

Evaluation of an improved daily solar radiation generator for the southeastern USA

Axel Garcia y Garcia*, Gerrit Hoogenboom

Department of Biological and Agricultural Engineering, The University of Georgia, 1109 Experiment Street, Griffin, Georgia 30223, USA

ABSTRACT: The availability of daily solar radiation (SR) data has been limited until recent years, restricting modeling applications when long-term data are required. The main goal of this study was to improve the WGENR (Weather Generator) SR generator and to evaluate its performance with recorded SR data for the southeastern USA. WGENR includes adjustment factors (AF) for obtaining proper average annual daily SR during dry (*RMD*) and wet (*RMW*) days and for amplification of its amplitude during dry and wet days. Thus, 3 sets of AFs, related to the occurrence of dry and wet days and precipitation greater than zero, were evaluated. Recorded SR data from 17 weather stations located in South Georgia, with records ranging from 1992 to 2003, were used for evaluation. The chi-square (χ^2) goodness-of-fit and the Kolmogorov-Smirnov (KS) tests were applied to analyze the frequency distributions and to determine significant differences between the cumulative distribution functions (CDFs) of generated and recorded SR. For most of the locations, generated SR differed from recorded data but improvements for *RMD* and *RMW* were obtained. Based on the χ^2 and the smallest difference between CDFs for most of the locations, a final set of AFs was selected that provided an overall improvement of the WGENR solar radiation generator. Further work will focus on the impact of these changes on decision support systems, specifically yield prediction with crop models using generated SR.

KEY WORDS: Weather generator · Automated weather stations · Decision Support System · Computer simulation model · Natural resource management

—Resale or republication not permitted without written consent of the publisher—

1. INTRODUCTION

Long-term daily weather data are required for most applications of decision support systems (DSS) in agricultural and natural resource management. Many of these systems have been used to determine the impact of climate change and climate variability on crop yield, water use and other environmental factors. However, most of the weather data that have been collected, such as those from the US National Weather Service's Cooperative Observer Network and similar national weather services, include only daily maximum and minimum temperature and total rainfall, which are often insufficient for agricultural or natural resource management studies (Hoogenboom 1996).

The availability of recorded daily solar radiation (SR) data has been limited to recent years, mainly due to

the high cost of the instrumentation, maintenance and calibration of pyranometers and other sensors (Elizondo et al. 1994). However, daily SR data have become more readily available with the introduction of automated weather stations (Hook & McClendon 1992). In the state of Georgia (USA), daily monitoring of SR was introduced with the development of the Georgia Automated Environmental Monitoring Network (GA-AEMN, www.georgiaweather.net). The first weather station was employed in 1991 (Hoogenboom 1996) and in 2004 there were 60 operational stations.

Automated weather stations (AWS) record air temperature, relative humidity, wind speed and direction, SR, precipitation, and soil temperature and other related meteorological and environmental variables. An automated station can avoid potential problems associated with a human observer, such as non-

*Email: agarcia@griffin.uga.edu

homogeneous data or site-to-site variability (Hubbard 1994). Thus, AWS are an excellent and appropriate source of data for modeling studies, especially related to the improvement and evaluation of algorithms that can estimate variables that have not been previously recorded, such as daily SR.

The prediction of SR has been explored in several studies that were based on a deterministic approach or regression analyses. The widely-used Bristow & Campbell (1984) method as well as the Hargreaves et al. (1985) method estimate daily SR as a function of the daily maximum and minimum air temperature and calculated daily extraterrestrial radiation. While the Bristow-Campbell method uses an exponential relationship, Hargreaves et al.'s method uses a simple linear regression. The Hook & McClendon (1992) approach estimates the daily SR from maximum and minimum air temperature, pan evaporation, precipitation, and calculated extraterrestrial radiation. Their method includes a random component to express the variability of the SR associated with the unmeasured degree of cloudiness.

Hunt et al. (1998) estimate the daily SR based on the daily maximum and minimum air temperature, precipitation, and calculated extraterrestrial radiation. When compared with 4 other models for a single location and 1 yr of data from Ontario, Canada, this model provided the best results. More recently, Donatelli et al. (2003) proposed a method, called RadEst3.00, for estimating daily SR as a function of air temperature and specific location information, including latitude, longitude, and elevation. RadEst3.00 uses the Bristow & Campbell (1984) approach as well as improvements that account for seasonality effects (Donatelli & Campbell 1998, Donatelli & Bellocchi 2001). Weiss & Hays (2004) proposed a generalized method, based on air temperature and precipitation, to estimate SR. The generalization was derived from empirical relationships of the bias error (difference between simulated and recorded SR) using data from several locations in the continental USA. This bias error, when associated with the day of the year, was linear during the early and later parts of the year and had a cosine shape during the middle of the year. Then, a location-specific linear and cosine correction factor was determined and introduced to their equation. Elizondo et al. (1994) used a neural network approach to estimate daily SR as a function of recorded weather variables, including daily precipitation, maximum and minimum air temperature, and calculated variables, including clear sky radiation and day length.

Other methods to obtain SR are through a stochastic or random process or through linear interpolation (Soltani et al. 2004). The stochastic method, usually called weather generator, is capable of producing one or more weather components with the same statistical

characteristics that naturally occur for a given location. However, the statistical parameters for the selected weather elements for a given time period must first be determined (Harmel et al. 2002). Several weather generators include SR as one of the parameters to be obtained, including Weather Generator (WGEN; Richardson 1981, Richardson & Wright 1984), Simulation of Meteorological Variables (SIMMETEO; Geng et al. 1988), Long Ashton Research Station-Weather Generator (LARS-WG; Racsco et al. 1991), MARKSIM (Jones & Thornton 2000), Disaggregated Variance Model (DVM; Hansen & Mavromatis 2001), and the Generation of weather Elements for Multiple applications (GEM; Harmel et al. 2002). Other generators have been developed that estimate a specific weather variable, such as the Weather Generator for Solar Radiation (WGENR). This is an adaptation of the solar radiation simulation model originally developed by Hodges et al. (1985) for crop growth and yield simulation models and is based on Richardson's (1981) WGEN algorithm.

The WGEN approach for generating maximum temperature, minimum temperature, and SR is based on a multivariate stochastic process following:

$$\chi_i = A \times \chi_{i-1} + B \times \epsilon \quad (1)$$

where χ_i and χ_{i-1} are (3×1) matrices for days i and $i - 1$ of a given year whose elements are residuals of maximum and minimum temperatures and SR; ϵ is a (3×1) matrix of independent random components that are normally distributed with a mean of zero and variance of unity; and A and B are (3×3) matrices whose elements are defined such that the new sequences have the desired serial correlation and cross-correlation coefficients. Then, the daily values of the 3 weather variables are found by multiplying the residuals by the standard deviation and adding the average. The average and the standard deviation are conditioned on the wet or dry status of the day determined by using the Markov chain model (Richardson 1981).

