

**MODELAMIENTO ESTOCÁSTICO DE VARIABILIDAD HIDROLÓGICA A  
ESCALAS DE ALTA RESOLUCIÓN TEMPORAL**

Trabajo presentado

Por

**JOHN ALEXANDER CHAVARRO DIAZ**

Ingeniero Agrícola

Trabajo dirigido por:

PhD. Efraín Antonio Domínguez Calle

Ingeniero Hidrólogo

Sometido a la Facultad de Ingeniería  
Pontificia Universidad Javeriana como requisito  
Parcial para la obtención del grado

**MAGISTER EN HIDROSISTEMAS**

Abril de 2017

Pontificia Universidad Javeriana – Sede Bogotá

Facultad de Ingeniería

Maestría en Hidrosistemas



MODELAMIENTO ESTOCÁSTICO DE VARIABILIDAD HIDROLÓGICA A ESCALAS DE  
ALTA RESOLUCIÓN TEMPORAL

Trabajo presentado

Por

**Ing. Agrícola. JOHN ALEXANDER CHAVARRO DIAZ**

Aprobado en cuanto a estilo y contenido por:

-----  
**PhD Jorge Alberto Escobar Vargas, Director Maestría en Hidrosistemas**

**PhD. Nelson Obregón Neira**

Jurado No. 1

**PhD. Hebert Rivera González**

Jurado No. 2

**PhD. Jorge Alberto Escobar**

Jurado No. 3

**M.Sc. Jorge Luis Sánchez Téllez**

Decano Facultad Ingeniería

## DEDICACIÓN

A Dios,  
A mi esposa e hija,

## AGRADECIMIENTOS

En primer lugar, a Dios por todas las bendiciones,

A mi familia, por su incondicional apoyo, mi hermano Jorge y la familia Cenigaa,

A mi tutor y amigo Efraín, por todos estos años de enseñanzas,

Al Instituto Geofísico, en cabeza del profesor y amigo Nelson Obregón, por sus consejos y por lo que se convirtió en mi segunda casa en estos últimos años,

A mis compañeros en esta aventura académica, en especial a Andrés Velasco, por el apoyo y amistad.

This research is partially supported by the Programmatic Cooperation between the Directorate-General for International Cooperation (DGIS) of the Dutch Ministry of Foreign Affairs and IHE Delft in the period 2016 - 2020, also called DUPC2.



Ministry of Foreign Affairs of the Netherlands

## TABLA DE CONTENIDO

1	OBJETIVOS GENERAL Y ESPECÍFICOS .....	11
1.1	OBJETIVO GENERAL.....	11
1.2	OBJETIVOS ESPECÍFICOS .....	11
1.3	PALABRAS CLAVE .....	11
2	MARCO TEÓRICO .....	12
2.1	VARIABILIDAD HIDROLÓGICA.....	12
2.2	CUENCA HIDROGRÁFICA.....	13
2.2.1	Características morfométricas .....	14
2.2.2	Caracterización de la red de drenaje.....	22
2.2.3	Coberturas.....	23
2.3	SISTEMAS COMPLEJOS .....	23
2.4	MAGNITUD, SISTEMA Y PROCESO ALEATORIO .....	25
2.5	PROBABILIDAD MULTIVARIADA, MARGINAL Y CONDICIONADA .....	26
2.6	CARACTERIZACIÓN DE LA VARIABILIDAD HIDROLÓGICA POR MOMENTOS ESTADÍSTICOS 28	28
2.7	PROCESOS ESTACIONARIOS Y NO ESTACIONARIOS.....	29
2.8	MODELACIÓN NUMÉRICA DE SISTEMAS DE COMPLEJIDAD ALGORÍTMICA SIGNIFICATIVA 30	30
2.8.1	Complejidad Algorítmica Mínima .....	31
2.8.2	Complejidad Algorítmica Máxima.....	31
2.8.3	Complejidad Algorítmica Significativa.....	31
2.9	ECUACIÓN DE LANGEVIN.....	32
2.10	ECUACIÓN DE FOKKER-PLANCK-KOLMOGOROV (MFPK) .....	33
2.10.1	La axiomática de Kolmogorov .....	33
2.10.2	Ecuación $n$ -dimensional de FPK .....	33
3	METODOLOGÍA .....	36
3.1	SELECCIÓN DE ÁREAS DE ESTUDIO Y FUENTES DE INFORMACIÓN.....	36
3.1.1	Determinación de la estructura del proceso Markoviano .....	37
3.1.2	Delimitación de áreas de drenaje para las estaciones seleccionadas .....	38
3.1.3	Selección de series de precipitación.....	38
3.2	ANÁLISIS DE CONSISTENCIA DE LA INFORMACIÓN.....	38
3.2.1	Identificación de datos anómalos .....	39
3.3	CARACTERIZACIÓN MORFOMÉTRICA.....	39
3.4	CARACTERIZACIÓN DEL PROCESO ESTOCÁSTICO EN LAS ÁREAS DE ESTUDIO.....	39
3.4.1	Pruebas de aleatoriedad.....	39
3.4.2	Ajuste de función de densidad de probabilidad.....	41
3.4.3	Estimación de momentos .....	41
3.5	APROXIMACIÓN NUMÉRICA DE LA ECUACIÓN 2D DE FOKKER-PLANCK-KOLMOGOROV ..	41
3.5.1	Condiciones iniciales y de frontera .....	44
3.5.2	Desarrollo y verificación del código .....	46

3.5.3	Optimización de parámetros .....	46
3.6	IDENTIFICACIÓN DE ESTRUCTURA MATEMÁTICA PARA LOS TÉRMINOS DE ADVECCIÓN Y DIFUSIÓN .....	48
3.7	RELACIÓN DE PARÁMETROS ÓPTIMOS CON FACTORES EXTERNOS E INTERNOS DEL SISTEMA .....	48
3.7.1	Regresiones lineales múltiples .....	49
3.7.2	Regresiones lineales simples .....	49
3.7.3	Estimación de parámetros .....	50
3.7.4	Estimación del error estándar de la regresión .....	51
3.7.5	Prueba de significancia de la regresión .....	51
4	RESULTADOS .....	53
4.1	ÁREAS DE ESTUDIO .....	53
4.2	CARACTERIZACIÓN MORFOMÉTRICA .....	63
4.3	ANÁLISIS DE INFORMACIÓN HIDROCLIMATOLÓGICA .....	67
4.3.1	Análisis visual .....	67
4.3.2	Análisis por el método de desviaciones absolutas de la mediana (MADS) .....	70
4.3.3	Análisis por el método de Grubbs .....	72
4.4	CARACTERIZACIÓN PROBABILÍSTICA .....	75
4.5	IDENTIFICACIÓN DE PARÁMETROS DE LA FPK .....	79
4.5.1	Pre-optimización .....	80
4.5.2	Optimización .....	81
4.5.3	Parámetros óptimos .....	83
4.6	ANÁLISIS DE LOS PARÁMETROS OBTENIDOS Y CARACTERÍSTICAS INTERNAS Y EXTERNAS DE LA CUENCA HIDROGRÁFICA .....	88
4.6.1	Parámetros FPK y morfometría .....	88
4.6.2	Parámetros FPK y clima .....	90
4.6.3	Parámetros FPK, clima y morfometría .....	91
5	ANÁLISIS DE RESULTADOS .....	94
6	CONCLUSIONES .....	104
7	BIBLIOGRAFÍA .....	107

## LISTA DE TABLAS

Tabla 1. Clasificación de unidades hidrográficas en función del área geométrica	15
Tabla 2. Clasificación de cuenca de la FAO en función del índice de compacidad	16
Tabla 3. Rangos de Pendientes de acuerdo con IGAC	19
Tabla 4. Análisis de varianza de la regresión	52
Tabla 5. Conjunto de estaciones seleccionadas para el ejercicio de investigación	56
Tabla 6. Resumen del proceso de delimitación de áreas aferentes	59
Tabla 7. Abreviatura y unidades de los índices morfométricos	64
Tabla 8. Resumen de la caracterización morfométrica (15/30)	65
Tabla 9. Resumen de la caracterización morfométrica (30/30)	66
Tabla 10. Datos anómalos potenciales de caudal medio diario detectados por MADS	70
Tabla 11. Datos anómalos potenciales de precipitación total diario detectados por MADS	70
Tabla 12. Datos anómalos potenciales de caudal medio diario detectados por Grubbs	72
Tabla 13. Datos anómalos potenciales de precipitación total diario detectados por Grubbs	72
Tabla 14. Realizaciones del proceso de optimización NO exitosas	84
Tabla 15. Resumen de los parámetros de la ecuación de FPK para análisis de regresión lineal múltiple (versión completa en el Anexo A8)	87
Tabla 16. Índices morfométricos incluidos en el análisis de regresiones	88
Tabla 17. Resultados de las regresiones lineales múltiples con nivel de significancia $\alpha = 0.10$	89
Tabla 18. Resultados de las regresiones lineales múltiples con nivel de significancia $\alpha = 0.05$	89
Tabla 19. Resultados de las regresiones lineales múltiples con nivel de significancia $\alpha = 0.01$	89
Tabla 20. Número de regresiones estadísticamente significativas agregados a escala mensual del parámetro "k3" de la ecuación de FPK y el valor esperado de la precipitación.	90
Tabla 21. Número de regresiones estadísticamente significativas agregados a escala mensual del parámetro "k3" de la ecuación de FPK y la desviación estándar de la precipitación total promedio	90
Tabla 22. Relaciones parámetros FPK versus índices morfométricos y clima con significancia de $\alpha = 0.10$	91
Tabla 23. Relaciones parámetros FPK versus índices morfométricos y clima con significancia de $\alpha = 0.05$	92
Tabla 24. Relaciones parámetros FPK versus índices morfométricos y clima con significancia de $\alpha = 0.01$	93
Tabla 25. Resumen de las unidades que obtuvieron en alguna de sus transiciones un error superior al 10 %	98
Tabla 26. Valores promedio, máximo y mínimo de los parámetros de la ecuación de FPK óptimos del componente advectivo	99
Tabla 27. Valores promedio, máximo y mínimo de los parámetros de la ecuación de FPK óptimos del componente difusivo	100



## LISTA DE FIGURAS

Figura 1 Diagrama conceptual propuesto de variabilidad hidrológica	13
Figura 2. Visualización de un proceso estacionario (izquierda) y no estacionario (derecha)	29
Figura 3. Proceso estocástico en sentido débil. Se puede apreciar el cambio en los valores esperados y varianzas de cada conjunto estadístico diario en un proceso hidrológico.	30
Figura 4. Complejidad algorítmica en modelación matemática	31
Figura 5. Evolución temporal de la PDF conjunta para dos variables aleatorias independientes. A la izquierda, la PDF inicial y a la derecha la PDF Final. Se puede observar la transformación de la PDF debido a los efectos de advección y difusión. En términos generales, el componente advectivo describe el comportamiento general del proceso, es decir la persistencia del sistema. Por otra parte, el componente difusivo describe las fluctuaciones o desviaciones de este comportamiento general.	34
Figura 6. Diagrama metodológico desarrollado	36
Figura 7. Ramillete de realizaciones [izquierda] y curvas de densidad de probabilidad de un proceso aleatorio: caudales medios diarios en la estación Pitalito 2 Automática [21017050].	40
Figura 8. Trilogía Modelo – Algoritmo – Programa	42
Figura 9. Curva de Densidad de Probabilidad Condicionada Empírica bidimensional	45
Figura 10. Esquema general de modelación matemática	47
Figura 11. Componentes de una regresión lineal simple	51
Figura 12. Red nacional de monitoreo hidrológico y evaluación del criterio de calidad de la información	54
Figura 13. Número de estaciones hidrológicas que cumplen el criterio de longitud de registro y completitud.	55
Figura 14. Función de auto-correlación para la estación 12017080 [Izq.] se observa que la memoria del proceso deja de ser significativa en el rezago 2, mientras la estación 44157010 [Der.] presenta una inercia de 30 rezagos.	55
Figura 15. Distribución del número de estaciones en función de su grado de inercia (magnitud de la ACF significativa)	56
Figura 16. Localización de las estaciones hidrológicas seleccionadas para la investigación	58
Figura 17. Distribución de las unidades de análisis en función de su área de drenaje en Km <sup>2</sup>	59
Figura 18. Delimitación de áreas aferentes a las estaciones hidrológicas seleccionadas.	60
Figura 19. Localización de las estaciones de precipitación seleccionadas para la investigación	62
Figura 20. Algunas salidas gráficas de la caracterización morfométrica. A la izquierda se observa la curva hipsométrica para dos unidades muy distintas como se puede apreciar, la primera [Izq. Arriba] clasifica como una cuenca con procesos de erosión intermedios dada la distribución de las elevaciones hacia la parte baja de la cuenca. Por otra parte, la segunda cuenca [izq. Abajo] presenta una distribución de elevaciones mejor distribuida lo cual indica alto potencial de erosión, esto es particular de cuencas en proceso de maduración. Los escenarios de ambas cuencas son corroborados en los gráficos de la derecha, en donde se muestran los perfiles longitudinales de los cauces principales.	64
Figura 21. Series de Caudales medios diarios [m <sup>3</sup> /seg] reportados por IDEAM	67
Figura 22. Series de precipitación total diaria [mm/día] reportados por IDEAM	69
Figura 23. Ajustes de las funciones de distribución teóricas para la estación hidrológica 21017050, arriba el conjunto del día juliano 1 (época seca), abajo día juliano 105 (época húmeda).	76
Figura 24. Ajustes de las funciones de distribución teóricas para la estación meteorológica (registros de precipitación total diaria) 21015030, arriba el conjunto del día juliano 1 (época seca), abajo día juliano 105 (época húmeda).	77
Figura 25. Resumen de los ajustes de las funciones de distribución para las series hidrológicas (caudal)	78
Figura 26. Resumen de los ajustes de las funciones de distribución para las series meteorológicas (precipitación)	78
Figura 27. Evolución de la curva de densidad de probabilidad bivariado mediante el esquema de diferencias finitas explícitas.	82
Figura 28. Resumen de los resultados del proceso de optimización	84
Figura 29. Valores de la función objetivo agregados a escala mensual para la unidad de análisis 11027050 y el régimen climático para la cuenca aferente a la estación hidrológica.	85
Figura 30. Desempeños obtenidos en el proceso de optimización matemática para la estación 11027050 sobre el río Atrato en los periodos seco (izquierda) y húmedo (derecha).	86
Figura 31. Desempeños obtenidos en el proceso de optimización matemática para la estación 11027050 sobre el río Atrato para un periodo de transición entre un régimen húmedo a seco	86
Figura 32. Análisis de componentes principales por índices morfométricos y por cuencas hidrográficas	95
Figura 33. Distribución espacial de las cuencas hidrográficas grupo 1 y 2 identificadas en el análisis por PCA.	96
Figura 34. Dispersión de los parámetros óptimos encontrados para las 30 UAH y 365 transiciones.	101

Figura 35. A la izquierda, valores del coeficiente de la regresión del parámetro  $k_3$  versus el clima [*precipitación*] agregado a escala mensual, y el régimen de precipitaciones promedio de las unidades de análisis hidrográfico. A la derecha, gráfico de dispersión de las mismas variables. La significancia confiabilidad de la regresión es del 99%.

102

## 1 OBJETIVOS GENERAL Y ESPECÍFICOS

De acuerdo con la propuesta de investigación aprobada por el comité de la maestría en Hidrosistemas, el objetivo general y los objetivos específicos del presente trabajo se definen a continuación:

### 1.1 OBJETIVO GENERAL

- Determinar si un modelo estocástico físicamente basado puede describir la variabilidad hidrológica de las cuencas hidrográficas a escalas de alta resolución temporal bajo condiciones de no estacionariedad.

### 1.2 OBJETIVOS ESPECÍFICOS

- Plantear el sistema de ecuaciones diferenciales estocásticas que permitan establecer las relaciones físicas del sistema hidrológico.
- Generar las curvas de densidad de probabilidad teóricas de acuerdo con el modelo estocástico planteado.
- Priorizar los factores que controlan la variabilidad de eventos extremos en periodos de alta y baja humedad.

### 1.3 PALABRAS CLAVE

Variabilidad hidrológica, procesos estocásticos, ecuación diferencial estocástica, modelación hidrológica, no estacionariedad, funciones de densidad probabilística, ecuación de Fokker-Planck-Kolmogorov, sistemas no-lineales, cuenca hidrográfica.

## 2 MARCO TEÓRICO

En la gestión de los recursos hídricos, según la disponibilidad de información, la toma de decisiones se realiza en condiciones: deterministas, de riesgo y de incertidumbre (Domínguez & Kovalenko, 2009). Una de las causas es la alta variabilidad presente en las cuencas hidrográficas, la cual es registrada por los sistemas de monitoreo hidrometeorológico.

Estadísticamente, la variabilidad es una medida de dispersión y se puede identificar como qué tan cerca, o lejos, se encuentran los datos en un conjunto estadístico con respecto a un valor central, (Mendoza & Bautista, 2002). A continuación, se expone el marco teórico básico para el desarrollo del presente trabajo de investigación.

### 2.1 VARIABILIDAD HIDROLÓGICA

De acuerdo al organismo de las Naciones Unidas para la Educación, la Ciencia y la Cultura, en adelante UNESCO y la Organización Meteorológica Mundial, en adelante WMO (UNESCO & WMO, 2012), la variabilidad hidrológica se refiere a la desviación de las estadísticas de las mediciones hidrológicas en un determinado periodo respecto de las estadísticas de largo plazo de dicho periodo. Según Woods, la variabilidad hidrológica son los cambios en las variables de estado que definen la hidrología (precipitación, temperatura del aire, humedad del suelo, nieve, niveles freáticos y caudales) entre dos puntos distintos y/o el mismo punto en diferentes intervalos de tiempo (Woods, 2006), es decir, corresponde a los cambios en los componentes del ciclo hidrológico en tiempo y espacio. También, la variabilidad hidrológica se refiere a los cambios sean de origen natural o antropogénicos (variación o modificación) en el estado medio de las variables del ciclo hidrológico y de sus estadísticos (media, desviación estándar, la ocurrencia de eventos extremos, etc.), observables en las escalas espacial y temporal de análisis (Domínguez, 2015b).

Sin embargo, para entender el comportamiento de la hidrología y en especial su variabilidad, es necesario conocer el comportamiento de los factores que la direccionan a diferentes escalas espaciotemporales. En este sentido, varios autores (Andrés-Doménech, García-Bartual, Montanari, & Marco, 2015; Domínguez, 2015b; Poff et al., 1997; Velasco, 2016; von Storch & Navarra, 1999; Woods, 2006) atribuyen como una de las principales causas de variabilidad hidrológica temporal a procesos meteorológicos los cuales controlan la precipitación y la evaporación principales entradas y salidas del sistema hidrológico.

Esta variabilidad climática debida a procesos meteorológicos se refiere a las variaciones en el estado medio y otras estadísticas del clima, en diferentes escalas espaciotemporales. Estas pueden obedecer a procesos internos naturales del sistema climático (variabilidad endógena), o a variaciones en fuerzas externas de origen antropogénico o natural (variabilidad exógena) (IPCC, 2013). Dichos componentes de variabilidad climática pueden agruparse en dos tipos principales: aleatorio y determinístico. En nuestro caso (el trópico), la variabilidad climática a diferentes escalas temporales está dominada por patrones cíclicos y recurrentes (MJO, Ondas Ecuatoriales, ENSO, QBO, AMO) (Laing & Evans, 2015).

Por otro lado, y retomando los componentes espaciotemporales de la variabilidad hidrológica, el componente espacial está gobernado externamente por patrones del clima, suelos, vegetación,

topografía y geología. Sin embargo, a escalas muy gruesas, se desarrollan estructuras espaciales complejas las cuáles son resultado de la dinámica interna del sistema hidrológico (Milly, Wetherald, Dunne, & Delworth, 2002; Velasco, 2016). Milly *et al.* (Milly et al., 2002) resaltan la existencia de otros factores más allá de la variabilidad climática que pueden inducir cambios significativos en la frecuencia de inundaciones como por ejemplo la variabilidad espacial que pueden promover las propiedades o los cambios en la cuenca hidrográfica, como por ejemplo cambios en la geometría de la corrientes y/o el uso del suelo. De hecho la escorrentía puede ser interpretada como una convolución suavizada de la precipitación (pasada y presente), donde el suavizado es operado (función filtro) por la cuenca de drenaje a través del tiempo de concentración (Andrés-Doménech et al., 2015). De hecho, en función de las características físicas de la cuenca y el clima, la respuesta hidrológica puede promediar (efecto suavizado) las variaciones espaciotemporales de la precipitación – por ejemplo, atenuando la variabilidad climática. Aunque Doménech asegura que la variabilidad climática no implica variabilidad hidrológica (Andrés-Doménech et al., 2015), esto puede ser discutido ya que indiscutiblemente la variabilidad hidrológica incluye la variabilidad climática lo cual puede ser más o menos importante en función de la escala temporal de análisis, como por ejemplo, en la escala mensual.

Por lo tanto, los fenómenos hidroclimáticos tienen componente determinista y de aleatoriedad (al igual que los fenómenos climáticos); de este modo, se tiene un sistema gobernado por la física, en el cual, iguales estados del sistema pueden generar diferentes realizaciones del clima y, consecuentemente, de la hidrología. Esa posibilidad de distintos resultados es entendida desde la estadística y la física como variabilidad (Domínguez, 2015b; Velasco, 2016). De esta forma, el mismo autor propone un esquema de caracterización de la variabilidad hidroclimática en función de los componentes estacionales (deterministas) y sus desviaciones (aleatoriedad).

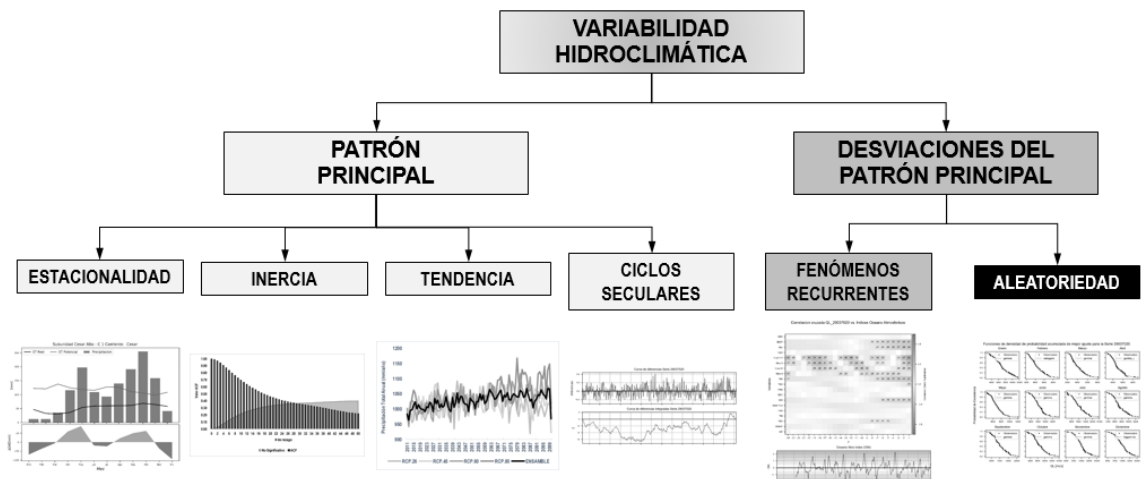


Figura 1 Diagrama conceptual propuesto de variabilidad hidroclimática  
Fuente: (Domínguez, 2015b; Velasco, 2016)

## 2.2 CUENCA HIDROGRÁFICA

La WMO y UNESCO definen el concepto general de cuenca como “área de drenaje de un curso de agua, río o lago” (UNESCO & WMO, 2012). De acuerdo al HIMAT, ahora IDEAM, “la cuenca es un espacio geográfico limitado por divisorias de agua donde se expresa el ciclo hidrológico en un

*volumen de control que permite, a partir del agua precipitada que entra a este dominio, determinar la parte que escurre a lo largo de las laderas y que puede ser interceptada por las depresiones naturales del terreno, la fracción que se evapora o infiltra y finalmente la cantidad que fluye a través de los drenajes naturales de la cuenca y forma el flujo superficial” (HIMAT, 1985).*

La geomorfometría es la ciencia del análisis cuantitativo de la superficie terrestre (Pike, Evans, & Hengl, 2009). Actualmente, es la agregación de distintas técnicas de análisis cartográfico que explican la topografía del terreno mediante representaciones digitales del mismo. El objetivo fundamental en geomorfometría es la extracción de parámetros y objetos desde los modelos digitales de terreno.

Teniendo en cuenta que la gravedad domina el movimiento del agua y una multitud de fenómenos asociados en o cerca de la superficie (Gruber & Peckham, 2009), y, debido a la amplia disponibilidad de los modelos digitales de terreno (MDT) a diferentes resoluciones, las técnicas geomorfométricas son extraordinariamente potentes en la cuantificación, análisis, previsión o parametrización de fenómenos relacionados con el flujo de agua superficial. Sin embargo, la elección de los métodos depende de la tarea a realizar (por ejemplo, la hidrología en grandes cuencas o la geomorfología en cuencas de alta montaña) y los datos disponibles (Gallant & Dowling, 2003; Gruber & Peckham, 2009; Nardi, Vivoni, & Grimaldi, 2006; Roberts, Dowling, & Walker, 1997).

### **2.2.1 Características morfométricas**

De acuerdo con lo anterior, las características de una cuenca y de las corrientes que forman el sistema hidrográfico pueden representarse cuantitativamente mediante índices de forma y relieve de la cuenca y de la conexión con la red fluvial. Muchos de los índices son razones matemáticas, por lo que pueden utilizarse para caracterizar y comparar cuencas de diferentes tamaños (Domínguez, 2010). El ingeniero e hidrólogo estadounidense Robert E. Horton fue el primero que estableció un método cuantitativo para analizar las redes de drenaje. Esta clasificación de las corrientes, desarrollada a comienzos de la década de 1940, establece una estructura jerárquica (Chow, 1994). Horton estableció en 1945 las leyes estadísticas de la composición de las redes de drenaje en las que relacionaba la categoría, número, longitud y área de drenaje de las corrientes. Las leyes de Horton, como se han denominado, fueron modificadas y ampliadas con posterioridad, principalmente por los investigadores estadounidenses A. N. Strahler y R. L. Shreve.

La morfometría hidrográfica actual tiende a centrarse en el área, longitud, forma, atributos del relieve y densidad de drenaje de la cuenca. La importancia de este tipo de caracterizaciones radica en el poder de inferencia que genera sobre fenómenos de inundaciones, por ejemplo. A continuación se describen algunos parámetros morfométricos para hidrología (Domínguez, 2010). Los parámetros pueden ser obtenidos en base a cualquier modelo digital de terreno de resolución variable, y además puede ser ajustado a la red hidrográfica “real” mediante el “quemado de corrientes” (stream burning), lo cual mejora la calidad de los índices de la red de drenaje y el relieve. A continuación, se explican algunos de ellos.

### 2.2.1.1 Área de la cuenca – Área aferente (A)

Este parámetro probablemente resulta ser el más importantes ya que está directamente relacionado con los procesos hidrológicos que se dan al interior de la cuenca. Está definida como la proyección ortogonal de toda la superficie de drenaje de un sistema de escorrentía dirigido directa o indirectamente a un mismo cauce natural (Reyes T., Barroso, & Carvajal E., 2010), el mismo autor propone una clasificación en función del área de la unidad hidrográfica.

Tabla 1. Clasificación de unidades hidrográficas en función del área geométrica

Área (Km <sup>2</sup> )	Unidad hidrológica
<5	Unidad
5-20	Sector
20-100	Microcuenca
100-300	Subcuenca
>300	Cuenca

Fuente: (Jiménez E. & Materón M., 1986)

### 2.2.1.2 Perímetro o longitud del parteaguas

El perímetro se estima mediante plataformas CAD y/o GIS, se calcula una vez se ha delimitado la cuenca bajo un sistema de geo-referenciación. Este parámetro junto al área permite inferir sobre la forma de la cuenca (Reyes T. et al., 2010).

### 2.2.1.3 Longitud del cauce principal (L<sub>t</sub>)

Es la distancia del recorrido que hace el cauce principal desde su nacimiento hasta su desembocadura medida en kilómetros. Usualmente L<sub>t</sub> presenta un grado de sinuosidad, el cual se representa a través del coeficiente de sinuosidad K<sub>s</sub> que constituye la relación entre la distancia total configurada por el recorrido de la corriente L<sub>t</sub>, sobre la distancia lineal desde el nacimiento hasta la desembocadura L<sub>l</sub> (Domínguez, 2010):

$$K_s = \frac{L_t}{L_l}$$

Ecuación 1. Coeficiente de Sinuosidad

Cauces con índice de sinuosidad mayores a 1.5 son considerados corrientes meandriformes, los cuales tienden a presentar pendientes suaves y mayor tiempo de tránsito hidrológico.

### 2.2.1.4 Ancho medio de la cuenca (B)

Se determina cómo el cociente entre el área de la cuenca y la longitud lineal de la corriente principal:

$$B = \frac{A}{L_t}$$

Ecuación 2. Ancho medio de la cuenca

### 2.2.1.5 Factor de forma (F)

Esta variable relaciona el área de la cuenca (A) con el cuadrado de la longitud total de la corriente principal ( $L_t$ ) (Ecuación 3), en casos donde la cuenca es perfectamente cuadrada el factor de forma es igual a la unidad, mientras cuando se trate de cuencas alargadas el factor de forma tomara valores inferiores a la unidad (Jardí, 1985).

$$F = \frac{A}{L_t^2}$$

Ecuación 3. Factor de forma

Este parámetro adimensional mide la tendencia de la cuenca hacia las crecidas, rápidas y muy intensas a lentas y sostenidas, si tiende hacia valores extremos grandes o pequeños, respectivamente. Un valor de F superior a la unidad dará el grado de achatamiento de ella o de un río principal corto, por consiguiente con tendencia a concentrar el escurrimiento de una lluvia intensa formando fácilmente grandes crecidas (Reyes T. et al., 2010).

### 2.2.1.6 Coeficiente de compacidad ( $K_c$ ) o índice de Gravelius

Es la relación entre la longitud del perímetro de la cuenca  $P_{cuenca}$  y el perímetro de un círculo con área equivalente a la de la cuenca  $P_{circulo}$ :

$$k_c = \frac{P_{cuenca}}{P_{circulo}} = \frac{P_{cuenca}}{2\sqrt{\pi A}} = 0.28 \frac{P_{cuenca}}{\sqrt{A}}$$

Ecuación 4. Coeficiente de compacidad

La Tabla 2 contiene la clasificación de cuenca en función del índice de compacidad según la FAO (FAO, 1985), cuencas con valores de compacidad superiores a 1.5 tienden a tener un tránsito más largo de sus crecientes y por lo tanto atenuación de los hidrogramas disminuyendo el riesgo de inundación.

Tabla 2. Clasificación de cuenca de la FAO en función del índice de compacidad

Tipo	Rango	Descripción
KC <sub>1</sub>	1.00 – 1.25	Redonda a oval redonda
KC <sub>2</sub>	1.26 – 1.50	Oval redonda a oval oblonga
KC <sub>3</sub>	1.51 – 1.75	Oval oblonga a rectangular oblonga
KC <sub>4</sub>	> 1.75	rectangular oblonga



Fuente: (FAO, 1985)

### 2.2.1.7 Índice de alargamiento (Ia)

Índice propuesto por Horton, relaciona la longitud máxima de la cuenca con su ancho máximo (Ecuación 5). Aquellas cuencas que registran valores mayores a uno presentan una relación largo / ancho a favor del numerador, obedeciendo a formas alargadas; asimismo este parámetro adimensional permite predecir la dinámica del movimiento del agua en los drenajes y su potencia erosiva o de arrastre (Reyes T. et al., 2010)

$$Ia = \frac{L_t}{B}$$

Ecuación 5. Índice de alargamiento

### 2.2.1.8 Perfil longitudinal del cauce principal

Este perfil muestra la evolución de las cotas de altitud del cauce principal. El corte longitudinal, se puede construir con base en batimetrías realizadas en campo o en su defecto con base en los modelos de elevación digital mundiales. El levantamiento de campo del perfil longitudinal puede realizarse para toda la corriente o en su defecto para tramos característicos. Los tramos característicos se pueden delimitar de acuerdo con los sitios de cambios fuertes en las profundidades del río, puntos de socavación o acumulación, lugares con cambio en el ancho superior del cauce, etc. Para construir el perfil longitudinal, con una línea recta horizontal, según la escala seleccionada, se expresa la longitud de la corriente. En el eje vertical se representan las cotas del fondo del río o del nivel del agua para cada punto de la corriente donde se tenga información (Domínguez, 2010).

### 2.2.1.9 Altura del cauce principal ( $H_{cauce}$ )

Se calcula como la altura promedio de los tramos característicos utilizados en la construcción del perfil longitudinal del cauce principal. Para las corrientes con muchos cambios de altura en el recorrido del cauce principal se calcula la pendiente promedio ponderada con base en las frecuencias relativas producto de la agrupación en clases (usualmente entre 8 a 12 clases) de los valores de altura tomados en los tramos característicos (Domínguez, 2010):

$$H_{cauce} = H_1f_1 + H_2f_2 + \dots + H_nf_n$$

Ecuación 6. Altura promedio ponderada del cauce

Aquí:  $H_i$  es la cota de altura en la mitad del intervalo de cada clase y  $f_i$  es la frecuencia que le corresponde a ese intervalo.

### 2.2.1.10 Pendiente del cauce principal ( $I_{cauce}$ )

Por pendiente se entiende el cociente entre los incrementos de alturas/cotas y los incrementos de longitud de la corriente para un tramo determinado. Se distingue la pendiente promedio y pendiente promedio ponderada. La pendiente promedio  $I_{cauce}$  es el cociente entre la caída de alturas/cotas  $\Delta H$  sobre la longitud del tramo  $\Delta L$  en la que esta caída ocurre (Domínguez, 2010):

$$I_{cauce} = \frac{\Delta H}{\Delta L}$$

Ecuación 7. Pendiente promedio del cauce principal

Para las corrientes con muchos cambios de pendiente se calcula pendiente promedio ponderada. Ésta utiliza las frecuencias relativas producto de la agrupación en clases (usualmente entre 8 a 12 clases) para ponderar los valores de pendiente calculados (Domínguez, 2010):

$$I_{cauce} = I_1f_1 + I_2f_2 + \dots + I_nf_n$$

Ecuación 8. Pendiente promedio ponderada del cauce principal

Aquí:  $I_i$  es la pendiente en la mitad del intervalo de cada clase y  $f_i$  es la frecuencia que le corresponde a ese intervalo.

La pendiente del cauce es uno de los factores importantes que inciden en la capacidad que tiene el flujo para transportar sedimentos, por cuanto está relacionada directamente con la velocidad del agua. En los tramos de pendiente fuerte los cauces tienen pendientes superiores al 5 %, y las velocidades de flujo resultan tan altas que pueden mover como carga de fondo sedimentos de diámetros mayores de 5 centímetros, además de los sólidos que ruedan por desequilibrio gracias al efecto de lubricación producido por el agua (Reyes T. et al., 2010).

### 2.2.1.11 Curva hipsométrica

Esta curva da una interpretación clara de la distribución de las áreas de la cuenca contenidas en distintas zonas altitudinales. También son un indicativo a los patrones evolutivos de las cuencas (Strahler, 1952), así, “*cuencas con mayor parte del área con elevación por encima de los valores medios de la misma, son consideradas cuencas con gran potencial erosivo (cuencas inmaduras), de lo contrario son consideradas cuencas maduras o sedimentarias*” (Strahler, 1952). Para cuencas donde esta relación se encuentre en equilibrio (área relativa cuenca – elevación ponderada) se consideran como cuencas en equilibrio.

Para construirla se determinan de 8 a 12 zonas altitudinales y de acuerdo con la información del modelo digital de terreno, se calcula el porcentaje de área contenido en cada una de ellas. En el eje horizontal se gráfica el porcentaje de área de la cuenca cubierto por cada zona altitudinal y en el vertical los intervalos de clase de cada zona altitudinal (Domínguez, 2010)

### 2.2.1.12 Altura de la cuenca ( $H_{cuenca}$ )

Se calcula con base al diagrama de frecuencias de la curva hipsométrica, homólogo al cálculo de  $H_{cauce}$ , pero usando las 8 a 12 zonas altitudinales definidas (Domínguez, 2010):

$$H_{cuenca} = H_1f_1 + H_2f_2 + \dots + H_nf_n$$

Ecuación 9. Altura promedio ponderada de la cuenca.

Aquí:  $H_i$  es la cota de altura en la mitad del intervalo de cada clase y  $f_i$  es la frecuencia que le corresponde a ese intervalo.

### 2.2.1.13 Pendiente de la cuenca ( $I_{cuenca}$ )

Para obtener la pendiente media de la cuenca con base en el DEM, se construye el diagrama de pendientes con las pendientes locales de cada celda calculada a partir del cambio de alturas de las 8 celdas vecinas (Arriba, Abajo, Derecha, Izquierda, Arriba derecha, Arriba izquierda, Abajo derecha, Abajo izquierda). De modo similar al aplicado para definir la altura media, la pendiente promedio ponderada de la cuenca se obtiene cómo (Domínguez, 2010):

$$I_{cuenca} = I_1f_1 + I_2f_2 + \dots + I_nf_n$$

Ecuación 10. Pendiente promedio ponderada de la cuenca

Aquí:  $I_i$  es la pendiente en la mitad del intervalo de cada clase y  $f_i$  es la frecuencia que le corresponde a ese intervalo. El Instituto Geográfico Agustín Codazzi propone un esquema de clasificación para este índice (IGAC, 2002). La Tabla 3 resume el sistema de clasificación para el mapa de pendientes.

Tabla 3. Rangos de Pendientes de acuerdo con IGAC

Símbolo	Gradiente (%)	Descripción
a	0 - 3	Plano
b	3 - 7	Ligeramente inclinado
c	7 - 12	Moderadamente inclinado
d	12 - 25	Fuertemente inclinado
e	25 - 50	Ligeramente escarpado
f	50 - 75	Moderadamente escarpado
g	> 75	Fuertemente escarpado

Fuente:(IGAC, 2002)

### 2.2.1.14 Profundidad de disección

La profundidad de disección de la cuenca  $H_{dis}$  representa la diferencia entre la altura media de la cuenca  $H_{cuenca}$  y la altura media del perfil longitudinal del Río principal  $H_{cauce}$  (Domínguez, 2010):

$$H_{dis} = H_{cuenca} - H_{cauce}$$

## Ecuación 11. Profundidad de disección de la cuenca

**2.2.1.15 Nivel de disección del relieve ( $N_{dis}$ )**

Resulta del doble producto de la densidad hidrográfica por la profundidad de disección (Domínguez, 2010):

$$N_{dis} = 2DH_{dis}$$

Ecuación 12. Nivel de disección del relieve

Estos dos últimos indicadores están relacionados con la modelación del terreno por parte del agua, es decir, en la creación de arroyos y ríos, específicamente con los procesos erosivos y de acumulación de material. También en contexto con otros indicadores como el relieve, la pendiente, densidad de drenaje, erosión y la profundidad del nivel freático, son claves en los factores geomorfológicos que controlan los flujos subterráneos y su acumulación en los acuíferos (Bhamare, Agone, Patil, & Science, 2012). Por lo tanto, valores altos de estos índices exponen mayor riesgo geomorfológico ya que estas cuencas estarán sujetas a procesos erosivos más intensos.

**2.2.1.16 Tiempos de Concentración ( $T_c$ )**

Es el tiempo transcurrido entre el final del hietograma de excesos y el final del escurrimiento directo, siendo ésta la definición que aparece reseñada en la literatura con mayor frecuencia. Sin embargo, otros autores reportan el tiempo de concentración como el tiempo comprendido entre el centroide del hietograma de excesos y el punto de inflexión sobre la curva de recesión del hidrograma de escurrimiento directo.

Además, se puede definir como el tiempo que demora en viajar una partícula de agua desde el punto más remoto hasta el punto de interés. Corresponde al lapso entre el final de la lluvia y el momento en que cesa el escurrimiento superficial. Existen una serie de fórmulas que permiten el cálculo de este tiempo desarrolladas por diversos autores. Algunas de las fórmulas que se emplean para el cálculo de este indicador se encuentran, Kirpich, Temez, Giandotti, Bransby-Williams, Dirección General Carreteras (DGC). Sus cálculos se muestran a continuación:

$$T_c = 0.02 * L^{0.77} * S^{-0.385}$$

Ecuación 13. Tiempo de Concentración por Kirpich  
Fuente: (Wanielista, 1997)

Donde:

- Tc      Tiempo de concentración (min)
- L      Longitud del cauce principal en (m)
- S      Pendiente del cauce principal (m/m)

$$T_c = 0.3 * \left[ \frac{L}{S^{0.25}} \right]^{0.75}$$

Ecuación 14. Tiempo de Concentración por Temez  
Fuente: (Chow, 1994)

Donde:

- Tc Tiempo de concentración (horas)  
L Longitud del cauce principal en (km)  
S Pendiente del cauce principal (m/m)

$$T_c = \left[ \frac{4\sqrt{A} + 1.5L}{0.8\sqrt{H}} \right]^{0.75}$$

Ecuación 15. Tiempo de Concentración por Giandotti  
Fuente: (Chow, 1994)

Donde:

- Tc Tiempo de concentración (Horas)  
L Longitud del cauce principal en (Km)  
A Área de la cuenca (km<sup>2</sup>)  
H Diferencia de alturas del cauce principal (m)

$$T_c = \frac{L}{1.5D} \sqrt[5]{\frac{A^2}{S}}$$

Ecuación 16. Tiempo de Concentración por Bransby-Williams  
Fuente: (Chow, 1994)

Donde:

- Tc Tiempo de concentración (Horas)  
L Longitud del cauce principal en (Km)  
A Área de la cuenca (km<sup>2</sup>)  
S Pendiente media del cauce principal (%)  
D Diámetro del círculo de área equivalente a la superficie de la cuenca (km)

$$T_c = 0.3 \left[ \frac{L}{J^{0.25}} \right]^{0.76}$$

Ecuación 17. Tiempo de Concentración por Dirección General de Carreteras  
Fuente: (Chow, 1994)

Donde:

- Tc Tiempo de concentración (Horas)  
L Longitud del cauce principal en (Km)  
J Pendiente media del cauce principal (H/L)  
H Diferencia de nivel entre el punto de salida y el punto hidrológicamente más cercano (m)

## 2.2.2 Caracterización de la red de drenaje

La red hidrográfica se constituye del cauce principal y sus tributarios, en los cuales se consideran corrientes permanentes, intermitentes y efímeras. Dicha red es función de la escala espacial del modelo digital de terreno. Algunos parámetros que definen su estructura son:

### 2.2.2.1 Densidad de la red hidrográfica (D)

Este parámetro se define como el cociente entre la longitud total de los cauces que conforman el sistema fluvial de la cuenca, expresados en kilómetros y el área total de la cuenca expresada en kilómetros cuadrados (Horton, 1945). Matemáticamente se expresa como:

$$D = \frac{\sum_{i=1}^n L_{t_i}}{A}$$

Ecuación 18. Densidad de Drenaje

Donde  $n$  es el número de corrientes visibles en la cartografía disponible y  $A$  es el área en kilómetros cuadrados la cuenca evaluada. La densidad de la red hidrográfica está controlada por el clima, la geología y la estructura del relieve (Domínguez, 2010). Con base en este parámetro se puede definir una aproximación de la longitud media de las laderas en la cuenca  $L_{lad}$ :

$$L_{lad} = \frac{1}{2D} = \frac{A}{2 \sum_{i=1}^n L_{t_i}}$$

Ecuación 19. Longitud media de las laderas

La longitud media de laderas, en conjunto con otros parámetros morfométricos como la orientación de la ladera y la pendiente, son un indicador del grado de amenaza al deslizamiento superficial en una cuenca hidrográfica (Vélez, 2004), por lo tanto, valores altos de este parámetro son indicativos de cuencas con mayor amenaza de deslizamiento superficial.

