

Multisite, multivariate weather generation based on generalised linear models

Richard E. Chandler

Department of Statistical Science, University College London, Gower Street, London, WC1E 6BT, UK

ARTICLE INFO

Keywords:

Climate impacts
Downscaling
Weather generator

ABSTRACT

This paper describes a methodology for constructing and simulating from models of daily weather time series at multiple locations, incorporating potential nonstationarities and suitable for use in those studies of climate impacts and adaptation where a detailed representation of local weather is required. The approach is based on generalised linear models (GLMs) and aims to allow for realistic representations of local weather structures including spatial, temporal and inter-variable dependencies. The theory is implemented in a software tool, Rglimclim, that runs in the R programming environment; and is illustrated using a case study involving generation of daily precipitation and temperature at 26 locations in northern Iberia.

1. Introduction

The 13th Sustainable Development Goal of the United Nations is ‘take urgent action to combat climate change and its impacts’ (United Nations, 2015). Most studies of future climate impacts and adaptation are based on projections obtained by running global climate models (GCMs) under specified scenarios of socio-economic development and greenhouse gas emissions (Collins et al., 2012). However, despite ongoing improvements in GCM spatial resolution, their outputs remain too coarse to represent the local-scale weather features that control some phenomena. Examples include, but are not limited to, urban flooding which may be sensitive to the timing and intensity of precipitation on spatial scales of a few hundreds of metres (e.g. Bruni et al., 2015; Gires et al., 2012); agricultural applications in which fine-scale weather inputs may be required to determine crop potential in regions with high spatial heterogeneity (Zhao et al., 2015); and biodiversity assessments in which individual species may occupy niches that are tightly defined in terms of habitat and climate (Lembrechts et al., 2019; Franklin et al., 2013). It is therefore common to downscale GCM outputs to a finer spatial resolution, to represent more realistically the weather variability that is relevant to the system of interest.

Many downscaling methodologies are available, varying in complexity: Maraun et al. (2010) give a relatively up-to-date review, with a more recent and detailed treatment in Maraun and Widmann (2018). A thorough evaluation of different aspects of downscaling performance, for a variety of methods, is reported by Gutiérrez et al. (2019), Maraun et al. (2019a, b) and Widmann et al. (2019). In situations

requiring a detailed representation of the system of interest however, it is usually necessary to consider some of the more complex approaches: when designing flood defences, for example, hydrological response can be sensitive to both the quantity and timing of precipitation in particular (e.g. Wheater, 2002) and it is also necessary to consider the inter-relationships between relevant weather variables so as to obtain realistic representations of evapotranspiration. Apart from methods that seek historical analogs of a given configuration of GCM outputs (and are therefore restricted in their application to settings where the large-scale structure is similar to situations that have occurred in the past), the only downscaling methodologies that meet the requirements of such demanding applications are those based on regional climate models (RCMs) and on weather generators. RCMs are high-resolution climate models that are run for specified regions, usually at a continental scale, and with boundary conditions taken from GCM simulations: modern RCMs can routinely produce output on spatial scales of around $10 \times 10 \text{ km}^2$, although their effective resolution is probably coarser than this (Maraun et al., 2010). Weather generators, on the other hand, are stochastic models or algorithms that exploit statistical relationships between large-scale climate and local-scale weather variables: future weather sequences can be generated by simulating from these models conditioned on GCM outputs, in such a way as to preserve the dependencies among the weather variables of interest.

In many respects, RCMs and weather generators are complementary tools. Weather generators are relatively cheap to develop and to simulate, and can thus be tailored specifically to the application of interest: this ensures their accessibility to those without access to

E-mail address: r.chandler@ucl.ac.uk.

<https://doi.org/10.1016/j.envsoft.2020.104867>

Accepted 3 September 2020

Available online 11 September 2020

1364-8152/© 2020 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

supercomputing facilities, and also that multiple simulations can be produced quickly to explore sensitivity to different sets of assumptions. On the other hand, a potential disadvantage is the implicit assumption that the embedded statistical relationships will persist into the future. To ensure the credibility of weather generators developed with climate change applications in mind therefore, a minimum requirement is that the statistical relationships must reflect, as far as possible, the known large-scale physical controls on weather in the region of interest.

Various heuristic schemes have been presented for capturing signals of climate change in a weather generator, for example via the use of additive or multiplicative ‘change factors’ derived from GCM outputs and applied to historical simulations (e.g. Kilsby et al., 2007). Such schemes are not appropriate for use in situations requiring a detailed representation of local weather however: for example, the application of multiplicative change factors to a historical precipitation simulation will not change the timings of wet and dry periods, and hence is not appropriate for use in situations where the timing of precipitation is important. A more principled approach is to incorporate indices of large-scale atmospheric structure directly into the weather generator specification. In this case however, the chosen indices must both capture the climate change signal and be well represented by GCMs. For more discussion of this and other considerations in the use of statistical downscaling, see Wilby et al. (2004) and Maraun and Widmann (2018, Section 11.5).

Many modern weather generators aim to produce weather sequences at a daily time scale, and are descended in some way from the WGEN model implementing the framework of Richardson (1981) in which precipitation occurrence is represented as a first-order Markov chain, wet-day precipitation intensities using independent exponential or gamma distributions, and other variables by fitting appropriate distributions separately for wet and dry days. Seasonality is represented by estimating separate sets of model parameters for different periods of the year. This type of generator has been studied extensively and, in its original form, has several known deficiencies. For example, it often fails to reproduce adequately the lengths of dry and wet spells which can be important in many applications (e.g. Semenov et al., 1998); it tends to underestimate the variability of seasonal means or totals (a phenomenon referred to as ‘overdispersion’ by Katz and Parlange, 1998); and it has a tendency to underestimate extreme events such as 100-year return levels (e.g. Katz et al., 2002). To address these deficiencies, many extensions to the basic structure have been proposed: for example, via the use of higher-order Markov chains to improve wet and dry spell performance (Wilks, 1999); the use of heavy-tailed distributions for precipitation intensity to improve extremal behaviour (Furrer and Katz, 2008); and the introduction of latent weather states with separate parameter sets, to increase variability in seasonal means (Katz and Zheng, 1999). More generally, other classes of weather generator have been proposed such as those based on direct modelling of wet and dry spell lengths (e.g. Semenov et al., 1998) and resampling methods (e.g. Buishand and Brandsma, 2001).

An alternative way to remedy the deficiencies of the original WGEN-type generator is to embed its structure within a more flexible class of models. One such class is that of generalised linear models (GLMs), which form a cornerstone of modern statistical practice. GLMs were first applied to the modelling of daily precipitation sequences by Coe and Stern (1982), and subsequently shown by Grunwald and Jones (2000) to include Markov-based structures as special cases. Moreover, GLMs provide a simple way to relax some of the independence assumptions in the basic WGEN structure that are arguably responsible for some of its deficiencies. For example, correlation between successive days’ precipitation intensities can be modelled easily: such correlation is to be expected, at least in regions and seasons where precipitation is predominantly associated with large-scale weather systems lasting several days. Moreover, seasonal totals from positively correlated daily series are more variable than totals from independent series (this follows from the standard formula for the variance of a sum: $\text{Var}(\sum_{i=1}^n Y_i) =$

$\sum_{i=1}^n \text{Var}(Y_i) + 2\sum_{i=1}^{n-1} \sum_{j=i+1}^n \text{Cov}(Y_i, Y_j)$ — the final term in this expression is zero for independent intensities). Thus the phenomenon of overdispersion is unsurprising in the context of a weather generator that allocates precipitation intensities independently. Similarly, it can be demonstrated that correlated intensity sequences can produce heavy-tailed distributions of extremes (Yang et al., 2005), thus addressing another deficiency of the basic generator. GLMs therefore provide an opportunity to improve performance by addressing the causes, rather than symptoms, of problems.

The earliest weather generators were designed to generate sequences at individual locations. For some applications however, spatially coherent sequences are needed at multiple sites: this is the case for hydrological modelling of large or rapidly responding catchments, for example (Segond et al., 2007). Even in situations where multisite modelling is not strictly needed, it can lead to increased precision by exploiting the larger volumes of data that are available from a collection of stations. However, a key challenge when developing multisite generators is to produce realistic levels of dependence both between sites and between variables. Some approaches to multi-site generation include simple extensions of WGEN-type generators along the lines of Wilks (1998) and, more recently, Dubrovsky et al. (2020); resampling methods (e.g. Beersma and Buishand, 2003); approaches based on transformed Gaussian variables (e.g. Stehlik and Bardsossy, 2002); and those based on Hidden Markov models (e.g. Ailliot et al., 2009) that postulate the existence of a sequence of unobserved ‘weather states’ on each day, each associated with characteristic spatial patterns. Multi-site GLMs have also been developed (Yang et al., 2005; Ambrosino et al., 2014). However, most GLM-based weather generators developed to date have been univariate: a notable exception is Asong et al. (2016), who used an earlier version of the software described herein.

In addition to the ‘scientific’ requirements of weather generators outlined above, there are often more mundane but equally important considerations relating to data availability and quality. Some multisite approaches become computationally infeasible if the number of sites grows to more than a handful; others require complete data (i.e. with no missing values) for calibration, which is rarely realistic. Moreover, simulated sequences are often required at ungauged locations — for example, when studying the hydrological response of a catchment when the nearest precipitation stations lie outside the catchment, or when a land surface model requires weather inputs on a regular spatial grid.

In summary, for applications where a detailed representation of weather inputs is needed, a weather generator must be capable of: producing sequences with realistic spatial, temporal and inter-variable dependence structures over a range of scales; capturing realistic levels of unstructured variability associated with phenomena such as extremes; representing systematic variation over space and time, including under scenarios of climate change; generating sequences at locations for which no data are available; and coping with missing values in the historical records used for model calibration. Of course, the relative importance of these requirements will be context-specific. Of the techniques reviewed above, GLMs are one of the few that can be made to meet all of them.

Against this background, the contribution of the present paper is to demonstrate how previous work on univariate GLM-based weather generators can be extended to the multivariate setting; and to present a software environment, Rglimclim (Chandler, 2018), implementing this framework. At present, there are few other multisite, multivariate weather generators for which software is publicly and freely available: to the author’s knowledge, the only exceptions are MuGETS (Chen et al., 2014) and RMAWGEN (Cordano and Eccel, 2017), both of which are designed to generate sequences of precipitation together with maximum and minimum temperature. The case study in the present paper involves generation of precipitation and mean temperature: there is no publicly-available software that can be used to provide a direct comparison with the results presented here, therefore. However, many

previous studies have evaluated the GLM framework in a wide range of scenarios with favourable results (e.g. Kenabatho et al., 2012; Mockler et al., 2016; Asong et al., 2016; Chun et al., 2017; Tosonoğlu and Onof, 2017); and others have carried out comparisons in which GLMs performed competitively alongside other leading approaches (e.g. Ayar et al., 2016; Frost et al., 2011).

A review of the GLM methodology is presented next: this covers model formulation as well as parameter estimation, model comparison, diagnostics and simulation. Section 3 then presents a case study to illustrate the ideas; the software design and philosophy is described in Section 4. Some open questions and practical issues are discussed in Section 5.

