Spatially-based reconstruction of daily precipitation instrumental data series

^{1,2,3} *Roberto Serrano-Notivoli, ^{1,3} Martín de Luis ^{1,3} Miguel Ángel Saz, ² Santiago
Beguiría

¹ *Department of Geography and Regional Planning, University of Zaragoza, Pedro Cerbuna 12,
50009. Zaragoza, Spain*

² *Estación Experimental de Aula Dei, Consejo Superior de Investigaciones Científicas (EEAD-
CSIC), Avda. Montañana 1005, 50059. Zaragoza, Spain*

³ *Environmental Sciences Institute (IUCA), University of Zaragoza, Pedro Cerbuna 12, 50009.
Zaragoza, Spain.*

Abstract: This work presents a method for the reconstruction of fragmentary daily precipitation datasets. The method aims to preserve the local and temporal variability characteristic of high-frequency precipitation data, and does not use the time-structure of the data. Based on the precipitation values recorded at closest neighbours during a target day, two reference values (RV) are computed: a binomial prediction (BP) expressing the probability of occurrence of a wet day; and a magnitude prediction (MP), referring to the amount of precipitation. Generalised linear models (GLM) are used to compute the RV using the precipitation data (occurrence and magnitude) of the 10 nearest neighbours as the dependent variable, and the geographic information of each station (latitude, longitude, and altitude) as the independent variables. The RV are then used to: (1) apply

* Corresponding author: rs@unizar.es

quality control to the data, flagging suspect records according to five predefined criteria;
(2) obtain serially complete time series by imputing RV to missing observations in the
original dataset; (3) create new time series at locations where there were no observations
or gridded datasets with even spatial coverage over the study area. The routines used were
compiled into an R-package called *reddPrec* (*reconstruction of daily data -
Precipitation*) available to any user. We applied the methods described to the complete
daily precipitation dataset of the island of Majorca in Spain, spanning the period from
1971 to 2014.

Keywords: Daily precipitation, spatial analysis, quality control, missing values, grid,
reddPrec.

1. Introduction

Quality-controlled and serially-complete datasets are required in most climate studies in order to compute long-term climatologies, assess the existence of trends, statistical downscaling of global and regional climate model data, or analyse the frequency and magnitude of extreme events, to name a few topics. Great effort has been expended to create datasets of climatic variables at different temporal resolutions, with spatial coverage ranging from the global scale (New et al. 1999, 2000 and 2002; Hijmans et al., 2005; Mitchell and Jones, 2005; Lawrimore et al., 2011; Becker et al., 2013), to the regional or local scale (Klein-Tank et al., 2002; Perry and Hollis, 2005; Ninyerola et al., 2007; Li et al., 2009; Vicente-Serrano et al., 2010; Chaney et al., 2014; Gonzalez-Hidalgo et al., 2011; Herrera et al., 2012). These datasets address a growing demand for climate data, in order to understand climate behaviour in the past and to serve as a basis for making good estimations of its future evolution (Trenberth, 2011; Zhang et al., 2011; Zwiers et al., 2013; Trenberth et al., 2014).

Among all the variables for which datasets have been created, precipitation is one of the most elusive, because of its considerable temporal and spatial variability. Most studies focusing on creating climatologies or determining trends in precipitation use datasets at a monthly time resolution (New et al. 1999, 2000, and 2002; Mitchell and Jones, 2005; Norrant and Douguédroit, 2005; Perry and Hollis, 2005; Caramelo and Manso, 2007; Ninyerola et al., 2007; Gonzalez-Hidalgo et al., 2011; Río et al., 2011).

While the monthly scale may be acceptable for climatological studies focused on precipitation totals, it is not sufficient for other crucial issues such as extreme event analysis (which usually have high impacts), for which daily and even higher time resolutions are necessary (Zhang et al., 2011). However, the development of high-

frequency precipitation datasets poses a number of difficulties that limit more widespread use. The increase in the amount of data compared to the monthly scale, the comprehensive quality control needed, and the high variability of the spatial fields requires considerable effort in terms of process automation and computing time. Quality-control and reconstruction methods developed for monthly data perform well for computing long-term averages or assessing trends in precipitation totals, and take advantage of the higher spatial correlation of the data at low time-resolutions. However, such procedures cannot be applied to daily-scale precipitation due to the much higher spatial variability and the interaction of local issues, such as orography, which, in combination with the specific atmospheric conditions, can lead to variability in spatial distribution patterns from one day to the next. These issues have caused daily precipitation to be one of the least studied variables in climate reconstruction studies, despite its great importance in extreme event analysis (Easterling et al., 2000; Griffiths et al., 2003; Alexander et al., 2006; López-Moreno et al., 2010; Min et al., 2011; Beguería et al., 2011; Orlowsky and Sereviratne, 2012).

Several studies have described daily precipitation reconstructions as a means to reach other objectives. Some of them created climatologies by a direct estimation of precipitation in grid points from raw data (Frei and Schär, 1998) or avoiding quality control and considering only elevation parameters (Daly et al., 2002). Other works were based on the interpolation of the anomalies from the climatological normal (Schneider et al., 2014), or improving downscaling methods (Wong et al., 2014). Different climatological studies based their analyses on daily precipitation values, but few of them include complete quality control of the information used (Feng et al., 2004; Vicente-Serrano et al., 2010; Castro et al., 2014). In addition, most of the methods are based on

the selection of a reduced number of stations (those whose data sets are longer and expected to be of higher quality), rejecting data of potential value.

Similar limitations can be found in different approaches used to create global or regional daily precipitation datasets. For instance, Eischeid et al. (2000) used only series with a minimum length and selected neighbouring stations by correlation, smoothing the local variability of precipitation. In a first version of the European Climate Assessment Dataset, Klein-Tank et al. (2002) applied quality control by considering only the temporal structure of the series. Later, Klok and Klein-Tank (2009) updated this dataset by adding stations, but they did not apply spatial quality control and did not do any kind of reconstruction, similar to the work of Menne et al. (2012) for a global dataset.

Substantial effort has been invested in the creation of daily gridded datasets. However, most of the previous works addressed problems of loss of spatial resolution due to the use of monthly totals in the estimation of daily precipitation (Haylock et al., 2008; Yatagai et al., 2012); the selection of series with a minimum length combined with global models in interpolation schemes (Herrera et al. 2012) that produced smoothed estimations with a low number of zeros and underestimated maximum events; or applying homogenisation procedures to reanalysis products (Chaney et al., 2014).

Other works tested methods to estimate daily precipitation to use in climatology (Neykov et al., 2012) and hydrology (Hwang et al., 2012), addressing similar problems of sub-optimal use of available data due to the need for the correlation of series (Verdin et al, 2015; Kleiber et al., 2012) and the incorporation of the error term in final estimations by using monthly empirical cumulative distribution functions (Clark et al., 2005). All of these approaches restrict the high-resolution and local results of daily precipitation datasets.

Only a few studies focused specifically on quality control and reconstruction of daily precipitation (Feng et al., 2004; Simolo et al., 2010; Vicente-Serrano et al., 2010; Castro et al., 2014). The creation of these datasets typically involves a process including quality-control tests for detecting and eliminating suspect data, correcting inhomogeneities within the data, and reconstructing fragmentary series. Reference series (RS) are commonly used in all these stages. A common characteristic of RS calculation is the use of data from neighbouring stations. However, the construction of RS requires that both candidate and neighbour series are long enough and overlap over a significant period. This is problematic, as short or non-overlapping data series need to be rejected from the final dataset, making sub-optimal use of the available information. Besides, two important assumptions are necessary for the construction of RS: (i) that the relationship among neighbour series does not vary over the reference period and/or during the overlap period of the two series; and (ii) that the series have a similar temporal structure throughout the period. These are problematic assumptions, especially with daily data, since the behaviour of neighbouring stations may be very different on daily scales, and because the evolution of precipitation can change in a different way to nearby locations.

