



UNIVERSITAT
POLITÈCNICA
DE VALÈNCIA

**PROGRAMA DE DOCTORADO EN INGENIERÍA DEL AGUA Y
MEDIOAMBIENTAL**

**PRONÓSTICO PROBABILÍSTICO DE EVENTOS DE SEQUÍAS Y
EVALUACIÓN DEL RIESGO EN LA GESTIÓN DE SISTEMAS DE
RECURSOS HÍDRICOS. CASO DE ESTUDIO EN UNA CUENCA
ANDINA REGULADA**

CANDIDATO:

ALEX AVILÉS AÑAZCO

DIRECTORES:

D. ABEL SOLERA SOLERA

D. JAVIER PAREDES ARQUIOLA

ENERO 2017

*A Dios, Jheimy,
Joaquín y Juan
Francisco, para ellos
mi amor y esfuerzo.*

*Lo que sabemos es una gota de
agua; lo que ignoramos es el
océano.*

Isaac Newton

AGRADECIMIENTOS:

El autor agradece a los Doctores Abel Solera Solera y Javier Paredes Arquiola por su guía para la realización de esta tesis, además un agradecimiento a la Facultad de Ciencias Químicas y el Departamento de Recursos Hídricos y Ciencias Ambientales de la Universidad de Cuenca por el apoyo en el transcurso de la fase de investigación de mi doctorado y por último agradecer al personal administrativo del Departamento de Ingeniería Hidráulica y Medio Ambiente de la Universidad Politécnica de Valencia por la ayuda en los trámites de mi formación.

RESUMEN

Las condiciones hidroclimáticas de escasez de agua pueden afectar a los servicios hidrológicos que ofrecen las cuencas de montaña causando severos impactos a la población que se beneficia de estos recursos. Estas condiciones se han convertido en un verdadero desafío para quienes toman las decisiones en cuencas andinas con respecto a la gestión de sistemas de recursos hídricos en épocas de sequías. Esta tesis evalúa el riesgo de fallo en el suministro de agua mediante la introducción de pronósticos probabilísticos de eventos de sequía que sirvan de soporte para la toma de decisiones con respecto a la satisfacción de las demandas consuntivas, no consuntivas y ambientales en tiempos de escasez. Mediante un índice de sequía (DI) creado a partir de la priorización de información disponible de variables relacionadas con el agua se comparó el rendimiento de dos métodos de pronóstico probabilístico de sequías basados en la probabilidad condicional. Un primer método comúnmente utilizado basado en cadenas de Markov (MC) presentó el mejor desempeño que un segundo método poco utilizado basado en redes bayesianas (BN). Los pronósticos obtenidos del mejor modelo fueron incorporados en la metodología de evaluación del riesgo de fallo de sistemas de recursos hídricos, la cual está basada en la simulación Montecarlo de múltiples escenarios estocásticos formados a partir de la generación sintética hidrológica y la modificación de algunas características de los sistemas de recursos hídricos. La cuenca del río Machángara ubicada al sur del Ecuador fue escogida para probar las herramientas y metodologías expuestas en esta tesis. Esta cuenca presenta cada vez periodos de sequías que conducen a la disminución de la disponibilidad del agua para los múltiples usos produciendo efectos negativos en la población y el ambiente. Mediante la utilización de información hidrometeorológica y datos de consumo de agua de las demandas se simuló 1728 escenarios estocásticos, los cuales incorporaron pronósticos probabilísticos de sequías para evaluar el riesgo de fallo en sistemas de recursos hídricos. Las salidas del proceso de simulación se consolidaron en un índice integral de satisfacción del sistema (DSI_G). El cual sirvió para obtener un banco de información que podría ayudar como soporte para la toma de decisiones de los gestores del agua para que puedan proponer múltiples alternativas de prevención y/o mitigación cuando la satisfacción sea baja. Esta tesis propone un enfoque integral de evaluación del riesgo de fallo de sistemas de recursos hídricos. Se analizan múltiples situaciones de escasez de disponibilidad de agua que podría servir para disminuir las incertidumbres en la toma de decisiones y perfeccionar las medidas orientadas para

prevenir o mitigar las consecuencias adversas de las sequías en sistemas de abastecimiento de agua.

RESUM

Les condicions hidroclimàtiques d'escassetat d'aigua poden afectar als serveis hidrològics que ofereixen les conques de muntanya causant severes impactes a la població que es beneficia d'aquests recursos. Aquestes condicions s'han convertit en un vertader desafiament per als qui prenen les decisions en conques d'andines pel que fa a la gestió de sistemes de recursos hídrics en èpoques de sequeres. Aquesta tesi avalua el risc de fallada en el subministrament d'aigua mitjançant la introducció de pronòstics probabilístics d'esdeveniments de sequera que servisquen de suport per a la presa de decisions pel que fa a la satisfacció de les demandes consumptives, no consumptives i ambientals en temps d'escassetat. Mitjançant un índex de sequera (DI) creat a partir de la prioritització d'informació disponible de variables relacionades amb l'aigua es va comparar el rendiment de dos mètodes de pronòstic probabilístic de sequeres basats en la probabilitat condicional. Un primer mètode comunament utilitzat basat en cadenes de Markov (MC) va presentar el millor acompliment que un segon mètode poc utilitzat basat en xarxes bayesianas (BN). Els pronòstics obtinguts del millor model de pronòstic van ser incorporats en la metodologia d'avaluació del risc de fallada de sistemes de recursos hídrics, la qual aquesta basada en la simulació Montecarlo de múltiples escenaris estocàstics formats a partir de la generació sintètica hidrològica i la modificació d'algunes característiques dels sistemes de recursos hídrics. La conca del riu Machángara situada al sud de l'Equador va ser escollida per a provar les eines i metodologies exposades en aquesta tesi. Aquesta conca presenta cada vegada períodes de sequeres que condueixen a la disminució de la disponibilitat de l'aigua per als múltiples usos produint efectes negatius en la població i l'ambient. Mitjançant la utilització d'informació hidrometeorològica i dades de consum d'aigua de les demandes es van simular 1728 escenaris estocàstics, els quals van incorporar pronòstics probabilístics de sequeres per a avaluar el risc de fallada en sistemes de recursos hídrics. Les eixides del procés de simulació es van consolidar en un índex integral de satisfacció del sistema (DSI_G), el qual va servir per a obtenir un banc d'informació de suport per a la presa de decisions dels gestors de l'aigua perquè puguin proposar múltiples alternatives de prevenció i/o mitigació quan la satisfacció siga baixa. Aquesta tesi proposa un enfocament integral d'avaluació del risc de fallada de sistemes de recursos hídrics analitzant múltiples situacions d'escassetat de disponibilitat d'aigua que podria servir d'ajuda per a disminuir les incerteses en la presa de decisions i perfeccionar les mesures orientades per a

previndre o mitigar les conseqüències adverses de les sequeres en sistemes de proveïment d'aigua.

ABSTRACT

Hydroclimatic water shortage conditions can affect hydrological services offered by mountain basins causing severe impacts to the population benefiting from these resources. These conditions have become a real challenge for decision-makers in Andean basins with respect to the management of water resources systems in drought times. This thesis evaluates the risk of failure in the water supply by introducing probabilistic forecasts of drought events that serve as support for decision-making regarding the satisfaction of the consumptive demands, non-consumptive and environmental in scarcity times. Through a drought index (DI) created from the prioritization of available information about water-related variables, performance of two methods of probabilistic drought forecasting based on the conditional probability were compared. A first commonly used method based on Markov chains (MC) presented the best performance than a second method based on Bayesian networks (BN). The forecasts obtained the model with the best performance were incorporated into the methodology of assessing the risk of failure of water resources systems, which is based on Monte Carlo simulation of multiple stochastic scenarios designed from hydrological synthetic generation and modification of some features of water resources systems. Machángara river basin located south of Ecuador was chosen to apply methods and tools proposed in this thesis. This basin has increasingly periods of drought leading to reduce water availability for multiple demands producing negative effects on the population and the environment. 1728 stochastic scenarios were simulated using hydro-meteorological information and data of water consumption demands, which incorporate probabilistic forecasts of drought events to assess the risk of failure in water resources systems. The outcomes of the simulation process were consolidated into a integrated satisfaction system index (DSI_G), which aided to obtain information to support decision-making by water managers so they can to propose several alternatives for prevention and / or mitigation when satisfaction is low. This thesis proposes a integrated approach to assessing the risk of failure of water resources systems analyzing multiple states of water availability shortages that could help to reduce uncertainty in decision-making and improve measures aimed to prevent or mitigate adverse effects of drought on water supply systems.

CONTENIDO

RESUMEN	V
RESUM.....	VI
ABSTRACT.....	VII
CONTENIDO	VIII
1. INTRODUCCIÓN	1
1.1. OBJETIVOS	5
1.2. PUBLICACIONES.....	6
1.2.1. EVALUACIÓN DE PRONÓSTICOS DE SEQUÍAS BASADOS EN CADENAS MARKOV EN UNA CUENCA ANDINA REGULADA USANDO LOS PUNTAJES DE HABILIDAD RPS Y GMSS..	6
1.2.2. PRONÓSTICO PROBABILÍSTICO DE EVENTOS DE SEQUÍAS USANDO MODELOS BASADOS EN CADENAS DE MARKOV Y REDES BAYESIANAS: UN CASO DE ESTUDIO DE UNA CUENCA ANDINA REGULADA	8
1.2.3. ESTIMACIÓN DEL RIESGO DE FALLO EN EL SUMINISTRO DE AGUA COMO AYUDA A LA PLANIFICACIÓN Y GESTIÓN DE RECURSOS HÍDRICOS	9
1.2.4. ANÁLISIS DE SISTEMAS DE RECURSOS HÍDRICOS DE LA CUENCA DEL RIO TOMBAMBA EN ECUADOR, MEDIANTE MODELOS ESTOCÁSTICOS Y DE GESTIÓN	10
1.3. ESTRUCTURA DE LA TESIS	11
2. CASO DE ESTUDIO	12
3. PROPUESTA DE MEJORA DE LA METODOLOGÍA GENERAL DE EVALUACIÓN DEL RIESGO DE FALLO DE SISTEMAS DE RECURSOS HÍDRICOS	15
3.1. ÍNDICE DE SEQUÍA	18
3.2. PRONÓSTICO PROBABILÍSTICO DE SEQUÍA	19
3.2.1. CADENAS DE MARKOV	19
3.2.2. REDES BAYESIANAS	21
3.2.2.1 Copulas	22
3.2.2.2. Ajuste de copulas	24
3.2.3. VERIFICACIÓN DEL PRONÓSTICO.....	25
3.3. MODELO ESTOCÁSTICO PARA LA GENERACIÓN DE SERIES SINTÉTICAS HIDROLÓGICAS	26
3.4. INTRODUCCIÓN DE LOS PRONÓSTICOS DE SEQUÍA EN LA GENERACIÓN DE SERIES SINTÉTICAS HIDROLÓGICAS	27
3.5. MODELO DE SIMULACIÓN MÚLTIPLE PARA LA EVALUACIÓN DEL RIESGO DE FALLO	28
4. RESULTADOS	30
4.1. CÁLCULO DEL DI	30
4.2. PRONÓSTICOS PROBABILÍSTICOS DE SEQUÍA.....	31
4.3. EVALUACIÓN DEL RIESGO DE FALLO.....	35
5. CONCLUSIONES	43
5.1. FUTURAS INVESTIGACIONES.....	44
REFERENCIAS.....	45

ANEXO 1: EVALUATION OF MARKOV CHAIN BASED DROUGHT FORECASTS IN AN ANDEAN REGULATED RIVER BASIN USING THE SKILL SCORES RPS AND GMSS.....	55
ANEXO 2: PROBABILISTIC FORECASTING OF DROUGHT EVENTS USING MARKOV CHAIN AND BAYESIAN NETWORK BASED MODELS. CASE STUDY: AN ANDEAN REGULATED RIVER BASIN.....	76
ANEXO 3: ESTIMACIÓN DEL RIESGO DE FALLO EN EL SUMINISTRO DE AGUA COMO AYUDA A LA PLANIFICACIÓN Y GESTIÓN DE RECURSOS HÍDRICOS	103
ANEXO 4: ANÁLISIS DEL SISTEMA DE RECURSOS HÍDRICOS DE LA CUENCA DEL RÍO TOMEBAMBA EN ECUADOR, MEDIANTE MODELOS ESTOCÁSTICOS Y DE GESTIÓN	118

FIGURAS

Figura 1. Localización de la cuenca del río Machángara.....	13
Figura 2. Promedio mensual de las series de tiempo de lluvia y volumen de agua de entrada a los embalses (1971 – 2010) de las microcuencas de los ríos (a) Machángara Alto y (b) Chulco	14
Figura 3. Esquema del sistema de recursos hídricos de la cuenca del río Machángara.....	15
Figura 4. Metodología de evaluación del riesgo de fallo en sistemas de recursos hídricos ..	16
Figura 5. Marco metodológico integral de evaluación del riesgo de fallo en sistemas de recursos hídricos	17
Figura 6. Series de tiempo del DI (1971 – 2010) de las microcuencas de los ríos (a) Machángara Alto y (b) Chulco	31
Figura 7. Probabilidad de fallo de las demandas de agua para: (a) ciudad de Cuenca, (b) canal de riego Machángara y (c) canal de riego La Dolorosa.....	39
Figura 8. DSI_G del sistema de recursos hídricos de la cuenca del río Machángara para un escenario	40
Figura 9. DSI_G del sistema de recursos hídricos de la cuenca del río Machángara para un escenario con diferentes volúmenes iniciales de almacenamiento de los embalses de Chanlud y El Labrado aplicando el marco metodológico integral de la Figura 5	41
Figura 10. $DSIG$ del sistema de recursos hídricos de la cuenca del río Machángara para un escenario con diferentes volúmenes iniciales de almacenamiento de los embalses de Chanlud y El Labrado aplicando la metodología de la Figura 4	42

TABLAS

Tabla 1. Autovalores y varianza explicada del PC1 de las microcuencas de los ríos (a) Machángara Alto y (b) Chulco	30
Tabla 2. Valores del RPSS de la verificación mensual de los pronósticos de sequía.....	32
Tabla 3. Pronósticos probabilísticos de sequía de la microcuenca del río Machángara Alto .	34
Tabla 4. Pronósticos probabilísticos de sequía de la microcuenca del río Chulco	34
Tabla 5. Matrices de parámetros ϕ y matrices de varianza-covarianza G del modelo MPAR1 de las dos series de aportaciones de agua normalizadas y estandarizadas de los dos embalses Chanlud y El Labrado	36
Tabla 6. Umbrales de las dos series históricas de aportaciones de agua normalizadas y estandarizadas de los dos embalses Chanlud y El Labrado	36

1. INTRODUCCIÓN

Las cuencas hidrográficas proveen múltiples servicios ambientales y si son manejados propiamente pueden contribuir significativamente al desarrollo socioeconómico de una región. Uno de los servicios ambientales más estratégicos en cuencas de montaña son los servicios hidrológicos, los mismos que proveen grandes beneficios a la población. Condiciones hidroclimáticas de escasez de agua tal como las épocas de sequía pueden afectar o incluso interrumpir completamente estos servicios causando severos impactos sociales, económicos y ecológicos en una cuenca. Esta situación va a depender de la vulnerabilidad de los habitantes de la zona y su capacidad para predecir, adaptarse o responder ante estas condiciones de déficit del recurso hídrico (Brown et al. 2010). Por lo tanto estas épocas se han convertido en un verdadero desafío para los gestores en una cuenca (Li et al. 2009), con respecto a la asignación del agua para los diferentes usuarios y el respeto por los caudales ecológicos. Además el incontrolable crecimiento poblacional, aumento del consumo de agua y la disminución de la oferta del agua por efectos del cambio climático y cambio del uso del suelo, están direccionando a complejas situaciones donde no son fáciles las decisiones (Huang et al. 2010). Asimismo todavía existe una brecha entre la ciencia y los gestores del agua (Steinemann 2006), estos últimos se muestran renuentes a aceptar innovaciones científicas que no se convierten en medidas aplicables en la toma de decisiones (Liu et al. 2008a; Viviroli et al. 2011). Una razón para que suceda esta situación es que las personas encargadas de tomar decisiones son escépticos acerca del reemplazo de su juicio por estrategias prescritas de solución (Labadie and Asce 2004). Esto es aún más evidente en cuencas de montaña donde se tiene experiencia de gestión de recursos hídricos en condiciones normales, pero poca experiencia en escenarios de escasez de agua. No obstante, algunos autores sostienen que los efectos del cambio climático pueden traer largos periodos de sequía en el futuro especialmente en cuencas de montaña (Buytaert et al. 2006b). Lo cual afecta a los residentes de las tierras bajas que se benefician de los bienes y servicios de estos ecosistemas (Beniston 2003) y eleva la conciencia y preocupación sobre la vulnerabilidad de los sistemas hídricos por la potencial disminución de la disponibilidad de agua (Celleri et al. 2007). El desarrollo de metodologías y herramientas que ayuden para una gestión de sistemas de recursos hídricos en épocas de sequía ofrece un buen soporte para tomar decisiones con respecto a la asignación de agua que satisfaga las demandas consuntivas, no consuntivas y ambientales. Para este fin la caracterización y el pronóstico de eventos de sequía podrían ayudar considerablemente para que los gestores de agua tomen las decisiones

más apropiadas para un suministro de agua confiable y adverso al riesgo de fallo (Avilés et al. 2016).

Actualmente la comunidad de gestores utilizan ampliamente información en forma de índices (Shukla and Wood 2008) para mejorar la capacidad de analizar, evaluar y predecir eventos de sequía (Svoboda et al. 2004). Para este último aspecto existen varios índices, tales como: el índice estandarizado de precipitación (SPI) (McKee et al. 1993; Mishra and Desai 2005; Mishra et al. 2007; Paulo and Pereira 2007; Cancelliere et al. 2007; Moreira et al. 2008; Nalbantis and Tsakiris 2009; Khalili et al. 2011), el índice estandarizado de precipitación y evapotranspiración (SPEI) (Vicente-Serrano et al. 2010; Banimahd and Khalili 2013; Vicente-Serrano et al. 2015; Vicente-Serrano et al. 2016), el índice de severidad de sequía de Palmer (PDSI) (Palmer 1965; Lohani and Loganathan 1997), el índice de caudales de sequía (SDI) (Nalbantis and Tsakiris 2009), el índice de reconocimiento de sequía (RDI) (Tsakiris and Vangelis 2005; Khalili et al. 2011; Banimahd and Khalili 2013), el índice agregado de sequía (ADI) (Keyantash and Dracup 2004; Barua et al. 2011; Barua et al. 2012), entre otros. Diferentes situaciones hidroclimáticas en una cuenca condicionan la utilización de algunos índices, esto debido a la información requerida y la especificidad del cálculo para desarrollar estos indicadores (Mishra and Singh 2010; Barua et al. 2012). De hecho, la caracterización de sequías requiere indicadores que sean generalmente aplicables, pero también específicos en una región para capturar el tipo de sequías con la información disponible (Staudinger et al. 2014). Además un índice debería captar el resultado de la sucesión de varios eventos de escasez de agua en diferentes periodos de tiempo (Kao and Govindaraju 2010). Por lo tanto, existe la necesidad de evaluar los estados de sequía mediante la información disponible de variables relacionadas con el agua en diferentes ventanas de tiempo (Avilés et al. 2016). Es por esto que la presente tesis utiliza el índice de sequía (DI) desarrollado por Avilés et al. (2015), el cual tiene la ventaja de agrupar información hidrometeorológica con diferentes intervalos de tiempo en un solo índice. Este tiene como función identificar la frecuencia y severidad de los diferentes eventos de sequía.

Por otro lado, pronósticos confiables y oportunos de eventos de sequías juegan un rol importante en la toma de decisiones para disminuir los impactos de este fenómeno en sistemas de recursos hídricos (Madadgar and Moradkhani 2013; Madadgar and Moradkhani 2014). Varias metodologías de pronósticos de sequías se han desarrollado en los últimos años, tal como los análisis de regresión (Kumar and Panu 1997; Liu and Juárez 2001; Leilah and Al-Khateeb 2005a), los modelos de series de tiempo (Mishra and Desai 2005; Modarres

2007; Fernández et al. 2009; Durdu 2010; Han et al. 2010) y las redes neuronales (Mishra and Desai 2006; Mishra et al. 2007; Morid et al. 2007) son algunos de estos enfoques. Estos modelos proveen una predicción de los estados de sequía sin considerar la incertidumbre asociada con el pronóstico (Hwang and Carbone 2009). Este aspecto puede ser manejado con pronósticos probabilísticos, los cuales ofrecen un pronóstico asociado con su incertidumbre en una forma cuantitativa (Murphy 1977). Existen varios autores que han desarrollado modelos de pronóstico probabilístico de sequías (Lohani and Loganathan 1997; Chung and Salas 2000; Anderson et al. 2000; Paulo and Pereira 2007; Cancelliere et al. 2007; Hwang and Carbone 2009; Nalbantis and Tsakiris 2009; Araghinejad 2010; Madadgar and Moradkhani 2013; Yuan et al. 2013; Madadgar and Moradkhani 2014; AghaKouchak 2014; Avilés et al. 2015; Avilés et al. 2016), pero pocos pronostican probabilísticamente los estados futuros de sequía dada la información de los eventos previos, es decir mediante una probabilidad condicional, tal como los modelos más comunes basados en cadenas de Markov (MC) (Lohani and Loganathan 1997; Paulo and Pereira 2007; Cancelliere et al. 2007; Nalbantis and Tsakiris 2009; Avilés et al. 2015; Avilés et al. 2016) y los modelos más sofisticados basados en redes bayesianas (BN) (Madadgar and Moradkhani 2013; Madadgar and Moradkhani 2014; Avilés et al. 2016). Estos dos enfoques han sido comparados en un estudio reciente de Avilés et al. (2016), en donde se concluye que los modelos basados en MC demostraron ser igualmente eficientes para pronosticar probabilísticamente eventos de sequías que los modelos basados en BN. Además estos autores resaltan el mejor desempeño del modelo MC de primer orden (MCFO) para pronosticar periodos húmedos y secos con respecto al resto de modelos que fueron evaluados. Por esta razón esta tesis utiliza el modelo MCFO para pronosticar probabilísticamente eventos de sequía, los cuales tienen la ventaja de ser los más utilizados en procesos estocásticos de series de tiempo discretas, además su método de cálculo es simple y requiere menos costo computacional.

La caracterización y el pronóstico de eventos de sequía podrían mejorar la gestión y operación de sistemas hídricos, sin embargo una evaluación del riesgo de fallo en el suministro de agua para las diferentes demandas podría complementar la información de soporte para la toma de decisiones orientadas a minimizar o mitigar los efectos de la sequía en sistemas de recursos hídricos en cuencas reguladas (Haro et al. 2014). El riesgo de fallo de un sistema depende de los caudales en los ríos, las reservas disponibles y los valores de consumo de las demandas en una cuenca, los cuales pueden ser variables en el tiempo por fluctuaciones de las demandas y el cambio climático (Rossi et al. 2012). Indicadores que

cuantifiquen el riesgo de fallo y la satisfacción de un conjunto de demandas podrían ser una buena opción para una gestión de sistemas de recursos hídricos a corto y mediano plazo con escenarios de escasez de agua. Para este fin los modelos de simulación son quizás los métodos más ampliamente utilizados para analizar y evaluar alternativas de gestión del agua con diferentes escenarios, dada su simplicidad matemática y su fácil comprensión por parte de los gestores del agua (Cancelliere et al. 2009). Estos modelos generalmente se basan en el principio de conservación de la masa del sistema, siguen un enfoque nodo-arco para describir la red del sistema, donde los nodos pueden representar las demandas (usos consuntivos y no consuntivos), los embalses, etc., y los arcos pueden representar ríos, canales, etc. Existen varias herramientas de este tipo, sin embargo pocos modelos evalúan el riesgo de fallo de un sistema y producen índices de una manera comprensible para identificar políticas de operación y gestión. Este es el caso de SIMGES (Andreu et al. 2007) y SIMRISK (Sánchez et al. 2001) del sistema de soporte de decisiones AQUATOOL (Andreu et al. 1996). El proceso de simulación en SIMGES consiste en redes de flujo conservativas que se optimizan mensualmente mediante programación lineal con el algoritmo Out-of-Kilter (Cattrysse 1991; Andreu 1993), para cumplir una función objetivo (satisfacción de las demandas, operación de embalses, etc.) sujeta a las restricciones de conservación de masa (principio de continuidad) y los límites físicos de transporte de flujo en conducciones (ríos, canales, etc.) y capacidades de embalses (reservas) y otros elementos si los hubiere. Este proceso se lo puede realizar múltiples veces utilizando SIMRISK, el cual está basado en la simulación Montecarlo y evalúa el riesgo de insatisfacción de las demandas en diferentes niveles mediante información probabilística que permite analizar el número de fallos en el suministro a las demandas y su severidad, y mediante esta información los tomadores de decisiones podrían formular medidas de prevención y/o mitigación para hacer frente al riesgo y maximizar el rendimiento del sistema (Cancelliere et al. 2009). La simulación Montecarlo consiste en la generación de múltiples escenarios probables mediante modelos de generación sintética dada la naturaleza estocástica de las entradas hidrológicas a un sistema de recursos hídricos, y de acuerdo a objetivos y criterios decididos con anticipación realizar la gestión del sistema a corto o mediano plazo mediante reglas, medidas y procedimientos que minimicen el riesgo de desabastecimiento de agua (Cancelliere et al. 2009; Haro et al. 2014).

Algunos gestores del agua suelen desconfiar alejarse demasiado de sus prácticas habituales (Gong et al. 2010). Esto puede causar que las decisiones se direccionen a las condiciones medias y un alejamiento de las condiciones extremas, con la consecuente disminución de la

efectividad en la toma de decisiones. En este sentido los gestores algunas veces prefieren incorporar pronósticos de variables hidrometeorológicas dentro de las herramientas de gestión con el objetivo de entender la sensibilidad del sistema hídrico con respecto a la satisfacción de las demandas y mejorar la evaluación de los posibles riesgos de desabastecimiento, logrando más certeza en sus decisiones (Brown et al. 2010; Gong et al. 2010). Estos pronósticos combinados con las herramientas de simulación múltiple podrían condicionar y limitar los posibles escenarios futuros, facilitando la predicción de la disponibilidad de agua y la simulación del suministro de agua a las diferentes demandas (Brown et al. 2010), adicionalmente ofrecen ideas sobre las dificultades y retos de utilizar la información del pronóstico para mejorar la gestión del sistema de recursos hídricos (Sankarasubramanian et al. 2009; Pouget et al. 2015).

Por lo expuesto, el propósito de esta tesis fue desarrollar un marco metodológico integral de evaluación del riesgo de fallo en el suministro de agua mediante la incorporación de pronósticos probabilísticos de sequía, que proporcione ayuda a los gestores del agua y se convierta en una herramienta que disminuya la incertidumbre en la toma de decisiones para perfeccionar las medidas orientadas a prevenir o mitigar las consecuencias cuando existe escasez de agua.

1.1. OBJETIVOS

El objetivo de esta tesis fue desarrollar un marco metodológico integral de evaluación del riesgo de fallo en el suministro de agua que incorpore pronósticos probabilísticos de eventos de sequía y pueda proporcionar un soporte en la toma de decisiones con respecto a la gestión de sistemas de recursos hídricos en tiempos de escasez.

Esta tesis puede dividirse en tres objetivos parciales de la siguiente manera.

- i. Caracterizar eventos de sequía, mediante un análisis y construcción de un índice de sequía (DI) priorizando información disponible de variables hidrometeorológicas a diferentes escalas de tiempo.
- ii. Modelar, evaluar y comparar herramientas de pronóstico probabilístico de sequías usando el DI, mediante modelos basados en cadenas de Markov (MC) y modelos basados en redes bayesianas (BN), comparando su desempeño y eligiendo el modelo de pronóstico que presente el mejor rendimiento.

- iii. Explorar y diseñar un marco metodológico general de evaluación del riesgo de fallo en sistema de recursos hídricos incorporando pronósticos probabilísticos de sequías, mediante la generación sintética hidrológica que admita las probabilidades de ocurrencia de los estados de sequía y la utilización de modelos de simulación de múltiples escenarios.

1.2. PUBLICACIONES

Esta tesis plantea un marco metodológico integral de evaluación del riesgo de fallo en el suministro de agua en sistemas de recursos hídricos en épocas de sequía acogiendo los aportes de cuatro publicaciones desarrolladas en el transcurso del periodo de investigación del programa de Doctorado en Ingeniería del Agua y Medioambiental de la Universidad Politécnica de Valencia. Las cuales se detallan a continuación:

1.2.1. EVALUACIÓN DE PRONÓSTICOS DE SEQUÍAS BASADOS EN CADENAS MARKOV EN UNA CUENCA ANDINA REGULADA USANDO LOS PUNTAJES DE HABILIDAD RPS Y GMSS

Avilés A, Célleri R, Paredes J, Solera A (2015) Evaluation of Markov Chain Based Drought Forecasts in an Andean Regulated River Basin Using the Skill Scores RPS and GMSS. *Water Resources Management* 29:1949–1963. doi: 10.1007/s11269-015-0921-2

El objetivo de la primera publicación que se presenta en el Anexo 1 fue desarrollar un DI para caracterizar y categorizar la frecuencia y el nivel de severidad de eventos de sequías. Además este artículo tuvo también el propósito de evaluar el rendimiento de dos modelos de pronóstico de sequía mensual y estacional basados en MC usando la serie de tiempo categorizada del DI. Se comparó el desempeño de los modelos cadena de Markov de primer orden (MCFO) y cadena de Markov de segundo orden (MCSO) usando dos puntajes de habilidad del pronóstico de variables discretas, tal como el puntaje jerarquizado de probabilidad (RPS) y el puntaje de habilidad Gandin-Murphy (GMSS). Los resultados de estudio mostraron la importancia de verificar la calidad de los pronósticos mediante uno o varios puntajes de habilidad previo a la utilización de esta información en la toma de decisiones con respecto a reglas y medidas preventivas de gestión del agua en periodos de escasez.

La metodología utilizada para el desarrollo del DI fue un enfoque flexible que engloba y prioriza información de variables relacionadas con el agua a diferentes escalas temporales brindando la posibilidad de caracterizar las sequías de una manera integral y facilitando la obtención de indicadores de déficit de agua con los datos disponible en una cuenca. Una serie de tiempo categórica del DI transformado se utilizó para pronosticar cuatro estados de sequía, extrema (categoría 3), severa (categoría 2), moderada (categoría 1) y no sequía (categoría 0).

Para tomar en cuenta la estacionalidad del ciclo anual se eligieron los modelos no-homogéneos MC con la finalidad de obtener un modelo de pronóstico para cada mes del año. El modelo MCFO utilizó la información del mes actual para pronosticar los estados de sequías del mes siguiente, mientras que el modelo MCSO utilizó la información de los meses actual y anterior para pronosticar los estados de sequía un mes adelante. Además se realizó predicciones tres meses adelante utilizando los modelos MCFO y MCSO mediante árboles de decisiones construidos por un encadenamiento de pronósticos en cada mes mediante la información de matrices de probabilidad de transición y el conocimiento de las predicciones en los meses previos. Los resultados de este proceso demostraron ser cómodamente manejables y de fácil entendimiento por la ventaja de mostrar la incertidumbre asociada con el pronóstico en una forma cuantitativa.

Los aportes de esta publicación para el contexto de esta tesis fueron:

- El desarrollo del DI a partir de la disponibilidad de información hidrológica y meteorológica a diferentes escalas temporales en una cuenca.
- La utilización del enfoque basado en MC para obtener pronósticos de sequías en un marco probabilístico, brindando información de la incertidumbre asociada con el pronóstico en una forma cuantitativa.
- La importancia de verificar la calidad de los pronósticos y el desempeño de los modelos mediante uno o varios puntajes de habilidad de la predicción antes de aceptar los resultados para la toma de decisiones en la gestión del agua.

1.2.2. PRONÓSTICO PROBABILÍSTICO DE EVENTOS DE SEQUÍAS USANDO MODELOS BASADOS EN CADENAS DE MARKOV Y REDES BAYESIANAS: UN CASO DE ESTUDIO DE UNA CUENCA ANDINA REGULADA

Avilés A, Célleri R, Solera A, Paredes J (2016) Probabilística Forecasting of Drought Events Using Markov Chain-and Bayesian Network-Based Models: A Case Study of an Andean Regulated River Basin. *Water* 8:1–16. doi:10.3390/w8020037

El propósito de la segunda publicación que se presenta en el Anexo 2 fue comparar modelos de pronósticos de sequías basados en la probabilidad condicional (pronóstico de eventos futuros dada la información de los eventos previos), mediante un enfoque clásico usado frecuentemente basado en MC y otro enfoque novedoso usado con menos frecuencia basado en BN. Ambos enfoques usaron el DI desarrollado en la primera publicación y verificaron la calidad de los pronósticos mediante el puntaje de habilidad ponderado de probabilidad (RPSS), derivado del puntaje ponderado de probabilidad (RPS) y los pronósticos de referencia. Las funciones copulas fueron usadas para resolver la compleja tarea de resolver los modelos basados en BN.

Mediante la agregación de información disponible hidrológica y meteorológica en diferentes escalas se construyó un DI utilizando el método expuesto en la primera publicación. La serie de tiempo resultante del DI fue categorizada en tres estados (no sequía, leve sequía y sequía), formando una nueva serie de datos categóricos que fueron la entrada de los modelos basados en MC, mientras que los valores de la serie de tiempo de la función de distribución normal acumulada del DI fueron las entradas para los modelos basados en BN, dado que la serie de tiempo del DI cumple la hipótesis de normalidad.

La calibración y validación de los modelos se realizó mediante validación cruzada dejando fuera uno. Se construyeron cuatro modelos basados en la probabilidad condicional para realizar los pronósticos de sequías, los modelos MC de primer orden (MCFO) y segundo orden (MCSO) y los modelos BN de primer orden (BNFO) y segundo orden (BNSO). Las funciones copulas fueron utilizadas para ayudar a resolver los cálculos de probabilidad conjunta en la formulación de los modelos basados en BN. Dos funciones copulas elípticas (normal y t) y dos funciones copulas arquimedianas (Clayton y Frank) se analizaron para definir la función más apropiada para las distribuciones conjuntas. El método máxima verosimilitud canónica (CML) fue utilizado para la estimación de parámetros y para seleccionar las mejores funciones ajustadas se utilizó el test de bondad de ajuste paramétrico

basado en bootstrap, mediante la comparación del valor de p con el nivel de significancia y escogiendo las copulas con el menor valor del estadístico Cramér-von-Mises (S).

Los resultados de una evaluación global de los pronósticos mostraron que los modelos basados en MC tienen mejor desempeño que los modelos basados en BN para pronosticar los eventos húmedos y secos, mientras que para los eventos de sequía los modelos basados en BN tienen un leve mejor rendimiento que los modelos basados en MC. Sin embargo, realizando una evaluación mensual se detectó que ambos enfoques tuvieron mejor rendimiento en pronosticar las sequías dependiendo del mes de análisis. Además en la mayoría de los análisis de verificación de la calidad de los pronósticos el modelo MCFO fue el que presentó mejor desempeño.

Los aportes de esta publicación para el contexto de esta tesis fueron:

- El análisis de la comparación del desempeño de modelos comúnmente utilizados (MC) y modelos complejos utilizados con menor frecuencia (BN) para el pronóstico probabilístico de sequías.
- La elección del mejor modelo con criterios probabilísticos y pronósticos de referencia.
- El mejor rendimiento del modelo MCFO según la evaluación de los pronósticos.
- La utilidad de comparar varios enfoques de pronóstico de sequía con la finalidad de elegir un modelo confiable que podría ser útil para asegurar una toma de decisiones oportuna de los gestores del agua para enfrentar los periodos de sequías

1.2.3. ESTIMACIÓN DEL RIESGO DE FALLO EN EL SUMINISTRO DE AGUA COMO AYUDA A LA PLANIFICACIÓN Y GESTIÓN DE RECURSOS HÍDRICOS

Avilés A, Solera A (2012) Estimación del riesgo de fallo en el suministro de agua como ayuda a la planificación y gestión de recursos hídricos. *Maskana* 3:77–86.

Este artículo se presenta en el Anexo 3 tiene el propósito de estimar el riesgo de fallo en el suministro de agua para explorar alternativas de gestión de la demanda en condiciones futuras mediante un análisis de modelos de series de tiempo y herramientas de simulación múltiple de escenarios estocásticos.

Se analizaron varios modelos de series de tiempo de caudales, tal como los autoregresivos de media móvil (ARMA) y modelos autoregresivos de media móvil periódicos (PARMA) con la finalidad de obtener un conjunto de modelos que sean capaces de reproducir las propiedades estadísticas de los datos históricos y cumplir los supuestos de normalidad e independencia de los residuos. Además se escogieron los mejores modelos mediante criterios de penalización por mayor número de parámetros. Estos modelos sirvieron para generar múltiples series sintéticas de caudales, las cuales fueron la base para la construcción de escenarios estocásticos futuros que ingresaron en las herramientas de simulación (SIMGES y SIMRISK). El proceso de simulación en SIMRISK se realizó mediante la introducción de los múltiples escenarios estocásticos y de acuerdo a criterios decididos con anticipación se valoró de una forma probabilística el riesgo de fallo del suministro de agua en diferentes niveles de déficit (diferencia entre el valor de la demanda y el valor del suministro).

Este banco de información probabilística sirvió para explorar alternativas de gestión de la demanda para la disminución del riesgo de fallo en la asignación de agua principalmente a las demandas prioritarias.

Los aportes de esta publicación para el contexto de esta tesis fueron:

- El análisis de la comparación entre los modelos ARMA y PARMA, en donde se muestra la ventaja de los modelos PARMA para captar la estacionalidad de los caudales y reproducir de mejor forma los estadísticos mensuales de la serie histórica. Y el mejor rendimiento de los modelos PAR1 con respecto a otros modelos PARMA con mayor número de parámetros.
- La necesidad de realizar una gestión a corto o mediano plazo y con actualizaciones constantes de la información de entrada a los modelos dada la incertidumbre de los cambios futuros en la oferta y demanda de agua cuando se realiza una simulación a largo plazo. Este aspecto podría realizarse mediante una predicción en periodos muy cortos de la distribución de todos los posibles valores de caudales y la introducción de información actualizada de las demandas.

1.2.4. ANÁLISIS DE SISTEMAS DE RECURSOS HÍDRICOS DE LA CUENCA DEL RIO TOMBAMBA EN ECUADOR, MEDIANTE MODELOS ESTOCÁSTICOS Y DE GESTIÓN

Avilés A, Solera A (2013) Análisis de sistemas de recursos hídricos de la cuenca del río Tomebamba en Ecuador, mediante modelos estocásticos y de gestión. En: Solera A, Paredes

J, Andreu J (editores) Aplicaciones de sistemas soporte a la decisión en planificación y gestión integradas de cuencas hidrográficas. Marcombo, Barcelona, España, pp 51–61

La cuarta publicación que presenta en el Anexo 4 es una aplicación de los métodos de la tercera publicación en la cuenca del río Tomebamba ubicada en el sur de los Andes ecuatorianos. El propósito de este estudio fue mostrar las bondades de trabajar con una metodología que incorpore un marco probabilístico en la toma de decisiones y que facilite la búsqueda de alternativas de gestión para disminuir el riesgo de desabastecimiento de agua a las demandas prioritarias.

Se utilizaron los mejores modelos PARMA encontrados en la tercera publicación para la generación de series sintéticas hidrológicas que sirvan para la construcción de escenarios estocásticos. Estos escenarios fueron ingresados en la simulación múltiple utilizando SIMRISK y se determinó las probabilidades de fallo en el suministro de agua a las distintas demandas y la severidad de los fallos (gravedad del déficit). El análisis de estos resultados sirvió para determinar el estado del sistema de recursos hídricos y para formular medidas de gestión en las demandas para que el riesgo de fallo sea el menor posible. Al final mediante un ejemplo se mostró que la formulación de una medida de restricción de las demandas menos prioritarias ayudó para que se reduzca el riesgo de fallo en las demandas con mayor prioridad.

Los aportes de esta publicación para el contexto de esta tesis fueron:

- La necesidad de tener indicadores que agrupen los resultados probabilísticos para una fácil comprensión del estado del sistema, dado que el análisis del riesgo de fallo cuando existen múltiples demandas puede llegar a convertirse en una tarea confusa y complicada.
- El desafío de buscar metodologías que ayuden para que las decisiones sean más objetivas y permitan obtener un marco amplio de información para tomar decisiones y por lo tanto disminuya la subjetividad para encontrar un valor adecuado del riesgo de fallo.

1.3. ESTRUCTURA DE LA TESIS

La tesis se estructura en cinco secciones. La sección 1 describe los antecedentes, la motivación y la importancia del estudio, los objetivos de la tesis y las publicaciones

realizadas en el transcurso del periodo de la investigación. La sección 2 detalla el caso de aplicación, mientras que la sección 3 presenta las herramientas utilizadas y los métodos desarrollados según los objetivos propuestos. Una descripción de los resultados obtenidos es expuesta en la sección 4. La sección 5 presenta las conclusiones y las futuras investigaciones. Los anexos finalizan el documento mostrando en detalle las publicaciones de la sección 1.

