AgMIP simulations for South India

Arthur M. Greene
International Research Institute
for Climate and Society,
Columbia University

1 General

The files in this collection were created as part of the AgMIP project (`www.agmip.org`), for the purpose of delineating climate uncertainty over the next few decades in terms of agricultural yields. The basic method of their creation, using the software suite *simgen,* is detailed in Greene et al. (2015).

2 Domain, subdomains and seasons

The initial South Indian (SI) Domain considered was all of the subcontinent south of 20°N (nominally the region bounded by 5°–20°N, 70°–85°E, including Sri Lanka). However, an EOF analysis of the seasonal cycle of precipitation over this domain revealed that climatologically it is composed of three subdomains having distinct seasonal precipitation maxima. These subdomains, denoted Regions 1, 2, and 3 are represented in data library subdirectories denoted R1–JJAS, R2–OND and R3–MAM, respectively, and will be referred to here as simply R1, R2 and R3. R1, with a June-September (JJAS) summer monsoonal maximum, corresponds to most of the main subcontinental area (in the nominal SI domain) excepting a strip along its eastern margin. R2, having a northeast monsoonal maximum in October-December (OND) comprises both this strip and the entirety of Sri Lanka. Finally, R3, with a precipitation peak in March-May (MAM), is limited to Sri Lanka. Precise subregion extents may be inspected by displaying map views of the data.

From these descriptions it is obvious that R2 and R3 overlap, but only with respect to OND. That is, the simulations for R2 actively model the OND rainy season, treating the other months (including MAM) as the off-season, whereas those for R3 model MAM, treating the other months (including OND) as the off-season. A user interested in Sri Lanka must therefore first decide which of these two seasonal rainfall peaks is of interest.

As in other modeling exercises, the off-season in each case is modeled as the climatological monthly pattern for the relevant months, superimposed on trends paralleling those assigned to the growing season. This procedure is used in order to (a) increase signal-to-noise in the modeled growing season and (b) insure that there is not an increasing discontinuity between modeled and off-season climates, particularly in the case of temperature but not excluding precipitation. It is assumed that the off-season months are less important for agricultural purposes than those modeled, and this treatment is applied mainly to provide off-season data for model spinup or multiple-year runs, where gaps between successive modeled seasons might be problematic.

3 Data reduction

Fitting of a simulation model involves detrending the observational dataset employed, here the UEA CRU TS3.21 product (Harris et al., 2013), followed by application of a vector autoregressive (VAR) model (see, e.g., Holden, 1995). Three variables are modeled jointly – precipitation and daily maximum and minimum temperatures. After detrending, the training data, extending from 1901 to 2011, are expressed in terms of principal components and it is to these PCs that the VAR model is applied. The number of retained modes depends in part on the need to reproduce within-region covariances in the simulated data. Twenty PCs are utilized here for the simulations for regions R1 and R2 and ten for R3. Because each EOF encodes variation over the entire domain there is no need for spatial downscaling (although there is a temporal disaggregation step, see below).

4 Future trends

Trends for the future cannot, of course, be inferred from the observational record. Rather, a distribution of future trends is obtained from an ensemble of Coupled Model Intercomparison Project, phase 5 (CMIP5) models (Taylor et al., 2011). Future precipitation is expressed as regional sensitivities, in terms of percent change per degree global warming, while future temperature trends are normalized by the global mean temperature. Then all that is required to project local precipitation and temperature trends into the future is a global mean temperature record. For this purpose a multimodel mean signal is utilized, based on the same CMIP5 models.

5 Downscaling

As mentioned above, the full spatial domain is represented in the PCA modeling, thus no spatial downscaling is required or performed. However the simulations produced by the VAR model are only annually resolved, representing mean wet-season values for each region (for the three variables). Again as described in Greene et al. (2015), a k-NN resampling step (Rajagopalan and Lall, 1999) is utilized to recover monthly patterns for the modeled season. Infilling of the off-season months was described above.

For daily data k-NN is also deployed. However the target dataset is now AgMERRA (Ruane et al., 2015), which also comprises a number of subsidiary variables of agricultural interest. These are (with their data library identifiers) dew point (dewp), relative humidity (hur), incoming shortwave radiation (rsds), vapor pressure (vp) and wind speed (wind). The primary variables, precipitation (pre) and maximum and minimum daily temperatures (tmx, and tmn, respectively) are also included.

6 Simulation IDs

For each subregion, and for both monthly and daily values, there are nine archived simulation files. These files, combining three precipitation trend values with each of three decadal fluctuations, span the ranges of both trend and fluctuation (or forced and natural) uncertainty, and constitute in this sense a "basic set." A typical file name might be "simall_JJAS_05_00964_SI_R1_001." Here, "JJAS" refers to the modeled season, while the two following digits ("05") identify the trend percentile (fifth). The following five-digit identifier encodes the percentile of the 10-year precipitation fluctuation, which occurs in each case during 2031-2040. Table 1 shows the fluctuation percentile associated with each of these codes. "SI" identifies the South Indian domain and "Rn," where n is 1, 2 or 3, the subregion. The final numerical identifier is an internal ID, maintained for possible future reference.

7 Further information

The user is encouraged to consult Greene et al. (2012, 2015) and the references cited therein for additional information on the methodology employed to generate these simulations.

Region	Code	Fluctuation percentile
R1	00964	5th
	04970	50th
	03950	95th
R2	06339	5th
	02030	50th
	09583	95th
R2	06728	5th
	01693	50th
	01359	95th

Table 1: Decadal fluctuation percentiles associated with each of the five-digit file name codes for each of the three SI subregions.

References

Greene, A. M., L. Goddard, P. L. Gonzalez, A. V. Ines, and J. Chryssanthacopoulos, 2015: A climate generator for agricultural planning in southeastern south america. *Agr. Forest Meteorol.*, **203 (0)**, 217–228, doi:10.1016/j.agrformet.2015.01.008, URL `http://www.sciencedirect.com/science/article/pii/S016819231500009X`.

Greene, A. M., M. Hellmuth, and T. Lumsden, 2012: Stochastic decadal climate simulations for the Berg and Breede Water Management Areas, Western Cape province, South Africa. *Water Resourc. Res.*, **48**, W06504, doi:10.1029/2011WR011152.

Harris, I., P. D. Jones, T. J. Osborn, and D. H. Lister, 2013: Updated high-resolution grids of monthly climatic observations – the CRU TS3.10 Dataset. *Int. J. Climatol.*, **34 (3)**, 623–642, doi:10.1002/joc.3711.

Holden, K., 1995: Vector autoregression modelling and forecasting. *J. Forecasting*, **14**, 159–166.

Rajagopalan, B. and U. Lall, 1999: A k-nearest-neighbor simulator for daily precipitation and other weather variables. *Water Resour. Res.*, **35 (10)**, 3089–3101.

Ruane, A., R. Goldberg, and J. Chryssanthacopoulos, 2015: AgMIP climate forcing datasets for agricultural modeling: Merged products for gap-filling and historical climate series estimation. *Agr. Forest Meteorol*, **200**, 233–248, doi:10.1016/j.agrformet.2014.09.016.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl, 2011: An overview of CMIP5 and the experiment design. *Bull. Am. Met. Soc.*, doi:10.1175/BAMS-D-11-00094.1.