The WGENR approach uses daily recorded maximum and minimum air temperature and rainfall as inputs to estimate daily SR. The algorithm can be simplified as follows:

$$SR = SRL \times SRSD + SRBAR \quad (2)$$

where SR is the generated daily SR. SRBAR corresponds to annual curves of long-term average daily values with separate curves for dry and wet days. The wet days are defined as those days with rainfall greater than zero. SRBAR is defined as:

$$SRBAR = R_M + R_A \{\cos[0.0172(j - 172) + a]\} \quad (3)$$

where R_M and R_A are the respective annual average and amplitude solar radiation values for either dry or

wet days, $j = 1, 2, 3, \dots, 365$, $a = 0.0$ for the northern hemisphere and $a = 183$ for the southern hemisphere. SRL, a random component that was added, is a function of two 3×3 matrices, which were originally developed by Richardson (1981) to describe the cross-correlations between daily minimum and maximum temperatures and SR for the continental United States. Daily deviations estimated from the empirical cross-correlation were then multiplied by the standard deviation (SRSD) conditioned to wet or dry days. Thus, the main difference between WGEN (Richardson 1981) and WGENR (Hodges et al. 1985) algorithms is that Richardson's approach uses a set of location specific constants to estimate daily rainfall, SR, and maximum and minimum temperature, while Hodges et al. (1985) uses recorded maximum and minimum temperature and rainfall to generate SR.

An early evaluation of the model outputs indicated that the daily variability of the SR was too small (Hodges et al. 1985). Therefore, a non-random factor (RNOISE) for wet or dry days was introduced to amplify the daily residuals for deviations above and below the annual curves:

$$SR = RNOISE \times SRL \times SRSD + SRBAR \quad (4)$$

where RNOISE = 4.4 for SRL > 0.0 on dry days, 11.4 for SRL > 0.0 on wet days, 13.2 for SRL ≤ 0.0 on dry days, and 34.32 for SRL ≤ 0.0 on wet days.

In addition to this early evaluation, the WGENR has also been used for climate change studies (Cooter 1990, Smith & Tirpack 1990) and for evaluating its performance to predict daily SR from interpolated climate records (Grant et al. 2004). However, its performance has not been evaluated with actual recorded daily SR from automated weather station networks for multiple locations. The most recent study to improve WGENR for use in biological applications was conducted by Cooter & Dhakhwa (1995). They used the Solar and Meteorological Surface Observation (SAMSON), a combination of generated and recorded SR data, to re-estimate the WGENR parameters. The meteorological-statistical model (METSTAT) developed by the National SR Data Base (NSRDB) project through the National Renewable Energy Laboratory (NREL 1992), Golden, Colorado, USA, was used to generate part of the SR included in the SAMSON data base (NREL 1992). Then the climatologic variability in radiation during dry and wet days, associated with cloud cover, through monthly adjustment factors, was geographically stratified. The impact of these improvements on corn yield was evaluated for only one location in Raleigh, North Carolina, USA, using the Erosion Productivity Impact Calculator (EPIC) model (Williams et al. 1989). They found that the time series of the simulated yield was statistically similar to the time series

of the recorded yield, obtaining 7873.6 kg ha⁻¹ and 7235.2 kg ha⁻¹ for simulated long-term average yield using the original and the improved Hodges et al. (1985) SR generator, respectively, against the 7336.3 kg ha⁻¹ of the long-term average observed yield (Cooter & Dhakhwa 1995).

The modification of empirical parameters in existing algorithms is a common practice for adapting models to specific conditions or for resolving specific problems. The generalization of these parameters should result in a perfect model that, theoretically, will generate data identical to those recorded. However, the complexity of the interactions associated with the natural phenomena makes it extremely difficult to create a perfect model at the stage of the current knowledge. Therefore, empirical parameters are used to obtain proper estimates at a specific degree of probability through a statistical test. Thus, the modification of empirical parameters in weather models has been used successfully in several studies. As an example, Wilks (1992) adapted Richardson's (1981) algorithm to generate a synthetic daily time series consistent with assumed future climate. Cooter & Dhakhwa (1995) re-estimated the parameters of the Hodges et al. (1985) approach to improve the SR during dry and wet days. Goodin et al. (1999) re-estimated the empirical coefficients of the Bristow & Campbell (1984) method using data from Manhattan (Kansas); this resulted in a greater improved SR estimates than the unmodified version when compared to the recorded values.

The WGENR was selected for this study because of the simplicity of its algorithm. It uses recorded maximum and minimum air temperature and rainfall for generating SR, variables which are normally available for recorded long-term weather data. WGENR was also selected because of its potential for use with crop model applications (Hodges et al. 1985). The main goal of this study was to improve the WGENR SR generator and to evaluate its performance with recorded SR data from the southeastern USA.

2. MATERIALS AND METHODS

2.1. Recorded weather data

Daily records from 17 automatic weather stations (GA-AEMN) with 16 stations located in Georgia and one in Florida were used. In the region that was selected agriculture is the dominant economic sector (Fig. 1, Table 1). The weather stations had complete records for daily SR, maximum and minimum temperature, and rainfall, for a combined total of 124 yr. The historical data for each location varied from 3 to 12 yr, covering the period from 1992 to 2003. The GA-AEMN

uses a LI-200SZ pyranometer sensor to measure SR with an accuracy of $\pm 5\%$ W m^{-2} (LI-COR 1991). Values equal or close to zero were removed from the recorded data sets, which accounted for only 0.043% of the total data available for the 17 locations. Then,

the recorded data sets were duplicated and the daily SR was removed from one set for which SR would be generated.

2.2. Selection of the adjustment factors

The average annual daily SR during dry (*RMD*) and wet days (*RMW*) obtained from the original WGENR algorithm were reported to be 5 and 15% greater than those from the recorded data (Hodges et al. 1985). Therefore, a set of adjustment factors (AF) was included in the algorithm to estimate proper values for *RMD* and *RMW*. These AFs, defined here as (F_D) and (F_W) for dry and wet days respectively, have an effect on the deviation from the annual curves (SRBAR), as well as on the RNOISE for precipitation greater than zero ($PP > 0$) defined here as F_{PP} .

The first set of AFs corresponded to the original values of WGENR ($F_D = 0.06$, $F_W = 0.19$, and $F_{PP} = 2.60$); it appears to have been determined by the difference between recorded and generated SR using central tendency descriptive statistics. The algorithm containing these values is referred to as WGENR00 in the remainder of this study. The second set of AFs, i.e. $F_D = -0.05$, $F_W = 0.35$, and $F_{PP} = 1.80$, was obtained by analyzing different combinations of F_D , F_W , and F_{PP} , e.g. $-0.10 < F_D < 0.10$, $0 < F_W < 0.40$, and $1.55 < F_{PP} < 2.60$, using 1 yr (1995) of recorded data (Camilla, Mitchell County, Georgia) in a response surface until obtaining values for *RMW* that were close to the recorded data. Then, the AFs were derived by fitting Eqs. (5) & (6) to the F_D , F_W , *RMD* or F_D , F_W , *RMW* data sets:

