Vol. 59: 61–75, 2014
doi: 10.3354/cr01201

CLIMATE RESEARCH
Clim Res

Published February 4

Validation of a stochastic temperature generator focusing on extremes, and an example of use for climate change

S. Parey[1,*], T. T. H. Hoang[1], D. Dacunha-Castelle[2]

[1]EDF/R&D, 6 quai Watier, 78 401 Chatou Cedex, France
[2]Laboratoire de Mathématiques, Université Paris 11, Orsay, France

ABSTRACT: We present a stochastic seasonal functional heteroscedastic auto-regressive model developed to simulate daily (minimum, maximum, or mean) temperature time series coherent with observed time series and designed to reliably reproduce extreme values through a careful study of extremes and the fact that the tails of the distribution are bounded. The model was first validated using different daily minimum and maximum weather station time series over Eurasia and the US in different climatic regions. The model was able to produce coherent results both for the bulk of the distribution and for its extremes and was able to produce higher or lower extreme values than observed. A possible use in the climate change context was then tested. We fit the model over the first part of a long temperature time series and then used it to simulate a large number of possible trajectories for the second part when temperature increased. Two approaches were tested to do so, one based on a simple mean change in mean and variance and the other in considering the full seasonalities and trends estimated over the observed second part of the time series. Both approaches yielded good results, both for the bulk and for the extremes of the temperature distribution over the second part of the period. However, the second approach allowed us to take interannual variability changes into account, which leads to more realistic results when this occurs. Our results support the use of this tool for statistical downscaling, enabling the reliable reproduction of temperature extremes.

KEY WORDS: Daily temperature · Stochastic modeling · Extremes

Resale or republication not permitted without written consent of the publisher

1. INTRODUCTION

Weather generators are commonly used in environmental or financial studies as a way to simulate key properties of observed meteorological records and then produce long series of daily weather parameters. Two main approaches can be found in those developments: weather generators are either based on randomly pooling out analog days in a database of past observations, or on statistically generating the desired variables with a stochastic model whose parameters are estimated on a database of past observations. The advantage of the first approach is a better reproduction of the observed distribution, but the main drawback is that it cannot reproduce non-observed values. Although the second approach is based on parametric or semi-parametric definitions of the distributions, its main advantage is its ability to produce physically realistic unobserved situations. This second approach is preferred here, as the focus is on extreme events.

Most efforts in weather generator developments have been devoted to precipitation (see Wilks & Wilby 1999 for a review). Precipitation is a crucial parameter in many environmental studies and its representation is complicated by its intermittent nature. Here again, different approaches can be found. Cowpertwait et al. (2007) proposed a model of storm cells whose occurrence follows a Poisson pro-

*Corresponding author: sylvie.parey@edf.fr

© Inter-Research 2014 · www.int-res.com

cess, during which rain cells occur as a secondary Poisson process. Other generators are based on different daily states, from the simple dry and wet days to more sophisticated weather type definitions, possibly introduced as a hidden state variable using hidden Markov models (Ailliot et al. 2009, Sansom & Thompson 2010). Furthermore, following Richardson (1981), weather generators are developed to simultaneously represent precipitation and other variables like temperature (daily minimum and maximum), solar radiation, or wind, for use in agricultural studies. Such models are increasingly used to downscale global climate model results in impact studies (Wilks 1992, Semenov & Barrow 1997, Wilks & Wilby 1999, Hansen 2002, Kyselý & Dubrovský 2005, Semenov 2008) because they allow changes in variability to be taken into account. The interest in extremes further motivates the use of such models; however, they generally must be improved to adequately reproduce extreme events (Furrer & Katz 2008). Semenov (2008) showed that if precipitation extremes are reasonably well represented by a Richardson type generator (called LARS-WG), temperature extremes are generally not, because the normality assumption used for the residuals is not universally true. Even with the use of weather types and skewed normal distributions (WACS-Gen), Flecher et al. (2010) reported difficulties in reproducing extreme events.

Stochastic temperature models are also used in the framework of weather derivatives. Weather derivative products provide protections against 'weather risk,' that is against the unpredictable component of weather fluctuations, called 'weather surprises' or 'weather noise.' This thus necessitates some knowledge of this 'weather noise' over space and time, which motivated the development of stochastic models (Campbell & Diebold 2005, Mraoua & Bari 2007, Benth & Saltyte-Benth 2011).

Extreme events are important for industrial adaptation, e.g. for the design and operation of installations, such as dams designed to withstand extreme flood events, or overhead lines designed to withstand storms. Our goal is then to propose a temperature generator able to correctly reproduce temperature extremes. The general principle of such stochastic models, whatever their usage, consists of modeling the temperature (daily maximum, minimum, or mean) as the summation of a deterministic part and a stochastic process, designed to represent the random fluctuations around the mean:

$$X(t) = \Lambda(t) + \Phi(t)Z(t) \quad (1)$$

where $\Lambda(t)$ and $\Phi(t)$ are deterministic and $Z(t)$ is stochastic. $\Lambda(t)$ contains at least a seasonal component, and usually also a trend component. $\Phi(t)$ is most often 1. The stochastic part generally presents an autoregressive structure, more or less sophisticated: from a first-order auto-regressive (AR1) to a general auto-regressive conditional heteroscedastic (GARCH).

For the proposed model, our basic idea comes from a preliminary analysis of the correlations and especially from the shape of the conditional variance of $Z(t)$ when $Z(t-1)$ is fixed. In particular, this conditional variance drastically decreases outside of a bounded interval. This leads to the use of a functional auto-regressive conditional heteroscedastic (FARCH) model, the simplest one able to take this behavior into account. FARCH processes are the first-order Euler scheme approximation of the discrete Markov chain given by the sequence of discrete observations of a diffusion. Furthermore, the coefficients (drift and diffusion) of the diffusion are those of the FARCH process. Thus we are led to consider temperature as a continuous time process with continuous trajectories. If $X(t)$ can be assumed as Markovian, then the continuous time process is a diffusion. The Markovian property can be tested. This mathematical justification is coherent with the physical interpretation of the heat equation as a diffusion of the thermal energy but also with more general considerations on nonlinearity and stochasticity which can be found in Sura (2012). The building of the model is based on discrete temperature observations at a given time interval, for instance every day, and the diffusive property has to be translated in this restrictive framework. The obtained seasonal functional heteroscedastic auto-regressive (SFHAR) model, with careful treatment of the extreme upper and lower bounds, is described in detail by Dacunha-Castelle et al. (2013) and briefly reviewed in Appendix 1. Here we focus on the validation of the model for different climates in Eurasia and the USA and propose a possible application in the climate change context. The model is calibrated on temperature time series starting in 1950 for the USA and Eurasia. It simulates the residuals after accounting for seasonalities and trends in mean and variance.

After a brief description of the model and the presentation of the temperature time series in Section 2, Section 3 is devoted to the validation of the model for different climates. In Section 4, the model is calibrated on the first part of the observed time series, and then different strategies are tested and validated to simulate the second part, warmer on average than the first one. Discussion and perspectives are proposed in Section 5.

2. MODEL AND OBSERVATIONS

In the following, $X(t)$ is the observed temperature time series (either daily minimum or daily maximum temperature), $m(t)$ its mean trend, $S_m(t)$ the seasonality of the mean, $s^2(t)$ its variance trend, $S_v(t)$ the seasonality of the variance, and $Z(t)$ the modeled residual time series.

2.1. Brief description of the model

2.1.1. Pre-processing

As stated before, the model is designed to simulate the residuals $Z(t)$ from a temperature time series $X(t)$ after accounting for seasonalities — ($S_m(t)$ and $S_v(t)$) — and trends — ($m(t)$ and $s(t)$) — in mean and standard deviation. The first step is then to identify and remove these deterministic parts from $X(t)$ to obtain $Z(t)$. This is done through the following succession of steps:

(1) Estimation of the seasonality of X(t): $\hat{S}_m(t)$

(2) Estimation of the trend $\hat{m}(t)$ from the time series $\left(X(t)-\hat{S}_m(t)\right)$

(3) Estimation of the seasonality of the variance from $\left[X(t)-\hat{S}_m(t)-\hat{m}(t)\right]^2 : \hat{S}_v^2(t)$

(4) Estimation of the trend $\hat{s}^2(t)$ from the time series $\left[X(t)-\hat{S}_m(t)-\hat{m}(t)\right]^2 / \hat{S}_v(t)$

(5) Finally, the residuals are estimated as

$$\hat{Z}(t)=\frac{X(t)-\hat{S}_m(t)-\hat{m}(t)}{\hat{S}_v(t)\hat{s}(t)} \quad (2)$$

Quantities with a hat (^) correspond to estimations. The identification of seasonality is based on the fitting of a trigonometric function of the form:

$$\theta_0+\sum_{i=1}^{p}\left(\theta_{i,1}\cos\frac{2\pi t}{365}+\theta_{i,2}\sin\frac{2\pi t}{365}\right),$$

where θ is the parameter of the function to be estimated and the number p of trigonometric terms is chosen through an Akaike criterion. This parametric identification has been compared to the nonparametric seasonal trends decomposition method (Cleveland et al. 1990), and both approaches have been found very similar.

