

# UNA PROPUESTA METODOLÓGICA PARA LA GENERACIÓN DE REDES DE PRECIPITACIÓN SIMULADAS A PARTIR DE REDES DE PRECIPITACIÓN OBSERVADAS EN EL MARCO DEL PROYECTO MULTITEST

José A. LÓPEZ<sup>1</sup>, José A. GUIJARRO<sup>1</sup>, Enric AGUILAR<sup>2</sup>,  
Peter DOMONKOS<sup>2</sup> y Manuela BRUNET<sup>2</sup>

<sup>1</sup>Agencia Estatal de Meteorología.

<sup>2</sup>Universitat Rovira y Virgili.

[jlopezd@aemet.es](mailto:jlopezd@aemet.es), [jguijarrop@aemet.es](mailto:jguijarrop@aemet.es), [Enric.Aguilar@urv.cat](mailto:Enric.Aguilar@urv.cat),  
[dpeterfree@gmail.com](mailto:dpeterfree@gmail.com), [manola.brunet@urv.cat](mailto:manola.brunet@urv.cat)

## RESUMEN

Una de las principales novedades que aportará el proyecto MULTITEST consiste en la comparación sistemática de procedimientos automáticos de homogenización aplicados a redes de series de precipitación representativas de climas muy diferentes, algo que no se había hecho hasta ahora. Para la generación de series de temperatura con inhomogeneidades prescritas se han desarrollado una cantidad importante de metodologías en los proyectos previos de intercomparación de procedimientos de homogeneización, como en la pasada Acción COST ES0601. Pero con la precipitación surgen complicaciones evidentes: la distribución (salvo en determinados climas y/o épocas del año) no es normal, sino que tiene habitualmente un sesgo positivo que puede llegar a ser muy pronunciado en climas secos; en estos climas es frecuente la presencia de numerosos ceros de precipitación mensual lo que requiere un tratamiento especial. La ausencia de normalidad complica además la deseable replicación en la red simulada de la estructura de covarianza espacial de los datos observados.

En este trabajo presentamos una nueva metodología que consigue, en principio, por una parte reproducir en la red de puntos simulados las características univariantes de las series que cabría esperar a partir de la climatología observada, y por otra conservar la estructura de dependencia espacial. Se muestran y analizan resultados de tres redes simuladas de clima mediterráneo, atlántico y monzónico.

**Palabras clave:** simulación de precipitaciones, técnica geoestadística

## ABSTRACT

One of the main innovations that MULTITEST introduces is that it will allow the systematic comparison of automatic homogenization procedures over networks representative of very disparate climates, something undone up to now. For temperature series generation with inhomogeneities added several methods have been developed under the past COST Action ES0601. But with precipitation there appear a number of complications: non-normality of the underlying distribution (except for some climates and/or year times) but instead it has a positive bias that can be quite great in dry climates; in these last climates it is also fairly common that one encounters

null values which requires a special treatment. The absence of normality also disturbs the desirable replication of the spatial covariance structure of the observations.

In this work a novel methodology is presented that achieves in principle both the reproduction on simulated points of the univariate characteristics according to the observed climatology and to conserve the spatial structure of dependence. Results of three simulated networks with Mediterranean, Atlantic and monsoon climates are presented and discussed.

**Key words:** precipitation simulations, geostatistical technique

## 1. INTRODUCCIÓN

Para conseguir simular series que repliquen razonablemente la estructura de correlación espacial presente en las observaciones de partida una posibilidad consiste en utilizar las técnicas de krigado en su vertiente de simulación de series. El krigado se usa generalmente como técnica de interpolación espacial que permite más flexibilidad que los modelos de regresión lineal basados en el correlograma, ya que el variograma se define también para campos espaciales sin covarianza bien definida. Aplicando una simulación basada en el krigado se puede replicar la estructura de dependencia obtenida por ajuste de modelos de variograma a las observaciones.

## 2. METODOLOGÍA

La técnica presentada aquí se puede dividir en varias fases:

### 2.1. Ajuste del modelo de estructura espacial

En este paso se estima la estructura de correlación espacial de cada mes por medio del variograma (Wackernagel, 1998). Como las series de precipitación pueden tener sesgos muy apreciables, conviene como primer paso gaussianizar la serie, o sea, transformar cada serie en una serie de valores distribuidos de normalmente y que tengan la misma ordenación por magnitud. A continuación debemos estimar un variograma medio entre los variogramas correspondientes a cada año de datos. Esta estimación de variogramas promedio o “pooled” se efectuó con el paquete “gstat” de R (Pebesma, 2010a). Como en las series puede haber valores nulos, por ejemplo en climas de tipo mediterráneo, en estos casos la serie gaussianizada no está unívocamente determinada, porque a los ceros se les puedan asignar varios rangos. Por eso para series con varios ceros se hizo un conjunto de asignaciones aleatorias de rangos a los ceros (permutando aleatoriamente el conjunto de rangos más pequeños que corresponden a los ceros) y tras hallar el variograma “pooled” para cada asignación se promediaron estos.

