Minimum data requirements for parameter estimation of stochastic weather generators

Afshin Soltani[1,*], Gerrit Hoogenboom[2]

[1]Department of Agronomy and Plant Breeding, Gorgan University of Agricultural Sciences, Gorgan, Iran
[2]Department of Biological and Agricultural Engineering, The University of Georgia, Griffin, Georgia 30223, USA

ABSTRACT: Long-term daily weather data are commonly required for the application of simulation models in agricultural systems. The objective of this study was to evaluate the impact of the length of input weather data on the quality of generated weather data and to determine the minimum data required for an accurate estimation of the climate coefficients for a weather generator. Five locations in Iran with contrasting climates were selected for this study. The WGEN and SIMMETEO weather generators were parameterized using 5, 10, 15, 20 and 30 yr of weather data as the base period (BP), and then 30 yr daily weather data sets were generated corresponding to each BP. While WGEN requires daily weather data for parameter estimation, SIMMETEO uses monthly summaries. The quality of the generated weather was measured as percentage significant differences between observed and generated data. Statistical tests were conducted, including t- and F-tests, to compare differences between generated weather data versus 30 yr historical weather data. The results showed that the number of years required to obtain generated data that were similar to observed data was a function of the weather generator and the weather variable. WGEN was more sensitive to the amount of input data than SIMMETEO. In 8 of the 14 cases WGEN required a specific minimum amount of data to be efficient in weather generation versus 6 of the 14 cases for SIMMETEO. There was a correlation between generator performance and its minimum data requirement. For the cases where the generators showed a good performance, i.e. low-percentage significant differences between observed and generated data, a specific minimum number of years was required. For generating precipitation, a smaller number of years was required than for generating temperature. Based on the results from this study, it can be concluded that at least 15 yr of input data are needed for WGEN and SIMMETEO to obtain generated sequences with statistical characteristics similar to the observed data. However, a greater length of input data may lead to an improved set of generated data as shown through more stable generated data.

KEY WORDS: Weather generation · Weather data · Climate · Simulation models · Decision support systems

1. INTRODUCTION

Weather is a key determinant in agricultural production, especially rainfed cropping systems commonly found in arid and tropical regions. Long-term daily weather data are normally required for the application of simulation models in agricultural or ecological systems. These include maximum and minimum air temperatures, total solar radiation, and total precipitation (Hoogenboom 2000). Some models also require rela-

tive humidity and wind speed as input. The accuracy of the data simulated by models is dependent upon the quality of the weather data that are used as input (Meinke et al. 1995). High-quality weather data are therefore needed to obtain reliable simulations. But in many cases the availability of these weather data might be limited. Common problems include (Hoogenboom 2000):

• Although the period of record is not too short, e.g. 10 to 20 yr, to conduct a modeling application, additional

years (>20 yr) are required to extend the data, because the probability distribution of simulated outcomes is important for understanding simulation results or decision making (Thornton & Hoogenboom 1994).

- Periods of daily records are missing within long-term historical data sets.
- Only monthly averages and totals are available.
- Only a few variables are recorded, such as precipitation and maximum and minimum temperatures, or in the worst case, only daily precipitation.

These problems have been recognized by many scientists, who have developed weather generators to address some of these problems.

The most commonly used weather generators are WGEN (Richardson & Wright 1984), SIMMETEO (Geng et al. 1986, 1988), LARS-WG (Semenov et al. 1998), MARKSIM (Jones & Thornton 2000), WM2 (Hansen & Mavromatis 2001), TAMSIM (McCaskill 1990) and others (e.g. Larson & Pense 1982, Guenni et al. 1991, Hayhoe 2000). WGEN and SIMMETEO are well known and widely used stochastic weather generators. These generators have been incorporated into the WeatherMan (Weather data Manager; Pickering et al. 1994), an application program of the Decision Support System for Agrotechnology Transfer (DSSAT) software system (Tsuji et al. 1994, Hoogenboom et al. 1999). DSSAT is a collection of crop models and computer programs integrated into a single software package to facilitate the application of crop simulation models in research and decision making. The main advantage of SIMMETEO in comparison to WGEN is that its input parameters can be estimated from monthly summaries instead of the daily data required for estimating the input parameters for WGEN.

While these weather generators have been evaluated in a number of studies (e.g. Richardson 1981, Richardson & Wright 1984, Meinke et al. 1995, Semenov et al. 1998, Mavromatis & Hansen 2001, Puche & Silva 2001, Warnock & Puche 2001, Soltani & Hoogenboom 2003), the outputs of the generators have not been evaluated when their required parameters are estimated from input data that are based on different durations. These evaluations will allow for the determination of the minimum number of years required for parameter estimation of the generators. In one study, Richardson (2000) evaluated the stability of estimates of weather generation parameters relative to the length of the observed input data from which the parameters were estimated. He found that, in general, at least 10 yr of weather data are needed for estimation of temperature parameters and 20 or more years of weather data are needed for estimation of the precipitation parameters. The objective of this study was to determine the minimum number of years required for obtaining generated weather series that have similar statistical characteristics similar to the observed data. A second objective was to compare the performance of WGEN and SIMMETEO with respect to the length of the input data used for determining the climate-input parameters.

2. METHODS

2.1. Weather data generators

The weather generators used in this study were adaptations of WGEN (Richardson & Wright 1984) and SIMMETEO (Geng et al. 1986, 1988) as implemented in WeatherMan (Pickering et al. 1994). The WGEN model provides daily values of precipitation, maximum and minimum temperatures and solar radiation for an *n* yr period at a given location. The precipitation is generated independently of other weather variables for a given day. The precipitation component of WGEN is a Markov-chain‐gamma-distribution model. Solar radiation and maximum temperature are sampled from normal distributions parameterized separately for dry and wet days. Minimum temperature is sampled from a normal distribution independently of precipitation occurrence. These 3 variables are modeled according to a first-order trivariate autoregression process. The major adaptations of WGEN relate to how Weather-Man handles input parameters. All parameters are estimated on a calendar-month basis. Daily values are computed internally, using linear interpolation. More details can be found in Richardson (1981), Richardson & Wright (1984), Semenov et al. (1998), Mavromatis & Hansen (2001) and Soltani & Hoogenboom (2003).

