

En español

Viabilidad para pronósticos hidrológicos de niveles diarios, semanales y decadales en Colombia

Efraín Antonio Domínguez Calle¹, Hector Angarita² y Hebert Rivera³

RESUMEN

El presente artículo analiza y concluye sobre la viabilidad de pronósticos hidrológicos de niveles diarios, semanales y decadales en 20 estaciones hidrológicas de la red de monitoreo hidrometeorológico que soporta al Servicio de Alertas del Instituto de Hidrología, Meteorología y Estudios Ambientales – Ideam en Colombia (www.ideam.gov.co). Esta viabilidad se determina a través de un conjunto de criterios de desempeño ortogonales y para el presente estudio recomienda la aplicación de combinaciones lineales adaptativamente óptimas (CLAO) como operador viable para la configuración de un sistema de pronóstico hidrológico en tiempo real de niveles diarios, semanales y decadales. En conclusión, se muestra que los pronósticos de niveles diarios, semanales y decadales tienen una viabilidad de pronóstico satisfactoria para el 70% de los casos estudiados.

Palabras clave pronósticos hidrológicos, modelación matemática, combinaciones lineales adaptativamente óptimas.

Recibido: mayo 18 de 2009

Aceptado: junio 21 de 2010

Introducción

La predicción hidrológica es una de las facetas más importantes de la hidrología. La viabilidad de realizar pronósticos hidrológicos en tiempo real es una característica de los tiempos modernos. Aparece con los inicios de la hidrología científica establecidos por Pierre Perrault en 1674. La observación del tiempo y del estado de las corrientes superficiales y de otras fuentes de agua data desde los egipcios, en el río Nilo (www.waterhistory.org/histories/cairo/). Sin embargo, sólo hasta el establecimiento de las mediciones sistemáticas y estandarizadas fue posible pensar en sistemas de alerta hidrológica. Colombia no ha sido ajena a la corriente internacional que dictó las pautas para establecer no sólo redes de monitoreo sobre el recurso hídrico, sino también sistemas de alerta temprana que ayudan a tomar ventaja de la información hidrológica en tiempo real y de su conjunción con los avances en ordenadores y medios de comunicación. En el año 1976 Colombia, con el Servicio Colombiano de Hidrología y Meteorología (SCMH), incursionó en predicción hidrológica con el modelo Sacramento, que se operaba con tarjetas perforadas y asimilaba información transmitida por radio y teléfonos convencionales.

¹ Ingeniero Hidrólogo. PhD. en Ciencias Técnicas, Departamento de Ecología y Territorio, Facultad de Estudios Ambientales y Rurales, Pontificia Universidad Javeriana, Bogotá, Colombia. CeIBA – Complejidad, Bogotá, Colombia.

² Ingeniero Civil. M.Sc., en Hidrosistemas, Facultad de Ingeniería, Pontificia Universidad Javeriana, Bogotá, Colombia. CeIBA – Complejidad, Bogotá, Colombia.

³ Ingeniero Hidrólogo. Ph.D., en Hidrología. Subdirección de Administración de los Recursos Naturales y Áreas Protegidas, CAR, Bogotá, Colombia.

In English

The feasibility of daily, weekly and ten-day water-level forecasting in Colombia

Efraín Antonio Domínguez⁴ Calle, Hector Angarita⁵ and Hebert Rivera⁶

ABSTRACT

This paper analyses the feasibility of forecasting daily, weekly and ten-day water-levels at 20 hydrological stations forming part of the monitoring network supporting the Institute of Hydrology, Meteorology and Environmental Studies' (IDEAM) Alert Service in Colombia (www.ideam.gov.co). Such viability was determined by a set of orthogonal performance criteria and implementing optimally adaptive linear combinations (OALC) was recommended for this study as a viable operator for configuring a real-time hydrological forecast system. It is shown that the forecast for daily, weekly and ten-day levels had satisfactory viability for 70% of the cases studied.

Keywords: hydrological forecasting, mathematical modelling, optimal linear adaptive combination.

Received: may 18th 2009

Accepted: jun 21th 2010

Introduction

Hydrological forecasting is one of the most important facets of hydrology. The feasibility of conducting real-time hydrological forecasting is a feature of modern times that appeared with the beginnings of scientific hydrology seated in 1674 by Pierre Perrault (Perrault, 1674). Observation of the weather and the state of surface runoff and other water sources dates from the Egyptians in the Nile valley (www.waterhistory.org/histories/cairo/). However, only until the establishment of systematic and standardised measurement did it become possible to think about a hydrological warning system setup. Colombia has not been far from the international trend that issued guidelines for establishing monitoring networks for water resources and early warning systems which would help take advantage of real-time hydrological information and its combination with advances in computers and the media. The Colombian Hydrological and Meteorological Service (CHMS) dabbled in hydrological forecasting with the Sacramento model in 1976; this was operated by punch cards and assimilating information transmitted by radio and telephone.

⁴ Hydrologist Engineer. PhD. Technical Sciences, Departamento de Ecología y Territorio, Facultad de Estudios Ambientales y Rurales, Pontificia Universidad Javeriana, Bogotá, Colombia. CeIBA – Complejidad, Bogotá, Colombia.

⁵ Civil Engineer. M.Sc., in Hydrosystems, School of engineering, Pontificia Universidad Javeriana, Bogotá, Colombia. CeIBA – Complejidad, Bogotá, Colombia.

⁶ Hydrologist Engineer. Ph.D., in Hydrology. Subdirección de Administración de los Recursos Naturales y Áreas Protegidas, CAR, Bogotá, Colombia.

En español

El sistema emitía pronósticos cuantitativos y apoyaba a los organismos de socorro en el manejo de la temporada invernal. En la actualidad, con mayor disponibilidad de tecnología en los ámbitos informáticos y de telecomunicaciones, Colombia no dispone de modelos hidrológicos que, en tiempo real, emitan pronósticos hidrológicos cuantitativos. Con miras al restablecimiento de capacidades institucionales en predicción hidrológica, el Ideam ha venido adelantando esfuerzos para implantar un servicio de pronóstico hidrológico cuantitativo en tiempo real.

Aquí se analiza la capacidad institucional del Ideam (fortalezas y debilidades) para emprender la emisión operativa de pronósticos cuantitativos de los niveles diarios, semanales y decadales. Entre otros aspectos, se define cuándo el pronóstico hidrológico, cuantitativo y en tiempo real es viable, lo cual se establece estudiando aspectos como: el nivel de capacitación del personal, las capacidades informáticas, la cadena de registro hidrológico continuo (red de estaciones automáticas), la disponibilidad de *software* integrador de flujos de información hidrometeorológica y la calidad de la información en tiempo real. En la etapa final se analiza con detenimiento la posibilidad de emitir pronósticos hidrológicos en tiempo real basándose en una técnica que contempla operadores matemáticos con componentes autorregresivos (variable endógena) y exógenos, y que además cuenta con un mecanismo expreso de optimización. En resumen, la viabilidad de los pronósticos acá planteados se concibe a través de la completitud en infraestructura y personal y la posibilidad de contar con una estructura matemática "optimizable" de acuerdo a criterios de desempeño seleccionados con anterioridad. Finalmente, se presenta un plan de trabajo para lograr concretar una cadena de producción de pronósticos de niveles diarios, semanales y decadales en tiempo real.

