

APLICACIÓN DE UN PROCESADOR BAYESIANO PARA LA EVALUACIÓN DE LA INCERTIDUMBRE PREDICTIVA MEDIANTE COMBINACIÓN DE MODELOS HIDROLÓGICOS EN LA PREDICCIÓN DE CRECIDAS EN TIEMPO REAL

Juan Camilo Múnera¹, Gabriele Coccia², Félix Francés¹, Ezio Todini²

¹Instituto de Ingeniería del Agua y el Medio Ambiente, Universidad Politécnica de Valencia, España

²Departamento de Ciencias de la Tierra y Geológico - Ambiental, Universidad de Bologna, Italia

juamuesl@doctor.upv.es, gabriele.coccia@gmail.com, f frances@hma.upv.es, ezio.todini@unibo.it

Introducción

En este trabajo se presentan los resultados derivados de la aplicación de un método Bayesiano para evaluar la incertidumbre predictiva en situaciones de predicción de crecidas en tiempo real, el cual se ha denominado *Model Conditional Processor* (MCP) (Todini, 2008). La predicción de crecidas en tiempo real requiere tener en cuenta la incertidumbre predictiva (IP) asociada a la variable de interés (el caudal o nivel futuro en el Río) por las razones siguientes: los modelos hidrológicos o hidráulicos de tipo determinístico o de otras características, como pueden ser los modelos basados en los datos, los modelos de redes neuronales artificiales y/o de lógica difusa, o cualquier modelo híbrido entre las tipologías anteriores, pueden proporcionar información útil sobre la ocurrencia futura de alguno de estos eventos, pero los resultados de dichos modelos no se pueden asumir como el valor futuro real de los mismos, como usualmente se suele hacer en la práctica, ya que tales previsiones están afectadas por diversas fuentes de incertidumbre que van desde los errores asociados a los datos observados a los errores de modelación y sus condiciones de contorno o de estado inicial.

Las previsiones obtenidas a partir de un modelo como los mencionados anteriormente son de naturaleza inherentemente incierta, pero al mismo tiempo, deben ser utilizadas por los tomadores de decisiones con la intención de reducir la incertidumbre sobre la ocurrencia futura de un evento de cierta magnitud. La IP se define, en efecto, como la probabilidad de ocurrencia de un valor futuro de un predictando (la variable a predecir, tal como el nivel, caudal, volumen, etc.) condicionada sobre todas las observaciones y conocimientos que se tienen a priori del fenómeno analizado, así como a todas las informaciones que se pueden obtener de este valor futuro específico a partir de las previsiones obtenidas con los modelos. Esta IP se debe cuantificar en términos de una función de distribución de probabilidad que podrá ser usada por los encargados de la gestión de emergencias en sus procesos de toma de decisiones, con el objetivo de mejorar la calidad y confiabilidad de las mismas.

La metodología propuesta del MCP posibilita evaluar la IP combinando tanto las incertidumbres de tipo meteorológico (previsión de la lluvia futura) como hidrológicas (modelos de transformación lluvia - escorrentía), mediante la derivación, para varios horizontes temporales de predicción, de la función de densidad de probabilidad conjunta multivariada entre la variable observada y las previsiones de uno o varios modelos. Esto se hace transformando las observaciones y las previsiones a un espacio multi-Normal mediante la operación *Normal Quantil Transform* (NQT), descrita en detalle por Kelly y Krzysztofowicz (1997), con la intención de generar nuevas variables con distribución marginal Gaussiana.

Según el Teorema de Bayes, a partir de esta distribución de

probabilidad conjunta en el campo transformado es posible calcular la densidad de probabilidad de la variable observada condicionada a las previsiones de los diferentes modelos:

$$f(\eta/\hat{\eta}_1, \dots, \hat{\eta}_n) = f(\eta, \hat{\eta}_1, \dots, \hat{\eta}_n) / f(\hat{\eta}_1, \dots, \hat{\eta}_n) \quad [1]$$

Esta distribución tiene como valor medio:

$$\mu_{\eta/\hat{\eta}} = \sum_{\eta} \sum_{\hat{\eta}} \eta \hat{\eta}^{-1} \hat{\eta} \quad [2]$$

Su varianza resulta ser:

$$\sigma^2_{\eta/\hat{\eta}} = 1 - \sum_{\eta} \sum_{\hat{\eta}} \hat{\eta} \hat{\eta}^{-1} \sum_{\eta} \eta \hat{\eta}^T \quad [3]$$

De la función de densidad condicionada se estima el valor esperado de la previsión en el campo Normal:

$$\eta^* = E[f(\eta/\hat{\eta}_1 = \hat{\eta}_1^*, \dots, \hat{\eta}_n = \hat{\eta}_n^*)] \quad [4]$$

Los intervalos de confianza para el nivel de significancia requerido, por ejemplo, para un 95%, se calculan como:

$$\eta^*_{95} = E[f(\eta/\hat{\eta}_1 = \hat{\eta}_1^*, \dots, \hat{\eta}_n = \hat{\eta}_n^*)] \pm 1.96 \sigma_{\eta/\hat{\eta}} \quad [5]$$

Para obtener una cuantificación real de la IP, es necesario convertir los valores de los cuantiles calculados en el campo Normal a los correspondientes valores en el campo original mediante la transformación inversa NQT⁻¹.

