



A conditional stochastic weather generator for seasonal to multi-decadal simulations



Andrew Verdin^{a,*}, Balaji Rajagopalan^{a,b}, William Kleiber^c, Guillermo Podestá^d, Federico Bert^e

^a Dept of Civil, Environmental, and Architectural Engineering, University of Colorado, Boulder, CO, United States

^b Cooperative Institute for Research in Environmental Sciences, University of Colorado, Boulder, CO, United States

^c Dept of Applied Mathematics, University of Colorado, Boulder, CO, United States

^d School of Marine & Atmospheric Sciences, University of Miami, Miami, FL, United States

^e Facultad de Agronomía, Universidad de Buenos Aires – CONICET, Buenos Aires, Argentina

ARTICLE INFO

Article history:

Available online 23 December 2015

Keywords:

Generalized linear models
Stochastic weather generator
Conditional simulation
Downscaling seasonal forecasts
Daily precipitation
Daily temperature

SUMMARY

We present the application of a parametric stochastic weather generator within a nonstationary context, enabling simulations of weather sequences conditioned on interannual and multi-decadal trends. The generalized linear model framework of the weather generator allows any number of covariates to be included, such as large-scale climate indices, local climate information, seasonal precipitation and temperature, among others. Here we focus on the Salado A basin of the Argentine Pampas as a case study, but the methodology is portable to any region. We include domain-averaged (e.g., areal) seasonal total precipitation and mean maximum and minimum temperatures as covariates for conditional simulation. Areal covariates are motivated by a principal component analysis that indicates the seasonal spatial average is the dominant mode of variability across the domain. We find this modification to be effective in capturing the nonstationarity prevalent in interseasonal precipitation and temperature data. We further illustrate the ability of this weather generator to act as a spatiotemporal downscaler of seasonal forecasts and multidecadal projections, both of which are generally of coarse resolution.

© 2015 Elsevier B.V. All rights reserved.

1. Introduction

Scientific and technological advances, together with awareness of the importance of climate on human endeavors, are creating increased worldwide demand for climate information. Fortunately, our ability to monitor and predict variations in climate has increased substantially (Barnston et al., 2010; Stockdale et al., 2010). A number of groups now forecast climate conditions a few seasons ahead (Goddard et al., 2003; Saha et al., 2006). Emerging developments may enable climate projections 10–20 years into the future, a scale intermediate between seasonal forecasts and manmade climate change projections (Haines et al., 2009; Hurrell et al., 2009; Meehl et al., 2009). These advances, however, must be matched by a better understanding of how science can inform climate-resilient planning and development (Stainforth et al., 2007).

To support public and private adaptation and mitigation responses, climate information must be credible, legitimate and, especially, salient – e.g., relevant to the needs of decision makers

(Cash et al., 2003). Needs include not only predictions or projections¹ (Bray and von Storch, 2009) of regional climate: *potential outcomes of adaptation actions are probably more relevant to stakeholders than raw climate information*. Thus, an enhanced capacity is needed to “translate” climate information into distributions of outcomes for risk assessment and management (Hansen et al., 2006).

Process models (e.g., crop biophysical models, hydrological models) can be useful tools to assess likely impacts on climate-sensitive sectors of society, and to evaluate the outcomes of alternative adaptive actions (Ferreira et al., 2001; Berger, 2001; Berger et al., 2006; Happe et al., 2008; Freeman et al., 2009; Schreinemachers and Berger, 2011; Bert et al., 2006, 2007, 2014). These models, however, typically require daily weather data. Although historical daily weather can be used, getting long-term daily weather is laborious and costly at best and, in some cases, impossible. Typically, historical observations have missing data

¹ Following Bray and von Storch (2009), **prediction** conveys a sense of certainty whereas **projection** is associated more with the possibility of something happening given a certain set of plausible, but not necessarily probable, circumstances. A prediction can be used to design specific response strategies, while a projection, or more precisely a series of projections, provides a range on which to consider a range of response strategies.

* Corresponding author.

E-mail address: andrew.verdin@colorado.edu (A. Verdin).

that are not accepted by impact models. Similarly, point measurements may not represent the true spatial variability of a nonstationary natural process (e.g., daily precipitation). Most importantly, observed sequences provide a solution based on only one realization of the weather process (Richardson, 1981).

The use of seasonal forecasts of regional climate and its impacts can help decision-makers to lessen the adverse effects of unfavorable conditions or, alternatively, to capitalize on favorable conditions. Nevertheless, a major obstacle to broader use of seasonal climate forecasts is their coarse spatial and temporal resolution. Similarly, 10–20 year projections of regional climate conditions have been identified as important to infrastructure planners, water resource managers, and many others (Hurrell et al., 2009). Unfortunately, projections of regional monthly precipitation and temperature from climate models not only are coarse in space and time – as seasonal forecasts – but also involve considerable uncertainty, which requires exploration of the impacts of alternative, plausible trajectories. Stochastic weather generators have long been used for risk assessment and adaptation, as they can provide a rich variety of plausible climatic scenarios. Moreover, weather generators can produce spatially consistent series that can be used to downscale larger-scale scenarios.

Traditional weather generators (stemming from Richardson, 1981) model precipitation occurrence as a chain-dependent process (Katz, 1977) and thus are capable of generating physically realistic prolonged wet and dry spells. The remaining weather variables (e.g., precipitation intensity and temperature) are parameterized using probability distributions (for precipitation intensity) and linear time series models (for temperature), which capture historical climatological variability and linear relationships between variables but fail to capture extremes (e.g., extreme drought or flooding). In order to capture the variability of weather attributes in any specific season, the simulations need to be conditioned on appropriate covariates. One approach is to estimate the parameters of the generator conditionally by considering ENSO (El Niño Southern Oscillation; Trenberth and Stepaniak, 2001) phase, or any other teleconnection to a region's climate, which enables simulation of skillful sequences (Grondona et al., 2000; Ferreyra et al., 2001; Wilby et al., 2002; Meza, 2005; Katz et al., 2002). Wilks (2008) illustrated the capability of interpolating weather generator parameters to arbitrary locations (e.g., on a grid) using local weighted regressions; Wilks (2009) subsequently offered a method to synchronize gridded synthetic weather series on observed weather data. Approaches to producing weather sequences that deviate from climatology have included the implementation of seasonal correction factors, perturbation of parameters or input data, and spectral approaches (Caron et al., 2008; Kilsby et al., 2007; Hansen and Mavromatis, 2001; Schoof et al., 2005; Qian et al., 2010).

Nonparametric weather generators have an improved ability to capture nonlinearities between variables and sites. Included in this subclass are the k-nearest neighbor (k-NN) bootstrap resampling method (Brandsma and Buishand, 1998; Rajagopalan and Lall, 1999; Buishand and Brandsma, 2001; Beersma and Buishand, 2003; Yates et al., 2003; Sharif and Burn, 2007) and kernel density based estimators (Rajagopalan et al., 1997; Harrold et al., 2003; Mehrotra and Sharma, 2007). Caraway et al. (2014) first applied a clustering algorithm to identify regions of similar climatology before applying the k-NN approach, which has shown good performance in regions of complex terrain. Apipattanavis et al. (2010) modified the k-NN approach to create a semi-parametric weather generator that better captures the duration of wet and dry spells via Markov chain modeling. Modifications of the k-NN based weather generator to incorporate seasonal precipitation forecasts (Apipattanavis et al., 2010) and multi-decadal projections (Podestá et al., 2009) have also been proposed. In these situations,

the resampling is weighted to reflect the projected distribution of regional climate conditions. These methods are simple and powerful, however their main drawback is that they cannot generate values outside the range of historical data. More importantly, it is not easy to generate weather sequences at locations other than those with historical observations.

Pioneered by Stern and Coe (1984), generalized linear models (GLMs) are able to straightforwardly model non-normal data through a suite of link functions. Relevant to this research, GLMs can be used to model and simulate daily weather sequences, and have paved the way for generating space–time weather sequences at any desired location (Kleiber et al., 2012, 2013; Furrer and Katz, 2007; Kim et al., 2012; Yan et al., 2002; Yang et al., 2005; Chandler, 2005; Verdin et al., 2015). Recently Verdin et al. (2015) incorporated these developments into a robust space–time weather generator and demonstrated its capability to generate realistic weather sequences at arbitrary locations in the Pampas of Argentina – also the region targeted by this paper. The GLM framework offers several advantages – mainly they reduce the effort in modeling non-normal variables and are parsimonious (McCullagh and Nelder, 1989), especially for discrete and skewed variables (e.g., precipitation occurrence and intensity, respectively). Coupled with spatial processes, GLMs can generate sequences at any spatial resolution – which is important for resource management. Furthermore, covariates such as ENSO information, seasonal climate forecasts, and annual cycles can easily be incorporated in the GLMs to refine or narrow the distribution of expected values (e.g., Chandler and Wheeler, 2002; Wheeler et al., 2005; Furrer and Katz, 2007; Kim et al., 2012).

As motivated earlier in this section, skillful and realistic sequences of daily weather in any given season are essential for efficient planning and management of agricultural resources. One method of obtaining such sequences requires generating space–time weather sequences that are consistent with, and conditioned on, coarse climate information from seasonal to decadal time scales. To this end, here we propose a modification to the stochastic weather generator presented in Verdin et al. (2015) to include the coarse scale information as covariates. We refer to the weather generator of Verdin et al. (2015) as “original”; that of this research will be called the “modified” weather generator. The paper is organized as follows: the study region and data are described in Section 2; Section 3 contains a brief summary of the modified methodology. In Section 4 we discuss the results, and in Section 5 we conclude with a summary of the research and future work.

2. Study region and data

Application of this methodology is focused on a network of seventeen weather stations located in and around the Salado A basin of the Pampas of Argentina (see Fig. 1). The Salado is part of the large Río de la Plata basin (Herzer, 2003). Note the study region differs from that of Verdin et al. (2015).

