TESTING THE PERFORMANCE OF DIFFERENT PROCESSES TO GENERATE TEMPERATURE AND SOLAR RADIATION: A CASE STUDY AT LLEIDA (NORTHEAST SPAIN)

F. Castellvi, C. O. Stockle, I. Mormeneo, J. M. Villar

ABSTRACT. *The type of weather data that is available largely conditions weather generation at any given location. Weather generators that provide flexibility in the selection of generation methods are advantageous. ClimGen, a multivariate autoregressive first–order weather generation model, was used to compare different methods to generate weather series of daily maximum (Tx) and minimum (Tn) temperatures and bright sunshine hours (Bsh). Two of these methods (A1 and A2) can be applied at stations where daily precipitation (Prec), Tx, Tn, and Bsh data are available. Another method (B) can be applied at stations where Bsh data are lacking. Method A1 uses Tx, Tn, and Bsh in the multivariate generation process. Method A2 uses Tx, $\Delta T = Tx - Tn$, and Bsh. Method B only uses Tx and Tn. Two temperature–based methods to estimate solar radiation were used to complement method B. In addition, the two–parameter Weibull and Gamma distributions were compared for the generation of precipitation amounts.*

The different methods were evaluated for agricultural applications at the Lleida region, in northeast Spain, a continental semiarid climate. Different climatological variables and agricultural indices were calculated using actual and generated data, and statistically compared at 5% and 10% levels of significance. It was concluded that: (1) The two–parameter Weibull distribution performed better in this region than the two–parameter Gamma distribution; (2) method A1 was slightly superior to method A2, mainly because the minimum temperature and the bright sunshine hours were better replicated; (3) method B was found satisfactory, but the complementary methods to estimate solar radiation from temperature did not perform well when they were calibrated at a given station and then used to generate for a station in a surrounding area; and (4) the methods tested well replicated chill units, growing degree days, the mean of heating degree days (but not its variance), and long hot or cold sequences, but they did not appear to reliably replicate short sequences as needed to assess problems such as crop frost potential. The generation of bright sunshine hours, as introduced in this study, is an alternative for locations where solar radiation data are not available to parameterize the weather generator.

Keywords. *Weather generators, Bright sunshine hours.*

Weather generators are often used to drive hydrological, agricultural, and environmental models. For example, Richardson (1985) addressed the use of generated weather in agricultural simulation models. In addition, weather generators are a practical tool to bypass problems like data unavailability, missing data, short records, or poor quality of data. Existing generators can be broadly categorized as either specific to a particular task, generating at most two or three weather variables, or specific to a particular site or range of sites (Peiris and McNicol, 1996).

WGEN (Richardson and Wright, 1984), a weather generator often used in the U.S., has constituted the basis for several other generators (e.g., Chineke et al., 1999; Johnson et al., 1996; Nicks et al., 1990; Wallis and Griffiths, 1995). WGEN requires as input a daily series of measured precipitation, maximum and minimum temperature, and solar radiation for some representative period of time. It preserves the interdependence among the generated weather variables as well as the seasonal characteristics of each variable. A first–order Markov chain is used to generate wet and dry days, and a two–parameter gamma distribution function is used to assign the amount of precipitation for wet days. The occurrence of precipitation (Prec) is the primary factor conditioning maximum (Tx) and minimum (Tn) temperature and solar radiation (Rs) values. These values result from a continuous multivariate stochastic process with the daily means and standard deviations conditioned by the dry or wet state of the day. WGEN uses the generation process introduced by Matalas (1967), implemented based on averaged correlation matrix coefficients obtained from a large U.S. weather database (Richardson and Wright, 1984).

For many applications in agriculture, Prec, Tx, and Tn data are often available at or near the site of interest. Rs data are less likely to be available, but series of bright sunshine hours (Bsh) may be available. In such cases, it is not possible to use the generation process in WGEN unless estimations of Rs are made. ClimGen is a weather generator that also uses

Article was submitted for review in July 2000; approved for publication by the Soil & Water Division of ASAE in January 2002.

The authors are **Francesc Castellvi**, Associate Professor, and **Josep M. Villar**, **ASAE Member Engineer**, Professor, ETSEA, Universitat de Lleida, Lleida, Spain; **Claudio O. Stockle**, Professor, Biological Systems Engineering Department, Washington State University, Pullman, Washington; and **Ines Mormeneo**, Associate Professor, Agronomia, Universidad Nacional del Sur, Bahía Blanca, Argentina. **Corresponding author:** Claudio O. Stockle, Biological Systems Engineering Dept., Washington State University, Pullman, WA 99164–6120; phone: 509–335–3826; fax: 509–335–2722; e–mail: stockle@wsu.edu.

a first–order Markov chain and the multivariate generation process proposed by Matalas (1967), but it includes features allowing greater flexibility. The coefficients of the required correlation matrices are determined for each location of interest. Therefore, when Bsh data are available, correlations can be determined with this variable to generate Tx, Tn, and Bsh. Relationships between bright sunshine hours, atmospheric transmissivity, and solar radiation are available (Monteith and Unsworth, 1990). In addition, when series of Rs or Bsh are lacking, ClimGen can generate daily series of Tx and Tn by reducing the generation process to two dimensions. Methods to estimate Rs from temperature can then be applied. To estimate Rs, the Donatelli and Campbell model (Donatelli and Campbell, 1998), based on the Bristow and Campbell model (Bristow and Campbell, 1984) is available in ClimGen. Other methods to estimate solar radiation from precipitation and temperature data have been reported in the literature (Hunt et al., 1998; Yin, 1999a, 1999b).

To assign the amount of precipitation for wet days, ClimGen uses a two–parameter Weibull distribution function. When sufficiently long Prec records are not available, the distribution function is parameterized following Selker and Haith (1990). The fact that ClimGen offers several optional methods to generate daily weather series, as discussed above, allows its application at sites with different levels of data availability. However, the question arises regarding the relative performance of the available methods.

