Stoch Environ Res Risk Assess (2015) 29:347–356
DOI 10.1007/s00477-014-0911-6

ORIGINAL PAPER

Coupled stochastic weather generation using spatial and generalized linear models

Andrew Verdin · Balaji Rajagopalan · William Kleiber · Richard W. Katz

Published online: 5 July 2014

Abstract We introduce a stochastic weather generator for the variables of minimum temperature, maximum temperature and precipitation occurrence. Temperature variables are modeled in vector autoregressive framework, conditional on precipitation occurrence. Precipitation occurrence arises via a probit model, and both temperature and occurrence are spatially correlated using spatial Gaussian processes. Additionally, local climate is included by spatially varying model coefficients, allowing spatially evolving relationships between variables. The method is illustrated on a network of stations in the Pampas region of Argentina where nonstationary relationships and historical spatial correlation challenge existing approaches.

Keywords Spatial correlation · Pampas · Precipitation · Temperature · Weather simulation

1 Introduction

Risk-based approaches are widely used in natural resources management such as water, land, crop, and ecology. Process-based models of these resources are driven with ensembles of input sequences, which are typically daily weather, resulting in ensembles of system variables and their probability density functions that provide estimates of risk that are useful for decision making. Historic data is often limited in space and time hence the risk estimates based solely on them do not accurately reflect the underlying variability. Therefore, robust generation of weather sequences that capture the underlying variability is essential. Generating random weather sequences that are statistically consistent with historical observations is known as stochastic weather generation.

Crop models for agriculture planning, hydrologic models for generating streamflow needed for water resources management, and erosion models for land erosion management (Wallis and Griffiths 1997; Richardson 1981; Richardson and Wright 1984; Wilks 1998; Wilks and Wilby 1999; Friend et al. 1997) have motivated the development of stochastic weather generators over the years. Traditional weather generators at a single location model the precipitation occurrence as a Markov Chain (Richardson 1981; Katz 1977; Stern and Coe 1984; Woolhiser 1992) or within a Poisson process framework (Foufoula-Georgiou and Georgakakos 1991; Furrer and Katz 2008). The daily rainfall amounts are modeled by fitting a gamma density function (Katz 1977; Buishand 1978; Yang et al. 2005; Furrer and Katz 2007). These models are traditionally estimated for each month or shorter to capture the seasonality. Conditioned on the rainfall state, temperatures are then simulated using autoregressive models (Richardson 1981).

Multi-site extensions of single-site weather generators can be unwieldy with large number of parameters to capture the statistics at each site and their spatial correlation (Mehrotra et al. 2006), more so with a large number of locations. Wilks (1998) proposed a multi-site precipitation

A. Verdin · B. Rajagopalan
Department of Civil, Environmental and Architectural Engineering, University of Colorado, Boulder, CO, USA
e-mail: andrew.verdin@colorado.edu

W. Kleiber (✉)
Department of Applied Mathematics, University of Colorado, Boulder, CO, USA
e-mail: william.kleiber@colorado.edu

R. W. Katz
Institute for Mathematics Applied to Geosciences, National Center for Atmospheric Research, Boulder, CO, USA

model with two-state Markov chain and mixed exponential distribution coupled with spatially correlated transformed normal variables to enable capturing the spatial correlation in precipitation. Later, Wilks (1999) extended this to multiple variables (temperature, solar radiation). Variations of this have been used in subsequent multi-site rainfall generators (Srikanthan and Pegram 2009; Brissette et al. 2007) and weather generators (Qian et al. 2002; Baigorria and Jones 2010; Khalili et al. 2009). Markov chains and direct acyclic graphs have been developed and proposed for stochastic multi-site rainfall simulation (Kim et al. 2008) as promising and less complex alternatives for space-time simulation. Bayesian hierarchical models for spatial rainfall have been developed in recent years (Lima and Lall 2009) which have the ability to provide robust estimation of uncertainties and are proving to be an attractive alternative with increase in computational power. Along this same line, Fassò and Finazzi (2011) offer a state of the art approach to space-time modeling using recent maximum likelihood advances based on the EM algorithm. For modeling sites with heavy-tailed precipitation distributions, a generalized Pareto distribution (Lennartsson et al. 2008) or a stretched exponential distribution (Furrer and Katz 2008) have been shown to be good alternatives to the more traditional methods.

Generalized linear models (GLMs) can greatly reduce the modeling effort of weather generators besides enabling the modeling of non-normal variables and being parsimonious (McCullagh and Nelder 1989). Herein, a GLM model (probit regression) is adopted for precipitation occurrence with a suite of covariates enabling the spatial modeling of occurrence with a single model—unlike a number of Markov chain models. A separate GLM is fitted to precipitation intensity, often using a gamma distribution and appropriate link function to capture non-Gaussian features. Early use of GLM for weather generation was by Stern and Coe (1984) with subsequent work by Yang et al. (2005) and Chandler (2005). Furrer and Katz (2007) developed this framework to include climate variables such as El Nino Southern Oscillation index for a location in the Pampas region of Argentina. Other methods to incorporate large scale climate information in weather generators include modeling the underlying climate process using a Hidden Markov Model and then conditionally generating stochastic weather sequences (Hauser and Demirov 2013). A limited extension of the GLM approach with a Poisson cluster model to multi-site precipitation generation was proposed by Wheater et al. (2005).

Semi-parametric approaches have been developed that resample historical data using an empirical distribution function—with precipitation occurrence modeled as a wet and dry process and seasonality addressed using Fourier components (Racsko et al. 1991; Semenov and Barrow 1997). These are relatively easy to implement and are widely used in climate change studies, especially in Europe (Calanca and Semenov 2013; Semenov and Calanca 2013). Multi-site extensions and enhancements to generate extremes have also been proposed in the above references and in (Semenov 2008). These weather generators have been shown to capture extreme events well over a region in New Zealand (Hashmi et al. 2011). In general, weather generators have difficulty capturing the properties of extreme events well—a formal way to enable this using extreme value distributions is proposed in Furrer and Katz (2008).

