

Análisis del riesgo de inundación mediante técnicas estadísticas avanzadas

Navas, S.^{a1}; del Jesus, M.^{a2}

^{a1}Instituto de Hidráulica Ambiental, Universidad de Cantabria - Avda. Isabel Torres, 15, Parque Científico y Tecnológico de Cantabria, 39011, Santander, España. E-mail: ^{a1}salvador.navas@unican.es, ^{a2}manuel.deljesus@unican.es.

Línea temática | Hidrología, usos y gestión del agua.

RESUMEN

En este estudio se plantea una metodología para analizar el riesgo de inundación a partir de la generación de series de precipitación sintéticas distribuidas por toda la cuenca y técnicas geoestadísticas, tomando como datos de partida series de precipitación real existentes. Realizada la reconstrucción de las series temporales sintéticas, se realizan simulaciones del comportamiento hidrológico de la cuenca. De las series de caudal obtenidas, se seleccionan aquellos eventos más significativos para realizar una simulación hidráulica y computar sus calados y velocidades de inundación. El resto de eventos se resuelven mediante técnicas híbridas de reducción de escala y métodos de minería de datos. Para calcular la inundación producida para un determinado periodo de retorno, la estadística de extremos se lleva a cabo sobre el calado y la velocidad de la inundación, no sobre la precipitación. La metodología propuesta muestrea de forma más exhaustiva el espacio de posibilidades, proporcionando mapas de riesgo de inundación que responden de manera más fidedigna a la información histórica.

Palabras clave | Geoestadística; Minería de datos; Regresión probabilística; Simulación sintética; Inundación; Riesgo; LIDAR; Base Topográfica Armonizada

INTRODUCCIÓN

Las inundaciones son el desastre natural que se manifiesta con mayor frecuencia en España, dando lugar a graves situaciones de riesgo colectivo. Para la protección de personas y bienes ante los efectos destructivos de las inundaciones, resulta necesario elaborar estudios preventivos de inundación que incluyan mapas de peligrosidad, vulnerabilidad y riesgo. La ocupación de terrenos inundables incrementa la afección que las inundaciones tienen sobre la población.

Los Corrales de Buelna es uno de los municipios españoles que más inundaciones ligadas a crecidas fluviales ha sufrido a lo largo de la historia. Los desbordamientos del río Besaya son los causantes de estos episodios de inundación. Estos desbordamientos están ligados a la dinámica natural del río, debido a la corta longitud de la cuenca y el fuerte desnivel que presenta desde su nacimiento, lo que reduce notablemente el tiempo de concentración en este tramo medio del río y agiliza su respuesta frente a episodios de precipitación. Esta problemática de Los Corrales de Buelna será el objeto de estudio.

La metodología habitualmente utilizada para caracterizar el riesgo de inundación, tiene como objetivo calcular la magnitud de una inundación tomando como datos de partida los hietogramas asociados a distintos periodos de retorno, obtenidos a partir de la estadística de extremos de la precipitación, asumiendo una respuesta homogénea en toda la cuenca. El problema de utilizar esta metodología surge al no incluir en el análisis variables que pueden afectar a la respuesta de la cuenca, como el grado de saturación de distintas zonas de la cuenca, la distribución espacial de la precipitación o la forma del hidrograma resultante del evento.

Objetivos

El presente estudio persigue los siguientes objetivos:

- Generación de series temporales de precipitación distribuidas espacialmente sobre la cuenca de estudio a partir de técnicas geoestadísticas.
- Generación sintética de series temporales de caudal a partir de métodos de clasificación y regresión probabilística.
- Selección de eventos sintéticos para la simulación hidráulica.
- Modelado hidráulico bidimensional de la zona de estudio.
- Reconstrucción de las manchas de inundación producidas por los eventos sintéticos generados.
- Comparativa de las inundaciones producidas con ambas metodologías.

MATERIAL Y MÉTODOS

Información disponible

- Información topográfica de datos LIDAR y BTA (Base Topográfica Armonizada).
- Series de precipitación diarias (Gobierno de Cantabria, 2014).
- Datos foronómicos diarios (Gobierno de Cantabria, 2014).

Área de estudio

El estudio de riesgo de inundación se centra en la localidad cántabra de Los Corrales de Buelna. Esta localidad se encuentra asentada en la cuenca del río Besaya, a 12 kilómetros de la ciudad de Torrelavega y a 32 kilómetros de la capital cántabra, Santander.

La población de Los Corrales de Buelna está muy influenciada en todos sus aspectos por el río Besaya. El fondo del valle es su territorio y en los aluviones del mismo se han ido desarrollando los principales asentamientos del municipio (Ayuntamiento de los Corrales de Buelna, 2009).

