Short communication

A brief assessment of a weather data generator (CLIMGEN) at Southern African sites

BSE Clemence
PO Box 619, Howick 3290, South Africa

Abstract

Weather data generators usually consist of two parts - a section that derives site-specific "data-generation coefficients", and a routine that creates the actual observations themselves. In this study, the two CLIMGEN routines were integrated into a single PASCAL program, together with a radiation estimator, and then tested at sites in Southern Africa which had more than 50 years of concurrent daily air temperature and rainfall data. Annual figures were compared for all sites, while monthly statistics were compared for "best" and "worst" cases. Additionally, CLIMGEN daily data were used with the crop yield model CERES-Maize. Yield estimates obtained were compared with those using real observations over a similar 25-year period.

CLIMGEN appears to work satisfactorily over much of Southern Africa. Its representation of climatic means and day to day variability are good. However, it requires more testing at places with shorter climatic records. Also, additional research using other crop yield models at other sites would be invaluable.

Introduction

Crop growth simulation models can be used in production risk studies provided sufficient long-term daily weather data are available. Crop models such as CERES-Maize and CERES-Wheat require daily values of at least rainfall, air temperature and solar radiation, although the latter can be estimated from potential surface values (Clemence, 1992a). However, it is common knowledge that long-term records (say 30 years or more) of concurrent daily weather data are scarce in Southern Africa. Thus, provided they are reliable, so-called "weather estimators" (IBSNAT Project, 1989) can play an important role in supplementing weather data records so that adequate climatic risk assessment can take place.

The "Daily Rainfall Model" (Zucchini and Adamson, 1984) is an example of a locally developed weather estimator. The author has tested the model informally, and it appears to work well at a number of sites here. However, it lacked the ability to generate concurrent air temperature observations, and hence investigation into this estimator was abandoned at an early stage.

Weather estimators usually work by computing certain probability/variability coefficients from existing real data. These coefficients are then used to generate the required data sequences. Provided sufficient real data have been used to create these, generated weather data from simulators such as TAMSIM (McCaskill, 1990a) and WGEN (Richardson and Wright, 1984) can emulate reality quite successfully over longer periods such as a month or more (Meinke et al., 1993). The latter authors do warn that because of the stochastic element present in estimators such as WGEN and TAMSIM, they are not recommended for the infilling of small gaps in weather data records. This is owing to the possibility of the occasional introduction of extreme events.

The CLIMGEN model (Campbell, 1993) is a modification of the well-known WGEN (Richardson and Wright, 1984) weather estimator. However, it is simply presented in the universal computer language PASCAL, which means the user can adapt the program and its input and output formats at will.

Method

To test the Campbell estimator in this study, weather stations with suitably lengthy concurrent records of at least rainfall and air temperature were selected. "Suitably lengthy" climatic records for a place were thought to be in the order of 30 years, especially if the climate of semi-arid areas was to be adequately represented (Lynch and Dent, 1990). However, to avoid auto-correlation, the period of real data used to evaluate the synthetic weather values had to be separate from that used to create the parameter files. In other words, at least 60 years of real, concurrent weather data were considered necessary for a station to be included in this test. Of the 430 odd weather stations for which daily records are kept on the PC-based weather data bank of the Grain Crops Institute (GCI) (Clemence, 1992b), only about 12 had concurrent records of air temperature and rainfall of 60 years or more. The stations selected appear in Table 1, and, fortunately, represented a wide variety of climates.

Next, parameter files of coefficients were created for each station, using the CLIMPAR routine of the CLIMGEN program. These were based on 30 years of data in each case, except for Newcastle and Jan Kempdorp. Each parameter file was then used to generate a 30-year record of synthetic daily data. These data sets were summarised statistically into monthly means and standard deviation values for certain weather elements. Since the key to realistic weather data estimation is the assessment of variability, extremes of daily air temperature were determined for each month, as was extreme high and low mean monthly rainfall. The mean monthly number of rain days were also computed, as rainfall distribution is important in crop growth simulation modelling.

The next aspect of the testing was to compare crop yield model outputs using both real and synthetic data. For this the CERES-Maize crop yield estimator (Ritchie et al., 1986) was used with 25 years of daily weather data at Cedara. A Hutton soil was assumed. Planting day was set at 327, the cultivar assumed

☎(0332) 303-023; fax (0332) 303-023; e-mail
Received 9 July 1996; accepted in revised form 9 January 1997.

TABLE 1
WEATHER STATIONS USED IN EVALUATING THE CAMPBELL (1993) WEATHER ESTIMATOR (ARRANGED BY LATITUDE). MAP IS THE MEAN ANNUAL PRECIPITATION IN mm.