Una vez hallados para cada mes los variogramas empíricos se procede con la función de gstat correspondiente al ajuste de los parámetros de un variograma teórico. Es preciso adaptarse a un modelo teórico aceptable de variograma, pues solo así se garantiza que la matriz de variogramas para localizaciones arbitrarias es semidefinida positiva, como teóricamente lo tiene que ser. Para las simulaciones de este trabajo se usó el modelo esférico. Se hizo la hipótesis habitual de estacionariedad e isotropía en este paso, lo que simplifica considerablemente el proceso.

## **2.2. Selección de lugares aleatorios para la localización de observatorios simulados**

Una ventaja importante de este método de generar series simuladas es que permite simular los emplazamientos de observatorios ficticios de forma arbitraria, pues con el modelo de variograma ajustado a los datos y conocidas las distancias entre observatorios así como entre puntos simulados se determina su matriz de valores del variograma. Para la selección de puntos aleatorios en el área de las observaciones se inscribió el área en un rectángulo y se simulaban puntos aleatorios en el interior del rectángulo, desechándose los que cayeran fuera del área. Para este paso se usó una función del paquete de R “sp” (Pebesma, 2010b) que permite averiguar si un punto de coordenadas arbitrarias está o no en el interior de un polígono dado.

## **2.3. Generación de valores gaussianizados en los puntos de simulación**

Con los valores gaussianizados de las observaciones y su variograma, “gstat” permite realizar simulaciones gaussianizadas en puntos arbitrarios. Se seleccionó aleatoriamente un número de años de observación igual al número de años simulados deseado, en este caso 60, y para cada año se generaron por meses los valores gaussianizados simulados. Para transformar esos valores gaussianizados a precipitación es preciso estimar la forma de la distribución de la precipitación en cada punto simulado.

## **2.4. Estimación de parámetros en puntos simulados**

Se han estimado en cada punto los parámetros: probabilidad de precipitación nula, y los parámetros de forma y de escala de la distribución gamma. Para los parámetros de la gamma, puesto que deben ser necesariamente positivos, se ha usado una simulación logarítmica de los parámetros, partiendo de los logaritmos de los parámetros en los puntos de observación. A continuación se aplicó simulación con krigeado como antes en la generación de valores gaussianizados de precipitación, y por último se deshizo la transformación logarítmica con una exponencial.

Para la probabilidad de precipitación nula se optó por una interpolación a los puntos simulados con ponderación por inverso de distancia que asegura automáticamente que valores en  $[0, 1]$  se transformen en valores en el mismo rango.

## **2.5. Generación de valores de precipitación simulados**

Como en cada punto de simulación hemos estimado los valores gaussianizados y los parámetros de la distribución el paso a valores finales de acuerdo a la distribución es inmediato en el caso de ausencia de valores nulos: basta pasar de valores gaussianizados a probabilidades acumuladas, y por último a cuantiles de la gamma correspondientes. Pero cuando hay valores nulos hay que ser cuidadosos para que al imponer una frecuencia de ceros de acuerdo a la probabilidad estimada para el punto a los valores de la gamma más pequeños, esto no distorsione la media final. Por ello se añadió un re-escalamiento de los valores no nulos para compensar este defecto. En realidad se aplicó un procedimiento más complejo para asegurar también que los valores finales no nulos se distribuyan de acuerdo a una gamma.

## 2.6. Efecto de la altitud en la precipitación

Para simular en los puntos de simulación el efecto topográfico en la precipitación, suponiendo que se asigna una altitud al punto (esto requiere disponer de una topografía del área), se ha supuesto que la altitud influye en la probabilidad de nulos y en el parámetro de escala de la distribución, no así en el parámetro de forma, lo que equivale básicamente a aplicar un factor multiplicativo. Para el parámetro de forma beta se ha hecho una regresión logarítmica de la beta de las observaciones con la altitud, se ha transformado beta a altitud 0 con esa regresión, y se ha estimado esta beta reducida en los puntos de simulación como arriba. Por último se deshizo la transformación según las altitudes de la simulación.

Para la probabilidad de nulos se siguió un procedimiento paralelo salvo que la regresión no fue logarítmica.

## 2.7. Efecto de las coordenadas geográficas

Cuando no hay valores nulos de precipitación se puede tener en cuenta el efecto de la longitud y latitud en los valores simulados aplicando un factor multiplicativo a las series observadas que las reduzca a una longitud/latitud estándar (se puede incluir también el efecto de la altitud, reduciendo a altitud cero). Para ello se hace una regresión logarítmica con las variables geográficas, se reducen los logaritmos a coordenadas estándar, se aplica la exponencial para estimar las precipitaciones reducidas y se simulan valores en los puntos deseados como se describe arriba. Al final se vuelve a usar la regresión logarítmica del comienzo para particularizar a las coordenadas geográficas del punto.

## 3. ANÁLISIS DE LAS SIMULACIONES

Se analizan las simulaciones realizadas correspondientes a los climas mediterráneo, monzónico y atlántico.