The purpose of this paper is to present a spatially-based method for daily precipitation reconstruction that preserves both the spatial and temporal variability. This approach is made without any preconceived assumption about the structure of the series or their relationship with near series, and maximising the use of available information. This method can be used to: (a) apply quality control to daily precipitation datasets; (b) reconstruct missing values; and (c) estimate precipitation values at specific locations of interest, even if no direct observations exist. All the calculations were made using the R language (R core team, 2016), using a new package called `reddPrec` (*reconstruction*

of daily data - Precipitation) compiled with all the required functions, and was made available to any interested user (Serrano-Notivoli et al., 2017).

The approach was tested and validated using daily precipitation data series from the island of Majorca in Spain. Located in the western Mediterranean, its orographic diversity (altitudes ranging between 0 and 1445 m a.s.l.) produces a varied climate subject to many influences. The island has quite an irregular and torrential precipitation regime, including long periods of consecutive dry days and frequent extreme precipitation events. In addition, the rain gauge network is relatively dense in space but strongly variable in the duration and expected quality of the series, representing a perfect framework to validate a daily precipitation reconstruction method.

2. Description of the method and the case study

2.1 The reference values (RV) as a general tool for quality control and reconstruction

The key difference between the method proposed and others traditionally used is the use of individual reference values (RV), instead of reference series (RS), to complete the quality control during the reconstruction process.

RV are calculated independently for each day and location according to the available data for that day from closer neighbouring stations. The geographic information of each station (latitude, longitude, and altitude) are used as the independent variables for modelling. Since data availability varies from day to day, selected neighbour stations can also vary accordingly. Also, since independent models are constructed for each location and day, the estimated parameters of the models (reflecting the influence of the altitude, latitude, and longitude on the probability of occurrence and the magnitude of the precipitation) may also vary in their sign and magnitude from neighbouring stations, or over different

days for the same location. Such an approach, including three regressors, allows the creation of highly flexible models able to reflect local precipitation conditions. In addition, since neighbour selection is local and independently determined each day, no restriction in data selection according to data set length or structural characteristics exists, allowing to use all available information.

Generalised linear models (GLM) are used to compute the RV using the precipitation data (occurrence and magnitude) of the 10 nearest neighbours as the dependent variable, and the geographic information of each station (latitude, longitude, and altitude) as the independent variables.

Computation of each individual RV is based on two predicted values: (i) a binomial prediction (BP) of the probability of occurrence of a wet day; and (ii) the magnitude prediction of precipitation (MP). Both are calculated with a multivariate logistic regression that uses topographic and geographic characteristics of the stations as independent variables:

$$P_{i,l} = \beta_{0,i,l} + \beta_{1,i,l}alt_l + \beta_{2,i,l}lat_l + \beta_{3,i,l}lon_l + \varepsilon_{i,l} \quad (1)$$

where $P_{i,l}$ is the BP or the MP in day i , and location l ; $\beta_{n,i,l}$ are the regression coefficients; alt , lat , and lon are the altitude, latitude, and longitude of the 10 nearest neighbours (NNS) with observations on that day; and $\varepsilon_{i,l}$ is the error term. The equation is least-squares fitted, based on the data. R-language nomenclature is used to describe the RV computation (R Core Team, 2016). The function “glm” from R-package “stats” is used to calculate both BP and MP (Serrano-Notivoli et al., 2017).

The probability of occurrence of precipitation on day i at location l ($BP_{i,l}$), is computed by using the binomial family (Eq. 2,3,4) to describe the error distribution (equivalent to a logistic regression):

178 $modelBP_{i,l} = glm(PCPb \sim ALT + LAT + LON,$
 179 $data = NEI, family = binomial)$ (2)
 180 $BP_{i,l} = predict(modelBP_{i,l}, newdata = CAN, type = "response")$ (3)

$$181 \quad err_{i,l} = \frac{sqrt\left(\sum(PCPb - BP_{i,l})^2\right)}{length(PCPb) - 3} \quad (4)$$

182 where $PCPb$ is a vector containing information on whether the day was dry or wet
 183 (codified as 0 or 1, respectively) in the 10 nearest neighbour stations that collected
 184 precipitation during day i ($NNS_{i,10}$); and ALT , LAT , and LON , are the vectors containing
 185 the geographic locations of these neighbouring stations. The $modelBP_{i,l}$ is then applied
 186 to the ALT , LAT , and LON of the candidate station l to calculate the probability of
 187 occurrence of precipitation on day i ($BP_{i,l}$). The uncertainty of $BP_{i,l}$ is also computed at
 188 each step as described in equation (3). $PCPb$ and the ALT , LAT , and LON of the
 189 neighbouring stations are stored in the dataframe NEI while the geographic information
 190 of candidate station l is stored in dataframe CAN .

191 Similarly, the magnitude prediction of precipitation on day i and location l ($BP_{i,l}$) is also
 192 computed using the glm function, but now using the quasibinomial family to describe the
 193 error distribution (Eq. 5,6,7).

194 $modelMP_{i,l} = glm(PCPq \sim ALT + LAT + LON,$
 195 $data = NEI, family = quasibinomial)$ (5)
 196 $MP_{i,l} = predict(modelMP_{i,l}, newdata = CAN, type = "response")$ (6)

$$197 \quad err_{i,l} = \frac{sqrt\left(\sum(PCPq - MP_{i,l})^2\right)}{length(PCPq) - 3} \quad (7)$$

198 where $PCPq$ includes the precipitation magnitudes observed for day i at the 10 nearest
 199 stations (Obs) rescaled between zero and one, with zero being half of the minimum

precipitation observed ($Obsmin$), and one being the maximum value ($Obsmax$) plus the difference between the maximum and minimum, minus half of the minimum (Eq. 8). This rescaling of the data in a quasi-binomial model context, allows the fixing of asymptotic limits to the predicted values according to the range of the local variation of observations. This transformation is especially useful in situations where the geographic conditions of the candidate stations are out of the range of altitudes, latitudes or longitudes of its NNS. Finally, the predicted value $MPq_{i,l}$ is rescaled inversely (Eq. 9) to obtain the final precipitation magnitude prediction ($MP_{i,l}$).

$$PCPq = \frac{(Obs - 0.5 * Obsmin)}{(2 * Obsmax - Obsmin) - 0.5 * Obsmin} \quad (8)$$

$$MP_{i,l} = MPq_{i,l} * ((2 * Obsmax - Obsmin) - 0.5 * Obsmin) + (0.5 * Obsmin) \quad (9)$$

$$PCPq_{i,l} = Obs_i - 0.5 * Obsmin \quad (10)$$

The final RV is determined by combining MP and BP (See Table S1 in the supplement), using a threshold value of $BP_{i,l} \geq 0.5$ to determine a wet day:

$$RV = \begin{cases} MP & \forall BP \geq 0.5 \\ 0 & \forall BP < 0.5 \end{cases} \quad (11)$$

The threshold of 0.5 allows the sum of the probabilities of occurrence of the predicted wet days in that area to be equal to the observed occurrence of that situation. For instance, given an area defined by 10 stations where the precipitation occurred at 7 of them (70% wet stations), the model assigns 70% of the area associated with these stations with probabilities of >0.5 , and the 30% of the area with probabilities of <0.5 .

2.1.1. Potential for quality control (QC)

The RV can be used to develop a quality control test to detect anomalous data. There are numerous different reasons to expect that anomalous or erroneous data are present in

original climate datasets, including: codification errors, incorrect transcriptions from manual to digital data, reading errors from analogue devices, problems in software or hardware at automatic stations, and many more (Reek et al., 1992; Feng et al., 2004, Su-Ping et al., 2012). For this reason, a comprehensive analysis of the data from each station on each day is necessary to be certain that the final dataset will be consistent and reliable. RV, calculated as described in the previous section was used to perform a QC of the original values. To do this, RV were calculated for each individual observed precipitation value and five criteria were defined to identify suspect values:

QC.1) Isolated wet conditions: Observed value is over zero and the precipitation at the 10 nearest neighbouring stations that measured precipitation during the same day ($NNS_{i,l10}$) is zero.

QC.2) Isolated dry conditions: Observed value is zero and all 10 $NNS_{i,l10}$ are over zero.

QC.3) Suspect outlier: The magnitude of the observed value is ten-times higher or lower than that predicted by its 10 $NNS_{i,l10}$. A value of 1 mm was added to each element to avoid conflicting situations when computing with zero values.

