



UNIVERSIDAD AGRARIA DE LA HABANA
"Fructuoso Rodríguez Pérez"
FACULTAD: CIENCIAS TÉCNICAS

Tesis presentada en opción al título de Máster en Biomatemática

Modelación lineal estocástica para el pronóstico de caudales a escala de cuencas

Autor: Ing. Anaily Rivero Villaverde

Tutor: Dr.C Gustavo R. Alonso Brito

Mayabeque, 2019

Dedicatoria

A mis padres, por su amor sin límites y su apoyo en esta gran carrera de la vida.

A mi esposo Leynier, por cautivar mi corazón y llenarlo con su amor.

Agradecimientos

Sirvan estas líneas para agradecer a todos aquellos que de una forma u otra que me han ayudado en el transcurso de esta etapa:

A mi tutor Gustavo Alonso Brito y a todos los colegas del GIAF, en especial a: Jorge Díaz, Jorge García, Maria Elena Ruíz, Andrés Lau y Norys, por el apoyo incondicional, por sus tan oportunos consejos y por ser más que compañeros: amigos.

A la Facultad de Ciencias Técnicas ya todos los profesores del Departamento de Matemática -Física, en especial a Lucía y Walkiria, por su comprensión, dedicación y esfuerzo demostrados en el proceso de realización de este gran proyecto.

A mis padres y a mi hermana, por enseñarme lo verdaderamente importante y tan difícil de apreciar, por su cariño, amor y dedicación sin límites, por su paciencia y por darme la oportunidad de ser un pedacito de ellos.

A mi esposo Leynier, por tener la maravillosa idea de llegar a mi vida y quedarse a mi lado, por hacer que nuestras diferencias nos mantengan fuertemente unidos, por ser paciente, alentarme y demostrarme cada día la gran persona que es.

A Dios y a la Virgen de la Caridad, que me han dado fe y fuerzas en los momentos más difíciles.

A mis suegros Alexis y a Aleida, por acogerme como un miembro más de su familia.

A mis tías Tere y Daysi, porque me han dado siempre su cariño maternal.

Al resto de mi tíos y primos, porque cada parte de mi ser tiene el aporte individual de cada uno de ellos.

A todos los educadores que me ayudaron a escalar, en sus etapas, la gran montaña del estudio.

A aquellos que hoy no me acompañan, pero que dejaron sus huellas imborrables en mí, y a todas esas personas que, de una forma u otra, me tendieron su mano y me permitieron transitar el camino con menos dificultad.

A todos los animalitos, en especial a mis mascotas, pues como bien expresó Anatole France, Premio Nobel de Literatura en 1921: *“hasta que uno no ha amado un animal, una parte del alma sigue sin despertar”* ya que *“el amor por las criaturas vivientes es el atributo más noble del hombre”* (Charles Darwin).

Resumen

Este estudio se enfoca en la subcuenca “V Aniversario”, perteneciente a la cuenca del Río Cuyaguaje, provincia de Pinar del Río. Su posición latitudinal favorece el desarrollo de un volumen de escurrimiento superior a otras cuencas en zonas cársticas del país, por lo que se producen numerosos daños relacionados fundamentalmente con las inundaciones. Por tanto, el objetivo de este trabajo es pronosticar probabilísticamente el gasto líquido en esta subcuenca a través de la modelación lineal estocástica. Para dar cumplimiento a lo anterior, se implementaron en el software R los modelos: Ruido Blanco, $AR(p)$, $MA(q)$ y $ARMA(p,q)$ para la serie de escurrimiento anual, ajustándose mejor a la misma el Ruido Blanco. Sin embargo, la serie de escurrimiento mensual fue modelada mediante un SARIMA, pues presentó un comportamiento estacional. Esta última estructura fue validada en los dos últimos años de la serie, 1989 y 1990, realizando una predicción acertada del gasto líquido en este período.

Abstract

This study focuses in the subwatershed “V Aniversario”, belonging to Cuyaguaje River Basin, Pinar del Río Province. Its latitudinal position leads to the development of a volume of runoff superior to other basins in kartic regions of the country, then numerous damages mainly related with floods are done. Therefore, the objective of this work is to forecast probabilistically the runoff at this subwatershed through linear stochastic modeling. In order to comply with the above, White Noise, AR (p), MA (q) and ARMA (p, q) models for the annual runoff series were implemented in R software, being the White Noise model the best adjusted to it. However, the monthly runoff series was modeled by a SARIMA, because it presented a seasonal behavior. This last structure was validated in the last two years of the series, 1989 and 1990, making an accurate prediction of the runoff in this period.

Índice

INTRODUCCIÓN.....	1
CAPÍTULO 1: Revisión bibliográfica.....	4
1.1 Proceso lluvia-escurrimiento.....	4
1.1.1 Factores que influyen en la respuesta hidrológica.....	8
1.1.1.1 Suelo.....	8
1.1.1.2 Vegetación.....	10
1.1.1.3 Relieve.....	11
1.1.1.4 Clima.....	12
1.2 Modelación del proceso	13
1.2.1 Modelación lineal estocástica.....	17
1.3 Sistemas de pronóstico.....	20
CAPÍTULO 2: Materiales y métodos	24
2.1 Área de estudio.....	24
2.2 Datos de precipitación.....	25
2.3 Serie de escurrimiento	26
2.4 Modelos lineales estocásticos	27
2.4.1 Ruido Blanco (WN).....	28
2.4.2 Modelo autorregresivo AR(p)	28
2.4.3 Modelo de media móvil MA(q)	29
2.4.4 Modelo autorregresivo de media móvil ARMA(p,q).....	30
2.4.5 Modelo autorregresivo integrado de media móvil ARIMA(p,d,q).....	31
2.4.6 Modelo estacional autorregresivo integrado de media móvil SARIMA (p,d,q)(P,D,Q)	31
2.5 Análisis estadístico	32
2.5.1 Criterio de Información de Akaike (AIC).....	33
2.5.2 Test de tendencia de Mann-Kendall	33
2.6 Metodología.....	35
CAPÍTULO 3: Resultados y discusión	39
3.1 Estadística descriptiva	39
3.1.1 Serie de escurrimiento anual	39
3.1.2 Serie de escurrimiento mensual.....	40
3.1.3 Serie de escurrimiento diaria.....	42
3.2 Calibración de modelos	44
3.2.1 Serie de escurrimiento anual	44
3.2.2 Serie de escurrimiento mensual.....	47

3.2.2.1 Calibración del modelo SARIMA (p,d,q)(P,D,Q).....	50
3.2.1 Serie de escurrimiento diaria.....	55
3.3 Validación de los modelos	56
3.3.1 Serie de escurrimiento anual	57
3.3.2 Serie de escurrimiento mensual.....	57
CONCLUSIONES.....	62
RECOMENDACIONES.....	63
REFERENCIAS BIBLIOGRÁFICAS.....	64

INTRODUCCIÓN

Las cuencas actúan como sistemas de procesos y respuestas en donde la escorrentía superficial es su consecuencia más evidente, siendo esta la responsable de catástrofes y desastres naturales, y de numerosos daños relacionados fundamentalmente con las inundaciones (Triviño y Ortiz, 2004).

Las inundaciones por avenidas constituyen ocurrencias naturales causantes en general de trastornos, riesgos y daños a múltiples actividades, bienes y hasta la misma vida del hombre (Rodríguez, 2011). Planos (1998) cita lo siguiente: Las estadísticas reflejan claramente la magnitud de las catástrofes de origen hídrico. Un informe del Comité Regional de Recursos Hidráulicos del Istmo Centroamericano (Chacón y Zarate, 1996), indica que anualmente mueren en el mundo alrededor de 20 mil personas por catástrofes de tipo hidrometeorológico y que las pérdidas económicas se elevan a miles de millones de dólares. Esta es una de las principales causas de muerte por desastres naturales y tecnológicos, donde destacan las tormentas tropicales (Alonso, 2018).

Según informaciones recogidas en el informe sobre el Panorama Ambiental de Cuba (ONEI, 2016), en el periodo 2001–2015 las pérdidas económicas por afectación de huracanes se calculan en 25 737 Millones CUC. Afortunadamente no se lamenta la pérdida de vidas humanas, lo cual se debe en gran medida a la gestión de la Defensa Civil (Alonso, 2018).

La comprensión de la dinámica de los sistemas de recursos hídricos y la compleja interacción de los factores y variables que gobiernan su comportamiento, exige el desarrollo de técnicas, modelos y algoritmos, mediante los cuales se pueda tener una aproximación a la realidad física (Avilés *et al.*, 2016), así como también una predicción y pronóstico del proceso.

Los modelos hidrológicos, entendidos como una aproximación al funcionamiento real del ciclo del agua en una cuenca vertiente, se pueden clasificar en: físicos y abstractos. A su vez, los modelos abstractos, de acuerdo con la aleatoriedad de las variables empleadas, pueden ser estocásticos o deterministas (Triviño y Ortiz, 2004). Los modelos estocásticos nacen por la marcada variabilidad temporal y espacial de muchos procesos y componentes en un sistema que dan cuenta que son funciones de muchos factores estocásticos, es por esto que el análisis de los sistemas va más allá de los métodos convencionales deterministas (Avilés *et al.*, 2016).

Actualmente, los modelos estocásticos son muy usados en diversas aplicaciones (Liang y Zhuang, 2014; Niezgoda *et al.*, 2014; Sun *et al.*, 2014; D'Amico *et al.*, 2015). Dentro de estas, la modelación lineal estocástica o series temporales es ampliamente usada, especialmente en varios estudios de cuencas de ríos (Wu y Chau, 2013; Sang, 2013; Fry *et al.*, 2013). Por otra parte, el análisis de las series de tiempo hidrológico se ha convertido en una herramienta importante en la hidrología, utilizándose para la construcción de modelos matemáticos, para generar registros sintéticos de hidrología, para pronosticar los eventos hidrológicos, para detectar tendencias, saltos, para llenar datos incompletos y para extender o ampliar los registros de datos (Díaz y Guevara, 2016).

En Cuba existe una marcada tendencia al empleo de los modelos determinísticos en los procesos a escala de cuencas (Rodríguez *et al.*, 2010; Estrada y Pacheco, 2012; Rodríguez y Marrero de León, 2015). En consecuencia, la dimensión estocástica en la modelación ha sido menos explorada, por lo que se necesita fomentar investigaciones donde esté presente la modelación lineal estocástica debido a sus grandes potencialidades.

Este estudio se llevará a cabo en la subcuenca V Aniversario, que es uno de los cierres hidrográficos naturales de la cuenca del Río Cuyaguatzeje, ubicada en la provincia de Pinar del Río. Esta subcuenca abarca parte de los municipios Viñales y Minas de Matahambre con una extensión de 157 km² y su posición latitudinal favorece el desarrollo de un volumen de escurrimiento superior a otras corrientes en zonas cársticas del país. En particular, dadas sus características geomorfológicas, de drenaje (suelo) y cobertura, presenta pronta respuesta en forma de avenidas (CTCH, 2000). Por tanto, es de vital importancia viabilizar la toma de decisiones basadas en sistema de pronóstico que permita alertar tempranamente a los actores de la subcuenca.

Es necesario destacar que esta investigación se encuentra en estrecha correspondencia con el proyecto internacional AGUACUBA, que lleva por título: "Plataforma interinstitucional para la gestión integrada de los recursos hídricos en un contexto de agricultura sostenible y cambio climático". A pesar de que este proyecto está en fase de revisión tiene, entre sus motivaciones fundamentales, la situación de la infraestructura relacionada con la gestión de los recursos hídricos. Además, los objetivos estratégicos que persigue son: contribuir a la seguridad alimentaria, a la adaptación al cambio climático y a la reducción de riesgos de desastre.

Problema científico: ¿Cómo mejorar el pronóstico de caudales en la subcuenca V Aniversarios a diferentes escalas temporales?

Hipótesis: Con el uso de herramientas estadísticas y el análisis de series temporales es posible mejorar el pronóstico de caudales, aportando la dimensión probabilística en la predicción.

Objeto de estudio: El proceso lluvia- escurrimiento y su variabilidad temporal.

Campo de acción: La variabilidad temporal del proceso lluvia-escurrimiento en la subcuenca V Aniversario, del Río Cuyaguaje.

Objetivo General: Pronosticar probabilísticamente el gasto líquido en la subcuenca V Aniversario a través de la modelación lineal estocástica.

Objetivos específicos:

1. Caracterizar el fenómeno del escurrimiento en la subcuenca V Aniversario, mediante el análisis estadístico de las series temporales del escurrimiento medio anual, mensual y diario.
2. Calibrar y validar diferentes estructuras de modelos lineales estocásticos para distintas escalas temporales haciendo uso del software R.
3. Seleccionar las mejores estructuras de modelos lineales estocásticos para el pronóstico del gasto líquido en la subcuenca V Aniversario.

CAPÍTULO 1: Revisión bibliográfica

1.1 Proceso lluvia-escurrimiento

En muchas partes del mundo, el rápido crecimiento de la población, la urbanización y la industrialización han incrementado la demanda de agua. Estos mismos motivos han resultado en cuencas hidrográficas y sistemas fluviales alterados, los cuales han contribuido a una mayor pérdida de vidas y daños a la propiedad debido fundamentalmente a las inundaciones. Por tanto, cada vez es más crítico planear, diseñar y administrar los sistemas de recursos de agua cuidadosa e inteligentemente (Tokar y Johnson, 1999).

Durante muchos años, los hidrólogos han intentado comprender la transformación de la precipitación en escorrentía, con el fin de pronosticar el flujo de agua para fines tales como suministro de agua, control de inundaciones, riego, drenaje, calidad del agua, generación de energía, recreación y propagación de los peces y de la vida silvestre. Sin embargo, la relación precipitación-escorrentía es uno de los fenómenos hidrológicos más complejos de entender debido a la tremenda variabilidad espacial y temporal de las características de la cuenca y de los patrones de precipitación y, por otra parte, debido al número de variables involucradas en el modelado de los procesos físicos (Tokar y Johnson, 1999). Por tanto, se hace imperiosa la necesidad de profundizar cada vez más en el proceso lluvia-escurrimiento, pues este es de vital importancia en Hidrología para determinar la cantidad de escurrimiento en una cuenca debido a la lluvia (Tarboton, 2003).

La lluvia y el escurrimiento son definidos como componentes esenciales del ciclo hidrológico. Si se tiene en cuenta el modelo conceptual del sistema hidrológico en este proceso, la lluvia es la entrada y el escurrimiento la salida (Chow *et al.*, 1988). El proceso lluvia-escurrimiento también es conocido como entrada de agua superficial-escorrentía, para hacer alusión a otros factores, además de la lluvia o precipitación, como fuentes del escurrimiento (Tarboton, 2003).

Para determinar cuánta escorrentía es generada a partir de la lluvia o, de forma general, de las aguas superficiales, se requiere la división de estas aguas en la superficie de la tierra en componentes de infiltración y componentes que fluyen sobre la superficie y entran directamente a los ríos. Las trayectorias tomadas por el agua determinan muchas de las características de un paisaje, la ocurrencia y el tamaño de las inundaciones, los usos que se le puede dar a la tierra y las estrategias requeridas para un buen manejo de la tierra. Entender y modelar el proceso de lluvia-escorrentía

es importante en muchos problemas de inundaciones y recursos hídricos (Tarboton, 2003).

En la Figura 1.1 se ilustran esquemáticamente muchos de los procesos implicados en la generación de escorrentía (Tarboton, 2003), por lo que es fundamental profundizar en algunos elementos que componen estos procesos, como:

Precipitación directa: La precipitación puede ser en forma de lluvia o nieve y la vegetación puede interceptar una fracción de precipitación. La precipitación que penetra en la vegetación se conoce como precipitación directa y puede consistir en ambas: precipitación que no entra en contacto con la vegetación o que cae o drena fuera de la vegetación después de ser interceptada. Una gran fracción de agua interceptada es comúnmente evaporada de vuelta a la atmósfera. También hay flujo de agua hacia la atmósfera por la transpiración de la vegetación.

Nivel Freático: Es la superficie debajo de la cual el suelo y la roca están saturados y a una presión mayor que la atmosférica. Esto sirve como límite entre la zona saturada que contiene el *agua subterránea* y la zona no saturada.

Recarga del acuífero: Se define de esta manera al agua añadida al agua subterránea.

Franja capilar: Es aquella región de suelo que está cerca de la saturación, debido a que el agua está retenida por las fuerzas capilares. Se encuentra localizada inmediatamente por encima del nivel freático.

Por otra parte, se puede apreciar en la Figura 1.1 que el agua superficial puede acumularse en la superficie en un *almacenamiento de depresión*, o fluir sobre la superficie hacia los ríos, o infiltrarse en el suelo. Esta agua infiltrada además, puede permanecer en el suelo para luego evaporarse o transpirarse de vuelta a la atmósfera. En dependencia de las trayectorias que siga el agua, se obtendrán las diferentes clasificaciones para los flujos de agua.

Flujo superficial: Es el agua que fluye sobre la superficie directamente hacia el río.

Flujo subsuperficial: No es más que el agua que se infiltra en el suelo y fluye lateralmente hacia el río. Este tipo de flujo es relativamente rápido.

Flujo base: El agua que se infiltra también puede percolar en aguas subterráneas profundas, que pueden mantener el flujo del caudal o la corriente constante en los ríos durante escalas de tiempo mucho más largas y fundamentalmente durante los

períodos sin lluvia. Este tipo de flujo es conocido como flujo base, el cual no es más que el drenaje lateral del agua subterránea en los ríos.

Flujo de retorno: El *agua subsuperficial*, ya sea de flujo subsuperficial o de agua subterránea, puede fluir de regreso a través de la superficie terrestre para aumentar el flujo superficial. Esto se conoce como flujo de retorno.

El flujo superficial y los procesos de flujo subsuperficial de menor profundidad transportan el agua al río dentro de la escala de tiempo de aproximadamente un día o más. El agua que se infiltra en las aguas subterráneas, se mueve a velocidades mucho más bajas y alcanza el río durante períodos más largos de tiempo tales como semanas, meses o incluso años (Tarboton, 2003).

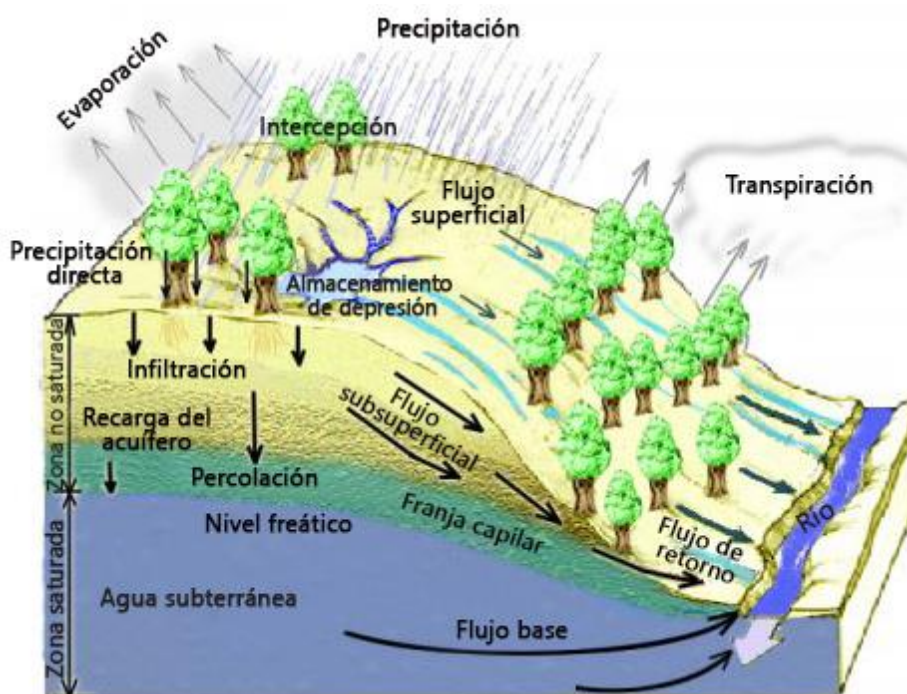


Figura 1.1 Procesos físicos involucrados en la generación de escurrimiento. Tomado de Tarboton (2003).

Es de interés destacar que existen dos mecanismos fundamentales de generación del escurrimiento: exceso de infiltración, conocido como flujo superficial hortoniano (Horton, 1933) (Figura 1.2a) y exceso de saturación (flujo superficial de saturación o flujo de Dunne) (Alonso, 2015) (Figura 1.2b). En el primero, se considera que el escurrimiento ocurrirá cuando la intensidad de la lluvia i sea mayor que la capacidad de infiltración del suelo f (pulgadas absorbidas por horas), por lo que la escorrentía ocurre a una tasa de $(i - f)$. En el segundo, el escurrimiento se produce cuando el flujo subsuperficial satura el suelo cerca de la parte inferior y entonces se presenta flujo

superficial a medida que la lluvia cae en un suelo saturado. La diferencia de este mecanismo con el anterior es que el suelo se satura desde abajo por el flujo subsuperficial, mientras que en el otro el suelo se satura desde arriba mediante infiltración (Chow *et al.*, 1988).

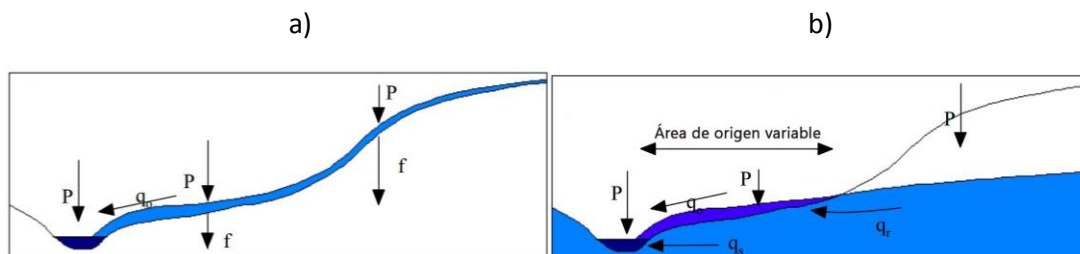


Figura 1.2 Mecanismos de generación del escurrimiento: a) Exceso de infiltración. b) Exceso de saturación. Tomado de Tarboton (2003).

Tarboton (2003) explica además que las escalas temporales y espaciales asociadas con el agua superficial, dadas como resultado de los procesos meteorológicos, tienen profundos efectos en el escurrimiento. Las precipitaciones de corta duración y de gran intensidad son mucho más propensas a exceder la capacidad de infiltración del suelo y resultar en un flujo superficial, en comparación con una lluvia de mayor duración y menos intensa. En los climas áridos con un nivel freático profundo, las lluvias concentradas espacialmente en una pequeña área pueden generar un escurrimiento local que luego se infiltra río abajo, mientras que un área más húmeda con un nivel freático poco profundo es menos probable que esté sujeta a pérdidas de infiltración e incluso, una lluvia moderada cuando se extiende y se acumula en grandes áreas puede dar lugar a grandes caudales.

Por último, es de interés destacar que los atributos geomorfológicos también son importantes para caracterizar la respuesta de escorrentía en una cuenca hidrográfica en particular. La escorrentía en grandes cuencas hidrográficas (> 1000 km², con redes de canales bien desarrolladas y fase de canal), por ejemplo, está dominada por el almacenamiento en los canales y estos son menos sensibles a las lluvias de corta duración y alta intensidad. Por el contrario, la escorrentía de cuencas pequeñas (<100 km², fase terrestre predominante) está dominada por el flujo superficial y estas son altamente sensibles a las precipitaciones de alta intensidad y corta duración (Alonso, 2015).

1.1.1 Factores que influyen en la respuesta hidrológica

En el epígrafe anterior se enunciaron dos mecanismos fundamentales para la generación de escorrentía: exceso de infiltración (flujo superficial hortoniano) y exceso de saturación (flujo superficial de saturación). Estos mecanismos dependen en gran medida de las vías seguidas por el agua infiltrada y cada uno de ellos tiene una respuesta diferente con respecto a la precipitación en el volumen de escorrentía producido, a la tasa de descarga máxima y al tiempo en que se contribuye al flujo del caudal en el canal. La importancia relativa de cada proceso se ve afectada por las características del suelo, la vegetación, el relieve y el clima (Tarboton, 2003). Por tanto, en este subepígrafe se analizará cómo influyen estos factores en la respuesta hidrológica, teniendo en cuenta los mecanismos anteriormente mencionados.