Name	Latitude	Longitude	Altitude	Record length	MAP
Messina	22° 16'	29° 54'	522	65 years	348.9
Warmbad	24° 54'	28° 20'	1143	55 years	625.8
Skukuza	24° 59'	31° 36'	263	93 years	526.4
Krugersdorp	26° 06'	27° 46'	1699	79 years	803.4
Potchefstroom	26° 44'	27° 05'	1345	78 years	617.8
Newcastle	27° 45'	29° 55'	1198	53 years **	893.0
Jan Kempdorp	27° 57'	24° 50'	1175	52 years **	450.4
Upington	28° 27'	21° 15'	793	93 years	110.0
Glen	28° 57'	26° 20'	1304	70 years	528.2
Cedara	29° 32'	30° 17'	1076	73 years	880.7
Mohaleshoek	30° 09'	27° 28'	1600	78 years	717.8
Middelburg	31° 29'	25° 01'	1270	77 years	361.0
Port St Johns	31° 38'	29° 33'	47	67 years	1168.8
Moorreesburg	33° 17'	18° 42'	177	62 years	390.5

** 23 and 22 years used for parameter file creation.

TABLE 2
YIELD STATISTICS FOR MAIZE IN TONNES OVER A 25-YEAR PERIOD AT CEDARA (REAL DATA AND CLIMGEN SYNTHETIC DATA BASED ON 16- AND 33-YEAR PARAMETER FILES RESPECTIVELY)

	Mean yield	Standard deviation	Highest yield in 25 years	Lowest yield in 25 years
Real data (1970-1994)	6004.9	2153.7	8181.5	642.2
Synthetic data based on 16 year parameter file (1945-1960)	5576.2	1656.2	8392.6	1490.5
Synthetic data based on 33 year parameter file (1960-1993)	6780.4	1230.0	8527.6	3095.2

was PANNAR 473 with a plant population density of 4.5 per square metre. Three separate model runs were conducted. Firstly, real data was used for the years 1970 to 1994. Then 25 years of daily weather observations were created with CLIMGEN using a parameter file based on 33 years of actual data (1960 to 1993), and then using a parameter file based on just 16 years of real data (1945 to 1960). The results are compared in Table 2.

Results

To determine how Campbell's simulated weather data compared with actual observations, a minimum of 30 years of real data were then extracted from the GCI bank for each of the 14 sites. The resulting data files were summarised in the same way as the synthetic values.

The statistics for real and synthetic data were first compared for all stations on a mean annual basis. Figure 1 compares mean annual rainfall and standard deviation values, and shows close agreement between actual and synthetic statistics. The difference in standard deviation values for real and synthetic data at Port St. Johns can be attributed to an unusually high rainfall figure of 530.1 mm in July, 1931. The annual means, and the standard deviations for the other months agree very well, however. This is true for all the other stations, except for Krugersdorp (annual means for real and synthetic data are 829.4 and 715.6 respectively).

Figure 2 shows the scatter of monthly standard deviations for all stations, with the outlier for Port St. Johns for July quite obvious. Willmott's (1982) Index of Agreement for the real versus the synthetic monthly standard deviations was 0.853.

Figures 3 and 4 compare the mean annual statistics of all stations for minimum and maximum daily air temperature. The differences between the real and the generated data are negligible,

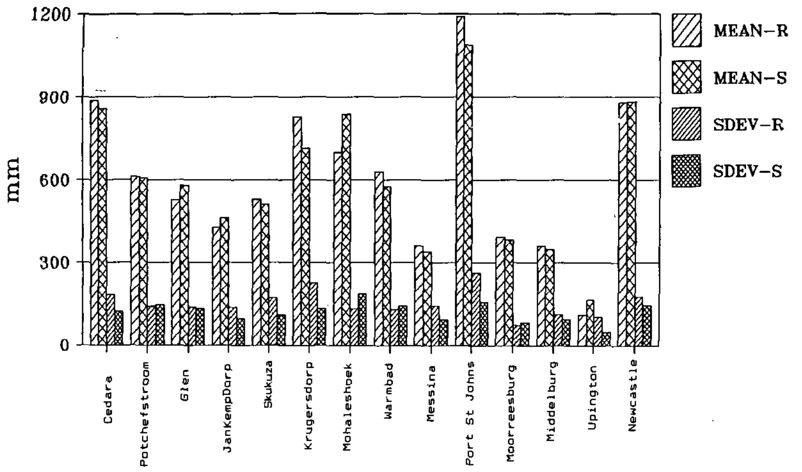

Figure 1
Mean annual rainfall and standard deviation for all sites over 30 years (R = real weather data, S = synthetic weather data)

