

Performance of stochastic weather generators LARS-WG and AAFC-WG for reproducing daily extremes of diverse Canadian climates

Budong Qian*, Samuel Gameda, Henry Hayhoe

Eastern Cereal and Oilseed Research Centre, Agriculture and Agri-Food Canada, 960 Carling Avenue, Ottawa, Ontario K1A 0C6, Canada

ABSTRACT: Stochastic weather generators are widely used for generating synthetic weather data, and constitute one of the techniques for developing local climate scenarios from large-scale climate changes simulated by global climate models. Since climate change impact models may be more sensitive to changes in climate extremes than to changes in climate means, there is a need to know the capability of stochastic weather generating techniques in reproducing climate extremes. An evaluation of 2 stochastic weather generators, namely LARS-WG and AAFC-WG, is presented in this study, from the perspective of reproducing observed climate extremes. Extreme daily values (highest daily maximum temperature, lowest daily minimum temperature and maximum daily precipitation) were analysed on a monthly and annual basis, as well as for the growing season (1 May to 30 September), for 9 stations across Canada. The evaluation was based on statistical tests for basic statistical properties and return values derived from fitted generalised extreme value distributions. Results showed that the weather generators LARS-WG and AAFC-WG could reproduce statistical properties of the extreme values of daily precipitation satisfactorily, including 10, 20 and 50 yr return values. However, the performance of the weather generators in reproducing extreme daily values of temperatures was not as good as for precipitation, especially that of LARS-WG. The deficiency in reproducing extremely low temperatures was also more noticeable than extremely high temperatures for LARS-WG. The mismatches of return values were often caused by a more extreme value estimated from synthetic data than that derived from observations.

KEY WORDS: Stochastic weather generator · Extreme daily value · Canada

Resale or republication not permitted without written consent of the publisher

1. INTRODUCTION

Stochastic weather generators are statistical models that can generate synthetic weather data mimicking the weather data to which they have been calibrated (Hutchinson 1986, Wilks & Wilby 1999). They were developed originally for 2 main purposes (Semenov & Barrow 1997). The first was the provision of a technical tool for simulating synthetic weather series with certain statistical properties mimicking the observed ones, while the second was to provide a tool for extending the simulation of weather data to locations where no observed weather data are available. A suffi-

ciently long synthetic weather series could be generated for use in assessments of risk in hydrological and agricultural applications. This is useful because the observed weather data are often of insufficient length to allow the estimation of the probability functions of rare events. According to the first purpose, the capability of a stochastic weather generator in simulating extremes is crucial, since extremes are rare events with a small probability of occurrence.

Stochastic weather generators have assumed a new role as one of the means of developing climate scenarios for climate change impact studies (Wilks 1992, Katz 1996, Semenov & Barrow 1997, Mearns et al. 1997,

*Email: qianb@agr.gc.ca

Qian et al. 2005). While change in the long-term climatic mean state will have many important consequences, the most acute effects of climate change may come about from changes in the intensity and frequency of climatic extremes (Zwiers & Kharin 1998). One of the principal means by which crops are affected is through changes in the frequency of extreme climate events (e.g. heat waves; Mearns et al. 1984). Katz & Brown (1992) used statistical theory for extremes to demonstrate that the frequency of such events is relatively more dependent on any changes in the variability (more generally, the scale parameter) than in the mean (more generally, the location parameter) of climate. Stochastic weather generators, such as LARS-WG (Racsko et al. 1991, Semenov et al. 1998, Semenov & Barrow 2002), a weather generator developed at the Long Ashton Research Station (UK), and AAFC-WG (Hayhoe 2000, Qian et al. 2004), a weather generator developed at Agriculture and Agri-Food Canada, apply the changes in variance as ratios to perturb weather generator parameters for generating future climate scenarios (Semenov & Barrow 2002, Qian et al. 2005). Therefore, future climate scenarios developed by stochastic weather generators are expected to represent changes in the intensity and frequency of climatic extremes associated with changes in both climate mean and variability.

Stochastic weather generators are used to generate a time-series of daily synthetic weather that is statistically 'identical' to observations. Semenov & Barrow (1997) warned that 'it must be borne in mind that statistical 'identity' depends on the number of statistics used for the comparison'. This implies that stochastic weather generators are expected to generate a synthetic weather series with a number of statistical properties resembling the observed weather. Such statistical properties are often relevant to a specific application. Extreme events occur with a small probability; thus a weather generator may be able to reproduce the observed extremes if the generated synthetic weather data have empirical distributions that are not statistically different from observed ones. For example, the extreme value corresponding to the 90th percentile of July daily maximum temperatures can be reasonably reproduced by a weather generator if the weather generator can generate July daily maximum temperatures following an empirical distribution that is not significantly different from that of observed July daily maximum temperatures. However, this does not assure that the return values (i.e. extreme values expected to be exceeded once every return period, e.g. 20 and 50 yr) of yearly maxima for July daily maximum temperatures for certain periods can be accurately reproduced by the weather generator. This is because the return values in this example are derived

from the probability distribution of yearly maxima of July daily maximum temperatures rather than the probability distribution from which the maxima were drawn. The yearly maxima of July daily temperatures here is a maximum value of the 31 daily maximum temperatures for each year. Stochastic weather generators are often known for the 'overdispersion' phenomenon, i.e. interannual variability in the synthetic series of aggregative statistics (such as monthly, seasonal and annual precipitation totals) is smaller than in observed ones (Katz & Parlange 1998). This overdispersion phenomenon may also be a concern for the yearly extreme daily values and their return values; however, a recent study (Semenov 2008) indicated that underestimation of variance for monthly means does not necessarily affect simulation of extremes.

Conventional models, such as the gamma distribution and the mixed exponential distribution for daily precipitation, were examined for obtaining a weather generator formulation with a better treatment of extreme values (Wilks 1999). Empirical distributions are used in some weather generators, such as LARS-WG and AAFC-WG, for their flexibility to diverse climates compared with conventional models. The methodology used to construct the empirical distribution, even pre-processing of the data, such as standardisation and transformation, may have effects on the derived empirical distribution. AAFC-WG applies a logarithmic transformation and standardisation to non-zero daily precipitation data before producing the empirical distribution. The transformed data more closely approximate the normal distribution; thus, modifying distributions with means and standard deviations seems more reasonable for generating future scenarios (Qian et al. 2005). It is worthwhile to mention that there are straightforward relations between parameters and moments for many distributions other than the normal distribution. Such relations make the methodology of modifying distributions with means and standard deviations practical in the development of future climate scenarios by statistical approaches. Kysely & Dubrovsky (2005) evaluated the performance of a Richardson-type (Richardson 1981) weather generator, Met&Roll (Dubrovsky 1997, Dubrovsky et al. 2004), regarding the simulation of extreme temperature events across Europe. They found that the introduction of a monthly generator into Met&Roll, which was intended to improve the overdispersion problem, could not improve the simulation of annual extreme daily values or the 20 yr return values. Furthermore, they noted that the discrepancies between the simulated and the observed temperature extremes were dependent upon the region in Europe, suggesting that daily

maximum and minimum temperatures in some regions depart considerably from a normal distribution assumed by the weather generator. Semenov (2008) also noted the possible effects of such an assumption for temperatures in the simulation of extremes.

The present paper aims to evaluate 2 stochastic weather generators—LARS-WG and AAFC-WG—in terms of their capability for reproducing climate extremes: monthly, annual and growing season (GS) extreme daily temperatures and precipitation, in diverse Canadian climates. It is of interest to investigate whether the assumption of a normal distribution often used in the simulation of daily temperatures and the ways used to construct empirical distributions, together with some other factors, have marked effects on the reproduction of extreme daily values. The findings from such a study may be useful in seeking appropriate methodologies for the improvement of weather generation with regards to climate extremes. It seemed appropriate for this study to use the 2 above-mentioned weather generators. LARS-WG uses the assumption of a normal distribution for daily temperatures, and it has been widely applied to weather generation around the world. In contrast, AAFC-WG uses empirical distributions of observed daily temperatures. The use of empirical distributions was proved effective with respect to basic statistical properties and agroclimatic indices, for temperature simulations in diverse Canadian climates (Qian et al. 2004). In addition, the ways used to construct empirical distributions are also somewhat different in the 2 weather generators. It should be kept in mind that the evaluation of a stochastic weather generator is always relevant to the targeted application and the properties under investigation. Therefore, we were not interested in finding a ‘best’ or ‘better’ weather generator; instead, we were interested carrying out an evaluation that would lead to improvements in weather generation algorithms.