The A basin is an agriculturally productive sub-basin within the Salado River basin where maize, soybean, and wheat are grown. The Salado Basin has very flat topography and a poorly developed and disintegrated drainage system. The western basin (Salado A) includes mega-parabolic dunes separated by depressions that constrain evacuation of surface water (Aragón et al., 2010; Viglizzo et al., 2009, 1997). Since colonial times, the Salado has shown alternating floods and droughts that displace populations and disrupt productive activities and livelihoods for extended periods. Floods were frequent during the late 19th and early 20th centuries, a relatively wet epoch. In contrast, extensive droughts were more frequent during the drier 1930s–1950s (Herzer, 2003; Seager et al., 2010). Partly in response to rain increases since the 1970s, severe

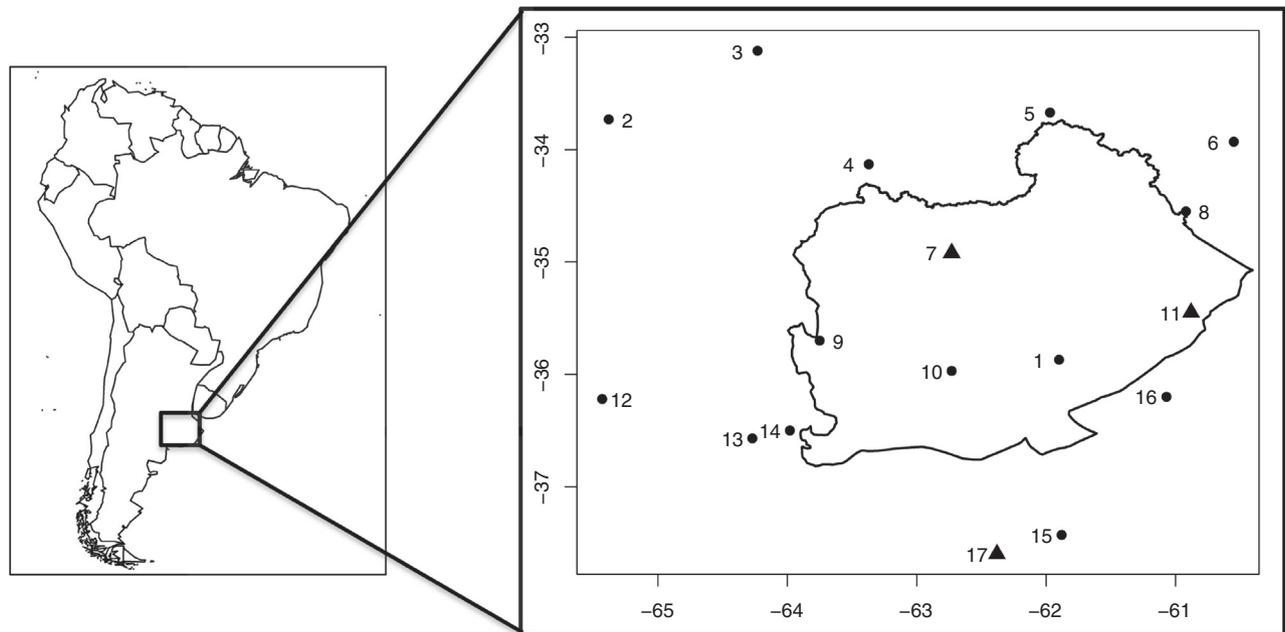


Fig. 1. Study region: weather stations shown as dots, numbers correspond to Table 1. The Salado 'A' basin domain is outlined. Three stations withheld in spatial validation shown as triangles.

floods have occurred in the Salado Basin in 1980, 1991–93, and 2000–01 (Herzer, 2003). Floods in the western half of the Pampas between 1997 and 2003 left 27% of the landscape under water, halved grain production, damaged infrastructure and soil quality, and transformed the few remaining natural areas (Viglizzo et al., 2009). In contrast, an almost unprecedented drought in 2008 (Skansi et al., 2009) decreased soybean and wheat production in the region by about 30% and 50% respectively.

We apply the proposed methodology on a sub-basin scale to illustrate its ability in downscaling coarse seasonal (multi-decadal) forecasts (projections) to local daily weather patterns while maintaining physically realistic climatic characteristics. As agriculture in the Pampas is entirely rainfed, it is of interest to provide a robust risk assessment for crop yields in this region.

During the last half of the 20th century, the study region experienced one of the most significant positive trends in annual precipitation amounts in the world (Giorgi, 2002). This overall increase in precipitation partly contributed to immense agricultural expansion to the semi-arid regions of the western Pampas (Bert et al., 2014). Since the turn of the 21st century, however, observed conditions suggest a significant decrease in regional annual precipitation, which begs the question: “Are the existing agricultural production systems viable in a drier future?” Analysis of a system’s response to an ensemble of possible futures that exhibit significant fluctuations in annual precipitation on the multi-decadal scale is of utmost importance for production risk analysis in climatically marginal regions such as the western Pampas.

Daily time series of precipitation, minimum temperature, and maximum temperature are available for a network of 17 weather stations from 1 January 1961 to near present (in this research we use data up to 31 December 2013). This data was collected and organized by associates at the Servicio Meteorológico Nacional (National Meteorological Service) of Buenos Aires, Argentina, and extensive quality control was carried out to ensure its validity. While there is a significant longitudinal gradient in precipitation and temperature (800 mm/year precipitation and 24 °C maximum temperature in the west, 1000 mm/year precipitation and 20 °C maximum temperature in east), the climatic tendencies

(e.g., trends) are similar between all weather stations, thus the A sub-basin serves as an optimal test bed for this methodology.

3. Methodology

3.1. Model structure

We follow the model structure defined in Verdin et al. (2015), a summary of which is provided below. In describing this methodology, we also develop modifications to improve flexibility by producing conditional weather sequences driven by seasonal forecasts, multi-decadal projections, climate drivers or variables, or any other relevant information introduced as time series of covariates. It should be noted that in Equations (2), (4), (7), and (8), the ellipses denote any number of relevant covariates the user wishes to include, such as seasonal characteristics (e.g., mean temperature or total precipitation), large-scale climate modes (e.g., El Niño-Southern Oscillation, Pacific Decadal Oscillation, Atlantic Multidecadal Oscillation), or any other climatic variables. Here we propose to use seasonal spatial average precipitation and temperatures as covariates. These additional covariates are calculated from the gauge data. It is acknowledged that a possible scale mismatch exists between the domain average calculated from 17 stations and the true domain average. However, the network of stations is evenly spaced throughout the domain, thus it is fair to assume the stations adequately represent the true domain average.

In the weather generator described here we define two explicit components of daily weather patterns: local climate and daily variability (as suggested by Kleiber et al., 2013). Local climate represents the expected value of a given meteorological process largely due to seasonal cycle; daily variability provides perturbations to local climate due to weather. Precipitation is considered the primary variable in that occurrence of precipitation tends to modify maximum and minimum temperatures on that day (e.g., due to cloud cover and latent heat transfer). Minimum and maximum temperatures are therefore conditional on precipitation occurrence; precipitation intensities are modeled and simulated independently from occurrence.

In this research, precipitation occurrence and intensity (e.g., amounts), and minimum and maximum temperatures at location $s \in \mathbb{R}^2$ for day $t = 1, \dots, T$, where T is the number of days in the observational record, are denoted as $O(s, t)$, $A(s, t)$, $Z_N(s, t)$, and $Z_X(s, t)$, respectively. As in Verdin et al. (2015), occurrence is modeled as a probit process driven by a latent Gaussian process $W_O(s, t)$ via:

$$O(s, t) = \mathbb{1}_{\{W_O(s, t) > 0\}} \quad (1)$$

If $W_O(s, t)$ is positive, this is indicative of rain on day t at location s and is assigned the value 1; if the latent Gaussian process is negative or equal to zero, day t at location s is dry and is assigned the value 0. The mean function of the latent Gaussian process is simply a regression on covariates that are appropriate for the domain of interest. Similar to Verdin et al. (2015), this regression has covariates

$$X_O(s, t) = \left(1, O(s, t-1), \cos\left(\frac{2\pi t}{365}\right), \sin\left(\frac{2\pi t}{365}\right), ST(t), \dots \right), \quad (2)$$

which are the intercept term, the previous day's occurrence, two harmonic terms to account for seasonality, and the domain-averaged seasonal total precipitation. The key modification to this regression is the seasonal total covariate, denoted by $ST(t)$. In practice this covariate is divided into four distinct covariates relating to each season; covariates are set to zero for times not included in their respective season. To maintain spatial correlations of precipitation occurrence in the domain, an explicit correlation function is defined for $W_O(s, t)$. A correlation function is used instead of a covariance function because probit regression has variance unity by definition. Precipitation amounts at any individual location are modeled as a Gamma random variable as in Kleiber et al. (2012) as follows:

$$A(s, t) = G_{s,t}^{-1}(\Phi(W_A(s, t))), \quad (3)$$

where $G_{s,t}^{-1}$ is the quantile function (e.g., inverse cumulative distribution function) of the Gamma distribution at location s and day t , and Φ is the cumulative distribution function of a standard normal. The simulated rainfall values maintain spatial correlation by applying a spatially varying copula function to the zero-mean Gaussian process $W_A(s, t)$ with correlation function $C_A(h, t)$ (Chilés and Delfiner, 1999). The shape parameter varies with space, such that each location has its own distinct value; the scale parameter varies with both space and time – its time dependence is based on the seasonal characteristics of precipitation, which are generally captured by annual harmonics. Similar to the occurrence process, the Gamma model parameters are informed by a set of covariates, including the areal seasonal total precipitation covariates as in the occurrence model, as follows:

$$X_A(s, t) = \left(1, \cos\left(\frac{2\pi t}{365}\right), \sin\left(\frac{2\pi t}{365}\right), ST(t), \dots \right) \quad (4)$$

Following Verdin et al. (2015), the minimum and maximum temperatures, $Z_N(s, t)$ and $Z_X(s, t)$, respectively, at location s and day t are decomposed as follows:

$$Z_N(s, t) = \beta_N(s)X_N(s, t) + W_N(s, t) \quad (5)$$

$$Z_X(s, t) = \beta_X(s)X_X(s, t) + W_X(s, t) \quad (6)$$

In each equation, the product on the right side of the equality is a regression on some covariates, $X_N(s, t)$ and $X_X(s, t)$ for minimum and maximum temperatures, respectively; these products represent the average behavior of temperatures over the observational period. In this, the key modification to the weather generator of

Verdin et al. (2015) is the inclusion of areal seasonal mean minimum ($SMN(t)$) and maximum ($SMX(t)$) temperature covariates, which are included in both temperature models, as follows,

$$X_N(s, t) = \left(1, \cos\left(\frac{2\pi t}{365}\right), \sin\left(\frac{2\pi t}{365}\right), r(t), Z_N(s, t-1), Z_X(s, t-1), O(s, t), SMN(t), SMX(t), \dots \right) \quad (7)$$

$$X_X(s, t) = \left(1, \cos\left(\frac{2\pi t}{365}\right), \sin\left(\frac{2\pi t}{365}\right), r(t), Z_N(s, t-1), Z_X(s, t-1), O(s, t), SMN(t), SMX(t), \dots \right), \quad (8)$$

which are the intercept term, two harmonic terms to account for seasonality, $r(t)$, which is a linear drift ranging from -1 to 1 to account for temperature trends over the observational period, the previous day's minimum and maximum temperatures, the current day's precipitation occurrence, and the seasonal mean minimum and mean maximum temperatures, respectively. Daily variability is denoted as $W_N(s, t)$ and $W_X(s, t)$ for minimum and maximum temperatures, respectively, and maintains spatial correlation by realizations from a mean zero Gaussian process with an empirical covariance structure defined by the residuals of the local regressions. Kleiber et al. (2013) found that the Gaussian assumption for temperature models was appropriate. The above are GLMs and are fitted hierarchically – we refer the reader to Verdin et al. (2015) and Kleiber et al. (2012, 2013) for details on implementation.

