



UNIVERSIDAD AUTÓNOMA CHAPINGO

**POSGRADO EN INGENIERÍA AGRÍCOLA
Y USO INTEGRAL DEL AGUA**

**PRONÓSTICO DE CAUDALES A CORTO PLAZO IMPLEMENTANDO EL
FILTRO DE KALMAN EN DOS CUENCAS HIDROGRÁFICAS DEL CENTRO Y
GOLFO DE MÉXICO**

TESIS

**QUE COMO REQUISITO PARCIAL PARA
OBTENER EL GRADO DE**

MAESTRO EN INGENIERÍA AGRÍCOLA Y USO INTEGRAL DEL AGUA

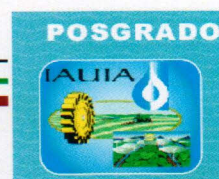
PRESENTA:

FERNANDO GONZÁLEZ LEIVA



DIRECCION GENERAL ACADEMICA
DEPTO. DE SERVICIOS ESCOLARES
OFICINA DE EXAMENES PROFESIONALES

Chapingo, Estado de México, Septiembre 2014.

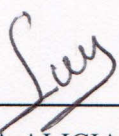


**PRONÓSTICO DE CAUDALES A CORTO PLAZO IMPLEMENTANDO EL
FILTRO DE KALMAN EN DOS CUENCAS HIDROGRÁFICAS DEL CENTRO
Y GOLFO DE MÉXICO**

Tesis realizada por **Fernando González Leiva** bajo la dirección del comité asesor indicado, aprobada por el mismo y aceptada como requisito parcial para obtener el grado de:

MAESTRO EN INGENIERÍA AGRÍCOLA Y USO INTEGRAL DEL AGUA

DIRECTORA:


Ph.D. LAURA ALICIA IBÁÑEZ CASTILLO

ASESOR:


DR. MARIO ALBERTO VÁZQUEZ PEÑA

ASESOR:


DR. AGUSTÍN RUIZ GARCÍA

AGRADECIMIENTOS

Al Consejo Nacional de Ciencia y Tecnología (CONACYT) por la beca otorgada para mis estudios de maestría y estancia de investigación.

A la Universidad Autónoma Chapingo, por brindarme la oportunidad de realizar mis estudios de maestría.

Al Posgrado en Ingeniería Agrícola y Uso Integral del Agua por el apoyo recibido durante el transcurso de la maestría.

A la Ph.D. Laura Alicia Ibáñez Castillo, por todo el apoyo académico, profesional y moral recibido durante la maestría. Infinitas gracias por su excelente orientación, profesionalismo y por brindarme la oportunidad de realizar la estancia de investigación.

Al Dr. Eugenio Romanchick Kriuchkova, por su espíritu de colaboración y apoyo.

Al Dr. Mario Alberto Vázquez Peña, por su disposición, colaboración y recomendaciones para llevar a cabo este proyecto.

Al Dr. Agustín Ruiz García por su constante apoyo, disposición y recomendaciones para llevar a cabo este proyecto.

Al Ph.D Juan B. Valdés en el Departamento de Hidrología de la Universidad de Arizona, U.S.A por haberme permitido realizar bajo su orientación la estancia de investigación. Gracias por compartir su inmenso conocimiento, sencillez y profesionalismo para llevar a cabo este proyecto. Ante todo, una gran persona y excelente investigador.

Al Dr. Mauricio Carrillo García, por su colaboración en la coordinación del posgrado.

A los profesores que hicieron parte de mi formación académica durante la maestría.

A Rossy y Mayra, por brindarme todo su apoyo como secretarias del programa.

DEDICATORIA

A Dios todo poderoso, por los dones recibidos durante mi estancia en México y por la oportunidad que me brindó para alcanzar este objetivo.

A mi santa madre Amalia Leiva, por su sacrificio y esmero en brindarme lo mejor para cumplir con mis sueños. A ti mama, te ofrezco este triunfo en tierras Mexicanas, porque a pesar de las dificultades que hemos pasado, siempre estuviste al frente con el apoyo necesario. Sólo una madre como tú, es capaz de entender lo que un hijo como yo, es capaz de valorar. Gracias Mama!

A mi hermana Karen Gonzalez Leiva, porque siempre me ha apoyado. Gracias por comprender que todo lo que hago, es por el bienestar de nosotros como familia.

A mis familiares en Colombia, porque siempre han creído en mí y gracias a su apoyo he vencido las dificultades que se me han presentado a lo largo de mi vida. En especial quiero dedicarles este logro a mis tíos Alfonso, Pedro, Noé y Martha.

A mi abuela Josefa Bastidas Q.E.P D., porque siempre creyó en mis capacidades y desafortunadamente me vio partir de casa sin conocer mi rumbo y nunca más la pude volver a ver. Estos logros son para ti abuela.

A mi pueblo del cual me siento muy orgulloso, Pacarní-Huila, Colombia.

A Marlio y Nadia porque fueron las personas que me brindaron todo el apoyo durante mis estudios de maestría. También para el pequeño Khalid, porque soy el tío lango.

A la paísa Ana María Toro, por su colaboración y compañía en el transcurso de la maestría. Paisalanga, usted sabe que yo siempre fui lo mejorcito que usted no quiso aprovechar.

A mi compañero de residencia de posgrado, Guillermo Cebada por la compañía y trasnochadas en nuestra área de estudio. Excelente amigo y persona, gracias por los buenos momentos como cuates y compas.

A Juan Carlos Torres, El Monsi, porque siempre estaba listo para salir a jugar futbol y por las desveladas jugando Xbox cuando no teníamos trabajo.

A todos los compañeros de residencias de posgrado desde el 2012 hasta el 2014.

DATOS BIOGRÁFICOS

Fernando González Leiva de nacionalidad Colombiano nació el 17 de julio de 1988 en la ciudad de Cali, Departamento del Valle del Cauca, pero fue en la Inspección de Pacarní, Departamento del Huila en el año de 1994 donde inicio sus estudios de primaria, secundaria y preparatoria finalizando en el año 2005.

En el año 2006 inicia sus estudios de educación superior ingresando a la Universidad Surcolombiana al programa de Ingeniería Agrícola, obteniendo reconocimientos académicos en el transcurso de su carrera. En febrero del año 2012 obtiene su título de Ingeniero Agrícola bajo la modalidad de mejor promedio académico de licenciatura en la Facultad de Ingeniería.

Se desempeñó como Ingeniero laborando para la empresa General Equipos de Colombia S. A, en el área de agricultura de precisión en cultivos de caña de azúcar, soya y maíz en los Departamentos de Valle, Vichada y Tolima, Colombia.

En Agosto del año 2012 ingreso a la Maestría en Ingeniería Agrícola y Uso Integral del Agua en la Universidad Autónoma Chapingo. Durante sus estudios de maestría, y gracias a las gestiones de la Ph.D Laura A. Ibáñez, realizó una estancia de investigación en el Department of Hydrology and Water Resources en The University of Arizona, Tucson-Estados Unidos, bajo asesoría del Ph.D Juan B. Valdés.

Pronóstico de caudales a corto plazo implementando el filtro de Kalman en dos cuencas hidrográficas del centro y Golfo de México.

Short-term streamflow forecasting using the Kalman filter in two basins in central México and the Gulf of México

Fernando González Leiva¹

Laura Alicia Ibáñez Castillo²

RESUMEN

En este trabajo se implementaron tres modelos de pronósticos de caudales a corto plazo utilizando el algoritmo del filtro de Kalman, evaluado para las cuencas del río Turbio y río Cazones pertenecientes al centro y golfo de México, vulnerables a inundaciones cuando se presentan las épocas de lluvia. Se propuso tres metodologías con el objetivo de establecer las diferencias en implementación y eficiencia de pronóstico a diferentes pasos en el tiempo para cada una de las cuencas evaluadas. La primera metodología implementada fue el algoritmo del filtro de Kalman Discreto (DKF). En la segunda se consideró un modelo autorregresivo ARX operado bajo el filtro de Kalman Discreto, y en la tercera, fue implementado el filtro de Kalman de Conjuntos (EnKF) igual bajo el modelo ARX. Se obtuvo que los pronósticos realizados con 24 h de anticipación presentaron mejores ajustes en términos de Nash-Sutcliffe y RMSE que los pronósticos realizados a 48 h, 72 h y 96 h de anticipación en todas las metodologías para las 2 cuencas evaluadas. Los resultados obtenidos para la cuenca del río Turbio fueron los mejores pronósticos comparados con los obtenidos en la cuenca del río Cazones, pues en esta cuenca la presencia de eventos meteorológicos extremos originados en el océano atlántico, generan mayor incertidumbre en los datos pronosticados debido al cambio repentino de las funciones de respuesta.

Palabras Clave: filtros de Kalman, pronósticos, inundaciones.

ABSTRACT

In this paper, three models for short-term streamflow forecasting were implemented using the Kalman filter algorithm, evaluated for the Turbio and Cazones basins in central Mexico and the Gulf of Mexico, which are vulnerable to flooding during the rainy season. Three methodologies were proposed to establish differences between implementation and forecast efficiency at different steps in time for each of the basins evaluated. The Discrete Kalman Filter (DKF) was the first methodology implemented. The second considered an autoregressive model ARX operated with the discrete Kalman filter and the third implemented the ensemble Kalman filter (EnKF) also with the autoregressive model ARX. It was found that the forecast made with 24 h ahead presented better settings for the Nash-Sutcliffe and RMSE than forecasts made with 48 h, 72 h y 96 h ahead for all methodologies in the two basins evaluated. The results obtained for the Turbio basin was the best forecast compared with that obtained for the Cazones river basin, because the extreme meteorological events originating in the Atlantic Ocean create more uncertainty for data forecast due to sudden changes in the response function.

Key words: Kalman filters, forecasting, flood.

TABLA DE CONTENIDO

AGRADECIMIENTOS	iii
DEDICATORIA.....	iv
DATOS BIOGRÁFICOS.....	v
ÍNDICE DE CUADROS.....	ix
ÍNDICE DE FIGURAS.....	xi
1. INTRODUCCIÓN	1
2. OBJETIVOS.....	4
2.1. Objetivo General	4
2.2. Objetivos Específicos	4
2.3. Hipótesis	5
3. REVISIÓN DE LITERATURA	6
3.1. Relaciones Lluvia Escurrimiento	6
3.2. Modelación de Procesos Hidrológicos	7
3.3. Predicción Hidrológica.....	10
3.4. Incertidumbre en la Predicción Hidrológica.....	13
3.5. El Filtro de Kalman	15
3.5.1. El Filtro de Kalman Discreto (DKF).....	16
3.5.2. Proceso de Estimación en el Filtro de Kalman Discreto	18
3.5.3. El Filtro de Kalman de Conjuntos (EnKF)	22
3.5.4. Algoritmo para el Filtro de Kalman Ensemble (EnKF).....	24
3.6. Modelos Autorregresivos y filtros de Kalman para el pronóstico de caudales	28
3.6.1. Modelos Autorregresivos (AR).....	28
3.6.2. Modelos Autorregresivos con Entradas Exógenas (ARX).....	29
3.6.3. Formulación en espacio de estados del modelo ARX.....	30
3.6.4. Filtros de Kalman en el pronóstico de caudales en cuencas hidrográficas	32

4. MATERIALES Y MÉTODOS	39
4.1. Cuenca 1: Río Turbio	39
4.1.1. Climatología.....	40
4.1.2. Precipitación.....	40
4.1.3. Red Hidrográfica y de Presas.....	42
4.1.4. Red de Estaciones Meteorológicas	44
4.2. Cuenca 2: Río Cazonces.....	46
4.2.1. Climatología.....	47
4.2.2. Precipitación.....	48
4.2.3. Red Hidrográfica.....	49
4.2.4. Red de Estaciones Meteorológicas	50
4.3. Metodología.....	53
4.3.1. Filtro de Kalman Discreto DKF	53
4.3.2. Modelo ARX en conjunto con el Filtro de Kalman Discreto	56
4.3.3. Modelo ARX a través del Filtro de Kalman de Conjuntos EnKF.....	60
5. RESULTADOS Y DISCUSIÓN	62
5.1. Cuenca 1: Río Turbio	62
5.1.1. Filtro de Kalman Discreto DKF	62
5.1.2. Enfoque ARX+DKF	70
5.1.3. Enfoque ARX+EnKF	87
5.2. Cuenca 2: Río Cazonces.....	93
5.2.1. Filtro de Kalman Discreto DKF	93
5.2.2. Enfoque ARX+DKF	99
5.2.3. Enfoque ARX+EnKF	114
6. CONCLUSIONES Y RECOMENDACIONES	126
7. BIBLIOGRAFÍA	129

ÍNDICE DE CUADROS

Cuadro 1. Clasificación de las cuencas según tamaño.....	7
Cuadro 2. Periodos de retorno para gastos máximos en la estación hidrométrica	43
Cuadro 3. Estaciones meteorológicas en la cuenca del río Turbio.	45
Cuadro 4. Factores de Thiessen para la cuenca del río Turbio.	45
Cuadro 5. Estaciones meteorológicas localizadas en la cuenca del río Cazones.....	51
Cuadro 6. Factores de Thiessen para la cuenca río Cazones.	52
Cuadro 7. Resumen estadístico para el pronóstico en la cuenca del río Turbio serie 2003.....	63
Cuadro 8. Resumen estadístico para el pronóstico de caudales en la cuenca del río Turbio serie 2004.	67
Cuadro 9. Resumen estadístico para diferentes pronósticos mediante ARX+DKF en la cuenca del río Turbio para la serie del año 2003.	71
Cuadro 10. Resumen estadístico para los pronósticos agrupados por clases en diferentes pasos del tiempo utilizando datos del año 2003.....	75
Cuadro 11. Resumen estadístico para los pronósticos a diferentes pasos mediante ARX+DKF en la cuenca del río Turbio para la serie del 2004.....	79
Cuadro 12. Resumen estadístico para los pronósticos agrupados en diferentes pasos del tiempo utilizando datos del año 2004.	83
Cuadro 13. Resumen estadístico para el pronóstico mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2003.....	87
Cuadro 14. Resumen estadísticas para pronóstico mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2004.....	88
Cuadro 15. Resumen estadístico para el pronóstico de caudales en la cuenca río Cazones para la serie 2008-2011.....	94
Cuadro 16. Resumen estadístico para el pronóstico de caudales en la cuenca del río Cazones para la serie 2005.	94
Cuadro 17. Resumen estadístico para diferentes pasos en el tiempo mediante ARX+DKF en la cuenca del río Cazones para la serie 2008-2011.....	99
Cuadro 18. Resumen estadístico para los caudales agrupados en la cuenca del río Cazones para la serie 2008-2011.....	103
Cuadro 19. Resumen estadístico para diferentes pronósticos mediante ARX+DKF en la cuenca del río Cazones para la serie del año 2005.....	108
Cuadro 20. Resumen estadístico para los pronósticos agrupados por clases en la cuenca del río Cazones para la serie 2008-2011.	110

Cuadro 21. Resumen estadístico para el pronóstico mediante ARX+EnKF en la cuenca del río Cazonés para la serie 2008-2011.	115
Cuadro 22. Resumen estadístico para el pronóstico mediante ARX+EnKF en la cuenca del río Cazonés para la serie del año 2005.	115
Cuadro 23. Resumen estadístico para el pronóstico de caudales a través del DKF en las dos cuencas evaluadas para diferentes periodos.	123
Cuadro 24. Resumen estadístico para el pronóstico de caudales a través del ARX+DKF en las dos cuencas evaluadas en diferentes periodos.	124
Cuadro 25. Resumen estadístico para el pronóstico de caudales según ARX+EnKF en diferentes periodos en las dos cuencas evaluadas.	125

ÍNDICE DE FIGURAS

Figura 1. Concepto de sistema y modelo hidrológico.....	8
Figura 2. Clasificación de los modelos matemáticos hidrológicos.....	9
Figura 3. Relación de convolución en tiempo continuo.....	12
Figura 4. Algoritmo para el filtro de Kalman Discreto.....	22
Figura 5. Esquema operativo del EnKF.....	27
Figura 6. Sistema discreto para única entrada y única salida.....	30
Figura 7. Localización de la cuenca del río Turbio.....	39
Figura 8. Isoyetas de precipitación en la cuenca del río Turbio.	41
Figura 9. Distribución comparativa de la precipitación media mensual en la cuenca río Turbio.....	42
Figura 10. Red hidrográfica y de presas en la cuenca del río Turbio.	44
Figura 11. Polígonos de Thiessen para la cuenca del río Turbio.	46
Figura 12. Localización de la cuenca del río Cazones.	47
Figura 13. Distribución comparativa de la precipitación media mensual en la cuenca río Cazones.	48
Figura 14. Isoyetas de precipitación para la cuenca del río Cazones.....	49
Figura 15. Red hidrográfica en la cuenca del río Cazones.....	50
Figura 16. Polígonos de Thiessen para la cuenca del río Cazones.	51
Figura 17. Esquema de conformación del vector $H(t)$	55
Figura 18. Descripción del pronóstico de caudales con L pasos adelante.	58
Figura 19. Descripción del proceso de pronóstico de caudales con base en periodos de estimación de parámetros del modelo ARX.....	59
Figura 20. Pronóstico de caudales en la cuenca río Turbio para la serie del año 2003. .	64
Figura 21. Caudales pronosticados y actualizados (a, b) y sus errores (c, d) generados en la cuenca del río Turbio para la serie del año 2003.....	65
Figura 22. Diagramas de dispersión para caudales pronosticados (a) y actualizados (b) para la cuenca del río Turbio para la serie del año 2003.....	65
Figura 23. Pronóstico de caudales en la cuenca del río Turbio para la serie del año 2004.	68
Figura 24. Caudales pronosticados y actualizados (a, b) y sus errores (c, d) en la cuenca del río Turbio para la serie del año 2004.....	69
Figura 25. Diagramas de dispersión para caudales pronosticados(a) y actualizados (b) en la cuenca del río Turbio para la serie del año 2004.....	69

Figura 26. Pronóstico de caudales mediante ARX+DKF en la cuenca del río Turbio para la serie del año 2003.....	72
Figura 27 Diagramas de dispersión comparativos para pronósticos con ARX+DKF a 24 h (a), 48 h (b), 72 h (c) y 96 h (c) en la cuenca del río Turbio para la serie del año 2003.	73
Figura 28. Diagramas de dispersión por clases para los pronósticos en diferentes pasos en el tiempo en la cuenca del río Turbio para la serie del año 2003.	74
Figura 29. Funciones de distribución de probabilidad para los errores de predicción con diferentes pasos en el tiempo en la cuenca del río Turbio para la serie del año 2003.	76
Figura 30. Diagrama de cajas para la distribución de los errores en diferentes pasos en el tiempo.....	77
Figura 31. Intervalo de predicción al 95% de probabilidad para L=1 o a 24 h en la cuenca del río Turbio para la serie del año 2003.	78
Figura 32. Pronóstico de caudales mediante ARX+DKF en la cuenca del río Turbio para la serie del año 2004.....	80
Figura 33. Diagramas de dispersión comparativos para pronósticos a diferentes pasos hacia adelante en la cuenca del río Turbio para la serie del año 2004.	81
Figura 34. Diagramas de dispersión agrupados para los pronósticos en diferentes pasos en el tiempo para la cuenca del río Turbio con la serie del año 2004.	82
Figura 35. Funciones de distribución de probabilidad para los errores de predicción con diferentes pasos en el tiempo en la cuenca del río Turbio para la serie del año 2004.	83
Figura 36. Intervalo de predicción al 95% de probabilidad para L=1 o a 24 h en la cuenca del río Turbio para la serie del año 2004.	85
Figura 37. Comparación de los errores pronosticados para diferentes pasos hacia adelante en la cuenca del río Turbio para la serie del año 2003.	86
Figura 38. Comparación de los errores pronosticados para diferentes pasos hacia adelante en la cuenca del río Turbio para la serie del año 2004.	86
Figura 39. Pronóstico de caudales mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2003.....	89
Figura 40. Conjunto de miembros generados para el pronóstico de caudales mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2003.....	90
Figura 41. Diagramas de dispersión para la media del conjunto de caudales pronosticados (a) y actualizados (b) mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2003.....	90
Figura 42. Pronóstico de caudales mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2004.....	91

Figura 43. Conjunto de miembros generados para el pronóstico de caudales mediante ARX+EnKF en el río Turbio para la serie año 2004.	92
Figura 44. Diagramas de dispersión para la media del conjunto de caudales pronosticados (a) y actualizados (b) mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2004.....	92
Figura 45. Pronóstico de caudales mediante DKF en la cuenca del río Cazones para la serie 2008-2011.....	95
Figura 46. Caudales pronosticados y actualizados (a, b) y sus errores (c, d) obtenidos mediante el DKF en la cuenca del río Cazones para la serie 2008-2011.....	96
Figura 47. Diagramas de dispersión para caudales pronosticados (a) y actualizados (b) en la cuenca del río Cazones para la serie 2008-2011.	96
Figura 48. Pronóstico de caudales utilizando el DKF en la cuenca del río Cazones para la serie del año 2005.....	97
Figura 49. Caudales pronosticados y actualizados (a, b) y sus errores (c, d) obtenidos en la cuenca del río Cazones para la serie del año 2005.....	98
Figura 50. Diagramas de dispersión para caudales pronosticados (a) y actualizados (b) en la cuenca del río Cazones para la serie del año 2005.	98
Figura 51. Pronóstico de caudales mediante ARX+DKF en la cuenca del río Cazones para la serie 2008-2011.	100
Figura 52. Errores generados para diferentes pasos en el tiempo en la cuenca del río Cazones para la serie 2008-2011.....	101
Figura 53. Diagramas de dispersión comparativos para pronósticos a diferentes pasos en el tiempo en la cuenca del río Cazones para la serie 2008-2011.	101
Figura 54. Funciones de distribución de probabilidad para los errores de predicción con diferentes pasos en el tiempo en la cuenca del río Cazones para la serie 2008-2011....	102
Figura 55. Diagrama de cajas para la distribución de los errores para diferentes pasos en el tiempo en la cuenca del río Cazones para la serie 2008-2011.	103
Figura 56. Diagramas de dispersión por clases para los pronósticos en diferentes pasos en el tiempo en la cuenca del río Cazones para la serie 2008-2011.....	105
Figura 57. Avance de la tormenta tropical Arlene el 30/Junio/2011 sobre la costa del Golfo de México	106
Figura 58. Intervalo de predicción al 95% de probabilidad para L=1 o 24 h en la cuenca del río Cazones para la serie 2008-2011.	107
Figura 59. Pronóstico de caudales mediante ARX+DKF en la cuenca del río Cazones para la serie del año 2005.....	109
Figura 60. Diagramas de dispersión comparativos para pronósticos a diferentes pasos hacia adelante en la cuenca del río Cazones para la serie del año 2005.	110

Figura 61. Diagramas de dispersión por clases para los pronósticos con diferentes pasos en el tiempo para la cuenca del río Cazonos utilizando la serie del año 2005.	111
Figura 62. Funciones de distribución de probabilidad para los errores de predicción con diferentes pasos en el tiempo en la cuenca del río Cazonos para la serie 2005.	112
Figura 63. Intervalo de predicción al 95% de probabilidad para L=1 o 24 h en la cuenca del río Cazonos para la serie del año 2005.	113
Figura 64. Avance del Huracán Stan sobre la costa del Golfo de México	114
Figura 65. Pronóstico de caudales mediante el ARX+EnKF en la cuenca del río Cazonos para la serie 2008-2011.....	116
Figura 66. Conjunto de miembros generados para el pronóstico caudales mediante ARX+EnKF en la cuenca del río Cazonos para la serie 2008-2011.	117
Figura 67. Diagrama de dispersión para la media del conjunto de caudales pronosticados (a) y actualizados (b) mediante el ARX+EnKF en la cuenca del río Cazonos para la serie 2008-2011.....	117
Figura 68. Pronóstico de caudales mediante el ARX+EnKF en la cuenca del río Cazonos para la serie del año 2005.	118
Figura 69. Conjunto de miembros generados en el pronóstico caudales mediante ARX+EnKF en la cuenca del río Cazonos para la serie del año 2005.....	119
Figura 70. Diagrama de dispersión para la media del conjunto de caudales pronosticados (a) y actualizados (b) mediante el ARX+EnKF en la cuenca del río Cazonos para la serie del año 2005.	119

1. INTRODUCCIÓN

Durante los últimos años, las proyecciones sobre cambio climático estiman incrementos significativos de lluvias extremas aumentando la probabilidad de que las inundaciones sean cada vez más constantes, ocasionando impactos negativos en la población civil involucrada.

El estudio de los procesos hidrológicos permite conocer la estructura operativa de un sistema natural, como lo es una cuenca hidrográfica. Analizar las interacciones que ocurren en este sistema, pueden brindar la información suficiente para modelar su comportamiento y plantear escenarios de ocurrencia con el objetivo de conocer con anticipación, los efectos que pueden ser ocasionados por la presencia de eventos hidrometeorológicos extremos.

El pronóstico de caudales en una cuenca hidrográfica, es uno de los procedimientos que permite la toma de decisiones para la planificación del riesgo en beneficio de mitigar los impactos que pueden ser ocasionados por las grandes avenidas en situaciones como la regulación de embalses y la emisión de alertas ante posibles inundaciones en zonas con presencia de población civil.

En este contexto, se hace necesario desarrollar modelos de pronósticos como herramientas que permitan integrar las entradas (lluvia) y mediante la convolución de procesos, conocer en un tiempo inmediato, las descargas (caudales) registrados en un punto a la salida de una cuenca.

Por esta razón, se propuso la implementación del algoritmo del filtro de Kalman para realizar el pronóstico de caudales, el cual, no es concebido como un modelo hidrológico de lluvia escurrimiento, sino un conjunto de ecuaciones matemáticas que proveen una solución recursiva óptima por el método de mínimos cuadrados para la estimación estocástica de estados. A través de un mecanismo de predicción, el algoritmo se provee con registros históricos de precipitación y caudal, el cual le permite estimar un nuevo estado con un grado de incertidumbre y la posibilidad de ir generando nuevas estimaciones en el tiempo mediante la integración de las mediciones observadas, estando cada vez más próximos a la realidad.

En este trabajo se implementaron tres modelos de pronósticos de caudales a corto plazo utilizando el algoritmo del filtro de Kalman, evaluado para las cuencas del río Turbio y río Cazonas localizadas en el centro y Golfo de México, vulnerables a inundaciones cuando se presentan las épocas de lluvia. Las tres metodologías tienen el objetivo de establecer las diferencias en implementación y eficiencia de pronóstico a diferentes pasos en el tiempo (24 h, 48 h, 72 h y 96 h) para cada una de las cuencas evaluadas a partir de datos de precipitación y caudal a nivel diario.

La primera metodología implementada fue el algoritmo del filtro de Kalman Discreto (DKF) en su version original. En la segunda se consideró un modelo autorregresivo con entrada exógena ARX operado bajo el filtro de Kalman Discreto, y en la tercera, se implementó el filtro de Kalman de Conjuntos (EnKF) bajo el mismo modelo ARX. Se obtuvo que los pronósticos realizados 1 paso hacia adelante, es decir con 24 h de anticipacion, presentaron los mejores ajustes en términos del índice de eficiencia de Nash-Sutcliffe y RMSE que los pronósticos realizados a 2, 3 y 4 pasos hacia adelante, es decir 48 h, 72 h y 96 h de anticipacion, en todas las metodologías propuestas para las 2 cuencas evaluadas. Los resultados obtenidos para la cuenca del río Turbio fueron los mejores pronósticos obtenidos comparados con los de la cuenca del río Cazones, pues en esta cuenca la presencia de eventos meteorológicos extremos originados en el Océano Atlántico, generan mayor incertidumbre en los datos pronosticados debido al cambio repentino de las funciones de respuesta.

2. OBJETIVOS

2.1. Objetivo General

- Evaluar la implementación del algoritmo del filtro de Kalman para el pronóstico de caudales a corto plazo con 24 h, 48 h, 72 h y 96 h de anticipación en dos cuencas vulnerables a inundaciones, localizadas en el centro y Golfo de México.

2.2. Objetivos Específicos

- Implementar el algoritmo del filtro de Kalman en sus variantes discreta (DKF) y de conjuntos (EnKF) bajo el esquema de un modelo autorregresivo de entrada exógena ARX para el pronóstico de caudales a corto plazo en las cuencas del río Turbio, en el Estado de Guanajuato y río Cazonas en el Estado de Veracruz.
- Comparar los pronósticos realizados a 24 h, 48 h, 72 h y 96 h de anticipación sin ejecutar la etapa de actualización en los modelos de ARX+DKF implementados en las dos cuencas con base en los registros de precipitación y caudal a nivel diario.
- Evaluar la incertidumbre de los pronósticos realizados en las cuencas seleccionadas mediante la estimación de intervalos de predicción con nivel de significancia establecido.

2.3. Hipótesis

Mediante el uso de técnicas de estimación recursiva de estados como el algoritmo del filtro de Kalman, es posible la implementación de modelos para el pronóstico de caudales a corto plazo en cuencas hidrográficas vulnerables a inundaciones, toda vez que exista una red de instrumentación adecuada para la medición de variables como lluvia y escurrimiento, disminuyendo así, el grado de incertidumbre que puede estar asociado por la ausencia de información que refleje el comportamiento real del sistema.

3. REVISIÓN DE LITERATURA

3.1. Relaciones Lluvia Esguerrimiento

Es común que no se cuente con registros adecuados de esguerrimiento en el sitio de interés para determinar los parámetros necesarios para el diseño y operación de obras hidráulicas. En general, los registros de precipitación son más comunes que los de esguerrimiento y además, no se afectan por cambios en la cuenca, como construcción de obras de almacenamiento y derivación, talas, urbanización etc. Por ello, es conveniente contar con métodos que permitan determinar el esguerrimiento en una cuenca mediante las características de la misma y de precipitación (Aparicio, 1989). Los principales parámetros que intervienen en el proceso de conversión de lluvia a esguerrimiento son:

Área de la cuenca: El tamaño de una cuenca determina las magnitudes de sus crecientes, gastos mínimos y esguerrimientos promedio. Una cuenca pequeña generará crecientes o avenidas repentinas pero pequeñas tanto en gasto máximo como en volumen. Una cuenca grande da origen a crecientes de larga duración y previsibles, pero con gastos y volúmenes extraordinarios.

Cuadro 1. Clasificación de las cuencas según tamaño.

Área (km ²)	Designación
< 25	Muy pequeña
25 - 250	Pequeña
250 - 500	Intermedia - Pequeña
500 - 2500	Intermedia - Grande
2500 - 5000	Grande
> 5000	Muy Grande

Fuente: Campos, 2007.

Características generales de la cuenca: La forma de la cuenca gobierna principalmente la magnitud y la velocidad con la cual el agua llega al cauce principal. Para caracterizar la forma de una cuenca generalmente se tienen parámetros como índice de compacidad y relación de elongación. De igual forma, se tienen parámetros importantes como el perfil del cauce principal, pendiente, altitud media, densidad de drenaje y orden de corrientes.

Distribución de la lluvia: La distribución y trayectoria de una tormenta pueden determinar la distribución del escurrimiento, pues para una cantidad fija de precipitación el volumen de escurrimiento que se origine es diferente si la lluvia fue uniforme sobre la cuenca (precipitación ciclónica o de frente cálido) o si ocurrió sobre una área pequeña (tormenta convectiva o de frente frío).

3.2.Modelación de Procesos Hidrológicos

El objetivo del análisis de un sistema hidrológico es estudiar la operación del sistema y predecir su salida. Un modelo de sistema hidrológico es una aproximación al sistema real; sus entradas y salidas son variables medibles y su estructura es un conjunto de ecuaciones que conectan las entradas y las salidas (Chow *et al.* 1994).

Los modelos matemáticos describen los procesos que ocurren dentro del ciclo hidrológico, que posteriormente son aplicados en cuencas hidrográficas para determinar su respuesta y así entender las interacciones que ocurren dentro del sistema. Los modelos pueden tener múltiples finalidades, se destacan entre ellos: la planificación y el diseño de prácticas de conservación de suelos, el manejo del agua en el riego, la restauración de humedales y la administración de aguas subterráneas, entre otros.

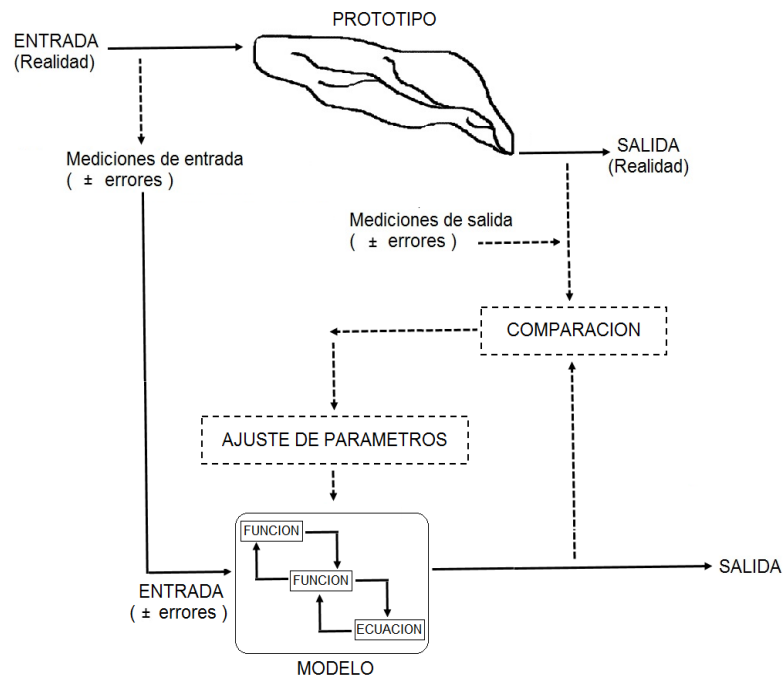


Figura 1. Concepto de sistema y modelo hidrológico Fuente: Campos, 2007.

En forma general, los modelos tienen un uso amplio en proyectos para el riesgo contra inundaciones, manejo de embalses, administración del recurso hídrico para el abastecimiento y evaluación de la calidad del agua para el consumo humano. Toda esta dinámica en los modelos es importante a la hora de evaluar impactos ambientales dentro de una cuenca, ya que se evalúan una serie de procesos que permiten relacionar el estado de un sistema (Singh y Woolhiser, 2002).

Los modelos matemáticos aplicados al manejo de los recursos naturales radican en la simulación de escenarios y predicción de los fenómenos físicos, a corto, mediano y largo plazo. Así mismo, a través de los modelos se pueden obtener relaciones de causa y efecto, sin haber realizado cambios en los sistemas reales de evaluación. La función de un modelo hidrológico es determinar con eficiencia y precisión los componentes del ciclo hidrológico en una cuenca para diferentes escenarios de manejo por eventos en un año o para eventos extremos. De esta manera, se convierten en herramientas para la toma de decisiones una vez planteados los escenarios y sus relaciones con el sistema analizado.

Según Singh (1988), Chow *et al.* (1994) y Jajarmizadeh *et al.* (2012) los modelos matemáticos en hidrología se pueden clasificar de la siguiente manera (figura 2).

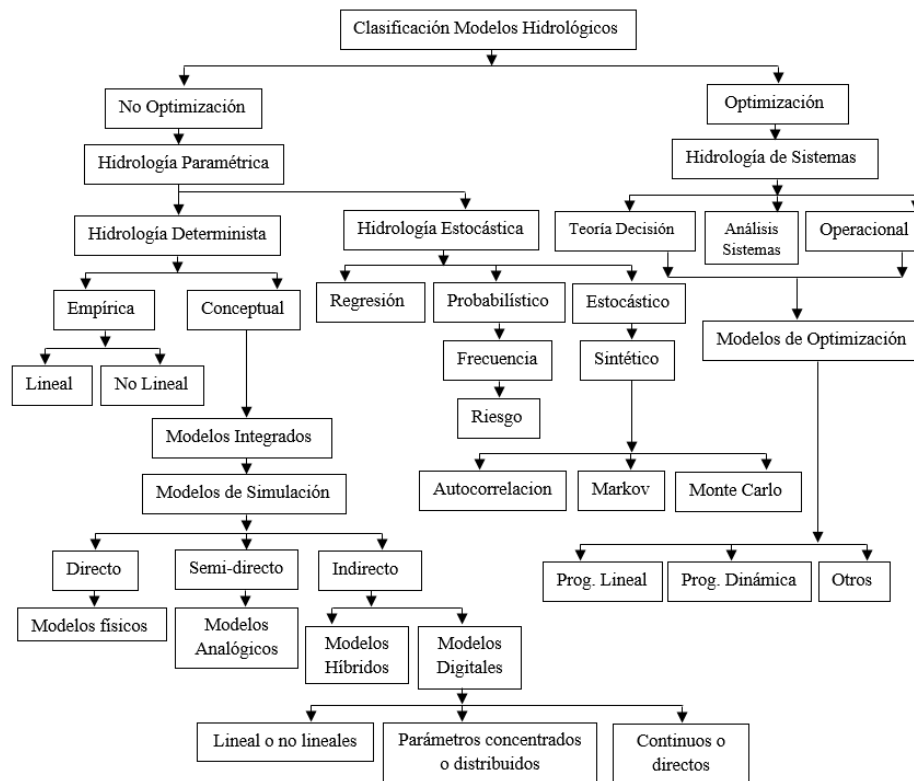


Figura 2. Clasificación de los modelos matemáticos hidrológicos.

3.3. Predicción Hidrológica

Las predicciones hidrológicas son estimaciones de los estados futuros de un fenómeno hidrológico. Dichas predicciones, son esenciales para el funcionamiento eficaz de la infraestructura hídrica y para la atenuación de desastres naturales provocados por las crecidas y sequías, según las condiciones ambientales. Las predicciones son cada vez más importantes como medio de apoyo a la gestión integrada de los recursos hídricos y para la reducción de las pérdidas causadas por las grandes avenidas (WMO, 2009).

La resolución temporal de una predicción puede variar de acuerdo al nivel de exactitud que se desee manejar. Por ejemplo, según la WMO (2009) establece la siguiente clasificación:

- a. Predicciones hidrológicas a corto plazo que abarcan periodos de hasta 2 días.
- b. Predicciones hidrológicas a mediano plazo, que comprenden periodos entre 2 y 10 días.
- c. Predicciones hidrológicas a largo plazo, que abarcan periodos mayores a 10 días.

Para obtener mejores predicciones, es necesario contar con equipos e instrumentos adecuados para realizar el monitoreo de las variables que influyen en la modelación hidrológica. La necesidad de instalar aforadores, pluviómetros y una red meteorológica asociada, es importante para contar con la disponibilidad de los datos provenientes de todas partes, incluso de productos que son provistos por la redes satelitales (WMO, 2009).

Según Valdés *et al.* (1980), uno de los problemas más importantes en hidrología es la predicción de caudales en tiempo real. Para poder realizar esta predicción es necesario postular una relación causal entre la precipitación en varios puntos de la cuenca, a diferentes instantes de tiempo y con un caudal medio para un día determinado. Esta relación causal denominada función de respuesta, puede ser representada como un modelo de caja negra, donde el interés solo radica en las entradas y salidas del sistema sin importar los procesos relacionados al interior del mismo.

De otra manera, la función de respuesta puede ser estudiada analizando los procesos lluvia-escurrimiento a gran escala y la relación puede ser determinística o estocástica, lineal o no lineal, variante o invariante en el tiempo. La teoría del hidrograma unitario propuesta por Sherman en 1932, se convierte entonces en la función de respuesta adecuada para explicar la relación causa-efecto entre precipitación y escurrimiento en una cuenca hidrográfica.

Desde el punto de vista de la teoría de los sistemas dinámicos, los sistemas hidrológicos pueden modelarse como sistemas lineales. Un sistema se denomina lineal si se aplica el principio de superposición, es decir, que la respuesta producida por la aplicación simultánea de entradas diferentes es igual a la suma de sus respuestas individuales (Ogata, 2003). Basado en lo anterior, para una lluvia dada de duración T el sistema puede ser descrito como el representado en la figura 3.

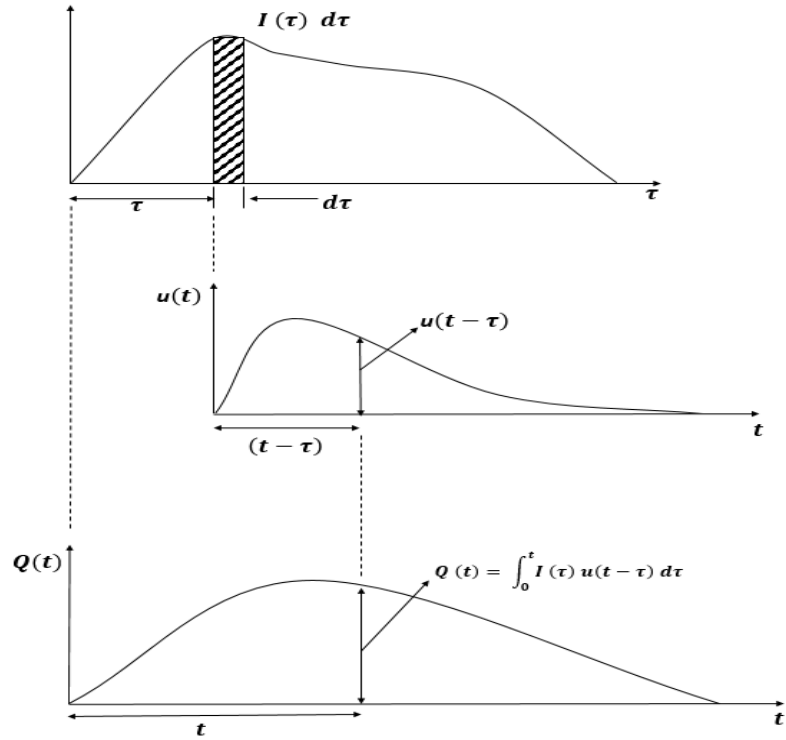


Figura 3. Relación de convolución en tiempo continuo.

La cantidad de lluvia que puede caer sobre el área (A) de una cuenca, puede describirse tal que:

$$P(\tau) = i(\tau) \times (A \times d\tau) \quad (1)$$

Donde $i(\tau)$ es la intensidad promedio en el intervalo $(\tau, \tau + d\tau)$ y $P(\tau)$ es la lluvia caída en el intervalo $(\tau, \tau + d\tau)$. La contribución de esa cantidad de lluvia a la generación de escurrimiento puede describirse en función del hidrograma unitario, según:

$$q(\tau) = i(\tau) \times u(t-\tau) d\tau \quad (2)$$

En este caso, $u(t-\tau)$ se interpreta como la ordenada en el hidrograma unitario e $i(\tau)$ la intensidad promedio en el intervalo $(t, \tau + d\tau)$

Debido a la suposición de que el sistema es lineal y aplicando el principio de superposición de causas y efectos, el caudal escurrido para un periodo (t), puede ser interpretado como:

$$Q(t) = A \int_0^t i(\tau) * u(t-\tau) d\tau \quad (3)$$

La ecuación (3) es conocida como la integral de convolución, fundamental para la solución de sistemas lineales en una escala de tiempo continuo (Gelb, 1974; Chow *et al.* 1994).

Para el caso discreto, la ecuación anterior es reducida a una sumatoria, tal que:

$$Q(t) = A \sum_{\tau=0}^t i(\tau) * u(t-\tau) \Delta\tau \quad (4)$$

3.4. Incertidumbre en la Predicción Hidrológica

El principal objetivo de la predicción hidrológica es proporcionar el periodo de anticipación máximo posible con un grado de exactitud suficiente para que los usuarios adopten medidas que les permitan atenuar las pérdidas y/o optimizar las decisiones para la gestión hídrica. Todas las predicciones contienen elementos de incertidumbre y una de las maneras más eficaces de abordar este problema es la utilización de conjuntos de estados posibles a ocurrir. La incertidumbre asociada a una predicción hidrológica comienza por la meteorología. Dado que todos los modelos atmosféricos en mesoescala tratan de modelar una atmósfera esencialmente caótica, se considera desde hace años que la meteorología es la fuente de incertidumbre más importante para los sistemas de predicción.

Por otra parte, los parámetros de los modelos hidrológicos y la mecánica de los mismos contribuyen también a la incertidumbre o al error vinculado a las predicciones (WMO, 2009).

Según Gouweleeuw *et al.* (2005) el enfoque probabilístico es ahora una práctica establecida en la comunidad meteorológica, pues los predictores sólo han comenzado a buscar maneras de interpretar y mitigar los productos de predicción obtenidos para los usuarios finales, mediante la combinación de los llamados sistemas de predicción por conjuntos (EPS) por sus siglas en inglés, con modelos de lluvia-escorrentía.

El principal mecanismo utilizado para incorporar directamente la incertidumbre ha consistido en perturbar las condiciones iniciales de las ecuaciones diferenciales en derivadas parciales no lineales que describen la atmósfera mediante métodos basados en sistemas convectivos de mesoescala. Sin embargo, la mayoría de los métodos actualmente utilizados distan de ser óptimos y dependen todavía de una valoración subjetiva del predictor. Por esta razón, para describir y propagar la incertidumbre de las predicciones han sido utilizadas técnicas recursivas como los denominados filtros de Kalman (WMO, 2009).