Fig. 1. Nine stations for diverse Canadian climates

2. DATA AND METHODOLOGY

2.1. Daily climate data

Historical weather observations, including daily maximum temperature (T_{\max}), daily minimum temperature (T_{\min}) and daily precipitation (P), were used in the present study. The historical weather data at 9 stations across Canada (Fig. 1), for the period from 1971 to 2000 were extracted from the historical weather data sets archived at AAFC. These 9 stations are the same as those used in a previous study (Qian et al. 2004). They are reasonably representative of diverse Canadian climates, especially agricultural regions. Climate statistics for 1971 to 2000 at these stations can be referred to in Qian et al. (2004). The 30 yr long historical weather data were used to calibrate and evaluate the weather generators. The 30 yr interval is sufficiently long to filter out many of the short-term interannual fluctuations and anomalies, but sufficiently short so as to be used to reflect longer term climatic trends.

2.2. LARS-WG and AAFC-WG

The weather generators LARS-WG and AAFC-WG were used as representatives of 2 major groups of stochastic weather generators according to their structure: the serial weather generator and the Richardson-type weather generator. Some major differences between the 2 weather generators are: (1) LARS-WG uses empirical distributions of wet and dry spells for the simulation of wet and dry day sequences, while AAFC-WG applies a second-order, 2-state Markov chain to generate wet and dry day sequences; (2) normal distributions are assumed for daily T_{\max} and T_{\min} in LARS-WG, whereas empirical distributions are used to simulate daily T_{\max} and T_{\min} in AAFC-WG; and (3) constant Lag-1 autocorrelation and preset cross-correlation between maximum and minimum temperatures are used in LARS-WG, while Lag-0 and Lag-1 correlation matrices estimated bimonthly from observations are applied in AAFC-WG. A brief description of the 2 weather generators is given below, but a detailed comparison of how LARS-WG and AAFC-WG simulate each weather variable can be found in Table 1 of Qian et al. (2004).

In both weather generators, daily maximum and minimum temperatures are simulated conditional on

the occurrence of precipitation. Therefore, precipitation occurrence on a day is determined first. LARS-WG (Semenov & Barrow 2002) generates a wet and dry day sequence by the use of empirical distributions of the lengths of wet and dry spells instead of Markov chains. The empirical distribution $\text{Emp} = \{a_0, a_i, h_i, i = 1, 2, \dots, 10\}$ is a histogram with 10 intervals $[a_{i-1}, a_i)$, where $a_{i-1} < a_i$ and h_i denotes the number of events from the observed data in the i th interval. The intervals $[a_{i-1}, a_i)$ are chosen based on the expected properties of the weather variables. They are derived from the observed data and are not pre-set. For the lengths of dry and wet series and for precipitation, the interval size gradually increases as i increases because there are typically many small values, but also a few very large ones, and this choice of interval structure prevents a very coarse resolution being used for the small values. Random values from the empirical distributions are drawn by first selecting one of the intervals (using the proportion of events in each interval as the selection probability) and then selecting a value within that interval from the uniform distribution.

AAFC-WG (Hayhoe 2000) adopts a second-order Markov chain approach in simulating sequences of wet and dry days. A day is wet or dry depending on the state (wet/dry) of its previous 2 d and the corresponding transition probabilities ($p_{001}, p_{011}, p_{101}$ and p_{111} : the probability of a wet day following 2 consecutive dry days, a dry day followed by 2 wet days, a wet day followed by a dry day and a wet day, and 3 consecutive wet days, respectively). The transition probabilities are estimated from bimonthly observations.

A random number is then generated as daily precipitation for a simulated wet day as follows. LARS-WG and AAFC-WG both employ the empirical distribution of daily precipitation with some different data treatment. LARS-WG uses the above-mentioned procedure for the construction of the empirical distribution and random number generation. Daily precipitation amounts on wet days are first log-transformed and then standardised in AAFC-WG. The empirical distribution (the cumulative frequency distribution) is then created from standardised log-transformed daily precipitation amounts. The data are sorted in ascending order, and then the probability of exceeding any specified level is estimated from the relative frequency that occurred in the observed data. A total of 101 grid points are used so that the relative frequencies are calculated to decimal places over the full range in probability from 0 to 1. Cubic spline interpolation is used to derive the continuous cumulative probability distribution. Therefore, a random number generated from the empirical distribution needs back-transformation to be a daily precipitation amount.

Daily maximum and minimum temperatures are considered as stochastic processes with daily means and standard deviations conditioned on the wet or dry status of a day in LARS-WG. The seasonal cycles of means and standard deviations are modelled by finite Fourier series, and the residuals are approximated by a normal distribution. Autocorrelations for the residuals of daily maximum and minimum temperatures are assumed constant through the year, with the average value from the observed data, while the cross-correlation between the residuals of daily maximum and minimum temperatures are pre-set to 0.6. Technical details and mathematical formulas can be found in Semenov & Barrow (2002) and Racsco et al. (1991).

The distributions of daily maximum and minimum temperatures are also conditional on precipitation occurrence in AAFC-WG and they are simulated by a generalized first-order multivariate autoregressive model as in the Richardson-type weather generator (Richardson 1981). Means and standard deviations of daily maximum and minimum temperatures for wet days and dry days, calculated for each month from observations, are interpolated to each calendar day in the year by a spline interpolation procedure (Press et al. 1992) instead of using the fitted Fourier series as in LARS-WG. The interpolated means and standard deviations for wet and dry days are then used to standardise the observed daily maximum and minimum temperature series by subtracting the daily mean and dividing by the standard deviation (Richardson 1981). A generalized first-order, autoregressive model is used to model the standardised daily maximum and minimum temperatures (sometimes referred as residuals). The Lag-1 and Lag-0 correlation matrices are calculated from the standardised series on a bimonthly basis. The bimonthly empirical distributions for the standardised daily maximum and minimum temperatures are created in the same way as for precipitation. The constructed bimonthly empirical distributions are then transformed using a normal score transformation (Johnson 1987) in order to be used in a first-order, autoregressive model based on the bimonthly correlation matrices (Hayhoe 2000). This approach is addressing the normality assumption of a simple autoregressive model, which is not always valid in diverse climates. Random numbers generated from the autoregressive models are back-transformed, multiplied by the standard deviation and then added to the daily mean of either daily T_{\max} or T_{\min} , depending on whether the day is wet or dry. A detailed description of the procedure and mathematical formulation can be found in Hayhoe (2000).

2.3. Extreme daily values

In addition to annual extreme values, extreme daily values were also examined on a monthly basis in order to determine seasonal differences in the performance of the stochastic weather generators. Growing season (1 May to 30 September) climatic conditions are often more important for agricultural applications. Extreme daily values for the growing season were considered, as stochastic weather generators are widely applied in agricultural studies. The climate extremes focused on in the present study are as follows: (1) Maximum temperature in a month, year and growing season, i.e. the highest daily T_{\max} in a month, year and growing season for each year; (2) Minimum temperature in a month, year and growing season, i.e. the lowest daily T_{\min} in a month, year and growing season for each year; (3) Maximum daily precipitation in a month, year and growing season, i.e. the maximal daily precipitation amount in a month, year and growing season for each year.

2.4. Evaluation strategy

Yearly extremes (maxima of daily T_{\max} , $-T_{\min}$ and P) generated by stochastic weather generators LARS-WG and AAFC-WG were compared, separately, with those from the observed climate data used to calibrate the weather generators. The maxima of $-T_{\min}$ corresponds to the minima of T_{\min} ; thus, the maxima of $-T_{\min}$ were used for computational convenience with the other 2 climate variables. A 30 yr long data set is typical for calibrating a weather generator, although a 30 yr period of yearly extreme values may not necessarily be sufficient for an accurate estimation of the extremes exceeding certain probability levels (e.g. 95th percentiles for the 20 yr return value, which is exceeded approximately once in 20 yr) from the observations. However, generating long time series for the estimation of rare events is one of the original purposes of the development of stochastic weather generators. The generated synthetic weather data are 300 yr long in this study, the same as in Semenov et al. (1998) and Qian et al. (2004). Time series of yearly maxima of the above-mentioned extreme daily values were derived for each station. The series are 30 and 300 yr long for the observed data and the synthetic data generated by the weather generators, respectively. It should be kept in mind that the synthetic data only represents the climate of 30 yr observations used to calibrate the weather generators no matter how long the synthetic data series is. It does not account for possible climate change in 100 yr, although a 100 yr return value can be estimated that is relevant only to the 30 yr observed

climate. It is also worth mentioning that a 30 yr data series in an unchanging climate will probably also not contain the order of extreme events that could be observed over 100 yr.

