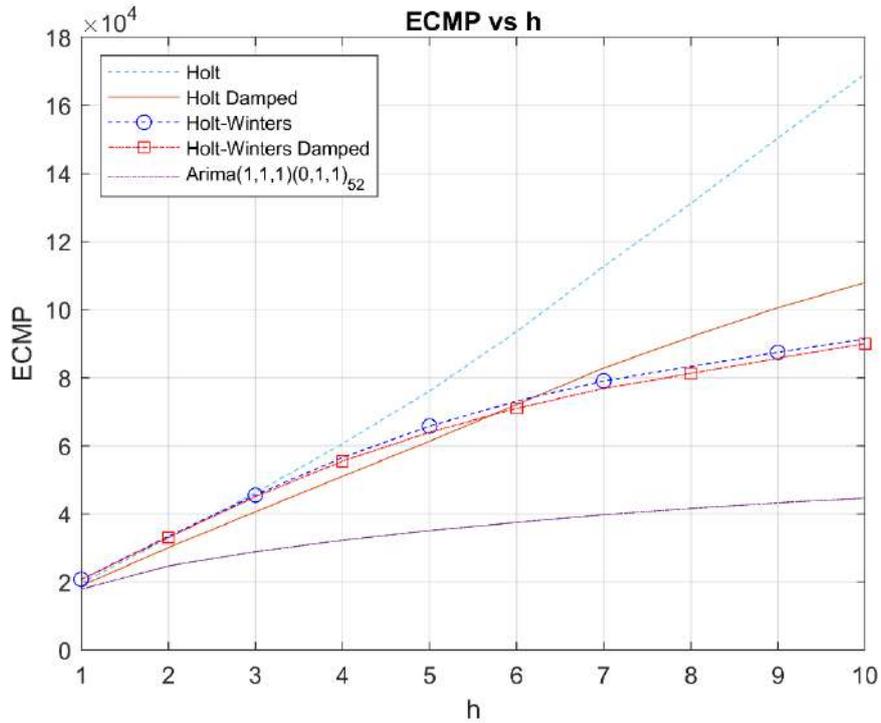


Figura 85. ECMP vs h - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

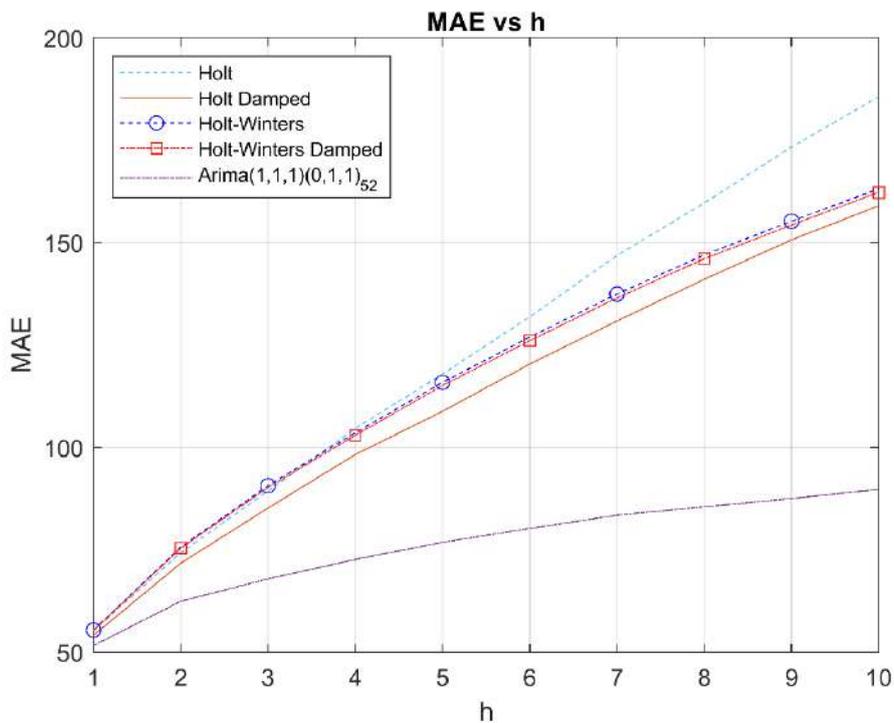


Figura 86. MAE vs h - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

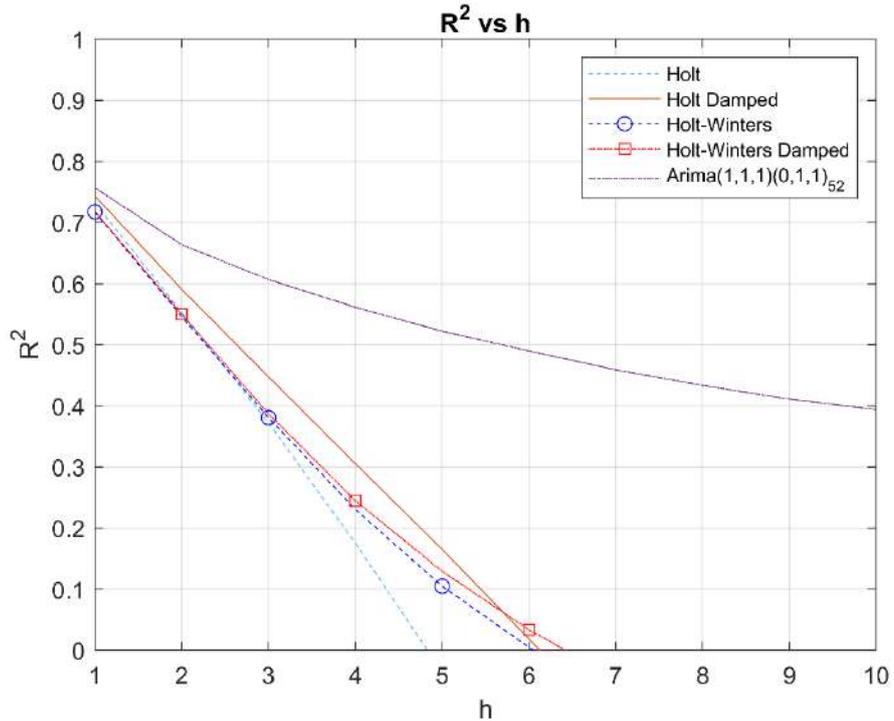


Figura 87. R² vs h- Comparativa frecuencia semanal.

Fuente: Elaboración propia.

Tabla 25. Resumen ECMP vs h - Comparativa frecuencia semanal.

h	Holt	Holt Damped	Holt-Winters	Holt-Winters Damped	ARIMA (1,1,1) (0,1,1) ₅₂
1	19688.41	18851.90	20759.93	20707.62	17875.03
2	32848.80	30093.13	33361.71	33096.80	24720.98
3	46150.89	40616.55	45513.43	45054.52	28902.69
4	60599.78	51021.94	56549.42	55496.03	32238.16
5	76071.48	61346.49	65767.70	63978.08	35118.26
6	93550.12	72115.63	73125.20	71012.40	37487.97
7	112699.62	82840.60	79054.31	76846.36	39754.37
8	131148.64	91988.36	83290.87	81242.83	41630.36
9	150444.91	100576.72	87465.09	85698.60	43279.44
10	169185.86	107894.13	91453.02	90038.67	44557.31

Fuente: Elaboración propia.

