Science and Technology Infusion Climate Bulletin
NOAA's National Weather Service
36th NOAA Annual Climate Diagnostics and Prediction Workshop
Fort Worth, TX, 3-6 October 2011

The Limits of Detecting Forced Responses on Seasonal and Continental Scales

Liwei Jia[1] and Timothy DelSole[1,2]

[1]*Center for Ocean-Land-Atmosphere Studies, Calverton, MD*
[2]*George Mason University, Fairfax, VA*

1. Introduction

A wide range of studies have concluded that anthropogenic greenhouse gas increases have very likely caused most of the warming on global and continental scales since the middle of the twentieth century, and that further warming over the next century is expected (Hegerl *et al.* 2007). This conclusion is not necessarily "actionable" by local governments and policy makers without more precise predictions on smaller spatial and temporal scales. The question arises as to what are the shortest space and time scales for which detection, attribution, and prediction are possible. One may also question whether the role of separate forcings, such as the role of greenhouse gases, aerosols, solar variability, can be investigated in specific climate events. Another complication is whether the indices for climate events have been selected specifically for their extreme nature, leading to selection bias. Also, pre-selecting indices (*e.g.*, based on spatial average) may lead us to overlook certain kinds of important events. This study proposes an objective framework for addressing the above questions by identifying components that maximize the signal-to-noise ratio of an externally forced event.

2. Method

Assuming forced climate variability (*i.e.*, response to external natural and anthropogenic forcing, including anthropogenic, volcanic and solar forcing) is an independent and additive perturbation to internal unforced variability, the total variance equals the sum of the variances due to forced and unforced components. Therefore, the variance of forced runs ought to be larger than the variance of unforced runs, since the forced runs contain an "extra" component of variability relative to the unforced runs. Moreover, components whose forced variance differs as much as possible from the unforced variance define components in which the forced response is most easily distinguished from unforced variability. Therefore, we seek the component that maximizes the ratio of forced variance to unforced variance. It can be shown that maximizing the variance ratio leads to the generalized eigenvalue problem (Noble and Daniel 1988; DelSole and Tippett 2009)

$$\hat{\Sigma}_F q = \lambda \hat{\Sigma}_U q, \qquad (1)$$

where $\hat{\Sigma}_F$ and $\hat{\Sigma}_U$ are sample covariance matrices of forced and unforced runs respectively. The eigenvalue λ turns out to be the variance ratio corresponding to eigenvector q. Equation (1) has more than one eigenvalue and eigenvector. Each eigenvector corresponds to a discriminant component. It is convention to order eigenvectors in decreasing order of their eigenvalues, such that the first eigenvector maximizes the variance ratio, the second eigenvector maximizes the variance ratio subject to being uncorrelated with the first, and so on. The time series associated with forced and unforced runs are

$$r_F = \mathrm{F}q \quad \text{and} \quad r_U = \mathrm{U}q\,, \qquad (2)$$

where F and U are matrices containing the (centered) time series for the forced and unforced simulations, respectively. It is shown in Jia and DelSole (2011) that the pattern given by

Correspondence to: Liwei Jia, Center for Ocean-Land-Atmosphere Studies, 4041 Powder Mill Road, Suite 302, Calverton, MD 20705; E-mail: liwei@cola.iges.org.

$$p = \hat{\Sigma}_U q \tag{3}$$

maximizes the mean statistic used to perform detection analysis in optimal fingerprinting analysis, and therefore maximizes detectability in the models. It follows that if no significant pattern can be found, the role of external forcing cannot be distinguished from internal variability.

3. Models and data

The data set used in this study is from the Coupled Model Intercomparison Project phase 3 (CMIP3) multimodel dataset. The 3-month means of surface air temperature and precipitation from the twentieth century runs (*i.e.*, forced runs) and pre-industrial control runs (*i.e.*, unforced runs) were analyzed. All fields were interpolated to a common 72 x 36 grid.

