WEATHER GENERATION WITH A NEW APPROACH TO RAINFALL VARIANCE ESTIMATION AND SEASONAL CORRELATION OF VARIABLES FOR CROP PRODUCTION MODELLING

L. Kuchar

Department of Mathematics, Wroclaw University of Environmental and Life Science
ul. Grunwaldzka 53, PL-50357 Wroclaw, Poland
E-mail: Leszek.Kuchar@gmail.com

Received: 15 December 2011. Accepted: 31 December 2011.

One of the key problems of crop modelling when studying the climate change impact is an application of quality mathematical models. Most of the models dealing with climate change and decision supporting systems for agriculture adaptation require weather generators producing meteorological data. In this paper improvement of the weather generator WGEN of Richardson as well as the validation test of the model are presented. The modified model (WGENK) introduces a set of parameters calibrating monthly rainfall variance, annual courses of transition probability, α parameter of Γ distribution correlations between solar radiation and temperatures described by trigonometric polynomials. The method was tested by comparing statistics of 500 years of generated data with 20 years of observed weather parameters at 30 meteorological stations. Validation tests showed that the modeled weather data were comparable with climatic characteristic of observed data. The monthly average rainfall variance was reduced about four times compared to Richardson model and the previous model version.

Keywords**:** Agricultural meteorology, mathematical model, weather generator, meteorological variables.

INTRODUCTION

One of the key problems of agriculture is the climate change impact on crop production. Decision supporting systems for agricultural adaptation including change of crops, crop varieties, zonation, nutrient and agrotechnical management, irrigation strategy are based on the procedure shown on Fig. 1.

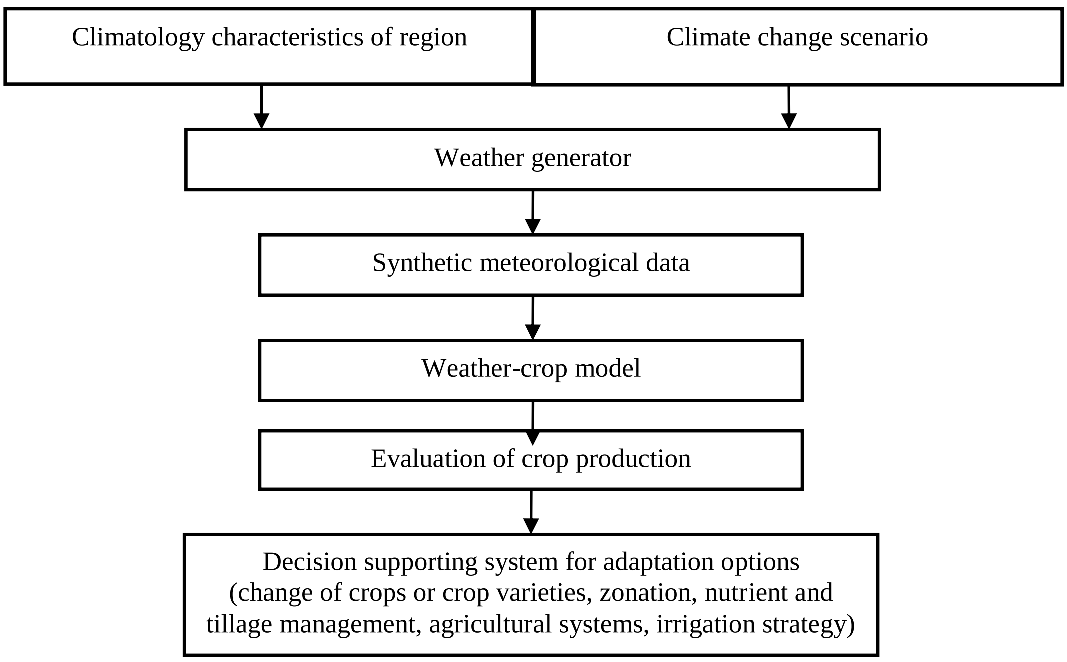

Fig. 1. Evaluation of crop production for the future climate

The procedure requires daily data of meteorological variables. Most often they are solar radiation (SR), maximum (T_{max}) and minimum air temperatures (T_{min}), and total precipitation (P) (Hunt et al., 1998; Richardson, 1985; Szulczewski et al., 2010). If there are no required data, then application of most of the models is very limited. Such a situation can occur if there is a lack of meteorological stations in the studied area, or if records of the weather data are not available (Apipattanavis et al., 2010). It is particularly common when predictions of the climate parameters in the future are necessary. The first weather generator (WGEN), which was generating meteorological data for the needs of agricultural modelling (mainly for crop simulations in new climate scenarios), was constructed by Richardson at the beginning of the 80-s, (Richardson, 1985; Semenov, 2006; Wilks, Wilby, 1999; Zhang, 2004). WGEN used the Markov chains to determine occurrence of wet/dry days, and gamma or exponential probability distribution for amount of rainfall. Daily values of solar radiation as well as maximum and minimum air temperatures were generated by general linear model (GLM) and were considered as weakly stationary processes (Hayhoe, 2000; Srikanthan, Mcmahon, 2001; Wilks, Wilby, 1999).

