Scalar Prediction in Climate Using Data and Model

Stephen Leroy*   John Dykema   James Anderson

Harvard School of Engineering and Applied Science, Harvard University

Abstract

We present an approach to predicting scalars of the climate system on inter-decadal timescales, and we call that approach generalized scalar prediction. Generalized scalar prediction combines the predictions of an ensemble of models that spans the community's uncertainty range in climate modeling with a timeseries of any data type, including hybrid data. The scalars can be any variable or linear combination of variables of the climate system, such as mass flux associated with the Hadley circulation or microscale precipitation. No preference need be given to any particular model. The resulting equations show that both natural variability and uncertainty in models physics must be considered when finding indicators in the given data type that optimize predicted change in arbitrary scalars. We demonstrate this approach using the CMIP3 ensemble of global warming simulations and the output of an independent model serving as a proxy for data to predict regional surface air temperature and regional precipitation trends, surface air temperature and precipitation representing end members in the physical certainty in our modeling of them. The approach gives accurate results, but the precision of the predictions depends on the certainty of model physics relevant to the scalar in question. We find a strong relationship between the mathematical techniques of signal detection and attribution and generalized scalar prediction, but some subtlety is required in the interpretation of results of the former.

1. Introduction

The fields of climate signal detection and attribution and regional trend prediction are two of the central concerns of climate research, the former answering questions regarding human influence on climate and the latter answering questions of how climate can be expected to change. A common approach to both of these problems involves scalar estimation. In climate signal detection and attribution, the scalars involved scale vectors describing the shapes with which signals are expected to emerge. That methodology is optimal detection, or linear multi-pattern regression, and it has been described at length elsewhere (Bell 1986; Hasselmann 1993; North et al. 1995; Hasselmann 1997). In regional trend prediction, the scalars are typically the rates at which surface air temperature and precipitation are predicted to increase (Houghton et al. 2001). The two fields thus far, even though both involve scalar prediction, have not been explicitly related. This paper addresses the relationship between the two.

We define generalized scalar prediction the field in which any linearly formed scalar of the climate system can be predicted based on any combination of data sets. The field of generalized scalar prediction in climate is the essential focus of the community involved in regional climate prediction. Kharin and Zwiers (2002) were among the first to point toward the necessity of using a large ensemble of sophisticated global climate models for use in regional trend prediction. A combination of the predictions of many models would average out some of the differences between them and give uncertainty estimates due to inter-model differences in prediction. Giorgi and Francisco (2000a) took a step back to evaluate the dominant sources of uncertainty in regional prediction and found that inter-model differences dominate over naturally occurring inter-annual variability and biases in model fields as sources of uncertainty. Giorgi and Francisco (2000b) in fact showed that a 3- or 4-member ensemble of realizations of transient runs of one climate model is enough to eliminate inter-annual variability as a significant source of uncertainty in regional climate prediction. Giorgi et al. (2001) point out that correlations between predictions of trends in different regions should contain important information on the

---

Corresponding author address: Stephen Leroy, Anderson Group, 12 Oxford St., Link Building, Cambridge, MA 02138.
E-mail: leroy@huarp.harvard.edu

robustness of trend prediction for any one region. At that point, several studies sought to assign weights to various models according to their performances in simulating historical and present climate. This is a natural conclusion when Bayesian inference is applied, as is done by Min and Hense (2006a,b); Min et al. (2007). In each of these papers, just as in Räisänen and Palmer (2001); Giorgi and Mearns (2002, 2003), each model in an ensemble of models is evaluated on criteria composed by the forecaster, a subjective weight is deduced, and a linear combination of the model members of the ensemble is combined to form a forecast. The Third Assessment Report of the Intergovernmental Panel on Climate Change (Houghton et al. 2001) sought to avoid controversy by equally weighting many climate models in deriving predictions of regional climate trends and associated uncertainties.

Here we also apply Bayesian inference but with different assumptions. We make no attempt to evaluate the relative performance of climate models. Instead, we take advantage of the strong relationships between variables and regions in the predictions of decadal time scale trends by different climate models. This assumption is the same as that made in linear multi-pattern regression, or optimal fingerprinting/detection (Bell 1986; Hasselmann 1993, 1997; North et al. 1995; Leroy 1998; Allen and Tett 1999), which has been applied to trend detection and attribution to specific causes for a variety of geophysical variables (Stott et al. 2001; Tett et al. 1999, 2002; Santer et al. 2003; Thorne et al. 2003; Allen et al. 2006). The central claim is that, while climate models have different sensitivities to anthropogenic radiative forcing, the patterns of change are mostly insensitive to these different climate sensitivities. We assume the same. More recently, the small differences in patterns of change have been taken into consideration in the application of optimal detection (Huntingford et al. 2006), an application of a methodology described in Bell (1986). The result of our application of Bayesian inference with the above assumptions is a method for using arbitrary data to arrive at predictions for trends in arbitrary scalar quantities of the climate system in a way that takes advantage of physical relationships of varying degrees of uncertainty within the climate system. The final equations are similar to those in Bell (1986) and Huntingford et al. (2006) but with the distinction that critical attention must be paid to the normalization of signal shapes.

In the second section we present a derivation of generalized scalar prediction that considers physical connections in the climate system by applying Bayesian inference on climate signal detection. A set of equations will result. In the third section we will present examples that illustrate various properties of generalized scalar

prediction. In the fourth we will summarize the results and discuss a few potential applications and limitations of this technique.

2. Generalized scalar prediction

This derivation is based on two levels of Bayesian inference and a simple model for trend prediction. We write down the well known Bayes's theorem,

$$P(x|D) \propto P(D|x)\,p(x), \tag{1}$$

where $p(x)$, the prior, is a probability density function on the scalar quantity x describing knowledge of x before data set D is obtained and analyzed; $P(D|x)$, the evidence function, is the conditional probability of obtaining data D should the scalar quantity be x; and $P(x|D)$, the posterior, is the probability density function for x after analysis of data D. The normalization constant for the posterior is $p(D)$, the likelihood of the data, which turns Bayes's theorem into an equality:

$$P(x|D) = P(D|x)\,p(x)/p(D). \tag{2}$$

This is a single level of Bayesian inference for the scalar x. In climate, a scalar x obtained from data D is typically obtained given the benefit of a model for the data M, which itself is uncertain. As a consequence, there are generally many models available that relate scalar quantity x to a data set D. We call that set of models $\mathcal{M}$, the members of which are M_i. The posterior probably density function for the scalar becomes

$$P(x|D,\mathcal{M}) \propto \sum_{M_i \in \mathcal{M}} P(D|x,M_i)\,p(x)\,p(M_i) \tag{3}$$

with $p(M_i)$ being the probability of model M_i relative to all the other models in $\mathcal{M}$. In the context of scalar prediction, this means that future trends can be estimated by timeseries analysis of historical data based on multiple models and then averaging together the results based on subjective weights.

In this section, the equations of generalized scalar prediction are derived from the foundational Eq. 3 and the method of climate forecast is presented based on standard error analysis.

a. Equations of scalar prediction

This is similar to the approach taken by Räisänen and Palmer (2001) but with some assumptions of optimal detection included. In truth, the relative probability of model M_i must be conditional on the existence of data D and consequently the weight $p(M_i)$ should be $p(M_i|D)$ and not entirely subjective. These are the marginal probabilities of the model given data D, and

how this can be deduced based on timeseries data was first presented in Leroy (1998) and thereafter by Min et al. (2007). For now we will only consider the subjective probability $p(M_i)$ in the derivation and discuss later how one can refine conclusions by instead using the marginal probabilities of the models as weights.

The standard assumption in optimal detection is that, while various models models show different sensitivities in greenhouse gas forcing, they nonetheless show very similar patterns of change. It was noted in Giorgi and Francisco (2000a) that patterns of surface air temperature change, when normalized by sensitivity of a model, are independent of greenhouse gas forcing scenario. Thus, it is reasonable to model a timeseries of data $\mathbf{d}(t)$ as

$$\mathbf{d}(t) = \alpha(t)\,\mathbf{s}_i + d\mathbf{n}(t) \qquad (4)$$

where the pattern $\mathbf{s}_i$ is the derivative of the model for the data in a scalar α. The $d\mathbf{n}(t)$ are realizations of natural variability of the climate system viewed through the filter of the data type represented by $\mathbf{d}$. Here, natural variability refers to all temporal fluctuations of the climate system not specifically related to trends $d\alpha/dt$ that result from external forcings of the system. In the most typical application of signal detection, $\mathbf{d}$ is a longitude-latitude map of surface air temperature, α is the global average surface air temperature, and the signal pattern $\mathbf{s}_i$ is given by

$$\mathbf{s}_i = \frac{d\mathbf{g}}{d\alpha}\Big|_{M_i} \qquad (5)$$

where $\mathbf{g}$ is a model of the data $\mathbf{d}$ that varies with the scalar quantity α. In one of the examples of section 3, the quantity on the right gives the trend in the longitude-latitude map of surface air temperature normalized by the trend in regional surface air temperature. The most likely fit to the timeseries of data $\mathbf{d}(t)$ will give a time derivative of $\alpha(t)$ that represents the secular trend in regional surface air temperature. The realizations of natural variability $d\mathbf{n}(t)$ are normally distributed as $d\mathbf{n} \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma_n})$.

We expand Eq. 3 using the trend in the data $\mathbf{d}(t)$ as the data and find that

$$P\Big(\frac{d\alpha}{dt} \,\Big|\, \frac{d\mathbf{d}}{dt}, \mathcal{M}\Big) \propto \sum_{M_i \in \mathcal{M}} (2\pi)^{-k/2} \left|\mathbf{\Sigma}_{d\mathbf{n}/dt}\right|^{-1/2} \exp\left[-\frac{1}{2}\left(\Big(\frac{d\mathbf{d}}{dt} - \mathbf{s}_i\frac{d\alpha}{dt}\Big)^T \mathbf{\Sigma}_{d\mathbf{n}/dt}^{-1}\Big(\frac{d\mathbf{d}}{dt} - \mathbf{s}_i\frac{d\alpha}{dt}\Big)\right)\right] p(M_i), \quad (6)$$

where k is the rank of $\mathbf{\Sigma}_{d\mathbf{n}/dt}$. We have assumed an uninformative prior for $d\alpha/dt$ ($p(x)$), an implicit assumption of optimal detection. The natural variability covariance $\mathbf{\Sigma}_{d\mathbf{n}/dt}$ derives from the normal distribution expected for a residual trend induced by the natural variability, $d\mathbf{n}/dt \sim \mathcal{N}(\mathbf{0}, \mathbf{\Sigma}_{d\mathbf{n}/dt})$, and is related to the natural variability covariance $\mathbf{\Sigma_n}$ by

$$\mathbf{\Sigma}_{d\mathbf{n}/dt} = \frac{12}{(N^3 - N)\,(\mathrm{yrs}^2)}\,\mathbf{\Sigma_n}, \qquad (7)$$

with N the number of years in a continuous timeseries of annual average data, for serially uncorrelated natural variability. At this point we assume a continuum of models $M_i \in \mathcal{M}$ such that the distribution of the signal $\mathbf{s}$ is normal: $\mathbf{s} \sim \mathcal{N}(\bar{\mathbf{s}}, \mathbf{\Sigma_s})$. Under this assumption, Eq. 6 becomes

$$P\Big(\frac{d\alpha}{dt} \,\Big|\, \frac{d\mathbf{d}}{dt}, \mathcal{M}\Big) \propto (2\pi)^{-k} \left|\mathbf{\Sigma}_{d\mathbf{n}/dt}\right|^{-1/2} \int d^k\mathbf{s}\,\left|\mathbf{\Sigma_s}\right|^{-1/2} \qquad (8)$$
$$\times \exp\left[-\frac{1}{2}\left(\Big(\frac{d\mathbf{d}}{dt} - \mathbf{s}\frac{d\alpha}{dt}\Big)^T \mathbf{\Sigma}_{d\mathbf{n}/dt}^{-1}\Big(\frac{d\mathbf{d}}{dt} - \mathbf{s}\frac{d\alpha}{dt}\Big)\right)\right] \times \exp\left[-\frac{1}{2}(\mathbf{s} - \bar{\mathbf{s}})^T \mathbf{\Sigma_s}^{-1}(\mathbf{s} - \bar{\mathbf{s}})\right].$$