Métodos y datos

El presente trabajo contempló cuatro pasos metodológicos: 1) revisión de antecedentes de modelación cuantitativa para pronósticos hidrológicos en el Ideam; 2) definición de viabilidad de pronóstico hidrológico; 3) aplicación de combinaciones lineales adaptativamente óptimas (CLAO) para pronósticos diarios, semanales y decadales de los niveles del agua en las estaciones de la red de alertas del Ideam; y 4) definición de un criterio para la calificación de la viabilidad de pronóstico en un punto de análisis y en el conjunto de estaciones hidrológicas estudiadas. La información hidrometeorológica utilizada se obtuvo de las bases de datos en tiempo real que almacena la información sobre niveles y precipitaciones de la red de alertas hidrometeorológica del Ideam. Los puntos de pronóstico analizados se presentan en la tabla 1. También se utilizaron series de tiempo obtenidas de la base de datos del Ideam que almacena la información de las estaciones que no operan en tiempo real. Estas últimas se utilizaron como predictores en aquellos puntos de pronóstico para los cuales no se encontraron predictores entre las estaciones que transmiten información en tiempo real.

Revisión de antecedentes de modelación cuantitativa para pronósticos hidrológicos en el IDEAM

Para el Servicio de Alertas del Ideam el pronóstico cuantitativo de niveles es una de las metas prioritarias. Se han adelantado varias investigaciones para la implementación de pronósticos cuantitativos, éstas incluyen desde modelos cualitativos basados en reglas semánticas y lógica difusa (Rivera *et al.*, 2004) y técnicas de series de tiempo (Mussy, 2005), hasta modelos avanzados capaces de

In English

The system emitted quantitative predictions and supported disaster-preventing agencies in managing season floods during winter. Colombia has no hydrological models for real-time quantitative hydrological forecasting today, even with the increased availability of technology in computing and telecommunications fields. IDEAM has made efforts to implement a quantitative realtime hydrologic forecasting service aimed at restoring institutional capacity in hydrological forecasting.

This paper examines IDEAM's institutional capacity (strengths and weaknesses) for undertaking operational quantitative forecasts for daily, weekly and ten-day frameworks. The feasibility of quantitative, real-time hydrologic forecasting is defined here, being set by studying aspects such as staff training level, computational capacity, continuous hydrological records (real-time hydrometrical network), the availability of software for integrating hydrometeorological information flows and the quality of real-time information. The final stage gives a detailed analysis of the possibility of issuing real-time hydrological forecasting based on a mathematical technique providing operators with endogenous (auto-regressive) and exogenous components and also has an explicit optimisation mechanism. The viability of forecasts raised here is conceived through completeness in infrastructure and staff and the possibility of having a mathematical structure optimised according to selected performance criterion. A work plan is presented for achieving concrete real-time forecast of daily, weekly and ten-day average water-levels.

Methods and data

This study was developed through four methodological steps: reviewing previous quantitative modelling for hydrological forecasting in IDEAM, defining hydrological forecasting feasibility, applying optimally-adapted linear combinations (OALC) for daily, weekly and ten-day water-level forecasting at IDEAM's hydrological warning system's stations and defining criteria for scoring the feasibility of a forecast at a given hydrological station and for all forecasting stations. Hydrometeorological information used was obtained from the IDEAM database which stores real-time information on rainfall and water-levels. The analysed forecast points are presented in Table 1. Time-series obtained from the IDEAM database from stations which do not operate in real-time were also used in some cases. The latter are used as forecast predictors in the stations for which no predictors were found amongst stations transmitting real-time information.

Review of quantitative modelling for water in IDEAM forecasts

Quantitative forecasting of water-levels is one of the priority goals for IDEAM's hydrological warning system. Several investigations have been carried out for implementing quantitative forecasts, ranging from qualitative models based on semantic and fuzzy logic rules (Rivera *et al.*, 2004), time-series technique (Mussy, 2005), to advanced models capable of predicting the probability density

En español

pronosticar la curva de densidad probabilística de los niveles diarios en lugar del dato mismo de nivel de agua (Domínguez, 2004a; Domínguez, 2004b; Domínguez, 2005). Estas experiencias presentan las siguientes dificultades: a) no se han realizado en forma articulada, por lo cual también han carecido de criterios comunes de desempeño que permitan medir las ventajas comparativas entre los diferentes operadores de pronóstico; b) por la anterior razón, no existe un estándar para documentar los modelos propuestos, que caracterice: entradas, salidas, operadores, variables de estado, parámetros, mecanismos de calibración; y c) es característica la ausencia de estandarización sobre horizontes de pronóstico, rangos de fecha para validación y calibración, y definición de reglas para el establecimiento de umbrales de pronóstico acertado.

Tabla 1. Puntos de pronóstico analizados.

No	Código	Nombre	Corriente	No	Código	Nombre	Corriente
No	Code	Name	Current	No	Code	Name	Current
1	2312702	Puerto Araujo	Carare	11	2315703	Barrancabermeja	Magdalena
2	2502720	Las Varas	Cauca	12	2502702	El Banco	Magdalena
3	2606701	Juanchito	Cauca	13	2205701	Piedras de Cobre	Saldaña
4	2610707	La Victoria	Cauca	14	2501701	Montelibano	San Jorge
5	2620708	Bolombolo	Cauca	15	2703703	San Juan	San Juan
6	2623704	Pto. Valdivia	Cauca	16	2501702	Pto. Libertador	San Pedro
7	2116706	San Pablo	Cuinde	17	1305701	Nueva Colombia	Sinú
8	2104701	Pte. Balseadero	Magdalena	18	1306702	Monteria	Sinú
9	2123701	Nariño	Magdalena	19	2406703	Puente La Paz	Sogamoso
10	2303701	Puerto Salgar	Magdalena	20	2119701	El Profundo	Sumapaz

In English

curve for daily levels of data itself rather than water-level (Domínguez, 2004a, Domínguez, 2004b; Domínguez, 2005). These have had the following difficulties. They have not been done in coordination and have also lacked common performance criteria for measuring comparative advantages amongst different forecast operators. There has thus been no standard for documenting the models proposed, characterising input, output, operators, state variables, parameters and calibration procedure. There has been a characteristic lack of standardisation regarding forecast horizons, date ranges for calibration and validation and defining rules establishing thresholds for accurate forecasting.

Table 1. Forecast points analysed

Definición de viabilidad de pronóstico hidrológico

Un sistema de pronóstico hidrológico cuantitativo y operativo será viable cuando (WMO, 1994): a) exista la infraestructura necesaria para la medición y transmisión de los niveles del agua en forma continua y en tiempo real; b) se cuente con información fisiográfica, hidrológica, meteorológica, topográfica, detallada y actualizada para los puntos de emisión de pronósticos; c) se disponga del personal capacitado para el análisis de la información asimilada del sistema de medición y transmisión en tiempo real, así como de los especialistas necesarios para la explotación de los modelos de pronóstico hidrológico y la infraestructura computacional para ejecutarlos; d) exista una tecnología de pronóstico/modelación que permita producir las predicciones en tiempo real, con niveles de desempeño aceptables según criterios de validación objetivos; e) estén formalizados los canales de difusión de los pronósticos hidrológicos; f) se disponga de una comunidad de usuarios, oficialmente registrada, capaz de asimilar los pronósticos emitidos y proveer retroalimentación al centro de emisión de pronósticos. Al analizar la infraestructura de mediciones, telecomunicaciones y personal se concluye que de los puntos señalados el Ideam incumple totalmente con el (d) y parcialmente con el (f). Por lo anterior, a continuación se propone una técnica de pronóstico optimizable en tiempo real y adecuado a la disponibilidad y flujos de información de la Red de Alertas del Ideam.

