

Potential seasonal predictability: Comparison between empirical and dynamical model estimates

Timothy DelSole,^{1,2} Arun Kumar,³ and Bhaskar Jha³

Received 9 April 2013; revised 21 May 2013; accepted 22 May 2013; published 18 June 2013.

[1] Methods for estimating potential seasonal predictability from a single realization of daily data are validated against an ensemble of simulations from an atmospheric model driven by observed sea surface temperature, sea ice extent, and greenhouse gas concentration. The methods give surprisingly good estimates of potential predictability of seasonal precipitation despite the fact that the methods assume Gaussian distributions. For temperature, the methods systematically underestimate weather noise variance over land, often by a factor of 2 or more. This bias can be reduced by taking account of precipitation-induced variability. These conclusions may be model dependent, and hence, confirmation in other models would be of interest. Nevertheless, for the state-of-the-art atmospheric model used in this study, the results strongly support the validity of the single time series approach to estimating potential predictability and enhances our confidence in previous estimates of potential predictability based on observations alone. **Citation:** DelSole, T., A. Kumar, and B. Jha (2013), Potential seasonal predictability: Comparison between empirical and dynamical model estimates, *Geophys. Res. Lett.*, 40, 3200–3206, doi:10.1002/grl.50581.

1. Introduction

[2] The atmosphere varies naturally independently of changes in radiative forcing or boundary conditions. This variability, called *weather noise*, arises from baroclinic and convective growth mechanisms and is predictable from days to weeks. Other components of the climate system, such as sea surface temperature (SST), soil moisture, or sea ice thickness, are predictable beyond a month owing to their longer memory. If these slower (“boundary”) components influence the atmosphere independently of weather noise, then they add variability that is predictable beyond a month. The difference between the actual variability and the variability expected from weather noise is identified with *potential predictability*.

[3] Two methods exist for estimating potential predictability. The first is to estimate it from an ensemble of

atmospheric model runs with the same boundary conditions and external forcing [Chervin, 1986; Kumar and Hoerling, 1995; Phelps *et al.*, 2004]. This ensemble-dynamical approach is limited by the ability of the dynamical model to capture all dynamical processes relevant to the potential predictability. The second is to estimate it from a single realization of a process [Madden, 1976; Zwiers, 1987; Zheng *et al.*, 2000; Feng *et al.*, 2011, 2012]. Empirical approaches are limited by the observational record and the validity of the assumptions used to model weather noise and boundary forcing.

[4] Whether empirical and dynamical approaches give consistent results has yet to be demonstrated. Shea and Madden [1990] validated an empirical method using time series generated by an atmospheric model. However, their model was run with “perpetual January” sea surface temperatures, implying no potential predictability. The purpose of this paper is to validate empirical methods when potential predictability is known to exist.

2. Empirical and Dynamical Methods

[5] Various methods for estimating potential seasonal predictability from a single realization of a process have been discussed by Madden [1976], Shukla and Gutzler [1983], Zwiers [1987], Zheng *et al.* [2000], Feng *et al.* [2011, 2012], and DelSole and Feng [2013]. These methods assume that the state on the d th day in the y th year can be modeled as

$$X_{d,y} = \mu_y + \epsilon_{d,y}, \quad (1)$$

where $\epsilon_{d,y}$ is a stationary, stochastic process representing weather noise and μ_y is the change in mean due to slowly varying boundary conditions or external forcing. This model assumes that weather noise varies on timescales much shorter than a season and that the boundary and external forcing varies so slowly that they are effectively constant during a season. While these assumptions may be questioned, some distinction between the behavior of weather noise and external predictability must be assumed in order to separate them. Let $d = 1, \dots, D$, where D is the number of days in a season, and $y = 1, \dots, Y$, where Y denotes the total number of years (e.g., number of winters).

[6] In terms of model (1), the hypothesis of no potential predictability corresponds to

$$H_0 : \mu_1 = \mu_2 = \dots = \mu_Y. \quad (2)$$

If $\epsilon_{d,y}$ is independently and identically distributed as a Gaussian, then the standard method for testing hypothesis (2) is analysis of variance (ANOVA). Unfortunately, the weather noise term $\epsilon_{d,y}$ is correlated on daily timescales, rendering ANOVA inappropriate.

Additional supporting information may be found in the online version of this article.

¹Department of Atmospheric, Oceanic, and Earth Sciences, George Mason University, Fairfax, Virginia, USA.

²Center for Ocean-Land-Atmosphere Studies, Calverton, Maryland, USA.

³Climate Prediction Center, NCEP, NWS, NOAA, College Park, Maryland, USA.