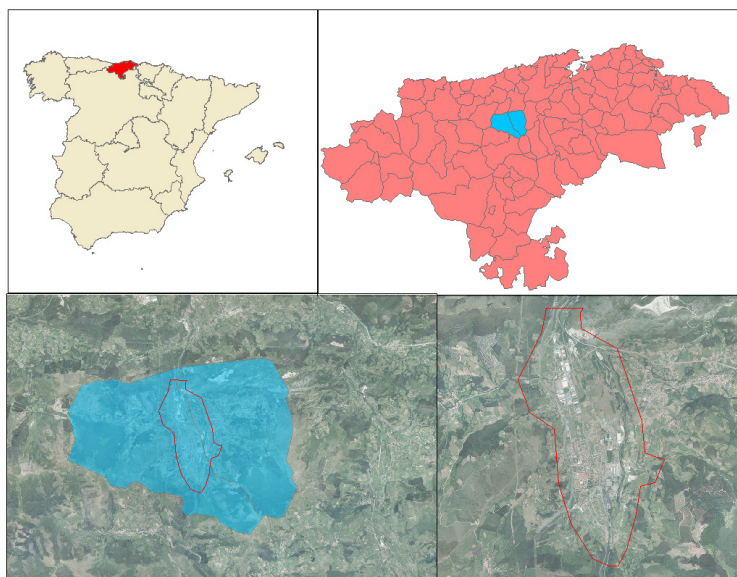


Figura 1 | Localización de la zona de estudio.

Metodología

La metodología parte de la mejora de la información de precipitación disponible mediante técnicas geoestadísticas con el objetivo de generar series pluviométricas sintéticas en puntos de la cuenca sin información. Para ello se analizan distintas técnicas y se elige la que mejores resultados proporciona. Una vez que se tienen datos de precipitación distribuidos, se realiza un modelado hidrológico para obtener series temporales de caudal en el cauce más cercano a la población de Los Corrales de Buelna. Tras obtener la serie de caudal diario, se seleccionan aquellos eventos que producen inundación, pero debido a que el número de eventos que se seleccionan es reducido, se generan sintéticamente miles de ellos con el objetivo de aumentar el abanico de eventos posibles que producen inundación. Ya que el número de eventos que se generan son demasiados para poder simular hidráulicamente uno a uno y obtener las manchas de inundación, se seleccionan aquellos eventos más representativos, obteniendo sus calados y velocidades a partir del modelado hidráulico bidimensional. Para poder obtener los calados y velocidades asociados a un período de retorno, se reconstruyen sintéticamente los calados y velocidades de los eventos sintéticos generados. Por último, se calculan las inundaciones asociadas a un período de retorno mediante su probabilidad empírica.

En la siguiente imagen se representa el proceso de forma esquemática:

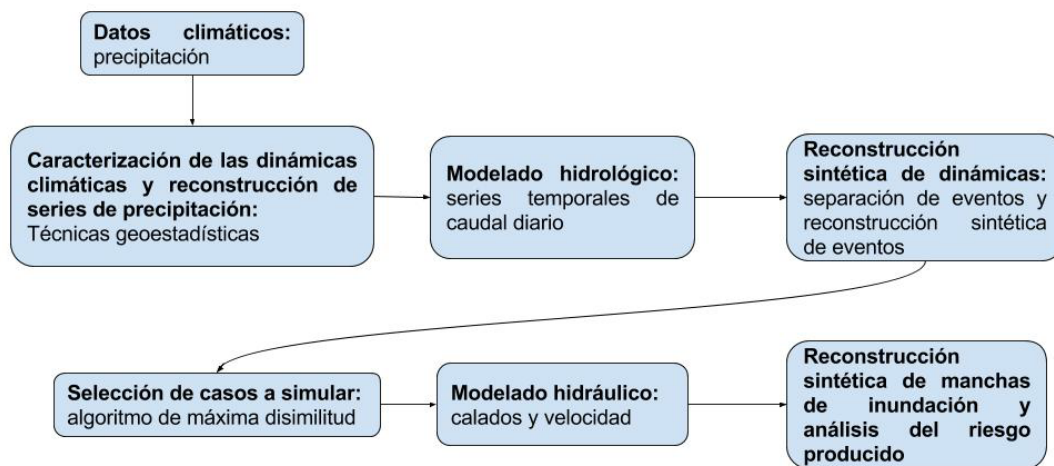


Figura 2 | Esquema del proceso a seguir en la metodología propuesta

Caracterización de las dinámicas climáticas

Para la implementación de un modelo hidrológico es necesario caracterizar las diferentes variables climáticas que definen el comportamiento hidrológico de una cuenca. En este estudio sólo se tiene en cuenta la variable de precipitación. Además, debido a la escasa densidad de pluviómetros sobre la cuenca, se utilizan diferentes técnicas geoestadísticas para obtener una mayor cantidad de puntos con información. Para el presente estudio se han analizado tres técnicas:

- Kriging (Willian, 2011)
- Universal-Kriging (Willian, 2011)
- IDW (Inverse Distance Weighting) (Villatoro et al., 2008)

Una vez realizada la interpolación, se eligen los resultados del modelo que mejor predice la precipitación. Para ello, se lleva a cabo la validación cruzada de los valores observados contra los valores predichos mediante la representación de la correlación, desviación típica y error cuadrático medio en gráficos de Taylor. En la Figura 2 se muestra un ejemplo del análisis realizado sobre una de las estaciones pluviométricas localizadas en la cuenca de estudio.

