

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA
Facultad de Ingeniería, Arquitectura y Diseño



Maestría y Doctorado en Ciencias e Ingeniería



**“MODELACIÓN ESTOCÁSTICA PARA EL ESTUDIO DE LA
VARIABILIDAD CLIMÁTICA EN LA CUENCA HIDROGRÁFICA
GUADALUPE B.C.”**

TESIS

que para cubrir parcialmente los requisitos necesarios para obtener el
grado de

MAESTRO EN INGENIERÍA

Presenta

SABTHA PAMELA CRESPO DÍAZ

Director de tesis

Dr. ALVARO ALBERTO LÓPEZ LAMBRAÑO

Ensenada, Baja California, junio del 2018.

UNIVERSIDAD AUTÓNOMA DE BAJA CALIFORNIA

FACULTAD DE INGENIERÍA
UNIDAD ENSENADA

Modelación estocástica para el estudio de la variabilidad climática en la cuenca hidrográfica Guadalupe B.C.

TESIS

Que para obtener el grado de maestría en ingeniería presenta:

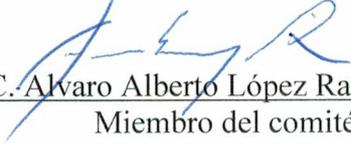
Sabtha Pamela Crespo Díaz

Aprobada por:


Dr. Alvaro Alberto López Lambrano
Director de tesis


Dr. Jorge Octavio Mata Ramirez
Miembro del comité


Dr. Miguel Enrique Martínez Rosas
Miembro del comité


M.C. Alvaro Alberto López Ramos
Miembro del comité

Ensenada Baja California, México. Junio 2018

ÍNDICE	
Índice de figuras	4
Índice de tablas	6
1 INTRODUCCIÓN	7
1.1 Estado del arte	7
1.1.1 Precipitación	12
1.1.2 Temperatura	28
1.2 Justificación.....	35
1.3 Descripción del Problema	36
1.4 Objetivos	38
1.4.1 Objetivo general	38
1.4.2 Objetivos específicos.....	38
1.5 Hipótesis	38
2 FUNDAMENTOS DE HIDROLOGÍA ESTOCÁSTICA.....	39
2.1 Introducción.....	39
2.2 Medidas de Tendencia Central	39
2.3 Medidas de Dispersión.....	40
2.4 Métodos de estimación de datos faltantes	42
2.4.1 Correlación	42
2.4.2 Método del U.S. Weather Service	43
2.4.3 Método de relación normalizada.....	43
2.4.4 Método de regresión lineal simple	44
2.5 Método de interpolación Distancia Inversa Ponderada (IDW).....	44
2.6 Regresión Lineal	46
2.7 Función de autocorrelación (ACF) y Función de autocorrelación parcial (PACF)	49
2.8 Matriz de correlación.....	51
2.9 Series de tiempo	52
2.9.1 Componentes de las series de tiempo	52
2.9.2 Test para análisis de consistencia de datos	53
2.10 Modelos autorregresivos.....	63

2.10.1	Modelo autorregresivo AR (p)	63
2.10.2	Modelo de media móvil MA (q)	64
2.10.3	Modelo autorregresivo de media móvil ARMA (p, q)	65
2.10.4	Modelo autorregresivo integrado de media móvil ARIMA (p, d, q)	66
2.11	Criterio de Información de Akaike (AIC).....	68
2.12	Principio de Parsimonia	71
2.12.1	Teoría del principio de Parsimonia	72
2.13	Estadístico Durbin-Watson (DW)	72
2.14	Transformación Box-Cox	73
3	METODOLOGÍA	75
3.1	Características del área de estudio.....	75
3.2	Recopilación de información hidrometeorológica.....	77
3.3	Conformación de Series de Tiempo.....	78
3.4	Estimación de Datos Faltantes.....	78
3.5	Espacialización de variables	81
3.6	Análisis estadístico.....	82
3.6.1	Momentos estadísticos	82
3.6.2	Gráficos de series de tiempo	83
3.6.3	Obtención de gradientes de series de tiempo	83
3.7	Autocorrelación y Autocorrelación Parcial.....	83
3.7.1	Autocorrelograma	84
3.8	Análisis de Consistencia	84
3.9	Proceso de normalización	84
3.10	Identificación y selección de modelos Autorregresivos.....	85
3.11	Espacialización de componentes de modelos Autorregresivos	85
3.12	Algebra de Mapas.....	85
4	RESULTADOS Y ANÁLISIS DE RESULTADOS	87
4.1	Condiciones climatológicas y fisiográficas de la cuenca Guadalupe, B.C.....	87
4.2	Análisis estadístico de series de precipitación, temperatura y evaporación.....	92
4.2.1	Momentos estadísticos y series de tiempo.....	92

4.2.2	Análisis de gradientes de variables climáticas en estudio	96
4.2.3	Análisis de Matriz de correlación de gradientes, variables climáticas y características fisiográficas de la cuenca Guadalupe	104
4.3	Análisis de función de autocorrelación ACF y función de autocorrelación parcial PACF	105
4.4	Resultados de análisis de consistencia.....	108
4.5	Normalización de series de tiempo	112
4.6	Análisis de la prueba Dickey-Fuller aumentada ADF	114
4.7	Identificación y selección de modelos para cuenca Guadalupe	115
4.7.1	Modelos Autorregresivos de la cuenca Guadalupe	121
4.8	Estimación de parámetros del modelo	122
4.9	Diagnóstico del modelo	123
4.10	Espacialización de componentes de modelos autorregresivos y de media móvil	125
5	CONCLUSIONES	126
6	REFERENCIAS.....	127

Índice de figuras

Figura 1.- Vecindad IDW del punto seleccionado. Fuente: (ArcGis, 2018)	45
Figura 2. Representación de una matriz de correlación de tres variables.....	51
Figura 3.- Esquema de algoritmo para Modelo Autorregresivo	68
Figura 4. Etapas de metodología empleada en la investigación	75
Figura 5. Localización del área de estudio	76
Figura 6. Interfaz del software SIHIMAX v1.56 de Hydrobits	80
Figura 7. Ventana para la estimación de datos faltantes.....	81
Figura 8. Ejemplo de capas de información tipo ráster para generar álgebra de mapas	86
Figura 9. Geoespacialización de las estaciones climatológicas en cuenca Guadalupe	88
Figura 10. Geoespacialización del promedio anual de la precipitación total mensual en cuenca Guadalupe	89
Figura 11. Geoespacialización de la temperatura mensual media anual en cuenca Guadalupe.....	90
Figura 12. Geoespacialización del promedio anual de evaporación total mensual en cuenca Guadalupe	91
Figura 13. Serie de tiempo para datos de precipitación total mensual para la estación Aguacaliente (2001)	96
Figura 14. Ajuste para obtención de gradientes de precipitación total mensual de la ...	98
Figura 15. Ajuste para obtención de gradientes de temperatura mensual media de la ...	98
Figura 16. Ajuste para obtención de gradientes de evaporación total mensual de la ...	99
Figura 17. Mapa de gradientes de precipitación total mensual ΔP de cuenca Guadalupe.....	101
Figura 18. Mapa de gradientes de temperatura mensual media ΔT de cuenca Guadalupe.....	102
Figura 19. Mapa de gradientes de evaporación total mensual ΔE de cuenca Guadalupe	103
Figura 20. Función de autocorrelación ACF de estación 2001 para serie de precipitación	107
Figura 21. Función de autocorrelación parcial PACF de estación 2001 para serie de	107
Figura 22. Gráfico de datos de precipitación mensual total de estación 2001, transformados mediante Box-Cox primera potencia	113
Figura 23. Gráfico de datos de precipitación mensual total de estación 2001, transformados mediante Box-Cox segunda potencia.....	113
Figura 24. Representación gráfica de modelos propuestos con base al valor mínimo de AIC de estación 2001 para serie de precipitación total mensual.....	120
Figura 25. Representación gráfica de residuales obtenidos con el modelo ARMA (2,5) para datos de precipitación total mensual en la estación 2001	123

Figura 26. Función de autocorrelación ACF para residuales de precipitación total mensual, modelo ARMA (2,5) para la estación 2001	125
Figura 27. Función de autocorrelación parcial PACF para residuales de precipitación total mensual, modelo ARMA (2,5) para la estación 2001	125

Índice de tablas

Tabla 1 Clasificación del coeficiente de correlación	52
Tabla 2 Estaciones climatológicas en el área de estudio y promedios anuales de variables de interés.	77
Tabla 3 Porcentaje de datos faltantes para series de precipitación, temperatura y evaporación.....	79
Tabla 4 Estadísticos de precipitación total mensual observada.	93
Tabla 5 Estadísticos de temperatura media mensual observada.....	94
Tabla 6 Estadísticos de evaporación total mensual observada.....	95
Tabla 7 Estadísticos de precipitación diaria observada.....	96
Tabla 8 Gradientes de series de tiempo.....	97
Tabla 9 Matriz de Correlación de gradientes de las variables climáticas en estudio, promedios anuales de las variables y características fisiográficas	104
Tabla 10 Función de autocorrelación ACF y autocorrelación parcial PACF de estación 2001 para serie de precipitación total mensual observada	106
Tabla 11 Análisis de consistencia para series de precipitación total mensual	108
Tabla 12 Análisis de consistencia para series de temperatura media mensual	110
Tabla 13 Análisis de consistencia para series de evaporación total mensual	111
Tabla 14 Estadísticos de serie de precipitación total mensual transformados mediante Box-Cox segunda potencia	114
Tabla 15 Prueba Dickey-Fuller aumentada para serie de precipitación total mensual transformada de la estación 2001	115
Tabla 16 Criterio de selección de modelo para serie de precipitación total mensual de estación 2001	116
Tabla 17 Modelos de series de tiempo seleccionados	121
Tabla 18 Parámetros de modelo seleccionado para serie de precipitación total mensual en estación 2001	122
Tabla 19 Función de autocorrelación ACF y autocorrelación parcial PACF de residuales de precipitación total mensual en estación 2001.....	124

1 INTRODUCCIÓN

En este trabajo se presentan los modelos generados mediante el análisis de series de tiempo, en donde las variables analizadas serán explicadas utilizando su propio pasado, es decir, en su propia evolución histórica. En la primera sección, se hace una breve revisión literaria sobre la evolución de las series de tiempo y los resultados obtenidos mediante la aplicación de las mismas en distintas investigaciones de otros países. Es importante destacar que el enfoque de los modelos autorregresivos y de media móvil se ha extendido a distintas áreas de la ciencia y de acuerdo a las necesidades de la investigación.

En la segunda sección, se presentan los fundamentos básicos de la hidrología estocástica, así como las pruebas necesarias para los modelos planteados. Además, se exhibe la problemática en el área de estudio, los objetivos y la hipótesis de investigación. En la siguiente sección, se describe la metodología empleada, misma que definirá los modelos de mayor ajuste. Una vez identificados y seleccionados los modelos, en la cuarta sección se presentan los resultados obtenidos, gráficos, tablas, mapas en base a los modelos propuestos, así como las técnicas de diagnóstico.

Por último, en el capítulo 5 se exhiben las conclusiones enfocadas no solo a los resultados obtenidos, sino también a algunas recomendaciones para mejorar la calidad en los resultados de los modelos.

1.1 Estado del arte

La precipitación, temperatura y evaporación son variables climáticas fundamentales para la caracterización del clima, por lo tanto, es importante contar con herramientas que permitan conocer el comportamiento de éstas a escala regional. Considerando su variabilidad en el tiempo y en el espacio, en especial cuando el objetivo es analizar cambios en el clima, es común estudiarlas y aplicar distintos métodos y modelos matemáticos existentes que permitan conocer y predecir su comportamiento. A través del tiempo se han desarrollado numerosas metodologías que facilitan la descripción de dichas variables a escala regional, mediante las cuales es posible generar registros sintéticos en zonas donde no se dispone de estos. Sin embargo, estas metodologías tienen sus limitaciones ya que suelen ser puramente estadísticas (ej. krigeado) o solo contemplan el dominio de las variables geográficas por la proximidad entre estaciones. Debido a lo anterior, surgió la necesidad de desarrollar métodos que tomen en cuenta variables que contemplen además del sentido geográfico, el sentido climático y a su vez la “proximidad entre puntos”, siendo capaz de mejorar los resultados obtenidos mediante las técnicas clásicas de interpolación espacial que existen actualmente (Johnson & Hanson, 1995; Moreno & Roldán, 1999; Teegavarapu & Chandramouli, 2005).

Actualmente los modelos experimentales utilizados son modelos estocásticos (Fortin, Bernier, & Bobée, 1997) basados en la teoría de la probabilidad para llevar a cabo la representación de la incertidumbre; dichos modelos se han utilizado desde el año de 1926; Yule (1926, 1927) fue el primero en describir los modelos AR y posteriormente Slutsky (1937) propuso los modelos MA; siendo Wold (1954) quien años después explica los modelos mixtos ARMA.

El desarrollo y la aplicación de esta variedad de modelos estocásticos continuó a través de Box y Jenkins (1976) y a través del tiempo estos han sido ampliamente utilizados en diversas áreas, una de ellas es la hidrología donde han sido útiles para pronosticar eventos hidrológicos, la cantidad de lluvia caída en una región y realizar análisis comparativos de modelos hidrológicos con el objetivo de generar simulaciones continuas de cuencas, además de generar registros sintéticos, etc.. Estos modelos abarcan los modelos lineales autorregresivos (AR), los modelos lineales de promedios móviles (MA), los modelos lineales autorregresivos de promedios móviles (ARMA), los modelos periódicos ARMA (PARMA), los modelos ARMA con entradas auxiliares (ARMAX) y los modelos autorregresivos integrados de promedios móviles (ARIMA), que permiten generar series sintéticas temporales.

Considerando lo anterior, el uso de modelos estocásticos del tipo autorregresivos ARMA (p,q) también conocidos como metodología Box & Jenkins son una de las herramientas más utilizadas en la modelación hidrológica, la cual es la base para la gestión integral del recurso hídrico así como en la mayoría de los procesos del ciclo hidrológico, y que además se puede encontrar en cualquier campo de la ciencia para el análisis y pronósticos de series de tiempo permitiendo la utilización de datos históricos para predecir datos futuros sin la necesidad de buscar otros datos relacionados con las series de tiempos (Momani, 2009).

La aplicación de los modelos autorregresivos de medias móviles (ARMA) se ha realizado principalmente a procesos hidrológicos estacionales univariados (Haltiner & Salas, 1988; Obeysekera & Salas, 1986; Padilla, Pulido-Bosch, & Calvache, 1996; Salas, 1980; Salas, Boes, & Smith, 1982; Ula, 1990), los cuales se basan en el análisis de la estructura estocástica de una serie temporal con el objetivo de generar pronósticos a corto plazo, completar datos faltantes y crear series sintéticas que preserven la correlación temporal de la serie original así como los momentos estadísticos de la misma.

No obstante, la aplicación de estos modelos también puede ser multivariada, como en Colombia donde Carvajal (2005), realizó teleconexiones lineales entre el fenómeno ENSO y los caudales mensuales del Río Cauca mediante dos técnicas estadísticas multivariadas: análisis de funciones ortogonales empíricas (FOES) y el análisis de correlación canónica (ACC), prediciendo los caudales a través del ajuste de modelos

ARMAX, tomando las primeras componentes y los coeficientes canónicos de las variables macroclimáticas, como variables auxiliares en los modelos de caudal obteniéndose mejoras en la predicción. Así mismo Domínguez (2001) realizó comparaciones con modelos autorregresivos aplicando la ecuación de Fokker-Planck-Kolmogorov con el fin de pronosticar las afluencias a embalses hidroeléctricos; obteniendo un modelo para pronosticar las curvas de densidad de probabilidad de los caudales de la represa de Betania en Colombia.

En lo que se refiere a la cuantificación de la recarga natural de acuíferos, Paris (2002) presenta una metodología mediante el análisis estocástico de las series temporales de precipitación y temperatura (Santa Fe, Argentina); cuya interpretación se realizó a través de gráficos de evolución temporal y comportamiento temporal medio, periodograma y periodograma integrado, así como de las funciones de autocorrelación y autocorrelación parcial. De esta forma, fue posible evaluar la homogeneidad, caracterizar la estacionariedad e identificar la estructura autorregresiva y estacional de las series, además de ajustar los parámetros de modelos ARIMA (Box et al., 1976) con el propósito de obtener el pronóstico de dichas variables en un periodo mensual. En cuanto a modelos ARIMA estacionales y con propósitos ilustrativos Etuk (2013) confirmó que estos son adecuados ya que se ajustan a las series de precipitación mensual debido a que la lluvia es un fenómeno estacional en todo el mundo.

Desde el punto de vista de cambio climático Piwowar (2002) utilizó modelos ARMA con el objetivo de estudiar el cambio climático a través del modelado de series de tiempo correspondientes a imágenes de sensores remotos y determinar los cambios y el impacto de las concentraciones de hielo en el océano Ártico; por esta razón, el análisis de sensibilidad remota se convirtió en una herramienta muy importante en esta área, por lo tanto el desarrollo de métodos estocásticos de datos climáticos se ha realizado a distintas escalas de tiempo ya sea anual, mensual y diaria (R. Srikanthan & McMahon, 2001). Con relación a la generación de caudales diarios, semanales, mensuales y anuales, así como a la estimación de parámetros de modelos para la predicción de caudales e inundaciones en tiempo real los modelos autorregresivos han mostrado resultados satisfactorios utilizando el método de mínimos cuadrados ponderados (Chao, Hua-sheng, Wei-min, & Luo-ping, 2008; Mohammadi, Eslami, & Kahawita, 2006); también han sido empleados en el estudio y el modelado de la escorrentía superficial proponiendo modelos conceptuales (Spolia & Chander, 1974; Spolia, 1972).

Ahora considerando la aplicación de herramientas informáticas, recientemente las redes neuronales artificiales (ANN) han sido ampliamente utilizadas en la predicción de variables hidroclimatológicas buscando alternativas diferentes para el aprovechamiento y conservación de recursos hídricos (Obregón, Fragala, & Prada, 2003), desde esta perspectiva también han sido aplicadas en la modelación hidrológica obteniendo resultados satisfactorios (Abrahart, See, & Kneale, 1999; Campolo, Andreussi, &

Soldati, 1999; Cheng & Noguchi, 1996; Daniell, 1991; Christian W. Dawson & Wilby, 1998; French, Krajewski, & Cuykendall, 1992; K. Kang, Oh, Kwon, & Park, 1993; Karunanithi, Grenney, Whitley, & Bovee, 1994; Lorrai & Sechi, 1995; Minns & Hall, 1997, 1996; S. & Minns, 1998; Smith & Eli, 1995; Yang, 1997).

De igual forma las ANN también han sido utilizadas para la estimación de gastos (Abrahart & See, 2000), donde las predicciones se han realizado basándose en las series de datos del caudal de cuencas mediante diferentes técnicas: modelos autorregresivos de promedios móviles (ARMA) y redes neuronales (NN), en donde las modelaciones han resultado similares; así mismo se han realizado predicciones de inundaciones repentinas en ríos (Sahoo, Ray, & De Carlo, 2006); para el desarrollo de series sintéticas de gastos mediante registros cortos (Jia & Culver, 2006); pronóstico de escurrimientos diarios mediante la aplicación de modelos híbridos de ANN (W. Wang, Gelder, Vrijling, & Ma, 2006); estimación de inundaciones en cuencas no canalizadas mediante modelos de regresión múltiple y ANN donde los resultados muestran que el rendimiento del modelo es influenciado por los factores geográficos (C. W. Dawson, Abrahart, Shamseldin, & Wilby, 2006); también han sido empleadas NN predictivas y algoritmos de aprendizaje de propagación flexible para generar pronósticos específicos de precipitación diaria en distintas localidades de Sao Paulo (Valverde Ramírez, De Campos Velho, & Ferreira, 2005) y, en México han sido aplicadas para analizar el tránsito de avenidas en cauces con el fin obtener resultados similares a los métodos hidráulicos e hidrológicos (Molina-Aguilar & Aparicio, 2006); para el pronóstico de sequías hidrológicas en el Organismo de Cuenca del Pacífico Norte considerando distintos indicadores de sequía (Douriet, Sanz, Ramirez, & Ravelo, 2015); para pronosticar alturas de precipitación en intervalos sucesivos de 15 minutos en una estación pluviométrica de una zona urbana, en donde agruparon los eventos en conjuntos mediante el método de *k-Nearest Neighbors* (k-NN) para mejorar el pronóstico e incluyeron el concepto de tendencia esperada que permite predecir si la altura de la precipitación aumentará o disminuirá (Cisneros-Iturbe & Pelczer, 2007).

Por otro lado, el empleo de métodos no paramétricos como *Nearest-Neighbors* ha sido útil para realizar comparaciones entre los modelos autorregresivos ARMA (p,q) y los procedimientos de ANN (Ochoa-Rivera, 2008; Toth, Brath, & Montanari, 2000). Por esta razón, es importante destacar que los modelos ARMA son un parámetro de referencia de los resultados que obtenidos mediante otras metodologías (Abrahart & See, 2000; Tingsanchali & Gautam, 2000).

En términos generales, los modelos ARMA (p,q) han permitido el estudio de variables hidroclimatológicas a través de las funciones de autocorrelación de las mismas con el fin de describir los distintos fenómenos dentro del ciclo hidrológico. No obstante, hoy en día existe una amplia variedad de modelos que se derivan del procedimiento tradicional y que considera diversas adecuaciones. En este sentido, se han creado esquemas

ARMA-GARCH/EGARCH/SR-SARV(2) teniendo como base series con valores diarios y en donde a través de modelos empíricos es posible evaluar las variaciones lineales que dependen de 2 factores autorregresivos (Karanasos, Kim, Journal, & Kim, 2003; Meddahi, 2003). En lo relacionado a la gestión de los recursos hídricos, el análisis de incertidumbre y de riesgos son áreas de vital importancia y el modelo de heteroscedasticidad condicional autorregresiva generalizada (GARCH) de procesos hidrológicos ha sido empleado para corregir residuos, y ha demostrado que en comparación con el modelo ARMA tradicional, el modelo GARCH tiene la capacidad de predecir intervalos de confianza más precisos bajo el mismo nivel de confianza (H. R. Wang, Gao, Qian, & Yu, 2012). Considerando que los enfoques tradicionales de series de tiempo se usan para modelar, simular y pronosticar la media condicional de las variables hidrológicas, pero descuidan su varianza en el tiempo, el enfoque de modelado de heteroscedasticidad condicional autorregresiva generalizada (MGARCH) muestra que la relación varianza-covarianza entre las variables hidrológicas varía en el tiempo y permite estimar la correlación condicional dinámica entre las variables para procesos de precipitación-escorrentía (Modarres & Ouarda, 2013). Es bien sabido que la predictibilidad de las variables meteorológicas no es constante, sin embargo, muestra variaciones regulares, en este caso los modelos GARCH permiten capturar esta característica ya que la varianza condicional de una observación depende linealmente de las varianzas condicionales de las observaciones previas y de los errores de predicción previos, la aplicación de este modelo junto con un modelo AR (2) demuestra que los modelos heteroscedásticos superan a los homoscedásticos para las temperaturas diarias y que funcionan mejor con la desviación estándar que con las varianzas (Tol, 1996).

En definitiva, algunos de los factores que han contribuido al desarrollo de métodos más eficientes son: la mayor disponibilidad de observaciones detectadas remotamente a través de radares y satélites y la necesidad de mejorar la parametrización de los procesos hidrológicos en los modelos climáticos (Foufoula-Georgiou & Krajewski, 1995). De este modo, la robustez y la aplicabilidad que muestran los modelos autorregresivos permite utilizarlos no solo en el área de hidrología, convirtiéndolos en una herramienta comúnmente utilizada a nivel global para la realización de pronósticos de diversas variables en otros campos de la ciencia. Vallverdú (1991) utilizó modelos NARMAX, MA y ARMAX con el fin de encontrar un modelo de control del Sistema Nervioso Central (CNS) del sistema cardiovascular humano basándose en ecuaciones diferenciales y en algoritmos ortogonales para diferentes grupos de pacientes; Valdes-Sosa (2004) utilizó un modelo de series temporales autorregresivas multivariadas (MAR) para explorar la conectividad funcional en Neuroimaging y analizar la causalidad de Granger de una región cerebral. Así mismo, McCall (2007) describe los movimientos de la respiración como una serie temporal no periódica superpuesta a una forma de onda periódica, considerando lo anterior implementó un nuevo enfoque al problema de

modelar y predecir el movimiento de la respiración, utilizando un algoritmo de promedio móvil autorregresivo periódico (ARMA) para definir un modelo matemático de los componentes periódicos y no periódicos de la respiración, demostrando que la precisión del modelo ARMA depende más de las variaciones en la duración del ciclo que en la amplitud de los ciclos respiratorios. Aguirre (2005) describe los modelos NARMAX como estructuras paramétricas tipo entrada-salida, mediante estos modelos investigó la aplicación de técnicas de identificación de sistemas dinámicos no lineales ya que han demostrado su capacidad para representarlos. Los modelos ARIMA-GARCH han sido ampliamente utilizados en aplicaciones financieras, Martínez (2014) lo aplicó para analizar el riesgo de mercado de fondos de inversión en México con el fin de examinar el comportamiento dinámico de su rendimiento y volatilidad; en este sentido la factibilidad de los modelos ARIMA para el análisis y la predicción permitió que fuera utilizado para evaluar la prevalencia y mortalidad de la parvovirus canina concluyendo que los modelos Box-Jenkins presentan una elevada precisión facilitándose la aplicación de estos en otros indicadores de salud animal (Aldaz Cárdenaz, García Díaz, & Quiñones Ramos, 2012). En lo que se refiere a estimaciones de mecanismos generadores del proceso psicológico Chorro (2000) estudió la posibilidad de aplicar modelos ateóricos como los modelos vectoriales de series temporales múltiples observando que el modelo explica un porcentaje importante de la variación de los datos. En el área de ciencias forenses los investigadores han comenzado a desarrollar técnicas forenses digitales para verificar la autenticidad de imágenes utilizando modelos autorregresivos unidimensionales (AR), mostrando que la técnica forense de filtrado propuesta logró mejoras en el rendimiento con respecto a métodos existentes (X. Kang, Stamm, Peng, & Liu, 2013).

1.1.1 Precipitación

El análisis del comportamiento de la precipitación es de suma importancia ya que los modelos matemáticos para representarla dependen de la escala de tiempo considerada, en otras palabras, cuando se habla de escalas de tiempo grandes ya sea anuales o mensuales, se puede asumir que no existe autocorrelación de la precipitación o que esta es despreciable (Roesner & Yevjevich, 1966). En este caso solo se requiere encontrar una distribución de probabilidad idónea para representar la variabilidad de la precipitación en la escala de tiempo considerada. Por el contrario, cuando se habla de escalas de tiempo menores, es decir semanas, días u horas, la autocorrelación de la precipitación ya no es despreciable, por lo que se requieren modelos más complejos que tomen en cuenta tanto la periodicidad anual, como la dependencia en el tiempo. Uno de estos modelos estocásticos, son las cadenas de Markov, las cuales han tenido suficiente aprobación para el modelado de la precipitación diaria debido al acierto en sus resultados (Katz, 1977b, 1977a). Una de

las características de estos modelos es que reproducen los resultados de ocurrencia y no ocurrencia de la precipitación, así como la cantidad o intensidad de lluvia, además mediante dichos modelos es posible describir la dependencia paso a paso también conocida como proceso de primer orden o correlación serial de un paso de exhibición (Viessman, 1989), dicho de otra manera, un evento no depende de otros procesos pasados. De este modo, los procesos de Markov están definidos en términos de probabilidades discretas y permiten generar series sintéticas de variables hidrológicas. Se destaca el trabajo realizado por Caskey (1963), donde las probabilidades teóricas derivadas de un modelo simple de cadena de Markov mostró ser adecuado para las precipitaciones diarias concordando estrechamente con los valores empíricos de la probabilidad de ocurrencia de la precipitación, sin embargo, es importante aclarar que el éxito de las cadenas de Markov para producir probabilidades teóricas que concuerden con las probabilidades observadas no debe interpretarse como que dichas distribuciones fueron explicadas. En lo que respecta a la descripción de la dependencia Subbaiah (2002) presentó un modelo para la precipitación de la temporada monzónica, basándose en la suposición de que la precipitación semanal en la temporada es un proceso Markoviano de primer donde, finalmente la comparación entre las series históricas y sintéticas mostró que ambas son estadísticamente comparables respecto a las medidas de tendencia central, dispersión y distribución. Dentro del esquema Cadenas de Markov Monte Carlo (MCMC) se ha presentado una nueva muestra denominada Metrópolis adaptativa de la evolución diferencial (DREAM), este enfoque proporciona nuevas formas para estimar la precipitación promedio de cuencas hidrográficas, información que es de suma importancia para diagnosticar errores estructurales en modelos y realizar comparaciones apropiadas de los dispositivos de medición de precipitaciones. De esta manera, los modelos MCMC permiten estimar y resumir la función de densidad de probabilidad posterior (de entrada, parámetro y error estructural) en el modelado hidrológico (Vrugt, ter Braak, Clark, Hyman, & Robinson, 2008).

Con relación a escalas de tiempo menores a 24 horas (lluvias horarias), también se han utilizado procesos Markovianos tipo estacionarios, generalmente los modelos basados en procesos puntuales (Cam, 1961; Kavvas & Delleur, 1975), los cuales modelan la precipitación en tiempo continuo, así como las que parten de una tormenta y dividen la precipitación a escalas de tiempo menores mediante el uso de hietograma (Huff, 1967). Sin embargo, hay que hacer notar que cuando las series temporales de lluvia se observan en una escala de días pueden ser tratadas como un proceso estocástico estacionario con la propiedad del ruido blanco, en este sentido, cuando se observan en una escala de horas, ya no es un proceso estocástico estacionario sino que pertenece a un proceso estocástico no estacionario lo que requiere modificaciones en la aplicación de la metodología y en el análisis de los datos hidrológicos (Hino & Hasebe, 1984).

Por otro lado, los modelos ARMA han sido utilizados para el estudio y desarrollo de modelos multivariados, permitiendo modelar series de tiempo de precipitaciones anuales, mensuales y diarias logrando el estudio del comportamiento de las variaciones de esta variable (Chapman, 1997; Sharma & Lall, 1999; R. Srikanthan & McMahon, 2001; R Srikanthan, 2004; R Srikanthan & others, 2005; Ratnasingham Srikanthan & MacMahon, 1985; Q. J. Wang, Nathan, & others, 2002); de igual forma los modelos clásicos multivariados se diseñaron para preservar la autocorrelación y la autocorrelación cruzada (Bras & Rodriguez-Iturbe, 1985; Chin, 1995; Salas, 1980). En relación a lo anterior Martínez y Salas (2004) realizaron un modelo para analizar y simular la precipitación horaria denominado modelo periódico discreto autorregresivo de orden 1 (PDAR), con el objetivo de generar las intensidades. En este caso la estimación de los parámetros se realizó mediante el método de momentos y para la estimación del modelo AR se hizo uso de dos métodos: el método 1 donde las medias y desviaciones típicas horarias de la variable transformada se estimaron para preservar las medias y las desviaciones típicas horarias obtenidas a través de la muestra histórica; y el método 2 en el que las medias y desviaciones típicas horarias de la variable se estimaron directamente de la muestra transformada analizando los datos de 4, 6, 12 y 24 horas demostrando que el modelo PDAR-AR es capaz de reproducir de manera adecuada las estadísticas históricas de la lluvia para las escalas de tiempo analizadas con el método 1.

1.1.1.1 Aplicación de Modelos Autorregresivos para predicción de la precipitación

Los métodos basados en la predicción de la precipitación generalmente consideran el análisis de series de tiempo debido a que sus pronósticos son bastante efectivos en comparación con otros métodos, teniendo en cuenta que el objetivo es modelar su comportamiento futuro de tal manera que permita llevar a cabo una correcta planeación y toma de decisiones. En esta sección se hace una revisión de investigaciones realizadas para las series de precipitación diaria, mensuales y/o anuales en distintos lugares del mundo, así como un análisis de los resultados obtenidos.

- *Precipitación diaria*

En lo que corresponde al área de Ingeniería, específicamente al diseño y operación de sistemas agrícolas, es necesario el estudio de inundaciones y sequías, así como el conocimiento de las características y secuencias de los días húmedos y secos. Éstas características pueden ser generadas a partir de series de precipitación diaria, sin embargo, debido a que la precipitación diaria no suele tener una distribución normal, se requiere de procesos más complejos mediante los cuales las series puedan ser

modeladas exitosamente, un ejemplo de esto es el proceso Binary Discrete Autoregressive Moving Average (B-DARMA) que incluye cadenas de Markov (Tiao J. Chang, M. L. Kavvas, 1984), el cual fue desarrollado con el fin de describir los periodos húmedos y secos en la ciudad de Cambridge en Indiana para la planeación óptima de actividades agrícolas en esta región geográfica. Una de las ventajas de este modelo es que su estructura de correlación puede especificarse independientemente de su densidad marginal y permanece invariante, esta propiedad es satisfecha por los datos de precipitaciones diarias, lo que facilita su modelado (Jacobs & Lewis, 1978). Considerando la flexibilidad de los modelos autorregresivos y de promedios móviles debido a que presentan un número menor de parámetros, su aplicabilidad resulta ser muy atractiva para modelar series temporales hidrológicas. Kumar (2004) desarrolla un modelo estocástico ARMA para la temporada de monzones del Suroeste de Bapantla, determinando el componente periódico en la serie temporal mediante análisis de Fourier y para evaluar la bondad de ajuste de los modelos hizo uso de la prueba de Box-Pierce Portmanteau y el Criterio de Información de Akaike. Con relación a las programaciones de riego, gestión agrícola y el funcionamiento de embalses en regiones donde se presentan lluvias monzónicas es posible aplicar modelos SARIMA a series de precipitación diaria que después se convierten en semanales y mensuales, como en Arunachal, Pradesh (India) en donde (Dabral & Murry, 2017) mediante las series de precipitación diaria aplican distintos modelos SARIMA $(0,1,1)(0,1,1)_{12}$, SARIMA $(2,1,5)(3,1,5)_{52}$ y SARIMA $(4,1,1)(0,1,1)_{153}$ para la predicción de 6 años de series de precipitación mensual, semanal y de días de monzón respectivamente. En dicho trabajo la selección del modelo de mejor ajuste se realizó con el criterio de información de Akaike y considerando el Schwarz Bayesian Criterion (SBC). Los valores pronosticados 14 años hacia adelante (2014-2027) obtenidos son utilizados para la planificación agrícola y de recursos hídricos, así como el estudio del cambio climático de dicha región.

Cuando se habla de precipitaciones a corto plazo (menores a 24 horas) se hace bajo el supuesto de que estas siguen un proceso de media móvil autorregresiva (ARMA), debido a que la estructura de algunos eventos puntuales como la precipitación horaria, son equivalentes a la estructura de autocovarianza de procesos de bajo orden, es decir, cuando el modelado está basado en modelos Poisson Rectangular Pulses (PRP), Neyman-Scott White Noise (NSWN) y Neyman-Scott Rectangular Pulses (NSRP), estos muestran que la estructura de correlación del proceso de precipitación media respecto a los intervalos de tiempo son equivalentes a la estructura de procesos ARMA en los cuales su espacio de parámetros están contenidos dentro del espacio de parámetros del modelo ARMA (Obeysekera, Tabios, & Salas, 1987); estas similitudes en la estructura de correlación de los modelos permiten que los modelos ARMA sean utilizados para la predicción de la precipitación.

Además, los pronósticos se pueden realizar con distintos enfoques ya sea de “estimación continua” o “estimación basada en eventos”, cada uno con distintas características, tal es el caso realizado en Denver y Morrison en Colorado y en la cuenca Arno en Italia central (Paolo Burlando, Renzo Rosso, 1993) donde utilizan un modelo ARMA (2,2) para pronosticar la precipitación por hora mediante enfoque de estimación continua, sin embargo, mediante enfoque continuo el modelo no fue capaz de realizar buenos pronósticos a futuro especialmente en verano, por otro lado, el enfoque de estimación basado en eventos demostró que puede reproducir pronósticos más precisos (2 horas adelante), esto se debe a la capacidad adaptativa del proceso de estimación, en consecuencia, el modelo ARMA es capaz de realizar mejores pronósticos *in situ* basándose en la precipitación promedio en la cuenca. Otro aspecto importante respecto al pronóstico de precipitaciones a corto plazo y que ha demostrado ser muy útil es en la estimación de las posibles zonas vulnerables y con riesgo a inundaciones, ya que contar con buenas predicciones de la entrada de lluvia permite ampliar el tiempo de avance (ej. un par de horas) de las previsiones de inundaciones, así como mejorar la estimación de una inundación para un plazo de pronóstico dado (Brath, Burlando, & Rosso, 1988) con el objetivo de realizar alertas tempranas a la población y reducir los daños, costos y pérdidas financieras que estas catástrofes naturales originan. En este sentido, los radares, satélites meteorológicos y Sistemas de Información Geográfica (SIG) son herramientas que adquieren mayor importancia ya que facilitan la generación de datos para realizar modelos hidrometeorológicos. Considerando las ventajas que aporta el uso de este tipo de herramientas para medir la precipitación y que esta puede ser pronosticada mediante procesos autorregresivos, la fusión de ambos permite una mayor eficiencia de los resultados, es decir, un modelo estocástico de tipo autorregresivo es más exitoso si los datos de entrada al modelo de precipitación se basan en mediciones de radar. En Italia del Norte el uso de los mapas de radar permitió la selección de aquellas estaciones donde las precipitaciones actuales y pasadas mostraron un mayor nivel de correlación con los datos observados en el sitio de pronóstico, basándose en el comportamiento cinemático de las tormentas con el objetivo de generar un Modelo Multivariate Autoregressive Integrated Moving Average (MARIMA-R) donde -R indica una estimación basada en radar del modelo y pronosticar una hora hacia adelante la variación horaria en la intensidad pluviométrica, así como la tendencia y transición de la precipitación creciente a decreciente y viceversa, convirtiéndose en una herramienta que puede ser utilizada para fines operacionales en cuencas urbanas y naturales (Burlando, Montanari, & Ranzi, 1996). También, se ha comparado el rendimiento de técnicas autorregresivas multivariadas (MAR) y los métodos autorregresivos (AR) univariados, Niedzielski (2007) aplicó dicha comparación a la modelación de precipitación-escorrentía a escala regional en la cuenca del río Odra en Polonia debido a la exposición de la región a eventos hidrológicos extremos (inundaciones a causa de la lluvia y derretimiento de nieve).

Ambos modelos se ajustaron a los datos obteniendo los pronósticos de flujos máximos con 2-4 días de anticipación tanto para picos de lluvia como para picos nevados, encontrando que cuando los datos de precipitación, descarga, capa de nieve y nivel de agua subterránea se analizan de forma combinada, la precisión de las predicciones de descarga aumenta. Además, demostró que los pronósticos basados en MAR solo son más precisos que las predicciones univariadas basadas en AR para un año cuando se produce una inundación, sin embargo, esta relación es inversa en el caso de un año sin inundaciones. Otro estudio similar llevado a cabo por Sulaiman (2014), en donde usando la precipitación acumulada máxima de 5 días, pudo predecir la magnitud de la precipitación en un periodo específico así como la precipitación extrema mensual, evaluando un método de ANN utilizando datos anteriores de MX5d e índices climáticos globales como el Índice de Oscilación del Sur (SOI), la Oscilación Madden Julian (MJO), y el Índice del Modo Dipole (DMI) en Malasia para posteriormente compararlos con un modelo ARIMA y con un análisis de regresión. La propuesta de Sulaiman fue utilizar entradas combinadas para investigar el efecto simultáneo de los valores rezagados de los datos de precipitación local y los índices climáticos globales sobre las precipitaciones estacionales extremas, encontrando que el modelo ARIMA es el mejor método para pronosticar la precipitación extrema cuando se usan como entrada los datos de MX5d.

Debido a que las series de lluvia se ven afectadas por una fuerte correlación espacial y una correlación de tiempo, se hace necesaria la generación y aplicación de series sintéticas diarias que conserven tanto la correlación en el tiempo (autocorrelación) como la correlación espacial (correlación cruzada), en estos casos hablamos en general de un modelo paramétrico lineal multivariable denominado VARMA(p,q) (Vector Autoregressive Moving Average) el cual permite expresar dichas correlaciones mediante combinaciones de parámetros lineales, tal es el caso de Grimaldi (2005), donde aplica dicho modelo para la simulación de escenarios de lluvias diarias en Tuscany Italia. Cuando se habla del estudio de fenómenos naturales como la precipitación, conocer sus magnitudes y frecuencias de ocurrencia es de gran importancia debido a las aplicaciones que se le pueden dar, tal es el caso de Hundecha (2009) el cual presenta un modelo autorregresivo multivariante para modelar la probabilidad local de ocurrencia de lluvia y la cantidad de lluvia diaria mediante una mezcla de dos funciones de distribución de probabilidad diferentes manteniendo la estructura de covarianza entre sitios utilizando una distribución normal truncada, así como la transformación integral de probabilidad en la cuenca Unstrut en el centro este de Alemania con el objetivo de caracterizar los extremos de la precipitación diaria tanto en estaciones individuales como en la escala de captación. Así también, Ruhf (2003) examinó datos de precipitación por hora a escala local, en este caso conocer la climatología de las precipitaciones diurnas fue un aspecto importante que permitió planificar actividades tales como el uso del suministro de agua o la gestión del control

del drenaje pluvial. Además, desarrolla un modelo ARIMA que describe la estructura de la cantidad, intensidad, duración e intervalo de la precipitación, encontrando que el cambio de cobertura y el uso de la tierra afectan la permeabilidad de la superficie, por lo tanto, se reduce la recarga de agua subterránea y se incrementa la evaporación.

Por otra parte, la variabilidad natural es un componente esencial de las observaciones de todas las variables geofísicas y climáticas, según Kim (2015) el análisis de función ortogonal empírica (EOF) cuenta con un conjunto de funciones ortogonales a partir de una función de covarianza espacial, estas funciones brindan información útil sobre los procesos físicos en los datos y sirven como una herramienta útil para desarrollar métodos estadísticos. Sin embargo, debido a que la suposición en EOF es la estacionalidad de los datos analizados (la función de covarianza no depende del origen del tiempo), ésta a menudo no es justificable para variables geofísicas y climáticas dando como resultado que las inferencias físicas y estadísticas basadas en EOF puedan ser engañosas. En ese sentido y debido a que algunas de estas variables muestran periódicamente estadísticas de covarianza dependientes del tiempo (los datos están correlacionados periódicamente o son cicloestacionarios). Un reconocimiento apropiado de las características de respuesta dependientes del tiempo es vital para extraer con precisión los modos físicamente significativos y sus evoluciones espacio-temporales a partir de los datos (evoluciones físicas consistentes y patrones de teleconexión en el análisis espectral de la variabilidad), ya que son objetivos importantes en muchos estudios climáticos y geofísicos, surgiendo estudios que consideran la base conceptual del análisis del EOF cicloestacionario (CSEOF) como una alternativa al análisis EOF regular u otras técnicas de análisis basadas en el supuesto de estacionariedad; tal es el caso en Corea del Sur donde Kim (2013) desarrolla un generador de clima para la precipitación en el verano utilizando la técnica CSEOF y aplica un modelo ARMA para generar cien conjuntos de series temporales de amplitud sintética, obteniendo ciclos estacionales en la precipitación sintética que se reproducen fielmente con picos bimodales típicos de la precipitación en verano y patrones de correlación espacial bastante similares a los de los datos observados.

Ahora, considerando que los datos de precipitación diarios de alta frecuencia son conocidos por su carácter intermitente en el tiempo (y el espacio), el modelado de este proceso se vuelve desafiante y por lo tanto, desde este punto de vista se ha presentado y aplicado una nueva metodología para el análisis de la precipitación diaria, Hannachi (2014) basa esta metodología en combinar teoría de procesos censurados con modelos autorregresivos continuos (de memoria corta) para modelar fenómenos intermitentes ya que las precipitaciones diarias tienen una ley de escalado compatible con los procesos de memoria corta. El procedimiento consiste en incorporar la teoría de los procesos censurados truncando el proceso AR (1) y analizar la serie truncada en

términos de un proceso censurado logrando que los parámetros del modelo capturen la estacionalidad de la precipitación en Irlanda del Norte.

- *Precipitación mensual*

El deseo de obtener predicciones en tiempo real de la precipitación para lograr una correcta gestión de los recursos hídricos ha sido uno de los objetivos de muchos investigadores en el área de la hidrología. Dentro de la amplia gama de modelos estocásticos aplicados a investigaciones de series de precipitación mensual se encuentra el modelo ARIMA, aplicado en 15 cuencas localizadas en Indiana, Illinois y Kentucky (Delleur & Kavvas, 1978), en este caso se hizo uso de 4 diferentes modelos ARIMA (0,0,0), (1,0,1), (1,1,1) y el modelo multiplicativo (1,0,1)X(1,1,1) para realizar pronósticos a 2 años. Se encontró que el modelo ARIMA (1,1,1) no es adecuado para las series de precipitación mensual, por el contrario del modelo ARIMA (1,0,1) generó un pronóstico más adecuado que se ajusta en gran manera a los datos históricos de las series de precipitación mensual. En lo que respecta al modelo multiplicativo se obtuvieron buenos resultados, sin embargo, la desventaja es que no permite la generación de series sintéticas debido a la pérdida de la desviación estándar mensual; como resultado estos modelos indican que son más útiles en pronósticos a corto plazo (1 o 2 años adelante) particularmente en predicción de la escorrentía en tiempo real. En tal sentido es fundamental contar con modelos adecuados para pronósticos de precipitación, ya que esto permite que se puedan desarrollar mejores modelos de pronósticos de escorrentía. Se destaca que, debido al aumento en la demanda de energía, a la creciente preocupación por el medio ambiente y el rápido agotamiento de las reservas de combustibles fósiles es necesario que los planificadores busquen maneras de complementar la base energética con fuentes de energía renovables. En este sentido Babu (2011) emplea un enfoque ARMA para modelar y simular los datos del flujo de lluvia en Vellore ya que este tipo de modelos representan un sistema de elementos que se mueven de un estado a otro a lo largo del tiempo, además permite la predicción de los parámetros del flujo de lluvia los cuales son importantes para diseñar estructuras sensibles al viento y para estudiar la contaminación del aire.

