UC Irvine
UC Irvine Previously Published Works

Title
Empirical orthogonal function analysis of the diurnal cycle of precipitation in a multi-scale climate model

Permalink
https://escholarship.org/uc/item/00k0t08w

Journal
Geophysical Research Letters, 36(5)

ISSN
0094-8276

Authors
Pritchard, Michael S
Somerville, Richard C. J

Publication Date
2009-03-11

DOI
10.1029/2008GL036964

Copyright Information

Peer reviewed

Empirical orthogonal function analysis of the diurnal cycle of precipitation in a multi-scale climate model

Michael S. Pritchard[1] and Richard C. J. Somerville[1]

[1] Long-term variability in the hydrologic cycle is poorly simulated by current generation global climate models (GCMs), partly due to known climatological biases at shorter timescales. We demonstrate that a prototype Multi-scale Modeling Framework (MMF) provides a superior representation of the spatial and temporal structure of precipitation at diurnal timescales than a GCM. Results from empirical orthogonal function (EOF) decomposition of the boreal summer climatological composite diurnal cycle of precipitation in an MMF are compared to a GCM and satellite data from the Tropical Rainfall Measuring Mission. The eigenspectrum, principal component time series, and the spatial structure of leading EOFs in an eigenmode decomposition of the MMF composite day are a much better match to observations than the GCM. Regional deficiencies in the MMF diurnal cycle are manifest as localized anomalies in the spatial structures of the first two leading EOFs. **Citation:** Pritchard, M. S., and R. C. J. Somerville (2009), Empirical orthogonal function analysis of the diurnal cycle of precipitation in a multi-scale climate model, *Geophys. Res. Lett.*, *36*, L05812, doi:10.1029/2008GL036964.

1. Introduction

[2] At a time when reliable projections about the hydrologic response to anthropogenic climate change are in demand by many stakeholders, a climate modeling framework that accurately represents the physical drivers of moist convection on multiple time scales is needed. The usual tools for climate forecasting, global climate models (GCMs), do not simulate realistic diurnal hydrologic variability [*Collier and Bowman*, 2004]. Diurnal precipitation in GCMs is too sinusoidal, peaks too early over continents and is too horizontally homogenous where spatial variations are observed in the phase and amplitude of the diurnal cycle in nature. Hydrologic biases at diurnal time scales can distort climate energetics on longer time scales [*Bergman and Salby*, 1997], casting doubt on GCM projections of future hydrologic variability.

[3] A new approach to climate modeling offers dramatic improvement in simulating diurnal hydrologic processes. Multi-scale Modeling Frameworks (MMFs) are GCMs that treat clouds and related sub-grid processes in a new way, using interactive embedded cloud resolving models instead of statistical approximations. Exploration of the diurnal cycle in MMFs has revealed:

[4] 1. Improved timing of maximum precipitation over continents [*Khairoutdinov et al.*, 2005].

[5] 2. Improved cross-coastal and intra-continental variability in the 24-hour harmonic and broadness of the mean summer day's precipitation (M. S. Pritchard and R. C. J. Somerville, Assessing the diurnal cycle of precipitation in a multi-scale climate model, submitted to *Journal of Advances in Modeling Earth Systems*, 2008).

[6] 3. Realistic relationships between the diurnal variability of cloud liquid water, longwave cooling, vertical velocity variance, inversion height and sub-cloud vertical velocity skewness along a transect displaying a range of cloud types [*Khairoutdinov et al.*, 2008].

[7] 4. Biases in the diurnal variability of longwave radiation proxies for precipitation [*Zhang et al.*, 2008].

[8] 5. Realistic diurnal precipitation variability over the southeastern United States, Gulf Stream, and western Atlantic, but no propagating organized convection over the central United States (Pritchard and Somerville, submitted manuscript, 2008).

[9] In this study we present new results demonstrating an overall improvement in the diurnal cycle of precipitation in an MMF relative to a conventional GCM, as diagnosed by empirical orthogonal function (EOF) decomposition of the climatological mean summer day. This EOF approach was recently advocated as a benchmark test for evaluating simulated hydrologic diurnal variability against new space-borne precipitation observations [*Kikuchi and Wang*, 2008]. Our results indicate that the MMF passes this test - the eigenstructure of the MMF's mean summer day's precipitation is more faithful to the observations than in conventional GCMs.