The trend identification is conducted in a non parametric way by using the LOESS technique (local regression, Stone 1977). The LOESS estimator is obtained by locally fitting a dth degree polynomial to the data via weighted least squares. Throughout this work, the local linear fit is used, which means $d = 1$. This method implies the choice of a smoothing parameter, which controls the balance between goodness of fit to the data and smoothness of the regression function. The smoothing parameter is obtained through an automated selection. This selection is difficult here as the data are correlated, non-stationary, and heteroscedastic. The modified partitioned cross-validation technique proposed by Hoang (2010) is used. It is based on the classical partitioned cross-validation technique of Marron (1987): the observations are partitioned into g subgroups by taking every gth observation, for example the first subgroup consists of the observations 1, 1 + g, 1 + 2g,…, the second subgroup consists of the observations 2, 2 + g, 2 + 2g, …. The observations in each subgroup are then independent for high g. Chu & Marron (1991) defined the optimal asymptotic bandwidth for partitioned cross validation in the case of constant variance as $h_{PCV} = h_0 g^{1/5}$, with h_0 estimated as the minimizer of

$$PCV_g(h)=\frac{1}{g}\sum_{k=1}^{g}CV_{0,k}(h) \quad (3)$$

where $CV_{0,k}$ is the ordinary cross-validation score for the kth group. This approach has been modified to take heteroscedasticity into account. The optimal g then corresponds to the minimum of a more complicated expression (Hoang 2010), and in practice, it is preferred to estimate h_{MPCV} (the optimal bandwidth of the modified partitioned cross validation) for different values of g and to retain the values of g for which h_{MPCV} is not too bad (i.e. not too close to 0 and not higher than 0.7). For each g, the trends m and s are estimated by LOESS with bandwidth $\hat{h}^g_{MPCV}$ to obtain an estimator of the expression to minimize. The value of g corresponding to the minimum value is retained, giving the corresponding optimal bandwidth h_{MPCV}.

The order of estimation of seasonality and trend is not important, as estimating trends first and then seasonality leads to similar results for $Z(t)$. The procedure is illustrated in Fig. 1.

Careful studies of $Z(t)$ have shown that although seasonality has been removed from the mean and variance, some seasonality remains in the higher-order moments, such as skewness and kurtosis of $Z(t)$ and in its autocorrelations. However, no significant remaining trends could have been found in high-order moments, autocorrelations, or extremes of $Z(t)$.

2.1.2. Model for $Z(t)$

The proposed model is described in detail by Dacunha-Castelle et al. (2013) and summarized in

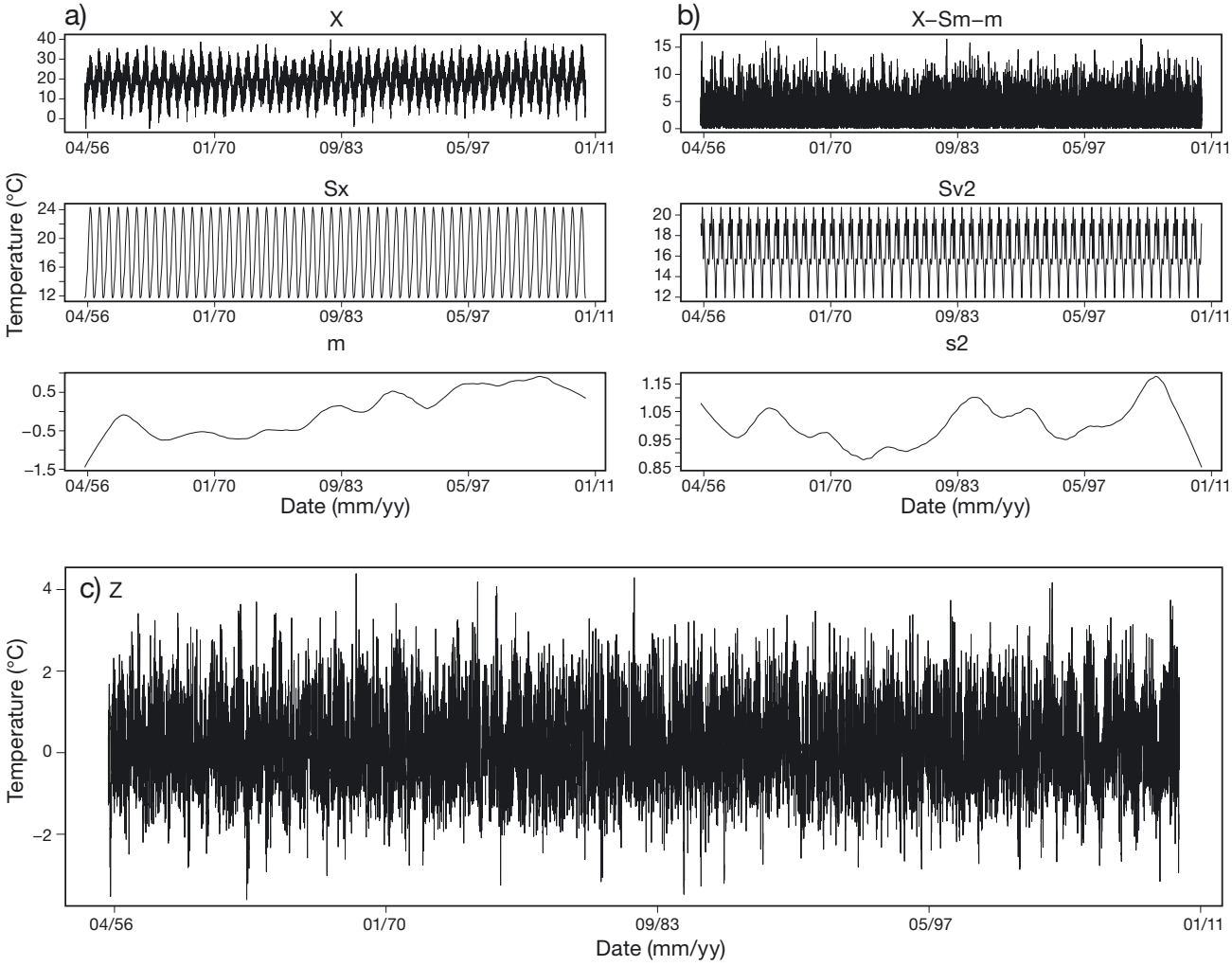

Fig. 1. Illustration of the derivation of the residuals from an observed daily temperature time series (Biarritz-Anglet, France, 1956–2009). (a) Original time series (X, top), its seasonality (S_X, middle), and trend (m, bottom), (b) time series of variance (top), its seasonality (middle), and trend (bottom), and (c) obtained time series of residuals (Z)

Appendix 1. The first step is to estimate the extremes of $Z(t)$. The upper and lower bounds r_1 and r_2, together with the corresponding shape parameters ξ_1 and ξ_2, are estimated by fitting a generalized extreme value (GEV) distribution to the minima and the maxima of $Z(t)$, respectively. The extremes of $Z(t)$ do not show any clear seasonality, and the fitting is done with 73 d blocks (5 blocks yr^{-1}). Sensitivity tests on the choice of block length showed that the results do not significantly differ. The shape parameter is negative, thus the distributions are bounded. However, if it is too close to 0, the simulation may be problematic. If this happens, it is advised to slightly change the block length in order to get a better estimate of this parameter.