2. CASO DE ESTUDIO

Los métodos expuestos posteriormente en este estudio se aplicaron en la cuenca del río Machángara (325 Km²) ubicada en el sur de los Andes ecuatorianos a una altitud de 2440 – 4398 m.s.n.m. (Figura 1). Esta cuenca tiene una singular importancia porque posee uno de los pocos sistemas de recursos hídricos multipropósito en el sur de Ecuador que beneficia a la economía y ecología local y regional. El río Machángara nace en la parte más alta de la cordillera occidental de los Andes para luego ser un tributario del río Paute. La parte alta y media de la cuenca son poco intervenidas y están principalmente formadas por páramos (llanuras con una gran variedad de lagos, humedales y pastizales húmedos entremezclados con matorrales y parches de bosques de baja estatura) (Buytaert et al. 2006a). En la parte alta se ubican dos embalses Chanlud (16 hm³) y El Labrado (6 hm³) que suministran agua para diferentes usos, el primero ubicado en la microcuenca del río Machángara Alto y el último ubicado en la microcuenca del río Chulco. Entre los usos más importantes está el de consumo humano para aproximadamente 240000 habitantes de la ciudad de Cuenca, la producción agropecuaria con 1300 Ha regadas y la producción energética con 40 MW. En la parte media de la cuenca existe producción ganadera, agrícola y en pequeñas proporciones producción maderera de bosques nativos. En la parte baja el uso de agua está destinada para producción agrícola, de consumo e industrial. Es así que la competencia por los diversos usos que ejercen los diferentes actores despliega una presión cada vez más importante en los recursos hídricos, esto sumado a los eventos de sequías, probablemente incrementados en su severidad, duración y frecuencia por los efectos de la variabilidad climática y el cambio climático, hacen prever menor disponibilidad del recurso hídrico en el futuro. Desde este punto de vista, se pudo observar potenciales investigaciones que se derivan de las características particulares de esta cuenca, y consecuentemente el propósito de la presente tesis.

Para el desarrollo del DI se utilizó series de tiempo mensuales (1971 – 2010) de precipitación media de las microcuencas de los ríos Machángara Alto y Chulco y el volumen de agua que ingresa a los embalses de Chanlud y El Labrado, recolectadas del Instituto Nacional de Meteorología e Hidrología de Ecuador (INAMHI) y del Consejo de la cuenca del río Machángara (CBRM). La Figura 2 muestra la variación del promedio mensual de la lluvia y el volumen de agua de entrada a los embalses. De una manera similar como la primera y segunda publicación (Anexos 1 y 2), se trabajó con cinco ventanas de tiempo (1, 3, 6, 9 y 12 meses) para generar cinco series de tiempo de la lluvia (PR1, PR3, PR6, PR9 y PR12) y cinco series de tiempo del volumen de agua de ingreso a los embalses (VS1 VS3, VS6, VS9 y VS12) con diferentes escalas temporales y para cada embalse. Para tomar en cuenta la estacionalidad mensual se dividió cada serie según cada mes del año, además toda la información fue estandarizada.

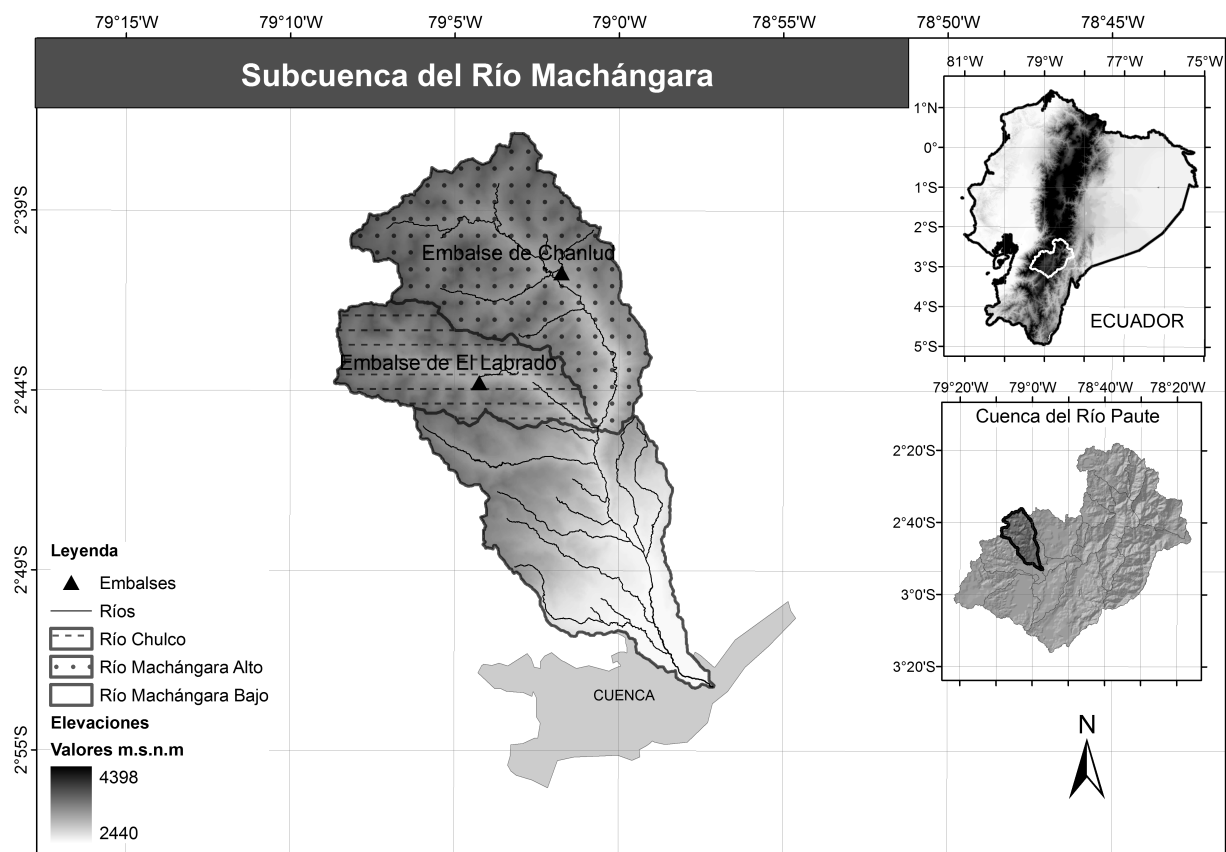


Figura 1. Localización de la cuenca del río Machángara

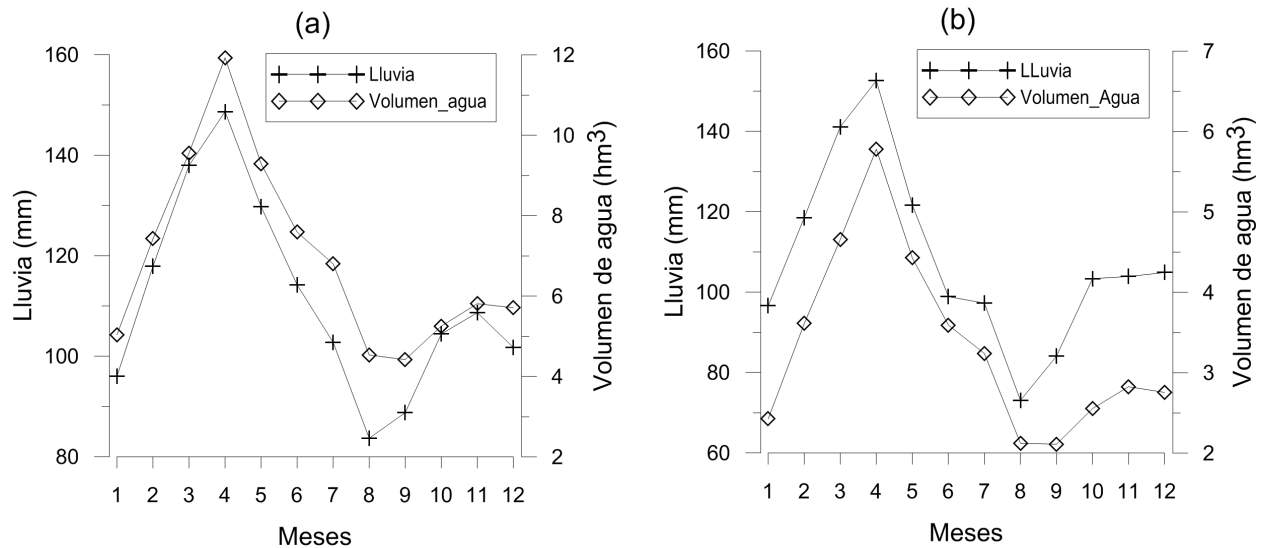


Figura 2. Promedio mensual de las series de tiempo de lluvia y volumen de agua de entrada a los embalses (1971 – 2010) de las microcuencas de los ríos (a) Machángara Alto y (b) Chulco

Para cuantificar las demandas de agua actuales se utilizó la información proporcionada por el Consejo de la cuenca del río Machángara (CBRM). Se obtuvo los datos de los tres usos de agua más importantes en la cuenca, tomando como demanda prioritaria el consumo humano, luego el agua para riego y por último la generación de energía. Se consideró un caudal ecológico equivalente al 10% de los caudales medios mensuales.

Los volúmenes mensuales de agua de entrada a los embalses se utilizaron para la generación de series sintéticas hidrológicas. Estas últimas conjuntamente con la información de las demandas de agua y las reservas de los embalses fueron los datos de entrada para el modelo de evaluación del riesgo de fallo del sistema, tomando en cuenta el mes inicial de simulación. El periodo de simulación fue de 12 meses, es decir los resultados del modelo servirán para gestión y operación del sistema resguardando las decisiones de todo el año hidrológico. La Figura 3 muestra un esquema del sistema de recursos hídricos de la cuenca del río Machángara.

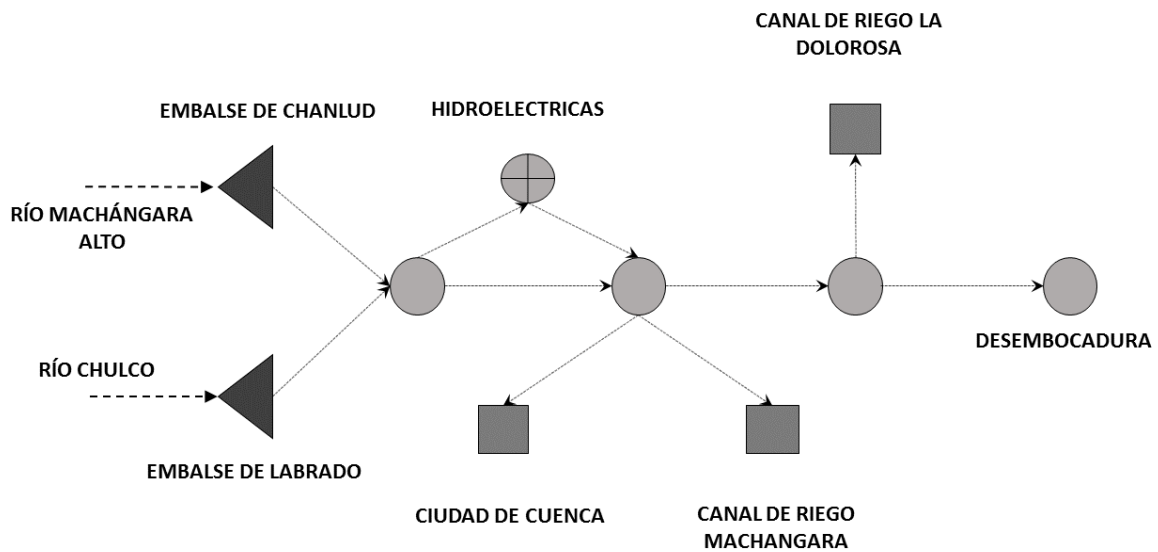


Figura 3. Esquema del sistema de recursos hídricos de la cuenca del río Machángara

3. PROPUESTA DE MEJORA DE LA METODOLOGÍA GENERAL DE EVALUACIÓN DEL RIESGO DE FALLO DE SISTEMAS DE RECURSOS HÍDRICOS

En este apartado se describe la metodología general de evaluación del riesgo de fallo en sistemas de recursos hídricos y la nueva propuesta de mejora con la incorporación de los pronósticos probabilísticos de sequía.

La metodología tradicional de estimación del riesgo de fallo en el suministro de demandas hídricas fue inicialmente desarrollada por Sánchez et al. (2001), aplicado por algunos estudios (Avilés and Solera 2012; Avilés and Solera 2013; Haro et al. 2014) y presentado de una forma similar en Cancelliere et al. (2009). Este método está representado en la Figura 4 y consiste en los siguientes pasos: i) Mediante un modelo estocástico de series de tiempo de caudales se realiza una generación sintética hidrológica condicionada a las observaciones previas, con lo cual se genera múltiples escenarios de los posibles caudales futuros; ii) Con los múltiples escenarios generados en el paso anterior, las características actuales del sistema de recursos hídricos y las reglas de gestión del sistema, se realiza la simulación múltiple de la gestión futura; iii) Con los resultados de la simulación múltiple se realiza un análisis

estadístico, obteniendo las probabilidades de fallo del suministro de agua de las demandas de agua del sistema de recursos hídricos; iv) La información del paso anterior determina el estado del sistema y se toma la decisión si el riesgo es admisible; v) Si el encargado de tomar decisiones no acepta el riesgo resultante, entonces se realiza la formulación de alternativas de gestión, las cuales ingresan nuevamente al modelo de simulación múltiple (paso 2) y se continua con los siguientes pasos hasta decidir nuevamente si el riesgo es aceptable. Este proceso se repite reiteradamente hasta conseguir que el riesgo asociado a la decisión sea aceptable.

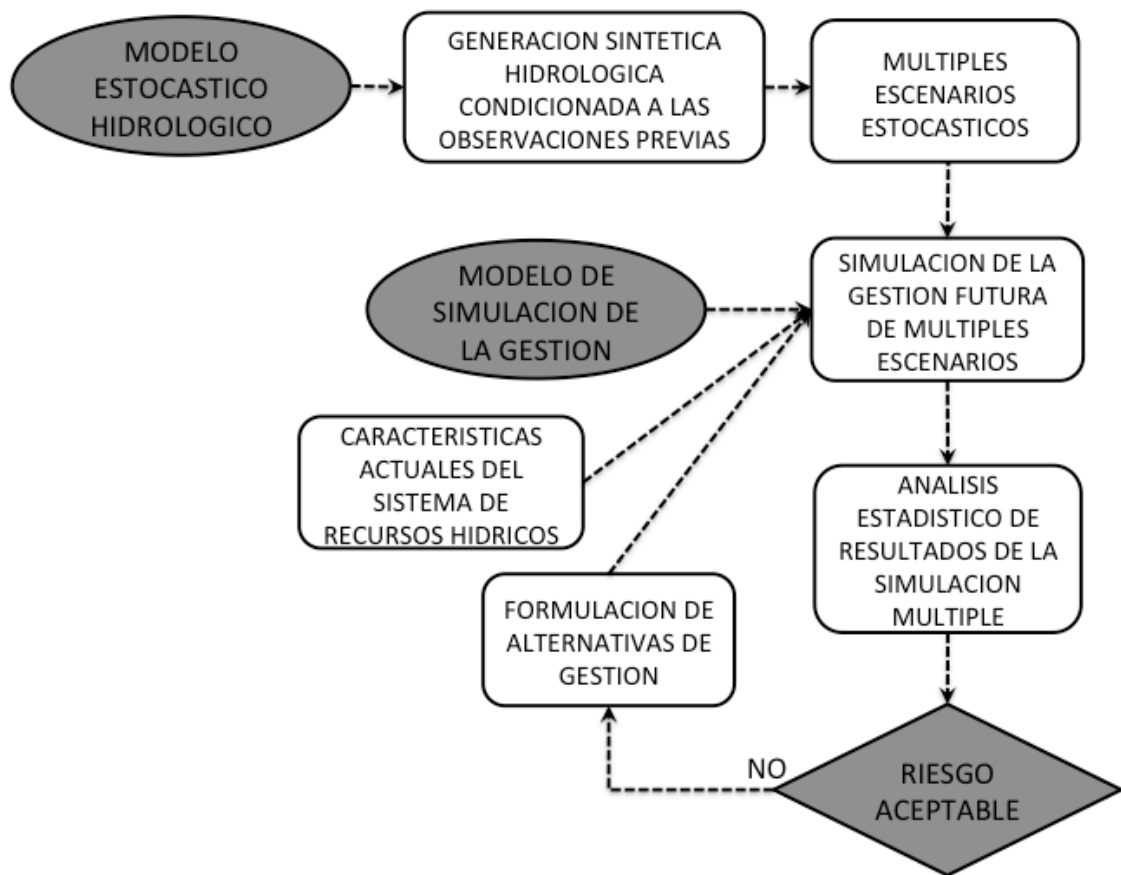


Figura 4. Metodología de evaluación del riesgo de fallo en sistemas de recursos hídricos

Esta metodología se utilizó en la tercera y cuarta publicación (Anexos 3 y 4) en donde se obtuvo varias enseñanzas, reflexiones e inquietudes que aportaron para desarrollar nuevos elementos para establecer un marco metodológico integral que obtenga resultados más efectivos para tomar decisiones en épocas de sequía.

En esta tesis doctoral se propone un marco metodológico integral que se muestra en la Figura 5 y se describe en los siguientes pasos: i) Mediante un modelo estocástico de generación sintética hidrológica se realiza la incorporación de los pronósticos probabilísticos de sequía desarrollados a partir de un DI; ii) Como resultado del paso previo se obtiene series sintéticas hidrológicas condicionadas a los datos previos, las cuales conjuntamente con las características actuales del sistema de recursos hídricos, los estados previos de sequía y las reglas de gestión del sistema, se utilizaron para realizar la simulación de la gestión futura de los múltiples escenarios; iii) Los resultados del proceso anterior generó varios indicadores del riesgo de fallo en el suministro de agua de las demandas del sistema hídrico que se agruparon para formar índices integrales de satisfacción, los cuales podrían servir para tomar decisiones sobre la gestión del sistema; iv) En caso de que la satisfacción no sea aceptable se podría formular alternativas de gestión y nuevamente realizar la simulación de la gestión futura con múltiples escenarios (paso 2); v) Este proceso se repetiría hasta conseguir una satisfacción aceptable.

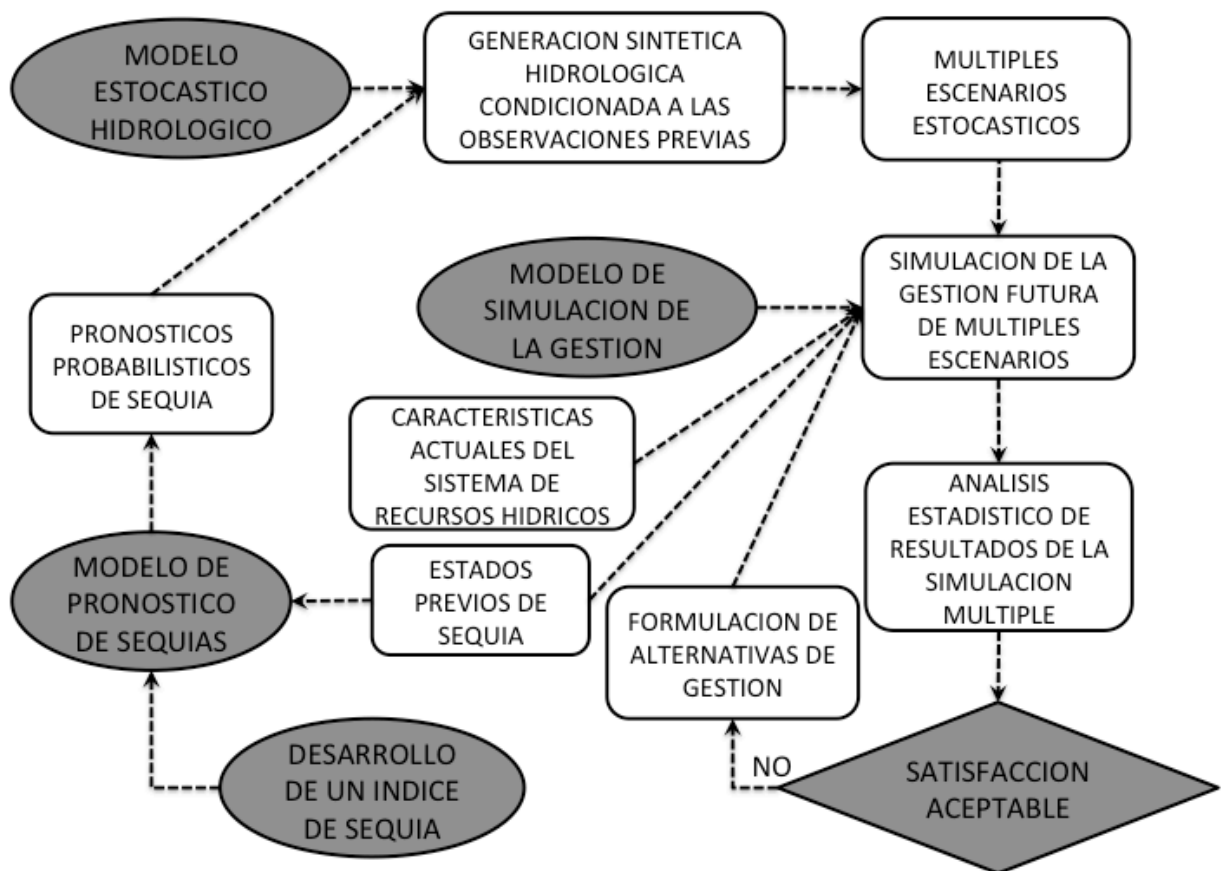


Figura 5. Marco metodológico integral de evaluación del riesgo de fallo en sistemas de recursos hídricos

Esta metodología general se aplicó en la cuenca del río Machángara y podría aplicarse en cuencas andinas reguladas con características similares con la finalidad de que sirva de ayuda para la toma de decisiones de los gestores del agua en épocas de sequía. La sección 3.1 describe el desarrollo del DI, la sección 3.2 comprende el proceso de modelación para obtener los pronósticos probabilísticos de sequías, la sección 3.3 describe el modelo estocástico para la generación sintética hidrológica, la sección 3.4 detalla la incorporación de los pronósticos de sequía en la generación de series sintéticas y finalmente la sección 3.5 describe el procedimiento de simulación múltiple para la evaluación del riesgo de fallo y el tratamiento estadístico de los resultados para la toma de decisiones.

3.1. ÍNDICE DE SEQUÍA

Para la construcción del DI se utiliza una metodología matemática similar a lo expuesto por (Keyantash and Dracup 2004), donde la información disponible de las r variables relacionadas con el agua se someten a un análisis de componentes principales (PCA). Los autovectores derivados de un PCA establecen la relación entre los componentes principales (PCs) y los datos originales:

$$S = D * E \quad (1)$$

donde S es la matriz ($w \times r$) de los PCs (en donde w es el número de observaciones), D es la matriz ($w \times r$) de la información original estandarizada, y E es la matriz ($r \times r$) de los autovectores. El DI es el primer componente principal (PC1), normalizado por su desviación estándar:

$$DI_{i,k} = \frac{S_{i,1,k}}{\sigma_k} \quad (2)$$

donde $DI_{i,k}$ es el valor del DI para el mes k en el año i , $S_{i,1,k}$ es el PC1 durante el año i , para el mes k , y σ_k es la desviación estándar de la muestra de $S_{i,1,k}$. El DI utiliza solo el PC1 porque este explica la más grande fracción de la varianza descrita por el conjunto completo de datos de las r variables, además por ser los PCs vectores ortogonales, no es apropiado matemáticamente combinarlos en una sola expresión (Keyantash and Dracup 2004). Una vez

que los valores del DI son calculados para cada año y cada mes, estos se reordenan en una sola serie de tiempo en orden cronológico.

El DI es un índice estandarizado capaz de captar las anomalías de las condiciones de humedad promedio de una cuenca en función de la información disponible de las variables relacionadas con el agua (Kao and Govindaraju 2010; Madadgar and Moradkhani 2013). Cualquier fenómeno que se puede cuantificar continuamente, como el índice de sequía, se lo puede tratar como una variable discreta mediante la categorización de la serie de tiempo teniendo en cuenta los umbrales para cada estado de la sequía (Avilés et al. 2016). Por lo tanto, el DI como una variable estandarizada fue dividido en categorías para caracterizar los estados de sequía, utilizando los mismos umbrales de la Organización Meteorológica Mundial (World Meteorological Organization 2012). Según esta última referencia las categorías que fueron derivadas son las siguientes: $DI > 0$ = categoría 0 (no sequía); $-1 < DI \leq 0$ = categoría 1 (leve sequía); $DI \leq -1$ = Categoría 2 (moderada, severa y extrema sequía). Los tres estados de la categoría 2 se tomaron como un solo estado llamado sequía. Esta nueva serie de tiempo de valores categóricos es la entrada de los modelos de cadenas de Markov y la serie de tiempo de la función de distribución normal acumulada de los valores del DI es la entrada de los modelos de la red bayesiana. La calibración y validación de los modelos se realizó aplicando el procedimiento de validación cruzada.

3.2. PRONÓSTICO PROBABILÍSTICO DE SEQUÍA

Remitiéndonos a lo realizado en la primera y segunda publicación (Anexos 1 y 2) y con la finalidad de presentar el proceso de investigación de la elección del modelo de pronóstico probabilístico de sequía que se utiliza en esta tesis, se detalla a continuación los métodos de cálculo desarrollados.

3.2.1. CADENAS DE MARKOV

El comportamiento de las cadenas de Markov está gobernado por un conjunto de matrices de probabilidades de transición que indican las probabilidades de ocurrencia de los estados de un sistema para un intervalo de tiempo del futuro a partir de la información del estado actual y/o estados de los intervalos pasados, dependiendo de la orden del modelo. La propiedad Markoviana del modelo de cadena de Markov de m^{th} orden es:

$$P(Y_{tn}|Y_{tn-1}, Y_{tn-2}, Y_{tn-3}, \dots, Y_1) = P(Y_{tn}|Y_{tn-1}, Y_{tn-2}, \dots, Y_{tn-m}) \quad (3)$$

Considerando una cadena de Markov de primer orden (MCFO), es decir, $m = 1$, las probabilidades de transición proporcionan el pronóstico probabilístico del estado un paso adelante basado en el estado actual, aplicando la siguiente fórmula:

$$p_{ij} = P(Y_{tn} = j|Y_{tn-1} = i) \quad (4)$$

donde p_{ij} representa la probabilidad de transición que Y_{tn} es igual a la categoría j dado que Y_{tn-1} es igual a la categoría i . La probabilidad de transición estimada (\hat{p}_{ij}) puede ser calculada tomando en cuenta las frecuencias relativas condicionales de las transiciones (f_{ij}):

$$\hat{p}_{ij} = \frac{f_{ij}}{\sum_j f_{ij}} \quad i, j = 1, \dots, s \quad (5)$$

donde f_{ij} es la frecuencia que Y es igual a la categoría i en el tiempo t_{n-1} e igual a la categoría j en el tiempo t_n . El valor de s es el número de estados del sistema. El numerador presenta el número de transiciones de categoría i a categoría j y el denominador representa la suma del número de transiciones de la categoría i a cualquier otra categoría.

Si $m = 2$, tendríamos una cadena de Markov de segundo orden (MCSO), las probabilidades de transición dependen de los estados del periodo de tiempo anterior y actual, proporcionando el pronóstico probabilístico del estado del siguiente paso de tiempo. Las probabilidades de transición se calculan como sigue:

$$p_{hij} = \Pr\{Y_{tn} = j|Y_{tn-1} = i, Y_{tn-2} = h\} \quad (6)$$

donde p_{hij} representa la probabilidad de transición que Y_{tn} es igual a la categoría j , dado que Y_{tn-1} es igual a la categoría que i y Y_{tn-2} es igual a la categoría h . Las probabilidades de transición estimadas (\hat{p}_{hij}) se obtienen del conteo de las frecuencias relativas condicionales de transición (f_{hij}):

$$\hat{p}_{hij} = \frac{f_{hij}}{\sum_j f_{hij}} \quad h, i, j = 1, \dots, s \quad (7)$$

donde el numerador es el número de transiciones de la categoría h en el tiempo t_{n-2} , categoría i en el tiempo t_{n-1} y la categoría j en el tiempo t_n y el denominador es la suma del número de

transiciones de las categorías h_i a cualquier otra categoría. Para esta condición, la probabilidad del estado de sequía del próximo período de tiempo depende de los estados de los dos períodos anteriores.

3.2.2. REDES BAYESIANAS

Heckerman (1998) asevera que una red bayesiana para un conjunto de variables $X = \{X_1, \dots, X_n\}$ consiste en una estructura de red (NS) que codifica un conjunto de afirmaciones de independencia condicional de las variables X y un conjunto P de distribuciones de probabilidad locales asociadas a cada variable. Conjuntamente estos componentes definen la distribución de probabilidad conjunta de X . El NS es un gráfico no cíclico dirigido (DAG), que consiste en la secuencia de eventos o variables aleatorias con arreglos directos, como los eventos que evolucionan en el tiempo (Thulasiraman and Swamy 2011; Madadgar and Moradkhani 2014). Dada la NS, la distribución de probabilidad conjunta de X , viene dada por:

$$P(X) = \prod_{i=1}^n P(X_i | P_{i1}) \quad (8)$$

donde X_i es la variable y su correspondiente nodo y P_{i1} son los padres del nodo X_i en la NS. La información probabilística de los nodos representa la influencia de sus padres en el gráfico sobre el nodo.

Siguiendo una similar formulación matemática como en Madadgar y Moradkhani (2013; 2014) y asumiendo un conjunto de variables X como variables aleatorias que evolucionan en el tiempo (por ejemplo un índice de sequía), $X = \{X_{t1}, \dots, X_{tn}\}$, donde el orden de dependencia sigue la secuencia temporal y los padres de X_{ti} son el conjunto de todas las variables previas (X_{t1}, \dots, X_{ti-1}), por lo tanto la ecuación 8 se pueden expresar de la siguiente manera:

$$P(X_{t1}, \dots, X_{tn}) = \prod_{i=1}^n P(X_{ti} | X_{t1}, \dots, X_{ti-1}) \quad (9)$$

La ecuación 9 representa la regla de la cadena en la teoría de las probabilidades, la misma que se puede resolver de la siguiente forma:

$$P(X_{t1}, \dots, X_{tn}) = P(X_{tn}|X_{t1}, \dots, X_{tn-1}) P(X_{t1}, \dots, X_{tn-1}) \quad (10)$$

Reordenando la Ecuación 10 se obtiene:

$$P(X_{tn}|X_{t1}, \dots, X_{tn-1}) = \frac{P(X_{t1}, \dots, X_{tn})}{P(X_{t1}, \dots, X_{tn-1})} \quad (11)$$

Ecuación 11 representa un modelo basado en BN, donde se calculan las probabilidades condicionales de la variable de pronóstico (X_{tn}) dado las variables predictoras (X_{t1}, \dots, X_{tn-1}). El cálculo de las distribuciones de probabilidad conjunta en el lado derecho de la ecuación 11 es relativamente tarea engorrosa, que puede simplificarse considerablemente usando las funciones copulas.

3.2.2.1 Copulas

Copulas son funciones que unen funciones de distribución multivariadas cuyas distribuciones marginales son uniformes en el intervalo (0,1) (Nelsen 2006). Aceptando la definición matemática de copulas de Yan (2006), con un vector aleatorio $(U_1, \dots, U_n)^T$, cuyas marginales U_i , $i=1, \dots, n$, son variables aleatorias uniformes en el intervalo unidad y suponiendo que la función de distribución acumulada conjunta (CDF) de $(U_1, \dots, U_n)^T$, es de la siguiente manera:

$$C(U_1, \dots, U_n) = P(U_1 \leq u_1, \dots, U_n \leq u_n) \quad (12)$$

donde la función C se denomina una copula n-dimensional y CDF define las probabilidades de que la variable aleatoria X no superará valores particulares (Wilks 2011), es decir, $U(X) = P(X \leq x)$. Por otra parte, el teorema de Sklar (Sklar 1959) explica el papel que juegan las copulas en la relación entre las funciones de distribución multivariadas y sus distribuciones marginales univariadas (Nelsen 2006). El teorema establece que una función de distribución multivariada conjunta $F(X_1, \dots, X_n)$, para todo x en el dominio de F, pueden expresarse por una copula n-dimensional (Yan 2007; Madadgar and Moradkhani 2013), como sigue:

$$F(X_1, \dots, X_n) = C\{U_1(X_1), \dots, U_n(X_n)\} \quad (13)$$

donde $U_i(X_i)$ constituye la i^{th} distribución marginal univariada en el intervalo unidad [0,1] y C es la función de distribución copula acumulada que representa la estructura de dependencia multivariada (Joe 1997). Si U_1, \dots, U_n son continuas y C es única (Yan 2006), entonces la Ecuación 11 puede expresarse como:

$$P(X_{tn}|X_{t1}, \dots, X_{tn-1}) = \frac{C(U_{t1}, \dots, U_{tn})}{C(U_{t1}, \dots, U_{tn-1})} \quad (14)$$

Si $n = 2$ (dependencia de primer orden), se calculan las probabilidades condicionales de la variable de pronóstico (X_{t2}) dada la variable predictiva (X_{t1}), es decir, el siguiente estado de sequía está condicionado al estado presente. Del mismo modo, si $n = 3$ (dependencia de segundo orden), se calculan las probabilidades condicionales de la variable de pronóstico (X_{t3}) dado las variables predictoras (X_{t1}, X_{t2}), es decir, el siguiente estado de sequía está condicionado a los estados actual y previo. Estos casos podrían ser llamados red bayesiana de primer orden (BNFO) y de segundo orden (BNSO), respectivamente. Sustituyendo $n = 2$ y $n = 3$ en la Ecuación 14 conduce a las siguientes expresiones:

$$P(X_{t2}|X_{t1}) = \frac{C(U_{t1}, U_{t2})}{U_{t1}} \quad (15)$$

$$P(X_{t3}|X_{t1}, X_{t2}) = \frac{C(U_{t1}, U_{t2}, U_{t3})}{C(U_{t1}, U_{t2})} \quad (16)$$

Para calcular la probabilidad de que el DI en un siguiente paso de tiempo no exceda los umbrales definidos en la sección 3.1 para los estados de sequía (sequía, leve sequía y no sequía), dada la información disponible del estado actual para el modelo BNFO y los estados presente y previo para el modelo BNSO, las ecuaciones 15 y 16 pueden ser redefinidas como:

$$P(X_{t2} \leq xd_s | X_{t1}) = \frac{C[U(X_{t1}), U(X_{t2} \leq xd_s)]}{U(X_{t1})} \quad (17)$$

$$P(X_{t3} \leq xd_s | X_{t1}, X_{t2}) = \frac{C[U(X_{t1}), U(X_{t2}), U(X_{t3} \leq xd_s)]}{C[U(X_{t1}), U(X_{t2})]} \quad (18)$$

donde xd_s es el índice de sequía que provoca los estados de sequía definidos según los umbrales mostrados en la sección 3.1 ($xd_0 = 0$ and $xd_1 = -1$). Aplicando las ecuaciones 17 y 18, las probabilidades de tener una categoría 0 (no sequía), es decir, $DI > 0$, será igual a:

$$P(X_{t2} > xd_0 | X_{t1}) = 1 - \frac{C[U(X_{t1}), U(X_{t2} \leq 0)]}{U(X_{t1})} \quad (19)$$

$$P(X_{t3} > xd_0 | X_{t1}, X_{t2}) = 1 - \frac{C[U(X_{t1}), U(X_{t2}), U(X_{t3} \leq 0)]}{C[U(X_{t1}), U(X_{t2})]} \quad (20)$$

De igual forma, las probabilidades de tener una categoría 1 (leve sequía), es decir, $-1 < DI \leq 0$, será igual a:

$$P(xd_1 < X_{t2} \leq xd_0 | X_{t1}) = \frac{C[U(X_{t1}), U(X_{t2} \leq 0)]}{U(X_{t1})} - \frac{C[U(X_{t1}), U(X_{t2} \leq -1)]}{U(X_{t1})} \quad (21)$$

$$P(xd_1 < X_{t3} \leq xd_0 | X_{t1}, X_{t2}) = \frac{C[U(X_{t1}), U(X_{t2}), U(X_{t3} \leq 0)]}{C[U(X_{t1}), U(X_{t2})]} - \frac{C[U(X_{t1}), U(X_{t2}), U(X_{t3} \leq -1)]}{C[U(X_{t1}), U(X_{t2})]} \quad (22)$$

Por último, las probabilidades de tener una categoría 2 (sequía), es decir, $DI \leq -1$, será igual a:

$$P(X_{t2} \leq xd_1 | X_{t1}) = \frac{C[U(X_{t1}), U(X_{t2} \leq -1)]}{U(X_{t1})} \quad (23)$$

$$P(X_{t3} \leq xd_1 | X_{t1}, X_{t2}) = \frac{C[U(X_{t1}), U(X_{t2}), U(X_{t3} \leq -1)]}{C[U(X_{t1}), U(X_{t2})]} \quad (24)$$

3.2.2.2. Ajuste de copulas

El procedimiento para la construcción de distribuciones conjuntas tiene los siguientes pasos: (1) identificación de distribuciones marginales en el intervalo unidad $[0,1]$; (2) selección de estructuras de dependencia adecuada; y (3) formación de distribuciones conjuntas (Kao and Govindaraju 2010). Las distribuciones marginales (U_i) del paso 1 se derivan cuando los valores de DI se transforman a una función de distribución normal acumulada, de acuerdo con la sección 3.1. Dado que las sequías son fenómenos que evolucionan lentamente, se espera que los DI tengan una fuerte autocorrelación temporal (Kao and Govindaraju 2010). Los pasos 2 y 3 consisten en modelar la estructura de dependencia temporal ajustando las funciones copulas dadas las distribuciones marginales.

Dos copulas elípticas (normal y t) y dos copulas arquimedianas (Clayton y Frank) se analizan en este estudio, en cuanto a definir la función más apropiada para las distribuciones conjuntas. Se aplicó el método máxima verosimilitud canónica (CML) (Cherubini et al. 2004) para la estimación de los parámetros de las funciones copula, usando la función de distribución acumulada empírica de cada distribución marginal (U_i) para transformar las observaciones en pseudo-observaciones en el intervalo unidad $[0,1]$ (Yan 2007). Las mejores copulas ajustadas fueron identificadas por medio del test de bondad de ajuste paramétrico basado en bootstrap (Genest and Rémillard 2008), que consiste en comparar una distancia (ΔC) entre la copula empírica (C_E) y la copula paramétrica estimada (C_θ) bajo la hipótesis nula $C_E \in C_\theta$. Esto último es evaluado por el valor de p; si el valor p es mayor que el nivel de

significancia α se acepta la hipótesis nula, en caso contrario se rechaza la hipótesis nula. Los valores de p se pueden obtener mediante simulación Monte Carlo en un procedimiento de bootstrapping paramétrico.

El estadístico Cramér-von-Mises (S) (Anderson 1962) se aplicó en el conjunto de copulas que cumplieron la hipótesis nula, es decir las copulas que son mayores que el nivel de significancia. Dentro de este grupo de copulas, la función copula con el menor valor de S es elegida como la mejor (Madadgar and Moradkhani 2013). La expresión del estadístico S es como sigue, donde d es el tamaño de la muestra:

$$S = \int \Delta C(U)^2 dC(U) = \sum_{i=1}^d \{C_E(U_i) - C_\theta(U_i)\}^2 \quad (25)$$

3.2.3. VERIFICACIÓN DEL PRONÓSTICO

La calidad de los pronósticos se verificó mediante validación cruzada, utilizando el puntaje de habilidad ponderado de probabilidad (RPSS), el cual es derivado del puntaje ponderado de probabilidad (RPS). Esta medida de habilidad penaliza los errores de pronóstico en términos de la probabilidad asignada a los eventos (Zhang and Casey 2000), y su cálculo se fundamenta en el error cuadrado entre las probabilidades acumuladas de los pronósticos y observaciones. El RPS es sensible a la distancia, es decir, incluye una penalización por los pronósticos que están más lejanos de las observaciones. Según Wilks (2011) el RPS se formula de la siguiente forma:

$$RPS = \sum_{m=1}^s (Y_m - O_m)^2 \quad (26)$$

donde Y_m es la probabilidad acumulada de pronósticos, O_m es la probabilidad acumulada de observaciones y s el número de categorías de eventos o estados del sistema. Un pronóstico perfecto tendría un valor RPS igual a cero y los peores pronósticos serían muy diferentes de cero.

Para la evaluación de la utilidad de los pronósticos podrían utilizarse el RPSS, el cual se refiere a la exactitud relativa de un conjunto de pronósticos con respecto a un cierto sistema de pronósticos de referencia. De hecho el RPSS puede interpretarse como el porcentaje de

mejora sobre el pronóstico de referencia (Wilks 2011) y su formulación se muestra a continuación:

$$RPSS = 1 - \frac{\langle RPS \rangle}{\langle RPS_{Clim} \rangle} \quad (27)$$

donde $\langle RPS \rangle$ es el promedio de los valores RPS para cada registro pronóstico-observación y $\langle RPS_{Clim} \rangle$ es el promedio de los valores RPS calculados con los pronósticos de referencia y las observaciones. En este estudio los pronósticos de referencia fueron las frecuencias relativas climatológicas de la variable de pronóstico. Si el RPSS es igual a 1 significa una mejora del 100% sobre los pronósticos de referencia, mientras que un RPSS igual a 0 o menor indica ninguna mejora sobre los pronósticos de referencia.

3.3. MODELO ESTOCÁSTICO PARA LA GENERACIÓN DE SERIES SINTÉTICAS HIDROLÓGICAS

Remitiéndonos a los aportes de la tercera publicación (Anexo 3) con respecto al modelo numérico para la síntesis de aportaciones de agua, en esta tesis se eligió los modelos autoregresivos periódicos multivariados de primer orden (MPAR1) para generar múltiples series sintéticas hidrológicas, de tal manera que cada serie generada sea igualmente probable que ocurra en el futuro (Sveinsson et al. 2007). Estos modelos tienen la ventaja de representar adecuadamente la correlación temporal (autocorrelación) y espacial (correlación cruzada) de series de tiempo, además de caracterizar diferentes estructuras de dependencia para cada estación del año (Sveinsson et al. 2007; Cancelliere et al. 2009).