Statistical tests were performed to verify if the difference in yearly maxima between the synthetic data and observations is statistically significant for basic statistical properties, such as means, variances and probability distributions. As yearly maxima may not follow a normal distribution, statistical tests for means and variances were conducted based on resampling (Wilks 1995, Qian et al. 2005). These resampling tests for 2 sample problems are also known as permutation tests. The test statistic for the means is the absolute difference between the means estimated from the 2 samples,

$$D = \left| \frac{1}{m} \sum_{i=1}^m x_{1i} - \frac{1}{n} \sum_{j=1}^n x_{2j} \right| \quad (1)$$

The test statistic for the variances is similar to the one for the F -test, where

$$\begin{aligned} F &= \frac{F_1}{F_2} \text{ if } F_2 \leq F_1 \text{ and } F = \frac{F_2}{F_1} \text{ if } F_2 > F_1, \\ F_1 &= \frac{1}{m-1} \sum_{i=1}^m \left(x_{1i} - \frac{1}{m} \sum_{i=1}^m x_{1i} \right)^2 \text{ and} \\ F_2 &= \frac{1}{n-1} \sum_{j=1}^n \left(x_{2j} - \frac{1}{n} \sum_{j=1}^n x_{2j} \right)^2 \end{aligned} \quad (2)$$

The test statistics D and F were estimated from the observed weather series X_1 ($x_{11}, x_{12}, \dots, x_{1m}$) and the synthetic series X_2 ($x_{21}, x_{22}, \dots, x_{2n}$) generated by the stochastic weather generators. Critical values of the test statistics at the 5% significance level were then determined from the empirical distributions of the test statistics. The null hypothesis is that the 2 samples are equal with respect to the test statistic of interest. If the test statistic is larger than the critical value, the null hypothesis is rejected. A detailed procedure for permutation tests is given in Wilks (1995).

The Anderson-Darling (A-D) test (Anderson & Darling 1952) was used to test if the empirical distributions of yearly maxima derived from observations and the synthetic data are different at the 5% significance level, since extremes were the foci of the present study. The A-D test is a modification of the Kolmogorov-Smirnov (K-S) test and gives more weight to the tails than does the K-S test. The A-D 2-samples test involves 2 samples X_1 and X_2 with, respectively, m and n data. The test statistic A^2 provides a comparison between the 2 empirical distribution functions F and G , involving also the empirical distribution function of the combined sample H . The test statistic:

$$A^2 = \frac{m+n-1}{(m+n)^2} \left[\frac{1}{m} \sum_{j=1}^L h_j \frac{((m+n)F_j - mH_j)^2}{H_j(m+n-H_j) - \frac{(m+n)h_j}{4}} + \frac{1}{n} \sum_{j=1}^L h_j \frac{((m+n)G_j - nH_j)^2}{H_j(m+n-H_j) - \frac{(m+n)h_j}{4}} \right] \quad (3)$$

where h_j is the number of data in the combined sample equal to z_j ; H_j is the number of data in the combined sample smaller than z_j plus one-half of the number of data in the combined sample equal to z_j ; F_j is the number of data in sample X_1 that are smaller than z_j plus one-half of the number of data in this sample equal to z_j ; G_j is the number of data in sample X_2 that are smaller than z_j plus one-half of the number of data in this sample equal to z_j ; z_1, z_2, \dots, z_L are the distinct values in the combined sample ($x_{11}, x_{12}, \dots, x_{1m}, x_{21}, x_{22}, \dots, x_{2n}$) ordered from the smallest to the largest (L is less than $m+n$ if there are tied values). The null hypothesis that the populations from which 2 samples of data were drawn are identical is rejected at the 5% significance level if the test statistic A^2 is greater than the critical level $A_{0.05}^2 = 1.992743$.

The generalized extreme value (GEV) distribution (Coles 2001) was fitted to both the observed and the synthetic series of yearly maxima for the estimation of return values. One of the advantages of the use of the GEV distribution is the possibility to obtain a confidence interval of a return value. This is useful for the verification of whether the return values estimated from the observed data and the synthetic series, respectively, are significantly different. The use of confidence intervals is especially important, because the sample size (30 yr) for observations is much smaller than the synthetic ones (300 yr). The GEV distribution:

$$F(x) = \begin{cases} \exp\{-[1 - \kappa(x - \xi) / \alpha]^{1/\kappa}\}, \kappa < 0, x > \xi + \alpha / \kappa \\ \exp\{-\exp[-(x - \xi) / \alpha]\}, \kappa = 0 \\ \exp\{-[1 - \kappa(x - \xi) / \alpha]^{1/\kappa}\}, \kappa > 0, x < \xi + \alpha / \kappa \end{cases} \quad (4)$$

combines the 3 possible asymptotic extreme value (EV) distributions depending on the shape parameter κ . For $\kappa = 0$, the GEV distribution reduces to the Gumbel (EV-I) distribution (Gumbel 1958), which is a limiting distribution of extreme maximal values drawn from a parent population that may follow one of several common types of distributions, including normal and exponential (Leadbetter et al. 1983). The GEV distribution has a wider (narrower) tail than the Gumbel distribution when the shape parameter $\kappa < 0$ (> 0). The shape parameter, together with the location parameter ξ and the scale parameter α , was estimated by the method of L-moments (Hosking 1990, 1992). The T -year return values X_T were estimated from the quantile function of the GEV distribution as

$$X_T = \begin{cases} \hat{\xi} + \hat{\alpha} \{1 - [-\ln(1 - 1/T)]^{\hat{\kappa}}\} / \hat{\kappa}, \hat{\kappa} \neq 0 \\ \hat{\xi} - \hat{\alpha} \ln[-\ln(1 - 1/T)], \hat{\kappa} = 0 \end{cases} \quad (5)$$

The sampling uncertainty of the estimates was determined by the parametric bootstrap procedure in Kharin & Zwiers (2000). In this procedure 1000 samples of size 30 for observations (300 for synthetic data) were generated from the fitted GEV distribution. A return value was then estimated from each generated sample by fitting and inverting a GEV distribution as derived above. The 2.5 and 97.5 percentiles of the resulting collection of return value estimates were then used as lower and upper 95% confidence bounds for the true T -year return value. If a return value estimated from the synthetic data falls within the 95% confidence interval of the corresponding return value estimated from observations, the return value estimated from the synthetic data was not considered different from the observed one at the 5% significance level. Similarly, we also estimated the 95% confidence intervals for each of the percentiles from the 1st to the 99th in steps of 1%, to graphically show how the fitted GEV distribution of the yearly maxima from the synthetic data matches that from observations, in addition to the 10, 20 and 50 yr return values, which are the 90th, 95th and 98th percentiles of the fitted GEV distribution.

3. RESULTS AND DISCUSSION

3.1. Means of yearly maxima of extreme daily values

The capacity of the 2 weather generators in simulating extreme daily values varied considerably across seasons as seen from the number of stations rejecting the null hypothesis that the observed means are equal to the ones from the synthetic data (Table 1). As there were 9 stations in this study, a number > 5 implies a failure for a majority of stations. Both weather generators performed very well in reproducing the means of yearly maxima of extreme daily precipitation. No rejection was found for LARS-WG, and only 3 rejections out of 126 cases (9 stations, 12 mo plus annual and GS) were found for AAFC-WG. A much higher percentage (31%, i.e. 39 out of 126 cases) of rejection was observed for T_{\max} for LARS-WG. The failure occurred more often for winter months than other seasons. AAFC-WG had relatively fewer cases of rejection for T_{\max} than LARS-WG. The null hypothesis was rejected for only 12% (i.e. 15 out of 126) of cases. Notably, more cases were rejected in the test for T_{\min} for LARS-WG, compared with T_{\max} . There were 38% (48 out of 126) cases of rejection, and the failure was found more often in winter months and in the transition seasons than in summer months. In contrast, AAFC-

Table 1. Number of stations (9 in total) rejecting the null hypothesis that the observed means are equal to the ones from the synthetic data generated by weather generators in statistical tests at the 5% significance level, for monthly, annual and growing season (GS) maximal daily maximum temperature (T_{\max}), minimal daily minimum temperature (T_{\min}) and maximal daily precipitation (P)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual	GS
T_{\max}														
LARS-WG	6	5	3	1	1	2	2	0	3	1	2	7	3	3
AAFC-WG	0	2	2	0	0	0	2	1	0	0	1	3	2	2
T_{\min}														
LARS-WG	6	6	1	4	6	0	1	1	1	2	5	2	7	6
AAFC-WG	1	1	1	0	0	0	0	0	1	0	1	4	1	1
P														
LARS-WG	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AAFC-WG	0	0	0	0	0	0	0	1	0	0	0	0	2	0

Table 2. Number of stations (9 in total) rejecting the null hypothesis that the observed variances are equal to the ones from the synthetic data generated by weather generators in statistical tests at the 5% significance level, for monthly, annual and growing season (GS) maximal daily maximum temperature (T_{\max}), minimal daily minimum temperature (T_{\min}) and maximal daily precipitation (P)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual	GS
T_{\max}														
LARS-WG	3	2	3	0	2	0	3	1	1	2	0	3	3	2
AAFC-WG	0	1	1	0	0	1	2	1	1	0	1	3	2	2
T_{\min}														
LARS-WG	4	3	4	2	2	2	1	0	2	4	1	1	5	4
AAFC-WG	0	2	3	0	0	1	1	0	2	1	0	0	0	0
P														
LARS-WG	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AAFC-WG	0	0	0	0	0	0	1	2	0	0	0	0	0	1

WG had a better performance than LARS-WG for the reproduction of the means of yearly minimal T_{\min} ; the rejection rate was 9% of total cases (11 out of 126).