It should be noted that the additional covariates are applied only to the local climate component, and not the daily weather component. The daily weather component is by definition random, temporally independent noise (see Kleiber et al., 2012 for validation of this assumption), thus the daily weather component is not conditional on the additional covariates, rather it is conditioned only by the calendar date – there are distinct correlation (and covariance) matrices for each month.

3.2. Significance testing

The inclusion of seasonal covariates could lead to a reduction in the significance of the harmonic covariates. For all 17 stations the seasonal covariates are highly significant, indicated by the respective p -values of their regression coefficients. For many of the stations both cosine and sine covariates remain highly significant, however, at few stations the sine covariate loses significance. The Akaike information criterion (AIC) of the modified models at each location for each climate variable are consistently lower than those for the original models that do not contain the seasonal covariates, implying that the modified weather generator more adequately describes the modeled processes. Table 1 reports the change in AIC value (original minus modified models – positivity implies a decrease) for all 17 stations, for the four variables that make up the weather generators.

4. Results from application in the Salado A basin

4.1. Covariate selection

We apply the methodology as described in the previous section to the network of 17 stations in and around the Salado A basin of the Argentine Pampas (see Fig. 1). Given the relative homogeneity of the basin area and scale at which seasonal climate forecasts are available, we propose three domain-averaged covariates: seasonal total precipitation, seasonal mean minimum temperature, and seasonal mean maximum temperature. The

Table 1

Differences between the AIC for occurrence, amounts, minimum temperature, and maximum temperature models of the original and modified weather generators (positivity implies a decrease).

Station	1	2	3	4	5	6	7	8	9
OCC	112	92	66	98	56	122	92	112	106
AMT	103	29	20	89	17	63	74	92	97
MIN	732	439	420	559	157	489	340	661	625
MAX	240	186	165	254	39	144	196	165	237
	10	11	12	13	14	15	16	17	
OCC	89	103	48	134	114	98	127	98	
AMT	82	87	66	50	63	54	66	58	
MIN	560	540	258	603	495	547	877	403	
MAX	176	181	104	231	264	244	194	219	

growing season for summer crops in the Salado A basin begins in October with harvest coming in late March or April, therefore we focus on the OND season. We define the seasons as January–March (JFM), April–June (AMJ), July–September (JAS), and October–December (OND).

The first principal component of OND seasonal total precipitation at each of the 17 stations explains 47% of the total variance; those of OND seasonal average minimum and maximum temperatures explain 71% and 77% of the total variance, respectively. The magnitudes of these first principal components are nearly constant across space, which further justifies the use of domain-averaged information. Fig. 2 shows the first principal component of the three variables along with the domain-averaged time series, the behavior of which are well described by their first principal components. Thus the four GLMs as described in the previous section were fitted with the additional covariates described above. These covariates were found to be highly significant at all the locations (e.g., regression assigns all additional covariates *p*-values < 0.001). Other seasons show similar results (not shown).

4.2. Validation

To assess the efficacy of the additional covariates, we employ both the original and modified weather generators in spatial and temporal validations, described in the following subsections.

4.2.1. Spatial validation

To assess the spatial performance of the modified weather generator, three stations were withheld from the model fitting process – these withheld stations are identified in Fig. 1. Spatial process models were used to estimate the model parameters at the withheld locations, and 100 realizations over the 53-year observational period were produced using the estimated parameters. Fig. 3 shows the relationship between the observed and ensemble mean OND probability of occurrence, seasonal total rainfall, mean maximum temperature, and mean minimum temperature for each of the three stations as produced by the original (top row) and modified (bottom row) weather generators. Simulations from the original weather generator show no relationship with the observations; this is to be expected, as only harmonic and autoregressive covariates are considered. Conversely, simulations from the modified generator capture the observations strongly, due to the inclusion of the domain-averaged seasonal covariates. Similar results were seen for other seasons (figures not shown).

4.2.2. Temporal validation

It is also worthwhile to investigate the temporal performance of the weather generator to validate its use for seasonal forecasts, multidecadal projections, and climate change scenarios. To this end, we fitted the original and modified weather generators on historic data for the calibration period: 1 January 1961–31 December 2000. Then 100 realizations were generated for the validation period: 1 January 2001–31 December 2013. Fig. 4 shows the difference between observed and ensemble mean simulated domain-averaged

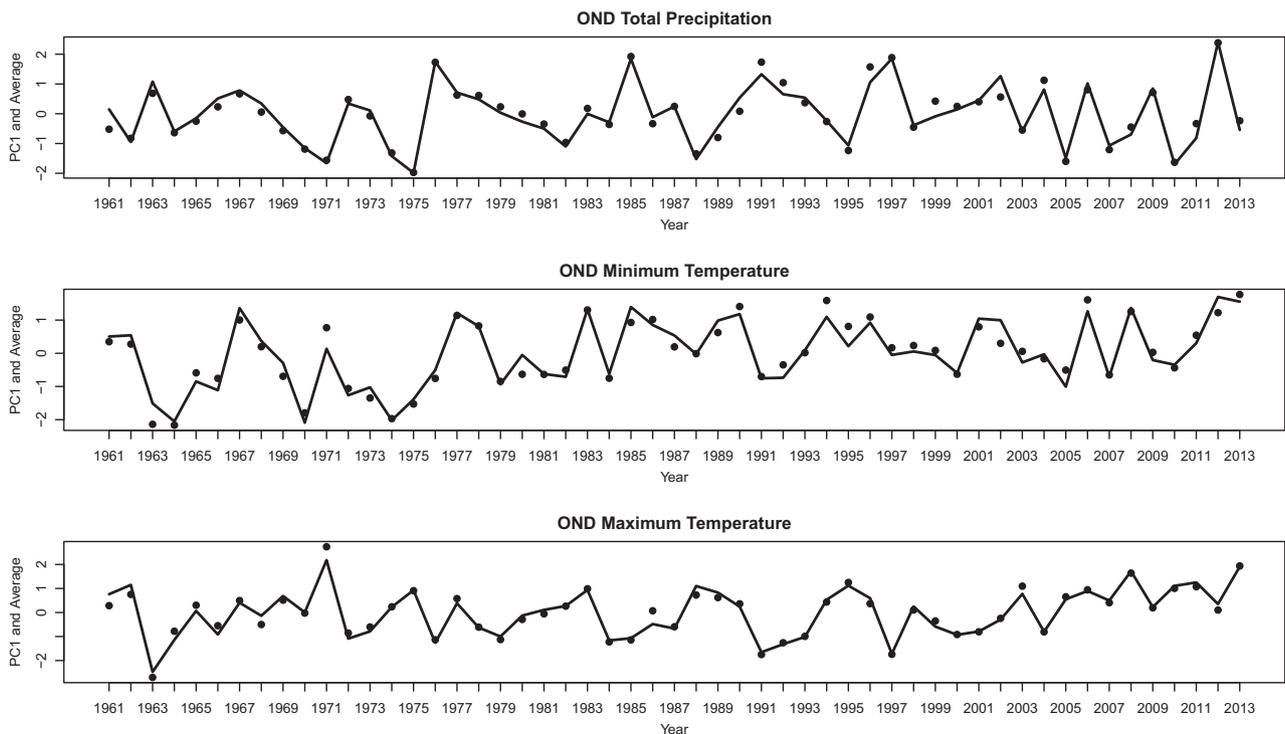


Fig. 2. First principal components of OND precipitation, minimum temperature, and maximum temperature, scaled and shown as points, and domain-averaged and scaled OND precipitation (top), minimum temperature (middle), and maximum temperature (bottom), shown as lines.

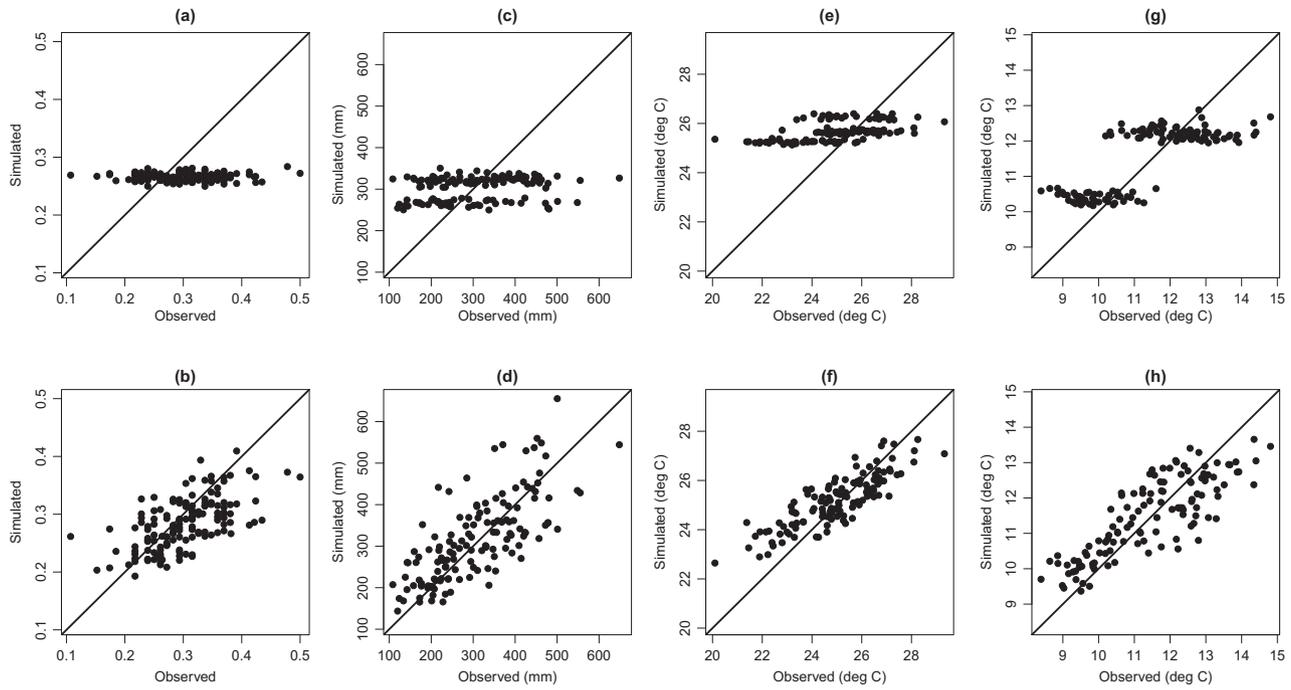


Fig. 3. Spatial validation: (a and b) OND 1961–2013 observed versus ensemble mean simulated probability of precipitation occurrence, (c and d) total precipitation, (e and f) mean maximum temperature, and (g and h) mean minimum temperature, for the three withheld stations. Top row (a, c, e, g) corresponds to simulations from the original generator and bottom row (b, d, f, h) is for simulations from the modified generator.