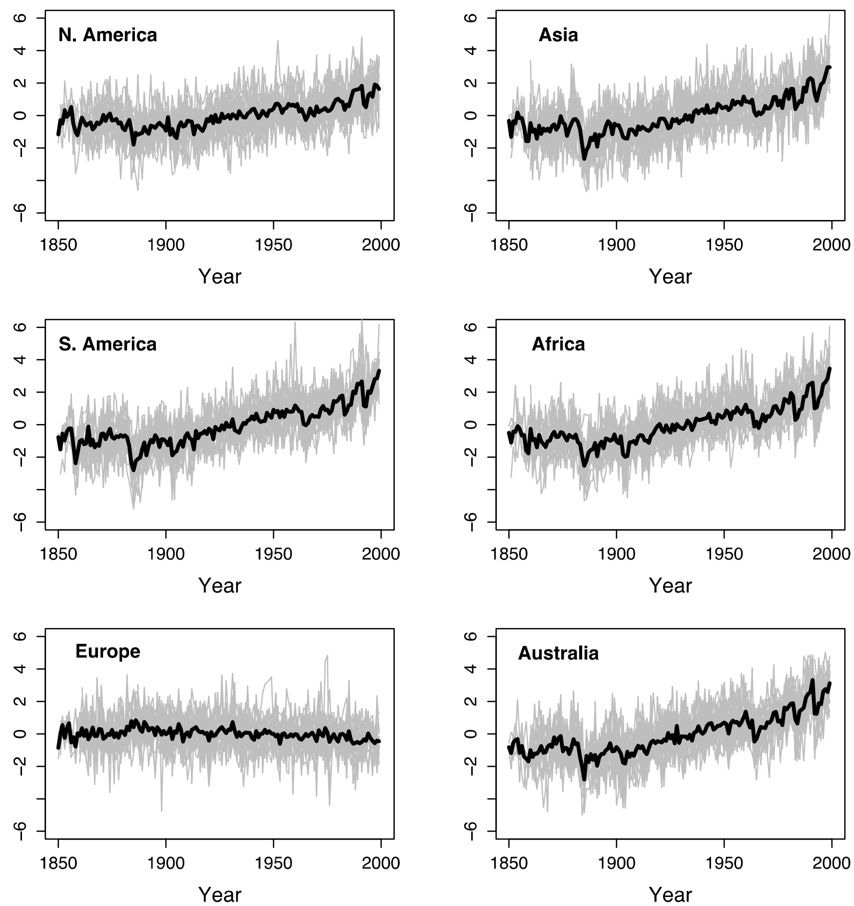

Fig. 1 Time series of the leading component of JFM mean surface air temperature in each forced ensemble member (thin grey curves) over six continents in independent verification data. The thick black curve in each panel shows the multi-model mean time series.

All statistical quantities are estimated in a multi-model sense; more precisely, the covariance matrices from different model simulations are averaged together. Only the last 300 years of unforced runs were used. The first half of the 300-year data was used as training data to maximize the variance ratio, and the second half was reserved for verification. Only one unforced run from each model was used as training. Models with significant trends in unforced runs, and significantly different variances compared to other models, were omitted. This screening procedure leads to a selection of eight models (GFDL-CM2.0, GFDL-CM2.1, IPSL-CM4, MIROC3.2 (medres), ECHO-G, MRI-CGCM2.3.2, CCSM3, UKMO-HadCM3). The selected unforced runs from each model were first centered with respect to each model's mean, and then lined up in temporal dimension to generate multi-model training and verification datasets.

For the forced runs, we used a maximum of five ensemble members in each model, and if the ensemble members are less than five in a model, we used all available members. One member of each model was used as training data to maximize the variance ratio, and the remaining members were used as verification data. Each member was centered with respect to the mean of the run. Similarly, members of eight models were lined up to form multi-model training and verification datasets.

To mitigate overfitting, we reduced the dimension of the data by projecting the data onto the leading 30 principal components. This study shows results only from independent verification data.

4. Results

4.1 Identifying forced response of continental surface air temperature in JFM

We first identify the component that maximizes the ratio of forced-to-unforced variance of JFM mean surface air temperature in training data over six continents. The resulting components were then projected onto the verification data to determine the variance ratios. According to the standard F-test, only the ratio of the leading component is well above the 5% significance level for all six continents except Europe (not shown). Thus, except for Europe, the forced response is distinguishable from unforced variability. The fact that Europe has no significant variance ratio implies that it is impossible to detect a forced response in these models over Europe based on JFM mean surface air temperature. Furthermore, since the response is not detectable in this "perfect model scenario", there is no reason to expect it to be detectable with real observations. Our conclusion pertains to seasonal mean response, whereas most previous studies, which claim that a forced response is detectable over Europe, employ longer-term means (for instance, Fig. 1 of FAQ 9.2 in Hegerl *et al.* (2007) is based on ten-year means).

