

**Predicción de caudales a escala semanal
en el río Madre Vieja (Guatemala)
mediante modelos estadísticos basados
en series de tiempo**

Elaborado por:

César Espinosa (ICC).

Supervisado por:

Ricardo Marroquín (ICC).

Cita bibliográfica

Instituto Privado de Investigación sobre Cambio Climático ICC (2026). *Predicción de caudales a escala semanal en el río Madre Vieja (Guatemala) mediante modelos estadísticos basados en series de tiempo*. Pág. 50.

Índice

1. Introducción	5
2. Planteamiento del problema	7
3. Justificación	8
4. Estado del arte	9
5. Objetivos	11
5.1. Objetivo general	11
5.2. Objetivos específicos	11
6. Área de estudio	12
7. Métodos y materiales	15
7.1. Datos Recopilados	16
7.1.1. <i>Datos hidrométricos:</i>	16
7.1.2. <i>Datos meteorológicos y de moduladores climáticos</i>	16
7.2. Procesamiento de datos y ejecución de los modelos	17
7.2.1. <i>Adecuación de las series de tiempo</i>	17
7.2.1.1. <i>Caudal</i>	18
7.2.1.2. <i>Datos meteorológicos</i>	18
7.3. Fundamentos estadísticos y formulación de los modelos	18
7.3.1. <i>Modelo autorregresivo integrado de medias móviles (ARIMA)</i>	19
7.3.2. <i>Modelo estacional autorregresivo integrado de medias móviles (SARIMA)</i>	19
7.3.3. <i>Modelo SARIMA con variables exógenas (SARIMAX)</i>	21
7.3.4. <i>Formulación preliminar de hiperparámetros de los modelos</i>	21
7.4. Calibración de los modelos	22
7.5. Validación de los modelos	23
7.6. Estimación de la incertidumbre en las predicciones	24

8. Resultados	26
8.1. Caracterización y resumen de las series de datos	26
8.1.1. <i>Caudal - variable endógena</i>	26
8.1.2. <i>Variables meteorológicas – variables exógenas</i>	27
8.2. Análisis de la variable objetivo (caudal) y su adecuación para los modelos	29
8.3. Resultados del modelo ARIMA.....	32
8.4. Resultados del modelo SARIMA	34
8.5. Resultados del modelo SARIMAX.....	37
9. Discusión	40
10. Conclusiones y recomendaciones	44
11. Bibliografía	46

1. Introducción

Centroamérica se caracteriza por una elevada variabilidad climática, estrechamente vinculada a su posicionamiento geográfico (Maldonado et al., 2018). Debido a su ubicación entre el mar Caribe y el océano Pacífico, la precipitación constituye el principal factor de variabilidad climática en la región, mientras que los cambios de temperatura tienden a ser relativamente más moderados (Taylor & Alfaro, 2005).

En Guatemala, particularmente en la vertiente del Pacífico, se han documentado cambios en los patrones de precipitación entre 1971 y 2014 (INSIVUMEH, 2018). En esta región se han reportado incrementos de lluvia de hasta $30 \text{ mm}\cdot\text{año}^{-1}$, acompañados de un aumento en la intensidad de los eventos y una reducción en el número de días lluviosos. En contraste, los cambios en temperatura han sido más moderados, con incrementos inferiores a los $0.05 \text{ }^\circ\text{C}\cdot\text{año}^{-1}$ (INSIVUMEH, 2018). Independientemente de su magnitud, los incrementos en precipitación y temperatura dentro de un mismo período pueden favorecer una mayor frecuencia e intensidad de fenómenos hidrometeorológicos, como inundaciones y sequías (IPCC, 2021).

En este contexto, comprender el comportamiento del caudal de los ríos (definido como el volumen de agua que fluye por unidad de tiempo) adquiere especial relevancia para la gestión del riesgo de desastres y del recurso hídrico. El caudal constituye un indicador directo tanto del riesgo de inundaciones como de la disponibilidad de agua durante períodos de sequía. En consecuencia, su análisis puede contribuir a fortalecer la capacidad de respuesta de poblaciones, sectores productivos y usuarios ante escenarios de escasez hídrica y eventos hidrológicos extremos.

El modelado y la predicción hidrológica constituyen un enfoque cuantitativo para identificar patrones, analizar y estimar el comportamiento futuro de variables hidrológicas en un área específica, como el caudal de un río (Zhao, 2024). Para este fin, existen distintos enfoques que pueden emplear información hidrometeorológica a escala de cuenca o a partir de monitoreo local, incluyendo modelos basados en procesos físicos y modelos basados en datos. Entre estos últimos, los modelos estadísticos describen la dinámica del sistema a partir de la información contenida en los datos observados. Dichos enfoques permiten representar el comportamiento hidrológico sin requerir una descripción explícita de los procesos físicos subyacentes, basándose en la identificación de patrones y relaciones estadísticas en las series de datos.

Según Hyndman & Athanasopoulos (2021), la modelación y predicción basadas en datos pueden abordarse mediante tres estrategias principales: (i) modelos de series de tiempo, que utilizan valores pasados de la variable de interés; (ii) modelos explicativos, que incorporan variables predictoras correlacionadas; y (iii) modelos mixtos, que combinan ambas aproximaciones. A estos enfoques se suman los métodos modernos de aprendizaje automático (machine learning), reconocidos por su capacidad para capturar relaciones no lineales y patrones complejos en los datos (Hastie et al., 2009; James et al., 2021). Entre ellos destacan las redes neuronales, ampliamente utilizadas en tareas de predicción hidrológica, especialmente en contextos donde los modelos tradicionales presentan limitaciones para representar dinámicas altamente no lineales (Goodfellow et al., 2016).

La presente investigación implementa y compara el desempeño de tres modelos estadísticos de series de tiempo, de naturaleza paramétrica y estocástica, para la predicción de caudales en un río de la vertiente del Pacífico de Guatemala. Estos modelos describen la evolución temporal del proceso mediante un conjunto finito de parámetros y representan el comportamiento del sistema como una realización de un proceso probabilístico sujeto a incertidumbre. Los tres modelos considerados pertenecen a la familia ARIMA, en sus variantes ARIMA, SARIMA y SARIMAX. Estos modelos se seleccionaron por su relativa simplicidad de implementación en entornos de programación y por compartir una estructura metodológica similar. Sus principales diferencias radican en la inclusión o exclusión de componentes estacionales y variables exógenas, lo que permite realizar una evaluación comparativa de su desempeño en la predicción del caudal.

El sitio de estudio corresponde al río Madre Vieja, donde se modeló el caudal promedio semanal en la sección media del cauce utilizando registros de la estación hidrométrica Acarigua, administrada por el Instituto Privado de Investigación sobre Cambio Climático (ICC). Los resultados obtenidos permiten evaluar el potencial de este tipo de modelos para aplicaciones de predicción hidrológica y su contribución al fortalecimiento de la gestión del recurso hídrico y la anticipación de eventos hidrológicos extremos en la región.

2. Planteamiento del problema

La creciente variabilidad climática en la vertiente del Pacífico de Guatemala se manifiesta en el aumento de la intensidad de las precipitaciones y de la temperatura. Estos cambios han favorecido la ocurrencia de eventos extremos, como crecidas repentinas e inundaciones, así como períodos de sequía que afectan la disponibilidad y calidad del agua, la producción agrícola y las actividades económicas (OCHA, 2018).

A pesar de la recurrencia de eventos hidrometeorológicos extremos en la región, son escasos los estudios orientados a la modelación de caudales. Existe un vacío en la evaluación sistemática de modelos que integren la información hidrometeorológica y climática disponible para describir la variabilidad temporal del caudal. Esta carencia genera incertidumbre en la gestión del recurso hídrico y en la planificación de medidas preventivas frente a escenarios de escasez o exceso de agua (WMO, 2021). En particular, existe una limitada evaluación de qué modelos resultan más adecuados para representar el comportamiento del caudal en la región, así como de su capacidad predictiva para anticipar escenarios de exceso o déficit hídrico.

3. Justificación

El desarrollo de herramientas de predicción de caudales es importante para fortalecer la gestión integrada del recurso hídrico y reducir la vulnerabilidad de las poblaciones ante eventos extremos (APFM, 2013). Anticipar las dinámicas de caudal de un río puede contribuir a una mejor planificación del uso del agua en los sectores agrícola, industrial y doméstico, además de facilitar la implementación de medidas de mitigación y respuesta temprana ante desastres como períodos de sequía e inundaciones (WMO, 2022).

Los modelos de la familia ARIMA, en sus variantes ARIMA, SARIMA y SARIMAX, constituyen un enfoque estadístico que ha sido utilizado para la estimación y predicción de caudales en hidrología (Montanari et al., 2013). Estos modelos tienen la capacidad de capturar la estructura temporal de las series, así como la influencia de variables externas o covariables, por ejemplo, de carácter meteorológico y climático. Esta capacidad los convierte en herramientas relativamente flexibles ampliamente aplicadas en estudios hidrológicos (Perazzolo et al., 2025; Al-Saati et al., 2021; Montanari et al., 1997). La implementación de esta familia de modelos en el río Madre Vieja permitirá evaluar su desempeño y potencial de predicción, así como generar información estratégica para la gestión hídrica regional.

4. Estado del arte

En el ámbito de la modelación y predicción hidrológica, particularmente en la estimación de caudales, autores como Montanari et al. (2013) resaltan la existencia de dos marcos metodológicos predominantes: (i) los modelos basados en procesos físicos, que describen de manera relativamente explícita la relación precipitación–escorrentía, y (ii) los modelos basados en datos, que emplean técnicas estadísticas y/o de aprendizaje automático para capturar patrones y dependencias temporales. Dentro de este último grupo se encuentra la familia de modelos estadísticos autorregresivos integrados de medias móviles (ARIMA) y sus extensiones estacionales o con covariables también conocidas como ‘variables exógenas’ (ARIMAX, SARIMA y SARIMAX).

Los modelos hidrológicos basados en procesos físicos suelen requerir mayores esfuerzos para alcanzar resultados aceptables, ya que dependen de numerosos parámetros (elevada parametrización), lo que puede dificultar su calibración y la adecuada representación espaciotemporal de estos en el área de estudio (Beven, 2012). En contraste, los modelos estadísticos y de aprendizaje automático han mostrado mayor flexibilidad operativa, al facilitar la integración de grandes volúmenes de datos y la automatización de los procesos de predicción (Perazzolo et al., 2025).

Dentro de los enfoques estadísticos, Adnan et al. (2017) evaluaron la capacidad de los modelos de medias móviles autorregresivas (ARMA) y de medias móviles autorregresivas integradas (ARIMA) para proyectar series mensuales de caudal en un río de la cuenca alta del Indo, en Pakistán. Los resultados mostraron que el modelo ARIMA presentó un mejor desempeño en la representación y proyección de la serie temporal.

En un estudio reciente, Perazzolo et al. (2025) compararon el desempeño de distintos enfoques para modelar una serie temporal horaria en una cuenca montana prístina del noreste de Italia, incluyendo un modelo estadístico con variables exógenas (ARIMAX), un modelo hidrológico basado en procesos (HEC-HMS) y un modelo de aprendizaje automático basado en redes neuronales de memoria a largo plazo (LSTM). Los resultados mostraron que ARIMAX tendió a subestimar los valores observados, mientras que HEC-HMS reprodujo adecuadamente los valores extremos. El modelo LSTM presentó el mejor desempeño general. Los autores señalaron que el ajuste de parámetros en HEC-HMS es fundamental para obtener simulaciones confiables.

En contexto nacional, Dobias & Salvinelli (2018) desarrollaron un modelo predictivo de caudales a partir de una serie temporal trimestral en la cuenca del río Chixoy, utilizando análisis de teleconexiones y regresión lineal múltiple. En dicho modelo, los índices oceánicos y atmosféricos asociados al fenómeno El Niño–Oscilación del Sur (ENSO) fueron utilizados como variables independientes.

Específicamente en la vertiente del Pacífico, el Instituto Privado de Investigación sobre Cambio Climático desarrolló un modelo hidrológico del río Madre Vieja mediante HEC-HMS (ICC, 2015). Sin embargo, debido a la ausencia de registros de caudal en ese momento, el modelo no pudo ser calibrado. Posteriormente, en el 2021 se realizaron estudios enfocados en la generación de mapas de riesgo por inundación y sequía para cuencas como la del río Coyolate (CONRED, 2021).

No obstante, estos estudios se orientan principalmente a la evaluación de escenarios hidrológicos extremos, más que a la predicción continua del caudal como serie temporal. Actualmente, la disponibilidad de registros continuos de caudal en el río Madre Vieja desde 2018 permite explorar la aplicación de enfoques basados en procesos físicos y estadísticos para la predicción hidrológica.

5. Objetivos

5.1. Objetivo general

Evaluar la capacidad predictiva de tres modelos estadísticos de series de tiempo de la familia ARIMA (ARIMA, SARIMA y SARIMAX) para estimar el caudal promedio semanal en el río Madre Vieja, ubicado en la vertiente del Pacífico de Guatemala.

5.2. Objetivos específicos

- Calibrar y validar los modelos ARIMA, SARIMA y SARIMAX para la parte media del cauce del río Madre Vieja.
- Evaluar el desempeño predictivo de los modelos mediante la comparación entre valores observados y estimados, mediante métricas como el error cuadrático medio (RMSE) y la eficiencia de Nash–Sutcliffe (NSE).
- Identificar cual es el modelo que mejor predice la dinámica de la variable de interés (caudal promedio semanal) en el río Madre Vieja.
- Discutir el potencial del mejor modelo para aplicaciones reales de pronóstico en el río Madre Vieja y su potencial para ser implementado en otros ríos de la Vertiente del Pacífico.

6. Área de estudio

El río Madre Vieja tiene un área de captación (cuenca) de un poco más de 880 km² y su cauce principal un largo aproximado de 116 km (Tabla 1). El río posee una dirección predominantemente sur (Figura 1), en línea con el gradiente topográfico de la vertiente del Pacífico de Guatemala. El río se origina en la Sierra Madre, en límites entre los departamentos de Sololá, Quiché y Chimaltenango. La elevación media de la cuenca es de 1110 msnm, mientras que su pendiente es bastante pronunciada (pendiente promedio de 33%). El río desemboca en el océano Pacífico en la zona limítrofe entre los municipios de Tiquisate y Nueva Concepción, en el departamento de Escuintla. El área de captación considerada corresponde a la delimitación oficial de la cuenca definida por el Ministerio de Agricultura, Ganadería y Alimentación (MAGA, 2004). Esta delimitación probablemente fue definida con fines de gestión territorial y no exclusivamente con base en criterios hidrológicos, ya que incluye sectores, particularmente en la parte baja de la cuenca, que no drenan hacia el cauce principal del río.

Tabla 1. Descriptores morfológicos básicos de la cuenca del río Madre Vieja.

Cuenca	A	P	L	I/Rf	D	E max	E min	E	Sg
	km ²	km	km		km/km ²	msnm	msnm	msnm	%
Madre Vieja	884	359	116	22	0.43	3070	0	1110	33

Los parámetros determinados son:

A: Área total de la cuenca (kilómetros cuadrados);

P: Perímetro de la cuenca (kilómetros);

L: Longitud del cauce principal (kilómetros);

I/Rf: Factor de forma (A/L^2);

D: Densidad de drenaje de la cuenca (kilómetros de cauce por kilómetros cuadrados de superficie);

E max: Elevación máxima de la cuenca (msnm);

E min: Elevación mínima de la cuenca (msnm);

E: Elevación promedio de la cuenca (msnm);

Sg: Pendiente media del terreno (porcentaje).

