LDRD precipitation

Abigail Snyder

Robert Link

Ben Bond-Lamberty

Corinne Hartin

ben.kravitz

February 6, 2018

Abstract

We propose a collaboration spanning earth science, mathematics, and statistics to extend a recently developed method for emulating Earth System Models (ESMs). The current method generates random realizations of global time series for *temperatures*. These realizations retain the spatial and temporal variance and covariance structures of the original input ESM temperature data, while running at a much lower computational cost compared to the ESMs being emulated.

The project described in this proposal will extend and adapt this technique to include *precipitation* as an output. Like the temperature case, this technique will capture the spatial and temporal covariance structure of the precipitation field. Additionally, the method will capture space and time correlation *between* temperature and precipitation.

Key staff Abigail Snyder (postdoc), Robert Link, Ben Bond-Lamberty, Corinne Hartin, Ben Kravitz

Motivation and background Two currently important topics in the scientific community are the effects of extreme events and uncertainty in climate impacts. Such questions are of interest to researchers in earth sciences, climate change, integrated assessment, and uncertainty quantification. Understanding the uncertainty around extreme events also has direct relevance for policymakers, the U.S. energy grid, and other stakeholders. Nowhere is this more important than in the water cycle, for which the precise distribution of precipitation over space and time is a primary driver of extreme events like drought and flooding and a key contributor to the energy-water-land nexus through its impact on agricultural performance.

Studying these topics requires large ensembles of high-resolution future climate scenarios, which are generally too expensive to produce directly by running Earth System Models (ESMs). Climate model emulators attempt to solve this problem by approximating the output a climate model *would have* produced had it been run for a specified scenario. These emulators are cheap to run, but capture only the mean response of the climate variable being modeled and little to none of the variability that would be present in a real ESM output (Mitchell et al., 1999; Mitchell, 2003; Tebaldi and Arblaster, 2014; MacMartin and Kravitz, 2016; Holden and Edwards, 2010; Neelin et al., 2010). Attempts to add variability to such emulations often lose important spatial and temporal correlations that fundamentally define ESM patterns and are critically important for understanding extreme events (such as the El Niño–Southern Oscillation), or else cannot produce the number of realizations necessary for extreme event or uncertainty studies (Castruccio and Stein, 2013; Osborn et al., 2015).

In recently completed work we have solved many of these problems for the case of global *temperature*. Our current method generates random realizations of global time series for temperatures that retain the spatial and temporal variance and covariance structures of the original input ESM data at a much lower computational cost compared to the ESMs being emulated. We have not yet demonstrated its robustness to use on other variables such as *precipitation*. Like the temperature case, this technique will capture the spatial and temporal covariance structure of the precipitation field. Additionally, the method will capture space and time correlation *between* temperature and precipitation.

Due to the mathematical details of our method, we expect there to be some adjustments necessary for the method to work for precipitation. Given the serious impacts of extreme precipitation events, we believe the additional work to extend our method will be valuable to researchers in a variety of fields housed in EBSD. Extending and testing our statistical approach to precipitation is perfectly suited for this LDRD call because it will enhance the impact of our established method, ultimately resulting in a method providing *joint realizations* of temperature and precipitation global time series. Preserving the space and time correlation between temperature and precipitation that is present in ESM outputs results in more realistic, and therefore more useful, realizations than would be achieved by arbitrarily pairing independent temperature and precipitation realizations.

Scientific basis and technical approach We will adapt our method for generating thousands of spatially spatially and temporally coherent realizations of global gridded temperatures in minutes (Link et al., 2018). The procedure fits a mean response model to input data and then generates thousands of new residual fields via empirical orthognal vector decomposition, discrete Fourier Transforms, and random phase perturbations, adding each residual field to the mean response to create a new temperature field. This ensures that the generated temperature fields have the same statistical properties as the ESM input used to train the emulator. At the same time, the realizations are statistically independent of one another and of the input data, allowing the generated data to express extremes that may not have been present in the training data. Note, however, that because the statistical properties of the ESM are inherited by our generated realizations, the extremes present in our realizations *could have been* produced in enough (expensive) runs of the ESM, and are therefore realistic.

Adapting this technique to produce realizations of precipitation global time series will require some additional work to account for the unique characteristics of the precipitation data, particularly its highly nonnormal distribution. Our method outputs normally distributed realizations of residuals, which would be inappropriate to add directly to mean precipitation responses. We propose to address this by transforming precipitation residuals to normally distributed Z-scores, applying our method, and inverting the Z-score transformation on the realizations before adding to the mean precipitation response.