2. Background

[10] MMFs are a response to a long-standing climate modeling dilemma - the fact that cloud processes cover a range of spatial and temporal scales that is too large for modern supercomputers to handle [*Randall et al.*, 2003]. The compromise in conventional GCMs is to use a crude statistical representation, or "parameterization", of the collective influence of physical processes below a certain truncation scale in space and time (about 100 km and 15 minutes respectively). In MMFs, parameterizations of sub-grid cloud processes are replaced by nested cloud-resolving model (CRM) integrations within small high resolution subdomains housed in each host GCM grid column. Statistics harvested from the CRM integration replace the conventional parameterizations for estimating sub-grid fluxes of heat and moisture. MMFs are hundreds of times more expensive to run than GCMs, but are quite scaleable on parallel supercomputers [*Khairoutdinov et al.*, 2005].

[1]Graduate Department, Scripps Institution of Oceanography, University of California, San Diego, La Jolla, California, USA.

[11] The prototype MMF used in this study is under development at the Center for Multiscale Modeling of Atmospheric Processes (CMMAP), and is described by *Khairoutdinov et al.* [2005]. It is identical to the National Center for Atmospheric Research Community Atmosphere Model v3.0 (CAM3) [*Collins et al.*, 2006] except that the parameterizations for deep and shallow sub-grid convection, clouds, and boundary layer variability have been replaced with a nested integration of a two dimensional $(x - z)$, laterally periodic realization of the CRM described by *Khairoutdinov and Randall* [2003]. The CRM solves the non-hydrostatic momentum equations of fluid dynamics subject to the anelastic approximation, using bulk micro-physics to track conversions between five categories of interactive precipitating and non-precipitating prognostic water condensate variables [*Khairoutdinov and Randall*, 2003].

[12] The simulations analyzed in this paper are four month (MJJA) boreal summer integrations of the MMF and CAM driven by monthly climatological sea surface temperatures. Only results from JJA are shown. In the MMF, the CRM time step is 20 seconds, and multi-scale coupling to the large-scale atmosphere and land surface model occurs on the 15 minute time step of the host GCM. We run the host GCM at T42 horizontal resolution ($2.8° \times 2.8°$) and configure the CRM to run on 32 nested grid columns spaced zonally 4 km apart on 26 vertical levels co-located with the interior of the GCM vertical grid. At this resolution the majority of vertical moisture transport in the CRM is explicitly resolved, with sub-grid transport responsible for less than 10% of zonal mean moisture fluxes [*Khairoutdinov et al.*, 2008].

[13] We use seven years of TRMM 3B42 data (2000-2006) as an observational baseline. TRMM 3B42 is a gridded ($0.25° \times 0.25° \times 3$ hours) dataset extending from 50 S to 50 N. This product is a "best estimate'" of precipitation that combines observations from several space-borne instruments, including the precipitation radar and microwave radiometer on board TRMM as well as infrared radiometers from other platforms [*Huffman et al.*, 2007]. The climatological composite diurnal cycle in TRMM 3B42 has been validated against independent surface radar and rain gauge data [*Dai et al.*, 2007].

3. Method

[14] Independent patterns of variability in the mean summer day's precipitation that account for high amounts of statistical variance can be identified using empirical orthogonal function (EOF) decomposition [*von Storch and Zwiers*, 1999].

[15] We apply EOF decomposition to a regularly gridded local solar time- (LST-) space matrix of the diurnal precipitation anomaly about its daily mean. Construction of this matrix is complicated by the fact that both the model output and observations are discretized at regular intervals in universal time (UTC), which means that discretization in local solar time (LST) varies as a function of longitude. Following *Kikuchi and Wang* [2008], we apply Fourier interpolation along the periodic LST dimension after conversion from UTC, to re-discretize LST consistently at each longitude.