Recientemente el Centro Europeo para el Pronóstico del Tiempo a Medio Plazo (ECMWF) ha venido desarrollando importantes investigaciones en la implementación de sistemas de predicción por conjuntos (EPS) debido a la capacidad de tener en cuenta las incertidumbres y dar lugar a múltiples predicciones mejorando la previsibilidad (Gouweleeuw *et al.*, 2005; Pappenberger *et al.*, 2008; Cloke and Pappenberger, 2009; Haiden *et al.*, 2014).

3.5. El Filtro de Kalman

El filtro de Kalman fue propuesto en 1960 como una solución recursiva por el método de mínimos cuadrados al problema de filtrado lineal de datos discretos. Esta solución permite calcular un estimador lineal, insesgado y óptimo del estado de un proceso en cada momento para un tiempo (t) con base en la información disponible en el tiempo $(t - 1)$, actualizando con la información adicional en el momento (t) , dichas estimaciones (Méndez, 2007).

El filtro es un procedimiento matemático que opera por medio de un mecanismo de predicción y corrección. Éste algoritmo pronostica el nuevo estado a partir de una estimación previa, añadiendo un término de corrección proporcional al error de predicción, de tal forma que este último es minimizado estadísticamente (Welch y Bishop, 2006). Por esta razón, si se tiene conocimiento del sistema dinámico, las estadísticas del sistema ruidoso, los errores de medición y las condiciones iniciales, se puede estimar el estado para un sistema dinámico (Gelb, 1974).

El filtro de Kalman proporciona un buen marco para la estimación incremental de una cantidad en una situación en la cual las mediciones relacionadas con la misma, están disponibles a lo largo del tiempo. Concretamente, se trata de una técnica de estimación bayesiana empleada para seguir sistemas estocásticos dinámicos observados mediante sensores ruidosos. El filtro es muy potente en varios aspectos: es compatible con las estimaciones del pasado, presente y futuro, incluso estados, y puede hacerlo además cuando la naturaleza exacta del sistema modelado es desconocida (Obregón y Rosas, 2003; Welch y Bishop, 2006).

Una de las aplicaciones más interesantes del filtro de Kalman fue el haber sido utilizado en la misión Apollo de la NASA que llevo por primera vez al hombre a la superficie lunar en la década de los sesenta del siglo XX. En años recientes, el filtro de Kalman ha presentado múltiples variaciones con respecto a su concepción inicial, lo cual ha permitido ser implementado en dispositivos de navegación por satélite, teléfonos inteligentes, incluyendo juegos de ordenador, entre otros (Faragher, 2012).

3.5.1. El Filtro de Kalman Discreto (DKF)

Se desempeña suponiendo que el sistema puede ser descrito a través de un modelo estocástico lineal, en donde el error asociado tanto al sistema como a la información adicional que se incorpora en el mismo, tiene una distribución normal con media cero y varianza determinada (Ramírez, 2003).

Según Kalman (1960), Simon (2001), Welch y Bishop (2006), Faragher (2012), el filtro trata de estimar el estado $X \in \mathbb{R}^n$ de un proceso controlado en tiempo discreto, el cual está descrito por la ecuación en diferencias lineales de tipo estocástica:

$$x_{t+1} = A_t x_t + B_t u_t + w_t \quad (5)$$

Donde:

x_{t+1} es el vector de estado con dimensiones $(n \times 1)$ que contiene los términos de interés para el sistema (ej. posición, velocidad, rumbo) en el tiempo $(t + 1)$.

A_t es la matriz con dimensiones $(n \times n)$ que relaciona el estado en el periodo de tiempo previo (t) al estado que ocurre en el momento $(t + 1)$.

x_t representa el vector de estados en el tiempo (t) con dimensiones (n x 1)

B_t representa la matriz de control opcional de entrada con dimensiones (n x k).

u_t es el vector que contiene las entradas del control con dimensiones (k x 1).

w_t es el vector con dimensiones (n x 1) que contiene los términos de ruido del proceso para cada parámetro en el vector de estado. El ruido del proceso se supone que se extrae de una distribución normal de media cero y varianza Q.

Las mediciones del sistema, se expresan de la siguiente manera:

$$Z_t = H_t x_t + v_t \quad (6)$$

Donde:

Z_t es el vector de las mediciones actuales del sistema con dimensiones (m x 1).

H_t es la matriz con dimensiones (m x n) de transformación que mapea los parámetros del vector de estado en el dominio de medición.

v_t es el vector con dimensiones (m x 1) que contiene los términos de ruido de medición para cada observación. El ruido de la medición tiene distribución normal de media cero y varianza R.

Las variables aleatorias w_t y v_t son completamente independientes, sin correlación una de la otra en cualquier tiempo. Finalmente, se puede observar que el sistema de ecuaciones (5) y (6) consiste de un componente determinístico y un componente estocástico que reflejan incertidumbre en el modelo (Lee y Singh, 1998).

3.5.2. Proceso de Estimación en el Filtro de Kalman Discreto

Como el verdadero estado del sistema x_{t+1} no puede ser observado directamente, el filtro de Kalman proporciona un algoritmo para determinar una estimación $\hat{x}_{t+1}$ mediante la combinación de los modelos de ruido tanto del sistema como de las mediciones, para ciertos parámetros o funciones lineales de los mismos. Las estimaciones de los parámetros de interés en el vector de estado son entonces proporcionadas por las funciones de densidad de probabilidad (PDF), en lugar de valores discretos. El filtro de Kalman está basado en funciones de distribución de probabilidades (PDFs) de tipo gaussiano, siendo necesario conocer todos sus parámetros, para luego ser utilizados en la predicción de nuevos estados.

Valdés *et al.* (1980) define el estimador óptimo como la media de la distribución condicional de X_t dado Z_t , es decir:

$$\hat{x}_{t|t} = E[x_t | z_1, z_2, z_3 \dots z_t] \quad (7)$$

Donde el término $(t | t)$ indica que se ha usado toda la información disponible hasta el tiempo (t) para obtener el estimador del estado del sistema. Para obtener $\hat{x}_{t|t}$, es necesario considerar las siguientes hipótesis:

- i. El estado inicial x_0 es un vector $(n \times 1)$ de componentes aleatorios normalmente distribuidos con media y matriz de covarianza conocida.

- ii. El vector w_t está compuesto por $(n \times 1)$ elementos aleatorios distribuidos normalmente con media cero y varianza Q .
- iii. El vector v_t está compuesto por $(m \times 1)$ elementos aleatorios normalmente distribuidos con media cero y varianza R definida.
- iv. Tanto w_t como v_t se asumen mutuamente independientes en cualquier instante del tiempo.
- v. El vector u_t es una secuencia determinística conocida.
- vi. Las matrices A_t, B_t, H_t son todas determinísticas y conocidas con anterioridad.

Estas seis hipótesis junto con la linealidad de las ecuaciones (5) y (6) implican que la distribución de probabilidad del estado del sistema x_t condicional en Z_t , es gaussiana y por tanto puede ser definido por sus dos primeros momentos:

$$E[x_t | z_t] = \hat{x}_{t|t} \quad (8)$$

$$E[(x_t - \hat{x}_t)(x_t - \hat{x}_t)^T] = P_t$$

La violación del supuesto de normalidad conlleva a no poder garantizar que el filtro produzca la media condicional del vector de estado. Sin embargo, el estimador sigue siendo óptimo dentro de los estimadores lineales (Ramírez, 2003).

Las matrices de covarianza para representar el ruido en el sistema y en las mediciones, pueden ser expresadas de acuerdo a la ecuación (9), siendo $E[\cdot]$ el valor esperado de la variable.

$$Q = E[w_t w_t^T] \quad R = E[v_t v_t^T] \quad (9)$$

Según Welch y Bishop (2006), los errores en el proceso de estimación para los estados a priori y a posteriori, se representan como:

$$\begin{aligned} \mathbf{e}_t^- &= \mathbf{x}_t - \hat{\mathbf{x}}_t^- \\ \mathbf{e}_t &= \mathbf{x}_t - \hat{\mathbf{x}}_t \end{aligned} \quad (10)$$

Donde $\mathbf{e}_t^-$ es el error a priori, $\mathbf{x}_t$ es el estado observado, $\hat{\mathbf{x}}_t^-$ el estado estimado a priori, $\mathbf{e}_t$ es el error a posteriori y $\hat{\mathbf{x}}_t$ es el estado estimado a posteriori. Siguiendo las estructuras definidas en las ecuaciones (10) y (11), las covarianzas de los errores a priori y a posteriori pueden ser definidas:

$$\begin{aligned} \mathbf{P}_t^- &= E[\mathbf{e}_t^- \mathbf{e}_t^{-T}] \\ \mathbf{P}_t &= E[\mathbf{e}_t \mathbf{e}_t^T] \end{aligned} \quad (11)$$

Los términos a lo largo de la diagonal principal de $\mathbf{P}_t^-$ y $\mathbf{P}_t$ son las varianzas asociadas a los términos correspondientes en el vector de estado. Los términos fuera de la diagonal de $\mathbf{P}_t^-$ y $\mathbf{P}_t$ proporcionan las covarianzas entre términos en el vector de estado. En el caso de un buen modelado, el sistema lineal unidimensional con errores de medición extraídos de una distribución gaussiana de media cero y varianza conocida, el filtro de Kalman ha demostrado ser un estimador óptimo (Faragher, 2012).

Para calcular el estado estimado a posteriori, se propone una combinación lineal del estado estimado a priori y una diferencia entre la medición real y la predicha, tal que:

$$\begin{aligned}\hat{x}_t &= \hat{x}_t^- + K(z_t - H\hat{x}_t^-) \\ r &= z_t - H\hat{x}_t^-\end{aligned}\tag{12}$$

El término r es conocido como las innovaciones o residuales, los cuales determinan la diferencia entre la medición actual y la predicha. El factor K representa el factor de ganancia de Kalman que minimiza la covarianza del error a posteriori. El factor K se expresa en los términos tal que:

$$\begin{aligned}K_t &= P_t^- H^T (H P_t^- H^T + R)^{-1} \\ K_t &= \frac{P_t^- H^T}{(H P_t^- H^T + R)}\end{aligned}\tag{13}$$

Para interpretar la importancia que representa la ecuación de ganancia en el filtro de Kalman, Simon (2001), Welch y Bishop (2006), realizaron el siguiente análisis:

- i. Si $R \rightarrow 0$ se tiene que las mediciones reales son más confiables con respecto a las mediciones predichas por el sistema $H\hat{x}_t^-$.

$$\lim_{R \rightarrow 0} K_t = H_t^{-1}\tag{14}$$

- ii. Si $P_t^- \rightarrow 0$ las mediciones predichas $H\hat{x}_t^-$ ganan confiabilidad con respecto a las medición reales.

$$\lim_{P_t^- \rightarrow 0} K_t = 0\tag{15}$$

El filtro de Kalman está dividido en dos procesos fundamentales, el primero, denominado de predicción y el segundo denominado de actualización. Según Kim (2011), el filtro de Kalman discreto puede ser ejecutado mediante el siguiente diagrama:

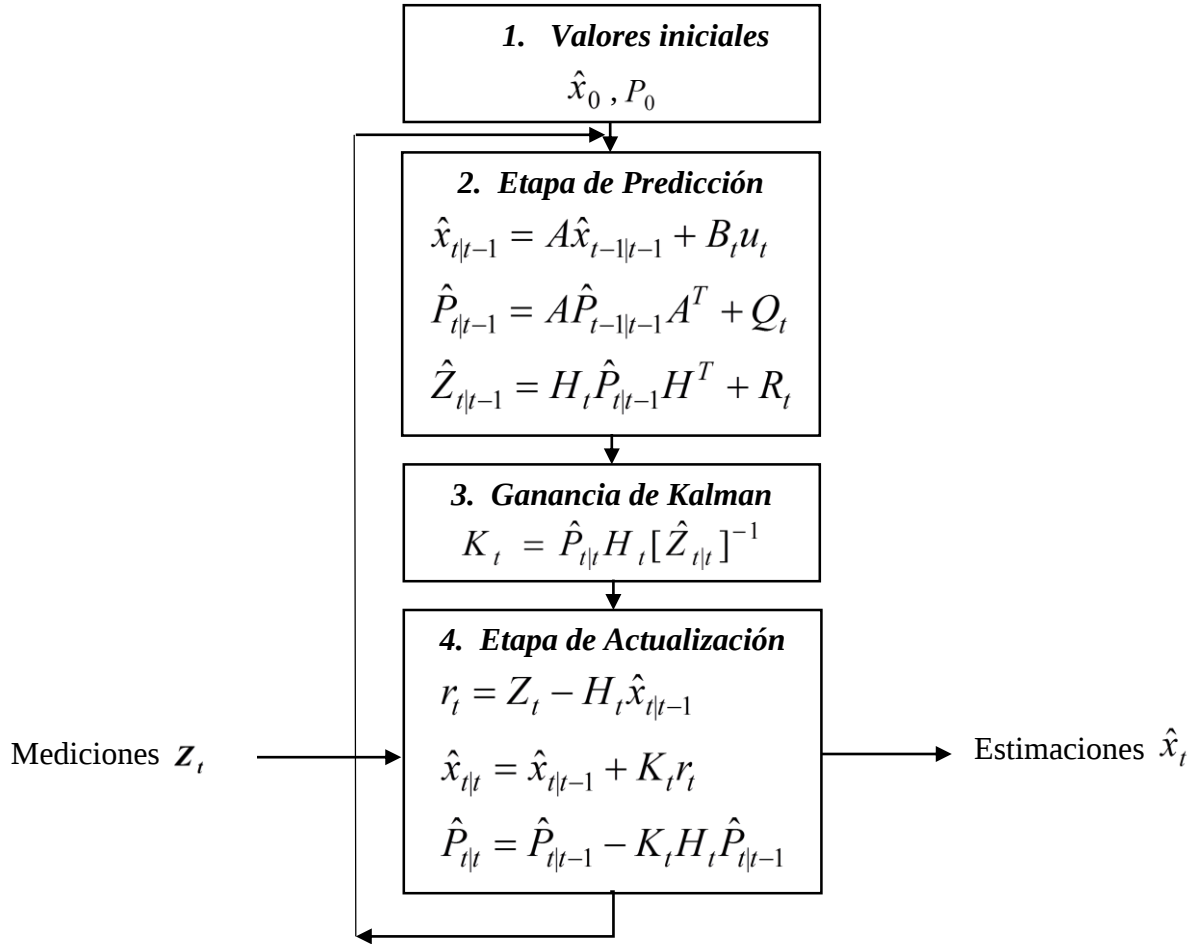


Figura 4. Algoritmo para el filtro de Kalman Discreto.

3.5.3. El Filtro de Kalman de Conjuntos (EnKF)

El filtro de Kalman de Conjuntos (EnKF) fue propuesto por Evensen (1994), basado en técnicas Monte Carlo para asimilación de datos en grandes sistemas. Fue desarrollado para sistemas no lineales y surgió como una variante del filtro de Kalman extendido (EKF) debido a la inestabilidad que éste podía presentar cuando la no linealidad en el sistema era muy fuerte. El EnKF proporciona una formulación conceptual simple y de fácil aplicación, ya a que no hay necesidad de derivar un operador lineal tangente o ecuaciones adjuntas, ni integraciones en tiempos previos.

En gran medida el EnKF supera dos problemas asociados con el filtro de Kalman discreto. En primer lugar, en el DKF la matriz de covarianza de error para el estado del modelo necesita ser almacenada y propagada en el tiempo, haciendo el método no muy práctico desde el punto de vista computacional para los modelos con vectores de estado de grandes dimensiones.

En segundo lugar, cuando la dinámica del modelo es no lineal, el EKF utiliza una ecuación para linealizar la evolución de la covarianza de error y esto puede resultar en inestabilidades lineales acotadas para la evolución de error (Reichle *et al.*, 2002; Evensen, 2009).

En el EnKF el estado está representado por un conjunto de posibles situaciones reales, donde la media del conjunto generado se considera como la estimación del estado y el error del modelo se aproxima por la varianza del conjunto generado. En la etapa de inicio, los conjuntos son muestreados a partir de una distribución a priori y luego durante una fase de predicción cada miembro evoluciona de forma independiente siguiendo el modelo de espacio de estado. Entonces, cada vez que las nuevas observaciones están disponibles, el estado posterior (información adicional del estado verdadero) se estima mediante la actualización de cada miembro generado en el conjunto, basado en la combinación de la información obtenida de la observación y del modelo propio. Estas secuencias se repiten de forma iterativa hacia adelante hasta el tiempo presente, por Kalman suavizado, o para el futuro, por filtrado de Kalman (Hirpa, 2013).

3.5.4. Algoritmo para el Filtro de Kalman Ensemble (EnKF)

Considerando un sistema dinámico no lineal

$$x_{t+1} = f(x_t, u_t) + w_t \quad (16)$$

Con mediciones representadas por

$$z_t = h(x_t) + v_t \quad (17)$$

Donde $x_t, w_k \in \mathbb{R}^n$, $u_t \in \mathbb{R}^m$, $z_t, v_k \in \mathbb{R}^p$, asumiendo además que w_t, v_t están correlacionados con ruido blanco de media cero y matrices de covarianzas Q_t, R_t respectivamente. El objetivo es obtener el mejor estimador x_t^a del estado x_t usando las mediciones z_t tal que $E[e_t^a (e_t^a)^T]$ sea mínimo, donde $e_t^a \in \mathbb{R}^n$ definido como:

$$e_t^a = x_t - x_t^a \quad (18)$$

De acuerdo a lo anterior, la matriz de error de covarianzas para el estado de análisis se define,

$$P_t^a = E[e_t^a (e_t^a)^T] \quad (19)$$

Según Gillijns *et al.* (2006) el método para el EnKF está representado en 3 etapas. La primera, para representar las estadísticas del error de predicción, se asume que en el tiempo (t) se tiene un conjunto de q estados predichos con errores de muestreo aleatorio.

De esta manera, este conjunto se puede denotar como $x_t^f \in \mathbb{R}^{n \times q}$.

$$\mathbf{x}_t^f = (x_t^{f_1}, \dots, x_t^{f_i}) \quad (20)$$

Donde el súper índice f_i indica el i -ésimo miembro del conjunto predicho. Por tanto, la media del conjunto $\bar{\mathbf{x}}_t^f \in \mathbb{R}^n$ está definido por:

$$\bar{\mathbf{x}}_t^f = \frac{1}{q} \sum_{i=1}^q \mathbf{x}_t^{f_i} \quad (21)$$

Como el estado verdadero de $\mathbf{x}_t$ no se conoce, se puede aproximar mediante la media del conjunto predicho. Para el sistema puede ser definido de la siguiente manera:

$$\mathbf{E}_t^f = [\mathbf{x}_t^{f_1} - \bar{\mathbf{x}}_t^f, \dots, \mathbf{x}_t^{f_i} - \bar{\mathbf{x}}_t^f] \quad (22)$$

Adicional, para el conjunto de errores a la salida del sistema $\mathbf{E}_{zt}^a \in \mathbb{R}^{p \times q}$ por

$$\mathbf{E}_{zt}^a = [\mathbf{z}_t^{f_1} - \bar{\mathbf{z}}_t^f, \dots, \mathbf{z}_t^{f_i} - \bar{\mathbf{z}}_t^f] \quad (23)$$

La matriz de error de covarianzas para los estados de predicción $\mathbf{P}_t^f$ se define por su estimador $\hat{\mathbf{P}}_t^f$, $\mathbf{P}_{xz_t}^f$ por $\hat{\mathbf{P}}_{xz_t}^f$ y $\mathbf{P}_{zz_t}^f$ por $\hat{\mathbf{P}}_{zz_t}^f$ respectivamente,

$$\hat{\mathbf{P}}_t^f = \frac{1}{q-1} \mathbf{E}_t^f (\mathbf{E}_t^f)^T \quad \hat{\mathbf{P}}_{xz_t}^f = \frac{1}{q-1} \mathbf{E}_t^f (\mathbf{E}_{z_t}^f)^T \quad \hat{\mathbf{P}}_{zz_t}^f = \frac{1}{q-1} \mathbf{E}_{z_t}^f (\mathbf{E}_{z_t}^f)^T \quad (24)$$

Se puede interpretar que la media del conjunto predicho es el mejor estimador del estado y la dispersión del conjunto de miembros alrededor de la media, como el error entre el mejor estimador y el estado actual. El segundo paso es la etapa de análisis y para obtener las estimaciones de estado, el EnKF realiza un conjunto de ciclos de asimilación de datos en paralelo, donde $i = 1, \dots, q$.

$$x_t^{a_i} = x_t^{f_i} + K_t(y_t^i - h(x_t^{f_i})) \quad (25)$$

Las observaciones perturbadas y_t^i están dadas por

$$y_t^i = y_t + v_t^i \quad (26)$$

Donde v_t^i es una variable aleatoria que proviene de una distribución normal con media cero y varianza R_t . A partir de la expresión clásica para la ganancia de Kalman y las aproximaciones de las covarianzas de error, se establece la ganancia para el conjunto de miembros por:

$$K_t = \hat{P}_{xz_t}^f (\hat{P}_{zz_t}^f)^{-1} \quad (27)$$

El último paso es la predicción del error estadístico en el paso de predicción:

$$x_{t+1}^{f_i} = f(x_t^{a_i}, u_t) + w_t^i \quad (28)$$

Donde w_t^i es una variable aleatoria que proviene de una distribución normal con media cero y varianza Q_t . De forma resumida, a continuación se muestra la metodología del EnKF para el proceso de análisis y predicción.

i) Proceso de Análisis

$$K_t = \hat{P}_{xz_t}^f (\hat{P}_{zz_t}^f)^{-1} \quad (29)$$

$$x_t^{a_i} = x_t^{f_i} + K_t(y_t^i - h(x_t^{f_i}))$$

$$\bar{x}_t^a = \frac{1}{q} \sum_{i=1}^q x_t^{a_i}$$

ii) Proceso de Predicción

$$x_{t+1}^{f_i} = f(x_t^{a_i}, u_t) + w_t^i \quad (30)$$

$$\bar{x}_{t+1}^f = \frac{1}{q} \sum_{i=1}^q x_{t+1}^{f_i}$$

$$E_t^f = [x_t^{f_1} - \bar{x}_t^f, \dots, x_t^{f_q} - \bar{x}_t^f]$$

$$E_{z_t}^a = [z_t^{f_1} - \bar{z}_t^f, \dots, z_t^{f_q} - \bar{z}_t^f]$$

$$\hat{P}_{xz_t}^f = \frac{1}{q-1} E_t^f (E_{z_t}^f)^T \quad \hat{P}_{zz_t}^f = \frac{1}{q-1} E_{z_t}^f (E_{z_t}^f)^T$$

De forma esquemática, el proceso del EnKF puede ser representado de la siguiente manera.

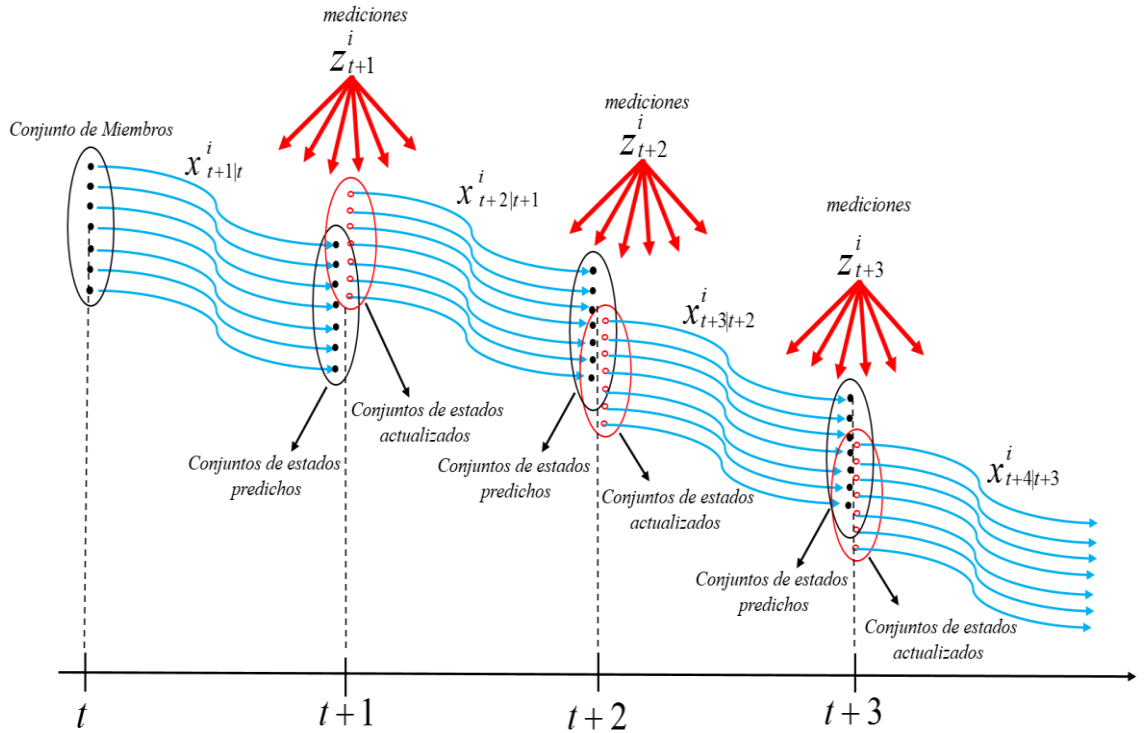


Figura 5. Esquema operativo del EnKF.

3.6. Modelos Autorregresivos y filtros de Kalman para el pronóstico de caudales

Un modelo autorregresivo lineal es una de las primeras aproximaciones para muchos procesos naturales. La aplicación del enfoque estocástico a escala de tiempos continuos o de corto registro para caudales ha sido limitada, debido a las complejas relaciones no lineales que caracterizan a los procesos de lluvia escurrimiento en esas escalas temporales. Los primeros intentos en modelar caudales a nivel horario y diario se basaron en modelos autorregresivos (AR) tras una normalización y transformación (Rajagopalan *et al.*, 2010).

Los modelos autorregresivos han sido ampliamente usados en la hidrología y los recursos hídricos desde inicios de los años sesenta del siglo XX, para la modelación anual y periódica en series de tiempo de tipo hidrológico. La aplicación de estos modelos han sido atractivos en la hidrología principalmente porque la forma autorregresiva tiene dependencia del tiempo y por la facilidad para ser usados (Salas *et al.*, 1980; Box *et al.*, 2013).

3.6.1. Modelos Autorregresivos (AR)

Un modelo estocástico que puede ser extremadamente útil en la representación de algunas series que pueden ocurrir, es un modelo autorregresivo. En este tipo de modelos, el valor actual del proceso se expresa como una serie de agregados finitos lineales de valores anteriores y un conjunto de valores aleatorios (Box *et al.*, 2013).

$$x_t = \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \alpha_3 x_{t-3} \dots + \alpha_{na} x_{t-na} + e(t) \quad (31)$$
$$x_t = \sum_{i=1}^{na} \alpha_i x_{t-i} + e(t)$$

El orden (p) de un proceso autorregresivo se establece de acuerdo al número de términos que establece la ecuación (31). Los modelos AR son de los modelos más populares en series de tiempo para simulación y predicción en hidrología. Este modelo es comúnmente llamado Thomas-Fiering y algunos autores lo expresan en términos de los momentos del proceso aleatorio y el término aleatorio se define como normalmente distribuido con media cero y varianza σ^2 (Bras y Rodríguez-Iturbe, 1985).

3.6.2. Modelos Autorregresivos con Entradas Exógenas (ARX)

Este tipo de modelo permite tener una entrada exógena que puede mejorar la predicción. La estructura que define este tipo de modelos es la siguiente, expresada en diferencias lineales:

$$x_t + \alpha_1 x_{t-1} + \alpha_2 x_{t-2} + \dots + \alpha_{na} x_{t-na} = \beta_1 u_{t-nk} + \beta_2 u_{t-nk-1} + \dots + \beta_{nb} u_{t-nk-nb+1} + e(t)$$

Comúnmente, en hidrología este modelo es conocido a través de la siguiente notación:

$$x_t = \sum_{i=1}^{na} \alpha_i x_{t-i} + \sum_{j=1}^{nb} \beta_j u_{t-j} + e(t) \quad (32)$$

Donde:

x_t es la salida del modelo ARX para $t = t - 1, t = t - na$.

u_{t-j} es la entrada exógena para el modelo ARX para $t = t - nk, t = t - nk - 1, t = t - nk - nb + 1$.

$e(t)$ es el ruido blanco o gaussiano asociado con la salida.

De la ecuación (32), el índice na hacen referencia al número de parámetros autorregresivos para la salida y nb hace referencia al número de parámetros para el vector de entradas exógenas y nk expresa el número de retrasos de la entrada exógena con respecto a la salida.

3.6.3. Formulación en espacio de estados del modelo ARX

Conocida la estructura general de un modelo ARX, la formulación en espacio de estados se convierte en una herramienta útil para utilizar el filtro de Kalman en cualquiera de sus versiones. De acuerdo con esta formulación, la dinámica de un sistema es representada por ecuaciones diferenciales de primer orden, permitiendo una representación más sencilla y proporcionando descripciones estadísticas del comportamiento de los sistemas (Gelb, 1974). La representación de un sistema dinámico puede ser de la siguiente manera:



Figura 6. Sistema discreto para única entrada y única salida.

La formulación para dicho sistema es:

$$\begin{aligned} x_{t+1} &= f(x_t, p_t, t) \\ z_t &= g(x_t, p_t, t) \end{aligned} \tag{33}$$

Los modelos en espacio de estados son esencialmente una notación conveniente para abordar el manejo de un amplio rango de modelos de series de tiempo. En la estimación y control de sistemas, esta metodología se basa en modelos estocásticos, dado el supuesto de la naturaleza errónea de las mediciones (Ramírez, 2003).

De acuerdo a la estructura del modelo ARX en la ecuación (32), éste puede ser escrito en notación de espacio de estados tal que cada parte que conforma A, B, H deben cumplir con los tamaños matriciales para poder operar bajo dicha formulación.

$$\begin{matrix} x_{t+1} \\ (nx1) \end{matrix} = \begin{matrix} A_t \\ (nxn) \end{matrix} \begin{matrix} x_t \\ (nx1) \end{matrix} + \begin{matrix} B_t \\ (nxk) \end{matrix} \begin{matrix} u_t \\ (kx1) \end{matrix} + \begin{matrix} w(t) \\ (nx1) \end{matrix} \quad (34)$$

$$\begin{matrix} z_{t+1} \\ (mx1) \end{matrix} = \begin{matrix} H_t \\ (mxn) \end{matrix} \begin{matrix} x_t \\ (nx1) \end{matrix} + \begin{matrix} v(t) \\ (mx1) \end{matrix}$$

Donde:

A: representa la matriz de parámetros α_i de tamaño (n x n), para la serie de caudales registrados en la estación hidrométrica seleccionada.

B: representa la matriz de parámetros exógenos β_j de tamaño (n x k), para la serie de precipitación media registrada en la cuenca.

u: representa el vector que contiene los valores de precipitación registrada en la cuenca de tamaño (k x 1).

De esta forma, el modelo ARX se ha escrito en espacio de estados facilitando la aplicación del filtro de Kalman, para obtener las predicciones de caudal en el tiempo (t + 1) a partir de la serie de caudales registrada en el tiempo (t) incluyendo la serie de precipitación media en la cuenca como un componente de entrada exógeno al modelo.

3.6.4. Filtros de Kalman en el pronóstico de caudales en cuencas hidrográficas

El filtro de Kalman desde su aparición en 1960, ha tenido múltiples aplicaciones en diferentes áreas de la ciencia. Para el caso de la hidrología, Hino (1973) publicó una de las primeras aplicaciones del filtro de Kalman para recursos hídricos. El problema planteado por Hino fue identificar las ordenadas del hidrograma unitario instantáneo de forma secuencial con base en puntos disponibles con información en la cuenca.

Según Valdés *et al.* (1980), la aplicación de modelos de caja negra basados en la teoría de análisis de sistemas para la simulación de series de caudales, les permitió identificar la función de respuesta a partir del hidrograma unitario instantáneo, con el fin de hacer predicciones en tiempo real en dos cuencas ubicadas en Venezuela.

Con el objetivo de mejorar los trabajos realizados con anterioridad, Husain (1985) evaluó la aplicación del filtro de Kalman en la predicción de caudales para la cuenca Carnation Creek ubicada en la región de Columbia Británica en Canadá. Este autor utilizó el índice de precipitación antecedente y las mediciones de caudal en las estaciones hidrométricas junto con registros de precipitación y escurrimiento obtenidos de las estaciones de la zona de estudio y determinó que la predicción hecha solamente con precipitación y escurrimiento generó un 85% de aproximación. Por el contrario, teniendo en cuenta la precipitación antecedente, la misma predicción proporcionó mejores resultados de hasta el 95 % de aproximación.

En estudios realizados por Awwad y Valdes (1992), utilizaron un modelo estocástico lineal ARMAX expresado en formulación de espacio de estados con el objetivo de realizar pronósticos de caudal en la cuenca del río Fraser en Canadá. Este trabajo presentó dos enfoques de pronósticos basados en el algoritmo de Bergman y Delleur (1985) con el objetivo de estimar de forma on-line los parámetros del modelo.

Awwad *et al.* (1994) retomaron el trabajo de Awwad y Valdes (1992), implementando la metodología en el río Han de Corea del Sur. En este caso, el modelo ARMAX tuvo como variables exógenas la precipitación y la humedad del suelo, para predecir caudales en 17 puntos de la cuenca. Para este caso, los coeficientes del modelo, los parámetros y las estadísticas del ruido, fueron actualizados en forma on-line utilizando un filtrado paralelo. Los resultados obtenidos indican muy buenos resultados debido a la actualización dinámica del sistema, ya que las condiciones tienden a cambiar con el paso del tiempo.

Ahsan y O'Connor (1994) argumentaron que cuando el modelo de predicción de caudal asumido correspondía a un modelo ARMA y los datos de caudal se consideraron libres de error en la medición, el cuadrado medio del error predicho que fue obtenido usando el método de Box-Jenkins para series de tiempo, fue idéntico a los resultados obtenidos utilizando el filtro de Kalman.

Durante el siglo XXI, el filtro de Kalman y sus aplicaciones en la hidrología, se han ido desarrollado en conjunto con modelos matemáticos programados en computadoras que han permitido simular muchos de los procesos del ciclo hidrológico.

Scherider *et al.* (2001) recurrieron al filtro de Kalman para la predicción en tiempo real basada en el modelo lluvia-escorrentía IHACRES, calibrado en diez subcuencas pertenecientes a la cuenca del alto Murray en Australia. El paso predictivo del algoritmo utilizó el módulo lineal del modelo IHACRES, con la entrada de la lluvia efectiva obtenida desde el módulo no lineal y el paso correctivo ajustó la estimación del caudal en función al error entre la predicción y las mediciones de caudal. Las conclusiones reportan que la integración del modelo IHACRES con el algoritmo de Kalman, mejoraron la eficiencia estadística con valores entre 0.839 a 0.967 con errores relativamente bajos entre 4% y 11% durante todo el periodo de observación para cada uno de las diez cuencas consideradas.

Otro de los trabajos reportados con una metodología similar fue el realizado por Madsen *et al.* (2003) al manipular la herramienta de pronóstico de avenidas del modelo MIKE 11 en conjunto con el filtro de Kalman ensemble (EnKF), como algoritmo actualizador del modelo. El algoritmo utilizó las mediciones disponibles en las estaciones de aforo ubicadas en la cuenca del río Sesia para una superficie de 3500 km², localizada en la región de Piamonte al noreste de Italia. Los resultados obtenidos establecen que la aplicación del procedimiento de actualización a través del filtro de Kalman, son significativamente mejores para un pronóstico de hasta 48 horas, en comparación con los pronósticos sin actualizar.

Aubert *et al.* (2003), desarrollaron en su trabajo una mejora para las herramientas hidrológicas de uso común en el contexto de la predicción operativa mediante la introducción de los datos de humedad del suelo en el modelado de caudales. Un procedimiento de asimilación secuencial, basado en un filtro de Kalman extendido, se desarrolló y se acopló a un modelo de precipitación-escorrentía conceptual agrupado. Estos autores concluyeron que la asimilación combinada de humedad del suelo y caudales fue la más adecuada en todo el período de predicción.

Kim *et al.* (2004) emplearon el filtro de Kalman de conjuntos (EnKF) con el modelo CDRMV3 desarrollado en la universidad de Kyoto para predecir caudales en tiempo real. El objetivo del EnKF fue actualizar las variables de estado del modelo distribuido. Concluyeron que utilizando el CDRMV3 junto con el filtro de Kalman mejoraron los resultados que utilizando el CDRMV3 sin el filtro de Kalman comparados en términos RMSE para los hidrogramas calculados.

Moradkhani *et al.* (2005), trabajaron el enfoque denominado doble estimación de parámetros de estado con base en el filtro de Kalman de Conjuntos (EnKF) para la estimación secuencial de parámetros y las variables de estado del modelo hidrológico HyMOD. Concluyeron que el pronóstico realizado para caudales un día hacia adelante fue muy consistente con los valores observados. De igual forma, resaltan la metodología como una técnica adecuada para modelos no lineales y junto con la técnica del kernel smoothing, resulta ser un algoritmo robusto y eficaz que puede hacer frente a las incertidumbres de los parámetros de entrada y salida de forma apropiada.

Neal *et al.* (2007) utilizaron un modelo de simulación hidráulico unidimensional actualizado con datos en tiempo real a través del EnKF. El objetivo fue pronosticar el nivel del agua con mayor precisión mediante la actualización de datos para el modelo, comparado con aquellos sin actualizar. Estos autores concluyeron que con la implementación de la técnica, el EnKF aumentó la exactitud del pronóstico entre el 50 y 70 %.

Méndez (2007), desarrolló un modelo de pronóstico de caudales en tiempo real que implementó el Filtro Kalman como herramienta de optimización de la predicción que realizó con un modelo lineal determinista, y después con un modelo no lineal, tomando como referencia la estación hidrométrica Puente Balseadero sobre el río Magdalena en el Departamento del Huila, Colombia. Los resultados mostraron una clara tendencia del modelo lineal determinista a subestimar los caudales simulados. Esta condición puede volverse crítica cuando el pronóstico subestima la magnitud de una creciente que está por suceder, manteniendo un riesgo latente.

El filtro de Kalman también ha sido trabajado en conjunto con el modelo SWAT (Soil and Water Assessment Tool), tal y como lo reportaron Xie y Zhang (2010) quienes evaluaron la combinación del filtro de Kalman de Conjuntos (EnKF) con éste modelo de base física, para entender la capacidad de estimación de los parámetros y estados. En este estudio se concluyó que mediante la asimilación de datos de escurrimiento, el EnKF puede actualizar eficazmente los estados hidrológicos y mejorar progresivamente la estimación de parámetros en el modelo SWAT.

Lü *et al.* (2013) emplearon el modelo lluvia-escorrentamiento Xinanjian para evaluar la sensibilidad de los parámetros del modelo y la correlación entre variables de estado con el caudal de salida utilizando 3 variantes del filtro de Kalman de Conjuntos (EnKF) para la asimilación de datos en el río Luo en China. Los resultados obtenidos indican que cuando los valores de los parámetros son determinados por el algoritmo de optimización utilizando 10 años de registro, el doble filtro de Kalman de Conjuntos (2EnKF) mejora notablemente los resultados simulados.

Morales *et al.* (2014), reportaron el uso del filtro de Kalman discreto para realizar el pronóstico de avenidas a corto plazo en la subcuenca inferior de la Presa Angel Albino Corso o Peñitas, ubicada en el estado de Chiapas, México. Los resultados obtenidos al considerar la función de respuesta de la cuenca como el hidrograma unitario instantáneo son aceptables, resaltando la importancia de mejorar la instrumentación en la cuenca para mejorar los pronósticos.

En los últimos años, los trabajos sobre predicción hidrológica se han enfocado en la estimación óptima de los parámetros de estado de los modelos utilizados a través de la asimilación de datos. Moradkhani *et al.* (2012) mencionan que la estimación de parámetros de estado de manera simultánea utilizando la asimilación de datos secuencial, es ideal para manejar el estrecho acoplamiento entre los estados del modelo y los parámetros relevantes y permite que los parámetros cambien dinámicamente con el tiempo, especialmente cuando algunos cambios funcionales en el sistema de interés se experimentan o se prevén.

En este contexto, Shi *et al.* (2014), presentaron la estimación de múltiples parámetros usando observaciones multivariantes a través del (EnKF) para un modelo hidrológico de base física. Un sistema de asimilación de datos fue desarrollado para el modelo hidrológico de base física (Flux-PIHM) mediante la incorporación del EnKF en el modelo de estimación de parámetros y estados del sistema. Los resultados muestran que todos los valores de aquellos parámetros estimados son muy cercanos a los valores verdaderos dentro del rango de incertidumbre estimada. Los autores resaltan que el sistema de asimilación de datos del EnKF, puede ser utilizado para proporcionar una guía en el diseño de sistemas de observación y es prometedor para establecer las probabilidades de inundación en tiempo real y pronóstico de sequías.

Para Samuel *et al.* (2014) la asimilación de datos ha sido ampliamente utilizado en los modelos hidrológicos para mejorar el estado del modelo y las estimaciones de caudales posteriores. Sin embargo, para las observaciones de estados pobres o inexistentes, la estimación del estado hidrológico puede causar problemas, dando lugar a cambios sin exactitud en los caudales actualizados.

Estos investigadores aplicaron tres métodos del filtro de Kalman de Conjuntos (EnKF) con estimación doble de parámetros de estado para los siguientes casos: asimilación de caudales, la asimilación de humedad del suelo y la asimilación combinada (humedad del suelo y el caudal) mediante la utilización del modelo sacramento (SAC-SMA). Concluyeron que la asimilación combinada de caudal y humedad del suelo proporciona pronósticos más exactos de ambos humedad del suelo y los caudales, en particular para tiempos cortos.

4. MATERIALES Y MÉTODOS

4.1. Cuenca 1: Río Turbio

La cuenca del río Turbio pertenece a la región hidrológica No. 12 Lerma-Santiago, ubicada en los estados de Guanajuato y Jalisco, entre los $20^{\circ}32'42''$ y los $21^{\circ}21'18''$ de Latitud Norte y entre los $101^{\circ}27'82''$ y los $102^{\circ}17'57''$ de Longitud Oeste. La cuenca del Turbio comprende gran parte de la llanura fértil del Bajío Guanajuatense, caracterizada por contar con obras de irrigación de gran magnitud que han convertido a esta región en una de las zonas agrícolas más importantes del país.

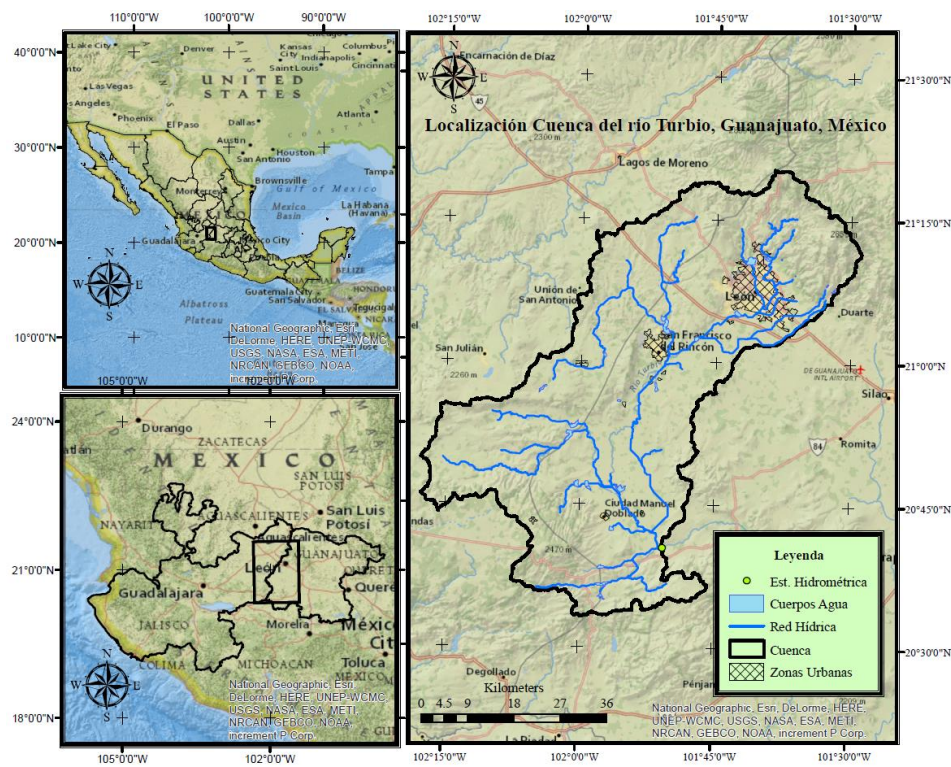


Figura 7. Localización de la cuenca del río Turbio.

La delimitación de la cuenca en mención, se hizo hasta la estación hidrométrica las Adjuntas ubicada en los 20°40' 32" de Latitud Norte y 101°50' 40" de Longitud Oeste, que es operada por la red nacional de estaciones hidrométricas de la Comisión Nacional del Agua. El área total de la cuenca es de 3322 km² de los cuales 58.8% pertenece al estado de Guanajuato y el 41.2% corresponde al estado de Jalisco. El parteaguas de la cuenca tiene su máxima elevación a los 2670 m y el punto de salida referenciado por la estación hidrométrica se encuentra a los 1723 m. La pendiente media para la cuenca del río Turbio es del 11%, el cauce principal presenta un desnivel de 323 m con un tiempo de concentración para la cuenca de 21 h.