Corresponding author: T. DelSole, Center for Ocean-Land-Atmosphere Studies, 4041 Powder Mill Rd., Suite 302, Calverton, MD, 20705, USA. (dels@cola.iges.org)

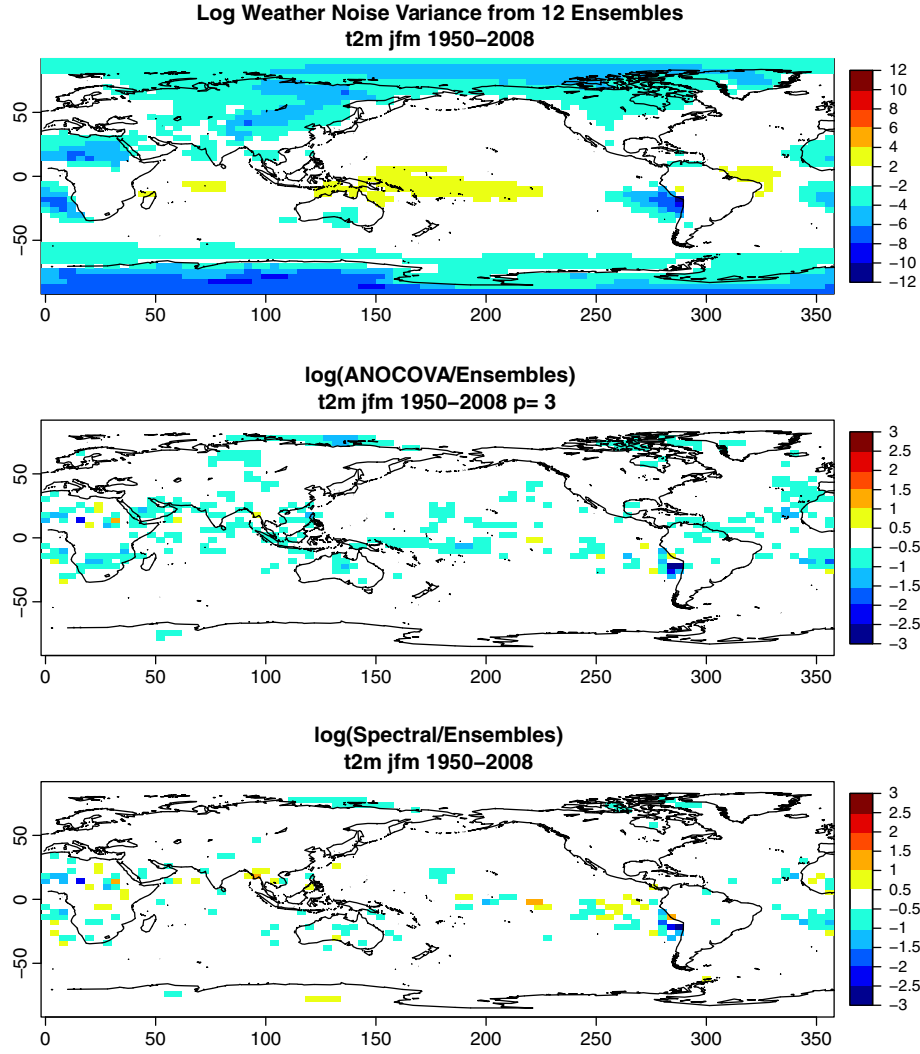


Figure 1. (Top) Logarithm of the weather noise variance of January–March mean precipitation (units: $(\text{mm/d})^2$) in an atmospheric model run with observed SST and external forcing during 1950–2008, estimated from 12 ensemble members. (Middle) The difference of weather noise log-variance between the variance calculated from a single realization of daily data using ANOCOVA, and the variance estimated from the 12 ensemble members. (Bottom) The difference in weather noise log-variance between the variance calculated from a single realization of daily data using spectral techniques, and the variance estimated from the 12 ensemble members.

2.1. Analysis of Covariance (ANOCOVA)

[7] *Feng et al.* [2012] suggested modeling the weather noise term $\epsilon_{d,y}$ by an autoregressive process. In this paper, we generalize the model slightly by including the annual cycle explicitly. Accordingly, we consider the model

$$X_{d,y} = \phi_1 X_{d-1,y} + \phi_2 X_{d-2,y} + \dots + \phi_p X_{d-p,y} + (\mu_y + a_1 d + a_2 d^2) + w_{d,y}, \quad (3)$$

where $w_{d,y}$ is the Gaussian white noise and $\phi_1, \phi_2, \dots, \phi_p, \mu_y, a_1, a_2$ are the parameters to be estimated from data. The terms involving a_1 and a_2 account for the annual cycle within a season. The model reduces to (1) when $\phi_1, \phi_2, \dots, \phi_p, a_1, a_2$ vanish.

[8] The hypothesis of no potential seasonal predictability is equivalent to (2) (where μ_y refers to the corresponding terms in (3)). The significance test for this hypothesis and the calculation of weather noise from the model is discussed in the auxiliary material but is essentially equivalent to that discussed in *Feng et al.* [2012].