The SIMMETEO generator of WeatherMan is embodied in the WGEN generator, but it uses different input data compared to WGEN. The modified SIMMETEO model uses monthly climatic means of the number of wet days, precipitation, solar radiation, and maximum temperature and minimum temperature as well as regression equations to compute conditional means, standard deviations and precipitation parameters (Geng et al. 1986, 1988, Pickering et al. 1994). Details on the parameter estimation methods used in SIMMETEO can be found in Soltani & Hoogenboom (2003).

2.2. Historical weather data sets

For this study 5 sites in Iran were selected that have weather stations with long-term reliable daily data. They are Gorgan, Kermanshah, Isfahan, Ahwaz and Shiraz. These stations cover a wide geographical area and represent several climatic zones in Iran. Climatic

characteristics of the selected weather stations can be found in Soltani & Hoogenboom (2003). For each location, precipitation and maximum and minimum temperatures were available for 30 yr (1966–1995). Solar radiation data were calculated from sunshine hours and extraterrestrial radiation. For days without sunshine-hours data, solar radiation was calculated based on locally determined relationships between precipitation, temperature and solar radiation. With estimated solar radiation data, a complete 30 yr data set of daily precipitation, maximum and minimum temperatures and solar radiation data was available for each location.

2.3. Analysis

For each location, the parameters used to generate the daily weather data were obtained from 5, 10, 15, 20 and 30 yr of weather data as BPs corresponding to 16.7, 33.3, 50, 66.7 and 100% of the historical weather data set. For each base period (BP), multiple sampling was used and the input data were selected from the early, middle and late sections of the historical weather data sets (Table 1). They overlapped for the BPs of 15 and 20 yr.

For each BP as shown in Table 1, a 30 yr series of daily weather data was generated with WGEN and SIMMETEO. For the 30 yr BP, three 30 yr weather data sets were generated using different 'seed numbers' to initialize each generation. For each location there were 30 generated weather data sets that consisted of 30 yr of daily weather data (15 generated with WGEN and 15 generated with SIMMETEO). Monthly means and variances for the number of wet days (WD), precipitation (P), solar radiation (R), maximum temperature (T_{max}), minimum temperature (T_{min}), number of days with a maximum temperature greater than 35°C ($T_{max} > 35$) and number of days with a minimum temperature less than 0°C ($T_{min} < 0$) were computed for the 2 groups of 15 data sets of generated weather data and also for the observed weather data. t- and F-tests were used to compare differences between observed and generated series with respect to mean and variance, respectively. When the distribution of the weather variables was not normal, suitable transformations were used before the t- and F-tests were applied. Normality of the data was tested using PROC UNIVARIATE of the Statistical Analysis System (SAS; SAS Institute 1989). t- and F-tests were also conducted with SAS, using a 0.01 level.

For each test, the generated data were compared with long-term (30 yr) historical weather data, not with

Table 1. Periods of input data used as base periods for parametrizing the WGEN and SIMMETEO weather generators. Historical data: 1966–1995

Base period length (yr)				
5	10	15	20	30
1971–1975	1966–1975	1966–1980	1966–1985	1966–1995
1981–1985	1976–1985	1973–1987	1971–1990	1966–1995
1991–1995	1986–1995	1981–1995	1976–1995	1966–1995

their corresponding input weather, except for the 30 yr period. There were approximately 1260 t-tests and 1260 F-tests for each location for a total of 12600 statistical analyses. For each location, the total number of tests (N), the number of tests where a significant difference was detected (N^*) and the percentage tests in the rejection region (= $N^*/N \times 100$) were separately recorded for the t- and F-tests. Hereafter, we use t^* to represent the percentage t-test in the rejection region, and F^* to represent the percentage F-tests in the rejection region.

A 2-piece segmented, non-linear regression model was used to describe the relationship between t^* and F^* (Y, %) with the length of input weather data (X, yr). The segmented model consisted of 2 intersecting lines, a sloping line for a linear decrease in Y and a horizontal line, which represented minimum obtainable Y (t^* or F^*). Mathematically, the segmented model can be expressed as

$$Y = a + bX \qquad \text{if } X < X_0$$
$$Y = a + bX_0 \qquad \text{if } X > X_0 \tag{1}$$

where a is the intercept with the vertical axis ($X = 0$, %), b is the rate (% yr^{-1}) of linear decrease in Y (t^* or F^*) with increasing the amount of input data, and X_0 is the intersecting point of the model. X_0 can be interpreted as the minimum number of years of input weather data required to obtain minimum differences (Y_{min}) between statistical characteristics of observed and generated data ($Y_{min} = a + bX_0$). This regression model was fitted to the data using the PROC NLIN (method DUD) of SAS. In some cases it was impossible to fit model (Eq. 1) to the data (e.g. for the number of cold days; see Fig. 8); hence, a simple, linear regression model was used.

3. RESULTS

3.1. General

Our earlier study showed that the performance of WGEN and SIMMETEO is highly independent of the location of the selected sites, which represent different climatic zones (Soltani & Hoogenboom 2003). In this

study similar results were found for the response of the percentage significant differences in the mean and variance comparisons ($t^\star$ and $F^\star$) to the length of input weather data (data not shown). Therefore, the results of the analyses for this study were combined across the 5 selected locations.

The percentage significant differences ($t^\star$ and $F^\star$) between observed and generated weather data when the length of input data is equal to the entire length of observed data (30 yr here; D_{30} in Table 2) indicates the adequacy and validity of the generators. Details on the performance of the generators for the selected locations of this study can be found in Soltani & Hoogenboom (2003).

