

# Stochastic simulation of daily air temperature and precipitation from monthly normals in North America north of Mexico

Jacques Régnière · Rémi St-Amant

Received: 24 May 2006 / Revised: 1 November 2006 / Accepted: 5 December 2006 / Published online: 16 January 2007  
© ISB 2007

**Abstract** A simple, stochastic daily temperature and precipitation generator (TEMPGEN) was developed to generate inputs for the study of the effects of climate change on models driven by daily weather information when climate data are available as monthly summaries. The model uses as input only 11 sets of monthly normal statistics from individual weather stations. It needs no calibration, and was parameterized and validated for use in Canada and the continental United States. Monthly normals needed are: mean and standard deviation of daily minimum and maximum temperature, first and second order autoregressive terms for daily deviations of minimum and maximum temperatures from their daily means, correlation of deviations of daily minimum and maximum temperatures, total precipitation, and the interannual variance of total precipitation. The statistical properties and distributions of daily temperature and precipitation data produced by this generator compared quite favorably with observations from 708 stations throughout North America (north of Mexico). The algorithm generates realistic seasonal patterns, variability and extremes of temperature, precipitation, frost-free periods and hot spells. However, it predicts less accurately the daily probability of precipitation, extreme precipitation events and the duration of extreme droughts.

**Keywords** Stochastic model · Climate change · Autoregressive model · Daily weather generator · Low-frequency variability

## Introduction

Progress in forecasting the impacts of climate change on forest ecosystems (stand growth, succession sequences) and disturbance regimes (insect and disease outbreaks, fire frequencies) involves the ability to simulate extreme weather events. Extreme events are often sequences ranging from a few to several successive days of unusually hot, cold, dry or wet weather. Many researchers are developing climate-sensitive models of ecological processes that require as input daily or even hourly time series of temperature and precipitation (Prentice et al. 1993; Kimball et al. 1997; Friend 1998; Lasch et al. 2002; Régnière and Nealis 2002; Logan et al. 2003; Gray 2004). Typically, however, climate change scenarios are generated from single runs of Global Circulation Models (CGMs) with daily output at coarse spatial resolution. Because this output represents a single realization from a large set of possible outcomes, the possibility of statistical treatment (replication) is severely limited (Oelschlägel 1995). However, the output of CGMs can be summarized as monthly statistics (normals) or as “anomalies” (deviations from base-line normals) and disaggregated through stochastic daily weather generators (Bürger 1997; Hansen and Ines 2005). The low spatial resolution of GCMs can be increased through regional models or other down-scaling approaches, and monthly statistics can be input into ecological models to take into account climate change scenarios (Fowler et al. 2005; Qian et al. 2005).

The major issue in the development of stochastic daily weather generators is to obtain a sequence of daily weather that conforms not only to observed (or expected) mean values, but also to natural variability, extremes, and time series characteristics (Wallis and Griffiths 1995). Several approaches to generation of daily weather information have

J. Régnière (✉) · R. St-Amant  
Natural Resources Canada, Canadian Forest Service,  
Laurentian Forestry Centre,  
1055 PEPS Street, P.O. Box 10380, Stn. Sainte-Foy,  
Quebec, QC G1V 4C7, Canada  
e-mail: jregniere@cfl.forestry.ca

been developed, and most are based on the simulation of alternating dry- and wet-day series using Markov-chain transition probabilities (Richardson 1981; Racsco et al. 1991; Hutchinson 1995a; Katz 1996; Wilks 1999). Many of these models, however, require considerable amounts of input information and often must be calibrated for application in specific geographical areas (Richardson and Wright 1984; Racsco et al. 1991; Semenov and Barrow 1997; Bürger 1997; Semenov et al. 1998). There has also been considerable attention focused on adding interannual variation in these generators for ecological applications (Mearns et al. 1997; Hansen and Mavromatis 2001; Mavromatis and Hansen 2001; Dubrovský et al. 2004).

Régnière and Bolstad (1994) developed an algorithm (TEMPGEN) for simulation of daily minimum and maximum air temperature based on normals (30-year monthly average and extreme minimum and maximum temperatures). This algorithm generates realistic distributions, autocorrelation and cross-correlation patterns of daily minimum, maximum temperatures and ranges. It has been quite useful in studying the responses of models driven by daily weather information under “normal” weather conditions and has been used in many applications through the BioSIM system (Régnière 1996). However, this generator has three weaknesses: (1) it does not take into account the interannual variation of monthly average temperatures, and therefore generates temperature traces that are not as variable as they are in nature; (2) it does not generate daily precipitation; and (3) it was parameterized from data from the Atlantic coastal area of the United States, and as such is not applicable in other areas.

The objective of the work reported here was to remedy these three weaknesses by generalizing and validating the Régnière and Bolstad (1994) algorithm for the entire North American continent (north of Mexico) and adding daily precipitation to its output. We sought a model that would require minimal input (in the form of local monthly normal statistics) would need no for calibration, and would generate realistic daily times series of minimum and maximum air temperature and precipitation, exhibiting levels of variability and statistical distributions close to those of observed time series.

**Methods**

**Algorithm structure**

Throughout this document, temperatures are expressed in °C and precipitation in millimeters of water. The symbol  $\sim$  above a term designates a 30-year normal while  $\bar{\phantom{x}}$  represents an average of some other type. Inputs of this model are 11 sets of monthly statistics extracted over a

standard normal period (SNP) of 30 years of daily observations from a given location. The first two sets of normals are mean daily minimum ( $\tilde{\mu}_n$ ) and maximum ( $\tilde{\mu}_x$ ) temperatures. From these, normal daily minimum and maximum temperatures are obtained by linear interpolation between the means of successive months, adjusted to correspond to the values occurring at mid-month. Seven additional monthly normals are calculated from the differences between observed daily minimum and maximum temperatures and these daily normals: their standard deviations ( $\tilde{\sigma}_\delta$  for minimum and  $\tilde{\sigma}_\varepsilon$  for maximum temperature); 2nd order autoregressive terms describing their time-series characteristics ( $\tilde{A}_\delta, \tilde{B}_\delta$  for minimum and  $\tilde{A}_\varepsilon, \tilde{B}_\varepsilon$  for maximum temperatures); and their cross-correlation ( $\tilde{r}_{\delta\varepsilon}$ ). Two additional monthly statistics describe the characteristics of precipitation: average monthly total ( $\tilde{P}$ ); and the standard deviation ( $\tilde{\sigma}_P$ ) of the ratio of observed monthly total to normal precipitation, given by:

$$\tilde{\sigma}_P = \sqrt{\sum_i \left(\frac{P_m}{\tilde{P}} - 1\right)^2 / (n - 1)} \tag{1}$$

where  $P_m$  is a month’s total precipitation in year  $m$  and  $n=30$  years.

Daily minimum and maximum temperatures are generated with the algorithm of Régnière and Bolstad (1994):

$$\begin{aligned} T_n(t) &= \tilde{T}_n(t) + \delta(t) \\ T_x(t) &= \tilde{T}_x(t) + \varepsilon(t) \end{aligned} \tag{2}$$

where  $\tilde{T}_n(t)$  and  $\tilde{T}_x(t)$  are daily normal minimum and maximum temperatures (obtained by linear interpolation as described above), and  $\delta(t)$  and  $\varepsilon(t)$  are auto-correlated (2nd order) and cross-correlated random numbers with means of 0 and variances  $\tilde{\sigma}_\delta^2$  and  $\tilde{\sigma}_\varepsilon^2$ . The autocorrelation and cross-correlation are generated through general linear stochastic processes expressed as:

$$\begin{aligned} \delta(t) &= \sum_i \alpha_i \gamma(t - i) \\ \varepsilon(t) &= \frac{\tilde{\sigma}_\varepsilon}{\tilde{\sigma}_\delta} \sum_i \beta_i [\rho \gamma(t - i) + (1 - |\rho|)\zeta(t - i)] \end{aligned} \tag{3}$$

where  $\alpha_i$  and  $\beta_i$  are infinite series of coefficients such that:

$$\begin{aligned} \alpha_1 &= 1, \alpha_2 = \tilde{A}_\delta, \text{ and } \alpha_i = \tilde{A}_\delta \alpha_{i-1} + \tilde{B}_\delta \alpha_{i-2} \text{ for } i > 2 \\ \beta_1 &= 1, \beta_2 = \tilde{A}_\varepsilon, \text{ and } \beta_i = \tilde{A}_\varepsilon \beta_{i-1} + \tilde{B}_\varepsilon \beta_{i-2} \text{ for } i > 2 \end{aligned} \tag{4}$$

and  $\gamma \in \text{Nor}(0, \sigma_\gamma^2), \zeta \in \text{Nor}(0, \sigma_\zeta^2)$ . The formulation of Eq. 3 is a generalization of Régnière and Bolstad’s (1994)

model 2, where  $\rho$  can be in the range  $[-1,1]$  and  $\xi$  can have any variance regardless of the values of  $\tilde{A}_\delta, \tilde{B}_\delta, \tilde{A}_\varepsilon, \tilde{B}_\varepsilon$  and  $\rho$ . The variances  $\tilde{\sigma}_\gamma^2$  and  $\tilde{\sigma}_\zeta^2$  are given by:

$$\sigma_\gamma^2 = \frac{\tilde{\sigma}_\delta^2}{\sum_i \alpha_i^2} \tag{5}$$

and

$$\sigma_\zeta^2 = \frac{\tilde{\sigma}_\delta^2}{(1 - |\rho|)^2} \left( \frac{1}{\sum_i \beta_i^2} - \frac{\rho^2}{\sum_i \alpha_i^2} \right) \tag{6}$$

and the value of  $\rho$  is:

$$\rho = \tilde{r}_{\delta\varepsilon} \frac{\sum_i \alpha_i^2}{\sum_i (\alpha_i \beta_i)} \tag{7}$$

(see Régnière and Bolstad 1994, Appendix 2, for the derivation of Eqs. 5–7). The sums  $\sum \alpha^2$ ,  $\sum \beta^2$  and  $\sum (\alpha\beta)$  converge within the first 20 terms. In the original model of Régnière and Bolstad (1994),  $\tilde{A}_\delta, \tilde{B}_\delta, \tilde{A}_\varepsilon, \tilde{B}_\varepsilon$  and  $r_{\delta\varepsilon}$  were calculated from station latitude and time of year. Observed extreme minimum and maximum observations were also needed to calculate  $\tilde{\sigma}_\delta$  and  $\tilde{\sigma}_\varepsilon$ , based on the analysis of data from the Atlantic area of the United States. In the present model, they are input station normals.