Tabla 26. Resumen MAE vs h - Comparativa frecuencia semanal.

h	Holt	Holt Damped	Holt-Winters	Holt-Winters Damped	ARIMA (1,1,1) (0,1,1) ₅₂
1	55.2233	54.1554	55.3910	55.3458	51.6489
2	74.3261	71.7545	75.6921	75.4301	62.4649
3	89.3740	85.2170	90.6623	90.2352	67.9260
4	104.6387	98.2748	103.5924	102.9726	72.6603
5	117.9188	108.8144	115.8814	115.1264	76.8625
6	131.8919	120.3661	126.9431	126.0533	80.2412
7	146.9014	130.8847	137.4581	136.5221	83.4741
8	159.6572	141.0557	147.0287	146.0804	85.5321
9	173.3506	150.6969	155.2323	154.3075	87.5567
10	185.6160	158.9967	163.1328	162.3060	89.7392

Fuente: Elaboración propia.

Tabla 27. Resumen R² vs h - Comparativa frecuencia semanal.

h	Holt	Holt Damped	Holt-Winters	Holt-Winters Damped	ARIMA (1,1,1) (0,1,1) ₅₂
1	0.7322	0.7435	0.7176	0.7183	0.7568
2	0.5531	0.5906	0.5461	0.5497	0.6637
3	0.3721	0.4474	0.3807	0.3870	0.6067
4	0.1754	0.3057	0.2305	0.2448	0.5613
5	-0.0352	0.1652	0.1050	0.1294	0.5221
6	-0.2731	0.0186	0.0049	0.0336	0.4899
7	-0.5337	-0.1273	-0.0758	-0.0458	0.4590
8	-0.7847	-0.2518	-0.1335	-0.1056	0.4335
9	-1.0473	-0.3687	-0.1902	-0.1662	0.4111
10	-1.3023	-0.4682	-0.2445	-0.2252	0.3937

Fuente: Elaboración propia.

Como se detalló en el apartado 4.1, la distribución de la serie de caudales varía en función de la estación del año. Es relevante la evaluación de la capacidad predictiva de los modelos diferenciando la estación del año en la cual se predice. El año se divide en trimestres, donde cada uno representa una estación del año.

El primer trimestre, que alude a la estación de verano, se representa desde la Figura 88 hasta la Figura 90. Los modelos de suavizamiento exponencial triple cuentan con el error cuadrático medio más alto en todos los horizontes de predicción, mismo comportamiento que se repite en el error absoluto medio y el coeficiente de determinación. Los modelos de

suavizamiento exponencial doble presentan el mejor ajuste en los dos primeros horizontes de predicción. Sin embargo, con el aumento del horizonte se reduce su capacidad de predicción, mientras que el error cuadrático medio y el error absoluto medio del Modelo ARIMA (1,1,1)(0,1,1)₅₂ se mantiene como el más bajo de la comparativa a partir del tercer horizonte y alcanza un valor ECMP máximo de 1.3110×10^5 en el último horizonte y un máximo MAE de 209.61. Además, el modelo cuenta con el coeficiente de determinación más alto en el último horizonte de predicción, con un valor de 0.1745.

En la Figura 91 a la Figura 93 se presenta el segundo trimestre, correspondiente al periodo de otoño. Los modelos de suavizamiento exponencial doble alcanzan los valores más altos de error cuadrático medio y error absoluto medio en todos los horizontes de predicción. El Modelo Holt-Winters Damped se presenta como el mejor de los modelos de suavizamiento exponencial propuestos, con los valores más bajos de ECMP y MAE. No obstante, el modelo que utiliza la metodología Box-Jenkins es superior en todos los horizontes de predicción, con valores de ECMP desde 1.6025×10^4 hasta 4.2775×10^4 , valores de MAE desde 59.43 hasta 105.21 y el mayor R^2 , desde 0.8113 hasta 0.4963.

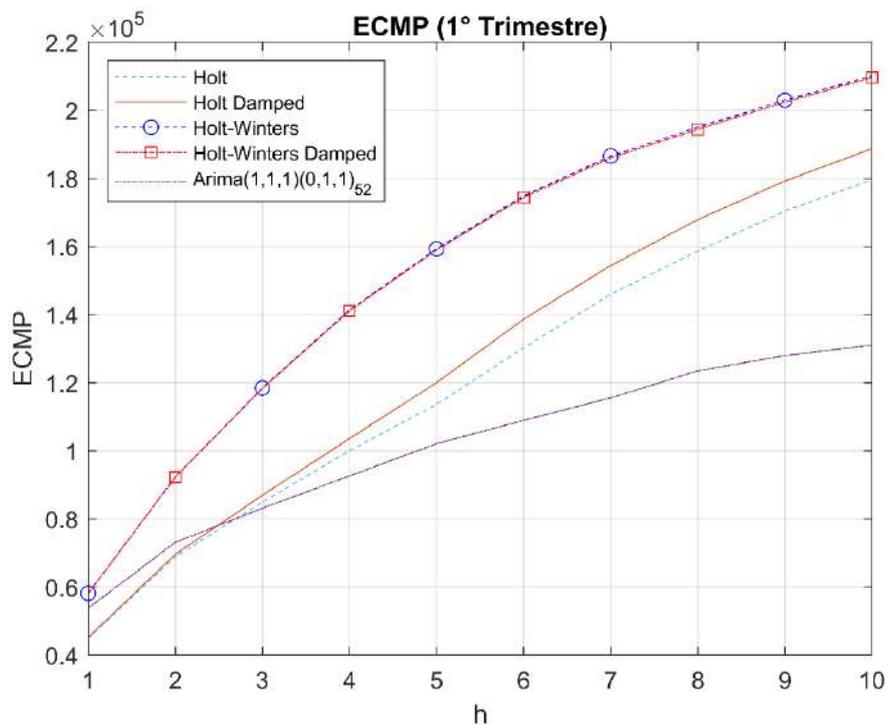


Figura 88. ECMP vs h (1° Trimestre) - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

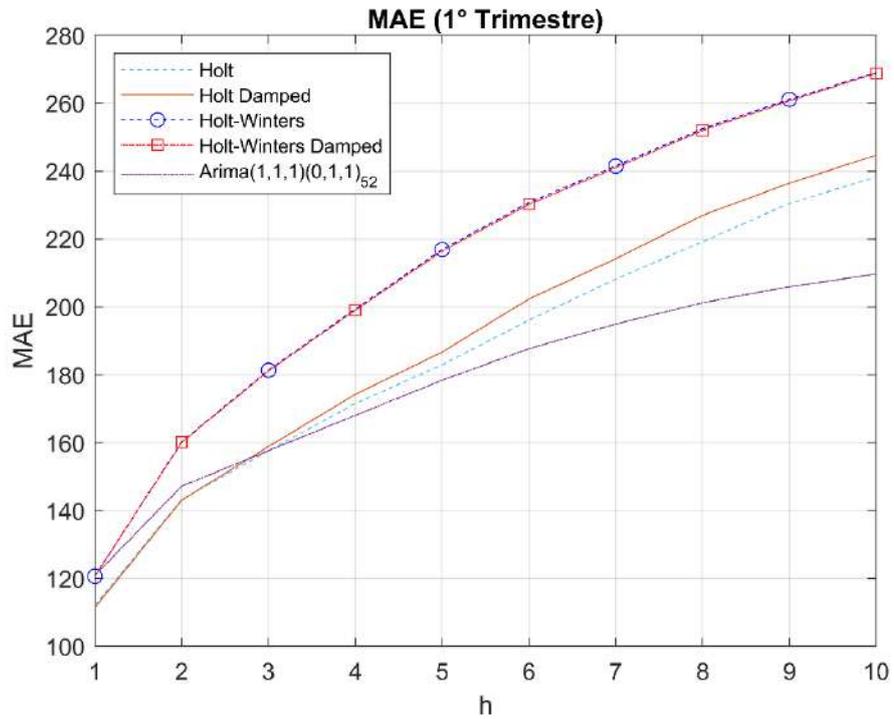


Figura 89. MAE vs h (1° Trimestre) - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