4.1.1. Climatología

El clima prevaleciente en la zona según la clasificación de Köppen, modificado por García, (1981) es el siguiente. En la parte Norte, sobre la sierra de Cuatralba el clima es C (w0), templado subhúmedo con temperaturas medias anuales que van de los 12°C a los 18°C. En la parte central, desde la ciudad de León hasta la ciudad Manuel Doblado en el estado de Guanajuato, prevalece el clima A(C) (w0), semicálido subhúmedo con temperaturas medias anuales mayores a 18°C. Al Suroeste, en dirección a San Julián y Arandas en el estado de Jalisco, el clima es C (w1), templado subhúmedo con temperaturas medias anuales entre los 12°C y 18°C.

4.1.2. Precipitación

La precipitación media anual registrada en las subregiones Lerma y Santiago, a la cual pertenece la cuenca del río Turbio entre los años 1945-1997 en diferentes puntos de la

región, varía entre los 500 y 900 mm anuales (según lecturas de estaciones climatológicas), con un promedio general de 705 mm (CONAGUA, 2003).

En la cuenca del río Turbio, la precipitación media anual para el periodo 1981-2010 fue de 659 mm obtenida mediante la ponderación por polígonos de Thiessen, con un periodo de lluvias que abarca desde el mes de junio hasta el mes de octubre, concentrándose en este periodo un total de 586 mm, es decir, un 88.8% del total de las precipitaciones anuales, caracterizadas por ser intensas y de corta duración.

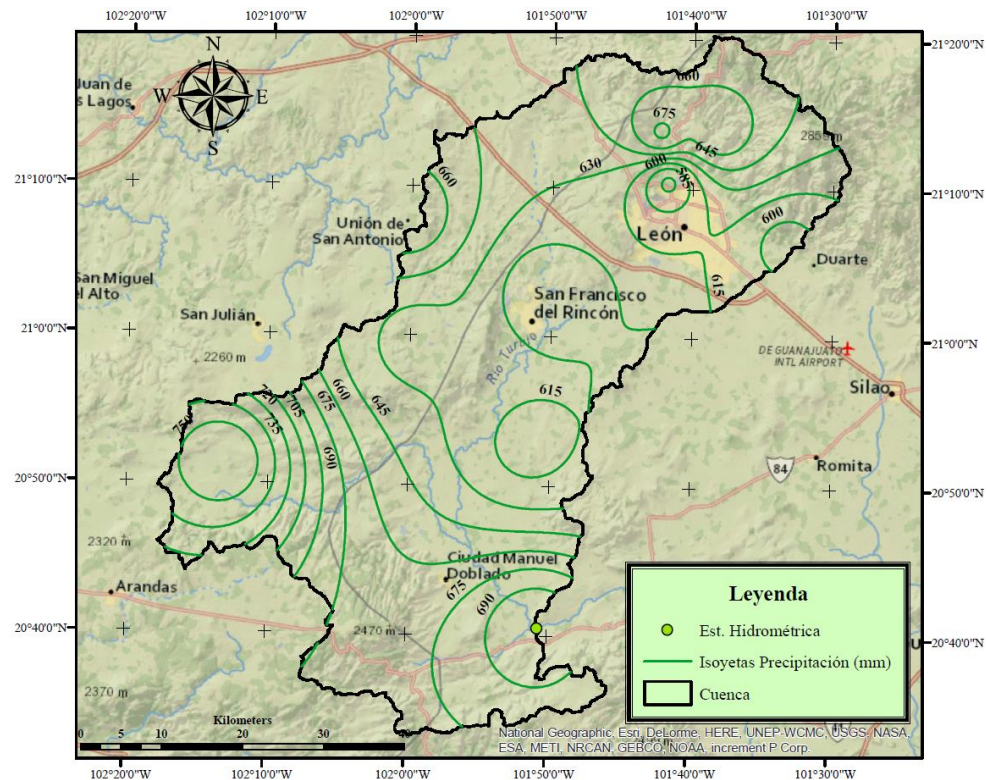


Figura 8. Isoyetas de precipitación en la cuenca del río Turbio.

Como dato comparativo, la precipitación media anual en México reportada por CONAGUA (2012) para el periodo 1971-2000 fue de 760 mm, ligeramente superior a la registrada en el periodo 1981-2010 en la cuenca del río Turbio. La figura 9, resume los datos comparativos por cuenca, región hidrológica y país.

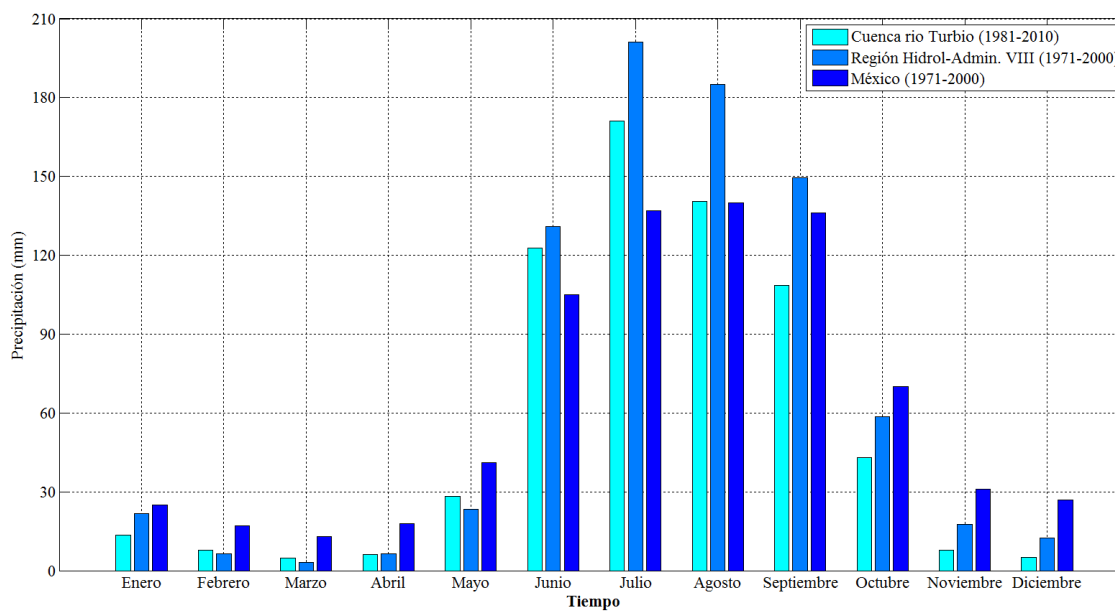


Figura 9. Distribución comparativa de la precipitación media mensual en la cuenca río Turbio.

4.1.3. Red Hidrográfica y de Presas

El río Turbio es un afluente derecho del río Lerma figurando entre los más importantes de la cuenca alta de éste río; tiene su origen en la sierra de Cuatralba a unos 20 kilómetros al Norte de la ciudad de León, con el nombre de arroyo La Patiña, uniéndose por la margen izquierda los arroyos Los Castillos, Ibarrilla, Ojo de Agua de los Reyes, Alfaro y El Salto; por la margen derecha confluye el arroyo Mariches, todos ellos en las inmediaciones por la parte este de la Ciudad de León. En este tramo el colector principal se conoce como río de los Gómez. El río continúa por los municipios de Purísima del Rincón y Manuel Doblado. A la altura de la localidad de las Adjuntas, se encuentra una estación hidrométrica que se encarga de medir los escurrimientos generados por el afluente. En el cuadro 2, se resumen los cuantiles para los gastos máximos registrados en la estación hidrométrica Las Adjuntas según la Comisión Nacional de Agua y reportados por Protección Civil Guanajuato (2012).

Cuadro 2. Periodos de retorno para gastos máximos en la estación hidrométrica Las Adjuntas.

Periodo de Retorno (Años)	Gasto (m ³ /s)
2	44.35
5	84.60
10	111.24
20	136.8
50	169.88
100	194.67
500	251.96
1000	276.59

Fuente: Protección Civil Guanajuato, 2012.

El río Turbio confluye con el río Lerma en una planicie de escasa pendiente, propiciando que en la ocurrencia de avenidas de consideración del Lerma, el agua de éste ingrese hacia el cauce del Turbio, ocasionando un retardo en el desalojo de avenidas de dicho río provocando desbordamientos e inundaciones en las planicies aledañas (López, 2008).

En lo que respecta a cuerpos de agua, existen un numero considerado de presas y/o reservorios que surten de agua a los cultivos que se desarrollan en la región del Bajío Guanajuatense y a su vez, sirven para el control de avenidas ocasionadas por las lluvias intensas en los periodos de mayor precipitación. Las presas más importantes dentro de la zona son El Palote y El Barrial. La primera, se encuentra en la ciudad de León y su propósito es agua potable y control de avenidas con una capacidad máxima de 17 Mm³ y la segunda se ubica sobre San Francisco del Rincón diseñada para el control de avenidas y riego con una capacidad de 55.26 Mm³. Otras presas que se destacan son Santa Efigenia con capacidad de 42.75 Mm³, Jalpa Vieja con 14.38 Mm³, Ciénaga de Galvanes 11 Mm³, Potrerillos con 15 Mm³ todas con doble propósito de riego y control de avenidas (CONAGUA, 2014b).

En lo que respecta a las estaciones ubicadas en la cuenca, un total de 11 estaciones fueron seleccionadas a lo largo y ancho de la cuenca del río Turbio. En la figura 10, se tiene la localización de las estaciones en toda la cuenca. En el cuadro 3, se resume el total de estaciones utilizadas para obtener los registros de precipitación diaria.

Cuadro 3. Estaciones meteorológicas en la cuenca del río Turbio.

Código Estación	Nombre	Municipio	Estado	Dependencia
11020	El Palote	León	Guanajuato	SMN
11025	Hacienda de Arriba	León	Guanajuato	SMN/CEAG
11035	La Sandia	León	Guanajuato	SMN
11036	Adjuntas	Manuel Doblado	Guanajuato	SMN/CEAG
11045	Media Luna	León	Guanajuato	SMN
11153	La Laborcita	León	Guanajuato	SMN
11157	San Fco. del Rincón	Peñuelas	Guanajuato	SMN
11159	San Fco. del Rincón	Presa el Barrial	Guanajuato	SMN
14123	San Diego Alejandría	San Diego Alejandría	Jalisco	SMN
14157	Unión de San Antonio	Unión de San Antonio	Jalisco	SMN
14369	La Vaquera	Arandas	Jalisco	SMN

Fuente: SMN, 2010.

Basado en la distribución espacial se obtuvieron los factores de ponderación para la lluvia diaria mediante los polígonos de Thiessen, según la figura 11. En el cuadro 4 se resumen los factores utilizados para la ponderación de la lluvia en la cuenca del río Turbio.

Cuadro 4. Factores de Thiessen para la cuenca del río Turbio.

Código Estación	Área km²	Factor Thiessen
11020	227.4	0.07
11025	233.8	0.07
11035	49.1	0.01
11036	600.3	0.18
11045	210.5	0.06
11153	180.6	0.05
11157	413.8	0.06
11159	368.7	0.11
14123	354.2	0.11
14157	191.5	0.12
14369	491.8	0.15

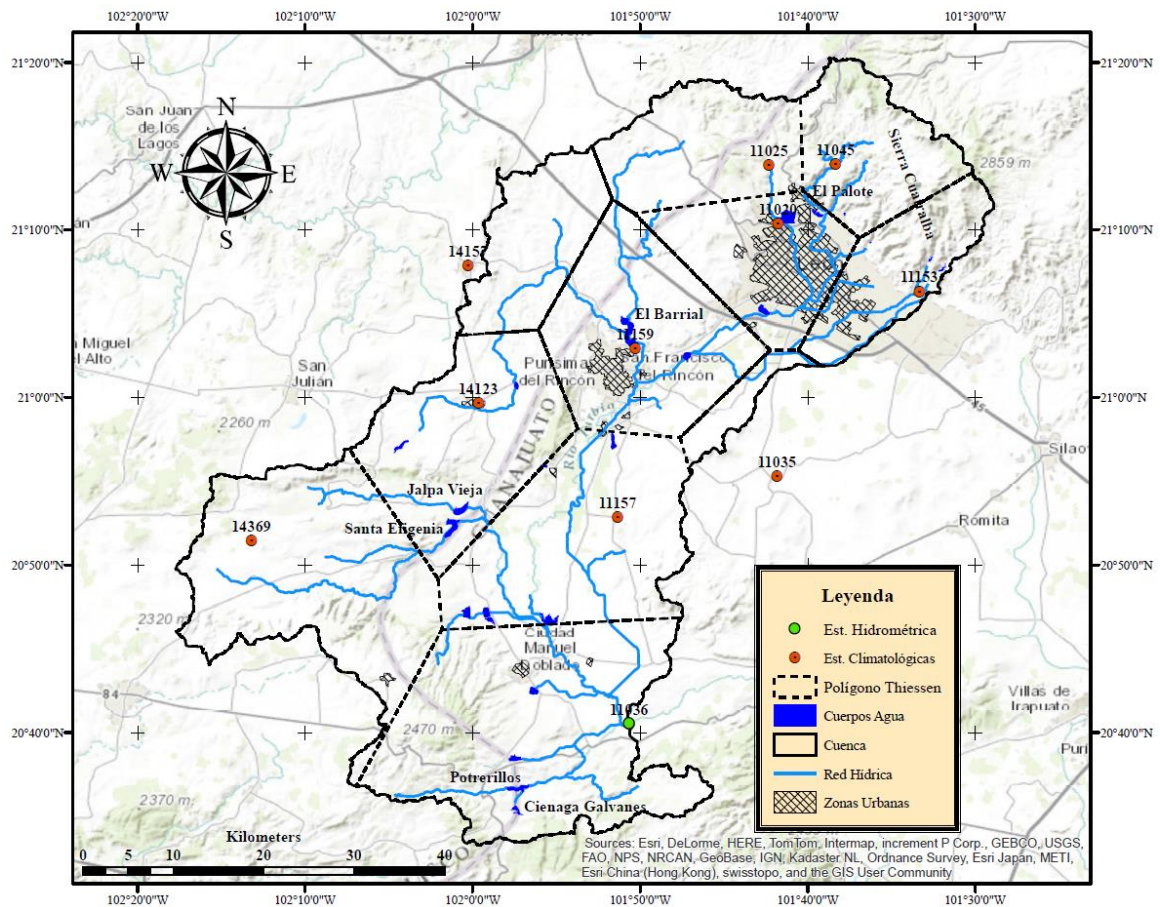


Figura 11. Polígonos de Thiessen para la cuenca del río Turbio.

4.2. Cuenca 2: Río Cazones

La cuenca del río Cazones se encuentra ubicada en la región hidrológica No. 27 denominada Tuxpan-Nautla, entre los 20°0' 00" y los 20°35' 00" de Latitud Norte y entre los 97°20' 00" y los 98°20' 00" de Longitud Oeste, incluyendo a los estados de Puebla, Hidalgo y Veracruz. La cuenca descrita fue delimitada hasta la estación hidrométrica Poza Rica en el Estado de Veracruz, registrando una altitud máxima del parteaguas a los 2750 m y pasando por los 50 m en la población de Poza Rica. El área total de la cuenca es de 1604.6 km² de los cuales 56% pertenece al estado de Puebla, el 13.5% al Estado de Hidalgo y el 30.5% restante al estado de Veracruz.

La cuenca del río Cazonos presenta una pendiente promedio de 27%, un desnivel para el cauce principal de 2100 m y un tiempo de concentración de 16 h.

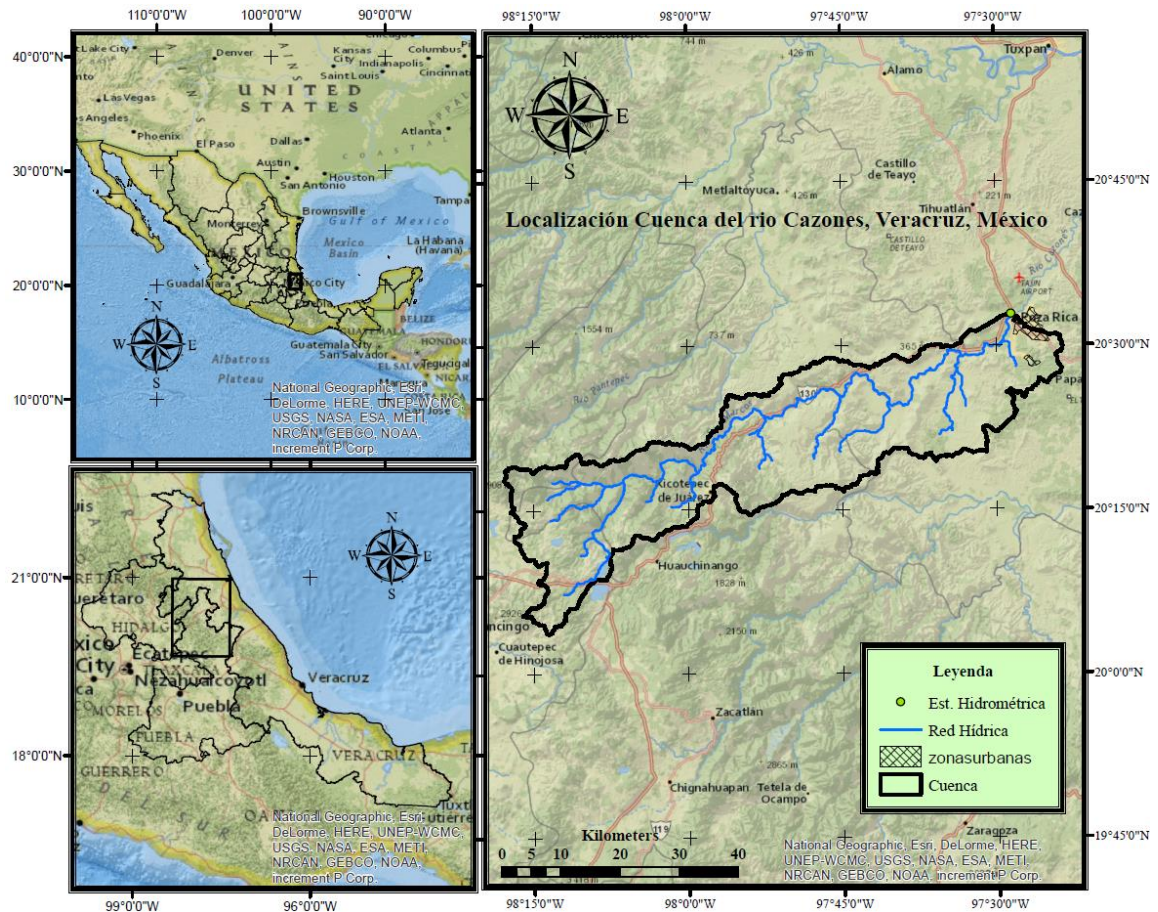


Figura 12. Localización de la cuenca del río Cazonos.

4.2.1. Climatología

El clima es de tipo cálido-húmedo, con una temperatura media anual de 23.6°C y con precipitaciones constantes a lo largo del año. En la zona donde se encuentra localizada la cuenca del río Cazonos, es constantemente afectada por fenómenos meteorológicos extremos los cuales suelen generar grandes cantidades de lluvia que al escurrir sobre suelos saturados producen inundaciones en las partes bajas de la cuenca (INEGI, 2005; Pereyra y Pérez, 2006; Santillán Gutiérrez *et al.* 2013).

4.2.2. Precipitación

Existen dos tipos de masas de aire que originan lluvia y se alternan sobre el estado de Veracruz, la tropical en verano y la polar en invierno. Durante los meses de transición estas masas de aire se van alternando y en ocasiones no hay una situación bien definida (Pereyra y Pérez, 2006). En la cuenca del río Cazones la precipitación media anual es de 1744 mm de acuerdo a las normales climatológicas reportadas en las estaciones consideradas en el estudio para el periodo 1981-2010. Este promedio es 12% superior al registrado por la región hidrológico-administrativa No. X Golfo Centro la cual reporta una precipitación media de 1558 mm en el periodo 1971-2000. La figura 13, resume la distribución de las lluvias a lo largo del año y de forma comparativa por cuenca, región hidrológico-administrativa y nacional.

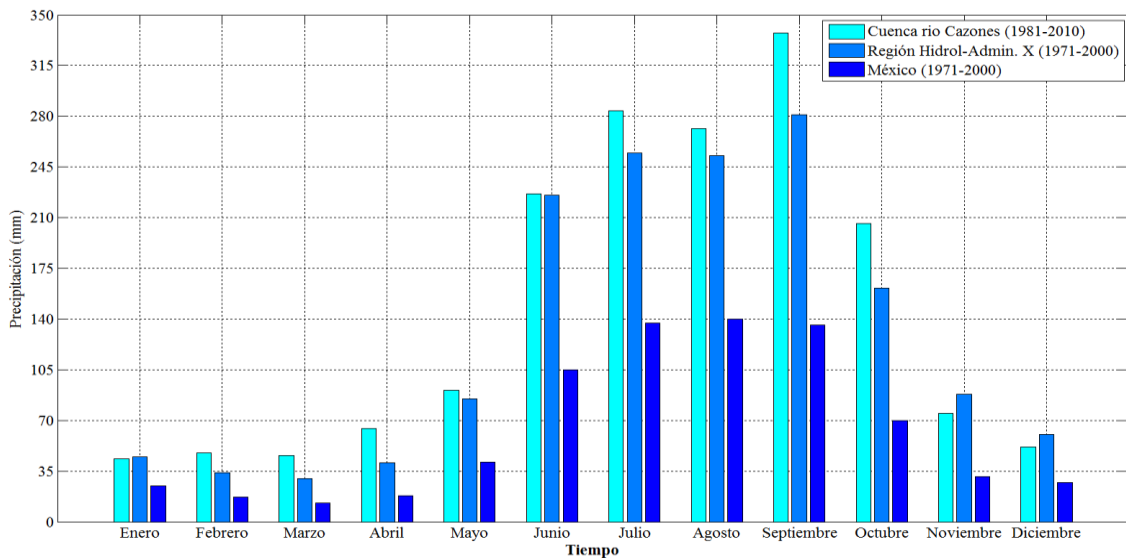


Figura 13. Distribución comparativa de la precipitación media mensual en la cuenca río Cazones.

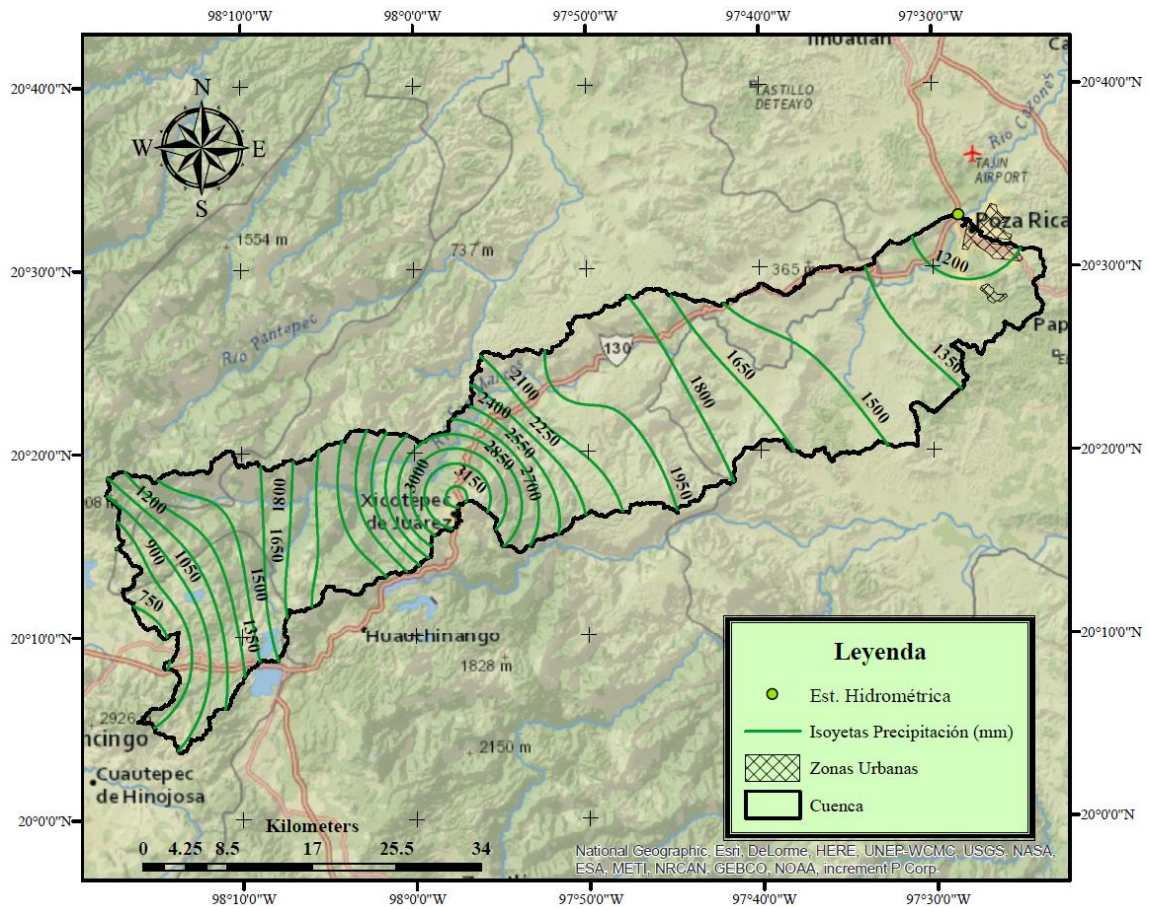


Figura 14. Isoyetas de precipitación para la cuenca del río Cazones.

4.2.3. Red Hidrográfica

Los arroyos formadores de esta corriente nacen y descienden de la sierra de Hidalgo. El parteaguas nace en los límites del río Tulancingo, 10 km al este de la ciudad del mismo nombre a una elevación de 2750 m, dando lugar a la formación del arroyo Chaltecontla, que en sus orígenes se denomina río Los Reyes. Es conocido en su tramo superior como río San Marcos; se interna en Veracruz y se dirige hacia el noreste; en su curso medio e inferior toma el nombre de Cazones hasta su desembocadura en el Golfo de México (INEGI, 2005; Pereyra y Pérez, 2010).

Según reportes de CONAGUA (2012), la cuenca del río Cazones tiene un escurrimiento natural medio superficial de 1712 Mm³.

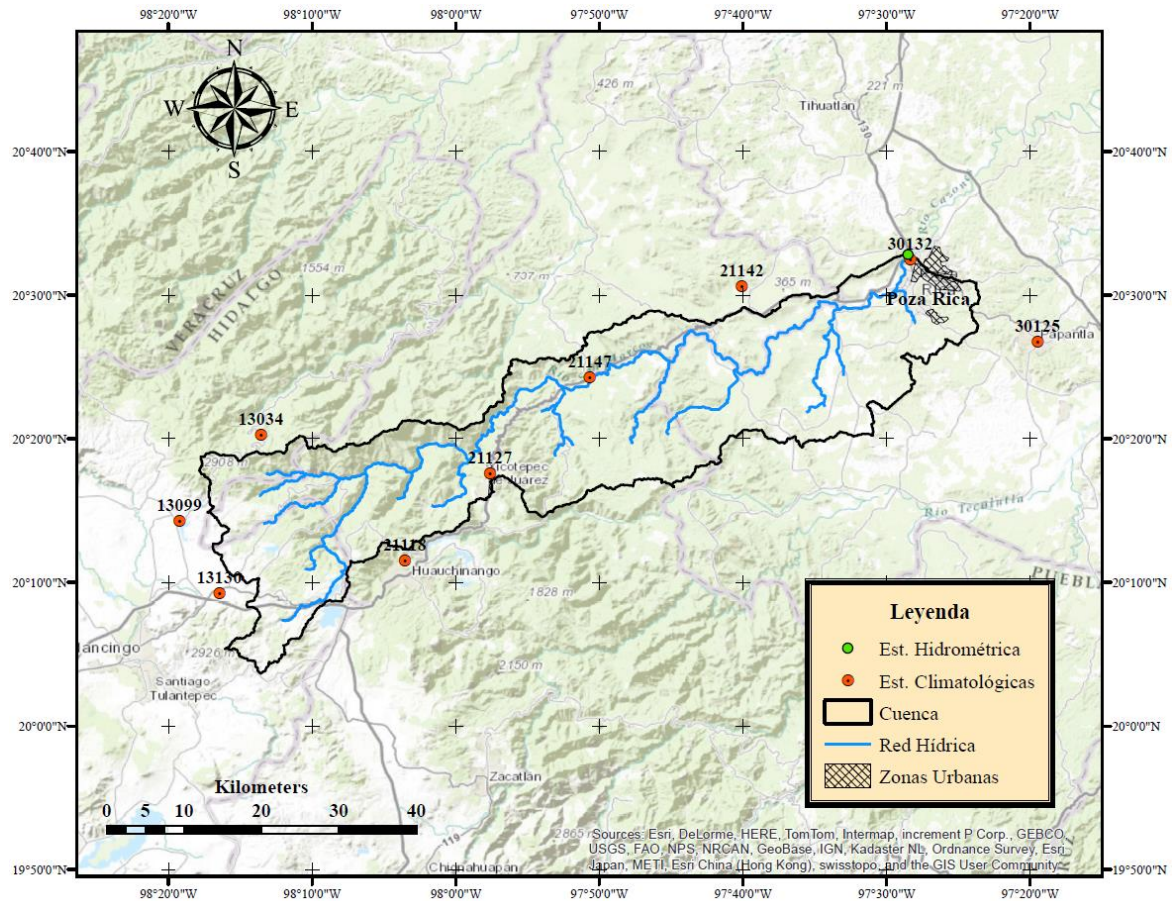


Figura 15. Red hidrográfica en la cuenca del río Cazones.

4.2.4. Red de Estaciones Meteorológicas

Un total de 9 estaciones meteorológicas fueron consideradas para determinar los registros de precipitación en la cuenca del río Cazones, todas pertenecientes y bajo la operación del Servicio Meteorológico Nacional tal como se puede observar en la figura 15.

En el cuadro 5, se resumen las estaciones utilizadas para obtener los registros de precipitación diaria.

Cuadro 5. Estaciones meteorológicas localizadas en la cuenca del río Cazones.

Código Estación	Nombre	Municipio	Estado	Dependencia
13034	Tenango de Doria	Tenango de Doria	Hidalgo	SMN
13099	Metepec	Metepec	Hidalgo	SMN
13130	Santa Ma. Asunción	Tulancingo	Hidalgo	SMN
21118	Huauchinango	Huauchinango	Puebla	SMN
21127	Xicotepec	Xicotepec	Puebla	SMN
21142	Venustiano Carranza	Venustiano Carranza	Puebla	SMN
21147	Apapantilla	Jalpan	Puebla	SMN
30125	Papantla	Papantla	Veracruz	SMN
30132	Poza Rica	Poza Rica	Veracruz	SMN

Fuente: SMN, 2010.

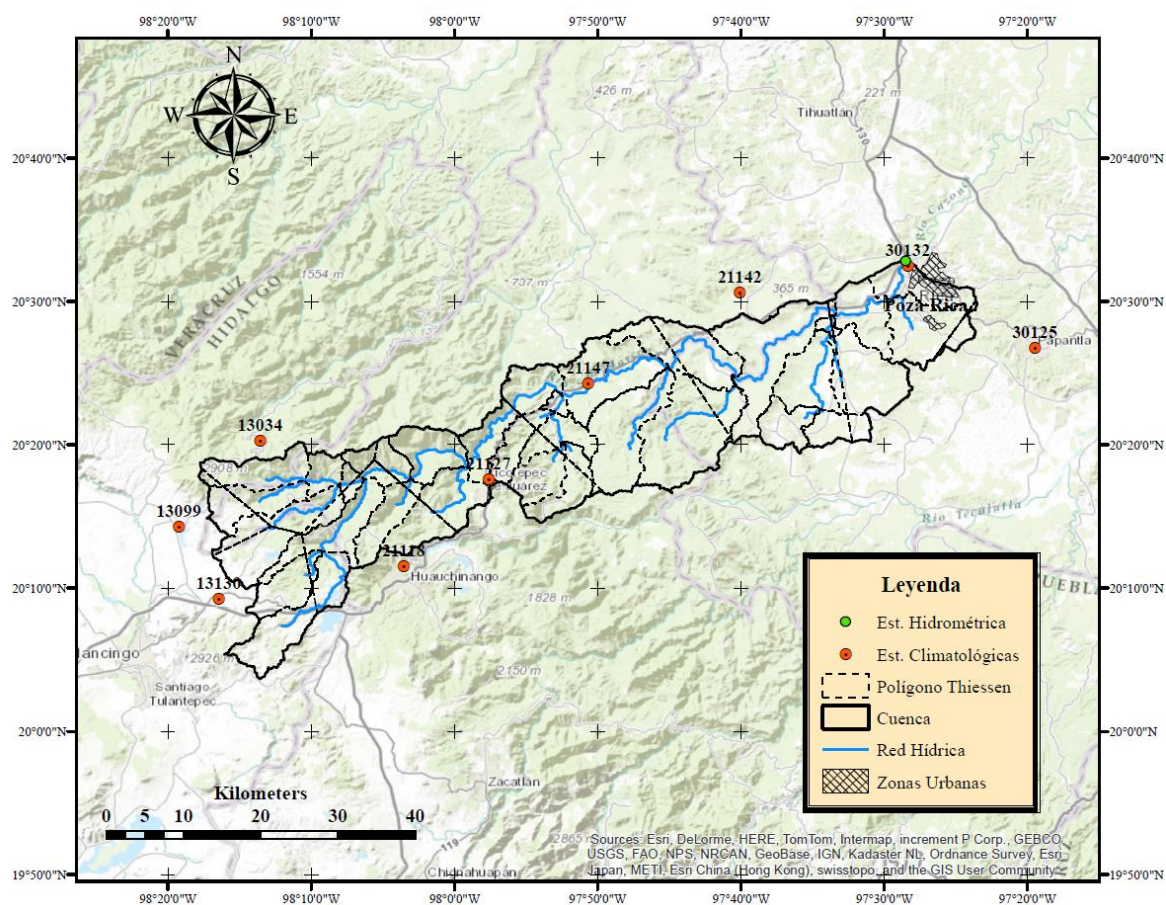


Figura 16. Polígonos de Thiessen para la cuenca del río Cazones.

Basados en la distribución espacial de las estaciones climatológicas, se obtuvo los factores de ponderación para la precipitación media diaria a través los polígonos de Thiessen, como se muestra en la figura 16. A continuación, en el cuadro 6 se resumen los factores utilizados para la ponderación de la precipitación en la cuenca del río Cazonés.

Cuadro 6. Factores de Thiessen para la cuenca río Cazonés.

Código Estación	Área km²	Factor Thiessen
30132	163.0	0.10
13034	129.1	0.08
13099	36.0	0.05
21142	295.4	0.18
13130	136.0	0.08
21118	164.1	0.10
21127	236.4	0.14
21147	374.9	0.23
30125	69.8	0.04

4.3. Metodología

La metodología que será descrita a continuación, fue desarrollada para realizar pronósticos de caudales a corto plazo, en las cuencas del río Turbio y río Cazonas, localizadas en el centro y Golfo de México, respectivamente. Se implementaron 3 metodologías con el objetivo de establecer las diferencias en operatividad, para cada una de las cuencas. En la primera metodología se utilizó únicamente el algoritmo del filtro de Kalman Discreto (DKF). En la segunda, se implementó un modelo autorregresivo con entrada exógena ARX operado bajo el filtro de Kalman Discreto y en la tercera, se utilizó el filtro de Kalman de Conjuntos (EnKF) bajo la misma estructura del modelo ARX.

Las series de datos utilizadas para la cuenca del río Turbio corresponden a los años 2003 y 2004 mientras que para la cuenca del río Cazonas, las series correspondientes fueron para el año 2005 y 2008-2011. Las series en cada cuenca fueron escogidas de acuerdo a la disponibilidad de los datos en las bases de datos BANDAS y CLICOM, respectivamente (CONAGUA, 2014b; SMN, 2010).

4.3.1. Filtro de Kalman Discreto DKF

El desarrollo de esta metodología tiene sus bases en los trabajos realizados por Hino (1973), Valdés *et al.* (1980) y Morales *et al.* (2014). Según Valdés *et al.* (1980), el filtro de Kalman puede ser empleado para identificar la función de respuesta de una cuenca, es decir, a través su hidrograma unitario.

De igual forma, el DKF es utilizado para predecir caudales en tiempo real en una cuenca dada, con registros de corta duración. En este sentido, el caudal a la salida de una cuenca puede ser representado en función de:

- i) Caudales pasados en el mismo sitio.
- ii) Caudales presentes y pasados en sitios aguas arriba de la estación de aforo.
- iii) Precipitaciones presentes y pasadas sobre la cuenca.

Un modelo para representar los caudales a la salida de una cuenca, con base en lo anterior, puede ser descrito como:

$$q(t) = \sum_{i=1}^{n_0} a_i q(t-1) + \sum_{k=1}^m \sum_{j=1}^{n_k} b_{kj} P_k q(t-j) + \delta(t) \quad (35)$$

Donde:

$q(t)$: es el caudal a la salida de la cuenca.

n_0 : es el número de términos autorregresivos.

m : es el número de estaciones de precipitación y de caudales aguas arriba.

n_k : es el número de términos de la estación.

$\delta(t)$: representa el ruido en la medición.

El modelo de caudales descrito anteriormente, puede ser implementado siguiendo la estructura de la ecuación de mediciones en el filtro de Kalman, tal que:

$$q(t) = H(t)X(t) + \delta(t) \quad (36)$$

Donde, $H(t)$: es el vector que contiene las precipitaciones y los caudales en el pasado. Ver representación en la figura 17.

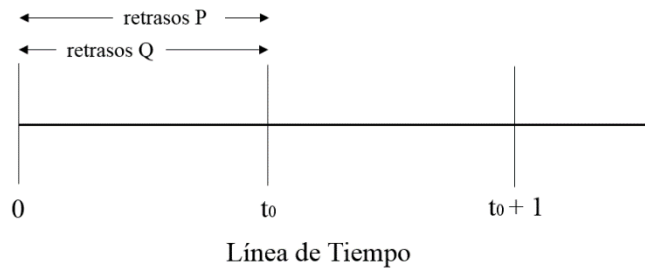


Figura 17. Esquema de conformación del vector $H(t)$.

$X(t)$: es el vector que contiene los elementos de todas las funciones de respuestas en todas las estaciones consideradas.

En el caso de predicción en tiempo real, la función de respuesta puede cambiar entre los periodos de lluvia y sequía o de acuerdo las condiciones climatológicas de la zona. Sin embargo, existen ríos en cuencas de gran tamaño, en el que las variaciones en las descargas no son abruptas sino que ocurren de forma suave y por tanto, las variaciones en la función de respuesta también ocurren de la misma manera.

Siguiendo la estructura del algoritmo del filtro de Kalman Discreto (DKF) de la figura 4, las ecuaciones en las etapas de predicción y corrección se conformaron de la siguiente manera:

- i) Para el estado inicial el valor esperado fue definido tal que $E[\hat{x}] = 0$. La matriz de covarianza del estado inicial se definió como $P_0 = K \cdot I$, donde K es un escalar lo suficientemente grande como para reflejar la incertidumbre de los supuestos iniciales, siendo I la matriz identidad de tamaño $(n \times n)$.

- ii) La matriz A en la ecuación de predicción se consideró invariante en el tiempo para cada relación del estado $(t+1)$ con el que ocurre en el estado (t) .
- iii) No fue considerada la matriz de control representada por el sistema $B * u$ debido a que se emplean los estimadores de estado.
- iv) El vector H en la ecuación de medición se conformó con los datos de precipitación y caudal registrados en la cuenca, presentando retrasos diferentes en caudal y precipitación, dependiendo de la respuesta de la cuenca (ver figura 17).
- v) La matriz que contiene el ruido en la medición fue representada como $\alpha * q(t-1)$ donde α es una constante de proporcionalidad que representa un error constante e igual a una fracción del caudal en el tiempo anterior, propuesta por Valdés *et al.* (1980).

Los registros de precipitación y caudal que son utilizados en la matriz de convolución para el pronóstico en el tiempo $(t+1)$, se encuentran a nivel diario para las dos cuencas descritas anteriormente, a diferencia del trabajo reportado por Morales *et al.* (2014) quienes obtuvieron registros a nivel diario soportados por Comisión Federal de Electricidad en México.

4.3.2. Modelo ARX en conjunto con el Filtro de Kalman Discreto

Se propuso la creación de un modelo autorregresivo con entrada exógena ARX para predecir los caudales medios diarios teniendo en cuenta las series del pasado, obtenidas a partir del registro BANDAS (CONAGUA, 2014b) para cada estación hidrométrica en las cuencas seleccionadas.

Para este caso, el modelo ARX tiene como entrada exógena la precipitación media diaria con el objetivo de mejorar los pronósticos, considerando la lluvia como un pulso sometido a convolución dentro del modelo de caja negra que se ha establecido.

Para crear el modelo ARX se programó una rutina en Matlab[®] utilizando el toolbox de identificación de sistemas y de esta forma, se obtuvo el orden (p) del proceso autorregresivo establecido de acuerdo al número de términos representado en la ecuación (31) y teniendo en cuenta principios de parsimonia. De esta forma, la representación del modelo en formulación de espacio-estado fue necesaria para operar el DKF mediante una rutina de programación y realizar el pronóstico en el tiempo $(t + 1)$ dado el tiempo (t) como lo describe la ecuación (34).

$$\begin{matrix} x_{t+1} \\ (nx1) \end{matrix} = \begin{matrix} A_t \\ (nxn) \end{matrix} \begin{matrix} x_t \\ (nx1) \end{matrix} + \begin{matrix} B_t \\ (nxk) \end{matrix} \begin{matrix} u_t \\ (kx1) \end{matrix} + \begin{matrix} w(t) \\ (nx1) \end{matrix} \quad (34)$$

$$\begin{matrix} z_{t+1} \\ (mx1) \end{matrix} = \begin{matrix} H_t \\ (mxn) \end{matrix} \begin{matrix} x_t \\ (nx1) \end{matrix} + \begin{matrix} v(t) \\ (mx1) \end{matrix}$$

Donde:

A: representa la matriz de parámetros α_i de tamaño $(n \times n)$, para la serie de caudales medios diarios registrados en la estación hidrométrica de la cuenca seleccionada.

B: representa la matriz de parámetros exógenos β_j de tamaño $(n \times k)$, para la serie de precipitación media diaria registrada en la cuenca.

u: representa el vector que contiene los valores de precipitación media diaria registrada en el tiempo (t) para la cuenca, con tamaño $(k \times 1)$.

$w(t)$ y $v(t)$ representan los ruidos gaussianos del proceso y de las mediciones con propiedades tal que $w(t) \sim N(0, Q)$ y $v(t) \sim N(0, R)$.

Con el objetivo de encontrar diferencias entre los pronósticos para L pasos a través del tiempo, es decir, $x_{t+1|t}$, $x_{t+2|t}$, ..., $x_{t+L|t}$ fue ejecutada una rutina en Matlab[®] que utilizó información del estado en el tiempo (t) para avanzar, sin actualizar, L pasos en la línea de tiempo. Una vez llegada a la posición L deseada, se realizó la actualización el primer estado que se había pronosticado. El ciclo fue repetido n veces hasta llegar al final de la serie de datos. El procedimiento es descrito gráficamente a continuación, donde la línea negra es el pronóstico y la roja es la actualización (ver figura 18).

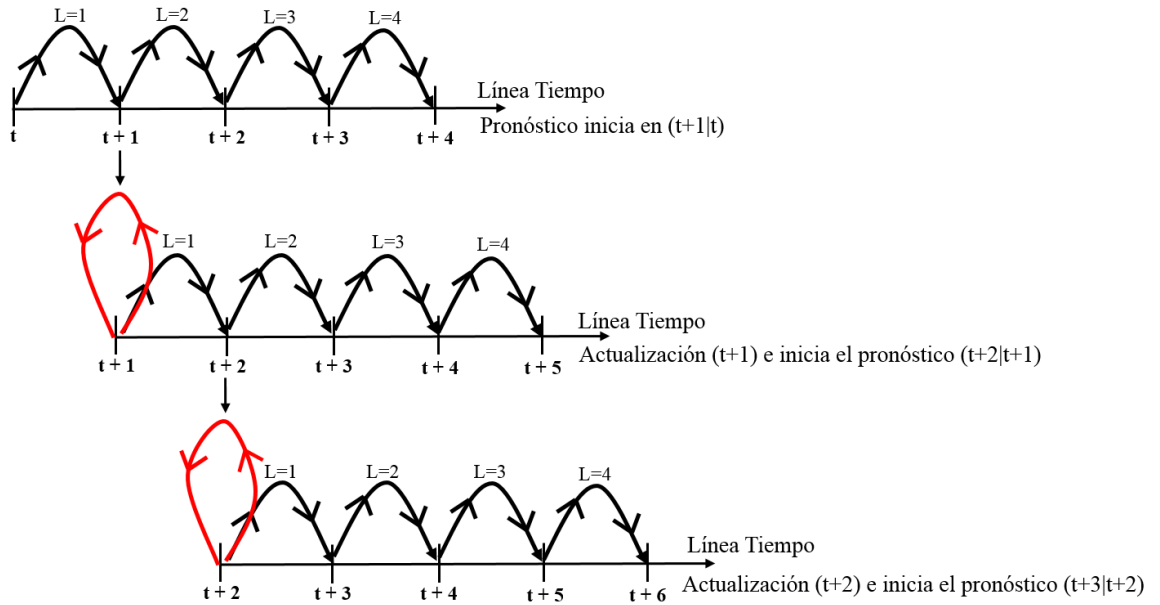


Figura 18. Descripción del pronóstico de caudales con L pasos adelante.