Contrary to most daily weather generators where daily precipitation is the basis for simulation of temperature values (Katz 1996), in this model air temperature is generated first, and precipitation is computed on the basis of daily temperatures, with the underlying assumption that precipitation is more likely on days with smaller temperature ranges. Total monthly precipitation ( $P$ ) is generated by drawing a multiplicative random deviation  $\Delta P \geq 0$  from a family of Weibull probability distribution functions entirely parameterized from  $\tilde{\sigma}_P$ :

$$P = \Delta P \cdot \tilde{P} \tag{8}$$

The probability of daily precipitation  $p(t)$  is generated from total monthly precipitation  $P$ , its variance  $\sigma_P^2$  and the daily temperature range  $R(t) = T_x(t) - T_n(t)$ . The value of  $p(t)$  is then compared with a uniformly distributed random number  $\psi \in \mathcal{U}(0, 1)$ . Whenever  $p(t) < \psi$ ,  $p(t)$  is reset to 0. For days where  $p(t) > 0$ , the amount of daily precipitation  $P(t)$  is calculated as follows. The proportion  $x(t)$  of total monthly precipitation falling on day  $t$  is obtained from a Beta distribution specified entirely from the month's  $\tilde{\sigma}_P$  and that day's precipitation is given by:

$$P(t) = x(t)P \tag{9}$$

where  $\sum_t x(t) = 1$  within a month.

### Data sources

Monthly normals ( $\tilde{\mu}_n, \tilde{\mu}_x, \tilde{\sigma}_\delta, \tilde{\sigma}_\varepsilon, \tilde{A}_\delta, \tilde{B}_\delta, \tilde{A}_\varepsilon, \tilde{B}_\varepsilon, \tilde{r}_{\delta\varepsilon}, \tilde{P}$  and  $\tilde{\sigma}_P$ ) for stations in North America were calculated over the 1961–1990 SNP from daily records contained in two databases. For Canada, data from Atmospheric Environment Service Daily Climate Data CD-ROMs (<http://www.msc-smc.ec.gc.ca/>) were used. For the United States (excluding Alaska, Atlantic and Pacific Island states), the NCDC COOP Summary of the Day TD3200 database (<http://www.ncdc.noaa.gov/>) was used. Only stations with at least 20 years of data for each month during the SNP were retained. For these stations, a month's data were retained only if it contained at least 20 days of non-missing temperature or precipitation data. Total precipitation in months with missing data was adjusted ( $P = P_{unadjusted} \times \text{days in month} / \text{non-missing days in months}$ ).

### Variation of monthly mean temperature and total precipitation

Annual variation of monthly mean temperature and precipitation was analyzed using data from four clusters of weather stations representing a wide range of North American climates: 156 in Quebec, 129 in British Columbia, 118 in California and 108 in Tennessee and North Carolina. Deviations of monthly mean temperature ( $\mu$ ) and total precipitation ( $P$ ) from 30-year normals [ $\tilde{\mu} = (\tilde{\mu}_n + \tilde{\mu}_x) / 2$ ] and  $\tilde{P}$  were calculated:

$$\begin{aligned} \Delta\mu &= \mu - \tilde{\mu} \\ \Delta P &= P / \tilde{P} \end{aligned} \tag{10}$$

To detect month-to-month autocorrelation, values of  $\Delta\mu$  and  $\Delta P$  were averaged by year and month over all stations in each cluster. This provided eight 360-month time series: four of  $\Delta\mu$  and four of  $\Delta P$ . Autocorrelation in each of these time series was calculated for lags of 1–12 months. Cross-correlations (without lag) between deviations in temperature and precipitation were also tested for significance.

The distribution of monthly precipitation deviations  $\Delta P$  changed consistently from month to month, and the resulting patterns were different between station clusters. The Weibull probability distribution function (see Hastings and Peacock 1975) was chosen as a suitably flexible model to describe these patterns:

$$p(\Delta P \leq X) = 1 - e^{-[(X+0.05)/\lambda]^p} \tag{11}$$

where  $X \geq 0$ . The value 0.05 was added to  $X$  to ensure that months with no precipitation could have a non-zero probability of occurring ( $p > 0$  when  $X = 0$ ).

Equation 11 was fitted to the cumulative distribution of  $\Delta P$  for each month, pooled separately over stations from

each cluster. Parameters  $\eta$  and  $\lambda$  were estimated for the 48 distributions (12 months, for stations in Quebec, British Columbia, California and Tennessee/North Carolina) by non-linear least squares regression (SAS release 8.04, PROC NLIN). Parameter  $\eta$  was highly correlated to the monthly values of  $\bar{\sigma}_P = \frac{1}{n} \sum \tilde{\sigma}_P$  averaged over all stations (by cluster and month). The relationship was expressed as:

$$\hat{\eta} = a_\eta + b_\eta / \bar{\sigma}_P \quad (12)$$

where  $\hat{\eta}$  is an estimate of  $\eta$  given  $\bar{\sigma}_P$ . Parameters  $a_\eta$  and  $b_\eta$  were estimated by linear regression. There was also a strong relationship between the values of  $\eta$  and  $\lambda$ , which we modeled with:

$$\hat{\lambda} = a_\lambda \left(1 - e^{-b_\lambda \hat{\eta}}\right)^{c_\lambda} \quad (13)$$

where  $\hat{\lambda}$  is an estimate of  $\lambda$ ;  $a_\lambda$ ,  $b_\lambda$  and  $c_\lambda$  are parameters that were estimated by non-linear least squares regression. The probability distribution of total monthly precipitation for a given station can be determined from  $\bar{P}$  and  $\tilde{\sigma}_P$  with Eqs. 11, 12 and 13.

### Generation of daily precipitation

For this analysis, daily air temperature and precipitation records were compiled from the four clusters of weather stations described above. Daily probabilities of precipitation were related through a purely empirical model to three variables: total monthly precipitation  $P$ , its interannual variability  $\tilde{\sigma}_P$ , and daily temperature range  $R(t)$ . The distinction between wet and dry days is made on the basis of these probabilities. Total monthly precipitation is partitioned among wet days through a family of Beta distributions.

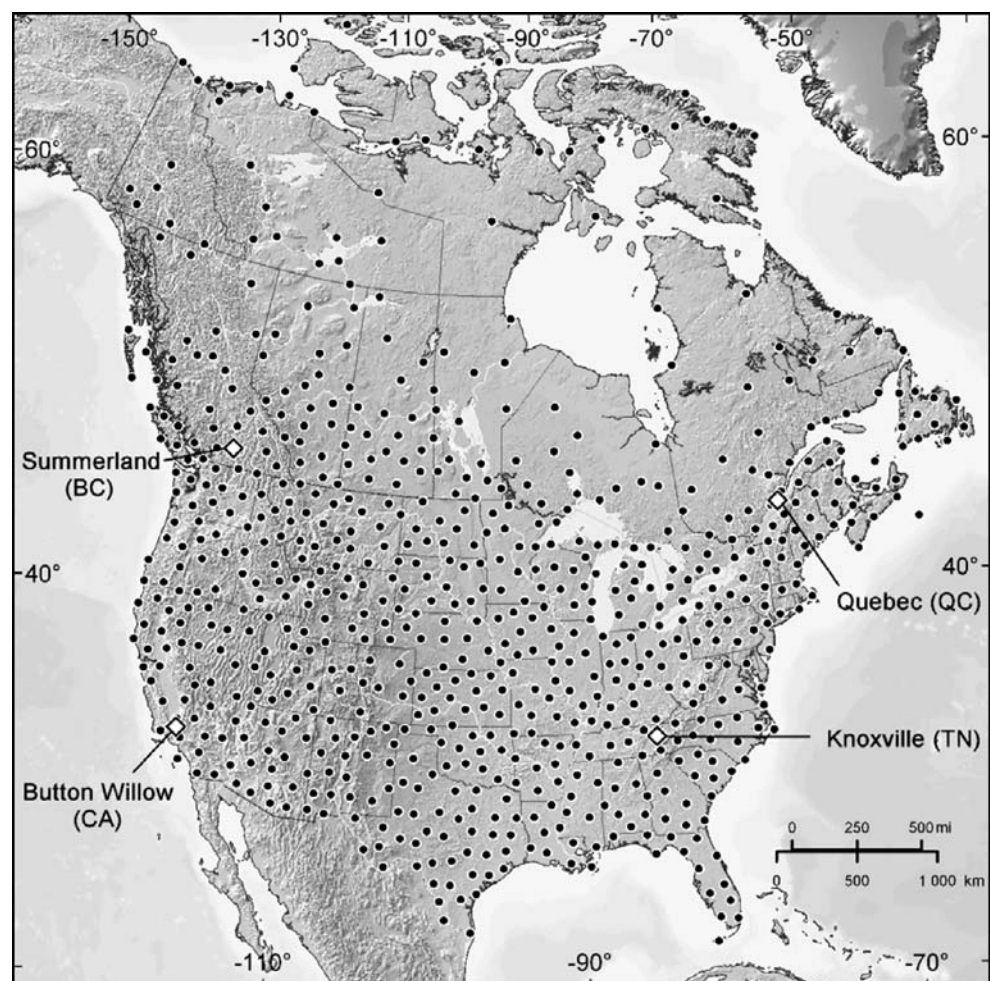
The probability of daily precipitation  $p(t)$  on day  $t$  was modeled with:

$$g(p) = k_0 + k_1 \log_{10}(P + 1) + k_2 \log_{10} W + k_3 \tilde{\sigma}_P \quad (14)$$

where

$$g(p) = \ln [p/(1 - p)] \quad (15)$$

**Fig. 1** Map of Canada and the United States.  $\diamond$  Data sources for detailed model output analysis,  $\bullet$  data sources for simulator validation



is the logistic transform of  $p$ , and  $W$  is a Weibull probability density function of daily temperature range  $R(t)$ :

$$W = \frac{\omega_2}{\omega_1} \left( \frac{R(t) + 1}{\omega_1} \right)^{\omega_2 - 1} e^{-\left( \frac{R(t) + 1}{\omega_1} \right)^{\omega_2}} \quad (16)$$

with four parameters:

$$\omega_1 = k_4 + k_5 \tilde{\sigma}_P \text{ and } \omega_2 = k_6 + k_7 \tilde{\sigma}_P. \quad (17)$$

Equations 14–17 were fitted to observations in the following manner. Days with precipitation exceeding 0.1 mm were assigned a precipitation index of 1 (otherwise 0). Precipitation indices were averaged, providing values of  $p(t)$  for classes of  $P$ ,  $R(t)$  and  $\tilde{\sigma}_P$ .  $P$  values were grouped into 5 mm/month classes from 0 to  $\geq 250$  mm/month.  $R(t)$  values were grouped into 1°C classes from  $\leq 3$  to  $\geq 25^\circ\text{C}$ .  $\tilde{\sigma}_P$  values were grouped into 14 classes: 0–0.3, 0.3–0.4, ..., 0.9–1.0, 1.0–1.5, ..., 3.0–3.5,  $>3.5$ . Classes containing fewer than 10 observations were dropped, leaving a total

of 9,821 distinct combinations of the independent variables. Parameters  $k_0$ – $k_7$  were estimated simultaneously by non-linear logistic regression analysis.