[16] Special attention must be paid to estimating uncertainty in the eigenvalue spectrum, since EOF patterns and their corresponding principal component (PC) time series provide meaningful information only if the uncertainty in their associated eigenvalues is sufficiently small that adjacent EOFs are not at risk of statistical degeneracy [*North et al.*, 1982]. Usually uncertainty in the eigenvalue spectrum is estimated using the "significance test" posed by *North et al.* [1982]. However this technique requires some knowledge of the degrees of freedom (DOF) in the space-time dataset [*Wallace et al.*, 1992] and introduces questionable assumptions about the linearity of error propagation. As *Kikuchi and Wang* [2008] point out, estimating the DOF in a climatological composite daily precipitation matrix is not at all straightforward: There are at least as many DOF as there are independent time samples in the mean summer day, but surely fewer DOF than there are total time samples in all of the days that went into the composite (92 for the model observations, 644 for the satellite data). *Kikuchi and Wang* [2008] chose to apply North's significance rule assuming 50 DOF, arguing this was a conservative choice between the former (8 time samples in the TRMM composite day) and the latter (many thousand time samples in the composite mean).

[17] We choose instead to estimate uncertainties in the eigenvalue spectra of the mean summer day using a statistical resampling technique. Our approach combines random re-sampling (bootstrapping) and under-sampling (jackknifing) of the composite diurnal cycle of precipitation. Repeated EOF decomposition operations are carried out on 2000 randomly populated subsets of the full mean summer day space-time matrix, and standard errors are estimated based on the resulting ensemble of eigenvalue spectra. This method is faithful to the nonlinearity of sampling error propagation through the sequence of EOF matrix operations. We choose to *undersample* the full space-time matrix because there is significant correlation between spatially adjacent grid points' diurnal variability. We limit the spatial dimension of each randomly populated bootstrap ensemble member's space-time matrix to a fraction f of the N spatial grid points x_i. The degree of undersampling is estimated as

$$f = \frac{1}{N} \sum_{1}^{N} f_i(R_i) \qquad (1)$$

where R_i is the cross-correlation map of the x_i-th point's mean summer day time series with every other point in space, and f_i is the fractional area occupied by the $R_i = 0.8$ contour that encloses the x_i-th point in space. Based on correlation analysis of TRMM 3B42 precipitation rate data bin-averaged to the models' T42 grid, we estimate $f \approx 7\%$.

4. Results

[18] Figure 1 shows the ensemble mean (left plot) and PDFs (right plots) of the 2000 eigenspectra computed as described above. The eigenspectrum for the MMF composite boreal summer day's precipitation is improved relative to CAM. In the observations, two primary EOFs are distinguishable in the ensemble, accounting for $63.0 \pm 10.4\%$ and $26.1 \pm 8.5\%$ of the variance in the TRMM 3B42 composite boreal summer day, respectively (note errors are the stan-
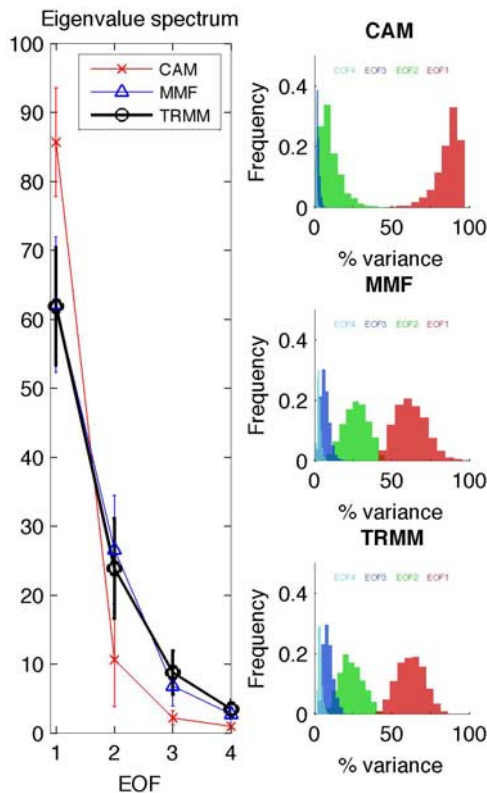
Figure 1. (left) Eigenvalue spectrum for the composite JJA diurnal cycle of precipitation as simulated by CAM and MMF, and as observed in TRMM 3B42. (right) Normalized histograms of the percent variance attributable to the first four EOFs in the 2000-member bootstrapping ensemble of EOF calculations, for each model; error bars in the eigenvalue spectrum are computed from these distributions.

dard deviation of the full ensemble). This is consistent with the findings of *Kikuchi and Wang* [2008]. The eigenspectrum for the MMF mean summer day is remarkably similar to the observations, with $61.7 \pm 8.6\%$ of the variance in EOF_1, and $24.1 \pm 7.2\%$ in EOF_2. But in CAM the eigenstructure of diurnal precipitation is distorted - the leading EOF explains over 85% of the variance, and higher modes are statistically indistinguishable from each other.