Siguiendo de una manera similar la formulación matemática de Sveinsson et al. (2007), los modelos MPAR1 pueden expresarse como:

$$\mathbb{Z}_{v,\tau} = \phi_{1,\tau} \mathbb{Z}_{v,\tau-1} + \epsilon_{v,\tau} \quad (28)$$

donde para las q aportaciones de agua a los embalses en el sistema hídrico, $\mathbb{Z}_{v,\tau}$ es un vector columna $[q \times 1]$ de elementos normalmente distribuidos con media cero para el año v y el mes τ . Los $\phi_{1,\tau}$, son las matrices $[q \times q]$ de parámetros autoregresivos periódicos de orden 1 para cada mes, y $\epsilon_{v,\tau}$ es el vector columna $[q \times 1]$ de los términos de ruido independientes normalmente distribuidos con media cero y matrices $[q \times q]$ de varianza-covarianza \mathbb{G}_τ .

Dado que la mayor parte de curvas de frecuencias de variables hidrológicas son asimétricamente distribuidas, algunas veces es necesario realizar un tratamiento previo a la información de entrada para conseguir series normalizadas, dado que la mayoría de técnicas estadísticas aplicadas en hidrología son desarrolladas asumiendo variables normalmente distribuidas (Salas et al. 1980). En este estudio se aplicó la función potencia para la normalización de las series originales históricas de aportaciones de los dos embalses Chanlud y El Labrado, de la siguiente manera:

$$YT_{v,\tau} = X_{v,\tau}^{1/2} \quad (29)$$

donde $X_{v,\tau}$ es la variable original para el año v de la serie de tiempo del mes τ y $YT_{v,\tau}$ es la variable transformada a la normal para el año v de la serie de tiempo del mes τ . Adicionalmente para conseguir series estacionarias mensuales con media y varianza constante, se procedió a estandarizar la variable transformada aplicando la siguiente fórmula:

$$Z_{v,\tau} = \frac{YT_{v,\tau} - \overline{YT}_{\tau}}{S_{\tau}(YT)} \quad (30)$$

donde $Z_{v,\tau}$ es la variable estandarizada con media cero y desviación estándar uno para el año v de la serie de tiempo del mes τ , \overline{YT}_{τ} y $S_{\tau}(YT)$ son la media y la desviación estándar de las series transformadas de cada mes τ . Una vez que se ajuste el modelo MPAR1 se deberá deshacer las operaciones efectuadas, mediante las operaciones inversas.

3.4. INTRODUCCIÓN DE LOS PRONÓSTICOS DE SEQUÍA EN LA GENERACIÓN DE SERIES SINTÉTICAS HIDROLÓGICAS

Una vez ajustado el modelo MPAR1 (estimación de parámetros) se procede a la generación de series sintéticas (en el mes τ) condicionadas a los valores previos ($\tau-1$). Para este proceso se utiliza la Ecuación 28 y se genera números aleatorios para $\mathcal{E}_{v,\tau}$ mediante una distribución normal multivariada con media cero y matrices de varianza-covarianza \mathbb{G}_{τ} . Para la generación de números aleatorios $\mathcal{E}_{v,\tau}$ se introduce los resultados de los pronósticos de los eventos de sequía en el mes τ dada la información del mes $\tau-1$. Esto es llevado a cabo identificando las probabilidades de ocurrencia de cada estado de sequía (no sequía, leve sequía y sequía), las cuales representan el porcentaje de la cantidad total de números aleatorios por generar; y mediante una distribución normal multivariada truncada en tres

intervalos (según cada estado de sequía) generar el conjunto de números aleatorios $\mathcal{E}_{v,\tau}$ en el mes τ . Los tres intervalos se dividen de la siguiente forma: 1) Desde el valor máximo de $Z_{v,\tau}$ de cada serie mensual hasta el valor de $Z_{v,\tau} = 0$, 2) Desde el valor de $Z_{v,\tau} = 0$ hasta el valor de $Z_{v,\tau} = -1$ y 3) Desde el valor de $Z_{v,\tau} = -1$ hasta el valor mínimo de $Z_{v,\tau}$ de cada serie mensual. Estos intervalos son análogos a los rangos de los estados no sequía, leve sequía y sequía, respectivamente, en la escala del DI (Sección 3.1). Una vez generados los números aleatorios se procedió a generar las múltiples series sintéticas con escenarios de escasez utilizando la Ecuación 28, para lo cual se asumieron tres opciones para los valores previos: 1) Valor del $Z_{v,\tau-1} = 0$, igual al valor promedio de cada serie mensual, 2) Valor del $Z_{v,\tau-1} = -1$, análogo al límite inferior del estado de leve sequía en la escala del DI (Sección 3.1) y 3) Valor mínimo de $Z_{v,\tau-1}$ de cada serie mensual. Con este proceso se obtendría una predicción de la función de distribución de los posibles valores de la variable en el mes τ condicionada a los valores del mes $\tau-1$. Con lo cual se atendería la recomendación de la tercera publicación (Anexo 3) con respecto a la necesidad de realizar una gestión a corto plazo. Las múltiples series de tiempo generadas serán información de entrada para el modelo de simulación.

3.5. MODELO DE SIMULACIÓN MÚLTIPLE PARA LA EVALUACIÓN DEL RIESGO DE FALLO

Remitiéndonos a la recomendación de la tercera publicación (Anexo 3) expuesta en el párrafo anterior, el periodo de simulación será de 12 meses (corto plazo), para poder operar y gestionar el sistema dentro de un año. Este periodo aprovecharía la estacionalidad de las variables hidrológicas para almacenar agua en tiempos de superávit y suministrar las reservas en tiempos de escasez. Además según lo que manifiestan Rossi et al. (2012), en periodos mayores a un año es más difícil formular medidas preventivas para ahorrar agua en una crisis de desabastecimiento. La simulación se lo realiza con diferentes condiciones previas hidrológicas, varias situaciones iniciales de los niveles de almacenamiento para cada mes, por lo tanto, la gestión u operación del sistema en el periodo de simulación están condicionados a los contextos previos e iniciales. Los datos de entrada para el modelo de evaluación del riesgo de fallo (SIMRISK) son el mes de inicio de la simulación (12 opciones de enero hasta diciembre), el volumen inicial de los embalses, que para este estudio se dividieron en 4 niveles para cada embalse (16 combinaciones), es decir, para Chanlud (4,8,12

y 16 hm³) y para El Labrado (1.5,3,4.5, y 6 hm³), las condiciones hidrológicas previas para cada embalse (3 valores previos para cada embalse según lo explicado en la sección 3.4, es decir 9 combinaciones), por lo tanto, el proceso de simulación se realizó con 1728 escenarios (12x16x9). Es así que el proceso de simulación considera la sugerencia de la cuarta publicación (Anexo4) con respecto a disminuir la subjetividad en la toma de decisiones cuando se tiene un marco amplio de información. Cada escenario consiste en 10000 series sintéticas hidrológicas generadas en cada uno de los embalses (en donde está incluida la información de los pronósticos de sequías según lo explicado anteriormente), cada una con una longitud igual a 12 meses (periodo de simulación).

Durante cada mes en el periodo de simulación para cada uno de los escenarios, las demandas podrán recibir un suministro mayor o igual al valor demandado (estado de satisfacción), o un valor menor (estado de insatisfacción). En este último caso se habrá producido un fallo en el suministro con un déficit (D) igual al valor de la demanda menos el valor de la cantidad suministrada. El nivel de severidad del déficit D dependerá de los niveles de suministro con respecto al valor de la demanda. La tolerancia al riesgo de fallo de las diferentes demandas puede llegar a convertirse en una tarea subjetiva. Sin embargo, como un apoyo a la objetividad esta información se puede representar en un solo índice de satisfacción de las demandas (DSI), tomando en consideración la sugerencia de la cuarta publicación (Anexo 4). Este índice es resultado del número de fallos en el suministro de las demandas mediante un índice de confiabilidad (RI) y la severidad de estos fallos mediante un índice de severidad (SI). Siguiendo de una forma similar como lo hicieron Hashimoto et al. (1982) y Sandoval-Solis et al. (2011), estos índices para una demanda particular y para cada mes en el periodo de simulación se pueden calcular de la siguiente manera:

$$RI = \frac{(\# \text{ total de simulaciones} - \# \text{ de fallos})}{\# \text{ total de simulaciones}} \quad (31)$$

$$SI = \frac{\sum_{j=1}^n (D_j)}{\# \text{ total de simulaciones} * \text{Valor de la demanda}} \quad (32)$$

$$DSI = RI * (1 - SI) \quad (33)$$

donde n es igual al número de niveles de suministro y D_j es la cantidad de déficit en cada nivel de suministro. Si existen varias demandas el DSI puede calcularse como un índice de satisfacción de un grupo de demandas DSI_G mediante una suma ponderada de los DSI particulares de la siguiente manera:

$$DSI_G = \left(\sum_{i=1}^k \frac{\text{Valor de la demanda } i}{\sum_{i=1}^k \text{Valor de la demanda } i} * DSI_i \right) * 100 \quad (34)$$

donde i es el contador de las demandas particulares y k es el número total de demandas del sistema de recursos hídricos.

El DSI_G podría servir para la toma de decisiones cada mes del año en un contexto operacional, además la información de entrada del modelo simulación debería ir actualizándose cada cierto número de años, dadas las variaciones en las condiciones hidroclimáticas y los cambios en los valores de las demandas. El valor del índice DSI_G puede variar desde 0% hasta el 100%, mientras más alto sea el valor del índice DSI_G significa una mayor satisfacción del sistema.

4. RESULTADOS

4.1. CÁLCULO DEL DI

Aplicando PCA para cada mes del año y para cada microcuenca (ríos Machángara Alto y Chulco), utilizando la información de las matrices de correlación de las diez series de tiempo (PR1, PR3, PR6, PR9, PR12, VS1, VS3, VS6, VS9 y VS12) en cada microcuenca, se obtuvo los autovalores y los autovectores. La tabla 1 muestra los autovalores y la varianza explicada del PC1 para cada mes y cada microcuenca

Tabla 1. Autovalores y varianza explicada del PC1 de las microcuencas de los ríos (a) Machángara Alto y (b) Chulco

PC1	ene	feb	mar	abr	may	jun	jul	ago	sep	oct	nov	dic
Autovalores (a)	6.08	6.67	7.64	7.03	7.59	7.08	7.15	7.01	7.00	6.59	5.78	5.56
Varianza explicada (a)	61%	67%	76%	70%	76%	71%	72%	70%	70%	66%	58%	56%
Autovalores (b)	6.21	7.26	7.89	7.13	7.64	7.08	7.10	7.16	7.08	6.77	5.86	5.71
Varianza explicada (b)	62%	73%	79%	71%	76%	71%	71%	72%	71%	68%	59%	57%

Entonces usando las ecuaciones 1 y 2 obtenemos los doce conjuntos de valores del DI. Estos valores se reordenan cronológicamente para obtener las series de tiempo de los valores del DI para cada microcuenca (1971-2010). La Figura 6 muestra los valores del DI y los umbrales de severidad de sequía, donde se puede observar la frecuencia y duración de cada evento de sequía (no sequía, leve sequía y sequía). Se puede observar en la figura que el periodo con mayor frecuencia de sequías severas fue desde 1983 hasta 1992 en ambas microcuencas, con una mayor persistencia de severas sequías desde 1989 hasta 1992 en la microcuenca del río Machángara Alto. Este periodo quizás pudo haber tenido influencia de fenómenos climáticos como El Niño de los años 1983, 1987 y 1992 (Vicente-Serrano et al. 2016). No obstante existen otros periodos con leves sequías como 1978-1979, 1994-1995 y 2000-2004, los cuales pueden deberse a varios factores climáticos y orográficos de la cuenca. Los mecanismos que producen sequías en cuencas andinas se manifiestan como futuros retos investigativos.

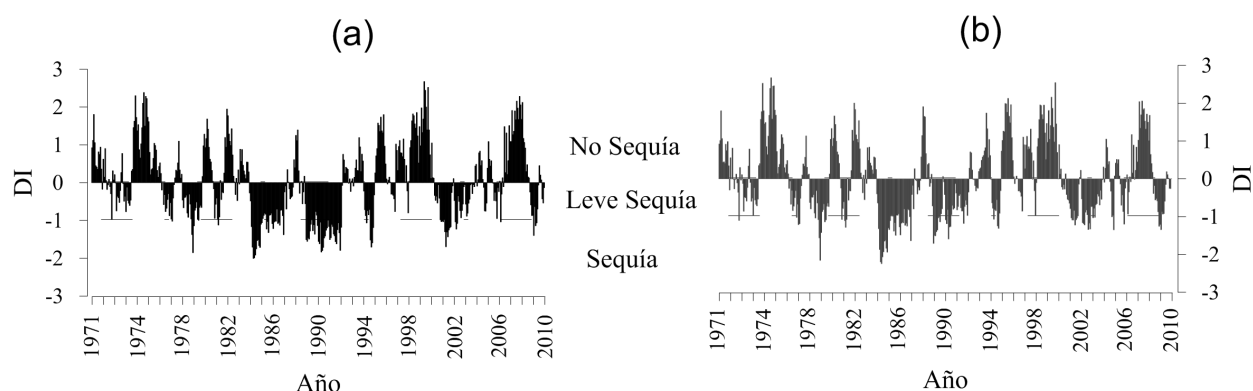


Figura 6. Series de tiempo del DI (1971 – 2010) de las microcuencas de los ríos (a) Machángara Alto y (b) Chulco

4.2. PRONÓSTICOS PROBABILÍSTICOS DE SEQUÍA

El análisis comparativo del desempeño de los modelos basados en MC y los modelos basados en BN para el pronóstico probabilístico de sequías se realizó en la microcuenca del río Chulco. Los valores del RPS_{CLIM} fueron calculados a partir de los pronósticos de referencia, que para este estudio fueron las frecuencias relativas climatológicas de la serie de tiempo categórica de los valores del DI. A partir de los valores del RPS y RPS_{CLIM} se calcularon los valores del RPSS mediante la Ecuación 27.

Considerando los valores del RPSS y tomando en cuenta todos los estados de sequía (no sequía, leve sequía y sequía) en todos los meses, el modelo MCFO fue el de mejor desempeño comparado con los otros modelos con el mas grande valor de mejora sobre los pronósticos de referencia igual a 0.29, seguido por los modelos MCSO y BNFO con valores del RPSS igual a 0.21 y -0.05 (valores negativos del RPSS significan que los pronósticos de referencia son mejores que los modelos evaluados). Tomando en cuenta los estados de leve sequía y sequía, el modelo BNFO se desempeña de mejor forma con el valor más grande del RPSS igual a 0.40, seguido por los modelos BNSO y MCFO con valores del RPSS igual a 0.33 y 0.19, respectivamente. Y solo considerando el estado de sequía más severo, el modelo BNFO produjo mejores resultados, con el valor más grande del RPSS igual a 0.44, seguido de los modelos MCSO y BNSO con valores iguales a 0.37 y 0.35. Estos resultados indican que en este caso de estudio el modelo MCFO produce mejores pronósticos en tiempos húmedos y secos, mientras tanto que en periodos secos los modelos BN puede ser levemente superiores.

La Tabla 2 presenta la verificación mensual de los pronósticos de sequía, con los valores del RPSS de todos los estados de sequía en la sección (a) y de los estados de leve sequía y sequía en la sección (b). Si analizamos la información de la sección (a) se puede concluir que el modelo MCFO tienen mejor desempeño en los meses de enero, febrero, marzo, mayo, junio, septiembre, octubre y noviembre, mientras que el modelo MCSO se desempeña de mejor forma en los meses de abril, julio, agosto y diciembre. En cambio si analizamos la información de la sección (b) se puede indicar que el modelo BNFO tiene mejor desempeño en los meses de enero, febrero, abril, mayo, junio, octubre, noviembre y diciembre, el modelo MCFO produjo mejores resultados en los meses de marzo y septiembre y el modelo MCSO se desempeña de mejor forma en los meses de julio y agosto.

Tabla 2. Valores del RPSS de la verificación mensual de los pronósticos de sequía

	ene	feb	mar	abr	may	jun	jul	ago	sep	oct	nov	dic
(a) No sequía, leve sequía y sequía												
MCFO	-0.40	0.27	0.63	0.52	0.37	0.41	0.48	0.45	0.33	0.20	0.15	0.15
MCSO	-0.46	-0.32	0.63	0.58	0.25	0.32	0.50	0.52	0.18	0.19	0.05	0.16
BNFO	-0.50	0.13	0.09	0.06	-0.20	0.08	0.04	0.00	-0.11	-0.15	0.02	-0.09
BNSO	-0.60	-0.33	-0.11	-0.08	-0.29	-0.08	-0.14	-0.13	-0.30	-0.33	-0.09	-0.27
(b) Leve sequía y sequía												
MCFO	-0.35	0.12	0.53	0.49	0.27	0.19	0.42	0.31	0.04	0.12	0.09	0.18
MCSO	-0.28	-0.28	0.53	0.53	0.15	0.07	0.48	0.47	-0.14	0.10	-0.03	0.16

BNFO	0.29	0.53	0.48	0.59	0.29	0.38	0.47	0.47	-0.07	0.32	0.45	0.62
BNSO	0.28	0.50	0.37	0.53	0.23	0.20	0.38	0.39	-0.27	0.18	0.42	0.53

La evaluación de los pronósticos tomando en cuenta todos los estados de sequía y todos los meses revelan que el modelo MCFO tienen el mejor rendimiento. En cambio si solo se toma en cuenta los pronósticos de los periodos secos los modelos basados en BN son ligeramente superiores. Por otro lado la evaluación mensual de los pronósticos tomando en cuenta todos los estados de sequías muestra que el modelo MCFO tiene mejor desempeño en ocho meses y el modelo MCSO es mejor en cuatro meses. Mientras si solo se consideran los periodos secos el modelo BNFO se desempeña de mejor forma en ocho meses y los modelos MCFO y MCSO tuvieron mejor rendimiento en dos meses cada uno. Lo expuesto coincide con lo expresado por Madadgar and Moradkhani (2013), en donde se declara que los pronósticos de modelos basados en BN no difieren significativamente de los pronósticos generados por otros modelos. Además estos autores (Madadgar and Moradkhani 2014) también expusieron que los modelos basados en BN en combinación con las funciones copulas prometen ser herramientas útiles para el pronóstico probabilístico de eventos de sequía.

Por lo tanto haciendo un análisis global de la comparación de los dos enfoques (MC y BN) para pronósticos de sequías, los modelos basados en MC parecieran ser una buena opción para este propósito dado que los pronósticos generados por estos modelos son mejores en periodos húmedos y secos, y también porque en periodos secos los pronósticos de estos modelos no presentan mayores diferencias que los pronósticos generados por los modelos basados en BN. Además hay que resaltar el comportamiento del modelo MCFO dado que en la mayoría de análisis realizados tiene el mejor rendimiento comparado con el resto de modelos. Por esta razón este modelo se ha escogido para realizar los pronósticos probabilísticos de sequía en el caso de estudio de esta tesis. Los resultados de este proceso se describe a continuación.

La información de la serie de tiempo categórica del DI se utilizó para el ajuste de los modelos MCFO, en donde los valores categóricos se encuentran distribuidos en tres estados 0, 1 y 2, correspondientes respectivamente a los estados de no sequía, leve sequía y sequía. Tomando en cuenta la estacionalidad mensual y utilizando la Ecuación 5, se construyeron 12 matrices de probabilidad de transición para cada microcuenca. Estas matrices permiten obtener el pronóstico probabilísticos del mes actual j dado la categoría del estado del mes anterior i .

Las Tablas 3 y 4 muestran los pronósticos probabilísticos de sequía de las microcuencas de los ríos Machángara Alto y Chulco.

Tabla 3. Pronósticos probabilísticos de sequía de la microcuenca del río Machángara Alto

Categoría mes anterior i	Categoría mes actual j	Pronóstico probabilístico del mes actual j											
		ene	feb	mar	abr	may	jun	jul	ago	sep	oct	nov	dic
0	0	0.75	0.80	0.94	0.95	0.86	0.80	1.00	0.86	0.84	0.82	0.78	0.89
	1	0.25	0.15	0.06	0.05	0.14	0.20	0.00	0.14	0.16	0.18	0.22	0.11
	2	0.00	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0	0.45	0.18	0.21	0.18	0.20	0.13	0.19	0.08	0.08	0.24	0.22	0.24
	1	0.45	0.55	0.58	0.73	0.80	0.74	0.75	0.84	0.92	0.76	0.67	0.47
	2	0.10	0.27	0.21	0.09	0.00	0.13	0.06	0.08	0.00	0.00	0.11	0.29
2	0	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
	1	0.11	0.56	0.25	0.11	0.44	0.20	0.00	0.00	0.25	0.33	0.25	0.20
	2	0.89	0.44	0.75	0.89	0.56	0.80	1.00	1.00	0.75	0.67	0.75	0.80

Tabla 4. Pronósticos probabilísticos de sequía de la microcuenca del río Chulco

Categoría mes anterior i	Categoría mes actual j	Pronóstico probabilístico del mes actual j											
		ene	feb	mar	abr	may	jun	jul	ago	sep	oct	nov	dic
0	0	0.75	0.79	0.95	0.89	0.85	0.79	0.94	0.89	0.72	0.86	0.78	0.78
	1	0.25	0.21	0.05	0.11	0.15	0.21	0.06	0.11	0.28	0.14	0.22	0.22
	2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
1	0	0.29	0.31	0.08	0.21	0.08	0.06	0.17	0.13	0.07	0.27	0.18	0.33
	1	0.50	0.38	0.84	0.65	0.84	0.76	0.72	0.68	0.86	0.59	0.53	0.54
	2	0.21	0.31	0.08	0.14	0.08	0.18	0.11	0.19	0.07	0.14	0.29	0.13
2	0	0.00	0.00	0.00	0.00	0.12	0.00	0.00	0.00	0.00	0.00	0.20	0.14
	1	0.17	0.38	0.33	0.14	0.50	0.25	0.33	0.33	0.57	0.50	0.40	0.29
	2	0.83	0.62	0.67	0.86	0.38	0.75	0.67	0.67	0.43	0.50	0.40	0.57

Por ejemplo si se desea pronosticar los eventos de sequías para el mes de agosto (mes con menor cantidad de lluvia) para la microcuenca del río Machángara Alto se tendría que observar la Tabla 3. Si en el mes de julio el estado de sequía corresponde a la categoría 0 (no sequía), entonces los pronósticos probabilísticos para el mes de agosto son: 86% para la categoría 0 (no sequía), 14% para la categoría 1 (leve sequía) y 0% para la categoría 2 (sequía). En cambio sí en el mes julio el estado de sequía corresponde a la categoría 1 (leve sequía), entonces los pronósticos para el mes de agosto son: 8% para la categoría 0 (no

sequía), 84% para la categoría 1 (leve sequía) y 8% para la categoría 2 (sequía). Finalmente si en el mes julio el estado de sequía corresponde a la categoría 2 (sequía), entonces los pronósticos para el mes de agosto son: 0% para la categoría 0 (no sequía), 0% para la categoría 1 (leve sequía) y 100% para la categoría 2 (sequía). Similarmente, para obtener los pronósticos probabilísticos del mes de agosto para la microcuenca del río Chulco se tendría que usar la información de la Tabla 4. Si en el mes de julio el estado de sequía corresponde a la categoría 0 (no sequía), entonces los pronósticos probabilísticos para el mes de agosto son: 89% para la categoría 0 (no sequía), 11% para la categoría 1 (leve sequía) y 0% para la categoría 2 (sequía). En cambio sí en el mes julio el estado de sequía corresponde a la categoría 1 (leve sequía), entonces los pronósticos para el mes de agosto son: 13% para la categoría 0 (no sequía), 68% para la categoría 1 (leve sequía) y 19% para la categoría 2 (sequía). Finalmente si en el mes julio el estado de sequía corresponde a la categoría 2 (sequía), entonces los pronósticos para el mes de agosto son: 0% para la categoría 0 (no sequía), 33% para la categoría 1 (leve sequía) y 67% para la categoría 2 (sequía). Hay que resaltar que en todos los casos para cada mes la suma de los pronósticos probabilísticos de cada categoría de sequía es igual a 100%.

4.3. EVALUACIÓN DEL RIESGO DE FALLO

El modelo MPAR1 resultó suficiente para preservar algunos estadísticos de las series históricas de aportaciones normalizadas (utilizando la Ecuación 29) y estandarizadas (utilizando la Ecuación 30) de los dos embalses de Chanlud y El Labrado. Los parámetros estimados y la varianza-covarianza del ruido para cada mes se muestran en la Tabla 5. Mientras que la Tabla 6 muestra los umbrales (promedio, mínimo, máximo, etc.) de las series históricas de aportaciones normalizadas y estandarizadas de los dos embalses Chanlud y El Labrado. Esta información nos indica las tres opciones de valores previos ($Z_{v,\tau-1}$) para cada mes y los umbrales de cada intervalo del truncamiento para la generación de números aleatorios (\mathcal{E}), utilizando los pronósticos probabilísticos de sequía presentados en las Tablas 3 y 4 (según lo expuesto en la sección 3.4). Por lo tanto, mediante una distribución normal multivariada con media cero y matrices de varianza-covarianza (\mathbb{G}) se generaron 10000 números aleatorios (\mathcal{E}), que sirvieron para la generación de 10000 series sintéticas hidrológicas con una longitud igual a 12 meses cada una (periodo de simulación), utilizando la Ecuación 28 y la información de las Tablas 3,4,5 y 6. Esta cantidad considerable de series

generadas podría capturar toda o gran parte de la variabilidad de los flujos de entrada de agua a los embalses.

Tabla 5. Matrices de parámetros ϕ y matrices de varianza-covarianza \mathbb{G} del modelo MPAR1 de las dos series de aportaciones de agua normalizadas y estandarizadas de los dos embalses Chanlud y El Labrado

MESES	MATRICES ϕ		MATRICES \mathbb{G}	
ENERO	-1.61	2.22	0.59	0.63
	-2.05	2.59	0.63	0.66
FEBRERO	-1.84	2.48	0.54	0.56
	-2.14	2.73	0.56	0.58
MARZO	0.48	0.35	0.30	0.32
	0.28	0.53	0.32	0.33
ABRIL	-0.14	1.01	0.24	0.26
	-0.50	1.34	0.26	0.28
MAYO	-4.64	5.28	0.52	0.54
	-5.27	5.86	0.54	0.58
JUNIO	-0.72	1.54	0.32	0.35
	-1.19	1.96	0.35	0.37
JULIO	-3.18	3.91	0.40	0.43
	-3.82	4.49	0.43	0.46
AGOSTO	-0.52	1.36	0.28	0.30
	-0.93	1.74	0.30	0.32
SEPTIEMBRE	2.33	-1.66	0.52	0.55
	2.15	-1.52	0.55	0.58
OCTUBRE	-1.74	2.36	0.59	0.61
	-2.07	2.65	0.61	0.63
NOVIEMBRE	-4.59	5.07	0.68	0.69
	-5.01	5.44	0.69	0.71
DICIEMBRE	-5.71	6.33	0.47	0.49
	-6.19	6.76	0.49	0.51

Tabla 6. Umbrales de las dos series históricas de aportaciones de agua normalizadas y estandarizadas de los dos embalses Chanlud y El Labrado

UMBRAL	ene	feb	mar	abr	may	jun	Jul	ago	sep	oct	nov	dic
EMBALSE DE CHANLUD												
PROMEDIO	0	0	0	0	0	0	0	0	0	0	0	0
NIVEL -1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1
MINIMO	-1.60	-1.71	-2.40	-3.01	-2.55	-1.77	-1.68	-1.69	-1.76	-1.68	-1.68	-1.94
MAXIMO	2.28	1.72	1.68	1.88	2.25	2.15	2.79	3.24	2.23	2.35	1.97	1.69
EMBALSE DE EL LABRADO												
PROMEDIO	0	0	0	0	0	0	0	0	0	0	0	0
NIVEL -1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1	-1

UMBRAL	ene	feb	mar	abr	may	jun	Jul	ago	sep	oct	nov	dic
MINIMO	-1.47	-1.78	-2.45	-3.00	-2.37	-1.60	-1.65	-1.66	-1.84	-1.66	-1.74	-2.09
MAXIMO	2.34	1.71	1.71	1.91	2.22	2.14	2.84	3.23	2.33	2.28	1.97	1.61

Por ejemplo para la generación de números aleatorios para el mes de agosto (mes de inicio de la simulación) y con un estado de sequía (categoría 2) para las dos microcuencas en el mes de julio (mes anterior), se utilizaría una distribución normal multivariada truncada con media cero y matriz de varianza-covarianza \mathbb{G} para el mes de agosto (presentada en la Tabla 5). Además se utilizaría la información de las Tablas 3 y 4 para obtener los pronósticos probabilísticos de sequía del mes de agosto dada la información del mes de julio (mes anterior) que servirían para identificar la cantidad de números aleatorios para cada intervalo de generación según los umbrales mostrados en la Tabla 6. Por lo tanto, la cantidad de números aleatorios para agosto en el embalse de Chanlud sería: 0 en el intervalo $[0, 3.24]$, 0 en el intervalo $[-1, 0]$ y 10000 en el intervalo $[-1.69, -1]$, y en el embalse de El Labrado sería: 0 en el intervalo $[0, 3.23]$, 3300 en el intervalo $[-1, 0]$ y 6700 en el intervalo $[-1.66, -1]$, sumando darían 10000 números aleatorios en ambos embalses para el mes de agosto. A partir de estos dos conjuntos de números aleatorios se generarían las series sintéticas hidrológicas para el mes de agosto, mediante la utilización de la Ecuación 28, los parámetros del modelo MPAR1 para el mes de agosto expuestos en la Tabla 5 y las condiciones previas hidrológicas del mes anterior (es decir del mes de julio) para los dos embalses expuestos en la Tabla 6, las cuales podríamos asumir que son las mismas condiciones de los estados de sequía para las dos microcuencas, es decir, una categoría 2 (sequía). Por lo tanto, las condiciones hidrológicas previas para el mes de agosto serían los valores mínimos del mes de julio en la Tabla 6, que para el embalse de Chanlud tendría un valor de $Z_{v,\tau-1} = -1.68$ y para el embalse de El Labrado tendría un valor de $Z_{v,\tau-1} = -1.65$.

Entonces, modificando el mes de inicio de la simulación, las condiciones hidrológicas previas en los dos embalses, los estados de sequía previos en las dos microcuencas y los volúmenes iniciales de los embalses se realizó la simulación de 1728 escenarios cada uno con 10000 series sintéticas hidrológicas de 12 meses cada una. Estos escenarios sirvieron de entrada para el modelo SIMRISK y los resultados que se obtienen son las probabilidades de fallo del sistema en 4 niveles de suministro y para cada mes del periodo de simulación. El nivel 1 (n1) es la situación más grave, quiere decir que el déficit supera el 75% de la demanda o sea el suministro se ubica entre el 0 y 25% del valor de la demanda; el nivel 2 (n2) quiere decir que

el suministro se ubica entre el 25 y 50% del valor de la demanda; el nivel 3 (n3) quiere decir que el suministro se ubica entre el 50 y 75% del valor de la demanda; y el nivel 4 (n4) es la situación menos grave, quiere decir que el suministro se ubica entre el 75 y 100%.

Para explicar los resultados, se toma como ejemplo un escenario con características similares a la sequía del año 1985: agosto como mes de inicio de la simulación, valores mínimos del mes de julio como las condiciones hidrológicas de entrada para cada embalse, una categoría 2 (sequía) en el mes de julio como el estado de sequía para ambas microcuencas, un volumen inicial para Chanlud igual a 4 hm³ y para El Labrado 1.5 hm³. La Figura 7 muestra los resultados de la simulación de este escenario con estas condiciones particulares. Se puede observar en esta figura que existe una probabilidad de fallo considerable de las demandas de riego en el mes de septiembre (probabilidad del n1 igual al 60% y probabilidad total igual a 80% aproximadamente), de igual forma en este mes la demanda urbana tiene una moderada probabilidad de fallo (probabilidad total igual a 34% aproximadamente). En el mes de octubre la probabilidad de las demandas de riego baja levemente (probabilidad total igual a 60% aproximadamente), y existe una probabilidad de fallo nula para la demanda urbana. En los meses de noviembre y diciembre las demandas de riego tienen bajas probabilidades de fallo (probabilidad total menor al 10%) y sigue manteniéndose una probabilidad de fallo nula para la demanda urbana. Esta información podría considerarse evidencia suficiente para que el gestor del sistema identifique medidas severas de prevención y/o mitigación para disminuir el riesgo de fallo de las demandas en los meses de septiembre y octubre y otras medidas menos severas para los meses de noviembre y diciembre. Sujetándonos a los aportes de la cuarta publicación (Anexo 4), la tolerancia al riesgo de fallo dependerá de la subjetividad de los tomadores de decisiones, no obstante el análisis condensado de los resultados de una cantidad considerable de escenarios con sus medidas respectivas podría apoyar para que la decisión sea más objetiva.

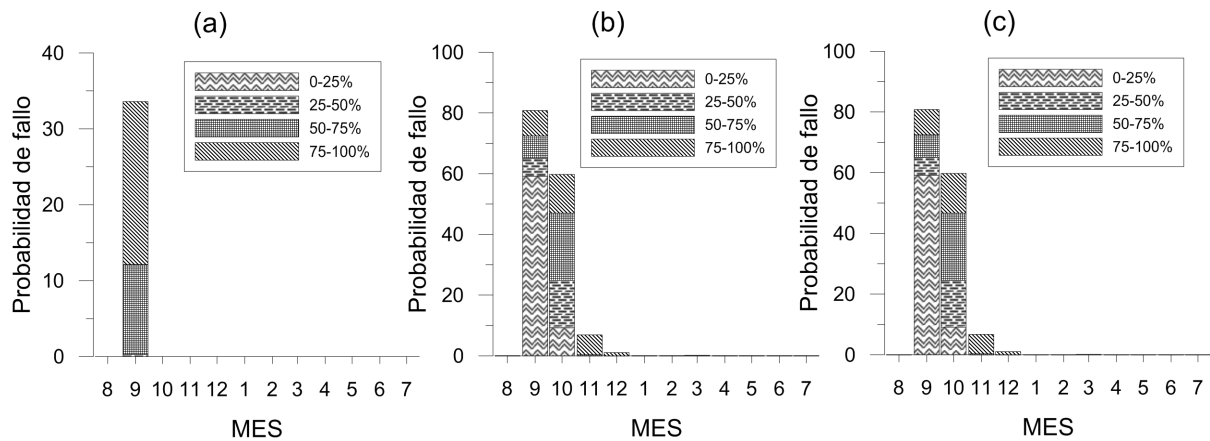


Figura 7. Probabilidad de fallo de las demandas de agua para: (a) ciudad de Cuenca, (b) canal de riego Machángara y (c) canal de riego La Dolorosa.

Quando se manejan múltiples escenarios como el caso de la presente tesis, la utilización del DSI_G podría ser de gran ayuda para concentrar los resultados de la simulación de un determinado escenario. Por esta razón mediante la utilización de las Ecuaciones 31, 32, 33 y 34 y los resultados del modelo SIMRISK se calculó el DSI_G para cada escenario y para cada mes del periodo de simulación. La Figura 8 muestra el DSI_G de la simulación del escenario descrito en el párrafo anterior. En esta figura podemos observar que el DSI_G es igual a 30% en el mes de septiembre y 60% en el mes de octubre y en todos los demás meses es mayor al 90%. Por lo tanto, la información de este gráfico evidencia de una forma más comprensible e integral que para los meses de septiembre y octubre se necesitarán formular algunas medidas de prevención y/o mitigación para operar y gestionar el sistema de tal forma que disminuya el riesgo de fallo.

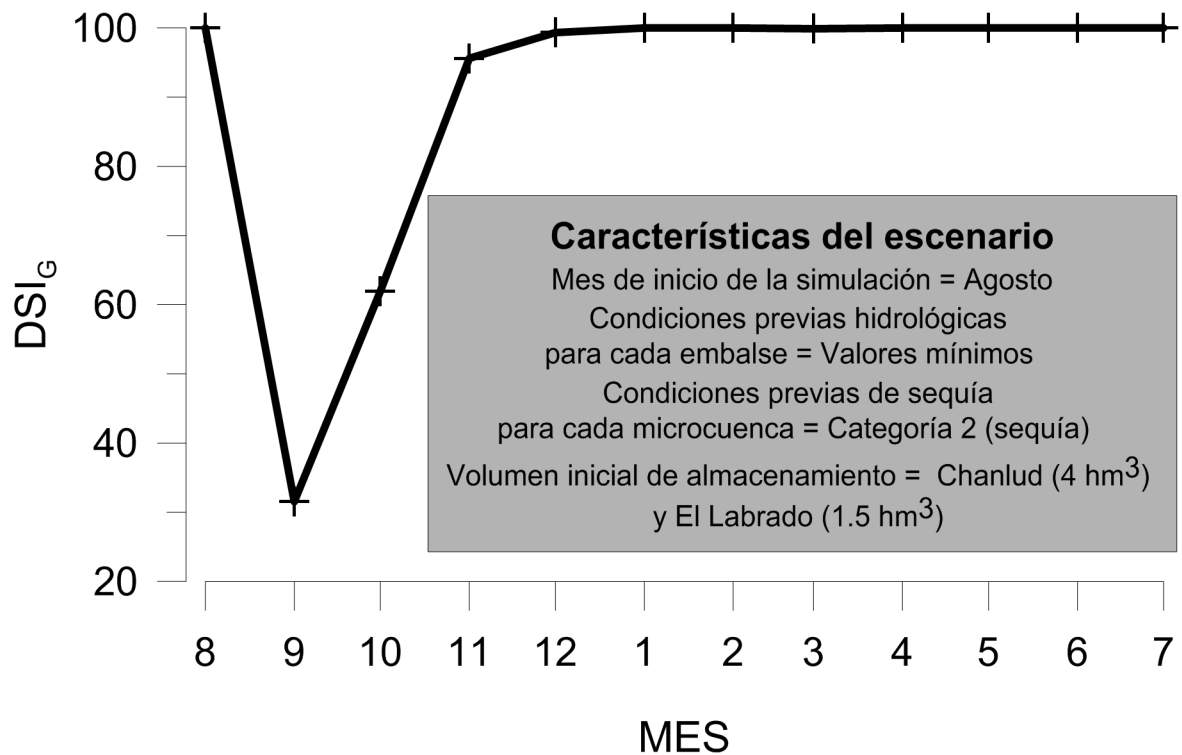


Figura 8. DSI_G del sistema de recursos hídricos de la cuenca del río Machángara para un escenario

No obstante manejar 1728 gráficos que son resultado de la simulación de todos los escenarios, sería poco útil, por lo tanto se podría agrupar los resultados de cada escenario por meses, por las condiciones previas hidrológicas de los flujos de entrada a los embalses o por las condiciones iniciales de almacenamiento para cada embalse. La Figura 9 muestra los DSI_G del conjunto de escenarios agrupados por las condiciones iniciales de almacenamiento para cada embalse, con las siguientes características comunes: agosto como mes de inicio de la simulación, valores mínimos del mes de julio como las condiciones hidrológicas de entrada para cada embalse, una categoría 2 (sequía) en el mes de julio como el estado de sequía para ambas microcuencas. Se puede observar en este gráfico que las condiciones iniciales de los embalses no es un aspecto muy sensible que modifique de manera considerable los DSI_G. Es evidente que si se tiene mayores reservas al inicio de la simulación, el índice DSI_G aumentará en los siguientes meses, todo lo contrario sucedería si las reservas son insuficientes.

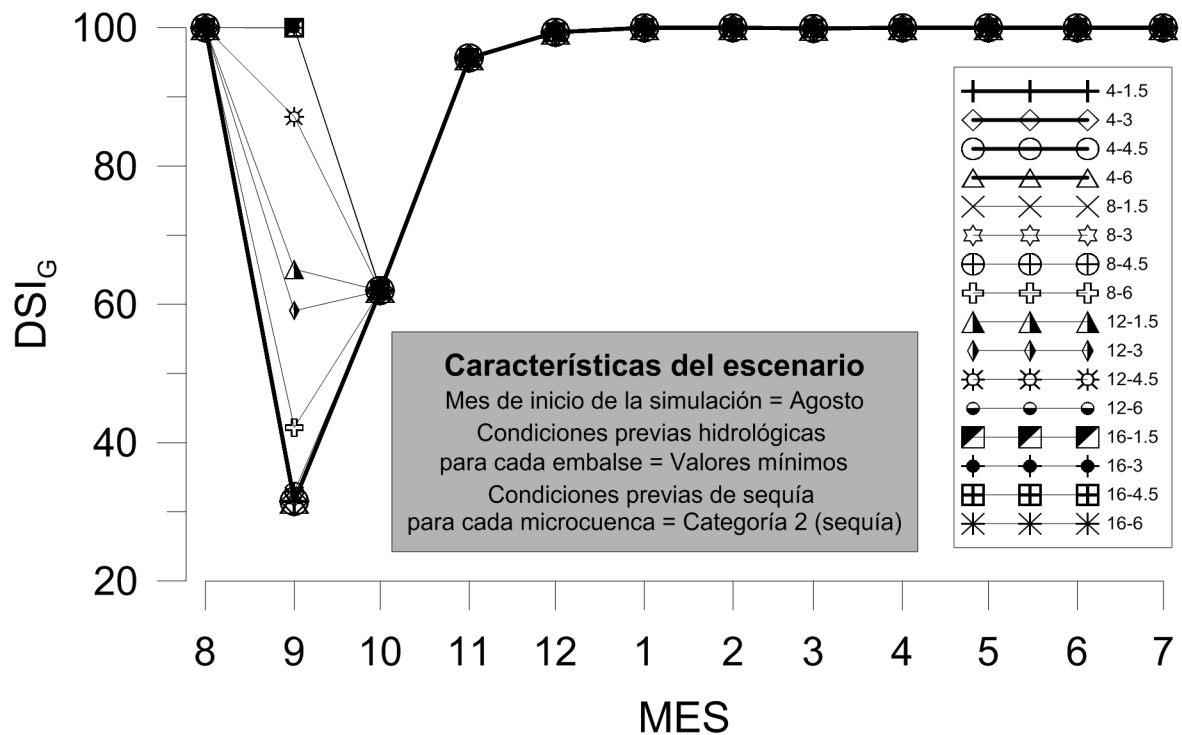


Figura 9. DSI_G del sistema de recursos hídricos de la cuenca del río Machángara para un escenario con diferentes volúmenes iniciales de almacenamiento de los embalses de Chanlud y El Labrado aplicando el marco metodológico integral de la Figura 5

Con esta agrupación se obtuvo 108 gráficos, los mismos que podrían agruparse nuevamente con otros criterios u ordenarlos de una forma sistemática para el análisis de los gestores del agua para que puedan tomar decisiones a corto plazo cuando los valores del DSI_G sean bajos. Por ejemplo se podría tener preparado un conjunto de alternativas de gestión y operación para hacer frente a cada una de las situaciones de escasez de disponibilidad de agua y poder reducir el riesgo de fallo.