### 3.1. Clima mediterráneo

Para el tipo de clima mediterráneo se utilizó como base una red de observatorios en Mallorca compuesta por 107 observatorios con datos mensuales del periodo 1951-2015.

#### 3.1.a. Distribución mensual de precipitación

En la fig. 1 se han dibujado los diagramas de cajas para todos los observatorios de la red por meses en abscisas, las medianas en la izquierda y los rangos intercuartílicos a la derecha. Claramente visible es el patrón de otoños lluviosos, seguidos por el invierno, con un verano muy seco. Los rangos intercuartílicos muestran una evolución inter-mensual muy similar a las medianas, aunque con una atenuación de la variación entre meses.

En la fig. 2 se reflejan diagramas de cajas de los valores de los parámetros de forma  $\gamma$  y escala  $\beta$  de la pdf gamma ajustada por el método de los momentos y la probabilidad de valores nulos por meses. Como cabría esperar, en los meses más secos los sesgos de la distribución mensual son mayores, como se aprecia por el menor valor del parámetro de forma  $\gamma$  (para la pdf gamma el sesgo varía con el inverso de la raíz

de  $\gamma$ ). Como  $\beta$  es el cociente entre la media y  $\gamma$  (en el método de los momentos) de lo anterior se deduce que como se puede apreciar el parámetro de escala es más estable que las medias entre meses de la fig. 1.

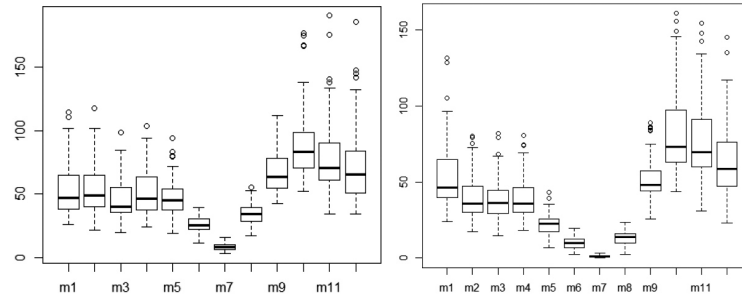


Fig. 1: Medianas (iq) y rango intercuartílico (der) por meses en los observatorios de la red de Mallorca

Una vez descrita la distribución mensual de las observaciones pasamos a la comparación con las simulaciones. En la fig. 3 se representan, para cada mes, los valores medios para todos los observatorios/puntos de simulación respectivamente de las medianas y rangos intercuartílicos.

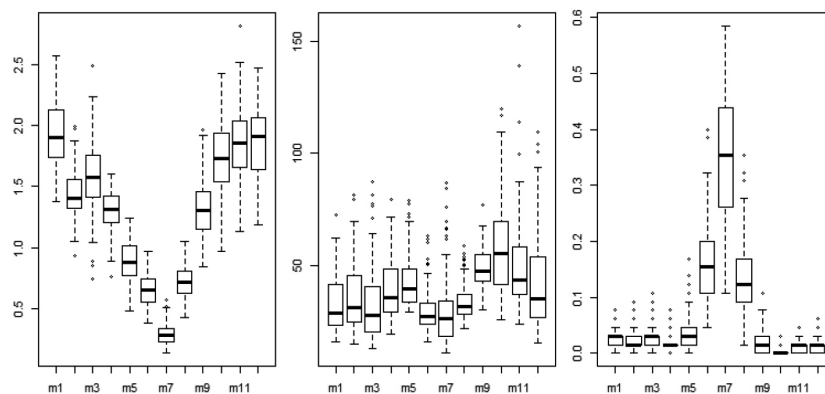


Fig. 2: Parámetros de la pdf gamma de la precipitación mensual, de izquierda a derecha: de forma de la gamma, de escala de la gamma y probabilidad de valores nulos.

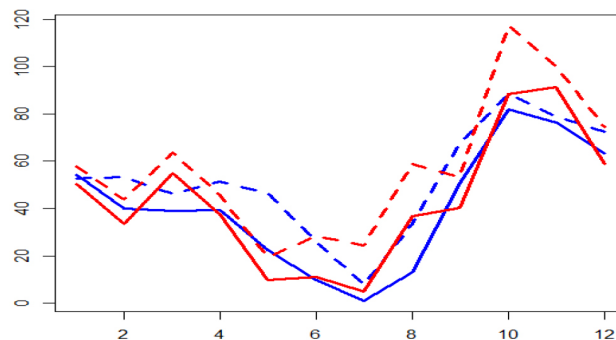


Fig. 3: Valores medios para todos los puntos de las medianas (líneas continuas) y rangos IQ (líneas de trazos) de observaciones (azul) y simulaciones (rojo) por meses para Mallorca.

Teniendo en cuenta que los puntos de simulación no coinciden con los de observación lógicamente la concordancia entre observaciones y simulaciones en cuanto a distribución medias mensuales podemos decir que es satisfactoria.