Defining the feasibility of hydrological forecasting

A quantitative hydrological forecasting system can become operational and viable when (WMO, 1994) the following conditions have been met: (a) The necessary infrastructure for real-time water-level measurement and transmission exists; (b) There is available detailed physiographic, hydrological, meteorological and topographic information regarding the points where the forecast is emitted; (c) Trained staff are available for analysing the assimilated real-time hydrological information, as well as operating the hydrological forecasting models and computational infrastructure needed to run these models; (d) There is forecasting technology for making real-time predictions, having acceptable performance levels; (e) There are formalised channels for broadcasting hydrological forecasting; and (f) An officially registered user community has been established and is able to assimilate the forecasts issued and provide feedback for the hydrological warning system. Analysing IDEAM's measurement infrastructure, telecommunications and staff it was concluded that IDEAM's hydrological warning system completely failed to fulfil requirement (d) and partially with condition (f). A real-time forecasting technique is thus proposed which is suitable for the flow of information currently available in IDEAM's hydrological warning system.

En español

In English

Combinaciones lineales adaptativamente óptimas (CLAO) para pronósticos diarios, semanales y decadales de los niveles del agua en las estaciones de la Red de Alertas del Ideam

Un elemento de gran importancia en la emisión de pronósticos hidrológicos es la disponibilidad de un operador matemático $\langle L \rangle$ que utilice predictores disponibles en tiempo real y casi real para la previsión del estado futuro de los niveles en un horizonte de pronóstico T . Existe una gran diversidad de operadores disponibles para tal fin, desde los más complejos hasta los más simples. Como ejemplos de operadores matemáticos para predicción hidrológica se pueden presentar: 1) modelos matemáticos en ecuaciones diferenciales ordinarias (Kovalenko, 1993; Kuchment, 1972b); 2) modelos matemáticos en ecuaciones diferenciales parciales para la descripción del tránsito hidráulico en una y dos dimensiones (Kuchment, 1972a; Rudkivi, 1979; WMO, 1975); 3) modelos matemáticos en ecuaciones diferenciales estocásticas (Gardiner, 1985; Kovalenko, 1993); 4) modelos basados en la teoría de extrainterpolación óptima de Kolmogorov (Kolmogorov, 1941) y en métodos regresivos y autorregresivos (Kazakievich, 1989; Popov, 1957; WMO, 1994); 5) modelos estadísticos/probabilísticos (Haan, 2002; Rozhdientsvienstkiy y Chievatariov, 1974); 6) modelos basados en lógica difusa (Ashu y Avadhnam, 2007; Luchetta y Manetti, 2003); 7) modelos basados en redes neuronales (Ashu y Avadhnam, 2007; Luchetta y Manetti, 2003) y otros.

De una u otra forma estos modelos pueden ser clasificados de la forma propuesta por Domínguez (2007), de modo que éstos, en general, se podrán diferenciar por sus habilidades para representar procesos dinámicos o estáticos, sistemas aglutinados o distribuidos, relaciones causa y efecto determinísticas o estocásticas. El abanico de posibilidades es diverso; por ello, para el caso presente es necesario establecer algunas características mínimas para el operador L . En régimen operativo L debe contar con ciertas características: a) cumplir con los objetivos planteados al sistema de pronósticos y satisfacer las expectativas de los usuarios de los pronósticos; b) ser lo más sencillo posible en su utilización por parte de los pronosticadores de turno, los códigos fuentes del algoritmo de aplicación ser lo suficientemente explícitos para su operación por parte del personal de pronósticos; c) ser coherente con las capacidades informáticas disponibles y con los niveles de precisión de los predictores registrados en tiempo real; contar con un algoritmo operativo de optimización dinámico y adaptable a los diferentes niveles de predictores asimilables en tiempo real y ajustable durante los eventos de pérdida de recepción de alguno de los predictores acoplados; d) ser aplicable a las diferentes condiciones fisiográficas de los puntos operativos de pronóstico.

Atendiendo a estas características, y retomando que según Kazakievich (1989) toda ecuación diferencial puede ser reducida en forma de modelo regresivo o autorregresivo, incluidas las ecuaciones en derivadas parciales (Zwillinger, 1997), a continuación se introduce una técnica de modelación y pronóstico basada en operadores adaptativamente óptimos. Esta técnica establece los operadores de pronóstico óptimos en un espacio lineal y define la ventana de parametrización óptima a través de búsqueda exhaustiva requiriendo tiempos computacionales mínimos, lo cual la hace adecuada como herramienta de pronóstico en sistemas de alerta en tiempo real.

Optimally adaptive linear combinations (OALC) for daily, weekly and ten-day water-level forecasts in IDEAM's hydrological warning system's stations

An important element in the emission of hydrological forecasts is the availability of a mathematical operator $\langle L \rangle$ using real or near-real-time predictors for forecasting the future state of levels for a forecast horizon T . There is a wide range of operators available for that purpose, from the simple to the most complex. Examples of hydrological forecasting mathematical operators would include: (a) Mathematical models in ordinary differential equations (Kovalenko, 1993; Kuchment, 1972b), (b) Mathematical models of partial differential equations for describing hydraulic transit in one and two dimensions (Kuchment, 1972a; Rudkivi, 1979, WMO, 1975); (c) Mathematical models in stochastic differential equations (Gardiner, 1985; Kovalenko, 1993), (d) Models based on the theory of optimal Kolmogorov extra / interpolation (Kolmogorov, 1941); (e) Auto-regressive and regressive methods (Kazakievich, 1989; Popov, 1957, WMO, 1994), (f) Statistical models (Haan, 2002; Rozhdientsvienstkiy and Chievatariov, 1974); (g) Fuzzy logic-based models (Ashu and Avadhnam, 2007; Luchetta and Manetti, 2003); and (h) Neural network-based models (Ashu and Avadhnam, 2007; Luchetta and Manetti, 2003).

These models can be classified as proposed by Domínguez (Domínguez, 2007) so that they can generally be differentiated by their ability to represent static or dynamic processes, lumped or distributed systems, deterministic or stochastic relationships. The range of possibilities is diverse, so some minimum requirements for operator L , had to be established. In fact: (a) L must meet the stated objectives of the hydrological warning system and fulfil forecast users' expectations; (b) It should be as simple as possible in its use by forecasters: the algorithm implementation source code should be explicit enough for operation by staff; (c) It should be consistent with the available computing power and precision levels of the predictors recorded in real-time, (d) It should have an operating optimisation algorithm which must be dynamic and adaptable to water-level data coming in in real-time and adjustable during the event of loss of reception of one of the predictor signals; and (e) It should be applicable to different physiographic conditions of forecast points.

Based on these characteristics, and according to Kazakievich (1989), any differential equation can be reduced to the form of regression or auto-regressive model, including partial differential equations (Zwillinger, 1997). A technique for modelling and forecasting based on adaptive optimal operators is given below. This technique provides optimal forecast operators in a linear space and defines the optimal parameterisation window through exhaustive search requiring minimum computational times, making it suitable as a forecasting tool in real-time early-warning systems.