Con la finalidad de mostrar las ventajas del marco metodológico integral expuesto en la Figura 5 se realizó la evaluación del riesgo de fallo siguiendo la metodología de la Figura 4 y con el mismo conjunto de escenarios de la Figura 9. Realizando este proceso se obtuvo los DSI_G agrupados por las condiciones iniciales de almacenamiento para cada embalse (Figura 10), en donde se puede observar que los valores del DSI_G aumentaron su valor comparados con los valores del DSI_G de la Figura 9. Por ejemplo, el valor mas bajo del DSI_G en la Figura 9 se presenta en el mes de septiembre y es aproximadamente igual a 30%, este valor al ser muy bajo puede considerarse suficiente para proponer alternativas de gestión en la oferta y demanda de agua para poder incrementar la satisfacción del sistema. Mientras que en la Figura 10 el valor mas bajo del DSI_G es aproximadamente igual a 75% en el mes de

septiembre, el mismo que podría considerarse admisible para la satisfacción del sistema. Por lo tanto, una incorporación de los pronósticos probabilísticos de sequía podrían direccionar de una mejor forma las proyecciones de los escenarios de simulación y contribuir a obtener resultados más efectivos para tomar decisiones en épocas de sequía.

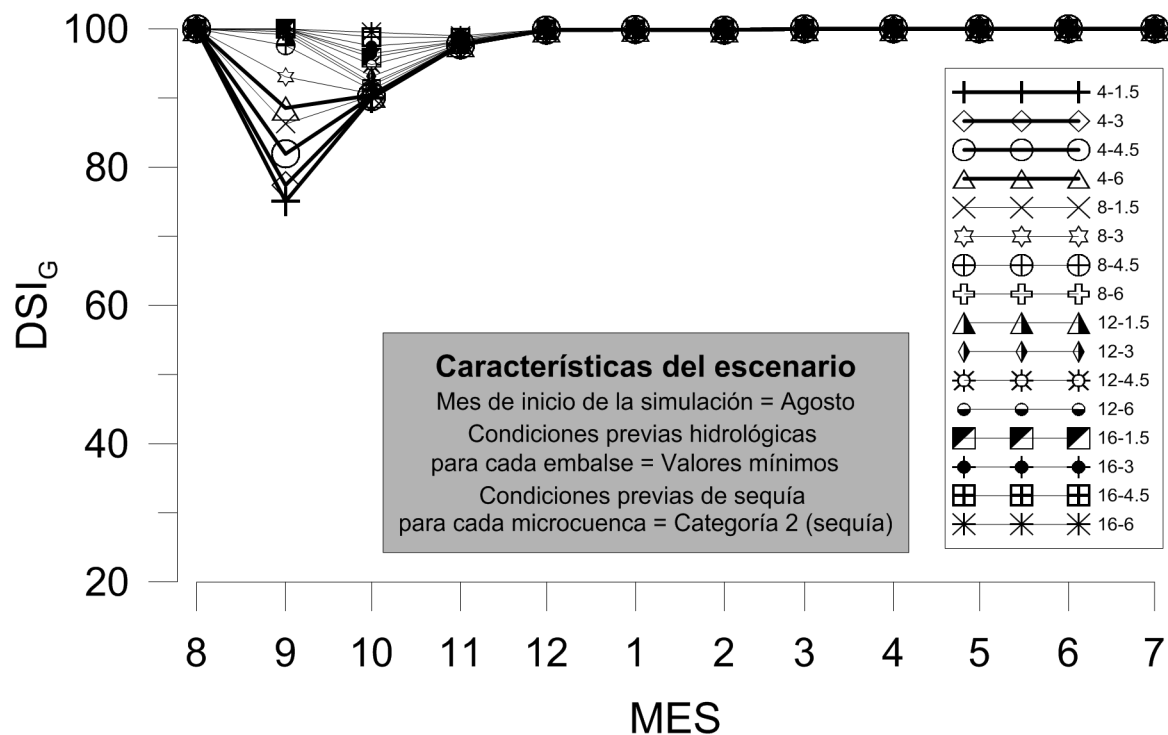


Figura 10. DSIg del sistema de recursos hídricos de la cuenca del río Machángara para un escenario con diferentes volúmenes iniciales de almacenamiento de los embalses de Chanlud y El Labrado aplicando la metodología de la Figura 4

Esta mejora en la toma de decisiones expuesta anteriormente coincide con algunos estudios:

Por ejemplo Sankarasubramanian et al. (2009) mejoró la asignación estacional e intraestacional del agua cuando utilizó los pronósticos de las probabilidades climatológicas de los flujos de entrada a los embalses y comprendió los retos y oportunidades en épocas de déficit y superávit de agua, respectivamente.

Por otro lado Pouget et al. (2015) mostró la mejora en la toma de decisiones cuando los pronósticos climáticos estacionales se integraron a las herramientas de gestión, sin embargo los pronósticos resultantes fueron presentados por medio de terciles (normal, bajo la normal y sobre la normal), por lo tanto no exploró una gran variedad de sequías (leves, moderadas, severas y extremas).

Igualmente los resultados del estudio de Gong et al. (2010) mostró también una mejora en las prácticas de la gestión del agua cuando los pronósticos de caudales basados en clima se incorporaron en las herramientas de operación de embalses, con lo cual redujeron el número de días de emergencia por las sequías.

5. CONCLUSIONES

Esta tesis propone un marco metodológico integral de evaluación del riesgo de fallo en el suministro a las demandas de un sistema de recursos hídricos mejorando metodologías tradicionales mediante la incorporación de pronósticos probabilísticos de sequía y proporcionando información que sirva de soporte para la toma de decisiones en la gestión del agua en tiempos de escasez. Los aportes de cuatro publicaciones (Anexos 1,2,3 y 4) realizadas en la fase de investigación del programa de Doctorado en Ingeniería del Agua y Medioambiental de la Universidad Politécnica de Valencia se han articulado con este marco general de evaluación del riesgo con la finalidad de presentar un trabajo completo y sistemático en la línea de investigación de análisis de riesgos de sequías.

La cuenca del río Machángara ubicada al sur del Ecuador fue escogida para aplicar este enfoque integral, los resultados obtenidos podrían ser útiles para la toma de decisiones con respecto a la formulación de medidas de prevención y mitigación para disminuir los impactos de los periodos de escasez de agua en el sistema de recursos hídricos.

La construcción de un DI agrupando la información disponible de las variables relacionadas con el agua a diferentes escalas ofrece la posibilidad de contemplar el fenómeno de manera integral y facilita la obtención de indicadores flexibles que vayan incorporando nuevas variables y permita actualizar la información de las variables existentes.

La comparación de los modelos basados en MC (comúnmente usado) y los modelos basados en BN (poco utilizado) proporcionó la oportunidad de evaluar los pronósticos probabilísticos de sequía de dos enfoques fundamentados en la probabilidad condicional y expuso la importancia de verificar la calidad de los pronósticos y el desempeño de los modelos antes de aceptar los resultados para la toma de decisiones en la gestión del agua.

La utilización de las funciones copulas para solucionar probabilidades conjuntas, expusieron una herramienta muy importante para captar la estructura de dependencia entre variables aleatorias.

La introducción de pronósticos probabilísticos de sequías en la generación de escenarios estocásticos constituye un paso importante para direccionar los posibles escenarios y tener situaciones más reales para analizar el riesgo de fallo de la asignación del agua para los diferentes usos.

Los resultados de la simulación de una gran cantidad de escenarios presentaron un panorama muy amplio para tomar decisiones, sin embargo poco manejable en algunas ocasiones. Es por esto que la utilización del índice integral DSI_G y la agrupación de estos indicadores mediante criterios específicos podrían ser de gran ayuda para almacenar los resultados de la simulación múltiple de una manera sistemática y pueda usarse esta información apropiadamente.

El riesgo de fallo en el sistema dependerá de los cambios en las demandas, variación de la disponibilidad de agua por las condiciones hidrometeorológicas cambiantes y las medidas llevadas a cabo para reducir los efectos de la sequía. Por lo tanto, cambios de estos factores en el futuro exigirá la actualización de los escenarios y por consiguiente variación de las medidas de gestión para la oferta y demanda de agua.

Cada uno de los escenario de simulación incluirá una o varias alternativas para disminuir el riesgo de fallo en los sistemas hídricos que dependerá del conocimiento y experiencia del operador o gestor del sistema. Sin embargo, estas medidas también estarán sometidas a la aceptación de los usuarios del agua en una cuenca, por lo que se hace necesario consensuar las mismas antes de su aplicación.

Finalmente indicar que algunos aspectos desarrollados en esta tesis servirá para elaborar otra publicación que se presentará para revisión bajo el título: “Marco metodológico integral de evaluación del riesgo de fallo en sistemas de recursos hidricos incorporando pronósticos de sequía. Caso de estudio en una cuenca andina regulada”.

5.1. FUTURAS INVESTIGACIONES

La trayectoria de investigación de esta tesis ha generado nuevos intereses en futuros trabajos. Algunos de ellos son:

- Investigar sobre los mecanismos que producen sequías en cuencas andinas para poder caracterizar de mejor forma los estados de sequía en el espacio y tiempo
- Evaluar el riesgo y la vulnerabilidad a los eventos de sequías con un enfoque integral que involucren aspectos sociales y económicos
- Explorar nuevos enfoques para el pronóstico de sequía que involucren predictores climáticos como la ocurrencia del Fenómeno del Niño, la Oscilación Decadal del Pacífico u otros fenómenos climáticos y su impacto en cuencas andinas.
- Indagar otros métodos de generación sintética y su influencia en la evaluación del riesgo de fallo de sistemas de recursos hídricos.
- Explorar otros métodos de optimización para formular reglas de gestión que involucren aspectos económicos, déficits de agua en las demandas u otros factores.
- Analizar los efectos del cambio climático y cambio del uso del suelo en la satisfacción futura de sistemas de recursos hídricos.

REFERENCIAS

- AghaKouchak A (2014) A baseline probabilistic drought forecasting framework using standardized soil moisture index: application to the 2012 United States drought. *Hydro Earth Syst Sci* 18:2485–2492. doi: 10.5194/hess-18-2485-2014
- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* 19:716–723. doi: 10.1109/TAC.1974.1100705
- Al-Faraj F, Scholz M, Tigkas D (2014) Sensitivity of Surface Runoff to Drought and Climate Change: Application for Shared River Basins. *Water* 6:3033–3048. doi: 10.3390/w6103033
- Anderson ML, Mierzwa MD, Kavvas ML (2000) Probabilistic seasonal forecasts of droughts with a simplified coupled hydrologic-atmospheric model for water resources planning. *Stoch Environ Res Risk Assess* 14:263–274. doi: 10.1007/s004770000049
- Anderson T (1962) On the distribution of the two-sample Cramer-von Mises criterion. *Ann Math Stat* 22:1148–1159. doi: 10.1214/aoms/1177704477.

- Andreu J (1993) Conceptos y métodos para la planificación hidrológica. CIMNE, BARCELONA, ESPAÑA
- Andreu J, Capilla J, Sanchís E (1996) AQUATOOL, a generalized decision-support system for water-resources planning and operational management. *J Hydrol* 177:269–291. doi: 10.1016/0022-1694(95)02963-X
- Andreu J, Solera A, Capilla J, Ferrer J (2007) Modelo SIMGES para simulación de cuencas. Manual de usuario v3. 00. Universidad Politécnica de Valencia, Valencia.
- Araghinejad S (2010) An Approach for Probabilistic Hydrological Drought Forecasting. *Water Resour Manag* 25:191–200. doi: 10.1007/s11269-010-9694-9
- Avilés A, Célleri R, Paredes J, Solera A (2015) Evaluation of Markov Chain Based Drought Forecasts in an Andean Regulated River Basin Using the Skill Scores RPS and GMSS. *Water Resour Manag* 29:1949–1963. doi: 10.1007/s11269-015-0921-2
- Avilés A, Célleri R, Solera A, Paredes J (2016) Probabilistic Forecasting of Drought Events Using Markov Chain-and Bayesian Network-Based Models: A Case Study of an Andean Regulated River Basin. *Water* 8:1–16.
- Avilés A, Solera A (2012) Estimación del riesgo de fallo en el suministro de agua como ayuda a la planificación y gestión de recursos hídricos. *MASKANA* 3:77–86.
- Avilés A, Solera A (2013) Análisis de sistemas de recursos hídricos de la cuenca del río Tomebamba en Ecuador, mediante modelos estocásticos y de gestión. In: Solera A, Paredes J, Andreu J (eds) *Aplicaciones de sistemas soporte a la decisión en planificación y gestión integradas de cuencas hidrográficas*. MARCOMBO, BARCELONA, ESPAÑA, pp 51–61
- Banimahd SA, Khalili D (2013) Factors Influencing Markov Chains Predictability Characteristics, Utilizing SPI, RDI, EDI and SPEI Drought Indices in Different Climatic Zones. *Water Resour Manag* 27:3911–3928. doi: 10.1007/s11269-013-0387-z
- Barua S, Asce SM, Ng AWM, Perera BJC (2011) Comparative Evaluation of Drought Indexes : Case Study on the Yarra River Catchment in Australia. *J Water Resour Plan Manag* 137:215–226. doi: 10.1061/(ASCE)WR.1943-5452.0000105.
- Barua S, Ng A, Perera B (2012) Drought assessment and forecasting: A case study on the Yarra River catchment in Victoria, Australia. *Aust J Water Resour* 15:95–108. doi: 10.7158/W10-848.2012.15.2

- Beniston M (2003) Climatic Change in Mountain Regions: A Review of Possible Impacts. *Clim Chang* 59:5–31.
- Brown C, Baroang KM, Conrad E, et al (2010) IRI Technical Report 10-15, Managing Climate Risk in Water Supply Systems. Palisades, NY
- Buytaert W, Céleri R, De Bièvre B, et al (2006a) Human impact on the hydrology of the Andean páramos. *Earth-Science Rev* 79:53–72. doi: 10.1016/j.earscirev.2006.06.002
- Buytaert W, Celleri R, Willems P (2006b) Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *J Hydrol* 329:413–421. doi: 10.1016/j.jhydrol.2006.02.031
- Buytaert W, Vuille M, Dewulf A, et al (2010) Uncertainties in climate change projections and regional downscaling in the tropical Andes: implications for water resources management. *Hydrol Earth Syst Sci* 14:1247–1258. doi: 10.5194/hess-14-1247-2010
- Cancelliere A, Mauro G Di, Bonaccorso B, Rossi G (2007) Drought forecasting using the Standardized Precipitation Index. *Water Resour Manag* 21:801–819. doi: 10.1007/s11269-006-9062-y
- Cancelliere A, Nicolosi V, Rossi G (2009) Assessment of Drought Risk in Water Supply Systems. In: *Coping with Drought Risk in Agriculture and Water Supply Systems*. Springer Netherlands, Dordrecht, pp 93–109
- Cattrysse D (1991) Linear programming and network flows M.S. BAZARAA, J.J. JARVIS and H.D. SHERALI Wiley, New York, 1990, xiv+684 pages, £15.95, second edition (first edition, 1977), ISBN 0-471-63681-9, cloth available at £43.20.
- Celleri R, Willems P, Buytaert W, Feyen J (2007) Space – time rainfall variability in the Paute Basin , Ecuadorian Andes. *Hydrol Process* 21:3316–3327. doi: 10.1002/hyp.6575
- Cherubini U, Luciano E, Vecchiato W (2004) Copula methods in finance. John Wiley & Sons
- Chung C, Salas JD (2000) Drought Occurrence Probabilities and Risks of Dependent Hydrologic Processes. *J Hydrol Eng* 5:259–268. doi: 10.1061/(ASCE)1084-0699(2000)5:3(259)
- Durdu ÖF (2010) Application of linear stochastic models for drought forecasting in the Büyük Menderes river basin, western Turkey. *Stoch Environ Res Risk Assess* 24:1145–1162. doi: 10.1007/s00477-010-0366-3

- Embrechts P, Lindskog F, McNeil A (2001) Modelling dependence with copulas. Zurich, Switzerland
- Fernández C, Vega J a., Fonturbel T, Jiménez E (2009) Streamflow drought time series forecasting: a case study in a small watershed in North West Spain. *Stoch Environ Res Risk Assess* 23:1063–1070. doi: 10.1007/s00477-008-0277-8
- Gandin LS, Murphy AH (1992) Equitable Skills scores for categorical forecast. *Mon Weather Rev* 120:361–370.
- Genest C, Rémillard B (2008) Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Ann l’Institut Henri Poincaré - Probab Stat* 44:1096–1127. doi: 10.1214/07-AIHP148
- Gerrity JP (1992) A note on Gandin and Murphy’s Equitable Skill Scores. *Mon Weather Rev* 120:2709–2712.
- Gong G, Wang L, Condon L, et al (2010) A Simple Framework for Incorporating Seasonal Streamflow Forecasts into Existing Water Resource Management Practices. *JAWRA J Am Water Resour Assoc* 46:574–585. doi: 10.1111/j.1752-1688.2010.00435.x
- Han P, Wang PX, Zhang SY, Zhu DH (2010) Drought forecasting based on the remote sensing data using ARIMA models. *Math Comput Model* 51:1398–1403. doi: 10.1016/j.mcm.2009.10.031
- Haro D, Solera A, Paredes J, Andreu J (2014) Methodology for drought risk assessment in within-year regulated reservoir systems. Application to the Orbigo River system (Spain). *Water Resour Manag* 28:3801–3814. doi: 10.1007/s11269-014-0710-3
- Hashimoto T, Loucks DP, Stedinger JR (1982) Reliability, Resiliency, and Vulnerability Criteria. *Water Resour Res* 18:14–20. doi: 10.1029/WR018i001p00014
- Heckerman D (1998) A tutorial on learning with Bayesian networks. In: *Learning in Graphical Models*. Springer Netherlands, pp 301–354
- Huang Y, Li YP, Chen X, et al (2010) Simulation-based optimization method for water resources management in Tarim River Basin, China. *Procedia Environ Sci* 2:1451–1460. doi: 10.1016/j.proenv.2010.10.158
- Hwang Y, Carbone GJ (2009) Ensemble Forecasts of Drought Indices Using a Conditional Residual Resampling Technique. *J Appl Meteorol Climatol* 48:1289–1301. doi:

10.1175/2009JAMC2071.1

Joe H (1997) *Multivariate Models and Multivariate Dependence Concepts*. CRC Press

Kao S-C, Govindaraju RS (2010) A copula-based joint deficit index for droughts. *J Hydrol* 380:121–134. doi: 10.1016/j.jhydrol.2009.10.029

Karavitis CA, Alexandris S, Tsesmelis DE, Athanasopoulos G (2011) Application of the Standardized Precipitation Index (SPI) in Greece. *Water* 3:787–805. doi: 10.3390/w3030787

Keyantash J a., Dracup J a. (2004) An aggregate drought index: Assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage. *Water Resour Res* 40:1–13. doi: 10.1029/2003WR002610

Khalili D, Farnoud T, Jamshidi H, et al (2011) Comparability Analyses of the SPI and RDI Meteorological Drought Indices in Different Climatic Zones. *Water Resour Manag* 25:1737–1757. doi: 10.1007/s11269-010-9772-z

Kolmogorov A (1933) Sulla determinazione empirica di una legge di distribuzione. *G dell' Ist Ital degli Attuari* 4:83–91.

Kumar V, Panu U (1997) Predictive assessment of severity of agricultural droughts based on agro-climatic factors. *J Am Water Resour Assoc* 33:1255–1264.

Labadie JW, Asce M (2004) Optimal Operation of Multireservoir Systems : State-of-the-Art Review. *J Water Resour Plan Manag* 130:93–111. doi: 10.1061/(ASCE)0733-9496(2004)130:2~93!

Lee S-E, Seo K-H (2013) The Development of a Statistical Forecast Model for Changma. *Weather Forecast* 28:1304–1321. doi: 10.1175/WAF-D-13-00003.1

Leilah a. a., Al-Khateeb S. (2005a) Statistical analysis of wheat yield under drought conditions. *J Arid Environ* 61:483–496. doi: 10.1016/j.jaridenv.2004.10.011

Leilah AA, Al-Khateeb S. (2005b) Statistical analysis of wheat yield under drought conditions. *J Arid Environ* 61:483–496. doi: 10.1016/j.jaridenv.2004.10.011

Li YP, Huang GH, Wang GQ, Huang YF (2009) FSWM: A hybrid fuzzy-stochastic water-management model for agricultural sustainability under uncertainty. *Agric Water Manag* 96:1807–1818. doi: 10.1016/j.agwat.2009.07.019

Liu WT, Juárez RIN (2001) ENSO drought onset prediction in northeast Brazil using NDVI.

- Int J Remote Sens 22:3483–3501. doi: 10.1080/01431160010006430
- Liu Y, Gupta H, Springer E, Wagener T (2008a) Linking science with environmental decision making: Experiences from an integrated modeling approach to supporting sustainable water resources management. *Environ Model Softw* 23:846–858. doi: 10.1016/j.envsoft.2007.10.007
- Liu Y, Gupta H, Springer E, Wagener T (2008b) Linking science with environmental decision making: Experiences from an integrated modeling approach to supporting sustainable water resources management. *Environ Model Softw* 23:846–858. doi: 10.1016/j.envsoft.2007.10.007
- Lohani VK, Loganathan G V (1997) An early warning system for drought management using the Palmer drought index. *J Am Water Resour Assoc* 33:1375–1386.
- Madadgar S, Moradkhani H (2013) A Bayesian framework for probabilistic seasonal drought forecasting. *J Hydrometeorol* 14:1685–1706. doi: 10.1175/JHM-D-13-010.1
- Madadgar S, Moradkhani H (2014) Spatio-temporal drought forecasting within Bayesian networks. *J Hydrol* 512:134–146. doi: 10.1016/j.jhydrol.2014.02.039
- Mason SJ (2004) On Using ““ Climatology ”” as a Reference Strategy in the Brier and Ranked Probability Skill Scores. *Mon Weather Rev* 132:1891–1895.
- Mauget S a., Ko J (2008) A Two-Tier Statistical Forecast Method for Agricultural and Resource Management Simulations. *J Appl Meteorol Climatol* 47:1573–1589. doi: 10.1175/2007JAMC1749.1
- McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scales. In: *Proceedings of the 8th Conference on Applied Climatology*. American Meteorological Society Boston, MA, pp 179–183
- Mishra a. K, Desai VR (2005) Drought forecasting using stochastic models. *Stoch Environ Res Risk Assess* 19:326–339. doi: 10.1007/s00477-005-0238-4
- Mishra a. K, Desai VR (2006) Drought forecasting using feed-forward recursive neural network. *Ecol Modell* 198:127–138. doi: 10.1016/j.ecolmodel.2006.04.017
- Mishra AK, Desai VR, Singh VP, Asce F (2007) Drought Forecasting Using a Hybrid Stochastic and Neural Network Model. *J Hydrol Eng* 12:626–638. doi: 10.1061/(ASCE)1084-0699(2007)12:6(626)

- Mishra AK, Singh VP (2010) Review paper A review of drought concepts. *J Hydrol* 391:202–216. doi: 10.1016/j.jhydrol.2010.07.012
- Mishra AK, Singh VP (2011) Drought modeling – A review. *J Hydrol* 403:157–175. doi: 10.1016/j.jhydrol.2011.03.049
- Modarres R (2007) Streamflow drought time series forecasting. *Stoch Environ Res Risk Assess* 21:223–233. doi: 10.1007/s00477-006-0058-1
- Moreira EE, Coelho C a., Paulo A a., et al (2008) SPI-based drought category prediction using loglinear models. *J Hydrol* 354:116–130. doi: 10.1016/j.jhydrol.2008.03.002
- Morid S, Smakhtin V, Bagherzadeh K (2007) Drought forecasting using artificial neural networks and time series of drought indices. *Int J Climatol* 27:2103–2111. doi: 10.1002/joc.1498
- Muller WA, Appenzeller C, Doblas-Reyes FJ, Liniger MA (2005) A Debaised Ranked Probability Skill Score to Evaluate Probabilistic Ensemble Forecasts with Small Ensemble Sizes. *J Clim* 18:1513–1523. doi: 10.1175/JCLI3361.1
- Murphy A (1971) A note on the ranked probability score. *J Appl Meteorol* 10:155–156.
- Murphy A (1993) What is a good forecast? An essay on the nature of goodness in weather forecasting.
- Murphy AH (1977) The Value of Climatological, Categorical and Probabilistic Forecasts in the Cost-Loss Ratio Situation. *Mon Weather Rev* 105:803–816. doi: 10.1175/1520-0493(1977)105<0803:TVOCCA>2.0.CO;2
- Nalbantis I, Tsakiris G (2009) Assessment of hydrological drought revisited. *Water Resour Manag* 23:881–897. doi: 10.1007/s11269-008-9305-1
- Nelsen RB (2006) *An Introduction to Copulas*, Second Edi. Springer, USA
- Palmer W (1965) Meteorological drought, Paper 45. US Department of Commerce, Weather Bureau Washington, DC, USA, Washington, DC, USA
- Panu US, Sharma TC (2002) Challenges in drought research: some perspectives and future directions. *Hydrol Sci J* 47:S19–S30. doi: 10.1080/02626660209493019
- Paulo A a., Pereira LS (2007) Prediction of SPI drought class transitions using Markov chains. *Water Resour Manag* 21:1813–1827. doi: 10.1007/s11269-006-9129-9

- Pearl J (1985) Bayesian networks: A model of self-activated memory for evidential reasoning. In: Proceedings of the 7th Conference of the Cognitive Science Society. University of California, Irvine, CA. pp 329–334
- Pouget L, Roldán T, Gómez M, et al (2015) Use of seasonal climate predictions in the water sector—preliminary results from the EUPORIAS project. In: Andreu J, Solera A, Paredes J, et al. (eds) Drought: Research and Science-Policy Interfacing. Taylor & Francis Group, London,UK, p 247
- Ries H, Schlünzen KH, Brümmer B, et al (2010) Impact of surface parameter uncertainties on the development of a trough in the Fram Strait region. *Tellus A* 62:377–392. doi: 10.1111/j.1600-0870.2010.00451.x
- Robertson DE, Wang QJ (2013) Seasonal Forecasts of Unregulated Inflows into the Murray River, Australia. *Water Resour Manag* 27:2747–2769. doi: 10.1007/s11269-013-0313-4
- Rossi G, Caporali E, Garrote L (2012) Definition of Risk Indicators for Reservoirs Management Optimization. *Water Resour Manag* 26:981–996. doi: 10.1007/s11269-011-9842-x
- Salas JD, Delleur JW, Yevjevich V, Lane W (1980) Applied Modeling of Hydrologic Time Series. Water Resources Publication, Littleton,Colorado
- Sánchez-Chóliz J, Sarasa C (2015) River Flows in the Ebro Basin: A Century of Evolution, 1913–2013. *Water* 7:3072–3082. doi: 10.3390/w7063072
- Sánchez S, Andreu J, Solera A (2001) Gestión de Recursos Hídricos con Decisiones Basadas en Estimación del Riesgo. UNIVERSIDAD POLITÉCNICA DE VALENCIA, VALENCIA, ESPAÑA
- Sandoval-Solis S, McKinney DC, Loucks M (2011) Sustainability index for water resources planning and management. *J Water Resour Plan Manag* 137:381–390. doi: 10.1061/(ASCE)WR.1943-5452.0000134.
- Sankarasubramanian A, Lall U, Devineni N, Espinueva S (2009) The role of monthly updated climate forecasts in improving intraseasonal water allocation. *J Appl Meteorol Climatol* 48:1464–1482. doi: 10.1175/2009JAMC2122.1
- Schwarz G (1978) Estimating the Dimension of a Model. *Ann Stat* 6:461–464.
- Shukla S, Wood AW (2008) Use of a standardized runoff index for characterizing hydrologic

- drought. *Geophys Res Lett* 35:1–7. doi: 10.1029/2007GL032487
- Sklar A (1959) Fonctions de répartition à n dimensions et leurs marges. *Publ Inst Stat Univ Paris* 8:229–231.
- Staudinger M, Stahl K, Seibert J (2014) A drought index accounting for snow. *Water Resour Res* 50:7861–7872. doi: 10.1002/2013WR015143
- Steinemann A (2006) Using Climate Forecasts for Drought Management. *J Appl Meteorol Climatol* 45:1353 – 1361. doi: 10.1175/JAM2401.1
- Steinemann A (2003) Drought indicators and triggers: A stochastic approach to evaluation. *J Am Water Resour Assoc* 39:1217–1233.
- Steinemann AC, Cavalcanti LF (2006) Developing multiple indicators and triggers for drought plans. *J Water Resour Plan Manag* 132:164–174. doi: 10.1061/(ASCE)0733-9496(2006)132:3(164)
- Sveinsson O, Salas JD, Lane W, Frevert D (2007) *Stochastic Analysis, Modeling, and Simulation (SAMS) Version 2007, User’s Manual*.
- Svoboda M, Hayes M, Wilhite D, Tadesse T (2004) Recent Advances in Drought Monitoring. *Drought Mitig Cent Fac Publ* 6:6.
- Thulasiraman K, Swamy MNS (2011) *Graphs: Theory and Algorithms*. John Wiley & Sons
- Todini E (2008) A model conditional processor to assess predictive uncertainty in flood forecasting. *Int J River Basin Manag* 6:123–137.
- Tsakiris G, Vangelis H (2005) Establishing a drought index incorporating evapotranspiration. *Eur Water* 9:3–11.
- Vicente-Serrano SM, Aguilar E, Martínez R, et al (2016) The complex influence of ENSO on droughts in Ecuador. *Clim Dyn* 1–23. doi: 10.1007/s00382-016-3082-y
- Vicente-Serrano SM, Beguería S, López-Moreno JI (2010) A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *J Clim* 23:1696–1718. doi: 10.1175/2009JCLI2909.1
- Vicente-Serrano SM, Chura O, López-Moreno JI, et al (2015) Spatio-temporal variability of droughts in Bolivia: 1955-2012. *Int J Climatol* 35:3024–3040. doi: 10.1002/joc.4190
- Viviroli D, Archer DR, Buytaert W, et al (2011) Climate change and mountain water

resources: overview and recommendations for research, management and policy. *Hydrol Earth Syst Sci* 15:471–504. doi: 10.5194/hess-15-471-2011

Westphal KS, Laramie RL, Borgatti D, Stoops R (2007) Drought Management Planning with Economic and Risk Factors. *J Water Resour Plan Manag* 133:351–362. doi: 10.1061/(ASCE)0733-9496(2007)133:4(351)

Wilks DS (2011) *Statistical Methods in the Atmospheric Sciences*, Third Edit. Academic Press, USA

World Meteorological Organization (2012) *Standardized Precipitation Index User Guide*.

Yan J (2006) Multivariate modeling with copulas and engineering applications. In: Pham H (ed) *Springer Handbook of Engineering Statistics*. Springer London, London, pp 973–990

Yan J (2007) Enjoy the Joy of Copulas: With a Package Copula. *J Stat Softw* 21:1–21.

Yuan X, Wood EF, Chaney NW, et al (2013) Probabilistic Seasonal Forecasting of African Drought by Dynamical Models. *J Hydrometeorol* 14:1706–1720. doi: 10.1175/JHM-D-13-054.1

Zhang H, Casey T (2000) Verification of categorical probability forecasts. *Weather Forecast* 15:80–89.

Zhong S, Shen L, Sha J, et al (2015) Assessing the Water Parallel Pricing System against Drought in China: A Study Based on a CGE Model with Multi-Provincial Irrigation Water. *Water* 7:3431–3465. doi: 10.3390/w7073431

ANEXO 1: Evaluation of Markov Chain based drought forecasts in an Andean regulated river basin using the skill scores RPS and GMSS¹

Abstract

On behalf of the decision-makers of Andean regulated river basins a drought index was developed to predict the occurrence and extent of drought events. Two stochastic models, the Markov Chain First Order (MCFO) and the Markov Chain Second Order (MCSO) model, predicting the frequency of monthly droughts were applied and the performance checked using two skill scores, respectively the ranked probability score (RPS) and the Gandin-Murphy skill score (GMSS). Data of the Chulco River basin (3200-4300 m.a.s.l.), situated in the Ecuadorian southern Andes, were employed to test the performance of both models. Results indicate that events with greater drought severity were more accurately predicted. The study also revealed the importance of verifying the quality of the forecasts and to have an assessment of the likely performance of the forecasting models before adopting any model and accepting the resulting information for decision-making.

Keywords Drought index, probabilistic forecast, Markov Chains, forecast evaluation, Andean basins

1 Introduction

Droughts occur in virtually all climate zones and are mostly related to a reduction in precipitation over an extended period of time (Mishra and Desai 2005; Mishra and Singh 2010). As stated by these authors, it is likely that the effects of climate change might increase the frequency of drought events and the level of severity with more devastating impacts. Droughts are likely more harsh in basins where human, agricultural, and industrial and ecosystem consumption exceeds the natural water yield. Mountain basins, which traditionally are appreciated for the provision and regulation of water to the local and downstream communities (Beniston 2003), are not exempted from droughts. Buytaert et al. (2006b) stated

¹ Avilés A, Célleri R, Paredes J, Solera A (2015) Evaluation of Markov Chain Based Drought Forecasts in an Andean Regulated River Basin Using the Skill Scores RPS and GMSS. *Water Resources Management* 29:1949–1963. doi: 10.1007/s11269-015-0921-2. La publicación final esta disponible en <http://link.springer.com/article/10.1007/s11269-015-0921-2>

that the high rainfall variability of mountain basins and the effects of climate change can result in short to long periods of drought. This could jeopardize the water supply, a concern of water managers, and raise the public awareness of the vulnerability of earth's water resources systems (Celleri et al. 2007).

Water management becomes a serious problem when the hydro-climatic variability exceeds the range of experience, i.e., when there are extreme events (e.g., floods, droughts, heat waves) which have not been recorded in the past. Traditionally, water managers rely on experiences or simple methods providing an answer on how to manage river basin water resources under normal to moderate extreme conditions. Additionally, Liu et al. (2008b) and Viviroli et al. (2011) state that water managers are reluctant to accept new scientific innovations in decision-making. One reason according to Labadie (2004) is that water managers are skeptical about the replacement of their judgment by prescribed solution strategies, even more in developing countries where there is little climatic and hydrologic information and a low culture of risk management present. A way of accommodating this problem is the introduction of simple approaches, providing reliable, timely and manageable information easily to integrate in decision-making.

As mentioned by Westphal et al. (2007) a preventive approach means the timely identification and prediction of droughts and their severity, providing accurate information to water managers on the likely impact on the basin water supply and reserves. This requires a better understanding of the drought phenomenon (Paulo and Pereira 2007) through analysis of existing data series, and the use of this information as input for decision-making methods (Steinemann and Cavalcanti 2006). For drought and flood management, the water management community traditionally uses a framework of indices (Shukla and Wood 2008). Indices not only provide an excellent basis for a better understanding of the risk of droughts, but also facilitate the ability to assess, predict and provide early warning of these events (Svoboda et al. 2004).

The Standardized Precipitation Index (SPI) (McKee et al. 1993) based on precipitation data is one of the most widely used in meteorological drought forecasting (Mishra and Desai 2005; Mishra et al. 2007; Paulo and Pereira 2007; Cancelliere et al. 2007; Moreira et al. 2008; Nalbantis and Tsakiris 2009; Khalili et al. 2011). The strength of SPI is that the index can be calculated for a variety of time scales, enabling the monitoring and prediction of short- and long-term basin water conditions (Mishra and Singh 2010). However, a disadvantage of SPI is that the index only is based on precipitation data (Tsakiris and Vangelis 2005), and as

stated by Mishra and Singh (2010) the value of SPI is affected by the length of the available precipitation record and the type of probability distribution used for the record. Lohani and Loganathan (1997) developed an early warning system of drought events using the Palmer Drought Severity Index (PDSI) (Palmer 1965), an index based on precipitation and temperature data. Although, PDSI is widely used as drought index, the main limitation is that the index is based on arbitrary rules and assumptions (Mishra and Singh 2010). Other indices used for the analysis and prediction of meteorological droughts are the Standardized Precipitation Evapotranspiration Index (SPEI) (Vicente-Serrano et al. 2010) and the Reconnaissance Drought Index (RDI) (Tsakiris and Vangelis 2005). Both these indices use as input precipitation and calculated evapotranspiration data (Khalili et al. 2011; Banimahd and Khalili 2013). In comparison to the rainfall based indices, the SPEI and RDI indices require information of many more meteorological variables. The Streamflow Drought Index (SDI), which uses as input streamflow data, is primarily employed for the evaluation and prediction of hydrological droughts. Concordant to Nalbantis and Tsakiris (2009) is SDI more representative for the deficient water balance conditions of a basin, than SPI. However, as the method of calculating the SPI and SDI are similar, the latter is subject to the same disadvantages as SPI, it is the length of the data record and the type of used probability distribution function (Mishra and Singh 2010). Finally, Barua et al. (2011) proposed the Aggregate Drought Index (ADI) (Keyantash and Dracup 2004) to evaluate and predict drought events. This index is based on water-related variables that include precipitation, evapotranspiration, streamflow, reservoir storage, soil moisture and snow water content. In comparison to other indices, the ADI describes uniquely droughts beyond the traditional meteorological, hydrological, and agricultural subcategories (Barua et al. 2011). A limitation hindering the use of this index is the broad spectrum of information needed to calculate the index.

Most indices are developed for a specific region and can therefore not blindly transferred to regions with different hydro-climatic and basin characteristics (Mishra and Singh 2010; Barua et al. 2012). Furthermore, it is noticed that more and more researchers in this area believe that the accurate estimation of droughts requires the integration of as many as possible hydro-meteorological variables (Keyantash and Dracup 2004; Barua et al. 2012). The main limitation however is the availability of the spectrum of hydro-meteorological data, particular in mountain basins (Buytaert et al. 2006a).

Notwithstanding the limitation of the availability of hydro-meteorological data, in this study a similar approach as the Aggregate Drought Index (ADI) for the construction of the monthly drought index (DI) was applied. Preference to a similar approach was inspired by the flexibility of the ADI method in the aggregation of basin water-related variables and the possibility to selectively use variables depending on the specific characteristics of the basin of interest (Keyantash and Dracup 2004). An additional element that affected our preference is the monthly scale of ADI, because as stated by Panu and Sharma (2002) the monthly scale is most suitable for the monitoring and prediction of drought effects in water supply situations.

For the categorization of the severity of drought events and for making input information and output of forecasting models easier to manage, some studies transform continuous time series to discrete time series. The Markov Chain theory (MC) lends itself excellent for this purpose (Lohani and Loganathan 1997; Paulo and Pereira 2007; Cancelliere et al. 2007; Nalbantis and Tsakiris 2009; Khalili et al. 2011; Banimahd and Khalili 2013). MC based models namely enable the characterization of the stochastic behavior of extreme events and its forecast in a probabilistic context, i.e. they possess the advantage of addressing the uncertainty associated with the forecast in a quantitative way; an important issue in decision-making (Murphy 1977).

A mayor issue in predictions is the verification of the quality of the forecast, involving comparison between observed and predicted values (Wilks 2011). A variety of evaluation methods of forecasting models using continuous variables exist (Mishra and Desai 2005; Mishra et al. 2007; Cancelliere et al. 2007; Barua et al. 2012; Robertson and Wang 2013), whereas in this study discrete variables are used. Zhang and Casey (2000) stated that it is necessary to apply different skill assessment approaches in model forecasting evaluation in order to obtain an objective assessment of any given forecast scheme. An additional objective of the research therefore was to identify verification methods with focus on both the forecasted value and the probability associated with it; methods that also penalizes forecasts that are farthest from the observation and rewarding correct forecasts, especially the correct forecasts of less frequent events. For those reasons, the Markov Chain First Order model (MCFO) and Markov Chain Second Order model (MCSO) were used as DI, valorizing the available hydro-meteorological information of the basin and performing probabilistic forecasts of the following four drought events: no-drought, moderate, severe and extreme drought event. The quality of the forecasts of discrete events was evaluated using the Ranked

Probability Score (RPS) (Murphy 1971) and the Gandin-Murphy Skill Score (GMSS) (Gandin and Murphy 1992). Several authors have been using those skill scores; among them Zhang and Casey (2000), Mason (2004), Muller et al. (2005), Mauget and Ko (2008), Ries et al. (2010) and Lee and Seo (2013). The RPS is sensitive to the distance between the predicted and observed drought event, it is the index possesses the capacity to increasingly penalize forecasts further away from the actual outcome (Wilks 2011). Whereas GMSS is also sensitive to distance, this index additionally has the property of being equitable, it implies that correct forecasts of less frequent events (such as extreme droughts) are weighted more strongly than correct forecasts of more common events, which discourages distortion of forecasts toward the more common event in order to artificially inflate the resulting score (Wilks 2011).

