



UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

Departamento de Obras Civiles

**PRONÓSTICOS ESTACIONALES DE CAUDALES
EN LA CUENCA ALTA DEL RÍO MAIPO
UTILIZANDO MODELOS ADITIVOS
GENERALIZADOS PARA UBICACIÓN, ESCALA Y
FORMA.**

Memoria de Título presentada por

Fernanda Monserrat Inostroza González

como requisito parcial para optar al título de la carrera de

Ingeniería Civil

Profesor Guía
Álvaro Ossandón

Septiembre de 2024



CONSTANCIA DE VALIDACIÓN Y CONFIDENCIALIDAD DE MONOGRAFÍA A REPOSITORIO ACADÉMICO

1.- IDENTIFICACIÓN DEL TRABAJO ACADÉMICO

Tipo de monografía (marcar una opción): ☒ Memoria o trabajo de título ☐ Tesis de Postgrado

Título del trabajo: Pronósticos estacionales de caudales en la cuenca alta del Río Maipo utilizando modelos aditivos generalizados para ubicación, escala y forma.

Nombre del candidato(a): Fernanda Monserrat Inostroza González

Carrera / Grado: Ingeniería Civil

Campus: San Joaquín

Departamento: Obras Civiles

2.- VALIDACIÓN DEL PROFESOR GUÍA/DIRECTOR DE TESIS

Yo, Álvaro Ossandón, en mi calidad de profesor(a) guía/director(a) del trabajo académico mencionado anteriormente **DEJO CONSTANCIA** que:

- He revisado esta versión del documento y corresponde a la versión final aprobada del trabajo.
- El trabajo cumple con los requisitos académicos y de formato establecidos por la institución.

3.- EVALUACIÓN DE CONFIDENCIALIDAD POR PROPIEDAD INDUSTRIAL (marcar una opción)

☒ El trabajo **NO contiene** información que amerite confidencialidad y puede ser publicado de inmediato en repositorio con acceso abierto.

☐ El trabajo **CONTIENE** información con potenciales implicancias de propiedad industrial o intelectual y requiere un periodo de confidencialidad (**embargo**) por (**marcar una opción**):

☐ 6 meses ☐ 12 meses ☐ 2 años ☐ 3 años ☐ 5 años ☐ 10 años

Fundamentación de la necesidad de confidencialidad (obligatorio si se solicita embargo):

4.- FIRMAS

Profesor(a) guía o director(a) de memoria o tesis:

Fecha: 29-10-2025

Firma:

Estudiante o Candidato(a):

Fecha: 29-10-2025

Firma:

Este formulario debe ser insertado como página 2 de la memoria o tesis, completado y firmado por estudiante y profesor(a) antes de la entrega en portal PRISMA de Biblioteca USM.



UNIVERSIDAD TECNICA
FEDERICO SANTA MARIA

TITULO DE LA TESIS:

**PRONÓSTICOS ESTACIONALES DE CAUDALES EN LA CUENCA
ALTA DEL RÍO MAIPO UTILIZANDO MODELOS ADITIVOS
GENERALIZADOS PARA UBICACIÓN, ESCALA Y FORMA.**

AUTOR:

FERNANDA MONSERRAT INOSTROZA GONZÁLEZ

TRABAJO DE MEMORIA, presentado como requisito parcial para optar al título de la carrera de INGENIERIA CIVIL de la Universidad Técnica Federico Santa María.

	<u>Nombre</u>	<u>Firma</u>
Prof. Guía		
Miembro 1 Comisión		
Miembro 2 Comisión		

Santiago, Chile, Septiembre de 2024

PRONÓSTICOS ESTACIONALES DE CAUDALES EN LA CUENCA ALTA DEL RÍO MAIPO UTILIZANDO MODELOS ADITIVOS GENERALIZADOS PARA UBICACIÓN, ESCALA Y FORMA

Fernanda Inostroza¹, Álvaro Ossandón¹

¹ Departamento de Obras Civiles, Universidad Técnica Federico Santa María, Santiago/Valparaíso, Chile.

Resumen

Este estudio evalúa modelos GAMLSS (Modelos Aditivos Generalizados para Localización, Escala y Forma) y los compara con un modelo de Regresión Lineal Multivariable (MRLM) para predecir caudales estacionales en la cuenca alta del río Maipo, durante las temporadas de primavera y verano. Se utilizaron diferentes tiempos de anticipación (0 y 3 meses) para calibrar y validar los modelos. Las distribuciones analizadas incluyen Gamma, Gamma Generalizada, Log Normal y Log Normal 2, utilizando predictores climáticos e hidrometeorológicos como el Niño 3.4, el AAO, caudales mensuales previos y anomalías de la temperatura superficial del mar (SST). Los resultados muestran que la distribución Gamma es la más adecuada, con un buen equilibrio entre ajuste del modelo, precisión y capacidad predictiva. Sin embargo, la capacidad predictiva de los modelos GAMLSS disminuye a mayor tiempo de desfase, especialmente en verano, aumentando la incertidumbre en las predicciones. Aunque los modelos GAMLSS capturan mejor la incertidumbre en comparación con el MRLM, no siempre son los más precisos para predecir los valores más probables de los caudales. Por tanto, es necesario explorar modelos alternativos que mejoren la precisión y reduzcan la incertidumbre en la predicción de caudales bajo diversos escenarios climáticos.

Palabras claves: Predicción de caudales estacionales, GAMLSS, Distribución Gamma, Incertidumbre, Regresión Lineal Multivariable.

1 Introducción

El cambio climático según el (Ministerio del Medio Ambiente, 2023) consiste en una variación que se está registrando en el clima del planeta, atribuido directa o indirectamente a la actividad humana, y que altera la composición de la atmósfera. En los últimos años, el cambio climático ha tenido un gran impacto en Chile, desencadenado una serie de consecuencias devastadoras, como la sequía (Garreaud et al., 2020), derretimiento de glaciares (Escobar et al., 1995), el desbordamiento de ríos (Gironas, 2013), una mayor intensidad de precipitaciones en eventos extremos (Valenzuela & Garreaud, 2019), incrementos del nivel de la isoterma 0° (Mardones, 2019), aumento de inundaciones (Rojas et al., 2014), entre otros.

Se presenta la necesidad de planificar los recursos hídricos (Sadoff & Muller, 2010) con el fin de adaptarse a las necesidades climáticas y humanas. En Chile, se creó la Estrategia Nacional de Recursos Hídricos (Ministerio del Medio Ambiente, 2012) para gestionar y regular los recursos de agua garantizando su uso sostenible y permitiendo una adecuada protección de la cantidad y calidad de aguas. Debido a lo anterior, resulta fundamental contar con herramientas robustas y precisas que permitan anticipar y abordar estos problemas, especialmente en lo que respecta al aumento y disminución del caudal en ríos (Feinstein et al., 2023). En este contexto, resulta fundamental desarrollar métodos que nos permitan predecir con precisión el caudal en un futuro cercano, para que así estos puedan ser utilizados en la planificación del recurso hídrico.

Los modelos lineales son una herramienta ampliamente utilizada para realizar pronósticos de caudales a distintas escalas temporales (Clark & Wells, 2023). Sin embargo, ellos presentan deficiencias tales como incapacidad para reproducir procesos no lineales involucrados en el proceso de escorrentía. Un ejemplo de esto son los modelos de series de tiempo (AR, ARIMA, SARIMA y PAR), los cuales son adecuados para pronosticar caudales dentro del rango de datos disponibles, pero presentan limitaciones fuera de dicho rango (Flamenco, 2010). También se realizó un estudio el cual permite comparar modelos lineales con modelos no lineales, el cual consiste en la predicción de caudales medios mensuales en ríos colombianos utilizando métodos no lineales, y para comparar se usan los modelos estocásticos lineales univariados AR(p) con métodos no lineales como MARS, Redes Neuronales y PREBEO. Los resultados indican que los métodos no lineales superaron en rendimiento al enfoque lineal, produciendo mejores pronósticos (Poveda et al., 2002).

Se han llevado a cabo diversos estudios orientados a identificar los modelos de predicción más efectivos, priorizando la evaluación de variables climatológicas con potencial predictivo. En un enfoque específico, el trabajo de Montalva (2022) se destacó al centrarse exclusivamente en la aplicación de variables climatológicas mediante métodos lineales. Los resultados de este estudio concluyen que las variables climatológicas seleccionadas demuestran ser predictores significativos dentro del marco de un modelo lineal; sin embargo, aún existe un vacío en la investigación al buscar modelos de pronóstico más avanzados que incorporen variables climatológicas y evaluar su desempeño en contextos no lineales.

Un área de investigación prometedora es la búsqueda de modelos más completos que tengan en cuenta una amplia gama de variables y aspectos al pronosticar. Los métodos GAMLSS, que incorporan ubicación, escala y forma, se destacan como la elección principal en la actualidad. Esta integración permite capturar la tendencia central, modelar patrones temporales, y abordar la variabilidad y la forma de la distribución. La ventaja distintiva de los GAMLSS radica en su capacidad para suavizar la función de ajuste a los datos, proporcionando resultados más precisos. Estos métodos se distinguen por su flexibilidad al permitir la inclusión de variables no estacionarias, y su efectividad ha sido sobradamente demostrada, especialmente en situaciones que se encuentran más allá del alcance tradicional de los datos de estudio (Debele et al., 2017).

Debido a los resultados positivos de la aplicación de modelos GAMLSS y la detección de variables climatológicas con potencial predictivo para el caudal estacional en la cuenca del Río Maipo, el objetivo es utilizar modelos GAMLSS para pronosticar caudales, considerando las variables climatológicas con potencial predictivo. Se realizarán análisis con diferentes modelos GAMLSS no lineales, con el propósito de encontrar la estructura que nos brinde la capacidad de pronosticar de manera precisa el caudal estacional. Esta memoria busca brindar herramientas sólidas y confiables para enfrentar los desafíos que surgen de este fenómeno global y avanzar hacia una gestión sostenible de los recursos hídricos.

2 Objetivos

2.1 Objetivo principal

El objetivo principal es implementar un modelo de pronóstico de caudales estacionales en la cuenca alta del Río Maipo utilizando modelos aditivos generalizados para ubicación, escala y forma.

2.2 Objetivos específicos

Para alcanzar el objetivo principal se presentan los siguientes objetivos específicos.

- 1) Construir y ajustar distintos modelos de pronóstico de caudales estacionales en la cuenca alta del Río Maipo utilizando las variables climáticas en los modelos aditivos generalizados para ubicación, escala y forma.
- 2) Comparar el rendimiento de los modelos aditivos generalizados para ubicación, escala y forma con los modelos tradicionales lineales para evaluar la eficiencia en la predicción.

3 Marco Teórico

3.1 Modelos GAMLSS

Los modelos aditivos generalizados para ubicación, escala y forma (GAMLSS) son modelos de regresión de tipo semiparamétricos (Stasinopoulos & Rigby, 2007). Se clasifican como paramétricos debido a la necesidad de asumir una distribución paramétrica para la variable de respuesta, y como 'Semi' porque los parámetros de dicha distribución pueden ser modelados de manera independiente mediante funciones no paramétricas.

Un modelo GAMLSS asume que, para $i=1,2,\dots,n$, observaciones independientes de la variable de respuesta Y_i , está tiene función de densidad $f_Y(y_i|\theta^i)$ condicional en $\theta^i = (\theta_{1i}, \theta_{2i}, \theta_{3i}, \theta_{4i}) = (\mu_i, \sigma_i, v_i, \tau_i)$, un vector de cuatro parámetros de distribución, donde cada uno puede ser una función de las variables explicativas. Esto se denota por:

$$Y \sim D(\mu_i, \sigma_i, v_i, \tau_i)$$

Donde D representa la distribución de la variable de respuesta con los parámetros de ubicación, escala y forma. El modelo GAMLSS asume un vector de cuatro parámetros de distribución $(\mu_i, \sigma_i, v_i, \tau_i)$, donde cada uno puede ser una función de las variables explicativas. Los parámetros μ y σ representan los parámetros de ubicación y escala respectivamente. Los parámetros v y τ son parámetros de forma como por ejemplo la asimetría y la curtosis.

Sea $Y^T = (y_1, y_2, \dots, y_n)$ el vector de longitud n de la variable de respuesta, y para $k = 1, 2, 3, 4$, se consideran $g_k()$, funciones de enlace monótonas bien conocidas que se relacionan con los parámetros de distribución para las variables explicativas, de acuerdo a:

$$g_k(\theta_k) = \eta_k = X_k \beta_k + \sum_{j=1}^{J_k} Z_{jk} \gamma_{jk} \quad (1)$$

Es decir,

$$\begin{aligned} g_1(\mu) &= \eta_1 = X_1 \beta_1 + \sum_{j=1}^{J_1} Z_{j1} \gamma_{j1} \\ g_2(\sigma) &= \eta_2 = X_2 \beta_2 + \sum_{j=1}^{J_2} Z_{j2} \gamma_{j2} \\ g_3(v) &= \eta_3 = X_3 \beta_3 + \sum_{j=1}^{J_3} Z_{j3} \gamma_{j3} \\ g_4(\tau) &= \eta_4 = X_4 \beta_4 + \sum_{j=1}^{J_4} Z_{j4} \gamma_{j4} \end{aligned}$$

Donde:

μ, σ, v, τ son parámetros de distribución.

η_k es la combinación lineal de las variables predictoras.

$\beta_k = (\beta_{1k}, \beta_{2k}, \dots, \beta_{J'_k k})$ es un vector de coeficientes de regresión de longitud J'_k asociados a las variables predictoras.

X_k es la matriz de predictores de orden $n \times J'_k$. Dentro de esta matriz se encuentran los términos lineales del modelo.

Z_{jk} es la matriz de peso de residuos de orden $n \times q_{jk}$. Esta se asigna para dar diferentes niveles de importancia o influencia a las observaciones del modelo estadístico.

γ_{jk} es una variable auxiliar que puede representar diferentes formas según el modelo, puede representar efectos aleatorios, funciones de bases, o cualquier otro tipo de efecto que no sea simplemente un predictor lineal.

