The RheaG Weather Generator Algorithm: Evaluation in Four Contrasting Climates from the Iberian Peninsula[✉]

DANIEL NADAL-SALA

*Ecology Section, Department of Evolutionary Biology, Ecology and Environmental Sciences,
University of Barcelona, Barcelona, Spain*

CARLOS A. GRACIA AND SANTIAGO SABATÉ

*Ecology Section, Department of Evolutionary Biology, Ecology and Environmental Sciences, University of Barcelona,
Barcelona, and Center for Ecological Research and Forestry Applications, Cerdanyola del Vallès, Spain*

(Manuscript received 26 June 2018, in final form 1 October 2018)

ABSTRACT

This paper describes the assumptions, equations, and procedures of the RheaG weather generator algorithm (WGA). RheaG was conceived for the generation of robust daily meteorological time series, whether in static or transient climate conditions. Here we analyze its performance in four Iberian locations—Bilbao, Barcelona, Madrid, and Sevilla—with differentiated climate characteristics. To validate the RheaG WGA, we compared observed and generated meteorological time series' statistical properties of precipitation, maximum temperature, and minimum temperature for all four locations. We also compared observed and simulated rain events spell length probabilities in all four locations. Finally, RheaG includes two weather generation procedures: one in which monthly mean values for meteorological variables are unconstrained and one in which they are constrained according to a predefined baseline climate variability. Here, we compare the two weather generation procedures included in RheaG using the observed data from Barcelona. Our results present a high agreement in the statistical properties and the rain spell length probabilities between observed and generated meteorological time series. Our results show that RheaG accurately reproduces seasonal patterns of the observed meteorological time series for all four locations, and it is even able to differentiate two climatic seasons in Bilbao that are also present in the observed data. We find a trade-off between generation procedures in which the unconstrained procedure better reproduces the variability of monthly and yearly precipitation than the constrained one, but the constrained procedure is able to keep the same climatic signal across meteorological time series. Thus, the first procedure is more accurate, but the latter is able to maintain spatial autocorrelation among generated meteorological time series.

1. Introduction

Ongoing human-induced climate change is already affecting both human societies and natural systems all across the globe. Its negative impacts are projected to increase in the coming decades (IPCC 2013). Simulation models are the most reliable ways to explore how changes in environmental conditions will likely modify present natural systems in the future. Future climate is unknown, but simulation models can be fed with climate

projections generated statistically from the meteorological record. Thus, developing tools to provide statistically robust future meteorological time series is crucial. Moreover, even when those projections are available, most of the time their temporal or spatial scale does not fit with simulation models' requirements. In that case, further transformations shall be applied upon global climate change projections to downscale them to a site-specific data. This fact is not trivial, as further downscaling needs to be also accurate and statistically robust both with global projections and local-scale observations. Furthermore, in most locations daily meteorological time series are not available. They also might have large gaps in their records. In both cases, current or past meteorological time series have to be reconstructed in order to feed simulation models with

[✉] Supplemental information related to this paper is available at the Journals Online website: <https://doi.org/10.1175/JAMC-D-18-0170.s1>.

Corresponding author: Daniel Nadal-Sala, d_nadal@ub.edu

reliable meteorological information to evaluate their performance under current climate conditions.

There have been many approaches in the development of algorithms for meteorological series generation since the mid-twentieth century (e.g., [Gabriel and Neumann 1962](#); [Richardson 1981](#); [Richardson and Wright 1984](#); [Bürger 1997](#))—see [Ailliot et al. \(2015\)](#) for a review. Generally speaking, weather generation algorithms (WGA) are based on summarizing observed climate data time series into a set of statistical descriptors in order to use them to generate artificial climate data statistically similar to the observed ones ([Wilks and Wilby 1999](#)). Traditionally, climate variables included in WGA have been temperature, incident radiation, wind speed, and the most important, precipitation. Most of WGA include temporal autocorrelation between their variables in their generated data. Many of them correlate climate variables to precipitation, as it is identified as the most difficult climate variable to be properly generated ([Hutchinson 1986](#)). In contrast to most global circulation models' (GCMs') fluid dynamics differential-equations-based projections, WGAs tend to represent better local climate variations because of their specificity to the statistical characteristics of a given local climate time series. However, the parameterization effort of those models is substantial for each location, and they need long observed climate time series to be parameterized, so their applicability to a large area is still challenging ([Semenov and Brooks 1999](#)). Still, WGAs also have to deal with the trade-off between the number of descriptive parameters and the relative information gained per added parameter ([Hawkins 2004](#)). The aim is to avoid irrelevant or roughly noninformative parameters that increase the minimum length of the observed meteorological time series required to calibrate them. In the end, each WGA is a trade-off between the likelihood and statistical robustness of its generated meteorological time series and the number of parameters required.

With this paper we propose a new Markovian transition probability-based WGA, and we evaluate its performance when compared to four contrasting climates from the Iberian Peninsula. We first compare the absolute values of observed and generated meteorological time series at four different locations representative of the different Iberian climates—that is, Barcelona, Bilbao, Madrid, and Sevilla—to account for the verisimilitude of the generated meteorological time series compared to the observed ones. Then, we compare generated versus observed precipitation patterns, focusing our attention to rain spell length distributions. Finally, we compare the two alternative procedures included in RheaG for meteorological data generation under different climate forcing scenarios for Barcelona.

2. Material and methods

RheaG is a WGA programmed in R coding language ([R Core Team 2017](#)) in order to make it freely and easily available. RheaG analyzes certain statistical properties of an input observed daily climate data and then uses these properties, along with a pseudorandom number generator, to produce simulated daily meteorological data one day at a time, thus reproducing observed climate statistical patterns. [Table 1](#) summarizes the meteorological inputs and its description in the RheaG.

Daily meteorological variables included in the RheaG are divided into two categories: the mandatory ones—that is, incident radiation Q ($\text{MJ ha}^{-1} \text{ day}^{-1}$), precipitation P (mm day^{-1}), and maximum and minimum temperatures (T_{max} and T_{min} ; $^{\circ}\text{C}$)—and the discretionary ones—wind speed (W_s ; m s^{-1}), atmospheric water pressure (VP ; kPa), and potential evapotranspiration (PET ; mm day^{-1}). The statistical properties obtained from the observed meteorological time series for T_{max} , T_{min} , Q , W_s , VP , and PET are the month daily mean and standard deviation values, along with the monthly first-order Markovian transition probability matrix. The temperature and radiation Markovian transition probability matrix is defined by the probability of observing a value above or below the monthly mean given that the value of the previous day was above or below the monthly mean. Statistical properties describing precipitation patterns in RheaG are as follows: annual mean precipitation during a rainy day; monthly mean precipitation during a rainy day; rain occurrence probability defined by 15 rain occurrence probability classes, following a $p[2^{(n-1)} < x < 2^n]$ distribution, where $n \in I\{1, 2, \dots, 15\}$; and last, the monthly Markovian transition probability matrix of rain occurrence given whether it rained the previous day.