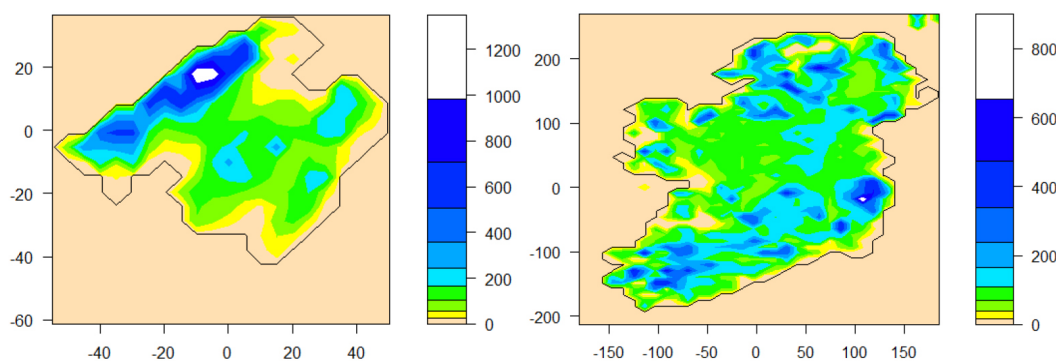


Fig. 4: Topografía de Mallorca (izq.) e Irlanda (der.).

### 3.1.b. Distribución espacial

La fig. 4 contiene un mapa de la topografía de Mallorca con una rejilla de aproximadamente 5 x 5 km en proyección cilíndrica que es la misma que se ha usado para los mapas de precipitación de Mallorca de este apartado. Para los mapas de precipitación anual, enero y julio de las observaciones y de las simulaciones de la fig. 5 se ha aplicado un krigeado universal con variables auxiliares las coordenadas x e y correspondientes a longitud y latitud. Los rasgos principales de la distribución espacial son similares en observaciones y simulaciones, con la peculiaridad de que en las alturas más altas de la sierra de la Tramontana al NO las simulaciones muestran un aumento con la altitud más acusado. Esto es congruente con el hecho de que en la simulación se ha incluido el efecto topográfico, junto con la escasez de observatorios en las zonas más altas. Parecidas consideraciones se aplican a un mes relativamente lluvioso como enero. Pero el acuerdo entre la distribución de las observaciones y las simulaciones es mucho peor en un mes muy seco como julio (nótese el cambio de escala entre los dos mapas de julio, última fila de la fig.5). El cambio de escala permite apreciar una correlación similar con los parámetros geográficos, pero las diferencias de valores son muy claras, con mayor precipitación en las simulaciones.

En la tabla 1 se resumen estadísticos con las diferencias entre los campos krigeados ploteados en la fig. 5. Las diferencias cuadráticas como porcentaje respecto a la media del campo de observaciones de la última fila muestran que para el total anual medio son de un 30%, lo cual puede considerarse aceptable, teniendo en cuenta que ya el hecho de que los puntos de simulación no coinciden con los observatorios introduce en cualquier caso un valor mínimo para este estadístico. Enero está en la línea de los totales, pero en julio vemos que el campo simulado es el doble en promedio que el observado, aunque la primera fila muestra que en términos absolutos la diferencia es pequeña, y el valor medio del campo observado es tan solo 7.4 mm.