$X_k \beta_k$ es un término que se puede representar de la siguiente forma:

$$X_k \beta_k = \beta_0 + \beta_1 \cdot x_1 + \beta_2 \cdot x_2 + \dots + \beta_n \cdot x_n$$

Donde β_0 corresponde al coeficiente del intercepto, $\beta_1, \beta_2, \dots, \beta_n$ son los coeficientes de regresión asociados a las variables predictoras y $x_1, x_2, \dots, x_n$ son los valores de las variables predictoras.

$\sum_{j=1}^{J_k} Z_{jk} \gamma_{jk}$ es un término adicional que introduce mayor complejidad al modelo, permitiendo que η_k dependa no solo de una combinación lineal simple de las variables predictoras, sino también de otras combinaciones posibles. B η β_0 x_1 x_2 x_n r_s d_i r_p

3.2 Coeficiente de Correlación de Spearman (r_s)

El coeficiente de correlación de Spearman es una prueba no paramétrica de correlación de rango, que permite analizar datos sin requerir una distribución normal (Hollander et al., 2015). Esta medida es útil para evaluar la relación monótona entre dos variables, especialmente cuando los datos presentan outliers o distribuciones asimétricas. Se calcula mediante la siguiente ecuación:

$$r_s = 1 - \frac{6 \sum_i d_i^2}{n(n^2 - 1)} \quad (2)$$

Donde r_s es la correlación de Spearman obtenida entre las variables en estudio, n es el número de datos de la muestra y d_i es la diferencia entre los rangos asignados a cada par de observaciones.

Los resultados oscilan entre -1 y 1, considerando -1 como una asociación perfectamente negativa, 1 como una asociación perfectamente positiva y 0 como una ausencia de relación entre las variables.

3.3 Métricas de Evaluación de Performance de Modelos Seleccionados

En esta sección, se presentan las métricas utilizadas para evaluar el rendimiento de los modelos seleccionados. Algunas de estas métricas (AIC, R de Pearson, BIAS y CRPSS) se usaron para cuantificar la capacidad predictiva de los modelos al comparar las predicciones generadas con los datos observados y otras (CRPSS y CE) se usaron para evaluar el rendimiento de nuestros modelos con respecto a otros modelos ya definidos.

3.3.1 Criterio de Información de Akaike (AIC)

El Criterio de Información de Akaike (AIC) constituye una métrica que evalúa la calidad de un modelo estadístico en relación con un conjunto de datos específico. Este indicador desempeña un papel fundamental en la selección de modelos (Akaike, 1974). La esencia del AIC se encuentra en encontrar un equilibrio entre la capacidad de ajuste del modelo y su complejidad, lo que brinda una estimación relativa de cuánta información se sacrifica al emplear un modelo en particular para representar el proceso que genera los datos. De manera general, el AIC se calcula de la siguiente manera:

$$AIC = 2k - 2 \ln(L) \quad (3)$$

Donde k corresponde a el número de parámetros en el modelo estadístico y L corresponde al valor máximo de la función de verosimilitud para el modelo que se desea estimar.

Entre un conjunto de modelos potenciales para un conjunto de datos, el modelo que se selecciona es aquel que exhibe el valor más bajo en el Criterio de Información de Akaike (AIC). Es importante tener en cuenta que este criterio no proporciona una medida absoluta de la calidad del modelo y, por lo tanto, no puede determinar de manera definitiva si los modelos propuestos se ajustan adecuadamente a los datos. Por esta razón, se hace imperativo complementar este criterio con otras consideraciones y factores para evaluar de manera más completa la idoneidad de los modelos.

3.3.2 Coeficiente de Correlación de Pearson (R de Pearson)

El coeficiente de correlación de Pearson evalúa la relación de covariación entre dos variables que están relacionadas de manera lineal, sin tener en cuenta la escala de medida de estas variables. Por lo tanto, resulta útil para cuantificar la intensidad de la relación entre dos variables cuando ambas son de naturaleza cuantitativa y continua (Moore et al., 2012). El cálculo de este coeficiente se lleva a cabo mediante la siguiente ecuación:

$$r_p = \frac{\sigma_{XY}}{\sigma_X \sigma_Y} = \frac{Cov(X, Y)}{\sqrt{Var(X)Var(Y)}} \quad (4)$$

Donde σ_{XY} representa la covarianza entre las variables X e Y , σ_X corresponde a la varianza del conjunto de datos X , y σ_Y corresponde a la varianza del conjunto de datos Y .

El resultado obtenido a partir de este cálculo para el coeficiente de correlación de Pearson varía en un rango que va desde -1 hasta +1. La magnitud de la relación entre las variables se expresa a través del valor numérico del coeficiente: en el primer caso, cuando el coeficiente es +1, indica una relación perfectamente positiva entre los datos; en el segundo caso, cuando el coeficiente es -1, indica una relación perfectamente negativa entre los datos.

3.3.3 Sesgo Porcentual (BIAS)

El sesgo porcentual (BIAS) se refiere a la discrepancia entre la media anticipada de un modelo propuesto y la media de los valores reales que se utilizan para calibrarlo (Gupta et al., 1999). Su cálculo se efectúa de la siguiente manera:

$$BIAS = \left(\frac{\overline{d_e} - \overline{d_r}}{\overline{d_r}} \right) * 100\% \quad (5)$$

Donde $\overline{d_r}$ representa la media de los valores reales medidos, mientras que $\overline{d_e}$ representa la media de los valores ajustados por el modelo.

Si la estimación del BIAS resulta en un valor positivo, indica que el sistema de aproximación tiende a sobrestimar el valor real. Por otro lado, si el BIAS es negativo, sugiere que el sistema tiende a subestimar el valor real. En consecuencia, se considera deseable que el valor del BIAS se acerque lo más posible a cero, lo que indicaría una menor tendencia del sistema a sobreestimar o subestimar los valores reales.

D de d_r

3.3.4 Puntuación de Habilidad de Probabilidad Clasificada Continua (CRPSS)

El CRPSS es una métrica probabilística que mide la habilidad del CRPS en relación con un pronóstico de referencia (Hersbach, 2000). El CRPS mide cuanto se asemeja la distribución de probabilidad pronosticada (F) a la distribución real observada (F_o) (Ossandón et al., 2022). El cálculo se lleva a cabo mediante la siguiente ecuación:

$$CRPSS = 1 - \left(\frac{CRPS_{fcst}}{CRPS_{ref}} \right) \quad (5)$$

$$CRPS = \frac{1}{T} \sum_{t=1}^T \int_{-\infty}^{\infty} [F(Q) - F_o(Q)]^2 dQ \quad (6)$$

$$F_o(Q) = \begin{cases} 0 & Q < Q_o \\ 1 & Q \geq Q_o \end{cases} \quad (7)$$

Esta métrica probabilística varía entre (0,1], un valor de 1 determina una predicción perfecta, un valor de 0 indica que el modelo es tan bueno como la referencia y un valor < 0 indica que el modelo es peor que la referencia.

3.3.5 Coeficiente de Eficiencia (CE)

El Coeficiente de eficiencia es una métrica determinista, la cual mide la asociación lineal entre los pronósticos y las observaciones independientes de la media y varianza de las distribuciones, en su proceso no se introducen elementos probabilísticos (Ossandón et al., 2022).

$$CE = 1 - \frac{\sum_{t=1}^N (Q_t - \hat{Q})^2}{\sum_{t=1}^N (Q_t - Q_{ref,t})^2} \quad (8)$$

Donde Q_t equivale a los datos de caudales observados, $\hat{Q}$ son los datos de caudales pronosticados y $Q_{ref,t}$ son datos de caudales referenciales. El CE varía entre [-1,1], con un valor perfecto de 1, cuando el CE=0 significa que el modelo no mejora la predicción sobre la media de las observaciones o los valores referenciales y cuando el CE es <0 quiere decir que el modelo es peor que los datos referenciales.

4 Zona de estudio

La cuenca del río Maipo se ubica en la región central de Chile, y es la fuente primordial de agua potable para la Región Metropolitana de Santiago, abasteciendo aproximadamente al 70% de la demanda de agua potable de la zona, de la cual el 90% se destina a regadío (García, 2008). La cuenca se divide en tres sectores principales, y en este estudio nos enfocaremos en la Cuenca Alta del Río Maipo.

La Cuenca Alta del Río Maipo, con un área aproximada de 4900 km², es una región de importancia crítica debido a su contribución al suministro de agua en la región. Como se muestra en la Figura 1, los recursos hídricos disponibles en esta cuenca provienen del río Maipo y de sus afluentes cordilleranos, que incluyen los ríos Yeso, Volcán, Colorado y Clarillo, así como los esteros Manzano, El Canelo, San José y Coyanco.

Los registros utilizados en este estudio se obtienen específicamente de la Cuenca del Río Olivares, una subcuenca de la Cuenca Alta del Río Maipo. Esta subcuenca tiene un área aproximada de 540 km², con una cota máxima de 6049 msnm, una cota media de 3690 msnm, y su punto de salida se encuentra a 1568 msnm.

Es importante destacar que tanto la Cuenca Alta del Río Maipo como sus subcuencas, como la del Río Olivares, son áreas de alta demanda sobre los recursos hídricos (Farías-Barahona et al., 2020), lo que subraya la relevancia de estudiar y gestionar adecuadamente estos recursos para satisfacer las necesidades de agua potable y de riego en la región.

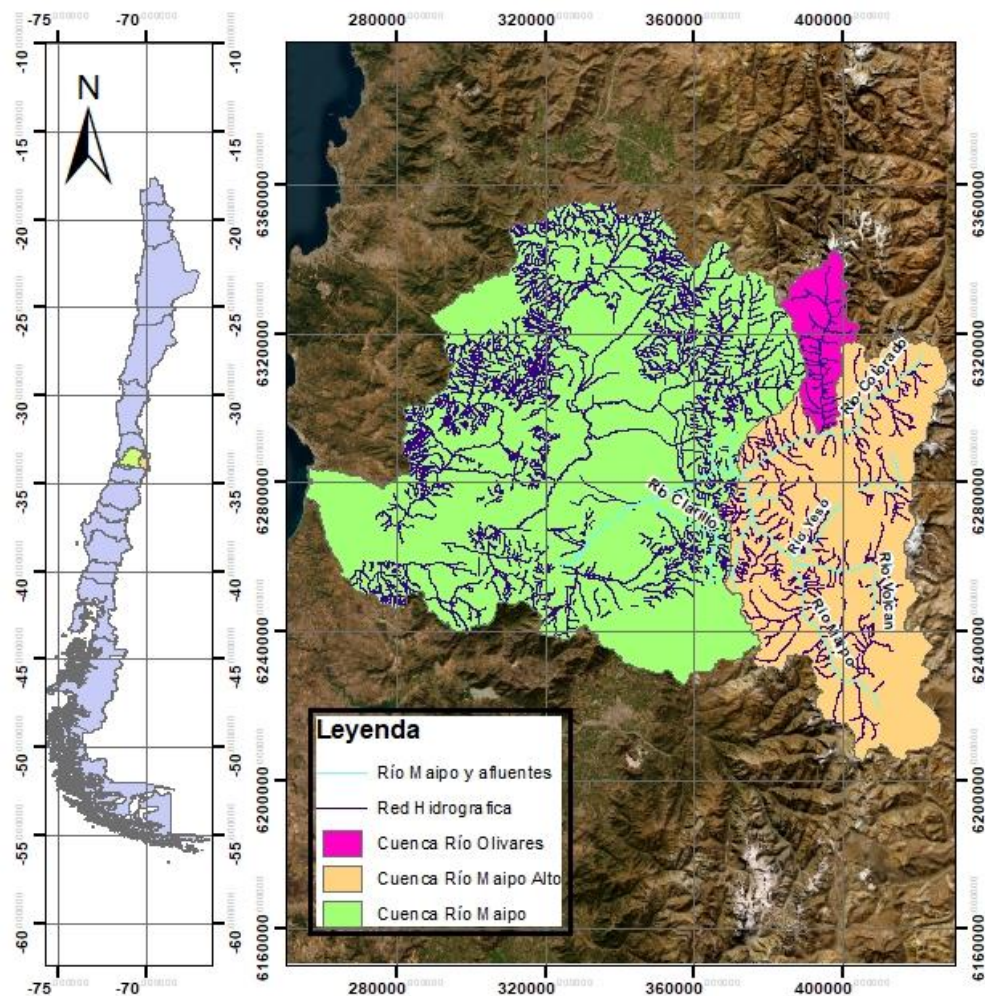


Figura 1. Mapa ilustrativo de la ubicación de cuenca de estudio. En la figura izquierda se muestra el mapa de Chile, donde se aprecia la cuenca Río Maipo, cuenca alta Río Maipo y la cuenca Río Olivares. En la figura derecha se aprecian de mejor forma las cuencas mencionadas con sus redes hidrográficas.

La cuenca del río Maipo presenta un régimen hidrológico dominado por el deshielo, como se evidencia en los gráficos de precipitación y caudal mensual presentados en las Figuras 2 y 3. En la Figura 2 se muestra que los meses de mayor precipitación son junio y julio, con valores que superan los 100 mm. La precipitación disminuye significativamente a partir de agosto, manteniéndose en niveles bajos de noviembre a marzo. Esto indica que la mayor parte de la precipitación ocurre durante los meses de invierno.

En contraste, en la Figura 3 donde se presentan los caudales medios, muestra que los meses de mayor flujo son diciembre y enero, con un aumento notable del caudal a partir de noviembre y un descenso gradual a partir de febrero. Este patrón de aumento de caudal en primavera y verano sugiere que el régimen hidrológico de la cuenca está dominado por el deshielo de nieve acumulada durante el invierno, lo que incrementa significativamente el flujo en los ríos y arroyos que alimentan la cuenca.