Performing the integral over $d^k\mathbf{s}$ gives

$$P\Big(\frac{d\alpha}{dt} \,\Big|\, \frac{d\mathbf{d}}{dt}, \mathcal{M}\Big) \propto (2\pi)^{-k/2} \left|\mathbf{\Sigma}\right|^{-1/2} \exp\left[-\frac{1}{2}\left(\Big(\frac{d\mathbf{d}}{dt} - \bar{\mathbf{s}}\frac{d\alpha}{dt}\Big)^T \mathbf{\Sigma}^{-1}\Big(\frac{d\mathbf{d}}{dt} - \bar{\mathbf{s}}\frac{d\alpha}{dt}\Big)\right)\right] \qquad (9)$$

where

$$\mathbf{\Sigma} = \mathbf{\Sigma}_{d\mathbf{n}/dt} + (d\alpha/dt)^2\,\mathbf{\Sigma_s}. \qquad (10)$$

The most likely estimate for the scalar trend $d\alpha/dt$ is $(d\alpha/dt)_{\mathrm{ml}}$, where

$$(d\alpha/dt)_{\mathrm{ml}} = \mathbf{f}^T(d\mathbf{d}/dt) \qquad (11)$$
$$\mathbf{f} = \mathbf{\Sigma}^{-1}\bar{\mathbf{s}}\,(\bar{\mathbf{s}}^T\mathbf{\Sigma}^{-1}\bar{\mathbf{s}})^{-1}. \qquad (12)$$

The vector $\mathbf{f}$ is the *contravariant fingerprint* to the fingerprint $\bar{\mathbf{s}}$ and is best interpreted as a vector which gives an optimal and unbiased most likely estimate for a trend in the climate system. Eqs. 11 and 12 are the equations for generalized scalar trend prediction for a single scalar.

In this derivation, no condition is placed on the data $\mathbf{d}$ or on the scalar α; hence, they are completely general, and any data set can be used to determine trends associated with any particular scalar quantity in climate. Data sets can be anything from in situ surface and upper air measurements to satellite hyperspectral measurements; the scalar can be anything from the mass flux associated with the Hadley circulation to microscale precipitation. The most standard application in the literature to date has been to define the data set $\mathbf{d}$ to be gridded maps of surface air temperature and the scalar α the global average surface air temperature (Huntingford et al. 2006). Generalized scalar prediction is well suited to regional climate prediction, though, because the scalar can be defined as a regional temperature and the data set chosen so as to minimize the influence of natural variability on detection. It is possible that, depending on the scalar chosen, there exists a strong correlation in the climate system between trends in the chosen scalar and other seemingly unrelated geophysical variables. Those correlations are found by using an ensemble of models $\mathcal{M}$ that span our subjective uncertainty in climate forecasting to define $\boldsymbol{\Sigma}_{\mathbf{s}}$:

$$\bar{\mathbf{s}} = \sum_{M_i \in \mathcal{M}} \mathbf{s}_i \, p(M_i) \tag{13}$$

$$\boldsymbol{\Sigma}_{\mathbf{s}} = \sum_{M_i \in \mathcal{M}} (\mathbf{s}_i - \bar{\mathbf{s}}) \, (\mathbf{s}_i - \bar{\mathbf{s}})^T \, p(M_i). \tag{14}$$

When the models in $\mathcal{M}$ show agreement in some dimensions of the signal space, the corresponding elements of $\boldsymbol{\Sigma}_{\mathbf{s}}$ will be small and detection will be directed preferentially toward those dimensions. Our interpretation of this effect is that uncertain climate system physics should be factored into forecasts of regional trends. When the physics, in the form of relationships between variables in the climate system, is uncertain, then a posterior estimate of trends takes only interannual variability into consideration. When it is more certain, however, a posterior estimate will begin to take other elements of the climate system into consideration.

The application of Eqs. 10 and 11 depends upon the length Δt of the timeseries of $\mathbf{d}(t)$ in question. Natural interannual variability in the space of the data $\boldsymbol{\Sigma}_{\mathbf{n}}$ is related to the influence of that variability on trend estimates $\boldsymbol{\Sigma}_{d\mathbf{n}/dt}$ by $\boldsymbol{\Sigma}_{d\mathbf{n}/dt} \approx (\Delta t)^{-3} \, \boldsymbol{\Sigma}_{\mathbf{n}}$ for a continuous timeseries, $\boldsymbol{\Sigma}_{d\mathbf{n}/dt} \approx (\Delta t)^{-2} \, \boldsymbol{\Sigma}_{\mathbf{n}}$ for a timeseries with major gaps. In either case, the longer the timeseries is

the less import natural variability becomes, because the signal physics uncertainty term—the second term on the right of Eq. 10—is independent of the length of the timeseries.

This form of detection ameliorates the difficulties of eigenmode truncation typically associated with signal detection. When a precise signal shape is prescribed in optimal detection, posterior estimates of trends become highly sensitive to the number of eigenvectors retained in the inversion of $\boldsymbol{\Sigma}_{\mathbf{n}}$. Allen and Tett (1999) solve this problem by demanding that post-fit residuals be consistent with estimates of natural variability and that eigenmodes be truncated accordingly, but sensitivity to eigenmode truncation remains. The eigenmode truncation problem derives from the fact that eigenvectors of $\boldsymbol{\Sigma}_{\mathbf{n}}$ associated with small eigenvalues occupy subspaces of the signal shape that are significantly more uncertain than the subspaces defined by eigenvectors associated with larger eigenvalues. As noted by Huntingford et al. (2006) and Leroy et al. (2008), introduction of $\boldsymbol{\Sigma}_{\mathbf{s}}$ "washes out" the subspaces associated with small eigenvalues of $\boldsymbol{\Sigma}_{\mathbf{n}}$, and the sensitivity to eigenmode truncation is almost entirely lost. The term $\boldsymbol{\Sigma}_{\mathbf{s}}$ in Eq. 10 serves that numerical purpose here.

Eqs. 10 and 11 are iterative in $d\alpha/dt$. In the first iteration for $d\alpha/dt$, one should introduce a good estimate for it into Eq. 10 and solve for a new value in Eq. 11. The quantity $d\alpha/dt$ in Eq. 10 plays the role of a weighting between the influence of natural variability and the signal physics uncertainty which, in most applications, strongly weights toward natural variability. After the first determination of $d\alpha/dt$ through Eq. 11, that determination can be inserted again into Eq. 10 to find a new determination of $d\alpha/dt$, and so on. Because of the strong weighting toward natural variability, though, convergence is expected to be rapid. For this reason, in the remainder of this paper we only apply Eqs. 10 and 11 one time with no iteration.

Some applications of generalized scalar prediction require the detection of multiple scalars simultaneously. In our example for forecasting trends in regional surface air temperature, the regional surface air temperature is expected to rise with increasing greenhouse gases but fall with increasing concentrations of tropospheric sulfate aerosols. Since future trends in these radiative forcings are expected to evolve differently, it makes sense to consider them separately when forecasting future trends, yet the historical record contains both radiative forcings simultaneously. In such instances, the model for the data (c.f. Eq. 4) becomes

$$\mathbf{d}(t) = \alpha_1(t) \, \mathbf{s}_i^{(1)} + \alpha_2(t) \, \mathbf{s}_i^{(2)} + d\mathbf{n}(t) \tag{15}$$

where $\mathbf{s}_i^{(1)}$ is the trend $d\mathbf{g}/d\alpha$ in the data expected of

increasing greenhouse gases with respect to the trend in global surface air temperature and $s_i^{(2)}$ is the trend $d\mathbf{f}/d\alpha$ in the data expected of increasing sulfate aerosols with respect to the trend in global surface air temperature. A derivation following that given above for just one pattern gives the following equations for scalar prediction:

$$\mathbf{\Sigma} = \mathbf{\Sigma}_{d\mathbf{n}/dt} + \sum_{M_i \in \mathcal{M}} p(M_i) \qquad (16)$$

$$\times \sum_{j,k=1}^{2} \left(\frac{d\alpha_j}{dt}\right)\left(\frac{d\alpha_k}{dt}\right) \delta\mathbf{s}_i^{(j)} \, \delta\mathbf{s}_i^{(k)^T}$$

$$(d\boldsymbol{\alpha}/dt)_{\mathrm{ml}} = \mathbf{F}^T (d\mathbf{d}/dt) \qquad (17)$$
$$\mathbf{F} = \mathbf{\Sigma}^{-1}\bar{\mathbf{S}} \, (\bar{\mathbf{S}}^T\mathbf{\Sigma}^{-1}\bar{\mathbf{S}})^{-1} \qquad (18)$$

in which the array $\boldsymbol{\alpha}$ is composed of the individual scalars α_j, the columns of $\bar{\mathbf{S}}$ are the individual $\bar{\mathbf{s}}^{(j)}$, and $\delta\mathbf{s}_i^{(j)}$ is the difference between the realization of signal j by model M_i from the mean signal j over all models in $\mathcal{M}$, or $\delta\mathbf{s}_i^{(j)} = \mathbf{s}_i^{(j)} - \bar{\mathbf{s}}^{(j)}$. The most likely estimate for the generalized scalar trend $d\boldsymbol{\alpha}/dt$ is $(d\boldsymbol{\alpha}/dt)_{\mathrm{ml}}$. The columns of the matrix $\mathbf{F}$ are the contravariant fingerprints to the fingerprints $\bar{\mathbf{S}}$ because, when multiplied by the fingerprints, yield the identity matrix (c.f. Eq. 12). Eqs. 16, 17, and 18 can be applied for, not just two, but any number of signals. These equations should be applied following the same methodology as Eqs. 10, 11, and 12.

b. Climate prediction

With a determination of most likely trends $d\boldsymbol{\alpha}/dt$ and an assumption of linear trends in climate, the probability density for the prediction of that scalar on interdecadal time scales becomes

$$\boldsymbol{\alpha}(t) \sim \mathcal{N}\big((d\boldsymbol{\alpha}/dt)_{\mathrm{ml}} \, (t-t_0), \, \mathbf{\Sigma}_{d\boldsymbol{\alpha}/dt}(t-t_0)^2 + \mathbf{\Sigma_n}\big). \qquad (19)$$

The probability density as a function of the scalar value shifts linearly in time at the rate $(d\boldsymbol{\alpha}/dt)_{\mathrm{ml}}$ beginning from a nominal time t_0. Its width is decided by two terms, the first being the uncertainty on the underlying "climate trend" $(d\boldsymbol{\alpha}/dt)_{\mathrm{ml}}$ and the second being an accounting for the perpetual natural fluctuations of the climate system $\delta\mathbf{n}$. Both are accounted as full covariance matrices, $\mathbf{\Sigma}_{d\boldsymbol{\alpha}/dt}$ and $\mathbf{\Sigma_n}$. The latter can be determined by a number of means involving combinations of detrended data and control runs of climate models. The former term requires more care.