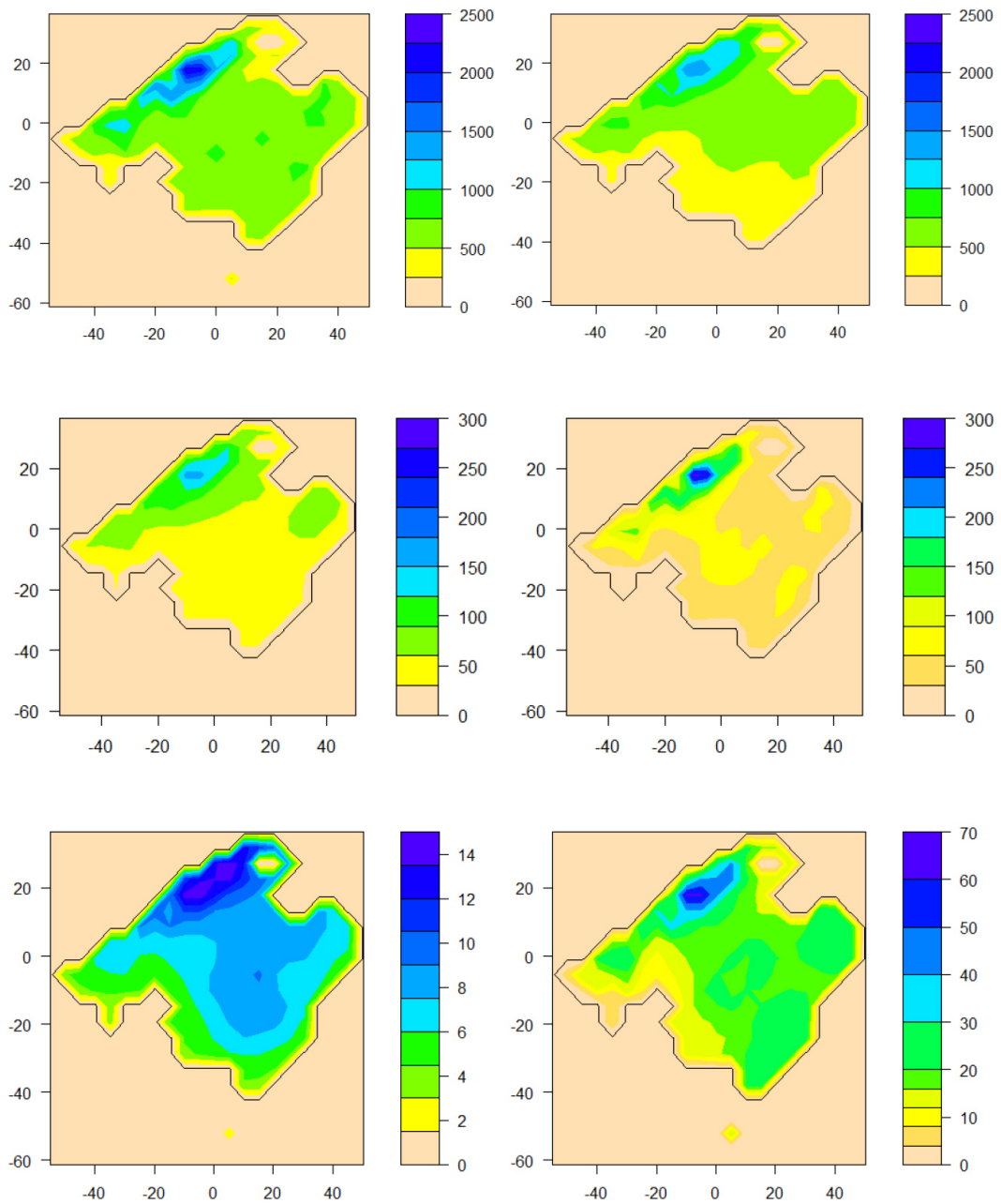


Fig. 5: De arriba abajo precipitaciones medias anual, en enero y en julio de las observaciones (columna derecha) y de las simulaciones (derecha.) en Mallorca.

|                    | AÑO   | ENERO | JULIO |
|--------------------|-------|-------|-------|
| RaízDifCuaMed (mm) | 184.5 | 23.0  | 14.8  |
| RDCM/media(%)      | 30.3  | 38.0  | 199.5 |

Tabla 1: Diferencias entre los valores krigeados de precipitación en las observaciones y las simulaciones para Mallorca. Valores cuadráticos medios y % respecto a observaciones.

### 3.1.c. Correlación espacial

Para comparar la estructura espacial de las correlaciones de las observaciones y simulaciones se ha escogido el correlograma (correlación en función de distancia) para las diferencias de precipitaciones estandarizadas entre meses sucesivos. La fig. 6 muestra una buena coherencia entre observaciones y simulaciones.

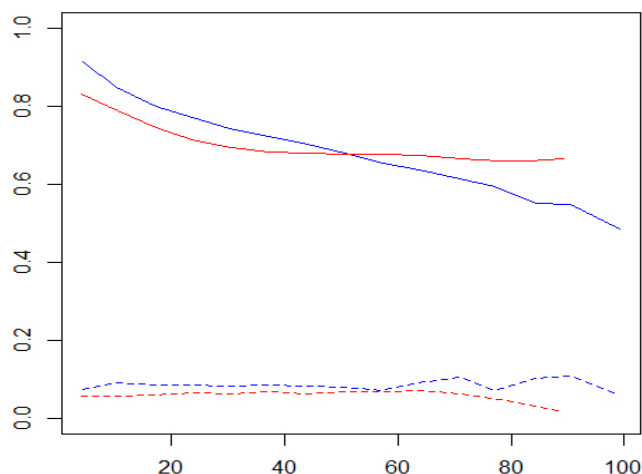


Fig. 6: Correlogramas de las diferencias estandarizadas entre meses sucesivos para las observaciones (azul) y las simulaciones (rojo); línea continua valores medios por cajas 7 km y a trazos valores de diferencias entre percentiles 10 y 90 en esas cajas en Mallorca. Eje X distancias en km.

## 3.2. Clima atlántico

Para el tipo de clima atlántico se utilizaron los datos de una red de red de observatorios de Irlanda formada por 172 observatorios con 70 años de observaciones mensuales en 1941-2010. La ausencia de ceros permitió aplicar el procedimiento descrito en 2.7, idóneo en este caso dada la extensión importante del área y por tanto la previsible dependencia con coordenadas geográficas.

### 3.2.a. Distribución mensual de precipitación

Los diagramas de cajas de la fig. 7 muestran el ritmo suave de cambios entre meses sucesivos de las medianas y rangos intercuartílicos de los observatorios irlandeses, con mayor pluviometría en invierno y mínima en verano. Las dispersiones se mueven de forma parecida, pero algo atenuada, a los totales mensuales.