### 2.2.2.2 Número de orden de la cuenca (N)

Se determina con base en el número de orden máximo obtenido para el cauce principal de la red hídrica de la cuenca. El orden del río se calcula con base en el método propuesto por Horton en 1945 (Horton, 1945) y modificado por Strahler en 1964, donde los ríos en su nacimiento tienen valor de orden 1 y estos aumentan de orden solo cuando 2 o más ríos de un mismo orden se cruzan. Por lo tanto, la intersección de un río de primer orden y un río de segundo orden mantendrá el río de segundo orden, mientras si se unen 2 o más ríos de segundo orden el río resultante será de tercer orden. En general, dos ríos o más de orden “ $n$ ” dan origen a uno de orden “ $n + 1$ ”, en caso de que una o más corrientes de orden inferior, la corriente principal conserva el mismo orden.

Horton en 1945 estableció tres leyes, llamadas las Leyes de Horton: la ley de los números de cauces, la ley de las longitudes y la ley de las áreas drenantes a los cauces. Dichas leyes establecen que las tasas de bifurcación, las relaciones de longitud y de áreas permanecen constantes de un orden a otro en la cuenca.

### 2.2.2.3 Radio o coeficiente de bifurcación ( $R_b$ )

La ley del número de los cauces de Horton (1945) establece la relación de bifurcación como el cociente entre el número de cauces  $N$  de orden  $i$  y el número de cauces  $N$  de orden inmediatamente superior  $i + 1$  (Domínguez, 2010):

$$R_{b_{i,i+1}} = \frac{N_i}{N_{i+1}}$$

Ecuación 20. Coeficiente de bifurcación

El valor promedio del radio de bifurcación para la cuenca se obtiene como la media aritmética de los radios de bifurcación obtenidos entre las combinaciones encontradas para los órdenes de cauces en la red hídrica de la cuenca. La relación de bifurcación de Horton describe el grado de ramificación entre corrientes de orden contiguo. La relación de bifurcación permite comprender algunas variaciones geo-ecológicas que se producen en el territorio de la cuenca, fundamentalmente cambios importantes en el sustrato rocoso y de los grupos de suelos dominantes. Las cuencas cuya relación de bifurcación permanece constante, indican homogeneidad en las características geo-ecológicas anteriores (Reyes T. et al., 2010).

### 2.2.3 Coberturas

Otro aspecto importante en la génesis de los procesos de escorrentía además de los ya mencionados (clima y características geomorfológicas) son las coberturas (Milly et al., 2002). Inclusive, bajo ciertas condiciones hidrológicas los cambios en las cobertura y uso del suelo pueden afectar por ejemplo, la frecuencia de inundaciones (Andrés-Doménech et al., 2015). Sin embargo, el presente estudio se enfoca en como la geomorfología puede explicar los aspectos de variabilidad hidrológica, y, por lo tanto, no se considera explícitamente el efecto de las coberturas y usos del suelo lo cual puede ser objeto de otros proyectos de investigación dada la importancia reportada por otros autores.

## 2.3 SISTEMAS COMPLEJOS

Una forma de clasificación de sistemas puede resumirse en sistemas simples, complicados y complejos (Campos, 2006). Con respecto a este último, el estudio de la complejidad y de los sistemas complejos ha devenido, desde mediados del siglo XX, en un objeto de estudio central para la ciencia contemporánea pero también para la reflexión filosófica, ética y política (Rodríguez & Aguirre, 2011). Morín (Morin, 2004) señala que lo que hoy suele llamarse teorías de la complejidad es en realidad el nombre de un campo con límites confusos que abarca en su formulación científica a las teorías de los sistemas complejos (sistemas dinámicos, sistemas no lineales, sistemas adaptativos y de propiedades emergentes) y la teoría del caos. La complejidad se introduce en el terreno de las ciencias, una racionalidad postclásica que habilita e incorpora problemas ignorados por el pensamiento científico moderno, por ejemplo, la incertidumbre, la emergencia y la auto-organización. En este sentido, puede entenderse como un paradigma científico emergente que involucra un nuevo modo de hacer y entender la ciencia (Rodríguez & Aguirre, 2011).

Por definición, un sistema complejo es un sistema abierto, formado por un gran número de elementos que interactúan de manera no lineal, que no se puede reducir a dos o más subsistemas distintos, y que presentan propiedades emergentes (Grigolini, Allegrini, & West, 2007).

Un ejemplo de sistema complejo lo configura el ciclo hidrológico y/o proceso hidrológico (Domínguez & Kovalenko, 2009; Koutsoyiannis, 2008). Este es un componente importante del sistema climático. En promedio cerca del 60% del total de la precipitación en tierra regresa a la atmósfera vía evapotranspiración, y el resto fluye a los océanos (incluye cuerpos de agua continentales) a través de la red de drenaje (Oki, 2006) en cada cuenca hidrográfica. A partir de aquí, las redes de drenaje y la estructura morfológica configuran una amplia gama de topologías en función del régimen de flujo y su complejidad estructural (Braga, 2000). En últimas, el caudal registrado en todos los ríos es derivado de las precipitaciones, sin embargo, en un determinado momento y ubicación el caudal es debido a una combinación de factores como la escorrentía superficial, humedad del suelo y almacenamiento de los acuíferos. El clima, la geología, la topografía, los suelos y la vegetación ayudan a determinar la oferta hídrica y la forma como la precipitación llega a la corriente principal conformando de esta manera patrones de flujo local y regional (Poff et al., 1997). Para ilustrar este concepto, podemos asumir que en cuencas de alta montaña los caudales altos debido a eventos torrenciales pueden ocurrir en periodos de horas a días (para cuencas con suelos permeables) o sobre minutos a horas (para cuencas con suelos impermeables), mientras que procesos de derretimiento de nieve ocurren sobre escalas diarias a semanales y, por lo tanto, configurándose una hidrógrafa más suave. Continuando el proceso aguas abajo dentro de la cuenca hidrográfica, el caudal en la corriente principal refleja la suma de los caudales generados aguas arriba y el tiempo de tránsito de una gran cantidad de corrientes menores tributarios al principal. De aquí en adelante el tiempo de viaje del caudal, es función de aportes no-sincronizados de otros tributarios aguas abajo y de las capacidades de almacenamiento de las planicies de inundación, hecho que atenúa y amortigua los flujos máximos instantáneos. De esta forma, las hidrógrafas anuales en grandes cuencas pueden mostrar grandes picos cuando se presentan grandes volúmenes de precipitación y/o eventos de derretimiento de nieve o grandes influencias estacionales que afectan muchos tributarios al mismo tiempo (Dunne & Leopold, 1978). Retomando, la estructura física de la cuenca y, en consecuencia, del hábitat, es definido en gran medida por el proceso físico, especialmente el movimiento del agua y los sedimentos en el cauce principal y la planicie de inundación. El régimen natural de caudales organiza y define los ecosistemas fluviales (Poff et al., 1997).

En general, la sociedad trata día a día con sistemas de esta índole, por lo tanto, tratar de entender su estructura, las múltiples relaciones entre sus partes y la emergencia de propiedades puede conllevar a herramientas de mayor valor para la gestión en los distintos campos de la ciencia (economía, biología, ecología, hidrología, etc.). De acuerdo con Campos (Campos, 2006) el alcance del anterior objetivo involucra el entendimiento, entre otros aspectos de:

- Las interacciones entre los elementos y su número,
- El carácter no lineal de esas interacciones,
- Los procesos de retroalimentación,
- Propiedades emergentes y de auto-organización,
- La existencia de estados críticos y transiciones de fase,
- Leyes de potencia inversa,



- y la percolación, que en sistemas ecológicos implica procesos epidémicos como fuego, enfermedades y parásitos.

Por lo tanto, modelar un sistema complejo indica en primer término definir un conjunto de relaciones, hechos, elementos, subsistemas, niveles, operaciones y variables con una delimitación externa y cuyos componentes son fundamentales ya que generan la estructura principal del sistema. Las relaciones de los elementos reflejan y determinan la manera como se ordenarán las fronteras y las intersecciones del sistema complejo (Maldonado & Gómez Cruz, 2010). Estos sistemas pueden ser vistos como un conjunto de acciones definidas por las fuerzas modeladoras del entorno (restricciones estructurales o contingentes) y del propio sistema (restricciones del sistema y adaptación al entorno). La selección del sistema está compuesta de parámetros de intercambio, de adaptación y de diferenciación. Los aspectos físicos del desarrollo del sistema complejo son pautas específicas de intercambio con el entorno, normalmente destinadas a la preservación de sus fronteras, estructura y función (Gómez and Villamil 2009).

Una manera de avanzar hacia el estudio de sistemas complejos como el descrito aquí, requiere desarrollar técnicas para inferir el sistema dinámico, de tal forma que reproduzca las propiedades observadas del sistema, bajo diferentes condiciones de los parámetros y condiciones iniciales (Campos, 2006).

El gran valor de los modelos matemáticos en el análisis de los sistemas complejos reside en su utilización para representar situaciones que ponen en evidencia los procesos característicos de los sistemas reales (Dendrinis & Sonis, 1990; Frigg & Hartmann, 2017; Giere, 2004; Moncho, 2008; Uusitalo, Lehikoinen, Helle, & Myrberg, 2015).

## 2.4 MAGNITUD, SISTEMA Y PROCESO ALEATORIO

De acuerdo a los elementos teóricos expuestos como variabilidad y sistemas complejos, la aplicación de los enfoques estocásticos para la toma de decisiones es válido (Domínguez & Kovalenko, 2009; Gottschalk, Gottschalk, & Lars, 2005). Por lo tanto, a continuación, se describen algunos conceptos necesarios para el buen entendimiento de este tema.

El concepto de probabilidad ha sido construido durante muchos años. Un enfoque que es fácil de entender es el axioma a priori: “*si un evento puede ocurrir  $n$  veces y son mutuamente exclusivos, y si  $n_a$  es un evento del subconjunto  $n$  con un atributo  $A$ , entonces la probabilidad de ocurrencia del evento con el atributo  $A$  es  $n_a/n$ , y se denota como:  $prob(A) = \frac{n_a}{n}$* ” (Haan, 1977).

El elemento base en la teoría de probabilidades es la magnitud aleatoria (notas de clase, Domínguez), la cual es un caso generalizado de evento aleatorio (A. A. Sveshnikov, 1966). A los diferentes resultados de una medición sistemática bajo las mismas condiciones se le denomina, magnitud aleatoria, la cual puede suceder en un orden impredecible (notas de clase, Domínguez). Haan, define una magnitud o variable aleatoria como una función definida en un espacio de muestreo, la cual puede ser continua o discreta (Haan, 1977) en el sentido en que ocupe todo o parcialmente el eje numérico del conjunto de los reales ( $\mathbb{R}$ ) respectivamente. Como convención universal se ha adoptado en la medida de lo posible las letras mayúsculas de los alfabetos griego y latino para denotar una variable o magnitud aleatoria, y letras minúsculas de los respectivos alfabetos para representar sus posibles valores (Haan, 1977; A. A. Sveshnikov, 1966). También, si

una variable o magnitud aleatoria  $X$ , y,  $Z = g(X)$ , entonces  $Z$  es una variable aleatoria también (Haan, 1977).

Un conjunto de variables o magnitudes aleatorias o un sistema de variables aleatorias puede ser descrito de forma exhaustiva por una función de distribución multi-dimensional. Esto es, para el caso de un sistema de  $n$  variables aleatorias  $[X_1, X_2, \dots, X_n]$  (como por ejemplo los sistemas hidrometeorológicos):

$$F(x_1, x_2, \dots, x_n) = P(X_1 < x_1, X_2 < x_2, \dots, X_n < x_n)$$

La ecuación anterior define la probabilidad de que conjuntamente  $X_i < x_i \forall i = \{1, 2, \dots, n\}$ .

La teoría de función aleatoria es una generalidad de la teoría de probabilidades. Según Sveshnikov, una función aleatoria es aquella donde para cada argumento de la función su imagen es una variable aleatoria y el argumento es una variable no-aleatoria,  $Z = g(t)$ .

En el estudio de funciones aleatorias, se presentan dos casos de estudio: a) donde el argumento de la función puede tomar cualquier valor en un intervalo dado (finito o infinito); b) donde el argumento de la función toma valores discretos. El caso “a” suele llamarse proceso aleatorio o proceso estocástico, mientras en “b” secuencia aleatoria (A. A. Sveshnikov, 1966). Gardiner (Gardiner, 2009), describe un proceso aleatorio como un sistema  $X(t)$ , en donde una variable aleatoria  $x(t)$  existe, y los valores de  $x(t)$  son medidos en diferentes momentos en el tiempo  $t_1, t_2, \dots, t_n$ . En este caso la función de densidad de probabilidad conjunta describe todo el sistema  $X(t)$ . En otras palabras, un proceso aleatorio puede ser entendido como una función que produce magnitudes aleatorias y, por lo tanto, configura un conjunto de variables aleatorias que describe el proceso.

Los sistemas hidrológicos, y por lo tanto sus observaciones, a las escalas apropiadas presentan patrones de variación complejos tanto en el tiempo como en el espacio reflejando alguna regularidad estadística. Por tal razón tiene sentido aplicar conceptos de la estadística y la teoría de probabilidades para describir tales procesos (Gottschalk et al., 2005; Haan, 1977).

Una forma de caracterizar un proceso aleatorio, es mediante funciones de distribución de frecuencia (caso unidimensional) (Domínguez & Kovalenko, 2009; Gottschalk et al., 2005), segundo momento estadístico (funciones de correlación y covarianza, funciones espectrales y semivariogramas), y por la expansión de Karhunen-Loève (análisis armónicos, wavelets, PCA, EOF) (Gottschalk et al., 2005).

## 2.5 PROBABILIDAD MULTIVARIADA, MARGINAL Y CONDICIONADA

Como se ha mencionado, un camino hacia la caracterización de procesos aleatorios es mediante las funciones de distribución. Si las variables  $[X_1, X_2, \dots, X_n]$  son independientes entonces el proceso puede ser descrito por la probabilidad multivariada como el producto de las funciones de probabilidad de cada variable aleatoria en el sistema (Ecuación 21).

$$F(x_1, x_2, \dots, x_n) = \prod_{i=1}^n P(X_i < x_i)$$

Ecuación 21 Función de probabilidad acumulada multivariada, caso variable independiente

De forma general, si para un sistema de magnitudes aleatorias existen todas las derivadas cruzadas de la función de distribución respecto a cada argumento de la función  $F(x_1, x_2, \dots, x_n)$ , entonces la densidad de distribución de probabilidades multivariada se define como (Domínguez, 2016; Haan, 1977):

$$p(x_1, x_2, \dots, x_n) = \frac{\partial^n F(x_1, x_2, \dots, x_n)}{\partial x_1 \partial x_2 \dots \partial x_n}$$

Ecuación 22. Caso general función de densidad de probabilidad multivariada

La función de distribución acumulada y curva de densidad de probabilidad están relacionadas de la siguiente manera:

$$\begin{aligned} F(x_1, x_2, \dots, x_n) &= P(-\infty < X_1 < x_1, -\infty < X_2 < x_2, \dots, -\infty < X_n < x_n) \\ &= \int_{-\infty}^{x_1} \int_{-\infty}^{x_2} \dots \int_{-\infty}^{x_n} p(x_1, x_2, \dots, x_n) dx_1 dx_2 \dots dx_n \end{aligned}$$

Ecuación 23. Relación entre la función de distribución acumulada y la función de densidad de probabilidad, caso  $n$ -dimensional.

La función de probabilidad conjunta o multivariada describe la probabilidad de ocurrencia de un evento  $\omega$  para un rango particular o conjunto discreto de valores especificados para cada variable aleatoria que compone el proceso aleatorio (Gardiner, 2009; Haan, 1977; Koutsoyiannis, 2008; A. A. Svishnikov, 1966) esto es:  $P(X \cap Y \cap V) = P\{(\omega \in X) \cap (\omega \in Y) \cap (\omega \in V)\}$ .

Para definir la densidad de distribución marginal de  $X_j$  se debe integrar la curva de densidad de probabilidad del sistema (Ecuación 23) cambiando los límites de integración por el rango  $[-\infty, +\infty]$  cumpliendo que  $x_i \forall i \llcorner j$  (Domínguez, 2016; Haan, 1977). Sin embargo, los eventos naturales son usualmente dependientes (Campos, 2006; Koutsoyiannis, 2008; Maldonado & Gómez Cruz, 2010) (sistemas complejos), en este caso, el sistema deberá ser descrito mediante funciones de distribución condicionadas o densidad de probabilidad condicional (Domínguez & Kovalenko, 2009). Una función de distribución de una variable aleatoria es condicionada cuando tiene restricciones o es “condicionada” a valores dados en las demás variables aleatorias que componen el sistema (Haan, 1977; A. A. Svishnikov, 1966). Por ejemplo, consideremos un proceso descrito por el conjunto de variables aleatorias  $[X_1, X_2, \dots, X_n]$  en donde se conoce la magnitud tomada por una de las variables aleatorias  $X_n = x_0$ .

$$p(x_1, x_2, \dots | x_n = x_0) = \frac{p(x_1, x_2, \dots, x_0)}{p(x_0)}$$

Ecuación 24 Densidad de probabilidad condicionada de un sistema  $n$  – dimensional

Muchos sistemas complejos están condicionados por una gran cantidad de factores aleatorios. La descripción exhaustiva de estos sistemas requiere no solo de las leyes de distribución de cada uno

de los factores aleatorios que afectan al sistema sino también de las inter-relaciones o dependencias entre estos factores.

## 2.6 CARACTERIZACIÓN DE LA VARIABILIDAD HIDROLÓGICA POR MOMENTOS ESTADÍSTICOS

Como se ha explicado, las variables hidrológicas (precipitación, la evapotranspiración, el contenido de humedad en el suelo, la infiltración en el suelo y sub-suelo, los almacenamientos en reservorios y/o embalses, la escorrentía, etc.) y por lo tanto, su variabilidad son variables aleatorias (Haan, 1977; A. A. Sveshnikov, 1966; Velasco, 2016). De esta forma, procesos cuyo resultado es una variable aleatoria es considerado un experimento, y cada simple salida de ese experimento de acuerdo a la sigma algebra definida se denomina realización (Haan, 1977). De esta forma, los registros históricos de caudales de una corriente determinada constituyen un conjunto de muestras u observaciones desde la cual se pueden realizar inferencias acerca del comportamiento de dicha corriente en cualquier momento en el tiempo, es decir, *realizaciones no generadas aún por el sistema*.

En los procesos de caracterización, es imperativo que el conjunto de observaciones (muestras de la población) sea válido, esto es, que las observaciones correspondan a la sigma algebra establecida y que el número de observaciones sea representativo del proceso. Es importante aclarar aquí que al menos los procesos hidrológicos, no se cuenta con todas las observaciones que componen el dominio espacial de la variable de estado, por lo tanto, no es conocida su “verdadera” función de densidad de probabilidad.

De esta forma, las funciones de distribución y de densidad probabilística caracterizar de forma exhaustiva las magnitudes aleatorias (Haan, 1977; A. A. Sveshnikov, 1966), sin embargo, no todo el tiempo es posible definirla y por ello, frecuentemente, para describir este tipo de magnitudes se utiliza otro tipo de atributos numéricos que representan características particulares de las magnitudes aleatorias, estos son los momentos estadísticos de distinto orden (Domínguez & Kovalenko, 2009). El primero de ellos, se denomina esperanza matemática de la magnitud aleatoria y se expresa matemáticamente como:

$$m_x = \int_{-\infty}^{+\infty} x f(x) dx$$

Ecuación 25. Momento al origen de primer orden magnitud continua

La desviación de la magnitud aleatoria  $X$  con respecto a su esperanza matemática se llama magnitud aleatoria centralizada, y se representa como:

$$\dot{X} = X - m_x$$

Ecuación 26. Magnitud aleatoria centralizada

Por lo tanto, el momento central de orden  $k$  es la esperanza matemática de orden  $k$  de la magnitud aleatoria centralizada, para una magnitud continua:

$$\mu_k[X] = \int_{-\infty}^{+\infty} (x - m_x)^k f(x) dx$$

Ecuación 27. Momento central de orden  $k$  magnitud continua

El momento central de primer orden siempre es igual a cero. Los momentos estadísticos al origen están referenciados respecto de la posición del eje de la ordenada mientras que los momentos estadísticos centrales están referenciados al centro de gravedad de la curva de densidad probabilística. El momento central de segundo orden ( $k = 2$ ) se denomina varianza de la magnitud aleatoria ( $D_x$ ) y representa la dispersión de la variable alrededor de su esperanza matemática. El tercer momento central se denomina asimetría de la distribución probabilística de la magnitud aleatoria. Si la curva de densidad probabilística es asimétrica en relación a la posición de la esperanza matemática todos sus momentos estadísticos de orden impar serán nulos. Para caracterizar este momento se utiliza el primer momento central de orden impar no nulo, es decir,  $\mu_3$ . La forma adimensional de este momento se denomina coeficiente de asimetría. El cuarto momento central se denomina kurtosis, y explica el apuntalamiento de la curva de densidad (Domínguez, 2016).

## 2.7 PROCESOS ESTACIONARIOS Y NO ESTACIONARIOS

La estacionariedad hace referencia a la conservación de los momentos estadísticos de una variable aleatoria. En la realidad las series de tiempo que representan un proceso estocástico en general presentan problemas de estacionariedad, es decir, que al realizar la caracterización probabilística de las variables aleatorias que componen el sistema estocástico se suelen afectar cualquiera de sus propiedades básicas (momentos estadísticos).

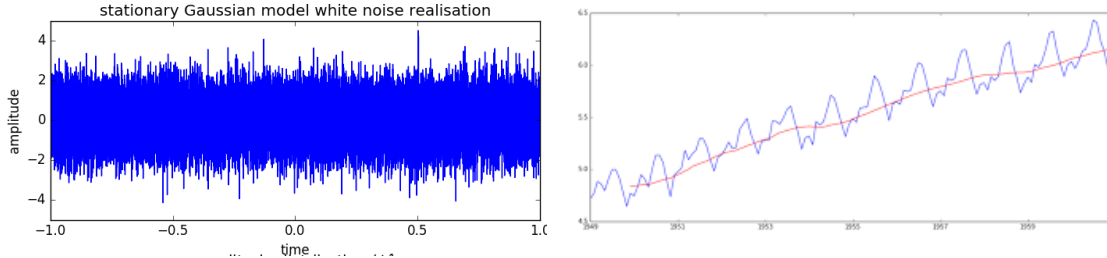


Figura 2. Visualización de un proceso estacionario (izquierda) y no estacionario (derecha)

Por lo tanto, un proceso estocástico es definido estacionario en el sentido estricto si al evolucionar la función de distribución que la caracteriza (Ruíz Abellón, 2017), esta no cambia en el tiempo, esto es:

$$F(X_{i_1}, X_{i_2}, \dots, X_{i_n}) = F(X_{i_1+j}, X_{i_2+j}, \dots, X_{i_r+j}) \forall \{i_1, i_2, \dots, i_r\} \bigwedge \forall j$$

Para el análisis, la condición anterior es bastante restrictiva y por lo tanto resulta práctico “debilitar” un poco la condición de estacionariedad, restringiendo la condición al sostenimiento constante de los dos primeros momentos estadísticos. Esto es:

- $\mu_t = \mu$ ,
- $\sigma_t^2 = \sigma^2$ ,
- $Cov(t, t + j) = Cov(s, s + j) = \gamma_j$ , en donde  $j = 0, \pm 1, \pm 2, \dots$

De acuerdo con otros libros la condición anterior recibe el nombre de estacionariedad en sentido amplio o estacionariedad de segundo orden. También, si sólo exigimos que el valor esperado sea constante se define que el proceso es estacionario de primer orden o en media. Por último, la estacionariedad en sentido débil no garantiza la estacionariedad estricta. La figura a continuación describe un proceso estacionario en el sentido débil o de segundo orden (cambio en media y varianza) para un proceso estocástico a escala diaria.

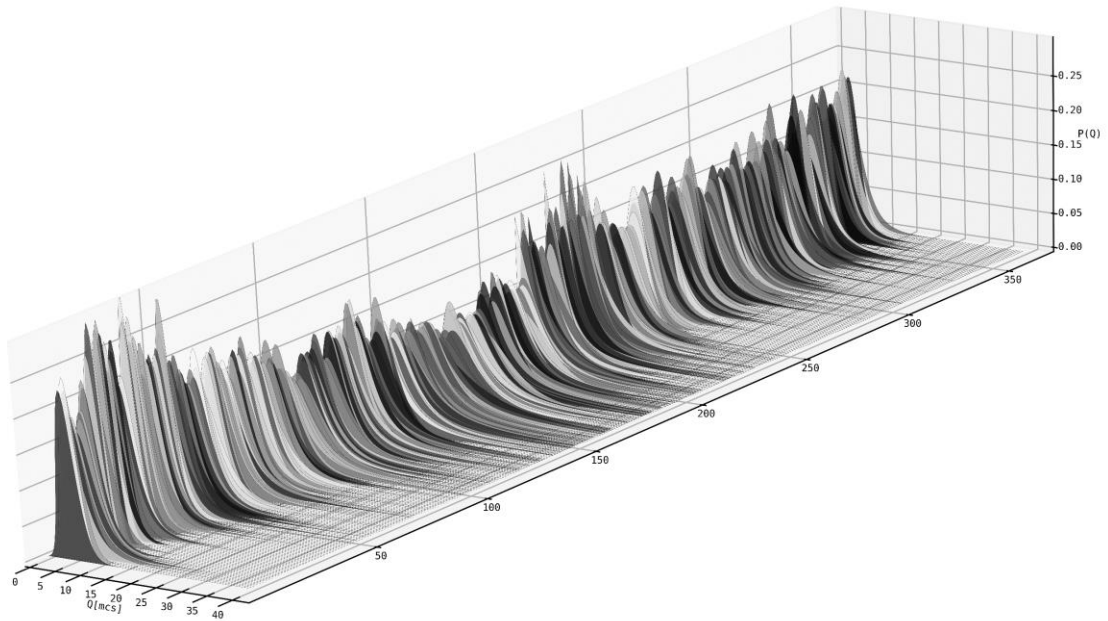


Figura 3. Proceso estocástico en sentido débil. Se puede apreciar el cambio en los valores esperados y varianzas de cada conjunto estadístico diario en un proceso hidrológico.

## 2.8 MODELACIÓN NUMÉRICA DE SISTEMAS DE COMPLEJIDAD ALGORÍTMICA SIGNIFICATIVA

Como se ha mencionado, un tema fundamental en el estudio de sistemas naturales (sistemas complejos) es la comprensión de los patrones o estructuras que definen la evolución del sistema. En la medida en que sea fácil o imposible el determinar los estados no observados de un sistema natural, o, dicho de otra forma, con la cantidad de incertidumbre que se definan los estados no observados de un sistema es posible generar una idea de la complejidad del sistema en estudio. A partir de este concepto de incertidumbre se puede desarrollar la definición de complejidad de un sistema. En función de la cantidad de información requerida para describir la evolución de un proceso aleatorio este será más o menos complejo (Angarita, 2008).

De acuerdo con Cover (Cover & Thomas, 2006), la complejidad puede medirse en términos informáticos como el número de bits que deben transmitirse para comunicar toda la secuencia sin que existan ambigüedades.



Figura 4. Complejidad algorítmica en modelación matemática. Imagen en medio tomado de <http://www.co.greene.pa.us/secured/gc2/images/conserv/what-is-watershed-action.gif>.

### 2.8.1 Complejidad Algorítmica Mínima

En sistemas simples, por ejemplo, el movimiento de un péndulo, la incertidumbre asociada a los estados futuros del sistema (ubicación del péndulo) es mínima, pues a partir de un conjunto de reglas matemáticas muy sencillas y el conocimiento del estado inicial del sistema será suficiente para estimar con gran precisión su estado futuro. En este caso, el observador será capaz en términos de programa de comprimir la señal del sistema en la solución de un sistema de ecuaciones diferenciales y su condición inicial, sin importar que el periodo de simulación tienda a infinito. A este tipo de sistema se le denomina de complejidad algorítmica mínima.

### 2.8.2 Complejidad Algorítmica Máxima

En el otro extremo, se encuentran aquellos sistemas en los cuales el observador no logra comprimir la señal del sistema, es decir, los estados observados son descritos por la misma cantidad de estados, como, por ejemplo, la secuencia de resultados de un par de dados.

### 2.8.3 Complejidad Algorítmica Significativa

En medio de los anteriores existen sistemas en los cuales la señal puede ser reducida en una cantidad menor al de los estados del mismo sistema en el periodo de evaluación. Esto es posible gracias a los componentes estructurales existentes en este tipo de sistemas, a través de estos el observador gana predictibilidad al estimar los estados futuros. A este tipo de sistema se le denomina de complejidad significativa, y la hidrología es una ciencia cuyos sistemas ocupan una posición en esta clase. En este contexto, además de los elementos estructurales existen elementos aleatorios que hacen que el sistema se aleje de ese patrón estructural y por lo tanto, se constituye en fuente de

incertidumbre dada la incapacidad del observador de establecer las reglas que determinan ese tipo de comportamiento aleatorio (Angarita, 2008).

## 2.9 ECUACIÓN DE LANGEVIN

Existen distintos enfoques para el estudio de los sistemas de complejidad algorítmica significativa, sin embargo, en todos los casos se requiere de una regla u operador matemático que describa la evolución del sistema.

La ecuación de Langevin ha sido ampliamente utilizada para la descripción de sistemas dinámicos estocásticos que tienen efectos de retroalimentación con su entorno (sistemas complejos). La ecuación de Langevin es una ecuación diferencial estocástica que describe la evolución probabilística de la variable de estado en el tiempo. De acuerdo con el marco conceptual expuesto este tipo de enfoque de modelación es apropiado para el objeto de estudio del presente trabajo de investigación, pues este operador combina un núcleo determinístico orientado a explicar el componente estructural del sistema hidrológico, mientras que el componente modulador del ruido en el operador matemático se orienta a explicar el componente aleatorio del sistema. Originalmente, Paul Langevin en 1908 desarrolló la ecuación para describir el movimiento Browniano (Gardiner, 2009) de una partícula en un líquido. Ahora, considerando una ecuación de este tipo con ruido blanco no gaussiano  $\xi(t)$ :

$$\dot{x} = f(x, t) + g(x, t)\xi(t)$$

Ecuación 28. Ecuación de Langevin

La solución de la ecuación anterior representa la evolución en el tiempo de la curva de densidad de probabilidad ( $\hat{x}$ ), dadas las condiciones iniciales y los moduladores del ruido en el sistema. En el contexto hidrológico, la ecuación de Langevin puede ser re-escrita como se muestra a continuación (Denisov, Horsthemke, & Hänggi, 2009):

$$\frac{dq(t)}{dt} = \psi[q(t), t] + \zeta[q(t), t]\xi(t)$$

Ecuación 29. Ecuación sobre-amortiguada de Langevin

En la ecuación anterior los términos  $\psi[q(t), t]$  y  $\zeta[q(t), t]$  son funciones deterministas que definen el sistema hidrológico y modulan el ruido  $\xi(t)$  respectivamente. El término  $\xi(t)$  es ruido blanco y representa las fluctuaciones del sistema.

Las ecuaciones del tipo Langevin (ecuaciones diferenciales estocásticas – SDE en inglés) son solucionadas por técnicas analíticas como el cálculo estocástico de Ito y/o la integral de Stratanovic, pero también es posible generar una ecuación equivalente como la ecuación de Fokker-Planck-Kolmogorov (FPK) (Denisov et al., 2009; Gardiner, 2009; Kumar & Narayanan, 2006). La solución analítica de la ecuación de FPK es limitada a casos específicos de sistemas lineales, sistemas escalares y algunos sistemas multidimensionales, desafortunadamente, para la mayoría de sistemas multidimensionales como las cuencas hidrográficas no existe una solución disponible (Domínguez & Rivera, 2010; Kumar & Narayanan, 2006). Sin embargo, muchos autores han mostrado la viabilidad de aproximaciones numéricas a la ecuación de FPK (Domínguez & Rivera, 2010;



Kozhevnikova, Shveikina, & Domínguez, 2012; Kumar & Narayanan, 2006; Masud & Bergman, 2005) con resultados aceptables.

Es importante resaltar que mientras que la ecuación de Langevin es una SDE que describe la evolución de una variable en el tiempo, la ecuación de FPK es una ecuación determinística que describe la evolución de la curva de densidad de probabilidad de la variable de estado en el tiempo. De esta forma, si el término de ruido de la ecuación de Langevin es producido por una función generadora de ruido es decir, es un proceso de difusión markoviano, entonces la curva de densidad de probabilidad y su evolución temporal puede representarse por la ecuación de FPK (Denisov et al., 2009; Gardiner, 2009; Yoon & Xu, 2010).

La derivación de la ecuación de FPK desde la ecuación sobre-amortiguada de Langevin es descrita en detalle en el trabajo de Denisov (Denisov et al., 2009; Jazwinski, 1966, 1970).

## 2.10 ECUACIÓN DE FOKKER-PLANCK-KOLMOGOROV (MFPK)

La ecuación de FPK, también conocida como la Ecuación Prospectiva de Kolmogorov (KFE, en inglés), fue introducida por Fokker (1914) y Planck (1917) para explicar el movimiento Browniano de partículas en la física moderna (Yoon & Xu, 2010). Posteriormente, Kolmogorov (Kolmogorov, 1931) describió analíticamente los fundamentos matemáticos en profundidad.

### 2.10.1 La axiomática de Kolmogorov

En 1931, el matemático soviético Andrey Nikolaevich Kolmogorov presentó sus axiomas en su trabajo Fundamentos de la teoría de probabilidades (Kolmogorov, 1931). Esta es basada en tres conceptos fundamentales y tres principales axiomas, como se presenta a continuación:

“Dado un conjunto  $\Omega$ , con elementos  $\omega$  que se denominan eventos/sucesos elementales, y un conjunto  $F$  de subconjuntos de  $\Omega$ . Los elementos de  $F$  se denominan sucesos aleatorios y  $\Omega$  es el espacio de sucesos elementales”

- I.  $F$ , es un algebra de conjunto ( $\sigma$  – algebra),
- II. Para cada subconjunto  $A \in F$  se contrapone una magnitud real no negativa  $p(A)$  que se denomina propiedad del evento/suceso  $A$ ,
- III.  $p(\Omega) = 1$ ,
- IV. Si  $A$  y  $B$  no se intersectan  $\rightarrow p(A + B) = p(A) + p(B)$ .

El conjunto de objetos  $[\Omega, F, p]$  que cumple con los axiomas anteriores es un campo de probabilidades, un espacio probabilístico.

### 2.10.2 Ecuación $n$ -dimensional de FPK

La ecuación de FPK es una ecuación diferencial parcial, determinística que utiliza la ley de conservación para describir la evolución temporal de la curva de densidad de probabilidad de un proceso aleatorio en sistemas  $n$ -dimensionales a través de los términos de advección y difusión.

Por lo tanto, la ecuación de FPK es un proceso difusivo que representa la densidad de probabilidad condicionada para cualquier sistema  $n$ -dimensional en diferentes momentos en el tiempo (Gardiner, 2009). Sean  $Q_i$  y  $Q_j$  dos variables aleatorias, entonces la ecuación de FPK toma la forma:

$$\frac{\partial p(Q_j, t|Q_i, t')}{\partial t} = - \sum_{i=1}^n \frac{\partial}{\partial Q_i} [A_i(Q_i, t)p(Q_j, t|Q_i, t')] + \frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2}{\partial Q_i \partial Q_j} [B_{ij}(Q_{ij}, t)p(Q_j, t|Q_i, t')]$$

Ecuación 30. Ecuación de FPK para un sistema  $n$ -dimensional

Donde  $A_i(Q, t)$  es el vector de advección,  $B_{ij}(Q, t)$  es la matriz de difusión y  $t'$  es el paso de tiempo anterior. Los términos de advección y difusión son funciones determinísticas que controlan el movimiento de la curva de densidad de probabilidad (PDF), el primero trasladando en el dominio espacial la curva, mientras el segundo apuntala o aplanla la curva de densidad.

La definición de la dimensionalidad de la ecuación de FPK radica (en el caso de este estudio) en un análisis de la memoria del proceso a través de la función de auto-correlación (caso endógeno) y correlación cruzada (caso exógeno). Como se describe más adelante, aquí el estudio se enfoca en procesos hidrológicos a resolución temporal diaria con memoria de orden 2, es decir, procesos hidrológicos cuyos valores de la función de auto-correlación dejan de ser significativos en el segundo rezago ( $\rho = 2$ ).

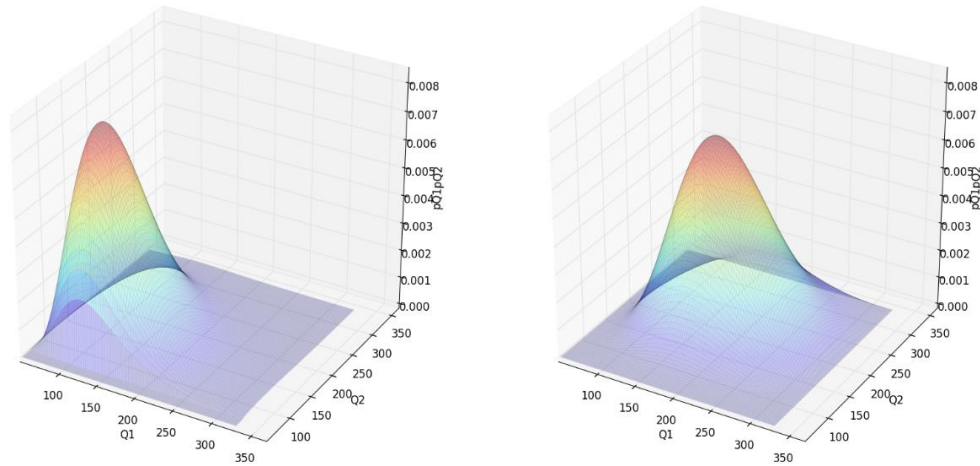


Figura 5. Evolución temporal de la PDF conjunta para dos variables aleatorias independientes. A la izquierda, la PDF inicial y a la derecha la PDF Final. Se puede observar la transformación de la PDF debido a los efectos de advección y difusión. En términos generales, el componente advectivo describe el comportamiento general del proceso, es decir la persistencia del sistema. Por otra parte, el componente difusivo describe las fluctuaciones o desviaciones de este comportamiento general.

Como ha mencionado el vector  $A_i(Q, t)$  y la matriz  $B_{ij}(Q, t)$  son funciones determinísticas, y sus parámetros son relacionados con características del sistema (Domínguez & Kovalenko, 2009), en este caso, variabilidad hidrológica, los parámetros hidrometeorológicos en la cuenca hidrográfica como por ejemplo la cantidad de agua precipitada y temperatura (factores externos), como también, las características morfométricas y del relieve, las coberturas y uso del suelo entre otros (factores internos). Este trabajo hace principal énfasis en establecer las relaciones existentes de los

parámetros de las funciones de advección y difusión con factores externos e internos, especialmente aquellos relacionados con la precipitación y las características morfométricas de las cuencas hidrográficas.

### 3 METODOLOGÍA

A continuación, se presentan los momentos metodológicos de este trabajo que contribuyeron al alcance de los objetivos. En general la metodología cumple con el protocolo de modelación planteado por Domínguez y Rivera (Domínguez, 2000; H. G. Rivera et al., 2008). La investigación consideró como dominio espacial las cuencas hidrográficas del territorio nacional que cumplieron con los criterios adoptados para el cumplimiento de los objetivos. Como base fundamental con relación al suministro de información hidrometeorológica, esta investigación recopiló y consolidó la información de series de tiempo de caudales medios diarios y precipitación total diaria de más de 700 estaciones hidrológicas y cerca de 4000 meteorológicas de la red nacional de monitoreo operada por el Instituto de Meteorología, Hidrología y Estudios Ambientales (IDEAM).