1.1.1.1 Suelo

Tarboton (2003) explica que el agua infiltrada puede fluir a través de la matriz del suelo por los poros inter-granulares y por pequeños huecos estructurales. El agua infiltrada también puede fluir a través de vacíos más grandes denominados macroporos, que incluyen los conductos abiertos en el suelo causados por raíces en descomposición y por la excavación de los animales. Los macroporos también incluyen vacíos estructurales más grandes dentro de la matriz del suelo que sirven como vías preferenciales para el flujo subsuperficial.

El flujo lateral rápido a través de una red de macroporos y el escurrimiento posterior en los canales de los ríos es el mecanismo principal para la generación de escorrentía en muchas regiones húmedas donde rara vez se observa flujo superficial. También existe una estrecha relación entre la profundidad de las aguas subterráneas y la escorrentía. Seibert *et al.* (2003) muestra cómo la escorrentía está más estrechamente relacionada con la profundidad del agua subterránea cerca del río que más arriba en la pendiente.

Por otra parte, los suelos naturales contienen heterogeneidades que conducen a la variabilidad en el proceso de infiltración en sí mismo. El agua que se infiltra sigue rutas preferenciales y macroporos, lo cual puede resultar en un incremento del contenido de humedad del suelo, a determinada profundidad y antes de la saturación, o en incrementos similares en el contenido de humedad en la parte superior del perfil del suelo. Muchos de los procesos de generación de escorrentía dependen del estado de humedad del suelo. Esto se conoce como las condiciones antecedentes. Los procesos de evaporación, transpiración, percolación y drenaje sirven para establecer las condiciones antecedentes de humedad del suelo (Tarboton, 2003).

Existe una tasa límite máxima a la cual un suelo, en una condición dada, puede absorber la entrada de agua superficial. Esto fue mencionado por Robert E. Horton en 1933 como la capacidad de infiltración del suelo y, por lo tanto, este mecanismo también se conoce como flujo superficial hortoniano, como ya ha sido mencionado con anterioridad. La capacidad de infiltración también se conoce como infiltrabilidad. Cuando el aporte de agua superficial excede la capacidad de infiltración, el exceso de agua se acumula en la superficie del suelo y llena pequeñas depresiones. Con la entrada continua de agua en la superficie, el agua almacenada en esas depresiones se derrama cuesta abajo por la pendiente. La cantidad de agua almacenada en la ladera en el proceso de fluir cuesta abajo se llama detención superficial. La transición del almacenamiento de depresión a la detención superficial y el flujo superficial no es nítida, ya que algunas depresiones pueden llenarse y contribuir al flujo superficial antes que otras.

Debido a la variabilidad espacial de las propiedades del suelo que afectan la capacidad de infiltración y debido a la variabilidad espacial del agua superficial, la escorrentía por exceso de infiltración no ocurre necesariamente sobre toda la cuenca de drenaje durante una tormenta o un evento de agua superficial. El flujo superficial por exceso de infiltración también se puede ver dónde se ha compactado el suelo o se ha eliminado la capa superficial del suelo. Este tipo de flujo es particularmente obvio en las áreas urbanas pavimentadas (Tarboton, 2003). Además, autores como Morin *et al.* (1981) plantean que la formación de costra o sello superficial, es el principal proceso responsable de la generación de escorrentía, pérdida de suelo y reducción de la infiltración.

Por otra parte, el flujo superficial puede ocurrir debido al aporte de agua superficial en áreas que están ya saturadas. Esto es conocido como flujo superficial por exceso de saturación, como ya ha sido abordado con anterioridad y sucede en lugares donde el agua que se infiltra satura por completo el perfil del suelo hasta que no hay espacio para infiltrarse más agua. La saturación completa del perfil de suelo resulta en una elevación del nivel freático a la superficie. Una vez que se produce la saturación desde abajo en un lugar, toda la entrada adicional de agua superficial en esa ubicación se convierte en escorrentía del flujo superficial.

Existen otros procesos de generación de escorrentía que se encuentran relacionados con las características del suelo, por lo que es necesario mencionar brevemente algunos de estos factores involucrados. En términos de composición mecánica del suelo, autores como Morgan (2005) afirman que las partículas más finas de la fracción

de arcilla dan mayor resistencia al suelo debido a su naturaleza cohesiva. La mineralogía, por su parte, tiene un efecto sustancial en la dispersión de arcillas. Este elemento afecta la estabilidad estructural, el escurrimiento y la pérdida de suelo (Lado y Ben-Hur, 2004). Por otra parte, el estado de congelación del suelo, en regiones donde se produce la congelación, juega un importante papel en la generación de escorrentía. La capacidad de infiltración se reduce debido al suelo congelado, dependiendo del contenido de humedad del suelo en el momento de la congelación.

El fuego es otro de los factores que genera escorrentía, pues este resulta en la repelencia del agua por los suelos, lo que reduce la capacidad de infiltración. Una de las causas de la repelencia al agua son los productos químicos liberados durante un incendio que son absorbidos por el suelo y pueden hacer que este sea repelente al agua durante meses o años después de un incendio. El calor del fuego también elimina las películas delgadas del agua adheridas a las partículas del suelo por fuerzas capilares, desconectando los posibles caminos potenciales para el flujo. La penetración del agua en los macroporos después de un incendio es limitada debido a este efecto. Las altas temperaturas en los desiertos tienen el mismo efecto, lo cual conlleva a que la tendencia de las capacidades de infiltración sea menor en las regiones áridas, lo que las hace más sujetas a los procesos de generación de escorrentía por exceso de infiltración. Esta repelencia al agua debido al fuego ha sido una de las causas fundamentales de muchas inundaciones luego de intensos incendios forestales.

1.1.1.2 Vegetación

Retomando lo abordado anteriormente, el flujo superficial por exceso de infiltración ocurre en cualquier parte donde el aporte de agua superficial excede la capacidad de infiltración de la superficie. Esto sucede con mayor frecuencia en áreas desprovistas de vegetación o que poseen solo una capa delgada. Las tierras de pastoreo semiáridas y los campos cultivados en regiones con alta intensidad de precipitaciones son lugares donde se puede observar este proceso.

En la mayoría de las regiones húmedas, las capacidades de infiltración son altas porque la vegetación protege el suelo del empapado y de la dispersión de la lluvia y porque el suministro de humus y la actividad de la micro fauna crean una estructura de suelo abierto. En tales condiciones, las intensidades del agua superficial generalmente no exceden las capacidades de infiltración y entonces, la escorrentía excesiva por infiltración es poco común.

La interceptación de la precipitación por la vegetación puede desempeñar un papel importante en la reducción de la escorrentía, especialmente en ambientes boscosos. Muchas aguas interceptadas eventualmente se evaporan nuevamente a la atmósfera. En algunos modelos hidrológicos, la interceptación se modela a veces como una abstracción inicial que es substraída de las entradas de precipitación antes de que las mismas sean usadas en cálculos de infiltración o escorrentía. En otros modelos hidrológicos se utilizan representaciones detalladas de la interceptación, el almacenamiento de agua o la precipitación directa.

1.1.1.3 Relieve

En las regiones húmedas se puede apreciar el flujo superficial por el exceso de saturación, pues los ríos típicamente están ganando agua por el drenaje del flujo base procedente de las aguas subterráneas. Lo anterior ocurre porque el nivel freático se encuentra cerca de la elevación superficial del agua de los ríos. Esto significa que el nivel freático cercano a los ríos está próximo a la superficie del suelo, especialmente en una topografía plana, lo que hace que las áreas cercanas a los ríos en una topografía plana sean particularmente susceptibles a la saturación desde abajo o por exceso de saturación. La extensión del área sujeta a la saturación desde abajo varía en el tiempo, tanto en las escalas de tiempo estacional como en las escalas de tiempo de eventos, debido a las fluctuaciones en la profundidad de los niveles freáticos poco profundos. Esta variabilidad del grado de saturación de la superficie se conoce como el área de origen variable (Hewlett y Hibbert, 1967).

Las consideraciones geométricas dictan que las zonas saturadas próximas al río serán más extensas en lugares con perfiles de pendiente cóncava y valles planos y anchos. Sin embargo, el flujo superficial por saturación no está restringido a las áreas cercanas al río. La saturación desde abajo también puede ocurrir: (1) donde las líneas del flujo subsuperficial convergen en las concavidades de la pendiente (pendientes muy empinadas) y el agua llega más rápido que si es transmitida por la pendiente como flujo subsuperficial; (2) en las roturas de las pendientes cóncavas donde el gradiente hidráulico que induce el flujo subsuperficial desde un punto superior de la pendiente es mayor que el que induce la transmisión en un punto inferior de la pendiente; (3) donde las capas del suelo que conducen al flujo subsuperficial son localmente delgadas; y (4) donde la conductividad hidráulica disminuye abruptamente o gradualmente con la profundidad y el agua de percolación se acumula por encima de las capas de baja conductividad para formar zonas de saturación que alcanzan la superficie.

En áreas con altas capacidades de infiltración, el flujo subsuperficial suele ser el factor que contribuye de manera dominante al flujo del caudal, especialmente en terrenos más empinados o en colinas más planas donde es menos probable que ocurra el exceso de saturación.

1.1.1.4 Clima

El régimen climático general controla el volumen total de escorrentía en cualquier región a través de su efecto sobre el balance de agua. Según Tarboton (2003), el balance de agua se define como:

$$P = Q + E \quad \text{Ec. 1.1}$$

donde P es la tasa de precipitación, Q la tasa de escorrentía y E la tasa de evapotranspiración. Esta ecuación indica que la entrada de precipitación se elimina ya sea en escorrentía o evapotranspiración. En general, el régimen climático controla la dosificación general. En este caso, se incluye Q en la recarga de agua subterránea. Debido a que las cantidades en la Ec. 1.1 deben ser positivas, esta ecuación coloca límites en los valores de Q y E dada cualquier P específica. Tanto Q como E están obligados a ser menores que P .

En general, existe un límite superior en la posible evapotranspiración, debido a las entradas de energía necesarias para evaporar el agua. Este límite está relacionado con la radiación solar y la capacidad de la atmósfera para transportar el agua evaporada lejos de la superficie (relacionada con el viento y la humedad). Este límite ha sido designado por Tarboton (2003) como E_p (evapotranspiración potencial). En general, si el volumen de precipitación es grande ($P \gg E_p$), E se aproximará a E_p y puede ocurrir escurrimiento. Además, si el volumen de precipitación es pequeño ($P \ll E_p$), el agua es muy limitada en la superficie de la tierra y puede que toda se evapore, por lo que es muy poco probable de que ocurra escurrimiento.

Por otra parte, Tarboton (2003) también explica que en las regiones áridas y semiáridas y las perturbadas por los humanos (a través de la agricultura, la urbanización y la minería), la capacidad de infiltración es un factor limitante, por lo que el exceso de infiltración es el proceso dominante de escorrentía. Sin embargo, en la mayoría de las regiones húmedas donde la infiltración no es un factor limitante, los dos principales procesos que generan escorrentía son: el flujo subsuperficial y el flujo superficial por el exceso de saturación.

1.2 Modelación del proceso

Según Pacheco (2012), un modelo hidrológico comprende un conjunto de abstracciones matemáticas que describen fases relevantes del ciclo hidrológico, con el objetivo de simular numéricamente los procesos identificados en el estudio. Sin embargo, según la percepción de la autora de este trabajo, Chow *et al.* (1988) ofrecen una definición más completa, pues plantean que un modelo del sistema hidrológico es una aproximación al sistema real; sus entradas y sus salidas son variables hidrológicas mensurables y su estructura es un conjunto de ecuaciones que conectan las entradas y las salidas.

Debido a que todos los modelos hidrológicos son aproximaciones de la realidad, la salida de un sistema nunca se puede predecir con exactitud y el problema es cómo lograr un modelo aceptable y operacional. Los numerosos modelos hidrológicos que ya existen varían en su construcción, en parte porque los modelos sirven de alguna manera diferentes propósitos. Existen modelos para el diseño de sistemas de drenaje, modelos para la predicción de inundaciones, modelos para la calidad del agua, etc (Jonsdottir *et al.*, 2006).

Como consecuencia de la naturaleza compleja del proceso lluvia-escurrimiento, determinado por varios procesos de agua, energía y vegetación altamente interconectados en diversas escalas espaciales, los hidrólogos confían en su propia comprensión del sistema, obtenida a través de la interacción con el mismo y de la observación y los experimentos. Este proceso se conoce como modelado perceptual. La percepción de un sistema hidrológico conduce a los modeladores a una variedad de formas de clasificar los modelos del proceso lluvia-escurrimiento: desde modelos deterministas a estocásticos, desde modelos físicos (caja blanca) a empíricos (caja negra) y conceptuales, y los más distintivos, desde modelos agregados a distribuidos (Moradkhani y Sorooshian, 2009).

De acuerdo con lo anterior, autores como Alonso (2015) manifiestan que los modelos pueden ser divididos en dos categorías: los primeros de base física o teórica, también conocidos como caja blanca, y modelos empíricos, conocidos como caja negra. A estos, es frecuente añadirles una tercera categoría: los modelos conceptuales.

Sin embargo, según Chow *et al.* (1988) y Triviño y Ortiz (2004) los modelos hidrológicos pueden dividirse en dos categorías: modelos físicos y modelos abstractos. Los primeros incluyen modelos que representan al sistema en una escala reducida; y modelos análogos, que usan otro sistema físico con propiedades similares a las del prototipo. Los modelos abstractos representan al sistema en forma matemática. La

operación del sistema se describe por medio de un conjunto de ecuaciones que relacionan las variables de entrada y de salida. Estas variables pueden ser funciones del espacio y del tiempo, y también pueden ser variables probabilísticas o aleatorias que no tienen un valor fijo en un punto particular del espacio y del tiempo, pero que están descritas a través de distribuciones de probabilidad (Chow *et al.*, 1988). A su vez, estos modelos, de acuerdo con la aleatoriedad de las variables empleadas, pueden ser estocásticos o deterministas (Triviño y Ortiz, 2004).

Un modelo determinístico no considera aleatoriedad; una entrada produce siempre una misma salida. Sin embargo, un modelo estocástico tiene salidas que son por lo menos parcialmente aleatorias. A pesar de que todos los fenómenos hidrológicos implican algún grado de aleatoriedad, la variabilidad resultante en la salida puede ser pequeña cuando se le compara con la variabilidad resultante de otros factores conocidos. En tales casos un modelo determinístico es más apropiado. Si la variación aleatoria es grande, un modelo estocástico es el más adecuado, porque la salida real podría ser bastante diferente del valor único producido por un modelo determinístico (Chow *et al.*, 1988).

Existen diferentes clasificaciones de los modelos estocásticos y determinísticos. Según Chow *et al.* (1988) dos criterios fundamentales son: la variación espacial y la variación temporal (Figura 1.3). De acuerdo al primero, los modelos determinísticos se dividen en agregados o distribuidos. En un modelo determinístico agregado, el sistema es promediado en el espacio o considerado como un punto único sin dimensiones en el espacio. En contraste un modelo determinístico distribuido considera que los procesos hidrológicos ocurren en varios puntos del espacio y define las variables del modelo como funciones de las dimensiones espaciales. A su vez, los modelos estocásticos se clasifican en independientes del espacio o correlacionados en él de acuerdo con la influencia que las variables aleatorias tengan entre ellas en diferentes puntos del espacio.

Considerando el segundo criterio: la variabilidad temporal, pueden clasificarse los modelos determinísticos en modelos de flujo permanente (la tasa de flujo no cambia con el tiempo) y modelos de flujo no permanente. Los modelos estocásticos siempre tienen salidas que son variables del tiempo. Estos se dividen en independientes del tiempo y correlacionados con él. Un modelo independiente del tiempo representa una secuencia de eventos hidrológicos que no influyen entre sí, mientras que un modelo correlacionado en el tiempo representa una secuencia en la cual el evento siguiente está parcialmente influido por el evento anterior y posiblemente por otros.

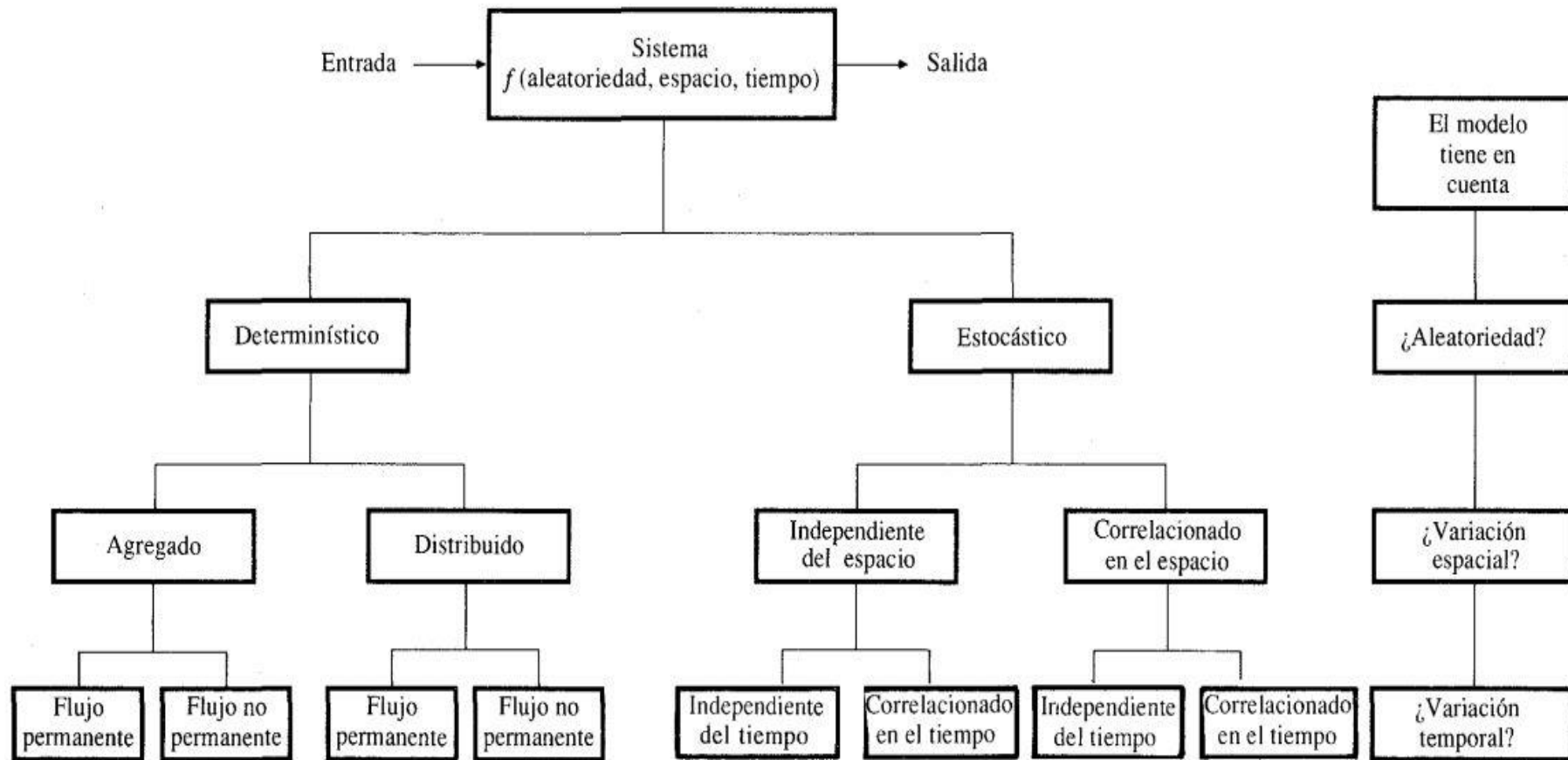


Figura 1.3 Clasificación de modelos hidrológicos de acuerdo con la forma, la aleatoriedad y la variabilidad espacial y temporal de los fenómenos hidrológicos. Tomado de Chow et al. (1988)

Como ejemplo de los modelos hidrológicos determinísticos, se destaca el HEC- HMS (Hydrologic Modeling System, por sus siglas en inglés) (Juárez *et al.*, 2010; Azam *et al.*, 2017). Este programa proporciona a sus usuarios varias opciones para simular el comportamiento de la escorrentía superficial (Triviño y Ortiz, 2004). Para trabajar con el HEC-HMS, además de las especificaciones técnicas de la simulación, son imprescindibles los datos de las precipitaciones y de las características morfológicas de la cuenca (Pacheco, 2012). Actualmente, en esta investigación, solamente se cuentan con los datos de escurrimiento y lluvia de la subcuenca V Aniversario, lo cual no constituye una desventaja, pues realizar simulaciones con los modelos estocásticos implica menor uso de recursos que puedan complejizar el proceso de arribar a los mismos resultados, en comparación con los modelos determinísticos (como el HEC-HMS).

Los modelos estocásticos han sido utilizados durante mucho tiempo para evaluar el rendimiento de sistemas de recursos hídricos, determinar capacidad de embalses, analizar la efectividad de reglas de gestión, planificar la implementación de infraestructura de regulación, generar escenarios para simular el comportamiento futuro de sistemas y otras aplicaciones más (Avilés *et al.*, 2016). También han sido muy usados en la generación de registros de precipitación sintética (Burton *et al.*, 2008), en la estimación y análisis del proceso lluvia-caudal (Jonsdottir *et al.*, 2006), en la simulación y pronóstico del escurrimiento en una cuenca (Montanari y Brath, 2004) y para evaluar la sensibilidad del riesgo de los cambios de escorrentía debido a diversas fuentes de incertidumbre (Preston y Jones, 2008).

Autores como Avilés *et al.* (2016) y Díaz y Guevara (2016) han expresado además, que los modelos estocásticos son adecuados para describir las series cronológicas o series de tiempo hidrológico. Los resultados de estos modelos lanzan un conjunto grande de series que conservan las propiedades estadísticas de la muestra histórica. Cada una de estas series genera los potenciales escenarios futuros que son considerados en los modelos de simulación (Avilés *et al.*, 2016).

Finalmente, los múltiples resultados de los modelos de simulación brindan la información necesaria de una manera probabilística para analizar la situación del sistema de recursos hídricos en estudio y permiten al gestor de la cuenca, planificar medidas de gestión de oferta y demanda de agua que maximicen el rendimiento del sistema (Avilés *et al.*, 2016). Debido a la importancia que posee la modelación lineal estocástica (o series temporales) (Díaz y Guevara, 2016), se dedicará el siguiente

subepígrafe a su estudio, por ser vital en la realización de este trabajo en torno al proceso lluvia- escurrimiento.

1.2.1 Modelación lineal estocástica

Autores como Contin (2001), Ríos (2008) y Bustamante (2014) coinciden en que las series temporales o cronológicas no son más que secuencia de datos numéricos u observaciones sobre valores que toma una variable, cada uno de los cuales se asocia con un instante específico del tiempo. Estas series se dividen en dos categorías: determinísticas y estocásticas. Las series determinísticas son aquellas cuyo valor en cualquier instante se puede determinar con certeza y estas series pueden ser periódicas, semiperiódicas, o presentar tendencias o saltos. Las series estocásticas pueden ser estacionarias o no según sus parámetros (media, varianza, etc.) sean independientes o no a través del tiempo (Díaz y Guevara, 2016).

Por otra parte, las series temporales también pueden clasificarse según la forma en que se ofrecen los valores de las mediciones en dos tipos: continuas y discretas. Las series temporales continuas son aquellas cuyos valores se ofrecen de forma permanente, de manera tal que cada uno de ellos representa el estado de la variable en un instante, el cual puede ser tan pequeño como teóricamente se quiera suponer. Sin embargo, son discretas: cuando los valores se ofrecen para intervalos de tiempo, generalmente homogéneos y donde representan la magnitud acumulada del estado de la variable durante ese intervalo (Contin, 2001).

Díaz y Guevara (2016) indican que históricamente tres son los tipos de análisis o procedimientos metodológicos para la evaluación o estudio de series temporales como son: los métodos de descomposición, método de análisis causal y el método de Box–Jenkins. El método de descomposición, consiste en dividir la serie temporal en sus componentes como son: tendencia, ciclo, componente estacional y componente irregular (aleatorio). El método de análisis causal, explica la evolución de la serie temporal mediante una ecuación que relaciona la variable en estudio con otras variables que inciden sobre ella y el Método de Box–Jenkins, explica la evolución de la serie temporal (variable temporal) en función del comportamiento en el pasado de la misma serie temporal.