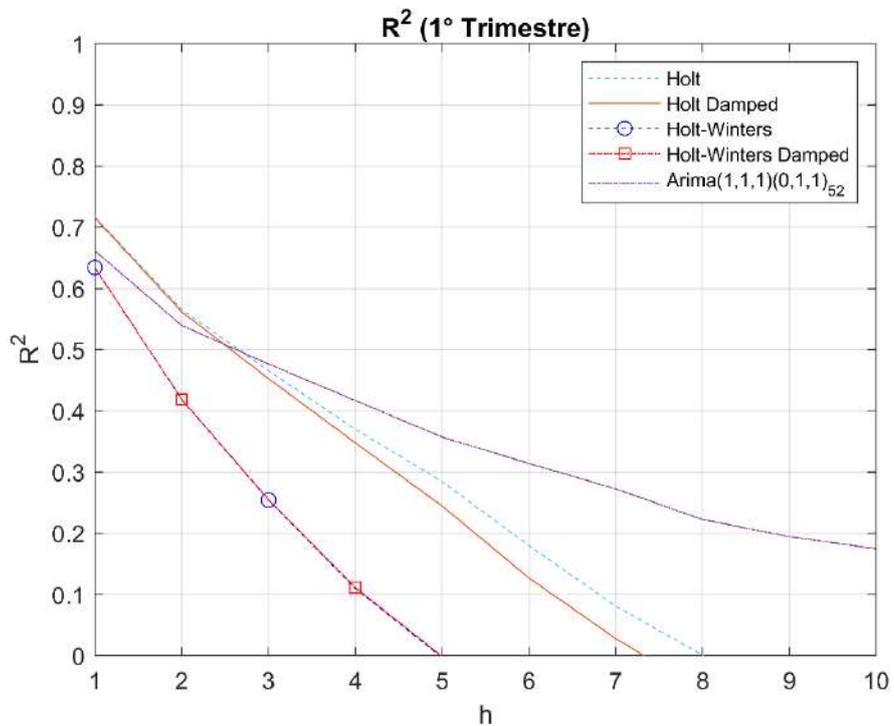


Figura 90. R² vs h (1° Trimestre) - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

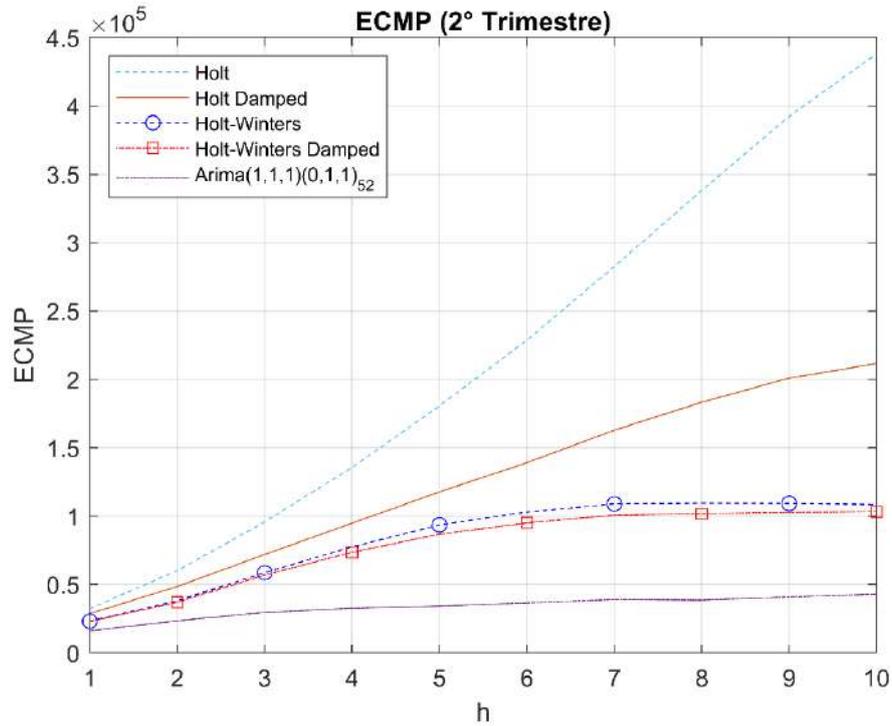


Figura 91. ECMP vs h (2° Trimestre) - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

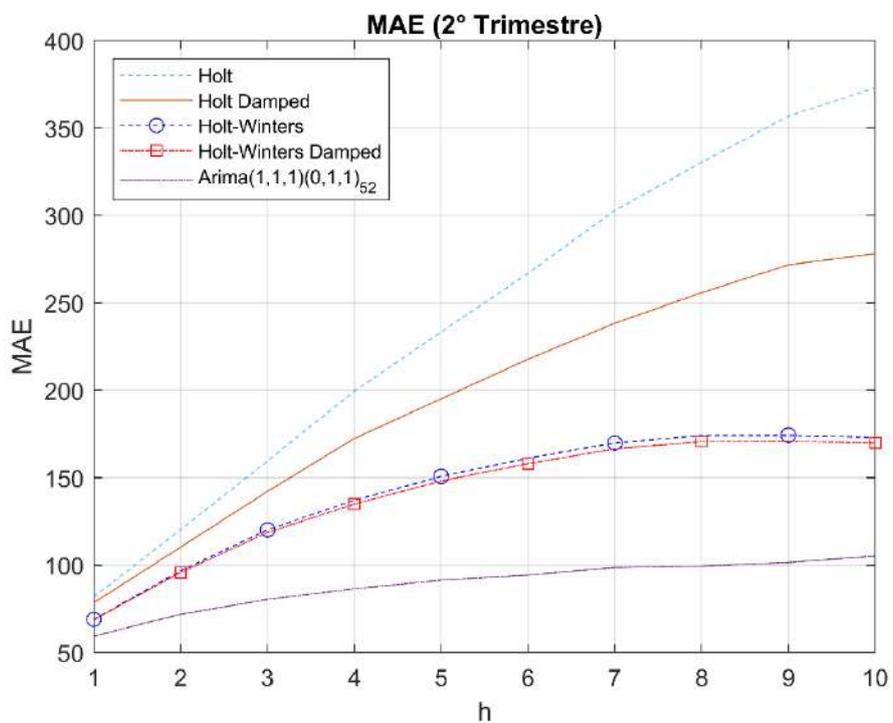


Figura 92. MAE vs h (2° Trimestre) - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

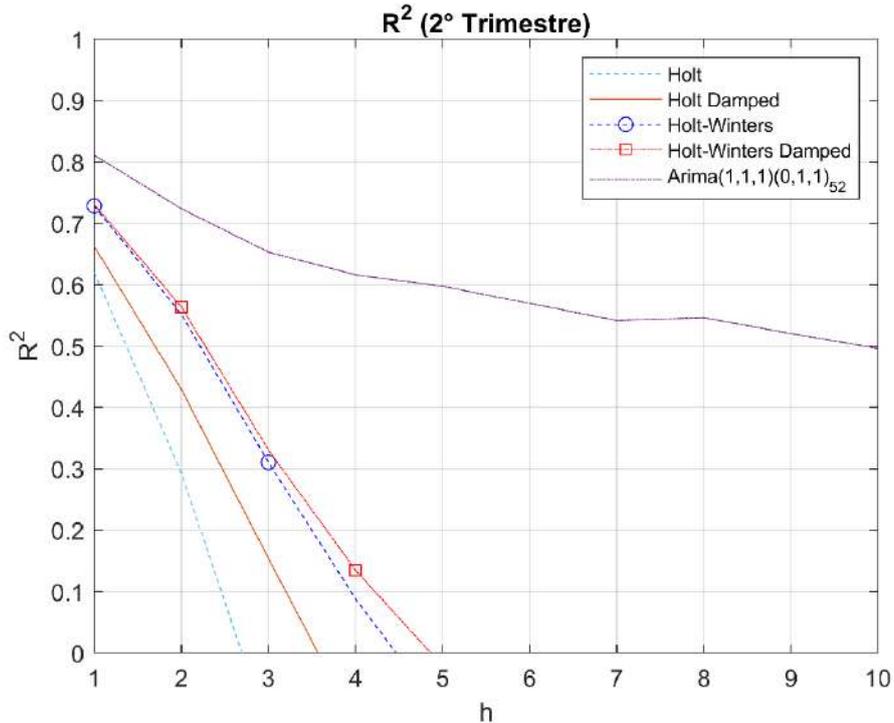


Figura 93. R² vs h (2º Trimestre) - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

