

ESTUDIO COMPARATIVO DE LOS MODELOS NAM Y DBMM APLICADO AL PRONÓSTICO DE CAUDALES

RESUMEN

En el siguiente trabajo de tesis se analiza dos sistemas de modelación hidrológica utilizados para el pronóstico y alerta temprana de eventos extremos. Se realizó primero una caracterización de los sistemas de modelación y luego una comparación de los mismos basándonos en la capacidad predictiva del esquema de validación del modelo encontrado.

El estudio se enmarca en la aplicación de los modelos a la cuenca del Paute, la misma que dispone de registros históricos diarios y cuyas condiciones climáticas son no estacionarias.

Los dos modelos representan a un sistema de modelación conceptual (NAM) y a un sistema de modelación híbrido (DBM). Como resultado se obtuvo una mejor descripción del sistema físico y un mejor rendimiento con el modelo DBM.

PALABRAS CLAVES

Data-Based Mechanistic Model, Nedbor Afstromning Model

1.. Introducción	1
1.1. Motivación	1
1.2. Estado de arte	2
1.3. Alcance y Objetivos	4
1.4. Metodología	4
1.5. Estructura del presente trabajo	5
2.. Marco teórico	6
2.1. Ciclo Hidrológico	6
2.2. Componentes del proceso Hidrológico	6
2.3. Modelos Hidrológicos	9
2.4. Modelo Hidrológico de series de tiempo	10
3.. Métodos	12
3.1. Modelo Mecanicista de Base de Datos	12
3.2. Nedbor-Afstromnings Model	30
3.3. Comparación de los modelos	38
4.. Experimentación Numérica	40
4.1. Ubicación y características geográficas de la cuenca de análisis	40
4.2. Modelación DBM de la cuenca	43
4.3. Modelo Conceptual Nedbor Afstromning Model NAM	60
4.4. Análisis comparativo de los modelos DBM y NAM	72
5.. Conclusiones	

ÍNDICE

1.. Introducción	1
1.1. Motivación	1
1.2. Estado de arte	2
1.3. Alcance y Objetivos	4
1.3.1. Objetivo General	4
1.3.2. Objetivos Específicos	4
1.4. Metodología	4
1.5. Estructura del presente trabajo	5
2.. Marco teórico	6
2.1. Ciclo Hidrológico	6
2.2. Componentes del proceso Hidrológico	6
2.2.1. Precipitación	6
2.2.2. Evaporación	6
2.2.3. Evapotranspiración	7
2.2.4. Infiltración	7
2.2.5. Escorrentía	7
2.3. Modelos Hidrológicos	9
2.3.1. Modelos Métricos	9
2.3.2. Modelos Conceptuales	9
2.3.3. Modelos Físicos	10
2.3.4. Modelos Híbridos	10
2.4. Modelo Hidrológico de series de tiempo	10
3.. Métodos	12
3.1. Modelo Mecanicista de Base de Datos	12
3.1.1. Introducción	12
3.1.2. Modelación Mecanicista de Base de Datos	13
3.1.3. Identificación y estimación	17
3.1.4. Análisis Mecanicista	22
3.1.5. Validación condicional	28
3.1.6. Pronóstico	29
3.2. Nedbor-Afstromnings Model	30
3.2.1. Introducción	30
3.2.2. Estructura del modelo NAM	31
3.2.3. Calibración	35
3.2.4. Validación	38
3.3. Comparación de los modelos	38



4.. Experimentación Numérica	40
4.1. Ubicación y características geográficas de la cuenca de análisis	40
4.2. Modelación DBM de la cuenca	43
4.2.1. Identificación y estimación	44
4.2.2. Interpretación física	50
4.2.3. Validación	54
4.2.4. Pronóstico	57
4.3. Modelo Conceptual Nedbor Asftroming Model NAM	60
4.3.1. Calibración	61
4.3.2. Validación	69
4.4. Análisis comparativo de los modelos DBM y NAM	72
5.. Conclusiones	80

LISTA DE FIGURAS

3.1. Modelo Base de Datos Mecanicista	15
3.2. Estructura del modelo DBM para el sistema precipitación-caudal	23
3.3. Esquema de reservorios lineales en paralelo.	26
3.4. Estructura Simplificada del NAM	31
4.1. Ubicación de la cuenca de estudio	41
4.2. Datos observados de precipitación-caudal y evapotranspiración estimada de la subcuenca del Paute	42
4.3. Datos observados de precipitación-caudal de la subcuenca del Paute	43
4.4. Diagrama de flujo para la modelación DBM del sistema precipitación-caudal	44
4.5. Parámetros Dependientes del Estado para la estimación de la precipitación efectiva.	47
4.6. Simulación utilizando la función de transferencia	49
4.7. Gráfico de dispersión	50
4.8. Descomposición de flujo total en subflujos	52
4.9. Suma de subflujos	53
4.10. Incertidumbre en la estimación de caudales	54
4.11. Descomposición de flujo total en subflujos	55
4.12. Suma de subflujos	56
4.13. Incertidumbre en la estimación de caudales	57
4.14. Grafico de dispersión	58
4.15. Representación gráfica de la variación temporal de los índices utilizados para evaluar el rendimiento del diferentes horizontes del pronóstico	59
4.16. Pronóstico obtenido para un horizonte de cinco días	60
4.17. Descomposición de flujo total en subflujos	64
4.18. Suma de subflujos	65
4.19. Descomposición de flujo total en subflujos	67
4.20. Suma de subflujos	68
4.21. Incertidumbre en la estimación de caudales	69
4.22. Incertidumbre en la estimación de caudales	70
4.23. Descomposición de flujo total en subflujos	71
4.24. Suma de subflujos	72
4.25. Caudal Acumulado	74
4.26. Caudal medio diario	75
4.27. Comparació estimación	76
4.28. Comparación de los modelos DBM y NAM para el periodo de validación	77
4.29. Índices para medir el rendimiento de los modelos	78
4.30. Scatter-plot	79

LISTA DE TABLAS

3.1. Efecto del cambio de los diferentes parámetros	36
3.2. Valores límites del espacio de parámetros	38
4.1. Identificación de la estructura	45
4.2. Comparación entre los modelos lineal y no lineal para identificar el modelo óptimo	45
4.3. Parametrización de los parámetros dependientes del estado por medio de varias funciones	48
4.4. Análisis Físico del modelo	51
4.5. Valores de los diferentes índices de eficiencia utilizados para evaluar el rendimiento del pronóstico	59
4.6. Condiciones iniciales	61
4.7. Parámetros del NAM encontrados mediante calibración automática y automática- manual	63
4.8. Parámetros encontrados del modelo NAM	65
4.9. Comparación del balance total y el caudal medio diario estimados de la cuenca del Paute para los períodos de calibración y validación.	73
4.10. Índices para medir el rendimiento del modelo	78



UNIVERSIDAD DE CUENCA

FACULTAD DE INGENIERÍA

ESTUDIO COMPARATIVO DE LOS
MODELOS NAM Y DBMM APLICADO AL
PRONÓSTICO DE CAUDALES

T E S I S

PREVIO A LA OBTENCIÓN DEL TÍTULO DE:
INGENIERO CIVIL

AUTOR:

EDISSON ANDRÉS QUICHIMBO M.

DIRECTOR DE TESIS:

PHD. ESTEBAN SAMANIEGO A.

2010-2011



Yo, Edison Andrés Quichimbo Miguitama, reconozco y acepto el derecho de la Universidad de Cuenca, en base al Art. 5 literal c) de su Reglamento de Propiedad Intelectual, de publicar este trabajo por cualquier medio conocido o por conocer, al ser este requisito para la obtención de mi título de Ingeniero Civil. El uso que la Universidad de Cuenca hiciere de este trabajo, no implicará afección alguna de mis derechos morales o patrimoniales como autor.

Edison Andrés Quichimbo Miguitama, certifica que todas las ideas, opiniones y contenidos expuestos en la presente investigación son de exclusiva responsabilidad de su autor/a.

INTRODUCCIÓN

1.1. Motivación

El análisis y pronóstico de caudales es uno de los problemas de mayor interés en hidrología. Comprender los procesos hidrológicos y tener la capacidad de predecir estos a tiempo, es esencial para la gestión sostenible de cuencas. Si el modelo es capaz de reflejar la esencia de cómo funciona una cuenca hidrológica (para toda la gama de sus estados posibles), entonces puede ser posible extrapolar con cierta confianza, más allá de las condiciones observadas y llegar a predicciones confiables.

La gestión adecuada de los recursos hídricos disponibles en el Ecuador es una herramienta para el manejo, planificación y conservación de recursos. Las fuentes hídricas permiten entre sus diferentes usos la generación hidroeléctrica; la misma que es la fuente más barata dentro de los medios de generación de electricidad en el Ecuador.

La correcta planificación y optimización en la generación de energía hidroeléctrica demanda del conocimiento anticipado de caudales de entrada a los embalses mediante la realización de un pronóstico hidrológico. La Región Austral es una área donde la alta variabilidad, típica de los Andes, se constituye en una limitante para la aplicación de modelos, por lo que se destaca aún más el desarrollo de técnicas más eficientes para dichos pronósticos, basadas fundamentalmente en el uso de herramientas numéricas de simulación a tiempo real de procesos ambientales y que tienen como insumos los datos provenientes de un sistema de monitoreo a tiempo real.

El uso de análisis de sistemas tiene una larga tradición en el campo de la hidrología. La nueva generación de potentes ordenadores ha asegurado una aplicación mucho más amplia de estas herramientas. Sistemas de apoyo para la toma de decisiones, en muchos casos, combinado con rutinas de optimización, permite realizar lo que hace unos años era imposible de aplicar en un plazo razonable, ahora están al alcance.

Es importante mencionar que la Facultad de Ingeniería y el Grupo de Ciencias de la Tierra y del Ambiente (CGTA), con la rectoría de la Dirección de Investigación de la Universidad de Cuenca (DIUC), está llevando a cabo el proyecto denominado "Sistema de Gestión de Datos y Modelación Hidrológica", para el soporte del pronóstico hidrológico de caudales del Sistema Paute Integral; dirigido por el Dr. Raúl Vázquez Z. En este proyecto



se está afrontando el problema del pronóstico temprano.

El presente trabajo de tesis surge de la necesidad que se presentó en dicho proyecto de investigación respecto al desarrollo de estos modelos propuestos. Se abordará una de las fases del proyecto que consiste en la exploración y preparación de estos modelos para obtener un adecuado nivel de confianza sobre la predicción.

1.2. Estado de arte

El debate actual sobre el desarrollo de modelos hidrológicos que pueden ser utilizados para realizar predicciones futuras se basa en las cualidades relativas de las dos filosofías principales de modelación [36].

En primer lugar, en la metodología *upward* o *bottom-up*; el paradigma dominante en la actualidad, considera que los modelos están basados en teorías hidrológicas de pequeña escala. Estos a menudo resultan en grandes y complejos modelos debido a la necesidad de capturar completamente todas las heterogeneidades y complejidades del proceso observado en pequeñas escalas de espacio-tiempo. Por el contrario, la metodología *downward* o *top-down* se concentra en la extracción de modelos paramétricamente eficientes o *parsimoniosos* que pueden reproducir las características observadas de la variabilidad hidrológica o pueden describir los modos dominantes del comportamiento dinámico del sistema que son *identificables* a partir de los datos disponibles de entrada-salida [37].

Los modelos pueden ser vistos como sistemas que transforman una entrada en una salida. El sistema es una formulación matemática-numérica de procesos hidrológicos que pueden estar basados en teorías o leyes físicas; además pueden analizar datos observados de entrada y de salida para construir modelos empíricos que describan las relaciones observadas, sea por métodos deterministas o estocásticos. Con este enfoque se pueden conceptualizar diferentes procesos que tienen cierto grado de contenido físico pero necesitan los datos observados para ajustar o calibrar los parámetros del modelo. El grado de consideraciones físicas en la descripción del proceso ha permitido clasificar de otra manera los modelos: modelos caja negra (empírico), caja gris (conceptual) y caja blanca (basados en la física) [34].

Sin embargo, para la elección de un modelo hidrológico se debe tomar en cuenta varias características que deben ser elegidas de acuerdo al uso previsto del modelo. En primer lugar, parte del ciclo hidrológico que se desea modelar; en segundo lugar, resultados esperados del proceso hidrológico elegido; y en tercer lugar, que datos disponemos o



cuales son las condiciones iniciales del proceso a modelar. En este sentido hay decisiones fundamentales que deben tomarse en relación tanto a la escala temporal como a la escala espacial de un modelo. En el régimen temporal el modelo puede ser basado en eventos (hidrograma unitario) o puede ser desarrollado para su aplicación en tiempo continuo si el modelo es descrito como estado-estacionario o es variable en el tiempo. En el dominio espacial los modelos pueden describir la cuenca hidrográfica como una sola unidad o considerar combinaciones de sub-modelos agrupados que pueden tener características dimensionales diferentes. Las opciones son múltiples, además cualquier modelo temporal puede combinarse con cualquier distribución espacial.

La modelación de ríos con la habilidad de realizar pronóstico en tiempo real tiene como principal utilidad la determinación de la alerta temprana de eventos de tal manera que permita una gestión para la emisión de alertas de emergencia. Modelos de este tipo son utilizados para modelar el sistema precipitación-escorrentía o en un sentido más amplio las cuencas hidrográficas, los mismos que pueden ser modelos basados en el hidrograma unitario o modelos conceptuales como el modelo Sacramento ([6]), el modelo HBV ([3]), modelo NAM ([28]) y el Data-Based Mechanistic (DBM) [45].

De este grupo, por un lado el NAM, modelo conceptual que calcula de forma continua el contenido de humedad en cuatro reservorios mutuamente interrelacionados; desarrollado originalmente en la Universidad Técnica de Dinamarca [10], ha sido modificado y aplicado extensamente por DHI, además, ha sido distribuido comercialmente en todo el mundo como parte del paquete DHI MIKE 11 NAM para la modelación de ríos.

Por otro lado, el modelo híbrido DBM [47], combina por un lado la habilidad de modelos que basados en los datos observados caracterizan la respuesta en base a esos datos usando para ello estimación u optimización estadística, modelos conocidos como *métricos*; y por otro lado los modelos *conceptuales*. Por lo tanto en el DBMM se pretende identificar y estimar un modelo conceptual a partir de los datos disponibles.

Con un grado de resolución espacial, los modelos distribuidos toman en cuenta la variación de las variables de estado de diversos lugares en la cuenca. Los requerimientos de datos, sin embargo, crecen rápidamente con la mayor resolución espacial. El problema de la distribución espacial puede apoyarse con un manejo adecuado de datos, como puede ser el uso de GIS (Sistemas de Información Geográfica); otros modelos distribuidos conocidos son: TOPMODEL ([4]), el modelo MIKE SHE ([1]) y el modelo SWAT ([2]). Estos modelos tienen en cuenta la fase espacial-terrestre del ciclo hidrológico, aunque con metodologías diferentes.

Finalmente, la reciente y creciente mejora de los recursos computacionales ha hecho



factible el desarrollo y aplicación de complejos modelos numéricos espacialmente distribuidos para evaluar diferentes procesos a varias escalas incluyendo la de la cuenca hidrográfica, sin embargo, las cuencas y subcuencas hidrográficas son sistemas heterogéneos y los sistemas hidro-bio-geo-químicos que allí se suceden son no lineales. El empleo de modelos distribuidos se ha cuestionado recientemente con respecto a los procesos de definición de parámetros y calibración tales como equifinidad y parametrización efectiva [4], lo que argumenta el empleo de modelos conceptuales y constituye una alternativa mucho más eficiente que no adolece de problemas tales como *equifinidad* en vista de que sus parámetros no representan entidades físicas y, como tal, se pueden identificar un solo grupo de parámetros que optimicen la bondad de ajuste de las predicciones del modelo a las observaciones pertinentes [39]. En este sentido, el presente trabajo pretende evaluar el rendimiento de dos modelos con características de conceptuales: el NAM y DBM; modelos que permiten realizar pronósticos de forma tal que permitan alertar de manera temprana eventos extremos.

1.3. Alcance y Objetivos

1.3.1. Objetivo General

Comparar los modelos hidrológicos NAM ("Nedbor-Afstromnings Model") y DBMM ("Data Based Mechanistic Model") para el pronóstico de caudales en tiempo real.

1.3.2. Objetivos Específicos

- Describir el ciclo hidrológico y sus principales componentes para el sistema precipitación-caudal.
- Caracterizar el sistema precipitación-caudal y su herramientas de modelación.
- Caracterizar los modelos NAM y DBMM como algoritmos útiles para el pronóstico de caudales.
- Implementar los modelos DBMM y NAM.
- Comparar y evaluar los algoritmos de los modelos NAM y DBMM de pronóstico de caudales en base a índices estadísticos para la eficiencia de los modelos.

1.4. Metodología

El presente trabajo se ha desarrollado utilizando principalmente la metodología de investigación experimental numérica. Los sistemas de modelación se han descrito y carac-



terizado, así como también sus herramientas metodológicas, por un lado, el DBM, se ha desarrollado numéricamente con herramientas de identificación y calibración estadística disponibles en el paquete de algoritmos CAPTAIN Toolbox para MATLAB; y por otro lado, para el NAM se ha utilizado un software comercial (DHI, MIKE 11) para su implementación.

El Método Experimental Numérico permite realizar una gran cantidad de pruebas con los distintos algoritmos de los modelos, de forma que permita evaluar el error obtenido de la modelación.

1.5. Estructura del presente trabajo

El capítulo 2 presenta una descripción general de las etapas principales del ciclo hidrológico y sus componentes. Se ha puesto más énfasis al proceso de escorrentía y sus características principales para el proceso de modelación.

El capítulo 3 caracteriza los modelos Data-Based Mechanistic (DBMM) y Nedbor Astronomic Model (NAM), y también se presenta la manera en la que se realiza la evaluación de los modelos. Este capítulo se divide en tres partes: la primera presenta en forma detallada las etapas de modelación mecanicista de base de datos, desde la identificación de la estructura y posterior estimación de parámetros del modelo, seguido de la interpretación física del modelo y su respectiva validación. Finalmente se enfoca el problema de predicción en tiempo real. La segunda parte se dedica a explicar el NAM, empezando por la descripción de su estructura y sus principales parámetros, posteriormente su calibración y finalmente su validación. Al final del capítulo 3 se presenta una descripción de los distintos índices que se utilizarán para evaluar el rendimiento de los modelos.

En el capítulo 4, una vez descrito los procesos de modelación, se presenta una aplicación numérica de los modelos descritos, se evalúa su rendimiento en la simulación y en el pronóstico de los resultados. El análisis del rendimiento del modelo se basa en los índices de eficiencia estadística apropiados para el modelo y el conjunto de datos [40].

Finalmente, el capítulo 5 presenta las conclusiones y además las recomendaciones y trabajo a futuro para el desarrollo de modelos.

MARCO TEÓRICO

2.1. Ciclo Hidrológico

El ciclo hidrológico es el principio fundamental de la hidrología; de las muchas representaciones que se pueden hacer de él la más ilustrativa es quizás la descriptiva. Como todo ciclo, el hidrológico no tiene ni principio ni fin y su descripción puede comenzar en cualquier punto. El agua que se encuentra en los océanos o sobre la superficie terrestre se *evapora* bajo el efecto de la radiación solar y el viento; el vapor de agua, que así se forma, se eleva y se transporta por la atmósfera en forma de nubes hasta que se condensa y cae hacia la tierra o los océanos en forma de *precipitación*. Durante su trayecto hacia la superficie de la tierra, el agua precipitada puede volver a evaporarse o ser *interceptada* por las plantas o las construcciones, convertirse en *flujo superficial* sobre el suelo o también *infiltrarse*. El agua infiltrada corre a través del suelo como *flujo subsuperficial* y luego desemboca en los ríos u océanos como *escorrentía superficial*. La mayor parte del agua interceptada y de la escorrentía superficial se vuelve a evaporar a la atmósfera. El agua infiltrada puede *percolar* profundamente para *recargar* el agua subterránea de donde emerge en manantiales o se desliza hacia los ríos para formar la escorrentía superficial y que finalmente desemboca en el mar o se evapora en la atmósfera de esta manera el ciclo continua [8, 23].

2.2. Componentes del proceso Hidrológico

2.2.1. Precipitación

La precipitación incluye la lluvia, la nieve y otros procesos mediante los cuales el agua cae a la superficie terrestre. La formación de precipitación requiere la elevación de una masa de agua a la atmósfera de tal manera que se enfríe y parte de su humedad se condense. Si la temperatura se encuentra por debajo del punto de congelamiento, se forman cristales de hielo [8].

2.2.2. Evaporación

La evaporación se produce básicamente por el aumento de energía cinética que experimentan las moléculas de agua cercanas a la superficie de un suelo húmedo o una



masa de agua producido por la radiación solar, el viento y las diferencias en presión de vapor. El aumento de energía cinética provoca que algunas moléculas de agua se eleven de manera continua a la atmósfera. Al mismo tiempo, algunas de las moléculas que ya se encuentran en la atmósfera se condensan y regresan al cuerpo de agua [23].

2.2.3. Evapotranspiración

Son los procesos de evaporación desde la superficie terrestre y de transpiración de la vegetación. La evapotranspiración es influida por los mismos factores de la evaporación desde una superficie abierta de agua, y también por un tercer factor, el suministro de humedad hacia la superficie de evaporación. La evapotranspiración potencial es la aquella que puede ocurrir desde una superficie bien cubierta por vegetación cuando el suministro de humedad es ilimitado, y se calcula de una forma similar a la que se aplica para la evaporación sobre una superficie abierta de agua. La evapotranspiración real cae por debajo de su nivel potencial a medida que el suelo se seca.

2.2.4. Infiltración

Se define como el movimiento del agua, a través de la superficie del suelo y hacia adentro del mismo, producido por la acción de las fuerzas gravitacionales y capilares [23]. Muchos factores influyen en la tasa de infiltración, incluyendo la condición de la superficie del suelo y su cubierta vegetal, las propiedades del suelo, tales como la porosidad y la conductividad hidráulica y el contenido de humedad presente en el suelo.

La tasa de infiltración, se expresa en unidades de longitud por tiempo, es la tasa a la cual el agua ingresa al suelo por la superficie. Si el agua se estanca en la superficie, la infiltración ocurre a la tasa de infiltración potencial. Si la tasa de suministro de agua en la superficie, por ejemplo por lluvia, es menor que la tasa de infiltración potencial, entonces la tasa de infiltración real también será menor que la tasa potencial.

2.2.5. Escorrentía

La escorrentía se define como el agua proveniente de la precipitación que circula sobre o bajo la superficie terrestre y que llega a una corriente para finalmente ser drenada hasta la salida de la cuenca. El agua proveniente de la precipitación que llega hasta la superficie terrestre una vez que una parte ha sido interceptada y evaporada sigue diversos caminos hasta llegar a la salida de la cuenca. Conviene dividir estos caminos en tres clases: escurrimiento superficial, escurrimiento subsuperficial y escurrimiento subterráneo.



Flujo Superficial

El flujo superficial también conocido como *flujo superficial de exceso de infiltración* o flujo Hortoniano ocurre cuando la tasa de precipitación es mayor que la tasa de infiltración del suelo [13]. Horton consideró que la escorrentía superficial tomaba la forma de un flujo en láminas cuya profundidad podía medirse; a medida que el flujo se acumulaba moviéndose a lo largo de la pendiente, su profundidad aumentaba hasta que descargaba en un canal. Con el flujo superficial existe un almacenamiento de presiones en los hundimientos de la superficie y un almacenamiento por detención superficial proporcional a la profundidad del mismo flujo superficial. El flujo superficial hortoniano es aplicable a superficies impermeables en áreas urbanas y a superficies naturales con capas delgadas de suelo y con baja capacidad de infiltración como ocurre en tierras semiáridas y áridas.

Flujo Subsuperficial

El flujo superficial hortoniano raramente ocurre en superficies con vegetación en regiones húmedas. Bajo estas condiciones, la capacidad de infiltración del suelo excede las intensidades de lluvias observadas para todos los casos con excepción de las lluvias extremas; en tal caso el flujo subsuperficial se convierte en el mecanismo primordial del transporte de agua lluvia hacia las corrientes. Antes de la lluvia, la superficie de la corriente se encuentra en equilibrio con el nivel freático y no ocurre flujo subsuperficial saturado. Cuando llueve se desarrolla un patrón de filtración en la superficie el cual sirve para aumentar el nivel freático; toda la lluvia se infiltra a lo largo de la superficie hasta que el suelo se satura por primera vez; a medida que el tiempo pasa disminuye la infiltración debido a que progresivamente se satura la superficie.

La velocidad del flujo subsuperficial es normalmente tan baja que por sí sola no contribuye en forma significativa a transmitir la precipitación directamente a la corriente, excepto bajo circunstancias especiales en que la conductividad hidráulica del suelo es muy alta.

Flujo Subterráneo

El suelo almacena el agua infiltrada y luego la libera lentamente como flujo subsuperficial para entrar en la corriente como flujo base durante los periodos secos. El flujo total de salida se compone en parte por flujo subterráneo saturado que contribuye directamente a la corriente y en parte por flujo subsuperficial no saturado que se filtra cuesta abajo por encima del nivel freático.

El flujo superficial de saturación se produce cuando el flujo subsuperficial satura el suelo próximo a la parte inferior de la pendiente; entonces se presenta flujo superficial a medida que la lluvia cae en un suelo saturado. El flujo superficial de saturación difiere del flujo superficial hortoniano en que en este último el suelo se satura desde arriba mediante infiltración, mientras que en el flujo superficial de saturación el suelo se satura desde abajo



por el flujo subsuperficial. El flujo superficial de saturación ocurre con más frecuencia en la parte inferior de las pendientes de colinas cerca a las bancas de la corriente. La velocidad del flujo subsuperficial es tan baja que no todo el flujo subsuperficial o el flujo superficial de saturación de una cuenca contribuye a incrementar la corriente durante una tormenta.

2.3. Modelos Hidrológicos

Los modelos hidrológicos son una herramienta indispensable para las pruebas de nuevas hipótesis y para mejorar la comprensión de los procesos hidrológicos y su interacción.

Un modelo es una abstracción de la realidad que intenta representar una realidad compleja en la forma más sencilla y que sea adecuada para el propósito de la modelación. El mejor modelo es siempre el que logra el mayor acuerdo entre los resultados del modelo y las observaciones del mundo real [24]. En la actualidad usar un modelo parsimonioso se ha convertido en un principio rector de las investigaciones científicas.