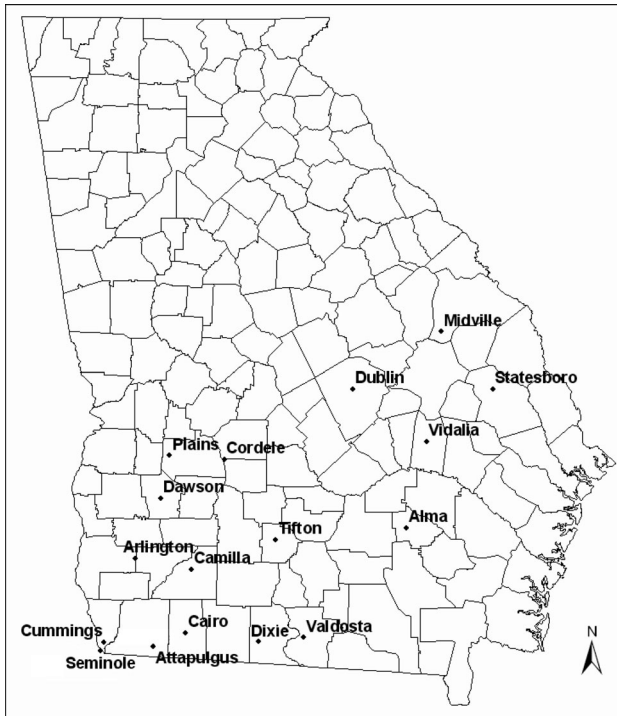


Fig. 1. Spatial distribution of the weather stations in Georgia and Florida, USA, used in this study

Table 1. Annual average climate parameters and other site information for the selected locations. SR: solar radiation; T_{\max} : maximum air temperature; T_{\min} : minimum air temperature; E: elevation above sea level

Location/County/State	SR ($\text{MJ m}^{-2} \text{d}^{-1}$)	T_{\max} ($^{\circ}\text{C}$)	T_{\min} ($^{\circ}\text{C}$)	Total rain (mm)	Dry days (%)	Wet days (%)	Lat. ($^{\circ}$)	Long. ($^{\circ}$)	Period of record (yr)	E (m)
Alma, Bacon, GA	14.5	25.61	12.95	1190	70	30	31.56	-82.51	10	65
Arlington, Baker, GA	16.2	25.82	11.87	1206	69	31	31.35	-84.63	7	67
Attapulgus, Decatur, GA	15.7	25.96	12.13	1238	68	32	30.76	-84.48	10	83
Cairo, Grady, GA	16.2	25.85	12.38	1082	70	30	30.85	-84.23	6	77
Camilla, Mitchell, GA	16.4	25.94	13.13	1125	70	30	31.28	-84.19	6	56
Cordele, Crisp, GA	16.6	25.20	12.88	986	70	30	32.02	-83.94	6	87
Cummings, Seminole, GA	16.7	25.44	15.23	1380	70	30	30.78	-84.87	3	25
Dawson, Terrell, GA	16.5	25.02	11.60	1146	69	31	31.75	-84.43	8	102
Dixie, Brooks, GA	16.9	26.08	13.43	1160	70	30	30.79	-83.66	5	61
Dublin, Laurens, GA	16.5	25.05	11.72	1045	70	30	32.49	-82.92	6	72
Midville, Burke, GA	16.1	24.41	12.02	1066	69	31	32.87	-82.21	12	81
Plains, Sumter, GA	15.8	24.20	11.61	1189	69	31	32.04	-84.37	11	162
Seminole, Jackson, FL	16.4	25.36	15.45	1357	68	32	30.72	-84.89	3	25
Statesboro, Bulloch, GA	15.5	24.65	11.41	1038	68	32	32.46	-81.76	8	64
Tifton, Tift, GA	16.5	24.78	13.55	1061	69	31	31.48	-83.53	11	113
Valdosta, Lowndes, GA	16.3	26.23	14.31	1009	70	30	30.82	-83.31	6	65
Vidalia, Toombs, GA	16.3	25.15	13.25	1080	70	30	32.14	-82.34	6	78

$$Z_D = Z_{0D} + a_D F_D + b_D F_W \quad (5)$$

$$Z_W = Z_{0W} + a_W F_D + b_W F_W \quad (6)$$

where Z_D and Z_W correspond to either *RMD* or *RMW*, and Z_{0D} and Z_{0W} , a_D and a_W , and b_D and b_W are empirical parameters for wet and dry days, respectively. Eqs. (5) & (6) were solved by simultaneous procedure, replacing Z_D and Z_W with the *RMD* and *RMW* values obtained from the 1995 recorded data. The AFs that provided the minimum simple difference (d) between recorded and generated *RMA* (average annual daily SR), *RMD*, and *RMW* were selected. The algorithm containing these values is referred to as WGENR01 in the remainder of this study. The third set of AFs was based on an arbitrary reduction of the original F_{pp} , from 2.60 to 2.00 ($F_D = 0.06$, $F_W = 0.19$, and $F_{pp} = 2.00$). This arbitrary value was used to evaluate the impact of changes on the daily SR if only the precipitation component is considered, as a control for local adjustments needs on F_{pp} . The algorithm containing these values is referred to as WGENR02 in the remainder of this study.

Thus, there were 4 data sets, one that included recorded data and 3 that included estimated data based on WGENR using 3 different AFs, e.g. WGENR00, WGENR01, and WGENR02. For the analysis, each data set was split into wet and dry days.

2.3. Statistical analysis of recorded and generated solar radiation

The generated daily SR data sets were analyzed by comparing their frequency distribution during dry and wet days with the frequency distribution of the recorded data sets. For each location, the chi-square goodness-of-fit (χ^2) at a 0.05 level of significance was used for this analysis. In addition, the daily recorded and generated data sets for each location during dry and wet days were compared through the cumulative distribution functions (CDFs). The Kolmogorov-Smirnov (KS) 2-sample, 2-sided test was used to compare the CDFs of recorded and generated daily SR. The KS test detects the largest difference that might exist between 2 distribution functions based on a statistic, called the D statistic. This statistic is a measure of the discrepancy between the empirical distribution and the hypothesized distribution:

$$D = \text{Max}_y |F_n(y) - F(Y)| \quad (7)$$

where $F_n(y)$ is the empirical cumulative distribution function and $F(Y)$ is the hypothesized cumulative distribution function. The D statistic is the maximum vertical distance between the 2 distribution functions. If D is sufficiently large, then the null hypothesis (identical

distribution) can be rejected. The smaller the D statistic, the smaller the difference between the 2 distributions at a given probability level (p-value). The statistical analyses were performed using the SAS Analyst and the SAS Insight procedure (SAS Institute 1999).

3. RESULTS

3.1. Analysis of recorded weather data

The recorded data for all locations and periods that were considered had similar values for the annual average SR, *RMA*, *RMD*, and *RMW* except for Alma (Bacon County, Georgia) (Table 1 and Fig. 2). While for most of the locations the *RMA* was slightly lower than $16 \text{ MJ m}^{-2} \text{ d}^{-1}$, for Alma it was $14.5 \text{ MJ m}^{-2} \text{ d}^{-1}$ (Table 1) and, while for most locations *RMD* and *RMW* were greater than $17 \text{ MJ m}^{-2} \text{ d}^{-1}$ and $12 \text{ MJ m}^{-2} \text{ d}^{-1}$, respectively, for Alma the *RMD* was slightly lower than $16 \text{ MJ m}^{-2} \text{ d}^{-1}$ and *RMW* was slightly greater than $11 \text{ MJ m}^{-2} \text{ d}^{-1}$ (Fig. 2). Since the maximum and minimum air temperatures (Table 1) and rainfall distribution for Alma were similar to the other locations (Fig. 3) and instrument siting was also the same, we could not find any cause for this difference.