There will also be software development required to collate the temperature and precipitation data (as they are provided separately) and to rearrange them into a form suitable for analysis. We will then develop and implement statistical tests suitable for testing the joint statistical properties of the temperature and precipitation fields and comparing them to the ESM input.

Summary of expected outcome This project will derive a computationally-efficient method to generate random joint time series of global precipitation and temperature at grid resolution. On average, across realizations, each generated time series will have the same variance and time, space, and inter-variable correlation structure as the ESM time series used to train the system. The method will be implemented as a package in the R programming language and released as open source software.

Timeline to expected manuscript Based on our experience developing the temperature field generation method, we expect to submit a manuscript on the extended technique in the summer of 2018.

Impacts for science This technique provides a computationally feasible way to produce the large ensembles of high-resolution future climate scenarios needed for studying the effects of extreme events and uncertainty in climate impacts. Furthermore, the lessons learned from extending the temperature method to precipitation will allow the method to be further extended to variables such as atmospheric aerosols or even globally gridded agricultural yield projections (e.g., to model bread basket failures).

Impacts for PNNL The papers and software produced by this project and anticipated follow on projects will help PNNL maintain its position as a leader in the fields of earth system modeling, uncertainty quantification, and integrated assessment modeling. The software produced will be directly useful to PNNL researchers interested in regional effects of future climate. Moreover, the results of this project will supplement ongoing PNNL research into applying machine learning techniques to climate model emulation and climate prediction.

No other funding The current JGCRI SFA is funded to *use* the software and data products described above, but not to *create* them. It would be very difficult to produce work of our desired high quality with the currently existing data methods and products. A successful outcome of this proposal could lead to expanded funding to explore the developments described above.

Estimated cost ACS: 56.35 x 80 hours = 4808.00
RPL: 137.32 x 40 hours = 5492.80
BBL: 137.32 x 20 hours = 2746.40
CH: 119.00 x 20 hours = 2380.00
BK: 119.00 x 20 hours = 2380.00
Total: 17507.20

References

S. Castruccio and M. Stein. Global space-time models for climate ensembles. *The Annals of Applied Statistics*, 7(3):1593–1611, 2013.

P. B. Holden and N. R. Edwards. Dimensionally reduced emulation of an AOGCM for application to integrated assessment modelling. *Geophysical Research Letters*, 37(21), 2010. doi: 10.1029/2010GL045137. URL http://dx.doi.org/10.1029/2010GL045137.

R. Link, C. Hartin, C. Lynch, B. Kravitz, B. Bond-Lamberty, and A. Snyder. Computationally Efficient Emulators for Earth System Models I: Producing Internal Variability. in preparation, 2018. in preparation.

D. G. MacMartin and B. Kravitz. Dynamic climate emulators for solar geoengineering. *Atmospheric Chemistry and Physics*, 16(24):15789–15799, 2016. doi: 10.5194/acp-16-15789-2016. URL http://www.atmos-chem-phys.net/16/15789/2016/.

J. Mitchell, T. C. Johns, M. Eagles, W. J. Ingram, and R. A. Davis. Towards the construction of climate change scenarios. *Climatic Change*, 41:547–581, 1999.

Timothy D. Mitchell. Pattern Scaling: An Examination of the Accuracy of the Technique for Describing Future Climates. *Climatic Change*, 60(3):217–242, 2003. doi: 10.1023/A:1026035305597. URL http://dx.doi.org/10.1023/A:1026035305597.

J David Neelin, Annalisa Bracco, Hao Luo, James C McWilliams, and Joyce E Meyerson. Considerations for parameter optimization and sensitivity in climate models. *Proceedings of the National Academy of Sciences*, 107(50):21349–21354, 2010.

Timothy J. Osborn, Craig J. Wallace, Ian C. Harris, and Thomas M. Melvin. Pattern scaling using ClimGen: monthly-resolution future climate scenarios including changes in the variability of precipitation. *Climatic Change*, pages 1–17, 2015. doi: 10.1007/s10584-015-1509-9. URL http://dx.doi.org/10.1007/s10584-015-1509-9.

Claudia Tebaldi and Julie M. Arblaster. Pattern scaling: Its strengths and limitations, and an update on the latest model simulations. *Climatic Change*, 122(3):459–471, 2014. doi: 10.1007/s10584-013-1032-9. URL http://dx.doi.org/10.1007/s10584-013-1032-9.