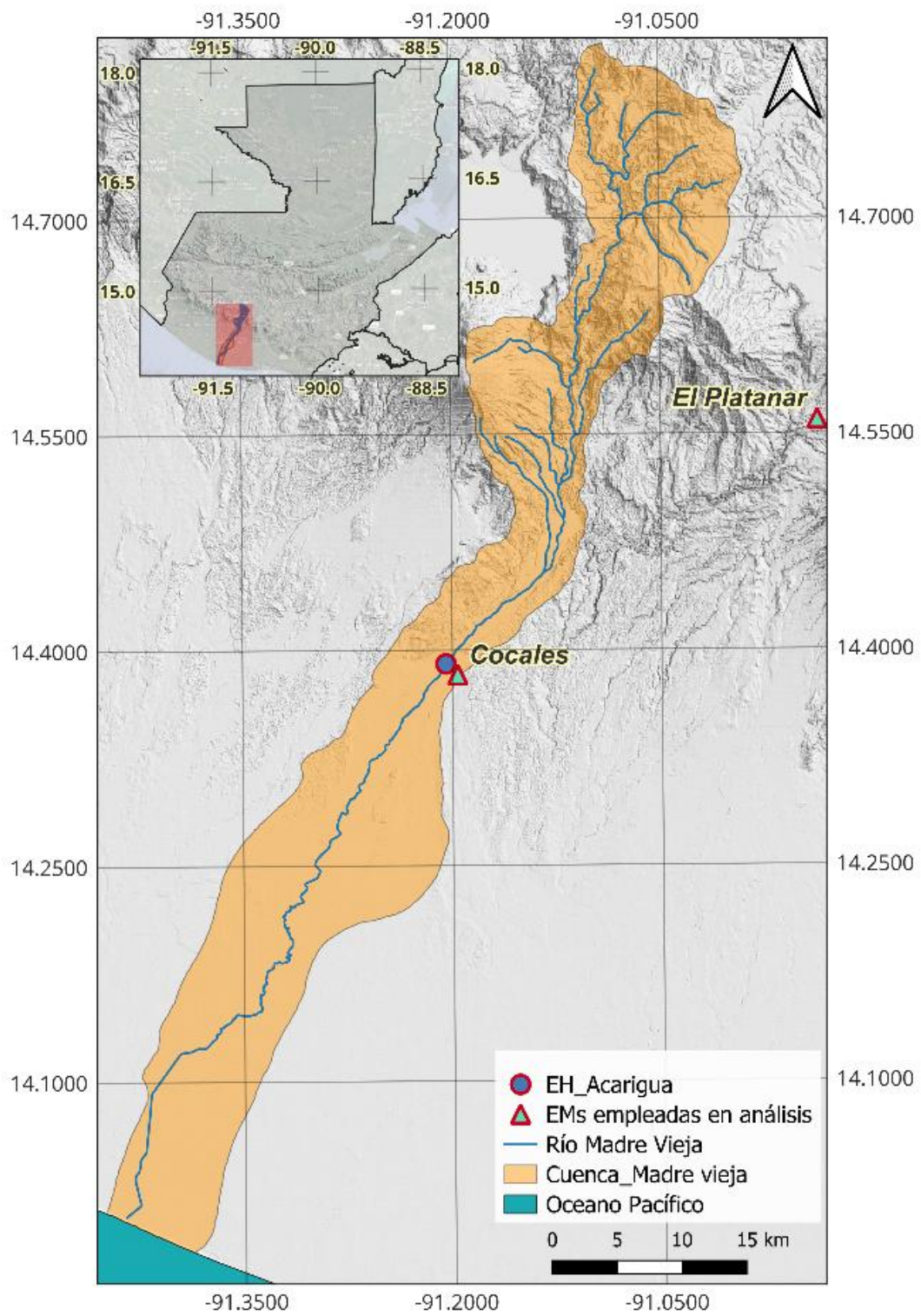


Figura 1. Ubicación de la cuenca del Río Madre Vieja, estaciones hidrométricas y estaciones meteorológicas automáticas empleadas en esta investigación.

La relación entre el área acumulada de la cuenca y su elevación relativa (curva hipsométrica) sugiere que la cuenca presenta un estado geomorfológico relativamente juvenil, con una alta proporción de superficie localizada en elevaciones mayores (Strahler, 1952) (Figura 2). Este patrón es característico de paisajes poco degradados, asociados a relieves abruptos, pendientes pronunciadas y procesos erosivos activos.

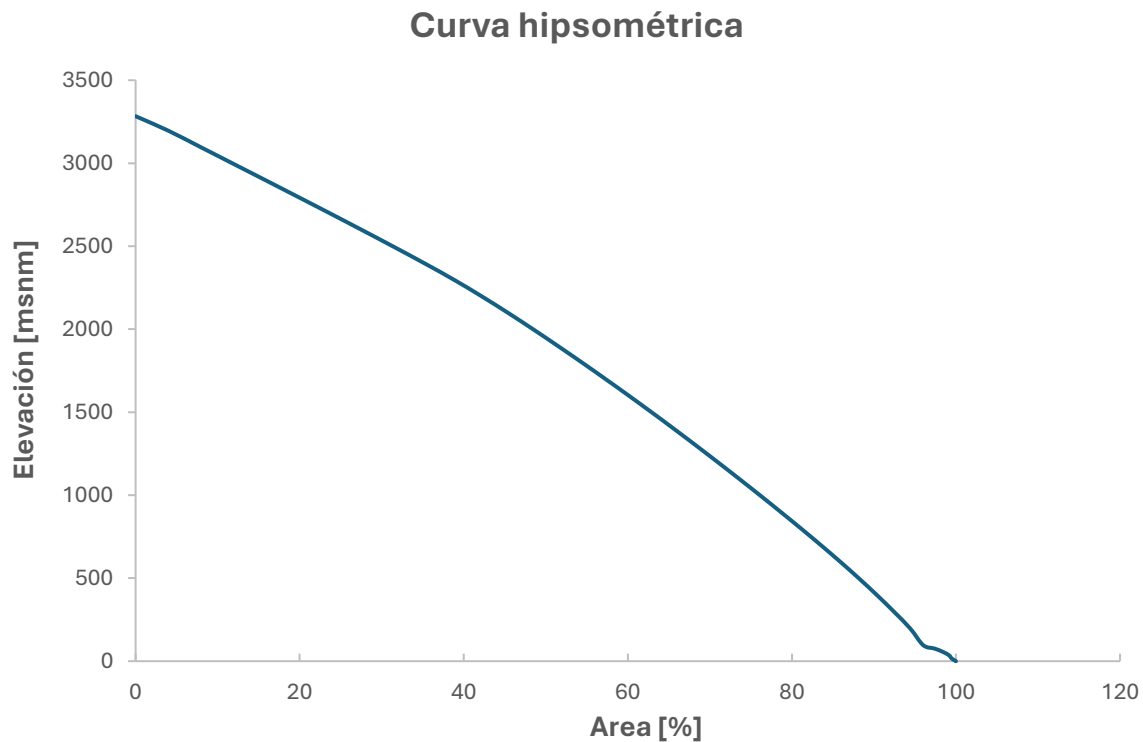


Figura 2. Curva hipsométrica para la cuenca del río Madre Viaje. El área se indica en porcentaje y la elevación en metros sobre el nivel del mar.

Según ICC (2015), los sistemas productivos agrícolas constituyen el uso de suelo predominante en la cuenca, ocupando aproximadamente el 60% de su superficie. En contraste, la cobertura forestal representa alrededor del 35% y se concentra principalmente en las zonas altas. Para un mayor detalle sobre la distribución espacial de los usos de suelo en el área de estudio, se recomienda consultar ICC (2015).

7. Métodos y materiales

Con el objetivo de modelar los caudales en la parte media del río Madre Vieja y avanzar hacia su predicción, se realizó un análisis estadístico de series de tiempo utilizando registros hidrométricos y meteorológicos disponibles. El flujo metodológico específico seguido en este trabajo (Figura 3) se basó en el marco general de predicción estadística propuesto por Hyndman & Athanasopoulos (2021).

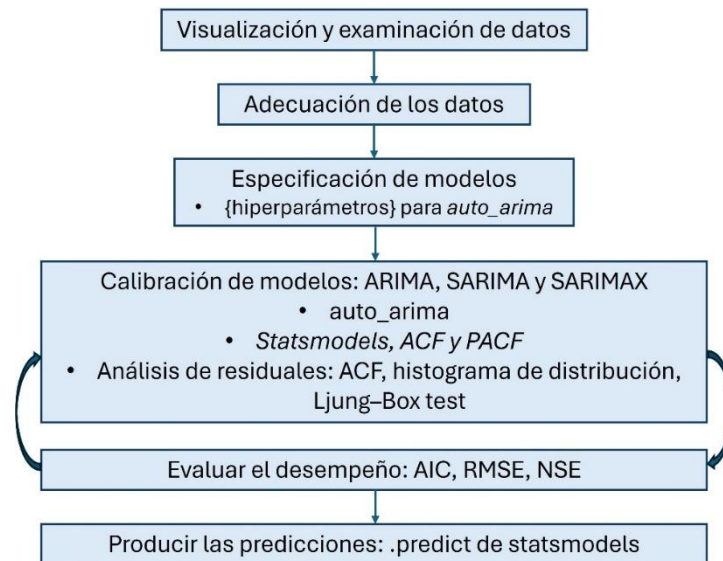


Figura 3. Flujo metodológico en este trabajo. Adaptado de Hyndman & Athanasopoulos (2021).

El enfoque metodológico adoptado enmarca la modelación en la sección media de la cuenca, donde las condiciones hidrológicas se conservan relativamente más prístinas en comparación con las zonas bajas de la misma, que pueden tener una mayor intervención de actividades humanas p.ej. actividades agroindustriales. Esta selección buscó que las relaciones entre caudal y variables meteorológicas reflejaran, en la medida de lo posible, la dinámica natural de la cuenca.

A partir de ello, se recopilaron los conjuntos de datos a ser utilizados en la modelación, los cuales fueron sometidos posteriormente a procesos de depuración, completado y agregación temporal. En etapas posteriores, se identificaron rangos de hiperparámetros y se evaluó el desempeño de diferentes combinaciones de hiperparámetros dentro de cada modelo mediante las métricas RMSE y NSE. Finalmente, se seleccionó el modelo con la relación más equilibrada entre el desempeño en calibración y validación.

7.1. Datos Recopilados

Los datos recopilados incluyen registros automáticos de estaciones hidrométricas y meteorológicas administradas por ICC (Tabla 2), así como información climática externa representada por el índice Niño-3.4 (N34), calculado por el NOAA Climate Prediction Center (CPC) y distribuido por el NOAA Physical Sciences Laboratory (PSL) (NOAA, 2025; Huang et al., 2017), basado en anomalías mensuales de la temperatura superficial del mar del conjunto ERSST v5. Estos datos constituyeron la base para la construcción de los modelos ARIMA, SARIMA y SARIMAX utilizados en este estudio.

7.1.1. Datos hidrométricos:

En la cuenca del río Madre Vieja operan dos estaciones hidrométricas automáticas (EH) administradas por el ICC, que registran el nivel del agua cada 15 minutos. Su funcionamiento se basa en la relación nivel-caudal en una sección específica del río, establecida mediante curvas de gasto (WMO, 2010b; Kennedy, 1990; Rantz et al., 1982). El nivel es medido mediante un sensor radar y transmitido a un servidor a través de una unidad de transmisión de datos remota. Posteriormente, los registros de nivel se convierten a caudal utilizando curvas de gasto, actualizadas periódicamente mediante aforos con molinete (ICC, 2017; WMO, 2010a). En este estudio se emplearon únicamente los datos de la EH Acarigua, ubicada en la parte media del cauce del río (véase la justificación en la Sección 7, Métodos y materiales).

7.1.2. Datos meteorológicos y de moduladores climáticos

Además de la serie objetivo (caudal), se incorporaron variables exógenas para la construcción del modelo SARIMAX (véase la estructura de cada modelo en la Sección 7.3). En este contexto, una variable exógena se entiende como una serie adicional, distinta de la variable de interés, que aporta información externa relevante y puede contribuir a mejorar la capacidad predictiva del modelo. En consecuencia, la serie hidrométrica objeto de modelación se denomina variable endógena.

Con base en la proximidad espacial y la correspondencia temporal con la estación hidrométrica Acarigua, se seleccionaron las estaciones meteorológicas (EM) Cocales y El Platanar, ambas administradas por el ICC (Figura 1). Como variables exógenas se consideraron la precipitación y la temperatura del aire a 2 m (mínima, media y máxima). Los registros de ambas estaciones fueron promediados, generando una serie combinada para cada variable.

Adicionalmente, se incorporó el índice Niño-3.4 (N34) como modulador climático externo, dada la reconocida influencia del ENSO sobre la variabilidad hidroclimática regional (Argueta Mayorga et al., 2025; Magaña et al., 1999). Este índice cuantifica las anomalías de la temperatura superficial del mar en la región 5°N–5°S y 170°–120°O del Pacífico ecuatorial, calculadas a partir del conjunto de datos ERSSTv5 (Huang et al., 2017), y es ampliamente utilizado para caracterizar la fase e intensidad del ENSO.

Tabla 2. Información general de las estaciones meteorológicas e hidrométricas utilizadas en este estudio.

Estación	Tipo	Longitud	Latitud	Altitud (msnm)	Frecuencia de registro
Acarigua	Hidrométrica	-91.204645	14.390637	215	15 minutos
Cocales	Meteorológica	-91.196198	14.382847	219	15 minutos
El Platanar	Meteorológica	-90.937931	14.559675	1578	15 minutos

7.2. Procesamiento de datos y ejecución de los modelos

Los análisis estadísticos se realizaron en el ambiente de Jupyter Notebook, un entorno interactivo cada vez más empleado por la comunidad científica (Kluyver et al., 2016). Este entorno permitió integrar en un mismo documento el código ejecutado, los resultados obtenidos y las visualizaciones generadas, contribuyendo a la reproducibilidad del proceso analítico.

Jupyter Notebook se combinó con librerías especializadas de Python: *pandas* para la manipulación de datos, *numpy* para cálculos numéricos, *statsmodels* para la implementación de modelos ARIMA/SARIMA/SARIMAX, *pmdarima* para la automatización de la selección de hiperparámetros, y *matplotlib* para la representación gráfica.

7.2.1. Adecuación de las series de tiempo

En una primera etapa, las series de datos crudos fueron depuradas mediante la identificación y eliminación de valores considerados inconsistentes con el comportamiento de la variable. Este proceso se basó en inspecciones estandarizadas y en el análisis exploratorio de las series, lo que permitió evaluar la coherencia de los registros en relación con su dinámica temporal.

Posteriormente, los datos faltantes en la serie fueron rellenados en su resolución original para, a partir de ella, proceder a la agregación semanal. La cantidad de datos faltantes fue menor al 7% del total de la serie en su resolución original. El proceso de completado de datos faltantes se realizó empleando promedios históricos calculados para la misma resolución temporal en

minutos, hora, día y mes considerando años anteriores (es decir, la climatología del sitio), garantizando así continuidad temporal en los registros del patrón hidrometeorológico.

7.2.1.1. Caudal

Se utilizó una serie de tiempo de caudal con resolución de 15 minutos correspondiente al período comprendido entre enero de 2018 y marzo de 2025 ($n = 251,509$). Dado el carácter exploratorio de esta aplicación, los análisis se desarrollaron a una escala temporal semanal ($n = 374$).