Existen muchos esquemas diferentes para la clasificación de los modelos hidrológicos que se han introducido con en el tiempo. A menudo, los modelos se han clasificado de acuerdo a las propiedades de contraste como deterministas versus estocásticas, agrupados frente a distribuidos, variables o estacionarios en función del tiempo. Sin embargo, uno de los problemas de modelado más interesantes de la hidrología es la caracterización de la relación dinámica-no lineal de la precipitación y la escorrentía. Esto ha recibido una atención considerable en los últimos treinta años, con modelos matemáticos y/o computacionales que van desde simples representaciones del tipo caja-negra, a los complejos modelos de cuencas hidrológicas basados en la física. [41] han clasificado estos modelos en los siguientes cuatro tipos:

2.3.1. Modelos Métricos

Los modelos métricos son aquellos que basados en los datos observados buscan caracterizar en gran parte la respuesta del sistema, utilizando para ello estimación u optimización estadística [47]. Los modelos métricos se caracterizan generalmente por ser tipo caja negra.

2.3.2. Modelos Conceptuales

Los modelos conceptuales son aquellos que se construyen a partir de una representación *a priori* del proceso hidrológico [22]. Se caracterizan porque su estructura y sus componentes son especificados por el hidrólogo o modelador. Los parámetros del modelo son posteriormente calibrados a través de los datos observados utilizando cualquier método de optimización [47].



2.3.3. Modelos Físicos

Los modelos físicos (Physics-Based) se caracterizan porque buscan representar los componentes del proceso físicamente, los componentes son matemáticamente definidos mediante elementos finitos o diferencias finitas [47]. Los modelos físicos buscan representar los procesos de la forma más próxima a la realidad, son modelos muy complejos y necesitan de un alto costo computacional.

2.3.4. Modelos Híbridos

Los modelos híbridos son modelos conceptuales identificados y estimados a través de los datos observados, que combinan la capacidad de los modelos métricos para caracterizar los datos observados eficientemente con la ventaja de la interpretación física prescrita por el modelador en los modelos conceptuales [47].

2.4. Modelo Hidrológico de series de tiempo

Existe una gran cantidad de información hidrológica, climatológica, etc. recopilada con fines de análisis para posteriormente llevar a cabo la planeación y toma de decisiones. A este registro metódico de la medición u observación numérica efectuada a intervalos de tiempo se le conoce como Serie de Tiempo (TS) [5].

Modelos no estacionarios de series de tiempo

La mayoría de los sistemas medioambientales son no lineales y no estacionarios. Los modelos no estacionarios y no lineales pueden representarse de varias formas, una de ellas es la siguiente [16]:

$$y_t = T_t + C_t + S_t + f(u_t) + N_t + e_t \quad (2.1)$$

en donde sus componentes se describen de la siguiente forma en el contexto de la modelación hidrológica:

- T_t Tendencia.- es el componente a largo plazo que constituye la base del crecimiento o declinación de una serie histórica. Las fuerzas básicas que producen o afectan la tendencia de una serie son: cambios climáticos, cambios en el uso del recurso hídrico, cambio tecnológico e incremento en la productividad.
- C_t Ciclicidad.- es un conjunto de fluctuaciones en forma de onda o ciclos, de más de un año de duración, producidos por cambios en las condiciones atmosféricas-ambientales; representan la diferencia entre los valores esperados de una variable (tendencia) y los valores reales (la variación residual que fluctúa alrededor de la tendencia).



- S_t Estacionalidad.- las fluctuaciones estacionales se encuentran típicamente en los datos clasificados por trimestre, mes o semana. La variación estacional se refiere a un patrón de cambio, regularmente recurrente a través del tiempo. El movimiento se completa dentro de la duración de un año y se repite a sí mismo año tras año.
- $f(u_t)$ Componente que captura la influencia de la variable exógena.
- N_t Ruido.- es una perturbación estocástica que se puede modelar con procesos autoregresivos (AR) o con modelos autoregresivos con promedio móvil (ARMA).
- e_t Aleatoriedad.- denominado también ruido blanco, presenta un comportamiento irregular compuesto por fluctuaciones causadas por sucesos impredecibles o no periódicos, como el clima poco usual.

MÉTODOS

3.1. Modelo Mecanicista de Base de Datos

3.1.1. Introducción

El análisis de las series de tiempo en la modelación de sistemas medioambientales ha sido ampliamente utilizado debido a su formulación estocástica y la estimación óptima de parámetros que permiten hacer fácilmente predicciones probabilistas. La mayoría de técnicas de modelación de series de tiempo son modelos tipo caja-negra por naturaleza. Estos proveen poco o ningún significado físico del modo de comportamiento que domina el sistema y además son menos robustos que un modelo basado en conceptos físicos. En el contexto de la modelación hidrológica, la utilización de funciones de transferencia para la modelación de las series de tiempo puede ser interpretada físicamente al relacionarla con los modelos hidrológicos conceptuales basados en reservorios lineales, de esta manera los modelos pueden tener una interpretación física; el resultado es un modelo mucho más robusto denominado: Modelo Mecanicista de Base de Datos (DBMM).

El DBMM es un modelo híbrido [47], combina por un lado características de modelos que basados en datos observados caracterizan la respuesta en base a esos datos, usando para ello estimación u optimización estadística, los cuales son conocidos como modelos *métricos*. Por otro lado, los modelos *conceptuales* los cuales basados en las percepciones del modelador intentan representar el sistema por medio de reservorios, cuyos componentes principales son considerados por la experiencia del modelador. En el DBMM se pretende identificar y estimar un modelo conceptual a partir de los datos disponibles.

La interpretación física es la parte esencial en la modelación DBM. No es suficiente con que el modelo explique bien los datos, pues este es considerado válido solo si se puede interpretar con un significado físico.

El DBM propone cuestionar los paradigmas actuales y aprovechar esta habilidad para promover un cambio de paradigmas basándose en los datos observados [47]. La naturaleza inductiva del DBMM evita en lo posible las preconcepciones teóricas en las etapas iniciales del análisis. La estructura del modelo no es pre especificada por el modelador, la misma es identificada estadísticamente de una manera directa a partir de los datos observados, para luego estimar los parámetros que caracterizan esta estructura. Segui-



damente, la interpretación busca un significado físico del modelo identificado con el fin de relacionarlo a una clase general de modelos conceptuales siempre que sea posible [36]. La modelación inductiva es fuertemente estadística, por lo que se asume que el modelo es estocástico. Por lo tanto, la incertidumbre en los parámetros y en las entradas estocásticas debe ser cuantificada de alguna manera [49].

3.1.2. Modelación Mecanicista de Base de Datos

El DBMM está basado en la identificación y estimación estadística de una relación dinámica estocástica entre unas variables de entrada y de salida usando herramientas avanzadas de análisis de series de tiempo, y si es necesario, se introduce no-linealidades por medio de una transformación no lineal de parámetros-dependientes del estado. El DBMM elige de todas las estructuras posibles aquella que es claramente identificable, que tenga una solución óptima y que los parámetros sean bien definidos en términos estadísticos [30]. El DBMM es obtenido de una manera inductiva a partir del análisis de los datos, las preconcepciones a priori son evitadas en lo máximo posible pero con el objetivo principal de obtener una interpretación física del mismo.

Básicamente, la modelación DBM está caracterizada por cuatro etapas: identificación de la estructura a partir de una base de datos disponible, estimación de los parámetros que caracterizan la estructura identificada, interpretación física del modelo estimado y una validación del modelo obtenido en un periodo de datos diferente a los utilizados en la identificación y estimación.

La etapa de identificación de la modelación DBM está basada en una estimación no-paramétrica de una función de transferencia no-lineal con parámetros dependientes del estado (SDP), donde las relaciones de los SDP que definen el comportamiento dinámico no-lineal del modelo son estimadas en una forma gráfica no-paramétrica, con los SDP dibujados contra los estados sobre los cuales son identificados como independientes.

En la etapa de estimación de la modelación DBM, los SDP linealmente no-paramétricos son parametrizados en una manera parsimoniosa (paramétricamente más eficiente); esta parametrización puede ser en términos de cualquier función, dependiendo sobretodo de los resultados de la identificación no-paramétrica. Para lograr este paso se puede utilizar herramientas desarrolladas para este tipo de modelación como el toolbox CAPTAIN para Matlab que se puede encontrar en <http://www.es.lancs.ac.uk/cres/captain/>.

La siguiente etapa es la interpretación física del modelo estimado, del adecuado sentido físico que se le pueda dar al modelo dependerá si el modelo es aceptado o no.

La etapa final, la validación, el modelo no puede declararse aceptable si no es bien validado. El modelo estimado de una parte de los datos disponibles es aplicado, sin reestimación de los parámetros, en un grupo de datos diferente al utilizado en la identificación y estimación. Si la validación es satisfactoria, si la incertidumbre del modelo simulado está en a nivel del modelo determinado en la etapa de estimación; el modelo es considerado condicionalmente valido.

La filosofía del DBM y sus herramientas metodológicas presentadas en [49] es descrita en las siguientes etapas:

1. El primer paso y el más importante es definir los objetivos de la modelación y considerar que tipo de modelo reúne esas características, el DBM requiere de datos adecuados; en esta etapa se debe incluir la escala y los datos disponibles a esta escala. Sin embargo, las asunciones *a priori* deben ser mínimamente mantenidas para evitar la imposición de percepciones sobre la naturaleza y complejidad del modelo.
2. La estructura apropiada del modelo se identifica a través de las series de tiempo mediante inferencia estadística, el modelo inicial es una función de transferencia cuyos parámetros pueden variar con el tiempo.
3. Si el modelo identificado es lineal, y sus parámetros varían lentamente o tiene intervalos lineales, entonces los parámetros del modelo identificado son estimados utilizando cualquier método de estimación de sistemas lineales.
4. Si los parámetros del modelo varían significativamente en el intervalo observado, los parámetros del modelo son estimados mediante métodos que permitan a los parámetros variar con el tiempo o variar con respecto al estado.
5. Si el sistema es no lineal, la relación no-paramétrica dependiente del estado debe ser parametrizada, y el modelo no lineal resultante debe ser estimado por cualquier método de optimización numérica.
6. Una vez identificado y estimado el modelo, ya sea lineal o no lineal, se busca una interpretación física del mismo que describa el comportamiento físico del sistema en estudio. Esta etapa de la modelación es la más importante, la interpretación física del modelo es la característica principal que le diferencia de los otros métodos.
7. Finalmente, el modelo encontrado deberá ser probado en un intervalo de tiempo distinto de tal forma que asegure su validez.

La metodología de modelación del DBMM descrita anteriormente, permite distinguir dos etapas principales y bien diferenciadas, estas etapas se pueden representar de forma gráfica como se indica en la figura 3.1 y permiten observar la combinación de los modelos métrico por un lado y conceptual por el otro.

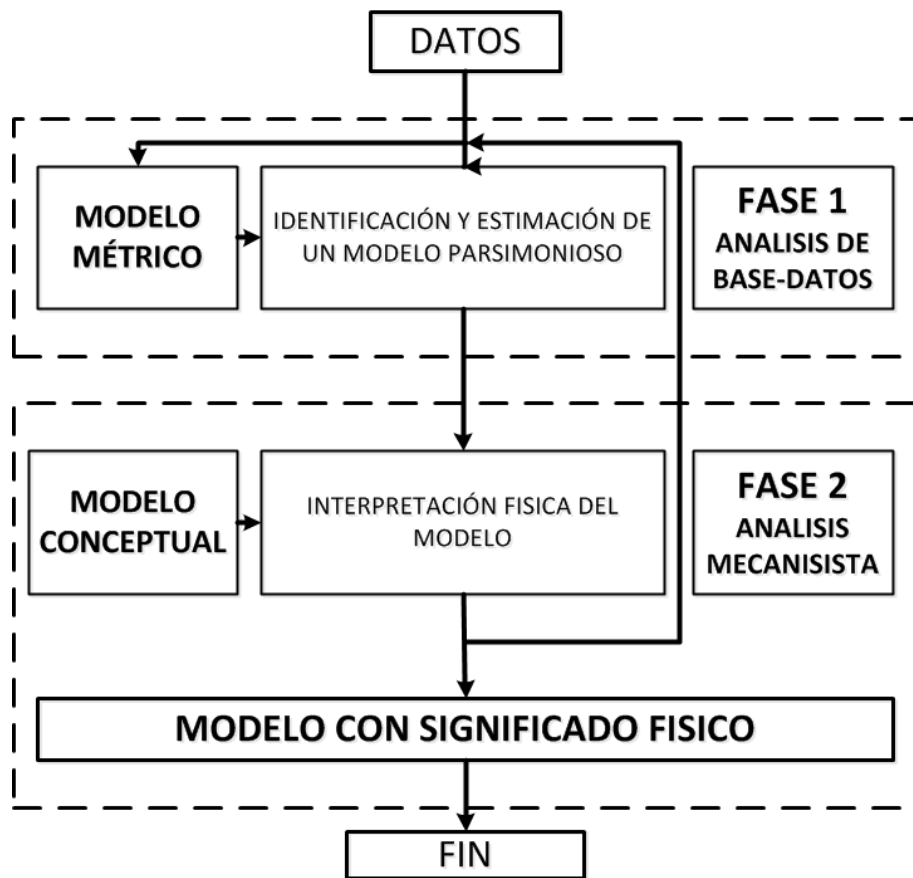


Fig. 3.1: Modelo Base de Datos Mecanicista

Modelos basados en funciones de transferencia

La modelación con funciones de transferencia (TF) se deriva del análisis y control de sistemas, donde ha sido utilizada como la herramienta principal para la modelación y diseño de sistemas de control para sistemas dinámicos lineales [5]. Sin embargo, las TF tienen su aparición en el campo de la hidrología desde que el concepto de impulso de respuesta unitario puede considerarse como una amplitud equivalente del hidrograma unitario instantáneo (IUH). Nash introdujo el concepto de la *cascada de Nash*, que es una cadena de funciones de transferencia de primer orden utilizada para encaminar el flujo [26]. Desde entonces, ha aparecido muchas referencias a los modelos TF en la literatura hidrológica.

En el contexto de la modelación hidrológica, las funciones de transferencia (TF) pueden ser utilizadas de dos maneras: el modelo no-lineal precipitación-caudal y dependiendo de las características de los datos, el modelo lineal o no-lineal de tránsito de caudales. La modelación de TF del sistema precipitación-caudal está basada primeramente en la estimación de parámetros dependientes de estado (SDP), estos SDP describen muchos aspectos hidrológicos y confirman la estructura no-lineal del modelo a identificar.

Los componentes del modelo de TF son dos: un componente lineal, el cual describe la relación entre la precipitación efectiva y el caudal; componente que le da la forma al hidrograma; y un componente no-lineal, el cual describe la relación entre la precipitación medida y la precipitación efectiva, que controla la magnitud del hidrograma a través del tiempo [48].

Para intervalos de tiempo constante y uniformes se puede utilizar la ecuación discreta o de convolución:

$$\begin{aligned} y(t) &= b_0x_t + b_1x_{t-1} + \dots + b_px_{t-m} && \text{Parte MA} \\ &+ a_1y_{t-1} + a_2y_{t-2} + \dots + a_ny_{t-n} && \text{Parte AR} \end{aligned} \quad (3.1)$$

en donde, el caudal y_t más reciente está relacionado a los caudales del tiempo pasado así como también los valores presentes y pasados de la precipitación efectiva en forma lineal. La variable $y(t)$ puede explicarse mediante un promedio móvil (MA) aplicado a la serie de entrada $x(t)$, combinada con una operación autoregresiva (AR) aplicada a la serie de salida [5].

La ecuación 3.2 se puede representar en la forma más común como es TF:

$$y_t = \frac{B(z^{-1})}{A(z^{-1})}u_{t-\delta} + \xi_t \quad (3.2)$$

donde, z^{-1} es un operador de retraso, $A(z^{-1})$ y $B(z^{-1})$ son polinomios cuyos coeficientes son constantes de la siguiente forma:

$$\left. \begin{aligned} A(z^{-1}) &= 1 + a_1z^{-1} + a_2z^{-2} + \dots + a_nz^{-n} \\ B(z^{-1}) &= b_0 + b_1z^{-1} + b_2z^{-2} + \dots + b_mz^{-m} \end{aligned} \right\} \quad (3.3)$$

El término δ indica el retraso, medido en intervalos de tiempo, el mismo que refleja cualquier retraso que puede ocurrir entre la incidencia de la entrada u_t y sus efectos a la salida y_t . El término ξ_t representa la incertidumbre en la relación precipitación efectiva-caudal derivada de la combinación de medidas del ruido, efectos de otras variables de entrada no medidas y del error de modelación. La variable ruido puede ser modelada explícitamente como un ruido coloreado [48], como son los modelos auto-regresivos (AR) o auto-regresivos con promedio móvil como lo describe [5].

La estructura del modelo TF es definida por la triada $[m \ n \ \delta]$, la cual es normalmente identificada estadísticamente a través de los datos durante la etapa de identificación y estimación del modelo. El orden del modelo es normalmente bajo, con $n \leq 2$, $m \leq 3$, y el valor de δ es definido por la naturaleza de la cuenca y la localización de los mecanismos de medida [15, 18, 48]. La parte lineal del modelo TF puede ser descompuesta en forma paralela, en funciones de transferencia de un orden menor. La descomposición le da un significado físico a la TF y además permite mejorar el rendimiento del pronóstico cuando el sistema es implementado en tiempo real como se verá más adelante.

Además, en el modelo TF, la precipitación puede acompañarse con otras variables que pueden ser relevantes como: la temperatura, la evapotranspiración y el caudal, que pueden ayudar a definir los cambios en el contenido de humedad y capacidad de almacenamiento de agua del suelo [47].

Para la estimación de los parámetros de la función de transferencia se pueden utilizar varios métodos, los cuales se pueden reducir a algoritmos recursivos simples utilizando el filtro de Kalman. Para utilizar estos métodos es conveniente realizar la siguiente transformación a la forma vector-matriz:

$$y_t = z_t^T p + n_t \quad (3.4)$$

donde z_t^T es un vector que contiene los valores pasados de entrada y salida,

$$z_t = [-y_{t-1} \dots -y_{t-n} u_{i,t-\delta_i} \dots u_{i,t-m_i-\delta_i}] \quad (3.5)$$

mientras que θ es el vector de parámetros desconocidos, el cual presenta la siguiente forma:

$$\theta = [a_1 \dots a_n b_0 \dots b_n] \quad (3.6)$$

3.1.3. Identificación y estimación

Existen gran número de algoritmos de estimación de parámetros, la mayor parte de ellos se basan en la minimización de funciones de costo de los residuos de las funciones de transferencia; la minimización permite obtener valores óptimos de los parámetros. El

método más utilizado como es el de mínimos cuadrados proporciona buenos parámetros del modelo, sin embargo, los datos proporcionados por las mediciones están generalmente corrompidos por un ruido aditivo coloreado. El método de mínimos cuadrados actúa amplificando los efectos del ruido durante el proceso de estimación pudiendo provocar que la estimación sea asintóticamente sesgada y estadísticamente inconsistente, independientemente de la longitud de la serie de tiempo [29, 18].

Método Simplificado de Refinado de la Variable Instrumental (SRIV)

El método simplificado de la variable instrumental refinada (SRIV), es una extensión del método de la variable instrumental (IV). En este método una nueva variable libre de ruido y fuertemente correlacionada con la variable independiente es usada como un sustituto de la misma para la estimación de los parámetros, mediante un proceso de optimización de máxima verosimilitud (ML) [50, 29].

En el método SRIV se asume que la TF lineal discreta de un sistema SISO (sistema con una sola entrada y una sola salida) toma la forma general 3.2 y su representación en forma de vector-matriz 3.43.5,3.6.

En el SRIV la estimación de \hat{p} es obtenida de la resolución de la siguiente ecuación:

$$\left[\sum_{t=1}^N \hat{x}_t^* z_t^{*T} \right] \hat{\theta} = \sum_{t=1}^N \hat{x}_t^* y_t^* \quad (3.7)$$

donde:

$$z^* = [-y_{t-1}^*, \dots, -y_{t-n}^*, u_t^*, \dots, u_{t-m}^*]^T \quad (3.8)$$

$$\hat{x}^* = [\hat{x}_{t-1}^*, \dots, \hat{x}_{t-n}^*, u_t^*, \dots, u_{t-m}^*]^T \quad (3.9)$$

En esta ecuación la variable instrumental es \hat{x} que es obtenida de la iteración del siguiente modelo auxiliar:

$$\hat{x} = \frac{\hat{B}(z^{-1})}{\hat{A}(z^{-1})} u_{t-\delta} \quad (3.10)$$

en donde $\hat{A}(z^{-1})$ y $\hat{B}(z^{-1})$ son estimaciones adaptadas de la TF; y el asterisco indica que las variables son pre filtradas utilizando $\hat{A}(z^{-1})$ obtenido de la iteración previa:

$$F = \frac{1}{\hat{A}(z^{-1})} \quad (3.11)$$

$$y_t = F_{j-1} y_t, \quad u_t^* = F_{j-1} u_t, \quad \hat{x}_t^* = F_{j-1} \hat{x} \quad (3.12)$$

El pre-filtrado permite remover componentes no deseables de gran frecuencia de la señal de entrada, las cuales reducen la eficiencia en la estimación de resultados, mientras se mantienen esas frecuencias que son esenciales para el análisis del sistema.

En primera instancia, una estimación a priori del vector \hat{x} puede ser obtenida del modelo

auxiliar 3.10 usando mínimos cuadrados; una vez determinado este, se realiza la primera iteración del algoritmo para obtener los parámetros $\hat{\theta}$ iniciales, los mismos que serán actualizados de una manera eficientemente estadística en cada iteración. [50, 29] han demostrado que resultados óptimos se obtienen a la tercera iteración.

Identificación del orden del modelo

Una vez que se hayan estimado los parámetros por cualquier método se procede a identificar la estructura óptima. La identificación del orden del modelo $[n \ m \ \delta]$ de la TF se realiza en base a criterios no solo estadísticos; en la modelación de series de tiempo se utiliza criterios cuyo objetivo debe ser medir cuan bien se ajusta el modelo a los datos e indicar si existe sobre parametrización; los criterios utilizados son los siguientes [16, 45]:

- Coeficiente de determinación:

$$R_T^2 = 1 - \frac{\sigma^2}{\sigma_y^2} \quad (3.13)$$

- Young Identification Criterion:

$$YIC = \ln \left(\frac{\sigma^2}{\sigma_y^2} \right) + \ln \left(\sum_{k=1}^N \frac{\sigma^2 p_{ii}}{\theta_i^2} \right) \quad (3.14)$$

- Akaike Information Criterion:

$$AIC(np) = N \ln \sigma^2 + 2np \quad (3.15)$$

en donde:

$$\sigma_y^2 = \frac{1}{N} + \sum_{k=1}^N [y(k) - \bar{y}]$$

$$\bar{y} = \frac{1}{N} + \sum_{k=1}^N y(k)$$

$$NEVN = \frac{1}{np} + \sum_{k=1}^{np} \frac{\sigma_y^2 \hat{p}_{ii}}{\hat{a}_i^2}$$

y $\hat{a}(N)$ son los parámetros estimados, $np = n + m + 1$ es el número de parámetros de $\hat{a}(N)$, p_{ii} es el i -ésimo elemento de la diagonal en la matriz de covarianza $P(N)$.

El coeficiente de determinación es una medida de cuan bien se ajusta el modelo a los datos, R_T^2 tiende a la unidad cuando la varianza de los residuos generados del modelo son bajos comparados con la varianza de los datos; y tiende a cero cuando ambos tienen

valores con magnitud similar.

El YIC consta de dos partes: la primera parte nos indica con un criterio estadístico cuan bien se ajusta el modelo a los datos y la segunda parte nos indica si el modelo está sobre-parametrizado. Cuando el modelo está sobre-parametrizado el segundo término tiende a dominar la función tomando valores cada vez mayores; por lo tanto el valor que debe tomar la función para determinar un modelo óptimo debería ser negativo con su valor absoluto cada vez mayor.

Por último; el AIC, criterio utilizado para identificar el orden de modelos AR, puede ser utilizado cuidadosamente para identificar el orden del modelo de TF; es parecido al YIC pues busca una relación entre el ajuste del modelo y su complejidad, por lo que ayuda a asegurar la parsimonia de la TF.

Estimación de parámetros de modelos no-lineales y no-estacionario

El comportamiento de una serie de tiempo discreta de un sistema dinámico no lineal puede representarse de la siguiente forma [15]:

$$y(t) = \mathfrak{J}\{\chi(t), \mu(t)\} \quad (3.16)$$

donde $y(t)$ son los datos medidos de salida; $\mathfrak{J}\{\cdot\}$ es la función que describe el comportamiento de las variables del sistema en un estado-espacio no-mínimo (NMSS), definido por un vector NMSS de la forma:

$$\chi(t) = [y(t-1), \dots, y(t-n), u(t)^T, \dots, u(t-m)^T, \dots, U(t)^T, \dots, U(t-q)^T]^T \quad (3.17)$$

donde $\chi(t)$ está compuesto por los valores de la variable de estado $y(t)$, de la variable exógena $u(t)$ y también de otras variables de entrada $U(t)$. Finalmente, la parte estocástica del modelo está dada por $\mu(t)$, el cual representa el ruido, con media cero y además es independiente de las variables de entrada $u(t)$, $U(t)$.