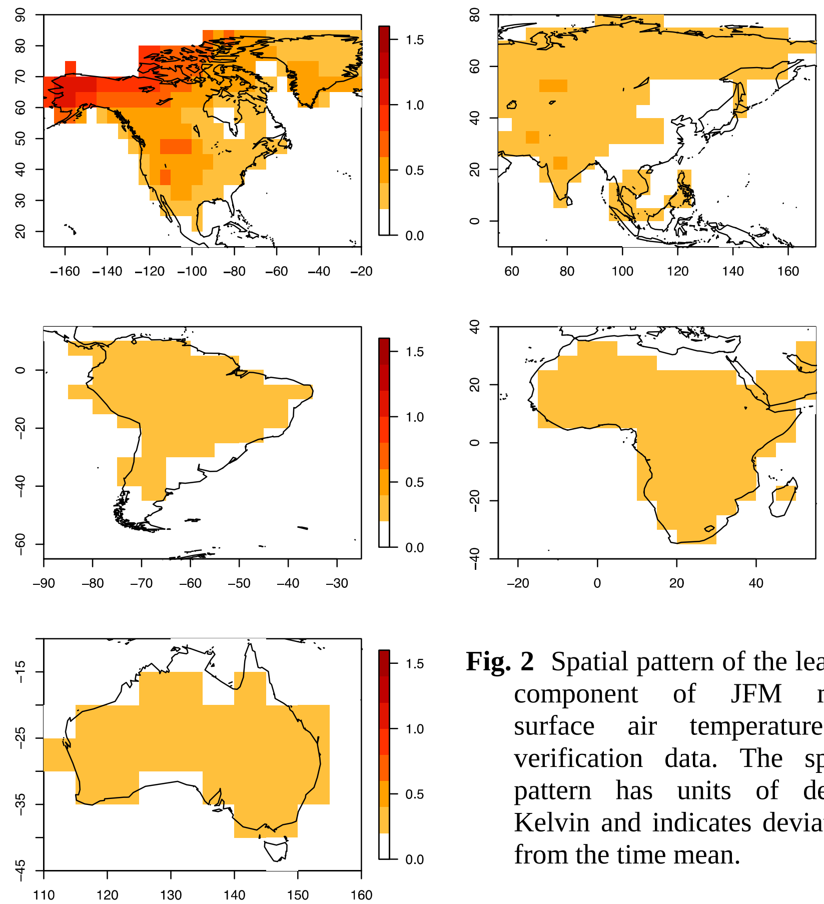

Fig. 2 Spatial pattern of the leading component of JFM mean surface air temperature in verification data. The spatial pattern has units of degree Kelvin and indicates deviations from the time mean.

The fact that only one significant variance ratio can be found in other continents implies that 1) detection of a forced seasonal response pattern in the other continents is possible, and 2) separating the response to different forcings using JFM mean surface air temperature pattern alone will prove difficult, because the similarity of the responses is so great that the responses can be compressed into a single pattern. Therefore, if there are more than one response patterns, they all project on the leading component, and will be collinear and hence difficult to separate.

We emphasize that our analysis is based only on JFM mean spatial structure, *i.e.*, no time lag information is taken into account. This allows us to apply detection and attribution on seasonal scale, but it limits our ability to attribute anomalies to specific forcings. Previous studies that attribute temperature changes to distinct forcings were based on both spatial and temporal information (Stott 2003; Zwiers and Zhang 2003).

The time series of the leading component for forced runs are shown in Fig. 1. The time series of the ensemble mean (thick black curve) shows an increasing trend in each continent except Europe. No significant trend in Europe is consistent with the fact that the corresponding variance ratio is insignificant. The spatial patterns of the leading component (Fig. 2) are of single sign. The positive sign associated with the increasing trend in each continent indicates warming on continental scales. Largest amplitudes are concentrated in high latitudes of North America.

4.2 Identifying forced response of continental surface air temperature in JAS

We repeated the above analysis except this time for JAS mean surface air temperature. Only the ratio of the leading component is well separated from the others and is statistically significant in each continent except Europe (not shown). The time series of the leading component for forced runs (Fig. 3) reveal increasing trends in all continents except Europe. The trends are generally larger in JAS than in JFM. The lack of obvious trend in Europe is consistent with the fact that Europe has no significant variance ratio. However, the forced time series in Europe does show an increasing trend in the last two decades of the twentieth century. It is possible that the trend is real, but that the short time for which the forced response is distinguishable from unforced variability leads to small (and statistically insignificant) variance ratio.