Nonparametric weather generators which make no assumption of the underlying distribution of the process and are data-driven have gained prominence in recent decades. Kernel density-based generators of precipitation (Lall and Sharma 1996; Harrold et al. 2003; Mehrotra and Sharma 2007) and other variables (Rajagopalan et al. 1997a, b) have been shown to perform very well at capturing non-normal and nonlinear features. Kernel methods perform poorly in high dimensions. To alleviate this, K-nearest neighbor (K-NN) time series bootstrap (Lall and Sharma 1996) based weather generators were developed (Rajagopalan and Lall 1999). In this, K-NNs of a weather vector on a current day are obtained from historical days within a window of the current day, and one of the neighbors (i.e., one of the historical days) is resampled with a weighted metric. The historical weather on the following day of the resampled neighbor becomes the simulated weather for the subsequent day. This is akin to resampling from a nonparametric estimation of the local conditional probability density function. For multi-site generation this is done on the site-averaged time series and the weather vector at all the locations of the selected day is taken to obtain multi-site simulation. As can be seen this is easy to implement and robust in capturing non-Gaussian features. This has been extended to multi-site and also has been conditioned on large scale climate information, climate forecasts, climate change projections, etc. (Yates et al. 2003; Apipattanavis et al. 2007; Buishand and Brandsma 2001; Beersma and Buishand 2003; Sharif and Burn 2007). Recently, Caraway et al. (2014) modified this approach for multi-site weather simulation by incorporating a cluster analysis wherein the sites are clustered and a single-site weather generator is applied to each cluster average. This modeling approach shows good performance in mountainous terrain.

One of the major drawbacks with the weather generators described above is their relative inability to generate weather sequences at any arbitrary locations, other than the locations with data. This is quite important for running hydrology, crop and ecology models which require weather sequences on a grid. It is in this context that the GLM-based methods offer a parsimonious and robust approach. Kleiber et al. (2012) extended this with latent Gaussian processes to model spatial occurrence and amounts of rainfall over the state of Iowa, US and the Pampas region of Argentina, and to temperature in complex terrain (Kleiber et al. 2013). Motivated by this drawback, in this paper we develop a GLM-based spatial weather generator which combines the precipitation and temperature generator of Kleiber et al. (2012) and Kleiber et al. (2013) and demonstrate it for application to the Pampas region. While we demonstrate the simulation of weather sequences at the locations with data, the model can generate sequences at any arbitrary location. The model, data and applications are described in the following sections.

2 Stochastic model

A basic full stochastic weather generator requires simultaneous simulation of minimum and maximum temperature, as well as precipitation, including both occurrence and intensity. The idea behind our approach is to condition the bivariate temperature process on precipitation occurrence. Although there is clearly a physical relationship between temperature and precipitation, precipitation largely occurs due to large scale atmospheric movement, while surface temperatures are highly controlled by local climate factors and by whether or not precipitation occurs. By maintaining a generalized linear modeling framework, we can straightforwardly condition temperature simulations on precipitation occurrence, thus allowing for distinct precipitation stochastic models to be used.

We follow the framework proposed by Kleiber et al. (2013) in focusing on the two components of local climate and weather. Local climate refers to the average behavior of a weather variable across time and space, while the weather component yields variability and individual realizations that deviate from climatology. For minimum and maximum temperatures at location $\boldsymbol{s} \in \mathbb{R}^2$ and day t, $Z_N(\boldsymbol{s}, t)$ and $Z_X(\boldsymbol{s}, t)$, respectively, we use the following decomposition,

$$Z_N(\boldsymbol{s}, t) = \beta_N(\boldsymbol{s})'\mathbf{X}_N(\boldsymbol{s}, t) + W_N(\boldsymbol{s}, t)$$
$$Z_X(\boldsymbol{s}, t) = \beta_X(\boldsymbol{s})'\mathbf{X}_X(\boldsymbol{s}, t) + W_X(\boldsymbol{s}, t).$$

The first component is a local regression on some covariate vector $\mathbf{X}_i(\boldsymbol{s}, t)$, while the weather component (denoted by W for weather) generates variability and spatial correlation via a multivariate normal Gaussian process. In our experience, temperature persistence is most straightforwardly accounted for by autoregressive terms in the mean function, and the weather component can then be viewed as temporally independent. It is worthwhile to note that the use of Gaussian models for temperature is justified for some climates [see Kleiber et al. (2013)]. Transformed variables can be used to produce stochastic realizations without assuming a Gaussian distribution, but these realizations have shown results that are consistent with this study.

The local climate component is a spatially varying coefficient model, where $\boldsymbol{\beta}_i(\boldsymbol{s}) = (\beta_{0i}(\boldsymbol{s}), \beta_{1i}(\boldsymbol{s}), \ldots, \beta_{pi}(\boldsymbol{s}))'$, for $i = N, X$ determines the influence of each covariate on temperature at a given location. For example, the intercept term $\beta_{0i}(\boldsymbol{s})$ accounts for the fact that, typically, temperatures at higher elevations tend to be lower than those at lower elevations or near oceans or seas. In our experience, it is useful and appropriate to include autoregressive terms in the covariate vectors $\mathbf{X}_i(\boldsymbol{s}, t) = (X_{0i}(\boldsymbol{s}, t), \ldots, X_{pi}(\boldsymbol{s}, t))'$.

Estimation of the coefficients $\boldsymbol{\beta}_i(\boldsymbol{s})$ rely on observations at a network of locations $\boldsymbol{s} = \boldsymbol{s}_1, \ldots, \boldsymbol{s}_n$ over a time period $t = 1, \ldots, T$ (note that an incomplete historical record does not affect estimation). A Bayesian approach would be to impose a prior distribution on the coefficients, viewing them as spatial processes, but for large networks of observation stations computations become infeasible using standard Bayesian techniques. Below, we allow the coefficients to vary with location, but suppress a stochastic representation.