3.2. Analysis of the adjustment factors

The impact of the AFs on generated SR using the 1995 data from Camilla showed that the lower the F_D values, the lower the values for *RMA*, with values for *RMA* close to recorded data for F_D values around zero (Fig. 4). This is due to the fact that during dry days SR is expected to be above the average (Fig. 2); for Georgia, in general, there are 2 dry days for every rainy day

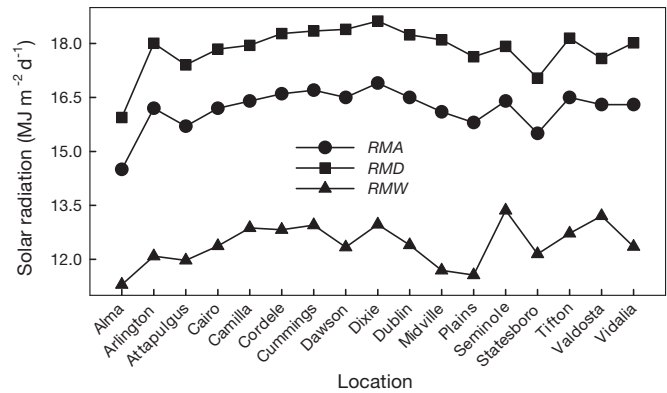


Fig. 2. Average annual daily recorded solar radiation (*RMA*) and average annual daily recorded solar radiation during dry (*RMD*) and wet (*RMW*) days for each location

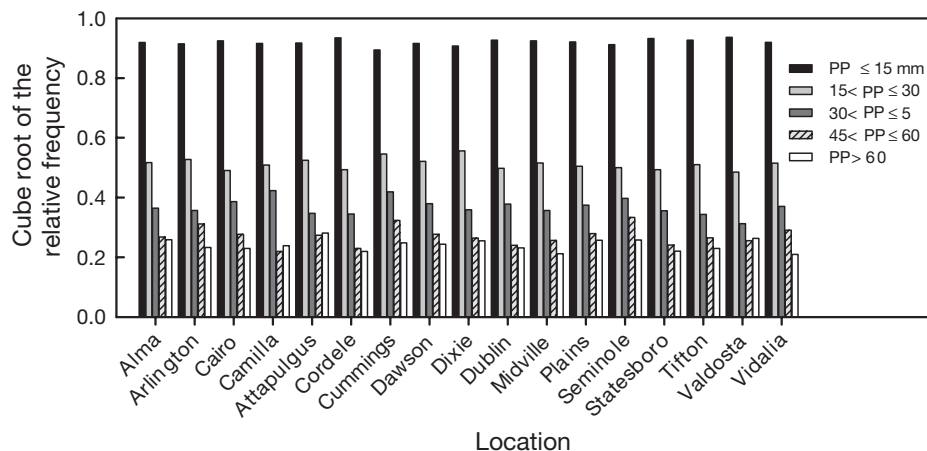


Fig. 3. Relative frequency of daily rainfall (PP) for each location

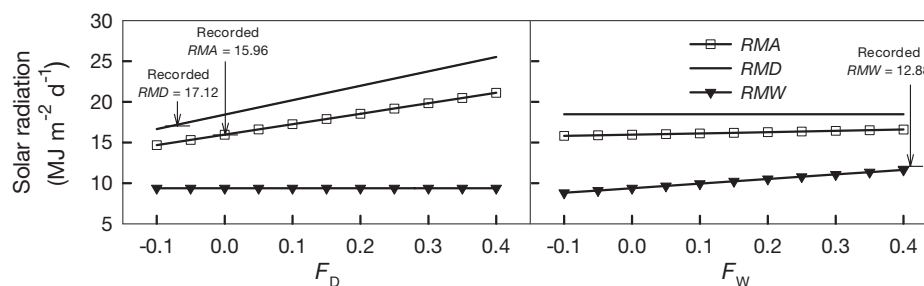


Fig. 4. Variation of the average annual daily solar radiation (RMA) and average annual daily solar radiation during dry (RMD) and wet (RMW) as a function of the adjustment coefficients F_D ($F_W = 0$) and F_W ($F_D = 0$)

(Table 1). On the other hand, the lower the F_W values the lower the estimated RMW . To obtain values for RMW that were close to the recorded data, it was necessary to set the values for F_W as high as 0.40. Beyond this value, RMW was overestimated. However, this set of coefficients did not provide adequate estimates for RMW , as the differences between the RMD and RMW from generated and recorded data were not the minimum (Fig. 5, AF number 37). Therefore, new values were assumed for F_{pp} , another response surface was obtained (Fig. 6), and Eqs. (4) & (5) were fitted again. A value of -0.05 for F_D and 1.8 for F_{pp} resulted in 0.35 as being the best value for F_W (Fig. 5, AF number 24). The use of this set of AFs (WGENR01) resulted in estimates for RMA and RMD that were slightly higher than those based on the recorded data. However, considerable improvements were obtained for RMW . The AFs for solar radiation during dry and wet days and the RNOISE for precipitation greater than zero ($PP > 0$) obtained in this study were in the range of the findings by Cooter & Dhakhwa (1995).

The use of the WGENR algorithm without AFs provided an adequate RMA with a d value of $0.013 \text{ MJ m}^{-2} \text{ d}^{-1}$. However, the d for the $RMDs$ was overestimated at

$1.35 \text{ MJ m}^{-2} \text{ d}^{-1}$, while the d for the $RMWs$ was underestimated at $3.5 \text{ MJ m}^{-2} \text{ d}^{-1}$. Also, several combinations of AFs provided an accurate RMA when compared to the RMA based on recorded data, suggesting that the WGENR algorithm is a robust tool for generating SR. Nevertheless, the $RMDs$ from the generated data were higher than those from the recorded data set, with d ranging from $-0.07 \text{ MJ m}^{-2} \text{ d}^{-1}$ to as high as $3.0 \text{ MJ m}^{-2} \text{ d}^{-1}$. In contrast, the RMW for the generated data was much lower than the RMW for the recorded data, with d ranging from -3.50 to $-0.46 \text{ MJ m}^{-2} \text{ d}^{-1}$. The original AFs (WGENR00) showed an overestimation of the RMA and the RMD of approximately 1 and $2.5 \text{ MJ m}^{-2} \text{ d}^{-1}$, respectively. However, a slight improvement was observed for the RMW (Fig. 5).

Although RMA and RMD were slightly overestimated when compared to the data obtained with the WGENR without adjustment, a great improvement was obtained with the AF set 24 which was selected as the 'best' because of its minimum d when compared to the RMW from the recorded data as well as a considerable reduction on the difference between the RMD from recorded and simulated data sets (Fig. 4). The adjustment factors evaluated in this study are summarized in Table 2.