2.2. Spectral Method

[9] *Madden* [1976] proposed a spectral approach to estimating weather noise variance. This method has been clarified by *Zwiers* [1987]. In particular, when the averaging period and spectral analysis period are the same, the weather noise variance reduces to $1/D$ times the power at the lowest frequency within the period [e.g., see *Zwiers*, 1987, equation (2.9)]:

$$\hat{\sigma}_{N,spec}^2 = \frac{1}{D^2 Y} \sum_{y=1}^Y \left(\left(\sum_{d=1}^D \cos(2\pi d/D) X_{d,y} \right)^2 + \left(\sum_{d=1}^D \sin(2\pi d/D) X_{d,y} \right)^2 \right). \quad (4)$$

We use $D = 90$ days, which is comparable to the value recommended by *Shea and Madden* [1990] based on atmospheric model simulations. The significance test for the hypothesis (2) is based on the F distribution and discussed in the supporting information.

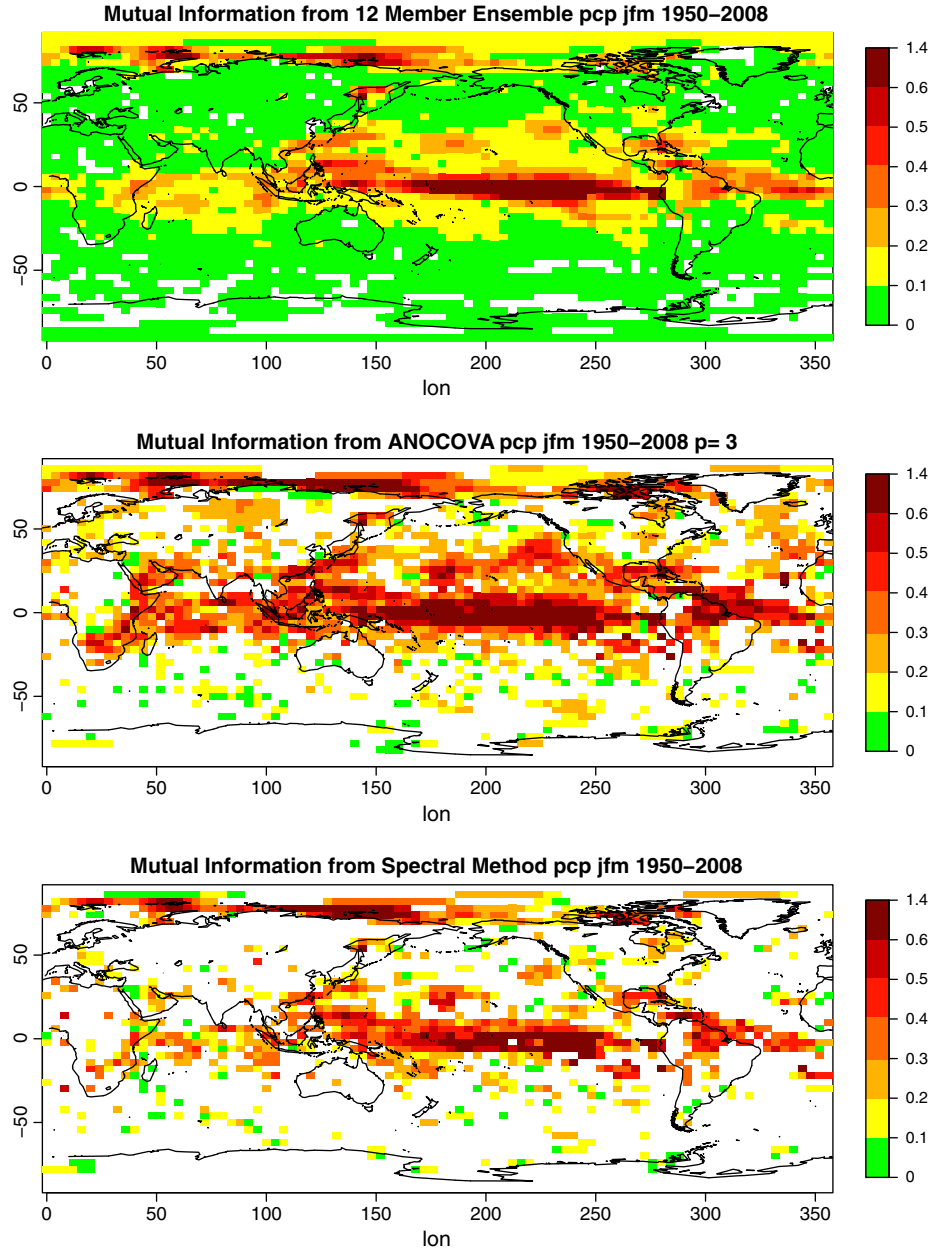


Figure 2. Mutual information of the potential predictability of JFM precipitation, as estimated directly (top) from the 12-member ensemble, (middle) by ANOCOVA, and (bottom) by a spectral method. Values insignificant at the 5% level, according to each respective method, are masked out; in particular, different masks are applied to the middle and bottom panels.