The precipitation process is broken into two components: the occurrence at location $\boldsymbol{s}$ on day t, $O(\boldsymbol{s}, t)$, and the intensity or amount, $A(\boldsymbol{s}, t)$, given that there is some precipitation. In particular, we follow Kleiber et al. (2012) in modeling the occurrence process as a probit process,

$$O(\boldsymbol{s}, t) = \mathbb{1}_{[W_O(\boldsymbol{s}, t) \geq 0]}$$

where the latent process $W_O(\boldsymbol{s}, t)$ is Gaussian. If the latent process is positive, it rains at location $\boldsymbol{s}$, whereas if the process is negative, it does not rain. The latent process is given a mean function that is a regression on some covariates, $\boldsymbol{\beta}_O(\boldsymbol{s})'\mathbf{X}_O(\boldsymbol{s}, \mathbf{t})$, with spatially varying coefficients as in the temperature model. Realizations are spatially correlated by imposing a nontrivial covariance structure for $W_O(\boldsymbol{s}, t)$. Kleiber et al. (2012) used an exponential covariance function to model spatial correlation, for example. Briefly, the precipitation intensity process follows the same approach as Kleiber et al. (2012). In particular, the intensity at a particular location and time is modeled as a gamma random variable, whose scale and shape parameters vary with location and time. Simulations are spatially correlated by imposing a zero-mean Gaussian process $W_A(\boldsymbol{s}, t)$ with covariance function $C_A(\boldsymbol{h}, t)$, such that

$$A(\boldsymbol{s}, t) = G_{\boldsymbol{s},t}^{-1}(\Phi(W_A(\boldsymbol{s}, t)))$$

where $G_{\boldsymbol{s},t}$ is the cumulative distribution function (CDF) of the gamma distribution at site $\boldsymbol{s}$ and time t, and Φ is the CDF of a standard normal. This transformation approach is called a spatially varying anamorphosis function (Chilès and Delfiner 1999), which retains the gamma distribution at individual locations but allows for spatial correlation between locations. We do not explore this model in detail here, acknowledging that other precipitation models can be swapped in easily.

2.1 Estimation and modeling choices

We take a conditional approach to estimation by first gathering local estimates $\hat{\beta}_{ji}(\boldsymbol{s})$ by ordinary least squares at each observation location for both minimum and maximum temperature. Spatial covariance often exhibits seasonal patterns, where, for example, temperatures tend to exhibit greater variability in summer than in winter. Additionally, length scale of spatial correlation can also vary across time. To account for these nonstationarities, we estimate the nonparametric spatial covariance structure for $W_i(\boldsymbol{s}, t)$ on a monthly basis, using the empirical covariance matrix of the residuals at the observation network. For each day t and spatial location s, we form the residuals $W_i(\boldsymbol{s}, t) = Z_i(\boldsymbol{s}, t)\beta_i(\boldsymbol{s})'X(\boldsymbol{s}, t)$, where $\beta_i(\boldsymbol{s})$ is the least squares estimate. These residuals are then assumed to be realizations from the W_i process, and from these values we form the empirical covariance matrix. Estimation for precipitation occurrence follows a similar strategy; we first estimate local mean coefficients $\hat{\beta}_{jO}(\boldsymbol{s})$ by probit regression, using all historical occurrence observations. The spatial structure for the latent process is estimated as the empirical correlation based on the probit model errors at all network stations, using occurrences as observations separately for each month.

For our coupled weather generator, we judiciously choose a multivariate autoregressive structure on the temperature process, conditional on precipitation occurrence. In particular, we set

$$\begin{aligned}\mathbf{X}_N(\boldsymbol{s}, t) = (&1, \cos(2\pi t/365), \sin(2\pi t/365),\\ &r(t), Z_N(\boldsymbol{s}, t-1), Z_X(\boldsymbol{s}, t-1), O(\boldsymbol{s}, t))'.\end{aligned}$$

The first three entries are an intercept and two harmonics to account for seasonal trends; $r(t)$ is a linear drift between -1 and 1 (for numerical stability), which we include to control for temperature trends over the period of our data set; the latter three entries imply a trivariate autoregressive structure. Note that temperature is conditioned on the coincidental occurrence; in practice this usually implies cooler temperatures on rainy days and warmer temperature on dry days. The same covariates are used for the maximum temperature process. The precipitation occurrence process is given the following covariates,

$$\mathbf{X}_O(\boldsymbol{s}, t) = (1, \cos(2\pi t/365), \sin(2\pi t/365), O(\boldsymbol{s}, t-1))',$$

where now precipitation uses a single autoregressive model. Note the linear drift is not included in the precipitation model because precipitation tends to exhibit epoch-like traits rather than gradual trends.

Simulation can proceed by simulating an entire trajectory of precipitation occurrence, with amounts if required for scientific purposes. Conditional on this realization, an initial temperature is chosen (we use the average of that calendar day's observations), and daily realizations are then available by simulating the weather component as multivariate normals, and adding the weather to the local climate.

3 Stochastic weather simulation in the Pampas

To illustrate the performance and capability of the proposed coupled model, we consider a dataset of minimum temperature, maximum temperature and precipitation observations at a network of 19 locations in the Pampas region of Argentina, shown in Fig. 1. The Pampas region covers much of northeastern Argentina, all of Uruguay, and very little of southeastern Brazil—covering more than 750,000 km^2—and is of utmost agricultural importance for much of the South American continent. With global food prices on the rise, the ability to quantify and forecast climate variability for the purposes of climate change impact assessment in the region is absolutely necessary. Observations are available over approximately an 80 year period, although the longest station record is from 1908 to 2010. Spatial precipitation simulation was previously explored by

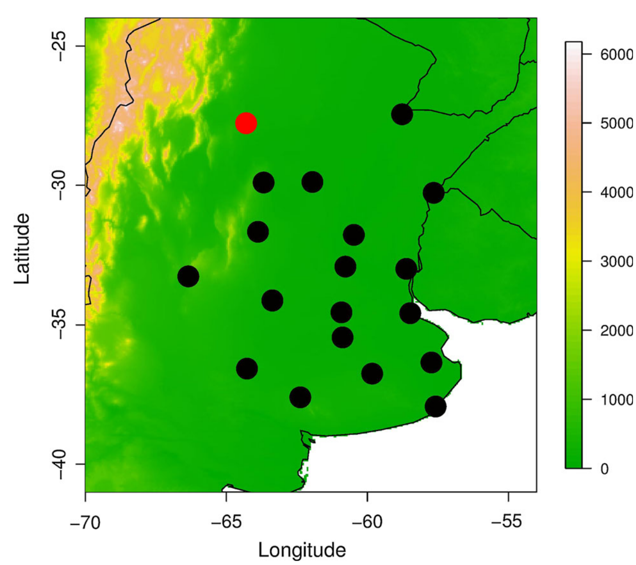


Fig. 1 Study region geography with elevation (in meters), *black dots* are observation station locations; *red dot* represents Santiago del Estero