Como ya se mencionaba con anterioridad, una serie de tiempo puede desglosarse en cuatro componentes fundamentales, que no son directamente observables y de las cuales únicamente se pueden obtener estimaciones (Ríos, 2008):

- **Tendencia:** representa el comportamiento predominante de la serie. Esta puede ser definida vagamente como el cambio de la media a lo largo de un extenso período de tiempo, lo cual concuerda con lo expresado por Contin (2001).
- **Ciclo:** caracterizado por oscilaciones alrededor de la tendencia con una larga duración, y sus factores no son claros. Según Contin (2001), la variación cíclica también es periódica. Aunque los intervalos de tiempo involucrados son mucho más largos, este componente solamente puede apreciarse, de existir, en series muy extensas.
- **Estacionalidad:** es un movimiento periódico que se produce dentro de un periodo corto y conocido. Según Montanari *et al.* (1997), cuando se remueve esta periodicidad de la media de la serie, entonces la serie puede ser considerada estacionaria, lo cual es corroborado por Avilés *et al.* (2016), que asumen que la serie a modelar debe ser estacionaria, cuando sus estadísticos son constantes en el tiempo al menos en media y varianza.
- **Aleatoriedad:** son movimientos erráticos que no siguen un patrón específico y que obedecen a causas diversas. Este componente es prácticamente impredecible. Este comportamiento es propio de todos los tipos de movimientos de una serie de tiempo que no son debido a la tendencia, ni a las variaciones estacionales o fluctuaciones cíclicas.

Por tanto, para el análisis de series hidrológicas es importante identificar y evaluar las tendencias, la estacionalidad y la no normalidad, según Díaz y Guevara (2016). En general, la mayoría de técnicas estadísticas y la teoría de probabilidades aplicados en hidrología son desarrolladas asumiendo variables normalmente distribuidas. Debido a que la mayoría de curvas de frecuencia de variables hidrológicas son asimétricamente distribuidas, se debe transformar estas variables a una distribución normal antes de iniciar la modelación, esto se consigue probando múltiples funciones normalizadoras como por ejemplo la función logaritmo y la función potencia (Avilés *et al.*, 2016).

Según Kwon (2002), el modelado de las series temporales no es más que la técnica o proceso para encontrar el modelo matemático que represente la muestra de la serie, para lo cual es necesario tener en cuenta los siguientes pasos:

1. Selección del tipo de modelo.
2. Identificación de la forma del modelo.
3. Estimación de los parámetros del modelo.
4. Chequeo de diagnóstico del modelo.

Por otra parte, Díaz y Guevara (2016) plantean que primeramente el objetivo no es seleccionar un modelo, sino identificar un conjunto de modelos posibles que sean compatibles con el gráfico de las funciones de autocorrelación simple y parcial de la serie. La identificación de los posibles modelos requiere detectar la estacionaridad (estabilidad) o la no estacionaridad de la serie de tiempo. Si sucede el último caso (series no estacionarias), es necesario realizar transformaciones para obtener series estacionarias. Posteriormente, del conjunto de modelos identificados y después de estimar sus parámetros, se selecciona al modelo que explica mejor al comportamiento de la serie observada (Díaz y Guevara, 2016). Una vez que se ha realizado el ajuste del modelo, se deberán deshacer las operaciones efectuadas a la serie original, mediante las operaciones inversas (Avilés *et al.*, 2016).

Díaz y Guevara (2016) agregan que es usual construir un modelo simplificado de las series de tiempo que explique su aleatoriedad, de manera que sea útil para pronosticar los valores futuros. Para lograr lo anterior, los modelos estocásticos son los más adecuados. Además, Jonsdottir *et al.* (2006) plantean que la modelación estocástica permite una estructura mucho más simple que la modelación determinística, pues algunas de las variaciones observadas en los datos son descritas por la parte estocástica del modelo.

Como ejemplos de modelos estocásticos, lo constituyen: *los modelos autorregresivos* (AR (p), por sus siglas en inglés) los cuales se han utilizado ampliamente en hidrología y recursos hídricos porque tienen un tipo intuitivo de dependencia del tiempo y son los modelos más simples de usar (Kwon, 2002); *los modelos de media móvil* (MA, por sus siglas en inglés), *los modelos autorregresivos de media móvil* (ARMA, por sus siglas en inglés) (Valipour *et al.*, 2013), siendo su estructura la combinación de los modelos AR y MA; *los modelos autorregresivos integrados de media móvil* (ARIMA, por sus siglas en inglés) (Wang *et al.*, 2015; Fernández *et al.*, 2011) y *los modelos ARIMA estacionales* (SARIMA, por sus siglas en inglés) (Durdu, 2010; Modarres, 2007). Todos estos modelos serán abordados, en profundidad, en el Capítulo 2 de este trabajo.

Es de interés destacar, que la modelación de las series temporales se usa generalmente para generar series de tiempo hidrológicas sintéticas y para pronosticar series de tiempo hidrológicas futuras en el campo de la hidrología y los recursos hídricos. El primer aspecto es útil para dimensionar el yacimiento, para determinar el riesgo de falla (o confiabilidad) del suministro de agua para los sistemas de riego y de las capacidades confiables de los sistemas hidroeléctricos y para planificar los estudios de la operación futura del yacimiento y la capacidad de expansión de los

sistemas de suministro de agua. El segundo aspecto es importante para la planificación a corto plazo de la operación del yacimiento, para operaciones en tiempo real y a corto plazo de cuencas o sistemas de ríos, para la operación de planificación durante una sequía en curso, y para aplicaciones similares (Kwon, 2002).

1.3 Sistemas de pronóstico

Según Krzysztofowicz (2001), la incertidumbre de los eventos futuros es la principal razón para predecir. En general, una predicción no elimina la incertidumbre, solamente la reduce y el grado de incertidumbre remanente varía de una ocasión a la otra. Por otra parte, plantea que los avances científicos y tecnológicos del siglo pasado se han estado utilizando para incrementar las soluciones espaciales y temporales, la precisión y tiempo de espera de las predicciones para todas las variables hidrológicas principales: precipitación, temperatura y escurrimiento. Desafortunadamente, para muchos eventos, especialmente para los desastres naturales: tormentas de lluvia intensas, inundaciones repentinas y otros, la predicción permanece lejos de la perfección y, en ocasiones, no alcanzan las crecientes expectativas de la sociedad en cuanto a las alertas oportunas y confiables. A pesar de que existen límites teóricos, tecnológicos y financieros para el pronóstico, existe una enorme fuente de información adicional que no ha sido conveniente explotada. Esto trae consigo que los pronósticos hidrológicos tengan que perfeccionarse con el objetivo de lograr potenciales beneficios para la sociedad.

Una de las aplicaciones más notorias e importantes que posee la predicción de los eventos futuros en el marco de la Hidrología son los sistemas de alerta temprana. Estos sistemas ayudan a preservar, fundamentalmente, la vida humana ante los desastres naturales.

Un sistema de alerta temprana, en general, ya sea contra inundaciones u otro tipo de evento hidrológico o meteorológico, es un paquete integrado de equipos de recopilación y transmisión de datos, modelos de previsión, advertencia y recursos humanos. La cantidad de muertes evitables e innecesarias y daños a la propiedad podría reducirse con un tiempo de advertencia efectivo (Azam *et al.*, 2017). Según Plate (2007), el sistema de alerta temprana no es más que un conjunto de procedimientos diseñados para proteger vidas humanas y minimizar los daños que se esperan de un fenómeno que excede un cierto nivel crítico. Consiste en una serie de partes relacionadas y conectadas: previsión, transformación del pronóstico en una advertencia, transmisión de la advertencia a los responsables de la toma de decisiones locales y conversión de la advertencia en acciones correctivas. Cada uno

de estos componentes debe funcionar de manera eficiente para que un sistema de alerta temprana sea exitoso.

El objetivo detallado de cualquier sistema de alerta, se desglosa en los siguientes elementos: 1) detectar y pronosticar los riesgos y desarrollar mensajes de advertencia de peligro 2) evaluar los riesgos potenciales e integrar la información de riesgo en los mensajes de advertencia 3) difundir mensajes de advertencia oportunos, confiables y comprensibles a las autoridades y a riesgo público 4) a la planificación, preparación y capacitación en emergencias basadas en la comunidad centradas en obtener una respuesta efectiva a las advertencias para reducir el impacto potencial en las vidas y los medios de subsistencia (Azam *et al.*, 2017).

Es de interés destacar que la primera parte de cualquier sistema de alerta es siempre hacer el pronóstico, o sea, estimar la salida futura en algún tiempo de pronóstico T de la forma más precisa posible, comenzando con las mediciones de las cantidades de entrada presentes y pasadas. Por lo tanto, estos sistemas están basados en datos que se necesitan para poder construir un modelo de pronóstico. Como los datos sobre precipitación y otros procesos naturales nunca son precisos, y dado que los modelos hidrológicos siempre simplifican las condiciones reales y añaden incertidumbres, los resultados también son cantidades inciertas, que en un enfoque científico sólido deberían describirse mediante distribuciones de probabilidad (Plate, 2007).

Según Krzysztofowicz (2001), los pronósticos probabilísticos ofrecen cuatro beneficios potenciales. En primer lugar, estos son científicamente más honestos que los pronósticos determinísticos: le permiten al que pronostica admitir la incertidumbre y expresar el grado de certeza. Por otra parte, permiten establecer criterios basados en el riesgo para avisar y advertir sobre determinado fenómeno como las inundaciones, con probabilidades de detección explícitamente establecidas. En tercer lugar, informan al usuario de la incertidumbre y además, proporcionan la información necesaria para tomar decisiones racionales, lo que le permite al usuario tener en cuenta los riesgos de forma explícita. Por último, ofrecen beneficios económicos adicionales que involucran a la sociedad en general.

Las predicciones también son importantes para la fase de respuesta del proceso de alerta temprana, ya que forman la base de los mapas de riesgo, que son mapas que muestran las áreas afectadas o los valores previstos de daño potencial como funciones de los intervalos de recurrencia. De hecho, el uso de tales mapas es una herramienta muy efectiva no solo para planificar medidas de protección en función de un fenómeno determinado, sino también para planificar medidas de socorro y rescate,

como la identificación de rutas de escape y la identificación de terreno seguro (Plate, 2007).

A nivel global se han realizado estudios en torno a los sistemas de alerta temprana, pero los trabajos de Plate (2007), Krzhizhanovskaya *et al.* (2011), Yu *et al.* (2016) y Azam *et al.* (2017), se han enfocado fundamentalmente en estos sistemas para alertar sobre las inundaciones a partir de los pronósticos de las mismas pues estas, en primer lugar, son una de las consecuencias del proceso lluvia-escurrimiento y por otra parte, causan importantes daños en la economía de los países y ponen en riesgo la vida de las personas, como ya ha sido expresado con anterioridad.

Según Azam *et al.* (2017), el sistema de alerta temprana contra inundaciones (EFWS, Early Flood Warning System por sus siglas en inglés) puede clasificarse en manual y automatizado. El EFWS manual es el enfoque más simple y menos costoso, compuesto por un sistema local de recolección de datos, un coordinador comunitario de inundaciones, un procedimiento de pronóstico de inundación fácil de usar, una red de comunicación para distribuir advertencias y un plan de respuesta. El procedimiento de pronóstico de inundación consiste normalmente en tablas, gráficos o cuadros que utilizan la lluvia pronosticada y un índice del potencial de inundación para estimar el pronóstico de inundación. El EFWS automatizado se compone de pluviómetros automáticos, sistema de comunicación, recopilación y procesamiento automatizado de datos, software de pronóstico y microprocesadores de análisis.

Nuestro país no se encuentra a salvo de este tipo de evento pues el archipiélago cubano está expuesto a diversas amenazas o peligros naturales debido a un conjunto de factores entre los que se encuentran su ubicación geográfica, clima y relieve. Según la «Guía para la realización de estudios de riesgo para situaciones de desastres» presentada por el Estado Mayor Nacional de la Defensa Civil de la República de Cuba en agosto del 2005 (Defensa Civil, 2005), existen peligros que se pueden considerar recurrentes, ya que están presentes cada año y en un periodo específico, como los huracanes, depresiones tropicales, penetraciones del mar e intensas sequías. Los eventos hidrometeorológicos severos son la principal amenaza del país, existiendo una gran población que habita en áreas de riesgo por huracanes, tormentas tropicales, lluvias intensas e inundaciones por diferentes causas (Garrido *et al.*, 2013).

En el estudio realizado por Garrido *et al.* (2013), se propone un sistema automatizado de alerta temprana ante el peligro de inundaciones, donde se integran la adquisición y procesamiento de la información con la simulación del proceso en tiempo real, a través

de cuatro componentes principales: modelos de simulación, sistema de información geográfica, adquisición y supervisión de datos en tiempo real y diferido y una base de datos histórica actualizada. Sin embargo, el pronóstico de las inundaciones es obtenido a través de los modelos de simulación determinísticos HEC-HMS, por lo que la modelación lineal estocástica no ha sido exhaustivamente explotada a pesar de los potenciales beneficios que los pronósticos probabilísticos poseen frente a los determinísticos (Krzysztofowicz, 2001).

CAPÍTULO 2: Materiales y métodos

2.1 Área de estudio

El estudio se centra en la subcuenca “V Aniversario”, que es uno de los cierres hidrográficos naturales de la cuenca del Río Cuyaguaje, ubicada en la provincia de Pinar del Río. Esta subcuenca abarca parte de los municipios de Viñales y Minas de Matahambre, con una extensión de 157 km² y se localiza en los 22°24'6"N – 22°35'40"N y 83°47'55"O – 83°56'9"O. La Figura 2.1 muestra la geografía del área de estudio. En su totalidad, la cuenca del Cuyaguaje es considerada una de las principales del país debido a su tamaño (723 Km²) e importancia económica a través de la agricultura (INRH, 2000). Su elongación latitudinal, similar al resto de las grandes cuencas del país, favorece una mayor generación de escorrentía en relación a otras regiones cársticas (CTCH, 2000).

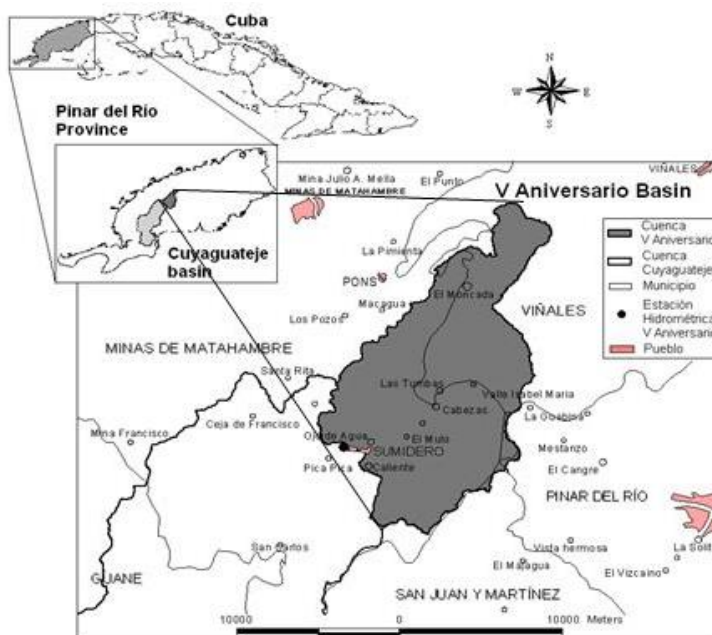


Figura 2.1 Localización geográfica de la subcuenca V Aniversario. Tomado de Alonso (2008).

La región está sujeta a un clima subtropical con períodos secos y húmedos bien definidos. El período seco se extiende desde noviembre a abril, y el resto corresponde con intensas precipitaciones (Alonso, 2018). Las características climáticas anuales (record climatológico 1965-1996) en la subcuenca son: precipitación 1766 mm, evaporación 1860 mm y temperatura 25.1 °C. La humedad relativa oscila alrededor del 77 % en el período seco y 82 % en el lluvioso (INRH, 2000). La Figura 2.2 muestra los regímenes mensuales de la precipitación y la evaporación en la subcuenca V Aniversario (Alonso, 2018).

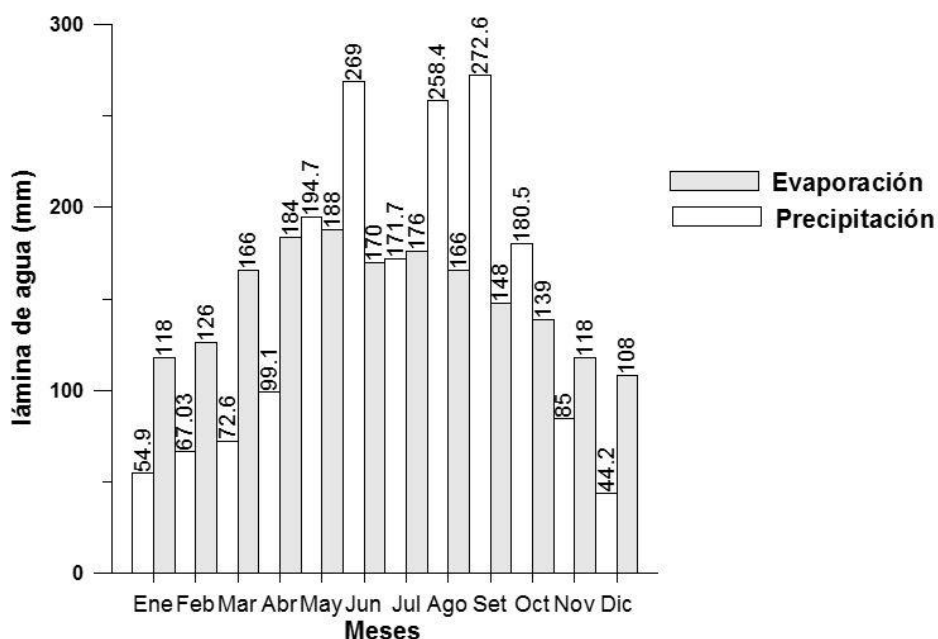


Figura 2.2 Regímenes mensuales de precipitaciones y evaporación (1965-1996). Tomado de Alonso (2008).

La región es comúnmente azotada por ciclones tropicales. Según Furrázola *et al.* (1972), estos fenómenos atmosféricos se originan principalmente en el Océano Atlántico y siguen trayectorias que cruzan frecuentemente sobre las provincias occidentales de Pinar del Río y La Habana. La duración e intensidad de los eventos de precipitaciones se incrementan drásticamente bajo la ocurrencia de un ciclón tropical (CITMA, 2005). Las condiciones meteorológicas bajo su influencia, son tales que, en pocas horas o días puede precipitar una lámina igual o superior a la del resto del año (Planos, 1998). Como resultado, los niveles de escurrimiento son también significativamente afectados. La situación hidrológica puede empeorar atendiendo a la frecuencia de ocurrencia de estos fenómenos, que en ocasiones, puede ser de solo días. Un ejemplo de esto constituye lo ocurrido en el año 2008, donde, la región fue azotada por dos ciclones tropicales de gran intensidad en un intervalo de 10 días. Las pérdidas ambientales y económicas fueron catastróficas (Ballester y Rubiera, 2008).

2.2 Datos de precipitación

Las precipitaciones en la subcuenca han sido monitoreadas desde los años 60s del pasado siglo. La red de estaciones está conformada por 8 pluviómetros que miden totales diarios de precipitación y, en 2 de ellas, se mide también la intensidad (acumulados en 5 minutos). Estas estaciones son manejadas por el Instituto Nacional de Recursos Hidráulicos (INRH) de Pinar del Río (Alonso, 2018). El periodo activo de mediciones cubre desde 1963 al 2006 en el caso de la lluvia diaria, con algunos años

perdidos para la mayoría de las estaciones. Las series de intensidad cubren desde 1964 al 2010, pero con solo 20 años de mediciones en este intervalo (Alonso, 2016).

La Figura 2.3a muestra los valores históricos de precipitación diaria disponible en la subcuenca, los cuales fueron calculados como la mediana de las 8 estaciones que monitorean la región. Se usó la mediana como medida de centro para un análisis más robusto y para evitar mayor influencia de posibles valores absurdos provocados por errores de medición o manejo de datos. El tiempo de referencia, fijado en día 0 en el eje x, corresponde al 1 enero de 1963. En la Figura 2.3b se muestra la función de densidad acumulada de esta serie, donde, el eje y usa una escala de visualización normal (Alonso, 2018).

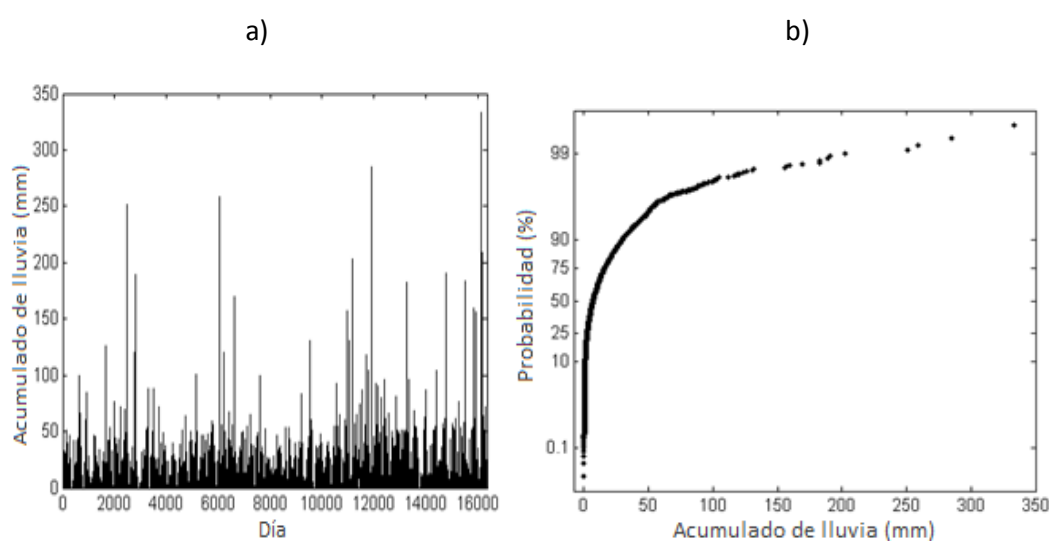


Figura 2.3 a) Precipitación diaria en la subcuenca V Aniversario en el período 1963-2006. b) Función de densidad acumulada de los valores de precipitación observados. Tomado de Alonso (2018).

2.3 Serie de escurrimiento

Estas variables son monitoreadas en la desembocadura de la subcuenca (estación hidrométrica V Aniversario) por los métodos tradicionales, usando molinete y la medición constante del nivel con un limnómetro. Las mediciones frecuentes se realizan en uno o pocos puntos de control de la sección del río, ubicaciones estas que han sido previamente relacionadas, mediante aforos más detallados, con la media de toda la sección. La frecuencia de las mediciones depende de las condiciones del flujo. En condiciones normales, los aforos se realizan diariamente a las 08:00 horas y durante el período de estiaje (período de seca) se pueden distanciar a 5 o 10 días. Por otra parte, en crecidas es necesario tomar varias muestras durante el evento y la cantidad y momento dependen del tipo de crecida (Alonso, 2018).

El dato disponible para este estudio es el escurrimiento medio diario, medido durante el período de 1971 a 1990 (Alonso, 2016). Es de interés destacar que, en este período, la estación hidrométrica de la subcuenca V Aniversario ha registrado un escurrimiento medio anual de $3,72 \text{ m}^3/\text{s}$ y un máximo observado de $453 \text{ m}^3/\text{s}$, la media mensual oscila entre $0,7$ y $8,9 \text{ m}^3/\text{s}$ (CTCH, 2000).

En este estudio, los valores medios diarios fueron convertidos en valores medios anuales y mensuales. En el caso de los valores medios anuales, la conversión se realizó hallando la media de todos los valores medios diarios en cada uno de los años. En el caso de los valores medios mensuales, se halló la media de los valores medios diarios en cada uno de los meses. En la Figura 2.4a se refleja el escurrimiento medio diario, así como la función empírica de densidad acumulada (Figura 2.4b).