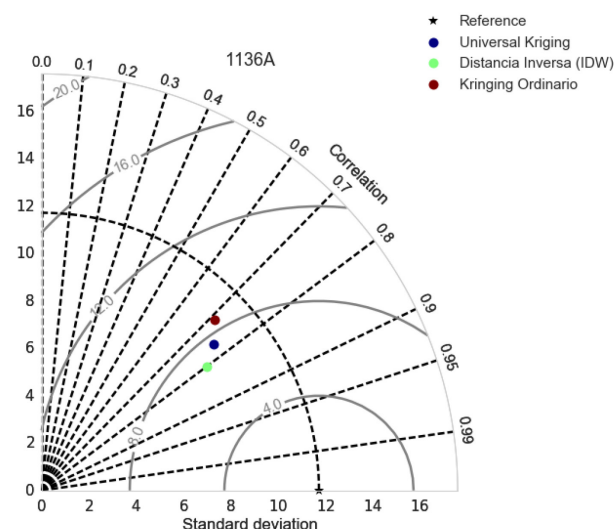


Figura 3 | Comparativa de las técnicas geoestadísticas utilizadas

Las diferencias entre los resultados obtenidos con los tres métodos utilizados son mínimas, debido principalmente a la escasez de pluviómetros y en el caso del método Universal Kriging en el que se utiliza la elevación como *drift*, a la relación inexistente entre la altura y la precipitación en las series disponibles. En este tipo de técnicas ofrecen buenos resultados cuando existen covariables altamente relacionadas con la precipitación en la zona de estudio (Wagner y col., 2012).

Analizando de forma más precisa los resultados de cada método, el mejor ajuste se obtiene en la interpolación mediante distancia inversa. Por lo tanto, se utilizan las lluvias obtenidas del método de interpolación IDW en el modelo hidrológico.

Modelado hidrológico

El modelo hidrológico utilizado en el presente proyecto ha sido HEC-HMS (Scharffenberg and Fleming, 2010). En este estudio se ha utilizado el método de Soil Moisture Accounting para pérdidas por infiltración. Este método considera la existencia de tres capas (capa de suelo, capa de agua subterránea 1 y capa de agua subterránea 2) las cuales permiten representar la dinámica del movimiento del agua entre ellas. Para transformar la lluvia neta en caudal, se ha utilizado el método de Clark, cuyo principio se basa en que el hidrograma total de una avenida es la suma de todos los hidrogramas aportados por las distintas subcuencas, debidamente modificados por el efecto de almacenamiento en el río. Y por último, el método de tránsito utilizado ha sido Muskingum. Este método calcula el hidrograma de una avenida en su tránsito por el cauce teniendo en cuenta el efecto de laminación. Tiene en cuenta dos parámetros, uno relacionado con la capacidad de almacenamiento del cauce en cada tramo, de carácter adimensional, (X), y otro relacionado con el tiempo que tarda la onda de avenida en desplazarse por el tramo considerado, con unidades de “tiempo”, (k).

Los parámetros del modelo hidrológico han sido calibrados y validados mediante las series de caudales instrumentales. En la Figura 4 se puede observar la comparativa entre los caudales reales en la estación de aforo más cercana a Los Corrales de Buelna y los caudales simulados.

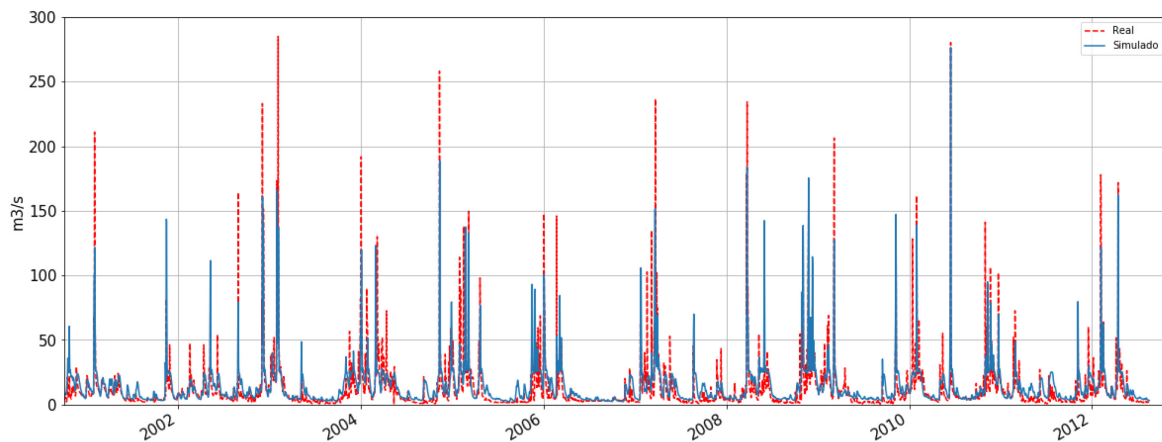


Figura 4 | Comparativa entre caudales reales del aforo más cercano a la zona de estudio y lo caudales simulados