It is noteworthy that, in absence of radiation inputs, the RheaG algorithm is still able to calculate them by following [Hargreaves and Samani \(1982\)](#), according to Eq. (1):

$$R_t = K_t(T_{\text{max}} - T_{\text{min}})^{0.5} R_a, \quad (1)$$

where R_t is the daily radiation ($\text{MJ ha}^{-1} \text{ day}^{-1}$); T_{max} and T_{min} are maximum and minimum temperatures, respectively, ($^{\circ}\text{C}$); and R_a is the extraterrestrial radiation for a given location and moment of year ($\text{MJ ha}^{-1} \text{ day}^{-1}$) and is calculated from the latitude of the location in degrees and the Julian day ([Spencer 1971](#); [Duffie and Beckman 1991](#)). The empirical coefficient K_t depends on the continentality of the site. For coastal sites, K_t equals 0.20, and for the interior regions it equals 0.17, according to [Allen \(1995\)](#).

TABLE 1. Meteorological variables considered in the RheaG weather generator. List of parameters needed for each variable in the model and the probability function used to obtain the daily values. Time autocorrelation indicates the probability matrix used to relate values of a given variable at day t to values obtained at day $t - 1$. If there is no time autocorrelation, values of day t are independent from the day $t - 1$. Inferred values column indicates when RheaG may generate meteorological values of this variable without a previous measurement.

Meteorological variable	Parameters	Probability function	Time autocorrelation	Inferred
Minimum temperature (MinT; °C)	Monthly mean MinT Monthly std dev MinT	$T_{\min_{\text{day}}} = N(\mu, \sigma^2)$	First-order Markovian	No
Maximum temperature (MaxT; °C)	Monthly mean MaxT Monthly std dev MaxT	$T_{\max_{\text{day}}} = N(\mu, \sigma^2)$	First-order Markovian	No
Vapor pressure (VP; kPa)	Monthly mean VP Monthly std dev VP	$VP_{\text{day}} = N(\mu, \sigma^2)$	First-order Markovian	Yes
Radiation (Q ; MJ m ⁻² day ⁻¹)	Monthly mean Q Monthly std dev Q	$Q_{\text{day}} = N(\mu, \sigma^2)$	First-order Markovian	No
Daily precipitation occurrence	None	Binomial (yes, no)	First-order Markovian	No
Precipitation amount (P_{day} ; mm day ⁻¹)	Probability distribution of rain events based on a $p[2^{(n-1)} < x \leq 2^n]P_{\text{day}}$ per rainy day distribution Monthly mean rainfall per rainy day Yearly mean rainfall per rainy day	$P_{\text{day}} = (1 + x^2)[2^{(n-2)}](P_{\text{month}}/P_{\text{year}}); n \in I\{1, 15\}$	No	No

In RheaG, discretionary variables, if not available, are calculated daily as follows:

Saturated vapor pressure (VPS; kPa) is calculated daily following Allen et al. (1998) [Eq. (2)]:

$$VPS = 0.614 \exp \left[\frac{T_{\min} \left(18.56 - \frac{T_{\min}}{254.4} \right)}{(T_{\min} + 255.57)} \right], \quad (2)$$

where VPS is the saturated water vapor pressure (kPa), and $T_{\min}$ is the minimum daily temperature (°C). Kimball et al. (1997) found that predawn air water vapor pressure does not reach saturation point in the most arid climates, and they proposed a correction upon VPS based on the proportion of annual precipitation related to potential evapotranspiration. If annual precipitation P_{365} is less than a half of potential evapotranspiration

(PET_{365}), a correction is applied upon VPS following Eq. (3):

$$VPS_{\text{corrected}} = 1.2 \left(\frac{P_{365}}{PET_{365}} + 0.4 \right) VPS, \quad (3)$$

where $VPS_{\text{corrected}}$ is the saturated water vapor pressure after applying Kimball's correction (kPa), VPS is the estimated saturated water vapor pressure (kPa), and P_{365} and PET_{365} are the last 365 days precipitation and potential evapotranspiration, respectively (mm yr⁻¹).

Actual VP is calculated by changing daily $T_{\min}$ mean daily temperature [i.e., $(T_{\max} + T_{\min})/2$] in Eq. (2).

Daily PET is calculated following the FAO Penman–Monteith equation described in Allen et al. (1998) [Eq. (4)]:

$$PET = \frac{0.408\Delta(Q - G) + \gamma \frac{900}{(T_{\text{mean}} + 273)} Ws(VPS - VP)}{\Delta + \gamma(1 + 0.34Ws)}, \quad (4)$$

where Δ is the slope vapor pressure curve (kPa °C⁻¹) given daily mean temperature; Q is the daily incident radiation (MJ ha⁻¹ day⁻¹); G is the soil heat flux density (MJ ha⁻¹ day⁻¹), fixed at a 10%; γ is the psychrometric constant (kPa °C⁻¹) given the altitude of the plot; T_{mean} is the daily mean temperature (°C); Ws is the daily mean wind speed (m s⁻¹); VPS is the saturation

vapor pressure (kPa); and VP is the actual vapor pressure (kPa).

Wind speed is often not available in original data. In that case, Ws RheaG generates daily Ws data using a Gaussian distribution $Ws_{\text{day}} = N(\mu, \sigma^2)$, with prescribed monthly mean and variance values according to the previous knowledge for each location.