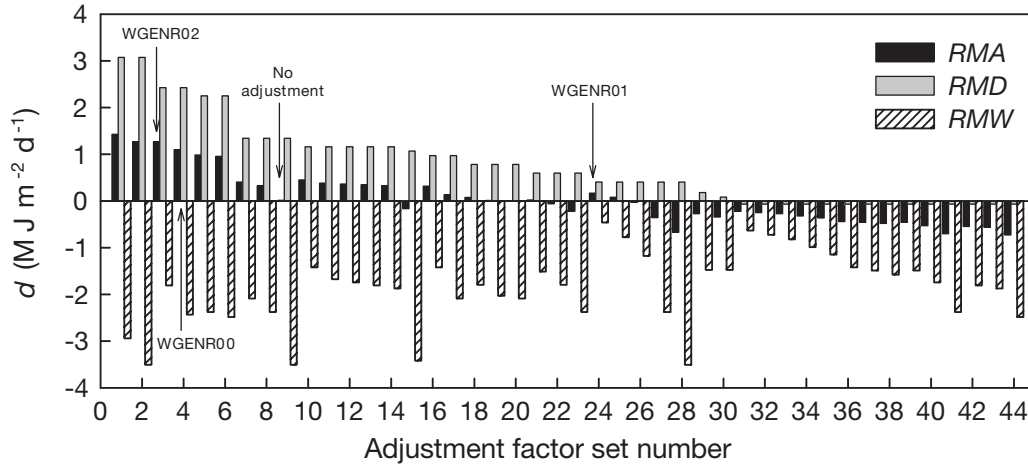


Fig. 5. Simple difference (d) between recorded and generated average annual daily solar radiation (RMA) and average annual daily solar radiation during dry (RMD) and wet (RMW) days for different combinations of F_D , F_W , and F_{PP}

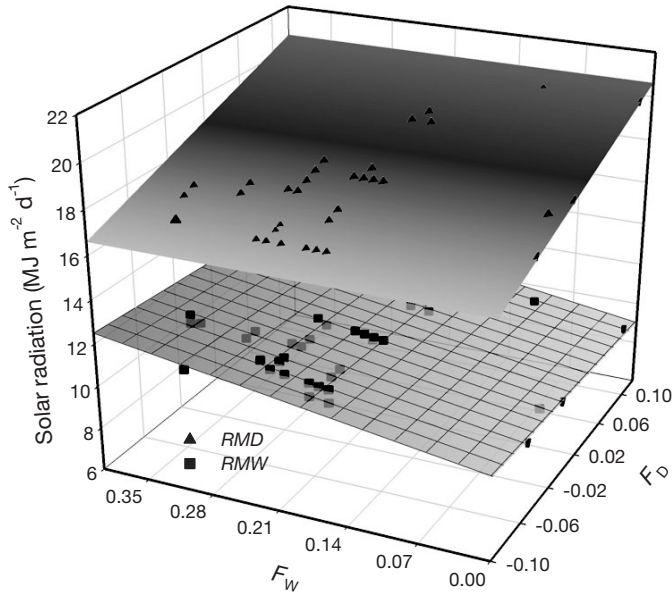


Fig. 6. Response surface of daily solar radiation as a function of F_D and F_W , the adjustment coefficients for the average annual daily solar radiation during dry (RMD) and wet (RMW) days

3.3. Evaluation of the recorded and generated daily solar radiation

For evaluation of the 3 different implementations (WGENR00, WGENR01, and WGENR02) of the WGENR algorithm, the recorded data from 17 weather stations (Table 1) located in south Georgia were used. These implementations provided results for the estimated SR that showed a similar pattern for RMD and RMW for each location when compared to those estimated from the recorded daily SR data (Fig. 7). Despite this similar pattern, the outputs of WGENR00 and

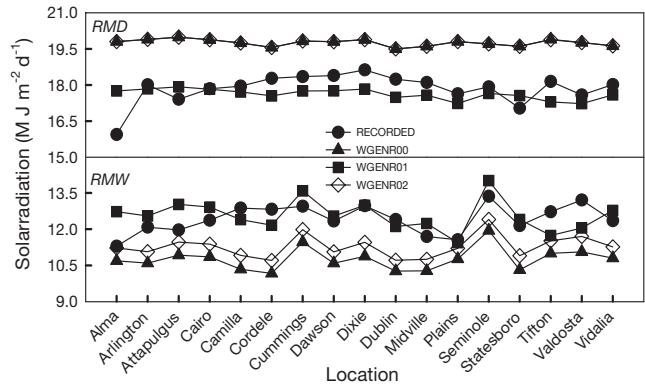


Fig. 7. Recorded and generated average annual daily solar radiation during dry (RMD) and wet (RMW) days for each location

WGENR02 overestimated RMD and underestimated RMW , while the results from the WGENR01 option were closely related to the recorded data for both dry and wet days. For WGENR00 and WGENR02, the value for d between recorded and generated RMD ranged from 1.26 to 3.9 $MJ\ m^{-2}\ d^{-1}$, while for WGENR01 the value for d between recorded and generated RMD ranged from 0 to 1.85 $MJ\ m^{-2}\ d^{-1}$. For WGENR00, the value for d between recorded and gen-

Table 2. Adjustment factors (AF) used for evaluation of WGENR. F_D , F_W , F_{PP} : adjustment factors for dry days, wet days and precipitation, respectively

Generator	F_D	F_W	F_{PP}
WGENR00	0.06	0.19	2.60
WGENR01	-0.05	0.35	1.80
WGENR02	0.06	0.19	2.00

erated *RMW* ranged from 0.57 to 2.65 MJ m⁻² d⁻¹, for WGENR01 from 0.03 to 1.43 MJ m⁻² d⁻¹, and for WGENR02 from 0.05 to 2.12 MJ m⁻² d⁻¹. These results showed that the values for the AFs introduced in this option and the importance of considering *F_{PP}* were adequate for improvements in the daily SR, mainly during wet days, and consequently for improvements of *RMD* and *RMW* (Fig. 7). These results are in agreement with the results obtained by Cooter & Dhakhwa (1995), who developed monthly adjustment factors that

were introduced in the WGENR algorithms. They obtained an improved performance of the model for rainy days in the humid and near-coastal regions of the southern USA.