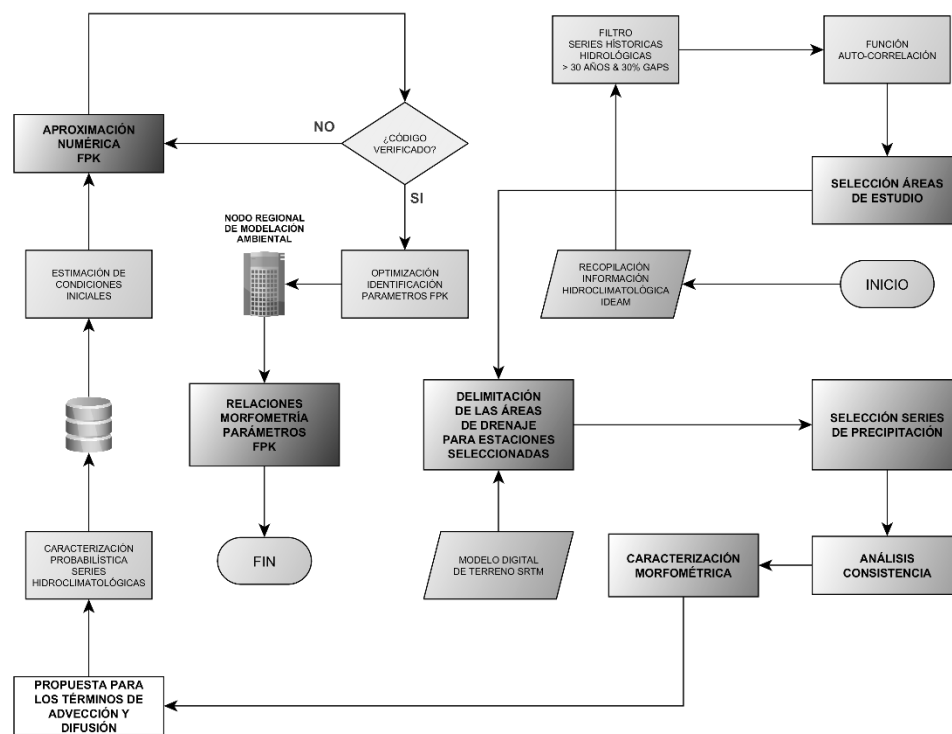


Figura 6. Diagrama metodológico desarrollado

#### 3.1 SELECCIÓN DE ÁREAS DE ESTUDIO Y FUENTES DE INFORMACIÓN

La red nacional de monitoreo hidrológica en Colombia está conformada por un conjunto de estaciones entre públicas operadas por el IDEAM y, otras (públicas y/o privadas) en donde resulta importante el papel de las corporaciones autónomas. En total la parte continental del país cuenta con cerca de 1233 puntos de monitoreo hidrológico entre la red pública y privada. La selección de las áreas de estudio se basó en las áreas de drenaje a los puntos de monitoreo hidrológico cuyos

registros históricos cumplieron con los criterios de calidad definidos para la presente investigación basado en lo dictado por la WMO. Estos son:

- a) Registro histórico con al menos 30 años de longitud (registros diarios medios de caudales),
- b) Completitud de la serie de al menos el 70% entre el primer y el último registro histórico,
- c) Series hidrológicas (caudales) con memoria de orden 2, esto es, que la función de autocorrelación tenga como máximo dos rezagos estadísticamente significativos.

Es importante recalcar la importancia de este último criterio, pues es aquí de acuerdo con el enfoque establecido (los forzadores del sistema serán variables endógenas, es decir, solo caudales) que se define la dimensionalidad del sistema representado en la dimensionalidad de la ecuación de FPK.

La fuente de información principal de este estudio es la información hidrometeorológica del IDEAM para el periodo histórico hasta 2014.

### 3.1.1 Determinación de la estructura del proceso Markoviano

Este paso metodológico define la suficiencia de la forma  $n$ -dimensional de la ecuación de FPK. Como se mencionó en los criterios de selección, se ha definido estudiar sistemas con memoria de orden 2.

La función de autocorrelación (ACF, en inglés) o dependencia secuencial es una herramienta estadística utilizada frecuentemente en el procesamiento de señales o series de tiempo (Bendat & Piersol, 2010) que permite determinar la estructura markoviana. La función de autocorrelación se define como la correlación cruzada de la señal consigo misma. Esta resulta de gran utilidad para encontrar patrones repetitivos dentro de una señal, como, por ejemplo, la periodicidad de una señal enmascarada bajo el ruido o para identificar la frecuencia fundamental de una señal que no contiene dicha componente, pero aparecen numerosas frecuencias armónicas de esta, los procesos estocásticos son ejemplos de autocorrelación. Su estimación se realizó de acuerdo con (Druzhinin & Sikan, 2001):

$$R_k = \frac{E[(x_i - \mu)(x_{i-k} - \mu)]}{\sigma_{x_i}^2 \sigma_{x_{i-k}}^2}$$

Ecuación 31. Función de autocorrelación (ACF)

Donde,  $E$  es el valor esperado,  $\mu$  es la media y  $k$  el desplazamiento temporal considerado (*lag o desfase*). La ACF varía en el rango de  $[-1, 1]$ , donde 1 y -1 indican correlación perfecta y anti-correlación perfecta, mientras  $R_k = 0$  indica independencia de entre  $X_i$  y su vector rezagado en  $k$  pasos de tiempo.

- Prueba de significancia

Para la estimación de la función del error estándar que define la significancia estadística de las magnitudes de la función de ACF, se aplicó la metodología descrita por Druzhinin and Sikan (Druzhinin & Sikan, 2001) así:

$$e = \frac{1 - r^2}{\sqrt{n - 1}} t_{1 - \frac{\alpha}{2}}$$

Ecuación 32. Prueba de significancia estadística

En la ecuación,  $n$  es el número total de observaciones para cada variable aleatoria,  $t$  es el valor crítico de la prueba de student con nivel de significancia  $\alpha$  y  $r$  es el coeficiente de correlación.

### 3.1.2 Delimitación de áreas de drenaje para las estaciones seleccionadas

El proceso de delimitación de cuencas hidrográficas se llevó a cabo mediante el procesamiento del modelo digital de terreno de 92 m. aproximadamente de la misión Shuttle Radar Topography Mission (SRTM) del Jet Propulsion Laboratory y NASA (NASA Jet Propulsion Laboratory (JPL), 2015), la hidrografía oficial nacional escala 1:100.000 derivada del Instituto Geográfico Agustín Codazzi (en adelante IGAC), el software de libre acceso GRASS GIS (Neteler, Bowman, Landa, & Metz, 2012) y por supuesto, los puntos de las estaciones seleccionadas definidas en el momento metodológico anterior.

Con fin de disminuir la incertidumbre en la delimitación de las áreas de drenaje, y posterior caracterización morfométrica, se realizó un proceso de ajuste de las corrientes derivadas del análisis digital con las encontradas en la hidrografía nacional mediante la metodología de quemado de corrientes (*stream burning*) (Jasiewicz & Metz, 2011), lo anterior obliga a que la red hidrográfica derivada del análisis digital coincida con la red hidrográfica real.

### 3.1.3 Selección de series de precipitación

Como se ha mencionado, esta investigación busca explicar las funciones de advección y difusión de la ecuación FPK a través de factores intrínsecos de los sistemas hidrológicos como por ejemplo la morfometría de las cuencas, y factores externos como las variables meteorológicas, por ejemplo, la precipitación. De esta forma, una vez definidas y delimitadas las áreas de estudio, se realizó una selección espacial de las estaciones de monitoreo de la precipitación presentes dentro de un búfer de 10 km de la cuenca hidrográfica en análisis. Inmediatamente, se aplicaron los criterios  $a$  y  $b$  descritos con anterioridad a las series de tiempo de precipitación presentes en el búfer. Por último, las series que cumplan dichos criterios serán tenidas en cuenta para los análisis posteriores.

## 3.2 ANÁLISIS DE CONSISTENCIA DE LA INFORMACIÓN

El análisis de consistencia de la información hidrometeorológica (caudales y precipitación) de las series seleccionadas de acuerdo a los criterios establecidos tuvo en cuenta las Guía de Prácticas Hidrológicas de la WMO y los estándares de calidad aplicables para Colombia (IDEAM, 2014). Los subsiguientes análisis se realizaron a los valores absolutos de las series de tiempo y a su análoga estandarizada.

### 3.2.1 Identificación de datos anómalos

Se realizó bajo tres metodologías, en primer lugar, se realizó la prueba de Grubbs ampliamente utilizada (Grubbs, 1950; Tietjen & Moore, 2012), esta es una prueba de student (prueba  $t$ ), donde se verifica que la diferencia absoluta entre el valor atípico y el promedio de la muestra dividida por la desviación estándar sea menor a 3 desviaciones. Donde para  $\alpha = 0,05$   $t_{\alpha} = 1,96$  (Domínguez, 2015a). Sin embargo, la prueba de Grubbs supone una distribución normal en los datos, por lo tanto, en series con distribuciones asimétricas puede no ser un buen criterio.

En segundo lugar, y teniendo en cuenta las restricciones del método de Grubbs se aplicó la metodología de las desviaciones absolutas en términos de medianas (MAD, median absolute deviation) (Leys, Ley, Klein, Bernard, & Licata, 2013; Rosenmai, 2013; Rousseeuw & Croux, 1993) en donde de manera similar a la prueba de Grubbs se busca identificar el número de desviaciones esta vez en términos de las medianas por derecha y por izquierda (cubriendo de esta forma distribuciones asimétricas) para cada conjunto estadístico (conjuntos de los días julianos).

En tercer lugar, se realizó la inspección visual mediante la evaluación de las gráficas de las series hidrometeorológicas en sus valores absolutos y estandarizados, identificando por criterio experto los valores potencialmente atípicos o anómalos.

Por último, se consideró como dato anómalo aquel en el cual convergieran al menos dos de las tres metodologías utilizadas. El dato etiquetado como anómalo es retirado de la serie y no es tenido en cuenta para los análisis siguientes.

### 3.3 CARACTERIZACIÓN MORFOMÉTRICA

Para cada una de las cuencas hidrográficas seleccionadas se aplicó la metodología descrita en el marco teórico de este documento [pág. 14], estimando así alrededor de 26 índices morfométricos e hidrográficos para cada cuenca. Posteriormente, se realiza un análisis de componentes principales (Jolliffe, Jolliffe, & Ian, 2014) para agrupar dichos índices en grupos ortogonales que permitan simplificar el análisis de la influencia de la morfometría en los procesos hidrológicos.

### 3.4 CARACTERIZACIÓN DEL PROCESO ESTOCÁSTICO EN LAS ÁREAS DE ESTUDIO

La caracterización del proceso estocástico implica el estudio de variables aleatorias de acuerdo a lo propuesto en la axiomática de Kolmogorov (Kolmogorov, 1931) anteriormente mencionada. Por lo tanto, se debe demostrar la aleatoriedad de las variables en análisis, en este caso, los caudales medios diarios.

#### 3.4.1 Pruebas de aleatoriedad

Para este propósito existente numerosas pruebas que determinan la aleatoriedad o no de un conjunto estadístico (Charmaine et al., 2001; Rukhin et al., 2010), sin embargo, para el presente estudio se han seleccionado 5 pruebas representativas:

- Runs test (Bradley, 1968),
- Wald Wolfowitz test (Lunneborg, Lunneborg, & E., 2014),
- Auto-correlation test (Druzhinin & Sikan, 2001),
- Serial test (Good, 1953) y,
- Gap test (Bofinger & Bofinger, 1961).

Una vez determinada la aleatoriedad de las variables, se continua con la aplicación de algunas técnicas estadísticas de caracterización que permiten describir el comportamiento de los conjuntos estadísticos de manera probabilística. Este paso metodológico busca capturar el componente estacional y aleatorio del concepto de variabilidad hidrológica descrito en el marco conceptual. Las técnicas estadísticas aplicadas para la caracterización estocástica consistieron entonces en el ajuste de funciones de densidad de probabilidad teóricas y posteriormente la estimación de momentos al origen y centrales.

Las pruebas aquí descritas aplican de acuerdo con la axiomática adoptada a conjuntos estadísticos válidos, es decir, la medición de la variable aleatoria bajo las mismas condiciones del experimento. La escala de análisis de la presente investigación se ha definido como diaria (alta resolución temporal), esto es, en términos de la axiomática el conjunto  $F$ , un sub-conjunto de sucesos aleatorios (realizaciones del sistema) de  $\Omega$  (espacio de sucesos elementales).  $F$  es la  $\sigma$ -álgebra. Las características anteriores permiten configurar nuestro sistema como un proceso estocástico y por lo tanto, aplicar la teoría moderna de probabilidades con sus respectivos beneficios, entre ellos la aplicación del modelo de FPK (M. E. Rivera, Correa H., Avendaño B., Rivera, & Fuentes B., 2015). En resumen, los conjuntos estadísticos a los cuales se aplicó las técnicas comentadas equivalen a los conjuntos de registros hidrológicos diarios indexados en el rango  $[1, 2, 3, \dots, 364, 365]$  (días julianos) como se presenta en la Figura 7.

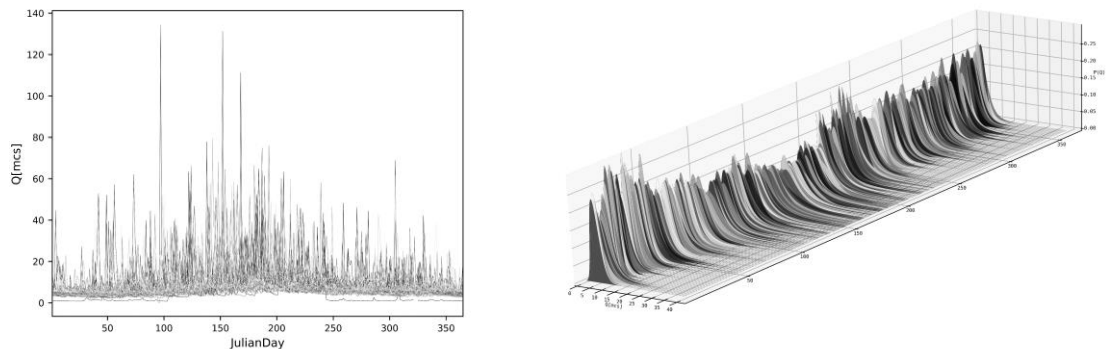


Figura 7. Ramillete de realizaciones [izquierda] y curvas de densidad de probabilidad de un proceso aleatorio: caudales medios diarios en la estación Pitalito 2 Automática [21017050].

En este caso, el sistema lo constituyen 365 conjuntos de variables aleatorias indexados en el tiempo que representan los 365 días del año (1 de enero a 31 de diciembre):

$$Q_t = \{Q_1, Q_2, \dots, Q_{365}\}$$

En forma estricta, en esta escala temporal se presenta un conjunto especial del sistema cada 4 años, en donde a la  $\sigma$ -álgebra se le agrega un sub-conjunto más  $Q_{366}$  el cual se genera por la consideración del día 29 de febrero. Sin embargo, en la generalidad del proceso estocástico esta



realización no cuenta con el mismo número de elementos que los demás sub-conjuntos, por lo tanto, la caracterización probabilística pierde robustez. Por la razón anterior, este sub-conjunto [29 de febrero] no fue incluido en la presente investigación y puede ser tema para futuras propuestas alrededor de procesos estocásticos en hidrología.

### **3.4.2 Ajuste de función de densidad de probabilidad**

La prueba de ajuste consiste en corroborar u objetar la tesis acerca de que la distribución empírica (datos observados) concuerda con una distribución teórica. La prueba de hipótesis utilizada en este estudio es la propuesta por Kolmogorov – Smirnov (Mathematics, 2012), y como criterios de selección último, en caso de que varias distribuciones aprueben el test de Kolmogorov-Smirnov, son el error medio y el error máximo.

Para variables hidroclimáticas se han probado 12 distribuciones teóricas dentro de las cuales se encuentran la normal, log-normal, exponencial, gamma, log-gamma, generalizada de extremos, Weibull máxima, Weibull mínima, Gumbel sesgada a la izquierda, Gumbel sesgada a la derecha, gamma generalizada y potencial.

Al ajustar un conjunto válido de datos observados a una distribución teórica, permite aproximarse al universo completo de magnitudes, con sus respectivas probabilidades, que puede tomar nuestra variable aleatoria (Haan, 1977; A. A. Sveshnikov, 1966). Además, permite establecer parámetros del comportamiento del proceso más cercanos a la realidad, estos parámetros se conocen como momentos estadísticos.

### **3.4.3 Estimación de momentos**

Una vez se comprobó la aleatoriedad de los conjuntos estadísticos y se determinó el mejor ajuste de distribución teórico, se aplicaron las metodologías descritas en el marco teórico para la estimación del primer momento al origen y segundo momento cruzado central [pág. 28].

## **3.5 APROXIMACIÓN NUMÉRICA DE LA ECUACIÓN 2D DE FOKKER-PLANCK-KOLMOGOROV**

Esta investigación busca generar nuevo conocimiento entorno a los factores físicos (internos y externos) que gobiernan la variabilidad hidrológica en las cuencas hidrográficas en escalas temporales finas (diarias). Para tal fin se han aplicado técnicas de modelación probabilística como herramienta cognitiva (Chater, Tenenbaum, & Yuille, 2006) para el entendimiento de sistemas complejos como lo son las cuencas hidrográficas.

El proceso de modelación matemática, en general, implica el desarrollo de tres etapas conocidas como la trilogía modelo – algoritmo – programa de acuerdo a Samarsky & Mikhailov (Domínguez, Ardila, & Bustamante, 2010).

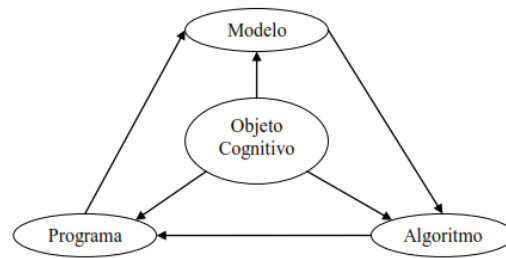


Figura 8. Trilogía Modelo – Algoritmo – Programa  
Fuente: Samarsky & Mikhailov, 1997 modificado por Domínguez 2.010

En este caso el objeto cognitivo (la cuenca hidrográfica) es reemplazada por un supra operador matemático equivalente de tipo determinístico dinámico distribuido como lo es la ecuación  $n$ -dimensional de FPK, y para el caso específico de esta investigación con dimensionalidad  $n=2$ .

Una vez definidos, el objeto de modelación, el operador, y de acuerdo con el protocolo de modelación matemática (Domínguez, 2000; H. G. Rivera et al., 2008), se procede con la selección y/o el desarrollo del algoritmo de cálculo que permita implementar el modelo en un ordenador.

Atrás se ha mencionado [en la página 32] que la viabilidad de una solución analítica para la ecuación diferencial propuesta es limitada a ciertos casos, y, por lo tanto, se requiere una aproximación a la solución mediante esquemas numéricos. En trabajos anteriores varios autores han reportado diferentes esquemas numéricos de solución para la ecuación de FPK  $n$ -dimensional, básicamente para una y dos dimensiones (Denisov et al., 2009; Domínguez & Rivera, 2010; Kumar & Narayanan, 2006; Kushner, 1976; Pawula, 1967; Rosmann, 2014; Sjöberg, 2005; Wojtkiewicz & Bergman, 2000), ya que para una mayor dimensionalidad, aunque se han planteado esquemas, el costo computacional ha sido la principal limitación (Masud & Bergman, 2005). En estos trabajos se han aproximado la solución de la ecuación de FPK mediante métodos numéricos de bajo orden como diferencias y elementos finitos con las desventajas que estos métodos conllevan. Recientemente, en 2016, Reyes (Reyes Gil, 2016) aplicó métodos discontinuos de alto orden para aproximar la ecuación 1D de FPK con el objetivo de eliminar la difusión numérica presente en los métodos en diferencias finitas. Sin embargo, la misma autora en sus conclusiones menciona que el efecto de la difusión numérica parece ser visible solo en los casos en donde la traslación/difusión es muy fuerte. También alerta del alto costo computacional de los métodos de alto orden en relación con las diferencias finitas. Por lo anterior, se definió utilizar el esquema de diferencias finitas explícitas debido a que la relación costo computacional/difusión numérica favorece los métodos de bajo orden.

El esquema numérico por diferencias finitas explícitas ha sido aplicado por algunos escasos autores en áreas diferentes a la hidrología como se ha referenciado arriba. En el campo de la hidrología en Colombia, el esquema ha sido aplicado para los casos unidimensional (Domínguez, 2001; Domínguez & Kovalenko, 2009; Domínguez & Rivera, 2010; Reyes Gil, 2016; Rosmann, 2014) con resultados aceptables. Por otra parte, para el caso bidimensional de la ecuación de FPK algunos autores han propuesto esquemas de aproximación por diferencias finitas explícitas (Masud & Bergman, 2005; Wojtkiewicz & Bergman, 2000), sin embargo, en dichos esquemas numéricos el término difusivo no contempla las derivadas cruzadas y por lo tanto, no aplica para nuestro caso de estudio.

$$\frac{1}{2} \sum_{i=1}^n \sum_{j=1}^n \frac{\partial^2}{\partial Q_i \partial Q_j} [B_{ij}(Q, t) p(Q_j, t | Q_i, t')]$$

Ecuación 33. Término difusivo de la ecuación de FPK. Cuando  $i \neq j$ , la derivada de segundo orden es de tipo cruzada o mixta.

El esquema de solución numérica en diferencias finitas para derivadas cruzadas de segundo orden aplicado a la ecuación de FPK se puede apreciar en el trabajo de Kushner (Kushner, 1976) para procesos condicionados. También Domínguez y Rivera (Domínguez & Rivera, 2010) proponen un método bidireccional en diferencias finitas. El método consiste en un esquema up-wind para la derivada espacial de primer orden (termino advectivo) y un esquema por diferencias centrales para la derivada espacial de segundo orden. Además, el esquema de aproximación numérico incluye un par de términos  $(\sigma, \varphi)$ . El primero  $[\sigma]$ , para la selección del esquema totalmente explícito o implícito. El segundo  $[\varphi]$ , para cambiar el esquema up-wind de acuerdo con la traslación del sistema.

Finalmente, el esquema numérico propuesto para la ecuación bidimensional de FPK presentada en la Ecuación 30 para  $n = 2$  es una combinación de las propuestas de Domínguez y Kushner como se presenta a continuación:

$$\begin{aligned} \frac{p_{i,j}^{k+1} - p_{i,j}^k}{\Delta t} = & - \left[ \frac{\varphi_L (A_{1i+1,j}^k p_{i+1,j}^k - A_{1i,j}^k p_{i,j}^k)}{\Delta Q_1} + \frac{\varphi_R (A_{1i,j}^k p_{i,j}^k - A_{1i-1,j}^k p_{i-1,j}^k)}{\Delta Q_1} \right. \\ & + \left. \frac{\varphi_L (A_{2i,j+1}^k p_{i,j+1}^k - A_{2i,j}^k p_{i,j}^k)}{\Delta Q_2} + \frac{\varphi_R (A_{2i,j}^k p_{i,j}^k - A_{2i,j-1}^k p_{i,j-1}^k)}{\Delta Q_2} \right] \\ & + \frac{1}{2} \left[ \frac{B_{11i+1,j}^k p_{i+1,j}^k - 2B_{11i,j}^k p_{i,j}^k + B_{11i-1,j}^k p_{i-1,j}^k}{\Delta Q_1^2} \right. \\ & + \frac{B_{22i,j+1}^k p_{i,j+1}^k - 2B_{22i,j}^k p_{i,j}^k + B_{22i,j-1}^k p_{i,j-1}^k}{\Delta Q_2^2} \\ & + \frac{B_{12i+1,j+1}^k p_{i+1,j+1}^k - B_{12i+1,j-1}^k p_{i+1,j-1}^k - B_{12i-1,j+1}^k p_{i-1,j+1}^k + B_{12i-1,j-1}^k p_{i-1,j-1}^k}{4\Delta Q_1 \Delta Q_2} \\ & \left. + \frac{B_{21i+1,j+1}^k p_{i+1,j+1}^k - B_{21i+1,j-1}^k p_{i+1,j-1}^k - B_{21i-1,j+1}^k p_{i-1,j+1}^k + B_{21i-1,j-1}^k p_{i-1,j-1}^k}{4\Delta Q_2 \Delta Q_1} \right] \end{aligned}$$

Ecuación 34. Esquema numérico en diferencias finitas explícitas para la aproximación de la ecuación de FPK bidimensional.

Donde,

$A = \begin{bmatrix} A_1 \\ A_2 \end{bmatrix}$  son los vectores de advección para cada dimensión,

$B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix}$  es la matriz de difusión para cada dimensión,

$i, j$  son los descriptores de la discretización espacial,

$\varphi_R, \varphi_L$  son los pesos direccionales para el esquema bidireccional de advección, si  $A_{n,i,j}^k < 0 \Rightarrow \varphi_R = 0, \varphi_L = 1$ , de otra forma,  $\varphi_R = 1, \varphi_L = 0$ ,

$k$  es el descriptor temporal,

$\Delta Q_i$  es la discretización espacial,

De acuerdo con el esquema numérico seleccionado, es imperativo establecer la condición de estabilidad. Para tal fin se utilizó el número de Péclet, el cual es una relación entre el proceso advectivo y el difusivo, y trata de establecer cuál de los dos procesos domina la evolución de la curva de densidad.

$$P_e = \frac{[|A| * \Delta Q]_{max}}{|B|_{max}}$$

Ecuación 35. Número de Péclet tomado de (Domínguez & Rivera, 2010)

En los casos donde la evolución es dominada por el proceso advectivo, esto es  $P_e > 3$  de acuerdo con el Danish Hydraulic Institute (DHI, 1998), la condición de estabilidad dada por la condición de Courant-Friedrichs-Lewy (en adelante CFL) fue definida de acuerdo a Domínguez y Rivera (Domínguez & Rivera, 2010), así:

$$\frac{\Delta q}{\Delta t} > |A|_{max}$$

Ecuación 36. Condición CFL para procesos dominados por traslación tomado de (Domínguez & Rivera, 2010)

De otra forma ( $P_e \leq 3$ ), se considera que el proceso difusivo es importante y es el que domina la transición de la curva, por lo tanto, la condición de estabilidad se estimó como:

$$|B|_{max} * \frac{\Delta t}{\Delta q^2} < \frac{1}{4}$$

Ecuación 37. Condición CFL para procesos dominados por difusión

A partir de la condición de estabilidad se estableció la discretización temporal del esquema, asegurando que se cumplan las desigualdades.

### 3.5.1 Condiciones iniciales y de frontera

Es importante recordar que la variable de estado en el presente ejercicio de modelación matemática es la curva de densidad de probabilidad condicionada (PDF Condicionada) de los caudales medios diarios de acuerdo con lo descrito en la Ecuación 30. Teniendo en cuenta la configuración del proceso estocástico, en la  $\sigma$ -álgebra aplicada aquí se construyeron 365 conjuntos estadísticos indexados en el tiempo por el día juliano  $F = \{1, 2, 3, \dots, 364, 365\}$ . Luego del análisis de consistencia de la información, remoción de datos anómalos y caracterización probabilística, se definió como condición inicial la curva de densidad de probabilidad condicionada empírica, la cual se estimó como se expone a continuación.

$$p(Q_2|Q_1) = \frac{p(Q_2, Q_1)}{p(Q_1)}$$

Ecuación 38. Densidad de probabilidad condicionada en un sistema bidimensional.

Donde  $p(Q_1, Q_2)$  es la probabilidad conjunta del sistema y  $p(Q_1)$  es la probabilidad marginal de  $Q_1$ . La aproximación a la probabilidad condicionada  $p(Q_2|Q_1)$  se realizó mediante la discretización de las variables  $[Q_1, Q_2]$  y posterior conteo del número de eventos que sucedieron dentro de un rango determinado (evento que condicionada). El conteo se realiza como se describe a continuación en función de las características de los rangos de las variables, así:

- Si los rangos son monótonos crecientes, entonces el rango es semi-abierto por derecho,  $bins_{i-1} \leq evento < bins_i$ , donde  $bins$  son los límites por izquierda y derecha de los rangos en función de la discretización realizada.
- Si los rangos son monótonos decrecientes, el rango es semi-abierto por izquierda,  $bins_{i-1} > evento \geq bins_i$ .

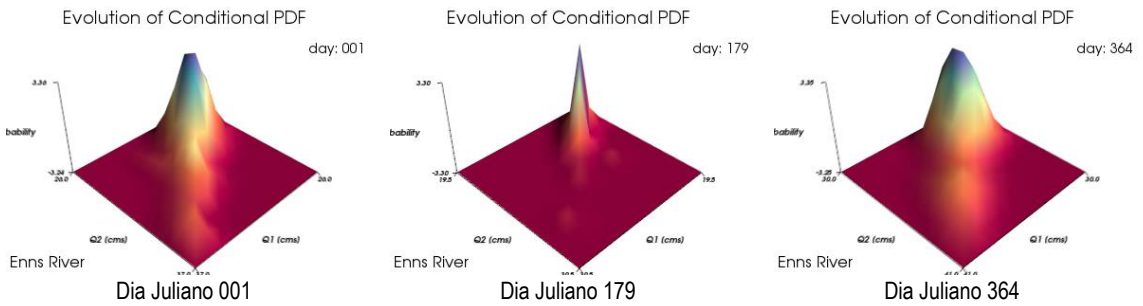


Figura 9. Curva de Densidad de Probabilidad Condicionada Empírica bidimensional

Se debe revisar que la sumatoria de los eventos condicionados sea igual al total de los eventos, y, que el número de eventos condicionados sea independiente de la discretización espacial realizada.

Adicionalmente, la condición de frontera seleccionada para el modelo es de tipo absorbente, esto significa que los valores de la variable de estado (PDF condicionada) para cualquier punto espacial en las fronteras o por fuera del dominio espacial de modelación es cero.

$$p(Q_2|Q_1) = 0 \forall q = \alpha, \beta$$

Ecuación 39. Definición de las condiciones de frontera

Donde  $\alpha, \beta$  son los valores límite del dominio de modelación. Sin embargo, y anticipando posibles situaciones de no conservación de masa debido a los procesos de traslación entre una transición y su siguiente, el dominio de modelación, es decir, sus valores límite ( $\alpha, \beta$ ) fueron ajustados en función de los valores extremos (y un poco más, alrededor de un +/- 20% más de sus valores extremos) de cada conjunto estadístico involucrado en cada transición, técnica que Yoon denomina dominio móvil (Yoon & Xu, 2010). La anterior técnica, además de disminuir el riesgo de pérdida de masa de la curva entre transición, mejora el costo computacional del modelo, pues se eliminan gran cantidad de nodos y, por lo tanto, el tiempo de ejecución disminuye de acuerdo con el esquema numérico explícito.

Las condiciones de frontera de este tipo (absorbentes) pueden ser explicadas a través de la axiomática de Kolmogorov y la técnica de dominio móvil implementada. De acuerdo con los registros históricos y la configuración de los conjuntos de caudales diarios que intervienen en una transición de la curva de densidad, se contraponen magnitudes reales no negativas a cada uno de los eventos registrados, esto es las probabilidades de los eventos, por lo tanto, al ampliar el dominio de modelación se genera un espacio probabilístico en donde no se han registrado eventos y, consecuentemente, las probabilidades en este espacio ampliado serán ceros justificando el uso de condiciones absorbentes.

### **3.5.2 Desarrollo y verificación del código**

Una vez definidas las unidades de análisis, realizado todo el tratamiento de los datos, la caracterización probabilística de los sistemas y construido el esquema numérico, se consolidó el anterior esfuerzo en un algoritmo de tal forma que pueda iniciarse el proceso de identificación de los parámetros de la ecuación de FPK.

Las herramientas computacionales utilizadas en el desarrollo de esta investigación hacen parte del mundo de licenciamiento gratuito. La herramienta principal y sobre la cual se realizaron distintos desarrollos es el lenguaje de programación científica Python (Van Rossum, 1991) sobre la cual se realizó el tratamiento de la información mediante sus módulos numéricos, estadísticos y de generación de gráficos avanzados como Numpy, SciPy, Matplotlib, Mayavi, y por su puesto para el manejo de series de tiempo, Pandas.

De manera especial, para el tratamiento de datos geográficos, y se resalta el proceso de caracterización morfométrica de cuencas hidrográficas, en el cual, se construyeron módulos mediante el uso de librerías especializadas como GDAL (GDAL Development Team, 2015) para el cálculo del total de indicadores descritos en el marco teórico, “el poder de los sistemas de información geográfica sin el uso de plataformas SIG”.

Finalmente, se realizaron pruebas al algoritmo que implementa un conjunto de rutinas desde la lectura de los datos hasta la generación de las salidas gráficas de la evolución espacio-temporal del campo de probabilidades condicionadas con curvas de densidad normales bivariados asumiendo independencia entre las variables para comprobar errores y estabilidad del esquema numérico.

Sin embargo, dado que en esta investigación no se identificó una solución analítica a la ecuación de FPK bidimensional que conservara los momentos de correlación cruzada en los términos difusivos, es decir, que la solución mantuviera las diferencias de segundo orden mixtas y, que varios autores (Domínguez & Rivera, 2010; Domínguez, Rosmann, & Chavarro, 2014; Kumar & Narayanan, 2006; Masud & Bergman, 2005; Reyes Gil, 2016; Sjöberg, 2005) han demostrado la viabilidad de los esquemas numéricos en diferencias finitas, se asume que la aproximación numérica implementada es aceptable.

### **3.5.3 Optimización de parámetros**

El esquema general de modelación matemática adaptado por Domínguez y Kovalenko (Domínguez & Kovalenko, 2009) fue adoptado en esta investigación para la implementación del operador

matemático, ecuación de FPK. Según el mismo esquema (ver Figura 10) se pueden configurar varios tipos de problema como sea la incógnita dentro del esquema (solución de problemas inversos).

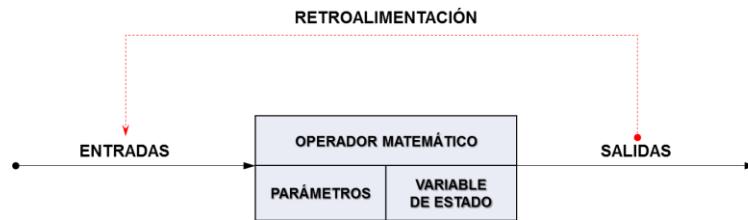


Figura 10. Esquema general de modelación matemática  
Tomado y adaptado de: (Domínguez & Kovalenko, 2009)

El objetivo en este caso es determinar las estructuras matemáticas de traslación y difusión que operadas en la ecuación de FPK reproduzcan de la mejor manera la variabilidad hidrológica del sistema condicionado definido por el conjunto de datos observados y simulados, esto es conocido como solución del problema inverso de modelación tipo 3. Para lograr este importante momento metodológico, es necesario definir dos aspectos:

- El algoritmo de optimización y,
- La función objetivo.

El primer ítem fue despejado utilizando las librerías disponibles para procesos de optimización matemático en Scipy, una de las librerías de acceso libre disponibles para Python. Dentro de Scipy, en los paquetes de optimización se encuentran una gran variedad de tipos de algoritmos para procesos de identificación de mínimos, máximos y raíces de funciones. Sin embargo, solo algunos paquetes permiten al usuario establecer o definir la función objetivo a minimizar (o maximizar), aspecto que se considera debe estar en el dominio del modelador. Por lo tanto, el paquete “minimize” fue seleccionado por dos razones fundamentales:

- Permite la construcción de la función objetivo,
- El paquete permite alternar bajo la misma estructura varios algoritmos para la minimización de la función escalar, como, por ejemplo: Nelder-Mead, Powell, Broyden, entre otras.

De esta manera, se seleccionó el método *BFGS* (Nocedal & Wright, 2006) que aplica el método Cuasi-Newton de Broyden, Fletcher, Goldfarb y Shanno y utiliza solo la primera derivada de la función objetivo, este método en experiencias previas del autor ha mostrado un buen desempeño.

La función objetivo fue seleccionada a partir de un conjunto de métricas entre las cuales están el RMSE (Root Mean Square Error), el R4MS4E (Fourth Root Mean Quadrupled Error), y la distancia de Kullback-Leibler.

Finalmente, el algoritmo implementado elimina el uso de los tradicionales ciclos para la solución del esquema numérico y ejecuta operaciones vectoriales lo cual acelera cerca de 160 veces la ejecución del esquema propuesto para la ecuación de FPK bidimensional.

### 3.6 IDENTIFICACIÓN DE ESTRUCTURA MATEMÁTICA PARA LOS TÉRMINOS DE ADVECCIÓN Y DIFUSIÓN

Como se mostró anteriormente en la Ecuación 28 (Ecuación de Langevin), las funciones  $f(x, t)$  y  $g(x, t)$  son funciones determinísticas que describen el fenómeno en estudio. También se ha expuesto que al reemplazar dichas funciones por ecuaciones diferenciales ordinarias y/o parciales (que generalmente representan sistemas complejos) la solución por métodos analíticos es limitada y/o complicada como por ejemplo el cálculo estocástico de Ito y/o la Integral de Stratonovich.

Una alternativa, para la aproximación a los núcleos deterministas de traslación y difusión es la solución del problema inverso de la ecuación de FPK, en donde mediante técnicas de optimización matemática es posible encontrar funciones aproximadas a los núcleos deterministas  $f(x, t)$  y  $g(x, t)$ .

En la Ecuación 30, las funciones deterministas análogas a las de Langevin, son  $A(q, t)$  y  $B(q, t)$ . En esta investigación  $A(q, t)$  y  $B(q, t)$  se han considerado como funciones polinómicas de alto orden de acuerdo a resultados en trabajos previos (Domínguez & Rivera, 2010; Reyes Gil, 2016; Rosmann, 2014), sin embargo, este aspecto y la evaluación de varios operadores puede ser objeto de investigaciones futuras. De forma simplificada las funciones propuestas para los términos de traslación y difusión son:

$$A_1(q, t) = k_1q + k_2q^2 + k_3$$

$$A_2(q, t) = k_4q + k_5q^2 + k_6$$

$$B_{11}(q, t) = g_1q + g_2q^2 + g_3$$

$$B_{22}(q, t) = g_4q + g_5q^2 + g_6$$

$$B_{12}(q, t) = x_1q + x_2q^2 + x_3$$

$$B_{22}(q, t) = y_1q + y_2q^2 + y_3$$

La forma de las funciones fue diseñada de tal manera que representen la naturaleza lineal del fenómeno de advección y la no linealidad de la difusión tal y como se muestra en los términos diferenciales de la ecuación de la FPK. De acuerdo con este supuesto, se configuran conjuntos de parámetros objeto del proceso de optimización, así:

$$\{k_1, k_2, k_3, k_4, k_5, k_6, g_1, g_2, g_3, g_4, g_5, g_6, x_1, x_2, x_3, y_1, y_2, y_3\}$$

Los cuales además son función del tiempo en la  $\sigma$ -álgebra establecida (365 sub-conjuntos), es decir, que al finalizar el proceso de optimización se habrán identificado 6570 parámetros que describen las 365 transiciones para cada unidad de estudio.

### 3.7 RELACIÓN DE PARÁMETROS ÓPTIMOS CON FACTORES EXTERNOS E INTERNOS DEL SISTEMA

La variabilidad en las condiciones climáticas y biofísicas del país, se refleja en la diversidad hidrológica y morfométrica de sus cuencas hidrográficas (Burbano-Girón, Domínguez, & Barón-



Ruiz, 2016). A su vez, las características morfométricas representan con claridad las características intrínsecas de cada cuenca hidrográfica (Baker, Kochel, & Patton, 1988; Biswas, Das Majumdar, & Banerjee, 2014; Patton & Baker, 1976).

Para abarcar este punto metodológico, se aplicaron regresiones lineales múltiples y simples, de manera adictiva y sustractiva entre los 22 índices morfométricos (se eliminaron los índices relacionados con las tasas de bifurcación de la red hidrográfica) y los 18 parámetros óptimos de la ecuación de FPK. Además, entre los valores esperados y varianzas de las magnitudes de precipitación total diaria antecedente en las cuencas de estudio y los parámetros óptimos de la ecuación de FPK, esto, con el fin de identificar los factores internos y externos que explican que pueden explicar la evolución de las curvas de densidad de probabilidad condicionadas.

Como resultado se espera obtener funciones que describan una a una las transiciones de las curvas de densidad condicionada en el conjunto de los 365 días julianos, es decir, que estas funciones serán dinámicas en principio.

### 3.7.1 Regresiones lineales múltiples

El análisis de regresión lineal es una técnica estadística para investigar y modelar la relación entre variables, para lo cual se construye la ecuación lineal que relaciona una variable aleatoria respuesta (dependiente)  $y$ , con una o múltiples variables predictoras o regresoras (independientes)  $x_i$ , y se estiman los parámetros desconocidos  $\beta_i$ .

La ecuación que representa el modelo se presenta a continuación:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \dots + \beta_k x_k + \varepsilon$$

Ecuación 40. Regresión lineal múltiple

Donde:

$y$	variable aleatoria de respuesta (dependiente)
$x_i$	variables predictoras o regresoras ortogonales (independientes)
$\beta_i$	parámetros del modelo
$\varepsilon$	error del modelo

### 3.7.2 Regresiones lineales simples

Al modelo en el cual se cuenta con un solo regresor se le conoce como regresión lineal simple y su representación gráfica es una línea recta.

$$y = \beta_0 + \beta_1 x + \varepsilon$$

Ecuación 41. Regresión lineal simple

En esta ecuación, a  $\beta_0$  y  $\beta_1$  se les conocen como coeficientes de la regresión, y son constantes desconocidas que usualmente tienen una interpretación simple y útil en el modelo.  $\varepsilon$  es un componente aleatorio del error. Se supone que los errores tienen promedio cero y varianzas constante desconocida ( $\sigma^2$ ), además, se supone que los errores no están correlacionados.

Se considera que el regresor  $x$  está controlado y se puede medir con error prácticamente despreciable, mientras que la respuesta  $y$  es una variable aleatoria. De este modo, hay una distribución de  $y$  para cada posible valor de  $x$ . La media de esta distribución se presenta a continuación:

$$E(y|x) = \beta_0 + \beta_1 x$$

y la varianza es:

$$\text{Var}(y|x) = \text{Var}(\beta_0 + \beta_1 x + \varepsilon) = \sigma^2$$

Así, la medida de  $y$  es una función lineal de  $x$ , aunque la varianza de  $y$  no depende del valor de  $x$ .