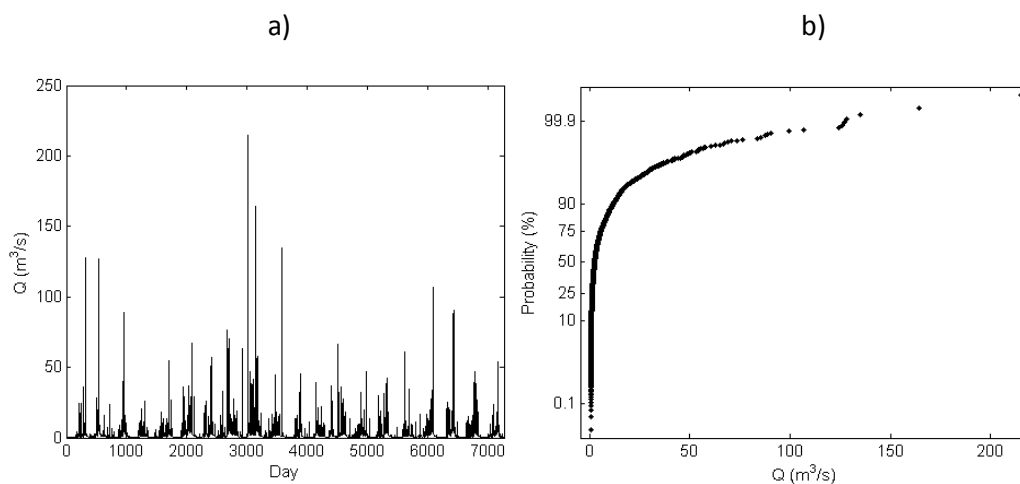


Figura 2.4 a) Escurrimiento medio diario. b) Función empírica de densidad acumulada del escurrimiento medio diario.

2.4 Modelos lineales estocásticos

La modelación lineal estocástica del escurrimiento en ríos da como resultado una respuesta hidrológica que se predice a partir de los pasos anteriores. Los modelos de Ruido Blanco y de la familia ARMA (Cowpertwait y Metcalfe, 2009): autorregresivos (AR(p), por sus siglas en inglés), de media móvil (MA(q), por sus siglas en inglés) y sus variantes, permiten obtener información adicional acerca de las series de tiempo, como su comportamiento estacional. Debido a la importancia que poseen estos modelos para arribar a los resultados en esta investigación, se abordarán brevemente a continuación.

2.4.1 Ruido Blanco (WN)

Un error residual es la diferencia entre el valor observado y el valor predicho por el modelo en un tiempo t . Si se supone que el modelo es definido por la variable y_t y $\hat{y}_t$ es el valor predicho por el modelo, el error residual x_t es:

$$x_t = y_t - \hat{y}_t \quad \text{Ec. 2.1}$$

Como los errores residuales ocurren en el tiempo, ellos forman una serie temporal: $x_1, x_2, \dots, x_n$.

Una serie de tiempo $\{w_t : t = 1, 2, \dots, n\}$ es un Ruido Blanco Discreto (DWN) si las variables $w_1, w_2, \dots, w_n$ son independientes e idénticamente distribuidas con media cero. Lo anterior implica que todas las variables tienen la misma varianza σ^2 y $\text{Cor}(w_i, w_j) = 0$ para toda $i \neq j$. Si, además, las variables siguen una distribución normal ($w_t \sim N(0, \sigma^2)$) la serie es nombrada Ruido Blanco Gaussiano (Cowpertwait y Metcalfe, 2009).

2.4.2 Modelo autorregresivo AR(p)

La serie $\{x_t\}$ es un modelo autorregresivo de orden p , abreviado como AR(p), si:

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_p x_{t-p} + w_t \quad \text{Ec. 2.2}$$

donde $\{w_t\}$ es un Ruido Blanco y los α_i son los parámetros del modelo con $\alpha_p \neq 0$ para un orden p del proceso. La Ec. 2.2 puede ser expresada como un polinomio de orden p en términos del operador de retroceso hacia atrás (**B**) (Cowpertwait y Metcalfe, 2009):

$$\theta_p(\mathbf{B})x_t = w_t \quad \text{Ec. 2.3}$$

El valor de la serie x_t es una combinación lineal de los p valores pasados más recientes de sí misma, más un término de “innovación” w_t que incorpora todo lo nuevo en la serie en un tiempo t , que no es explicado por los valores pasados. Por tanto, para cada t , se asume que w_t es independiente de $x_{t-1}, x_{t-2}, x_{t-3}, \dots$ (Cryer y Chan, 2008).

Los siguientes puntos deben tenerse en cuenta:

- El modelo es una regresión de x_t en los términos pasados de la misma serie, de ahí el uso del término ‘autorregresivo’ (Cowpertwait y Metcalfe, 2009).
- p determina el número de pasos en el pasado necesarios para predecir el actual valor de la serie x_t (Shumway y Stoffer, 2011).
- Los parámetros del modelo pueden ser estimados minimizando la suma de los errores cuadrados (Cowpertwait y Metcalfe, 2009).

Es de interés destacar que los modelos AR pueden tener parámetros constantes para el modelado de las series temporales a escala anual, mientras que, por otro lado, ellos pueden tener parámetros que varían con el tiempo o una combinación de ambos (parámetros variables y constantes) para el modelado de series temporales que tengan una escala que sea una fracción de un año. Sin embargo, estos modelos presentan algunas limitaciones. En primer lugar, necesita ser chequeado si los modelos AR pueden mantener la estacionaridad para la estadística de los valores históricos o no, esto es importante cuando el Método de Máxima Verosimilitud (ML, por sus siglas en inglés) es usado para estimar los parámetros y el retardo de tiempo es incrementado. Otras limitaciones adicionales surgen si los datos están normalmente distribuidos y si la estacionaridad del primer y/o segundo momento es mantenida o no (Kwon, 2002).

2.4.3 Modelo de media móvil MA(q)

Un modelo de media móvil (MA) de orden q es una combinación lineal del término actual de Ruido Blanco y de los q términos pasados más recientes de Ruido Blanco. Este es definido como:

$$x_t = w_t + \beta_1 w_{t-1} + \dots + \beta_q w_{t-q} \quad \text{Ec. 2.4}$$

donde $\{w_t\}$ es un Ruido Blanco con media cero y varianza σ_u^2 . La Ec. 2.4 puede ser nuevamente escrita en términos del operador de retroceso hacia atrás $\mathbf{B}$ y del polinomio ϕ_q de orden q como:

$$x_t = \phi_q(\mathbf{B})w_t \quad \text{Ec. 2.5}$$

Debido a que los procesos MA consisten en una suma finita de términos estacionarios de Ruido Blanco, son entonces estacionarios y, por lo tanto, tienen una media y autocovarianza invariantes en el tiempo. La media es simplemente cero porque es una suma de términos que tienen una media de cero. La varianza es $\sigma_u^2(1 + \beta_1^2 + \dots + \beta_q^2)$ porque cada uno de los términos de Ruido Blanco tiene la misma varianza y son mutuamente independientes. La función de autocorrelación para $k \geq 0$, está dada por:

$$p(K) = \begin{cases} 1 & k = 0 \\ \sum_{i=0}^{q-k} \beta_i \beta_{i+k} / \sum_{i=0}^q \beta_i^2 & k = 1, \dots, q \\ 0 & k > q \end{cases} \quad \text{Ec. 2.6}$$

donde β_0 es la unidad. La función es cero cuando $k > q$ porque x_t y x_{t+k} son una suma de términos independientes de Ruido Blanco y que tienen covarianza cero.

Un proceso MA es reversible si puede ser expresado como un proceso autorregresivo estacionario de un orden infinito, sin un término de error. En general, la propiedad de reversibilidad se cumple cuando todas las raíces de $\phi_q(\mathbf{B})$ exceden la unidad en un valor absoluto. La función de autocovarianza solamente identifica un único proceso MA(q), si la condición de que este sea reversible es impuesta (Cowpertwait y Metcalfe, 2009).

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

Un modelo muy útil es obtenido cuando los términos del AR y MA son juntados en una sola expresión. Una serie de tiempo $\{x_t\}$ sigue un modelo autorregresivo de media móvil (ARMA) de orden (p, q), denotado como ARMA(p, q), cuando:

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_p x_{t-p} + w_t + \beta_1 w_{t-1} + \beta_2 w_{t-2} + \dots + \beta_q w_{t-q} \quad \text{Ec. 2.7}$$

donde $\{w_t\}$ es un Ruido Blanco. La Ec. 2.7 puede ser representada en términos del operador de retroceso hacia atrás $\mathbf{B}$ y puede ser reajustada en la forma de un polinomio más conciso:

$$\theta_p(\mathbf{B})x_t = \phi_q(\mathbf{B})w_t \quad \text{Ec. 2.8}$$

Los siguientes puntos deben tenerse en cuenta en un modelo ARMA(p, q):

- El proceso es estacionario cuando todas las raíces de θ exceden la unidad en un valor absoluto.
- El proceso es reversible cuando todas las raíces de ϕ exceden la unidad en un valor absoluto.
- El modelo AR(p) es el caso especial cuando ARMA(p, 0).
- El modelo MA(q) es el caso especial cuando ARMA(0, q).
- Cuando se ajustan los modelos a los datos, un modelo ARMA será frecuentemente más eficiente en los parámetros que un modelo MA o AR por separado.
- Cuando θ y ϕ comparten un factor común, un modelo estacionario puede ser simplificado (Cowpertwait y Metcalfe, 2009).

Los supuestos para los modelos ARMA son que las series temporales sean estacionarias y distribuidas de forma normal aproximadamente. Por consiguiente, si las series temporales no cumplen con el último aspecto, deben ser normalizadas usando la ecuación de transformación adecuada. Una ventaja de usar modelos ARMA es tener menos parámetros que estimar comparado con los altos órdenes de los modelos AR debido a la flexibilidad (Kwon, 2002).

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

Durante más de medio siglo, la metodología ARIMA de Box-Jenkins ha sido un procedimiento popular para aplicaciones de análisis y pronóstico de series temporales. El modelo ARIMA se origina a partir del modelo autorregresivo (AR), el modelo de media móvil (MA) y la combinación de los modelos AR y MA (ARMA) (Wang *et al.*, 2015).

Una serie temporal $\{x_t\}$ sigue un proceso ARIMA(p,d,q) si las diferencias d-ésimas de la serie $\{x_t\}$ son un proceso ARMA(p,q). Si se introduce $y_t = (1 - B)^d x_t$, entonces $\theta_p(B)y_t = \phi_q(B)w_t$. Ahora se está en condiciones de sustituir y_t para obtener la forma más sucinta de un proceso ARIMA(p,d,q) como:

$$\theta_p(B)(1 - B)^d x_t = \phi_q(B)w_t \quad \text{Ec. 2.9}$$

donde θ_p y ϕ_q son polinomios de orden p y q, respectivamente (Cowpertwait y Metcalfe, 2009).

Si el proceso no contiene términos autorregresivos, se denomina modelo integrado de media móvil y se abrevia como IMA(d,q). Si, de lo contrario, no presenta términos de media móvil, se denota como modelo autorregresivo integrado ARI(p,d) (Cryer y Chan, 2008).

2.4.6 Modelo estacional autorregresivo integrado de media móvil SARIMA (p,d,q)(P,D,Q)

Los modelos estacionales ARIMA son herramientas poderosas en el análisis de series temporales, pues son capaces de modelar una amplia gama de series. Gran parte de la metodología fue iniciada por Box y Jenkins en los años 70.

Un modelo SARIMA (o ARIMA estacional) utiliza la diferenciación en un retardo igual al número de estaciones o temporadas para eliminar los efectos estacionales aditivos. Al igual que con la diferenciación en un retardo 1 para eliminar una tendencia, la diferenciación en un retardo s introduce un término de media móvil, es por ello que estos modelos incluyen términos autorregresivos y de media móvil en un retardo s. De forma más sucinta, el modelo SARIMA (p, d, q)(P, D, Q) puede expresarse, usando el operador de retroceso hacia atrás, como:

$$\theta_p(B^s)\theta_p(B)(1 - B^s)^D(1 - B)^d x_t = \phi_q(B^s)\phi_q(B)w_t \quad \text{Ec. 2.10}$$

donde θ_p , θ_p , ϕ_q , y ϕ_q son polinomios de órdenes P, p, Q, y q, respectivamente (Cowpertwait y Metcalfe, 2009). Durdu (2010) plantea que el modelo SARIMA puede ser explicado brevemente como ARIMA(p, d, q)(P, D, Q)_s, donde (p, d, q) es la parte

no estacional del modelo y (P, D, Q) constituye el componente estacional. Además, agrega que p es el orden de la autorregresión no estacional, d es el número de diferenciación regular, q es el orden no estacional MA, P es el orden de la autorregresión estacional, D es el número de diferenciación estacional, Q es el orden MA estacional, s es la duración de la estación o temporada, θ es el parámetro estacional AR de orden P y ϕ es el parámetro estacional MA de orden Q .

En general, el modelo es no estacionario, aunque si $D = d = 0$ y las raíces de la ecuación característica (los términos polinomiales en el lado izquierdo de la Ec. 2.10) exceden la unidad en valor absoluto, el modelo resultante puede ser estacionario.

Es importante estar al tanto de que la diferenciación en un retardo s eliminará una tendencia lineal, por lo que existe la opción de incluir o no la diferenciación en un retardo 1. Si se incluye la diferenciación en un retardo 1, cuando una tendencia lineal es apropiada, esta introducirá términos de media móvil en una serie de Ruido Blanco.

Los modelos estacionales ARIMA pueden tener potencialmente una gran cantidad de parámetros y combinaciones de términos. Por lo tanto, es apropiado probar una amplia gama de modelos cuando se ajustan a los datos y elegir un modelo que se adapte mejor usando un criterio apropiado como el Criterio de Información de Akaike (AIC, por sus siglas en inglés). Una vez que se ha sido encontrado un modelo que se ajuste mejor, el correlograma de los residuos debe ser verificado como Ruido Blanco. Adicionalmente, se puede ganar confianza en el modelo que mejor se ajusta, si deliberadamente se sobre ajusta el modelo incluyendo parámetros adicionales y observando un incremento en el AIC (Cowpertwait y Metcalfe, 2009).

Una ventaja inherente de los modelos SARIMA es que se requieren pocos parámetros para describir series de tiempo que muestren no estacionaridad, tanto dentro como a través de las estaciones o temporadas (Durdu, 2010).

2.5 Análisis estadístico

Para decidir el modelo que describa el comportamiento de los datos estudiados, es necesario realizar elecciones basadas en criterios y test estadísticos. A continuación, se describen el Criterio de Información Akaike (AIC, por sus siglas en inglés) (Shumway y Stoffer, 2011) y el test de tendencia de Mann-Kendall (Wilks, 2011), los cuales son fundamentales en este trabajo.

2.5.1 Criterio de Información de Akaike (AIC)

Este criterio está dado por la siguiente definición:

$$AIC = \log \hat{\sigma}_k^2 + \frac{n + 2k}{n} \quad \text{Ec. 2.11}$$

donde k es el número de parámetros en el modelo y $\hat{\sigma}_k^2$ está dado por:

$$\hat{\sigma}_k^2 = \frac{SSE_k}{n} \quad \text{Ec. 2.12}$$

SSE_k denota la suma residual de cuadrados en el modelo con k coeficientes de regresión.

El valor de k que arroje el AIC mínimo, es el mejor modelo. La idea es minimizar $\hat{\sigma}_k^2$, lo cual sería un objetivo razonable si este no decreciera monótonamente a medida que k se incrementa. Entonces, se necesita penalizar la varianza del error por un término proporcional al número de parámetros. La elección del término de penalización, en este caso, está dada por la Ec. 2.11.

Formalmente, el AIC es definido como $AIC = -2 \log L_k + 2k$ donde L_k es la máxima verosimilitud y k es el número de parámetros en el modelo. Para el problema de regresión normal, el AIC puede ser reducido a la forma dada por la Ec. 2.11. Es de interés destacar que el AIC es un estimado de la discrepancia de Kullback-Leibler entre un modelo verdadero y un modelo candidato.

2.5.2 Test de tendencia de Mann-Kendall

El test de tendencia de Mann-Kendall es una alternativa no paramétrica popular para probar la presencia de una tendencia, o que la tendencia central sea no estacionaria, de una serie temporal. En el contexto de examinar la posibilidad de tendencia que corresponde a una serie temporal x_i , donde i es el índice de tiempo (por ejemplo: el año de observación de cada dato) y toma valores $i = 1, \dots, n$, la tendencia es, por definición, monótonamente creciente, lo que simplifica los cálculos.

El test estadístico para la tendencia de Mann-Kendall es

$$S = \sum_{i=1}^{n-1} \text{sgn}(x_{i+1} - x_i) \quad \text{Ec. 2.13}$$

donde

$$\text{sgn}(\Delta x) = \begin{cases} +1, & \Delta x > 0 \\ 0, & \Delta x = 0 \\ -1, & \Delta x < 0 \end{cases} \quad \text{Ec. 2.14}$$

La estadística en la Ec. 2.13 cuenta el número de pares de datos adyacentes en los que el primer valor es más pequeño que el segundo y subtrae el número de pares de datos en el que el primero es más grande que el segundo. Si los datos x_i son consecutivamente independientes y son obtenidos de la misma distribución (en particular, si el proceso de generación tiene la misma media a lo largo de la serie temporal), entonces los números de pares de datos adyacentes para los cuales $sgn(\Delta x)$ es positivo y negativo deberían ser casi igual.

Para longitudes de series, ya sean moderadas (n aproximadamente 10) o mayores, la distribución de la muestra del test estadístico en la Ec. 2.13 es aproximadamente gaussiana, y si la hipótesis nula (no existe tendencia) es cierta, esta distribución gaussiana nula tendrá media cero. La varianza de esta distribución depende de si todas las x son distintas o si algunas son valores repetidos. Si no hay vínculos, la varianza de la distribución de la muestra de S es:

$$Var(S) = \frac{n(n-1)(2n+5)}{18} \quad Ec. 2.15$$

De lo contrario, la varianza es

$$Var(S) = \frac{n(n-1)(2n+5) - \sum_{j=1}^J t_j(t_j-1)(2t_j+5)}{18} \quad Ec. 2.16$$

donde J indica el número de grupos de valores repetidos, y t_j es el número de valores repetidos en el grupo j -ésimo. El test del p-valor es evaluado usando el valor gaussiano estándar.

$$z = \begin{cases} \frac{S-1}{[Var(S)]^{1/2}}, & S > 0 \\ \frac{S+1}{[Var(S)]^{1/2}}, & S < 0 \end{cases} \quad Ec. 2.17$$

La media de S en la hipótesis nula es cero; sin embargo, el ± 1 en el numerador de la Ec. 2.17 representa una corrección de continuidad.

Si todos los n valores de datos en la serie son distintos, la relación de la Ec. 2.13 con el valor τ de Kendall que caracteriza la relación entre las x y el índice de tiempo i es

$$S = \binom{n}{2} \tau = \frac{n(n-1)}{2} \tau \quad Ec. 2.18$$

2.6 Metodología

La metodología de trabajo que se presenta, sienta las bases para la obtención de los resultados en esta investigación, los cuales serán posteriormente discutidos en el Capítulo 3. La Figura 2.5 resume gráficamente las etapas propuestas. Es de interés destacar que esta metodología tributa al objetivo de este trabajo, pues una vez que se obtiene el modelo adecuado para las series de escurrimiento en las diferentes escalas de tiempo, se está en condiciones de pronosticar con antelación el flujo en canales de la subcuenca V Aniversario.

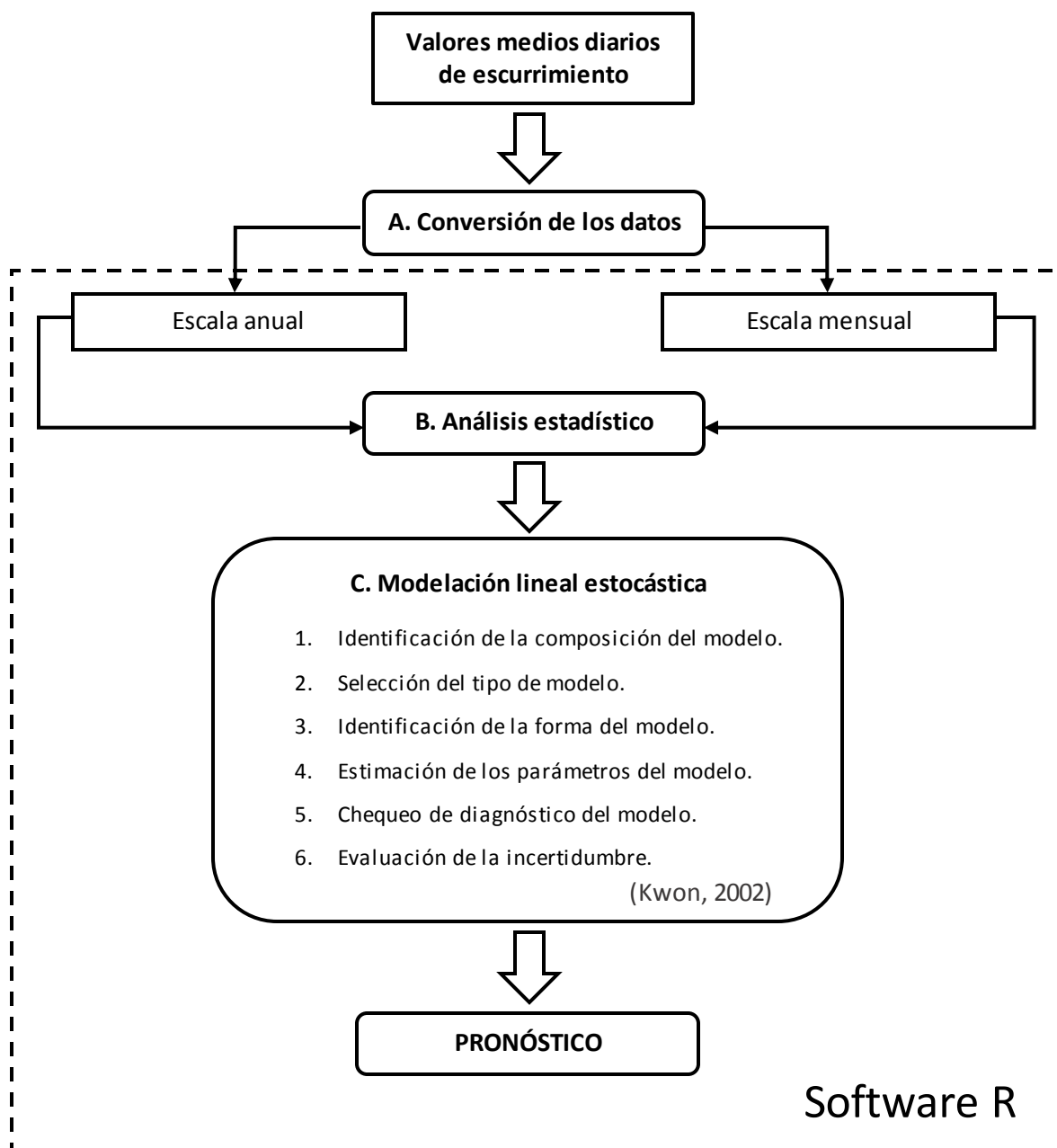


Figura 2.5 Metodología de trabajo.

De forma breve, se describen a continuación las etapas por las que transita esta metodología de trabajo:

A. Conversión de los datos.

En esta primera etapa, los valores medios diarios de escurrimiento son convertidos en valores medios anuales y mensuales, tal y como se explica en el epígrafe 2.3 de este mismo capítulo, conformándose de esta manera las series de escurrimiento a escalas anual y mensual. Para lograr lo anterior, se utilizó una hoja de cálculo.

B. Análisis estadístico.

Una vez que se obtienen las series de escurrimiento a escala anual y mensual, se procede a hacer el análisis de la estadística básica de cada una de ellas en el software R versión 3.0.2. Para este análisis, se debe hacer especial énfasis en los estadísticos de media, varianza, desviación estándar, coeficiente de variación y de asimetría, así como también en los gráficos de normalidad.

C. Modelación lineal estocástica.

Los seis pasos que son propuestos en esta etapa constituyen una aproximación esquemática para modelar apropiadamente las series de tiempo hidrológico. Además, estos pasos pueden ayudar a disminuir las diferencias entre la teoría y la aplicación práctica resultantes de una estimación aproximada por el uso de datos limitados (Kwon, 2002). Debido a la importancia que tienen estos pasos en el desarrollo de este trabajo, así como de investigaciones afines, se abordarán brevemente a continuación:

1. Identificación de la composición del modelo:

En este primer paso, debe ser tomada la decisión de que el modelo sea univariado, multivariado o una combinación de ambos aspectos.

