

Caracterización de regímenes extremos multivariados de inundación considerando la variabilidad natural de las dinámicas inductoras

Manuel del Jesus, Paula Camus, Fernando J. Méndez, Iñigo J. Losada
*Instituto de Hidráulica Ambiental de la Universidad de Cantabria
Calle Isabel Torres 15, PCTCAN. 39011 Santander*

J.D. Fernández Bethencourt, P. Delgado Melián
*Consejo Insular de Aguas de Tenerife
Calle Leoncio Rodríguez 7, 38003 Santa Cruz de Tenerife*

1. Introducción

La teoría de valores extremos proporciona una herramienta de gran utilidad para la caracterización de los regímenes extremos de variables ambientales ya que permite mitigar las dificultades inducidas por dos de los principales problemas que pueden encontrarse al tratar con extremos: la escasez de medidas y la limitada longitud de las series. La teoría de valores extremos permite, a partir de una serie limitada de observaciones, predecir los valores extremos de una variable ligados a un determinado periodo de retorno mayor que el periodo de observación.

En aplicaciones ingenieriles normalmente no se dispone de medidas de la variable de interés (la escorrentía superficial si tratamos con un problema de drenaje o la profundidad de inundación en evaluaciones de riesgo, por ejemplo). La metodología más habitual consiste en buscar medidas de alguna variable relacionada (por ejemplo la precipitación en el caso de la escorrentía), hacer la caracterización de extremos sobre dicha variable y mediante simulación transformar dichos extremos en los de la variable de interés.

Cuando la relación entre la variable objetivo y la variable relacionada (normalmente variable inductora) es sencilla, y siempre que esta última tenga un régimen estocástico univariado, es decir, que su variabilidad dependa únicamente de un único mecanismo, el procedimiento descrito en el párrafo anterior puede conducir a buenas aproximaciones al régimen extremal real de la variable de interés.

Sin embargo, las relaciones entre la variable objetivo y la inductora no siempre son sencillas, ni la variabilidad de la variable inductora puede explicarse por un único mecanismo. En estos casos, la aplicación del procedimiento habitual, en que el régimen extremal de la variable de

interés se determina mediante el de la variable inductora, puede llevar a subestimar los periodos de retorno de la variable de interés con el consiguiente riesgo que esto supone.

En este estudio presentamos una metodología para la caracterización del régimen extremo de inundación en un canal que descarga en la costa. En este caso la cota de inundación va a depender de dos variables, la descarga del canal y el nivel del mar en el desagüe; la cota de inundación será por tanto tratada como una variable bivariada y la caracterización de valores extremos será realizada sobre esta última variable y no sobre las inductoras.

Como una de las variables inductoras a caracterizar, la precipitación, presenta un patrón de variabilidad complejo, la metodología hace uso de técnicas de minería de datos para separar las observaciones iniciales en tormentas similares (tipos de tiempo (Bermejo y Ancell 2009)), lo que a su vez permite construir subconjuntos de observaciones homogéneas para un mejor ajuste de las distribuciones marginales.

La disgregación en tipos de tiempo permite a su vez, mediante simulación de Monte Carlo, generar series temporales de la longitud deseada que reproducen la estadística de las observaciones en cada una de las distribuciones marginales, pero incluyendo también eventos que si bien no han sido observados son plausibles en base a las observaciones. Este proceso de simulación permite extender la cobertura temporal de las observaciones y generar series sintéticas de mayor duración, dónde la estimación de grandes periodos de retorno es más estable. Ha de destacarse también que la disgregación en tipos de tiempo permite, mediante la caracterización estadística de patrones atmosféricos espacio-temporales, incluir la variabilidad natural atmosférica y su influencia sobre la variable analizada, lo que dota de gran robustez al procedimiento.

La metodología hace uso de un método de “downscaling” híbrido, dinámico y estadístico, para transformar la serie inicial de eventos combinados (precipitación y nivel del mar) en una serie temporal de la variable de interés (cota de inundación), sobre la que se aplican las distintas técnicas matemáticas para derivar el régimen extremo que se busca.

En este trabajo se hace además una aplicación de esta metodología al canal de desagüe de los barrancos de Huertas y el Cercado en San Andrés, Tenerife. En esta aplicación destaca el valor añadido de la metodología al comparar las estimaciones obtenidas mediante nuestra metodología y mediante un método clásico.

