An update on the geospatial weather generator GiST

Guillermo A. Baigorria and James W. Jones

Justification

Weather generators are tools developed to create synthetic daily weather data over long periods of time. These tools have also been used for downscaling monthly to seasonal forecasts, produced by global and regional circulation models, to daily values in order to provide inputs for crop and other environmental models. One main limitation of weather generators is that they do not take into account the spatial structure of weather and climate in a given region or watershed. This spatial correlation is important when one spatially aggregates, for example, simulated crop yields or water resources in a watershed or region.

Objective

The objective of this study was to design a simple rainfall event and rainfall amount weather generator capable of reproducing both the daily spatial correlation among weather stations as well as the monthly statistics of each individual weather station.

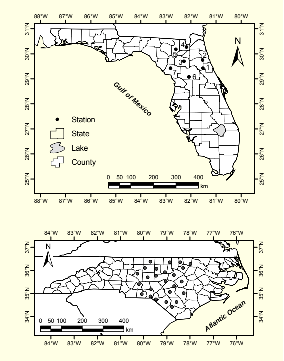

Fig 1. Study areas in Florida and North Carolina, respectively, with 7 and 25 weather stations

The GeoSpatio-Temporal weather generator (GiST)

The first step in generating rainfall events is to calculate all parameters and initial conditions (i.e. degree of association among locations for the amoeba-like assimilation, identification of core locations, total monthly number of rainfall events), whereas the second step creates the spatially and temporally correlated rainfall events. In the second step, the application of the two-state orthogonal Markov transition probabilities requires previously-generated rainfall events in two locations for the day on which rainfall is being generated. By definition, the first two core locations are not generated using the two-state orthogonal Markov transition probabilities; therefore they are generated first using an alternative method also described in the study. Rainfall amounts for each location are generated by rescaling a vector of correlated random numbers from a gamma distribution ($\mathbf{R_g}^{\Psi} \sim \Gamma[\alpha]$). This vector is obtained from uniformly distributed random numbers ranging from 0 to 1 ($\mathbf{R_u} \sim U[0,1]$) matrix-multiplied by a Toeplitz-Cholesky factorization matrix (**F**).

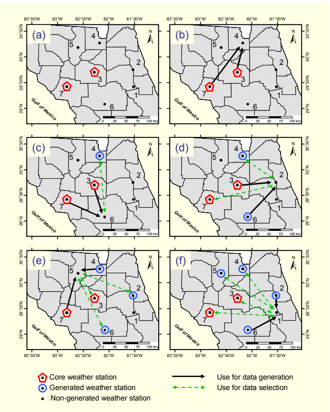

Fig 2. Example of the amoeba-like assimilation

In this case, **F** is calculated based on the correlation matrix of transformed (gamma to Gaussian) daily rainfall amounts. The size of **F** (e x e) and the random vector's element number (e) vary daily depending on the number of locations where rainfall events occur. The resulting correlated vector, which follows a Gaussian distribution, is next transformed into a gamma distribution ranging in value from 0 to 1 by using cumulative probability functions.

Results

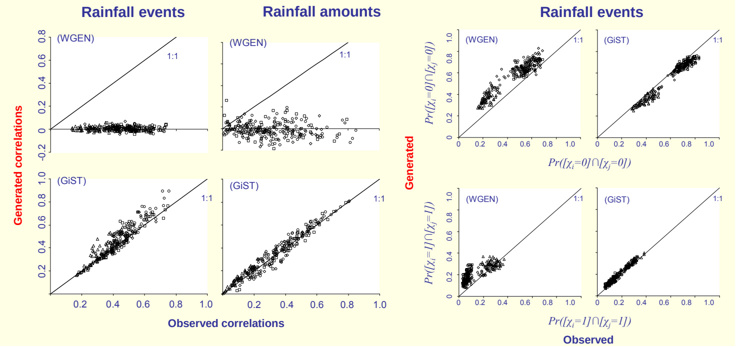

Fig 3. Comparisons between performances of WGEN and GiST in Florida. Observed versus generated correlations and joint probabilities among all pairs of weather stations for each month

Results from one thousand year-long replications individually for both study areas

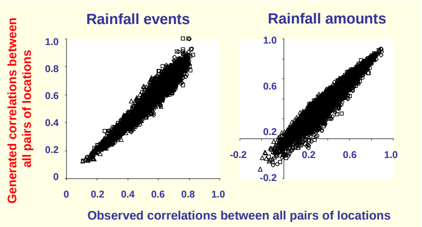

Fig 4. Comparison between correlations among all pairs of locations calculated based on observed and GiST generated rainfall events and amounts in North Carolina

Table 1. Observed versus generated monthly average (%) region-wide days without rainfall at any of the seven weather stations in Florida

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Obs	47	51	55	58	55	19	10	10	24	45	50	50
WGEN	10	9	19	20	15	1	0	1	2	13	18	12
GiST	50	52	54	62	52	22	14	13	27	49	51	51

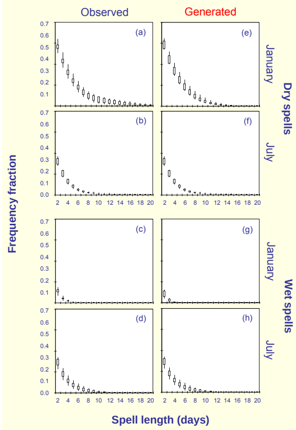

Fig 5. Whisker plots (5th, 25th, 75th and 95th percentiles) comparing observed and generated dry and wet spell length frequencies in Florida

Proposed applications

- Use for weather and seasonal climate operative forecast (NCSU)
- Multi-ensemble seasonal climate forecast (FSU)
- Drought and flood analysis at regional scales (AU)
- Downscaling Climate Change Scenarios in the SE-USA (NASA)

Conclusions

For both study areas in Florida and North Carolina, GiST reproduced the two-state first-order Markov transition probabilities, joint probabilities, distributions of length of dry and wet spells, total number of monthly number of rainy days, and total rainfall amounts with statistical significance, as well as the region-wide number of days without rainfall at any location.

Reference

- Baigorria, GA, JW Jones, JJ O'Brien. 2007. Understanding rainfall spatial variability in the Southeast USA at different time scales. *Int. J. Climatol*. **27**:749-760
- Baigorria, GA, JW Jones. 2009. GiST, A model for generating spatial-temporal daily rainfall data. SECC-09-002 Technical Report Series. 32p.

Acknowledgments

The research was supported in part through grants from the National Oceanic and Atmospheric Administration – Applied Research Center (NOAA-ARC), NOAA – Climate Program Office (NOAA-CPO), US Department of Agriculture – National Institute for Food and Agriculture (USDA-NIFA) and USDA – Risk Management Agency (USDA-RMA), 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, USDA or any of its sub-agencies.