Una serie de tiempo está compuesta generalmente de tendencia, periodicidad y componentes estocásticos, una de las formas de realizar el modelado de series mensuales de precipitación es mediante su descomposición, es decir, obtener sus componentes periódicos y estocásticos, de esta manera, el modelado proporciona una metodología mejorada para pronósticos cortos e intermedios. Utilizando dicha metodología Bhakar, et al.(2006) generaron un modelo AR(2) para pronosticar la precipitación mensual futura y estimar los requerimientos de agua de cultivo en la región de Kota al norte de la India, obteniendo un modelo con un coeficiente de

correlación de $r=0.9961$ y un coeficiente de determinación $R^2=0.9001$. Dabral P.P., et al. (2008) desarrollaron un modelo estocástico mensual en la región húmeda de Doimukh, Arunachal Pradesh en la India para predecir 2 años futuros de precipitación con el objetivo de mejorar las actividades de manejo de cuencas así como la predicción de eventos futuros, en este caso el modelo AR(6) fue considerado el modelo más adecuado y con alto grado de ajuste de acuerdo a los datos observados debido a sus valores mínimos en los estadísticos AICC/BIC/FPE. En relación a la planificación y desarrollo de proyectos y a la simulación en el diseño de los sistemas de recursos hídricos de cuencas, Mallikarjuna (2002) elige los modelos autorregresivos y Thomas-Fiering para generar secuencias mensuales de lluvia de forma independiente para la cuenca Kalangi, simultáneamente emplea modelo multisitio de unificación de Young y Pisano para la generación de secuencias sintéticas de lluvia, obteniendo buenos resultados ya que los modelos se ajustaron satisfactoriamente a las precipitaciones mensuales de las cuatro estaciones pluviales de la cuenca.

Debido a que el recurso más valioso en regiones con climas áridos o semi-áridos es el agua, el auge de investigaciones en estas zonas ha sido considerable, por lo limitado del recurso y por ser la base para el desarrollo y estabilidad de los ecosistemas. Un ejemplo, es el caso en Jordania (Momani, 2009), un país donde además de que el recurso hídrico es limitado, la calidad del mismo se ha ido deteriorando debido al desarrollo urbano. Es importante conocer el futuro de los recursos hídricos para una correcta gestión de la demanda del agua, y teniendo en cuenta esta problemática se desarrolló un modelo ARIMA (1,0,0) (0,1,1) usado para pronosticar la precipitación mensual con el objetivo de obtener los valores máximos de precipitación de los próximos 10 años. Con base a los resultados obtenidos, el modelo ARIMA fue adecuado para representar los datos y logró pronosticar los datos de precipitación próximos, permitiendo establecer estrategias, prioridades y toma de decisiones con respecto a los recursos hídricos en dicho país. Así mismo, en el Lago Urmia al noroeste de Irán (Farajzadeh, Fakheri Fard, & Lotfi, 2014) se ha utilizado una Feed-forward neural network y un Modelo Autorregresivo de Promedios Móviles Integrados (ARIMA) para la predicción de la precipitación en la cuenca, con el fin de investigar el escurrimiento producido hacia el lago. En términos generales, la capacidad para desarrollar el procesamiento y aprendizaje de un proceso (mediante patrones correspondientes a las variables de entrada) en forma paralela, adaptativa y no lineal es la ventaja principal al aplicar ANN.

Se destaca que al investigar las habilidades de diferentes modelos de series de tiempo y comparar su capacidad de predicción en condiciones climáticas semiáridas, Dastorani M. (2016) dice que seleccionar una técnica adecuada para modelar un fenómeno como la precipitación depende de varios factores como lo son: la precisión de los datos, tiempo, costo, facilidad de uso de los resultados del modelo así como la interpretación

de los datos, por lo que realizaron el análisis de los datos de 9 estaciones en la provincia de Khorasan del norte en Irán, en donde los modelos MA(2), AR(1), ARMA(2,1) y ARIMA(1,1,2) obtuvieron el mejor rendimiento de pronósticos mensuales de lluvia.

Sin embargo, no solo el problema de la escasez de recurso hídrico es un motivo para prever o predecir el futuro mediante modelos de series temporales, en regiones con climas tropical y subtropical como lo es la ciudad de Dhaka en Bangladesh (Mahsin, Akhter, & Begum, 2012) este tipo de análisis y pronóstico mediante series de tiempo son muy utilizados, debido a la susceptibilidad de la ciudad a inundaciones durante las épocas de monzones, así como a fuertes lluvias y ciclones. En este caso se hizo uso de un modelo multiplicativo Seasonal Autoregressive Integrated Moving Average Model (SARIMA) el cual contiene un orden regular y un orden estacional, con el propósito de pronosticar la precipitación mensual para un periodo de 2 años hacia adelante, los resultados del modelo ARIMA (0,0,1) (0,1,1)¹² fueron lo suficientemente precisos para pronosticar las próximas lluvias y establecer estrategias, prioridades, además del uso adecuado de los recursos hídricos en Dhaka. Ahora, hablando del análisis de la tendencia en la precipitación, ésta ha generado interés debido a que es un componente fundamental del sistema climático, por lo cual ha surgido la necesidad de estudiar su variabilidad espacial y temporal dentro del contexto del cambio climático global (Guenni, Degryze, & Alvarado, 2008). Así pues, Chang, et al. (2012) emplea un modelo SARIMA (1,0,1)(0,1,1)₁₂ en Yantai, China con el objetivo de predecir la tendencia de la precipitación. Las predicciones basadas en el modelo indican que durante los próximos tres años la precipitación disminuirá, sugiriendo que la aplicación del modelo permitió realizar estrategias apropiadas para el manejo del recurso hídrico en respuesta a la escasez de agua. Generalmente, la mayoría de las investigaciones son direccionadas al modelado y al pronóstico de la precipitación, sin embargo, hay poca información sobre la frecuencia de la precipitación, por lo que Akpanta, et al. (2015) utilizan la técnica de modelado de series de tiempo para obtener de manera concisa esa información. La aplicación llevada a cabo en Umuahia, la capital del estado de Abia en Nigeria mediante un modelo SARIMA (0,0,0) (0,1,1)₁₂ resultó favorable para pronosticar la frecuencia de lluvias mensuales, es decir, la flexibilidad, sistematicidad y facilidad de captación de información original de series de tiempo son ventajas que los modelos ARIMA estacionales ofrecen y por lo tanto son ampliamente utilizados. Considerando las ventajas antes mencionadas y tomando en cuenta los efectos del cambio climático observados en diversas variables ambientales en muchas regiones del mundo, la lluvia es la variable más afectada debido a sus distribuciones no homogéneas en el tiempo y en el espacio, los modelos SARIMA han sido utilizados también para el pronóstico de precipitaciones mensuales con el propósito de proveer a las autoridades el desarrollo de planes estratégicos, políticas y el uso apropiado de los recursos hídricos disponibles (Nirmala & Sundaram, 2010; Papalaskaris, Panagiotidis, & Pantrakis, 2016). Yusof

(2012) por ejemplo, aplico dos técnicas de análisis de series de tiempo: los modelos espaciales de estado basados en el suavizado exponencial ETS y el método de modelado SARIMA; demostrando que ambos modelos fueron adecuados para pronosticar las precipitaciones y brindar información que permita establecer estrategias para la planificación adecuada de la agricultura, el sistema de drenaje y otras aplicaciones de recursos hídricos.

Ahora bien, la aplicación de modelos autorregresivos también ha sido utilizada para la predicción de precipitaciones de embalses hidroeléctricos debido a la influencia de las precipitaciones en los recursos hídricos de la cuenca. Tal es el caso de Indonesia en donde muchas de las centrales hidroeléctricas no funcionan de manera óptima después de operar más de 15 años, por la degradación de los recursos hídricos con el tiempo, siendo indispensable realizar un estudio de investigación sobre la sostenibilidad de los recursos hídricos para hacer predicciones de la ocurrencia de lluvias en la hidroeléctrica de la cuenca Mamasa, la cual suministra el recurso de los planes hidroeléctricos de Bakaru (Said, Manjang, Tjaronge, & Thaha, 2013).

Otro aspecto es que la precipitación juega un papel importante en la determinación del clima de una región, por lo cual se requiere una estimación precisa para gestionar y planificar los recursos hídricos, así como otras aplicaciones relacionadas con la hidrología, la climatología, la meteorología y la agricultura. Lo anterior ha originado la necesidad de innovar para lograr el modelado preciso de la precipitación. Soltani (2007) por ejemplo, empleó modelos de series de tiempo para determinar e identificar los climas regionales de Irán mediante las precipitaciones mensuales de 28 ciudades principales; para determinar los climas regionales, aplicando un análisis jerárquico de conglomerados sobre los coeficientes de autocorrelación encontrando tres grupos climáticos principales. Los resultados del modelado mostraron una alta variación del patrón temporal de la precipitación mensual sobre Irán, excepto en los márgenes del Mar Caspio y el Golfo Pérsico y los diferentes modelos presentaron una gran variación en el mecanismo de producción de precipitaciones espaciales y diferentes características de estacionalidad y estacionariedad del patrón temporal de lluvia sobre Irán. Valipour (2012) estimó la capacidad de los modelos de series de tiempo ARMA y ARIMA para pronosticar la precipitación e investigar el efecto de las condiciones climáticas sobre el número de observaciones apropiadas en 4 estaciones con diferentes condiciones climáticas (templado, húmedo, árido y semiárido). Se encontró que los modelos de series de tiempo son más apropiados para la predicción de lluvia en clima semiárido y que un número de 60 datos de observación de lluvia son suficientes.

Ahora bien, tomando en cuenta que el principal problema del modelado de variables hidrológicas utilizando métodos de computación es ignorar la parte estocástica de las series de tiempo hidrológicas, Mehdizadeh (2017) introduce tres nuevos modelos

híbridos integrando métodos informáticos que incluyen splines de regresión adaptativa multivariada (MARS), redes bayesianas (BN) y programación de expresión génica (GEP) con un modelo de series temporales, en específico el de heterocedasticidad condicional autorregresiva generalizada (GARCH) para el modelado de la precipitación mensual, es decir, las partes determinísticas (obtenidas mediante los métodos informáticos) y estocásticas (obtenidas por el modelo de la serie temporal GARCH) de la precipitación se combinan entre sí, obteniendo modelos híbridos MARS-GARCH, BN-GARCH y GEP-GARCH. En general, los modelos propuestos arrojaron una mayor precisión que los modelos únicos tradicionales y confirmaron la idoneidad de la metodología propuesta para el modelado de la precipitación.

Considerando que la mayoría de los datos hidrológicos son de naturaleza no estacionaria y que los modelos autorregresivos se emplean en procesos estacionarios y lineales, una de las limitantes es que la serie de datos analizados sean no estacionarios y no lineales. Para lo anterior, la utilización de Wavelet (Particularly Discrete Wavelet Transform DWT) y el EMD (Empirical Mode Decomposition) basados en modelos ARMA son métodos que permiten descomponer los datos no estacionarios y no lineales en conjuntos de componentes más simples que al ser combinados con componentes aleatorios pueden ser modelados fácilmente permitiendo producir predicciones futuras de series temporales, como es el caso de Karnataka en la India (Karthikeyan & Nagesh Kumar, 2013) donde el enfoque de la investigación es determinar la predictibilidad de series temporales no estacionarias con el objetivo de pronosticar sequías con mayor precisión. Finalmente, ambos modelos pueden ser utilizados para modelar sequías a pesar de la no estacionariedad de países tropicales como la India, sin embargo, el método basado en Wavelets demostró mayor precisión en comparación al EMD aumentado la posibilidad de utilizarlo para el modelado de inundaciones mediante ciertas modificaciones en los parámetros. De igual forma, Kwon (2007) presenta un esquema de simulación de series temporales basado en la descomposición de wavelets acoplado a un modelo autorregresivo (WARM) para series hidroclimáticas que exhiben variabilidad de baja frecuencia para la precipitación y temperatura. El objetivo consistió en generar un simulador de series temporales para sistemas con comportamiento cuasiperiódico de memoria larga, dinámicas no lineales con comportamiento similar al régimen persistente o transiciones de régimen estocástico; en este caso la transformación wavelet continua se aplicó para descomponer la serie temporal univariada en varios componentes estadísticamente significativos y luego un modelo AR lineal es empleado para simular cada componente extraído desde el análisis de transformada wavelet. El modelo WARM propuesto arrojó mejores resultados que el modelo tradicional en lo relacionado a la reproducción de las propiedades tiempo-frecuencia para la precipitación observada. Una de las principales ventajas para las transformadas wavelets es que permiten la descomposición ortogonal de la señal original en el tiempo y en el dominio de la frecuencia, lo cual las hace aptas

para ser utilizadas en este tipo de variables. En el mismo contexto, países cuya economía depende de la agricultura, durante décadas se han enfrentado a sequías recurrentes convirtiéndose en fenómenos normales, un ejemplo de esto es la ciudad de Gadaref en Sudan en donde Hassan (2015) presentó un modelo lineal estocástico SARIMA para simular sequías mediante las series de índice de precipitación mensual estandarizada (SPI). Los resultados mostraron que el modelo ajustado es adecuado para ayudar a la toma de decisiones ante la necesidad de establecer prioridades para la demanda, el almacenamiento y la distribución del agua.

Con relación al análisis de los parámetros climáticos para minimizar efectos adversos asociados a cambios en las temperaturas y en la precipitación (variabilidad climática) debido al cambio climático, el análisis de series de tiempo también es una de las herramientas más importantes. Dabral, et al.(2016) aplicaron el análisis de series de tiempo para evaluar el cambio de la precipitación mensual y las temperaturas máximas y mínimas en ambientes húmedos de la India, específicamente en Jorhat (Assam), el objetivo de este estudio fue ayudar a los organismos gubernamentales para que incluyeran el impacto del cambio climático en la toma de decisiones con la finalidad de proteger la rica biodiversidad de esta región, un modelo AR(12) fue el adecuado pronosticando 41 años (2009-2050) e indicando que el clima en el área de estudio se mantendrá más o menos estable durante las próximas cuatro décadas.

Otro campo de aplicación donde se utilizan los modelos autorregresivos es en la medición sistemática de la composición del agua de lluvia, las cuales surgieron en Estados Unidos hace aproximadamente cuatro décadas como parte de un programa Nacional de Deposición Atmosférica, en donde la información puede ser utilizada para estudios que aborden la relación entre emisiones de azufre y los óxidos de nitrógeno y las concentraciones de varios analitos principales observadas en agua de lluvia. Según Nickerson (2005) los modelos ARIMA de datos de precipitación son una mejora con respecto a los modelos no lineales para la predicción de la composición química de la precipitación. En este caso el modelo ARIMA con tendencia lineal se ajustó a la concentración y a los valores de deposición húmeda de cada analito y luego se usó para pronosticar la concentración del analito y la deposición húmeda funcionando considerablemente mejor que los modelos NLR.

- *Precipitación anual*

Una parte importante respecto al estudio de la modificación climática es comprender su evolución, para esto la construcción de modelos de series de tiempo permiten conocer las características temporales de las variables de interés, así como predecir su evolución en el tiempo, convirtiéndolo en un desafío y en un tema de gran interés

científico ya que frecuentemente, la estructura del modelo de una serie se puede ver afectada por cambios, generando que las propiedades estadísticas se vean modificadas. Dichas especificaciones inapropiadas en el modelo propuesto arrojan representaciones que no reflejan el comportamiento real de la variable estudiada y por consecuencia la obtención de resultados inexactos e incorrectos y una errónea interpretación de los mismos (Junttila, 2001; Tsay, 1988). Hay que destacar que la predicción de las precipitaciones a largo plazo se utiliza para tomar decisiones económicas y agrícolas en muchos países y que además es una herramienta que permite minimizar la devastación causada por sequías recurrentes, tal es el caso de El-Fandy (1994) que utilizó dos modelos ARIMA, uno para los niveles de las inundaciones estacionales en el río Nilo (El Cairo) y el segundo para modelar la precipitación anual en Etiopía.

Pocos estudios se han dedicado a la investigación de la evolución de la precipitación en distintas regiones, uno de ellos es el realizado en Dobrudja (Rumania), en donde el periodo más largo sin lluvia registrado en la zona ha llegado al 89% debido a la frecuencia de los años de sequía, por lo cual (Bărbulescu & Pelican, 2003) investigaron las características temporales de la evolución de la precipitación mediante modelos ARIMA. Considerando los cambios que pueden darse en la estructura de las series, realizaron pruebas de ruptura tales como *Pettit Test* y *Segmentation procedure of Hubert* para detectar algún cambio promedio en las series de tiempo y los instantes en los que suceden.

Partiendo desde el punto de vista de las principales actividades económicas desarrolladas en el mundo, la agricultura ha sido seriamente afectada debido a las sequías y escasez de agua, convirtiéndose el agua de lluvia en un importante recurso que requiere de pronósticos precisos para la planificación de la asignación de agua con fines agrícolas. Un ejemplo, es el caso de Tailandia en donde existe una alta variación interanual de las precipitaciones debido a los monzones, causando problemas relacionados con el agua, como inundaciones o sequías que dañan los cultivos. Ante esta problemática (Weesakul & Lowanichchai, 2005) desarrollaron un modelo para pronosticar la precipitación anual en el país mediante modelos ARMA y ARIMA que provee pronósticos aceptables para cumplir con los requerimientos de planificación de asignación de agua a distintas zonas agrícolas, además mediante los cálculos estadísticos realizados con base en la distribución espacial para cada estación pluviométrica, pudieron determinar las zonas de ocurrencia de precipitaciones baja-moderada y moderada-abundante en todas las regiones del país. Caso similar se presenta en la India en donde el éxito o la falla de los cultivos y la escasez de agua es un asunto de mucha preocupación, estos problemas están altamente asociados con el comportamiento de las lluvias monzónicas de verano, por lo tanto el pronóstico del promedio de lluvias monzónicas es necesario para el futuro agrícola; por lo que

(Chattopadhyay & Chattopadhyay, 2010) desarrollaron un modelo híbrido al que denominaron Modelo Autorregresivo de Redes Neuronales ARNN. Este modelo toma la parte autorregresiva de un modelo ARIMA y una red neuronal, obteniendo una red neuronal artificial autorregresiva cuyo propósito es investigar la estacionariedad dentro de las series de tiempo de lluvias monzónicas, ya que los monzones de verano son el periodo más productivo con respecto a las prácticas agrícolas, en donde además de conservar el agua de lluvia para estos periodos, también se utiliza para futuros propósitos de irrigación en la ciudad. Sin embargo, debido a la ubicación geográfica de la India, se presentan zonas tropicales y sub-tropicales, por lo tanto, los monzones gobiernan casi todo el territorio, siendo esta la razón por la cual las precipitaciones monzónicas se convierten en tema de investigación debido a la gran variabilidad de la misma y al impacto económico que estas generan. Teniendo en cuenta lo anterior, se ha empleado el análisis de series de tiempo de precipitaciones monzónicas interanuales, con el objetivo de ajustar modelos estocásticos adecuados mediante la aplicación de un filtro de 3 puntos al proceso autorregresivo (AR) permitiendo que el modelo pueda ser usado para pronosticar monzones con tiempo de anticipación de casi 1 año. Estas predicciones pueden combinarse con otras para aumentar la habilidad de los pronósticos (Dahale & Singh, 1993), así también Valipour (2016) comparó la capacidad de tres estructuras diferentes para detectar alarmas de sequía y año húmedo: la red neuronal autorregresiva no lineal (NARNN), la entrada-salida no lineal (NIO) y la NARNN con entrada exógena (NARNNX) prediciendo la precipitación anual por codificación en el entorno de software MATLAB para 27 estaciones pluviométricas. Los resultados mostraron que la precisión de la red NARNNX es mejor que la de NARNN y NIO en función de los valores de r . Sin embargo, el NARNN fue adecuado para los casos en que el número de datos disponibles para el análisis de precipitación fue limitado o no se observaron puntos máximos. Considerando que el clima y las precipitaciones son fenómenos altamente no lineales y complicados, Somvanshi (2006) aplicó dos enfoques diferentes para el diseño de un modelo para modelar y predecir el patrón de comportamiento de la lluvia basado en observaciones pasadas, el primero un método estadístico ARIMA y el segundo una ANN, revelando que el modelo de ANN se puede utilizar como herramienta de pronóstico adecuada para predecir la precipitación.

Finalmente, es relevante identificar la relación que existe entre el cambio climático y los recursos del agua. Según Fan, et al. (2011) en la zona del río Tarim en Xinjiang China, el aumento en las temperaturas tuvo un impacto positivo ya que la precipitación y la escorrentía en las cuatro corrientes del río aumentaron, generando una entrada de flujo mayor. En este caso se realizó un análisis de las características del proceso hidrológico y la respuesta de la escorrentía a los cambios de la temperatura y precipitación en la cuenca del río mediante un modelo *Holt double exponential smoothing* el cual fue utilizado para ajustar las dos variables (precipitación y temperatura), además un modelo ARIMA (0,1,1) fue utilizado para pronosticar 50 años de escorrentía anual en el

cual las estimaciones de los parámetros y los valores pronosticados muestran un buen ajuste generando predicciones precisas con el fin de hacer frente al problema respecto a la disponibilidad del agua.

- *Compuesta*

La metodología del análisis de series de tiempo brinda la opción de realizar el modelado de series de forma multivariante, las cuales proporcionan una descripción efectiva de las series de tiempo analizadas. En el sur de la Florida, Chin D A. (1995) realizó el modelado de series de tiempo multivariante mediante un enfoque de modelo a escala y un modelo de precipitación CARMA en 12 cuencas hidrológicas. Los resultados obtenidos mostraron poco favorable el enfoque del modelado CARMA, debido a que los modelos a escala (regional o local) consideran que las desviaciones de los fenómenos a gran escala cuentan con una estructura temporal, sin embargo, los modelos ARMA no consideran dichos fenómenos y asumen que la varianza se asocia a fenómenos aleatorios distribuidos (a gran escala) que no tienen dicha estructura, por lo cual el enfoque de modelado a escala multivariante resultó ser una alternativa efectiva.

Dentro del análisis de series de tiempo para periodos mensuales y anuales (Ünal, Aksoy, & Akar, 2004) generaron series sintéticas de precipitación a través de la aplicación de modelos autorregresivos (AR), modelo Thomas-Fiering (TF), método de fragmentos (F) y su versión modificada (MF), método de dos colas (TT) y el modelo basado en acercamiento de ondas (W), obteniéndose mejores resultados en la correlación con el método W. Lo anterior indica que los modelos estocásticos pueden ser relacionados con características morfológicas, variables y fenómenos atmosféricos ya sea de forma espacial o temporal, destacándose la variable de la precipitación; en lo que se refiere a la construcción de modelos estocásticos para la predicción de la misma.

Es importante destacar que algunas de las ventajas de estos modelos es permitir el análisis de ocurrencia de incertidumbres, eventos extremos, cambios climáticos, así como estrategias para determinar políticas de operación de presas, etc. Con relación a la caracterización de cambios mensuales de niveles de agua en lagos y embalses, ha sido utilizado el modelo Thomas-Fiering ARMA (Somlyódy & Honti, 2005). Actualmente los modelos autorregresivos facilitan el estudio de modelos conceptuales para procesos lluvia-escorrentía desarrollados para distintos escenarios climáticos (Lidén & Harlin, 2000).

1.1.2 Temperatura

La temperatura es el parámetro meteorológico más importante, a pesar de esto, los pronósticos que proporcionan distintos servicios meteorológicos no especifican la hora de ocurrencia, impidiendo o limitando el uso de la información y, en otros casos esta información no se encuentra disponible siendo a veces necesario que los pronósticos de temperatura dependan sólo de los parámetros disponibles o medidos en el sitio (Abdel-Aal, 2004). Frente a la naturaleza compleja y no lineal de las variaciones de temperatura y la abundancia de datos históricos se requiere de técnicas de modelado basadas en datos que permitan resolver el problema de pronóstico de la temperatura.

De igual forma que con la precipitación, los modelos autorregresivos han sido utilizados en diversos países latinoamericanos con el fin de estudiar varios fenómenos y componentes hidrológicos, entre los más comunes se encuentran los relacionados con el fenómeno del Niño, ya que es un suceso que conecta el clima de manera global, los ecosistemas, así como las actividades socioeconómicas. Dado que es un evento complejo, los modelos estadísticos y dinámicos son insuficientes, por lo tanto, el intento de predecirlo se remonta a la década de los 80. En este sentido, Quispe (2007) utilizó modelos autorregresivos para estudiar el efecto del fenómeno del Niño en la temperatura superficial del mar (TSM) frente a la costa peruana, dicho trabajo se desarrolló en 5 etapas basadas en la metodología Box & Jenkins: identificación, estimación, verificación diagnóstica, previsión y validación; logrando simular condiciones mensuales similares a las observadas en el litoral peruano, condiciones frías a finales del 2004, y condiciones neutrales a principios del 2005.

En la región de Centroamérica, se realizó el ajuste de un modelo Vectorial Autorregresivo de Medias Móviles (VARMA) para el estudio de las condiciones anómalas de la temperatura superficial del mar (ATSM) de los océanos Pacífico y Atlántico Tropical, con el objetivo de cuantificar la interacción océano-atmósfera entre distintos índices oceanográficos y los regímenes de precipitación en Centroamérica, encontrando que en el Atlántico Tropical Norte las ASTM influyen más fuertemente sobre la precipitación que aquellas del Pacífico Tropical Este (Alfaro & Cid, 1999).

1.1.2.1 Aplicación de Modelos Autorregresivos para predicción de la temperatura

Como ya se mencionó anteriormente, uno de los parámetros básicos que caracterizan el clima de una región es la temperatura, por esta razón se presenta una revisión de investigaciones realizadas para las series de temperatura diaria, mensuales y/o anuales en distintos lugares del mundo.

- *Temperatura diaria*

Uno de los objetivos principales por el que los investigadores emplean los modelos autorregresivos para el análisis de las variables meteorológicas es que permite el análisis de la estructura interna de las series temporales de la variable en cuestión y revelar las diferencias climatológicas entre distintos sitios, Ceschia (1986) por ejemplo analizo las temperaturas medias diarias utilizando procesos estocásticos para obtener el componente estacional desarrollando un método computacional que involucra la expansión de Fourier y en donde la serie residual estacionaria se llevó a cabo mediante filtros recursivos ARMA, para realizar dicho análisis se utilizó el programa MICMET con el objetivo de obtener el análisis detallado de los fenómenos micrometeorológicos en la región de Friuli-Venecia, específicamente para Udine y Lignano.

Una de las ventajas de centrarse en el análisis de datos diarios es que resulta más útil debido a que el proceso de serie temporal puede cambiar a lo largo de un año, en el caso del modelado de series temporales de datos meteorológicos diarios permite comprender mejor una serie de actividades dependientes del clima (por ejemplo, la demanda de electricidad máxima diaria). Haciendo referencia a lo anterior, Bowden (2012) muestra que es posible que múltiples series temporales independientes del mismo proceso ARMA pueden representarse mediante una única serie temporal ARMA univariante mediante un entrelazado de la serie original, utilizando las series de temperaturas máximas diarias de Perth (Australia Occidental) demostró que dicha técnica de entrelazado puede extenderse a series diferenciadas, modelos ARMA estacionales, modelos VARMA y modelos GARCH, con la restricción de que las series deben de ser de la misma longitud. Ahora bien, argumentando que la volatilidad del registro de temperatura tiende a agruparse, es decir, la predictibilidad del clima no es constante, pero muestra una variación sistémica, Tol (1996) empleo un modelo GARCH para capturar esa característica en las temperaturas medias diarias de verano e invierno; en este modelo la varianza condicional de una observación depende linealmente de las varianzas condicionales de las observaciones previas y de los errores de predicción previos. En base a los resultados obtenidos se demostró que esos modelos heteroscedásticos superan a sus versiones homoscedásticas y que se prefiere el modelo que actualiza la desviación estándar condicional.

En relación a simular datos climáticos sintéticos mediante generadores de clima estocástico, Breinl (2015) desarrollo un algoritmo semi-paramétrico para la precipitación diaria de sitios múltiples usando un proceso de Markov univariado; para simular la temperatura diaria promedio en tres sitios empleo modelos ARMA, en donde las transformaciones de potencia fueron utilizadas para reducir el sesgo de temperaturas extremas simuladas, logrando que el generador meteorológico multi-sitio propuesto sea adecuado para evaluar peligros hidrometeorológicos ya que reproduce muy bien la variabilidad espacial, puede generar precipitaciones y temperaturas extremas,

reproduce de forma aceptable la autocorrelación y es sencillo. Otro punto importante es la generación de secuencias sintéticas de temperaturas ya que son útiles cuando es necesario simular el efecto de ésta en un sistema, Magnano (2008) presento modelos potentes que pueden usarse como herramientas para generar secuencias sintéticas de temperaturas de media hora, tomando en cuenta la estacionalidad intra-día e intra-anual así como las correlaciones seriales a corto y largo plazo se generaron las temperaturas empleando una combinación de modelos (ARMA para temperaturas diarias). El enfoque utilizado en este caso tiene en cuenta la inercia térmica que se produce cuando la temperatura es superior al promedio mensual durante días sucesivos hasta que se produce un cambio frío y la temperatura está por debajo del promedio mensual durante un período corto. El modelo obtenido fue utilizado para generar diferentes patrones de demanda de electricidad permitiendo un análisis detallado de como la demanda de electricidad varia con los diferentes perfiles de temperatura.

Los modelos de series temporales son aplicables a una diversidad de casos, uno de los que actualmente resulta interesante es la producción de energías renovables, en el caso de la temperatura la energía solar fotovoltaica cuenta con innumerables ventajas, desde este punto de vista la aplicación de un modelo autorregresivo con modelos de entradas exógenas (ARX) en patrones de radiación y temperatura así como la utilización de instrumentos virtuales para medir y analizar datos en Bogotá permitió evaluar el potencial solar disponible para diseñar sistemas fotovoltaicos (pv) y obtener un polinomio para el comportamiento de la radiación solar representando una fuente confiable de información para interpretar y predecir a corto plazo la irradiancia y la temperatura (Chica, Rey, & Aristizábal, 2012).

En lo que se refiere a la ciencia de la contaminación del aire, por ejemplo, las emisiones, la formación y/o el transporte de contaminantes comprender la variabilidad espacial y temporal a corto plazo (diaria), por lo tanto la temperatura es importante ya que los aspectos mencionados anteriormente pueden variar en función del clima; por estas razones, Im (2009) propuso un proceso parsimonioso para describir la estructura de covarianza espacio-temporal de las temperaturas máximas y mínimas diarias en el área metropolitana de Chicago, para realizar el modelo consideraron tres covariables físicas: la latitud, la distancia al Lago Michigan y los vientos descompuestos en NE, SW, NW y componentes SE. Para describir la correlación temporal se empleó un modelo ARMA en donde los campos de innovación de la serie temporal se modelaron como campos aleatorios gaussianos espacialmente correlacionados, pero temporalmente no correlacionados con la función de covarianza de Matérn.

Es importante destacar que el estudio de las características estadísticas de eventos extremos en la temperatura diaria es un aspecto importante en el análisis del clima, ya que estas características afectan varias actividades humanas como lo son la

producción agrícola, el diseño de invernaderos agrícolas y el consumo de energía doméstica por mencionar algunas, frente a esta situación Macchiato (1993) investigo las series temporales de temperatura en el sur de Italia debido a la extrema variabilidad climática que presenta la zona. Los componentes determinísticos y estocásticos de las series de 50 estaciones se identificaron mediante un modelo dinámico-estocástico que es periódico en la parte determinista y Markoviano en la parte estocástica, con base en el modelo ajustado se lograron identificar patrones espaciales que hacen posible extrapolar la probabilidad de eventos fríos (calientes) en lugares donde no hay observaciones directas disponibles.

Con respecto a áreas de aplicación como la agricultura se requiere de datos de temperatura y lluvia tan precisos como sea posible, por lo tanto se han tomado prestadas técnicas geoestadísticas como el kriging (con deriva externa o regresión y kriging de residuos RK) o el modelado estadístico con el objetivo de mejorar las estimaciones de las variables relevante, con referencia a lo anterior Joly (2013) se propuso analizar los residuos (que no son necesariamente impredecibles porque pueden estar correlacionados en el tiempo) mediante procesos autorregresivos empleando primero la técnica RK, los resultados mostraron que los procesos autorregresivos mejoraron significativamente las estimaciones evaluadas mediante validaciones cruzadas. Así pues, determinaron que la mejora de la calidad de las estimaciones a través de la metodología propuesta depende tanto de la linealidad entre los residuos de d día, como de los días $d-1$ y la desviación estándar de los residuos en día d .

Actualmente los procesos que ocurren en campos como el medio ambiente a menudo exhiben una mezcla de diferentes dinámicas en donde la estimación de modelos es difícil debido a la dimensión y a la cantidad de parámetros involucrados, buscando una solución a esta problemática Monbet (2017) propone el uso de *Markov switching vector autoregressive models* (MSVAR) y una *Smoothly Clipped Absolute Deviation penalization of the likelihood* para reducir los parámetros hacia ceros y regularizar el problema de inferencia para datos de temperatura media diaria.

Finalmente demostró que los modelos MSVAR dispersos no homogéneos permiten reproducir no linealidades existentes en los datos tales como covarianzas no separables o cruces no simétricos y que la introducción de transiciones no homogéneas permite mejorar los resultados comparados a modelos con transiciones homogéneas entre los regímenes. Sin embargo, a pesar de que se logra reducir significativamente la cantidad de parámetros en los modelos y ayuda a la interpretación, el algoritmo propuesto es computacionalmente costoso y está lejos de ser exhaustivo.

- *Temperatura mensual*

Frente al calentamiento global observado y proyectado en los siglos XX y XXI el cual ha afectado la agricultura, el ciclo hidrológico, las condiciones del medio ambiente y sistemas ecológicos; la lluvia, la evaporación, la transpiración y la evapotranspiración son componentes vitales e importantes para los procesos de riego y las prácticas agrícolas. Desde esa perspectiva, Chattopadhyay (2011) adoptó un enfoque autorregresivo para inspeccionar las series de tiempo de la temperatura máxima mensual en la ciudad de la India que depende económicamente de cultivos de arroz, te y productos forestales; el modelo autorregresivo de la serie de tiempo T_{max} reducida generado mediante ecuaciones de Yule-Walker fue comparado con el modelo de red neuronal autorregresivo con perceptrón multicapa, red neuronal de alimentación anticipada generalizada y red neuronal modular (AR-NN (4))-MNN mostrando que la fusión de dichos métodos es una mejor opción para pronosticar una serie temporal ya que requiere de menos predictores para un pronóstico superior. Así mismo, motivado por los problemas que ocurren en áreas como el medio ambiente y geografía, Gao (2011) propuso un modelo de regresión multivariado con errores ARMA espacio-temporales (STARMA) aplicándolo a las temperaturas medias mensuales en cinco ubicaciones, los resultados obtenidos muestran que tras el análisis de variabilidad temporal y espacial de la temperatura, las temperaturas medias mensuales en todas las ubicaciones están aumentando, encontrando las mayores tendencias en Faraday / Vernadsky y Rothera. Gil-Alana (2012) también empleó un modelo ARMA para examinar las características de tendencia, persistencia a largo plazo y estacionalidad para las temperaturas mensuales en Alaska enfocándose en el cambio climático y la hipótesis del calentamiento en las anomalías de las series mediante la eliminación de la estructura estacional encontrando evidencia de calentamiento climático en la mayoría de las estaciones analizadas.

En China, Gu (2005) estableció un modelo autorregresivo complejo basado en derivación matemática de mínimos cuadrados en la predicción de las anomalías mensuales de temperatura en 160 estaciones meteorológicas, el modelo propuesto difiere del convencional en que el número real y el número imaginario se calculan por separado; la aplicación de este modelo demostró que usar un número complejo para ajustarse a un campo de elementos meteorológicos y predecir con el modelo autorregresivo complejo es efectivo para mejorar los resultados del pronóstico y para ser usado en campos de elementos similares en las ciencias geofísicas. Recientemente Nury (2017) propuso un modelo alternativo para la predicción de la temperatura combinando la técnica wavelet con un modelo ARIMA y una red neuronal artificial ANN aplicada a los datos mensuales de temperatura máxima y mínima, en donde los datos se dividen en un conjunto de datos de capacitación para construir los modelos y un conjunto de datos de prueba para estimar su rendimiento. Los resultados obtenidos

indican que el modelo wavelet-ARIMA es más efectivo que el modelo wavelet-ANN, concluyendo que el cambio climático está afectando la temperatura en el noreste de Bangladesh. Iqelan (2015) por su parte, decidió comparar la eficiencia de los modelos ARIMA y AR-GARCH para predecir una serie de temperatura promedio mensual en Jerusalén/Palestina; el modelo GARCH lo empleo para proporcionar una medida de volatilidad de la serie; en base a los resultados obtenidos se detectó que el modelo AR-GARCH tiene un mejor rendimiento que el ARIMA, esto debido a su capacidad para cautivar la volatilidad o las variaciones (errores) condicionales de las variables en el tiempo además de considerar la posibilidad de poder usar otro tipo de modelos como el *Integrated* GARCH (IGARCH) y *GARCH exponencial* (EGARCH) con la asistencia del software R.

Khedhiri (2016) estudio las propiedades estadísticas de temperaturas medias mensuales mínimas y máximas utilizando métodos de prueba estadísticos alternativos para identificar el tipo de patrón estacional en los datos y desarrollar un modelo con la capacidad de producir pronósticos precisos; empleando pruebas de raíces unitarias alternativas estacionales determinaron un modelo ARIMA estacional moderado para cada serie demostrando que los modelos sugeridos funcionan bien y realizan buenas predicciones, confirmando los hallazgos anteriores de que las temperaturas medias en la región del Atlántico canadiense continuaran aumentando.

Otro punto importante por el cual se investigan las tendencias de cambio en las variables climáticas se debe al problema de detectar y pronosticar una señal de cambio en el registro climatológico para comprender los cambios regionales y globales, Lee (2007) realizo un análisis de la temperatura del aire de la superficie en una región de China, en este caso utilizo un método de descomposición en el análisis de series temporales representada por un modelo ARMA, para la estimación de parámetros emplearon el método *Ordinary Least Square* (OLS) debido a que proporciona estimaciones consistentes incluso en el caso en que los errores no sean estacionarios y se modulen uniformemente; encontrando que no detectaron tendencias significativas de aumento en las series de temperatura de anomalías mensuales regionales, sin embargo, muestran características interesantes que tienen relevancia para los problemas del cambio global, concluyendo que un modelo climático regional es una forma prometedora de reducir la incertidumbre en el análisis de la variación climática regional.

- *Temperatura anual*

Tradicionalmente se considera que un modelo ARMA es capaz de generar oscilaciones lineales, sin embargo, a menudo estos modelos no reproducen la firma espectral de las

series de tiempo, específicamente la modulación de frecuencia de amplitud a lo largo del tiempo, desde esta perspectiva Kwon (2007) presenta un esquema de simulación de series temporales basado en la descomposición de wavelets acoplado a un modelo autorregresivo para series hidrológicas que exhiben variabilidad de baja frecuencia de banda limitada. El uso de la descomposición wavelet seguido por un modelo autorregresivo de cada componente principal se exploró como un modelo para tales series de tiempo, con el fin de simular series temporales para sistemas con comportamiento cuasiperiódico de memoria larga, dinámicas no lineales que pueden conducir a un comportamiento similar al régimen persistente o transiciones de régimen estocástico. El modelo WARM propuesto arrojó mejores resultados en términos de reproducción de las propiedades de tiempo y frecuencia, y a su vez conserva las estadísticas generalmente reproducidas por los modelos AR.

Así mismo, dada la preocupación actual sobre evidencia de calentamiento global y debido a que investigaciones anteriores sugieren que la evidencia de memoria prolongada se puede eliminar a causa del ajuste de modelos con tendencias que cambian lentamente, Mills (2007) muestra que es posible modelar características de memoria larga de temperatura mediante un proceso autorregresivo de promedio móvil fraccionalmente integrado (ARFIMA) concluyendo que el modelo se puede utilizar para pronosticar las temperaturas del hemisferio norte.

Por otra parte, el caracterizar los cambios climáticos en una región es de gran ayuda ya que permite comprender los efectos de dichos cambios, Croitoru (2012) propuso un análisis punto-cambio para determinar el tiempo y la magnitud de los cambios de temperatura en verano en las cumbres de los Cárpatos rumanos, para ello empleó un método estadístico basado en modelos de regresión con errores ARMA correlacionados en serie; dicho modelo permite obtener los errores estándar correctos de pendiente e interceptar los valores estimados, de esta manera, los intervalos de confianza resultantes tienen la longitud correcta lo cual es crucial en el modelo. Finalmente, se confirmó el calentamiento general de la región, identificando puntos de cambio para cada estación tanto en áreas altas como bajas, concluyendo que las regiones más altas de esa región han experimentado 3 veces el aumento de la temperatura promedio mundial.

No obstante años atrás se desarrolló el análisis de conjuntos de observaciones meteorológicas (temperatura y precipitación) representándolas mediante un modelo autorregresivo de orden 45 en donde a través de la serie temporal se encontraron con una tendencia secular de aproximadamente $0,6^{\circ}\text{C}$ durante el siglo pasado y un calentamiento definitivo en la década de 1980 y principios de la de 1990. Dicho análisis permitió caracterizar la homogeneidad, la tendencia y la persistencia, y simular las series temporales de los elementos climáticos más importantes en Lisboa (Leite & Peixoto, 1996).

Hay que destacar que los estudios sobre la persistencia son importantes para comprender la dinámica de los sistemas que crean series temporales. En climatología, el análisis de la función de autocorrelación ha sido la principal herramienta para investigar la persistencia de una serie temporal. Por esta razón, Triacca (2014) propone utilizar un modelo ARIMA como una herramienta para obtener una estimación de la persistencia en series temporales de temperatura media anual terrestre y oceánica y hemisférica global. Mediante los resultados obtenidos, mostro que las anomalías de la temperatura del océano parecen ser más persistentes que las de la tierra, además las anomalías de la temperatura para el hemisferio sur parecen ser más persistentes que las del hemisferio norte. De este modo, Triacca pudo interpretar dichos resultados estadísticos con un mecanismo físico concluyendo que, debido a la inercia térmica oceánica, las anomalías de temperatura pueden persistir durante varios años. La gran inercia térmica de los océanos permite una menor variabilidad interanual, y esto conduce a una mayor persistencia en las temperaturas oceánicas y del hemisferio sur.

1.2 Justificación

El cambio climático y la variabilidad climática son áreas de investigación que presentan oportunidades y retos muy importantes para el medio científico, ya que son fenómenos que se han registrado a escala global mediante el incremento de la temperatura (Wurl, Nayeli, García, & Imaz, 2017) o en el incremento o decremento de la precipitación debido a la interconexión entre las distintas variables involucradas y sus efectos se encuentran relacionados con la actividad humana. En este sentido, la aplicación de modelos autorregresivos en eventos climatológicos permite obtener información para identificar, estimar y diagnosticar la ocurrencia de los mismos.

En lo que respecta a Baja California, la ausencia de investigaciones que evalúen la variabilidad climática a escala local y la vulnerabilidad de la zona frente al cambio climático, apoya a la necesidad de analizar este fenómeno no solo en el ámbito global sino también a nivel local, así como el creciente interés por determinar si los ciclos climáticos naturales responden a patrones repetitivos predecibles de escala global a regional o, si es posible detectar componentes que permitan generar modelos de previsión de fenómenos extremos (Zotelo, 2011) de forma que los escenarios generados puedan emplearse en evaluaciones de impactos y estrategias de adaptación hacia la protección de los recursos naturales, así como estabilidad y seguridad a la población (Gonzalez Barrios & Descroix, 2012).

De esta manera, la aplicación de modelos autorregresivos a partir de información climatológica en Hidrología es una herramienta indispensable que permite analizar el comportamiento de las variables estudiadas y a su vez generar escenarios con la finalidad de predecir eventos a futuro y desarrollar programas de adaptación/mitigación

(López-Lambraño, Fuentes, López-Ramos, Pliego-Díaz, & López-L, 2016) que ayuden a contrarrestar las modificaciones negativas a consecuencia de este fenómeno, principalmente en la escasez de agua y la agricultura.

En este caso de estudio, la precipitación, la temperatura y la evaporación son el punto de partida, ya que son las variables meteorológicas de mayor relevancia porque mediante ellas es posible sintetizar el comportamiento del clima de una región, y permite realizar una comparación con el clima simulado por los modelos climáticos de manera global o regional (Pabón-Caicedo, Eslava-Ramírez, & Gómez-Torres, 2001).

Frente a esta situación, se requieren análisis y desarrollo de modelos a nivel local, con el objetivo de conocer a detalle la forma en la que la variabilidad climática se presenta en determinadas regiones y/o poblaciones.

1.3 Descripción del Problema

Mediante los distintos resultados obtenidos a través del estudio del cambio climático se han identificado distintas consecuencias generadas por éste, las cuales son cuantificadas respecto a las anomalías presentadas en distintas variables climáticas como la precipitación, la temperatura y la evaporación (debido al impacto de los gases efecto invernadero), así como también por la disponibilidad del agua y la disminución de la producción agrícola (Sánchez-Cohen et al., 2008). En relación a esto se ha encontrado que la actividad solar, la actividad volcánica y la concentración de gases efecto invernadero juegan un papel muy importante en la variabilidad climática mundial, siendo para el siglo XX la concentración de gases de efecto invernadero el factor principal (Mann, Bradley, & Hughes, 1998). Estos cambios se encuentran relacionados con cambios en otros factores climáticos como el ciclo hidrológico, los cuales han generado cambios en la precipitación, nubosidad, velocidad del viento y en la frecuencia de eventos climáticos extremos (Stocker et al., 2001), que son causados debido a la combinación de la variabilidad climática natural y a la que tiene su origen en actividades antropogénicas (Barnett et al., 1999).