It is noteworthy that, despite in some of the other WGA there are some kinds of correlations among Tmax and Tmin, Q , PET, and VPS with daily precipitation (e.g., Wilks and Wilby 1999; Kilsby et al. 2007), RheaG does not include any a priori correlations among them, in spite of maintaining a low number of parameters required to run the algorithm—but see the constrained generation method below. We are aware that independent meteorological data generation for each variable in some cases may lead to physically unrealistic daily combinations (e.g., high Q values generated during extended rain events), while still presenting consistent monthly and yearly variable values. Other approaches (e.g., Kilsby et al. 2007), based on hourly observed data, have proven to significantly improve the agreement between generated and observed meteorological time series, while they consistently kept correlations between precipitation, temperature, and radiation. However, the main purpose of RheaG WGA is to feed simulation vegetation models with site-specific meteorological time series. That kind of meteorological record is often found in the form of daily values. Besides, in the past they covered too few years to capture straightforward correlations between daily meteorological variables and daily rain intensity. Slightly negative correlations between daily precipitation during the days when $P > 0$ and Tmax ($r_{\text{Barcelona}} = -0.22$, $n_{\text{Barcelona}} = 2065$; $r_{\text{Sevilla}} = -0.15$, $n_{\text{Sevilla}} = 1672$), Tmin ($r_{\text{Barcelona}} = -0.17$, $n_{\text{Barcelona}} = 2065$; $r_{\text{Sevilla}} = 0.01$, $n_{\text{Sevilla}} = 1672$) and Q ($r_{\text{Barcelona}} = -0.13$, $n_{\text{Barcelona}} = 2065$; $r_{\text{Sevilla}} = -0.18$, $n_{\text{Sevilla}} = 1672$) were observed in our data, with elevated intra-annual variability (see Figs. SM1 and SM2 and Table SM1 in the online supplemental material). Furthermore, it has been observed that sensitivity to biases in input precipitation and temperature for simulation models occurs mostly at monthly—rather than daily—levels (e.g., Wood et al. 2004; Fowler and Kilsby 2007). Hence, we opted to apply the most parsimonious criteria of not accounting for correlations within variables at a daily scale in order to reduce parameter number input.

a. Precipitation

Precipitation tends to be the main driver variable in most of WGA (Kilsby et al. 2007). First-order Markovian transition probability matrices (MM) have been largely used to model the daily occurrence of climate events (Gabriel and Neumann 1962; Caskey 1963; Stern 1980; Sonnadara and Jayewardene 2015). They are a simple but effective way to model wet and dry spells' length, frequency, and distribution if the original meteorological series analyzed are long enough (Katz 1981). However, they only generate binomial information—rain or no rain—for a single day. Thus, rain occurrence

models must be coupled to rain amount distribution models to generate the meteorological time series. Exponential, potential, gamma, or mixed exponential rainfall amount distribution models have been successfully applied in several weather generator algorithms (Wilks and Wilby 1999; Ailliot et al. 2015). RheaG uses monthly based MM to determine if it rains during a given day, given that it rained the previous day. In the case of rain presence, RheaG describes the amount of precipitation during a rainy day from 15 probability classes in $p[2^{(n-2)}] < P_{\text{amount}} \leq 2^{(n-1)}$, where $n \in I\{1, 2, \dots, 15\}$, and P_{amount} is the amount of precipitation during a rainy day. RheaG first randomly determines the nP_{amount} class for a given day, from the yearly probability distribution of belonging to n precipitation class, obtained from observed data. Then, RheaG generates actual precipitation for this day randomly, following Eq. (5):

$$P_{\text{day}} = (1 + x^2)[2^{(n-2)}] \frac{P_{\text{month}}}{P_{\text{year}}}, \quad (5)$$

where P_{day} is the amount of precipitation during a given day (mm day^{-1}), x is a random number between 0 and 1, n refers to the precipitation class— $n \in I\{1, 2, \dots, 15\}$, P_{month} is the mean daily precipitation for a given month, and P_{year} is the yearly mean daily precipitation.

Here we evaluate generated precipitation in two ways: first, we analyze the ability of RheaG to reproduce the statistical properties of the generated meteorological time series—that is, Markovian transition probabilities, rain occurrence probability, rain intensity, and probability distribution. Second, we compare the wet spell length probability between reference climate time series and generated time series. It is not explicitly included in the input for the RheaG, although is closely linked to the Markovian transition matrix. Hence, it is a good estimator of the agreement among statistical properties of observed and generated meteorological time series. As most of the complexity in weather generation relies on accurately describing precipitation patterns (Wilks and Wilby 1999), we set the main focus of this work on precipitation patterns in order to validate RheaG meteorological time series.

b. Maximum and minimum daily temperatures and daily radiation

Daily maximum and minimum temperatures ($^{\circ}\text{C}$) in a monthly basis—hereafter referred as “temperature values”—have been successfully modeled from Gaussian distributions (e.g., Mearns et al. 1984; Kilsby et al. 2007), and so have been implemented in the RheaG WGA, coupled to a first-order Markovian transition

matrix to account for daily autocorrelation in temperatures, following Eq. (6):

$$T_j = N(\text{Ref}_j, \sigma^2), \quad (6)$$

where T_j is the temperature value for a given j day ($^{\circ}\text{C}$), Ref_j is the reference mean temperature on the j th day ($^{\circ}\text{C}$), and σ^2 is the standard deviation of monthly temperature values. In addition, RheaG includes a MM-based probability to account for temporal autocorrelation trends in temperature. Therefore, a given day is going to add or subtract the standard deviation to the Ref_j values according to a MM-based probability.

Ref_j is a correction upon reference mean monthly temperature. It is calculated following Eq. (7) or Eq. (8), depending on being the day j within the first or the second half of the month, respectively:

$$\text{Ref}_j = \mu_m + \frac{\mu_m - \mu_{(m-1)}}{d_m}(d_j - d_{c,m}) \quad \text{and} \quad (7)$$

$$\text{Ref}_j = \mu_m + \frac{\mu_m - \mu_{(m+1)}}{d_m}(d_j - d_{c,m}), \quad (8)$$

where Ref_j is the reference mean temperature on the j th day; μ_m is the mean monthly temperature of that given month ($^{\circ}\text{C}$); $\mu_{(m-1)}$ and $\mu_{(m+1)}$ are the mean monthly temperature values for the previous and the following month ($^{\circ}\text{C}$), respectively; d_m is the number of days within the month m ; d_j is the current Julian day; and $d_{c,m}$ is the Julian day at the middle of the current month. These corrections are used to smooth the transition of the generated climate time series between one month and the previous or following month, alternatively. Last, RheaG recalculates daily temperature values if $T_{\text{max}} \leq T_{\text{min}}$ for a given day in order to avoid unrealistic output values.