[10] Removal of the annual cycle is critical to the spectral method; otherwise, power spectra will be artificially inflated, thereby inflating weather noise variance. We approximate the annual cycle by a second-order polynomial in time, just as for ANOCOVA. Alternative definitions of the annual cycle also were explored, including using the first 2–5 annual harmonics estimated from the full 365 day year, and higher order polynomials in time, but the results (from both the spectral and ANOCOVA methods) were insensitive to the definition of the annual cycle. The advantage of using a low-order polynomial is that the uncertainty in estimating the annual cycle can be taken into account in ANOCOVA.

2.3. Ensemble Method

[11] If an ensemble of realizations for the same boundary and external forcing are available, such as from a dynamical model, then potential predictability can be estimated under much less restrictive assumptions than those used in empirical methods [Shukla, 1981; Zwiers, 1987, 1996]. Let $\bar{X}_{y,e}$ denote realizations of the seasonal mean for the same boundary and external forcing, where $e = 1, 2, \dots, E$ is an index denoting ensemble member. An unbiased estimate of weather noise variance can be calculated as

$$\hat{\sigma}_{N,ENS}^2 = \frac{1}{Y(E-1)} \sum_{e=1}^E \sum_{y=1}^Y \left(\bar{X}_{y,e} - \langle \bar{X} \rangle_y \right)^2, \quad (5)$$

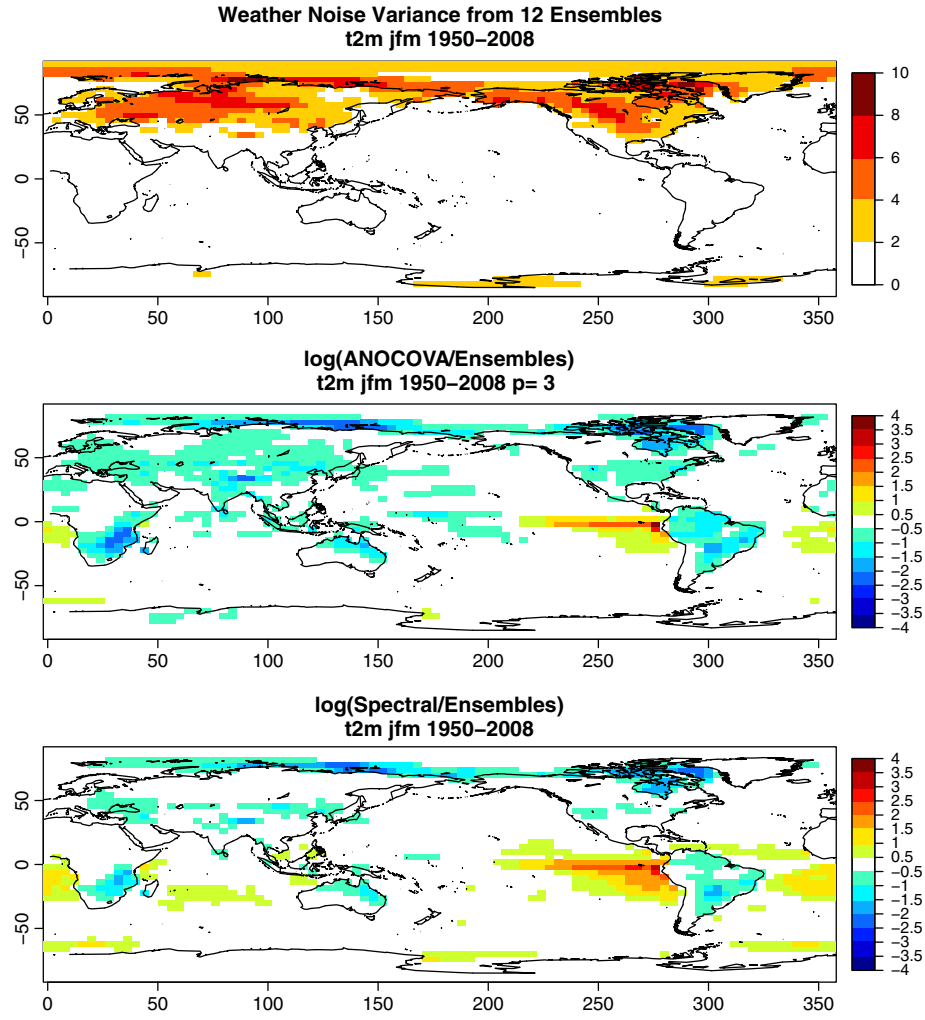


Figure 3. (Top) Logarithm of the weather noise variance of January - March mean 2m temperature (units: K²) in an atmospheric model run with observed SST and external forcing during 1950-2008, estimated from 12 ensemble members. (Middle) The difference of weather noise log-variance between the variance calculated from a single realization of daily data using ANOCOVA, and the variance estimated from the 12 ensemble members. (Bottom) The difference in weather noise log-variance between the variance calculated from a single realization of daily data using spectral techniques, and the variance estimated from the 12 ensemble members.