The overall goal of this work was to construct a suitable weather generator model for the Lleida region (northeast Spain) and its associated agricultural area as a study case. However, the methodology presented here can be applied to any location. The objective of this study was to evaluate three methods for multivariate generation, two distribution functions to assign the amount of precipitation on a wet day, and two methods to estimate solar radiation in this region. The study region is located in the Ebro river basin, characterized by a semiarid continental climate. Details of the region can be found in Castellvi et al. (1996).

THEORY

The daily generation of weather variables is conditioned by the dry or wet status of the day. A first–order Markov chain with two states was used to determine wet or dry days. Good results with this method have been reported (Arnold and Elliot, 1996; Gabriel and Neumann, 1962; Johnson et al., 1996; Richardson and Wright, 1984). The generation of Prec, Tx, Tn, and Bsh, and the estimation of Rs values were performed as follows.

GENERATION OF THE AMOUNT OF PRECIPITATION

The amount of precipitation assigned to a wet day can be generated by sampling from a two–parameter Weibull (eq. 1) or a two–parameter gamma distribution (eq. 2) function:

$$F(p) = 1 - \exp[-\beta\, p]^{\alpha} \tag{1}$$

$$F(p) = [\beta^{\alpha}\, \Gamma(\alpha)]^{-1}\, p^{\alpha-1}\, e^{-p/\beta} \tag{2}$$

where

F(p) = cumulative probability of a precipitation amount equal or less than p

β and α = two parameters specific for each distribution function on a monthly basis

$\Gamma(\alpha)$ = gamma function of α.

These distributions are sampled for each precipitation event using the inverse method.

GENERATION OF TEMPERATURE AND BRIGHT SUNSHINE HOURS

The time series of a weather variable *j* is reduced to a time series of residual elements as follows:

$$\chi_{n,i}(j) = \frac{X_{n,i}(j) - X_i^k(j)}{\sigma_i^k(j)} \tag{3}$$

where

$\chi_{n,i}(j)$ = residual component for the weather variable *j*, day *i*, and year *n*

$X_{n,i}(j)$ = daily value of the variable

$X_i^k(j)$ and $\sigma_i^k(j)$ = daily mean and standard deviation, with $k = 0$ to indicate dry days and $k = 1$ for wet days.

The residual series for each variable *j* are supposed to be normally distributed with mean zero and variance of one, and described by an auto–regressive linear model. The weakly stationary generating multivariate process proposed by Matalas (1967) is used to generate the residual series as follows:

$$\chi_{n,i}(j) = A\chi_{n,i-1}(j) + B\varepsilon_{n,i}(j) \tag{4}$$

where

$\chi_{n,i}(j)$ and $\chi_{n,i-1}(j)$ = $j \times 1$ matrices for day *i* and $i - 1$ of year *n* whose elements are the residuals of the weather variables for day *i* and $i - 1$ of year *n*, respectively

$\varepsilon_{n,i}(j)$ = $j \times 1$ matrix of independent random components normally distributed with mean zero and variance one

A and B = $j \times j$ matrices determined by the lag–0 cross–correlation and the lag–1 cross–correlation between residuals.

Daily generated values for each weather variable are determined by rearranging terms in equation 3, where their respective daily means and standard deviations are obtained from monthly means using spline functions. The periodical smoothing is conditioned on wet or dry status.

The versatility of ClimGen was used to implement the following two scenarios:

Scenario A

Long records of Prec, Tx, Tn, and Bsh are available. Two different methods to generate temperatures and bright sunshine hours were applied:

Method A1: The weather variables used in the generation process (eq. 4) were $j = 1$ for Tx, $j = 2$ for Tn, and $j = 3$ for Bsh.

Method A2: This is based on the inter–dependency between the daily temperature amplitude ($\Delta T = Tx - Tn$) and Bsh. The weather variables used were $j = 1$ for Tx, $j = 2$ for

ΔT, and $j = 3$ for Bsh. The minimum temperature was obtained from the generated maximum temperature and the amplitude of the temperature, Tn = Tx – ΔT.

Scenario B

Only long records of Prec, Tx, and Tn are available. In this case, the generation process was reduced to two dimensions, using $j = 1$ for Tx and $j = 2$ for Tn. This case will be referred to as method B.

Estimation of Bright Sunshine Hours and Solar Radiation

When only precipitation and temperature data are available (scenario B), methods to estimate solar radiation can be applied. Two methods were evaluated, a modified TAMSIM method (McCaskill, 1990) and the Donatelli and Campbell (1998) method. The selection of these methods was partially based on previous results obtained by Meinke et al. (1995), indicating that the TAMSIM method (McCaskill, 1990) appears to give better performance than the Bristow and Campbell (1984) method, and those from Hunt et al. (1998), suggesting that methods based on the difference between maximum and minimum temperature give a better performance.

The modified TAMSIM method was calibrated to estimate daily sunshine fraction (i.e., bright sunshine hours over total daylength), which was converted to solar radiation (Allen et al., 1994). The Donatelli and Campbell (1998) method estimates the atmospheric transmissivity. After calculation of daily extraterrestrial solar radiation, the method was used to estimate solar radiation.

Modified TAMSIM Method

Following Hunt et al. (1998), the daily difference between the maximum and minimum temperature was introduced in the TAMSIM method (McCaskill, 1990), resulting in a modified version for predicting sunshine fraction as follows:

$$(n/N) = a + b\cos(\theta) + c\sin(\theta) + d\cos(2\theta) + e\sin(2\theta) + f\,P_{Jday-1} + g\,P_{Jday} + h\,P_{Jday+1} + i\,\Delta T_{Jday} \quad (5)$$

where

n/N = sunshine fraction
a, b, c, d, e, f, g , h, and i = calibrated coefficients
θ = day of the year expressed in radians (θ = Jday $2\pi/365$)
P_{Jday-1}, P_{Jday}, and P_{Jday+1} = precipitation converted to a binary form (0 for dry days and 1 for wet days) for the previous, current, and following day, respectively
ΔT_{Jday} = difference between maximum and minimum temperatures for the current day.