La Figura 94, Figura 95 y Figura 96 muestran el comportamiento de los modelos en el tercer trimestre, que corresponde al periodo de invierno. En esta gráfica se excluye a los modelos de suavizamiento exponencial doble por presentar valores de ECMP y MAE muy altos, dificultando la comparativa entre los modelos con menor error. Es evidente que el Modelo ARIMA (1,1,1)(0,1,1)₅₂ obtiene el mejor rendimiento, con un valor máximo de ECMP de 3.0908×10^3 en el último horizonte de predicción y un valor MAE máximo de 23.02. El error cuadrático medio producido por los modelos en este trimestre es alto en relación a la varianza de la muestra, razón por la cual el valor del coeficiente de determinación es menor a cero a partir del tercer horizonte de predicción. En los meses de invierno, la serie temporal de caudales medio toma valores bajos, con nivel muy constante. Siendo la serie tan horizontal en este trimestre, las predicciones de los modelos resultan interesantes hasta el segundo o tercer horizonte. En horizontes posteriores es más efectivo predecir con el valor medio del caudal en el trimestre. Para valores positivos del coeficiente de determinación, se interpreta que las predicciones realizadas por el modelo son superiores a predecir utilizando la media de los datos. Cuando el coeficiente es negativo, indica que la predicción utilizando la media es más efectiva que la predicción realizada con el modelo.

El último trimestre, que hace referencia al periodo de primavera, se presenta en la Figura 97 a la Figura 99. Se excluyen de esta gráfica los modelos de suavizamiento exponencial

triple por dificultar la comparativa al presentar valores muy altos de error cuadrático medio y error absoluto medio. En este periodo, los modelos de suavizamiento exponencial doble muestran un mejor rendimiento en la mayoría de horizontes de predicción. El Modelo Holt consigue los menores valores de ECMP en los dos primeros horizontes. El Modelo ARIMA $(1,1,1)(0,1,1)_{52}$ cuenta con el ECMP más bajo en el tercer y cuarto horizonte de predicción. Por último, el Modelo Holt Damped obtiene el menor error a partir del quinto horizonte hasta el décimo. Al observar el error absoluto medio, se puede notar que el Modelo Holt Damped presenta los valores más bajos, a excepción del último horizonte, donde es superado por el Modelo ARIMA $(1,1,1)(0,1,1)_{52}$. Al igual que el trimestre anterior, los valores de ECMP son altos respecto a la varianza, y el coeficiente de determinación es menor a cero después del quinto horizonte de predicción. A partir de este punto, la predicción debe realizarse utilizando el valor medio del caudal en el trimestre.

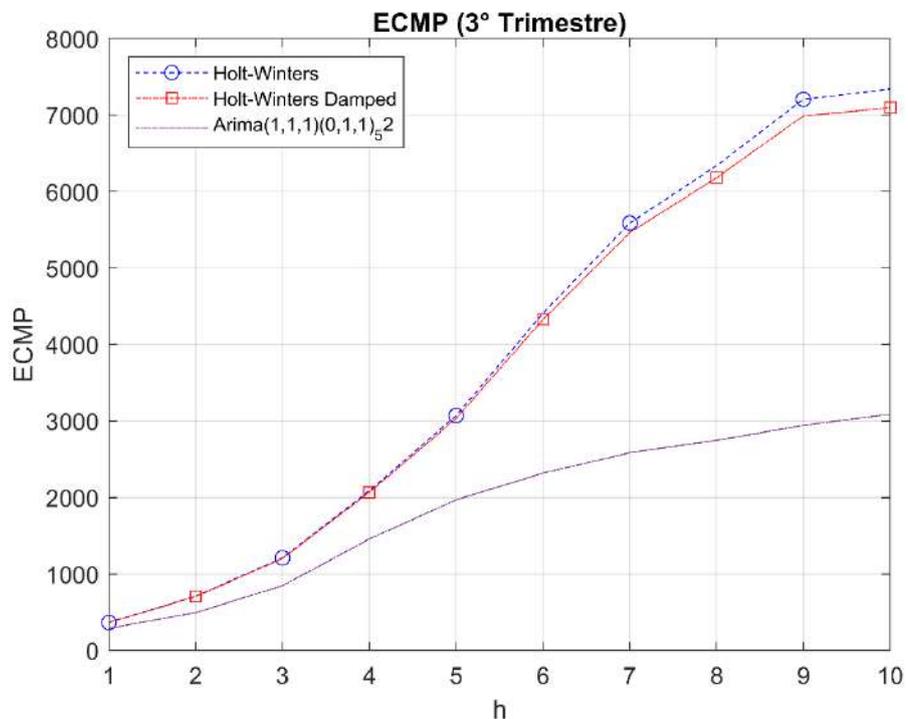


Figura 94. ECMP vs h (3º Trimestre) - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

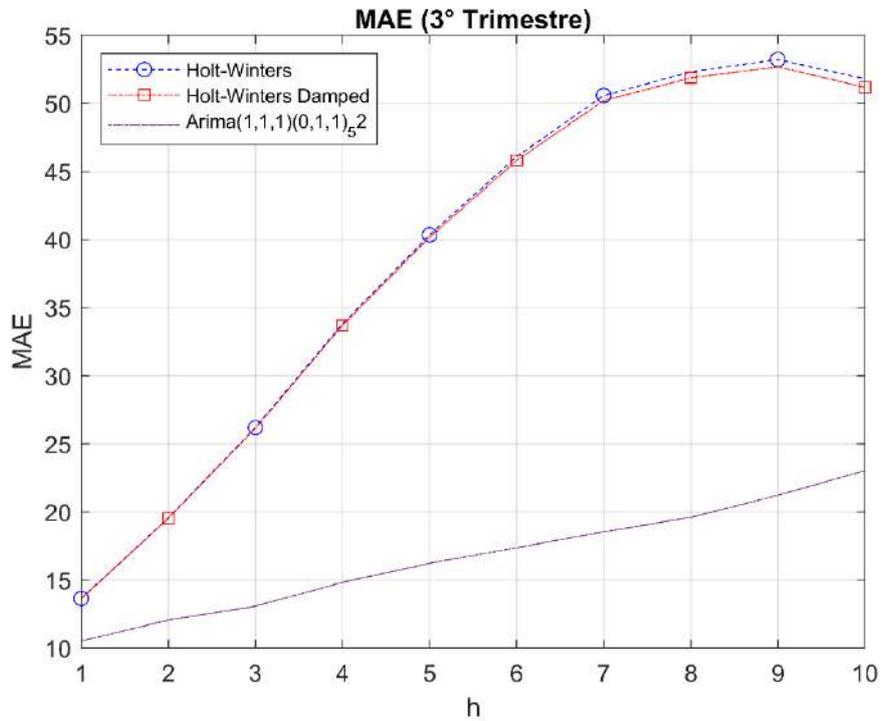


Figura 95. MAE vs h (3° Trimestre) - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