[19] Figure 2a shows corresponding improvement in the structure and amplitude of the MMF bootstrap ensemble mean principal component (PC) time series for EOF_1 relative to CAM. Note that the precipitation variability associated with each eigenmode is the product of the PC time series with its spatial EOF. Both the MMF and TRMM PC1 time series are quasi-sinusoidal with a minimum around 0900 LST; the maximum of PC1 occurs at 2100 LST in the MMF, compared to 1800 LST in TRMM. The PC1 time series of the leading EOF in CAM has a very different structure.

[20] Figures 2b−2d compares the spatial structure of the leading EOF in the models and the observations. In the observations, EOF_1 is mostly positive over land and negative over the oceans. This mode captures the phase difference of marine vs. continental diurnal precipitation cycles, with a tendency for higher amplitudes in the tropics. Although this broad structure is also present in EOF_1 for

both MMF and CAM, the amplitude of EOF_1 is dramatically improved in the MMF. The MMF captures observed continental maxima in EOF_1 over northeastern Brazil, Central America, central Africa, and Thailand as well as oceanic minima in the tropics, over the Atlantic and Pacific storm tracks, and the South Pacific Convergence Zone (SPCZ). But the MMF EOF_1 is not perfect: the (positive) magnitude of EOF_1 is underestimated over the central United States, western equatorial Africa, and central Brazil and the (negative) magnitude of EOF_1 is overestimated over the oceans. The MMF also exhibits excessive EOF_1 amplitude in the vicinity of the Indian monsoon and over the Western Pacific. But in comparison to CAM, which has excessive EOF_1 amplitude over all continental land masses, the global structure of the leading diurnal EOF in the MMF is improved.

[21] Figure 3a shows corresponding improvement in the MMF EOF_2 principal component time series (PC2). As in the observations, PC2 in the MMF has a broad local
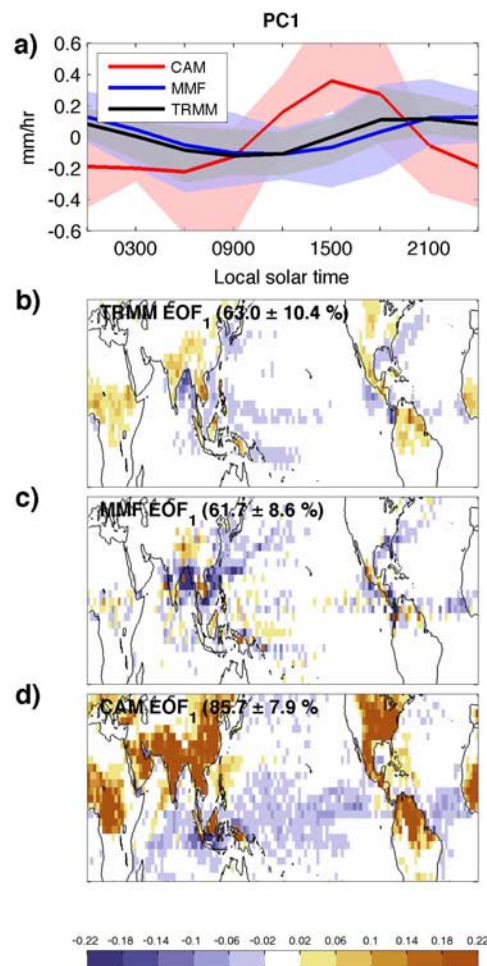
Figure 2. (a) Principle component (PC) time series of the leading EOF of the composite boreal summer diurnal cycle of precipitation as simulated by CAM (red) and MMF (blue), compared to TRMM observations (black). Solid lines and shading denote the mean ± one standard deviation of a 2000 member ensemble of randomly undersampled bootstrap EOF calculations. Ensemble mean spatial structure of EOF_1 for (b) TRMM data, (c) the MMF, and (d) CAM.
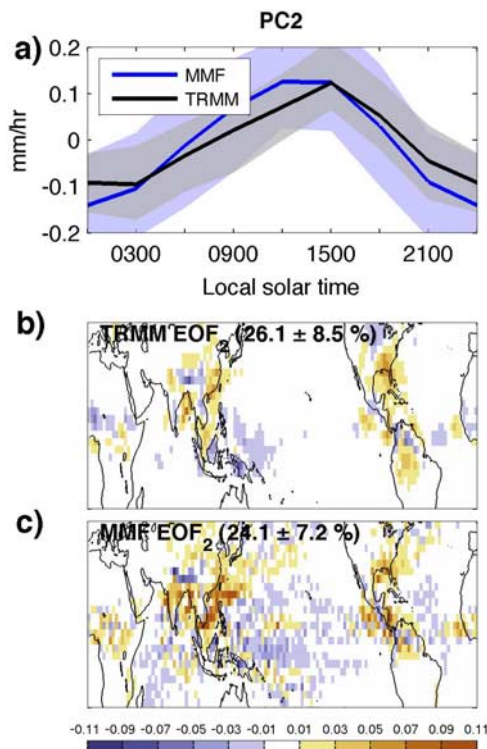
Figure 3. As in Figure 2, but for EOF$_2$.

