Quantifying Uncertainty for Climate Change and Long-Range Forecasting Scenarios with Model Errors. Part I: Gaussian Models

BORIS GERSHGORIN AND ANDREW J. MAJDA
Courant Institute of Mathematical Sciences, New York University, New York, New York

(Manuscript received 17 August 2011, in final form 24 December 2011)

ABSTRACT

Information theory provides a concise systematic framework for measuring climate consistency and sensitivity for imperfect models. A suite of increasingly complex physically relevant linear Gaussian models with time periodic features mimicking the seasonal cycle is utilized to elucidate central issues that arise in contemporary climate science. These include the role of model error, the memory of initial conditions, and effects of coarse graining in producing short-, medium-, and long-range forecasts. In particular, this study demonstrates how relative entropy can be used to improve climate consistency of an overdamped imperfect model by inflating stochastic forcing. Moreover, the authors show that, in the considered models, by improving climate consistency, this simultaneously increases the predictive skill of an imperfect model in response to external perturbation, a property of crucial importance in the context of climate change. The three models range in complexity from a scalar time periodic model mimicking seasonal fluctuations in a mean jet to a spatially extended system of turbulent Rossby waves to, finally, the behavior of a turbulent tracer with a mean gradient with the background turbulent field velocity generated by the first two models. This last model mimics the global and regional behavior of turbulent passive tracers under various climate change scenarios. This detailed study provides important guidelines for extending these strategies to more complicated and non-Gaussian physical systems.

1. Introduction

The climate is an extremely complex coupled dynamical system that evolves over a whole range of spatial and temporal scales, from millimeters to thousands of kilometers and from minutes to centuries (Neelin et al. 2006; Emanuel et al. 2005). The main focus of climate science is predicting the coarse-grained planetary-scale long time behavior of a climate system under various scenarios involving variations of external or internal parameters such as carbon dioxide concentration. The development of models that accurately recover crucial features of the earth’s climate system and that are able to forecast its response to the changes in external or internal parameters is one of the ultimate goals of contemporary climate science. An important feature of all the current computer Atmosphere Ocean Science (AOS) models (Neelin et al. 2006; Emanuel et al. 2005; Majda and Gershgorin 2010) is that they have significant model errors compared with the true signal in nature through either lack of understanding of the underlying physical processes or the limitations of computing power with the necessary parameterization of subgrid processes. Examples of important physical phenomena that are not always resolved by climate models include clouds, mesoscale and submesoscale oceanic eddies, and sea ice cover. Recently, the authors proposed (Majda and Gershgorin 2010, 2011a,b) a conceptual framework intermediate between detailed dynamical physical modeling and purely statistical analysis based on empirical information theory to address model fidelity and sensitivity of imperfect models. One advantage of the information-theoretic approach is that it is based on a skill measure, relative entropy, which is unbiased and invariant under the general change of variables unlike other metrics that are mainly based on RMS errors and that strongly depend on a particular choice of variables. In climate change science, information theory has been utilized to systematically improve model fidelity and sensitivity (Majda and Gershgorin 2010, 2011a), to quantify the role of coarse-grained initial states in long-range...
We address the role of seasonality in making ensemble predictions. Using information-theoretic metric, we find that the initial conditions sampled in the winter seasons, when the zonal jet is stronger, have longer memory than the initial conditions sampled in the summer season when the zonal jet is weaker. We monitor the distribution of uncertainty between the signal and dispersion components in the model error scenario before and after the optimization is performed. The second model is a spatially extended system of linear Rossby waves governed by a linear stochastic partial differential equation (PDE). One of the characteristic physical features of this model is the existence of a turbulent energy spectrum that allows us to study the role of coarse graining in quantifying uncertainty for various spatial scales. We show that for any coarse graining, if the optimal value of the stochastic forcing parameter for climate consistency is utilized, we greatly improve the predictive skill of the model. The third model is a special case of an advection–diffusion equation for passive tracers with a deterministic cross sweep, turbulent stochastic shear velocity (given by the second model), and strong mean gradient for the tracer. Using exact solutions, we demonstrate the nontrivial dependence of the mean and variance of the tracer on the temporal structure of the zonal jet: the tracer turns out to have stronger mean and variance when the zonal jet is weaker (in the summer) and weaker statistics otherwise. This interesting behavior translates to the uncertainty of the imperfect model with increased dissipation and eddy-diffusivity approximation so that the uncertainty rises in the summer when the tracer has weaker statistics and falls in the winter when the statistics take larger values. We also study how the optimal stochastic forcing depends on the coarse graining with a particular emphasis on the large scales, the main focus of climate science. These models have already been utilized by the authors and their collaborators in a more general nonlinear form in the context of real-time filtering of noisy signals (Majda et al. 2010b; Gershgorin and Majda 2011) and uncertainty quantification in climate change science Majda and Gershgorin (2010, 2011a,b).