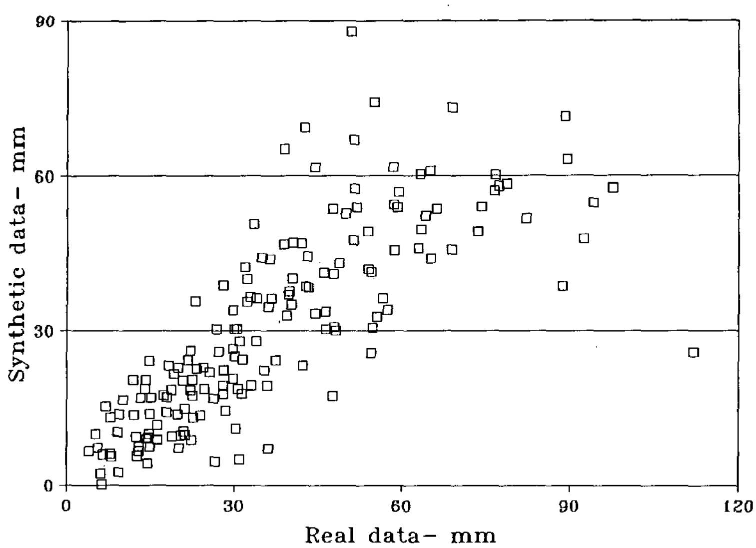

Figure 2
Mean monthly standard deviation for rainfall for all sites- real and synthetic weather data

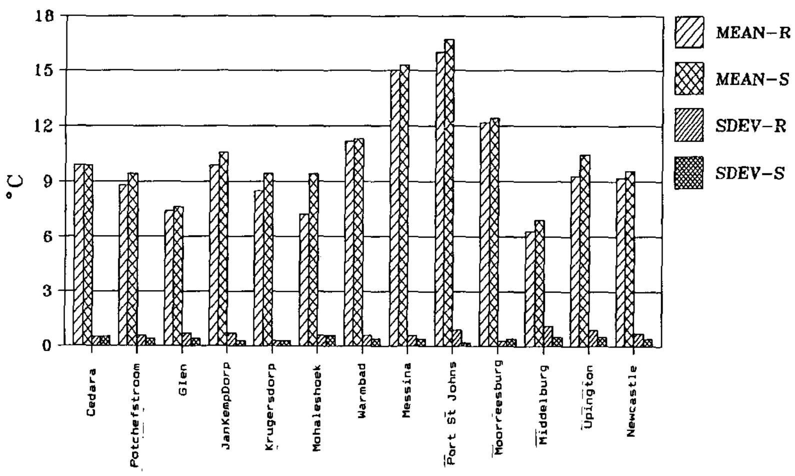

Figure 3
Mean anuual statistics for minimum daily air temperature for all sites (R = real weather data, S = synthetic weather data)

except for minimum air temperature at Mohaleshoek (the mean for the real data was 7.2°C, vs. 9.4°C). Whether this was random error would need to be investigated.

Figure 5 compares the estimated number of rainy days per month with the real figures for Moorreesburg (a winter rainfall station) and Krugersdorp (a summer rainfall station). At Moorreesburg, in only one month did the estimated number of rain days differ from the actual mean by more than one day. At Krugersdorp, the number of rain days differed from the mean actual figures by two days for five months of the year. The balance of the estimates was all within one day of the real values. This result is particularly encouraging, since rainfall distribution is a critical factor in crop growth.

As far as the crop yield model runs using real and synthetic data were concerned (Table 2), mean yields for both sets of synthetic data varied by roughly 10% from the average yield based on actual data. The standard deviation of the mean was more realistic in the case of the 16-year parameter file, as was the absolute lowest yield over the 25-year period. This appears to be random error, since one would have expected more variation in the data set derived from a "longer term" parameter file with a greater chance of including droughts and floods in the computation of the coefficients.

Conclusions

The Campbell weather estimator appears to work satisfactorily over much of Southern Africa. Mean daily and monthly values for synthetic rainfall and air temperature agreed consistently with those for real observations over different periods. Except for one or two weather stations, standard deviations based on synthetic data correlated well with those computed from actual data. This suggests that the more extreme weather events were described adequately by CLIMGEN. It must be remembered, however, that the coefficients created for weather estimation in this study were derived in nearly all instances from about 30 years of actual observations. Whether this criterion can be lowered for moister, more maritime-type localities with less climatic variability would require additional research. The CERES-maize experiment at Cedara with weather data from a CLIMGEN parameter file based on just 16 years of real data had realistic results for this relatively moist site.