### 3.7.3 Estimación de parámetros

La estimación de los parámetros se realiza a través de los datos de la muestra, provenientes de un experimento controlado, diseñado en forma específica para recolectarlos, o en un estudio observacional, o a partir de registros históricos existentes (estudio retrospectivo). El método empleado para la estimación de los valores de  $\beta_0$  y  $\beta_1$  es mínimos cuadrados, en el cual se busca que la suma de los cuadrados de las diferencias entre las observaciones  $y_i$  y la línea recta sea mínima. Las ecuaciones que solucionan el sistema de mínimos cuadrados se presentan a continuación:

$$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n y_i x_i - \frac{(\sum_{i=1}^n y_i)(\sum_{i=1}^n x_i)}{n}}{\sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}}$$

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i$$

$$\bar{y} = \frac{1}{n} \sum_{i=1}^n y_i$$

El modelo ajustado de regresión simple es, entonces:

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

La ecuación anterior produce un estimado puntual de la media de  $y$  para una determinada  $x$ .

La diferencia entre el valor observado  $y_i$  y el valor ajustado correspondiente  $\hat{y}_i$  se llama residual. Los residuales tienen un papel importante para saber si el modelo de regresión ajustado es adecuado, y para detectar diferencias respecto a las hipótesis básicas. La ecuación para el cálculo del  $i$ -ésimo residual se presenta a continuación:

$$e_i = y_i - \hat{y}_i = y_i - (\hat{\beta}_0 + \hat{\beta}_1 x)$$

### 3.7.4 Estimación del error estándar de la regresión

Además de los parámetros, se requiere estimar el error estándar  $\sigma$  de la regresión para probar hipótesis y formar intervalos estimados pertinentes al modelo de la regresión. El valor estimado de  $\sigma^2$  se obtiene a partir de la suma de cuadrados de residuales, o suma de cuadrados de error:

$$SSE = \sum_{i=1}^n e_i^2 = \sum_{i=1}^n (y_i - \hat{y}_i)^2$$

$$SST = SSE + SSR$$

Donde,  $SST$  es la suma de cuadrados totales y  $SSR$  es la suma de cuadrados de la regresión. En la Figura 11 se pueden observar los componentes de la regresión lineal simple.

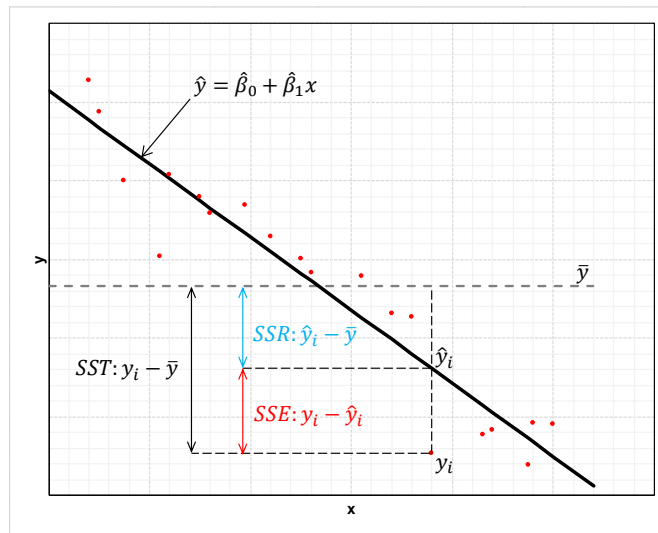


Figura 11. Componentes de una regresión lineal simple

La suma de cuadrados de los errores tiene  $n - 2$  grados de libertad, porque dos grados de libertad se asocian con los estimados de  $\beta_0$  y  $\beta_1$  que se usan para obtener  $\hat{y}_i$ . De aquí se obtiene el cuadrado medio residual ( $MSE$ ), el cual se calcula mediante la siguiente ecuación:

$$\hat{\sigma}^2 = \frac{SSE}{n - 2} = MSE$$

A la raíz cuadrada del  $\hat{\sigma}$  se le llama error estándar de la regresión y tiene las mismas unidades que la variable de respuesta  $y$ .

### 3.7.5 Prueba de significancia de la regresión

Las hipótesis planteadas para establecer la significancia de la regresión son las siguientes:

$$H_0: \beta_1 = 0$$

$$H_1: \beta_1 \neq 0$$

El no rechazar  $H_0$  implica que no hay relación lineal entre  $x$  y  $y$ . Esto es que  $x$  tiene muy poco valor para explicar la variación de  $y$  y que el mejor estimador para cualquier  $x$  es  $\hat{y} = \bar{y}$ , o que la verdadera regresión entre  $x$  y  $y$  no es lineal. Por consiguiente, si no se rechaza  $H_0$ , equivale a decir que no hay relación lineal entre  $x$  y  $y$ .

También se utilizó la prueba F de Fisher del análisis de varianza para probar la hipótesis nula. Aquí, Se requiere conocer los grados de libertad de los componentes de la regresión, los cuales se presentan a continuación:

Tabla 4. Análisis de varianza de la regresión

COMPONENTE	GRADOS DE LIBERTAD	SUMA DE CUADRADOS	CUADRADO MEDIO
REGRESIÓN(SSR)	1	SST - SSE	MSR
ERRORES (SSE)	$n-2$	$(y_i - \hat{y}_i)$	MSE
TOTAL (SST)	$n-1$	$(y_i - \bar{y})$	

El estadístico  $F_0$  de la regresión se calcula mediante la siguiente ecuación:

$$F_0 = \frac{MSR}{MSE}$$

El valor de  $F_0$  debe ser grande si  $\beta_1 \neq 0$ , por consiguiente, para probar la hipótesis  $H_0$ , se calcula el estadístico  $F_0$  de prueba y se rechaza  $H_0$  si:

$$F_0 > F_{\alpha,1,n-2}$$

Donde,  $\alpha$  es el grado de significancia de la prueba para  $n - 2$  grados de libertad en el numerador.

## 4 RESULTADOS

### 4.1 ÁREAS DE ESTUDIO

Como fuente de información primaria esta investigación utilizó los datos y registros hidrometeorológicos de la red nacional de monitoreo operada por el IDEAM y disponibles, para el caso de los registros hidrológicos, en el Sistema de Información del Recurso Hídrico (SIRH) (Instituto de Hidrología y Meteorología y Estudios Ambientales - IDEAM, 2014). De esta base de datos disponible en línea se extrajeron alrededor de 10 millones de datos correspondientes a registros diarios de caudales en 1233 estaciones hidrológicas en todo el territorio nacional. La información se almacenó en archivos planos (\*.csv) en forma de series de tiempo. La información completa de los metadatos de las estaciones se presenta en el anexo **A1**. A continuación, se recuerda el criterio de selección de las áreas de estudio:

- Registro histórico con al menos 30 años,
- Cantidad de datos faltantes no superior al 30%, entre el primer y último registro,
- Series con memoria de orden 2, esto es, que la función de auto-correlación tenga como máximo dos rezagos estadísticamente significativos.

La distribución espacial de la red nacional de monitoreo hidrológico, y, la distribución de las estaciones que cumplen los dos primeros criterios se presenta en la Figura 12. En la figura, se aprecia que la mayor densidad de monitoreo se encuentra en la cuenca Magdalena – Cauca (área hidrográfica 2 de acuerdo con la zonificación realizada por el IDEAM).

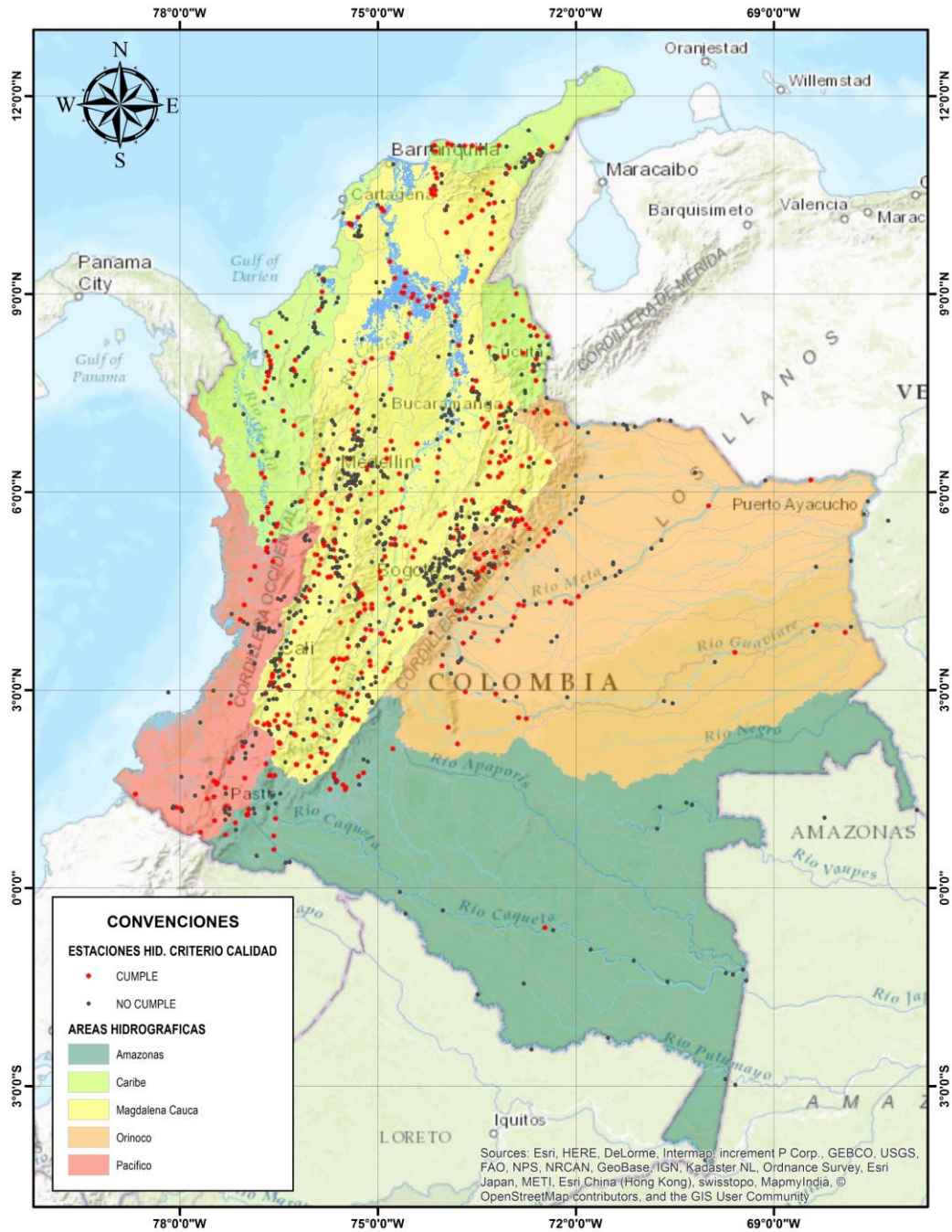


Figura 12. Red nacional de monitoreo hidrológico y evaluación del criterio de calidad de la información

La aplicación de los primeros dos criterios a los registros permitió identificar los puntos de cierre para la delimitación de las áreas de drenaje, los resultados se resumen a continuación en la Figura 13. Se puede observar que resultan 436 estaciones hidrológicas con un registro histórico igual o mayor a 30 años con un porcentaje menor al 30% en cantidad de datos faltantes en el periodo de registro.

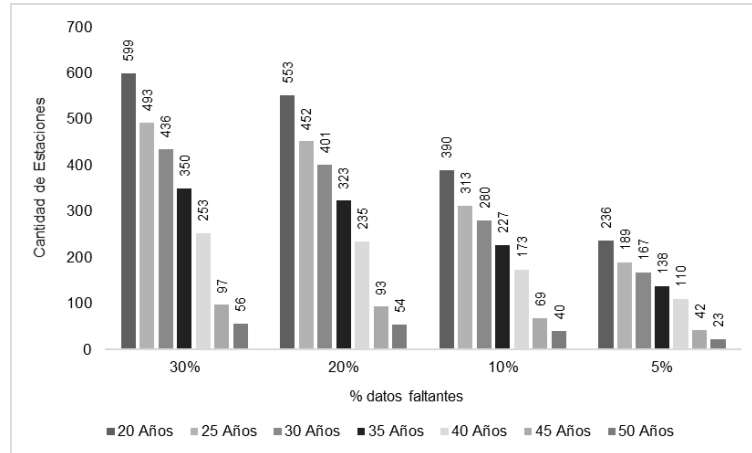


Figura 13. Número de estaciones hidrológicas que cumplen el criterio de longitud de registro y completitud.

El último filtro de selección, lo completa la función de auto-correlación. La metodología de cálculo fue descrita previamente [pág. 37]. Para efectos de considerar solo las anomalías del proceso en la estimación de la ACF, las series hidrológicas se desestacionalizaron, esto es, remover el componente estacional a la serie de tiempo. En este caso, las series se centralizaron restando el valor medio de cada conjunto estadístico a su respectivo conjunto.

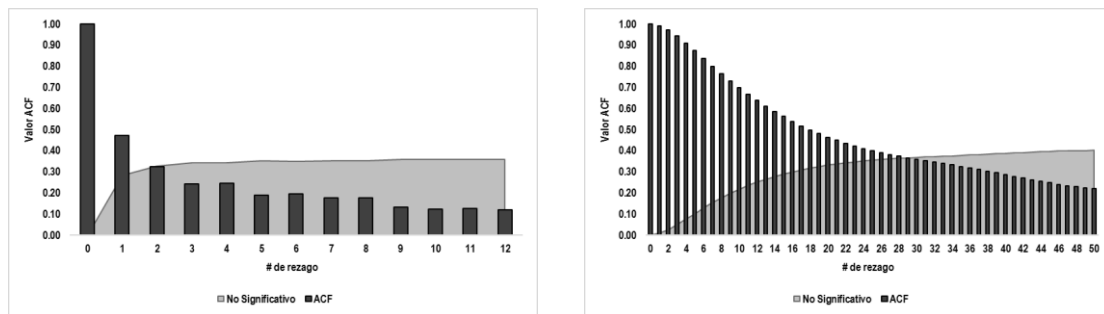


Figura 14. Función de auto-correlación para la estación 12017080 [Izq.] se observa que la memoria del proceso deja de ser significativa en el rezago 2, mientras la estación 44157010 [Der.] presenta una inercia de 30 rezagos.

La aplicación del filtro de ACF sobre las 436 estaciones identificadas con los primeros filtros, reduce a 20 puntos de monitoreo que cumplen perfectamente los tres criterios adoptados para la selección.

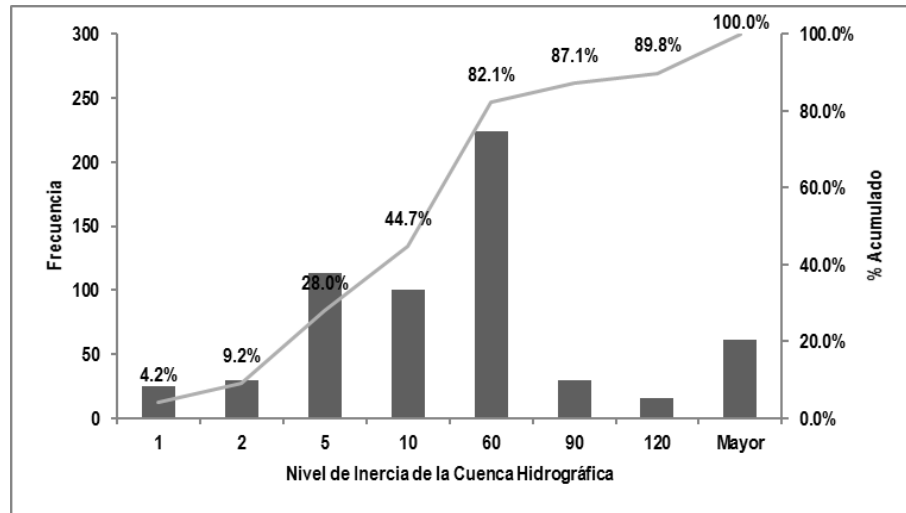


Figura 15. Distribución del número de estaciones en función de su grado de inercia (magnitud de la ACF significativa)

Este resultado presenta lógica en el sentido que corrientes con una corta memoria (2 días), en general, está asociada a cuencas hidrográficas relativamente pequeñas, y, por lo tanto, es lógico pensar que la mayor cantidad de estaciones de la red de monitoreo está distribuida en mayor medida hacia cuencas de mayor tamaño como se puede apreciar en la Figura 15. De acuerdo al resultado obtenido bajo los tres criterios estipulados inicialmente se definió reducir el umbral del primer criterio (longitud del registro) a 20 años de registro dado que: a) el número de áreas que cumplen no garantizan un conjunto estadístico representativo al momento de obtener regresiones significativas entre los índices morfométricos y los parámetros de la FPK, b) las 20 estaciones que cumplen estrictamente los criterios, obedecen a en su mayoría a la cuenca Magdalena – Cauca y en menor proporción a las zonas de Orinoquía, Amazonía y Caribe, faltando aún la zona pacífica.

De acuerdo con el nuevo criterio, y al realizar los filtros nuevamente, se obtienen en esta ocasión 30 estaciones que cumplen, esta vez, con representaciones en todas las áreas hidrográficas del país, lo cual es importante pues registra toda la variabilidad morfométrica del territorio nacional, y, por lo tanto, enriquece el análisis de las relaciones entre estos índices intrínsecos de la cuenca versus los parámetros del operador matemático que la representa. De esta forma, el número de cuencas objeto de esta investigación queda formada por un conjunto de 30 estaciones como se resume en la Tabla 5. El consolidado de los resultados obtenidos para la aplicación del tercer filtro (Función de auto-correlación) puede ser consultado en el Anexo A2-Funciones de auto-correlación.

Tabla 5. Conjunto de estaciones seleccionadas para el ejercicio de investigación

CÓDIGO	INICIO REGISTRO	FIN REGISTRO	LONGITUD REGISTRO (AÑOS)	% GAPS	NOMBRE	DEPARTAMENTO	CORRIENTE
21017050	abr-71	dic-14	43	0.5%	PITALITO 2 AUTOMÁTICA	HUILA	GUARAPAS
21167050	may-71	dic-14	43	2.1%	PTE CUNDAY	TOLIMA	PRADO
44037080	abr-71	dic-14	43	6.5%	MORELIA	CAQUETÁ	BODOQUERO
21017030	abr-71	dic-14	43	26.1%	CASCADA SIMÓN BOLI	HUILA	MAGDALENA
22067010	feb-71	dic-14	43	26.2%	PTE ORTEGA	TOLIMA	ORTEGA
21227010	ene-73	dic-14	42	9.4%	PIEDRAS	TOLIMA	OPIA
21037020	may-71	nov-13	42	23.8%	SAN MARCOS	HUILA	SUAZA
23087160	mar-74	jun-15	41	8.7%	CARAMANTA	ANTIOQUIA	NUS
35057010	ene-69	abr-10	41	9.5%	CABLE EL	CUNDINAMARCA	HUMEA
26187010	mar-42	abr-80	38	13.8%	MAYORIA LA BP-1	ANTIOQUIA	BUEY



CÓDIGO	INICIO REGISTRO	FIN REGISTRO	LONGITUD REGISTRO (AÑOS)	% GAPS	NOMBRE	DEPARTAMENTO	CORRIENTE
15017030	ene-78	dic-14	37	8.7%	MINCA	MAGDALENA	GAIRA
24057050	dic-73	dic-10	37	18.5%	PTE COMUNA	SANTANDER	CHUCURI
12027050	ago-78	jun-15	36	1.1%	PUEBLO NUEVO	ANTIOQUIA	MULATOS
21027010	jun-80	dic-15	35	0.0%	PERICONGO	HUILA	MAGDALENA
35017070	ene-80	dic-14	35	19.6%	RANCHO ALEGRE	META	OROTOY
12017080	mar-77	sep-11	34	9.2%	CAREPA	ANTIOQUIA	CAREPA
47017150	jun-81	dic-14	33	10.2%	JOYA LA	PUTUMAYO	GUINEO
44017050	ene-83	dic-14	32	0.1%	RIOGRANDE - CAUCA	CAUCA	GRANDE
35037130	may-78	dic-10	32	9.3%	PTE DEL AMOR	META	OCOA
11027050	ene-82	nov-13	31	18.7%	GINDRAMA	CHOCÓ	ATRATO
27017130	ene-61	dic-88	28	25.3%	PALMICHAL G-8	ANTIOQUIA	GUADALUPE
53077020	ene-90	dic-14	25	3.3%	GUAYABAL	CAUCA	MICAY
27017070	ago-54	jul-80	25	4.2%	PTE BELMIRA RG-6	ANTIOQUIA	GRANDE
53077060	ene-91	dic-14	24	13.7%	MONO EL	CAUCA	CHUARE
21017060	sep-91	dic-14	23	1.1%	MAGDALENA LA	HUILA	MAGDALENA
26137120	sep-71	oct-94	23	0.1%	REINA LA	RISARALDA	SAN EUGENIO
21057040	ago-71	feb-94	22	0.0%	BOCATOMA	CAUCA	ULLUCOS
21167090	ago-92	dic-14	22	1.9%	HDA EL RUBY AUTOM	TOLIMA	CUNDAY
26187020	ene-59	dic-79	21	0.4%	RIO PIEDRAS BP-2B	ANTIOQUIA	PIEDRAS
23147040	nov-92	oct-14	21	17.3%	AYACUCHO	SANTANDER	LA COLORADA

En la Figura 16 se presenta el mapa con la localización espacial de las estaciones que finalmente cumplieron los criterios de selección. Como se puede observar, el 60% de las estaciones se ubica en la macrocuenca Magdalena-Cauca sobre las partes altas del Río Magdalena, los afluentes medios sobre la parte media de la cuenca del Magdalena, seguidos de las áreas hidrográficas Caribe, Orinoquía y Amazonía. Por último, con 2 estaciones localizadas sobre la región Pacífico completan el cuadro de áreas experimentales.

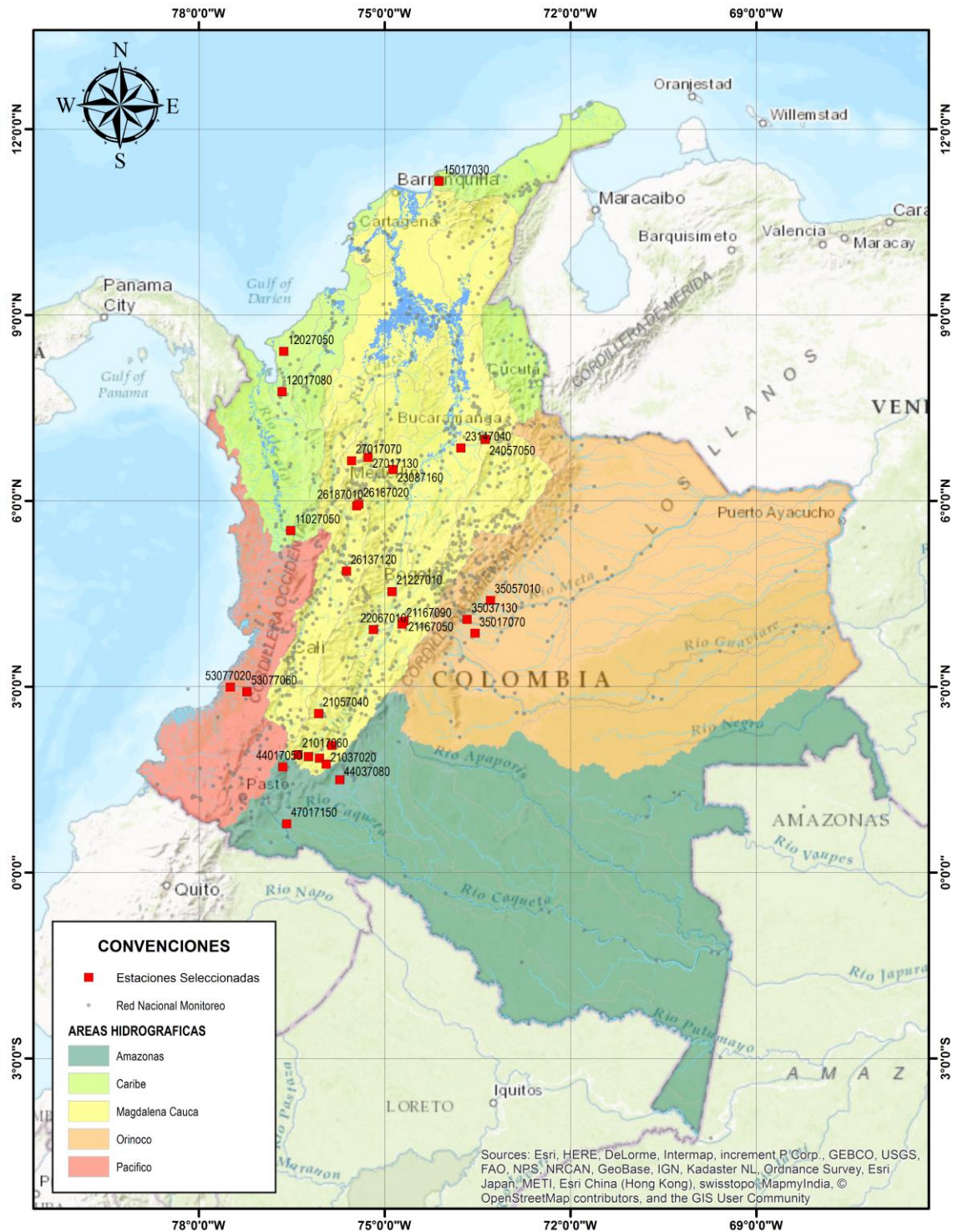


Figura 16. Localización de las estaciones hidrológicas seleccionadas para la investigación. La identificación de las estaciones hidrológicas que cumplen los criterios de selección permitió delimitar las áreas de drenaje a dichas estaciones mediante el uso del modelo digital de terreno de 92 metros aproximadamente (NASA Jet Propulsion Laboratory (JPL), 2015) y la hidrografía nacional escala 1:100.000. En la

Figura 18 se muestran los resultados del proceso de delimitación de las 30 áreas correspondientes a los puntos seleccionados. Por su parte, la Tabla 6 consolida la información del proceso de delimitación, en resumen, las áreas de estudio son alrededor de 20,000 Km<sup>2</sup> concentrados en la macrocuenca Magdalena Cauca el 54% del área de estudio, sin embargo, el área hidrográfica pacífico [5] con solo dos unidades representa el 22% de las áreas de estudio, esto es importante para nuestro análisis pues a pesar de ser cuencas que comparten el mismo grado de inercia (representado por la función de autocorrelación) evidentemente presentan características internas (incorporadas aquí por el área de drenaje) que claramente difieren entre sí. La Figura 17 muestra la distribución de las áreas de drenaje, se observa que más del 50% de las unidades presenta áreas entre 50 a 400 Km<sup>2</sup>.

Tabla 6. Resumen del proceso de delimitación de áreas aferentes

Área Hidrográfica	Unidades de Análisis Hidrográficas (UAH)	Área Drenaje (Km <sup>2</sup> )	% del Total de Unidades	% del total del Área
1	4	2,879	13%	15%
2	18	10,609	60%	54%
3	3	1,122	10%	6%
4	3	852	10%	4%
5	2	4,307	7%	22%
<b>Total</b>	<b>30</b>	<b>19,769</b>	<b>100%</b>	<b>100%</b>

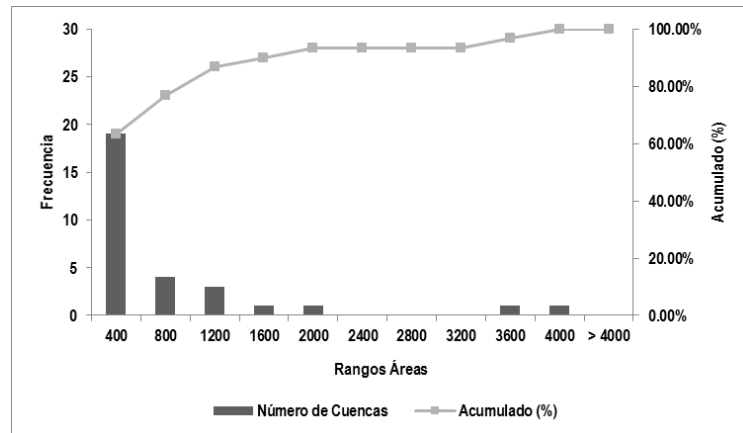


Figura 17. Distribución de las unidades de análisis en función de su área de drenaje en Km<sup>2</sup>

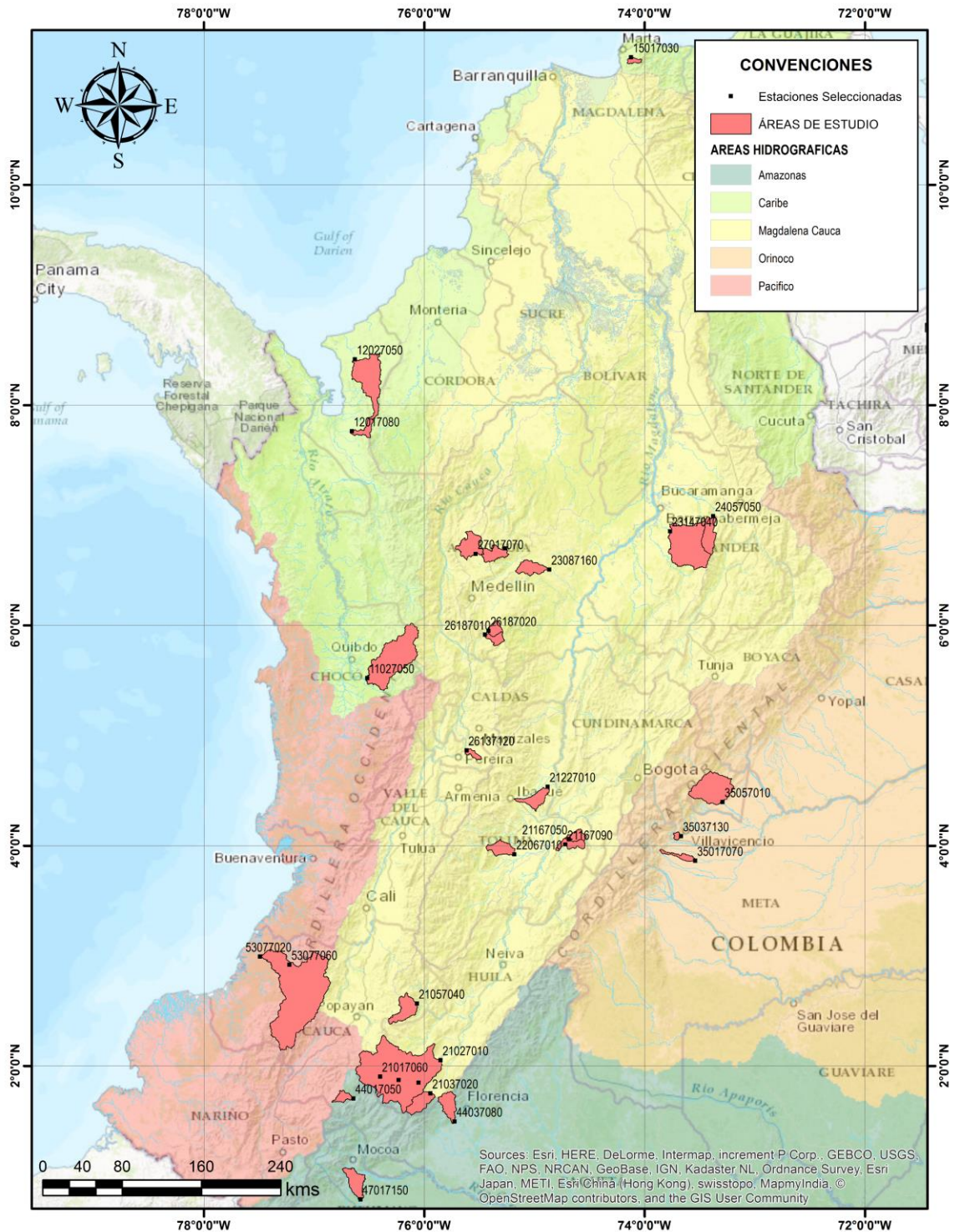


Figura 18. Delimitación de áreas aferentes a las estaciones hidrológicas seleccionadas.

Una vez delimitadas las áreas de drenaje a los puntos seleccionados anteriormente, se determinaron las estaciones de precipitación según la metodología planteada en el capítulo anterior. El filtro espacial aplicado (buffer de 10 km a las cuencas hidrográficas) permitió identificar cerca de 287

estaciones de tipo meteorológico, sin embargo, al aplicar el filtro de calidad de los datos (longitud de registro mayor a 30 años y cantidad de datos faltantes inferior al 30% entre el primer y último registro) esta cantidad disminuyó a 129 estaciones. La distribución espacial de las estaciones se presenta en la Figura 19. La relación de estaciones seleccionadas para la caracterización climática y las series hidrológicas seleccionadas se adjunta en el Anexo A4-Series hidroclimatológicas. Además, el Anexo contiene las series diarias de precipitación total de las 287 estaciones identificadas con el buffer, y la relación de las estaciones de precipitación que caracterizaron las entradas al sistema para cada cuenca hidrográfica.

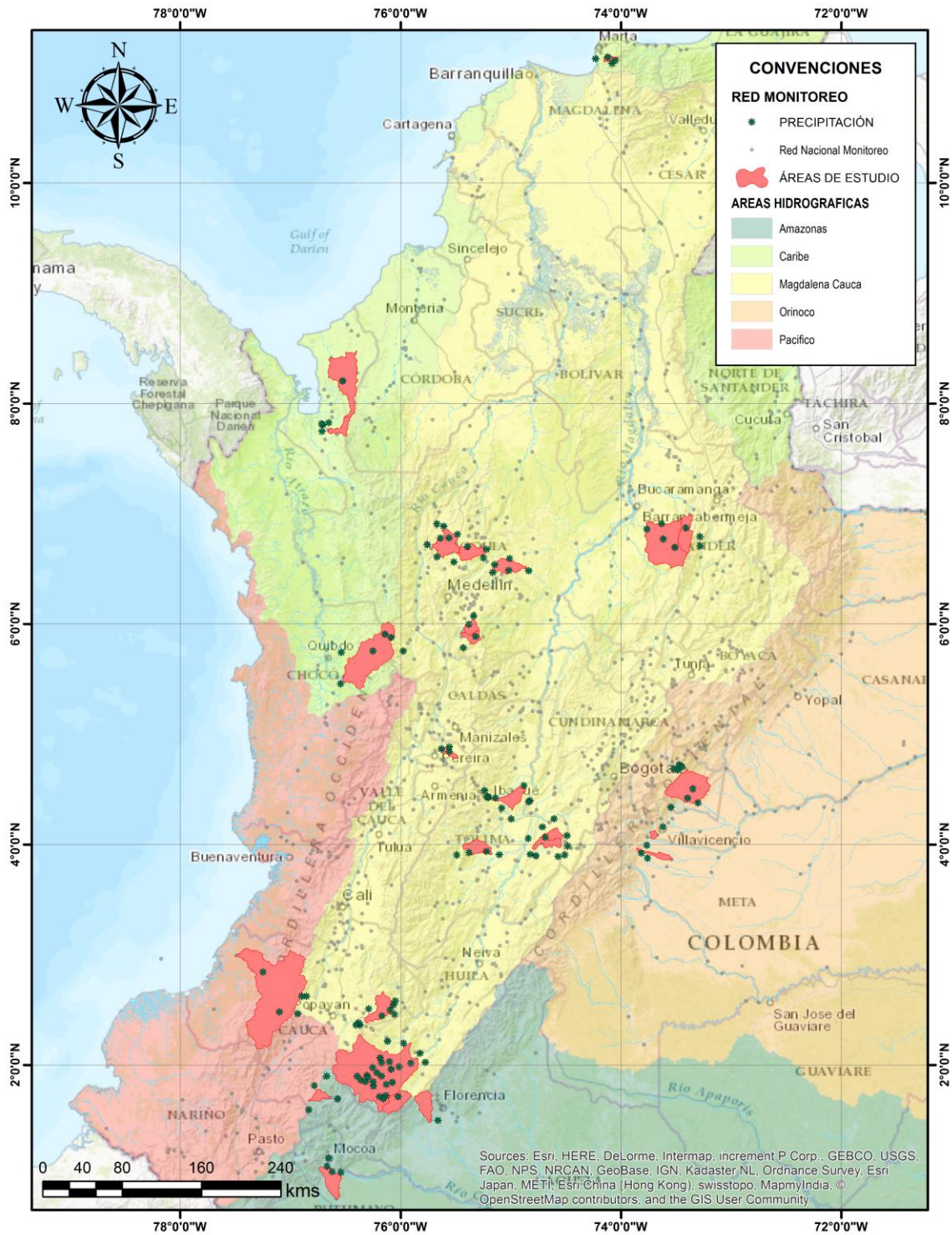


Figura 19. Localización de las estaciones de precipitación seleccionadas para la investigación

## 4.2 CARACTERIZACIÓN MORFOMÉTRICA

Continuando con el flujo de trabajo, y de acuerdo a la metodología descrita, se estimaron 23 índices morfométricos, entre los que se encuentran, altura ponderada de la cuenca, altura ponderada del cauce principal, ancho medio de la cuenca, área, coeficiente de sinuosidad, coeficiente de compacidad, coeficiente de forma, densidad de drenaje, longitud cauce más largo, longitud lineal del cauce más largo, longitud media de laderas, nivel de disección del relieve, orden jerárquico (Strahler), pendiente ponderada de la cuenca, pendiente ponderada del cauce principal, perímetro, profundidad de disección de la cuenca, tiempos de concentración por distintas metodologías como, Bransby-Williams, Dirección. General de Carreteras (España), Giandotti, Kirpich, Temez, y, las tasas de bifurcación. La estimación de los índices fue implementada en el lenguaje de programación Python. El código toma como insumos ráster el modelo digital de terreno, la dirección de flujo y la acumulación del flujo, y como insumos vectoriales, el archivo vectorial de las unidades de análisis delimitadas en el paso anterior y el modelo digital de corrientes derivado del proceso de delimitación. Es importante recalcar aquí que mediante un pre-proceso del modelo digital de terreno y la hidrografía, las corrientes generadas digitalmente concuerdan con la hidrografía nacional lo cual hace que el procedimiento tanto de delimitación como de caracterización morfométrica sea de gran confiabilidad metodológica. La Figura 20 muestra algunos de los resultados de la caracterización morfométrica.

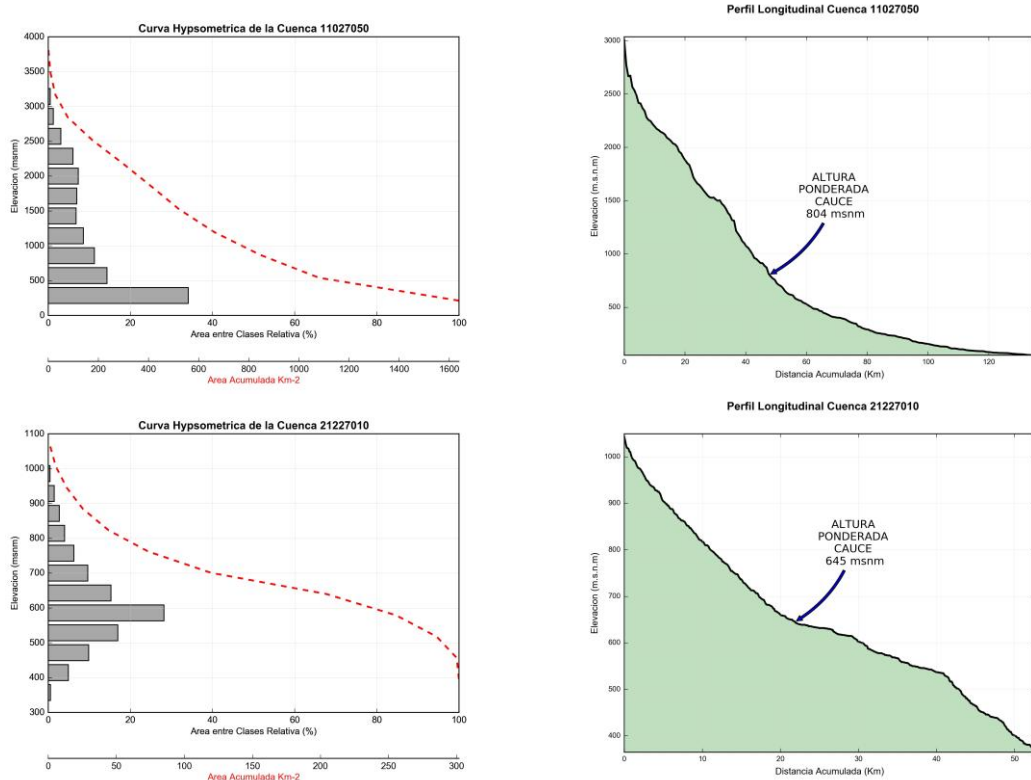


Figura 20. Algunas salidas gráficas de la caracterización morfométrica. A la izquierda se observa la curva hipsométrica para dos unidades muy distintas como se puede apreciar, la primera [Izq. Arriba] clasifica como una cuenca con procesos de erosión intermedios dada la distribución de las elevaciones hacia la parte baja de la cuenca. Por otra parte, la segunda cuenca [Izq. Abajo] presenta una distribución de elevaciones mejor distribuida lo cual indica alto potencial de erosión, esto es particular de cuencas en proceso de maduración. Los escenarios de ambas cuencas son corroborados en los gráficos de la derecha, en donde se muestran los perfiles longitudinales de los cauces principales.

Los resultados en detalle de la caracterización morfométrica se presentan en el Anexo A3- Caracterización Morfométrica. La Tabla 7 contiene el listado de índices morfométricos estimados para cada cuenca hidrográfica, su sigla y sus unidades.