QC.4) Suspect dry: Observed value is zero, wet probability is over 99%, and predicted magnitude is over 5 mm.

QC.5) Suspect wet: Observed value is over 5 mm, dry probability is over 99%, and predicted magnitude is under 0.1 mm.

The five QC criteria were applied to all observed data and the suspect observations were removed. This process was repeated using the cleaned database from the previous step until no more observations were flagged (Fig. 1).

2.1.2. Potential for gap filling

The RV can also be used for gap filling in order to obtain serially complete data series. In this case, the RV can be directly used to fill the missing values. To preserve the peculiarities of the original series and to avoid including inhomogeneity in the time dimension, a correction is required consisting of multiplying the RV by a correction coefficient computed as the ratio between the monthly means of daily precipitation of all observed values, and the same calculation of predicted data. This allows RV series to have the same mean as that of the observed data in the candidate series. Correction coefficients are computed for each monthly value independently to avoid biases due to seasonal variation:

$$RV_{c,i,m} = RV_{i,m} * (\overline{Obs_m} / \overline{RV_m}) \quad (11)$$

To compute the corrected value of day i of a specific year, the RV of the same day $RV_{i,m}$ is multiplied by the ratio between the average of the observations in that month for all lengths of the series Obs_m , and the average of the estimations of that month in the same period RV_m .

2.1.3. Potential for inference

The computation of the RV can also be used for estimating precipitation values at locations where no direct observations exist. Using the coordinates and altitude of the location of interest, and based on the observations at the 10 NNS, complete daily precipitation series can be computed.

This is useful for creating new data series at single ungauged locations, and it can also be used for computing gridded datasets by applying the same approach to a set of regularly distributed pairs of coordinates.

The process of inferring a new series is the same as that used to fill the missing values in the original series. In the first stage, all the data series with observations are filled. As the new estimations cannot be corrected by the original values (because there are no observations at the candidate location), they need to use the same 10 NNS each day to avoid including inhomogeneity. In a second stage, the new RV are computed for each day using the filled original stations. These RV are the final estimations.

2.2. Software implementation

An R package called `reddPrec` (*reconstruction of daily data - Precipitation*) was developed containing the previously described procedures. This package contains two main functions: `qcPrec` flags suspect data, as explained in Section 2.2, resulting in a cleaned dataset; and `gapFilling` computes the RV to fill missing data (Section 2.3) and/or create new series (Section 2.4), depending on the set of locations included as input. The package is freely available from the official R repository (cran.r-project.org/web/packages/reddPrec) and its operation is explained in Serrano-Notivol et al. (2017).

2.3. Case study: evaluating the accuracy of the reconstruction process

The method was evaluated using all available precipitation data covering the island of Majorca for the period between 1971 and 2014 (Fig. 2). The purpose of the validation was the evaluation of the agreement between the observations registered in these stations and their corresponding predictions, computed for each location and time of observed values.

In Majorca, most rainfall occurs during the autumn and from November to April, it is generally associated with the Atlantic Westerlies and other cyclogenetic processes within the western Mediterranean Basin (Sumner et al., 1995a). Convective storms are more frequent in the warmer season. Intense precipitation events mostly occur from September to November (Sumner et al., 1995a). Annual precipitation exceeds 1000 mm on the highest zones of the Tramuntana Range and falls gradually with altitude to amounts under 400 mm in the south-east coastal fringe (Sumner et al., 1995b).

During the period 1971 to 2014, a total of 325 precipitation stations have existed in Majorca. Of these, only 2.8% of the stations registered more than 99% of the daily observations during the whole period, and a total of 216 stations included data from more than 10 years. The average length of time series is 21 years, and the number of stations available on any day ranges from 110 to 170 (Fig. 3 lower panel). Globally, data gaps accounted for 53.05% of the original dataset, and they were regularly distributed over time, not focused in a specific season or group of years (Fig. 3 upper panel).

The data from the 325 stations were subject to QC and all the suspect data were flagged and discarded before further analysis. Gap filling was performed by imputing the RV for any missing data. Finally, the set of 216 stations with more than 10 years of original data were reconstructed and used to build a 1x1 km grid across the island.

The accuracy of the reconstruction method was evaluated based on three different statistical analyses:

A contingency table analysis was used to assess the accuracy of the binomial (wet/dry) prediction. Accuracy was computed for global observations-estimation comparison and separated by month.

A correlation analysis between observed and predicted data allowed comparison of the RV and the observations in two ways: (i) using daily results, the correlation between observations and predictions of daily precipitation means (considering all days) was calculated for each day, medians of the wet days and the 95th percentile of the wet days; and (ii) using the same methods but applied to each station.

Goodness of fit (GOF) was assessed through four different statistics: the Mean Absolute Error (MAE), the Mean Error (ME), the Ratio of Means (RM) and the Ratio of Standard Deviations (RSD).

The MAE is calculated as the mean of the absolute differences between predicted and observed values, and provides information about the magnitude of the errors irrespective of their sign:

$$MAE = \frac{1}{n} \sum_{i=1}^n |p_i - o_i| \quad (12)$$

where p is the predicted value; o is the observed value, and n is the number of cases. The ME is the mean of the differences between predicted and observed values, and gives the bias (the tendency towards under- or over-predicting):

$$ME = \frac{1}{n} \sum_{i=1}^n (p_i - o_i) \quad (13)$$

These statistics were applied to daily values by splitting the whole dataset (only observations over zero) into deciles according to the distribution of observed values, as well as to monthly aggregated values (only considering predicted values for available observations) in order to check the performance of the method for aggregated precipitation.

Finally, the RM and the RSD were performed. These techniques give the bias on the estimation of the mean and the variance of the data, with one indicating a non-biased estimate:

$$RM = \frac{\bar{p}}{\bar{o}} \quad RSD = \frac{\sigma_p}{\sigma_o} \quad (14)$$

where $\bar{p}$ is the mean of estimates, $\bar{o}$ is the mean of the observations, σ_p is the standard deviation of the predictions, and σ_o is the standard deviation of the observations. These two statistics were applied only to monthly and annual totals, since the distribution of the daily precipitation data is highly non-normal, so these techniques would be misleading in that case.

Daily precipitation climatologies were developed by computing several indices at the annual scale to assess the accuracy of the spatial distribution of climatic variables. They were analysed through the annual total precipitation, the daily precipitation intensity (PMED), the maximum precipitation in one day (RX1), the number of days with precipitation over 20 mm (R20mm), the maximum consecutive dry days (CDD), and the maximum consecutive wet days (CWD). These variables were mapped using the grid computed at a 1x1 km spatial distribution based on the reconstructed stations.

3. Results

3.1. Quality control

From the 2,412,810 original daily precipitation observations, 1.59% were flagged (and removed) according to the rules outlined in Section 2.2. The rate of flagged data was higher during the first twenty years when data availability was lowest (Fig. 4a). QC.2 (isolated dry conditions) was the most frequent criterion (34.8% of total), especially in the first fifteen years and during the months with a larger number of wet days (in winter

and spring Fig. 4b). These high values are probably due to days when the accumulated precipitation does not reach the minimum to be noted (if the observation is taken manually) or detected (if at an automatic station). QC.1 (isolated wet conditions) was more evenly distributed over the period and was also a large source of flagged data (23.2%). This precipitation probably occurred, but the climatic context suggests that they are far from representative of the local climatic situation and most of the removed data were low values (Fig. 4c). The frequency of flagged data due to QC.4 (suspect dry) (20.5%) and QC.5 (suspect wet) (12.5%) decreased from the start to the end of the time series. They represent similar situations to QC.2 and QC.1, respectively, by removing mainly zero or low precipitation values when predicted values were very low in the case of QC.4 (Fig. 4e) and removing values under 20 mm in almost all cases with QC.5 (Fig. 4f). QC.3 (9%) represented the detection of outliers. These outliers were observations that differ from predictions by an order of magnitude, and are probably due to codification problems and confusion with decimal places, which produces the removal of a great number of low values (Fig. 4d). However, a wide range of daily precipitation values were flagged as suspect. For instance, high extremes (>50 mm) represented 6.5% of removed data in this category, and their inclusion in this category represented situations where the observed value was very different from that predicted by its nearest neighbours. QC.3 (suspect outlier) was more frequent in months with intense precipitation (October to December), but it was lower in the warmer season, when it is expected that more convective storms will occur (Fig. 4b). This means that the quality control process is more sensitive to very extreme precipitation than to regular convective situations.