2. Metodología

2.1 Descripción general

El objetivo de la presente metodología es la caracterización del régimen extremal de la cota de inundación en un canal que descarga en el mar a partir de la información recopilada sobre precipitación y nivel del mar. Un esquema de la metodología se muestra en la **Figura 1**.



Para incluir la variabilidad natural de estas variables en el análisis se caracterizan las distribuciones marginales de cada variable condicionadas a la ocurrencia de unos patrones sinópticos de referencia (Bermejo y Ancell 2009). Estos patrones sinópticos agrupan las duplas precipitación y nivel del mar en subgrupos homogéneos, mejorando la bondad del ajuste de las observaciones a las distribuciones teóricas utilizadas. La distribución conjunta de las variables inductoras (precipitación y nivel del mar) se modela por medio de cópulas Gaussianas de las distribuciones marginales ajustadas.

Una vez ajustadas las distribuciones conjuntas (Ben Alaya, Chebana, y Ouarda 2014) condicionadas a cada patrón sinópticos de referencia (o tipo de tiempo) se procede mediante simulación de Monte Carlo a generar una serie sintética de eventos combinados de precipitación y nivel del mar de la longitud deseada y que será representativa de las condiciones del periodo

de observación escogido.

Figura 1 Esquema de la metodología Mediante la técnica de máxima disimilitud (Camus et al. 2011) (MDA por sus siglas en inglés) se seleccionan de la serie completa los 250 eventos más representativos de toda la serie, es decir, los 250 eventos que mejor cubren todo el rango de variabilidad de ambas variables.

Se realiza una simulación numérica (“downscaling” dinámico) de estos 250 eventos. La precipitación se transforma en esorrentía y caudal mediante un modelo hidrológico basado en el número de curva siguiendo las indicaciones marcadas en la Guía Metodológica del Consejo Insular de Aguas de Tenerife. Los caudales de descarga y niveles del mar son transformados en cotas de inundación en el canal mediante la aplicación del modelo “InfoWorks Integrated Catchment Modeling” (Jamieson et al. 2012).

Estas 250 simulaciones sirven para construir mediante la técnica de funciones de base radiales (Buhmann 2003) (RBF por sus siglas en inglés) una función de transferencia que transforma las duplas precipitación – nivel del mar en cota de inundación. Aplicando esta función de transferencia a la serie completa de eventos combinados precipitación – nivel del mar (“downscaling” estadístico) se genera una serie sintética de cotas de inundación en el canal de la misma longitud que la serie de eventos sintéticos generada mediante Monte Carlo. El régimen extremal de la cota de inundación puede computarse mediante la función de distribución empírica de la serie sintética generada.

2.2 Bases de datos

Los datos pluviométricos utilizados en el presente estudio han sido extraídos de la base de datos del Consejo Insular de Aguas de Tenerife. Esta base de datos combina información de la Agencia Estatal de Meteorología (AEMET) y de AgroCabildo, del Cabildo de Tenerife. Se han utilizado los registros de precipitación acumulada diaria para el periodo que comprende desde 1980 hasta 2010. Debido a la reducida extensión de la cuenca para el caso de estudio se ha interpolado una serie de precipitación diaria en el centroide de la cuenca.

Los datos de nivel del mar se han extraído de la base de datos océano-meteorológicos del proyecto MARUCA (MARUCA web) (Caracterización Climática del Medio Físico Marítimo Español para la Optimización de la Explotación Portuaria y Navegación Marítima). Los datos proporcionan tanto información de marea astronómica como sobreelevación no astronómica ligada a la ocurrencia de eventos de marea de tormenta y de oleaje. Se han generado serie de nivel diarias con el máximo nivel alcanzado en cada una de las componentes.

Para la generación de los tipos de tiempo se ha utilizado la base de datos de reanálisis SEAWIND(Menendez et al. 2014) de 30Km de resolución espacial, generada mediante el modelo WRF(Done, Davis, y Weisman 2004) y los datos del reanálisis global NCEP-NCAR(Kalnay et al. 2010). Concretamente se han utilizado los datos diarios de presión en superficie y altura del geopotencial de 500hPa para definir los tipos de tiempo de referencia.

2.3 Generación de tipos de tiempo

La generación de tipos de tiempo se basa en la aplicación del algoritmo K-Medias(Camus et al. 2011) (K-Means). Este es un algoritmo de agrupamiento que parte una muestra de datos en un número predeterminado de subgrupos homogéneos. Sigue un algoritmo iterativo que minimiza la varianza intra-grupo. Cada grupo queda representado por su centroide, que se computa como la media de todo el conjunto de datos que componen el grupo.