where the ensemble mean is defined as

$$\langle \bar{X} \rangle_y = \frac{1}{E} \sum_{e=1}^E \bar{X}_{y,e}. \quad (6)$$

The null hypothesis (2) can be tested by ANOVA, as discussed in the supporting information.

2.4. Differences Between Methods

[12] The above methods differ in their underlying assumptions. The spectral method assumes that the power spectrum is flat on timescales longer than a season. ANOCOVA assumes the whole power spectrum has a shape characteristic of a low-order AR model. In contrast to these methods, the ensemble-dynamical method, which is based on ANOVA, makes no assumptions about the shape of the power spectrum. All three methods assume that weather noise is Gaussian and independent of the potentially predictable signal.

3. Data

[13] The data used in this study come from the atmospheric component of the National Centers of Environmental Prediction Climate Forecast System version 2 model [Kumar *et al.*, 2012]. An ensemble of 12 simulations with a horizontal resolution of T126 (100 km) and 64 vertical levels extending from the surface to 0.26 hPa was run from 1950 to 2008. For each simulation, the same observed evolution of sea surface temperature, sea ice extent, and CO₂ concentration was specified as external forcings; however, each started with a different atmospheric initial condition. In addition to saving the monthly means of relevant variables for all 12 runs, daily data was saved for one of the simulations.

[14] The atmospheric general circulation model used here is imperfect and hence generates time series that differ from those of nature. This discrepancy is not relevant for validation purposes, since our goal is to test the consistency of techniques for time series generated by the same realistic model.

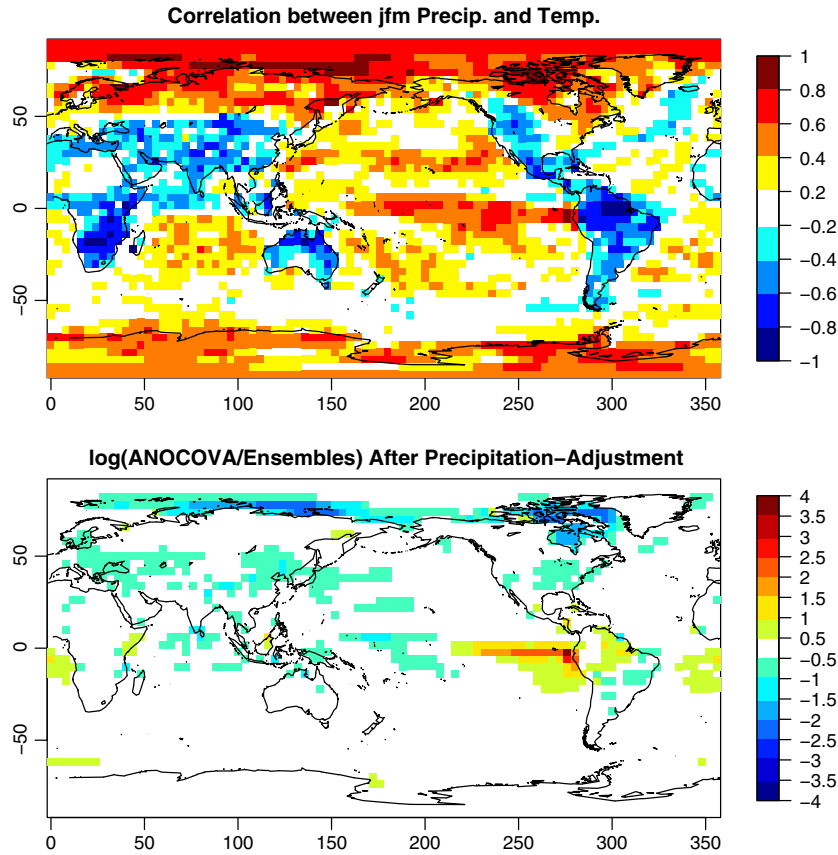


Figure 4. (top) Correlation between JFM precipitation and 2 m temperature in an atmospheric model run with observed SST and external forcing during 1950–2008. The correlations were computed from a single ensemble member. (bottom) Also shown is the log of the ratio of weather noise variance between the ANOCOVA and dynamical-ensemble estimates, after noise variance associated with precipitation variability has been added to the ANOCOVA estimate.