Reconstrucción sintética de hidrogramas

Una vez simulada la serie de caudales mediante el modelado hidrológico, en primer lugar, es necesario separarla en eventos estableciendo un umbral a partir del cual comienza a producirse el desbordamiento del río Besaya. Se considera que un evento comienza en el punto en la pendiente se transforma en creciente siempre que en dicha subida se sobrepase el umbral prefijado. El punto final del evento es el punto donde termina la primera curva de recesión que cruce el umbral. La separación de eventos se realiza ya que es inviable la simulación completa de una serie de 42 años de datos diarios en un modelo hidráulico bidimensional debido al tiempo computacional requerido. Cada evento se caracteriza por 4 parámetros:

- Caudal máximo (Q_{max})
- Caudal medio (Q_{med})
- Duración (T)
- Tipo de hidrograma según la forma

En la Figura 3 se muestra un ejemplo de la separación de eventos realizada sobre la serie de caudal.

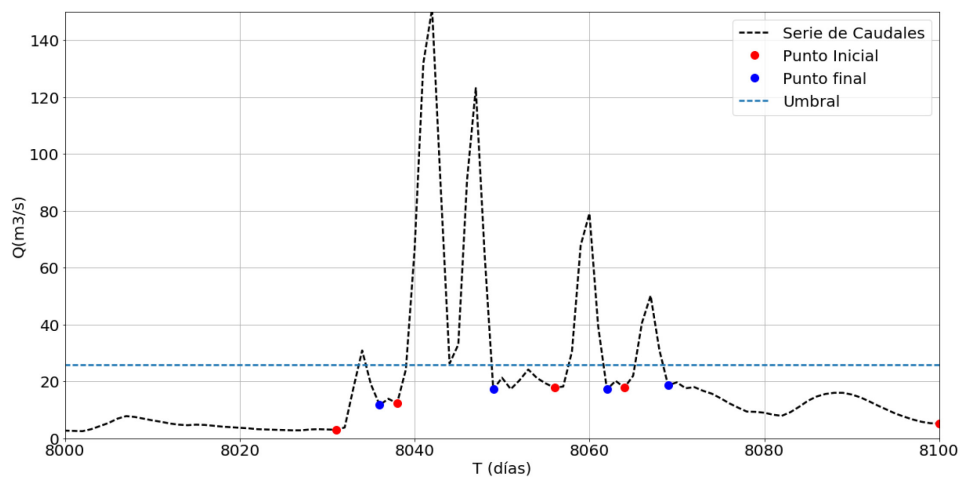


Figura 5 | Ejemplo separación de la serie de caudales en eventos

Tras la separación en eventos, éstos se clasifican en grupos según la forma del hidrograma combinando dos métodos:

- Método de las componentes principales (PCA) (Smith, 2002)
- Algoritmo K-Means (Camus et al., 2011)

La aplicación de ambos métodos da como resultado una colección de X hidrogramas tipo, que son los que mejor representan la variabilidad en la forma de los hidrogramas presentes en la serie simulada. Una vez separados los eventos de la serie de caudal y clasificados según la forma del hidrograma, se dispone de una colección de eventos representativos de la serie simulada. Estos eventos pueden utilizarse para generar eventos sintéticos a través de regresión probabilística mediante Cópulas Gaussianas las cuales son capaces de guardar la correlación existente entre los cuatro parámetros seleccionados (Ben Alaya et al., 2014); de esta forma se simulan series sintéticas de mayor longitud que las observadas.

Selección de casos a simular

Puesto que el número de eventos sintéticos generados es demasiado grande y no pueden simularse hidráulicamente cada uno de ellos, es necesario seleccionar un número más reducido, asegurando que los eventos sintéticos seleccionados abarcan todo el rango de datos extremos probables. La selección se realiza a través del algoritmo de máxima disimilitud (MaxDiss) (Camus et al., 2011).

En la Figura 4 se muestran los eventos sintéticos generados mediante las cópulas gaussianas (puntos azules) y los seleccionados para la simulación hidráulica (puntos rojos), caracterizados por el caudal medio, caudal máximo y duración del evento. Puede observarse como el conjunto de puntos rojos cubre adecuadamente el rango de variación total con un número de puntos más reducido.