En español

Usualmente el modelador matemático realiza la calibración de los parámetros de su modelo matemático tratando de utilizar el máximo de información histórica disponible. A continuación se contradice este paradigma demostrando que la calibración dinámica de combinaciones lineales puede crear un operador matemático más eficiente que aquel en el cual sus coeficientes toman valores estáticos basados en la información de toda la serie temporal de la variable que se predice. Esto puede resultar particularmente útil cuando en la serie temporal se registran oscilaciones de diferente escala y frecuencia y para las cuales la predicción está orientada a una prelación de pronóstico no mayor a $T \leq \rho$ con $\rho \ll Ns$ donde T representa el horizonte de pronóstico (prelación del pronóstico), ρ el radio de correlación del proceso estudiado y Ns es la longitud total de la serie de tiempo de la variable a pronosticar.

Dado un intervalo de tiempo $[t - N - 1, t] \in \mathfrak{R}$, en el cual están registrados N valores de la variable $Y(t)$. El pronóstico $Y(t + T)$ puede ser expresado de la siguiente forma:

$$Y(t + T) = L[W(t)] \tag{1}$$

En (1) L es un operador matemático que actúa sobre un polinomio $W(t)$ de orden, que combina predictores endógenos con "m" exógenos en la siguiente estructura:

$$L[W(t)] = \sum_{i=t}^{t-\rho_Y} a_i Y_i^{\alpha_i} + \sum_{k=1}^m \sum_{j=t}^{t-\rho_{X_{kj}Y}} b_{kj} X_{kj}^{\beta_{kj}} \tag{2}$$

donde ρ_Y y $\rho_{X_{kj}Y}$ son el radio de autocorrelación de la variable endógena $Y(t)$ y el radio de correlación cruzada entre ésta y la variable exógena $X(t)_k$. A su vez a_i, b_{kj}, α_i y β_{kj} son coeficientes y exponentes en el polinomio $W(t)$.

Sea la diferencia entre $Y(t + T)^{Observado} - Y(t + T)^{Pronosticado} = \Delta$, se puede denominar a L operador óptimo si éste minimiza alguna función de Δ , por ejemplo, la esperanza matemática de su valor cuadrático:

$$E[\Delta^2] = E\{[Y(t + T) - L[W(t)]]^2\} \tag{3}$$

Cuando es conocido un vector finito de los valores tomados por $Y(t)$, y si su radio de autocorrelación es igual a ρ_Y , el valor buscado para $Y(t + T)$ puede ser representado como una combinación de los valores de $Y(t)$ y de la variable exógena $X(t)$ tomados desde t hasta $t - \rho_Y$ para $Y(t)$ y desde t hasta $t - \rho_{XY}$ de forma que:

$$Y(t + T) = \sum_{i=t}^{t-\rho_Y} a_i Y_i^{\alpha_i} + \sum_{k=1}^m \sum_{j=t}^{t-\rho_{X_{kj}Y}} b_{kj} X_{kj}^{\beta_{kj}} \tag{4}$$

donde a_i, b_i, α_i y β_i son coeficientes y exponentes variables que minimizan

In English

Usually a mathematical modeller calibrates a mathematical model's parameters, trying to use the maximum available historical information. Here, this paradigm has been contradicted by showing that a dynamic calibration of linear combinations can yield a more efficient mathematical operator than that in which the coefficients take static values obtained from information from the entire time-series for the variable being predicted. This can be particularly useful when the recorded time-series exhibit different scale and frequency oscillations and for which prediction is aimed for a lead time no higher than $T \leq \rho$ with $\rho \ll Ns$ where « T » represents the forecast horizon, ρ the process auto-correlation radius and Ns is the total length of the time series of the variable being predicted.

Given a time interval $[t - N - 1, t] \in \mathfrak{R}$ in which N was record length, then forecast $Y(t + T)$ could be expressed as follows:

$$Y(t + T) = L[W(t)] \tag{1}$$

In (1) L is a mathematical operator working on a polynomial $W(t)$ of order $k = \max(\alpha_i, \beta_j)$ using endogenous and "m" exogenous predictors in the form (Kolmogorov, 1941):

$$L[W(t)] = \sum_{i=t}^{t-\rho_Y} a_i Y_i^{\alpha_i} + \sum_{k=1}^m \sum_{j=t}^{t-\rho_{X_{kj}Y}} b_{kj} X_{kj}^{\beta_{kj}} \tag{2}$$

Where ρ_Y and $\rho_{X_{kj}Y}$ were the radius of autocorrelation of the endogenous variable $Y(t)$ and the radius of cross-correlation between the last and the exogenous variable $X(t)_k$. In turn, a_i, b_{kj}, α_i and β_{kj} were the coefficients and exponents in polynomial $W(t)$.

If the difference between $Y(t + T)^{Observado} - Y(t + T)^{Pronosticado} = \Delta$, L could be referred to as an optimal operator if it minimised some function of Δ , for example the expected value of the square mean error:

$$E[\Delta^2] = E\{[Y(t + T) - L[W(t)]]^2\} \tag{3}$$

When a finite vector of values taken by $Y(t)$ was known and if its correlation radius was equal to ρ_Y , then forecast value $Y(t + T)$ could be represented as a combination of values $Y(t)$ and the exogenous variable $X(t)$ taken from t until $t - \rho_Y$ for $Y(t)$ and from t to $t - \rho_{XY}$ so that:

$$Y(t + T) = \sum_{i=t}^{t-\rho_Y} a_i Y_i^{\alpha_i} + \sum_{k=1}^m \sum_{j=t}^{t-\rho_{X_{kj}Y}} b_{kj} X_{kj}^{\beta_{kj}} \tag{4}$$

where a_i, b_i, α_i and β_i were the coefficients and exponents that minimised:

En español

$$\sigma^2(\mathbf{a}, \mathbf{b}, \mathbf{\alpha}, \mathbf{\beta}) = E \left\{ \left[Y(t+T) - \left(\sum_{i=t}^{t-\rho_Y} a_i Y_i^{\alpha_i} + \sum_{k=1}^m \sum_{j=t}^{t-\rho_{X_{kj}Y}} b_{kj} X_{kj}^{\beta_{kj}} \right) \right]^2 \right\}$$

en una ventana de tiempo de ancho θ . De lo anterior se desprende que en el dominio mencionado:

$$\mathbf{a} = \{a_t, a_{t-1}, \dots, a_{t-\rho}\},$$

$$\mathbf{b} = \{b_{k,t}, b_{k,(t-1)}, \dots, b_{k,(t-\rho_{X_{kj}Y})}\},$$

$$\mathbf{\alpha} = \{\alpha_t, \alpha_{t-1}, \dots, \alpha_{t-\rho_Y}\} \quad \mathbf{\beta} = \{\beta_{k,t}, \beta_{k,(t-1)}, \dots, \beta_{k,(t-\rho_{X_{kj}Y})}\}$$

deben ser las soluciones de los sistemas de ecuaciones

$$\begin{aligned} \frac{\partial \sigma^2}{\partial a_i} = 0, \quad i = t, t-1, \dots, t-\rho_Y; \quad \frac{\partial \sigma^2}{\partial b_j} = 0, \quad i = t, t-1, \dots, t-\rho_{X_{kj}Y}; \\ \frac{\partial \sigma^2}{\partial \alpha_i} = 0, \quad k = t, t-1, \dots, t-\rho_Y; \quad \frac{\partial \sigma^2}{\partial \beta_j} = 0, \quad l = t, t-1, \dots, t-\rho_{X_{kj}Y}; \end{aligned} \quad (6)$$