Sin embargo, el problema no radica en que se estén presentando cambios, puesto que en el clima terrestre esto es natural, el problema real se debe a la rapidez con la que estos cambios se están presentando, sin dar tiempo a la adaptación de los seres vivos a los mismos. Según Sánchez-Cohen (2008) uno de los principales efectos negativos causados por el cambio climático es la incertidumbre climática, espacial y temporal definida como la incapacidad de predecir el comportamiento de las variables climáticas como la precipitación y la temperatura. Por lo tanto, el impacto de estos cambios en la

variabilidad de patrones de las diferentes variables climáticas es un riesgo que puede afectar de manera directa el entorno físico y el económico de cualquier población; de aquí la necesidad de aplicación de modelos que sean de utilidad para la predicción del comportamiento de las variables climáticas a futuro.

El uso de herramientas como los modelos autorregresivos permiten obtener información sobre el comportamiento de la precipitación, temperatura y evaporación a través del tiempo, así como sobre posibles eventos a futuro con el objetivo de identificar y de crear escenarios de cambio climático y variabilidad climática. Esta información permite a los organismos encargados (entidades públicas o privadas) de la gestión integral de los recursos la toma de decisiones con mayor conocimiento para desarrollar y emprender proyectos que aseguren el uso eficiente de los recursos.

En este caso el uso de modelos autorregresivos para la cuenca hidrográfica Guadalupe, B.C. no ha sido aplicado considerando la mayor cantidad de variables hidroclimáticas y meteorológicas posibles, que permitan predecir la ocurrencia de eventos climatológicos en dicha zona de estudio. Como consecuencia, existe incertidumbre respecto a escenarios de cambio climático y de la variabilidad climática en la zona, optando por decisiones basadas en información generada por modelos que solo contemplan una parte de la cuenca o por modelos incompletos que no permiten evaluar todas las consecuencias, lo que ha derivado en un impacto sobre la población. Así mismo, la falta de estrategias ante posibles eventos futuros y su impacto en la cuenca se ha visto reflejada en efectos negativos (escasez de agua en periodos de sequía, erosión, etc.) que han afectado a la población cercana, ya que los organismos encargados de gestionar la cuenca no cuentan con ninguna herramienta de planificación y análisis que permita modelar de manera fidedigna el comportamiento de las variables estudiadas.

De ahí la importancia de contar con modelos que permitan obtener información más cercana a la realidad, ya que según González (2003) actualmente uno de los retos para la ciencia se encuentra en poder distinguir el cambio climático provocado por actividades antropogénicas de la variabilidad climática natural, y predecir mediante distintos modelos escenarios climáticos futuros.

Por lo tanto, en este proyecto se desea obtener la modelación estocástica de la precipitación, temperatura y evaporación basado en los datos históricos de estas variables a través de la metodología Box & Jenkins, con el fin de determinar si existe variabilidad climática en la cuenca Guadalupe, B.C., México y permita ver las afectaciones y su impacto, cuando se tomen decisiones sobre el manejo de la cuenca

con el objetivo de propiciar estrategias de adaptación-mitigación así como generar posibles soluciones.

1.4 Objetivos

1.4.1 Objetivo general

Realizar modelación estocástica para el análisis de la variabilidad climática en la cuenca hidrográfica Guadalupe, B.C. México.

1.4.2 Objetivos específicos

- Elaborar un estado del arte sobre la aplicación de modelos autorregresivos en el estudio y comprensión del cambio climático.
- Estimar los gradientes de precipitación, temperatura y evaporación para determinar regiones de incremento y decremento dentro de la zona de estudio.
- Proponer modelos estocásticos que permitan la modelación espacio temporal de variables hidroclimatológicas.

1.5 Hipótesis

Los parámetros $\phi_{i,j}$ y $\theta_{i,j}$ presentan una correlación temporal con los parámetros de las series vecinas o contenidas dentro de una región hidrológicamente homogénea; que puede medirse con el correlograma de las series.

2 FUNDAMENTOS DE HIDROLOGÍA ESTOCÁSTICA

En este capítulo se introducen los conceptos fundamentales para el estudio de las series de tiempo, la estructura probabilística de las mismas, así como los momentos estadísticos de primer y segundo orden. Además, se definen algunos métodos de estimación de datos faltantes, criterio de selección de modelos y proceso de normalización de datos.

2.1 Introducción

A cualquier fenómeno relacionado con la lluvia y las consecuencias directas que ésta puede ocasionar son procesos aleatorios, esto se debe a que cuentan con un nivel de incertidumbre (componente aleatoria del evento) o una probabilidad de ser igualados o excedidos a través de los años (componente determinístico del evento), por lo cual se dice que se manejan por leyes de probabilidad y se les llama procesos estocásticos (García G., 2010). Por esta razón, la Hidrología estocástica aplica la teoría de los procesos estocásticos, al estudio de los eventos hidrológicos.

De esta manera, realizar el análisis hidrológico de una región o de un lugar determinado implica contar con una serie de datos relacionados con la variable de interés (temperatura, precipitación, evaporación, etc.), por lo tanto, es esencial la obtención de las medidas de tendencia central y las medidas de dispersión debido a que permiten describir, entender y explicar los fenómenos, además de poder cuantificarlo. A continuación, se presentan los conceptos fundamentales de dichas medidas.

2.2 Medidas de Tendencia Central

Las medidas de tendencia central ofrecen información cuyo propósito es resumir en un valor numérico una imagen representativa de la distribución de los datos. Además de brindar alguna medida cuantitativa de donde está el centro de los datos en una muestra (Vilalta Perdomo, 2016; Walpole, 2007).

De entrada, la media representa el primer momento con respecto al origen de la variable aleatoria, según Vilalta (2016) es simplemente un promedio numérico, es decir, se obtiene de la suma de todos los valores observados (x) dividida entre el número de observaciones (n). La media de una muestra de tamaño n , suponiendo que las observaciones en una muestra son X_1, X_2, \dots, X_n , está definida por:

$$\bar{x} = \sum_{i=1}^n \frac{X_i}{n} = \frac{X_1 + X_2 + \dots + X_n}{n} \quad (1)$$

De esta forma, la media es un punto de localización en una distribución que representa el valor más cercano u óptimo a todos los valores en tal distribución o conjunto de datos.

La mediana es el valor intermedio de todos los valores observados en la distribución ordenados en forma creciente o decreciente. Si el número de observaciones en una muestra es impar, la mediana se define por:

$$\tilde{x} = X_{(n+1)/2} \quad (2)$$

Si el número de observaciones es par, la mediana es el promedio de los dos valores centrales, es decir, divide las frecuencias de una distribución en dos partes iguales.

$$\tilde{x} = \frac{1}{2} \left(X_{\frac{n}{2}} + X_{\frac{n}{2}+1} \right) \quad (3)$$

El propósito de ésta, es reflejar la tendencia central de la muestra, de manera que no esté influida por los valores extremos. Para el análisis y solución de problemas hidrológicos es recomendable utilizar la mediana, debido a que es la medida de tendencia central que menos afectación por los valores extremos (Vilalta Perdomo, 2016; Walpole, 2007).

Finalmente, la moda es el valor más frecuente en una distribución de datos, la cual suma más información descriptiva a la media y mediana y refleja el pico en la distribución de los datos.

2.3 Medidas de Dispersión

Estas tienen por finalidad cuantificar la variabilidad de los datos, es decir, que tan separados o disímiles son uno de otro (Rustom Jabbaz, 2012). Las medidas de dispersión más utilizadas se describen a continuación:

La varianza es la media aritmética de los cuadrados de las desviaciones de cada observación con respecto a la media aritmética del conjunto de datos. Según Walpole (2007) la varianza de la muestra, denotada con S^2 , está dada por

$$S^2 = \sum_{i=1}^n \frac{(x_i - \bar{X})^2}{n-1} \quad (4)$$

Se llama desviación estándar a la raíz cuadrada positiva de la varianza respectiva, ésta dependerá del tipo de varianza que se esté empleando (Panteleeva, 2005). Por lo tanto, la desviación estándar de la muestra, denotada con s , está dada por

$$s = \sqrt{S^2} \quad (5)$$

Cuando se desea saber que tan grande o pequeña es la desviación respecto a su media estamos hablando del coeficiente de variación, el cual expresa la desviación estándar como porcentaje de la media aritmética, mostrando una mejor interpretación porcentual del grado de variabilidad que la desviación estándar. Definido como el cociente entre la desviación típica y la media

$$Cv = \frac{s}{\bar{X}} * 100 \quad (6)$$

El coeficiente de asimetría como su propio nombre nos indica, es una medida de la asimetría de la distribución de los valores respecto a la media, es decir, nos indica si existe el mismo número de elementos a la izquierda y a la derecha de la media (Pedroza & Dicovskyi, 2006). Se define mediante el estadístico:

$$Cs = \frac{n * \sum_{i=1}^n (x_i - \bar{X})^3}{(n-1)(n-2) * S^3} \quad (7)$$

El coeficiente de curtosis, es una medida de la concentración de la distribución en torno a la media. Si la variable sigue una distribución normal, su valor de curtosis será cero. Valores mayores a cero, indican que la distribución tiende a concentrarse en torno a la media más que a una distribución normal; mientras que valores menores a cero, indican que tiende a dispersarse más. Definido por la siguiente expresión:

$$Ck = \frac{n^2 * \sum_{i=1}^n (x_i - \bar{x})^4}{(n-1)(n-2)(n-3) * S^4} \quad (8)$$

2.4 Métodos de estimación de datos faltantes

Para la mayoría de los análisis hidrológicos como lo son la evaluación de los recursos hídricos, periodos hidrológicos y la variabilidad climática, entre muchos otros, son de suma importancia los registros históricos (Aly, Pathak, Teegavarapu, Ahlquist, & Fuelberg, 2009), sin embargo, rara vez se encuentran completos debido a errores en la lectura o daño en equipos de medición. Dado lo anterior, se presentan distintos métodos de estimación para llenado de datos faltantes en los registros de las variables.

2.4.1 Correlación

Este método es un análisis sencillo de la relación entre dos variables o dos series de datos. En el caso de variables como la precipitación se refiere a la relación que existe entre una estación climatológica y otra (alturas en diferentes sitios). Dicho análisis procura medir la fuerza de las relaciones entre dos variables por medio de un simple número llamado coeficiente de correlación y se calcula empleando la siguiente expresión:

$$y = \alpha + \beta x \quad (9)$$

Donde para obtener α y β :

$$\alpha = \frac{\sum y_i \sum x_i^2 - \sum x_i y_i \sum x_i}{n \sum x_i^2 - (\sum x_i)^2} \quad (10) \quad \text{y} \quad \beta = \frac{n \sum x_i y_i - \sum x_i \sum y_i}{n \sum x_i^2 - (\sum x_i)^2} \quad (11)$$

De manera que x_i es el valor de la variable x , y_i es el valor de la variable y , \bar{x} el valor medio de la variable x , \bar{y} el valor medio de la variable y , y n es el número total de valores.

2.4.2 Método del U.S. Weather Service

También conocido como el método del inverso de la distancia al cuadrado, se basa en una metodología fundamentada teórica y empíricamente, considerándose una de las mejores y más comunes para el completado de series de precipitación. Se define como:

$$P_x = \sum_{i=1}^n W_i P_i \quad (12)$$

Donde P_x es el dato faltante para el día, mes o año en la estación x , P_i es la precipitación en la estación i y W_i es el peso de la estación i , el cual se calcula empleando la siguiente expresión:

$$W_i = \frac{(\text{inv. dist. Eucl. } i)^2}{\sum_{i=1}^n (\text{inv. dist. Eucl.})^2} \quad (13)$$

2.4.3 Método de relación normalizada

Este método consiste en multiplicar la precipitación de cada estación por el cociente de los promedios de la serie de trabajo y la auxiliar, es considerado uno de los métodos más sencillos para completar series climáticas a nivel mensual y anual, y se define como:

$$P_x = \frac{1}{n} \left[\left(\frac{N_x}{N_A} \right) P_A + \left(\frac{N_x}{N_B} \right) P_B + \dots + \left(\frac{N_x}{N_n} \right) P_n \right] \quad (14)$$

Donde P_x es el dato perdido para el día, mes o año en la estación x , N_x es la precipitación media diaria, mensual o anual en la estación de dato faltante x , N_A, N_B, \dots, N_n es la precipitación media diaria, mensual o anual en las estaciones auxiliares (medias de todas las históricas) y P_A, P_B, \dots, P_n : precipitación mostrada en las estaciones auxiliares el día o mes o año en el que falta el dato en la estación x .

2.4.4 Método de regresión lineal simple

Este es el método de llenado más divulgado por su sencillez, en él se plantea que existe una relación lineal entre los datos de una estación y otra, de manera que se define por la siguiente expresión:

$$P_x = mP + b \quad (15)$$

Donde m y b son los parámetros correspondientes a la ecuación de la recta y P es la precipitación en la estación auxiliar.

2.5 Método de interpolación Distancia Inversa Ponderada (IDW)

El método IDW está basado básicamente en la inversa de la distancia elevada a una potencia matemática. El parámetro "potencia" permite controlar la significancia de puntos conocidos en los valores interpolados basándose en la distancia desde el punto de salida (Figura 1). Es un número real positivo y su valor predeterminado es 2.

Cuando se define un valor de potencia más alto, se pone mayor énfasis en los puntos más cercanos, por lo cual, los datos cercanos tendrán más influencia y la superficie tendrá más detalles (superficie menos suave). Conforme aumenta la potencia, los valores interpolados se acercan al valor del punto de muestra más cercano. Al definir un valor de potencia más bajo, los puntos circundantes adquieren mayor influencia que los que están más lejos, generando una superficie más suave.

En conclusión, el valor óptimo para la “potencia” es en donde el error absoluto medio mínimo se encuentra en su punto más bajo. En términos generales, este método determina los valores de celda a través de una combinación ponderada linealmente de un conjunto de puntos de muestra, dicha ponderación es una función de la distancia inversa, es importante considerar que la superficie que se interpola debe corresponder a la variable dependiente de la ubicación (Philip & Watson, 1982; Watson, 1985).

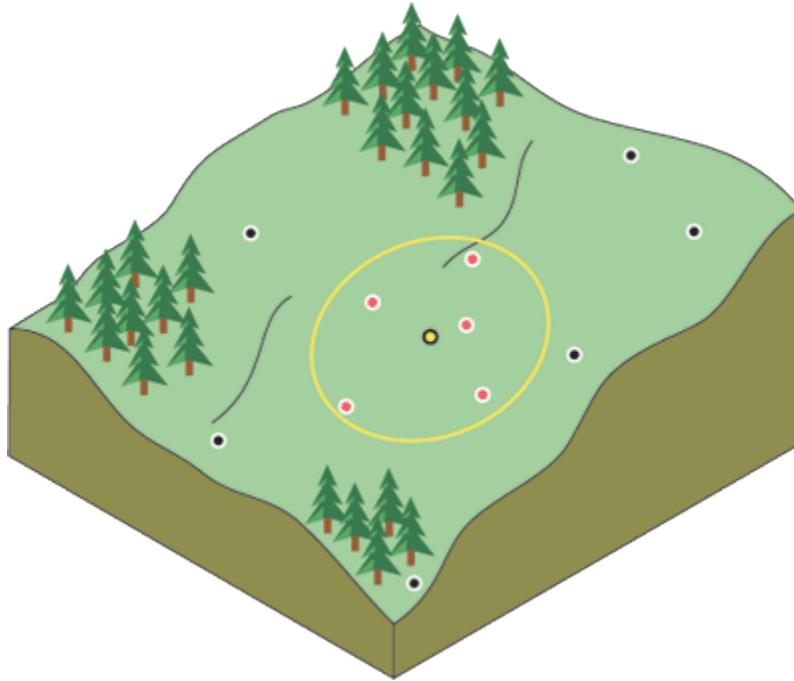


Figura 1.- Vecindad IDW del punto seleccionado. Fuente: (ArcGis, 2018)

El método IDW da más peso a los valores cercanos a un punto y asume que las predicciones son una combinación lineal de los datos, ofrece la ventaja de proyectar mapas o superficies continuas a través de datos, lo que lo convierte en uno de los más utilizados en estudios de análisis de la variación espacial (Villatoro, Henríquez, & Sancho, 2008), y se define como:

$$z(\mathbf{s}_0) = \sum_{i=1}^n \lambda_i z(\mathbf{s}_i) \quad i = 1, \dots, n \quad (16)$$

Donde $\hat{z}(\mathbf{s}_0)$ es el valor estimado en el punto interpolado \mathbf{s}_0 ; n es el número de observaciones vecinas usadas para la estimación y λ_i es el peso dado al valor observado $z(\mathbf{s}_i)$ en las cercanías del valor \mathbf{s}_0 .

2.6 Regresión Lineal

El estudio de los gradientes en series de tiempo hidrológicas es esencialmente importante ya que permite conocer, de forma general, la influencia que los distintos factores geográficos ejercen en el clima, la variabilidad climática de la zona, así como entender su relación con las alteraciones climáticas (Martín, 1995). Uno de los modelos más sencillos y comunes está basado en la suposición de que dos variables se relacionan en forma lineal, el objetivo de este modelo es tratar de explicar la relación que existe entre la variable de respuesta Y y una única variable explicativa X , es decir, estimar el valor de la variable dependiente a partir del valor de la otra (Carollo L., n.d.). Dicho modelo se representa mediante la siguiente expresión:

$$Y = \alpha + \beta X \quad (17)$$

En donde α es la ordenada en el origen (el valor que toma Y cuando X vale 0) y β es la pendiente de la recta (indica cómo cambia Y al incrementar X en una unidad. Y y X son variables aleatorias, por lo que no se puede establecer una relación lineal exacta entre ellas.

Si se grafica la ecuación (17), se logrará observar que se comete un error e_i en la predicción para el valor x_i , cuyo valor es

$$e_i = \hat{y}_i - y_i \quad (18)$$

Donde \hat{y}_i es el valor estimado de y . Para estimar los valores de α y β conviene que:

$$M = \sum \mathbf{e}_i^2 \quad (19)$$

Sea el mínimo posible. De esta manera se evita que los errores positivos y negativos se compensen. Para que M sea mínimo se debe cumplir:

$$\frac{\partial M}{\partial \alpha} = 0 \quad (20)$$

$$\frac{\partial M}{\partial \beta} = 0 \quad (21)$$

Sustituyendo las ecuaciones (17), (18) y (19) en las ecuaciones (20) y (21) se tiene

$$\frac{\partial M}{\partial \alpha} = 2 \sum [\alpha + \beta \mathbf{x}_i] = n\alpha + \beta \sum \mathbf{x}_i - \sum \mathbf{y}_i = 0 \quad (22)$$

$$\frac{\partial M}{\partial \beta} = 2 \sum \mathbf{x}_i [\alpha + \beta \mathbf{x}_i - \mathbf{y}_i] = \alpha \sum \mathbf{x}_i + \beta \sum \mathbf{x}_i^2 - \sum \mathbf{x}_i \mathbf{y}_i = 0 \quad (23)$$

Desarrollando las ecuaciones se obtiene

$$\alpha = \frac{\sum \mathbf{y}_i \sum \mathbf{x}_i^2 - \sum \mathbf{x}_i \mathbf{y}_i \sum \mathbf{x}_i}{n \sum \mathbf{x}_i^2 - (\sum \mathbf{x}_i)^2} \quad (24)$$

$$\beta = \frac{n \sum \mathbf{x}_i \mathbf{y}_i - \sum \mathbf{x}_i \sum \mathbf{y}_i}{n \sum \mathbf{x}_i^2 - (\sum \mathbf{x}_i)^2} \quad (25)$$

Cuando se desarrolle un ajuste por regresión como el anterior, conviene determinar la calidad de ajuste de la recta a los puntos medidos, es decir, que tanta dispersión existe

entre los puntos en torno a la recta. En este sentido se calcula el coeficiente de correlación mediante la siguiente expresión

$$R = \beta \frac{\sigma_x}{\sigma_y} \quad (26)$$

Donde σ es la desviación estándar de la variable especificada. Otra forma de calcularlo se conoce como el coeficiente de correlación de Pearson definido mediante la siguiente expresión:

$$R_{x,y} = \frac{\sigma_{xy}}{\sigma_x \sigma_y} = \frac{E\left[(X - \mu_x)(Y - \mu_y)\right]}{\sigma_x \sigma_y} \quad (27)$$

Donde σ_{xy} es la covarianza de (x,y) . Finalmente

Si $R = 1$, indica una dependencia total entre las dos variables, es decir, cuando una de ellas aumenta, la otra también lo hace en proporción constante, por lo tanto, existe una correlación positiva perfecta.

Si $0 < R < 1$, existe una correlación positiva.

Si $R = 0$, no existe relación lineal, sin embargo, esto no implica que las variables son independientes ya que pueden existir relaciones no lineales entre ellas.

Si $-1 < R < 0$, existe una correlación negativa.

Si $R = -1$, indica una dependencia total entre las dos variables denominada relación inversa, es decir, cuando una de ellas aumenta, la otra disminuye en proporción constante, por lo tanto, existe una correlación negativa perfecta.

En lo referente a R^2 o coeficiente de determinación, este tiene como propósito determinar la calidad del modelo, así como la proporción de variación de los resultados que pueden explicarse a través del modelo.

Por otra parte, cuando se habla de predecir variables aleatorias mediante su media, en donde el error cuadrático medio es su varianza, el máximo error cuadrático medio aceptable en un modelo para una variable aleatoria es la varianza. Así, la diferencia entre el valor observado y el valor predicho es llamado residuo. De esta manera, la media cuadrática de los residuos es la varianza residual.

Entonces, representando σ^2 como la varianza de la variable dependiente y σ_r^2 como la varianza residual, el coeficiente de determinación esta dado por la siguiente expresión:

$$R^2 = 1 - \frac{\sigma_r^2}{\sigma^2} \quad (28)$$

En términos generales, si la varianza residual es igual a cero, el modelo explica al 100% el valor de la variable; si por el contrario el valor de R^2 es del 0%, el modelo no explica nada.

2.7 Función de autocorrelación (ACF) y Función de autocorrelación parcial (PACF)

En lo que respecta a describir el comportamiento de procesos estacionarios y considerando que en ciertos casos los valores que toman las variables en una serie de tiempo no son independientes entre sí, sino que estos dependen de valores anteriores (Box et al., 1976), se hace necesario medir la dependencia de estas variables, en este sentido, se emplea la función de autocorrelación.

La ACF permite verificar el tipo de componente de una serie de tiempo observando el grafico de la serie original y el correlograma generado mediante dicha función (Murillos, Trejos, & Carvajal, 2003). En este caso, el propósito es identificar los modelos autorregresivos que describan de forma adecuada las variables estudiadas y para lograrlo el primer paso es determinar si las series son o no estacionarias, ya que muchas de las series no se pueden considerar estacionarias debido a que presentan tendencias muy marcadas y una dispersión creciente. En relación a lo anterior es esencial calcular la función de autocorrelación, la cual se define como

$$r_k = \frac{C_k}{C_0} = \frac{\frac{1}{N} \sum_{t=1}^{N-k} (x_t - x_{med})(x_{t+k} - x_{med})}{\frac{1}{N} \sum_{t=1}^N (x_t - x_{med})^2} \quad (29)$$

Donde C_0 es la varianza de la muestra, C_k es la autocovarianza de retraso k , k representa el tiempo de retraso o la distancia entre los pares (x_t, x_{t+k}) , x_{med} es la media de la muestra y N es el número de datos. La ACF tiene las siguientes propiedades:

- Es adimensional
- Solo toma valores $[-1,1]$
- Relación lineal perfecta entre dos variables $r = +1$
 $r = -1$

En cuanto a la función de autocorrelación parcial PACF, esta mide la correlación entre dos variables separadas por k periodos sin considerar la dependencia creada por los retardos intermedios existentes entre ambas. Se entiende que la PACF de una serie

temporal formada por: $\phi_{11}, \phi_{22}, \phi_{33}, \phi_{44}, \dots, \phi_{\tau\tau}, \dots$ en donde cada uno de los

valores de PACF por ejemplo el de orden τ , es decir, $\phi_{\tau\tau}$ se define como la

interrelación entre las variables Y_t y $Y_{t-\tau}$, eliminando los efectos lineales de las

variables: $Y_{t-1}, Y_{t-2}, Y_{t-3}, \dots, Y_{t-\tau+1}$, por lo tanto está dada por

$$\phi_{\tau\tau} = \text{corr}(Y_t, Y_{t-\tau} \cdot Y_{t-1}, Y_{t-2}, Y_{t-3}, \dots, Y_{t-\tau+1}) \quad (30)$$

Así pues, se entiende por PACF a los sucesivos coeficientes de correlación parcial de los distintos órdenes de una variable con ella misma desfasada diferentes ordenes o periodos.

2.8 Matriz de correlación

La matriz de coeficientes de correlación entre varias variables cuantitativas se representa de la siguiente manera:

	X	Y	Z
X	r_{xx}	r_{xy}	r_{xz}
Y	r_{yx}	r_{yy}	r_{yz}
Z	r_{zx}	r_{zy}	r_{zz}

Figura 2. Representación de una matriz de correlación de tres variables

Cada columna de la Figura 2 representa una variable y las filas son observaciones. A la medida del grado de asociación lineal entre dos variables cuantitativas se le denomina coeficiente de correlación, generalmente varía entre $[-1,1]$ y se define con la siguiente expresión:

$$\rho(A,B) = \frac{1}{N-1} \sum_{i=1}^N \left(\frac{A_i - \mu_A}{\sigma_A} \right) \left(\frac{B_i - \mu_B}{\sigma_B} \right) \quad (31)$$

Donde

μ = media

σ = desviación estándar

$$\rho(A,B) = \frac{\text{cov}(A,B)}{\sigma_A \sigma_B} \quad (32)$$

Para determinar el tipo de correlación entre una variable y otra se considera la siguiente clasificación:

Tabla 1 Clasificación del coeficiente de correlación

Rango		Tipo de Correlación
±0.00	→ ±0.09	Nula
±0.10	→ ±0.19	Muy débil
±0.20	→ ±0.49	Débil
±0.50	→ ±0.69	Moderada
±0.70	→ ±0.84	Significativa
±0.85	→ ±0.95	Fuerte
±0.96	→ ±1.00	Perfecta

2.9 Series de tiempo

Las series de tiempo son parte de un proceso estocástico, lo que las hace tan comunes e importantes es que permiten reconocer cualquier cambio del comportamiento típico de manera fácil. Al realizar un análisis de series de tiempo, el objetivo principal es desarrollar un modelo que brinde una descripción aceptable de los datos que conforman la serie y proporcionar un pronóstico de observaciones futuras que se puedan comparar con lo que realmente se observa, para finalmente tomar las medidas apropiadas (Chatfield, 2000; Janert, 2011). Por lo cual, es importante familiarizarse con los conceptos básicos.

En este sentido, según Chatfield (2000) una serie de tiempo es un conjunto de observaciones medidas secuencialmente a través del tiempo. Estas mediciones pueden realizarse ya sea de forma continua a lo largo del tiempo o en un conjunto discreto de puntos de tiempo, es decir, a intervalos iguales en momentos determinados de tiempo. Así mismo una serie puede ser univariante recogiendo la información de una sola variable, analizando su pasado a partir de la información que aporta en el tiempo o multivariante en donde los datos que la conforman constan de observaciones de varias variables (Scheidereiter & Faure, n.d.). En este caso las series analizadas son univariantes.

2.9.1 Componentes de las series de tiempo

Muchos análisis estadísticos de datos de series de tiempo hidrológicas, principalmente en estudios de planeación de los recursos del agua están basados en un conjunto de componentes fundamentales que son: homogeneidad, estacionariedad, tendencia, periodicidad y persistencia, mismos que deben ser verificados previamente en las series de tiempo para asegurar que cumplan con dichas condiciones y proceder a la identificación de modelos autorregresivos a evaluar (Komorník, Komorníková, Mesiar, Szökeová, & Szolgay, 2006; Machiwal & Jha, 2008). Por esta razón se describen de forma breve los conceptos de los componentes antes mencionados.

- *Tendencia*
Se dice que existe tendencia en un conjunto de datos si existe una correlación significativa entre las observaciones y el tiempo, y se puede definir como el componente a largo plazo que representa el crecimiento o disminución de una serie temporal durante un periodo de tiempo prolongado.
- *Estacionariedad*
La estacionariedad implica que los parámetros estadísticos (media, varianza, etc.) no cambian excepto debido a variaciones en las series, es decir, permanecen constantes a través del tiempo. Por el contrario, una serie es no estacionaria cuando los parámetros estadísticos cambian a lo largo del tiempo.
- *Homogeneidad*
Cuando se habla de que una serie es homogénea, nos referimos a que los datos en la serie pertenecen estadísticamente a la misma población y por lo tanto tienen una media de tiempo invariante.
- *Periodicidad*
En series hidrológicas se da generalmente debido a ciclos astronómicos (rotación de la tierra, etc.), por lo tanto, es un patrón de cambio más o menos estable que se da en los datos y que aparece de año en año, generalmente estas fluctuaciones se presentan series de datos inferiores al anual (trimestrales, mensuales o semanales).
- *Persistencia*
Esta se refiere a la tendencia para la magnitud de un evento que depende de la magnitud de eventos previos.

Vale la pena decir que cuando las series hidrológicas presentan tendencia en la media o en la varianza, o cuando ocurren cambios negativos o positivos, son anomalías producidas por la pérdida de homogeneidad e inconsistencia y afectan las características estadísticas de la serie, como la media, la desviación estándar y los coeficientes de correlación (Escalante-Sandoval & Reyes-Chávez, 2002).

2.9.2 Test para análisis de consistencia de datos

Existen distintos métodos para el análisis de consistencia de los datos, cuyo objetivo principal es asegurar la calidad de los mismos, y a su vez que cumplan las condiciones

descritas anteriormente. En el análisis que se hizo para asegurar la confiabilidad de los datos se emplearon los métodos descritos a continuación.

2.9.2.1 Test de Helmer

Esta prueba comprueba la homogeneidad de los datos observados en un procedimiento que consiste en ordenar la serie de datos cronológicamente y analizar el signo de las desviaciones de cada evento P_i^j de la serie j para $i = 1, 2, \dots, n_j$ con respecto a la media central \bar{P}^j . Si una desviación de cierto signo es seguida por otra del mismo signo, entonces se dice que se forma una secuencia S , de lo contrario se considera un cambio C . La serie se considera homogénea si se cumple el siguiente criterio:

$$-\sqrt{n-1} \leq (S-C) \leq \sqrt{n-1} \quad (33)$$

2.9.2.2 Test de Secuencias

Esta prueba consiste en analizar el signo de las desviaciones con respecto a la mediana muestral para cada dato y comparar el número de cambios de signo con el número de cambios permitidos con base al tamaño de la muestra. La serie se considera homogénea cuando el número de cambios está entre los valores establecidos de la serie.

2.9.2.3 Test de Mann-Kendall

Este test ha sido ampliamente utilizado en series hidrometeorológicas y otras series ambientales para la identificación de tendencia. Es un test no paramétrico, estadísticamente se basa en el número de orden de las variables, si una de esas variables es el tiempo y tiene la cualidad de presentar mayor potencia estadística en la medida que la serie de datos presenta un mayor sesgo que la aleje de una distribución normal. Su desarrollo consiste en lo siguiente:

1. Registrar los valores de las variables (temperatura, precipitación, etc.) de manera ordenada (X_1, X_2, \dots, X_n) .

2. Obtener el signo de la diferencia de cada par de valores al comparar sus magnitudes $(\mathbf{x}_j - \mathbf{x}_k)$ con $(j > k)$ de acuerdo a:

$$\text{signo}(\mathbf{x}_j - \mathbf{x}_k) = \begin{cases} 1 & \text{si } (\mathbf{x}_j - \mathbf{x}_k) > 0 \\ 0 & \text{si } (\mathbf{x}_j - \mathbf{x}_k) = 0 \\ -1 & \text{si } (\mathbf{x}_j - \mathbf{x}_k) < 0 \end{cases} \quad (34)$$

3. Obtener estadístico S de Mann-Kendall:

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^n \text{signo}(\mathbf{x}_j - \mathbf{x}_k) \quad (35)$$

Se dice que la serie cuenta con tendencia creciente si S es positivo, si S es negativo se dice que existe una tendencia decreciente.

4. Se estima la varianza para el estadístico S de Mann-Kendall, considerando el caso de empate $\text{signo}(\mathbf{x}_j - \mathbf{x}_k) = 0$ obtenidos en el paso 2, con la siguiente ecuación:

$$\text{Var}[S] = \frac{1}{18} \left[n(n-1)(2n+5) - \sum_{p=1}^q t_p(t_p-1)(t_p+5) \right] \quad (36)$$

5. Calcular estadístico Z_{MK} mediante:

$$Z = \begin{cases} \frac{S-1}{[\text{Var}(S)]^{1/2}} & \text{si } S > 0 \\ 0 & \text{si } S = 0 \\ \frac{S+1}{[\text{Var}(S)]^{1/2}} & \text{si } S < 0 \end{cases} \quad (37)$$

6. Evaluar la hipótesis de interés a partir del estadístico Z :

H_0 : no existe tendencia en las observaciones

H_1 : existe tendencia en las observaciones

2.9.2.4 Test t de Student

Si se considera una serie de datos P_i^j para $i = 1, 2, \dots, n_j$, del sitio j , la cual se divide en dos partes $n_1 = n_2 = \frac{n}{2}$, entonces el estadístico para la prueba se define mediante la expresión:

$$t_d = \frac{\bar{X}_1 - \bar{X}_2}{\left[\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2} \left(\frac{1}{n_1} + \frac{1}{n_2} \right) \right]^{0.5}} \quad (38)$$

Donde:

\bar{X}_1, s_1 : son la media y la varianza de la primera parte del registro tamaño n_1

\bar{X}_2, s_2 : son la media y la varianza de la primera parte del registro tamaño n_2

El valor absoluto de t_d se compara con el valor de la distribución t de Student de dos colas y con $v = n_1 + n_2 - 2$ grados de libertad para un nivel de significancia: $\alpha = 0.05$.

Si y solo si el valor absoluto de t_d es mayor que aquel de la distribución t de Student, se concluye que la diferencia entre las medias es evidencia de inconsistencia, y por ende la serie se considera no homogénea. En caso contrario la serie es considerada homogénea.

2.9.2.5 Test de Anderson

Para comprobar que una muestra P_i^j de la serie para $i = 1, 2, \dots, n_j$ está compuesta por variables aleatorias y determinar los límites de probabilidad de series independiente, se utiliza el test de independencia de Anderson (Haltiner & Salas, 1988), este hace uso del coeficiente de autocorrelación serial r_k^j para diferentes tiempos de retraso k . Si se analiza un solo registro $j = 1$.

El coeficiente de autocorrelación serial de retraso k se obtiene mediante la siguiente expresión:

Para $k = 1, 2, \dots, \frac{n_j}{3}$

$$r_k^j = \frac{\sum_{i=1}^{n_j-k} (P_i^j - \bar{P}^j)(P_{i+k}^j - \bar{P}^j)}{\sum_{i=1}^{n_j} (P_i^j - \bar{P}^j)^2} \quad (39)$$

Donde

$$\bar{P}^j = \sum_{i=1}^{n_j} \frac{P_i^j}{n_j} \quad (40)$$

El valor de los límites para un 95% de confianza para r_k^j se obtienen con:

$$r_k^j = \frac{-1 \pm 1.96 \sqrt{n_j - k - 1}}{n_j - k} \quad (41)$$

Si menos del 10% de los valores de autocorrelación r_k^j calculados superan los límites de confianza, se dice que la serie P_i^j es independiente y por lo tanto es una variable que sigue las leyes de la probabilidad. La grafica de los valores estimados de r_k^j (ordenadas) contra los tiempos de retraso k (abscisas), junto con los límites de confianza correspondientes se denomina correlograma de la muestra.

2.9.2.6 Test Dickey-Fuller *DF*

Al igual que los test anteriores, el propósito de este test es determinar el número de raíces unitarias de la serie, ya que la aplicación de la metodología de modelos autorregresivos requiere la utilización de series estacionarias en varianza y en media. Específicamente en el caso de los modelos ARIMA, cuando la serie no es estacionaria en la media, es decir, cuando no es integrada de orden cero $I(0)$, se dice que presenta al menos una raíz unitaria. Cuando esto ocurre, es posible obtener una serie estacionaria mediante la transformación de la serie original con una diferenciación adecuada. En este sentido, el número de diferencias que toma la serie para que sea estacionaria en la media está dada, por el número de raíces unitarias que presente la serie original, las cuales se determinan mediante los procedimientos de Dickey-Fuller *DF* (con su versión aumentada *ADF*).

Este test se basa en asumir que la serie se puede aproximar por un proceso $AR(1)$ con tres variantes: media cero, media diferente de cero y tendencia lineal. Inicialmente se asume que X_t sigue un modelo $AR(1)$ y se procede a transformar el modelo de la siguiente manera:

$$X_t = \phi_1 X_{t-1} + \varepsilon_t$$

$$X_t - X_{t-1} = (\phi_1 - 1) X_{t-1} + \varepsilon_t$$

$$\Delta X_t = \rho X_{t-1} + \varepsilon_t$$

Donde $\rho = \phi_1$. La existencia de una raíz unitaria equivale a $\phi_1 = 1$, es decir, a $\rho = 0$.

Caso 1: Test DF suponiendo que $\mathbf{X}_t \sim AR(1)$ con media cero, entonces

$$\Delta \mathbf{X}_t = \rho \mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t \quad (42)$$

La hipótesis nula es $H_0: \rho = 0$ versus la alterna $H_1: \rho < 0$. El estadístico de la prueba se denota por τ y su distribución bajo H_0 permite calcular los valores críticos, de tal forma que el criterio de rechazo es $\hat{\tau} < \tau_{0.05}$, con $\hat{\tau}$ es el valor calculado del estadístico.

Caso 2: Test DF suponiendo que $\mathbf{X}_t \sim AR(1)$ con media diferente de cero, entonces

$$\Delta \mathbf{X}_t = \alpha + \rho \mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t \quad (43)$$

Con la misma hipótesis de Caso 1.

Caso 3: Test DF suponiendo que $\mathbf{X}_t \sim AR(1)$ con tendencia lineal, entonces

$$\Delta \mathbf{X}_t = \alpha + \beta_t + \rho \mathbf{X}_{t-1} + \boldsymbol{\varepsilon}_t \quad (44)$$

Con la misma hipótesis de Caso 1 y 2.

2.9.2.7 Test Dickey-Fuller Aumentado (*ADF*)

El test Dickey-Fuller aumentado (*ADF*) requiere de una estrategia de análisis. Según Elder (2001), esta estrategia es necesaria para determinar si un intercepto, un intercepto más una tendencia con el tiempo, o ninguna de las dos anteriores debería incluirse al correr la regresión para la prueba de raíz unitaria. Considerar demasiados regresores puede generar una pérdida de potencia, por el contrario, no considerar los suficientes puede generar resultados sesgados hacia el no rechazo de la hipótesis nula. Considerar intercepto, o intercepto más tendencia, es necesario para permitir una representación de la hipótesis alterna que pueda competir contra la hipótesis nula de raíz unitaria.

En relación a lo anterior, se consideran las siguientes normas de carácter general. Si la serie original presenta tendencia, incluir ambos términos en la regresión, es decir, considerar como regresores al término independiente y al término de tendencia lineal. Si la serie no parece presentar tendencia y tiene un valor medio distinto de cero, incluir un término constante en la regresión. Finalmente, si la serie parece fluctuar en torno al valor medio cero, no se considera necesario incluir ningún regresor adicional en la regresión, es decir, no se incluyen ni el término constante ni el término de tendencia.

En lo que se refiere a la estrategia para el test (*ADF*), esta consiste en determinar cuál de los tres casos determina una mejor aproximación a la serie original, \mathbf{x}_t . Es importante destacar, que los casos dependen del orden autorregresivo p , y la decisión se toma considerando el valor menor del Criterio de Información de Akaike (*AIC*).

Caso 1: Suponiendo que $\mathbf{x}_t \sim AR(p)$ con media cero, entonces

$$\mathbf{x}_t = \sum_{j=1}^p \phi_j \mathbf{x}_{t-j} + \varepsilon_t = \phi_1 \mathbf{x}_{t-1} + \sum_{j=2}^p \phi_j \mathbf{x}_{t-j} + \varepsilon_t \quad (45)$$

Defina

$$\rho_1 = \sum_{j=1}^p \phi_j, \rho_i = -\sum_{j=1}^p \phi_j, i = 2, \dots, p$$

Entonces con $\rho = \rho_1 - 1$, la ecuación (45) se transforma en

$$\Delta \mathbf{X}_t = \rho \mathbf{X}_{t-1} + \sum_{j=2}^p \rho_j \Delta \mathbf{X}_{t-j+1} + \varepsilon_t \quad (46)$$

Si hay una raíz unitaria se cumple $1 - \phi_1 - \phi_2 - \dots - \phi_p = 0$, es decir, $\rho_1 = 1$. En este caso, se tiene que el modelo (46) equivale a

$$\begin{aligned} \mathbf{X}_t - \mathbf{X}_{t-1} &= \sum_{j=2}^p \rho_j (\mathbf{X}_{t-j+1} - \mathbf{X}_{t-j}) + \varepsilon_t \\ \Delta \mathbf{X}_t &= \sum_{j=1}^{p-1} \rho_{j+1} \Delta \mathbf{X}_{t-j} + \varepsilon_t \end{aligned} \quad (47)$$

Es decir, un $AR(p-1)$ en la variable $\mathbf{Z}_t = \Delta \mathbf{X}_t$.

La hipótesis nula es $H_0: \rho = 0$ y la alterna $H_1: \rho < 0$. El estadístico DF , $\hat{\tau}$ tiene la misma distribución asintótica que el estadístico DF para $\rho = 0$, ver caso 1 (42).

Caso 2: Suponiendo $\mathbf{X}_t \sim AR(p)$ con media diferente de cero. Con la misma notación de Caso 1,

$$\begin{aligned} \mathbf{X}_t - \mu &= \sum_{j=1}^p \phi_j (\mathbf{X}_{t-j} - \mu) + \varepsilon_t \\ \mathbf{X}_t &= \alpha + \phi_1 \mathbf{X}_{t-1} + \sum_{j=2}^p \rho_j (\mathbf{X}_{t-j+1} - \mathbf{X}_{t-j}) + \varepsilon_t \end{aligned}$$

$$\Delta \mathbf{X}_t = \alpha + \rho \mathbf{X}_{t-1} + \sum_{j=2}^p \rho_j \Delta \mathbf{X}_{t-j+1} + \varepsilon_t \quad (48)$$

Donde $\alpha = \mu \left(1 - \sum_{j=1}^p \phi_j \right)$, con los ρ_j definidos como en el caso anterior.

La hipótesis nula es $H_0: \rho = 0$ y la alterna $H_1: \rho < 0$. Bajo $H_0: \rho = 0$ el termino α se anula porque $\sum_{j=1}^p \phi_j = 1$. La distribución asintótica del estadístico DF es igual a la del caso $AR(1)$ con media, ver caso 2 (43).

Caso 3: Modelo $AR(p)$ con tendencia lineal, se define como

$$\mathbf{X}_t = a + bt + \sum_{j=1}^p \phi_j (\mathbf{X}_{t-j} - a - b(t-j)) + \varepsilon_t \quad (49)$$

Que se puede reordenar de la siguiente forma

$$\Delta \mathbf{X}_t = k_1 + k_2 t + \rho \mathbf{X}_{t-1} + \sum_{j=2}^p \rho_j \Delta \mathbf{X}_{t-j+1} + \varepsilon_t \quad (50)$$

Donde

$$k_1 = a \left(1 - \sum_{i=1}^p \phi_i \right) + b \sum_{i=1}^p i \phi_i$$

$$k_2 = b \left(1 - \sum_{i=1}^p \phi_i \right)$$

$$\rho_1 = \sum_{i=1}^p \phi_i \rho = \rho_1 - 1$$

La hipótesis nula es $H_0: \rho = 0$ y la alterna $H_1: \rho < 0$. Bajo $H_0: \rho = 0$ se tiene que $k_2 = 0$, $k_1 = b \sum_{i=1}^p i \phi_i$ y el estadístico DF tiene la misma distribución asintótica del estadístico en el caso $AR(1)$, ver caso 3 (44).

En conclusión, el test ADF corrige la correlación serial de orden elevado añadiendo más retardos del término diferenciado de la serie original en el lado derecho de la ecuación.

2.10 Modelos autorregresivos

Un modelo que describe la estructura de probabilidad de una secuencia de observaciones se denomina proceso estocástico. De esta forma, una serie temporal de N observaciones sucesivas se considera una muestra de una población infinita de muestras que puede ser representada mediante un modelo (Box et al., 1976). Dichos modelos se utilizan para explicar la estructura teórica temporal de la serie y predecir su comportamiento futuro a corto plazo, y se construyen a partir de la información que aportan la función de Autocorrelación (ACF) y la función de Autocorrelación Parcial (PACF). El proceso de modelación ARMA y ARIMA tiene como base los procesos autorregresivos AR y de media móvil MA. Debido a lo anterior, se introduce la definición de los modelos antes mencionados.

2.10.1 Modelo autorregresivo $AR(p)$

Un modelo estocástico que puede ser extremadamente útil en la representación de ciertas series que ocurren en la práctica es el modelo autorregresivo. En este modelo, el valor actual del proceso se expresa como un agregado finito y lineal de valores previos del proceso y un residual ε_t (Box et al., 1976). En estos modelos, la observación actual es la suma ponderada de una cantidad finita p de observaciones precedentes más un residual aleatorio independiente.

Se dice que un proceso $\{X_t\}$ es un proceso autorregresivo de orden p , es decir $(AR(p))$ si

$$X_t = \phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p} + \varepsilon_t \quad (51)$$

Donde

p = orden del modelo autorregresivo

x = operador de retraso (variables independientes o valores rezagados de la variable dependiente con rezagos de tiempo 1,2,3,..., p periodos).

ϕ = coeficiente autorregresivo

ε_t = residual (representa perturbaciones aleatorias que no pueden ser explicadas por el modelo).

X_t = variable aleatoria indexado por el tiempo (variable dependiente)

Por lo tanto, el valor en el tiempo t depende linealmente de los últimos valores p y el modelo se ve como un modelo de regresión, de ahí el término autoregresión.