The proposed model is then justified. It consists of a modification of a SFHARl of the form:

$$Z(t) = bZ(t-1) + a(t, Z(t-1))\varepsilon_t,\ \varepsilon_t \qquad (4)$$

being a normal distribution with 0 mean and unit variance, and:

$$b = \theta_0 + \sum_{j=1}^{p_1}(\theta_{j,1}\cos\frac{2\pi t}{365} + \theta_{j,2}\sin\frac{2\pi t}{365}) \qquad (5)$$

p_1 being chosen by an Akaike criterion, because seasonality remains in the autocorrelation, and a is estimated as a degree 5 trigonometric polynomial:

$$a^2(t, Z(t-1)) = (\hat{r}_2 - t)(t - \hat{r}_1)\sum_{k=0}^{5}\sum_{j=1}^{p_2}(\alpha_{1,k}^j\cos\frac{2\pi t}{365} + \alpha_{2,k}^j\sin\frac{2\pi t}{365})Z(t-1)^k \qquad (6)$$

under constraints

$$(a^2)'(r_1) = \frac{2b(r_1)}{1 - 1/\xi_1};\ (a^2)'(r_2) = \frac{2b(r_2)}{1 - 1/\xi_2} \qquad (7)$$

and $a^2(t) > 0 \forall t$, with p_2 chosen by an Akaike criterion, r_1 and r_2 being respectively the lower and upper bound of the extreme value distributions of $Z(t)$, and ξ_1 and ξ_2 the corresponding shape parameters. The

form of a and the constraints are given by the extreme value theory of the continuous time process (Davis 1982). In practice, the autoregressive part of $Z(t)$ is first estimated, then a is estimated from $(Z(t)-\hat{b}Z(t-1))^2$ by maximum likelihood with constraints.

Once the parameters have been estimated, as many sequences of $Z(t)$ as desired can be simulated with the model. A sequence consists of a certain number of years and each day t, $Z(t)$ is computed from $Z(t-1)$. The initial value is randomly selected from the observed residuals. A condition is added to insure that each $Z(t)$ remains inside the limit bounds r_1 and r_2: if the simulated value at time t exceeds the upper bound or is lower than the lower bound, it is disregarded and another value for $Z(t)$ is computed from $Z(t-1)$. This is equivalent to a modified model where the distribution of ε_t is a truncated normal distribution whose truncation depends on the value of $Z(t-1)$ (its values are $r_1 - bZ(t-1)/a(Z(t-1))$ and $r_2 - bZ(t-1)/a(Z(t-1))$). Thus the obtained simulated residuals are bounded.

Then a simulation of the initial temperature time series is obtained by re-introducing the estimated deterministic parts $\hat{S}_m(t), \hat{m}(t), \hat{S}_v(t)$ and $\hat{s}(t)$. As an indication, 100 simulations of a 60 yr daily time series need around 7 min of computing time on a standard laptop.

Compared to most generators found in the literature, our model differs in its bounded property and in the careful retrieval of the smoothing parameter to compute the nonparametric trends in both mean and variance to obtain the simulated residuals. The main consequence is thus that the length of the simulated time series is at most that of the observed one used to determine the trends. As many equivalent time series as desired can be computed, giving a similarly rich sample. The optimal smoothing parameter is linked to interannual variability, which allows an indirect consideration of this property of temperature time series besides daily variance. Furthermore, the autocorrelations are fully seasonal and the behavior of the extremes is carefully introduced in the volatility (or lag 0 auto-correlation) coefficients $a(t)$. This is expected to substantially improve the ability of the model to reproduce extremes, which will be examined in this paper.

2.2. Observed time series

Model validation is conducted for different climates in Eurasia and the US. For Eurasia, weather station time series of daily minimum temperature (TN) and daily maximum temperature (TX) are obtained from the ECA&D project database. The project gives indications of homogeneity through the results of different break identification techniques (Klein Tank et al. 2002). First, the series which could be considered as homogeneous (stated as 'useful' in the database) over the period 1950–2009 have been selected for both TN and TX. Then, only the time series with less than 5% missing data are kept, leading to 106 series for TX and 120 for TN (many TX series, mostly in Russia, have missing values from 2007 onward, whereas the corresponding TN ones have missing values only in 2009).

For the US, weather station TX and TN time series are obtained from the Global Historical Climatology Network – Daily Database (GHCN daily; Menne et al. 2012). A similar selection procedure left us with 86 series for TX and 85 for TN.

Among these time series, 4 weather stations corresponding to different climates, in terms of mean annual temperature, have been chosen for each continent, as listed in Table 1. As stated before, the weather station of Olekminsk in Russia cannot be considered for TX as it exhibits too many missing values. No other station with a similar climate to that of Olekminsk is available for TX.

Table 1. Temperature time series considered in this study, showing period lengths and mean observed temperatures. TN (TX) daily minimum (maximum) temperature

Weather station	TN		TX	
	Period	Mean annual mean (°C)	Period	Mean annual mean (°C)
Biarritz	1956–2009	10.1	1956–2009	17.7
Berlin	1950–2009	5.1	1950–2009	13.4
Petropavlovsk	1950–2009	–3.3	1950–2009	6.9
Olekminsk	1950–2009	–11.3	–	–
Death Valley	1962–2009	17.0	1962–2009	32.8
Charleston	1950–2009	15.4	1950–2009	23.0
Jacksonville	1950–2009	5.2	1950–2009	17.5
Glasgow	1950–2009	–0.7	1950–2009	12.5

3. VALIDATION

For each of the 7 (for TX) or 8 (for TN) temperature time series, the parameters of the model are fitted over the whole period. We then compute 100 simulations of the model for each location, and the results are compared to the observed time series both for the representation of the bulk of the distribution and of its warm and cold extremes.

3.1. Bulk of the distribution

Tables 2 & 3 summarize the comparison of the mean, variance, skewness, and kurtosis of the distributions of each temperature time series obtained from the observations and from the 100 model simulations (mean value and 95% confidence interval). The results show that these different moments of the observed distribution of daily maximum or minimum temperature are correctly reproduced by the stochastic simulations, although the higher moments are sometimes less accurately reproduced. This good result may be linked to the domination of the annual cycle, thus the seasonal distributions are also compared. Fig. 2 shows the Q-Q plots of observed and simulated winter and summer distributions of TN in Olekminsk and TX in Death Valley. Similar results for the other stations confirm that the model reasonably reproduces the seasonal temperature distributions.

Fig. 3 shows that the mean annual cycle, as well as that of the standard deviation, is faithfully represented. Fig. 3 is for TN in Berlin and TX in Jacksonville, but similar results are found for each individual time series. Kolmogorov-Smirnov tests were applied to compare the distributions obtained for each day of the year between observations on the

Table 2. Mean, variance, skewness, and kurtosis estimated from observed (Obs) and simulated (Sim) daily maximum temperature (TX) time series. Simulations: values are mean with 95% confidence interval in brackets

	Mean		Variance		Skewness		Kurtosis	
	Obs	Sim	Obs	Sim	Obs	Sim	Obs	Sim
Berlin	13.4	13.4 (13.3; 13.5)	84.1	83.2 (80.1; 85.9)	−0.03	−0.02 (−0.06; 0.02)	−0.78	−0.79 (−0.85; −0.72)
Biarritz	17.7	17.7 (17.6; 17.8)	37.1	37.2 (35.9; 38.7)	0.07	0.05 (−0.01; 0.10)	−0.06	−0.23 (−0.33; −0.10)
Petropavlovsk	6.9	6.9 (6.7; 7.1)	237.8	238.2 (233.6; 243.9)	−0.18	−0.16 (−0.18; −0.13)	−1.03	−1.08 (−1.11; −1.05)
Olekminsk	–	–	–	–	–	–	–	–
Death Valley	32.8	32.8 (32.6; 32.9)	112.2	112.2 (109.5; 114.8)	−0.08	−0.07 (−0.10; −0.04)	−1.19	−1.17 (−1.21; −1.11)
Jacksonville	17.5	17.4 (17.3; 17.6)	137.1	136.3 (133.2; 139.7)	−0.43	−0.39 (−0.41; −0.36)	−0.82	−0.87 (−0.91; −0.82)
Glasgow	12.5	12.5 (12.2; 12.7)	204.3	202.8 (196.6; 208.7)	−0.38	−0.33 (−0.37; −0.30)	−0.59	−0.63 (−0.69; −0.58)
Charleston	23	23 (22.9; 23.1)	50.9	50.7 (49.6; 52.0)	−0.46	−0.41 (−0.44; −0.39)	−0.48	−0.57 (−0.63; −0.51)

Table 3. As in Table 2, but for daily minimum temperature (TN)