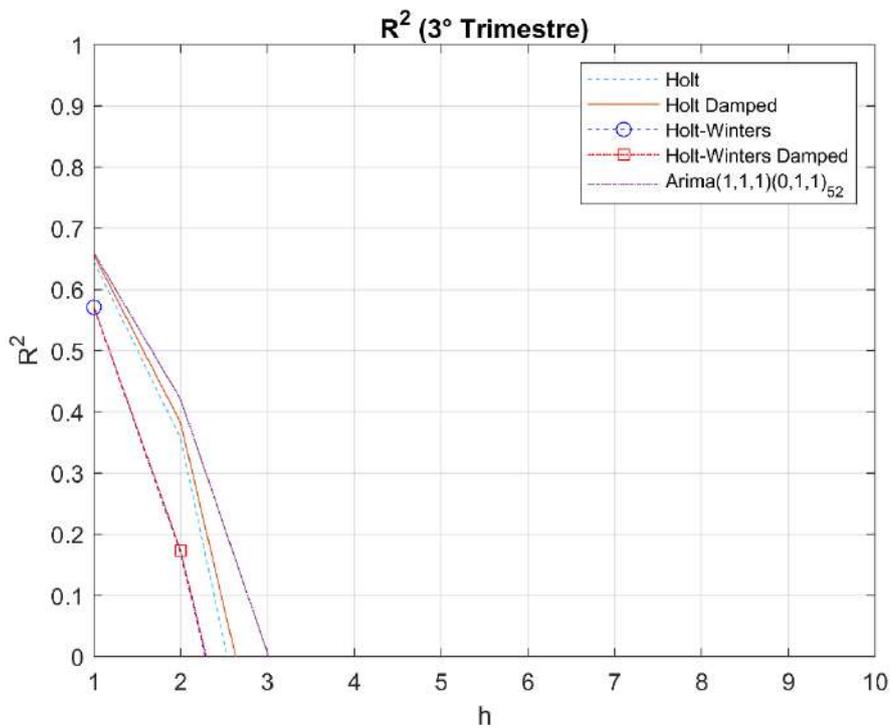


Figura 96. R² vs h (3° Trimestre) - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

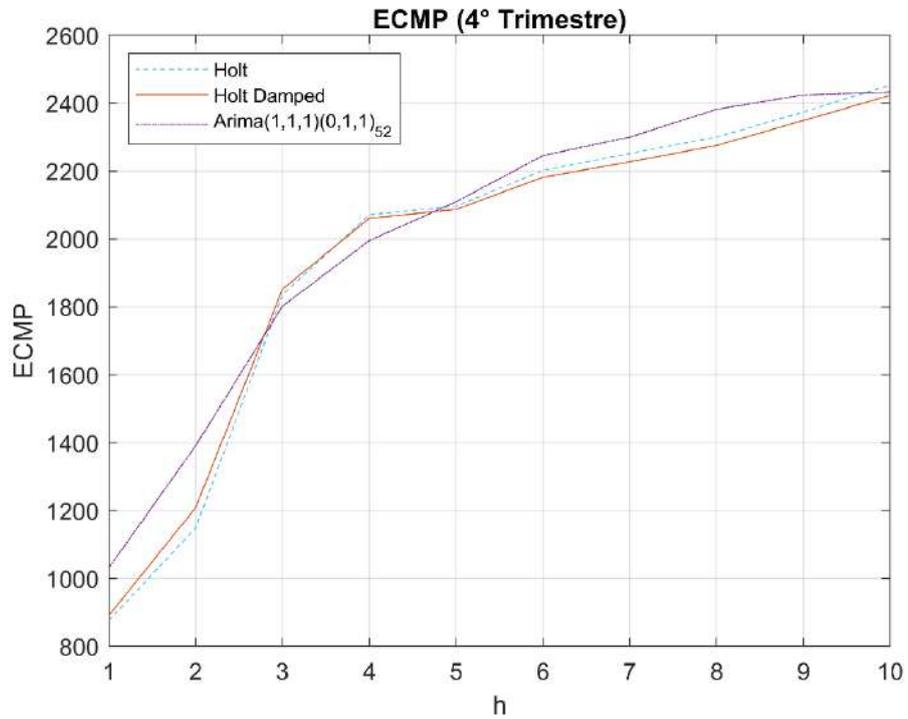


Figura 97. ECMP vs h (4° Trimestre) - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

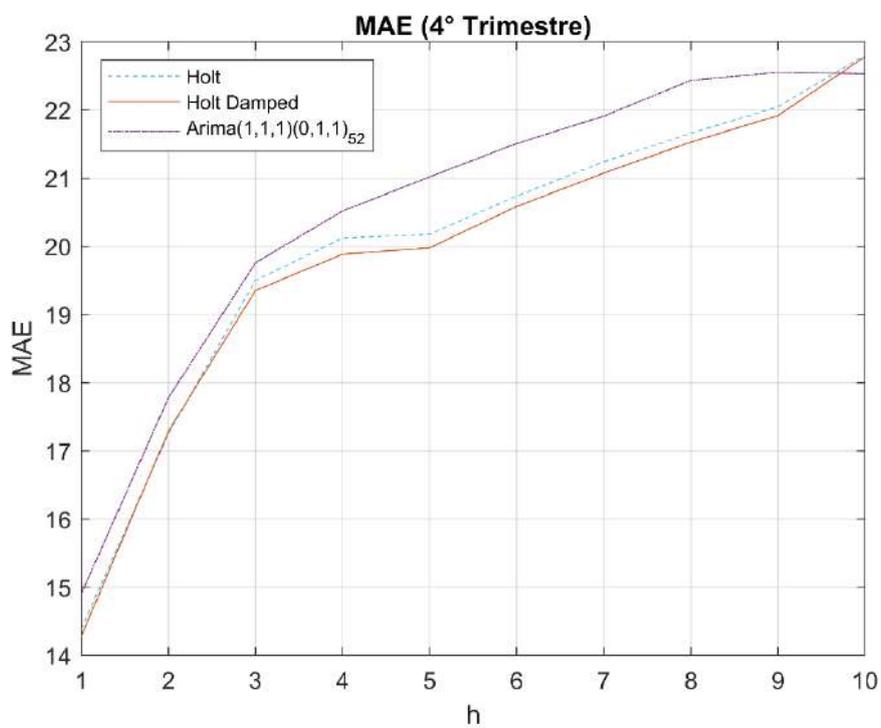


Figura 98. MAE vs h (4° Trimestre) - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