3.2. Agreement between observed and predicted precipitation values

Wet/dry prediction

The percentage of true and false positives and negatives of predictions gives information about the accuracy of the wet and dry day prediction (Table 1).

A rate of 82.84% of true positives, representing the correctly-predicted wet days, was obtained. The false positives corresponded in most cases with low values (median: 1.2 mm; mean: 2.2 mm), with the Binomial Prediction (BP) in most of these cases near to 1 but also lower than 0.5 (Fig. 5a, c). The rate of true negatives, representing well-predicted dry days, was very high (96.36%).

False negatives (wet predictions on dry days) were identified in 3.58% of observed dry days (median: 0.99; mean: 1.7 mm) but only in 0.22% of cases did magnitude predictions exceed 5 mm, being BP values higher than 0.5 in a few cases (Fig. 5b, d). The accuracy of the wet/dry prediction by month (Table 1) showed that, overall, the percentage of well-predicted dry days was higher than the well-predicted wet days that obtained values under 80% in summer months (Jun, Jul, and Aug).

Magnitude prediction

Overall, the distribution of observations and predictions was very similar, with a slight overestimation of the lowest values and underestimation of the highest (Fig. 6). However, predictions of lower values were more frequent, showing a global displacement that is not reflected in correlations when aggregating by daily means (Pearsons' $r = 0.999$, Fig. 7a). This strong correlation is influenced by the prevalence of days with no precipitation in the dataset. However, even considering only the wet days the agreement between RV and observations was also high ($r = 0.992$). Similar results were obtained when days with precipitation exceeding the 95th percentile were considered ($r = 0.991$). The density

curves constructed for observed and predicted values also show a highly significant agreement (Fig. 7d, e, f). Overall, there were discrepancies in the low values, showing a slight overestimation in estimated values. This is mainly due to the higher number of observations with low values.

By analysing the same statistics when averaging precipitation by stations instead of days, quite high correlations were also obtained ($r > 0.96$ in all cases, Fig. 8), and the density curves of predicted values matched those of the observed values quite well.

The comparison between observed (%OBS) and predicted (%PRED) data by deciles showed small differences (Table 2). These values are related to the error statistics of ME and MAE. The first groups (up to decile seven) represent less than 10% of the total precipitation in each class, and the values of MAE are higher than the range of their corresponding decile. MAE increased from low to high deciles with the inverse behaviour found in the ME statistics. In decile eight, the percentage of precipitation over the total is more than 10%, and the ME is zero. Decile 10 groups included values over 26 mm, including 41% of the total precipitation in the observed values, and 36.1% in the predicted values. This decile has the highest MAE (12.46 mm) and the lowest ME (−5.45 mm).

The largest discrepancies between observed and estimated values occur in the lowest deciles (See Fig. S1 in the supplement), as shown in Table 2. The relative weight of the errors in the precipitation totals in these ranges is low. These differences were due to the prediction of lower values of precipitation, also shown in daily means (Fig. 7d, e, f).

Predicted values showed a smoother distribution than observed, which reveals stepped magnitudes (due to the observation ranges). The differences in the ratio of means (RM) by stations did not show a defined spatial pattern over the study area (Fig. 9). Most of the stations were between 0.9 and 1.1. However, 19 of the 26 stations with overestimated

means ($RM > 1.10$) were distributed along the northwest of the island, coinciding with elevated areas of the Tramuntana Range.

Monthly precipitation totals showed good agreement between observed and predicted values (Table 3). The best agreement corresponded with the summer months (June and July), with MAE values below 5 mm; while the highest MAE corresponded to autumn (October to December), coinciding with the highest mean rainfall amounts and the lowest ME. Estimation of the mean and standard deviation of monthly totals showed no significant biases.

Gap Filling

After the quality control stage the dataset contained 55.97% missing values over the 1971 to 2014 period. These missing values were imputed using the RV, resulting in complete series after the reconstruction. An example of the gap filling process for a station (B_249) is given in Fig. 10. The time series show the original and reconstructed data at daily and monthly time scales. Correlation between observed and RV (predictions on the same days and at the same locations as the observations) was very high in both cases (0.906 and 0.967 for daily and monthly values, respectively), and the density curves were also very similar (See Fig. S2 in the supplement).

Inference

As an example of inference, a random pair of coordinates was selected (ETRS_89_UTM_30N X1016969, Y4410641, ALT 109m) corresponding to one location where no observations existed. Daily precipitation series and monthly and annual totals

were estimated using data from nearby stations located at distances between 2 and 7 km (Fig. 11).

When applying the method for creating new data, as there are no observations with which to make direct comparisons, the standard error of the regression can be used as a measure of uncertainty. The shaded areas in Fig. 11 show the 95% prediction confidence bands. As with the predicted magnitudes themselves, the prediction errors depend on the particular situation on each day, so two days with a similar predicted magnitude can have very different errors. Besides being informative, the uncertainty range can be used in further analyses or simulations to determine the uncertainty of computed variables by error propagation analysis.

Newly created data fitted within the distribution function of their nearest stations at daily, monthly, and annual scales. Observed values showed similar variability between neighbouring stations, and predicted values faithfully reproduced precipitation behaviour at all temporal scales (See Fig. S3 in the supplement).

The same procedure was applied to a set of regularly distributed points making up a grid covering Majorca with a spatial resolution of 1 km. Results are shown for annual precipitation and five extreme precipitation indices, showing a coherent spatial distribution of precipitation (Fig. 12). The final maps of climatic indices showed a high accuracy in spatial distribution of the variable. This is due to the use of all the available daily precipitation data and the independent estimation for each grid point and day, using the nearest observations to maintain the local variability of precipitation distribution.

The annual precipitation map (Fig. 12a) showed a clear south-north gradient from lower to higher values. The southern half of the island averaged lower values of annual accumulated precipitation, coinciding with the lower values of daily and extreme

precipitation indices. Furthermore, the maximum length of dry spells (Fig. 12d) in this area exceeded 60 days, with the maximum length of wet events being less than five days. On the other hand the north of the island, which corresponds to the most elevated zone, reached the highest values in annual totals, PMED, RX1, R20mm, and CWD (Fig. 12a, b, c, e, and f, respectively).

4. Discussion

Having a reliable and consistent daily precipitation dataset is key to any climatological study performing extreme events analysis, a study area characterisation, or to test relationships with other natural elements. For this reason, the method applied to obtain a quality-controlled and complete series should reflect the spatial and temporal variability of the variable of interest.

In the approach demonstrated here, the use of local models with three sources of variations (ALT, LAT and LON) requires a minimum number of observations for each prediction. The use of fewer NNS could reduce the strength of the models and the predictions. Using only a few NNS makes contextualisation of the general precipitation situation difficult, and therefore the options for applying the quality control are reduced. On the other hand, the use of many more NNS will limit the desired local character of the predictions. For these reasons 10 NNS were used, based on previous testing. This number allows for a contextualisation of the surrounding precipitation, and a model with three independent variables could in some cases represent real local conditions. However, this is not a magic number, and it is a value that must be set according to the climatic characteristics of the study area and to the density of available observatories.

With respect to the set QC thresholds, isolated wet and dry days (QC.1 and QC.2) that could be real local elements were removed. Although these observations could be correct, they are related to less representative situations with respect to their environment, and in most cases, they corresponded to low precipitation values. On the other hand, the suspect dry and wet days (QC.4 and QC.5) could also correspond to real situations, but they did not fit the spatial distribution of local precipitation in their corresponding days. The same circumstance of unrepresentative observations was considered when removing outliers (QC.3). A ratio of 10 times over or under the predictions was used to eliminate the observed values, because most of these cases were related to problems of digitalisation or decimal errors. Any threshold used to remove observations is necessarily arbitrary. For this reason, the values used here are not necessarily adequate for all situations. These thresholds are user-defined and this method allows them to be flexibly defined.