The distribution of the proportion of total monthly precipitation falling on a given rainy day  $t$ , the ratio  $P(t)/P$  where  $P(t) > 0.1$  mm, was approximated by a family of Beta distributions  $\text{Beta}(v, w)$  where shape parameters  $v$  and  $w$  are estimated from the mean and uncorrected variance of the  $P(t)/P$  ratio (see Hastings and Peacock 1975). The data set was subdivided into 14 groups according to the values of  $\tilde{\sigma}_P$  (classes listed above),  $v$  and  $w$  were calculated by month within each class and estimates of  $w$  were related to  $\tilde{\sigma}_P$  (the average  $\tilde{\sigma}_P$  within each group) with:

$$w = m_0 + \frac{m_1}{(\tilde{\sigma}_P + 1)^{m_2}} \quad (18)$$

Parameters  $m_0$ – $m_2$  were estimated by non-linear least-squares regression. There was also a curvi-linear relationship between  $v$  and  $w$  expressed with:

$$v = m_3 + m_4 w + m_5 / w \quad (19)$$

where  $m_3$ – $m_5$  are parameters estimated by linear regression. From relationships in Eqs. 18 and 19, a Beta distribution of the  $P(t)/P$  ratio can be assigned for any value of  $\tilde{\sigma}_P$ .

Model validation

Data from one station in each of the four clusters used in model parameterization were used for a detailed comparison of observed and simulated patterns of monthly variation

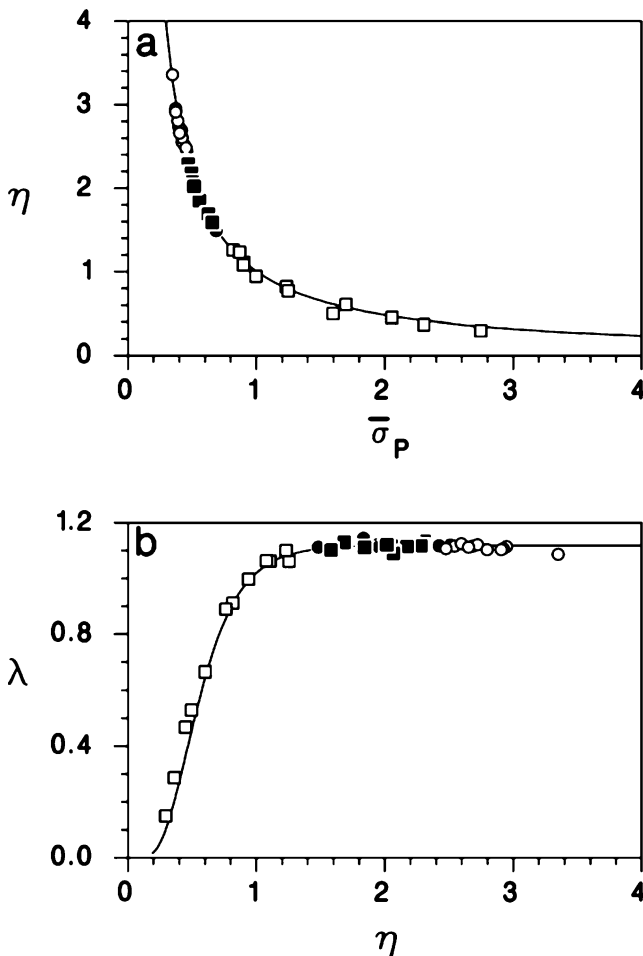


Fig. 2a, b Variation of parameters  $\eta$  and  $\lambda$  of Eq. 11 describing the distribution of  $\Delta P = P_i / \bar{P}$ , relative monthly precipitation for stations in (○) Quebec, (●) British Columbia, (□) California and (•) Tennessee/North Carolina. a Relationship between  $\eta$  and the standard deviation of monthly precipitation  $\tilde{\sigma}_P$  (line: Eq. 12). b Relationship between  $\lambda$  and  $\eta$  (line: Eq. 13)

Table 1 List of parameter estimates for Eqs. 12–19 of the simulator

Equation	Parameter	Estimate±SE
Eq. 12 ( $R^2=0.995$ )	$a_\eta$	0.075±0.009
	$b_\eta$	0.928±0.020
	$c_\lambda$	7.509±0.960
Eq. 13 ( $R^2=0.985$ )	$a_\lambda$	1.117±0.005
	$b_\lambda$	4.411±0.226
Eqs. 14–17 ( $R^2=0.937$ )	$k_0$	1.544±0.072
	$k_1$	1.023±0.013
	$k_2$	1.459±0.075
	$k_3$	-1.140±0.025
	$k_4$	-0.925±0.221
	$k_5$	4.875±0.085
	$k_6$	-0.166±0.011
	$k_7$	1.617±0.019
Eq. 18 ( $R^2=0.997$ )	$m_0$	0.2 <sup>a</sup>
	$m_1$	29.6±1.2
	$m_2$	4.025±0.118
Eq. 19 ( $R^2=0.904$ )	$m_3$	0.284±0.027
	$m_4$	0.039±0.004
	$m_5$	0.035±0.011

<sup>a</sup> Parameter value fixed

and distributions. These stations were selected because they had the least missing data: Summerland BC (49°34'N, 119°39'W), Quebec QC (45°48'N, 71°23'W), Button Willow CA (35°24'N, 119°28'W) and Knoxville TN (35°49'N, 83°59'W) (Fig. 1). For each station, ten observed and simulated monthly statistics were compared: extreme and mean minimum and maximum temperature, the standard deviations of mean temperature and total precipitation, average daily precipitation and probability of precipitation, and the duration of droughts and hot spells. Droughts were defined as periods of successive days with precipitation  $< 0.1$  mm. Hot spells were defined as periods of successive days with maximum temperature  $\geq 30^\circ\text{C}$ . Seven observed and simulated frequency distributions were compared: daily minimum and maximum temperatures, daily temperature range, frost-free period, duration of droughts and hot spells, and amount of daily precipitation.

Normals and daily data over the entire 1961–1990 SNP were obtained for a set of 708 North American weather stations. This set was selected at random by a thinning algorithm eliminating neighboring stations within 100 km of each other, so that selected stations were distributed as evenly as possible across Canada and the continental United States, conditional on station density. Station density was very low in northern Canada (Fig. 1). From the monthly normals of each station ( $\tilde{\mu}_n, \tilde{\mu}_x, \tilde{\sigma}_\delta, \tilde{\sigma}_\varepsilon, \tilde{A}_\delta, \tilde{B}_\delta, \tilde{A}_\varepsilon, \tilde{B}_\varepsilon, \tilde{r}_{\delta\varepsilon}, \tilde{P}$  and  $\tilde{\sigma}_P$ ), daily temperature and precipitation

data (300 years) were generated with the daily weather generator and simulator-generated monthly normals for each station were compiled. Six additional observed and simulated monthly statistics were calculated by station: mean extreme minimum and maximum temperatures, average daily temperature range, average daily precipitation and probability of precipitation, and the standard deviation of mean temperature, designated  $\sigma_\mu$ . Observed and simulated annual frost-free periods were compiled. Five extreme events were extracted from observed and simulated frequency distributions compiled by station over the entire observation (30 years) or simulation (300 years) period: the coldest and hottest temperatures (5th percentile lowest minimum and 95th percentile highest maximum temperatures), the heaviest daily precipitation events (95th percentile), the longest droughts (95th percentile) and the longest hot spells (95th percentile).

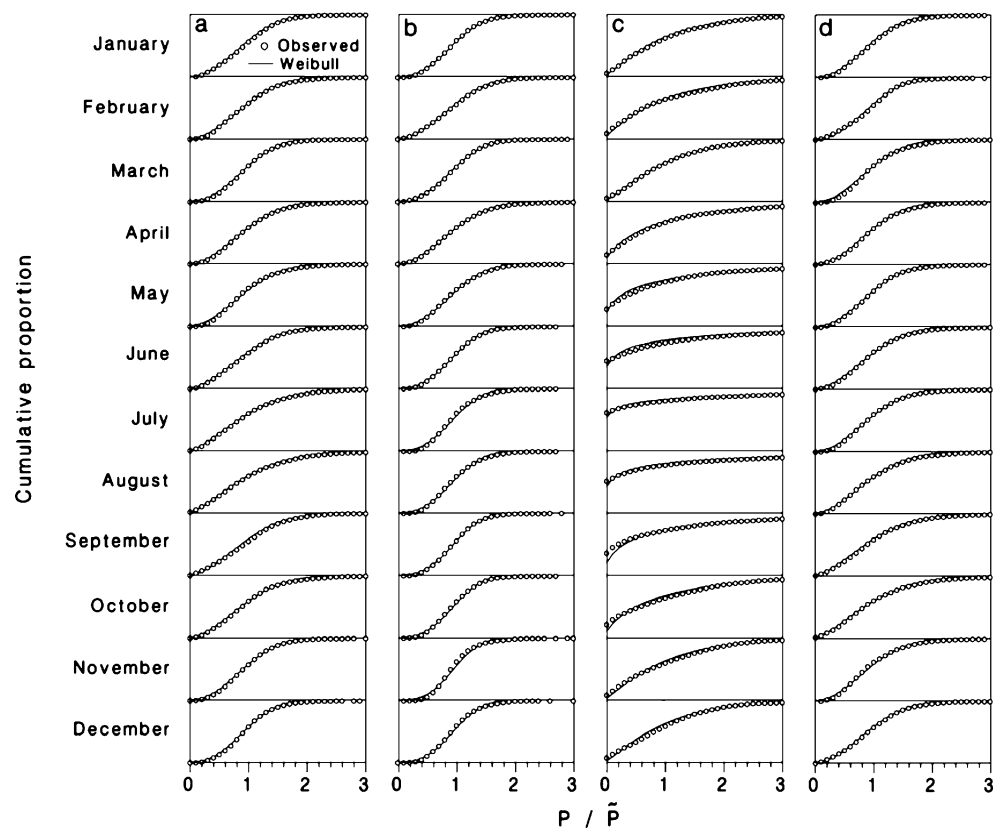
## Results

### Data analysis

#### Autocorrelation

There was a weak but significant 1-month lag autocorrelation in the monthly mean deviations from normals

**Fig. 3** Cumulative distribution of  $\Delta P = P_i / \tilde{P}$ , the ratio of monthly precipitation on normal precipitation, by month, for stations in **a** Quebec, **b** British Columbia, **c** California and **d** Tennessee/North Carolina.  $\circ$  Observations, — Weibull distribution function (Eq. 11)



temperature ( $\Delta\mu$ ) among stations from British Columbia ( $r=0.229, P<0.001$ ), California ( $r=0.159, P=0.003$ ), Tennessee/North Carolina ( $r=0.131, P=0.014$ ), but not from Quebec ( $r=0.067, P=0.201$ ). There was no significant autocorrelation in deviations of total monthly precipitation from normals ( $\Delta P$ ) among stations of any cluster normals at any time lag. Thus, the simulator was designed without a built-in autocorrelation structure in monthly deviations from monthly mean temperature or total precipitation. There was a weak correlation between deviations from monthly mean temperatures and deviations from monthly total precipitation among time series from British Columbia ( $r=-0.149, P<0.001$ ), Quebec ( $r=0.045, P<0.001$ ) and California ( $r=-0.130, P<0.001$ ), but none from Tennessee/North Carolina ( $r=-0.004, P=0.454$ ). Because of the weakness and ambiguity of this association, no attempt