The paper is organized as follows. In section 2, we discuss the general principles of information theory and, in particular, how it can be used to improve both model fidelity and model sensitivity simultaneously. Next, we apply the developed theoretical results to mathematically tractable test models that reproduce some statistical features of the earth’s climate. In section 3, we apply the general principles of empirical information theory to the scalar Gaussian SDE. Here, we discuss the role of seasonal cycle in quantifying uncertainty under various model error scenarios. In section 4, we present the second test model, which is a spatially extended system of forecasting (Giannakis and Majda 2012a,b), and to make a quantitative analytical (Majda and Gershgorin 2011b) and empirical link between model fidelity and forecasting skill (DelSole 2005; DelSole and Shukla 2010). This recent use of relative entropy as an optimization principle to improve models builds on earlier contributions where various aspects of the relative entropy as a skill measure for predictions have been developed and applied both in perfect model (Kleeman 2002; Majda et al. 2002b; Kleeman et al. 2002; Abramov and Majda 2004; Abramov et al. 2005; Kleeman 2011) and imperfect model scenarios (Roulston and Smith 2002; DelSole 2004, 2005). Throughout this paper, the use of the relative entropy to measure discrepancies in probability measure will be called the information metric (Kullback and Leibler 1951; Majda et al. 2005; Majda and Wang 2006). For the important special case of Gaussian distribution, the relative entropy consists of the signal and dispersion parts. It is important to point out that the optimization principles based on the information metric utilized here do not necessarily minimize the error in either the mean or the covariance individually but rather minimize the “distance” between the perfect and imperfect probability distributions as whole, which is precisely given by the relative entropy. The use of relative entropy here to improve imperfect models in a dynamic climate change context builds on earlier use of such concepts by statisticians to improve imperfect models (Akaike 1974; Burnham and Anderson 2002).

In this paper, we demonstrate how the information-theoretic approach can be used to systematically improve imperfect models for climate consistency and simultaneously increase the model’s predictive skill. We utilize a series of instructive, progressively more complex test models that mimic certain features of the earth’s climate system, to demonstrate this approach. All three models considered here are chosen to be linear and Gaussian, which makes them easily tractable analytically without any Monte Carlo estimates of statistics, yet they are rich enough to possess such important physical properties as seasonal cycle in both mean and covariance, a turbulent energy spectrum, and eddy diffusivity. The first model is a time-periodic linear Gaussian stochastic differential equation (SDE) (Gardiner 2010), whose solution is an Ornstein–Uhlenbeck process that mimics some features of the seasonal variation of the zonal jet. With this model, we apply the general principle of model improvement through the minimization of the relative entropy of the imperfect model compared to the perfect model. The optimal model tuned for climate consistency via this principle turns out also to have significantly improved predictive skill in the climate change scenario. Moreover, we address the role of seasonality in making ensemble predictions. Using information-theoretic metric, we find
linear Rossby waves governed by a linear Gaussian PDE. Here, we focus on the role of coarse graining in making accurate statistical description of a climate state of the system and in improving forecasting skill of the imperfect models. In section 5, we develop the model for tracers in a turbulent velocity field. We consider the practically relevant type of model error due to over dissipation of the turbulent velocity field and an eddy-diffusivity approximation for the tracer. Then we use the developed information-theoretic approach for simultaneously improving climate consistency and predictive skill of the imperfect model and specifically focus on the role of coarse graining. Conclusions and future work are discussed in section 6.

2. Empirical information theory

With a subset of variables \( u \in \mathbb{R}^N \) and a family of measurement functionals, \( \{ E_j(u) \} \), \( 1 \leq j \leq L \), empirical information theory (Jaynes 1957; Majda and Wang 2006) builds the least biased probability measure \( \pi_L(u) \) consistent with the \( L \) measurements of the current climate. There is a unique functional on probability densities (Jaynes 1957; Majda and Wang 2006), called entropy

\[
S = - \int \pi \ln \pi, \tag{1}
\]

and \( \pi_L \) is a unique probability density consistent with the measured information \( E_L \) that maximizes \( S \). For example, if \( u \) is a scalar variable representing the global mean temperature, then natural choices for \( E_L \) as measurements are the mean and variance of the temperature. Standard calculations using the method of Lagrange multipliers (see Majda and Wang 2006 for examples) show that \( \pi_L(u) \) is given by

\[
\pi_L(u) = e^{-\alpha_0 - \alpha_L E_L(u)}, \tag{2}
\]

where the Lagrange multipliers \( \alpha_L = (\alpha_1, \ldots, \alpha_L) \) are chosen so that the probability density is consistent with the \( L \) measurements

\[
E_L = \int E_L(u) e^{-\alpha_0 - \alpha_L E_L(u)}, \tag{3}
\]

while \( \alpha_0 \) is determined by normalization needed to guarantee that \( \pi_L(u) \) is a probability density. A natural choice for \( E(u) \) are all the moments up to some order \( p \), that is,

\[
E_j(u) = \int (u)^j \pi(u), \quad |j| \leq p. \tag{4}
\]