Generated data series for solar radiation and temperature are required to have the same statistics as measured climatology data including means, variations and cross, lag and lag-cross correlations. Generated data on precipitation and its variation are also expected to be similar to the observed data (Bannayan, Hoogenboom, 2008; Hayhoe, 1998; Schoof, Pryor, 2008; Schuol, Abbaspour, 2007). While means and variations of generated data for solar radiation and temperature sufficiently estimate moments of theoretical distribution, there are still poor fitting in precipitation variation, precipitation extremes and correlations between variables (Abdulharis et al., 2010; Apipattanavis et al., 2010; El-Sadek et al., 2011; Kuchar, 2004; Srikanthan, Mcmahon, 2001; Wilks, Wilby, 1999). In the WGEN weather generator given by Richardson, the cross, the lag and the cross-lag correlation illustrating seasonal and spatial relation between variables are constant through locations and over the year (Richardson, 1985; Schoof, Robeson, 2003). Spatial and monthly course of correlation are introduced to the model by staircase function (Hayhoe, 2000; Schoof, Robeson, 2003). However, correlations which differed from month to month and at different locations were constant within the given month. Transition probabilities and parameters of rainfall probability distribution are fixed monthly or biweekly as a set of 12 or 26 values (Richardson, 1985) or estimated by strait function (Hayhoe, 2000; Rivington et al., 2006; Srikanthan, Mcmahon, 2001).

In this paper evaluation of rainfall variance estimation and correlation fitting are studied by application of the parallel forcing to the model: (1) – parametrization of serial correlation, transition probability, and α parameter of Γ probability distribution and, (2) – a new idea of calibration of β parameter of Γ probability distribution for monthly precipitation (the modified model was called WGENK).

The above procedure when applied to the fitting of daily values of correlation, transition probability, and α parameter of Γ probability distribution by a trigonometric polynomial function using continuous function versus staircase or strait function preserves continuous changes in generated data and their climatology statistics, while a calibration of β parameter (responsible for a rainfall variance), as a simple way, preserves precipitation simulation well.

MATERIALS AND METHODS

Evaluation of WGENK model (rainfall variance estimation and variable seasonally correlation fitting in particular) was made for Polish Lowland with comprehensive climatology statistics derived at 30 meteorological stations over a 20-year period.

Above mentioned climatology information was obtained from the Department of Meteorology and Climatology at the University of Warmia and Mazury in Olsztyn (originally collected the Institute of Meteorology and Water Management (IMGW)) collected from meteorological stations across the country. Stations were selected so that there would be no data gaps. In case of missing data on solar radiation they were estimated using values of precipitation, minimum and maximum air temperatures, and calculated values for clear sky radiation according to Hunt et al. (1998).

According to Kuchar (2004) the WGENK model – a modification of the well-known WGEN weather generator of Richardson - generates daily values of P, SR, T_{max}, T_{min}. The occurrence of precipitation has an influence on the solar radiation and the temperature for each day by determining the day status (wet or dry) and independently generating solar radiation and temperature for the given day (Kuchar, 2004). Precipitation data are generated at first by means using Markov chain to determine occurrence of wet and dry days, and then with two-parameter gamma distribution for the amount of rainfall (Soltani et al., 2000; Taulis, Milke, 2005). The transition probabilities and α parameter of Γ distribution are continuous during one year for a given location owing to the application of a trigonometric polynomial function (Eq. 1) which describes the course of seasonal changes (Kuchar, 2004; Walpole et al., 2002):

$$W_n(t) = \sum_{k=1}^{n} (a_k \sin(kt) + b_k \cos(kt)) \quad (1)$$

t – the day of the year $1 \le t \le 365$
n – the degree of the polynomial
a_k and b_k – coefficients.

Estimation of the a_k and b_k coefficients is made by the method of least square. Selection of the polynomial degree *n* is halted when the proportion of the total variation in the response variable is explained by the fitted model in 90%, or *n* arbitrary exceeds the number of months. Functions $W_n(t)$ are fitted by non linear approximation using NLIN procedure given by SAS Institute Inc. (2008).