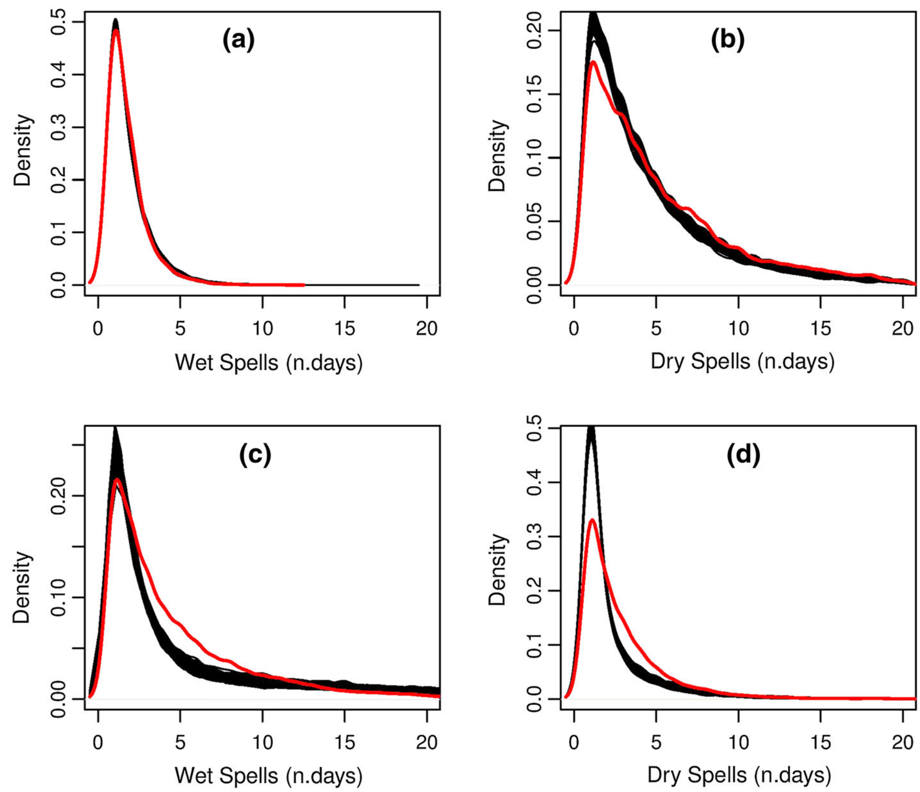


Fig. 2 **a**, **b** Densities for simulated local wet and dry spells at Santiago del Estero, **c**, **d** regional dry and wet spells; observed densities shown in *red*

Kleiber et al. (2012) on this dataset, but these authors did not consider the temperature observations.

We adopt the model setup as outlined in Sect. 2, and estimate local regression coefficients (linear regression for temperature, probit regression for precipitation occurrence). Conditional on these estimates, the spatial covariance structure is estimated empirically. In our model, temperature simulations are conditional on precipitation occurrence, but the dependence is one-directional. Thus, we begin by simulating precipitation occurrence each day over the 19 locations throughout the Pampas. We simulate 100 trajectories of occurrence independently, thus producing an ensemble of daily weather patterns. To consistently compare simulations to observations, we necessitate that output from the coupled weather generator be masked to match the pattern of missing values from the observed precipitation and temperature time series.

Validation of the precipitation occurrence model is carried out through spells analysis—local and regional wet and dry spells. A regional dry spell occurs when all 19 locations report no rain—occurrence at any location breaks the regional dry spell. Figure 2 shows the density of spells from the 100 trajectories as well as that from the observed time series.

As can be seen in Fig. 2, both wet and dry spells are reproduced with good skill by the trajectories at Santiago del Estero. Simulated wet spells are very nearly perfect, while there is slight discrepancy in the density occurrence of dry spells—simulations are producing systematically shorter dry spells than the observations. Due to this underrepresentation of local dry spells, there is even greater discrepancy between observed and simulated regional dry spells. The trajectories imply that regional dry spells are very rarely longer than one day, while observations show they have a better chance of lasting at least three days. However, the frequency of longer domain dry spells is adequately reproduced, particularly above spells of eight or more days. To illustrate the validity of this model in reproducing the frequency of long dry spells, two dry spells analysis were carried out for an independent station. We simulated weather for a station not included in the model—thus validating the ability of this model to simulate weather at any arbitrary location—and analyzed the ability to reproduce long dry spells. These long dry spells are crucial to capture for impact assessment planning. As can be seen in Fig. 3, the model is quite impressive in its ability to reproduce the frequency and longevity of dry spells, especially considering this station was not included in the model fit.

Allowing the coupled relationship between temperature and precipitation to vary with location is important over large domains, such as in the Pampas. Figure 4a illustrates the relationship that precipitation occurrence has with minimum temperature. It can be seen there is little spatial

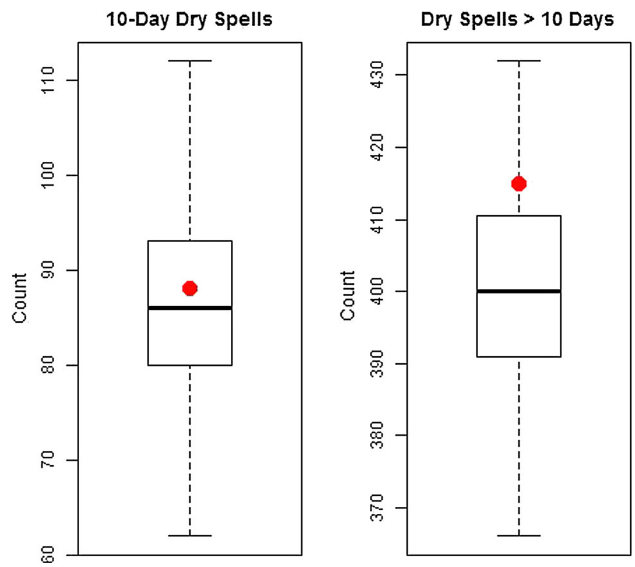


Fig. 3 Dry spells validation for Pergamino, a station not included in the model. *Boxplots* show the count of dry spells equal to (*left*) or greater than (*right*) ten days for the 100 realizations. *Red dots* represent the count of these dry spells for the historical data