Definidos los pronósticos hacia adelante y teniendo en cuenta que los parámetros del modelo ARX que han sido estimados pueden variar en función de las épocas de máximas precipitaciones y/o de las variaciones abruptas en los caudales, se propuso un período previo para la estimación de los parámetros del modelo ARX antes de iniciar el pronóstico, con el objetivo de obtener parámetros variando en función del tiempo ajustándose a las condiciones dadas por las series de datos (caudal y precipitación).

De esta manera, la implementación del DKF se hace dinámica y con matrices de estado que varían en cada periodo definido, realizando una calibración con el objetivo de disminuir el ruido generado en el proceso representado como $w(t)$ en la ecuación de (34). En la figura 19, se describe el proceso para la calibración y pronósticos por periodos.

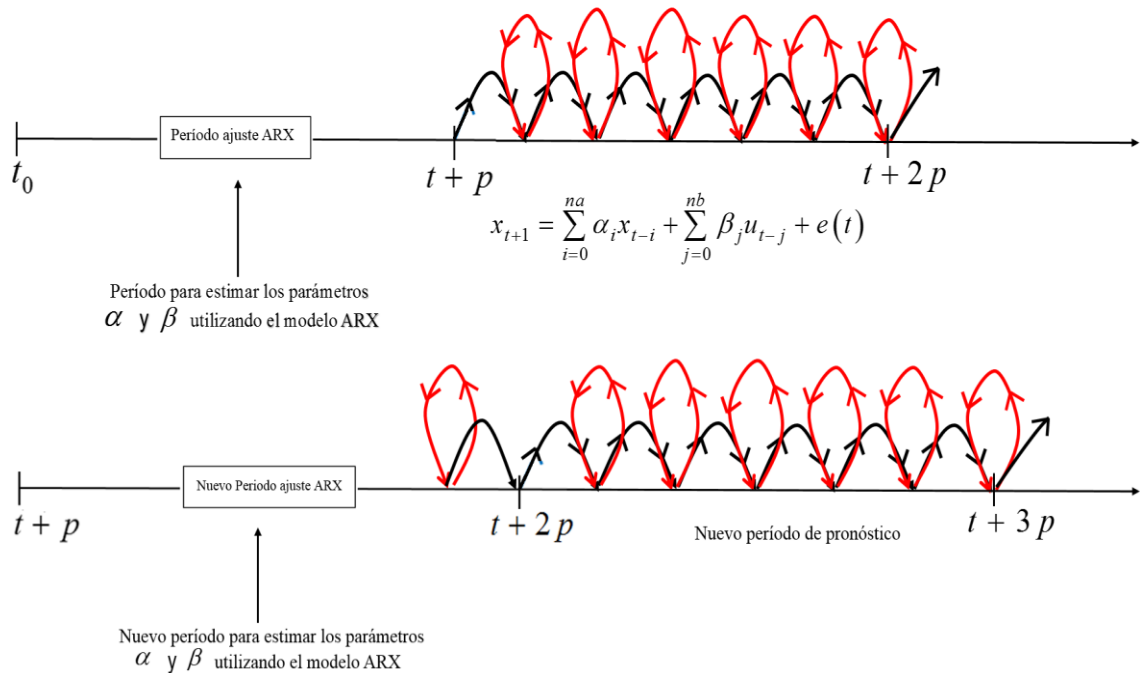


Figura 19. Descripción del proceso de pronóstico de caudales con base en periodos de estimación de parámetros del modelo ARX.

4.3.3. Modelo ARX a través del Filtro de Kalman de Conjuntos EnKF

En esta variante metodológica, se utilizó el esquema propuesto anteriormente para el modelo ARX y su periodo de calibración con la diferencia de realizar los pronósticos mediante el filtro de Kalman de conjuntos (EnKF).

Esta técnica de asimilación de datos fue propuesta para sistemas dinámicos no lineales y como alternativa al filtro de Kalman Extended (EKF) debido a su formulación conceptual simple y de fácil aplicación, ya a que no hay necesidad de derivar un operador lineal tangente o ecuaciones adjuntas, ni integraciones en tiempos previos como lo hace el EKF. A pesar de las consideraciones planteadas anteriormente y de la linealidad existente dada en el modelo ARX, se implementó la rutina en Matlab® como ejercicio académico y evaluar así su respuesta, en comparación con ARX +DKF.

Siguiendo la metodología propuesta por Gillijns *et al.* (2006), fue necesario generar un número aleatorio de miembros o posibles trayectorias, donde la media del conjunto generado se considera como la estimación del estado y el error del modelo se aproxima por la varianza del conjunto generado. De esta manera, se realizaron pronósticos con un total de 100 miembros, como conjunto lo suficientemente grande para generar una media que se aproxime a los valores reales.

Se calcularon las principales estadísticas según Gupta *et al.* (2009) para cada una de las 3 metodologías descritas anteriormente. El error cuadrático medio MSE (Mood *et al.* 1974) y su normalización relacionada, la eficiencia de Nash-Sutcliffe (Nash y Sutcliffe, 1970) fueron los dos criterios a tener en cuenta, por ser los más utilizados para la calibración y evaluación de modelos hidrológicos con los datos observados.

De igual forma, a partir de los errores de predicción obtenidos en cada intervalo de tiempo tanto para el DKF como para el ARX+DKF, se calculó un intervalo de predicción al 95% de probabilidad para las series generadas según Chatfield (2004), para conocer el rango de probabilidad de que los valores reales oscilen dentro del mismo, como fundamento para el enfoque de pronósticos de caudales en tiempo real.

Para comprobar los supuestos de la teoría general del filtro de Kalman, se realizaron ajustes a funciones de distribución de probabilidad para los errores provenientes de los datos observados y pronosticados para las series evaluadas empleando la metodología de ARX+DKF.

5. RESULTADOS Y DISCUSIÓN

5.1. Cuenca 1: Río Turbio

5.1.1. Filtro de Kalman Discreto DKF

Se realizaron varios pronósticos teniendo en cuenta la disponibilidad de las series de caudal y precipitación para los años 2003 y 2004 registrados en las bases de datos BANDAS y CLICOM (CONAGUA, 2014b; SMN, 2010), utilizando el enfoque DKF. A continuación, se muestran los resultados obtenidos para los datos que inician en el mes de Junio hasta Diciembre para la serie del año 2003.

En la parte superior de la figura 20, se muestra la precipitación media de la cuenca obtenida a través de los polígonos de Thiessen para la serie en mención. En la parte inferior de la misma, se puede observar el pronóstico realizado a través del filtro de Kalman utilizando el enfoque DKF. En este pronóstico, fue necesario considerar 2 días de lluvia y 1 día de caudal previo para conformar la matriz H tal como se describió en la figura 17, considerando los valores de $\alpha=0.3$ y la matriz de covarianza inicial $P_0 = 1000 \cdot I$. Se puede apreciar cerca de la ordenada junio/27/2003 un sobre salto en la línea roja de predicción, la cual puede ser atribuida a las primeras convoluciones del filtro que generan cambios abruptos y tienden a ser suavizados a medida que el sistema incorpora las nuevas mediciones.

En la figura 21, se muestran los caudales pronosticados y actualizados con sus respectivos errores $e_{t+1|t} = (Q_{t+1} - \hat{Q}_{t+1|t})$. Como se observa en la gráfica de caudales actualizados, estos presentan un mejor ajuste con los datos observados y sus errores tienden a ser de menor magnitud debido a que las mediciones del sistema real son asimiladas por el algoritmo

En el cuadro 7 se resume las estadísticas para el pronóstico de caudales, los cuales presentan muy buenos resultados en cuanto al grado de ajuste entre el caudal observado y el pronosticado con coeficientes de Nash-Sutcliffe de 0.97 respectivamente. La media de los valores pronosticados y actualizados se mantiene con respecto a los observados, el RMSE de los valores actualizados se encuentra ligeramente por debajo de los valores pronosticados y con una tendencia leve del modelo a sobreestimar los caudales pronosticados. Es de aclarar, que aunque el filtro de Kalman presenta dos etapas (pronostico y actualizacion), las estadísticas que se tendrán en cuenta para evaluar la utilidad de la metodología, será la de caudales pronosticados.

Cuadro 7. Resumen estadístico para el pronóstico en la cuenca del río Turbio serie 2003.

Estadísticas	Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
Caudales Pronosticados	21.62	23.37	4.83	0.97	0.98	1.75
Caudales Actualizados	21.52	18.09	4.25	0.98	0.99	1.31
Caudales Observados	21.23	-	-	-	-	-

*Las estadísticas son con respecto a los caudales observados

Las siglas de las estadísticas descritas en el cuadro anterior, son las siguientes:

MSE cuadrado medio del error, RMSE raíz del cuadrado medio del error, R² coeficiente de correlación de los datos y PBE porcentaje de error sistemático.

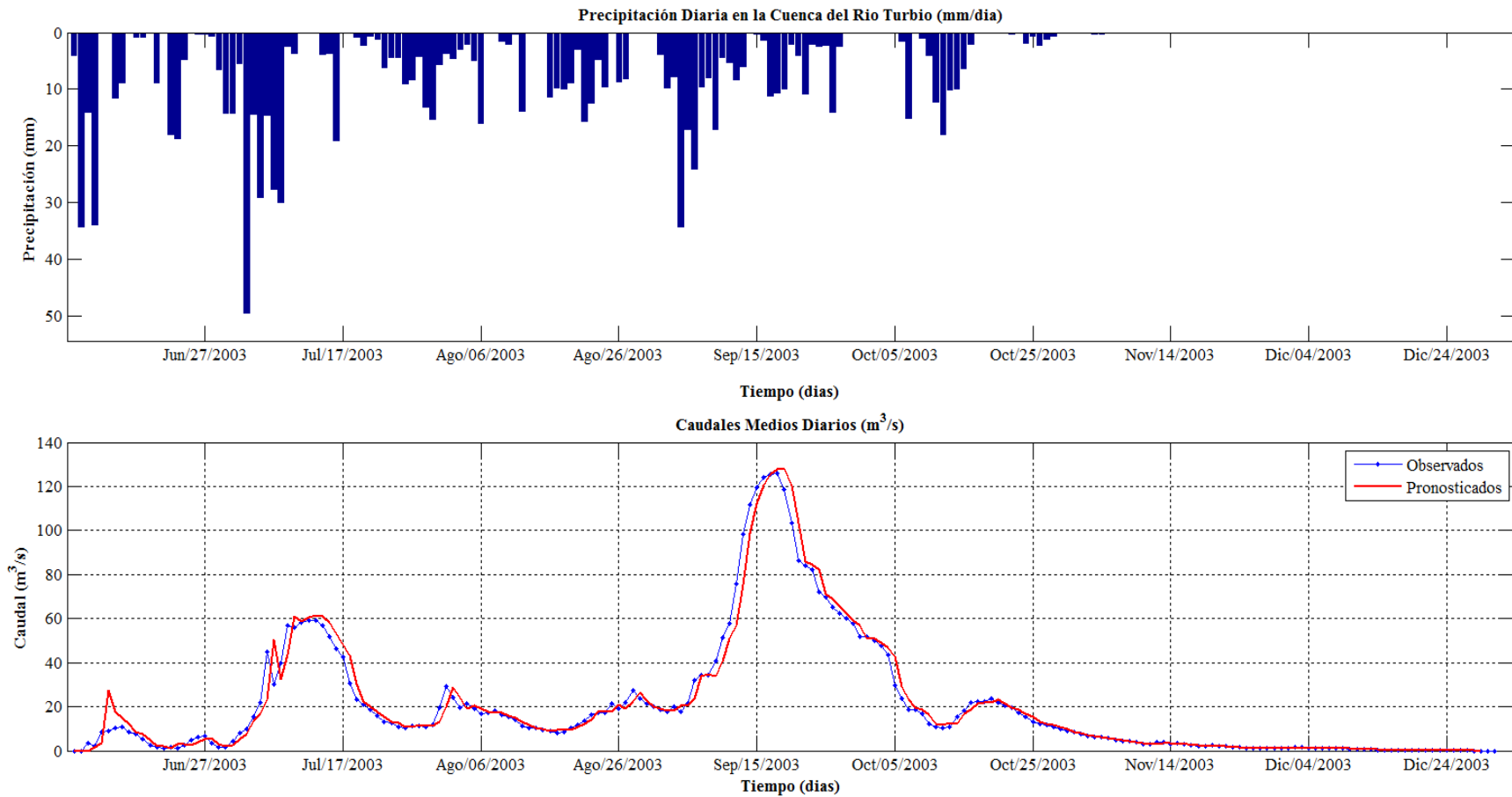


Figura 20. Pronóstico de caudales en la cuenca río Turbio para la serie del año 2003.

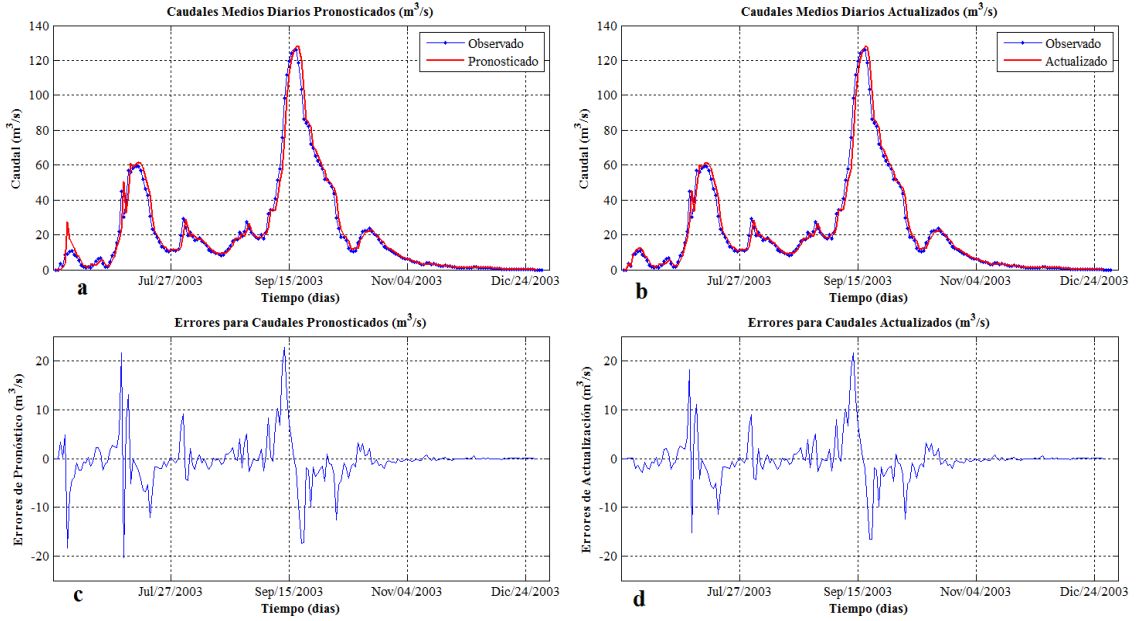


Figura 21. Caudales pronosticados y actualizados (a, b) y sus errores (c, d) generados en la cuenca del río Turbio para la serie del año 2003.

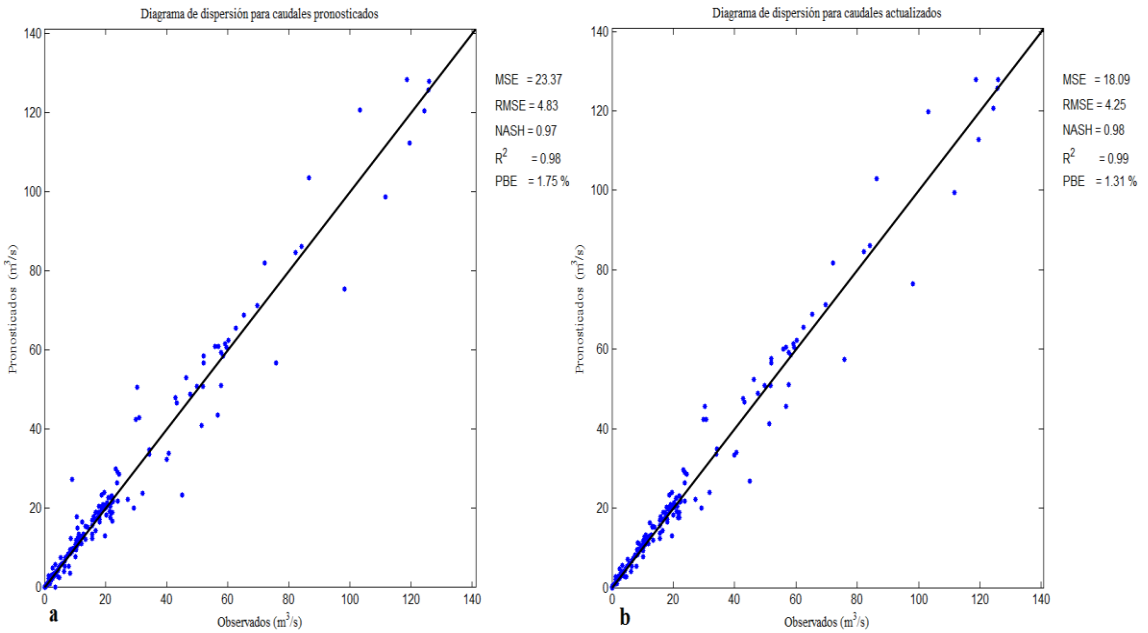


Figura 22. Diagramas de dispersión para caudales pronosticados (a) y actualizados (b) para la cuenca del río Turbio para la serie del año 2003.

En la figura 23, se muestra el pronóstico de caudales para la serie que comprende el año 2004. La conformación de la matriz H que contiene los valores de precipitación y caudal fueron los mismos que los reportados en la serie del 2003 ($\alpha=0.3$, 2 días de precipitación previa y 1 día para caudal previo, $P_0 = 1000 \cdot I$ y matriz de covarianza del proceso $[Q]_{n \times n}$ de valor nulo).

En la figura 25 se muestra un buen ajuste entre los caudales pronosticados y los caudales observados, reflejados en el coeficiente de Nash-Sutcliffe de 0.97. De igual forma, se puede apreciar por el valor de PBE (Percentage Bias Error o medida del error sistemático) que existe una leve tendencia a subestimar los caudales pronosticados con respecto a los observados.

Los resultados obtenidos en las dos series 2003 y 2004, son muy semejantes a los reportados por Morales *et al.* (2014), en su aplicación del filtro de Kalman Discreto a nivel diario para la cuenca peñitas, en Chiapas México, pues también considero 2 días previos de precipitación y 1 día previo de caudal, como vectores a conformar la matriz H en la ecuación de mediciones del filtro. El valor de la media que se muestra en el cuadro 8, tiene como objetivo conocer que tanto se preserva el estadístico en las series generadas, siguiendo la metodología propuesta por Morales *et al.* (2014).

En el cuadro 8, se muestra el resumen estadístico de los pronósticos realizados para la serie 2004.

Cuadro 8. Resumen estadístico para el pronóstico de caudales en la cuenca del río Turbio serie 2004.

Estadísticas	Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
Caudales Pronosticados	9.51	11.89	3.45	0.97	0.94	-2.59
Caudales Actualizados	9.50	10.64	3.26	0.97	0.95	-2.47
Caudales Observados	9.72	-	-	-	-	-

*Las estadísticas son con respecto a los caudales observados

Los resultados para la serie de 2004 presentan muy buen ajuste en términos de MSE y Nash-Sutcliffe, con la tendencia a mantenerse el valor promedio de los caudales generados en comparación con la serie de caudales observados, además de que el RMSE para el pronóstico de caudales en la serie del año 2004 es menor que para los valores obtenidos en la serie del año 2003. En este caso, según los valores del PBE el modelo tiende a subestimar los caudales pronosticados.

En la figura 23, se puede apreciar que la lluvia no presenta un comportamiento de respuesta directa, ya que se encuentran desfasados algunos puntos máximos del hidrograma con respecto a los hietogramas de precipitación. Esto puede ser atribuido a las condiciones en las que se encontraba la cuenca al momento de presentarse el evento (humedad del suelo, cobertura vegetal, usos del suelo, etc) aunque también puede ser posible, debido a que en la cuenca del río Turbio existe una alta presencia de reservorios (por ser una zona agrícola en la mayor parte de su extensión), que ejercen un control sobre las avenidas provocando un retraso en el caudal máximo de respuesta para una evento dado.

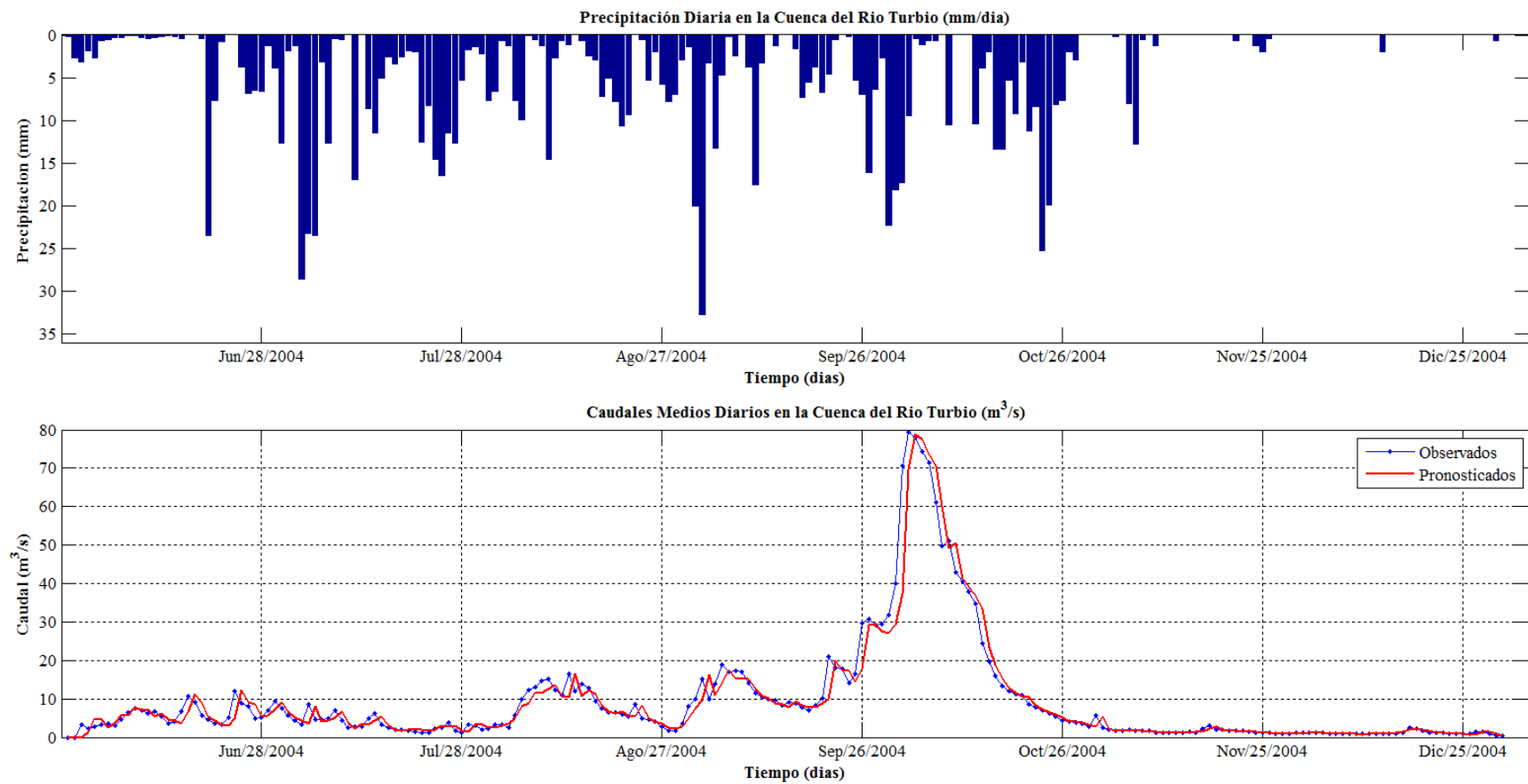


Figura 23. Pronóstico de caudales en la cuenca del río Turbio para la serie del año 2004.

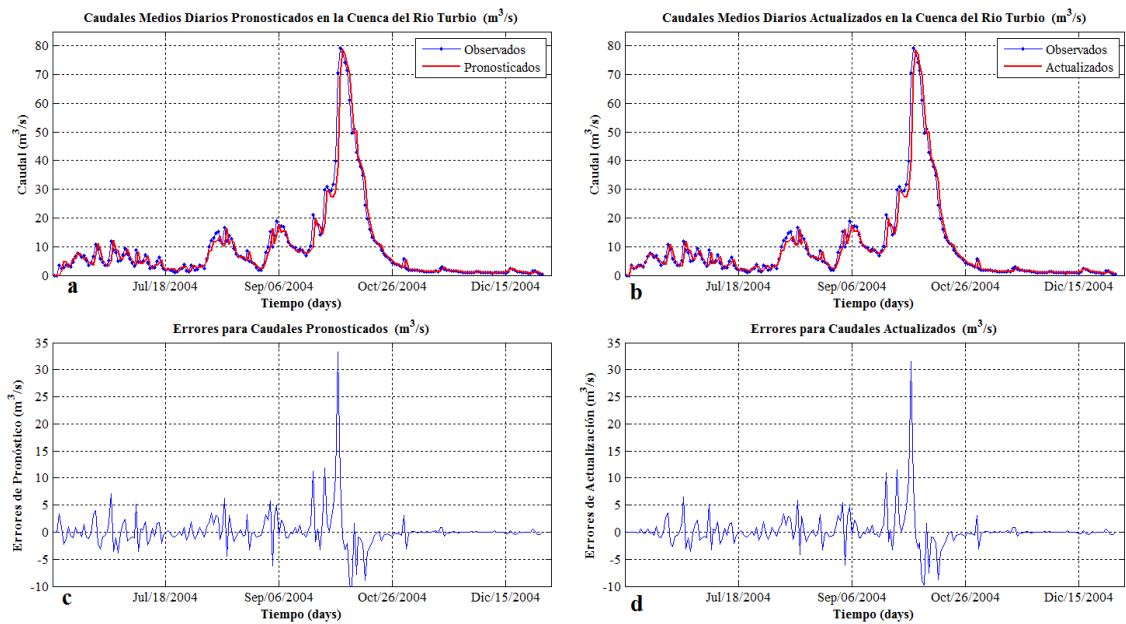


Figura 24. Caudales pronosticados y actualizados (a, b) y sus errores (c, d) en la cuenca del río Turbio para la serie del año 2004.

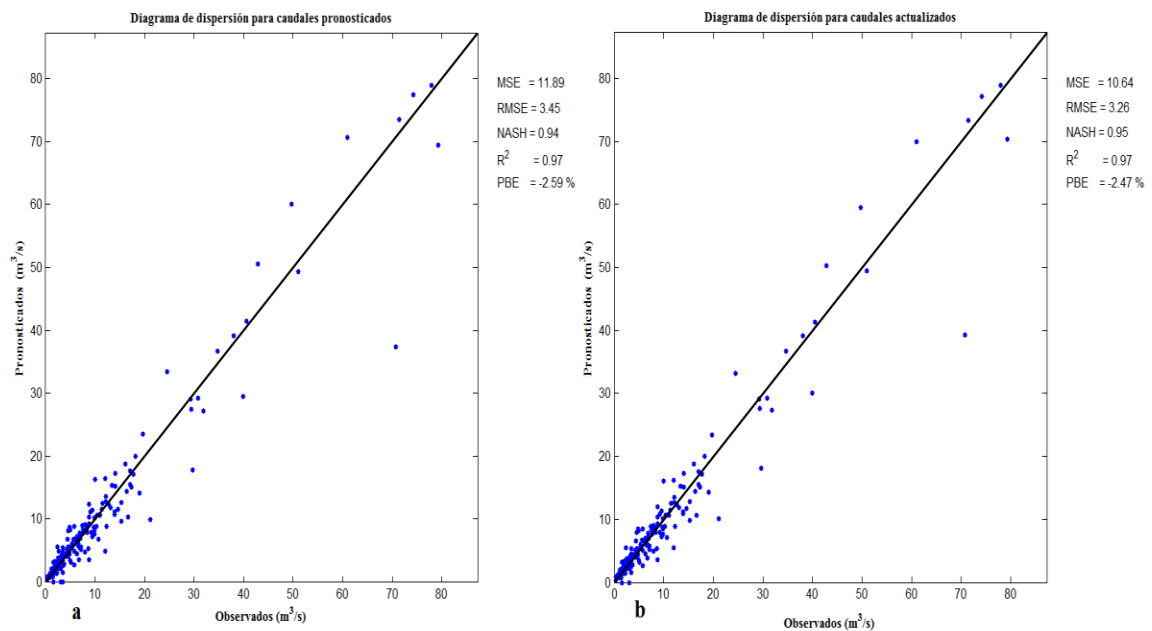


Figura 25. Diagramas de dispersión para caudales pronosticados(a) y actualizados (b) en la cuenca del río Turbio para la serie del año 2004.

5.1.2. Enfoque ARX+DKF

A continuación se muestran los resultados obtenidos para la serie del año 2003 con la implementación de un modelo ARX tal como se describió en la metodología del apartado (4.3.2.) para la cuenca del río Turbio. El periodo que utilizó el modelo para su calibración previa fue de 20 días con la siguiente estructura en cada periodo generado.

$$q_t = \sum_{i=1}^{na=2} \alpha_i q_{t-i} + \sum_{j=1}^{nb=1} \beta_j u_{t-j} + e(t)$$

Donde $na = 2$, se consideró el orden autorregresivo para caudales y $nb=1$ el número de días de precipitación previa. El periodo escogido fue el que mejor resultados en términos estadísticos arrojó luego de probar con varios periodos de calibración. La figura 26 muestra los resultados obtenidos a través de la metodología ARX+DKF para $L=1$, es decir, para el pronóstico realizado en el tiempo $(t+1|t)$ o a 24 h de anticipación.

De igual forma, fueron realizados los pronósticos para diferentes pasos hacia adelante, $x_{t+1|t}$, $x_{t+2|t}$, ..., $x_{t+L|t}$, con el objetivo de cuantificar los errores de predicción teniendo en cuenta la información previa y sin actualizar, como se describió en la figura 18 del apartado (4.3.2). En el cuadro 9, se resumen las estadísticas para los diferentes pronósticos en el tiempo utilizando la serie del año 2003. Los pronósticos con $L=1$ corresponden a los realizados a 24 h, $L=2$ para 48 h, $L=3$ para 72 h y $L=4$ para 96 h respectivamente.

Cuadro 9. Resumen estadístico para diferentes pronósticos mediante ARX+DKF en la cuenca del río Turbio para la serie del año 2003.

Estadísticas	Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE %
Caudales Pronosticados L=1	22.6	41.37	6.43	0.97	0.95	-3.82
Caudales Pronosticados L=2	22.0	103.63	10.18	0.93	0.87	-6.14
Caudales Pronosticados L=3	21.6	188.29	13.72	0.87	0.76	-7.87
Caudales Pronosticados L=4	21.2	291.01	17.06	0.80	0.63	-9.35
Caudales Observados	20.1	-	-	-	-	-

*Las estadísticas son con respecto a los caudales observados

De acuerdo al cuadro 9, la media para los caudales pronosticados tiende a preservarse con respecto a la serie de los caudales observados, el coeficiente de eficiencia de Nash-Sutcliffe disminuye conforme el RMSE aumenta a medida que el pronóstico se hace L pasos hacia adelante, generando pérdida de confiabilidad en los pronósticos realizados. Los resultados obtenidos por el PBE para todos los pronósticos en el tiempo, indican que los caudales tienden a ser subestimados por el modelo.

Como se observa en la figuras 26 y 27, los pronósticos realizados presentan un comportamiento característico para caudales bajos, medios y altos de acuerdo a la dispersión que presentan a lo largo de la línea a 45 grados. Por esta razón, los datos se agruparon en 3 clases $Q < 20 \text{ m}^3/\text{s}$, $20 \text{ m}^3/\text{s} \leq Q \leq 60 \text{ m}^3/\text{s}$ y $Q > 60 \text{ m}^3/\text{s}$ (ver figura 28) siguiendo el patrón de dispersión de la figura 27.

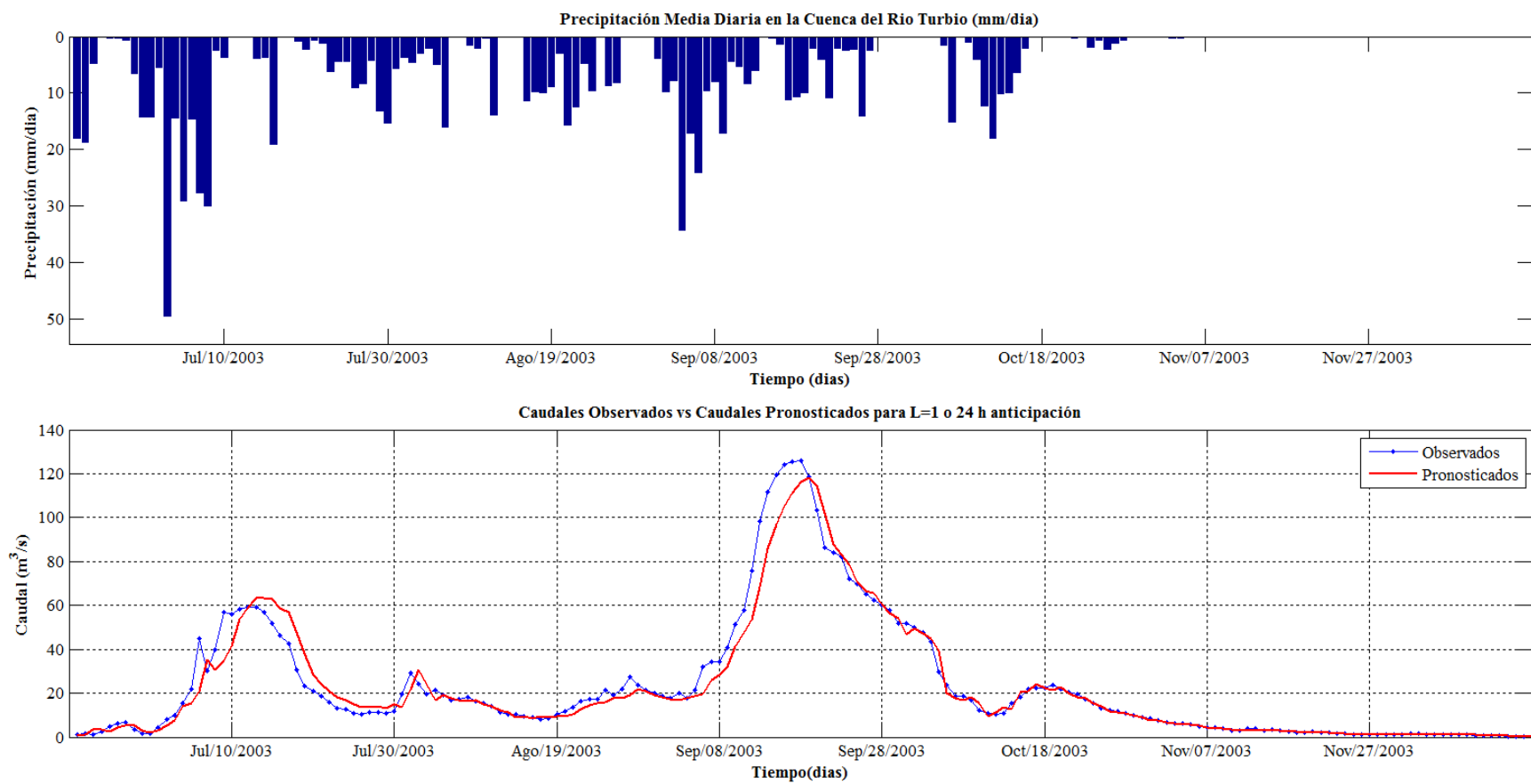


Figura 26. Pronóstico de caudales mediante ARX+DKF en la cuenca del río Turbio para la serie del año 2003.

Diagrama de dispersión para pronósticos con L=1 o 24 h anticipación

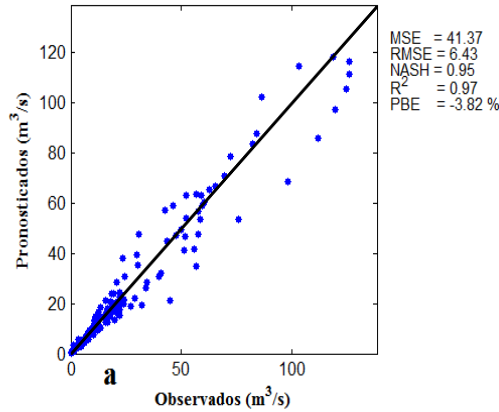


Diagrama de dispersión para pronósticos con L=2 o 48 h anticipación

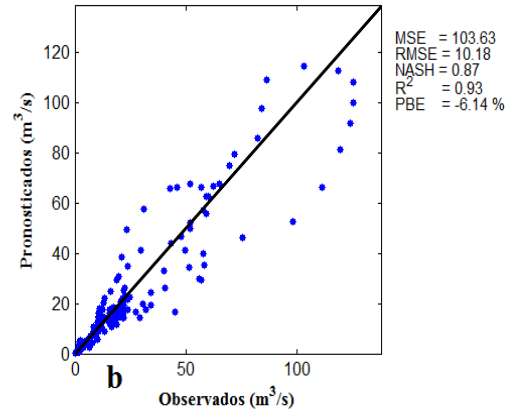


Diagrama de dispersión para pronósticos con L=3 o 72 h anticipación

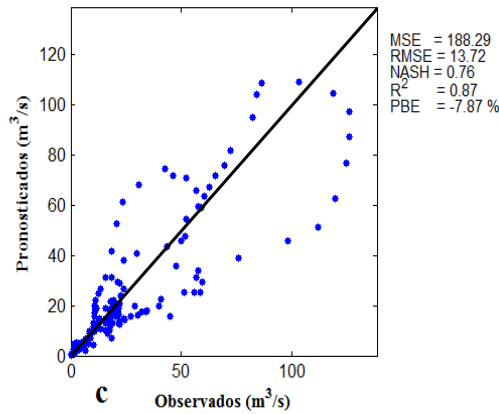


Diagrama de dispersion para pronósticos con L=4 o 96 h anticipación

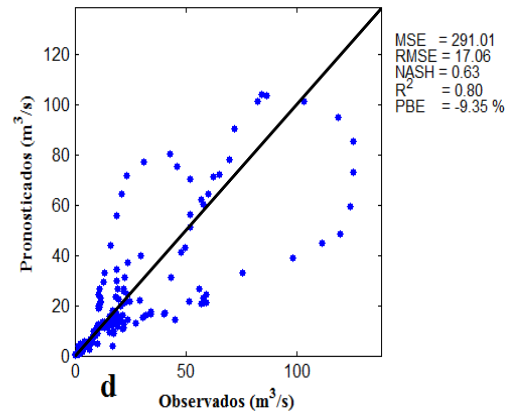


Figura 27 Diagramas de dispersión comparativos para pronósticos con ARX+DKF a 24 h (a), 48 h (b), 72 h (c) y 96 h (c) en la cuenca del río Turbio para la serie del año 2003.

El orden de los diagramas de dispersión en la figura 28 es para L=1 gráficas (a, b, c), para L=2 gráficas (d, e, f), para L=3 gráficas (g, h, i) y para L=4 gráficas (j, k, l). Se puede apreciar que para la serie del año 2003, el ARX+KFD realizó un mejor pronóstico para caudales menores a $20 \text{ m}^3/\text{s}$, en diferentes pasos del tiempo (ver gráficas a,d,g,j). Por el contrario, para los pronósticos de caudales entre $20\text{-}60 \text{ m}^3/\text{s}$ (ver gráficas b,e,h,k) el ARX+KFD tiende a perder el grado de ajuste con respecto a los valores observados, siendo mayor la incertidumbre para los caudales pronosticados por arriba de los $60 \text{ m}^3/\text{s}$ (ver gráficas c,f,i,l).

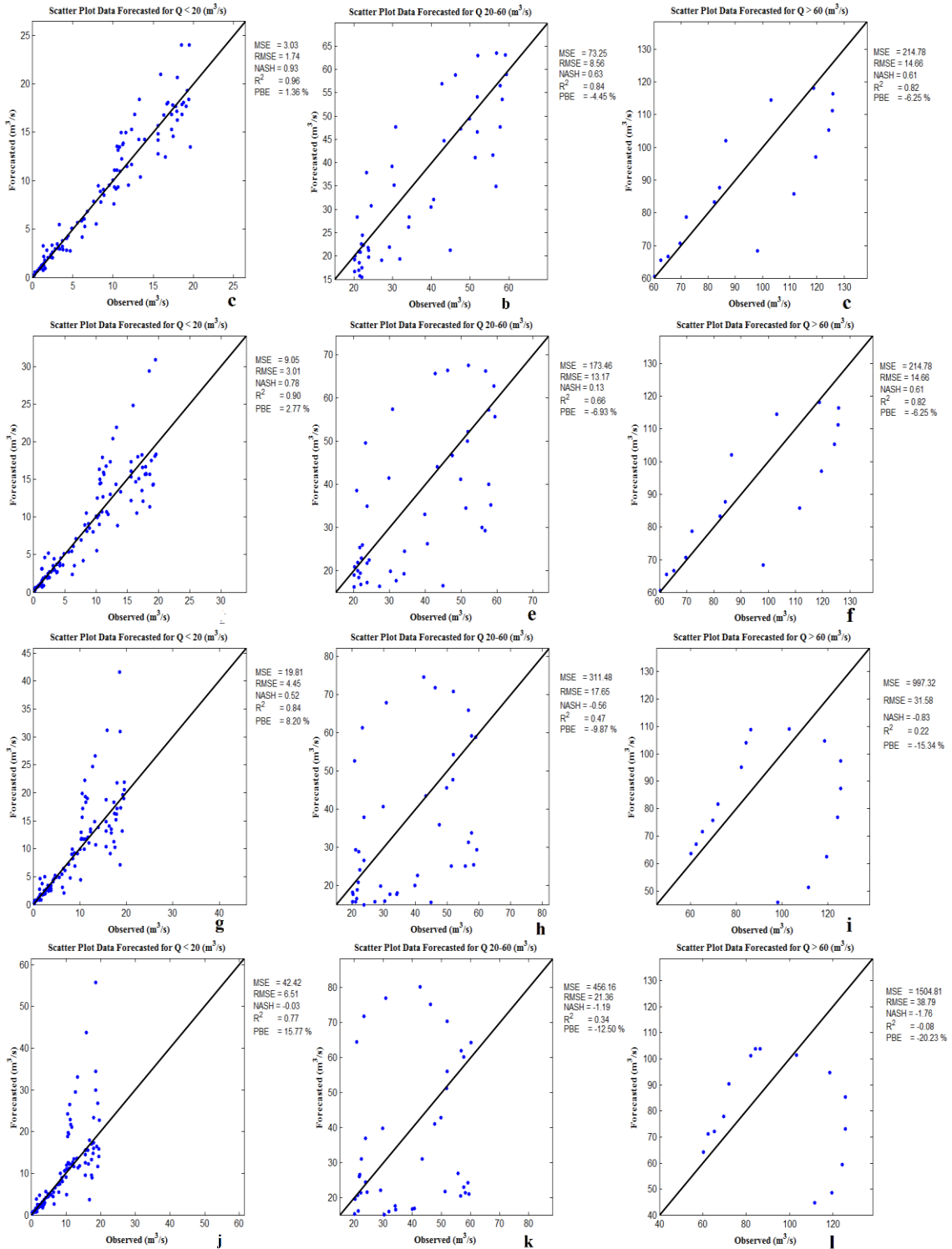


Figura 28. Diagramas de dispersión por clases para los pronósticos en diferentes pasos en el tiempo en la cuenca del río Turbio para la serie del año 2003.

Cuadro 10. Resumen estadístico para los pronósticos agrupados por clases en diferentes pasos del tiempo utilizando datos del año 2003.

Pronósticos	Rango	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
L=1 24 h	Q < 20 m3/s	3.03	1.74	0.96	0.93	1.36
	20 ≤ Q ≤ 60 m3/s	73.25	8.56	0.84	0.63	-4.45
	Q > 60 m3/s	214.78	14.66	0.82	0.61	-6.25
L=2 48 h	Q < 20 m3/s	9.05	3.01	0.9	0.78	2.77
	20 ≤ Q ≤ 60 m3/s	173.46	13.17	0.66	0.13	-6.93
	Q > 60 m3/s	214.78	14.66	0.82	0.61	-6.25
L=3 72 h	Q < 20 m3/s	19.81	4.45	0.84	0.52	8.20
	20 ≤ Q ≤ 60 m3/s	311.48	17.65	0.47	0.56	-9.87
	Q > 60 m3/s	997.32	31.58	0.22	-0.83	-15.34
L=4 96 h	Q < 20 m3/s	42.42	6.51	0.77	-0.03	15.77
	20 ≤ Q ≤ 60 m3/s	456.16	21.36	0.34	-1.19	-12.50
	Q > 60 m3/s	1504.81	38.79	0.08	-1.76	-20.23

Con el objetivo de establecer el grado de incertidumbre asociado al realizar los pronósticos para diferentes pasos en el tiempo, se determinaron los intervalos de predicción para cada valor puntual siguiendo los criterios establecidos por Chatfield (2004). En términos generales, la teoría asume que los errores de predicción $e = (x_{obs} - \hat{x}_{pron})$ provienen de una distribución normal y ente contexto, éste supuesto fue verificado en las diferentes series pronosticadas teniendo en cuenta cada paso en el tiempo (L=1, L=2, L=3, L=4).