	Mean		Variance		Skewness		Kurtosis	
	Obs	Sim	Obs	Sim	Obs	Sim	Obs	Sim
Berlin	5.1	5.1 (5.0; 5.2)	49.1	48.5 (46.4; 50.9)	−0.36	−0.33 (−0.39; −0.26)	−0.21	−0.28 (−0.48; −0.06)
Biarritz	10.1	10.1 (10.0; 10.2)	30.2	30.1 (29.1; 31.1)	−0.35	−0.35 (−0.40; −0.31)	−0.3	−0.31 (−0.41; −0.21)
Petropavlovsk	−3.3	−3.3 (−3.4; −3.1)	196.6	196.7 (191.4; 202.9)	−0.41	−0.39 (−0.42; −0.37)	−0.85	−0.87 (−0.91; −0.81)
Olekminsk	−11.4	−11.3 (−11.5; −11.1)	335.4	334.9 (325.6; 344.7)	−0.3	−0.29 (−0.31; −0.27)	−1.14	−1.13 (−1.17; −1.09)
Death Valley	16.9	16.9 (16.8; 17.0)	102.9	102.7 (101.3; 104.3)	0.02	0.03 (0.00; 0.05)	−1.13	−1.11 (−1.14;−1.09)
Jacksonville	5.2	5.2 (5.0; 5.3)	110	109.7 (106.3; 113.9)	−0.33	−0.31 (−0.35; −0.27)	−0.58	−0.62 (−0.71; −0.53)
Glasgow	−0.7	−0.7 (−0.9; −0.5)	147.1	146.9 (141.1; 152.0)	−0.58	−0.54 (−0.58; −0.50)	−0.25	−0.3 (−0.40; −0.19)
Charleston	15.4	15.4 (15.3; 15.5)	60	60 (58.6; 61.6)	−0.4	−0.37 (−0.40; −0.34)	−0.86	−0.9 (−0.94; −0.85)

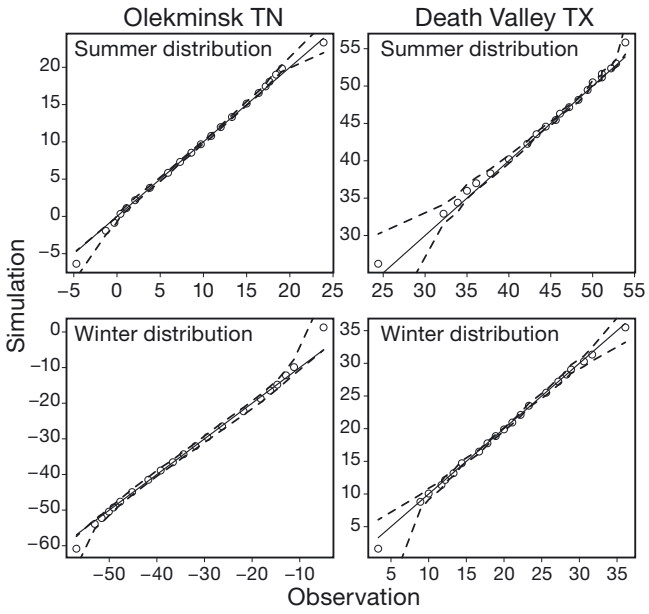

Fig. 2. Q-Q plots of the summer and winter distributions for (a) daily minimum temperature in Olekminsk and (b) daily maximum temperature in Death Valley. Solid line: diagonal (i.e. 1:1 relationship); circles: mean simulation; dashed lines: 95% confidence interval of the simulations

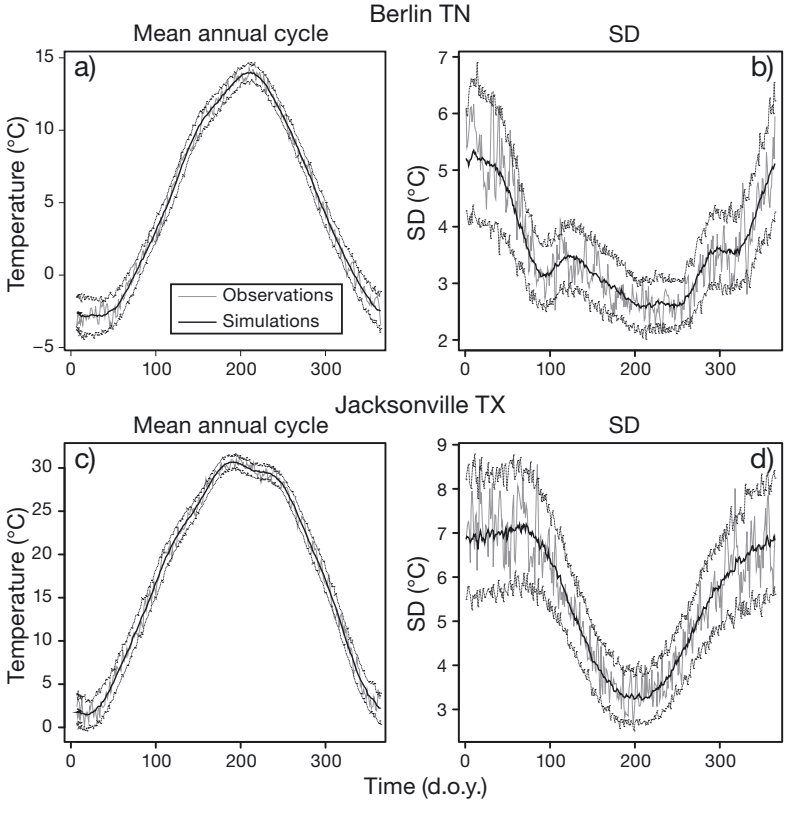

Fig. 3. Observed (grey) and simulated (black, where the solid line shows the mean and the dashed line shows the 95% confidence interval) mean annual cycle and daily standard deviation annual cycle for (a,b) daily minimum temperature in Berlin and (c,d) daily maximum temperature in Jacksonville. d.o.y.: day of the year

one hand and simulations on the other, and they show that the distributions can be considered similar with a 95% confidence level. The proposed stochastic model is thus able to correctly reproduce the bulk of the daily minimum or maximum temperature distributions for different climates.

3.2. Extremes

The model is constructed for a bounded variable and the simulations are made in such a way that each simulated value remains inside the estimated bounds of the residuals. Thus first, the GEV distribution parameters for the simulated residuals, are compared to those of the observed ones, both for the lowest and the highest extremes. Fig. 4 shows the distributions for each parameter (location μ, scale σ, and shape ξ) obtained from the 100 simulations for the highest (warm) extremes and the lowest (cold) extremes together with the same parameters obtained from the observed residuals (grey line) for TN in Berlin and TX in Death Valley. The results show that the shape parameter is generally better reproduced in the simulations than the location and scale parameters; this is also true for the other temperatures and locations. It can be mathematically proven that the proposed stochastic model is able to produce the correct shape parameter when a truncated normal distribution is used for ε_t.

Table 4 compares the 50 yr return levels (RLs) of the maxima of TX and the minima of TN for the different locations over the whole observation period. The estimation is made by fitting a GEV to the block maxima of summer TX or winter TN (Coles 2001) with the maximum likelihood method and considering the choice of 2 blocks season^{-1} as a reasonable bias/variance compromise. The estimation is conducted as if the extremes would not present trends over the entire period, which is of course wrong, but it simplifies the computations and is sufficient to give a first view of the representation of the extremes by the proposed model. For each of the 100 simulations, the 50 yr RLs are computed. The given confidence interval is obtained as the 2.5th and 97.5th percentiles of the dis-