Para la aplicación del algoritmo K-Means ha de procederse en primer lugar a la reducción de la dimensionalidad de los datos mediante la técnica de análisis de componentes principales(Camus et al. 2011) (PCA por sus siglas en inglés). El PCA toma la serie temporal de patrones atmosféricos (que es en realidad una matriz con unas dimensiones iguales al número de pasos de tiempo multiplicadas por el número de pixels espaciales) y devuelve una matriz con el mismo número de pasos de tiempo pero con un número de componentes espaciales reducidos tales que aseguran que se conserva un porcentaje determinado de la variabilidad de los datos en la serie de partida. El algoritmo PCA sirve para reducir el número de componentes con que se representa una determinada información, eliminando principalmente ruido para mantener la mayor parte de la señal.

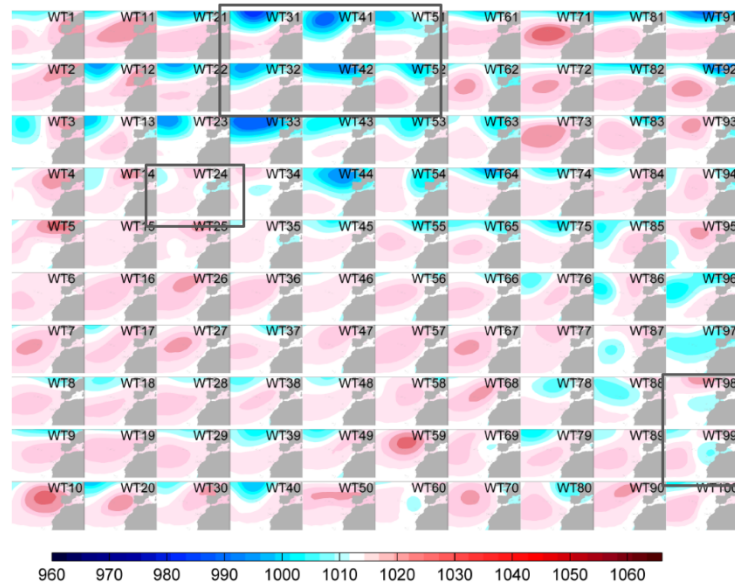


Figura 2 Tipos de tiempo para la presión en superficie. Presiones medidas en hPa.

Una vez se ha aplicado el algoritmo PCA a los datos, se aplica el algoritmo K-Means y se obtiene un agrupamiento de las situaciones sinópticas originales en el número de grupos preestablecidos. Esta agrupación permite asignar probabilidades de ocurrencia (mediante el simple conteo de elementos por grupo) de cada tipo de tiempo.

En este estudio los tipos de tiempo se han caracterizado a través de los campos combinados de presión en superficie y altura del geopotencial de 500hPa. La **Figura 2** muestra los tipos de tiempo para la presión en superficie, mientras que la **Figura 3** muestra los tipos de tiempo en la altura del geopotencial de 500hPa. Es importante remarcar que cada tipo de tiempo es la combinación del mapa de presión en superficie y el de altura del geopotencial. Situaciones similares en presión en superficie, por ejemplo WT10 y WT59, se diferencian por sus patrones en la altura del geopotencial.

2.4 Distribuciones marginales y cópulas Gaussianas

Para cada tipo de tiempo se procede al ajuste de las distribuciones marginales de la precipitación, que se ajustan a distribuciones exponenciales, y del nivel del mar, suma de la componente astronómica y la meteorológica, que se ajustan a distribuciones generalizadas de extremos. El ajuste se realiza por medio del método de máxima verosimilitud.

La distribución conjunta de nivel del mar y precipitación se construye por medio de una cópula Gaussiana (Ben Alaya, Chebana, and Ouarda 2014) que hace uso de las distribuciones marginales anteriormente comentadas y del coeficiente de correlación entre ambas variables.