En la figura 29, se presentan los errores de predicción ajustados para un conjunto de distribuciones continuas y diferentes pasos en el tiempo utilizando la serie del año 2003. Se puede observar que los errores de predicción para diferentes pasos en el tiempo, se ajustan a una distribución t de student escalada.

Esta distribución pertenece a la familia de las t student y se caracteriza por presentar colas pesadas debido a la presencia de valores atípicos o outliers, que no se presentan en una distribución normal.

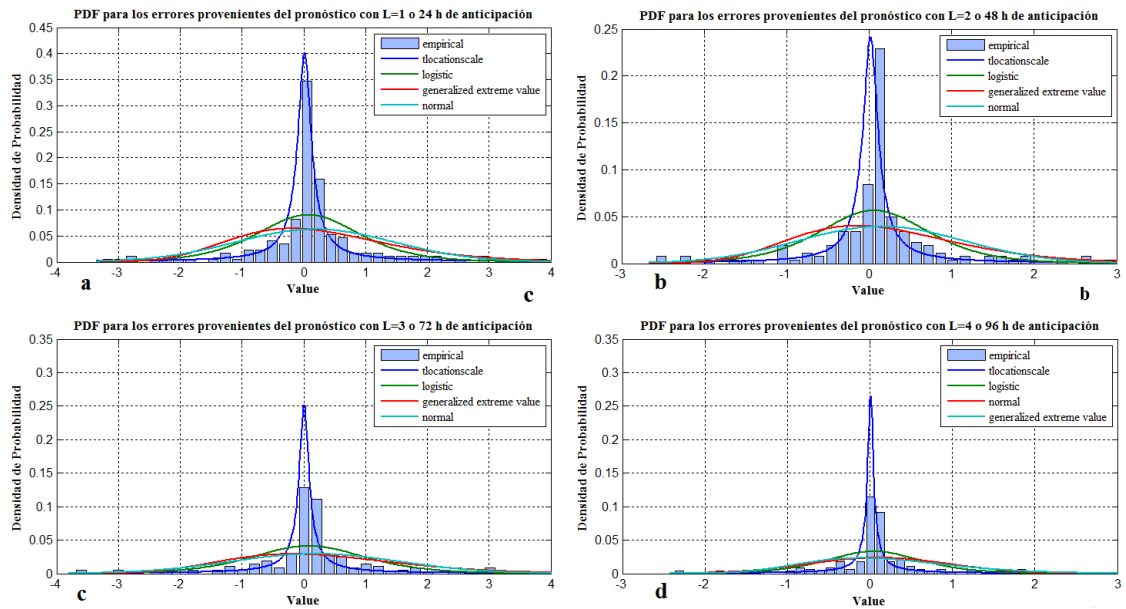


Figura 29. Funciones de distribución de probabilidad para los errores de predicción con diferentes pasos en el tiempo en la cuenca del río Turbio para la serie del año 2003.

Según Chatfield (2001), en la práctica la distribución de los errores no siempre se ajusta a una distribución normal, debido a la asimetría que pueda presentarse o a la presencia de datos atípicos que generan distribuciones con colas pesadas y por errores asociados a los datos involucrados que contaminan la distribución de los errores en la predicción.

Wang (2006) sugiere que la estimación de la incertidumbre para el pronóstico hidrológico se realice a través del análisis estadístico de las propiedades de acuerdo al modelo de errores asociado o implementando técnicas de simulación Monte Carlo a través del método GLUE de Beven y Binley (1992), teoría de estimación Bayesiana y desarrollo de modelos estocásticos tipo AR, ARX, ARMA, ARMAX etc., entre otros.

En la figura 30, se presenta un diagrama de cajas que representa los errores que fueron ajustados mediante las distribuciones mostradas en la figura 29, confirmando la tendencia a incrementar las colas de las distribuciones a medida que se hace el pronóstico L pasos en el tiempo, interpretando que los mejores pronósticos son los realizados en el tiempo inmediato ($t+1$). Los puntos en azul, indican los datos atípicos o outliers para el conjunto de errores.

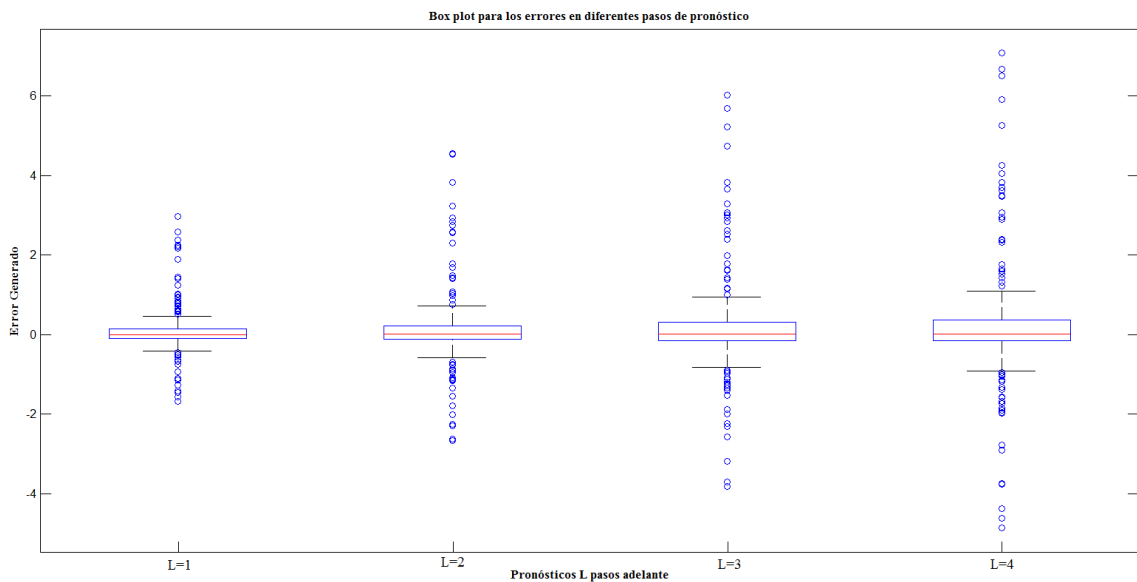


Figura 30. Diagrama de cajas para la distribución de los errores en diferentes pasos en el tiempo.

En la figura 31, se presenta el intervalo de confianza al 95% de probabilidad para el pronóstico realizado con $L=1$, es decir ($x_{t+1|t}$) o a 24 h de anticipación, para la serie del año 2003. Se puede observar que el pronóstico es aceptable ya que los caudales observados se encuentran en el rango de los intervalos de confianza a lo largo de la serie, a diferencia del intervalo generado en el evento de Sep/15/2003 donde el caudal observado queda fuera del rango para el intervalo de confianza, el cual guarda relación con los datos discretizados para $Q > 60 \text{ m}^3/\text{s}$ con $L=1$.

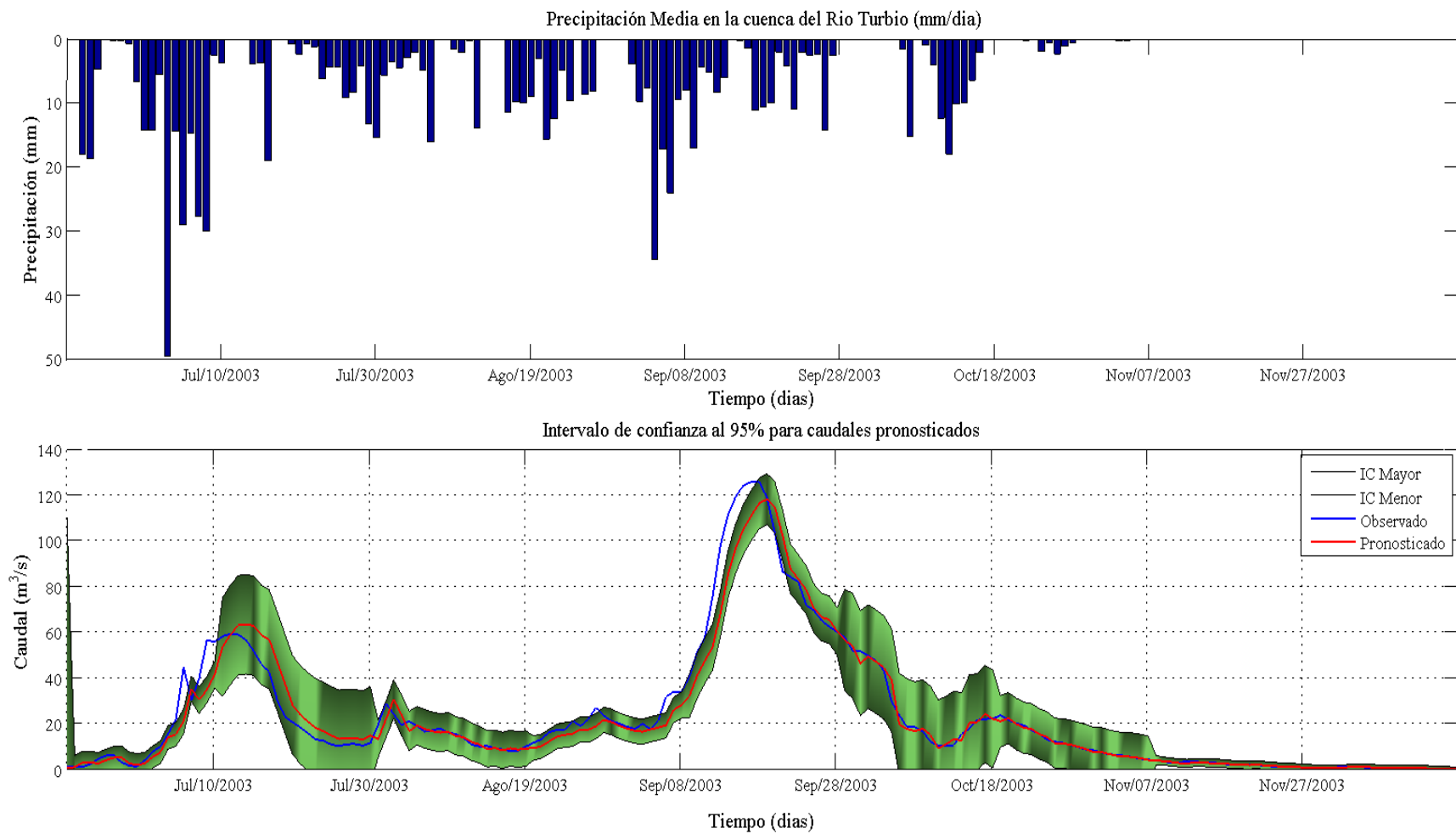


Figura 31. Intervalo de predicción al 95% de probabilidad para $L=1$ o a 24 h en la cuenca del río Turbio para la serie del año 2003.

La figura 32, muestra el pronóstico realizado para la serie del año 2004 utilizando el ARX+DKF para $(x_{t+1|t})$. En esta serie pronosticada, el período que mejor se ajustó fue de 36 días considerando en 2 el orden autorregresivo para caudales y 1 para el número de días de precipitación. Las estadísticas para el pronóstico de esta serie 2004 se resumen en el cuadro 11, teniendo en cuenta que se realizaron 4 pasos adelante en la predicción.

Cuadro 11. Resumen estadístico para los pronósticos a diferentes pasos mediante ARX+DKF en la cuenca del río Turbio para la serie del 2004.

Estadísticas	Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE %
Caudales Pronosticados L=1	10.12	16.32	4.04	0.96	0.93	-5.64
Caudales Pronosticados L=2	9.64	39.13	6.26	0.91	0.82	-10.02
Caudales Pronosticados L=3	9.18	62.16	7.88	0.86	0.72	-14.14
Caudales Pronosticados L=4	8.76	83.13	9.12	0.80	0.62	-17.84
Caudales Observados	8.16	-	-	-	-	-

*Las estadísticas son con respecto a los caudales observados.

Los mejores resultados se obtienen realizando el pronóstico para L=1, es decir, $(x_{t+1|t})$ o a 24 h de anticipación (ver los diagramas figura 33). En este caso, el paso adelante se traduce como el pronóstico realizado a un día, por la resolución temporal de los datos utilizados. A medida que se avanza en el tiempo, el pronóstico tiende a perder eficiencia debido a que no se actualizan los datos asimilados en los días anteriores para el caso de L=4 o para 96 h de anticipación, reflejado en los valores de Nash-Sutcliffe y RMSE, con la tendencia a subestimar los datos según los valores negativos proporcionados por el PBE.

Para tener un mejor estimador del ajuste en cada pronóstico se agrupó la información, tal como se realizó para la serie del año 2003. En este caso, solo se tuvo en cuenta dos clases para agrupar los datos, siendo los $Q < 20 \text{ m}^3/\text{s}$ y los $Q \geq 20 \text{ m}^3/\text{s}$ debido al comportamiento de los datos en las gráficas de la figura 33.

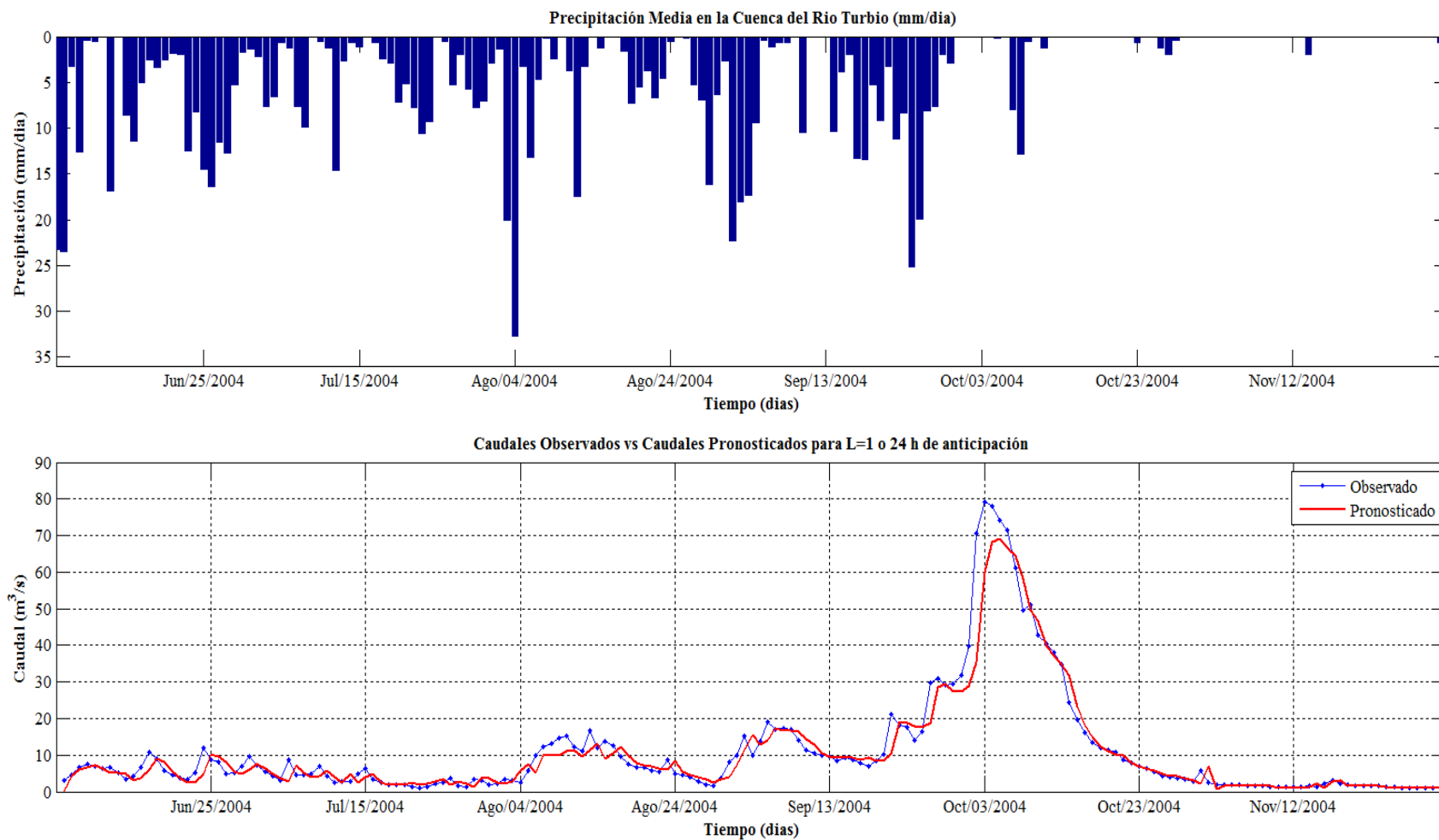


Figura 32. Pronóstico de caudales mediante ARX+DKF en la cuenca del río Turbio para la serie del año 2004.

Diagrama de dispersion para pronósticos con L=1 o 24 h de anticipación

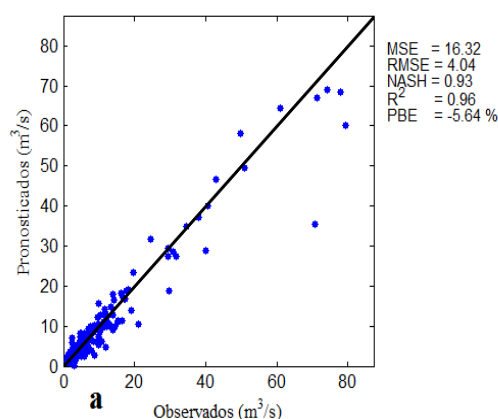


Diagrama de dispersion para pronósticos con L=2 o 48 h de anticipación

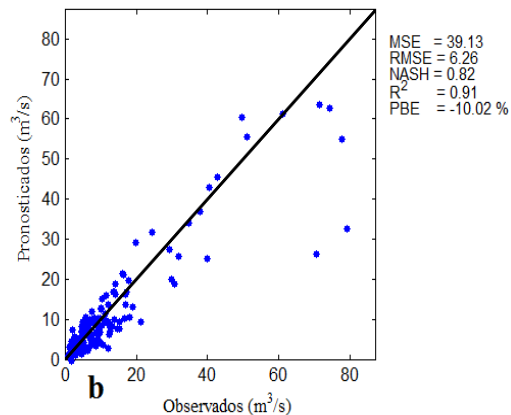


Diagrama de dispersion para pronósticos con L=1 o 72 h de anticipación

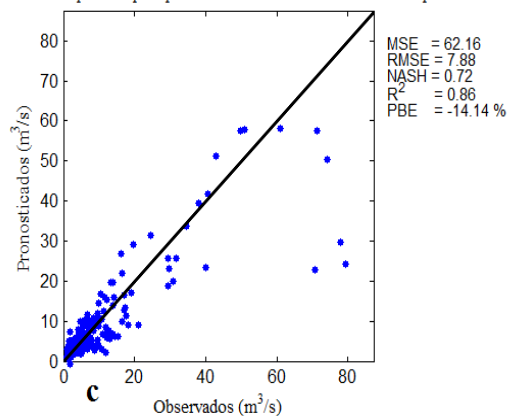


Diagrama de dispersion para pronósticos con L=1 o 96 h de anticipación

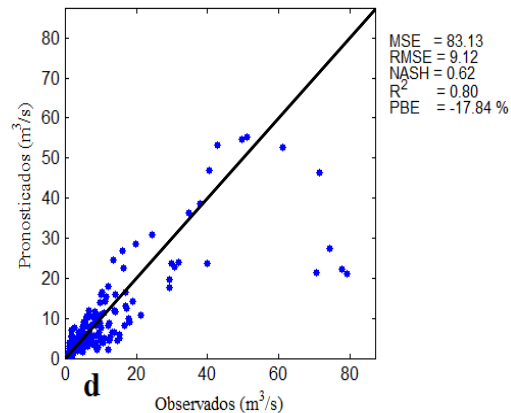


Figura 33. Diagramas de dispersión comparativos para pronósticos a diferentes pasos hacia adelante en la cuenca del río Turbio para la serie del año 2004.

El orden de los diagramas de dispersión en la figura 34 es para L=1 gráficas (a, b), para L=2 gráficas (c, d), para L=3 gráficas (e, f) y para L=4 gráficas (g, h). Se observa que los mejores ajustes para esta serie se encuentran para los caudales menores a 20 m³/s, tal como se observó en la serie del año 2003, ver gráficas (a, c, e, g). Los caudales mayores iguales a 20 m³/s tienden a ser subestimados en todos los pasos de tiempo, lo que implica que el modelo de pronóstico para caudales que se encuentren en este rango pierde confiabilidad, ver gráficas (b, d, f, h).

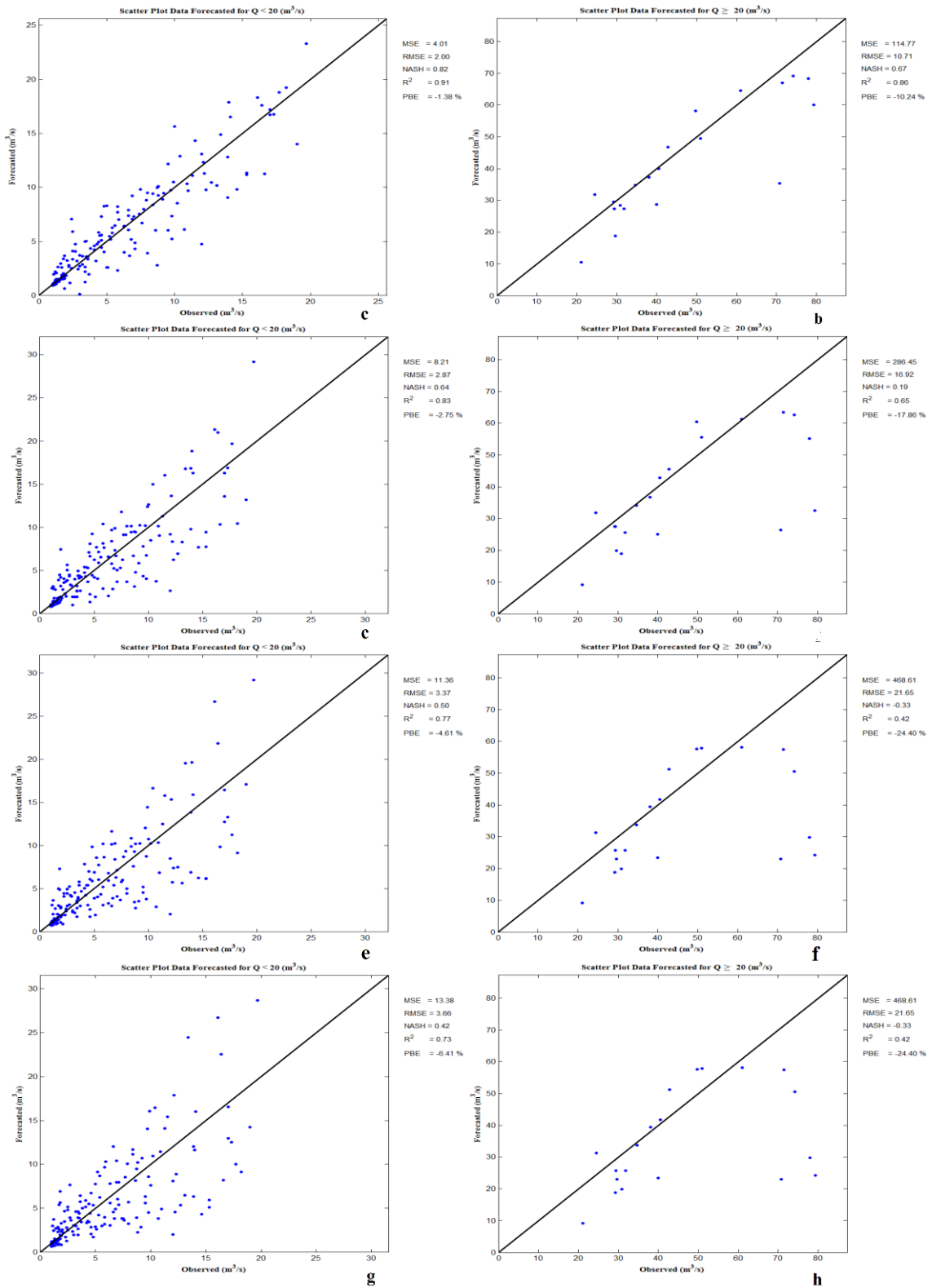


Figura 34. Diagramas de dispersión agrupados para los pronósticos en diferentes pasos en el tiempo para la cuenca del río Turbio con la serie del año 2004.

Cuadro 12. Resumen estadístico para los pronósticos agrupados en diferentes pasos del tiempo utilizando datos del año 2004.

Pronósticos	Rango	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
L=1	Q < 20 m3/s	4.01	2.00	0.91	0.82	-1.38
24 h	Q ≥ 20 m3/s	114.77	10.71	0.86	0.67	-10.24
L=2	Q < 20 m3/s	8.21	2.87	0.83	0.64	-2.75
48 h	Q ≥ 20 m3/s	286.45	16.92	0.65	0.19	-17.86
L=3	Q < 20 m3/s	11.36	3.37	0.77	0.50	-4.61
72 h	Q ≥ 20 m3/s	468.61	21.65	0.42	-0.33	-24.40
L=4	Q < 20 m3/s	13.38	3.66	0.73	0.42	-6.41
96 h	Q ≥ 20 m3/s	468.61	21.65	0.42	-0.33	-24.40

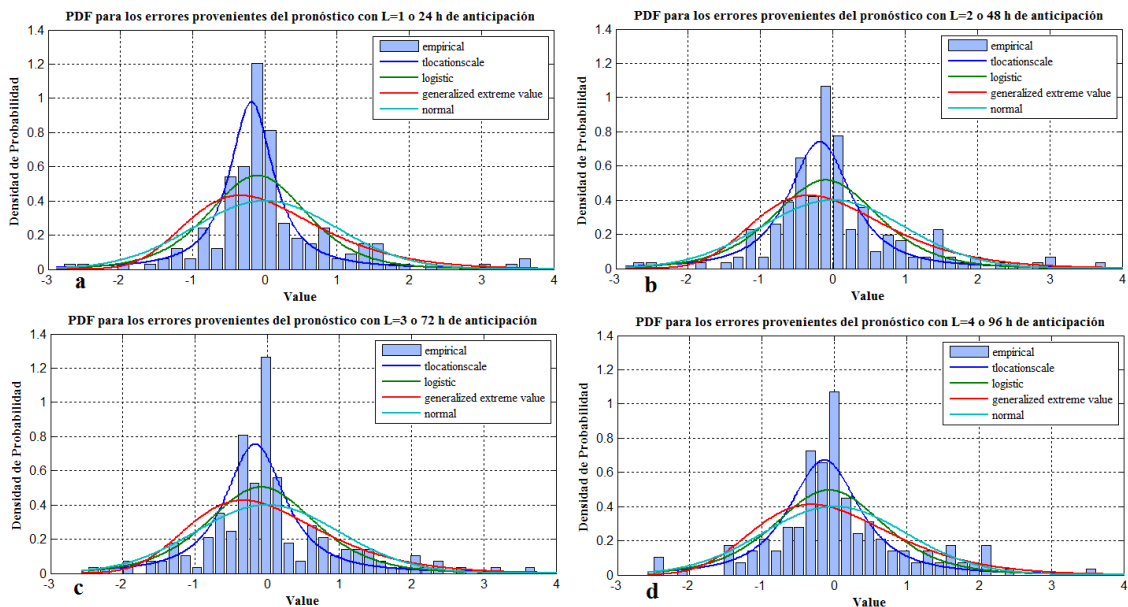


Figura 35. Funciones de distribución de probabilidad para los errores de predicción con diferentes pasos en el tiempo en la cuenca del río Turbio para la serie del año 2004.

La función de densidad de probabilidad obtenida para cada conjunto de errores en diferentes pasos del tiempo en la serie de 2004, sigue la tendencia a presentar colas pesadas conformadas por datos atípicos presentes.

A diferencia de las funciones obtenidas para la serie del año 2003, las distribuciones de la figura 35 presentan una mayor varianza ya que la forma de campana tiende a ser más amplia, sustentado esto, en los valores obtenidos por los ajustes realizados. En todos los 4 casos de pronóstico, las distribuciones se ajustaron a una t student escalada, generada por datos atípicos discutidos anteriormente. Obtenidos los parámetros de la función de probabilidad, se construyó el intervalo de predicción para el 95% de probabilidad, tal y como se muestra en la figura 36.

El intervalo de predicción para el 95% de probabilidad se realizó con los datos generados para $L=1$ o con 24 h de anticipación, mostrando ajustes aceptables con respecto a los valores observados de la serie en mención. En congruencia con los resultados obtenidos agrupando los caudales pronosticados en los rangos descritos anteriormente, los intervalos de predicción para caudales mayores iguales a $20 \text{ m}^3/\text{s}$ tienden a ser subestimados por el modelo ARX+DKF. Por esta razón, para el evento registrado entre Oct/01/2004 y Oct/04/2004 los valores observados quedan por fuera de las envolventes para el intervalo de predicción.

De forma general, en las figuras 37 y 38 se comparan los errores generados al realizar los pronósticos a 24 h, 48 h, 72 h y 96 h para las series 2003 y 2004 en la cuenca del río Turbio. En cada figura se presentan una tendencia de aumento de rango a medida que se avanza en el pronóstico L pasos en el tiempo, con un comportamiento tal que $L=1 < L=2 < L=3 < L=4$.

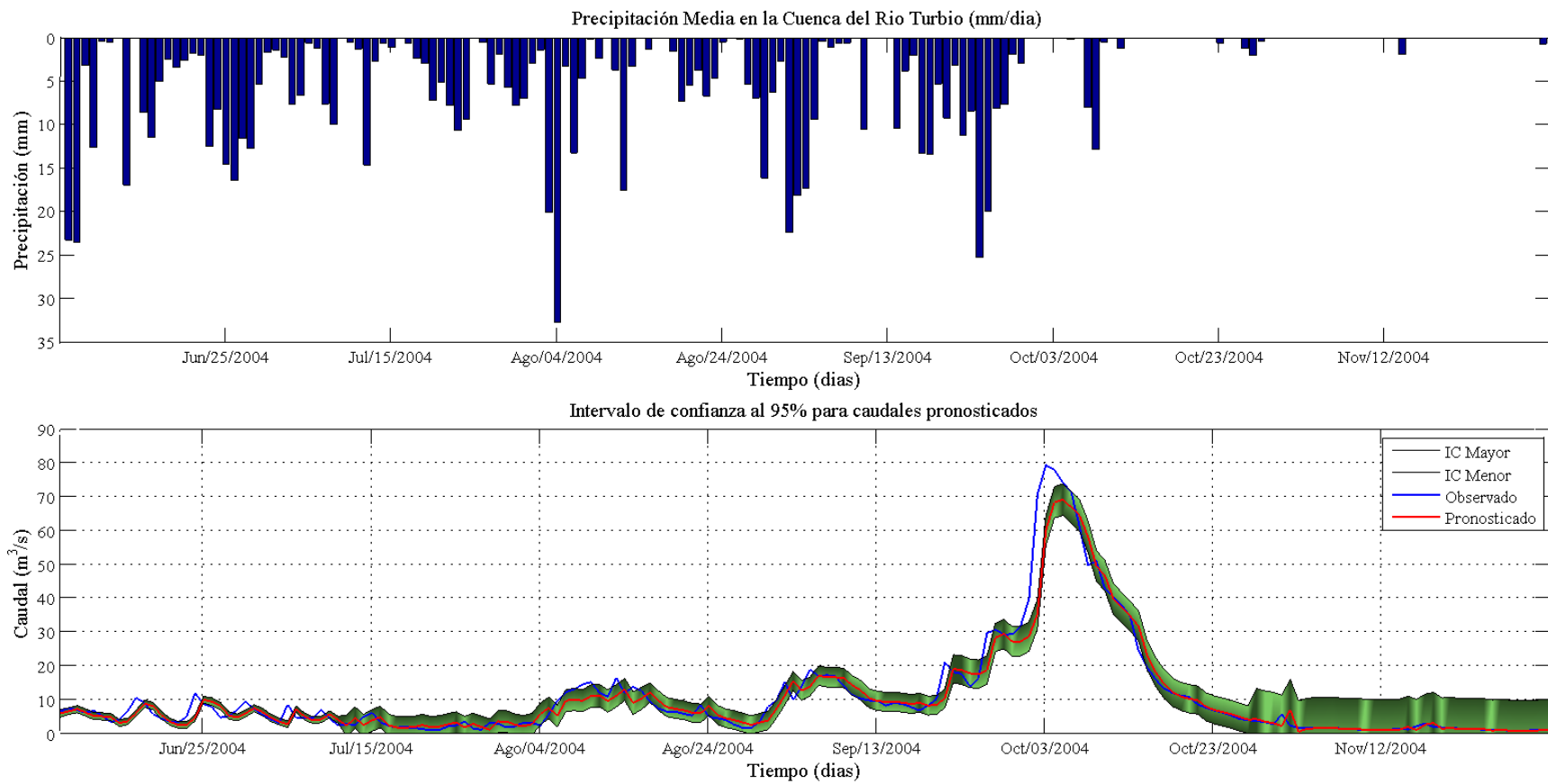


Figura 36. Intervalo de predicción al 95% de probabilidad para $L=1$ o a 24 h en la cuenca del río Turbio para la serie del año 2004.

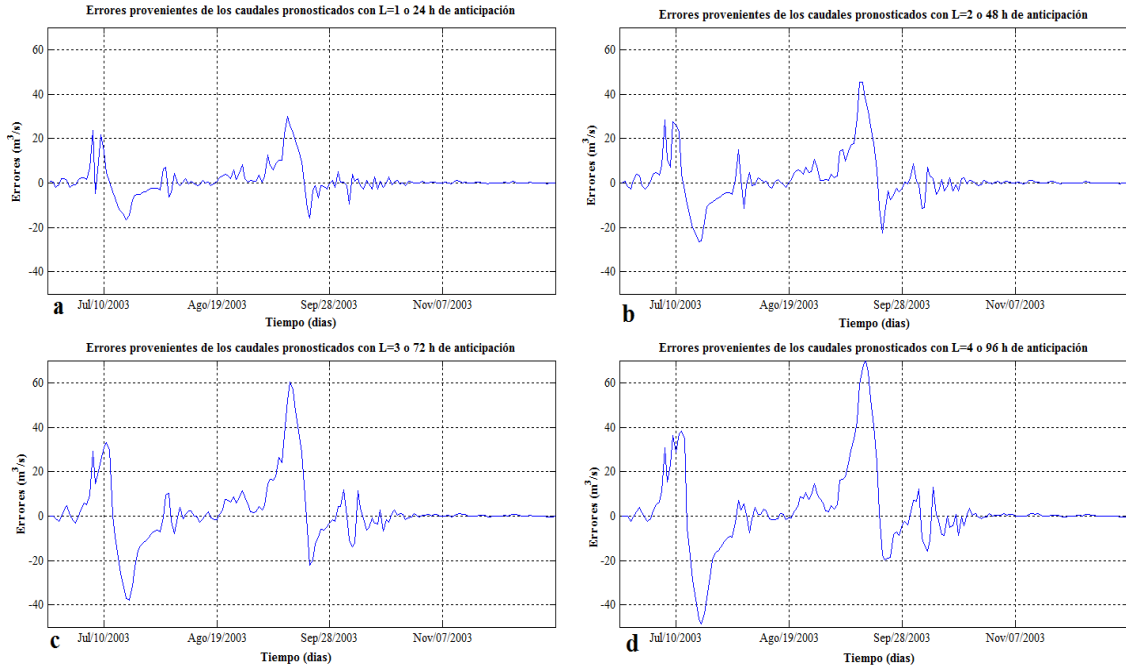


Figura 37. Comparación de los errores pronosticados para diferentes pasos hacia adelante en la cuenca del río Turbio para la serie del año 2003.

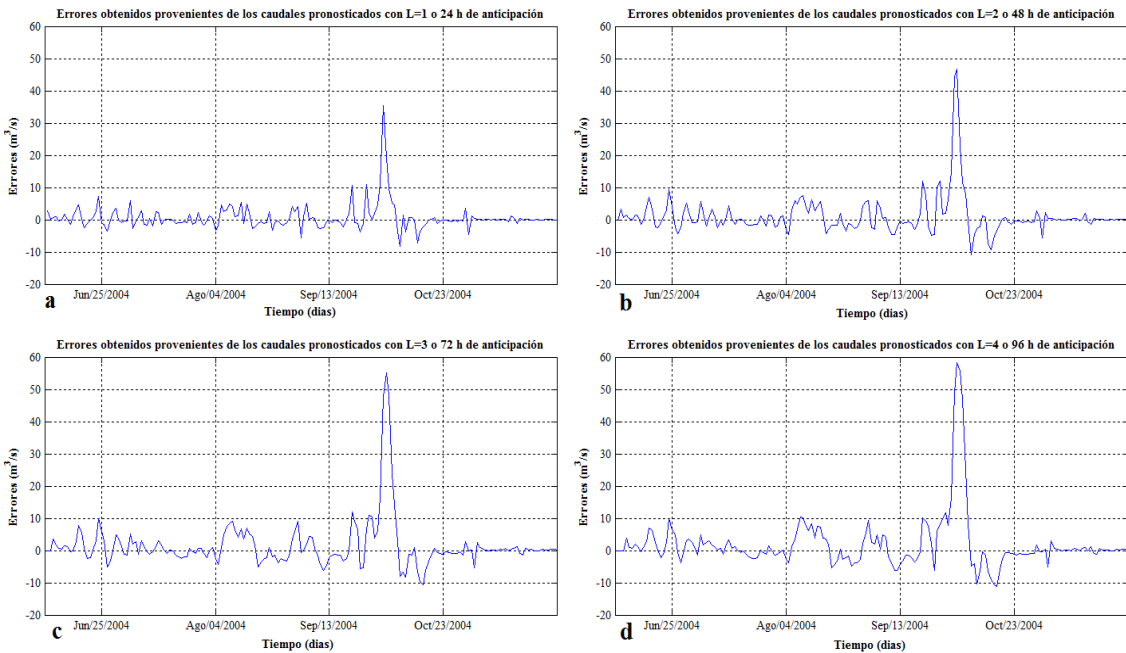


Figura 38. Comparación de los errores pronosticados para diferentes pasos hacia adelante en la cuenca del río Turbio para la serie del año 2004.

5.1.3. Enfoque ARX+EnKF

En este caso, el pronóstico de caudales se realizó mediante el EnKF utilizando la misma estructura operativa del modelo ARX definidas anteriormente y siguiendo la metodología propuesta por Gillijns *et al.* (2006) en la conformación del algoritmo del EnKF.

La figura 39, muestra los caudales pronosticados utilizando el ARX+EnKF utilizando periodos de 20 días para la calibración del modelo ARX, con una estructura autorregresiva de 2 coeficientes para caudal y 1 coeficiente para la lluvia para la serie del año 2003.

La figura 40, muestra el total de miembros del conjunto generados para la serie del año 2003, con un número total de 100, de los cuales, la línea roja representa el valor medio del conjunto de trayectorias pronosticadas. Se puede observar que la tendencia para las estimaciones de los caudales máximos en las avenidas de consideración, no se ajustan como se esperaba al considerar un número razonable de trayectorias recomendada por Gillijns *et al.* (2006), como suficientes para obtener un valor promedio óptimo.

Cuadro 13. Resumen estadístico para el pronóstico mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2003.

Estadísticas	Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
Q Pronosticados EnKF	21.88	61.93	7.87	0.97	0.92	-6.68
Q Actualizados EnKF	22.32	22.16	4.70	0.99	0.97	-4.83
Q Observados	23.45	-	-	-	-	-

*Las estadísticas son con respecto a los caudales observados.

*Los Q pronosticados hacen referencia a la media de los miembros generados.

En las figuras 42, 43 y 44 se presentan los resultados para el pronósticos de caudales utilizando el ARX+EnKF para la serie del año 2004. Las estadísticas se resumen en el cuadro mostrado a continuación.

Cuadro 14. Resumen estadísticas para pronostico mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2004.

Estadísticas	Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE %
Q Pronosticados EnKFs	9.91	19.79	4.45	0.96	0.91	-7.60
Q Actualizados EnKF	10.47	2.54	1.59	0.99	0.99	-2.34
Q Observados	10.73	-	-	-	-	-

*Las estadísticas son con respecto a los caudales observados

*Los Q pronosticados hacen referencia a la media de los miembros generados

Los resultados obtenidos mediante el ARX+EnKF son muy similares a los obtenidos mediante ARX+DKF para las series 2003 y 2004 en la cuenca del río Turbio. Se esperaba que el ARX+EnKF brindara mejores resultados en términos de RMSE considerando que las trayectorias generadas fueron las recomendadas en la literatura como óptimas para obtener una media representativa de los estados. En términos de Nash-Sutcliffe las dos metodologías presentan eficiencias similares en el orden de 0.92 ± 2 .

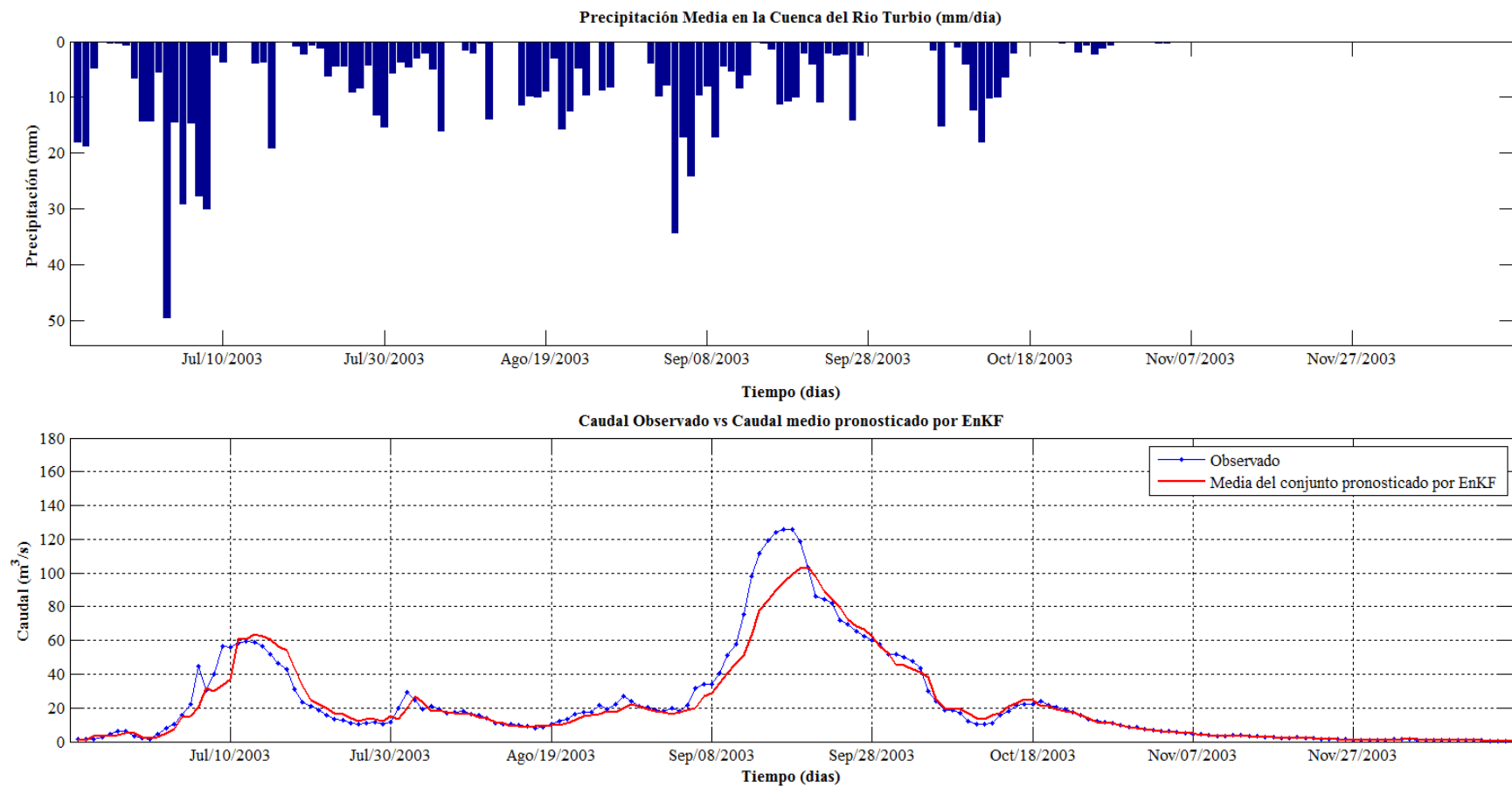


Figura 39. Pronóstico de caudales mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2003.