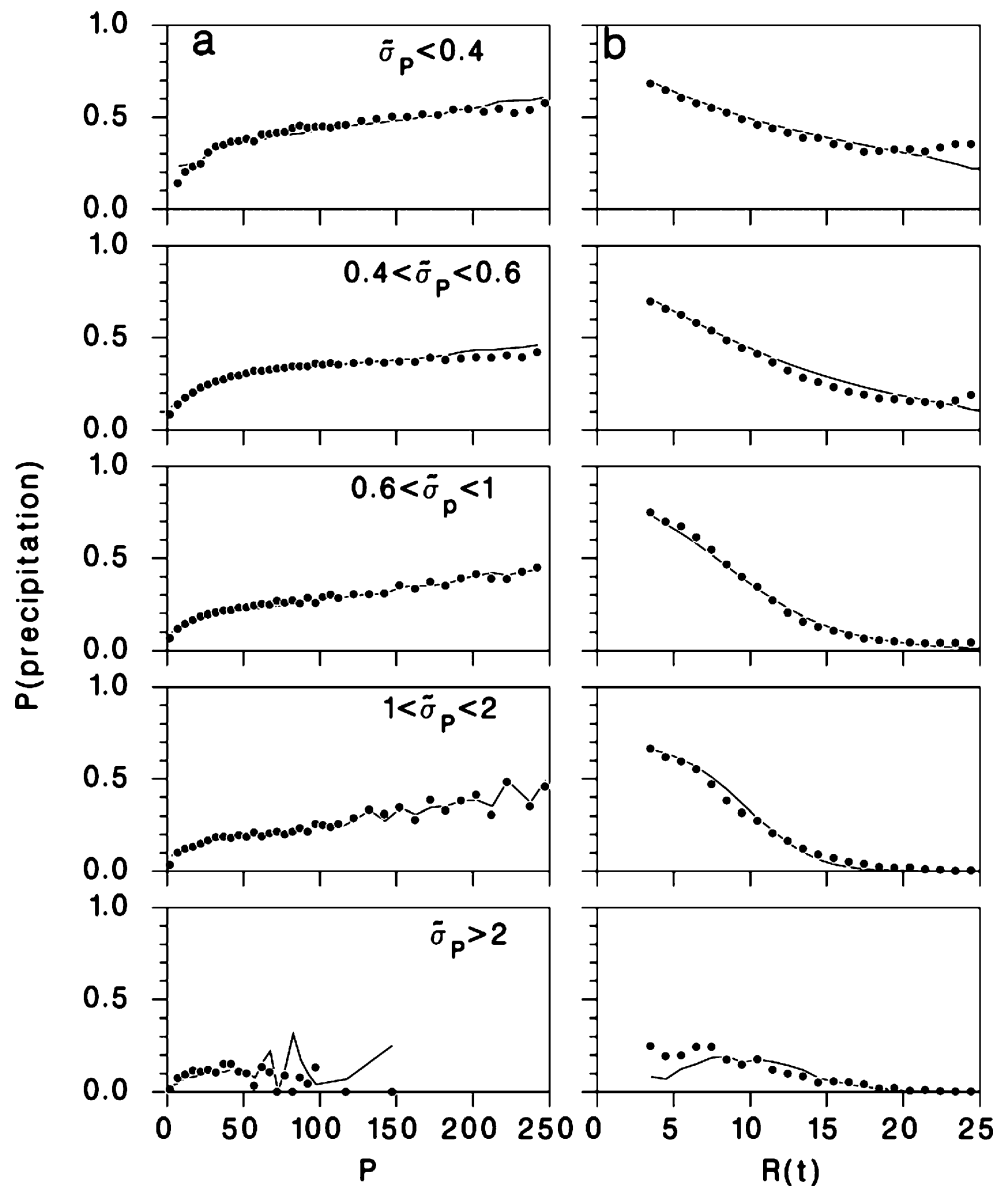
was made to correlate average monthly temperature and total precipitation in the simulator.

Precipitation

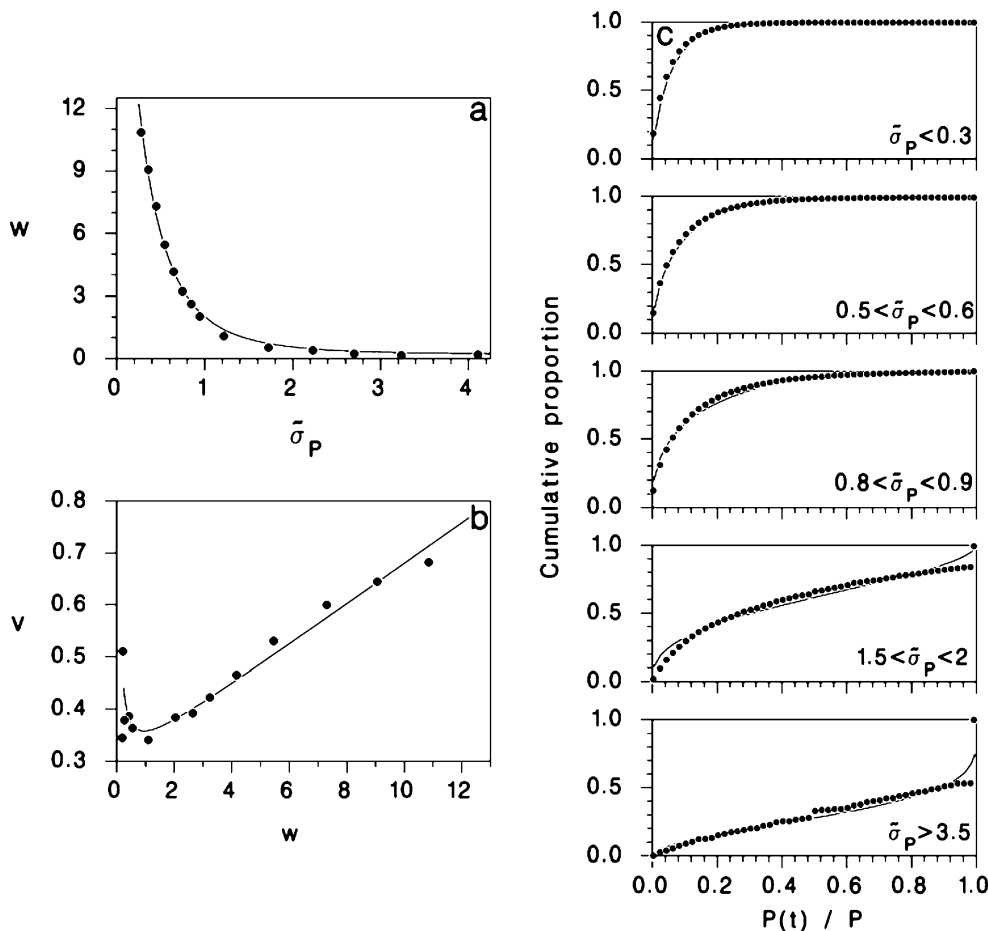
Equations 12 and 13 described precisely the relationships between  $\eta$  and  $\bar{\sigma}_p$  (Fig. 2a), and  $\lambda$  and  $\eta$  (Table 1; Fig. 2b). The family of Weibull distribution functions (Eq. 11) generated by these relationships approximated quite well the patterns of monthly deviations of total precipitation from normals,  $\Delta P$ , among data from stations from the four clusters: British Columbia (Fig. 3a), Quebec (Fig. 3b), California (Fig. 3c) and Tennessee/North Carolina (Fig. 3d).

There was a clear diminishing-return relationship between the daily probability of precipitation  $p(t)$  and total

**Fig. 4** Relationships between the daily probability of precipitation  $p(t)$  and **a** total monthly precipitation  $P$  or **b** daily temperature range  $R(t)$  for different values of precipitation variability  $\bar{\sigma}_p$ . • Averaged observations, — Eqs. 14–17



**Fig. 5a–c** Variation of the parameters  $\nu$  and  $w$  of the family of Beta functions describing the distribution of daily precipitation. **a** Eq. 18 relating  $w$  to  $\tilde{\sigma}_P$ . **b** Eq. 19 relating  $\nu$  to  $w$ . **c** Cumulative distribution of the amount of daily precipitation  $P(t)$  relative to total monthly precipitation  $P$ , for five classes of precipitation variability  $\tilde{\sigma}_P$ . — expected values, • observations



monthly precipitation  $P$  (Fig. 4a). There was also a strong relationship between  $p(t)$  and daily temperature range  $R(t)$  (Fig. 4b), confirming the initial underlying assumption that the likelihood of precipitation is inversely proportional to  $R(t)$ . These relationships were modulated strongly by variability of monthly precipitation ( $\tilde{\sigma}_P$ ). Equation 14 provided a good description of these relationships (Table 1).

The Beta distribution captured well the changes in distribution of the ratio  $P(t)/P$ , the proportion of a month's total precipitation falling on a given day, with respect to  $\tilde{\sigma}_P$ . Equations 18 and 19 accurately described the relationships between the  $\nu$  and  $w$  parameters of the Beta( $\nu$ ,  $w$ ) distribution and  $\tilde{\sigma}_P$  (Table 1; Fig. 5).