structure relating these processes, implying that precipitation occurrence is not the only contributing factor in simulating minimum temperatures in the region. Note that the coefficient on occurrence is generally positive, indicating that the presence of precipitation implies that minimum temperatures tend to be between 1–2 °C warmer. Conversely, Fig. 4b shows that there is a much stronger spatial structure relating precipitation occurrence and maximum temperature. Indeed, inland maximum temperatures tend to be reduced with the presence of precipitation, while the maximum temperature at locations near the ocean are less affected by precipitation.

Capturing the spatial coherence of daily weather patterns is of utmost importance in producing realistic weather generator output. To this end, pairwise correlations are considered for minimum and maximum temperatures at all 19 locations, producing $\frac{19 \times 18}{2} = 171$ pairs, as can be seen in Fig. 5. These pairwise correlations are somewhat consistent between the simulations and historical observations, although there is evidence of slightly reduced model correlation, on the order of 5 %.

Figure 5 illustrates that minimum and maximum temperatures are positively spatially correlated, in that neighboring locations have similar daily temperature patterns. It follows to analyze the output from the coupled temperature models further, thus assessing its ability to reproduce cold and hot spells. A cold spell is defined as the number of consecutive days that the minimum temperature at a location is less than 5 °C. Similarly, a hot spell is defined as the number of consecutive days that the maximum temperature at a location exceeds 30 °C. A regional cold spell occurs when the minimum temperatures at all 19 locations are less than 5 °C. For a regional hot spell, the maximum temperatures at all 19 locations must exceed 30 °C. For consistency, and because simulated precipitation occurrence is used as input in the minimum and maximum temperature models, local cold and hot spells are

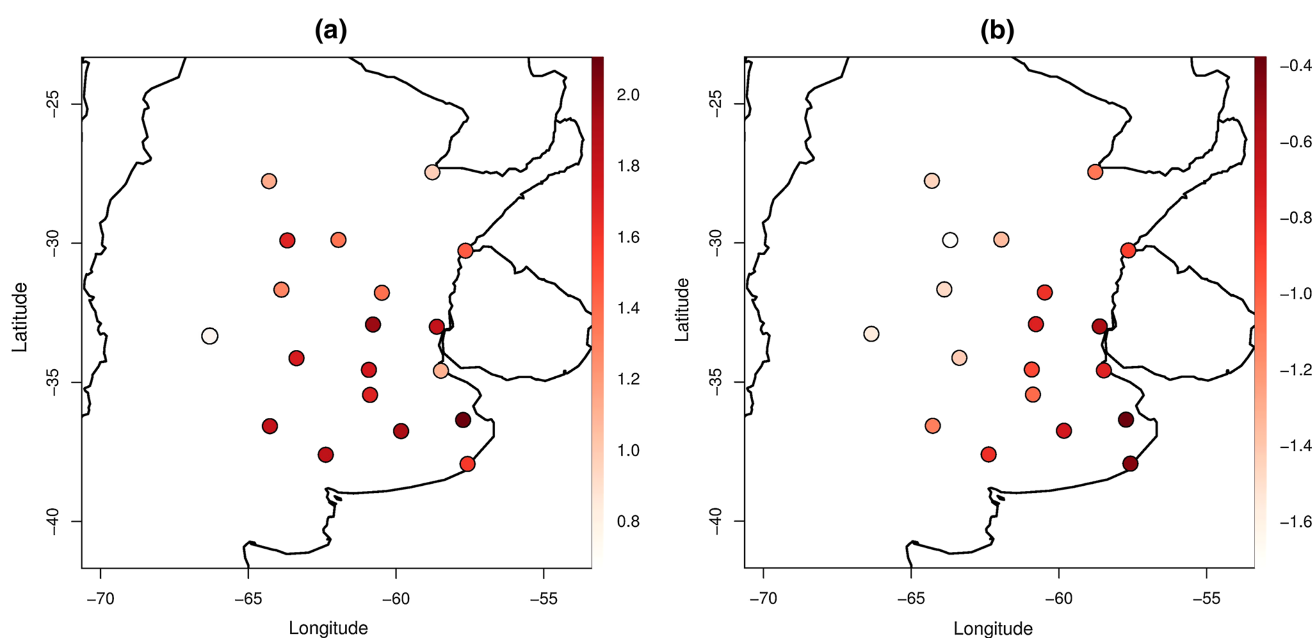


Fig. 4 Precipitation occurrence coefficient for **a** minimum temperature and **b** maximum temperature