3.2. Variances of yearly maxima of extreme daily values

Although an overdispersion problem is common to stochastic weather generators, differences in variances between observed and weather-generator-simulated extreme values were fewer than differences in means (Table 2). This seems reasonable as a maximum is not an aggregated value such as monthly means and totals. Both weather generators reproduced the variances of yearly maxima of extreme daily values satisfactorily for precipitation. No rejection of the null hypothesis that the observed variances are equal to the ones from the synthetic data was found for LARS-WG and only 4 rejections out of 126 cases (3%) were seen for AAFC-WG. The rejection rate was 20% (25 out of 126) for T_{\max} for LARS-WG and 12% (15 out of 126) for AAFC-WG. The seasonal difference was not remarkable. There were also more cases of rejection (35 out of

126 or 28%) for T_{\min} for LARS-WG, compared with the test for T_{\max} . The rejection rate was 8% (10 out of 126) for T_{\min} for AAFC-WG. The rejection rates for AAFC-WG were similar for both the mean and the variance. Surprisingly, the rejection of the null hypothesis for the variances was caused more by overestimation of observed variances than by underestimation. All cases of rejection for precipitation by AAFC-WG resulted from overestimation. The cases of rejection caused by overestimation were more than twice the cases caused by underestimation for T_{\max} and T_{\min} for LARS-WG. However, the cases of rejection caused by overestimation were almost the same as those by underestimation for T_{\max} and T_{\min} for AAFC-WG.

3.3. Probability distributions of yearly maxima of extreme daily values

The rejection rate in the A-D test to the null hypothesis that yearly maxima derived from observations and synthetic data were drawn from identical populations was higher than those of the tests for means and

Table 3. Number of stations (9 in total) rejecting the null hypothesis that the observed and the synthetic yearly maxima generated by weather generators were drawn from identical populations in the Anderson-Darling test at the 5% significance level, for monthly, annual and growing season (GS) maximal daily maximum temperature (T_{\max}), minimal daily minimum temperature (T_{\min}) and maximal daily precipitation (P)

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Annual	GS
T_{\max}														
LARS-WG	6	5	6	1	3	2	3	2	4	2	5	7	5	5
AAFC-WG	1	1	3	0	1	2	2	2	0	0	2	4	3	3
T_{\min}														
LARS-WG	7	8	5	6	7	0	2	3	2	4	6	4	7	6
AAFC-WG	2	3	2	0	3	0	0	0	1	0	1	5	2	3
P														
LARS-WG	0	0	0	0	0	0	0	0	0	0	0	0	0	0
AAFC-WG	1	1	0	0	0	0	2	1	0	0	1	0	3	3

variances (Table 3). This is understandable, because a significant difference in means or variances, as well as higher order moments, could result in the mismatch of the probability distributions.

No rejection was found for P for LARS-WG, while a rejection rate of 10% (12 out of 126) was seen for AAFC-WG, where notably the null hypothesis in the A-D test was rejected for one-third of the stations for either annual or GS maximal daily P . Fig. 2 shows the 95% confidence intervals of the percentiles estimated from the GEV distributions fitted to the annual maximal daily precipitation amounts derived from observations and synthetic data. These plots were fairly consistent with the results from the A-D test. Goose Bay, Toronto and Truro were the 3 stations where the null hypothesis in the A-D test was rejected for AAFC-WG. The 95% confidence intervals for the GS maximal daily precipitation amounts are shown in Fig. 3. The 3 stations rejecting the null hypothesis in the A-D test for the GS maximal daily P for AAFC-WG were Goose Bay, Regina and Toronto, although the plots showed a larger deviation for Truro than for Regina.

The rejection rate was 44% (56 out of 126) for T_{\max} for LARS-WG, and it was higher in winter months than other seasons. The corresponding rejection rate was lower for AAFC-WG, constituting 19% (24 out of 126). The null hypothesis in the A-D test was rejected for more than half of the stations for the annual and the GS maximal daily T_{\max} for LARS-WG, while this ratio was one-third for AAFC-WG. The 95% confidence intervals of the percentiles estimated from the fitted GEV distributions were plotted in Fig. 4 for the annual maximal daily T_{\max} . The results from these were often consistent with the results from the A-D test, although some inconsistent cases could be observed. For example, the right tail simulated by LARS-WG significantly deviated from the observed one at Beaverlodge, but the null hypothesis in the A-D test was not rejected for the case. Since annual maximal daily maximum tem-

peratures occurred most often during the warm season, the GS maximal daily maximum temperatures were almost always the same as annual maxima. Therefore, it is not necessary to show the plots for the GS maxima of daily T_{\max} .

The rejection rate for minimal daily T_{\min} was higher than that for maximal daily T_{\max} for LARS-WG, ranging as high as 53% (67 out of 126). The performance of the weather generators was better in summer months than in other seasons. The performance of AAFC-WG was considerably better, with a rejection rate of 17% (22 out of 126). The rejection rate was two-thirds in the A-D test for LARS-WG, while the rejection rate was about one-third for AAFC-WG, for annual and GS minimal daily T_{\min} . The 95% confidence intervals of the percentiles estimated from the fitted GEV distributions are shown in Fig. 5 for annual minimal daily T_{\min} and in Fig. 6 for GS minimal daily T_{\min} . The results shown from the plots were often consistent with the results from the A-D test. However, mismatches of the results between the plots and the A-D test were also seen for a few cases. For example, the fitted right tail from the synthetic annual maximal daily $-T_{\min}$ generated by either LARS-WG or AAFC-WG significantly deviated from the fitted right tail of the observed annual maximal daily $-T_{\min}$ at Beaverlodge, but the null hypothesis in the A-D test was not rejected for either synthetic series itself.

3.4. Return values of 10, 20 and 50 yr

As was observed in the plots of the 95% confidence intervals of the percentiles, there were cases rejecting the null hypothesis in the A-D test, although the return values for some return periods might not be significantly different, or vice versa, in comparisons between the observed maxima and those derived from synthetic data generated by the