Los parámetros  $\gamma$  y  $\beta$  de las pdf gamma ajustadas por el método de los momentos de la fig. 8 muestran una esperable disminución del parámetro de forma con la reducción de la precipitación (sesgo mayor); también está claro por comparación con la fig. 2 que los valores de  $\gamma$  para Irlanda son más altos que para el clima mediterráneo (por tanto mejor aproximación a la distribución normal).

La comparación sumaria de los estadísticos de las distribuciones mensuales de precipitación contenida en la fig. 10 muestra a las claras el excelente acuerdo entre observaciones y simulaciones en este aspecto.



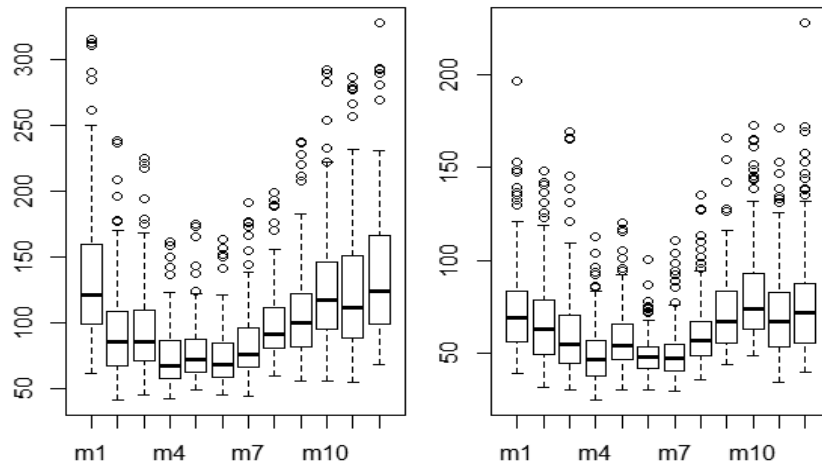


Fig. 7: Medianas (izq) y rango intercuartílico (der) por meses en los observatorios de la red de Irlanda

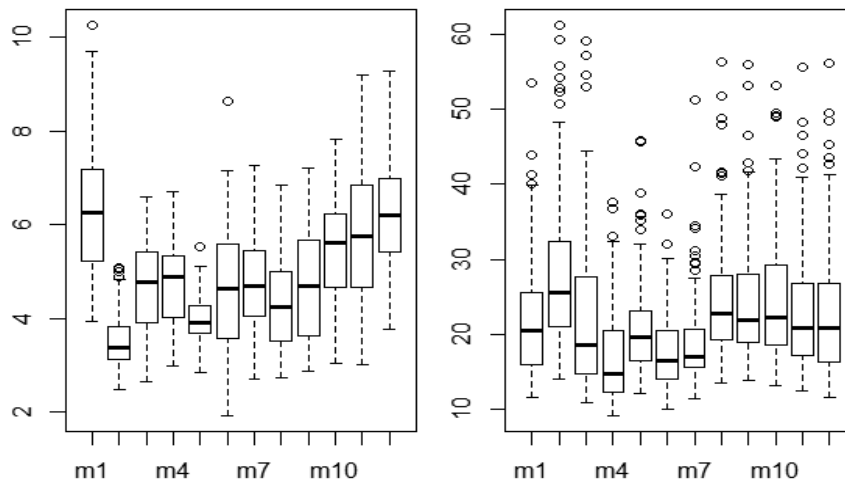


Fig. 8: Parámetros de la pdf gamma de la precipitación mensual, de forma (der) y de escala (izq).