The manuscript presents subsequently the method used for the calculation of the DI and the identification of drought events, a description of probabilistic forecast models and tools for the verification of the quality of the forecasts, followed by a description of the case study area, the Chulco river basin, the available data, the presentation and interpretation of results, and conclusions. Forecasts in the basin are meant to improve preventively the monthly based management of the El Labrado reservoir.

2 Methods

2.1 Calculation of the DI and identification of events

For the construction of the DI a methodology similar to Keyantash and Dracup (2004) was applied. Herein, information available from water-related variables are subjected to a principal components analysis (PCA) to discern dominant patterns of the observed data by transforming the original variables into new uncorrelated variables that represent the maximum fraction of the variability in the original data. For further details, the reader is referred to Wilks (2011). The first step in the calculation of the DI is the construction of square symmetric matrices of correlations between time series of r available hydro-meteorological datasets for each month of the year. These correlation matrices were subjected to a PCA, where the principal components are a transformation of the dataset of the original variables in terms of orthogonal uncorrelated components z_j ($1 < j \leq r$). The eigenvectors derived from the PCA establishes the relationship between the PCs and the original data:

$$Z = S * E \quad (1)$$

where Z is the $n \times r$ matrix of PCs, n the number of observations, S the $n \times r$ matrix of the standardized original information, and E the $r \times r$ matrix of eigenvectors. Using the approach of Keyantash and Dracup (2004) the DI is the first principal component (PC1), normalized by its standard deviation, or:

$$DI_{i,k} = \frac{Z_{i,1,k}}{\sigma} \quad (2)$$

where $DI_{i,k}$ is the DI value for month k in year i , $Z_{i,1,k}$ the PC1 for month k in year i , and σ the sample standard deviation of $Z_{i,1,k}$ for month k of all years. The DI utilizes only PC1 because it explains the largest fraction of the variance described by the full data set of r variables. Since PCs are orthogonal, it is mathematically incorrect to combine them into a single expression (Keyantash and Dracup 2004). Once the DI values are calculated for each month and year, they are in chronologic order rearranged into a single time series.

Identification of drought events requires that the DI values are converted into an equivalent categorization system based on percentiles of the empirical cumulative normal distribution function. The thresholds division categories used are similar as in Steinemann (2003), however, herein only 4 drought events are considered: extreme (category 3), severe (category 2), moderate (category 1) and no-drought (category 0). Further, the following division of drought events and categories based on the percentile values was applied: $0.2 < PV \leq 1 =$ no drought (category 0); $0.1 < PV \leq 0.2 =$ moderate drought (category 1); $0.05 < PV \leq 0.1 =$ severe drought (category 2); and $PV \leq 0.05 =$ extreme drought (category 3). In this way a new time series of categorical values, serving as input to the probabilistic forecast models, was obtained.

2.2 Probabilistic forecast models

The most widely used models of stochastic processes in discrete time series are derived from the Markov Chains (MC), which can be defined as a set of states of a system corresponding to elements mutually exclusive and collectively exhaustive (MECE) of a sample space describing a random variable. Wilks (2011) provides a detailed description of the mathematical formulation of MC models. The behavior of this type of model is governed by a set of matrices of transition probabilities indicating the probability of occurrence of the states of a system for a time interval of the future, based on the current status information

and/or statements of past intervals, depending on the model order. For the selection of the appropriate order dependence of MC models representing categorical time series use was made of the following common criteria: the Akaike Information Criterion (AIC) (Akaike 1974) and the Bayesian Information Criterion (BIC) (Schwarz 1978). Both criteria try to find the order dependence of the model by producing a balance between the goodness of fit, as reflected in a log-likelihood function, and a penalty that increases with the number of fitted parameters. The order dependency of the model is considered suitable when the AIC and BIC criteria are lowest. Depiction of the mathematical routines of the AIC and BIC are not detailed herein, but can be found in Wilks (2011). Since the MCFO and MCSO models most reduce the AIC and BIC criteria, both were used in the probabilistic forecasting and the verification of the forecasts. Additionally, the non-homogeneous model version of MC offers the possibility to adjust the model for each month of the year to account for seasonality.

2.2.1 MCFO model

According to the mathematical formulation of MCFO models (Steinemann (2003)), p_{ij} , represents the transition probability that X_{t+1} is equal to category j given X_t is equal to category i , which can be expressed as:

$$p_{ij} = \Pr\{X_{t+1} = j | X_t = i\} \quad (3)$$

The estimate of the transition probabilities (\hat{p}_{ij}) can be calculated to account for the conditional relative frequencies of transitions (m_{ij}), or:

$$\hat{p}_{ij} = \frac{m_{ij}}{\sum_j m_{ij}} \quad i, j = 1, \dots, s \quad (4)$$

where m_{ij} is the frequency that X is equal to the category i at time t and equal to category j at the time $t+1$. The numerator presents the number of transitions of category i to category j , and the denominator stands for the sum of the number of transitions of category i to any other category. The value of s is the number of categories of drought events or states of the system. T is the matrix of transition probabilities (p_{ij}) for multiple states of MCFO.

The transition matrix provides the information necessary to obtain the probabilistic forecast for the next step from the current information. For example at monthly time scale, if an initial state i is known in a given month, this state seeks in the transition probability matrix of the month in question and the transition probabilities to obtain the state of next month j . The

higher value of the transition probabilities of the system states predicts the pattern for state j in the next month. There may be forecast states with equal probability of occurrence. In this case the same category for the following month could be kept, since usually the managers do not change their practices unless there is a significant change in the probabilities of normal conditions (Zhang and Casey 2000). Managers could also choose the category with worst state to secure the management of the system.

Forecasts can also be formulated several times ahead by the development of decision trees, built by a sequence of forecasts at each time step corresponding to the knowledge of the state of the current period and the probabilistic information of transition matrices, as done by Lohani and Loganathan (1997).

2.2.2 MCSO model

Let p_{hij} represent the transition probability that X_{t+1} is equal to category j , given that X_t is equal to the category i and X_{t-1} equal to the category h , expressed as:

$$p_{hij} = \Pr\{X_{t+1} = j | X_t = i, X_{t-1} = h\} \quad (5)$$

As in the case of the MCFO model, the estimated transition probabilities (\hat{p}_{hij}) depend on the accounting for conditional relative frequencies of transitions (m_{hij}):

$$\hat{p}_{hij} = \frac{m_{hij}}{\sum_j m_{hij}} \quad h, i, j = 1, \dots, s \quad (6)$$

Or, given the variable X is equal to the category h at time $t-1$, and equal to the category i at time t , then the estimated probability that the future value of the variable X is equal to the category j at the time $t+1$ is \hat{p}_{hij} . The numerator is the number of transitions categories h,i to category j , and the denominator is the sum of the number of transitions categories h,i to any other category.

In this model the transition matrix T provides the information necessary to obtain the probabilistic forecast of the state of the next time step from the information of the states of the current and the previous time periods. Similarly as for the MCFO model, forecasts several times ahead can be made by decision trees encompassing the information of the states of the current and previous periods and the values of the transition matrices.

2.3 Forecast verification

In this study the RPS and GMSS are used to verify the quality of the probabilistic forecasts of both MC models. The routine calculation of the two methods are detailed in Wilks (2011).

2.3.1 Ranked Probability Score (RPS)

The RPS is a measure of the accuracy of the forecast in terms of assigned probabilities. This method is appropriate for evaluating probabilistic forecasts of ordinal categorical variables, based on the square error between the cumulative probabilities of forecasts and observations, including a penalty for the forecasts that are farther of the observation. It can be seen as the probabilistic extension of the mean squared error (MSE). Let s be the number of event categories and the number of probabilities included in each forecast, i.e., in each time interval. This would correspond to a vector of probabilities of forecasts for each category (y_1, y_2, \dots, y_s) where the sum of all its components equals 1, and a vector of probabilities of observations for each category (o_1, o_2, \dots, o_s) where one of these components corresponds to the event that occurred, equal to 1, and the other $s-1$ components will be equal to zero. The cumulative probabilities forecasts (Y_m) and observations (O_m), are respectively defined as a function of the components of the forecast and observation vectors, according to:

$$Y_m = \sum_{j=1}^m y_j \quad m = 1, \dots, s \quad (7)$$

and

$$O_m = \sum_{j=1}^m o_j \quad m = 1, \dots, s \quad (8)$$

Since Y_m and O_m are cumulative probability function components that by definition must add to one, the final sums of Y_s and O_s are always equal to one.

RPS is the sum of the squared differences between the components of the cumulative forecast (Eq. 7) and observations (Eq. 8) vectors, given by:

$$RPS = \sum_{m=1}^s (Y_m - O_m)^2 \quad (9)$$

A perfect forecast would result in a value of $RPS = 0$ and the worst forecasts would be very much different from zero.

2.3.2 Gandin-Murphy Skill Score (GMSS)

This method, also called the Gerrity score, is based on a weighing of the forecasts of droughts events, rewarding correct forecasts and penalizing incorrect ones, also rewarding further to less frequent events. This method checks the quality of forecasts around the set of pairs of forecasts and observations of the verification time series by constructing contingency tables $I \times J$, where I is the number of categories of events predicted and J is the number of categories of events observed, i.e., the table is composed of the relative frequencies of pairs of predicted (i) and observed (j) events.

The computation of GMSS involves first defining a set of scoring weights s_{ij} , $i = 1, \dots, I$, $j = 1, \dots, J$; each of which is applied to one of the joint probabilities of forecasts and observations $p(y_i, o_j)$. A general formulation for the calculation of the GMSS is:

$$GMSS = \sum_{i=1}^I \sum_{j=1}^J p(y_i, o_j) s_{ij} \quad (10)$$

Scoring weights are defined following the approach of Gerrity (1992) based on the observed sample information $p(o_j)$. First, the sequence of $J-1$ odds ratio (ratio of a probability p to its complementary probability $1-p$) is defined:

$$D(j) = \frac{1 - \sum_{r=1}^j p(o_r)}{\sum_{r=1}^j p(o_r)} \quad j = 1, \dots, J-1 \quad (11)$$

where r is a dummy summation index. The scoring weights for correct forecasts are then calculated according to:

$$s_{jj} = \frac{1}{J-1} \left[\sum_{r=1}^{j-1} \frac{1}{D(r)} + \sum_{r=j}^{J-1} D(r) \right] \quad j = 1, \dots, J \quad (12)$$

and the weights for the incorrect forecasts according to:

$$s_{ij} = \frac{1}{J-1} \left[\sum_{r=1}^{i-1} \frac{1}{D(r)} + \sum_{r=j}^{J-1} D(r) - (j-i) \right] \quad 1 \leq i < j \leq J \quad (13)$$

The summations in the Eqs. 12 and 13 are equal to zero if the lower index is larger than the upper index. Both these equations fully define the $I \times J$ scoring weights when symmetric errors are equally penalized; that is when $s_{ij} = s_{ji}$. Equation 12 gives more credit for correct

forecasts of rarer events and less credit for correct forecasts of common events. On the other hand, Eq. 13 also accounts for the intrinsic rarity of the J events, and increasingly penalizes errors for greater differences between the forecasted categories i and the observed categories j , through the penalty term $(j-i)$. Higher values of GMSS represent the best forecasts.

3 Case study

The methodology outlined in previous section was applied to the river basin Chulco, located in the Ecuadorian southern Andes. The basin is situated in the altitudinal range of 3200 and 4300 m.a.s.l. The anthropogenic impact is small and the basin consists of a succession of plains containing a large variety of lakes, peat bogs and wet grasslands, intermingled with shrubland and low-saturated forest patches. Most of the basin area is covered by páramo vegetation (Buytaert et al. 2006a). It is one of the few basins in the region containing a water regulation infrastructure. The El Labrado reservoir is situated at the basin outlet, storing water for human consumption, irrigation and hydroelectricity for use by the downstream city of Cuenca. Security in water delivery is threatened by drought periods challenging water managers, especially in making decisions regarding the operation of the El Labrado reservoir. The Chulco basin was selected to test the proposed stochastic methodology and to verify the usability of the information in decision-making.

Monthly rainfall and streamflow data for the 1970-2010 period were provided by the National Institute of Meteorology and Hydrology of Ecuador (INAMHI) and the Council Basin River Machangara (CBRM). Basin average monthly rainfall and monthly volumes of water entering the El Labrado reservoir were used for the stochastic prediction and the evaluation of predicted values. The flexibility of the DI calculation permitted the temporal aggregation of both variables as to compose additional time series at different time scales, such as 3, 6, 9 and 12 month interval periods. This process enabled the generation of five variables of total rainfall (PR1, PR3, PR6, PR9 and PR12) in mm, and five variables of volume of water entering the reservoir (VS1, VS3, VS6, VS9 and VS12) in hm^3 . Considering monthly seasonality, each time series of the 10 variables was divided for each month of the year, resulting in 10 time series for each month (10 time series for January, 10 time series for February, etc.). Figure 1 shows the variation of the average of monthly rainfall and average of monthly streamflow volume, respectively for 1, 3, 6, 9 and 12 months. Due to the

existence of different units and different scales of variables, the information was standardized by subtracting the monthly mean and dividing by the monthly standard deviation.

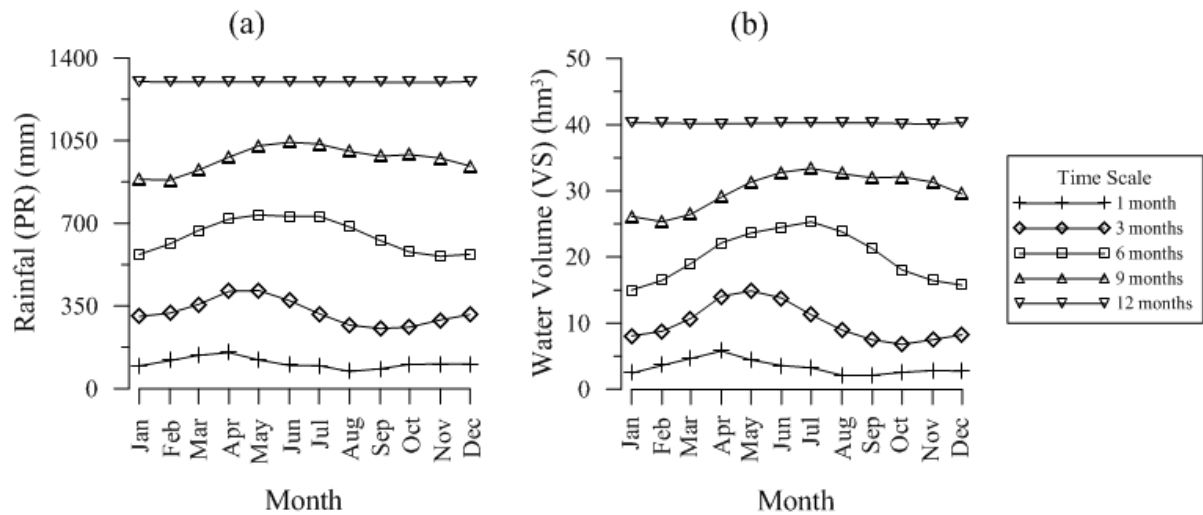


Fig 1 Monthly average of the time series of the ten variables of original information (1970-2010): (a) total rainfall and (b) total streamflow volume

4 Results

4.1 DI values and categories of drought events

With the standardized information of the time series of the 10 variables (PR1 to PR12, and VS1 to VS12) the correlation matrices for each month of the year were derived. Eigenvalues and eigenvectors were obtained for each month by applying the PCA to the correlation matrices. The results of the eigenvalues of the PC1 explain between 57 and 79% of the variance in the original dataset, with an average of 69% considering all the months of the hydrological year. Also were obtained twelve unit column vectors corresponding to the uncorrelated eigenvectors PC1 for each month, e.g. the eigenvector for April (rainiest month) is $e_1 = [0.13 \ 0.34 \ 0.35 \ 0.35 \ 0.32 \ 0.24 \ 0.34 \ 0.35 \ 0.35 \ 0.33]$.

Application of the Eqs. 1 and 2 yield the twelve sets of values of the DI. These values are chronologically rearranged to obtain the time series of the values of DI from December 1970 to December 2010. As explained to obtain a time series with categorical values of drought severity, each value of the DI was transformed to its respective percentile of the empirical cumulative normal distribution function. Figure 2 shows the values of the DI and thresholds

of severity of drought, from which the frequency and duration of each drought event (extreme, severe, moderate and non-drought) was derived.

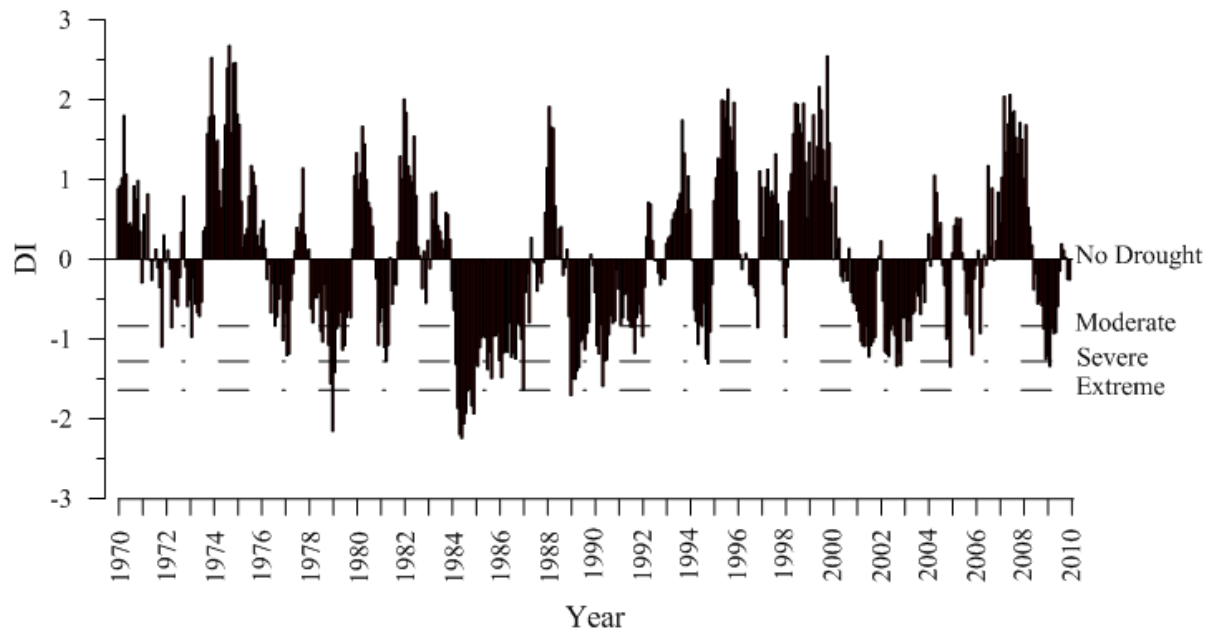


Fig 2 Time series (1970-2010) of the DI values in the Chulco river basin with thresholds of drought severity

4.2 Probabilistic forecast models

The information of the first 30 years of the categorical monthly time series X (1970-2000) was used for the adjustment of the MC models, where the categorical values are distributed in four states 0, 1, 2 and 3, corresponding respectively to no-drought, moderate, severe and extreme events. First the order dependence of models was examined by the AIC and BIC, with the lowest values for the MCFO model (AIC=358 and BIC=357) and MCSO model (AIC=407 and BIC=406). The obtained values do not have a significant difference compared with the values of the AIC and BIC of the MC models of zero and third order. The MCFO and MCSO models were adopted to probabilistic forecasts and used to test the methods applied for the evaluation of the quality of the forecasts. Twelve transition probability matrices were obtained taking into account the seasonality of the annual cycle. The transition probability matrix estimated for April using the MCFO model and the transition probability matrix estimated for March-April using the MCSO model are depicted in Table 1.

Table 1 Transition probability matrices, T, using the MCFO model for the month April and the MCSO model for the period March-April

Transition Probabilities		State j				
MCFO model for the month of April						
State i	0	1	2	3	Σ	
0	0.96	0.04	0.00	0.00	1.00	
1	0.75	0.25	0.00	0.00	1.00	
2	0.00	1.00	0.00	0.00	1.00	
3	0.00	0.00	0.00	1.00	1.00	
MCSO model for the period March-April						
States h-i	0	1	2	3	Σ	
0-0	0.95	0.05	0.00	0.00	1.00	
0-1	1.00	0.00	0.00	0.00	1.00	
0-2	0.00	1.00	0.00	0.00	1.00	
0-3	0.25	0.25	0.25	0.25	1.00	
1-0	1.00	0.00	0.00	0.00	1.00	
1-1	0.50	0.50	0.00	0.00	1.00	
1-2	0.25	0.25	0.25	0.25	1.00	
1-3	0.25	0.25	0.25	0.25	1.00	
2-0	0.25	0.25	0.25	0.25	1.00	
2-1	1.00	0.00	0.00	0.00	1.00	
2-2	0.00	1.00	0.00	0.00	1.00	
2-3	0.25	0.25	0.25	0.25	1.00	
3-0	0.25	0.25	0.25	0.25	1.00	
3-1	0.25	0.25	0.25	0.25	1.00	
3-2	0.25	0.25	0.25	0.25	1.00	
3-3	0.00	0.00	0.00	1.00	1.00	

Probabilistic forecasts of states of the next month j are shown in Table 2, respectively composed from the information of the current month i applying the MCFO model and derived from the information of previous month h and current month i applying the MCSO model. For example in Table 2 for the MCFO model, if in April there was an initial state with category 2, the most likely forecast for May would be a state with category 1. However for the MCSO model, if for May the forecast is equal to a state with category 1, the following states would occur in the months of March-April: 0-2, 1-1, 2-2 and 3-1.

Table 2 Probabilistic forecasts of the states of the following month (j) applying respectively the MCFO and the MSCO model

MCFO model												
Initial State i	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
0	0	0	0	0	0	0	0	0	0	0	0	0
1	1	0	1	0	1	1	1	0	1	0	0	0
2	1	2	2	1	2	1	2	1	3	2	3	2
3	3	3	3	3	3	3	3	2	3	3	2	2

MSCO model												
Initial States h-i	Jan-Feb	Feb-Mar	Mar-Apr	Apr-May	May-Jun	Jun-Jul	Jul-Aug	Aug-Sep	Sep-Oct	Oct-Nov	Nov-Dec	Dec-Jan
0-0	0	0	0	0	0	0	0	0	0	0	0	0
0-1	2	0	0	1	1	1	2	1	0	0	0	1
0-2	3	2	1	2	2	2	2	2	2	2	2	2
0-3	3	3	3	3	3	3	3	3	3	3	2	3
1-0	0	0	0	1	0	0	0	0	0	0	0	0
1-1	0	1	1	2	1	1	0	1	1	1	2	1
1-2	2	1	2	2	1	2	1	1	2	3	0	1
1-3	3	3	3	3	3	3	3	3	3	3	3	3
2-0	0	0	0	0	0	0	0	0	0	0	0	0
2-1	1	1	0	1	1	2	1	1	0	1	1	1
2-2	2	2	1	2	2	2	2	2	2	2	2	1
2-3	3	3	3	3	3	3	3	3	3	3	2	3
3-0	0	0	0	0	0	0	0	0	0	0	0	0
3-1	1	1	1	1	1	1	1	1	1	1	1	1
3-2	2	2	2	2	2	2	2	3	2	2	2	2
3-3	3	3	3	3	3	3	2	3	3	2	3	3

4.3 Decision trees

Making an analysis of frequencies across the categorical monthly time series X (1970-2010), it was found that the months of December, January and February showed the highest frequency of drought events. Therefore, this period of interest was chosen for the forecasting of drought events in a three month period by building decision trees with the probabilistic information of the transition matrices from the state of November for the MCFO model and the states of October and November for the MCSO model (Fig. 3).

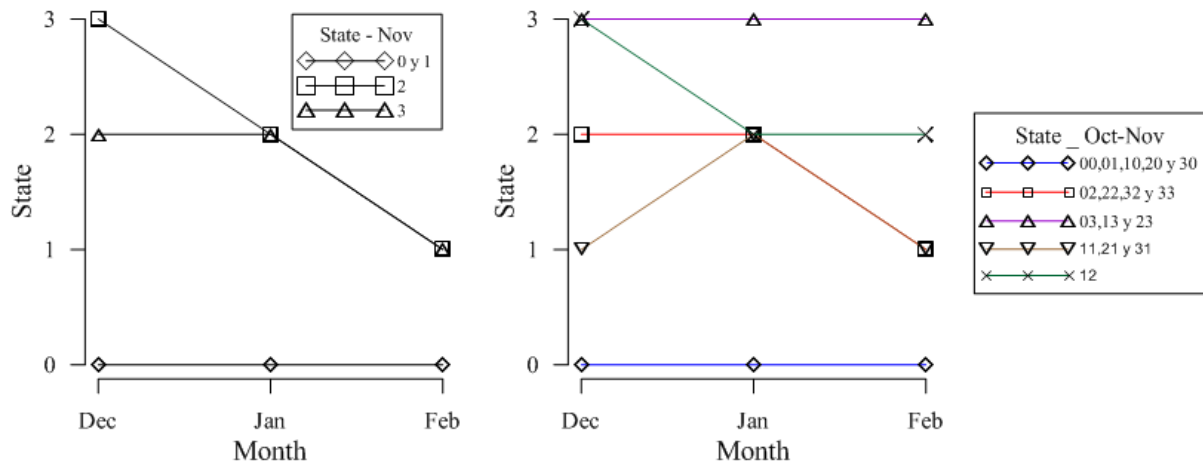


Fig 3 Tree based decisions of the MCFO model using state information of the month of November (left); and of the MCSO model using state information of the months October and November (right)

4.4 Forecast verification

The information for the last 10 years (2001-2010) of the categorical monthly time series was used to verify the quality of the forecasts. With monthly categorical information and the information in Table 2, event forecasts for 120 months and their respective probabilities of occurrence were derived. To assess a global check in the entire time series the RPS and the GMSS skill scores were calculated for each model. The RPS value was derived for all months in the time series resulting in a monthly average of 0.20 for the MCFO model and 0.25 for the MCSO model. If only the months with greater severity of droughts were considered in the assessment, the value of the monthly average of the RPS is 0.52 for the MCFO model and 0.50 for the MCSO model. Analyzing the forecasts with false alarms, i.e., forecasts indicating that there will be an event of drought, but in reality there was no drought or events with minor severity, both the MCFO model and MCSO model yielded respectively 6 and 12 months, less than 10% of the entire time series. Following the definition of the RPS method the forecasts of the MCFO model are better, although RPS values for both models did not differ much. Based on the RPS score of the most severe events, it is correct to state that both models perform equally, with a minor better score for the MCSO model. Instead the GMSS score method gave a value of 0.09 for the MCFO model and a value of 0.24 for the MCSO model, showing that the latter produces better forecasts, forecasting even better less frequent events, i.e., events with greater severity of drought. The MCSO model produced

47% of hits (correct forecasts) of moderate and severe droughts, while the MCFO model produced only 30% of hits; explaining why the GMSS score is higher for the MCSO than the MCFO model. Nevertheless, the overall performance shows that the proportions of correct forecasts were 77% for both models. For the assessment of the quality of the forecasts using the decision tree approach ten periods of three months (December, January and February) were selected from the verification time series, yielding a RPS equal to 0.42 and a GMSS equal to -0.07 for the MCFO model and a RPS equal to 0.35 and a GMSS equal to 0.39 for the MCSO model. Both models yielded 3 months with false alarms, less than 10% of all months of the assessment time series. Furthermore, the MCSO model indicated 43% of hits of severe and moderate droughts, while the MCFO model did not reveal such hits. Congruent to the skill score of both methods and the analysis of hits and false alarms, the MCSO model forecasts better in periods with higher frequency of drought events. Whereas, this conclusion suggests that the MCSO model performs better, the order dependence criteria given by the AIC and BIC values indicate that the MCFO model is the most appropriate for modeling. This illustrates the importance of verifying the quality of the forecasts before adopting any model and accepting the resulting information for decision-making.

5 Conclusions

The methodology described in this article provides the ability to forecast short-term drought events in a probabilistic framework, reducing the risk of subjectivity in decision-making with respect to preventive measures to manage water shortages for different uses. A DI, encompassing and prioritizing hydrological and meteorological information at different scales, was developed using basin information of the Chulco river basin; a regulated river basin in the southern Andean mountain range of Ecuador. The Markov Chain based probabilistic forecast models proved to be effective. Results show that the MCSO model performs better in forecasting drought events at monthly time scale, especially events with greater severity of drought, based on a better GMSS score. Therefore, the probabilistic information generated by this model could help the monthly or seasonal (three time period) decision-making of the management of the El Labrado reservoir. The proposed methodology seems easily adaptable for any system of water resources confronted with droughts events.

Acknowledgements

The research was conducted within the frame of the projects “Meteorological Cycles and Evapotranspiration along the Altitudinal Gradient of the Cajas National Park” and “Identification of hydro-meteorological processes that trigger extreme floods in the city of Cuenca using precipitation radar”. Both projects were funded by the University of Cuenca and the Public Municipal Company of Water Supply from Cuenca (ETAPA). Thanks are due to INAMHI and CBRM for providing the information of the Chulco river basin.

References

- Akaike H (1974) A new look at the statistical model identification. *IEEE Trans Automat Contr* 19:716–723. doi: 10.1109/TAC.1974.1100705
- Banimahd SA, Khalili D (2013) Factors Influencing Markov Chains Predictability Characteristics, Utilizing SPI, RDI, EDI and SPEI Drought Indices in Different Climatic Zones. *Water Resour Manag* 27:3911–3928. doi: 10.1007/s11269-013-0387-z
- Barua S, Asce SM, Ng AWM, Perera BJC (2011) Comparative Evaluation of Drought Indexes : Case Study on the Yarra River Catchment in Australia. *J Water Resour Plan Manag* 137:215–226. doi: 10.1061/(ASCE)WR.1943-5452.0000105.
- Barua S, Ng A, Perera B (2012) Drought assessment and forecasting: A case study on the Yarra River catchment in Victoria, Australia. *Aust J Water Resour* 15:95–108. doi: 10.7158/W10-848.2012.15.2
- Beniston M (2003) Climatic Change in Mountain Regions: A Review of Possible Impacts. *Clim Chang* 59:5–31.
- Buytaert W, Céleri R, De Bièvre B, et al. (2006a) Human impact on the hydrology of the Andean páramos. *Earth-Science Rev* 79:53–72. doi: 10.1016/j.earscirev.2006.06.002
- Buytaert W, Celleri R, Willems P (2006b) Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *J Hydrol* 329:413–421. doi: 10.1016/j.jhydrol.2006.02.031
- Cancelliere A, Mauro G Di, Bonaccorso B, Rossi G (2007) Drought forecasting using the Standardized Precipitation Index. *Water Resour Manag* 21:801–819. doi: 10.1007/s11269-006-9062-y

- Celleri R, Willems P, Buytaert W, Feyen J (2007) Space – time rainfall variability in the Paute Basin , Ecuadorian Andes. *Hydrol Process* 21:3316–3327. doi: 10.1002/hyp.6575
- Gandin LS, Murphy AH (1992) Equitable Skills scores for categorical forecast. *Mon Weather Rev* 120:361–370.
- Gerrity JP (1992) A note on Gandin and Murphy’s Equitable Skill Scores. *Mon Weather Rev* 120:2709–2712.
- Keyantash J a., Dracup J a. (2004) An aggregate drought index: Assessing drought severity based on fluctuations in the hydrologic cycle and surface water storage. *Water Resour Res* 40:1–13. doi: 10.1029/2003WR002610
- Khalili D, Farnoud T, Jamshidi H, et al. (2011) Comparability Analyses of the SPI and RDI Meteorological Drought Indices in Different Climatic Zones. *Water Resour Manag* 25:1737–1757. doi: 10.1007/s11269-010-9772-z
- Labadie JW, Asce M (2004) Optimal Operation of Multireservoir Systems : State-of-the-Art Review. *J Water Resour Plan Manag* 130:93–111. doi: 10.1061/(ASCE)0733-9496(2004)130:2~93!
- Lee S-E, Seo K-H (2013) The Development of a Statistical Forecast Model for Changma. *Weather Forecast* 28:1304–1321. doi: 10.1175/WAF-D-13-00003.1
- Liu Y, Gupta H, Springer E, Wagener T (2008) Linking science with environmental decision making: Experiences from an integrated modeling approach to supporting sustainable water resources management. *Environ Model Softw* 23:846–858. doi: 10.1016/j.envsoft.2007.10.007
- Lohani VK, Loganathan G V (1997) An early warning system for drought management using the Palmer drought index. *J Am Water Resour Assoc* 33:1375–1386.
- Mason SJ (2004) On Using ““ Climatology ”” as a Reference Strategy in the Brier and Ranked Probability Skill Scores. *Mon Weather Rev* 132:1891–1895.
- Mauget S a., Ko J (2008) A Two-Tier Statistical Forecast Method for Agricultural and Resource Management Simulations. *J Appl Meteorol Climatol* 47:1573–1589. doi: 10.1175/2007JAMC1749.1

- McKee TB, Doesken NJ, Kleist J (1993) The relationship of drought frequency and duration to time scales. Proc. 8th Conf. Appl. Climatol. American Meteorological Society Boston, MA, pp 179–183
- Mishra a. K, Desai VR (2005) Drought forecasting using stochastic models. *Stoch Environ Res Risk Assess* 19:326–339. doi: 10.1007/s00477-005-0238-4
- Mishra AK, Desai VR, Singh VP, Asce F (2007) Drought Forecasting Using a Hybrid Stochastic and Neural Network Model. *J Hydrol Eng* 12:626–638. doi: 10.1061/(ASCE)1084-0699(2007)12:6(626)
- Mishra AK, Singh VP (2010) Review paper A review of drought concepts. *J Hydrol* 391:202–216. doi: 10.1016/j.jhydrol.2010.07.012
- Moreira EE, Coelho C a., Paulo A a., et al. (2008) SPI-based drought category prediction using loglinear models. *J Hydrol* 354:116–130. doi: 10.1016/j.jhydrol.2008.03.002
- Muller WA, Appenzeller C, Doblas-Reyes FJ, Liniger MA (2005) A Debiased Ranked Probability Skill Score to Evaluate Probabilistic Ensemble Forecasts with Small Ensemble Sizes. *J Clim* 18:1513–1523. doi: 10.1175/JCLI3361.1
- Murphy A (1971) A note on the ranked probability score. *J Appl Meteorol* 10:155–156.
- Murphy AH (1977) The Value of Climatological, Categorical and Probabilistic Forecasts in the Cost-Loss Ratio Situation. *Mon Weather Rev* 105:803–816. doi: 10.1175/1520-0493(1977)105<0803:TVOCCA>2.0.CO;2
- Nalbantis I, Tsakiris G (2009) Assessment of hydrological drought revisited. *Water Resour Manag* 23:881–897. doi: 10.1007/s11269-008-9305-1
- Palmer W (1965) Meteorological drought, Paper 45. 65.
- Panu US, Sharma TC (2002) Challenges in drought research: some perspectives and future directions. *Hydrol Sci J* 47:S19–S30. doi: 10.1080/02626660209493019
- Paulo A a., Pereira LS (2007) Prediction of SPI drought class transitions using Markov chains. *Water Resour Manag* 21:1813–1827. doi: 10.1007/s11269-006-9129-9
- Ries H, Schlünzen KH, Brümmer B, et al. (2010) Impact of surface parameter uncertainties on the development of a trough in the Fram Strait region. *Tellus A* 62:377–392. doi: 10.1111/j.1600-0870.2010.00451.x

- Robertson DE, Wang QJ (2013) Seasonal Forecasts of Unregulated Inflows into the Murray River, Australia. *Water Resour Manag* 27:2747–2769. doi: 10.1007/s11269-013-0313-4
- Schwarz G (1978) Estimating the Dimension of a Model. *Ann Stat* 6:461–464.
- Shukla S, Wood AW (2008) Use of a standardized runoff index for characterizing hydrologic drought. *Geophys Res Lett* 35:1–7. doi: 10.1029/2007GL032487
- Steinemann A (2003) Drought indicators and triggers: A stochastic approach to evaluation. *J Am Water Resour Assoc* 39:1217–1233.
- Steinemann AC, Cavalcanti LF (2006) Developing multiple indicators and triggers for drought plans. *J Water Resour Plan Manag* 132:164–174. doi: 10.1061/(ASCE)0733-9496(2006)132:3(164)
- Svoboda M, Hayes M, Wilhite D, Tadesse T (2004) Recent Advances in Drought Monitoring. *Drought Mitig Cent Fac Publ* 6.
- Tsakiris G, Vangelis H (2005) Establishing a drought index incorporating evapotranspiration. *Eur Water* 9:3–11.
- Vicente-Serrano SM, Beguería S, López-Moreno JI (2010) A multiscalar drought index sensitive to global warming: The standardized precipitation evapotranspiration index. *J Clim* 23:1696–1718. doi: 10.1175/2009JCLI2909.1
- Viviroli D, Archer DR, Buytaert W, et al. (2011) Climate change and mountain water resources: overview and recommendations for research, management and policy. *Hydrol Earth Syst Sci* 15:471–504. doi: 10.5194/hess-15-471-2011
- Westphal KS, Laramie RL, Borgatti D, Stoops R (2007) Drought Management Planning with Economic and Risk Factors. *J Water Resour Plan Manag* 133:351–362. doi: 10.1061/(ASCE)0733-9496(2007)133:4(351)
- Wilks DS (2011) *Statistical Methods in the Atmospheric Sciences*, Third Edit. 704.
- Zhang H, Casey T (2000) Verification of categorical probability forecasts. *Weather Forecast* 15:80–89.

ANEXO 2: Probabilistic forecasting of drought events using Markov Chain and Bayesian Network based models. Case study: An Andean regulated river basin²

Abstract

The scarcity of water resources in mountain areas can distort normal water application patterns with among other effects, a negative impact on water supply and river ecosystems. Knowing the probability of droughts might help to optimize a priori the planning and management of the water resources in general and of the Andean watersheds in particular. This study compares Markov chain (MC) and Bayesian Network (BN) based models in drought forecasting using a recently developed drought index with respect to their capability to characterize different drought severity states. The copulas functions were used to solve the BNs and the Ranked Probability Skill Score (RPSS) to evaluate the performance of the models. Monthly rainfall and streamflow data of the Chulco River basin, located in southern Ecuador, were used to assess the performance of both approaches. Global evaluation results revealed that the MC-based models predict better wet and dry periods and BN-based models generate slightly more accurately forecasts of extreme droughts. However evaluation of monthly results reveals that for each month of the hydrological year either the MC- or BN-based model provides better forecasts. The presented approach could be of assistance to water managers to ensure that timely decision-making on drought response is undertaken.

Keywords: Probabilistic drought forecasting; Drought index; Markov Chains; Bayesian Networks; Copulas; Andean watersheds

1. Introduction

Droughts are recognized as an environmental disaster, being the consequence of a reduction in precipitation over an extended period of time (Mishra and Singh 2010), negatively affecting agriculture, domestic water supply and economic growth (Zhong et al. 2015), and even in some cases altering the functioning of natural ecosystems (Karavitis et al. 2011).

² Avilés A, Célleri R, Solera A, Paredes J (2016) Probabilistic Forecasting of Drought Events Using Markov Chain-and Bayesian Network-Based Models: A Case Study of an Andean Regulated River Basin. *Water* 8:1–16. doi:10.3390/w8020037. La publicación final esta disponible en <http://www.mdpi.com/2073-4441/8/2/37>

Droughts are especially important in regions where economic activities are highly dependent on water resources (Vicente-Serrano et al. 2015), such as in mountain regions which traditionally provide water resources and services to local communities and lowland residents (Beniston 2003). For example, the Andean basins which have been identified as excellent providers of water for multiple uses (Buytaert et al. 2006a; Buytaert et al. 2010) could be affected by droughts and climate change, putting the water supply at risk and augmenting the vulnerability of the basin's water resources systems and eco-services (Buytaert et al. 2006b; Celleri et al. 2007; Staudinger et al. 2014; Al-Faraj et al. 2014). The increasing tendency of water shortage is a concern of water managers (Celleri et al. 2007), wrestling with questions such as which drought indicator (Steinemann 2003) and threshold of the indicators to use to qualify the drought status. In fact, drought characterization requires indicators that are generally applicable, but specific enough to capture the type of drought relevant to the region and the variables of interest (Staudinger et al. 2014). Droughts may occur at different moments in the hydrological year, with the consequence that hydrologic variables might experience quite different levels of drought, complicating the overall assessment of the drought status (Madadgar and Moradkhani 2013). Furthermore, droughts can be the result of a succession of water shortages over different periods of time (Kao and Govindaraju 2010). Consequently, there is a need to assess the drought status using indices based on multiple variables monitored during different time windows. The present study uses the drought index (DI) developed by Avilés et al. (2015), which is based on water-related variables of different window-sizes in the hydrologic year, enabling the capturing of the drought status for short, medium and long periods. The main advantages of this index is the flexibility in aggregation of basin water-related variables and the possibility to selectively use variables depending of the availability hydro-meteorological data of the basin of interest.