Conversion from sunshine fraction to solar radiation was made using the following equation (Allen et al., 1994):

$$Rs = (a + b\,n/N)\,Ra \quad (6)$$

where a and b are calibrated coefficients, and Ra is extraterrestrial radiation for a 24–hour period (MJ m^{-2} d^{-1}), calculated as follows:

$$Ra = 37.6[1 + 0.033\cos(0.0172\,J)] \times [\omega\sin\phi\sin\delta + \cos\phi\cos\delta\sin\omega] \quad (7)$$

where

J = number of the day in the year
δ = solar declination (radians), calculated as $\delta = 0.409 \sin(0.0172\,J - 1.39)$
ϕ = latitude (in radians, negative for southern hemisphere)
ω = sunset hour angle (in radians), calculated as $\omega = \arccos(-\tan\phi\tan\delta)$.

Simulated annealing (Goffe et al., 1994) was used to optimize the modified TAMSIM method (eq. 5). Simulated annealing is a global optimization method that distinguishes between different local optima. The minimum root mean square error was used as the function to be minimized.

Donatelli and Campbell (1998) Method

The Bristow and Campbell (1984) method is based on the relationship between the daily atmospheric transmissivity and the difference between the daily maximum and the average of the minimum temperatures for the current and following day (e.g., $\Delta T_{Jday} = Tx_{Jday} - 0.5\,Tn_{Jday} + Tn_{Jday+1}$). However, some correction must be made when calculating ΔT_{Jday} in sites where cold and warm air masses often move through the area. The relationship depends on three coefficients that must be locally calibrated. The Donatelli and Campbell (1998) method is based on Bristow and Campbell (1984), introducing the parameterization of these three empirical coefficients from maximum and minimum temperature.

The solar radiation is estimated according to the following equation:

$$Rs = \tau_a\,Ra \quad (8)$$

where Ra is the extraterrestrial radiation (eq. 8), and τ_a is the daily atmospheric transmissivity estimated using the following equation:

$$\tau_a = \tau_c\left[1 - \exp\left(\frac{b f_1 f_2 \Delta T_{Jday}^2}{\Delta T_m}\right)\right] \quad (9)$$

where

τ_c = clear sky transmissivity
ΔT_{Jday} and ΔT_m = daily (as described above) and monthly amplitude of the temperature difference, respectively
f_1 = function that depends on daily average temperature (T): $f_1 = 0.017 * \exp[\exp(-0.053\,T)]$
f_2 = function that depends on daily minimum temperature: $f_2 = \exp(Tn/Tnc)$
b and Tnc = empirical parameters.

Following Donatelli and Campbell (1998), the parameters of equation 9 are optimized when the slope of estimated (eqs. 8 and 9) versus measured Rs data is close to one, the root mean square error is low, and the coefficient of residual mass (Loague and Green, 1991) does not depart noticeably from zero.

MATERIALS AND METHODS

Two stations, Lleida and Monte Julia, were selected to apply the methods. Six years of daily Rs data and thirty years of Prec, Tx, Tn, and Bsh data were available at Lleida. At Monte Julia, ten years of Prec, Tx, Tn, and Rs data were available. Monte Julia is located 40 km from Lleida. However, Monte Julia is windier and drier than Lleida.

Daily estimations of Rs using equation 5 were obtained after calibration of equation 6 using the six years of actual Rs data from Lleida, which was assumed to play the role of a first–order station. Equation 9 was also calibrated at the first–order station. The empirical parameters needed by the two methods to estimate solar radiation were assumed valid in the surroundings of the first–order station. The two estimation equations were tested at Monte Julia, which was assumed to play the role of a surrounding thermo–pluviometric station.

STATISTICAL ANALYSIS FOR GENERATED DATA

Ten series (samples) of 30 years of daily Prec, Tx, Tn, and Bsh were generated for the first–order station (Lleida). The following variables and statistics were used to compare generated versus actual data:

- Cumulative distribution of the length of dry periods.
- Mean and variance of the lengths of dry and wet periods.
- Monthly means and variances of the amount of precipitation, maximum and minimum temperatures, and relative sunshine hours.
- Cumulative distribution of the length of hot and cold periods. For agricultural purposes, sequences of days with Tx > 30° C were defined as hot periods. Sequences with Tn < 0° C were defined as cold periods.
- The mean and variance of temperature–based agricultural indices including:
  Frost–free period: The length of the frost–free period was obtained by taking the first of July as a demarcation between spring and fall (Vestal, 1971).
  Growing degree days for maize, calculated for the period between DOY 105 and 258 and with a base temperature of 10° C, and for barley, calculated for the period from DOY 350 to 166 (next year) and with a base temperature of 0° C.
  Chill units and heating degree days were calculated with a base temperature of 7° C and 18° C, respectively.

The following statistical tests were used to compare actual and generated data: t–test for means, F–test for standard deviations, and the Kolmogorov–Smirnoff test for cumulative distribution functions. All tests were conducted at the 5% and 10% levels of significance.

RESULTS

PRECIPITATION

Table 1 shows, as an example, the cumulative distribution function of the length of dry days for the actual and generated data for one of the ten generated 30–year series. Also shown are the total number and the maximum length of dry periods observed and obtained by generation. The results were similar for the other generated series. The first–order Markov chain well reproduced the actual cumulative wet and dry events since all ten runs replicated the observed sequences at a 10% level of significance.

Table 2 shows, for the same sample presented in table 1, the critical level of significance of the t–test and F–test applied to the mean length of dry and wet periods and their respective standard deviations. These results were similar for the other nine generated series (not shown). The Markov chain did not reproduce the variance of the length of dry periods (only passed the test at the 1% level of significance). Otherwise, the performance of the generation process was good.

Table 3 shows monthly values of interest, such as mean precipitation (Prec) and the frequency of wet days (fwet). The percentage of t–tests and F–tests rejected for monthly means

Table 1. Cumulative probability distribution of the length of dry days for 30 years of actual and generated data.