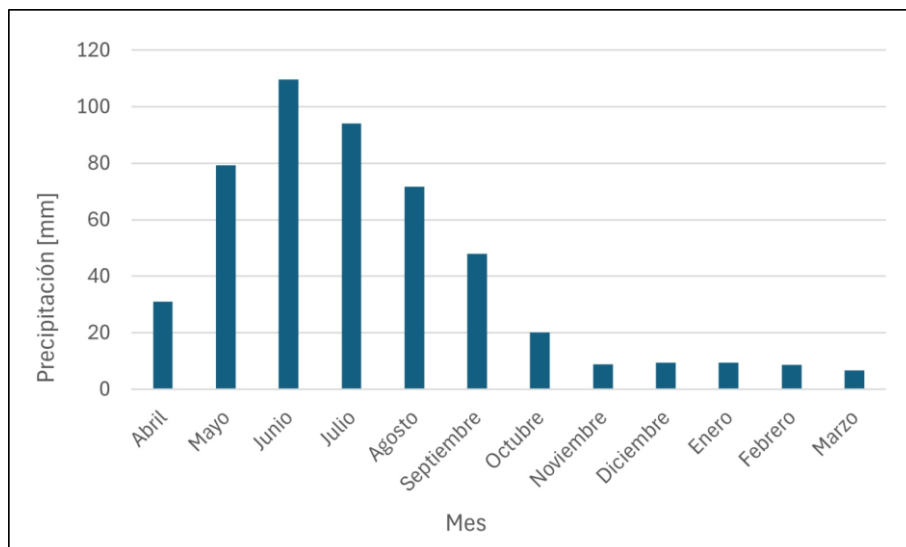


Figura 2. Precipitaciones medias mensuales en la cuenca alta del Río Maipo para el periodo 1978-2019.

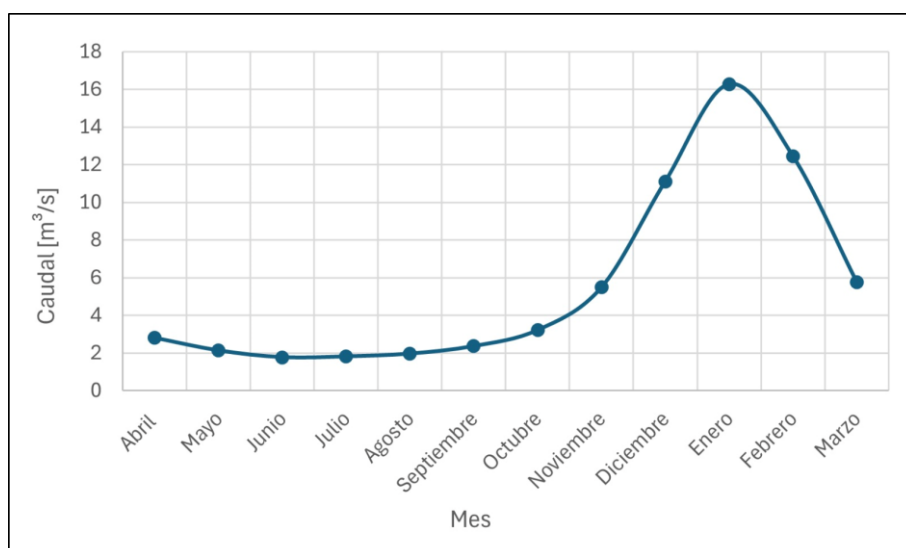


Figura 3. Caudales medios mensuales en la cuenca alta del Río Maipo para el periodo 1978-2019.

En resumen, la cuenca alta del río Maipo depende significativamente de un régimen hidrológico dominado por el deshielo, donde el derretimiento de nieve durante la primavera y el verano es el principal contribuyente al aumento del caudal, mientras que las precipitaciones invernales juegan un papel fundamental en la acumulación de agua que se liberará en los meses más cálidos. Este comportamiento destaca la importancia de un manejo adecuado de los recursos hídricos para satisfacer las demandas de agua potable y riego en la región.

5 Datos

5.1 Desfases temporales

En este estudio, se ha considerado la predicción de caudales para las estaciones de primavera y verano, utilizando dos desfases temporales: 0 meses y 3 meses previos. Para ambos desfases, se toman en cuenta los datos de los 3 meses anteriores a la estación en cuestión. En la Tabla 1 se detallan los meses específicos a considerar para cada estación y desfase.

Tabla 1. Meses considerados para cada desfase y estación.

Desfases	Estación	Meses
0-meses	Primavera	Junio, Julio y Agosto
	Verano	Septiembre, Octubre y Noviembre
3-meses	Primavera	Marzo, Abril y Mayo
	Verano	Junio, Julio y Agosto

5.2 Predictores

Entre los predictores seleccionados se incluyen variables climáticas, índices climáticos, y anomalías globales de la Temperatura Superficial del Mar (SST). A continuación, se describen en detalle cada uno de estos componentes y su relevancia en el análisis.

5.2.1 Variables Climáticas

Las variables climáticas son los factores o elementos que describen las condiciones del clima en una región particular. Estas variables se utilizan para estudiar, analizar y predecir el comportamiento del clima y sus efectos en el ambiente.

5.2.1.1 Caudal

El caudal es la medida del volumen de agua que fluye a través de una superficie o sección transversal en un tiempo determinado. En el contexto de ríos, el caudal indica la cantidad de agua que pasa por un punto específico, estos son medidos en estaciones fluviométricas. El caudal es fundamental en la ingeniería hidráulica y en la gestión de recursos hídricos (Valdivielso, 2023), ya que determina la capacidad de transporte de agua y es crucial para el diseño de infraestructuras como represas, canales, y sistemas de distribución de agua.

5.2.1.2 Precipitación

La precipitación es el término que se utiliza para describir cualquier forma de agua que cae desde la atmósfera hacia la superficie terrestre. Este fenómeno incluye lluvia, llovizna, nieve, aguanieve, granizo, pero no virga, neblina ni rocío, que son formas de condensación y no de precipitación. La precipitación es una parte importante del ciclo hidrológico, llevando agua dulce a la parte emergida de la corteza terrestre y, por ende, favoreciendo la vida en nuestro planeta, tanto de animales como de vegetales, que requieren agua para vivir (Instituto Geográfico de Aragón, 2017).

5.2.2 Índices Climáticos

Los índices climáticos son valores que se derivan de datos meteorológicos y climáticos para representar patrones complejos del clima. Pueden basarse en observaciones directas, como temperaturas y precipitaciones, o en modelos que combinan múltiples variables climáticas. Se utilizan para monitorear, analizar y predecir variaciones en el clima a lo largo del tiempo y en diferentes regiones (Trenberth & Fasullo, 2013).

5.2.2.1 El Niño – Oscilación del Sur

El Niño o también llamado Oscilación del Sur (ENOS) corresponde a un evento climático natural que se desarrolla en el océano Pacífico ecuatorial central. La fase cálida del ENOS es conocida como el fenómeno del Niño, el cual se manifiesta principalmente por un aumento de la Temperatura Superficial del Mar (TSM) y una disminución de los vientos alisios en el lado oriental del océano Pacífico. La fase

fría del ENOS, conocida como el fenómeno de la Niña, cuenta con TSM más frías de lo normal e intensifica los vientos alisios en el lado oriente del océano Pacífico (Maturana et al., 2004).

Para obtener mediciones estandarizadas y más precisas, se ha dividido el Pacífico ecuatorial en diferentes regiones, que corresponden a: El Niño 1+2 (0°-10°S, 90°O-80°O), Niño 3 (5°N-5°S, 150°O-90°O), Niño 3.4 (5°N-5°S, 170°O-120°O) y Niño 4 (150°O-160°E, 5°N-5°S) (Maturana et al., 2004). Montalva (2022) en su memoria de título, selecciona como predictores El Niño 1+2 y El Niño 3.4.

5.2.2.2 Oscilación Antártica

La Oscilación Antártica (AAO), también conocida como Modo Anular del Sur (SAM), es un patrón de variabilidad climática que describe las fluctuaciones en la presión atmosférica en el hemisferio sur. Se define como un cinturón de vientos del oeste de baja presión que rodea la Antártida. En su fase positiva, este modo anular se contrae y se acerca más a la Antártida (hacia el Sur), mientras que, en su fase negativa, el cinturón de vientos se desplaza hacia el ecuador (hacia el Norte) (Thompson & Wallace, 2000).

5.2.3 Anomalías Globales de la Temperatura Superficial del Mar (SST)

Las anomalías globales de la Temperatura Superficial del Mar (SST) son variaciones en la temperatura de la superficie del océano que se desvían de la media histórica para una región específica. Estas anomalías juegan un papel crucial en la climatología global y regional, ya que pueden influir significativamente en los patrones climáticos y en los eventos meteorológicos.

(Montalva, 2022) identificó dos zonas de alta correlación en la región del Pacífico Sur con el fin de evitar la detección de teleconexiones sin significado. La primera zona se ubica en el Pacífico Sureste, cerca de Sudamérica (SEPZ), y está delimitada por las coordenadas 260-280° E y 15-25° S. La segunda zona, en el Pacífico Sur Medio (MSPZ), está definida por las coordenadas 185-215° E y 10-45° S.

5.3 Correlaciones con predictores

Los predictores presentados anteriormente fueron escogidos por un análisis de correlación exploratorio realizado por (Montalva, 2022). En la Figura 4 presentada en el mismo estudio, se observan las correlaciones de todos los meses del año para los predictores AAO, Niño 3.4 y Niño 1+2.

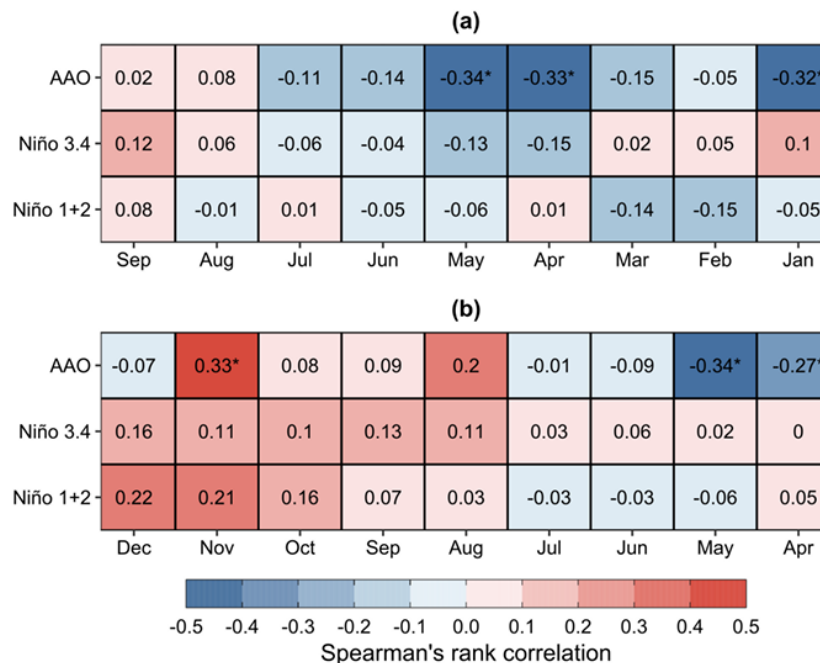


Figura 4. Correlación por rangos de Spearman entre el caudal estacional en la cuenca ORB y los índices climáticos tradicionales a gran escala con diferentes retrasos mensuales. (a) Caudal estacional de primavera. (b) Caudal estacional de verano. El símbolo * denota correlaciones significativas al nivel de confianza del 90%, (adaptada de Montalva (2022)).

En la Figura 4, se observa que el predictor AAO muestra correlaciones significativas, mientras que los otros predictores presentan correlaciones bajas, en algunos casos llegando a ser 0. Basado en estos resultados y en los obtenidos para los distintos predictores, se presentan a continuación la Tabla 2 y Tabla 3, que resumen los meses y las correlaciones a utilizar para cada predictor en las estaciones de primavera y verano, respectivamente.

Tabla 2. Resumen de meses y correlaciones para cada predictor en la estación de primavera.

PRIMAVERA				
Predictores	0-meses		3-meses	
	Meses	Correlación	Meses	Correlación
Caudal medio mensual	Agosto	0.901	Mayo	0.695
Precipitación acumulada	Abril-Agosto	0.502	Abril-Mayo	0.253
Niño 1+2	Febrero	-0.15	Febrero	-0.15
Niño 3.4	Abril	-0.15	Abril	-0.15
AAO	Mayo	-0.34	Mayo	-0.34
SEPZ	Junio-Agosto	0.370	Marzo-Mayo	0.344
MSPZ	Junio-Agosto	0.710	Marzo-Mayo	-0.571

Tabla 3. Resumen de meses y correlaciones para cada predictor en la estación de verano.

VERANO				
Predictores	0-meses		3-meses	
	Meses	Correlación	Meses	Correlación
Caudal medio mensual	Noviembre	0.695	Agosto	0.641
Precipitación acumulada	Abril-Agosto	0.181	Abril-Agosto	0.181
Niño 1+2	Noviembre	0.21	-	-
Niño 3.4	Septiembre	0.13	Agosto	0.11
AAO	Noviembre	0.33	Mayo	-0.34
SEPZ	Septiembre-Noviembre	0.320	Junio-Agosto	0.322
MSPZ	Septiembre-Noviembre	-0.523	Junio-Agosto	-0.541

En la estación de primavera, con un desfase de 0 meses, se decidió utilizar el caudal medio de agosto elevado al cuadrado como predictor. Se ha demostrado que esta transformación no lineal del caudal tiene una efectividad significativa en la reducción de sesgos no sistemáticos en las predicciones (Ossandón et al., 2022). Específicamente, se ha observado que el uso del caudal cuadrático ayuda a mitigar la subestimación de caudales altos y la sobrestimación de caudales bajos. Esto se traduce en una mejora en la precisión del modelo, ya que permite capturar de manera más fiel la variabilidad extrema del caudal en la cuenca.

6 Metodología

En esta sección se detallan las metodologías utilizadas para prever los caudales en la cuenca alta del río Maipo. Se implementaron dos enfoques principales: primero, la evaluación de modelos GAMLSS utilizando predictores seleccionados en base a su relevancia para la zona de estudio; y segundo, la comparación de estos modelos con un modelo de regresión lineal multivariable (MRLM) previamente identificado. Ambos enfoques permiten analizar y comparar la precisión y efectividad de los métodos de predicción en diferentes escenarios.