4.1 Reference values (RV) vs. reference series (RS)

Most reconstruction methods rely on the statistical characteristics of the data over a long period of time, such as climatologies (long-term averages) or correlations between neighbouring stations. This imposes restrictions on data selection, since there is a minimum length required for any climatology or correlation to be reliable, and subsequently any data series shorter than this minimum length has to be discarded. Considering this restriction, an important proportion of the original data is not used in the final reconstruction, which represents a significant reduction in the information content of the process, and may even lead to spatial smoothing. A major advantage of the method described here is that all of the computations are done independently for each day and location, so there are no limitations on the use of all available observations, irrespective

of their length. Daily and site-independent modelling also allows for a highly flexible reconstruction process, to adapt to the specific and changing characteristics of different days and sites.

GLM were used for estimating the probability of wet day (BP) and the magnitude of precipitation (MP). These kinds of regressions have been used before for precipitation prediction (Hay et al., 2002; Helsel and Hirsch, 2002; Syed et al., 2003; Hwang et al., 2012) using climatologies as independent variables. However, none of these studies used adaptive models for each location and time to represent the local precipitation variability as fully as possible.

Simolo et al. (2010) used a similar two-step method using a Gamma-GLM and time-dependent regression coefficients for modelling the occurrence and magnitude of precipitation. Neykov et al. (2014) also separated the precipitation occurrence and intensity, but in their work, which is based on that of Furrer and Katz (2008), they fitted a hybrid Gamma-generalised Pareto (GP) and a hybrid Weibull-GP distribution to the whole data series to improve previous methods, which in most cases underestimated heavy precipitation values. Wong et al. (2014) used logistic models to predict wet/dry days and a mixture of probability distributions over the whole series for precipitation intensities, and Hwang et al. (2011) used a two-step approach with a logistic regression to compute the spatial occurrence of precipitation, but using different interpolation methods to estimate precipitation amounts that also considered the monthly climatologies. Moncho et al. (2012) did not separate precipitation occurrence from intensity, and used a model with four parameters to predict the probability of precipitation with best results on wet days rather than dry days. Other works used a similar approach. This is, for instance, the case for Verdin et al. (2015), who used all the historical

occurrence observations to fit a probit model to estimate the local coefficients, and then set the spatial structure through the empirical correlation of these probit models' errors. This required a minimum length of data series to apply the models and, furthermore, they did not estimate precipitation amounts. These problems also affected Kleiber et al. (2012), who estimated the precipitation occurrence with a Gaussian process based on a spatiotemporal approach. Finally, Clark et al. (2005) set the precipitation occurrence using the conditional cumulative distribution function of each series and the precipitation intensity based on ensemble grids by using realisations from correlated random fields. Although they estimated the error term, it was incorporated into the final predictions, while in the method presented here it was provided separately as a measure of uncertainty. However, despite the similarities, all these approaches were based on using the complete data series of the candidate and neighbouring stations, assuming that the relationship between them remains constant over time. Apart from this assumption, which may not hold for daily data, they require a minimum length of both data series and a minimum overlapping period between them in order for this relationship to be captured. In contrast, the method used in this study treats each daily record independently from the rest of the series. One direct advantage is that no overlapping or minimum data lengths are required and also, perhaps more importantly, it avoids assumptions of stationarity of the relationship between candidate and neighbouring series. At the daily scale this is important because the relationships between any two stations (e.g. one station registering higher precipitation than the other due to their exposure to the dominant winds) may change dramatically from one day to the next due to changing meteorological conditions.

4.2 Estimation thresholds

The setting of a lower and an upper limit for each daily prediction by using an adaptive asymptote based on the data recorded at neighbouring stations prevented overshooting predictions, especially at high elevations where all observations occur at lower altitudes than that of the candidate location. This is a common problem in data prediction, as noted by Daly et al. (1994, 2008) when using standard linear regression. Using a quasi-binomial GLM and setting reasonable upper and lower limits to the predictions, overestimation was minimised while being able to estimate precipitation at altitudes (and latitudes and longitudes) outside the range of the observations. However, by changing the thresholds imposed to set the asymptotes the predictions would change, especially in extreme situations where the candidate location is out of range of any of the covariates (latitude, longitude, or altitude) and extrapolation occurs. Thresholds were selected that kept estimated magnitudes within a range compatible with the observations. However, in datasets with larger differences between available stations these thresholds should probably be changed to increase the variability in estimations.

4.3 Quality control improvements and limitations

The process of identifying and removing suspect data was made only considering the observed precipitation and the influence of geographic variables, and was not based on the correlation structure of the stations in the long-term nor on departures from computed climatologies. The five criteria used were adapted to the climatic behaviour of Mediterranean areas, but other settings could be used in other climates or geographical and data settings.

Among the studies reporting QC of daily precipitation data (Reek et al., 1992; Griffiths et al., 2003; Feng et al., 2004; Hubbard et al., 2005; Vicente-Serrano et al., 2010; Picarreta

et al., 2013), only a few gave final percentages of removed data even though the study area, the station density, and the number of observations were different from this research in each case. Using the present method, 1.59% of the original data were flagged by the five criteria and removed prior to further analysis. This value is similar to the 1.04% maximum proportion of data rejected in any one series registered by Vicente-Serrano et al. (2010) and to the 2% of data flagged as erroneous by Hubbard et al. (2005) but far more than the 0.04% suspect data of Reek et al. (1992). These authors dedicated the most criteria to temperature and only considered whether the daily precipitation values were inside predefined limits, and the relation with snowfall and snow depth data. Feng et al. (2004) found a very low figure of suspect data (0.02%), which may be at least partially attributed to their use of a highly-correlated series. Nonetheless, the percentage of flagged data will depend ultimately on how restrictive the thresholds used during the QC stage are. For instance, more permissive thresholds (lower required probability of precipitation occurrence or higher ratios to detect outliers) will result in a lower flagging of suspect data at the cost of increasing the probability of leaving problematic data un-flagged. It is important to set appropriate thresholds (using the proper criteria) adapted to the climate behaviour of study area. The threshold values used here aimed not only to remove erroneous data but also anomalous observations that could have a negative effect (excess leverage) on the reconstruction of other series or on the construction of the grid. Observations that could be considered as outliers with respect to their neighbours will have a negative effect on the attribution process since they will have an excessive influence on the models, known as leverage. By removing these observations from further processing, it is possible to attain more robust estimations at the expense of smoothing the predicted spatial field. However, it is unlikely to remove original data corresponding

to extreme events because the QC criteria are applied individually for each day, considering the nearest observations, with adaptive thresholds based on the input data. A compromise between detection of suspect data and maintaining the original variability of the data is therefore required. The values used in this work ensured that the vast majority of the local variability of precipitation was retained and that only very exceptional situations were removed.

4.4. Reconstruction and inference potential

The reconstruction of original data series and the creation of new ones make the method described here a useful tool to any user that needs to estimate precipitation data at a specific location. The creation of new series is a key factor in developing climatologies in new locations that could support other studies that do not depend on the availability of daily data for any location and period. In reference to gridded datasets, the method is able to estimate new data at any spatial resolution, always considering that the result is not a continuous surface of precipitation data over the territory but predictions at specific locations. Moreover, functions based on the latitude, longitude, and altitude of the 10 nearest observations are able to capture the local variability, which may change from one day to the next. Ensor and Robertson (2008) found that while gridding increased the frequency of low-precipitation events, the frequency of heavy-precipitation was greatly reduced. Beguería et al. (2015) highlighted that most gridding methods result in a reduction of the variance with respect to the observed data. This method, aimed at retaining the local variability, may help in reducing this smoothing effect associated with other gridding methods. The examples of the applications in the present work show that, despite the daily precipitation estimations being computed independently for each grid

point, the construction of different climatologies and extreme precipitation indices produced coherent spatial patterns and, as an improvement over other methods, the improvement in accuracy is noticeable. This would be impossible to detect without using all the available information and considering local variability of precipitation.

5. Conclusions

This paper describes a spatially-based method for quality control, gap filling, and inference of daily precipitation data.