Length of Dry Sequences	Actual Data	Generated Data
1	0.186	0.152
2	0.332	0.287
3	0.436	0.389[a]
4	0.512	0.472
5	0.573	0.551
6	0.625	0.623
7	0.670	0.678
8	0.706	0.710
9	0.738	0.744
10	0.776	0.782
11	0.796	0.813
12	0.820	0.836
13	0.844	0.857
14	0.864	0.867
15	0.878	0.881
16	0.892	0.890
17	0.907	0.904
18	0.913	0.913
19	0.920	0.922
20	0.929	0.929
25	0.958	0.966
30	0.973	0.981
35	0.985	0.988
40	0.988	0.994
45	0.993	0.997
50	0.993	0.998
Total number of sequences	1212	1222
The longest sequence (days)	61	81

[a] Maximum difference (%) 4.7 $D_{0.05}$ (%): 5.5 $D_{0.10}$ (%): 4.9

Table 2. Mean and standard deviation (Std) values for dry and wet periods (days) calculated using 30 years of actual and generated data, and critical levels of significance obtained when t–tests (for means) and F–tests (for Stds) were applied.

	Actual Data		Generated Data		Critical Level of Significance	
	Dry	Wet	Dry	Wet	Dry	Wet
Mean	7.37	1.63	7.33	1.61	0.90	0.63
Std	8.31	0.999	7.71	1.02	0.01	0.45

and standard deviations are also shown. Both the Weibull and gamma distributions well reproduced the annual mean and standard deviation of precipitation amount (not shown). On a monthly basis, the Weibull distribution function well reproduced the means, passing all t–tests at the 5% level of significance and failing only one run in January and April at the 10% level of significance. The standard deviations were well reproduced except for February and September, where 70% and 60% of the runs, respectively, did not replicate the observed data at the 5% level of significance.

The results indicate that the two–parameter Weibull distribution was superior to the two–parameter gamma distribution function, both for means and standard deviations. Overall, the replication of precipitation using a first–order Markov chain and a Weibull distribution function can be considered adequate for long–term agricultural analyses in the study region. The poor replication of the variance in February and September is not critical in the region. February is the driest month of the year, and in September most crops and fruit trees are ready for harvest.

Temperature and Bright Sunshine Hours

Figure 1 shows the percentage of t–tests rejected at the 5% and 10% levels of significance for Tx, Tn, and Bsh (converted to Rs using eq. 6) for each method. Figure 2 shows similar information for the F–tests.

Maximum Temperature

All methods performed well for the means at the 5% level of significance. At the 10% level of significance, 65% of the months were successfully replicated with a probability of 70% or better (i.e., seven or more from the ten runs). Methods A1 and A2 tended to perform poorly in early spring and fall. For method B, fall and winter were the most difficult seasons to replicate. Although all methods performed similarly, method B was slightly better in the summer.

The variances were more difficult to reproduce, especially at the 10% level of significance. Only about 40% of the months were successfully reproduced with a probability of 50%. The most difficult months to reproduce were those characterized by the presence of moving fronts across the area (mainly during the fall).

Overall, the methods performed similarly, although method B tended to be slightly better, with a lower total number of t–tests and F–tests rejected and a slightly better performance in the summer, an important season for agricultural practices.

Minimum Temperature

Method A1 had the best performance, and method A2 had the worst. Method A1 had the lowest number of t–tests and F–tests rejected and tended to perform better in those months when minimum temperature is important for agriculture. The minimum temperature from June to September has less relative importance.

For means, method A1 performed well at both levels of significance. Runs were rejected with a probability greater than 30% at the 10% level of significance only for July and September. Method B also performed well for means at the 5% level of significance, except for March (50% of t–tests rejected), a critical month for growing degree days for barley and for development of fruit trees. However, the proportion of runs rejected in some months dramatically increased at the 10% level of significance. Unfortunately, those months included October and December, which play a critical role for frost–free period, chill units, and growing degree days of barley. Despite the fact that method B tended to have a lower percent of rejections than method A2, by taking into account the critical months, the performance of both methods were comparable.

For variances, at the 5% level of significance, method A1 did not perform well in April, which is the most critical month for frost. At the 10% level of significance, the number of F–tests rejected increased noticeably. However, the winter season was well replicated by this method. Method B also performed well in the critical months, except for January (only at the 10% significance level) and October with up to 50% of runs rejected at both levels of significance. Methods A1 and B were clearly superior to method A2.

Bright Sunshine Hours

For means, method A2 had the worst performance since the number of t–tests rejected was greater and was concen–

Table 3. Monthly values of precipitation (Prec), frequency of wet days (fwet) and percent rejection of t–tests and F–tests applied to monthly precipitation generated with the Gamma and Weibull distribution functions.

			Number of Tests Rejected (%)							
			Gamma				Weibull			
Month	Prec (mm)	fwet	t–test 5%	F–test 5%	t–test 10%	F–test 10%	t–test 5%	F–test 5%	t–test 10%	F–test 10%
1	27.77	0.227	100	100	100	100	0	0	10	0
2	14.23	0.168	100	100	100	100	0	70	0	70
3	28.12	0.158	40	0	40	20	0	20	0	30
4	39.37	0.216	20	40	30	50	0	0	10	0
5	50.18	0.247	50	90	70	90	0	0	0	10
6	33.58	0.192	100	100	100	100	0	10	0	30
7	14.44	0.092	10	40	60	60	0	20	0	50
8	20.45	0.108	40	100	80	100	0	10	0	20
9	39.67	0.143	10	10	10	10	0	60	0	70
10	40.41	0.176	40	20	60	20	0	30	0	30
11	26.56	0.19	0	10	0	20	0	10	0	30
12	26.67	0.244	100	100	100	100	0	10	0	20