4. Results

[15] The weather noise variance of January–March (JFM) mean precipitation estimated directly from the 12-member ensemble is shown in Figure 1. The log of the weather noise variance is plotted because precipitation varies by orders of magnitude. The maximum weather noise variance occurs along the western equatorial Pacific, and minimum variance occurs over Antarctica and the western coasts of tropical South America and Africa. For comparison, weather noise variance is estimated empirically from the daily time series of a single ensemble member. The difference between empirical estimates and the 12-member ensemble estimates, expressed as the log of the ratio of weather noise variance, is shown in the bottom panels of Figure 1. The figure shows that the discrepancies are relatively small, with the largest discrepancies occurring in the tropics. The spectral method tends to give more accurate estimates than ANOCOVA. Despite the questionable use of Gaussian-based techniques, empirical estimates of potential seasonal predictability of precipitation agree fairly well with the 12-member ensemble estimates.

[16] Potential predictability depends on the amount by which total variance, σ_T^2 , exceeds weather noise variance, σ_N^2 . To measure potential predictability, we use mutual information:

$$M = \frac{1}{2} \log \frac{\sigma_T^2}{\sigma_N^2}. \quad (7)$$

The mutual information of JFM precipitation calculated from all three methods is shown in Figure 2. Note that the difference in mutual information between the empirical and ensemble estimates equals half the bottom panels of Figure 1, owing to the properties of logarithms and the fact that each noise variance is compared to the same total variance. The top panel of Figure 2 shows that precipitation is predictable mostly along the equatorial ocean basins. Although predictability is statistically significant over most of the globe, the magnitude of the predictability is relatively small (e.g., $M = 0.1$ corresponds to a predictable variance of about 20%). Empirical methods tend to overestimate potential predictability, consistent with the underestimates of weather noise shown in Figure 1, but capture the spatial structure of the ensemble-dynamical estimate. Note that even though M is similar for the two methods, its statistical significance differs. As a result, most differences between M in the bottom two panels arise from different significance masks.

[17] Turning now to temperature, the weather noise variance of JFM mean 2 m temperature estimated from the 12 ensemble members is shown in the top panel of Figure 3. The figure shows the familiar enhancement of temperature variance over midlatitude land areas. The middle and bottom panels show the log-ratio of the empirical estimates relative to the 12-member ensemble estimate. Both empirical methods tend to underestimate weather noise variance over land and polar regions and overestimate weather noise