In the first phase, a prediction of the probability of precipitation occurrence (BP) was computed by binomial GLM against altitude, longitude, and latitude data using the information of the 10 nearest neighbours (NNS). Then, the magnitude of precipitation (MP) was calculated using quasi-binomial GLM, after rescaling the data of the NNS to a range between zero and one to reduce the risk of over-shooting predictions in out-of-range situations. A combination of BP and MP resulted in the final RV for each location and day.

To flag suspect data, the RV values were compared with their corresponding observations, which were flagged according to five criteria aimed at detecting: (1) isolated wet conditions, (2) isolated dry conditions, (3) suspect outliers, (4) suspect wet days, and (5) suspect dry days. In the final stage, RV were computed again with the clean dataset after removing all suspect data to reconstruct a complete daily precipitation dataset that was used to calculate a 1x1 km spatial resolution grid over Majorca island.

Many examples of the application of this procedure to the reconstruction of a fragmentary station-based dataset are shown here, with satisfactory results in terms of goodness-of-fit and errors, bias, distribution density, monthly and annual totals, and daily and extreme

precipitation indices. The potential for inference of daily precipitation at ungauged locations, and for grid construction, was also shown. In addition to the expected data, the method estimates the uncertainty of the predictions, which is also locally variable and may change from day to day.

Acknowledgements

This study was supported by research projects CGL2012-31668, CGL2015-69985-R, CGL2011-24185 and CGL2014-52135-C3-1-R, financed by the Spanish Ministerio de Economía y Competitividad (MINECO) and FEDER-ERDF funds. The researchers were supported by the Government of Aragón through the ‘Programme of research groups’ (groups H38, ‘Clima, Cambio Global y Sistemas Naturales’ and ‘E68, Geomorfología y Cambio Global’).

6. References

- Alexander L, Zhang X, Peterson T, Caesar J, Gleason B, Tank A, Haylock M, Collins D, Trewin B, Rahimzadeh F, Tagipour A, Kumar K, Revadekar J, Griffiths G, Vincent L, Stephenson D, Burn J, Aguilar E, Brunet M, Taylor M, New M, Zhai P, Rusticucci M, Vazquez-Aguirre J (2006) Global observed changes in daily climate extremes of temperature and precipitation. *Journal of Geophysical Research* 111(D5). doi: 10.1029/2005JD006290.
- Beguiría S, Angulo-Martínez M, López-Moreno JI, El-Kenawy AE, Vicente-Serrano S (2011) Assessing trends in extreme precipitation by non-stationary peaks-over-threshold analysis, NE Spain 1930–2006. *International Journal of Climatology* 31: 2102–2114. doi: 10.1002/joc.2218
- Beguiría S, Vicente-Serrano SM, Tomás-Burguera M, Maneta M (2015) Bias in the variance of gridded data sets leads to misleading conclusions about changes in climate variability. *International Journal of Climatology*, early view. doi: 10.1002/joc.4561.
- Becker A, Finger P, Meyer-Christoffer A, Rudolf B, Schamm K, Schneider U, Ziese M (2013) A description of the global land-surface precipitation data products of the Global Precipitation Climatology Centre with sample applications including centennial (trend) analysis from 1901-present. *Earth Syst. Sci. Data* 5: 71–99. doi: 10.5194/essd-5-71-2013

- Brunetti M, Colacino M, Maugeri M, Nanni T (2001) Trends in the daily intensity of precipitation in Italy from 1951 to 1996. *International Journal of Climatology* 21(3): 299–316. doi: 10.1002/joc.613.
- Camberlin P, Gitau W, Oettli P, Ogallo L, Bois B (2014) Spatial interpolation of daily rainfall stochastic generation parameters over East Africa. *Climate Research*. 59: 39-60. doi: 10.3354/cr01198.
- Caramelo L and Manso MD (2007) A study of precipitation variability in the Duero Basin (Iberian Peninsula). *International Journal of Climatology* 27(3): 327–339. doi: 10.1002/joc.1403.
- Castro L, Gironás J, Fernández B (2014) Spatial estimation of daily precipitation in regions with complex relief and scarce data using terrain orientation. *Journal of Hydrology*. 517: 481-492. doi: 10.1016/j.jhydrol.2014.05.064.
- Chaney NW, Sheffield J, Villarini G (2014) Developement of a High-Resolution Gridded Daily Meteorological Dataset over Sub-Saharan Africa: Spatial Analysis of Trends in Climate Extremes. *Journal of Climate* 27, 5815–5835. doi: 10.1175/JCLI-D-13-00423.1
- Clark MP, Slater AG (2005) Probabilistic Quantitative Precipitation Estimation in Complex Terrain. *Journal of Hydrometeorology*. 7(1):3-22. doi: 10.1175/JHM474.1
- Daly C, Neilson RP, Phillips DL (1994) A statistical-topographic model for mapping climatological precipitation over mountainous terrain. *Journal of Applied Meteorology* 33: 140–158.
- Daly C, Gibson WP, Taylor GH, Johnson GL, Pasteris P (2002) A knowledge-based approach to the statistical mapping of climate. *Climate Research*. 22: 99-113.
- Daly C, Halbleib M, Smith J, Gibson W, Doggett M, Taylor G, Curtis J, Pasteris P (2008) Physiographically sensitive mapping of climatological temperature and precipitation across the conterminous United States. *International Journal of Climatology* 28(15): 2031–2064. doi: 10.1002/joc.1688.
- Easterling DR, Meehl GA, Parmesan C, Changnon SA, Karl TR, Mearns LO (2000) Climate extremes: observations, modeling, and impacts. *Science* 289(5487): 2068–74. doi: 10.1126/science.289.5487.2068.
- Eischeid JK, Pasteris PA, Diaz HF, Plantico MS, Lott NJ (2000) Creating a serially complete, national daily time series of temperature and precipitation for the western United States. *Journal of Applied Meteorology*, 39 (9): 1580-1591.
- Ensor L and Robeson S (2008) Statistical Characteristics of Daily Precipitation: Comparisons of Gridded and Point Datasets. *Journal of Applied Meteorology and Climatology* 47(9): 2468-2476. doi: 10.1175/2008JAMC1757.1.

- 754 Feng S, Hu Q, Qian W (2004) Quality control of daily meteorological data in China,
755 1951–2000: a new dataset. *International Journal of Climatology* 24: 853–870.
756
- 757 Frei C and Schär C (1998) A precipitation climatology for the Alps from high-resolution
758 rain-gauge observations. *International Journal of Climatology* 18: 873-900.
759
- 760 Furrer E, Katz R (2008) Improving the simulation of extreme precipitation events by
761 stochastic weather generators. *Water Resources Research*. 44(12). doi:
762 10.1029/2008WR007316.
763
- 764 Griffiths GM, Salinger MJ, Leleu I (2003) Trends in extreme daily rainfall across the
765 South Pacific and relationship to the South Pacific converge zone. *International Journal*
766 *of Climatology* 23:847-869. DOI: 10.1002/joc.923.
767
- 768 González-Hidalgo JC, Brunetti M, de Luis M (2011) A new tool for monthly precipitation
769 analysis in Spain: MOPREDAS database (monthly precipitation trends December 1945-
770 November 2005). *International Journal of Climatology* 31: 715-731. doi:
771 10.1002/joc.2115.
772
- 773 Hay LE, Clarck MP, Wilby RL (2002) Use of regional climate model output for
774 hydrologic simulations. *Journal of Hydrometeorology* 3:571-590
775
- 776 Haylock M, Hofstra N, Tank A, Klok E, Jones P, New M (2008) A European daily high-
777 resolution gridded data set of surface temperature and precipitation for 1950–
778 2006. *Journal of Geophysical Research* 113, D20119. doi: 10.1029/2008JD010201.
779
- 780 Helsel, DR and Hirsch, RM (2002) Statistical methods in water resources techniques of
781 water resources investigations. Book 4, chapter A3. U.S. Geological Survey. 522 p.
782
- 783 Herrera S, Gutiérrez J, Ancell R, Pons M, Frías M, Fernández J (2012) Development and
784 analysis of a 50-year high-resolution daily gridded precipitation dataset over Spain
785 (Spain02). *International Journal of Climatology* 32: 74-85. doi: 10.1002/joc.2256.
786
- 787 Hijmans RJ, Cameron SE, Parra JL, Jones PG, Jarvis A (2005) Very high resolution
788 interpolated climate surfaces for global land areas. *International Journal of Climatology*.
789 25(15): 1965–1978. doi: 10.1002/joc.1276.
790
- 791 Hubbard KG, Goddard S, Sorensen WD, Wells N, Osugi TT (2005) Performance of
792 quality assurance procedures for an applied climate information system. *Journal of*
793 *Atmospheric Oceanic Technology*, 22: 105–112. doi: 10.1175/JTECH-1657.1
794
- 795 Hwang Y, Clark M, Rajagopalan B, Leavesley G (2012) Spatial interpolation schemes of
796 daily precipitation for hydrologic modeling. *Stochastic Environmental Research and Risk*
797 *Assessment*. 26(2): 295–320. doi: 10.1007/s00477-011-0509-1.
798
- 799 Klok E and Klein-Tank A (2009) Updated and extended European dataset of daily climate
800 observations. *International Journal of Climatology*.29(8): 1182–1191. doi:
801 10.1002/joc.1779.