Surface incident radiation Q ($\text{MJ ha}^{-1} \text{ day}^{-1}$) for a given day is obtained following exactly the same procedure.

c. Generation of daily meteorological time series under no climate forcing

To validate RheaG-generated time series, we selected four climate time series from the Agencia Estatal de Meteorología (AEMET) Spanish Meteorological Service (<http://www.aemet.es/es/portada>). This climate time series are representative of different climates present in the Spanish mainland. Bilbao (43.17°N , 25.42°W) represents the wet, northwestern Atlantic climate; Barcelona (41.25°N , 20.73°E) represents the moist Mediterranean-type climate; Sevilla (37.09°N , 53.66°W) represents the arid Mediterranean-type climate; and Madrid (40.24°N , 34.04°W) represents the continental climate. We used AEMET climate time series from 1989

to 2009 for all locations, hereafter referred as validation locations, to account for at least 20 years (Richardson and Wright 1984) of meteorological recorded data. In addition, we selected this period to avoid the longer gaps present in the register. Gaps in temperature and radiation were filled as follows: gaps shorter than 2 days were filled by linear interpolation. Gaps larger than 2 days were interpolated linearly from nearby meteorological stations. In addition, gaps in precipitation were all interpolated linearly from nearby meteorological stations. Only meteorological stations with $R^2 \geq 0.8$ were used to fill those gaps. Under current climate change conditions, it is very likely (IPCC 2013) a modification of climate statistical properties. So, by using only the last 21 years available, we made sure to obtain both the current means and standard deviations, as well as the most recent Markovian probability distribution. For each location, five independent 90-yr-long daily climate data series were then generated without any monthly data restriction, or any further climate forcing, to account for the ability of RheaG to reproduce the observed climate pattern in the original meteorological time series. We refer to these runs as reference (R) runs, as they do not include further forcing than observed in generated time series.

d. Generating time series under climate forcing

RheaG also includes the option to generate meteorological data with climate forcing. RheaG includes two procedures for doing so. The first one, called the unconstrained one because monthly values for different meteorological time series are independent, generates climate just as described above for all years required, and at the end it adds three monthly cumulative anomalies for the values of T_{max} , T_{min} , and P , respectively. These anomalies are the monthly increment per year projected in T_{max} and T_{min} and the yearly percentage reduction in precipitation on a monthly basis.

The second procedure for meteorological time series generation under climate forcing, so-called the constrained one, first calculates a baseline monthly variability to be included in all time series to be generated. This monthly variation is generated randomly for T_{max} , T_{min} , Q , and P from a truncated normal distribution centered in the mean monthly values, with a variation that is function of the observed monthly standard deviation for this variable multiplied by a Gaussian-probability-based factor. It includes the option to account for autocorrelations among variables (e.g., months with positive increase in maximum temperature with respect to the mean are more likely to have positive increases in minimum temperature with respect to the mean). Final monthly values for a given meteorological

TABLE 2. Climate forcing scenarios considered, following [IPCC \(2013\)](#). SPC indicates the stationary-at-present climate scenario, with no further climate forcing. RCP4.5 is the moderate-climate-change scenario (climate forcing equal to 4.5 W m^{-2}), and RCP8.5 is the extreme-climate-change scenario (climate forcing equal to 8.5 W m^{-2}). In the columns, ΔT_{max} and ΔT_{min} are the anomaly in T_{max} and T_{min} projected by the year 2100 with respect to the reference period (1990–2010), and ΔP is the anomaly in precipitation projected by the year 2100 with respect to the reference period.

Climate scenario	$\Delta T_{\text{max}} (\text{°C century}^{-1})$	$\Delta T_{\text{min}} (\text{°C century}^{-1})$	$\Delta P (\% \text{ century}^{-1})$
SPC	0	0	0
RCP4.5	2.225	2.225	−25
RCP8.5	5.6	5.6	−35

variable are calculated as follows: for a given month i and during a given year, the reference value ($\text{Ref}_{\text{month}}$) for a given variable x follows the equation $\text{Ref}_{\text{month}} = N(\mu_{x,i} + \Delta x_i, a_{x,i}\sigma_i^2)$, where $\mu_{x,i}$ is the observed mean monthly value for this given variable, Δx_i is the anomaly of this variable under climate change conditions for the given month, $a_{x,i}$ is the variability for a given x variable and i month—with values ranging between -1.5 and $+1.5$ and following a truncated Gaussian distribution centered in 0; and σ_i^2 is the observed standard deviation for the given x climate variable during i month. The intention here is to maintain the same variation signal among generated meteorological time series for a given month when they are spatially autocorrelated. The $a_{x,i}$ coefficient is, then, used to maintain the same underlying climate signal for the same month among different meteorological time series. It is noticeable here that absolute individual monthly values can also be included as an input for RheaG, so this WGA is able to disaggregate the more common monthly meteorological projections into the daily climate time series required by several simulation models.

Once reference monthly values are obtained, daily T_{max} , T_{min} , Q , and P values are generated iteratively for each month. Then, if mean T_{max} , mean T_{min} , mean Q , or cumulative P is equal to reference monthly values for a given month $\pm 0.5^\circ\text{C}$, $1 \text{ MJ m}^{-2} \text{ month}^{-1}$, or 2 mm month^{-1} , depending on the variable, generated daily values are accepted. Conversely, if they are not within this range, generated daily values are rejected and RheaG iteratively generates another daily values

set for the given month, until generated values match the monthly specific reference monthly values.

To assess how climate data generated from a given validation location vary under different climate forcing conditions—that is, climate change—we generated climate data following three climate forcing scenarios for Barcelona ([Table 2](#)) for the 2010–2100 period, and for three climate forcing scenarios: no static present climate (SPC), moderate climate forcing according to the RCP4.5 climate forcing scenario, and strong climate forcing according to the RCP8.5 climate forcing scenario ([IPCC 2013](#)). We kept all monthly anomalies for both temperature and precipitation equal, to make the results comparison more straightforward. We then compared the original daily, monthly and annual climate data probability distributions, with the newly generated climate time series under both RCP4.5 and RCP8.5 climate forcing scenarios.

All data processing and statistical analyses have been performed using R software ([R Core Team 2017](#)).

3. Results

a. Validation of generated time series statistical properties

Madrid presented the lowest mean annual precipitation from the four validation locations climate data time series analyzed, with values around 370 mm yr^{-1} . Madrid also presented the widest range of temperature values, with 13°C between mean minimum and maximum annual temperatures ([Table 3](#)). Conversely, Barcelona

TABLE 3. Comparison for all four validation locations for observed (21 years) and generated with RheaG (100 years) annual mean precipitation (P ; mm yr^{-1}), annual mean maximum temperature (T_{max} ; $^\circ\text{C}$), and annual mean minimum temperature (T_{min} ; $^\circ\text{C}$). Values are noted as mean plus and minus standard error of five generated climate time series.