Finalmente, en este trabajo se ha explorado la posibilidad de utilizar una función de densidad de probabilidad conjunta multivariada truncada, buscando mejorar la adaptación al separar los datos en dos muestras correspondientes a los caudales bajos y altos, respectivamente. Por otro lado, la combinación de modelos de diferente naturaleza busca sacar ventaja de sus características particulares y de sus diferentes capacidades para reducir la incertidumbre predictiva.

Caso de estudio

La segunda fase del proyecto de intercomparación de modelos hidrológicos distribuidos DMIP2 (Smith et al, 2004), propuesto por el National Weather Service (NOAA/NWS) surge del interés de esta institución en incorporar modelos distribuidos para la predicción de avenidas en la gestión de los ríos en EE. UU.

Para la aplicación de la metodología propuesta se ha seleccionado la subcuenca Baron Fork, Eldon del Río Illinois con un área de drenaje de 795 km² en el estado de Oklahoma. Esta cuenca se caracteriza por un clima semiárido con muy poca precipitación en forma de nieve y con presencia de tormentas convectivas significativas.

Descripción de los modelos de previsión

A continuación se describen los tres modelos hidrológicos de previsión utilizados en el estudio para la evaluación de la incertidumbre predictiva:

TETIS es un modelo lluvia – escorrentía de tipo conceptual con base física y distribuido en el espacio, que permite simular los procesos hidrológicos más relevantes en la cuenca, incluyendo la variabilidad espacial de las componentes del balance, y se puede utilizar para un amplio espectro de escalas temporales. Adicionalmente, tiene una novedosa estructura de parámetros que hace una distinción entre la variabilidad espacial de las características físicas a escala de cuenca, representadas mediante mapas de parámetros estimados a priori a partir de la información medioambiental disponible, y por otro lado, los parámetros efectivos que se deben emplear en el modelo a escala de celda, considerando los posibles errores en la información de entrada, las características particulares de la cuenca y los efectos de escala espacial y temporal. La relación entre los mapas de parámetros estimados a priori y los mapas efectivos se considera como una función de corrección, que en su forma más simple, constituye un factor corrector (FC) que modifica globalmente los primeros. Esta estructura diferenciada reduce drásticamente el número de parámetros a calibrar al conjunto de FC posibilitando su calibración automática (Vélez, 2003; Francés et al, 2007; Vélez et al, 2009).

TOPKAPI es un modelo lluvia – escorrentía físicamente basado y distribuido en el espacio que se basa en la agregación de diferentes aproximaciones de onda cinemática en el suelo, en la superficie del terreno y en la red de drenaje, que conducen a simular los procesos de producción de escorrentía y traslación de la misma mediante tres ecuaciones diferenciales de embalses no lineales. La percolación hacia la capa de suelo profundo se simula mediante la introducción de una segunda capa de suelo con características diferentes a las del estrato más superficial que involucra movimiento de agua en dirección vertical hacia el acuífero (Liu et al, 2005).

El tercer modelo es un modelo modular de Redes Neuronales Artificiales (ANN) como el propuesto por Pujol (2009), que combina una red auto-organizativa de Kohonen (*Self Organizing Map* SOM) con alguno de 3 modelos ANN del tipo *Multi Layer Perceptron* (MLP), correspondientes a condiciones hidrológicas diferenciadas utilizados para hacer la previsión. El primer modelo se utiliza para realizar la clasificación automática en grupos de datos, y está basado en una técnica estadística a partir de las variables de entrada elegidas sin una definición previa de las clases. Los 3 modelos ANN MLP tienen arquitecturas similares pero con diferente configuración de las variables en la capa de entrada y número de nodos en la capa oculta.

Referencias Bibliográficas

- Francés F., Vélez J. I., Vélez J. J. (2007). Split-parameter structure for the automatic calibration of distributed hydrological models. *Journal of Hydrology*, 332, pp226– 240.
- Kelly, K. S., and R. Krzysztofowicz. (1997), A bivariate meta-Gaussian density for use in hydrology. *Stochastic Hydrol. Hydraul.*, 11, 17– 31.
- Liu, Z., Martina, M.V.L. and Todini, E. (2005). Flood Forecasting Using a Fully Distributed Model: Application of the TOPKAPI model to the Upper Xixian Catchment. *Hydrol. Earth Syst. Sci.*, 9, 347–364.

Pujol, L. (2009). Predicción de caudales en tiempo real en grandes cuencas utilizando redes neuronales artificiales. Tesis doctoral, Universidad Politécnica de Valencia, Departamento de Ingeniería Hidráulica y Medio Ambiente. 288P.

Smith, M.B., Seo, D.-J., Koren, V.I., et al. (2004). The distributed model intercomparison project (DMIP): an overview. *Journal of Hydrology* 298(1–4), 4–26.

Todini, E. (2008). A model conditional processor to assess predictive uncertainty in flood forecasting. *Intl. J. River Basin Management*. Vol. 6 (2), 123-137.

Vélez, J. I. (2003). Desarrollo de un modelo distribuido de predicción en tiempo real para eventos de crecidas. Tesis doctoral, Universidad Politécnica de Valencia, Departamento de Ingeniería Hidráulica y Medio Ambiente. 266P.

Vélez, J.J., Puricelli, M., López Unzu, F., and Francés, F. (2009). Parameter extrapolation to ungauged basins with a hydrological distributed model in a regional framework, *Hydrol. Earth Syst. Sci.*, 13, 229-246.