The frequency distribution of the estimated daily SR for both dry and wet days was similar to the frequency distribution of the recorded data. Persistent lower estimations were observed around the recorded values of 8 and 16 MJ m⁻² d⁻¹ as well as overestimations for values higher than 24 MJ m⁻² d⁻¹. During dry days the 3

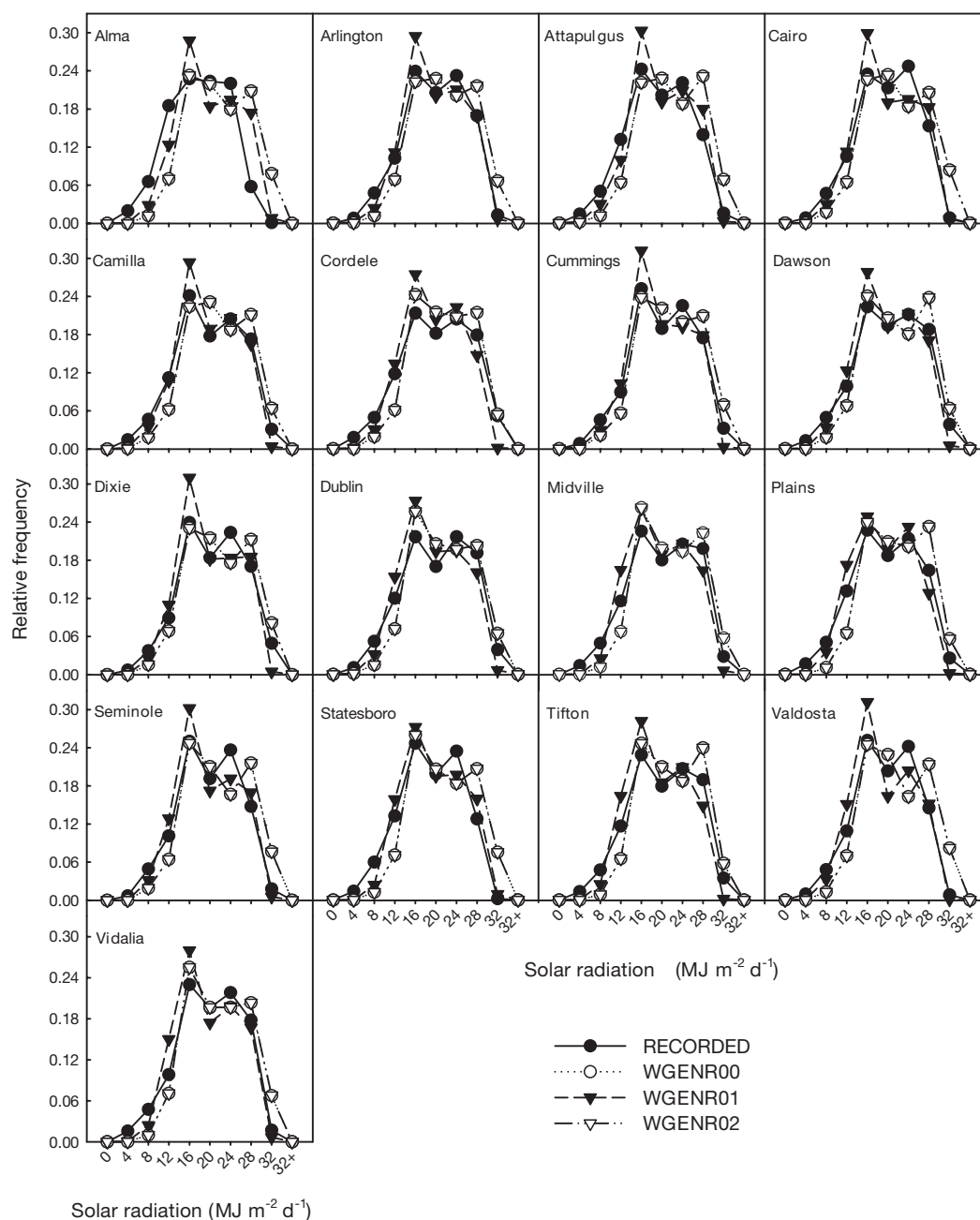


Fig. 8. Frequency distribution of recorded and generated daily solar radiation during dry days for each location

generators overestimated the daily SR for observations between 14 to 24 MJ m⁻² d⁻¹, especially for WGENR01, while an overestimation for recorded data above 24 MJ m⁻² d⁻¹ was more frequent for the WGENR00 and WGENR02 options (Fig. 8). Similar discrepancies were observed for SR generated during wet days. The 3 generators overestimated SR for recorded data between 4 and 12 MJ m⁻² d⁻¹ and there was a persistent overestimation at a recorded SR of 8 MJ m⁻² d⁻¹, especially for WGENR00 and WGENR02. An overestimation was

more frequent for the WGENR01 option for recorded data between 20 and 28 MJ m⁻² d⁻¹ (Fig. 9). Nevertheless, these results demonstrate that the AF modifications for both dry and wet days resulted in an improvement of *RMD* and *RMW*, especially when WGENR01 was used.

Based on the χ^2 , WGENR00 showed a poor performance in generating daily SR across the 17 selected locations: for more than 80% of the locations the frequency distribution between recorded and generated