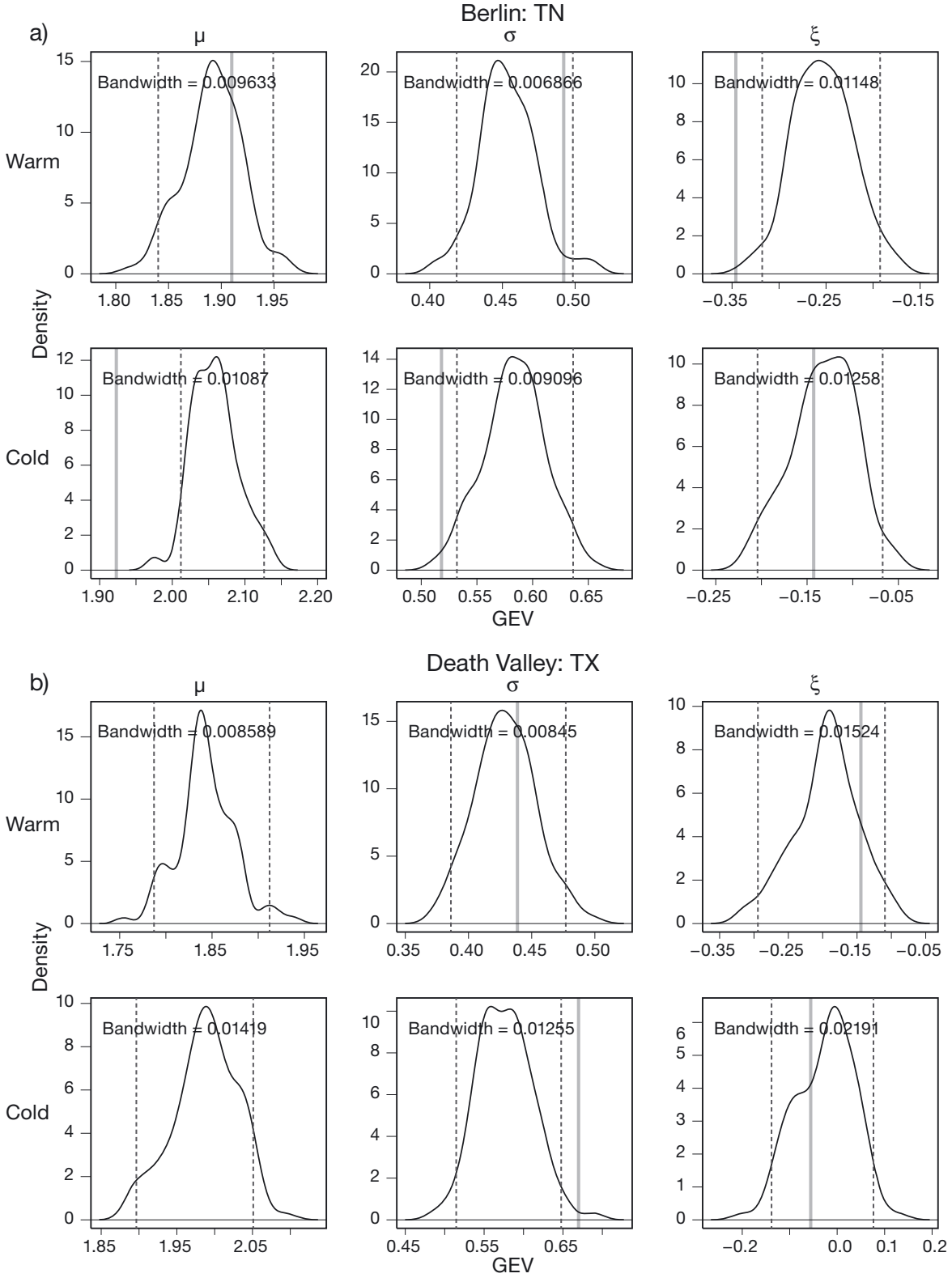

Fig. 4. Distributions of the parameters of the generalized extreme value (GEV) distribution fitted to the 100 simulations of the residuals (black): warm and cold extremes (location μ, scale σ, and shape ξ from left to right) with their 2.5th and 97.5th percentiles (black dotted lines) and value of the same parameters obtained from the observations (grey line). (a) Daily minimum temperature in Berlin and (b) daily maximum temperature in Death Valley

Table 4. 50 yr return levels (RLs) estimated from observed and simulated time series. For observations, the 95% confidence interval (in brackets) is obtained with the delta-method; for simulations, the given interval corresponds to the 2.5th and 97.5th percentiles of the distribution of the 100 obtained 50 yr RLs. TX (TN) daily maximum (minimum) temperature

	TX		TN	
	Observations	Simulations	Observations	Simulations
Berlin	38.2 (37.1; 39.2)	39.8 (38.8; 41.0)	−23.4 (−25.5; −21.0)	−26.5 (−31.5; −22.9)
Biarritz	39.6 (38.8; 40.4)	41.0 (39.0; 43.5)	−9.4 (−12.2; −6.6)	−11.0 (−12.6; −9.7)
Petropavlovsk	38.5 (37.6; 39.5)	41.5 (39.3; 44.8)	−43.7 (−45.2; −42.1)	−48.7 (−52.5; −45.3)
Olekminsk	–	–	−56.3 (−57.8; −54.8)	−58.8 (−61.4; −56.2)
Death Valley	54.3 (53.5; 55.1)	55.2 (54.3; 56.1)	−6.4 (−7.5; −5.3)	−7.4 (−8.8; −6.0)
Jacksonville	41.8 (40.3; 43.3)	43.1 (41.5; 44.5)	−29.5 (−31.3; −27.7)	−33.8 (−38.5; −30.6)
Glasgow	42.0 (41.1; 42.8)	45.5 (44.3; 46.9)	−42.9 (−44.4; −41.4)	−46.9 (−50.4; −44.0)
Charleston	39.5 (38.6; 40.4)	40.3 (39.5; 41.2)	−11.3 (−13.7; −9.0)	−8.8 (−10.0; −7.5)

tribution of the 100 RLs, whereas for the observations the confidence interval is the 95% one given by the delta-method (that is based on the asymptotic normality of the maximum likelihood estimators). Generally, the simulations give higher warm RLs and lower cold RLs than observed, but the confidence intervals obtained from the observations generally show some overlapping with the 2.5th and 97.5th percentiles of the distribution obtained from the simulations (except for TX in Glasgow and TN in Petropavslosk). The fact that the model produces higher (or lower for cold temperature) extremes than observed is not surprising because the simulations produce 100 possible realities, among which higher or lower extremes could have been observed. This thus shows that the model is not only able to produce extremes, but it can also produce more extreme extremes than observed.

Finally, we investigate the ability of the model to produce heat or cold waves. Cold waves are defined as periods of consecutive days with daily minimum temperature lower than the 2nd percentile and heat waves as periods of consecutive days with daily maximum temperature above the 98th percentile. The number of consecutive days varies between 1 and 15, the last class corresponding to the few episodes with more than 15 d, where these occur. Thus for each location, the 2nd and 98th percentiles of the observed time series are computed and the distribution of episodes in the observed time series is compared to the minimum, maximum, and mean frequencies of such a distribution in the 100 simulations. Fig. 5 shows the results for cold waves in Petropavslovsk and heat waves in Charleston. Even though the stochastic model tends to overestimate the proportion of 1 d cold spells compared to the observations, it is still able to produce longer episodes in a reasonable proportion, even for the longest events. This tendency to overestimate the frequency of 1 d events is less systematic for heat waves.

4. POSSIBLE USE IN THE CLIMATE CHANGE CONTEXT

The previous section has shown that the proposed stochastic model, when fitted on a temperature time series, is able to correctly reproduce the bulk of the distribution as well as the extremes of the studied time series. This is an interesting result as far as the model allows reliable simulations of a high number of possible temperature evolutions at a given location giving access to potential unobserved but still possible levels. In the climate change context, it could also be very interesting to produce possible temperature evolutions for the future, given that climate is warming. General or regional climate models are designed to allow such projections for the different climatic variables, but their ability to represent extreme values for a precise location is still questionable. Thus different downscaling techniques, from simple bias corrections to full dynamical downscaling with limited area models, are explored (Maraun et al. 2010). The aim here is to check whether the proposed stochastic model can be used as a statistical downscaling tool giving reliable indications of temperature extremes.

4.1. Simulation procedure

To do so, among the previously used temperature time series, 2 have been selected as showing an identifiable break in the evolution of mean temperature, splitting the time series in 2 sub-series of roughly similar length. The break is identified using the Mudelsee (2009) method which consists of selecting the date for which the standard deviation of the residuals resulting from the 2-phase regression model is the minimum, after having considered all dates (except the first and last 5, to avoid edge effects) as

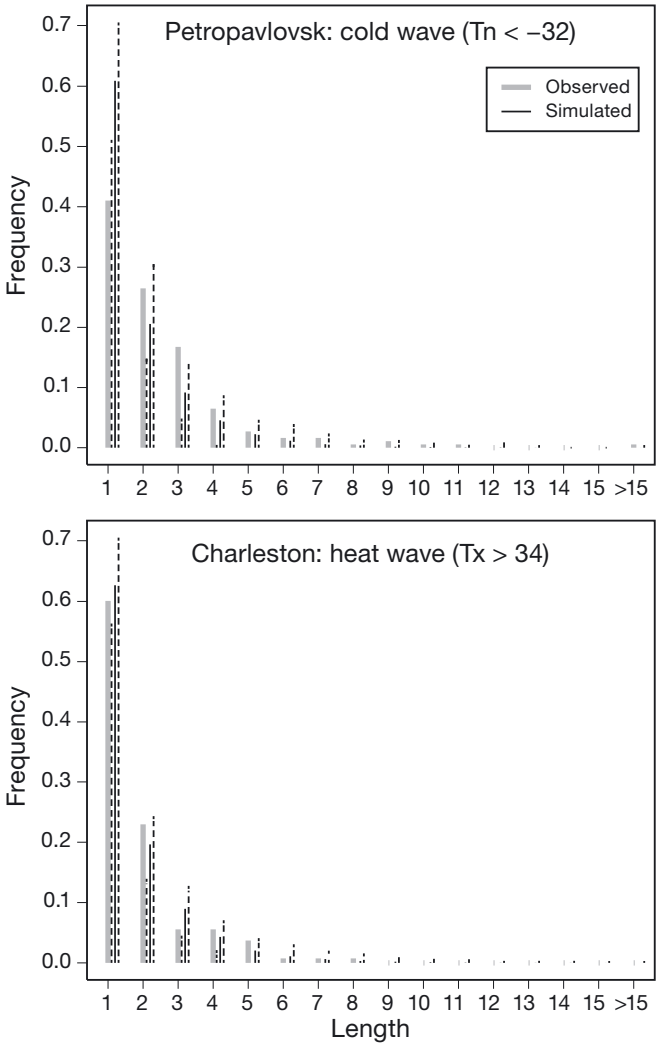

Fig. 5. Frequencies of the 1 to >15 d long (a) cold waves in Petropavlovsk and (b) heat waves in Charleston. A cold wave is defined as consecutive days with daily minimum temperature lower than the 2nd percentile of the observations and a heat wave as consecutive days with daily maximum temperature higher than the 98th percentile of the observations. Solid black lines: mean frequencies obtained from the simulations; dashed lines: minimum and maximum frequencies; grey lines: observed frequencies