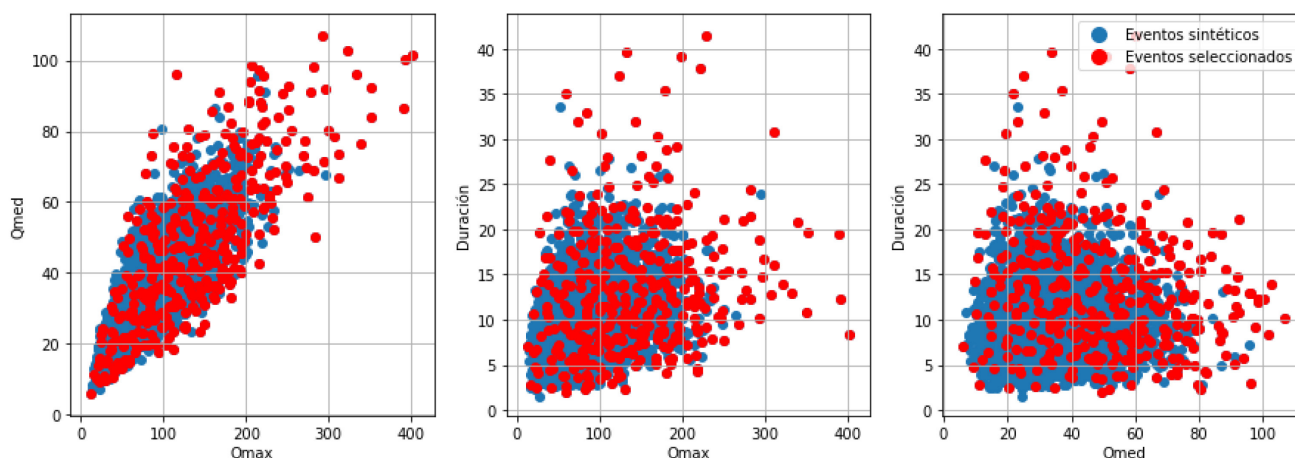


Figura 6 | Reconstrucción y selección de eventos.

Una vez seleccionados los eventos sintéticos a simular hidráulicamente, es necesario reconstruir los hidrogramas en función de los cuatro parámetros que lo definen (Q_{max} , Q_{med} , T y Tipo de hidrograma). El primer objetivo es dar la forma del hidrograma tipo al evento sintético.

Para reconstruir los hidrogramas de cada evento atendiendo a su forma y que los parámetros Q_{med} , Q_{max} y T sigan caracterizando a cada uno de ellos, se recurre a técnicas de ajuste mediante polinomios de grado 2. Los pasos a realizar para reconstruir los hidrogramas son los siguientes:

1. Discretización de hidrogramas originales en unidades de tiempo de segundos.

En primer lugar deben discretizarse los hidrogramas de los eventos iniciales de cada tipo, ya que son éstos los que van a dar la forma a cada uno de los sintéticos. Para que el ajuste de forma sea más preciso, es necesario discretizar la duración del hidrograma en segundos y realizar una interpolación lineal de los caudales para tener el mismo número de datos.

2. Interpolación mediante polinomios de grado 2 para la obtención de los caudales sintéticos instantáneos.

El polinomio de grado dos que se utiliza para realizar la reconstrucción es la siguiente:

$$Q_i^* = a(Q_i)^2 + b(Q_i) \quad (1)$$

Siendo Q_i^* el caudal para un instante i del evento sintético y Q_i el caudal del evento inicial tipo. a y b son los parámetros que definen el polinomio, obtenidos a partir de la relación existente entre ecuaciones:

$$b = \frac{Q_{max}^* - a(Q_{max})^2}{Q_{max}} \quad (2)$$

$$b = \frac{NQ_{med}^* - a \sum_{i=1}^N Q_i^2}{\sum_{i=1}^N Q_i} \quad (3)$$

Donde N es igual al número total de datos de caudal diario correspondiente a cada evento sintéticos generado, n el números de datos de caudal de cada uno de los eventos reales tipo. Igualando ambas ecuaciones y despejando la variable a se obtiene:

$$a = \frac{NQ_{med}Q_{max} - Q_{max}^* \sum_{i=1}^N Q_i^2}{\sum_{i=1}^N Q_i^2 Q_{max} - Q_{max}^2 \sum_{i=1}^N Q_i} \quad (4)$$

3. Duración del hidrograma sintético.

Tras obtener los caudales de los eventos sintéticos para cada instante de tiempo, es necesario establecer la duración total del hidrograma. Para ello se aplica la siguiente expresión:

$$t_i^* = \frac{t_i}{T_{max}} T_{max}^* \quad (5)$$

Donde t_i^* es el instante i del evento sintético, t_i el instante i del evento inicial tipo, T_{max}^* la duración total del evento sintético y T_{max} la duración total del evento inicial tipo.

En la Figura 7 se muestra un ejemplo de las reconstrucciones realizadas para distintas tipologías de hidrograma según su forma.

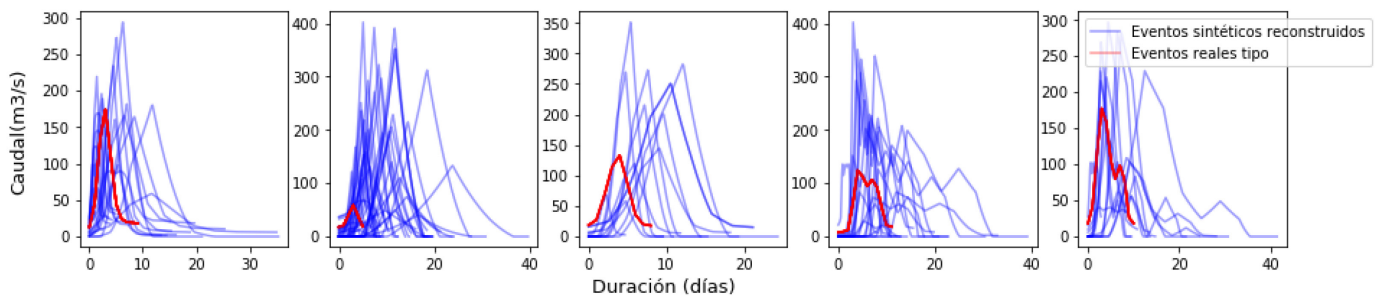


Figura 7 | Reconstrucción de la forma de los eventos sintéticos seleccionados.