Location	Observed			Generated		
	$P (\text{mm yr}^{-1})$	$T_{\text{max}} (^\circ\text{C})$	$T_{\text{min}} (^\circ\text{C})$	$P (\text{mm yr}^{-1})$	$T_{\text{max}} (^\circ\text{C})$	$T_{\text{min}} (^\circ\text{C})$
Bilbao	1100	19.6	10	1090 ± 9.3	19.7 ± 0.05	9.8 ± 0.01
Barcelona	603	19.9	12	607 ± 8.6	19.7 ± 0.01	11.9 ± 0.05
Madrid	373	21.3	8.1	375 ± 2.7	21.4 ± 0.02	8 ± 0.01
Sevilla	528	25.7	13.3	536 ± 6.2	25.7 ± 0.01	13.3 ± 0.01

FIG. 1. Comparison between monthly rain probability for observed (solid line) and mean (dots) and standard deviation (bars: ± 1 std dev) of the five simulated climate time series for each of the four validation locations. Spearman's correlation coefficient is noted for each location.

presented the narrowest range, with only 8°C between minimum and maximum mean annual temperatures. Bilbao presented the highest annual precipitation values, around 1100 mm yr^{-1} for the 1989–2010 period. Annual mean maximum and minimum temperatures were the highest in Sevilla, with mean values of 13° and 26°C , respectively.

Differences between observed and generated climate time series in mean annual P , $T_{\min}$, and $T_{\max}$ values were minimal for all four locations. They diverged less than 1% for all three variables considered, except for mean annual precipitation in Sevilla, which diverged about 1.5%, with slightly higher generated precipitation values than observed ones (Table 3).

Monthly daily rain probabilities between observed and generated climate time series were almost identical (Fig. 1, $0.96 < R^2 < 0.99$). However, in Sevilla, generated monthly daily rain probabilities diverged more than the observed ones (RMSE = 0.025, $n = 60$). Madrid (RMSE = 0.01, $n = 60$) was the validation location where generated monthly daily rain probabilities were

most similar to the observed ones, followed by Barcelona (RMSE = 0.012, $n = 60$), and Bilbao (RMSE = 0.016, $n = 60$).

MM transition probabilities of generated climate time series were also consistent with observed properties for all locations (Fig. 2; $0.91 < R^2 < 0.995$). RMSE of the probability of observing a nonrainy day after a nonrainy day $p(\text{DD})$ ranged from 0.007 in Madrid to 0.031 in Barcelona. It is noteworthy that the relatively high divergence in Barcelona is attributable to an anomalous February $p(\text{DD})$ in one out of five of the meteorological time series generated, which can occur when Markov chains are used to randomly generate a finite number of iterations. Additionally, the RMSE among the observed and generated probabilities of observing a rainy day after a rainy day $p(\text{WW})$ reached its lowest values in Bilbao (0.015) and its highest values in Sevilla (0.048).

Newly generated climate time series maintained the observed negative linear correlation between $p(\text{DD})$ and daily rain probability (Fig. 3). They maintained both the overall trend and the climate time series' specific

FIG. 2. Comparison between observed monthly Markovian dry-dry [$p(DD)$] and wet-wet [$p(WW)$] transition probabilities (solid line and dotted line, respectively) and mean (triangle) and standard deviation (bars; ± 1 std dev) of the five R time series for each of the four validation locations. Spearman's correlation coefficient is noted for each location and for both $p(WW)$ and $p(DD)$.

trend for all four sites. The analysis of Bilbao meteorological time series presented two different season-related patterns of this relationship (Fig. 4): the first one comprises the cold and wet months (from November to May), and the latter one includes the warm months (from June to October). This indicates that, given the same daily rain probability, during warm months Bilbao presents a lower $p(DD)$ than during the wet ones within the ranges of daily rain probability registered in the reference meteorological time series. Hence, an increase in the dispersion of the rain events during the warm season is observed in Bilbao.

b. Comparison between observed and generated rain spell length

For all four validation locations, simulated rain spell length probability distribution matched the observed rain spell length (Fig. 5). In all cases, 1-day-long spell

lengths were found the most probable case. A nonlinear decrease in occurrence probability was observed at increasing spell lengths. One-day-long spell length probabilities ranged from 0.35 in Bilbao to 0.55 in Barcelona, and Madrid and Sevilla being the middle cases, with values ranging from 0.45 to 0.5, respectively. Furthermore, the steeper decreases in spell length frequency were observed in the Mediterranean-type climates, Sevilla and Barcelona.

Spearman correlation scores between observed and simulated rain spell lengths were higher than 0.99 in all cases, with differences ranging from less than 2% in Bilbao to a maximum of 6% in 1-day-long rain spell length in Sevilla. No trend was observed in the differences among observed and generated climate time series. Nevertheless, RheaG tends to better reproduce low-frequency rain spell lengths in all locations and slightly fails to reproduce the observed frequency of

FIG. 3. Correlation between daily rain probability and dry-to-dry probability transition. (left) Observed climate data values. (right) Five generated climate data values. Linear regression coefficients and Spearman's correlation coefficient are noted both for the overall values and for the separate four locations (inner tables). The overall linear regression trend is noted as a solid line. The Bilbao (Δ) trend is noted as a dotted line, the Barcelona (+) trend is noted as a dashed line, the Madrid ($\times$) trend is noted as an $\times$ -based line, and the Sevilla ($\blacksquare$) trend is noted as a small-dashed line. For each coefficient, the significance level is noted: ns = not significant; * = $p < 0.05$; *** $p < 0.01$. See Fig. 4 for the reason for the low correlation found in Bilbao.

short rain spell lengths, with no apparent trend among locations.

c. Weather generation under climate forcing

Interannual variation in cumulated precipitation under SPC climate forcing (Fig. 6) better represented observed interannual variability in precipitation in Barcelona ($602 \pm 144 \text{ mm yr}^{-1}$) by using the unconstrained procedure ($620 \pm 133 \text{ mm yr}^{-1}$) than the constrained one ($614 \pm 72 \text{ mm yr}^{-1}$). The same pattern was also observed in mean annual maximum temperature: unconstrained values ($19.4^\circ \pm 0.22^\circ\text{C}$) were more variable than constrained ones ($19.5^\circ \pm 0.2^\circ\text{C}$), thus better reproducing the observed pattern ($19.4^\circ \pm 0.54^\circ\text{C}$).