Para la agregación semanal, se consideraron tres enfoques, cada uno produciendo una serie semanal distinta. (i) Caudal promedio semanal: Se calcularon promedios diarios de caudal y luego se obtuvo el promedio semanal de dichos valores diarios (ii) Caudal mínimo semanal: Se extrajo el caudal mínimo diario y posteriormente se promedió semanalmente, generando una serie de promedios semanales de los mínimos diarios (iii) Caudal máximo semanal: se derivaron los caudales máximos diarios y se calculó el promedio semanal de dichos máximos diarios.

Cabe recalcar que el presente estudio se centra en la serie de caudal promedio semanal (serie i). Las otras dos series se utilizaron únicamente para la caracterización del régimen de descarga de la cuenca y no fueron consideradas en el análisis predictivo.

7.2.1.2. Datos meteorológicos

Utilizando la serie meteorológica combinada a escala diaria (ver Sección 7.1.2), se generaron agregaciones y promedios a escala semanal dependiendo la naturaleza de la variable.

- Precipitación: se calcularon acumulados semanales, obtenidos como la suma de los valores diarios registrados dentro de cada semana.
- Temperaturas: se derivaron promedios semanales de los valores diarios, generando así tres series semanales; temperatura mínima diaria semanal, temperatura promedio diaria semanal, temperatura máxima diaria semanal.

7.3. Fundamentos estadísticos y formulación de los modelos

Una vez establecidos los conjuntos de datos y definido el entorno de análisis, el siguiente paso consistió en la formulación e implementación de los tres modelos estadísticos seleccionados. En este estudio se emplearon los modelos estadísticos autorregresivos integrados de medias móviles (ARIMA), así como sus extensiones estacionales (SARIMA) y su variante con variables exógenas (SARIMAX).

A continuación, se describe brevemente la estructura que sustenta cada modelo, los componentes que los integran, los criterios utilizados para su identificación y parámetros de ajuste del modelo.

7.3.1. Modelo autorregresivo integrado de medias móviles (ARIMA)

ARIMA es el acrónimo para modelo autorregresivo integrado de medias móviles. Como su nombre lo indica, permite modelar procesos estacionarios y no estacionarios mediante diferenciación, componentes autorregresivos y de medias móviles (Box & Jenkins, 1976).

Uno de los principales supuestos del modelo es que la serie de tiempo bajo análisis es estacionaria, en el sentido de que la forma en que se dan los cambios en la serie temporal se mantienen constantes en el tiempo, si no, un proceso de diferenciación es requerido. El modelo puede ser escrito como la Ecuación (1).

$$y'_t = c + \sum_{i=1}^p \phi_i y'_{t-i} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \varepsilon_t \quad (1)$$

donde:

- y'_t denota la serie después de aplicar d diferenciaciones (orden d);
- ϕ_i son los coeficientes autorregresivos (AR);
- θ_j son los coeficientes de medias móviles (MA);
- c es una constante o término de deriva;
- ε_t representa el error aleatorio independiente e idénticamente distribuido.

En consecuencia, la estructura del modelo tiene los siguiente hiperparámetros ARIMA (p, d, q) donde:

- p = orden de la parte autorregresiva;
- d = orden de la diferenciación involucrada;
- q = orden de la parte de la media móvil.

7.3.2. Modelo estacional autorregresivo integrado de medias móviles (SARIMA)

El modelo SARIMA (Seasonal Autoregressive Integrated Moving Average) extiende la formulación clásica del modelo ARIMA incorporando explícitamente la estructura estacional presente en muchas series temporales (Box et al., 2016). Al igual que ARIMA permite modelar procesos estacionarios y no estacionarios bajo una diferenciación; además, incluye términos

adicionales que capturan la dependencia estacional con un periodo m , normalmente asociado a ciclos específicos (por ejemplo, ciclos anuales, mensuales o semanales).

La estructura del modelo se define mediante los órdenes (p, d, q) para la parte no estacional y $(P, D, Q)_m$ para la parte estacional, lo que permite capturar dinámicas complejas de corto y largo plazo dentro del mismo marco probabilístico. El modelo puede ser expresado mediante la Ecuación (2).

$$y'_t = c + \sum_{i=1}^p \phi_i y'_{t-i} + \sum_{k=1}^P \Phi_k y'_{t-km} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \sum_{l=1}^Q \Theta_l \varepsilon_{t-lm} + \varepsilon_t \quad (2)$$

donde:

- y'_t es la serie ya diferenciada (regular y estacional si aplica), por ejemplo:

$$y'_t = (1 - L)^d (1 - L^m)^D y_t$$

- ϕ_i : coeficientes AR no estacionales;
- Φ_k : coeficientes AR estacionales con rezagos múltiplos de m ;
- θ_j : coeficientes MA no estacionales;
- Θ_l : coeficientes MA estacionales;
- m : periodo estacional;
- ε_t : ruido blanco.

En consecuencia, la estructura del modelo tiene los siguientes hiperparámetros SARIMA $[(p, d, q)(P, D, Q)_m]$ donde:

- p, d, q = componentes AR, diferenciación y MA no estacionales;
- P = orden de la parte autorregresiva estacional (SAR), correspondiente a dependencias entre valores separados por múltiplos del periodo m ;
- D = orden de la diferenciación estacional, utilizada para eliminar tendencias o patrones repetitivos cada m observaciones;
- Q = orden de la parte de medias móviles estacional (SMA), que modela dependencias entre los errores separados también por múltiplos de m ;
- m = periodicidad estacional, es decir, el número de observaciones necesarias para completar un ciclo (por ejemplo, $m = 52$ para estacionalidad semanal en datos semanales con patrón anual).

7.3.3. Modelo SARIMA con variables exógenas (SARIMAX)

El modelo SARIMAX es una extensión del modelo SARIMA que incorpora el efecto de variables exógenas sobre la serie temporal de interés (Hyndman & Athanasopoulos, 2021). Su estructura combina componentes autorregresivas, de diferenciación simple y estacional, y de medias móviles, permitiendo representar tanto la dinámica interna de la serie como la influencia directa de factores externos. La serie transformada y'_t se obtiene al aplicar la diferenciación necesaria para lograr estacionariedad, mientras que el conjunto de variables exógenas X_t se integra a través de un término de regresión lineal.

Esta formulación pretende que el modelo represente adecuadamente procesos influenciados por condiciones externas conocidas, manteniendo la capacidad de describir patrones estacionales y de corto plazo propios de la serie. El modelo puede ser expresado mediante la Ecuación (3)

$$y'_t = c + \beta^T X_t + \sum_{i=1}^p \phi_i y'_{t-i} + \sum_{k=1}^P \Phi_k y'_{t-km} + \sum_{j=1}^q \theta_j \varepsilon_{t-j} + \sum_{l=1}^Q \Theta_l \varepsilon_{t-lm} + \varepsilon_t \quad (3)$$

donde:

- y'_t : serie ya diferenciada (regular y/o estacional);
- X_t : vector de variables exógenas en el tiempo t ;
- β : vector de coeficientes asociados a las variables exógenas;
- m : periodo estacional (p. ej., 52 para series semanales);
- $\phi_i, \Phi_k, \theta_j, \Theta_l$: coeficientes AR/AR_m y MA/MA_m;
- ε_t : error blanco.

Al igual que SARIMA, el modelo SARIMAX utiliza los mismos hiperparámetros $[(p, d, q)(P, D, Q)_m]$ que describen la estructura autorregresiva, de diferenciación y de medias móviles, tanto regular como estacional.

7.3.4. Formulación preliminar de hiperparámetros de los modelos

Para evaluar la idoneidad de las series temporales originales para su modelación, se aplicaron pruebas estadísticas y análisis gráficos orientados a caracterizar su estacionariedad (ausencia de cambios sistemáticos en media y varianza a lo largo del tiempo) y estacionalidad (patrones recurrentes que se repiten a intervalos regulares). Asimismo, se analizó la estructura de dependencia temporal mediante las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF), que permiten estimar, respectivamente, la correlación total entre una observación y sus rezagos, y la correlación directa eliminando la influencia de rezagos intermedios. Estos análisis

permitieron definir la pertinencia de las series y las transformaciones necesarias para cumplir los supuestos del modelado, tales como diferenciaciones temporales o ajustes de escala.

Para identificar y separar los componentes de tendencia y estacionalidad se empleó una descomposición STL. Adicionalmente, se inspeccionaron los correlogramas ACF y PACF tanto de la serie original como de las series diferenciadas (estacional y estacional combinada con diferenciación de primer orden). Con base en estos diagnósticos, y siguiendo los lineamientos de Hyndman & Athanasopoulos (2021) y Box et al. (2016), se preseleccionaron valores iniciales de los hiperparámetros utilizados posteriormente en las etapas de calibración y validación.

7.4. Calibración de los modelos

En primer lugar, se definió el conjunto de datos de calibración. Para ello, se utilizaron los primeros 304 de los 374 registros disponibles (aproximadamente el 80 % de los datos) de las series temporales, correspondientes al periodo comprendido entre enero de 2018 y octubre de 2023. Una transformación Box-Cox se estimó y aplicó exclusivamente sobre este subconjunto de calibración, obteniéndose un valor de $\lambda = 0.56$. Este parámetro fue posteriormente utilizado durante el proceso de modelado, con el fin de evitar el uso de información del conjunto de validación.

En una primera instancia de modelado, se efectuó una búsqueda automática de combinaciones de hiperparámetros mediante la función *auto_arima*. Esta herramienta explora distintas configuraciones posibles de órdenes autorregresivos (p), de diferenciación (d) y de medias móviles (q), así como sus equivalentes estacionales cuando corresponde (P , D , Q), seleccionando aquellas que minimizan el Criterio de Información de Akaike (AIC). El AIC permite comparar modelos en términos de parsimonia, evaluando el equilibrio entre calidad de ajuste y complejidad (Akaike, 1974), de modo que valores menores indican modelos relativamente más eficientes desde el punto de vista explicativo.

A partir de las cinco combinaciones con menor AIC y de una combinación adicional derivada del proceso de formulación preliminar descrito en la Sección 7.3.4, se generaron diez nuevas configuraciones de hiperparámetros dentro de rangos acotados. Estas variaciones permitieron ajustes de hasta ± 2 en los órdenes del modelo, sin admitir valores negativos ni configuraciones repetidas. Todas ellas fueron ajustadas utilizando la librería *statsmodels*, que proporciona un mayor control sobre la especificación estructural y el proceso de estimación.

El desempeño se evaluó mediante el error cuadrático medio (RMSE) y la eficiencia de Nash–Sutcliffe (NSE). Mientras el AIC evalúa parsimonia, estas métricas permiten cuantificar directamente la capacidad predictiva del modelo (Hyndman & Athanasopoulos, 2021; Moriasi et al., 2007).

La interpretación general de las métricas es la siguiente: el RMSE se expresa en las mismas unidades que la variable simulada y puede tomar valores entre 0 y ∞ , donde valores más bajos indican un mejor ajuste del modelo. Por su parte, el NSE es una métrica adimensional que puede variar entre $-\infty$ y 1, siendo los valores cercanos a 1 indicativos de un mejor desempeño predictivo (Krause et al., 2005).

Una clasificación cualitativa general de los valores de RMSE y NSE se presenta en la Tabla 3. Esta clasificación se basa en criterios comúnmente utilizados en estudios hidrológicos y refleja un criterio general, aunque su interpretación puede variar según el contexto de modelación. En este estudio, se utiliza como referencia para evaluar el grado de concordancia entre las predicciones de los modelos y las observaciones, sin embargo, no constituye un criterio absoluto.

Finalmente, el procedimiento se implementó de forma iterativa. A partir de la mejor combinación identificada según RMSE y NSE, se generaron nuevas variantes en su entorno paramétrico inmediato. Este proceso se repitió hasta completar la evaluación de 50 configuraciones adicionales por modelo. En conjunto con las seis configuraciones iniciales, se analizaron 56 combinaciones (por modelo), de las cuales se seleccionaron las 12 combinaciones con mejor desempeño para cada modelo. Estas 12 combinaciones para cada modelo fueron las consideradas para el proceso de validación y evaluación predictiva.

Tabla 3. Interpretación cualitativa de los rangos de valores de RMSE y NSE.

Métrica	Muy bueno	Bueno	Satisfactorio	Insatisfactorio	Referencia
NSE	(0.75,1.00]	(0.65,0.75]	(0.50,0.65]	($-\infty$,0.65]	Moriasi et al. 2007
RMSE	≤ 0.31 SD*	≤ 0.45 SD	≤ 0.83 SD	> 0.83 SD	Dazzi et al. 2021

Elaboración propia con base en Moriasi et al. (2007) y Dazzi et al. (2021), adaptado de Costa et al. (2023).

* SD = desviación estándar

7.5. Validación de los modelos

La validación se realizó mediante un esquema de validación temporal, utilizando un conjunto de datos observados independiente y cronológicamente posterior al período de calibración. Esta fase tuvo como objetivo verificar la capacidad predictiva de los modelos fuera de la muestra

utilizada en la calibración y evaluar la consistencia de su desempeño respecto a dicha fase, contribuyendo a la selección final del modelo más robusto.

El conjunto de validación consistió en los últimos 70 registros de las series temporales, correspondientes al período comprendido entre noviembre de 2023 y marzo de 2025. Dicho subconjunto representa aproximadamente el 20 % del total de datos disponibles a escala semanal (374 registros totales) para cada serie temporal.

Para cada modelo (ARIMA, SARIMA y SARIMAX), únicamente las 12 mejores combinaciones de hiperparámetros obtenidas en la fase de calibración fueron simuladas en la fase de validación. Su desempeño se evaluó utilizando las mismas métricas de desempeño utilizadas en la calibración, es decir, mediante RMSE y NSE.

Adicionalmente, se aplicaron pruebas diagnósticas sobre los residuos con el fin de seleccionar la mejor especificación (combinación de hiperparámetros) de cada modelo. En particular, se evaluó la ausencia de autocorrelación mediante la prueba de Ljung–Box, la normalidad de los residuos a través de análisis gráficos y la consistencia del desempeño entre las fases de calibración y validación. Con base en estos criterios, se seleccionó la especificación óptima de cada modelo entre las 12 combinaciones evaluadas en la fase de validación. Este enfoque permite que los modelos seleccionados cumplan con los supuestos teóricos del análisis de series de tiempo y representan de la mejor manera posible la dinámica hidrológica de la cuenca del río Madre Vieja, priorizando su capacidad predictiva fuera de la muestra, que es donde el modelo adquiere mayor valor práctico y potencial de aplicación en contextos reales (Shmueli, 2010).

7.6. Estimación de la incertidumbre en las predicciones

Con el fin de complementar la evaluación del desempeño de los modelos, se estimaron intervalos de confianza (IC) para las predicciones. Estos permiten evaluar la precisión y confiabilidad de las estimaciones al delimitar el rango esperado de las observaciones futuras para un nivel de confianza dado. Su uso resulta relevante en modelos de series de tiempo, donde la incertidumbre asociada a las predicciones tiende a incrementarse conforme aumenta el horizonte de pronóstico (Hyndman & Athanasopoulos, 2021).

Los intervalos de confianza se obtuvieron a partir de la función *get_forecast ()* de la librería *statsmodels*, la cual asume una distribución aproximadamente normal de los errores de predicción y considera la varianza asociada al horizonte de pronóstico. Para cada paso temporal, el intervalo se calcula como la media pronosticada $\pm z_{1-\alpha/2}$ veces la desviación estándar del error de predicción, para un nivel de confianza especificado, en este caso del 75 %.

Los modelos fueron ajustados utilizando la serie transformada mediante la transformación Box-Cox. En consecuencia, las predicciones y los intervalos de confianza obtenidos con la función *get_forecast()* de la librería *statsmodels* se generaron inicialmente en la escala transformada. Posteriormente, tanto las predicciones como los límites inferior y superior de los intervalos fueron convertidos nuevamente a la escala original del caudal, mediante la aplicación de la transformación inversa de Box-Cox.