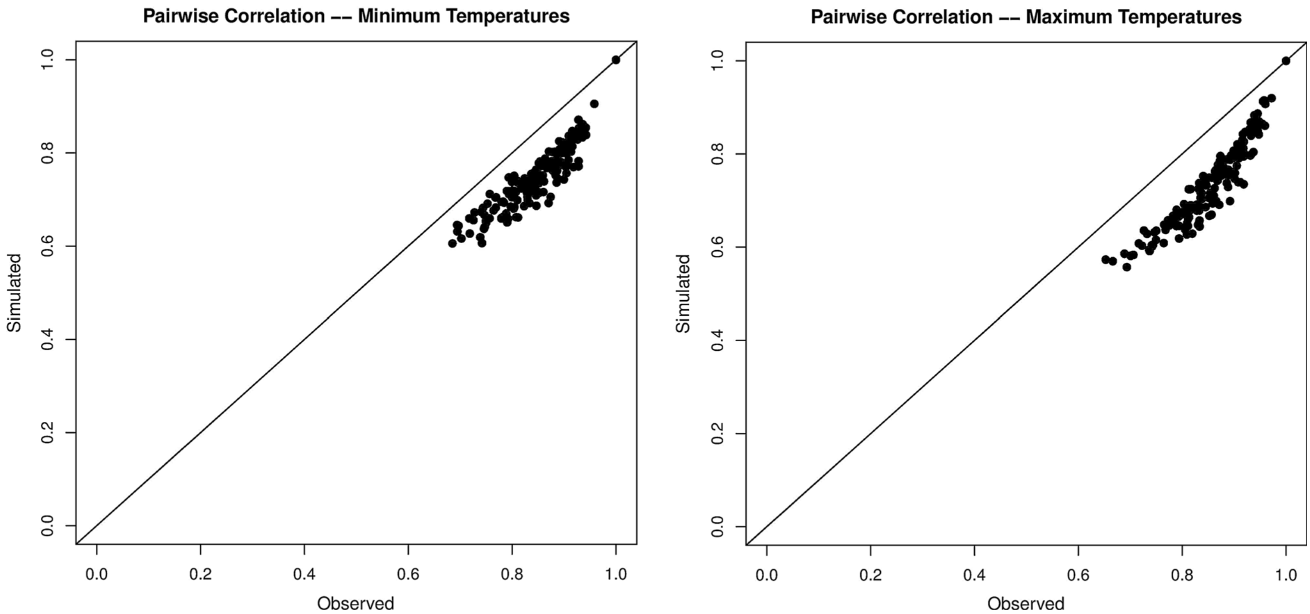
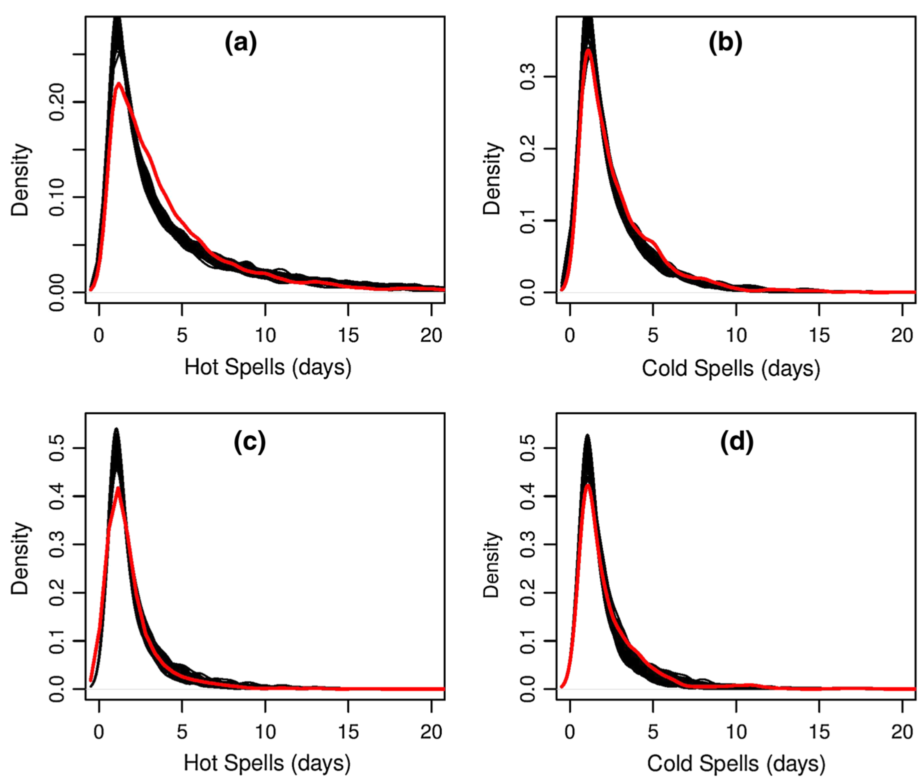


Fig. 5 Observed versus simulated pairwise correlations for minimum and maximum temperatures

Fig. 6 **a**, **b** Densities for simulated local cold and hot spells at Santiago del Estero, **c**, **d** regional cold and hot spells. Observed density shown in *red*



analyzed for Santiago del Estero, Argentina. Figure 6a–b shows that the 100 trajectories reproduce the density of cold and hot spells with good skill. The shape and peak of simulated cold spells are nearly perfect, while those of hot spells show slight discrepancy with respect to observations. In Fig. 6c–d, it can be seen there is not systematic underrepresentation of local hot spells—as was the case for local dry spells at this location—as the simulated and observed densities of regional cold and hot spells are reproduced with very good skill.

The space-time aspect of this stochastic weather generator may be examined by obtaining the covariance of