For example, measurements of the mean and second moments of a subset of variables of the perfect system necessarily lead to a Gaussian approximation (Majda and Wang 2006; Majda et al. 2002a) to the perfect system \( \pi_L(u) = \pi_{\text{G}}(u) \). This observation motivates us to consider Gaussian models as the first natural step in the hierarchy of models that mimic the behavior of climate systems. Hence, the three examples that we present in sections 3, 4, and 5 are increasingly complex linear Gaussian models, with a dynamic behavior ranging from a simple scalar motion to spatially extended turbulent mixing of passive tracers.

The natural way (Majda and Wang 2006; Kullback and Leibler 1951) to measure the lack of information in one probability density \( q(u) \) compared with the other \( p(u) \) is through the relative entropy \( \mathcal{P}(p,q) \) given by

\[
\mathcal{P}(p,q) = \int p \ln \frac{p}{q}, \tag{5}
\]

The relative entropy \( \mathcal{P}(p,q) \) is always positive unless \( p = q \) and is invariant under any invertible change of variables (Majda and Wang 2006; Majda et al. 2002a; Kullback and Leibler 1951). Because of these attractive properties, the relative entropy characterizes the “distance” between two probability densities despite the lack of symmetry in the arguments \( p \) and \( q \). Thus, relative entropy provides a natural framework for assessing model error in AOS applications (Kleeman 2002; Majda et al. 2002a; DelSole 2004, 2005; Abramov et al. 2005; Majda and Gershgorin 2010; Branstator and Teng 2010; Teng and Branstator 2011; Giannakis and Majda 2012a,b). In the context of model error, \( \mathcal{P}(\pi, \pi^M) \) characterizes the lack of information in the imperfect distribution \( \pi^M \) compared to the true distribution \( \pi \). Alternatively, the relative entropy \( \mathcal{P}(p,q) \) could be interpreted as a gain of information in the distribution \( p \) beyond the information given by the distribution \( q \). This interpretation is useful if we assume that \( q = \pi_{\text{attr}} \) is the climate (attractor) distribution of a system, while \( p = \pi_{\text{obs}} \) is a distribution of a system with a given initial condition at time \( t_0 \). Then, \( \mathcal{P}(\pi_{\text{obs}}, \pi_{\text{attr}}) \) quantifies the role of initial conditions in the forecast of a future state of a system (Kleeman 2002; Giannakis and Majda 2012a,b). Finally, we utilize yet another interpretation of relative entropy, when \( p = \pi_{\delta} \) characterizes the distribution of a perturbed system with perturbation of a small size \( \delta \), while \( q = \pi_{\text{attr}} \) again characterizes the climate distribution. In this context, the relative entropy \( \mathcal{P}(\pi_{\delta}, \pi_{\text{attr}}) \) measures the sensitivity of a system to external perturbations (Majda and Gershgorin 2011a,b).

Next, we focus on the role of relative entropy in improving imperfect models. Consider an imperfect model
with its associated probability density $\pi^M(u)$. Then, as we mentioned above, $P(\pi, \pi^M)$ exactly characterizes the intrinsic error in the climate statistics. Consider a class of imperfect models $\mathcal{M}$; the best imperfect model for the coarse-grained variable $u$ is the $M_u \in \mathcal{M}$ so that the perfect model has the smallest additional information beyond the imperfect model distribution (Burnham and Anderson 2002; Majda and Gershgorin 2010) $\pi^M(u)$, that is,  

$$P(\pi, \pi^M) = \min_{M \in \mathcal{M}} P(\pi, \pi^M). \quad (6)$$

Also, actual improvements in a given climate model with distribution $\pi^M(u)$ either through higher resolution or improved parameterization resulting in a new distribution $\pi^M_{\text{post}}(u)$ should result in improved information for the actual climate

$$P(\pi, \pi^M_{\text{post}}) \leq P(\pi, \pi^M). \quad (7)$$

otherwise, objectively, the model has not been improved compared with the original climate model (Majda and Gershgorin 2010). The following general principle (see Majda and Gershgorin 2010; Majda et al. 2005 for details) facilitates the practical calculation of Eq. (6):