6.1 Evaluación de modelos GAMLSS utilizando predictores

En esta sección, se evaluaron Modelos Aditivos Generalizados para Ubicación, Escala y Forma (GAMLSS) empleando los predictores propuestos por Montalva (2022) para la cuenca alta del río Maipo. Estos predictores, seleccionados en función de su relevancia para la zona de estudio, fueron integrados en los modelos GAMLSS para prever los caudales durante las estaciones de primavera y verano.

Se llevó a cabo un análisis exhaustivo de las distribuciones de probabilidad adecuadas para modelar los caudales. En la Tabla 4 se muestran las distribuciones continuas con variables de respuestas positivas que se consideraron en los GAMLSS (Stasinopoulos et al., 2017): Gamma, Gamma Generalizada, Log Normal y Log Normal 2.

Tabla 4. Distribuciones utilizadas en los modelos GAMLSS.

Distribución	Gamlss name	Función de enlace de un parámetro			
		μ	σ	ν	τ
Gamma Generalizada	GG	log	log	ident	-
Gamma	GA	log	log	-	-
Log Normal	LOGNO	ident	log	-	-
Log Normal 2	LOGNO2	log	log	-	-

Para cada una de estas distribuciones, se realizó la predicción de caudales utilizando un parámetro de localización (μ) no estacionario, mientras que el parámetro de escala (σ) se mantuvo estacionario.

Para identificar los modelos más adecuados, se realizaron 1000 simulaciones para cada combinación de predictores y distribuciones dentro de los modelos GAMLSS. Estas simulaciones permiten predecir 1000 veces los caudales esperados, proporcionando un conjunto amplio de resultados posibles que reflejan la variabilidad e incertidumbre inherentes en los datos hidrológicos. A partir de estas simulaciones, se calcularon los percentiles para cada predicción, lo cual permitió establecer los intervalos de confianza (IC) que indican el rango dentro del cual es más probable que se encuentren los caudales reales. Estos IC son cruciales para evaluar la incertidumbre y la precisión de los modelos, ya que muestran la dispersión de las predicciones y ayudan a identificar los modelos que mejor capturan la variabilidad observada en los caudales durante las estaciones de primavera y verano.

Para determinar cuál de las cuatro distribuciones era la más adecuada para cada desfase temporal y estación, se emplearon métricas de evaluación de desempeño (AIC, R de Pearson, BIAS y CRPSS), donde se usó como referencia la climatología de la cuenca alta del río Maipo. Esta metodología permitió cuantificar la habilidad predictiva de los modelos GAMLSS (y cada distribución) en comparación con un escenario climatológico, permitiendo comparar la efectividad de cada distribución en la predicción de los caudales y facilitando así la selección del modelo más preciso para cada caso específico.

6.2 Validación cruzada de los modelos

La validación cruzada es un método fundamental para evaluar la robustez y generalización de los modelos estadísticos. Para esta memoria se implementó la “Leave One Out Cross Validation” (LOOCV), un método en el que cada observación del conjunto de datos se deja fuera una vez para probar el modelo, mientras el resto de los datos se utiliza para entrenarlo. Este proceso se repite hasta que cada observación ha sido utilizada como conjunto de prueba una vez, permitiendo obtener una evaluación completa del desempeño del modelo en base a todas las observaciones disponibles. LOOCV maximiza el uso de los datos, ya que casi todo el conjunto de datos se emplea en el entrenamiento en cada iteración, lo que reduce el sesgo y mejora la capacidad del modelo para predecir los caudales estacionales.

Una vez seleccionado el mejor modelo, se generaron gráficos comparativos que muestran los caudales observados y pronosticados, junto con intervalos de confianza del 5%-95% y 25%-75%. Estos intervalos de confianza son fundamentales para evaluar la incertidumbre asociada con las predicciones, proporcionando una comprensión visual de qué tan bien los modelos capturan la variabilidad en los caudales. Esta información permite realizar una evaluación más precisa de la efectividad del modelo en cada situación específica.

6.3 Comparación con modelo de regresión lineal multivariable (MRLM)

Después de identificar la distribución más adecuada en la sección anterior, se realizó una comparación con un modelo de regresión lineal multivariable (MRLM) identificado por Montalva (2022) para la cuenca alta del río Maipo. Esta comparación tuvo como objetivo evaluar el desempeño predictivo de la distribución seleccionada en los modelos GAMLSS frente al modelo MRLM. Para esta comparación, se llevaron a cabo dos enfoques distintos:

En el primer enfoque se calcularon y compararon las métricas de desempeño de cada modelo utilizando la climatología de la cuenca alta del río Maipo como referencia. Estas métricas, que incluyen el R de Pearson, BIAS, y el CRPSS, permitieron evaluar la habilidad de cada modelo para mejorar sobre las predicciones climatológicas estándar, proporcionando un punto de referencia común para medir la precisión y fiabilidad de las predicciones de caudales.

En el segundo enfoque, se utilizaron dos métricas fundamentales: el coeficiente de eficiencia (CE) y el Continuous Ranked Probability Skill Score (CRPSS). En este caso, el modelo MRLM se utilizó directamente como referencia. El coeficiente de eficiencia permitió medir la capacidad de cada enfoque para replicar los caudales observados, mientras que el CRPSS evaluó la habilidad predictiva de la distribución GAMLSS seleccionada en comparación con el MRLM, proporcionando una medida clara del rendimiento relativo entre ambos modelos.

7 Resultados y Análisis

7.1 Análisis Modelos GAMLSS

7.1.1 Modelo Calibrado

En este estudio, se calibraron un total de 127 modelos para cada una de las cuatro distribuciones de probabilidad analizadas: Gamma, Gamma Generalizada, Log Normal y Log Normal 2. Estos modelos se construyeron utilizando combinaciones de los siete predictores: Caudal medio, Precipitación acumulada, Niño 1+2, Niño 3.4, AAO, SEPZ y MSPZ. Para cada distribución, se evaluaron 127 modelos distintos, lo que resulta en un total de 508 modelos calibrados.

La selección del mejor modelo dentro de cada distribución, la cual se llevó a cabo mediante un análisis detallado de cuatro métricas de evaluación: el Criterio de Información de Akaike (AIC), el coeficiente de correlación de Pearson (R), el sesgo (BIAS) y el Continuous Ranked Probability Skill Score (CRPSS).

En la Tabla 5 se presentan los mejores modelos calibrados para cada distribución, teniendo en cuenta la temporada (Primavera y Verano) y el desfase (0-meses y 3-meses):

Tabla 5. Predictores y métricas de evaluación de desempeño para el mejor modelo GAMLSS calibrado de cada distribución obtenido en las dos diferentes estaciones y los tiempos de desfase.

Temporada	Desfase	Distribución	Predictores	AIC	R	BIAS [%]	CRPSS
Primavera	0-meses	Gamma Generalizada	$Q_{Ago}^2, Niño\ 3.4, AAO, MSPZ$	88,59	0,96	2,63	0,71
		Gamma	$Q_{Ago}^2, Niño\ 3.4, AAO, MSPZ$	89,36	0,96	0,28	0,83
		Log Normal	$Q_{Ago}^2, Niño\ 3.4, AAO, MSPZ$	112,81	0,89	-1,14	0,71
		Log Normal 2	$Q_{Ago}^2, Niño\ 3.4, AAO, SEPZ, MSPZ$	112,57	0,88	-0,25	0,69
	3-meses	Gamma Generalizada	$Q_{May}, P_{A-M}, MSPZ$	147,72	0,61	-19,18	0,54
		Gamma	$Q_{May}, P_{A-M}, MSPZ$	142,94	0,61	-3,99	0,53
		Log Normal	$Q_{May}, P_{A-M}, MSPZ$	146,40	0,59	-15,39	0,53
		Log Normal 2	$Q_{May}, P_{A-M}, MSPZ$	146,40	0,57	-15,10	0,52
Verano	0-meses	Gamma Generalizada	$Q_{Nov}, Niño\ 3.4, MSPZ$	239,12	0,72	-12,36	0,56
		Gamma	$Q_{Nov}, Niño\ 3.4, MSPZ$	235,35	0,75	-8,49	0,60
		Log Normal	$Q_{Nov}, Niño\ 3.4, MSPZ$	237,20	0,71	-11,18	0,57
		Log Normal 2	$Q_{Nov}, Niño\ 3.4, MSPZ$	237,20	0,71	-11,50	0,57
	3-meses	Gamma Generalizada	$Q_{Ago}, P_{A-A}, AAO, SEPZ$	241,04	0,73	-10,63	0,59
		Gamma	$Q_{Ago}, P_{A-A}, AAO, MSPZ$	238,85	0,79	-9,60	0,62
		Log Normal	$Q_{Ago}, P_{A-A}, AAO, SEPZ$	239,26	0,73	-10,2	0,59
		Log Normal 2	$Q_{Ago}, P_{A-A}, AAO, SEPZ$	239,26	0,74	-10,45	0,59

El análisis de los resultados obtenidos para los modelos GAMLSS muestra un rendimiento diverso según la temporada y el desfase temporal considerado. Para la temporada de primavera con un desfase de 0 meses, las distribuciones Gamma Generalizada y Gamma presentan un desempeño similar en términos del coeficiente de correlación de Pearson ($R \approx 0,96$), lo que indica una alta correlación entre los valores predichos y observados. Sin embargo, la distribución Gamma Generalizada presenta un AIC ligeramente menor (88,59 frente a 89,36), lo que sugiere un mejor ajuste del modelo. En cuanto al sesgo (BIAS), la distribución Gamma se destaca por su baja sobrestimación (0,28%), siendo más precisa que la Gamma Generalizada (2,63%). Además, el CRPSS es más alto para la distribución Gamma (0,83), lo que sugiere un mejor rendimiento probabilístico. Los predictores seleccionados para ambas distribuciones en primavera con un desfase de 0 meses incluyen el caudal de agosto al cuadrado (Q_{Ago}^2), Niño 3.4, AAO y MSPZ, lo que refleja la importancia de estos factores hidrometeorológicos y climáticos para capturar la variabilidad de los caudales en esta estación.

Para la primavera con un desfase de 3 meses, las distribuciones Gamma Generalizada y Gamma nuevamente destacan por sus buenos resultados, con un AIC menor en comparación con las distribuciones Log Normal. La distribución Gamma tiene un sesgo menor (-3,99%) y un CRPSS similar al de la Gamma Generalizada, lo que la convierte en la mejor opción para este desfase temporal. Por

otro lado, las distribuciones Log Normal y Log Normal 2 muestran mayores sesgos (-15,39% y -15,10% respectivamente), lo que indica una mayor subestimación en los pronósticos y, por lo tanto, menor precisión. Los predictores más relevantes para este desfase incluyen el caudal de mayo (Q_{May}), la precipitación acumulada en abril-mayo (P_{A-M}) y MSPZ, lo que sugiere que las condiciones previas de caudal y precipitación son fundamentales para prever los caudales con este desfase.

En la temporada de verano con un desfase de 0 meses, la distribución Gamma sobresale con el mejor AIC (235,35) y el mayor CRPSS (0,60), lo que indica no solo un buen ajuste, sino también una mejor capacidad predictiva en comparación con las otras distribuciones. Aunque las distribuciones Log Normal y Log Normal 2 presentan un CRPSS menor (0,57), el sesgo es también mayor, reforzando la robustez de la distribución Gamma para esta estación. Los predictores claves para verano con un desfase de 0 meses son el caudal de noviembre (Q_{Nov}), Niño 3.4 y MSPZ, lo que indica que las condiciones de caudal reciente, junto con factores climáticos como Niño 3.4 y patrones oceánicos en el Pacífico, juegan un papel crucial en la modelación de los caudales.

Finalmente, en verano con un desfase de 3 meses, la distribución Gamma nuevamente se presenta como la mejor opción, con un AIC más bajo (238,85) y el mayor CRPSS (0,62), superando a las otras distribuciones. El sesgo también es menor en comparación con las distribuciones Log Normal y Log Normal 2, lo que sugiere que la Gamma proporciona pronósticos más precisos para este desfase. En este caso, los predictores más significativos incluyen el caudal de agosto (Q_{Ago}), la precipitación acumulada en abril-agosto (P_{A-A}), AAO y SEPZ, lo que destaca la relevancia de las condiciones climáticas y oceánicas más amplias para los pronósticos a mayor anticipación.

7.1.2 Validación Cruzada

La validación cruzada es una etapa fundamental en la evaluación de modelos predictivos, ya que permite verificar la capacidad del modelo para generalizar a datos no utilizados en su calibración inicial. En este contexto, la validación cruzada ayuda a asegurar que los modelos no estén sobreajustados a los datos de entrenamiento y que puedan ofrecer predicciones fiables y precisas en situaciones reales. En la Tabla 6 se presentan los resultados de la validación cruzada para las diferentes distribuciones evaluadas en las temporadas de primavera y verano, considerando dos desfases temporales (0 y 3 meses). Las métricas utilizadas incluyen el coeficiente de correlación de Pearson (R), el sesgo (BIAS) y el Continuous Ranked Probability Skill Score (CRPSS).

Tabla 6. Predictores y métricas de evaluación de desempeño para el mejor modelo GAMLSS en validación cruzada de cada distribución obtenido en las dos diferentes estaciones y los tiempos de desfase.