The response of absolute difference between observed and generated weather variables (expressed in their physical unit) to the length of input data was similar to those of $t^\star$ (Table 2). We prefer to use $t^\star$, because it is impossible to apply the segmented model (Eq. 1) to the differences, due to their fan-like shape. Therefore, some examples were selected from the data of Kermanshah and are shown in Fig. 1. For the number of wet days generated with SIMMETEO, the range of differences decreased to less than 4 d when the period length was increased to 15 yr or more. Similarly, for maximum temperature generated

with WGEN, the range of differences stabilized at around 2.5°C with the period length greater than 12 yr. However, for solar radiation generated with WGEN and the number of hot days generated with SIMMETEO, the range of differences was approximately constant for all selected BPs.

3.2. Precipitation

The mean monthly number of wet days for the data generated by WGEN and SIMMETEO were similar to those of the observed data. No significant differences were found between the observed and generated number of wet days for both generators (D_{30} in Table 2). However, the percentage significant difference for the variance of the number of wet days was 7% for WGEN and 10% for SIMMETEO. To obtain the minimum difference between the observed and generated number of wet days ($t^\star$), 12 yr of weather data were required for WGEN and 15 yr for SIMMETEO (Table 2, Fig. 2). For $F^\star$, only WGEN showed a significant response to the length of input data, with a minimum of 18 yr required. For SIMMETEO, $F^\star$ was small and independent of the number of years of input data (Table 2, Fig. 2).

Table 2. Parameter estimates for the 2-piece, segmented regression model (Eq. 1) relating the percentage significant difference of the mean and variance between observed and generated weather data with WGEN and SIMMETEO to the length of input data (n = 75). The percentage significant difference for base period of 30 yr (D_{30}) is also included. The parameter estimates are: a the intercept (%), b the slope (% yr^{-1}) and X_0 the intersecting point of the segmented model (yr) ± standard error. T_{max}: maximum temperature, T_{min}: minimum temperature. Significance: *p < 0.05, **p < 0.01. nd: not determinable

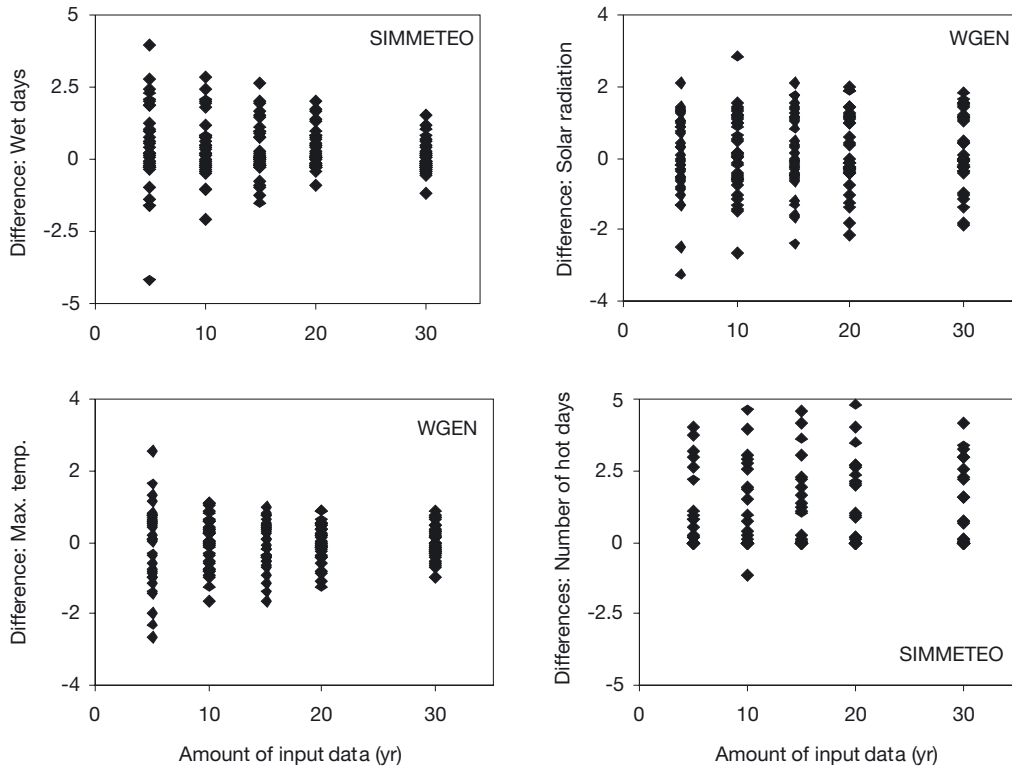| | WGEN | | | | SIMMETEO | | | |
	a	b	X_0	D_{30}	a	b	X_0	D_{30}
Number of wet days								
Mean	28.9**	−2.33**	11.7 ± 0.5	0.0	17.0**	−1.11**	15.3 ± 0.8	0.0
Variance	25.2**	−0.94*	17.8 ± 2.0	7.2	10.9**	−0.08	nd	10.0
Precipitation								
Mean	27.2**	−2.00**	13.0	0.6	35.0**	−3.00**	11.5 ± 0.3	0.0
Variance	36.8**	−0.38	nd	27.2	45.0**	−1.56**	15.2 ± 1.6	21.1
Solar radiation								
Mean	63.0**	−0.19	nd	56.7	33.0**	−1.11*	17.2 ± 4.2	15.0
Variance	38.8**	0.14	nd	43.3	18.9**	−1.67*	9.5 ± 0.4	1.1
Maximum temperature								
Mean	41.1**	−2.78**	12.1 ± 0.5	5.0	37.5**	−0.27	nd	31.2
Variance	40.6**	−2.22**	9.8 ± 3.4	19.9	13.7**	0.09	nd	14.9
Minimum temperature								
Mean	45.6**	−1.72**	25.5 ± 3.3	1.6	40.0**	−1.52**	23.7 ± 3.7	3.8
Variance	23.9**	−1.11*	13.7 ± 2.4	7.0	7.5**	0.03	nd	8.8
Number of hot days ($T_{max} > 35$)								
Mean	27.9**	−0.10	nd	25.3	40.7**	0.21	nd	46.6
Variance	28.3**	−0.07	nd	29.0	31.8**	−0.08	nd	31.2
Number of cold days ($T_{min} < 0$)								
Mean	27.8**	−0.85**	30.0 ± nd	2.4	41.1**	−0.12	nd	36.9
Variance	25.1**	−0.30	nd	16.2	24.3**	−0.06	nd	26.9