8. Resultados

8.1. Caracterización y resumen de las series de datos

8.1.1. Caudal - variable endógena

La serie de caudal presenta una estacionalidad marcada, coherente con el régimen hidrometeorológico de la costa sur de Guatemala, con valores relativamente bajos durante la época seca (noviembre–abril) y mayores caudales en la época lluviosa (mayo–octubre) (Figura 4a y 4b). Este patrón intra-anual se mantiene de forma consistente a lo largo de los siete años analizados. Las series de tiempo de caudal semanal (máximo, promedio y mínimo) se muestran en la Figura 4a.

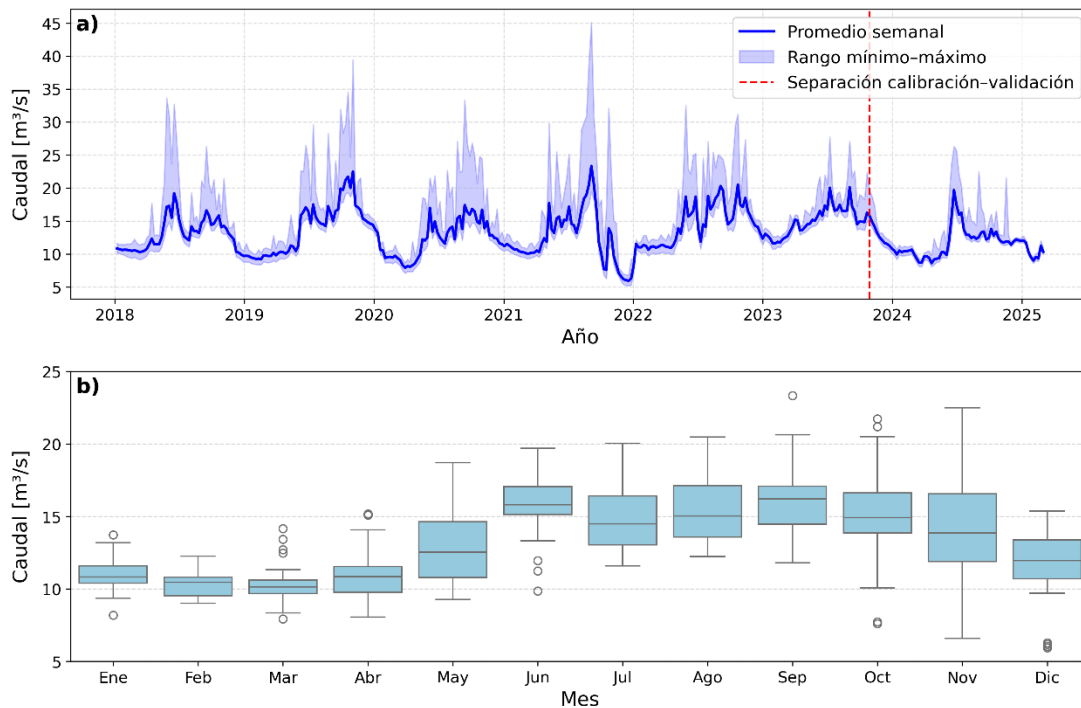


Figura 4. a) Series semanales de caudales mínimo, máximo y promedio en la EH Acarigua y b) Distribución mensual de los valores semanales de caudal promedio, por medio de grafico de cajas.

La cercanía de los caudales promedio al rango de los mínimos sugiere una distribución generalmente asimétrica a escala semanal. En particular, la serie objetivo (caudal promedio semanal) presenta un coeficiente de asimetría de 0.43, indicando un sesgo hacia valores altos (Figura 5a), un comportamiento común en series hidrológicas debido a eventos extremos que generan colas largas hacia la derecha (Helsel et al., 2020).

Con el propósito de reducir esta asimetría y aproximar la distribución a condiciones más cercanas a la normalidad, se aplicó una transformación de Box–Cox. Como resultado, el coeficiente de asimetría se redujo a 0.005, lo que evidencia una mejora sustancial en la simetría de la serie y un mayor cumplimiento de los supuestos requeridos por los modelos estadísticos empleados (Figura 5b).

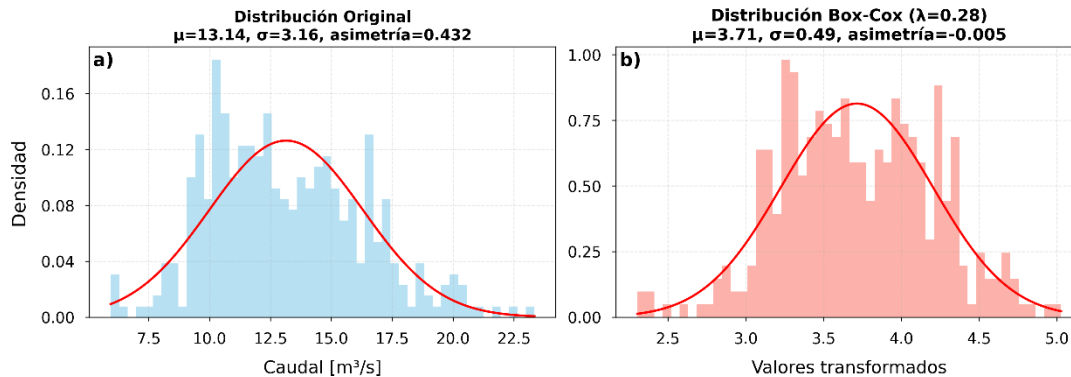


Figura 5. Distribución de caudales promedio con curvas de distribución normal ajustadas (curvas rojas).
a) datos originales, b) datos transformados por medio de Box-Cox.

8.1.2. Variables meteorológicas – variables exógenas

Las series temporales semanales de las variables meteorológicas consideradas como exógenas, aplicables solo en el modelo SARIMAX, se presentan en la Figura 6. En el caso de la precipitación, se registran mayores valores en la estación meteorológica (EM) Cocales en comparación con la EM El Platanar, ubicada a una mayor altitud. Como consecuencia de su forma de cálculo, la serie sintética unificada presenta valores intermedios respecto a ambas estaciones.

Tal como podría anticiparse a partir del comportamiento de la variable endógena, la precipitación presenta una estacionalidad marcada, con una época seca extendida entre noviembre y abril y una época lluviosa entre mayo y octubre (Figura 6a). En contraste, las series de temperatura máxima, promedio y mínima no evidencian de manera visual una estacionalidad tan pronunciada (Figura 6b - 6d). Sin embargo, se observa de forma recurrente un ligero descenso de la temperatura mínima durante los últimos y primeros meses del año calendario. Las temperaturas más altas se observan generalmente durante abril y mayo, período de transición entre la época seca y lluviosa.

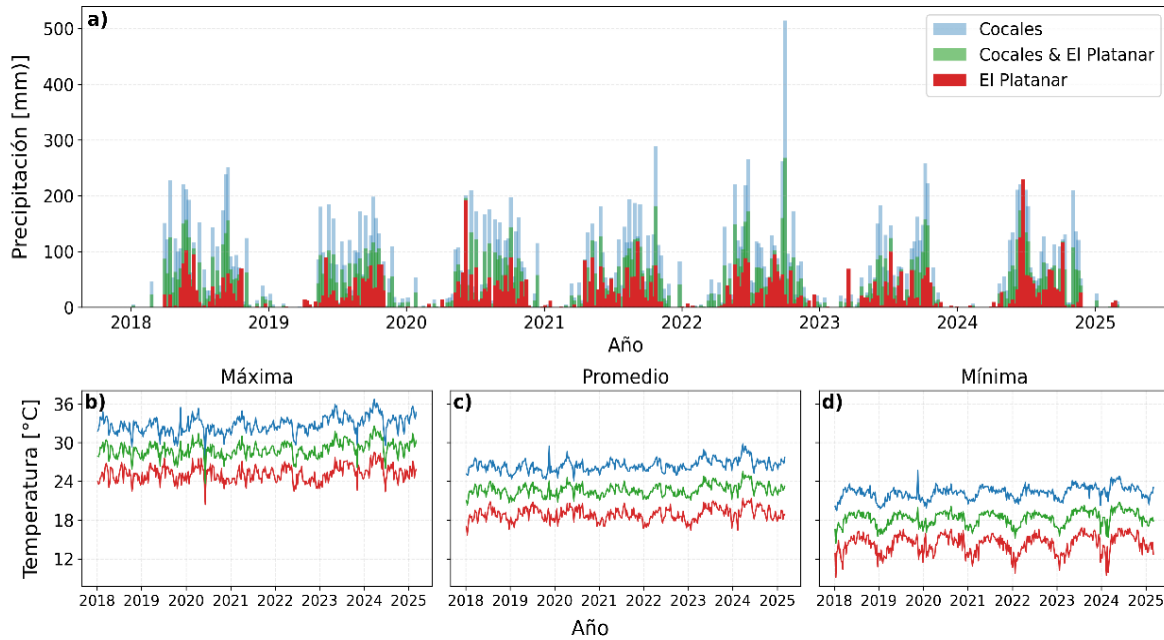


Figura 6. Series semanales de variables meteorológicas empleadas en el modelo SARIMAX. a) Precipitación. b) Temperatura Máxima. c) Temperatura Promedio. d) Temperatura Mínima.

La agrupación de los datos utilizada en el proceso de modelación, correspondiente al periodo total, así como a las fases de calibración y validación, se resume en la Tabla 4. En total, se consideraron 374 semanas; las primeras 304 se emplearon para la calibración y las últimas 70 para la validación de los modelos.

Asimismo, se presenta un resumen estadístico básico de la serie modelada, definida como el caudal promedio semanal (Tabla 4). En la estación hidrométrica (EH Acarigua), la serie de caudales promedio semanales del río Madre Vieja presenta un valor medio de $13.1 \text{ m}^3 \text{ s}^{-1}$, un mínimo de $5.9 \text{ m}^3 \text{ s}^{-1}$ y un máximo de $23.3 \text{ m}^3 \text{ s}^{-1}$. La desviación estándar de esta serie es de $3.1 \text{ m}^3 \text{ s}^{-1}$ (Tabla 4).

Tabla 4. Resumen de rangos de subconjunto de datos con estadísticos básicos del caudal promedio semanal.

Subconjunto de datos	Periodo	n	Promedio (m ³ s ⁻¹)	Mín (m ³ s ⁻¹)	Máx (m ³ s ⁻¹)	Std* (m ³ s ⁻¹)
Completo	01/2018 – 03/2025	374	13.1	5.9	23.3	3.1
Calibración	01/2018 – 10/2023	304	13.4	5.9	23.3	3.2
Validación	11/2023 – 03/2025	70	11.6	8.6	19.6	2.1

*desviación estándar

8.2. Análisis de la variable objetivo (caudal) y su adecuación para los modelos

La descomposición STL (Figura 7) revela una tendencia variable a lo largo del período analizado y un patrón estacional bien definido, lo que indica que la serie en su resolución semanal no es estacionaria. En consecuencia, resulta necesaria la aplicación de transformaciones y diferenciaciones previas a su modelación (Hyndman & Athanasopoulos, 2021; Box et al., 2016).

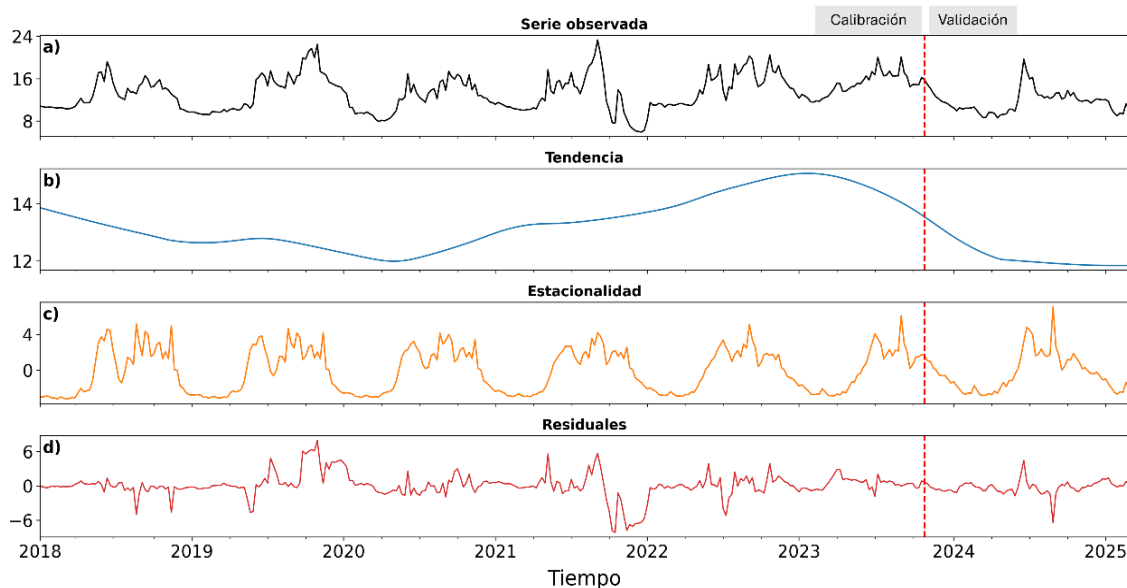


Figura 7. Descomposición STL de la serie original. (a) Serie observada. (b) Componente de tendencia. (c) Componente estacional. (d) Residuales.

Por ello, el subconjunto de calibración de la serie fue sometida a una transformación y dos diferenciaciones. Primero se aplicó una transformación Box–Cox ($\lambda = 0.56$) para estabilizar la varianza y aproximar la serie a condiciones de homocedasticidad en el subconjunto de calibración (Hyndman & Athanasopoulos, 2021). Posteriormente, se aplicaron una diferenciación estacional de 52 semanas y una de primer orden, las cuales atenuaron la estructura temporal remanente y generaron un componente residual cercano al ruido blanco en la descomposición STL (Figura 8).

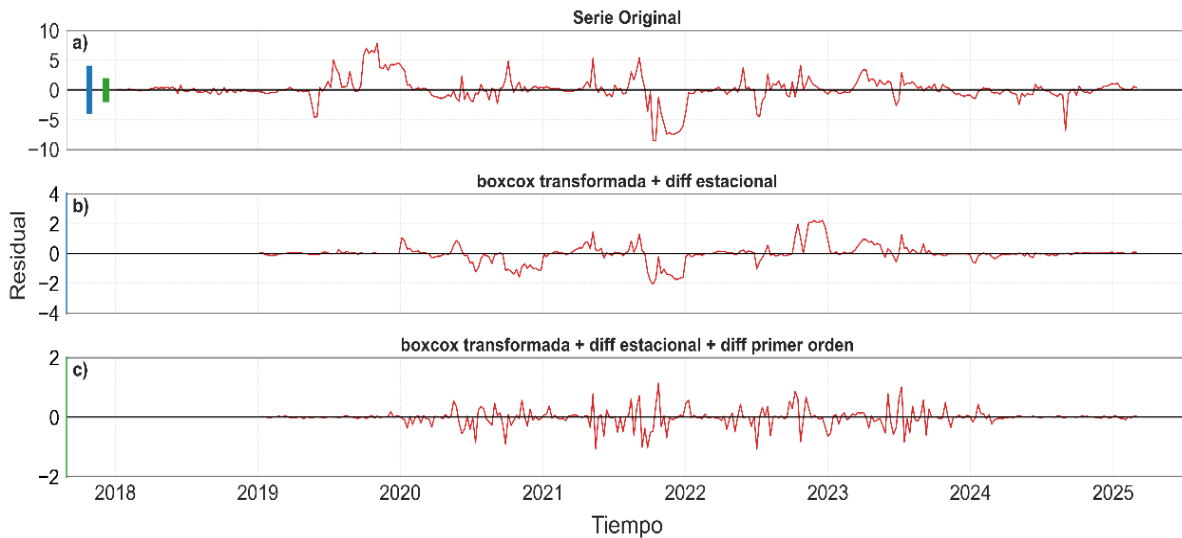


Figura 8. Residuales de descomposición STL de: (a) serie original, (b) serie Box-Cox transformada + diferenciación estacional y (c) serie Box-Cox transformada, diferenciación estacional y diferenciación de primer orden.