Tabla 7. Abreviatura y unidades de los índices morfométricos

.	UNIDAD	DESCRIPCIÓN
<b>A</b>	Km <sup>2</sup>	Área
<b>P</b>	Km	Perímetro
<b>L<sub>t</sub></b>	Km	Longitud Cauce Más Largo
<b>L<sub>l</sub></b>	Km	Longitud Lineal del Cauce Más Largo
<b>K<sub>s</sub></b>	Ad	Coefficiente Sinuosidad
<b>D</b>	KmKm <sup>-2</sup>	Densidad de Drenaje
<b>L<sub>lad</sub></b>	Km <sup>2</sup> Km	Longitud Media de Laderas
<b>B</b>	Km	Ancho Medio de la Cuenca
<b>s</b>	Ad	Coefficiente de Forma
<b>K<sub>c</sub></b>	Ad	Coefficiente de Compacidad
<b>H<sub>cauce</sub></b>	Msnm	Altura Ponderada Cauce Principal
<b>I<sub>cauce</sub></b>	%	Pendiente Ponderada Cauce Principal
<b>H<sub>cuenca</sub></b>	Msnm	Altura Ponderada de la Cuenca
<b>I<sub>cuenca</sub></b>	%	Pendiente Ponderada de la Cuenca
<b>H<sub>dis</sub></b>	m.	Profundidad de Disección de la Cuenca
<b>N<sub>dis</sub></b>	Ad	Nivel de Disección del Relieve
<b>T<sub>c</sub></b>	Horas	Tiempo Concentración - Kirpich
<b>T<sub>c</sub></b>	Horas	Tiempo Concentración - Temez
<b>T<sub>c</sub></b>	Horas	Tiempo Concentración - Giandotti
<b>T<sub>c</sub></b>	Horas	Tiempo Concentración - Bransby-Williams
<b>T<sub>c</sub></b>	Horas	Tiempo Concentración - Dir. Gen. Carreteras
<b>N</b>	Ad	Numero de Orden Máximo - Strahler

La Tabla 8 y la Tabla 9 resumen los índices morfométricos para cada una de las 30 unidades objeto de esta investigación.



Tabla 8. Resumen de la caracterización morfométrica (15/30)

<b>Símbolo</b>	<b>Unidad</b>	<b>11027050</b>	<b>12017080</b>	<b>12027050</b>	<b>15017030</b>	<b>21017030</b>	<b>21017050</b>	<b>21017060</b>	<b>21027010</b>	<b>21037020</b>	<b>21057040</b>	<b>21167050</b>	<b>21167090</b>	<b>21227010</b>	<b>22067010</b>	<b>23087160</b>
<b>A</b>	<b>Km<sup>2</sup></b>	1,638.6	160.4	1,017.0	58.8	1,124.0	269.3	557.8	3,562.3	352.7	423.5	148.1	348.1	301.7	281.4	303.4
<b>P</b>	<b>Km</b>	305.1	109.7	285.8	52.0	260.8	130.9	150.4	466.7	120.0	162.0	80.0	135.9	139.0	116.9	128.0
<b>L<sub>t</sub></b>	<b>Km</b>	135.6	37.6	114.1	17.2	73.7	42.4	44.3	141.0	39.4	45.8	22.4	29.5	52.6	45.5	45.4
<b>L<sub>l</sub></b>	<b>Km</b>	70.3	20.5	60.4	13.1	47.2	31.6	30.5	81.8	30.4	33.4	17.1	21.6	34.8	29.0	34.0
<b>K<sub>s</sub></b>	<b>Ad</b>	1.9	1.8	1.9	1.3	1.6	1.3	1.5	1.7	1.3	1.4	1.3	1.4	1.5	1.6	1.3
<b>D</b>	<b>KmKm<sup>-2</sup></b>	0.5	0.6	0.5	0.5	0.6	0.5	0.6	0.5	0.5	0.5	0.6	0.5	0.8	0.6	0.4
<b>L<sub>lad</sub></b>	<b>Km<sup>2</sup>Km</b>	0.9	0.9	1.0	1.0	0.9	0.9	0.9	0.9	1.0	1.0	0.9	0.9	0.6	0.8	1.2
<b>B</b>	<b>Km</b>	23.3	7.8	16.8	4.5	23.8	8.5	18.3	43.5	11.6	12.7	8.7	16.1	8.7	9.7	8.9
<b>σ</b>	<b>Ad</b>	3.0	2.6	3.6	2.9	2.0	3.7	1.7	1.9	2.6	2.6	2.0	1.3	4.0	3.0	3.8
<b>K<sub>c</sub></b>	<b>Ad</b>	2.1	2.4	2.5	1.9	2.2	2.2	1.8	2.2	1.8	2.2	1.8	2.0	2.2	2.0	2.1
<b>l<sub>cauce</sub></b>	<b>%</b>	2.7	2.7	1.0	11.5	5.1	3.8	5.3	2.3	4.8	4.7	8.2	6.5	1.5	5.8	4.4
<b>l<sub>cuena</sub></b>	<b>%</b>	31.1	21.0	13.2	36.5	31.1	24.8	35.1	27.2	29.0	31.8	21.8	22.1	9.7	25.4	24.0
<b>H<sub>dis</sub></b>	<b>m.</b>	231.2	165.9	27.4	183.3	554.8	163.0	457.7	533.7	212.0	291.9	131.0	237.7	19.1	-53.9	240.8
<b>N<sub>dis</sub></b>	<b>Ad</b>	0.3	0.2	0.0	0.2	0.6	0.2	0.5	0.6	0.2	0.3	0.1	0.3	0.0	-0.1	0.2
<b>T<sub>c</sub></b>	<b>Horas</b>	6.9	2.6	7.5	1.1	3.9	2.7	2.6	7.1	2.4	2.7	1.4	1.8	3.8	2.6	2.9
<b>N</b>	<b>Ad</b>	5.0	3.0	5.0	3.0	5.0	3.0	4.0	6.0	4.0	4.0	4.0	4.0	4.0	4.0	3.0

Tabla 9. Resumen de la caracterización morfométrica (30/30)

<b>Símbolo</b>	<b>Unidad</b>	<b>23147040</b>	<b>24057050</b>	<b>26137120</b>	<b>26187010</b>	<b>26187020</b>	<b>27017070</b>	<b>27017130</b>	<b>35017070</b>	<b>35037130</b>	<b>35057010</b>	<b>44017050</b>	<b>44037080</b>	<b>47017150</b>	<b>53077020</b>	<b>53077060</b>
<b>A</b>	<b>Km<sup>2</sup></b>	1,458.7	383.1	80.7	139.6	144.2	446.7	275.4	125.8	41.5	954.5	130.1	341.3	379.7	3,691.6	603.4
<b>P</b>	<b>Km</b>	271.3	139.4	66.2	85.1	71.1	165.4	129.6	118.2	36.9	183.2	83.1	127.9	150.4	505.0	172.9
<b>L<sub>t</sub></b>	<b>Km</b>	82.2	54.1	23.0	26.5	27.9	35.6	36.4	45.4	11.2	47.4	28.3	49.3	47.6	178.9	60.6
<b>L<sub>l</sub></b>	<b>Km</b>	50.1	38.9	17.4	12.6	8.8	20.0	25.1	36.0	8.2	30.9	22.0	29.8	35.7	96.5	38.5
<b>K<sub>s</sub></b>	<b>Ad</b>	1.6	1.4	1.3	2.1	3.2	1.8	1.4	1.3	1.4	1.5	1.3	1.7	1.3	1.9	1.6
<b>D</b>	<b>KmKm<sup>-2</sup></b>	0.5	0.5	0.6	0.5	0.5	0.4	0.4	0.9	0.6	0.5	0.5	0.5	0.6	0.5	0.5
<b>L<sub>lad</sub></b>	<b>Km<sup>2</sup>Km</b>	1.0	1.1	0.8	1.1	1.1	1.2	1.1	0.5	0.9	0.9	0.9	1.0	0.8	1.0	1.1
<b>B</b>	<b>Km</b>	29.1	9.8	4.6	11.0	16.4	22.3	11.0	3.5	5.1	30.9	5.9	11.5	10.6	38.3	15.7
<b>σ</b>	<b>Ad</b>	1.7	4.0	3.7	1.1	0.5	0.9	2.3	10.3	1.6	1.0	3.7	2.6	3.4	2.5	2.5
<b>K<sub>c</sub></b>	<b>Ad</b>	2.0	2.0	2.1	2.0	1.7	2.2	2.2	2.9	1.6	1.7	2.0	1.9	2.2	2.3	2.0
<b>I<sub>cauce</sub></b>	<b>%</b>	5.3	6.6	12.1	3.8	2.1	2.9	3.1	2.8	8.0	9.0	5.4	5.9	8.3	3.1	7.0
<b>I<sub>cuenca</sub></b>	<b>%</b>	22.1	37.6	28.8	21.7	15.5	14.8	17.0	5.6	20.0	32.4	22.6	36.3	10.0	41.6	42.3
<b>H<sub>dis</sub></b>	<b>m.</b>	-35.1	384.2	47.0	151.1	95.3	82.2	119.8	-23.8	58.9	310.8	130.6	413.4	-37.7	652.4	335.9
<b>N<sub>dis</sub></b>	<b>Ad</b>	0.0	0.4	0.1	0.1	0.1	0.1	0.1	0.0	0.1	0.3	0.1	0.4	0.0	0.6	0.3
<b>T<sub>c</sub></b>	<b>Horas</b>	4.3	2.9	1.3	1.9	2.2	2.5	2.6	3.0	0.9	2.6	1.8	2.8	2.8	8.9	3.2
<b>N</b>	<b>Ad</b>	5.0	4.0	3.0	4.0	3.0	4.0	4.0	3.0	3.0	5.0	3.0	4.0	3.0	6.0	5.0

### 4.3 ANÁLISIS DE INFORMACIÓN HIDROCLIMATOLÓGICA

El análisis de la calidad y consistencia de la información hidroclimatológica se realizó entonces a 30 series de tiempo de caudales medios diarios y a 129 series meteorológicas de precipitación total diaria que fueron seleccionados bajo los criterios descritos en el capítulo anterior.

#### 4.3.1 Análisis visual

La revisión detalla de las series mediante la inspección visual permitió identificar valores evidentemente anómalos, y periodos de tiempo donde se observan alteración en la estructura de los registros como se puede observar en las gráficas abajo.

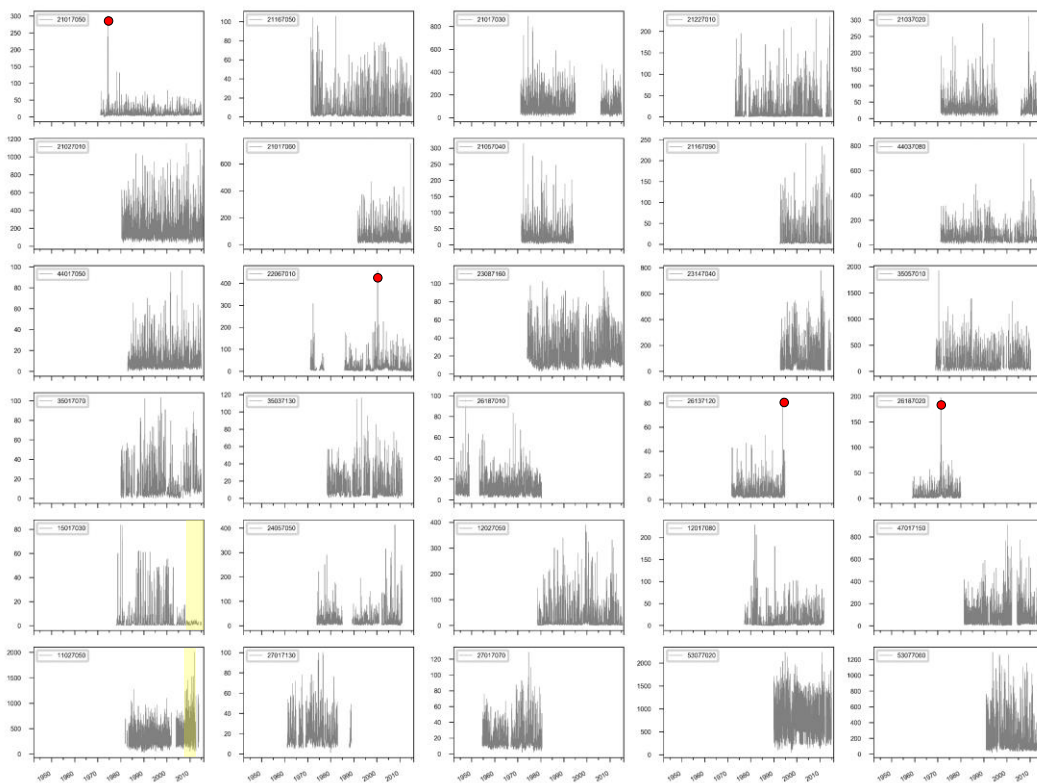
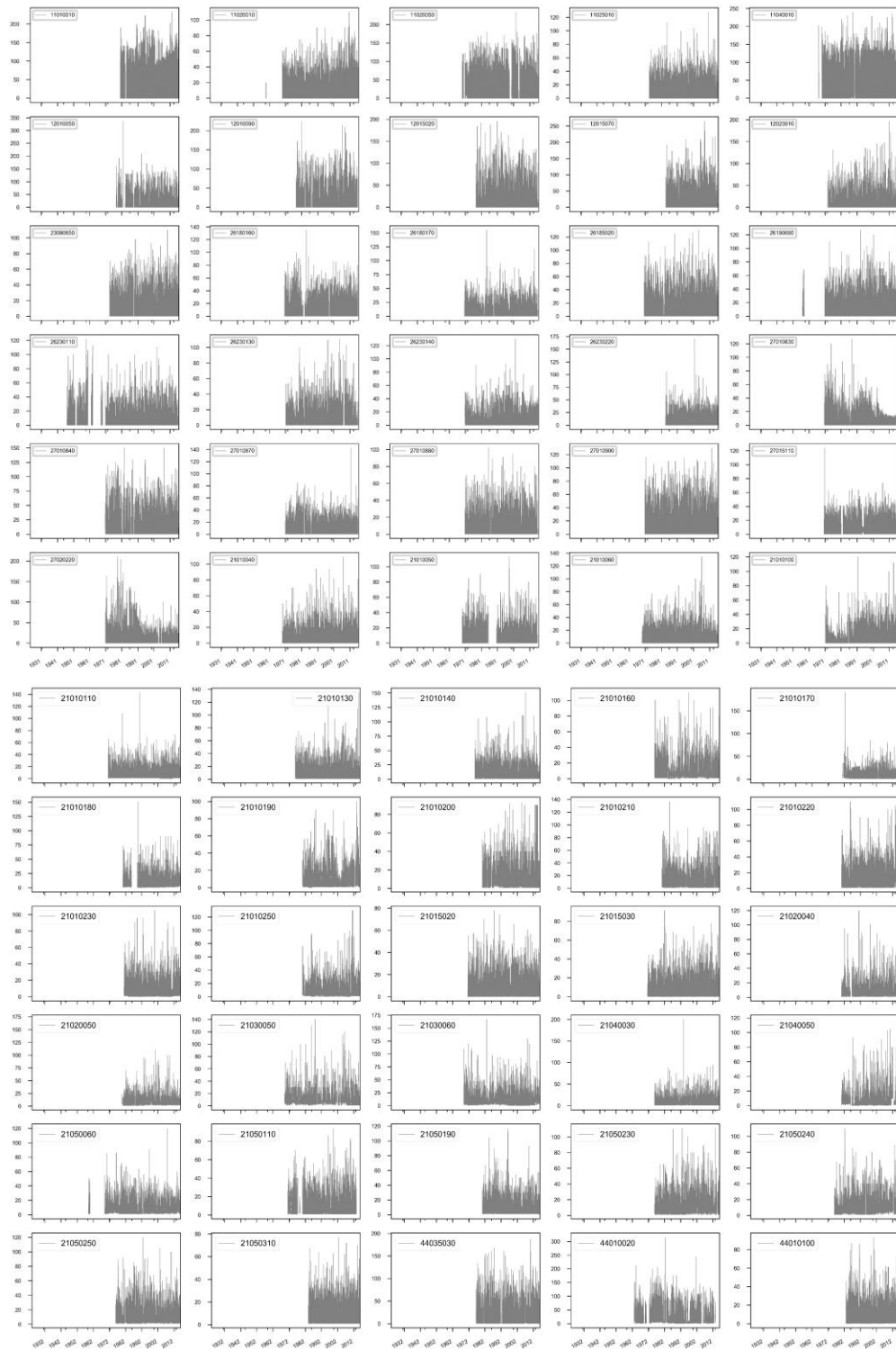


Figura 21. Series de Caudales medios diarios [m<sup>3</sup>/seg] reportados por IDEAM

Dichos datos fueron marcados como potenciales [puntos rojos], mientras que los cambios en la estructura [bandas amarillas] fueron removidos.



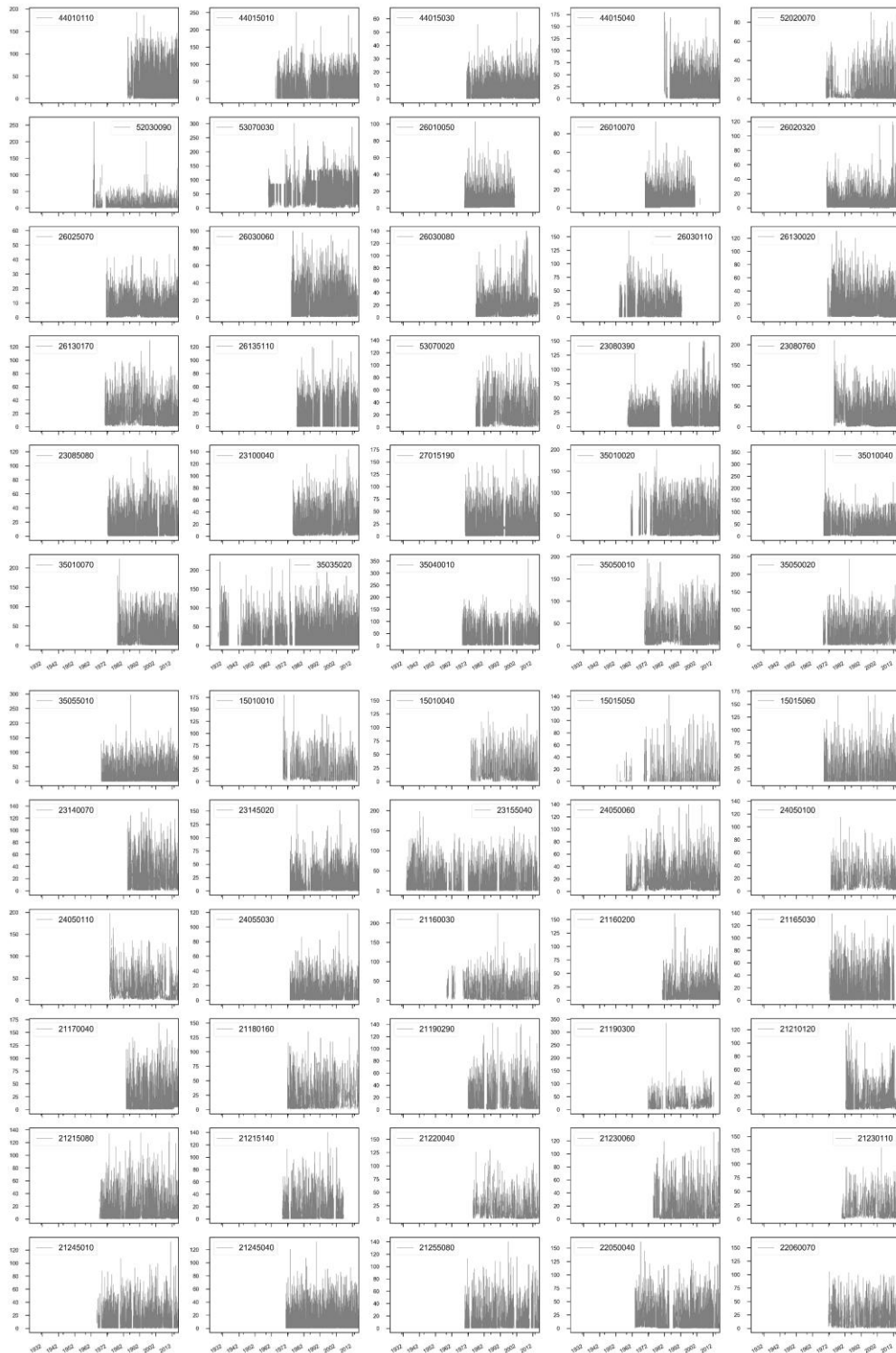


Figura 22. Series de precipitación total diaria [mm/día] reportados por IDEAM

### 4.3.2 Análisis por el método de desviaciones absolutas de la mediana (MADS)

Los resultados obtenidos por este método se resumen en la Tabla 10. En general, para las series hidrológicas, el método detectó en promedio 0.71% de datos potencialmente anómalos con un máximo detectado de 3.79% en la estación 21227010.

Tabla 10. Datos anómalos potenciales de caudal medio diario detectados por MADS

Estación	Registros iniciales	Registros Finales	% datos removidos	Estación	Registros iniciales	Registros Finales	% datos removidos
21017050	15,890	15,751	0.87%	35017070	10,269	10,230	0.38%
21167050	15,603	15,276	2.10%	35037130	10,790	10,753	0.34%
21017030	11,800	11,785	0.13%	26187010	11,995	11,981	0.12%
21227010	13,889	13,363	3.79%	26137120	8,448	8,428	0.24%
21037020	11,830	11,758	0.61%	26187020	7,634	7,618	0.21%
21027010	12,988	12,981	0.05%	15017030	12,331	12,114	1.76%
21017060	8,420	8,388	0.38%	24057050	11,026	10,908	1.07%
21057040	8,237	8,212	0.30%	12027050	13,332	13,018	2.36%
21167090	8,028	7,797	2.88%	12017080	11,446	11,297	1.30%
44037080	14,927	14,865	0.42%	47017150	11,011	10,986	0.23%
44017050	11,664	11,612	0.45%	11027050	9,460	9,459	0.01%
22067010	11,833	11,785	0.41%	27017130	7,635	7,631	0.05%
23087160	13,769	13,767	0.01%	27017070	9,072	9,052	0.22%
23147040	6,616	6,611	0.08%	53077020	8,824	8,824	0.00%
35057010	13,652	13,596	0.41%	53077060	7,564	7,546	0.24%

Para el caso de las series de precipitación total diaria, el método presenta una leve modificación debido a que en esta escala temporal se registran una gran cantidad de valores nulos que representan los eventos de no lluvia. Para efectos de ejecución del algoritmo se removieron los datos correspondientes a los eventos no lluviosos, por lo que los resultados presentados a continuación corresponden a datos potencialmente anómalos de los eventos lluviosos.

Los resultados para las series meteorológicas indican una detección levemente superior a las series hidrológicas con un promedio de 1.19% alcanzando valores de hasta 3.33% en la estación 12015070.

Tabla 11. Datos anómalos potenciales de precipitación total diario detectados por MADS

Estación	Registros iniciales	Registros Finales	% datos removidos	Estación	Registros iniciales	Registros Finales	% datos removidos
11010010	12,779	12,712	0.52%	52020070	16,155	15,890	1.64%
11020010	17,059	16,916	0.84%	52030090	17,768	17,556	1.19%
11020050	16,077	15,916	1.00%	53070030	16,805	16,780	0.15%
11025010	15,361	15,191	1.11%	26010050	11,206	11,022	1.64%
11040010	16,850	16,813	0.22%	26010070	11,150	11,015	1.21%
12010050	12,454	12,329	1.00%	26020320	16,330	16,006	1.98%
12010090	13,606	13,502	0.76%	26025070	15,063	14,876	1.24%
12015020	13,540	13,151	2.87%	26030060	14,935	14,778	1.05%
12015070	11,683	11,294	3.33%	26030080	13,972	13,856	0.83%
12020010	15,769	15,665	0.66%	26030110	12,859	12,779	0.62%
23080650	15,327	15,227	0.65%	26130020	15,896	15,647	1.57%
26180160	16,045	15,940	0.65%	26130170	16,211	16,049	1.00%
26180170	16,037	15,932	0.65%	26135110	11,378	11,180	1.74%
26185020	16,283	16,046	1.46%	53070020	12,850	12,740	0.86%

Estación	Registros iniciales	Registros Finales	% datos removidos	Estación	Registros iniciales	Registros Finales	% datos removidos
26190090	16,765	16,655	0.66%	23080390	17,610	17,360	1.42%
26230110	21,454	21,272	0.85%	23080760	14,504	14,372	0.91%
26230130	15,829	15,718	0.70%	23085080	14,584	14,361	1.53%
26230140	16,289	16,172	0.72%	23100040	14,510	14,412	0.68%
26230220	11,575	11,508	0.58%	27015190	16,480	16,218	1.59%
27010830	16,043	15,855	1.17%	35010020	16,734	16,584	0.90%
27010840	15,778	15,718	0.38%	35010040	16,439	16,276	0.99%
27010870	16,126	16,043	0.51%	35010070	13,251	13,117	1.01%
27010880	16,336	16,239	0.59%	35035020	26,804	26,418	1.44%
27010900	16,376	16,278	0.60%	35040010	15,278	15,148	0.85%
27015110	15,998	15,834	1.03%	35050010	16,457	16,337	0.73%
27020220	16,059	15,873	1.16%	35050020	16,673	16,565	0.65%
21010040	17,035	16,856	1.05%	35055010	16,793	16,566	1.35%
21010050	14,893	14,699	1.30%	15010010	15,930	15,820	0.69%
21010060	16,967	16,792	1.03%	15010040	14,371	14,282	0.62%
21010100	16,056	15,829	1.41%	15015050	19,443	19,112	1.70%
21010110	16,300	16,153	0.90%	15015060	16,296	16,147	0.91%
21010130	14,570	14,383	1.28%	23140070	11,209	11,073	1.21%
21010140	14,592	14,442	1.03%	23145020	13,974	13,646	2.35%
21010160	14,165	13,956	1.48%	23155040	25,842	25,578	1.02%
21010170	12,367	12,261	0.86%	24050060	20,922	20,742	0.86%
21010180	11,523	11,368	1.35%	24050100	15,188	15,130	0.38%
21010190	12,455	12,323	1.06%	24050110	15,184	15,115	0.45%
21010200	12,641	12,515	1.00%	24055030	14,664	14,448	1.47%
21010210	12,590	12,458	1.05%	21160030	17,595	17,483	0.64%
21010220	12,883	12,728	1.20%	21160200	12,477	12,239	1.91%
21010230	12,696	12,503	1.52%	21165030	14,542	14,164	2.60%
21010250	12,652	12,459	1.53%	21170040	11,672	11,446	1.94%
21015020	15,952	15,632	2.01%	21180160	16,047	15,968	0.49%
21015030	16,007	15,722	1.78%	21190290	14,503	14,389	0.79%
21020040	12,710	12,503	1.63%	21190300	12,443	12,397	0.37%
21020050	13,115	12,959	1.19%	21210120	11,299	11,111	1.66%
21030050	17,115	16,940	1.02%	21215080	16,974	16,695	1.64%
21030060	16,916	16,688	1.35%	21215140	12,363	12,041	2.60%
21040030	14,650	14,473	1.21%	21220040	14,613	14,521	0.63%
21040050	12,478	12,307	1.37%	21230060	14,649	14,476	1.18%
21050060	17,324	17,234	0.52%	21230110	12,792	12,696	0.75%
21050110	14,370	14,206	1.14%	21245010	16,675	16,448	1.36%
21050190	12,916	12,709	1.60%	21245040	16,279	16,006	1.68%
21050230	14,024	13,921	0.73%	21255080	14,716	14,370	2.35%
21050240	14,145	13,979	1.17%	22050040	17,620	17,382	1.35%
21050250	14,436	14,199	1.64%	22060070	15,571	15,479	0.59%
21050310	11,590	11,510	0.69%	22060100	13,042	12,961	0.62%
44035030	13,709	13,486	1.63%	22065040	11,351	11,152	1.75%
44010020	15,969	15,746	1.40%	21190090	20,262	19,988	1.35%
44010100	11,935	11,705	1.93%	21190240	15,454	15,164	1.88%
44010110	11,192	11,046	1.30%	21190330	12,347	12,130	1.76%
44015010	15,971	15,732	1.50%	35060090	18,343	18,136	1.13%
44015030	15,698	15,479	1.40%	35060120	15,314	15,138	1.15%
44015040	10,993	10,764	2.08%	35060130	14,409	14,183	1.57%
				35060140	15,675	15,539	0.87%

### 4.3.3 Análisis por el método de Grubbs

La prueba de Grubbs arrojó en general una mayor cantidad de datos anómalos en relación con la prueba de las medianas. Para el caso de las series hidrológicas, en promedio la prueba detecto un 1.35% de datos potencialmente anómalos, alcanzando valores máximos de hasta el 5.39% en la estación 21227010, la misma con mayor cantidad de datos detectados en MADS.

Tabla 12. Datos anómalos potenciales de caudal medio diario detectados por Grubbs

Estación	Registros iniciales	Registros Finales	% datos removidos	Estación	Registros iniciales	Registros Finales	% datos removidos
21017050	15,890	15,649	1.52%	35017070	10,269	10,135	1.30%
21167050	15,603	15,030	3.67%	35037130	10,790	10,716	0.69%
21017030	11,800	11,740	0.51%	26187010	11,995	11,943	0.43%
21227010	13,889	13,141	5.39%	26137120	8,448	8,380	0.80%
21037020	11,830	11,673	1.33%	26187020	7,634	7,563	0.93%
21027010	12,988	12,898	0.69%	15017030	12,331	12,122	1.69%
21017060	8,420	8,335	1.01%	24057050	11,026	10,817	1.90%
21057040	8,237	8,171	0.80%	12027050	13,332	12,885	3.35%
21167090	8,028	7,780	3.09%	12017080	11,446	11,149	2.59%
44037080	14,927	14,760	1.12%	47017150	11,011	10,900	1.01%
44017050	11,664	11,528	1.17%	11027050	9,460	9,437	0.24%
22067010	11,833	11,679	1.30%	27017130	7,635	7,609	0.34%
23087160	13,769	13,719	0.36%	27017070	9,072	8,983	0.98%
23147040	6,616	6,562	0.82%	53077020	8,824	8,824	0.00%
35057010	13,652	13,558	0.69%	53077060	7,564	7,502	0.82%

Para el caso de las estaciones meteorológicas, a nivel general los resultados fueron similares con un 1.49% de datos potencialmente anómalos en la prueba de Grubbs (frente al 1.19% de MADS). Sin embargo, se presentan una mayor cantidad de estaciones con valores de detección por encima del 5% en comparación con la prueba de MADS. Esta situación puede obedecer a: a) la serie puede presentar una fuerte asimetría en su función de densidad de probabilidad, y/o b) la serie presenta cambios drásticos en su estructura. Es conocido que la prueba de Grubbs es diseñada para series con distribución simétrica (Grubbs, 1950; Tietjen & Moore, 2012), por lo tanto, como criterio último de identificación de datos anómalos se seleccionaron el criterio de inspección visual y de las medianas.

Tabla 13. Datos anómalos potenciales de precipitación total diario detectados por Grubbs

Estación	Registros iniciales	Registros Finales	% datos removidos	Estación	Registros iniciales	Registros Finales	% datos removidos
11010010	12,779	12,770	0.07%	52020070	16,155	15,693	2.86%
11020010	17,059	16,968	0.53%	52030090	17,768	17,600	0.95%
11020050	16,077	16,063	0.09%	53070030	16,805	16,794	0.07%
11025010	15,361	15,310	0.33%	26010050	11,206	11,169	0.33%
11040010	16,850	16,844	0.04%	26010070	11,150	11,108	0.38%
12010050	12,454	12,209	1.97%	26020320	16,330	16,189	0.86%
12010090	13,606	13,457	1.10%	26025070	15,063	14,985	0.52%
12015020	13,540	13,238	2.23%	26030060	14,935	14,876	0.40%
12015070	11,683	11,467	1.85%	26030080	13,972	13,899	0.52%
12020010	15,769	15,174	3.77%	26030110	12,859	12,826	0.26%
23080650	15,327	15,274	0.35%	26130020	15,896	15,783	0.71%
26180160	16,045	15,948	0.60%	26130170	16,211	16,053	0.97%



Estación	Registros iniciales	Registros Finales	% datos removidos	Estación	Registros iniciales	Registros Finales	% datos removidos
26180170	16,037	15,977	0.37%	26135110	11,378	11,314	0.56%
26185020	16,283	16,218	0.40%	53070020	12,850	12,820	0.23%
26190090	16,765	16,685	0.48%	23080390	17,610	17,498	0.64%
26230110	21,454	21,306	0.69%	23080760	14,504	14,414	0.62%
26230130	15,829	15,770	0.37%	23085080	14,584	14,498	0.59%
26230140	16,289	16,230	0.36%	23100040	14,510	14,423	0.60%
26230220	11,575	11,538	0.32%	27015190	16,480	16,407	0.44%
27010830	16,043	15,943	0.62%	35010020	16,734	16,660	0.44%
27010840	15,778	15,725	0.34%	35010040	16,439	16,365	0.45%
27010870	16,126	16,074	0.32%	35010070	13,251	13,183	0.51%
27010880	16,336	16,281	0.34%	35035020	26,804	26,626	0.66%
27010900	16,376	16,318	0.35%	35040010	15,278	15,260	0.12%
27015110	15,998	15,962	0.23%	35050010	16,457	16,353	0.63%
27020220	16,059	15,867	1.20%	35050020	16,673	16,591	0.49%
21010040	17,035	16,885	0.88%	35055010	16,793	16,634	0.95%
21010050	14,893	14,740	1.03%	15010010	15,930	15,802	0.80%
21010060	16,967	16,857	0.65%	15010040	14,371	14,293	0.54%
21010100	16,056	15,924	0.82%	15015050	19,443	15,652	19.50%
21010110	16,300	16,137	1.00%	15015060	16,296	16,188	0.66%
21010130	14,570	14,446	0.85%	23140070	11,209	11,064	1.29%
21010140	14,592	14,490	0.70%	23145020	13,974	13,840	0.96%
21010160	14,165	14,039	0.89%	23155040	25,842	25,433	1.58%
21010170	12,367	12,286	0.65%	24050060	20,922	20,543	1.81%
21010180	11,523	11,435	0.76%	24050100	15,188	14,985	1.34%
21010190	12,455	12,369	0.69%	24050110	15,184	14,956	1.50%
21010200	12,641	12,544	0.77%	24055030	14,664	14,465	1.36%
21010210	12,590	12,491	0.79%	21160030	17,595	17,247	1.98%
21010220	12,883	12,815	0.53%	21160200	12,477	12,344	1.07%
21010230	12,696	12,563	1.05%	21165030	14,542	14,164	2.60%
21010250	12,652	12,509	1.13%	21170040	11,672	11,380	2.50%
21015020	15,952	15,773	1.12%	21180160	16,047	15,351	4.34%
21015030	16,007	15,859	0.92%	21190290	14,503	14,088	2.86%
21020040	12,710	12,489	1.74%	21190300	12,443	12,240	1.63%
21020050	13,115	12,991	0.95%	21210120	11,299	11,124	1.55%
21030050	17,115	16,823	1.71%	21215080	16,974	15,806	6.88%
21030060	16,916	16,786	0.77%	21215140	12,363	11,532	6.72%
21040030	14,650	14,487	1.11%	21220040	14,613	13,980	4.33%
21040050	12,478	12,187	2.33%	21230060	14,649	13,296	9.24%
21050060	17,324	17,193	0.76%	21230110	12,792	10,132	20.79%
21050110	14,370	14,222	1.03%	21245010	16,675	16,278	2.38%
21050190	12,916	12,829	0.67%	21245040	16,279	16,026	1.55%
21050230	14,024	13,925	0.71%	21255080	14,716	13,796	6.25%
21050240	14,145	14,033	0.79%	22050040	17,620	17,337	1.61%
21050250	14,436	14,315	0.84%	22060070	15,571	15,191	2.44%
21050310	11,590	11,564	0.22%	22060100	13,042	12,304	5.66%
44035030	13,709	13,612	0.71%	22065040	11,351	11,252	0.87%
44010020	15,969	15,888	0.51%	21190090	20,262	20,043	1.08%
44010100	11,935	11,871	0.54%	21190240	15,454	15,231	1.44%
44010110	11,192	11,154	0.34%	21190330	12,347	12,218	1.04%
44015010	15,971	15,900	0.44%	35060090	18,343	18,234	0.59%
44015030	15,698	15,604	0.60%	35060120	15,314	15,245	0.45%
44015040	10,993	10,919	0.67%	35060130	14,409	14,335	0.51%

Estación	Registros iniciales	Registros Finales	% datos removidos	Estación	Registros iniciales	Registros Finales	% datos removidos
				35060140	15,675	15,618	0.36%

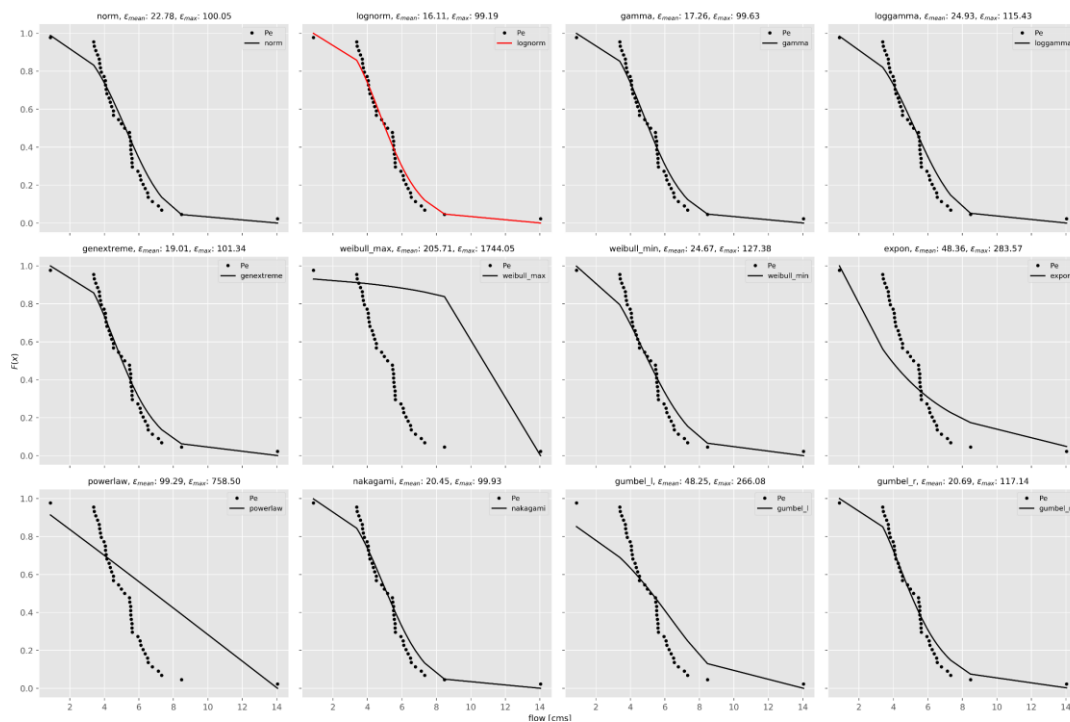
En el Anexo **A5** se adjuntan los resultados obtenidos, así como también las series tratadas y utilizadas finalmente en los ejercicios de modelación y correlación.

#### 4.4 CARACTERIZACIÓN PROBABILÍSTICA

Los resultados de los ajustes de las funciones de distribución teóricas marginales para las 30 estaciones hidrológicas y 129 estaciones meteorológicas se realizaron mediante programación en Python y uso de la librería Scipy. Se realizaron alrededor de 700.000 evaluaciones que representan algo más de 50.000 salidas gráficas de los ajustes de la forma en que se presentan en la Figura 23 y la Figura 24 para las series hidrológicas y meteorológicas respectivamente.

El consolidado de los mejores ajustes y sus parámetros se adjuntan a este documento en el Anexo **A6** Ajustes CDF tanto para las estaciones de precipitación como para las de caudal.

En general, los ajustes presentaron desempeños aceptables en los periodos de alta humedad dado que el número de datos nulos o magnitudes cercanas a cero es mucho menor en comparación con las épocas de baja humedad, especialmente para la variable precipitación, en donde los eventos no lluviosos en los periodos de baja humedad pueden representar alrededor del 40% al 50% del conjunto estadístico.