La ecuación 3.16 puede aproximarse a una TF con parámetros dependientes del tiempo (TVP) [18, 47], la cual puede ser escrita de la siguiente forma:

$$y_t = \frac{B(t, z^{-1})}{A(t, z^{-1})} u_{t-\delta} + \xi_t \quad (3.18)$$

donde el ruido ξ_t es independiente de las variables de entrada, $A(t, z^{-1})$ y $B(t, z^{-1})$ son definidos de la siguiente manera:

$$\left. \begin{aligned} A(z^{-1}) &= 1 + a_1(t)z^{-1} + a_2(t)z^{-2} + \dots + a_n(t)z^{-n} \\ B(z^{-1}) &= b_0(t) + b_1(t)z^{-1} + b_2(t)z^{-2} + \dots + b_m(t)z^{-m} \end{aligned} \right\} \quad (3.19)$$

Con el propósito de estimar los TVP, es conveniente expresar 3.18 en forma de vector matriz:

$$y_t = z_t^T \theta + n_t \quad (3.20)$$

donde z_t^T es un vector que contiene los valores pasados de entrada y salida,

$$z_t = [-y_{t-1} \dots -y_{t-n} u_{i,t-\delta_i} \dots u_{i,t-m_i-\delta_i}] \quad (3.21)$$

mientras que θ es el vector de TVP desconocidos y el cual presenta la siguiente forma:

$$\theta_t = [a_1(t) \dots a_n(t) b_{0_i}(t) \dots b_{n_i}(t)] \quad (3.22)$$

en donde cada parámetro del vector θ está definido por un vector estado estocástico de dos dimensiones $x(i, t)$, el mismo que tiene la siguiente forma:

$$x(i, t) = \begin{bmatrix} l_t \\ d_t \end{bmatrix} \quad (3.23)$$

donde, l_t indica el cambio de nivel y d_t indica el cambio de pendiente de los TVP. La evolución de los parámetros es descrita por un proceso Random Walk Generalized (GRW) de la siguiente manera [46]:

$$x(i, t) = Fx(i, t-1) + g\eta(i, t) \quad (3.24)$$

donde:

$$F = \begin{bmatrix} \alpha & \beta \\ 0 & \lambda \end{bmatrix}, \quad g = \begin{bmatrix} \chi \\ \varepsilon \end{bmatrix} \quad (3.25)$$

y $\eta(i, t)$ es un ruido blanco con varianza q_i .

Combinando: 3.33 y 3.24 puede expresarse en un sistema estado-espacio Gauss-Markov (GM) [18, 16]:

$$\begin{aligned} \theta(t) &= F\theta(t-1) + G\eta(k) \\ y(t) &= H\theta(t-1) + n(k) \end{aligned} \quad (3.26)$$

donde $\theta(t)$ es un sustituto del vector estado que caracteriza la evolución de los parámetros y que está formado por la concatenación de los vectores $x(i, t)$ de una forma adecuada; F y G son matrices de transición y de entrada respectivamente que pueden ser dependientes del tiempo.

Una vez formulado el sistema estado-espacio (SS), se puede utilizar el filtro de Kalman ([14] para estimar los parámetros desconocidos del modelo. El mismo que puede ser escrito de la siguiente forma:

Predicción:

$$\begin{aligned}X_{(t|t-1)} &= FX_{(t-1)} + Bu_t \\P_{(t|t-1)} &= FP_{(t-1)}F' + Qr \\Y_{(t|t-1)} &= HX_{(t|t-1)}\end{aligned}\tag{3.27}$$

Corrección:

$$\begin{aligned}G &= P_{(t|t-1)}H'(HP_{(t|t-1)}H' + R_t) \\v &= 1 + HP_{(t|t-1)}H' \\X_t &= X_{(t|t-1)} + G(y_t - Y_{(t|t-1)}) \\P_t &= P_{(t|t-1)} - G * HP_{(t|t-1)} \\Y_t &= HX_t\end{aligned}\tag{3.28}$$

donde, P_t es la matriz de covarianza normalizada de los parámetros, v_t es la varianza normalizada de la variable de salida, G es un factor denominado ganancia de Kalman y Q_r es el radio de varianza del ruido (NVR) [14, 12, 16].

Una vez determinado los parámetros, si estos parecen no variar o varían muy poco con respecto al tiempo, se puede decir que el sistema es predominantemente lineal. Si por el contrario los parámetros exhiben cambios temporales significativos, entonces se debe investigar la naturaleza de estas variaciones e intentar identificar si estos cambios pueden estar relacionados a una de las variables del vector NMSS 3.17. Los métodos para investigar la dependencia de estado se pueden encontrar en [32, 51].

La modelación SDP puede ir acompañada de técnicas de clasificación de datos, como puede ser el ordenamiento no temporal antes de la estimación de los TVP. Esta forma de ordenamiento ayuda a mitigar la variación natural entre la entrada y la salida [50, 47, 29].

3.1.4. Análisis Mecanicista

El principal objetivo de la modelación DBM es obtener un modelo que explique bien los datos además de proporcionar una interpretación física aceptable de la cuenca a una gran escala. La característica más importante de la modelación DBM, que la diferencia de los otros modelos, es su interpretación física.

La interpretación física en el contexto hidrológico es bastante subjetiva y depende de la percepción del modelador [4], sin embargo, en el DBMM estos paradigmas son evitados tanto como sea posible o son mantenidos lo menos posible; la característica inductiva del modelo hace que esto se posible al encontrar un modelo conceptual a partir de la base de datos disponibles de las observaciones o mediciones.

Modelo precipitación-caudal

El modelo precipitación-caudal representado por la TF de tiempo discreto se basan en la estimación SDP para identificar la estructura no-lineal del DBM que puede ser representado en forma gráfica como en la figura 3.2; en donde los dos componentes del modelo TF son: el componente lineal que describe el comportamiento representado por el hidrograma; y el componente no lineal que describe la relación entre la precipitación r_t y la precipitación efectiva u_t controlando así la magnitud de la contribución del hidrograma a través del tiempo.

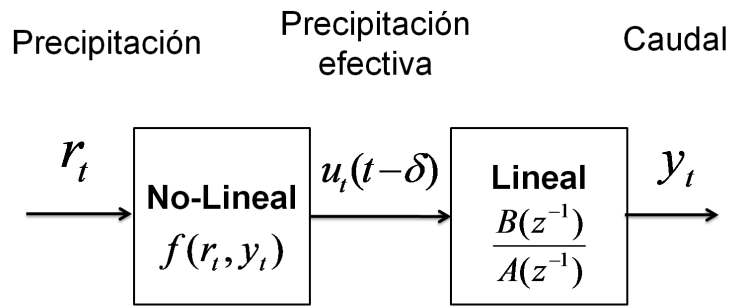


Fig. 3.2: Estructura del modelo DBM para el sistema precipitación-caudal

El modelo DBM resultante tiene la forma de la ecuación:

$$y_t = \frac{B(z^{-1})}{A(z^{-1})} u_{t-\delta} + e_t$$

cuya estructura $[n \ m \ \delta]$ y componentes δ y e fueron definidos en 3.1.2.

La función SDP de entrada puede interpretarse como una precipitación efectiva no lineal, en la cual el SDP es dependiente del caudal. En esta función el caudal actúa simplemente como un sustituto del contenido de humedad del suelo, de lo contrario esta relación no tendría sentido físico debido a que el caudal es función del contenido de humedad del suelo y su patrón de cambio temporal tendría sus mismas características. Por lo tanto, cuando el flujo es bajo indica que el contenido de humedad del suelo también es bajo, esto implica que la precipitación es multiplicada por un factor pequeño y la precipitación efectiva es atenuada; mientras que una medida de caudal alta implicaría también un contenido de humedad de suelo también alto, lo que reflejaría un aumento en la precipitación efectiva y por lo tanto tendría un efecto mucho mayor sobre el caudal [30].

El componente no-lineal del u_t de la figura 3.2 toma la forma:

$$u_t = f(r_t, y_t) \tag{3.29}$$

donde $f(r_t, y_t)$ es una función no lineal desconocida, que relaciona componentes no-observados del estado cuenca-almacenamiento como es la humedad del suelo; en esta función no solo está involucrada la precipitación, sino otras variables como pueden ser la temperatura o la evapotranspiración que pueden causar cambios en la humedad y en la capacidad de almacenamiento del suelo.

La forma de la función no lineal es estimada inicialmente de los datos a través de una estimación SDP de una manera no paramétrica, sin asumir una estructura a priori conservando así la filosofía del DBM. La parametrización puede tomar la forma de cualquier función; [45, 47, 48, 49, 33, 18, 16, 44] han definido a esta como una función potencial de la forma:

$$f(y_t) = cy_t^\beta \quad (3.30)$$

donde el parámetro β es estimado a partir de los datos.

El modelo precipitación-caudal completo toma la siguiente forma:

$$Y_t = \frac{B(z^{-1})}{A(z^{-1})} u_{t-\delta} + e_t, \quad u_k = cf(y_t)r_t \quad (3.31)$$

donde el parámetro c es un factor de escala definido por la siguiente expresión [47]:

$$c = \frac{\sum_{t=1}^N y_t}{\sum_{t=1}^N y_t^\beta r_t} \quad (3.32)$$

La precipitación efectiva es la entrada a la función de transferencia que es la parte lineal, la cual caracteriza la dinámica dominante de la cuenca y define la escala fundamental del hidrograma. La TF puede descomponerse en funciones de transferencia de primer orden; la descomposición puede ser en forma paralela como se verá más adelante en 3.1.4, donde los TF de primer orden son representaciones de reservorios lineales. Cuando los datos disponibles para la modelación son horarios la TF se descompone usualmente en dos componentes que tienen características dinámicas muy diferentes, como un componente de flujo rápido y un componente de flujo lento. En el caso de datos diarios pueden ser tres componentes: subflujo instantáneo, subflujo rápido y subflujo lento; el componente instantáneo puede ser un caudal de respuesta en el mismo día de la precipitación probablemente debido procesos superficiales. El flujo rápido provoca un incremento rápido del hidrograma que probablemente es debido a procesos superficiales o de poca profundidad en la cuenca y con un tiempo de residencia corto. El componente de flujo lento el cual se distingue por su prolongada recesión y que se puede deber a procesos de mayor profundidad o también debido a la reposición del agua de los acuíferos y el consecuente desplazamiento del agua subterránea con un tiempo de residencia largo. El componente de flujo lento nos permite estimar el flujo base total del río, así como también cuantificar y extraer el flujo base.

Descomposición de la función de transferencia

La descomposición matemática de la TF es posible en todos los casos y sin perder su forma original, esta puede ser en forma paralela, en serie o cualquier otra forma. En la modelación hidrológica y particularmente en la modelación de sistemas precipitación-caudal (RR) se puede aceptar una configuración en paralelo [44] la misma que se puede presentar en forma gráfica como en la figura 3.3.

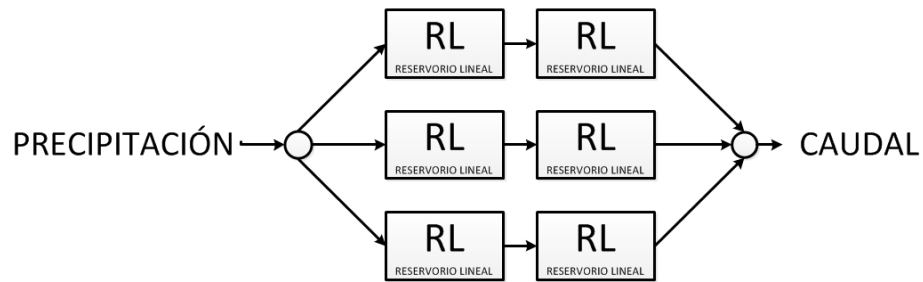


Fig. 3.3: Esquema de reservorios lineales en paralelo.

La TF puede ser descompuesta en forma paralela utilizando fracciones parciales, la misma que puede tomar la siguiente forma:

$$\frac{B(z)}{A(z)} = \frac{\beta_1}{1 - \alpha_1} + \frac{\beta_2}{1 - \alpha_2} + \frac{\beta_3}{1 - \alpha_3} \dots + k(z) \quad (3.33)$$

en donde α_i y β_i con $i = 1, 2, \dots, n$ son los eigenvalores o raíces del polinomio $A(z^{-1})$. Este resultado se interpreta como si la precipitación efectiva se fraccionara en i -ésimos subflujos, los que pueden representar estimaciones de los estados no *observados* de flujo que están asociados a la dinámica del sistema precipitación-caudal.

Estos componentes se pueden interpretar físicamente a partir de las siguientes definiciones paramétricas de las i^{th} TF de primer orden resultantes de la descomposición [44]:

1. Steady State Gain: Define la cantidad de agua que atraviesa por cada reservorio,

$$G_i = \frac{\beta_i}{1 + \alpha_i} \quad (3.34)$$

2. Tiempo de Residencia: Es el tiempo de residencia del agua en la i^{th} ruta. El Steady State Gain y el tiempo de retraso δ definen la dinámica del agua atravesando la i^{th} ruta de flujo. El tiempo de viaje se obtiene a partir de la siguiente expresión:

$$t_i = T_i + \delta.$$

$$T_i = \frac{-\Delta t}{\log_e(-\alpha_i)} \quad (3.35)$$

3. Fracción del volumen del Componente: Es la fracción del volumen atravesando por la i^{th} ruta.

$$p_i = \frac{G_i}{\sum_{j=1}^{j=n} G_j} \quad (3.36)$$

4. Volumen Activo: Volumen asociado con el agua transferida a través de la i^{th} ruta para el t^{th} instante.

$$v_{i,t} = p_i T_i y_t \quad (3.37)$$

5. Fracción de Volumen: Volumen del agua atravesando la i^{th} ruta en el instante t^{th} .

$$d_{i,t} = v_{i,t} V \quad (3.38)$$

Estimación del estado del sistema

Si el modelo del pronóstico es lineal o no lineal, donde la no linealidad toma la forma de una entrada no lineal; entonces la actualización del estado puede ser realizada usando el filtro de Kalman en su forma original, donde los estados son ahora modelos de estado más que parámetros y los parámetros constantes están contenidos en la matriz de transformación en la ecuación de estado.

La función de transferencia para un modelo general puede ser formulada de la siguiente manera:

$$\hat{x}_t = Fx_{t-1} + Gu_t + \eta_t \hat{y}_t = Hx_t + e_t \quad (3.39)$$

donde x_t es el vector estado; u_t es la matriz con los datos de entrada; H es el vector observación; F es una matriz que contiene los parámetros del denominador de la TF; G contiene los parámetros del numerador de la función de transferencia, y η_t y e_t son definidos como ruido blanco.

El mayor problema de utilizar el filtro de Kalman para la actualización en tiempo real del estado es la determinación de los NVRs. El valor de los NVRs controla el comportamiento del filtro, si los NVRs son altos entonces los estados estimados también serán altos y viceversa, además de ser influenciado por el ruido de salida, resultando así en un pronóstico deficiente. Sin embargo, los métodos de optimización numérica basados en RMSE (f) proveen un método objetivo de calibrado de los NVRs sobre datos históricos. Este procedimiento tiene un costo computacional alto, pero la optimización de los NVRs es una operación no incluida en el sistema de pronóstico por lo que no afecta el rendimiento en el uso del filtro de Kalman.

Estimación de hyper-parámetros

Para las aplicaciones de pronóstico utilizando el filtro de Kalman se necesita determinar los hyper-parámetros desconocidos Qr que generalmente conforman una matriz

diagonal. La estimación de hyper-parámetros se la puede realizar utilizando métodos de optimización numérica local como el simplex [27] o mediante métodos de optimización global como el Shuffled Complex Evolution (SCE) [11, 38].

Para aplicaciones que requiere el pronóstico de un intervalo se puede utilizar el método de estimación de máxima verosimilitud (ML) [16] para lo cual se puede utilizar la expresión 3.40; la misma que debe ser minimizada para obtener el criterio de máxima verosimilitud:

$$J = (N - k) \log \left(\frac{1}{N - k} \sum_{t=k+1}^N \frac{d_t^2}{v_t} \right) + \sum_{t=k+1}^N \log v_t \quad (3.40)$$

donde k es el número de metaparámetros.

Para aplicaciones en las cuales el pronóstico requerido es de f -intervalos se puede utilizar la expresión 3.41 propuesta por [33]:

$$J = \sum_{k=1}^{T-f} \frac{(\hat{y}_{k+f|k} - y_{k+f|k})^2}{cov(y_k)} + \left(\frac{\sum i}{T-f} - 0.95 \right)^2 \quad (3.41)$$
$$i = \begin{cases} 0 & y_{k+f} \notin [\hat{y}_{l,k+f|k}, \hat{y}_{u,k+f|k}] \\ 1 & y_{k+f} \in [\hat{y}_{l,k+f|k}, \hat{y}_{u,k+f|k}] \end{cases} \quad k = 1, \dots, T - f$$

donde $\hat{y}_{l,k+f|k}$ y $\hat{y}_{u,k+f|k}$ son los límites superior e inferior respectivamente con un intervalo de confianza del 95 % obtenidos con la expresión 3.42:

$$\hat{y}_{l,k+f|k} = \hat{y}_{k+f|k} - 1.96 \sqrt{var(\hat{y}_{k+f|k})} \quad (3.42)$$
$$\hat{y}_{u,k+f|k} = \hat{y}_{k+f|k} + 1.96 \sqrt{var(\hat{y}_{k+f|k})}$$

3.1.5. Validación condicional

La validación condicional es un proceso mediante el cual se evalúa el potencial de predicción del modelo, esta se realiza en un grupo de datos diferente al utilizado en las etapas de identificación y estimación seleccionado mediante un ensayo *split-sample* [31]. La validación condicional predictiva permite definir o medir la eficacia del modelo. Los modelos son aproximaciones de la realidad diseñadas con un objetivo específico, por lo tanto la validación demuestra si la aproximación es satisfactoria en el limitado sentido predictivo [47].

En el contexto de la modelación precipitación-caudal, la validación condicional permite determinar si las predicciones realizadas con el modelo se encuentran dentro de los límites de incertidumbre asociados con el modelo. En la modelación de las series de tiempo, la validación condicional predictiva es simplemente un diagnostico estadístico que asegura que el modelo tiene propiedades de certeza aceptables [47].

3.1.6. Pronóstico

En el contexto del pronóstico, el objetivo del modelo estimado no solo es explicar bien los datos sino también que el pronóstico tenga el mejor rendimiento para el intervalo de

tiempo que se desea predecir. Si el modelo encontrado no presenta ningún retraso, no se podría realizar un pronóstico debido a la imposibilidad de encontrar la precipitación efectiva en el contexto del sistema precipitación-caudal. Sin embargo, se podría dar un tiempo de retraso al modelo disminuyendo su capacidad de explicar bien los datos pero agregándole la posibilidad de utilizar el modelo para realizar pronósticos.

Asimilación de Datos

Una de las principales ventajas de la TF es la posibilidad de realizar pronósticos a intervalos de tiempo corto, y además, si se combina con técnicas de predicción avanzadas como es el filtro de Kalman [14], que sumado a la recolección en tiempo real de los datos pueden conformarse en un sistema de pronóstico en tiempo real. Sin embargo, se debe tener en cuenta la disposición de instrumentos de medición adecuados para realizar las observaciones y transmitirlos en tiempo real.

Para realizar un pronóstico en línea por medio del filtro de Kalman se puede emplear el siguiente algoritmo para f -intervalos: $f = 1, 2, \dots$:

$$\begin{aligned}x_{t+f} &= F\hat{x}_{t+f-1} \\ P_{t+f} &= FP_{t+f-1}F^T + gQ_r g^T \\ \hat{y}_{t+f} &= H_{t+f}\hat{x}_{t+f} \\ v_{t+f} &= 1 + H_{t+f}P_{t+f}H_{t+f}^T\end{aligned}\tag{3.43}$$

donde la varianza del pronóstico dada por $v_{t+f}^* = \sigma^2 v_{t+f}$. σ^2 debe ser estimada recursivamente si no es conocida *a priori*.

La corrección del pronóstico se puede realizar fuera de línea si es necesario; técnicas como la actualización de los parámetros pueden ser utilizadas cuando se presentan grandes cambios en el sistema causados por cambios físicos como pueden ser el uso del suelo en sistemas precipitación-escorrentía o cambios en las propiedades de canales en el caso de regulación de crecidas [4, 42].

Sin embargo, debido a la limitada información contenida en los datos disponibles de las series de tiempo, no es posible estimar robustamente un gran número de parámetros del modelo en línea que mejore los resultados del pronóstico. Para mejorar el rendimiento en el pronóstico se puede utilizar otros métodos como pueden ser: la estimación de un parámetro de ganancia escalar o del pronóstico del error.

Una simplificación común para la técnica del filtro de Kalman de actualización de estado, es el reemplazo del último estado disponible observado:

$$\hat{Q}_{t+1} = a_1 Q_t + b_0 u_{t-1}$$

$$\hat{Q}_{t+2} = a_1 \hat{Q}_{t+1} + b_0 u_t$$

Este método es equivalente al filtro de Kalman de actualización de estado con la covarianza del ruido observado establecido con valores relativamente bajos a el sistema de la covarianza del ruido así el filtro de Kalman sigue de cerca los resultados. Aunque este método es muy simple en comparación a el filtro de Kalman, tiene la desventaja que el pronóstico probabilístico no puede ser producido.

La actualización de parámetros se refiere a la estimación en línea de los mismos usando un estimador recursivo como es el filtro de Kalman con el sistema de ecuaciones reformulado con el vector estado conteniendo los parámetros.

3.2. Nedbor-Afstromnings Model

3.2.1. Introducción

El sistema de modelación hidrológico NAM, "Nedbor-Afstromnings-Model" fue desarrollado originalmente por el Departamento de Hidrodinámica y Recursos Hídricos de la Universidad Técnica de Dinamarca [27]. Este permite simular los procesos de precipitación-escorrentía que ocurren a escala de la cuenca.

Un modelo hidrológico matemático como el NAM es un conjunto de afirmaciones matemáticas vinculadas entre sí, que describen de forma cuantitativa simplificada el comportamiento de la fase terrestre del ciclo hidrológico [10]. El NAM representa los distintos componentes del proceso de precipitación-escorrentía en forma continua por el contenido de agua en cuatro depósitos diferentes e interrelacionados entre sí. Cada uno representa diferentes elementos físicos de almacenamiento de la cuenca.

El NAM puede ser caracterizado como un modelo determinista, concentrado y conceptual, con requisitos de entrada moderado [31].

3.2.2. Estructura del modelo NAM

El sistema de modelación conceptual NAM está basado en estructuras físicas y ecuaciones semi-empíricas. Al ser un modelo concentrado, el NAM considera cada cuenca como una sola unidad, por lo tanto, los parámetros y las variables representan valores promedio para toda la cuenca; como resultado, algunos de los parámetros del modelo se puede evaluar a partir de datos físicos de la cuenca, pero la estimación de parámetros finales se debe realizar mediante un calibrado con datos hidrológicos observados de la cuenca.

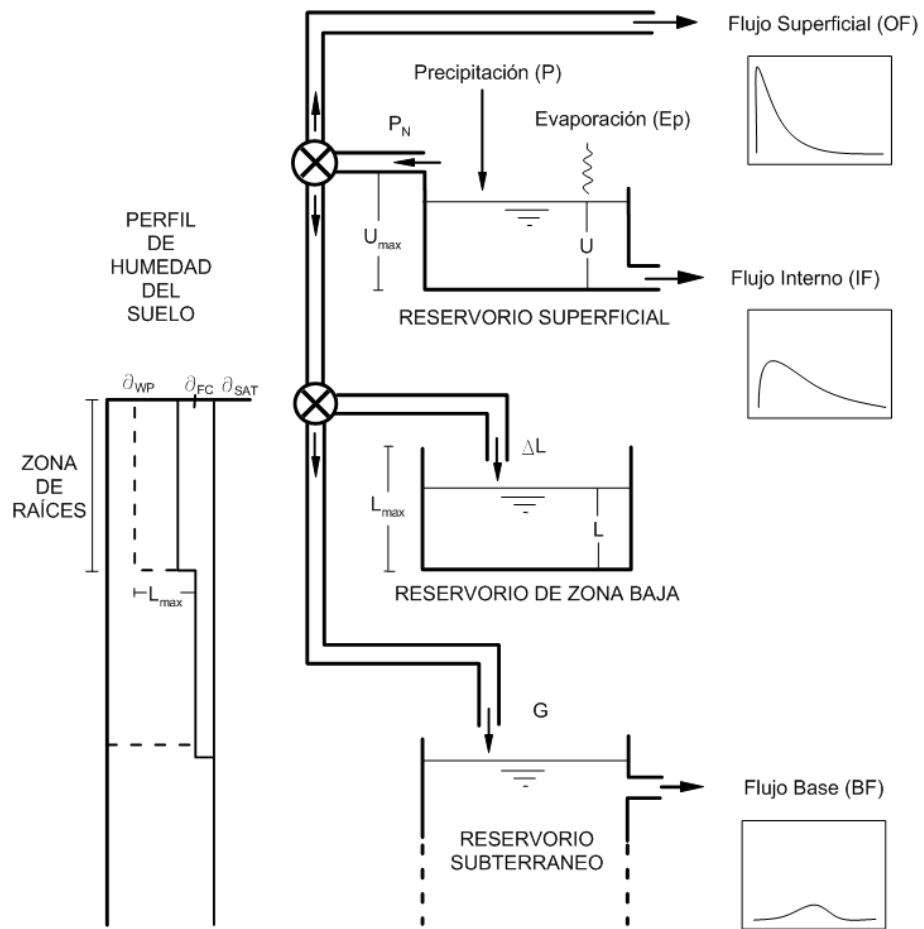


Fig. 3.4: Estructura Simplificada del NAM

La estructura del modelo se muestra en la figura 3.4. Se trata de una representación de la fase terrestre del ciclo hidrológico. El NAM en su versión completa simula el proceso lluvia-escorrentía mediante la representación en forma continua del contenido de agua en cuatro reservorios diferentes e interrelacionados entre sí, que representan diferentes elementos físicos de la cuenca. Estos reservorios son los siguientes [27]:

- Reservorio de nieve,
- Reservorio superficial,
- Reservorio de la zona inferior o zona de raíces,
- Reservorio subterráneo

Basados en los datos meteorológicos de entrada, el NAM es capaz de predecir el escurrimiento de la cuenca, así como la información sobre otros elementos de la fase terrestre del ciclo hidrológico, tales como la variación temporal de la evapotranspiración, la humedad del suelo, recarga de acuíferos y aguas subterráneas. El escurrimiento resultante de la cuenca se divide conceptualmente en tres componentes: flujo superficial, flujo intermedio y flujo de base.

En el presente trabajo no se considera el reservorio de nieve debido a que las características meteorológicas de la cuenca que se va a modelar no presenta condiciones de nieve razón por lo cual se omite este reservorio.

Reservorio Superficial

El reservorio superficial está constituido por la humedad interceptada en la vegetación, el agua atrapada en las depresiones y en la superficie. El límite superior de este reservorio es denotado por U_{max} el mismo que indica la capacidad máxima de almacenamiento de este reservorio.

La cantidad de agua U del reservorio superficial disminuye constantemente debido al consumo por evaporación así como también por el flujo intermedio. Cuando el reservorio superficial excede su capacidad de almacenamiento U_{max} , el exceso de agua P_N empieza a fluir; parte de este produce escorrentía superficial o flujo superficial mientras que la otra parte se infiltra al reservorio de la zona de raíces y al reservorio subterráneo.