Modelado hidráulico

El modelo utilizado para la simulación hidráulica es Iber, modelo numérico bidimensional de simulación de flujo turbulento en lámina libre en régimen no permanente y de procesos medioambientales en hidráulica fluvial (FLUMEN; CIMNE; CEDEX; GEAMA y CIMEN, 2010). Para comenzar el modelado hidráulico es necesario incorporar la topografía de la zona a partir de un modelo digital del terreno. En este caso, se utiliza el MDT generado a partir de datos LIDAR con una densidad de puntos de 0.5 puntos/m². La generación de la malla de cálculo con IBER se realiza a través de la herramienta “RTIN”, capaz de crear e importar una geometría formada por una red de triángulos rectángulos a partir de un Modelo Digital del Terreno en formato ASCII. En este sentido, se fijó como límite una tolerancia máxima de importación de 0.5 m, acorde con la precisión altimétrica del vuelo LIDAR. Asimismo, se limitó el valor del lado máximo de los triángulos generados a 8 m.

Los hidrogramas de entrada son los obtenidos a partir de la selección de los eventos sintéticos más representativos tras ser reconstruidos con paso de tiempo diario. Dichos hidrogramas son introducidos de forma tabulada en la selección de elementos perimetrales de la malla correspondiente entrada.

Reconstrucción sintética de manchas de inundación

Para reconstruir las manchas de inundación ligadas a un periodo de retorno, han de reconstruirse en primer lugar las series temporales de calado de inundación y velocidad de flujo en cada celda de la malla. Para ello se utiliza un método de interpolación basado en la distancia inversa de los k vecinos más próximos. En este caso se toman N vecinos. De esta forma, en cada punto, se toman los resultados de la simulación hidráulica de Iber de los vecinos más próximos y se asigna dicho valor al punto analizado. Una vez que este proceso se ha repetido para todos los eventos, se dispone de las series temporales de cota de inundación y de velocidad del flujo para toda la serie sintética.

La obtención del número de vecinos más próximos a utilizar, para reconstruir las manchas de inundación, se realiza a través de un proceso iterativo en el cual son reconstruidos las manchas de inundación de los 10 últimos eventos simulados en Iber, a partir de los 150 primeros, variando el número de vecinos cercanos. En cada iteración se va comprobando el error total, es decir, la suma de los errores de cada pixel. Este error en cada pixel es calculado como la diferencia entre el valor real de inundación simulado y el valor reconstruido. El número óptimo de vecinos cercanos, será aquel que produzca un menor error acumulado. Como se puede observar en la Figura 8 el número óptimo es 6.

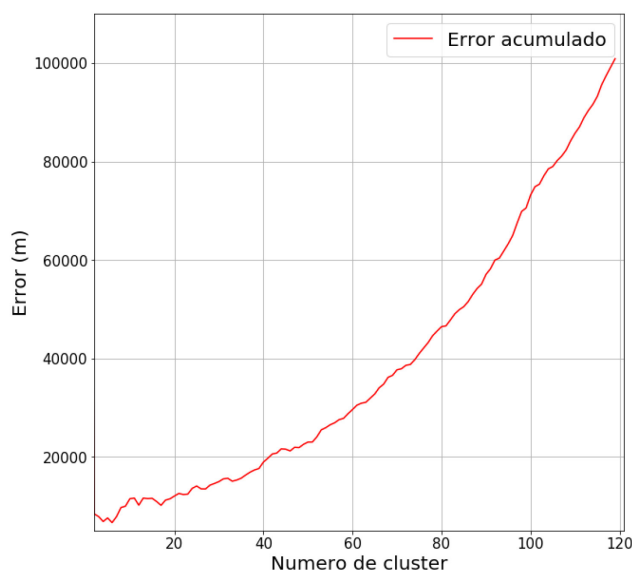


Figura 8 | Error acumulado según el número de vecinos más próximos (clúster)

Tras realizar la reconstrucción de las manchas de inundación, se analiza cual hubiese sido el número de simulaciones máximas a realizar en Iber. Se ha seguido el mismo proceso que para la obtención del número de vecinos cercanos; a partir de las 150 primeras simulaciones, se reconstruyen las 10 últimas, en este caso variando el número de simulaciones de 1 a 150. Se puede observar en la Figura 9 que a partir de las 110 simulaciones el error comienza a estabilizarse.