En el dominio de potencia θ también se puede exigir, incluso para combinaciones no lineales, que la relación entre $S = \sqrt{E[\Delta^2]}$ – desviación del error cuadrático de los pronósticos y σ_Δ – desviación estándar de los incrementos de $Y(t)$ para intervalos de tiempo iguales a T sea menor o igual que 0,8, exigiendo a priori el cumplimiento del criterio de desempeño S/σ_Δ tal como éste es utilizado en Domínguez (2004a). Otros criterios de desempeño que pueden ser utilizados como función objetivo son presentados en (Dawson et al., 2007); no obstante, es recomendable definir un juego de criterios de desempeño que sea ortogonal. Por ello, para este trabajo se utilizan el criterio S/σ_Δ , el porcentaje de pronósticos acertados, el error cuadrático estándar y el coeficiente de determinación entre observados y simulados

Selección de predictores

Como predictores se pueden postular la variable endógena $Y(t)$ y las exógenas, por ejemplo $X_i(t)$, rezagadas en $t - m\Delta t$, donde $m = 0 \dots \rho$ con ρ como el radio de correlación cruzada característico con cada predictor. Para tener una visión del bosque y no solo de los árboles es aconsejable construir la matriz de correlación cruzada entre los predictores seleccionados y la variable a pronosticar. En adelante consideraremos que

$$\alpha_t = \alpha_{t-1} = \dots = \alpha_{t-\rho_Y} = \beta_t = \beta_{t-1} = \dots = \beta_{t-\rho_{X_{kj}Y}} = 1 \quad (7)$$

lo que reduce la búsqueda del operador óptimo a un espacio de polinomios de primer orden. Para filtrar los predictores y utilizar sólo los que aportan información no redundante se escogen aquellos cuyos coeficientes $\{a_t, a_{t-1}, \dots, a_{t-\rho}\}$ o $\{b_t, b_{t-1}, \dots, b_{t-\rho_Y}\}$

cumplen la siguiente condición: $(a_k/\sigma_{a_k} \geq 2)$ ó $(b_k/\sigma_{b_k} \geq 2)$, es decir, aquellos con coeficientes a_k o b_k que en magnitud superan dos veces el error cuadrático estándar de definición del coeficiente (σ_{a_k} o σ_{b_k}). Como predictores exógenos, para el caso de pronósticos hidrológicos, se puede utilizar la información de

In English

$$\sigma^2(\mathbf{a}, \mathbf{b}, \mathbf{\alpha}, \mathbf{\beta}) = E \left\{ \left[Y(t+T) - \left(\sum_{i=t}^{t-\rho_Y} a_i Y_i^{\alpha_i} + \sum_{k=1}^m \sum_{j=t}^{t-\rho_{X_{kj}Y}} b_{kj} X_{kj}^{\beta_{kj}} \right) \right]^2 \right\}$$

In a time window of length θ . That stated above would result in the aforementioned domain

$$\mathbf{a} = \{\alpha_t, \alpha_{t-1}, \dots, \alpha_{t-\rho_Y}\} \quad \text{and}$$

$$\mathbf{\beta} = \{\beta_{k,t}, \beta_{k,(t-1)}, \dots, \beta_{k,(t-\rho_{X_{kj}Y})}\}$$

being solutions for the following system of equations

$$\begin{aligned} \frac{\partial \sigma^2}{\partial a_i} = 0, \quad i = t, t-1, \dots, t-\rho_Y; \quad \frac{\partial \sigma^2}{\partial b_j} = 0, \quad i = t, t-1, \dots, t-\rho_{X_{kj}Y}; \\ \frac{\partial \sigma^2}{\partial \alpha_i} = 0, \quad k = t, t-1, \dots, t-\rho_Y; \quad \frac{\partial \sigma^2}{\partial \beta_j} = 0, \quad l = t, t-1, \dots, t-\rho_{X_{kj}Y}; \end{aligned} \quad (6)$$

For a time interval of length θ it might also be required (even for non-linear combinations) that the relationship between

$S = \sqrt{E[\Delta^2]}$, the deviation from the mean square error of forecasts and σ_Δ , the standard deviation of increments of $Y(t)$ for lead time « T » to be less than or equal to 0.8, requiring an a-priori compliance of performance criterion S/σ_Δ as used in Domínguez (Domínguez, 2004a). Other performance criteria which can be used as objective function have been presented by Dawson (Dawson et al., 2007). However, it is advisable to define a set of performance criteria as being orthogonal. Therefore, criterion S/σ_Δ , the percentage of accurate forecasts, the standard square error and the coefficient of determination between observed and simulated data as such set have been used for this study.

Selecting predictors

Endogenous variable $Y(t)$ and exogenous, for example $X_i(t)$, delayed in $t - m\Delta t$, where $m = 0 \dots \rho$ with ρ characteristic cross-correlation radius for each predictor could be postulated as predictors. It was advisable that the cross-correlation matrix between selected predictors and the variable to be predicted should be constructed to get an overview of the forest and not just the trees. Henceforth, the following was considered:

$$\alpha_t = \alpha_{t-1} = \dots = \alpha_{t-\rho_Y} = \beta_t = \beta_{t-1} = \dots = \beta_{t-\rho_{X_{kj}Y}} = 1 \quad (7)$$

This reduced the search for optimal operator to a space of first-order polynomials. Predictors having coefficients $\{a_t, a_{t-1}, \dots, a_{t-\rho}\}$ and $\{b_t, b_{t-1}, \dots, b_{t-\rho_Y}\}$, satisfying constrains $(a_k/\sigma_{a_k} \geq 2)$ or $(b_k/\sigma_{b_k} \geq 2)$ had to be chosen to filter the predictors and only use those that provided non-redundant information. To be precise, only coefficients a_k or b_k greater than twice the standard definition square error of the coefficient (σ_{a_k} or σ_{b_k}) were used. As exogenous predictors, in the case of hydrological forecasts,

En español

precipitaciones actuales y antecedentes y la información de afluencias sobre el tramo antecedente a la estación hidrológica para la que se emiten los pronósticos. En general un algoritmo de búsqueda exhaustiva es capaz de establecer el número de predictores óptimos entre todas las combinaciones disponibles utilizando el criterio S/σ_{Δ} como función objetivo. La minimización del funcional presentado en la ecuación (5) se puede adelantar por el método de mínimos cuadrados, el de gradiente conjugado o inclusive por técnicas bioinspiradas (Yang et al., 2003; Боглаев, 1990). Ejemplos de la potencia del método del gradiente conjugado pueden ser consultados en Press et al. (1986) y en Fylstra et al. (1998). Otros análisis realizados para establecer el número óptimo de predictores son: 1) determinación del número de observaciones equivalentemente independientes y del error corregido del coeficiente de determinación múltiple; 2) evaluación de la evolución del criterio S/σ_{Δ} ; y 3) evaluación del nivel de informatividad de cada grupo de predictores y construcción de la función de saturación. El número de observaciones equivalentemente independientes m_1 se determinó como:

$$m_1 = \frac{m}{\tau_1} \tag{8}$$

donde el radio de autocorrelación τ_1 se define como:

$$\tau_1 = \frac{2}{R(0)} \sum_{\tau=0}^{\infty} R(\tau) \tag{9}$$

Aquí $R(\tau)$ es la función de autocorrelación y $R(0)$ su máximo valor. De este modo, el error del coeficiente de determinación Δ_R , toma en cuenta el número de observaciones equivalentemente independientes, como:

$$\Delta_R = (1 - R_0^2) / \sqrt{m_1} \tag{10}$$

De esta forma el grupo óptimo de predictores se escoge como aquel que minimiza el error del coeficiente de determinación (Comité Hidrometeorológico Estatal de la URSS, 1989). Otra forma de escoger el grupo óptimo de predictores consiste en la evaluación del nivel de informatividad de los predictores por separado y de los distintos grupos de predictores. El grupo óptimo de predictores está conformado por el número mínimo de ellos que logran alcanzar el punto de saturación en la función de ésta. La función de saturación se construye como el nivel de informatividad del grupo de predictores en función del número de variables en el grupo. El nivel de informatividad se define como:

$$I(a_j) = 0,5 \ln \left(\frac{D'}{D} \right) \tag{11}$$

donde D es el determinante de la matriz de correlación entre el nivel y sus predictores y D' es el determinante de la matriz de correlación de los predictores; $I(a_j)$ es el nivel de informatividad del grupo de j predictores (Comité Hidrometeorológico Estatal de la [disuelta] URSS, 1989).

In English

current and lagged rainfall information and the information on river inflows may be used. An exhaustive search algorithm is usually able to establish the optimal number of predictors amongst all available combinations using criterion S/σ_{Δ} as objective function. The minimisation of the function presented in equation (5) may be advanced by the least squares method, the conjugate gradient or even by bio-inspired techniques (Yang et al., 2003; Боглаев, 1990). Examples of the power of the conjugate gradient method can be found in Press (Press et al., 1986) and Fylstra (Fylstra et al., 1998). Other analysis conducted to establish the optimal number of predictors could be by determining the number of independent observations and error equivalently corrected coefficient of multiple determination, evaluating the evolution of criterion S/σ_{Δ} and evaluating the level of informativeness for each predictor group and constructing ion the saturation function. The number of equivalently independent observations m_1 determined as:

$$m_1 = \frac{m}{\tau_1} \tag{8}$$

Where the radius of autocorrelation τ_1 was thus defined as:

$$\tau_1 = \frac{2}{R(0)} \sum_{\tau=0}^{\infty} R(\tau) \tag{9}$$

Here $R(\tau)$ was the autocorrelation function and $R(0)$ its maximum value. Thus, the error of the coefficient of determination Δ_R taking into account the number of equivalently independent observations, was:

$$\Delta_R = (1 - R_0^2) / \sqrt{m_1} \tag{10}$$

The optimal group of predictors was thus chosen as being the group that minimised the standard error for the coefficient of determination (State Hydrometeorological Committee of the USSR, 1989). Another way to choose the best predictors' group was by evaluating the level of informativeness of predictors separately and in different sets of predictors. The best group of predictors was formed by the minimum number of those achieving saturation point in the saturation function. The saturation function was constructed as being the level of informativeness for the predictors' groups with a different number of predictor variables in the group. The level of informativeness was defined as being:

$$I(a_j) = 0,5 \ln \left(\frac{D'}{D} \right) \tag{11}$$

Where D was the determinant of the correlation matrix between water-level values and their predictors and D' was the determinant of the predictors' correlation matrix. $I(a_j)$ represented the informativeness for a group of j predictors (State Hydrometeorological Committee of the USSR, 1989).

En español

In English

Criterio para la calificación de la viabilidad de pronóstico en un punto de análisis y en el conjunto de estaciones hidrológicas estudiadas

Se establecieron distintos niveles de viabilidad para valorar la posibilidad de predecir los niveles diarios, semanales y decadales en términos del criterio S/σ_{Δ} y del porcentaje de aciertos por error máximo permitido. Adicionalmente, se tomó en cuenta si la información de los predictores se encuentra en línea al momento de emitir el pronóstico o no. De acuerdo a los resultados de S/σ_{Δ} la viabilidad de pronóstico se valoró según se presenta en la tabla 2. A la lectura de viabilidad se le anexa la cantidad de aciertos y el nivel de error con el que se definen estos aciertos. Una lectura puede ser "viable con un nivel de aciertos del 70% para errores absolutos promedios del 15%". Este apéndice no influye en la conclusión de viabilidad de la implementación de operadores adaptativos para la predicción de niveles, dado que se desconocen las exigencias de los usuarios de los pronósticos del Ideam. Por otra parte, la predictibilidad del proceso también debe ser tomada en cuenta al momento de establecer el error máximo permitido para establecer el nivel de aciertos de un punto de pronósticos. El error máximo permitido, acorde a la variabilidad natural del proceso, se establece (Апполов et al., 1974) como: $EMP = 0,674\sigma_{\Delta}$.

Tabla 2. Escala de valoración de la viabilidad de aplicación de modelos autorregresivos.

Criterio	Valoración de la viabilidad
$S/\sigma_{\Delta} \leq 0.5$	Alta
$0.5 < S/\sigma_{\Delta} \leq 0.80$	Buena
$0.80 < S/\sigma_{\Delta} \leq 0.90$	Satisfactoria
$S/\sigma_{\Delta} > 0.90$	No viable por el momento

Resultados

Para concluir sobre la viabilidad de los pronósticos de niveles promedios diarios, semanales y decadales se llevaron a cabo más de 300 experimentos numéricos de optimización gruesa. En cada uno de estos experimentos se establecían grupos de predictores de menor potencia que el total disponible al inicio del experimento. Cada experimento representa una búsqueda de entre 540 combinaciones posibles para cada caso, es decir, en total se efectuaron 162.000 ensayos de optimización gruesa. Esta optimización se hizo de forma no supervisada y en promedio tomó alrededor de 4 horas de cómputo en un ordenador de doble procesador con frecuencia de tacto de 2,33 GHz (Intel Pentium D) y dos GB de memoria RAM. No se utilizaron algoritmos de paralelización (vectorización de cálculos) por lo que los tiempos expresados son susceptibles a reducción. Después de la optimización gruesa se realizaron 20 experimentos de sintonización para afinar el grupo de predictores óptimos deducidos de la optimización gruesa. Esta optimización es supervisada y se hace en forma semi-automática. Cada ejercicio de optimización revisa grupos de, en promedio, 3 predictores que se pueden armar en un conjunto de 10 predictores posibles si se toman en cuenta los componentes autorregresivos y los rezagos de correlación cruzada. En teoría, aquí se deberían revisar 120 posibles combinaciones; sin embargo, siguiendo las recomendaciones del parágrafo de selección de predictores este número se reduce drásticamente a unas 10-20 combinaciones. En total el ejercicio de sintonización, para las 20 estaciones seleccionadas, revisó alrededor de 400 combinaciones. En promedio, por estación, en la sintonización se gastaron cerca de 15 minutos, lo que suma un tiempo total para todas las estaciones igual a 5 horas en promedio.