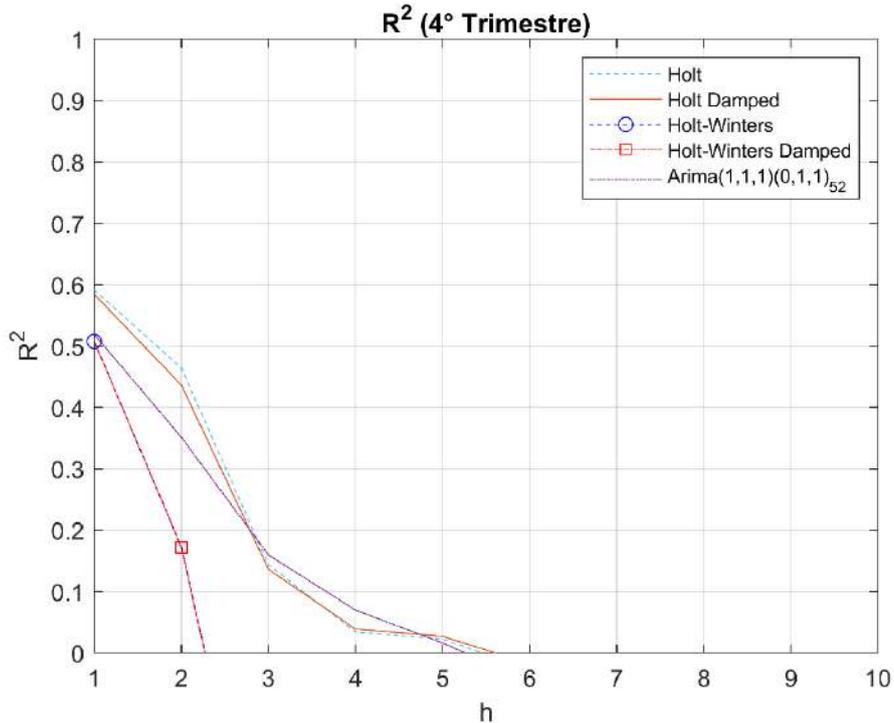


Figura 99. R^2 vs h (4º Trimestre) - Comparativa frecuencia semanal.

Fuente: Elaboración propia.

Por último, se realiza el análisis de la capacidad predictiva de los modelos con relación al valor del caudal medido. La distribución de la serie de caudal medio semanal se muestra en la Figura 16 y se aprecia que la mayor frecuencia de la serie se encuentra en el intervalo de 0 a 500 m³/s. Por esta razón, la evaluación del error cuadrático medio y el error absoluto medio respecto del caudal real se centra en este intervalo. El análisis se realiza diferenciando los horizontes de predicción.

El error cuadrático medio respecto del caudal en el primer horizonte de predicción se visualiza en la Figura 100. Los valores de ECMP se incrementan ante el aumento en los valores del caudal. En un primer intervalo de 0 a 100 m³/s, el comportamiento de los modelos es similar, con valores de ECMP bajos. Las diferencias se manifiestan con el aumento del caudal. Los modelos de suavizamiento doble presentan los valores más altos de ECMP, mientras que los modelos de suavizamiento exponencial triple tienen un comportamiento similar, con valores de ECMP más bajos. El modelo ARIMA supera a los anteriores, manteniendo el menor error en tanto que que los valores de caudal aumentan. En la Figura 101 se observa el error absoluto medio, donde el modelo ARIMA junto con los modelos de suavizamiento exponencial triple tienen los valores más bajos con el aumento del caudal.

En la Figura 102 y la Figura 103 se observan las mismas gráficas, esta vez para el quinto horizonte de predicción. El Modelo Holt presenta valores de ECMP y MAE muy altos, con un

crecimiento exponencial del ECMP conforme aumenta el valor del caudal, y seguido por el Modelo Holt Damped. El modelo ARIMA (1,1,1)(0,1,1)₅₂ consigue un rendimiento superior a los modelos de suavizamiento exponencial, obteniendo el menor error cuadrático medio y menor error absoluto medio en el intervalo de evaluación.

La Figura 104 y Figura 105 muestra la comparativa en el décimo horizonte de predicción. En este gráfico se excluye a los modelos de suavizamiento exponencial doble por presentar valores muy altos de ECMP y MAE, dificultando la comparativo de los otros modelos. El modelo que utiliza la metodología ARIMA es el que obtiene el mejor rendimiento, al presentar el menor error absoluto medio y error cuadrático medio a lo largo de todo el intervalo de evaluación, superando al Modelo Holt-Winters y Modelo Holt-Winters Damped.

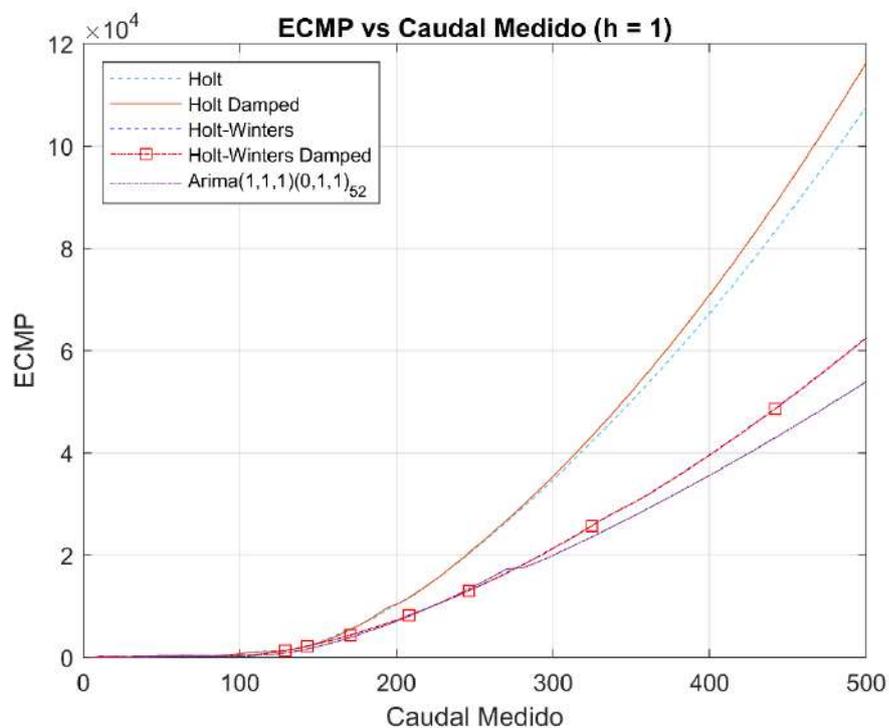


Figura 100. ECMP vs Caudal Medido - Primer horizonte de predicción. Frecuencia semanal.

Fuente: Elaboración propia.

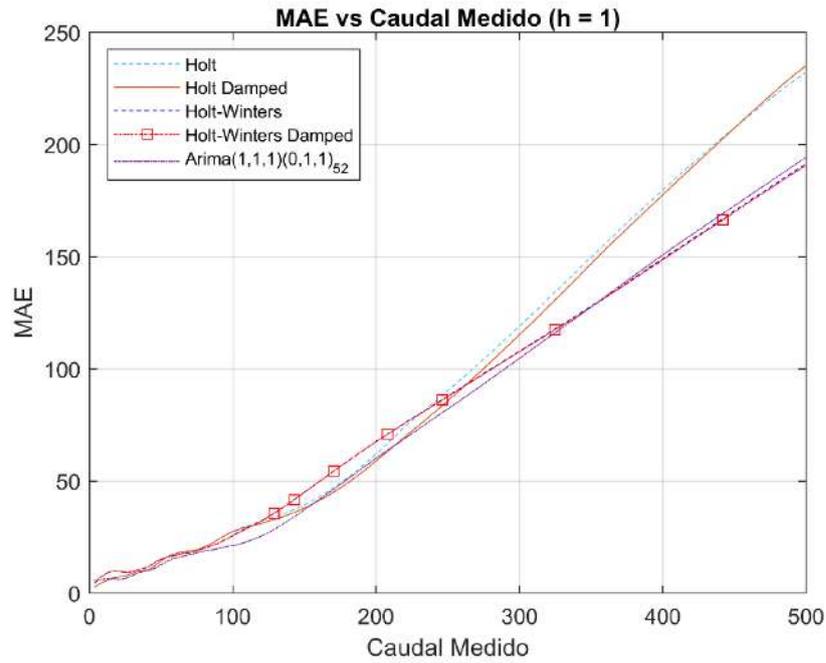


Figura 101. MAE vs Caudal Medido - Primer horizonte de predicción. Frecuencia semanal.