Reservorio de la zona inferior o zona de raíces

Este reservorio representa la capa que se encuentra por debajo de la superficie, en donde las raíces de las plantas aprovechan la humedad del suelo para extraer agua para

realizar el proceso de transpiración. El límite máximo de almacenamiento de este reservorio se representa como L_{max} .

La humedad L en el reservorio de la zona inferior está sujeto a la pérdida de consumo por la transpiración. El contenido de humedad controla la cantidad de agua que ingresa en el reservorio subterráneo como recarga, además del flujo intermedio y el flujo superficial.

Evapotranspiración

La demanda de evapotranspiración hasta completar la evapotranspiración potencial es obtenida primero del reservorio superficial. Si el contenido de humedad U del reservorio superficial es menor que la demanda por evapotranspiración ($U < E_p$), la fracción restante es obtenida conceptualmente por las raíces de las plantas del reservorio de la zona inferior a una tasa real de E_a . E_a es proporcional a la evapotranspiración potencial y varía linealmente con la relación del contenido de humedad del suelo, L/L_{max} , del reservorio de la zona baja:

$$E_a = (E_p - U) \frac{L}{L_{max}} \quad (3.44)$$

Flujo Superficial

Cuando el reservorio superficial llega a su máxima capacidad $U > U_{max}$, el exceso de agua P_N empieza a fluir produciendo de esta manera la escorrentía o flujo superficial, así como también la infiltración. La parte de P_N que contribuye a la escorrentía se denota por QOF . QOF se define como proporcional a P_N y varía linealmente con el contenido de humedad del suelo en la relación, L/L_{max} , del reservorio de la zona baja:

$$QOF = \begin{cases} (CQOF) \frac{L}{L_{max}} \frac{1 - TOF}{1 - TOF} P_N & \text{si } \frac{L}{L_{max}} > TOF \\ 0 & \text{si } \frac{L}{L_{max}} \leq TOF \end{cases} \quad (3.45)$$

donde $CQOF$ es el coeficiente de escorrentía del flujo superficial y que es un valor entre ($0 \leq CQOF \leq 1$). TOF es un valor límite para el flujo superficial y puede estar entre ($0 \leq TOF \leq 1$). El valor de $CQOF$ permite determinar que porción del exceso de precipitación se transforma en flujo superficial y la magnitud de la infiltración. El valor de TOF controla la generación del flujo superficial, si la relación $\frac{L}{L_{max}}$ es menor que TOF , el flujo superficial es generado.

La proporción del exceso de agua P_N que no ingresa a la corriente superficial se infiltra al reservorio de la zona baja. Una fracción ΔL , de la cantidad de agua disponible para la infiltración ($P_N - QOF$), se considera que aumenta el contenido de humedad L del reservorio inferior. La parte restante de la humedad de infiltración, G , se filtra hacia las profundidades y recarga el reservorio de agua subterránea.



Flujo Interno

El flujo interno QIF se asume que es proporcional a U y varía linealmente con la humedad relativa del reservorio de la zona de raíces:

$$QIF = \begin{cases} (CKIF)^{-1} \frac{L}{L_{max}} \frac{1-TIF}{1-TIF} U & \text{si } \frac{L}{L_{max}} > TIF \\ 0 & \text{si } \frac{L}{L_{max}} \leq TIF \end{cases} \quad (3.46)$$

donde $CKIF$ es una constante de tiempo para el flujo interno, y TIF es un valor límite del flujo interno ($0 \leq TIF \leq 1$). El valor de TIF cumple la misma función de TOF pero para el flujo interno.

Escorrentía del flujo superficial y flujo interno

El flujo interno es transitado a través de dos reservorios lineales en serie que tienen la constante CK_{12} similar. El flujo superficial también está basado en el concepto de reservorio lineal, pero con una constante de tiempo variable:

$$CK = \begin{cases} CK_{12} & \text{si } OF < OF_{min} \\ CK_{12} \frac{OF}{OF_{min}}^{-\beta} & \text{si } OF \geq OF_{min} \end{cases} \quad (3.47)$$

donde OF es el flujo superficial, OF_{min} es el límite superior para el tránsito lineal, y $\beta = 0.4$. β es una constante que corresponde al coeficiente Manning para modelar el flujo superficial.

La constante CK_{12} determina la forma de los picos en el hidrograma, sus valores dependen principalmente del tamaño de la cuenca y la respuesta de esta a la precipitación, el CK_{12} se encuentra entre 3 – 48 horas, [10]

Recarga de acuíferos

La infiltración del agua G que llega a recargar los acuíferos depende del contenido de humedad del reservorio de la zona de raíces:

$$G = \begin{cases} (P_N - QOF) \frac{L}{L_{max}} \frac{1-TG}{1-TG} & \text{si } \frac{L}{L_{max}} > TG \\ 0 & \text{si } \frac{L}{L_{max}} \leq TG \end{cases} \quad (3.48)$$

donde TG es el valor umbral límite para la recarga del acuífero.

Contenido de humedad del suelo

El reservorio de la zona inferior representa el contenido de agua de la zona de raíces. La humedad L de este reservorio se incrementa en ΔL con el valor residual de la precipitación neta después de repartirla a la escorrentía superficial y a la infiltración hacia el



reservorio subterráneo.

$$\Delta L = P_N - QOF - G \quad (3.49)$$

Flujo base

El flujo base BF descargado del reservorio subterráneo es calculado como un reservorio lineal con tiempo constante CK_{BF} . CK_{BF} determina la forma del hidrograma en periodos de sequía, según la descripción de reservorio lineal, la descarga de va disminuyendo exponencialmente en el tiempo.

Condiciones Iniciales

El proceso de simulación con el sistema de modelación NAM consiste en la estimación de caudales a partir de las medidas de precipitación y evapotranspiración; el caudal que es escurrido hacia la fuente es la suma de tres tipos de flujos con características diferentes y su valor depende del contenido de humedad del suelo (reservorio de la zona de raíces). El contenido de humedad del suelo es función del tiempo, del caudal descargado el día anterior y de la precipitación. Por lo tanto, el modelo matemático es representado por un sistema de ecuaciones diferenciales de la forma 3.50:

$$\begin{aligned} \frac{dU}{dt} &= f_1(U, L, t) = P - IF - P_N - E_a \\ \frac{dL}{dt} &= f_2(U, L, t) = \Delta L - E_a \end{aligned} \quad (3.50)$$

Para resolver este sistema de ecuaciones diferenciales de primer orden (EDO) se puede utilizar algoritmos de integración numérica (Trapezio, Runge-Kutta, etc...) introduciendo sus respectivas condiciones de frontera o valores iniciales ($U(t_0) = U_0$ y $L(t_0) = L_0$).

3.2.3. Calibración

En el NAM, los parámetros y las variables representan los valores promedio para toda la cuenca. Aunque en algunos casos, una serie de valores de los parámetros probablemente pueden ser estimados; no es posible, en general determinar los valores de los parámetros del NAM sobre la base de las características físicas, climáticas y de suelo de la cuenca, puesto que la mayoría de los parámetros son de naturaleza empírica y conceptual. Por lo tanto, la estimación de parámetros finales se debe realizar mediante una calibración con las series de tiempo observadas disponibles.

En el proceso de calibración, se debe tener en cuenta los objetivos de la modelación de acuerdo al uso previsto del modelo. Si los objetivos son de igual importancia, se debe



encontrar un equilibrio entre todos; caso contrario se debe favorecer la prioridad a un determinado objetivo.

La evaluación del rendimiento incluye la comparación gráfica de los hidrogramas simulados y observados, y la comparación de la escurriencia acumulada simulada y observada. Las medidas de rendimiento numérico incluye el error del balance hídrico global (es decir, la diferencia entre el promedio de escurrimiento simulado y observado), una medida de la forma general del hidrograma basado en el coeficiente Nash-Sutcliffe [25], y también el RSME, tanto para el error total como para el rendimiento de los caudales picos y caudales bajos [20].

Es importante tener en cuenta que la bondad de ajuste del modelo calibrado se ve afectada por diferentes fuentes de error como: error en la medición de datos meteorológicos, datos observados, errores conceptuales debido a la estructura del modelo y debido a la utilización de parámetros con valores que no son óptimos.

El NAM puede calibrarse de una manera manual y de una manera automática mediante una función multi-objetivo [21, 10]. La calibración manual requiere de ensayos prueba-error hasta que el rendimiento del modelo se considere aceptable de acuerdo a los objetivos planteados; se puede tomar en cuenta algunas consideraciones presentadas en [10] para este efecto como son:

- Realizar un ajuste del balance de agua del sistema. La precipitación acumulada neta menos la escurriencia debe ser igual a la evapotranspiración acumulada para un periodo determinado; la evapotranspiración podría incrementarse aumentando la capacidad de los reservorios superficial U_{max} y de la zona de raíces L_{max} .
- Los caudales picos son causados por un mayor flujo superficial. El caudal pico puede ajustarse mediante el cambio del coeficiente del flujo superficial ($CQOF$) y su forma por medio de la constante de escurriencia CK_{12} .
- El caudal base depende de los demás caudales, cualquier cambio en el caudal superficial o caudal interno afecta al caudal base. La forma de recesión del caudal base se ve afectada por la constante de tiempo de caudal base $CKBF$
- El valor inicial de los parámetros TOF , TIF , y TG puede ser cero para la primera simulación.

Un resumen del efecto de cambio en los parámetros se presenta en la siguiente tabla [35]:

Parámetro	Cambio	Efecto
Lmax	Incremento	Reducción del flujo pico, disminución del volumen de escorrentía
Umax	Incremento	Reducción del flujo pico, disminución del volumen de escorrentía
CQOF	Incremento	Reducción del flujo pico, disminución del volumen de escorrentía
TOF	Incremento	Reducción del flujo pico, disminución del volumen de escorrentía
CK1, CK2	Incremento	Reducción del flujo pico, la forma triangular se expande horizontalmente
CKBF	Incremento	Reducción del flujo base
Profundidad del agua subterránea	Incremento	Reducción del flujo pico, disminución del volumen de escorrentía

Tab. 3.1: Efecto del cambio de los diferentes parámetros

Por otro lado, se puede realizar una calibración automática de los parámetros mediante una función multi-objetivo [21, 10]. Para este tipo de calibración se han definido medidas numéricas del rendimiento como:

- Error medio Absoluto (MAE) [17]

$$MAE = F_1(\theta) = \left| \frac{1}{N} \sum_{i=1}^N [y_i - y_i^*] \right| \quad (3.51)$$

- Error medio cuadrático ($RMSE$)

$$RMSE = F_2(\theta) = \left[\frac{1}{N} \sum_{i=1}^N [y_i - y_i^*]^2 \right]^{\frac{1}{2}} \quad (3.52)$$

- Error medio cuadrático de los eventos picos ($RMSE$)

$$RMSE = F_3(\theta) = \frac{1}{M_p} \sum_{j=1}^{M_p} \left[\frac{1}{n_j} \sum_{i=1}^{n_j} [y_i - y_i^*]^2 \right]^{\frac{1}{2}} \quad (3.53)$$

en donde M_p es el número de eventos de flujo pico, y n_j es el número de intervalos de tiempo en el evento j .

- Error medio cuadrático de los eventos de flujo bajo ($RMSE$)

$$RMSE = F_4(\theta) = \frac{1}{M_l} \sum_{j=1}^{M_l} \left[\frac{1}{n_j} \sum_{i=1}^{n_j} [y_i - y_i^*]^2 \right]^{\frac{1}{2}} \quad (3.54)$$

en donde M_l es el número de eventos de flujo bajo, y n_j es el número de intervalos de tiempo en el evento j .

Los eventos se definen como picos o bajos de acuerdo a un umbral especificado por el modelador.

Los resultados del problema de calibración usando múltiples objetivos no proporciona una única solución, el resultado es un pareto de soluciones óptimas. Una posible solución a este problema puede ser la utilización de una función objetivo concentrada [19]:

$$F_{con}(\theta) = \left[\sum_{i=1}^4 [F_i(\theta) + A_1]^2 \right]^{\frac{1}{2}} \quad (3.55)$$

donde A_i son constantes de transformación cuyo objetivo es reflejar la importancia de los diferentes objetivos.

La función objetivo 3.55 se puede optimizar automáticamente usando el SCE ([11]).

El problema de optimización está restringido en el sentido de θ , el mismo que está limitado al espacio de parámetros Θ :

$$\theta \in \Theta$$

El espacio de los parámetros se define como un hipercubo mediante la definición de los límites inferior y superior de cada parámetro. Estos deben ser elegidos de acuerdo a las restricciones físicas y matemáticas en el modelo y de las experiencias del modelador; la tabla 3.2.3 desarrollada por [10] muestra los posibles límites de espacio de parámetros.

Parámetro	Unidad	Límite Inferior	Límite Superior
U_{max}	[mm]	5	35
L_{max}	[mm]	50	400
CQOF	[-]	0	1
CKIF	[hora]	200	2000
CK _{1,2}	[hora]	3	72
TOF	[-]	0	0.9
TIF	[-]	0	0.9
TG	[-]	0	0.9
CKBF	[hora]	500	5000

Tab. 3.2: Valores límites del espacio de parámetros

3.2.4. Validación

La validación del modelo encontrado se realiza de la misma forma que en el DBM descrito en secciones anteriores. Tomando un grupo de datos diferente al utilizado en las etapas de identificación y estimación, seleccionado mediante un ensayo *split-sample* [31]. La validación condicional predictiva permite definir o medir la eficacia del modelo, es simplemente un diagnóstico estadístico que asegura que el modelo tiene propiedades de certeza aceptables [47].

3.3. Comparación de los modelos

Para evaluar el rendimiento de los modelos de la prueba SS, se han establecidos varios criterios con el objetivo principal de medir cuanto se ajustan las series tiempo simuladas a las series de tiempo medidas.

Se ha establecido un conjunto de criterios que comprenden una combinación de las siguientes representaciones gráficas y medidas numéricas: (1) hidrogramas de caudal de

los datos observados y y simulados y^* ; (2) Scatter-plot de los datos observados vs. simulados de los caudales diarios; (3) Caudal medio diario; (4) volumen total de descarga por periodo; (5) el coeficiente de variación (CV); (6) el coeficiente de Nash-Sutcliffe ($R2$); (7) Coeficiente de Determinación (CD); (8) coeficiente de correlación (CC) y (9) el índice de eficiencia de grado dos d_2 y de grado tres d_3 .

Coeficiente de Variación se define como [43]:

$$CV = \frac{\sqrt{\sum_{k=1}^N (y_k^* - y_k)^2}}{\bar{y}} \quad (3.56)$$

El coeficiente de ($R2$), introducido por [25], se obtiene de la base de datos diarios observados y simulados durante el período de prueba y en el periodo de validación:

$$R2 = 1 - \frac{\sum_{k=1}^N y_k - y_k^{*2}}{\sum_{k=1}^N y_k - \bar{y}^2} \quad (3.57)$$

donde:

\bar{y} = caudal medio diario;

cuando el rendimiento del modelo es bueno el $R2$ tiende a la unidad.

El coeficiente de determinación CD , es la medida de la proporción de la varianza total de los datos observados explicada por los datos predichos [40]:

$$CD = \frac{\sum_{k=1}^N (y_k - \bar{y})^2}{\sum_{k=1}^N (y_k^* - \bar{y})^2} \quad (3.58)$$

El valor óptimo de CD es 1, este coeficiente es útil para caracterizar caudales picos; valores entre $0 < CD < 1$ nos indica que el modelo es sobreestimado especialmente en los caudales picos. Valores entre $1 < CD < +\infty$ nos indican subestimación del modelo.

Otros índices utilizados para medir el rendimiento de los modelos, son los coeficientes de bondad de ajuste de grado dos y de grado tres, los mismos que son sensibles a los picos:

$$d_2 = 1 - \frac{\sum_{k=1}^N (y_k - y_k^*)^2}{\sum_{k=1}^N (|y_k^* - \bar{y}| + |y_k - \bar{y}|)^2} \quad (3.59)$$

$$d_3 = 1 - \frac{\sum_{k=1}^N |y_k - y_k^*|^3}{\sum_{k=1}^N (|y_k^* - \bar{y}| + |y_k - \bar{y}|)^3} \quad (3.60)$$

Los coeficientes que nos permiten medir el ajuste de los datos son calculados para los periodos de calibración y validación de los dos modelos, lo cual permite evaluar el rendimiento del modelo y su capacidad para pronosticar.

EXPERIMENTACIÓN NUMÉRICA

4.1. Ubicación y características geográficas de la cuenca de análisis

La cuenca hidrográfica del presente estudio se localiza en la región centro sur del Ecuador, y consiste de la parte alta de la cuenca del río Paute. Se ubica entre los paralelos $2^{\circ}15'$ y $3^{\circ}15'$ de latitud sur y entre los meridianos $78^{\circ}30'$ y $79^{\circ}20'$ de latitud oeste. El área de estudio comprende tres subcuencas: la subcuenca del río Tomebamba con una extensión aproximada de 380.418km^2 , la subcuenca del río Yanuncay con una área aproximada de 418.888km^2 y la subcuenca del río Tarqui con una área aproximada de 476.129km^2 . La zona de estudio es accidentada con valles y llanuras, con un rango de elevación que va desde 2440 hasta los 4420 m.s.n.m.

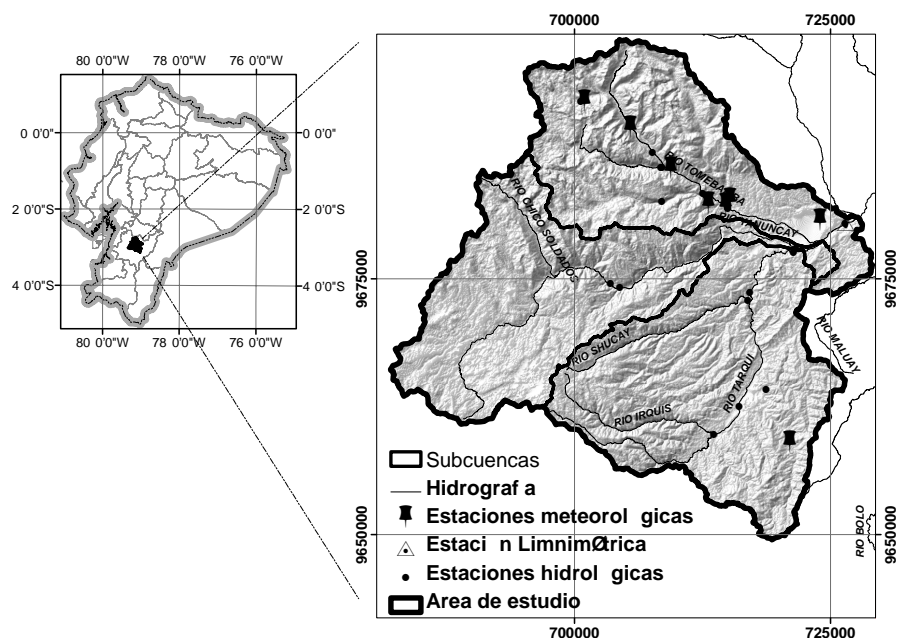


Fig. 4.1: Ubicación de la cuenca de estudio

El área de estudio total es aproximadamente de 1275.4km^2 la misma que cuenta con una de sus estaciones hidrométricas a la salida del área de estudio y además cuenta con varias estaciones meteorológicas en toda su área como se muestra en la figura 4.1.

Para la presente modelación, se contó con series temporales de caudal y lluvia diarios, así como también de una variable bio-edafo-climática como es la evapotranspiración potencial diaria (ET_p) en el período de [1 de julio de 1981 - 27 de marzo de 1985]. En este período, se ha registrado una lluvia media diaria de $2.39mm$ (aproximadamente $778.70mm$ por año en promedio). La ET_p media diaria de $2.27mm$ (aproximadamente $824.24mm$ por año en promedio). El caudal diario fue del orden de $15.43m^3$ (aproximadamente $3588.19m^3$ por año en promedio).

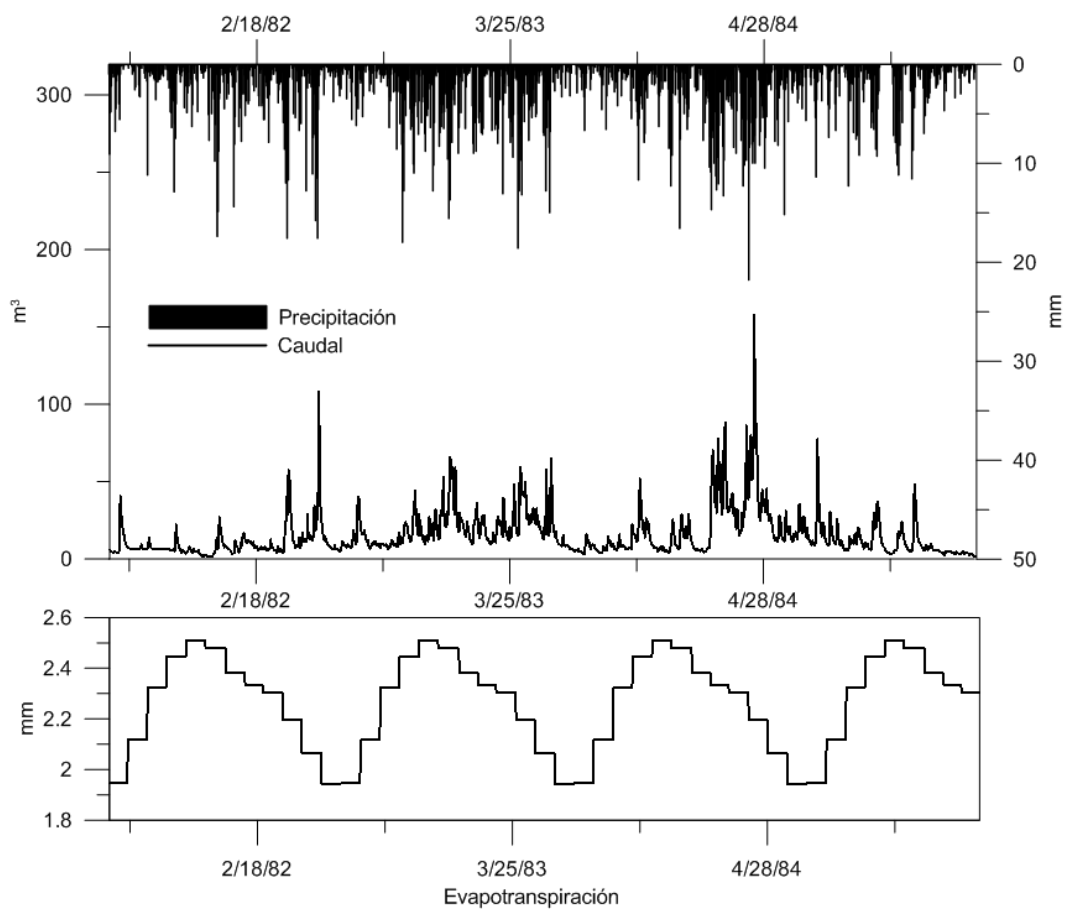


Fig. 4.2: Datos observados de precipitación-caudal y evapotranspiración estimada de la subcuenca del Paute

4.2. Modelación DBM de la cuenca

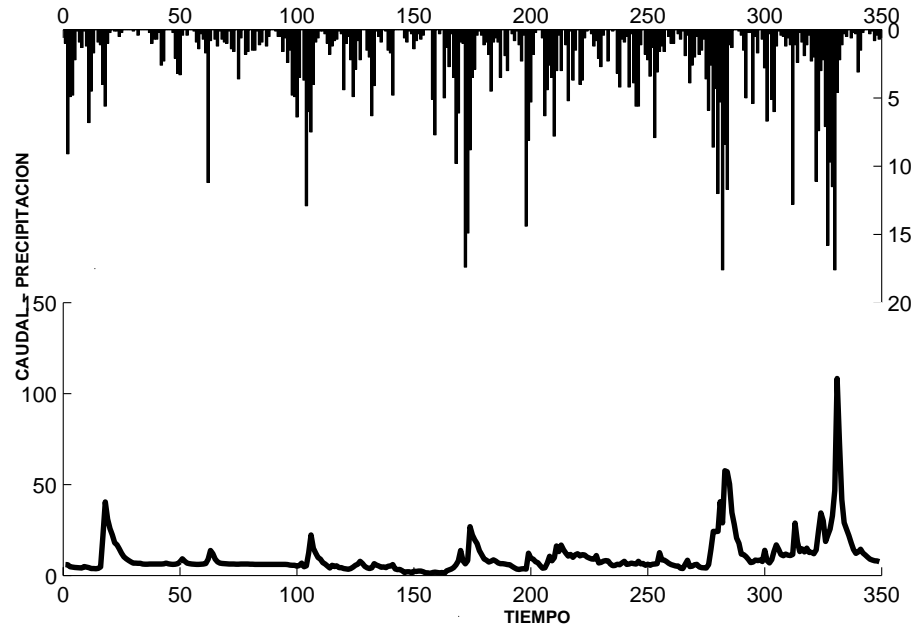


Fig. 4.3: Datos observados de precipitación-caudal de la subcuenca del Paute

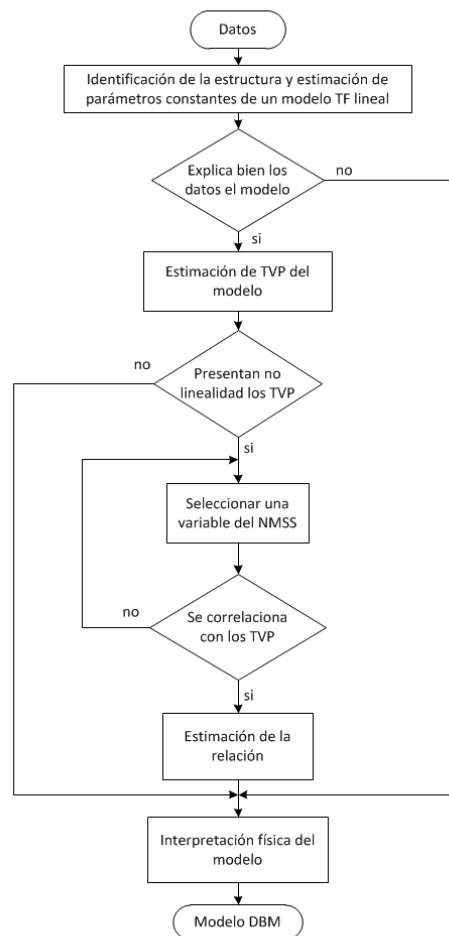


Fig. 4.4: Diagrama de flujo para la modelación DBM del sistema precipitación-caudal

4.2.1. Identificación y estimación

Como se vió en capítulos anteriores, para la identificación y estimación del modelo se utilizó un ensayo "split-sample". De esta forma, y en función de la disponibilidad de datos para la siguiente modelación, se definieron intervalos: [1 de julio de 1981 - 3 de octubre de 1983] para identificación y estimación, y [4 de octubre de 1983 - 27 de marzo de 1985] para validación.