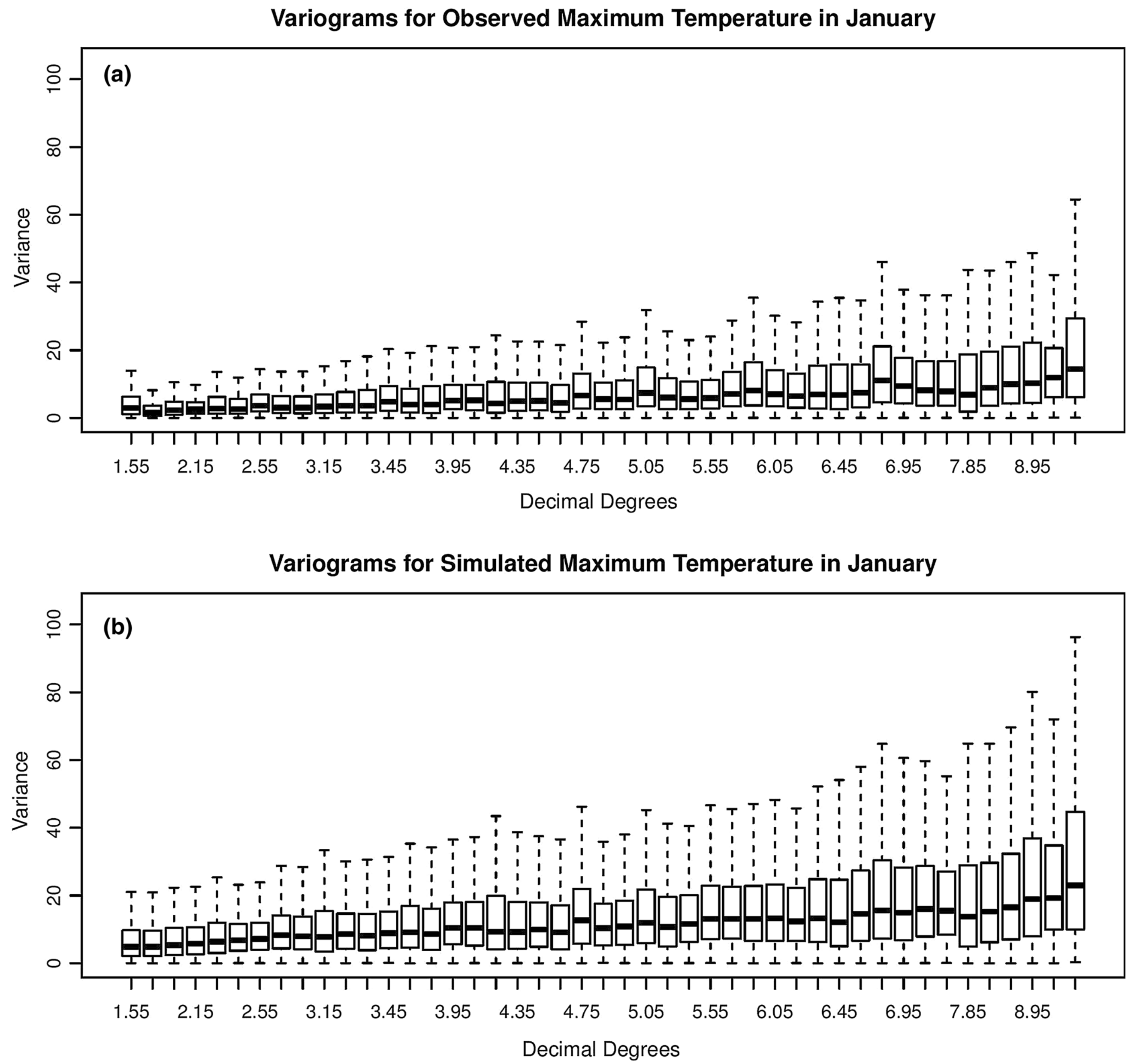


Fig. 7 Variograms for **a** observed and **b** simulated daily January maximum temperatures

daily weather on a monthly scale, resulting in an ensemble of empirical variograms per month. It is assumed the monthly covariance is isotropic, such that weather patterns vary on the same scale in all directions to a given lag. For a given month, the daily weather is used to obtain an empirical variogram; the process is repeated for all days within the given month with non-missing data, and the ensemble of empirical variograms may be visualized as boxplots. Ensemble variograms of daily maximum temperature for January are shown in Fig. 7. Note that Fig. 7a shows the ensemble of empirical variograms for observed maximum temperature, while Fig. 7b shows that for a randomly-selected trajectory. We see that the spatial correlation structure based on model realizations are similar to those observed in the historical data, indicating that the model adequately captures the spatial behavior as well as the local behavior of temperature.

4 Discussion

We have introduced a conditional approach to daily space-time stochastic weather simulation wherein temperature is conditioned on precipitation occurrence. Although this idea has previously been explored, we endow the model with additional flexibility by allowing model coefficients to vary with location within a local climate framework. Simulations are correlated via Gaussian process residual terms, yielding spatially and temporally consistent realizations.

In this manuscript we have focused on a spells analysis for model output assessment, as the purpose of this research is to provide a stochastic weather generator that generates daily weather ensembles where the minimum and maximum temperatures are conditioned on precipitation occurrence. The precipitation intensities have not been validated nor reported on because they are produced from the same technique as in Kleiber et al. (2012). Validation of the model's ability to simulate weather at any arbitrary location was validated by producing statistically consistent simulations at an independent station. Not only were short spells reproduced with near perfection, the longevity and frequency of dry spells were maintained throughout all 100 realizations. The ability to reproduce long dry spells is crucial for impact assessment planning.

Future work will consider gridded simulation, the output of which can be used to run hydrologic models for decision support. The authors have also considered a more sophisticated relationship between temperature and precipitation, as well as the introduction of other important variables such as solar radiation.

Acknowledgments Research partially supported by NSF EaSM grant 1049109. Thanks to Guillermo Podesta for providing daily weather data for Argentine Pampas.

References

Apipattanavis S, Podestá G, Rajagopalan B, Katz RW (2007) A semiparametric multivariate and multisite weather generator. Water Resour Res 43:1–19. doi:10.1029/2006WR005714

Baigorria GA, Jones JW (2010) GiST: a stochastic model for generating spatially and temporally correlated daily rainfall data. J Clim 23(22):5990–6008

Beersma JJ, Buishand TA (2003) Multi-site simulation of daily precipitation and temperature conditional on the atmospheric circulation. Clim Res 25(2):121–133

Brissette FP, Khalili M, Leconte R (2007) Efficient stochastic generation of multi-site synthetic precipitation data. J Hydrol 345:121–133

Buishand T (1978) Some remarks on the use of daily rainfall models. J Hydrol 36(3):295–308

Buishand TA, Brandsma T (2001) Multisite simulation of daily precipitation and temperature in the Rhine basin by nearest-neighbor resampling. Water Resour Res 37(11):2761–2776

Calanca P, Semenov MA (2013) Local-scale climate scenarios for impact studies and risk assessments: integration of early 21st century ENSEMBLES projections into the ELPIS database. Theor Appl Climatol 113:445–455

Caraway NM, McCreight JL, Rajagopalan B (2014) Multisite stochastic weather generation using cluster analysis and k-nearest neighbor time series resampling. J Hydrol 508:197–213

Chandler RE (2005) On the use of generalized linear models for interpreting climate variability. Environmetrics 16:699–715

Chilès JP, Delfiner P (1999) Geostatistics: modeling spatial uncertainty. Wiley, New York

Fassò A, Finazzi F (2011) Maximum likelihood estimation of the dynamic coregionalization model with heterotopic data. Environmetrics 22(6):735–748