Fig. 1. Difference between observed and WGEN-generated solar radiation and maximum temperature and SIMMETEO-generated numbers of wet and hot days. The difference is shown as the generated minus the observed value
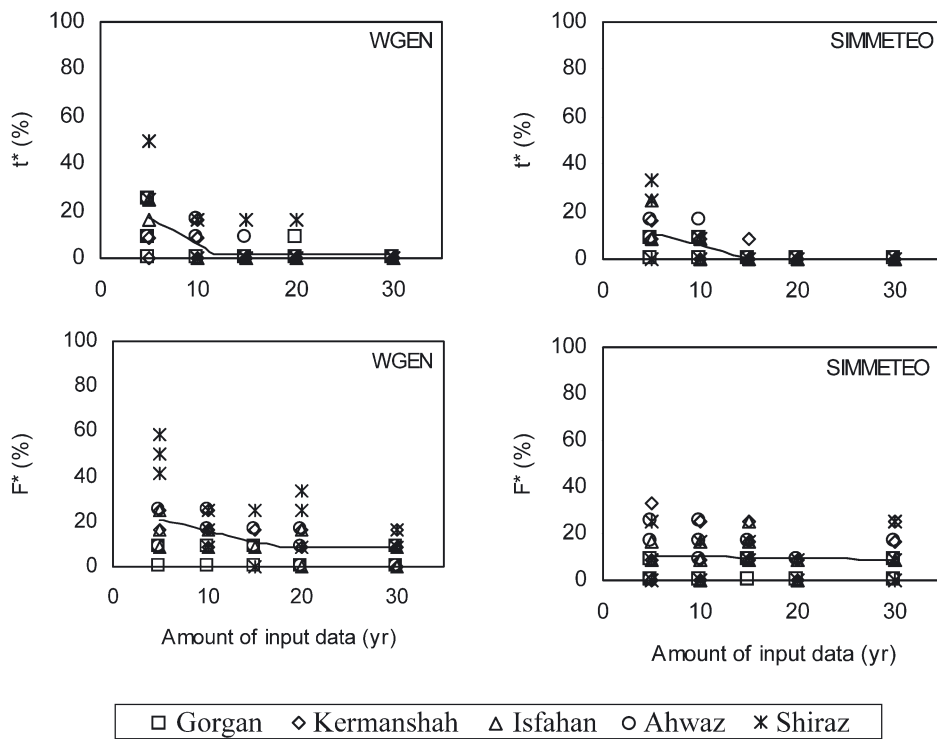

Fig. 2. Percentage t-test (t^*) and F-tests (F^*) between the observed and generated number of wet days that were rejected as a function of the length of the input data

WGEN and SIMMETEO performed well for generating the amount of precipitation. The values for t^* were 1% for WGEN and 0% for SIMMETEO. Both generators had a somewhat greater F^*, i.e. 27% for WGEN and 21% for SIMMETEO (D_{30} in Table 2). The minimum number of years required to obtain minimum t^* was 13 yr for WGEN and 12 yr for SIMMETEO (Fig. 3). For F^*, WGEN was not sensitive to the amount of input data, but SIMMETEO required a minimum of 15 yr (Table 2).

3.3. Solar radiation

WGEN showed a poor performance in generating daily solar radiation: 57% of the means and 43% of the variances of generated data were significantly different from the observed data. However, the daily solar radiation data generated with SIMMETEO were closer to the observed data, with only 15% of the means and 1% of the variances significantly different from observed data. This indicated a better performance for SIMMETEO compared to WGEN (D_{30} in Table 2). For WGEN, increasing the number of years of input data did not affect t^* or F^*. SIMMETEO, on the other hand, required a minimum of 17 yr for a minimum t^* and 10 yr for a minimum F^* (Table 2, Fig. 4).

3.4. Temperature

The comparison of the observed and WGEN-generated daily maximum temperature data rejected 5% of the t-tests and 20% of the F-tests. The analysis for SIMMETEO showed that 31% of the t-tests and 15% of the F-tests were significantly different between the generated and observed daily maximum temperature data (D_{30} in Table 2). These results indicated a good performance of WGEN and moderate performance of SIMMETEO. WGEN required 12 yr of input data to obtain a minimum t^* and 10 yr of weather data to obtain a minimum F^*. However, no significant decrease in t^* or F^* was found for SIMMETEO in response to increasing the number of input data (Table 2, Fig. 5).