2. Selección del tipo de modelo:

Una vez que se haya realizado lo anterior, debe decidirse el modelo adecuado entre los diversos modelos tales como: $AR(p)$, $ARMA(p,q)$, $ARIMA(p,d,q)$, etc., en relación con las características de los procesos físicos hidrológicos, las características de las series de tiempo hidrológicas y las experiencias del modelador (Kwon, 2002).

Para encontrar el mejor modelo que represente el comportamiento de las series temporales, las funciones de autocorrelación (ACF, por sus siglas en inglés) y de autocorrelación parcial (PACF, por sus siglas en inglés) son usualmente graficadas. La función ACF es la correlación lineal entre la serie temporal y la misma serie con lapsos de tiempo. Por otro lado, la función

PACF es la autocorrelación que se mantiene después de que el ACF es explicado (Fernández *et al.*, 2011). Del comportamiento de las funciones ACF y PACF, se observa si una serie es estacionaria o no, estacional o no. Es de interés destacar que en el caso de que una serie sea no estacionaria, la diferenciación debe ser hecha para obtener series de tiempo estacionarias (Modarres, 2007).

3. Identificación de la forma del modelo:

En este paso, se deben determinar los órdenes de los modelos lineales estocásticos, por ejemplo: los órdenes p y q para un modelo ARMA (Kwon, 2002). Para dar cumplimiento a lo anterior, las funciones ACF y PACF son empleadas. Sin embargo, una simple inspección de los gráficos de las funciones ACF y PACF, en general, no deben dar valores claros, por ejemplo, de p y q . Es por ello que se han empleado adicionalmente otros métodos de identificación de modelos como el Criterio de Información de Akaike (AIC) (Wang *et al.*, 2015).

4. Estimación de los parámetros del modelo:

Después de que los modelos son identificados, se necesita obtener una estimación eficiente de los parámetros. Estos parámetros deben satisfacer dos condiciones nombradas: estacionaridad e invertibilidad para los modelos autorregresivos y de media móvil, respectivamente. Por otra parte, también debería ser comprobado si los parámetros son estadísticamente significativos o no (Modarres, 2007), lo cual puede ser evidenciado a partir de un test de significación de los coeficientes (si el p -valor es inferior a 0.05, los coeficientes son estadísticamente significativos en un nivel de confianza del 95%) (Fernández *et al.*, 2011).

En adición a lo anterior, se debe seleccionar el método adecuado de estimación de parámetros entre el Método de los Momentos (MOM, por sus siglas en inglés), Método de Máxima Verosimilitud (ML, por sus siglas en inglés), etc (Kwon, 2002).

5. Chequeo de diagnóstico del modelo:

El modelo estimado necesita ser chequeado para verificar si cumple con algunos de los supuestos y para conocer qué tan bien este representa las series temporales. Los supuestos son corroborados a partir de los tests de independencia y normalidad de los residuos del modelo. Además, se debe verificar si el modelo reproduce adecuadamente la estadística histórica de la serie como: la media, varianza, asimetría, correlaciones, etc., a través de la generación de datos.

6. Evaluación de la incertidumbre:

Para poder cumplimentar este último paso, es necesario conocer que la incertidumbre del modelo es el resultado de las aproximaciones en la selección del modelo y la incertidumbre de los parámetros se obtiene a partir de la estimación de los parámetros con una cantidad limitada de datos.

CAPÍTULO 3: Resultados y discusión

3.1 Estadística descriptiva

A continuación, se realiza el análisis estadístico de las series de escurrimiento a escalas anual, mensual y diaria. El período seleccionado para llevar a cabo la investigación, en las diferentes escalas, comprende desde el año 1971 a 1990, como es reflejado en el epígrafe 2.3, constituyendo un total de 20 valores de escurrimiento anuales, 240 mensuales y 7300 valores diarios.

3.1.1 Serie de escurrimiento anual

La representación de la serie de datos del escurrimiento medio anual de la subcuenca “V Aniversario” es mostrada en la Figura 3.1a, donde cada punto simboliza el escurrimiento en cada uno de los años y la línea continua: la media de la serie. En la Figura 3.1b, se observa el gráfico de normalidad, donde la distribución de la serie es muy cercana a la normal, con pequeñas desviaciones en la cola.

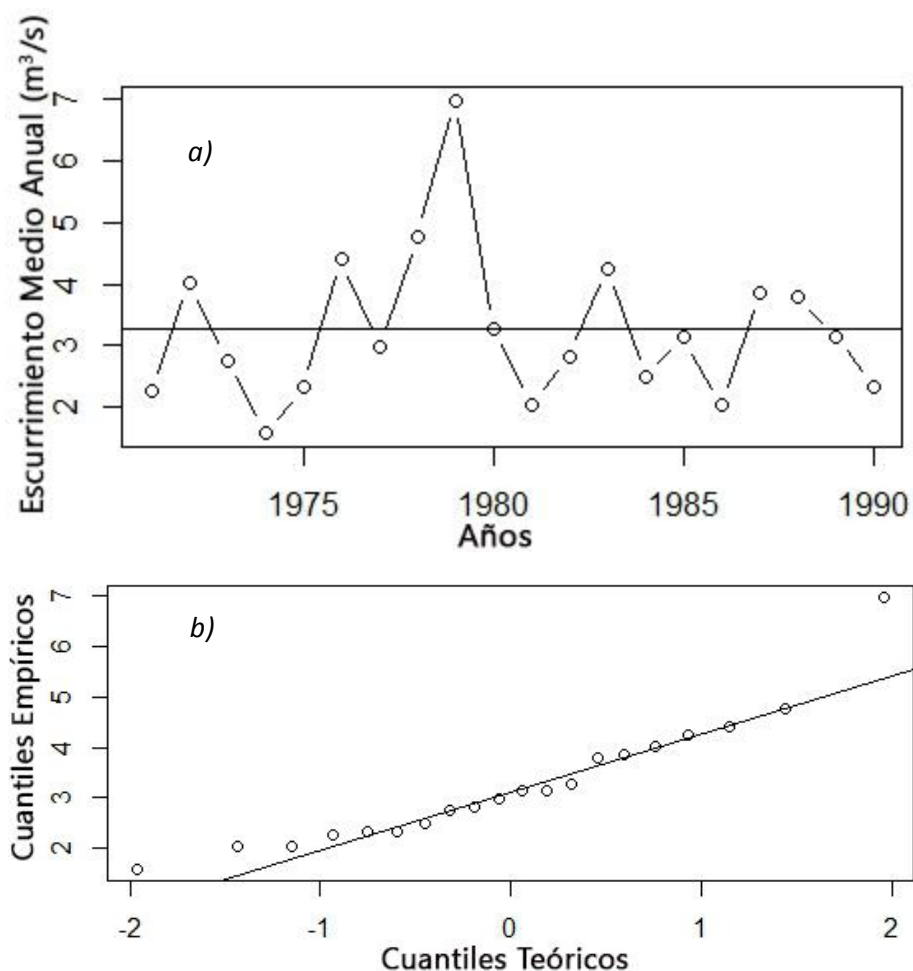


Figura 3.1 a) *Ecurrimiento medio anual de la subcuenca V Aniversario (período 1971-1990).* b) *Diagrama Q-Q.*

La Tabla 3.1 resume la estadística descriptiva básica para esta serie. Según el valor del coeficiente de asimetría (1,41), se determina que la serie es asimétrica a la derecha, patrón típico de las series hidrológicas (Chow *et al.*, 1988) y esto es corroborado por el diagrama Q-Q de la Figura 3.1b. Estas series se caracterizan por poseer unos pocos valores anómalos, o extremos, que se alejan de la zona de mayor densidad.

Tabla 3.1 Estadística del escurrimiento medio anual para la subcuenca V Aniversario.

<i>Estadística</i>										
Anual	Mín	Q ₁	Q ₂	Media	Q ₃	Máx.	S ²	S	Cv	CA
V Aniversario	1,58	2,33	3,06	3,26	3,89	6,98	1,55	1,24	0,38	1,41

Q₁ primer cuartil de la distribución, Q₂ segundo cuartil, Q₃ tercer cuartil, S² varianza, S desviación típica, Cv coeficiente de variación y CA coeficiente de asimetría.

En concordancia con la Figura 3.1a, la Tabla 3.1 refleja, además, que la serie alcanza su valor mínimo de escurrimiento (1,58 m³/s) en el año 1974 y su valor máximo (6,98 m³/s) en 1978. Por otra parte, la media de la serie es de 3,26 m³/s, la mediana: 3,06 m³/s, la varianza: 1,55 m⁶/s² y la desviación estándar es de 1,24 m³/s. Es de interés destacar que el 25 % y el 75% de los datos se encuentran por debajo de 2,33 y 3,89 m³/s, respectivamente.

3.1.2 Serie de escurrimiento mensual

En la Figura 3.2 se muestra la serie del escurrimiento medio mensual de la subcuenca V Aniversario. En esta gráfica no se observa una posible tendencia en los valores, pero sí un comportamiento periódico provocado por la estacionalidad de las precipitaciones. Las sospechas a priori pueden ser comprobadas desglosando la serie en sus componentes fundamentales, tal y como se expuso en el subepígrafe 1.2.1.

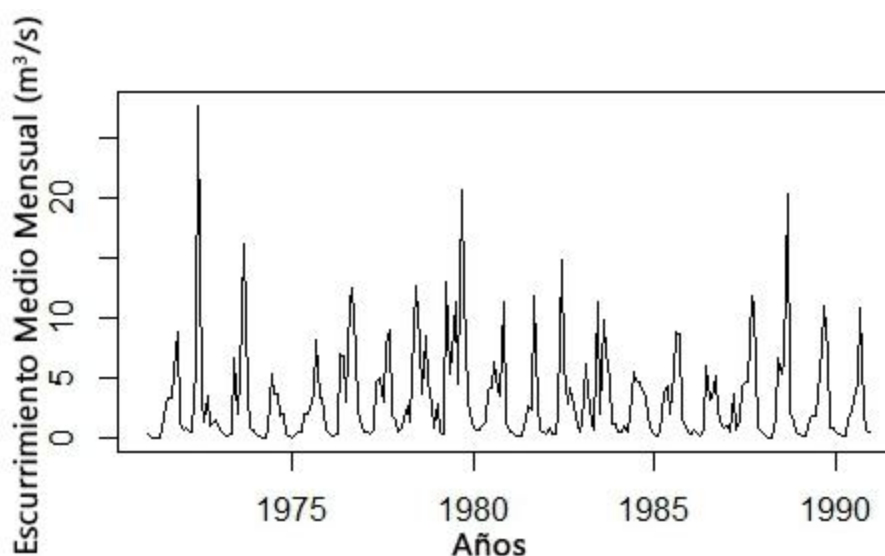


Figura 3.2 Serie de escurrimiento medio mensual de la subcuenca V Aniversario (1971-1990).

En la Figura 3.3a están representados los valores observados y los elementos en los que se descompone la serie temporal: tendencia, estacionalidad y aleatoriedad. Como es esperado, la principal componente a tener en cuenta en esta serie es su carácter estacional. De la Figura 3.3b se evidencia que los datos se distribuyen aproximadamente normal, solamente con pequeñas desviaciones en las colas, tal y como ocurre a escala anual.

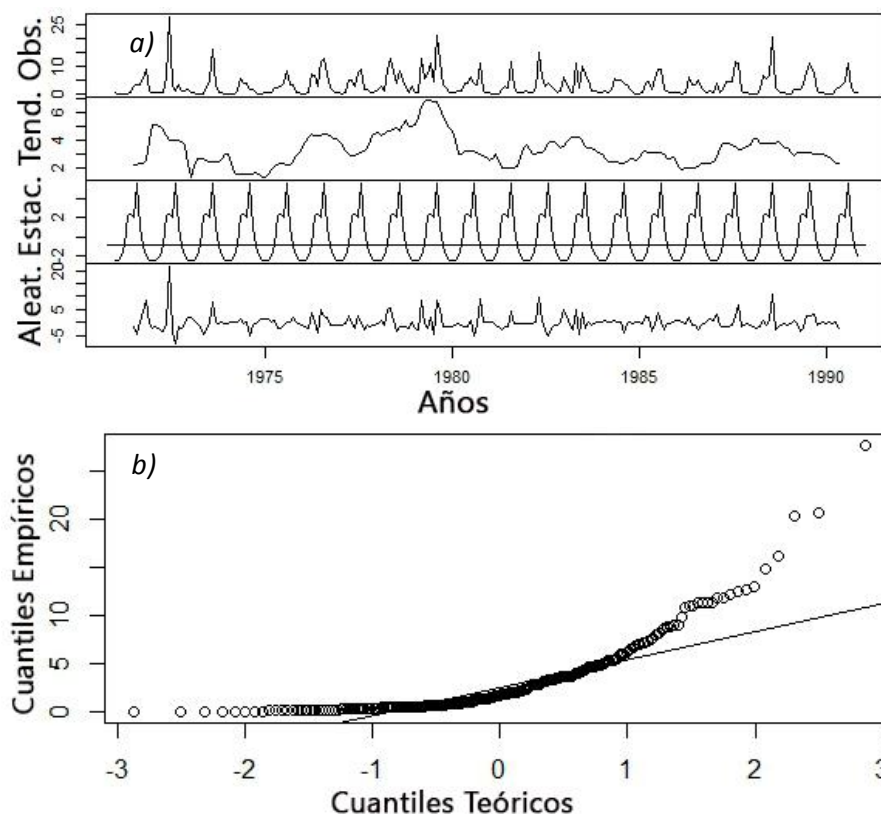


Figura 3.3 Serie de escurrimiento mensual, a) componentes de la serie temporal y b) diagrama Q-Q.

La Tabla 3.2 resume la estadística descriptiva fundamental para analizar la serie temporal a escala mensual, siendo esta serie asimétrica a la derecha en concordancia con la anual. Sin embargo, su coeficiente de asimetría (2,40) es superior al de la serie anual debido a que al promediar se suavizan los valores. Esto sucede de igual manera con el resto de los estadísticos de la tabla.

Tabla 3.2 Estadística del escurrimiento medio mensual para la subcuenca V Aniversario.

Estadística										
Mensual	Mín	Q ₁	Q ₂	Media	Q ₃	Máx.	S ²	S	Cv	CA
V Aniversario	0,05	0,60	1,82	3,29	4,61	27,78	16,10	4,01	1,22	2,40

Otra de las características significativas que arroja la Tabla 3.2, es que la serie mensual logra su valor mínimo de escurrimiento ($0,05 \text{ m}^3/\text{s}$) en marzo de 1971 y su valor máximo ($27,78 \text{ m}^3/\text{s}$) en junio de 1972 (Figura 3.2). Además, el 25 % los datos se encuentran por debajo de $0,60 \text{ m}^3/\text{s}$ y del 75% al 100 %, se localizan por encima de $4,61 \text{ m}^3/\text{s}$. Es de interés destacar, que esta serie tiene una media muy similar a la de la serie anual. Sin embargo, no sucede lo mismo con los estadísticos de mediana y desviación estándar y la varianza, en la serie mensual, es notablemente superior con respecto a la anual.

3.1.3 Serie de escurrimiento diaria

La serie del escurrimiento medio diario es mostrada en la Figura 3.4. Como puede apreciarse, al igual que en la escala mensual, esta serie no presenta una tendencia significativa.

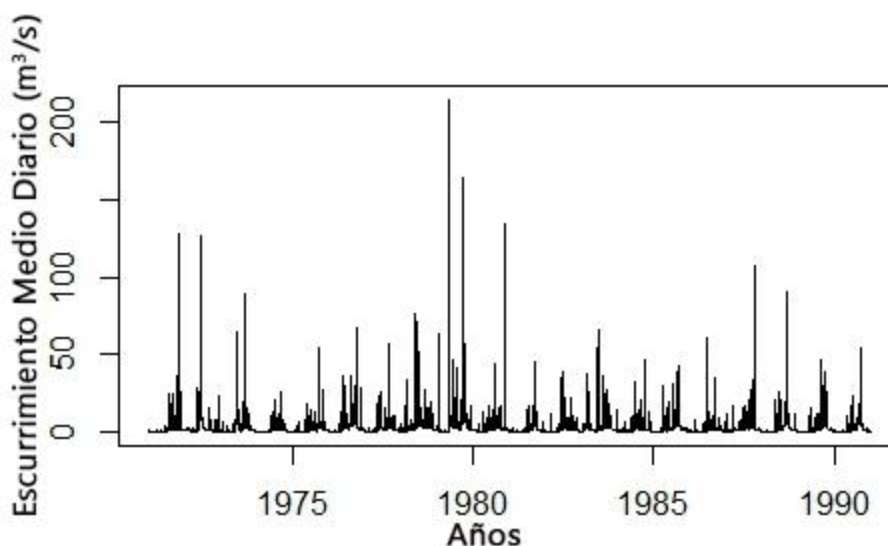


Figura 3.4 Serie de escurrimiento medio diario de la subcuenca V Aniversario (1971-1990).

En la Tabla 3.3 se recoge, de manera concisa, la información de la serie a escala diaria para poder realizar un correcto análisis acerca del comportamiento de la misma. Esta serie alcanza su valor de máximo escurrimiento ($215 \text{ m}^3/\text{s}$) el 24 de abril de 1979 y su valor mínimo ($0,004 \text{ m}^3/\text{s}$), consecutivamente, del 1^{ero} al 5 de febrero de 1987. Por otra parte, el 25 % y el 75% de los datos se encuentran por debajo de $0,42$ y $3,032 \text{ m}^3/\text{s}$, respectivamente. Es de interés destacar, que esta serie coincide prácticamente en la media con respecto a la serie anual y mensual, no cumpliéndose lo anterior en los otros valores estadísticos. Su desviación estándar ($8,07 \text{ m}^3/\text{s}$), por ejemplo, es el doble de la desviación estándar de la serie mensual y su varianza es significativamente

elevada en comparación con las dos primeras series analizadas, siendo este el resultado de promediar.

Tabla 3.3 Estadística del escurrimiento medio diario para la subcuenca V Aniversario.

<i>Estadística</i>										
Diaria	Mín	Q ₁	Q ₂	Media	Q ₃	Máx.	S ²	S	Cv	CA
V Aniversario	0,004	0,42	1,04	3,32	3,032	215	65,17	8,07	2,43	9,56

En la Figura 3.5a se muestra el desglose de la serie en sus principales componentes, siguiendo el mismo procedimiento de la escala mensual. Se aprecia que la estacionalidad encontrada en la serie mensual se mantiene y es aún más marcada cuando el análisis se realiza con valores diarios.

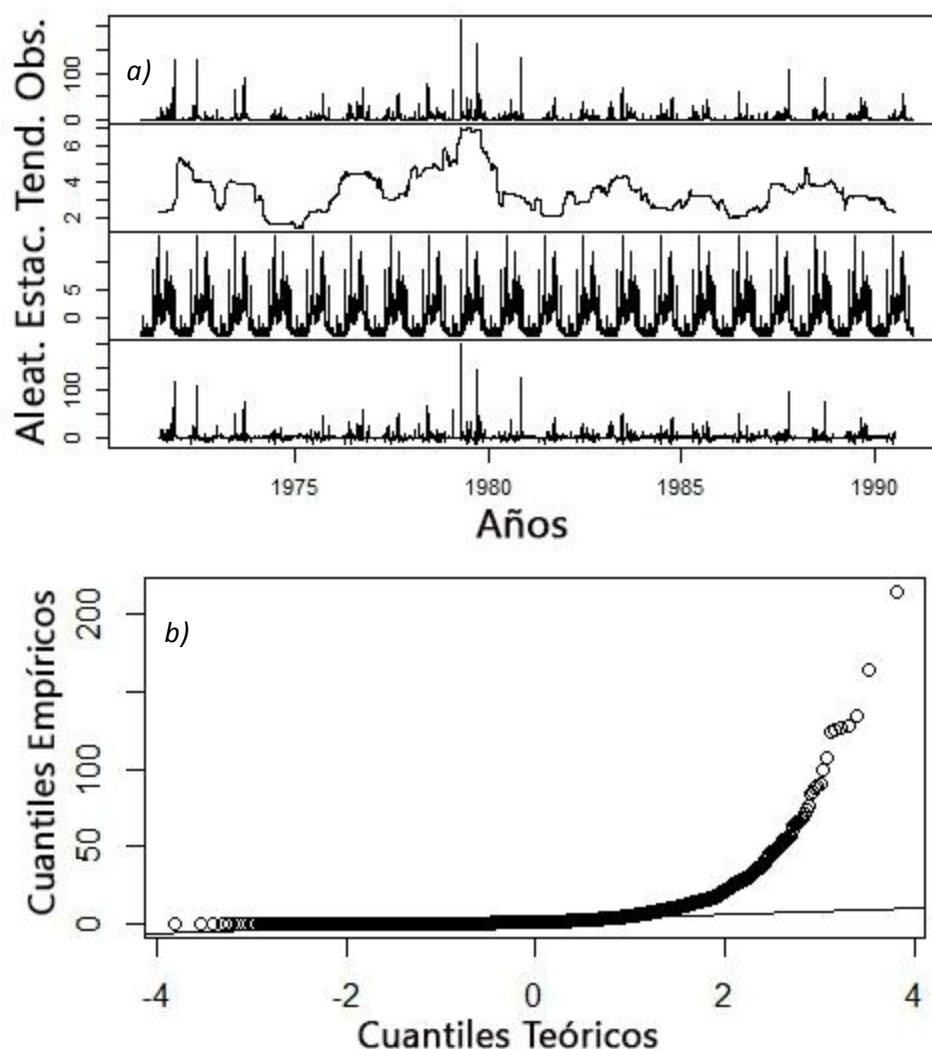


Figura 3.5 Serie de escurrimiento medio diario, a) componentes de la serie temporal y b) diagrama Q-Q.

En la Figura 3.5b, se observa el gráfico de normalidad donde, si bien la mayoría de los valores siguen una distribución aproximadamente normal, en el extremo de la cola derecha se aprecia un cambio en este patrón. Esto último se encuentra en correspondencia con el elevado valor del coeficiente de asimetría (9,56), reflejado en la Tabla 3.3, lo cual confirma que la serie diaria es asimétrica a la derecha de manera similar a como ocurre en las escalas anual y mensual.

3.2 Calibración de modelos

En este epígrafe, se identificarán los modelos que mejor representan el comportamiento de las series de tiempo en las diferentes escalas. Para lograr lo anterior, se hará uso fundamentalmente de las funciones de autocorrelación y del Criterio de Información Akaike (AIC). En el caso especial de la serie del escurrimiento medio mensual, se estimarán además los parámetros del modelo y se analizarán si los residuos cumplen con algunos de los supuestos establecidos.

3.2.1 Serie de escurrimiento anual

En la Figura 3.6, se muestra la serie del escurrimiento medio anual, donde también aparecen los resultados del test no paramétrico Mann-Kendall para analizar la significación de la tendencia en los valores ($pval = 0,87$).

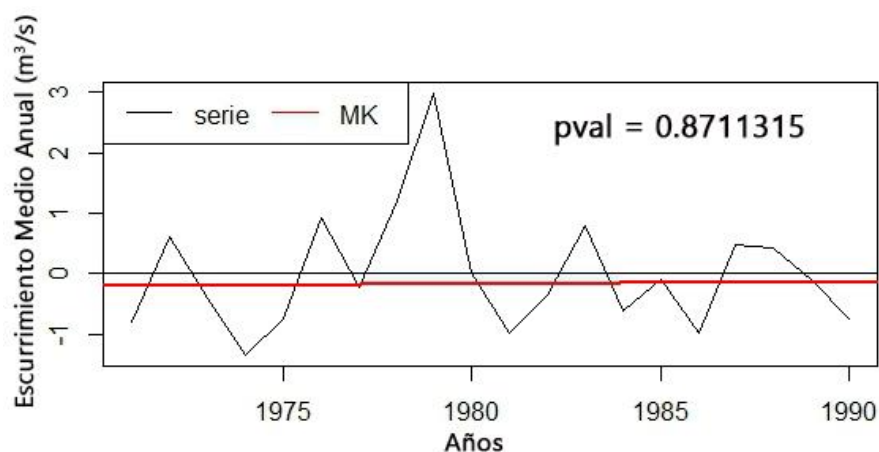


Figura 3.6 Serie de escurrimiento medio anual y test no paramétrico Mann-Kendall.

En la serie de tiempo graficada puede observarse que no existe una tendencia significativa. En el test de Mann-Kendall, la hipótesis nula (H_0) corresponde a que no haya tendencia, lo cual es corroborado por el resultado del $pval$ (mucho mayor que el umbral 0,05). En general, la serie puede ser considerada como estacionaria.