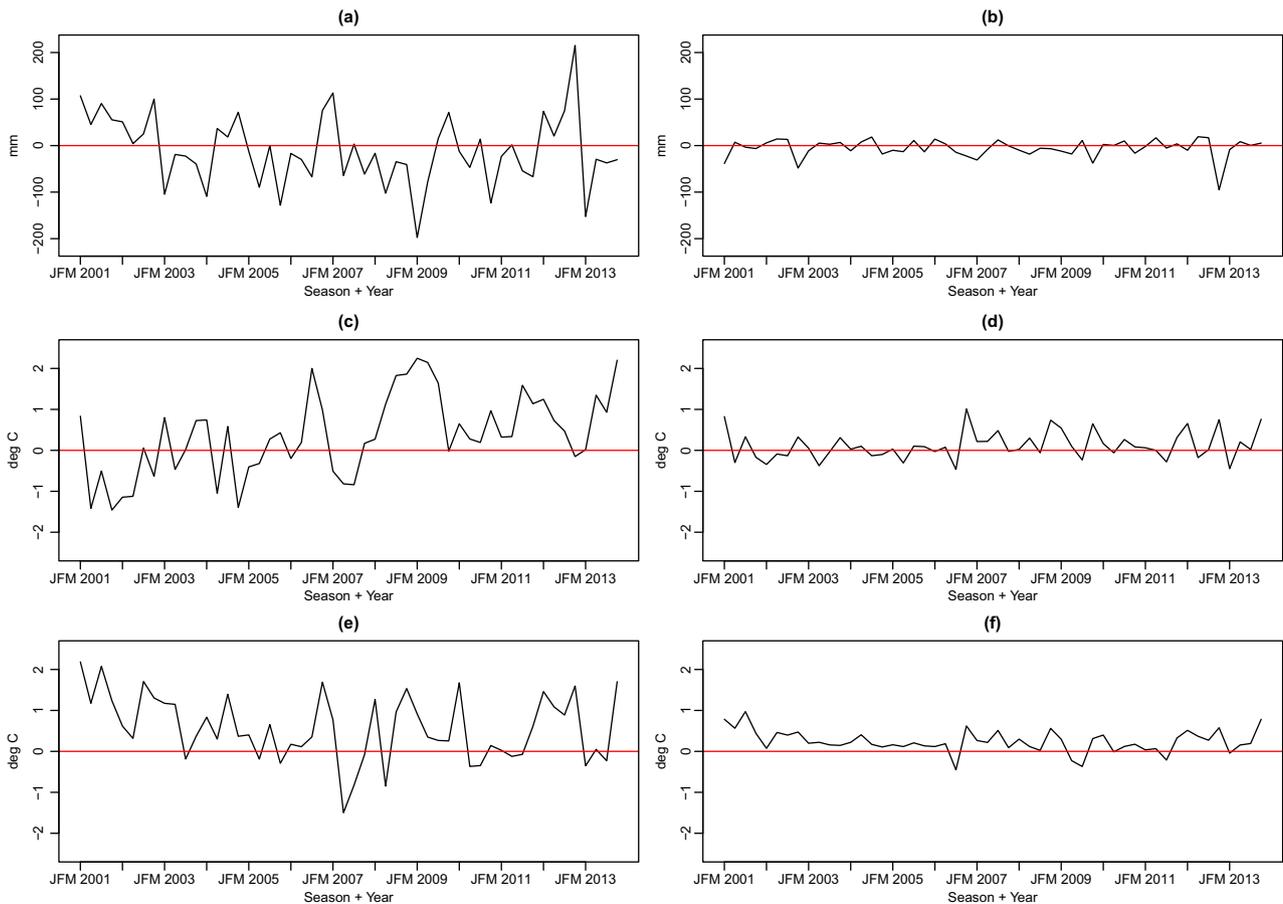


Fig. 4. Temporal validation: JFM 2001 – OND 2013 observed minus ensemble mean simulated (a and b) seasonal total precipitation, (c and d) mean maximum temperature, and (e and f) mean minimum temperature. Left panels (a, c, e) are for the original weather generator and right panels (b, d, f) for the modified weather generator.

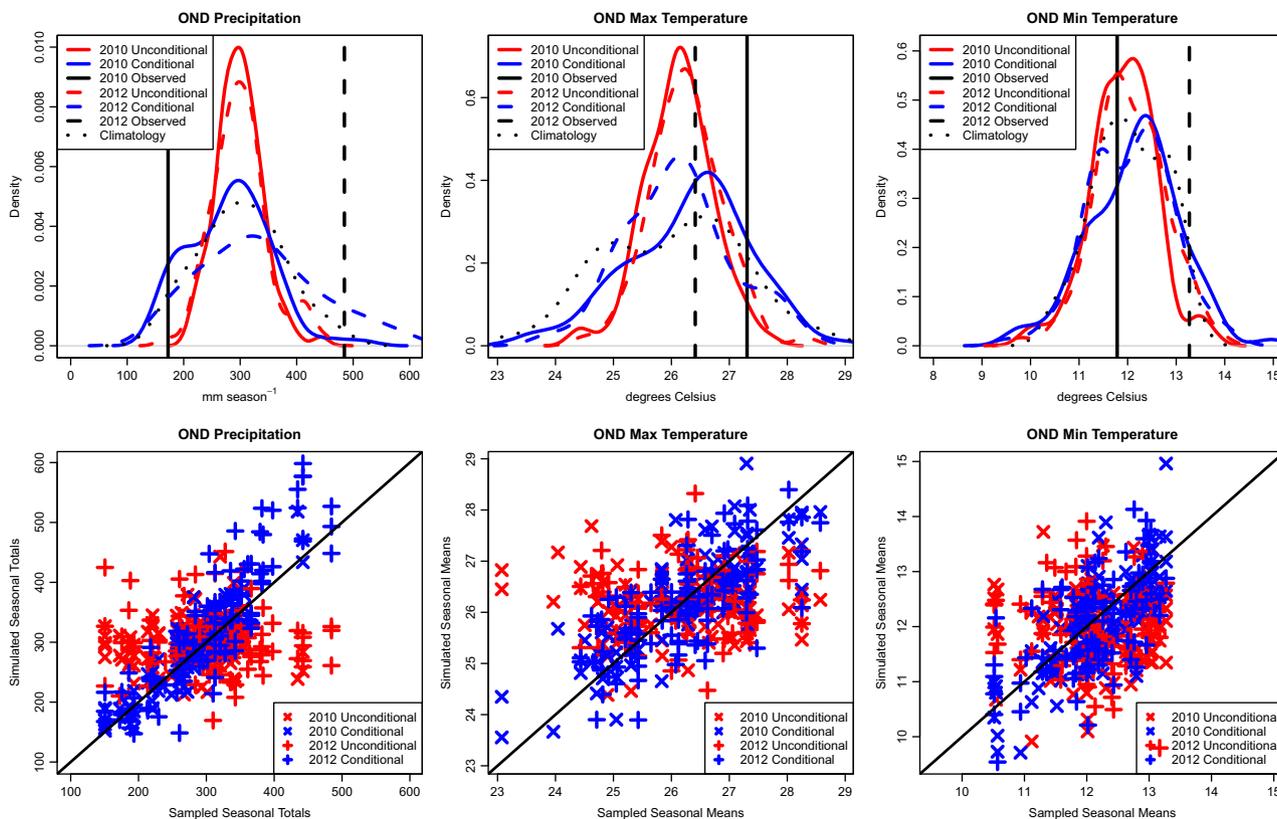


Fig. 5. Top panels: Kernel density estimates of PDF of domain-averaged seasonal precipitation, maximum temperature and minimum temperature from 100 simulated weather scenarios from the modified (blue) and original (red) weather generators (OND 2010 and OND 2012 denoted as solid and dashed lines, respectively), along with the climatological PDF (dotted black line). Observed values are shown as vertical lines. Bottom panels: Sampled seasonal precipitation and temperatures from the categorical probabilistic forecasts with the domain-averaged values generated from the two weather generators – modified (blue) and original (red). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

seasonal total precipitation, mean maximum temperature, and mean minimum temperature: a “perfect fit” would show a horizontal line of ordinate zero. Root mean square error (RMSE) is calculated between simulated and observed seasonal values for all 100 realizations. The RMSE is greatly reduced by including the domain-averaged seasonal covariates in the validation period. For seasonal total precipitation the RMSE is reduced from 77 (± 2.4) mm to 21 (± 3.6) mm; for seasonal mean maximum temperature the RMSE is reduced from 1.05 (± 0.05) °C to 0.37 (± 0.05) °C; and for seasonal mean minimum temperature the RMSE is reduced from 0.99 (± 0.04) °C to 0.37 (± 0.04) °C.

4.3. Seasonal forecasts

Often times, seasonal climate forecasts are provided as probabilities of precipitation and temperature being within different ranges (e.g., terciles) for a large region – this is a common format for presenting uncertain climate information. Among other agencies around the world, the International Research Institute for Climate and Society (IRI, www.iri.columbia.edu) provides seasonal (three-month) probabilistic forecasts with one to four months lead-time. The IRI presents these forecasts in terms of A:N:B probabilities, where “A” is above-normal, “N” is near-normal, and “B” is below-normal. The three categories span an equal range and are defined with respect to climatological terciles (e.g., 33rd and 67th percentiles). For example, a 15:35:50 precipitation forecast implies there is a 15% chance of experiencing above-normal conditions, a 35% chance of experiencing near-normal conditions, and a 50% chance of experiencing below-normal precipitation in the upcoming season.