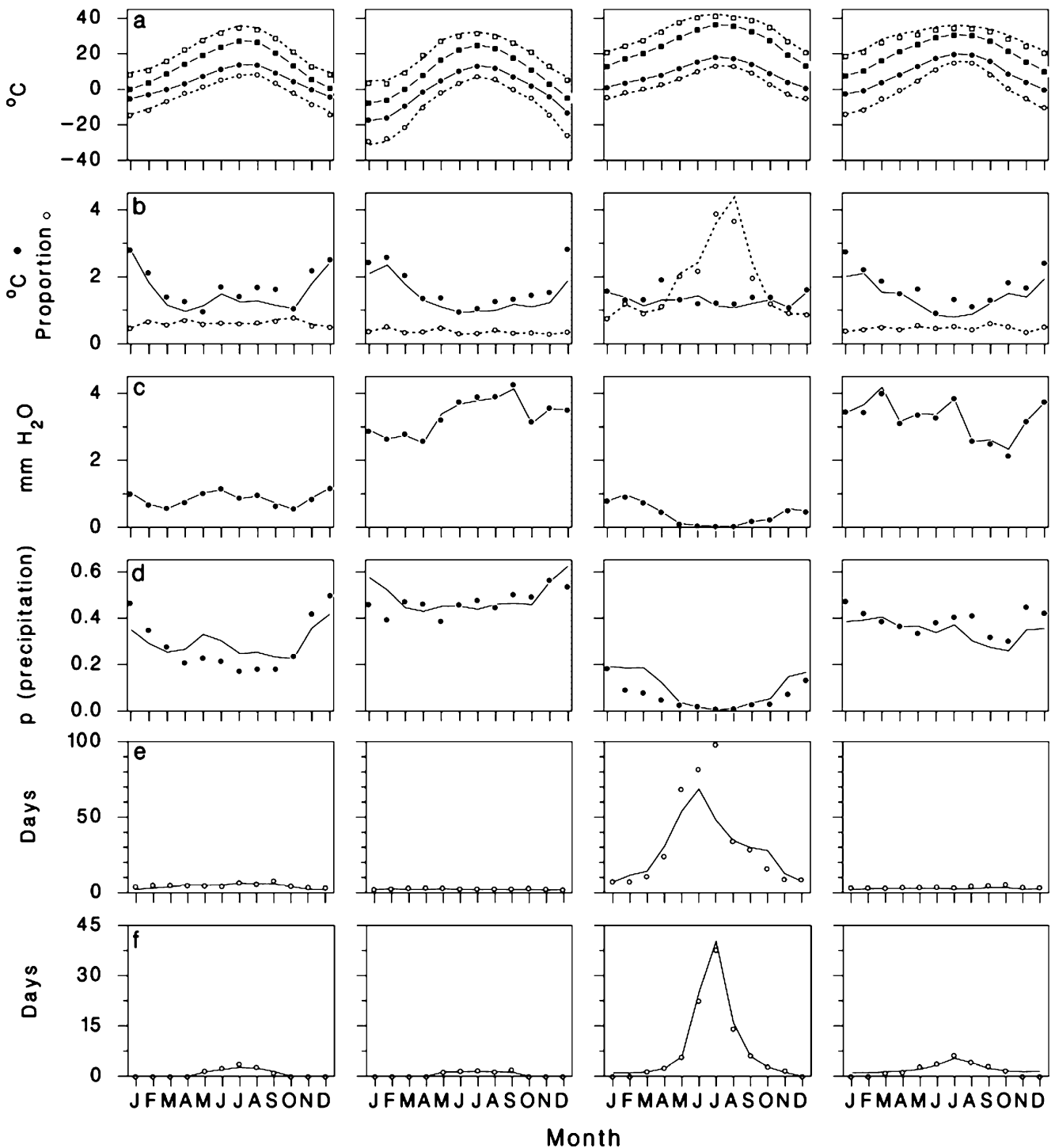
**Model validation**

Simulated and observed monthly temperature fluctuations ( $\tilde{\mu}_n, \tilde{\mu}_x$  and mean extreme minimum and maximum) were nearly identical at the four stations selected for detailed examination (Summerland BC, Quebec QC, Knoxville TN and Button Willow CA; Fig. 6a). Seasonal fluctuations of the variance of mean monthly temperature and total precipitation ( $\sigma_\mu$  and  $\tilde{\sigma}_P$ ) were also well mimicked by the

model (Fig. 6b). However, the simulated values of  $\sigma_\mu$  tended to be lower than observed (closed circles, Fig. 6b). Simulated and observed daily precipitation averages were nearly identical (Fig. 6c). Monthly variations of the probability of simulated daily precipitation varied in good general agreement with observations (Fig. 6d). The simulated seasonal patterns of drought and hot-spell duration were very similar to observations, although simulated droughts tended to be shorter than observed in Button Willow CA (Fig. 6e,f).

The distributions of observed and simulated daily minimum and maximum temperatures and of daily ranges were nearly identical (Fig. 7a,b). The distribution of the frost-free period was well simulated by the model in three of the four stations (Fig. 7c). In Button Willow CA, observed frost-free periods were about 10 days longer than simulated. The distribution of hot spell duration was well mimicked by the model in all four locations (Fig. 7d). While the distributions of simulated and observed single-day precipitation amounts (Fig. 7e) and droughts (Fig. 7f) were quite similar in three of the four stations, both simulated distributions tended toward higher values than observed in Button Willow CA.





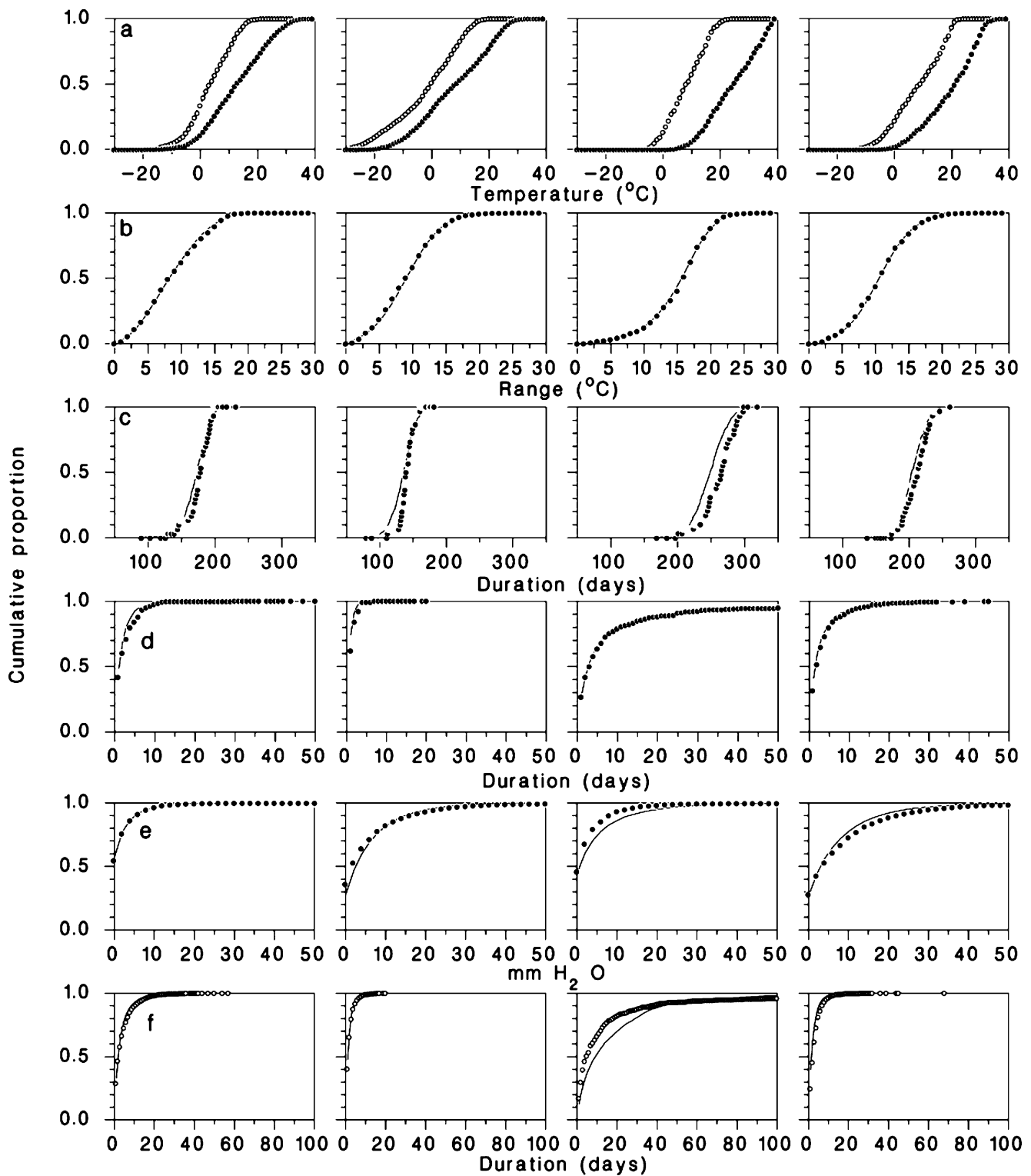
**Fig. 6a–f** Comparison of simulated and observed monthly statistics for four selected stations. Columns from left to right: 1 Summerland BC, 2 Quebec QC, 3 Button Willow CA, 4 Knoxville TN. Symbols observations, lines model. **a** Extreme and mean minimum and

**Month**

maximum temperatures. **b** Standard deviation of mean temperature (—•—) and precipitation (—○—). **c** Average daily precipitation. **d** Daily probability of precipitation. **e** Drought duration. **f** Hot-spell duration

Over the 708 stations in the validation data set, 9 of the 11 monthly normals calculated from model output were for all practical purposes identical to observed normals in terms both of means and variances ( $\tilde{\mu}_n, \tilde{\mu}_x, \tilde{\sigma}_\delta, \tilde{\sigma}_\varepsilon, \tilde{A}_\delta, \tilde{B}_\delta, \tilde{A}_\varepsilon, \tilde{B}_\varepsilon, \tilde{r}_{\delta\varepsilon}$ ; Table 2). Simulated total monthly precipitation  $\tilde{P}$

was slightly lower than observed, with a mean error (ME) of  $0.3 \pm 2.7$  mm/month and a mean absolute error (MAE) of  $1.7 \pm 1.9$  mm/month. However, these differences were very small (0.5% on average) in comparison to the observed average monthly precipitation of  $63.6 \pm$



**Fig. 7a–f** Comparison of cumulative distributions of simulated and observed statistics for four selected stations. Columns from left to right: 1 Summerland BC, 2 Quebec QC, 3 Button Willow CA, 4 Knoxville TN. Symbols observations, lines model. **a** Daily minimum

( $\circ$ ) and maximum ( $\bullet$ ) temperatures. **b** Daily temperature range. **c** Frost-free period. **d** Hot-spell duration. **e** Amount of daily precipitation. **f** Drought duration