La elección de la estructura óptima del modelo se encontró analizando varias estructuras de un rango escogido sin restricción alguna, desde [1 1 0] a [3 3 8], las estructuras analizadas son presentadas en la tabla 4.1, las mismas que fueron encontradas utilizando la función SRIV; esta y otras funciones utilizadas en el presente trabajo se encuentran dentro del toolbox de Matlab denominado CAPTAIN el mismo que se encuen-

tra disponible de manera gratuita en www.es.lancs.ac.uk/cres/captain.

La tabla ha sido estimada para la relación precipitación efectiva-caudal en la cual de una manera implícita se toma en cuenta el almacenamiento de la cuenca, como se muestra en 3.1.4. Dicha tabla muestra que el modelo de orden tres [3 3 0] puede ser utilizado para la modelación debido a que explica de mejor manera los datos con un R^2 de 0.803, sin embargo, este modelo muestra evidencia de que podría no ser el más óptimo al tener un YIC de 0.802. El modelo de primer orden muestra un valor del índice YIC con el menor valor lo cual mostraría que es el modelo parametrizado de la mejor manera, sin embargo, su índice R^2 con un valor de 0.725 muestra no ser el modelo que explica de mejor manera los datos.

n	m	r	δ	R^2	YIC	AIC	CD	D2	D3	EF	CC
3	3	0	0	0.802	3.463	3.299	0.923	0.943	0.987	0.774	0.904
2	3	0	0	0.8	-1.007	3.308	0.917	0.943	0.987	0.773	0.904
3	2	0	0	0.79	-4.549	3.355	0.821	0.936	0.986	0.732	0.905
1	3	0	0	0.785	-4.352	3.379	0.867	0.939	0.986	0.749	0.9
2	3	1	0	0.774	-4.34	3.43	1.031	0.936	0.981	0.761	0.885
3	2	1	0	0.758	-1.284	3.499	0.976	0.932	0.98	0.739	0.879
1	3	1	0	0.752	-4.745	3.521	0.95	0.929	0.979	0.726	0.878
3	1	0	0	0.743	-4.527	3.557	0.751	0.923	0.982	0.665	0.891
2	2	1	0	0.74	-6.163	3.567	0.891	0.929	0.979	0.716	0.877
1	2	0	0	0.736	-5.377	3.583	0.782	0.92	0.979	0.659	0.884
1	2	1	0	0.73	-7.283	3.602	0.912	0.922	0.976	0.693	0.869
1	1	0	0	0.725	-7.974	3.618	0.73	0.915	0.979	0.626	0.885
1	1	1	0	0.63	-7.339	3.918	0.747	0.884	0.96	0.503	0.836
1	3	2	0	0.577	-3.529	4.055	1.08	0.872	0.935	0.555	0.779
1	2	2	0	0.561	-6.333	4.091	1.092	0.865	0.928	0.533	0.769
2	2	2	0	0.544	-4.871	4.131	0.868	0.879	0.945	0.528	0.788
2	3	2	0	0.459	-3.586	4.303	0.673	0.878	0.944	0.451	0.797

Tab. 4.1: Identificación de la estructura

El modelo determinado anteriormente (modelo no-lineal) se puede comparar con un modelo lineal, es decir, un modelo que relacione la precipitación con el caudal directamente sin ninguna transformación. Los resultados se muestran simultáneamente en la tabla 4.2, donde el modelo lineal de orden [3 3 0] no parece ser el modelo que explique mejor los datos como lo es el modelo no-lineal a pesar de que los dos modelos están sobre parametrizados. De la tabla 4.2 se puede confirmar que el sistema precipitación-caudal se comporta de forma no-lineal como lo proponen la mayoría de autores.

Siendo el modelo no lineal, se ha escogido la estructura [2 3 0] que si bien no es la que explica mejor los datos su diferencia es de solo dos milésimas del mejor, pero además, el mejor modelo esta sobre parametrizado por lo cual estaría en contra de la filosofía del DBM. Sin embargo, el modelo escogido tampoco es el más parsimonioso, la razón por la que se ha seleccionado el modelo se basa también en las recomendaciones presentadas en [42] en el cual sugiere, basado en la experimentación, que si se dispone de una base de datos precipitación-caudal diarios los modelos podrían ser generalmente de tercer orden; lo que implica que el comportamiento del sistema precipitación-caudal se



Modelo			Precipitación-Caudal						Precipitación efectiva-Caudal					
n	m	delay	R2T	YIC	AIC	CD	EF	CC	R2T	YIC	AIC	CD	EF	CC
3	3	0	0.60	0.65	4.02	1.34	0.59	0.78	0.80	3.48	3.30	0.92	0.63	0.89
2	3	0	0.62	-2.36	3.95	1.40	0.62	0.79	0.80	-1.00	3.31	0.92	0.50	0.84
3	2	0	0.66	0.20	3.85	1.53	0.66	0.81	0.79	-4.55	3.36	0.82	0.32	0.67
1	3	0	0.66	-2.75	3.84	1.53	0.66	0.81	0.79	-4.35	3.38	0.87	0.23	0.59
2	3	1	0.63	-1.43	3.92	1.43	0.63	0.80	0.77	-4.34	3.43	1.03	0.18	0.53
3	2	1	0.63	-1.00	3.92	1.44	0.63	0.80	0.76	-1.30	3.50	0.98	0.14	0.49
1	3	1	0.66	-1.97	3.84	1.54	0.66	0.81	0.75	-4.75	3.52	0.95	0.10	0.45
3	1	0	0.66	-5.20	3.85	1.57	0.66	0.81	0.74	-4.53	3.56	0.75	-0.03	0.40
2	2	1	0.61	-4.58	3.97	1.40	0.61	0.78	0.74	-6.16	3.57	0.89	-8.94	-0.39
1	2	0	0.65	-3.56	3.88	1.52	0.65	0.80	0.74	-5.38	3.58	0.78	0.66	0.88
1	2	1	0.66	-4.35	3.85	1.52	0.66	0.81	0.73	-7.28	3.60	0.91	0.69	0.87
1	1	0	0.63	-7.11	3.92	1.67	0.63	0.79	0.73	-7.97	3.62	0.73	0.53	0.77
1	1	1	0.64	-7.26	3.90	1.49	0.64	0.80	0.63	-7.34	3.92	0.75	0.40	0.68
1	3	2	0.57	-2.45	4.07	1.73	0.57	0.76	0.58	-3.53	4.06	1.08	0.29	0.61
1	2	2	0.57	-4.75	4.08	1.70	0.57	0.75	0.56	-6.33	4.09	1.09	-29.16	-0.29
2	2	2	0.52	-4.40	4.18	1.64	0.52	0.72	0.54	-4.87	4.13	0.87	-4.36	-0.15
2	3	2	0.56	-2.06	4.09	1.71	0.56	0.75	0.47	-3.60	4.29	0.68	0.03	0.45

Tab. 4.2: Comparación entre los modelos lineal y no lineal para identificar el modelo óptimo

representa de la mejor manera por medio de tres reservorios los mismos que tiene tiempos de respuesta diferentes. En el modelo escogido también se puede observar que no presenta ningún retraso, es decir, el tiempo que transcurre desde que la precipitación es medida hasta que el caudal es registrado en las estaciones limnimétricas es menor a un día, lo cual también se puede confirmar visualmente en el hidrograma precipitación-caudal de la figura 4.3.

El modelo no-lineal identificado anteriormente, consiste en dos partes, las mismas que fueron mencionadas con mayor detalle en capítulos anteriores. El siguiente paso es la estimación y calibración de los parámetros del modelo; la estimación es realizada utilizando el SDP [9] para la parte no-lineal, cuyo resultado es la estimación no paramétrica de una función no-lineal entre el caudal y el parámetro $b_{0,t}$, la misma que es representada en la figura 4.5 con una línea continua delgada, también se encuentra graficada la incertidumbre de la estimación no lineal, representada con líneas de puntos.

La estimación final de la parte no-lineal del modelo es parametrizar esta relación utilizando cualquier función; para el presente trabajo se utilizó una función potencial de la forma $f(y_t) = cy_t^\beta$; para optimizar esta función se puede utilizar cualquier método de optimización como por ejemplo el simplex [27] el cual fue utilizado en el presente trabajo.

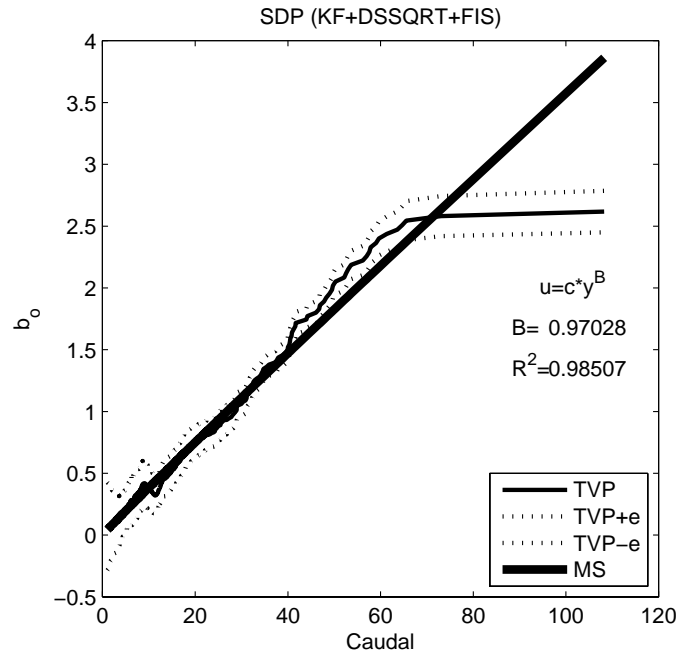


Fig. 4.5: Parámetros Dependientes del Estado para la estimación de la precipitación efectiva.

La función no lineal puede ser de cualquier forma, en el presente trabajo se ha escogido la función potencial siguiendo recomendaciones de la bibliografía, sin embargo se ha analizado otro tipo de funciones no lineales para comparar su bondad de ajuste; en este caso mediante el índice R^2 ; los resultados se presentan en la tabla 4.3, donde se puede observar que la función polinomial de segundo grado proporciona un índice $R^2 = 0.997$ mejor que el obtenido con la función potencial $R^2 = 0.991$; sin embargo, la función potencial requiere menos parámetros para obtener un valor de ajuste elevado lo cual permite concluir que el modelo es parsimonioso. Finalmente, la función pueden ser cualquiera de las presentadas en la tabla o de otra forma con el único objetivo de representar de manera adecuada los datos.

Luego de haber estimado la función no lineal, el siguiente paso es la calibración de la función de transferencia, la TF se calibra reestimando los parámetros de la misma después de haber hecho la transformación no lineal de la precipitación a precipitación efectiva, la misma que es la entrada al TF. Los parámetros de la TF son estimados utilizando el RIV, función disponible también en CAPTAIN [9]. Los parámetros obtenidos para la estructura seleccionada anteriormente se muestran en la ecuación siguiente:

$$y_t = \frac{0.113(0.0018) - 0.016(0.0016)z^{-1} - 0.089(0.0001)z^{-2}}{1 - 1.454(0.0004)z^{-1} + 0.464(0.0002)z^{-2}} u_{t-\delta} \quad (4.1)$$

En la función anterior, la varianza de los parámetros estimados se encuentran entre paréntesis; se observa que el orden de magnitud de la varianza es muy baja, además del



Modelo	R^2
$a_0 + \frac{a_3}{1 + (\frac{y_t}{a_2})^{a_1}}$	0.997
$a_0 + a_1 y_t + a_2 y_t^2$	0.996
$a_0(1 - e^{-a_1 y_t})$	0.995
$\frac{a_0 y_t}{a_1 + y_t}$	0.995
$a_0 + \frac{a_3 - a_0}{1 + e^{a_1(\ln y_t - a_2)}}$	0.991
$\frac{a_0 y_t^{a_1}}{a_0 y_t^{a_1} + a_2}$	0.991
$a_0 e^{a_1 y_t} + a_2$	0.988
$\frac{a_2}{1 + e^{-a_0 - a_1 y_t}}$	0.952
$\frac{10 y_t - a_0}{1 + 10 y_t - a_0} a_1$	0.648
$a_0 * \ln y_t + a_1$	0.632
$a_0 \log_{10} y_t + a_1$	0.632
$a_0 e^{(a_1 y_t)}$	0.578
$\frac{1}{1 + e^{(-a_0 - a_1 y_t)}}$	0.286

Tab. 4.3: Parametrización de los parámetros dependientes del estado por medio de varias funciones

valor de R^2 es alto del orden de 0.93, por lo que el modelo puede considerarse robusto. La figura 4.6 presenta el modelo simulado y los datos medidos, en la misma se puede hacer un análisis visual del modelo, se aprecia que el modelo presenta un buen ajuste a la forma del hidrograma de los datos observados. Sin embargo, el modelo en general está sobreestimado especialmente en los picos, en donde es más notable; en los valles en cambio se observa que se ajusta de mejor manera.

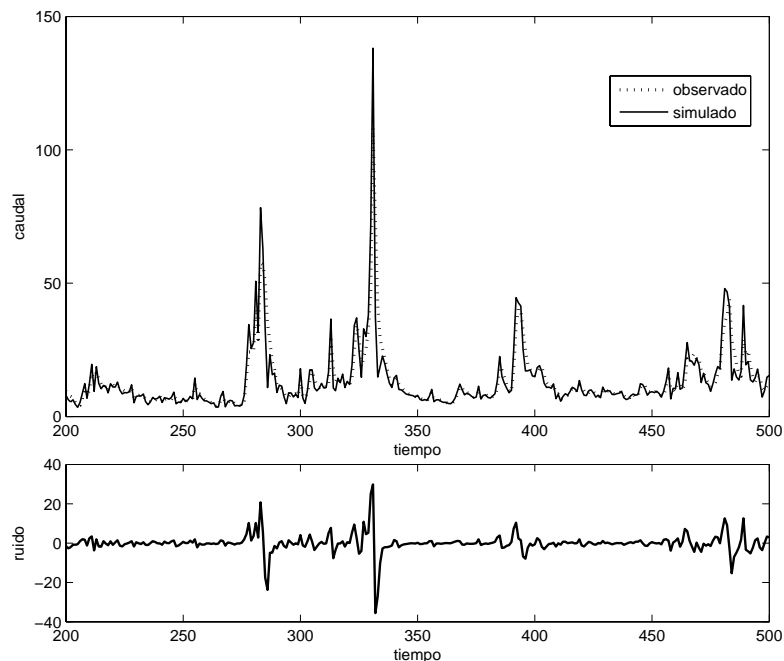


Fig. 4.6: Simulación utilizando la función de transferencia

El análisis se puede hacer cuantitativamente mediante índices que nos permitan de-

terminar la bondad de ajuste del modelo; los índices utilizados y su respectivo valor se encuentran incrustados en la figura 4.7, que es la gráfica de dispersión del modelo simulado vs. el observado. En esta gráfica se observa claramente que el modelo es sobreestimado especialmente en los picos, el índice CD lo confirma al tener una magnitud menor a la unidad; también muestra que el modelo presenta un buen ajuste en los valles, mientras que en eventos de magnitud intermedia el modelo presenta una dispersión notable, la misma que debe ser considerada de acuerdo al objetivo de la modelación.

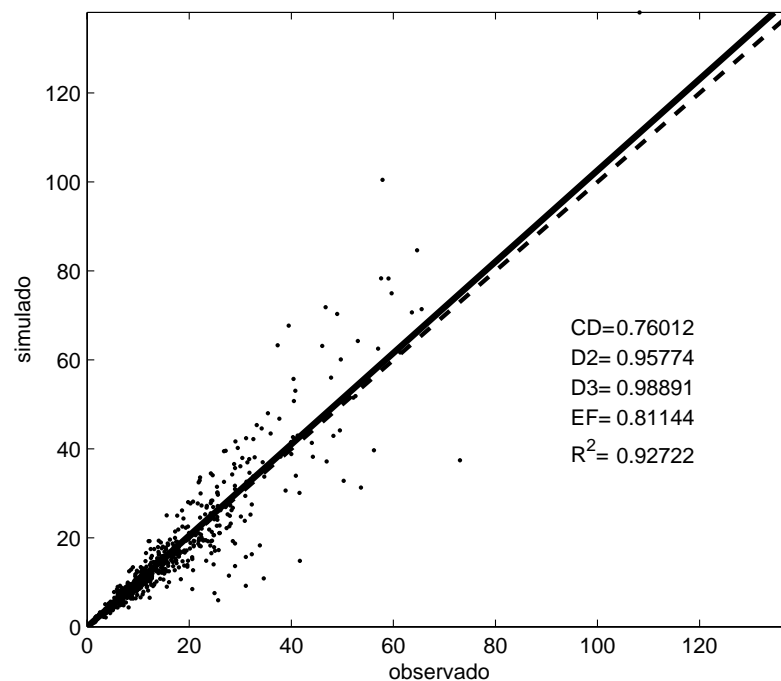


Fig. 4.7: Gráfico de dispersión

4.2.2. Interpretación física

En términos físicos, la función no lineal define la conexión entre la precipitación medida y la *precipitación efectiva*, como se describió en los capítulos anteriores, la precipitación efectiva es la causante de las variaciones del caudal. La relación entre la precipitación efectiva y el caudal es identificada como lineal, dinámica y estocástica, esta relación es expresada mediante una función de transferencia lineal (TF) discreta cuyo impulso de respuesta representa el hidrograma unitario.

La TF del modelo puede ser descompuesta y escrita de la siguiente forma:

$$\hat{y}_t = 0.113u_{t-\delta} + \frac{0.982}{1 - 0.009z^{-1}}u_{t-\delta} + \frac{0.472}{1 - 0.140z^{-1}}u_{t-\delta} \quad (4.2)$$



donde \hat{y} considera al modelo como determinista. El modelo descompuesto se representa gráficamente como en la figura 3.3 mostrada en capítulos anteriores; los detalles de esta descomposición se presentan en la tabla 4.4.

Flujo	Tiempo de Residencia (días)	Fracción del Volumen Total (%)	SS	Eigenvalores
Flujo Lento	58.32	58	0.510	0.010
Flujo Rápido	1.331	30	0.260	0.140
Flujo Instantáneo	<1	13		0.11

Tab. 4.4: Análisis Físico del modelo

La ecuación anterior se puede interpretar de la siguiente forma: la precipitación efectiva u_t llega hasta el río y afecta su caudal aportando al mismo por medio de tres caminos diferentes: uno muy rápido *instantáneo* que representa un 13% del caudal total y que se debe probablemente a procesos superficiales; un componente de tipo *flujo rápido* con un tiempo de residencia de 1.331 días, que aporta en un 30% al caudal y que probablemente se debe a procesos de poca profundidad; y un componente de *flujo lento* con un tiempo de residencia de 58,32 días que aporta en un 58% al cauce total del río siendo además su componente principal, el mismo que es asociado a procesos sub-superficiales y subterráneos debido a la recarga de los acuíferos. También se puede notar que los componente de flujo *rápido* y *lento* de la ecuación 4.2 pueden ser interpretados directamente como ecuaciones de balance de masa o ecuaciones de reservorios.

La figura 4.8 muestra la descomposición paralela en subflujos del caudal: componente de flujo lento, el flujo rápido y el flujo instantáneo.

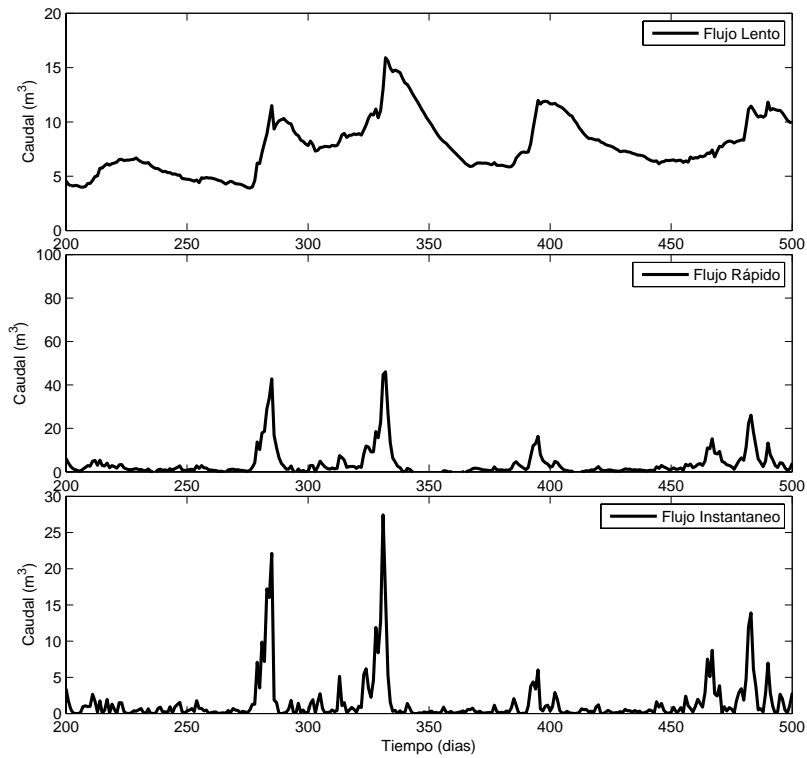


Fig. 4.8: Descomposición de flujo total en subflujos

El efecto de los subflujos se encuentran agrupados en la figura 4.9; en la misma se muestra como aportan al caudal total cada subflujo y como definen la forma del hidrograma: flujo lento, flujo rápido y flujo instantáneo; en la misma gráfica la línea entrecortada representa los datos observados con los cuales se puede comparar de manera visual el ajuste del modelo.

Cada componente tiene una característica temporal como es su tiempo de recesión que permite describir su comportamiento físico como se mencionó en capítulos anteriores.

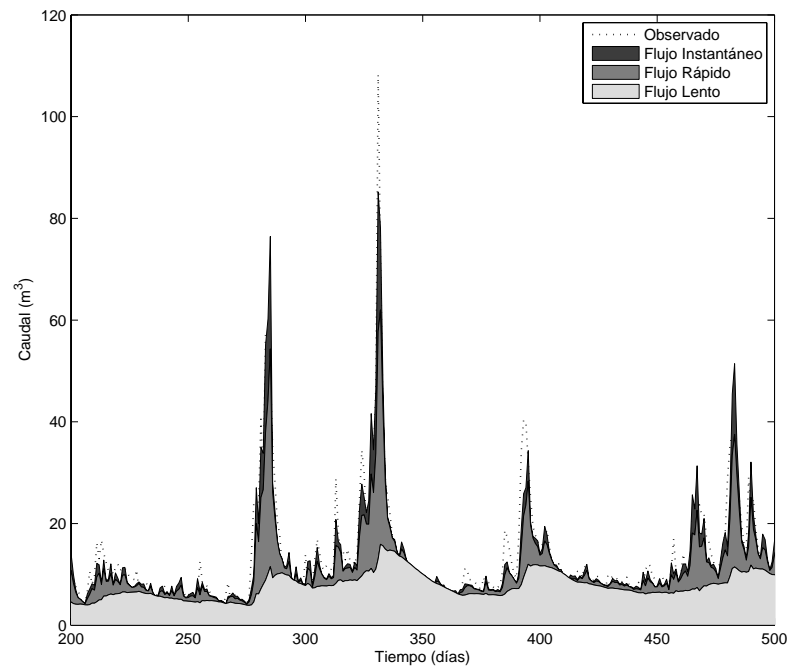


Fig. 4.9: Suma de subflujos

Finalmente, en la gráfica 4.10 se presenta el modelo ajustado y su incertidumbre; el modelo simulado está representado por una línea continua y su incertidumbre está representada por la franja gris. También se encuentra graficado los datos observados.

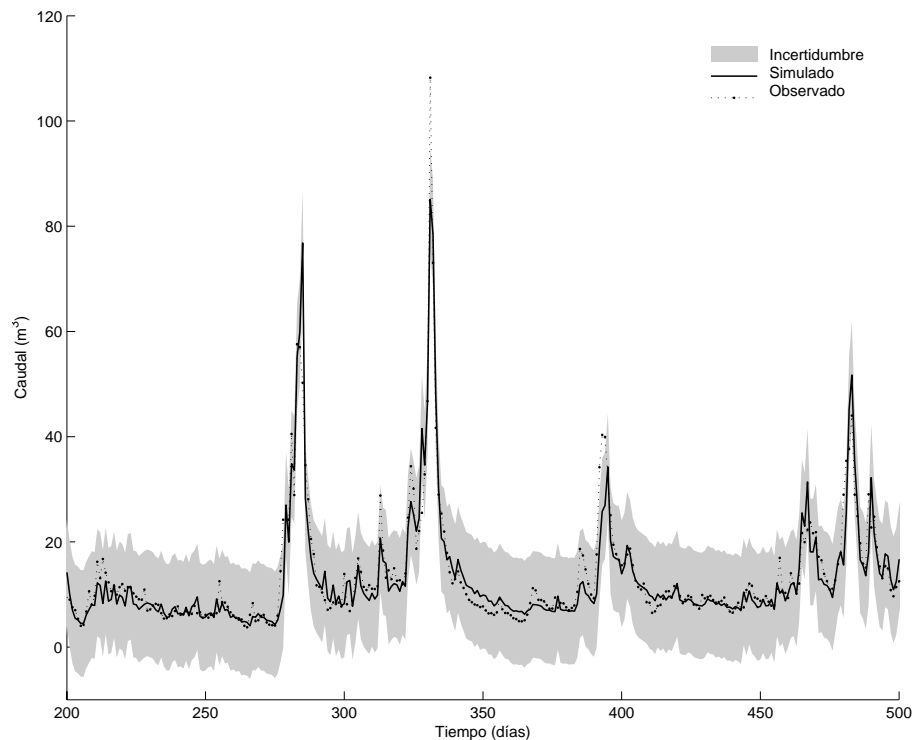


Fig. 4.10: Incertidumbre en la estimación de caudales

4.2.3. Validación

El modelo ha sido validado en un sentido predictivo, como se indicó en los capítulos anteriores; la validación se ha realizado en un segmento de datos obtenidos del ensayo *split sample* el mismo que contiene una muestra de 545 observaciones. El coeficiente de determinación R^2 obtenido es 0.91. Ya que este resultado muestra que el modelo, sin ningún tipo de reestimación continua explicando los datos a un nivel acorde con el obtenido durante el período estimación original de una muestra de 1180 datos observados; se puede concluir que el modelo es *condicionalmente válido* y puede ser utilizado para aplicaciones de pronóstico en tiempo real. En estas aplicaciones, el modelo se transformó en el mejor modelo de tiempo discreto de la forma estocástica estado-espacio integrado dentro de un mecanismo de filtro de Kalman.