Agricultural decisions in the Salado A basin are typically made before the beginning of the summer growing season (1 October) every year, thus we focus on OND seasonal forecasts. The OND season is also a critical period in terms of crop yield generation, and has shown tendencies towards skillful climate predictions, in part due to significant ENSO signals (Grimm et al., 1998, 2000; Montecinos et al., 2000; Ropelewski and Halpert, 1987; Barros and Silvestri, 2002; Boulanger et al., 2005; Ropelewski and Bell, 2008; Grimm, 2011; Barreiro, 2010). We select IRI forecasts for OND 2010 (a dry and hot forecast; e.g., 15:35:50 for precipitation, 40:35:25 for temperature) and OND 2012 (a wet and hot forecast; e.g., 40:35:25 for both precipitation and temperature) as case studies for this methodology, issued on 1 September 2010 and 1 September 2012, respectively.

To generate space-time weather sequences for the two OND seasons from the modified generator, ensembles of domain-averaged seasonal precipitation and temperature are needed to use as covariates. To this end, 100 observed OND domain-averaged values of precipitation and maximum and minimum temperature are sampled with replacement. This is accomplished by first categorizing the observed domain-averaged seasonal weather as above-, near-, or below-normal based on the empirical terciles, then assigning the categorical forecasts as probabilities (e.g., 15:35:50 and 40:35:25 for precipitation and temperature, respectively) to the values in each category and resampling with these assigned weights as the probability metric. For instance, there is a 15% chance of sampling an above-normal precipitation value; there are 35% and 50% chances of sampling near-normal and below-normal precipitation values, respectively. The result of this resampling scheme is 100 values that are used as covariates

Table 2

Kolmogorov–Smirnov tests comparing the original and modified weather generator output. *P*-values lower than 0.05 indicate the output from original and modified generators may come from different distributions.

	2010			2012		
	Precip	Max temp	Min temp	Precip	Max temp	Min temp
<i>p</i> -values	0.0039	0.0014	0.0243	<0.0001	0.0541	0.5806

to drive the modified weather generator 100 separate times. The output of these 100 independent runs is essentially a downscaled ensemble of weather patterns that exhibit the traits of the seasonal forecasts.

The top row of Fig. 5 shows the probability density functions (PDFs) of domain-averaged OND precipitation and temperatures from the original and modified weather generators, the PDF of OND climatology, and the observed values of OND 2010 and 2012. The precipitation PDFs from the modified generator have shifted towards the observed values in both 2010 and 2012. This shift is indicative not only of forecast skill, but also the effectiveness of the modified generator in simulating scenarios representative of the forecasts. Mean maximum (minimum) temperature during OND 2010 (2012) was observed to be above-normal, and the PDF from the modified generator gives greater probability to above-average temperatures than that of the original generator. OND 2012 (2010) experienced near-normal maximum (minimum) temperatures, so the original generator gives highest probability to observed conditions. However, the range of possible scenarios offered by the original generator is limited and will give near-zero probability to above- and below-normal conditions, which for planning purposes can be misleading. The domain-averaged seasonal totals of precipitation and seasonal averages of temperatures that are generated from the two weather generators are plotted with the observed in the bottom row of Fig. 5.

Table 2 reports *p*-values from Kolmogorov–Smirnov tests comparing the distributions of original and modified generator output. The differences between the original and modified distributions for OND 2010 weather scenarios and OND 2012 precipitation are significant at the 95% level; maximum and minimum temperature scenarios for OND 2012 are not significantly different, indicating the covariate values sampled from the IRI probabilistic forecast (thus the scenarios produced by the modified weather generator) do not deviate significantly from climatology.

Weather simulations on a regular grid are of particular interest, as they are used to drive hydrologic and agriculture models for agricultural planning to mitigate crop failure. To simulate daily weather on a grid, the β coefficients for each covariate of the weather generator models are estimated in space from their

respective spatial models to the desired spatial resolution (5 km \times 5 km). These gridded coefficients are then used to obtain the mean function, and the daily weather processes are simulated via mean-zero Gaussian random fields.

Fig. 6 shows the difference between the ensemble mean of gridded seasonal total precipitation, mean maximum temperature, and mean minimum temperature for OND 2010 from the original and modified weather generator. The modified generator simulates a drier and hotter domain than the original generator, which is consistent with the seasonal forecast. Notably, the modified weather generator simulates a cooling for mean minimum temperature in the southern part of the basin, which is inconsistent with the seasonal forecast.

The differences between the 95% ensemble spread (97.5th percentile minus 2.5th percentile) produced by the original and modified weather generators are shown in Fig. 7. As can be seen, the ensemble spread difference for seasonal total precipitation is mostly red and yellow, while those for mean maximum and minimum temperature are mostly blue and yellow, which illustrates that the modified generator produced wider ensemble spreads than the original generator. The uncertainty in the probabilistic seasonal climate forecast is propagated to the modified weather generator, resulting in a wider distribution than that of the original generator. Similar findings can be seen for OND 2012 (figures not shown).

4.4. Multi-decadal projections

Multi-decadal projections are useful in a number of applications, including environmental impact studies, agricultural decision-making, and water resources management, to name a few. In agriculture, multi-decadal projections help in making informed investment decisions (e.g., whether or not to buy a farm in a climatically marginal area, invest in irrigation, etc.). Specifically, the climate of the Pampas has shown significant decadal variability, and since the 1970s has exhibited a steady increase in both annual and extreme precipitation. This trend in precipitation has in part promoted significant expansion of agricultural area to climatically marginal regions of the Pampas. Given the uncertainty of future climate, coupled with a known decadal variability, it is unclear if existing agricultural systems may remain viable if climate reverts to a drier epoch.

Specific to this research, future climate scenarios can be used to drive hydrologic and crop simulation models, thus providing an assessment of the viability of existing agricultural production systems in climatically marginal regions of the Salado A basin. However, future climate projections from climate models are generally of coarse spatial (e.g., on a grid) and temporal (e.g., monthly) resolutions, and therefore cannot provide reliable

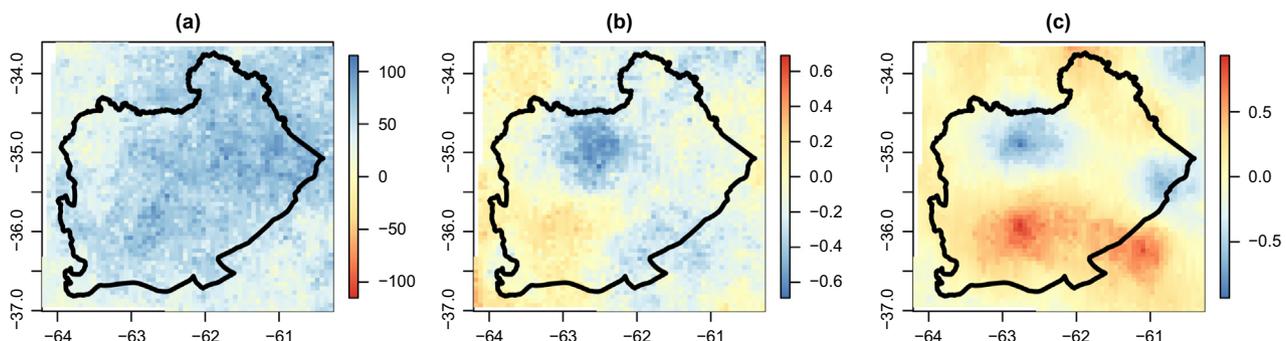


Fig. 6. OND 2010 differences in ensemble mean of seasonal (a) total precipitation (mm season⁻¹), (b) mean maximum temperature (°C), and (c) mean minimum temperature (°C). Differences calculated as original minus modified generators. Salado A basin is outlined.

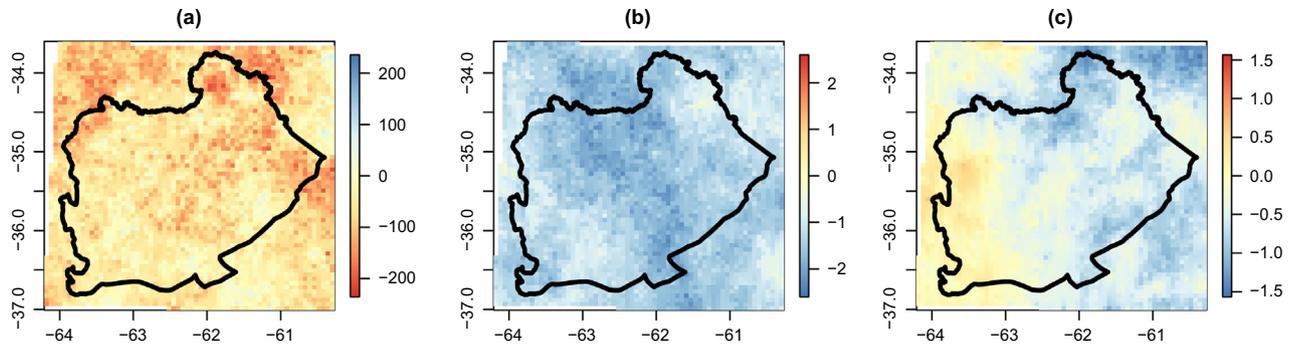


Fig. 7. OND 2010 differences in 95% ensemble spread for (a) seasonal total precipitation (mm season^{-1}), (b) seasonal mean maximum temperature ($^{\circ}\text{C}$), and (c) seasonal mean minimum temperature ($^{\circ}\text{C}$). Differences calculated as original minus modified generators. Salado A basin is outlined.