- Kleiber W, Katz RW, Rajagopalan B (2012) Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. *Water Resources Research* 48: W01523. doi:10.1029/2011WR011105
- Klein-Tank AMG, Wijngaard JB, Können GP, Böhm R, Demarée G, Gocheva A, Mileta M, Pashiardis S, Hejkrlik L, Kern-Hansen C, Heino R, Bessemoulin P, Müller-Westernmeier G, Tzanakou M, Szalai S, Pálsdóttir T, Fitzgerald D, Rubin S, Capaldo M, Maugeri M, Leitass A, Bukantis A, Aberfeld R, van Engelen AFV, Forland E, Miletus M, Coelho F, Mares C, Razuvaev V, Nieplova E, Cegnar T, López JA, Dahlström B, Moberg A, Kirchhofer W, Ceylan A, Pachaliuk O, Alexander LV, Petrovic P (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *International Journal of Climatology*. 22:1441–1453. doi: 10.1002/joc.773
- Lawrimore JH, Menne MJ, Gleason BE (2011) An overview of the Global Historical Climatology Network monthly mean temperature data set, version 3. *Journal of Geophysical Research*. 116: D19121. doi: 10.1029/2011JD016187.
- Li Q, Zhang H, Liu X, Chen J, Li W (2009) A mainland China homogenized historical temperature dataset of 1951-2004. *Bulletin of the American Meteorological Society*. 90, 1062–1065. doi: 10.1175/2009BAMS2736.1
- López-Moreno J.I., Vicente-Serrano S.M., Angulo-Martínez M., Beguería S., Kenawy A (2010) Trends in daily precipitation on the northeastern Iberian Peninsula, 1955–2006. *International Journal of Climatology* 30, 1026–1041.
- Menne MJ, Durre I, Vose RS, Gleason BE, Houston TG (2012) An overview of the global historical climatology network-daily database. *Journal of Atmospheric and Oceanic Technology*. doi: 10.1175/JTECH-D-11-00103.1
- Min SK, Zhang K, Zwiers FW, Hegerl GC (2011) Human contribution to more-intense precipitation extremes. *Nature*. 470, 378-381. doi: 10.1038/nature09763
- Mitchell TD and Jones PD (2005.) An improved method of constructing a database of monthly climate observations and associated high-resolution grids. *International Journal of Climatology*. 25:693–712.
- Moncho R, Caselles V, Chust G (2012) Alternative model for precipitation probability distribution: application to Spain. *Climate Research*. 51: 23-33. doi: 10.3354/cr01055
- New M, Hulme M, Jones PD (1999) Representing twentieth-century space-time climate variability. Part I: development of a 1961-90 mean monthly terrestrial climatology. *Journal of Climate*. 12:829–856.
- New M, Hulme M, Jones PD (2000) Representing twentieth century space-time climate variability. II: development of 1901–1996 monthly grids of terrestrial surface climate. *Journal of Climate*. 13:2217–2238.

- New M, Lister D, Hulme M, Makin I (2002) A high-resolution data set of surface climate over global land areas. *Climate Research*. 21:1-25.
- Neykov N, Neytchev P, Zucchini W (2014) Stochastic daily precipitation model with a heavy-tailed component. *Natural Hazards and Earth System Science* 14(9): 2321-2335. doi: 10.5194/nhess-14-2321-2014.
- Ninyerola M, Pons X, Roure JM (2007) Monthly precipitation mapping of the Iberian Peninsula using spatial interpolation tools implemented in a Geographic Information System. *Theoretical and Applied Climatology* 89(3):195-209. doi: 10.1007/s00704-006-0264-2.
- Norrant C and Douguédroit A (2005) Monthly and daily precipitation trends in the Mediterranean (1950–2000). *Theoretical and Applied Climatology*. 83(1-4): 89–106. doi: 10.1007/s00704-005-0163-y.
- Orlowsky B and Sereviratne SI (2012) Global changes in extreme events: regional and seasonal dimension. *Climatic Change*. 110(3): 669-696. doi: 10.1007/s10584-011-0122-9
- Perry M, Hollis D (2005) The generation of monthly gridded datasets for a range of climatic variables over the UK. *International Journal of Climatology*. 25: 1041–1054.
- Picarreta M, Pasini A, Capolongo D, Lazzari M (2013) Changes in daily precipitation extremes in the Mediterranean from 1951 to 2010: the Basilicata region, southern Italy. *International Journal of Climatology*. 33(15):3229-3248. doi: 10.1002/joc.3670.
- R Core Team (2016) A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0. URL <https://www.R-project.org/>.
- Reek T, Doty SR, Owen TW (1992) A deterministic approach to the validation of historical daily temperature and precipitation data from the cooperative network. *Bulletin of the American Meteorological Society*. 73: 753–762.
- Río S, Herrero L, Fraile R, Penas A (2011) Spatial distribution of recent rainfall trends in Spain (1961-2006). *International Journal of Climatology*. 31(5): 656–667. doi: 10.1002/joc.2111.
- Schneider U, Becker A, Finger P, Meyer-Christoffer A, Ziese M, Rudolf B (2014) GPCC's new land surface precipitation climatology based on quality-controlled in situ data and its role in quantifying the global water cycle. *Theoretical Applied Climatology*. doi: 10.1007/s00704-013-0860-x
- Serrano-Notivol R, de Luis M, Beguería S (2017) An R package for daily precipitation climate series reconstruction. *Environmental Modelling & Software*. 89:190-195. doi: 10.1016/j.envsoft.2016.11.005

- Simolo C, Brunetti M, Maugeri M, Nanni T (2010) Improving estimation of missing values in daily precipitation series by a probability density function-preserving approach. *International Journal of Climatology*. 38: L16701 DOI: 10.1002/joc.1992.
- Su-Ping N, Yong L, Wei-Ping L, Tong-Wen W, Xue-Li S, Zai-Zhi W (2012) Quality Control and Analysis of Global Gauge-Based Daily Precipitation Dataset from 1980 to 2009. *Advances in Climate Change Research*. 3(1): 45-53. doi: 10.3724/SP.J.1248.2012.00045.
- Sumner G, Ramis C, Guijarro JA (1995a) Daily rainfall domains in Mallorca. *Theoretical and Applied Climatology*. 51: 199-221. doi: 10.1007/BF00867280.
- Sumner G, Guijarro JA, Ramis C (1995b) The impact of surface circulation on significant daily rainfall patterns over Mallorca. *International Journal of Climatology* 15: 673-696. doi: 10.1002/joc.3370150607
- Syed K, Goodrich D, Myers D, Sorooshian S (2003) Spatial characteristics of thunderstorm rainfall fields and their relation to runoff. *Journal of Hydrology* 271(1-4):1-21. doi: 10.1016/S0022-1694(02)00311-6.
- Trenberth K (2011) Changes in precipitation with climate change. *Climate Research* 47: 123-138. doi: 10.3354/cr00953.
- Trenberth K, Dai A, van der Schrier G, Jones PD, Barichivich J, Briffa KR, Sheffield J (2014) Global warming and changes in droughts. *Nature Climate Change* 4: 17-22. doi: 10.1038/NCLIMATE2067
- Verdin A, Rajagopalan B, Kleiber W, Katz RW (2015) Coupled stochastic weather generation using spatial and generalized linear models. *Stochastic Environmental Research Risk Assessment*. 29:347–356. doi:10.1007/s00477-014-0911-6
- Vicente-Serrano S, Beguería S, López-Moreno J, García-Vera M, Stepanek P (2010) A complete daily precipitation database for northeast Spain: reconstruction, quality control, and homogeneity. *International Journal of Climatology*. 30: 1146-1163. doi: 10.1002/joc.1850.
- Wong G, Maraun D, Vrac M, Widmann M, Eden J, Kent T (2014) Stochastic Model Output Statistics for Bias Correcting and Downscaling Precipitation Including Extremes. *Journal of Climate* 27(18): 6940-6959. doi: 10.1175/JCLI-D-13-00604.1.
- Yatagai A, Kamiguchi K, Arakawa O, Hamada A, Yasutomi N, Kitoh A (2012) APHRODITE: Constructing a Long-Term Daily Gridded Precipitation Dataset for Asia Based on a Dense Network of Rain Gauges. *Bulletin of the American Meteorological Society*. 93(9): 1401-1415. doi: 10.1175/BAMS-D-11-00122.1.
- Zhang X, Alexander L, Hegerl G, Jones P, Tank A, Peterson T, Trewin B, Zwiers F (2011) Indices for monitoring changes in extremes based on daily temperature and