The minimum temperature was generated successfully by both generators. For minimum temperature generated with WGEN, 2% of the t-tests and 7% of the F-tests were significantly different from the observed data. For SIMMETEO these values were 4% for the t-tests and 9% for the F-tests (D_{30} in Table 2). WGEN required 26 yr of input data to be efficient in reproducing the daily minimum temperature with respect to central tendency and 14 yr of input data to be efficient in reproducing minimum temperature with respect to variability (Table 2, Fig. 6). For SIMMETEO, only t^*
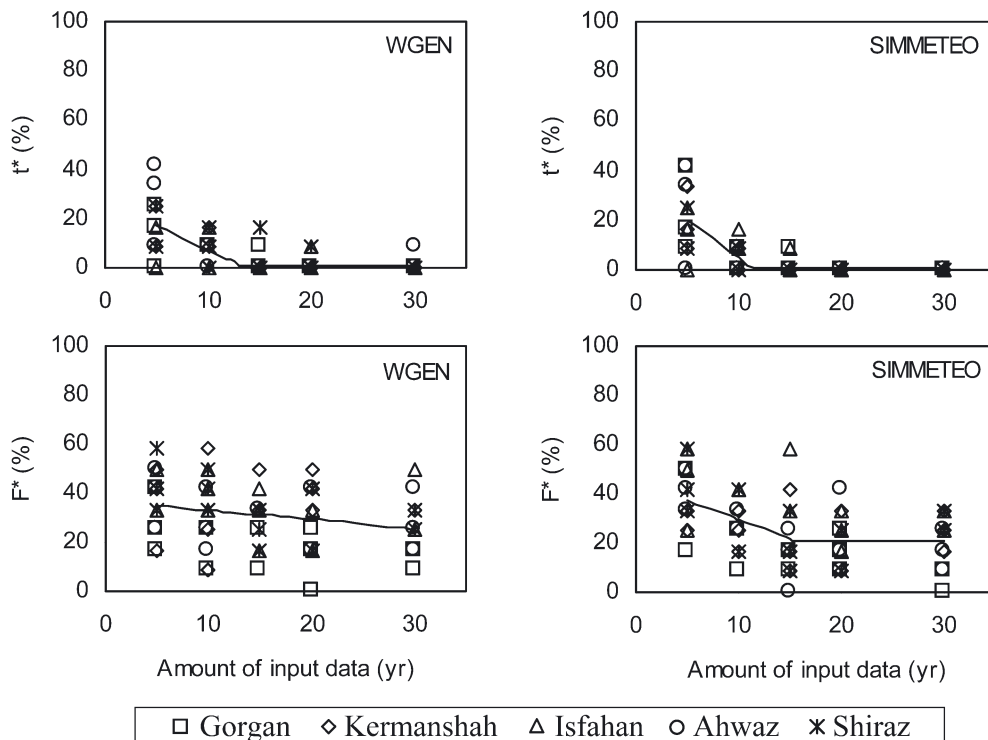

Fig. 3. Percentage t-test (t^*) and F-tests (F^*) between the observed and generated precipitation amount that were rejected as a function of the length of the input data
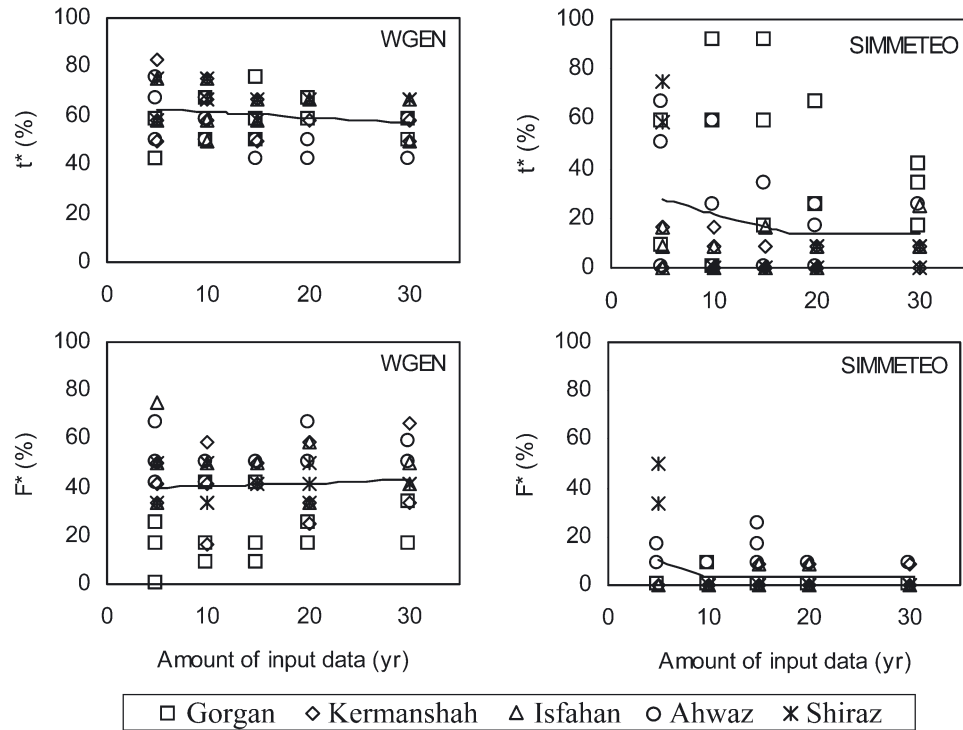
Fig. 4. Percentage *t*-test (*t**) and *F*-tests (*F**) between the observed and generated solar radiation that were rejected as a function of the length of the input data
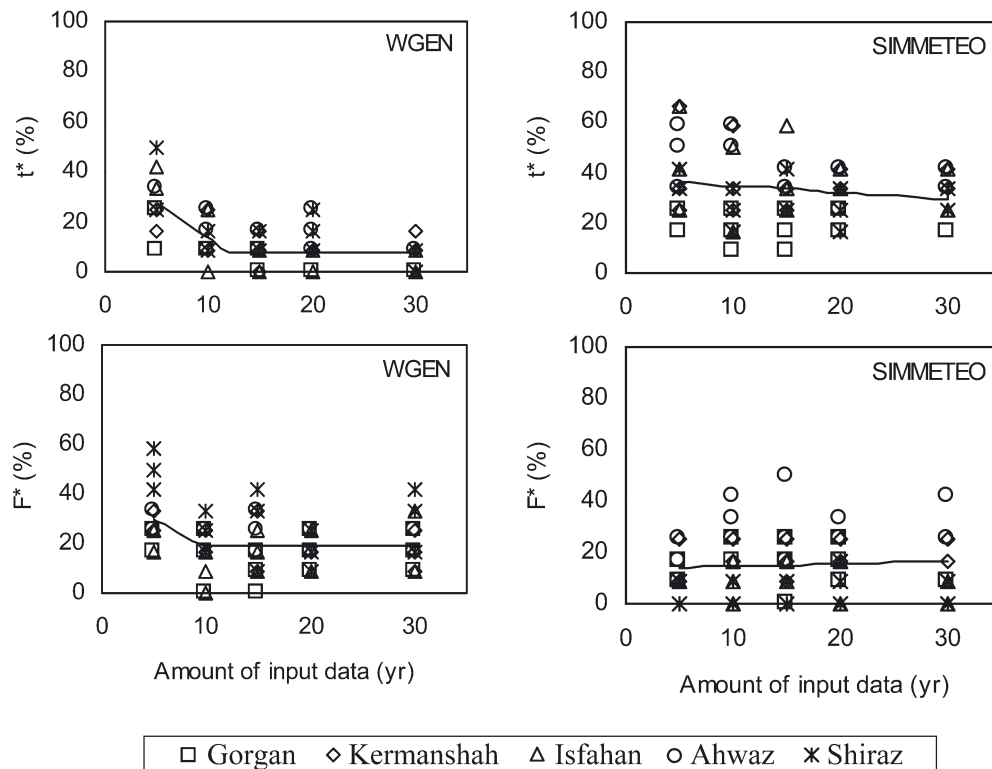