Standard optimal fingerprinting techniques provide a simple estimate for the uncertainty covariance in the most likely trend: $\mathbf{\Sigma}_{d\boldsymbol{\alpha}/dt} = (\bar{\mathbf{S}}^T\mathbf{\Sigma}^{-1}\bar{\mathbf{S}})^{-1}$, obtainable

from error propagation based on the model in Eq. 4 in which error enters through the formulation of $\bar{\mathbf{s}}$ and the natural variability $d\mathbf{n}$. The natural variability is generally prescribed by a long control run of a climate model and is rarely realistic, so a more relevant error analysis should be based on data alone. In the approach, we multiply the contravariant fingerprints $\mathbf{F}$ to each temporal realization of the climate data vector $\mathbf{d}(t)$ to arrive at a timeseries of "detectors" $\boldsymbol{\alpha}(t)$:

$$\boldsymbol{\alpha}(t) = \mathbf{F}^T \, \mathbf{d}(t). \qquad (20)$$

The contravariant fingerprints $\mathbf{F}$ are derived based on a prescription of natural variability, but it is possible nevertheless to deduce an uncertainty which is semi-independent of the model-prescribed natural variability. We do this by linear regression on the timeseries of detectors $\alpha(t)$ (now working with just one scalar):

$$m = \Big(\sum_{i=1}^{N}(t_i - \bar{t})^2\Big)^{-1} \sum_{i=1}^{N} \alpha(t_i) \, (t_i - \bar{t}) \qquad (21)$$
$$b = \bar{\alpha} - m\bar{t} \qquad (22)$$

with $\bar{t}$ the mean of the t_i, $\bar{\alpha}$ the mean of the detectors, N the number of time intervals (generally years), and m and b the slope and intercept of the fit. If σ_α^2 is the mean square post-fit residual of the detectors, then the squared uncertainties in the slope $\langle \delta m^2 \rangle$, the intercept $\langle \delta b^2 \rangle$, and the covariance of their uncertainties $\langle \delta m \, \delta b \rangle$ are

$$\langle \delta m^2 \rangle = \Big(\sum_{i=1}^{N}(t_i - \bar{t})^2\Big)^{-1} \sigma_\alpha^2 \qquad (23)$$
$$\langle \delta b^2 \rangle = \langle \delta m^2 \rangle \, \bar{t}^2 + \sigma_\alpha^2/N \qquad (24)$$
$$\langle \delta m \, \delta b \rangle = -\langle \delta m^2 \rangle \, \bar{t}. \qquad (25)$$

The fitted slope m is exactly the same as the most likely estimate of the trend $(d\alpha/dt)_{\mathrm{ml}}$, but the uncertainty in the slope is not the same as the uncertainty in the scalar trend. The uncertainty in the slope, given in Eq. 23, only accounts for the uncertainty introduced by natural variability and not the uncertainty due to uncertainty in the shape of the emerging signal. Because the most likely estimate of the scalar trend will be in error in part because of error in the estimation of the signal fingerprint $\bar{\mathbf{s}}$, this latter error must be included. This is done by adding it to the uncertainty in the slope, $\langle \delta m^2 \rangle$:

$$\sigma_{d\alpha/dt}^2 = \langle \delta m^2 \rangle + \mathbf{f}^T\mathbf{\Sigma_s}\mathbf{f} \, (d\alpha/dt)^2 \qquad (26)$$

with the $d\alpha/dt$ being a prior estimate of the scalar trend. It is this expression for uncertainty in trend which enters into the climate forecast. In the case of a forecast

of a single scalar, the forecast has mean $\alpha_{\text{forecast}}(t)$ and standard deviation $\sigma^2_{\text{forecast}}$ given by

$$\alpha_{\text{forecast}}(t) = mt + b \qquad (27)$$
$$\sigma^2_{\text{forecast}}(t) = \sigma^2_{d\alpha/dt}(t - \bar{t})^2 + \sigma^2_{\alpha}/N. \qquad (28)$$

Thus, the climate prediction for the scalar trend has the probability density function $\alpha(t) \sim \mathcal{N}\big(\alpha_{\text{forecast}}(t), \sigma^2_{\text{forecast}}(t) + \sigma^2_{\mathbf{n}}\big)$, as in Eq. 19.

3. Examples

In this section we demonstrate the methodology's predictive capability within a modeling framework. We utilize an ensemble of different models run with the same boundary conditions for the same simulated time period to prescribe the trend patterns $\mathbf{s}_i$ (signals). One model from this ensemble is withheld and is used to represent an independent physical realization and associated set of simulated observations. We study two familiar state variables, surface air temperature and precipitation. We utilize the World Climate Research Programme's (WCRP's) Coupled Model Intercomparison Project phase 3 (CMIP3) multi-model dataset for the SRES A1B emissions scenario, which provides a total of 55 realizations for 24 different models for temperature and 53 realizations for 24 models for precipitation. We use generalized scalar prediction to make forecasts of area-weighted regional quantities, choosing the Central United States, Northern Europe, and the Sahel as representative (Table 1). Temperature and precipitation provide demonstrative examples of the extremes of application of this type of multi-model ensemble prediction because the models predict positive, significant trends in temperature but the precipitation trends are distributed nearly symmetrically about zero (Fig. 1). We proceed first with the discussion of temperature followed by that for precipitation.

a. Temperature

For each of the three regions, the following analysis procedure is followed. First we compute a signal pattern $\mathbf{s}_i$ for each model realization by linear regression analysis of the first 50 years temperature data for the 21st century. We calculate an estimate of the scalar trend $d\alpha/dt$ (regional temperature in this case) by finding the area-weighted average of the atmospheric surface temperature trend at the model gridpoints lying within the target region. We normalize the global signal patterns for each model (Eq. 5) by the corresponding estimate of the scalar trend. For models with multiple realizations, we average the normalized signals obtained from all realizations together to compute a single signal pattern per model, resulting in a set of 24 normalized global

signals. Each of these normalized signals is then regridded onto a uniform 32×64 (latitude by longitude) reduced-resolution grid (Fig. 2). The mean signal $\bar{\mathbf{s}}$ and covariance matrix $\boldsymbol{\Sigma}_{\mathbf{s}}$ (Fig. 3) are computed from these normalized, regridded signal patterns.

A prescription of natural variability is obtained from the present-day control run of the NCAR CCSM 3.0 model. The annual mean from each of the 600 model years of monthly data is found, and the covariance for this annual average timeseries is computed. This covariance matrix is normalized by $12/(N^3 - N)$ as in Eq. 7, where N is the number of years of "observations" available to generate the prediction, to obtain the covariance of the realization of natural variability, $\boldsymbol{\Sigma}_{d\mathbf{n}/dt}$ (Fig. 3). Like signal uncertainty, natural variability is also confined largely to high latitudes, though it extends further into continental regions. The optimization provided by the method we present here narrows the uncertainty of the prediction by minimizing the influence of the areas of high signal uncertainty and natural variability that do not contribute proportionally to the scalar trend of interest.

The quantities $\bar{\mathbf{s}}$, $\boldsymbol{\Sigma}_{\mathbf{s}}$ and $\boldsymbol{\Sigma}_{d\mathbf{n}/dt}$ (Eqs. 10,12) allow the computation of $\mathbf{f}$, the contravariant fingerprint for regional prediction. We averaged together the 31 estimates of $\mathbf{f}$ resulting from inversions of $\boldsymbol{\Sigma}$ (Eq. 10) with eigenvalue 20 to 50 kept to obtain the value used for prediction (Fig. 4). In this exercise, we designated one realization for one model as simulated observations $\mathbf{d}$ and withheld all realizations for this model from the computation of this contravariant fingerprint. The application of $\mathbf{f}$ to the timeseries $\mathbf{d}_i$ of annual average temperature observations produces a timeseries of detectors α. The prediction of $d\alpha/dt$ is made through linear regression analysis of these detectors. For these regional demonstrations, we choose GISS Model E-R as simulated observations because it exhibits regional temperature changes far from the ensemble mean (c.f. Fig. 1). From the designated GISS model data, we derive both a true trend value, from all 50 years, and a simulated measurement dataset, limited to 10 or 20 years, from which we make a forecast. Henceforth these simulated measurement datasets of limited length will be referred to as "observations."

1) CENTRAL UNITED STATES

In the analysis of the regional forecast for the Central United States, we retain only Northern hemispheric data. The contravariant fingerprint ($\mathbf{f}$) for the Central United States shows the importance of both local and oceanic effects for determining regional temperature trends (Fig. 4). This contravariant fingerprint shows a strong peak in amplitude over the central United States,
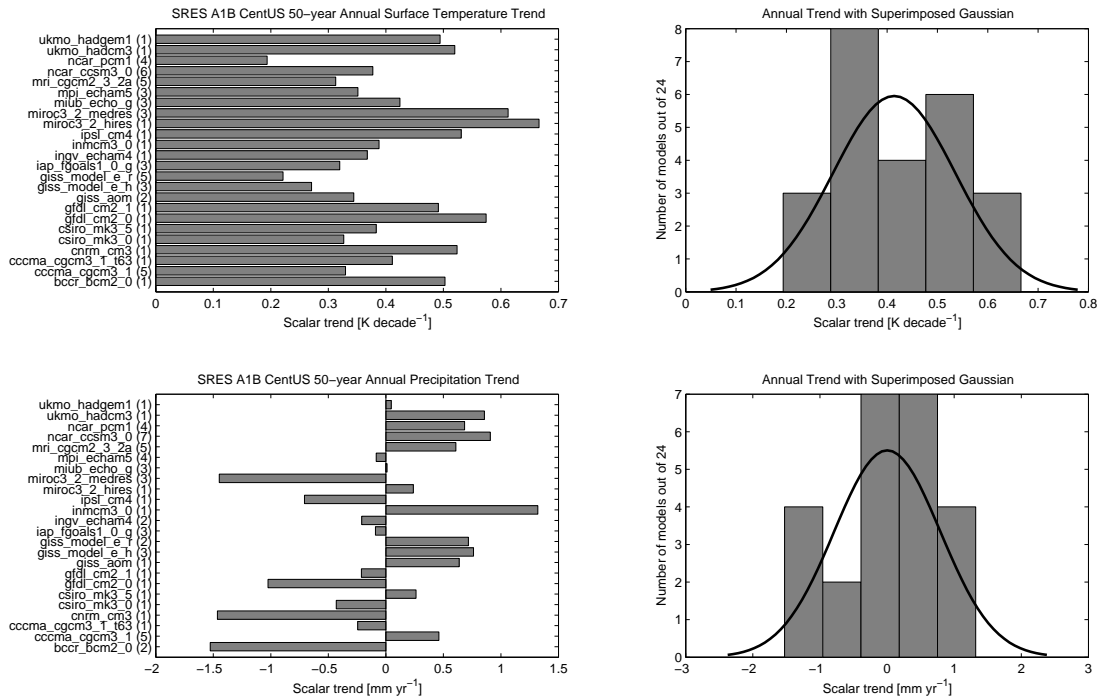
6

FIG. 1. Distributions of Central United States temperature (upper) and precipitation (lower) trends from all available model realizations of SRES-A1B. The left hand plots are bar graphs showing the regional trend predicted for each model (computed from a simple average for models submitting multiple realizations). The names of the models are given on the vertical axis, with the corresponding number of realizations in parenthesis. The right hand plots are frequency distributions for the same data as the left hand plots. The heavy black line shows the best fit of a normal distribution to the data. Temperature and precipitation provide a useful set of extremes for regional prediction because models show relatively good agreement for temperature change but much variation for precipitation.

along with a signicant correlation with sea surface temperatures in the northern central Pacific, offset by an anticorrelation with north Atlantic temperatures. This fingerprint also indicates the dependence of the regional surface temperature trend on high-latitude temperatures strongly affected by sea ice variability off the east coast of Greenland and in the Bering Sea.

We computed the results for Δt equal to 10 and 20 years—as if we had 10 or 20 years of data—to study the forecast accuracy gained as longer timeseries of data are utilized (Fig. 5). For 10 years, the estimate for the temperature increase resulting from trend analysis of the detectors (obtained from application of the contravariant fingerprint, Eq. 20) is 0.43 K with a one-sigma uncertainty of 0.40 K for the 30th year of the forecast. For 20 years, the estimate of the increase is 1.19 K and has a substantially reduced uncertainty of 0.30 K for the end

of a 30-yr forecast. The result for 20 years is very close to the true value, obtained by trend analysis of all 50 years of the model output of 1.31 ± 0.03 K. The 10 year timeseries of "observations" is too short to sufficiently resolve the relationship between the global temperature and regional temperature trends to make an optimized forecast, but remarkably generalized scalar prediction recovers a positive trend from these 10 years of observations when, in the same time period, the actual regional average surface air temperature trend over the Central United States exhibits a net negative trend of -0.13 K yr^{-1}.