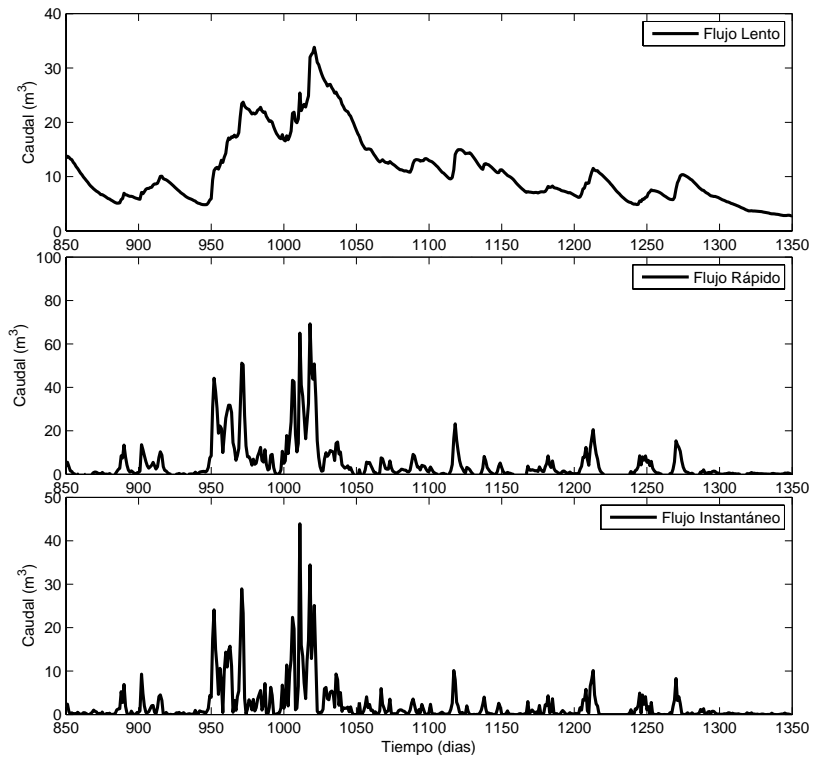


Fig. 4.11: Descomposición de flujo total en subflujos

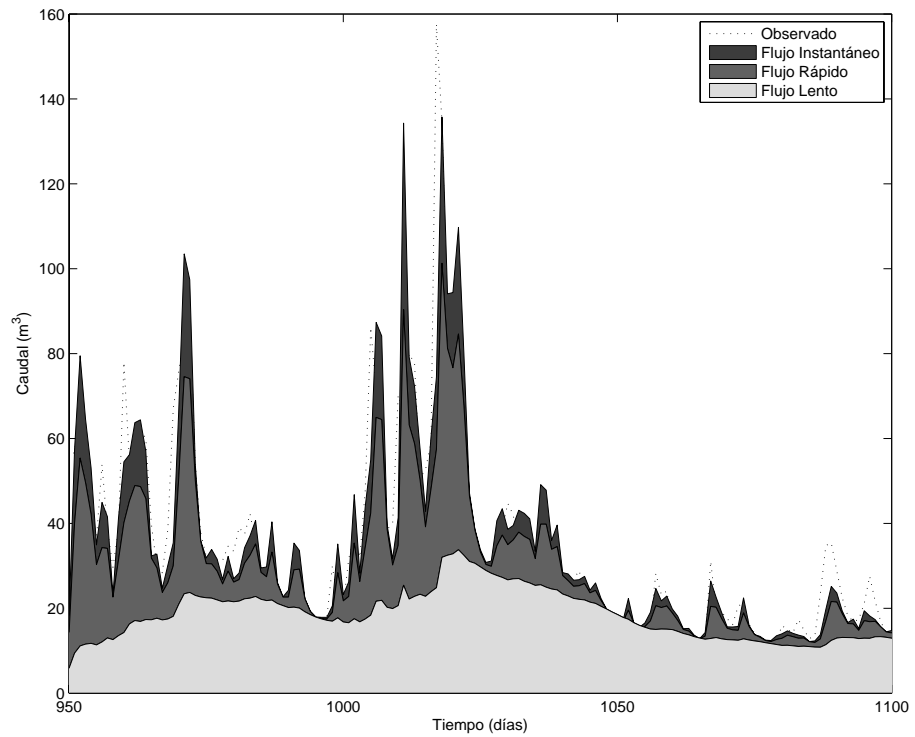


Fig. 4.12: Suma de subflujos

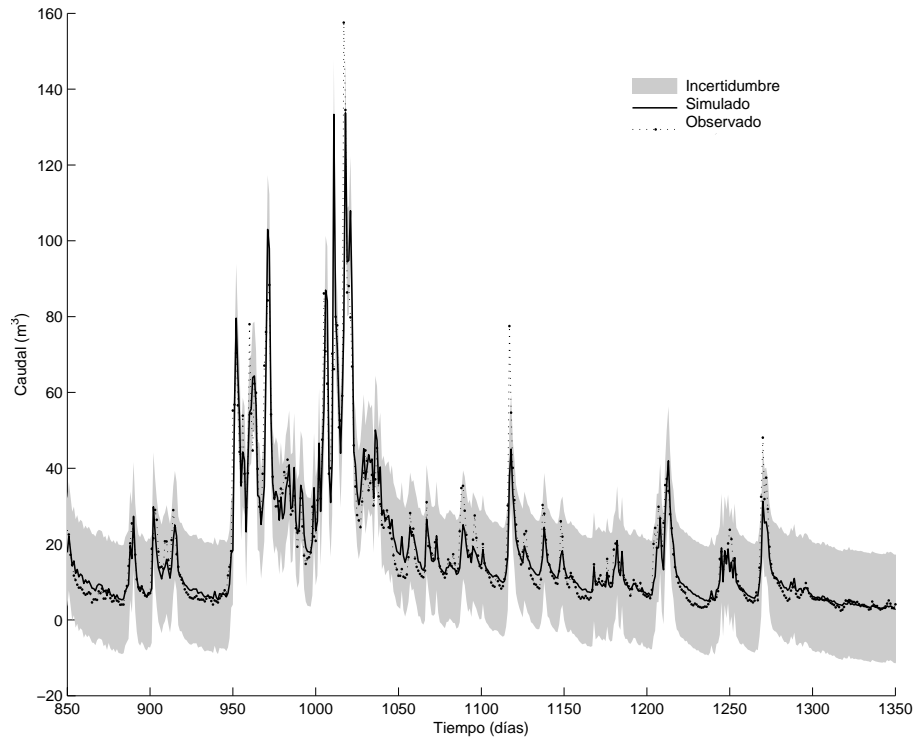


Fig. 4.13: Incertidumbre en la estimación de caudales

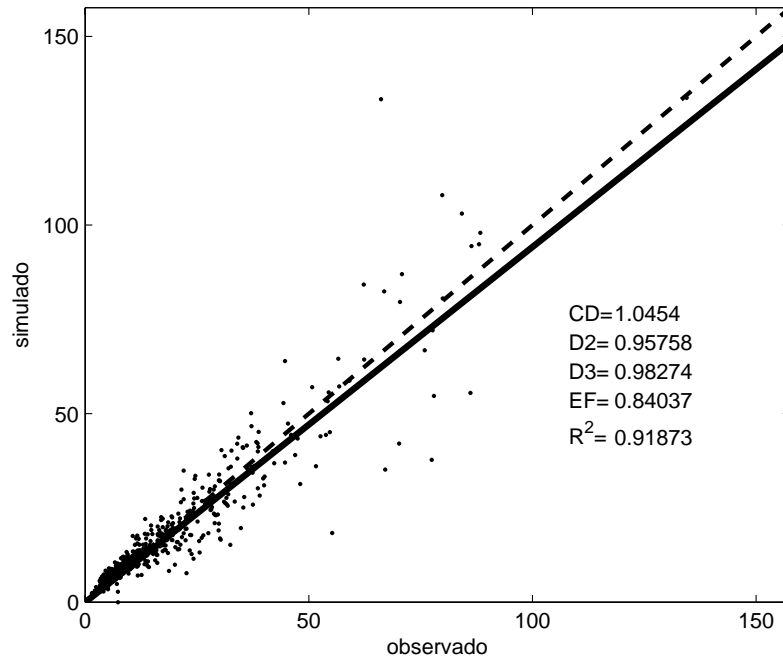


Fig. 4.14: Grafico de dispersión

4.2.4. Pronóstico

Siguiendo el objetivo principal del presente presente trabajo: utilizar el modelo para realizar un pronóstico, como se mencionó en la sección 1.3; el modelo desarrollado anteriormente, aunque sea el modelo que mejor explique los datos de precipitación-caudal, no tiene la capacidad de realizar pronóstico; para que el modelo tenga la capacidad de realizar un pronóstico este debe tener por lo menos un día de retraso en la precipitación que necesita como dato de entrada de lo contrario sería imposible pronosticar la precipitación efectiva. Una alternativa para dotar al modelo de la capacidad de predecir es insertando un tiempo de retraso *virtual* de al menos un día al modelo, como se propuso en la sección 3.1.6. Con el propósito de ilustrar la capacidad predictiva del modelo DBM se han impuesto varios tiempos de retraso, que van desde uno a cinco días. Los parámetros del modelo se han re-estimado para cada día de retraso impuesto (1, 2, 3, 4, 5 días), y se han descompuesto de la misma forma que en el modelo 4.2. Además de pronosticar mediante la función de transferencia, se ha implementado también el sistema estado-espacio dentro del filtro de kalman, de esta manera se pretende analizar el rendimiento de las dos métodos de pronóstico.

Los resultados se presentan resumidos en la tabla 4.5, en la misma se han estimado índices que permiten de alguna manera medir el rendimiento del modelo para producir pronósticos aceptables.

Filtro de Kalman							
Tiempo	RMSE	R2T	CD	D2	D3	EF2	CC
1	9.02	1.00	0.81	0.94	0.98	0.75	0.89
2	12.54	1.00	0.61	0.90	0.96	0.52	0.84
3	14.64	0.99	0.52	0.88	0.95	0.34	0.81
4	15.06	1.00	0.50	0.87	0.95	0.30	0.81
5	13.94	0.99	0.55	0.88	0.95	0.40	0.82

Función de Transferencia							
Tiempo	RMSE	R2T	CD	D2	D3	EF2	CC
1	7.51	1.00	0.99	0.96	0.98	0.83	0.91
2	7.68	0.72	0.96	0.95	0.98	0.82	0.91
3	10.66	1.00	0.67	0.93	0.97	0.65	0.88
4	13.78	1.00	0.54	0.89	0.96	0.42	0.83
5	12.55	0.97	0.57	0.90	0.97	0.52	0.86

Tab. 4.5: Valores de los diferentes índices de eficiencia utilizados para evaluar el rendimiento del pronóstico

Los valores de la tabla anterior se presentan gráficamente en la figura 4.15. En la misma se puede observar como el rendimiento de los modelos en el pronóstico disminuye conforme aumenta el intervalo del pronóstico, resultado que es el esperado, sin embargo se puede ver que el rendimiento del modelo para pronosticar al cuarto día es menor que el de cinco días; esta disminución del rendimiento se puede atribuir a los datos observados de precipitación-caudal, debido a que los eventos de precipitación coinciden de una mejor manera con los caudales observados al quinto día que con los caudales del cuarto día, es decir que puede ocurrir un evento semejante cada cinco días.

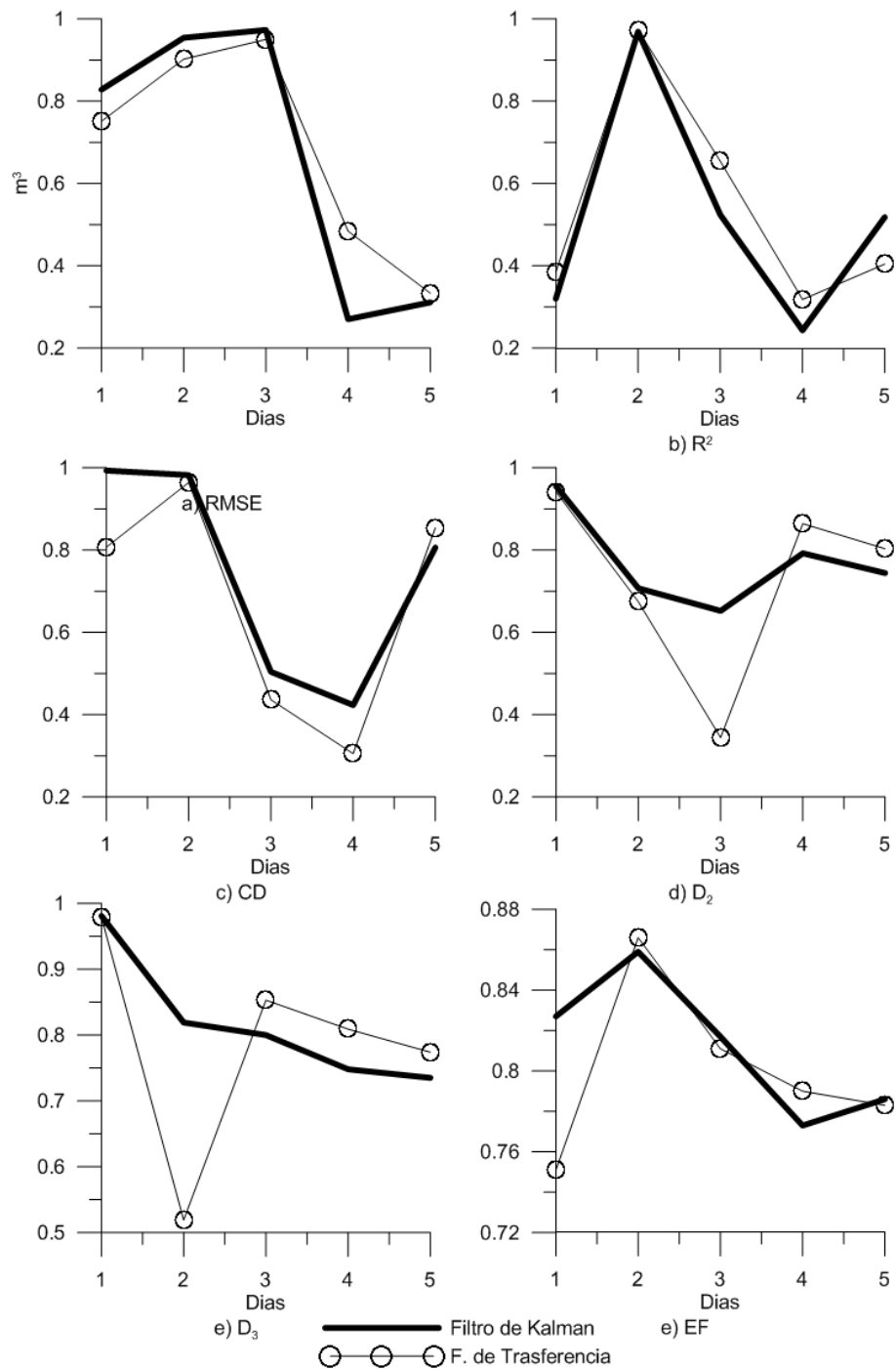


Fig. 4.15: Representación gráfica de la variación temporal de los índices utilizados para evaluar el rendimiento del diferentes horizontes del pronóstico

Finalmente, en la figura 4.16 se muestra el pronóstico realizado con todos los modelos, es decir, el pronóstico para el primer día se ha realizado con el modelo calibrado a un día, para el segundo día se ha utilizado el modelo calibrado a dos días; y de esta manera se lo ha hecho para los siguientes cinco días.

El pronóstico realizado presenta un retraso de un día en el evento pico, también se muestra que el evento pico es pronosticado tanto en el modelo de función de transferencia como en el modelo implementado con el filtro de Kalman.

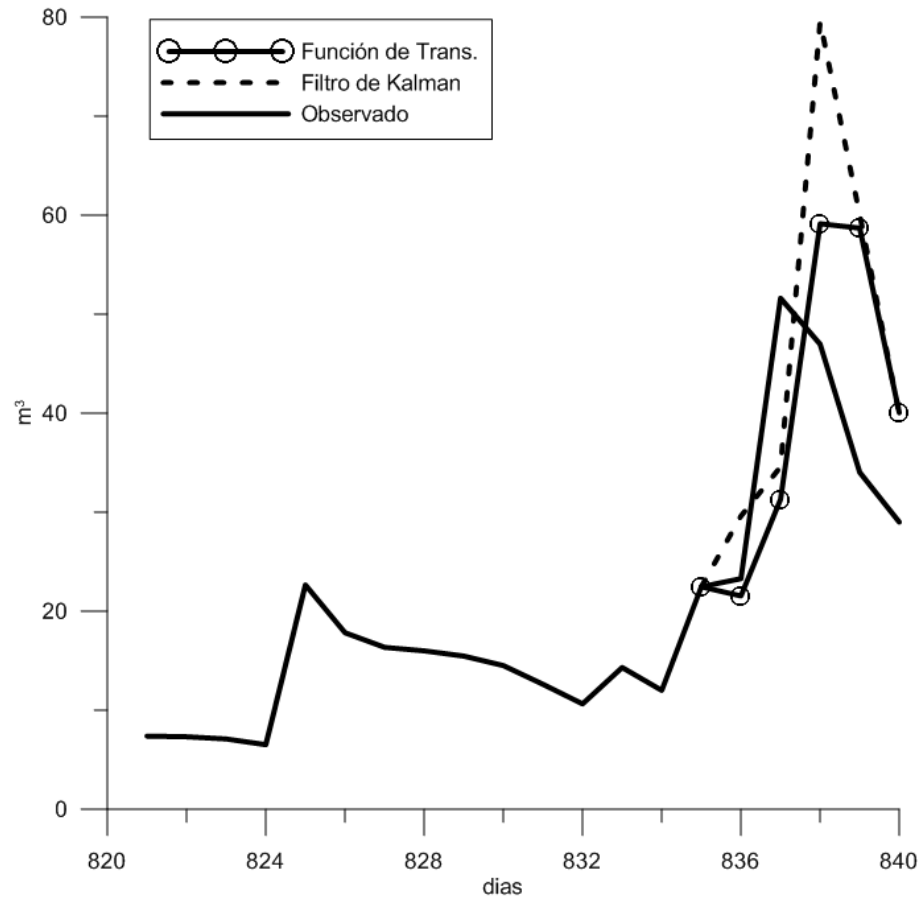


Fig. 4.16: Pronóstico obtenido para un horizonte de cinco días

4.3. Modelo Conceptual Nedbor Asftromning Model NAM

El sistema de modelación NAM aplicado a la cuenca del Paute se ha realizado con el mismo ensayo split-sample utilizado para la modelación DBM de la sección anterior. Los datos para la calibración son del período de [1 de julio de 1981 - 3 de octubre de 1983], y para la validación el período de [4 de octubre de 1983 - 27 de marzo de 1985].

El algoritmo utilizado en el presente trabajo corresponde al NAM ([10]) el mismo que forma parte del software MIKE 11. El sistema NAM se aplicó principalmente debido a su capacidad para simular los procesos físicos, además, se trata de un software de modelado

completo y efectivo el mismo que dispone de módulos que permite la flexibilidad para futuras investigaciones. El MIKE 11 se aplicó debido a su disponibilidad en el Departamento de Investigación de la Universidad de Cuenca. MIKE11 NAM es un paquete de software profesional para ingeniería, desarrollado por el Instituto de Hidráulica de Dinamarca (DHI). Esta herramienta de modelación unidimensional desarrollada desde 1972 ha sido aceptada en todo el mundo para la gestión y manejo de los recursos hídricos; es utilizado para la simulación de: caudales, calidad del agua y transporte de sedimentos en los estuarios, ríos, sistemas de riego, canales y otros cuerpos de agua.

Para encontrar el modelo, el MIKE 11 NAM requiere de los siguientes datos: el área de la cuenca, las series de tiempo meteorológicas de precipitación y evaporación potencial, y adicionalmente las serie de tiempo de los caudales que son necesarios para realizar la calibración del modelo. También necesita de valores iniciales: contenido inicial de agua reservorio superficial y reservorio de la zona raíz, junto con los valores iniciales de flujo superficial, subsuperficial y el flujo de base; el valor para estas condiciones iniciales fueron de cero excepto para el flujo base que se adoptó un valor de $0.76m^3/s$ que corresponde al menor caudal observado el mismo que pudo haber ocurrido en un período de sequía. Un resumen de las condiciones iniciales se presenta en la tabla 4.6.

Condiciones Iniciales		
Condiciones Iniciales	Valor	Unidad
U_ini	0	mm
L_ini	0	mm
OF_ini	0	m ³ /s
IF_ini	0	m ³ /s
BF_ini	0.76	m ³ /s

Tab. 4.6: Condiciones iniciales

4.3.1. Calibración

Los datos de entrada para el sistema de modelación NAM consisten en series de tiempo diarias de precipitación, descarga y evapotranspiración potencial. Los parámetros fueron encontrados por medio de calibración a través de las series de tiempo diarias de descarga a la salida de la cuenca.

La prueba SS es suficiente para obtener una calibración y evaluación confiable para modelos concentrados ([31]), por lo tanto, para la calibración y evaluación se llevó a cabo una prueba SS. La calibración se realizó mediante dos pasos principales: una calibración automática y luego una calibración manual. La calibración automática fue usada para acelerar el proceso de calibración y para restringir los parámetros más importantes a un cierto rango de valores aceptables. La rutina de calibración está basada en una estrategia de calibración multi-objetivo en la cual cuatro diferentes objetivos pueden ser optimizados simultáneamente. La rutina utiliza el error total del volumen, la raíz media cuadrada del



error total (para definir la forma del hidrograma), la (RMSE) de los eventos picos y la RMSE de los eventos valle.

Después de haber realizado el calibrado automático se realizó una calibración manual, se realizaron ensayos prueba-error hasta obtener resultados satisfactorios. La estrategia de calibración manual paso a paso considera las diferentes descripciones del proceso precipitación-escorrentía para calibrar los parámetros que más afectan a los procesos que van a ser descritos.

El primer paso en la calibración paso a paso fue ajustar el balance total del agua, ajustando el balance entre la evapotranspiración y la diferencia entre la precipitación neta y la escorrentía. La evapotranspiración podría incrementarse cuando incrementamos el contenido de agua máximo del reservorio superficial (U_{max}) y el reservorio de la zona de raíces (L_{max}) y viceversa. Usualmente este se ajusta implícitamente pero el mismo resultado podría ser llevado a cabo explícitamente controlando el balance de agua de salida. Esto se puede generar después de cada corrida del modelo de forma tal que se evalúe el progreso del balance.

El segundo paso fue ajustar a la forma del hidrograma los eventos de escorrentía picos, el ajuste del volumen pico es controlado por el coeficiente de escorrentía superficial ($CQOF$) y la constante de tiempo de tránsito de escorrentía ($CK_{1,2}$) respectivamente. El tercer paso es ajustar el flujo base, esto se puede lograr modificando la constante de tiempo del flujo base ($CKBF$). Si la recesión del flujo base cambia a un valor muy bajo después de cierto tiempo, entonces se puede agregar un reservorio de agua subterránea más bajo.

Un total de 9 parámetros fueron calibrados, sin considerar el riego y módulos de nieve. La Tabla 3.2.3 muestra un resumen de los principales parámetros del modelo y su impacto en los diferentes componentes del ciclo hidrológico. La evaluación cuantitativa de los resultados del modelo se realizó utilizando el coeficiente de Nash y el coeficiente de eficiencia (EF).

El primer paso de la calibración del modelo (calibración automática) utiliza el algoritmo SCE, el mismo que está incluido en el paquete MIKE 11 NAM, el algoritmo de optimización de parámetros SCE necesita del espacio de posibles soluciones de los parámetros, este espacio solución recomendado por DHI se encuentra en la tabla 3.2.3.

Los parámetros obtenidos de la calibración automática se presentan en la tabla 4.7, el rendimiento resultante de los mismos medido con el coeficiente de Nash se encuentra en la parte inferior de la misma tabla.

Una vez realizado la calibración automática se llevó a cabo la calibración manual, para ello se siguió el protocolo descrito en capítulos anteriores; los resultados se encuentra



también en la tabla 4.7.

Parámetro	Unidad	Espacio de Parámetros	
		Automática	Automática-manual
U_max	mm	10.950	10.000
L_max	mm	51.590	50.000
CQOF		0.645	0.650
CKIF	hrs.	602.950	500.000
CK_12	hrs.	35.080	40.400
TOF		0.070	0.07
TIF		0.106	0.11
TG		0.028	0.030
CKBF	hrs.	1133.000	1150.000
R²		0.366	0.411

Tab. 4.7: Parámetros del NAM encontrados mediante calibración automática y automática-manual

Los parámetros encontrados describen el sistema de la siguiente manera: la escorrentía total es la suma de tres tipos de caudales con diferentes características: escorrentía superficial, escorrentía subsuperficial y caudal base. En la figura 4.17 se presentan los hidrogramas de los tres diferentes caudales que aportan a la escorrentía total; se puede observar en los hidrogramas como estos caudales cambian con el tiempo, es decir, los caudales presentan curvas de recesión más largas cuanto más lento es el flujo, se puede ver claramente como en el hidrograma de caudales de flujo base como el caudal disminuye lentamente en el transcurso del tiempo; lo contrario ocurre en flujo superficial, donde el hidrograma presenta curvas de recesión muy cortas que indican que el flujo es muy rápido; en el término medio se encuentra el flujo subsuperficial, en el cual la forma del hidrograma es más suavizada.

El modelo encontrado no parece tener un buen rendimiento ($R^2 = 0.411$), esto puede deberse a una calibración deficiente o tal vez el modelo no tiene la capacidad de representar conceptualmente el comportamiento físico apegado a la realidad de la cuenca de estudio, es importante recordar que la cuenca de estudio está conformada en un 80 % por páramos, y no existen estudios suficientes que describan el ciclo hidrológico en este tipo de cuencas.