2.10.2 Modelo de media móvil $MA(q)$

Otro tipo de modelo, de gran importancia práctica en la representación de series temporales observadas, es el proceso de media móvil finito. En este modelo, la observación actual se define como la suma del residual actual y de los residuales aleatorios anteriores con un determinado peso, esto se debe a que en este modelo una observación está condicionada por los residuales aleatorios de las observaciones anteriores. Por esta razón, la variable dependiente X_t depende de valores previos del residual más que de la variable misma.

Se dice que un proceso $\{X_t\}$ es un proceso de media móvil de orden q , es decir $(MA(q))$ si

$$X_t = \varepsilon_t + \theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \dots + \theta_q Z_{t-q} \quad (52)$$

Donde

q = orden de promedios móviles

z = operador de retraso (valores anteriores del residual o error)

θ_q = coeficiente de media móvil

ε_t = residual o error en el periodo t

X_t = variable aleatoria indexado por el tiempo (variable dependiente)

2.10.3 Modelo autorregresivo de media móvil ARMA (p, q)

Para lograr una mayor flexibilidad en el ajuste de series de tiempo reales, a veces es ventajoso incluir términos autorregresivos y de medias móviles en el modelo. Esto conduce al modelo autorregresivo de media móvil, que es la combinación de las estructuras del modelo AR(p) y MA(q), en donde una observación está determinada por observaciones anteriores y por residuales aleatorios de observaciones pasadas.

Se dice que un proceso estocástico $\{X_t\}$ sigue un modelo autorregresivo de media móvil de orden (p, q) si

$$X_t = \underbrace{\phi_1 X_{t-1} + \phi_2 X_{t-2} + \dots + \phi_p X_{t-p}}_{\text{Comp. autorregresiva}} + \varepsilon_t + \underbrace{\theta_1 Z_{t-1} + \theta_2 Z_{t-2} + \dots + \theta_q Z_{t-q}}_{\text{Comp. media móvil}} \quad (53)$$

Donde

p = orden del modelo autorregresivo (valores tomados por la variable durante los p periodos anteriores)

q = orden de promedios móviles (errores o residuos de la variable durante los q momentos anteriores)

x y z = operador de retraso

ϕ_p = coeficiente autorregresivo

θ_p = coeficiente de media móvil

ε_t = residual

X_t = variable aleatoria indexado por el tiempo (variable dependiente)

2.10.4 Modelo autorregresivo integrado de media móvil ARIMA(p, d, q)

Un modelo poderoso para describir series temporales estacionarias y no estacionarias es el modelo autorregresivo integrado de media móvil ARIMA, generalmente se construyen a partir de los modelos ARMA, pero este incorpora la diferenciación (integración) dentro del mismo (para que la serie se considere estacionaria en media), lo que representa una ventaja con respecto a los modelos ARMA (Alarcón Valverde & Chuñir Panjón, 2012).

Se dice que un proceso estocástico $\{X_t\}$ sigue un modelo autorregresivo integrado de media móvil de orden (p, d, q) si

$$X_t^d = \phi_1 X_{t-1}^d + \phi_2 X_{t-2}^d + \dots + \phi_p X_{t-p}^d + \varepsilon_t^d + \theta_1 Z_{t-1}^d + \theta_2 Z_{t-2}^d + \dots + \theta_q Z_{t-q}^d \quad (54)$$

Donde

p = orden del modelo autorregresivo (valores tomados por la variable durante los p periodos anteriores)

d = orden de integración (diferenciación) del modelo.

q = orden de promedios móviles (errores o residuos de la variable durante los q momentos anteriores)

x y z = operador de retraso

ϕ_p = coeficiente autorregresivo

θ_q = coeficiente de media móvil

ε_t^d = residual (serie de los errores que se cometen en la serie anterior)

X_t^d = variable aleatoria indexado por el tiempo (serie de las diferencias de orden d)

El proceso básico para la construcción de modelos autorregresivos consta de los siguientes pasos (Figura 3):

A. Identificación del modelo

Esta se logra sobre la base del patrón de los datos, el gráfico de la serie temporal, el AIC, su función de autocorrelación y función de autocorrelación parcial.

B. Estimación de parámetros

Después de identificar el modelo tentativo, los parámetros son estimados y probados para significancia estadística. Si los parámetros no cumplen la condición de estacionariedad, se debe identificar un nuevo modelo y se deben estimar y probar sus parámetros.

C. Diagnóstico

Después de encontrar el modelo correcto, debe ser diagnosticado. En el proceso de diagnóstico, la autocorrelación de los residuos del modelo estimado debe ser lo suficientemente pequeña y debe parecerse al ruido blanco. Si los residuos permanecen significativamente correlacionados entre ellos, se debe identificar y diagnosticar un nuevo modelo.

D. Pronóstico

Una vez que se selecciona el modelo, se utiliza para pronosticar la serie de la variable en estudio.

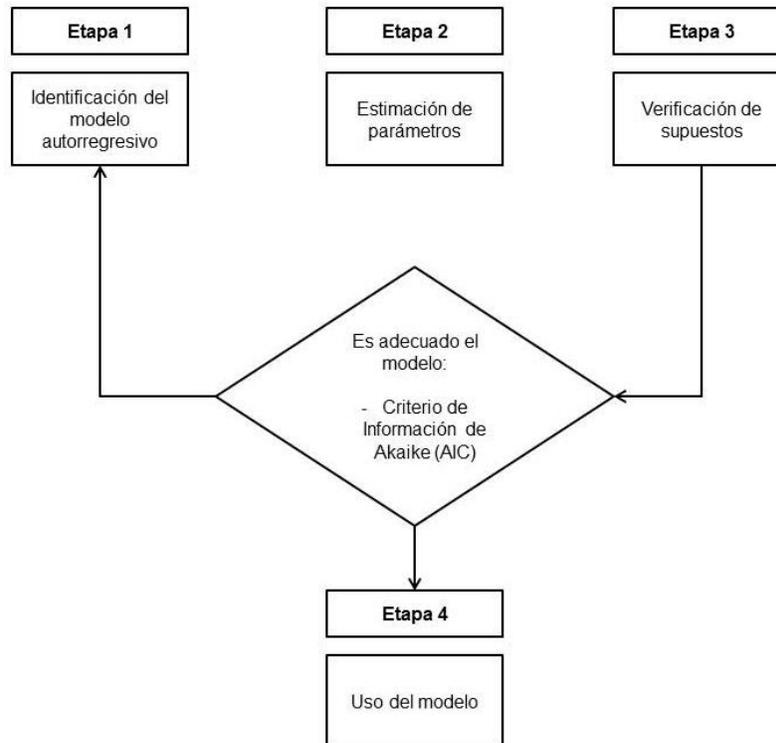


Figura 3.- Esquema de algoritmo para Modelo Autorregresivo

2.11 Criterio de Información de Akaike (AIC)

En el análisis de series de tiempo es común encontrar que más de un modelo logra representar adecuadamente la estructura del proceso generador de la serie. En este sentido, la selección del modelo que mejor represente la dinámica de los datos dentro de un conjunto de modelos propuestos y las características de las series hidrológicas se convierten en factores importantes. En este sentido, existen diversas técnicas y metodologías llamadas criterios de selección de modelos que permiten evaluar la calidad del modelo adecuado. Uno de los métodos más utilizados para seleccionar el modelo de mejor ajuste es un estadístico basado en las propiedades del modelo de máxima verosimilitud, denominado criterio de información de Akaike (AIC).

Considerando lo anterior, dada una distribución de probabilidad discreta (x_i, p_i) , donde x_i son los valores de la variable aleatoria y p_i las probabilidades; se define el concepto de entropía mediante la siguiente expresión:

$$H = -\sum p_i \ln p_i = -E\{\ln p_i\} \quad (55)$$

Donde, la función H es positiva y alcanza su máximo si $p_i = p_j$, siendo cero cuando para algún valor x , $p_i = 1$. Así pues, la entropía se considera una medida natural de la incertidumbre asociada a una distribución de probabilidad discreta.

Por el contrario, cuando se trata de una distribución de probabilidad continua, se define mediante:

$$H = -E\{Ln f(x)\} = -\int Ln f(x) f(x) dx \quad (56)$$

De este modo, la ecuación anterior muestra que la entropía, para algunas distribuciones puede ser negativa y para otras, infinita. En el caso de probabilidad continua, un inconveniente es que la entropía no es invariante respecto a transformaciones monotonas de la variable, como en el caso de probabilidad discreta. Así pues, la variación de información se define al pasar de la distribución inicial a la final, considerando la noción de información como idea central (Kullback & Leibler, 1951). Por lo tanto, la variación de información al pasar de f a g se define mediante:

$$I(g;f) = \int g(x) \ln \frac{g(x)}{f(x)} dx \quad (57)$$

Donde según Kullback (1951) al pasar de la distribución f a la g es conocida como la cantidad de información y es positiva mientras ambas distribuciones no coincidan en casi todo punto.

Partiendo de la expresión anterior, el criterio propuesto por Akaike (1974), supone que $g(x)$ es la distribución verdadera de una variable y siendo $f\left(\frac{x}{\theta}\right)$ una familia paramétrica de modelos, θ representa un vector de parámetros. La idea es seleccionar el modelo que minimiza la cantidad de información requerida para pasar de $f\left(\frac{x}{\theta}\right)$ a $g(x)$, es decir, el modelo más próximo al modelo teórico. Considerando que:

$$I\left(g(\bullet); f\left(\frac{\bullet}{\theta}\right)\right) = \int g(x) \text{Ln}g(x) dx - \int g(x) \text{Ln}f\left(\frac{x}{\theta}\right) dx \quad (58)$$

Minimizar $I(g : f)$ es equivalente a maximizar la segunda integral, que es la esperanza de $\text{Ln}f\left(\frac{x}{\theta}\right)$. Donde un estimador consistente de esta cantidad en su media muestral este dado por:

$$\frac{1}{N} \sum \text{Ln}f\left(\frac{x}{\theta}\right) \quad (59)$$

El fin del método de máxima verosimilitud es minimizar dicha expresión, lo que lleva a minimizar $(g : f)$. Por lo tanto, en base al resultado obtenido se puede considerar el método de máxima verosimilitud como un criterio natural de ajuste de un modelo. Ahora bien, tomando en cuenta el sesgo del estimador anterior, minimizar la integral lleva a escoger el modelo con un valor menor de:

$$AIC = -2 \left(\sum_{j=1}^N \text{Ln}f\left(\frac{\mathbf{X}_j}{\hat{\theta}}\right) \right) + 2K \quad (60)$$

Donde K representa el número de parámetros estimados, dimensión del vector θ y $\hat{\theta}$ es el estimador de máxima verosimilitud, conocido como MAIC.

La aplicación de este criterio a un modelo ARMA es inmediato y según Box (1970) la función de verosimilitud está dada por:

$$\text{Ln}(\theta, \sigma_a^2) = f(\theta) - \frac{N}{2} \text{Ln} \sigma_a^2 - \frac{\sum \mathbf{a}^2}{2\sigma_a^2} \quad (61)$$

Donde θ es el vector de parámetros del modelo y $f(\theta)$ es despreciable cuando el tamaño de la muestra es grande. Esto es:

$$AIC = -2 \left(-\frac{N}{2} \ln \sigma_a^2 - \frac{\sum a_t^2}{2\sigma_a^2} \right) + 2K \quad (62)$$

$$AIC = N \ln \sigma_a^2 + N + 2K \quad (63)$$

Donde N representa el número total de datos empleados en la estimación de los parámetros del modelo ARMA, y K representa el número de parámetros estimados que será, para un modelo estacional igual a la suma de los órdenes de los operadores más uno (debido a σ_a^2), más uno (solamente cuando la media de la serie estacionaria sea distinta de cero).

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

La “racionalidad” para la selección de modelos mediante este criterio depende de su eficacia empírica, esto se debe a que el establecimiento de este criterio está basado en un procedimiento heurístico (Peña & Arnáiz Tovar, 1981).

2.12 Principio de Parsimonia

El principio general del modelado de datos parsimoniosos establece que si dos modelos modelan adecuadamente un conjunto dado de datos, el que se describe con un número menor de parámetros tendrá una mejor capacidad predictiva dados los nuevos datos (Seasholtz & Kowalski, 1993).

2.12.1 Teoría del principio de Parsimonia

El concepto de parsimonia se puede considerar en un contexto estadístico limitado en problemas de regresión lineal múltiple con selección de variables. Para un problema de regresión

$$c = Rb + e \quad (64)$$

donde c es un vector que contiene la propiedad física o química de interés para una serie de muestras, R es la matriz de respuesta que contiene λ medidas para cada una de las muestras, b es un vector de parámetros de regresión y e es vector de errores.

Finalmente, el principio de parsimonia sugiere que el número de factores comunes debe ser considerablemente menor que el de variables, es decir, lo bastante reducido para ser interpretable y bastante completo para dar cuenta de las relaciones entre variables (Ferrando & Anguiano-Carrasco, 2010).

2.13 Estadístico Durbin-Watson (DW)

El estadístico Durbin-Watson DW mide la asociación lineal entre residuos adyacentes de un modelo de regresión, es decir, el grado de autocorrelación entre el residuo correspondiente a cada observación y la anterior, por lo tanto, determina si la correlación entre los residuos es o no igual a cero. La autocorrelación significa que los errores de las observaciones adyacentes están correlacionados, lo que puede generar que los predictores parezcan significativos cuando en realidad no lo son (Durbin & Watson, 1951; Savin & White, 1977). El DW es una prueba de la hipótesis en la especificación:

$$U_t = \rho U_{t-1} + \varepsilon_t \quad (65)$$

Si no hay una correlación serial, el estadístico DW será alrededor de 2. El estadístico DW caerá por debajo de 2 si hay una correlación serial positiva (el peor escenario, será cerca de cero). Si hay una correlación negativa, la estadística se ubicará entre 2 y 4.

La correlación serial positiva es la forma más común de dependencia observada. Como regla general, con 50 o más observaciones y solo unas pocas variables

independientes, un estadístico DW por debajo de 1.5 es una fuerte indicación de correlación serial positiva de primer orden (Dinardo, Johnston, & Johnston, 1997).

Limitaciones principales de la prueba DW como prueba para la correlación serial:

- 1) La distribución del estadístico DW bajo la hipótesis nula depende de la matriz de datos. El enfoque habitual para manejar este problema es colocar límites en la región crítica, creando una región donde los resultados de la prueba no son concluyentes.
- 2) Si hay variables dependientes rezagadas en el lado derecho de la regresión, la prueba DW ya no es válida.
- 3) Solo puede probar la hipótesis nula de ninguna correlación serial contra la hipótesis alternativa de la correlación serial de primer orden.

2.14 Transformación Box-Cox

Antes de realizar el análisis principal en series de tiempo, es necesario decidir si una variable necesita ser transformada, en el caso de que exista una tendencia ascendente en el tiempo, o que la varianza parezca estar aumentando con el nivel medio, entonces se puede pensar que se requiere de una transformación para estabilizar la varianza. También se puede indicar una transformación para hacer que los datos se distribuyan más normalmente si las observaciones parecen sesgadas (por ejemplo, con más 'picos' hacia arriba que hacia abajo). De igual forma, se puede pensar en transformar los datos cuando el efecto estacional parezca ser multiplicativo, ya que los efectos lineales son más fáciles de manejar (Chatfield, 2000).

Generalmente, se utilizan dos tipos de transformaciones, una para estabilizar la dispersión conocida como transformación de Box-Cox y otra para inducir estacionariedad en la media. Estas transformaciones son empleadas para corregir sesgos en la distribución de errores, para corregir varianzas desiguales y principalmente para corregir la no linealidad (mejorar la correlación entre las variables). La transformación de Box-Cox (1ra. Potencia), dada una serie de tiempo observada $\{x_i\}$, la serie transformada se define mediante la siguiente expresión:

$$W_i = \begin{cases} \frac{(x_i^\lambda - 1)}{\lambda} & \lambda \neq 0, \quad x_i > 0 \\ \ln(x_i) & \lambda = 0, \quad x_i > 0 \end{cases} \quad (66)$$

Donde:

X_i = serie a normalizar

W_i = serie transformada de primera potencia

λ = parámetros de transformación de potencia

El valor de λ que normaliza a la muestra es encontrado al tanteo hasta que el coeficiente de asimetría de la serie pueda suponerse estadísticamente igual a cero.

La transformación Box-Cox de primera potencia nos garantiza el coeficiente de asimetría (λ normaliza el coeficiente de asimetría ($g \approx 0$) pero no el de curtosis), por lo tanto, la corrección de la curtosis se puede lograr utilizando la transformación Box-Cox de segunda potencia, la cual se define mediante la siguiente expresión:

$$t_i = (|W_i - \bar{W}|)^\tau \quad (67)$$

Donde:

t_i = serie transformada 2da. Potencia

\bar{w} = valores medios de W_i

W_i = serie normalizada mediante 1ra. Potencia

τ = parámetro de transformación de 2da. Potencia

Los valores de τ estarán dentro de un rango: $0 < \tau < 1$ si la curtosis de $W_i > 3$, si es menor a 3 ($W_i < 3$) entonces τ tiende a 1, $\tau > 1$ si la curtosis (de los valores transformados en 1ra. Potencia) $W_i < 3$.

Hay que destacar que, las transformaciones deben ser empleadas solo cuando sea necesario o cuando la asimetría y los cambios en la varianza sean severos, caso contrario, siempre que sea posible se debe trabajar con los datos observados.

3 METODOLOGÍA

A lo largo de este capítulo se presenta la metodología aplicada en la presente investigación que implica desde la recopilación de información hasta la generación de algebra de mapas, así mismo se describe detalladamente cada una de las etapas (Figura 4) ejecutadas para la obtención de los resultados.

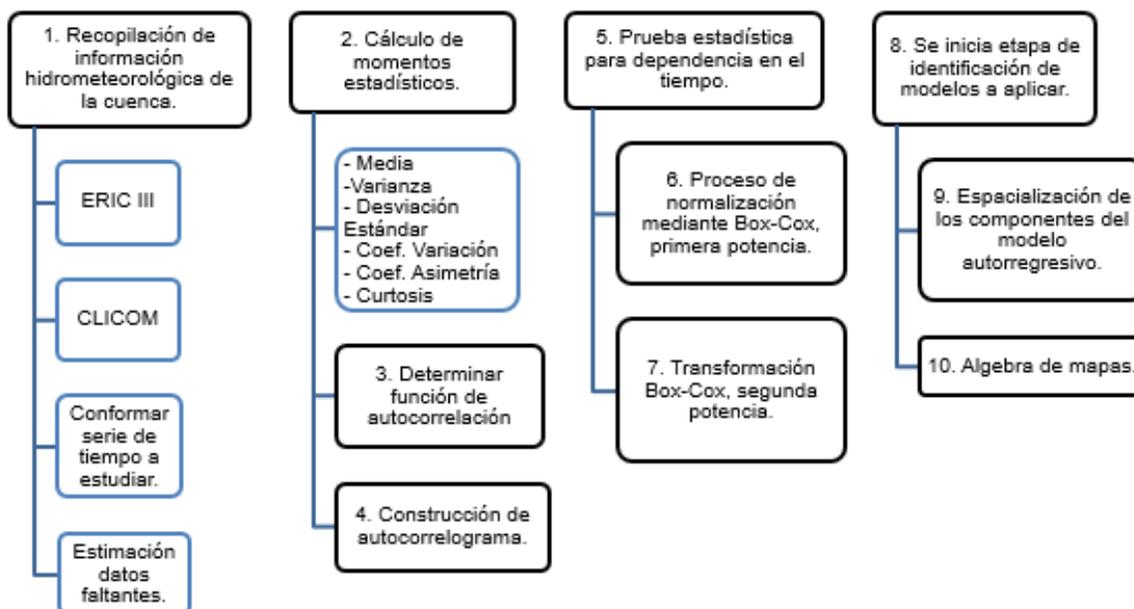


Figura 4. Etapas de metodología empleada en la investigación

3.1 Características del área de estudio

La cuenca Guadalupe se ubica al noroeste del estado de Baja California, en México. Conformada por tres sub-cuencas, forma parte de la Región Hidrológica No. 1 (RH 1) de la vertiente del Océano Pacífico. En ella se localizan tres valles, Ojos Negros, Guadalupe y La Misión (

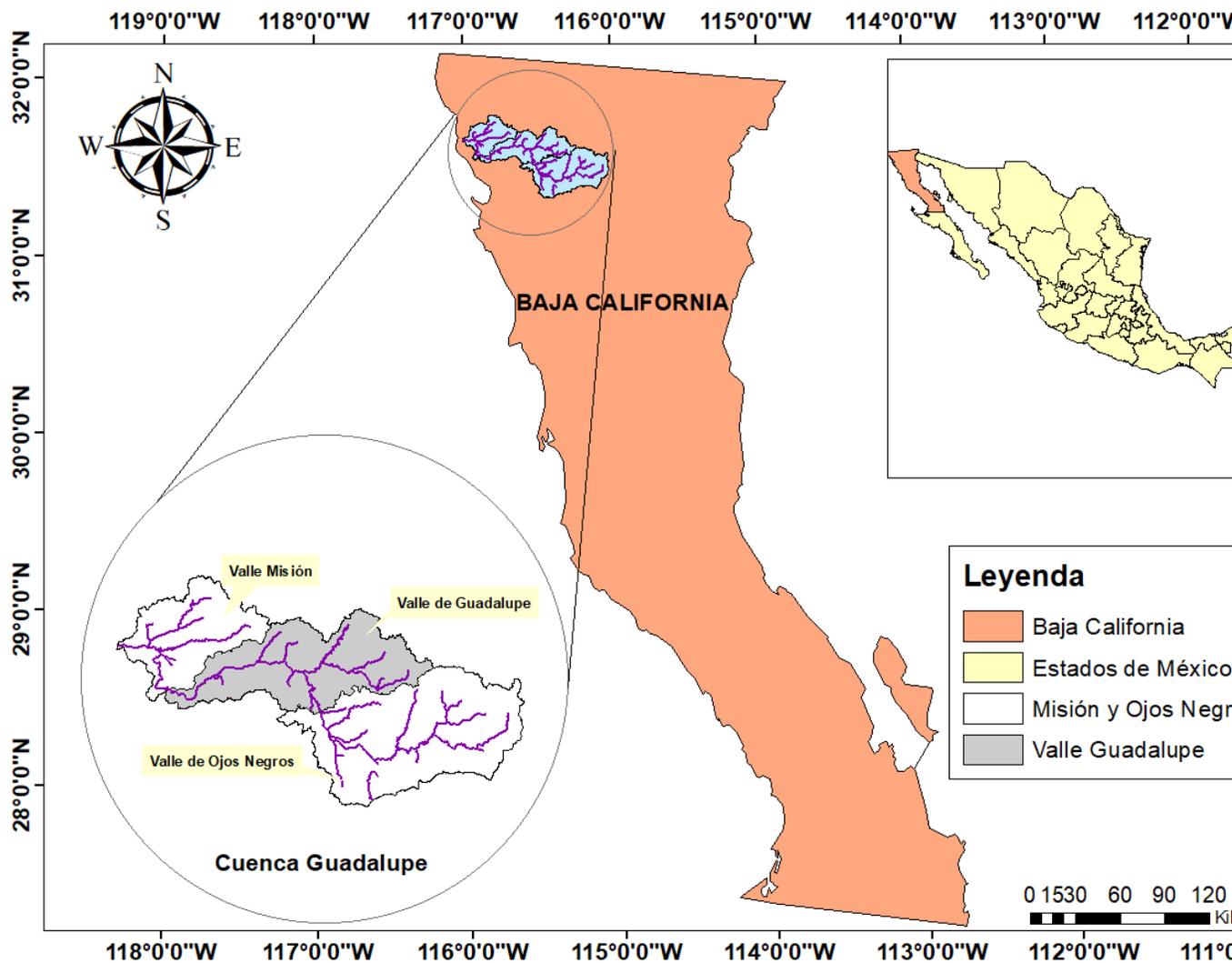


Figura 5), lo que la posiciona en una de las áreas agrícolas más productivas de la región y de gran importancia para la economía local.

Es una cuenca tipo exorreica, la cual cuenta con un área de aproximadamente 2,572.18 km² que inicia en la vertiente Occidental de la Sierra de Juárez a una altitud de 1700 msnm, cuyo clima se caracteriza por su variación entre mediterráneo y árido.

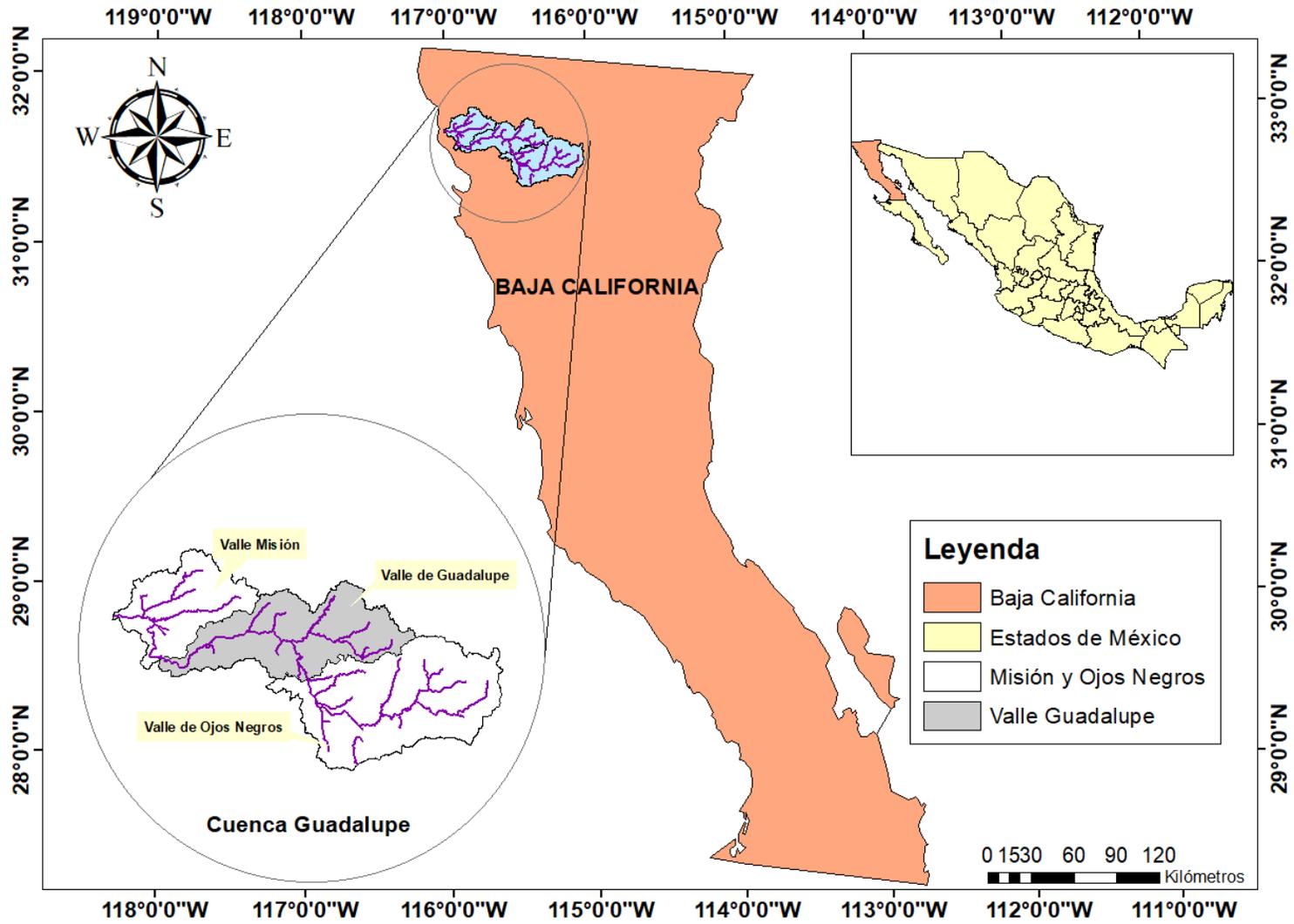


Figura 5. Localización del área de estudio

3.2 Recopilación de información hidrometeorológica

La presente investigación se enfoca en el análisis de las variables de precipitación, temperatura y evaporación para determinar la variabilidad climática en la cuenca hidrográfica Guadalupe B.C., por esta razón, el primer paso es identificar las estaciones climatológicas ubicadas dentro de la zona de estudio y sus alrededores (Tabla 2) para posteriormente iniciar con la primera etapa de la metodología (Figura 4).

Tabla 2 Estaciones climatológicas en el área de estudio y promedios anuales de variables de interés.

Clave	Nombre Estación	Long. (W)	Latitud	Altitud m.s.n.m	Promedio anual Prec. Mensual Tot. (mm)	Promedio anual Temp. Mensual media (°C)	Promedio anual Evap. Mensual Tot. (mm)
2001	Agua Caliente	116.454	32.107	400	272.83	12.68	2102.92
2004	Ignacio Zaragoza-Belen	116.486	32.195	540	301.72	12.48	1329.71
2005	Boquilla Sta. Rosa de la Misión	116.777	32.022	250	266.00	12.98	1228.95
2010	Compostela	115.767	31.783	625	196.05	14.51	1879.15
2014	El Alamo	116.054	31.593	1115	269.96	13.66	1687.62
2019	El Compadre	116.254	32.338	1110	302.29	18.84	-
2021	El Pinal	116.292	32.183	1320	460.84	9.40	-
2024	El Testerazo	116.534	32.296	380	263.35	11.31	1499.58
2025	Ensenada (Obs)	116.606	31.858	21	251.52	13.37	1583.90
2031	La Rumorosa	116.046	32.549	1232	169.65	14.33	-
2035	Ojos Negros	116.265	31.912	680	247.72	12.80	1463.39
2036	Olivares Mexicanos	116.681	32.049	340	288.85	14.27	1490.86
2042	Rosarito	117.002	32.288	150	121.78	17.16	-
2045	San Carlos	116.464	31.785	164	291.45	14.34	1459.35
2049	San Juan de Dios Norte	116.165	32.132	1280	375.30	12.02	1262.67
2057	Sta. Catarina Norte	115.824	31.657	1150	251.13	14.54	1710.34
2062	Sta. Isabel	115.817	31.871	1540	338.92	13.61	-
2065	Santo Tomas	116.406	31.792	180	261.04	16.36	1673.34
2066	Sierra de Juárez	115.948	32.004	1580	359.40	9.65	1261.44
2069	Valle de las Palmas	116.654	32.37	280	216.02	12.77	1797.33
2077	La Misión	116.811	32.102	20	289.73	13.25	1525.67
2079	El Alamar	116.204	31.836	710	268.84	12.92	1618.29
2088	Ej. Héroes de la Independencia	115.938	31.61	1000	257.19	14.19	-
2093	Ej. Valle de la Trinidad	115.749	32.356	780	228.35	11.60	2020.70
2094	El Farito	116.667	31.983	250	318.73	16.35	1075.07
2097	Laguna Salada	115.691	32.224	25	118.42	25.73	2012.99
2104	El Ciprés	116.588	31.790	8	230.63	14.80	1288.29
2106	Manadero	116.573	31.696	50	241.34	14.85	919.67
2108	Punta Banda	116.666	31.714	15	256.02	14.79	1490.28
2114	Ej. Carmen Serdán	116.584	32.244	560	265.40	10.98	104.85
2118	Valle de San Rafael	116.234	31.919	721	173.29	25.05	1400.24
2119	Cancio	116.600	32.31	450	298.61	12.17	1608.47
2122	Real del Castillo Viejo	116.746	31.951	610	260.71	11.47	838.47
2152	Ej. José María Pino Suarez	116.068	32.373	1380	278.43	10.03	136.85
2164	Ej. El Porvenir	115.852	32.106	330	256.47	13.07	1709.94

Una vez seleccionadas las estaciones climatológicas, se procedió a la extracción de los registros de las series antes mencionadas, considerando todas aquellas estaciones dentro de la zona de estudio que poseen al menos 25 años de registro aportando información suficiente para evaluar las características de las mismas. La recopilación de la información se logró con la ayuda de las siguientes bases de datos:

- a. Extractor rápido de información climatológica (ERIC III Versión 3.2) del Instituto Mexicano de Tecnología del Agua.
- b. Datos climáticos mensuales del CLICOM (Clima Computarizado) del Sistema Meteorológico Nacional del Centro de Investigación Científica y de Educación Superior de Ensenada (CICESE).

Específicamente se obtuvieron las series de las siguientes variables:

- 1) Precipitación Total Mensual
- 2) Temperatura Media Mensual
- 3) Evaporación Total Mensual

3.3 Conformación de Series de Tiempo

A partir de los registros obtenidos, se procedió a transferir los datos de cada una de las estaciones a documentos con formato Excel (un documento para cada una de las estaciones) para conformar las series de tiempo de las variables a estudiar. Esta conformación es la base para el análisis de consistencia de las series cuyo fin es asegurar que los datos cumplan y se basen en el conjunto de requerimientos estadísticos fundamentales que se deben considerar (Francisco, 2015; Machiwal & Jha, 2008).

3.4 Estimación de Datos Faltantes

Ahora bien, debido al porcentaje de datos faltantes observado en las series (Tabla 3) y con el propósito de cumplir con el enfoque de la investigación, se requiere de series de datos históricos completos y amplios para que las estimaciones estadísticas sean confiables y exactas, por lo anterior, se procedió a realizar la estimación de datos faltantes con el objetivo de completarlas y mejorar la estimación de los parámetros.

Tabla 3 Porcentaje de datos faltantes para series de precipitación, temperatura y evaporación.

Clave	Nombre Estación	Datos Faltantes Precipitación (%)	Datos Faltantes Temperatura (%)	Datos Faltantes Evaporación (%)
2001	Agua Caliente	0	0.75	2.381
2004	Ignacio Zaragoza-Belen	33.3333	15.9722	70.2703
2005	Boquilla Sta. Rosa de la Misión	20.6897	20.2777	46.6667
2010	Compostela	15.3846	0.59	0
2014	El Alamo	2.2222	2.6595	0
2019	El Compadre	38.4615	7.0512	-
2021	El Pinal	16.6667	20	-
2024	El Testero	40	2.7435	-
2025	Ensenada (Obs)	0	0	0
2031	La Rumorosa	40.7407	30.2511	-
2035	Ojos Negros	26.1905	1.5503	0
2036	Olivares Mexicanos	37.6623	8.3333	2.8571
2042	Rosarito	18.75	6.9444	-
2045	San Carlos	26.1905	27.1241	37.2549
2049	San Juan de Dios Norte	0	6.6666	23.0769
2057	Sta. Catarina Norte	0	5.9523	6.25
2062	Sta. Isabel	20	6.6666	-
2065	Santo Tomas	1.5385	6.2820	21.7391
2066	Sierra de Juárez	4	5.7692	31.25
2069	Valle de las Palmas	15.873	1.3227	0
2077	La Misión	0	0	0
2079	El Alamar	3.2258	3.125	0
2088	Ej. Héroes de la Independencia	38.4615	22.0085	-
2093	Ej. Valle de la Trinidad	0	2.5641	23.6842
2094	El Farito	0	33.3333	-
2097	Laguna Salada	25	8.8235	27.2727
2104	El Ciprés	33.3333	5.7142	42.4242
2106	Manadero	14.2857	5.8823	44.4444
2108	Punta Banda	14.2857	2.5462	0
2114	Ej. Carmen Serdán	23.3333	18.1818	-
2118	Valle de San Rafael	45.4545	26.2254	15.1515
2119	Cancio	20	14.7058	6.6667
2122	Real del Castillo Viejo	9.0909	10.6060	-
2152	Ej. José María Pino Suarez	38.7097	19.8924	-
2164	Ej. El Porvenir	0	4.7619	0

Para esto, se empleó el Sistema de Información Hidrológico para Intensidades Máximas SIHIMAX (programa informático diseñado para la búsqueda de información climatológica, consistencia, estimación de datos faltantes, funciones de distribución, así como obtención y tránsito de hidrogramas, etc.) desarrollado para la elaboración de estudios hidrológicos en la República Mexicana (Figura 6).

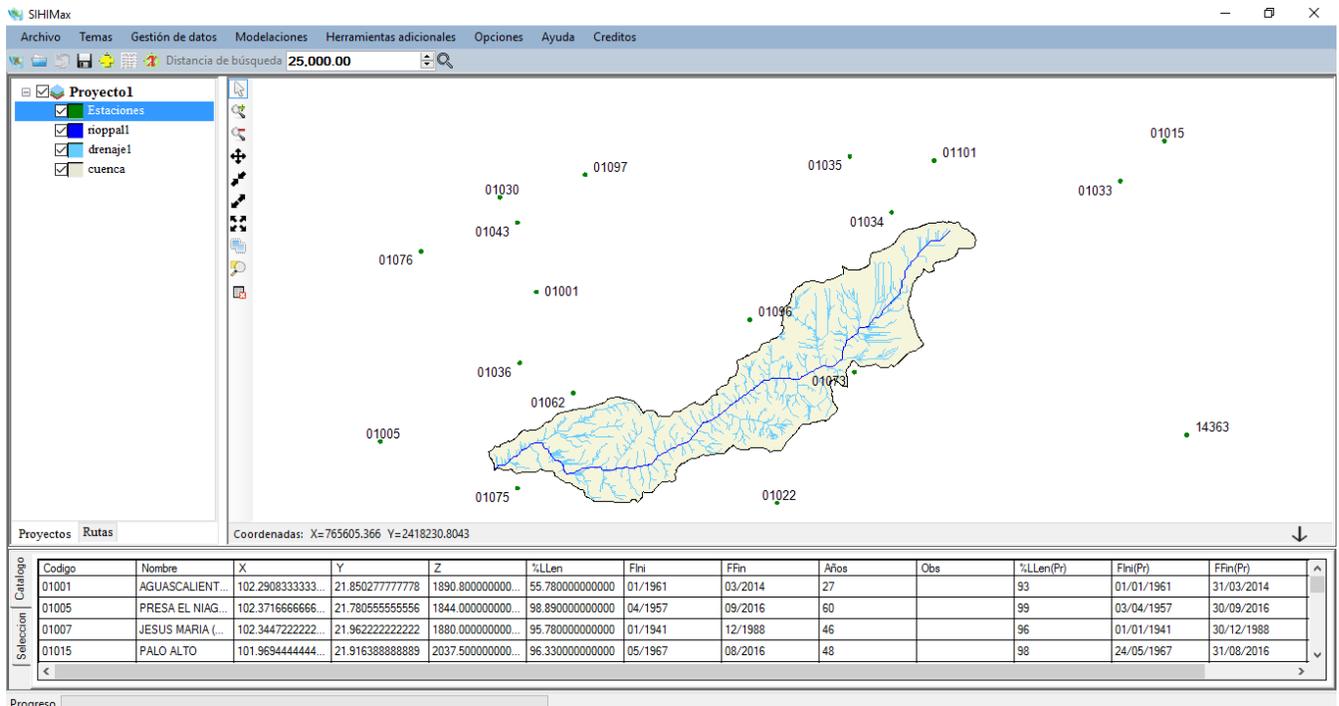


Figura 6. Interfaz del software SIHIMAX v1.56 de Hydrobits

La estimación de datos faltantes mediante SIHIMAX (Figura 7), se realizó bajo las siguientes metodologías:

- *Método del U.S. Weather Service* (ecuación (12)) debido a su capacidad para reproducir mediciones precisas, además de ser eficiente y proveer información fiable.
- *Método de relación normalizada* (ecuación (14)) que se caracteriza por su capacidad para estimar datos faltantes anuales y mensuales, el cual obtiene información de registros faltantes mediante los registros ya existentes y es ampliamente utilizado para el diseño de obras hidráulicas, sistemas de almacenamiento o balances hídricos, etc..

Cabe resaltar que, para la estimación de los datos faltantes, se consideraron las cuatro estaciones auxiliares con una correlación mayor a 0.7 con respecto a la estación a la que se le están generando los datos faltantes. Posteriormente, los datos faltantes obtenidos bajo ambas metodologías generaron 2 vectores que fueron promediados mediante matrices con Matlab, de esta manera, los promedios estimados son los

valores utilizados para completar las series de cada una de las variables (datos faltantes).

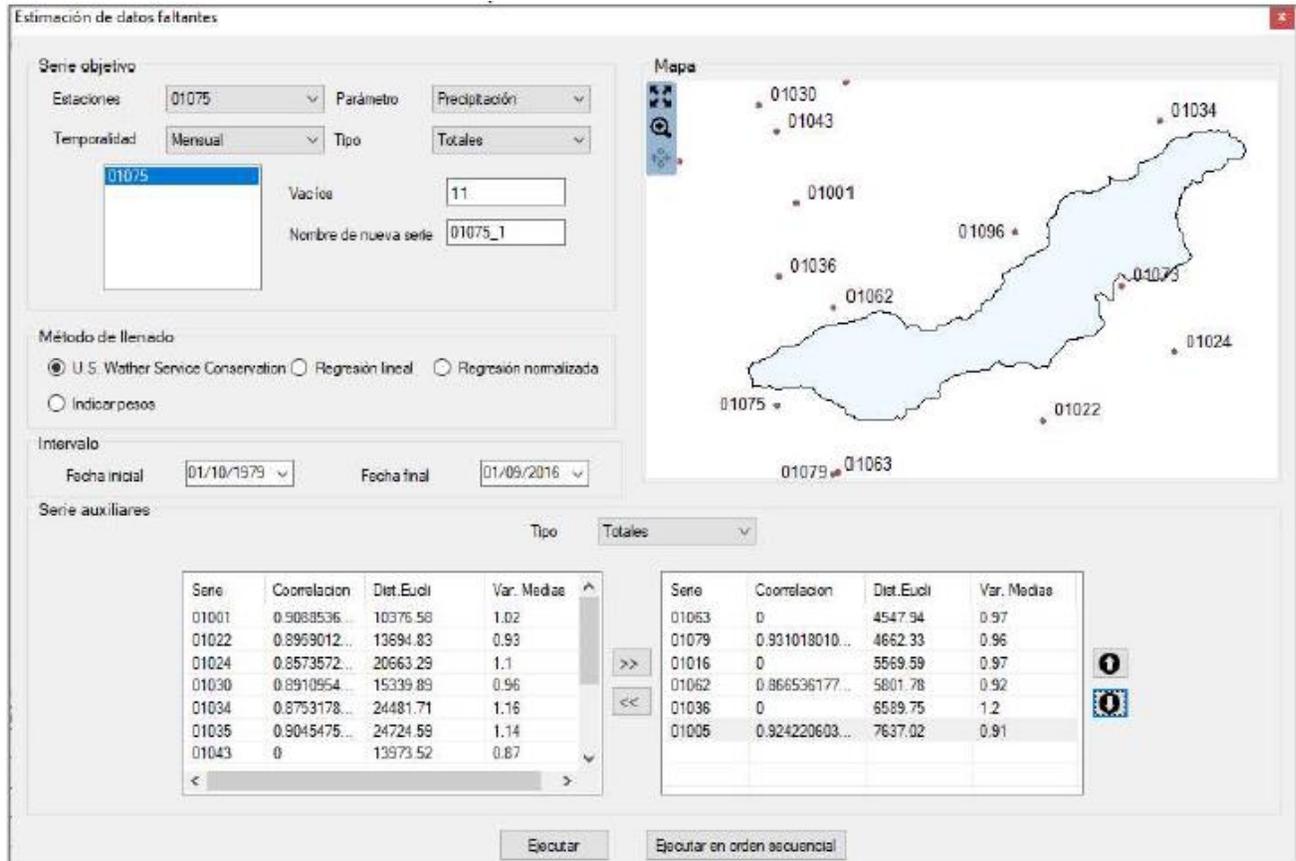


Figura 7. Ventana para la estimación de datos faltantes

3.5 Espacialización de variables

En esta etapa de la investigación se desarrollan los mapas de distribución de las estaciones climatológicas, gradientes de precipitación mensual total (ΔP), gradientes de temperatura mensual media (ΔT) y gradientes de evaporación mensual total (ΔE); de esta manera es posible evaluar de forma espacial las condiciones climáticas de la cuenca Guadalupe. Los mapas han sido elaborados a través de un sistema de información geográfica (SIG) como se describe a continuación:

- a. Georeferenciación de las estaciones climatológicas de la Tabla 2.
- b. Obtención de valores de gradientes a partir de los datos de precipitación mensual total de las 35 estaciones climatológicas. Posteriormente, se realiza la

georeferenciación de los valores obtenidos de acuerdo a la estación climatológica correspondiente. Para esto se empleó la técnica IDW mediante la ecuación (16) para la obtención del mapa de interpolación de ΔP en la cuenca Guadalupe.

- c. El procedimiento del paso *b.*, se realizó para las series de temperatura mensual media y evaporación mensual total.

El propósito de esta etapa es espacializar las variables consideradas en el análisis y a su vez descubrir las características climáticas y fisiográficas de la cuenca Guadalupe.

3.6 Análisis estadístico

Un análisis estadístico permite interpretar y comprender el comportamiento de las series de tiempo desde un enfoque estocástico, según Box-Jenkins (1976) el objetivo principal de un análisis estadístico es el poder inferir propiedades de las series de tiempos.

De esta manera, con el propósito de analizar la precipitación, temperatura y evaporación, en esta etapa se desarrolla el análisis estadístico de las series en cada una de las estaciones consideradas en el área de estudio. De esta manera el cálculo de los momentos estadísticos, series de tiempo y gradientes, muestran un indicio del comportamiento de estas variables a través del tiempo.

3.6.1 Momentos estadísticos

Generar el cálculo de los momentos estadísticos de una serie de tiempo es esencial ya que permiten obtener gran cantidad de información para comprender su comportamiento. Por esta razón, se desarrolló el cálculo de los siguientes momentos estadísticos: media, mediana, desviación estándar, varianza, coeficiente de asimetría, curtosis y coeficiente de variación. Dichos momentos estadísticos fueron obtenidos para las series de cada una de las estaciones climatológicas mediante las ecuaciones (1),(2),(3),(5),(4),(7),(8) y (6).

- a. Series de precipitación total mensual
- b. Series de temperatura promedio mensual
- c. Series de evaporación total mensual

3.6.2 Gráficos de series de tiempo

Por otra parte, se generaron los gráficos de las series de precipitación, temperatura y evaporación, con el fin de analizar la tendencia de las mismas a través del tiempo.