2) NORTHERN EUROPE

The case of the contravariant fingerprint for Northern European temperature stands in contrast with that of the central United States in that it shows very little de-

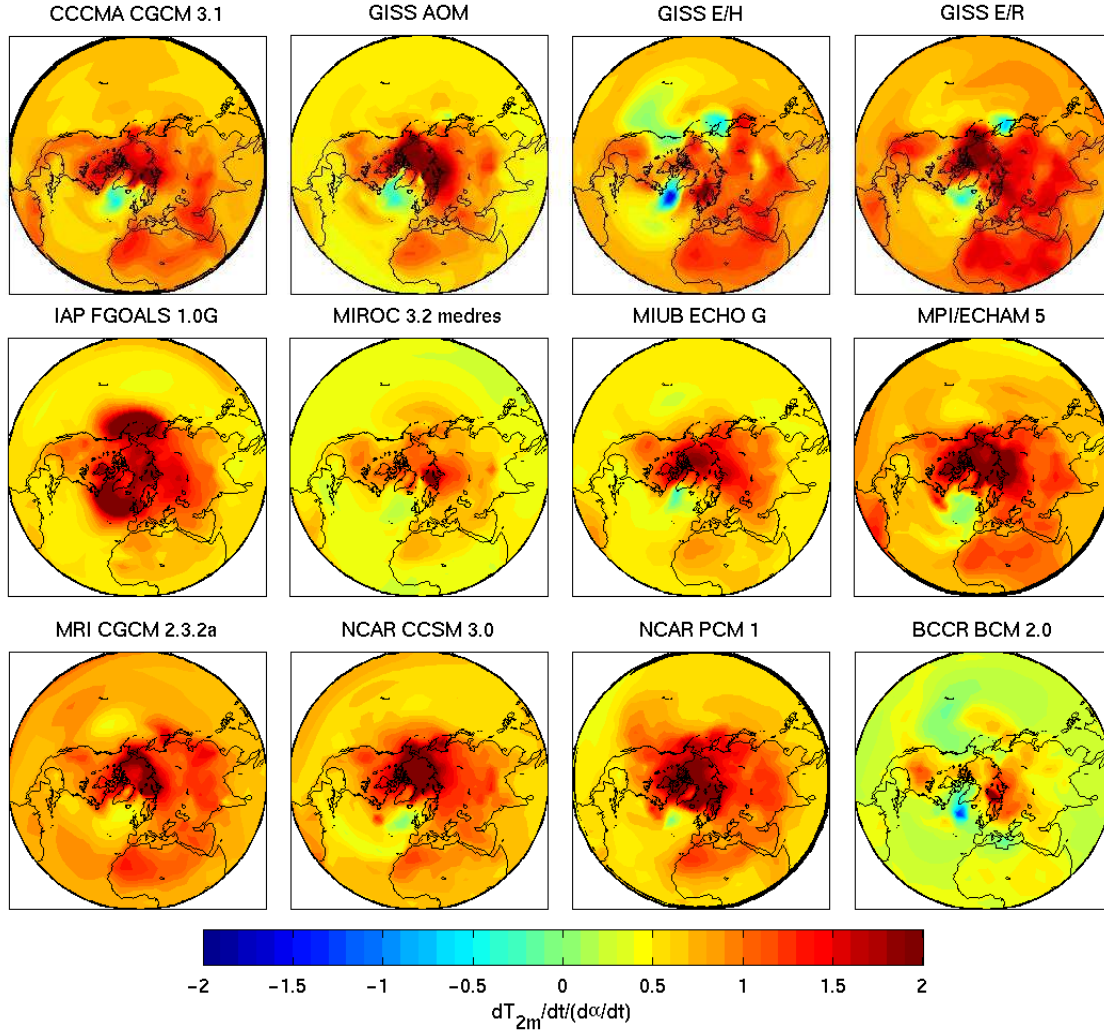| Target Region	Latitude Range	Longitude Range
Central United States	35°N−45°N	110°W−85°W
Northen Europe	48°N−55°N	0°E−18°E
Sahel	23°N−25°N	13°E−33°E

TABLE 1.

FIG. 2. Trends of atmospheric surface air temperature normalized by trend in Central United States temperature for twelve models. Eleven of the twelve models contributed multiple realizations for the emissions scenario and thus provide a more robust estimation of the signal pattern.

pendence on the local temperature, indicating the long-term trend is determined mostly by large-scale effects. Like the Central United States, the surface temperature trend is also dependent on ocean regions subject to a great deal of sea ice variability, including the eastern shore of Greenland, the Bering Sea, and the western portion of Hudson Bay and its environs. For Northern Eu-

8
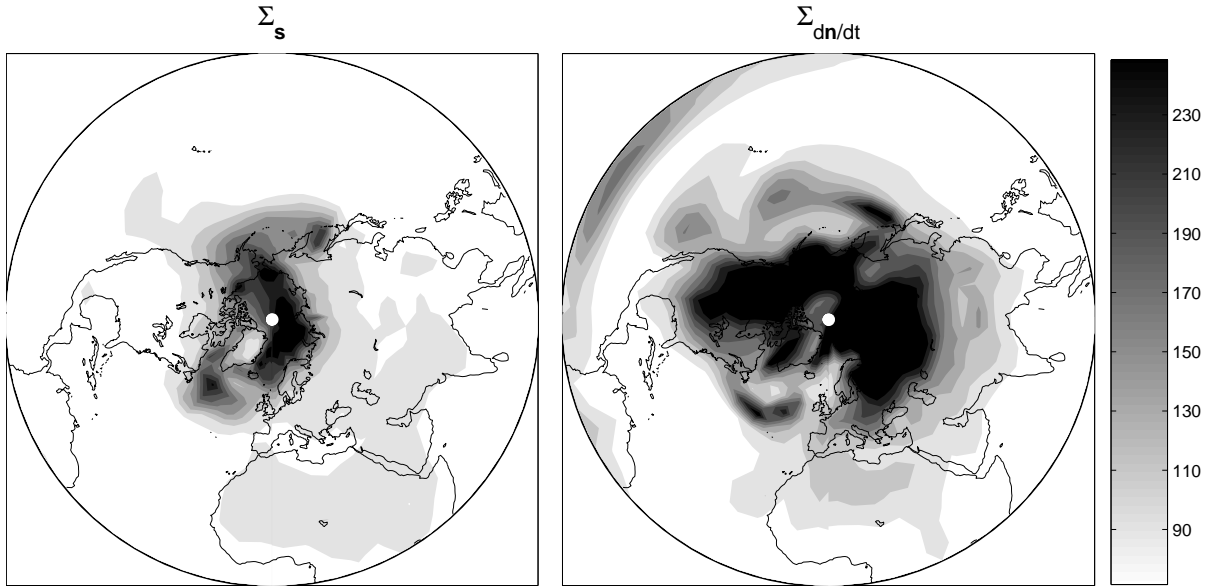
FIG. 3. The variance of natural variability $\Sigma_{d\mathbf{n}/dt}$ and signal uncertainty $\Sigma_\mathbf{s}$ for prediction of Central U.S. surface air temperature trend $d\alpha/dt$. The signal uncertainty is established using signals from 21 models that contributed to CMIP3. Those signals were determined by linear regression Northern Hemisphere surface air temperature of the first 50 years of SRES-A1B runs divided by the trend in surface air temperature over the Central U.S. over the same time interval. Both are dimensionless, the natural variability $\Sigma_{d\mathbf{n}/dt}$ having been normalized by the estimated trend $(d\alpha/dt)^2$.

rope, analysis of a ten year timeseries of "observations" ($\Delta t = 10$ yrs) yields a predicted warming of 0.35 ± 0.43 K, significantly lower than the true value of 1.26 ± 0.03 K obtained from model output. Increasing the length of the observational timeseries to $\Delta t = 20$ yrs provides a prediction of 0.59 ± 0.23 K, a substantial improvement in qualitative agreement with the true value but quantitatively it is three-sigma away from truth. This large disagreement indicates that fundamental information is missing from this prediction. Most likely this shortcoming is due to a combination of small areal extent of the target region, which contains only three model gridpoints for the reduced grid, and the relative paucity of models used to prescribe the mean signal shape $\bar{\mathbf{s}}$ and the signal shape uncertainty $\Sigma_\mathbf{s}$. The contravariant fingerprint reduces the natural variability effectively but makes only a marginally quantitatively robust temperature forecast.

3) THE SAHEL

The contravariant fingerprint for temperature in the Sahel shows an intermediate sensitivity to local effects, between the sensitivities exhibited by the Central United States and Northern Europe. On the large scale, the

fingerprint exhibits a subtropical-equatorial dipole that measures the expansion of the Hadley circulation. The increase in cold air mass over the central and eastern tropical Pacific is consistent with decreased subsidence, a measure of the weakening Walker circulation. The analysis of $\Delta t = 10$ years of "observations" yields a prediction of 1.37 ± 0.46 K, so that the one-σ uncertainty envelope encompasses the true value of 1.73 K. For $\Delta t = 20$ years, the prediction is 1.52 ± 0.16 K.

b. Mixed data types

We have argued that this method of scalar prediction is generalizable to arbitrary data types. This property can be particularly useful when homogeneous, accurate measurements of a scalar quantity of interest are not available, but measurements of a different sort with a strong physical relationship to the desired scalar exist. We demonstrate that behavior by utilizing simulated observations of microwave brightness temperature representative of the mid-troposphere (hereafter abbreviated TMT), as would be measured by channel 2 of a Microwave Sounding Unit-type (MSU) satellite sensor. We compute TMT using a radiative transfer model applied to upper air temperature, specific humidity, sur-
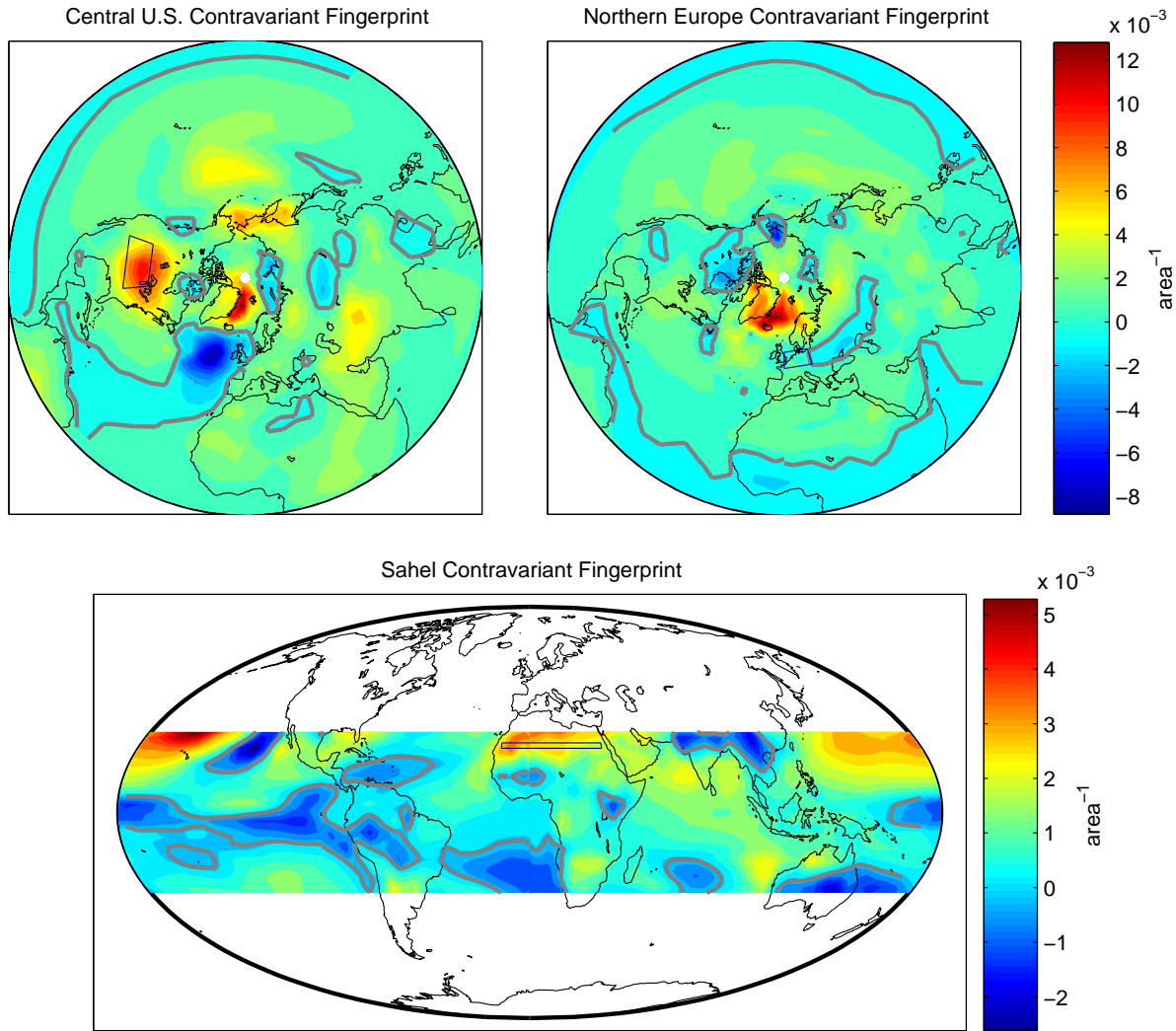
9

FIG. 4. The contravariant fingerprints **f** for scalar prediction of (a) Central U.S. surface air temperature trend, (b) Northern Europe surface air temperature trend, and (c) Sahel surface air temperature trend.

face pressure, surface air temperature, and surface emission temperature for a subset (14 models, one realization each) of the same model runs we used in the surface air temperature-only exercise above. This calculation results in two-dimensional spatial maps of TMT, which are regridded to the same reduced resolution as the surface air temperature maps.