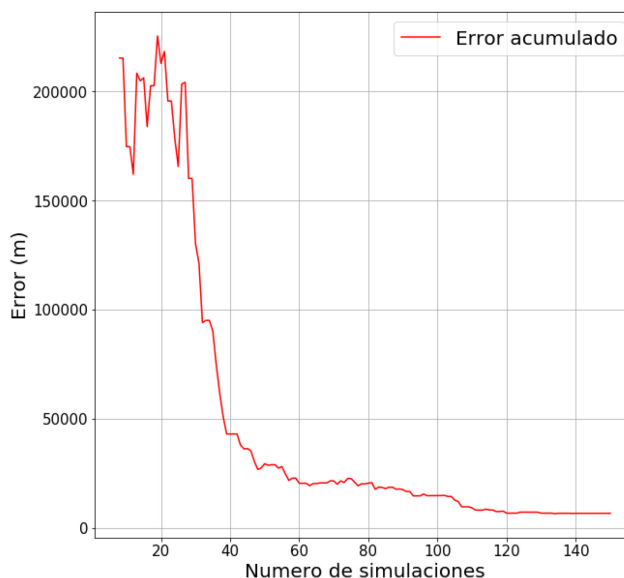


Figura 9 | Error acumulado según el número de simulaciones en Iber

A partir de aquí, se realiza un ajuste de extremos empírico pixel por pixel en toda la zona de estudio, definiendo la cota de inundación y la velocidad del flujo en cada pixel del área de estudio de forma independiente.

El riesgo de inundación es analizado una vez que se tienen los resultados de calados y velocidad. Aplicando la metodología propuesta en (MAGRAMA, 2013) y utilizando tratamientos de datos raster, se analiza el riesgo de inundación como puede observarse en la Figura 10.

RESULTADOS

En los resultados obtenidos del modelo hidrológico, se puede observar la reducción de caudales máximos que se produce en la serie simulada. Esta reducción es de en torno al 50%-60% como puede observarse en la Figura 4. Por esta razón la generación de eventos sintéticos mediante cópulas es imprescindible para cubrir el rango de caudales máximos que no se refleja en el resultado del modelo hidrológico.

La utilización de cópulas Gaussianas permite obtener eventos con caudales máximos similares a los de la serie real, que son los que influyen directamente sobre la magnitud de la inundación.

Las inundaciones obtenidas con la metodología propuesta tienen una mayor extensión que las generadas con la metodología habitual para un mismo período de retorno, lo que permite quedar por el lado de la seguridad. Esta diferencia se debe principalmente a la escasa densidad de pluviómetros y a la corta longitud de las series de precipitación utilizadas para obtener las intensidades de lluvia correspondientes a cada período de retorno en el uso de la metodología habitual, además de no incluirse variables explicativas como la forma del hidrograma de los eventos de inundación y la distribución espacial de la precipitación.

Por ello, la estadística aplicada no recoge al completo las dinámicas posibles del río, dando lugar a manchas de inundación reducidas. Sin embargo, en la metodología propuesta se recogen todas las dinámicas posibles al realizar la reconstrucción sintética generan manchas más extensas y por lo tanto zonas con mayor riesgo de inundación.

Es muy interesante observar cómo la mancha generada para un período de retorno de 10 años coincide con la llanura de inundación del río Besaya en el año 1946. De esta forma se demuestra que la inundación generada es realmente probable.

En la Figura 7 se muestra la comparativa entre los resultados obtenidos mediante la metodología habitual y la que se plantea.

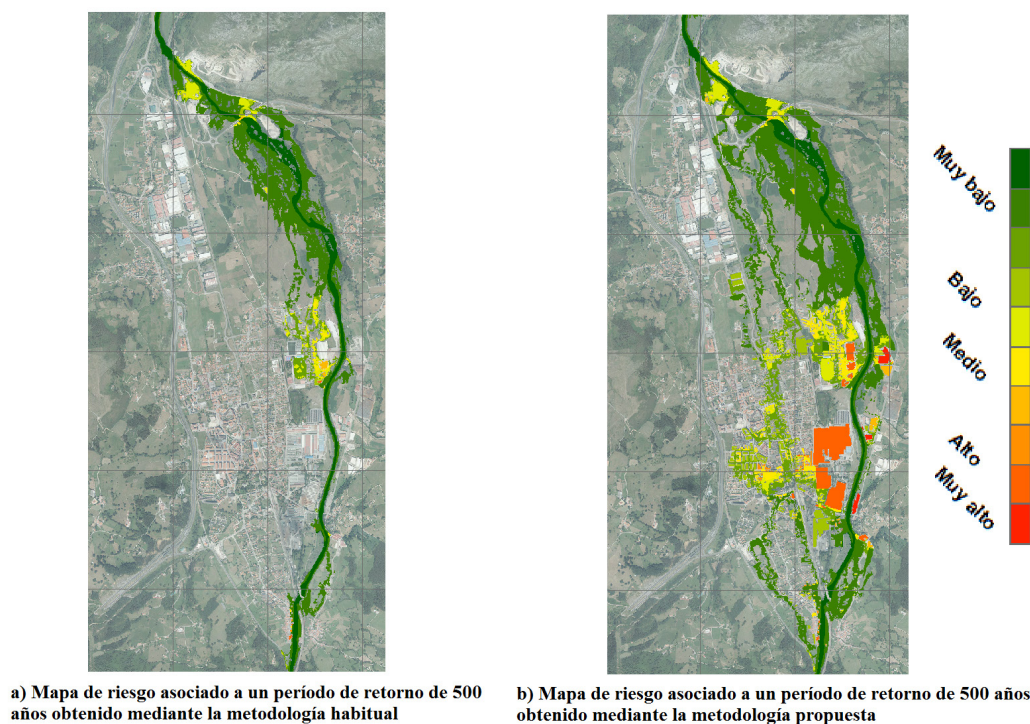


Figura 10 | Comparativa en los resultados de los mapas de riesgo obtenidos mediante la metodología planteada y la metodología habitual.