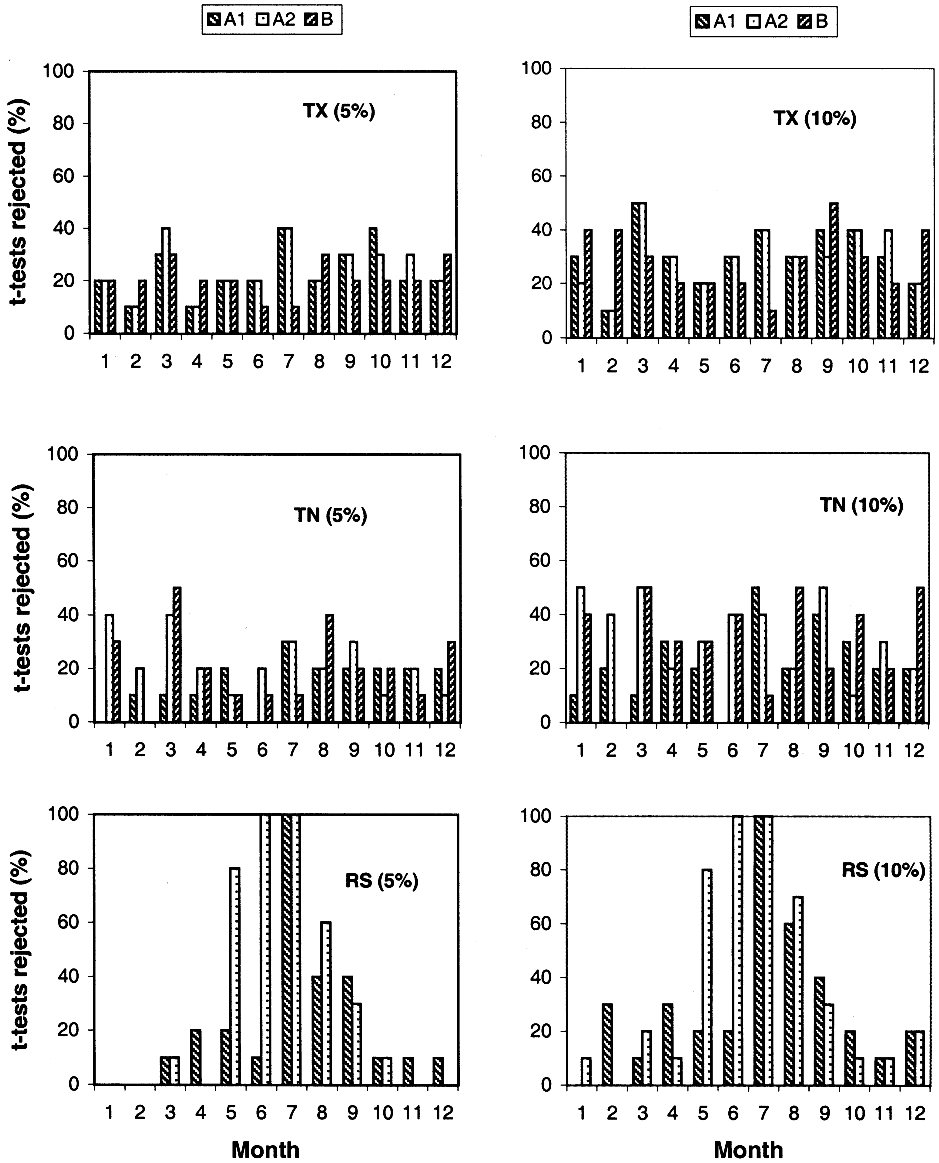

Figure 1. Monthly percentage of t–tests rejected for three methods when generating maximum (Tx) and minimum (Tn) temperature and bright sunshine hours converted to solar radiation (Rs) using equation 6. Results are shown for tests performed at the 5% (left) and the 10% (right) levels of significance.

trated in late spring, early fall, and summer seasons. The number of t–tests rejected using method A1 were lower and were mainly concentrated in July, August, and September. For variances, method A1 was also superior since the number of F–tests rejected was lower. Overall, method A1 was superior to method A2. However, neither method had a reasonable performance for means or variances during the critical months (summer season).

Cumulative Distributions of Hot and Cold Sequences

Table 4 shows the cumulative distribution of hot and cold sequences from actual and generated data using one sample run as an example. Table 5 shows the fraction of runs rejected at the 5% and 10% levels of significance.

All methods had large percentage of rejections for hot and cold sequences. The test tended to fail for sequence lengths of one to three days (table 4). Overall, results in table 5 show that none of the three methods evaluated gave enough confidence in replicating short hot or cold sequences. However, the methods seem more adequate for analyses involving long sequences (four days or longer). For agricultural applications, good replication of long sequences of hot weather are of interest for processes such as potential evapotranspiration and crop water stress. However, the methods may not be adequate to analyze phenomena such as crop frost damage, where short sequences are of interest.

Temperature–Based Indices

Table 6 shows the percentage of t–tests and F–tests rejected for each of the three generation methods evaluated. Method A2 was the only method able to replicate the frost–free period (only two rejections), but the interannual variability was more difficult to reproduce since the F–test was rejected for 40% and 50% of the runs at the 5% and 10% levels of significance, respectively. The variance of the