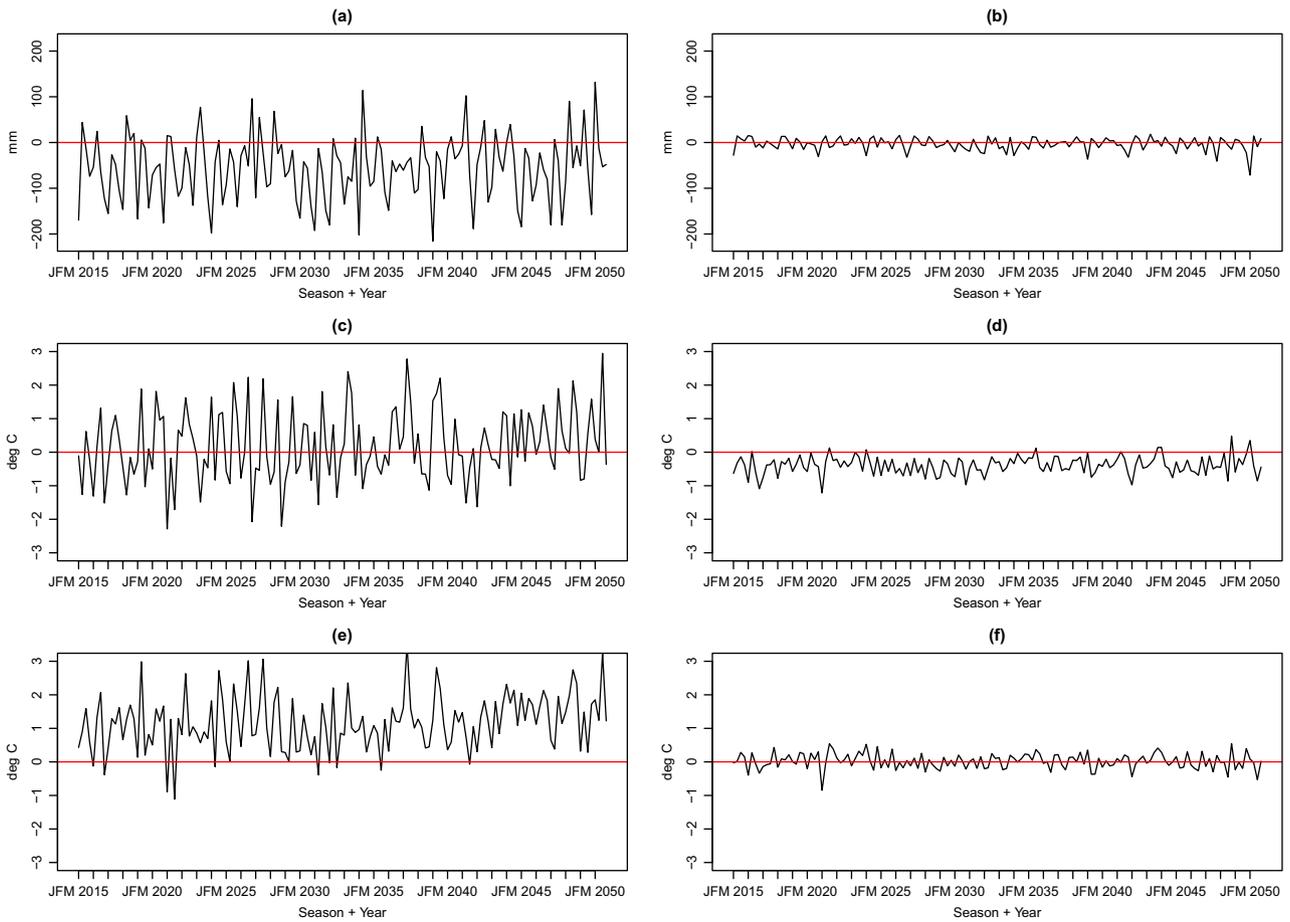


Fig. 8. JFM 2015 – OND 2050 projected minus ensemble mean (a and b) simulated seasonal total precipitation, (c and d) mean maximum temperature, and (e and f) mean minimum temperature.

projections of weather at the local scale. These monthly and consequently seasonal projections can be incorporated into the modified generator and thus enable the generation of daily weather sequences conditioned on the projections at any desired location – both monitored and unmonitored – in the study region.

To this end, we explored the ability of the modified generator to downscale medium-term projections in the Salado A basin. A regional climate model projection, experiment RCP8.5, was obtained for the period 1 January 2015 to 31 December 2050 (a 36-year projection), produced using the CORDEX-CMIP5 regional climate model (EC-Earth Consortium, 2014) and bias-corrected (McGinnis et al.,

2015) using the CLARIS-LPB dataset (Penalba et al., 2014). This projection focuses on South America and is gridded at 0.44° . No notable long-term trends in annual precipitation totals are projected, but the magnitudes are significantly lower than seen in the historic record; both maximum and minimum annual average temperatures show positive trends, and are projected to increase by approximately 1°C by the year 2050. Seasonal values of areal precipitation and temperature for the Salado A basin were computed to use as covariates to drive the modified weather generator. Only the grids that cover the Salado A basin and the 17 station data are considered when computing domain-averaged precipitation totals

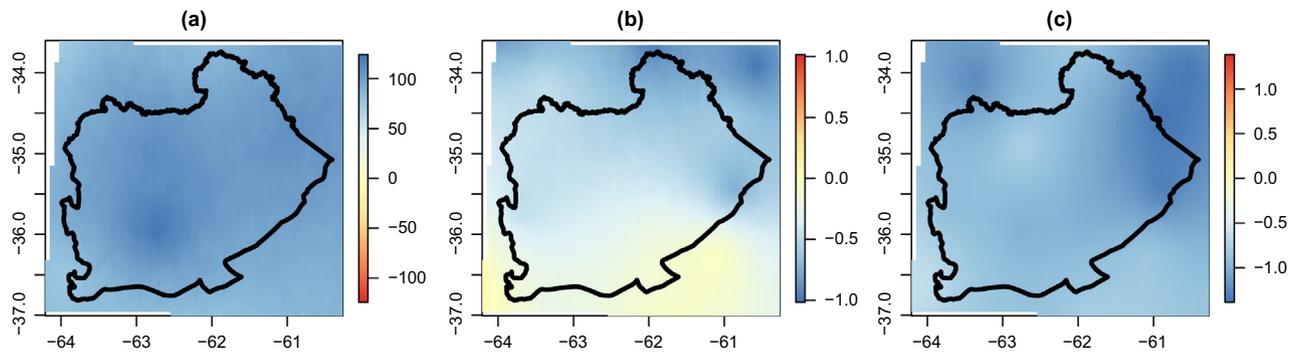


Fig. 9. OND 2015–2050 differences in ensemble mean of seasonal (a) total precipitation (mm season^{-1}), (b) mean maximum temperature ($^{\circ}\text{C}$), and (c) mean minimum temperature ($^{\circ}\text{C}$). Differences calculated as original minus modified generators. Salado A basin is outlined.

and temperature means. 100 realizations of daily weather sequences were simulated using both the original and modified weather generators.

Fig. 8 shows seasonal residuals (projected minus simulated) of the ensemble mean of the original and modified weather generator simulations for the 36-year projection period. As was seen previously, the original generator shows much larger and more variable residuals than does the modified generator. Consistent with the temporal validation, RMSE is greatly reduced by including the domain-averaged seasonal covariates in the models. RMSE for seasonal total precipitation is reduced from 90 mm to 14 mm; RMSE for seasonal mean maximum temperature is reduced from 1.09 $^{\circ}\text{C}$ to 0.48 $^{\circ}\text{C}$; RMSE for seasonal mean minimum temperature is reduced from 1.43 $^{\circ}\text{C}$ to 0.23 $^{\circ}\text{C}$. There is a slight warm bias in seasonal mean maximum temperature simulated by the modified generator as compared to that projected by the RCM.

To illustrate the spatial ability of the weather generators, we simulated 100 realizations of daily weather sequences for the period 2015–2050, conditioned on the projected seasonal characteristics. Fig. 9 shows the difference in ensemble mean OND total precipitation, maximum temperature, and minimum temperature as simulated by the original and modified generators. Consistent with the climate model trends, the modified generator simulates a drier and hotter future across the domain.

5. Summary and future work

We have proposed and validated the use of a parametric stochastic weather generator in a nonstationary context, such as climate change impact studies, with application in the Salado A basin of the Argentine Pampas. This region was selected due to its status as one of the most productive agricultural regions in South America, and its strong climatic variability that is experienced at multiple time scales. Agriculture in the Pampas is predominantly rainfed, thus high quality seasonal forecast information could greatly impact the outcome (e.g., crop yield, risk of failure) of a growing season. The modified weather generator presented in this research has flexibility in its GLM framework such that any number of covariates can be included in the model fit, effectively conditioning the weather generator to produce downscaled weather sequences.

For example, in this research we used areal average seasonal total precipitation and mean minimum and maximum temperatures as additional covariates, which were shown to be highly significant in the model fit. The use of areal averages was justified via principal component analysis; for non-homogeneous or mountainous regions, consider site-specific averages or a clustering algorithm. The coarse information provided by these additional covariates successfully trickled from seasonal (regional) down to daily (local) scales, such that wet (dry) days are more prevalent

during seasons with above-normal (below-normal) seasonal total precipitation. It is with the conditioned output of the weather generator that research teams may provide a more robust estimate of production risk for a region, by running the daily weather sequences through process based (e.g., crop simulation, hydrologic) models. The output of process based models may be interpreted and provided to a farmer or decision maker, who then will have seasonal forecast information that is relevant to the decisions they must make (e.g., probability of not meeting a crop yield goal, where and when to plant a certain crop) as opposed to spatially coarse probabilistic statements as are typically reported. Similarly, multidecadal projection information can be used to generate conditional weather sequences to assist in assessing the viability of existing agricultural infrastructure in climatically marginal regions. In this, a regional climatic trend may be extracted and used to produce conditional weather sequences, which may be used to drive any relevant process based models.

The output of the modified weather generator presented in this manuscript has been validated by direct comparison to the original weather generator of Verdin et al. (2015). It has been shown that using simple covariates such as domain-averaged seasonal total (mean) precipitation (temperatures) improves the skill of the generator in producing daily weather sequences that exhibit the traits (and trends) of a seasonal forecast or multi-decadal projection. In representing domain-averaged behavior for the validation period (2001–2013), this modification to the weather generator reduced RMSE values from 77 mm to 21 mm for precipitation, 1.05 $^{\circ}\text{C}$ to 0.37 $^{\circ}\text{C}$ for maximum temperature, and 0.99 $^{\circ}\text{C}$ to 0.37 $^{\circ}\text{C}$ for minimum temperature. Similarly, the modified generator faithfully reproduced the trends and variability of historic precipitation and temperature at individual sites, while the original generator replicates the expected behavior of (e.g., climatology) of each season with little to no interannual variability.

In generating sequences consistent with a seasonal forecast, the Kolmogorov–Smirnov tests suggest the output from original and modified weather generators exhibit significantly different traits with 95% confidence, unless the seasonal forecast is similar to climatology. The modified weather generator was shown to produce PDFs that better represent the range of possible futures, while the PDFs from the original weather generator give near-zero probability to the upper and lower terciles (e.g., wet (hot) and dry (cold) conditions). On the multi-decadal scale, the modified weather generator is flexible in its ability to capture the considerable interannual and decadal variability prevalent in the projected precipitation totals, as well as the increase in both minimum and maximum temperatures.

Application of this methodology to other areas is called for. However, careful attention must be paid to the spatial and temporal climatic variability in the region of interest. Local climate,

regional teleconnections, and global climate drivers should be identified for optimal skill in downscaling seasonal forecasts and multi-decadal projections. Principal component analysis on seasonal attributes such as seasonal total precipitation and mean temperatures can help decide if domain-averaged, clustered, or site-specific covariates should be considered. However, the model setup, as defined in Verdin et al. (2015), should be considered a baseline model for use in any basin – the additional covariates as described in this manuscript need be fine tuned to successfully generate skillful weather scenarios.