**Table 2** Comparison of observed and simulated statistics from 708 randomly-selected weather stations

Normal <sup>a</sup>	Observed Mean±SD	Simulated Mean±SD	Error Mean±SD	Absolute error Mean±SD	r
<b>Input normals</b>					
$\tilde{\mu}_n(^{\circ}\text{C})$	1.9±11.7	1.9±11.7	0.0±0.1	0.1±0.1	0.999
$\tilde{\mu}_x(^{\circ}\text{C})$	14.5±13.1	14.5±13.1	0.0±0.1	0.1±0.1	0.999
$\tilde{\sigma}_{\delta}(^{\circ}\text{C})$	4.7±1.8	4.7±1.8	0.0±0.1	0.1±0.1	0.999
$\tilde{\sigma}_{\varepsilon}(^{\circ}\text{C})$	5.1±1.6	5.1±1.6	0.0±0.1	0.1±0.1	0.999
$\tilde{A}_{\delta}$	0.66±0.14	0.66±0.14	0.00±0.01	0.01±0.01	0.997
$\tilde{B}_{\delta}$	-0.05±0.09	-0.05±0.09	0.00±0.01	0.01±0.01	0.993
$\tilde{A}_{\varepsilon}$	0.73±0.12	0.73±0.12	0.00±0.01	0.01±0.01	0.996
$\tilde{B}_{\varepsilon}$	-0.09±0.08	-0.09±0.08	0.00±0.01	0.01±0.01	0.991
$\tilde{r}_{\delta\varepsilon}$	0.576±0.198	0.585±0.196	0.008±0.014	0.012±0.011	0.997
$\tilde{P}(\text{mm})$	63.6±47.8	63.3±47.5	-0.3±2.7	1.7±1.9	0.998
$\tilde{\sigma}_P$	0.69±0.32	0.69±0.35	0.00±0.07	0.03±0.06	0.984
<b>Non-input monthly statistics</b>					
Mean extreme min ( $^{\circ}\text{C}$ )	-6.9±14.0	-7.2±14.3	-0.3±0.9	0.7±0.7	0.998
Mean extreme max ( $^{\circ}\text{C}$ )	23.1±11.3	24.1±11.5	1.1±1.0	1.2±0.8	0.996
St. Dev. of mean ( $^{\circ}\text{C}$ )	1.9±0.9	1.6±0.8	-0.3±0.3	0.3±0.3	0.966
Daily range ( $^{\circ}\text{C}$ )	12.1±3.1	12.1±3.1	0.0±0.1	0.0±0.0	1.000
Daily pptn. (mm)	2.1±1.6	2.1±1.6	0.0±0.1	0.1±0.1	0.998
Daily pptn. prob. (mm)	0.287±0.133	0.267±0.152	-0.020±0.083	0.067±0.052	0.840
<b>Annual or extreme events</b>					
5% lowest min ( $^{\circ}\text{C}$ )	-16.4±11.3	-16.2±11.0	0.1±0.7	0.5±0.5	0.998
95% highest max ( $^{\circ}\text{C}$ )	29.8±5.3	29.8±6.1	0.0±0.5	0.2±0.4	0.997
Heaviest pptn. (mm)	25.0±12.0	27.7±11.6	2.7±9.5	6.9±7.0	0.677
Frost-free period (days)	150.7±67.6	151.2±68.0	0.5±6.6	4.9±4.5	0.995
Longest hot-spell (days)	15.0±18.8	13.1±19.2	-1.9±3.9	2.6±3.5	0.979
Longest drought (days)	18.1±13.5	18.2±13.3	0.1±7.5	4.8±5.8	0.846

<sup>a</sup> First group: input normals. Second group: additional monthly statistics ( $n=8,496$ ). Third group: mean annual value (frost-free period) or station extreme values ( $n=708$ )

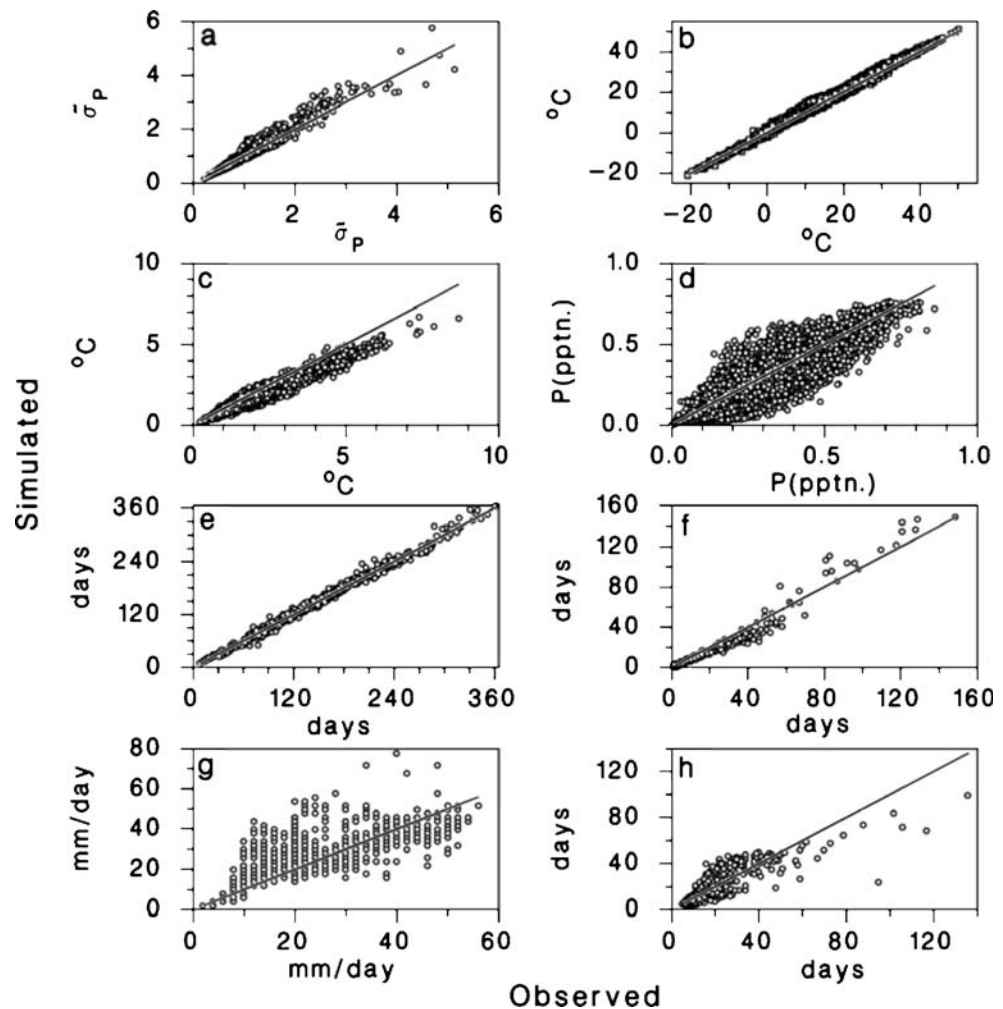
47.8 mm/month and the correlation between observed and average total monthly precipitation was very high ( $r=0.998$ ). While the ME and MAE between observed and simulated variation of monthly total precipitation  $\tilde{\sigma}_P$  were very small (Table 2), the correlation between them was not as high as for other normal statistics ( $r=0.984$ ). This indicates that the model realistically mimics this variation, but does not do this exactly, especially for higher values of  $\tilde{\sigma}_P$  (Fig. 8a).

The model generated accurate average monthly mean extreme minimum temperatures (Table 2) that were nearly unbiased (ME:-0.3±0.9°C; MAE: 0.7±0.7°C) and were highly correlated with observations ( $r=0.998$ ). Simulated average monthly maximum temperature was clearly biased (ME:1.1±1.0°C, MAE 1.2±1.2°C), despite a high correlation between observed and simulated values ( $r=0.996$ ; Fig. 8b). To explain this bias, the distributions of observed and simulated deviations of maximum temperature from daily averages,  $\varepsilon(t)$  in equation [2], were compared for the four stations selected for detailed analysis. While the distributions of simulated deviations were nearly normal, observed deviations had lower than expected frequencies at

the higher extremes (> 98th percentile), in all four stations and throughout the year (Fig. 9). This difference in distributions between simulated and observed  $\varepsilon(t)$  is the cause of the bias in mean monthly extreme maximum temperatures output by the model. Limiting the values of  $\varepsilon(t)$  to the range  $\pm 1.96\tilde{\sigma}_{\varepsilon}$  in the calculation of  $T_x(t)$  (Eq. 2) corrects this bias very efficiently without impacting other output characteristics.

The standard deviations of mean monthly temperature  $\sigma_{\mu}$  generated by the model were smaller than observed (ME:-0.3±0.3; MAE:0.3±0.3; Table 2), although the correlation between observed and simulated values was high ( $r=0.966$ ). This bias represents a 15% difference on average (Fig. 8c) and was most pronounced at higher values of  $\sigma_{\mu}$  (Fig. 8c). In the model, this variation results entirely from the variances ( $\tilde{\sigma}_{\varepsilon}, \tilde{\sigma}_{\delta}$ ), autocorrelation ( $\tilde{A}_{\delta}, \tilde{B}_{\delta}, \tilde{A}_{\varepsilon}, \tilde{B}_{\varepsilon}$ ) and cross-correlation ( $\tilde{r}_{\delta\varepsilon}$ ) structure of the stochastic processes  $\delta$  and  $\varepsilon$  in Eq. 3. There was a strong spatial pattern in observed values of  $\sigma_{\mu}$  over the North American continent, with the higher values occurring in the Prairie Provinces and around Hudson’s Bay in Canada (Fig. 10a). The differences between model output

**Fig. 8a–h** Comparison of selected observed (ordinate) and simulated (abscissa) statistics at 708 randomly selected weather stations, either a single annual value ( $n=708$ ) or 12 monthly values per station ( $n=8496$ ) (see Fig. 1 for geographical distribution). Lines: equality. **a** Standard deviation of monthly precipitation ( $n=9,496$  records). **b** Mean extreme maximum ( $n=8,496$ ). **c** Standard deviation of monthly mean ( $n=8,496$ ). **d** Average daily probability of precipitation ( $n=8,496$ ). **e** Frost-free period ( $n=708$ ). **f** Extreme (95th percentile) hot spell duration ( $n=708$ ). **g** Extreme (95th percentile) daily precipitation ( $n=708$ ). **h** Extreme (95th percentile) drought duration ( $n=708$ )