The characterization of droughts and its reliable prediction would help water managers to make the decisions that make water supplies safe and reliable, and risk-adverse. To this end water managers need the instruments that enable them to mimic the water demand and supply systems using adequate models that also address uncertainties and vulnerabilities, and helps to prioritize risks (Viviroli et al. 2011). A major restriction in the formulation of strategies addressing the availability of water is the forecast of droughts (Sánchez-Chóliz and Sarasa 2015). It is generally accepted that the reliable forecast of drought events plays an important role in the development of adequate strategies for the planning and distribution of water resources (Madadgar and Moradkhani 2013; Madadgar and Moradkhani 2014). Several

methodologies to forecast drought saw daylight in recent years, and Mishra and Singh (2011) describe some of these approaches. Regression analysis (Kumar and Panu 1997; Liu and Juárez 2001; Leilah and Al-Khateeb 2005b) is a commonly used method; however it has the limitation of assuming linearity between predictor and predictand which makes the technique less suitable for long-lead forecasting (Mishra and Singh 2011). Time series models (Mishra and Desai 2005; Modarres 2007; Fernández et al. 2009; Durdu 2010; Han et al. 2010) are advantageous since they provide a framework for the identification, estimation and the diagnostic check for model development (Mishra and Desai 2005), but they are also, as the regression models, linear. Neural networks (Mishra and Desai 2006; Mishra et al. 2007; Morid et al. 2007) are nonlinear models having the capacity to discover patterns from data and estimate any complex functional relationship with high accuracy. However, major disadvantages of neural networks are their black-box nature and computational burden (Mishra and Singh 2011). The above listed models provide a deterministic prediction of the drought status, and does not consider the uncertainty associated with the forecast (Hwang and Carbone 2009). This aspect can be handled by probabilistic models, which forecast in a quantitative way droughts and the associated uncertainty (Murphy 1977). A variety of probabilistic models for drought forecasting have been developed (Lohani and Loganathan 1997; Chung and Salas 2000; Anderson et al. 2000; Paulo and Pereira 2007; Cancelliere et al. 2007; Hwang and Carbone 2009; Nalbantis and Tsakiris 2009; Araghinejad 2010; Madadgar and Moradkhani 2013; Yuan et al. 2013; Madadgar and Moradkhani 2014; AghaKouchak 2014; Avilés et al. 2015), but not that many calculate the conditional probability if there are multiple events, such as the Markov chains and Bayesian networks. Those approaches generate probabilistic forecasts of future droughts in function of earlier drought conditions. According to the probability conditional theory are the models based on MC most common for drought forecasting (Lohani and Loganathan 1997; Paulo and Pereira 2007; Cancelliere et al. 2007; Nalbantis and Tsakiris 2009; Avilés et al. 2015), while the BN-based models are more sophisticated. The latter have not been widely used for the probabilistic forecasting of drought events, notwithstanding they seem to have the ability of better forecasting droughts (Madadgar and Moradkhani 2013; Madadgar and Moradkhani 2014).

A BN is a graphical model that encodes the joint probability for a large set of random variables. They offer a natural way of describing the conditional dependencies of dependent variables in time, such as drought indices (DIs) with probabilities (Heckerman 1998). Since droughts are slowly evolving phenomena, strong temporal autocorrelation among DIs is

expected (Kao and Govindaraju 2010), and therefore they can be expressed within a BNs (Pearl 1985; Madadgar and Moradkhani 2014). A considerable effort is required to solve BNs, a task that can be simplified using copulas functions. Those functions are capable to combine several variables with different levels of correlation and dependence structures (Madadgar and Moradkhani 2013). The two most frequently used copulas families are Elliptic and Archimedean copulas (Yan 2007); and a commonly used list of copulas families can be found in (Embrechts et al. 2001; Nelsen 2006; Yan 2006). In this study, two types of each family, Elliptic and Archimedean are tested. To enhance the confidentiality of the forecasts, we used for the verification of forecasts the RPSS, which provides a measure of the accuracy of the forecast in terms of assigned probabilities (Murphy 1993; Mason 2004).

This study uses the Drought Index (DI) developed on the basis of rainfall and streamflow data of the Chulco River basin in southern Ecuador, as described in Aviles et al (2015). In this study, the Markov Chain First Order and the Markov Chain Second Order stochastic models were used to predict the frequency of monthly droughts. The aim of the present study is to compare MC-based models, a more classical approach, with BN-based models, a novel technique, for forecasting future droughts given knowledge of earlier droughts, and to assess the benefits of BNs over the first and second order MCs. Section 2 of the manuscript describes the methodology, including a brief description of the developed DI, the forecast models and the method used to verify forecasts. The 3rd section of the paper gives a snapshot of the study area, while the 4th and 5th section respectively present the results and conclusions.

2. Materials and Methods

2.1. Case Study

Data of the Andean Chulco mountain basin, situated in southern Ecuador at an altitude of 3200-4300 m.a.s.l (Figure 1), were used to examine the performance of the MC and BN model approaches. Most of the basin area is covered by páramo (tussock grass). The Andean hilly páramo region consist of glacier formed valleys and plains with a large variety of lakes, peat bogs and wet grasslands, intermingled with shrubland and low-statured montane forest patches (Buytaert et al. 2006a). The study basin is one of the few regulated basins in the region; with the El Labrado reservoir situated at the basin outlet. The reservoir stores water for urban, agricultural and industrial uses, and power generation. The basin is strategic

because it is one of the few multi-purpose water resources systems in southern Ecuador that benefits local and regional ecology and economies. From this point of view, research results might be useful for water managers working in similar environmental conditions.

For development of the DI, monthly time series of basin average rainfall and stored volume in the El Labrado reservoir, for the period 1971 to 2010 were collected from the National Institute of Meteorology and Hydrology of Ecuador (INAMHI) and the Council Basin River Machangara (CBRM). Figure 2 shows the variation of monthly average of the two time series of the variables. Similarly as in Avilés et al. (2015), to capture short, medium and large term droughts, time windows of 1, 3, 6, 9 and 12 months were selected for the generation of ten time series with varying time scales, respectively five time series of the total rainfall (PR1, PR3, PR6, PR9 and PR12) in mm, and five time series of the water volume entering the reservoir (VS1, VS3, VS6, VS9 and VS12) in hm^3 . The monthly seasonality is accounted by the division of the ten time series for each month. All data was standardized to jointly considered variables with different units.

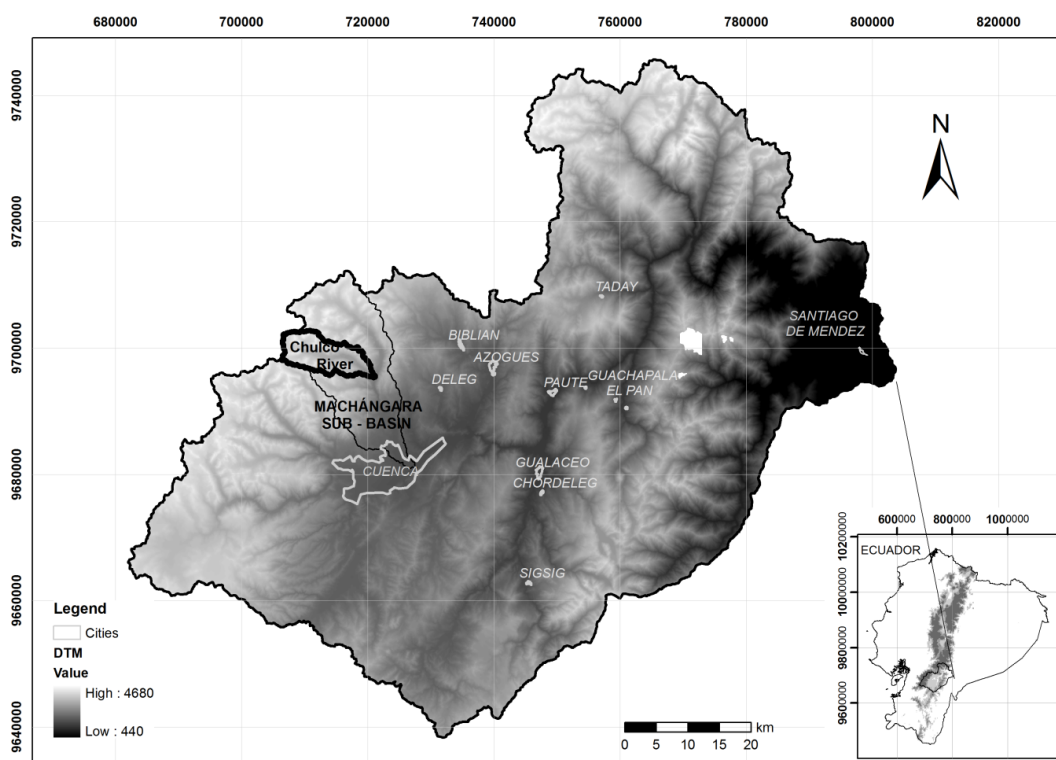


Figure 1. Location of the Chulco river basin in the Paute river basin.

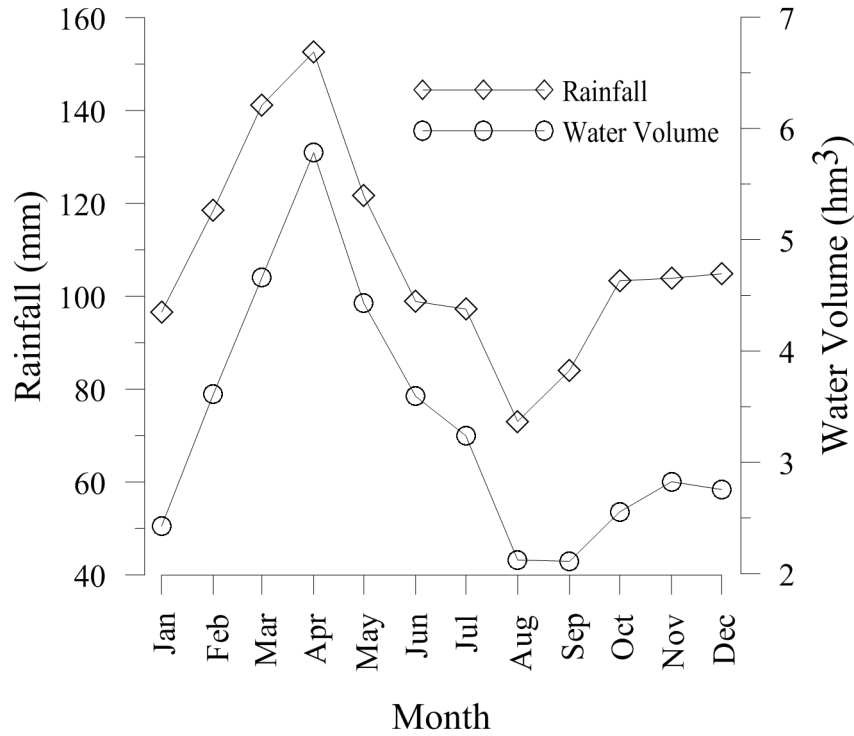


Figure 2. Monthly average of the time series of original information (1971-2010)

2.2. Drought Index

The DI used herein is derived subjecting available water-related variables of the study area to Principal Components Analysis (PCA), calculating the correlation matrices between time series of r available hydro-meteorological datasets, yielding the eigenvalues and eigenvectors. The eigenvectors establish the relationship between the principal components (PCs) and the original information as expressed by Avilés et al. (2015):

$$Z = D * E \quad (1)$$

where Z is the $o \times r$ matrix of PCs, o the number of observations, D the $o \times r$ matrix of the standardized original information, and E the $r \times r$ matrix of eigenvectors. DI is the first principal component (PC1), normalized by its standard deviation. Hence we have:

$$DI_{i,k} = \frac{Z_{i,1,k}}{\sigma_k} \quad (2)$$

where $DI_{i,k}$ is the DI value for the k^{th} month in the i^{th} year, $Z_{i,1,k}$ the PC1 for the k^{th} month in the i^{th} year, and σ_k the sample standard deviation of $Z_{i,1,k}$ of all years. Once the DI values for each month and year were calculated, they were chronological rearranged into a single time series.

The DI is a standardized index (SI) capable of capturing the anomalies from the average moisture condition of a basin as a function of the available information of water-related variables (Kao and Govindaraju 2010; Madadgar and Moradkhani 2013). Any phenomenon that can be quantified continuously, such as the drought index, is likely treated as a discrete variable by categorizing the time series considering thresholds for each state of drought. Hence, DI being a normal score (satisfied the null hypothesis of Kolmogorov-Smirnov test (Kolmogorov 1933) for normality), it was divided into the categories using the same thresholds indicating the World Meteorological Organization (World Meteorological Organization 2012), used to categorize the states of drought, we obtain the following drought categories: $DI > 0 =$ category 0 (no drought); $-1 < DI \leq 0 =$ category 1 (mild drought); $DI \leq -1 =$ category 2 (moderate, severe and extreme drought); the three states of category 2 can be taken as a single state called drought. This new time series of categorical values (variable Y) are the inputs of the Markov Chain models and the time series of cumulative normal distribution function of the DI values are the inputs of the Bayesian Network models. The calibration and validation of the models is performed with leave-one-out cross-validation procedure.

2.3. Markov Chain models

The behavior of a Markov Chain (MC) is governed by a set of probabilities of transitions that specify probabilities for the system being in each of its possible states during the next time period (Wilks 2011). A m^{th} order Markov chain is one where the transition probabilities depend on the states in the previous m time periods. The Markovian property to the m^{th} order Markov Chain is (for more details see Wilks (2011)):

$$P(Y_{tn} | Y_{tn-1}, Y_{tn-2}, Y_{tn-3}, \dots, Y_1) = P(Y_{tn} | Y_{tn-1}, Y_{tn-2}, \dots, Y_{tn-m}) \quad (3)$$

Considering a first-order Markov Chain (MCFO), i.e., $m=1$, the transition probabilities provide the probabilistic forecast for the status of the next step based on the current status, applying the following formula:

$$p_{ij} = P(Y_{tn} = j | Y_{tn-1} = i) \quad (4)$$

where p_{ij} represents the transition probability that Y_{tn} is equal to category j given Y_{tn-1} is equal to category i . The estimate of the transition probabilities (\hat{p}_{ij}) can be calculated to account for the conditional relative frequencies of transitions (f_{ij}):

$$\hat{p}_{ij} = \frac{f_{ij}}{\sum_j f_{ij}} \quad i, j = 1, \dots, s \quad (5)$$

where f_{ij} is the frequency that Y is equal to the category i at time t_{n-1} and equal to category j at the time t_n . The value of s is the number of states of the system. The numerator presents the number of transitions of category i to category j , and the denominator stands for the sum of the number of transitions of category i to any other category.

If $m=2$, we would have a second-order Markov Chain (MCSO), the transition probabilities depend on the states of the current and the previous time period, providing the probabilistic forecast of the status of the next time step. The transition probabilities are calculated as follow:

$$p_{hij} = \Pr\{Y_{tn} = j | Y_{tn-1} = i, Y_{tn-2} = h\} \quad (6)$$

where p_{hij} represent the transition probability that Y_{tn} is equal to category j , given that Y_{tn-1} is equal to the category i and Y_{tn-2} equal to the category h . The transition probability estimates (\hat{p}_{hij}) are obtained from conditional relative frequencies of transition counts (f_{hij}):

$$\hat{p}_{hij} = \frac{f_{hij}}{\sum_j f_{hij}} \quad h, i, j = 1, \dots, s \quad (7)$$

where the numerator is the number of transitions of category h at time t_{n-2} , category i at time t_{n-1} , and category j at the time t_n , and the denominator is the sum of the number of transitions categories h,i to any other category. Thus, the probability of the drought status of the next time period depends on the status of the two previous time periods.

2.4. Bayesian Network models

Heckerman (Heckerman 1998) claims that a Bayesian network for a set of variables $X = \{X_1, \dots, X_n\}$ consists of: (1) a network structure (NS) that encodes a set of conditional independence assertions about variables in X ; and (2) a set P of local probability distributions associated with each variable. Together, these components define the joint probability distribution for X . The NS is a directed acyclic graph (DAG), consisting of the sequence of events or random variables with direct ordering, such as time evolving events (Thulasiraman and Swamy 2011; Madadgar and Moradkhani 2014). Given the NS, the joint probability distribution for X , is given by:

$$P(X) = \prod_{i=1}^n P(X_i | \text{Pai}) \quad (8)$$

where X_i is both the variable and its corresponding node, and Pai are the parents of node X_i in NS. The probabilistic information on the nodes stands for the influence of their parents in the graph on the node. For further details, the reader is referred to Heckerman (1998).

Following a similar mathematical formulation as in Madadgar and Moradkhani (2013; 2014), and given the set of variables X is a unique time-evolving random variable (e.g., drought index), $X = \{X_{t1}, \dots, X_{tn}\}$, where the dependency ordering follows the temporal sequence and the parents of X_{ti} is the set of all prior variables (X_{t1}, \dots, X_{ti-1}), then Equation 8 can be expressed as:

$$P(X_{t1}, \dots, X_{tn}) = \prod_{i=1}^n P(X_{ti} | X_{t1}, \dots, X_{ti-1}) \quad (9)$$

Equation 9 represents the chain rule in probability theory, which can be solved as:

$$P(X_{t1}, \dots, X_{tn}) = P(X_{tn} | X_{t1}, \dots, X_{tn-1}) P(X_{t1}, \dots, X_{tn-1}) \quad (10)$$

Reordering Equation 10 we obtain:

$$P(X_{tn} | X_{t1}, \dots, X_{tn-1}) = \frac{P(X_{t1}, \dots, X_{tn})}{P(X_{t1}, \dots, X_{tn-1})} \quad (11)$$

Equation 11 represents a BN-based model, where the conditional probabilities of the forecast variable (X_{tn}) are calculated given the predictor variables (X_{t1}, \dots, X_{tn-1}). For more details read Todini (2008). The calculation of the joint probability distributions in the right hand side of Equation 11 is a relative cumbersome task, which can be considerably simplified using copulas functions.

2.5. Copulas

Copulas are functions that join multivariate distribution functions whose one-dimensional margins are uniform on the interval (0,1) (Nelsen 2006). Accepting the mathematical definition of copulas in Yan (2006), one obtains a random vector $(U_1, \dots, U_n)^T$, where each margin U_i , $i=1, \dots, n$, is a uniform random variable over the unit interval. Suppose the joint cumulative distribution function (CDF) of $(U_1, \dots, U_n)^T$, is as follow:

$$C(U_1, \dots, U_n) = P(U_1 \leq u_1, \dots, U_n \leq u_n) \quad (12)$$

where the function C is called a n -dimensional copula and CDF specifies the probabilities that the random quantity X will not exceed particular values (Wilks 2011), i.e., $U(X) = U(X \leq x)$. On the other hand, the Sklar's theorem (Sklar 1959) explains the role that copulas play in the relationship between multivariate distribution functions and their univariate margins (Nelsen 2006). The theorem states that a joint multivariate distribution functions $F(X_1, \dots, X_n)$, for all x in the domain of F , can be expressed by a n -dimensional copula, as follows (Yan 2007; Madadgar and Moradkhani 2013):

$$F(X_1, \dots, X_n) = C\{U_1(X_1), \dots, U_n(X_n)\} \quad (13)$$

where $U_i(X_i)$ represents the i^{th} univariate marginal distribution on the unit interval $[0,1]$ and C is the cumulative copula distribution function that represents the multivariate dependence structure (Joe 1997). If U_1, \dots, U_n are all continuous and C is unique (Yan 2006), Equation 11 can be expressed as:

$$P(X_{tn} | X_{t1}, \dots, X_{tn-1}) = \frac{C(U_{t1}, \dots, U_{tn})}{C(U_{t1}, \dots, U_{tn-1})} \quad (14)$$

If $n=2$ (dependence of first order), the conditional probabilities of the forecast variable (X_{t2}) are calculated given the predictor variable (X_{t1}), i.e., the next drought status is conditional to the current status. Similarly, if $n=3$ (dependence of second order), the conditional probabilities of the forecast variable (X_{t3}) are calculated given the predictor variables (X_{t1}, X_{t2}), i.e., the next drought status is conditional to the current status and the status of a previous step. These cases could be called Bayesian Network of first order (BNFO) and second order (BNSO), respectively. Thus, replacing $n=2$ and $n=3$ in Equation 14 leads to the following expressions:

$$P(X_{t2} | X_{t1}) = \frac{C(U_{t1}, U_{t2})}{U_{t1}} \quad (15)$$

$$P(X_{t3} | X_{t1}, X_{t2}) = \frac{C(U_{t1}, U_{t2}, U_{t3})}{C(U_{t1}, U_{t2})} \quad (16)$$

To calculate the probability that DI in the next time step does not exceed the thresholds defined in section 2.2, in which drought states are categorized as drought, mild drought and non-drought, given the availability of the information for the current status for the BNFO model and the states of the current and previous time steps for the BNSO model, the Equations 15 and 16 can be rewritten as:

$$P(X_{t2} \leq x_{d_s} | X_{t1}) = \frac{C[U(X_{t1}), U(X_{t2} \leq x_{d_s})]}{U(X_{t1})} \quad (17)$$

$$P(X_{t3} \leq x_{d_s} | X_{t1}, X_{t2}) = \frac{C[U(X_{t1}), U(X_{t2}), U(X_{t3} \leq x_{d_s})]}{C[U(X_{t1}), U(X_{t2})]} \quad (18)$$

where x_{d_s} is the drought index that causes a drought status according to the threshold defined in section 2.2 ($x_{d_0} = 0$ and $x_{d_1} = -1$). Applying the Equations 17 y 18, the probabilities of having a category 0 (no drought), i.e., $DI > 0$, will be equal to:

$$P(X_{t2} > x_{d_0} | X_{t1}) = 1 - \frac{C[U(X_{t1}), U(X_{t2} \leq 0)]}{U(X_{t1})} \quad (19)$$

$$P(X_{t3} > x_{d_0} | X_{t1}, X_{t2}) = 1 - \frac{C[U(X_{t1}), U(X_{t2}), U(X_{t3} \leq 0)]}{C[U(X_{t1}), U(X_{t2})]} \quad (20)$$

Analogous, the probabilities of having a category 1 (mild drought), i.e., $-1 < DI \leq 0$, can be formulated as:

$$P(x_{d_1} < X_{t2} \leq x_{d_0} | X_{t1}) = \frac{C[U(X_{t1}), U(X_{t2} \leq 0)]}{U(X_{t1})} - \frac{C[U(X_{t1}), U(X_{t2} \leq -1)]}{U(X_{t1})} \quad (21)$$

$$\begin{aligned} P(x_{d_1} < X_{t3} \leq x_{d_0} | X_{t1}, X_{t2}) \\ = \frac{C[U(X_{t1}), U(X_{t2}), U(X_{t3} \leq 0)]}{C[U(X_{t1}), U(X_{t2})]} - \frac{C[U(X_{t1}), U(X_{t2}), U(X_{t3} \leq -1)]}{C[U(X_{t1}), U(X_{t2})]} \end{aligned} \quad (22)$$

Finally, the probabilities of having a category 2 (drought), i.e., $DI \leq -1$, can be expressed as:

$$P(X_{t2} \leq x_{d_1} | X_{t1}) = \frac{C[U(X_{t1}), U(X_{t2} \leq -1)]}{U(X_{t1})} \quad (23)$$

$$P(X_{t3} \leq x_{d_1} | X_{t1}, X_{t2}) = \frac{C[U(X_{t1}), U(X_{t2}), U(X_{t3} \leq -1)]}{C[U(X_{t1}), U(X_{t2})]} \quad (24)$$

2.6. Copulas fitting

The procedure for constructing joint distributions includes: (1) identifying marginal distributions on the unit interval $[0,1]$; (2) selecting suitable dependence structures; and (3) forming joint distributions (Kao and Govindaraju 2010). The marginal distributions (U_i) of step 1 can be derived when the DI values are transformed to a cumulative normal distribution function, in accordance to section 2.2. Since droughts are slowly evolving phenomena, strong temporal autocorrelation among DI is expected (Kao and Govindaraju 2010). Steps 2 and 3

consist in modeling the temporal dependence structure by fitting the copulas functions given the marginal distribution.

Two Elliptical copulas (normal and t) and two Archimedean copulas (Clayton and Frank) were tested in this study, as to define the most appropriate function for the joint distributions. The Canonical Maximum Likelihood (CML) (Cherubini et al. 2004) was used for the estimation of the parameters in the copulas functions, it uses the empirical cumulative distribution function of each marginal distribution (U_i) to transform the observations into pseudo-observations on the unit interval $[0,1]$ (Yan 2007). The best fitted copulas function was identified by the parametric bootstrap-based goodness-of-fit test (Genest and Rémillard 2008), which consists in comparing a distance (ΔC) between the empirical copula (C_E) and the estimated parametric copula (C_θ) function under the null hypothesis that $C_E \in C_\theta$. The latter is evaluated by the p-value; if the p value is greater than the significance level α the null hypothesis is accepted, conversely the null hypothesis is rejected. P values can be obtained via the Monte Carlo method embedded in a parametric bootstrapping procedure.

The Cramér-von-Mises statistic (S) (Anderson 1962) was applied on the group of copulas that are in agreement with the null hypothesis, as to derive the best alternative among the fitted copulas. In the group of copulas that are greater than the significance level, the smallest S is chosen as the best copula function (Madadgar and Moradkhani 2013). The expression of the statistic is as follow, where d is the sample size:

$$S = \int \Delta C (U)^2 dC(U) = \sum_{i=1}^d \{C_E(U_i) - C_\theta(U_i)\}^2 \quad (25)$$

2.7. Forecast verification

The quality of forecasts is assessed using cross-validation procedure. RPSS is used to provide some overall measure of forecast performance. Despite RPSS is a skill score appropriate for evaluating probabilistic forecasts of categorical variables, it is used in this study to verify the quality of the forecasts of both models, bearing in mind that the results of the Bayesian network model should be divided by categories before the forecast verification.

The Ranked Probability Score (RPS) is a skill measurement which penalizes forecast errors in terms of the probability assigned to the events (Zhang and Casey 2000), and is based on the square error between the cumulative probabilities of forecasts and observations. This

score is sensitive to distance, i.e., it includes a penalty for the forecasts that are further away of the observations. According to Wilks (2011) is RPS defined as:

$$RPS = \sum_{m=1}^s (Y_m - O_m)^2 \quad (26)$$

where Y_m is the cumulative probabilities of forecasts, O_m is the cumulative probabilities of observations (O_m) and s the number of event categories or states of the system. As RPS is seen as the probabilistic extension of the square error, a perfect forecast would have a RPS value equal to 0 and the worst forecasts would be very much different from zero.

When we want evaluate the useful of the forecasts, we could use the RPSS, it refers to the relative accuracy of a set of forecasts, with respect to some set of reference forecasts, it can be interpreted as a percentage improvement over the reference forecast (Wilks 2011) . The RPSS is computed as:

$$RPSS = 1 - \frac{\langle RPS \rangle}{\langle RPS_{Clim} \rangle} \quad (27)$$

Where $\langle RPS \rangle$ is the average of the RPS values for each forecast-observation record and $\langle RPS_{Clim} \rangle$ is the average of the RPS values computed with reference forecasts and the observations. In this study the reference forecast is the climatological relative frequencies of the predictand. If the RPSS is equal to 1, indicating a 100% of improvement over the reference forecasts, while if the RPSS is less than or equal to 0, indicating no improvement over the reference forecast.

3. Results and Discussion

3.1 Drought Index

Applying PCA for each month of the year, using information of the correlation matrices of the ten time series (PR1, PR3, PR6, PR9, PR12, VS1, VS3, VS6, VS9 and VS12), the eigenvalues and eigenvectors were obtained. Table 1 shows the eigenvalues and the explained variance of the PC1 for each month.

Table 1. Eigenvalues and explained variance of the PC1

PC1	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Eigenvalues	6.21	7.26	7.89	7.13	7.64	7.08	7.10	7.16	7.08	6.77	5.86	5.71
Explained variance	62%	73%	79%	71%	76%	71%	71%	72%	71%	68%	59%	57%

By means of the Equations 1 and 2, and using information from the eigenvectors and the standardized original data, the DI values were derived for each month and rearranged in chronological order, forming the DI time series for the period 1971 to 2010. The values of the DI were converted to the cumulative normal distribution function that serve as input for the Bayesian Network models, while that the Markov Chain models require categorized time series of the DI according to the thresholds listed in section 2.2. Table 2 shows the frequencies of each category (drought states) for the 480 months of the categorical time series, the Figure 3 shows the time series of the DI and the Figure 4 shows an example of the correlation between the time series of the DI of the months of July, August and September by the graphics of scatterplots. The models were fitted using leave-one-out cross-validation, this procedure allows forecast and assess their accuracy with respect of all events between 1971 and 2010.

Table 2. Frequency of drought state categories

Category	Drought state	Frequency
0	no drought	218
1	mild drought	185
2	drought	77

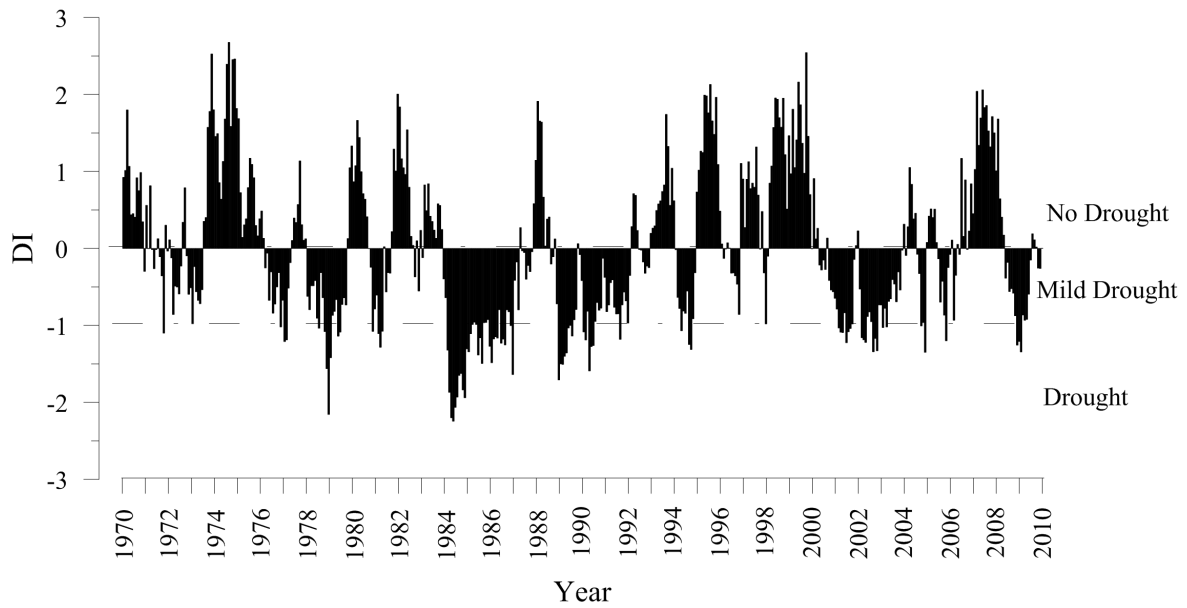


Figure 3. Time series of the DI (1971-2010)

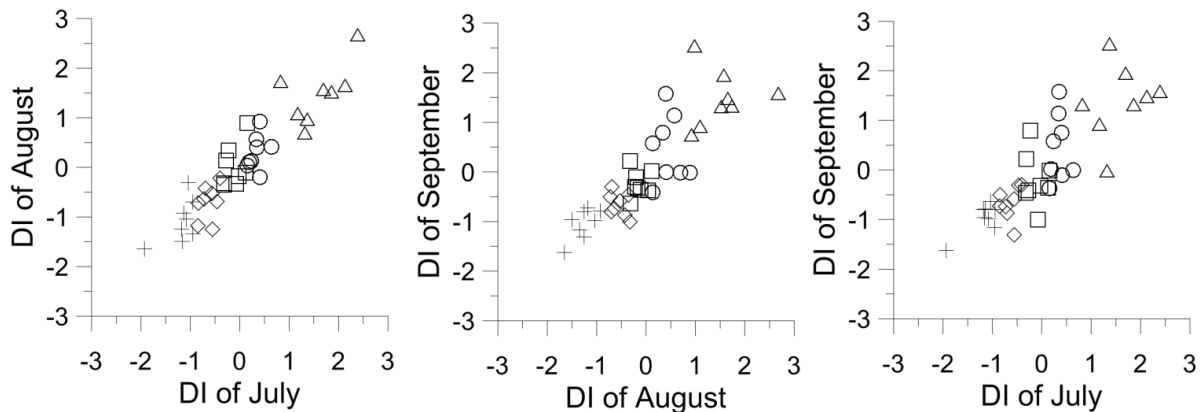


Figure 4. Scatter plots between the time series of the DI of July, August and September (1971-2010)

3.2 Markov Chain Models

In order to take into account the seasonality, the non-homogeneous model version of MC was used, i.e., the calculation of transition probabilities was performed for each month of the year. Therefore twelve matrices of transition probabilities estimation for each model (MCFO and MCSO) were obtained. These matrices offer the possibility of obtaining probabilistic forecasts of the status of next month j given the current status in month i by applying the MCFO model, and given the states in the current month i and previous month h applying the

MCSO model. The transition probability matrix leaving the third year out in cross-validation procedure for August (month with least rain) using the MCFO model is shown in Table 3 and the transition probability matrix for July-August applying the MCSO model is listed in Table 4. For example, if in August the drought status corresponds to category 1 (mild drought), then according to Table 3 the probabilities drought forecasting for the month of September are: for category 0 (no drought) 7%, for category 1 (mild drought) 86% and for category 2 (drought) 7%, adding gives 100%. Similarly, the probabilities drought forecasting for September using the information in Table 4 when the states of July and August are category 0 (no-drought), are: for category 0 (no drought) 75%, for category 1 (mild drought) 25% and for category 2 (drought) 0%, adding gives 100%. When there were no cases of transitions between states, the transition probability for each state is equally likely equal to one divided by the number of categories, e.g. the states 0-2 and 2-0 of the table 4 have a probability equal to 0.3333 for each category.

Table 3. Transition probability matrix of the MCFO model leaving the third year out in cross-validation for the month August

State i	State j		
	0	1	2
0	0.72	0.28	0.00
1	0.07	0.86	0.07
2	0.00	0.57	0.43

Table 4. Transition probability matrix of the MCSO model leaving the third year out in cross-validation for the period July-August

States h-i	State j		
	0	1	2
0-0	0.75	0.25	0.00
0-1	0.00	1.00	0.00
0-2	0.33	0.33	0.33

States h-i	State j		
	0	1	2
1-0	0.50	0.50	0.00
1-1	0.09	0.82	0.09
1-2	0.00	0.33	0.67
2-0	0.33	0.33	0.33
2-1	0.00	1.00	0.00
2-2	0.00	0.75	0.25

3.3 Bayesian Network Models

The BN models consider seasonality and copulas fittings were developed for each month. For testing the null hypothesis the significance level α was taken equal to 0.05 and the number of bootstrap in the Monte Carlo method equal to 1000. For example, the tables 5 and 6 depict the P-values and the S-statistic leaving the third year out in cross-validation of the BNFO and BNSO models, respectively, for each month and the four types of copulas (normal, t, Clayton and Frank). The four copula types were tested since the symmetrical dependence between the most of the monthly time series of DI and the stronger dependence in the lower tail between few monthly time series of DI as is showing in the Figure 4.

Table 5. P-value and S-statistic leaving the third year out in cross-validation for the BNFO model

P-value												
Copulas	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Normal	0.292	0.233	0.077	0.329	0.368	0.527	0.369	0.727	0.189	0.222	0.524	0.664
t	0.057	0.108	0.066	0.358	0.338	0.313	0.464	0.432	0.127	0.322	0.326	0.326
Clayton	0.007	0.022	0.052	0.022	0.009	0.034	0.005	0.020	0.004	0.002	0.021	0.017
Frank	0.363	0.910	0.854	0.534	0.535	0.565	0.804	0.864	0.207	0.170	0.426	0.193

S-statistic												
Copulas	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Normal	0.028	0.027	0.033	0.024	0.023	0.021	0.022	0.018	0.029	0.029	0.022	0.020
t	0.040	0.033	0.034	0.024	0.024	0.025	0.022	0.024	0.034	0.028	0.027	0.027
Clayton	-	-	0.049	-	-	-	-	-	-	-	-	-
Frank	0.027	0.018	0.018	0.023	0.022	0.023	0.019	0.019	0.030	0.034	0.026	0.031

Table 6. P-value and S-statistic leaving the third year out in cross-validation for the BNSO model

P-value												
Copulas	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Normal	0.083	0.164	0.103	0.582	0.698	0.651	0.400	0.276	0.459	0.542	0.433	0.437
t	0.007	0.079	0.087	0.406	0.523	0.254	0.236	0.110	0.305	0.294	0.064	0.088
Clayton	0.001	0.006	0.009	0.010	0.005	0.008	0.004	0.000	0.001	0.000	0.000	0.000
Frank	0.314	0.718	0.648	0.588	0.661	0.605	0.865	0.136	0.108	0.068	0.074	0.259

S-statistic												
Copulas	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Normal	0.047	0.038	0.042	0.027	0.026	0.026	0.031	0.035	0.032	0.030	0.032	0.032
t	-	0.048	0.046	0.033	0.031	0.036	0.039	0.046	0.038	0.039	0.051	0.049
Clayton	-	-	-	-	-	-	-	-	-	-	-	-
Frank	0.045	0.032	0.034	0.035	0.033	0.034	0.028	0.055	0.061	0.064	0.063	0.049

The shaded P-values indicate the copulas that reject the null hypothesis and the shaded S-statistics highlight the copulas with the best fit for each month. The latter copulas were chosen to calculate the conditional probabilities of the drought status for next month given the drought status of the current month for the BNFO model and the current and previous

drought states for the BNSO model, applying the Equations 19-26. Figure 5 shows the probability distributions leaving the third year out in cross-validation of the drought forecasting for the month of September given the drought index of the month of August for the BNFO model. For example, if there is a DI of -1.65 in August corresponding to a category 2 (drought), the probabilities drought forecasting for the month of September are: for category 0 (no drought) 0%, for category 1 (mild drought) 5% and for category 2 (drought) 95%. In a similar way, if there is a DI of 2.68 in August corresponding to a category 0 (no drought), the probabilities drought forecasting for the month September are: for category 0 (no drought) 50%, for category 1 (mild drought) 34%, and for category 2 (drought) 16%. In brief, a low DI value in August leads to higher probabilities of having more severe droughts in September, conversely high values of DI in August leads to low probabilities of drought in September.

A similar analysis can be performed for the BNSO model, with the difference that for the forecast of next month one has to take into consideration the information of the current and the previous month. Despite that the results of the BN models are continuous variables; the division for categories of the resulting time series results into a discrete time series, comparable with the forecast results of the MC models.

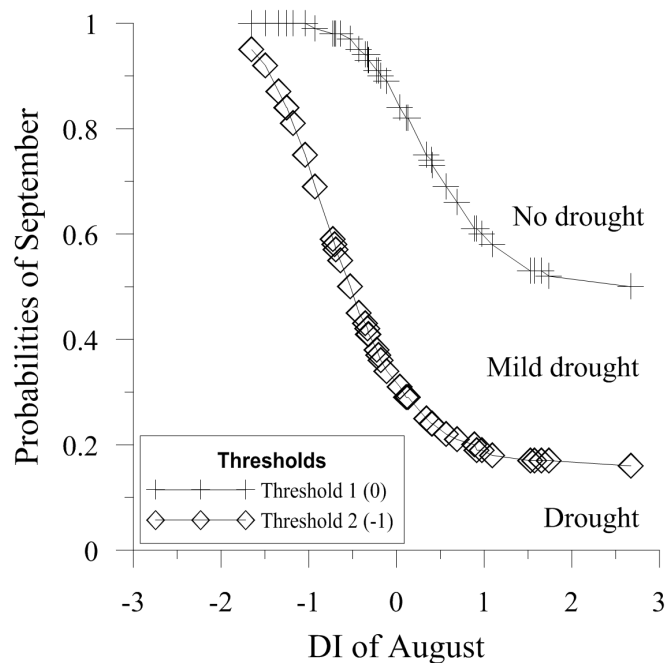


Figure 5. Probabilities of drought forecasts leaving the third year out in cross-validation using the BNFO model for the month of September given the DI of August corresponding respectively to the drought threshold values 0 and -1

3.4 Forecast Verification

Cross-validation was used to evaluate the probabilistic forecasts. Forty RPS values were obtained for each month and each model after to perform leave-one-out procedure. The RPS_{Clim} was calculated by reference forecasts equal to climatological relative frequencies of the time series of categorical values of the DI. Considering the average of the RPS and RPS_{Clim} values for the entire time series, i.e., taking into account all states of drought (drought, mild drought and no-drought) and all months, and applying the equation 27 for calculate the RPSS values, the MCFO model performed better compared to the other models with the greatest value of improvement over reference forecast equal to 0.29, followed by the MCSO and BNFO models with values of RPSS equal to 0.21 and -0.05 (negative RPSS values mean that the reference forecasts are better than the model tested), respectively. However, considering only the observed drought states (drought and mild drought), the BNFO model performed better with the greatest values of RPSS equal to 0.40, followed by the BNSO and MCFO models with RPSS values equal to 0.33 and 0.19, respectively. Even considering only the most severe drought status, the BNFO model yielded better results, with the greatest RPSS value equal to 0.44, followed by the MCSO and BNSO models with RPSS values equal to 0.37 and 0.35, respectively. These results indicate that for the given case study, the MCFO model performed better for the probabilistic forecast of dry and wet periods, while for the probabilistic forecast of dry periods the BN-based models are a better option.