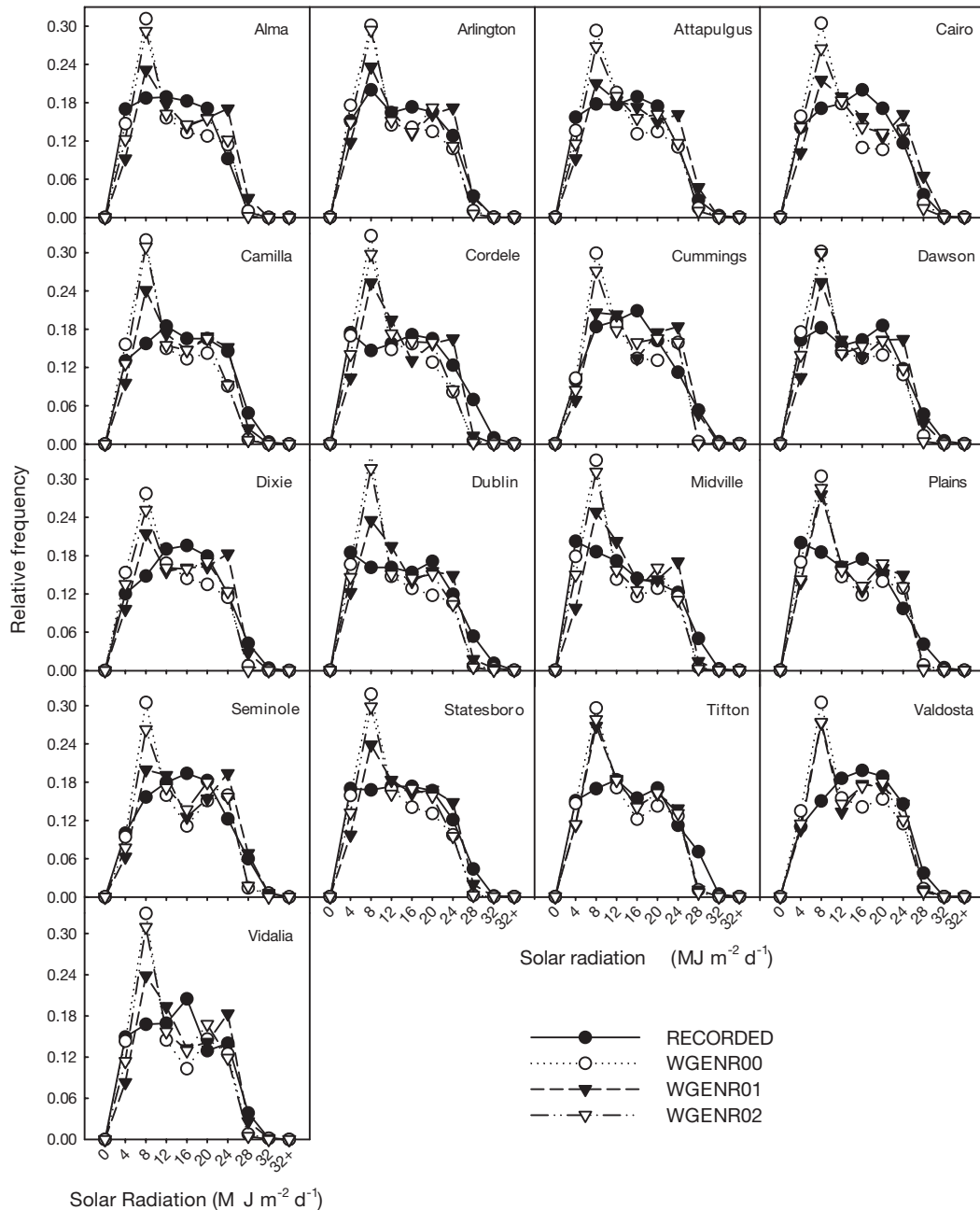


Fig. 9. Frequency distribution of recorded and generated daily solar radiation during wet days for each location

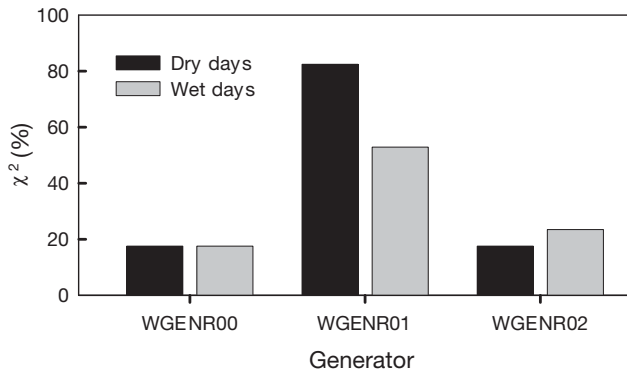


Fig. 10. Percentage chi-square goodness-of-fit (χ^2) of generated solar radiation distributions that were statistically significant ($\alpha = 0.05$) when compared with observed solar radiation distributions

data sets during dry and wet days was statistically different. The frequency distribution for the daily SR generated with WGENR02 was statistically different in 76 and 82% of the locations for dry and wet days, respectively. Nonetheless, the daily SR data generated with WGENR01 had a frequency distribution that was statistically different from the recorded data for only 47 and 18% of the locations during dry and wet days (Fig. 10).

The Kolmogorov-Smirnov test is an indication of any potential difference between the CDFs of generated and recorded SR data sets. During dry days, the D values ranged from 0.120 to 0.245 for WGENR00 and WGENR02 and from 0.042 to 0.136 for WGENR01.

During wet days, the D values ranged from 0.110 to 0.202 for WGENR00, from 0.071 to 0.124 for WGENR01, and from 0.097 to 0.155 for WGENR02. However, this statistical analysis showed that for most of the locations, the CDFs of the SR generated by the 3 WGENR implementations for both dry and wet days were significantly different from the recorded SR. The poorest performance was observed for the data obtained from WGENR00, with statistically significant differences for all locations for both dry and wet days. A slightly better performance was found for the WGENR02 option, mainly during wet days. Meanwhile, the data generated with the WGENR01 option had the smallest difference, e.g. the lowest value for D , from the recorded data as well as no statistical differences for several locations for both, dry and wet days (Fig. 11). The statistical differences found between CDFs from recorded and generated daily SR are not uncommon. Studies conducted by Meinke et al. (1995) and Soltani & Hoogenboom (2003) also reported that stochastic weather generators do not adequately reproduce the empirical distribution of the daily SR.

Even if most of the CDFs for the SR generated with the WGENR01 option were significantly different when compared to the CDFs from the recorded data, 5 of the 17 selected locations; including Arlington, Cairo, Camilla, Cummings, and Seminole for dry days and Camilla, Cummings, Dixie, Seminole, and Vidalia for wet days, had evidence of no statistically significant discrepancies between the CDFs. Also, no statistical significant discrepancies were observed for the comparison of the CDF of recorded and generated SR with

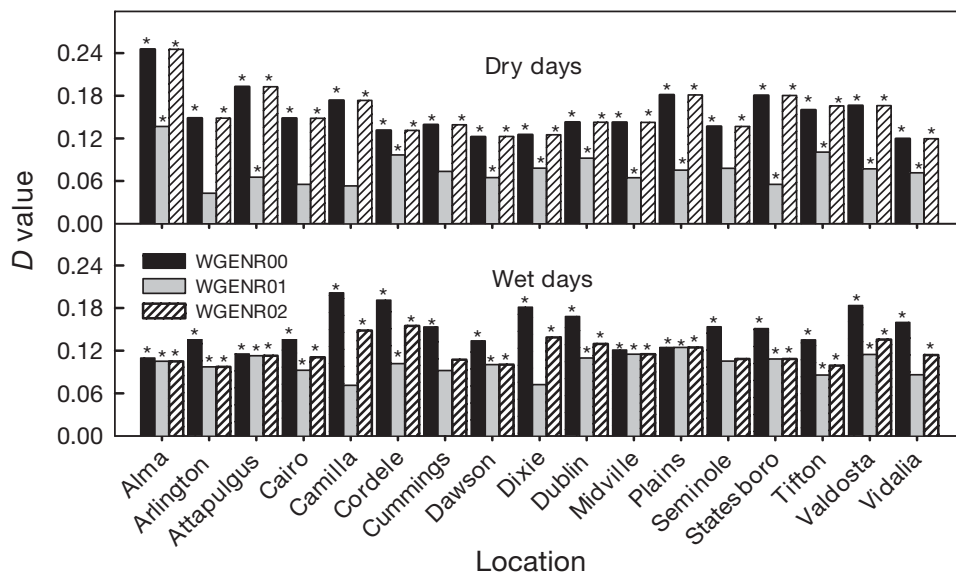


Fig. 11. Kolmogorov-Smirnov D test values for daily solar radiation during dry and wet days. Values with * have generated data distributions that differ significantly from distribution of recorded data at $p < 0.05$

the WGENR02 option for 2 locations, i.e. Cummings and Seminole, for wet days. Because the SR generated with the WGENR01 option differed the least from recorded for all locations for dry days and for most of the locations for wet days, e.g. it had the lowest D value, these results suggest a slightly better stability of the WGENR01 generator for the selected locations when compared to the performance of WGENR00 and WGENR02.

4. DISCUSSION AND CONCLUSIONS

Adjustment factors were determined to improve the daily SR generated with WGENR for the southeastern USA and other regions that have similar climatic characteristics. The daily SR generated with the proposed AFs for the WGENR provided better results compared to the existing method. Moreover, the estimate of the goodness-of-fit of the frequency of distributions for the 17 locations provided an adequate confidence for the daily SR generated through the AFs that were ultimately selected.