3.6.3 Obtención de gradientes de series de tiempo

Con el fin de evaluar la tasa de cambio a través del tiempo, es decir, la tendencia de cada variable en cada una de las estaciones, se procedió a obtener los gradientes ($\Delta P, \Delta T, \Delta E$) correspondientes a un ajuste lineal (Figura 14) a partir de las series de tiempo analizadas mediante la ecuación (17). Una vez calculados los gradientes (Tabla 8), se generaron los mapas de interpolación de los mismos (Figura 17, Figura 18 y Figura 19) mediante la técnica de interpolación IDW (ecuación (16)), esto con el objetivo de observar el comportamiento y los patrones de cambio de la precipitación, temperatura y evaporación en la cuenca Guadalupe.

3.6.3.1 Matriz de correlación de gradientes

Una vez obtenidos los gradientes $\Delta P, \Delta T, \Delta E$ (Tabla 8), se procedió a obtener la matriz de correlación de los mismos, para lograr esto se consideraron las características fisiográficas de la cuenca y los promedios anuales de las variables climáticas en estudio (Tabla 9).

El objetivo principal es analizar la relación que existe entre las variables climáticas y las características fisiográficas de la cuenca y determinar si existe cierto efecto de las características fisiográficas sobre las variables climáticas, lo cual podría tener una relación con la variabilidad climática del área de estudio (Tabla 1).

3.7 Autocorrelación y Autocorrelación Parcial

Con el propósito de medir la correlación entre los valores de la serie distanciados un lapso de tiempo k , se determinó la ACF para cada una de las estaciones climatológicas de cada una de las variables climáticas mediante la ecuación (29). Así mismo, se obtuvo la PACF para medir la correlación entre las observaciones de cada serie separadas τ periodos mediante la ecuación (30), para posteriormente iniciar con la identificación de los modelos a utilizar.

3.7.1 Autocorrelograma

Una vez determinadas las ACF y PACF, se realizaron los autocorrelogramas para cada una de las variables climáticas en estudio en cada una de sus estaciones. Estos son la representación gráfica de ACF y PACF, los cuales permiten visualizar si la serie es aleatoria, determinar si existe estacionariedad o no, así como patrones repetitivos dentro de la serie.

3.8 Análisis de Consistencia

Cuando los datos de las variables en investigación se basan en el tiempo es de vital importancia realizar pruebas de estacionariedad y a su vez determinar la forma de la tendencia de los datos ya que generalmente estos pueden contener tendencias estocásticas que se caracterizan por la raíz unitaria (Augmented Dickey Fuller Test Rizwan Mushtaq). Uno de los distintos procedimientos para determinar dicha propiedad en las series de tiempo es la prueba Dickey Fuller Aumentado (ADF). Por lo tanto, una vez completadas las series de tiempo, el siguiente paso fue realizar la prueba ADF mediante las ecuaciones (45), (48) y (49).

Así también se realizaron las pruebas Helmer (33), Secuencias y t de Student (38) para detectar homogeneidad en las series y la prueba Mann-Kendall para determinar la tendencia de las mismas mediante las ecuaciones (35), (36) y (37).

3.9 Proceso de normalización

Antes de iniciar el proceso de identificación y selección de modelos, generalmente es necesario normalizar los datos de las variables en estudio. El propósito del proceso de normalización es garantizar que los valores de la serie de tiempo se mantengan dentro de cierto rango y sigan una distribución normal. En términos generales, normalizar los datos es realizar un ajuste a fin de que tengan una pendiente constante en el tiempo, es decir, remover el sesgo (coeficiente de asimetría igual a cero y coeficiente de curtosis igual a 3) mediante transformaciones normalizantes (logarítmica, Box-Cox o Box-Cox segunda potencia) en el caso de que las series de datos de cada estación no sean normales (Escalante & Reyes, 1998; Medina Hurtado & García Aguado, 2005).

En este caso, debido a que usualmente las series de tiempo hidrológicas no siguen una distribución normal, se procedió al proceso de normalización de los datos mediante las transformaciones Box-Cox (66) y Box-Cox segunda potencia (67), para posteriormente iniciar con la etapa de construcción de modelos para las variables en estudio.

3.10 Identificación y selección de modelos Autorregresivos

Una vez que las series de datos fueron transformadas, se inicia el proceso de identificación de modelos para cada una de las estaciones. Para esto, primeramente se realiza la prueba Dickey-Fuller DF (ecuación(42)) o Dickey-Fuller aumentada ADF (ecuación (45), cuando la serie no resulta estacionaria en la prueba DF) a las series de datos transformados, con el propósito de determinar si son estacionarias.

El segundo paso es identificar los modelos, mediante la utilización del paquete estadístico Eviews, se generaron 169 modelos distintos para cada estación y variable en estudio, finalmente la selección del modelo se hizo en base al criterio de información de Akaike (AIC), eligiendo el que tiene el valor mínimo de AIC conforme la ecuación (60).

3.11 Espacialización de componentes de modelos Autorregresivos

En esta etapa se desarrollan los mapas de los componentes ϕ y θ obtenidos de cada modelo seleccionado para cada una de las estaciones climatológicas y de las variables en estudio, de esta manera es posible visualizar de forma espacial el comportamiento de la variable analizada en cada una de las estaciones de la cuenca Guadalupe. De igual manera que los mapas realizados anteriormente, estos también han sido elaborados a través de un sistema de información geográfica (SIG) como se describe a continuación:

- a. Georeferenciación de los parámetros ϕ y θ obtenidos para los modelos de precipitación mensual total de acuerdo a la estación climatológica correspondiente. Para esto se empleó la técnica IDW mediante la ecuación (16) para la obtención del mapa de interpolación en la cuenca Guadalupe.
- b. El procedimiento del paso a., se realizó también para los parámetros de los modelos obtenidos de las series de temperatura mensual media y evaporación mensual total.

3.12 Algebra de Mapas

Una vez geoespacializados cada uno de los parámetros de los modelos, los mapas obtenidos se convierten en nuevas capas de información tipo ráster con las cuales es posible realizar distintas operaciones algebraicas (Figura 8). El objetivo final de estas operaciones algebraicas mejor conocidas como algebra de mapas es producir una o varias capas de salida, en este caso, mapas de pronóstico para la precipitación mensual total, temperatura mensual media y evaporación mensual total, que permitan conocer el comportamiento futuro en la cuenca Guadalupe, es decir, realizar análisis de

la zona para identificar áreas de decremento o incremento de las variables en estudio y predecir escenarios futuros de riesgo o vulnerabilidad climática.

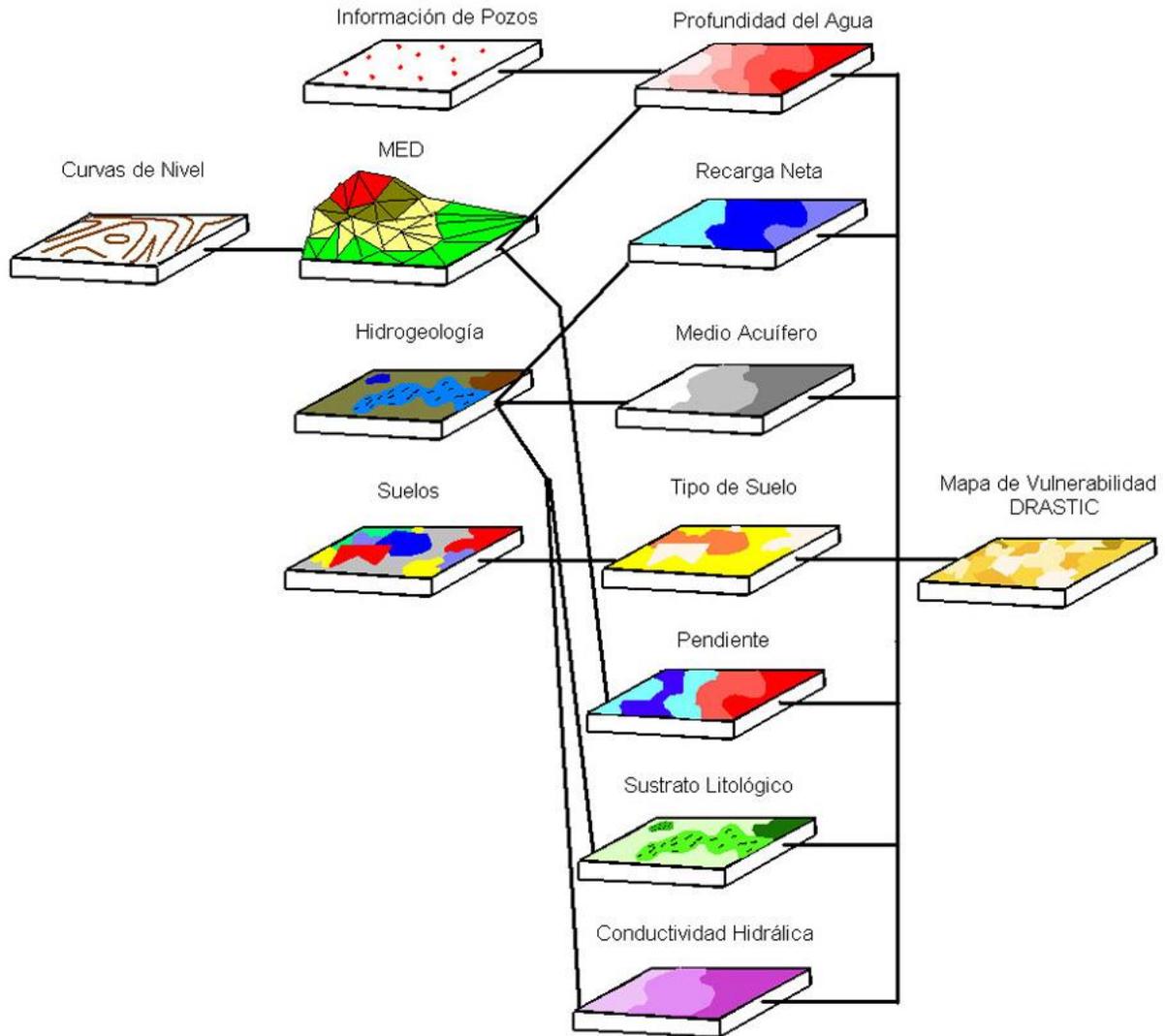


Figura 8. Ejemplo de capas de información tipo ráster para generar algebra de mapas. Fuente: (Fonseca, 2013)

4 RESULTADOS Y ANÁLISIS DE RESULTADOS

A continuación, se presentan los resultados obtenidos durante la investigación, así como el análisis de los mismos con el objetivo de visualizar y detectar comportamientos, relaciones y efectos de las variables climatológicas estudiadas en la cuenca Guadalupe.

4.1 Condiciones climatológicas y fisiográficas de la cuenca Guadalupe, B.C.

En lo que respecta al mapa de geolocalización de estaciones climatológicas (Figura 9) generado, la idea principal fue considerar primeramente el mayor número de estaciones, de igual forma, se optó por que dichas estaciones presenten una distribución uniforme sobre la cuenca, ya que esto repercute de manera positiva a los resultados de las variables climáticas al momento de realizar la geoespacialización y permite clasificar áreas de acuerdo a la similitud de estaciones vecinas, además de obtener una mejor definición del área de influencia de cada estación y las diferencias entre patrones de las variables en estudio en distintos periodos.

De igual forma, se generaron los mapas de geolocalización de los promedios anuales de las variables analizadas en esta investigación (Tabla 2), con base en los mapas generados y la información consultada en el Servicio Meteorológico Nacional, se puede decir que la precipitación total mensual de la cuenca Guadalupe está en los 118.42 mm y 470.60 mm al año (Figura 10), mostrando que las estaciones ubicadas en el área de Ojos Negros y Valle de Guadalupe son las que presentan registros de precipitación más altos que en La Misión.

En cuanto a la temperatura mensual media esta se encuentra entre los 9.4°C y los 25.79°C (Figura 11), en donde algunas de las estaciones que se encuentran en Ojos Negros y Valle de Guadalupe presentan las temperaturas más bajas, presentando un comportamiento contrario a la precipitación. Con respecto a la evaporación total mensual en el área de estudio, esta se encuentra entre los 920.23 mm y 2102.66 mm (Figura 12), siendo evidente los tres valles que conforman la cuenca presentan cambios significativos entre una estación y otra. El comportamiento de cada una de las variables analizadas en la cuenca podría estar representando características distintas a las de la zona en que se ubican.

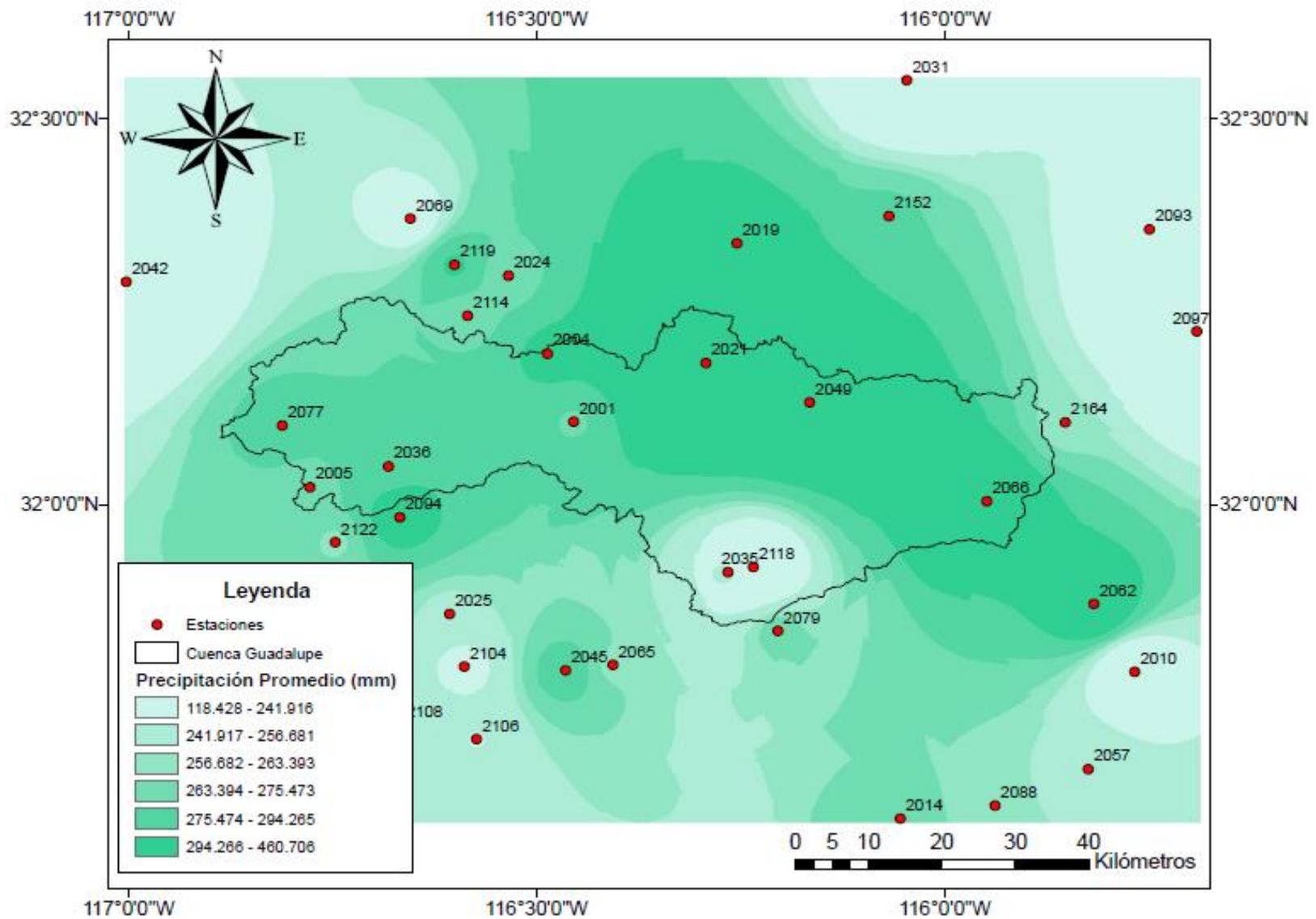


Figura 10. Geoespacialización del promedio anual de la precipitación total mensual en cuenca Guadalupe

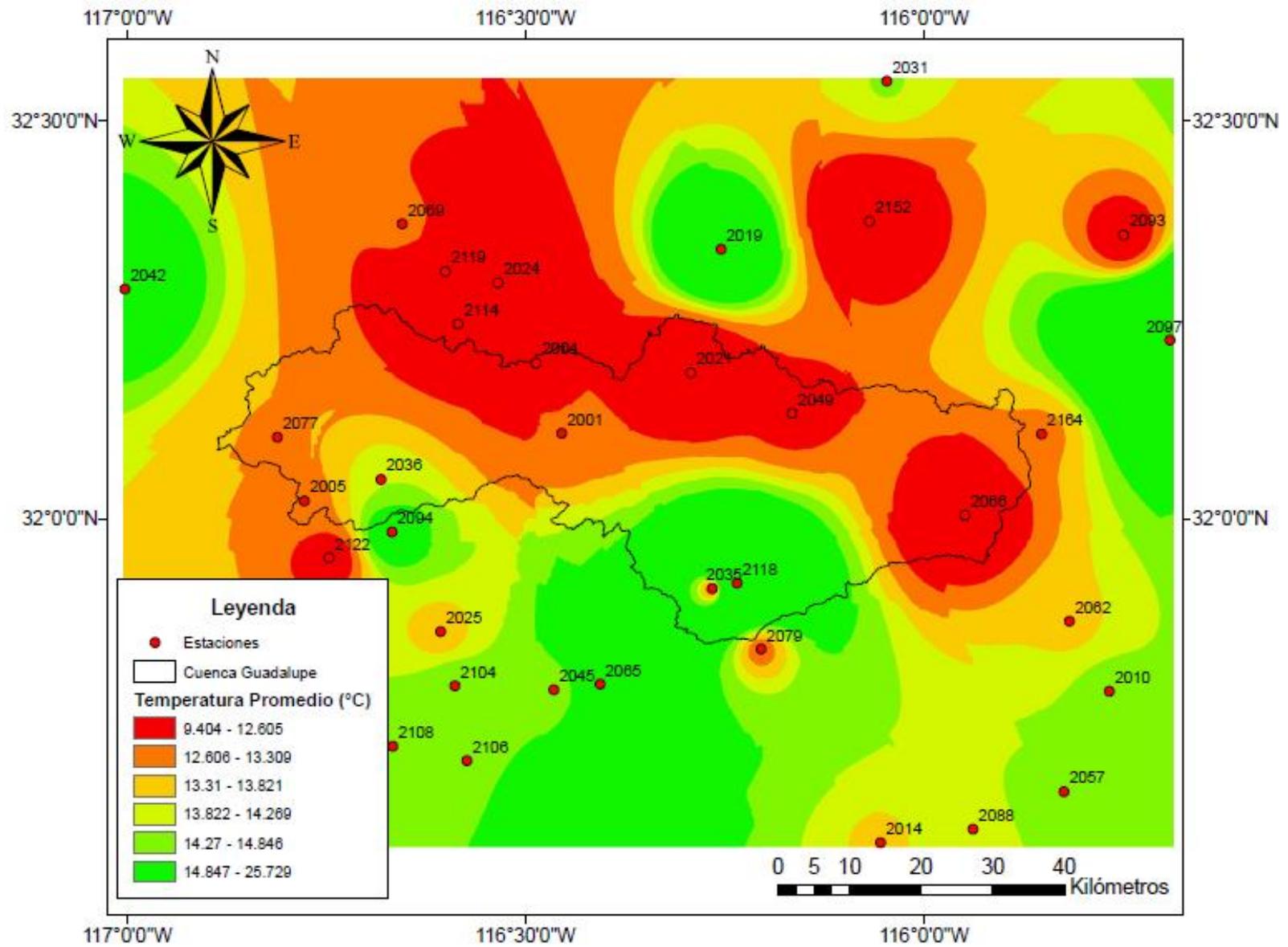


Figura 11. Geoespacialización de la temperatura mensual media anual en cuenca Guadalupe

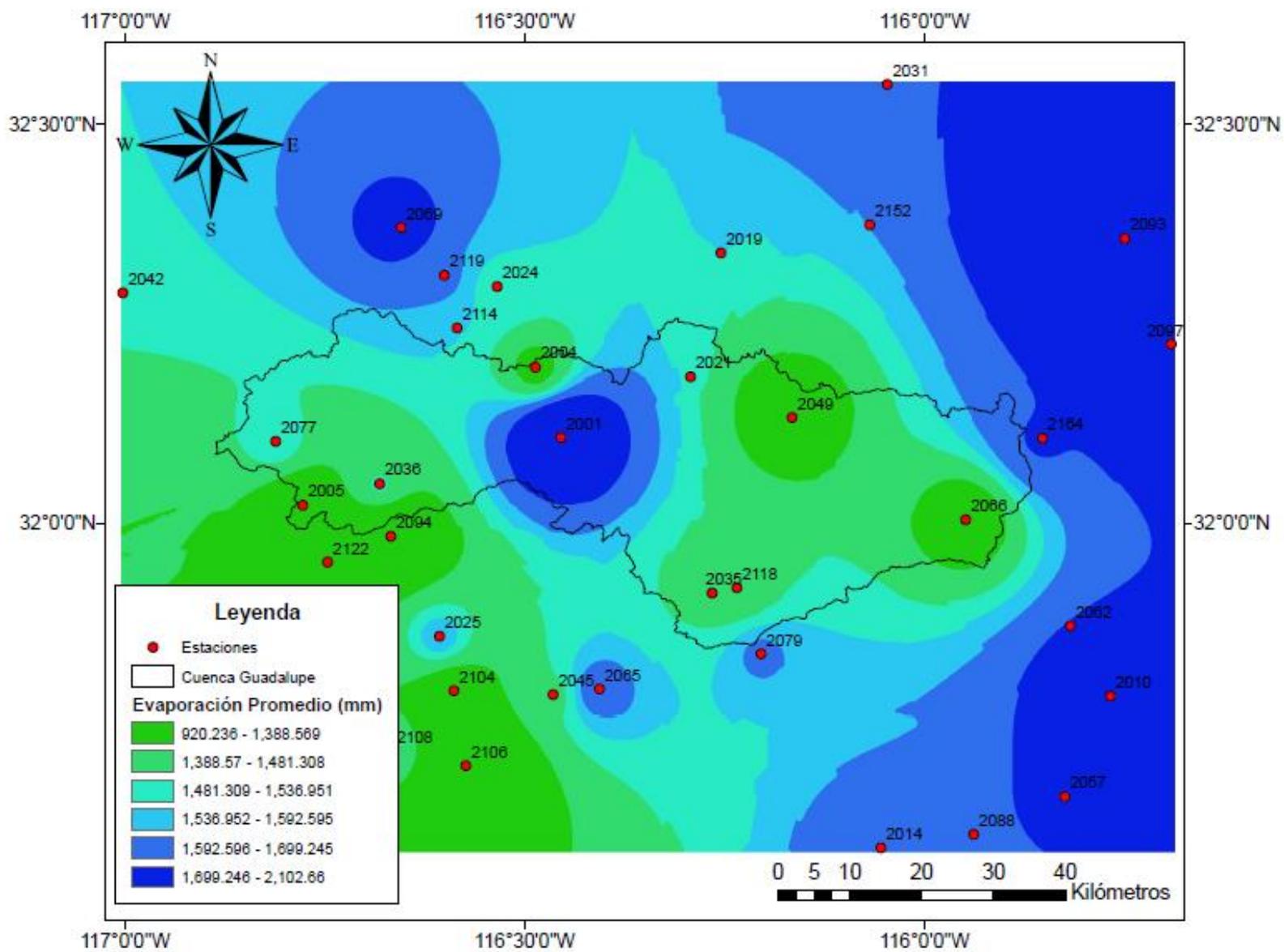


Figura 12. Geoespacialización del promedio anual de evaporación total mensual en cuenca Guadalupe

4.2 Análisis estadístico de series de precipitación, temperatura y evaporación

4.2.1 Momentos estadísticos y series de tiempo

Se obtuvieron los momentos estadísticos de las 35 estaciones climatológicas, con el fin de determinar el comportamiento de las series de tiempo de precipitación, temperatura y evaporación obtenidas de los registros de las estaciones analizadas. En la Tabla 4, Tabla 5 y Tabla 6 se muestra la información estadística obtenida.

En primera instancia se analiza el comportamiento de los momentos estadísticos de la variable de precipitación total mensual, de acuerdo a la Tabla 4, se observa que estos no presentan una tendencia, sino que fluctúan. En lo que respecta a la media y la varianza, estas presentan cambios significativos, lo que nos indica que la serie presenta un comportamiento no estacionario. El coeficiente de asimetría refleja que los valores tienden a ubicarse del lado izquierdo de la media, es decir, presentan asimetría positiva. Por otro lado, el hecho de que la curtosis para todas las estaciones sea mayor a 3, refleja que la mayoría de los valores se concentran en torno a la media, sin embargo, también nos indica que la serie de precipitación no sigue un comportamiento normal.

Caso similar se presenta en la temperatura mensual media (Tabla 5), en donde también se presentan fluctuaciones en la media y la varianza de las 35 estaciones, indicando que no presentan tendencia y que su comportamiento es no estacionario. En lo que respecta al coeficiente de asimetría, se puede observar que existen valores negativos en algunas estaciones, lo que nos dice que se concentran del lado derecho de la media, pero en su mayoría, los valores se encuentran del lado izquierdo de la media, presentando una asimetría positiva. De igual manera que para la precipitación, la curtosis nos está indicando que la serie de temperatura no sigue un comportamiento normal, sin embargo, en este caso, el valor de la curtosis es menor a 3, lo que indica que existe una baja concentración de valores en torno a la media.

En lo que respecta a la evaporación mensual total, con base en la Tabla 6, se puede concluir que los datos de la media y la varianza presentan fluctuaciones, lo que indica que existe una tendencia y que siguen un comportamiento no estacionario. Así mismo el coeficiente de asimetría, indica que los datos se concentran del lado izquierda de la media, es decir, presentan asimetría positiva. Ahora bien, considerando los valores obtenidos para la curtosis, se muestra que es menor a 3, lo que indica que existe una baja concentración de valores en torno a la media, es decir, presenta una curtosis platicúrtica.

Tabla 4 Estadísticos de precipitación total mensual observada.

Clave	Nombre de estación	Media	Mediana	Des. Est.	Coef. Asimetría	Curtosis	Coef. Variación	Varianza
2001	Agua Caliente	22.90	6.25	36.12	2.46	10.34	1.57	1302.78
2004	Ignacio Zaragoza-Belen	25.80	7.00	41.60	2.46	9.81	1.61	1727.65
2005	Boquilla Sta. Rosa de La Misión	22.70	5.30	38.83	2.89	13.59	1.71	1506.28
2010	Compostela	17.20	3.90	29.37	2.52	9.67	1.70	857.34
2014	El Alamo	23.68	7.00	39.19	2.90	14.01	1.65	1533.49
2019	El Compadre	27.48	11.60	38.45	2.09	7.92	1.39	1471.74
2021	El Pinal	42.03	19.80	60.30	2.40	9.82	1.43	3625.31
2024	El Testerazo	22.73	5.70	38.77	3.08	16.34	1.70	1500.32
2025	Ensenada	20.96	5.10	33.33	2.08	7.19	1.59	1104.14
2031	La Rumorosa	15.33	3.00	27.44	3.02	15.00	1.79	752.41
2035	Ojos Negros	20.88	6.50	33.11	2.48	9.94	1.58	1094.22
2036	Olivares Mexicanos	24.20	5.95	40.54	2.64	11.44	1.67	1641.97
2042	Rosarito	10.79	0.00	20.43	3.60	21.65	1.89	415.38
2045	San Carlos	24.60	6.00	39.98	2.43	9.69	1.62	1595.92
2049	San Juan de Dios norte	32.63	14.00	48.22	2.48	10.53	1.47	2319.35
2057	Santa Catarina Norte	22.09	11.00	29.17	2.17	8.98	1.32	839.13
2062	Santa Isabel	29.98	14.00	39.39	1.80	6.59	1.31	1495.04
2065	Santo Tomas	22.06	3.00	40.50	3.03	14.84	1.83	1629.71
2066	Sierra de Juárez	32.33	16.00	47.47	2.83	13.78	1.46	2246.40
2069	Valle de Las Palmas	18.17	4.20	31.17	3.13	17.30	1.71	970.80
2077	La Misión	24.14	5.60	39.19	2.00	6.28	1.62	1526.29
2079	El Alamar	22.83	7.50	39.25	3.80	25.56	1.71	1536.40
2088	Ej. Héroes de la Independencia	23.49	9.40	35.21	2.70	13.34	1.49	1236.69
2093	Ej. Valle de la Trinidad	19.40	7.30	31.73	3.17	16.97	1.63	1004.72
2094	El Farito	28.48	7.30	47.48	2.55	10.10	1.66	2214.35
2097	Laguna Salada	10.37	0.00	23.85	3.17	13.39	2.29	566.21
2104	El Ciprés	20.31	3.79	36.28	2.93	13.97	1.78	1313.50
2106	Maneadero	21.14	4.36	35.66	2.54	10.11	1.68	1268.60
2108	Punta Banda	22.20	3.00	41.02	3.16	16.56	1.84	1679.23
2114	Ej. Carmen Serdán	23.35	3.69	41.81	2.57	10.08	1.79	1743.95
2118	Valle de San Rafael	21.91	8.50	32.85	2.47	10.50	1.49	1075.10
2119	Cancio	28.78	7.85	45.59	2.52	10.85	1.58	2066.34
2122	Real del Castillo Viejo	23.20	7.50	36.07	2.36	9.70	1.55	1291.77
2152	Ej. José María Pino Suárez	25.73	7.38	42.53	2.78	12.85	1.65	1803.70
2164	Ej. El Porvenir	22.67	6.00	41.06	3.49	19.98	1.81	1668.29

Tabla 5 Estadísticos de temperatura media mensual observada.

Clave	Nombre de estación	Media	Mediana	Des. Est.	Coef. Asimetría	Curtosis	Coef. Variación	Varianza
2001	Agua Caliente	12.67	12.21	4.45	0.29	1.92	0.35	19.77
2004	Ignacio Zaragoza-Belen	12.78	12.19	5.90	0.29	2.45	0.46	34.79
2005	Boquilla Sta. Rosa de La Misión	13.12	13.22	3.87	-0.02	2.29	0.29	15.00
2010	Compostela	14.52	13.81	5.91	0.28	2.05	0.40	34.73
2014	El Alamo	13.66	12.52	6.40	0.27	1.92	0.46	40.92
2019	El Compadre	19.95	20.00	6.74	-0.13	2.60	0.33	44.97
2021	El Pinal	9.60	8.80	4.81	0.72	3.04	0.50	23.06
2024	El Testerazo	11.39	10.45	5.27	0.64	2.80	0.46	27.78
2025	Ensenada	13.37	12.93	3.55	0.17	1.78	0.26	12.53
2031	La Rumorosa	14.86	14.03	7.00	0.23	1.90	0.47	48.95
2035	Ojos Negros	12.82	11.65	6.77	0.49	2.55	0.52	45.84
2036	Olivares Mexicanos	14.35	14.10	4.58	0.33	2.63	0.31	21.01
2042	Rosarito	17.12	17.00	6.30	0.66	3.78	0.36	39.59
2045	San Carlos	14.32	14.68	4.20	-0.13	2.55	0.29	17.68
2049	San Juan de Dios norte	11.95	11.58	5.75	0.38	2.32	0.48	32.99
2057	Santa Catarina Norte	14.49	14.06	5.55	0.30	2.15	0.38	30.72
2062	Santa Isabel	13.89	12.79	6.16	0.43	2.04	0.44	37.78
2065	Santo Tomas	16.23	15.94	5.25	0.19	2.37	0.32	27.59
2066	Sierra de Juárez	9.93	8.33	5.81	0.64	2.56	0.58	33.67
2069	Valle de Las Palmas	12.79	12.60	4.87	0.31	2.36	0.38	23.73
2077	La Misión	13.25	13.45	4.20	0.09	2.08	0.31	17.58
2079	El Alamar	12.98	12.87	3.58	0.09	1.91	0.27	12.83
2088	Ej. Héroes de la Independencia	14.23	13.37	5.64	0.24	2.10	0.39	31.72
2093	Ej. Valle de la Trinidad	11.59	10.39	5.44	0.41	2.14	0.46	29.61
2094	El Farito	14.01	14.32	3.37	-0.01	1.80	0.24	10.93
2097	Laguna Salada	17.08	16.18	7.92	0.43	2.20	0.46	62.42
2104	El Ciprés	14.89	15.22	3.61	-0.25	2.26	0.24	13.04
2106	Manadero	14.82	15.26	3.98	-0.37	2.93	0.26	15.85
2108	Punta Banda	14.76	14.90	3.32	-0.26	2.60	0.22	11.05
2114	Ej. Carmen Serdán	11.15	10.54	4.96	0.52	2.98	0.44	24.60
2118	Valle de San Rafael	12.21	10.85	5.03	0.43	1.98	0.41	25.23
2119	Cancio	12.59	12.11	4.66	0.30	2.02	0.37	21.67
2122	Real del Castillo Viejo	12.01	10.70	5.81	0.21	1.89	0.48	33.49
2152	Ej. José María Pino Suárez	10.19	9.44	6.42	0.39	2.32	0.63	41.16
2164	Ej. El Porvenir	12.98	12.59	3.32	0.12	2.23	0.25	10.99

Tabla 6 Estadísticos de evaporación total mensual observada.

Clave	Nombre de estación	Media	Mediana	Des. Est.	Coef. Asimetría	Curtosis	Coef. Variación	Varianza
2001	Agua Caliente	181.57	172.50	82.88	0.21	1.95	0.45	6855.41
2004	Ignacio Zaragoza-Belen	151.35	145.23	70.18	0.46	2.32	0.4637	4905.77
2005	Boquilla Sta. Rosa de La Misión	115.33	115.70	43.53	0.20	2.08	0.3774	1889.27
2010	Compostela	157.53	147.90	69.83	0.32	1.91	0.4432	4847.45
2014	El Alamo	154.69	143.75	78.12	0.43	2.02	0.5050	6103.69
2019	El Compadre	-	-	-	-	-	-	-
2021	El Pinal	-	-	-	-	-	-	-
2024	El Testerazo	-	-	-	-	-	-	-
2025	Ensenada	133.70	132.70	37.97	0.00	1.87	0.28	1432.55
2031	La Rumorosa	-	-	-	-	-	-	-
2035	Ojos Negros	154.04	139.22	88.38	0.38	1.72	0.57	7674.18
2036	Olivares Mexicanos	130.90	124.54	54.26	0.32	2.13	0.41	2938.01
2042	Rosarito	-	-	-	-	-	-	-
2045	San Carlos	139.16	137.30	51.98	0.11	1.97	0.37	2695.32
2049	San Juan de Dios norte	149.52	147.74	61.69	0.25	2.14	0.41	3789.19
2057	Santa Catarina Norte	156.78	156.37	57.92	0.14	2.14	0.36	3345.96
2062	Santa Isabel	-	-	-	-	-	-	-
2065	Santo Tomas	153.82	148.30	73.78	0.35	2.06	0.47	5431.33
2066	Sierra de Juárez	112.27	101.90	50.83	0.79	3.20	0.45	2570.84
2069	Valle de Las Palmas	160.30	147.30	68.77	0.35	2.19	0.42	4717.82
2077	La Misión	128.79	121.64	44.49	0.23	2.05	0.34	1967.23
2079	El Alamar	148.34	144.60	57.37	0.19	1.88	0.38	3277.90
2088	Ej. Héroes de la Independencia	-	-	-	-	-	-	-
2093	Ej. Valle de la Trinidad	190.84	177.63	81.21	0.32	1.95	0.42	6577.34
2094	El Farito	-	-	-	-	-	-	-
2097	Laguna Salada	182.99	178.95	82.45	0.50	2.41	0.45	6750.93
2104	El Ciprés	117.45	109.66	51.13	0.87	3.44	0.43	2607.62
2106	Manadero	109.48	109.90	42.51	0.34	2.35	0.38	1798.49
2108	Punta Banda	140.17	132.00	56.02	0.27	2.03	0.39	3122.99
2114	Ej. Carmen Serdán	-	-	-	-	-	-	-
2118	Valle de San Rafael	143.95	133.66	69.60	0.86	3.12	0.48	4829.81
2119	Cancio	164.13	155.69	80.55	0.78	2.82	0.49	6445.35
2122	Real del Castillo Viejo	-	-	-	-	-	-	-
2152	Ej. José María Pino Suárez	-	-	-	-	-	-	-
2164	Ej. El Porvenir	147.62	147.00	57.00	0.10	1.71	0.38	3226.08

A manera de ejemplificar que, en efecto los datos no presentan una distribución normal, se presentan los momentos estadísticos de la serie de precipitación diaria para la estación 2001, con base en la Tabla 7, y al valor que presenta la curtosis se puede concluir que al ser mayor a 3 presenta una distribución leptocúrtica, es decir, la mayoría de los datos están concentrados en torno a su media.

Tabla 7 Estadísticos de precipitación diaria observada

Clave	Nombre de estación	Media	Mediana	Des. Est.	Coef. Asimetría	Curtosis	Coef. Variación	Varianza
2001	Agua Caliente	0.75	0	3.86	9.06	123.04	5.14	14.94

4.2.2 Análisis de gradientes de variables climáticas en estudio

Para iniciar con el análisis de gradientes $\Delta P, \Delta T$ y ΔE , es importante analizar el comportamiento de la serie de tiempo observada, para el caso de la precipitación total mensual en la estación 2001, la Figura 13 muestra claramente una tendencia no lineal, además puede observarse que la serie presenta un incremento seguido de un decremento y viceversa, lo que nos dice que los datos presentan una variación aleatoria, es decir, existe la misma probabilidad de registrar un incremento o un decremento independientemente del registro anterior.

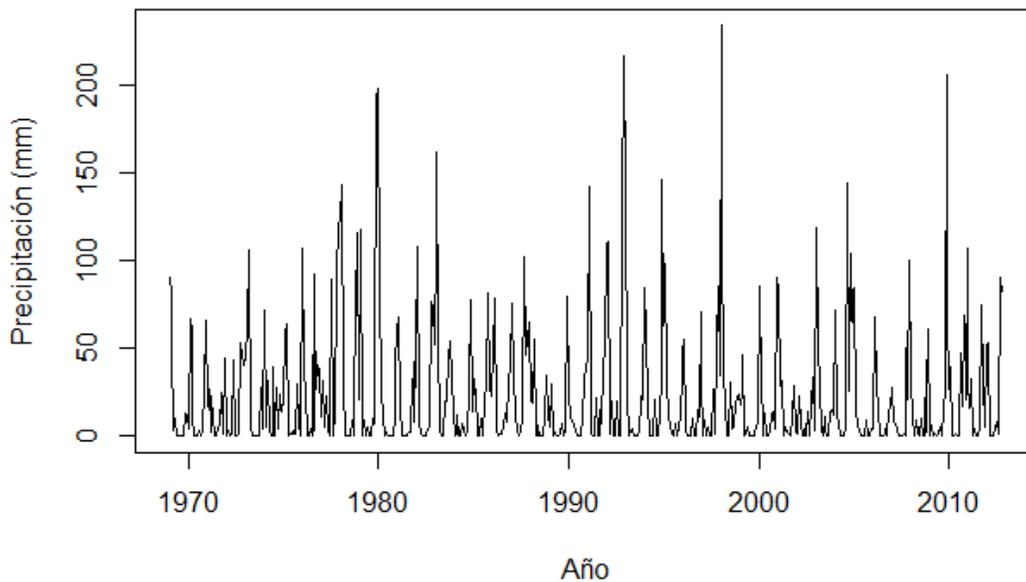


Figura 13. Serie de tiempo para datos de precipitación total mensual para la estación Aguacaliente (2001)

De este modo es posible afirmar que se conoce el comportamiento de la serie de tiempo de la estación Aguacaliente (2001), sin embargo, es necesario un análisis comparativo con el resto de las estaciones que se encuentran espacializadas en la cuenca Guadalupe, con el fin de detectar patrones de comportamiento o tendencias que permitan generalizar entre el comportamiento de las series de precipitación,

temperatura y evaporación de las estaciones climatológicas y las características fisiográficas que presentan la zona donde están ubicadas.

Tabla 8 Gradientes de series de tiempo

Clave	Nombre de Estación	Longitud	Latitud	Gradiente Temperatura (ΔT)	Gradiente de Precipitación (ΔP)	Gradiente Evaporación (ΔE)
2001	Aguacaliente	-116.454	32.107	0.0005	-0.0106	0.0955
2004	Ignacio Zaragoza-Belen	-116.486	32.195	-0.0088	-0.0142	0.012
2005	Boquilla Sta. Rosa de La Misión	-116.777	32.022	0.0009	0.0114	0.0957
2010	Compostela	-115.767	31.783	0.0215	0.0059	0.024
2014	El Alamo	-116.054	31.593	-0.007	0.0083	-0.1778
2019	El Compadre	-116.254	32.338	0.0227	0.0106	-
2021	El Pinal	-116.292	32.183	-0.0089	-0.004	-
2024	El Testerazo	-116.534	32.296	0.0035	-0.0236	-1.3157
2025	Ensenada	-116.606	31.858	0.0029	0.0534	-0.0056
2031	La Rumorosa	-116.046	32.549	-0.0032	0.0132	-
2035	Ojos Negros	-116.265	31.912	-0.0127	0.0212	-0.4401
2036	Olivares Mexicanos	-116.681	32.049	-0.008	0.0019	0.0151
2042	Rosarito	-117.002	32.288	-0.0532	0.0513	-
2045	San Carlos	-116.464	31.785	-0.0002	-0.0141	-0.0709
2049	San Juan de Dios Norte	-116.165	32.132	-0.0157	0.092	0.1043
2057	Sta. Catarina Norte	-115.824	31.657	-0.001	0.0315	0.0442
2062	Sta. Isabel	-115.817	31.871	0.0136	0.111	-
2065	Santo Tomas	-116.406	31.792	0.0059	0.0074	-0.1029
2066	Sierra de Juárez	-115.948	32.004	0.0114	0.1305	0.285
2069	Valle de Las Palmas	-116.654	32.37	-0.0106	0.0113	-0.0974
2077	La Misión	-116.811	32.102	-0.0254	0.065	0.0674
2079	El Alamar	-116.204	31.836	-0.0023	-0.0087	-0.0655
2088	Ej. Héroes de la Independencia	-115.938	31.61	0.0016	-0.0201	-
2093	Ej. Valle de la Trinidad	-115.749	32.356	-0.0004	-0.0316	0.1677
2094	El Farito	-116.667	31.983	-0.1182	-0.0381	0.8407
2097	Laguna Salada	-115.691	32.224	-0.0098	-0.0521	-0.1347
2104	El Ciprés	-116.588	31.79	-0.0016	-0.0622	-0.0774
2106	Maneadero	-116.573	31.696	0.0142	-0.0525	0.2019
2108	Punta Banda	-116.666	31.714	-0.001	-0.0361	-0.0232
2114	Ejido Carmen Serdán	-116.584	32.244	-0.0223	-0.0339	-
2118	Valle de San Rafael	-116.234	31.919	0.0032	-0.0392	0.0638
2119	Cancio	-116.6	32.31	0.0176	-0.0038	-0.4325
2122	Real del Castillo Viejo	-116.746	31.951	-0.0049	-0.2107	-
2152	Ej. José María Pino Suarez	-116.068	32.373	0.0333	-0.0776	-
2164	Ej. El Porvenir	-115.852	32.106	0.0022	0.0705	-0.0279

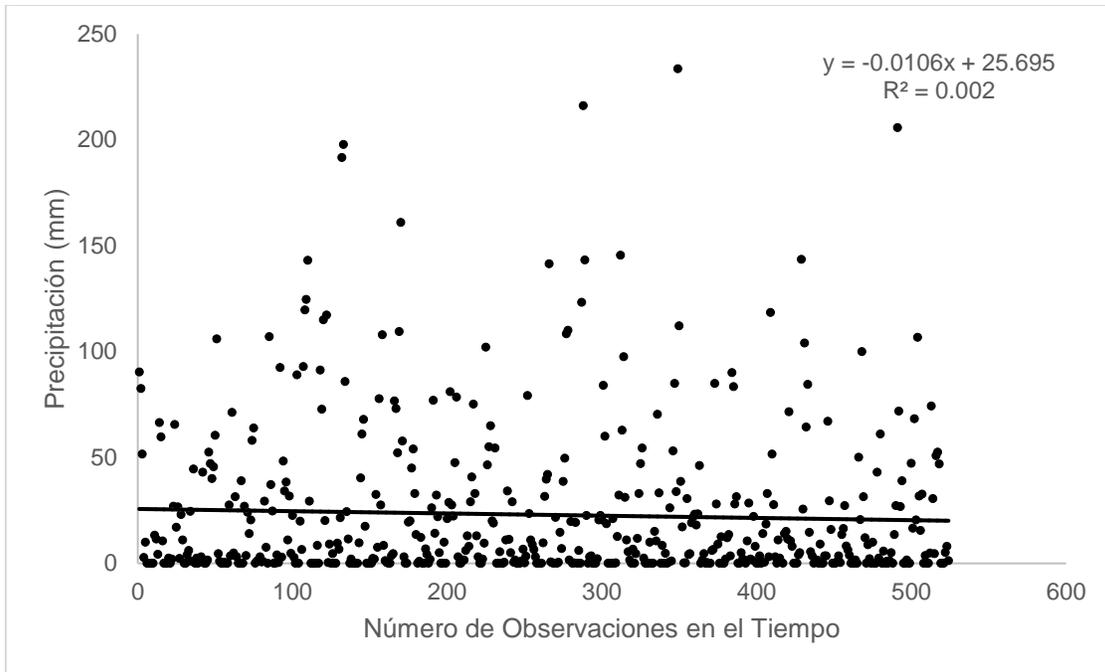


Figura 14. Ajuste para obtención de gradientes de precipitación total mensual de la estación Aguacaliente (2001)

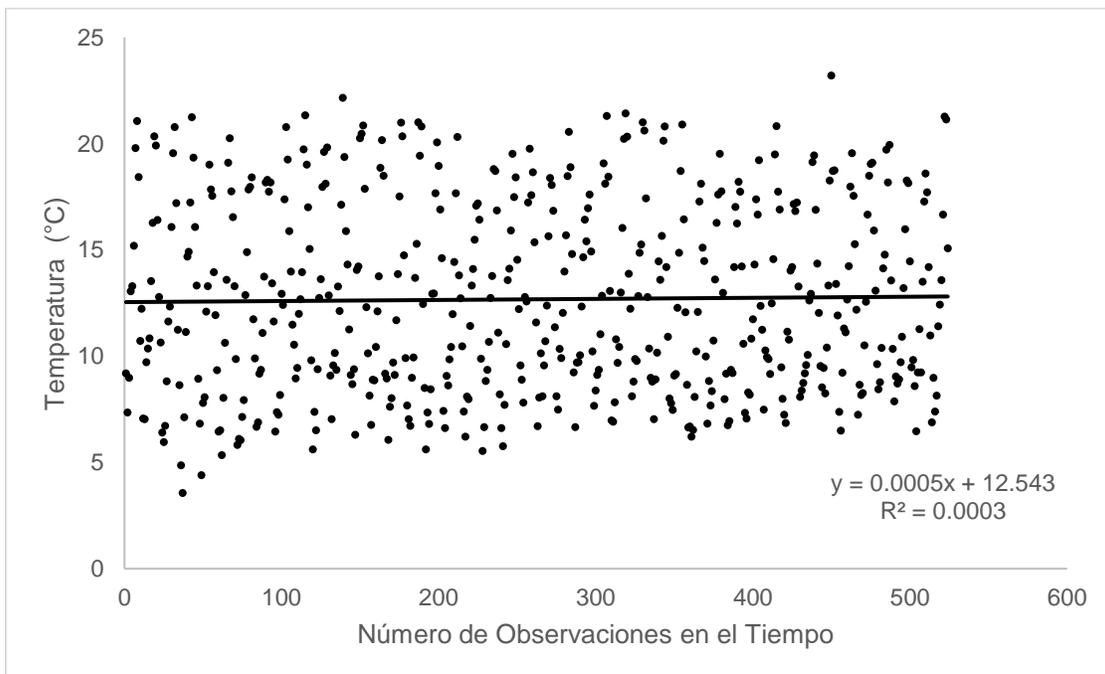


Figura 15. Ajuste para obtención de gradientes de temperatura mensual media de la estación Aguacaliente (2001)

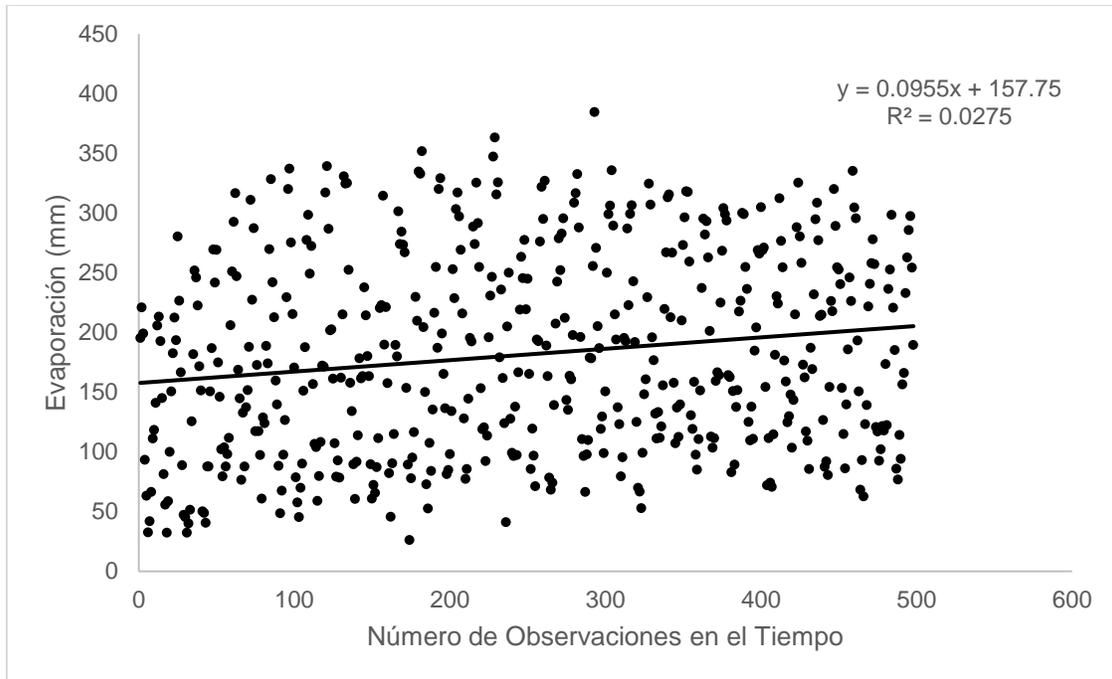


Figura 16. Ajuste para obtención de gradientes de evaporación total mensual de la estación Aguacaliente (2001)

Con base en los gráficos de ajuste para la obtención de gradientes ($\Delta P, \Delta T, \Delta E$) de la estación 2001 (Figura 14, Figura 15 y Figura 16), se puede afirmar que en esta estación las variables de temperatura y evaporación tienden a aumentar a través del tiempo, caso contrario se presenta en la precipitación, en donde el gradiente negativo, nos indica que esta variable tiende a disminuir a través del tiempo (Figura 17).