Fig. 5. Percentage *t*-test (*t**) and *F*-tests (*F**) between the observed and generated maximum temperature that were rejected as a function of the length of the input data
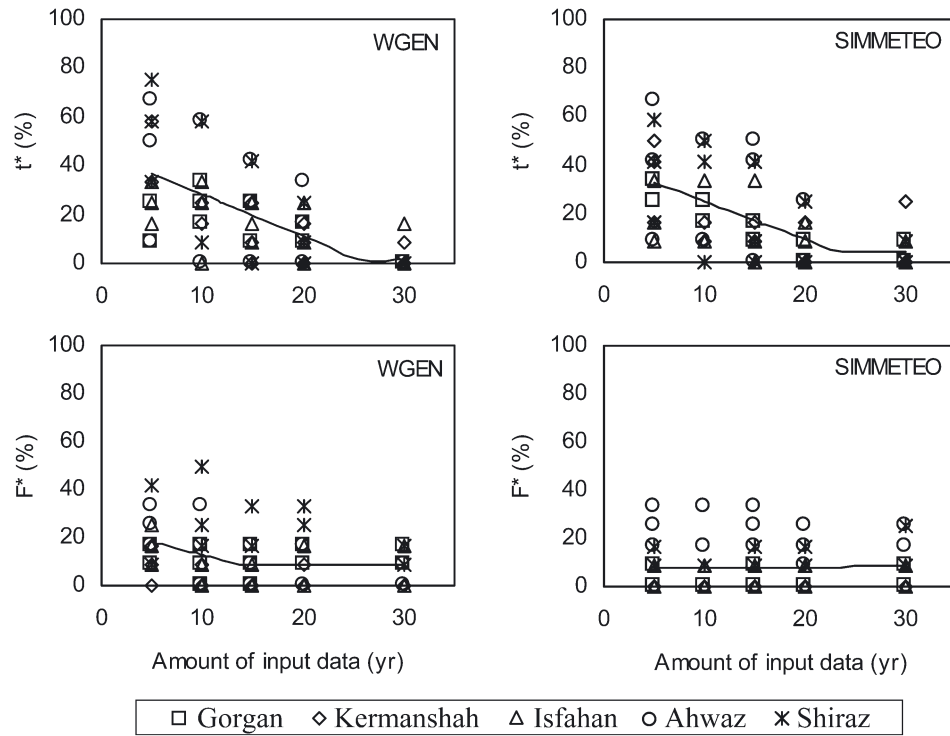
Fig. 6. Percentage *t*-test (t^*) and *F*-tests (F^*) between the observed and generated minimum temperature that were rejected as a function of the length of the input data
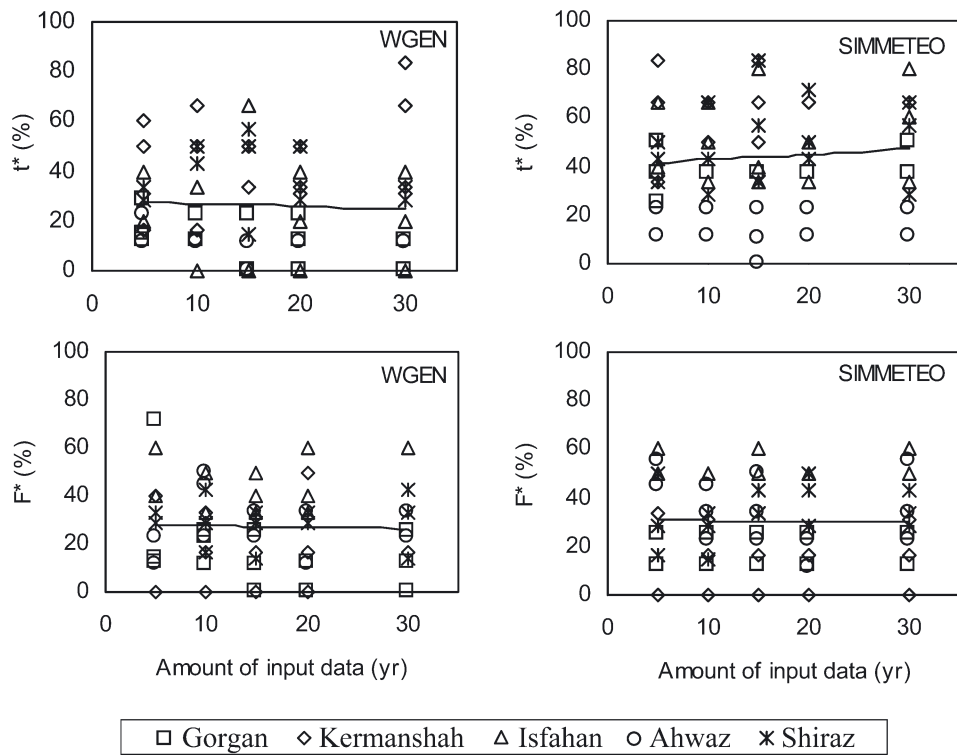