2. Weather generation using generalised linear models

This methodological review is necessarily brief. The technical details are described fully in the cited references, and in the Rglimclim package manual.

2.1. Univariate weather generation

We first consider a daily multisite weather generator for a single variable. Let Y_{st} denote the variable of interest — referred to below as the *response variable* — at site s on day t . Then, in the GLM framework, the $\{Y_{st}\}$ are considered to be drawn from a common class of distributions which, for technical reasons, is restricted to those falling in the *exponential family* (McCullagh and Nelder, 1989). This family includes many common distributions including the normal, gamma, Poisson and Bernoulli. The distributions of the individual $\{Y_{st}\}$ can vary between sites and days, with the variation determined by row vectors $\{\mathbf{x}_{st}\}$ of covariates (see below) that determine the expected values of the $\{Y_{st}\}$. Formally, a GLM relates the conditional expectation $\mathbb{E}(Y_{st}|\mathbf{x}_{st}) = \mu_{st}$ (say) to a linear function of the covariate values:

$$g(\mu_{st}) = \mathbf{x}_{st}\boldsymbol{\beta} = \eta_{st} \text{ say,}$$

where $\boldsymbol{\beta}$ is a column vector of coefficients and $g(\cdot)$ is a monotonic *link function*. The quantities $\{\eta_{st}\}$ are known as *linear predictors* (not to be confused with the covariates themselves, which are often referred to as ‘predictors’ in the downscaling literature).

For some choices of distribution, an additional *dispersion parameter* is needed to specify the distributions of the $\{Y_{st}\}$ completely. For models based around normal distributions for example, the dispersion parameter is the variance; for those based around gamma distributions, it is the inverse of the shape parameter. Dispersion parameters are usually assumed constant in GLMs, so that they do not depend on covariates. This assumption can be relaxed however, as described below in the context of models based on normal distributions.

The most familiar example of a GLM is a linear regression model, in which the expectation of the response variable is a linear function of covariates and the distributions of the responses are normal with constant variance. In the framework outlined above, this corresponds to a normal GLM with an identity link function: $Y_{st} \sim N(\mu_{st}, \sigma^2)$ with $\mu_{st} = \mathbf{x}_{st}\boldsymbol{\beta} = \eta_{st}$ so that $\boldsymbol{\beta}$ is the vector of regression coefficients. The normality assumption may be approximately justified for some daily weather variables such as temperature, but not for variables with skewed distributions. The GLM framework provides an opportunity to apply ‘regression-like’ thinking to the modelling of such variables. For example, as noted by Coe and Stern (1982) and Stern and Coe (1984), daily precipitation sequences can be modelled using a combination of logistic regression (a GLM with a Bernoulli response distribution) to model precipitation occurrence, and a gamma GLM for non-zero precipitation intensities. Wind speed is another variable that typically has a highly skewed distribution at the daily time scale, and can be modelled using a gamma GLM (Yan et al., 2002).

In the context of multisite weather generation, the possible

covariates typically fall into a small number of categories: a constant term, along with covariates representing systematic variation in space (e.g. site altitude, and basis functions representing systematic regional variation as described in Chandler, 2005), temporal trends, seasonal variation (often represented using sine and cosine functions of the time index t), day-to-day dependence (represented by including functions of previous days’ responses as covariates in a GLM), and indices of large-scale climate. By including previous days’ (‘lagged’) responses as covariates, autocorrelation can be represented easily: this offers the potential to overcome some of the deficiencies of the standard WGEN-type generator, as discussed above.

A useful feature of GLMs, as with other regression models, is the ability to represent interactions between covariates: these occur when the effect of one covariate on the response is modulated by the value of another. For example, in temperate latitudes it is common for the strength of autocorrelation in daily weather sequences to vary seasonally: this variation can be represented in a GLM as an interaction between ‘seasonal’ and ‘autocorrelation’ covariates. Formally, an interaction is included by defining an additional covariate whose value is the product of the interacting covariates. For justification of this, see Chandler and Scott (2011, Section 3.2.2).

To develop a GLM it is necessary to choose an appropriate distributional family and link function, to choose an appropriate set of covariates and interactions, and to estimate the corresponding coefficient vector. The choice of distribution may be based on the nature of the response variable as in the examples given above, whereas the link function is often chosen to ensure that the modelled means $\{\mu_{st}\}$ satisfy any constraints arising from the context. For example, wind speeds cannot take negative values, so in a gamma GLM for wind speed it is natural to consider a log link and to write $\log \mu_{st} = \mathbf{x}_{st}\boldsymbol{\beta}$ so that $\mu_{st} = \exp(\mathbf{x}_{st}\boldsymbol{\beta})$ is itself guaranteed non-negative. In principle, a variety of possible link functions could be considered in any given application. In applied statistical practice however, standard ‘default’ choices are often made, as in the models developed in Section 3.2.1 below: alternatives can be considered if these standard choices lead to some systematic lack of model fit (see Dobson and Barnett, 2018, Section 7.3 for an example).

2.1.1. Model fitting and comparison

Having chosen a distributional family and link function, covariates can be chosen by fitting different models to a data set and comparing them. If the data values can all be considered as statistically independent conditional on the covariates, then the coefficient vector $\boldsymbol{\beta}$ can be estimated using maximum likelihood estimation which, for most GLMs, is carried out using an efficient iterative weighted least-squares algorithm (Davison, 2003, Section 10.2). The exception is the standard multiple linear regression model, for which there is no need for iteration: the usual least-squares procedure delivers maximum likelihood estimates directly in this case.

For models involving a dispersion parameter, this is estimated separately (McCullagh and Nelder, 1989) using the Pearson residuals (defined below). This ability to estimate the coefficients separately from the dispersion parameter is a key feature of the exponential family of distributions: the simplest example is the linear regression model, where the regression coefficients are estimated first and then the residuals are used to estimate the error variance.

Having estimated the model coefficients and, if necessary, dispersion parameter, standard large-sample theory then allows the construction of standard errors and confidence intervals for the coefficients, and formal model comparisons can be done using techniques such as likelihood ratio tests. Specifically, suppose that models \mathcal{M}_0 and \mathcal{M}_1 have been fitted to the same data using maximum likelihood, and that \mathcal{M}_0 can be obtained by fixing p of the coefficients in \mathcal{M}_1 to pre-specified values (often zero, so that the corresponding covariates have no effect on the response in \mathcal{M}_0). Suppose in addition that the data were generated from model \mathcal{M}_0 , so that \mathcal{M}_1 is overfitted. Denote by L_0 and L_1 the maximised likelihoods for the two models. Then, in large samples, the likelihood

ratio statistic

$$\Lambda = 2(\log L_1 - \log L_0) \quad (1)$$

has approximately a chi-squared distribution with p degrees of freedom. If the value of Λ exceeds some high percentile of this chi-squared distribution therefore, we can conclude that \mathcal{M}_1 offers a statistically significant improvement in model fit compared with \mathcal{M}_0 .

For a typical daily multisite dataset however, independence is usually an unrealistic assumption: dependence in both time (autocorrelation between successive days' values) and space (dependence between neighbouring sites) is to be expected. It can be shown (Chandler et al., 2006) that if a model contains an adequate representation of temporal autocorrelation by including functions of previous days' responses as covariates, then standard likelihood theory can be applied in the absence of inter-site dependence. Inter-site dependence is harder to deal with, however: there are few modelling options available that are simultaneously tractable, computationally undemanding and physically defensible (for a critique of modern statistical approaches to the problem, see Chandler and Ambrosino, 2015). An alternative approach is adopted here, therefore: models are fitted as though sites are independent, and empirical adjustments are made to the standard errors and likelihood ratios to account for the neglected dependence. It can be shown (Chandler et al., 2006; Chandler and Bate, 2007) that this approach delivers valid estimates and accurate inference: the estimates are less precise than would be obtained from a correctly specified fully spatiotemporal model, but with the large datasets that are usually available for calibrating weather generators (often tens or hundreds of thousands of observations) the precision is invariably adequate for all practical purposes.

It is worth noting that GLMs can be fitted to data sets with missing observations, although only the observations with complete cases (i.e. with non-missing values for the response variable and all covariates) will be used in the fitting. This allows the use of data from sites with short or patchy records when fitting models, which is helpful in data-sparse regions. A caveat, however, is that when comparing two models using likelihood ratio tests, the models must be fitted to the same set of observations: if model \mathcal{M}_1 contains covariates that do not appear in \mathcal{M}_0 and for which some values are missing therefore, the corresponding observations must also be omitted when fitting \mathcal{M}_0 . This situation arises most frequently when comparing models involving different numbers of lagged response terms: to compare models containing 1 and 2 previous days' values for example, both models must be fitted to the subset of observations for which 2 previous days' responses are available.

2.1.2. Diagnostics and model checking

In addition to formal model comparisons, various diagnostic measures can be used to assess the fit of a GLM. Many of these measures are graphical in nature and are based around model residuals, constructed in such a way as to have a homogeneous structure if the model is well specified. In the statistical literature it is common to use deviance residuals, which are constructed as a direct theoretical analogue of the residuals usually encountered in regression models (Davison, 2003, Section 10.2). These can be hard to communicate to non-specialists however, so the diagnostics in Rglimclim are mostly based on the Pearson residuals $\{r_{st}^{(P)}\}$ say, which are dimensionless quantities defined as

$$r_{st}^{(P)} \propto \frac{y_{st} - \mu_{st}}{\sigma_{st}}.$$

Here, y_{st} is the observed response at site s on day t , and μ_{st} and σ_{st} are respectively the mean and standard deviation of the generating distribution according to the fitted GLM. If the model captures the systematic structure in the data then the Pearson residuals should all have zero mean and a constant variance: in this case, plots of residuals against potential or actual covariates should reveal no obvious structure. In practice, it can be hard to see structure in plots with large numbers of

observations: Rglimclim therefore produces summary plots showing the mean and standard deviation of Pearson residuals in different subgroups of observations. For example, if seasonality is well captured by a model then a plot of mean residuals for each month of the year will show no systematic structure. To aid interpretation of these plots, 95% reference bands are added to show where the mean residuals should fall if the model is reasonable. The width of these bands is adjusted where necessary to account for inter-site dependence.

When modelling response variables with continuous distributions, a further use of residuals is to produce quantile-quantile plots for checking distributional assumptions: if the assumed family of distributions (e.g. normal or gamma) is reasonable, then the points on a quantile-quantile plot should fall close to a straight line, although some sampling variation is to be expected at the extreme ends of such a plot. Such plots are not useful, however, when the response variable is highly discrete: in the context of weather generation, this situation arises when modelling precipitation occurrence. In this case, other checks on the probability structure are needed. A common approach (Dawid, 1986) is based on the observation that among all occasions when the modelled probability of precipitation is p , precipitation should have occurred roughly 100% of the time. In practice, we collect together groups of days for which the modelled probabilities are in the intervals $(0.0, 0.1)$, $(0.1, 0.2)$, ..., $(0.9, 1.0)$ and compute observed and expected proportions of 'wet' days within each of these groups. Unless there is agreement across the whole range of probabilities, there is something wrong with the probability structure of the model.