El análisis del correlograma de la función de autocorrelación (ACF) y autocorrelación parcial (PACF) juega un papel fundamental en la identificación de modelos autorregresivos (Machiwal & Jha, 2012). La Figura 9a muestra los correlogramas de la ACF y PACF de la serie de caudales semanales promedio.

La ACF presenta un patrón oscilatorio con decaimiento gradual y rezagos estadísticamente significativos hasta aproximadamente 52 periodos, lo que indica una dependencia temporal persistente y sugiere la presencia de componentes autorregresivos ($\rho > 0$). Por otro lado, la PACF muestra que solo los términos autorregresivos inmediatamente anteriores son relevantes o pueden tener influencia en un valor actual de caudal. Por su parte, la PACF muestra significancia principalmente en los primeros rezagos, lo que indica que los valores recientes ejercen mayor influencia sobre el comportamiento actual de la serie (Box et al., 2016).

Tras aplicar la transformación Box-Cox y una diferenciación estacional (Figura 9b) el número y la magnitud de las autocorrelaciones significativas disminuyen notablemente. Al incorporar además una diferenciación de primer orden (Figura 9c), la señal estacional se atenúa casi por completo y la ACF se concentra en rezagos de baja magnitud, lo cual es consistente con que el preprocesamiento aplicado aproxima la serie a condiciones de estacionariedad.

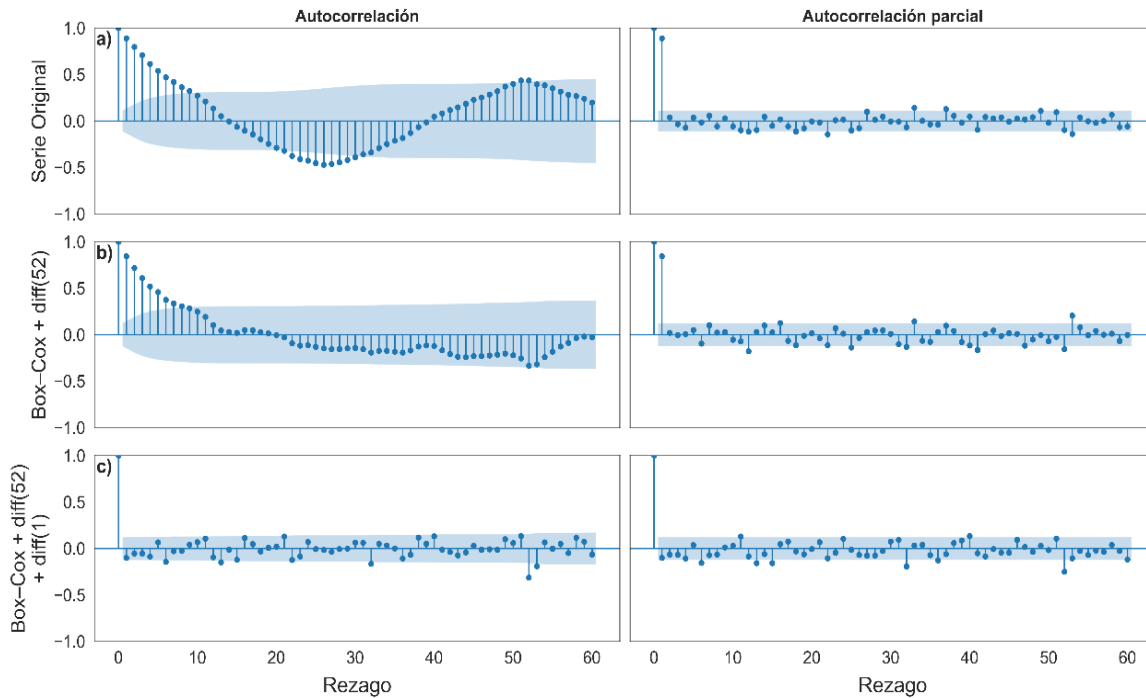


Figura 9. Correlogramas ACF y PACF de la serie de caudal promedio semanal: (a) serie original, (b) serie con transformación Box-Cox y diferenciación estacional ($s = 52$), (c) serie con transformación Box-Cox y diferenciaciones estacional y de primer orden.

La Figura 10 presenta los valores p de la prueba estadística de Ljung-Box aplicados a (a) la serie original de caudales promedio semanales y (b) la serie transformada mediante Box-Cox y diferenciaciones estacional y de primer orden. En la serie original (Figura 10a), los valores p se mantienen por debajo del umbral del 5 % en todos los rezagos evaluados, evidenciando autocorrelación estadísticamente significativa. Esto confirma que la serie no se comporta como ruido blanco y presenta una estructura temporal que justifica su modelación mediante enfoques ARIMA/SARIMA.

Por otra parte, en la serie transformada y diferenciada (Figura 10b) los valores p superan el umbral en los primeros rezagos, indicando una reducción de la autocorrelación de corto plazo. No obstante, en rezagos mayores se mantiene evidencia de dependencia temporal, lo que

sugiere que la serie conserva parte de su estructura. Este resultado es consistente con el objetivo del preprocesamiento: estabilizar la varianza y aproximar la serie a estacionariedad sin eliminar completamente su dinámica temporal (Hyndman & Athanasopoulos, 2021). Una vez analizada la estructura temporal de la serie, se procedió al modelado.

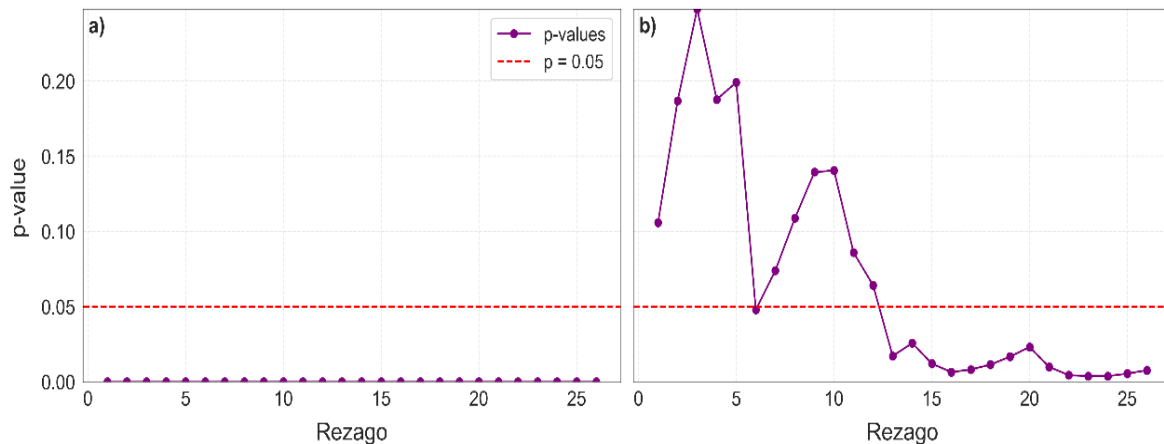


Figura 10. Resultados de la prueba de Ljung–Box para un rango de rezagos: (a) serie original y (b) serie Box-Cox transformada con diferenciación estacional y de primer orden.

8.3. Resultados del modelo ARIMA

En términos generales, las combinaciones del modelo ARIMA ajustados sobre la serie transformada mediante Box–Cox y diferenciación (estacional y de primer orden) durante la etapa de calibración fueron las que alcanzaron mejor desempeño, evidenciado por los valores más altos de NSE y los RMSE más bajos. La leyenda de la Figura 11 presenta las doce mejores combinaciones ARIMA obtenidas durante la fase de calibración.

Entre ellas, las series resaltadas en color corresponden a las combinaciones del modelo cuyos residuales presentaron valores p superiores al nivel de significancia de la prueba de Ljung–Box para un horizonte de 52 rezagos (equivalente a un año hidrológico) (Figura 11). En otras palabras, en fase de calibración no se detectó autocorrelación significativa en los residuales, lo que sugiere que el modelo capturó adecuadamente la estructura temporal de la serie. Este criterio constituye un requisito fundamental para la validez estadística del modelo y respalda su potencial capacidad predictiva (Ljung & Box, 1978). Considerando esto, únicamente tres combinaciones de las 12 evaluadas superaron este filtro tanto en fase de calibración como de validación. El modelo ARIMA (6,1,8) presentó el mejor desempeño, con valores de RMSE de 1.9 en calibración y 1.2 en validación, y valores de NSE de 0.650 y 0.655, respectivamente (Tabla 5). Estos resultados permiten determinar que su capacidad predictiva es satisfactoria a buena (ver Tabla 3).

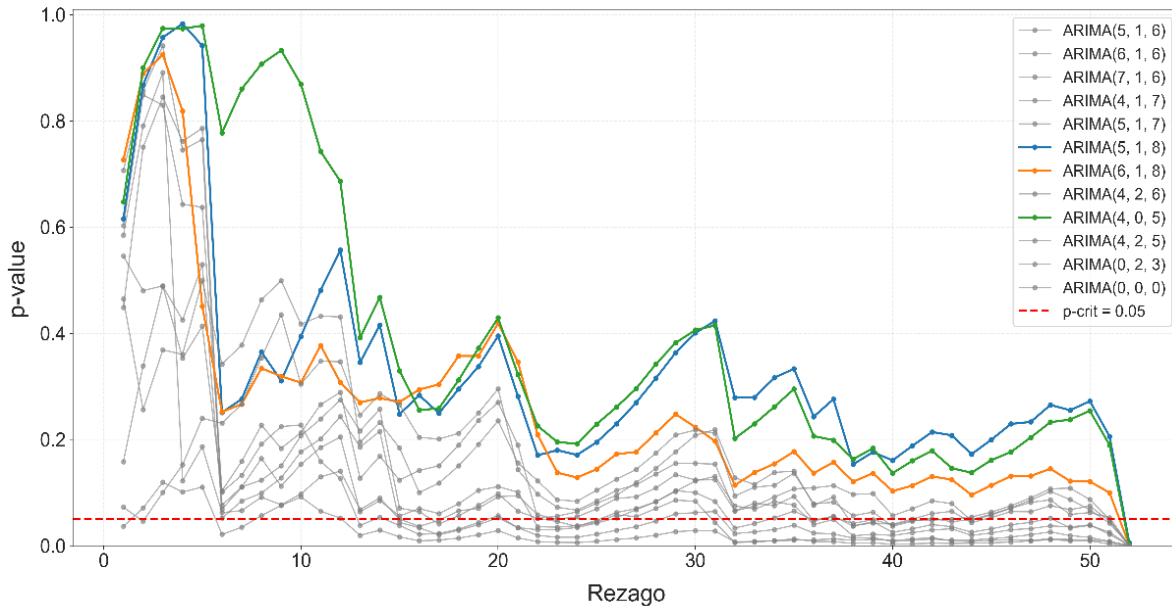


Figura 11. Resultados de la prueba de Ljung–Box para un rango de rezagos (1–52) de los residuales de los 12 mejores modelos ARIMA en calibración.

Tabla 5. Desempeño de los tres mejores modelos ARIMA en calibración y validación.

Modelo	RMSE cal	Calif.*	NSE cal	Calif.*	RMSE val	Calif.*	NSE val	Calif.*
(5, 1, 8)	2.033	Satisfact.	0.635	Satisfact.	1.243	Satisfact.	0.653	Bueno
(6, 1, 8)	1.989	Satisfact.	0.650	Bueno	1.239	Satisfact.	0.655	Bueno
(4, 0, 5)	1.995	Satisfact.	0.650	Bueno	1.250	Satisfact.	0.649	Satisfact.

La representación gráfica de los residuales del mejor modelo ARIMA identificado (Figura 12) indica que la serie simulada captura de manera adecuada la estructura observada en la fase de calibración. En efecto, los residuales presentan un comportamiento compatible con ruido blanco, sin autocorrelación significativa y con una distribución aproximadamente normal y media cercana a cero.

La serie observada y la simulada por el mejor modelo ARIMA (6,1,8) durante las fases de calibración y validación se muestran en la Figura 13. Visualmente, el modelo reproduce con una precisión relativamente buena los caudales mínimos de la época seca, mientras que tiende a sobreestimar algunos picos de la época lluviosa y a subestimar ciertos descensos dentro del mismo periodo.

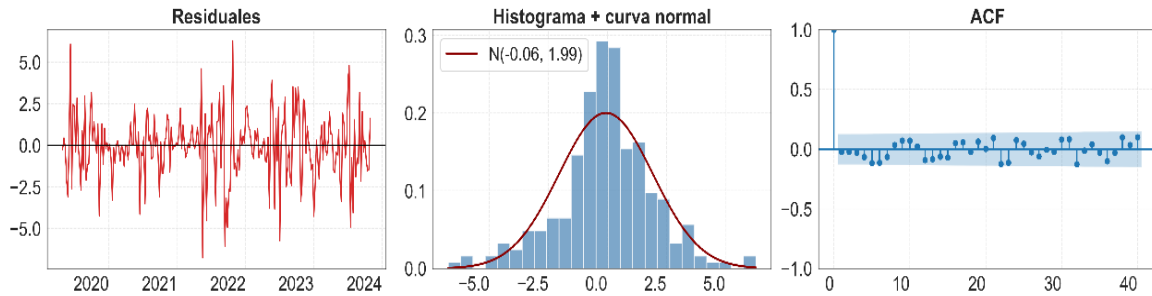


Figura 12. Análisis gráfico de los residuales del modelo ARIMA (6,1,8) seleccionado como el de mejor desempeño.

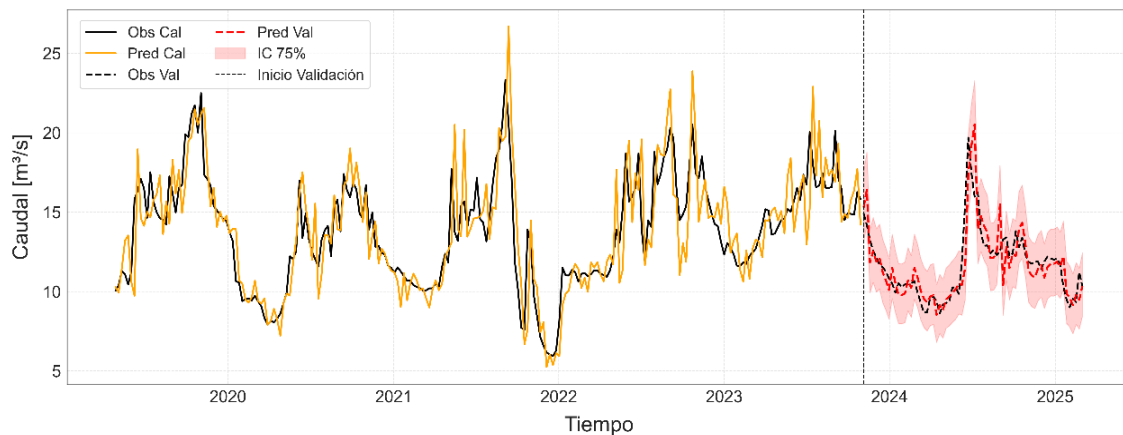


Figura 13. Comparación entre valores observados y simulados del modelo ARIMA (6,1,8) en calibración y validación, con intervalo de confianza del 75 % para la fase predictiva.