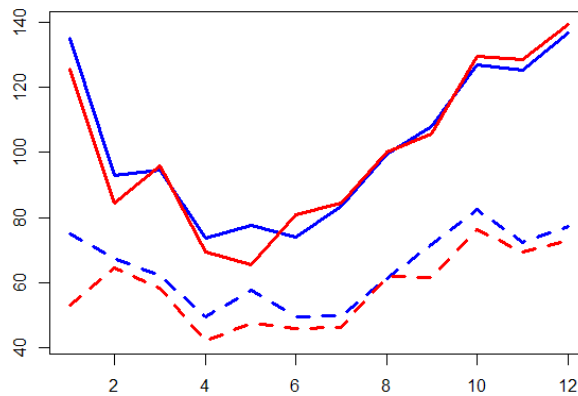
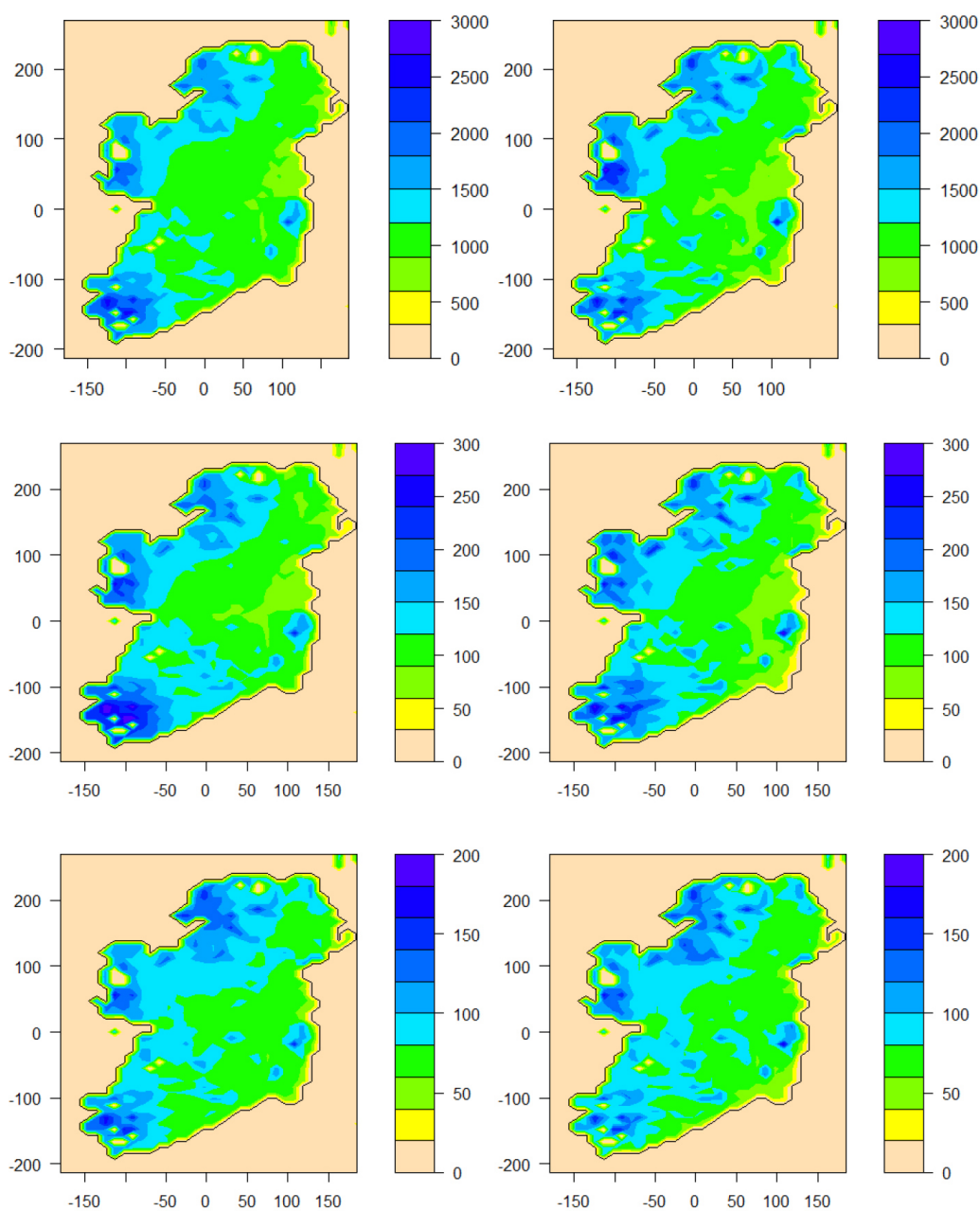


Fig. 9: Valores medios para todos los puntos de las medianas (líneas continuas) y rangos IQ (líneas de trazos) de observaciones (azul) y simulaciones (rojo) por meses para Irlanda.

### 3.2.b. Distribución espacial

Para los mapas de Irlanda de este apartado se ha empleado una rejilla de aproximadamente 10 x 10 km en proyección cilíndrica. El mapa de la fig. 4 con la topografía de Irlanda muestra el relieve suave de la isla con casi un 95% de las tierras emergidas por debajo de la cota de 300 m.

Los mapas de la distribución de medias anuales de la fig. 10 manifiestan una muy buena concordancia entre las observaciones y las simulaciones, que se traslada a los meses individuales, incluso los más secos (última fila de la fig. 10 correspondiente a julio).



*Fig. 10: De arriba abajo precipitaciones medias anuales, en enero y en julio de las observaciones (columna derecha) y de las simulaciones (derecha) en Irlanda.*

|                    | AÑO   | ENERO | JULIO |
|--------------------|-------|-------|-------|
| RaízDifCuaMed (mm) | 114.8 | 16.5  | 7.4   |
| RDCM/media(%)      | 9.3   | 12.6  | 8.4   |

Tabla 2: Diferencias entre los valores krigeados de precipitación en las observaciones y las simulaciones para Irlanda. Valores cuadráticos medios y % respecto a observaciones.

Los estadísticos sumarios de la tabla 2 corroboran la capacidad de las simulaciones para capturar los rasgos de la distribución espacial de precipitación en Irlanda, incluso a escala mensual (comparar con tabla 1 para Mallorca). Los errores en porcentaje respecto de los valores medios oscilan alrededor del 10%.