The modifications of the AFs resulted in improvements of the SR generated with WGENR, providing a more accurate estimation of daily SR during both dry and wet days. These results are encouraging for crop model applications, especially for studies that require long-term weather data, which normally include minimum and maximum air temperature and precipitation, and require SR estimations. A modification of the RNOISE parameter, a non-random factor associated with the occurrence of SR during wet days, was necessary in order to obtain daily SR values that were similar to those obtained from recorded data.

The comparison of the frequency distribution between recorded and generated daily SR during dry and wet days provided an adequate performance of the WGENR01 option throughout the 17 locations, indicating a better performance of this option for generating SR compared to WGENR00 and WGENR02. Although for most of the locations the generated daily SR differed statistically from the recorded data, the F_D , F_W , and F_{PP} used in the WGENR01 option provided the most accurate daily SR when compared to the recorded data. Moreover, when this option was used, the smallest difference through the Kolmogorov-Smirnov test, as well as the best values for the average annual daily SR during dry (RMD) and wet (RMW) days were obtained. Thus, values of -0.05 for F_D , 0.35 for F_W , and 1.80 for F_{PP} were selected as the best adjustment factors for daily generation of SR during dry and wet days. Additional studies are needed to determine the impact of using the daily SR data based on different implementations of the WGENR algorithm. These

studies should focus on the prediction of yield and other agronomic variables with crop simulation models as well as the impact of climate change and variability on agricultural production.

Acknowledgements. This work was conducted under the auspices of the Southeast Climate Consortium (SECC; www.secc.coaps.fsu.edu) and supported by a partnership with the United States Department of Agriculture-Risk Management Agency (USDA-RMA), by grants from the US National Oceanic and Atmospheric Administration-Office of Global Programs (NOAA-OGP) and USDA-Cooperative State Research, Education and Extension Services (USDA-CSREES) and by State and Federal funds allocated to Georgia Agricultural Experiment Stations Hatch project GEO00895. The authors would like to thank Mr. J. Davis for assistance with the statistical analysis.

LITERATURE CITED

- Bristow CL, Campbell GS (1984) On the relationship between incoming solar radiation and daily maximum and minimum temperature. *Agric For Meteorol* 31:159–166
- Cooter EJ (1990) The impact of climate change on continuous corn production in the southern USA. *Clim Change* 16: 53–82
- Cooter EJ, Dhakhwa GB (1995) A solar radiation model for use in biological applications in the south and southeastern USA. *Agric For Meteorol* 78:31–51
- Donatelli M, Bellocchi G (2001) Estimate of daily global solar radiation: new developments in the software RadEst3.00. In: *Proceedings of the 2nd International Symposium Modelling Cropping Systems*, 16–18 July. Florence, Italy, p 213–214
- Donatelli M, Bellocchi G, Fontana F (2003) RadEst3.0: software to estimate daily radiation data from commonly available meteorological variables. *Eur J Agron* 18:363–367
- Donatelli M, Campbell GS (1998) A simple model to estimate global solar radiation. p. 133–134. In: *Proceedings of the 5th European Society of Agronomy Congress*, 28 June–2 July 1998, Nitra, Slovak Republic. The Slovak Agricultural University, Nitra
- Elizondo D, Hoogenboom G, McClendon RW (1994) Development of a neural network model to predict daily solar radiation. *Agric For Meteorol* 71:115–132
- Geng S, Auburn JS, Brandstetter E, Li B (1988) A program to simulate meteorological variables: documentation for SIMMETEO. *Agronomy Progress Rep* 204, Department of Agronomy and Range Science, University of California, Davis, CA
- Goodin DG, Hutchinson JMS, Vanderlip, RL, Knapp MC (1999) Estimating solar irradiance for crop modeling using daily air temperature data. *Agron J* 91:845–851
- Grant RH, Hollinger SE, Hubbard KG, Hoogenboom G, Vanderlip RL (2004) Ability to predict daily solar radiation values from interpolated climate records for use in crop simulation models. *Agric For Meteorol* 127:65–75
- Hansen JW, Mavromatis T (2001) Correcting low-frequency bias in stochastic weather generators. *Agric For Meteorol* 109:297–310
- Hargreaves GL, Hargreaves GH, Riley JP (1985) Irrigation water requirement for Senegal River basin. *J Irrig Drain Eng* 111:265–275

- Harmel RD, Johnson G, Richardson CW (2002) The GEM experience: weather generator technology development in the USDA. *Bull Am Meteorol Soc* 83:954–957
- Hodges T, French V, LeDuc SK (1985) Estimating solar radiation for plant simulation models. *Agricultural and Resources Inventory Surveys Through Aerospace Remote Sensing*. AgRISTARS Tech Rep JSC-20239; YM-15-00403
- Hoogenboom G (1996) The Georgia Automated Environmental Monitoring Network. In: *Proceedings of the 22th Agricultural and Forest Meteorology Conference*, Atlanta, Georgia. *Am Meteorol Soc*, p 343–346
- Hook JE, McClendon RW (1992) Estimation of solar radiation data missing from long-term meteorological records. *Agron J* 84:739–742
- Hubbard KG (1994) Spatial variability of daily weather variables in the high plains of the USA. *Agric For Meteorol* 68: 29–41
- Hunt LA, Kuchar L, Swanton CJ (1998) Estimation of solar radiation for use in crop modeling. *Agric For Meteorol* 91: 293–300
- Jones PG, Thornton PK (2000) Marksim: software to generate daily weather data for Latin America and Africa. *Agron J* 92:445–453
- LI-COR (1991) *Terrestrial Radiation Sensors, type SZ: Instruction Manual*. Publication No. 8609–60, LI-COR, Lincoln, NE
- NREL (1992) *User's manual, National Solar Radiation Data Base (1961–1990)*. National Renewable Energy Laboratory, Golden, CO
- Racsco P, Szeidl L, Semenov MA (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
- Richardson CW, Wright DA (1984) WGEN: A model for generating daily weather variables. USDA, Agricultural Research Service. Publ. No. ARS-8. National Technical Information Service, Springfield, VA
- SAS Institute (1999) *SAS Online Doc, Version 8*. SAS Institute, Cary, NC
- Smith JB, Tirpack DA (eds) (1990) *The potential effect of global climate change on the United States*. Hemisphere, New York
- Soltani A, Hoogenboom G (2003) A statistical comparison of the stochastic weather generators WGEN and SIMMETEO. *Clim Res* 24:215–230
- Soltani A, Meinke H, de Voil P (2004) Assessing linear interpolation to generate daily radiation and temperature data for use in crop simulations. *Eur J Agron* 21:133–148
- Weiss A, Hays CJ (2004) Simulation of daily solar irradiance. *Agric For Meteorol* 123:187–199
- Wilks DS (1992) Adapting stochastic weather generation algorithms for climate change studies. *Clim Change* 22:67–84
- Williams JR, Jones CA, Kiniry JR, Spaniel DA (1989) The EPIC crop growth model. *Trans ASAE* 32:497–511

*Editorial responsibility: Otto Kinne,
Oldendorf/Luhe, Germany*

*Submitted: December 28, 2004; Accepted: July 7, 2005
Proofs received from author(s): August 5, 2005*