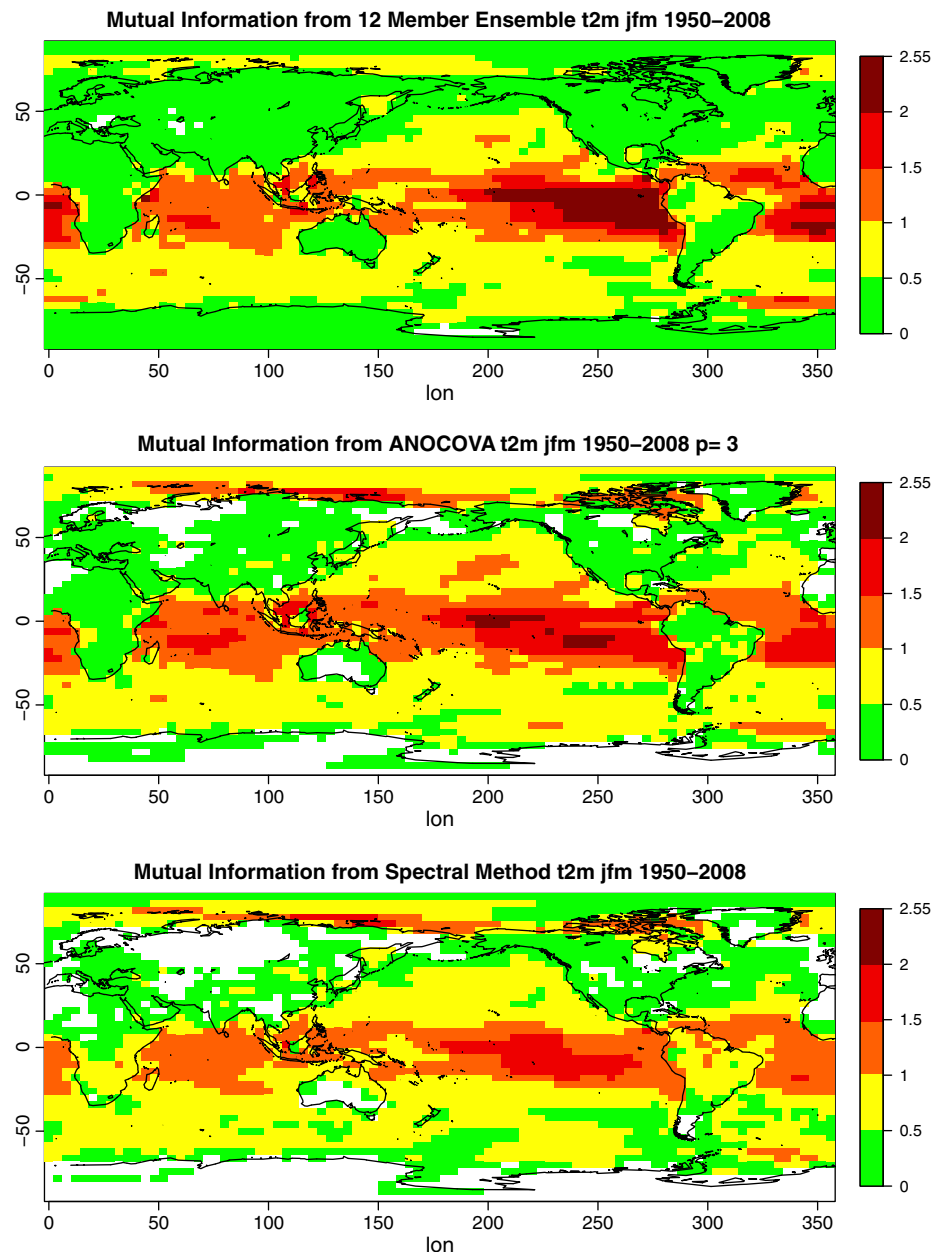


Figure 5. Mutual information of the potential predictability of JFM 2 m temperature, as estimated by (top) the 12-member ensemble, (middle) precipitation-adjusted ANOCOVA, and (bottom) the spectral method. Values insignificant at the 5% level, according to each respective method, are masked out.

variance over the equatorial west Pacific Ocean. The spectral method generally produces more accurate estimates than ANOCOVA. The cause of the discrepancies over the equatorial eastern Pacific is discussed in the supporting information and shown to be confined to that narrow region.

[18] To explore the biases over land more closely, we examined individual time series; see supporting information for representative examples. Remarkably, the variation of seasonal means across ensemble members often exceeds the variation of daily time series within the same year. This result is remarkable because 90 day averages should have *less* variance than daily time series. A clue for understanding this result is Figure 4, which shows the correlation between JFM precipitation and 2 m temperature derived from a single ensemble member. Comparison with Figure 3

suggests a tendency for the most extreme underestimates in weather noise to occur where precipitation and temperature are most strongly correlated. This correlation exists not only across different years but also across ensemble members within the same year (see supporting information). Significant correlations imply that precipitation and temperature covary. Although correlation does not imply causation, we hypothesize that, in fact, precipitation variability causes temperature variability. For instance, over most land areas, particularly in the tropics, the correlation is strongly negative. This negative correlation is plausibly explained by a land-atmosphere evaporative feedback: Over land, excess liquid precipitation leads to enhanced evaporative cooling, while a deficit of precipitation leads to reduced evaporative cooling. (Precipitation also influences the land memory, but this

effect is considered to be small [Delworth and Manabe, 1988] and, moreover, does not consistently explain our results.) The figure also shows that the northern most land areas have a positive correlation with precipitation. These areas probably are characterized by frozen precipitation, which can insulate the ground and prevent ground temperatures from reaching very cold temperatures, thereby producing relatively warmer temperatures. In both cases, precipitation variability can be interpreted as causing temperature variability. Moreover, since land precipitation is only weakly predictable in the model, temperature variability induced by precipitation is mostly unpredictable. Thus, we hypothesize that precipitation induces an additional source of unpredictable variance of temperature through land-atmosphere coupling.

[19] The above hypothesis can be verified quantitatively as follows. The variance explained by precipitation can be estimated as $\rho^2\sigma_T^2$, where ρ is the correlation between mean temperature and precipitation shown in Figure 4. Adding this variance to the weather noise variance derived from ANOCOVA yields the result shown in the bottom panel of Figure 4. Comparison between Figures 3 and 4 shows that accounting for precipitation-induced variability substantially improves the weather noise variance estimate from ANOCOVA. Nevertheless, the noise variance still is underestimated in some land regions. It is possible that precipitation noise variance might be enhanced further in some regions due to land surface feedbacks, an effect that is not taken into account in the above regression analysis.

[20] Estimates of mutual information of JFM 2 m temperature by all three methods are shown in Figure 5. As expected, empirical methods underestimate predictability in the tropical eastern Pacific and slightly overestimate predictability over land. The spectral method deems many more land points to be unpredictable compared to the other methods.