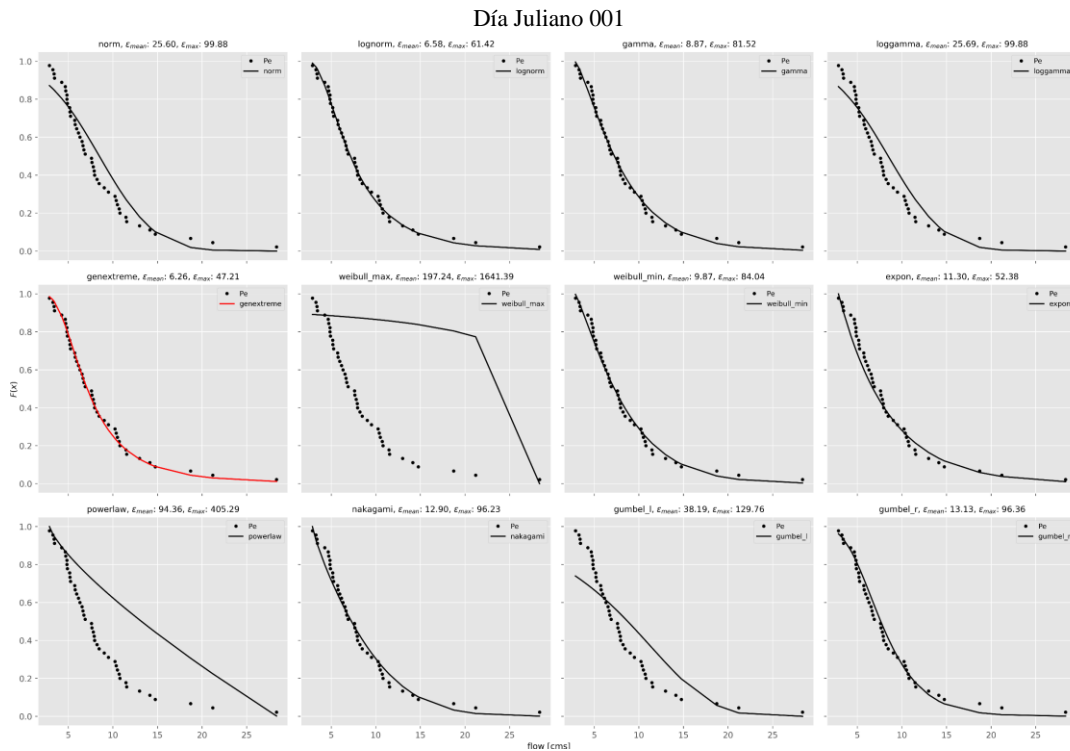
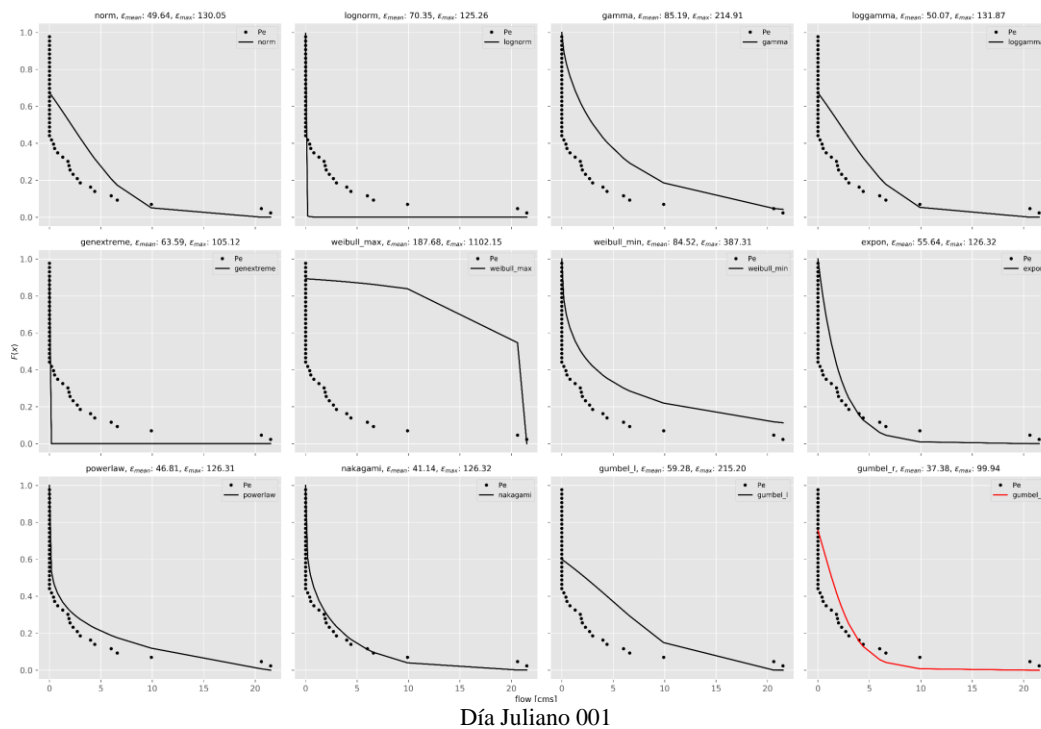


Figura 23. Ajustes de las funciones de distribución teóricas para la estación hidrológica 21017050, arriba el conjunto del día juliano 1 (época seca), abajo día juliano 105 (época húmeda).



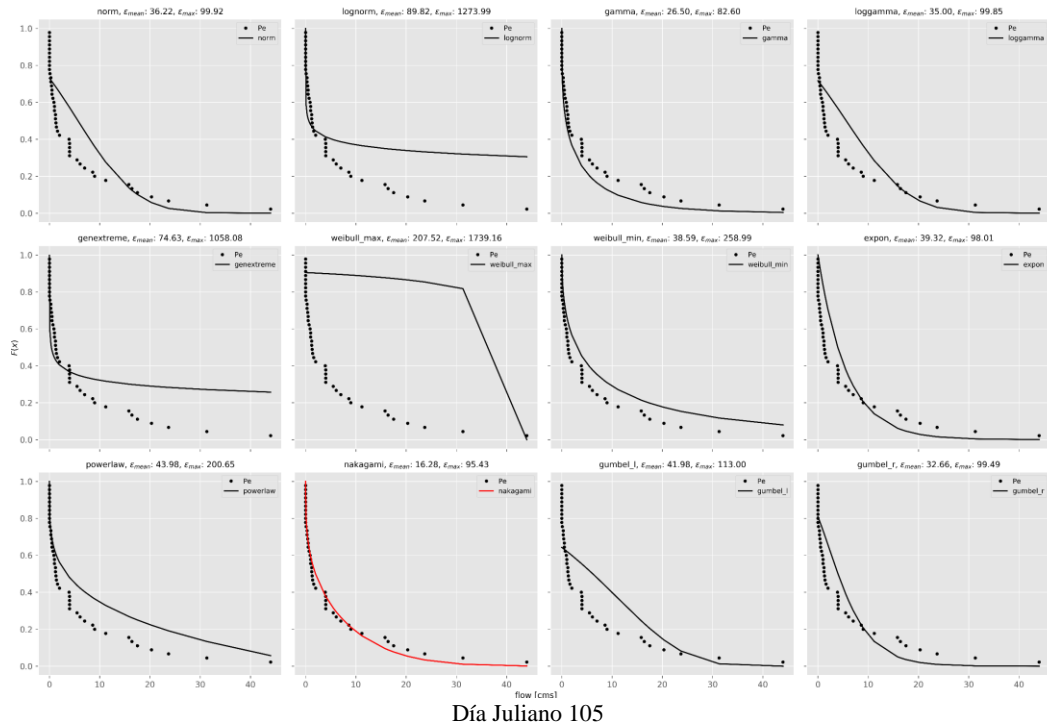


Figura 24. Ajustes de las funciones de distribución teóricas para la estación meteorológica (registros de precipitación total diaria) 21015030, arriba el conjunto del día juliano 1 (época seca), abajo día juliano 105 (época húmeda).

En general, la distribución que mejor ajustó a los datos de caudal medio diario fue la generalizada de extremos en más del 50% de las veces, es decir, en más de la mitad del año, seguida de las distribuciones log-normal, exponencial y gamma. Este resultado se mantuvo para las estaciones de las áreas hidrológicas Caribe (AH 1), Magdalena-Cauca (AH 2), Orinoco (AH 3) y Amazonia (AH 4). En las estaciones hidrológicas del Pacífico (AH 5), esta situación no es tan clara a favor de la distribución generalizada de extremos, a pesar de que en la mayoría de los conjuntos presentó el mejor desempeño, otras distribuciones como log-normal, gumbel con asimetría derecha, exponencial, gamma y nakagami se alternan con una relativa homogeneidad durante el año hidrológico.

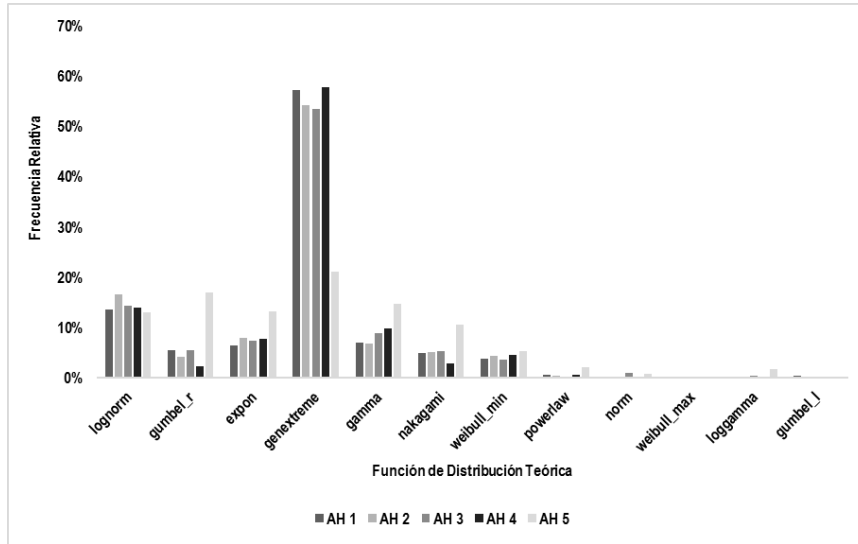


Figura 25. Resumen de los ajustes de las funciones de distribución para las series hidrológicas (caudal)

Los mismos resultados para las series meteorológicas de precipitación diaria total, indican un predominio de la función gumbel con asimetría derecha de alrededor del 35% al 45% del tiempo seguida de la función nakagami con un 17% al 30% de las veces. La distribución nakagami es un caso especial de la distribución gamma (Julian Cheng & Beaulieu, 2001), en menor medida aparecen las funciones log-gamma, gumbel con asimetría izquierda y generalizada de extremos.

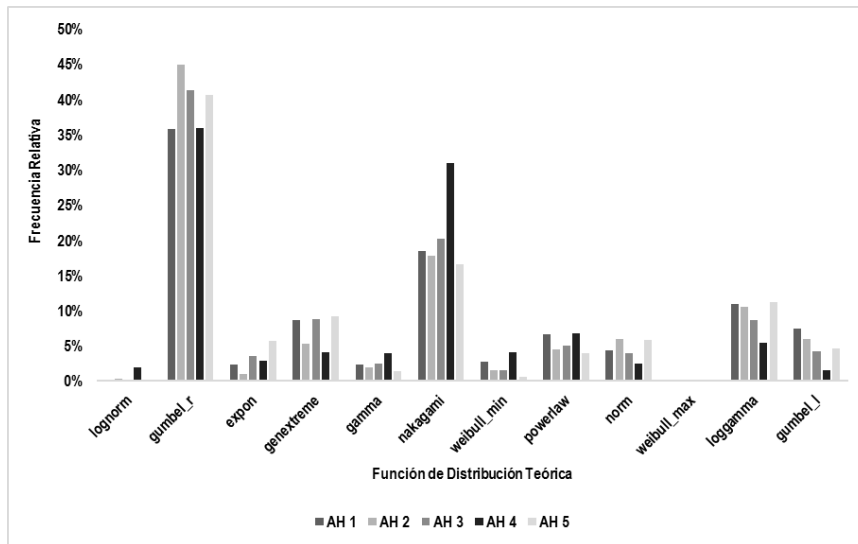


Figura 26. Resumen de los ajustes de las funciones de distribución para las series meteorológicas (precipitación)

#### 4.5 IDENTIFICACIÓN DE PARÁMETROS DE LA FPK

Los ejercicios de optimización para la identificación de los parámetros de la ecuación de FPK se realizaron en el lenguaje de programación científica Python mediante el uso de librerías especializadas como Scipy, Pandas y Numpy.

Los primeros intentos por encontrar soluciones al ejercicio de modelación demandaron una gran cantidad de recursos computacionales y tiempos de procesamiento que hacían inviable la estimación de parámetros que cumplieran con criterios de desempeño aceptables. Vale la pena recalcar que en este momento metodológico se estiman un conjunto de 18 parámetros que solucionan el problema inverso para cada evolución temporal de la curva de densidad de probabilidad empírica condicionada, es decir, 365 conjuntos de parámetros por unidad de análisis. El escenario anterior configura entonces 10.950 realizaciones del algoritmo de optimización matemática, esto sin contar, la cantidad de llamados de la función objetivo lo cual puede entonces significar millones de ejecuciones del algoritmo de optimización.

La situación anterior implicó meses de procesamiento en equipos de alto rendimiento sin llegar a resultados satisfactorios. En las distintas pruebas realizadas se logró identificar que ciertas aproximaciones de las magnitudes de los parámetros iniciales generaban una gran cantidad de llamados de la función objetivo ( $> 50.000$  llamados de la función), sin alcanzar los criterios de convergencia de la solución establecidos en esta investigación en el orden  $1E-03$  tanto del valor de la función objetivo como el de los parámetros. De esta forma fue necesario el diseño de un método que permitiera realizar una aproximación al conjunto de parámetros semilla, es decir, establecer un procedimiento para la estimación de los valores iniciales de la ecuación de FPK que definen los coeficientes de advección y difusión como se describe en la Ecuación 34, esto es optimización supervisada. De acuerdo con Gardiner y Sveshnikov (Gardiner, 2009; A. Sveshnikov, 1968) y para el caso unidimensional, estos coeficientes pueden ser determinados como:

$$A(Q, t) = \lim_{\Delta t \rightarrow 0} \frac{E[\Delta Q | Q]}{\Delta t}$$

$$B(Q, t) = \lim_{\Delta t \rightarrow 0} \frac{E[\Delta Q^2 | Q]}{\Delta t}$$

Ecuación 42. Definición de los coeficientes de advección y difusión  
Tomado de: (Domínguez & Rivera, 2010)

A partir de los enunciados de Gardiner y Sveshnikov, se planteó entonces un proceso de optimización de dos pasos, en el primero, se busca identificar los parámetros que definen al vector de advección ( $A_1, A_2$ ) y la matriz de difusión ( $B_{11}, B_{12}, B_{21}, B_{22}$ ). En el segundo, a partir de las aproximaciones de  $A$  y  $B$  del paso anterior, se identifican los valores del conjunto de 18 parámetros de la ecuación de FPK que explican la transición para dos pasos de tiempo consecutivos.

La aplicación del esquema anterior de optimización redujo en más de 160 veces el tiempo de procesamiento computacional e identificación de los parámetros óptimos, y de esta manera viabilizar el esquema numérico planteado. A continuación, se detalla el procedimiento de optimización:

### 4.5.1 Pre-optimización

El proceso de pre-optimización busca aproximar las magnitudes del conjunto de los 18 parámetros  $\{k_1, k_2, \dots, y_2, y_3\}$  que definen los vectores  $A_1, A_2, \dots, B_{22}$ , de tal forma que la discretización temporal resultante debido a la aplicación de la condición de estabilidad (Ecuación 36 y Ecuación 37) del esquema numérico no afecte la viabilidad del proceso de optimización, es decir, que la solución numérica se obtenga en el menor número de pasos temporales.

$$\begin{aligned}
 A_1(q, t) &= k_1q + k_2q^2 + k_3 \\
 A_2(q, t) &= k_4q + k_5q^2 + k_6 \\
 B_{11}(q, t) &= g_1q + g_2q^2 + g_3 \\
 B_{22}(q, t) &= g_4q + g_5q^2 + g_6 \\
 B_{12}(q, t) &= x_1q + x_2q^2 + x_3 \\
 B_{22}(q, t) &= y_1q + y_2q^2 + y_3
 \end{aligned}$$

Ecuación 43. Conjunto de parámetros a optimizar

Para los propósitos de esta investigación el número máximo de pasos temporales entre una transición de la curva de densidad condicional empírica y otra consecutiva fue establecido en 40,000. El valor anterior fue determinado a partir de un análisis del costo computacional de la solución de una transición con valores controlados de  $A$  y  $B$ , modificando las demás variables que son función de la condición de estabilidad como el  $\Delta Q_1$  y  $\Delta Q_2$ . Como resultado se obtuvo que la discretización espacial que lograba tiempos de computo asequibles se daba con  $nQ_1 = nQ_2 = 50$  nodos.

A continuación, se construyeron los vectores  $A$  y  $B$  a partir de los postulados de Gardiner y Sveshnikov, así:

$$\begin{aligned}
 A_i(q_i, t) &= \frac{E[Q_i^{k+1}] - E[Q_i^k]}{q_i} \\
 B_{ij}(q_{ij}, t) &= \frac{p(Q_i^{k+1}|Q_j^{k+1}) - p(Q_i^k|Q_j^k)}{q_i * q_j}
 \end{aligned}$$

Ecuación 44. Aproximación del vector de advección y la matriz de difusión para la estimación de los parámetros  $\{k_1, k_2, \dots, y_2, y_3\}$

Donde  $i, j$  son los descriptores espaciales en este caso varían entre 1 y 2; y  $k$  es el descriptor temporal. A partir de las aproximaciones de  $A$  y  $B$  se estiman los parámetros  $\{k_1, k_2, \dots, y_2, y_3\}$  de la Ecuación 43 mediante la aplicación del algoritmo de optimización “*minimize*” de Scipy implementado en Python y descrito en la metodología.

Como en la mayoría de los esquemas de optimización, la función exige se especifiquen los supuestos valores iniciales de los parámetros a optimizar. Como se ha reiterado, se busca establecer relaciones entre los parámetros de la ecuación de FPK y agentes externos e internos en la cuenca hidrográfica como el clima y la geomorfología respectivamente. Por lo tanto, el espacio de búsqueda para la fase de pre-optimización se definió entre los valores mínimos y máximos de los índices morfométricos estimados y expuestos en este capítulo en el apartado de caracterización



morfométrica. Teniendo en cuenta que los polinomios que explican la advección y difusión (Ecuación 43) contienen términos de segundo orden, el espacio de búsqueda se amplió a la parte negativa de la recta de los reales en el mismo dominio de las magnitudes de los índices morfométricos. Estos ejercicios de optimización resultan mucho menos complejos dado que la expresión matemática (polinomio de segundo orden) es simple y, por lo tanto, los tiempos de pre-optimización son significativamente menores en comparación a los de optimización. De esta forma, para los valores iniciales de los 18 parámetros a optimizar en esta fase, se generaron 5,000 conjuntos de parámetros semilla de forma aleatoria mediante una distribución uniforme en el dominio de búsqueda descrito. Sin embargo, en ejercicios numéricos preliminares, se identificó que al menos para las cuencas en análisis, los conjuntos de parámetros en la fase de pre-optimización existen en un rango entre  $[-100, 100]$  para valores de error aceptable (en esta ocasión se utilizó la raíz cuarta del promedio de los errores a la cuarta [R4MS4E] como medida del desempeño de la optimización). Consecuentemente, el espacio de búsqueda para la fase de pre-optimización fue reducida a este intervalo.

#### 4.5.2 Optimización

La verificación del código desarrollado se evaluó con la curva de densidad de probabilidad conjunta empírica para un sistema bivariado asumiendo independencia entre las variables. La Figura 27 muestra la evolución temporal de la curva de densidad de probabilidad conjunta de datos de una estación hidrológica a resolución mensual, en donde las dimensiones que definen al sistema son los valores del caudal en el mes  $i - 1$  ( $Q_1$ ) y en el mes  $i$  ( $Q_2$ ),  $p_{Q_1 Q_2}$  es la probabilidad de que sucedan los eventos  $Q_1$  y  $Q_2$ . Como se puede apreciar el esquema numérico planteado en diferencias finitas reproduce la evolución de la curva en forma estable y además en tiempos factibles a el desarrollo de este trabajo. Los análisis preliminares incluyeron la verificación del volumen bajo la superficie en el ejercicio controlado, en donde bajo diferentes esquemas como por ejemplo variaciones de la malla espacial, el esquema numérico conserva la masa en alrededor de un 96%.

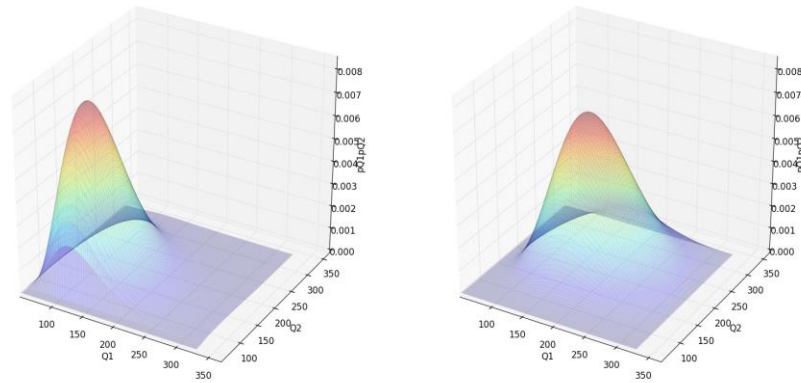


Figura 27. Evolución de la curva de densidad de probabilidad bivariado mediante el esquema de diferencias finitas explícitas.

Los parámetros obtenidos en el procedimiento anterior se constituyen en los valores iniciales para el proceso definitivo de optimización en donde se identifica el mejor conjunto de parámetros del operador matemático que reproduce con desempeños aceptables la evolución de la curva de densidad de probabilidad condicionada.

Por otra parte, de la Ecuación 36 y Ecuación 37, fácilmente se deduce que la discretización temporal es función de la discretización espacial asumiendo constantes la advección y difusión máximas, la primera de manera lineal y la segunda en forma cuadrática. Por lo tanto, es importante la definición del dominio espacial y su discretización ( $\Delta q$ ) ya que en el caso más estricto (cuando el proceso difusivo define la evolución de la curva,  $Pe \leq 3$ ) los nodos temporales se incrementan de forma cuadrática aumentando en dos órdenes de magnitud el tiempo de computo. Para esta investigación se definió la discretización temporal como:

$$nq_1 = nq_2 = 50$$

Ecuación 45. Discretización espacial del problema de modelación.

Es decir, que la malla espacial estará definida por 2,500 celdas, esta resolución espacial permite caracterizar con buen detalle las curvas de densidad de probabilidad condicionada.

Otro aspecto importante en la realización del código es la solución del esquema explícito mediante operaciones vectoriales, o como se denomina en Python, “*Fancy Slicing*”, en donde aprovechando que el esquema es explícito se puede ejecutar de manera vectorial cada paso temporal, y, por lo tanto, se elimina el costoso recorrido de la matriz celda a celda. Algunas evaluaciones del costo computacional de la aplicación de diferentes esquemas de solución para la ecuación de FPK en el caso unidimensional para esquemas numéricos explícitos e implícitos se pueden consultar en el trabajo de Domínguez (Domínguez et al., 2014).

Todos los aspectos comentados fueron implementados en un código desarrollado en Python 2.7 mediante el uso de librerías especializadas como Numpy, SciPy, Pandas, GDAL, Mayavi & MultiProcessing todas con licencia OpenSource y disponibles en la distribución de Continuum

Analytics Inc (<https://www.continuum.io/>) Anaconda versión 4.3.1 (<https://www.continuum.io/downloads>) a excepción de Mayavi que está disponible en la distribución de Enthought Scientific Computing Solutions Inc (<https://www.enthought.com/products/canopy/>) Canopy.

Los códigos desarrollados fueron ejecutados en los equipos de alto rendimiento “*Excelsior I y II*” en el Nodo Regional de Modelación Ambiental (NRMA) operado por el grupo de investigación Dinámica de Sistemas GeoAgroAmbientales (DSGAA) del Instituto de investigación CENIGAA en Neiva departamento del Huila. Las características de los equipos relevantes para el desarrollo de esta investigación son:

- 2 CPU Intel® Xeon® E5-2683v3 con 14 núcleos lógicos y 28 hilos de procesamiento, velocidad de reloj de 2.0 GHz y 35 MB de SmartCache, para un total de 112 hilos de procesamiento.
- 256 GB de memoria RAM DDR4 disponibles,
- 10 TB de almacenamiento en discos HDD de 7200 rpm.

El tiempo de cómputo para la estimación del conjunto de parámetros de la ecuación de FPK para las 365 transiciones temporales y las 30 unidades de análisis seleccionadas son estimados en cerca de 120 horas en una de las máquinas descritas (*Excelsior II*), es decir, en un equipo con características promedio (8 hilos de procesamiento) le tardaría cerca de 35 días el desarrollar la misma tarea.

Sin embargo, llegar a este resultado implicó alrededor de 2,160 horas de los equipos descritos en alrededor de 15 ejercicios numéricos y cerca de 750 millones de evaluaciones de la función objetivo, lo que a un computador con características promedio le tomaría cerca de 27 años.

La función objetivo fue construida a partir de la métrica de desempeño raíz cuarta del promedio de los errores a la cuarta (R4MSE4) como se describe a continuación:

$$R4MSE4 = \sqrt[4]{\frac{\sum_{i=1}^n (pQ_{obs} - pQ_{sim})^4}{n}}$$

Ecuación 46. Función objetivo utilizada en la solución del problema inverso de modelación.

### 4.5.3 Parámetros óptimos

Como resultado se obtuvieron un total de 197,100 parámetros que describen las 365 evoluciones de la curva de densidad de probabilidad empírica condicionada para los 30 casos de estudio (18 parámetros de la ecuación de FPK por 365 transiciones por 30 casos de estudio).

De las 10,950 transiciones realizadas, en 26 ocasiones (0.24%) el esquema de pre-optimización no logró proporcionar buenas semillas, de tal forma que cada llamado de la función objetivo superaba el límite establecido de la discretización temporal (máximo 40, 000 pasos de tiempo) y/o superaban el criterio de máximo llamado de la función objetivo por parte del algoritmo de optimización (fijado en 20,000 llamados de la función) y, por lo tanto, no se considera una optimización exitosa. El resultado se presenta en la siguiente tabla:

Tabla 14. Realizaciones del proceso de optimización NO exitosas

UAH	Optimización NO Exitosa	Años de registro	% Datos Faltantes	Calidad	NOMBRE	CORRIENTE
21167090	1	22	1.9%	NO CUMPLE	HDA EL RUBY AUTOM	CUNDAY
26137120	1	23	0.1%	NO CUMPLE	REINA LA	SAN EUGENIO
15017030	22	37	8.7%	CUMPLE	MINCA	GAIRA
12027050	1	36	1.1%	CUMPLE	PUEBLO NUEVO	MULATOS
27017130	1	28	25.3%	NO CUMPLE	PALMICHAL G-8	GUADALUPE

El restante 99.976% correspondiente a *10,924 transiciones exitosas* se adjuntan en el Anexo **A7** – Resultados de Optimización. El archivo se estructura en función de cada transición ejecutada, la cual es indexada de acuerdo con el código de la estación (que define el área de estudio, la cuenca en estudio) y el número de transición de la curva. Para cada índice, se presentan los predictores de la ecuación de FPK representados en los días julianos y que se debe leer de la siguiente manera: *la curva condicionada por lo días julianos 5 y 6, es pronosticada por la curva condicionada por los días julianos 4 y 5, así, cada transición de la curva involucra 3 conjuntos estadísticos*. Para cada transición se presentan el conjunto de 18 parámetros  $\{k_1, k_2, \dots, y_2, y_3\}$ , y, por último, los desempeños obtenidos y algunos resultados del proceso de optimización como, por ejemplo, el resultado de la optimización (True o False, para indicar si la optimización es exitosa), el número de llamados de la función objetivo, y el número de cambio en el conjunto de parámetros probados antes de la convergencia del problema.

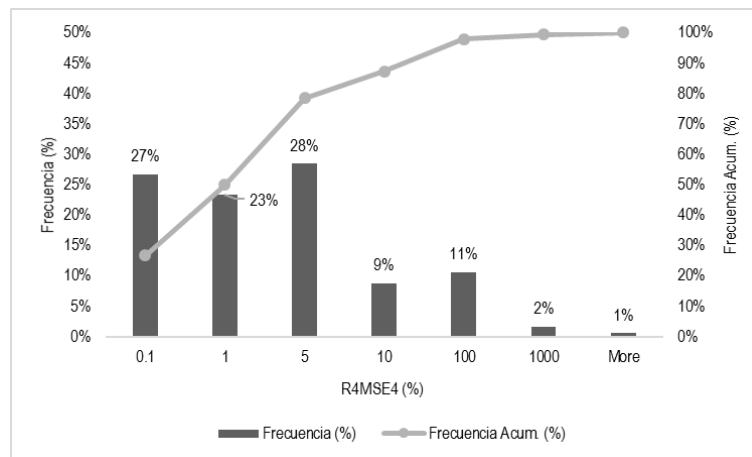


Figura 28. Resumen de los resultados del proceso de optimización

En general, los resultados obtenidos son satisfactorios según los desempeños obtenidos como se puede apreciar en la Figura 28. El 87.2% de las optimizaciones realizadas presentan un error relativo menor al 10 %. No obstante, cerca del 11 % de las optimizaciones presentan errores entre el 10 % y el 100%, y una franja marginal del 3 % presentan errores inadmisibles.

A continuación, se presentan algunos resultados para la unidad de análisis hidrográfica Gindrama (11027050) en el río Atrato. La estación Gindrama presentó las mejores métricas de desempeño

como se muestra en la figura abajo, en promedio las 365 transiciones se lograron con un error de optimización del 0.47%. Sin embargo, se aprecia que los resultados presentan un patrón temporal que se sincronizan con el régimen de precipitaciones en la cuenca, es decir, que las magnitudes absolutas de los errores de la función objetivo son mayores en las épocas húmedas que en las épocas secas, al menos para el segundo semestre del año.

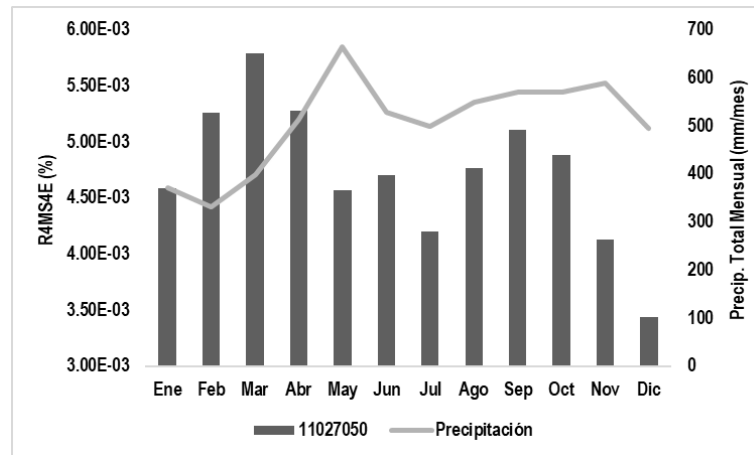


Figura 29. Valores de la función objetivo agregados a escala mensual para la unidad de análisis 11027050 y el régimen climático para la cuenca aferente a la estación hidrológica.

La Figura 30 presenta los resultados de simulación de las curvas de densidad de probabilidad condicionada y las funciones marginales para los datos observados (en azul) y simulados (en rojo) para dos momentos de tiempo en la estación 11027050, a la izquierda en la figura el día juliano 29 correspondiente al mes de febrero (época seca) y a la derecha el día juliano 147 (época húmeda). En la figura, se puede apreciar como el esquema numérico intenta reproducir la variabilidad de los datos observados concentrando su masa en las zonas de mayor probabilidad observada, lo anterior se puede corroborar con las funciones marginales.

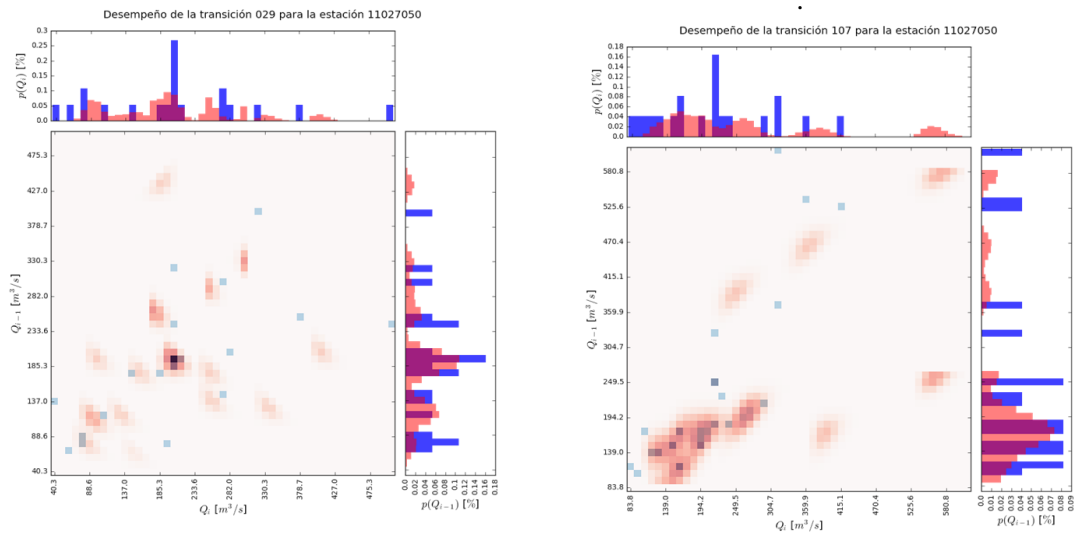


Figura 30. Desempeños obtenidos en el proceso de optimización matemática para la estación 11027050 sobre el río Atrato en los periodos seco (izquierda) y húmedo (derecha).

De la misma forma, la Figura 31 presenta los resultados obtenidos para el día juliano 195 (transición de época húmeda a seca) en la misma estación. Se puede apreciar que, en este caso, y para el 13% de los procesos de optimización, la optimización no presenta buenos desempeños con valores relativos de la métrica mayores al 100%.

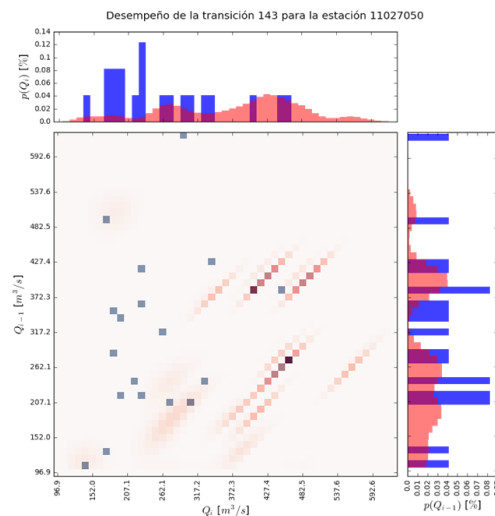


Figura 31. Desempeños obtenidos en el proceso de optimización matemática para la estación 11027050 sobre el río Atrato para un periodo de transición entre un régimen húmedo a seco

Esto puede obedecer a diferentes fuentes de incertidumbre como, por ejemplo, el mismo algoritmo de optimización, corto registro de los datos históricos, valores atípicos que los métodos de detección de anomalías no identificaron y hacen que el esquema numérico se desestabilice, entre otros. La tabla a continuación resume el orden de magnitudes encontrados en el proceso de optimización.

Tabla 15. Resumen de los parámetros de la ecuación de FPK para análisis de regresión lineal múltiple (versión completa en el Anexo A8)

<b>Parámetro</b>	<b>Promedio</b>	<b>Máximo</b>	<b>Mínimo</b>
<b>k1</b>	-5.248E-03	2.308E-02	-4.709E-02
<b>k2</b>	1.017E-04	1.177E-03	-1.538E-04
<b>k3</b>	2.488E-01	4.400E+00	-3.972E+00
<b>k4</b>	-4.063E-03	3.202E-02	-4.848E-02
<b>k5</b>	5.778E-05	7.674E-04	-3.989E-04
<b>k6</b>	-1.438E-01	4.472E+00	-1.111E+01
<b>g1</b>	-7.094E-05	1.444E-04	-1.010E-03
<b>g2</b>	2.751E-06	4.415E-05	-5.926E-06
<b>g3</b>	4.012E-04	5.085E-03	-1.391E-03
<b>g4</b>	-5.106E-05	2.760E-04	-4.068E-04
<b>g5</b>	1.466E-06	1.509E-05	-1.541E-05
<b>g6</b>	3.697E-04	2.788E-03	-6.861E-04
<b>x1</b>	-7.104E-05	1.444E-04	-1.010E-03
<b>x2</b>	2.751E-06	4.415E-05	-5.926E-06
<b>x3</b>	4.740E-04	5.085E-03	-1.391E-03
<b>y1</b>	-5.109E-05	2.760E-04	-4.068E-04
<b>y2</b>	1.466E-06	1.509E-05	-1.541E-05
<b>y3</b>	3.963E-04	2.788E-03	-6.861E-04

#### 4.6 ANÁLISIS DE LOS PARÁMETROS OBTENIDOS Y CARACTERÍSTICAS INTERNAS Y EXTERNAS DE LA CUENCA HIDROGRÁFICA

Mediante análisis por regresiones lineales múltiples se establecieron las relaciones entre los parámetros óptimos de la ecuación de FPK y las características internas y externas de las 30 cuencas hidrográficas seleccionadas para cada una de las transiciones que componen los conjuntos estadísticos de los días julianos. Como se ha expuesto, las características internas de los sistemas estudiados están representadas por las características morfométricas de las mismas unidades, mientras que las características externas son explicadas por los factores que perturban el sistema como por ejemplo la precipitación. A continuación, se presentan los resultados de las regresiones lineales múltiples del conjunto de 18 parámetros de la ecuación de FPK versus 26 índices morfométricos, y, del mismo conjunto de parámetros versus las esperanzas matemáticas de la precipitación total diaria para cada una de las 365 transiciones.

Las regresiones lineales fueron implementadas mediante la librería SciPy, específicamente el paquete estadístico *Ordinary Linear Regression* (OLS), el cual construye y evalúa la significancia de la regresión de acuerdo con lo descrito en el marco metodológico.

Tabla 16. Índices morfométricos incluidos en el análisis de regresiones

ÍNDICE MORFOMÉTRICO	CÓDIGO	ÍNDICE MORFOMÉTRICO	CÓDIGO
Ancho Medio de la Cuenca	P3	Pendiente Ponderada Cauce Principal	P15
Área	P4	Perímetro	P16
Coefficiente Sinuosidad	P5	Profundidad de Disección de la Cuenca	P17
Coefficiente de Compacidad	P6	Bransby-Williams	P18
Coefficiente de Forma	P7	Dir. Gen. Carreteras	P19
Densidad de Drenaje	P8	Giandotti	P20
Longitud Cauce Más Largo	P9	Kirpich	P21
Longitud Lineal del Cauce Más Largo	P10	Temez	P22
Longitud Media de Laderas	P11	Tiempo de Concentración Promedio	P23
Nivel de Disección del Relieve	P12	Tasa Bifurcación Promedio	P24
Strahler	P13	Tasa Bifurcación Nivel 01	P25
Pendiente Ponderada Cuenca	P14	Tasa Bifurcación Nivel 02	P26

##### 4.6.1 Parámetros FPK y morfometría

Con el fin de simplificar el análisis, los parámetros optimizados han sido promediados para cada sistema, es decir, el conjunto de datos para las regresiones es conformado por 30 conjuntos de parámetros (igual número al de las unidades de análisis) y cada conjunto por 18 parámetros (igual al número de parámetros de la ecuación de FPK). Vale la pena aclarar que, de acuerdo con los desempeños obtenidos en la optimización, el resumen de los parámetros presentados a continuación, corresponden exclusivamente a aquellos que obtuvieron un desempeño menor al 10%, el cual corresponde a un conjunto de 9,526 transiciones. Lo anterior con el fin de eliminar el ruido que pueden ofrecer las optimizaciones con malos desempeños. El conjunto de parámetros completo puede ser consultado en el Anexo A8 – **Parámetros FPK Promedio**.



En las pruebas realizadas, se evaluaron tres niveles de significancia para la prueba estadística de la regresión así:  $\alpha = \{0.10, 0.05, 0.01\}$ . Los resultados se consolidan a continuación en las Tabla 17 a Tabla 19. Los colores rojos indican significancia estadística entre el índice morfométrico y el parámetro de la ecuación de FPK respectivo.

Tabla 17. Resultados de las regresiones lineales múltiples con nivel de significancia  $\alpha = 0.10$

FPK PARÁMETROS		ADVECCIÓN						DIFUSIÓN											
Índices Morfométricos		k1	k2	k3	k4	k5	k6	g1	g2	g3	g4	g5	g6	x1	x2	x3	y1	y2	y3
P3	Ancho Medio de la Cuenca																		
P4	Área																		
P5	Coefficiente Sinuosidad																		
P9	Longitud Cauce Más Largo																		
P10	Longitud Lineal del Cauce Más Largo																		
P11	Longitud Media de Laderas																		
P13	Strahler																		
P15	Pendiente Ponderada Cauce Principal																		

Con una confiabilidad del 90 %, varios índices morfométricos como el ancho de la cuenca, el área de drenaje, algunos factores de forma del cauce como el coeficiente de sinuosidad, la longitud lineal del cauce más largo, la longitud más larga del cauce, y la pendiente ponderada de la corriente presentan relación lineal significativa con todos los 18 parámetros de la FPK. También, algunos parámetros que tienen relación con la morfología de la cuenca y sus suelos como la longitud media de laderas y el número de orden jerárquico máximo de las corrientes (Strahler).

Tabla 18. Resultados de las regresiones lineales múltiples con nivel de significancia  $\alpha = 0.05$

FPK PARÁMETROS		ADVECCIÓN						DIFUSIÓN											
Índices Morfométricos		k1	k2	k3	k4	k5	k6	g1	g2	g3	g4	g5	g6	x1	x2	x3	y1	y2	y3
P3	Ancho Medio de la Cuenca																		
P4	Área																		
P5	Coefficiente Sinuosidad																		
P10	Longitud Lineal del Cauce Más Largo																		
P13	Strahler																		
P15	Pendiente Ponderada Cauce Principal																		

Para un nivel de confiabilidad del 95 %, se mantienen como predictores válidos estadísticamente los parámetros de forma de la cuenca como el ancho medio y su área de drenaje, estos últimos exclusivamente con los parámetros de advección de la ecuación de FPK. Los índices relacionados con procesos hidráulicos como el coeficiente de sinuosidad, la longitud lineal del cauce más largo y la pendiente ponderada del mismo explican varios parámetros de la advección y difusión.

Tabla 19. Resultados de las regresiones lineales múltiples con nivel de significancia  $\alpha = 0.01$

FPK PARÁMETROS		ADVECCIÓN						DIFUSIÓN											
Índices Morfométricos		k1	k2	k3	k4	k5	k6	g1	g2	g3	g4	g5	g6	x1	x2	x3	y1	y2	y3
P4	Área																		
P5	Coefficiente Sinuosidad																		

Finalmente, con una confiabilidad del 99 % son el área de drenaje de la cuenca y el índice hidráulico coeficiente de sinuosidad los que explican los parámetros  $k_2$  y  $k_6$  de la ecuación de FPK respectivamente. El resultado detallado de las regresiones realizadas es presentado en el Anexo **A9 – Regresiones Lineales**.

#### 4.6.2 Parámetros FPK y clima

Para la evaluación de la influencia del clima sobre la variabilidad hidrológica en las cuencas hidrográficas representada en este estudio a través de la ecuación de FPK, se construyeron regresiones lineales entre el conjunto de 18 parámetros promedio de las 365 transiciones correspondientes a las  $\sigma$ -álgebra establecida, y, los valores esperados y varianzas promedio de las estaciones de precipitación seleccionadas para cada unidad de análisis descritas al comienzo del capítulo de resultados en la descripción de las áreas de estudio [pág. 53]. Los resultados de las regresiones de los parámetros de la ecuación versus el valor esperado promedio de la precipitación de las estaciones para cada unidad de análisis se presentan en la tabla a continuación (Tabla 20). Los resultados indican que el único parámetro de la ecuación de FPK que puede ser explicado por las variables climáticas (valor esperado y varianza), es el parámetro  $k_3$  asociado al vector de advección, pues en cerca del 49 % de los conjuntos estadísticos se establecieron regresiones válidas con una confiabilidad del 90 %, mientras que para una confiabilidad del 95% solo el 24 % de los conjuntos estadísticos presentó regresiones estadísticamente significativas. Para un nivel de confiabilidad del 99 %, el análisis de regresión no estableció ninguna relación con ningún parámetro de la ecuación de FPK.