We incorporate the TMT data into the analysis in three different ways. The first method includes large-scale maps of both surface air temperature observations and TMT observations to form a joint observational dataset. The second method includes surface air temperature observations drawn from target region alone together with large-scale spatial maps of TMT. The final methods includes TMT observations alone. In each

case, the TMT observations were treated according to the same procedure as surface temperature. First a signal in TMT is found from linear trend analysis of 50 years of annual average 21st century forecast data. Then a normalized signal is computed by dividing this TMT trend map by the trend in the scalar surface air temperature for the region of interest.

The resulting contravariant fingerprint for the first method, consisting of a map corresponding to TMT observations and a map corresponding to contemporaneous surface air temperature observations, in the case of the Central United States, is shown in Fig. 6. The surface air temperature component of the joint contravariant fingerprint preserves the key features of the surface air temperature-only contravariant fingerprint: peak
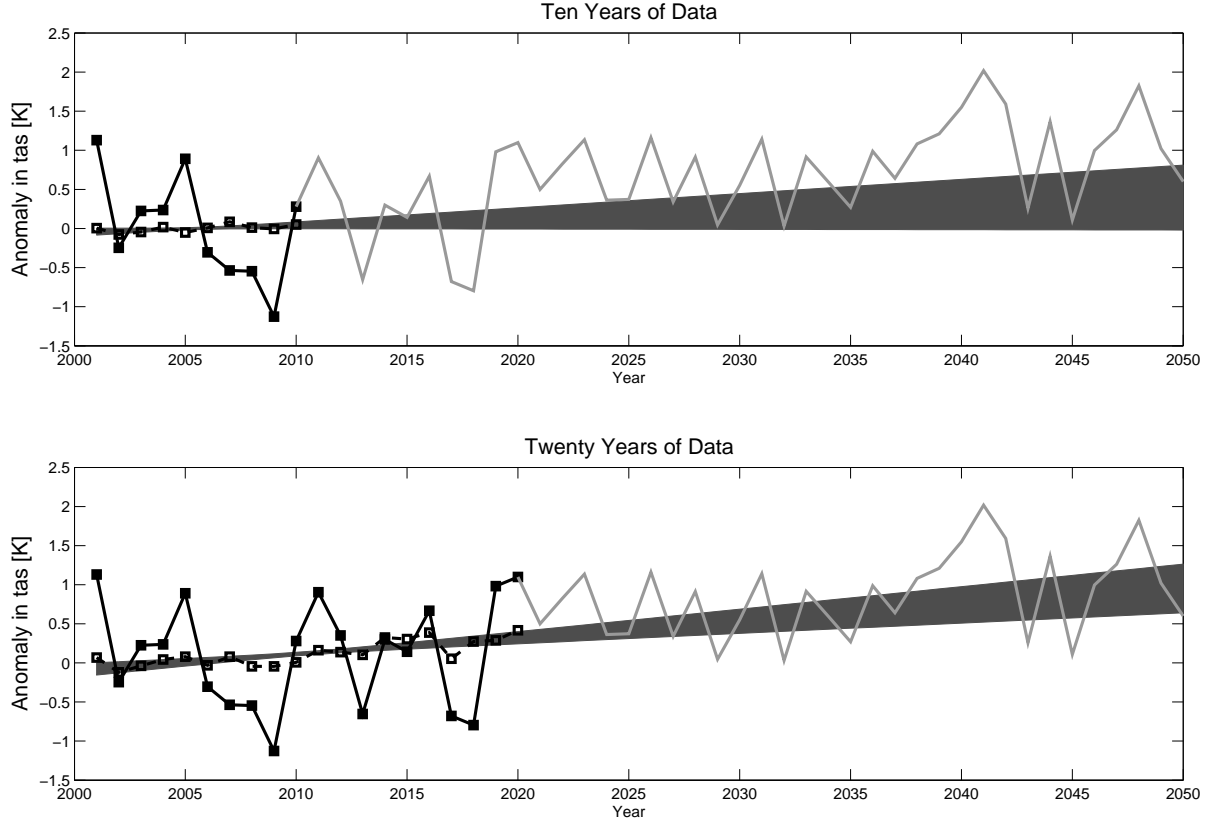
FIG. 5. Application of Central U.S. surface air temperature prediction to "observations" taken from the output of a forced run of a climate model. The climate model used as a stand-in for data was not included in the ensemble used to compose signal uncertainty $\Sigma_{\mathbf{s}}$. The natural variability was composed from a present day control run of NCAR CCSM3. The black line and filled squares are the "observations", defined to be surface air temperature averaged over the Central U.S. The dashed line and open squares are the detectors $\mathbf{f}^{T}\,\mathbf{d}(t)$, the application of the contravariant fingerprint to Northern Hemisphere surface air temperature year by year. The gray envelope shows the forecast for the underlying trend in surface air temperature over the Central U.S. computed according to Eqs. 27 and 28, and the light gray line shows the future evolution of surface air temperature over the Central U.S. Plot (a) shows the result for ten years of data and plot (b) for twenty years of data. Notice the ability of this technique to suppress natural variability and produce a forecast that is accurate and four times more precise than would have been obtained from extrapolating the surface air temperature over the Central U.S. only.

sensitivity over the central United States and strong sensitivity to temperature in Arctic regions correlated with areas of ice extent. The most prominent feature in the TMT fingerprint is a nearly annular pattern centered around 70°N which corresponds to a prominent warming around 500 hPa that may be seen in zonal cross-sections of upper air temperature trends for these models under consideration here. This feature probably is a measure of poleward migration of the polar jet.

Both ten-year and twenty-year prediction see degraded precision when adding TMT to surface air temperature. Addition of information in generalized scalar prediction, inasmuch as it is a type of optimal estima-

tion, should improve the precision of prediction. Because the ratio of TMT variability to surface air temperature variability realized in the "observations" differ substantially than the same ratio in the prescription of natural variability by a different model, a proper understanding of information content is not realized and the suppression of natural variability is unable to handle the two data types appropriately. This is borne out by an increased in the post-fit residuals of the detectors σ_{α}^{2} in adding the Northern Hemisphere TMT field to the Northern Hemisphere surface air temperature field. This points toward the necessity of having the prescription of natural variability capturing structures of variability that
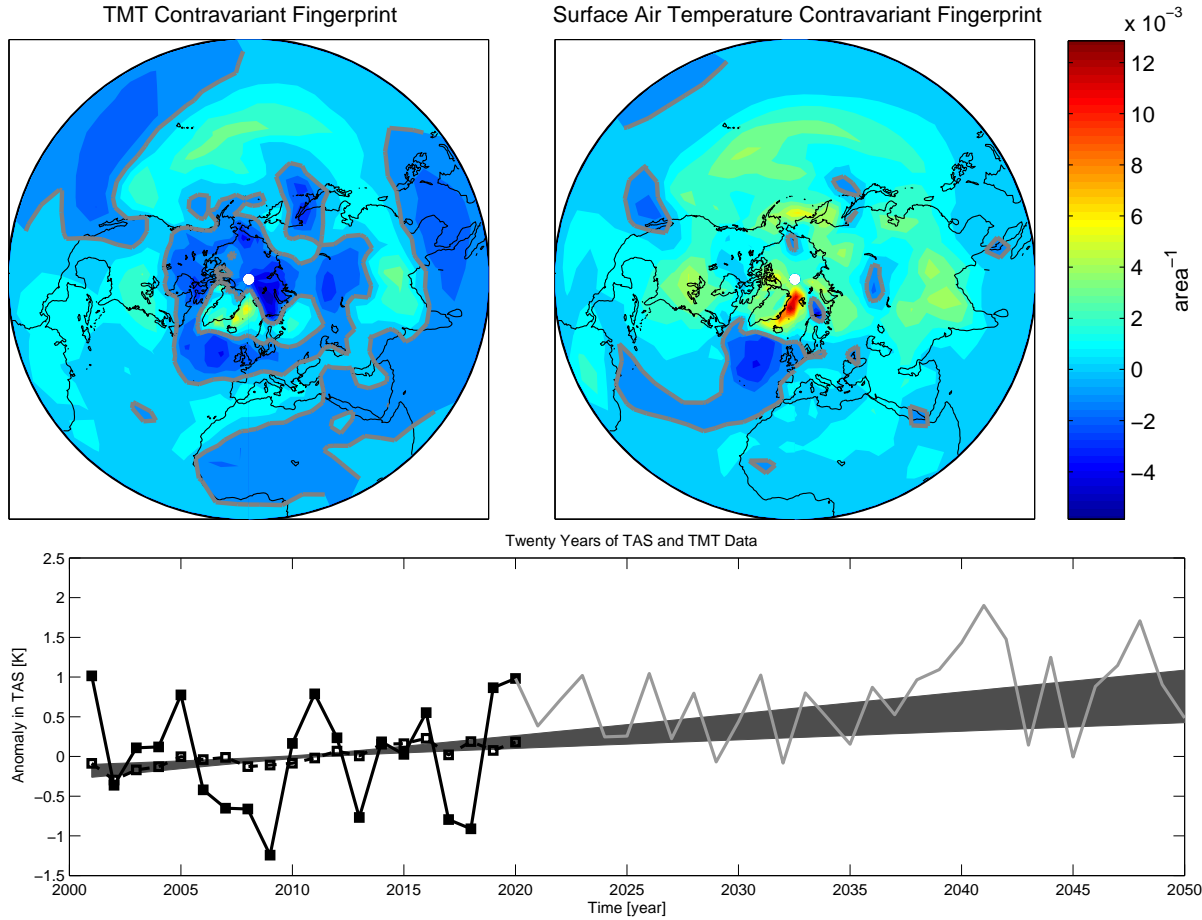
FIG. 6. Joint surface air temperature-TMT fingerprint for Central U.S. surface air temperature prediction. The data field is taken to be Northern Hemisphere surface air temperature (TAS) and Northern Hemisphere mid-tropospheric temperature (TMT) as measured by the Microwave Soundings Units aboard the NOAA/TOVS satellites. The top two plots show the fingerprints in surface air temperature space and MSU TMT space. The lower plot follows the same scheme as in Fig. 5. Notice that inclusion of TMT into the data space adds little information to the forecast given 20 years of data. This means that long time records of Northern Hemisphere TAS and Northern Hemisphere mid-tropospheric temperature contain substantially redundant information.

are useful in optimization.

For the first method, utilizing global surface air temperature and TMT data together, the forecast with $\Delta t = 10$ years for year 50 is 1.12 ± 0.62 K, a substantial improvement over the result from the inclusion of surface air temperature observations alone. For $\Delta t = 20$ years, the forecast for year 50 is 1.10 ± 0.35 K, which is a minor degradation in accuracy and uncertainty over the case without TMT. For the second method, which joins TMT with surface air observations for the just target region, the forecast for year 50 is 1.20 ± 1.26 K, also improved relative to the case with no TMT, but with much larger uncertainty than for the case with global surface

observations. From this we conclude that important additional information about the regional trend is embedded within TMT variable. In the third method, utilizing just TMT with no surface observations, the forecast for year 50 is 1.40 ± 1.33 K. This uncertainty is still a substantial improvement from that obtained from a pure extrapolation of the surface observations themselves for the target region (equal to ± 2.96 K), and much more accurate as well, capturing the correct sign for the trend.