Following the unconstrained procedure, higher variability of monthly precipitation was observed in Barcelona on a monthly basis (Fig. 7). In addition, using the unconstrained procedure, the probability distribution of monthly anomaly of precipitation $[(P - \text{PMean})/\text{PMean}]$ was more in accordance with observed values than using the constrained one. There were peaks in autumn precipitation that reached $\sim 320 \text{ mm month}^{-1}$ for all three climate forcing scenarios in the unconstrained procedure, while maximum monthly precipitation values following the constrained procedure were lower, about $150 \text{ mm month}^{-1}$ during autumn. However, as climate change intensity increased, the high precipitation months became less frequent for the two procedures (Fig. 8). On the other hand, the constrained procedure generated less variability in monthly precipitation values, with

both maximum and minimum monthly precipitation values constrained within predefined ranges. In addition, under SPC conditions, daily probability of no rain was 0.844 and 0.827 for the unconstrained and the constrained

FIG. 4. Bilbao's correlation between daily rain observed probabilities and dry-to-dry Markov probability transition. Two different seasons have been differentiated: the mild season (November–May; solid line) and the warm season (June–October; dashed line). Numbers represent the month of the year. Big numbers are observed values; small numbers are generated values. Meteorological time series are analyzed for each season. Table indicates a and b regression coefficients for all four conditions: obs = observed and gen = generated. For each regression, the Spearman's correlation coefficient is also noted. For all regressions and regression coefficients, $p < 0.001$.

FIG. 5. Comparison between spell length probability in days for observed probability (dashed line) and mean (dots) and standard deviation (± 1 std dev; bars) of the five simulated climate time series for each of the four validation locations. Spearman's correlation coefficient is noted in each case. Inner plot indicates the difference of probability in spell length, calculated as the observed climate series spell length probability minus the mean simulated climate series spell length probability.

procedures, respectively, while it was 0.846 in the observed data. In the RCP4.5 scenario, daily probability of no rain increased to 0.844 and 0.841 for the unconstrained and the constrained procedures, respectively. Finally, in the extreme RCP8.5 scenario, both procedures presented similar values of 0.844.

Daily Tmax probability distribution matched almost completely observed daily Tmax distribution for both the constrained and the unconstrained procedures (Fig. 9). Median observed daily Tmax values were 18.8°C. In a SPC scenario, observed median daily Tmax values were 19.1° and 18.8°C following the unconstrained and the constrained procedures, respectively. They increased to 20.4° and 20.6°C, respectively, under an RCP4.5 forcing scenario, and finally reached 22.4° and 22.9°C, respectively, under the extreme RCP8.5 climate forcing scenario.

4. Discussion and conclusions

We have described and tested RheaG WGA in four contrasting climates from Iberian Peninsula. RheaG

accurately reproduces the statistical properties of the reference time series and is able to generate self-consistent series of meteorological variables following preestablished climate forcing scenarios. Given the intention behind the RheaG model, which is to generate meteorological time series to feed forest simulation models that run on a daily basis, our results have shown that both the simulated precipitation trends and the temperature probability distributions matched the observed values in the field when no further climate forcing—that is, SPC scenario—is considered.

In addition, RheaG WGA is able to generate meteorological time series under climate forcing. Meteorological time series generated under climate forcing take into account both the observed climate data statistical properties and the projected changes in those statistical properties—that is, anomalies in mean monthly temperature and precipitation. When new climate change projections will become available, RheaG outputs can be easily updated accordingly.

It is noteworthy the overall good agreement of generated time series to the observed monthly $p(DD)$ and

FIG. 6. (top) Evolution of annual cumulated precipitation (mm yr^{-1}) and (bottom) annual mean maximum temperature ($^{\circ}\text{C}$) for Barcelona, and for the two procedures of weather generation implemented in RheaG: (left) the constrained one and (right) the unconstrained one. Different climate forcing scenarios are noted as follows in each panel: SPC, solid black line; RCP4.5, solid gray line; and RCP8.5, broken gray line. Dotted lines indicate reference time series mean values: mean annual precipitation = 602 mm yr^{-1} , annual mean maximum temperature = 19.4°C .

$p(\text{WW})$ values in all four locations, with correlations ranging between $0.97 \leq R^2 \leq 0.99$ (Fig. 2). Moreover, this good agreement in MM transition probabilities for all locations also derive in highly correlated monthly $p(\text{Rain})$ in all locations [$0.96 \leq R^2 \leq 0.99$ (Fig. 1)]. Interestingly, in the case of Bilbao, RheaG is even able to reproduce observed bimodal seasonal trend (Fig. 4) that differentiates two climate seasons within Bilbao meteorological time series. These two seasons, the warm—from June to October—and the mild—from November to May—affect the relationship between the month $p(\text{Rain})$ and the month $p(\text{DD})$. During the coldest months, there is a higher $p(\text{DD})$ for the same $p(\text{Rain})$. As $p(\text{DD})$ and $p(\text{WW})$ are negatively correlated (Geng 1986), the trend observed in Bilbao indicates a higher rain concentration in fewer rain events during the mild period than during warm period, in which rain spells are shorter and more

frequent. This different seasonal behavior in Bilbao leads us to emphasize the importance of monthly characterization of meteorological time series, rather than just working with annual averages.

Rain spell length directly depends on the Markov transition probability matrix (Sonnadara and Jayewardene 2015), but it is not a direct input for the RheaG WGA. Therefore, it can be used as an indicator of suitability for the generated meteorological time series. Our results show that observed and generated meteorological time series present a very close pattern of rain spell length probabilities. Most of the inconsistencies occur at 1-day spell length probabilities, with the Sevilla climate having lower agreement and the Bilbao climate having greater agreement. On the other hand, low-probability events are well represented in all locations.

FIG. 7. Monthly anomalies in precipitation $[(P - P_{\text{mean}})/P_{\text{mean}}]$ for Barcelona and for three meteorological time series generated following the (a) unconstrained and (b) constrained procedures ($n = 200$). Different independent time series are noted with different symbols in each panel ($+$, Δ , $\circ$). Inner tables indicate the Spearman's correlation coefficient between monthly values for all three meteorological time series analyzed. Interestingly, Spearman's correlation in the unconstrained procedure is close to zero, while in the constrained one it is close to 1. (c) The three unconstrained (black dashed lines) and the three constrained (gray dotted lines) meteorological time series monthly anomalies in precipitation density distribution compared to the observed density distribution (black solid line).