2.1.3. Inter-site dependence

The discussion so far has focused on fitting, choosing and checking models, for which purpose it is not necessary to specify the inter-site dependence structure in detail as noted above. However, when generating synthetic weather sequences, a multi-site weather generator must be capable of producing realistic levels of dependence between neighbouring sites. This can be achieved by modelling separately the residual dependence that is not accounted for in the mean and variance structure of the GLM. For quantities with continuous distributions, it is often convenient to specify the dependence via the correlation structure of residual measures that are defined in such a way as to be approximately Gaussian distributed: these quantities are sometimes called *Anscombe residuals* (McCullagh and Nelder, 1989, Section 2.4). In this approach, which can be regarded as an application of copula-based methods (see for example Davison et al., 2012; Patton, 2012), a variety of standard spatial or geostatistical correlation models can be used (see Webster and Oliver, 2001 for a review of such models). For highly discrete quantities however, a correlation-based approach may be less appealing and alternatives may be preferable. Spatial dependence in precipitation occurrence is particularly problematic: in Rglimclim, the options available include models based on the distribution of the total number of sites experiencing precipitation on a given day (Yang et al., 2005), as well as models based on the correlation structure of latent Gaussian processes (Ambrosino et al., 2014). The former is designed for application to relatively small study areas, whereas the latter is more appropriate in larger areas where dependence between distant pairs of sites is weaker.

There are, of course, alternatives to the approaches outlined above. However, the choice of methodology in Rglimclim has been made throughout with incomplete data sets in mind. Specifically, a user who wishes to evaluate the performance of a weather generator will want to compare the properties of synthetic sequences with those that have been observed. If the observed record contains many missing observations, then its properties cannot be compared directly with those of a complete synthetic sequence. One possible approach to this is to remove the corresponding values from the synthetic sequence before making the comparison; a preferable alternative is to impute the missing values by sampling from their distributions conditional on all available observations. By carrying out multiple imputations, it is possible to characterise the uncertainty arising from missing observations in historical weather

properties. Obviously, this requires that the conditional distributions can be calculated: this cannot always be done efficiently, but it is possible for all of the dependence structures implemented in R*Glimclim*. The use of imputations is illustrated in the case study below.

2.2. Multivariate weather generation

In moving from univariate to multivariate weather generators, the main challenge is to preserve inter-variable dependencies — the strength of which may depend on other factors (as an obvious example, in temperate latitudes the correlation between daily temperature and daily precipitation is typically negative in summer and positive in winter, due to the effect of cloud cover on net radiation). A flexible strategy for modelling the joint distributions of the variables is therefore needed. Unfortunately, tractable models for joint distributions are scarce (see, however, [Fahrmeir and Tutz, 2001](#) for some options); and there are no standard models for situations where, as often occurs with weather generators, the marginal distributions of the individual variables are of different types.

In a multivariate setting, the collection of variables at site s and day t can be collected into a random vector $\mathbf{Y}_{st} = (Y_{1st} \dots Y_{Kst})^T$, say. If all components of \mathbf{Y}_{st} have continuous distributions then, as discussed above in the context of inter-site dependence, it is natural to consider copula-based approaches involving transformations to Gaussianity, so as to model the inter-variable dependencies via the covariance structure of the transformed variables. For binary or discrete variables however, such approaches are less appealing. One possibility is to treat such variables as discretised versions of latent Gaussian quantities and to consider the latent, rather than observed, variables as components of the joint distribution. Inter-relationships among assumed latent variables can be hard to identify, however. Within the GLM framework it is easier, more flexible and often more intuitive to build models for each variable in turn, at each stage conditioning on the variables that have already been modelled. This is possible because any joint probability density function can be factorised into a product of conditional densities: the joint density for \mathbf{Y}_{st} can thus be written as $f(\mathbf{y}_{st}) = f_1(y_{1st})f_2(y_{2st}|y_{1st}) \dots f_K(y_{Kst}|y_{1st}, \dots, y_{(K-1)st})$, the vertical bar ‘|’ denoting a conditional distribution. Each of these conditional densities can be modelled using its own GLM, by incorporating the conditioning variables into the respective covariate set. This approach is particularly useful when the individual components of \mathbf{Y}_{st} behave very differently, because the models for each variable can be constructed separately using whatever distributional assumptions are appropriate. The approach also simplifies the modelling task by breaking it down into smaller steps, each involving a single response variable.

When building a multivariate model using successive conditioning, an obvious question is how to order the variables. In almost all of the weather generator literature, precipitation is treated as the ‘primary’ variable with everything else conditioned on precipitation via, for example, regression relationships. However, the reasons for this are historical: in the early 1980s, when the first weather generators were developed, few if any tractable models were available for highly non-Gaussian quantities (such as precipitation) conditional on other variables. The increased flexibility of GLMs forces us to confront the question.

Mathematically the ordering of variables in a multivariate model is unimportant: the joint density always factorises. In practice however, the models for each variable are at best approximations to the corresponding conditional distributions, whence the quality of the overall model depends on the adequacy of the cumulative approximation. All other things being equal therefore, it seems sensible to order the variables so as to reflect any subject-matter understanding of the causal relationships between them: if one variable can be considered as a ‘driver’ of the others, then it is natural to consider this first and then to use it as a covariate in the models for the remaining variables. Unfortunately however, this approach is not always possible or optimal. This

is particularly true in situations where ‘primary’ variables have substantial amounts of missing data, because this potentially limits the number of cases available for modelling other variables as well. A further consideration ([Ni, 2019](#)) is the simulation performance of the univariate models: if the model for one variable has some deficiencies, then these will potentially be inherited by the simulations of other variables that depend on it. In practice therefore, decisions about variable ordering must be informed by a mixture of subject-matter knowledge and pragmatism.

3. Example: a bivariate weather generator for northern iberia

To illustrate the ideas, we construct and test a bivariate weather generator for precipitation and temperature in northern Iberia, defined here as the region of Spain and Portugal lying between 42.1 °N–44 °N, 9.5 °W–0°W (see [Fig. 1](#): the dimensions of the region shown here are around 770 km × 210 km). Part of France also falls within these latitude and longitude limits, but is excluded from the present study because it falls on the opposite side of the Pyrenean mountain range. Northern Iberia experiences dry summers and wet winters, with a climate that is strongly influenced both by local topography and by its proximity to the Atlantic Ocean ([Vicente-Serrano et al., 2009](#)). All code and data used in the study are provided in the online electronic supplement.

3.1. Data

3.1.1. Precipitation and temperature data

The precipitation and temperature data are from the European Climate Assessment & (EC&A) Dataset ([Klein Tank et al., 2002](#)), available from <http://www.ecad.eu>. The EC&A dataset provides a ‘non-blended’ data product in which some values are missing, and a ‘blended’ product in which these missing values have been infilled. The non-blended data are used here, because complete datasets are not required for the GLM methodology and because the infilling of missing values creates its own problems, such as the potential for the infilled values to have different statistical properties from the remainder of the data. However, for the precipitation data in particular, there are some differences in recording resolution between stations and over time: such inconsistencies can appear as model deficiencies in some of the diagnostics produced by *Rglimclim*. To minimise the impact on the model-building process therefore, all precipitation values have been rounded to the nearest 0.5 mm.

We work with data spanning the period 1979 to 2018, for which relevant covariate data are publicly available. There are 34 stations within the study area providing data for some or all of this period, and for which all required covariate data are available. 25 of the 34 stations (solid squares in [Fig. 1](#)) were operational at some point during the period from 1979 to 2000, which is used to build and calibrate models; and 31 of the stations were operational during the period from 2001 onwards which is used for model validation. Unfortunately, no data are available before 2001 for the cluster of easternmost stations shown in [Fig. 1](#): thus there is no information from which to model the systematic regional variation in this eastern part of the region. In view of the complex topography here, it would be unwise to extrapolate a statistical description of regional variation that has been developed using information from the remaining sites. These eastern sites have not been considered further therefore; and the longitude range for the developed models has been restricted to 9.5 °W–1°W.

3.1.2. Topographic data

The EC&A dataset provides geographical coordinates (latitude and longitude) and altitude for each station. These are supplemented with topographic data, at roughly 1 km² resolution, from the GTOPO30 Digital Elevation Model at http://webmap.ornl.gov/wcsdown/dataset.jsp?ds_id=10003. Following [Ambrosino et al. \(2014\)](#), these have been used to compute several other topographic indices for each station: the

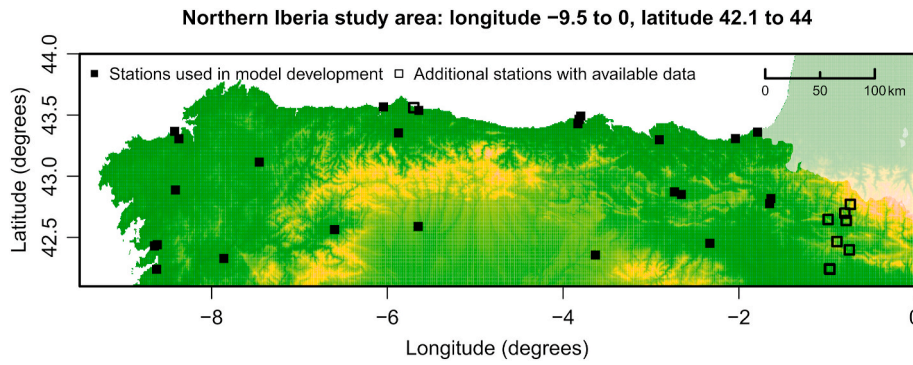


Fig. 1. Topographic map of northern Iberia, with stations for which data are available for all or part of the period 1979–2018. The semi-opaque area is part of France, which is not considered in the study.

mapped altitude (i.e. the GTOPO30 altitude for the pixel containing the station) together with measures of aspect (west-east and south-north average slopes over domains of size 3×3 , 10×10 and 30×30 km² centred on each station) and topographic variability (altitude standard deviations computed for the same domains).

3.1.3. Atmospheric covariates

To incorporate climate change signals into the developed weather generator, we condition on spatially averaged values of mean sea level pressure, 10-m u - and v -wind velocity components and wind speed, 2-m air temperature and dewpoint temperature, for the region 27.5°N – 45°N , 10°W – 15°E . These variables are derived from the ERA-INTERIM Reanalysis dataset (Dee et al., 2011), available from <https://www.ecmwf.int/en/forecasts/datasets>. ERA-INTERIM provides six-hourly fields of each variable (except wind speed, which is calculated from the u - and v -components) on a 0.75° grid; each variable has been averaged first over the selected region for each six-hourly field, and then over time to provide daily and monthly series.

This choice of atmospheric covariates was arrived at by considering the criteria mentioned in Section 1 regarding the selection of large-scale indices for use in statistical downscaling applications. In particular, the choice builds on the work of Gutiérrez et al. (2013) who carried out a thorough and physically-informed evaluation of potential covariates for downscaling with climate change applications in mind, and who concluded that for downscaling in Spain one should use sea level pressure and mean temperature from their region Z8: this is the region considered here. The additional inclusion of dewpoint temperature and wind information provides information on moisture availability and airflow, which are considered important for precipitation downscaling (Wilby and Wigley, 2000).