En la Figura 3.7, se muestran las funciones de autocorrelación (ACF, por sus siglas en inglés) y de autocorrelación parcial (PACF, por sus siglas en inglés) de la serie. La

función de autocorrelación es una herramienta estadística muy útil que mide si los primeros valores de la serie tienen alguna relación con los últimos. En general, el test para la autocorrelación debe ser hecho desde el retardo uno hasta el retardo $n/4$, donde n es el número total de observaciones. Por otra parte, la función de autocorrelación parcial permite conocer la cantidad de correlación entre una variable y su retardo correspondiente que no es explicada por las correlaciones en los retardos inferiores (Durdu, 2010).

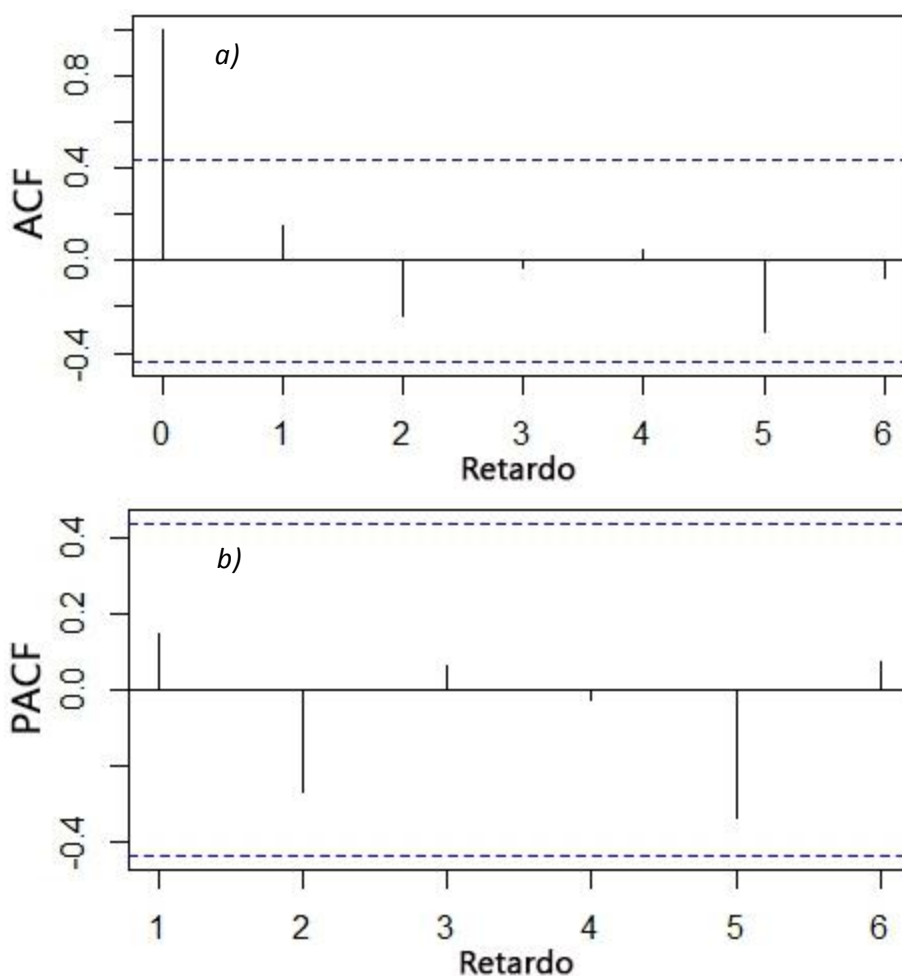


Figura 3.7 a) Función de autocorrelación (ACF) y b) función de autocorrelación parcial (PACF) para el escurrimiento medio anual de la subcuenca V Aniversario.

De acuerdo con lo anterior, se puede concluir a partir de estas gráficas (Figura 3.7), que esta serie se comporta como un Ruido Blanco puro, pues no existe una correlación temporal significativa entre los valores de escurrimiento medio anual, es decir, ninguno de los retardos posee un valor de correlación que exceda las líneas discontinuas (prueba de significación), con valores aproximadamente de $\pm 0,4$ (valor determinado en función de la cantidad de observaciones de la muestra). El valor de correlación 1 (Figura 3.7a), obtenido como resultado de comparar la serie con ella

misma, no ofrece información en este análisis temporal. La ausencia de valores significativos de correlación para otros órdenes indica que no existe una memoria temporal en el sistema que pueda ser explotada. Por tanto, el proceso, a esta escala, está explicado por el componente aleatorio.

El resultado arrojado por las gráficas anteriores (Figura 3.7), es corroborado mediante la Figura 3.8, donde se representa el Criterio de Información Akaike (AIC) para una estructura de modelo $AR(p)$ y $MA(q)$, respectivamente.

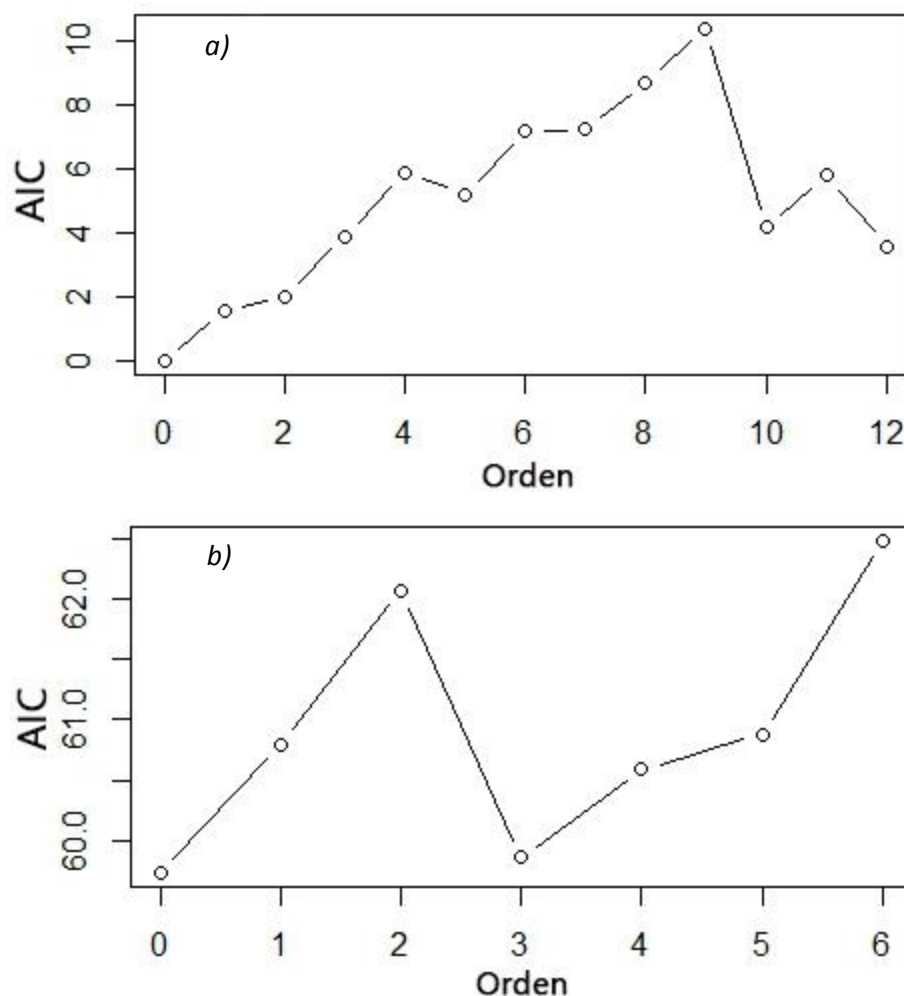


Figura 3.8 Criterio de Información Akaike (AIC), a) estructura $AR(p)$ y b) $MA(q)$.

En el caso del $AR(p)$ (Figura 3.8a), el único valor significativo se obtiene para el orden cero, por lo cual carece de sentido usar esta estructura de modelación. Sin embargo, para el $MA(q)$ en el orden tres (Figura 3.8b), pareciera haber indicio de una posible explicación de la variabilidad de la serie. Esta variabilidad puede estar asociada a la variabilidad interanual de las precipitaciones y por ende, del escurrimiento como respuesta de las mismas. Específicamente, los efectos del ENOS (El Niño Oscilación

Sur), fenómeno atmosférico que influye en las condiciones climáticas del Caribe, provocan ciclos de entre tres y cuatro años, como promedio. Los resultados en el AIC pueden estar reflejando este contexto.

Sin embargo, no existen elementos suficientes para concluir esto, la cantidad de valores analizados parece ser aún insuficiente para detectar una correlación en este fenómeno que es cíclico, pero errático. Por otra parte, Akaike penaliza a los modelos por la cantidad de parámetros, que en este caso es tres, y como puede observarse en la Figura 3.8b, el modelo MA(3) no logra mejorar el AIC para el MA(0) o Ruido Blanco. Además, los resultados de este criterio para el orden tres son corroborados mediante las gráficas de ACF y PACF, donde se demuestra que no existe prácticamente correlación para este orden.

Por tanto, se considera que el mejor modelo que representa el comportamiento de esta serie es el Ruido Blanco, como ya se había expresado anteriormente. Sin embargo, esto no está en correspondencia con lo obtenido por Avilés *et al.* (2016), donde se probaron varios modelos a escala anual y según el AIC, el mejor que se ajustó a su serie de datos fue el ARMA (1,1). Por otra parte, según Díaz y Guevara (2016), el modelo adecuado para describir las descargas medias anuales en las subcuencas del río Santa, Perú es el AR(1).

3.2.2 Serie de escurrimiento mensual

Atendiendo a la escasa memoria temporal encontrada en la serie de escurrimiento a escala anual, se decidió explorar la escala mensual, para lo cual se graficó primeramente el test de tendencia no paramétrico de Mann-Kendall (Figura 3.9).

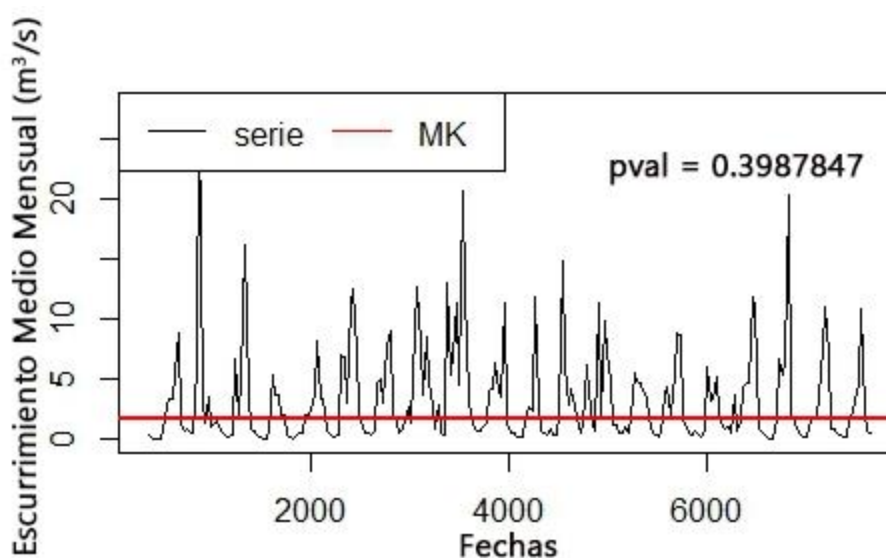


Figura 3.9 Serie de escurrimiento medio mensual y test de tendencia no paramétrico Mann-Kendall.

Una de las similitudes entre esta serie y la serie a escala anual, es que ambas poseen un p valor superior a 0.05, por lo que se puede comprobar que no existe una tendencia en la serie, en el período estudiado, de acuerdo con este test.

Por otra parte, la Figura 3.10 muestra las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF). En el eje x, los retardos representan fracciones de años, correspondiendo cada retardo de 0.5 a 6 meses. Como es de esperar, la memoria o persistencia es mucho mayor en períodos más cortos en las observaciones de escurrimiento, lo cual está en correspondencia con el estudio realizado por Valipour *et al.* (2013) a esta misma escala. Por tanto, la correlación lineal significativa, con diferentes intervalos de desfase temporal, es muy superior en las mediciones mensuales comparadas con las anuales.

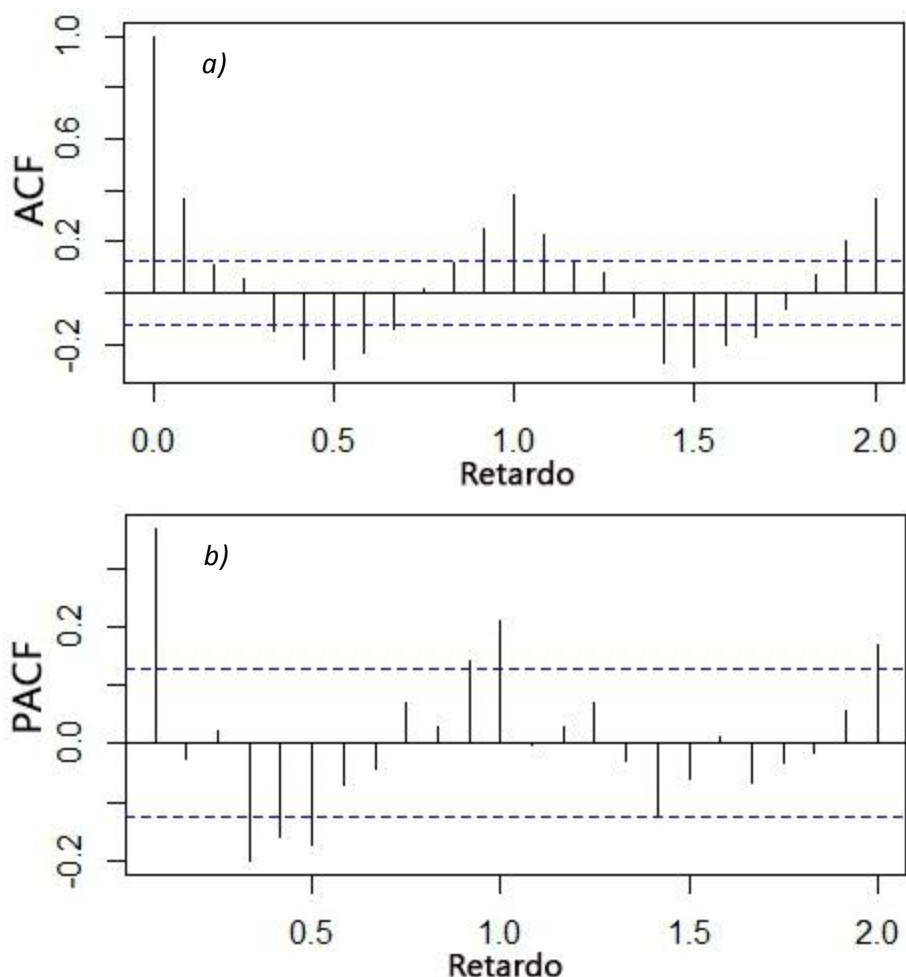


Figura 3.10 a) Función de autocorrelación (ACF). b) Función de autocorrelación parcial (PACF) para el escurrimiento medio mensual de la subcuenca V Aniversario.

La estacionalidad de la serie mensual puede ser fácilmente reconocida en estas gráficas. El coeficiente de correlación lineal, ya sea positivo o negativo, se alterna cada 6 meses, es por ello que un modelo adecuado para la misma sería un modelo

Estacional Autorregresivo Integrado de Media Móvil (SARIMA, por sus siglas en inglés).

Los modelos SARIMA pueden poseer hasta seis parámetros: p , d , q , P , D y Q , por lo que son más complejos con respecto a los ARMA debido a que incorporan, en su estructura, una flexibilización tal que permite modelar directamente la componente estacional y la tendencia en los valores. Por tanto, inicialmente, se debe intentar modelar la serie a partir de los procesos ARMA con valores de p y q que sean iguales o inferiores a 3 (Dettling, 2014), para garantizar una composición más sencilla del modelo.

Los resultados de la modelación ARMA serán comparados con los de SARIMA para encontrar el modelo que mejor represente el comportamiento de la serie de tiempo. La Figura 3.11 muestra el gráfico de AIC para diferentes combinaciones de p y q desde el orden 1 hasta el 3. Según este criterio, los modelos más adecuados son: el ARMA (1, 2), ARMA (2, 1) y el ARMA (3, 2), siendo este último el de menor AIC.

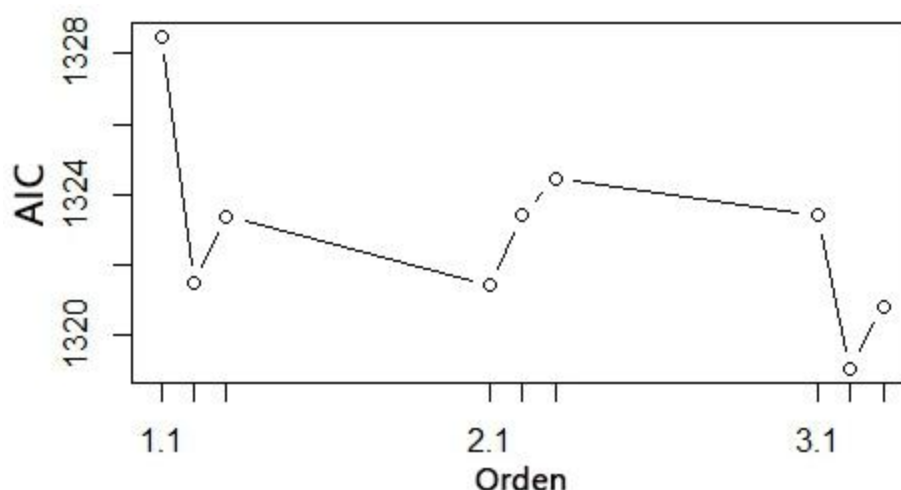


Figura 3.11 (AIC) para diferentes combinaciones de p y q de un modelo ARMA para simular el escurrimiento medio mensual.

Adicionalmente, en el software STATGRAPHICS Plus 5.0 se simulaban los modelos ARMA anteriormente mencionados y los resultados se muestran en la Tabla 3.4. En esta tabla se comparan los resultados en la predicción de estos tres modelos según la raíz del error cuadrático medio (RMSE, por sus siglas en inglés) y de los cinco tests corridos en los residuos: test para corridas excesivas arriba y abajo (RUNS), test para corridas excesivas por encima y debajo de la media (RUNM), test de Box-Pierce para la autocorrelación excesiva (AUTO), test para la diferencia en media desde la primera hasta la segunda mitad (MEAN) y el test para la diferencia en la varianza desde la primera hasta la segunda mitad (VAR).

Tabla 3.4 RMSE de los modelos y tests estadísticos en los residuos.

<i>Modelos</i>	<i>RMSE</i>	<i>RUNS</i>	<i>RUNM</i>	<i>AUTO</i>	<i>MEAN</i>	<i>VAR</i>
ARIMA (1,0,2)	3,82619	OK	***	***	OK	*
ARIMA (2,0,1)	3,81585	**	***	***	OK	*
ARIMA (3,0,2)	3,99538	*	***	***	OK	*

Cuando alguno de los tests anteriores tiene un OK, significa que el modelo pasa el test (no significativo: $p \geq 0.05$). Por otra parte, un asterisco (*) representa que el modelo falla en el 95% del nivel de confianza (marginamente significativo: $0.01 < p \leq 0.05$), dos asteriscos (**) en el 99% (significativo: $0.001 < p \leq 0.01$), y tres asteriscos (***) en el 99.9% (altamente significativo: $p \leq 0.001$).

En la Tabla 3.4 se puede apreciar que el mayor valor de RMSE se alcanza en el ARIMA (3,0,2) y el menor en el ARIMA (2,0,1), en ambos modelos se cumple solamente con el test RUNS. Por otro lado, el ARIMA (1,0,2), a pesar de poseer un valor de RMSE ínfimamente superior en comparación con el ARIMA (2,0,1), logra pasar el test MEAN y cumple también con RUNS por lo que, a consideración de la autora, este es el modelo más adecuado de los tres. Sin embargo, debido a que los tests RUNM, AUTO y VAR, en este último modelo, son estadísticamente significativos en un 95 o mayor porcentaje del nivel de confianza, hay que considerar seleccionar otro modelo.

3.2.2.1 Calibración del modelo SARIMA (p,d,q)(P,D,Q)

Atendiendo a las limitaciones de las estructuras ARMA, se decidió comprobar un modelo SARIMA (p,d,q)(P,D,Q), estructura que incluye la estacionalidad presente en la serie. Para llevar a cabo este proceso, primeramente, fue necesario truncar la serie del escurrimiento medio mensual en el año 1989 con el propósito de estimar, en el período comprendido entre los años 1971 y 1988, los parámetros del modelo. En los dos años restantes: 1989 y 1990, se validarán las capacidades predictivas del modelo, aspecto que será abordado en el subepígrafe 3.3.2.

En el software R (versión 3.0.2) se implementó una función que permite calcular de forma iterativa el AIC para las diferentes combinaciones de p, d, q, P, D y Q. Estos parámetros varían desde cero hasta su valor máximo que es configurable en la función. Por tanto, a partir de este procedimiento, se logra escoger el modelo que tiene el menor valor de AIC.

Luego de ejecutar la función implementada, tomando a 2 como el valor máximo de p, d, q, P, D y Q (Dettling, 2014), el mejor modelo que se obtuvo para representar el

comportamiento de la serie fue el SARIMA (0,1,1)(2,1,1), el cual posee un RMSE = 3.136384 que es inferior con respecto a los modelos mencionados con anterioridad. Esto se encuentra en concordancia con la investigación desarrollada por Modarres (2007), donde el modelo SARIMA (1,1,0)(0,1,1) fue seleccionado para pronosticar el escurrimiento mensual en el Río Zayandehrud, provincia Isfahan, Iran.

Sin embargo, autores como Avilés *et al.* (2016) utilizan otras estructuras de modelación en la escala mensual. En el estudio que estos autores realizaron en la cuenca del río Tomebamba, en Ecuador, los mejores modelos escogidos fueron: el autorregresivo de media móvil periódico PARMA (1,0) para los meses de enero, febrero, marzo, mayo, junio, septiembre y noviembre, el PARMA (1,1) para los meses de abril, julio, agosto y diciembre y el modelo PARMA (1,2) para el mes de octubre.

En la Tabla 3.1 se resumen los coeficientes del modelo SARIMA (0,1,1)(2,1,1) y los p-valores, para lo cual fue necesario implementar otra función en el software que permite obtener los resultados mostrados. Como puede apreciarse, todos los parámetros son estadísticamente significativos, de acuerdo con lo planteado por (Modarres, 2007; Fernández *et al.*, 2011), pues en todos los casos el p-valor es inferior a 0.05.

Tabla 3.5 Parámetros del modelo SARIMA (0,1,1)(2,1,1) calibrado para el escurrimiento medio mensual de la subcuenca V Aniversario.

Parámetros del modelo	Valor del parámetro	Error Estándar	t-valor	p-valor
ϕ_1	-0,9056	0,0343	-26,44	p < 0,0001
θ_1	-0,1410	0,0553	-2,55	0,011
θ_2	-0,1681	0,0615	-2,73	0,006
Φ_1	-0,8262	0,0461	-17,94	p < 0,0001

ϕ_1 término MA de primer orden, θ_1 término AR estacional de primer orden, θ_2 AR estacional de segundo orden, Φ_1 término MA estacional de primer orden.

La estructura definitiva del modelo con sus coeficientes, según el desglose de la Ec. 2.10 en cada uno de sus polinomios (Abudu *et al.*, 2010), queda de la siguiente manera:

$$\theta_p(B^s) = 1 + 0,1410B^{12} + 0,1681B^{24} \quad \text{Ec. 3.1}$$

$$\theta_p(B) = 1 \quad \text{Ec. 3.2}$$

$$\Phi_Q(B^s) = 1 + 0,8262B^{12} \quad \text{Ec. 3.3}$$

$$\phi_q(B) = 1 + 0,9056B \quad \text{Ec. 3.4}$$

y la ecuación general para el modelo SARIMA (0,1,1)(2,1,1) es:

$$(1 + 0,1410B^{12} + 0,1681B^{24})(1 - B^{12})^1(1 - B)^1x_t = (1 + 0,8262B^{12})(1 + 0,9056B)w_t \quad \text{Ec. 3.5}$$

donde: x_t es la variable dependiente, B el operador de retroceso hacia atrás y w_t es el Ruido Blanco o el error residual del modelo. Despejando x_t se obtiene:

$$\begin{aligned} x_t = & x_{t-1} + 0,859(x_{t-12} - x_{t-13}) - 0,0271(x_{t-24} - x_{t-25}) \\ & + 0,1681(x_{t-36} - x_{t-37}) + w_t + 0,9056(w_{t-1} + w_{t-12}) \\ & + 0,7482w_{t-13} \end{aligned} \quad \text{Ec. 3.6}$$

Una vez que el modelo apropiado es escogido y sus parámetros son estimados, se necesita examinar los residuos del modelo para verificar que este es adecuado para la serie de tiempo (Durdu, 2010). El gráfico de los residuos entre los valores observados y los simulados por el modelo de la Ec. 3.6, se muestra en la Figura 3.12. Estos cumplen con los supuestos establecidos de media cero y varianza constante.