5. Summary and Conclusion

[21] This paper compared the potential seasonal predictability estimated empirically from a single realization of daily data to the predictability estimated from ensemble methods, using data generated by a state-of-the-art atmospheric model run with observed SST and external paper. The empirical methods include an analysis of covariance method proposed by Feng et al. [2012] and a spectral method proposed by Madden [1976].

[22] For JFM precipitation, empirical methods produce reasonable, though negatively biased, estimates of weather noise variance. This consistency is somewhat surprising given that precipitation is non-Gaussian, while the empirical methods assume Gaussian distributions. Consistent with previous studies, potential predictability of precipitation is found to be concentrated along the equatorial oceans and only weakly predictable over land.

[23] For JFM temperature, empirical methods tend to underestimate weather noise variance over land and overestimate weather noise variance over the equatorial

eastern Pacific. The overestimation is discussed in the supporting information and attributed to slow variations in the potentially predictable component within a season, contradicting the assumptions made in the empirical methods. The bias over land was argued to be caused by neglecting precipitation-induced variability through land-atmosphere coupling. When the variance explained by the correlation between seasonal mean temperature and precipitation is added to the ANOCOVA weather noise variance, the resulting estimate agrees much more closely with the true weather noise variance. This modified estimate assumes that seasonal mean precipitation is mostly unpredictable, which is a reasonable assumption for this dynamical model in most land areas, but may not be for other models, or for reality.

[24] **Acknowledgments.** We thank Xia Feng, J. Shukla, and Ahmed Tawfik for useful discussions and two anonymous reviewers for helpful comments. We also thank Xia Feng for finding a coding error in our ANOCOVA calculations, which significantly affected an early version of this paper. Support is gratefully acknowledged from grants from the NSF (0830068), the National Oceanic and Atmospheric Administration (NA09OAR4310058), and the National Aeronautics and Space Administration (NNX09AN50G).

[25] The Editor thanks two anonymous reviewers for their assistance in evaluating this paper.

References

- Chervin, R. M. (1986), Interannual variability and seasonal climate predictability, *J. Atmos. Sci.*, **43**, 233–251.
- DelSole, T., and X. Feng (2013), The “Shukla-Gutzler” method for estimating potential seasonal predictability, *Mon. Weather Rev.*, **141**, 822–831.
- Delworth, T. L., and S. Manabe (1988), The influence of potential evaporation on the variabilities of simulated soil wetness and climate, *J. Clim.*, **1**, 523–547.
- Feng, X., T. DelSole, and P. Houser (2011), Bootstrap estimated seasonal potential predictability of global temperature and precipitation, *Geophys. Res. Lett.*, **38**, L07702, doi:10.1029/2010GL046511.
- Feng, X., T. DelSole, and P. Houser (2012), A method for estimating potential seasonal predictability: Analysis of covariance, *J. Clim.*, **25**, 5292–5308.
- Kumar, A., M. Chen, L. Zhang, W. Wang, Y. Xue, C. Wen, L. Marx, and B. Huang (2012), An analysis of the nonstationarity in the bias of sea surface temperature forecasts for the NCEP Climate Forecast System (CFS) version 2, *Mon. Weather Rev.*, **140**, 3003–3016.
- Kumar, A., and M. P. Hoerling (1995), Prospects and limitations of seasonal atmospheric GCM predictions, *Bull. Am. Meteorol. Soc.*, **76**, 335–345.
- Madden, R. A. (1976), Estimates of the natural variability of time-averaged sea-level pressure, *Mon. Weather Rev.*, **104**, 942–952.
- Phelps, M. W., A. Kumar, and J. J. O’Brien (2004), Potential predictability in the NCEP CPC dynamical seasonal forecast system, *J. Clim.*, **17**, 3775–3785.
- Shea, D. J., and R. A. Madden (1990), Potential for long-range prediction of monthly mean surface temperature over North America, *J. Clim.*, **3**, 1444–1451.
- Shukla, J. (1981), Dynamical predictability of monthly means, *Mon. Weather Rev.*, **38**, 2547–2572.
- Shukla, J., and D. S. Gutzler (1983), Interannual variability and predictability of 500 mb geopotential height over the northern hemisphere, *Mon. Weather Rev.*, **111**, 1273–1279.
- Zheng, X., H. Nakamura, and J. A. Renwick (2000), Potential predictability of seasonal means based on monthly time series of meteorological variables, *J. Clim.*, **13**, 2591–2604.
- Zwiers, F. (1996), Interannual variability and predictability in an ensemble of AMIP climate simulations conducted with the CCC GCM2, *Clim. Dyn.*, **12**, 825–847.
- Zwiers, F. W. (1987), A potential predictability study conducted with an atmospheric general circulation model, *Mon. Weather Rev.*, **115**, 2957–2974.