As an alternative to the use of spatially-averaged weather variables, it would in principle be possible to develop a weather generator based on discrete ‘weather types’ defined in terms of the large-scale atmospheric structure (Maraun et al., 2010). Weather state information can be incorporated into a GLM, by defining ‘dummy’ binary covariates taking the value 0 or 1 to indicate whether or not a particular day belongs to a specific weather type. Some experimentation suggests that, for the region considered here, weather generators based on spatially averaged atmospheric variables outperform those based on weather types. The latter are not considered further, therefore.

3.2. Model-building and diagnostics

The construction of a bivariate, multi-site weather generator for northern Iberia is done in three stages. The first is to build multi-site models for precipitation and temperature individually, without including atmospheric covariates, to characterise the climatology of each variable. Next, the individual models are linked to incorporate inter-variable dependence. Finally, the resulting joint model is extended

to include the effects of large-scale atmospheric covariates.

3.2.1. First stage: modelling precipitation and temperature individually

Following standard practice (Chandler, 2005; Yang et al., 2005; Ambrosino et al., 2014), we use logistic regression models for precipitation occurrence, gamma distributions for precipitation intensity on ‘wet’ days, and normal distributions for temperature. Specifically, let R_{st} and T_{st} denote the precipitation and temperature respectively, at site s on day t . Then the model structure is:

$$P(R_{st} > 0) = \pi_{st}; \quad (2)$$

$$R_{st} | R_{st} > 0 \sim \Gamma(\alpha, \lambda_{st}), \text{ with } \mathbb{E}(R_{st} | R_{st} > 0) = \alpha / \lambda_{st} = \mu_{st}^{(R)}, \text{ say}; \quad (3)$$

$$\text{and } T_{st} \sim N(\mu_{st}^{(T)}, \sigma_{st}^2). \quad (4)$$

The quantities π_{st} (the probability of experiencing precipitation at site s on day t), $\mu_{st}^{(R)}$ (the expected precipitation intensity on a ‘wet’ day), $\mu_{st}^{(T)}$ and σ_{st}^2 (the expected value and variance of the temperature distribution) are related to linear predictors via logistic, logarithmic, identity and logarithmic link functions respectively. Thus, denoting by $\eta_{st}^{(\cdot)}$ a generic linear predictor (i.e. linear combination of covariate values), we set

$$\log\left(\frac{\pi_{st}}{1 - \pi_{st}}\right) = \eta_{st}^{(\pi)}; \quad (5)$$

$$\log \mu_{st}^{(R)} = \eta_{st}^{(\mu^{(R)})}; \quad (6)$$

$$\mu_{st}^{(T)} = \eta_{st}^{(\mu^{(T)})}; \quad (7)$$

$$\text{and } \log \sigma_{st}^2 = \eta_{st}^{(\sigma^2)}. \quad (8)$$

The parameter α in equation (3) (the coefficient of variation of precipitation intensity distributions) is assumed to be constant — this assumption is usually made in the literature. However, the variance of the temperature distribution is allowed to depend on covariates via equation (8): this accounts for features such as seasonality in variance, which is common in temperature data.

In each stage of model-building, it is necessary to choose covariates and estimate their coefficients for each of the linear predictors defined by equations (5)–(8). For the first stage, this was done by starting with simple models and gradually increasing their complexity to represent seasonal and topographic variation along with temporal dependence, using diagnostic plots and measures of predictive performance (such as root mean squared error) together with formal model comparisons to inform the model development. Throughout, attention was restricted to mathematically coherent model structures: for example, sine and cosine components were always considered in pairs when developing Fourier

representations of seasonality; polynomial representations of specified degree were always complete (for example, a complete quadratic function of latitude and longitude includes terms in latitude, longitude, latitude², longitude² and latitude \times longitude); and the inclusion of an interaction term required the inclusion of the corresponding main effects as well. For more on these considerations when building statistical models, see Faraway (2014, Chapter 10). The process is documented fully in the online supplement. The first section of Table 1 provides summary information about the models developed by the end of this first stage; Figs. 2 and 3 show some specimen diagnostics for the precipitation occurrence and temperature models respectively.

By comparison with simpler weather generators, the models developed here appear complex: for example, a first-order Markov model uses just two parameters to represent precipitation occurrence, whereas the first-stage occurrence model in Table 1 has 23 (six representing regional variation as a function of longitude, latitude and altitude; four providing a simplified Fourier representation of seasonality; three representing dependence on previous days' precipitation occurrence, a nonlinear parameter representing an exponential decay rate of previous days' dependence on inter-site distance; an intercept; and several interaction terms). In fact, however, the models in Table 1 are relatively parsimonious because they account for seasonal and regional variation at all sites simultaneously. A 'standard' WGEN-type generator might fit separate pairs of Markov Chain parameters at each site and for each of four seasons: with 25 sites providing data in the fitting period, this would require $2 \times 4 \times 25 = 200$ parameters in total — almost ten times as many as the GLM fitted here. Similar comments apply to the remaining first-stage models.

Table 1 also shows that temperature is much more predictable than precipitation: 88% of the variation in daily temperature is attributable to autocorrelation and to regional and seasonal variation, whereas just 10% of the variation in precipitation intensity is similarly attributable. This is typical in our experience.

Fig. 2(a) shows the spatial pattern of residuals from the final first-stage precipitation occurrence model. The pattern of positive and negative mean residuals (indicated by solid and dashed circles respectively) appears fairly random, indicating that the model has captured most of the systematic large-scale regional variation in precipitation occurrence. At many sites however, the mean residual differs significantly from zero at the 5% level (circles drawn in thick lines): this should occur at around 5% of locations in a perfect model, so the plot indicates some mismatch — albeit unsystematic — between model and data. For precipitation occurrence, this phenomenon is also typical in our experience. It is probably caused either by very local-scale controls on precipitation, or by differences in recording practice or measurement technology (see, for example, Yang et al., 2006). This kind of plot can

Table 1

Summaries of models fitted at each stage of the weather generator development. The '(R)MSE' column gives the root mean squared error for the precipitation intensity (in mm) and temperature mean (in degrees Celsius) models, and the mean squared error (equivalent to the Brier score) for the precipitation occurrence models. The ' $R^2(\%)$ ' column gives the percentage of variance explained for the precipitation intensity and temperature mean models.

	Model	# parameters	(R)MSE	$R^2(\%)$
First stage	Precipitation occurrence	23	0.175	–
	Precipitation intensity	22	10.00	9.9
	Temperature mean	21	1.92	87.9
	Temperature variance	11	–	–
	Precipitation occurrence	27	0.174	–
Second stage	Precipitation intensity	26	9.94	11.0
	Temperature mean	28	1.88	88.4
	Temperature variance	15	–	–
	Precipitation occurrence	43	0.156	–
Third stage	Precipitation intensity	41	9.82	13.2
	Temperature mean	38	1.78	89.6
	Temperature variance	29	–	–

help to identify potential data quality problems, and also to highlight specific locations at which the simulation performance of a weather generator should be checked carefully.

Fig. 2(b) provides a check on the chosen inter-site dependence structure for the first-stage precipitation occurrence model. In this case, dependence is represented using the correlation structure of a latent Gaussian process: when simulating from the fitted model at time t , non-zero precipitation is allocated to site s if $Z_{st} > \tau_{st}$ where Z_{st} is a standard Gaussian variable and the threshold τ_{st} is chosen to ensure that $P(R_{st} > 0) = \pi_{st}$ as defined by equations (2) and (5). 'Empirical' correlations between the Gaussian variables at each pair of sites are estimated by matching the observed and expected proportions of days on which both sites are wet: these are plotted against inter-site distance in Fig. 2 (b), and a spatial correlation model (here, a simple exponential model) is fitted using weighted least-squares as described in Ambrosino et al. (2014) and in the Rglimclim manual. In the plot, colour intensity indicates the number of pairs of observations contributing to each empirical correlation estimate: darker points represent more precise estimates. This is an aid to interpretation, ensuring that a visual comparison of the empirical correlations with the fitted curve is predominantly based on the most informative points. The exponential model provides a reasonable overall representation of the empirical structure here, except for a handful of points at the bottom of the plot. Some of these points correspond to negative correlations, which is unexpected: closer inspection of the Rglimclim output (not shown) reveals that they are all associated with a single site, coded S14, which also has a large negative mean residual according to Fig. 2(a). The negative correlations suggest the possibility of undocumented data quality problems at this site, although a careful inspection of the data revealed no obvious errors. For the present analysis, the data are therefore taken at face value. In this case, Fig. 2 suggests a model deficiency at site S14, which may lead to poor performance of the weather generator at this site.

Fig. 3 shows further diagnostics, this time for the final first-stage temperature model. The top two plots show no systematic seasonal structure in the mean or variance of the residuals (the mean residual for March lies slightly outside the 95% variability bands, but an occasional point outside these bands is to be expected): the model appears to have captured the seasonality in temperature mean and variance, therefore. The bottom left plot, however, shows an increasing trend over time in the mean residuals: this is a signal of long-term change that is not captured by the covariates in this first stage of model-building. The standard deviations in the bottom right plot show no such trend.

3.2.2. Second stage: bivariate modelling

The second stage of model-building is to link the precipitation and temperature models, to produce a bivariate model incorporating inter-variable dependence. As described earlier, this is done by including one of the variables as a covariate in the model(s) for the other. Unfortunately, in the present study it is hard to choose the 'primary' variable either on grounds of data availability (for the 'fitting' period 1979–2000 there are around 200 000 observations for each variable) or physical considerations. For example: one 'physical' argument is that precipitation is the primary variable because it influences temperature indirectly, particularly in winter when precipitation is associated with extensive cloud cover and hence warmer temperatures. Another is that temperature is the primary variable because it influences precipitation directly in summer, due to enhanced convection on warm days. This does not help. Both options have been explored, therefore: again, the online supplement provides details. The final second-stage models are summarised in the second section of Table 1. They all involve a handful of additional parameters compared with their first-stage counterparts, and there are some improvements in predictability as measured by (R) MSE. In practical terms however, these improvements are small: it is therefore unsurprising that the diagnostics for these models (not shown) are similar to those from the first stage.

In the absence of clear guidance based on either data availability or

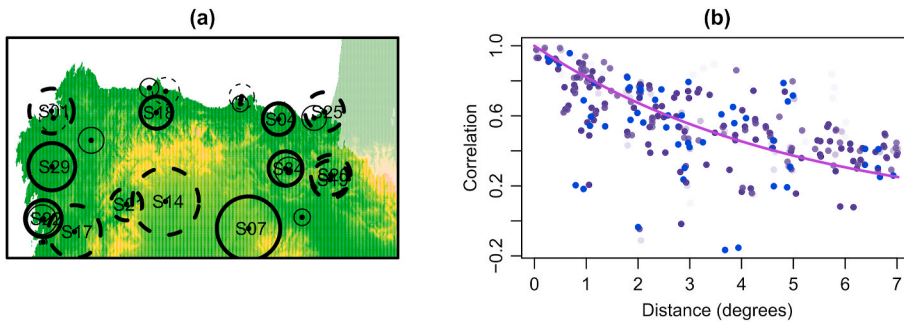


Fig. 2. Specimen diagnostic plots for first-stage precipitation occurrence model. In panel (a), the circle sizes show the magnitudes of mean Pearson residuals at each site. Solid circles indicate positive values (observed proportion of wet days higher than expected), and dashed circles indicate negative values. Thick circles indicate locations where mean residuals differ significantly from zero at the 5% level: the site identifiers for these locations are also shown. Panel (b) shows the estimated correlations between latent Gaussian variables at each pair of sites, as a function of inter-site distance (calculated from the geographical co-ordinates in degrees). The intensity of shading for each point corresponds to the number of days' data available for calculating the corresponding correlation. The smooth curve is the fitted inter-site correlation model: $\rho(d) = \exp(-0.1973d)$ where d is the inter-site distance.