Fuente: Elaboración propia.

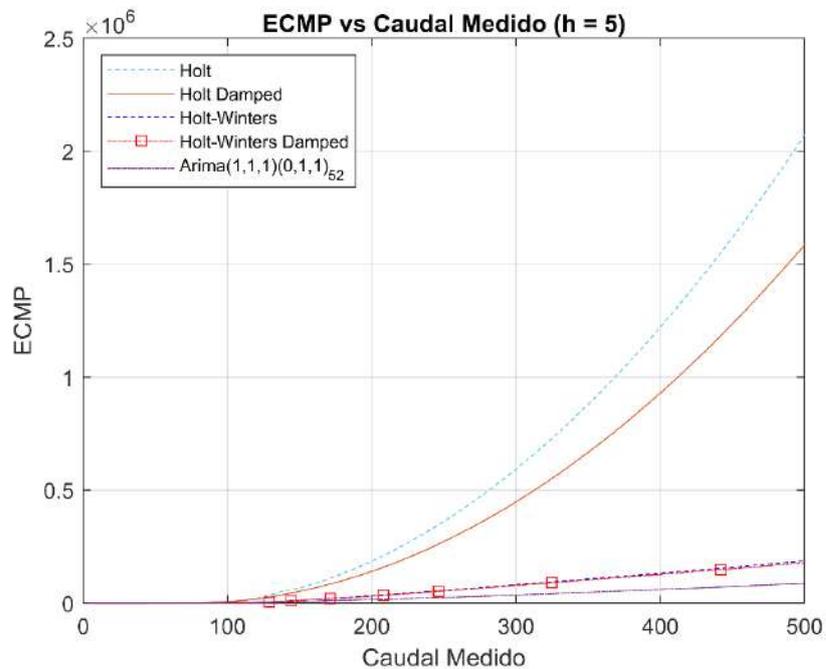


Figura 102. ECMP vs Caudal Medido - Quinto horizonte de predicción. Frecuencia semanal.

Fuente: Elaboración propia.

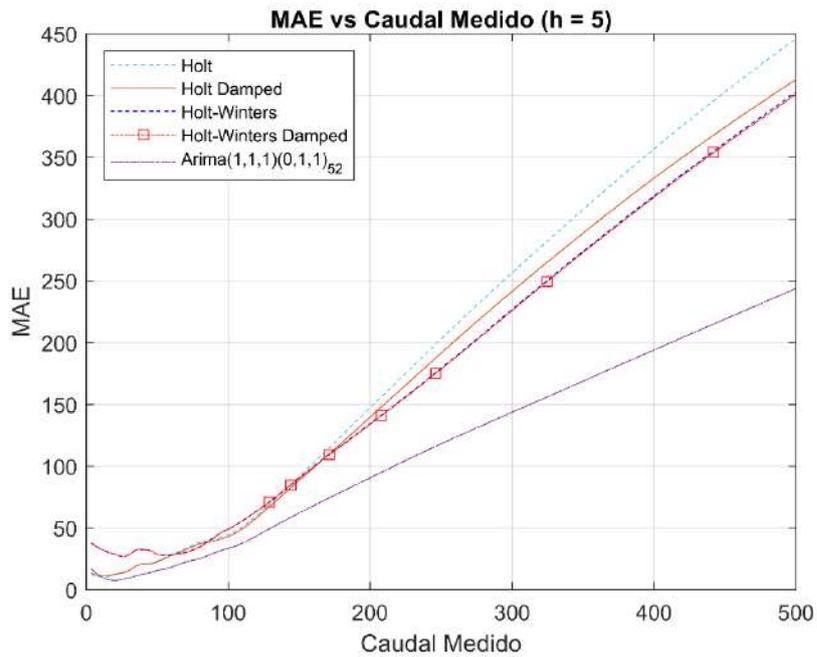


Figura 103. MAE vs Caudal Medido - Quinto horizonte de predicción. Frecuencia semanal.

Fuente: Elaboración propia.

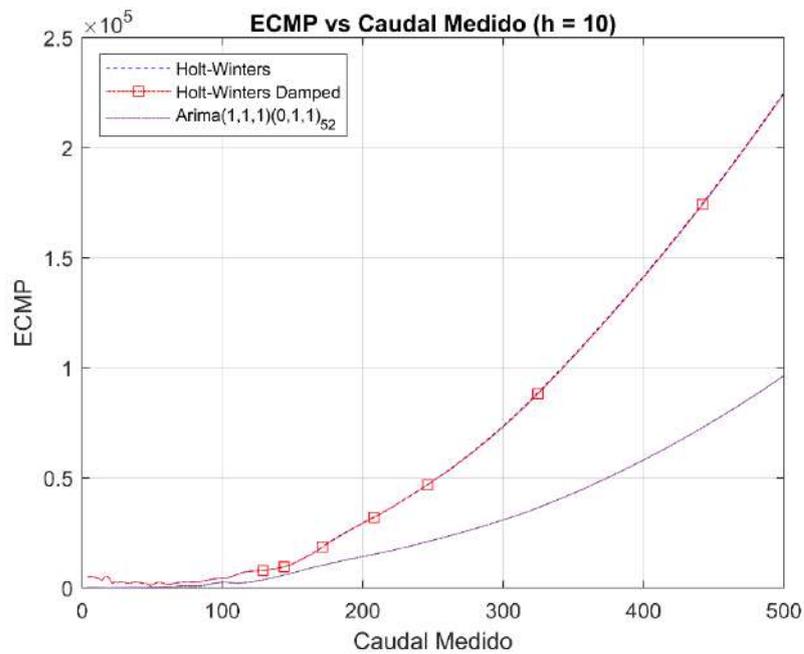


Figura 104. ECMP vs Caudal Medido - Décimo horizonte de predicción. Frecuencia semanal.

Fuente: Elaboración propia.

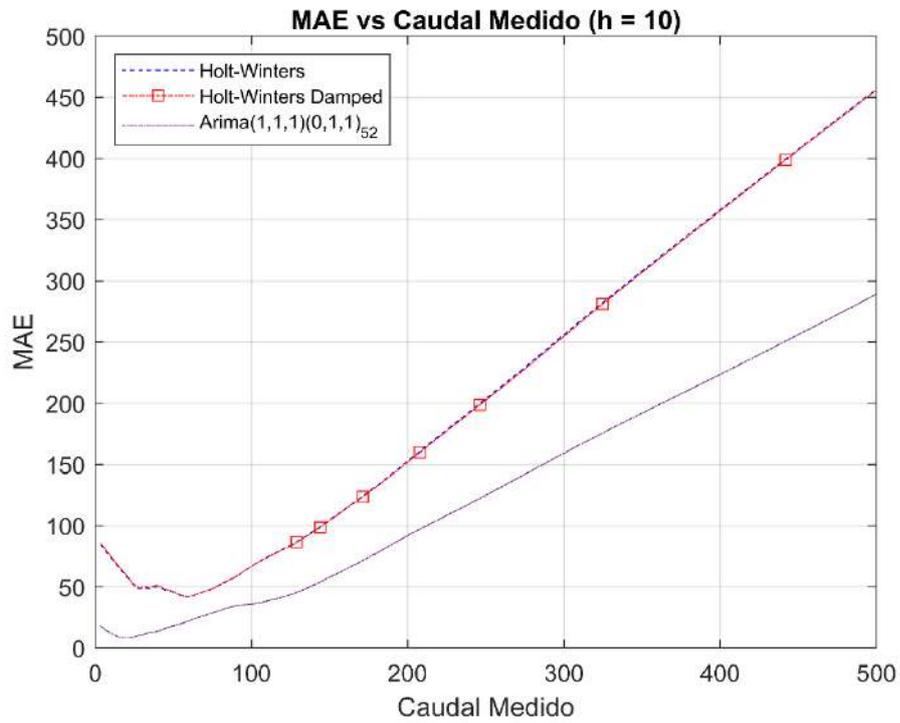


Figura 105. MAE vs Caudal Medido - Décimo horizonte de predicción. Frecuencia semanal.