Traditionally, temporal autocorrelations in precipitation generation have been implemented in different WGA using a broad range of algorithms: the simplest ones, such as the used in RheaG WGA, are based on first-order Markov probabilities. Including second-order Markov probabilities has been found to better describe precipitation patterns and to minimize the inconsistencies in 1-day rain spell lengths observed when using only first-order Markovian matrices (e.g., Buishand 1978; Chen et al. 2012), at the expenses of higher input parameter number. Recently, more complex rain-cell clustering approaches (Rodríguez-Iturbe et al. 1987), such as the Neyman–Scott rectangular pulses (NSRP; e.g., Kilsby et al. 2007; Burton et al. 2010) or the Bartlett–Lewis rectangular pulses (Onof et al. 2000) have also been developed. The latter models have proven to better reproduce the 1-day autocorrelation lag between rainy days than MM transition-based models (Kilsby et al. 2007). Hence, they better represent the 1-day lag between rain events than the ones reported by RheaG WGA (Fig. 5). On the other hand, no differences in mean monthly daily rainfall and monthly proportion of dry days were observed between MM-based and NSRP rain generation procedures (Kilsby et al. 2007). Additionally, NSRP models, compared with the most simple first-order MM-based models, are more expensive both in terms of parameter number and in the required length of the input observed meteorological time series (Ailliot et al. 2015). Therefore, in RheaG WGA we opted to maintain an input parameter number that was as low as possible in order to maximize parsimony, as well as to allow the WGA to run with shorter input observed climate time series.

Concerning to the two procedures of weather generation included in RheaG, the weather generation procedure that better reproduces both the annual and the monthly variability in precipitation of the observed time series is the unconstrained one. However, annual and monthly values for meteorological time series generated following this option are independent among sites and scenarios, as there is no spatial or temporal autocorrelation among time series generated. This leads to uncorrelated monthly and yearly values for each one, despite all of them maintaining equal statistical properties on the long run. Thus, we consider that the unconstrained procedure is the best and fastest option for independent one-site meteorological time series generation, as their meteorological time series are more accurate and realistic.

On the other hand, the generation of meteorological time series following the constrained procedure is more computationally costly—it takes 2 min to generate a 100-yr time series—than the unconstrained one, which takes about 30 s per time series. In addition, monthly

FIG. 8. Monthly generated precipitation (mm month^{-1}) for Barcelona and for the two procedures considered in RheaG for the three climate forcing scenarios considered: (top) SPC, (middle) RCP4.5, (bottom) RCP8.5. (left) The values for the constrained procedure; (right) the results for the unconstrained one. For each month, rain frequency distribution (shadowed area), mean monthly precipitation (black dots), and mean monthly precipitation in the observed climate time series (black dashed lines) are shown.

precipitation values variability is more restricted, thus obtaining monthly precipitation values only within a given range (Figs. 7 and 8). Thus, this method does not allow us to generate the low-frequency high monthly precipitation events registered. However, it maintains

the same background yearly climatic variability signal among time series. Keeping this background signal is important when there is spatial autocorrelation within generated meteorological time series, so monthly and yearly climate variability background has to be equalized

FIG. 9. Probability distribution of daily maximum temperatures ($T_{\max}$; $^{\circ}\text{C}$), for the two weather generation procedures included in RheaG [(left) constrained; (right) unconstrained] under the three climate forcing conditions considered and for Barcelona. The observed distribution is noted as a solid black line, SPC distribution is noted as a dashed black line, RCP4.5 climate forcing distribution is noted as a solid gray line, and RCP8.5 climate forcing distribution is noted as a dashed gray line.

to all generated meteorological time series. A further improvement of the constrained procedure, according to previous work on spatial autocorrelation of daily precipitation in the United Kingdom (Burton et al. 2013), would be the training and implementation of lag-0 cross-correlation-distance (XCD) double exponential algorithms for the mandatory variables (i.e., P , $T_{\max}$, $T_{\min}$, and Q), thus effectively describing the spatial autocorrelation of meteorological data for a given region based on the distance between meteorological stations. This opens an interesting field of research, as we hypothesize that this algorithm should be able to be trained at a monthly basis for Spain, if the amount of meteorological stations available for this training is large enough. If so, the capacity of RheaG WGA to generate spatially autocorrelated monthly reference time series at stations where its monthly values are unknown will be substantially increased.

Hence, the two RheaG weather generation procedures present an interesting trade-off between keeping the background climate signal similar in a multiplot generation with the constrained procedure, and the higher consistency with the observed annual precipitation variability obtained with the unconstrained one. Deciding to use the constrained or the unconstrained version of the RheaG WGA has to be carefully considered before each RheaG application—that is, we recommend applying the unconstrained procedure for generating climate time series for individual plots. When considering multiple plots where climate is spatially autocorrelated, the constrained procedure could be better as all meteorological time series generated maintain the same baseline monthly and

annual climate variability. Thus, the procedure selection must be made wisely in accordance with the spatial scale to which RheaG WGA is intended to work.

Acknowledgments. This work was funded by MEDFORESTREAM (CGL2011-30590-C02-01) project and MEDSOUL (CGL2014-59977-C3-1-R) project. Meteorological data were provided by Spanish Meteorological Agency AEMET in the frame of the MONTES-CONSOLIDER (CSD2008-00040) project. DN-S is supported by a FPI grant of the Spanish Ministerio de Economía y Competitividad (BES-2015-072983). DNS, SS, and CG are members of the research group FORESTREAM (AGAUR, Catalonia 2017SGR976). We the authors want to acknowledge the insightful comments provided by two anonymous reviewers. Specific author contributions: DN-S, CG, and SS designed the research; DN-S wrote the main text; SS and CG provided comments to the text; CG, DN-S, and SS formulated the RheaG algorithm; DN-S coded the RheaG weather generator algorithm in R language.

REFERENCES

- Ailliot, P., D. Allard, V. Monbet, and P. Naveau, 2015: Stochastic weather generators: An overview of weather type models. *J. Soc. Fr. Stat.*, **156**, 101–113.
- Allen, R. G., 1995: Evaluation of procedures for estimating mean monthly solar radiation from air temperature. Food and Agricultural Organization.
- , L. S. Pereira, D. Raes, and M. Smith, 1998: Crop evapotranspiration: Guidelines for computing crop water requirements. FAO Irrigation and Drainage Paper 56, 290 pp.