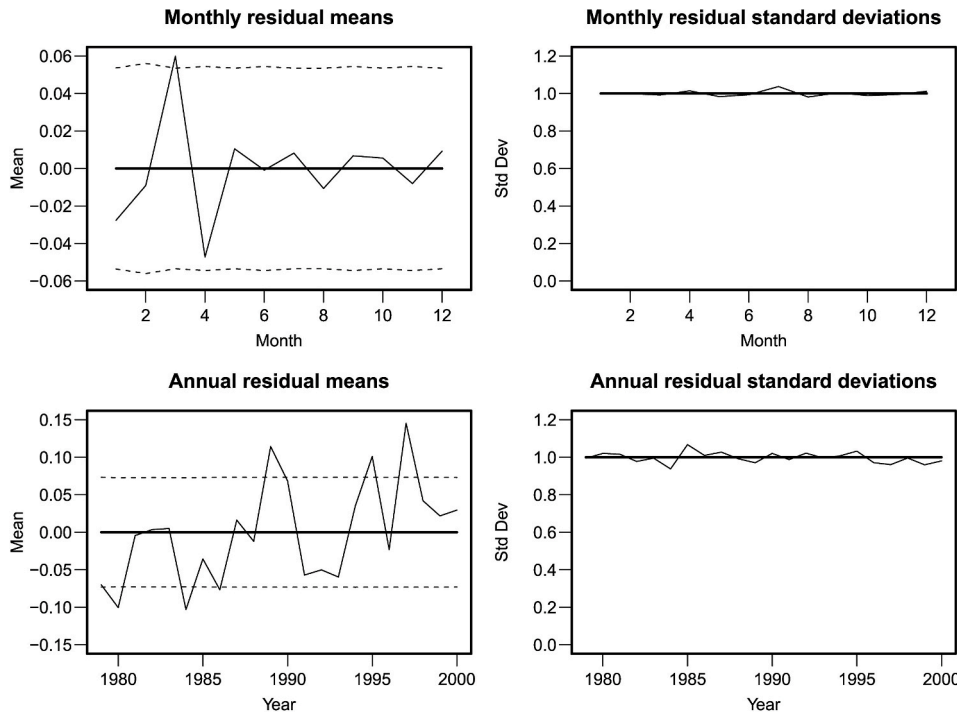


Fig. 3. Specimen diagnostic plots for first-stage temperature model. Means and standard deviations of Pearson residuals for each month of the year and for each year of the fitting period. Dashed lines in the left-hand plots show the limits within which 95% of the mean residuals should lie according to the model. Horizontal lines in the right-hand plots show the theoretical standard deviations.

physical considerations, the model diagnostics for both the first- and second-stage models can be used to decide whether temperature or precipitation should be considered as the primary variable in a bivariate weather generator. These diagnostics suggest that the performance of the marginal (i.e. first-stage) temperature model is better overall than that of the precipitation intensity model (see online supplement). Moreover, the temperature model residuals are small so that any violations of model assumptions will have a relatively minor effect on simulation performance; and any small deficiencies in the temperature model are unlikely to compromise the precipitation simulations, because the total percentage of explained variability in the second-stage precipitation intensity model is low (see Table 1). A final consideration is the very clear upward trend in the mean residuals from the temperature model (Fig. 3): we might expect that this can be explained by conditioning on large-scale temperatures, thus providing a relatively simple, parsimonious and physically intuitive way to incorporate the majority of any climate change signal into downscaled temperature

simulations. The effect of this signal will be inherited by the simulated precipitation if this is conditioned on local temperatures, but not otherwise.

3.2.3. Final stage: incorporating large-scale atmospheric covariates

The arguments above suggest treating temperature as the primary variable in this case study. The final stage of model-building is therefore to add the effects of large-scale atmospheric covariates to both the first-stage temperature and second-stage precipitation models. Initially, monthly covariate data were used: however, the residuals from the resulting temperature model still showed a clear increasing trend over time. This trend almost completely disappeared when daily covariate data were used (see Fig. 4), suggesting that daily large-scale information is needed to capture the local climate change signal in this region. Fig. 4 also shows, however, that 1995 and 1997 were both unusually warm years given the covariates considered: this may be due to other drivers of long-term change in the region that have not been considered here, or it

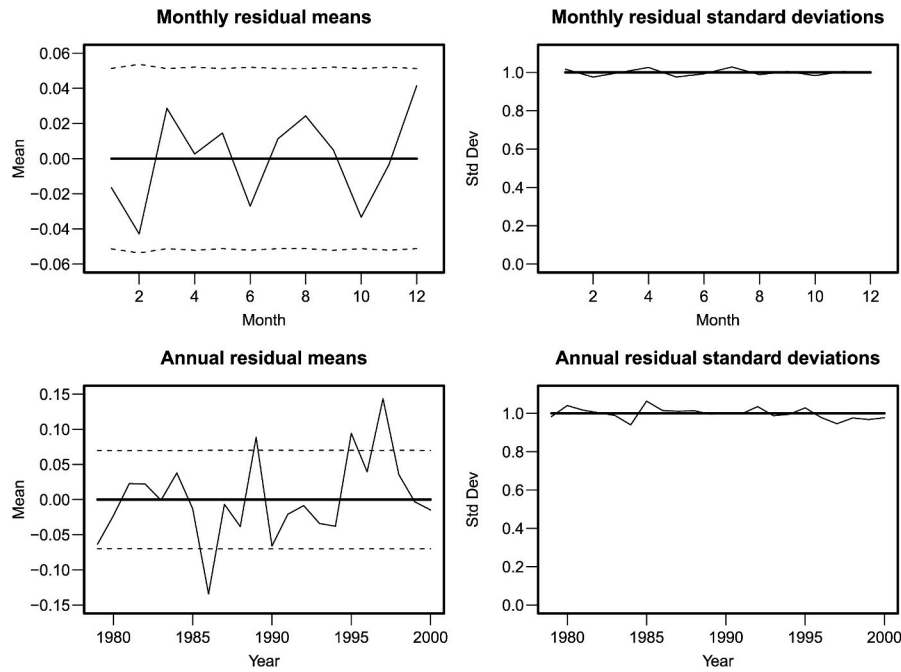


Fig. 4. As Fig. 3, but for the final temperature model including the effects of daily large-scale covariates.

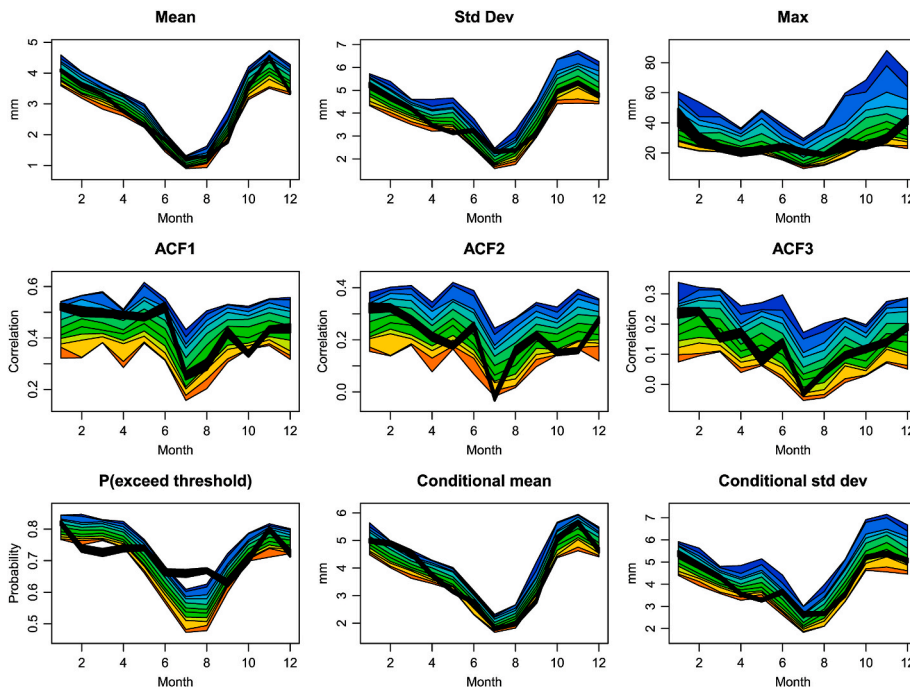


Fig. 5. Observed and simulated monthly summary statistics for areally averaged daily precipitation time series, 2001–2018. Coloured bands indicate the range and 1st, 5th, 10th, 25th, 50th, 75th, 90th, 95th and 99th percentiles of simulated distributions; black bands are uncertainty envelopes for observed quantities obtained from 39 imputations. Top row: mean, standard deviation and maximum daily rainfall. Second row: autocorrelation at lags 1 to 3. Bottom row: proportion of wet days, wet-day mean and standard deviation.

may indicate that the spatial domain chosen for the covariates is suboptimal.

The final section of Table 1 shows that the inclusion of large-scale atmospheric covariate information typically involves up to 15 additional parameters for each model, with relatively modest gains in predictive performance (e.g. the RMSE in the temperature model drops from 1.92°C to 1.78°C, and that for precipitation intensity drops from 9.94 mm to 9.82 mm). Most of the final-stage models include the effects of five of the six large-scale covariates (wind speed is omitted from all models, because the u and v wind velocity components are both included), in some cases with seasonal and regional variation in the effects. As usual, full details are in the online supplement.

The small gains in predictive performance in this final model-building stage may initially appear to contradict suggestions in the literature that large-scale atmospheric covariates can explain a substantial proportion of variation in local-scale weather (e.g. Maraun et al., 2019b). In reality however, there is no contradiction. There are indeed techniques that implicitly attribute systematic local-scale variation entirely to large-scale covariates. However, this systematic variation includes both climatological and interannual components, the former involving seasonality and temporal autocorrelation — although autocorrelation can be poorly represented by techniques that rely solely on the large-scale covariates (Maraun et al., 2019a). In the present framework, the climatological component of variation is represented by the

first- and second-stage models: the primary additional contribution of the atmospheric covariates is thus to explain the interannual component of systematic variation. At a daily time scale, it is unsurprising that this contribution is small.

3.3. Simulation

The fitted models can be tested by using them to simulate multi-site daily sequences of temperature and precipitation, and comparing the properties of these sequences with those of observations. This complements the diagnostics used during model development, which focused on distributions conditional on previous days' weather (due to the inclusion of lagged terms in the models) rather than on overall properties that are more directly relevant in applications.