Foufoula-Georgiou E, Georgakakos KP (1991) Hydrologic advances in space-time precipitation modeling and forecasting. In: Bowles DS, O'Connell (eds) Recent advances in the modeling of hydrologic systems. Springer, Netherlands, pp 47–65

Friend AD, Stevens AK, Knox RG, Cannell MGR (1997) A process-based terrestrial biosphere model of ecosystem dynamics. Ecol Model 95:249–287

Furrer EM, Katz RW (2007) Generalized linear modeling approach to stochastic weather generators. Clim Res 34:129–144

Furrer EM, Katz RW (2008) Improving the simulation of extreme precipitation events by stochastic weather generators. Water Resour Res 44:1–13. doi:10.1029/2008WR007316

Harrold TI, Sharma A, Sheather SJ (2003) A nonparametric model for stochastic generation of daily rainfall amounts. Water Resour Res 39(12):1–11

Hashmi MZ, Shamseldin AY, Melville BW (2011) Comparison of SDSM and LARS-WG for simulation and downscaling of extreme precipitation events in a watershed. Stoch Environ Res Risk Assess 25(4):475–484

Hauser T, Demirov E (2013) Development of a stochastic weather generator for the sub-polar North Atlantic. Stoch Environ Res Risk Assess 27(7):1533–1551

Katz RW (1977) Precipitation as a chain-dependent process. J Appl Meteorol 16:671–676

Khalili M, Brissette F, Leconte R (2009) Stochastic multi-site generation of daily weather data. Stoch Environ Res Risk Assess 23(6):837–849

Kim T-W, Ahn H, Chung G, Yoo C (2008) Stochastic multi-site generation of daily rainfall occurrence in south Florida. Stoch Environ Res Risk Assess 22(6):705–717

Kleiber W, Katz RW, Rajagopalan B (2012) Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. Water Resour Res 48:1–17. doi:10.1029/2011WR011105

Kleiber W, Katz RW, Rajagopalan B (2013) Daily minimum and maximum temperature simulation over complex terrain. Ann Appl Stat 7:588–612

Lall U, Sharma A (1996) A nearest neighbor bootstrap for resampling hydrological time series. Water Resour Res 32:679–693

Lennartsson J, Baxevani A, Chen D (2008) Modelling precipitation in Sweden using multiple step Markov chains and a composite model. J Hydrol 363(1–4):42–59. URL http://linkinghub.elsevier.com/retrieve/pii/S0022169408004848

Lima CHR, Lall U (2009) Hierarchical Bayesian modeling of multisite daily rainfall occurrence: rainy season onset, peak and end. Water Resour Res 45:1–14. doi:10.1029/2008WR007485

McCullagh P, Nelder JA (1989) Generalized linear models. Chapman and Hall, London

Mehrotra R, Sharma A (2007) A semi-parametric model for stochastic generation of multi-site daily rainfall exhibiting low-frequency variability. J Hydrol 335:180–193

Mehrotra R, Srikanthan R, Sharma A (2006) A comparison of three stochastic multi-site precipitation occurrence generators. J Hydrol 331:280–292

Qian B, Corte-Real J, Xu H (2002) Multisite stochastic weather models for impact studies. Int J Climatol 2002:1377–1397

Racsko P, Szeidl L, Semenov M (1991) A serial approach to local stochastic weather models. Ecol Model 57:27–41

Rajagopalan B, Lall U (1999) A k-nearest neighbor simulator for daily precipitation and other weather variables. Water Res Res 35(10):3089–3101

Rajagopalan B, Lall U, Tarboton DG (1997a) Evaluation of kernel density estimation methods for daily precipitation resampling. Stoch Hydrol Hydraul 11(6):523–547

Rajagopalan B, Lall U, Tarboton DG, Bowles DS (1997b) Multivariate nonparametric resampling scheme for generation of daily weather variables. Stoch Hydrol Hydraul 11:523–547

Richardson CW (1981) Stochastic simulation of daily precipitation, temperature, and solar radiation. Water Resour Res 17(1): 182–190

Richardson CW, Wright DA (1984) WGEN: a model for generating daily weather variables. ARS (USA) 1–83

Semenov MA (2008) Simulation of extreme weather events by a stochastic weather generator. Clim Res 35:203–212

Semenov MA, Barrow EM (1997) Use of a stochastic weather generator in the development of climate change scenarios. Clim Change 35:397–414

Semenov P-BS, Pilkington-Bennett S, Calanca P (2013) Validation of ELPIS 1980–2010 baseline scenarios using the European Climate Assessment observed dataset. Clim Res 51:1–9

Sharif M, Burn DH (2007) Improved k-nearest neighbor weather generating model. J Hydrol Eng 12(1):42–51

Srikanthan R, Pegram GGS (2009) A nested multisite daily rainfall stochastic generation model. J Hydrol 371:142–153

Stern RD, Coe R (1984) A model fitting analysis of daily rainfall data. J R Stat Soc Ser A Gen 147:1–34

Wallis TW, Griffiths JF (1997) Simulated meteorological input for agricultural models. Agric Forest Meteorol 88:241–258

Wheater H, Chandler R, Onof C, Isham V, Bellone E, Yang C, Lekkas D, Lourmas G, Segond M-L (2005) Spatial-temporal rainfall modelling for flood risk estimation. Stoch Environ Res Risk Assess 19(6):403–416

Wilks DS (1998) Multisite generalization of a daily stochastic precipitation generation model. J Hydrol 210:178–191

Wilks DS (1999) Simultaneous stochastic simulation of daily precipitation, temperature and solar radiation at multiple sites in complex terrain. Agric Forest Meteorol 96:85–101

Wilks DS, Wilby RL (1999) The weather generation game: a review of stochastic weather models. Prog Phys Geogr 23:329–357

Woolhiser DA (1992) Modeling daily precipitation: progress and problems. Stat Environ Earth Sci 5:71–89

Yang C, Chandler RE, Isham VS, Wheater HS (2005) Spatial-temporal rainfall simulation using generalized linear models. Water Resour Res 41:1–13. doi:10.1029/2004WR003739

Yates D, Gangopadhyay S, Rajagopalan B, Strzepek K (2003) A technique for generating regional climate scenarios using a nearest-neighbor algorithm. Water Resour Res 39:1199