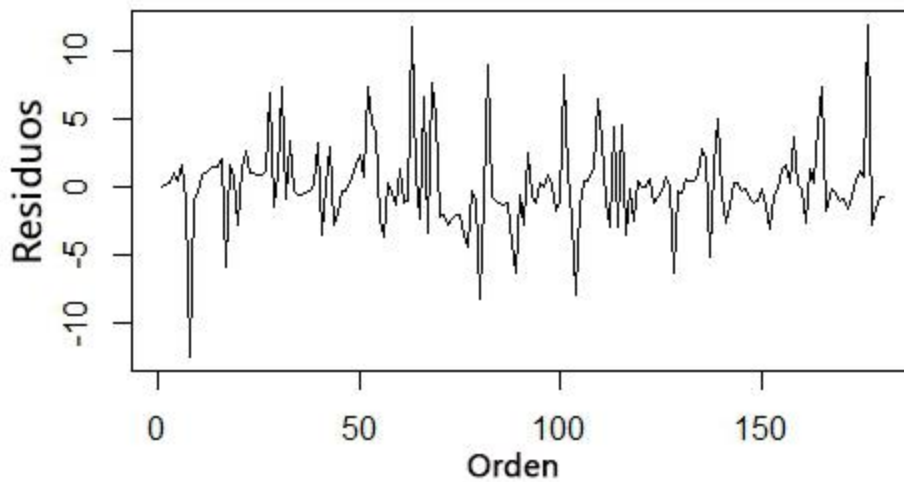


Figura 3.12 Gráfico de los residuos para el modelo de la Ec. 3.6.

Por otra parte, la Figura 3.13 muestra el gráfico del histograma y el diagrama Q-Q de los residuos del SARIMA (Ec. 3.6). A pesar de que en los extremos los residuos sobresalen de la campana, en rojo (Figura 3.13a), y en las colas de la Figura 3.13b se alejan de la línea normal, se evidencia que la mayoría de los valores se distribuyen normalmente.

En el estudio realizado por Farajzadeh *et al.* (2014), en el que se simuló el escurrimiento mensual en la cuenca del lago Urmia a partir de redes neuronales y de los modelos lineales estocásticos, se determinó que la estructura SARIMA (2,1,2)(4,0,0) mostraba mejor desempeño. Al realizarse el análisis de los residuos del modelo seleccionado, se encontró que el histograma tenía la forma de la campana normal y que en el diagrama Q-Q, los residuos se ajustaban a la línea normal por lo

que, en concordancia con la presente investigación, estos se distribuyeron normalmente.

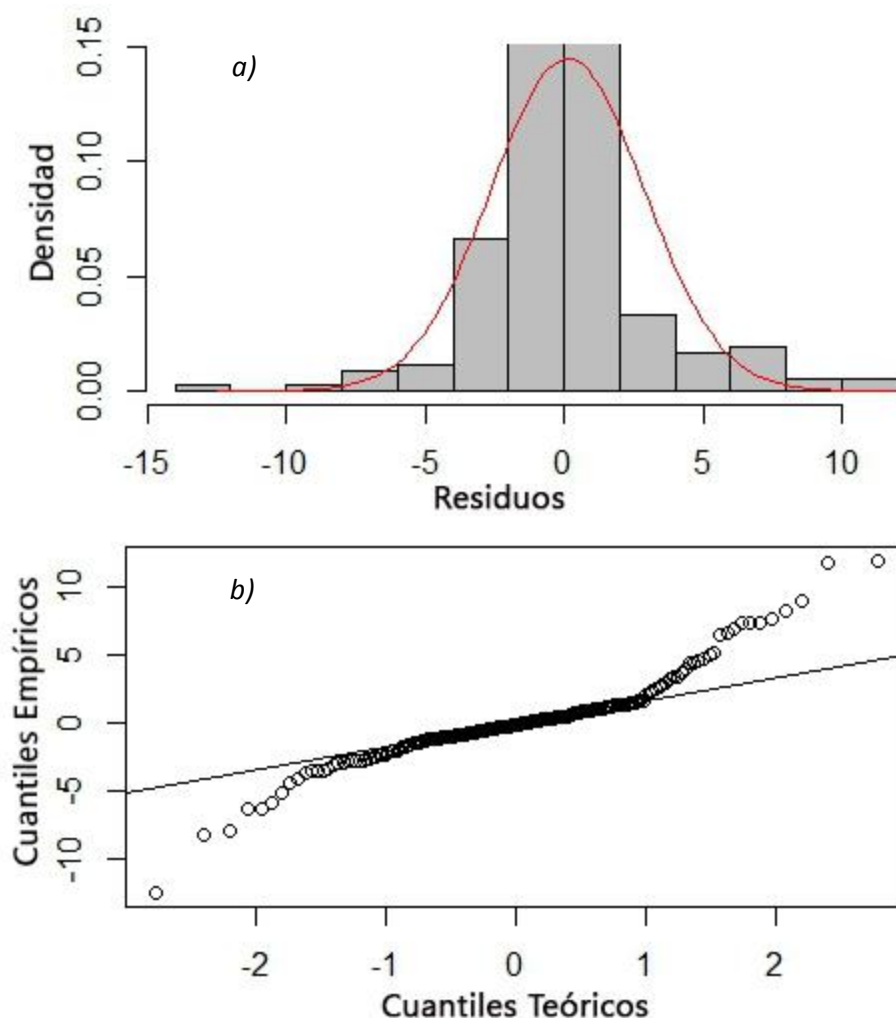


Figura 3.13 a) Histograma. b) Diagrama Q-Q de los residuos del SARIMA (Ec. 3.6).

Las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF) de los residuos son mostradas en la Figura 3.14. De las mismas se concluye que los residuos presentan un comportamiento aleatorio en torno a cero y que estos no se encuentran correlacionados, por lo que su patrón es típico de un Ruido Blanco. Lo anterior se encuentra en correspondencia con lo planteado por autores como Durdu (2010) y (Fernández *et al.*, 2011), estos últimos hacen especial énfasis en que ninguno de los picos deben exceder las bandas de confianza.

En la investigación realizada por Farajzadeh *et al.* (2014), mencionada con anterioridad, se arribó a que los errores o residuos del modelo SARIMA que los autores seleccionaron para representar el escurrimiento mensual en el lago Urmia, se ajustaban perfectamente a un Ruido Blanco. Esta conclusión fue alcanzada a partir del

análisis de las correlaciones de los errores de su modelo, las cuales no excedieron las bandas de confianza en el gráfico del ACF, similarmente como ocurrió en el presente estudio.

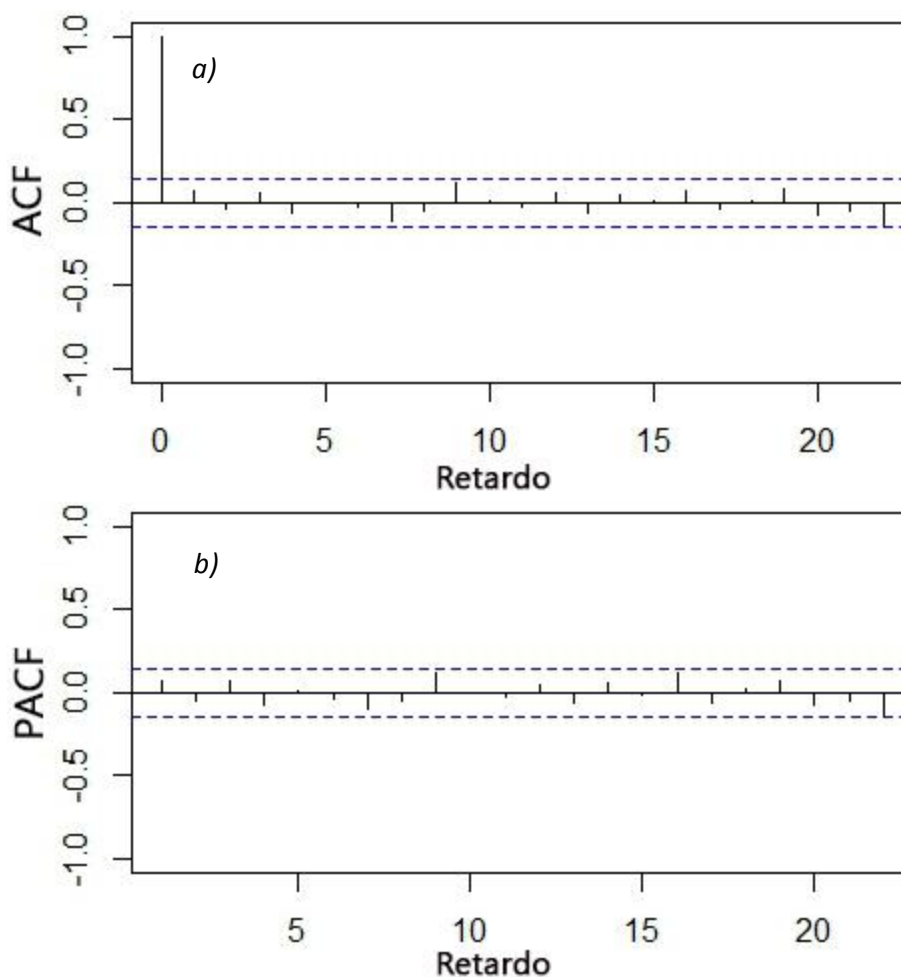


Figura 3.14 a) Función de autocorrelación (ACF). b) Función de autocorrelación parcial (PACF) de los residuos del modelo SARIMA (Ec. 3.6).

Para conocer el desempeño estadístico de un modelo se utilizan dos índices estadísticos fundamentales: el coeficiente de determinación (R^2) y la eficiencia del modelo (E). La eficiencia del modelo es un criterio de evaluación del poder predictivo de los modelos hidrológicos, propuesto por Nash y Sutcliffe en 1970 (Abudu *et al.*, 2010). Una eficiencia del 80% o por encima, indica un desempeño excelente del modelo. Un valor en los rangos del 60 al 80% es muy bueno, del 40 al 60% es bueno y del 20 al 40% es un desempeño satisfactorio. Sin embargo, un valor por debajo del 20% representa un ajuste insuficiente (Burlando, 2014).

Para este modelo en particular, tanto la eficiencia como el R^2 son iguales a 0.51, lo que indica que este se ajusta adecuadamente a la serie temporal. Sin embargo, el desempeño de este modelo puede mejorar si se introducen nuevas variables como la

lluvia (Pektaş y Cigizoglu, 2013; Kwon, 2002). Por otro lado, es recomendable probar, en investigaciones posteriores, modelos multivariados que expliquen de manera más acertada el comportamiento del escurrimiento observado.

3.2.3 Serie de escurrimiento diaria

En la Figura 3.15, se muestra la serie del escurrimiento medio diario, donde también aparecen los resultados del test de tendencia no paramétrico Mann-Kendall. Se puede comprobar, a partir del resultado del p-valor (mayor que 0,05), que en esta serie no existe una tendencia significativa, igualmente como sucedía a escalas anual y mensual.

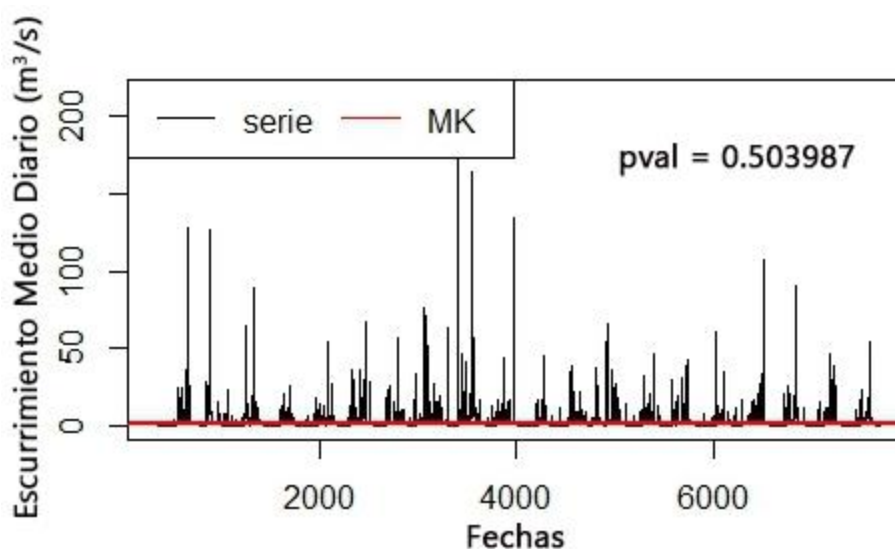


Figura 3.15 Serie de escurrimiento medio diario y test de tendencia no paramétrico Mann-Kendall.

Por otra parte, la Figura 3.16 muestra las funciones de autocorrelación (ACF) y autocorrelación parcial (PACF). Como es de esperar, la memoria o persistencia es mucho mayor en las observaciones del escurrimiento diario en comparación con las mensuales. Por tanto, la correlación lineal significativa, con diferentes pasos en el tiempo hacia atrás, es muy superior en las mediciones diarias comparadas con las mensuales y anuales. Sin embargo, la estacionalidad descubierta en la serie mensual puede ser fácilmente reconocida en estas gráficas a escala diaria, manteniéndose que el coeficiente de correlación lineal, ya sea positivo o negativo, se alterna cada 6 meses.

Por tanto, un modelo adecuado para simular el comportamiento de esta serie sería el modelo Fraccional Autorregresivo Integrado de Media Móvil (FARIMA, por sus siglas en inglés). Este modelo fue introducido por Granger y Joyeux (1980) y Hosking (1981) y su composición es muy similar al ARIMA (p,d,q) dado que sus parámetros son los

mismos. Sin embargo, al permitir que el orden de diferenciación d sea fraccional, el modelo exhibe larga memoria o dependencia en un intervalo largo. Autores como Montanari *et al.* (1997) y Grimaldi (2004) confirman en sus investigaciones que estos modelos son necesarios para este tipo de serie.

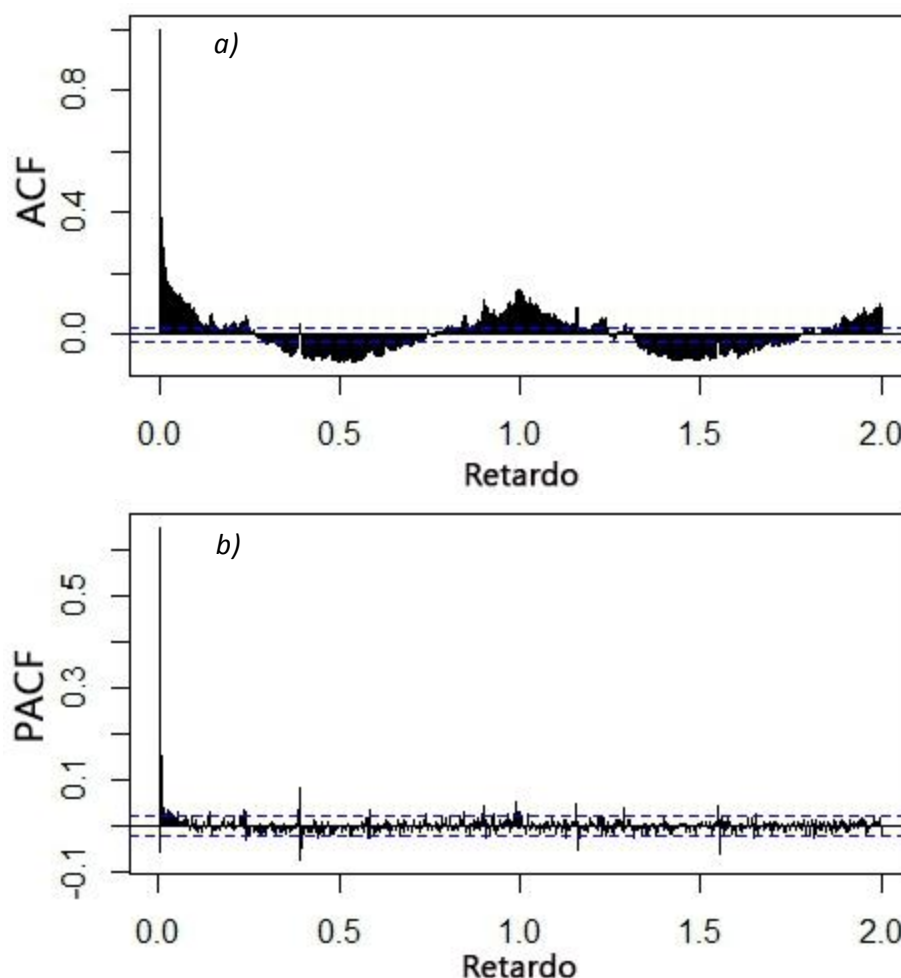


Figura 3.16 a) Función de autocorrelación (ACF). b) Función de autocorrelación parcial (PACF) para el escurrimiento medio diario de la subcuenca V Aniversario.

3.3 Validación de los modelos

En este último epígrafe se validan los modelos seleccionados en las escalas anual y mensual, que resultaron del epígrafe anterior. Este procedimiento se realiza ya sea a partir de la comparación de los principales estadísticos: media, varianza y desviación estándar de la serie observada y de las estimaciones del modelo o a partir de diferentes criterios en las predicciones. El modelo lineal estocástico para la serie de escurrimiento medio diario no fue ajustado en esta investigación, al igual que la relación multivariada con variables como la lluvia, por lo que se recomienda su implementación en investigaciones posteriores.

3.3.1 Serie de escurrimiento anual

Para poder validar el ajuste del modelo de Ruido Blanco a la serie de escurrimiento anual, fue necesario generar 100 series (Alonso, 2018). La Tabla 3.6 muestra la estadística del escurrimiento observado y de las generaciones del Ruido Blanco. En este caso, el modelo reproduce muy bien las estadísticas de las mediciones, coincidiendo prácticamente en media, varianza y desviación estándar (S).

La Figura 3.17, conocida como gráfico “espagueti”, representa las 100 series generadas por el modelo de Ruido Blanco, donde en rojo aparece la serie de datos observada. Por tanto, pueden compararse visualmente las realizaciones sintéticas y la observada, las cuales provienen del mismo proceso.

Tabla 3.6 Estadística de la serie de escurrimiento medio anual observado en la subcuenca V Aniversario y de las realizaciones del Ruido Blanco (WN).

	Media					S				
	Media	S ²	S	Cv	CA	Media	S ²	S	Cv	CA
Datos	3,26	1,55	1,24	0,38	1,41	-	-	-	-	-
WN	3,27	1,50	1,21	2,77	-0,04	0,28	0,45	0,18	0,55	0,46

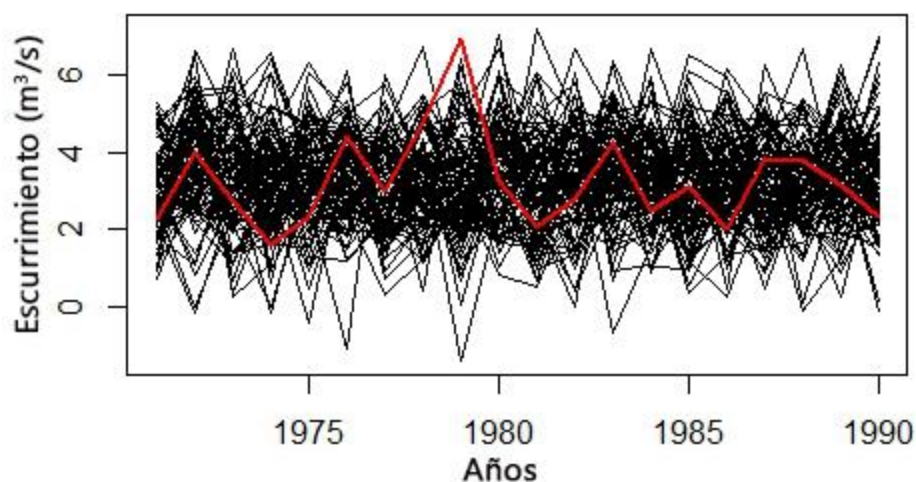


Figura 3.17 Gráfica de las series de escurrimiento medio anual observada y generadas por el Ruido Blanco.

3.3.2 Serie de escurrimiento mensual

Para evaluar las potencialidades del SARIMA (0,1,1)(2,1,1) en cuanto a la predicción, se decidió truncar la serie de tiempo mensual, contemplada en el período 1971-1990, en el año 1989, como se explicó en el subepígrafe 3.2.2.1. Por ende, la validación del modelo para los dos últimos años de la serie: 1989 y 1990 se muestra en la Figura 3.18, donde las líneas discontinuas representan la serie de datos observada y las

líneas continuas en rojo: la predicción del modelo. Se arriba a la conclusión, a partir de esta gráfica, de que el modelo logra replicar los dos picos anuales y también los valores mínimos de la serie. Por tanto, se considera que esta estructura da como resultado un pronóstico acertado del escurrimiento a esta escala.

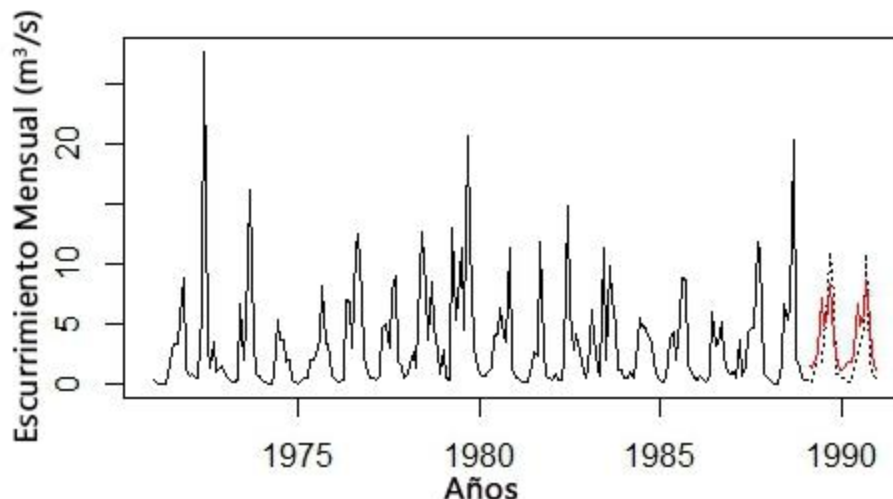


Figura 3.18 Serie de escurrimiento observada (líneas discontinuas) y simulada (rojo) con el modelo SARIMA para el período 1989-1990.

El modelo de Holt-Winters es un método popular de predicción, pues desde que surgió aún se usa en las investigaciones (Da Veiga *et al.*, 2014). Este fue sugerido por Holt en 1957 y Winters en 1960 (Cowpertwait y Metcalfe, 2009). Este método no posee una ecuación general, pero sí un conjunto de ecuaciones que utilizan condiciones iniciales o parámetros de suavizamiento. Este se basa en las medias móviles ponderadas exponencialmente para suavizar la tendencia, estacionalidad y otras componentes de una predicción (Da Veiga *et al.*, 2014). Por tanto, puede adaptarse fácilmente a cambios y tendencias, así como a patrones estacionales, y el tiempo necesario para calcular el pronóstico es considerablemente rápido.

Sin embargo, a pesar de que el método de Holt-Winters es relativamente sencillo, los modelos de la familia ARIMA proporcionan predicciones sin necesidad de la existencia de ningún tipo de condición previa (Guerrero *et al.*, 2006), por lo que en este sentido estos son más ventajosos con respecto al Holt-Winters.

En la Figura 3.19, puede apreciarse cómo el método de Holt-Winters (líneas continuas en rojo), a medida que pasa el tiempo, se acerca a la serie de datos observada (líneas continuas en negro), a partir del cálculo de los pesos o variaciones en la media de la serie, para poder finalmente hacer la predicción (líneas discontinuas en azul) en los años 1989 y 1990.