A wide range of simulation tests is reported in the online supplement: this section presents some representative results, for a period (2001–2018) that was not used for model development. There is one 'new' station (coded S10) with data for this period, but where no data were available during the 1979–2000 period used for model development. The results at this station can therefore be regarded as indicative of weather generator performance at ungauged locations within the study area. For illustrative purposes, performance is assessed by inspecting a range of statistical properties including means, variances, extremes and inter-variable correlations, separately for each month of the year; and also by inspecting time series of seasonal means or totals.

Due to the stochastic nature of a weather generator, an exact match between observed and simulated properties is not to be expected: any 'observed' quantity must be regarded as coming from an underlying probability distribution, so performance assessment seeks to determine whether the observed quantities of interest could plausibly have been drawn from the corresponding distributions. In the absence of missing observations during a specified time period, one way to achieve this is by simulating a large number of multi-site daily time series over the same period (we use 100 simulations below) and computing the quantities of interest separately for each simulation. The result is a simulated distribution for each quantity. If the simulations are realistic then the

observed values of each quantity should span the range of the simulated distributions.

This approach cannot be applied directly in the presence of missing data however, because in this case the 'observed' quantities of interest cannot be calculated exactly. As discussed in Section 2, multiple imputation provides a convenient way to characterise the resulting uncertainty. In the work reported below, 39 independent imputations of missing daily values have been produced, to produce 39 'completed' datasets: for each of these, the corresponding properties have been calculated. The range of the results is a 95% uncertainty interval. To see this, note that if 39 values are imputed from a distribution that was used to generate an observation, then the observation has a 1/40 chance of being smaller than all of the imputed values and a 1/40 chance of being larger: hence it has a 5% chance of lying outside the range of the imputations.

Fig. 5 shows some results relating to monthly summary statistics of areally averaged daily precipitation. The imputation bands show that there is some uncertainty due to missing observations, but it is never excessively large. Subject to this uncertainty, the observed quantities span the ranges of the simulated distributions well overall: the main exception is the proportion of wet days (bottom left panel) for which the seasonal cycle in the observations is weaker than that in the simulations. This presumably relates to some aspect of the precipitation occurrence model. Elsewhere, the simulations fail to capture fully the very low lag 2 autocorrelation in July. Whether this is a problem in practice will, of course, be application-dependent.

Fig. 6 shows the performance of a similar set of monthly summary statistics for temperature, this time for a single location rather than an areal average. The location is station S10, which was not used for model fitting. There is little variability in mean monthly temperatures between simulations (as expected, given the low RMSE for the fitted temperature model), and there is a tendency to slightly underestimate the observed mean and minimum temperature in July through September. Conversely, the simulated standard deviations at this station tend to be too high in summer. Most other statistics are reproduced reasonably well, including the correlation with precipitation. The imputation

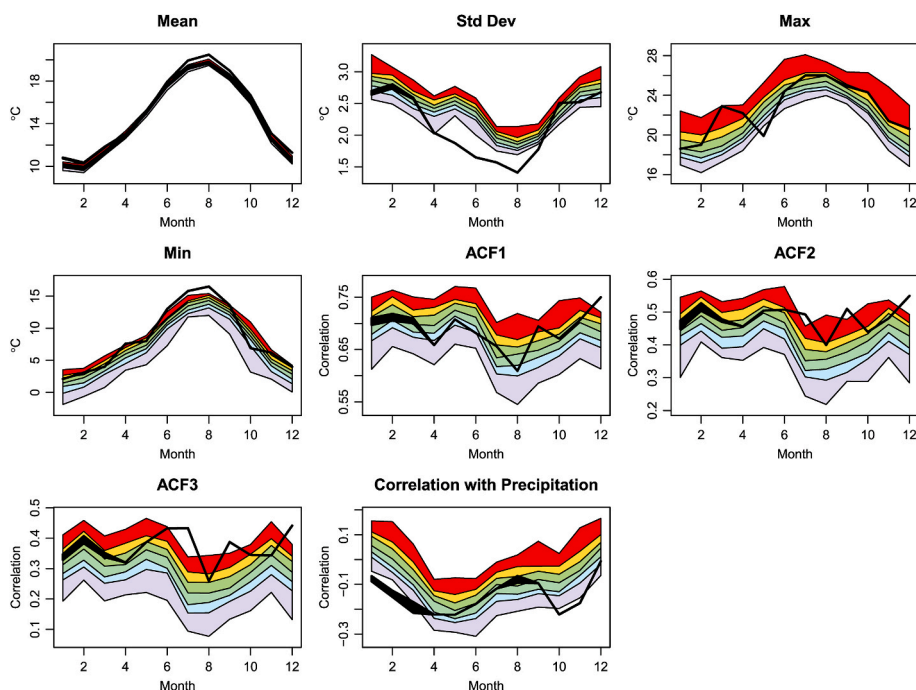


Fig. 6. As Fig. 5, but for daily temperature time series at station S10. Top row: mean, standard deviation and maximum daily temperature. Second row: minimum daily temperature and autocorrelation at lags 1 and 2. Bottom row: lag 3 autocorrelation, and correlation with daily precipitation at the same station.

uncertainties tend to be relatively smaller than for precipitation, albeit with wider bands (presumably corresponding to more missing observations) for some statistics in the first part of the year.

Finally, Fig. 7 shows the distributions of annual time series of areally averaged seasonal mean daily precipitation. These distributions were calculated by first averaging the daily time series across sites and then, for each year, computing the mean precipitation within each of the standard climatological seasons. As before, the observations more or less span the range of the simulated distributions. The substantial interannual variability in the simulations is of particular interest: the only possible source of this variability is the large-scale covariates in the models, despite the fact that they explain a small additional proportion of the daily variation according to Table 1. Moreover, the simulated interannual variation mostly tracks that in the observations: this provides some reassurance that the models have captured the relationships with large-scale conditions and hence might be expected to provide a reasonable climate change response as these conditions change in the future. It is particularly noteworthy that the simulations capture (just) the severe summer (June, July, August) drought of 2005, which was one of the worst on record and has been associated with an unusual combination of synoptic conditions (Garcia-Herrera et al., 2007).

Earlier, it was noted that the precipitation occurrence model performed poorly at site S14. As expected, this affects the simulation performance at this site, including for the 1980–2000 model development period (see the online supplement). For example, the observed proportion of wet days at this site is just under 0.25 in January and February whereas the median of the simulated distributions is just under 0.3. Coupled with a tendency for the simulations to overestimate mean wet-day rainfall intensities, this produces an upward bias of around 30% in overall mean precipitation compared with observations. This would often be considered inadequate in applications. Strategies for resolving the problem include: inspect the data from this site to look for symptoms of systematic under-recording, for example by comparing with alternative data sources such as reanalysis; and, if the data really are beyond reproach, add an indicator variable to the models that takes the value 1 for any observation from this site, and zero otherwise. The effect of this

will be to adjust the model fit so as to match the observed properties more closely at this site.

3.4. Design and use of the software

The Rglimclim software is designed to be efficient and flexible, and to encourage a structured approach to weather generator construction. This section provides a brief overview of its design philosophy and implications for users.

3.5. Design considerations

The software is provided as a library to be run in the R programming environment (R Core Team, 2019), taking advantage of R's user-friendly interface and graphical capabilities. As an interpreted rather than compiled language however, R is relatively inefficient for computationally intensive work: the model-fitting and simulation code is therefore written in Fortran, and automatically compiled into the library on installation. Moreover, to allow processing of large datasets, as far as possible large arrays are stored in direct-access temporary files rather than in computer memory: fairly complicated models can thus be fitted to datasets containing millions of observations on a modest modern desktop or laptop computer. The software is also designed with computational efficiency in mind, for example by storing rather than recalculating quantities that are used repeatedly. To give an indication of computational speed: for the case study in Section 3, it took around 3 s to produce each 20-year simulation of two variables at 26 sites on a laptop with 2.2 GHz processor, running R version 3.6.1 in RStudio 1.1.456 under the Ubuntu 18.04.2 operating system. On the same machine running Windows 7, it took around 12 s per simulation. However, simulation at more than a few hundred sites is likely to be slow due to the matrix factorisations needed to sample pseudo-random numbers from high-dimensional multivariate normal distributions (see, for example, Monahan, 2001).

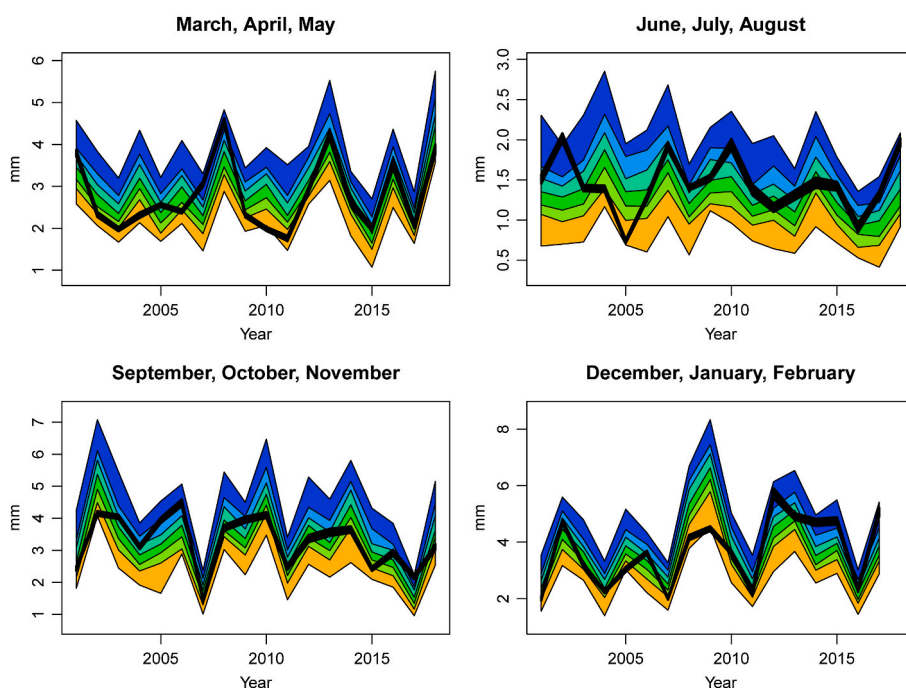


Fig. 7. Observed and simulated time series of areally averaged seasonal mean daily precipitation, 2001–2018. Coloured and black bands show percentiles/range of simulation and imputation distributions as in Fig. 5.

3.6. Use of the software

As discussed in Section 2, covariates in a weather generator can be split into a small number of categories: a model can be broken down into components corresponding to each of these categories, therefore. Rglimclim offers a menu of options for representing the structure of each component. Flexibility in the representation of regional and interannual variation is achieved by linking to user-defined databases of site attributes and potential drivers, such as the atmospheric covariates considered in the case study above. Moreover, the software is predominantly object-oriented in its approach so that plots (and other outputs) such as those in Figs. 2–7 can each be produced using a single command.