One shortcoming to the weather generator presented in this research is that it uses only precipitation and temperature covariates to condition the weather generator. There has been great progress in the identification of teleconnections and climate drivers for regions around the world – the Pampas are no exception. While it was mentioned in this manuscript that such teleconnections could be used as covariates to condition the weather generator output, this approach was not investigated. A second shortcoming to this methodology is that the uncertainties associated with the parameters of the weather generator are not propagated to the simulations, as the maximum likelihood estimates of the parameters are kept fixed. As a result the variability of simulations can be underestimated. Bayesian methods that explicitly quantify the parameter uncertainties are attractive options.

The methodology of this weather generator is inherently hierarchical, thus considering the use of a Bayesian hierarchical framework is a natural extension to this problem – the authors are currently exploring this approach. In a Bayesian context, the parameters are treated as random variables and are sampled from appropriate distributions (typically via Markov chain Monte Carlo) based on likelihood acceptance criteria, which results in posterior distributions of all model parameters. These posterior distributions better represent the uncertainty involved in traditional parameter estimation techniques, and when used in a weather generation framework will provide a more realistic range of uncertainty in synthetic weather sequences.

Acknowledgments

This research is funded by National Science Foundation Award # 1211613. The authors would like to thank the Met Service of Argentina for providing the weather data. The authors are grateful for Drs. Linda Means and Seth McGinnis of the National Center for Atmospheric Research (NCAR) for providing the CMIP5 regional climate model projections. This work utilized the Janus supercomputer, which is supported by the National Science Foundation (award number CNS-0821794) and the University of Colorado Boulder. The Janus supercomputer is a joint effort of the University of Colorado Boulder, the University of Colorado Denver and the National Center for Atmospheric Research. We thank the two anonymous reviewers for their insightful comments, which significantly improved the manuscript.

References

Apipattanavis, S., Bert, F., Podestá, G., Rajagopalan, B., 2010. Linking weather generators and crop models for assessment of climate forecast outcomes. *Agric. For. Meteorol.* 150 (2), 166–174. <http://dx.doi.org/10.1016/j.agrformet.2009.09.012>.

Aragón, R.M., Jobbágy, E.G., Viglizzo, E.F., 2010. Surface and groundwater dynamics in the sedimentary plains of the Western Pampas (Argentina). *Ecohydrology* 4, 433–447.

Barnston, A.G. et al., 2010. Verification of the first 11 years of IRI's seasonal climate forecasts. *J. Appl. Meteorol. Climatol.* 49 (3), 493–520.

Barreiro, M., 2010. Influence of ENSO and the South Atlantic Ocean on climate predictability over Southeastern South America. *Clim. Dyn.* 35 (7), 1493–1508.

Barros, V.R., Silvestri, G.E., 2002. The relation between sea surface temperature at the subtropical South-Central Pacific and precipitation in southeastern South America. *J. Clim.* 15, 251–267.

Beersma, J., Buishand, T., 2003. Multi-site simulation of daily precipitation and temperature conditional on the atmospheric circulation. *Clim. Res.* 25, 121–133. http://www.leg.ufpr.br/~eder/Artigos/Wather/Beersma_2003.pdf.

Berger, T., 2001. Agent-based spatial models applied to agriculture: a simulation tool for technology diffusion, resource use changes, and policy analysis. *XXIV Int. Conf. Agric. Econom. (IAAE)* 25, 245–260.

Berger, T., Schreinemachers, P., Woelcke, J., 2006. Multi-agent simulation for the targeting of development policies in less-favored areas. *Agric. Syst.* 88, 28–43. <http://dx.doi.org/10.1016/j.agry.2005.06.002>.

Bert, F.E., Satorre, E.H., Toranzo, F.R., Podestá, G.P., 2006. Climatic information and decision-making in maize crop production systems of the Argentinean Pampas. *Agric. Syst.* 88 (2–3), 180–204. <http://dx.doi.org/10.1016/j.agry.2005.03.007>.

Bert, F.E., Laciána, C.E., Podestá, G.P., Satorre, E.H., Menéndez, A.N., 2007. Sensitivity of CERES-Maize simulated yields to uncertainty in soil properties and daily solar radiation. *Agric. Syst.* 94 (2), 141–150. <http://dx.doi.org/10.1016/j.agry.2006.08.003>.

Bert, F.E., Rovere, S.L., Macal, C.M., North, M.J., Podestá, G.P., 2014. Lessons from a comprehensive validation of an agent based-model: the experience of the Pampas Model of Argentinean agricultural systems. *Ecol. Model.* 273, 284–298. <http://dx.doi.org/10.1016/j.ecolmodel.2013.11.024>.

Boulanger, J.-P. et al., 2005. Observed precipitation in the Paraná-Plata hydrological basin: long-term trends, extreme conditions and ENSO teleconnections. *Clim. Dyn.* 24 (4), 393–413.

Brandsma, T., Buishand, T.A., 1998. Simulation of extreme precipitation in the Rhine basin by nearest-neighbour resampling. *Hydrol. Earth Syst. Sci.* <http://dx.doi.org/10.5194/hess-2-195-1998>.

Bray, D., von Storch, H., 2009. "Prediction" or "Projection"? The nomenclature of climate science. *Sci. Commun.* 30 (4), 534–543.

Buishand, T., 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. <http://onlinelibrary.wiley.com/doi/10.1029/2001WR002291/full>.

Caraway, N.M., McCreight, J.L., Rajagopalan, B., 2014. Multisite stochastic weather generation using cluster analysis and k-nearest neighbor time series resampling. *J. Hydrol.* 508, 197–213. <http://dx.doi.org/10.1016/j.jhydrol.2013.10.054>.

Caron, A., Leconte, R., Brissette, F., 2008. An improved stochastic weather generator for hydrological impact studies. *Can. Water Resour.* 33 (April 2007), 233–256. <http://www.tandfonline.com/doi/abs/10.4296/cwrj3303233>.

Cash, D.W., Clark, W.C., Alcock, F., Dickson, N.M., Eckley, N., Gaston, D.H., Jager, J., Mitchell, R.B., 2003. Science and technology for sustainable development special feature: knowledge systems for sustainable development. *Proc. Natl. Acad. Sci.* 100 (14), 8086–8091.

Chandler, R.E., 2005. On the use of generalized linear models for interpreting climate variability. *Environmetrics* 16 (7), 699–715. <http://dx.doi.org/10.1002/env.731>.

Chandler, R.E., Wheeler, H.S., 2002. Analysis of rainfall variability using generalized linear models: a case study from the west of Ireland. *Water Resour. Res.* 38 (10). <http://dx.doi.org/10.1029/2001WR000906>.

Chilés, J.P., Delfiner, P., 1999. *Geostatistics: Modeling Spatial Uncertainty*. Wiley, New York.

EC-Earth Consortium, 2014. EC-EARTH Model Output Prepared for CMIP5 rcp85, Served by ESGF World Data Center for Climate. CERAD-B "IHECr8". <http://cera-www.dkrz.de/WDCC/CMIP5/Compact.jsp?acronym=IHECr8>.

Ferreira, R.A., Podestá, G.P., Messina, C.D., Letson, D., Dardanelli, J., Guevara, E., Meira, S., 2001. A linked-modeling framework to estimate maize production risk associated with ENSO-related climate variability in Argentina. *Agric. For. Meteorol.* 107 (3), 177–192. [http://dx.doi.org/10.1016/S0168-1923\(00\)00240-9](http://dx.doi.org/10.1016/S0168-1923(00)00240-9).

Freeman, T., Nolan, J., Schoney, R., 2009. An agent-based simulation model of structural change in Canadian prairie agriculture, 1960–2000. *Can. J. Agric. Econom. – Rev. Can. D Agroec.* 57, 537–554.

Furrer, E., Katz, R., 2007. Generalized linear modeling approach to stochastic weather generators. *Clim. Res.* 34, 129–144. <http://www.image.ucar.edu/~eva/research/files/CR2007FurrerKatz.pdf>.

Giorgi, F., 2002. Variability and trends of sub-continental scale surface climate in the twentieth century. Part I: Observations. *Clim. Dyn.* 18 (8), 675–691. <http://dx.doi.org/10.1007/s00382-001-0204-x>.

Goddard, L., Barnston, A.G., Mason, S.J., 2003. Evaluation of the IRI's "Net Assessment" seasonal climate forecasts: 1997–2001. *Bull. Am. Meteorol. Soc.* 84 (12), 1761–1781.

Grimm, A.M., 2011. Interannual climate variability in South America: impacts on seasonal precipitation, extreme events, and possible effects of climate change. *Stoch. Env. Res. Risk Assess.* 25 (4), 537–554.

Grimm, A.M., Ferraz, S.E.T., Gomes, J., 1998. Precipitation anomalies in southern Brazil associated with El Niño and La Niña events. *J. Clim.* 11, 2863–2880.

Grimm, A.M., Barros, V.R., Doyle, M.E., 2000. Climate variability in southern South America associated with El Niño and La Niña events. *J. Clim.* 13, 35–58.

Gronzona, M.O., Podesta, G.P., Bidegain, M., Marino, M., Hordij, H., 2000. A stochastic precipitation generator conditioned on ENSO phase: a case study in southeastern South America. *J. Clim.* 13 (Richardson 1981), 2973–2986. [http://dx.doi.org/10.1175/1520-0442\(2000\)013<2973:ASPGCO>2.0.CO;2](http://dx.doi.org/10.1175/1520-0442(2000)013<2973:ASPGCO>2.0.CO;2).

Haines, K., Hermanson, L., Liu, C., Putt, D., Sutton, R., Iwi, A., Smith, D., 2009. Decadal climate prediction (project GCEP). *Philos. Trans. Roy. Soc. A: Math., Phys. Eng. Sci.* 367 (1890), 925–937.

Hansen, J.W., Mavromatis, T., 2001. Correcting low-frequency variability bias in stochastic weather generators 109, 297–310.