Tabla 20. Número de regresiones estadísticamente significativas agregados a escala mensual del parámetro " $k_3$ " de la ecuación de FPK y el valor esperado de la precipitación.

$\alpha$	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	Total
0.10	28	18	13	3	3			3	22	31	30	27	178
0.05	7	5	4						6	25	22	19	88
0.01													0

El mismo ejercicio para los valores de los parámetros de la ecuación versus la varianza promedio de precipitación total diaria para las unidades de análisis arrojó como resultados muy pocas regresiones válidas, menos de 50 regresiones válidas con una significancia del 10% y 9 regresiones con una significancia del 5%, aunque las regresiones válidas se presentan en los periodos de baja y alta humedad.

Tabla 21. Número de regresiones estadísticamente significativas agregados a escala mensual del parámetro " $k_3$ " de la ecuación de FPK y la desviación estándar de la precipitación total promedio

$\alpha$	Ene	Feb	Mar	Abr	May	Jun	Jul	Ago	Sep	Oct	Nov	Dic	Total
0.10	6	4	4	2				1	7	9	6	8	47
0.05	2	2	2						1		1	1	9
0.01													0

### 4.6.3 Parámetros FPK, clima y morfometría

A partir de los resultados obtenidos en las relaciones lineales independientes encontradas entre los parámetros de la ecuación de FPK y, los índices morfométricos (agente interno) y el clima (agente externo), se realizaron evaluaciones para cada conjunto estadístico {1, 2, 3, ..., 364, 365} con los respectivos valores de los parámetros óptimos identificados de la FPK y los índices morfométricos junto con el valor esperado de la precipitación y la varianza diaria de cada conjunto estadístico. Dado que las magnitudes de los predictores en la regresión varían en distintos órdenes de magnitud, este análisis utilizó la versión normalizada de los datos, tanto de los predictores (valor esperado de la precipitación, varianza de la precipitación y 24 índices morfométricos) como de la variable dependiente (18 parámetros de la FPK). En la Tabla 22 se presentan los resultados de las regresiones lineales bajo los parámetros expuestos, los parámetros P3 a P26 corresponden a los índices morfométricos de acuerdo a la Tabla 16,  $PT$  y  $\sigma (PT)$  son el clima representado por el valor esperado de la precipitación total promedio diaria y su varianza respectivamente para cada conjunto estadístico y  $k1$  a  $y3$  son los parámetros de la ecuación de FPK. Los valores indican el porcentaje de veces en que se encontraron regresiones con significancia estadística ( $\alpha = 0.10$  en este caso) sobre el conjunto de días julianos {1, 2, 3, ..., 364, 365}, es decir, que entre mayor sea la magnitud, mayor es el número de veces en que el índice morfométrico o el clima explicaron el parámetro de la ecuación de FPK.

Tabla 22. Relaciones parámetros FPK versus índices morfométricos y clima con significancia de  $\alpha = 0.10$

FPK	ADVECCIÓN						DIFUSIÓN											
	k1	k2	k3	k4	k5	k6	g1	g2	g3	g4	g5	g6	x1	x2	x3	y1	y2	y3
P3	0.08	0.01	<b>0.38</b>	0.09	0.01	<b>0.39</b>	0.01	0.02	0.10	0.01	0.02	0.09	0.01	0.02	0.07	0.01	0.02	0.09
P4	<b>0.18</b>	0.01	<b>0.50</b>	<b>0.15</b>	0.01	<b>0.53</b>	0.00	0.01	<b>0.18</b>	0.01	0.00	<b>0.22</b>	0.00	0.01	<b>0.18</b>	0.01	0.00	<b>0.19</b>
P5	0.03	<b>0.07</b>	0.02	0.02	<b>0.07</b>	0.01	<b>0.10</b>	<b>0.12</b>	0.06	<b>0.10</b>	<b>0.11</b>	0.07	<b>0.10</b>	<b>0.12</b>	0.06	<b>0.10</b>	<b>0.11</b>	0.07
P6	0.01	0.02	0.01	0.02	0.03	0.01	0.02	0.04	0.01	0.01	0.04	0.01	0.02	0.04	0.01	0.01	0.04	0.01
P7	0.02	<b>0.07</b>	0.01	0.02	0.04	0.01	<b>0.07</b>	<b>0.09</b>	0.04	<b>0.07</b>	<b>0.08</b>	0.03	<b>0.07</b>	<b>0.09</b>	0.04	<b>0.07</b>	<b>0.08</b>	0.04
P8	0.02	0.02	0.05	0.02	0.02	0.05	0.01	0.02	0.02	0.03	0.05	0.03	0.01	0.02	0.02	0.03	0.05	0.03
P9	<b>0.05</b>	0.01	<b>0.11</b>	0.05	0.01	<b>0.12</b>	0.01	0.00	0.02	0.00	0.01	0.02	0.01	0.00	0.02	0.00	0.01	0.02
P10	0.02	0.01	0.01	0.02	0.01	0.03	0.04	0.03	0.02	0.02	0.04	0.04	0.04	0.03	0.02	0.02	0.04	0.03
P11	0.03	<b>0.03</b>	0.02	0.02	0.04	0.01	<b>0.06</b>	<b>0.05</b>	0.04	<b>0.06</b>	<b>0.06</b>	0.05	<b>0.06</b>	<b>0.05</b>	0.04	<b>0.06</b>	<b>0.06</b>	<b>0.06</b>
P12	<b>0.06</b>	0.01	0.08	<b>0.07</b>	0.02	0.07	0.01	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.01
P13	0.02	<b>0.05</b>	0.01	0.02	<b>0.05</b>	0.00	<b>0.06</b>	<b>0.10</b>	0.03	<b>0.06</b>	0.04	0.01	<b>0.05</b>	<b>0.10</b>	0.04	<b>0.06</b>	<b>0.04</b>	0.02
P14	0.03	<b>0.03</b>	0.05	0.02	0.02	0.03	0.03	0.04	0.03	0.03	0.03	0.04	0.03	0.04	0.03	0.03	0.03	0.06
P15	0.03	<b>0.04</b>	0.04	0.02	0.04	0.02	<b>0.10</b>	<b>0.11</b>	0.07	<b>0.12</b>	<b>0.12</b>	0.08	<b>0.10</b>	<b>0.11</b>	0.07	<b>0.12</b>	<b>0.12</b>	0.08
P16	0.03	0.01	0.07	0.03	0.01	0.07	0.00	0.01	0.01	0.00	0.01	0.02	0.01	0.01	0.02	0.00	0.01	0.02
P17	<b>0.05</b>	<b>0.02</b>	0.09	<b>0.07</b>	0.01	0.06	0.01	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.02	0.00	0.00	0.01
P18	0.01	0.00	0.02	0.00	0.01	0.02	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
P19	0.02	0.00	0.04	0.00	0.00	0.04	0.00	0.00	0.01	0.00	0.01	0.01	0.00	0.00	0.01	0.00	0.01	0.01
P20	0.01	0.00	0.02	0.01	0.00	0.01	0.01	0.01	0.01	0.02	0.02	0.01	0.02	0.01	0.02	0.02	0.02	0.01
P21	0.02	0.01	0.03	0.01	0.01	0.03	0.01	<b>0.03</b>	0.01	0.01	0.01	0.01	0.01	<b>0.03</b>	0.01	0.01	0.01	0.01
P22	0.02	0.01	0.04	0.01	0.00	0.03	0.01	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01	0.00	0.01	0.01
P23	0.01	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P24	0.02	<b>0.02</b>	0.02	0.02	0.02	0.02	0.02	0.02	0.01	0.02	0.03	0.02	0.02	0.02	0.02	0.02	0.03	0.02
P25	0.01	<b>0.04</b>	0.01	0.02	0.02	0.01	0.02	0.03	0.01	0.03	0.04	0.02	0.02	0.03	0.03	0.03	0.04	0.04
P26	0.04	<b>0.05</b>	0.05	0.03	<b>0.05</b>	0.04	<b>0.06</b>	<b>0.08</b>	0.02	<b>0.06</b>	<b>0.08</b>	0.03	<b>0.06</b>	<b>0.08</b>	0.01	<b>0.06</b>	<b>0.08</b>	0.04

FPK	ADVECCIÓN						DIFUSIÓN											
	k1	k2	k3	k4	k5	k6	g1	g2	g3	g4	g5	g6	x1	x2	x3	y1	y2	y3
PT	0.13	0.02	0.29	0.12	0.03	0.24	0.01	0.00	0.03	0.02	0.02	0.06	0.01	0.00	0.04	0.02	0.02	0.06
$\sigma$ (PT)	0.04	0.02	0.09	0.04	0.03	0.06	0.02	0.02	0.03	0.03	0.03	0.03	0.03	0.02	0.02	0.03	0.03	0.02

Los resultados obtenidos ratifican los resultados obtenidos en los ejercicios anteriores (parámetros FPK versus morfometría y parámetros FPK versus clima), en donde con una confiabilidad de la regresión del 90 %, los índices de forma ancho medio de la cuenca, área de drenaje e índices hidráulicos como la longitud del cauce más largo explican parcialmente los parámetros de advección, especialmente los parámetros de forma. En el caso de los parámetros difusivos, los índices de mayor relevancia (con regresiones válidas entre el 10 % y el 22 % de las transiciones) fueron el área de drenaje (índice de forma), el coeficiente de sinuosidad y la pendiente ponderada del cauce (índices hidrográficos relacionados con la hidráulica). Además, en al menos el 30% de las 365 transiciones el valor esperado de la precipitación diaria presentó influencia directa sobre los parámetros advectivos.

Tabla 23. Relaciones parámetros FPK versus índices morfométricos y clima con significancia de  $\alpha = 0.05$

FPK	ADVECCIÓN						DIFUSIÓN											
	k1	k2	k3	k4	k5	k6	g1	g2	g3	g4	g5	g6	x1	x2	x3	y1	y2	y3
P3	0.04	0.00	0.24	0.05	0.00	0.25	0.01	0.01	0.07	0.00	0.01	0.05	0.01	0.01	0.04	0.00	0.01	0.04
P4	0.15	0.01	0.44	0.11	0.00	0.49	0.00	0.00	0.15	0.00	0.00	0.12	0.00	0.00	0.14	0.00	0.00	0.13
P5	0.01	0.04	0.01	0.01	0.02	0.00	0.07	0.05	0.05	0.05	0.06	0.03	0.07	0.05	0.05	0.05	0.06	0.02
P6	0.00	0.01	0.01	0.01	0.01	0.00	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.02	0.00	0.01	0.01	0.00
P7	0.02	0.05	0.00	0.01	0.02	0.00	0.05	0.06	0.03	0.04	0.05	0.01	0.05	0.06	0.02	0.04	0.05	0.01
P8	0.01	0.01	0.04	0.01	0.00	0.03	0.01	0.00	0.01	0.02	0.04	0.02	0.01	0.00	0.01	0.02	0.04	0.01
P9	0.03	0.01	0.07	0.03	0.01	0.07	0.01	0.00	0.01	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.01
P10	0.01	0.01	0.01	0.01	0.01	0.01	0.03	0.01	0.01	0.02	0.01	0.01	0.02	0.01	0.00	0.01	0.01	0.01
P11	0.01	0.01	0.01	0.01	0.00	0.01	0.03	0.03	0.02	0.03	0.04	0.02	0.03	0.03	0.02	0.02	0.04	0.02
P12	0.04	0.00	0.04	0.05	0.01	0.04	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P13	0.01	0.03	0.00	0.00	0.01	0.00	0.04	0.05	0.02	0.03	0.02	0.00	0.04	0.05	0.02	0.03	0.02	0.01
P14	0.02	0.01	0.02	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.01	0.01	0.01	0.03
P15	0.01	0.02	0.02	0.01	0.02	0.01	0.05	0.06	0.03	0.06	0.06	0.05	0.05	0.06	0.02	0.06	0.06	0.04
P16	0.01	0.00	0.04	0.01	0.00	0.04	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01
P17	0.02	0.01	0.04	0.05	0.00	0.03	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01
P18	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00
P19	0.01	0.00	0.01	0.00	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01
P20	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.02	0.01	0.01	0.00
P21	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.00
P22	0.01	0.00	0.02	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01
P23	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P24	0.01	0.01	0.01	0.01	0.01	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
P25	0.00	0.02	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
P26	0.01	0.04	0.02	0.02	0.02	0.02	0.03	0.03	0.01	0.02	0.04	0.01	0.02	0.03	0.01	0.02	0.04	0.01
PT	0.10	0.01	0.21	0.09	0.01	0.16	0.00	0.00	0.02	0.01	0.00	0.04	0.00	0.00	0.02	0.01	0.00	0.04
$\sigma$ (PT)	0.02	0.01	0.04	0.03	0.01	0.03	0.01	0.01	0.00	0.02	0.01	0.01	0.01	0.01	0.00	0.02	0.01	0.01

La Tabla 23 presenta los resultados para un nivel de confiabilidad del 95 %. Con este nivel de significación de la regresión, se confirman los índices morfométricos de forma ancho de la cuenca y su área de drenaje sobre los componentes de advección y difusión de la ecuación de FPK. También el clima con el valor esperado de la precipitación total diaria sobre el componente advectivo.

Tabla 24. Relaciones parámetros FPK versus índices morfométricos y clima con significancia de  $\alpha = 0.01$

FPK	ADVECCIÓN						DIFUSIÓN											
	k1	k2	k3	k4	k5	k6	g1	g2	g3	g4	g5	g6	x1	x2	x3	y1	y2	y3
P3	0.01	0.00	0.05	0.01	0.00	0.05	0.00	0.00	0.02	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01
P4	0.08	0.01	0.36	0.06	0.00	0.37	0.00	0.00	0.07	0.00	0.00	0.07	0.00	0.00	0.07	0.00	0.00	0.06
P5	0.00	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.02	0.01	0.01	0.00	0.01	0.00	0.01	0.01	0.01	0.01
P6	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P7	0.01	0.02	0.00	0.00	0.01	0.00	0.02	0.02	0.01	0.02	0.02	0.00	0.02	0.02	0.01	0.02	0.02	0.01
P8	0.01	0.00	0.01	0.01	0.00	0.02	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01
P9	0.01	0.00	0.04	0.01	0.00	0.03	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00
P10	0.01	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
P11	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00
P12	0.02	0.00	0.02	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P13	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00
P14	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
P15	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01	0.01
P16	0.01	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P17	0.00	0.00	0.01	0.02	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P18	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P19	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P20	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.01	0.01	0.00	0.00	0.00	0.01	0.01	0.01	0.00
P21	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P22	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P23	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P24	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P25	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
P26	0.00	0.01	0.01	0.00	0.00	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00	0.00	0.00
PT	0.05	0.00	0.12	0.05	0.00	0.09	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01	0.00	0.00	0.01
$\sigma$ (PT)	0.00	0.00	0.01	0.01	0.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00

Para un nivel de confiabilidad del 99 %, solo el área de drenaje y el valor esperado de la precipitación total diaria explican el componente advectivo en al menos un 36 % de las veces para el caso del índice morfométrico y un 12 % para la precipitación total diaria promedio (Tabla 24).

## 5 ANÁLISIS DE RESULTADOS

Los criterios de selección de las áreas de estudio, las cuencas hidrográficas, permitieron identificar 20 cuencas en el territorio colombiano, sin embargo, de acuerdo con la zonificación hidrográfica del país realizada por el IDEAM, la zona hidrográfica del pacífico bajo los criterios iniciales planteados no aporta ningún área para este estudio. Por lo tanto, se definió flexibilizar el periodo de registro mínimo de la estación a 20 años con el fin de: *i*) incluir muestras de todas las zonas hidrográficas del país, y *ii*) aumentar el conjunto de datos para robustecer los análisis posteriores. De esta forma se logró consolidar un conjunto de 30 cuencas hidrográficas para el desarrollo de la presente investigación. Si bien el número de áreas de estudio es reducido, permitió establecer con soporte estadístico relaciones entre los componentes de la ecuación estocástica y parámetros físicos de la cuenca, así como también con agentes exógenos a ella.

Los resultados de la caracterización morfométrica fueron objeto de un análisis de componentes principales (PCA por sus siglas en ingles) el cual permitió identificar grupos de índices morfológicos que resumen la mayor cantidad de información de la variabilidad morfológica de las cuencas en estudio. Los índices morfométricos seleccionados a partir del PCA son el área de drenaje, el tiempo de concentración, coeficiente de compacidad, tasas de bifurcación, pendiente ponderada del cauce principal, coeficiente de sinuosidad y pendiente ponderada de la cuenca como se puede apreciar en la Figura 32 (izquierda). El análisis se repitió esta vez para verificar la heterogeneidad de las cuencas hidrográficas. En la misma figura a la derecha, se observan claramente dos grupos ortogonales entre sí. El grupo de la derecha (grupo 1) corresponde en su gran mayoría a cuencas en el área hidrográfica dos (magdalena cauca), mientras las estaciones del grupo 2 (izquierda) se ubican en las cinco áreas hidrográficas del país como se puede verificar en la Figura 33. Las cuencas del grupo uno son en general cuencas con un área de drenaje de entre 10 a 400 Km<sup>2</sup> y corresponden a cuencas de alta montaña en su mayoría, mientras que las cuencas del grupo 2 presentan áreas de drenaje mucho mayores por encima de los 800 Km<sup>2</sup> y hasta los 4000 Km<sup>2</sup> y, por lo tanto, presentan características morfométricas de alta montaña y valle ponderadas. La presencia de unidades de análisis con características morfométricas ortogonales enriquece el análisis de las relaciones entre la morfometría y los parámetros de la ecuación de FPK.

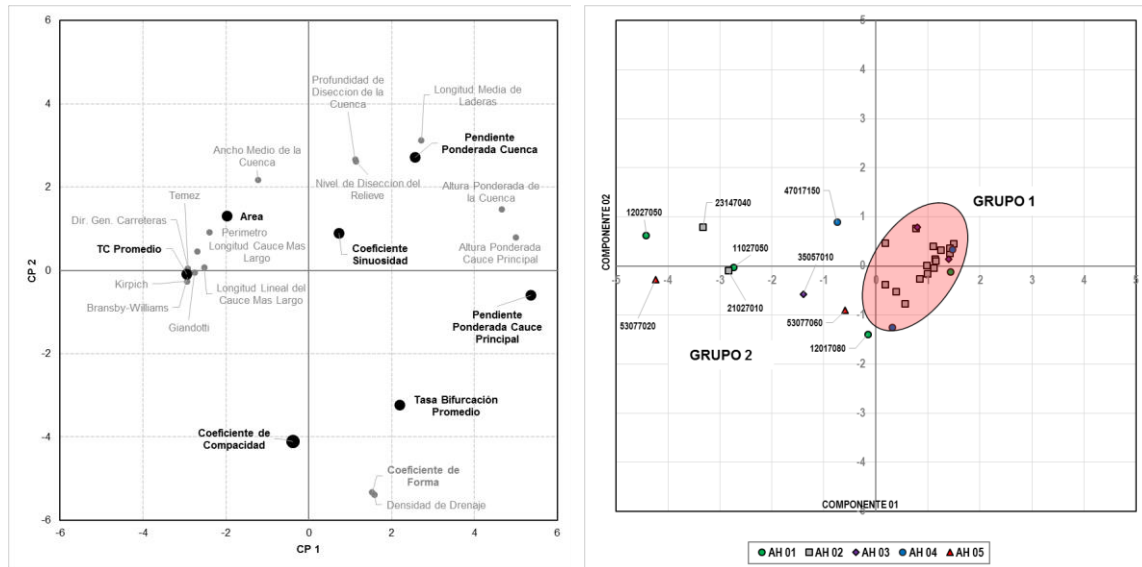


Figura 32. Análisis de componentes principales por índices morfométricos y por cuencas hidrográficas

En el país, la mayor parte de la red de monitoreo hidrológico inició en los años 70s y tuvo su máximo desarrollo en los 80s disminuyendo paulatinamente en las décadas siguientes, la situación anterior limita los registros a menos de cuarenta años, un poco más del mínimo exigido por la Organización Meteorológica Mundial (OMM, 2011) para la caracterización formal del clima. Lo anterior sin contar con los registros que por algún motivo la red deja de registrar y, además, los registros anómalos o espurios. De esta forma, la disponibilidad de información con la calidad requerida para este tipo de análisis es limitada en dos sentidos, el primero el número de estaciones que reúnen características de longitud de registro y calidad lo cual restringió el análisis a solo 20 unidades que cumplían los criterios estrictamente, y 10 más que cumplían en forma débil. En segundo lugar, el corto registro y la discontinuidad en las series temporales dificultan la caracterización de los eventos condicionados y dificultando el proceso de modelación matemática.

A pesar de que la cantidad de datos anómalos en las series de tiempo hidrológicas de las 30 estaciones seleccionadas fue muy baja, resulta de suma importancia este proceso ya que valores atípicos dentro de los conjuntos estadísticos inducen una “falsa variabilidad” en la serie lo cual se traduce en mayor gasto computacional y ruido en el proceso parametrización del operador estocástico.

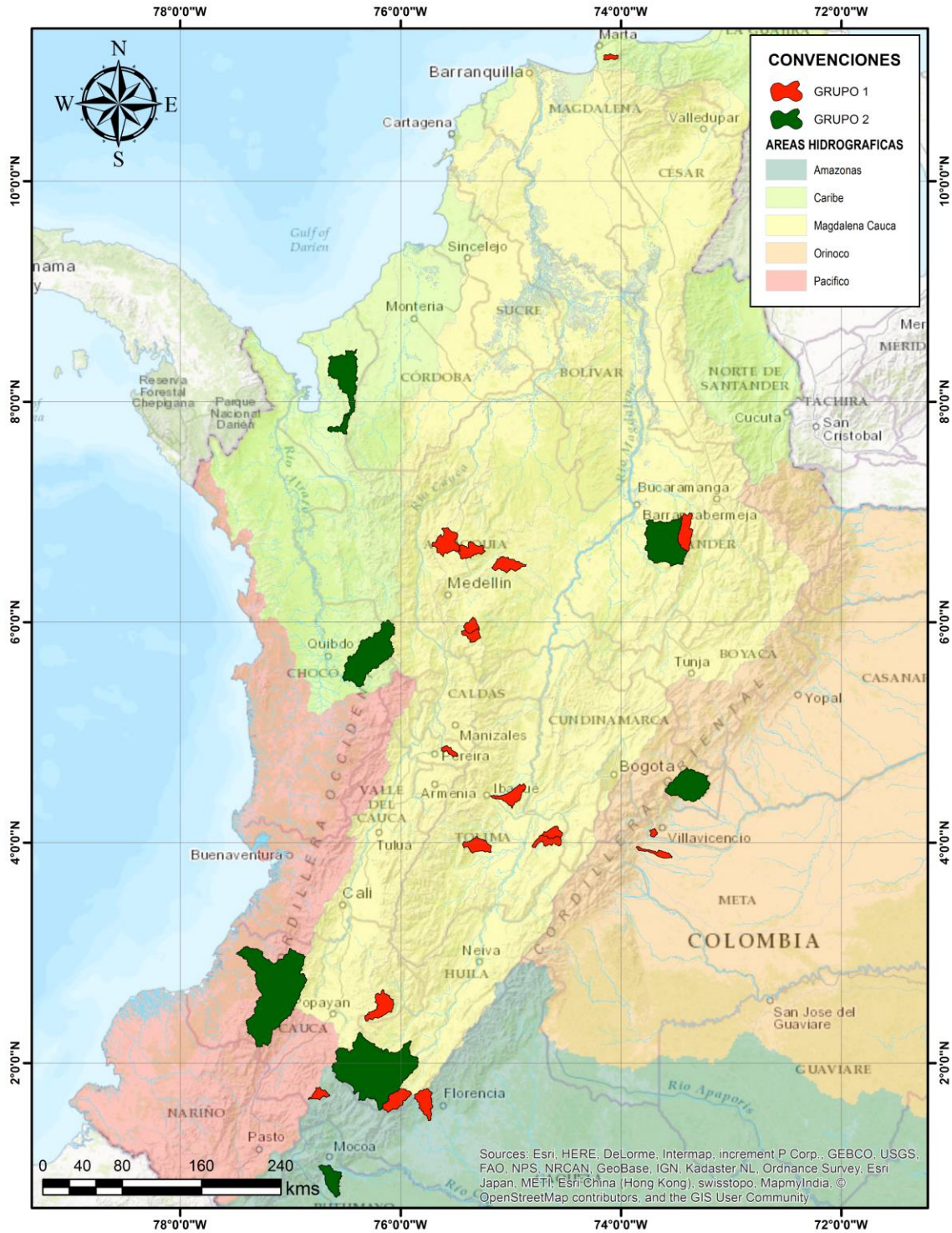


Figura 33. Distribución espacial de las cuencas hidrográficas grupo 1 y 2 identificadas en el análisis por PCA.



La caracterización probabilística de las series hidrológicas y meteorológicas (precipitación) arrojó que la distribución teórica predominante en las cuencas en estudio son la generalizada de extremos y, gumbel con asimetría derecha y la nakagami (un caso especial de la distribución gamma), lo anterior implica un proceso de transformación de la señal de entrada al sistema a través de los elementos en la cuenca hidrográfica que están involucrados con el fenómeno de transformación lluvia escorrentía como por ejemplo la morfología, la geología, los suelos, las redes de drenaje, las coberturas y por supuesto, la presión antrópica. Sin embargo, a la resolución temporal de este estudio, los conjuntos estadísticos, en espacial el caso de la precipitación diaria presenta gran cantidad de valores nulos lo cual dificulta en primer lugar, la definición como variable aleatoria pues el conjunto estadístico presenta estructura y, por lo tanto, la caracterización probabilística resulta inviable. Por tal razón, el proceso de la precipitación total diaria podría analizarse mediante la construcción de dos  $\sigma$ -álgebras, una en donde se estudien conjuntos de eventos de precipitación y no precipitación, y otra en donde solo se analiza el segundo conjunto (los eventos de lluvia). Es posible que debido al uso de la precipitación promedio en lugar del valor esperado se pierda información en relación con la explicación de los parámetros de la ecuación de FPK.

En general el esquema numérico planteado permitió identificar los conjuntos de parámetros de mejor desempeño para la reproducción de las curvas de densidad de probabilidad empírica condicionada mediante la ecuación de FPK. Así también, los polinomios algebraicos de segundo orden propuestos para la descripción del componente advectivo y difusivo e incluidos en el esquema numérico lograron representar con niveles aceptables de error la variabilidad hidrológica de las escalas diarias en al menos las unidades de análisis seleccionadas. El esquema de optimización en dos pasos permitió reducir en más de 160 veces los tiempos de procesamiento del esquema numérico en diferencias finitas explícito.

Extractando, los aspectos incluidos en este trabajo que permitieron viabilizar la ejecución del esquema numérico planteado para el complejo problema de modelación desarrollado son:

- Inclusión del proceso de *pre-optimización*, que permitió establecer valores de parámetros iniciales al proceso de optimización de forma tal que la condición de estabilidad nunca fue restrictiva para el esquema, esto es, que la discretización temporal se mantuvo en valores aceptables con relación al costo computacional,
- Solución del esquema numérico mediante operaciones vectoriales (*fancy slicing*) que permitieron eliminar el recorrido celda a celda en el dominio espacial mediante el uso de los tradicionales ciclos,
- Definición de un *dominio espacial móvil* en función de los límites del problema para cada transición de la curva de densidad de probabilidad,
- *Selección de la cantidad de nodos espaciales* que permitieran una adecuada discretización del problema de modelación mediante el análisis de la condición de estabilidad,
- *Aplicación de esquemas en multiproceso (MP & MPI)*, que permiten explotar todo el poder de procesamiento de las nuevas arquitecturas CPU en equipos de alto rendimiento (High Performance Computing, HPC).

Los errores obtenidos para el 87.2 % de las transiciones son menores al 10 %. Sin embargo, para algo menos del 11% de las transiciones los resultados de optimización presentan errores superiores

al 100 %, inclusive para un 1% de este último conjunto los errores son mayores al 1000 %. La Tabla 25 presenta las características de las estaciones hidrológicas que arrojaron un error superior al 10 % en alguna de sus transiciones.

Tabla 25. Resumen de las unidades que obtuvieron en alguna de sus transiciones un error superior al 10 %

UAH	Mínimo	Máximo	Promedio	Número de Transiciones	Longitud Registro (Años)	% Faltantes	Criterio de calidad	Nombre Estación	Corriente
26137120	10.1%	110910%	589%	259	23	0.1%	NO CUMPLE	REINA LA	SAN EUGENIO
15017030	10.2%	54036373151%	275468685%	198	37	8.7%	CUMPLE	MINCA	GAIRA
26187020	10.0%	191549%	1589%	130	21	0.4%	NO CUMPLE	RIO PIEDRAS BP-2B	PIEDRAS
21167050	10.6%	280827%	5628%	107	43	2.1%	CUMPLE	PTE CUNDAY	PRADO
26187010	10.3%	2701%	96%	81	38	13.8%	CUMPLE	MAYORÍA LA BP-1	BUEY
27017130	10.0%	137%	28%	80	28	25.3%	NO CUMPLE	PALMICHAL G-8	GUADALUPE
35017070	10.5%	9224%	203%	73	35	19.6%	CUMPLE	RANCHO ALEGRE	OROTOY
21167090	10.2%	1342%	106%	72	22	1.9%	NO CUMPLE	HDA EL RUBY AUTOM	CUNDAY
44017050	10.0%	3495%	93%	70	32	0.1%	CUMPLE	RIOGRANDE - CAUCA	GRANDE
35037130	10.1%	299%	35%	64	32	9.3%	CUMPLE	PTE DEL AMOR	OCOA
12017080	10.7%	205%	57%	60	34	9.2%	CUMPLE	CAREPA	CAREPA
24057050	11.2%	744%	56%	42	37	18.5%	CUMPLE	PTE COMUNA	CHUCURI
27017070	10.4%	1103%	66%	39	25	4.2%	NO CUMPLE	PTE BELMIRA RG-6	GRANDE
21017050	10.1%	57%	17%	30	43	0.5%	CUMPLE	PITALITO 2 AUTOMÁTICA	GUARAPAS
21227010	10.6%	63%	19%	21	42	9.4%	CUMPLE	PIEDRAS	OPIA
21057040	10.3%	326%	39%	21	22	0.0%	NO CUMPLE	BOCATOMA	ULLUCOS
22067010	10.1%	228%	34%	13	43	26.2%	CUMPLE	PTE ORTEGA	ORTEGA
23087160	11.3%	352%	66%	12	41	8.7%	CUMPLE	CARAMANTA	NUS
21037020	12.0%	371%	75%	11	42	23.8%	CUMPLE	SAN MARCOS	SUAZA
12027050	10.2%	24%	15%	8	36	1.1%	CUMPLE	PUEBLO NUEVO	MULATOS
21017060	16.6%	242%	113%	4	23	1.1%	NO CUMPLE	MAGDALENA LA	MAGDALENA
21017030	16.2%	16%	16%	1	43	26.1%	CUMPLE	CASCADA SIMON BOLI	MAGDALENA
23147040	27.4%	27%	27%	1	21	17.3%	NO CUMPLE	AYACUCHO	LA COLORADA
53077020	98.3%	98%	98%	1	25	3.3%	NO CUMPLE	GUAYABAL	MICAY

Como se puede apreciar 24 de las 30 unidades presentaron en al menos una vez una transición con un error superior al 10 %. Sin embargo, solo 6 estaciones (La Reina, Minca, río Piedras, Puente Cunday, Rancho Alegre y Hacienda Ruby Automática) presentaron más de 50 transiciones con errores promedio por encima del 100 %.

Sin embargo, unidades con un registro superior a 30 años y con una cantidad de datos faltantes al 10 % también presentaron errores superiores al 10 %, inclusive la estación Minca (15017030) con un registro neto de 34 años (descontando los datos faltantes) fue la estación que presentó el error relativo de ajuste más grande ( $> 1E06$  %), y en 198 ocasiones de las 365 transiciones presentó errores por encima del 10 %. Lo anterior indica que el desempeño del algoritmo de optimización no solo es función de la cantidad de datos y que existen otras variables que inciden en el logro de buenos ajustes. Estas variables pueden ser explicadas por temas como difusión numérica la cual ha sido identificada por Rosmann en la ecuación de FPK unidimensional (Rosmann, 2014), aunque en trabajos previos a este en donde se evaluaron los resultados de optimización de la ecuación de FPK (caso unidimensional) bajo diferentes esquemas de aproximación numérica se concluyó que tal efecto puede no ser tan significativo (Reyes Gil, 2016). Debido a que en más del 80% los resultados

obtenidos son aceptables, se definió seguir adelante con el análisis de estos buenos “ajustes”, y dejar el análisis de las “malas” transiciones como tema de nuevas investigaciones.

En adelante, los resultados y conclusiones de este trabajo de investigación se realizaron en base a el conjunto de datos que obtuvieron un error menor al 10 % en el proceso de optimización, los cuales conforman el 87 % del conjunto de información procesada.

La Tabla 26 y Tabla 27 resumen las magnitudes de los parámetros de la ecuación de FPK identificados mediante el proceso de optimización para el componente advectivo y difusivo. Para ambos componentes se puede

Tabla 26. Valores promedio, máximo y mínimo de los parámetros de la ecuación de FPK óptimos del componente advectivo

UAH	k1	k2	k3	k4	k5	k6
11027050	1.61E-02	-3.23E-05	6.39E-01	2.21E-02	-3.51E-05	-1.92E+00
12017080	-7.14E-03	1.25E-04	1.14E-01	-6.62E-03	9.99E-05	1.16E-01
12027050	3.91E-05	1.02E-05	-3.61E-02	-2.58E-05	1.79E-05	-6.07E-02
15017030	-7.04E-03	1.37E-04	1.04E-01	-8.40E-03	1.96E-04	1.02E-01
21017030	2.31E-02	-7.73E-05	-3.97E+00	1.69E-03	7.85E-06	5.61E-01
21017050	-2.77E-03	9.97E-05	1.57E-02	-5.06E-03	1.80E-04	2.74E-02
21017060	1.97E-02	-1.28E-04	-2.07E-01	-4.54E-03	7.82E-05	1.92E-01
21027010	4.46E-03	-2.70E-05	1.10E+00	-5.41E-03	1.35E-05	1.04E+00
21037020	-4.70E-03	2.26E-05	-9.89E-02	-1.88E-02	8.11E-05	7.70E-01
21057040	-2.54E-02	3.52E-04	1.20E+00	5.32E-03	-3.99E-04	1.38E+00
21167050	-3.82E-03	1.15E-04	2.71E-02	-3.05E-03	7.42E-05	3.26E-02
21167090	-7.75E-03	7.36E-05	3.48E-01	-5.68E-03	1.00E-04	1.88E-01
21227010	4.46E-04	5.76E-06	-1.09E-02	-1.44E-04	7.46E-06	3.25E-04
22067010	-7.64E-03	2.01E-04	2.45E-01	-2.09E-03	-1.37E-04	2.90E-01
23087160	4.76E-03	-9.13E-05	-1.41E-01	1.55E-03	-8.69E-05	-6.11E-02
23147040	-7.30E-03	1.19E-06	2.30E+00	3.43E-03	-3.19E-05	-2.44E-01
24057050	-1.97E-02	3.74E-04	2.53E-01	-1.55E-02	5.65E-05	4.25E-01
26137120	-2.39E-02	1.27E-04	4.80E-01	-2.87E-02	4.99E-04	4.02E-01
26187010	-2.59E-02	2.56E-04	6.53E-01	-1.43E-02	2.32E-05	4.25E-01
26187020	-4.71E-02	1.18E-03	6.25E-01	-4.85E-02	7.67E-04	9.37E-01
27017070	-3.58E-02	4.92E-04	5.05E-01	-1.11E-02	8.89E-05	3.15E-01
27017130	4.59E-03	-1.54E-04	2.50E-02	3.69E-03	-1.88E-04	1.31E-01
35017070	7.93E-04	-5.97E-05	8.64E-03	-6.81E-03	8.52E-05	1.58E-01
35037130	1.56E-03	-6.36E-05	1.03E-02	-7.19E-03	8.42E-05	1.81E-01
35057010	3.51E-03	-4.61E-06	-7.87E-01	1.04E-02	-1.48E-05	-2.20E+00
44017050	-5.54E-03	1.88E-04	-5.50E-02	-9.13E-03	2.34E-04	7.03E-03
44037080	7.87E-04	-3.40E-05	1.97E-01	2.98E-04	-3.32E-05	2.71E-01
47017150	1.77E-02	-4.16E-05	-2.30E+00	1.04E-02	-1.78E-05	-1.14E+00
53077020	-2.28E-02	1.42E-05	4.40E+00	3.20E-02	-2.06E-05	-1.11E+01
53077060	-7.17E-04	-6.96E-06	1.84E+00	-1.17E-02	2.24E-06	4.47E+00
<b>Promedio</b>	<b>-5.25E-03</b>	<b>1.02E-04</b>	<b>2.49E-01</b>	<b>-4.06E-03</b>	<b>5.78E-05</b>	<b>-1.44E-01</b>
<b>Max</b>	<b>2.81E+00</b>	<b>1.13E-01</b>	<b>1.64E+03</b>	<b>2.53E+00</b>	<b>5.29E-02</b>	<b>1.16E+03</b>
<b>Min</b>	<b>-2.56E+00</b>	<b>-4.17E-02</b>	<b>-1.10E+03</b>	<b>-2.69E+00</b>	<b>-8.05E-02</b>	<b>-1.49E+03</b>

Los resultados para el componente advectivo indican que las magnitudes de los parámetros oscilan entre  $[-1.49E03, 1.64E03]$ . Sin embargo, como se muestra en la Figura 34, la magnitud del término

independiente es dos órdenes de magnitud mayor a los demás términos del polinomio. Es decir, los parámetros  $k3$  y  $k6$  resumen la información de la advección de cada transición de la curva de densidad de probabilidad empírica condicionada. Los valores de las magnitudes de los parámetros de advección de  $10E-1$ , especialmente  $k3$  y  $k6$  pueden obedecer a la escala de análisis diaria y al orden del proceso markoviano (radio de correlación 2), esto es que se presentan pequeños cambios en los valores esperados entre conjuntos de caudales diarios consecutivos, lo cual puede diferir de forma significativa en las diferentes escalas de análisis como la mensual, en donde los cambios entre un conjunto y otro se registran en un orden mayor.