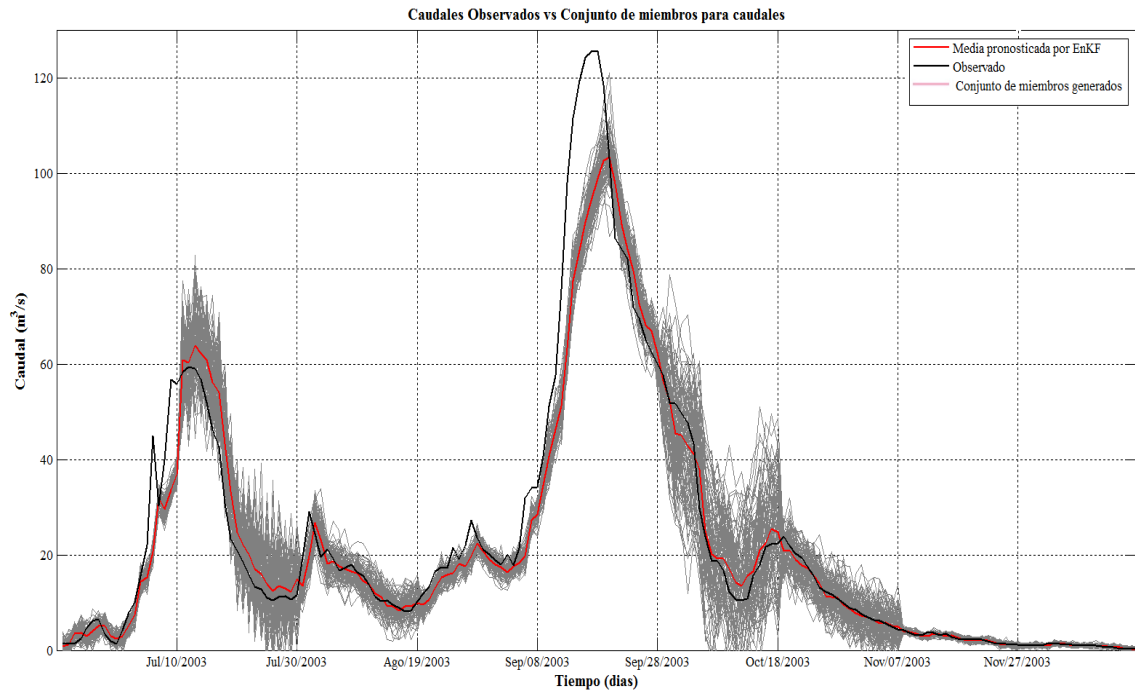


Figura 40. Conjunto de miembros generados para el pronóstico de caudales mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2003.

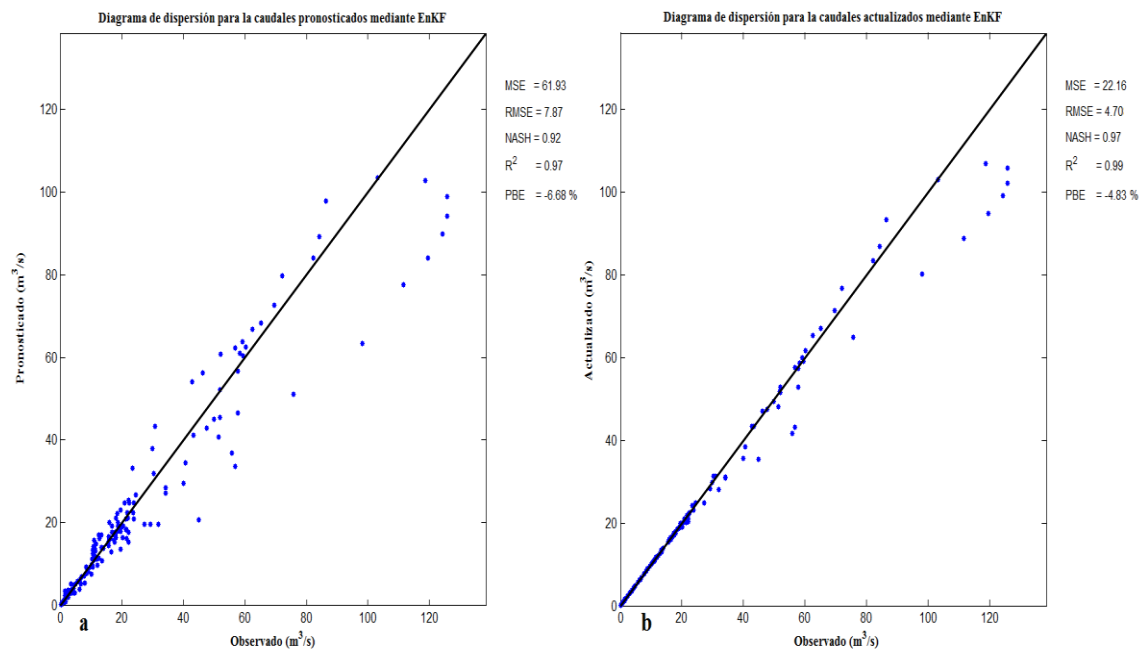


Figura 41. Diagramas de dispersión para la media del conjunto de caudales pronosticados (a) y actualizados (b) mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2003.

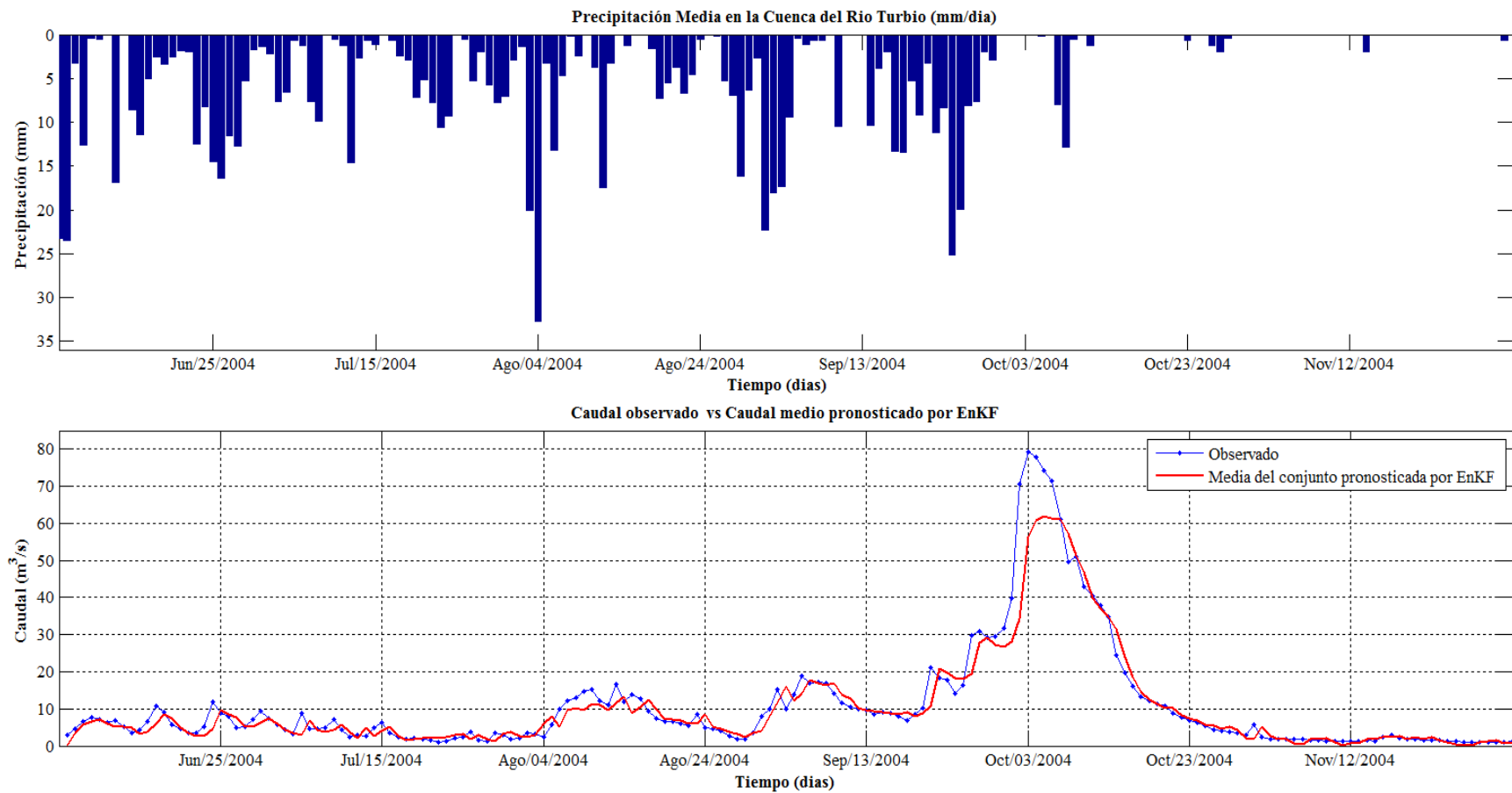


Figura 42. Pronóstico de caudales mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2004.

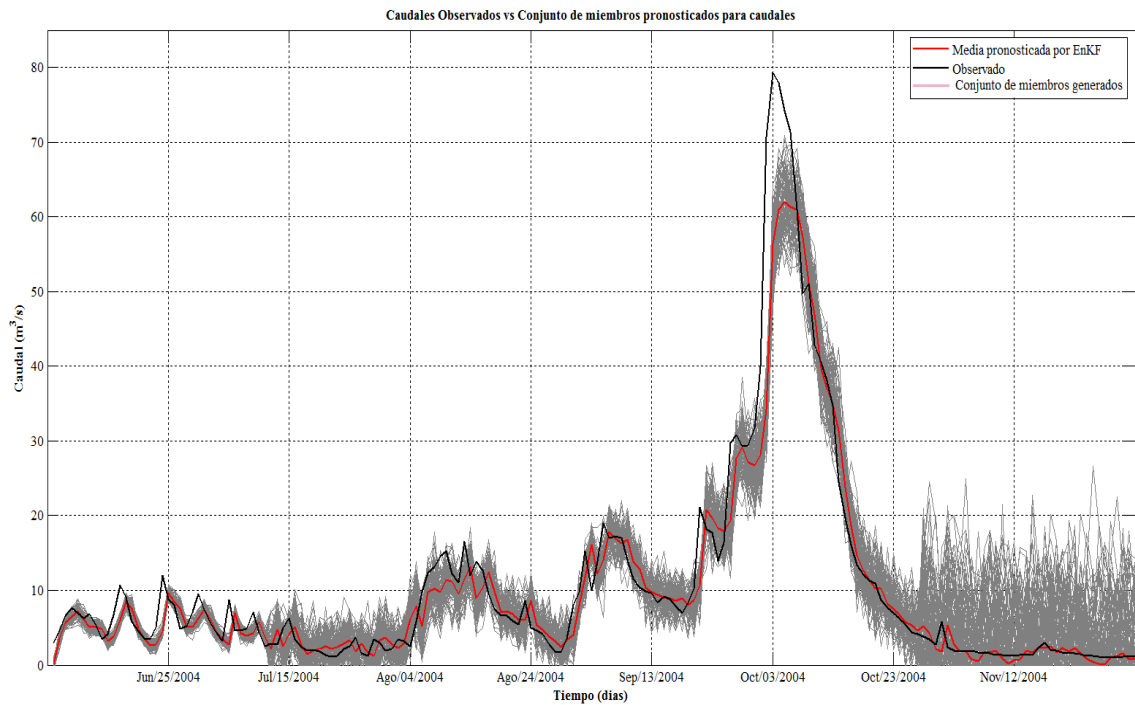


Figura 43. Conjunto de miembros generados para el pronóstico de caudales mediante ARX+EnKF en el río Turbio para la serie año 2004.

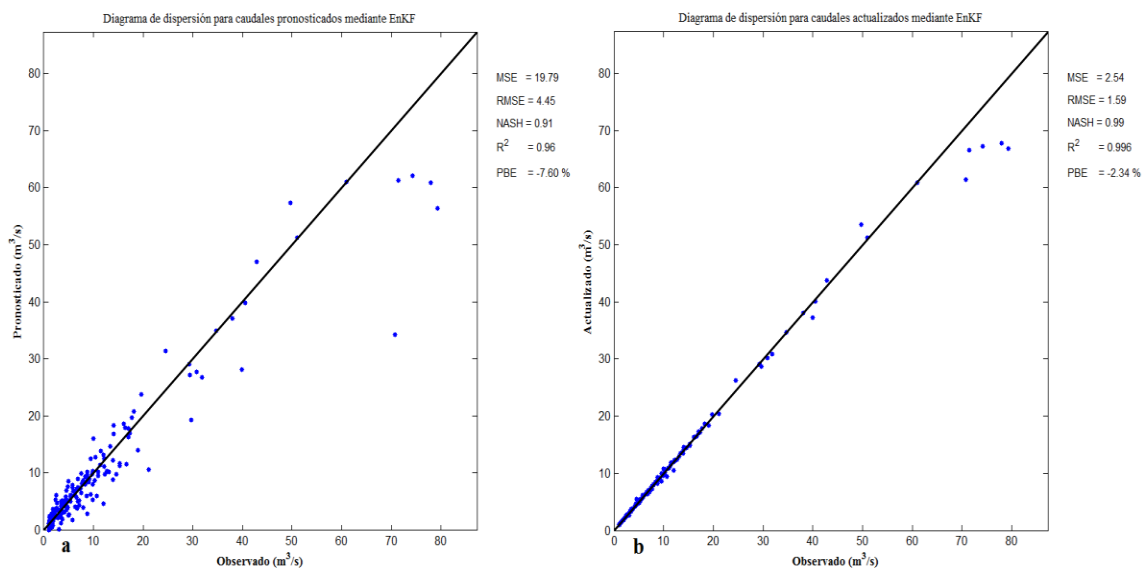


Figura 44. Diagramas de dispersión para la media del conjunto de caudales pronosticados (a) y actualizados (b) mediante ARX+EnKF en la cuenca del río Turbio para la serie del año 2004.

5.2. Cuenca 2: Río Cazonas

5.2.1. Filtro de Kalman Discreto DKF

Con el objeto de contar con un registro más amplio para realizar los pronósticos de caudales utilizando las mismas metodologías que se emplearon para la cuenca del río Turbio, se escogió la cuenca del río Cazonas por contar con registros de datos más amplios para llevar a cabo la propuesta, además de ser una cuenca con mayores registros pluviométricos, que los reportados en la cuenca del río Turbio. Las series de caudal y precipitación que fueron estudiadas en esta cuenca corresponden a 2005 y 2008-2011 obtenidas de las bases de datos BANDAS y CLICOM (CONAGUA, 2014b; SMN, 2010).

En la figura 45, se presenta los pronósticos realizados bajo el DKF para la serie de datos (precipitación y caudal) correspondiente a la serie de datos que abarca los años 2008-2011. El pronóstico se realizó conformando la matriz H con valores de 1 día previo de precipitación y 1 día previo de caudal, estableciendo el valor de $\alpha=0.35$, matriz de covarianza inicial de $P_0 = 1000 \cdot I$ y matriz de covarianza del proceso $[Q]_{n \times n}$ de valor nulo.

De acuerdo a los resultados obtenidos en el cuadro 15, muestran que existe un ajuste satisfactorio en terminos de Nash-Sutcliffe para los caudales pronosticados con valor de 0.69 y la tendencia a mantenerse el valor de la media para cada una de las series pronosticadas, actualizadas y observadas. No existe una tendencia a la subestimación o sobreestimación en los pronosticos según los resultados arrojados por el PBE para el pronóstico de esta serie.

En la figura 46, se presentan los errores a lo largo del pronóstico, los cuales tienden a ser mayores para los caudales de las grandes avenidas en cada periodo de lluvias, según el registro 2008-2011.

Cuadro 15. Resumen estadístico para el pronóstico de caudales en la cuenca río Cazonas para la serie 2008-2011.

Estadísticas	Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
Caudales Pronosticados	59.87	2889.35	53.75	0.84	0.69	0.15
Caudales Actualizados	60.0	2722.68	52.18	0.85	0.70	0.49
Caudales Observados	59.79	-	-	-	-	-

*Las estadísticas son con respecto a los caudales observados

Según Pereyra y Pérez (2005), el año 2005 fue uno de los años más azotados para el estado de Veracruz de manera severa por algunos fenómenos meteorológicos como ondas tropicales, tormentas tropicales y huracanes que provocaron desbordamientos en los principales ríos de la región e inundaciones en las poblaciones aledañas a los mismos. Con estos antecedentes, se implementó un modelo de pronósticos para la serie del año 2005 y de esta manera, observar el comportamiento predictivo en las fechas en que ocurrieron los eventos antes mencionados, utilizando los diferentes enfoques (DKF, ARX+DKF, ARX+EnKF). Los resultados obtenidos para el DKF se muestran en las figuras 45, 46 y 47 y en el cuadro 16.

Cuadro 16. Resumen estadístico para el pronóstico de caudales en la cuenca del río Cazonas para la serie 2005.

Estadísticas	Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
Caudales Pronosticados	47.2	619.42	24.89	0.92	0.84	-4.18
Caudales Actualizados	47.2	561.13	23.69	0.92	0.85	-4.22
Caudales Observados	49.3	-	-	-	-	-

*Las estadísticas son con respecto a los caudales observados.

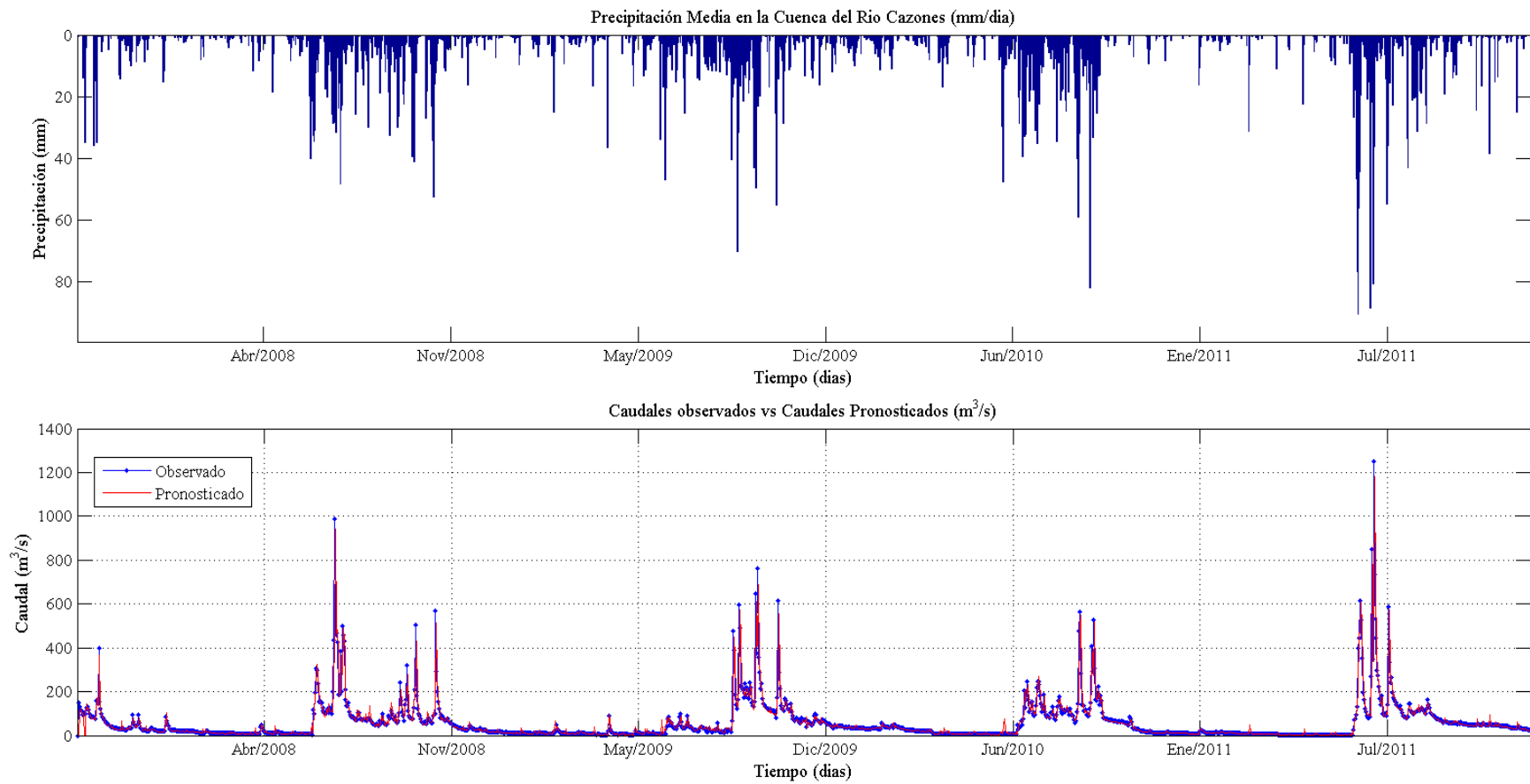


Figura 45. Pronóstico de caudales mediante DKF en la cuenca del río Czones para la serie 2008-2011.

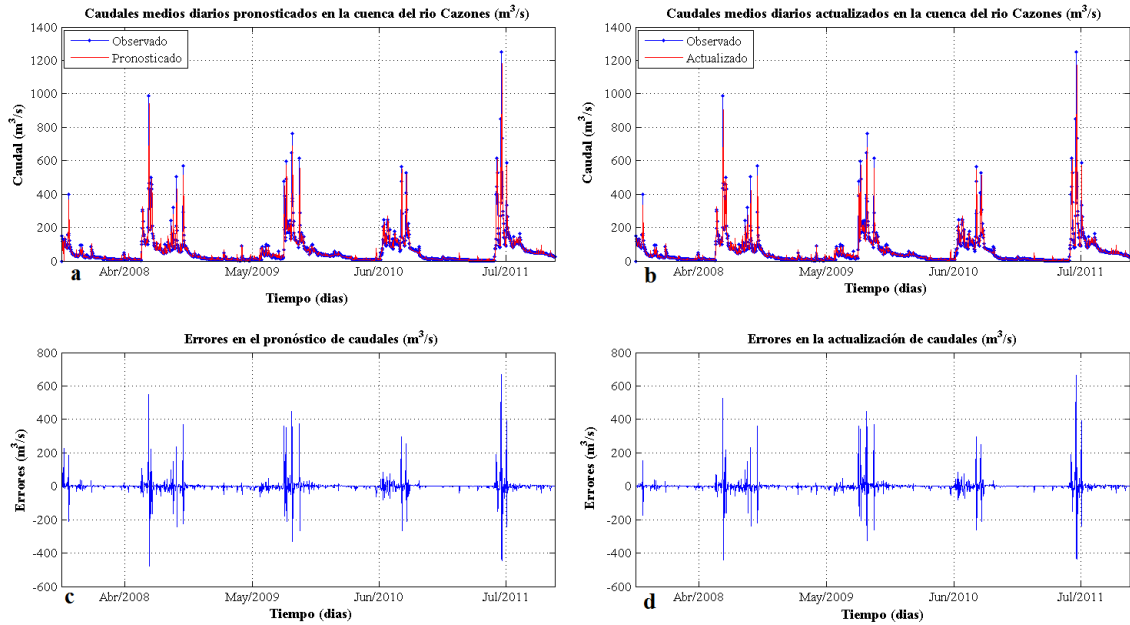


Figura 46. Caudales pronosticados y actualizados (a, b) y sus errores (c, d) obtenidos mediante el DKF en la cuenca del río Cazonces para la serie 2008-2011.

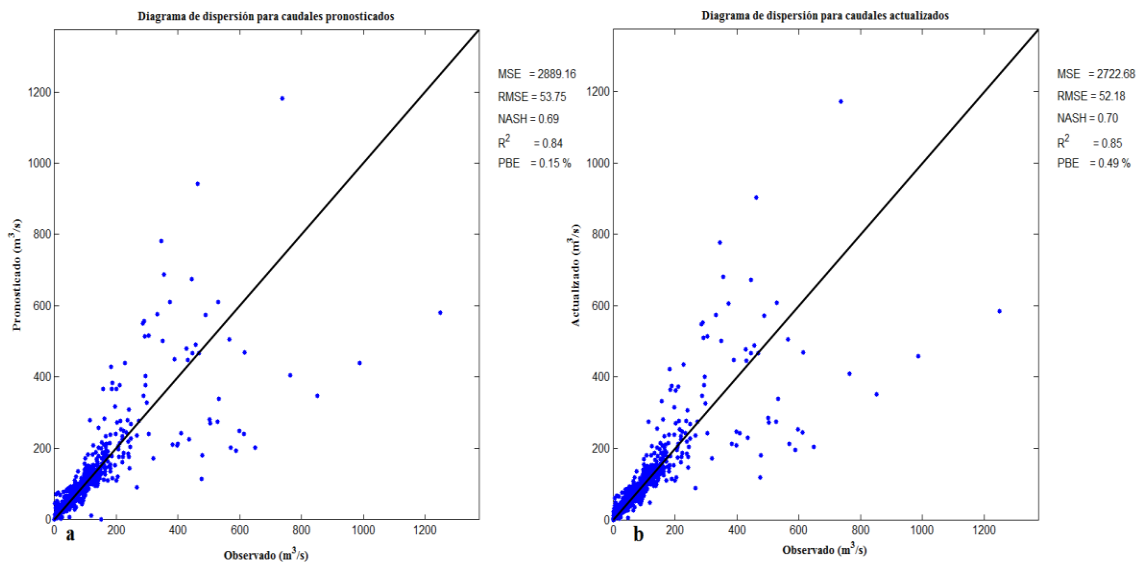


Figura 47. Diagramas de dispersión para caudales pronosticados (a) y actualizados (b) en la cuenca del río Cazonces para la serie 2008-2011.

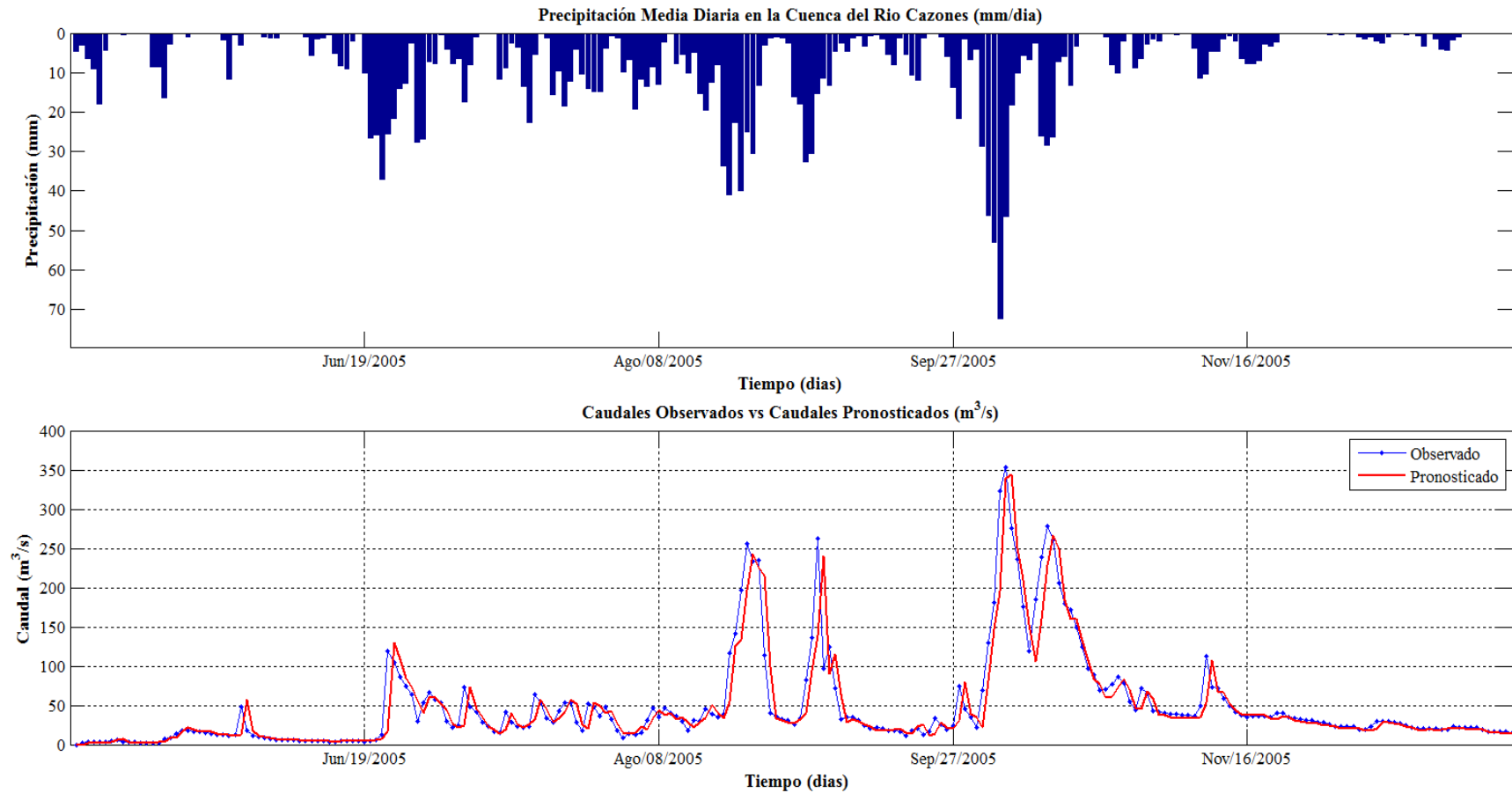


Figura 48. Pronóstico de caudales utilizando el DKF en la cuenca del río Cazonés para la serie del año 2005.

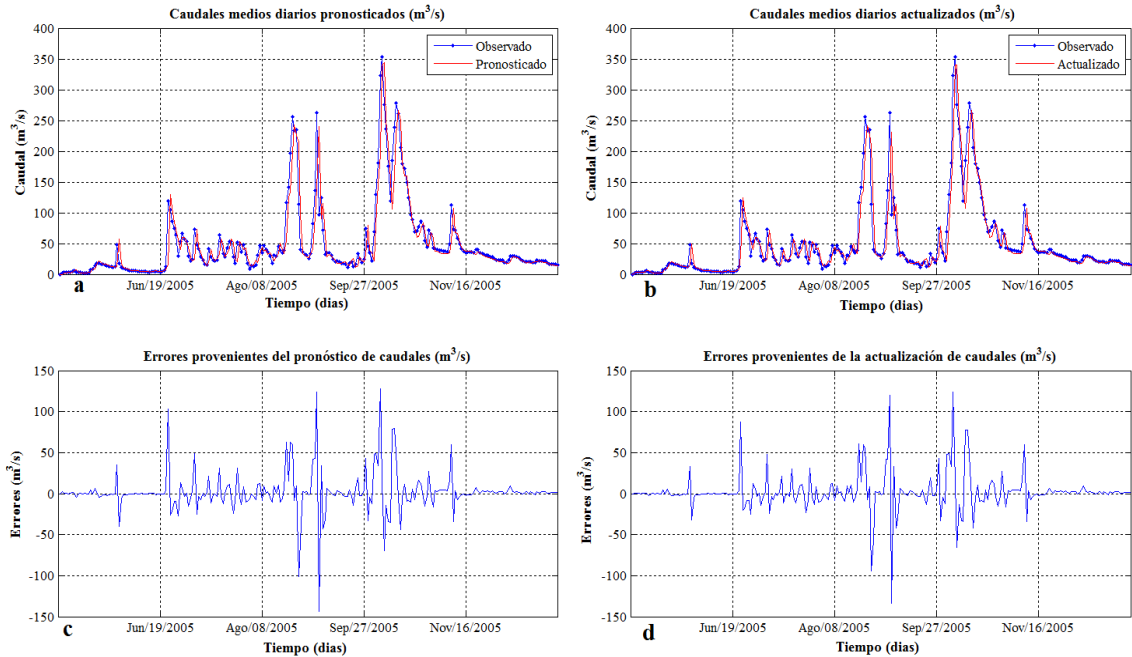


Figura 49. Caudales pronosticados y actualizados (a, b) y sus errores (c, d) obtenidos en la cuenca del río Cazon para la serie del año 2005.

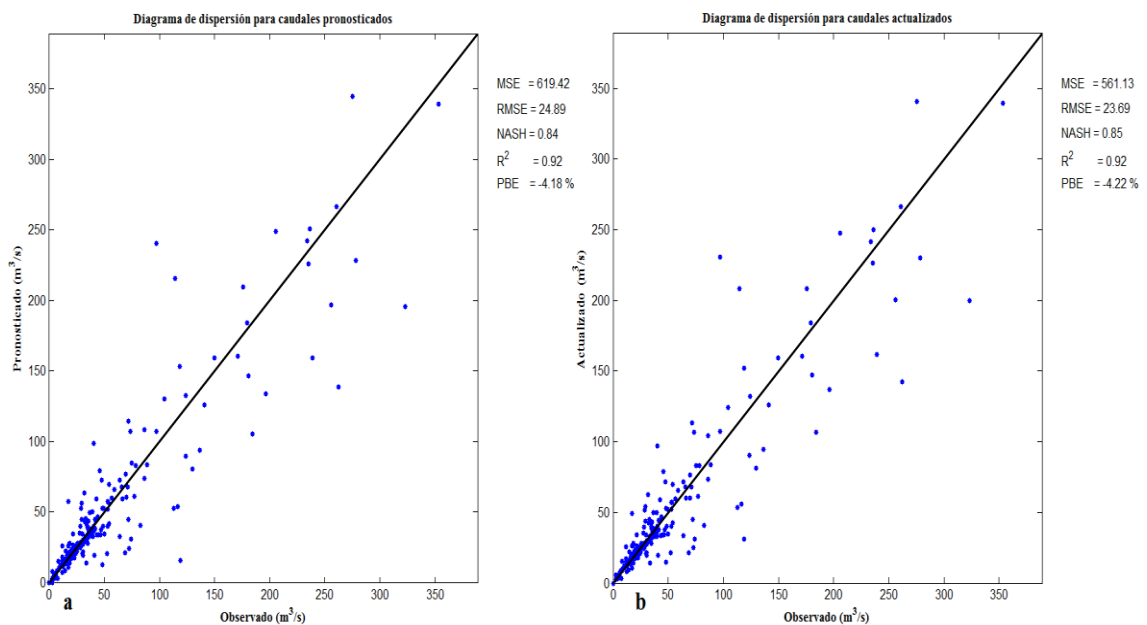


Figura 50. Diagramas de dispersión para caudales pronosticados (a) y actualizados (b) en la cuenca del río Cazon para la serie del año 2005.

5.2.2. Enfoque ARX+DKF

Para la serie 2008-2011 se ejecutó el algoritmo ARX+DKF, el cual se conformó con un periodo de 360 días, con una estructura autorregresiva de orden 2 para la precipitación y 1 para caudal. El periodo escogido aumentó con respecto al promedio de las series anteriores, debido a una longitud de registro obtenida de 1461 días, siendo este periodo el que mejor resultados en términos estadísticos arrojó. Los resultados se resumen en el cuadro 17.

Cuadro 17. Resumen estadístico para diferentes pasos en el tiempo mediante ARX+DKF en la cuenca del río Cazones para la serie 2008-2011.

Estadísticas	Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
Caudales Pronosticados L=1	56.91	1984.91	44.55	0.90	0.81	-3.50
Caudales Pronosticados L=2	55.64	2680.36	51.77	0.86	0.74	-5.49
Caudales Pronosticados L=3	54.80	2841.68	53.31	0.85	0.72	-6.35
Caudales Pronosticados L=4	54.37	3036.20	55.10	0.84	0.70	-7.05
Caudales Observados	58.97	-	-	-	-	-

*Las estadísticas son con respecto a los caudales observados.

De acuerdo a los resultados del cuadro 17, los ajustes del pronóstico para un paso adelante en términos de Nash-Sutcliffe son buenos, registrando un valor de 0.81 para toda la serie pronosticada. El RMSE y el coeficiente de correlación para los pronósticos disminuyen a medida que se avanza L pasos en el tiempo, además de preservarse el valor de la media en cada serie pronosticada con respecto a la observada. En términos generales, los pronósticos en todos los pasos del tiempo tienden a ser subestimados por el modelo implementado. La tendencia de presentar grandes diferencias entre los caudales pronosticados y observados para las máximas avenidas se sigue presentando tal como se puede observar en la figuras 51 y 52.

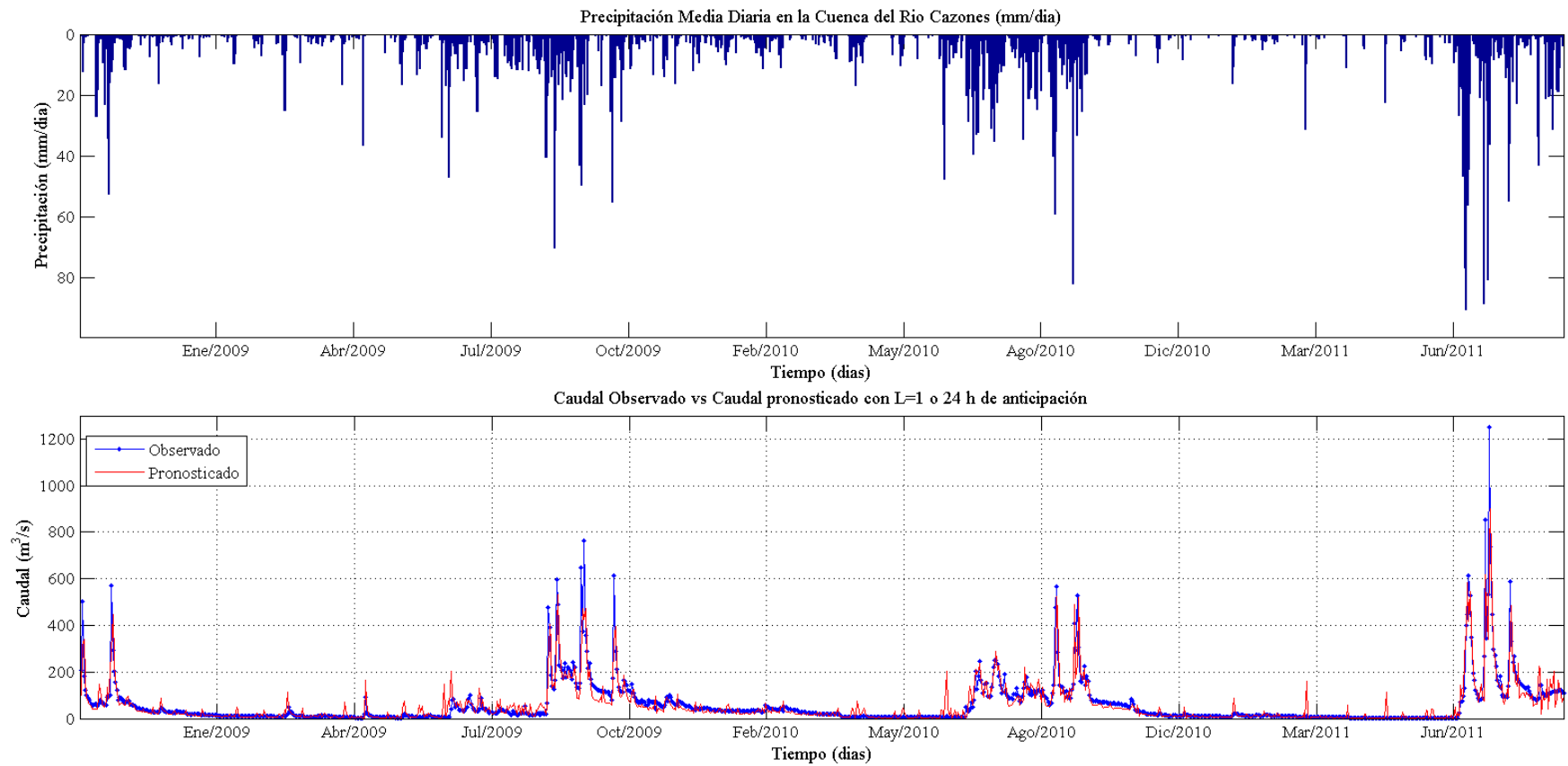


Figura 51. Pronóstico de caudales mediante ARX+DKF en la cuenca del río Cazones para la serie 2008-2011.

Los errores generados a través del tiempo para cada predicción, son mostrados en la figura 52, confirmando que éstos tienden a ser mayores a medida que el pronóstico se realiza L pasos adelante $x_{t+L|t}$, condicionados por un valor previo con el que se ha realizado el pronóstico, antes de ser actualizados por el filtro de Kalman para el nuevo ciclo.

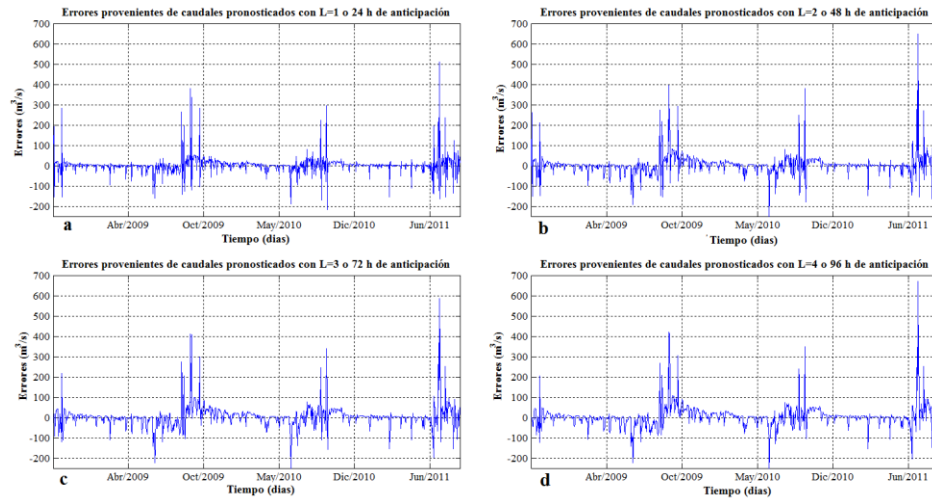


Figura 52. Errores generados para diferentes pasos en el tiempo en la cuenca del río Cazones para la serie 2008-2011.

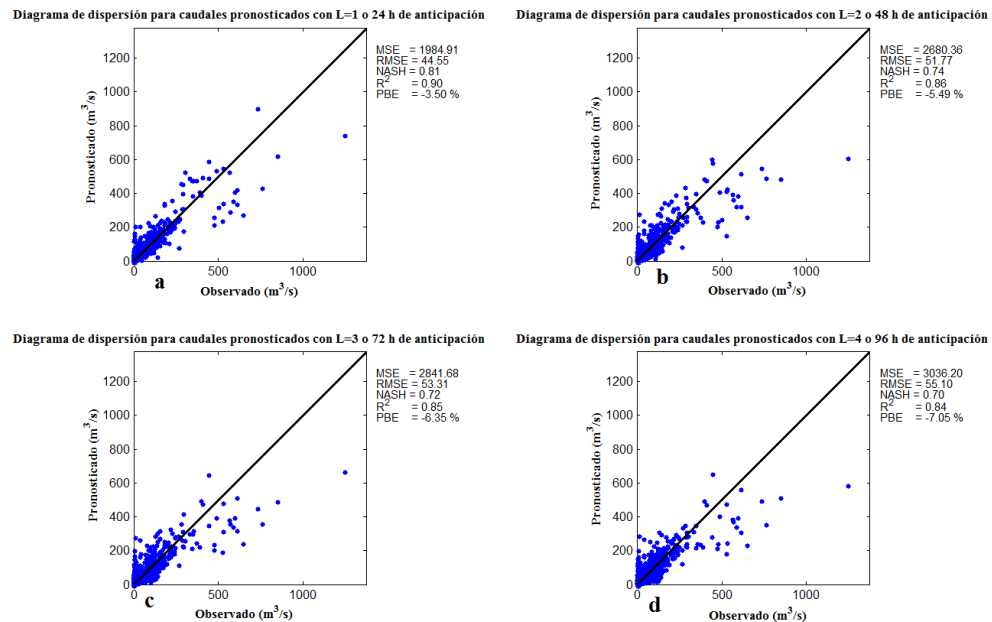


Figura 53. Diagramas de dispersión comparativos para pronósticos a diferentes pasos en el tiempo en la cuenca del río Cazones para la serie 2008-2011.

En la figura 54 se presentan los errores generados en el modelo de pronóstico para la cuenca de río Cazonos utilizando la serie 2008-2011 con la característica que los errores siguen la tendencia de ajuste a distribuciones de tipo t student escalada debido a los valores atípicos que inciden en colas pesadas en los extremos de la distribución.