Fig. 7. Percentage *t*-test (t^*) and *F*-tests (F^*) between the observed and generated number of hot days that were rejected as a function of the length of the input data
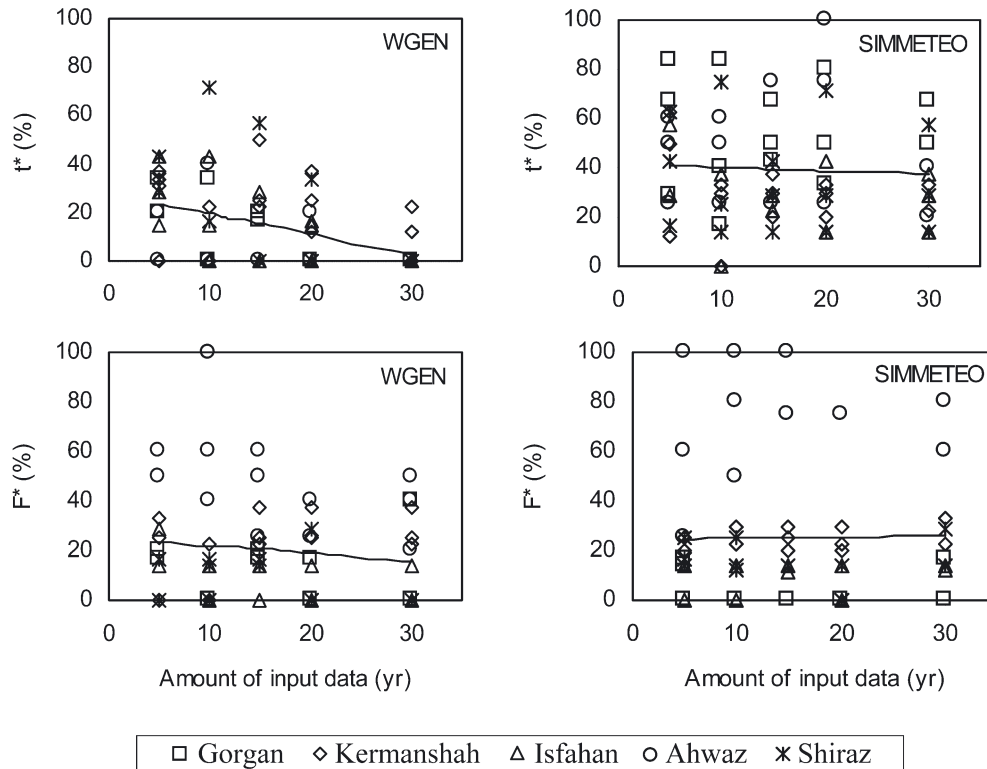
Fig. 8. Percentage *t*-test (*t**) and *F*-tests (*F**) between the observed and generated number of cold days that were rejected as a function of the length of the input data

was sensitive to the number of years of input data with a minimum requirement of 24 yr. Increasing the number of years of input data did not reduce F^* for SIMMETEO (Table 2, Fig. 6).

3.5. Extreme temperatures

The comparison of $T_{max} > 35$ generated by WGEN and observed data showed that 25% of the means and 29% of the variances were significantly different. The values for SIMMETEO were somewhat greater, i.e. 47 and 31%, respectively (D_{30} in Table 2). For $T_{min} < 0$, 2% of the *t*-tests and 16% of the *F*-tests were significantly different for WGEN, while 37% of the *t*-tests and 27% of the *F*-tests were significantly different for SIMMETEO (D_{30} in Table 2). SIMMETEO performed poorly in generating the number of extreme temperature events, i.e. hot and cold days, while WGEN showed a good performance in reproducing the number of frost days and a fairly good performance in reproducing the number of hot days.

With one exception, no reduction was observed in t^* and F^* between observed and WGEN- and SIMMETEO-generated number of hot ($T_{max} > 35$) and cold

($T_{min} < 0$) days as a result of a larger amount of input data (Table 2, Figs. 7 & 8). The exception was t^* for the number of cold days ($T_{min} < 0$) generated with WGEN (Fig. 8). The percentage significant differences in mean comparison (t^*) between observed and WGEN-generated $T_{min} < 0$ decreased linearly by 0.85% yr^{-1}, and all input weather data were required in order to obtain the best performance of this weather generator (Fig. 8).

4. DISCUSSION

Generated weather data must have similar statistical characteristics as the observed data to be able to have any value for applications (Semenov et al. 1998, Richardson 2000). In general, WGEN showed a good performance in reproducing the number of wet days and the total amount of precipitation, maximum and minimum temperature and the number of cold days, a moderate performance in reproducing the number of hot days and a weak performance in generating solar radiation. SIMMETEO showed a good performance in reproducing the number of wet days and the total amount of precipitation, solar radiation and minimum

temperature, a moderate performance in generating maximum temperature and a poor job in generating extreme temperatures (cold and hot days) (Table 2). A slightly poorer performance of SIMMETEO compared to WGEN is caused by its parameter estimation method because the generation algorithms are the same in both generators. While WGEN uses daily weather data for estimating the required generation parameters, SIMMETEO uses monthly summaries and regression relations for this purpose. The superiority of SIMMETEO in generating solar radiation was also due to its parameter estimation method that compensated for errors in its generation method. The poorer performance of both generators for extreme temperatures compared to other variables might be expected, as the extremes are not used as input parameters of the generators. Additional details and a full discussion on the efficiency of the generators can be found in Soltani & Hoogenboom (2003).

The main objective of this study was to determine the minimum number of years required for parameter estimation of the WGEN and SIMMETEO weather generators. The results from this study showed this was a function of both the weather generator as well as the weather variable that was generated. WGEN was more responsive to the number of years used as input data than SIMMETEO. WGEN required a specific minimum amount of data in 8 of the 14 cases (7 weather variables × 2 statistical characteristics, mean and variance) to be efficient in weather generating (Table 2). However, for SIMMETEO this requirement was observed in 6 of the 14 cases. For generating precipitation, including the number of wet days and total amount of precipitation, a smaller number of years was needed than for generating maximum and minimum temperature (Table 2). For WGEN, the minimum number of years was 14 for precipitation and 18 (15 when the number of hot days were ignored) for temperature. This was not in agreement with the findings of Richardson (2000), who reported that longer data records were needed for precipitation (20 yr) than for temperature (10 yr) for 2 locations in the USA (Temple, Texas, and Spokane, Washington). However, Richardson (2000) conducted a qualitative evaluation for the minimum data requirements in order to obtain stable, representative estimates of the generated parameters, while in this study we conducted a qualitative assessment for the number of years required to obtain generated data that had similar characteristics as the observed data.