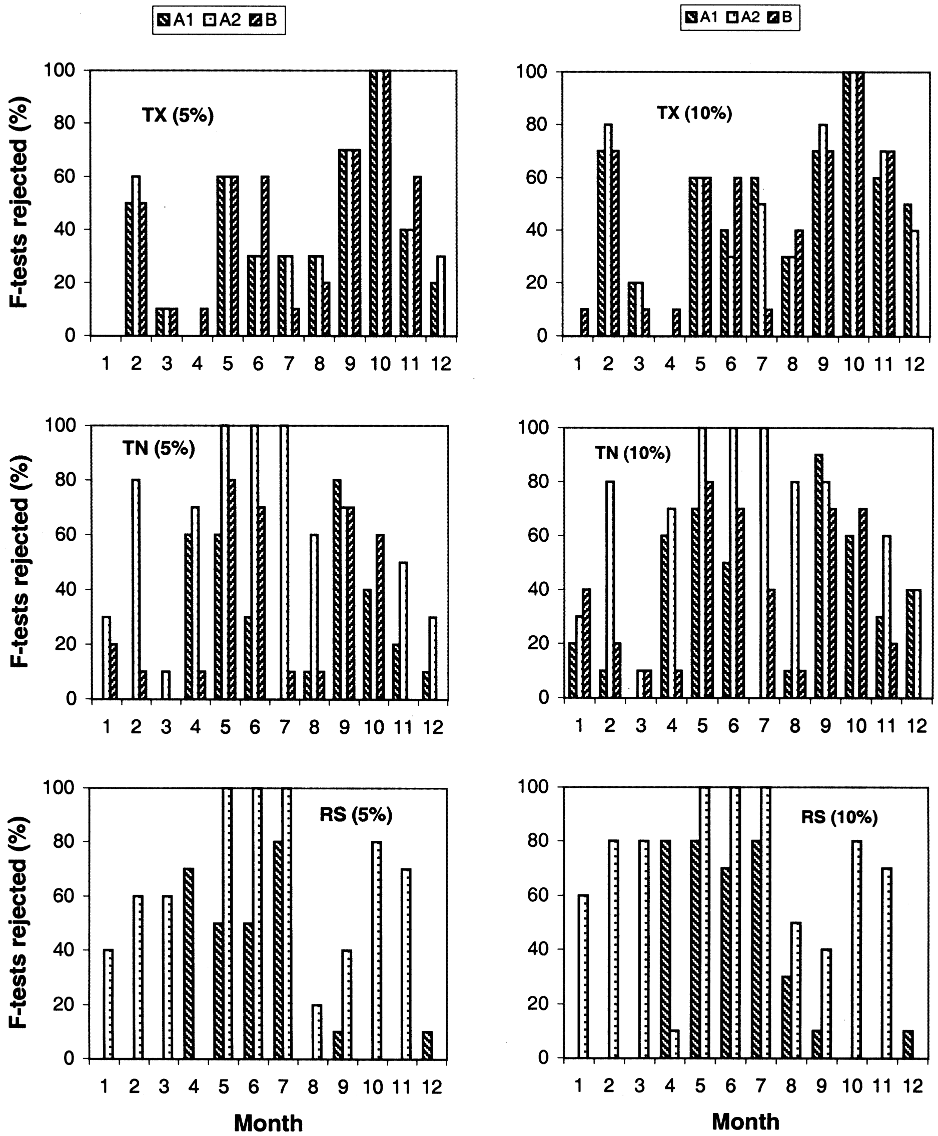

Figure 2. Monthly percentage of F–tests rejected for three methods when generating maximum (Tx) and minimum (Tn) temperature and bright sunshine hours converted to solar radiation (Rs) using equation 6. Results are shown for tests performed at the 5% (left) and the 10% (right) levels of significance.

frost–free period was well replicated using method B. All methods well replicated the means of growing degree days for barley and the heat degree days, but a poor performance was obtained in reproducing their annual variability. The mean and variance of growing degree days for maize and chill units were well replicated by all methods.

Solar Radiation Estimation

Equation 6 was calibrated at the first–order station. The coefficients obtained from the linear fit were a = 0.23 and b = 0.51, with a correlation coefficient (R^2) = 0.91 and a standard error of the estimate (SEE) = 2.52 MJ m^{-2} (N = 2079).

Modified TAMSIM Method: The coefficients b, c, d, and e (eq. 5) were zero since the trigonometric functions did not contribute to improve the predicted values .The optimized a, f, g, h, and i parameters were a = 0.24, f = –0.08, g = –0.15, h = 0.02, and i = 0.035. The root mean square error obtained was 0.016 (unitless).

Donatelli and Campbell Method: The parameters calibrated using solar radiation from Lleida were the following: b = 0.056, Tnc = 8.00, and τ_c = 0.76. The root mean square error obtained was 4.71 MJ m^{-2}.

The two methods, with parameters calibrated at Lleida, were applied at Monte Julia, which was assumed to play the role of a surrounding thermo–pluviometric station, and compared with measured data. Figure 3 shows the actual and estimated means and standard deviations on a monthly basis. The TAMSIM method had eight rejections for the means at the 10% level of significance but well replicated February, April, May, and September. These last three months are important for agricultural purposes. However, this method was not able to replicate the interannual variability since all the standard deviations were too low (fig. 3, bottom).

The Donatelli and Campbell method was not able to reproduce the mean monthly values. However, it was able to replicate two standard deviations (March and June), and the interannual variability was more realistic (fig. 3, bottom).

Table 4. Comparison of cumulative probability distribution functions of hot and cold sequences from actual and generated values using three methods. Bold numbers indicate in which sequence the test was rejected. Only one representative sample out of 10 was used for illustration.

Length of Sequence (days)	Actual Data		Method A1		Method A2		Method B	
	Hot	Cold	Hot	Cold	Hot	Cold	Hot	Cold
1	0.294	0.445	**0.393**	0.484	0.362	0.456	0.396	0.468
2	0.474	0.622	0.578	0.689	**0.563**	0.696	**0.581**	0.671
3	0.583	0.725	0.689	**0.815**	0.672	0.808	0.691	0.799
4	0.677	0.804	0.763	0.879	0.752	0.839	0.763	0.875
5	0.732	0.849	0.811	0.920	0.792	0.900	0.806	0.913
6	0.784	0.908	0.829	0.947	0.818	0.937	0.827	0.946
7	0.813	0.930	0.854	0.957	0.848	0.960	0.852	0.953
8	0.826	0.938	0.885	0.970	0.876	0.965	0.882	0.969
9	0.836	0.950	0.901	0.977	0.895	0.977	0.901	0.976
10	0.867	0.961	0.928	0.991	0.921	0.988	0.928	0.986
11	0.883	0.972	0.934	0.995	0.927	0.993	0.934	0.991
12	0.904	0.975	0.942	0.995	0.940	0.993	0.942	0.991
13	0.924	0.980	0.957	0.998	0.953	0.998	0.957	0.995
14	0.927	0.980	0.963	0.998	0.959	0.998	0.963	0.995
15	0.938	0.980	0.963	0.998	0.961	0.998	0.963	0.995
16	0.943	0.986	0.965	0.998	0.966	1.000	0.965	0.995
17	0.951	0.992	0.967	0.998	0.968		0.967	0.995
18	0.953	0.992	0.971	0.998	0.974		0.971	0.998
19	0.958	0.992	0.973	0.998	0.974		0.973	0.998
20	0.958	0.992	0.973	0.998	0.974		0.973	0.998
25	0.974	0.994	0.986	0.998	0.983		0.986	0.998
30	0.984	1.000	0.994	1.000	0.9940		0.994	1.000
35	0.992		0.996		.996		0.996	
40	0.995		0.996		0.996		0.996	
45	0.995		0.998		0.998		0.998	
50	1.000		1.000		1.000		1.000	
Total number of sequences	384	357	486	438	526	428	467	428
Longest sequence (days)	47	28	46	28	47	28	46	16
Maximum difference (%)			10.6	9.0	8.9	8.3	10.8	7.4
$D_{0.05}$ (%)			9.3	9.7	9.1	9.7	9.4	9.7
$D_{0.1}$ (%)			8.3	8.7	8.1	8.7	8.4	8.7