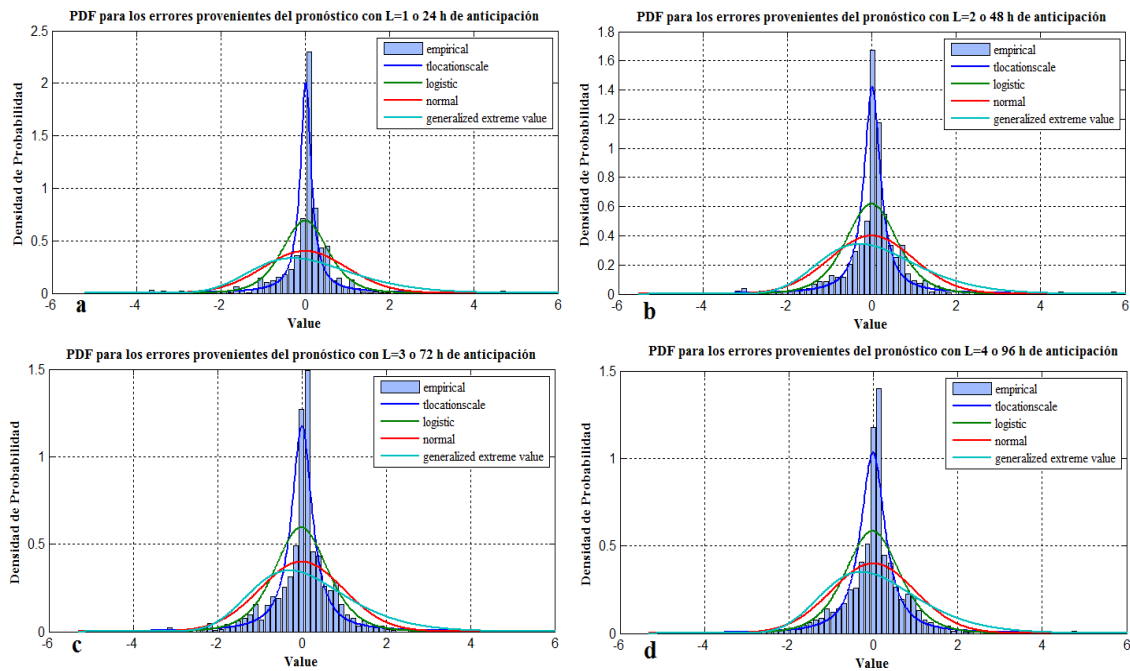


Figura 54. Funciones de distribución de probabilidad para los errores de predicción con diferentes pasos en el tiempo en la cuenca del río Cazonos para la serie 2008-2011.

En particular, la varianza para los errores en el pronóstico con $L=1$ es menor que para los errores cuando se tiene $L=4$. A continuación, se presenta un box plot sustentando los resultados descritos anteriormente.

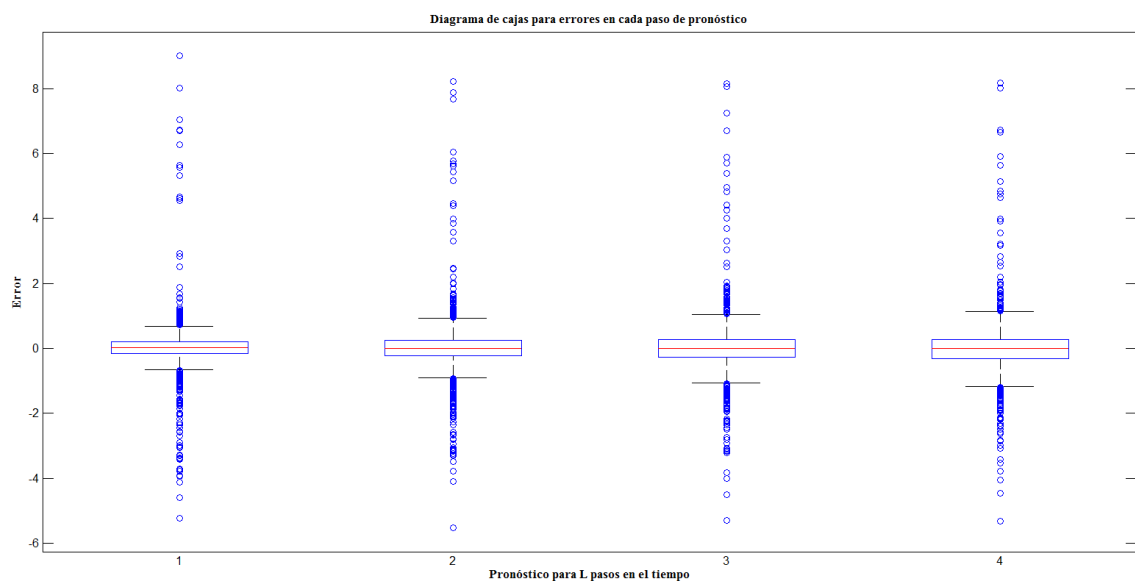


Figura 55. Diagrama de cajas para la distribución de los errores para diferentes pasos en el tiempo en la cuenca del río Cazonos para la serie 2008-2011.

Se analizaron los resultados obtenidos en la figura 53, agrupando la información en dos clases según la tendencia de los caudales pronosticados contra los observados para $L=1$ gráficas (a, b), $L=2$ gráficas (c, d), $L=3$ gráficas (e, f), $L=4$ gráficas (g, h). En el cuadro 18, se resumen las estadísticas para cada paso en el tiempo teniendo en cuenta las clases $Q < 300 \text{ m}^3/\text{s}$ y $Q \geq 300 \text{ m}^3/\text{s}$.

Cuadro 18. Resumen estadístico para los caudales agrupados en la cuenca del río Cazonos para la serie 2008-2011.

Pronósticos	Rango	MSE	RMSE	R^2	Nash-Sutcliffe	PBE
L=1 24 h	$Q < 300 \text{ m}^3/\text{s}$	787.76	28.07	0.88	0.75	0.83
	$Q \geq 300 \text{ m}^3/\text{s}$	45371.39	213.01	0.37	-0.26	-16.98
L=2 48 h	$Q < 300 \text{ m}^3/\text{s}$	1228.45	35.05	0.83	0.6	2.2
	$Q \geq 300 \text{ m}^3/\text{s}$	55299.31	235.16	0.44	-0.54	-29.39
L=3 72 h	$Q < 300 \text{ m}^3/\text{s}$	1365.32	36.95	0.80	0.56	1.88
	$Q \geq 300 \text{ m}^3/\text{s}$	58310.61	241.48	0.53	-0.57	-32.84
L=4 96 h	$Q < 300 \text{ m}^3/\text{s}$	1443.57	37.99	0.78	0.53	1.16
	$Q \geq 300 \text{ m}^3/\text{s}$	62873.74	250.75	0.46	-0.69	-33.47

De los resultados obtenidos para los datos agrupados por clases, se puede concluir que el modelo de pronóstico para caudales mediante el ARX+DKF en la cuenca del río Cazones, se ajusta satisfactoriamente para caudales que se encuentran por debajo de los 300 m³/s con una tendencia a sobreestimar los caudales pronosticados cuando se realiza para L=4 o 96 h de anticipación.

De la misma manera, para aquellos pronósticos de caudales que son iguales o mayores a los 300 m³/s la tendencia es a la subestimación, presentando grados mínimos de ajuste con respecto a los valores observados reflejados a través del coeficiente de Nash-Sutcliffe, pues a medida que se avanza L pasos con el pronóstico la eficiencia tiende cada vez a valores negativos más grandes, indicando que la media de los valores observados puede ser usada como el mejor predictor que los valores predichos por el modelo implementado.

A diferencia de la cuenca del río Turbio, la cuenca del río Cazones está expuesta a la ocurrencia de fenómenos meteorológicos ocasionados por ondas tropicales, incluso huracanes, incrementando la probabilidad de que se presenten avenidas de mayor consideración lo cual incide en el modelo de pronóstico implementado, debido a que los datos de entrada son valores extremos que al ser modelados a través de un proceso autorregresivo, no se ajustan a las condiciones reales que normalmente se presentan dada la serie de tiempo, perdiendo confiabilidad y obteniendo valores subestimados que incluso pueden no pertenecer a un intervalo de predicción a un nivel probabilidad dada.

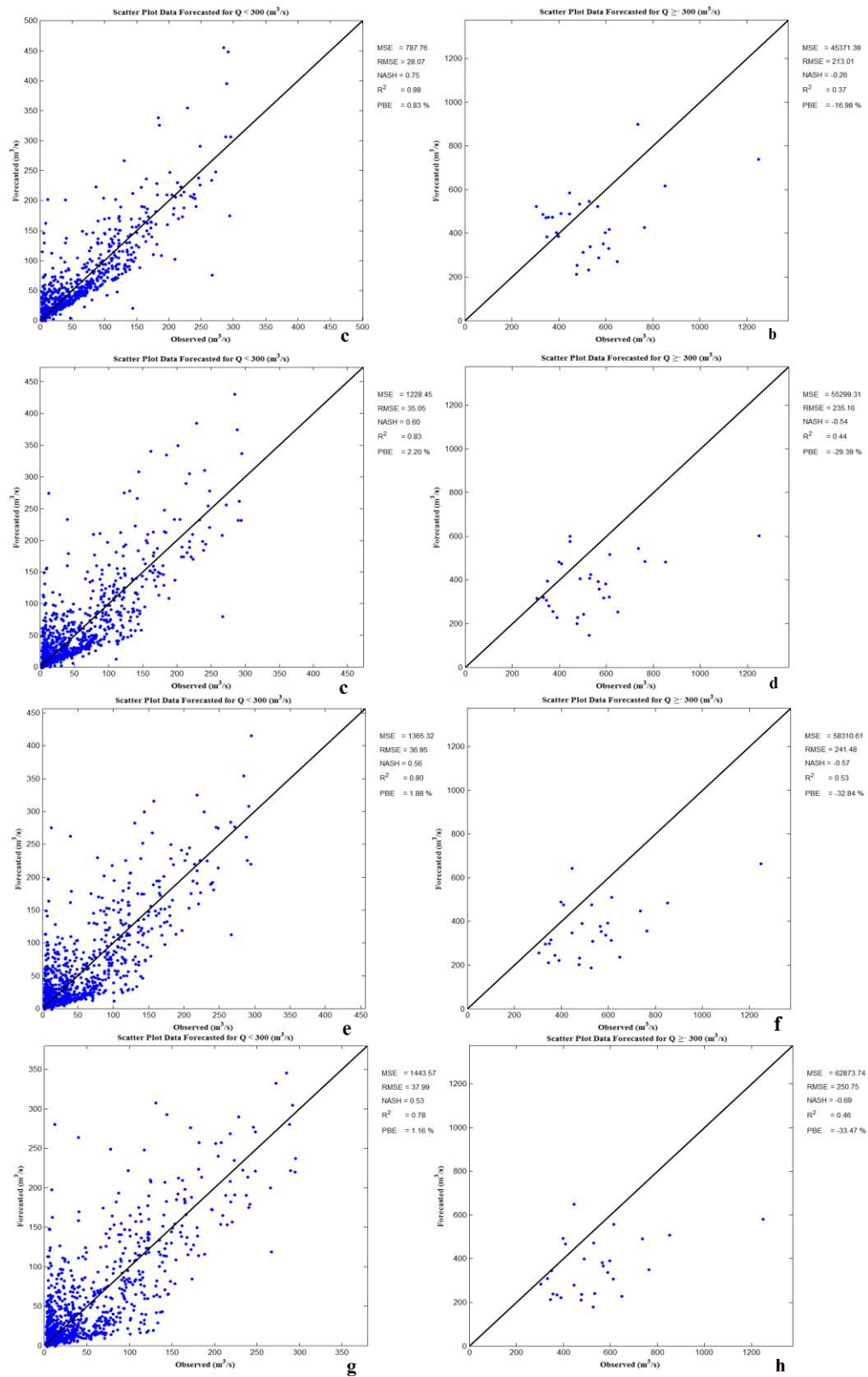


Figura 56. Diagramas de dispersión por clases para los pronósticos en diferentes pasos en el tiempo en la cuenca del río Czones para la serie 2008-2011.

Con los errores ajustados a una función de distribución t student escalada para $L=1$ y conocimiento sus parámetros (μ, σ, ν) , se calculó el intervalo de predicción para la serie 2008-2011 con una probabilidad del 95% como se muestra en la figura 58.

En términos generales, el modelo de pronóstico para la cuenca del río Cazones es aceptable, haciendo énfasis en que dada la serie 2008-2011 el modelo tiende a presentar mejores ajustes para los caudales menores a los $300 \text{ m}^3/\text{s}$, un límite superior a la media de los valores registrados en dicha serie con $58.9 \text{ m}^3/\text{s}$, siendo el valor máximo $1249.6 \text{ m}^3/\text{s}$ provocado por el avance de las lluvias de la tormenta tropical Arlene en el año 2011.

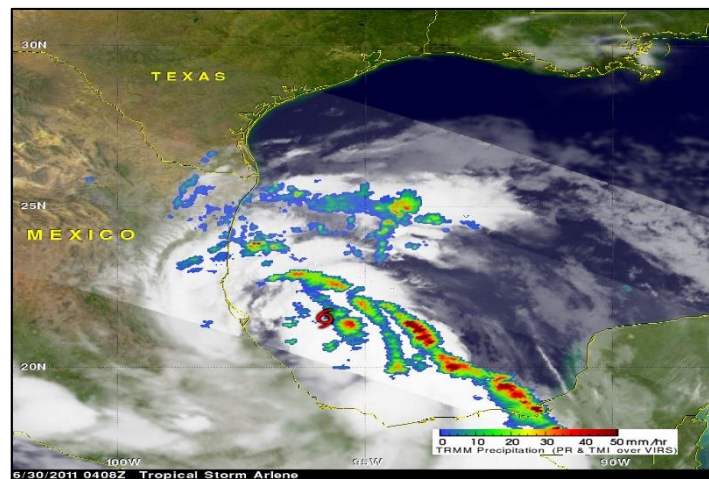


Figura 57. Avance de la tormenta tropical Arlene el 30/Junio/2011 sobre la costa del Golfo de México. Fuente: NASA/SSAI, Hal Pierce (2014).

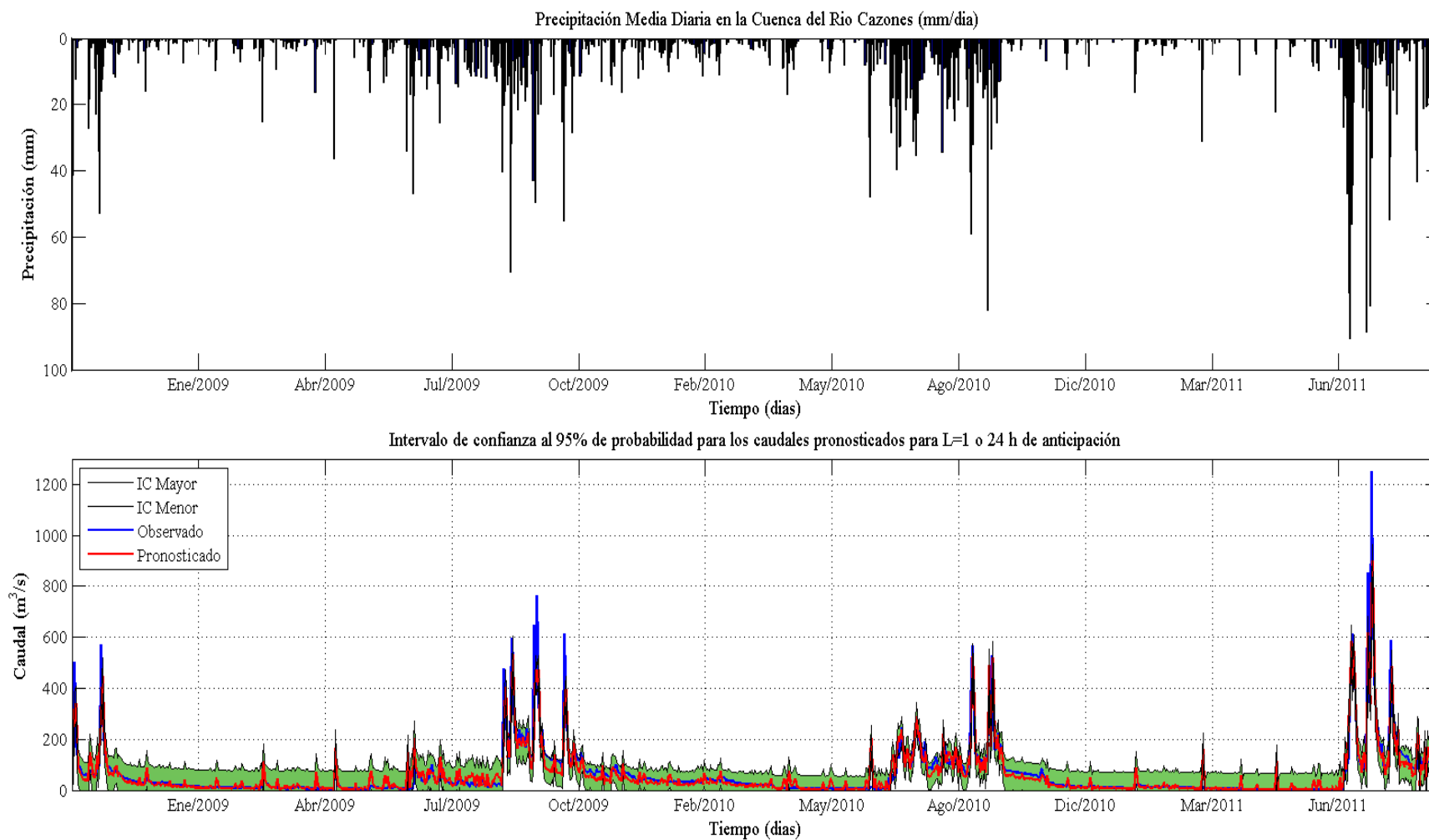


Figura 58. Intervalo de predicción al 95% de probabilidad para L=1 o 24 h en la cuenca del río Cuzco para la serie 2008-2011.

Con el propósito de evaluar los pronósticos para la serie del año 2005, en el cuadro 19 se resumen las estadísticas obtenidas luego de la implementación del modelo ARX+DKF con un periodo de 26 días para calibración y un orden autorregresivo de 2 para precipitación y 1 para caudal (Ver figuras 59 y 60).

Cuadro 19. Resumen estadístico para diferentes pronósticos mediante ARX+DKF en la cuenca del río Cazonas para la serie del año 2005.

Estadísticas	Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
Caudales Pronosticados L=1	51.25	640.21	25.30	0.92	0.84	-8.65
Caudales Pronosticados L=2	48.04	1018.09	31.91	0.88	0.75	-14.45
Caudales Pronosticados L=3	46.10	1202.45	34.68	0.86	0.71	-17.98
Caudales Pronosticados L=4	44.78	1286.02	35.86	0.85	0.69	-20.15
Caudales Observados	56.10	-	-	-	-	-

*Las estadísticas son con respecto a los caudales observados.

Los mejores resultados para el pronóstico de caudales en la serie 2005 son para L=1 o con 24 h de anticipación, tendencia que se ha mantenido en todas las series utilizadas tanto para la cuenca del río Turbio, como para la cuenca del río Cazonas.

Según Kim (2004), quien realizó la implementación del filtro de Kalman con el modelo CDRMV3 para el pronóstico de caudales, encontró que los mejores resultados pertenecían a los pronósticos realizados a 1 hora que los obtenidos a 12 horas después, evaluados en términos de RMSE. Por otra parte, Hirpa *et al.* (2013) encontró que en términos del índice de eficiencia de Nash-Sutcliffe, los mejores resultados fueron para los pronósticos de caudales realizados a 1 día que para aquellos que fueron realizados a 15 días, en el río Ganges y Brahmaputra de Asia meridional.

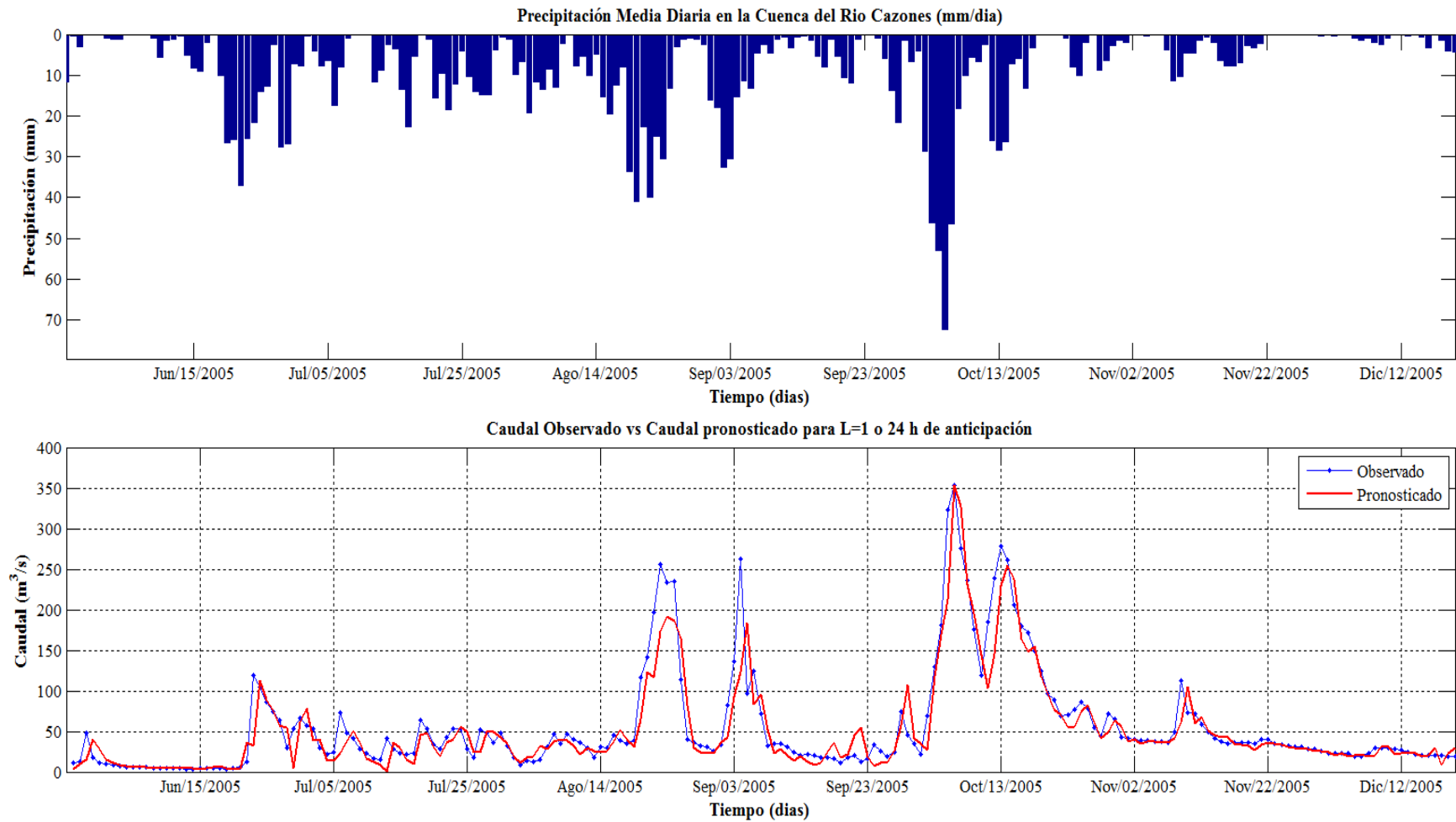


Figura 59. Pronóstico de caudales mediante ARX+DKF en la cuenca del río Cazonés para la serie del año 2005.

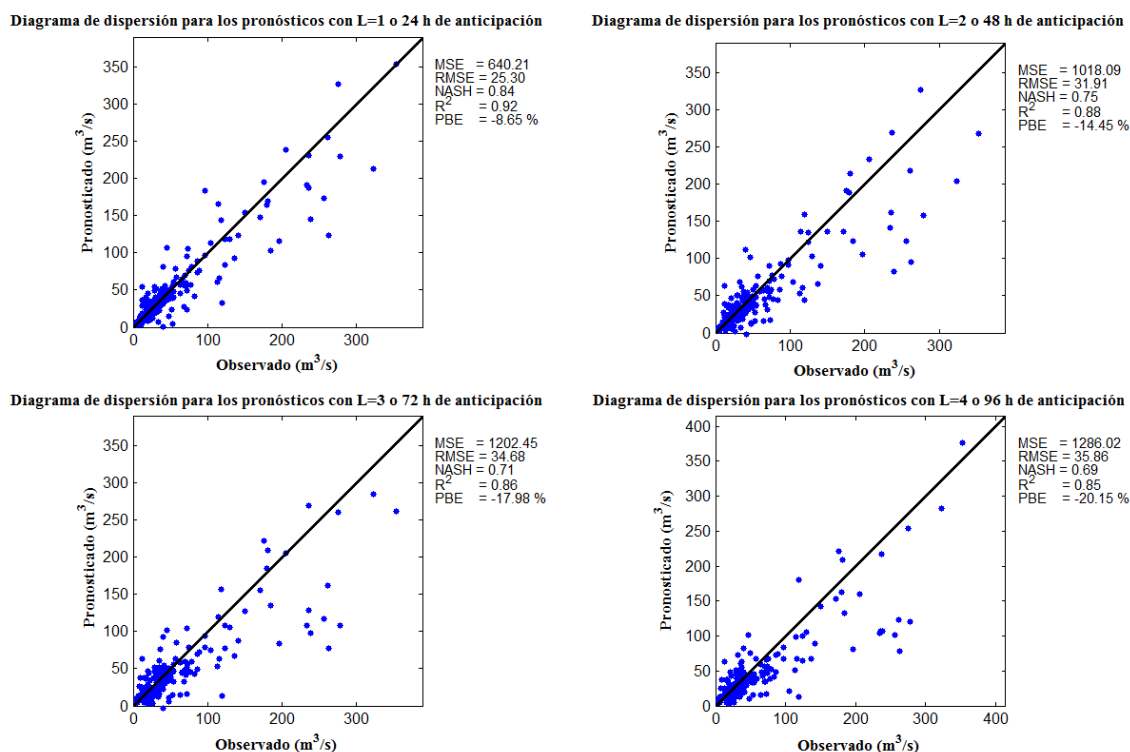


Figura 60. Diagramas de dispersión comparativos para pronósticos a diferentes pasos hacia adelante en la cuenca del río Cazones para la serie del año 2005.

Los caudales pronosticados en la serie del año 2005, fueron agrupados en dos clases, la primera para los $Q < 100 \text{ m}^3/\text{s}$ y la segunda para los $Q \geq 100 \text{ m}^3/\text{s}$.

Cuadro 20. Resumen estadístico para los pronósticos agrupados por clases en la cuenca del río Cazones para la serie 2008-2011.

Pronósticos	Rango	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
L=1 24 h	$Q < 100 \text{ m}^3/\text{s}$	231.65	15.22	0.78	0.47	-1.81
	$Q \geq 100 \text{ m}^3/\text{s}$	3064.31	55.36	0.78	0.33	-15.73
L=2 48 h	$Q < 100 \text{ m}^3/\text{s}$	262.71	16.21	0.71	0.39	-6.2
	$Q \geq 100 \text{ m}^3/\text{s}$	5500.05	74.16	0.62	-0.20	-23
L=3 72 h	$Q < 100 \text{ m}^3/\text{s}$	281.87	16.79	0.68	0.35	-8.19
	$Q \geq 100 \text{ m}^3/\text{s}$	6664.56	81.64	0.60	-0.45	-28.15
L=4 96 h	$Q < 100 \text{ m}^3/\text{s}$	280.55	16.75	0.66	0.35	-10.20
	$Q \geq 100 \text{ m}^3/\text{s}$	7251.76	85.16	0.64	-0.58	-30.44

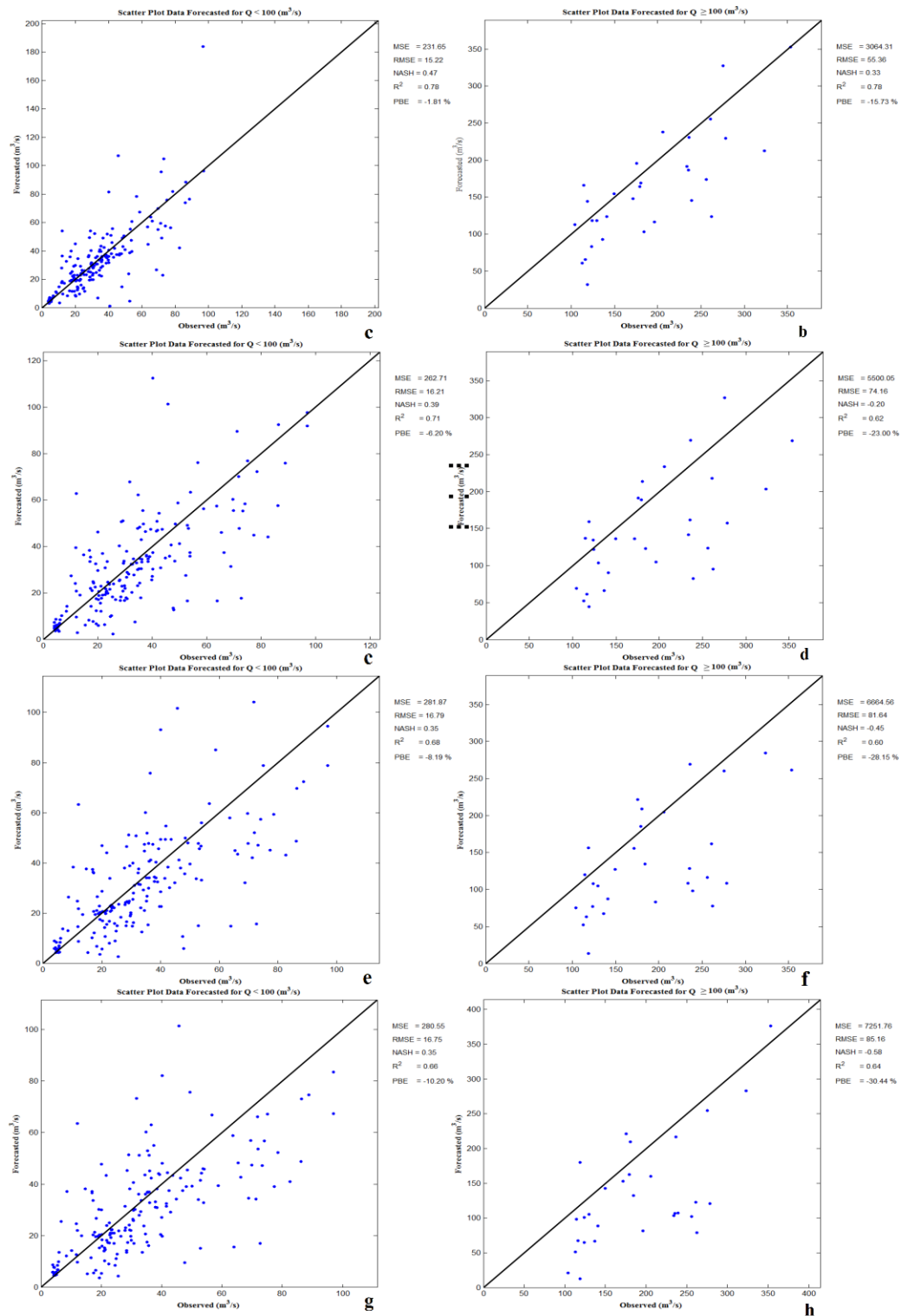


Figura 61. Diagramas de dispersión por clases para los pronósticos con diferentes pasos en el tiempo para la cuenca del río Cazon utilizando la serie del año 2005.

El orden de los diagramas de dispersión en la figura 61 es para $L=1$ gráficas (a, b), para $L=2$ gráficas (c, d), para $L=3$ gráficas (e, f) y para $L=4$ gráficas (g, h).

El ajuste de los errores a una distribución de probabilidad para la serie 2005 se presentan a continuación, con la tendencia a aumentar la varianza para la distribución de los errores con $L=4$. Todas las distribuciones presentadas en la figura 62 corresponden a una *t* student escalada.

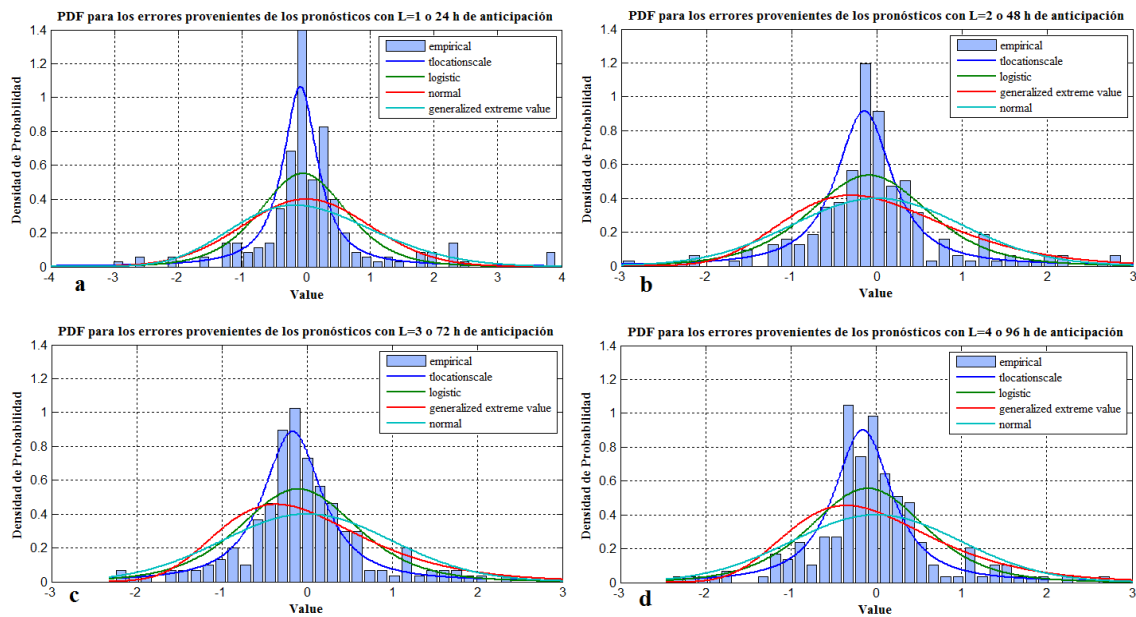


Figura 62. Funciones de distribución de probabilidad para los errores de predicción con diferentes pasos en el tiempo en la cuenca del río Czones para la serie 2005.

En la figura 63, se presentan los pronósticos de caudales para la serie del año 2005. Se pueden observar dos eventos importantes registrados entre el 15 al 25 de Agosto de 2005, conocido como la tormenta tropical José y del 1 al 7 de octubre conocido como el huracán Stan. En general, los pronósticos son aceptables aunque el intervalo de predicción presenta un menor grado de incertidumbre para el evento registrado por el huracán Stan.

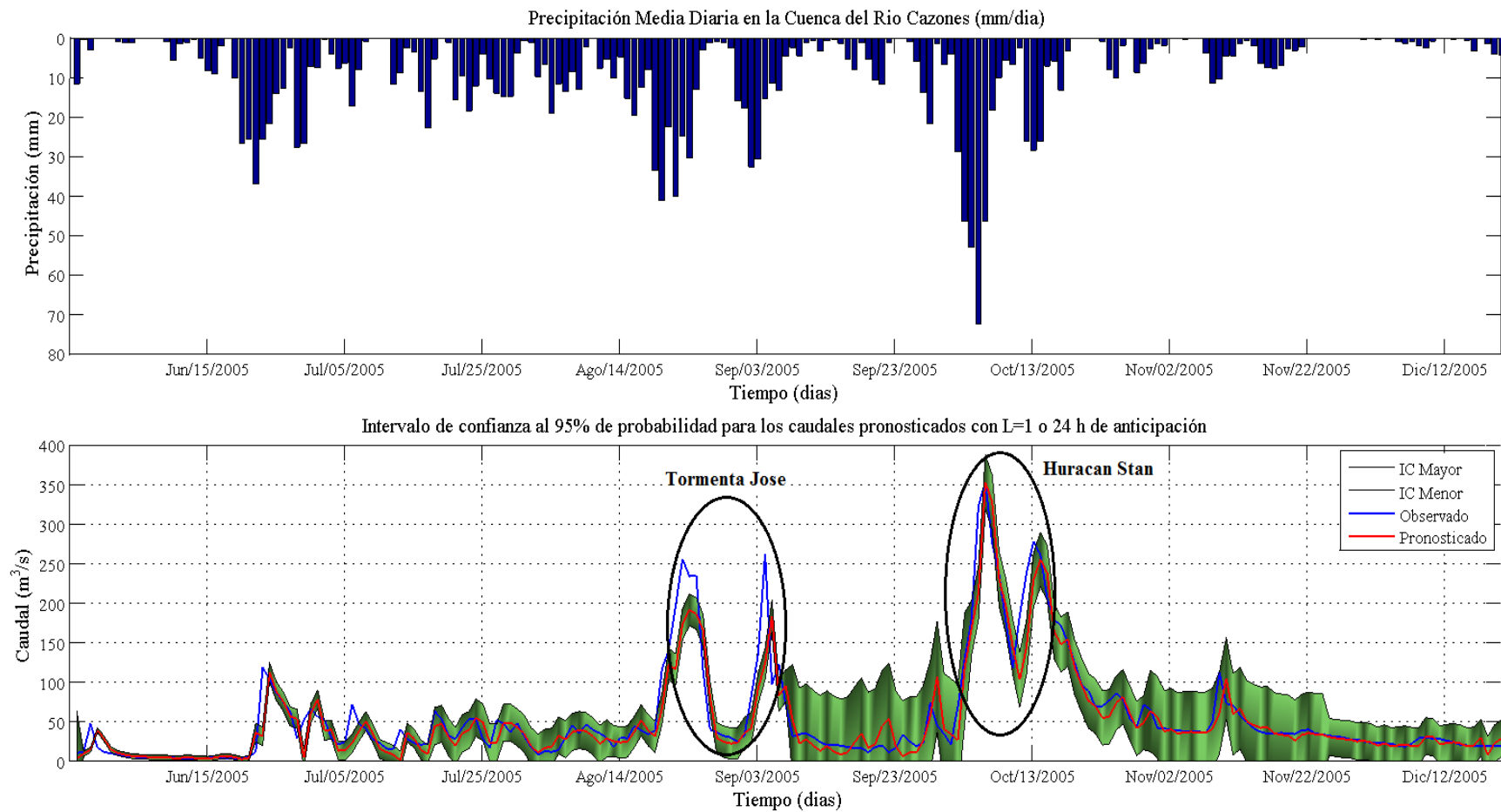


Figura 63. Intervalo de predicción al 95% de probabilidad para L=1 o 24 h en la cuenca del río Cazones para la serie del año 2005.

Stan fue un huracán de categoría 1 que tocó tierra en la costa sur del Golfo de México y llevaba consigo vientos fuertes y lluvias habituales asociados a los huracanes de baja intensidad. A su paso por Centroamérica y sobre la Península de Yucatán causó muertes y destrucción, así como deslizamientos de tierra provocados por las lluvias durante varios días en Ciudad de México, a más de 400 kilómetros de distancia del Golfo de México, como consecuencia de la banda exterior de lluvia del huracan (NASA/GSFC, 2014).

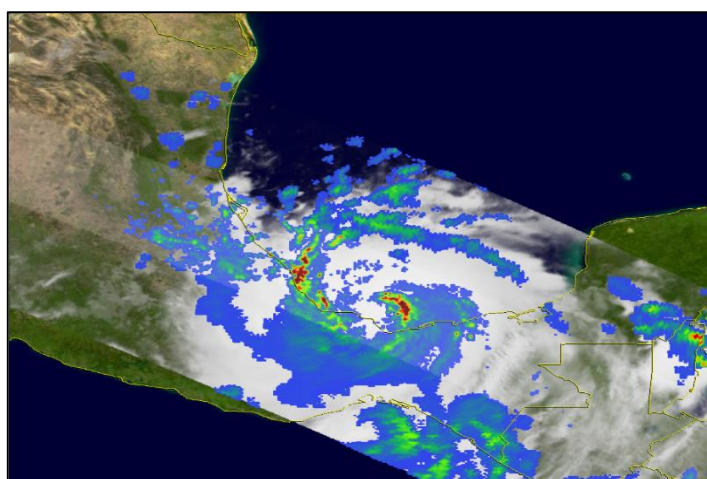


Figura 64. Avance del Huracán Stan sobre la costa del Golfo de México. Fuente: NASA/SSAI, Hal Pierce (2014).

5.2.3. Enfoque ARX+EnKF

La figura 65, muestra el pronóstico realizado para la serie 2008-2011 utilizando el ARX+EnKF. Para este caso, el algoritmo se conformó con un periodo de 360 días, con orden 2 para los coeficientes de caudal y 2 para de lluvia en el modelo ARX. Se generó aleatoriamente un conjunto de 100 miembros o trayectorias posibles de los estados como requerimiento inicial para la operación del EnKF. Los resultados se resumen en el cuadro 21.

Cuadro 21. Resumen estadístico para el pronóstico mediante ARX+EnKF en la cuenca del río Cazones para la serie 2008-2011.

Estadísticas	Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE %
Q Pronosticados EnKFs	57.77	1949.51	44.15	0.90	0.81	-2.04
Q Actualizados EnKF	59.11	44.82	6.69	0.99	0.99	0.23
Q Observados	58.97	-	-	-	-	-

*Las estadísticas son con respecto a los caudales observados

*Los Q pronosticados hacen referencia a la media de los miembros generados

La media del conjunto de miembros generados tiende a ajustarse a los caudales observados con valores de Nash-Sutcliffe de 0.81 tal como se observa en la figura 65, con algunos cambios abruptos en los caudales máximos provocados por los impulsos de lluvia asimilados y que en la convolución hace que el valor pronosticado tienda a estar por encima del real en algunas casos o por debajo del mismo, situación común que se ha presentado en todas las series evaluadas.

En la figuras 68, 69 y 70 se presentan los resultados de los pronósticos para la serie 2005 utilizando el ARX+EnKF con un periodo de 26 días y orden 2 para el modelo autorregresivo. Los resultados muestran un buen ajuste para la media de los estados con la tendencia a subestimar los pronósticos. En el cuadro 22 se presenta el resumen de las estadísticas.

Cuadro 22. Resumen estadístico para el pronóstico mediante ARX+EnKF en la cuenca del río Cazones para la serie del año 2005.

Estadísticas	Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE %
Q Pronosticados EnKFs	52.1	649.12	25.48	0.92	0.84	-7.15
Q Actualizados EnKF	55.4	59.90	7.73	0.99	0.98	-1.23
Q Observados	56.1	-	-	-	-	-

*Las estadísticas son con respecto a los caudales observados.

*Los Q pronosticados hacen referencia a la media de los miembros generados.

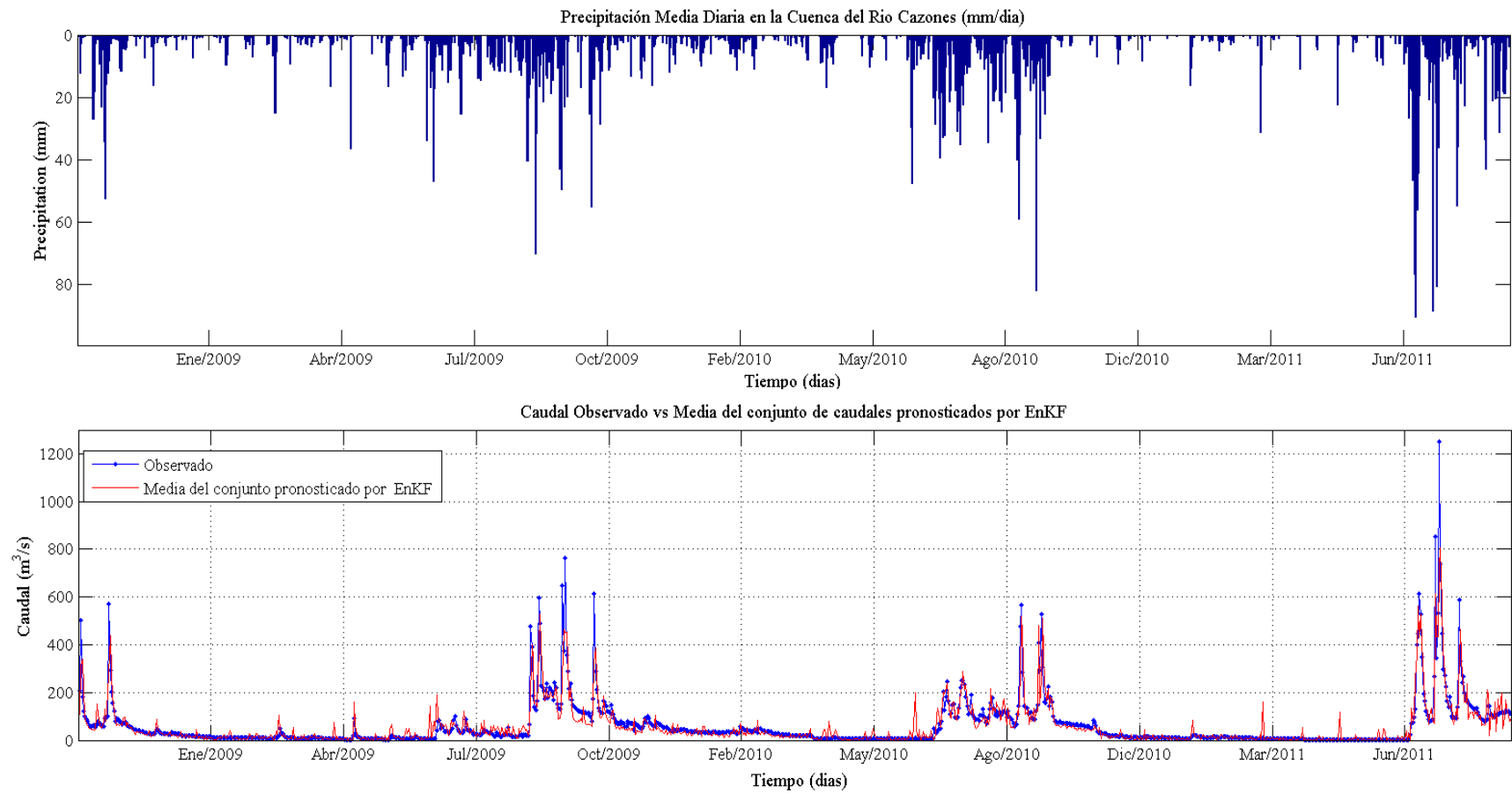


Figura 65. Pronóstico de caudales mediante el ARX+EnKF en la cuenca del río Cazones para la serie 2008-2011.