Daily values of solar radiation, maximum and minimum air temperatures are considered as weakly stationary processes but include seasonal changes of correlations (cross, lag and lag-cross) between factors in the form:

$$d_i = A \cdot d_{i-1} + B \cdot e_i, \quad (2)$$

where d_i – a vector (3×1) of normalized residuals for all three variables for day i; e_i – a vector (3x1) of independent random components normally distributed with mean zero and variance 1; A and B – matrices (3×3) obtained from equations: $A = M_1 M_o^{-1}$, $BB^T = M_o^{-1} AM^T$; M_o, M_1 - matrices (3x3) of cross correlations (M_o) and lag-correlation, lag-cross correlations (M_1) of variables SR, T_{max}, T_{min} (Richardson, 1985).

Seasonal correlation (for each correlation i.e. components of M_o, M_1 matrices) is described by trigonometric polynomial function (1) in the same manner as transition probability and α parameter of Γ distribution. Solar radiation and temperature are generated by (Eq. 2) which uses daily values of correlation estimated by function (Eq. 1). To prevent poor estimation of transition probabilities, α parameter and correlations smoothing them by 11-day moving window before function fitting and Richardson's suggestion of twenty-year observation are applied (Richardson, 1985; Srikanthan, Mcmahon, 2001).

New approach presented in this paper provides better fitting of rainfall variance generated by the WGEN model. Set of coefficients k_i(i = 1, 2,..., 12) for month i, reshapes Γ probability distribution for a given day t($1 \le t \le 365$). Parameters β determining the shape of $\Gamma(\alpha_t, \beta_t)$ probability distribution are obtained by using the formula: $\beta_t = \mu_t/(w_t \cdot \alpha_t)$, for a given scale parameter – α_t, average total rainfall – μ_t and weights – w_t of shape parameter β_t, where

$w_t = k_1$ for $t \in [1, 31]$, $w = k_2$ for $t \in [32, 59]$, ..., $w_t = k_{12}$ for $t \in [334, 365]$.

In other words, the shape parameters β (estimated on the basis of data sample) of Γ distribution (applied for data generation) are multiplied by coefficients k_i to minimize the difference between the observed and generated variances of total precipitation.

RESULTS AND DISCUSSION

The WGENK weather generator was examined by parallel comparison: observed climatology data vs. WGEN generated data, and observed climatology data vs. WGENK generated data (Castellvi, Stöckle, 2001; Schuol, Abbaspour, 2007; Soltani et al., 2000; Taulis, Milke, 2005). The evaluation was done using measured data from meteorological stations and the data generated by both models for 500-year data series. For all data, the values of means, variances and correlations in different time periods – annual, growing season (March – October) and monthly – were computed and evaluated by absolute differences between observed and generated parameters as well as relative differences in form (Hayhoe, 1998; Rivington et al. 2006): abs (observed – estimated)·100% /observed.

Data generated by WGEN as well WGENK models have shown low errors for all means and variances of SR, T_{min}, T_{max} (Table 1).

Table 1. Average absolute and relative error for means and standard deviations for solar radiation, maximum and minimum air temperature generated by the WGEN and WGENK models for different time periods (average for 30 meteorological stations).

WGEN							WGENK						
Time period	SR (MJ m^{-2})		T_{max} (ºC)		T_{min} (ºC)		Time period	SR (MJ m^{-2})		T_{max} (ºC)		T_{min} (ºC)	
	Mean	SD	Mean	SD	Mean	SD		Mean	SD	Mean	SD	Mean	SD
Annual							Annual						
(*)Abs	0.2	0.1	<0.1	<0.1	<0.1	<0.1	(*)Abs	0.2	<0.1	<0.1	<0.1	<0.1	<0.1
Rel.	2.1	1.6	2.0	1.4	2.0	1.5	Rel.	2.1	1.5	2.1	1.5	2.0	1.5
Growing season							Growing season						
(*)Abs	0.2	0.1	<0.1	<0.1	<0.1	<0.1	(*)Abs	0.2	<0.1	<0.1	<0.1	<0.1	<0.1
Rel.	2.1	1.5	2.1	1.6	2.0	1.6	Rel.	2.1	1.5	2.0	1.5	2.0	1.5
Month							Month						
(*)Abs	0.2	0.1	<0.1	<0.1	<0.1	<0.1	(*)Abs	0.2	<0.1	<0.1	<0.1	<0.1	<0.1
Rel.	2.0	1.6	2.1	1.6	2.2	1.6	Rel.	2.0	1.4	1.9	1.5	2.0	1.5

(*) Abs. = abs (observed – estimated); Rel. = (abs (observed – estimated)/observed)·100%).