The results for the application of these methods to the target regions of the Central United States and Northern Europe are summarized in Table 2. The overall accuracy of the scalar prediction is degraded some-

12

Target Region	Data Fields	Baseline (years)	Prediction at 50 years (K) ± one-sigma uncertainty
Central United States	TAS NH	10	0.43 ± 0.40
Central United States	TAS and TMT NH	10	1.12 ± 0.62
Central United States	regional TAS and TMT NH	10	1.20 ± 1.26
Central United States	TMT NH only	10	1.40 ± 1.33
Central United States	TAS NH	20	1.19 ± 0.30
Central United States	TAS and TMT NH	20	1.10 ± 0.35
Central United States	regional TAS and TMT NH	20	2.40 ± 0.64
truth			1.31 ± 0.03
Northern Europe	TAS NH	10	0.35 ± 0.43
Northern Europe	TAS and TMT NH	10	0.82 ± 0.43
Northern Europe	regional TAS and TMT NH	10	0.88 ± 0.94
Northern Europe	TAS NH	20	0.59 ± 0.23
Northern Europe	TAS and TMT NH	20	0.69 ± 0.27
Northern Europe	regional TAS and TMT NH	20	1.79 ± 0.46
truth			1.26 ± 0.03

TABLE 2.

what for the case of Northern Europe, which is limited to only 3 model gridpoints, as opposed to 8 for the Central United States. The results for Northern Europe still show great improvement, however, over the results of extrapolation from the observations alone. For both target regions, the prediction is improved by additional observations for the case of $\Delta t = 10$ years. For the case of $\Delta t = 20$ years, the prediction is not substantially degraded by additional observations, although it is not helped. From this pattern, we conclude that additional measurements improve the prediction's uncertainty contributed by natural variability that is particularly pronounced for a short timeseries of observations. When a longer timeseries of observations is available, the influence of natural variability diminishes with respect to signal uncertainty; thus, TMT contributes less information but nevertheless still does not degrade the prediction. In the case of the extended observational period, however, when measurements are removed that contain the global spatial details that correctly constrain the scalar trend, the predictions including subsidiary measurements lose accuracy and precision. This property serves to inform prediction design based on the availability and quality of measurements.

c. Precipitation

For precipitation, we will also consider scalar prediction for the trend in annual average precipitation for the same three geographical regions: the Central United States, Northern Europe, and the Sahel, as defined in Table 1. The models do not predict a consistent sign for the trend in Central United States precipitation, so we choose two models as "observations" that capture the strongest trends, both positive and negative (c.f. Fig. 1). From this exercise we conclude that the formalism does not have a substantial difference in predictive performance based on sign. The model ensemble distributions of precipitation for Northern Europe and the Sahel are not so nearly symmetric about zero, so we present the results for just a single model as "observations" for these regions. For these regions, we employ the same model that provides the large negative trend case for the Central United States.

1) CENTRAL UNITED STATES

Annual average precipitation contrasts temperature in that the models do not predict a consistent sign for the trend for the Central United States (Fig.1). We consider two end members of the corresponding multi-model distribution of precipitation trends as test cases for scalar prediction. The INM CM 3.0 model predicts a linear trend in annual precipitation of about $+13$ mm decade^{-1} in the Central United States based on a single realization. The MIROC 3.2 medium resolution model predicts a linear trend in annual precipitation of -15 mm decade^{-1} based on three realizations.

The inconsistency in sign among the model precipitation trends manifests itself in the substantially larger ratio of signal shape uncertainty to natural variability relative to temperature. In Fig. 7 we show the natural variability and signal uncertainty fields associated with prediction of precipitation in the Central United States for a 20-yr timeseries. Additionally, the precipitation case is distinguished by the dominance of the
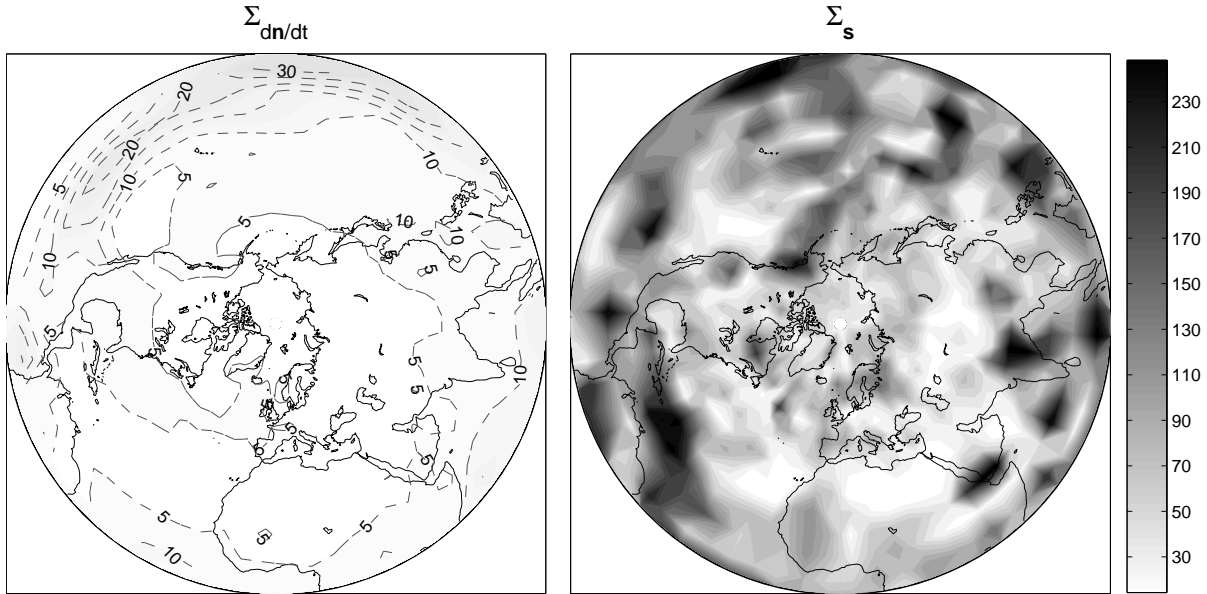
13

FIG. 7. The signal uncertainty $\Sigma_{\mathbf{s}}$ and natural variability $\Sigma_{d\mathbf{n}/dt}$ associated with the Central United States (see Table 1 for precise definition) precipitation. The diagonal elements of the dimensionless covariances are shown. The same procedures were followed to form these covariances as were used for temperature, and the same CMIP3 models were used (although the number of realizations per model varied slightly). In contrast to temperature, the signal uncertainty is much larger than the natural variability. Both the signal uncertainty and the natural variability are dominated by the Tropics. The signal shape uncertainty has significantly more regional scale spatial structure than temperature.

Tropics in the case of both natural variability and signal shape uncertainty, even when considering a precipitation trend at mid-latitudes. These differences between precipitation and temperature are also strongly evident in the contravariant fingerprint $\mathbf{f}$ for precipitation prediction (Fig. 8). In the case of the Central United States, there is moderate sensitivity to the local precipitation behavior and substantial sensitivity to precipitation on the Pacific coast of North America centered near $45°$N. Beyond these correlations, however, the major predictors are to be found in the subtropical eastern Pacific, the tropical western Pacific, and the Indian Ocean basin. This contravariant fingerprint efficiently suppresses natural variability by a factor of five or more for the case of both increasing and decreasing precipitation.

Determination of a contravariant fingerprint for precipitation is hindered by the limited size of the CMIP3 ensemble we use to estimate signal uncertainty. The mottled pattern of signal uncertainty as it appears in Fig. 7 indicates that the physics of precipitation is highly uncertain in climate modeling. The fact that patches of low uncertainty are intertwined with patches of high uncertainty does not mean that the patches of low uncertainty point toward confidence in emergence of a pre-

cipitation signal. With an ensemble with as few as 24 members, one should expect a mottled pattern of uncertainty rather than a uniform blanket of uncertainty. With an ensemble of hundreds to thousands of runs, only then should a uniform blanket of uncertainty appear. The limited size of the CMIP3 ensemble should, then, permit some suppression of natural variability in the regions with relatively certain model physics that would be inappropriate to apply to real data.

As seen in Fig. 9, the 30-yr forecast for INM CM3.0 is $+7 \pm 21$ mm, nearly three-sigma from the true value of $+66 \pm 1$ mm. This statistically improbable occurrence is likely an indication that both the ensemble lacks enough members to fully span the space of possible precipitation patterns, and that important information for precipitation is missing from the analysis, information on changes in the large-scale mid-latitude circulation. The forecast for the first realization of the MIROC 3.2 medium resolution model is -3 ± 13 mm, which encompasses the true value of -5.5 ± 0.6 mm. Note that we arbitrarily chose the first realization which shows a substantially reduced drying relative to the mean of the three realization. This prediction is accurate but the precision is very low compared to the case of the temper-
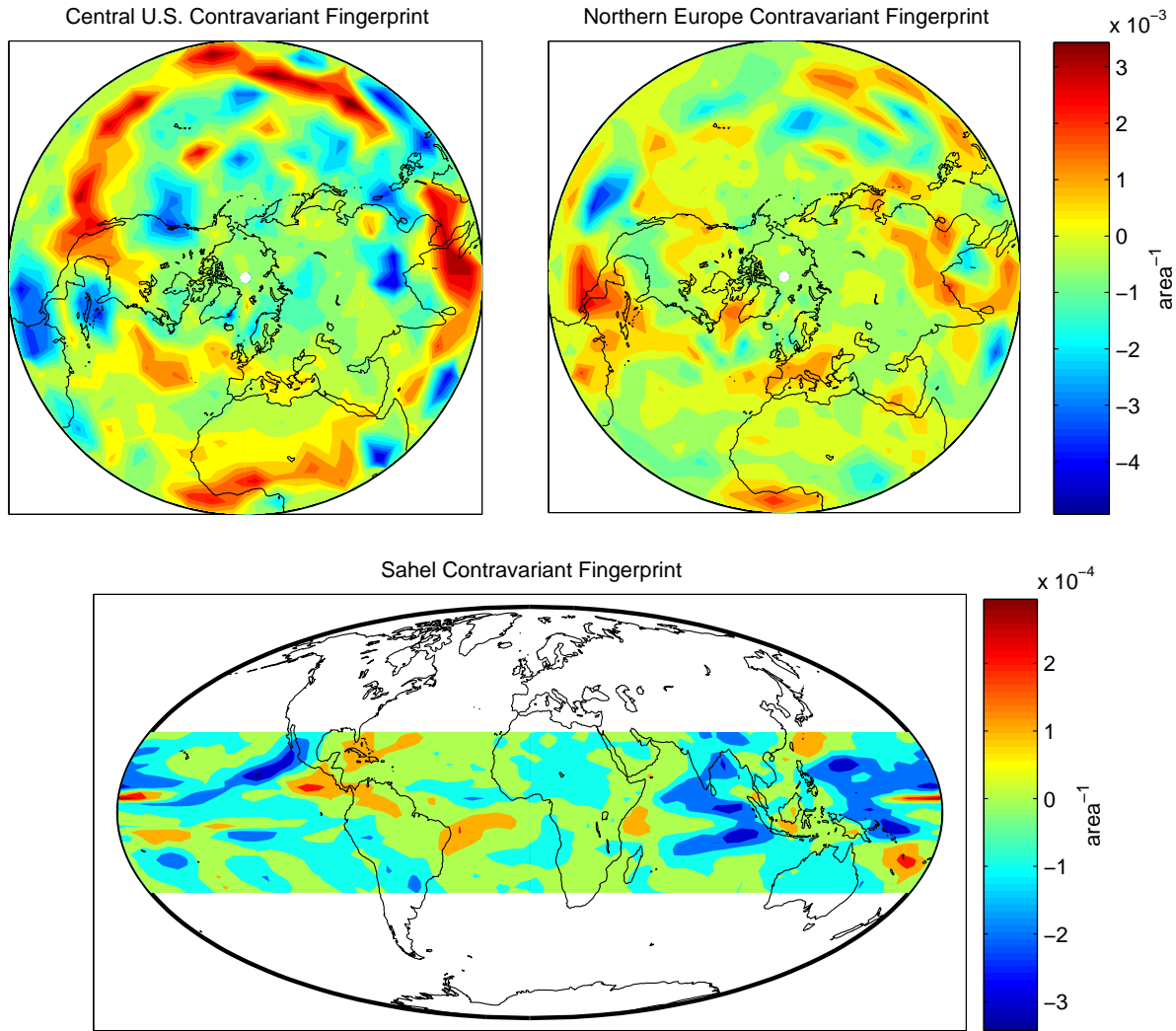
14

FIG. 8. Contravariant fingerprints for regional precipitation prediction of the Central United States (upper left), Northern Europe (upper right), and the Sahel (lower).

ature forecasts. This result is a direct consequence of models' extreme uncertainty in the physics of precipitation relative to the more certain physics of temperature change due to anthropogenically induced greenhouse forcing.