Temporada	Desfase	Distribución	Predictores	R	BIAS [%]	CRPSS
Primavera	0-meses	Gamma Generalizada	$Q_{Ago}^2, Niño\ 3.4, AAO, MSPZ$	0,95	2,66	0,63
		Gamma	$Q_{Ago}^2, Niño\ 3.4, AAO, MSPZ$	0,95	0,29	0,81
		Log Normal	$Q_{Ago}^2, Niño\ 3.4, AAO, MSPZ$	0,84	2,48	0,65
		Log Normal 2	$Q_{Ago}^2, Niño\ 3.4, AAO, SEPZ, MSPZ$	0,83	2,76	0,63
	3-meses	Gamma Generalizada	$Q_{May}, P_{A-M}, MSPZ$	0,48	-18,07	0,45
		Gamma	$Q_{May}, P_{A-M}, MSPZ$	0,56	-3,85	0,49
		Log Normal	$Q_{May}, P_{A-M}, MSPZ$	0,46	-11,16	0,45
		Log Normal 2	$Q_{May}, P_{A-M}, MSPZ$	0,46	-11,19	0,45
Verano	0-meses	Gamma Generalizada	$Q_{Nov}, Niño\ 3.4, MSPZ$	0,61	-8,94	0,49
		Gamma	$Q_{Nov}, Niño\ 3.4, MSPZ$	0,72	-8,46	0,56
		Log Normal	$Q_{Nov}, Niño\ 3.4, MSPZ$	0,61	-8,41	0,50
		Log Normal 2	$Q_{Nov}, Niño\ 3.4, MSPZ$	0,63	-8,23	0,50
	3-meses	Gamma Generalizada	$Q_{Ago}, P_{A-A}, AAO, SEPZ$	0,65	-11,11	0,51
		Gamma	$Q_{Ago}, P_{A-A}, AAO, MSPZ$	0,73	-9,27	0,56
		Log Normal	$Q_{Ago}, P_{A-A}, AAO, SEPZ$	0,65	-9,50	0,52
		Log Normal 2	$Q_{Ago}, P_{A-A}, AAO, SEPZ$	0,65	-9,57	0,52

Los resultados de la validación cruzada confirman en gran medida las conclusiones obtenidas en la calibración inicial. En general, la distribución Gamma sigue destacándose por su rendimiento superior en la mayoría de los casos. Para la primavera con un desfase de 0 meses, la distribución Gamma presenta un coeficiente de correlación (R) de 0,95, un BIAS de 0,29%, y un CRPSS de 0,81, reafirmando su capacidad para capturar adecuadamente la variabilidad de los caudales sin introducir un sesgo significativo. Este rendimiento consistente en la validación cruzada subraya la robustez del modelo Gamma, no solo en términos de ajuste a los datos de entrenamiento, sino también en su capacidad para predecir con precisión caudales futuros.

En la primavera con un desfase de 3 meses, aunque las correlaciones disminuyen en todas las distribuciones debido a la mayor incertidumbre asociada con un mayor desfase temporal, la distribución Gamma sigue mostrando un rendimiento ligeramente superior con un R de 0,56 y un CRPSS de 0,49, lo que indica que, a pesar de las dificultades, este modelo mantiene una ventaja sobre las demás distribuciones.

Para la temporada de verano con un desfase de 0 meses, la distribución Gamma nuevamente se destaca con un R de 0,72 y un CRPSS de 0,56. Aunque las distribuciones Log Normal y Log Normal 2 presentan valores similares de R, su CRPSS es ligeramente inferior, lo que indica una menor capacidad predictiva y sugiere que estos modelos pueden ser menos confiables en la predicción probabilística de caudales.

Finalmente, en verano con un desfase de 3 meses, la distribución Gamma sigue superando a las otras distribuciones con un R de 0,73 y un CRPSS de 0,56. La validación cruzada muestra que, incluso cuando se predice con un desfase temporal más largo, la distribución Gamma logra mantener un buen equilibrio entre precisión y fiabilidad en la predicción de caudales.

7.1.3 Selección del mejor modelo para cada temporada y desfase

A continuación, se presenta el mejor modelo para cada combinación de temporada y desfase temporal, basado tanto en los resultados de calibración como en la validación cruzada. Dado que la distribución Gamma ha mostrado consistentemente el mejor rendimiento en ambas etapas, se utilizará esta distribución para todos los modelos seleccionados.

La Tabla 7 detalla los valores obtenidos durante la calibración de los modelos, mientras que la Tabla 8 resume los resultados de la validación cruzada. Estas tablas permitirán una comparación clara y directa de cómo los modelos se comportan tanto en la fase de calibración como en la validación, proporcionando una visión integral de su desempeño.

Los modelos con un desfase de 0 meses son superiores tanto en la calibración como en la validación cruzada para la primavera. La precisión y la capacidad predictiva disminuyen significativamente con un mayor tiempo de desfase, sugiriendo que el modelo de 3 meses tiene dificultades para captar la variabilidad de los caudales con mayor anticipación.

Tabla 7. Resumen de predictores y métricas de evaluación de desempeño para el mejor modelo GAMLSS calibrado y mejor distribución obtenidos en las dos diferentes estaciones y los tiempos de desfase.

Temporada	Desfase	Distribución	Predictores	AIC	R	BIAS [%]	CRPSS
Primavera	0-meses	Gamma	$Q_{Ago}^2, Niño\ 3.4, AAO, MSPZ$	89,36	0,96	0,28	0,83
	3-meses	Gamma	$Q_{May}, P_{A-M}, MSPZ$	142,94	0,61	-3,99	0,53
Verano	0-meses	Gamma	$Q_{Nov}, Niño\ 3.4, MSPZ$	235,35	0,75	-8,49	0,60
	3-meses	Gamma	$Q_{Ago}, P_{A-A}, AAO, MSPZ$	238,85	0,79	-9,60	0,62

En la comparación entre los modelos de verano muestra que, aunque el modelo con 3 meses de desfase tiene una correlación ligeramente mejor (R más alto), presenta un mayor sesgo en la subestimación de los caudales, lo que podría ser problemático en situaciones donde se necesita una predicción precisa de caudales altos. Los modelos de 0 meses muestran un mejor equilibrio entre precisión (menor BIAS) y capacidad predictiva en la mayoría de los casos.

Tabla 8. Resumen de predictores y métricas de evaluación de desempeño para el mejor modelo GAMLSS en validación cruzada y mejor distribución obtenidos en las dos diferentes estaciones y los tiempos de desfase.

Temporada	Desfase	Distribución	Predictores	R	BIAS [%]	CRPSS
Primavera	0-meses	Gamma	$Q_{Ago}^2, Niño\ 3.4, AAO, MSPZ$	0,95	0,29	0,81
	3-meses	Gamma	$Q_{May}, P_{A-M}, MSPZ$	0,56	-3,85	0,49
Verano	0-meses	Gamma	$Q_{Nov}, Niño\ 3.4, MSPZ$	0,72	-8,46	0,56
	3-meses	Gamma	$Q_{Ago}, P_{A-A}, AAO, MSPZ$	0,73	-9,27	0,56

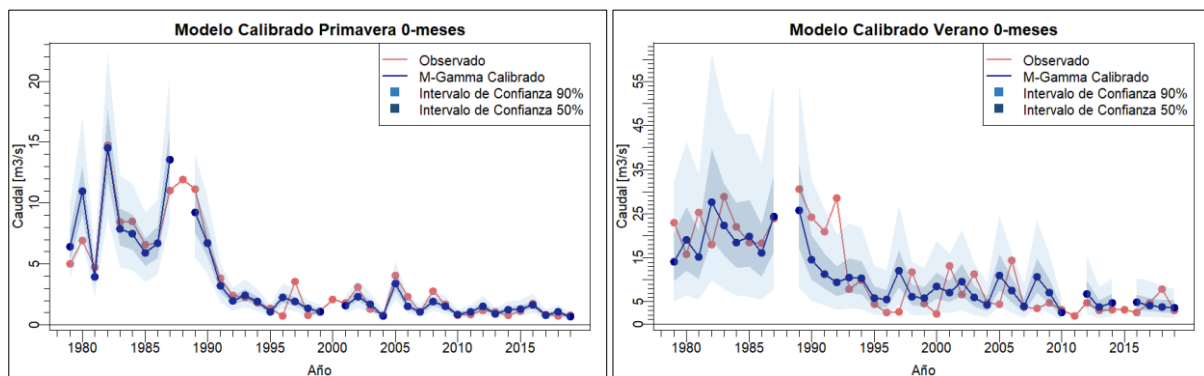


Figura 5. Serie temporal del caudal modelado y observado para primavera y verano con un desfase de 0 meses. El modelo GAMLSS con la mejor distribución Gamma presenta un parámetro de localización (μ) no estacionario y un parámetro de escala (σ) estacionario.

Como se observa en Figura 5, se presentan los gráficos para las estaciones de primavera y verano con un desfase de 0 meses. Para primavera, el modelo de desfase 0 muestra intervalos de confianza más estrechos hacia los años recientes, especialmente en el intervalo del 50%, lo que refleja una mayor certeza en la predicción de caudales medios. Sin embargo, durante los primeros años y en periodos de alta variabilidad (1980-1990), los intervalos del 90% se amplían, indicando mayor incertidumbre en esos casos. En comparación, el modelo de desfase de 3 meses mantiene intervalos del 90% más amplios hasta los años 2000, y su precisión se ve reducida, especialmente en los caudales altos (Anexo 1).

Para el verano, el análisis es similar: el modelo con desfase de 0 meses tiene intervalos más estrechos en años recientes y muestra una precisión razonable en la mediana de los caudales pronosticados. Los modelos de 3 meses, aunque presentan mejor correlación (R), muestran mayor sesgo en la estimación, lo que podría ser problemático para predicciones de caudales altos. En ambos casos, los intervalos del 90% son más amplios en la validación cruzada, sugiriendo una mayor incertidumbre cuando se aplican a datos nuevos (Anexo 1).

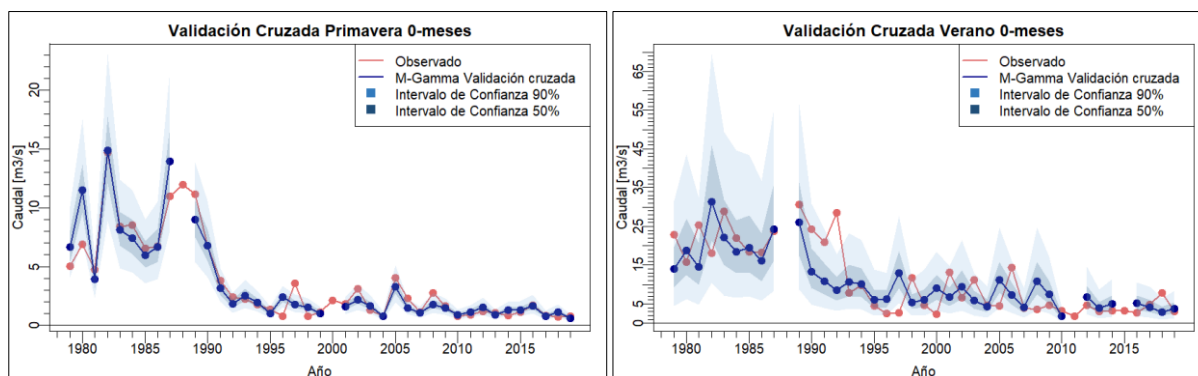


Figura 6. Serie temporal del caudal modelado y observado para primavera y verano con un desfase de 0 meses en validación cruzada. El modelo GAMLSS con la mejor distribución Gamma presenta un parámetro de localización (μ) no estacionario y un parámetro de escala (σ) estacionario.

Como se muestra en la Figura 6, se presentan los gráficos para las estaciones de primavera y verano con un desfase de 0 meses en validación cruzada. En primavera, el modelo con desfase de 0 meses presenta intervalos de confianza del 90% más amplios en comparación con el modelo calibrado, especialmente en años con mayor variabilidad de caudal, lo cual sugiere una incertidumbre adicional cuando el modelo es validado en datos nuevos. No obstante, el intervalo de confianza del 50% permanece relativamente estrecho, lo que indica que el modelo retiene su habilidad para capturar adecuadamente la mediana de los caudales. En contraste, el modelo de 3 meses muestra una mayor amplitud en ambos intervalos de confianza, reflejando una menor precisión en la predicción de la mediana y una mayor incertidumbre en años de alta variabilidad de caudal (Anexo 2).

Para el verano, el comportamiento es similar, el modelo de 0 meses presenta intervalos de confianza del 90% más amplios bajo validación cruzada en comparación con el calibrado, aunque sigue manteniendo cierta precisión en la mediana. En el caso del desfase de 3 meses, los intervalos de confianza son aún más amplios, particularmente en periodos de caudales altos, lo que indica una menor fiabilidad en la predicción de estos valores extremos (Anexo 2).

7.2 Comparación con Modelo de Regresión Lineal Multivariable (MRLM)

En esta sección, se realiza una comparación entre los modelos GAMLSS y un modelo de Regresión Lineal Multivariable (MRLM) (adaptado por Montalva, 2023), utilizando los datos de calibración para diferentes estaciones y tiempos de anticipación. El análisis se centra en la evaluación del rendimiento de los modelos usando métricas como el R, BIAS, y CRPSS, con el MRLM actuando como modelo de referencia en lugar de la climatología, como se hizo en secciones anteriores.

Para realizar una comparación detallada entre los modelos GAMLSS y el modelo de Regresión Lineal Multivariable (MRLM), se analiza cómo se comportan ambos tipos de modelos en términos de sus métricas de evaluación.

Comparando los resultados de la Tabla 9 (adaptada de Montalva (2022)) y los de la Tabla 5, se puede mencionar que, en primavera con un desfase de 0 meses, tanto los modelos GAMLSS como el MRLM muestran altos coeficientes de correlación (R). El MRLM presenta un R de 0,96, igual al mejor modelo GAMLSS (Gamma), lo que indica una precisión similar en la correlación entre los caudales observados y predichos. Sin embargo, hay diferencias importantes en otras métricas. En cuanto al BIAS, el MRLM tiene un valor de 1,10%, lo que indica una ligera sobreestimación. Por otro lado, el modelo GAMLSS Gamma presenta un BIAS mucho menor de 0,28%, indicando un mejor equilibrio en las predicciones. Además, el modelo GAMLSS Gamma tiene un CRPSS de 0,83, superior al CRPSS de 0,71 del MRLM, lo que refleja una mejor habilidad predictiva en la distribución completa de los caudales.

Para primavera con un desfase de 3 meses, el rendimiento de ambos modelos disminuye, pero el impacto es más pronunciado en el MRLM. El modelo GAMLSS con distribución Gamma tiene un R de 0,61,

similar al del MRLM ($R = 0,67$), pero el CRPSS del MRLM baja considerablemente a 0,25, mientras que el del modelo GAMLSS Gamma es de 0,53. Esta diferencia en CRPSS sugiere que, aunque ambos modelos pierden precisión con el tiempo de anticipación, el modelo GAMLSS Gamma maneja mejor la predicción de caudales a largo plazo en términos de habilidad predictiva global. El BIAS del MRLM para este desfase es de 3,91%, lo que implica una tendencia a la sobreestimación, mientras que el modelo GAMLSS Gamma tiene un BIAS de -3,99%, indicando una ligera subestimación, pero con menor variabilidad en la predicción.