Fuente: Elaboración propia.

En resumen, en la mayoría de los análisis expuestos, el Modelo ARIMA(1,1,1)(0,1,1)₅₂ presenta el menor error en las predicciones respecto a los otros modelos de la comparativa.

Conclusiones

La presente investigación realiza un análisis de la serie temporal de caudales promedio, con la información recopilada en la estación meteorológica Ardilla ubicada en el río Chira, aguas arriba del Reservorio de Poechos. Los datos se agruparon con frecuencia diaria y semanal, analizando un total de 67 años, comprendidos en un periodo desde enero de 1951 a agosto de 2017.

En base a las series obtenidas, se elaboró distintos modelos de predicción univariante que satisfagan la necesidad de prever el valor del caudal para fines de operación del Reservorio de Poechos. Para el desarrollo de modelos, se consideró los métodos de suavizamiento exponencial, en sus variantes de doble ecuación de suavizado y triple ecuación, y la metodología ARIMA desarrollada por Box-Jenkins.

Para reducir el impacto de la variabilidad de las series, se transformó las series haciendo uso del logaritmo natural. Los modelos se desarrollaron utilizando las series transformadas, a excepción de los métodos de suavizamiento exponencial triple multiplicativos, que consideran en su estructura la variabilidad estacional proporcional al nivel de la serie.

El ejercicio de predicción se realizó considerando un horizonte de diez días para la serie de caudal medio diario y diez semanas para la serie de caudal medio semanal. En ambos casos se consideró un periodo de muestra inicial, comprendido entre 1951 a 1979. Las predicciones se realizan desde 1980 hasta al final de la data recaudada en 2017. En cada iteración, el periodo de muestra incorpora una unidad de tiempo. Se realizó con este método 13731 iteraciones para el análisis diario y 1953 iteraciones para el análisis semanal.

El valor del error cuadrático medio, el error absoluto medio y el coeficiente de determinación se utilizaron como indicadores de la capacidad predictiva. Se realizó la comparativa entre los modelos, encontrando los mejores resultados para cada horizonte de predicción en los métodos ARIMA. El Modelo ARIMA (1,1,1) obtuvo el mejor rendimiento en el análisis diario y el Modelo ARIMA (1,1,1)(0,1,1)₅₂ el mejor del análisis semanal.

Con el objetivo de complementar el estudio, se analizaron los indicadores en los diferentes horizontes de predicción para cada trimestre del año, que corresponden a las

estaciones del año. De igual manera, se incluyó el análisis del error cuadrático medio y el error absoluto medio respecto del valor del caudal, para evaluar la capacidad predictiva de los modelos respecto al valor del caudal medido.

En conclusión, el Modelo ARIMA (1,1,1) se selecciona como método de predicción para la serie temporal del caudal promedio con frecuencia diaria y el Modelo ARIMA (1,1,1)(0,1,1)₅₂ es seleccionado como el método de predicción óptimo para la serie de caudal promedio semanal.



Recomendaciones

Considerar la utilización de los modelos propuestos realizando estimaciones adaptativas. La serie temporal del caudal del río Chira cambia en función al tiempo y se aprecian variaciones importantes en su comportamiento como los detectados en el análisis trimestral. El uso de técnicas de estimación adaptativa puede incrementar la exactitud de las predicciones y se debe evaluar en base a las evaluaciones desarrolladas en esta investigación, considerando la capacidad de adaptación, la facilidad de uso y el coste computacional.

Incluir en el análisis parámetros externos como las precipitaciones o temperatura registrados en la estación Ardilla y también en estaciones aguas arriba como las estaciones El Ciruelo, Visín, Puente Internacional y Paraje Grande, para mejorar la capacidad de predicción de los modelos.

En base al presente trabajo, investigar el empleo de nuevas metodologías de machine learning, tal como métodos no paramétricos basados en knn (k-nearest neighbors) o la utilización de redes neuronales. A través de estos métodos se obtienen modelos con menor error en las predicciones.

Referencias bibliográficas

- Aguilar Villena, R. M. (2016). *Predicción de caudales en el Río Chira con fines de descolmatación del embalse de Poechos (Tesis de Master en Ingeniería Civil con Mención en Recursos Hídricos)*. Piura: Universidad de Piura. Facultad de Ingeniería.
- ANA. (2012). *Compendio de los recursos hídricos superficiales de Piura*.
- Box, G., & Jenkins, G. (1976). *Time Series Analysis forecasting and control*. Oakland, California: Holden-Day.
- Chinchay Alza, L., Santana Vera, G., Panta Samillán, V., Portocarrero Lau, C., & Vásquez Arca, P. (2007). *Plan maestro de gestión integrada de los recursos hídricos en las cuencas hidrográficas de la región Piura*. Piura.
- CONSORCIO ATA-UNL-UNP. (2002). *Caracterización hídrica y adecuación entre la oferta y la demanda en el ámbito de la cuenca binacional Catamayo-Chira*.
- Enríquez Beck, J. A. (2016). *Estado hidrológico, hidráulico y sedimentológico de los embalses en la Región Piura (Tesis de pregrado en Ingeniería Civil)*. Piura: Universidad de Piura. Facultad de Ingeniería. Programa Académico de Ingeniería Civil.
- Hyndman, R. J., & Athanasopoulos, G. (2018). *Forecasting: principles and practice* (Segunda ed.). Melbourne, Australia: OTexts. Obtenido de <https://otexts.com/fpp2/>
- Peña, D. (2005). *Análisis de series temporales*. Madrid: Alianza Editorial.
- Proyecto Especial Chira Piura. (14 de Agosto de 2020). Obtenido de <http://www.chirapiura.gob.pe/index.php?pag=13&id=43>
- Proyecto Especial Chira Piura I. (14 de Agosto de 2020). Obtenido de <http://www.chirapiura.gob.pe/proyecto/primera-etapa>
- Proyecto Especial Chira Piura II. (14 de Agosto de 2020). Obtenido de <http://www.chirapiura.gob.pe/proyecto/segunda-etapa>
- Proyecto Especial Chira Piura III. (14 de Agosto de 2020). Obtenido de <http://www.chirapiura.gob.pe/proyecto/tercera-etapa>
- Reyes Salazar, J., Farías de Reyes, M., Guerrero Asmad, L., & Tesén Arámbulo, K. (2018). Modelación numérica del proceso de descolmatación del reservorio Poechos utilizando la misma energía del agua. *Innovation in Education and Inclusion : Proceedings of the 16th LACCEI International Multi-Conference for Engineering, Education and Technology*, (pág. 38). Lima, Perú.

Universidad de Piura. (28 de Enero de 2014). *UDEP Académico*. Recuperado el 10 de Julio de 2019, de <http://udep.edu.pe/hoy/2014/nos-hemos-convertido-en-dependientes-del-reservorio-de-pochos/>

Varona Alama, M. C. (2018). *Modelo hidrológico de la cuenca Catamayo-Chira hasta el ingreso al reservorio Poechos usando Hec-Hms (Tesis para optar el título de Ingeniero Civil)*. Piura: Universidad de Piura, Facultad de Ingeniería. Programa Académico de Ingeniería Civil.