In many modern software environments, models are specified using a structured formula syntax as with the native `glm()` command in R, for example. Typically however, the use of such formulae requires significant amounts of data manipulation. For example, if autocorrelation is represented via covariates computed from lagged values of the response variable, these covariates must usually be pre-calculated: this is inconvenient, and sometimes impossible for covariates that depend on the values of unknown parameters as with the precipitation occurrence (and other) models in the case study above. In Rglimclim, such calculations are all handled internally: however, this makes it hard to develop a concise syntax that can capture all possible model structures. Models are therefore specified using definition files, the format of which is described in the software manual. The online supplement provides definition files for all models used in the case study.

Most software packages for fitting GLMs offer a wide range of different distributions. The range available in Rglimclim is currently limited to the models and distributions (logistic regression, gamma GLM with log link, normal linear model and heteroscedastic normal model) used in the case study. These models provide sufficient flexibility to deal with most situations in which weather generators are required: for example, gamma GLMs are often found to provide a reasonable representation of positive-valued quantities such as precipitation intensity or wind speed, while normal or normal-heteroscedastic models may be more appropriate for quantities such as temperature, pressure and humidity. Further distributions may be added in the future: it is straightforward in principle to estimate coefficient vectors for additional distributions, but harder to devise tractable and realistic models for inter-site dependence and imputation.

4. Discussion

The Rglimclim software uses the GLM framework to unify a range of modelling techniques for multi-site, multivariate weather generation incorporating scenarios of climate change. By comparison with some other weather generators it is relatively unrestrictive in its data requirements, in the sense that it does not require complete datasets for model calibration and is able to generate sequences at ungauged locations providing that sufficient data are available from ‘similar’ locations for model calibration. Its multiple imputation facilities provide a useful way to examine the uncertainties associated with incomplete observational records. It is also relatively scalable, such that model-fitting and simulation at a few hundred locations is feasible on a typical laptop computer.

There are, nonetheless, opportunities to extend the software capabilities in the future. One is to offer a wider range of distributions, as noted above. Another is to relax the assumption of a constant shape parameter when using gamma distributions: this would potentially improve the reproduction of seasonal precipitation extremes, which has

been noted as an issue in previous studies (e.g. Yang et al., 2005). Moreover, there is scope to extend the range of inter-site dependence structures available, for example by allowing for seasonal variation in the strength and spatial scale of dependence.

As with any powerful modelling tool, the appropriate use of Rglimclim requires care and awareness. It is developed primarily for use in situations where the potential impacts of climate may be sensitive to the detailed spatial, temporal and inter-variable structure of local weather: in less complex applications, it may be preferable to use a simpler weather generator requiring less effort to implement. The case study above has illustrated some of the issues to be considered when developing a weather generator using Rglimclim: these include decisions about the ordering of the variables in a multivariate model, the appropriate choice of large-scale atmospheric covariates, and the need to consider and compare a variety of different options that are informed both by diagnostics and by an understanding of the underlying modelling principles. The last point is worth emphasising: Rglimclim is not itself a weather generator, rather a tool for developing application-specific weather generators. This has implications for the reporting of results, and for studies that aim to compare different weather generation techniques (e.g. Frost et al., 2011; Fu et al., 2018): to enable a critical evaluation of results and model performance, it is necessary to provide full details of the model structures that have been used and the rationale for them (in the present paper, this information is in the online supplement).

The potential for overfitting is often raised as a concern, in the context of the modelling methodology used here. This stems partly from the fact that the models often have relatively large numbers of parameters, as in Table 1. However, as noted in Section 3, the approach in fact requires fewer parameters than the common practice of fitting relatively simple model structures with separate parameter sets for individual sites and/or seasons. Nonetheless, care is needed to ensure that the available data contain enough information to support the models being fitted: for example, a detailed representation of systematic regional variation will only be supported if sufficient data are available from a relatively large number of locations that are well distributed over the study area. A useful operational guideline is to monitor the standard errors of the estimated model coefficients as an initial simple model is gradually expanded: a sudden increase in one or more standard errors, by an order of magnitude or more, suggests that a model has become too complex for the data available.

When building complex models, a further consideration is that any statistical measure of model ‘quality’ is by definition conditioned on the available data; and estimation methods such as maximum likelihood are designed to find a model that matches the data as closely as possible in some sense. When developing a weather generator based on station data, the data may not always represent precisely the phenomenon of interest: more seriously perhaps, data from different time periods or locations may represent slightly different phenomena due to changes in measurement technology or variation in observer practice. This is worth bearing in mind when monitoring overall measures of performance during the model-building process. For example, if recording practice for small precipitation amounts at one location differs substantially from that elsewhere, then overall performance measures may suggest that a very complex precipitation occurrence model fits the data much better than a relatively simple one — but this may be almost entirely due to an improved fit at the single location, which may even be detrimental elsewhere. Diagnostics such as Fig. 2 (a) can help to identify this kind of situation; and subject-matter considerations can be used to determine whether a particular model structure is reasonable. Model-building is

thus a holistic process, requiring informed judgements about how to balance multiple aspects of performance.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The development of Rglimclim was partially funded by the Natural Environment Research Council, award NE/I006656/1.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.envsoft.2020.104867>.

References

- Ailliot, P., Thompson, C., Thomson, P., 2009. Space-time modelling of precipitation using a hidden Markov model and censored Gaussian distributions. *Appl. Statist.* 58, 405–426.
- Ambrosino, C., Chandler, R.E., Todd, M.C., 2014. Rainfall-derived growing season characteristics for agricultural impact assessments in South Africa. *Theor. Appl. Climatol.* 115, 411–426. <https://doi.org/10.1007/s00704-013-0896-y>.
- Asong, Z.E., Khaliq, M.N., Wheeler, H.S., 2016. Multisite multivariate modeling of daily precipitation and temperature in the Canadian Prairie Provinces using generalized linear models. *Clim. Dynam.* 47, 2901–2921. <https://doi.org/10.1007/s00382-016-3004-z>.
- Ayar, P.V., Vrac, M., Bastin, S., Carreau, J., Deque, M., Gallardo, C., 2016. Intercomparison of statistical and dynamical downscaling models under the EURO- and MED-CORDEX initiative framework: Present climate evaluations. *Clim. Dynam.* 46 (3–4), 1301–1329.
- Beersma, J.J., Buishand, T.A., 2003. Multi-site simulation of daily precipitation and temperature conditional on the atmospheric circulation. *Clim. Res.* 25 (2), 121–133.
- Bruni, G., Reinoso, R., van de Giesen, N.C., Clemens, F.H.L.R., ten Veldhuis, J.A.E., 2015. On the sensitivity of urban hydrodynamic modelling to rainfall spatial and temporal resolution. *Hydrol. Earth Syst. Sci.* 19 (2), 691–709.
- Buishand, T., Brandsma, T., 2001. Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling. *Water Resour. Res.* 37, 2761–2776.
- Chandler, R.E., 2005. On the use of generalized linear models for interpreting climate variability. *Environmetrics* 16 (7), 699–715.
- Chandler, R.E., 2018. Rglimclim: A Multisite, Multivariate Weather Generator Based on Generalised Linear Models. R package version 1.3- 6. Available from: <http://www.homepages.ucl.ac.uk/~ucajarc/work/glimclim.html>.
- Chandler, R.E., Ambrosino, C., 2015. Spatiotemporal modelling for environmental applications: Challenges and opportunities. In: Friedl, H., Wagner, H. (Eds.), *Proceedings of the 30th International Workshop on Statistical Modelling*. Linz, Austria, pp. 3–22.
- Chandler, R.E., Bate, S.M., 2007. Inference for clustered data using the independence log-likelihood. *Biometrika* 94, 167–183.
- Chandler, R.E., Isham, V., Bellone, E., Yang, C., Northrop, P.J., 2006. Space-time modelling of rainfall for continuous simulation. In: Finkenstadt, B., H. L., Isham, V. (Eds.), *Statistics of Spatial-Temporal Systems*. CRC Press, pp. 169–207.
- Chandler, R.E., Scott, E.M., 2011. *Statistical Methods for Trend Detection and Analysis in the Environmental Sciences*. John Wiley & Sons, Chichester.
- Chen, J., Brissette, F.P., Zhang, X.J., 2014. A multi-site stochastic weather generator for daily precipitation and temperature. *Trans. ASABE (Am. Soc. Agric. Biol. Eng.)* 57 (5), 1375–1391. <https://doi.org/10.13031/trans.57.10685>.
- Chun, K.P., Mamet, S.D., Metsaranta, J., Barr, A., Johnstone, J., Wheeler, H., 2017. A novel stochastic method for reconstructing daily precipitation time-series using tree-ring data from the western Canadian Boreal Forest. *Dendrochronologia* 44, 9–18. <https://doi.org/10.1016/j.dendro.2017.01.003>.
- Coe, R., Stern, R.D., 1982. Fitting models to daily rainfall. *J. Appl. Meteorol.* 21, 1024–1031.
- Collins, M., Chandler, R.E., Cox, P.M., Huthnance, J.M., Rougier, J., Stephenson, D.B., 2012. Quantifying future climate change. *Nat. Clim. Change* 2, 403–409. <https://doi.org/10.1038/nclimate1414>.
- Cordano, E., Eccel, E., 2017. RMAWGEN: Multi-site auto-regressive weather GENERator. R package. Available from: <https://CRAN.R-project.org/package=RMAWGEN>.
- Davison, A.C., 2003. *Statistical Models*. Cambridge University Press, Cambridge.
- Davison, A.C., Padoan, S.A., Ribatet, M., 2012. Statistical modeling of spatial extremes. *Stat. Sci.* 27, 161–186.
- Dawid, A.P., 1986. Probability forecasting. In: Kotz, S., Johnson, N. (Eds.), *Encyclopedia of Statistical Sciences*. Wiley, New York, pp. 210–218.
- Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae, U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J., Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, I., Kållberg, P., Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.-J., Park, B.-K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.-N., Vitart, F., 2011. The ERA-Interim reanalysis: Configuration and performance of the data assimilation system. *Q. J. R. Meteorol. Soc.* 137 (656), 553–597. <https://doi.org/10.1002/qj.828>.
- Dobson, A.J., Barnett, A.G., 2018. *An Introduction to Generalized Linear Models*, fourth ed. Chapman and Hall/CRC Press, Boca Raton.
- Dubrovsky, M., Huth, R., Dabhi, H., Rotach, M.W., 2020. Parametric gridded weather generator for use in present and future climates: Focus on spatial temperature characteristics. *Theor. Appl. Climatol.* 139, 1031–1044. <https://doi.org/10.1007/s00704-019-03027-z>.
- Fahrmeir, L., Tutz, G., 2001. *Multivariate Statistical Modelling Based on Generalized Linear Models*, second ed. Springer-Verlag, New York.
- Faraway, J.J., 2014. *Linear Models with R*, second ed. Chapman and Hall/CRC, Boca Raton.
- Franklin, J., Davis, F.W., Ikegami, M., Syphard, A.D., Flint, L.E., Flint, A.L., Hannah, L., 2013. Modeling plant species distributions under future climates: how fine scale do climate projections need to be? *Global Change Biol.* 19 (2), 473–483.
- Frost, A.J., Charles, S.P., Timbal, B., Chiew, F.H.S., Mehrotra, R., Nguyen, K.C., Chandler, R.E., McGregor, J.L., Fu, G., Kirono, D.G.C., Fernandez, E., Kent, D.M., 2011. A comparison of multi-site daily rainfall downscaling techniques under Australian conditions. *J. Hydrol.* 408, 1–18. <https://doi.org/10.1016/j.jhydrol.2011.06.021>.
- Fu, G., Chiew, F.H.S., Shi, X., 2018. Generation of multi-site stochastic daily rainfall with four weather generators: a case study of Gloucester catchment in Australia. *Theor. Appl. Climatol.* 134, 1027–1046. <https://doi.org/10.1007/s00704-017-2306-3>.
- Furrer, E.M., Katz, R.W., 2008. Improving the simulation of extreme precipitation events by stochastic weather generators. *Water Resour. Res.* 44, W12439. <https://doi.org/10.1029/2008WR007316>.
- García-Herrera, R., Paredes, D., Trigo, R.M., Trigo, I.F., Hernandez, E., Barriopedro, D., Mendes, M.A., 2007. The outstanding 2004/05 drought in the Iberian Peninsula: Associated atmospheric circulation. *J. Hydrometeorol.* 8 (3), 483–498.
- Gires, A., Onof, C., Maksimovic, C., Schertzer, D., Tchiguirinskaia, I., Simoes, N., 2012. Quantifying the impact of small scale unmeasured rainfall variability on urban runoff through multifractal downscaling: A case study. *J. Hydrol.* 442, 117–128.
- Grunwald, G.K., Jones, R.H., 2000. Markov models for time series with mixed distribution. *Environmetrics* 11, 327–339.
- Gutiérrez, J.M., Maraun, D., Widmann, M., Huth, R., Hertig, E., Benestad, R., Roessler, O., Wibig, J., Wilcke, R., Kotlarski, S., San Martín, D., Herrera, S., Bedia, J., Casanueva, A., Manzanar, R., Iturbide, M., Vrac, M., Dubrovsky, M., Ribalaygua, J., Pórtoles, J., Rätty, O., Räisänen, J., Hingray, B., Raynaud, D., Casado, M.J., Ramos, P., Zerenner, T., Turco, M., Bosshard, T., Štěpánek, P., Bartholy, J., Pongrácz, R., Keller, D.E., Fischer, A.M., Cardoso, R.M., Soares, P.M.M., Czernecki, B., Pagé, C., 2019. An intercomparison of a large ensemble of statistical downscaling methods over Europe: Results from the VALUE perfect predictor cross-validation experiment. *Int. J. Climatol.* 39, 3750–3785. <https://doi.org/10.1002/joc.5462>.
- Gutiérrez, J.M., San-Martín, D., Brands, S., Manzanar, R., Herrera, S., 2013. Reassessing statistical downscaling techniques for their robust application under climate change conditions. *J. Clim.* 26, 171–188. <https://doi.org/10.1175/JCLI-D-11-00687.1>.
- Katz, R., Parlange, M., 1998. Overdispersion phenomenon of stochastic modeling of precipitation. *J. Clim.* 11, 591–601.
- Katz, R., Parlange, M., Naveau, P., 2002. Statistics of extremes in hydrology. *Adv. Water Resour.* 25, 1287–1304.
- Katz, R., Zheng, X., 1999. Mixture model for overdispersion of precipitation. *J. Clim.* 12, 2528–2537.
- Kenabatho, P.K., McIntyre, N.R., Chandler, R.E., Wheeler, H.S., 2012. Stochastic simulation of rainfall in the semi-arid Limpopo basin, Botswana. *Int. J. Climatol.* 32 (7), 1113–1127. <https://doi.org/10.1002/joc.2323>.
- Kilsby, C., Jones, P., Burton, A., Ford, A., Fowler, H., Harpham, C., James, P., Smith, A., Wilby, R.L., 2007. A daily weather generator for use in climate change studies. *Environ. Model. Software* 22 (12), 1705–1719.
- Klein Tank, A.M.G., Wijngaard, J.B., Können, G.P., Böhm, R., Demarée, G., Gocheva, A., Mileta, M., Pashiardis, S., Hejkrlik, L., Kern-Hansen, C., Heino, R., Bessemoulin, P., Müller-Westermeier, G., Tzanakou, M., Szalai, S., Pálásdóttir, T., Fitzgerald, D., Rubin, S., Capaldo, M., Maugeri, M., Leitass, A., Bukantis, A., Aberle, R., Engelen, A. F.V.V., Forland, E., Miletus, M., Coelho, F., Mares, C., Razuvaev, V., Nieplová, E., Cegnár, T., Antonio López, J., Dahlström, B., Moberg, A., Kirchhofer, W., Ceylan, A., Pachaliuk, O., Alexander, L.V., Petrovic, P., 2002. Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. *Int. J. Climatol.* 22, 1441–1453. <https://doi.org/10.1002/joc.773>.
- Lembrechts, O., Nijs, I., Lenoir, J., 2019. Incorporating microclimate into species distribution models. *Ecography* 42, 1267–1279.
- Maraun, D., Huth, R., Gutiérrez, J., San Martín, D., Dubrovsky, M., Fischer, A., Hertig, E., Soares, P.M.M., Bartholy, J., Pongrácz, R., Widmann, M., Casado, M.J., Ramos, P., Bedia, J., 2019a. The VALUE perfect predictor experiment: Evaluation of temporal variability. *Int. J. Climatol.* 39, 3786–3818. <https://doi.org/10.1002/joc.5222>.
- Maraun, D., Wetterhall, F., Ireson, A.M., Chandler, R.E., Kendon, E.J., Widmann, M., Brienen, S., Rust, H.W., Sauter, T., Themeßl, M., Venema, V., Chun, K., Goodess, C., Jones, R., Onof, C., Vrac, M., Thiele-Eich, I., 2010. Precipitation downscaling under climate change — recent developments to bridge the gap between dynamical models and the end user. *Rev. Geophys.* 48, RG3003. <https://doi.org/10.1029/2009RG000314>.
- Maraun, D., Widmann, M., 2018. *Statistical Downscaling and Bias Correction for Climate Research*. Cambridge University Press, Cambridge.