Table 5. Percentage of Kolmogorov–Smirnoff tests rejected (at 5% and 10% significance levels) when comparing the distribution of hot and cold sequences from actual data and values generated using three methods.

	Hot		Cold	
	5%	10%	5%	10%
Method A1	60	80	90	100
Method A2	60	80	90	100
Method B	50	70	80	90

Overall, the methods analyzed to estimate solar radiation for an area surrounding a first–order station were not able to replicate the daily variability with confidence. On the other hand, the two methods can be used to fill missing data. This partially corroborates results obtained by other authors (Hunt et al., 1998; Meinke et al., 1995).

Table 6. Percent of t–tests and F–tests rejected at 5% and 10% levels of significance for frost–free period, growing degree days for maize and barley, chill units, and heating degree days when comparing indices calculated from actual and generated data using three methods.

	Frost–Free Period (days)		Growing Degree Days: Maize		Growing Degree Days: Barley		Chill Units		Heat Degree Days	
	5%	10%	5%	10%	5%	10%	5%	10%	5%	10%
Method A1										
Mean	80	100	0	0	0	0	10	10	0	0
Std	40	70	0	0	40	70	0	20	100	100
Method A2										
Mean	30	30	20	20	0	0	0	20	0	0
Std	40	50	10	20	40	60	0	10	90	100
Method B										
Mean	70	80	0	0	0	0	10	10	0	0
Std	10	10	0	0	70	80	10	10	100	100
Observed										
Mean	247.57		1463.08		2132.18		2504.52		2940.56	
Std	27.77		126.11		154.72		451.22		198.95	

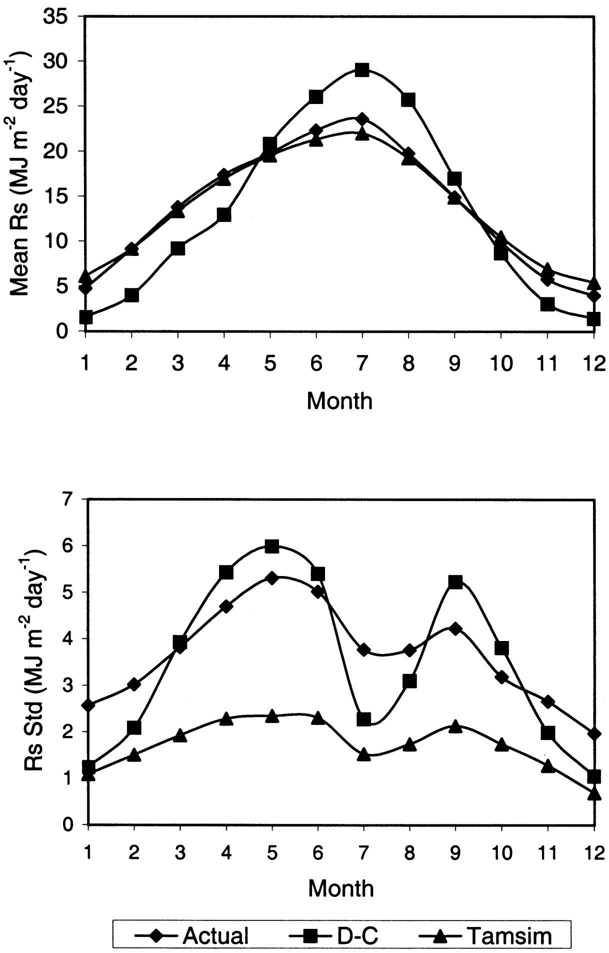

Figure 3. Comparison of solar radiation measured at Monte Julia (10 years of data) with estimates from the modified TAMSIM and Donatelli and Campbell methods after calibration using six years of data from Lleida: monthly means (top) and standard deviations (bottom).

Summary and Conclusions

The weather generator ClimGen was applied at Lleida (in northeast Spain, a semiarid continental Mediterranean climate) where the longest series of available data are precipitation, temperature, and bright sunshine hours. The Lleida station is surrounded by thermo–pluviometric stations. This scenario is common in other agricultural locations around the world. Using options in ClimGen that allow the implementation of alternative generation methods, the overall objective of this study was to investigate the best approach to generate weather data for agricultural use in the Lleida region. Three different methods were evaluated. Method A1 used maximum and minimum temperature and bright sunshine hours in the multivariable generation process. Method A2 used maximum temperature, thermal amplitude, and bright sunshine hours. Method B used only maximum and minimum temperature. To complement method B, two empirical equations, a modified TAMSIM model (Hunt et al., 1998) and the Donatelli and Campbell model (Donatelli and Campbell, 1998) were used to estimate solar radiation from temperature. In addition, the two–parameter gamma and Weibull distribution functions were tested to analyze their performance in generating precipitation amounts.

Using a first–order Markov chain for predicting wet days and the Weibull distribution for precipitation amounts resulted in good replication of the distribution function of the length of dry periods, the annual means of length of dry and wet periods, and the annual and monthly means of the amount of precipitation, showing a better performance than the Gamma distribution.