En la fase de validación se empleó un intervalo de confianza (IC) del 75 % asociado a la predicción (Figura 13) lo que implica que, bajo los supuestos del modelo, se espera que aproximadamente el 75 % de las observaciones futuras se ubiquen dentro del rango estimado (Hyndman & Athanasopoulos, 2021). En este caso, el intervalo es relativamente estrecho y contiene la totalidad de la serie observada durante el período analizado, lo que sugiere que el modelo representa de forma consistente la variabilidad en la fase de validación.

8.4. Resultados del modelo SARIMA

En el modelo SARIMA, la incorporación explícita de componentes estacionales, diferenciaciones y términos autorregresivos y de medias móviles permitió trabajar tanto con la serie original como con la serie transformada mediante Box–Cox. Las combinaciones de hiperparámetros ajustadas sobre esta última mostraron un mejor desempeño relativo en las fases de calibración y validación. La Figura 14 presenta las doce mejores combinaciones en SARIMA obtenidas tras el proceso de búsqueda y selección inicial.

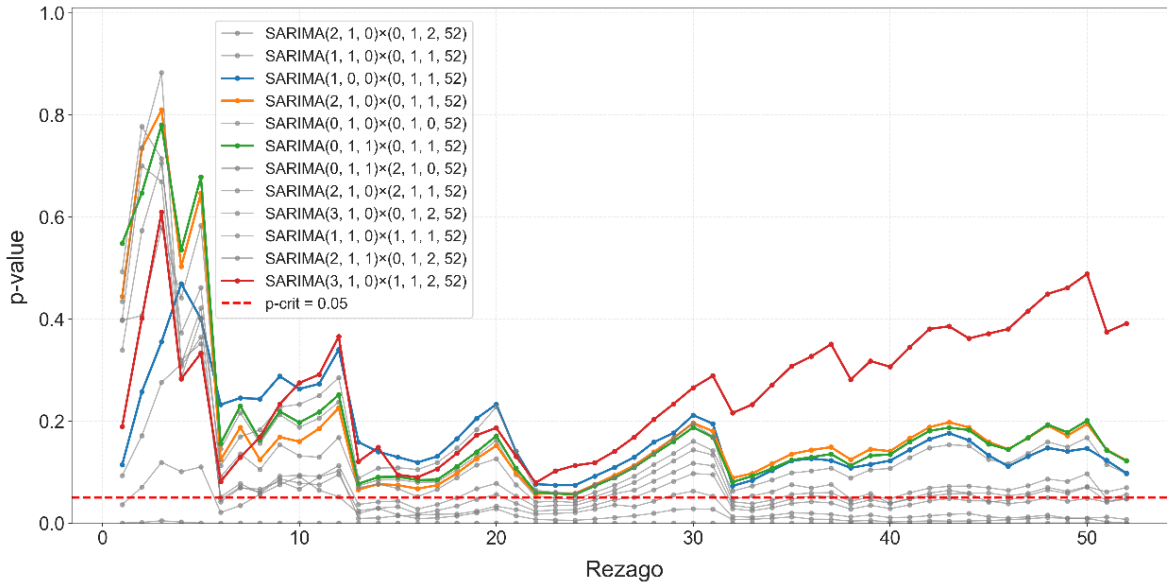


Figura 14. Resultados de la prueba de Ljung–Box para un rango de rezagos (1–52) de los residuales de los 12 mejores modelos SARIMA en calibración.

Al igual que en el análisis ARIMA, se aplicó la prueba de Ljung–Box con un horizonte de 52 rezagos (equivalente a un año hidrológico) para identificar aquellas combinaciones cuyos residuales no presentan autocorrelación acumulada significativa. Las que superaron el nivel de significancia (Figura 14) fueron considerados como los que capturaron con mayor fidelidad la estructura temporal de la serie en la fase de calibración.

Posteriormente, las cuatro mejores combinaciones que cumplieron la prueba, junto con un quinto modelo que no la superó, pero destacó por sus métricas de desempeño, fueron evaluadas con mayor detalle (Tabla 6). El modelo seleccionado como mejor SARIMA fue el (2,1,0), (0,1,2,52). Este presentó métricas entre satisfactorias y buenas en calibración; sin embargo, en la fase de validación su desempeño fue insatisfactorio. Otro modelo, como el (1,0,0), (0,1,1,52), mostró un ajuste muy bueno en calibración, pero un rendimiento limitado fuera de muestra, evidenciando problemas de generalización. Aún siendo el mejor dentro del grupo SARIMA, el mejor modelo seleccionado no alcanza un desempeño adecuado en validación.

La Figura 15 muestra que los residuales del modelo presentan un comportamiento compatible con ruido blanco, sin autocorrelación significativa y con una distribución aproximadamente normal, lo que indica que la estructura temporal fue capturada adecuadamente en la fase de calibración. Por su parte, la Figura 16 compara la serie observada y la simulada en ambas fases.

En calibración, el modelo reproduce adecuadamente los valores más pronunciados de la época seca; sin embargo, durante la época lluviosa el ajuste resulta menos satisfactorio.

Tabla 6. Desempeño de los cinco mejores modelos SARIMA en calibración y validación.

Modelo	RMS E cal	Calif.	NSE cal	Calif.	RMS E val	Calif.	NSE val	Calif.
(1,0,0), (0,1,1,52)	1.693	Satisfact.	0.753	Muy bueno	2.295	Insatisfac.	-0.180	Insatisfac.
(2,1,0), (0,1,1,52)	1.743	Satisfact.	0.737	Bueno	2.035	Insatisfac.	0.072	Insatisfac.
(0,1,1), (0,1,1,52)	1.739	Satisfact.	0.739	Bueno	2.064	Insatisfac.	0.045	Insatisfac.
(2,1,0), (0,1,2,52)	1.839	Satisfact.	0.707	Bueno	1.901	Insatisfac.	0.190	Insatisfac.
(3,1,0), (0,1,2,52)	1.841	Satisfact.	0.706	Bueno	1.907	Insatisfac.	0.184	Insatisfac.

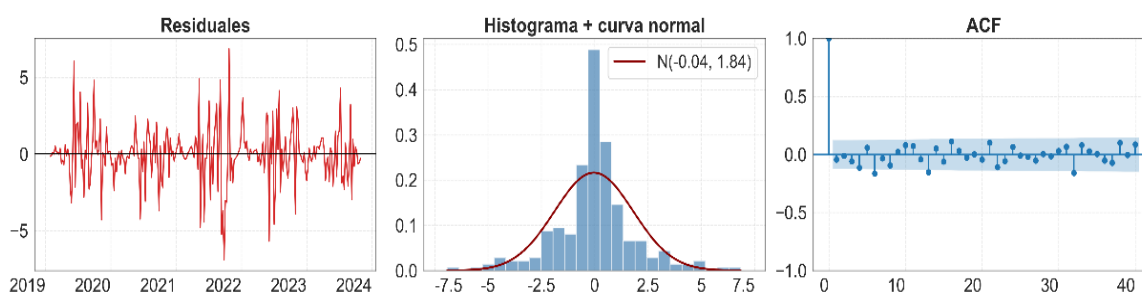


Figura 15. Análisis gráfico de los residuales del modelo SARIMA (2,1,0), (0,1,2, 52) seleccionado como el de mejor desempeño.

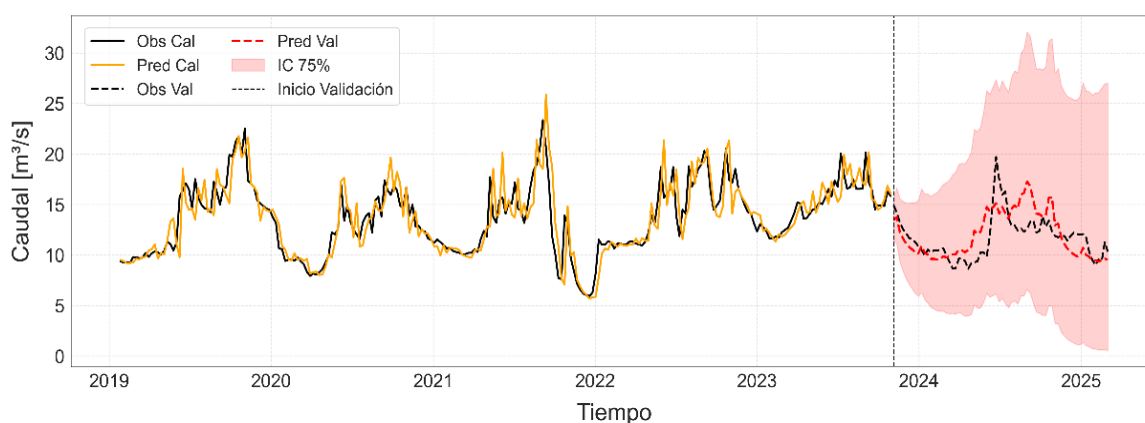


Figura 16. Comparación entre valores observados y simulados del modelo SARIMA (2,1,0), (0,1,2, 52) en calibración y validación, con intervalo de confianza del 75 % para la fase predictiva.

En validación, el desempeño es insatisfactorio ($RMSE = 1.9$; $NSE = 0.190$), evidenciando limitaciones para reproducir tanto la magnitud como la temporalidad de los picos y valles característicos del año hidrológico evaluado. El intervalo de confianza presenta un rango amplio y tiende a incrementarse con el horizonte de predicción (Figura 16), reflejando la acumulación progresiva de la incertidumbre del modelo, comportamiento habitual en modelos SARIMA con componentes estacionales activos (Hyndman & Athanasopoulos, 2021), es decir, cuando al menos uno de los términos estacionales (P, D o Q) es distinto de cero.

8.5. Resultados del modelo SARIMAX

De forma análoga al modelado SARIMA, la especificación explícita de componentes estacionales en SARIMAX permitió trabajar con la serie original de caudales y su versión transformada mediante Box–Cox, sin aplicar diferenciaciones adicionales a la variable objetivo. Las variables exógenas se utilizaron en su forma original, sin transformaciones ni diferenciaciones, con el fin de preservar su interpretación hidrometeorológica.

La Figura 17 presenta las doce mejores combinaciones SARIMAX seleccionadas con base en las métricas RMSE y NSE. Posteriormente, la aplicación de la prueba de Ljung–Box con un horizonte de 52 rezagos permitió identificar aquellos modelos cuyos residuales no presentan autocorrelación significativa. Las configuraciones que superaron este criterio fueron consideradas como las que capturaron con mayor fidelidad la estructura temporal de la serie. Siete combinaciones de hiperparámetros superaron la prueba y fueron comparadas mediante las métricas RMSE y NSE en las fases de calibración y validación. Entre ellas, el modelo SARIMAX (1,1,2), (0,1,1,52) mostró el mejor desempeño global, con RMSE de 1.5 en calibración y 1.4 en validación, y valores de NSE de 0.800 y 0.558, respectivamente (Tabla 7).

Este modelo incluyó como variables exógenas la precipitación, la temperatura promedio y la temperatura máxima representativas de las condiciones río arriba de la EH Acarigua. Asimismo, se incorporó el índice Niño 3.4 (N34), tanto sin rezago como con un rezago de seis meses, con el objetivo de capturar posibles efectos climáticos interanuales sobre la dinámica del caudal.

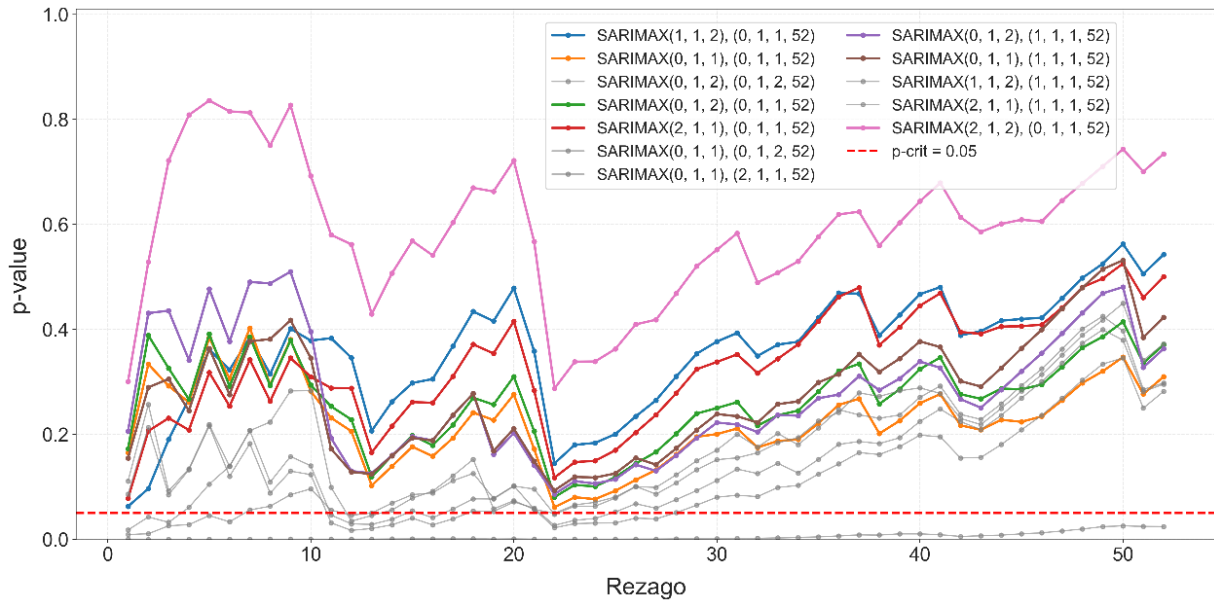


Figura 17. Resultados de la prueba de Ljung–Box para un rango de rezagos (1–52) de los residuales de los 12 mejores modelos SARIMAX en calibración.

Tabla 7. Desempeño de los mejores modelos SARIMAX en calibración y validación.

Modelo	RMSE cal	Calif.	NSE cal	Calif.	RMSE val	Calif.	NSE val	Calif.
(1, 1, 2), (0, 1, 1, 52)	1.518	Satisfact.	0.800	Muy bueno	1.404	Satisfact.	0.558	Satisfact.
(0, 1, 1), (0, 1, 1, 52)	1.529	Satisfact.	0.798	Muy bueno	1.501	Satisfact.	0.495	Insatisfa.
(0, 1, 2), (0, 1, 1, 52)	1.534	Satisfact.	0.797	Muy bueno	1.533	Satisfact.	0.473	Insatisfa.
(2, 1, 1), (0, 1, 1, 52)	1.555	Satisfact.	0.790	Muy bueno	1.545	Satisfact.	0.465	Insatisfa.
(0, 1, 2), (1, 1, 1, 52)	1.618	Satisfact.	0.774	Muy bueno	1.986	Insatisfa.	0.116	Insatisfa.
(0, 1, 1), (1, 1, 1, 52)	1.616	Satisfact.	0.775	Muy bueno	1.998	Insatisfa.	0.106	Insatisfa.
(2, 1, 2), (0, 1, 1, 52)	1.502	Satisfact.	0.804	Muy bueno	2.139	Insatisfa.	-0.025	Insatisfa.

La Figura 18 muestra que los residuales del modelo presentan un comportamiento compatible con ruido blanco, sin autocorrelación significativa, con media cercana a cero y con una distribución aproximadamente normal, lo que sugiere que el modelo capturó adecuadamente la estructura temporal de la serie durante la fase de calibración.