potential break points. This simple technique is used because the identification of the break is not the ultimate goal of the work but is only made for the sake of illustration. More general regression techniques exist, such as the segmented regression proposed by Muggeo (2003). Such a break is identified in the year 1980 for TN in Berlin and 1985 for TX in Death Valley for both the mean and variance evolutions.

Each time series is then split into 2 shorter time series: 1950–1980 and 1981–2009 for TN in Berlin and 1962–1985 and 1986–2009 for TX in Death Valley. For both sub-series, the residuals $Z(t)$, after removing trends and seasonalities in mean and variance, are estimated. The parameters of the stochastic model defined to simulate $Z(t)$ are fitted over the first sub-series in each case. The reconstruction of the desired temperature time series for each period necessitates that trends and seasonalities are added to the simulated residuals. Two approaches are compared to compute the desired temperature time series over the second (and warmer) sub-period:

(1) Average mean and variance changes are added to the trends computed from the first sub-period: if m_1 is the mean over the first period, s_1 the standard deviation, and m_2 and s_2 the same quantities for the second period:

$$X_2(t) = \hat{S}_{m_1}(t) + \hat{m}_1(t) + (m_2 - m_1) + \hat{S}_{v_1}(t) \times \hat{s}_1(t) \times \frac{s_2}{s_1} \times Z(t) \quad (8)$$

where $\hat{S}_{m_1}(t) + \hat{m}_1(t)$, $\hat{S}_{v_1}(t)$ and $\hat{s}_1(t)$ are the seasonalities and trends estimated over the first sub-period. Table 5 summarizes the means and variances of each sub-series.

(2) Seasonalities and trends are those computed over the second sub-period:

$$X_2(t) = \hat{S}_{m_2}(t) + \hat{m}_2(t) + \hat{S}_{v_2}(t) \times \hat{s}_2(t) \times Z(t) \quad (9)$$

where $\hat{S}_{m_2}(t) + \hat{m}2(t)$, $\hat{S}_{v_2}(t)$ and $\hat{s}_2(t)$ are the seasonalities and trends estimated over the second sub period.

In the first approach, interannual variability, included in the smoothing parameter of the nonparametric trends, remains that of the first period, whereas the second approach allows interannual variability of the second period to be taken into account.

4.2. Results

4.2.1. Bulk of the distribution

As before, the first comparisons aim at validating the reproduction of the main characteristics of the bulk of the distribution. Table 6 gives the observed

Table 5. Mean and SD estimated for the first (m_1 and s_1) and second part (m_2 and s_2) of the time series. The first part corresponds to 1950–1980 for Berlin and 1962–1985 for Death Valley and the second part corresponds to 1981–2009 for Berlin and 1986–2009 for Death Valley

	m_1 (°C)	m_2 (°C)	s_1 (°C)	s_2 (°C)
TN Berlin	4.7	5.5	7.0	6.9
TX Death Valley	32.3	33.2	10.4	10.7

Table 6. Observed (Obs) and simulated (Sim) winter and summer mean and variance for the second period (1981–2009 for Berlin, 1986–2009 for Death Valley) according to each of the 2 approaches used to reconstruct temperature (Sim1 and Sim2: mean with 95% confidence interval in brackets)

Winter						Summer					
Mean			Variance			Mean			Variance		
Obs	Sim1	Sim2	Obs	Sim1	Sim2	Obs	Sim1	Sim2	Obs	Sim1	Sim2
TN Berlin											
−1.7	−2.0	−1.8	25.0	25.2	23.2	13.1	13.2	13.0	8.7	8.2	9.1
	(−2.7; −1.4)	(−2.5; −1.2)		(20.8; 29.8)	(18.9; 28.4)		(12.9; 13.4)	(12.8; 13.2)		(7.5; 8.9)	(8.2; 9.9)
TX Death Valley											
20.2	20.7	20.2	18.7	18.2	18.4	45.7	45.4	45.7	15.0	13.6	14.5
	(20.4; 21.1)	(19.9; 20.6)		(15.5; 20.8)	(15.9; 21.2)		(45.1; 45.8)	(45.3; 46.0)		(11.7; 16.1)	(12.6; 17.1)

and simulated mean and variance obtained for the second period in winter and in summer with each of the above 2 approaches (see Section 4.1) for each location and variable. As expected, approach 2, which takes trends and seasonalities of the second period into account, gives better results, but the results given by the first approach are also close to the observations. Fig. 6 gives a better view of the entire distribution: it presents, for different percentiles (from the very low 1% to the very high 99% through the median), the distribution of such percentiles obtained from the 100 simulations in black and the values obtained from the observations in grey. It shows that for all percentiles, the observed estimates fall inside the distributions of the simulated estimates, regardless of the approach taken for the simulations. This thus validates the 2 approaches to compute the distribution of temperature for a future period when mean and variance have changed.

4.2.2. Extremes

We now consider the extremes, in terms of 50 yr return levels and of heat or cold waves. Table 7 gives the obtained 50 yr return levels for period 2, again considering the series as stationary, and estimated from the observations and from each type of simulation. As in the previous section, the 95% confidence interval for the observations is computed with the delta-method, while for the simulations, the 2.5th and 97.5th percentiles of the distribution of the 100 estimated 50 yr RLs are taken. The results show that for Berlin, approach 2 gives slightly better results than approach 1, whereas for Death Valley, this is not the case. This can be explained by the fact that the smoothing parameter computed to estimate the mean and variance trends is the same for both periods for Death Valley (0.08), whereas for Berlin it changes from 0.32 in the first period to 0.08 in the second. Thus, in Berlin, interannual variability for daily minimum temperature is higher in the second period, and taking this into account logically improves the simulations. Fig. 7 shows the distributions of cold waves in Berlin and heat waves in Death Valley according to each simulation procedure in the same way as Fig. 5 in the previous section. Here, both approaches give similarly good results.

5. CONCLUSION AND PERSPECTIVES

In this study, we present and validate a stochastic seasonal functional heteroscedastic auto-regressive model for daily temperature for different climates in Eurasia and the US.

When fitted over a long temperature series (daily minimum or maximum) and used to simulate a large number of equivalent trajectories, the model is able to correctly reproduce both the bulk and the extremes of the observed distribution. In particular, it is able to produce higher or lower extremes than observed.