Classification criterion for forecast feasibility assessment at each forecast point and for a set of hydrological stations

Different forecast feasibility levels were established for assessing the viability of predicting daily, weekly and ten-day averaged water-levels in terms of S/σ_{Δ} criterion and taking care of the percentage of successful forecasts using the maximum permitted error (MPE) as threshold. It also took into account whether the information from the predictors could be found online at the time of issuing the forecast. Table 2 gives the results of S/σ_{Δ} criterion forecasting viability. The number of successful forecasts given a maximum permitted error is appended to the feasibility reading. One outcome might have been that the forecasting procedure was viable (70% successfulness) given 15% MPE (mean absolute error). This appendix does not affect the conclusion about the feasibility of implementing optimal adaptive operators for predicting water-levels, since IDEAM forecast users' requirements remain unknown. Moreover, the natural process predictability must also be taken into account when setting the maximum error allowed for setting the level of correct forecasts point. The maximum permissible error, according to the natural variability of the process, was set to (Апполов et al., 1974): $MPE = 0.674\sigma_{\Delta}$.

Table 2. Scale for assessing the feasibility of applying autoregressive models

Criterion	Assessing viability
$S/\sigma_{\Delta} \leq 0.5$	High
$0.5 < S/\sigma_{\Delta} \leq 0.80$	Good
$0.80 < S/\sigma_{\Delta} \leq 0.90$	Satisfactory
$S/\sigma_{\Delta} > 0.90$	Not viable at the moment

Results

More than 300 numerical coarse-tuning experiments were conducted on forecasting feasibility for daily, weekly, and tenday average levels. Smaller sets of predictors were present in each experiment. Each experiment represented a search from 540 possible combinations for each case (i.e. a total of 162,000 coarse optimisation tests were made). Such optimisation was unsupervised and took an average 4 hours computation time on a dualcore processor computer having 2.33 GHz frequency and two GB of RAM. No parallelisation algorithms were used (vectorisation) so specified times were subject to reduction. After the coarse-tuning, 20 experiments were then conducted to refine the set of optimal predictors. Such optimisation was supervised and was performed semi-automatically. Each optimisation exercise reviewed groups of three predictors which could be assembled into a set of 10 potential predictors by taking into account the autoregressive components and cross-correlation lags. In theory this requires reviewing 120 possible combinations; however, following the recommendations given in the predictor selection paragraph, this number was drastically reduced to about 10-20 combinations. The tuning exercise for the 20 selected stations reviewed about 400 combinations. On average, per station, about 15 minutes was spent in tuning, thereby adding 5 hours to setting up the forecast for all stations.

En español

Teniendo una plataforma de integración de información y pronosticadores bien entrenados estos tiempos se pueden reducir a la mitad; no obstante, si se desea emisión de pronósticos en régimen operativo, estas tareas no pueden estar a cargo de un solo pronosticador. Como mínimo se esperaría mantener un grupo de dos pronosticadores. Los tiempos aquí expresados son válidos para el conjunto de 20 estaciones aquí seleccionadas. Un conjunto mayor de estaciones exigiría más tiempo pero con computación en paralelo la tarea es escalable.

En general se puede concluir que la viabilidad de pronóstico disminuye con el aumento del periodo de agregación, es decir, la mejor predictibilidad la tienen los niveles diarios y la peor los decadales. Este juicio es soportado por las magnitudes del criterio S/σ_{Δ} , que en promedio, para todas las 20 estaciones, a nivel diario es igual a 0,740, mientras que a nivel pentadal y decadal es de un valor de 0,803 y 0,797. A su vez, el porcentaje de aciertos, según el error máximo permitido del CHR, es del 81,5% para los decadales, 82,9% para los pentadales y 86,5% para los niveles diarios. El porcentaje de aciertos con errores menores al 15% es del 89,9% para los diarios, 75,7% para los pentadales y 70,3% para los decadales. Para los distintos tipos de pronósticos, en promedio la viabilidad de implementación de operadores tipo CLAO es buena. Si se clasifican los resultados por periodo de agregación, las cuencas con mejor viabilidad de implementación de operadores tipo CLAO para el pronóstico de niveles diarios son las de los ríos Sogamoso, Cauca, Magdalena y Sinú, y las más complicadas las de los ríos San Pedro y Carare, en las que apenas se logró una clasificación satisfactoria de la viabilidad. En el ámbito de los pronósticos semanales los mejores índices los presenta el río Sogamoso, mientras que el Carare se mantiene como la corriente de más baja viabilidad (satisfactoria, según la escala de la tabla 2). Para el caso de los pronósticos decadales los ríos Sogamoso, Cuinde y Cauca se mantienen como los más predecibles, y el San Pedro ocupa el lugar de la predictibilidad más baja. Finalmente, desde el punto de vista del número de pronósticos acertados con un error máximo permitido del 15%, alcanza la cifra del 77% para todas las agregaciones, situación que refuerza la conclusión sobre la viabilidad de implementación de operadores adaptativos para el pronóstico en tiempo real de niveles en la red de estaciones automáticas del Ideam.

In English

Taking an information integration platform and trained forecasters, this computational time could become reduced by half; however, if it were required to set up an operative forecast system, these tasks could not be performed by a single person. As a minimum, it would be expected that a staff of two trained forecasters should be used. The times presented here were valid for all 20 stations selected here. A larger set of stations would require more time; however, a parallel computing approach would be desirable.

It can thus be concluded that the feasibility of prediction decreased by increasing the aggregation period (i.e. the best predictability of levels was obtained on daily data and the worst on ten-day averaged data). This view was supported by the magnitude of criterion S/σ_{Δ} which was 0.740 on average, for all 20 stations, on a daily basis, while being 0.803 and 0.797 at weekly and ten-day aggregations, respectively. In turn, the percentage of successful forecasts according to maximum permitted error was 81.5% for a ten-day averaged water-level forecast, 82.9% for a weekly one and 86.5% for daily water-level forecasts. The percentage of success with errors less than 15% was 89.9% for daily, 75.7% for weekly and 70.3% for the ten-day forecasts. The feasibility of implementing OALC operators was good for the different types of forecasts if the results were sorted by aggregation level, basins having improved feasibility for implementing such operators for forecasting daily levels would be those for the Sogamoso, Cauca, Magdalena and Sinu rivers and the worse would be for the San Pedro and Carare rivers in which only a satisfactory rating of viability was achieved. In the field of weekly forecasting the best feasibility rate was reached by the Sogamoso River, while the Carare river retained the lowest viability (satisfactory according to the scale shown in Table 2). In the case of ten-day forecasts, the Sogamoso, Cuinde and Cauca rivers remained as being the most predictable and the San Pedro river was ranked as having the lowest predictability. A 77% level of successfulness was reached for all forecast types from the standpoint of the number of successful forecasts having a 15% MPE. This fact reinforces the conclusion regarding the feasibility of implementing adaptive operators for real-time forecasting of water-levels using information from IDEAM's real-time hydrological network.

Tabla 3. Viabilidad de pronósticos por cuenca - agregación diaria.
Table 3. Feasibility of daily forecast by basin

Cuenca River basin	S/σ_{Δ}	Errores <15% Errors <15%
Sogamoso	0591	97
Cauca	0679	96
Magdalena	0705	92
Sinu	0767	91
Cuinde	0801	96
Saldaña	0839	69
San Jorge	0848	83
Sumapaz	0854	80
San Pedro	0888	95
Carare	0890	61
Average	0.7861	86

Tabla 4. Viabilidad de pronósticos por cuenca - agregación semanal.
Table 4. Feasibility of weekly forecast by basin

Cuenca River basin	S/σ_{Δ}	Errores <15% Errors <15%
Sogamoso	0405	99
Cauca	0798	67
Cuinde	0807	97
Magdalena	0809	80
Sinu	0824	74
San Pedro	0873	91
Saldaña	0876	64
San Jorge	0878	66
Sumapaz	0879	72
Carare	0881	51
Average	0.8030	76

Conclusiones

La viabilidad de pronósticos diarios, semanales y decadales en los nodos de la red de alertas hidrológicas del Ideam es buena. Más del 70% de las estaciones hidrológicas analizadas reportaron valores de desempeño $S/\sigma_{\Delta} \leq 0,85$ y con un porcentaje de aciertos mayor al 70% para un error máximo permitido del 15%.