Para el verano con 0 meses de desfase, el modelo MRLM presenta un R de 0,83, que es ligeramente inferior al R de 0,75 del modelo GAMLSS Gamma. Por otro lado, el CRPSS del MRLM es de 0,46, menor que el 0,60 del GAMLSS Gamma, lo que indica que, aunque el MRLM se ajusta mejor en términos de simplicidad del modelo, el GAMLSS ofrece una mejor habilidad predictiva. En términos de BIAS, el MRLM tiene un 3,81%, lo que indica una sobreestimación, mientras que el modelo GAMLSS Gamma muestra un BIAS de -8,49%, lo que representa una subestimación más significativa. Esto sugiere que el MRLM puede proporcionar predicciones más equilibradas, mientras que el GAMLSS Gamma podría ser más adecuado para capturar la variabilidad del caudal en una gama más amplia de escenarios.

Para verano con un desfase de 3 meses, ambos modelos experimentan una reducción en el rendimiento, pero el MRLM mantiene un R de 0,80, que es comparable al R de 0,79 del modelo GAMLSS Gamma. En términos de CRPSS, el MRLM tiene un 0,41, mientras que el GAMLSS Gamma muestra un CRPSS de 0,62, indicando que el GAMLSS es significativamente más efectivo para manejar la incertidumbre y proporcionar predicciones más útiles para la toma de decisiones en este escenario. El BIAS del MRLM (4,86%) sugiere una sobreestimación en comparación con el BIAS de -9,60% del modelo GAMLSS Gamma, la cual indica una subestimación más significativa.

Tabla 9. Predictores y métricas de rendimiento para el mejor MRLM obtenido de la calibración para diferentes estaciones y tiempos de anticipación (adaptada de Montalva (2022)).

Temporada	Desfase	Predictores	R	BIAS [%]	CRPSS
Primavera	0-meses	$Q_{Ago}^2, P_{A-A}, MSPZ$	0,96	1,10	0,71
	3-meses	$Q_{May}, MSPZ$	0,67	3,91	0,25
Verano	0-meses	$Q_{Nov}, P_{A-A}, AAO, MSPZ$	0,83	3,81	0,46
	3-meses	$Q_{Ago}, P_{A-A}, AAO, MSPZ$	0,80	4,86	0,41

La Tabla 10 presenta la comparación directa entre los modelos GAMLSS y MRLM utilizando el coeficiente de eficiencia (CE) y el CRPSS, tomando el MRLM como modelo de referencia. Se muestra que, en primavera con un desfase de 0 meses, el modelo GAMLSS tiene un CE de -0,11 y un CRPSS de 0,27. El coeficiente de eficiencia (CE) negativo sugiere que el modelo GAMLSS es menos eficiente que el MRLM en términos de precisión de las predicciones. Sin embargo, el CRPSS positivo de 0,27 indica que, en este caso, el modelo GAMLSS tiene una mejor capacidad predictiva en comparación con el MRLM, especialmente en capturar la variabilidad de los caudales primaverales en un corto tiempo de anticipación.

Con un desfase de 3 meses para la primavera, el modelo GAMLSS presenta un CE de -0,21 y un CRPSS de 0,27. Aunque el CE negativo indica una menor eficiencia comparada con el MRLM, el CRPSS positivo de 0,27 sugiere que el modelo GAMLSS sigue superando al MRLM en términos de habilidad predictiva global, incluso cuando aumenta el tiempo de anticipación. Esto refleja la capacidad del modelo GAMLSS para capturar la incertidumbre y la variabilidad de los caudales en la primavera a medida que el desfase temporal se extiende.

En el caso del verano con un desfase de 0 meses, el modelo GAMLSS tiene un CE de -0,42, lo que indica una eficiencia significativamente menor en comparación con el MRLM. A pesar de esto, el CRPSS positivo de 0,17 sugiere que el modelo GAMLSS aún posee una mejor capacidad predictiva que el MRLM. Sin embargo, la diferencia es menos pronunciada que en los casos de primavera, lo que

podría indicar que el modelo GAMLSS es menos efectivo para capturar la variabilidad en los caudales de verano con un tiempo de anticipación corto.

Para el verano con un desfase de 3 meses, el modelo GAMLSS muestra un CE de -0,10, indicando una eficiencia ligeramente menor que la del MRLM. Sin embargo, con un CRPSS positivo de 0,27, se demuestra que el modelo GAMLSS tiene una mejor capacidad predictiva que el MRLM para este tiempo de anticipación, a pesar de la disminución en la eficiencia general del modelo. Esto resalta la capacidad del GAMLSS para manejar la variabilidad e incertidumbre en los caudales de verano incluso con un aumento en el tiempo de desfase.

Tabla 10. Comparación entre modelos GAMLSS y el modelo MRLM. Utilizando métricas de evaluación usando como modelo de referencia el modelo MRLM.

Temporada	Desfase	CE	CRPSS
Primavera	0-meses	-0,11	0,27
	3-meses	-0,21	0,27
Verano	0-meses	-0,42	0,17
	3-meses	-0,10	0,27

8 Conclusiones

En este estudio, se han evaluado distintos modelos GAMLSS y se ha comparado con un modelo de Regresión Lineal Multivariable (MRLM) para prever caudales en las temporadas de primavera y verano, utilizando diferentes desfases temporales (0 y 3 meses). A lo largo del análisis, se han identificado varias conclusiones clave sobre el rendimiento de estos modelos y sus limitaciones.

Al comparar las distintas distribuciones GAMLSS, se observa que las distribuciones Log Normal y Log Normal 2 presentan un problema inherente: la transformación logarítmica aplicada a la variable durante el ajuste. Esta transformación puede mejorar el ajuste en el espacio logarítmico, pero cuando se revierte a su escala original, la precisión del modelo disminuye. Esto se debe a que la transformación tiende a amplificar las diferencias, lo que resulta en un mayor sesgo y menor habilidad predictiva cuando los resultados se interpretan en la escala original. Este efecto se evidencia en los mayores valores de BIAS y CRPSS más bajos para las distribuciones Log Normal en comparación con las distribuciones Gamma y Gamma Generalizada, especialmente en los desfases de 3 meses.

Por otro lado, la distribución Gamma ha demostrado ser la mejor opción en la mayoría de los escenarios analizados. Esta distribución ha mostrado consistentemente un buen equilibrio entre el ajuste del modelo (AIC bajo), la precisión (R alto y BIAS bajo) y la capacidad predictiva global (CRPSS alto). La razón por la cual la distribución Gamma es superior se debe a su flexibilidad para modelar datos positivos sesgados, como es el caso de los caudales, sin necesidad de transformaciones que puedan distorsionar los resultados en su escala original. La distribución Gamma es capaz de capturar de manera más precisa la variabilidad natural de los caudales, lo que resulta en una mejor representación tanto de la media como de la dispersión de los datos observados.

Los resultados indican que la capacidad predictiva de los modelos GAMLSS es generalmente mejor para la primavera que para el verano, especialmente para desfases de 0-meses. En primavera con un desfase de 0 meses, los modelos muestran altos coeficientes de correlación y bajos sesgos, mientras que, en verano, la capacidad predictiva disminuye notablemente, con menores valores de CRPSS y mayores sesgos. Esto sugiere que los patrones climáticos y los factores hidrometeorológicos de la primavera son más predecibles o están mejor representados en los datos utilizados que los del verano, lo cual podría estar relacionado con una mayor estabilidad en las condiciones de caudal durante la primavera.

A medida que aumenta el tiempo de desfase (de 0 a 3 meses), se observa una pérdida significativa en la capacidad predictiva de los modelos, tanto en términos de precisión (R y BIAS) como de habilidad predictiva (CRPSS). Esta tendencia sugiere que la incertidumbre y la variabilidad en los factores que afectan los caudales se incrementan con el tiempo, haciendo más difícil para los modelos prever correctamente los caudales a largo plazo. Este comportamiento es consistente tanto en primavera como en verano, aunque es más pronunciado en verano, lo que refuerza la necesidad de mejorar los modelos para escenarios de pronóstico a mayor anticipación.

El análisis de los gráficos muestra que los intervalos de confianza (IC) del 90% y 50% son sorprendentemente similares tanto en la calibración como en la validación cruzada. Esto indica que los modelos GAMLSS están generalizando bien, manteniendo niveles de incertidumbre comparables cuando se aplican a datos no vistos previamente. La razón de esta similitud se debe a que las características de los predictores seleccionados se mantienen consistentes, lo que permite que el modelo mantenga un rendimiento estable y una incertidumbre similar en ambas etapas.

Los picos observados en los gráficos, especialmente en los intervalos de confianza del 90%, reflejan períodos de alta variabilidad en los caudales, lo que aumenta la incertidumbre en las predicciones del modelo. Esta variabilidad puede estar asociada a fluctuaciones interanuales en las condiciones hidrológicas, haciendo que el modelo tienda a reflejar intervalos de confianza más amplios en respuesta a estas variaciones, lo que sugiere una menor precisión en la predicción de ciertos períodos con alta variabilidad.

En la comparación entre los modelos GAMLSS y MRLM, se observa que los coeficientes de eficiencia (CE) para los modelos GAMLSS son negativos cuando se usa el MRLM como referencia. Esto se debe a que el CE evalúa principalmente el error medio, sin considerar la representación de la incertidumbre.

Por lo tanto, aunque los modelos GAMLSS muestran una menor eficiencia según el CE, esto no significa que sean inferiores en todos los aspectos. De hecho, el CRPSS positivo en varias configuraciones indica que los modelos GAMLSS tienen una mejor capacidad para capturar la variabilidad e incertidumbre de los caudales que el MRLM, lo que es crucial para la gestión de recursos hídricos en contextos de alta variabilidad climática. Además, permite una toma de decisiones más informada, especialmente cuando se enfrentan escenarios de riesgo climático.

A pesar de que los modelos GAMLSS han demostrado ser superiores en algunos aspectos comparados con el MRLM, no son necesariamente los mejores modelos para la predicción de caudales en términos absolutos. La existencia de grandes incertidumbres en ciertos escenarios y la sobreconfianza que esto puede generar al no capturar correctamente todos los factores influyentes sugiere que se necesita seguir buscando nuevos modelos que ofrezcan mayor precisión y menor incertidumbre. Los modelos GAMLSS, no logran predecir con la exactitud y consistencia deseada, y en algunos casos, pueden inducir a errores significativos en la toma de decisiones si no se consideran adecuadamente sus limitaciones.

En conclusión, el trabajo futuro debe centrarse en desarrollar y probar modelos alternativos que puedan mejorar tanto la precisión como la representación de la incertidumbre en la predicción de caudales, particularmente para desfases temporales más largos y para condiciones de alta variabilidad climática.