For 2 temperature time series for which a break in the evolution of both mean and variance was identified around the middle of the period, the model was constructed over the first part of the period and used to reproduce the second part. As the model simulates the residuals after accounting for trends and seasonalities in mean and variance, the reconstruction of the observed variable for any period consists of reintroducing this information on trends and seasonalities. Two approaches have been tested: firstly taking global mean and variance changes between both periods into account (such as in the so-called 'delta method') and secondly introducing the real trends and seasonalities computed over the second period. The second approach allows interannual changes to

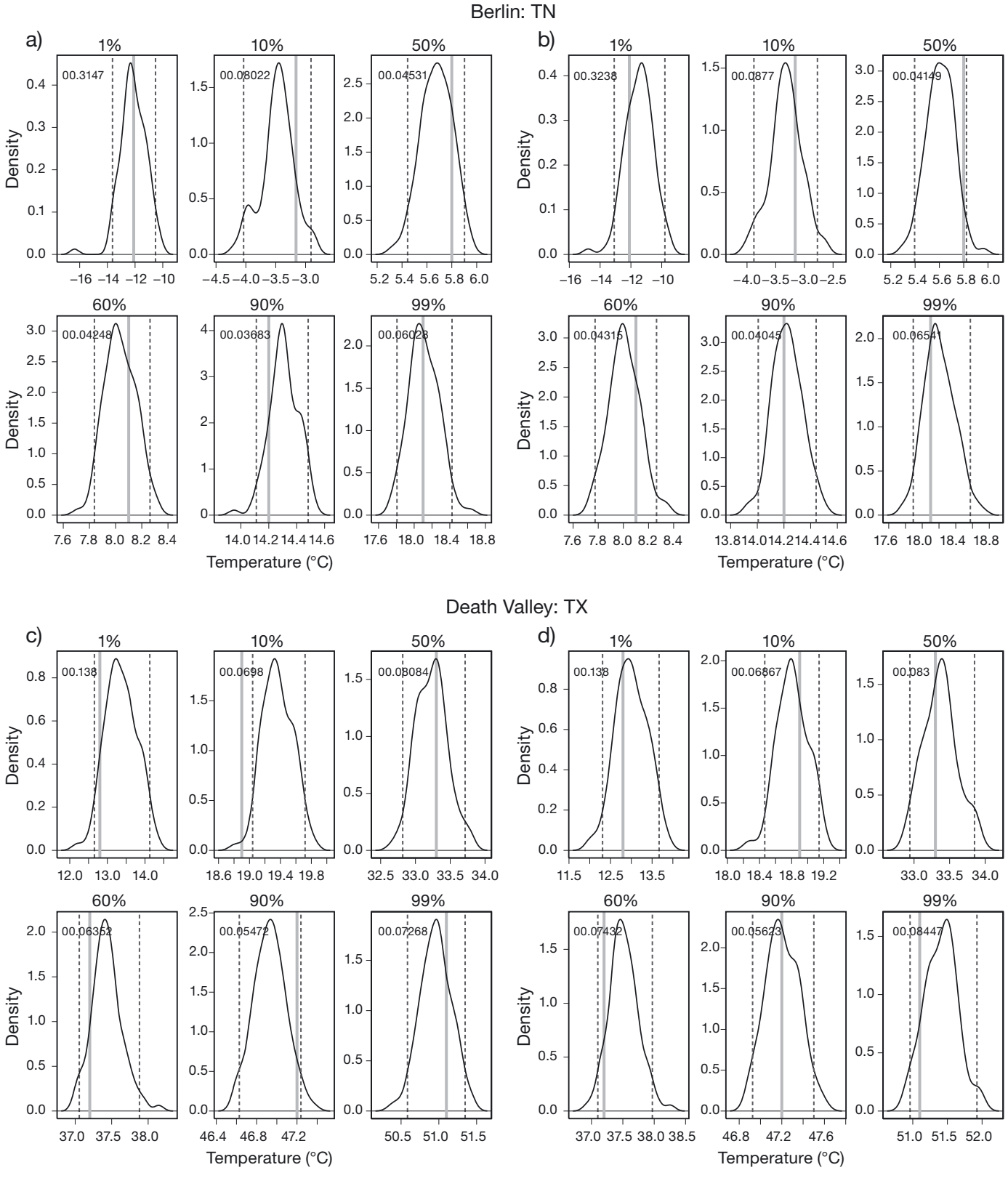

Fig. 6. Distributions of the 1st, 10th, 50th, 60th, 90th, and 99th percentiles of the 100 simulated temperature distributions (represented by solid black lines; dashed lines: 95% confidence interval [2.5th and 97.5th percentiles of the 100 simulations]) estimated for (a,b) daily minimum temperature in Berlin over the second period (1981–2009) and (c,d) daily maximum temperature in Death Valley over period 1986–2009, based on the first approach (a,c) and the second approach (b,d), together with the estimation of the same percentiles from the observations over the same period (grey line). Nos. within panels: bandwidths

Table 7. 50 yr return levels (RLs) of winter cold minimum daily temperature (TN) in Berlin and summer warm maximum daily temperature (TX) in Death Valley estimated from observed and simulated time series for the second period (1981–2009 for Berlin, 1986–2009 for Death Valley) and according to both approaches to reconstruct temperature (Simulations 1 and 2). For observations, the 95% confidence interval (in brackets) was obtained with the delta-method; for simulations, the given interval corresponds to the 2.5th and 97.5th percentiles of the distribution of the 100 obtained 50 yr RLs

Observations	Simulation 1	Simulation 2
Cold extremes Berlin		
−21.6 (−24.7; −18.5)	−28.2 (−37.3; −22.3)	−26.7 (−37.5; −21.5)
Warm extremes Death Valley		
53.2 (52.2; 54.1)	54.3 (53.2; 55.6)	55.1 (54.1; 56.2)

be taken into account if any occur. This is the case for the daily minimum temperature time series in Berlin, for which this last approach improves the results. Both approaches, however, give equivalently good results, both in terms of the bulk of the distribution and in terms of extremes.

Our results are encouraging from the perspective of using this tool as a downscaling technique suitable to deal with temperature extremes. The second approach in particular makes it possibile to take into account interannual variability changes. We can imagine, for example, that the model is fitted over an observed temperature time series representative of a location of interest, and then future temperatures for this location can be obtained by introducing the seasonalities and trends estimated over a corresponding, suitably corrected, grid point time series produced by different climate models with different scenarios.

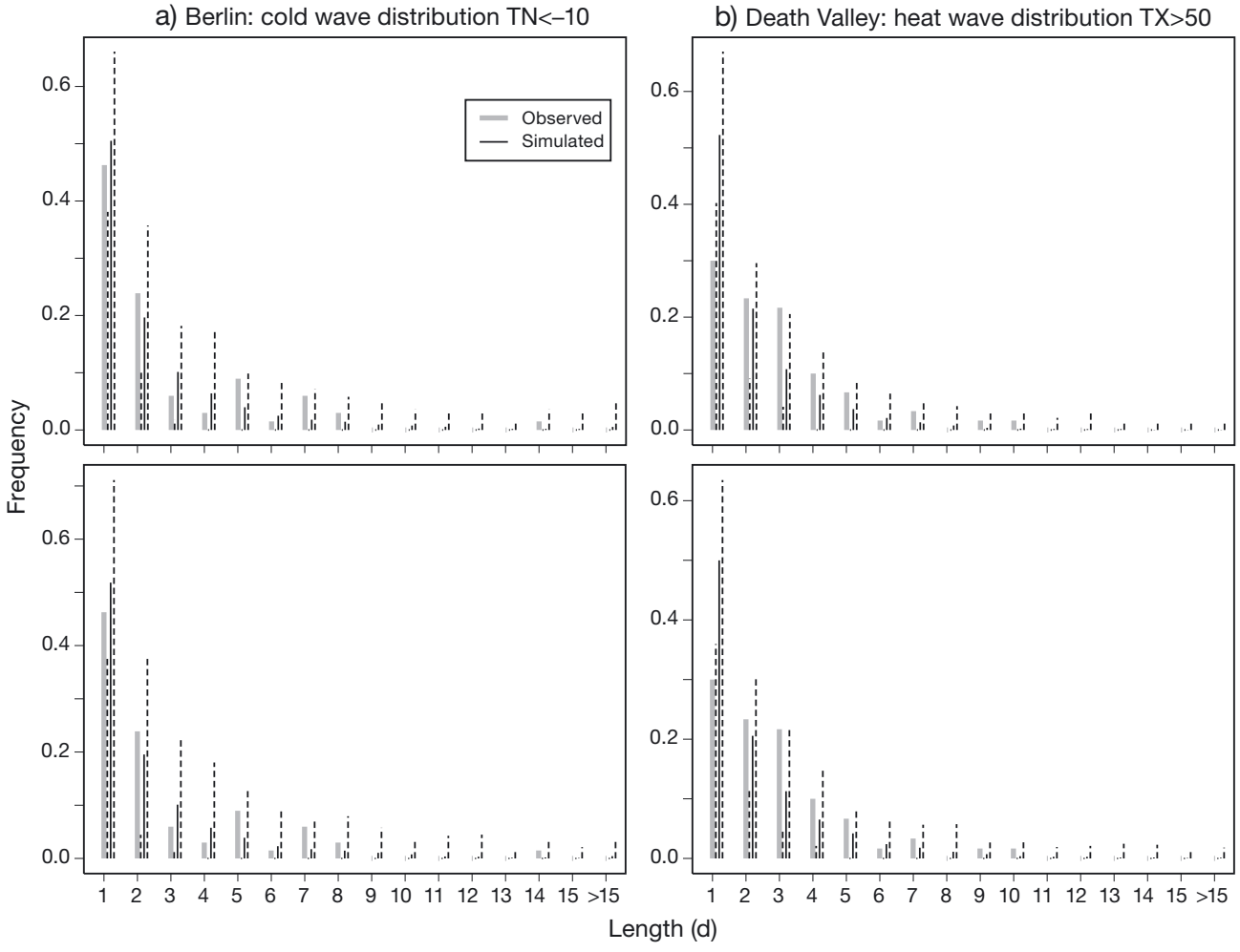

Fig. 7. Frequencies of the 1 to >15 d long (a) cold waves in Berlin and (b) heat waves in Death Valley. The definitions of cold and heat waves are the same as in Fig. 5. Solid black lines: mean frequencies obtained from the simulations; dashed lines: minimum and maximum frequencies; grey lines: observed frequencies. For each location, the top and bottom panels correspond to the first and second simulation approach, respectively