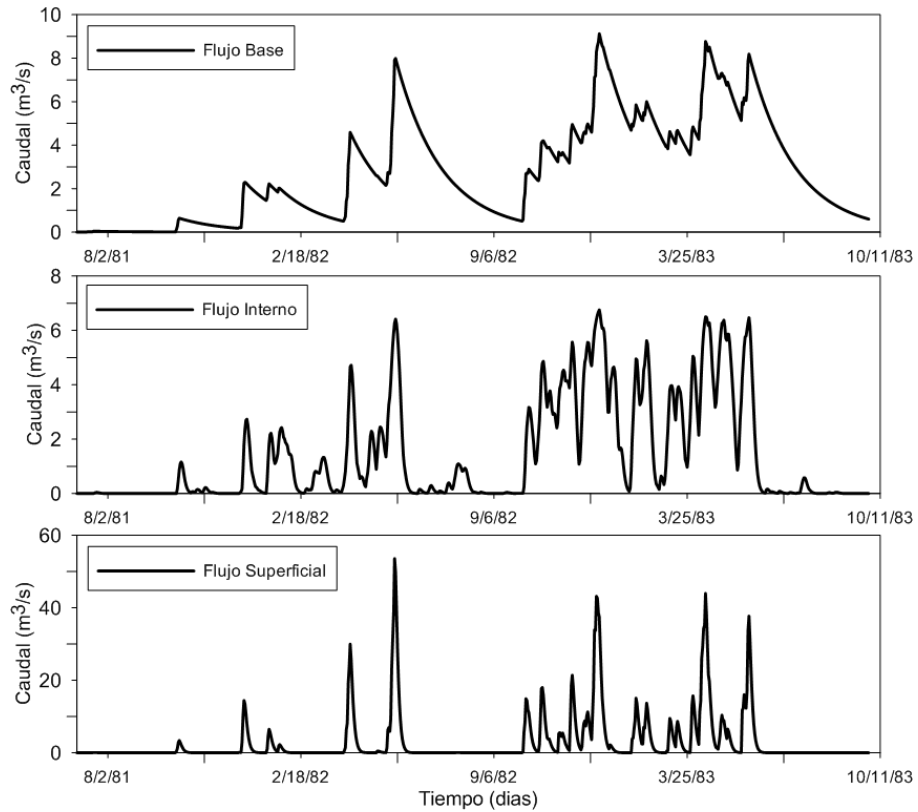


Fig. 4.17: Descomposición de flujo total en subflujos

La composición del caudal total es suma de los distintos flujos, la contribución de cada uno de estos se muestra en la gráfica 4.18, donde se vé claramente como los distintos caudales se van acumulando hasta completar el caudal total, además, como influyen en el tiempo los distintos caudales.

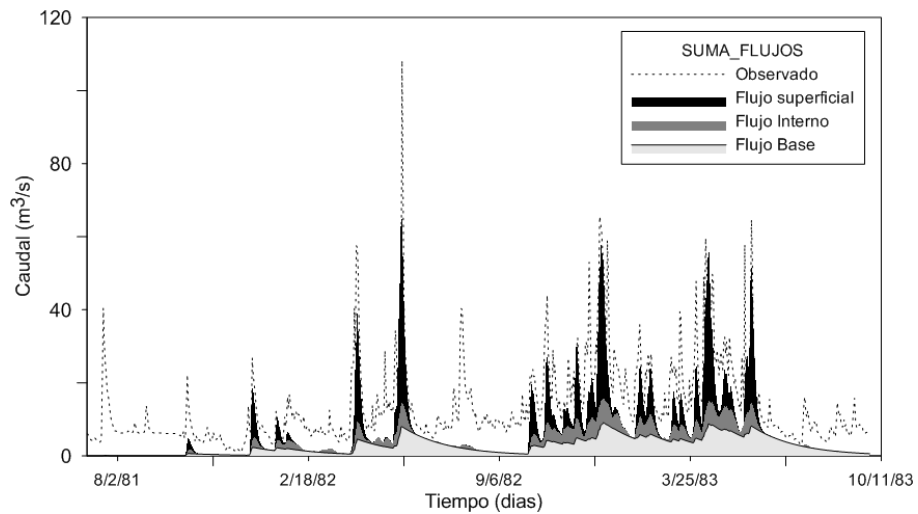


Fig. 4.18: Suma de subflujos



El modelo encontrado anteriormente presenta un bajo rendimiento razón por la cual se ha cambiado el rango de posibles soluciones, teniendo en cuenta que el cambio puede ocasionar que el modelo pierda la capacidad de describir el sistema precipitación-caudal, condicionando al mismo probablemente a un modelo caja-negra.

Los parámetros encontrados en el rango fuera del recomendado por DHI se muestran en la tabla 4.8, y son resultado de una calibración automática-manual.

Parámetros óptimos para el modelo NAM			
Reservorio	Parámetro	Valor	Unidad
Reservorio Superficial y Zona de raíces	U_max	0.791	mm
	L_max	0.526	mm
	CQOF	0.276	
	CKIF	23.704	hrs.
	CK_12	46.927	hrs.
	TOF	0.422	
	TIF	0.514	
Reservorio de agua Subterránea	TG	0.43	
	CKBF	1972.232	hrs.

Tab. 4.8: Parámetros encontrados del modelo NAM

Los parámetros encontrados indican que la mayor parte de la precipitación se transforma en flujo base, lo que implicaría que el suelo que conforma la cuenca no tiene capacidad de almacenar agua y que la fuente se abastece en gran parte solo de acuíferos. Los suelos de los páramos debido a su baja densidad aparente y su estructura abierta y porosa, tienen una conductividad hidráulica alta y una capacidad de retención de agua muy alta [7].

Los parámetros encontrados explican de manera adecuada los datos, sin embargo no describen la realidad física del sistema precipitación-caudal para esta cuenca, sus valores están fuera de la realidad y no caracterizan a los elementos conceptuales que describen la realidad física de la cuenca de estudio.

Si se intenta buscar una interpretación física del modelo basados en los parámetros encontrados se podría decir que el caudal total del río en punto de medición está compuesto de los tres tipos de flujo que aportan de la siguiente manera, el mayor aporte a la escorrentía total se debe al flujo base, cuyo tiempo de recesión es de aproximadamente de 82 días; el flujo intermedio y el flujo superficial aporta de forma reducida al caudal total con tiempos de recesión de aproximadamente de 2 y 1 días respectivamente.

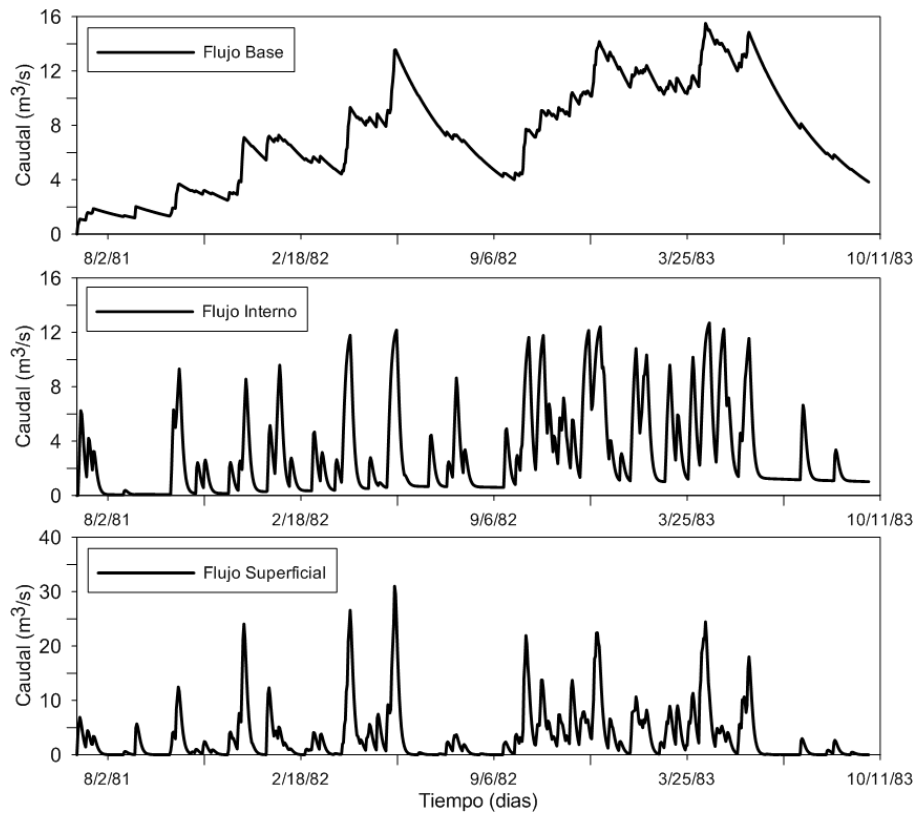


Fig. 4.19: Descomposición de flujo total en subflujos

La suma de los distintos flujos resulta en el caudal total, la contribución de cada uno de estos se muestra en la gráfica 4.20, donde se ve claramente como los distintos caudales se van acumulando hasta completar el caudal total, además, como influyen en el tiempo.

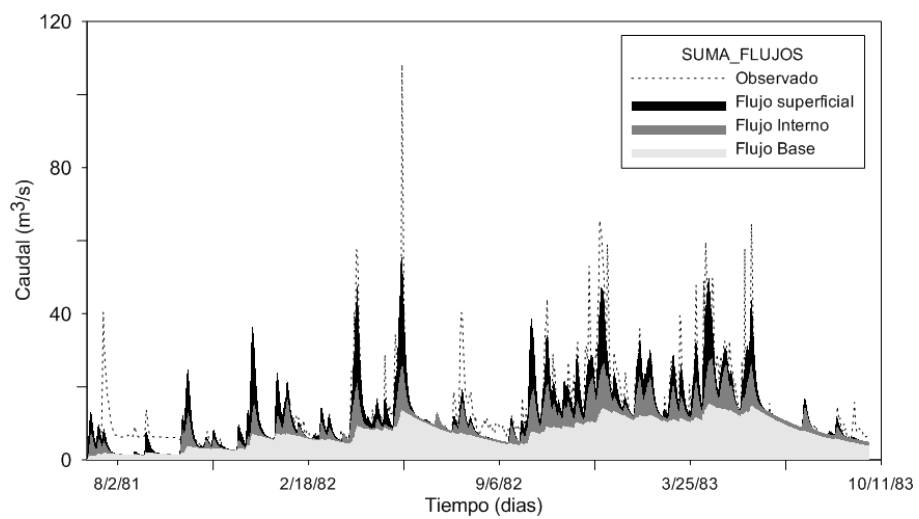


Fig. 4.20: Suma de subflujos

Finalmente, para realizar una evaluación visual de rendimiento general del modelo

se presenta la gráfica 4.21, que permite observar como el caudal simulado se aproxima al caudal observado; un análisis más profundo de estos resultados se presenta en la sección siguiente, en donde se comparan los modelos NAM y DBM.

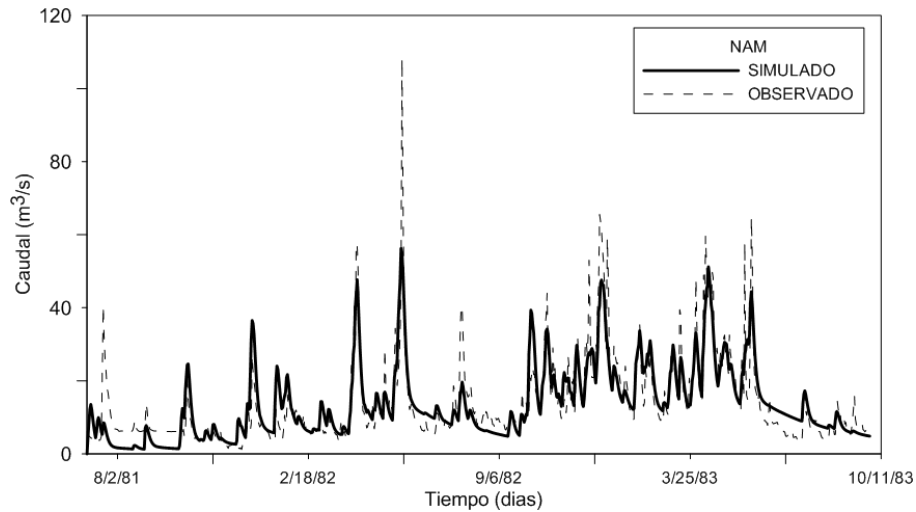


Fig. 4.21: Incertidumbre en la estimación de caudales

4.3.2. Validación

Siguiendo la misma metodología empleada para la validación del modelo DBM, se ha validado en un sentido predictivo el NAM con el modelo cuyos parámetros están fuera de una representación física, como se ha indicado en la sección anterior; la validación se ha realizado en un segmento de datos obtenido del ensayo *Split-sample* el mismo que contiene una muestra de 545 observaciones. El coeficiente de determinación R^2 obtenido para el período de validación es 0.66; el índice muestra que el rendimiento obtenido se reduce en 0.05 con respecto a al periodo de calibración. La gráfica 4.22 muestra el hidrograma de caudales simulados y observados, se puede ver claramente como el modelo no predice los caudales picos de una manera adecuada, los mismos que presentan subestimación. Un análisis más profundo de estos resultados se presenta en la sección correspondiente a comparación de los modelos.

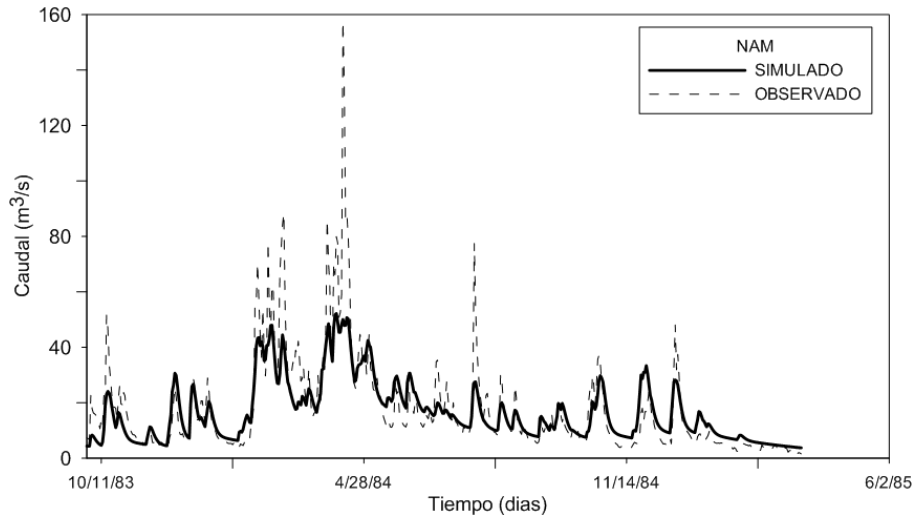


Fig. 4.22: Incertidumbre en la estimación de caudales

Los caudales simulados son el resultado de la suma de tres tipos de subflujos como se explicó en secciones anteriores, las gráficas 4.23 muestran los hidrogramas de los subflujos encontrados; en los mismos se puede observar que su comportamiento es muy parecido al periodo de calibración, siguiendo el mismo patrón de cambio en el tiempo.

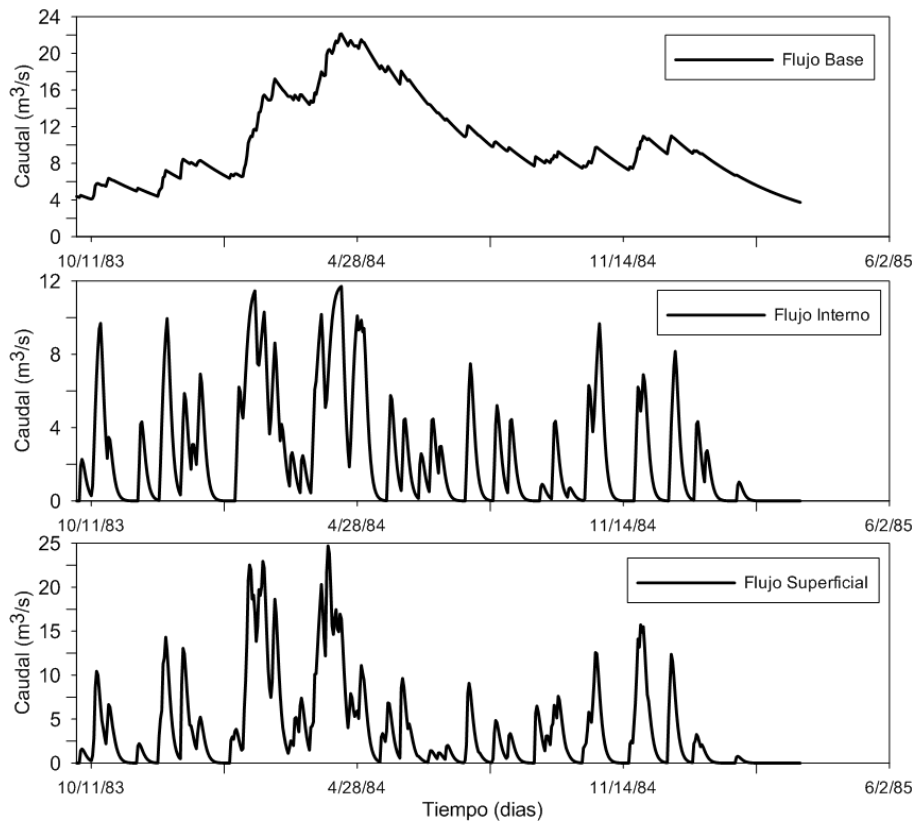


Fig. 4.23: Descomposición de flujo total en subflujos

Aunque el modelo no presenta un buen rendimiento para el período de validación, su composición de caudales es muy parecida al período de calibración como se puede observar en la gráfica 4.24.

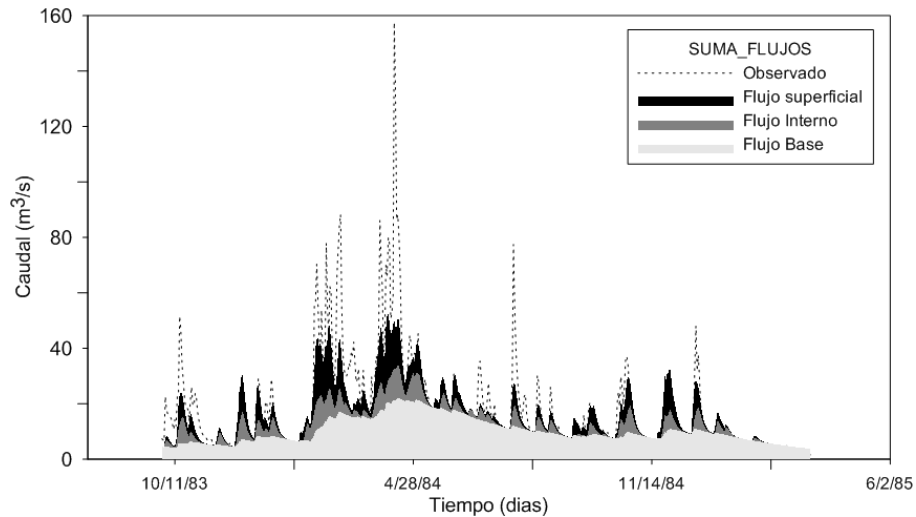


Fig. 4.24: Suma de subflujos

4.4. Análisis comparativo de los modelos DBM y NAM

La modelación realizada a la cuenca del Paute comprendió dos sistemas de modelación, el DBM y el NAM; con el sistema de modelación DBM se ha realizado una variante la misma que es analizada en esta sección. Una de las variantes posibles en el DBM es el filtrado final de los datos utilizando el filtro de Kalman. En el caso del modelo NAM se tomó en cuenta el modelo que tiene un mejor rendimiento, dejando de lado el modelo cuyos parámetros tendrían un significado físico.

En los sistemas de modelación conceptual se intenta describir el ciclo del agua en la fase terrestre por medio de reservorios, los mismos que necesitan de una entrada y una salida de forma tal que exista un balance agua; la cantidad total de agua que es descargada por la fuente debe ser igual a la precipitación neta menos la evapotranspiración. El volumen acumulado observado en un periodo determinado debe ser igual al volumen acumulado en el mismo período del modelo simulado; en este sentido, con el modelo DBM se obtiene mejores resultados en el balance de agua con respecto al NAM tanto en los periodos de calibración como en los de validación como lo demuestra la gráfica 4.25 y la tabla 4.9. En ambos modelos el caudal acumulado se reduce en el periodo de validación pero es mucho más notable en el NAM cuya diferencia es alta, lo cual quiere decir que no



existe un balance de agua correcto. Este resultado afecta también al caudal medio diario (tabla 4.9 y figura 4.26; aunque en el periodo de calibración esta diferencia es insignificante en ambos modelos; para la validación los resultados muestran que el modelo DBM tiene un mejor rendimiento, lo cual condiciona el uso del NAM para realizar pronósticos a corto plazo.

Modelo	Calibración (m ³ /s)		Validación (m ³ /s)	
	Caudal medio diario	Caudal Acumulado	Caudal medio diario	Caudal Acumulado
Obs	13.85	11355.80	17.80	9717.10
DBM TF	13.57	11126.64	17.77	9704.89
DBM KF	13.84	11346.93	17.43	9514.09
NAM	13.84	11352.31	16.63	9079.41

Tab. 4.9: Comparación del balance total y el caudal medio diario estimados de la cuenca del Paute para los períodos de calibración y validación.

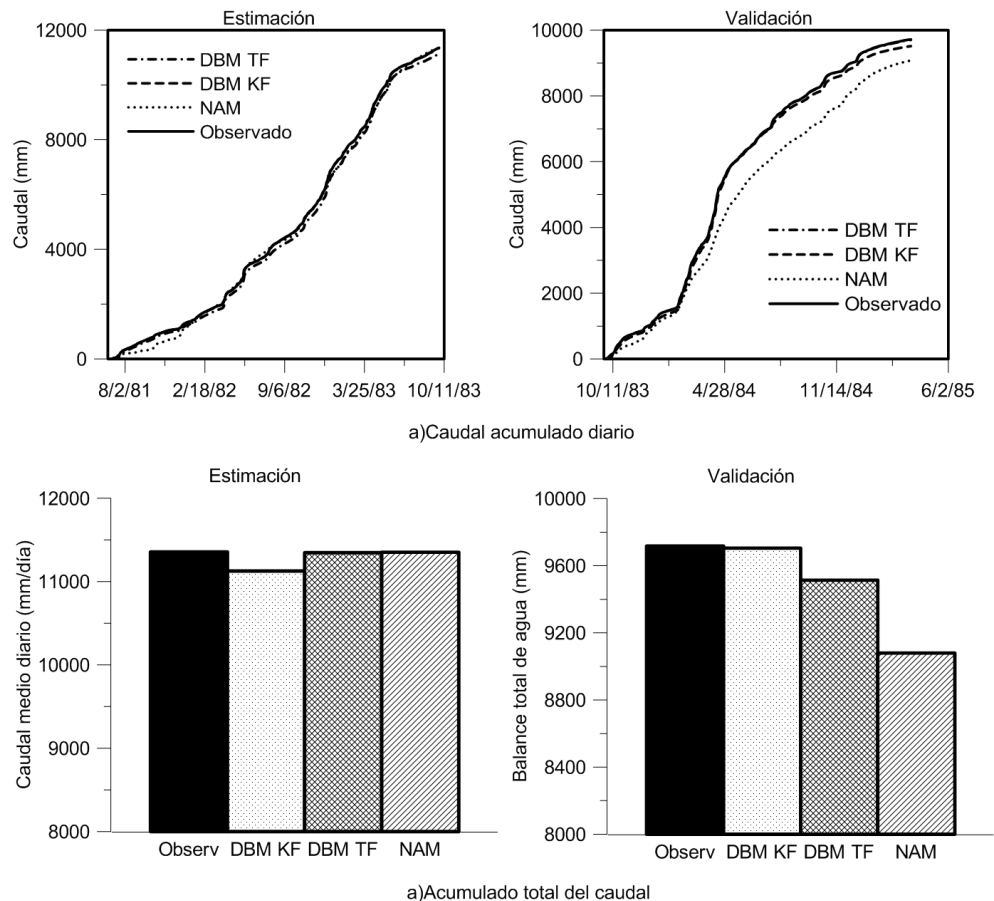


Fig. 4.25: Caudal Acumulado

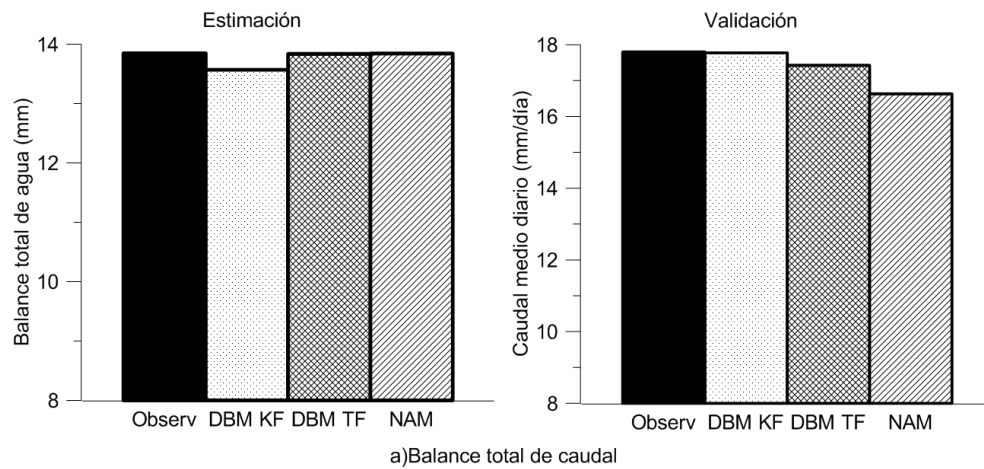


Fig. 4.26: Caudal medio diario

El análisis anterior nos da una idea de cómo se están comportando los modelos, sin embargo, es un resultado global del sistema precipitación-caudal que no permite un análisis del comportamiento del modelo con respecto al tiempo. Con este fin, realizar un análisis visual de las series de tiempo de los caudales simulados con los caudales observados permite ver de manera más clara como el modelo simula los caudales picos y los caudales valle en el tiempo.