The only reason that we found for the differences between the weather generators and weather variables with respect to minimum of year requirement was the relationship between the performance of the generator and its response to the number of years used

as input data (Table 2). When the WGEN and SIMMETEO generators showed a good performance, a significant response to the number of years used as input data was observed and a minimum number of years was required to obtain a minimum difference between observed and generated data, e.g. for precipitation amount (Fig. 3). In contrast, for the weather variables for which the generators did not mimic the observed data well, increasing the number of years used as input data was not effective in reducing the values of the t^* or F^* between observed and generated weather data, e.g. the number of hot days (Fig. 7). Similarly, slightly more responsiveness of WGEN to the amount of input data was due to a slightly better performance of the generator compared to SIMMETEO.

In most cases a larger number of years than the minimum requirement led to less fluctuation in the significant differences between observed and generated data, as demonstrated by the t^* analysis for the number of wet days (Fig. 2) and by the F^* analysis for solar radiation (Fig. 4), both generated with SIMMETEO. Therefore, it is recommended that users of weather generators should not limit themselves to the minimum year requirements and instead should use all available long-term historical weather data to parameterize the weather generator.

For practical applications, at least 15 yr of input data are needed to obtain the input parameters for both the WGEN and SIMMETEO weather generators. This will result in sequences of generated weather data that are similar to the actual data and have similar climatological characteristics. In conditions where the length of records is less than 15 yr, the generated data might have different climatological characteristics than the historical weather data. However, the statistical characteristics of the generated data might be similar to their corresponding input data, because weather generators by definition will generate the same (in terms of statistical characteristics) data as their input provided that the generator structure is adequate.

LITERATURE CITED

Geng S, Penning de Vries FWT, Supit I (1986) A simple method for generating daily rainfall data. Agric For Meteorol 36:363–376

Geng S, Auburn JS, Brandstetter E, Li B (1988) A program to simulate meteorological variables: documentation for SIMMETEO. Agronomy Progress Rep. 204. Dept of Agronomy and Range Science, University of California, Davis, CA

Guenni L, Charles-Edwards D, Rose R, Braddock R, Hogarth W (1991) Stochastic weather modelling: a phenomenolog-

ical approach. Math Comput Sim 32:113–118

Hansen JW, Mavromatis T (2001) Correcting low-frequency variability bias in stochastic weather generators. Agric For Meteorol 109:297–310

Hayhoe HN (2000) Improvements of stochastic weather data generators for diverse climates. Clim Res 14:75–87

Hoogenboom G (2000) Contribution of agrometeorology to the simulation of crop production and its applications. Agric For Meteorol 103:137–157

Hoogenboom G, Wilkens PW, Tsuji GY (eds) (1999) DSSAT v3 Vol 4. University of Hawaii, Honolulu, HI

Jones PG, Thornton PK (2000) MarkSim: software to generate daily weather data for Latin America and Africa. Agron J 92:445–453

Larsen GA, Pense RB (1982) Stochastic simulation of daily climatic data for agronomic models. Agron J 74:510–514

Mavromatis T, Hansen JW (2001) Interannual variability characteristics and simulated crop response of four stochastic weather generators. Agric For Meteorol 109:283–296

McCaskill MR (1990) TAMSIM—a program for preparing meteorological records for weather driven models. Tropical Agronomy Technical Memorandum No. 65, 1990. CSIRO, Division of Tropical Crops and Pastures, Brisbane

Meinke H, Carberry PS, McCaskill MR, Hills MA, McLeod I (1995) Evaluation of radiation and temperature data generators in the Australian tropics and sub-tropics using crop simulation models. Agric For Meteorol 72:295–316

Pickering NB, Hansen JW, Jones JW, Wells CM, Chan VK, Godwin DC (1994) WeatherMan: a utility for managing and generating daily weather data. Agron J 86:332–337

Puche M, Silva O (2001) Assessments of generated rainfall data (EPIC) for a seasonal rainfall site. In: Proc 2nd International Symposium on Modelling Cropping Systems, 16–18 July, Florence, Italy. European Society for Agronomy, p 213–214

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

Richardson CW (2000) Data requirement for estimation of weather generation parameters. Trans Am Soc Agricult Eng 43:877–882

Richardson CW, Wright DA (1984) WGEN: A model for generating daily weather variables. ARS-8, Agricultural Research Service, US Dept of Agriculture, Washington, DC

SAS Institute (1989) SAS/STAT user's guide, Ver 6, 4th edn. SAS Institute, Cary, NC

Semenov MA, Brooks RJ, Barrow EM, Richardson CW (1998) Comparison of WGEN and LARS-WG stochastic weather generators for diverse climates. Clim Res 10:95–107

Soltani A, Hoogenboom G (2003) A statistical comparison of the stochastic weather generators WGEN and SIMME-TEO. Clim Res 24:215–230

Thornton PK, Hoogenboom G (1994) A computer program to analyze single-season crop model outputs. Agron J 86: 860–868

Tsuji GY, Uehara G, Balas S (1994) Decision Support System for Agrotechnology Transfer, Ver 3. International Benchmark Sites Network for Agrotechnology Transfer, University of Hawaii, Honolulu, HI

Warnock R, Puche M (2001) Assessing generated weather data to explore management scenarios in Venezuela. In: Proc 2nd International Symposium on Modelling Cropping Systems, 16–18 July, Florence, Italy. European Society for Agronomy, p 255–256