DISCUSIÓN

Para caracterizar adecuadamente el clima de una zona, es necesario tener series de datos climáticos largas y cierta densidad de estaciones meteorológicas distribuidas por el territorio. Series cortas y un escaso número de estaciones impiden realizar estudios precisos.

En estudios realizados en Estados Unidos se ha demostrado que una alta resolución espacial de datos de precipitación es esencial para poder simular picos de inundación. Series de precipitación cortas y escasas pueden llegar a subestimar en un 50% - 60% el valor de los picos de inundación (Michaud and Sorooshian, 1994). A pesar de realizar la reconstrucción de series climáticas en puntos donde no existe dato, sigue existiendo cierta incertidumbre que va a influir sobre el modelado hidrológico. Además, aunque se utilicen modelos más complejos, no quiere decir que la predicción en los puntos sin datos vaya a ser más precisa, sino que depende de la longitud de series y el número de pluviómetros; cuanto mayor es la calidad de los datos mejor puede ser el ajuste. Por esta razón, es imprescindible la generación de eventos sintéticos para cubrir el abanico de caudales máximos que producen inundación y evitar que se produzca la subestimación de ésta.

La metodología propuesta requiere de más medios informáticos y tiempo computacional que la metodología habitual; se han utilizado diversos equipos informáticos trabajando simultáneamente durante dos semanas, sin embargo, un estudio de inundaciones preciso y detallado puede evitar afecciones graves en una población durante episodios de inundación, siempre y cuando la distribución territorial futura se desarrolle fuera de las zonas inundables delimitadas en los estudios de inundación.

Cuando se tienen series amplias de precipitación ambas metodologías deberían de parecerse. En futuras líneas de investigación se estudiará este tipo de casos verificando el ajuste entre ambas metodologías.

CONCLUSIONES

En los estudios de inundación se requiere de datos de precipitación con alta resolución espacial y temporal para disminuir la incertidumbre en el cálculo de la extensión de inundación. Un método para disminuir la incertidumbre en este tipo de estudios es recurrir a la generación sintética de caudales para cubrir el abanico de caudales máximos que producen inundación y evitar que se produzca la subestimación de la inundación, en el caso en el que los datos pluviométricos hayan sido escasos en el modelado hidrológico.

Cuando las inundaciones no se encuentran del lado de la seguridad y se comparan con las extensiones fluviales mediante ortofotos históricas se observa cómo las manchas de inundación quedan muy por debajo de las llanuras de inundación, sin embargo, con una metodología como la que se propone en este estudio, se reproduce mejor las extensiones probables de la macha de inundación.

REFERENCIAS

- Ayuntamiento de los Corrales de Buelna. *Plan General de Ordenación Urbana de Los Corrales de Buelna*. 2009.
- M. A. Ben Alaya, F. Chebana, and T. B M J Ouarda. Probabilistic gaussian copula regression model for multisite and multivariable downscaling. *Journal of Climate*, 27(9):3331–3347, 2014.
- Paula Camus, Fernando J. Mendez, Raul Medina, and Antonio S. Cofiño. Analysis of clustering and selection algorithms for the study of multivariate wave climate. *Coastal Engineering*, 58(6):453–462, 2011.
- Gobierno de Cantabria. *Investigación, desarrollo e innovación en materia de agua (2012-2014). Estudio de recursos*. 2014.
- Daniel T. Larose. *k-Nearest Neighbor Algorithm*, pages 90–106. 2005.
- MAGRAMA. *Propuesta de mínimos para la metodología de realización de los mapas de riesgo de inundación*, 2013.
- Jene Michaud and Soroosh Sorooshian. Comparison of simple versus complex distributed runoff models on a midsized semiarid watershed. *Water Resources Research*, 30(3):593–605, 1994.
- William a. Scharffenberg and Matthew J. Fleming. Hydrologic Modeling System User ' s Manual. page 318, 2010.
- Lindsay I Smith. A tutorial on Principal Components Analysis Introduction. *Statistics*, 51:52, 2002.
- Mario Villatoro, Carlos Henríquez, and Freddy Sancho. Comparación de los interpoladores idw y kriging en la variación espacial de ph, ca, cice y p del suelo. *Agronomia Costarricense*, 32(1):95–105, 2008.
- Caballero Guardo Willian. Kriging Universal para Datos Funcionales. 2011.
- FLUMEN; CIMNE; CEDEX; GEAMA y CIMEN. *Manual Básico de Usuario de Iber*, 2010.
- Wagner, P.D. y col. «Comparison and evaluation of spatial interpolation schemes for daily rainfall in data scarce regions». En: *Journal of Hydrology* 464-465, páginas 388-400. 2012