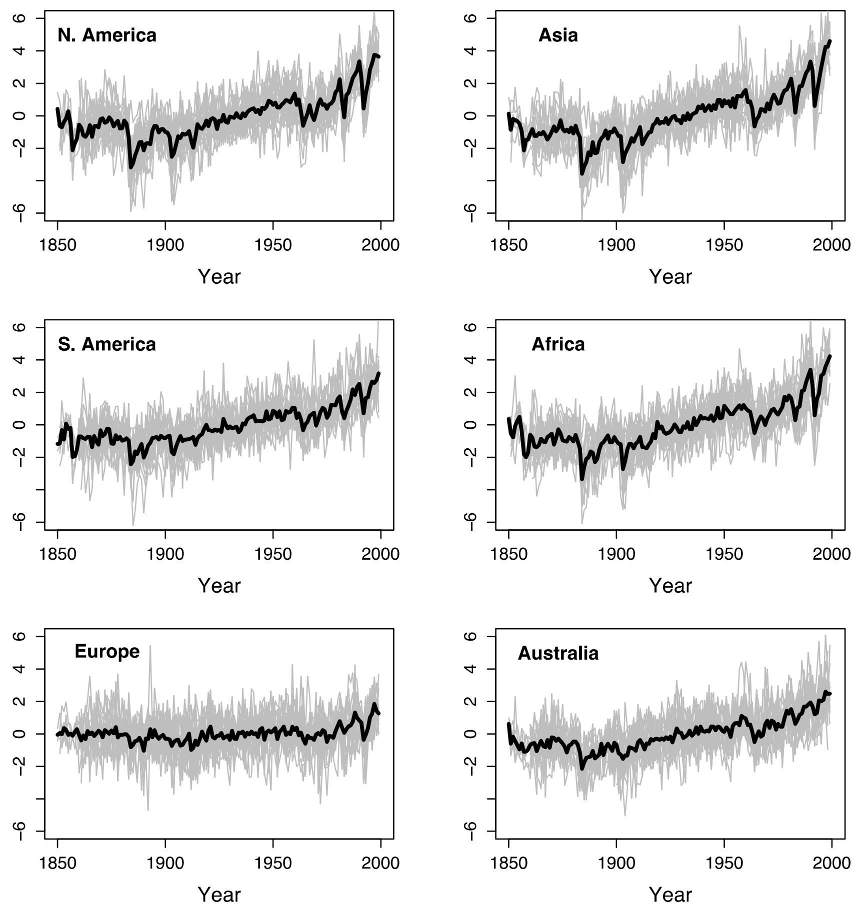

Fig. 3 Time series of the leading component in forced runs over six continents in independent verification data, as in Fig. 1, but for JAS mean surface air temperature.

The spatial patterns of the leading component, shown in Fig. 4, are of positive sign, as those in JFM. The positive sign in spatial pattern associated with the increasing trend in forced time series indicates warming on continental scales. The largest amplitudes are concentrated in high latitudes of North America in JFM, but in the interior of the continent in JAS.

We have tested that the variance ratios determined by projecting vector *q* onto verification data are larger than the ratios of continental averages for all seasons and all continents, except for Europe (not shown).

4.3 Identifying forced response of seasonal precipitation

As for precipitation, none of the variance ratios of JFM and JAS mean precipitation are significant at a 5% level in any continent, implying that the forced response of seasonal mean precipitation is not detectable. This conclusion is somewhat at odds with Zhang *et al.* (2007), who claim that anthropogenic forcing has had a detectable influence on observed changes in precipitation. Although Zhang *et al.* (2007) use land averages in zonal bands while we use continental patterns, we have repeated our analysis for global domains and zonal bands and still find only marginal-to-no significant component. A key difference between the two studies is that Zhang *et al.* (2007) test a trend pattern, which includes decadal scale information of the response, whereas here we test a seasonal mean pattern. Nevertheless, the fact that no significant forced response for precipitation can be found suggests that the precipitation trend must be weak, if it exists at all. This is in fact the case, as Zhang et al. (2007) use scaling factors around 5-10 to match modeled trends with observed trends. It is not surprising that different statistical procedures produce different conclusions for weak signals.