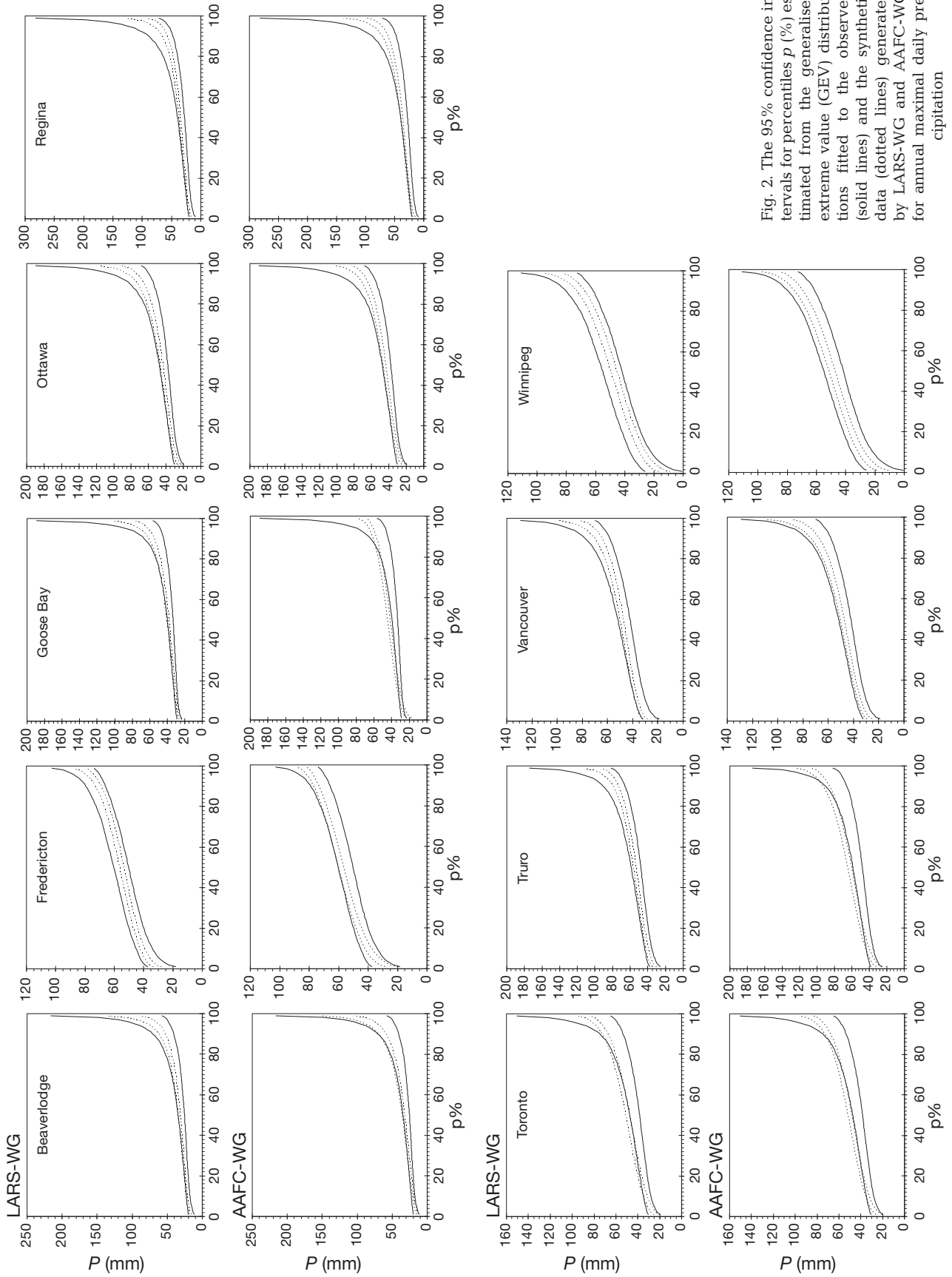


Fig. 2. The 95% confidence intervals for percentiles p (%) estimated from the generalised extreme value (GEV) distributions fitted to the observed data (dotted lines) and the synthetic data (solid lines) generated by LARS-WG and AAFC-WG for annual maximal daily precipitation

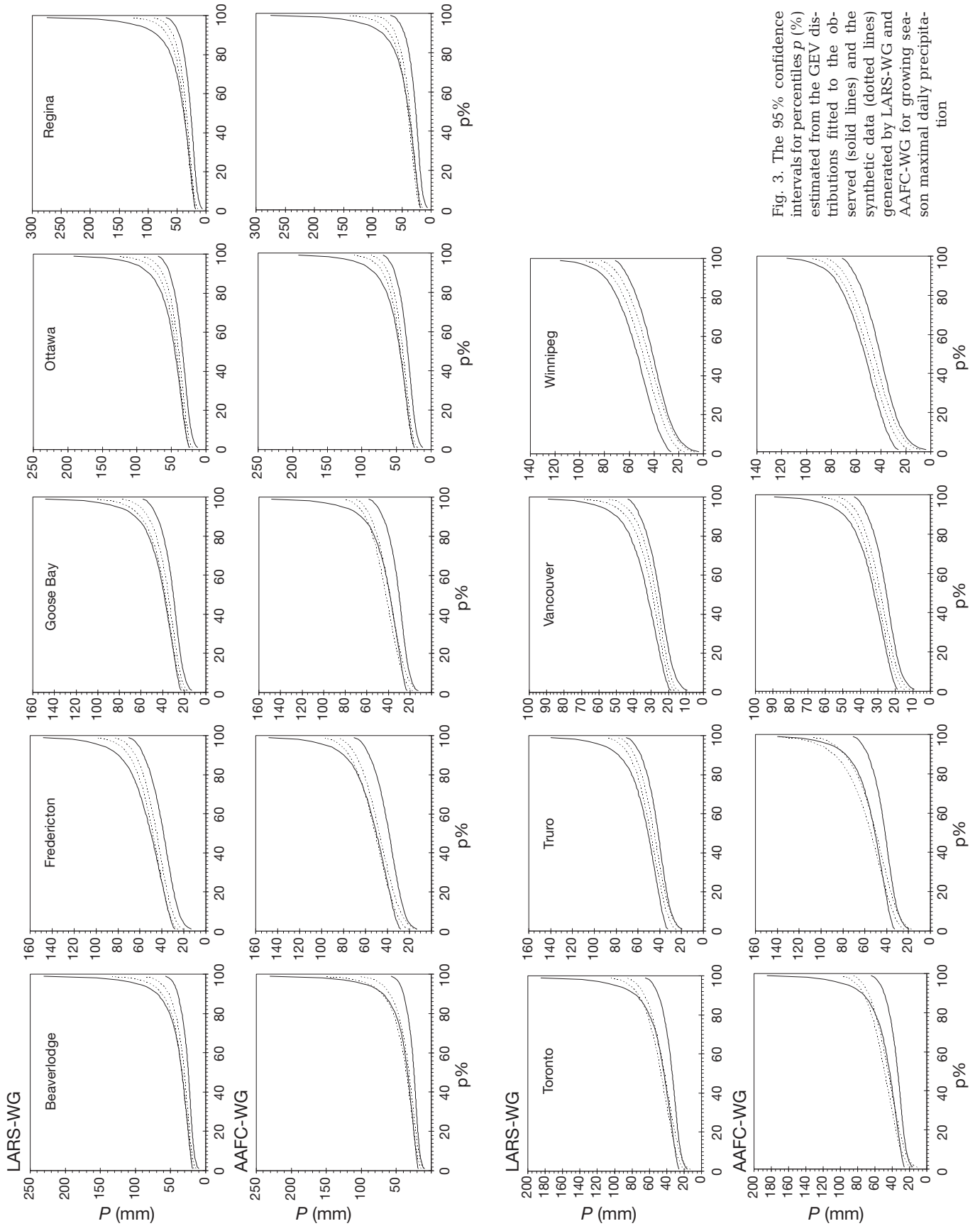


Fig. 3. The 95% confidence intervals for percentiles p (%) estimated from the GEV distributions fitted to the observed (solid lines) and the synthetic data (dotted lines) generated by LARS-WG and AAFC-WG for growing season maximal daily precipitation

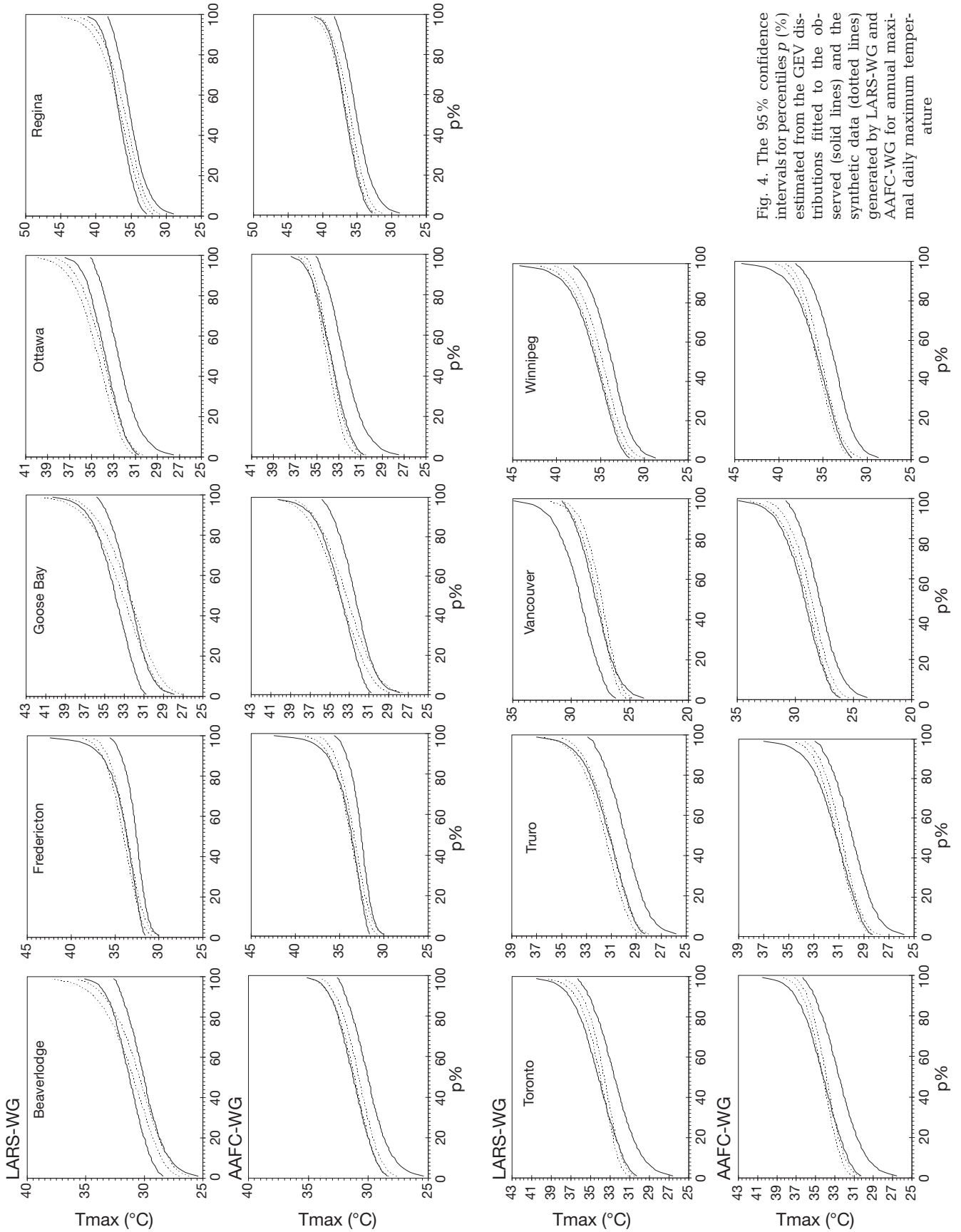


Fig. 4. The 95% confidence intervals for percentiles p (%) estimated from the GEV distributions fitted to the observed data (dotted lines) and the synthetic data (solid lines) generated by LARS-WG and AAFC-WG for annual maximum daily maximum temperature