maximum peaking at 1500 LST, and a minimum from 0000-0300 LST. The PC2 time series in CAM is not shown, since EOF$_2$ in CAM is statistically degenerate with subsequent EOFs (Figure 1).

[22] Figures 3b and 3c show that the spatial structure of EOF$_2$ in the MMF is broadly similar to the observations, but is far from perfect. Local maxima in EOF$_2$ over the southeastern United States, northeastern South America, and northeastern Asia are reasonably well reproduced by the MMF. However, the MMF does not capture observed minima in the spatial structure of EOF$_2$ over the central United States and equatorial Africa. As for EOF$_1$, the amplitude of oceanic minima in EOF$_2$ is exaggerated in the MMF, as is the EOF$_2$ amplitude over the Western Pacific and Indian monsoon region.

[23] Taken together, Figures 1−3 are compelling evidence of improved diurnal hydrologic variability in the MMF relative to CAM. The correspondence of a realistic eigenspectrum and PC time series in the MMF with spatial structures of the leading EOFs that are in broad agreement with observations strongly suggests overall improved performance at diurnal time scales. However regional biases in the spatial projections of EOF$_1$ and EOF$_2$ demonstrates that in certain parts of the globe, the MMF diurnal cycle needs improvement.

[24] These findings are consistent with complementary diagnostics of diurnal variability in the MMF. Pritchard and Somerville (submitted manuscript, 2008) applied a suite of diurnal cycle diagnostics to evaluate regional features of the MMF's diurnal cycle of precipitation. Harmonic analysis indicated that too much of the variance in the CAM diurnal cycle could be explained by a single 24-hour harmonic, just as EOF analysis indicates that too much of the variance in

CAM is in a single EOF. The amplitude of the 24-hour harmonic in the MMF is less than in the CAM and more in line with observations. An analogous improvement (reduction) is also evident in the simulated diurnal range in the MMF (not shown). Harmonic analysis also showed excessive diurnal amplitude in the MMF over its superactive Indian Monsoon. Similar biases can be seen in the spatial structure of the leading EOFs over this region. Animation of the full spatial and temporal structure of the mean summer day's precipitation over North America indicated that the MMF is unable to capture observed diurnal eastward propagation of organized convection over the central United States but that overall the MMF diurnal cycle seems well resolved over the eastern seaboard and Gulf Stream. These facts are consistent with regions of good and poor agreement in the spatial structure of the leading EOFs over North America in the MMF. EOF analysis thus provides a consistent and compact framework for evaluating simulated model biases on diurnal timescales.

5. Conclusions

[25] The development of Earth system models capable of reproducing observed hydrologic variability on diurnal time scales is a high priority and an important step towards making projections about hydrologic variability on longer time scales. Using empirical orthogonal function (EOF) decomposition of the climatological composite boreal summer diurnal cycle of rainfall, we have demonstrated that a prototype multi-scale modeling framework (MMF) outperforms a conventional global climate model (GCM) at simulating the diurnal cycle of precipitation.