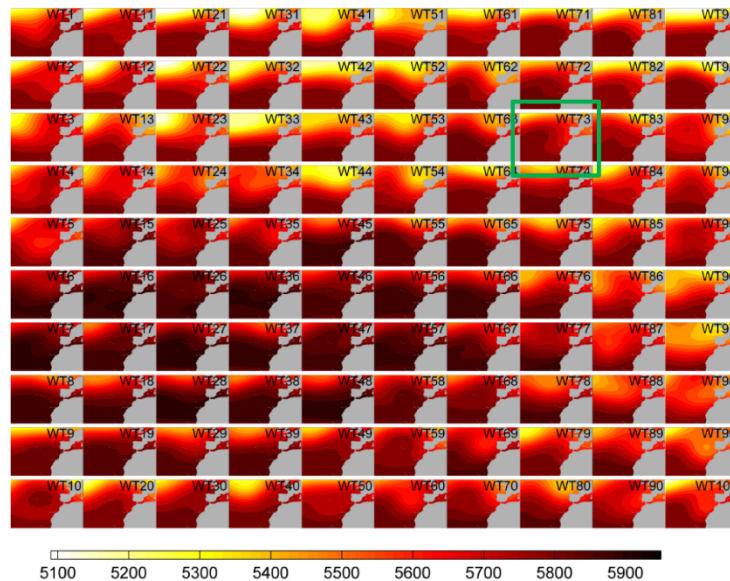


Figura 3 Tipos de tiempo para la altura del geopotencial de 500hPa. Altura medida en metros.

2.5 Generación de series sintéticas mediante simulación de Monte Carlo

Para la generación de series sintéticas de eventos combinados de precipitación y nivel del mar se procede de la forma siguiente. Para cada día se genera un valor aleatorio para determinar el tipo de tiempo reinante durante dicho día. Se genera un segundo valor aleatorio para determinar si existe precipitación en dicha día o no. Nótese que la probabilidad de precipitación puede ser diferente para cada tipo de tiempo. Si no existe precipitación se determina el nivel existente mediante la generación de un nuevo valor aleatorio y la distribución conjunta condicionada a la no existencia de precipitación. En el caso de que exista precipitación se determinan los valores combinados de precipitación y nivel a partir de la cópula Gaussiana y un vector aleatorio de dos componentes.

De esta forma se genera una serie temporal de ocurrencia de eventos combinados de precipitación y nivel del mar que reproducen la estadística contenida en las medidas utilizadas, lo que permite alargar la serie, explorando exhaustivamente el espacio de los extremos y obteniendo unas estimaciones más robustas y estables.

La generación sintética, tal y como se muestra en la **Figura 4**, permite explorar mejor el espacio de posibilidades de combinación de las distintas variables. Esto permite considerar dentro del régimen estadístico situaciones plausibles, relativamente similares a situaciones acontecidas, que mejoran la estabilidad de las predicciones, sobre todo de los periodos de retorno.

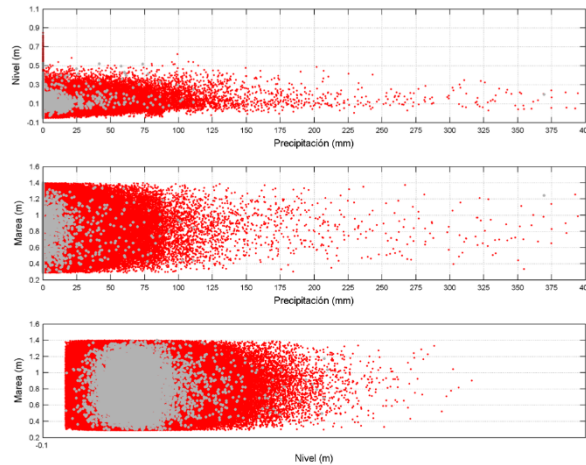


Figura 4 Datos históricos (en gris) y sintéticos (en rojo).

2.6 “Downscaling” híbrido

Mediante la simulación de Monte Carlo pueden generarse series temporal sintéticas de longitud arbitraria. En este estudio la longitud de las series es 10.000 años. Debido a la gran longitud no resulta práctico realizar una simulación numérica de 10.000 años de eventos de precipitación y nivel por su alta demanda computacional. Es por ello que se utiliza un “downscaling” híbrido, un procedimiento que combina el “downscaling” dinámico o simulación numérica, con un “downscaling” estadístico basado en comportamientos medios esperables.

El “downscaling” híbrido hace uso del algoritmo de máxima disimilitud (Camus et al. 2011) (MDA, por sus siglas en inglés) para seleccionar de los 10.000 años de eventos, un número predeterminado (250 en este caso particular) de eventos. El algoritmo MDA asegura que los casos seleccionados (tras haber determinado el caso 0 que es impuesto por el operador) son los que mejor representan la variabilidad de la serie original.