Ahora bien, el ajuste para la obtención de gradientes (ecuación (17)) de las variables climáticas en estudio (Tabla 8), refleja que a lo largo de la cuenca existen ciertas estaciones en las que los valores de los gradientes ($\Delta P, \Delta T, \Delta E$) son negativos, lo que indica que en estas estaciones (2045, 2079, 2097, 2104 y 2108) las tres variables en estudio tienden a disminuir, lo que se confirma en los mapas de gradientes generados (Figura 17, Figura 18 y Figura 19). Siendo el Valle de Guadalupe la zona en donde se presentan la mayor cantidad de estaciones con valores negativos de ΔT , es decir, que la temperatura mensual media tiende a disminuir a través del tiempo (Figura 18); seguida de la zona de La Misión en donde las estaciones localizadas en esa área también presentan una tendencia negativa, lo que indica la disminución de la temperatura a través del tiempo. Estas disminuciones de temperatura tanto en la zona de La Misión como en Valle de Guadalupe se podrían asociar a la influencia del

Océano Pacífico debido a que genera vientos frescos y fríos. Por esta razón, la interacción de este es muy significativa en el clima de la región, lo que se refleja en un clima muy variante.

Sin embargo, en Ojos Negros se presentan tanto ΔT negativos como positivos, en la Figura 18 es posible observar que las estaciones que colindan con Valle de Guadalupe presentan tendencias negativas, lo que indica que en esas áreas la temperatura tiende a disminuir, sin embargo, las estaciones que se encuentran alejadas del Valle de Guadalupe presentan ΔT positivos, es decir, que en estas estaciones la temperatura tiende a aumentar a través del tiempo.

En relación a los ΔP al igual que en los ΔT , Valle de Guadalupe es la zona en donde se encuentran la mayoría de las estaciones con tendencia negativa, lo que indica que la precipitación en esta área de la cuenca tiende a disminuir a través del tiempo. La zona de La Misión también cuenta con estaciones de tendencia negativa, de igual manera que Valle de Guadalupe, en el mapa de ΔP (Figura 17) se puede observar que el área de Ojos Negros es la que presenta aumento en las precipitaciones a través del tiempo, esto con respecto a las estaciones con ΔP positivo.

En cuanto a los ΔE , Ojos Negros es el área de la cuenca que presenta mayor disminución de la evaporación a través del tiempo, es decir, la mayoría de las estaciones que se ubican en dicha área presentan ΔE negativos. Sin embargo, el mapa de ΔE (Figura 19) muestra que las estaciones cercanas a la Sierra presentan una tendencia positiva al aumento de evaporación a través del tiempo. El área de Valle de Guadalupe también presenta ambos casos, las estaciones 2001, 2005, 2036, 2024 y 2049 presentan una tendencia positiva, lo que indica el aumento de la evaporación a través del tiempo, sin embargo, en la Figura 19 es posible visualizar que también existen áreas con tendencia negativa dentro de Valle de Guadalupe, al igual que para la zona de La Misión, en donde las estaciones que colindan con el Océano Pacífico presentan un aumento en la evaporación a través del tiempo y las que se ubican hacia el norte presentan tendencia negativa, indicando disminución de la evaporación a través del tiempo.

En términos generales, con base en la Tabla 8 se puede concluir que para el ΔT , el 57% de las estaciones climatológicas consideradas en el estudio reflejan una tendencia negativa, indicando que las temperaturas tienden a disminuir a través del tiempo, de la misma manera el 51% de las estaciones presentan un ΔP negativo, indicando de la misma manera que en la temperatura, que la precipitación en la zona de la cuenca tiende a disminuir a través del tiempo, y finalmente el ΔE en la zona refleja que el 50% de las estaciones en dicha zona también presentan valores negativos, indicando así también una disminución en la evaporación en la cuenca Guadalupe.

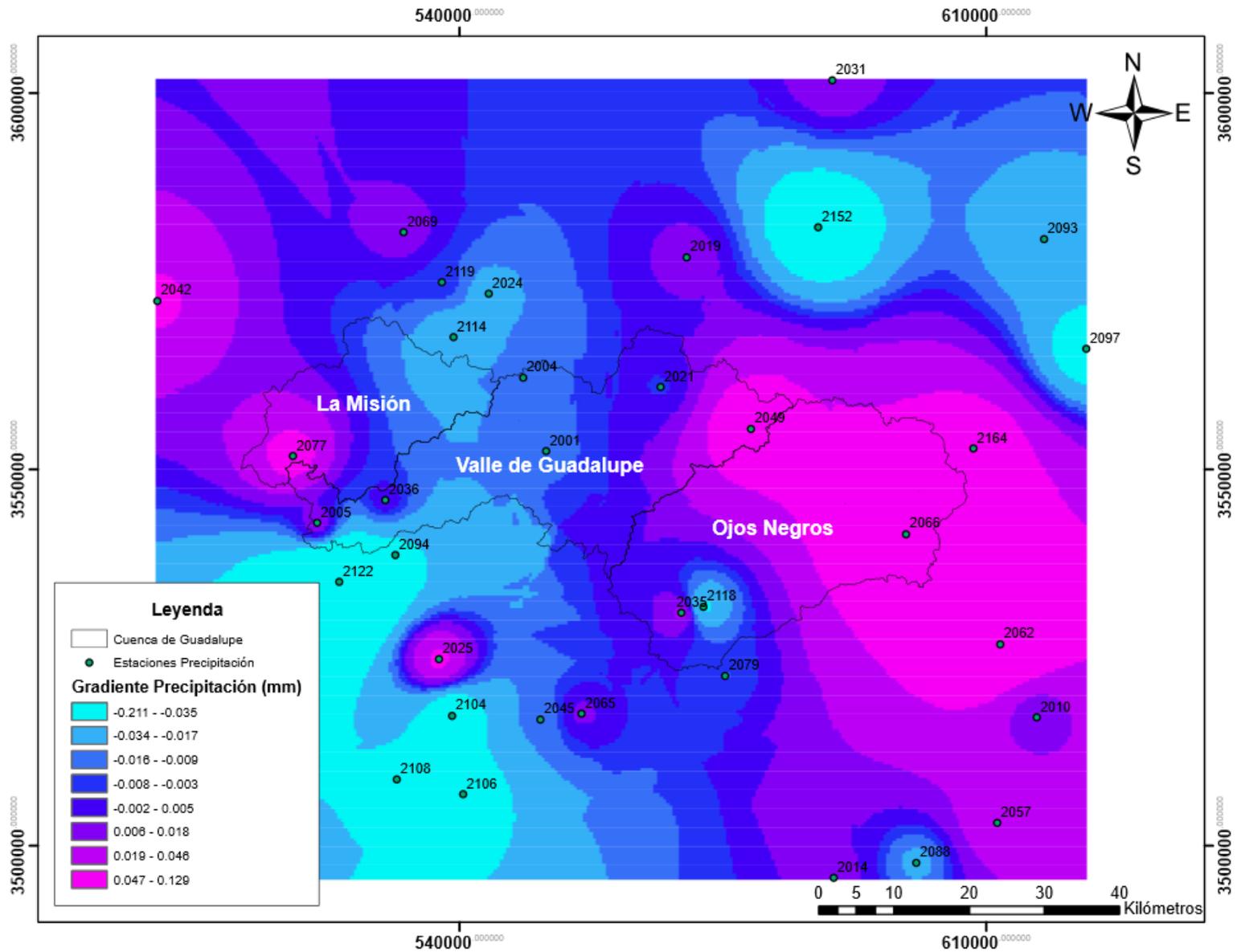


Figura 17. Mapa de gradientes de precipitación total mensual ΔP de cuenca Guadalupe

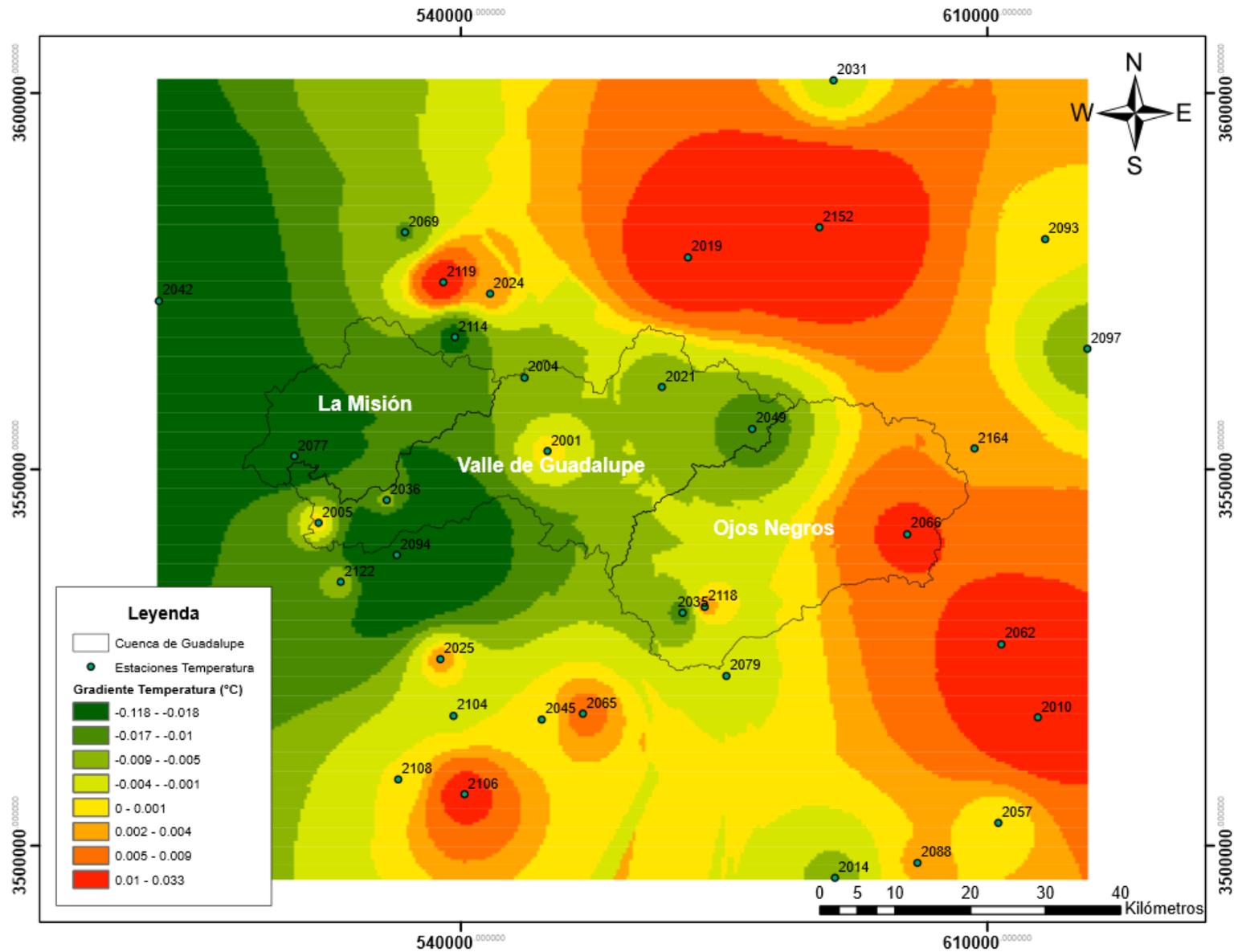


Figura 18. Mapa de gradientes de temperatura mensual media ΔT de cuenca Guadalupe

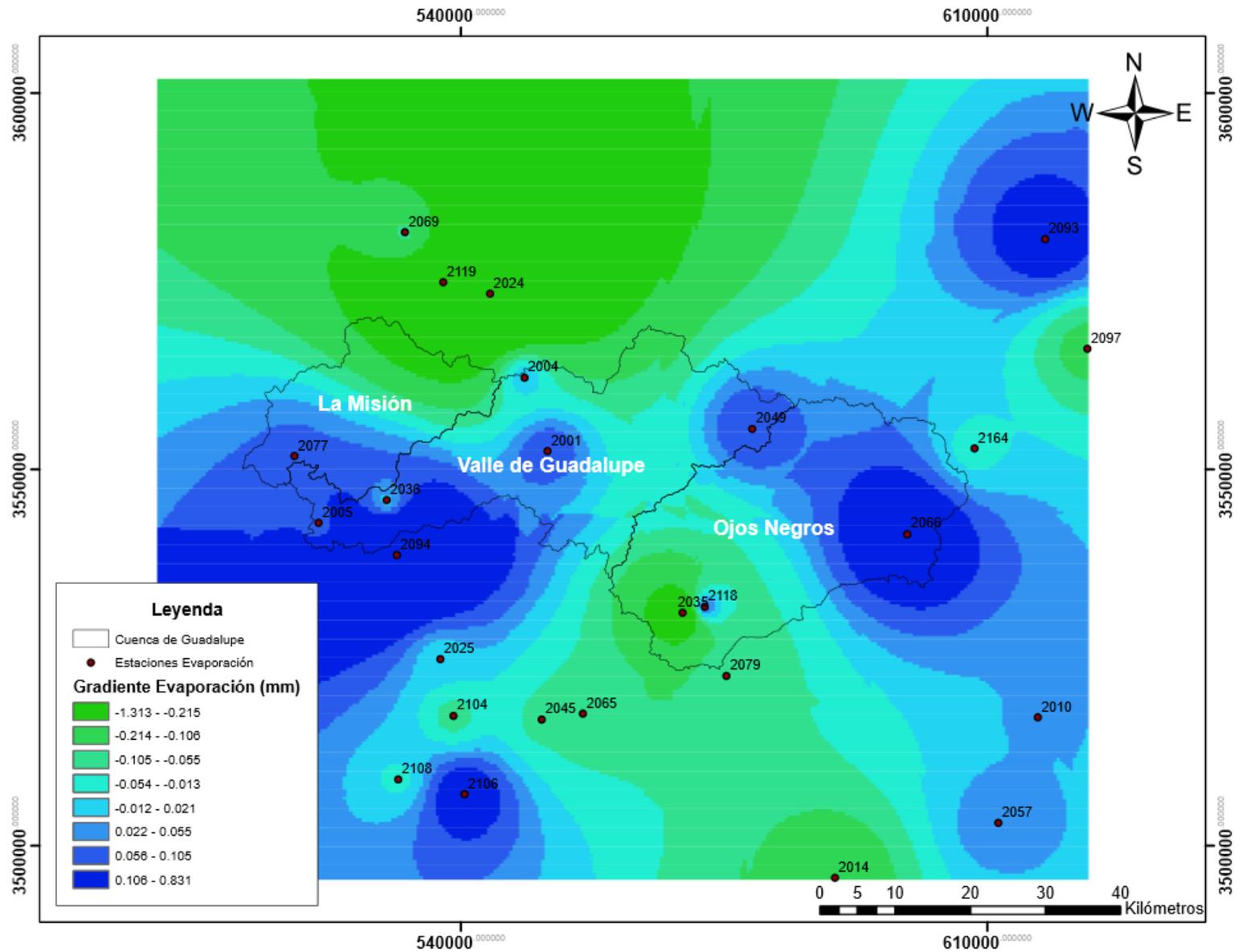


Figura 19. Mapa de gradientes de evaporación total mensual ΔE de cuenca Guadalupe

4.2.3 Análisis de Matriz de correlación de gradientes, variables climáticas y características fisiográficas de la cuenca Guadalupe

Tabla 9 Matriz de Correlación de gradientes de las variables climáticas en estudio, promedios anuales de las variables y características fisiográficas

VARIABLE	Altitud	Longitud	Latitud	ΔT	ΔP	ΔE	Prec.	Temp.	Evap.
Altitud	1.0000	0.6007	0.1114	0.2845	0.2736	0.0706	0.4393	-0.3012	0.0196
Longitud	0.6007	1.0000	-0.0843	0.3813	0.2318	0.0364	-0.0245	0.1145	0.4875
Latitud	0.1114	-0.0843	1.0000	-0.0530	0.0195	-0.217	-0.0622	-0.1069	0.2653
ΔT	0.2845	0.3813	-0.0530	1.0000	0.0136	-0.4746	0.0149	-0.1323	0.2873
ΔP	0.2736	0.2318	0.0195	0.0136	1.0000	0.0948	0.2497	-0.1199	-0.0298
ΔE	0.0706	0.0364	-0.2170	-0.4746	0.0948	1	0.1841	0.1361	-0.227
Prec.	0.4393	-0.0245	-0.0622	0.0149	0.2497	0.1841	1	-0.5935	-0.4651
Temp.	-0.3012	0.1145	-0.1069	-0.1323	-0.1199	0.1361	-0.5935	1	0.1016
Evap.	0.0196	0.4875	0.2653	0.2873	-0.0298	-0.227	-0.4651	0.1016	1

En términos generales y con base a los resultados obtenidos en la matriz de correlación (Tabla 9), se puede observar que no existe una relación entre la altitud, el ΔE y la evaporación, caso contrario, se presenta entre la altitud, el ΔP y la precipitación en donde a pesar de ser coeficientes de correlación pequeños nos muestran que existe una relación, en la que la altitud influye en el ΔP , es decir que a mayor altitud, mayor será el ΔP , de la misma manera para la variable de precipitación, en donde existe una correlación positiva, lo que nos indica que mientras exista mayor altitud la precipitación tendera a aumentar, en este sentido, en la Figura 17 este comportamiento es evidente ya que las estaciones que presentan altitudes grandes se encuentran ubicadas en zonas donde el ΔP es positivo y tiende a aumentar, por ejemplo:

- San Juan de Dios norte (2049) a 1280 msnm
- Santa Catarina norte (2057) a 1150 msnm
- Santa Isabel (2062) a 1540 msnm y
- Sierra de Juárez (2066) a 1580 msnm

En lo que corresponde a la altitud, el ΔT y la temperatura, también se puede observar que existe una relación, para el caso de altitud- ΔT la relación es positiva, indicando que a mayor altitud, el ΔT tiende a aumentar, para altitud-temperatura la relación es inversa lo que nos dice que si la altitud es mayor, la temperatura tiende a disminuir y viceversa. Esto apoya el comportamiento observado en la Figura 18, en donde las estaciones 2057, 2062, y 2066 se encuentran en áreas donde el ΔT es positivo y tiende a aumentar. Ahora bien, considerando las características fisiográficas es evidente que existe una relación positiva entre altitud-longitud y altitud-latitud, es decir,

que a mayor altitud la longitud también será mayor, de igual manera a mayor altitud, mayor latitud.

En lo que respecta a la longitud, la matriz muestra que no existe relación longitud-latitud, caso contrario, es el que presenta la longitud- ΔT en donde el coeficiente de correlación muestra una relación positiva, que a pesar de no ser muy grande nos dice que la longitud influye en el ΔT , es decir, a mayor longitud el ΔT tiende a crecer. Esta relación se refleja entre longitud-temperatura, en donde también la relación es positiva lo que apoya a la relación longitud- ΔT , de esta manera, mientras aumente la longitud, la temperatura también aumentara. En cuanto a la relación longitud- ΔP , se puede observar que es positiva, por lo tanto, si aumenta la longitud aumenta el ΔP , sin embargo, para la longitud-precipitación la relación es despreciable. Con respecto al coeficiente obtenido en longitud- ΔE se puede decir que no existe una relación, sin embargo, a pesar de que estas no presentan una relación, la longitud-evaporación si la presentan, lo que indica que la longitud influye de manera positiva en la evaporación, es decir, el aumento de la longitud genera un aumento en la evaporación.

Por otra parte, la matriz nos muestra que no existe relación latitud-longitud, latitud- ΔP , latitud-precipitación y latitud- ΔT . En el caso de latitud-altitud la relación es positiva, al igual que la relación latitud-evaporación, es decir que la latitud influye en la altitud y en la evaporación, a medida que esta aumente, la altitud y la evaporación también aumentarán, sin embargo, la relación latitud- ΔE es negativa, indicando que a mayor latitud, el ΔE tiende a disminuir en el tiempo. En relación a latitud-temperatura se observa una relación inversa, a mayor latitud, la temperatura tendera a disminuir y viceversa.

Considerando los gradientes de cada variable, existe una relación positiva entre ΔT -altitud y ΔT -longitud, es decir, la altitud y longitud influyen de manera positiva para el ΔT . Sin embargo, en el caso de la temperatura y el ΔE , el ΔT presenta una relación inversa. Para ΔT -evaporación la relación es positiva, por lo tanto, se puede decir que el ΔT influye de manera positiva en la evaporación.

4.3 Análisis de función de autocorrelación ACF y función de autocorrelación parcial PACF

Después de realizar el ajuste y análisis de gradientes, se procedió a estimar la ACF y PACF para cada una de las variables en estudio en cada una de las estaciones. La Tabla 10 muestra los resultados obtenidos para ACF y PACF de la estación Aguacaliente (2001) para la serie de precipitación total mensual, así como los valores del estadístico Q de Ljung-Box y las probabilidades asociadas al mismo.

Tabla 10 Función de autocorrelación ACF y autocorrelación parcial PACF de estación 2001 para serie de precipitación total mensual observada

Lag	ACF	PACF	Q-Stat	Prob.	Lag	ACF	PACF	Q-Stat	Prob.
1	0.43	0.43	97.631	0	51	-0.05	0.005	1118.6	0
2	0.216	0.037	122.2	0	52	-0.151	-0.012	1131.9	0
3	-0.02	-0.154	122.4	0	53	-0.202	0.012	1155.8	0
4	-0.133	-0.101	131.77	0	54	-0.232	-0.065	1187.2	0
5	-0.22	-0.122	157.41	0	55	-0.198	0.008	1210.2	0
6	-0.245	-0.108	189.27	0	56	-0.134	0.002	1220.8	0
7	-0.228	-0.087	217.06	0	57	-0.022	0.005	1221.1	0
8	-0.148	-0.029	228.83	0	58	0.14	0.079	1232.6	0
9	-0.051	0.001	230.23	0	59	0.259	0.019	1272.3	0
10	0.122	0.123	238.24	0	60	0.345	0.125	1343	0
11	0.266	0.156	276.17	0	61	0.29	0.036	1393	0
12	0.286	0.066	320.29	0	62	0.184	0.047	1413.2	0
13	0.266	0.077	358.48	0	63	0.013	0.02	1413.3	0
14	0.141	-0.009	369.21	0	64	-0.102	0.002	1419.5	0
15	-0.042	-0.105	370.17	0	65	-0.211	-0.048	1446.3	0
16	-0.177	-0.083	387.1	0	66	-0.192	0.047	1468.5	0
17	-0.206	0.002	410.22	0	67	-0.179	0.01	1487.8	0
18	-0.23	-0.044	438.93	0	68	-0.118	0.005	1496.1	0
19	-0.211	-0.049	463.14	0	69	-0.023	-0.009	1496.5	0
20	-0.177	-0.066	480.31	0	70	0.067	-0.079	1499.2	0
21	-0.074	-0.03	483.28	0	71	0.162	-0.015	1515.2	0
22	0.137	0.129	493.52	0	72	0.253	0.035	1554.3	0
23	0.249	0.072	527.57	0	73	0.173	-0.048	1572.5	0
24	0.362	0.14	599.99	0	74	0.072	-0.026	1575.7	0
25	0.293	0.038	647.55	0	75	-0.044	-0.012	1576.9	0
26	0.13	-0.051	656.91	0	76	-0.138	-0.007	1588.7	0
27	-0.034	-0.051	657.55	0	77	-0.203	-0.048	1614.1	0
28	-0.173	-0.068	674.26	0	78	-0.213	0.009	1642.1	0
29	-0.223	-0.006	701.84	0	79	-0.195	-0.004	1665.7	0
30	-0.236	-0.014	732.8	0	80	-0.117	-0.006	1674.2	0
31	-0.239	-0.054	764.62	0	81	-0.03	-0.003	1674.8	0
32	-0.138	0.019	775.34	0	82	0.106	0.022	1681.8	0
33	-0.054	-0.028	776.96	0	83	0.255	0.084	1722.6	0
34	0.086	0.014	781.15	0	84	0.258	-0.001	1764.3	0
35	0.27	0.111	822.39	0	85	0.206	-0.042	1790.8	0
36	0.303	0.016	874.14	0	86	0.103	0	1797.5	0
37	0.301	0.076	925.5	0	87	-0.015	0.025	1797.7	0
38	0.11	-0.071	932.39	0	88	-0.135	-0.034	1809.2	0
39	-0.039	-0.043	933.25	0	89	-0.211	-0.013	1837.5	0
40	-0.133	0.026	943.38	0	90	-0.229	-0.023	1870.7	0
41	-0.197	-0.01	965.43	0	91	-0.207	-0.024	1897.9	0
42	-0.223	-0.012	993.79	0	92	-0.118	0.016	1906.8	0
43	-0.21	-0.023	1019.1	0	93	-0.05	-0.06	1908.4	0
44	-0.128	0.032	1028.5	0	94	0.032	-0.076	1909	0
45	-0.035	0.003	1029.2	0	95	0.156	0.018	1924.5	0
46	0.111	0.014	1036.3	0	96	0.21	-0.023	1952.9	0
47	0.197	-0.02	1058.7	0	97	0.219	0.019	1983.7	0
48	0.266	0.009	1099.6	0	98	0.064	-0.088	1986.4	0
49	0.164	-0.097	1115.3	0	99	-0.029	-0.004	1987	0
50	0.057	-0.039	1117.2	0	100	-0.142	-0.023	2000.1	0

La ACF de la estación 2001 (Figura 20) muestra un pico grande que sobrepasa los límites significativos lo cual indica que la correlación para ese desfase (lag) no es igual a cero, dicho pico esta seguido de una oscilación amortiguada que se alterna entre correlaciones positivas y negativas en donde varios coeficientes sobrepasan los límites.

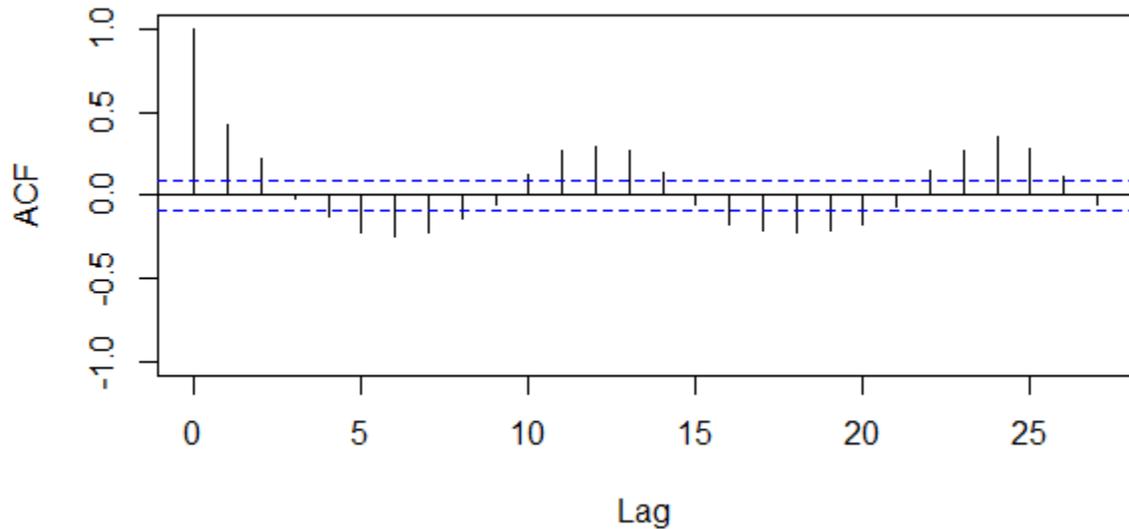


Figura 20. Función de autocorrelación ACF de estación 2001 para serie de precipitación total mensual observada

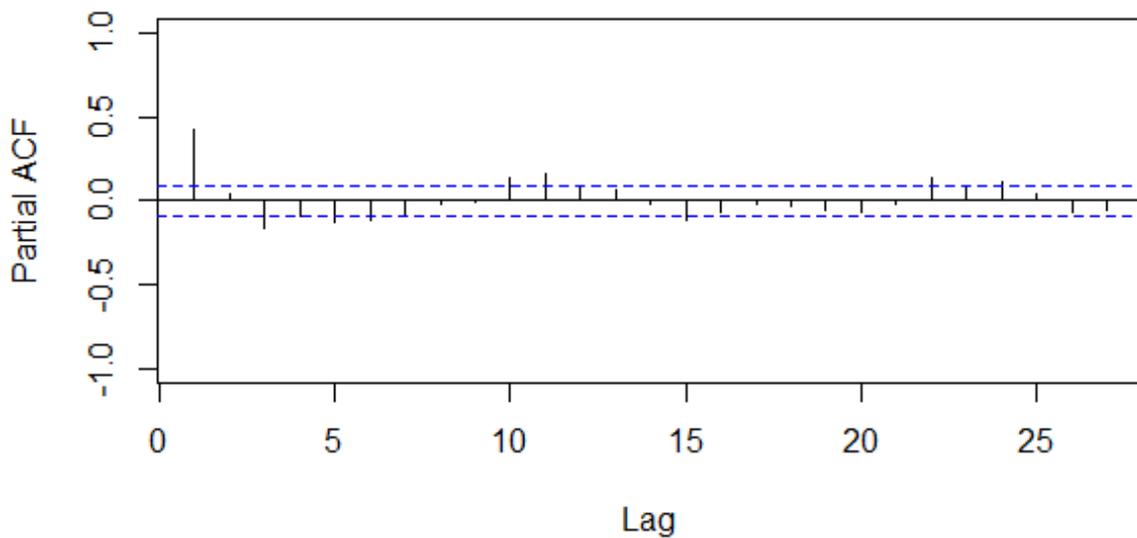


Figura 21. Función de autocorrelación parcial PACF de estación 2001 para serie de precipitación total mensual observada

En la Figura 21, la PACF muestra un pico grande en el desfase 1 seguido de una oscilación amortiguada que se alterna entre correlaciones positivas y negativas, en este caso la estación muestra en su mayoría coeficientes distintos a cero, lo que nos confirma al igual que la Figura 20, que la precipitación total mensual para la estación 2001 no es estacionaria.

4.4 Resultados de análisis de consistencia

A continuación, se presentan los resultados obtenidos en el análisis de consistencia realizado para cada una de las estaciones con cada una de las variables en estudio. La Tabla 11, Tabla 12 y Tabla 13 presentan las diferentes pruebas estadísticas aplicadas para determinar algunas de las características principales de las series como lo son la homogeneidad, tendencia e independencia, etc.

Al analizar las tablas, específicamente para la serie de precipitación total mensual (Tabla 11), la prueba *Helmert* muestra que el 100% de las estaciones consideradas son homogéneas, la de *secuencias* nos dice que el 62.85% de las estaciones no son homogéneas y *t de Student* nos indica que el 57.14% no son homogéneas. En lo que respecta a la tendencia la prueba *Mann-Kendall* indica que el 100% de las estaciones analizadas presentan tendencia. Finalmente, la prueba *Anderson* muestra que solo el 25.72% de las estaciones presentan independencia.

Tabla 11 Análisis de consistencia para series de precipitación total mensual

Estación	Helmert	Secuencias	Mann-Kendall	Anderson	t de Student
2001	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2004	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2005	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2010	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2014	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2019	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2021	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2024	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2025	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2031	Homogéneo	No homogéneo	No existe error de tendencia	Independiente	No homogéneo
2035	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo

2036	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2042	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2045	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2049	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2057	Homogéneo	No homogéneo	No existe error de tendencia	Independiente	No homogéneo
2062	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2065	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2066	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2069	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2077	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2079	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2088	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2093	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2094	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2097	Homogéneo	No homogéneo	No existe error de tendencia	Independiente	Homogéneo
2104	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2106	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2108	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2114	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2118	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2119	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2122	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2152	Homogéneo	No homogéneo	No existe error de tendencia	Independiente	No homogéneo
2164	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo

En lo que respecta a las series de temperatura, la Tabla 12 nos muestra, que la prueba *Helmert* considera homogéneas al 94.28% de las estaciones analizadas, la de *secuencias* indica que el 82.85% son homogéneas y t de Student muestra que solo el 54.28 de las estaciones son homogéneas. Para la característica de independencia, la prueba Mann-Kendall nos dice que el 94.28% de las estaciones no presentan error de

tendencia, por último, para la independencia de las series, la prueba Anderson indica que solo el 22.85% de las estaciones presentan dicha característica.

Tabla 12 Análisis de consistencia para series de temperatura media mensual

Estación	Helmert	Secuencias	Mann-Kendall	Anderson	t de Student
2001	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2004	No homogéneo	Homogéneo	No existe error de tendencia	Independiente	No homogéneo
2005	Homogéneo	Homogéneo	Error de tendencia	Error de independencia	Homogéneo
2010	Homogéneo	No homogéneo	Error de tendencia	Error de independencia	No homogéneo
2014	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2019	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2021	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2024	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2025	No homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2031	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2035	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2036	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2042	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2045	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2049	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2057	Homogéneo	No homogéneo	No existe error de tendencia	Independiente	No homogéneo
2062	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2065	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2066	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2069	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2077	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2079	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2088	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2093	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo

2094	Homogéneo	No homogéneo	No existe error de tendencia	Independiente	No homogéneo
2097	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2104	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2106	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2108	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2114	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2118	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2119	Homogéneo	No homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2122	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	No homogéneo
2152	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2164	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo

Finalmente, en el análisis de consistencia de las series de evaporación total mensual, la Tabla 13 nos muestra que la prueba Helmert arrojó que el 42.30% de las estaciones no son homogéneas, la de secuencias indica que solo el 19.23% no lo son, mientras que la prueba t de Student muestra que el 42.30% no son homogéneas. Para la detección de tendencia, la prueba Mann-Kendall muestra que el 11.53% de las estaciones presentan error en tendencia, y para la característica de independencia, la prueba Anderson indica que el 34.61% de las series de evaporación en las estaciones son independientes.

Tabla 13 Análisis de consistencia para series de evaporación total mensual

Estación	Helmert	Secuencias	Mann-Kendall	Anderson	t de Student
2001	No homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2004	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2005	No homogéneo	Homogéneo	Error de tendencia	Error de independencia	No homogéneo
2010	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2014	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2024	Homogéneo	No homogéneo	Datos insuficientes	Datos insuficientes	No homogéneo
2025	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2035	Homogéneo	No homogéneo	No existe error de tendencia	Independiente	Homogéneo

2036	No homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2045	No homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2049	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2057	No homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2065	No homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2066	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2069	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2077	No homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2079	No homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2093	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2094	Homogéneo	No homogéneo	Datos insuficientes	Datos insuficientes	No homogéneo
2097	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2104	Homogéneo	Homogéneo	No existe error de tendencia	Independiente	Homogéneo
2106	No homogéneo	Homogéneo	Error de tendencia	Error de independencia	No homogéneo
2108	No homogéneo	No homogéneo	Error de tendencia	Datos insuficientes	No homogéneo
2118	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	Homogéneo
2119	Homogéneo	Homogéneo	No existe error de tendencia	Error de independencia	No homogéneo
2164	No homogéneo	No homogéneo	No existe error de tendencia	Independiente	Homogéneo

En este sentido, y debido a los resultados obtenidos en el análisis de consistencia, posteriormente se desarrolla el proceso de transformación de datos con el propósito de normalizarlos, proceso necesario para la construcción de modelos.

4.5 Normalización de series de tiempo

Cada una de las series de las variables en estudio para cada una de las estaciones fueron sometidas al proceso de transformación mediante Box-Cox y Box-Cox segunda potencia.

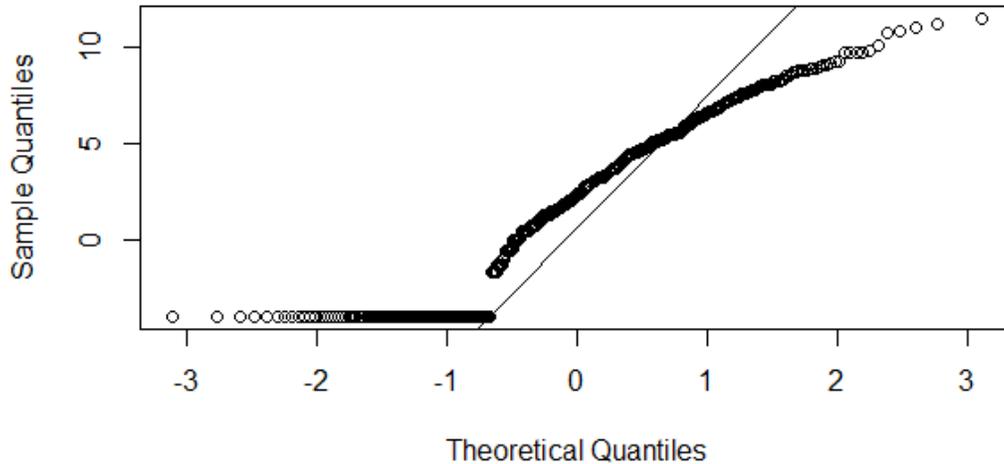


Figura 22. Gráfico de datos de precipitación mensual total de estación 2001, transformados mediante Box-Cox primera potencia

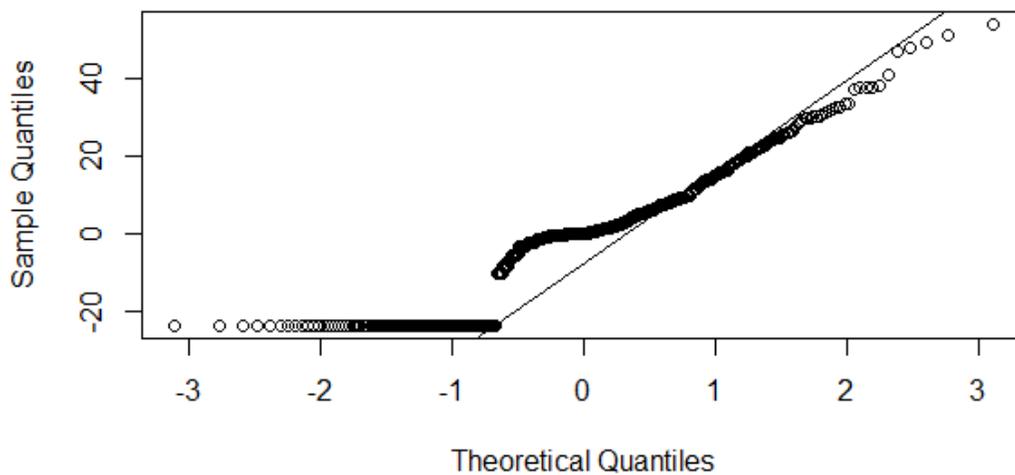


Figura 23. Gráfico de datos de precipitación mensual total de estación 2001, transformados mediante Box-Cox segunda potencia

La Figura 22 muestra los datos transformados mediante Box-Cox (ecuación(66)) para la serie de precipitación mensual total de la estación 2001. Como es evidente los datos en apariencia parecen no presentar una distribución normal, por esta razón, fueron sometidos a un segundo proceso de transformación Box-Cox segunda potencia mediante la ecuación (67), la Figura 23 muestra dicho proceso de transformación para la misma estación. Es importante destacar que, a pesar de que en la Figura 23 los datos no son normales aparentemente, los momentos estadísticos de la serie de datos transformados (Tabla 14) muestran un coeficiente de asimetría igual a cero y una curtosis igual a 3, lo que nos indica lo contrario.

Tabla 14 Estadísticos de serie de precipitación total mensual transformados mediante Box-Cox segunda potencia

Clave	Nombre de estación	Media	Mediana	Des. Est.	Coef. Asimetría	Curtosis	Coef. Variación	Varianza
2001	Agua Caliente	-0.3064	0.1222	15.86	0.3216	2.9282	-51.76	278.27

4.6 Análisis de la prueba Dickey-Fuller aumentada ADF

En cuanto los datos de las series son normalizados, se realizó la prueba Dickey-Fuller aumentada ADF para la serie de datos de precipitación total mensual transformados mediante Box-Cox de segunda potencia de la estación 2001. La Tabla 15 nos muestra el resultado de la prueba ADF en donde se observa que, en términos absolutos el estadístico de prueba (-13.81180) es mayor que los valores críticos de MacKinnon al 1%, 5% y 10%, lo que indica que la variable no presenta raíz unitaria, en consecuencia, se concluye que la serie de precipitación total mensual para la estación 2001 es estacionaria.

Así mismo, el ADF presenta una probabilidad menor al 5%, lo que indica que la serie es estacionaria, en lo que respecta a otras especificaciones de la prueba, se aprecia que el estadístico Durbin-Watson (1.98) es muy bajo, lo que nos dice que no hay evidencia de autocorrelación, por lo tanto, es posible generar conclusiones razonables en base a estos resultados y se puede iniciar el proceso de identificación y selección de modelos.