All three multivariate generation methods reproduced satisfactorily temperature means at the 5% level of significance. In general, for methods A1 and B, maximum temperature was more difficult to replicate than minimum temperature. Method A2 did not perform well for minimum temperature since the variance could not be reproduced for most months. The methods tested well replicated chill units, growing degree days, and the mean of heating degree days (but not its variance). Good performance was also obtained in replicating long hot or cold sequences at the 5% level of significance. Method A2 well replicated the mean frost–free period but not its variance. The other methods performed poorly for both mean and variance. None of the models evaluated can be recommended with confidence to analyze problems involving short–term freezing events.

Overall, for scenario A, method A1 is more recommendable than method A2 for temperature generation. For scenario B, method B is also recommended since in general it is comparable to the other methods and superior to method A2 in replicating the minimum temperature.

Method A1 replicated the bright sunshine hours better than method A2 did. Unfortunately, the worst performance was obtained during the summer (July was never successfully reproduced). For the rest of the year, the performance was good using method A1. The generation of bright sunshine hours is an alternative for locations where solar radiation data are not available to parameterize the weather generator.

Although method B generally performed well in predicting temperature, estimation of solar radiation using two models with parameters calibrated at a near first–order station was not adequate. The Donatelli and Campbell model is not recommended as a complement to method B because its parameters tend to be local. The parameters of the TAMSIM model are less sensitive to extrapolation to surrounding stations. On the other hand, the Donatelli and Campbell model gave a more realistic replication of the standard deviation than the modified TAMSIM model. Overall, to obtain a more complete and suitable weather generator for the Lleida region, a better method to estimate solar radiation is desirable.

References

Allen, R. G., M. Smith, L. S. Pereira, and A. Perrier. 1994. An update for the calculation of reference evapotranspiration. *ICID Bulletin* 43(2): 1–29.

Arnold, C. D., and W. J. Elliot. 1996. Cligen weather generator predictions of seasonal wet and dry spells in Uganda. *Trans. ASAE* 39(3): 969–972.

Bristow, K. L., and G. S. Campbell. 1984. On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agric. Forest Meteorology* 31: 150–166.

Castellvi, F., P. J. Perez, J. M. Villar, and J. I. Rosell. 1996. Analysis of methods for estimating vapor pressure deficit and relative humidity. *Agric. Forest Meteorology* 82: 29–45

Chineke, T. C., S. S. Jagtap, and J. I. Aina. 1999. Applicability of a weather simulation model based on observed daily meteorological data in humid tropical climate. *Theor. Appl. Climatology* 64: 15–25.

Donatelli, M., and G. S. Campbell. 1998. A simple model to estimate global solar radiation. In *Proc. 5th ESA Congress*, 133–134. Nitra, Slovak Republic. Wageningen, The Netherlands: European Society of Agronomy.

Gabriel, H. R., and J. Neumann. 1962. A Markov chain model for daily precipitation occurrence at Tel Aviv. *Q.J.R. Meteorol. Soc*. 88: 90–95.

Goffe, W. L., G. D. Ferrier, and J. Rogers. 1994. Global optimization of statistical functions with simulated annealing. *J. Econometrics* 60: 65–100.

Hunt, L. A., L. Kuchar, and C. J. Swanton. 1998. Estimation of solar radiation for use in crop modeling. *Agric. Forest Meteorology* 91: 293–300.

Johnson, G. L., C. L. Hanson, S. P. Hardegree, and E. B. Ballard. 1996. Stochastic weather simulation: Overview and analysis of two commonly used models. *J. Applied Meteorology* 35: 1878–1896.

Loague, K., and R. E. Gren. 1991. Statistical and graphical methods of evaluating solute transport models: Overview and application. *J. Contam. Hydrology* 7: 51–73.

Matalas, N. C. 1967. Mathematical assessment of synthetic hydrology. *Water Res. Research* 3(4): 937–945.

McCaskill, M. R. 1990. An efficient method for generation of full climatological records from daily precipitation. *Australian J. Agric. Res.* 41: 595–602.

Meinke, H., P. S. Camberry, M. R. McCaskill, M. A. Hills, and I. MacLeod. 1995. Evaluation of radiation and temperature data generators in the Australian tropics and sub–tropics using crop simulation models. *Agric. Forest Meteorology* 72: 295–316.

Monteith, J. L., and M. H. Unsworth. 1990. *Principles of Environmental Physics*. New York, N.Y.: Edward and Arnold.

Nicks, A. D., C. W. Richardson, and J. R. Williams. 1990. Evaluation of the EPIC model weather generator. In *Erosion / Productivity Impact Calculator*, 1. A. N. Sharpley and J. R. Williams, eds. Model Documentation. USDA–ARS Technical Bulletin 1768. Washington, D.C.: USDA Agricultural Research Service.

Peiris, D. R., and J. W. McNicol. 1996. Modeling daily weather with multivariate times series. *Agric. Forest Meteorology* 79: 219–231.

Richardson, C. W. 1985. Weather simulation for crop management models. *Trans. ASAE* 28(5): 1602–1606.

Richardson, C. W., and D. A. Wright. 1984. WGEN: A model for generating daily weather variables. ARS–8. Washington, D.C.: USDA Agricultural Research Service.

Selker, J. S., and D. A. Haith. 1990. Development and testing of single–parameter precipitation distributions. *Water Res. Research* 26: 2733–2740.

Vestal, C. K. 1971. First and last occurrences of low temperatures during the cold season. *Mon. Weather Rev.* 99: 650–652.

Wallis, T. W. R., and J. F. Griffiths. 1995. An assessment of the weather generator (WXGEN) used in the erosion/productivity impact calculator (EPIC). *Agric. Forest Meteorology* 73: 115–133.

Yin, X. 1999a. Bright sunshine duration in relation to precipitation, air temperature, and geographic location. *Theor. Appl. Climatology* 64: 61–68.

_____. 1999b. Evaluation of solar irradiance models with a special reference to globally–parameterized and land cover–sensitive solar 123. *Theor. Appl. Climatology* 64: 249–261.