### 3.2.c. Correlación espacial

De forma a priori quizá un tanto sorprendente, los correlogramas de la fig. 11 muestran una concordancia entre las observaciones y las simulaciones claramente inferior a lo encontrado en Mallorca (ver fig. 6): en concreto las correlaciones no decrecen con la distancia en las simulaciones en forma análoga a las observaciones. Una posible vía de explicación podría estibar en la mayor dispersión de los valores del correlograma de las observaciones (línea azul a trazos de la fig. 11) en comparación con Mallorca, que además crece de forma marcada con la distancia.

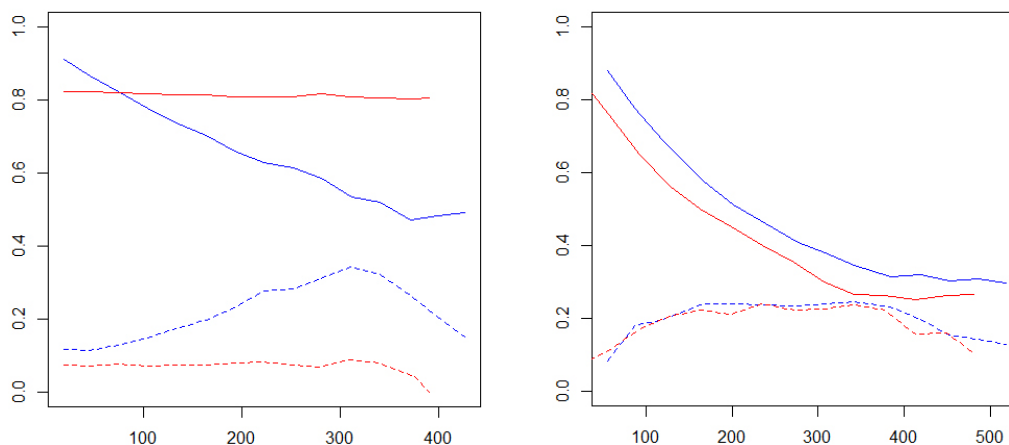


Fig. 11: Correlogramas entre diferencias mensuales sucesivas en Irlanda (der.) y la India (izq.) (clave como en la fig. 6)

Se han efectuado también simulaciones para una zona de la India de clima monzónico a partir de datos en rejilla sin topografía provenientes del grid global de precipitación del GPCC (Rudolf *et al.*, 2005). En la fig. 11 derecha se han reflejado los correlogramas que muestran que para esa área de clima monzónico la estructura de correlación espacial sí que se replica en las simulaciones.

## 4. RESULTADOS Y DISCUSIÓN

Se ha presentado un método novedoso para la generación de datos simulados de precipitación mensual en puntos aleatorios a partir de una red de observaciones que

es capaz de imitar de forma aceptable las distribuciones mensuales de estadísticos de centralidad y dispersión incluyendo el efecto topográfico. El método se basa en la técnica geoestadística del krigado que posibilita replicar también la estructura de correlación espacial.

La técnica de simulación de precipitación en redes de puntos aleatorios es razonablemente satisfactoria bajo todos los aspectos analizados para el área de clima mediterráneo (*a priori* la más complicada). Para un área de clima monzónico a partir de datos en rejilla y sin topografía, con resultados no mostrados también se consigue replicar satisfactoriamente tanto las propiedades de distribución univariante mensuales como la distribución espacial y la estructura de dependencia espacial.

En la zona de clima atlántico replica muy bien las propiedades de distribución por meses y la distribución espacial de valores medios mensuales, lo que es congruente con la mayor normalidad y suavidad de variación inter-mensual de precipitaciones del clima atlántico. En cambio la estructura espacial en este último caso, de forma un tanto sorprendente, no se consigue replicar satisfactoriamente. La razón para esto requiere posteriores investigaciones por el equipo para intentar introducir mejoras que lo resuelvan.

## AGRADECIMIENTOS

El proyecto MULTITEST está financiado por el Ministerio de Economía y Competitividad (CGL2014-52901-P).

Damos gracias al servicio nacional de meteorología de Irlanda (Met Éireann) por facilitarnos las series de precipitación observadas. Enric Aguilar y Manola Brunet agradecen el apoyo del proyecto europeo titulado: Uncertainties in Ensembles of Regional Reanalyses (UERRA. FP7-SPACE-2013-1, No. de referencia: 607193).

## REFERENCIAS

- Pebesma, E.J. (2010a). The meuse data set: a tutorial for the gstat R package (en <https://cran.r-project.org/web/packages/gstat/vignettes/gstat.pdf>)
- Pebesma, E.J. (2010b). Package sp (en <https://cran.r-project.org/web/packages/sp/sp.pdf>)
- Rudolf, B., Schneider U. (2005): Calculation of gridded precipitation data for the global land-surface using in-situ gauge observations. Proceedings of the 2<sup>nd</sup> Workshop of the International Precipitation Working Group IPWG, Monterey, October 2004,231-247.
- Wackernagel, H. (1998). Multivariate Geostatistics, an introduction with applications Springer, Berlin.