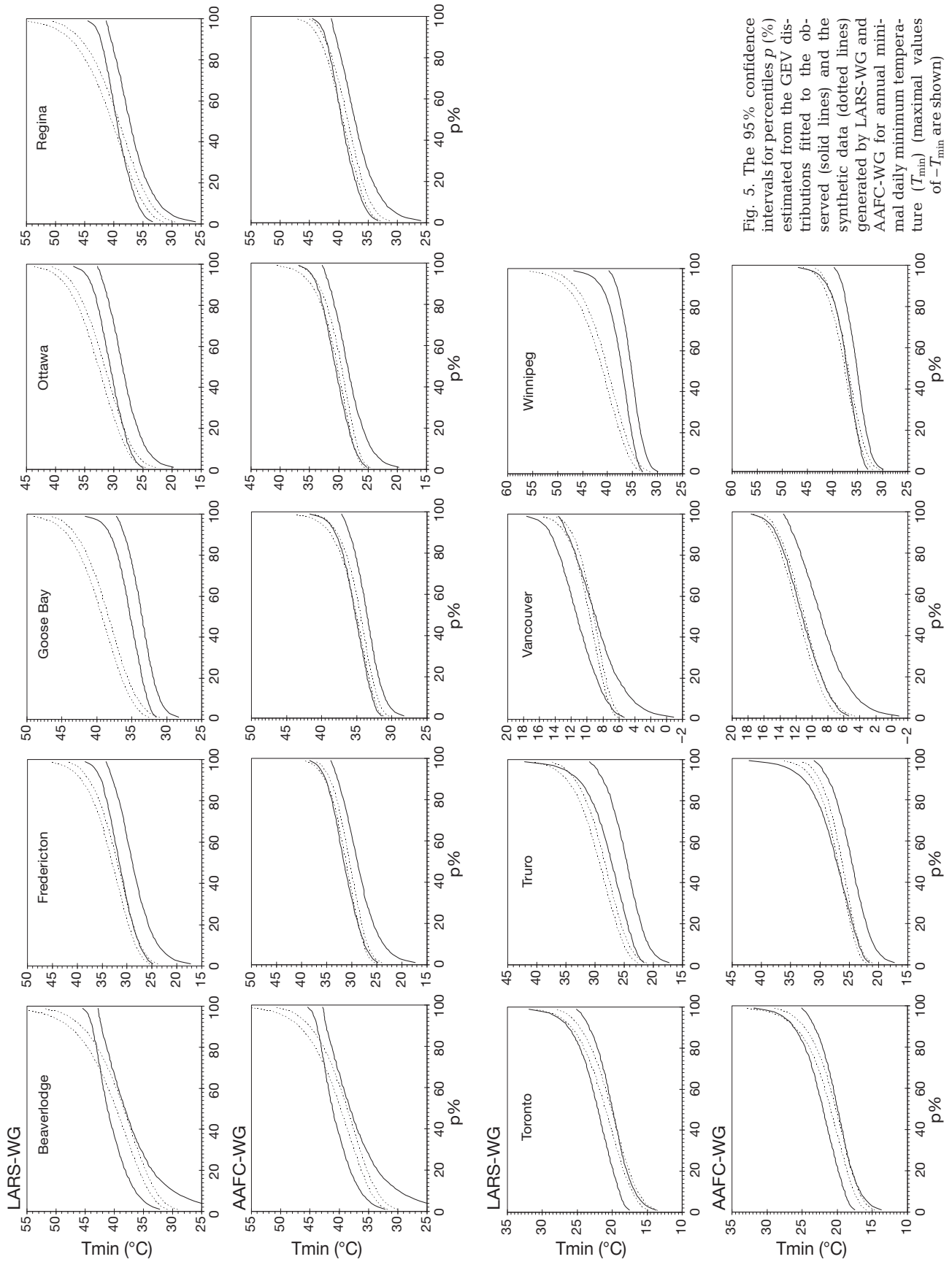


Fig. 5. The 95% confidence intervals for percentiles p (%) estimated from the GEV distributions fitted to the observed (solid lines) and the synthetic data (dotted lines) generated by LARS-WG and AAFC-WG for annual minimum daily minimum temperature (T_{\min}) (maximal values of $-T_{\min}$ are shown)

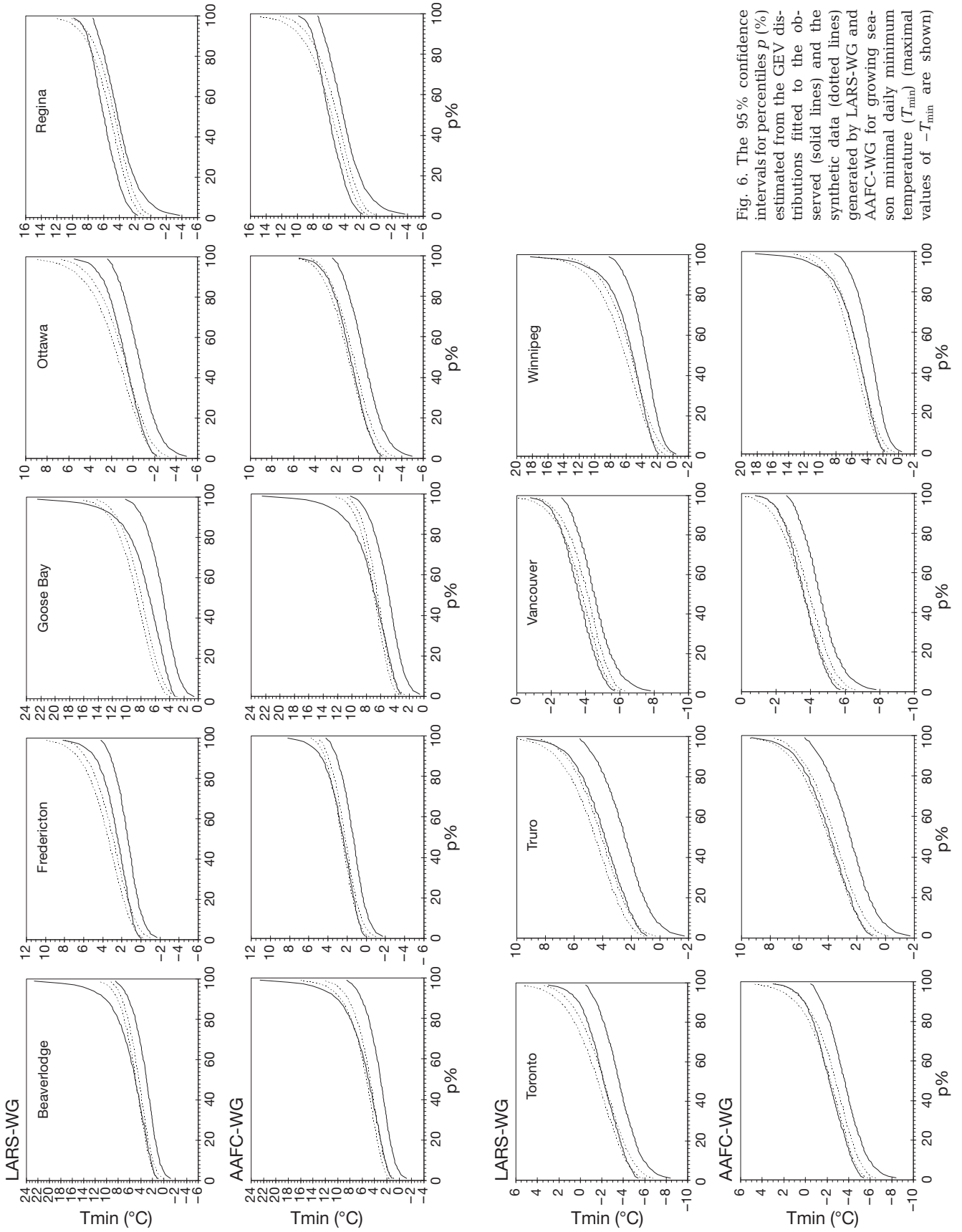


Fig. 6. The 95% confidence intervals for percentiles p (%) estimated from the GEV distributions fitted to the observed (solid lines) and the synthetic data (dotted lines) generated by LARS-WG and AAFC-WG for growing season minimal daily minimum temperature (T_{\min}) (maximal values of $-T_{\min}$ are shown)