Estos casos seleccionados serán simulados numéricamente para determinar la cota de inundación mediante “downscaling” dinámico. El resto de casos no son simulados numéricamente, sino que su cota de inundación puede determinarse mediante una función de transferencia construida a partir de la simulación de los casos provistos por el algoritmo MDA. La función de transferencia hace uso de las funciones de base radiales (Buhmann 2003) (RBF) que permiten la interpolación de funciones multidimensionales sin la necesidad de construir una malla para la interpolación.

La metodología de “downscaling” híbrido permite una importante reducción del coste computacional de la generación de series largas de variables derivadas con una reducción de la precisión que puede ser fácilmente controlada. La selección del número óptimo de casos se lleva a cabo controlando el error de interpolación como función del número de componentes

simuladas mediante “downscaling” dinámico. Cuando este error se estabiliza en un mínimo aceptable, se acepta este como el óptimo del número de casos a simular.

2.7 Modelado hidrológico e hidráulico

Para el modelado hidrológico se han utilizado las especificaciones establecidas en la Guía metodológica para el cálculo de caudales de avenida en la isla de Tenerife del Consejo Insular de Aguas de Tenerife. La Guía hace uso del método del número de curva (Boughton 1989) para la simulación hidrológica y proporciona los parámetros óptimos para la simulación para cada cuenca de la isla.

Para el modelado hidráulico se ha utilizado el modelo “InfoWorks Integrated Catchment Modeling” (Jamieson et al. 2012) que resuelve el flujo bidimensional mediante las ecuaciones de Saint-Venant sobre una malla irregular que comprende toda la geometría del canal en estudio.

3. Resultados

La metodología descrita en los puntos anteriores permite caracterizar el régimen extremal de las variables estudiadas. La **Figura 5** muestra la distribución extremal de los datos históricos y de los datos sintéticos.

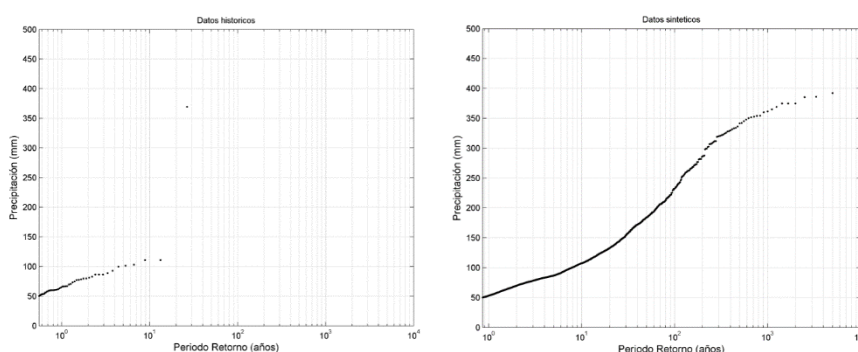


Figura 5 Distribución extremal de los datos históricos de precipitación (en la izquierda) y de los datos sintéticos (en la derecha)

La figura muestra como los datos históricos y los datos sintéticos representan de forma similar los menores periodos de retorno. Sin embargo los datos sintéticos representan mejor la precipitación de mayor periodo de retorno.

Es importante destacar que si se realiza un ajuste extremal a los datos históricos (véase figura **Figura 6**), debido al evento de mayor periodo de retorno, el ajuste va a sobredimensionar las precipitaciones para periodos de retorno mayores.

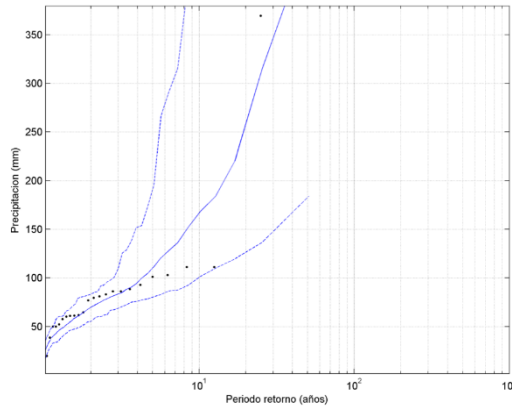


Figura 6 Ajuste extremal a los datos históricos

Sin embargo, la simulación sintética, con su mejor representación de los mayores periodos de retorno, es capaz de captar mejor la forma de la distribución real, y por tanto permite una estimación más precisa de los valores asociados a periodos de retorno mayores que la longitud de la serie histórica.