Tabla 15 Prueba Dickey-Fuller aumentada para serie de precipitación total mensual transformada de la estación 2001

Null Hypothesis: BOXCOX has a unit root
 Exogenous: Constant
 Lag Length: 6 (Automatic - based on SIC, maxlag=18)

	t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic	-13.81180	0.0000
Test critical values:		
1% level	-3.443123	
5% level	-2.867066	
10% level	-2.569775	

*Mackinnon (1996) one-sided p-values.

Augmented Dickey-Fuller Test Equation
 Dependent Variable: D(BOXCOX)
 Method: Least Squares
 Date: 05/21/18 Time: 17:11
 Sample (adjusted): 1969M08 2012M10
 Included observations: 503 after adjustments

Variable	Coefficient	Std. Error	t-Statistic	Prob.
BOXCOX(-1)	-1.209992	0.087606	-13.81180	0.0000
D(BOXCOX(-1))	0.489207	0.074057	6.605774	0.0000
D(BOXCOX(-2))	0.619305	0.066071	9.373342	0.0000
D(BOXCOX(-3))	0.482481	0.061756	7.812743	0.0000
D(BOXCOX(-4))	0.390574	0.056641	6.895623	0.0000
D(BOXCOX(-5))	0.314755	0.053438	5.890052	0.0000
D(BOXCOX(-6))	0.143074	0.044915	3.185473	0.0015
C	-0.294208	0.586454	-0.501673	0.6161

R-squared	0.402806	Mean dependent var	0.025023
Adjusted R-squared	0.394361	S.D. dependent var	16.89017
S.E. of regression	13.14441	Akaike info criterion	8.005647
Sum squared resid	85523.84	Schwarz criterion	8.072773
Log likelihood	-2005.420	Hannan-Quinn criter.	8.031980
F-statistic	47.69669	Durbin-Watson stat	1.985083
Prob(F-statistic)	0.000000		

4.7 Identificación y selección de modelos para cuenca Guadalupe

Determinada la estacionariedad de la serie de datos transformados, se inicia con el proceso de construcción de modelos. En esta etapa se estimaron 169 modelos ARMA para la estación 2001, la Tabla 16 muestra los modelos construidos para la variable de precipitación total mensual, así como los valores de distintos criterios de selección (Log-likelihood “LogL”, Akaike Information Criterion “AIC”, Schwarz Criterion “BIC” y Hannan-Quinn criterion “HQ”) para cada uno de los modelos.

Tabla 16 Criterio de selección de modelo para serie de precipitación total mensual de estación 2001

Model Selection Criteria Table				
Dependent Variable: Serie Transformada Precipitación total mensual 2001				
Sample: 1969M01 2012M10				
Included Observations:524				
Model	LogL	AIC*	BIC	HQ
(2,5) (0,0)	-2052.877765	7.839839	7.91282	7.868414
(8,8) (0,0)	-2043.908537	7.839956	7.985917	7.897107
(10,12) (0,0)	-2037.941725	7.840083	8.034697	7.916283
(4,4) (0,0)	-2052.479514	7.842127	7.923217	7.873877
(7,10) (0,0)	-2043.660239	7.842815	7.996884	7.90314
(9,10) (0,0)	-2041.782545	7.84328	8.013567	7.909955
(6,9) (0,0)	-2045.81304	7.843396	7.981248	7.897371
(4,2) (0,0)	-2054.828614	7.843455	7.908326	7.868855
(9,4) (0,0)	-2048.1997	7.844866	7.9665	7.892491
(8,4) (0,0)	-2049.321231	7.845328	7.958853	7.889778
(3,6) (0,0)	-2052.468928	7.845889	7.935088	7.880815
(4,5) (0,0)	-2052.638169	7.846533	7.935731	7.881458
(6,10) (0,0)	-2045.674611	7.846672	7.992632	7.903822
(9,6) (0,0)	-2046.768403	7.847028	7.98488	7.901003
(6,4) (0,0)	-2051.850944	7.847342	7.944649	7.885442
(9,8) (0,0)	-2045.029222	7.84802	8.00209	7.908345
(2,3) (0,0)	-2057.169821	7.848554	7.905317	7.870779
(3,5) (0,0)	-2054.628584	7.850299	7.931388	7.882049
(8,2) (0,0)	-2052.642634	7.850352	7.947659	7.888452
(5,4) (0,0)	-2053.850009	7.851141	7.940339	7.886066
(4,10) (0,0)	-2048.949691	7.85152	7.981263	7.90232
(8,5) (0,0)	-2050.132174	7.852214	7.973848	7.899839
(4,11) (0,0)	-2048.158775	7.852315	7.990167	7.90629
(11,9) (0,0)	-2043.232964	7.852597	8.030993	7.922447
(2,9) (0,0)	-2052.370598	7.85312	7.958536	7.894395
(4,9) (0,0)	-2050.480215	7.853537	7.975171	7.901162
(7,6) (0,0)	-2050.54158	7.85377	7.975404	7.901395
(7,9) (0,0)	-2047.574992	7.853897	7.999858	7.911047
(11,4) (0,0)	-2048.617223	7.854058	7.99191	7.908033
(3,9) (0,0)	-2051.862654	7.854991	7.968516	7.899441
(12,12) (0,0)	-2040.166936	7.856148	8.06698	7.938698
(3,4) (0,0)	-2057.169127	7.856156	7.929137	7.884731
(10,2) (0,0)	-2052.255198	7.856484	7.970009	7.900934
(5,8) (0,0)	-2051.325824	7.856752	7.978386	7.904377
(4,12) (0,0)	-2048.517008	7.857479	8.00344	7.914629
(9,12) (0,0)	-2043.751036	7.858369	8.044875	7.931394
(5,11) (0,0)	-2049.187545	7.860029	8.00599	7.917179
(10,4) (0,0)	-2051.239846	7.860228	7.989971	7.911028
(8,7) (0,0)	-2050.344729	7.860626	7.998478	7.914601
(2,2) (0,0)	-2061.820183	7.862434	7.911088	7.881484
(11,8) (0,0)	-2046.831253	7.862476	8.032764	7.929151
(2,12) (0,0)	-2051.875317	7.862644	7.992387	7.913444
(2,5) (0,0)	-2052.877765	7.839839	7.91282	7.868414
(8,8) (0,0)	-2043.908537	7.839956	7.985917	7.897107
(10,12) (0,0)	-2037.941725	7.840083	8.034697	7.916283
(4,4) (0,0)	-2052.479514	7.842127	7.923217	7.873877
(7,10) (0,0)	-2043.660239	7.842815	7.996884	7.90314
(9,10) (0,0)	-2041.782545	7.84328	8.013567	7.909955

(6,9) (0,0)	-2045.81304	7.843396	7.981248	7.897371
(4,2) (0,0)	-2054.828614	7.843455	7.908326	7.868855
(11,12) (0,0)	-2043.382549	7.864572	8.067296	7.943948
(5,12) (0,0)	-2049.574185	7.865301	8.019371	7.925626
(5,6) (0,0)	-2055.993125	7.866894	7.97231	7.908169
(7,8) (0,0)	-2052.477567	7.868736	8.006588	7.922711
(11,10) (0,0)	-2047.068257	7.870982	8.057488	7.944007
(11,1) (0,0)	-2076.572921	7.948946	8.062472	7.993397
(12,1) (0,0)	-2075.907535	7.950219	8.071853	7.997844
(12,0) (0,0)	-2077.307168	7.951738	8.065263	7.996188
(6,12) (0,0)	-2071.402338	7.9521	8.114279	8.0156
(10,1) (0,0)	-2079.558139	7.956495	8.061911	7.99777
(9,1) (0,0)	-2083.362948	7.967159	8.064467	8.00526
(8,1) (0,0)	-2086.324265	7.974617	8.063815	8.009542
(7,0) (0,0)	-2088.955021	7.977015	8.049996	8.00559
(10,0) (0,0)	-2086.038119	7.977331	8.074639	8.015431
(8,0) (0,0)	-2088.903018	7.98062	8.061709	8.01237
(7,1) (0,0)	-2088.920613	7.980687	8.061776	8.012437
(9,0) (0,0)	-2088.668186	7.983529	8.072728	8.018454
(6,1) (0,0)	-2090.805181	7.98405	8.057031	8.012625
(6,0) (0,0)	-2094.663196	7.994917	8.059789	8.020317
(5,1) (0,0)	-2095.014872	7.996254	8.061126	8.021654
(4,1) (0,0)	-2098.150448	8.004374	8.061137	8.026599
(3,1) (0,0)	-2100.328518	8.008854	8.057507	8.027904
(3,10) (0,0)	-2097.326036	8.031658	8.153292	8.079283
(1,12) (0,0)	-2097.726112	8.033179	8.154813	8.080804
(5,0) (0,0)	-2105.742382	8.033241	8.090004	8.055466
(1,11) (0,0)	-2098.832887	8.033585	8.14711	8.078035
(0,12) (0,0)	-2101.551692	8.043923	8.157448	8.088373
(0,10) (0,0)	-2103.877746	8.045163	8.14247	8.083263
(0,9) (0,0)	-2104.920568	8.045325	8.134524	8.08025
(1,10) (0,0)	-2102.947273	8.045427	8.150843	8.086702
(0,11) (0,0)	-2103.518723	8.0476	8.153016	8.088875
(1,7) (0,0)	-2107.237998	8.050335	8.131424	8.082085
(0,8) (0,0)	-2107.789072	8.05243	8.133519	8.08418
(0,7) (0,0)	-2109.118792	8.053684	8.126664	8.082259
(1,6) (0,0)	-2109.293484	8.054348	8.127328	8.082923
(1,9) (0,0)	-2107.272948	8.058072	8.155379	8.096172
(4,0) (0,0)	-2113.875486	8.060363	8.109017	8.079413
(1,8) (0,0)	-2109.006625	8.060862	8.15006	8.095787
(1,5) (0,0)	-2113.589566	8.06688	8.131752	8.092281
(1,4) (0,0)	-2116.45517	8.073974	8.130737	8.096199
(0,6) (0,0)	-2117.145855	8.080402	8.145274	8.105803
(1,3) (0,0)	-2121.81483	8.090551	8.139204	8.109601
(3,0) (0,0)	-2122.876669	8.090786	8.131331	8.106661
(2,1) (0,0)	-2123.756408	8.094131	8.134676	8.110006
(0,3) (0,0)	-2126.653563	8.105147	8.145691	8.121022
(1,2) (0,0)	-2126.852895	8.105905	8.146449	8.12178
(0,5) (0,0)	-2125.256499	8.107439	8.164202	8.129664
(0,2) (0,0)	-2128.528866	8.108475	8.140911	8.121175
(0,4) (0,0)	-2126.584089	8.108685	8.157338	8.127735
(2,8) (0,0)	-2121.155948	8.110859	8.208166	8.148959
(2,0) (0,0)	-2134.956374	8.132914	8.16535	8.145614
(1,0) (0,0)	-2136.910217	8.136541	8.160868	8.146066
(1,1) (0,0)	-2135.964932	8.136749	8.169185	8.149449
(2,6) (0,0)	-2131.932943	8.144232	8.225321	8.175982

(4,3) (0,0)	-2138.583919	8.165718	8.238699	8.194293
(3,8) (0,0)	-2140.585704	8.188539	8.293955	8.229814
(3,3) (0,0)	-2147.345006	8.195228	8.2601	8.220628
(0,1) (0,0)	-2152.925877	8.197437	8.221764	8.206962
(10,10) (0,0)	-2136.73436	8.208115	8.386512	8.277966
(11,3) (0,0)	-2144.917205	8.216415	8.346158	8.267215
(7,5) (0,0)	-2148.853627	8.223778	8.337303	8.268228
(8,11) (0,0)	-2145.908082	8.239194	8.409482	8.305869
(3,2) (0,0)	-2160.490433	8.241408	8.298171	8.263634
(12,8) (0,0)	-2148.225186	8.251807	8.430203	8.321657
(10,9) (0,0)	-2149.870902	8.254262	8.42455	8.320937
(9,2) (0,0)	-2159.947452	8.262158	8.367574	8.303433
(6,5) (0,0)	-2160.071835	8.262631	8.368047	8.303906
(11,0) (0,0)	-2160.951753	8.265976	8.371392	8.307251
(10,3) (0,0)	-2159.92712	8.269685	8.391319	8.31731
(6,3) (0,0)	-2164.294486	8.271082	8.36028	8.306007
(9,9) (0,0)	-2159.612115	8.287499	8.449677	8.350999
(12,3) (0,0)	-2162.771643	8.288105	8.425957	8.34208
(7,4) (0,0)	-2172.133883	8.308494	8.41391	8.349769
(11,2) (0,0)	-2171.727589	8.314554	8.436188	8.362179
(12,6) (0,0)	-2168.306529	8.320557	8.482736	8.384057
(12,2) (0,0)	-2175.439105	8.332468	8.462211	8.383268
(11,11) (0,0)	-2167.984864	8.334543	8.529158	8.410743
(11,7) (0,0)	-2172.667117	8.337137	8.499316	8.400637
(0,0) (0,0)	-2191.55725	8.340522	8.35674	8.346872
(12,5) (0,0)	-2176.789951	8.349011	8.503081	8.409336
(2,4) (0,0)	-2190.062926	8.357654	8.422525	8.383054
(10,7) (0,0)	-2179.334221	8.358685	8.512755	8.41901
(6,7) (0,0)	-2183.824328	8.360549	8.482183	8.408174
(5,2) (0,0)	-2190.502098	8.363126	8.436106	8.391701
(5,3) (0,0)	-2190.544883	8.367091	8.44818	8.398841
(6,2) (0,0)	-2191.017229	8.368887	8.449976	8.400637
(8,12) (0,0)	-2179.053928	8.369026	8.547423	8.438876
(2,7) (0,0)	-2191.471812	8.374418	8.463616	8.409343
(7,2) (0,0)	-2191.481526	8.374454	8.463653	8.40938
(7,12) (0,0)	-2181.492651	8.374497	8.544784	8.441172
(8,9) (0,0)	-2184.090745	8.376771	8.530841	8.437096
(5,5) (0,0)	-2191.235062	8.37732	8.474627	8.41542
(4,6) (0,0)	-2191.320938	8.377646	8.474953	8.415746
(3,7) (0,0)	-2191.519	8.378399	8.475706	8.416499
(7,3) (0,0)	-2191.523927	8.378418	8.475725	8.416518
(8,3) (0,0)	-2191.779912	8.383194	8.48861	8.424469
(4,7) (0,0)	-2191.793165	8.383244	8.48866	8.424519
(5,9) (0,0)	-2189.131856	8.384532	8.514275	8.435332
(2,10) (0,0)	-2191.514599	8.385987	8.499512	8.430437
(9,3) (0,0)	-2191.802321	8.387081	8.500606	8.431531
(2,0) (0,0)	-2134.956374	8.132914	8.16535	8.145614
(1,0) (0,0)	-2136.910217	8.136541	8.160868	8.146066
(1,1) (0,0)	-2135.964932	8.136749	8.169185	8.149449
(2,6) (0,0)	-2131.932943	8.144232	8.225321	8.175982
(4,3) (0,0)	-2138.583919	8.165718	8.238699	8.194293
(3,8) (0,0)	-2140.585704	8.188539	8.293955	8.229814
(3,3) (0,0)	-2147.345006	8.195228	8.2601	8.220628
(0,1) (0,0)	-2152.925877	8.197437	8.221764	8.206962
(10,10) (0,0)	-2136.73436	8.208115	8.386512	8.277966
(11,3) (0,0)	-2144.917205	8.216415	8.346158	8.267215

(7,5) (0,0)	-2148.853627	8.223778	8.337303	8.268228
(8,11) (0,0)	-2145.908082	8.239194	8.409482	8.305869
(3,2) (0,0)	-2160.490433	8.241408	8.298171	8.263634
(12,8) (0,0)	-2148.225186	8.251807	8.430203	8.321657
(10,9) (0,0)	-2149.870902	8.254262	8.42455	8.320937
(9,2) (0,0)	-2159.947452	8.262158	8.367574	8.303433
(6,5) (0,0)	-2160.071835	8.262631	8.368047	8.303906
(11,0) (0,0)	-2160.951753	8.265976	8.371392	8.307251
(10,3) (0,0)	-2159.92712	8.269685	8.391319	8.31731
(6,3) (0,0)	-2164.294486	8.271082	8.36028	8.306007
(9,9) (0,0)	-2159.612115	8.287499	8.449677	8.350999
(12,3) (0,0)	-2162.771643	8.288105	8.425957	8.34208
(7,4) (0,0)	-2172.133883	8.308494	8.41391	8.349769
(11,2) (0,0)	-2171.727589	8.314554	8.436188	8.362179
(12,6) (0,0)	-2168.306529	8.320557	8.482736	8.384057
(12,2) (0,0)	-2175.439105	8.332468	8.462211	8.383268
(11,11) (0,0)	-2167.984864	8.334543	8.529158	8.410743
(11,7) (0,0)	-2172.667117	8.337137	8.499316	8.400637
(0,0) (0,0)	-2191.55725	8.340522	8.35674	8.346872
(12,5) (0,0)	-2176.789951	8.349011	8.503081	8.409336
(2,4) (0,0)	-2190.062926	8.357654	8.422525	8.383054
(10,7) (0,0)	-2179.334221	8.358685	8.512755	8.41901
(6,7) (0,0)	-2183.824328	8.360549	8.482183	8.408174
(5,2) (0,0)	-2190.502098	8.363126	8.436106	8.391701
(5,3) (0,0)	-2190.544883	8.367091	8.44818	8.398841
(6,2) (0,0)	-2191.017229	8.368887	8.449976	8.400637
(8,12) (0,0)	-2179.053928	8.369026	8.547423	8.438876
(2,7) (0,0)	-2191.471812	8.374418	8.463616	8.409343
(7,2) (0,0)	-2191.481526	8.374454	8.463653	8.40938
(7,12) (0,0)	-2181.492651	8.374497	8.544784	8.441172
(8,9) (0,0)	-2184.090745	8.376771	8.530841	8.437096
(5,5) (0,0)	-2191.235062	8.37732	8.474627	8.41542
(4,6) (0,0)	-2191.320938	8.377646	8.474953	8.415746
(3,7) (0,0)	-2191.519	8.378399	8.475706	8.416499
(7,3) (0,0)	-2191.523927	8.378418	8.475725	8.416518
(8,3) (0,0)	-2191.779912	8.383194	8.48861	8.424469
(4,7) (0,0)	-2191.793165	8.383244	8.48866	8.424519
(5,9) (0,0)	-2189.131856	8.384532	8.514275	8.435332
(2,10) (0,0)	-2191.514599	8.385987	8.499512	8.430437
(9,3) (0,0)	-2191.802321	8.387081	8.500606	8.431531
(4,8) (0,0)	-2192.057821	8.388053	8.501578	8.432503
(2,11) (0,0)	-2191.063887	8.388076	8.50971	8.435701
(5,7) (0,0)	-2192.226983	8.388696	8.502221	8.433146
(6,6) (0,0)	-2192.274285	8.388876	8.502401	8.433326
(11,5) (0,0)	-2188.321224	8.389054	8.535015	8.446204
(3,11) (0,0)	-2191.11123	8.392058	8.521801	8.442858
(12,9) (0,0)	-2184.125348	8.392112	8.578617	8.465137
(3,12) (0,0)	-2190.61199	8.393962	8.531814	8.447937
(9,5) (0,0)	-2192.49777	8.39733	8.527073	8.44813
(12,4) (0,0)	-2190.876154	8.398769	8.54473	8.455919
(8,6) (0,0)	-2192.999621	8.399238	8.528981	8.450038
(6,8) (0,0)	-2193.015649	8.399299	8.529042	8.450099
(7,7) (0,0)	-2193.219914	8.400076	8.529819	8.450876
(10,5) (0,0)	-2192.24959	8.400189	8.53804	8.454164
(5,10) (0,0)	-2192.269786	8.400265	8.538117	8.45424
(10,6) (0,0)	-2192.772567	8.405979	8.55194	8.463129

(11,6) (0,0)	-2192.322351	8.40807	8.56214	8.468395
(6,11) (0,0)	-2192.337857	8.408129	8.562199	8.468454
(9,7) (0,0)	-2193.500494	8.408747	8.554708	8.465897
(7,11) (0,0)	-2192.807237	8.413716	8.575894	8.477216
(12,7) (0,0)	-2192.297362	8.415579	8.585867	8.482254
(10,8) (0,0)	-2193.519535	8.416424	8.578603	8.479924
(8,10) (0,0)	-2193.522904	8.416437	8.578616	8.479937
(9,11) (0,0)	-2193.091998	8.422403	8.6008	8.492253
(10,11) (0,0)	-2192.845104	8.425267	8.611772	8.498292
(12,10) (0,0)	-2192.346653	8.427174	8.621788	8.503374
(12,11) (0,0)	-2191.896298	8.429263	8.631987	8.508639

Una vez estimados los modelos, se procedió a seleccionar el modelo que mejor se ajuste a la variable en estudio tomando en consideración el valor mínimo de AIC (ecuación (60)) y el principio de Parsimonia. La Figura 24 muestra los veinte modelos que cuentan con el valor mínimo de AIC para la estación 2001. En términos generales, con base al AIC el modelo elegido para dicha estación es el modelo ARMA (2,5) que presenta un AIC de 7.8398 y con la menor cantidad de parámetros.

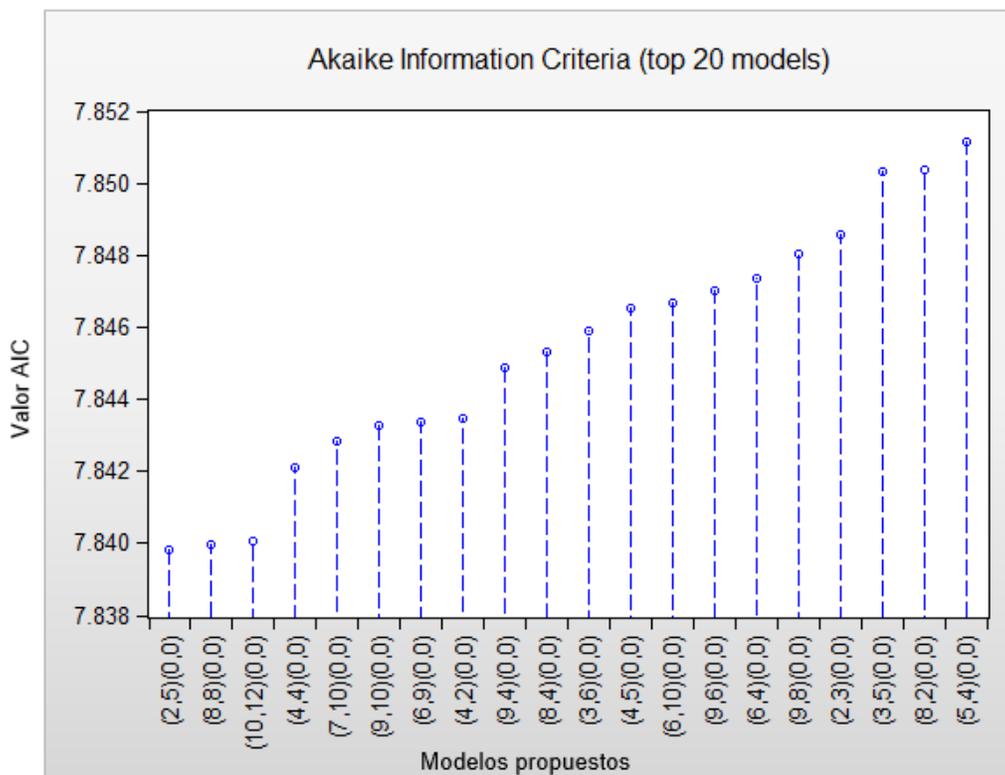


Figura 24. Representación gráfica de modelos propuestos con base al valor mínimo de AIC de estación 2001 para serie de precipitación total mensual

4.7.1 Modelos Autorregresivos de la cuenca Guadalupe

La selección de cada uno de los modelos para las variables en estudio y las estaciones consideradas en esta investigación se presentan en la Tabla 17. En lo que respecta a la precipitación total mensual, el 51.42% de las estaciones están representadas por modelos ARMA, de la misma manera para la variable de temperatura mensual media con el 51.42% y el 65.38% de las estaciones para la evaporación total mensual, lo que indica que estos modelos reflejan de mejor manera el comportamiento de las tres variables en estudio.

Tabla 17 Modelos de series de tiempo seleccionados

#	Clave	Nombre de Estación	Modelos de Prec. Tot mensual	Modelos de Temp. Mensual media	Modelos de Evap. Tot mensual
1	2001	Aguacaliente	ARMA (2,5)	ARMA (10,11)	ARIMA (6,1,9)
2	2004	Ignacio Zaragoza-Belen	ARMA (6,9)	ARIMA (11,1,9)	ARMA (7,2)
3	2005	Boquilla Sta. Rosa de La Misión	ARIMA (3,1,3)	ARMA (11,10)	ARIMA (10,1,12)
4	2010	Compostela	ARMA (4,6)	ARMA (8,11)	ARMA (3,2)
5	2014	El Alamo	ARMA (9,9)	ARIMA (8,1,11)	ARMA (9,10)
6	2019	El Compadre	ARMA (8,9)	ARMA (6,10)	-
7	2021	El Pinal	ARMA (5,4)	ARIMA (11,1,12)	-
8	2024	El Testerazo	ARIMA (8,1,9)	ARMA (4,5)	ARMA (4,3)
9	2025	Ensenada	ARMA (4,4)	ARMA (6,3)	ARMA (2,2)
10	2031	La Rumorosa	ARIMA (10,1,12)	ARMA (8,5)	-
11	2035	Ojos Negros	ARIMA (7,1,7)	ARIMA (12,1,11)	ARMA (5,5)
12	2036	Olivares Mexicanos	ARMA (3,2)	ARIMA (9,1,9)	ARMA (9,10)
13	2042	Rosarito	-	ARIMA (9,1,12)	-
14	2045	San Carlos	ARMA (3,2)	ARMA (9,9)	ARIMA (6,1,11)
15	2049	San Juan de Dios Norte	ARIMA (10,1,10)	ARIMA (9,1,7)	ARMA (6,4)
16	2057	Sta. Catarina Norte	ARMA (10,8)	ARMA (2,4)	ARMA (6,12)
17	2062	Sta. Isabel	ARIMA (11,1,3)	ARMA (10,6)	-
18	2065	Santo Tomas	ARMA (3,2)	ARIMA (10,1,11)	ARIMA (9,1,10)
19	2066	Sierra de Juárez	ARIMA (9,1,12)	ARIMA (3,1,3)	ARIMA (5,1,6)
20	2069	Valle de Las Palmas	ARIMA (7,1,8)	ARIMA (7,1,9)	ARIMA (2,1,4)
21	2077	La Misión	ARMA (2,2)	ARIMA (9,1,8)	ARMA (9,2)
22	2079	El Alamar	ARMA (3,2)	ARMA (9,10)	ARMA (3,2)
23	2088	Ej. Héroes de la Independencia	ARIMA (8,1,11)	ARMA (10,11)	-
24	2093	Ej. Valle de la Trinidad	ARIMA (10,2,12)	ARMA (11,10)	ARIMA (10,1,7)
25	2094	El Farito	ARMA (4,4)	ARMA (2,2)	AR (1)
26	2097	Laguna Salada	-	ARMA (3,2)	ARMA (3,3)
27	2104	El Ciprés	ARIMA (3,1,6)	-	ARMA (11,10)
28	2106	Manadero	ARIMA (4,1,3)	ARIMA (8,1,7)	ARIMA (2,1,5)
29	2108	Punta Banda	ARIMA (5,1,5)	ARMA (12,10)	ARMA (9,5)
30	2114	Ejido Carmen Serdán	ARIMA (4,1,3)	ARIMA (7,1,5)	-
31	2118	Valle de San Rafael	ARMA (3,2)	ARIMA (3,1,3)	ARMA (5,6)
32	2119	Cancio	ARMA (4,2)	ARMA (11,4)	ARMA (12,4)
33	2122	Real del Castillo Viejo	ARMA (2,3)	ARMA (3,2)	-
34	2152	Ej. José María Pino Suarez	ARIMA (4,1,6)	ARIMA (10,1,10)	-
35	2164	Ej. El Porvenir	ARMA (3,2)	ARMA (5,2)	ARMA (4,3)

4.8 Estimación de parámetros del modelo

La estimación de parámetros del modelo ARMA (2,5) para la estación 2001, seleccionado con base al AIC y el principio de Parsimonia para la serie de precipitación total mensual se muestra en la Tabla 18. Representándolo conforme a la estructura de modelo ARMA mediante la ecuación (53) se define de la siguiente manera

$$X_t = \underbrace{-0.2595 + 1.7321X_{t-1} + (-0.9999X_{t-2})}_{\text{Comp. autorregresiva}} + \varepsilon_t + \underbrace{(-1.5914Z_{t-1}) + 0.8706Z_{t-2} + (-0.1160Z_{t-3}) + 0.2180Z_{t-4} + (-0.0609Z_{t-5})}_{\text{Comp. media móvil}}$$

Tabla 18 Parámetros de modelo seleccionado para serie de precipitación total mensual en estación 2001

Dependent Variable: AR2MA5BOXCOX
Method: ARMA Maximum Likelihood (BFGS)
Date: 06/10/18 Time: 22:12
Sample: 1969M01 2012M10
Included observations: 524
Failure to improve objective (non-zero gradients) after 84 iterations
Coefficient covariance computed using outer product of gradients

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	-0.259533	0.642300	-0.404069	0.6863
AR(1)	1.732107	0.000425	4078.267	0.0000
AR(2)	-0.999997	7.83E-05	-12777.46	0.0000
MA(1)	-1.591406	0.232521	-6.844139	0.0000
MA(2)	0.870629	0.252870	3.442988	0.0006
MA(3)	-0.116086	0.092954	-1.248859	0.2123
MA(4)	0.218042	0.112403	1.939829	0.0529
MA(5)	-0.060926	0.053361	-1.141767	0.2541
SIGMASQ	145.2803	26.23141	5.538412	0.0000
R-squared	0.421999	Mean dependent var	-0.306413	
Adjusted R-squared	0.413021	S.D. dependent var	15.86916	
S.E. of regression	12.15809	Akaike info criterion	7.869697	
Sum squared resid	76126.89	Schwarz criterion	7.942891	
Log likelihood	-2052.861	Hannan-Quinn criter.	7.898360	
F-statistic	47.00028	Durbin-Watson stat	1.992203	
Prob(F-statistic)	0.000000			
Inverted AR Roots	.87-.50i	.87+.50i		
Inverted MA Roots	.87+.50i	.87-.50i	.27	-.21+.43i
	-.21-.43i			

En lo que se refiere al estadístico Durbin-Watson con valor igual a 1.99 (ecuación (65), Tabla 18), este es ligeramente inferior al valor de referencia que es 2, lo que sugiere

que no existe un problema de autocorrelación importante en los coeficientes de los residuales del modelo, porque se puede concluir que el modelo es idóneo para realizar predicciones a futuro. Sin embargo, para confirmar dicha aseveración se requiere del diagnóstico de los residuales del modelo.

4.9 Diagnóstico del modelo

Al evaluar la significancia del modelo, en la Tabla 18, el estadístico de la F (de Fisher) es igual a 47.00, valor que sobrepasa al valor de tablas (convencionalmente se suele tomar como 4). Revisando además la probabilidad de error asociada, que presenta un valor de 0.0000, valor inferior al 5%. En consecuencia, se puede concluir que el modelo es significativo para explicar el comportamiento de la precipitación total mensual.

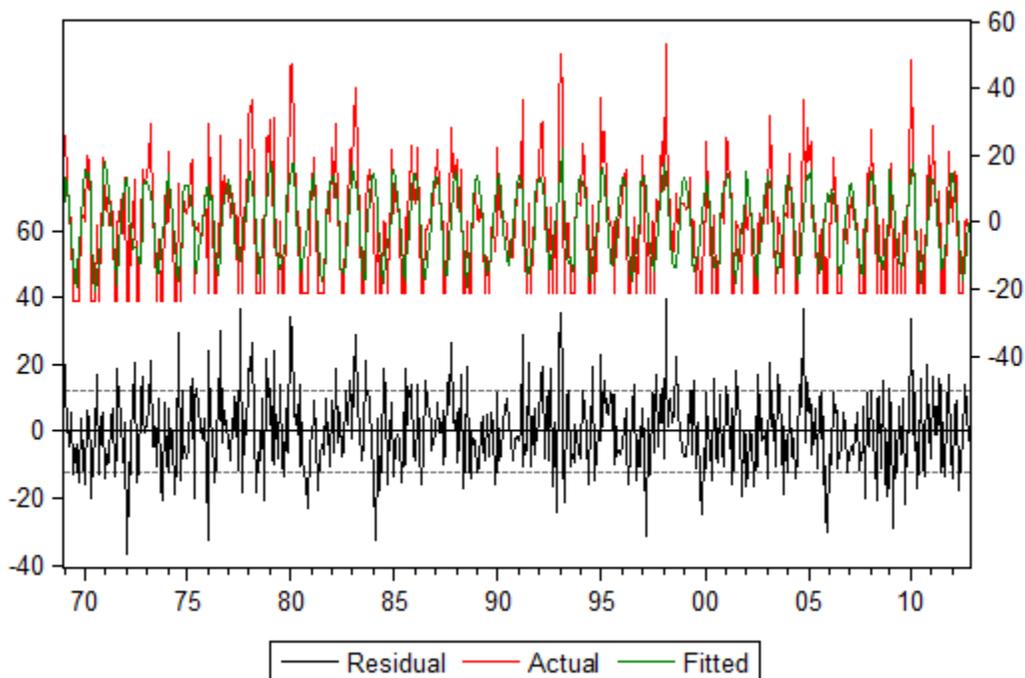


Figura 25. Representación gráfica de residuales obtenidos con el modelo ARMA (2,5) para datos de precipitación total mensual en la estación 2001

Ahora bien, en la Figura 25 se presenta la comparación entre el valor (transformado Box-Cox 2da. Potencia) de precipitación mensual total de la estación (Actual) y el valor estimado (Fitted) por el modelo ARMA (2,5), mostrando el residual obtenido para cada registro de la serie de precipitación total mensual de la estación 2001, tal como se puede apreciar, existen casos para los cuales el error de estimación sale de la banda

de aceptación deseable. Estos resultados corroboran la afirmación anterior en sentido de que, si bien el modelo explica buena parte del comportamiento de la precipitación total mensual, existe aún un porcentaje de dicho comportamiento que no ha logrado ser capturado en el modelo. Considerando las estimaciones de la ACF y PACF de los residuales a manera de evaluarlos, en la Tabla 19 se muestra que las probabilidades asociadas al estadístico Q son mayores que 0.05, por lo tanto, se concluye que los coeficientes de autocorrelación no son significativos y que los residuales son puramente aleatorios, lo cual se puede corroborar en la Figura 26 y Figura 27.

Tabla 19 Función de autocorrelación ACF y autocorrelación parcial PACF de residuales de precipitación total mensual en estación 2001

Lag	ACF	PACF	Q-Stat	Prob.
1	-0.000	-0.000	6.E-05	0.994
2	-0.006	-0.006	0.0174	0.991
3	0.001	0.001	0.0176	0.999
4	-0.008	-0.009	0.0559	1.000
5	0.034	0.034	0.6522	0.985
6	0.001	0.001	0.6526	0.995
7	-0.031	-0.031	1.1760	0.991
8	0.006	0.006	1.1952	0.997
9	-0.032	-0.031	1.7266	0.995
10	-0.012	-0.013	1.8047	0.998
11	0.014	0.013	1.9078	0.999
12	-0.021	-0.019	2.1559	0.999
13	0.052	0.052	3.6229	0.995
14	-0.007	-0.006	3.6458	0.997
15	-0.036	-0.034	4.3602	0.996
16	-0.103	-0.107	10.108	0.861
17	0.000	0.001	10.108	0.899
18	0.009	0.005	10.154	0.927
19	0.055	0.054	11.775	0.895
20	-0.020	-0.015	11.997	0.916
21	-0.110	-0.107	18.600	0.611
22	-0.015	-0.017	18.731	0.662
23	-0.003	-0.008	18.734	0.717
24	0.049	0.046	20.069	0.693
25	0.040	0.037	20.938	0.696
26	-0.028	-0.024	21.380	0.722
27	0.007	0.006	21.405	0.767
28	-0.034	-0.040	22.046	0.779
29	0.037	0.048	22.814	0.785
30	-0.025	-0.039	23.167	0.808
31	0.026	0.025	23.541	0.829
32	0.064	0.051	25.813	0.772
33	-0.032	-0.031	26.384	0.786
34	-0.004	0.013	26.395	0.821
35	0.010	0.014	26.446	0.850
36	-0.038	-0.047	27.278	0.852

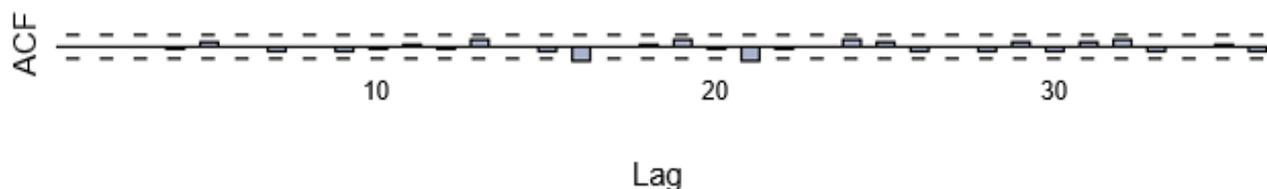


Figura 26. Función de autocorrelación ACF para residuales de precipitación total mensual, modelo ARMA (2,5) para la estación 2001

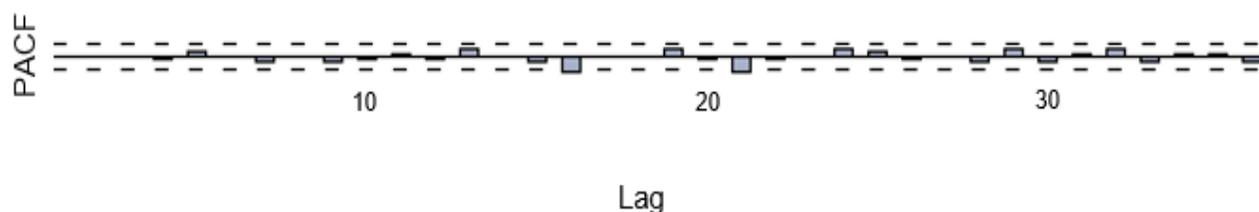


Figura 27. Función de autocorrelación parcial PACF para residuales de precipitación total mensual, modelo ARMA (2,5) para la estación 2001

Los resultados presentados en la Tabla 19 confirman que el modelo AR (2) MA (5) es una opción apropiada para efectuar pronósticos, sin embargo, podrían hacerse ajustes en el modelo que permitan explicar en mayor medida el comportamiento de la precipitación total mensual.

4.10 Espacialización de componentes de modelos autorregresivos y de media móvil

Mediante la Espacialización de los componentes de los modelos se visualiza el comportamiento de las variables analizadas en cada una de las estaciones de la cuenca Guadalupe considerando cada uno de los componentes de los modelos seleccionados.

5 CONCLUSIONES

En definitiva, el análisis y estudio de las variables meteorológicas a través del tiempo se ha convertido en una tarea muy importante, tanto para caracterizar el clima como para proyectar comportamientos futuros, que permitan dimensionar y evaluar escenarios de cambio con el fin de generar estrategias de adaptación/mitigación ante la variabilidad climática y el cambio climático. La utilización de modelos autorregresivos para dichos propósitos ha sido una pieza clave ya que en la actualidad las distintas técnicas empleadas para la generación de pronósticos se basan en la metodología Box-Jenkins, ya sea para evaluar la eficiencia de las técnicas propuestas, para entrenar algoritmos o para fusionarlas y obtener mejores resultados.

Es importante destacar que conocer el comportamiento y la evolución de las variables meteorológicas es esencial en regiones con cualquier tipo de clima, debido a que de ello dependerá la gestión integral de los recursos.

Desde esta perspectiva, el análisis de series de tiempo empleando los fundamentos de modelos autorregresivos y de media móvil, permiten caracterizar la estructura temporal y la ocurrencia de eventos de precipitación, temperatura y evaporación. La función de autocorrelación y autocovarianza representan una medición del grado de dependencia de las series en mención y pueden explicar el comportamiento temporal de dichas variables. La resolución temporal y la longitud de las series de tiempo de las variables en estudio constituyen factores determinantes al momento de realizar análisis de series temporales a partir de modelos autorregresivos.

La modelación anual y mensual de las series de precipitación mediante la utilización de modelos autorregresivos, permite establecer las siguientes conclusiones:

La planificación racional exige prever los sucesos del futuro que probablemente vayan a ocurrir y la previsión, a su vez, se suele basar en lo que ha ocurrido en el pasado. El análisis de series de tiempo es una técnica que permite hacer previsiones sobre el futuro con base en lo ocurrido en el pasado, ya que se han logrado preservar los estadísticos en los eventos de la precipitación anual y mensual con modelos del tipo ARMA.

Finalmente, el cambio climático y sus posibles consecuencias deben ser el origen minucioso de estudios, reflexiones y consideraciones de cambio de modelos de desarrollo.

6 REFERENCIAS

- Abdel-Aal, R. E. (2004). Hourly temperature forecasting using abductive networks. *Engineering Applications of Artificial Intelligence*, 17(5), 543–556. <https://doi.org/10.1016/j.engappai.2004.04.002>
- Abrahart, R. J., & See, L. (2000). Comparing neural network and autoregressive moving average techniques for the provision of continuous river flow forecasts in two contrasting catchments. *Hydrological Processes*, 14(11–12), 2157–2172.
- Abrahart, R. J., See, L., & Kneale, P. E. (1999). Using pruning algorithms and genetic algorithms to optimise network architectures and forecasting inputs in a neural network rainfall-runoff model. *Journal Of Hydroinformatics*, (Figure 1), 103–114.
- Aguirre Madrid, L. G., & Correal, M. E. (2005). Análisis de sistemas dinámicos no lineales utilizando modelos Narmax Polinomiales.
- 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>
- Akaike, H. (1976). Canonical Correlation Analysis of Time Series and the Use of an Information Criterion. *Mathematics in Science and Engineering*, 126(C), 27–96. [https://doi.org/10.1016/S0076-5392\(08\)60869-3](https://doi.org/10.1016/S0076-5392(08)60869-3)
- Akaike, H. (1978a). A New Look at the Bayes Procedure. *Biometrika*, 65(1), 53–59. Retrieved from <http://www.jstor.org/stable/2335276>
- Akaike, H. (1978b). On the Likelihood of a Time Series Model. *Journal of the Royal Statistical Society. Series D (The Statistician)*, 27(3/4), 217–235. Retrieved from <http://www.jstor.org/stable/2335276>
- Akpanta, A. C., Okorie, I. E., & Okoye, N. N. (2015). SARIMA Modelling of the Frequency of Monthly Rainfall in Umuahia , Abia State of Nigeria. *American Journal of Mathematics and Statistics*, 5(2), 82–87. <https://doi.org/10.5923/j.ajms.20150502.05>
- Alarcón Valverde, L. A., & Chuñir Panjón, L. (2012). *Evaluación de modelos econométricos alternativos de series de tiempo para el pronóstico de la inflación en el Ecuador en el corto plazo: periodo 2000-2010*.
- Aldaz Cárdenaz, J. W., García Díaz, J. R., & Quiñones Ramos, R. (2012). Parvovirus canina en la provincia Bolívar, Ecuador. Utilidad de los modelos Box-Jenkins para su análisis y predicción. *Rev. Salud Anim*, 34(3), 165–172.
- Alfaro, E. J., & Cid, L. (1999). Ajuste de un modelo VARMA para los campos de anomalías de precipitación en Centroamérica y los índices de los océanos Pacífico y Atlántico Tropical. *Atmósfera*, 12(4), 205–222.
- Aly, A., Pathak, C., Teegavarapu, R. S. V, Ahlquist, J., & Fuelberg, H. (2009). Evaluation of Improvised Spatial Interpolation Methods for Infilling Missing Precipitation Records. *World Environmental and Water Resources Congress*,

5914–5923. [https://doi.org/10.1061/41036\(342\)598](https://doi.org/10.1061/41036(342)598)

- ArcGis. (2018). Vencidad de IDW del punto seleccionado. Retrieved from <http://desktop.arcgis.com/es/arcmap/10.3/tools/spatial-analyst-toolbox/how-idw-works.htm>
- Babu, S. K., Karthikeyan, K., Ramanaiah, M. V., & Ramanah, D. (2011). Prediction of Rain-fall flow Time Series using Auto-Regressive Models. *Advances in Applied Science Research*, 2(2), 128–133. Retrieved from www.pelagiaresearchlibrary.com
- Bărbulescu, A., & Pelican, E. (2003). ARIMA models for the analysis of the Precipitation Evolution (pp. 221–226).
- Barnett, T. P., Hasselmann, K., Chelliah, M., Delworth, T., Hegerl, G., Jones, P., ... Tett, S. (1999). Detection and attribution of recent climate change: A status report. *Bulletin of the American Meteorological Society*, 80(12), 2631–2659. [https://doi.org/10.1175/1520-0477\(1999\)080<2631:DAAORC>2.0.CO;2](https://doi.org/10.1175/1520-0477(1999)080<2631:DAAORC>2.0.CO;2)
- Bhakar, S. R., Singh, R. V., Chhajed, N., & Bansal, A. K. (2006). Stochastic modeling of monthly rainfall at kota region. *ARPJ Journal of Engineering and Applied Sciences*, 1(3), 36–44.
- Bowden, R. S., & Clarke, B. R. (2012). A single series representation of multiple independent ARMA processes. *Journal of Time Series Analysis*, 33(2), 304–311. <https://doi.org/10.1111/j.1467-9892.2011.00766.x>
- Box, G. E. P., & Jenkins, G. M. (1970). *Time Series Analysis, Forecasting, and Control*. Francisco Holden-Day.
- Box, G. E. P., Jenkins, G. M., & Reinsel, G. C. (1976). *Time Series Analysis Forecasting and Control* (Third Edit).
- Bras, R. L., & Rodriguez-Iturbe, I. (1985). *Random functions and hydrology*. Courier Corporation.
- Brath, A., Burlando, P., & Rosso, R. (1988). Sensitivity analysis of real-time flood forecasting to on-line rainfall predictions. In *Selected Papers from the Workshop on Natural Disasters in European-Mediterranean Countries, Perugia, Italy* (pp. 469–488).
- Breinl, K., Turkington, T., & Stowasser, M. (2015). Simulating daily precipitation and temperature: A weather generation framework for assessing hydrometeorological hazards. *Meteorological Applications*, 22(3), 334–347. <https://doi.org/10.1002/met.1459>
- Burlando, P., Montanari, A., & Ranzi, R. (1996). Forecasting of storm rainfall by combined use of radar, rain gages and linear models. *Atmospheric Research*, 42(1), 199–216. [https://doi.org/10.1016/0169-8095\(95\)00063-1](https://doi.org/10.1016/0169-8095(95)00063-1)
- Cam, L. Le. (1961). A Stochastic Description of Precipitation. *On, Berkeley Symposium*

Statistics, Mathematical, 1960, xii+767.