En las figuras 4.27 y 4.28 se presentan los hidrogramas de las series de tiempo simuladas y observadas para los períodos de calibración y validación respectivamente. Teniendo en cuenta que en el modelo NAM los parámetros no representan entidades físicas: en la parte a) de la figura 4.27 ambos modelos parecen ajustarse favorablemente a los datos observados con una pequeña ventaja a favor del DBM en sus dos variantes; pero en el período de validación el modelo NAM no se ajusta de manera adecuada a los datos observados especialmente en los caudales picos, que son de mucha importancia en la prevención de inundaciones y eventos extremos, lo cual confirma con que el modelo no trabaja de manera adecuada y condiciona su uso para el pronóstico. Los caudales picos son sub-estimados por el NAM, esto podría explicar el por que del desbalance del sistema precipitación-caudal.

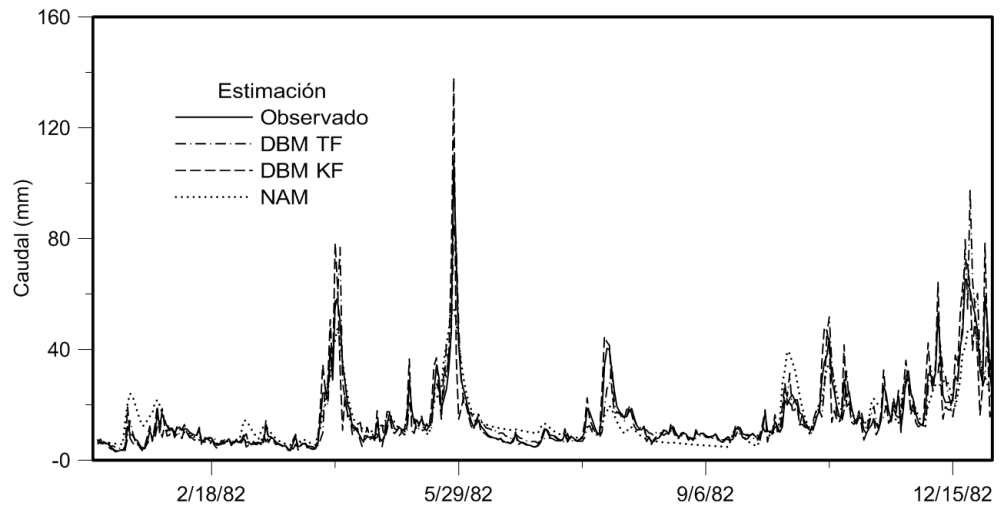


Fig. 4.27: Comparació estimación

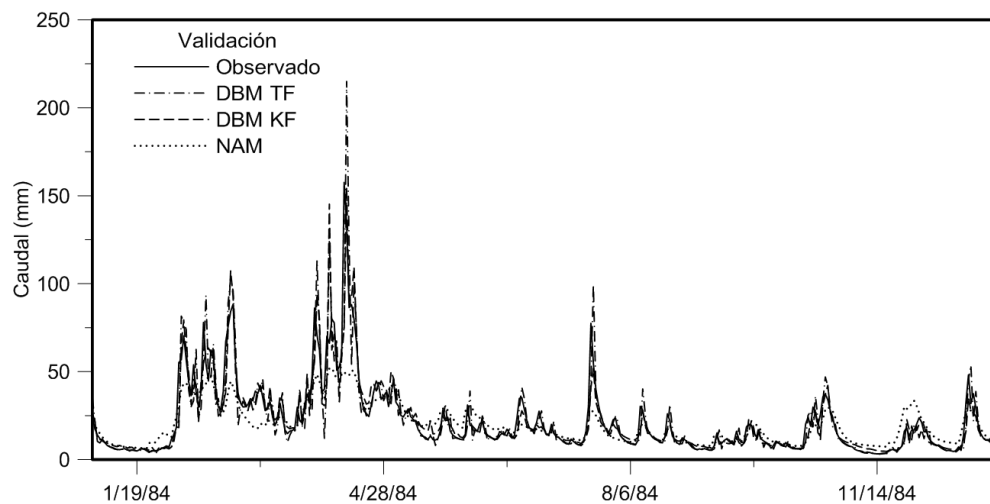


Fig. 4.28: Comparación de los modelos DBM y NAM para el periodo de validación

El examen visual realizado anteriormente no permite medir como los modelos están simulando los caudales, esto puede solucionarse mediante el uso de algunos índices que permiten calificar el rendimiento del modelo; los índices utilizados se describieron con más detalle en la sección 3.3 los mismos que son: coeficiente de variación (CV), coeficiente de Nash (R^2), coeficiente de determinación (CD), coeficiente de correlación (CC), y los índices de eficiencia de grado dos y tres (EF_2 y EF_3 respectivamente).

El valor de los diferentes índices calculados para los modelos son presentados en forma resumida en la tabla 4.10 y de forma gráfica en la figura 4.29. Los índices indican un mejor rendimiento del modelo híbrido DBM con respecto al modelo conceptual NAM, tanto en los periodos de calibración como en los periodos de validación; ambos reducen su capacidad de explicar los datos para el periodo de validación, esto es muy normal debido a que los



parámetros encontrados son óptimos para el período de calibración. Existe una evidente sub-estimación de caudales en el NAM (4.29 parte c)) característica que es más notable en la validación gracias al parámetro CD.

Índice	Calibración			Validación		
	DBM TF	DBM KF	NAM	DBM TF	DBM KF	NAM
CV	0.37	0.36	0.45	0.51	0.42	0.59
R2	0.81	0.81	0.71	0.75	0.83	0.66
CD	0.92	0.76	1.32	0.81	0.99	2.68
CC	0.91	0.93	0.84	0.89	0.92	0.85
EF2	0.95	0.96	0.91	0.94	0.96	0.86
EF3	0.98	0.99	0.97	0.98	0.98	0.90

Tab. 4.10: Índices para medir el rendimiento del modelo

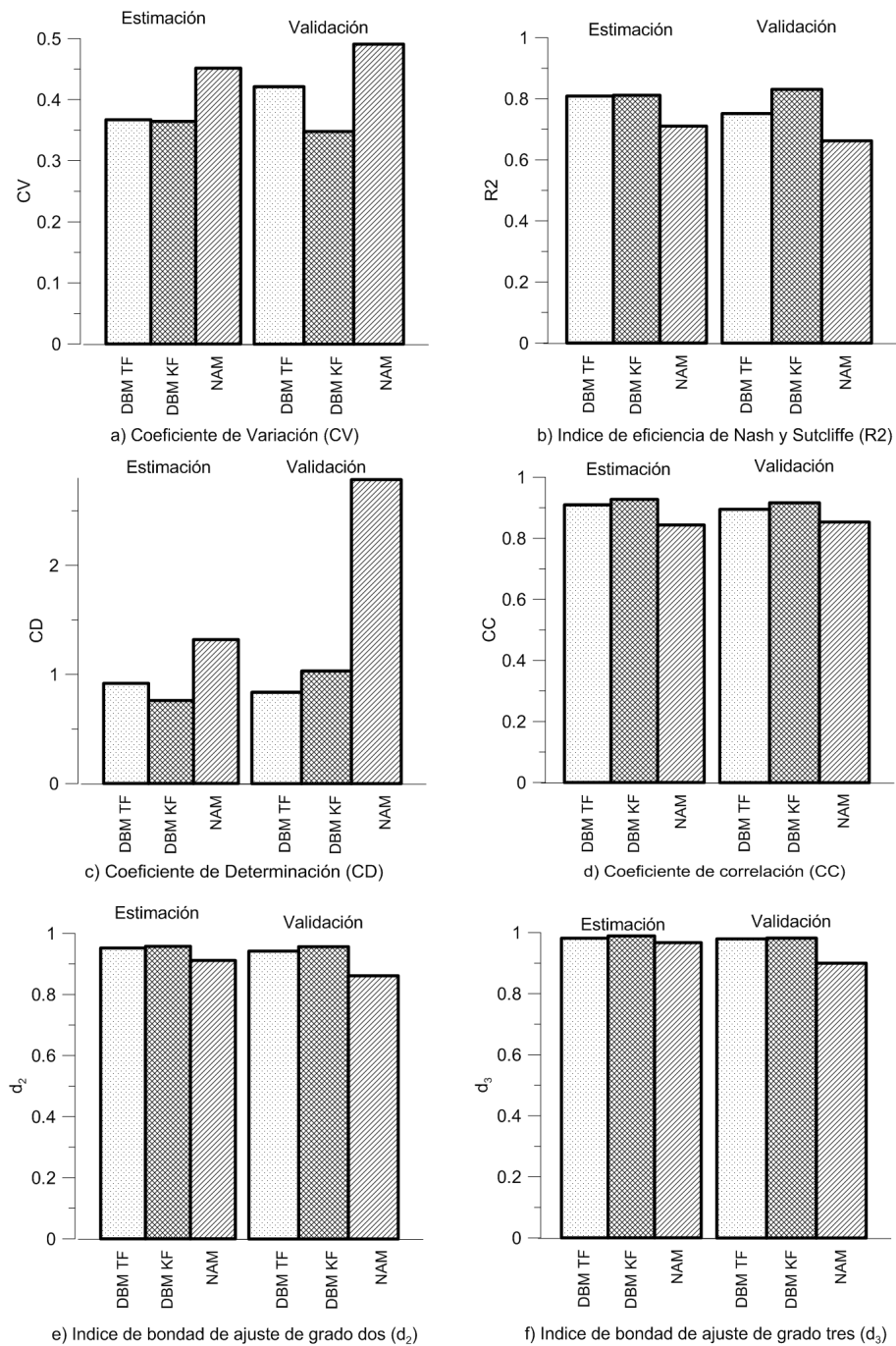


Fig. 4.29: Índices para medir el rendimiento de los modelos

Por último se presenta los scatter-plots de los caudales observados vs. los simulado: en la parte a) de la gráfica 4.30 se muestra los datos simulados con el modelo DBM mediante funciones de transferencia (TF), en el mismo se puede analizar en forma visual como los valores de los caudales picos están sobreestimados, los caudales medios tienen mayor incertidumbre y los caudales bajos parecen ajustarse de mejor manera. Para



el periodo de validación ocurre algo similar ajustándose los picos de mejor manera. En la parte b) de la misma gráfica se presentan los resultados obtenidos con el modelo DBM utilizando el filtro de Kalman, el modelo parece comportarse igual manera que sin el filtro kalman en el periodo de calibración; igualmente los caudales picos parecen estar sobreestimados, los caudales medios parecerían también tener la misma incertidumbre, y los caudales valle parecen ajustarse de mejor manera; para el periodo de validación ocurre lo mismo sin mayor cambio aparente. Por último, el NAM es presentado en la parte c) de la gráfica 4.30 en este se puede diferenciar la existencia de una mayor incertidumbre con respecto al DBM en la etapa de estimación y también en la etapa de validación. El modelo no representa de una manera adecuada los flujos picos y los flujos valles pues los mismos están subestimados, lo cual confirma el bajo rendimiento de los índices y la reducida capacidad de pronosticar con respecto al modelo DBM.

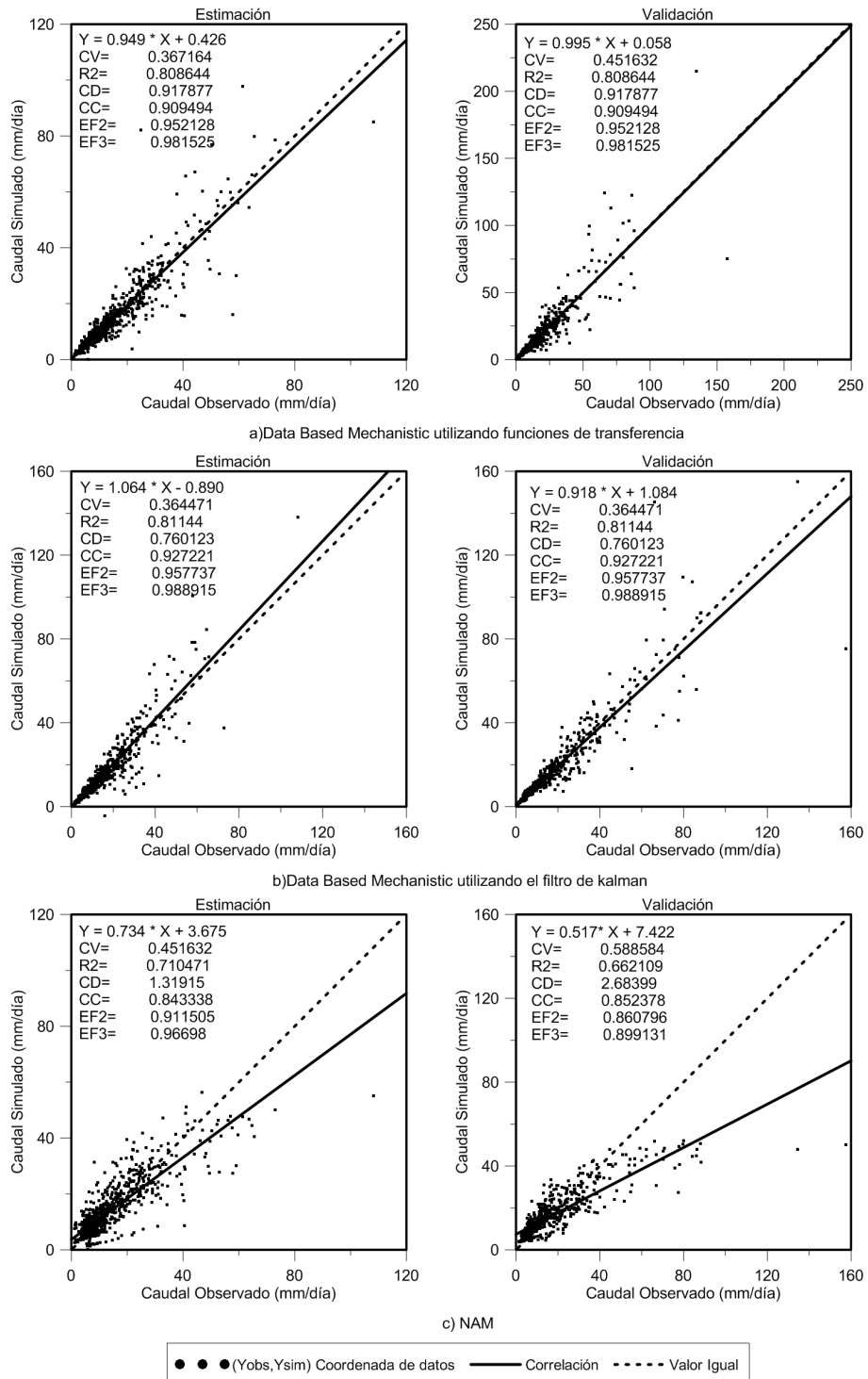


Fig. 4.30: Scatter-plot

CONCLUSIONES

Los sistemas de modelación DBM y NAM han sido sometidos a un riguroso programa de ensayos con los datos de la cuenca del Paute. Por un lado el NAM es un típico representante de los modelos conceptuales y por otro lado el DBM es un modelo que combina características de un modelo conceptual con un modelo black-box.

Los resultados obtenidos con el DBM demuestran una gran ventaja con respecto a los modelos conceptuales, el mismo por sus características es capaz adaptarse cualquier tipo de cuencas al permitir una interpretación física después de haber realizado la simulación, lo cual no ocurre con el modelo NAM, que en esta caso no se pudo ajustar a los datos de esta cuenca que tiene características particulares.

Sin embargo, a pesar de que los resultados obtenidos están de acuerdo con el objetivo principal del presente trabajo se debe considerar que los mismos están influenciados por condiciones exógenas al sistema de modelación como pueden ser: la disponibilidad de datos, y la calidad de estos, así como la evaluación subjetiva realizada por el usuario (por ejemplo, la interpretación de la información disponible para determinar los parámetros del modelo). Por lo tanto los resultados obtenidos no son solo una función del sistema de modelación, sino también de los usuarios y muchos otros factores; y por tanto para llegar a una conclusión general se requiere de muchas más pruebas de validación.

Con esta observación se pueden llegar a conclusiones específicas derivadas de este estudio:

- En primer lugar, debido al bajo rendimiento y a la reducida capacidad de explicar los componentes físicos del sistema precipitación-caudal, el sistema de modelación NAM no es adecuado para este tipo de cuencas; por las características de la misma es necesario perder la capacidad de alguna interpretación física del modelo para mejorar el rendimiento del mismo, esto consecuentemente involucra clasificar al modelo NAM dentro del grupo de modelos del tipo caja-negra para este tipo de cuencas, caracterizadas por estar conformadas predominantemente por páramos.
- En segundo lugar, el DBM tiene en general un rendimiento muy bueno con los registros disponibles, se ajusta con cierta precisión al caudal registrado en el periodo de calibración sin reducir severamente el rendimiento durante un periodo independiente



utilizado para la validación. Por lo tanto de esta prueba sugiere que si se dispone de un periodo adecuado de registros de unos pocos años, el sistema de modelación DBM podría ser utilizado como una herramienta confiable para llenar los vacíos en los registros.

- En tercer lugar, el modelo DBM podría considerarse útil para pronosticar eventos extremos, pues los resultados demuestran que el modelo es sensible a los mismos.
- En cuarto lugar, debido a la utilización de algoritmos avanzados para la calibración del modelo, el modelo DBM sería mucho más fácil calibrar ya que existen menos parámetros y se requiere menos experiencia del modelador, contrastando con el sistema de modelación NAM en el cual los parámetros pueden representar entidades físicas promedio de la cuenca.
- Finalmente, la estructura del DBM y sus herramientas de modelación permiten concluir que el modelo tiene mejor capacidad para realizar pronósticos debido a su capacidad de actualización continua de información observada en tiempo real, lo que le permite obtener resultados más ajustados a los datos medidos, sumado a esto la utilización de algoritmos que mejoren el rendimiento como el filtro de Kalman.

REFERENCIAS

- [1] M.B. Abbott, J.C. Bathurst, J.A. Cunge, P.E. O'Connell, and J. Rasmussen, *An introduction to the european hydrological system systeme hydrologique europeen, she, 1: History and philosophy of a physically-based, distributed modelling system*, Journal of Hydrology **87** (1986), 45–59.
- [2] J.G. Arnold, R. Srinivasan, R. S. Muttiah, and J. R. Williams, *Large area hydrologic modeling and assessment , part ii: Model application*, Journal of the American Water Association American Water Resources Association **34(1)** (1998), 91–101.
- [3] Sten Bergstrom and Arne Forsman, *Devepment of a conceptual deterministic rainfall-runoff model*, Nordic Hydrology **4** (1973), 147–170.
- [4] Keith J. Beven, *Rainfall-runoff modeling-the premier*, Wiley and Chichester, 2001.
- [5] G. E. P. Box and G. M. Jenkins, *Time series analysis, forecasting and control*, 1976.
- [6] R. J. C. Burnash, *The nws river forecast system - catchment modeling.*, Computer models of watershed hydrology. **2** (1995), 311–366.
- [7] W. Buytaert, R.Céleri, B. De Bièvre, F. Cisneros, Guido Wyseure, Jozef Deckers, and Robert Hofstede, *Human impact on the hydrology of the andean páramos*, Earth-Science Reviews **79** (2006), 53 – 72.
- [8] Ven Te Chow, David R. Maidment, and Larrys W. Mays, *Hidrología aplicada*, Mc Graw Hill, 1994.
- [9] Peter C.Young, C. J. Taylor, W. Tych, and Pedregal, *The captain toolbox*.
- [10] DHI, *A modelling system for rivers and channels*, DHI, 2007.
- [11] Q. Duan, V. K. Gupta, and S. Sorooshian, *A shuffled complex evolution approach for effective and efficient global minimization*, J. Optimiz. Theory Applic **76** (1993), 501521.
- [12] Simon Haykin, *Kalman filtering and neural networks.*, John Wiley and Sons, INC., 2001.



- [13] Robert Horton, *The role of infiltration in the hydrological cycle*, Transactions of the American Geophysical Union **14** (1933), 446460.
- [14] R. E. Kalman, *A new approach to linear filtering and prediction problems*, Transactions of the ASME Journal of Basic Engineering **82** (1960), 35–45.
- [15] Matthew J. Lees, *Advances in transfer function based flood forecasting*, In Flood forecasting: What does current research offer the practitioner? ed. M. J. Lees & P. Walsh (BHS occasional paper no. 12, produced by the Centre for Ecology and Hydrology on behalf of the British Hydrological Society.) **7** (2000), 40–54.
- [16] _____, *Data-based mechanistic modelling and forecasting of hydrological systems*, Journal of Hydroinformatics (2000).
- [17] Dennis P. Lettenmaier and Eric F. Wood, *Handbook of hydrology*, ch. Hydrologic forecast, pp. 26.1–26.30, McGraw-Hill, 1993.
- [18] Zhulu Lin, *Modeling environmental systems under uncertainty: Towards a synthesis of data-based and theory-based models*, Ph.D. thesis, The University of Georgia, 2003.
- [19] Henrik Madsen, *Automatic calibration for rainfall-runoff model using multi-objective*, Journal of Hydrol **235** (2000), 276–278.
- [20] Henrik Madsen and Michael Kristensen, *A multi-objective calibration framework for parameter estimation in the mike she integrated hydrological modelling system*, Proceedings of the 4th International Conference on Calibration and Reliability in Groundwater Modelling, 2002.
- [21] Henrik Madsen, Geoffrey Wilson, and Hans Christian Ammentorp, *Comparison of different automated strategies for calibration of rainfall-runoff models*, Journal of Hydrology **261** (2002), 48–59.
- [22] David Maréchal, *A soil-based approach to rainfall-runoff modelling in ungauged catchments for england and wales*, Ph.D. thesis, Cranfield University at Silsoe Institute of Water and Environment, 2004.
- [23] Francisco Javier Aparicio Mijares, *Fundamentos de hidrología de superficie*, Limusa, 1989.
- [24] Mark Mulligan and Jhon Wainwright, *Environmental modelling: Finding simplicity in complexity*, ch. Modelling and Model Building, p. Mark Mulligan and Jhon Wainwright, John Wiley and Sons, 2004.
- [25] I. E. Nash and I. V. Sutcliffe, *River flow forecasting through conceptual models*, Journal of Hydrology **10** (1970), 282290.



- [26] J. E. Nash, *Systematic determination of unit hydrograph parameters*, JOURNAL OF GEOPHYSICAL RESEARCH **64** (1959), 111–115.
- [27] J. A. Nelder and R. Mead., *A simplex method for function minimization.*, Computer **7** (1965).
- [28] Steen Asger Nielsen and Eggert Hansen, *Numerical simulation of the rainfall-runoff process on daily basis.*, Nordic Hydrology **4** (1973), 171–190.
- [29] D. J. Pedregal, C. J. Taylor, and P. C. Young, *System identification, time series analysis and forecasting.*, Centre for Research on Environmental Systems and Statistics (CRES), and Lancaster University, Lancaster, LA1 4YQ, United Kingdom., 2007.
- [30] M. Ratto., P. C. Young., R. Romanowicz., F. Pappenberger., A. Saltelli., and A. Paganó., *Uncertainty, sensitivity analysis and the role of data based mechanistic modeling in hydrology*, Hydrology and Earth System Sciences **11** (2007), 12491266.
- [31] Jens Christian Refsgaard and Jesper Knudsen, *Operational validation and intercomparison of different types of hydrological models*, Water Resources Research **32** (1996), no. 7, 21892202.
- [32] Renata J. Romanowicz, Adam Kiczko, and Florian Pappenberger, *A state dependent nonlinear approach to flood forecasting*, PUBLS. INST. GEOPHYS. POL. ACAD. SC., **7** (2007), 223–230.
- [33] Renata J. Romanowicz, Peter C. Young, Keith J. Beven, and Florian Pappenberger, *A data based mechanistic approach to nonlinear flood routing and adaptive flood level forecasting*, Elsevier Ltd. (2008).
- [34] Dan Rosbjerg and Henrik Madsen, *Concepts of hydrologic modeling*, Encyclopedia of Hydrological Sciences **1** (2005), 155–165.
- [35] Supiah Shamsudin and Normala Hashim, *Rainfall runoff simulation using mike11 nam*, JURNAL KEJURUTERAAN AWAM **2** (2002), 1–13.
- [36] Murugesu Sivapalan and Peter C. Young, *Downward approach to hydrological model development*, John Wiley and Sons, Ltd **134** (2005), 2–16.
- [37] E. Todini, *Hydrological catchments modelling: past, present and future.*, Hydrology and Earth System Sciences (2007).
- [38] Jasper A. Vrugt, Hoshin V. Gupta, Willem Bouten, and Soroosh Sorooshian, *A shuffled complex evolution metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters*, Water Resources Research **39** (2003), 1–16.



- [39] Raúl F. Vázquez, *Modelación hidrológica de una microcuenca altoandina ubicada en el austro ecuatoriano*, *Maskana* **1** (2010), 79–90.
- [40] Raul Fernando Vázquez, *Assessment of the performance of physically based distributed codes simulating medium size hydrological systems*, Ph.D. thesis, Katolieke Universiteit Leuven, 2003.
- [41] H. S. Wheater, A. J. Jakeman, and K. J. Beven, *Modelling change in environmental systems*, ch. Progress and directions in rainfall-runoff modelling, pp. 101–132, John Wiley and Sons Ltd: Chichester, 1993.
- [42] Howard Wheater, Soroosh Sorooshian, and Irvine K. D. Sharma, *Hydrological modelling in arid and semi-arid areas*, Cambridge University Press, 2008.
- [43] WMO, *Intercomparison of conceptual models used in operational hydrological forecasting.*, World Meteorological Organization **7** (1975).
- [44] P. C. Young and C. M. Tomlin, *Data-based mechanistic modelling and adaptive flow forecasting*, Flood forecasting: What does current research offer the practitioner? (Matthew Lees and Peter Walsh, eds.), 2000, pp. 25–39.
- [45] Peter C. Young, *Data-based mechanistic modelling of environmental, ecological, economic and engineering systems.*, *Environmental Modelling and Software* **13** (1998), 105–122.
- [46] _____, *Nonstationary time series analysis and forecasting*, *Progress in Environmental Science* **1** (1999), 348.
- [47] _____, *Data-based mechanistic modelling and validation of rainfall-flow processes.*, J. Wiley (2001).
- [48] _____, *Advances in real-time flood forecasting*, The Royal Society (2002).
- [49] _____, *Data-based mechanistic and top-down modelling*, IEMSs. International Environmental Modelling and Software Society, (2002).
- [50] _____, *Recursive estimation and time-series analysis: An introduction for the student and practitioner*, Springer, 2011.
- [51] Peter C. Young, P. McKenna, and J. Bruun, *Identification of non-linear stochastic systems by state dependent parameter estimation*, *International Journal of Control* **74** (2001), 1837–1857.