Table 7 depicts the assessment of the monthly drought forecasts, with in the top section of the table (a) the RPSS values of all observed states (drought, mild drought and no-drought) and in the bottom section (b) the RPSS values for the states mild drought and drought. On the basis of the RPSS values in the top section of Table 8 (a) it can be concluded that the MCFO model produces better forecasts with respect to reference forecasts in the months of January, February, March, May, June, September, October and November, while in the months of April, July and August the MCSO performed better. The RPSS values in the bottom section of Table 7 (b) reveal that the BNFO model produces better forecasts with respect to reference forecasts in the months January, February, April, May, June, October, November and December; the MCFO model performed better in March and September; and the MCSO model performs best in July and August.

Table 7. RPSS values of the monthly assessment of drought forecasts

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
(a) No Drought, Mild Drought and Drought												
MCFO	-0.40	0.27	0.63	0.52	0.37	0.41	0.48	0.45	0.33	0.20	0.15	0.15
MCSO	-0.46	-0.32	0.63	0.58	0.25	0.32	0.50	0.52	0.18	0.19	0.05	0.16
BNFO	-0.50	0.13	0.09	0.06	-0.20	0.08	0.04	0.00	-0.11	-0.15	0.02	-0.09
BNSO	-0.60	-0.33	-0.11	-0.08	-0.29	-0.08	-0.14	-0.13	-0.30	-0.33	-0.09	-0.27
(b) Mild Drought and Drought												
MCFO	-0.35	0.12	0.53	0.49	0.27	0.19	0.42	0.31	0.04	0.12	0.09	0.18
MCSO	-0.28	-0.28	0.53	0.53	0.15	0.07	0.48	0.47	-0.14	0.10	-0.03	0.16
BNFO	0.29	0.53	0.48	0.59	0.29	0.38	0.47	0.47	-0.07	0.32	0.45	0.62
BNSO	0.28	0.50	0.37	0.53	0.23	0.20	0.38	0.39	-0.27	0.18	0.42	0.53

Whereas global assessment of drought forecasts permits evaluation of the overall performance of the models, the assessment of drought forecasts on a monthly basis might reveal which of the models performs best for a given month of the hydrologic year. As for example our study disclosed that the MC-based models also better perform in predicting dry periods in specific months. Despite our study used the monthly scale for drought forecasting, findings are in agreement with the results obtained by Madadgar and Moradkhani (2013); it is drought forecasts using the BN-based models did not significantly differ from the forecast of other models. Those authors (2014) also stated that the BN model in combination with copulas functions is a useful procedure in the probabilistic forecasting of drought events.

4. Conclusions

The performance of the MC- and BN-based models in drought forecasting was tested. Both conditional based approaches offer the ability to derive probabilistic forecasts based on information of previous droughts. Copulas functions were used to simplify the solution of the mathematical formulation of the Bayesian network, enabling the derivation of joint distribution functions. Using a recently developed drought index, both modeling approaches

were applied and tested using 40-year rainfall and streamflow data of the Chulco River basin; an Andean regulated river basin in southern Ecuador. The leave-one-out cross-validation procedure was used for forecast verification, it shows that BN-based models were slightly better in forecasting severe drought events than the MC-based models; however the monthly verification of forecasts suggests that the four models distinctively perform better for given drought categories in given months of the hydrological year. The tested approaches can be applied to other spatial and temporal scales, and serve other purposes as simulation, characterization, classification and evaluation of drought events for decision making. It is evident that monthly probabilistic forecasting will help water managers in the development of appropriate mitigation measures.

Acknowledgments

This work has been supported by the University of Cuenca and the Public Municipal Company of Water Supply from Cuenca (ETAPA) through the projects “Meteorological Cycles and Evapotranspiration along the Altitudinal Gradient of the Cajas National Park” and “Identification of hydro-meteorological processes that trigger extreme floods in the city of Cuenca using precipitation radar”. The authors are grateful to INAMHI and CBRM for providing the data. We would like to thank Prof. Jan Feyen for his valuable comments and suggestions that improved the final version of the manuscript and a recognition to Lenin Camposano who helped develop the cross-validation procedure.

Author Contributions

A.A. conceived the initial idea of the study, constructed the models and prepared the manuscript. R.C. participated in the design of the study, and provided a critical revision of manuscript. A.S. and J.P. provided many useful suggestions to improve the analysis of the results and supervised the research.

Conflicts of Interest

The authors declare no conflict of interest.

References

- AghaKouchak A (2014) A baseline probabilistic drought forecasting framework using standardized soil moisture index: application to the 2012 United States drought. *Hydrol Earth Syst Sci* 18:2485–2492. doi: 10.5194/hess-18-2485-2014
- Al-Faraj F, Scholz M, Tigkas D (2014) Sensitivity of Surface Runoff to Drought and Climate Change: Application for Shared River Basins. *Water* 6:3033–3048. doi: 10.3390/w6103033
- Anderson ML, Mierzwa MD, Kavvas ML (2000) Probabilistic seasonal forecasts of droughts with a simplified coupled hydrologic-atmospheric model for water resources planning. *Stoch Environ Res Risk Assess* 14:263–274. doi: 10.1007/s004770000049
- Anderson T (1962) On the distribution of the two-sample Cramer-von Mises criterion. *Ann Math Stat* 22:1148–1159. doi: 10.1214/aoms/1177704477.
- Araghinejad S (2010) An Approach for Probabilistic Hydrological Drought Forecasting. *Water Resour Manag* 25:191–200. doi: 10.1007/s11269-010-9694-9
- Avilés A, Célleri R, Paredes J, Solera A (2015) Evaluation of Markov Chain Based Drought Forecasts in an Andean Regulated River Basin Using the Skill Scores RPS and GMSS. *Water Resour Manag* 29:1949–1963. doi: 10.1007/s11269-015-0921-2
- Beniston M (2003) Climatic Change in Mountain Regions: A Review of Possible Impacts. *Clim Chang* 59:5–31.
- Buytaert W, Célleri R, De Bièvre B, et al. (2006a) Human impact on the hydrology of the Andean páramos. *Earth-Science Rev* 79:53–72. doi: 10.1016/j.earscirev.2006.06.002
- Buytaert W, Celleri R, Willems P (2006b) Spatial and temporal rainfall variability in mountainous areas: A case study from the south Ecuadorian Andes. *J Hydrol* 329:413–421. doi: 10.1016/j.jhydrol.2006.02.031
- Buytaert W, Vuille M, Dewulf A, et al. (2010) Uncertainties in climate change projections and regional downscaling in the tropical Andes: implications for water resources management. *Hydrol Earth Syst Sci* 14:1247–1258. doi: 10.5194/hess-14-1247-2010
- Cancelliere A, Mauro G Di, Bonaccorso B, Rossi G (2007) Drought forecasting using the Standardized Precipitation Index. *Water Resour Manag* 21:801–819. doi: 10.1007/s11269-006-9062-y
- Celleri R, Willems P, Buytaert W, Feyen J (2007) Space – time rainfall variability in the

- Paute Basin , Ecuadorian Andes. *Hydrol Process* 21:3316–3327. doi: 10.1002/hyp.6575
- Cherubini U, Luciano E, Vecchiato W (2004) *Copula methods in finance*. John Wiley & Sons
- Chung C, Salas JD (2000) Drought Occurrence Probabilities and Risks of Dependent Hydrologic Processes. *J Hydrol Eng* 5:259–268. doi: 10.1061/(ASCE)1084-0699(2000)5:3(259)
- Durdu ÖF (2010) Application of linear stochastic models for drought forecasting in the Büyük Menderes river basin, western Turkey. *Stoch Environ Res Risk Assess* 24:1145–1162. doi: 10.1007/s00477-010-0366-3
- Embrechts P, Lindskog F, McNeil A (2001) *Modelling dependence with copulas*. Zurich, Switzerland
- Fernández C, Vega J a., Fonturbel T, Jiménez E (2009) Streamflow drought time series forecasting: a case study in a small watershed in North West Spain. *Stoch Environ Res Risk Assess* 23:1063–1070. doi: 10.1007/s00477-008-0277-8
- Genest C, Rémillard B (2008) Validity of the parametric bootstrap for goodness-of-fit testing in semiparametric models. *Ann l'Institut Henri Poincaré - Probab Stat* 44:1096–1127. doi: 10.1214/07-AIHP148
- Han P, Wang PX, Zhang SY, Zhu DH (2010) Drought forecasting based on the remote sensing data using ARIMA models. *Math Comput Model* 51:1398–1403. doi: 10.1016/j.mcm.2009.10.031
- Heckerman D (1998) *A tutorial on learning with Bayesian networks*. Learn. Graph. Model. Springer Netherlands, pp 301–354
- Hwang Y, Carbone GJ (2009) Ensemble Forecasts of Drought Indices Using a Conditional Residual Resampling Technique. *J Appl Meteorol Climatol* 48:1289–1301. doi: 10.1175/2009JAMC2071.1
- Joe H (1997) *Multivariate Models and Multivariate Dependence Concepts*. CRC Press
- Kao S-C, Govindaraju RS (2010) A copula-based joint deficit index for droughts. *J Hydrol* 380:121–134. doi: 10.1016/j.jhydrol.2009.10.029
- Karavitis CA, Alexandris S, Tsesmelis DE, Athanasopoulos G (2011) Application of the Standardized Precipitation Index (SPI) in Greece. *Water* 3:787–805. doi: 10.3390/w3030787
- Kolmogorov A (1933) Sulla determinazione empirica di una legge di distribuzione. *G dell'*

Ist Ital degli Attuari 4:83–91.

Kumar V, Panu U (1997) Predictive assessment of severity of agricultural droughts based on agro-climatic factors. *J Am Water Resour Assoc* 33:1255–1264.

Leilah AA, Al-Khateeb S. (2005) Statistical analysis of wheat yield under drought conditions. *J Arid Environ* 61:483–496. doi: 10.1016/j.jaridenv.2004.10.011

Liu WT, Juárez RIN (2001) ENSO drought onset prediction in northeast Brazil using NDVI. *Int J Remote Sens* 22:3483–3501. doi: 10.1080/01431160010006430

Lohani VK, Loganathan G V (1997) An early warning system for drought management using the Palmer drought index. *J Am Water Resour Assoc* 33:1375–1386.

Madadgar S, Moradkhani H (2013) A Bayesian framework for probabilistic seasonal drought forecasting. *J Hydrometeorol* 14:1685–1706. doi: 10.1175/JHM-D-13-010.1

Madadgar S, Moradkhani H (2014) Spatio-temporal drought forecasting within Bayesian networks. *J Hydrol* 512:134–146. doi: 10.1016/j.jhydrol.2014.02.039

Mason SJ (2004) On Using “ Climatology ” as a Reference Strategy in the Brier and Ranked Probability Skill Scores. *Mon Weather Rev* 132:1891–1895.

Mishra a. K, Desai VR (2005) Drought forecasting using stochastic models. *Stoch Environ Res Risk Assess* 19:326–339. doi: 10.1007/s00477-005-0238-4

Mishra a. K, Desai VR (2006) Drought forecasting using feed-forward recursive neural network. *Ecol Modell* 198:127–138. doi: 10.1016/j.ecolmodel.2006.04.017

Mishra AK, Desai VR, Singh VP, Asce F (2007) Drought Forecasting Using a Hybrid Stochastic and Neural Network Model. *J Hydrol Eng* 12:626–638. doi: 10.1061/(ASCE)1084-0699(2007)12:6(626)

Mishra AK, Singh VP (2010) Review paper A review of drought concepts. *J Hydrol* 391:202–216. doi: 10.1016/j.jhydrol.2010.07.012

Mishra AK, Singh VP (2011) Drought modeling – A review. *J Hydrol* 403:157–175. doi: 10.1016/j.jhydrol.2011.03.049

Modarres R (2007) Streamflow drought time series forecasting. *Stoch Environ Res Risk Assess* 21:223–233. doi: 10.1007/s00477-006-0058-1

Morid S, Smakhtin V, Bagherzadeh K (2007) Drought forecasting using artificial neural

networks and time series of drought indices. *Int J Climatol* 27:2103–2111. doi: 10.1002/joc.1498

Murphy A (1993) What is a good forecast? An essay on the nature of goodness in weather forecasting. *Weather Forecast*. 8:

Murphy AH (1977) The Value of Climatological, Categorical and Probabilistic Forecasts in the Cost-Loss Ratio Situation. *Mon Weather Rev* 105:803–816. doi: 10.1175/1520-0493(1977)105<0803:TVOCCA>2.0.CO;2

Nalbantis I, Tsakiris G (2009) Assessment of hydrological drought revisited. *Water Resour Manag* 23:881–897. doi: 10.1007/s11269-008-9305-1

Nelsen RB (2006) *An Introduction to Copulas*, Second Edi. Springer, USA

Paulo A a., Pereira LS (2007) Prediction of SPI drought class transitions using Markov chains. *Water Resour Manag* 21:1813–1827. doi: 10.1007/s11269-006-9129-9

Pearl J (1985) Bayesian networks: A model of self-activated memory for evidential reasoning. *Proc. 7th Conf. Cogn. Sci. Soc. Univ. California, Irvine, CA*. pp 329–334

Sánchez-Chóliz J, Sarasa C (2015) River Flows in the Ebro Basin: A Century of Evolution, 1913–2013. *Water* 7:3072–3082. doi: 10.3390/w7063072

Sklar A (1959) Fonctions de répartition à n dimensions et leurs marges. *Publ Inst Stat Univ Paris* 8:229–231.

Staudinger M, Stahl K, Seibert J (2014) A drought index accounting for snow. *Water Resour Res* 50:7861–7872. doi: 10.1002/2013WR015143

Steinemann A (2003) Drought indicators and triggers: A stochastic approach to evaluation. *J Am Water Resour Assoc* 39:1217–1233.

Thulasiraman K, Swamy MNS (2011) *Graphs: Theory and Algorithms*. John Wiley & Sons

Todini E (2008) A model conditional processor to assess predictive uncertainty in flood forecasting. *Int J River Basin Manag* 6:123–137.

Vicente-Serrano SM, Chura O, López-Moreno JI, et al. (2015) Spatio-temporal variability of droughts in Bolivia: 1955-2012. *Int J Climatol* 35:3024–3040. doi: 10.1002/joc.4190

Viviroli D, Archer DR, Buytaert W, et al. (2011) Climate change and mountain water resources: overview and recommendations for research, management and policy. *Hydrol*

Earth Syst Sci 15:471–504. doi: 10.5194/hess-15-471-2011

Wilks DS (2011) *Statistical Methods in the Atmospheric Sciences*, Third Edit. Academic Press, USA

World Meteorological Organization (2012) *Standardized Precipitation Index User Guide*.

Yan J (2007) Enjoy the Joy of Copulas: With a Package Copula. *J Stat Softw* 21:1–21.

Yan J (2006) Multivariate modeling with copulas and engineering applications. In: Pham H (ed) *Springer Handb. Eng. Stat.* Springer London, London, pp 973–990

Yuan X, Wood EF, Chaney NW, et al. (2013) Probabilistic Seasonal Forecasting of African Drought by Dynamical Models. *J Hydrometeorol* 14:1706–1720. doi: 10.1175/JHM-D-13-054.1

Zhang H, Casey T (2000) Verification of categorical probability forecasts. *Weather Forecast* 15:80–89.

Zhong S, Shen L, Sha J, et al. (2015) Assessing the Water Parallel Pricing System against Drought in China: A Study Based on a CGE Model with Multi-Provincial Irrigation Water. *Water* 7:3431–3465. doi: 10.3390/w7073431

© 2015 by the authors; licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).

ANEXO 3: Estimación del riesgo de fallo en el suministro de agua como ayuda a la planificación y gestión de recursos hídricos³

RESUMEN

En este estudio modelos numéricos y escenarios estocásticos se utilizaron para analizar si sistemas de recursos hídricos en la cuenca del río Paute bajo ciertas condiciones de infraestructura (embalses de almacenamiento) son capaces de satisfacer demandas de agua en los siguientes 30 años. Los análisis revelaron la necesidad de implementar un embalse con una capacidad de al menos de 21 Hm³ en la subcuenca del río Tomebamba, mientras que en la subcuenca del río Pindilig no se requerirá ningún embalse. En caso de que el suministro de agua durante los periodos secos esté por debajo de la demanda de agua podría requerirse la aplicación de medidas de restricción en el uso del agua. Los resultados muestran claramente la utilidad del enfoque utilizado y su capacidad para generar alternativas en la oferta y demanda de agua en condiciones futuras. Como tal, el método será muy útil para los tomadores de decisiones en la definición de las políticas más adecuadas para la planificación y gestión del recurso hídrico.

Palabras clave: Recursos hídricos, modelación numérica, escenarios estocásticos, tomadores de decisiones, gestión y planificación.

ABSTRACT

In this study numerical models and stochastic scenarios are used to verify if the water resource systems of the Paute River Basin under the given conditions of infrastructure (storage reservoirs) are capable of meeting the water demand in 30 years from today. The analysis revealed the need to implement a reservoir with a capacity of at least 21 Hm³ in the Tomebamba subbasin, whereas the Pindilig subbasin does not require any storage reservoir. In case the water supply during dry periods ceases below the water demand the implementation of water use restriction measures might be required. The results clearly show the usefulness of used approach and its capacity in generating alternatives in water supply and demand for future conditions. As such the method will be very helpful to decision-

³ Avilés A, Solera A (2012) Estimación del riesgo de fallo en el suministro de agua como ayuda a la planificación y gestión de recursos hídricos. Maskana 3:77–86. La publicación final esta disponible en <http://www.ucuenca.ec/ojs/index.php/maskana/article/view/403>

makers in defining the most appropriate policies for the planning and management of the water resources.

Keywords: Water resources, numerical modeling, stochastic scenarios, decision-making, management, and planning.

1. INTRODUCCIÓN

Los sistemas de recursos hídricos son una mezcla de muchos aspectos ambientales, sociales, hidrológicos, económicos y administrativos, su administración requiere amplias e integradas toma de decisiones, que requieren información de las diferentes necesidades de los usuarios y los conflictivos intereses económicos y sociales por acceder al agua (Alemu y col., 2011). Variaciones espaciales y temporales en los componentes del sistema dan cuenta que son funciones de muchos factores estocásticos, es por esto que los análisis de los sistemas van más allá de los métodos convencionales deterministas (Huang y col., 2010). Durante la explotación de los sistemas de recursos hídricos, pueden darse inevitablemente situaciones en las que no sea posible suministrar de forma satisfactoria los usos establecidos (fallo del sistema, denominado también sequía operativa), esta situación puede minimizarse mediante una anticipación de la sequía en el sistema en las fases de planificación y explotación, minimizando los valores esperados de duración y magnitud de sequías, mediante la adopción de reglas de gestión (Sánchez y col., 2000).

Las políticas y reglas de operación de un sistema de recursos hídricos son herramientas fundamentales para la toma de decisiones durante la fase de explotación de un sistema, el uso de una regla u otra, o de ninguna regla, es determinante para el rendimiento del sistema (Solera, 2003). Debe destacarse que no existen reglas o estrategias de gestión universales para el manejo de los recursos hídricos, sin embargo algunas reglas son extrapolables a sistemas con topologías y propósitos similares, proponiendo una amplia gama de experiencias en el manejo de sistemas complejos que pueden utilizarse como ejemplos para adoptar reglas para un sistema particular (Lund y Guzmán, 1999). Dolling (2001) y Ochoa (2002), hacen una descripción de algunas herramientas informáticas destinadas a resolver distintos problemas del área de la hidrología y de la gestión de los sistemas hídricos, en donde se incluyen modelos desarrollados para el diseño, planificación y explotación de un sistema hídrico, estos modelos tienen incorporados modelos matemáticos complejos de

simulación y optimización para los distintos análisis en los sistemas de recursos hídricos, cuyo objetivo final es ayudar a los tomadores de decisiones para identificar políticas o reglas de operación en la planificación y gestión del agua. A continuación se detallan algunos modelos con la finalidad de mostrar las ventajas de utilización de unos u otros.

REZES (Simonovic y Savic, 1989) es un programa diseñado como herramienta para la asistencia al análisis de embalses (dimensionamiento, planificación de operación, operación a tiempo real), en donde están incluidos métodos de optimización (programación determinística, programación estocástica). AQUARIUS (Diaz y Brown, 1997) es un sistema orientado a la distribución temporal y espacial de agua entre los usuarios en una cuenca hidrográfica, utilizando programación no lineal. HEC-5 (Hydrologic Engineering Center, 1998) es un modelo que asigna el agua a las demandas mediante el balance hídrico de una cuenca, teniendo en cuenta las reglas de operación dadas por el usuario, además admite la consideración de múltiples propósitos. MODSIM (Labadie y col., 2000) tiene una estructura de trabajo basado en algoritmos de optimización de programación dinámica, la función objetivo está dada en términos de costos, y los repartos y almacenamientos de agua se hacen según una escala de prioridades dada por el modelador. WEAP (SEI, 2001) es un modelo enfocado a la evaluación y planificación hidrológica de cuencas; dicho modelo trabaja con algoritmos de balance hídrico mediante redes de flujo, en las que tiene la posibilidad de considerar como elementos físicos del sistema, embalses, canales, tramos de ríos, tuberías y demandas, etc. AQUATOOL (Andreu y col., 1996) es un software para el diseño y desarrollo de sistemas de soporte a la decisión (SSD) para gestores de recursos hídricos, que permite, entre otras cosas diseñar e introducir gráficamente la configuración de un sistema de recursos hídricos y sus características físicas y condiciones de gestión, para su posterior optimización y simulación.

Como se puede ver en el párrafo anterior, existe una amplia variedad de modelos que resuelven problemas y necesidades de gestión y planificación de agua mediante modelos de simulación y optimización, no obstante el uso u otro modelo va a depender de los objetivos y recursos disponibles de cada estudio. En el presente documento se utiliza dos modelos SIMGES (Andreu y col., 2010) y SIMRISK (Andreu y col., 2003) que están integrados en el SSD AQUATOOL, son modelos amigables con el usuario y su utilización no representa costos computacionales excesivos. SIMGES realiza la simulación de sistemas de recursos hídricos utilizando una red de flujo interna, basándose en el principio de conservación de masa; el modelo optimiza la red interna con un algoritmo lineal, dando como resultado la

asignación de agua para los distintos usos, según un orden de prioridades de las demandas previamente establecido. En cambio SIMRISK es un modelo muy útil y eficaz para la simulación con escenarios estocásticos, el modelo proporciona al decisor de herramientas para la gestión, como la determinación de las probabilidades de fallos en los sistemas de recursos hídricos y la evaluación de los riesgos de las decisiones en el sistema y las medidas de mitigación según el riesgo admisible impuesto.

Con estos antecedentes, el presente trabajo pretende estimar el riesgo de fallo en el suministro de agua para generar medidas adecuadas en la oferta y la demanda de agua en los próximos 30 años en dos subcuencas del río Paute (Tomebamba y Pindilig) con diferente régimen de precipitación. Para lo cual se recurrió a métodos numéricos de simulación utilizando escenarios estocásticos como entrada.

2. MATERIALES Y MÉTODOS

2.1. Materiales

La Figura 1 muestra un mapa de la cuenca del río Paute, la misma que tiene gran importancia por sus abastecimientos urbanos e hidroeléctricos, sin embargo adolece de una ausencia de infraestructuras de regulación en la mayoría de subcuencas, esto sumado a la variabilidad climática en la cuenca que conlleva a que existan épocas de sequía severas que afectan al suministro de agua de las demandas consuntivas, no consuntivas y ambientales.

Para los modelos numéricos estocásticos para la síntesis de aportaciones se utilizó registros históricos de caudales mensuales de estaciones hidrológicas ubicadas al final de las áreas consideradas naturales de las subcuencas de estudio, es decir, donde no existe una explotación considerable del recurso hídrico. La estación hidrológica Matadero en Sayausí (ubicada en la subcuenca del río Tomebamba) y la estación hidrológica Dudas en Pindilig (ubicada en la subcuenca del río Pindilig) fueron las escogidas para extraer la información de aportaciones mensuales para aplicar los distintos modelos. Estas series también sirvieron de datos de entrada para simular el escenario actual en el modelo SIMGES conjuntamente con las demandas de agua actuales.

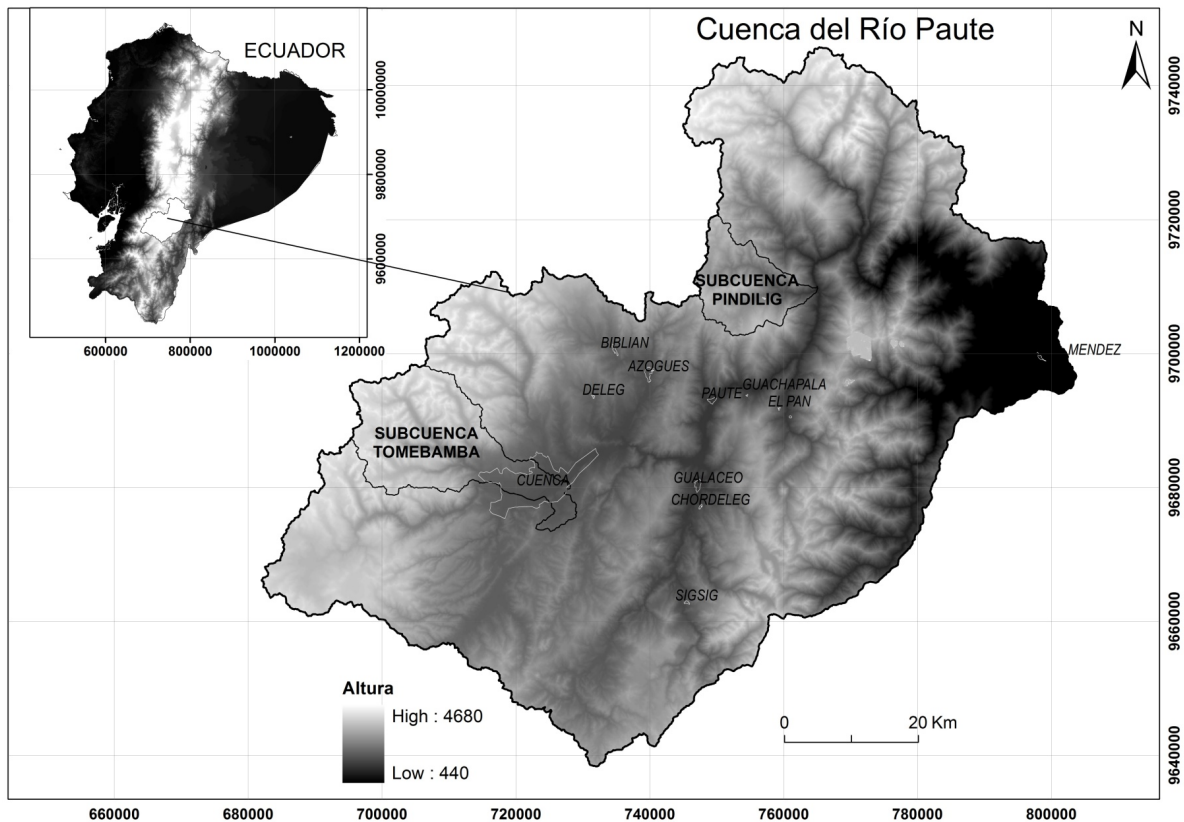


Figura 1. Ubicación del ámbito de estudio.

Para cuantificar las demandas actuales de agua se utilizaron las bases de datos de la Secretaría Nacional del Agua (SENAGUA) y de la Empresa Municipal de Agua Potable de la Ciudad de Cuenca (ETAPA), en donde está expuesta información de los diferentes usos de agua. Estas demandas fueron agrupadas según una prioridad en el uso del recurso (primero el agua para consumo doméstico, luego el agua para riego y al final los otros usos del agua) y según el peso que tienen las demandas que utilizan un mayor volumen de agua, con el propósito de simplificar los esquemas de los sistemas de recursos hídricos construidos para los modelos de simulación. Para atender los requerimientos ambientales de los ecosistemas fluviales, se consideró un caudal ecológico equivalente al 10% de los caudales medios mensuales. El cálculo de la demanda futura se realizó de una manera sencilla, considerando proyecciones lineales de la población, hectáreas de riego, semovientes, etc., y posibles dotaciones futuras según las tendencias potenciales de uso del agua en las zonas de estudio. La demanda futura y las series sintéticas de aportaciones fueron la información de entrada para el modelo SIMRISK.

2.2. Modelos estocásticos de series de tiempo

Los modelos estocásticos se fundamentan principalmente en el máximo aprovechamiento de la información de la serie histórica de una variable determinada, es decir reproducir las estadísticas históricas de las series de tiempo; el concepto clave en este tipo de modelos es la dependencia entre los datos, su aplicación presenta algunas ventajas, por ejemplo tiene mucha flexibilidad para trabajar en diferentes escalas, además son de fácil entendimiento y no requieren altos requerimientos computacionales.

En el presente estudio se aplicaron modelos autoregresivos de media móvil (ARMA) y modelos autoregresivos de media móvil periódicos (PARMA), con diferentes órdenes (p,q), se utilizó como ayuda el software SAMS (Sveinsson *y col.*, 2007), cuyo objetivo fue conseguir un conjunto de modelos que sean capaces de reproducir las propiedades estadísticas históricas de los datos (estadísticos básicos, sequía y almacenamiento) y esencialmente las suposiciones claves de los residuos como la normalidad y la independencia. Finalmente se escogió los mejores modelos con los menores valores del criterio de información de Akaike (1974) corregido (AICC) (Hurvich y Tsai, 1989), el mismo que se basa en una penalización de los modelos con mayor número de parámetros. Estos modelos sirvieron para la generación de series sintéticas anuales (utilizadas para la implementación de infraestructura para la regulación del agua) y la generación de series sintéticas mensuales (utilizadas para la construcción de escenarios futuros para la simulación de la gestión de la infraestructura planificada).

2.3. Simulación de los sistemas de recursos hídricos utilizando SIMGES

El proceso de simulación realiza una optimización de una red de flujo conservativa, mediante el cumplimiento de una función objetivo (satisfacción de las demandas, operación de embalses, etc.) sujeta a las restricciones de conservación de masa (continuidad) y los límites físicos de transporte de flujo en conducciones (ríos, quebradas, etc.) y capacidades de embalses (reservas) y otros elementos si los hubiere.

A partir de las series históricas de caudales y la agrupación de las demandas del agua, se construyeron diseños de cada uno de los sistemas de recursos hídricos para cada subcuenca un ejemplo del esquema construido para la subcuenca del río Pindilig se muestra en la Figura 2.

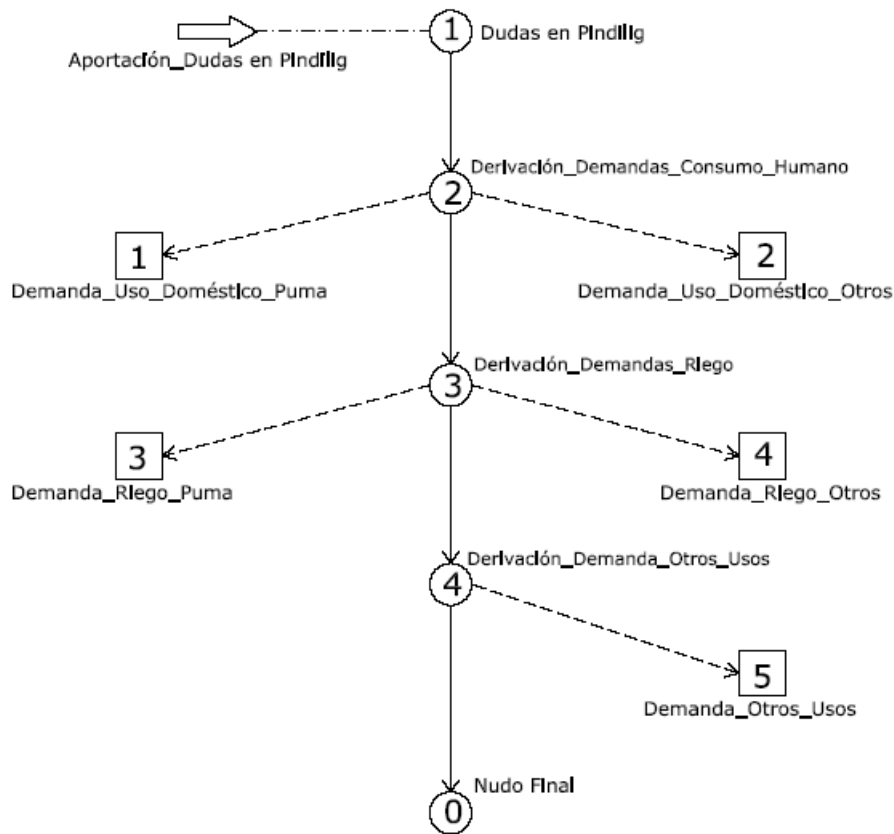


Figura 2. Esquema del sistema de recursos hídricos en la subcuenca del río Pindilig.

Una vez diseñados los esquemas, mediante el modelos SIMGES, se procedió a simular el escenario actual (1979-2009) mediante la introducción de la información de la serie de aportaciones mensuales históricas y la demanda actual, luego el escenario futuro (2010-2040) con la introducción de la demanda futura y la información de la serie de aportaciones mensuales históricas, suponiendo que el régimen natural de flujo de las subcuencas no variará con el tiempo. En caso de un incumplimiento en la satisfacción de las demandas se planificó la implementación de un embalse con una capacidad ideal para la satisfacción total de las demandas, sin embargo, esta capacidad ideal supone que no existirán restricciones físicas, económicas, sociales y ambientales en la construcción de la infraestructura, es por esto que se probó la implementación de un embalse con una capacidad menor, dando como resultado una insatisfacción de las demandas y el inicio de una gestión de los sistemas con un riesgo de desabastecimiento de agua.

2.4. Simulación de la gestión con escenarios estocásticos utilizando SIMRISK

Con la ayuda de SIMRISK se simuló la gestión de 100 escenarios estocásticos, creados por cada serie sintética mensual, tomando en cuenta la implementación de un embalse con una capacidad no ideal, luego se verificó de una manera probabilística la satisfacción de las demandas, determinando el estado del sistema mediante la evaluación de las probabilidades de fallo mensual y la magnitud del déficit en el suministro (dividido en cuatro niveles). Por ejemplo el nivel 1 (n1) es la situación más grave, quiere decir que el déficit supera el 75% de la demanda o sea el suministro se ubica entre el 0 y 25% del valor de la demanda; el nivel 2 (n2) quiere decir que el suministro se ubica entre el 25 y 50% del valor de la demanda; el nivel 3 (n3) quiere decir que el suministro se ubica entre el 50 y 75% del valor de la demanda; y el nivel 4 (n4) es la situación menos grave, quiere decir que el suministro se ubica entre el 75 y 100%. Todo este marco probabilístico sirvió de ayuda para generar alternativas de gestión de las demandas para que el riesgo de fallo sea el menor posible.

3. RESULTADOS

3.1. Modelos estocásticos de series de tiempo

Según el AICC los mejores modelos ARMA para la subcuenca del río Tomebamba fueron el ARMA(1,1) para la escala anual y el AR(1) para la escala mensual; y para la subcuenca del río Pindilig fueron el MA(3) para la escala anual y mensual. De igual forma los mejores modelos PARMA mensuales según el AICC se muestran sombreados en la Tabla 1 para la subcuenca del río Tomebamba y en la Tabla 2 para la subcuenca del río Pindilig. Estos modelos sirvieron para la generación de series sintéticas anuales y mensuales, cuyos estadísticos resultantes fueron similares con los estadísticos de las series históricas. Se escogió trabajar con los modelos PARMA a escala mensual debido a que los estadísticos de sequía y almacenamiento de las series generadas tuvieron una mejor representación de los estadísticos de la muestra histórica.

Tabla 1. Valores AICC de los modelos PARMA en la subcuenca del río Tomebamba.

Mes	PARMA(1,0)	PARMA(1,1)	PARMA(1,2)
ENE	31.39	33.06	34.24
FEB	26.97	28.57	31.75
MAR	28.21	29.58	31.41
ABR	22.30	22.08	25.42

Mes	PARMA(1,0)	PARMA(1,1)	PARMA(1,2)
MAY	28.29	28.45	30.81
JUN	33.91	36.36	37.79
JUL	33.99	32.58	33.72
AGO	30.32	20.82	23.30
SEP	21.75	23.79	26.46
OCT	33.18	35.41	31.66
NOV	35.43	37.87	36.20
DIC	24.73	22.68	27.27

Tabla 2. Valores AICC de los modelos PARMA en la subcuenca del río Pindilig.

Tabla	PARMA(1,0)	PARMA(3,0)	PARMA(0,2)	PARMA(0,3)	PARMA(1,1)
ENE	6.87	11.72	20.77	6.59	9.29
FEB	1.43	3.57	20.10	-1.60	3.88
MAR	14.68	18.48	14.57	23.77	16.47
ABR	21.26	25.72	25.46	22.40	23.32
MAY	19.63	14.12	31.02	20.19	15.29
JUN	20.96	24.75	24.58	26.27	22.57
JUL	15.35	20.14	22.46	22.80	17.70
AGO	11.55	13.64	20.73	18.25	11.15
SEP	-2.61	1.83	7.17	10.13	-0.16
OCT	19.24	23.80	26.39	24.60	21.38
NOV	5.19	8.38	14.96	15.49	6.56
DIC	-11.76	-13.67	-0.49	-3.04	-15.33

3.2. Simulación de los sistemas de recursos hídricos utilizando SIMGES

Los resultados de la simulación de los escenarios actual y futuro en la subcuenca del río Tomebamba revelaron que todas las demandas tienen déficits en algunos meses (incluyendo la demanda ambiental), además no cumplían con la garantía (es decir, no existía satisfacción de las demandas). Por lo tanto se intentó incluir una infraestructura de regulación virtual (embalse), de tal manera que las garantías de las demandas sean cumplidas. En cambio en la subcuenca del río Pindilig los resultados de la simulación de los escenarios actual y futuro revelaron el cumplimiento total de las garantías de las demandas (incluida la demanda ambiental); por lo tanto, en este sistema no haría falta una infraestructura de regulación al menos en los siguientes 30 años (con la información disponible y estimada).

Con las 100 series sintéticas anuales generadas en la estación de caudales de la subcuenca del río Tomebamba y con un nivel de demanda igual a la demanda anual total futura en la subcuenca, se calculó 100 valores de capacidad de almacenamiento, donde 40 valores fueron iguales a cero y 60 valores diferentes de cero. Con un valor medio de los

valores no nulos de capacidad de almacenamiento que corresponde a un valor aproximadamente igual a 21 Hm³ se procedió a simular el sistema de recursos hídricos de la subcuenca del Tomebamba en el escenario futuro, dando como resultado un cumplimiento total de las garantías de las demandas mensuales y anuales. También hay que indicar que la demanda ambiental cumplió satisfactoriamente en todos los meses.

Probando bajar la capacidad del embalse a la mitad aproximadamente (10 Hm³), suponiendo que esta capacidad será factible según unas restricciones físicas, ambientales, sociales y económicas, que se podrían presentar en la construcción de una infraestructura de tal magnitud, se incumpliría nuevamente las garantías de las demandas y daría inicio a una gestión futura del sistema, con el embalse factible planificado, con escenarios estocásticos y con enfoque de riesgo de desabastecimiento de agua a las demandas.

3.3. Simulación de la gestión con escenarios estocásticos utilizando SIMRISK

Para que sean más entendibles los resultados se realizó un promedio mensual de los valores de probabilidad de fallo para cada nivel de suministro (n1, n2, n3 y n4) y para cada mes en el periodo de simulación, es decir por cada mes del año hidrológico se obtendrá una probabilidad de fallo de la satisfacción total de las demandas. Por ejemplo el mes de Febrero es el más desfavorable para el suministro de agua para la ciudad de Cuenca y tiene una probabilidad de fallo total de 6.33% (Figura 3). De igual forma las demandas de riego presentan el mes de Febrero como el más desfavorable para el suministro de agua, con una magnitud de probabilidad de fallo total de 8.47% y una probabilidad de fallo en el nivel 1 de 6.90%, o sea en el nivel con un déficit de 75% de la demanda (Figura 4).

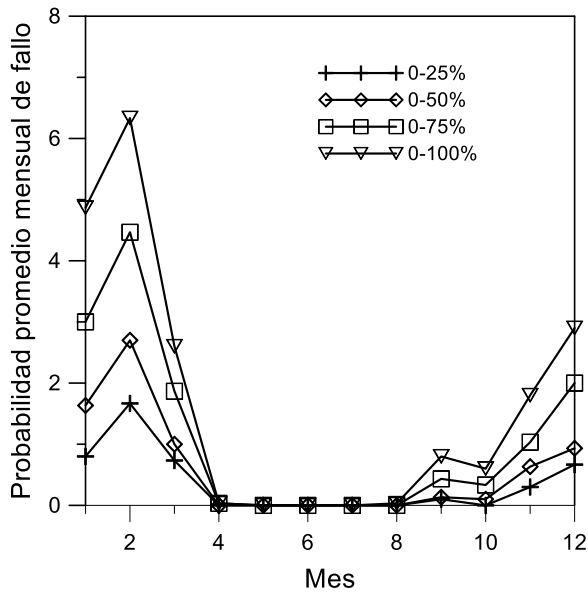


Figura 3. Probabilidad promedio mensual de fallo de la demanda futura de la ciudad de Cuenca.