Conclusions

The feasibility of forecasting daily, weekly and ten-day averaged water-levels at the nodes of IDEAM's real-time hydrological network was good. Over 70% of the analysed hydrological stations reported $S/\sigma_{\Delta} \leq 0.85$ performance values and 70% successfulness for a 15% MPE.

En español

La técnica de pronóstico analizada, tanto en los tiempos de cómputo como en el error de pronóstico, ofrece buen desempeño para las distintas agregaciones analizadas. La metodología de pronóstico propuesta permite determinar que las estaciones de la red hidrológica deben ser modernizadas e incluidas en la red de transmisión en tiempo real. Para los casos puntuales en los que la viabilidad de pronóstico no sea satisfactoria, aún existen posibilidades de mejoramiento mediante la inclusión de predictores no considerados.

Tabla 5. Viabilidad de pronósticos por cuenca - agregación decadal.

Cuenca	S/σ_{Δ}	Errores <15%
Sogamoso	0.426	91
Cuinde	0.736	97
Cauca	0.770	63
Carare	0.812	56
Magdalena	0.826	75
San Jorge	0.848	36
Saldaña	0.852	65
Sinú	0.862	63
Sumapaz	0.867	72
San Pedro	0.903	88
Promedio	0.7901	70

Bibliografía / References

- Ashu, J., Avadhnam, M. K., Hybrid Neural Network Models for Hydrologic Time Series Forecasting., Applied Soft Computing, 2007, pp. 585-592.
- Comité Hidrometeorológico Estatal de la URSS, C., Directrices para pronósticos hidrológicos - pronósticos de corto plazo de caudales y niveles del agua en ríos., 1989, pp. 246.
- Dawson, C., W., Abraham, R., J., See, L., M., HydroTest: A web-based toolbox of evaluation metrics for the standardised assessment of hydrological forecasts., Environmental Modelling & Software, No. 22, 2007, pp.1034-1052.
- Domínguez, E., Aplicación de la ecuación de Fokker-Planck-Kolmogorov para el pronóstico de afluencias a embalses hidroeléctricos (caso práctico de la represa de Betania), Meteorología Colombiana, No 8, 2004a, pp.17-26.
- Domínguez, E., Stochastic forecasting of streamflow to Colombian hydropower reservoirs., PhD. Thesis, Russian State Hydrometeorological University, San Petersburg, 2004b, pp. 235.
- Domínguez, E., Pronóstico probabilístico de afluencias para la evaluación de riesgos en embalses hidroeléctricos., Avances en Recursos Hidráulicos, 2005, pp. 12-25.
- Domínguez, E., Introducción a la modelación matemática., Googlepages., Bogotá, 2007.
- Fylstra, D., Lasdon, L., Watson, J., Waren, A., Design and use of the Microsoft Excel Solver., Computers/Computer Science-Software, No. 28, 1998, pp. 29-55.
- Gardiner., Handbook of stochastic methods., Springer-Verlag, Berlin, 1985, pp. 442.
- Haan, T. C., Statistical methods in hydrology., Iowa state press, Iowa, 2002, pp.378.
- Kazakievich, D. I., Osnovi teorii sluchainij funktsiiv v zadachax gidrometeorologii., Guidrometeoizdat, Leningrad, 1989, pp. 230.
- Kolmogorov, A. N., Interpolirovanie y Extrpolirovanie Statsionarnij Sluchainij Posliedovatielnostiei., Bulletin De l' Academie Des Sciences De l'URSS, No. 5, 1941, pp. 3-14.
- Kovalenko, V., Modelling of hydrological processes. Guidrometeoizdat., Saint Petersburg, 1993, pp. 255.

In English

The forecasting technique analysed for both computational time and forecast error offered good performance for the different aggregations. The proposed forecasting methodology led to determining which hydrological stations should be upgraded and included in the real-time transmission system. For specific cases in which the viability of forecast was not satisfactory there is still space for improvement by including predictors which were not considered here.

Table 5. Feasibility of ten-day forecast by basin

River basin	S/σ_{Δ}	Errors <15%
Sogamoso	0.426	91
Cuinde	0.736	97
Cauca	0.770	63
Carare	0.812	56
Magdalena	0.826	75%
San Jorge	0.848	36
Saldaña	0.852	65
Sinu	0.862	63
Sumapaz	0.867	72
San Pedro	0.903	88
Average	0.7901	70

- Kuchment, L. S., Matematicheskoe Modelirovanie Rechnova Stoka., Guidrometeoizdat, Leningrad, 1972a.
- Kuchment, L. S., Modelación matemática de la escorrentía fluvial., Guidrometeoizdat, Leningrado, 1972b, pp. 191.
- Luchetta, A., Manetti, S., A Real-time Hydrological Forecasting System using a Fuzzy Clustering Approach., Computers & Geosciences, No. 29, 2003, pp. 1111-1117.
- Mussy, A., Short Term Hydrological Forecasting Model In Colombia: Simulation For The Magdalena River., IDEAM., Lausanne, 2005.
- Perrault, P., De l'Origine Des Fontaines. Pierre Le Petit., Paris, 1674.
- Popov, E. G., Guidrologicheskie Prognozi, Guidrometeoizdat., Leningrad, 1957.
- Press, W. H., Teukolsky, S. A., Vetterling, W. T., Flannery, B. P., Numerical Recipes in Fortran 77 The art of Scientific Computing., Cambridge University Press, No. 2, pp. 999.
- Rivera, H., Zamudio, E., Romero, H., Modelación con fines de pronósticos hidrológicos de los niveles diarios en periodo de estiaje en los sitios de Calamar, El Banco y Puerto Berrio del Magdalena., Avances en Recursos Hidráulicos, No11, 2004.
- Rozhdientsvienskiy, A. B., Chievatariov, A. I., Statisticheskii Metodi v Guidrologuii., Guidrometeoizdat, Leningrad, 1974, pp. 424.
- Rudkivi, A. J., Hydrology. An advanced introduction to hydrological modelling., Pergamon Press, Sydney, 1979, pp. 479.
- WMO., Intercomparison of Conceptual Models Used in Operational Hydrological Forecasting., Operational Hydrology Report, No7, WMO-No. 429, WMO, Geneva, 1975.
- WMO., Guide to hydrometeorological practices., 168, WMO, Geneva, 1994, pp. 770.
- Yang, Z. R., Thomson, R., Hodgman, T.C., Dry, J., Doyle, A. K., Narayanan, A., Wu, X. Searching for discrimination rules in protease proteolytic cleavage activity using genetic programming with a min-max scoring function., Biosystems, 72(1-2), 2003, pp. 159-176.
- Zwillinger, D., Handbook of differential equations., Academic press, Boston, 1997.
- Апполов, Б., Калинин, Г., Комаров, В., Курс гидрологических прогнозов, 1. Гидрометеоиздат, Ленинград, 1974, pp. 419.
- Боглаев, Ю.П., Вычислительная математика и программирование, 1. Высшая школа, Москва, 1990, pp. 544.