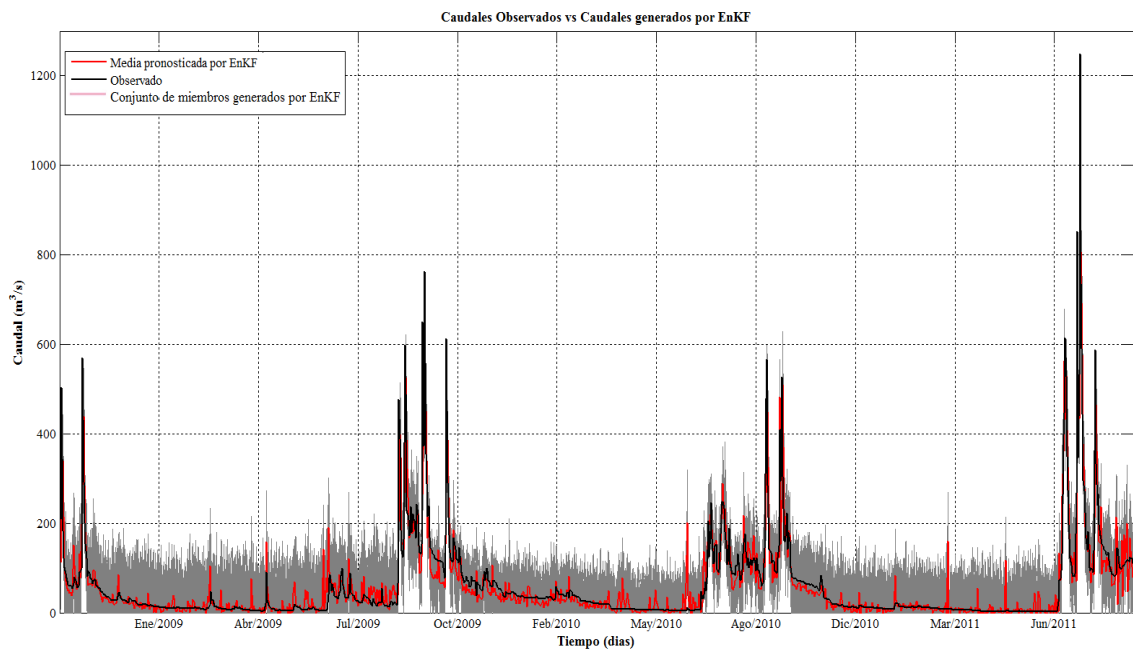


Figura 66. Conjunto de miembros generados para el pronóstico caudales mediante ARX+EnKF en la cuenca del río Cazones para la serie 2008-2011.

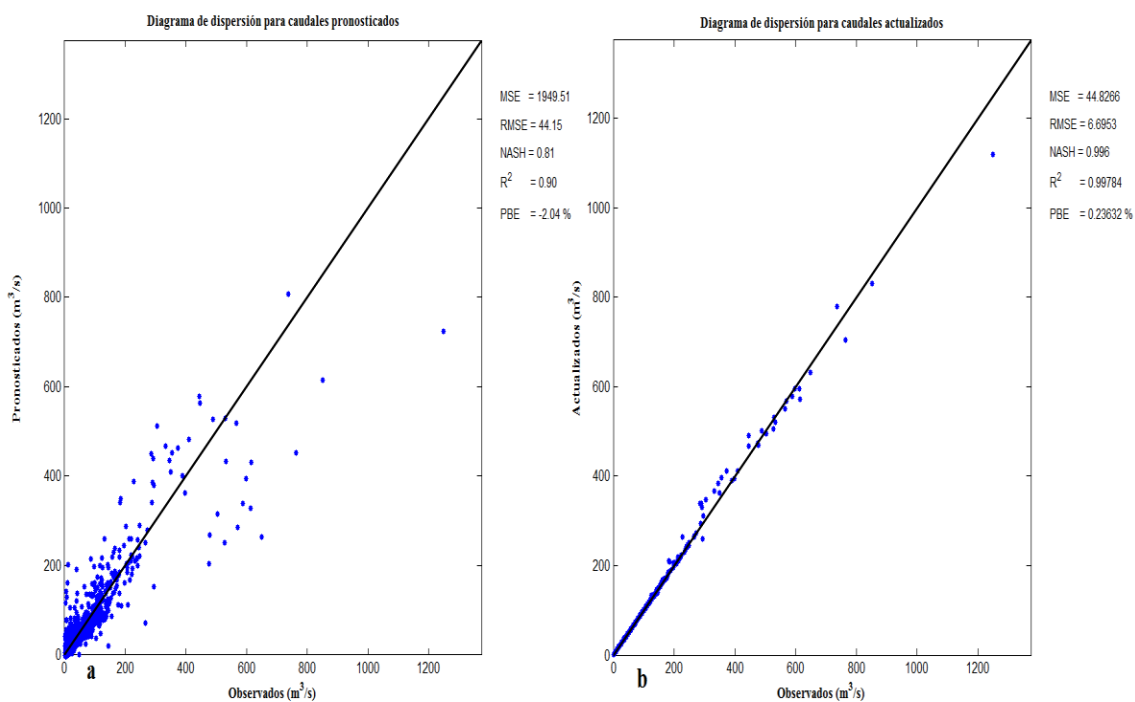


Figura 67. Diagrama de dispersión para la media del conjunto de caudales pronosticados (a) y actualizados (b) mediante el ARX+EnKF en la cuenca del río Cazones para la serie 2008-2011.

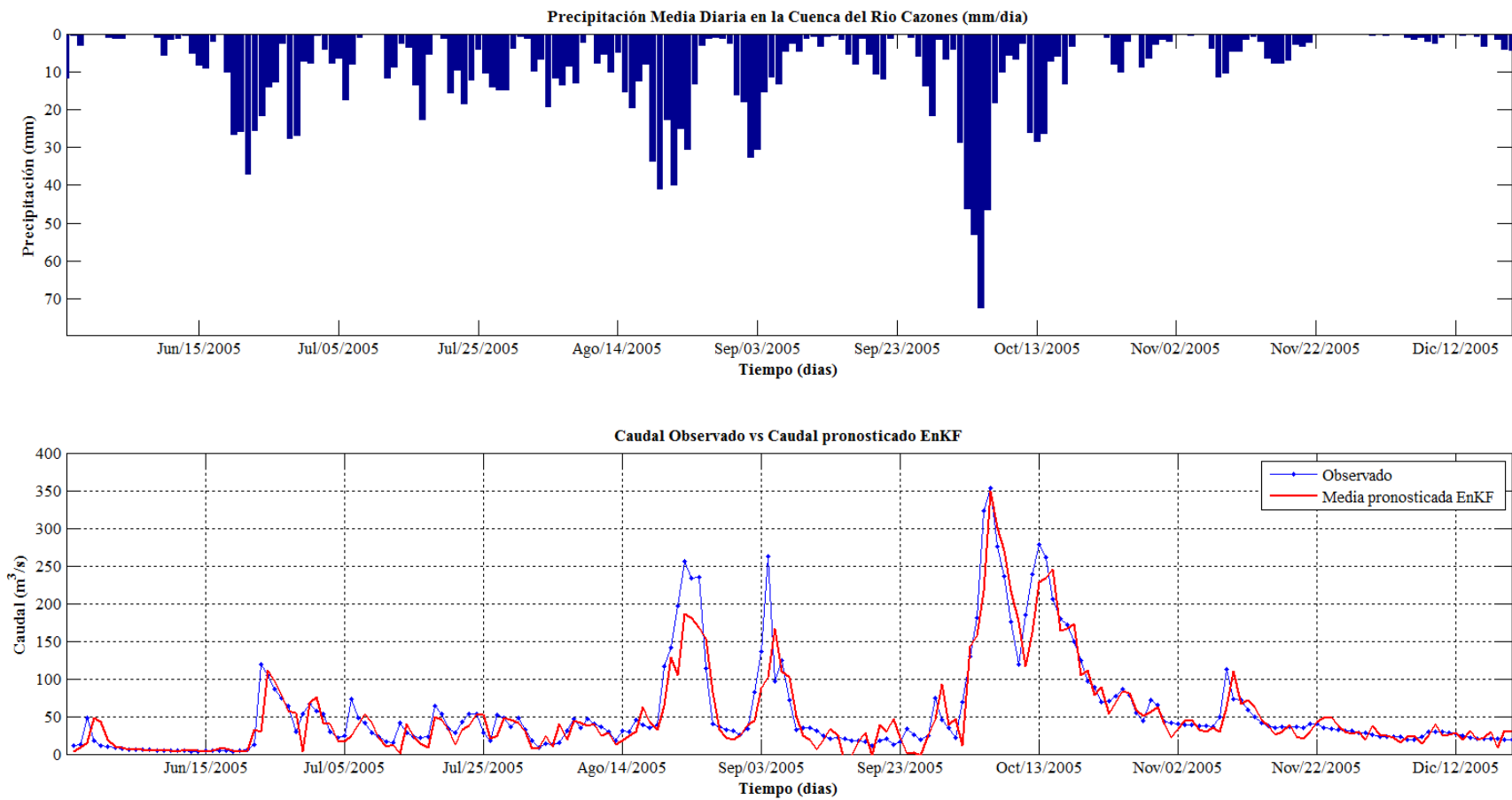


Figura 68. Pronóstico de caudales mediante el ARX+EnKF en la cuenca del río Cazonés para la serie del año 2005.

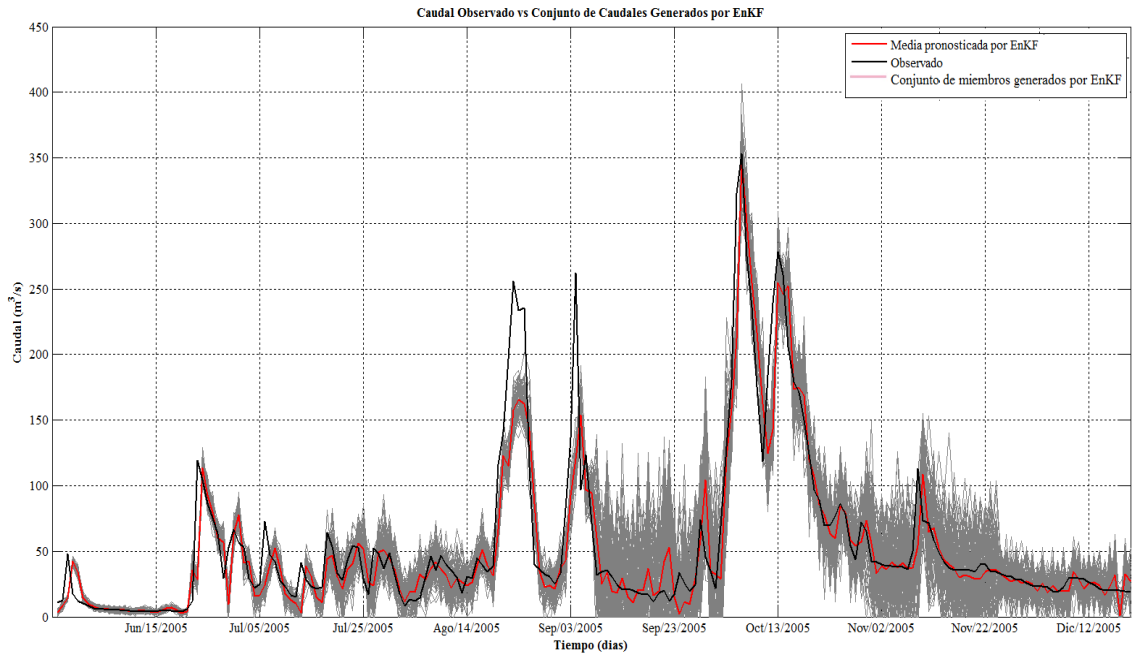


Figura 69. Conjunto de miembros generados en el pronóstico caudales mediante ARX+EnKF en la cuenca del río Cazones para la serie del año 2005.

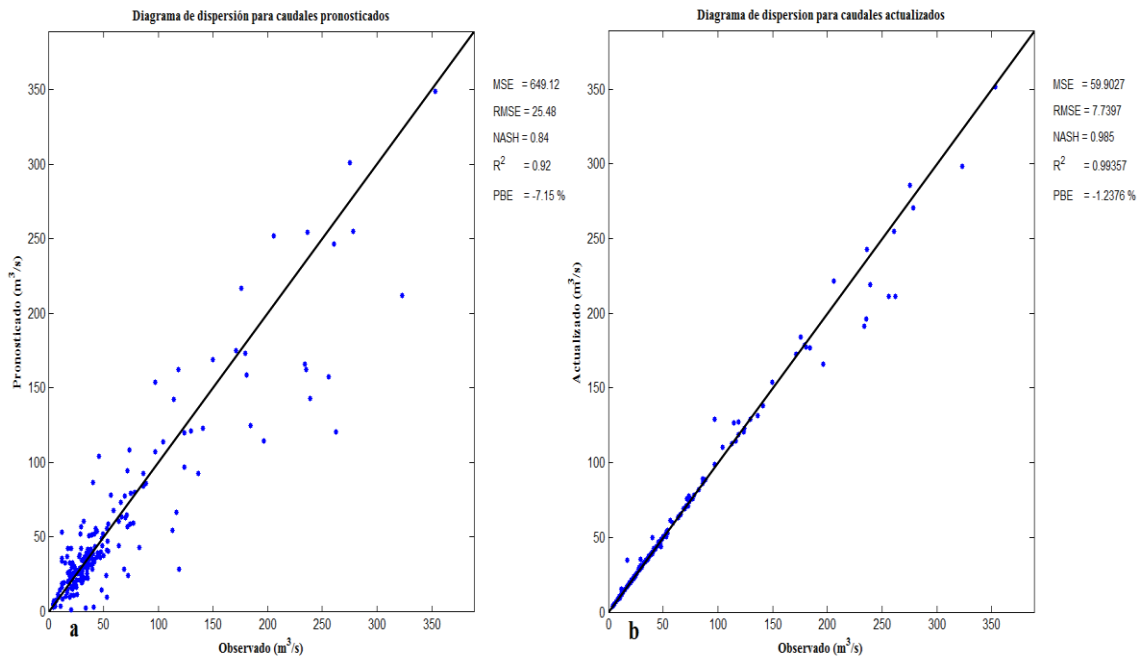


Figura 70. Diagrama de dispersión para la media del conjunto de caudales pronosticados (a) y actualizados (b) mediante el ARX+EnKF en la cuenca del río Cazones para la serie del año 2005.

Los pronósticos realizados utilizando el ARX+EnKF para la serie del año 2005 presentan un ajuste muy similar a los obtenidos mediante el ARX+DKF para la misma serie. Es de aclarar que los caudales actualizados presentan un mejor ajuste ya que las mediciones son asimiladas por el sistema. Se realizaron algunas pruebas para comprobar que el aumento en el número de miembros disminuiría el RMSE de la media de los caudales pronosticados. Los resultados no fueron satisfactorios, ya que los incrementos en el número de miembros no reflejó una disminución considerable en los valores de RMSE para la media del conjunto de estados pronosticados.

Según las condiciones climatológicas presentes en cada una de las cuencas estudiadas, los modelos de pronóstico que fueron implementados se consideran satisfactorios para su implementación en una primera etapa de un sistema de pronóstico en tiempo real, como una herramienta que contribuya a la gestión del riesgo para prevenir inundaciones. Se pudo evidenciar que para cuencas como la del río Cazonas que se encuentran expuestas a fenómenos meteorológicos extremos, los modelos implementados presentaron ajustes inferiores en términos de eficiencia predictiva con respecto a los presentados en la cuenca del río Turbio. La WMO (2011) ha sugerido que en zonas húmedas existen dificultades para la implementación de modelos lluvia-escorrentía por la variabilidad de la precipitación y la rápida respuesta de los ríos.

En términos generales, tanto para la cuenca del río Turbio como para la cuenca del río Cazonas, la metodología del ARX+EnKF presenta resultados muy similares a los obtenidos mediante el ARX+DKF.

Como se mencionó en la revisión de literatura, el EnKF brinda los mejores resultados para modelos con una fuerte relación de no linealidad, debido a que fue propuesto como variante del Extended Kalman Filter (EKF), aunque su implementación con el modelo autorregresivo a pesar de la linealidad del mismo, fue propuesto como una variante académica para su evaluación.

Algunos autores como Pedregal *et al.* (2009) abordaron el problema del pronóstico de caudales y niveles del agua para el río Ebro en España, utilizando modelos de tránsito de avenidas de tipo no lineal en conjunto con filtro de Kalman, que generalmente son tratados como un problema de tipo lineal. Los resultados mostraron la evidente mejoría cuando el modelo no lineal fue utilizado con respecto a la versión lineal.

Por otra parte, es importante mencionar que existen otras técnicas para el pronóstico de caudales a través de modelos lineales o no lineales que pueden ser implementadas en cuencas hidrográficas mexicanas. Recientemente las investigaciones en modelación hidrológica de sistemas no lineales han reportado muy buenos resultados en los que se han implementado el uso de modelos NARX (Chang *et al.* 2014; Ruslan *et al.* 2014), NARMAX (Tabrizi *et al.* 1998; Bartolini *et al.* 2008; Amisigo *et al.* 2008), modelos neuro difusos tipo ANFIS (Khadangi *et al.* 2009; Sanikhani *et al.* 2012; Shiri *et al.* 2012; Rezaeianzadeh *et al.* 2013), redes neuronales artificiales ANN (Zealand *et al.* 1999; Salas *et al.* 2000; Moradkhani *et al.* 2004; Wang *et al.* 2006; Toth *et al.* 2007; Kisi, 2010) y programación genética (Charhat *et al.* 2009; Ni *et al.* 2010).

Finalmente, en los cuadros 23, 24 y 25 se resumen todos los pronósticos realizados en cada cuenca para las diferentes series de caudales. En general, todas las metodologías implementadas con pronósticos de caudales para $L=1$ o 24 h de anticipación, presentan los mejores resultados en cada una de las cuencas evaluadas.

Cuadro 23. Resumen estadístico para el pronóstico de caudales a través del DKF en las dos cuencas evaluadas para diferentes periodos.

Cuenca	Periodo	Parámetros	Series	Estadísticos					
				Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
Río Turbio	2003	dpP = 2	Q Observado	21.2	-	-	-	-	-
		dpQ = 1	Q Pronosticado	21.6	23.37	4.83	0.97	0.98	1.75
		$\alpha = 0.3$	Q Actualizado	21.5	18.09	4.25	0.98	0.99	1.31
	2004	dpP = 2	Q Observado	9.7	-	-	-	-	-
		dpQ = 1	Q Pronosticado	9.5	11.89	3.45	0.97	0.94	-2.59
		$\alpha = 0.3$	Q Actualizado	9.5	10.64	3.26	0.97	0.95	-2.47
Río Cazones	2005	dpP = 1	Q Observado	49.3	-	-	-	-	-
		dpQ = 1	Q Pronosticado	47.2	619.42	24.89	0.92	0.84	-4.18
		$\alpha = 0.25$	Q Actualizado	47.2	561.13	23.69	0.92	0.85	-4.22
	2008-2011	dpP = 1	Q Observado	59.8	-	-	-	-	-
		dpQ = 1	Q Pronosticado	59.9	2889.35	53.75	0.84	0.69	0.15
		$\alpha = 0.35$	Q Actualizado	60.0	2722.68	52.18	0.85	0.70	0.49

*dpP = días de precipitación previa.

*dpQ = días de caudal previo.

* α = Constante de proporcionalidad para la matriz de ruido en la medición R, según Valdés *et al.*, (1980)

Cuadro 24. Resumen estadístico para el pronóstico de caudales a través del ARX+DKF en las dos cuencas evaluadas en diferentes periodos.

Cuenca	Período	Parámetros ARX	Series	Estadísticos					
				Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
Río Turbio	2003		Q Observado	20.1					
		na= 2	Q Pronosticado 24 h	22.6	41.37	6.43	0.97	0.95	-3.82
		nb= 1	Q Pronosticado 48 h	22.0	103.63	10.18	0.93	0.87	-6.14
		P= 20	Q Pronosticado 72 h	21.6	188.29	13.72	0.87	0.76	-7.87
			Q Pronosticado 96 h	21.2	291.01	17.06	0.8	0.63	-9.35
	2004		Q Observado	8.2					
		na= 2	Q Pronosticado 24 h	10.1	16.32	4.04	0.96	0.93	-5.64
		nb= 1	Q Pronosticado 48 h	9.6	39.13	6.26	0.91	0.82	-10.02
		P= 36	Q Pronosticado 72 h	9.2	62.16	7.88	0.86	0.72	-14.14
			Q Pronosticado 96 h	8.8	83.13	9.12	0.8	0.62	-17.84
Río Cazones	2005		Q Observado	56.1					
		na= 1	Q Pronosticado 24 h	51.3	640.21	25.3	0.92	0.84	-8.65
		nb= 2	Q Pronosticado 48 h	48.0	1018.09	31.91	0.88	0.75	-14.45
		P= 26	Q Pronosticado 72 h	46.1	1202.45	34.68	0.86	0.71	-17.98
			Q Pronosticado 96 h	44.8	1286.02	35.86	0.85	0.69	-20.15
	2008-2011		Q Observado	59.0					
		na= 1	Q Pronosticado 24 h	56.9	1984.91	44.55	0.90	0.81	-3.5
		nb= 2	Q Pronosticado 48 h	55.6	2680.36	51.77	0.86	0.74	-5.49
		P= 360	Q Pronosticado 72 h	54.8	2841.68	53.31	0.85	0.72	-6.35
			Q Pronosticado 96 h	54.4	3036.2	55.1	0.84	0.70	-7.05

*na = orden autorregresivo del modelo para caudal.

*nb = número de días previos precipitación.

*P = periodo de calibración del modelo ARX en días.

Cuadro 25. Resumen estadístico para el pronóstico de caudales según ARX+EnKF en diferentes períodos en las dos cuencas evaluadas.

Cuenca	Período	Parámetros	Series	Estadísticos					
				Media	MSE	RMSE	R ²	Nash-Sutcliffe	PBE
Río Turbio	2003	na= 2	Q Observado	21.9	-	-	-	-	-
		nb= 1	Q Pron EnKF	22.3	61.93	7.87	0.97	0.92	-6.68
		P= 20	Q Act EnKF	23.5	22.16	4.70	0.99	0.97	-4.83
	2004	na= 2	Q Observado	10.7	-	-	-	-	-
		nb= 1	Q Pron EnKF	9.9	19.79	4.45	0.96	0.91	-7.6
		P= 36	Q Act EnKF	10.5	2.54	1.59	0.99	0.99	-2.34
Río Cazones	2005	na= 2	Q Observado	56.1	-	-	-	-	-
		nb= 2	Q Pron EnKF	52.1	649.12	25.48	0.92	0.84	-7.15
		P= 26	Q Act EnKF	55.4	59.9	7.73	0.99	0.98	-1.23
	2008-2011	na= 2	Q Observado	59.0	-	-	-	-	-
		nb= 2	Q Pron EnKF	57.8	1949.51	44.15	0.90	0.81	-2.04
		P= 360	Q Act EnKF	59.1	44.82	6.69	0.99	0.99	0.23

*na = orden autorregresivo del modelo para caudal.

*nb = número de días previos precipitación.

*P = periodo de calibración del modelo ARX en días.

6. CONCLUSIONES Y RECOMENDACIONES

De acuerdo a los resultados obtenidos para las dos cuencas evaluadas siguiendo la metodología de ARX+DKF, los pronósticos realizados un paso hacia adelante $L=1$ o 24 h de anticipación, presentan un mejor ajuste en términos de Nash-Sutcliffe y RMSE que los pronósticos realizados 4 pasos hacia adelante, es decir con 48 h, 72 h y 96 h de anticipación, pues a medida que se avanza L pasos en el tiempo, éstos tienden a presentar un menor grado de ajuste con los datos observados.

Las series pronosticadas para el año 2003 y 2004 siguiendo la metodología del DKF en la cuenca del río Turbio, presentaron muy buenos ajustes con respecto a los datos observados. De esta manera, se puede concluir que el modelo de pronósticos a corto plazo implementado en esta cuenca, se ajusta a las condiciones de régimen hidrológico presentando un nivel bajo de incertidumbre.

Los periodos previos de calibración para la metodología del ARX+DKF son fundamentales para mejorar los pronósticos a lo largo del tiempo, pues la función de respuesta en la cuenca puede ser variante según los periodos del ciclo hidrológico o a la presencia de eventos meteorológicos extremos.

La aplicación del ARX+EnKF presentó resultados muy similares con respecto a los obtenidos a través de ARX+KFD para cada una de las cuencas teniendo en cuenta los pronósticos realizados con $L=1$ o 24 h de anticipación.

Los índices de eficiencia de Nash-Sutcliffe para la serie caudales pronosticados a través del ARX+EnKF no difieren en más del 5% con respecto a los mismos estadísticos reportados para la serie pronosticada a través de ARX+DKF correspondiente a los años 2003 y 2004 en la cuenca del río Turbio.

Para el caso de la cuenca del río Cazonas, los índices de Nash-Sutcliffe obtenidos a través de ARX+EnKF en las series del año 2005 y 2008-2011 son muy similares a los obtenidos mediante ARX+KFD con $L=1$ o 24 h de anticipación, implementado en la misma cuenca. Se esperaban los mejores resultados para los pronósticos realizados a través ARX+EnKF, aunque éstos no se presentaran debido a la linealidad que describen los procesos autorregresivos, recordando que el EnKF fue desarrollado para modelos con fuertes relaciones de no linealidad, aunque no se descarta su finalidad como modelo de pronóstico como una propuesta del ejercicio académico.

Existe una relación entre el orden autorregresivo de los coeficientes que mejor describe la predicción de caudales según el modelo ARX y las descargas registradas en la estación de aforo en la cuenca del río Turbio, pues en este caso, el modelo considera un mayor orden autorregresivo para caudales que para precipitación (ver cuadro 20). Este comportamiento puede guardar una estrecha relación con el sistema real, debido a que existe gran cantidad de reservorios que se comportan como amortiguadores de las grandes avenidas, que desde el punto de vista del tránsito hidrológico regulan los escurrimientos, modificando los hidrogramas a la salida de la cuenca.

Los modelos autorregresivos son herramientas de pronóstico que en la hidrología son fundamentales para la representación de un sistema y que pueden considerarse de tipo exploratorio para futuros análisis con niveles de detalle más profundos.

Los pronósticos de caudales implementados en las 3 metodologías, partieron del supuesto de que el sistema se representa como un modelo de caja negra que relaciona a la precipitación media de la cuenca en diferentes instantes del tiempo como entrada principal y el caudal medio registrado en la estación de aforo como la respuesta obtenida. Esta función de respuesta presenta un componente estocástico lineal, invariante en el tiempo para el caso de la metodología del DKF y variante para las metodologías del ARX+DKF y ARX+EnKF respectivamente.

La implementación de técnicas robustas para el pronóstico de caudales en cuencas con instrumentación muy pobre o deficiente, puede ser una tarea compleja si la información adquirida es de mala calidad o no existen registros de larga duración.

Se propone avanzar en la implementación de técnicas de asimilación de datos a través de las versiones del filtro de Kalman utilizando modelos hidrológicos de base física agrupados o distribuidos, como también el uso de modelos no lineales tipo NARX, redes neuro difusas tipo ANFIS, redes neuronales ANN y programación genética GP, con el objetivo de evaluar su operatividad en cuencas mexicanas.

Se recomienda considerar la humedad antecedente como variable exógena en el modelo de pronóstico que se desee implementar, ya que puede proporcionar mejores resultados.

7. BIBLIOGRAFÍA

Ahsan, M. and O'Connor, K. M. 1994. A reappraisal of the Kalman filtering technique, as applied in river flow forecasting. *Journal of Hydrology* 161(1): 197-226.

Amisigo, B. A., Van de Giesen, N., Rogers, C., Andah, W. E. and Friesen, J. 2008. Monthly streamflow prediction in the Volta Basin of West Africa: A SISO NARMAX polynomial modelling. *Physics and Chemistry of the Earth, Parts A/B/C* 33(1-2):141-150.

Aparicio, F. J. 1989. *Fundamentos de hidrología de superficie*. México, Editorial Limusa S.A. De C.V. 303 pp.

Aubert, D., Loumagne, C. and Oudin, L. 2003. Sequential assimilation of soil moisture and streamflow data in a conceptual rainfall-runoff model. *Journal of Hydrology* 280(1): 145-161.

Awwad, H. M. and Valdés, J. B. 1992. Adaptive parameter estimation for multisite hydrologic forecasting. *Journal of Hydraulic Engineering* 118(9): 1201-1221.

Awwad, H. M., Valdés, J. B. and Restrepo, P. J. 1994. Streamflow forecasting for Han river basin, Korea. *Journal of Water Resources Planning and Management* 120(5): 651-673.

Bartolini, P., Carcano, E.C., Piroddi, L., and Valdés, J.B. 2008. Forecasting Daily Streamflows Using NARMAX Models: How Disturbances May Affect Model Performance World Environmental and Water Resources Congress. pp. 1-13. doi: 10.1061/40976(316)613.

Bergman, M. and Delleur, J. 1985. Kalman filter estimation and prediction of daily Streamflow II. Application to the Potomac River 1, Wiley Online Library.

Beven, K.J. and Binley, A.M., 1992. The future of distributed models: model calibration and uncertainty prediction. *Hydrological Processes* 6, 279-298 pp.

Box, G. E. P., Jenkins, G. M. and Reinsel, G. C. 2013. *Time series analysis: forecasting and control*. Four edition. John Wiley & Sons. 746 pp.

Bras, R. L. and Rodríguez-Iturbe, I. 1985. *Random Functions and Hydrology*, Addison-Wesley. 559 pp.

Campos D.F. 2007. Estimación y aprovechamiento del escurrimiento. Editorial Campos, San Luís de Potosí, México. 440p.

Chang, F.-J., Chen, P.A., Lu, Y.R., Huang, E. and Chang, K.Y. 2014. Real-time multi-step-ahead water level forecasting by recurrent neural networks for urban flood control. *Journal of Hydrology*, 517(0): 836-846.

Charhat, S. B., Dandawat, Y. H. and Londh, S. N. 2009. Genetic Programming to Forecast Stream Flow. *Advances in Water Resources and Hydraulic Engineering*, Springer Berlin Heidelberg: 29-34.

Chatfield, C. 2001. Prediction intervals for time-series forecasting. In *Principles of forecasting*. Springer US. pp. 475-494.

Chatfield, C. 2004. *The analysis of time series: An introduction*. Six edition. Chapman & Hall/ CRC press. 313 pp.

Chow, V. T., Maidment, D. and Mays, L. 1994. *Hidrología Aplicada*. Bogotá D.C, McGraw-Hill Interamericana, S.A. Colombia. 584 pp.

Cloke, H. L., and Pappenberger, F. 2009. Ensemble flood forecasting: a review. *Journal of Hydrology*, 375(3), 613-626.

Comisión Nacional del Agua, CONAGUA. 2003. Programa Hidráulico Regional 2002-2006 Región VIII Lerma Santiago Pacífico.

Comisión Nacional del Agua, CONAGUA. 2012. Atlas del Agua en México. Disponible en: <http://www.conagua.gob.mx/atlas/>.

Comisión Nacional del Agua, CONAGUA. 2014a. Sistema de Seguridad de Presas. Disponible en: <http://www.conagua.gob.mx/atlas/>.

Comisión Nacional del Agua, CONAGUA. 2014b. Banco Nacional de Aguas Superficiales BANDAS. Disponible en: ftp://ftp.conagua.gob.mx/Bandas/Bases_Datos_Bandas. Recuperado 10 de Febrero 2014.

Esri. 2013. ArcGis® Desktop. Release 10.1.

Evensen, G. 1994. Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics,” *Journal of Geophysical Research*. vol. 99, no. C5, pp. 10143–10162.

Evensen, G. 2009. The ensemble Kalman filter for combined state and parameter estimation. *Control Systems*, IEEE 29(3): 83-104.

Faragher, R. 2012. Understanding the basis of the Kalman filter via a simple and intuitive derivation. *IEEE Signal Processing Magazine* 29(5): 128-132.

García, E. 1981. Modificaciones al sistema de clasificación climática de Köppen para adaptarlo a las condiciones de la República Mexicana. Instituto de Geografía, Universidad Nacional Autónoma de México. México D.F. 220 pp.

Gelb, A. 1974. Applied Optimal Estimation. Massachusetts, MIT Press. 374 pp.

Gillijns, S., Mendoza, O. B., Chandrasekar, J., De Moor, B., Bernstein, D. and Ridley, A. 2006. What is the ensemble Kalman filter and how well does it work? American Control Conference, IEEE.

Gouweleeuw, B. T., Thielen, J., Franchello, G., De Roo, A. P. J. and Buizza, R. 2005. Flood forecasting using medium-range probabilistic weather prediction. Hydrology and Earth System Sciences, 9(4): 365-380.

Gupta, H. V., Kling, H., Yilmaz, K. and Martinez, G.F. 2009. Decomposition of the mean squared error and NSE performance criteria: Implications for improving hydrological modelling: Journal of Hydrology, v. 377, p. 80-91.

Haiden, T., Magnusson, L., Tsonevsky, I., Wetterhall, F., Alfieri, L., Pappenberger, F., de Rosnay, P., Muñoz-Sabater, J., Balsamo, G. and Albergel, C. 2014. ECMWF forecast performance during the June 2013 flood in Central Europe. European Center for Medium-Range Weather Forecast. Technical Memorandum 723.

Hino, M. 1973. On-line prediction of hydrologic system. Proceedings of the 15th Congress of the International Association for Hydraulic Research, Istanbul.

Hirpa, F. A., Hopson, T. M., De Groeve, T., Brakenridge, G. R., Gebremichael, M., & Restrepo, P. J. 2013. Upstream satellite remote sensing for river discharge forecasting: Application to major rivers in South Asia. Remote Sensing of Environment. 131, 140-151.

Husain, T. 1985. Kalman filter estimation model in flood forecasting. Advances in water resources 8(1): 15-21.

Instituto Nacional de Estadística y Geografía, INEGI. 2005. Síntesis de información geográfica del Estado de Veracruz.

Jajarmizadeh, M., Harun, S. and Salarpour, M. 2012. A review on theoretical consideration and types of models in hydrology. Journal of Environmental Science and Technology 5(5): 249-261.

Kalman, R. E. 1960. A new approach to linear filtering and prediction problems. Journal of Fluids Engineering 82(1): 35-45.

Khadangi, E., Madvar, H. R. and Ebadzadeh, M. M. 2009. Comparison of ANFIS and RBF models in daily stream flow forecasting. Computer, Control and Communication. IC4 2009. 2nd International Conference on.

Kim, P. 2011. Kalman filter for beginners: with MATLAB examples. CreatesSpace. 232 pp.

Kim, S., Tachikawa, Y. and Takara, K. 2004. Embedding Kalman filter into a distributed Hydrological Model. IHP Conference with WSUD .pp. 21-25.

Kisi, O. 2010. Wavelet regression model for short-term streamflow forecasting. Journal of Hydrology 389(3-4): 344-353.

Lee, Y. and Singh, V. 1998. Application of the Kalman filter to the Nash model. Hydrological Processes 12(5): 755-767.

López, M. A. 2008. Estudio de mitigación de riesgos de inundación del río Turbio. Aqua Forum Año 12(48).

Lü, H., Hou, T., Horton, R., Zhu, Y., Chen, X., Jia, Y., Wang, W. and Fu, X. 2013. The streamflow estimation using the Xinanjiang rainfall runoff model and dual state-parameter estimation method. Journal of Hydrology 480: 102-114.

Madsen, H., Rosbjerg, D., Damgaard, J. and Hansen, F. S. 2003. Data assimilation in the MIKE 11 flood forecasting system using Kalman filtering. International Association of Hydrological Sciences, Publication (281): 75-81.

Matlab® R2013b. MathWorks, Inc.

Méndez, Z. C. 2007. Implementación del filtro de Kalman en el pronóstico de caudales caso de estudio: estación puente Balseadero río Magdalena. Tesis para obtener el título de Máster en Hidrosistemas. Pontificia Universidad Javeriana. Bogota D.C, Colombia.

Mood, A.; Graybill, F.; Boes, D. 1974. Introduction to the Theory of Statistics. Third Edition. McGraw-Hill. 229 pp.

Moradkhani, H., DeChant, C. M. and Sorooshian, S. 2012. Evolution of ensemble data assimilation for uncertainty quantification using the particle filter-Markov chain Monte Carlo method. Water Resources Research 48(12).

Moradkhani, H., Hsu, K. L., Gupta, H. V., & Sorooshian, S. 2004. Improved streamflow forecasting using self-organizing radial basis function artificial neural networks. Journal of Hydrology 295(1): 246-262.

Moradkhani, H., Sorooshian, S., Gupta, H. V. and Houser, P. R. 2005. Dual state-parameter estimation of hydrological models using ensemble Kalman filter. Advances in Water Resources 28(2): 135-147.

Morales, M. I., Aparicio Mijares, F. J. and Valdés, J. B. 2014. Pronóstico de avenidas utilizando el filtro de Kalman discreto. Tecnología y Ciencias del Agua 5(2): 85-110.

- Nash, J. E., and J. V. Sutcliffe. 1970. River flow forecasting through conceptual models part I. A discussion of principles: *Journal of Hydrology*, v. 10, p. 282-290.
- National Aeronautics and Space Administration NASA. 2014. Archives Hurricane Stan. Recuperado 15 Agosto 2014. Disponible en: http://www.nasa.gov/mission_pages/hurricanes/archives/2005/h2005_stan.html.
- Neal, J. C., Atkinson, P. M. and Hutton, C. W. 2007. Flood inundation model updating using an ensemble Kalman filter and spatially distributed measurements. *Journal of Hydrology* 336(3): 401-415.
- Ni, Q., Wang, L., Ye, R., Yang, F. Sivakumar, M. 2010. Evolutionary modeling for streamflow forecasting with minimal datasets: a case study in the West Malian River, China. *Environmental Engineering Science* 27 (5): 377-385.
- Obregón, N. and Rosas, J. 2003. Hacia la predicción de precipitación puntual en tiempo real mediante el filtro Kalman. *International Water Association*. Cartagena de Indias, Colombia. 1-4 pp.
- Ogata, K. 2003. *Ingeniería de Control Moderna*. Cuarta Edición. España, Pearson Educación. 965 pp.
- Pappenberger, F., Bartholmes, J., Tsonevsky, I., Thielen, J. and Anghel, E. 2008. TIGGE: medium range multi model weather forecast ensembles in flood forecasting (a case study) *European Center for Medium-Range Weather Forecast*. Technical Memorandum 557.
- Pedregal, D. J., Rivas, R., Feliu, V., Sánchez, L., and Linares, A. 2009. A non-linear forecasting system for the Ebro River at Zaragoza, Spain. *Environmental Modelling & Software*, 24(4): 502-509.
- Pereyra, D. and Pérez, J. A. 2006. Hidrología de superficie y precipitaciones intensas 2005 en el estado de Veracruz, Inundaciones. Disponible en: <http://cdigital.uv.mx/bitstream/123456789/9650/1/04HIDROLOGIA.pdf>. Recuperado 11 Agosto 2014.
- Pereyra, D. and Pérez, J. A. 2010. Atlas del patrimonio natural, histórico y cultural de Veracruz. *Hidrología*. Comisión del Estado de Veracruz: 85-122.
- Protección Civil Guanajuato. 2012. Online. Disponible en: http://proteccioncivil.guanajuato.gob.mx/atlas/hidrometeorologico/hidro_cuenca_turbio.php. Recuperado Junio 24 de 2014.
- Rajagopalan, B., Salas, J. D. and Lall, U. 2010. Stochastic methods for modeling precipitation and streamflow. *Advances in data based approaches for hydrologic modeling and forecasting* 17-52 pp.

- Ramírez, Á. S. 2003. El filtro de Kalman. Nota técnica. Banco Central de Costa Rica, División Económica, Departamento de Investigaciones Económicas. 33 pp.
- Reichle, R. H., McLaughlin, D. B. and Entekhabi, D. 2002. Hydrologic data assimilation with the ensemble Kalman filter. *Monthly Weather Review* 130(1): 103-114.
- Rezaeianzadeh, M., Tabari, H., Yazdi, A. A., Isik, S., and Kalin, L. 2013. Flood flow forecasting using ANN, ANFIS and regression models. *Neural Computing and Applications*, 1-13 pp.
- Ruslan, F.A.; Samad, A.M.; Zain, Z.M.; Adnan, R. Flood water level modeling and prediction using NARX neural network: Case study at Kelang river. 2014. *Signal Processing & its Applications (CSPA), IEEE 10th International Colloquium*. On page(s): 204-207.
- Samuel, J., Coulibaly, P., Dumedah, G. and Moradkhani, H. 2014. Assessing model state and forecasts variation in hydrologic data assimilation. *Journal of Hydrology* 513: 127-141.
- Sanikhani, H. and Kisi, O. 2012. River Flow Estimation and Forecasting by Using Two Different Adaptive Neuro-Fuzzy Approaches. *Water Resources Management* 26(6): 1715-1729.
- Santillán Gutiérrez, E., Dávila-Vázquez, G., Sánchez, J. and Díaz Torres, J. 2013. Estimación del balance hídrico mediante variables climáticas, en la cuenca del río Cazones, Veracruz, México. *Revista Ambiente & Água* 8: 104-117.
- Salas, J. D., Delleur, J. W., and Yevjevich, V. 1980. *Applied modeling of hydrologic time series*. Water Resources Publications. Colorado. 484 pp.
- Salas, J. D., Markus, M. and Tokar, A. S. 2000. Streamflow Forecasting Based on Artificial Neural Networks. In *Artificial Neural Networks in Hydrology*. Springer Netherlands 36: 23-51.
- Servicio Meteorológico Nacional, SMN. 2014. Sistema de información climática computarizada CLICOM.
- Schreider, S. Y., Young, P. and Jakeman, A. 2001. An application of the Kalman filtering technique for streamflow forecasting in the Upper Murray Basin. *Mathematical and Computer Modelling* 33(6): 733-743.
- Shi, Y., Davis, K. J., Zhang, F., Duffy, C. J. and Yu, X. 2014. Parameter estimation of a physically based land surface hydrologic model using the ensemble Kalman filter: A synthetic experiment. *Water Resources Research* 50(1): 706-724.
- Simon, D. 2001. Kalman filtering. *Embedded Systems Programming* 14(6): 72-79.

Shiri, J., Kişi, Ö., Makarynsky, O., Shiri, A.A. and Nikoofar, B. 2012. Forecasting daily stream flows using artificial intelligence approaches. *ISH Journal of Hydraulic Engineering* 18(3): 204-214.

Singh, V. P. 1988. *Hydrologic Systems: Watershed Modeling V1. Rainfall-runoff Modeling*, Prentice Hall PTR. 320 pp.

Singh, V. P. and Woolhiser, D. A. 2002. Mathematical modeling of watershed hydrology. *Journal of Hydrologic Engineering* 7(4): 270-292.

Tabrizi, M. H. N., Said, S. E., Badr, A. W., Mashor, Y. and Billings, S. A. 1998. nonlinear modeling and prediction of a river flow system1. *Journal of the American Water Resources Association* 34(6): 1333-1339.

Toth, E. and Brath, A. 2007. Multistep ahead streamflow forecasting: Role of calibration data in conceptual and neural network modeling, *Water Resources Research.*, 43(11), doi: 10.1029/2006WR005383.

U.S Army Corps of Engineers. Hec-GeoHMS. Release 10.1 for ArcGis® 10.1.

Valdés, J. B., Velázquez, M. J. and Rodríguez-Iturbe, I. 1980. Filtros de Kalman en la Hidrología: Predicción de descargas fluviales para la operación optima de embalses. Informe Técnico No 80-2. Universidad Simón Bolívar, Decanato de estudios de Posgrado, Posgrado en Planificación e Ingeniería de Recursos Hídricos. Venezuela. 70 pp.

Wang, W. 2006. *Stochasticity, nonlinearity and forecasting of streamflow processes*. Ios Press. 206 pp.

Wang, W., Van Gelder, P. H., Vrijling, J. K., and Ma, J. 2006. Forecasting daily streamflow using hybrid ANN models. *Journal of Hydrology* 324(1): 383-399.

Welch, G. and Bishop, G. 2006. *An introduction to the Kalman filter*. University of North Carolina, Chapel Hill.

World Meteorological Organization WMO 2009. *Guide to Hydrological Practices Volume II. Management of Water Resources and Application of Hydrological Practices*. WMO N° 168. Geneva.

World Meteorological Organization WMO 2011. *Manual on Flood Forecasting and Warning*. WMO-No. 1072.

Xie, X. and Zhang, D. 2010. Data assimilation for distributed hydrological catchment modeling via ensemble Kalman filter. *Advances in Water Resources* 33(6): 678-690.

Zealand, C. M., Burn, D. H., and Simonovic, S. P. 1999. Short term streamflow forecasting using artificial neural networks. *Journal of hydrology* 214(1): 32-48.