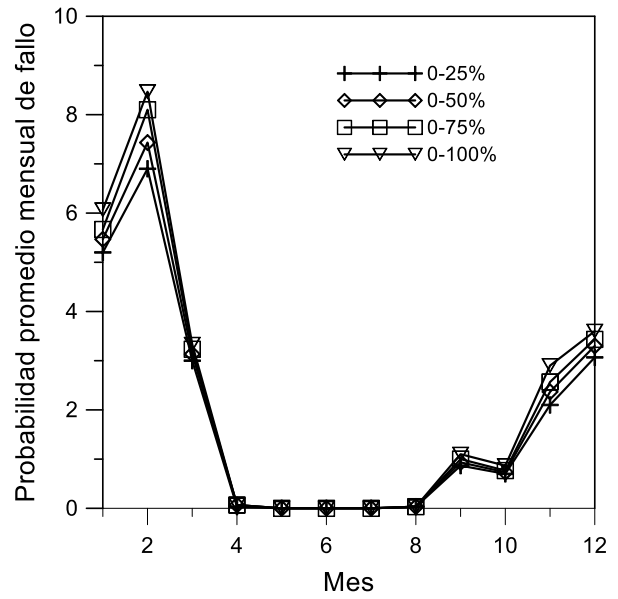


Figura 4. Probabilidad promedio mensual de fallo de las demandas futuras de riego.

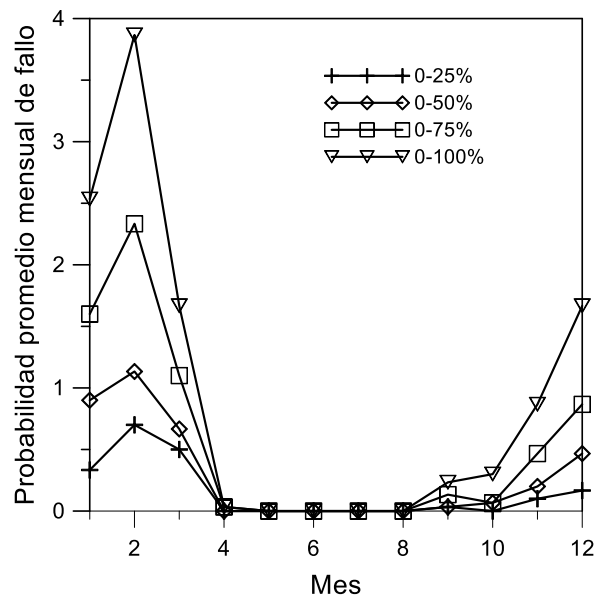


Figura 5. Probabilidad promedio mensual de fallo de la demanda futura de la ciudad de Cuenca en el escenario con restricciones a las demandas menos prioritarias.

Podría interesar para la gestión futura del sistema bajar el valor de estimación de probabilidad de fallo en las demandas, sobre todo de las prioritarias, por lo tanto se analizó la posibilidad de implementar una restricción en los usos del agua destinadas para el riego, abrevaderos de ganado y las piscícolas, de tal manera de bajar la probabilidad de fallo en el suministro para las demandas de consumo humano, especialmente de la ciudad de Cuenca. Entonces se realizó la simulación múltiple con una restricción del 50% de las demandas con menor prioridad, manteniendo fija la capacidad factible del embalse (10 Hm³). Este procedimiento dio como resultado la disminución de los valores promedios mensuales de probabilidad total, el mes de febrero considerado el mes más crítico bajo de 6.33% a 3.87 (Figura 5), demostrando la utilidad de la metodología para disminuir el riesgo de fallo en el suministro de agua, sobre todo de las demandas de agua más prioritarias.

4. CONCLUSIONES

La metodología expuesta en el presente artículo proporciona la posibilidad de generar alternativas en la oferta y demanda de agua, muy necesarias para una planificación y gestión de sistemas de recursos hídricos en cuencas deficitarias de agua. Esta metodología podría ser una herramienta eficaz para inferir reglas y políticas del uso de agua en cuencas hidrográficas a nivel local, regional y/o nacional. Además la introducción de reglas de gestión de sistemas de recursos hídricos con una introducción del riesgo asociado a la decisión, brinda a los gestores una seguridad y confianza de contemplar los sistemas de manera global, así como también a cada uno de sus componentes.

La aplicación de la metodología con la información disponible y estimada, muestra que existirán sistemas de recursos hídricos con una inminente necesidad de implementación de infraestructuras de regulación del agua para satisfacer a las demandas de agua, como el sistema de la subcuenca del río Tomebamba, mientras el sistema de la subcuenca del río Pindilig no requiere de embalses al menos en los próximos 30 años, sin embargo no hay que descartar esta posibilidad, dado que puedan surgir en el futuro nuevas demandas que no fueron consideradas en este estudio. Además existirán situaciones en donde no se pueda implementar embalses de gran magnitud, dando como resultado la necesidad de una gestión de sistemas de recursos hídricos con un riesgo de desabastecimiento de agua.

Los resultados en el sistema de la subcuenca del río Tomebamba muestran una disminución en el riesgo de desabastecimiento de agua en las demandas de consumo humano de la ciudad de Cuenca cuando se toma la decisión de restringir el suministro de agua a la mitad de las demandas con una prioridad menor (riego, abrevaderos de ganado, hidroeléctricas y piscícolas). Esta regla de gestión planificada supone un beneficio para las demandas prioritarias, sin embargo supone un costo en las demandas de riego y otros usos, dado que no tendrán el suministro total para sus fines. No obstante se pudo también poner restricciones de pequeña magnitud a las demandas prioritarias y así disminuir restricciones muy severas en las demandas menos prioritarias, pero esto ya dependerá de la toma de decisiones del gestor del sistema. Hacer notar que se pudo haber tomado como alternativa de gestión, la implementación de un embalse con una capacidad ideal (satisfacción total de las demandas), pero la capacidad de almacenamiento se ha pretendido limitarla para mostrar la utilidad de una metodología de gestión de sistemas de recursos hídricos cuando existe riesgo de desabastecimiento de agua.

En la aplicación de los casos de estudios se han extraído algunas enseñanzas, por ejemplo los resultados de la planificación de la gestión en la subcuenca del río Tomebamba mostraron que las situaciones más graves se producen en el mes de febrero, por lo cual este mes es de suma importancia para la gestión mensual futura de este sistema de recursos hídricos. Por otro lado, las reglas de gestión planificadas en la subcuenca del río Tomebamba deberán ser revisadas y perfeccionadas cuando el sistema esté en la fase de explotación. En esta fase se podría realizar una gestión del sistema en periodos muy cortos para la toma de decisiones, explorando modelos para predicción de caudales a corto plazo.

5. AGRADECIMIENTOS

El autor agradece al Profesor Jan Feyen por sus constructivas y valiosas sugerencias, las cuales mejoraron la estructura de este artículo.

BIBLIOGRAFÍA

Akaike, H., 1974. A new look at the statistical model identification. IEEE Transactions on Automatic Control, AS19 (6), 716-723.

- Alemu E.T., R.N. Palmer, A. Polebitski, B. Meaker, 2011. Decision support system for optimizing reservoir operations using ensemble streamflow predictions. *J. Water Res. Pl.-ASCE*, 237(1), 72-82.
- Andreu, J., J. Capilla, E. Sanchís, 1996. AQUATOOL, a generalized decision-support system for water-resources planning and operational management. *J. Hydrol.*, 177(3-4), 269-291.
- Andreu, J., A. Solera, J. Paredes, 2003. Sistema de estimación de riesgos en la gestión de sequías. In: *Jornadas sobre sistemas de ayuda a la decisión ante problemas hidrológicos e hidráulicos en tiempo real*. CEDEX, Madrid, España.
- Andreu, J., A. Solera, J. Capilla, J. Ferrer, 2010. SIMGES Simulation Model of Water Resource Management including Conjunctive Use. Version 3.02. User manual. Polytechnic University of Valencia, España, 100 pp. Descargado de <http://www.upv.es/aquatool/docs/ManSIMGESI.pdf> en septiembre 2012.
- Diaz, G.E., T. Brown, 1997. AQUARIUS: A modeling system for river basin water allocation. General Technical Report RM-GTR-299, USDA Forest Service, Washington D.C., USA.
- Dolling, O.R., 2001. Sistema de apoyo a la gestión integral de cuencas hidrográficas. *Disertación Doctoral Ingeniero de Caminos, Canales y Puertos*, Pontificia Universidad Católica de Chile, Santiago, Chile, 238 pp.
- Huang, Y., Y.P. Li, X. Chen, A.M. Bao, M. Zhou, 2010. Simulation-based optimization method for water resources management in Tarim River Basin, China. *Energ. Policy*, 2, 1451-1460.
- Hurvich, C.M., C.L. Tsai, 1989. Regression and time series model selection in small samples. *Biometrika*, 76(2), 297-307.
- Hydrologic Engineering Center, 1998. HEC-5: Simulation of flood control and conservation systems, Version 8.0. Department of the Army Corps of Engineers, Davis, CA, USA. Descargado de <http://www.hec.usace.army.mil/software/legacysoftware/hec5/hec5.htm> en septiembre 2012.
- Labadie, J.W., M.L. Baldo, R. Larson, 2000. MODSIM: Decision Support System for River Basin Management: Documentation and User Manual, Colorado State University and U.S. Bureau of Reclamation, Ft Collins, CO, USA.

Lund, J.R., J. Guzman, 1999. Some derived operating rules for reservoirs in series or in parallel. *J. Water Res. Pl.-ASCE*, 125(3), 143-153.

Ochoa, J.C., 2002. Modelo estocástico de redes neuronales para la síntesis de caudales aplicados a la gestión probabilística de sequías. Tesis Doctoral. Escuela Técnica Superior de Ingenieros de Caminos Canales y Puertos, Departamento de Ingeniería Hidráulica y Medio Ambiente, Universidad Politécnica de Valencia, Valencia, España, 205 pp.

Sánchez, S.T., J. Andreu, A. Solera, 2000. Gestión de recursos hídricos con decisiones basadas en estimación del riesgo. Publicaciones Universidad Politécnica de Valencia, Valencia, España, 158 pp.

SEI Stockholm Environment Institute, 2001. WEAP: Water Evaluation and Planning System. Stockholm Environment Institute. Descargado de <http://www.weap21.org/> en septiembre 2012.

Simonovic, S. P., Savic, D. , 1989. Intelligent Decision Support and Reservoir management and Operations, *J. Comput. Civ. Eng.*, 3(4), 367–385.

Solera, A., 2003. Herramientas y métodos para la ayuda a la decisión en la gestión sistemática de recursos hídricos. Aplicación a las cuencas de los ríos Tajo y Júcar. Tesis Doctoral. Escuela Técnica Superior de Ingenieros de Caminos, Canales y Puertos, Departamento de Ingeniería Hidráulica y Medio Ambiente, Universidad Politécnica de Valencia, Valencia, España, 278 pp.

Sveinsson, O.G.B., J.D. Salas, W.L. Lane, D.K. Frevert, 2007. Stochastic analysis modeling and simulation (SAMS). User's Manual. Version 2007. Computing Hydrology Laboratory, Department of Civil and Environmental Engineering, Colorado State University, Fort Collins, CO, USA, 123 pp.

ANEXO 4: Análisis del sistema de recursos hídricos de la cuenca del río Tomebamba en Ecuador, mediante modelos estocásticos y de gestión⁴

RESUMEN

La cuenca del río Tomebamba ubicada en el centro sur del Ecuador adolece de algunas problemas como el contante crecimiento de la población que demanda más recursos hídricos, intervenciones humanas en ecosistemas naturales proveedores de agua en la zona alta de la cuenca, épocas de sequía que afectan el suministro normal de agua y deficiencias en la planificación y gestión del sistema de recursos hídricos de la cuenca. Esto ha motivado realizar el presente estudio, mediante el análisis del sistema aplicando modelos hidrológicos estocásticos y realizando simulaciones con un enfoque de riesgo de la gestión de oferta y demanda con múltiples escenarios futuros con la finalidad de realizar una planificación de acciones futuras como la implementación de infraestructuras de regulación y reglas de operación en el suministro de agua a las demandas futuras del sistema. Los resultados muestran que trabajar con una metodología que incorpore componentes probabilísticos para la toma de decisiones facilita a los gestores del sistema tener una visión más integral del sistema para identificar reglas adecuadas que disminuyan el riesgo de desabastecimiento de agua a las demandas, sobre todo de las demandas prioritarias.

Palabras clave: Modelación Estocástica, Simulación, Gestión, Planificación, Riesgo

1. INTRODUCCIÓN

El recurso hídrico se encuentra bajo presión creciente debido al incremento de la población, las actividades económicas y la mayor competencia entre usuarios. Además la contaminación del agua agudiza aún más esta situación, provocando que exista escasez del recurso.

Es evidente que un mayor desarrollo implica un impacto del recurso hídrico, y esto sumado a los escenarios potenciales futuros de los modelos de cambio climático, muestran que existirá una disminución de los recursos hídricos.

⁴ Avilés A, Solera A (2013) Análisis de sistemas de recursos hídricos de la cuenca del río Tomebamba en Ecuador, mediante modelos estocásticos y de gestión. En: Solera A, Paredes J, Andreu J (editores) Aplicaciones de sistemas soporte a la decisión en planificación y gestión integradas de cuencas hidrográficas. Marcombo, Barcelona, España, pp 51–61

Es por esto que se hace necesaria una gestión del agua tanto desde el punto de vista de la oferta como también de la demanda, de tal manera que cubra los requerimientos de los usuarios.

En las cuencas hidrográficas ecuatorianas existe la necesidad de una gestión y planificación hídrica, debido al constante crecimiento de la población que demandará más recursos hídricos para nuevos servicios de agua potable, actividades productivas, actividades recreativas, etc.; esto sumado a intervenciones del hombre que afectan ecosistemas naturales proveedores de agua en la parte alta de la cuenca y épocas de sequía que afecta el suministro normal de agua.

La cuenca del río Tomebamba está ubicada en la zona centro sur del Ecuador (Figura 1), tiene gran importancia por su abastecimiento de agua a la ciudad de Cuenca (tercera ciudad del Ecuador en importancia), sin embargo adolece de algunos problemas acotados en el párrafo anterior y tiene algunas deficiencias en lo que respecta a infraestructuras de regulación y monitoreo de caudales. Esta realidad hace necesario una planificación y gestión de acciones futuras, de tal manera de enfrentar sobre todo las épocas de sequía.

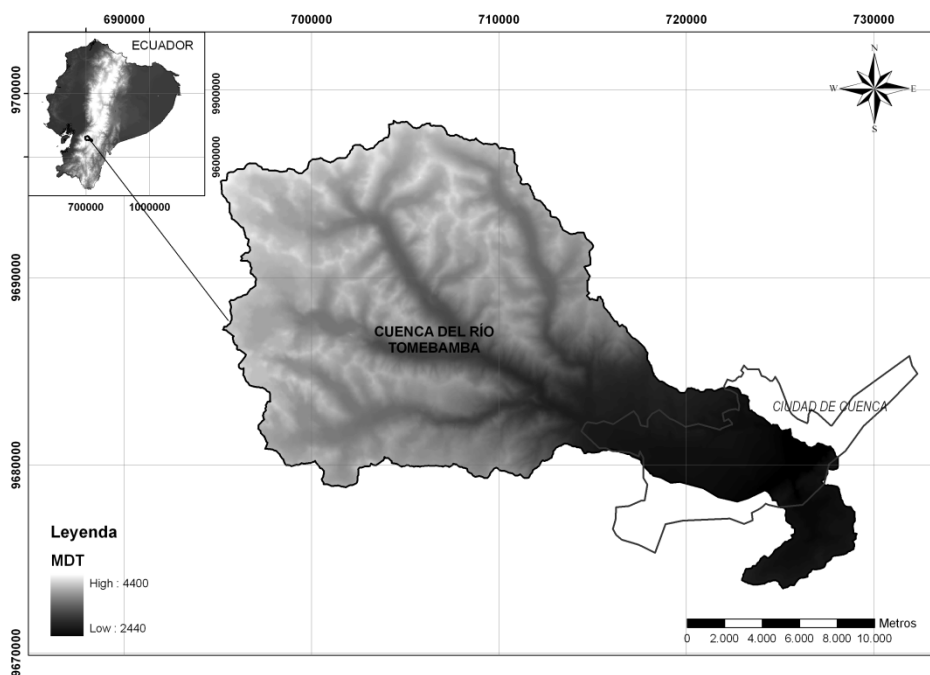


Figura 1. Ubicación cuenca del río Tomebamba

Por lo expuesto, para el presente trabajo se ha planteado analizar el sistema de recursos hídricos de la cuenca del río Tomebamba mediante modelos hidrológicos estocásticos y la simulación de la gestión con múltiples escenarios futuros, a partir de una planificación de infraestructuras de regulación y las posibles acciones de gestión en la demanda futura del sistema.

2. ÁMBITO DE ESTUDIO

La cuenca del río Tomebamba tiene un área de 380 Km² (Figura 2), su parte alta (3000 – 4400 m.s.n.m.) es una zona natural formada en su gran mayoría por humedales, lagunas, páramos y bosque andino, la parte media (2800 – 3000 m.s.n.m.) es una mezcla de zonas agrícolas, zonas urbanas y en poca cantidad zonas con cobertura vegetal natural y la parte baja (2400 – 2800 m.s.n.m.) está formada por zonas urbanas y agrícolas.

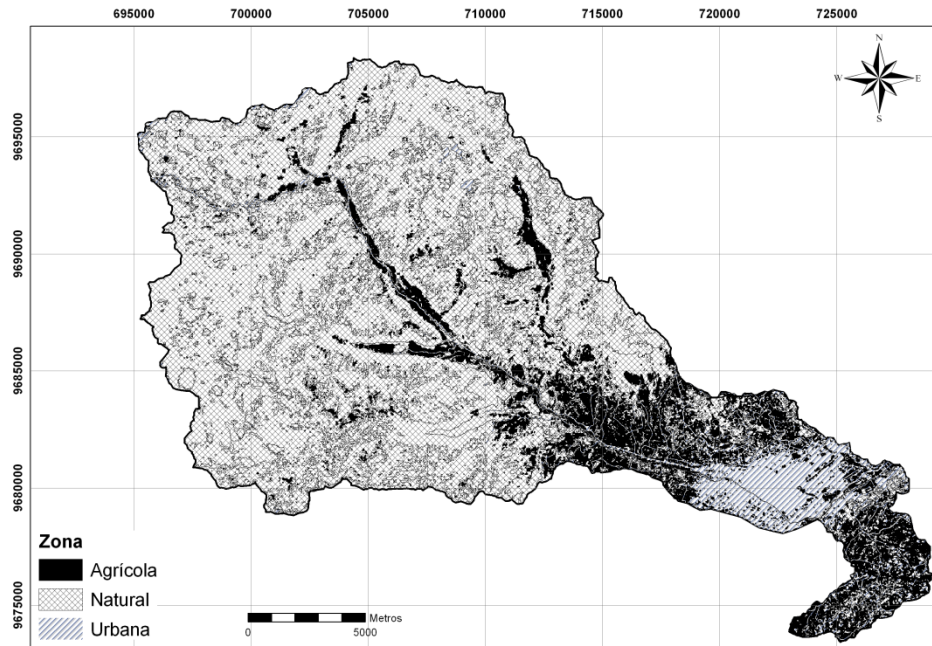


Figura 2. Uso del suelo en la cuenca del río Tomebamba

La cuenca del río Tomebamba cuenta con información histórica de caudales mensuales en la estación hidrológica Matadero en Sayausí (1979 – 2009), ubicada en la desembocadura de la parte alta de la cuenca, es decir, aguas abajo de la zona considerada natural (áreas con escasa utilización del agua). Esta información fue facilitada por el Instituto Nacional de Meteorología e Hidrología del Ecuador – INAMHI.

También se pudo obtener información de la demanda de agua de la cuenca, mediante la base de datos de las concesiones (permisos) de agua de los diferentes usos, facilitada por la Secretaría Nacional del Agua del Ecuador – SENAGUA y también los datos de consumo de agua de la Empresa municipal de agua potable de la ciudad de Cuenca – ETAPA.

2.1 Oferta de agua

La serie de tiempo de aportaciones anuales de la estación hidrológica Matadero en Sayausí consta de 31 años de datos (372 meses), tiene una media anual de 194.7 Hm³/año, con un valor máximo anual de 299.6 Hm³/año y un valor mínimo anual de 119.4 Hm³/año. La Tabla 1 muestra un resumen de los estadísticos básicos de la serie anual y mensual.

Mes	Media (Hm ³)	Desviación Estándar	Coefficiente Variación (CV)	Sesgo	Valor Min (Hm ³)	Valor Max (Hm ³)	ACF (1)	ACF (2)
ANUAL	194.7	45.74	0.2349	0.2459	119.4	299.6	0.2146	-0.0083
ENE	9.542	4.991	0.5231	0.2296	2.073	21.58	0.1126	-0.1636
FEB	14	8.729	0.6237	0.9517	0.8903	41.82	0.0228	-0.2175
MAR	20.68	9.746	0.4714	0.2694	3.361	40.82	-0.1288	-0.1016
ABR	26.02	9.498	0.365	0.4144	6.488	50.65	0.2158	-0.0898
MAY	23.1	10.3	0.446	1.0418	7.69	48.88	0.2234	0.0427
JUN	19.16	8.108	0.4231	0.6917	3.914	40.1	0.1589	0.0629
JUL	19.37	6.015	0.3105	-0.0553	6.265	34.71	0.2468	0.2443
AGO	12.91	6.204	0.4805	0.7133	2.092	31.69	-0.1782	-0.0629
SEP	10.11	5.054	0.4998	0.3673	0.8891	21.05	0.03	-0.2111
OCT	14.26	6.491	0.4552	0.1466	2.204	27.48	-0.0753	0.172
NOV	12.19	8.189	0.6717	0.7523	1.159	33.72	-0.0067	0.0547
DIC	13.33	7.948	0.5963	0.6677	1.953	36.75	0.0454	0.262

Tabla 1. Estadísticas básicas serie anual y mensual original - Matadero en Sayausí

2.2 Demanda de agua

En la cuenca del río Tomebamba existen usos diversos (uso doméstico, riego, abrevaderos de ganado, etc.), el uso que predomina es el uso doméstico para la ciudad de Cuenca, seguido por el uso de riego y en menor cantidad los otros usos.

Con la finalidad de que sea un esquema de fácil introducción en el modelo de simulación de la gestión, se realizó una agrupación de las demandas por prioridad en uso y por el peso que tienen las demandas que utilizan un mayor volumen de recurso. La demanda para consumo doméstico de la ciudad de Cuenca es el mayor usuario de agua de la cuenca del río Tomebamba. Hay que hacer notar que la prioridad de utilización del agua, se considero primero el consumo humano, luego las demandas para riego y finalmente la demanda para otros usos (abrevaderos de ganado y piscícolas) en tercer lugar.

También se tomó en cuenta los requerimientos ambientales de los ecosistemas del entorno mediante la consideración de un caudal ecológico. Este caudal debería ser el que imite de mejor manera el régimen natural, por lo tanto su cálculo debe ser detallado y bien fundamentado, sin embargo debido al alcance del presente trabajo se ha asumido un caudal de 10% de los caudales medios mensuales, considerando el registro histórico de caudales.

Además se realizó un análisis de la demanda futura con la finalidad de trabajar con escenarios futuros para la gestión de los sistemas de recursos hídricos en la cuenca de estudio. La proyección de la demanda futura se hizo de una manera muy sencilla sin realizar un análisis profundo, dado que no es el objetivo del presente trabajo, sin embargo se podría afinar los cálculos en próximos trabajos.

Para la proyección de la demanda futura se consideró las posibles proyecciones de la población, hectáreas de riego, semovientes y piscícolas según criterios de conocimiento de las zonas de estudio y también de datos históricos de población que se han logrado conseguir. Además se consideraron las posibles proyecciones en las dotaciones de los usos del recurso hídrico (consumo humano, riego, abrevaderos de ganado y piscícolas), según criterios de conocimiento de las zonas de estudio. Las proyecciones se realizaron para el año 2040, tomando como año de inicio el año 2010.

Entonces, realizando los cálculos respectivos para las proyecciones futuras, se construyó la Tabla 2 para la demanda futura, haciendo notar que el caudal ecológico se mantuvo para el futuro ($0.617 \text{ m}^3/\text{seg}$).

MES	Demanda consumo humano ciudad de Cuenca (HM3/MES)	Demanda consumo humano otros usos (HM3/MES)	Demanda riego parroquia San Joaquín (HM3/MES)	Demanda riego parroquia Sayausí (HM3/MES)	Demanda otros usos (HM3/MES)	TOTAL (HM3/MES)
ENERO	4.8313	0.2024	0.5770	0.2029	1.3462	7.1598
FEBRERO	4.4435	0.2024	0.5770	0.2029	1.3462	6.7720
MARZO	4.9754	0.2024	0.5770	0.2029	1.3462	7.3040
ABRIL	4.8585	0.2024	0.5770	0.2029	1.3462	7.1870
MAYO	5.1087	0.2024	0.5770	0.2029	1.3462	7.4372
JUNIO	4.9773	0.2024	0.5770	0.2029	1.3462	7.3058
JULIO	5.1672	0.2024	0.5770	0.2029	1.3462	7.4957
AGOSTO	5.1726	0.2024	0.5770	0.2029	1.3462	7.5011
SEPTIEMBRE	5.1443	0.2024	0.5770	0.2029	1.3462	7.4729
OCTUBRE	4.6577	0.2024	0.5770	0.2029	1.3462	6.9863
NOVIEMBRE	5.1880	0.2024	0.5770	0.2029	1.3462	7.5166
DICIEMBRE	5.1408	0.2024	0.5770	0.2029	1.3462	7.4693

Tabla 2. Valores de demandas futuras en la cuenca del río Tomebamba

2.3 Garantías de las demandas

Se puede definir garantía como la probabilidad de que el sistema se halle en una situación satisfactoria, entendiendo por tal aquella en que el suministro es superior a la demanda solicitada (Balairón, 2002). A partir de esta definición se encuentra lo que se denomina fallo del sistema, entendiéndose que es la situación de insatisfacción en el suministro [Déficit > % demanda (mensual o anual)].

Dado que el suministro de agua para consumo humano es muy susceptible en la población de la cuenca del río Tomebamba, se asumió que existe fallo cuando el déficit sea mayor que el 1% de la demanda mensual. Para las demandas anuales se consideró el enfoque similar al departamento de recursos hídricos de Utah (UTAH-DWR) (Andreu et al, 2007), con los siguientes criterios de fallo (existirá fallo si se cumple al menos un criterio):

Déficit anual a 1 año $>$ D% demanda anual

Déficit anual a 2 años consecutivos $>$ E% demanda anual

Déficit anual a 10 años consecutivos $>$ F% demanda anual

Donde D, E y F son iguales a 5, 10 y 16 para las demandas anuales de uso doméstico, y 20, 30 y 40 para las demandas anuales de uso para riego y otros usos.

3. MODELACIÓN ESTOCÁSTICA

Los modelos estocásticos han sido utilizados para diversos problemas y necesidades relacionados con la gestión y planificación de los recursos hídricos, como por ejemplo, la determinación de la dimensión de un embalse, construcción de reglas de gestión para la operación de un sistema de embalses, estrategias de gestión para la adecuada distribución del recurso hídrico para los diferentes usos, realización de balances hídricos (oferta vs demanda) en cuencas hidrológicas, y otras más.

Estos se fundamentan principalmente en el máximo aprovechamiento de la información de la serie histórica de una variable determinada (caudales, lluvias, etc.), es decir reproducir las estadísticas históricas de las series de tiempo. El concepto clave en este tipo de modelos es la dependencia entre los datos.

En la literatura existen varios modelos estocásticos que se han utilizado en hidrología, los modelos que se han aplicado para las series de tiempo de los casos de estudio son los autoregresivos de media móvil - ARMA [incluyendo los modelos autoregresivos (AR) y los modelos de media móvil (MA)] y los modelos autoregresivos de media móvil periódicos – PARMA [incluyendo los modelos autoregresivos periódicos (PAR) y los modelos de media móvil (PMA)]. Los mismos que se detallaran a continuación; hay que hacer notar que en algunas actividades de la modelación se utilizó como ayuda el software SAMS 2007 (Sveinsson et al, 2007).

Se probaron modelos autoregresivos de media móvil (ARMA) (Escala anual) y modelos autoregresivos de media móvil periódicos (PARMA) (Escala mensual), con diferentes órdenes (p,q). Se escogió los mejores modelos con los menores valores del criterio de información de Akaike corregido (AICC) (Hurvich y Tsai, 1989), este criterio se basa en una penalización de los modelos con mayor número de parámetros.

Los mejores modelos sirvieron para la generación de series sintéticas anuales y series sintéticas mensuales, que servirán respectivamente para una planificación en la implementación de infraestructura para la regulación del agua y para la construcción de escenarios futuros para la simulación de la gestión de la infraestructura planificada. Lo esencial de esta generación es la conservación de las propiedades estadísticas de la muestra histórica, como resultado cada muestra generada son igualmente probables que ocurran en el futuro (Sveinsson et al, 2007).

A escala anual se probaron modelos ARMA con diferentes órdenes: (1,1), (1,2), (1,3), (2,1), (2,3) y (3,3), resultando el mejor modelo según el AICC el ARMA (1,1). Este modelo fue utilizado para generar 100 series sintéticas de 31 años cada una. A escala mensual se probaron 11 modelos PARMA con diferentes órdenes: (1,0); (1,1); (1,2); (1,3); (1,4); (2,1); (2,2); (3,2); (3,3); (4,3); (4,4), resultando los mejores modelos según el AICC el modelo PARMA (1,0) para los meses de enero, febrero, marzo, mayo, junio, septiembre y noviembre, el modelo PARMA (1,1) para los meses de abril, julio, agosto y diciembre y el modelo PARMA (1,2) para el mes de octubre. Con estos modelos se generaron 100 series sintéticas de 372 meses cada una. No fue necesario generar un número elevado de series sintéticas dado que el sistema es sencillo y con un solo punto de generación (Sánchez S., et al., 2001).

4. SIMULACIÓN DE LA GESTIÓN DEL SISTEMA

Para la simulación de la gestión del sistema de recursos hídricos se trabajará con el modelo SIMGES (Andreu et al., 2007) desarrollado por el Área de Ingeniería de Recursos Hídricos del Instituto de Ingeniería Hidráulica y Medio Ambiente, Universidad Politécnica de Valencia. Este modelo forma parte del sistema de soporte de decisiones AQUATOOL.

A partir de las series históricas de caudales y la agrupación de las demandas del agua, se construye un esquema del sistema de recursos hídricos de la cuenca en SIMGES (Figura 3). El proceso de simulación realiza una optimización de una red de flujo conservativa, mediante el cumplimiento de una función objetivo (satisfacción de las demandas, operación de

embalses, etc.) sujeta a las restricciones de conservación de masa (continuidad) y los límites físicos de transporte de flujo en conducciones (ríos, quebradas, etc.) y capacidades de embalses (reservas) y otros elementos si los hubiere.

Se desarrolló la simulación mensual del escenario actual (1979-2009) mediante la introducción de la información de la serie de aportaciones mensuales históricas y la demanda actual, luego se hizo la simulación del escenario futuro (2010-2040) con la introducción de la demanda futura y la información de la serie de aportaciones mensuales históricas, suponiendo que estas sucederán en el futuro.

Los resultados de la simulación de los dos escenarios muestran que todas las demandas tienen déficits en algunos meses, además no cumplen con la garantía (satisfacción de las demandas), por lo tanto se intentará incluir una infraestructura de regulación virtual (embalse), de tal manera que las garantías de las demandas sean cumplidas, estamos hablando de un embalse con una capacidad ideal, sin embargo suponiendo restricciones de tipo físico como la topografía de la zona de estudio, el uso del suelo en las misma y el impacto ambiental que podría causar una infraestructura de esta magnitud, se probará con una capacidad menor acorde a la realidad de la cuenca, en cuyo caso se incumplirían nuevamente las garantías, por lo que haría falta la elaboración de acciones de gestión de la demanda en el sistema.

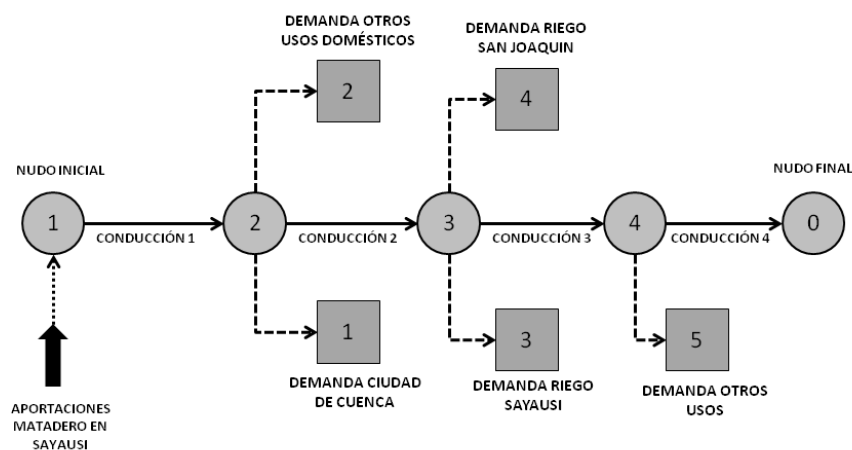


Figura 3. Esquema del sistema de recursos hídricos de la cuenca del río Tomebamba

Simulando el sistema en el periodo de 2010-2040 con un embalse con una capacidad de 21 Hm³ cumple totalmente las garantías de las demandas, siendo esta capacidad la ideal para la satisfacción del sistema, sin embargo pensando en una capacidad factible se disminuyó esta capacidad a la mitad (10 Hm³) suponiendo las restricciones expuestas anteriormente, por lo tanto se pasó de una gestión normal del sistema a una gestión con riesgo de desabastecimiento de agua.

5. SIMULACIÓN DE LA GESTIÓN CON ENFOQUE DE RIESGO

Durante la explotación de los sistemas de recursos hídricos, pueden darse inevitablemente situaciones en las que no sea posible suministrar de forma satisfactoria los usos establecidos (fallo del sistema, denominado también sequía operativa), esta situación puede minimizarse mediante una anticipación de la sequía en el sistema en las fases de planificación y explotación, minimizando los valores esperados de duración y magnitud de sequías, mediante la adopción de reglas de gestión (Sánchez S., et al., 2001).

Por lo expuesto en el presente estudio se realizó el análisis del sistema hídrico del río Tomebamba con la proyección de implementar un embalse de regulación con una capacidad factible y planificando la gestión con escenarios estocásticos y valore el riesgo de la gestión planificada. Para la ejecución de esta metodología se ha requerido la ayuda del modelo SIMRISK (Sánchez S., et al., 2001) del sistema de soporte de decisiones AQUATOOL. El mismo que proporciona al decisor de herramientas para la gestión, como la determinación de las probabilidades de fallos en los sistemas de recursos hídricos y la simulación de medidas de mitigación según el riesgo admisible impuesto.

Para la simulación de la gestión con enfoque de riesgo se implementaron los siguientes pasos:

- a. Generación de múltiples escenarios mediante las series sintéticas mensuales generadas anteriormente.
- b. Simulación estocástica mensual del comportamiento del sistema con los escenarios del punto anterior desde el año 2010 hasta el año 2040 y las demandas futuras, para lo cual se utilizará el modelo SIMRISK.
- c. Análisis de los resultados del modelo y determinación del estado del sistema mediante la evaluación de las probabilidades de fallo mensual y la magnitud del déficit en el suministro (dividido en cuatro niveles, 0-25%, 25-50%, 50-75% y 75-100%, por ejemplo el nivel más grave es cuando el suministro se ubica entre 0-25% del valor total de la demanda, es decir, el déficit supera los 75% del valor total

de la demanda). d. Decisión sobre el riesgo asumible de fallo del sistema, para lo cual se genera alternativas de gestión de demandas para que el riesgo de fallo sea el menor posible.

Los resultados con la planificación del embalse de capacidad factible muestran que para el mes de febrero (época más desfavorable) la demanda de la ciudad de Cuenca tiene una probabilidad promedio mensual de fallo aproximada de 6% (Figura 4), para las demandas de riego existe una probabilidad promedio mensual de fallo aproximada de 8% y para las demandas de otros usos tienen una probabilidad promedio mensual de fallo aproximada del 11%.

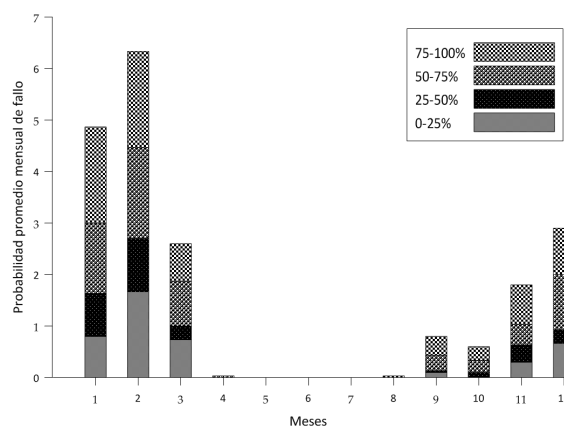


Figura 4. Probabilidad promedio mensual de fallo de la demanda futura ciudad de Cuenca

Estas probabilidades de fallo pueden ser consideradas bajas, sin embargo el gestor del sistema podría considerar no adecuadas y le puede interesar para la gestión futura del sistema bajar el valor de estimación de probabilidad de fallo en las demandas, sobre todo de las prioritarias, en el caso de la cuenca del Tomebamba el uso prioritario es el consumo humano, luego el riego y luego los otros usos, por lo tanto se intentará implementar una restricción en los usos de riego y otros usos de tal manera de bajar la probabilidad de fallo en el suministro para las demandas prioritarias de consumo humano.

Los valores estimados promedios mensuales de probabilidades de fallo de la demanda de la ciudad de Cuenca después de simular el sistema con una restricción del 50% en las demandas de riego y otros usos, manteniendo fija la capacidad del embalse factible (10 Hm³) se muestran en la Figura 5.

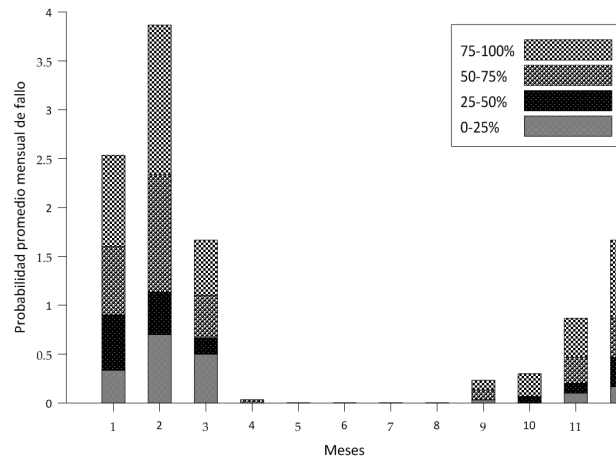


Figura 5. Probabilidad promedio mensual de fallo de la demanda futura ciudad de Cuenca con restricciones en las otras demandas

Los valores promedios mensuales de probabilidad total de fallo bajaron, el valor del mes de febrero considerado el mes más crítico en esta cuenca bajo de 6.33% a 3.87%. Podríamos asumir estas probabilidades de fallos como factibles para el sistema, no obstante va a depender del riesgo que se quiera imponerse los gestores del sistema para la insatisfacción en las demandas y el sistema en su conjunto.

6. CONCLUSIONES

El presente trabajo optó por modelos estocásticos, dado su sencillez y rapidez en la modelación de la realidad de un sistema donde predominan los procesos aleatorios, sin embargo conocer la dinámica de variables físicas en un sistema, ayuda considerablemente en la formulación de modelos más competitivos para representar la realidad de un sistema, los mismos que pueden ser objeto de próximos estudios.

Por otro lado los resultados de la simulación del sistema mostraron que la cuenca del río Tomebamba necesitaría de un embalse con capacidad ideal para cumplir con las garantías de las demandas, sin embargo, suponiendo restricciones en la implementación de un embalse de gran capacidad, se planificó un embalse con una capacidad factible, incluyendo la

planificación de la gestión futura mediante reglas en el suministro de las demandas menos prioritarias con la finalidad de disminuir el riesgo de fallo en las demandas prioritarias.

La incorporación de componentes probabilísticos para generar alternativas en la oferta y demanda del agua en sistemas de recursos hídricos favorece a una visión más integral de los problemas y necesidades que se puedan presentar en el futuro y ayuda a los tomadores de decisiones para la identificación de reglas adecuadas en la gestión y planificación del agua en cuencas.

Como reflexión final de la simulación estocástica del sistema hídrico de la cuenca del río Tomebamba, se puede indicar que las reglas de gestión planificadas deberán ser revisadas y perfeccionadas cuando el sistema esté en la fase de explotación. En esta fase se podría realizar una gestión del sistema en periodos muy cortos para la toma de decisiones, como por ejemplo 1, 2 o 3 años, dependiendo de la longitud máxima de las sequías hidrológicas observadas y generadas.

7. REFERENCIAS

Balairón Luis. 2002. “Gestión de recursos hídricos”, Cataluña – España, Edición UPC, Universidad Politécnica de Cataluña.

Andreu, J.Capilla, J.Ferrer y A.Solera. 2007. “Modelo Simges para simulación de cuencas : manual de usuario V 3.00”, Editorial UPV, Universidad Politécnica de Valencia, Valencia.

Hurvich, C.M. and Tsai, C.-L. 1989. “Regression and time series model selection in small samples”, *Biometrika*, 76(2):pp. 297–307.

Sánchez, S. T., Andreu, J. y Solera, A. 2001. Gestión de recursos hídricos con decisiones basadas en estimación del riesgo, Editorial Universidad Politécnica de Valencia, Valencia.

Sveinsson, O. G. B., Salas, J. D., Lane, W. L., and Frevert, D. K. 2007. Stochastic analysis modeling and simulation (SAMS) - Version 2007 - User's Manual, Colorado State University, Fort Collins, Colorado.