Tabla 27. Valores promedio, máximo y mínimo de los parámetros de la ecuación de FPK óptimos del componente difusivo

UAH	g1	g2	g3	g4	g5	g6	x1	x2	x3	y1	y2	y3
11027050	-1.72E-06	3.10E-09	1.32E-04	-7.61E-06	9.47E-09	1.38E-03	5.61E-06	-5.64E-09	-1.23E-03	-3.00E-06	3.80E-09	4.55E-04
12017080	-9.49E-06	3.60E-08	2.68E-04	2.22E-05	-1.68E-06	7.55E-05	-9.49E-06	3.60E-08	2.68E-04	2.22E-05	-1.68E-06	7.55E-05
12027050	6.93E-05	-2.98E-06	-2.63E-04	-6.11E-05	3.30E-06	1.81E-04	6.93E-05	-2.98E-06	-2.63E-04	-6.11E-05	3.30E-06	1.81E-04
15017030	-1.88E-04	4.35E-06	1.84E-03	1.53E-06	-1.24E-06	3.11E-04	-1.88E-04	4.35E-06	1.84E-03	1.53E-06	-1.24E-06	3.11E-04
21017030	-3.68E-08	-1.31E-08	1.04E-04	5.21E-06	-2.06E-08	-2.00E-04	2.26E-07	-1.36E-08	7.11E-05	5.48E-06	-2.10E-08	-2.36E-04
21017050	1.30E-04	-5.93E-06	-7.78E-04	-8.57E-05	4.62E-06	2.56E-04	1.30E-04	-5.93E-06	-7.78E-04	-8.57E-05	4.62E-06	2.56E-04
21017060	8.10E-06	-1.55E-07	3.70E-06	-8.61E-06	1.31E-07	1.28E-04	8.70E-06	-1.55E-07	-1.08E-04	-9.13E-06	1.32E-07	2.01E-04
21027010	5.05E-06	-1.64E-09	-1.17E-03	3.14E-06	-3.21E-09	-5.67E-04	-7.19E-07	3.79E-09	3.74E-05	1.54E-06	-1.92E-09	-1.55E-04
21037020	-1.12E-05	2.77E-07	2.44E-04	2.31E-05	-4.16E-07	-1.44E-04	-1.12E-05	2.77E-07	2.44E-04	2.31E-05	-4.16E-07	-1.44E-04
21057040	1.44E-04	-4.04E-06	-1.39E-03	-2.26E-04	7.43E-06	1.49E-03	1.44E-04	-4.04E-06	-1.39E-03	-2.26E-04	7.43E-06	1.49E-03
21167050	1.04E-04	-3.42E-06	-5.89E-04	1.74E-04	-7.28E-06	-6.86E-04	1.04E-04	-3.42E-06	-5.89E-04	1.74E-04	-7.28E-06	-6.86E-04
21167090	-2.26E-05	-1.19E-08	6.55E-04	-2.09E-05	1.63E-06	5.20E-05	-2.26E-05	-1.19E-08	6.55E-04	-2.09E-05	1.63E-06	5.20E-05
21227010	-4.05E-05	3.41E-06	-1.82E-04	-4.14E-05	2.56E-06	3.86E-05	-4.05E-05	3.41E-06	-1.82E-04	-4.14E-05	2.56E-06	3.86E-05
22067010	2.89E-06	1.69E-07	1.45E-04	-6.46E-05	2.31E-06	4.76E-04	2.89E-06	1.69E-07	1.45E-04	-6.46E-05	2.31E-06	4.76E-04
23087160	-5.44E-05	2.27E-06	2.16E-04	-7.78E-06	3.50E-07	8.53E-06	-5.44E-05	2.27E-06	2.16E-04	-7.78E-06	3.50E-07	8.53E-06
23147040	-2.90E-06	2.65E-08	1.13E-04	-3.48E-06	5.44E-08	-9.91E-07	-3.47E-06	2.76E-08	1.75E-04	-3.63E-06	5.46E-08	1.76E-05
24057050	-1.06E-04	3.68E-06	7.00E-04	-1.50E-04	4.21E-06	8.85E-04	-1.06E-04	3.68E-06	7.00E-04	-1.50E-04	4.21E-06	8.85E-04
26137120	-1.01E-03	4.41E-05	5.08E-03	2.76E-04	-1.54E-05	-6.37E-04	-1.01E-03	4.41E-05	5.08E-03	2.76E-04	-1.54E-05	-6.37E-04
26187010	-2.71E-04	8.30E-06	1.94E-03	-2.85E-04	6.63E-06	2.79E-03	-2.71E-04	8.30E-06	1.94E-03	-2.85E-04	6.63E-06	2.79E-03
26187020	-7.72E-05	1.52E-06	7.86E-04	-3.23E-04	1.14E-05	1.80E-03	-7.72E-05	1.52E-06	7.86E-04	-3.23E-04	1.14E-05	1.80E-03
27017070	-1.45E-04	4.88E-06	1.04E-03	-9.85E-05	1.62E-06	9.95E-04	-1.45E-04	4.88E-06	1.04E-03	-9.85E-05	1.62E-06	9.95E-04
27017130	-1.30E-04	4.36E-06	9.30E-04	-4.52E-05	1.64E-06	1.60E-04	-1.30E-04	4.36E-06	9.30E-04	-4.52E-05	1.64E-06	1.60E-04
35017070	4.95E-05	-1.76E-06	-4.86E-04	-1.12E-04	2.87E-06	9.08E-04	4.95E-05	-1.76E-06	-4.86E-04	-1.12E-04	2.87E-06	9.08E-04
35037130	-8.67E-05	3.59E-06	4.12E-04	-9.72E-05	4.24E-06	2.62E-04	-8.67E-05	3.59E-06	4.12E-04	-9.72E-05	4.24E-06	2.62E-04
35057010	1.65E-06	-2.58E-08	9.58E-05	-2.77E-06	2.63E-08	1.68E-05	2.08E-06	-2.63E-08	6.80E-05	-4.99E-07	2.33E-08	-4.49E-04
44017050	-4.98E-04	2.00E-05	3.01E-03	-4.07E-04	1.51E-05	2.37E-03	-4.98E-04	2.00E-05	3.01E-03	-4.07E-04	1.51E-05	2.37E-03
44037080	1.16E-05	-2.35E-07	-1.55E-04	6.24E-06	-1.76E-08	-1.77E-04	1.16E-05	-2.35E-07	-1.53E-04	6.24E-06	-1.76E-08	-1.76E-04
47017150	2.59E-07	1.98E-09	-5.96E-05	3.03E-06	-2.04E-08	-1.38E-04	-4.76E-07	2.74E-09	8.50E-05	3.01E-06	-2.04E-08	-1.33E-04
53077020	-2.61E-06	1.70E-09	-4.22E-05	-3.04E-08	3.39E-10	-5.01E-04	-2.44E-06	7.22E-10	1.41E-03	-2.62E-06	1.54E-09	7.44E-04
53077060	3.38E-06	-4.28E-09	-5.76E-04	2.30E-06	-3.44E-09	-4.45E-04	-1.34E-06	1.83E-09	2.76E-04	-6.51E-07	3.35E-10	2.36E-05
<b>Promedio</b>	<b>-7.09E-05</b>	<b>2.75E-06</b>	<b>4.01E-04</b>	<b>-5.11E-05</b>	<b>1.47E-06</b>	<b>3.70E-04</b>	<b>-7.10E-05</b>	<b>2.75E-06</b>	<b>4.74E-04</b>	<b>-5.11E-05</b>	<b>1.47E-06</b>	<b>3.96E-04</b>
<b>Max</b>	<b>1.27E-03</b>	<b>2.52E-01</b>	<b>2.10E-02</b>	<b>8.39E-04</b>	<b>2.57E-01</b>	<b>2.64E-02</b>	<b>1.27E-03</b>	<b>2.45E-01</b>	<b>2.10E-02</b>	<b>8.39E-04</b>	<b>2.56E-01</b>	<b>9.99E-02</b>
<b>Min</b>	<b>-9.35E-04</b>	<b>-1.57E-01</b>	<b>-2.36E-02</b>	<b>-7.37E-04</b>	<b>-2.46E-01</b>	<b>-2.40E-02</b>	<b>-9.35E-04</b>	<b>-1.59E-01</b>	<b>-2.36E-02</b>	<b>-7.37E-04</b>	<b>-2.20E-01</b>	<b>7.00E-06</b>

El componente difusivo del proceso estocástico mostró un comportamiento similar al caso advectivo, aquí los parámetros  $g3$ ,  $g6$ ,  $x3$ ,  $y3$  que representan los términos independientes de los polinomios difusivos y el término de primer orden, resumen la información del proceso difusivo.

Estos son sensiblemente menores a los valores encontrados en los parámetros advectivos en tres (para el término independiente) y cuatro órdenes de magnitud (para el término de primer orden). En promedio para las 30 UAH, la matriz de difusión puede definirse de la siguiente manera:

$$B = \begin{bmatrix} -7.09 \times 10^{-5} * Q_1 + 4.01 \times 10^{-4} & -7.10 \times 10^{-5} * Q_1 + 4.74 \times 10^{-4} \\ -5.11 \times 10^{-5} * Q_2 + 3.96 \times 10^{-4} & -5.11 \times 10^{-5} * Q_2 + 3.701 \times 10^{-4} \end{bmatrix}$$

Las magnitudes en los caudales diarios en las UAH varían entre  $1 \times 10^{-2} \text{ m}^3/\text{seg}$  y  $1 \times 10^4 \text{ m}^3/\text{seg}$ , con un valor promedio de  $1 \times 10^2 \text{ m}^3/\text{seg}$ , es decir que los órdenes de las magnitudes del término difusivo pueden estar en el orden  $1 \times 10^{-3}$  y  $1 \times 10^{-1}$ .

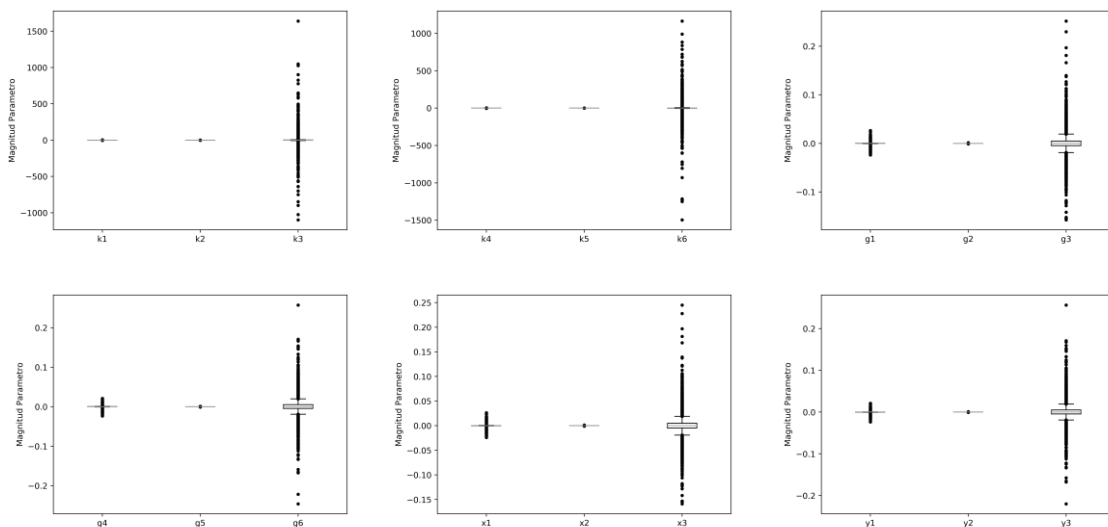


Figura 34. Dispersión de los parámetros óptimos encontrados para las 30 UAH y 365 transiciones.

De los resultados se puede inferir que: a) dado los términos que resumen la información de los componentes advectivos y difusivos recaen en su mayoría en los términos independientes de los polinomios sugeridos, las estructuras matemáticas propuestas pueden entonces ser simplificadas a estructuras más simples, b) la estructura determinista polinómica no logra capturar la complejidad del proceso hidrológico y, por lo tanto, como efecto del artefacto de optimización se resume la información en los términos independientes.

Por otra parte, los parámetros morfométricos que inciden en componente advectivo son índices de forma como el área de drenaje y el ancho medio de la cuenca, resultado lógico pues es claro que, a mayor área de captación, las probabilidades de encontrar aumento en los caudales son mayores. El componente difusivo es explicado parcialmente por índices morfométricos relacionados con aspectos hidráulicos en el sistema como son el coeficiente de sinuosidad, la longitud lineal del cauce más largo y la pendiente ponderada del cauce principal. Los resultados para las relaciones clima y parámetros de la ecuación de FPK, indican que las relaciones se dan principalmente con los parámetros advectivos, específicamente con el parámetro  $k3$  que es el término independiente

del polinomio para la dimensión con rezago de un día, es decir, que para la evolución de las magnitudes de los caudales a esta escala temporal (diaria) y para sistemas hidrológicos con memoria de orden dos, el clima representado por el valor esperado de la precipitación es un factor clave en la determinación del componente advectivo.

Otro aspecto importante resultó al agregar los coeficientes que resultaron de las regresiones lineales entre el parámetro  $k3$  (parámetro que resume en mayor medida la información del proceso advectivo) y el valor esperado de la precipitación total diaria, y compararlos con el régimen de precipitaciones promedio de las UAH (en general, las UAH presentan un régimen similar con dos máximos al año). La Figura 35, presenta la situación expuesta, en ella se observa claramente que la relación directa entre el parámetro  $k3$  (advección) y el régimen de precipitaciones para las épocas de alta humedad (abril a mayo y septiembre a noviembre) y baja humedad (diciembre a marzo). Para la época de transición del régimen de alta a baja humedad no se encontraron relaciones estadísticamente significativas. *Se recuerda que el termino advectivo en la ecuación de FPK es negativo, por lo tanto, la relación entre  $k3$  y la precipitación es directa.*

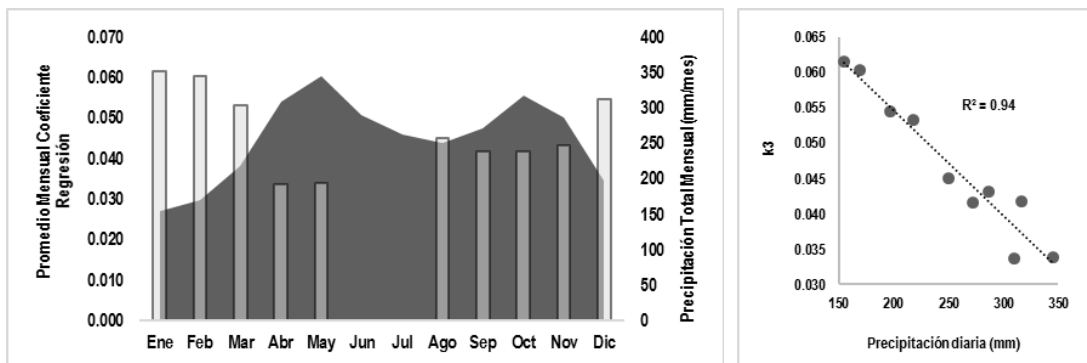


Figura 35. A la izquierda, valores del coeficiente de la regresión del parámetro  $k3$  versus el clima [precipitación] agregado a escala mensual, y el régimen de precipitaciones promedio de las unidades de análisis hidrográfico. A la derecha, gráfico de dispersión de las mismas variables. La significancia confiabilidad de la regresión es del 99%.

Finalmente, las relaciones de los parámetros versus el clima y la morfometría analizadas mediante regresiones lineales múltiples, confirmaron los análisis parciales: parámetros FPK versus morfometría y parámetros FPK versus clima, por lo tanto, como resultado de esta investigación es posible reproducir la variabilidad hidrológica a resolución temporal diaria mediante enfoque estocástico en los periodos de alta y baja humedad a través de aproximaciones a los núcleos deterministas los cuales involucran el clima e índices morfométricos para explicar la evolución de las curvas de densidad de probabilidad condicionada.

Por ejemplo, para el caso a la transición de la curva condicionada  $p(Q_{334}|Q_{333})$  para cualquier UAH, y de acuerdo con los resultados obtenidos:

$$k3 = 24.13 * P3_{UAH} - 14.95 * P12_{UAH} + 10.35 * P15_{UAH} - 18.51 * PT_{UAH}$$

$$k6 = -6.82 * P3_{UAH} + 7.21 * P4_{UAH}$$

Vale la pena recordar que en el caso de los términos de traslación los análisis de regresión indican que la mayor información se encuentra en los términos constantes del polinomio ( $k_3$  y  $k_6$ ). Para la construcción de los 365 conjuntos de parámetros, el Anexo **A9** contiene los resultados obtenidos mediante el análisis de regresiones lineales para los diferentes grados de significancia ( $\alpha = 1\%$ ,  $5\%$  y  $10\%$ ).

## 6 CONCLUSIONES

Según los resultados obtenidos, la aproximación numérica propuesta a la ecuación de Fokker-Planck-Kolmogorov reprodujo con un error aceptable la variabilidad hidrológica (representada aquí como la evolución de la curva de densidad de probabilidad condicionada) a escalas de alta resolución temporal como la resolución diaria.

El esquema numérico propuesto aproxima la solución de la ecuación bidimensional de FPK por diferencias finitas explícitas para el análisis de sistemas complejos, en este caso sistemas hidroclimáticos. El costo computacional para el caso bidimensional fue reducido en 160 veces (en relación con un esquema donde los valores iniciales de los parámetros en el proceso de optimización son asumidos por el mismo algoritmo) mediante un esquema que combina aproximaciones de los parámetros iniciales en función de los cambios en los momentos estadísticos de cada predictor y operaciones vectoriales en la solución temporal del esquema de diferencias finitas explícitas.

De acuerdo a las estructuras matemáticas identificadas para los términos de advección y difusión, las relaciones detectadas entre los parámetros que la definen y, aspectos endógenos (índices morfométricos) y exógenos (precipitación) en las unidades de análisis hidrográfico, es posible concluir que aspectos de forma como el área de drenaje y el ancho medio de la cuenca hidrográfica explican el componente advectivo en la evolución temporal de la curva de densidad de probabilidad condicionada, mientras que los índices relacionados con aspectos hidráulicos como el coeficiente de sinuosidad, la longitud lineal del cauce principal y su pendiente ponderada explican el componente difusivo de la ecuación de FPK. También, variables exógenas como la precipitación total diaria explican el componente advectivo en la evolución de la curva en los periodos de alta y baja humedad. Por lo tanto, es posible definir la evolución de la curva de densidad de probabilidad condicionada de un sistema hidrológico y, por lo tanto, describir su variabilidad a escalas de alta resolución temporal como la diaria mediante la identificación de los núcleos deterministas de la ecuación de FPK que involucren características endógenas y exógenas tales como índices morfométricos y la precipitación total antecedente.

Los resultados de esta investigación implementan un cambio en el paradigma mecanicista basado en las ideas de Descartes predominante unas décadas atrás y, por el contrario, propone el enfoque sistémico de Bertalanffy analizando el sistema como un todo a través de la evolución de su curva de densidad de probabilidad condicionada ya que la misma restricción de la curva de densidad (condicionada) permite inferir sobre las *relaciones entre las partes del sistema complejo* en análisis y que son estas mismas relaciones, en parte, las que agregan complejidad al mismo sistema y definen su evolución. Por lo tanto, mediante el enfoque propuesto aquí es posible abordar una gran cantidad de problemas del ámbito geofísico e inclusive de otras ramas de la ciencia que involucran complejidad como las ciencias sociales, entre ellas la economía.

En el contexto de la hidroclimatología y, teniendo en cuenta las mismas restricciones de esta investigación en cuanto a la selección de las unidades de análisis y que tienen que ver con la disponibilidad de información hidrológica, estos resultados contribuyen con un método para el estudio de cuencas hidrográficas no instrumentadas en condiciones no estacionarias. Esto es posible, una vez se concluya el proceso de recolección de información de las entradas y salidas del sistema, mediante la identificación de las estructuras que gobiernan los procesos advectivos y



difusivos y, de los valores de los parámetros asociados a estas estructuras en la ecuación de FPK que sean regionalizables a través de los índices morfométricos, de modo que en cuencas donde solo exista información de precipitación se pueda mediante la ecuación de FPK bidimensional producir información probabilística de los caudales solucionando el problema de la evaluación del riesgo en cuencas no instrumentadas.

También es importante, al lograr generar estructuras matemáticas basadas en condiciones físicas intrínsecas del sistema (factores endógenos) y aspectos que lo perturban (factores exógenos) reproducir la variabilidad del sistema en análisis (en este caso la variabilidad hidrológica), entonces, es posible estudiar el modelo y, por lo tanto, el sistema bajo diferentes escenarios como por ejemplo cambios en los factores exógenos y endógenos, los cuales en esta investigación pueden obedecer a cambios en el clima (aumento o disminución del valor esperado de la precipitación) y/o cambios en las características hidráulicas del canal principal (aumento de la sección hidráulica, reducción de la sinuosidad por rectificación y enderezamiento del cauce, entre otros).

Se simuló en forma práctica la variabilidad hidrológica a escala diaria en sistemas bidimensionales. Esta dimensionalidad está definida en este caso por la estructura de autocorrelación. También, pueden existir casos en los cuales la bidimensionalidad se determina por la presencia de dos (2) variables de estado del sistema, como, por ejemplo, para el caso de balance hídrico donde los principales componentes son los excedentes y la evapotranspiración en la cuenca hidrográfica, no obstante, el esquema numérico es el mismo con la diferencia que el lugar de  $Q_1$  y  $Q_2$  lo toman el caudal antecedente y la evapotranspiración respectivamente. Aunque el esquema es bidimensional, este es fácilmente generalizable a  $n$  dimensiones lo que facilitaría el estudio de sistemas de complejidad aún mayor a los aquí analizados. Por ejemplo, el sistema del ciclo hidrológico incluyendo escorrentía, evaporación e infiltración, o procesos con autocorrelación mayor a dos días. Así el esquema numérico propuesto puede ser generalizado para cualquier sistema  $n$  – dimensional.

Algunos aspectos importantes en el desarrollo de esta investigación que no fueron profundizados debido al alcance de este trabajo se listan a continuación ya que pueden ser fuente para futuros trabajos:

- La variable de precipitación en la escala de análisis diaria presenta gran cantidad de valores nulos (fenómeno de intermitencia), lo cual dificulta el proceso de caracterización probabilística y sugiere un tratamiento diferencial para los eventos de lluvia efectiva y las rachas de eventos de lluvia/no lluvia,
- Si bien los resultados de las regresiones entre los parámetros identificados mediante optimización matemática y, los índices morfométricos y el clima muestran con algún grado relaciones significativas entre estos, es posible que se presenten efectos de mínimos locales en el proceso de minimización de la función objetivo introduciendo ruido en los parámetros y, por lo tanto, en la explicación de estos a través de aspectos como la morfometría y el clima,
- La propagación de los errores desde distintas fuentes (modelos digitales de terreno, medición de variables hidroclimáticas, efectos numéricos) puede ser objeto de estudio más aún cuando la dimensionalidad de los sistemas en estudio puede ser mayor a los aquí analizados, teniendo en cuenta que las variables dimensionales del sistema complejo están correlacionadas,

- De acuerdo con los resultados obtenidos en términos de la reducción de los tiempos de computo en el proceso de la solución del problema inverso y, las nuevas tecnologías disponibles (CPU con frecuencias base de 4.2 GHz, GPU y arquitecturas MP & MPI), es importante la exploración de esquemas de aproximación numérica como por ejemplo esquemas implícitos bajo el enfoque de optimización implementado en este estudio.

## 7 BIBLIOGRAFÍA

- Andrés-Doménech, I., García-Bartual, R., Montanari, A., & Marco, J. B. (2015). Climate and hydrological variability: The catchment filtering role. *Hydrology and Earth System Sciences*, 19(1), 379–387. <https://doi.org/10.5194/hess-19-379-2015>
- Angarita, H. (2008). *Extracción de operadores de transición de estado en sistemas complejos. Aplicaciones en hidrología*. Pontificia Universidad Javeriana.
- Baker, V. R., Kochel, R. C., & Patton, P. C. (Peter C. (1988). *Flood geomorphology*. Wiley.
- Bendat, J. S., & Piersol, A. G. (2010). *Random Data*. Hoboken, NJ, USA: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118032428>
- Bhamare, S. M., Agone, V. M., Patil, S. P., & Science, S. S. V. P. S. L. K. P. R. G. (2012). Delineation of geomorphic aquatones in watershed. In *World Congress on Water, Climate and Energy* (pp. 1–9). Dublin, Ireland.
- Biswas, A., Das Majumdar, D., & Banerjee, S. (2014). Morphometry Governs the Dynamics of a Drainage Basin: Analysis and Implications. *Geography Journal*, 2014, 1–14. <https://doi.org/10.1155/2014/927176>
- Bofinger, E., & Bofinger, V. J. (1961). The Gap Test for Random Sequences. *The Annals of Mathematical Statistics*, 32(2), 524–534. Retrieved from <http://www.jstor.org/stable/2237761>
- Bradley, J. V. (1968). *Distribution-free statistical tests*. Prentice-Hall. Retrieved from <https://books.google.com.co/books?id=QKFqAAAAAAAJ>
- Braga, M. I. J. (2000). *Integración de la función y los servicios de los ecosistemas de agua dulce a los proyectos de desarrollo hídrico*. Washintong D.C. Retrieved from <http://services.iadb.org/wmsfiles/products/Publication/s/1441454.pdf>
- Burbano-Girón, J., Domínguez, E., & Barón-Ruiz, O. (2016). Análisis de la relación entre variables morfométricas y biofísicas en la estimación de características probabilísticas para la oferta hídrica superficial en Colombia. *Revista de La Academia Colombiana de Ciencias Exactas, Físicas Y Naturales*, 40(156), 514. <https://doi.org/10.18257/raccefyn.291>
- Campos, D. (2006). Modelos en la ciencia: Sistemas simples, complicados y complejos. In *Día de la Complejidad*. Bogota, D.C.: Universidad Nacional de Colombia.
- Charmaine, K., ECOVERSA, Friedman, J. H., Hurst, H., Ideam, C., McKinney, W., ... Д. Казакевич. (2001). Multivariate Adaptive Regression Splines. *Proceedings of the 9th Python in Science Conference*, 800–22(1), 5543–5551. <https://doi.org/10.1002/wics.1198>
- Chater, N., Tenenbaum, J. B., & Yuille, A. (2006). Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Sciences*, 10(7), 287–291. <https://doi.org/10.1016/j.tics.2006.05.007>
- Chow, V. T. (1994). *Handbook of Applied Hydrology*. New York: McGraw-Hill.
- Cover, T., & Thomas, J. (2006). *Elements of information theory* (Second Edi). Wiley.
- Dendrinov, D. S., & Sonis, M. (1990). *Chaos and Socio-Spatial Dynamics* (Vol. 86). New York, NY: Springer New York. <https://doi.org/10.1007/978-1-4612-0991-1>
- Denisov, S. I., Horsthemke, W., & Hänggi, P. (2009). Generalized Fokker-Planck equation: Derivation and exact solutions. *The European Physical Journal B*, 68(4), 567–575. <https://doi.org/10.1140/epjb/e2009-00126-3>
- DHI. (1998). *Mike - SHE, Advection-Dispersion, Particle Tracking, Sorption-Degradation User Guide and Technical Reference*. Copenhagen.
- Domínguez, E. (2000). Protocolo para la modelación matemática de procesos hidrológicos. *Meteorología Colombiana*, 2, 33–38.
- Domínguez, E. (2001). Modelo matemático para la formulación de escenarios de escorrentía ante posibles variantes del cambio climático. *Meteorología Colombiana*, 3(391), 93–102.
- Domínguez, E. (2010). *Cálculo de parámetros morfométricos*. Bogota, D.C.
- Domínguez, E. (2015a). Análisis de consistencia.
- Domínguez, E. (2015b). Introducción a la hidroclimatología. Bogota, D.C. Retrieved from [www.mathmodelling.org](http://www.mathmodelling.org)
- Domínguez, E. (2016). Procesos estocásticos aplicados a la hidrología. Bogota, D.C. Retrieved from [www.mathmodelling.org](http://www.mathmodelling.org)
- Domínguez, E., Ardila, F., & Bustamante, S. (2010). System Solver: an open source tool for mathematically modelling dynamical systems. *Ingeniería E Investigación*, 30(3), 157–164.
- Domínguez, E., & Kovalenko, V. (2009). Modelación estocástica de sistemas complejos: adaptación y bifurcación como mecanismos de evolución. In *Complejidad: revolución científica y teoría* (p. 180). Bogota, Colombia.
- Domínguez, E., & Rivera, H. (2010). A Fokker-Planck-Kolmogorov equation approach for the monthly affluence forecast of Betania hydropower reservoir. *Journal of Hydroinformatics*, 12(4), 486. <https://doi.org/10.2166/hydro.2010.083>

- Domínguez, E., Rosmann, T., & Chavarro, J. (2014). A Fokker-Planck-Kolmogorov approach for inverse modeling of complex processes applied to a hydrological system. *Geophysical Research Abstracts*, 16(EGU2014), 14689.
- Druzhinin, V. C., & Sikan, A. B. (2001). *Statistical Methods for the treatment of hydrometeorological information*. St. Petersburg: RGGMU.
- Dunne, T., & Leopold, L. B. (1978). *Water in Environmental Planning*. W. H. Freeman. Retrieved from <https://books.google.com.co/books?id=d7WEkcTNk6EC>
- FAO. (1985). Revista Internacional de Silvicultura e Industrias Forestales Unasyliwa. In *IX Congreso forestal mundial México*. México.
- Frigg, R., & Hartmann, S. (2017). Models in Science. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Spring 201). Metaphysics Research Lab, Stanford University.
- Gallant, J. C., & Dowling, T. I. (2003). A multiresolution index of valley bottom flatness for mapping depositional areas. *Water Resources Research*, 39(12), 14. <https://doi.org/10.1029/2002WR001426>
- Gardiner, C. (2009). *Stochastic Methods* (4th ed.). Springer-Verlag Berlin Heidelberg.
- GDAL Development Team. (2015). GDAL 2015 Geospatial Data Abstraction Library Versión 1.1. Open Source Geospatial Foundation. Retrieved from <http://gdal.osgeo.org>
- Giere, R. N. (2004). How Models Are Used to Represent Reality. *Philosophy of Science*, 71(5), 742–752. <https://doi.org/10.1086/425063>
- Gómez, N., & Villamil, N. (2009). *Ingeniería de los sistemas complejos. Complejidad: revolución científica y teoría*. Bogota: Universidad del Rosario.
- Good, I. J. (1953). The serial test for sampling numbers and other tests for randomness. *Mathematical Proceedings of the Cambridge Philosophical Society*, 49(2), 276. <https://doi.org/10.1017/S030500410002836X>
- Gottschalk, L., Gottschalk, & Lars. (2005). Methods of Analyzing Variability. In *Encyclopedia of Hydrological Sciences*. Chichester, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/0470848944.hsa006>
- Grigolini, P., Allegrini, P., & West, B. J. (2007). In search of a theory of complexity: An overview on the Denton workshop. *Chaos, Solitons and Fractals*, 34(1), 3–10. <https://doi.org/10.1016/j.chaos.2007.01.044>
- Grubbs, F. E. (1950). Sample Criteria for Testing Outlying Observations. *The Annals of Mathematical Statistics*, 21(1), 27–58. <https://doi.org/10.1214/aoms/1177729885>
- Gruber, S., & Peckham, S. (2009). Chapter 7 Land-Surface Parameters and Objects in Hydrology. In *Developments in Soil Science* (Vol. 33, pp. 171–194). [https://doi.org/10.1016/S0166-2481\(08\)00007-X](https://doi.org/10.1016/S0166-2481(08)00007-X)
- Haan, C. T. (1977). *Statistical Methods in Hydrology* (1st ed.). Iowa: The Iowa State University Press.
- HIMAT. (1985). Inventario de cuencas hidrográficas en Colombia. In *III Congreso de Cuencas Hidrográficas*. Cali, Colombia.
- Horton, R. ~E. (1945). Erosional development of streams and their drainage basin: hydrophysical approach to quantitative morphology. *Bulletin of Geological Society of America*, 56(1), 275–370.
- IDEAM. (2014). *Estudio Nacional del Agua. Estudio Nacional del Agua 2014*.
- IGAC. (2002). *Clasificación del Uso Actual de las Tierras*. Bogota, D.C.
- Instituto de Hidrología y Meteorología y Estudios Ambientales - IDEAM. (2014). Sistema de Información del Recurso Hídrico - SIRH. Retrieved from <http://sirh.ideam.gov.co:8230/Sirh/pages/inicio.html>
- IPCC. (2013). Annex III: Glossary. *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*, 1447–1466. <https://doi.org/10.1017/CBO9781107415324.031>
- Jardí, M. (1985). Forma de una cuenca de drenaje. Análisis de las variables morfométricas que nos la definen. *Revista de Geografía*, 19, 41–68.
- Jasiewicz, J., & Metz, M. (2011). A new GRASS GIS toolkit for Hortonian analysis of drainage networks. *Computers & Geosciences*, 37(8), 1162–1173. <https://doi.org/10.1016/j.cageo.2011.03.003>
- Jazwinski, A. (1966). Filtering for nonlinear dynamical systems. *IEEE Transactions on Automatic Control*, 11(4), 765–766. <https://doi.org/10.1109/TAC.1966.1098431>
- Jazwinski, A. (1970). *Stochastic Processes and Filtering Theory* (First Edit). Academic Press.
- Jiménez E., H., & Materón M., H. (1986). *Hidrología Básica Tomo III*. Cali, Colombia: Universidad del Valle.
- Jolliffe, I., Jolliffe, & Ian. (2014). Principal Component Analysis. In *Wiley StatsRef: Statistics Reference Online*. Chichester, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118445112.stat06472>
- Julian Cheng, & Beaulieu, N. C. (2001). Maximum-likelihood based estimation of the Nakagami m parameter. *IEEE Communications Letters*, 5(3), 101–103. <https://doi.org/10.1109/4234.913153>
- Kolmogorov, A. (1931). Über die analytischen Methoden in der Wahrscheinlichkeitsrechnung. *Mathematische Annalen*, 104(1), 415–458. <https://doi.org/10.1007/BF01457949>
- Koutsoyiannis, D. (2008). *Probability and statistics for geophysical processes*. National Technical University of Athens. <https://doi.org/10.13140/RG.2.1.2300.1849/1>

- Kozhevnikova, I., Shveikina, V., & Domínguez, E. (2012). Modelling fluctuations of Caspian Sea levels using a mixed probability distribution. *Journal of Flood Risk Management*, 5(1), 3–13. <https://doi.org/10.1111/j.1753-318X.2011.01116.x>
- Kumar, P., & Narayanan, S. (2006). Solution of Fokker – Planck equation by finite element and finite difference methods for nonlinear systems, 31(August), 445–461.
- Kushner, H. J. (1976). Finite Difference Methods for the Weak Solutions of the Kolmogorov Equations for the Density of Both Diffusion and Conditional Diffusion Processes. *Journal of Mathematical Analysis and Applications*, 53, 251–265.
- Laing, A., & Evans, J. L. (2015). *Introduction to Tropical Meteorology* (2nd ed.). The COMET Program.
- Lays, C., Ley, C., Klein, O., Bernard, P., & Licata, L. (2013). Detecting outliers: Do not use standard deviation around the mean, use absolute deviation around the median. *Journal of Experimental Social Psychology*, 49(4), 764–766. <https://doi.org/10.1016/j.jesp.2013.03.013>
- Lunneborg, C. E., Lunneborg, & E., C. (2014). Runs Test: Examples. In *Wiley StatsRef: Statistics Reference Online*. Chichester, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118445112.stat06553>
- Maldonado, C. E., & Gómez Cruz, N. A. (2010). Modelamiento y simulación de sistemas complejos. *Documentos de Investigación*, (66), 32.
- Masud, A., & Bergman, L. (2005). Solution of the Four Dimensional Fokker-Planck Equation: Still a Challenge. *Icosar*, 2005, 1911w1916. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Solution+of+the+Four+Dimensional+Fokker-Planck+Equation++Still+a+Challenge#0>
- Mathematics, E. of. (2012). Kolmogorov-Smirnov test.
- Mendoza, H., & Bautista, G. (2002). Probabilidad y Estadística.
- Milly, P. C. D., Wetherald, R. T., Dunne, K. A., & Delworth, T. L. (2002). Increasing risk of great floods in a changing climate. *Nature*, 415, 514–517. <https://doi.org/10.1038/415514a>
- Moncho, A. C. (2008). *Modelización Y Simulación De Sistemas Complejos*. Retrieved from <http://www.uv.es/caselles/Mod1.pdf>
- Morin, E. (2004). *Introducción al pensamiento complejo*.
- Nardi, F., Vivoni, E. R., & Grimaldi, S. (2006). Investigating a floodplain scaling relation using a hydrogeomorphic delineation method. *Water Resources Research*, 42(9). <https://doi.org/10.1029/2005WR004155>
- NASA Jet Propulsion Laboratory (JPL). (2015). *Shuttle Radar Topography Mission United States 1 arc second. Version 3. 6oS, 69oW*. Sioux Falls, South Dakota. <https://doi.org/10.5067/MEASURES/SRTM/SRTMUS1.003>
- Neteler, M., Bowman, M. H., Landa, M., & Metz, M. (2012). GRASS GIS: A multi-purpose open source GIS. *Environmental Modelling & Software*, 31, 124–130. <https://doi.org/10.1016/j.envsoft.2011.11.014>
- Nocedal, J., & Wright, S. J. (2006). *Numerical Optimization*. Springer New York. <https://doi.org/10.1007/978-0-387-40065-5>
- Oki, T. (2006). Global Hydrological Cycles and World Water Resources. *Science*, 313(5790), 1068–1072. <https://doi.org/10.1126/science.1128845>
- OMM. (2011). *Guía de prácticas hidrológicas Volumen I*.
- Patton, P. C., & Baker, V. R. (1976). Morphometry and floods in small drainage basins subject to diverse hydrogeomorphic controls. *Water Resources Research*, 12(5), 941–952. <https://doi.org/10.1029/WR012i005p00941>
- Pawula, R. (1967). Generalizations and Extensions of the Fokker-Planck-Kolmogorov Equations. *Information Theory, IEEE Transactions on, IT-13*(2), 33–41. Retrieved from [http://ieeexplore.ieee.org/xpls/abs\\_all.jsp?arnumber=1053955](http://ieeexplore.ieee.org/xpls/abs_all.jsp?arnumber=1053955)
- Pike, R. J., Evans, I. S., & Hengl, T. (2009). Chapter 1 Geomorphometry: A Brief Guide. In *Geomorphometry - Concepts, Software, Applications* (Vol. 33, pp. 3–30). [https://doi.org/10.1016/S0166-2481\(08\)00001-9](https://doi.org/10.1016/S0166-2481(08)00001-9)
- Poff, N. L., Allan, J. D., Bain, M. B., Karr, J. R., Prestegard, K. L., Richter, B. D., ... Stromberg, J. C. (1997). The Natural Flow Regime. *BioScience*, 47(11), 769–784. <https://doi.org/10.2307/1313099>
- Reyes Gil, N. C. (2016). *Solución del problema inverso de la ecuación de Fokker-Planck-Kolmogorov empleando métodos discontinuos de alto orden*. Pontificia Universidad Javeriana. Retrieved from [http://www.mathmodelling.org/home/trabajos-de-tesis/Trabajo\\_de\\_grado\\_NCRG\\_Ed04.pdf?attredirects=0&d=1](http://www.mathmodelling.org/home/trabajos-de-tesis/Trabajo_de_grado_NCRG_Ed04.pdf?attredirects=0&d=1)
- Reyes T., A., Barroso, F., & Carvajal E., Y. (2010). *Guía básica para la caracterización morfológica de cuencas hidrográficas* (Primera Ed). Cali, Colombia: Editorial Universidad del Valle.
- Rivera, H. G., Marentes, H., Martínez Sarmiento, Ó., Domínguez Calle, E., Romero Pinzón, H., Fajardo Sierra, M., ... Carvajal Contreras, M. (2008). *Protocolo Para La Emisión De Los Pronósticos Hidrológicos*. (Imprenta Nacional de Colombia, Ed.). Bogota, Colombia: Imprenta Nacional de Colombia. Retrieved from <http://www.ideam.gov.co/documents/14691/15016/ProtocoloEmision.pdf/7a801741-6e42-435b-aed4-669e30ef0e16>
- Rivera, M. E., Correa H., J. A., Avendaño B., A., Rivera, H. G., & Fuentes B., J. B. (2015). Construcción de un proceso estocástico para simular el movimiento de caudales medios en el río Fonce (San Gil - Santander). *AVANCES Investigación En Ingeniería*, 12, 9.
- Roberts, B. D. W., Dowling, T. I., & Walker, J. (1997). *FLAG : A Fuzzy Landscape Analysis GIS Method for*

*Dryland Salinity Assessment.*

- Rodríguez, L., & Aguirre, J. (2011). Teorías de la complejidad y ciencias sociales, Nuevas Estrategias Epistemológicas y Metodológicas. *Nómadas*, 30, 20. [https://doi.org/10.5209/rev\\_NOMA.2011.v30.n2.3](https://doi.org/10.5209/rev_NOMA.2011.v30.n2.3)
- Rosenmai, P. (2013). Using the median absolute deviation to find outliers. Retrieved from <http://eurekastatistics.com/using-the-median-absolute-deviation-to-find-outliers>
- Rosmann, T. (2014). *Stochastic modeling of hydrometeorological extremes and their possible Relation with global change*. Pontificia Universidad Javeriana. Retrieved from [http://www.mathmodelling.org/home/trabajos-de-tesis/Rosmann%2C Thomas 2014.pdf?attredirects=0&d=1](http://www.mathmodelling.org/home/trabajos-de-tesis/Rosmann%2C%202014.pdf?attredirects=0&d=1)
- Rousseeuw, P. J., & Croux, C. (1993). Alternatives to the Median Absolute Deviation. *Journal of the American Statistical Association*, 1273–1283.
- Ruíz Abellón, M. C. (2017). Procesos Estocásticos. Retrieved from [http://www.dmae.upct.es/~mcruiz/Telem06/Teoria/apuntes\\_procesos.pdf](http://www.dmae.upct.es/~mcruiz/Telem06/Teoria/apuntes_procesos.pdf)
- Rukhin, A., Soto, J., Nechvatal, J., Miles, S., Barker, E., Leigh, S., ... Vo, S. (2010). *A statistical test suite for random and pseudorandom number generators for cryptographic applications*. *National Institute of Standards and Technology* (Vol. 800). Retrieved from <http://oai.dtic.mil/oai/oai?verb=getRecord&metadataPrefix=html&identifier=ADA393366>
- Sjöberg, P. (2005). *Numerical Solution of the Fokker-Planck Approximation of the Chemical Master Equation*. UPPSALA University. Retrieved from <http://scholar.google.com/scholar?hl=en&btnG=Search&q=intitle:Numerical+Solution+of+the+Fokker-Planck+Approximation+of+the+Chemical+Master+Equation#0>
- Strahler, A. N. (1952). Hypsometric (area-altitude) analysis of erosional topography. *Bulletin of Geological Society of America*, 63, 1117–1142.
- Sveshnikov, A. (1968). *Applied methods of the theory of random functions*. Nauka.
- Sveshnikov, A. A. (1966). *Applied Methods of the Theory of Random Functions*. Retrieved from <http://www.sciencedirect.com/science/book/9781483197609>
- Tietjen, G. L., & Moore, R. H. (2012). Some Grubbs-Type Statistics for the Detection of Several Outliers. *Technometrics*.
- UNESCO, & WMO. (2012). *International glossary of hydrology*.
- Uusitalo, L., Lehtikainen, A., Helle, I., & Myrberg, K. (2015). An overview of methods to evaluate uncertainty of deterministic models in decision support. *Environmental Modelling & Software*, 63, 24–31. <https://doi.org/10.1016/j.envsoft.2014.09.017>
- Van Rossum, G. (1991). Python. Netherlands: Python Software Foundation. Retrieved from [www.python.org](http://www.python.org)
- Velasco, A. (2016). *Integración del concepto de variabilidad hidrológica en pronósticos hidrológicos de largo plazo de resolución mensual en Colombia*. Pontificia Universidad Javeriana. Retrieved from [http://www.mathmodelling.org/home/trabajos-de-tesis/TG\\_VelascoSanchezAndres\\_MM.pdf?attredirects=0&d=1](http://www.mathmodelling.org/home/trabajos-de-tesis/TG_VelascoSanchezAndres_MM.pdf?attredirects=0&d=1)
- Vélez, J. I. (2004). Modelo hidrológico y geotécnico para estimar la amenaza de deslizamiento superficial. *XVI Seminario Nacional de Hidráulica E Hidrología*, 11.
- von Storch, H., & Navarra, A. (Eds.). (1999). *Analysis of Climate Variability*. Berlin, Heidelberg, Heidelberg: Springer Berlin Heidelberg. <https://doi.org/10.1007/978-3-662-03744-7>
- Wanielista, M. P. (1997). *Hydrology and Water Quality Control*. (Wiley, Ed.) (Segunda).
- Wojtkiewicz, S., & Bergman, L. (2000). Numerical solution of high dimensional Fokker-Planck equations. *The 8th ASCE Specialty Conference on Probabilistic Mechanics and Structural Reliability*, 1–6. Retrieved from <http://www.nd.edu/~pmc2000/pmc2000/sessions/papers/p167.pdf>
- Woods, R. (2006). Hydrologic Concepts of Variability and Scale. In *Encyclopedia of Hydrological Sciences*. Chichester, UK: John Wiley & Sons, Ltd. <https://doi.org/10.1002/0470848944.hsa002>
- Yoon, J., & Xu, Y. (2010). Alternating Direction Implicit Method-Based Nonlinear Filtering for Relative Orbit Estimation. *Journal of Aerospace Engineering*, 23(3), 186–196. [https://doi.org/10.1061/\(ASCE\)AS.1943-5525.0000024](https://doi.org/10.1061/(ASCE)AS.1943-5525.0000024)