- Buishand, T. A., 1978: Some remarks on the use of daily rainfall models. *J. Hydrol.*, **36**, 295–308, [https://doi.org/10.1016/0022-1694\(78\)90150-6](https://doi.org/10.1016/0022-1694(78)90150-6).
- Bürger, G., 1997: On the disaggregation of climatological means and anomalies. *Climate Res.*, **8**, 183–194, <https://doi.org/10.3354/cr008183>.
- Burton, A., H. J. Fowler, S. Blenkinsop, C. G. Kilsby, 2010: Downscaling transient climate change using a Neyman–Scott rectangular pulses stochastic rainfall model. *J. Hydrol.*, **381**, 18–32, <https://doi.org/10.1016/j.jhydrol.2009.10.031>.
- , V. Glenis, M. R. Jones, and C. G. Kilsby, 2013: Models of daily rainfall cross-correlation for the United Kingdom. *Environ. Modell. Softw.*, **49**, 22–33, <https://doi.org/10.1016/j.envsoft.2013.06.001>.
- Caskey, J., 1963: A Markov chain model for the probability of precipitation occurrence in intervals of various length. *Mon. Wea. Rev.*, **91**, 298–301, [https://doi.org/10.1175/1520-0493\(1963\)091<0298:AMCMFT>2.3.CO;2](https://doi.org/10.1175/1520-0493(1963)091<0298:AMCMFT>2.3.CO;2).
- Chen, J., F. Brissette, and R. Leconte, 2012: WeaGETS—A Matlab-based daily scale weather generator for generating precipitation and temperature. *Procedia Environ. Sci.*, **13**, 2222–2235, <https://doi.org/10.1016/j.proenv.2012.01.211>.
- Duffie, J. A., and W. A. Beckman, 1991: *Solar Engineering of Thermal Processes*. Wiley, 944 pp.
- Fowler, H. J., and C. G. Kilsby, 2007: Using regional climate model data to simulate historical and future river flows in northwest England. *Climatic Change*, **80**, 337–367, <https://doi.org/10.1007/s10584-006-9117-3>.
- Gabriel, K. R., and J. Neumann, 1962: A Markov chain model for daily rainfall occurrence at Tel Aviv. *Quart. J. Roy. Meteor. Soc.*, **88**, 90–95, <https://doi.org/10.1002/qj.49708837511>.
- Geng, S., 1986: A simple method for generating daily rainfall data. *Agric. For. Meteorol.*, **36**, 363–376, [https://doi.org/10.1016/0168-1923\(86\)90014-6](https://doi.org/10.1016/0168-1923(86)90014-6).
- Hargreaves, G. H., and Z. A. Samani, 1982: Estimating potential evapotranspiration. *J. Irrig. Drain. Div.*, **108**, 225–230.
- Hawkins, D. M., 2004: The problem of overfitting. *J. Chem. Inf. Comput. Sci.*, **44**, 1–12, <https://doi.org/10.1021/ci0342472>.
- Hutchinson, M. F., 1986: Methods of generation of weather sequences. *Agricultural Environments*, A. H. Bunting, Ed., C.A.B. International, 149–157.
- IPCC, 2013: *Climate Change 2013: The Physical Science Basis*. Cambridge University Press, 1535 pp., <https://doi.org/10.1017/CBO9781107415324>.
- Katz, R. W., 1981: On some criteria for estimating the order of a Markov chain. *Technometrics*, **23**, 243–249, <https://doi.org/10.2307/1267787>.
- Kilsby, C. G., and Coauthors, 2007: A daily weather generator for use in climate change studies. *Environ. Modell. Software*, **22**, 1705–1719, <https://doi.org/10.1016/j.envsoft.2007.02.005>.
- Kimball, J. S., S. W. Running, and R. Nemani, 1997: An improved method for estimating surface humidity from daily minimum temperature. *Agric. For. Meteorol.*, **85**, 87–98, [https://doi.org/10.1016/S0168-1923\(96\)02366-0](https://doi.org/10.1016/S0168-1923(96)02366-0).
- Mearns, L. O., R. W. Katz, and S. H. Schneider, 1984: Extreme high-temperature events: Changes in their probabilities with changes in mean temperature. *J. Climate Appl. Meteorol.*, **23**, 1601–1613, [https://doi.org/10.1175/1520-0450\(1984\)023<1601:EHTECI>2.0.CO;2](https://doi.org/10.1175/1520-0450(1984)023<1601:EHTECI>2.0.CO;2).
- Onof, C., R. E. Chandler, A. Kakou, P. Northrop, H. S. Wheeler, and V. Isham, 2000: Rainfall modeling using Poisson-cluster processes: A review of developments. *Stochastic Environ. Res. Risk Assess.*, **14**, 384–411, <https://doi.org/10.1007/s004770000043>.
- R Core Team, 2017: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, <https://www.R-project.org/>.
- Richardson, C. W., 1981: Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resour. Res.*, **17**, 182–190, <https://doi.org/10.1029/WR017i001p00182>.
- , and D. A. Wright, 1984: WGEN: A model for generating daily weather variables. U.S. Department of Agriculture Agricultural Research Service ARS-8, 83 pp.
- Rodríguez-Iturbe, I., D. R. Cox, and V. Isham, 1987: Some models for rainfall based on stochastic point processes. *Proc. Roy. Soc. London*, **410A**, 269–288, <https://doi.org/10.1098/rspa.1987.0039>.
- Semenov, M., and R. Brooks, 1999: Spatial interpolation of the LARS-WG stochastic weather generator in Great Britain. *Climate Res.*, **11**, 137–148, <https://doi.org/10.3354/cr011137>.
- Sonnadara, D. U., and R. D. Jayewardene, 2015: A Markov chain probability model to describe wet and dry patterns of weather at Colombo. *Theor. Appl. Climatol.*, **119**, 333–340, <https://doi.org/10.1007/s00704-014-1117-z>.
- Spencer, J. W., 1971: Fourier series representation of the position of the sun. *Search*, **2**, 172.
- Stern, R. D., 1980: The calculation of probability distributions for models of daily precipitation. *Arch. Meteor. Geophys. Bioklimatol.*, **28**, 137–147, <https://doi.org/10.1007/BF02243841>.
- Wilks, D. S., and R. L. Wilby, 1999: The weather generation game: A review of stochastic weather models. *Prog. Phys. Geogr.*, **23**, 329–357, <https://doi.org/10.1177/030913339902300302>.
- Wood, A. W., L. R. Leung, V. Sridhar, and D. P. Lettenmaier, 2004: Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. *Climatic Change*, **62**, 189–216, <https://doi.org/10.1023/B:CLIM.0000013685.99609.9e>.