- Campolo, M., Andreussi, P., & Soldati, a. (1999). River flood forecasting with a neural network model. *Water Resources Research*, 35(4), 1191–1197. <https://doi.org/10.1029/1998WR900086>
- Carollo L., M. C. Regresión Lineal Simple, Estadística FBA I Departamento de Estadística e Investigación Operativa 1–31. Universidad de Santiago de Compostela.
- Carvajal Escobar, Y., & Segura, J. B. M. (2005). Modelos multivariados de predicción de caudal mensual utilizando variables macroclimáticas. Caso de estudio Río Cauca, Colombia. *Revista Ingeniería y Competitividad*, 7(1), 18–32.
- Caskey, J. E. (1963). A Markov Chain Model for the Probability of Precipitation Occurrence in Intervals of Various Length. *Monthly Weather Review*, 91(6), 298–301. [https://doi.org/10.1175/1520-0493\(1963\)091<0298:AMCMFT>2.3.CO;2](https://doi.org/10.1175/1520-0493(1963)091<0298:AMCMFT>2.3.CO;2)
- Ceschia, M., Garfagnini, R., Lavenia, A., & Toppano, E. (1986). An analysis by ARMA stochastic processes of the daily average temperatures collected during 1984 in Udine and Lignano Sabbiadoro. *Il Nuovo Cimento C*, 9(4), 859–884. <https://doi.org/10.1007/BF02558084>
- Chang, X., Gao, M., Wang, Y., & Hou, X. (2012). Seasonal Autoregressive Integrated Moving Average Model for Precipitation Time Series. *Journal of Mathematics and Statistics*, 8(4), 500–505. <https://doi.org/10.3844/jmssp.2012.500.505>
- Chao, Z., Hua-sheng, H., Wei-min, B., & Luo-ping, Z. (2008). Robust recursive estimation of auto-regressive updating model parameters for real-time flood forecasting. *Journal of Hydrology*, 349(3–4), 376–382. <https://doi.org/10.1016/j.jhydrol.2007.11.011>
- Chapman, T. G. (1997). Stochastic models for daily rainfall in the Western Pacific. *Mathematics and Computers in Simulation*, 43(3–6), 351–358. [https://doi.org/10.1016/S0378-4754\(97\)00019-0](https://doi.org/10.1016/S0378-4754(97)00019-0)
- Chatfield, C. (2000). *Time-Series Forecasting*. Retrieved from <https://books.google.no/books?id=PFHMBQAAQBAJ>
- Chattopadhyay, S., & Chattopadhyay, G. (2010). Univariate modelling of summer-monsoon rainfall time series: Comparison between ARIMA and ARNN. *Comptes Rendus - Geoscience*, 342(2), 100–107. <https://doi.org/10.1016/j.crte.2009.10.016>
- Chattopadhyay, S., Jhajharia, D., & Chattopadhyay, G. (2011). Univariate modelling of monthly maximum temperature time series over northeast India: Neural network versus Yule-Walker equation based approach. *Meteorological Applications*, 18(1), 70–82. <https://doi.org/10.1002/met.211>
- Cheng, X., & Noguchi, M. (1996). Rainfall-runoff modelling by neural network approach. In *Proc. Of int. conf. on water resources and environment research: towards the*

21st century (pp. 29–31).

- Chica, A., Rey, F., & Aristizábal, J. (2012). Application of autoregressive model with exogenous inputs to identify and analyse patterns of solar global radiation and ambient temperature. *International Journal of Ambient Energy*, 33(4), 177–183. <https://doi.org/10.1080/01430750.2012.686197>
- Chin, D. A. (1995). A scale model of multivariate rainfall time series. *Journal of Hydrology*, 168(1), 1–15. [https://doi.org/10.1016/0022-1694\(94\)02668-2](https://doi.org/10.1016/0022-1694(94)02668-2)
- Chorro Gascó, J. L. (2000). Modelos autorregresivos de series temporales múltiples. *Psicothema*, 12(2), 132–134. Retrieved from <http://www.redalyc.org/articulo.oa?id=72797034>
- Cisneros-Iturbe, H. L., & Pelczer, I. J. (2007). Redes neuronales artificiales para pronosticar alturas de precipitación cada 15 minutos. *Ingeniería Hidráulica En México*, XXII(2003), 5–21.
- Croitoru, A. E., Driginei, D., Holobaca, I. H., & Dragota, C. S. (2012). Change-point analysis for serially correlated summit temperatures in the Romanian Carpathians. *Theoretical and Applied Climatology*, 108(1–2), 9–18. <https://doi.org/10.1007/s00704-011-0508-7>
- Dabral, P. P., & Murry, M. Z. (2017). Modelling and Forecasting of Rainfall Time Series Using SARIMA. *Environmental Processes*, 4(2), 399–419. <https://doi.org/10.1007/s40710-017-0226-y>
- Dabral, P. P., Pandey, A., Baithuri, N., & Mal, B. C. (2008). Stochastic modelling of rainfall in humid region of North East India. *Water Resources Management*, 22(10), 1395–1407. <https://doi.org/10.1007/s11269-007-9232-6>
- Dabral, P. P., Saring, T., & Jhajharia, D. (2016). Time Series Models of Monthly Rainfall and Temperature to detect climate change for Jorhat (Assam), India. *Global NEST Journal*, 18(3), 494–507.
- Dahale, S. D., & Singh, S. V. (1993). Modelling of Indian Monsoon Rainfall Series by Univariate Box --Jenkins Type of Models. *Advances in Atmospheric Sciences*, 10(2), 211–220.
- Daniell, T. M. (1991). Neural networks. Applications in hydrology and water resources engineering. In *National Conference Publication- Institute of Engineers. Australia*.
- Dastorani, M., Mirzavand, M., Dastorani, M. T., & Sadatinejad, S. J. (2016). Comparative study among different time series models applied to monthly rainfall forecasting in semi-arid climate condition. *Natural Hazards*, 81(3), 1811–1827. <https://doi.org/10.1007/s11069-016-2163-x>
- Dawson, C. W., Abrahart, R. J., Shamseldin, A. Y., & Wilby, R. L. (2006). Flood estimation at ungauged sites using artificial neural networks. *Journal of Hydrology*, 319(1), 391–409. <https://doi.org/10.1016/j.jhydrol.2005.07.032>

- Dawson, C. W., & Wilby, R. (1998). An artificial neural network approach to rainfall-runoff modelling. *Hydrological Sciences Journal*, 43(1), 47–66. <https://doi.org/10.1080/02626669809492102>
- Delleur, J. W., & Kavvas, M. L. (1978). Stochastic Models for Monthly Rainfall Forecasting and Synthetic Generation. *Journal of Applied Meteorology*, 17(10), 1528–1536. [https://doi.org/10.1175/1520-0450\(1978\)017<1528:SMFMRF>2.0.CO;2](https://doi.org/10.1175/1520-0450(1978)017<1528:SMFMRF>2.0.CO;2)
- Dinardo, J., Johnston, J., & Johnston, J. (1997). Econometric methods. *Forth Edition) McGraw-Hill Companies, Inc*, 204–326.
- Domínguez, R., Fuentes, M. G., & Arganis, M. L. (2001). Procedimiento para generar muestras sintéticas de series periódicas mensuales a través del método de Svanidze modificado aplicado a los datos de las presas La Angostura y Malpaso. Series del Instituto de Ingeniería, UNAM, Serie Blanca CI-19. Agosto.
- Douriet, J. C., Sanz, R., Ramirez, A. I., & Ravelo, A. C. (2015). Pronóstico de sequías hidrológicas utilizando redes neuronales univariadas y multivariadas en el Organismo de Cuenca del Pacífico.
- Durbin, J., & Watson, G. S. (1951). Testing for serial correlation in least squares regression. II. *Biometrika*, 38(1/2), 159–177.
- El-Fandy, M. G., Ashour, Z. H., & Taniel, S. M. M. (1994). Time series models adoptable for forecasting Nile floods and Ethiopian rainfalls. *Bulletin - American Meteorological Society*.
- Elder, J., & Kennedy, P. E. (2001). Testing for unit roots: What should students be taught? *Journal of Economic Education*, 32(2), 137–146. <https://doi.org/10.1080/00220480109595179>
- Escalante-Sandoval, C., & Reyes-Chávez, L. (2002). Técnicas Estadísticas en Hidrología. *Facultad de Ingeniería. Universidad Nacional Autónoma de México*.
- Escalante, C., & Reyes, L. (1998). Identificación y análisis de sequías en la Región Hidrológica Número 10, Sinaloa. *Ingeniería Hidráulica En México, XIII(2)*, 23–43.
- Etuk, E. H., Moffat, I. U., & Chims, B. E. (2013). Modelling Monthly Rainfall Data of Port Harcourt , Nigeria by Seasonal Box-Jenkins Methods. *International Journal of Sciences*, 2(1), 1–8.
- Fan, Y., Chen, Y., Li, W., Wang, H., & Li, X. (2011). Impacts of temperature and precipitation on runoff in the Tarim River during the past 50 years. *Journal of Arid Land*, 3(3), 220–230. <https://doi.org/10.3724/SP.J.1227.2011.00220>
- Farajzadeh, J., Fakheri Fard, A., & Lotfi, S. (2014). Modeling of monthly rainfall and runoff of Urmia lake basin using “feed-forward neural network” and “time series analysis” model. *Water Resources and Industry*, 7–8(2014), 38–48. <https://doi.org/10.1016/j.wri.2014.10.003>

- Ferrando, P. J., & Anguiano-Carrasco, C. (2010). El Análisis Factorial Como Técnica De Investigación En Psicología. *Papeles Del Psicólogo*, 31(1), 18–33. Retrieved from <http://www.redalyc.org/articulo.oa?id=77812441003>
- Fonseca, L. (2013). Aplicacion del Algebra de mapas en la Elaboracion de un Mapa de Susceptibilidad de Acuíferos por el método Drastic, usando ArcGis 10.1. Retrieved from <https://luisfonsecablog.wordpress.com/aplicacion-del-algebra-de-mapas-en-la-elaboracion-de-un-mapa-de-susceptibilidad-de-acuiferos-por-el-metodo-drastic-usando-arcgis-10-1/>
- Fortin, V., Bernier, J., & Bobée, B. (1997). Simulation, Bayes, and bootstrap in statistical hydrology. *Water Resources Research*, 33(3), 439–448. <https://doi.org/10.1029/96WR03355>
- Foufoula-Georgiou, E., & Krajewski, W. (1995). Recent advances in rainfall modeling, estimation, and forecasting. *U.S. National Report to International Union of Geodesy and Geophysics 1991-1994*, 1125–1137.
- Francisco, C.-A. D. (2015). Búsqueda de tendencias en la precipitación anual del estado de Zacatecas, México; en 30 registros con más de 50 años. *Ingeniería, Investigación y Tecnología*, 16(3), 355–368. <https://doi.org/10.1016/j.riit.2015.05.008>
- French, M. N., Krajewski, W. F., & Cuykendall, R. R. (1992). Rainfall forecasting in space and time using a neural network. *Journal of Hydrology*, 137(1–4), 1–31. [https://doi.org/10.1016/0022-1694\(92\)90046-X](https://doi.org/10.1016/0022-1694(92)90046-X)
- Gao, X., & Rao, T. S. (2011). Regression Models with STARMA Errors: An Application to the Study of Temperature Variations in the Antarctic Peninsula. *Advances in Directional and Linear Statistics*, 1–321. <https://doi.org/10.1007/978-3-7908-2628-9>
- García G., F. P. (2010). Modelación hidrológica estocástica: Desarrollo de un modelo de generación sintética de series temporales. *Universidad Autónoma Gabriel Rene Moreno. Santa Cruz--Bolivia*.
- Gil-Alana, L. A. (2012). Long memory, seasonality and time trends in the average monthly temperatures in Alaska. *Theoretical and Applied Climatology*, 108(3–4), 385–396. <https://doi.org/10.1007/s00704-011-0539-0>
- Gonzalez Barrios, J. L., & Descroix, L. (2012). Adaptación a los impactos del cambio climático en la zona semiárida Mexicana. *Parcerias Estratégicas*, 16(33), 09-22.
- González E, M., Jurado, E., González E, S., Aguirre C, Ó., Jiménez P, J., & Navar, J. (2003). Cambio climático mundial: origen y consecuencias. *Ciencia Uanl*, 3, 377–385.
- Grimaldi, S., Serinaldi, F., & Tallorini, C. (2005). Advances in Geosciences Multivariate linear parametric models applied to daily rainfall time series. *Advances In Geosciences*, 2, 87–92.

- Gu, X., & Jiang, J. (2005). A complex autoregressive model and application to monthly temperature forecasts. *Annales Geophysicae*, 23(10), 3229–3235. <https://doi.org/10.5194/angeo-23-3229-2005>
- Guenni, L., Degryze, E., & Alvarado, K. (2008). Análisis de la tendencia y la estacionalidad de la precipitación mensual en Venezuela. *Revista Colombiana de Estadística*, 31(1), 41–65.
- Haltiner, J. P., & Salas, J. D. (1988). Development and testing of a multivariate, seasonal arma(1,1) model. *Journal of Hydrology*, 104, 247–272.
- Hannachi, A. (2014). Intermittency, autoregression and censoring: a first-order AR model for daily precipitation. *Meteorological Applications*, 21, 384–397. <https://doi.org/10.1002/met.1353>
- Hassan Ali, H. M., & Mahgoub Mohamed, T. (2015). Rainfall Drought Simulating Using Stochastic SARIMA Models for Gadaref Region, Sudan, (August 2014), 1–11.
- Hino, M., & Hasebe, M. (1984). Identification and prediction of nonlinear hydrologic systems by the filter-separation autoregressive (AR) method: Extension to hourly hydrologic data. *Journal of Hydrology*, 68(1–4), 181–210. [https://doi.org/10.1016/0022-1694\(84\)90211-7](https://doi.org/10.1016/0022-1694(84)90211-7)
- Huff, F. (1967). Time Distribution Rainfall in Heavy Storms. *Water Resources Research*, 3(4), 1007–1019. <https://doi.org/10.1029/WR003i004p01007>
- Hundecha, Y., Pahlow, M., & Schumann, A. (2009). Modeling of daily precipitation at multiple locations using a mixture of distributions to characterize the extremes. *Water Resources Research*, 45(12). <https://doi.org/10.1029/2008WR007453>
- Im, H.-K., Rathouz, P. J., & Frederick, J. E. (2009). Space-time modeling of 20 years of daily air temperature in the Chicago metropolitan region. *Environmetrics*, 20(5), 494–511. <https://doi.org/10.1002/env.940>
- Iqelan, B. M. (2015). Time Series Modeling of Monthly Temperature Data of Jerusalem / Palestine. *Matematika*, 31(2), 159–176.
- Jacobs, P. A., & Lewis, P. A. . W. (1978). Discrete Time Series Generated by Mixtures. I: Correlational and Runs Properties. *Journal of the Royal Statistical Society. Series B (Methodological)*, 40(1), 94–105. <https://doi.org/10.1002/jae.1>
- Janert, P. K. (2011). *Data Analysis with Open Source Tools*. (M. Loukides, Ed.) (First Edit). Sebastopol, CA: O'Reilly Media Inc.
- Jia, Y., & Culver, T. B. (2006). Bootstrapped artificial neural networks for synthetic flow generation with a small data sample. *Journal of Hydrology*, 331(3–4), 580–590. <https://doi.org/10.1016/j.jhydrol.2006.06.005>
- Johnson, G. L., & Hanson, C. L. (1995). Topographic and Atmospheric Influences on Precipitation Variability over a Mountainous Watershed. *Journal of Applied*

Meteorology. <https://doi.org/10.1175/1520-0450-34.1.68>

- Joly, D., Cardot, H., & Schaumberger, A. (2013). Improving spatial temperature estimates by resort to time autoregressive processes. *International Journal of Climatology*, 33(10), 2289–2297. <https://doi.org/10.1002/joc.3601>
- Junttila, J. (2001). Structural breaks, ARIMA model and Finnish inflation forecasts. *International Journal of Forecasting*, 17(2), 203–230. [https://doi.org/10.1016/S0169-2070\(00\)00080-7](https://doi.org/10.1016/S0169-2070(00)00080-7)
- Kang, K., Oh, J., Kwon, C., & Park, Y. (1993). Two-Layer Neural Network, 48(6), 4805–4809.
- Kang, X., Stamm, M. C., Peng, A., & Liu, K. J. R. (2013). Robust median filtering forensics using an autoregressive model. *IEEE Transactions on Information Forensics and Security*, 8(9), 1456–1468. <https://doi.org/10.1109/TIFS.2013.2273394>
- Karanasos, A. M., Kim, J., Journal, E., & Kim, J. (2003). Moments of the ARMA — EGARCH model. *Econometrics Journal*, 6(1), 146–166.
- Karthikeyan, L., & Nagesh Kumar, D. (2013). Predictability of nonstationary time series using wavelet and EMD based ARMA models. *Journal of Hydrology*, 502, 103–119. <https://doi.org/10.1016/j.jhydrol.2013.08.030>
- Karunanithi, B. N., Grenney, W. J., Whitley, D., & Bovee, K. (1994). Neural Networks for River Flow Prediction. *Journal of Computing in Civil Engineering*, 8(2), 201–220.
- Katz, R. W. (1977a). An Application of Chain-Dependent Processes to Meteorology. *Journal of Applied Probability*, 14(3), 598–603. <https://doi.org/10.2307/3213463>
- Katz, R. W. (1977b). Precipitation as a Chain-Dependent Process. *Journal of Applied Meteorology*. [https://doi.org/10.1175/1520-0450\(1977\)016<0671:PAACDP>2.0.CO;2](https://doi.org/10.1175/1520-0450(1977)016<0671:PAACDP>2.0.CO;2)
- Kavvas, M. L., & Delleur, J. W. (1975). The Stochastic and Chronologic Structure Of Rainfall Sequences: Application To Indiana. *IWRRC Technical Reports.*, (56), 242. Retrieved from <http://docs.lib.purdue.edu/watertech/56/>
- Khedhiri, S. (2016). Forecasting Temperature Records in PEI, Canada. *Letters in Spatial and Resource Sciences*, 9(1), 43–55. <https://doi.org/10.1007/s12076-014-0135-x>
- Kim, J. W., Kim, K. Y., Kim, M. K., Cho, C. H., Lee, Y., & Lee, J. (2013). Statistical multisite simulations of summertime precipitation over South Korea and its future change based on observational data. *Asia-Pacific Journal of Atmospheric Sciences*, 49(5), 687–702. <https://doi.org/10.1007/s13143-013-0061-7>
- Kim, K. Y., Hamlington, B., & Na, H. (2015). Theoretical foundation of cyclostationary EOF analysis for geophysical and climatic variables: Concepts and examples.

- Komorník, J., Komorníková, M., Mesiar, R., Szökeová, D., & Szolgay, J. (2006). Comparison of forecasting performance of nonlinear models of hydrological time series. *Physics and Chemistry of the Earth*, 31(18), 1127–1145. <https://doi.org/10.1016/j.pce.2006.05.006>
- Kullback, S., & Leibler, R. A. (1951). On Information and Sufficiency. *The Annals of Mathematical Statistics*, 22(1), 79–86. Retrieved from <http://www.jstor.org/stable/2236703> .
- Kumar, R. K., & Kumar, D. (2004). Stochastic Modelling of Daily Rainfall for Sout West Monsoon Season of Bapatla, Andhra Pradesh. *Journal of Agricultural Engineering*, 41(3), 41–45.
- Kwon, H. H., Lall, U., & Khalil, A. F. (2007). Stochastic simulation model for nonstationary time series using an autoregressive wavelet decomposition: Applications to rainfall and temperature. *Water Resources Research*, 43(5), 1–15. <https://doi.org/10.1029/2006WR005258>
- Lee, J.-H., & Sohn, K.-T. (2007). Prediction of monthly mean surface air temperature in a region of China. *Advances in Atmospheric Sciences*, 24(3), 503–508. <https://doi.org/10.1007/s00376-007-0503-1>
- Leite, S. M., & Peixoto, J. P. (1996). The Autoregressive Model of Climatological Time Series: An application to the longest Time Series in Portugal. *International Journal of Climatology*, 16(March), 1165–1173. [https://doi.org/10.1002/\(SICI\)1097-0088\(199610\)16](https://doi.org/10.1002/(SICI)1097-0088(199610)16)
- Lidén, R., & Harlin, J. (2000). Analysis of conceptual rainfall-runoff modelling performance in different climates. *Journal of Hydrology*, 238(3–4), 231–247. [https://doi.org/10.1016/S0022-1694\(00\)00330-9](https://doi.org/10.1016/S0022-1694(00)00330-9)
- López-Lambraño, A., Fuentes, C., López-Ramos, A., Pliego-Díaz, M., & López-L, M. (2016). Rainfall Series Fractality in the Baja California State. In *Recent Advances in Fluid Dynamics with Environmental Applications* (pp. 139–152). Springer.
- Lorrai, M., & Sechi, G. M. (1995). Neural nets for modelling rainfall-runoff transformations. *Water Resources Management*, 9(4), 299–313. <https://doi.org/10.1007/BF00872489>
- Macchiato, M., Serio, C., Lapenna, V., & La Rotonda, L. (1993). Parametric time series analysis of cold and hot spells in daily temperature: an application in southern Italy. *Journal of Applied Meteorology*, 32(7), 1270–1281. [https://doi.org/10.1175/1520-0450\(1993\)032<1270:PTSAOC>2.0.CO;2](https://doi.org/10.1175/1520-0450(1993)032<1270:PTSAOC>2.0.CO;2)
- Machiwal, D., & Jha, M. K. (2008). Comparative evaluation of statistical tests for time series analysis: application to hydrological time series. *Hydrological Sciences*,

53(April), 353–366. <https://doi.org/10.1623/hysj.53.2.353>

- Magnano, L., Boland, J. W., & Hyndman, R. J. (2008). Generation of synthetic sequences of half-hourly temperatures. *Environmetrics*, 19(8), 818–835. <https://doi.org/10.1002/env.905>
- Mahsin, M., Akhter, Y., & Begum, M. (2012). Modeling Rainfall in Dhaka Division of Bangladesh Using Time Series. *Journal of Mathematical Modelling and Application*, 1(5), 67–73.
- Mallikarjuna, P., & Vardhan, G. V. (2002). Stochastic modelling of monthly rainfall—a case study. *ISH Journal of Hydraulic Engineering*, 8(2), 60–72. <https://doi.org/10.1080/09715010.2002.10514717>
- Mann, M. E., Bradley, R. S., & Hughes, M. K. (1998). Global-scale temperature patterns and climate forcing over the past six centuries. *Nature*, 392, 779–788.
- Martín, J. M. S. (1995). Propuesta metodológica para la obtención de gradientes termohídricos anuales. *Lurralde: Investigación y Espacio*, (18), 137–154.
- Martínez, A., & Salas, J. D. (2004). Modelación estocástica de lluvias horarias. *Ingeniería Del Agua*, 11(1), 29–39.
- Martínez Preece, M. R., & Martínez, F. V. (2014). Análisis del riesgo de mercado de los fondos de pensión en México Un enfoque con modelos autorregresivos. *Contaduría y Administración*, 59(3), 165–195. [https://doi.org/10.1016/S0186-1042\(14\)71269-0](https://doi.org/10.1016/S0186-1042(14)71269-0)
- McCall, K. C., & Jeraj, R. (2007). Dual-component model of respiratory motion based on the periodic autoregressive moving average (periodic ARMA) method. *Physics in Medicine and Biology*, 52(12), 3455–3466. <https://doi.org/10.1088/0031-9155/52/12/009>
- Meddahi, N. (2003). ARMA representation of integrated and realized variances. *Econometrics Journal*, 6(2), 335–356.
- Medina Hurtado, S., & García Aguado, J. (2005). Predicción de demanda de energía en Colombia mediante un sistema de inferencia difuso neuronal. *Energética*, (33).
- Mehdizadeh, S., Behmanesh, J., & Khalili, K. (2017). A comparison of monthly precipitation point estimates at 6 locations in Iran using integration of soft computing methods and GARCH time series model. *Journal of Hydrology*, 554, 721–742. <https://doi.org/10.1016/j.jhydrol.2017.09.056>
- Mills, T. C. (2007). Time series modelling of two millennia of northern hemisphere temperatures: Long memory or shifting trends? *Journal of the Royal Statistical Society. Series A: Statistics in Society*, 170(1), 83–94. <https://doi.org/10.1111/j.1467-985X.2006.00443.x>
- Minns, A. W., & Hall, M. J. (1996). Artificial neural networks as rainfall-runoff models.

- Minns, A. W., & Hall, M. J. (1997). Living with the ultimate black box: more on artificial neural networks. In *BHS 6th National Hydrology Symposium* (pp. 9–45).
- Modarres, R., & Ouarda, T. B. M. J. (2013). Modeling rainfall-runoff relationship using multivariate GARCH model. *Journal of Hydrology*, 499, 1–18. <https://doi.org/10.1016/j.jhydrol.2013.06.044>
- Mohammadi, K., Eslami, H. R., & Kahawita, R. (2006). Parameter estimation of an ARMA model for river flow forecasting using goal programming. *Journal of Hydrology*, 331(1–2), 293–299. <https://doi.org/10.1016/j.jhydrol.2006.05.017>
- Molina-Aguilar, J. P., & Aparicio, J. (2006). Tránsito de avenidas en cauces mediante redes neuronales artificiales. *Ingeniería Hidráulica En México*, 21(4), 65–86.
- Momani, P. E. N. M. (2009). Time series analysis model for rainfall data in Jordan: Case study for using time series analysis. *American Journal of Environmental Sciences*, 5(5), 599–604. Retrieved from <http://www.scopus.com/inward/record.url?eid=2-s2.0-70350129727&partnerID=40&md5=5e356da8cc4381cf6d1ee25214abe8b0>
- Monbet, V., & Ailliot, P. (2017). Sparse vector Markov switching autoregressive models. Application to multivariate time series of temperature. *Computational Statistics and Data Analysis*, 108, 40–51. <https://doi.org/10.1016/j.csda.2016.10.023>
- Moreno, F., & Roldán, J. (1999). Regionalization of Daily Precipitation Stochastic Model Parameters. Application to the Guadalquivir Valley in Southern Spain. *Science*, 24(1), 65–71.
- Murillos, J., Trejos, a, & Carvajal, P. (2003). Estudio del pronóstico de la demanda de energía eléctrica, utilizando modelos de series de tiempo. *Scientia et Technica*, 23(23), 37–40.
- Nickerson, D. M., & Madsen, B. C. (2005). Nonlinear regression and ARIMA models for precipitation chemistry in East Central Florida from 1978 to 1997. *Environmental Pollution*, 135, 371–379. <https://doi.org/10.1016/j.envpol.2004.11.010>
- Niedzielski, T. (2007). A data-based regional scale autoregressive rainfall-runoff model: A study from the Odra River. *Stochastic Environmental Research and Risk Assessment*, 21(6), 649–664. <https://doi.org/10.1007/s00477-006-0077-y>
- Nirmala, M., & Sundaram, S. M. (2010). A Seasonal Arima Model for Forecasting Monthly Rainfall in Tamilnadu. *National Journal on Advances in Building Sciences and Mechanics*, 1(2), 43–47.
- Nury, A. H., Hasan, K., & Alam, M. J. Bin. (2017). Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh. *Journal of King Saud University - Science*, 29(1), 47–61. <https://doi.org/10.1016/j.jksus.2015.12.002>

- Obeysekera, J. T. B., & Salas, J. D. (1986). Modeling of aggregated hydrologic time series. *Journal of Hydrology*, 86(3–4), 197–219. [https://doi.org/10.1016/0022-1694\(86\)90165-4](https://doi.org/10.1016/0022-1694(86)90165-4)
- Obeysekera, J. T. B., Tabios, G. Q., & Salas, J. D. (1987). On parameter estimation of temporal rainfall models. *Water Resources Research*, 23(10), 1837–1850. <https://doi.org/10.1029/WR023i010p01837>
- Obregón, N., Fragala, F., & Prada, L. (2003). Redes neuronales artificiales en hidroinformática. In *Seminario Internacional La Hidroinformática en la Gestión Integrada de los Recursos Hídricos* (pp. 1–5). Retrieved from <http://cinara.univalle.edu.co/archivos/pdf/61.pdf>
- Ochoa-Rivera, J. C. (2008). Prospecting droughts with stochastic artificial neural networks. *Journal of Hydrology*, 352(1–2), 174–180. <https://doi.org/10.1016/j.jhydrol.2008.01.006>
- Pabón-Caicedo, J. D., Eslava-Ramírez, J. A., & Gómez-Torres, R. E. (2001). Generalidades De La Distribución Espacial Y Temporal De La Temperatura Del Aire Y De La Precipitación En Colombia. *Meteorología Colombiana*, 4, 47–59. <https://doi.org/ISSN-0124-6984>
- Padilla, A., Pulido-Bosch, & Calvache, M. L. (1996). Análisis de series temporales de dos manantiales kársticos mediante la aplicación de modelos ARMA. *Geogaceta*, 19, 138–141.
- Panteleeva, O. V. (2005). *Fundamentos de Probabilidad y Estadística*. Universidad Autónoma del Estado de México. Retrieved from <https://books.google.com.mx/books?id=cbdromy2XrwC>
- Paolo Burlando, Renzo Rosso, L. G. C. and J. D. S. (1993). Forecasting of short-term rainfall using ARMA models. *Journal of Hydrology*, 144, 193–211.
- Papalaskaris, T., Panagiotidis, T., & Pantrakis, A. (2016). Stochastic Monthly Rainfall Time Series Analysis, Modeling and Forecasting in Kavala City, Greece, North-Eastern Mediterranean Basin. *Procedia Engineering*, 162, 254–263. <https://doi.org/10.1016/j.proeng.2016.11.054>
- Paris, M., D'Elia, M., Perez, M., & Tujchneider, O. (2002). Análisis estocástico de variables hidrometeorológicas para la estimación de recarga de acuíferos. *Groundwater and Human Development*, (March 2016), 684–692.
- Pedroza, H., & Dicovskyi, L. (2006). *Sistema de Analisis Estadístico con SPSS*. IICA Biblioteca Venezuela. Retrieved from <https://books.google.com.mx/books?id=sE0qAAAAYAAJ>
- Peña, D., & Arnáiz Tovar, G. (1981). Criterios de selección de modelos ARIMA. *Trabajos de Estadística e Investigación Operativa*, ISSN 0041-0241, Vol. 32, Nº. 1, 1981, Pags. 70-93, 32.

- Philip, G. M., & Watson, D. F. (1982). A precise method for determining contoured surfaces. *The APPEA Journal*, 22(1), 205–212.
- Piwowar, J. M., & Ledrew, E. F. (2002). ARMA time series modelling of remote sensing imagery: A new approach for climate change studies. *International Journal of Remote Sensing*, 23(24), 5225–5248. <https://doi.org/10.1080/01431160110109552>
- Quispe, C., & Purca, S. (2007). Previsión de la temperatura superficial del mar frente a la costa peruana mediante un modelo autorregresivo integrado de media móvil. *Revista Peruana de Biología*, 14(1), 109–115. Retrieved from http://www.scielo.org.pe/scielo.php?script=sci_arttext&pid=S1727-99332007000200020
- Roesner, L. A., & Yevjevich, V. M. (1966). Mathematical Models For Time Series of Monthly Precipitation and Monthly Runoff. *Hydrology Papers* 15, 13(15), 57.
- Ruhf, R. J., & Cutrim, E. M. C. (2003). Time series analysis of 20 years of hourly precipitation in Southwest Michigan. *Journal of Great Lakes Research*, 29(2), 256–267. [https://doi.org/10.1016/S0380-1330\(03\)70431-6](https://doi.org/10.1016/S0380-1330(03)70431-6)
- Rustom Jabbaz, A. (2012). *Estadística Descriptiva, Probabilidad e Inferencia. Una visión conceptual y aplicada*. Santiago de Chile. Retrieved from <http://www.captura.uchile.cl/handle/2250/50772>
- S., & Minns, A. W. (1998). Artificial Neural Networks as Subsymbolic Process Descriptors. *Thesis*, 140.
- Sahoo, G. B., Ray, C., & De Carlo, E. H. (2006). Use of neural network to predict flash flood and attendant water qualities of a mountainous stream on Oahu, Hawaii. *Journal of Hydrology*, 327(3–4), 525–538. <https://doi.org/10.1016/j.jhydrol.2005.11.059>
- Said, S. M., Manjang, S., Tjaronge, M. W., & Thaha, M. A. (2013). Arima Application as an Alternative Method of Rainfall Forecasts In Watershed Of Hydro Power Plant. *International Journal of Computational Engineering Research*, 03(9), 68–73. Retrieved from <http://www.ijceronline.com>
- Salas, J. D. (1980). *Applied modeling of hydrologic time series*. Water Resources Publication.
- Salas, J. D., Boes, D. C., & Smith, R. A. (1982). Estimation of ARMA Models With Seasonal Parameters. *Water Resources Research*, 18(4), 1006–1010.
- Sánchez-Cohen, I., Díaz-Padilla, G., Ojeda-Bustamante, W., Orona-Castillo, I., González-Barríos, J. L., & González-Cervantes, G. (2008). Variabilidad climática en México: algunos impactos hidrológicos, sociales, y económicos Variabilidad climática en México: algunos impactos hidrológicos, sociales y económicos. *Ingeniería Hidráulica En México*, XXIII(4), 5–24.
- Savin, N. E., & White, K. J. (1977). The Durbin-Watson test for serial correlation with

- extreme sample sizes or many regressors. *Econometrica: Journal of the Econometric Society*, 1989–1996.
- Scheidereiter, G. D., & Faure, O. R. (n.d.). Aplicación de los Modelos Autorregresivos Integrados de Media Móvil (ARIMA) a las Series de Precipitaciones de Lluvia.
- Seasholtz, M. B., & Kowalski, B. (1993). The parsimony principle applied to multivariate calibration. *Analytica Chimica Acta*, 277(2), 165–177. [https://doi.org/10.1016/0003-2670\(93\)80430-S](https://doi.org/10.1016/0003-2670(93)80430-S)
- Sharma, A., & Lall, U. (1999). A nonparametric approach for daily rainfall simulation. *Mathematics and Computers in Simulation*, 48(4–6), 361–371. [https://doi.org/10.1016/S0378-4754\(99\)00016-6](https://doi.org/10.1016/S0378-4754(99)00016-6)
- Slutzky, E. (1937). The Summation of Random Causes as the Source of Cyclic Processes. *ECONOMETRICA Journal of the Econometric Society Stable*, 5(2), 105–146. Retrieved from <http://www.jstor.org/stable/1907241>
- Smith, J., & Eli, R. N. (1995). Neural-Network Models of Rainfall-Runoff Process. *Journal of Water Resources Planning and Management*, 121(6), 499–508. [https://doi.org/10.1061/\(ASCE\)0733-9496\(1995\)121:6\(499\)](https://doi.org/10.1061/(ASCE)0733-9496(1995)121:6(499))
- Soltani, S., Modarres, R., & Eslamian, S. S. (2007). The use of time series modeling for the determination of rainfall climates of Iran. *International Journal of Climatology*, 27(October 2006), 819–829. <https://doi.org/10.1002/joc>
- Somlyódy, L., & Honti, M. (2005). The case of Lake Balaton: how can we exercise precaution? *Water Science and Technology*, 52(6), 195–203.
- Somvanshi, V. K., Pandey, O. P., Agrawal, P. K., Kalanker, N. V., Prakash, M. R., & Chand, R. (2006). Modelling and prediction of rainfall using artificial neural network and ARIMA techniques Modelling and prediction of rainfall using artificial neural network and ARIMA techniques. *J. Ind. Geophys. Union*, 10(2), 141–151. Retrieved from https://www.researchgate.net/publication/255587986_Modelling_and_prediction_of_rainfall_using_artificial_neural_network_and_ARIMA_techniques
- Spolia, & Chander. (1974). Modelling of Surface Runoff Systems by an ARMA Moel. *Journal of Hydrology*, 22, 317–332. <https://doi.org/10.1017/CBO9781107415324.004>
- Spolia, S. K. (1972). *Study on Stream Flow Synthesis*. M. Tech. Thesis, Dept. of Civil Eng., IIT, Delhi.
- Srikanthan, R. (2004). Stochastic generation of daily rainfall data using a nested model. In *57th Canadian water resources association annual congress* (pp. 16–18).
- Srikanthan, R., & MacMahon, T. A. (1985). *Stochastic generation of rainfall and evaporation data*. Australian Government Publ. Service.

- Srikanthan, R., & McMahon, T. A. (2001). Stochastic generation of annual, monthly and daily climate data: A review. *Hydrology and Earth System Sciences*, 5(4), 653–670. <https://doi.org/10.5194/hess-5-653-2001>
- Srikanthan, R., & others. (2005). Stochastic generation of daily rainfall data using a nested transition probability matrix model. In *29th Hydrology and Water Resources Symposium: Water Capital, 20-23 February 2005, Rydges Lakeside, Canberra* (p. 26).
- Stocker, T. F., Clarke, G. K. C., Le Treut, H., Lindzen, R. S., Meleshko, V. P., Mugara, R. K., ... Willebrand, J. (2001). Physical Climate Processes and Feedbacks. *Climate Change 2001: The Scientific Bases. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*, 881. <https://doi.org/10.1256/004316502320517344>
- Subbaiah, R., & Sahu, D. D. (2002). Stochastic model for weekly rainfall of Junagadh. *Journal of Agrometeorology*, 4(1), 65–73.
- Sulaiman, J. B., Darwis, H., & Hirose, H. (2014). Monthly maximum accumulated precipitation forecasting using local precipitation data and global climate modes. *Journal of Advanced Computational Intelligence and Intelligent Informatics*, 18(6), 999–1006.
- Teegavarapu, R. S. V, & Chandramouli, V. (2005). Improved weighting methods, deterministic and stochastic data-driven models for estimation of missing precipitation records. *Journal of Hydrology*, 312(1–4), 191–206. <https://doi.org/10.1016/j.jhydrol.2005.02.015>
- Tiao J. Chang, M. L. Kavas, J. W. D. (1984). Modeling of Sequences of Wet and Dry Days by Binary Discrete Autoregressive Moving Average Processes. *Journal of Climate and Applied Meteorology*, 12.
- Tingsanchali, T., & Gautam, M. R. (2000). Application of tank, NAM, ARMA and neural network models to flood forecasting. *Hydrological Processes*, 14(14), 2473–2487. [https://doi.org/10.1002/1099-1085\(20001015\)14:14<2473::AID-HYP109>3.0.CO;2-J](https://doi.org/10.1002/1099-1085(20001015)14:14<2473::AID-HYP109>3.0.CO;2-J)
- Tol, R. S. J. (1996). Autoregressive Conditional Heteroscedasticity in daily wind speed measurements. *Environmetrics*, 7, 67–75. <https://doi.org/10.1007/BF00863788>
- Toth, E., Brath, A., & Montanari, A. (2000). Comparison of short-term rainfall prediction models for real-time flood forecasting. *Journal of Hydrology*, 239(1–4), 132–147. [https://doi.org/10.1016/S0022-1694\(00\)00344-9](https://doi.org/10.1016/S0022-1694(00)00344-9)
- Triacca, U., Pasini, A., & Attanasio, A. (2014). Measuring persistence in time series of temperature anomalies. *Theoretical and Applied Climatology*, 118(3), 491–495. <https://doi.org/10.1007/s00704-013-1076-9>
- Tsay, R. S. (1988). Outliers, Level Shifts, and Variance Changes in Time Series:

ABSTRACT - ProQuest, 7(May 1987). Retrieved from <http://0-search.proquest.com.oasis.unisa.ac.za/docview/225203712/fulltextPDF/E119F641B9434E19PQ/1?accountid=14648>

Ula, T. A. (1990). Periodic covariance stationarity of multivariate periodic autoregressive moving average processes. *Water Resources Research*, 26(5), 855–861. <https://doi.org/10.1029/WR026i005p00855>

Ünal, N. E., Aksoy, H., & Akar, T. (2004). Annual and monthly rainfall data generation schemes. *Stochastic Environmental Research and Risk Assessment*, 18(4), 245–257. <https://doi.org/10.1007/s00477-004-0186-4>

Valdes-Sosa, P. A. (2004). Spatio-Temporal Autoregressive Models Defined Over Brain Manifolds. *Neuroinformatics*, 2, 239–250. Retrieved from <http://www.springerlink.com/content/t3t73113753t3223/fulltext.pdf?page=1>

Valipour, M. (2012). Number of Required Observation Data for Rainfall Forecasting According to the Climate Conditions. *American Journal of Scientific Research*, (74), 79–86.

Valipour, M. (2016). Optimization of neural networks for precipitation analysis in a humid region to. *Royal Meteorological Society*, 100(November 2015), 91–100. <https://doi.org/10.1002/met.1533>

Vallverdu, M., Korenberg, M. J., & Caminal, P. (1991). Model identification of the neural control of the cardiovascular system using NARMAX models. [1991] *Proceedings Computers in Cardiology*, 13(5), 585–588. <https://doi.org/10.1109/CIC.1991.168978>

Valverde Ramírez, M. C., De Campos Velho, H. F., & Ferreira, N. J. (2005). Artificial neural network technique for rainfall forecasting applied to the São Paulo region. *Journal of Hydrology*, 301(1–4), 146–162. <https://doi.org/10.1016/j.jhydrol.2004.06.028>

Viessman, B. W. (1989). Technology, society, and water management, 115(1), 48–51.

Vilalta Perdomo, C. J. (2016). *Análisis de datos*: CIDE. Retrieved from <https://books.google.com.mx/books?id=9W84DgAAQBAJ>

Villatoro, M., Henríquez, C., & Sancho, F. (2008). Comparación de los interpoladores IDW y Kriging en la variación espacial de pH, Ca, CICE y P del suelo. *Agronomía Costarricense*, 32(1).

Vrugt, J. A., ter Braak, C. J. F., Clark, M. P., Hyman, J. M., & Robinson, B. A. (2008). Treatment of input uncertainty in hydrologic modeling: Doing hydrology backward with Markov chain Monte Carlo simulation. *Water Resources Research*, 44(12), 1–15. <https://doi.org/10.1029/2007WR006720>

Walpole, R. E. (2007). *Probabilidad y estadística para ingeniería y ciencias*. Pearson Educación. Retrieved from <https://books.google.com.mx/books?id=jBnRMQAACAAJ>

- Wang, H. R., Gao, X., Qian, L. X., & Yu, S. (2012). Uncertainty analysis of hydrological processes based on ARMA-GARCH model. *Science China Technological Sciences*, 55(8), 2321–2331. <https://doi.org/10.1007/s11431-012-4909-3>
- Wang, Q. J., Nathan, R. J., & others. (2002). A daily and monthly mixed algorithm for stochastic generation of rainfall time series. In *Water Challenge: Balancing the Risks: Hydrology and Water Resources Symposium 2002* (p. 698).
- Wang, W., Gelder, P. H. A. J. M. Van, Vrijling, J. K., & Ma, J. (2006). Forecasting daily streamflow using hybrid ANN models. *Journal of Hydrology*, 324(1–4), 383–399. <https://doi.org/10.1016/j.jhydrol.2005.09.032>
- Watson, D. F. (1985). A refinement of inverse distance weighted interpolation. *Geoprocessing*, 2, 315–327.
- Weesakul, U., & Lowanichchai, S. (2005). Rainfall forecast for agricultural water allocation planning in Thailand. *Thammasat Int. J. Sc. Tech*, 10(3), 18–27.
- Wold, H. O. A. (1954). *The analysis of stationary time series*. Almqvist & Wiksell.
- Wurl, J., Nayeli, C., García, M., & Imaz, Á. (2017). Respuesta hidrológica al cambio climático en regiones áridas : caso de estudio en los Comondú , Baja California Sur , México. *Hidrobiologica*, 27(1), 13–22. Retrieved from <http://148.206.42.192/ojs/index.php/revHidro/article/view/1208>
- Yang, R. (1997). *Application of neural networks and genetic algorithms to modelling flood discharges and urban water quality*. University of Manchester.
- Yule G. Udny. (1926). Why do we Sometimes get Nonsense-Correlations between Time-Series ? --A Study in Sampling and the Nature of Time-Series Author (s): G . Udny Yule Source : Journal of the Royal Statistical Society , Vol . 89 , No . 1 (Jan . , 1926), pp . 1-63 Published. *Journal of the Royal Statistical Society*, 89(1), 1–63. Retrieved from <http://www.jstor.org/stable/2341482>
- Yule G. Udny. (1927). On a Method of Investigating Periodicities in Disturbed Series, with special reference to Wolfer's Sunspot Numbers. *Journal of the Royal Statistical Society*, 226, 267–298.
- Yusof, F., & Kane, I. L. (2012). Modeling Monthly Rainfall Time Series using ETS State Space and Sarima models. *International Journal of Current Research*, 4(09), 195–200.
- Zotelo, C. (2011). Variabilidad Climática y Ciclos Naturales. *Anales 2011 de La Academia Nacional de Agronomía y Veterinaria*, Tomo LXV, 374–381. <https://doi.org/10.13140/RG.2.2.30028.41602>