The variance of precipitation and lag, cross and cross-lag correlations received when using the WGEN model proved a poor estimation and not acceptable errors (Table 2, 3). To improve the estimation, annual courses of transition probabilities, α parameter of Γ distribution, variables correlation described by trigonometric polynomial approximation and β parameter calibration was introduced to the model. To make a better rainfall variance approximation, time series of transitions probability and α parameter are smoothed by moving average (five points on either side of the target) and fitting by trigonometric polynomial, and calibrating shape parameter β by w_t weights.

The new approach based on the measured data from thirty lowland stations has reduced absolute errors of variances for precipitation (generated vs. observed) 2–3 times, while relative error was reduced close to 7% compared to the original WGEN method (Table 2), depending on the considered period. Annual absolute error of standard deviation was reduced from 7.7 mm to 3.2 mm and relative error from 8.5% to 3.2%. The best results were obtained for monthly periods: absolute error of standard deviation was reduced from 5.9 mm to 1.9 mm while relative error from 25.3% to 6.8%. No significant differences were noted for precipitation sums (Table 2). As expected, the new method provided no significant improvement in fitting of means and variances for solar radiation, maximum and minimum air temperatures.

Table 2. Average absolute and relative error for sums and standard deviations for precipitation generated by WGEN and WGENK models for different time period (average for 30 meteorological stations).

WGEN					WGENK				
Time period	Absolute error (mm)*		Relative error (%)**		Time period	Absolute error (mm)*		Relative error (%)**	
	Sum	SD	Sum	SD		Sum	SD	Sum	SD
Annual	1.3	7.7	0.2	8.5	Annual	1.0	3.2	0.1	3.2
Growing season	1.2	5.5	0.2	6.6	Growing season	1.1	1.6	0.1	1.9
Monthly	2.9	5.9	5.8	25.3	Monthly	1.5	1.9	2.2	6.8

(*) abs (observed – estimated), (**) (abs (observed – estimated)/observed)·100%.

The same trigonometric approach was also used to fit a seasonal correlation between solar radiation, maximum and minimum air temperatures (Kuchar, 2004). Each correlation given by M_o, M_1 matrices was smoothed by 11-day moving window before function fitting. The approach used for annual course of correlations reduced errors of correlations about three times (comparing to WGEN model); new correlations were statistically examined (at 0.05 level) and accepted in 98% for all 5040 computed tests (correlation type **x** time period **x** number of stations) while for original method accepted in 50 to 55%. For the annual period, highest average reduction error was observed for cross correlations (from 0.09 to 0.02). In these cases the number of significantly different correlations between observed and generated data was 5.2% while for original WGEN model – 56%.

Table 3. Average absolute error for cross, lag and cross-lag correlations between solar radiation, maximum and minimum air temperature, and percentage of significantly different correlations from measured meteorological parameters at α=0.05 for different time periods (average for 30 meteorological stations).

WGEN				WGENK			
Correlation type	Time period	Average absolute error (*)	Sig diff correlations at α=0.05	Correlation type	Time period	Average absolute error (*)	Sig diff correlations at α=0.05
Lag	Annual	0.06	45.3	Lag	Annual	0.02	0.0
	Growing season	0.05	74.3		Growing season	0.01	0.1
	Monthly	0.13	22.2		Monthly	0.04	1.1
Cross	Annual	0.09	56.0	Cross	Annual	0.02	5.2
	Growing season	0.07	89.1		Growing season	0.03	0.2
	Monthly	0.21	41.0		Monthly	0.05	0.6
Lag-Cross	Annual	0.07	80.2	Lag- Cross	Annual	0.03	5.2
	Growing season	0.08	83.0		Growing season	0.02	5.2
	Monthly	0.18	62.1		Monthly	0.04	1.9

(*) abs (observed – estimated)/number of observations.

There is a significant reduction of errors for growing seasons. This fact is well shown by absolute difference between models, but first of all shown by the percentage of statistically different (at alpha 0.05) tests. The evidence of method improvement is observed for monthly correlations. For this period (i.e. monthly) the biggest error reduction for absolute differences was observed for cross correlation (error reduction from 0.21 to 0.05). Percentage of significantly different correlations was reduced from 22.2% to 1.1% for lag correlation, from 41.0% to 0.6% for cross correlation, and from 62.1% to 1.9% for lag-cross correlation.

Between different types of correlation the best improvement was obtained for cross correlations, and the worst – for lag correlations. The structure of errors depending on the type of correlation, on the time period and on the method is shown in Table 3.