[26] The eigenspectrum for the MMF mean summer day, as well as the principal component time series of its two leading EOFs, are a remarkable match to observations. This is very promising. The spatial structure of EOF$_1$ and EOF$_2$ in the MMF are in broad agreement with observations, but exhibit several regional biases consistent with areas of known diurnal cycle problems in the model. MMFs are still in their infancy, and further work is needed to correct these regional biases, which often occur at both diurnal and seasonal time scales (not shown).

[27] We find EOF decomposition to provide a meaningful and compact way to appraise the overall space-time variability of simulated climatological diurnal variability in climate models. We support the recent recommendation of *Kikuchi and Wang* [2008] that this approach be used as a litmus test to evaluate the simulated diurnal cycle of precipitation in climate models against gridded precipitation products.

References

Bergman, J. W., and M. L. Salby (1997), The role of cloud diurnal variations in the time-mean energy budget, *J. Clim.*, *10*(5), 1114–1124.

Collier, J. C., and K. P. Bowman (2004), Diurnal cycle of tropical precipitation in a general circulation model, *J. Geophys. Res.*, *109*, D17105, doi:10.1029/2004JD004818.

Collins, W., et al. (2006), The formulation and atmospheric simulation of the Community Atmosphere Model version 3 (CAM3), *J. Clim.*, *19*(11), 2144–2161.

Dai, A. G., X. Lin, and K. L. Hsu (2007), The frequency, intensity, and diurnal cycle of precipitation in surface and satellite observations over low- and mid-latitudes, *Clim. Dyn.*, *29*(7–8), 727–744.

Huffman, G., R. Adler, D. Bolvin, G. Gu, E. Nelkin, K. Bowman, Y. Hong, E. Stocker, and D. Wolff (2007), The TRMM multisatellite precipitation analysis (TMPA): Quasi-global, multiyear, combined-sensor precipitation estimates at fine scales, *J. Hydrometeorol.*, *8*(1), 38–55.

Khairoutdinov, M. F., and D. A. Randall (2003), Cloud resolving modeling of the ARM summer 1997 IOP: Model formulation, results, uncertainties, and sensitivities, *J. Atmos. Sci.*, *60*(4), 607–625.

Khairoutdinov, M., D. Randall, and C. DeMott (2005), Simulations of the atmospheric general circulation using a cloud-resolving model as a super-parameterization of physical processes, *J. Atmos. Sci.*, *62*(7), 2136–2154.

Khairoutdinov, M., C. DeMott, and D. Randall (2008), Evaluation of the simulated interannual and subseasonal variability in an AMIP-style simulation using the CSU multi-scale modeling framework, *J. Clim.*, *21*(3), 413–431.

Kikuchi, K., and B. Wang (2008), Diurnal precipitation regimes in the global tropics, *J. Clim.*, *21*(11), 2680–2696.

North, G. R., T. L. Bell, R. F. Cahalan, and F. J. Moeng (1982), Sampling errors in the estimation of empirical orthogonal functions, *Mon. Weather Rev.*, *110*(7), 699–706.

Randall, D., M. Khairoutdinov, A. Arakawa, and W. Grabowski (2003), Breaking the cloud parameterization deadlock, *Bull. Am. Meteorol. Soc.*, *84*(11), 1547–1562.

von Storch, H., and F. W. Zwiers (1999), *Statistical Analysis in Climate Research*, chap. 13, pp. 293–312, Cambridge Univ. Press, Cambridge, U. K.

Wallace, J. M., C. Smith, and C. S. Bretherton (1992), Singular value decomposition of wintertime sea-surface temperature and 500-mb height anomalies, *J. Clim.*, *5*(6), 561–576.

Zhang, Y., S. A. Klein, C. Liu, B. Tian, R. T. Marchand, J. M. Haynes, R. B. McCoy, Y. Zhang, and T. P. Ackerman (2008), On the diurnal cycle of deep convection, high-level cloud, and upper troposphere water vapor in the Multiscale Modeling Framework, *J. Geophys. Res.*, *113*, D16105, doi:10.1029/2008JD009905.

————————————

M. S. Pritchard and R. C. J. Somerville, Scripps Institution of Oceanography, University of California, San Diego, Mail Code 0224, 9500 Gilman Drive, La Jolla, CA 92093-0224, USA. (mikepritchard@ucsd.edu)