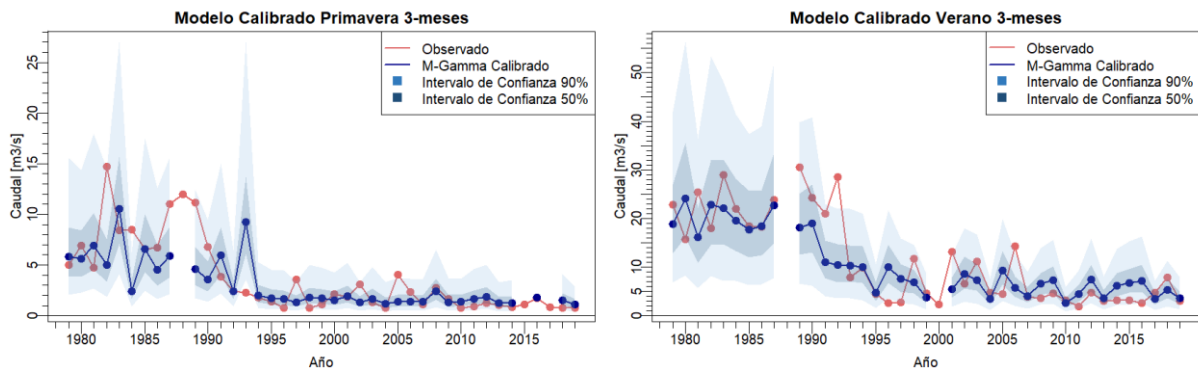
9 Referencias

- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723. <https://doi.org/10.1109/TAC.1974.1100705>
- Clark, N. J., & Wells, K. (2023). Dynamic generalised additive models (<scp>DGAMs</scp>) for forecasting discrete ecological time series. *Methods in Ecology and Evolution*, 14(3), 771–784. <https://doi.org/10.1111/2041-210X.13974>
- Debele, S. E., Bogdanowicz, E., & Strupczewski, W. G. (2017). Around and about an application of the GAMLSS package to non-stationary flood frequency analysis. *Acta Geophysica*, 65(4), 885–892. <https://doi.org/10.1007/s11600-017-0072-3>
- Escobar, F., Casassa, G., & Pozo, V. (1995). *Variaciones de un glaciar de montaña en los andes de Chile central en las últimas dos décadas*. 683–695.
- Farías-Barahona, D., Ayala, Á., Bravo, C., Vivero, S., Seehaus, T., Vijay, S., Schaefer, M., Buglio, F., Casassa, G., & Braun, M. (2020). 60 Years of Glacier Elevation and Mass Changes in the Maipo River Basin, Central Andes of Chile. *Remote Sensing*, 12(10), 1658. <https://doi.org/10.3390/rs12101658>
- Feinstein, J., Ploussard, Q., Veselka, T., & Yan, E. (2023). Using Data-Driven Prediction of Downstream 1D River Flow to Overcome the Challenges of Hydrologic River Modeling. *Water*, 15(21), 3843. <https://doi.org/10.3390/w15213843>
- Flamenco, E. A. (2010). *PRONÓSTICO ESTACIONAL DE CAUDALES EN LA CUENCA DEL RÍO JACHAL*.
- García, R. (2008). *Gestión de cuencas y Cambio Climático: El caso del Maipo*.
- Garreaud, R. D., Boisier, J. P., Rondanelli, R., Montecinos, A., Sepúlveda, H. H., & Veloso-Aguila, D. (2020). The Central Chile Mega Drought (2010–2018): A climate dynamics perspective. *International Journal of Climatology*, 40(1), 421–439. <https://doi.org/10.1002/joc.6219>
- Gironas, J. (2013). *Exploring possible connections between hydrological extreme events and climate change in central south Chile*. <https://www.cigiden.cl/exploring-possible-connections-between-hydrological-extreme-events-and-climate-change-in-central-south-chile/>
- Gupta, H. V., Sorooshian, S., & Yapo, P. O. (1999). Status of Automatic Calibration for Hydrologic Models: Comparison with Multilevel Expert Calibration. *Journal of Hydrologic Engineering*, 4(2), 135–143. [https://doi.org/10.1061/\(ASCE\)1084-0699\(1999\)4:2\(135\)](https://doi.org/10.1061/(ASCE)1084-0699(1999)4:2(135))
- Hersbach, H. (2000). Decomposition of the Continuous Ranked Probability Score for Ensemble Prediction Systems. *Weather and Forecasting*, 15(5), 559–570. [https://doi.org/10.1175/1520-0434\(2000\)015<0559:DOTCRP>2.0.CO;2](https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2)
- Hollander, M., A. Wolfe, D., & Chicken, E. (2015). *Nonparametric Statistical Methods*. Wiley. <https://doi.org/10.1002/9781119196037>
- Instituto Geográfico de Aragón. (2017). *Indicador continental. Precipitación - Versión Beta*.
- Mardones, P. (2019). *Impactos del cambio climático en la altura de la isoterma 0°C sobre Chile Central* [Universidad de Chile]. <https://repositorio.uchile.cl/handle/2250/174529>
- Maturana, J., Bello, M., & Manley, M. (2004). Antecedentes históricos y descripción del fenómeno El Niño, Oscilación del Sur. *El Niño-La Niña 1997-2000. Sus Efectos En Chile.*, 13–27.
- Ministerio del Medio Ambiente. (2012). *Estrategia Nacional de Recursos Hídricos*.
- Ministerio del Medio Ambiente. (2023). *¿Qué es el cambio climático?* Ministerio Del Medio Ambiente. <https://cambioclimatico.mma.gob.cl/que-es-el-cambio-climatico/>

- Montalva, J. (2022). *Estudio de Teleconexión entre Caudal Estacional y Variables Climatológicas e Hidrometeorológicas en la cuenca alta del Río Maipo para su aplicación en Modelo de Pronóstico Estacional*.
- Moore, D. S., McCabe, G. P., & Craig, B. A. (2012). *Introduction to the Practice of Statistics* (7th ed.). W.H. Freeman and Company.
- Ossandón, Á., J. S., N., Mendoza, P. A., Rajagopalan, B., & Mishra, V. (2022). A Bayesian Hierarchical Framework for Postprocessing Daily Streamflow Simulations across a River Network. *Journal of Hydrometeorology*, 23(6), 947–963. <https://doi.org/10.1175/JHM-D-21-0167.1>
- Poveda, G., Mesa, O., Carvajal, L. F., Hoyos, C. D., Mejía, J. F., Cuartas, A., & Pulgarín, A. (2002). *PREDICCIÓN DE CAUDALES MEDIOS MENSUALES EN RÍOS COLOMBIANOS USANDO MÉTODOS NO LINEALES*. <http://www.cdc.noaa.gov/~kew/MEI>.
- Rojas, O., Mardones, M., Arumí, J. L., & Aguayo, M. (2014). Una revisión de inundaciones fluviales en Chile, período 1574-2012: causas, recurrencia y efectos geográficos. *Revista de Geografía Norte Grande*, 57, 177–192. <https://doi.org/10.4067/S0718-34022014000100012>
- Stasinopoulos, D. M., & Rigby, R. A. (2007). *Journal of Statistical Software Generalized Additive Models for Location Scale and Shape (GAMLSS) in R*. <http://www.jstatsoft.org/>
- Stasinopoulos, M. D., Rigby, R. A., Heller, G. Z., Voudouris, V., & Bastiani, F. (2017). *Flexible Regression and Smoothing: Using GAMLSS in R*.
- Thompson, D. W. J., & Wallace, J. M. (2000). Annular Modes in the Extratropical Circulation. Part I: Month-to-Month Variability*. *Journal of Climate*, 13(5), 1000–1016. [https://doi.org/10.1175/1520-0442\(2000\)013<1000:AMITEC>2.0.CO;2](https://doi.org/10.1175/1520-0442(2000)013<1000:AMITEC>2.0.CO;2)
- Trenberth, K. E., & Fasullo, J. T. (2013). *Regional climate trends and projections*. In *Climate Change: The Physical Science Basis*.
- Valdivielso, A. (2023). *¿Qué es un caudal?* [https://www.iagua.es/respuestas/que-es-caudal#:~:text=El%20termino%20caudal%20significa%3A%20volumen,volumen\)%20y%20t%20\(tiempo\)](https://www.iagua.es/respuestas/que-es-caudal#:~:text=El%20termino%20caudal%20significa%3A%20volumen,volumen)%20y%20t%20(tiempo).).
- Valenzuela, R., & Garreaud, R. (2019). *Extreme Daily Rainfall in Central-Southern Chile and Its Relationship with Low-Level Horizontal Water Vapor Fluxes*.

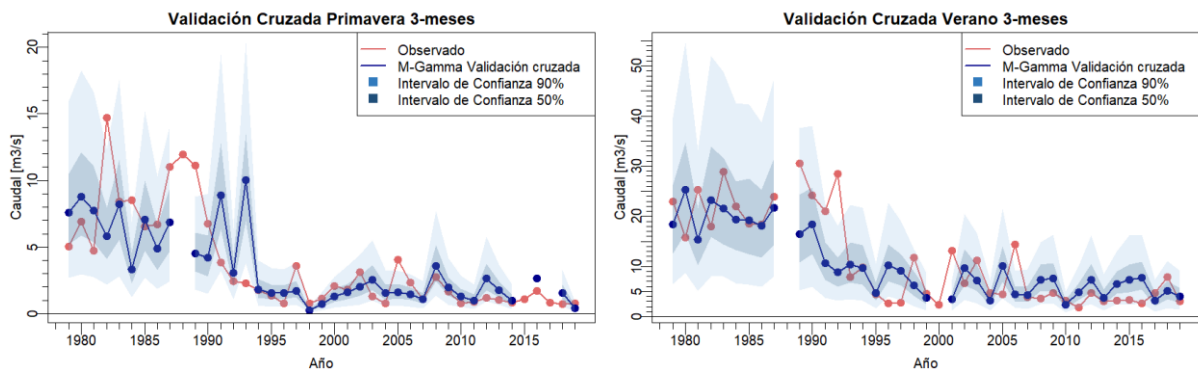
Anexos

A. Anexo 1



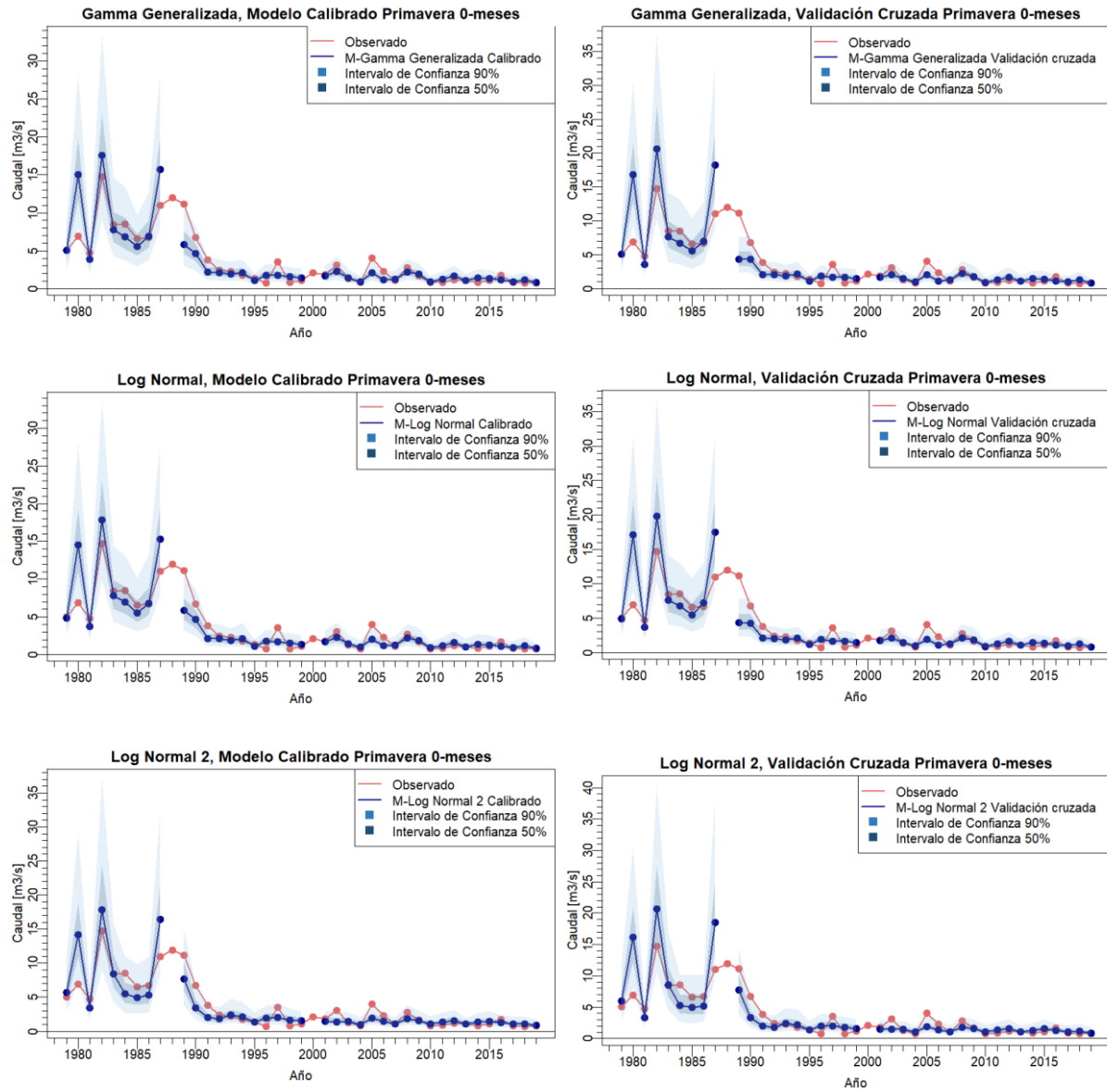
Anexo 1. Serie temporal del caudal modelado y observado para primavera y verano con un desfase de 3 meses. El modelo GAMLSS con la mejor distribución Gamma presenta un parámetro de localización (μ) no estacionario y un parámetro de escala (σ) estacionario.

B. Anexo 2



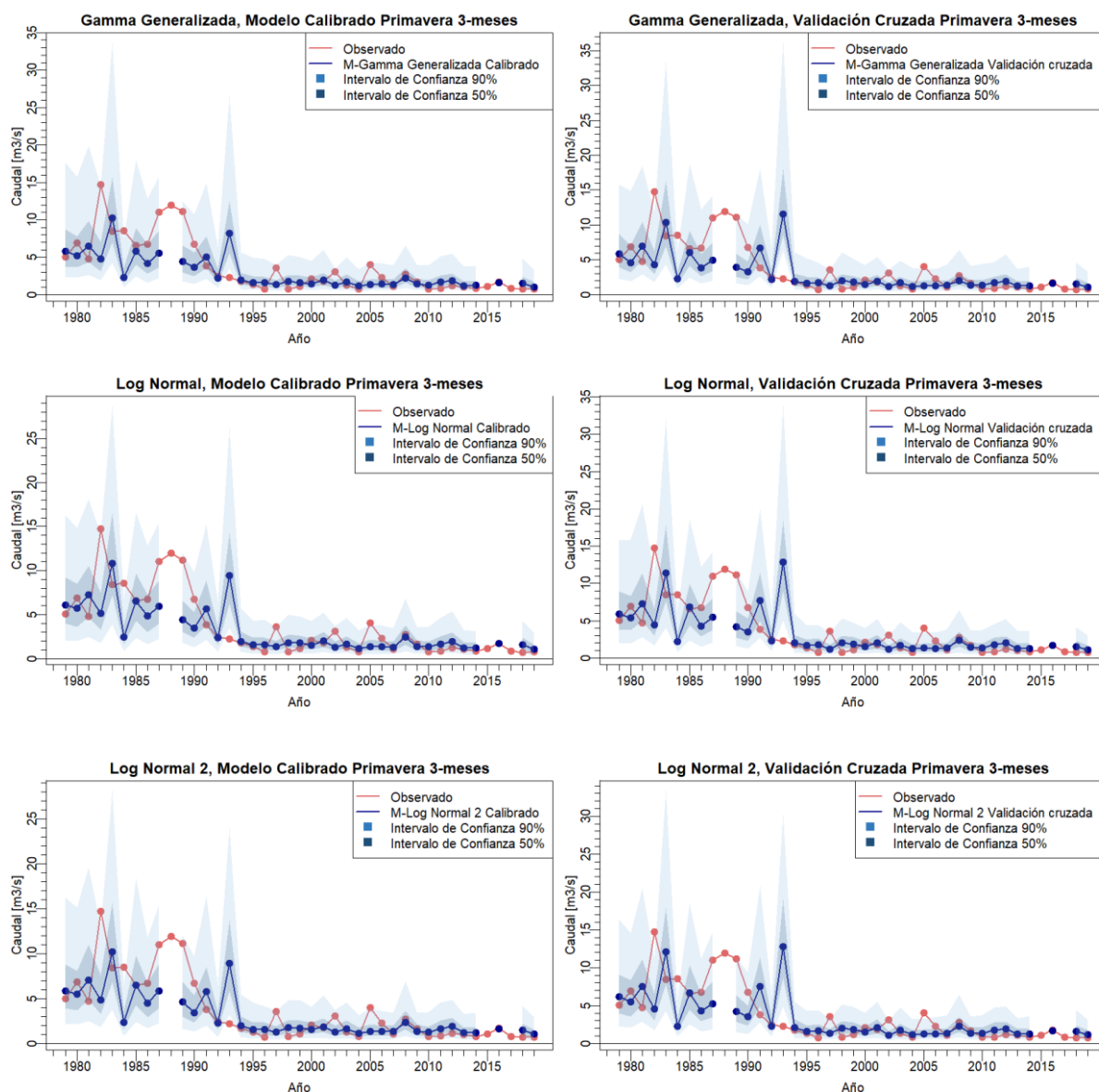
Anexo 2. Serie temporal del caudal modelado y observado para primavera y verano con un desfase de 0 meses en validación cruzada. El modelo GAMLSS utiliza la mejor distribución Gamma, con un parámetro de localización (μ) no estacionario y un parámetro de escala (σ) estacionario.