Our results show that this technique is able to give reliable information for the temperature extremes, for highest or lowest values as well as episodes. However, further studies should be devoted to hot and cold episodes. Although the model is able to produce long cold or heat waves, it should be able to produce more of such events among 100 simulations. Here the autocorrelation coefficient has been considered periodic, but it is suspected that it may increase once a certain high or low threshold is crossed. This will be further investigated. In a broader perspective, the model could be part of a more general weather generator in addition to a rainfall generator, for example.

Acknowledgements. We acknowledge the data providers from the ECA&D project (http://eca.knmi.nl) and the National Climatic Data Center, NOAA (www.ncdc.noaa.gov). We also thank the reviewers, whose comments helped to improve the paper.

LITERATURE CITED

Ailliot P, Thompson C, Thomson P (2009) Space–time modelling of precipitation using a hidden Markov model and censored Gaussian distributions. J R Stat Soc Ser C Appl Stat 58:405–426

Benth EF, Saltyte Benth J (2011) Weather derivatives and stochastic modelling of temperature. Int J Stoch Anal 2011:576791

Campbell SD, Diebold FX (2005) Weather forecasting for weather derivatives. J Am Stat Assoc 100:6–16

Chu CK, Marron JS (1991) Comparison of two bandwidth selectors with dependent errors. Ann Stat 19:1906–1918

Cleveland RB, Cleveland WS, McRae JE, Terpenning I (1990) STL: a seasonal-trend decompostion procedure based on loess (with discussion). J Off Stat 6:3–73

Coles S (2001) An introduction to statistical modeling of extreme values. Springer Series in Statistics. Springer Verlag, London

Cowpertwait P, Isham V, Onof C (2007) Point process models of rainfall: developments for fine-scale structure. Proc R Soc Lond A Math Phys Eng Sci 463:2569–2587

Dacunha-Castelle D, Hoang TTH, Parey S (2013) Modeling of air temperatures: preprocessing and trends, reduced stationary process, extremes, simulation. J Fr Stat Soc

Davis RA (1982) Maximum and minimum of one-dimensional diffusions. Stoch Proc Appl 13:1–9

Flecher C, Naveau P, Allard D, Brisson N (2010) A stochastic weather generator for skewed data. Water Resour Res 46:W07519 doi:10.1029/2009WR008098

Furrer EM, Katz RW (2008) Improving the simulation of extreme precipitation events by stochastic weather generators. Water Resour Res 44:W12439, doi:10.1029/2008WR007316

Hansen JW (2002) Realizing the potential benefits of climate prediction to agriculture: issues, approaches, challenges. Agric Syst 74:309–330

Hoang TTH (2010) Modélisation de séries chronologiques non stationnaires, non linéaires: application à la définition des tendances sur la moyenne, la variabilité et les extrêmes de la température de l'air en Europe. PhD thesis, University of Paris 11, Orsay (in English)

Klein Tank AMG, Wijngaard JB, Können GP, Böhm R and others (2002) Daily dataset of 20th-century surface air temperature and precipitation series for the European Climate Assessment. Int J Climatol 22:1441–1453

Kyselý J, Dubrovský M (2005) Simulation of extreme temperature events by a stochastic weather generator: effects of interdiurnal and interannual variability reproduction. Int J Climatol 25:251–269

Maraun D, Wetterhall F, Ireson AM, Chandler RE and others (2010) Precipitation downscaling under climate change: recent developments to bridge the gap between dynamical models and the end user. Rev Geophys 48:RG3003, doi:10.1029/2009RG000314

Marron JS (1987) Partitioned cross-validation. Econom Rev 6:271–284

Menne MJ, Durre I, Vose RS, Gleason BE, Houston TG (2012) An overview of the Global Historical Climatology Network — Daily Database. J Atmos Ocean Technol 29: 897–910

Mraoua M, Bari D (2007) Temperature stochastic modelling and weather derivatives pricing: empirical study with Moroccan data. Afr Stat 2:22–43

Mudelsee M (2009) Break function regression. Eur Phys J Spec Top 174:49–63

Muggeo VMR (2003) Estimating regression models with unknown break-points. Stat Med 22:3055–3071

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

Sansom J, Thompson P (2010) A hidden seasonal switching model for high-resolution breakpoint rainfall data. Water Resour Res 46:W08510, doi:10.1029/2009WR008602

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

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

Stone CJ (1977) Consistent nonparametric regression. Ann Stat 5:595–620

Sura P (2012) Stochastic models of climate extremes: theory and observations. In: A. AghaKouchak A, Easterling D, Hsu K, Schubert S, Sorooshian S (eds) Extremes in a changing climate: detection, analysis and uncertainty. Water Sci Tech Library 65:181–222

Wilks DS (1992) Adapting stochastic weather generation algorithms for climate change studies. Clim Change 22: 67–84

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

Appendix 1. Model description

Before choosing a model for the reduced process $Z(t)$, after removal of trends and seasonalities in mean and variance, its correlations and conditional variance were analyzed. The nonparametric analysis of the conditional variance of $Z(t)$ given $Z(t-1)$ shows a particular behavior: linear in the core of the distribution, and close to 0 for very high and low values of $Z(t-1)$, with the conditional mean being close to a linear function. The first step is thus to choose a FARCH model with finite bounds for the distribution. The application of the extreme theory is not justified at this step (because a mathematical theory does not exist for these processes) but once done, it gives a negative shape parameter ($\xi < 0$) that suggests a bounded distribution.

The idea is then to choose a modified FARCH model $Z(t) = b(Z(t-1)) + a(Z(t-1))\varepsilon_t$ where ε_t is a truncated Gaussian noise whose bounds depend on the value of $Z(t-1)$. The second step is then to represent the temperature as a continuous time process (with continuous trajectories). The FARCH processes are the first-order Euler scheme approximation of the discrete Markov chain M, where $M(t)$ is the observation at time t of the continuous diffusion given by: $dY(t) = b(t, Y(t)) + a(t, Y(t))dW(t)$ where b is the drift, a the diffusion coefficient, and $W(t)$ a Brownian motion. The estimation of the coefficients of such a continuous stationary diffusion is commonly done using its first-order Euler scheme Z, thus a FARCH process with the same functional coefficients. Technically this situation is very informative in relation to the extremes theory. From the geometric ergodicity of the diffusion, the extreme parameters and the bounds of the continuous time process can be estimated using only the chain M. Z is from now considered as an approximation of M. Now we use the continuous process as a tool. The coefficients of extremes and thus the bounds r_1 and r_2 are estimated by fitting a GEV distribution to the maxima of the reduced series modeled here as $M(t)$. The domain of $M(t)$, say (r_1, r_2), is bounded so that r_1 and r_2 are inaccessible boundary points for Y. At the boundary, we have:

(1) a and b are defined and continuous on $[r_1, r_2]$

(2) $b(r_1)b(r_2) \neq 0$

Under hypotheses 1 and 2 and $\xi < 0$, Dacunha-Castelle et al. (2013) proved the following theorem:

If the distribution of the maximum of the diffusion Y is in the domain of attraction of a GEV distribution with $\xi < 0$, then the marginal distribution is common to the chain M and to Y, and so they are in the same domain of max attraction.

We have the following behavior of a as $x \to r_2$:

$a^2(x) = -2b(r_2)\xi'(r_2 - x) + o(r_2 - x)$ where $\xi' = \dfrac{\xi}{\xi - 1}$

This information is then plugged in as constraints in the likelihood of the Euler scheme to estimate coefficients a and b with bound constraints.

Editorial responsibility: Richard Katz, Boulder, Colorado, USA

Submitted: July 8, 2013; Accepted: October 6, 2013
Proofs received from author(s): January 22, 2014