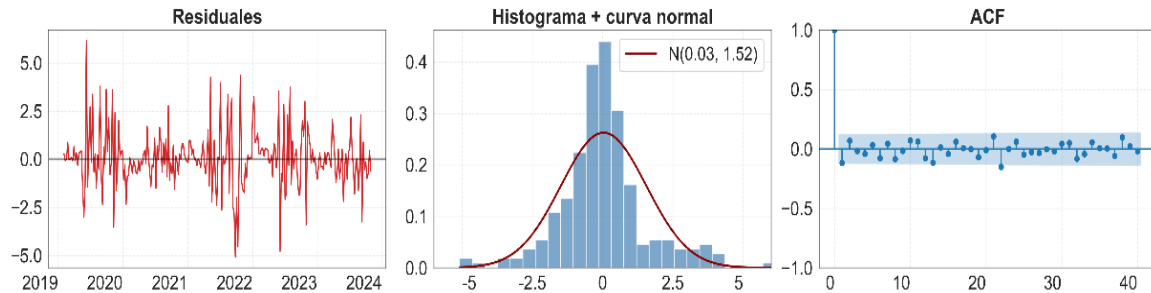


Figura 18. Análisis gráfico de los residuales del modelo SARIMAX (1,1,2), (0,1,1, 52) seleccionado como el de mejor desempeño.

La Figura 19 muestra la comparación entre la serie observada y la simulada por el mejor modelo SARIMAX (1,1,2), (0,1,1, 52) en las fases de calibración y validación. En calibración, el modelo reproduce adecuadamente los valles más pronunciados asociados a la época seca; sin embargo, durante la época lluviosa el ajuste es menos satisfactorio, ya que tiende a representar fluctuaciones de baja magnitud no presentes en la serie observada. El RMSE (1.5) y el NSE (0.800) reflejan un desempeño entre satisfactorio y muy bueno en calibración.

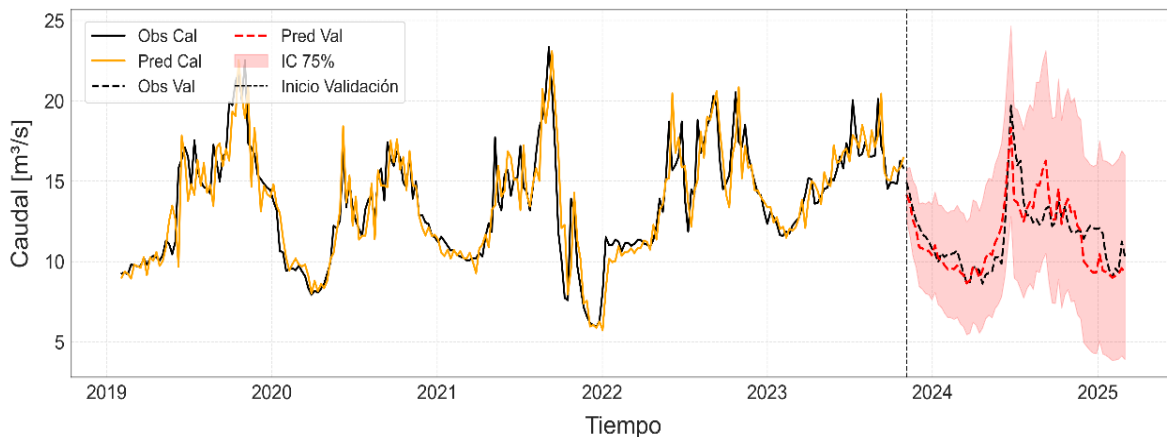


Figura 19. Comparación entre valores observados y simulados del modelo SARIMAX (1,1,2), (0,1,1, 52) en calibración y validación, con intervalo de confianza del 75 % para la fase predictiva.

En la fase de validación, el modelo presenta un desempeño satisfactorio (RMSE = 1.4; NSE = 0.558), lo que indica una capacidad predictiva moderadamente buena, aunque con ciertas limitaciones. Estas se evidencian principalmente durante la época lluviosa y hacia horizontes de predicción más lejanos (Figura 19). Asimismo, el intervalo de confianza se amplía rápidamente con el horizonte de predicción, evidenciando la acumulación de incertidumbre inherente al modelo.

9. Discusión

En primera instancia, los análisis gráficos y estadísticos evidenciaron que la serie de caudal promedio semanal considerada presenta una estacionalidad marcada y dependencia temporal. Estas condiciones respaldan su modelación mediante enfoques estadísticos, en este caso, mediante modelos autorregresivos integrados de medias móviles.

Los resultados de las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF), por ejemplo, evidencian la presencia de dependencia temporal, indicando que los valores pasados de la serie tienen la capacidad de representar algunos aspectos del comportamiento de los datos en el futuro. Otro ejemplo es que la prueba de Ljung–Box indicó que persistía autocorrelación estadísticamente significativa, aún después del preprocesamiento de la serie (por medio de transformación Box–Cox, seguida de una diferenciación estacional y una diferenciación de primer orden). Estos resultados respaldaron el enfoque de modelación adoptado y motivan la comparación entre los modelos ARIMA, SARIMA y SARIMAX (Hyndman & Athanasopoulos, 2021; Box et al., 2016).

En términos generales, las métricas RMSE y NSE muestran que el ajuste fue mejor en calibración que en validación para la gran mayoría de combinaciones de hiperparámetros evaluadas para cada uno de los tres modelos. Sin embargo, el modelo ARIMA (6,1,8) presentó resultados muy similares en ambas fases, e incluso ligeramente mejores en validación (NSE = 0.650 en calibración y 0.655 en validación). La **Figura 20** presenta la serie de tiempo de caudales promedio semanales junto con las predicciones generadas por cada uno de los tres modelos evaluados, considerando en cada caso la combinación de hiperparámetros que mostró el mejor desempeño según las métricas RMSE y NSE.

En la fase de calibración, los tres modelos alcanzaron desempeños entre satisfactorios y muy buenos ($1.51 \leq \text{RMSE} \leq 1.98$ y $0.65 \leq \text{NSE} \leq 0.80$). No obstante, todos presentaron ciertas limitaciones, evidenciadas en pequeñas oscilaciones en las series simuladas, particularmente durante periodos de alta variabilidad de caudales en la época lluviosa. En contraste, los descensos más pronunciados de la época seca fueron reproducidos con mayor estabilidad.

Cabe destacar que la inclusión explícita de componentes estacionales y variables exógenas contribuyó a reducir esta variabilidad intraestacional, lo que explica el mejor desempeño en fase de calibración del modelo SARIMAX frente a SARIMA y ARIMA en este trabajo. Esto ha sido documentado en estudios previos (Costa et al., 2023; Kim et al., 2019). Desde una perspectiva

hidrológica, los componentes estacionales permiten capturar patrones recurrentes del régimen anual, mientras que las variables exógenas incorporan información climática adicional que mejora la representación de la respuesta del sistema.

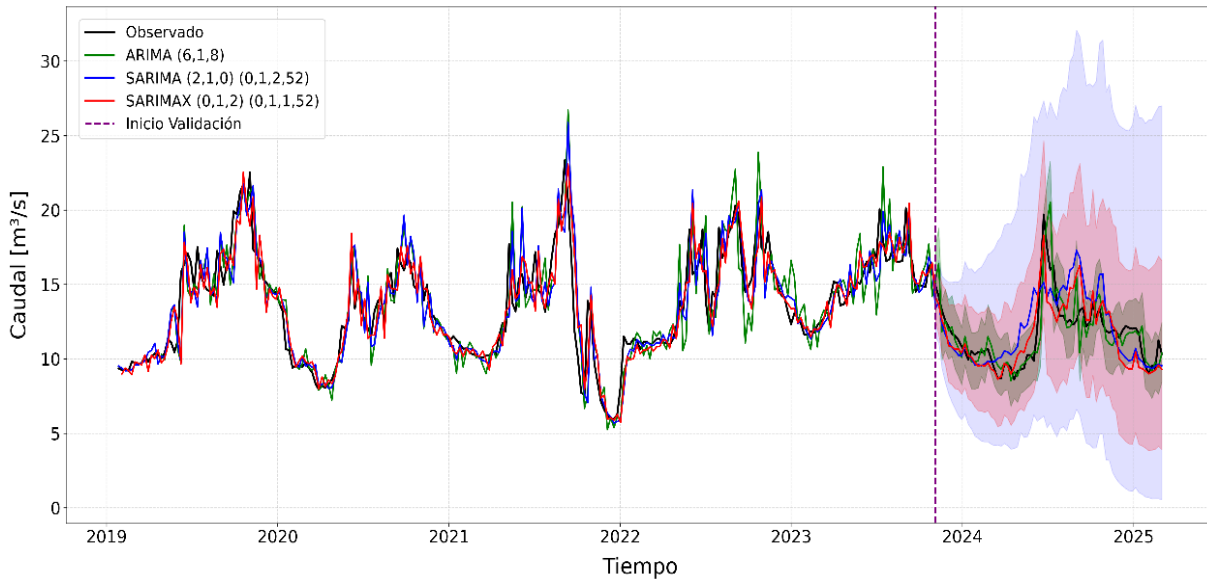


Figura 20. Comparación entre serie observada, modelos ARIMA, SARIMA y SARIMAX en calibración y validación.

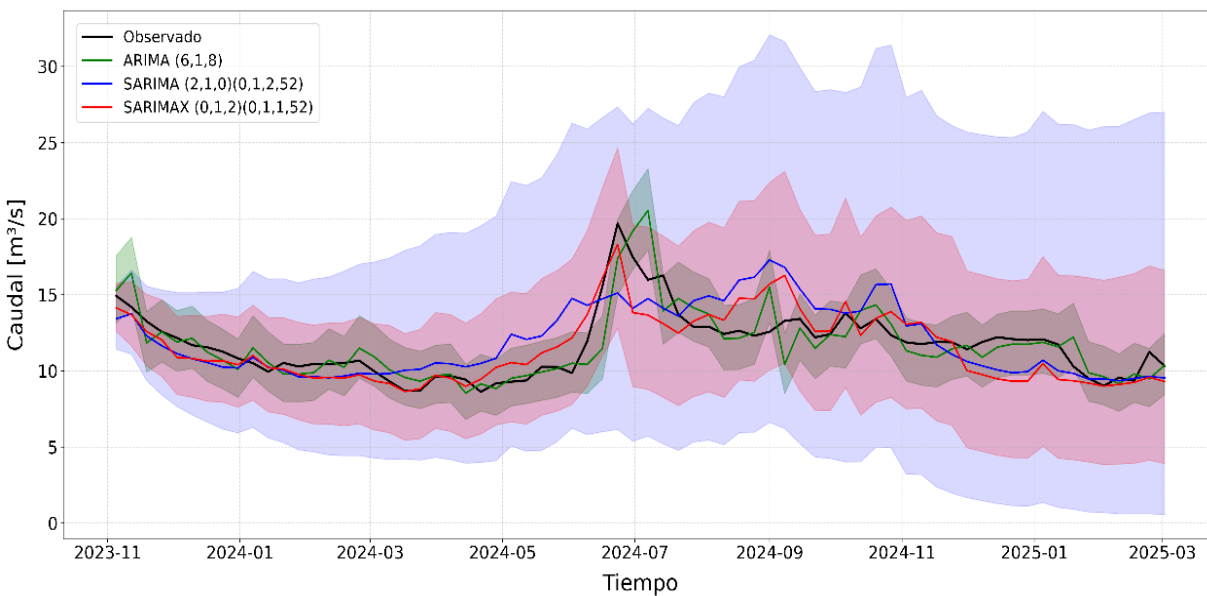


Figura 21. Comparación entre serie observada, modelos ARIMA, SARIMA y SARIMAX en validación con un intervalo de confianza (IC) del 75 %.

En la fase de validación, el desempeño de los modelos fue variable, con ajustes que oscilaron entre insatisfactorios y buenos ($1.23 \leq \text{RMSE} \leq 1.90$; $0.190 \leq \text{NSE} \leq 0.655$). Esto evidencia que la capacidad predictiva fuera de muestra no fue homogénea. En general, los modelos con componente estacional explícito (SARIMA y SARIMAX) no superaron al modelo ARIMA, que es relativamente más simple, lo que sugiere que la incorporación de estacionalidad estructural no garantiza por sí sola una mejora en la predicción.

No obstante, entre los modelos estacionales, el SARIMAX mostró un mejor desempeño que el SARIMA, lo que indica que la inclusión de variables exógenas aporta información relevante para explicar la dinámica del caudal y mejorar el ajuste fuera de muestra. En consecuencia, la incorporación de forzantes climáticos emerge como un factor que puede fortalecer la capacidad predictiva en esta aplicación.

A pesar de ello, es importante considerar que fue necesario incorporar hasta ocho series exógenas, lo que plantea el debate sobre el equilibrio entre robustez y complejidad del modelo. Esta consideración resulta especialmente relevante ante una posible implementación operativa, donde la dependencia de múltiples variables externas puede limitar su aplicabilidad práctica. La literatura en predicción hidrológica señala que un aumento en la complejidad no necesariamente se traduce en una mayor capacidad predictiva y puede afectar la implementación real del modelo (Hyndman & Athanasopoulos, 2021; Montanari & Koutsoyiannis, 2014).

Asimismo, la disponibilidad futura de las variables exógenas constituye un aspecto crítico. No todas están garantizadas en tiempo real y, cuando provienen de modelos globales o productos por celda, su precisión e incertidumbre deben contextualizarse a escala local. Desde esta perspectiva operativa, el modelo menos complejo evaluado (ARIMA) fue el que presentó el mejor desempeño en validación, lo que le confiere mayor potencial de aplicabilidad para un sistema de pronóstico de caudales en la sección media del río Madre Vieja.

Adicionalmente, el análisis de los intervalos de confianza asociados a las predicciones de cada modelo mostró que el ARIMA (6,1,8) presentó los niveles más bajos de incertidumbre alrededor de la predicción (Figura 21). En este sentido, el intervalo de confianza del 75 % evidenció un desempeño relativo superior del modelo ARIMA (6,1,8), seguido por el SARIMAX (1,1,2), (0,1,1,52) y, finalmente, el SARIMA (2,1,0), (0,1,2,52).

En síntesis, para la predicción de caudales promedio semanales en la sección media del río Madre Vieja, el modelo ARIMA (6,1,8) constituye la alternativa más adecuada dentro del conjunto de modelos evaluados en este estudio. Aunque los modelos con estacionalidad explícita y

variables exógenas mostraron ventajas puntuales, especialmente en calibración, el ARIMA destacó por su mayor consistencia entre calibración y validación, así como por menores niveles de incertidumbre asociados a la predicción. Estas características, junto con su menor dependencia de información externa, le otorgan mayor robustez y viabilidad operativa.

El hecho de que fuera necesaria la aplicación de una transformación Box–Cox, junto con diferenciaciones estacional y de primer orden, para lograr un ajuste satisfactorio del modelo ARIMA evidencia la presencia de heterocedasticidad y no estacionariedad en la serie original de caudal. Esto sugiere que la variabilidad del caudal no es estrictamente constante en el tiempo y que existen cambios sistemáticos en su nivel medio, rasgos característicos de sistemas hidrológicos influenciados por controles climáticos interanuales (Ward et al., 2014) y variaciones estacionales (Dettinger & Diaz, 2000).

Asimismo, los órdenes elevados del modelo ARIMA (6,1,8) indican una dinámica temporal compleja, donde el caudal depende de múltiples valores pasados y de una propagación amortiguada de perturbaciones. Este comportamiento es consistente con respuestas acumulativas y retardadas del sistema, posiblemente asociadas a procesos hidrológicos a escala cuenca como el almacenamiento de agua, intercepción, atenuación de escorrentía superficial y subterránea, y la propia respuesta de los diferentes usos de suelo a las condiciones hidrometeorológicas (de Lavenne et al., 2022).