Table 4. Number of stations (9 in total) showing significant differences in the return values of 10, 20 and 50 yr estimated from the observed and the synthetic yearly maxima generated by weather generators at the 5% significance level, for monthly, annual and growing season (GS) maximal daily maximum temperature (T_{\max}), minimal daily minimum temperature (T_{\min}) and maximal daily precipitation (P)

		T_{\max}			T_{\min}			P		
		10 yr	20 yr	50 yr	10 yr	20 yr	50 yr	10 yr	20 yr	50 yr
Jan	LARS-WG	7	8	7	7	7	6	0	0	0
	AAFC-WG	1	3	3	0	1	1	0	1	1
Feb	LARS-WG	6	6	5	5	5	5	0	0	0
	AAFC-WG	1	2	2	0	0	0	0	1	1
Mar	LARS-WG	4	3	2	4	3	2	0	0	0
	AAFC-WG	2	2	2	2	2	2	0	0	0
Apr	LARS-WG	1	1	1	0	1	1	0	0	0
	AAFC-WG	0	0	0	0	0	0	0	0	0
May	LARS-WG	3	2	2	7	7	5	0	0	0
	AAFC-WG	2	2	2	3	3	3	0	1	1
Jun	LARS-WG	1	1	0	3	3	3	0	0	0
	AAFC-WG	0	1	1	2	2	2	1	1	1
Jul	LARS-WG	5	4	3	5	4	3	0	0	0
	AAFC-WG	1	1	1	2	2	3	1	1	1
Aug	LARS-WG	2	2	3	2	2	2	0	0	0
	AAFC-WG	0	0	1	0	1	1	2	2	2
Sep	LARS-WG	4	3	2	3	3	4	0	0	0
	AAFC-WG	0	0	0	3	3	3	1	1	1
Oct	LARS-WG	1	1	1	4	4	4	0	0	0
	AAFC-WG	0	1	1	0	0	0	1	1	1
Nov	LARS-WG	1	1	3	2	2	2	0	0	0
	AAFC-WG	0	0	1	2	2	2	0	1	1
Dec	LARS-WG	8	7	7	2	2	1	0	0	0
	AAFC-WG	7	7	5	1	2	2	1	1	1
Annual	LARS-WG	4	5	4	8	7	6	0	0	0
	AAFC-WG	1	2	1	5	4	4	0	0	0
GS	LARS-WG	4	5	3	6	7	6	0	0	0
	AAFC-WG	1	2	1	4	4	3	1	1	0
Total	LARS-WG	51	49	43	58	57	50	0	0	0
	AAFC-WG	16	23	21	24	26	26	8	12	11

stochastic weather generators. The return values of 10, 20 and 50 yr return periods were examined in detail. The number of stations showing significant differences between the return values estimated from the observed and the synthetic yearly maxima generated by the weather generators at the 5% significance level are listed in Table 4, for monthly, annual and GS maximal daily T_{\max} , minimal daily T_{\min} and maximal daily P .

No significant difference was found for LARS-WG, while several cases (8 to 12 out of 126 or 6 to 10%) were found for AAFC-WG regarding the return values of 10, 20 and 50 yr periods for yearly maxima of daily P . The simulation of return values for yearly maxima of extreme daily temperatures was not as successful as that for extreme daily precipitation, especially

for LARS-WG. Approximately 40% of the simulated return values of 10, 20 and 50 yr periods by LARS-WG were significantly different from the values estimated from observations for maximal daily T_{\max} . The simulation was worse in winter months, in general. A higher rate of failure was found for minimal daily minimum temperature, when >40% of the return values simulated by LARS-WG were significantly different from those estimated from observations. The worst cases were for January and February, resulting in failures for annual minimal values. AAFC-WG performed slightly better than LARS-WG for reproducing return values of extreme daily temperatures (Table 4). The rate of failure for AAFC-WG was <20% for extremely high daily T_{\max} and about 20% for extremely low daily T_{\min} . It also appeared that, for both weather generators, simu-

Table 5. Return values of 10, 20 and 50 yr return periods estimated from observations and synthetic data generated by the weather generators LARS-WG and AAFC-WG for annual maximum daily T_{\max} , T_{\min} and P . The values with asterisks estimated from the synthetic data are significantly different from the values estimated from observations at the 0.05 significance level

		T_{\max} (°C)			T_{\min} (°C)			P (mm)		
		10 yr	20 yr	50 yr	10 yr	20 yr	50 yr	10 yr	20 yr	50 yr
Beaverlodge	Observed	32.6	33.0	33.5	-43.1	-43.5	-43.8	53.9	68.2	92.8
	LARS-WG	33.8*	34.8*	36.0*	-46.6*	-49.2*	-52.4*	56.7	69.2	88.8
	AAFC-WG	32.9	33.5	34.1	-46.0*	-48.3*	-51.0*	63.6	78.8	102.2
Fredericton	Observed	35.4	36.2	37.4	-34.1	-34.9	-35.7	72.6	77.6	83.0
	LARS-WG	36.1	36.8	37.6	-38.2*	-40.1*	-42.3*	70.6	75.3	80.4
	AAFC-WG	35.9	36.7	37.6	-34.8	-36.1	-37.4	75.3	79.2	83.0
Goose Bay	Observed	35.7	36.4	37.3	-37.1	-37.9	-38.7	53.8	63.9	80.9
	LARS-WG	36.6	37.8*	39.2*	-43.6*	-45.0*	-46.6*	57.2	65.7	78.0
	AAFC-WG	36.9*	37.8*	38.8	-38.3*	-39.6*	-41.2*	59.5	64.3	69.8
Ottawa	Observed	35.1	35.5	35.9	-32.8	-33.6	-34.3	66.3	77.7	94.8
	LARS-WG	36.7*	37.5*	38.4*	-37.6*	-39.2*	-40.9*	68.1	77.9	91.3
	AAFC-WG	35.5	35.8	36.2	-34.1*	-35.6*	-37.4*	66.5	74.6	85.1
Regina	Observed	38.2	38.8	39.3	-41.5	-42.0	-42.5	63.9	81.6	111.3
	LARS-WG	39.9*	41.1*	42.6*	-46.5*	-48.7*	-51.1*	61.0	73.1	91.3
	AAFC-WG	39.0	39.8*	40.7*	-42.9*	-44.0*	-45.2*	66.1	79.3	98.7
Toronto	Observed	36.2	36.9	37.6	-24.8	-25.9	-27.2	63.6	72.9	86.2
	LARS-WG	36.2	36.9	37.9	-25.2	-26.6	-28.3	69.3	75.8	83.4
	AAFC-WG	36.1	36.7	37.3	-25.3	-27.0	-29.2	70.1	76.6	84.2
Truro	Observed	32.8	33.5	34.2	-30.8	-32.4	-34.4	77.3	87.6	102.1
	LARS-WG	33.6	34.4	36.3	-34.0*	-35.7*	-37.7	75.7	83.6	93.7
	AAFC-WG	32.8	33.3	33.8	-30.6	-32.0	-33.6	89.2	97.9	108.1
Vancouver	Observed	30.8	31.5	32.2	-13.6	-14.3	-15.0	67.5	75.7	86.3
	LARS-WG	29.5*	30.1*	30.8	-12.2*	-13.1	-14.1	66.7	73.8	82.9
	AAFC-WG	31.1	31.8	32.7	-14.6	-15.4	-16.2	70.7	79.3	90.4
Winnipeg	Observed	37.8	38.8	40.0	-39.5	-40.6	-41.9	70.8	77.4	84.6
	LARS-WG	38.2	39.2	40.4	-46.9*	-49.1*	-51.7*	69.7	76.4	83.7
	AAFC-WG	37.9	38.6	39.3	-41.3*	-42.5	-43.8	72.4	79.0	86.2

lations for extremely low daily T_{\min} were poorer than for extremely high daily T_{\max} .