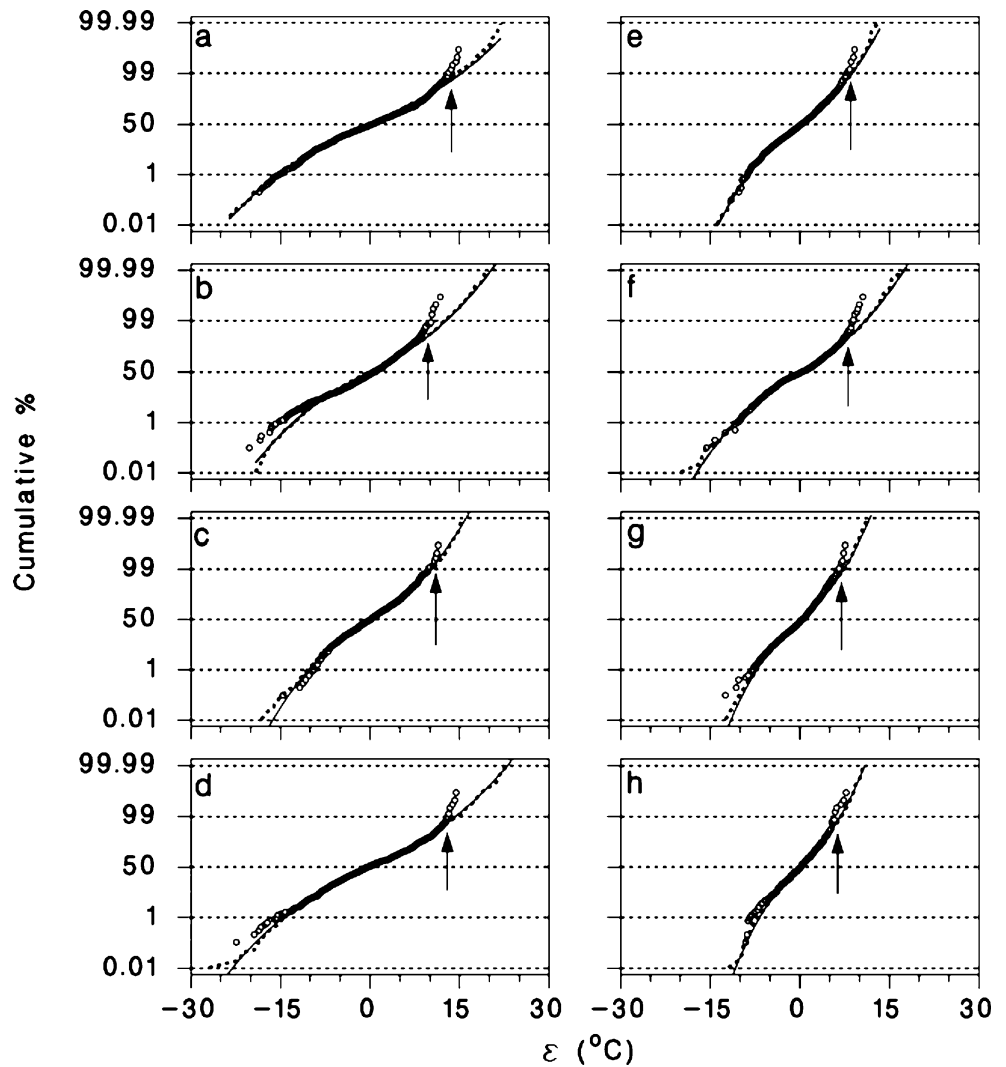


and observations being proportional to  $\sigma_{\mu}$ , they had a very similar spatial pattern (Fig. 10b).

The average daily ranges and daily precipitation generated by the model were nearly identical to observations, with no bias and very high correlation coefficients (Table 2). However, average simulated daily probability of precipitation was lower than observed (ME:  $0.020 \pm 0.083$ ; MAE:  $0.067 \pm 0.052$ ; Table 2). Because of the high variability of daily precipitation probability averaged by month and station, the correlation between simulated and observed means was only 0.84 (Fig. 8d). The daily probability of precipitation in North America is strongly patterned: it is highest in the Pacific Northwest and the Northeast (especially northern Ontario and Quebec), and lowest in the Southwest (Fig. 10c). Differences between model output and observations were also patterned spatially (Fig. 10d). The largest negative differences (simulated < observed) were located in the mountains of the West, the Canadian Prairies and at the latitude of James Bay in eastern Canada. Positive differences were concentrated in the eastern half of the continent, south of the Great Lakes, and along the Southern half of Canada's Atlantic Coast.

The model generated highly realistic distributions of extreme temperatures (5th percentile lowest daily minimum and 95th percentile highest daily maximum). The values of these two statistics generated by the model were highly correlated with observations, with low ME and MAE (Table 2). Average frost-free periods were also quite well predicted, with no bias and a small MAE of  $4.9 \pm 4.5$  days (Fig. 8e). The model predicted reasonably well the duration of extreme hot spells (95th percentile longest) with an MAE of only  $2.6 \pm 3.5$  days, relative to the observed average of  $15.0 \pm 18.8$  days. The correlation between simulated and observed extreme hot spells was high ( $r=0.979$ ; Fig. 8f). However, the model did not perform so well in predicting extreme daily precipitation (95th percentile highest values) or extreme drought durations (95th percentile longest). The model generated extreme precipitation events that were only slightly larger than observed on average (ME:  $2.7 \pm 9.5$  mm/day), but the the MAE was quite high ( $6.9 \pm 7.0$  mm/day) relative to the observed average of  $25.0 \pm 12.0$  mm/day (Fig. 8g). Extreme precipitation events have a strong spatial pattern in North America, and are largest in the Pacific Northwest and

**Fig. 9a–h** Comparison of the distribution of observed and simulated daily deviations of maximum temperature from their means at four selected weather stations. Rows: 1 Quebec QC, 2 Summerland BC, 3 Button Willow CA, 4 Knoxville TN. **a–d** January. **e–h** July.  $\circ$  Observed, — simulated, — normal distribution. Arrows indicate the systematic shortage of very high values among observations



around the Gulf of Mexico (Fig. 10e). The model tended to predict heavier extreme precipitation than observed in the western half of the United States, and lighter than observed in the Southeast (Fig. 10f). Average predicted extreme drought duration was unbiased (ME:  $0.1 \pm 7.5$  days) but absolute differences between model output and observations were high (MAE:  $4.8 \pm 5.8$  days) compared to the observed average ( $18.1 \pm 13.5$  days; Fig. 8h). The model underestimated the extreme summer droughts observed in the southwestern corner of the continent, but overestimated those over the high-elevation terrain of much of the western United States (Fig. 10g,h).

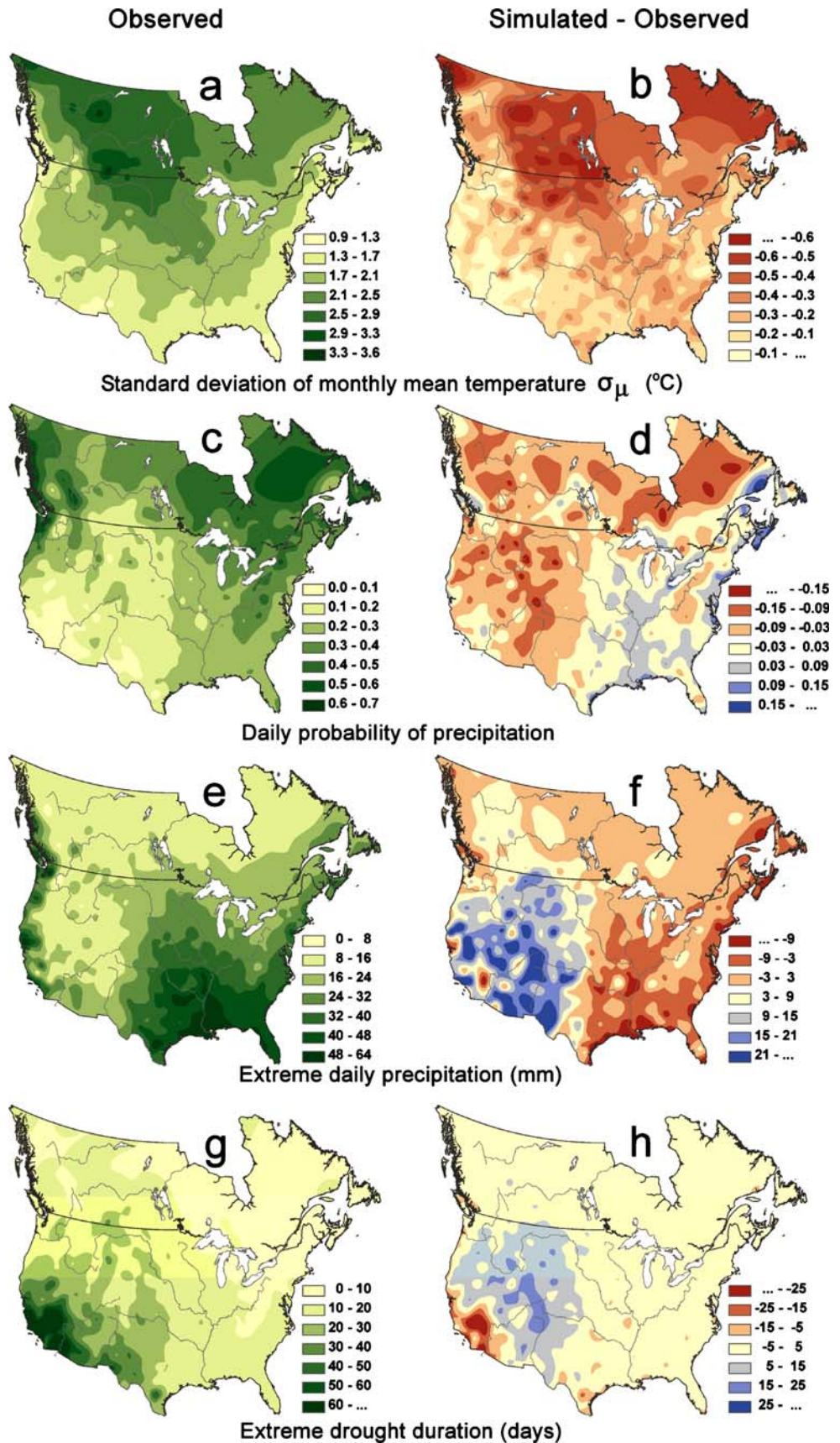
**Discussion**

Two of the objectives of this modification of Régnière and Bolstad’s (1994) TEMPGEN algorithm were to generalize it to the entire North American continent, north of Mexico, and to generate realistic interannual variability in monthly

means. These objectives were achieved in two ways: (1) by replacing empirical relationships between station latitude and the variance, autocorrelation and cross correlation of daily minimum and maximum temperature time series by input station-specific monthly statistics ( $\tilde{\sigma}_\delta, \tilde{\sigma}_\varepsilon, \tilde{A}_\delta, \tilde{B}_\delta, \tilde{A}_\varepsilon, \tilde{B}_\varepsilon, \tilde{r}_{\delta\varepsilon}$ ); and (2) by generalizing the stochastic processes (Eq. 3) that generate the daily fluctuations around normal monthly (mean) temperatures ( $\tilde{\mu}_m, \tilde{\mu}_x$ ). Comparison of the characteristics of observed and simulated daily temperature time series indicates that these objectives have been met successfully in terms of both seasonality and accuracy. The model also produces highly realistic frost-free periods and extreme events such as very low or very high temperatures and hot spells.

The model-generated interannual variation of monthly mean temperatures, which results entirely from the stochastic processes generating daily minimum and maximum temperature time series, is 15% smaller than observed variability. This suggests that other sources of variability are at play in observed temperature fluctuations, such as

**Fig. 10a–h** Maps of selected statistics from 708 randomly selected weather stations (see Fig. 1 for geographical distribution). Columns: *Left* Observed values, *right* simulated-observed differences. **a, b**  $\sigma_{\mu}$ . **c, d** Daily probability of precipitation. **e, f** Extreme (95th percentile) daily precipitation. **g, h** Extreme (95th percentile) drought duration



low-frequency fluctuations including the El Niño/Southern Oscillation (Philander 1990), the North Atlantic Oscillation (Watanabe and Nitta 1999) and the Pacific Decadal Oscillation (Latif and Barnett 1994). No attempt has been made at this point to include such synoptic low-frequency sources of temperature variability in TEMPGEN. Yet low-frequency interannual variability can have important consequences in ecological systems and needs to be taken into consideration in the generation of daily weather data (Hansen and Mavromatis 2001; Mavromatis and Hansen 2001; Dubrovský et al. 2004). One simple way to achieve this without recourse to models of those complex climate oscillation processes would be to apply a suitably small amount of stochastic annual variation simultaneously to monthly mean minimum and maximum temperatures ( $\tilde{\mu}_n, \tilde{\mu}_x$ ).