2) NORTHERN EUROPE

The precipitation trend in Northern Europe, as revealed by its contravariant fingerprint, is subject to more local influence than the trend in temperature. It is strongly affected by the precipitation trend over the majority of the European landmass and an eastward shift in precipitation over the subtropical eastern Pacific indicated by adjacent areas of increasing and decreasing am-

plitude. Overall there are scattered contributions from the Gulf of Guinea, southeast Asia, and the western tropical Pacific, but these contributions are much lower in amplitude relative to the low-latitude dependences in the Central United States fingerprint. The analysis of 20 years of "observations" generates a prediction of $+14.8 \pm 15.6$ mm increase in precipitation (Fig. 10), compared to a true value of $+13.0 \pm 0.8$ mm.

3) THE SAHEL

The contravariant fingerprint for Sahelian precipitation shows adjacent areas of increasing and decreasing precipitation in the tropical Indian and Pacific Ocean basins that is largely the result of inconsistent spatial
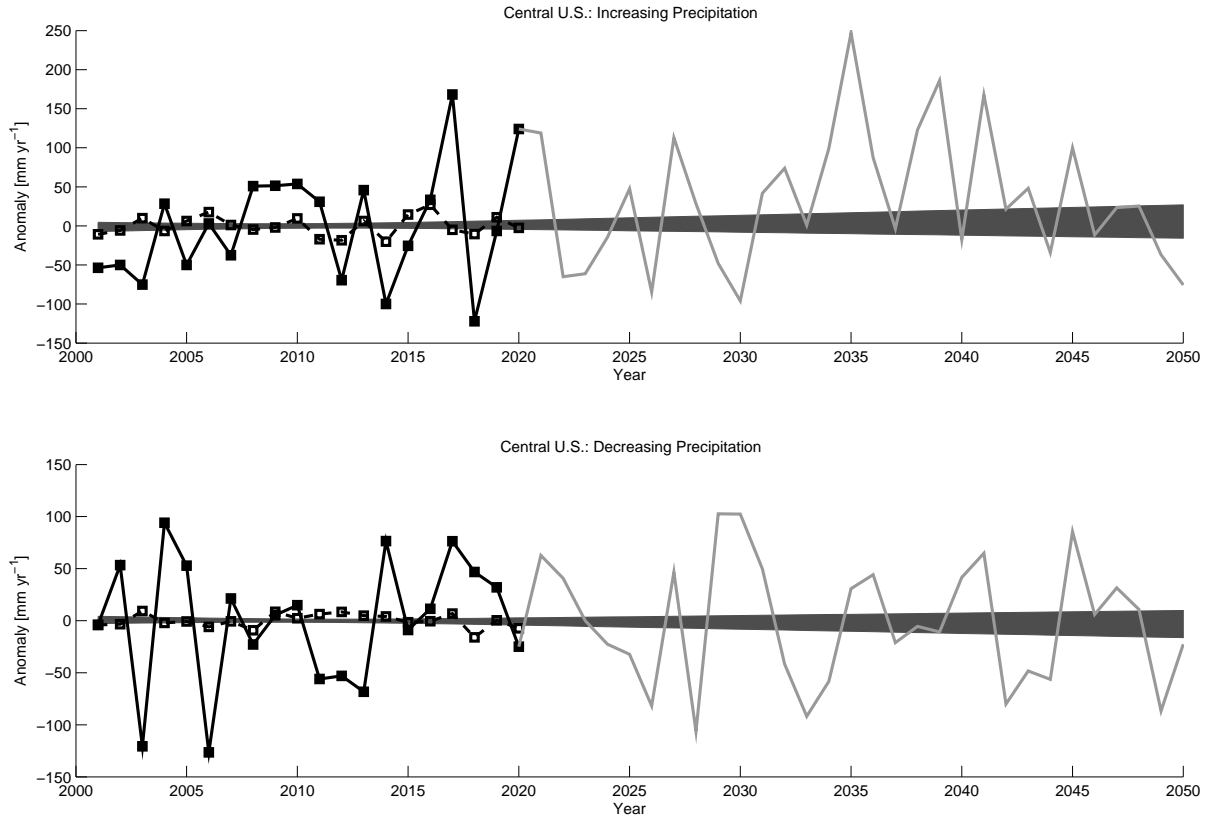
15

FIG. 9. Two cases of "observations" are presented, following the same scheme as Figs. 5 and 6, but with precipitation in the Central United States as the target scalar. The upper panel shows a model realization of the 21st century with increasing precipitation over fifty years drawn from INM CM 3.0, and the lower panel shows decreasing precipitation for the same period drawn from MIROC 3.2 (medium resolution). The timeseries of regional precipitation are significantly more variable than temperature, but the contravariant fingerprint still provides substantial reduction in interannual variability.

patterns of precipitation trends among the multi-model ensemble. Most of the individual models show a correlation between overall decreases in precipitation in the ITCZ and decreases in precipitation in the Sahel. The spatial disparity in the pattern of ITCZ precipitation change among the models results in the areas of alternating sign in precipitation change about the Equator seen in the fingerprint. Application of the contravariant fingerprint for $\Delta t = 20$ years generates a prediction of $+0.8 \pm 8.6$ mm (c.f. Fig. 10), compared with the true value of -8.1 ± 0.2 mm. The true value lies just outside the one-sigma uncertainty envelope.

The contrast between the performance in temperature and precipitation illustrates a key points about our method of scalar prediction. The method relies on the relative robustness of trend patterns, as opposed to the magnitudes of the trends themselves, in making accurate predictions. The performance is therefore strong in

the case of temperature because the agreement among the model patterns is good compared with the amplitude of natural variability (Fig. 3). On the other hand, for precipitation, signal shape uncertainty strongly outweighs natural variability, and trend forecasts are substantially more uncertain. A path to potential improvement for the precipitation forecasts, based on the results of the inclusion of TMT in the temperature analysis, is the addition of more data types. In particular, recent work has shown that sea surface temperature plays a strong role in precipitation trends in the Sahel (Yoshioka et al. 2007), and that upper air wind patterns are highly correlated with extreme precipitation events in Northern Europe. Thus a future study incorporating surface temperature and geopotential height into a precipitation analysis would be enlightening in this regard.
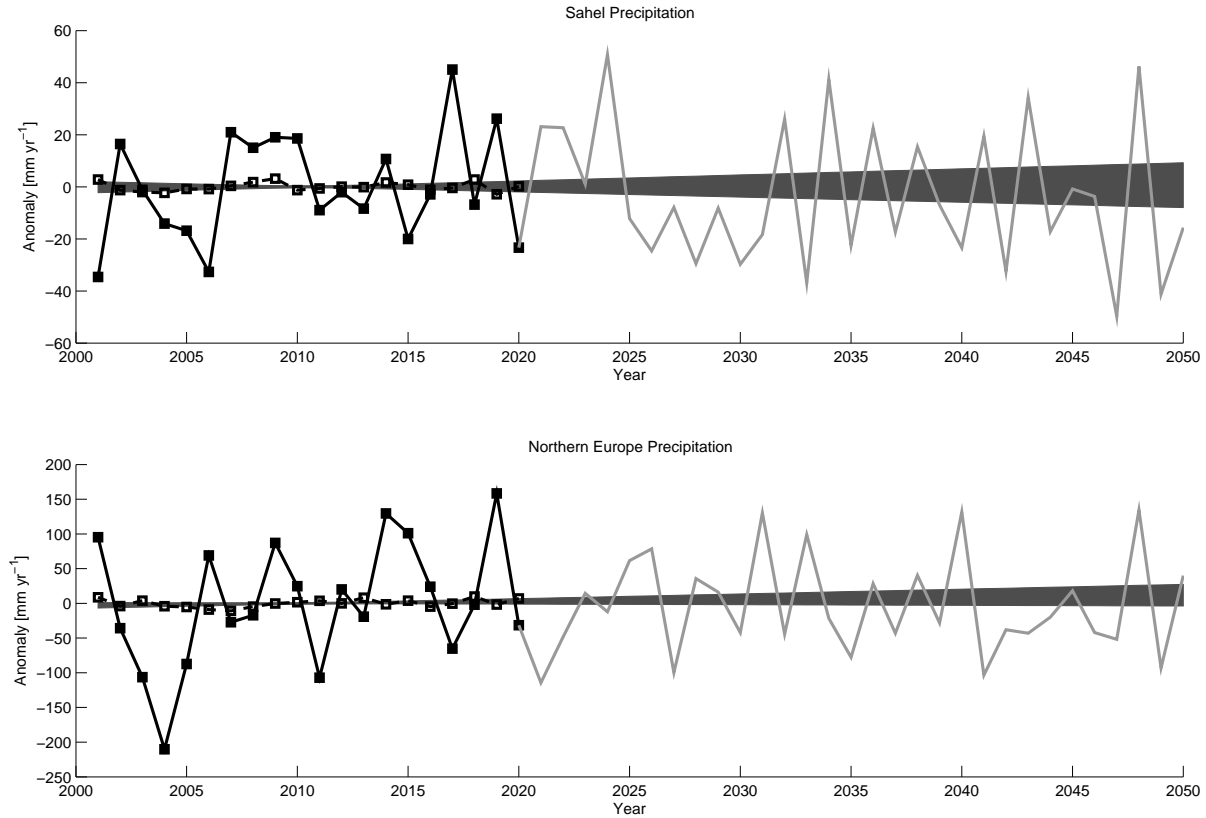
16

FIG. 10. The upper panel follows the same scheme as Figs. 5 and 6 but for 20 years of data on precipitation in the Sahel. The true change in precipitation at year 50 of -7.76 mm is consistent with the scalar prediction of $+0.86 \pm 8.61$ mm. The lower panel shows the same but for precipitation in Northern Europe. Here the true change of $+13.02$ mm is easily consistent with the predicted value of $+14.80 \pm 15.57$ mm. For both these cases, the scalar prediction method strongly reduces uncertainty due to nautral variability and represents a qualitative improvement over extrapolation of local observations.

4) Summary and discussion

We have presented a method for predicting arbitrary scalars of the climate system given a timeseries of any data type. The equations are derived from the standard approaches of ensemble prediction of climate change but subject to the assumption that climate models predict patterns of change better than they predict the magnitude of change, a fundamental assumption of optimal fingerprinting/linear multi-pattern regression. This assumption is reasonable because there are very few observable variables whose pattern of change explains the physics underlying the transient sensitivity of the climate. Consequently, it is safe to assume the transient sensitivity of the climate is only weakly constrained by climate models and quantifiable only from timeseries of data. Eqs. 10, 11, and 12 are the equations of generalized scalar prediction for a single scalar, and Eqs. 16,

17, and 18 are the equations of generalized scalar prediction for multiple scalars. A contravariant fingerprint is determined such that, when it is convolved with the data, an accurate and precise estimate of the underlying multi-decadal trend in the scalar is found. In optimal detection, natural variability as determined by a climate model is implicitly used to estimate the uncertainty and confidence bounds on detection. It is frequently the case that climate models only poorly simulate inter-annual climate variability, so in Eqs. 23 and 26 we present a method for deriving scalar trend uncertainty that is only minimally dependent upon a climate model.

Generalized scalar prediction preferentially weights those subspaces of a data type in which there is little natural variability and there is strong certainty in the physics relating the data to the scalar of interest. A long control run of a climate model is used to estimate the natural variability, and an ensemble of forced runs of

a climate model is used to estimate the uncertainty of the physics of climate change. The ensemble of forced runs is generated by either varying the tunable free parameters of the climate model within physically reasonable values or incorporating different physically plausible parameterizations of sub-gridscale processes. When the physics of climate change relevant to a specified scalar is relatively certain with respect to natural variability, generalized scalar detection selects subspaces of the data type where the physics relating the scalar and the data is strong and the influence of natural variability is minimal. If the data space includes the scalar of interest, a significant component of the contravariant fingerprint strongly considers the past history of that scalar (surface air temperature of the Central United States) unless that scalar is naturally highly variable (surface air temperature of Northern Europe). In other cases where the physics of scalar change is less certain, the weighting is strongly toward the historical trend in that scalar.