5. Summary

This study addresses the limit to which the response to anthropogenic and natural forcing can be distinguished from unforced variability on seasonal and continental scales. Only one statistically significant forced pattern of seasonal mean surface air temperature can be identified in each season and continent (except Europe, which has no significant forced response), implying that detection of anthropogenic and natural forcing of temperature on seasonal and continental scales is possible. The pattern in each continent is of single sign and consistent with long-term warming, but varies with season. However, the fact that only one significant pattern was obtained implies that different forcings produce similar patterns that may be difficult to separate in an attribution analysis on seasonal and continental scales. No significant forced pattern of seasonal mean precipitation could be identified, implying that detection of anthropogenic and natural forcing of precipitation is not generally possible on seasonal and continental scales. The forced response identified in this study provides the basis for detection and attribution studies on seasonal scales, for instance, in the detection and attribution of observed extreme events.

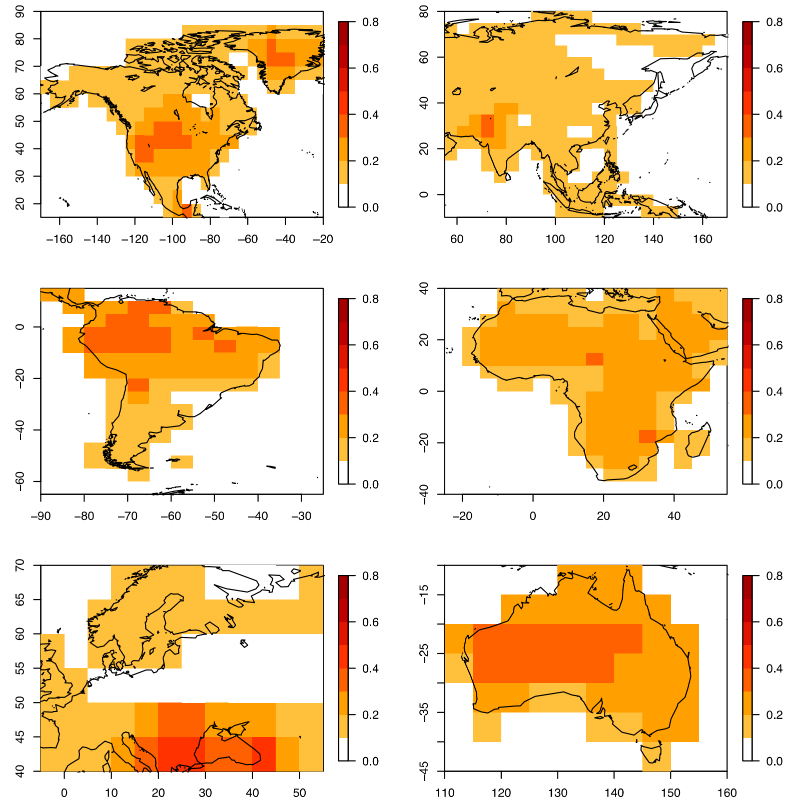

Fig. 4 Spatial pattern of the leading component of JAS mean surface air temperature in verification data. The spatial pattern has units of degree Kelvin and indicates deviations from the time mean.

Reference

DelSole, T., and M. K. Tippett, 2009: Average predictability time. Part II: Seamless diagnoses of predictability on multiple time scales. *J. Atmos. Sci.*, **66**, 1188–1204.

Hegerl, G. C., and Coauthors, 2007: Understanding and attributing climate change. *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. Averyt, M. Tignor, and H. Miller, Eds., Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Jia, L. and T. DelSole, 2011: Diagnosis of multiyear predictability on continental scales. *J. Climate*, **24**, 5108-5124.

Noble, B., and J. W. Daniel, 1988: *Applied Linear Algebra*. 3rd ed., Prentice-Hall, 521 pp.

Stott, P. A., 2003: Attribution of regional-scale temperature changes to anthropogenic and natural causes. *Geophys. Res. Lett.*, **30**, 1728, doi:10.1029/2003GL017 324.

Zhang, X., F. W. Zwiers, G. C. Hegerl, F. H. Lambert, N. P. Gillett, S. Solomon, P. A. Stott, and T. Nozawa, 2007: Detection of human influence on twentieth-century precipitation trends. *Nature*, **448**, 461–465.

Zwiers, F. W., and X. Zhang, 2003: Toward regional-scale climate change detection. *J. Climate*, **16**, 793–797.