Geospatio-temporal Daily Rainfall Data Generation

Guillermo A. Baigorria* and James W. Jones
Agricultural and Biological Engineering Department, University of Florida, Gainesville 32611 (* E-mail: gbaigorr@ifas.ufl.edu)

Justification

Agricultural systems have evolved in response to spatial and temporal rainfall variability but, in most regions of the world, rainfall variability continues to be a major source of risk for farmers. Adequate quantification of this variability requires rainfall records over a long time period. However, the availability of daily rainfall data over long time periods continues to be a problem. To increase rainfall record lengths, statistical methods are used to generate daily realizations of rainfall for many applications.

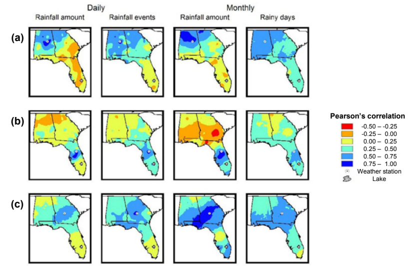


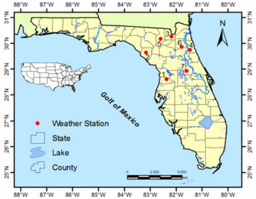
Fig. 1: Daily and monthly spatial correlations of observed rainfall events and amounts between one weather station [(a) Sylacauga, Alabama; (b) Mountain Lake, Florida; and (c) Hawkinville, Georgia] and 522 weather stations across the three states (Baigorria et al., 2007)

One issue with most currently available weather generators is that they create daily realizations for points in space without considering spatial correlation or persistence of rainfall events and amounts over space. Spatial variability may not be a problem if one's interest is in temporal properties of rainfall and its effects on crop production at points or fields. But, if spatially independent generated data are used to aggregate rainfall or model outputs over space for subsequent analyses, spatial correlations of the variables must be taken into account for the same time scale at which the data are used as inputs to models.

Objective

To develop a weather generator capable to preserve both the temporal and the spatial structure of the observed rainfall data.

Materials



- Daily rainfall data (1974 – 2004) of seven weather stations located in North Central Florida obtained from the National Climate Data Center (NCDC).

Fig. 2: Location of the seven weather stations used in the present study

Methods

1. Correlation matrices and Markov chain parameters:

- Correlation matrices of daily rainfall events and daily rainfall amounts were calculated at monthly steps.
- Probabilities of dry-wet and wet-wet days were calculated individually for each weather station.

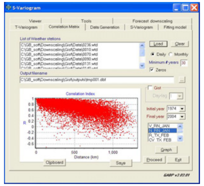


Fig. 3: Software used to produce correlation matrices and to generate rainfall data

2. Generating rainfall events:

- Monthly Choleski factorization matrices were calculated based on the rainfall event correlation matrices.
- For each day to be generated, seven random numbers, ranging from 0 to 1, were sampled from a Uniform Distribution.
- The matrix of random numbers was multiplied by the Choleski factorization matrix, yielding correlated values from 0 to 1.

$$\begin{bmatrix} Event_{ws_1} \\ \cdots \\ \cdots \\ \cdots \\ Event_{ws_7} \end{bmatrix} = \begin{bmatrix} Rnd_1 \\ \cdots \\ \cdots \\ \cdots \\ Rnd_7 \end{bmatrix} \times \begin{bmatrix} Ch_{1,1} & Ch_{1,2} & \ldots & \ldots & Ch_{1,7} \\ 0 & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \ldots \\ 0 & 0 & 0 & 0 & Ch_{7,7} \end{bmatrix}$$

- Rainfall events were assigned using a First Order Markov chain.

3. Generating rainfall amounts:

- Monthly Choleski factorization matrices were calculated based on the rainfall amount correlation matrices considering only those n weather stations with rainfall events generated in the step before.
- For each day to be generated, n random numbers, ranging from 0 to 1, were sampled from a 2-parameter Gamma distribution.
- The matrix of random numbers was multiplied by the Choleski factorization matrix yielding correlated values from 0 to 1.

$$\begin{bmatrix} Rnd_1^c \\ \cdots \\ \cdots \\ \cdots \\ Rnd_n^c \end{bmatrix} = \begin{bmatrix} Rnd_1 \\ \cdots \\ \cdots \\ \cdots \\ Rnd_n \end{bmatrix} \times \begin{bmatrix} Ch_{1,1} & Ch_{1,2} & \ldots & \ldots & Ch_{1,n} \\ 0 & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots & \ldots \\ 0 & 0 & 0 & 0 & Ch_{n,n} \end{bmatrix}$$

- The resulting matrix was scaled and multiplied by the natural logarithm of the corresponding 2-parameter Gamma function.
- 100 years of rainfall data were generated using this new method and also using an existing weather generator (WGEN; Richardson and Wright, 1984).
- Results were compared to the observed correlations.

Results

1. Rainfall events:

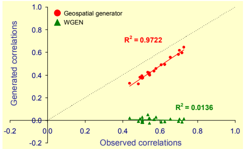


Fig. 4: Observed versus Generated correlations of daily rainfall events in January

- Observed spatial correlations of rainfall events among weather stations were significantly correlated (α=0.05) to the spatial correlations produced by the generated rainfall events from the geospatial rainfall data generation.

- Observed spatial correlations of rainfall events among weather stations were non-significantly correlated (α=0.05) to the spatial correlations produced by the generated rainfall events from the standard weather generator (WGEN).

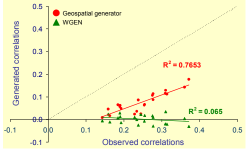


Fig. 5: Observed versus Generated correlations of daily rainfall events in July

2. Rainfall amounts:

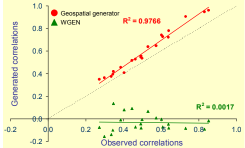


Fig. 6: Observed versus Generated correlations of daily rainfall amounts in January

- Observed spatial correlations of rainfall amounts among weather stations were significantly correlated (α=0.05) to the spatial correlations produced by the generated rainfall amounts from the geospatial rainfall data generation.

- Observed spatial correlations of rainfall amounts among weather stations were non-significantly correlated (α=0.05) to the spatial correlations produced by the generated rainfall amounts from the standard weather generator (WGEN).

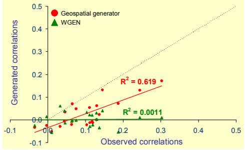


Fig. 7: Observed versus Generated correlations of daily rainfall amounts in July

- Correlations among daily geospatial generated rainfall events and amounts in the month of January were closer to the observed correlations than in those in the month of July.

Conclusions

Analysis of the main statistics obtained from individual weather stations by using both data generation methodologies, matched those from the observed climatology. However, analysis of the spatial structure of the generated data was only preserved by the new geospatial method. Analysis of a generated time series longer than the 100 years performed in this study will provide more stable and reliable statistics.

Reference

- Baigorria, GA, JW Jones, JJ O'Brien. 2007. Understanding rainfall spatial variability in the southeast USA. Int. J. Climatology, 27:749-760.
- Richardson, CW, DA Wright. 1984. WGEN: A model for generating daily weather variables. U.S. Department of Agriculture, Agricultural Research Service, Washington, D.C., ARS-8, 83 p.

Acknowledgments

The research was supported by the National Oceanic and Atmospheric Administration (NOAA) and developed under the auspices of the Southeast Climate Consortium (SECC). The views expressed in this poster are those of the authors and do not necessarily reflect the views of NOAA or any of its sub-agencies.