- Hansen, J.W., Challinor, A., Ines, A., Wheeler, T.y., Moron, V., 2006. Translating climate forecasts into agricultural terms: advances and challenges. *Clim. Res.* 33, 27–41.
- Happe, K., Balmann, A., Kellermann, K., Sahrbacher, C., 2008. Does structure matter? The impact of switching the agricultural policy regime on farm structures. *J. Econ. Behav. Organ.* 67, 431–444. <http://dx.doi.org/10.1016/j.jebo.2006.10.009>.
- Harrold, T.I., Sharma, A., Sheather, S.J., 2003. A nonparametric model for stochastic generation of daily rainfall occurrence. *Water Resour. Res.* 39 (10). <http://dx.doi.org/10.1029/2003WR002182>.
- Herzer, H., 2003. Flooding in the pampean region of Argentina: the Salado basin. In: Kreimer, A., Arnold, M., Carlin, y.A. (Eds.), *Building Safer Cities: The Future of Disaster Risk. Disaster Risk Management Series. The World Bank, Disaster Management Facility*, Washington, D.C., pp. 137–147.
- Hurrell, J.W. et al., 2009. Decadal Climate Prediction: Opportunities and Challenges, *OceanObs'09*, Lido, Venice.
- Katz, R.W., 1977. Precipitation as a chain-dependent process. *J. Appl. Meteorol.* 16 (7), 671–676. [http://journals.ametsoc.org/doi/abs/10.1175/1520-0450\(1977\)016%3C0671%3APAACDP%3E2.0.CO%3B2](http://journals.ametsoc.org/doi/abs/10.1175/1520-0450(1977)016%3C0671%3APAACDP%3E2.0.CO%3B2).
- Katz, R.W., Parlange, M.B., Naveau, P., 2002. Statistics of extremes in hydrology. *Adv. Water Resour.* 25, 1287–1304. [http://dx.doi.org/10.1016/S0309-1708\(02\)00056-8](http://dx.doi.org/10.1016/S0309-1708(02)00056-8).
- Kilsby, C.G., Jones, P.D., Burton, a., Ford, a.C., Fowler, H.J., Harpham, C., et al., 2007. A daily weather generator for use in climate change studies. *Environ. Modell. Softw.* 22 (12), 1705–1719. <http://dx.doi.org/10.1016/j.envsoft.2007.02.005>.
- Kim, Y., Katz, R.W., Rajagopalan, B., Podestá, G.P., Furrer, E.M., 2012. Reducing overdispersion in stochastic weather generators using a generalized linear modeling approach. *Clim. Res.* 53, 13–24. <http://dx.doi.org/10.3354/cr101071>.
- Kleiber, W., Katz, R.W., Rajagopalan, B., 2012. Daily spatiotemporal precipitation simulation using latent and transformed Gaussian processes. *Water Resour. Res.* 48 (1), 1–17. <http://dx.doi.org/10.1029/2011WR011105>.
- Kleiber, W., Katz, R.W., Rajagopalan, B., 2013. Daily minimum and maximum temperature simulation over complex terrain. *Ann. Appl. Stat.* 7 (1), 588–612. <http://dx.doi.org/10.1214/12-AOAS602>.
- McCullagh, P., Nelder, J.A., 1989. *Generalized Linear Models*. Chapman and Hall, London.
- McGinnis, S., Nychka, D., Mearns, L.O., 2015. A new distribution mapping technique for climate model bias correction. In: *Machine learning and data mining approaches to climate science: proceedings of the Fourth International Workshop on Climate Informatics*, Springer (to appear).
- Meehl, G.A., Goddard, L., Murphy, J., Stouffer, R.J., Boer, G., Danabasoglu, G., et al., 2009. Decadal prediction: can it be skilful? *Bull. Am. Meteorol. Soc.* 90, 1467–1485. <http://dx.doi.org/10.1175/2009BAMS2778.1>.
- Mehrotra, R., Sharma, A., 2007. A semi-parametric model for stochastic generation of multi-site daily rainfall exhibiting low-frequency variability. *J. Hydrol.* 335 (1–2), 180–193. <http://dx.doi.org/10.1016/j.jhydrol.2006.11.011>.
- Meza, F.J., 2005. Variability of reference evapotranspiration and water demands. Association to ENSO in the Maipo river basin, Chile. *Glob. Planet. Change* 47, 212–220. <http://dx.doi.org/10.1016/j.gloplacha.2004.10.013>.
- Montecinos, A., Díaz, A.F., Aceituno, P., 2000. Seasonal diagnostics and predictability of rainfall in subtropical South America based on tropical Pacific SST. *J. Clim.* 13, 746–758.
- Penalba, O.C., Rivera, J.a., Pántano, V.C., 2014. The CLARIS LPB database: constructing a long-term daily hydro-meteorological dataset for La Plata Basin, Southern South America. *Geosci. Data J.* 1, 20–29. <http://dx.doi.org/10.1002/gdj3.7>.
- Podestá, G., Bert, F., Rajagopalan, B., Apipattanavis, S., Laciana, C., Weber, E., et al., 2009. Decadal climate variability in the Argentine Pampas: regional impacts of plausible climate scenarios on agricultural systems. *Clim. Res.* 40, 199–210. <http://dx.doi.org/10.3354/cr00807>.
- Qian, B., Gameda, S., De Jong, R., Falloon, P., Gornall, J., 2010. Comparing scenarios of Canadian daily climate extremes derived using a weather generator. *Clim. Res.* 41, 131–149. <http://dx.doi.org/10.3354/cr00845>.
- Rajagopalan, B., Lall, U., 1999. A k-nearest-neighbor simulator for daily precipitation. *Water Resour. Res.* 35 (10), 3089–3101.
- Rajagopalan, B., Lall, U., Tarboton, D.G., 1997. Evaluation of kernel density estimation methods for daily precipitation resampling. *Stoch. Hydrol. Hydraul. Res.* 11 (6), 523–547. <http://dx.doi.org/10.1007/BF02428432>.
- Richardson, C.W., 1981. Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resour. Res.* 17 (1), 182–190. <http://onlinelibrary.wiley.com/doi/10.1029/WR017i001p00182/full>.
- Ropelewski, C.F., Bell, M.A., 2008. Shifts in the statistics of daily rainfall in South America conditional on ENSO phase. *J. Clim.* 21 (5), 849–865.
- Ropelewski, C.F., Halpert, S., 1987. Global and regional scale precipitation patterns associated with the El Niño-Southern Oscillation. *Mon. Weather Rev.* 115, 1606–1626.
- Saha, S. et al., 2006. The NCEP climate forecast system. *J. Clim.* 19 (15), 3483–3517.
- Schoof, J.T., Arguez, a., Brolley, J., O'Brien, J.J., 2005. A new weather generator based on spectral properties of surface air temperatures. *Agric. For. Meteorol.* 135, 241–251. <http://dx.doi.org/10.1016/j.agrformet.2005.12.004>.
- Schreinemachers, P., Berger, T., 2011. An agent-based simulation model of human-environment interactions in agricultural systems. *Environ. Modell. Softw.* 26 (7), 845–859. <http://dx.doi.org/10.1016/j.envsoft.2011.02.004>.
- Seager, R., Naik, N., Baethgen, W.E., Robertson, A.W., Kushnir, Y., Nakamura, J.y., Jurburg, S., 2010. Tropical oceanic causes of interannual to multidecadal precipitation variability in southeast South America over the past century. *J. Clim.* 23, 5517–5539.
- Sharif, M., Burn, D.H., 2007. Improved K-nearest neighbor weather generating model. *J. Hydrol. Eng.* 12 (1), 42–52.
- Skansi, M.M., Núñez, S.E., Podestá, G.P., Garay, N., 2009. La sequía del año 2008 en la región húmeda argentina descripta a través del Índice de Precipitación Estandarizado. *CONGREGMET 2009*, Buenos Aires, Argentina.
- Stainforth, D.A., Allen, M.R., Tredger, E.R.y., Smith, L.A., 2007. Confidence, uncertainty and decision-support relevance in climate predictions. *Philos. Trans. Roy. Soc. A: Math., Phys. Eng. Sci.* 365 (1857), 2145–2161.
- Stern, R.D., Coe, R., 1984. A model fitting analysis of daily rainfall data. *J. Roy. Stat. Soc.* 147 (1), 1–34. <http://www.jstor.org/stable/2981736>.
- Stockdale, T.N. et al., 2010. Understanding and predicting seasonal-to-interannual climate variability – the producer perspective. *Proc. Environ. Sci.* 1, 55–80.
- Trenberth, K.E., Stepaniak, D.P., 2001. Indices of El Niño evolution. *J. Clim.* 14, 1697–1701.
- Verdin, A., Rajagopalan, B., Kleiber, W., Katz, R.W., 2015. Coupled stochastic weather generation using spatial and generalized linear models. *Stoch. Env. Res. Risk Assess.* 29, 347–356. <http://dx.doi.org/10.1007/s00477-014-0911-6>.
- Viglizzo, E.F., Roberto, Z.E., Lertora, F., Gay, E.L.y., Bernardos, J., 1997. Climate and land-use change in field-crop ecosystems of Argentina. *Agric. Ecosyst. Environ.* 66 (1), 61–70.
- Viglizzo, E.F., Jobbágy, E.G., Frank, F.C., Aragón, R., De Oro, L.y., Salvador, V., 2009. The dynamics of cultivation and floods in arable lands of Central Argentina. *Hydrol. Earth Syst. Sci.* 13, 491–502.
- Wheater, H.S., Chandler, R.E., Onof, C.J., Isham, V.S., Bellone, E., Yang, C., et al., 2005. Spatial-temporal rainfall modelling for flood risk estimation. *Stoch. Env. Res. Risk Assess.* 19 (6), 403–416. <http://dx.doi.org/10.1007/s00477-005-0011-8>.
- Wilby, R.L., Conway, D., Jones, P.D., 2002. Prospects for downscaling seasonal precipitation variability using conditioned weather generator parameters. *Hydrol. Process.* 16 (October 2001), 1215–1234. <http://dx.doi.org/10.1002/hyp.1058>.
- Wilks, D.S., 2008. High-resolution spatial interpolation of weather generator parameters using local weighted regressions. *Agric. For. Meteorol.* 148 (1), 111–120. <http://dx.doi.org/10.1016/j.agrformet.2007.09.005>.
- Wilks, D.S., 2009. A gridded multisite weather generator and synchronization to observed weather data. *Water Resour. Res.* 45 (10), 1–11. <http://dx.doi.org/10.1029/2009WR007902>.
- Yan, Z., Bate, S., Chandler, R.E., Isham, V., Wheater, H., 2002. An analysis of daily maximum wind speed in northwestern Europe using generalized linear models. *Am. Meteorol. Soc.* 15, 2073–2088.
- Yang, C., Chandler, R.E., Isham, V.S., Wheater, H.S., 2005. Spatial-temporal rainfall simulation using generalized linear models. *Water Resour. Res.* 41 (11). <http://dx.doi.org/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 (7). <http://dx.doi.org/10.1029/2002WR001769>.