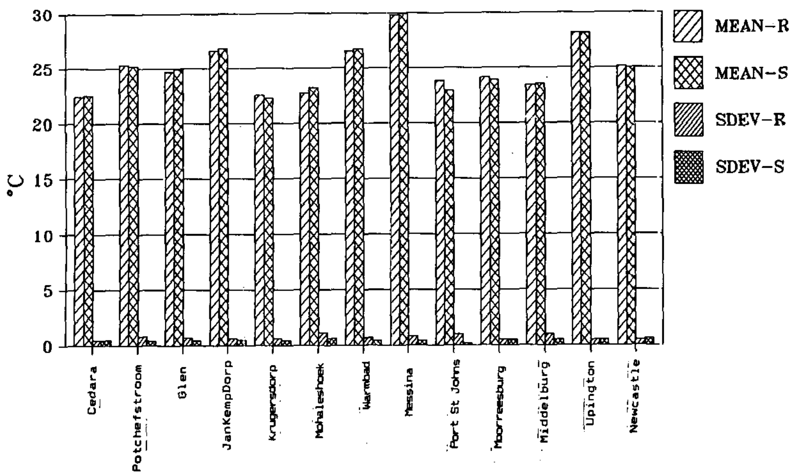

Figure 4
Mean annual statistics for maximum daily air temperature for all sites (R = real weather data, S = synthetic weather data)

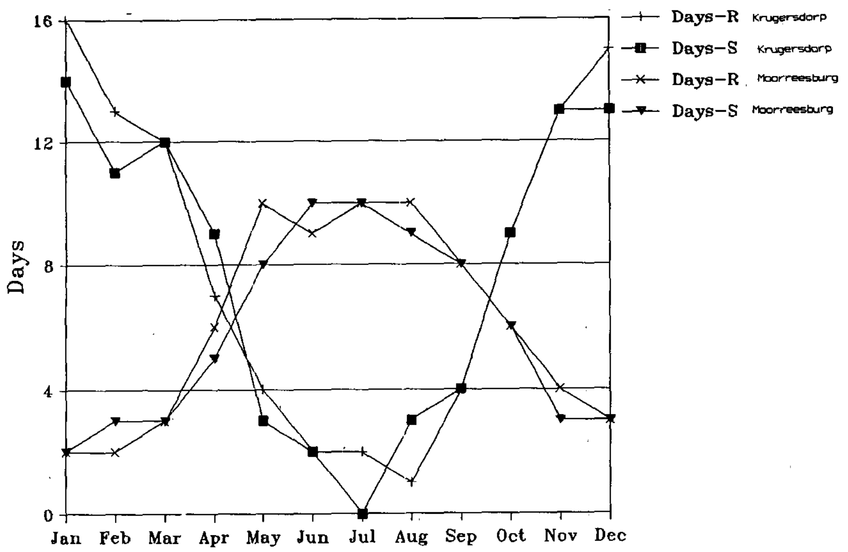

Figure 5
Mean monthly number of rain days for Krugersdorp and Moorreesburg (R = real weather data, S = synthetic weather data)

References

CAMPBELL G (1993) Personal communication.

CLEMENCE BSE (1992a) An attempt at estimating solar radiation at South African sites which measure air temperature only. *S. Afr. J. Plant Soil* **9** (1) 40-42.

CLEMENCE BSE (1992b) A PC-based weather data bank for crop growth modelling. *Water SA* **18** 293-297.

IBSNAT PROJECT (1989) Decision Support System for Agrotechnology Transfer V2.1 (DSSAT V2.1). Department of Agronomy and Soil Science, College of Tropical Agriculture and Human Resources, University of Hawaii, Honolulu.

LYNCH SD and DENT MC (1990) Appropriate record lengths for the estimation of mean annual and mean monthly precipitation in Southern Africa. *Water SA* **16** 93-98.

McCASKILL MR (1990a) TAMSIM - A program for preparing meteorological records for weather-driven models. Tropical Agronomy Technical Memorandum No. 65, 1990. CSIRO, Div. of Tropical Crops and Pastures, Brisbane. 26 pp.

MEINKE H, CARBERRY PS, McCASKILL MR, HILLS MA and McLEOD I (1993) Performance evaluation of radiation and temperature data generators using dynamic crop simulation models. Draft manuscript prepared for Agricultural and Forest Meteorology.

RICHARDSON CW and WRIGHT DA (1984) WGEN: A model for generating daily weather variables. US Dept. of Agriculture, Agricultural Res. Service, ARS-8. 83 pp.

RITCHIE JT, KINIRY JR, JONES CA and DYKE PT (1986) CERES-Maize: A Simulation Model of Maize Growth and Development. Texas A&M University Press, College Station.

WILLMOTT CJ (1982) Some comments on the evaluation of model performance. *Bull. Am. Met. Soc.* **63** 1309-1313.

ZUCCHINI W and ADAMSON PT (1984) The occurrence and severity of droughts in South Africa. Water Research Commission Report No. 91/1/84. 191 pp.