The relative weighting of physical relationships of climate system and natural variability in generalized scalar prediction is also a function of the length of the timeseries of data available. Natural variability is anticipated to be a stationary random process in time, so the longer the timeseries, the less important natural variability should be in trend estimation. As a consequence, a shorter timeseries of data will more strongly weight toward consideration of natural variability in the construction of the contravariant fingerprint so that the influence of natural variability can be suppressed in climate prediction. Longer timeseries of data will more preferentially consider uncertainty in model physics, because uncertainty in model physics is not impacted by length of the timeseries while natural variability becomes less significant. Therefore, the timeseries of detectors—the result of multiplying individual timesteps of data by the contravariant fingerprint—for shorter timeseries will show more suppression of natural variability than for longer timeseries of data. In the end, longer timeseries of data always give more precise forecasts of climate change than shorter timeseries.

Generalized scalar prediction handles not only many different types of data, it also handles them jointly. The addition of new data types is expected to add new information to the scalar prediction. Just how much information depends on the data type. If the data type reveals new aspects of climate change physics relevant to the scalar in question, the data type should contribute significantly to the reduction of uncertainty in scalar prediction. Also, if the data type is sensitive to components of the climate system that are less naturally variable, it too should contribute to the reduction of uncertainty in scalar prediction. In the latter case, the contribution of the new data type becomes less valuable with a longer timeseries of data because the increase in length of the timeseries by itself acts to reduce the influence of natural variability in scalar prediction. This is illustrated by the addition of mid-tropopheric microwave brightness temperature (TMT) to the surface air temperature record in predicting trends in surface air temperature.

Practical application of generalized scalar prediction depends on a satisfactory prescription of naturally occurring inter-annual variability of the climate system. The suppression of natural variability, or optimization, depends on the prescription of natural variability correctly capturing patterns of variability and not necessarily on the overall amplitude of variability. If patterns of variability, represented by the dominant eigenmodes of $\Sigma_\mathbf{n}$, are incorrect, spurious optimization will take place, and addition of new data types to scalar prediction might degrade scalar prediction. This indeed is what occurred when adding the Northern Hemisphere TMT field to the Northern Hemisphere surface air temperature field as the data in predicting Central United States surface air temperature trends. The amplitude of natural variability enters only in establishing the relationship between signal uncertainty and natural variability in Eqs. 10 and 16, and this is only relevant at the level of order of magnitude. In most circumstances, either signal uncertainty is much greater than natural variability or vice versa, thus establishing dominance in evaluating Σ without regard to the actual amplitude of natural variability. Moreover, the amplitude of natural variability plays little role in the estimate of uncertainty in the scalar trend, because natural variability's contribution to the uncertainty is determined by convolution of the contravariant fingerprint onto the actual data (c.f. Eqs. 23 and 26).

Practical application of generalized scalar prediction depends strongly on a complete accounting for the community's collective uncertainty in the physics of anthropogenic climate change. In the examples presented in this paper, we used the CMIP3 ensemble of 24 climate models which are somewhat but not fully independent. As a consequence, not all of our actual uncertainty is accounted for in the CMIP3 ensemble: real data might exhibit physics which is not included in the physical space spanned by the CMIP3 models. A more appropriate ensemble would include several thousands of forced runs of a climate model, each run incorporating different yet reasonable values for the physical parameters representing processes that are unresolved by the model's spatio-temporal grid.

Practical application most likely also depends upon prediction using multiple scalars. Perhaps the most obvious application is forecasting of inter-decadal climate trends due to anthropogenic forcing. Anthropogenic

forcing of the climate system can take many forms, the most prominent of which are increasing well-mixed greenhouse gases, sulfate aerosols, black carbon resulting from agricultural burning, and changing land use. Social policy can influence each of these types of forcing differently, so generalized scalar prediction should incorporate scalars that distinguish between these different types of forcing independently. Then, given varying scenarios for how these different forcings will evolve over the coming decades, we would have sufficient information to formulate a climate forecast that considers all of these types of forcings. In this application, then, it is essential that, for each prescription of model physics, individual runs be executed for radiative forcing by increasing well-mixed greenhouse gases alone, increasing sulfate aerosols alone, increasing black carbon alone, and changing land use alone.

The equations of generalized scalar detection are nearly the same as those of signal detection and attribution (Bell 1986; Huntingford et al. 2006). Both signal detection and attribution and generalized scalar prediction assume a model of a slow secular trend underlying a timeseries of data with large naturally occurring interannual variability superimposed. Mathematically, there is nothing inherent to signal detection and attribution that dictates how signals should be normalized before the signal uncertainty matrix Σ_s is evaluated. Generalized scalar detection shows that normalization depends on the application, namely that the normalization should be by the scalar for which a *prediction* is desired.

Results of the application of generalized scalar prediction, then, can be interpreted in the light of both signal detection and attribution and scalar prediction. In the context of signal detection and attribution, for instance, trends in data relevant to regional scale can be attributed to anthropogenic greenhouse warming with an interesting paradox. It is possible that a region exhibits a cooling trend in a timeseries of regional data yet have generalized scalar detection find a positive trend, as in the example of surface air temperature in the Central United States and just ten years of data (Fig. 5, top). While it might seem odd to declare that global warming is occurring in this region even though it has cooled, nonetheless, the technique has not only considered the regional data but also physically reliable indicators external to the region that are useful in tying trends to the physics of global warming. The individual researcher who is uncomfortable attributing regional cooling to an anthropogenically enhanced greenhouse when regional warming is expected may feel more comfortable applying signal detection and attribution with the data domain restricted to just the region of interest, as in Kharin and Zwiers (2002). Restricting signal detection and attri-

bution strictly to the domain of interest, though, would likely produce less appealing results for prediction of future trends.

Finally, we have assumed throughout that the underlying trends in climate will be linear in time over the coming decades. This is partially justified by the simulations of CMIP3. In reality, though, we do not yet know how the anthropogenic radiative forcings of climate will evolve in the coming decades. Thus, some research will be necessary to determine how past trends in climate translate to prediction if rates of anthropogenic radiative forcing change in the future.

References

Allen, M., N. Gillett, J. Kettleborough, G. Hegerl, R. Schnur, P. Stott, G. Boer, C. Covey, T. Delworth, G. Jones, J. Mitchell, and T. Barnett, 2006: Quantifying anthropogenic influence on recent near-surface temperature change. *Surv. Geophys.*, **27**, 491–544.

Allen, M. and S. Tett, 1999: Checking for model consistency in optimal fingerprinting. *Climate Dyn.*, **15 (6)**, 419–434.

Bell, T., 1986: Theory of optimal weighting to detect climate change. *J. Atmos. Sci.*, **43**, 1694–1710.

Giorgi, F. and R. Francisco, 2000a: Evaluating uncertainties in the prediction of regional climage change. *Geophys. Res. Lett.*, **27 (9)**, 1295–1298.

———, 2000b: Uncertainties in regional climate change prediction: a regional analysis of ensemble simulations with the HADCM2 coupled AOGCM. *Climate Dyn.*, **16**, 169–182.

Giorgi, F. and L. Mearns, 2002: Calculations of average, uncertainty range, and reliability of regional climate changes from AOGCM simulations via the "reliability ensemble averaging" (REA) method. *J. Climate*, **15**, 1141–1158.

———, 2003: Probability of regional climate change based on the reliability ensemble averaging (REA) method. *Geophys. Res. Lett.*, **30 (12)**, 1629, doi:10.1029/2003GL017130.

Giorgi, F., P. Whetton, R. Jones, J. Christensen, L. Mearns, B. Hewitson, H. von Storch, R. Francisco, and C. Jack, 2001: Emerging patterns of simulated regional climatic changes for the 21st century due to anthropogenic forcings. *Geophys. Res. Lett.*, **28 (17)**, 3317–3320.

Hasselmann, K., 1993: Optimal fingerprints for the detection of time-dependent climate change. *J. Climate*, **6 (10)**, 1957–1971.

———, 1997: Multi-pattern fingerprint method for detection and attribution of climate change. *Climate Dyn.*, **13 (9)**, 601–611.

Houghton, J., Y. Ding, D. Griggs, M. Noguer, P. van der Linden, X. Dai, K. Maskell, and C. Johnson, (Eds.) , 2001: *Climate Chance 2001: The Scientific Basis, Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, New York, 881 pp.

Huntingford, C., P. Stott, M. Allen, and F. Lambert, 2006: Incorporating model uncertainty into attribution of observed temperature change. *Geophys. Res. Lett.*, **33**, doi:10.1029/2005GL024831.

Kharin, V. and F. Zwiers, 2002: Climate predictions with multimodel ensembles. *J. Climate*, **15**, 795–799.

Leroy, S., 1998: Detecting climate signals: Some Bayesian aspects. *J. Climate*, **11 (4)**, 640–651.

Leroy, S., J. Anderson, J. Dykema, and R. Goody, 2008: Testing climate models using thermal infrared spectra. *J. Climate*, In Press.

Min, S. and A. Hense, 2006a: A Bayesian approach to climate model evaluation and multi-model averaging with an application to global mean surface temperatures from IPCC AR4 coupled climate models. *Geophys. Res. Lett.*, **33**, L108708, doi:10.1029/2006GL025779.

———, 2006b: A Bayesian assessment of climate change using multimodel ensembles. Part I: Global mean surface temperature. *J. Climate*, **19**, 3237–3256.

Min, S., D. Simonis, and A. Hense, 2007: Probabilistic climate change predictions applying Bayesian model averaging. *Phil. Trans. R. Soc. A*, **365**, doi:10.1098/rsta.2007.2070, 2103–2116.

North, G., K. Kim, S. Shen, and J. Hardin, 1995: Detection of forced climate signals: I. Filter theory. *J. Climate*, **8 (3)**, 401–408.

Räisänen, J. and T. Palmer, 2001: A probability and decision-model analysis of a multimodel ensemble of climate change simulations. *J. Climate*, **14**, 3212–3226.

Santer, B., T. Wigley, G. Meehl, M. Wehner, C. Mears, M. Schabel, F. Wentz, C. Ammann, J. Arblaster, T. Bettge, W. Washington, K. Taylor, J. Boyle, W. Bruggemann, and C. Doutriaux, 2003: Influence of satellite data uncertainties on the detection of externally forced climate change. *Science*, **300 (5623)**, 1280–1284.

Stott, P., S. Tett, G. Jones, M. Allen, W. Ingram, and J. Mitchell, 2001: Attribution of twentieth century temperature change to natural and anthropogenic causes. *Climate Dyn.*, **17 (1)**, 1–21.

Tett, S., G. Jones, P. Stott, D. Hill, J. Mitchell, M. Allen, W. Ingram, T. Johns, C. Johnson, A. Jones, D. Roberts, D. Sexton, and M. Woodage, 2002: Estimation of natural and anthropogenic contributions to twentieth century temperature change. *J. Geophys. Res.*, **107**, D164306, doi:10.1029/2000JD000028.

Tett, S., P. Stott, M. Allen, W. Ingram, and J. Mitchell, 1999: Causes of twentieth-century temperature change near the earth's surface. *Nature*, **399 (6736)**, 569–572.

Thorne, P., P. Jones, S. Tett, M. Allen, D. Parker, P. Stott, G. Jones, T. Osborn, and T. Davies, 2003: Probable causes of late twentieth century tropospheric temperature trends. *Climate Dyn.*, **21 (7-8)**, 573–591.

Yoshioka, M., N. Mahowald, A. Conley, W. Collins, D. Fillmore, C. Zender, and D. Coleman, 2007: Impact of Desert Dust Radiative Forcing on Sahel Precipitation : Relative Importance of Dust Compared to Sea Surface Temperature Variations, Vegetation Changes, and Greenhouse Gas Warming. *J. Climate*, **20 (8)**, 1445–1467.