It is interesting that the significant differences between return values estimated from observations and from synthetic data were often caused by more intense extreme values in the weather generator simulations. Return values of 10, 20 and 50 yr for annual extreme daily values are given in Table 5. They were estimated from observations and synthetic data generated by LARS-WG and AAFC-WG, respectively. There were 51 return values estimated from synthetic data that significantly differed with those estimated from observations. Only 3 out of the 51 values were smaller than those from observations, i.e. the 10 and 20 yr return values of annual maximal daily T_{\max} and the 10 yr return value of annual minimal daily T_{\min} by LARS-WG at Vancouver. Similar cases were observed for monthly and GS maximal values. This is consistent with the results pertaining to variances, where significant difference between the variances of observed and simulated yearly maxima was often caused by overestimation of the observed variance.

4. CONCLUSIONS

The stochastic weather generators LARS-WG and AAFC-WG were evaluated for the capability of reproducing extreme daily values of precipitation and maximum and minimum temperatures in diverse Canadian climates. Results from the statistical analysis showed that both weather generators, especially LARS-WG, were capable of reproducing yearly maxima of monthly, annual and GS maximal daily precipitation. However, the performances of the weather generators for reproducing extreme daily temperatures were not as good as for daily precipitation, especially in the case of LARS-WG. The deficiency in reproducing extremely low temperatures was also more noticeable than extremely high temperatures for LARS-WG. The mismatches of return values were often caused by a more extreme value estimated from synthetic data than the one derived from observations.

The better performance of AAFC-WG on reproducing extreme daily temperatures than LARS-WG might be associated with the assumption of a normal

distribution for daily temperatures by LARS-WG. Therefore, using empirical distributions from observed data or a better fitted conventional distribution may be helpful for improving the performance of weather generators in reproducing extreme daily temperature values. The performance of AAFC-WG in reproducing extremely high and low temperatures was fairly similar, but the differences for LARS-WG were notable. As LARS-WG uses a pre-set cross-correlation between daily maximum and minimum temperatures, the results may imply that the difference was related to an unrealistic cross-correlation between daily maximum and minimum temperatures. Using cross-correlations estimated from observations might also be a measure for the improvement of the weather generator in reproducing extreme daily temperatures.

It is interesting to note why AAFC-WG had several cases of failure in reproducing extreme daily precipitation when empirical distributions were used, as they were in LARS-WG. LARS-WG uses only 10 intervals while AAFC-WG uses 100 intervals in histograms for the construction of empirical distributions. Semenov (2008) indicated the drawback of the flexibility of an empirical distribution that, in some cases, a semi-empirical distribution can 'overfit' the observed data. Overfitting is a typical problem encountered where a statistical distribution or model has many parameters and high complexity in comparison of the data available. This seems very likely for precipitation in AAFC-WG, as the number of days with precipitation is very limited in dry seasons at some stations, even if data are combined bimonthly. The empirical distributions of temperatures could also be overfitted, since they are conditional on the occurrence of precipitation. Outliers can also be an issue to address in the construction of empirical distributions as an outlier in a data set may strongly affect the distribution parameters (Semenov 2008). It is worthwhile to investigate the effects of the method used in the construction of empirical distributions on the performance of weather generators, as this has an impact on how to reasonably perturb the parameters of empirical distributions for climate change studies. The results from this study indicated that the failure of reproducing return values was often associated with more intensive extreme values in synthetic data than in observations. This phenomenon might be associated with the 'overfitting' problem and the issue of 'outliers' in the data set. Further studies may be required to address the 'overfitting' problem; however, it might be a good aspect of a weather generator in view of a possible need to extrapolate beyond the range of the data, since we only have 30 yr of data, if more intensive extremes derived from the synthetic data are associated with 'outliers'.

Despite the noted limitations, both weather generators could reproduce extreme daily precipitation very well, and the capability of AAFC in reproducing extreme daily temperatures was reasonably satisfactory, though some improvement may be necessary for both weather generators, especially LARS-WG, regarding extreme daily temperatures. The 95% confidence intervals of return values estimated from synthetic data were always smaller than the ones estimated from observations, as expected, indicating a more accurate estimation from synthetic data than from the observed data, as the observed data were available often for short periods. This fulfils the original goal of the development of stochastic weather generators in view of generating long synthetic data for the estimation of extreme values.

Acknowledgements. The authors are indebted to 2 internal reviewers Reinder de Jong and Brian Grant, as well as Elizabeth Pattey of the ECORC Publication Committee, for their comments, which were helpful in the improvement of the manuscript. We acknowledge the anonymous reviewers and the editor, Dr. Bryson Bates, for their critique and constructive suggestions on earlier versions of this paper. Dr. Xuebin Zhang is also acknowledged for a useful discussion. This is ECORC Contribution No. 06-747.

LITERATURE CITED

- Anderson TW, Darling DA (1952) Asymptotic theory of certain goodness-of-fit criteria based on stochastic processes. *Ann Math Stat* 23:193–212
- Coles S (2001) An introduction to statistical modeling of extreme values. Springer, London
- Dubrovsky M (1997) Creating daily weather series with use of the weather generator. *Environmetrics* 8:409–424
- Dubrovsky M, Buchtele J, Zalud 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
- Gumbel EJ (1958) Statistics of extremes. Columbia University Press, New York
- Hayhoe HN (2000) Improvements of stochastic weather data generators for diverse climates. *Clim Res* 14:75–87
- Hosking JRM (1990) L-moments: analysis and estimation of distributions using linear combinations of order statistics. *J R Stat Soc [Ser B]* 52:105–124
- Hosking JRM (1992) Moments or L-moments? An example comparing the two measures of distributional shape. *Am Stat* 46:186–189
- Hutchinson MF (1986) Methods of generation of weather sequences. In: Bunting AH (ed) *Agricultural environments*. CAB International, Wallingford, p 149–157
- Johnson M (1987) *Multivariate statistical simulation*. John Wiley & Sons, New York
- Katz RW (1996) Use of conditional stochastic models to generate climate change scenarios. *Clim Change* 32:237–255
- Katz RW, Brown BG (1992) Extreme events in a changing climate: variability is more important than averages. *Clim Change* 21:289–302
- Katz RW, Parlange MB (1998) Overdispersion phenomenon in stochastic modeling of precipitation. *J Clim* 11:591–601

- Kharin VV, Zwiers FW (2000) Changes in the extremes in an ensemble of transient climate simulations with a coupled atmospheric–ocean GCM. *J Clim* 13:3760–3788
- Kysely J, Dubrovsky 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
- Leadbetter MR, Lindgren G, Rootzen H (1983) *Extremes and related properties of random sequences and processes*. Springer-Verlag, New York
- Mearns LO, Katz RW, Schneider SH (1984) Extreme high-temperature events: changes in their probabilities with changes in mean temperature. *J Clim Appl Meteorol* 23:1601–1613
- 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
- Press WH, Teukolsky SA, Vetterling WT, Flannery BP (1992) *Numerical recipes in FORTRAN, the art of scientific computing*, 2nd edn. Cambridge University Press, Cambridge
- Qian B, Gameda S, Hayhoe H, De Jong R, Bootsma A (2004) Comparison of LARS-WG and AAFC-WG stochastic weather generators for diverse Canadian climates. *Clim Res* 26:175–191
- 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
- Racsco P, Szeidl L, Semenov M (1991) A serial approach to local stochastic weather models. *Ecol Model* 57:27–41
- Richardson CW (1981) Stochastic simulation of daily precipitation, temperature, and solar radiation. *Water Resour Res* 17:182–190
- Semenov MA (2008) Simulation of extreme weather events by a stochastic weather generator. *Clim Res* 35:203–212
- Semenov MA, Barrow EM (1997) Use of a stochastic weather generator in the development of climate change scenarios. *Clim Change* 35:397–414
- Semenov MA, Barrow EM (2002) LARS-WG: a stochastic weather generator for use in climate impact studies (Version 3.0). User manual. Available at www.rothamsted.ac.uk/mas-models/download/LARS-WG-Manual.pdf
- 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
- Wilks DS (1992) Adapting stochastic weather generation algorithms for climate change studies. *Clim Change* 22:67–84
- Wilks DS (1995) *Statistical methods in the atmospheric sciences*. Academic Press, San Diego, CA
- Wilks DS (1999) Interannual variability and extreme-value characteristics of several stochastic daily precipitation models. *Agric Meteorol* 93:153–169
- Wilks DS, Wilby RL (1999) The weather generator game: a review of stochastic weather models. *Prog Phys Geogr* 23:329–358
- Zwiers FW, Kharin V (1998) Changes in the extremes of the climate simulated by CCC GCM2 under CO₂ doubling. *J Clim* 11:2200–2222

*Editorial responsibility: Bryson Bates,
Wembley, WA, Australia*

*Submitted: December 5, 2006; Accepted: June 16, 2008
Proofs received from author(s): July 25, 2008*