945 precipitation data. Wiley Interdisciplinary Reviews: Climate Change. doi:
946 10.1002/wcc.147.
947
948 Zwiers FW, Alexander LV, Hegerl GC, Knutson TR, Kossin JP, Naveau P, Nicholls N,
949 Schär C, Seneviratne SI, Zhang X (2013) Climate Extremes: Challenges in Estimating
950 and Understanding Recent Changes in the Frequency and Intensity of Extreme Climate
951 and Weather Events. Climate Science for Serving Society. 339–389 doi:10.1007/978-94-
952 007-6692-1_13
953

Note: Iterative process. One loop means doing these tasks in all stations for each day. After every loop the 10 NNS (Near Stations with observed value) are recomputed. When a loop do not detect anomalies in any station, the process stops.

Figure 1. Workflow of quality control process.

Figure 2. Spatial structure of the daily precipitation dataset: location of the stations (black dots) used in this study.

Figure 3. Temporal structure and evolution of the daily precipitation dataset. Upper panel: Black lines indicate the days with observed data in each station. In an ideal situation with no gaps in observations, a black surface will be shown. Lower panel: Number of daily available data (grey lines) and 365 days moving average (black line).

Figure 4. Quality control results. Annual (a) and monthly (b) percentage of removed data using the five criteria (stacked bars). Frequency distribution of observations removed by QC.1 (Isolated wet conditions) (c), QC.3 (Suspect outliers) (d) and QC.5 (Suspect wet) (f). Frequency distribution of estimations used to remove observations by QC.4 (Suspect dry) (e).

Figure 5. Frequency distribution of binomial predictions of occurrence of precipitation (BP) in dry (a) and wet (b) days and spatial distribution of percentage of wet days estimated as dry (c) and percentage of dry days estimated as wet (d).

Figure 6. Frequency distribution of observations and estimations (days with zero in observations and estimations are not shown). Precipitation classes are created in logarithmic scale.

Figure 7. Comparison of predicted reference values (RV) and observed values, by days:

a) means of all available stations, considering all wet and dry days; b) medians of all available stations, considering only wet days; c) same as b) but considering only days exceeding the 95th percentile of the wet days; d), e), and f), density curves corresponding to a), b), and c).

Figure 8. Comparison of predicted RV and observed values, by stations: a) means of all available days, considering all wet and dry days; b) medians of all available days, considering only wet days; c) same as b), but considering only days exceeding the 95th percentile of the wet days; d), e), and f), density curves corresponding to a), b), and c).

Figure 9. Ratio of means (mean of estimations / mean of observations).

1003
1004
1005
1006

Figure 10. Reconstruction of “B249” observatory. Time series of daily observed and reconstructed values and correlation between them (upper panel); time series of monthly observed and reconstructed totals (lines) (lower panel) with the information about the months when any missing value was filled (lines on top) and correlation between them.

Figure 11. Newly created series: a) daily values; b) monthly totals; c) annual totals.

Shaded areas show the 95% confidence bands.

Figure 12. Maps of precipitation indices (left of each panel) and uncertainty of predictions (right of each panel) computed from the grid. Annual precipitation (a), number of days with precipitation over 20 mm (R20mm) (b), median of daily precipitation in wet days (PMED) (c), maximum consecutive dry days (CDD) (d), maximum precipitation in one day (RX1) (e), and maximum consecutive wet days (CWD) (f).

Table 1. Contingency table with the percentage of true and false zero and positive. In bold are the accuracy of true positives. True zero: predicted and observed precipitation values are zero. False zero: observed value is zero but predicted value is positive. True positive: predicted and observed precipitation values are positive. False positive: observed value is positive but predicted value is zero.

	True zero	False zero	True positive	False positive
All	96.36 %	3.58 %	82.84 %	17.16 %
JAN	95.37 %	4.63 %	84.45 %	15.55 %
FEB	94.96 %	5.04 %	83.47 %	16.53 %
MAR	95.88 %	4.12 %	82.70 %	17.30 %
APR	95.37 %	4.63 %	82.00 %	18.00 %
MAY	96.94 %	3.06 %	82.00 %	18.00 %
JUN	98.25 %	1.75 %	77.70 %	22.30 %
JUL	99.06 %	0.94 %	70.78 %	29.22 %
AUG	98.41 %	1.59 %	78.15 %	21.85 %
SEP	96.77 %	3.23 %	84.06 %	15.94 %
OCT	95.33 %	4.67 %	84.42 %	15.58 %
NOV	94.62 %	5.38 %	84.86 %	15.14 %
DEC	94.13 %	5.87 %	82.54 %	17.46 %

Table 2. Statistics for each decile (1 to 10) and the range of values included in each one (in mm). %OBS, %PRED: percentage of total precipitation corresponding to each decile, in observed and predicted RV data, respectively; MAE: Mean Absolute Error (in mm) and ME: Mean Error (in mm).

Decile	Range (mm)	%OBS	%PRED	MAE (mm)	ME (mm)
1	> 0–1.0	0.6	1.7	1.30	1.12
2	> 1.0–2.0	1.6	2.6	1.65	1.08
3	> 2.0–3.0	2.1	2.9	1.96	1.02
4	> 3.0–4.4	3.2	4	2.28	0.91
5	> 4.4–6.0	4.8	5.6	2.74	0.86
6	> 6.0–8.4	6.5	7.1	3.23	0.72
7	> 8.4–11.6	9	9.4	3.81	0.44
8	> 11.6–16.5	12.6	12.6	4.69	0
9	> 16.5–26.1	18.7	17.9	6.27	–0.89
10	> 26.1	41	36.1	12.30	–5.45

Table 3. Goodness-of-fit by month and the range of values included in each one (in mm).
 MAE (in mm), ME (in mm), RM (Ratio of Means), and RSD (Ratio of Standard
 Deviations).

	Range (mm)	MAE (mm)	ME (mm)	RM	RSD
JAN	0 to 407.2	8.17	0.65	1.01	1.01
FEB	0 to 428.0	7.10	0.27	1.01	1.02
MAR	0 to 408.1	6.99	0.42	1.01	1.00
APR	0 to 429.0	7.93	0.45	1.01	1.01
MAY	0 to 365.5	7.26	0.26	1.01	1.01
JUN	0 to 172.0	4.16	0.09	1.01	1.01
JUL	0 to 213.0	2.75	0.03	1.00	1.01
AGO	0 to 244.0	6.38	0.08	1.00	1.02
SEP	0 to 430.0	12.45	0.39	1.01	1.00
OCT	0 to 478.0	13.29	1.08	1.01	1.01
NOV	0 to 687.5	14.21	-3.65	0.96	0.96
DEC	0 to 537.2	12.94	-3.49	0.95	0.96