- Maraun, D., Widmann, M., Gutiérrez, J., 2019b. Statistical downscaling skill under present climate conditions: A synthesis of the VALUE perfect predictor experiment. *Int. J. Climatol.* 39, 3692–3703. <https://doi.org/10.1002/joc.5877>.
- McCullagh, P., Nelder, J., 1989. *Generalized Linear Models*, second ed. Chapman and Hall, London.
- Mockler, E.M., Chun, K.P., Sapriza-Azuri, G., Bruen, M., Wheeler, H.S., 2016. Assessing the relative importance of parameter and forcing uncertainty and their interactions in conceptual hydrological model simulations. *Adv. Water Resour.* 97, 299–313. <https://doi.org/10.1016/j.advwatres.2016.10.008>.
- Monahan, J.F., 2001. *Numerical Methods of Statistics*. Cambridge University Press, Cambridge.
- Ni, J., 2019. *Daily Weather Generation for Impact Assessments of Climate Change on Rice Production in Hokkaido*. B.Sc. Dissertation. Department of Statistical Science, University College London.
- Patton, A.J., 2012. A review of copula models for economic time series. *J. Multivariate Anal.* 110, 4–18.
- R Core Team, 2019. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resour. Res.* 17, 182–190.
- Segond, M.-L., Wheeler, H.S., Onof, C., 2007. The significance of spatial rainfall representation for flood runoff estimation: A numerical evaluation based on the Lee catchment, UK. *J. Hydrol.* 347, 116–131.
- Semenov, M., Brooks, R., Barrow, E., Richardson, C., 1998. Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. *Clim. Res.* 10, 95–107.
- Stehlik, J., Bárdossy, A., 2002. Multivariate stochastic downscaling model for generating daily precipitation series based on atmospheric circulation. *J. Hydrol.* 256, 120–141.
- Stern, R.D., Coe, R., 1984. A model fitting analysis of rainfall data (with discussion). *J. Roy. Stat. Soc. A147*, 1–34.
- Tosonoglu, F., Onof, C., 2017. Joint modelling of drought characteristics derived from historical and synthetic rainfalls: application of generalized linear models and copulas. *J. Hydrol.: Reg. Stud.* 14, 167–181. <https://doi.org/10.1016/j.ejrh.2017.11.001>.
- United Nations, 2015. *Transforming our world: the 2030 agenda for sustainable development*. UN resolution A/RES/70/1. URL <https://undocs.org/en/A/RES/70/1>.
- Vicente-Serrano, S.M., Begueria, S., Lopez-Moreno, J.I., El Kenawy, A.M., Angulo-Martinez, M., 2009. Daily atmospheric circulation events and extreme precipitation risk in northeast Spain: role of the north atlantic oscillation, the western mediterranean oscillation, and the mediterranean oscillation. *J. Geophys. Res. Atmos.* 114.
- Webster, R., Oliver, M.A., 2001. *Geostatistics for Environmental Scientists*. John Wiley & Sons, Chichester.
- Wheeler, H., 2002. Progress in and prospects for fluvial flood modelling. *Proc. Roy. Soc. Lond. A360*, 1409–1432.
- Widmann, M., Bedia, J., Gutiérrez, J.M., Bosshard, T., Hertig, E., Maraun, D., Casado, M. J., Ramos, P., Cardoso, R.M., Soares, P.M.M., Ribalaygua, J., Pagé, C., Fischer, A.M., Herrera, S., Huth, R., 2019. Validation of spatial variability in downscaling results from the VALUE perfect predictor experiment. *Int. J. Climatol.* 39, 3819–3845. <https://doi.org/10.1002/joc.6024>.
- Wilby, R., Wigley, T., 2000. Precipitation predictors for downscaling: observed and general circulation model relationships. *Int. J. Climatol.* 20 (6), 641–661.
- Wilby, R.L., Charles, S.P., Zorita, E., Timbal, B., Whetton, P., Mearns, L., 2004. Guidelines for use of climate scenarios developed from statistical downscaling methods. Tech. Rep. Intergovernmental Panel on Climate Change, Geneva, Switzerland.
- Wilks, D., 1998. Multisite generalization of a daily stochastic precipitation generation model. *J. Hydrol.* 210, 178–191.
- Wilks, D., 1999. Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agric. For. Meteorol.* 93 (3), 153–169.
- Yan, Z., Bate, S., Chandler, R.E., Isham, V.S., Wheeler, H.S., 2002. An analysis of daily maximum windspeed in northwestern Europe using Generalized Linear Models. *J. Clim.* 15 (No.15), 2073–2088.
- Yang, C., Chandler, R.E., Isham, V.S., Wheeler, H.S., 2005. Spatial-temporal rainfall simulation using generalized linear models. *Water Resour. Res.* 41 <https://doi.org/10.1029/2004WR003739>.
- Yang, C., Chandler, R.E., Isham, V.S., Wheeler, H.S., 2006. Quality control for daily observational rainfall series in the UK. *Water Environ. J.* 20, 185–193. <https://doi.org/10.1111/j.1747-6593.2006.00035.x>.
- Zhao, G., Siebert, S., Enders, A., Rezaei, E.E., Yan, C., Ewert, F., 2015. Demand for multi-scale weather data for regional crop modeling. *Agric. For. Meteorol.* 200, 156–171.