CONCLUSIONS

A test of WGENK weather generator for weather large data set from Lowland Poland showed low errors for means and variances of generated data, and accepted errors for lag, cross, and cross-lag correlations. Computed tests have shown the significant improvement in the method fitting of annual seasonality of transition probability, α parameter of Γ probability distribution and correlation course of variables by application of the trigonometric polynomial function. Application of β parameter calibration reduced standard deviation relative error of precipitation 3–4 times compared to original WGEN model. The results of the validation suggest that the weather generator WGENK performs adequately for many applications.

ACKNOWLEDGEMENTS

The author thanks unknown reviewers for their constructive comments and the vice-editor of the journal for helpful assistance and advice concerning this work.

REFERENCES

Abdulharis A., Khan M. A., Chhabra V., Biswas S., Pratap A. 2010. Evaluation of LARS-WG for generating long term data for assessment of climate change impact in Bihar. Journal of Agrometeorology. 12:198-201.

Apipattanavis S., Bert F., Podesta G., Rajagopalan B. 2010. Linking weather generators and crop models for assessment of climate forecast outcomes. Agricultural and Forest Meteorology. 150:166-174.

Bannayan M., Hoogenboom G. 2008. Predicting realizations of daily weather data for climate forecasts using the non-parametric nearest-neighbour re-sampling technique. International Journal of Climatology. 28:1357-1368.

Castellvi F., Stöckle C.O. 2001. Comparing the performance of WGEN and CLIMGEN in the generation of temperature and solar radiation. Transactions of the ASAE. 44:1683-1687.

El-Sadek A., Bleiweiss M., Shukla M., Guldan S., Fernald A. 2011. Alternative climate data sources for distributed hydrological modelling on a daily time step. Hydrological Processes. 25:1542-1557.

Hayhoe H. N. 1998. Relationship between weather variables in observed and WXGEN generated time series. Agricultural and Forest Meteorology. 90:203-214.

Hayhoe H. N. 2000. Improvements of stochastic weather data generators for diverse climates. Climate Research. 14:75-87.

Hunt L.A., Kuchar L., Swanton C.J. 1998. Estimation of solar radiation for use in crop modelling. Agricultural and Forest Meteorology. 91:293-300.

Kuchar L. 2004. Using WGENK to generate synthetic daily weather data for modelling of agricultural processes. Mathematics and Computers in Simulation. 65:69-75.

Richardson C.W. 1985. Weather simulation for crop management models. Transactions of the ASAE. 28:1602-1606.

Rivington M., Matthews K. B., Bellocchi G., Buchan K. 2006. Evaluating uncertainty introduced to process-based simulation model by alternative sources of meteorological data. Agricultural Systems. 88:451-461.

SAS/STAT 9.2 User's guide. 2008. SAS Institute Inc., Cary.

Schoof J. T., Pryor S. C. 2008. On the proper order of Markov Chain Model for daily precipitation occurrence in the contiguous United States. Journal of Applied Meteorology. 47:2477-2486.

Schoof J. T., Robeson S. M. 2003. Seasonal and spatial variation of cross-correlation matrices used by stochastic weather generators. Climate Research. 24:95-102.

Schuol J., Abbaspour K.C. 2007. Using monthly weather statistics to generate daily data in a SWAT model application to West Africa. Ecological Modelling. 201:301-311.

Semenov M.A. 2006. Using weather generators in crop modelling. Acta Horticulturae. 707:93-100.

Soltani A., Latifi N., Nasiri M. 2000. Evaluation of WGEN for generating long term weather data for crop simulations. Agricultural and Forest Meteorology. 102:1-12.

Srikanthan R., Mcmahon T. A. 2001. Stochastic generation of annual, monthly and daily climate data: A review. Hydrology and Earth System Sciences. 5:653-670.

Szulczewski W., Zyromski A., Biniak-Pierog M., Machowczyk A. 2010. Modelling of the effect of dry periods on yielding of spring barley. Agricultural Water Management. 97:587-595.

Taulis M. E., Milke M. W. 2005. Estimation of WGEN weather generation parameters in arid climates. Ecological Modelling. 184:177-191.

Walpole R. E., Myers R. H., Myers S. L., Ye K. 2002. Probability and statistics for
engineers and scientists. Prentice Hall, 7th Ed., New Jersey.

Wilks D. S., Wilby R. L. 1999. The weather generation game: a review of stochastic weather models. Progress in Physical Geography. 23:329-357.

Zhang Q., Singh B., Gagnon S., Rousselle J., Evora N., Weyman S., Roy L. 2004. The application of WGEN to simulate daily climatic data for several Canadian stations. Canadian Water Resources Journal. 29:1-14.