The third objective of this work was to add daily precipitation to TEMPGEN's output. Many of the daily weather generators developed since the 1950s (Bruhn 1980; Richardson 1981; Racsco et al. 1991; Johnson et al. 1996; Kimball et al. 1997; Friend 1998; Skiles and Richardson 1998; Jones and Thornton 2000; Yu 2003; Hansen and Ines 2005) are based on the simulation of precipitation using Markov-chain transitions between wet and dry states (Chin 1977), and Alpha distributions of the daily amount of precipitation. The TEMPGEN generator here is not based on a distinction between wet and dry days. It approaches precipitation on the basis of an empirical relationship between probability of precipitation and daily temperature range. Rather than describing transitions between wet and dry periods, TEMPGEN's wet and dry spells are the result of the autocorrelation and cross-correlation structure of minimum and maximum daily temperatures. The distribution of dry spell durations is an emergent property, conditioned by expected total monthly rainfall and its interannual variability, both station-specific model inputs ( $\tilde{P}, \tilde{\sigma}_P$ ). The variability  $\tilde{\sigma}_P$  is used to generate stochastic values of total monthly precipitation. From these and daily temperature ranges, the daily probability of precipitation is calculated. The proportion of total monthly precipitation falling on any given day is a Beta-distributed random variable with parameters entirely specified by the input value of  $\tilde{\sigma}_P$ . In very dry climates (high  $\tilde{\sigma}_P$ ), such as in the southwestern United States in summer, most monthly precipitation falls within a few days. In less variable climates (low  $\tilde{\sigma}_P$ ), precipitation is more evenly distributed. TEMPGEN generates realistic amounts and seasonal patterns of monthly precipitation. The average daily amount of precipitation is also accurately simulated. However, the daily probability of precipitation generated by the model is slightly lower than observed (by about 2%). The model generates fairly accurate drought durations, except in the driest climates where it underestimates extreme drought

durations. The extreme daily precipitation events generated by the model are unbiased, but their correlation with observations is low ( $r = 0.0677$ ). It must be noted, however, that these extreme precipitation events are rare and that observations used for comparison with model output were based on only 30 years of data per station.

Many authors have used daily weather generators as a means of down-scaling the outputs of Global Circulation Models (CGMs) for climate change impact studies (Oelschlägel 1995; Bürger 1997; Dubrovský 1997; Semenov and Barrow 1997). This involves summarizing CGM outputs and generating daily weather from the resulting distributions (Qian et al. 2005), or altering monthly statistics by fixed amounts (Mearns et al. 1997). One of the problems with these approaches is that CGM outputs often do not reflect many of the statistical characteristics of actual daily weather fluctuations, especially precipitation patterns at regional scales. This is a particularly problematic issue with respect to dry-wet day models, as discussed by Katz (1996). Many daily weather generators require calibration for use at specific locations (Soltani and Hoogenboom 2003; Fowler et al. 2005). The TEMPGEN algorithm has the main advantage of requiring no input data other than the 11 sets of station-specific monthly normals. It can be used to generate daily minimum and maximum temperature and precipitation time series for any location in Canada or the continental United States when provided with local climate normals (whether actual or as a synthesis of CGM output). To provide such normals for locations not close to a weather station or at a finer spatial scale than CGM output, several spatial disaggregation methods are available, such as application of regional climate models or spatial interpolation on the basis of a network of weather station records and topography (Hutchinson 1991, 1995b; Gignac 2000; Price et al. 2000).

**Acknowledgements** Thanks to M. Lindner for ideas in the development of this model. Thanks also to D. Houle, D.W. McKenney and two anonymous reviewers for very helpful comments on an earlier draft. This work was funded in part by members of SERG-International, including Forest Protection Limited, the Quebec SOPFIM, the Province of Ontario, and the USDA Forest Service.

## References

- Bruhn JA (1980) A stochastic model for the simulation of daily weather. *Prot Ecol* 2:199–208
- Bürger G (1997) On the disaggregation of climatological means and anomalies. *Clim Res* 8:183–194
- Chin EH (1977) Modelling daily precipitation occurrence process with Markov chain. *Water Resour Res* 13:949–956
- Dubrovský M (1997) Creating daily weather series with use of the weather generator. *Environmetrics* 8:409–424

- Dubrovský M, Buchtele J, Žalud Z (2004) High-frequency and low-frequency variability in stochastic daily weather generator and its effect on agricultural and hydrologic modelling. *Clim Change* 63:145–179
- Fowler HJ, Kilsby CG, O'Connell PE, Burton A (2005) A weather-type conditioned multi-site stochastic rainfall model for the generation of scenarios of climatic variability and change. *J Hydrol* 308:50–66
- Friend AD (1998) Parameterisation of a global daily weather generator for terrestrial ecosystem modelling. *Ecol Model* 109:121–140
- Comparaison de la régression spatiale et du krigeage avec dérive pour interpoler les extrants de modèles de simulation de développement d'insectes au Québec en fonction de l'échelle, de la topographie et de l'influence maritime. M.Sc. Thesis, Faculté de foresterie et géomatique, Université Laval, Québec, Québec, Canada
- Gray DR (2004) The gypsy moth life stage model: landscape-wide estimates of gypsy moth establishment using a multi-generational phenology model. *Ecol Model* 176:155–171
- Hansen JW, Ines AVM (2005) Stochastic disaggregation of monthly rainfall data for crop simulation studies. *Agric For Meteorol* 131:233–246
- Hansen JW, Mavromatis T (2001) Correcting low-frequency variability bias in stochastic weather generators. *Agric For Meteorol* 109:297–310
- Hastings NAJ, Peacock JB (1975) *Statistical distributions: A handbook for students and practitioners*. Butterworth, London
- Hutchinson MF (1991) The application of thin-plate smoothing splines to continent-wide data assimilation. In: Jasper JD (ed) *Data assimilation systems*. BMRC Res Rep No. 27, Melbourne, pp 104–113
- Hutchinson MF (1995a) Stochastic space-time weather models from ground-based data. *Agric For Meteorol* 73:237–264
- Hutchinson MF (1995b) Interpolating mean rainfall using thin-plate smoothing splines. *Int J GIS* 9:385–403
- Johnson GL, Hanson CL, Hardegree SP, Ballard EB (1996) Stochastic weather simulation: overview and analysis of two commonly used models. *J Appl Meteorol* 35:1878–1896
- Jones PG, Thornton PK (2000) MarkSim: software to generate daily weather data for Latin America and Africa. *Agron J* 92:445–453
- Katz RW (1996) The use of stochastic models to generate climate change scenarios. *Clim Change* 32:237–255
- Kimball JS, Thornton PE, White MA, Running SW (1997) Simulating forest productivity and surface-atmosphere carbon exchange in the BOREAS study region. *Tree Physiol* 17:589–599
- Lasch P, Lindner M, Erhard M, Suckow F, Wenzel A (2002) Regional impact assessment on forest structure and functions under climate change - the Brandenburg case study. *For Ecol Manag* 162:73–86
- Latif M, Barnett TP (1994) Causes of decadal climate variability over the North Pacific and North America. *Science* 266:634–637
- Logan JA, Régnière J, Powell JA (2003) Assessing the impacts of global warming on forest pest dynamics. *Front Ecol Environ* 1:130–137
- Mavromatis T, Hansen JW (2001) Interannual variability characteristics and simulated crop response of four stochastic weather generators. *Agric For Meteorol* 109:283–296
- Mearns LO, Rosenzweig C, Goldberg R (1997) Mean and variance change in climate scenarios: methods, agricultural applications, and measures of uncertainty. *Clim Change* 35:367–396
- Oelschlägel B (1995) A method for downscaling global climate model calculations by a statistical weather generator. *Ecol Model* 82:199–204
- Philander SGH (1990) *El Niño, La Niña and the Southern Oscillation*. Academic, San Diego CA
- Prentice IC, Sykes MT, Cramer W (1993) A simulation model for the transient effects of climate change on forest landscapes. *Ecol Model* 65:51–70
- Price DT, McKenney DW, Nalder IA, Hutchinson MF, Kesteven JL (2000) A comparison of two statistical methods for spatial interpolation of Canadian monthly mean climate data. *Agric For Meteorol* 101:81–94
- Qian B, Hayhoe H, Gameda S (2005) Evaluation of the stochastic weather generators LARS-WG and AAFC-WG for climate change impact studies. *Clim Res* 29:3–21
- Racsko P, Szeidl L, Semonov M (1991) A serial approach to local stochastic weather models. *Ecol Model* 57:27–41
- Régnière J (1996) A generalized approach to landscape-wide seasonal forecasting with temperature-driven simulation models. *Environ Entomol* 25:869–881
- Régnière J, Bolstad P (1994) Statistical simulation of daily air temperature patterns in eastern North America to forecast seasonal events in insect pest management. *Environ Entomol* 23:1368–1380
- Régnière J, Nealis V (2002) Modelling seasonality of gypsy moth, *Lymantria dispar* (Lepidoptera: Lymantriidae), to evaluate probability of its persistence in novel environments. *Can Entomol* 134:805–824
- Richardson CW (1981) Stochastic simulation of daily precipitation, temperature and solar radiation. *Water Resour Res* 17:182–190
- Richardson CW, Wright DA (1984) WGEN: A model for generating daily weather variables. United States Department of Agriculture, Washington, ARS-8
- Semenov MA, Barrow EM (1997) Use of a stochastic weather generator in the development of climate change scenarios. *Clim Change* 35:397–414
- Semenov MA, Brooks RJ, Barrow EM, Richardson CW (1998) Comparison of the WGEN and LARS-WG stochastic weather generators for diverse climates. *Clim Res* 10:95–107
- Skiles JW, Richardson CW (1998) A stochastic weather generation model for Alaska. *Ecol Model* 110:211–232
- Soltani A, Hoogenboom G (2003) Minimum data requirements for parameter estimation of stochastic weather generators. *Clim Res* 25:109–119
- Wallis TWR, Griffiths JF (1995) An assessment of the weather generator (WXGEN) used in the erosion /productivity impact calculator (EPIC). *Agric For Meteorol* 73:115–133
- Watanabe M, Nitta T (1999) Decadal changes in the atmospheric circulation and associated surface climate variations in the Northern Hemisphere winter. *J Clim* 12:494–510
- Wilks DS (1999) Simultaneous stochastic simulation of daily precipitation, temperature and solar radiation at multiple sites in complex terrain. *Agric For Meteorol* 96:85–101
- Yu B (2003) An assessment of uncalibrated CLIGEN in Australia. *Agric For Meteorol* 119:131–148