El desempeño relativamente inferior de SARIMA y SARIMAX en validación puede relacionarse con la forma en que incorporan la estacionalidad explícita, basada en patrones observados en años previos. Si el conjunto de entrenamiento no representa adecuadamente la diversidad de escenarios hidrológicos posibles, la capacidad del modelo para generalizar y reproducir años atípicos se ve limitada, como parece haber ocurrido en este estudio.

Finalmente, la resolución temporal semanal, en un contexto estacional de hasta seis meses, introduce alta variabilidad intraestacional y reduce la correspondencia directa entre semanas equivalentes de distintos años. En este sentido, la escala temporal emerge como un factor clave en la idoneidad de modelos con estacionalidad explícita. Es posible que SARIMA y SARIMAX presenten mejor desempeño a escalas más agregadas, como la mensual, donde la estacionalidad tiende a ser más estable y representativa.

10. Conclusiones y recomendaciones

Los resultados obtenidos evidencian que la serie temporal de caudal promedio semanal analizada presenta una estructura temporal relativamente definida, caracterizada por una marcada estacionalidad asociada a la alternancia entre la época seca y la lluviosa. Estas características permitieron su representación mediante modelos autorregresivos integrados de medias móviles, una vez verificado el cumplimiento de los supuestos de estacionariedad requeridos. En términos generales, los modelos ARIMA, SARIMA y SARIMAX lograron reproducir adecuadamente la estacionalidad anual (mediante predicciones a escala semanal), particularmente la dinámica de la época seca, aunque presentaron mayores limitaciones durante la época lluviosa, especialmente ante episodios de alta fluctuación en caudales.

En particular, los modelos que incorporan términos estacionales explícitos en su estructura (SARIMA y SARIMAX) mostraron un mejor ajuste durante la fase de calibración en comparación con el modelo ARIMA. Sin embargo, en la fase de validación el modelo ARIMA presentó un desempeño superior, evidenciando mayor capacidad de generalización y una reproducción más consistente de la dinámica del caudal correspondiente al año hidrológico evaluado.

En consecuencia, tras evaluar y comparar el desempeño predictivo de los modelos ARIMA, SARIMA y SARIMAX en este estudio, se concluye que el modelo ARIMA (6,1,8) es el más adecuado para predecir el caudal promedio semanal en la sección media del río Madre Vieja. Dada la representatividad temporal de las condiciones climáticas e hidrológicas observadas durante el período de estudio, un modelo basado en la dependencia temporal de valores pasados resultó más eficaz que aquellos con estacionalidad explícita, la cual, en esta aplicación, fue capturada de manera implícita mediante la diferenciación previa de la serie.

No obstante, la aplicabilidad de modelos como SARIMA o SARIMAX no debe descartarse. A medida que se disponga de series más extensas y representativas de la variabilidad climática e hidrológica, estos enfoques podrían incrementar su robustez, particularmente en la representación de episodios de alta variabilidad. En este sentido, se recomienda continuar la ampliación y consolidación de los registros hidrométricos y meteorológicos, así como reevaluar periódicamente el desempeño de modelos con estacionalidad explícita e inclusión de variables exógenas adecuadamente vinculadas al comportamiento del caudal.

Bajo el estado actual de disponibilidad de datos, se recomienda priorizar modelos de la familia ARIMA sin componentes estacionales explícitos, evaluando la incorporación selectiva de variables exógenas, por ejemplo, mediante enfoques ARIMAX para capturar de mejor forma la variabilidad asociada a eventos puntuales. Asimismo, se sugiere explorar técnicas de *machine learning* como alternativas complementarias y/o modelos hidrológicos basados en procesos físicos. Estos últimos pueden ser una alternativa muy interesante debido a que su estructura generalmente incorpora diversas condiciones hidrológicas que permiten representar de manera más explícita la dinámica del sistema en estudio.

En una fase operativa del sistema de predicción considerando el modelo ARIMA, se recomienda evaluar un horizonte de pronóstico de varias semanas, por ejemplo, de un mes (cuatro semanas), que contribuya a anticipar niveles bajos durante la época seca y apoyar la toma de decisiones. Asimismo, se sugiere adoptar esquemas de actualización continua como el *rolling forecast origin* (Hyndman & Athanasopoulos, 2021), en los cuales el modelo se recalibra de manera iterativa con nuevas observaciones, generando predicciones sobre una ventana fija y facilitando la evaluación permanente de su desempeño ante cambios en la dinámica hidrológica.

Por otro lado, si el objetivo es el pronóstico de eventos de crecida, se recomienda explorar modelos orientados a escalas temporales más finas (diaria u horaria), especialmente durante la época lluviosa. En este contexto, además de los enfoques estadísticos, los modelos basados en procesos físicos constituyen una alternativa adecuada cuando se busca una mejor comprensión del sistema y de las dinámicas hidrológicas subyacentes, las cuales pueden ser particularmente complejas en la vertiente del Pacífico. Por su parte, los métodos de aprendizaje automático pueden ofrecer ventajas cuando el objetivo principal es maximizar la precisión predictiva, especialmente en sistemas caracterizados por relaciones no lineales.

11. Bibliografía

- Adnan, R. M., Yuan, X., Kisi, O., & Curtef, V. (2017). Application of Time Series Models for Streamflow Forecasting. *Civil and Environmental Research*, 9(3), 56-63.
- Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- Al-Saati, N. H., Omran, I. I., Salman, A. A., Al-Saati, Z., & Hashim, K. S. (2021). Statistical modeling of monthly streamflow using time series and artificial neural network models: Hindiya Barrage as a case study. *Water Practice and Technology*, 16(2), 681–691. doi:<https://doi.org/10.2166/wpt.2021.012>
- Argueta Mayorga, J. L., Castillo montes, M. V., Bardales Espinoza, W. A., & Gálvez Sinibaldi, A. S. (2025). Regionalization of precipitation in Guatemala: ENSO influence and patterns. *Atmósfera*, 39, 321-347.
- Associated Programme on Flood Management (APFM). (2013). *Integrated Flood Management: A Tool for Integrated Water Resources Management*. World Meteorological Organization & Global Water Partnership.
- Beven, K. (2012). *Rainfall–runoff modelling: The primer*. Wiley-Blackwell.
- Box, G. E., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. San Francisco: Holden-Day.
- Box, G. E., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time series analysis: Forecasting and control* (5th ed.). Wiley.
- CONRED. (2021). *Sistematización de información actualizada para la generación del mapa de riesgo por inundaciones y sequía en la cuenca del río Coyolate*. Guatemala.
- Costa, G. E., Menezes Filho, F. C., Canales, F. A., Fava, M. C., Brandão, A. R., & de Paes, R. P. (2023). Assessment of time series models for mean discharge modeling and forecasting in a sub-basin of the Paranaíba River, Brazil. *Hydrology*, 10(11), 208.
- Dazzi, S., Vacondio, R., & Mignosa, P. (2021). Flood stage forecasting using machine-learning methods: A case study on the Parma River (Italy). *Water*, 13(12), 1612.

- de Lavenne, A., Andréassian, V., Crochemore, L., Lindström, G., & Arheimer, B. (2022). Quantifying multi-year hydrological memory with catchment forgetting curves. *Hydrology and Earth System Sciences*, 26(10), 2715–2732. doi:<https://doi.org/10.5194/hess-26-2715-2022>
- Dettinger, M. D., & Diaz, H. F. (2000). Global Characteristics of Stream Flow Seasonality and Variability. *Journal of Hydrometeorology*, 1(4), 289–310.
- Dobias, J., & Salvinelli, C. (2018). Linear regression analysis and statistical models using oceanic and atmospheric variables for streamflow forecasting in Chixoy River, Guatemala. *World Environmental and Water Resources Congress 2018*, (págs. 280-286).
- Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press. Obtenido de <http://www.deeplearningbook.org>
- Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* (2nd ed.). New York, NY: Springer. doi:<https://doi.org/10.1007/978-0-387-84858-7>
- Helsel, D. R., Hirsch, R. M., Ryberg, K. R., Archfield, S. A., & Gilroy, E. J. (2020). *Statistical methods in water resources* (2nd ed.). U.S. Geological Survey. doi:<https://doi.org/10.5066/P9JWL6XR>
- Huang, B., Thorne, P., Banzon, V., Boyer, T., G., C., Lawrimore, J., . . . Vose, R. (2017). Extended Reconstructed Sea Surface Temperature, Version 5 (ERSSTv5): Upgrades, Validations, and Intercomparisons. *Journal of Climate*, 8179–8205. doi:<https://doi.org/10.1175/JCLI-D-16-0836.1>
- Hyndman, R. J., & Athanasopoulos, G. (2021). *Forecasting: principles and practice* (3rd ed.). Melbourne, Australia: OTexts. Recuperado el 15 de 10 de 2025
- ICC (Instituto Privado de Investigación sobre Cambio Climático) . (2017). *Manual de medición de caudales*. Guatemala: ICC.
- INSIVUMEH. (2018). *Variabilidad y Cambio Climático en Guatemala*. Guatemala: INSIVUMEH.
- Instituto Privado de Investigación sobre Cambio Climático (ICC). (2015). *Estudio hidrológico de la cuenca del río Madre Vieja*. Guatemala.

- IPCC. (2021). *Climate change 2021: The physical science basis. Contribution of Working Group I to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press. doi:<https://doi.org/10.1017/9781009157896>
- James, G., Witten, D., Hastie, T., & Tibshirani, R. (2021). *An Introduction to Statistical Learning: With Applications in R* (2nd ed.). New York, NY: Springer. doi:<https://doi.org/10.1007/978-1-0716-1418-1>
- Kennedy, E. (1990). *Techniques of Water-Resources Investigations of the USGS. Chapter A 10, DISCHARGE RATINGS AT GAGING STATIONS*. USGS.
- Kim, T., Shin, J. Y., Kim, H., Kim, S., & Heo, J. H. (2019). The use of large-scale climate indices in monthly reservoir inflow forecasting and its application on time series and artificial intelligence models. *Water*, 11(2), 374. doi:<https://doi.org/10.3390/w11020374>
- Kluyver, T., Ragan-Kelley, B., Pérez, F., Granger, B., Bussonnier, M., Frederic, J., . . . team, J. d. (2016). Jupyter Notebooks—a publishing format for reproducible computational workflows. En F. Loizides, & B. . Schmidt, *Positioning and power in academic publishing: Players, agents and agendas* (págs. 87–90). IOS Press. doi:<https://doi.org/10.3233/978-1-61499-649-1-87>
- Krause, P., Boyle, D., & Bäse, F. (2005). Comparison of different efficiency criteria for hydrological model assessment. *Advances in Geosciences*, 89–97.
- Ljung, G. M., & Box, G. E. (1978). On a measure of lack of fit in time series models. *Biometrika*, 65(2), 297–303. doi:<https://doi.org/10.1093/biomet/65.2.297>
- Machiwal, D., & Jha, M. (2012). *Hydrologic Time Series Analysis: Theory and Practice*. New Delhi: Springer.
- MAGA (Ministerio de Agricultura, Ganadería y Alimentación). (2004). *MAPA DE CUENCAS HIDROGRAFICAS, REPUBLICA DE GUATEMALA*. Guatemala: MAGA.
- Magaña, V., Amador, J. A., & Medina, S. (1999). The midsummer drought over Mexico and Central America. *Journal of Climate*, 12(6), 1577–1588.
- Maldonado, T., Alfaro, E. J., & Hidalgo, H. G. (2018). A review of the main drivers and variability of Central America's Climate and seasonal forecast systems. *Rev. Biol. Trop*, 66(1), S153-S175,.

- Montanari, A., & Koutsoyiannis, D. (2014). Modeling and mitigating natural hazards: Stationarity is immortal! *Water Resources Research*, *50*(12), 9748–9756.
doi:<https://doi.org/10.1002/2014WR016092>
- Montanari, A., Rosso, R., & Taqqu, M. S. (1997). Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation, and simulation. *Water Resources Research*, *35*(5), 1035–1044. doi:<https://doi.org/10.1029/97WR00043>
- Montanari, A., Young, G., Savenije, H. H., Hughes, D., Wagener, T., Ren, L., . . . Belyaev, V. (2013). Panta Rhei—Everything flows: Change in hydrology and society—The IAHS scientific decade 2013–2022. *Hydrological Sciences Journal*, *58*(6), 1256–1275.
doi:<https://doi.org/10.1080/02626667.2013.809088>
- Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D., & Veith, T. L. (2007). Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *Transactions of the ASABE*, *50*(3), 885–900.
doi:<https://doi.org/10.13031/2013.23153>
- National Oceanic and Atmospheric Administration (NOAA). (10 de 12 de 2025). *Niño 3.4 index*. NOAA Climate Prediction Center & Physical Sciences Laboratory. Obtenido de <https://psl.noaa.gov/data/climateindices/list/>
- OCHA. (2018). *Guatemala, daños por el fenómeno del El Niño. Años 2015/16 y 2018*. United Nations Office for the Coordination of Humanitarian Affairs.
- Perazzolo, D., Lazzaro, G., Fiume, A., Fanton, P., & Grisan, E. (2025). Streamflow forecasting: A comparative analysis of ARIMAX, rolling forecasting LSTM neural network and physically based models in a pristine catchment. *Water*, *17*(15), 2341.
doi:<https://doi.org/10.3390/w17152341>
- Rantz et al. (1982). *Measurement and Computation of Streamflow: Volume 2. Computation of Discharge*. Washington: USGS.
- Shmueli, G. (2010). To Explain or to Predict? *Statistical Science*, *25*(3), 289-310.
doi:<https://doi.org/10.1214/10-STS330>
- Strahler, A. N. (1952). Hypsometric (area-altitude) analysis of erosional topography. *Geological Society of America Bulletin*, *63*(11), 1117–1142. doi:[https://doi.org/10.1130/0016-7606\(1952\)63%5B1117:HAAOET%5D2.0.CO;2](https://doi.org/10.1130/0016-7606(1952)63%5B1117:HAAOET%5D2.0.CO;2)

- Taylor, M. A., & Alfaro, E. J. (2005). Climate of Central America and the Caribbean. (Springer, Ed.) *Encyclopedia of World Climatology*.
- Ward, P. J., Jongman, B., Sperna Weiland, F. C., Bouwman, A., van Beek, R., Bierkens, M. F., & Winsemius, H. C. (2014). Strong influence of El Niño–Southern Oscillation on flood risk around the world. *Proceedings of the National Academy of Sciences*, 111(44), 15659–15664. doi:<https://doi.org/10.1073/pnas.1409822111>
- World Meteorological Organization. (2010a). *MANUAL ON STREAM GAUGING, VOLUME I – Field work*. Geneva: WMO.
- World Meteorological Organization. (2010b). *MANUAL ON STREAM GAUGING, VOLUME II – Computation of discharge*. Geneva: WMO.
- World Meteorological Organization. (2021). *2021 State of Climate Services: Water (WMO-No. 1278)*. Geneva: WMO.
- World Meteorological Organization. (2022). *Early warnings for all: Executive action plan 2023–2027*. WMO.
- Zhao, X., Wang, H., Bai, M., Xu, Y., Dong, S., Rao, H., & Ming, W. A. (2024). Comprehensive Review of Methods for Hydrological Forecasting Based on Deep Learning. (MDPI, Ed.) *Water*, 16, 1407-1439. doi:<https://doi.org/10.3390/w16101407>



www.icc.org.gt