La combinación de generación sintética de eventos con simulación numérica permite una definición muy precisa de las leyes de comportamiento extremal de la cota de inundación en diversas localizaciones del canal en estudio.

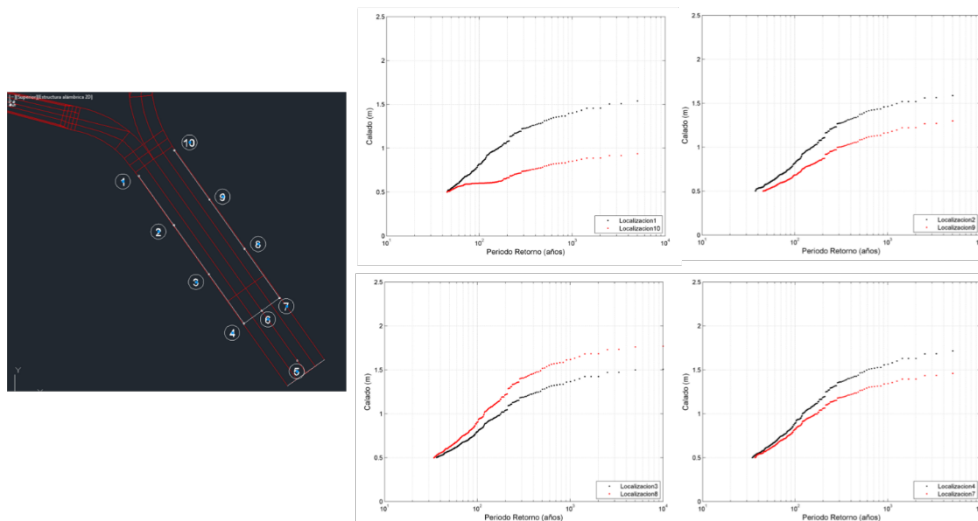


Figura 7 Localización de los puntos de caracterización de la cota del canal y ajustes extremales del calado para las localizaciones 1 y 10 (superior izquierda), 2 y 9 (superior derecha), 3 y 8 (inferior izquierda) y 4 y 7 (inferior derecha).

La Figura 7 muestra los ajustes del régimen extremal de la cota de inundación en diversas localizaciones del canal. Debido a las características geométricas del mismo existe una diferencia importante entre las cotas esperables a uno y otro lado del canal, diferencia que se reduce a medida que nos desplazamos aguas abajo del mismo.

4. Referencias

Ben Alaya, Mohamed Ali, Fateh Chebana, and Taha B M J Ouarda. 2014.

- “Probabilistic Gaussian Copula Regression Model for Multisite and Multivariable Downscaling.” *Journal of Climate* 27(9): 3331–47.
- Bermejo, M, and R Ancell. 2009. “Observed Changes in Extreme Temperatures Over Spain During 1957–2002, Using Weather Types.” *Rev Climatol.* 9: 45-61
- Boughton, W C. 1989. “ A Review of the USDA SCS Curve Number Method.” *Soil Research* 27(3): 511–23.
- Buhmann, Martin D. 2003. *Radial Basis Functions*. Cambridge University Press.
- Camus, Paula, Fernando J Mendez, Raul Medina, and Antonio S Cofiño. 2011. “Analysis of Clustering and Selection Algorithms for the Study of Multivariate Wave Climate.” *Coastal Engineering* 58(6): 453–62.
- MARUCA. Caracterización climática del medio físico marino español para la optimización de la explotación portuaria y navegación marítima. <http://maruca.ihcantabria.es/>.
- Done, James, Christopher A Davis, and Morris Weisman. 2004. “The Next Generation of NWP: Explicit Forecasts of Convection Using the Weather Research and Forecasting (WRF) Model.” *Atmospheric Science Letters* 5(6): 110–17.
- Jamieson, Sam R et al. 2012. “A Highly Efficient 2D Flood Model with Sub-Element Topography.” *Proceedings of the ICE - Water Management* 165(10): 581–95.
- Kalnay, E et al. 2010. “The NCEP/NCAR 40-Year Reanalysis Project.” *Bulletin of the American Meteorological Society* 77(3): 437–71.
- Menendez, M et al. 2014. “High-Resolution Sea Wind Hindcasts Over the Mediterranean Area.” *Climate dynamics* 42(7-8): 1857–72.