C. Anexo 3



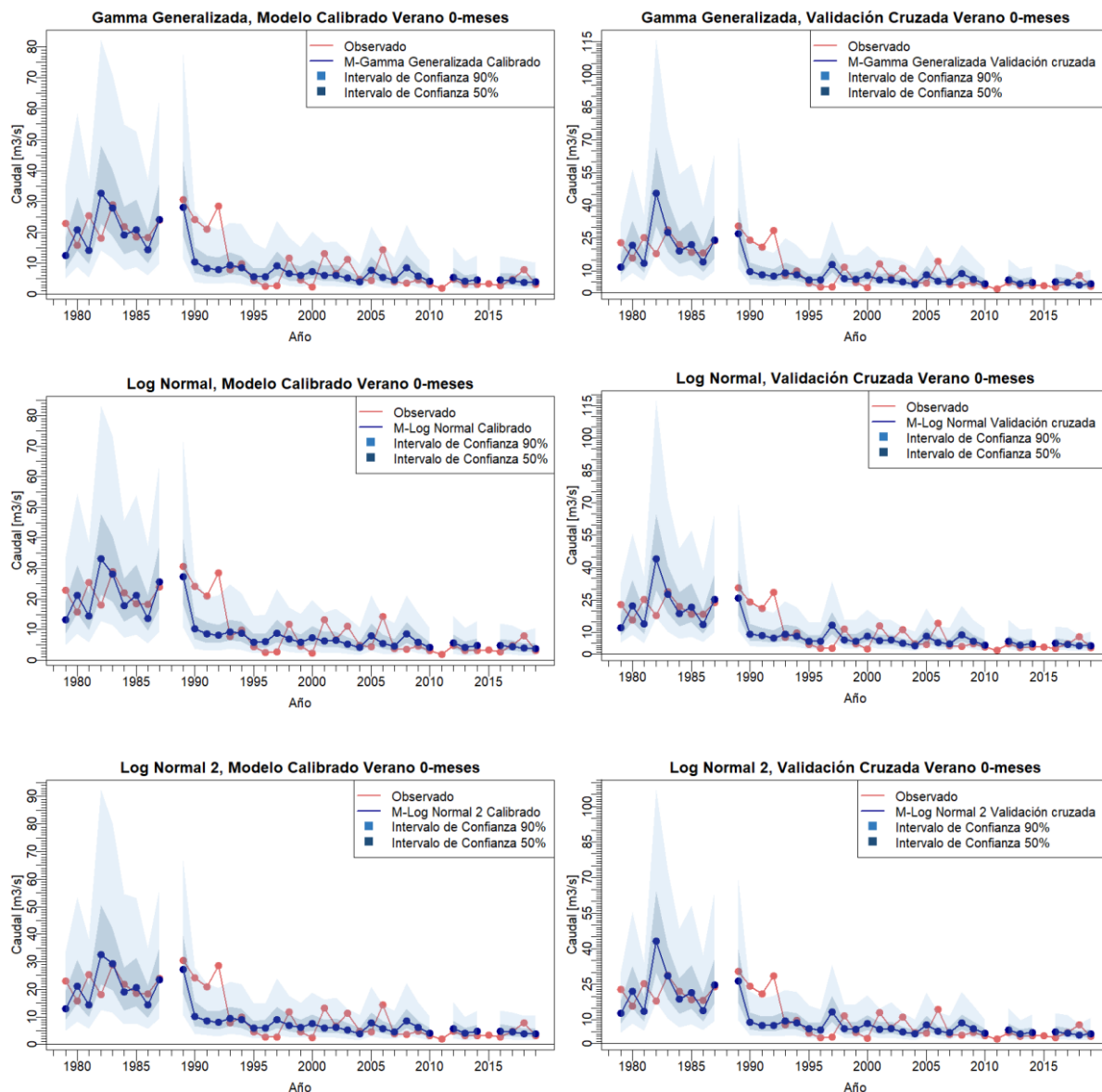
Anexo 2. Series temporales de caudales modelados, validados por validación cruzada y observados para primavera con un desfase de 0 meses. Los modelos GAMLSS presentados son para las distribuciones: Gamma Generalizada, Log Normal y Log Normal 2, con un parámetro de localización (μ) no estacionario y un parámetro de escala (σ) estacionario.

D. Anexo 4



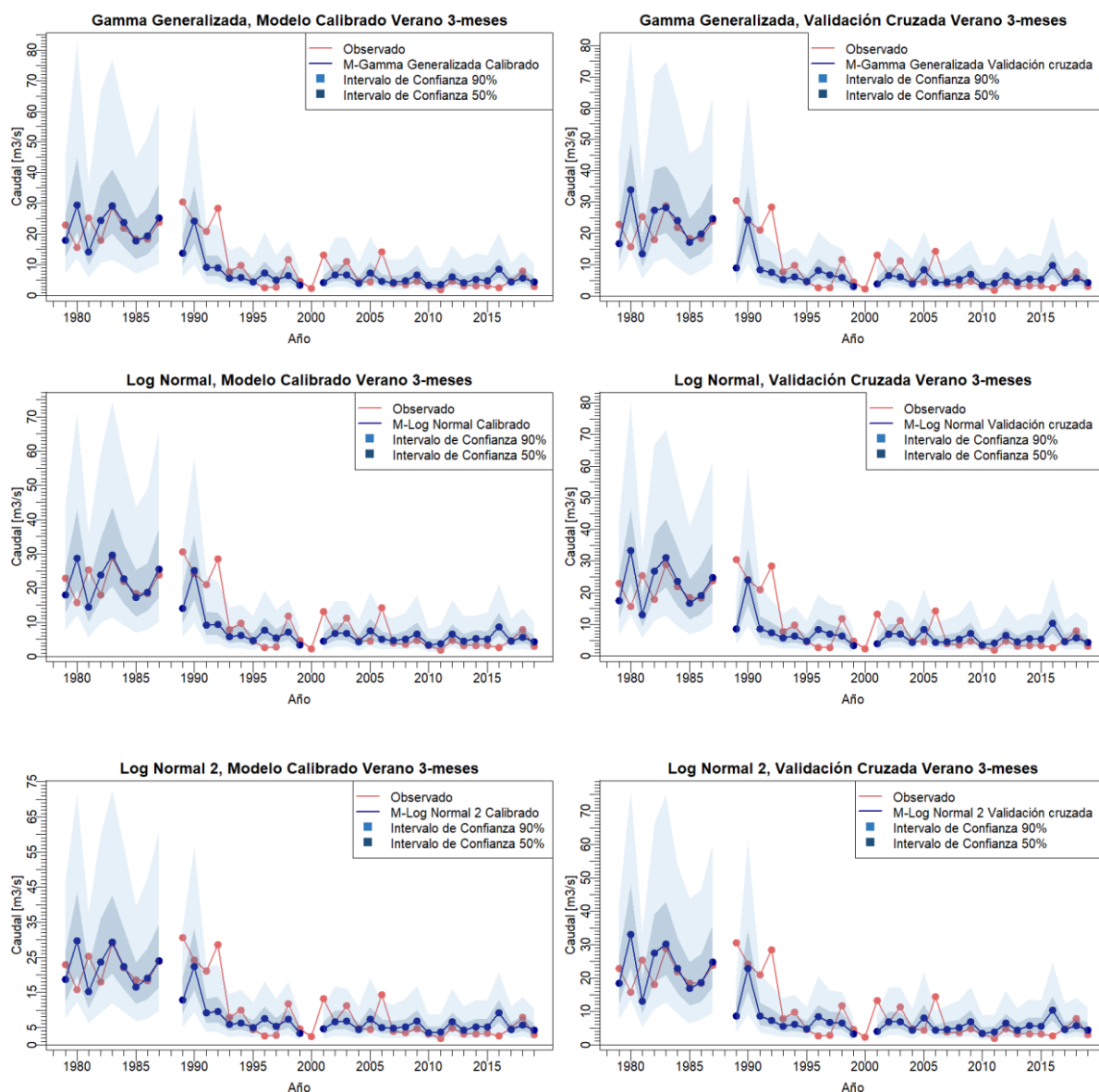
Anexo 3. Series temporales de caudales modelados, validados por validación cruzada y observados para primavera con un desfase de 3 meses. Los modelos GAMLSS presentados son para las distribuciones: Gamma Generalizada, Log Normal y Log Normal 2, con un parámetro de localización (μ) no estacionario y un parámetro de escala (σ) estacionario.

E. Anexo 5



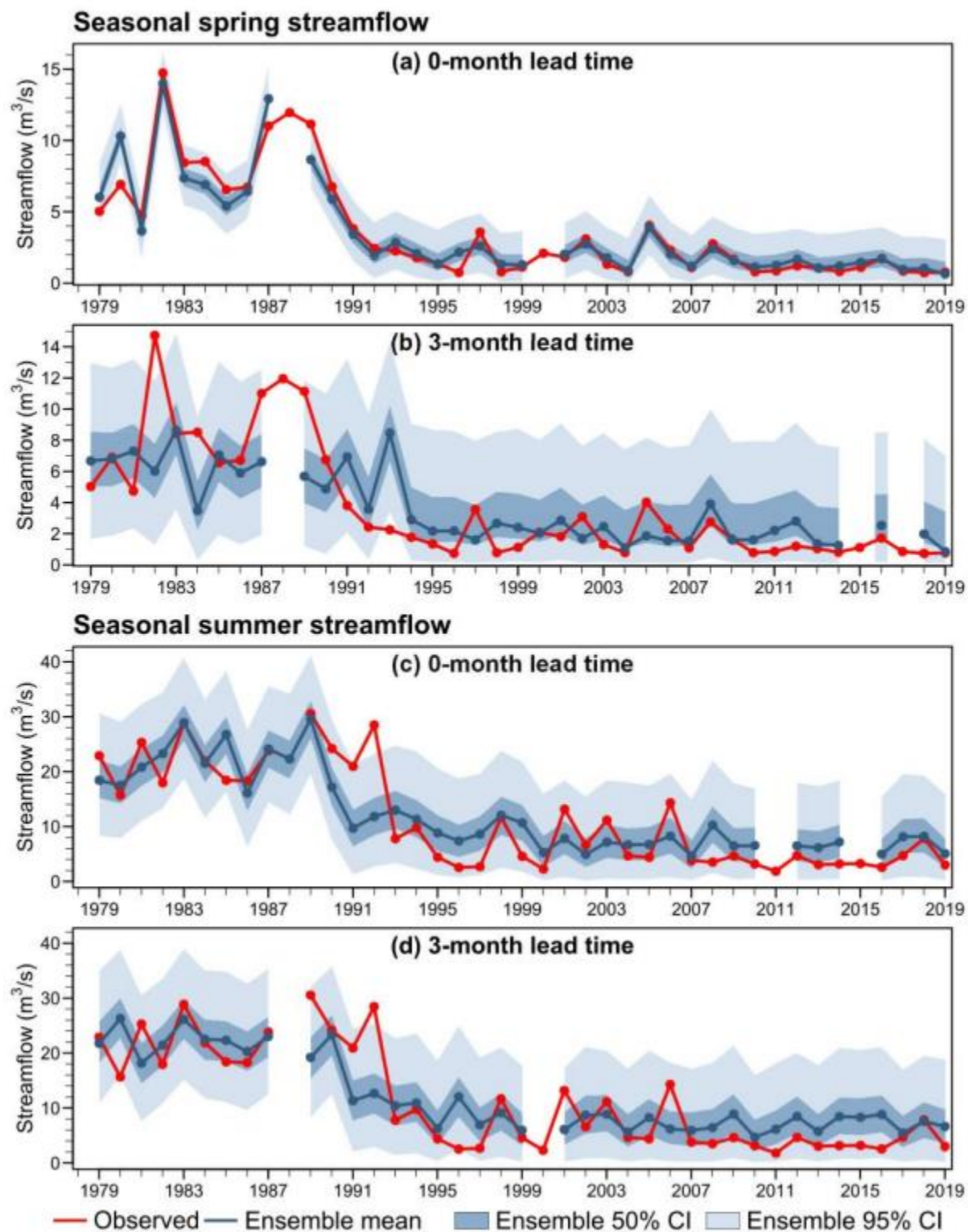
Anexo 4. Series temporales de caudales modelados, validados por validación cruzada y observados para verano con un desfase de 0 meses. Los modelos GAMLSS presentados son para las distribuciones: Gamma Generalizada, Log Normal y Log Normal 2, con un parámetro de localización (μ) no estacionario y un parámetro de escala (σ) estacionario.

F. Anexo 6



Anexo 5. Series temporales de caudales modelados, validados por validación cruzada y observados para verano con un desfase de 3 meses. Los modelos GAMLSS presentados son para las distribuciones: Gamma Generalizada, Log Normal y Log Normal 2, con un parámetro de localización (μ) no estacionario y un parámetro de escala (σ) estacionario.

G. Anexo 7



Anexo 6. Series temporales de caudales modelados y observados para primavera y verano con los desfases 0-meses y 3-meses para el MRLM.