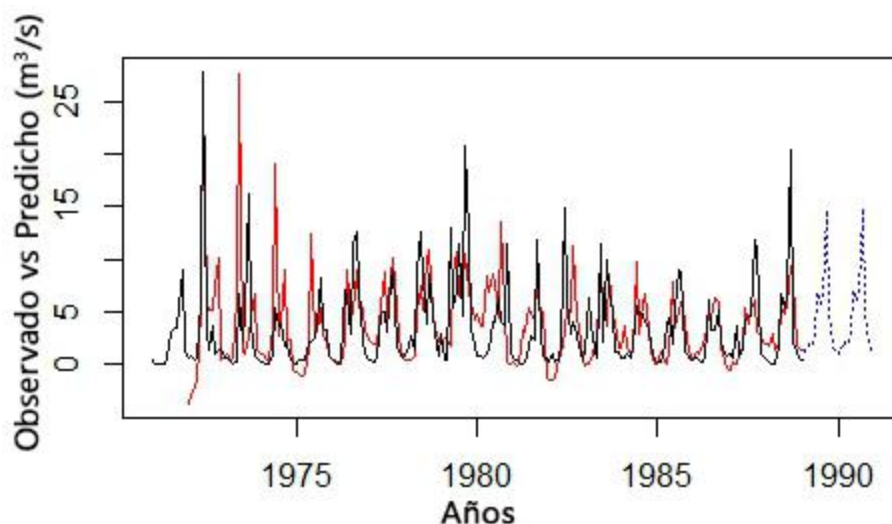


Figura 3.19 Predicción del Método Holt-Winters en la serie de escurrimiento mensual.

Para establecer una comparación visual entre los dos métodos de predicción (Holt-Winters y SARIMA), se graficó la serie de datos observada (líneas y puntos en negro) con las predicciones de los modelos: Holt-Winters (líneas y triángulos en azul) y SARIMA (0,1,1)(2,1,1) (líneas y cuadrados en rojo) como se muestra en la Figura 3.20.

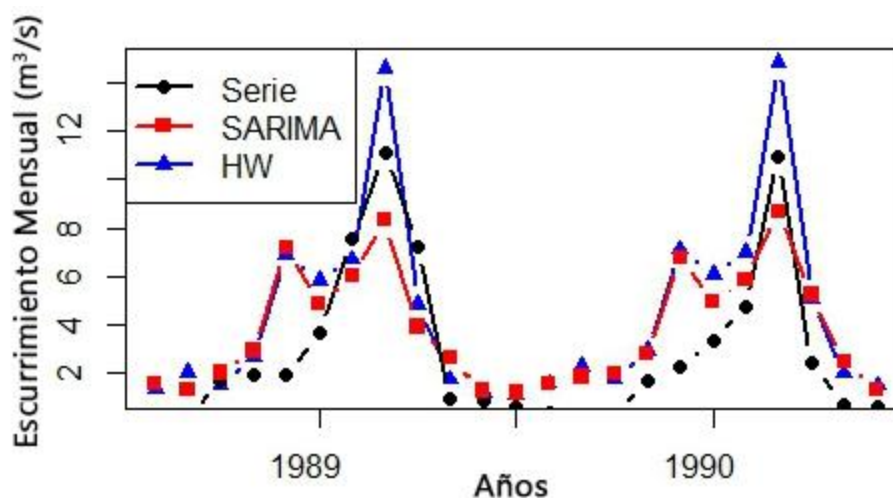


Figura 3.20 Predicción de la serie de datos observada por los modelos: Holt-Winters y SARIMA (0,1,1)(2,1,1).

Como puede apreciarse en esta figura, el modelo SARIMA predice mejor lo que ocurrió en los dos últimos años (1989 y 1990), pues los picos anuales de esta estructura son más cercanos a los de la serie. En este sentido, el valor máximo en el año 1989 es: 11.1, 8.33 y 14.60 m^3/s para la serie observada, el valor predicho por el SARIMA y por el Holt-Winters, respectivamente, evidenciando mayor semejanza el SARIMA. Igualmente ocurre en el año 1990, donde los valores son: 10.94, 8.68 y 14.85 m^3/s , respetando el orden anterior. Por tanto, los picos anuales de la predicción

del Holt-Winters, se alejan de los valores observados, sobreestimado lo que ocurrió en este período.

Sin embargo, la representación gráfica en la Figura 3.20 no es suficiente para evaluar la capacidad predictiva de ambos métodos. Autores como Guerrero *et al.* (2006) y Makatjane y Moroke (2016) utilizan cuatro estadísticos para cuantificar globalmente los errores de predicción: la raíz del error cuadrático medio (RMSE, por sus siglas en inglés), el error absoluto medio porcentual o relativo (MAPE, por sus siglas en inglés), el error absoluto medio (MAE, por sus siglas en inglés) y el error cuadrático medio (MSE, por sus siglas en inglés). Estos valores se encuentran resumidos en la Tabla 3.7.

Tabla 3.7 Comparación en los errores de la predicción de los modelos Holt-Winters y SARIMA

<i>Criterios</i>	<i>Modelo Holt-Winters</i>	<i>Modelo SARIMA</i>
RMSE	6,166525	5,351782
MAPE	5,492751	5,179282
MAE	3,958702	3,319915
MSE	38,02603	28,64157

El error cuadrático medio (MSE) calcula el promedio de los residuos al cuadrado, es decir, mide las diferencias entre el estimador y lo que se estima. Como puede apreciarse en la Tabla 3.7, el MSE en el modelo Holt-Winters excede en un 25% al del SARIMA, debido a que los valores predichos por este primer método en los años 1989 y 1990, se alejan más de los valores observados, como fue expresado con anterioridad. Los valores del RMSE son consecuentes con este contexto, pues estos se obtienen a partir de la raíz del MSE.

Por otro lado, con respecto al error absoluto medio (MAE), el cual calcula el promedio de la diferencia absoluta entre los valores predichos y observados, se obtiene que en el modelo SARIMA este parámetro es inferior comparado con el método Holt-Winters. De igual manera se comporta el error absoluto medio porcentual (MAPE) que expresa la media del porciento de los residuos en las observaciones.

Debido a que en la Tabla 3.7, los valores de todos los errores en la predicción son inferiores en el modelo SARIMA en comparación con el Holt-Winters, se arriba a la conclusión de que el SARIMA (0,1,1)(2,1,1) es el modelo más adecuado para realizar la predicción de la serie del escurrimiento medio a escala mensual. Estos resultados son muy importantes debido a que la integración de este último modelo en un sistema de alerta temprana contra inundaciones, ayudaría a preservar la vida humana y los

bienes materiales y económicos en las regiones cercanas a la subcuenca V Aniversario, ante la posibilidad de ocurrencia de inundaciones.

CONCLUSIONES

1. El fenómeno del escurrimiento en la subcuenca V Aniversario tiene un valor medio que oscila entre 3,26 y 3,32 m³/s en las diferentes escalas y en todas estas es reconocida la asimetría a la derecha, patrón típico de las series hidrológicas. Cuando la escala es mensual o diaria, la presencia de la estacionalidad se hace visible.
2. La serie de escurrimiento medio anual de la subcuenca V Aniversario no presentó buena memoria temporal y está explicada, en su mayor parte, por una elevada aleatoriedad. El modelo de Ruido Blanco es el que mejor se ajusta a esta serie, considerando las semejanzas encontradas en la estadística de ambos, coincidiendo en media, varianza y desviación estándar.
3. En la serie de escurrimiento mensual, el coeficiente de correlación lineal es tanto positivo como negativo y se alterna cada 6 meses. Esta serie fue modelada mediante un modelo estacional autorregresivo integrado de media móvil: SARIMA (0,1,1)(2,1,1).
4. El modelo SARIMA es la estructura más adecuada para realizar la predicción de la serie de escurrimiento medio mensual en los años 1989 y 1990, por ser sus errores inferiores a los obtenidos con el método Holt-Winters, en este mismo período.

RECOMENDACIONES

1. Extender la serie de escurrimiento medio diario desde el año 1991 hasta la actualidad, 2019.
2. Modelar la serie de escurrimiento mensual a partir de la relación multivariada con la precipitación.
3. Ajustar la serie de escurrimiento medio diario mediante un modelo lineal estocástico, en investigaciones posteriores, y utilizar como una primera aproximación el modelo Fraccional Autorregresivo Integrado de Media Móvil.
4. Divulgar los resultados de este estudio con los campesinos y las instituciones de toma de decisiones de la región del Cuyaguateje.

REFERENCIAS BIBLIOGRÁFICAS

ABUDU, S.: ; CUI, C.: ; KING, J.P.: ; ABUDUKADEER, K.: "Comparison of performance of statistical models in forecasting monthly streamflow of Kizil River, China", *Water Science and Engineering*, 3(3): 269–281, 2010.

ALONSO, G.R.: *Estimación del riesgo de erosión hídrica en la subcuenca V Aniversario del río Cuyaguatzeje*, 102pp., Tesis (en opción al Master en Ciencias del Suelo), Facultad de Mecanización Agropecuaria, Universidad Agraria de La Habana (UNAH), La Habana, Cuba, 2008.

ALONSO, G.R.: *Modeling discharge and sediment yield for extreme events in western Cuba*, 81pp., Tesis (en opción al Master of Advanced Studies in Sustainable Water Resources (MAS-SWR)), Swiss Federal Institute of Technology Zurich (ETH), Zurich, Switzerland, 2015.

ALONSO, G.R.: "Predicción probabilística del escurrimiento superficial y la pérdida de sedimento para eventos extremos. Parte II", *Revista Ciencias Técnicas Agropecuarias*, 25(4): 4–16, 2016.

ALONSO, G.R.: *Predicción probabilística del escurrimiento y la pérdida de sedimentos para eventos extremos de precipitación en la subcuenca V Aniversario, río Cuyaguatzeje*, 120pp., Tesis (en opción al grado científico de Doctor en Ciencias Técnicas Agropecuarias), Departamento Matemática-Física, Facultad de Ciencias Técnicas, Universidad Agraria de La Habana (UNAH), Mayabeque, 2018.

AVILÉS, A.: ; SOLERA, A.: ; PAREDES, J.: "Análisis del rendimiento de sistemas hídricos en desarrollo mediante el acople de modelos estocásticos hidrológicos y optimización de redes de flujo", *Ingenius*, (15): 48–57, 2016.

AZAM, M.: ; SAN KIM, H.: ; MAENG, S.J.: "Development of flood alert application in Mushim stream watershed Korea", *International journal of disaster risk reduction*, 21: 11–26, 2017.

BALLESTER, M.: ; RUBIERA, J.: *Temporada ciclónica de 2009 en el Atlántico Norte*, [en línea] Instituto de Meteorología (INSMET), 2008. Disponible en: <http://www.insmet.cu>

BURLANDO, P.: *Water Resources Management*, Institute of Environmental Engineering, Swiss Federal Institute of Technology Zurich, 2014.

BURTON, A.: ; KILSBY, C.G.: ; FOWLER, H.J.: ; COWPERTWAIT, P.S.P.: ; O'CONNELL, P.E.: "RainSim: A spatial-temporal stochastic rainfall modelling system", *Environmental Modelling & Software*, 23(12): 1356–1369, 2008.

BUSTAMANTE, R.: *Econometría de series de tiempo*, 2014,

CHACÓN, J.: ; ZARATE, E.: *Inundaciones y desastres naturales producidos por tormentas tropicales y huracanes.*, Conferencia OMM/BID, Evaluación y estrategias de gestión de recursos hídricos en América Latina y el Caribe, Ginebra, Suiza, 1996.

CHOW, V.T.: ; MAIDMENT, D.R.: ; MAYS, L.W.: *Applied Hydrology*, Ed. MacGraw-Hill, pp. 572, ISBN-0-07-010810-2, 1988.

CITMA: *Informe de Rendición de Cuenta del CTCH de Pinar del Río ante el CNCH*, Ministerio de Ciencia, Tecnología y Medioambiente, 22pp., Pinar del Río, Cuba, 2005.

CONTIN, G.: *Las Series Temporales*, 2001,

COWPERTWAIT, P.S.P.: ; METCALFE, A.V.: *Introductory time series with R*, 2009.

CRYER, J.D.: ; CHAN, K.-S.: *Time series analysis with applications in R*, Ed. Springer, Second Edition ed. , pp. 491, ISBN-978-0-387-75958-6, 2008.

CTCH: *Catálogo de Cuencas Hidrográficas Río Cuyaguaje*, Consejo Territorial de Cuencas Hidrográficas, Pinar del Río, Cuba, 2000.

DA VEIGA, C.P.: ; DA VEIGA, C.R.P.: ; CATAPAN, A.: ; TORTATO, U.: ; DA SILVA, W.V.: "Demand forecasting in food retail: A comparison between the Holt-Winters and ARIMA models", *WSEAS transactions on business and economics*, 11(1): 608–614, 2014.

D'AMICO, G.: ; PETRONI, F.: ; PRATTICO, F.: "Economic performance indicators of wind energy based on wind speed stochastic modeling", *Applied energy*, 154: 290–297, 2015.

DEFENSA CIVIL: "Guía para la realización de estudios de riesgo para situaciones de desastres", *Ciudad de la Habana*, extraído de: <http://www.sld.cu/galerias/pdf/sitios/desastres/gu%C3%A1%20para%20la%20realizaci%C3%B3n%20de%20estudios%20de%20riesgo%20para%20situaciones%20de%20desastres.pdf>, en enero 2013, , 2005.

DETLING, M.: *Applied Time Series Analysis*, Institute for Data Analysis and Process Design, Zurich University of Applied Sciences, Zurich, Switzerland, 2014.

DÍAZ, A.M.: ; GUEVARA, E.: "Modelación estocástica de los caudales medios anuales en la cuenca del río Santa, Perú", *Revista INGENIERÍA UC*, 23(2), 2016.

DURDU, Ö.F.: "Application of linear stochastic models for drought forecasting in the Büyük Menderes river basin, western Turkey", *Stochastic Environmental Research and Risk Assessment*, 24(8): 1145–1162, 2010.

ESTRADA, V.: ; PACHECO, R.M.: "Modelación hidrológica con HEC-HMS en cuencas montañosas de la región oriental de Cuba", *Ingeniería Hidráulica y ambiental*, 33(1): 71–80, 2012.

FARAJZADEH, J.: ; FARD, A.F.: ; LOTFI, S.: "Modeling of monthly rainfall and run off of Urmia lake basin using "feed-forward neural network" and "time series analysis" model", *Water Resources and Industry*, 7: 38–48, 2014.

FERNÁNDEZ, A.: ; QUINTANO, C.: ; FERNÁNDEZ, O.: "Forecast of NDVI in coniferous areas using temporal ARIMA analysis and climatic data at a regional scale", *International journal of remote sensing*, 32(6): 1595–1617, 2011.

FRY, L.M.: ; HUNTER, T.S.: ; PHANIKUMAR, M.S.: ; FORTIN, V.: ; GRONEWOLD, A.D.: "Identifying streamgage networks for maximizing the effectiveness of regional water balance modeling", *Water Resources Research*, 49(5): 2689–2700, 2013.

FURRAZOLA, G.: ; JUDOLEY, C.M.: ; MIJAILOVSKAYA, M.S.: ; MIROLIUBOV, Y.S.: ; NOVOJATSKY, I.P.: ; JIMÉNEZ, A.: ; SOLSONA, J.: *Geología de Cuba*, Ed. Instituto Cubano del libro, pp. 239, La Habana, Cuba, 1972.

GARRIDO, M.: ; GÓMEZ, M.R.: ; LEÓN, A.: "Sistema automatizado de alerta temprana ante el peligro de inundaciones", *Ingeniería Hidráulica y Ambiental*, 34(3): 30–43, 2013.

GRIMALDI, S.: "Linear parametric models applied to daily hydrological series", *Journal of Hydrologic Engineering*, 9(5): 383–391, 2004.

GUERRERO, J.F.J.: ; FERNÁNDEZ, R.S.: ; ABAD, J.C.G.: "La capacidad predictiva en los métodos Box-Jenkins y Holt-Winters: una aplicación al sector turístico", *Revista Europea de Dirección y Economía de la Empresa*, 15(3): 185–198, 2006.

HEWLETT, J.D.: ; HIBBERT, A.R.: "Factors affecting the response of small watersheds to precipitation in humid areas", *Forest hydrology*, 1: 275–290, 1967.

HORTON, R.E.: "The role of infiltration in the hydrologic cycle", *Eos, Transactions American Geophysical Union*, 14(1): 446–460, 1933.

INRH: *Catálogo de Cuencas Hidrográficas (Río Cuyaguatzeje)*, Instituto Nacional de Recursos Hidráulicos, 27pp., Pinar del Río, Cuba, 2000.

JONSDOTTIR, H.: ; MADSEN, H.: ; PALSSON, O.P.: "Parameter estimation in stochastic rainfall-runoff models", *Journal of hydrology*, 326(1-4): 379–393, 2006.

KRZHIZHANOVSKAYA, V.V.: ; SHIRSHOV, G.S.: ; MELNIKOVA, N.B.: ; BELLEMAN, R.G.: ; RUSADI, F.I.: ; BROEKHUIJSEN, B.J.: ; GOULDBY, B.P.: ; LHOMME, J.: ; BALIS, B.: ; BUBAK, M.: "Flood early warning system: design, implementation and computational modules.", pp. 106–115, 2011.

KRZYSZTOFOWICZ, R.: "The case for probabilistic forecasting in hydrology", *Journal of hydrology*, 249(1-4): 2–9, 2001.

KWON, J.-H.: *Streamflow forecasting for Geum River basin by using ARMAX model*, Tesis (en opción al Master of Science), Department of Civil Engineering, Graduate School of Yonsei University, 2002.

LADO, M.: ; BEN-HUR, M.: "Soil mineralogy effects on seal formation, runoff and soil loss", *Applied Clay Science*, 24(3-4): 209–224, 2004.

LIANG, H.: ; ZHUANG, W.: "Stochastic modeling and optimization in a microgrid: A survey", *Energies*, 7(4): 2027–2050, 2014.

MAKATJANE, K.: ; MOROKE, N.: "Comparative study of holt-winters triple exponential smoothing and seasonal Arima: forecasting short term seasonal car sales in South Africa", *Makatjane KD, Moroke ND*, , 2016.

MODARRES, R.: "Streamflow drought time series forecasting", *Stochastic Environmental Research and Risk Assessment*, 21(3): 223–233, 2007.

MONTANARI, A.: ; BRATH, A.: "A stochastic approach for assessing the uncertainty of rainfall-runoff simulations", *Water Resources Research*, 40(1), 2004.

MONTANARI, A.: ; ROSSO, R.: ; TAQQU, M.S.: "Fractionally differenced ARIMA models applied to hydrologic time series: Identification, estimation, and simulation", *Water resources research*, 33(5): 1035–1044, 1997.

MORADKHANI, H.: ; SOROOSHIAN, S.: "General review of rainfall-runoff modeling: model calibration, data assimilation, and uncertainty analysis", *Hydrological modelling and the water cycle*, pp. 1–24, Ed. Springer, 2009.

MORGAN, R.P.: *Soil erosion and Soil Conservation*, Ed. Blackwell Publishing Ltd., Oxford, UK, Third edition ed. , 2005.

MORIN, J.: ; BENYAMINI, Y.: ; MICHAELI, A.: "The effect of raindrop impact on the dynamics of soil surface crusting and water movement in the profile", *Journal of Hydrology*, 52: 321–335, 1981.

NIEZGODA, S.R.: ; KANJARLA, A.K.: ; BEYERLEIN, I.J.: ; TOMÉ, C.N.: "Stochastic modeling of twin nucleation in polycrystals: an application in hexagonal close-packed metals", *International journal of plasticity*, 56: 119–138, 2014.

ONEI: *Panorama Ambiental de Cuba 2015*, Centro de Gestión de la Información Económica, Medioambiental y Social. Oficina Nacional de Estadística e Información, República de Cuba, 2016.

PACHECO, R.M.: *Modelación hidrológica con Hec-Hms en cuencas montañosas de la región oriental de Cuba.*, 128pp., Tesis (en opción al Master en Ingeniería Hidráulica), Instituto Superior Politécnico José Antonio Echeverría (CUJAE), La Habana, Cuba, 2012.

PEKTAŞ, A.O.: ; CIGIZOGLU, H.K.: "ANN hybrid model versus ARIMA and ARIMAX models of runoff coefficient", *Journal of Hydrology*, 500: 21–36, 2013.

PLANOS, E.O.: *Características Hidrológicas de las Grandes Precipitaciones: Región Occidental – Central. PNCT: Cambios Climáticos y Evolución del Medio Ambiente Cubano*, Instituto de Meteorología, 1998.

PLATE, E.J.: "Early warning and flood forecasting for large rivers with the lower Mekong as example", *Journal of Hydro-environment Research*, 1(2): 80–94, 2007.

PRESTON, B.L.: ; JONES, R.N.: "Evaluating sources of uncertainty in Australian runoff projections", *Advances in Water Resources*, 31(5): 758–775, 2008.

RÍOS, G.: *Series de Tiempo*, 2008,

RODRÍGUEZ, Y.: *Modelación hidrológica de avenidas. Caso de estudio: Cuenca del río Zaza*, 132pp., Tesis (en opción al grado científico de Doctor en Ciencias Técnicas), Instituto Superior Politécnico José Antonio Echeverría, La Habana, Cuba, 2011.

RODRÍGUEZ, Y.: ; MARRERO DE LEÓN, N.: "Simulación hidrológica en dos subcuencas de la cuenca del río Zaza de Cuba", *Ingeniería Hidráulica y Ambiental*, 36(2): 109–123, 2015.

RODRÍGUEZ, Y.: ; MARRERO DE LEÓN, N.: ; GIL, L.: "Modelo lluvia-escurrimiento para la cuenca del río Reno", *Revista Ciencias Técnicas Agropecuarias*, 19(2): 31–37, 2010.

SANG, Y.-F.: "A review on the applications of wavelet transform in hydrology time series analysis", *Atmospheric research*, 122: 8–15, 2013.

SEIBERT, J.: ; BISHOP, K.: ; RODHE, A.: ; MCDONNELL, J.J.: "Groundwater dynamics along a hillslope: A test of the steady state hypothesis", *Water Resources Research*, 39(1), 2003.

SHUMWAY, R.H.: ; STOFFER, D.S.: *Time series analysis and its applications: With R examples*, Ed. New York: Springer, Third Edition ed. , pp. 596, ISBN-978-1-4419-7864-6, 2011.

SUN, K.: ; YAN, D.: ; HONG, T.: ; GUO, S.: "Stochastic modeling of overtime occupancy and its application in building energy simulation and calibration", *Building and Environment*, 79: 1–12, 2014.

TARBOTON, D.G.: "Rainfall-runoff processes", *Utah State University*, , 2003.

TOKAR, A.S.: ; JOHNSON, P.A.: "Rainfall-runoff modeling using artificial neural networks", *Journal of Hydrologic Engineering*, 4(3): 232–239, 1999.

TRIVIÑO, A.: ; ORTIZ, S.: "Metodología para la modelación distribuida de la escorrentía superficial y la delimitación de zonas inundables en ramblas y ríos-rambla mediterráneos", , 2004.

VALIPOUR, M.: ; BANIHABIB, M.E.: ; BEHBAHANI, S.M.R.: "Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir", *Journal of hydrology*, 476: 433–441, 2013.

WANG, W.: ; CHAU, K.: ; XU, D.: ; CHEN, X.-Y.: "Improving forecasting accuracy of annual runoff time series using ARIMA based on EEMD decomposition", *Water Resources Management*, 29(8): 2655–2675, 2015.

WILKS, D.S.: *Statistical Methods in the Atmospheric Sciences*, ser. INTERNATIONAL GEOPHYSICS SERIES, 100, Ed. Academic Press, Thrid Edition ed. , pp. 676, ISBN-978-0-12-385022-5, 2011.

WU, C.L.: ; CHAU, K.-W.: "Prediction of rainfall time series using modular soft computing methods", *Engineering applications of artificial intelligence*, 26(3): 997–1007, 2013.

YU, W.: ; NAKAKITA, E.: ; JUNG, K.: "Flood Forecast and Early Warning with High-Resolution Ensemble Rainfall from Numerical Weather Prediction Model", *Procedia Engineering*, 154: 498–503, 2016.