$$P(\pi, \pi^M_{L'}) = P(\pi, \pi^M_L) + P(\pi^M_{L}, \pi^M_{L'}). \quad (8)$$

where $L' \leq L$, and $\pi^M_L$ is given by Eq. (2). In particular, it is possible to improve the mean of an imperfect model through improved parameterization but objectively not improve the imperfect model as required in (7) because of a poorer variance. Note that $P(\pi, \pi^M_L)$ exactly gives the lack of information in a “coarse-grained” probability density associated with fewer constraints, $\pi^M_L$. Now the optimization principles in Eqs. (6) and (7) can be computed explicitly by replacing the unknown $\pi$ by the hypothetically measured and known $\pi_L$, so that the optimal model $M_u$ from 6 is now calculated by

$$P(\pi^M_{L}, \pi^M_{L'}) = \min_{M \in \mathcal{M}} P(\pi^M_{L}, \pi^M_{L'}). \quad (9)$$

The most practical setup used here for applying the framework of empirical information theory developed above arises when both the perfect system measurements and the model measurements involve only the mean and covariance of the variables $u$ so that $\pi^M_L$ is Gaussian with climate mean $\mu$ and covariance $R$, whereas $\pi^M$ is Gaussian with model mean $\mu^M_L$ and covariance $R^M$. In this case, $P(\pi^M_L, \pi^M)$ has explicit formula (Majda and Wang 2006; Kleeman 2002)

\[
P(\pi^M_L, \pi^M) = \left[ \frac{1}{2} (\mu - \mu^M)^T (R^M)^{-1} (\mu - \mu^M) \right]
+ \left[ -\frac{1}{2} \text{ln det}(RR^M) + \frac{1}{2} \text{tr}(RR^M) - N \right].
\]

(10)

The same Eq. (10) applies with $\pi^M$ replaced by $\pi^M_k$ in (9). Note that the first term in brackets in Eq. (10) is the signal, reflecting the model error in the mean but weighted by the inverse of the model variance $R^M_k$, whereas the second term in brackets, the dispersion, involves only the model error covariance ratio $RR^M$. There is extensive use of this signal-dispersion decomposition and its generalizations in quantifying uncertainty in ensemble predictions with perfect models (Abramov and Majda 2004; Kleeman 2002; Kleeman et al. 2002; Abramov et al. 2005; Kleeman 2011). The role of these different contributions to the model error in the climate calibration metric is illustrated below in prototype Gaussian test models. The intrinsic metric in Eq. (10) is invariant under any (linear) change of variables, which maps Gaussian distributions to Gaussians; moreover, the signal and dispersion terms are individually invariant under these transformations—this property is very important for unbiased AOS model calibration. In this paper, we utilize the metric in Eq. (10) in application to linear Gaussian models, that is, when the true probability distribution is fully determined by its first two moments $\pi = \pi_2 = \pi_G$. In particular, we are interested in the distribution of the total uncertainty between the signal and dispersion parts in various scenarios such as the climate consistency study when imperfect models are used, the role of initial conditions in the forecast of a future state of a system, and the sensitivity of a system to external perturbation with and without model error.

One important special case involves a family of variables $u$ where simultaneously both $R$ and $R^M_k$ are block diagonal. For simplicity assume that all the blocks in these two matrices are of the same size $J \times J$. For example, this situation can arise when the system is represented by a number of Fourier modes such that the components of the system that correspond to the same Fourier mode are correlated while different Fourier modes are independent, which will be the case for the Gaussian tracer model considered in section 5. Then the whole state space vector $u \in \mathbb{R}^N$ consists of $K \leq N$ components, $u_k \in \mathbb{R}^J$ for $1 \leq k \leq K$ and $N = KJ$ (note that we assume here that the zeroth Fourier mode vanishes and thus we exclude it from consideration) with $R_k$ and $R_{M,k}$ denoting covariance matrices of $u_k$ and $u_{M,k}$, respectively. Using a simple fact from linear algebra that the inverse of a block-diagonal matrix is also a
block-diagonal matrix where each block is an inverse of the original blocks, we simplify Eq. (10)

$$P(\pi_L, \pi^M) = \sum_{k=1}^{K} \left[ \frac{1}{2} (u_k - u_{M,k})^2 (R_{M,k})^{-1} (u_k - u_{M,k}) \right]$$

$$+ \sum_{k=1}^{K} \left\{ -\frac{1}{2} \ln \det (R_k R^{-1}_{M,k}) \right\} + \frac{1}{2} [\text{tr}(R_k R^{-1}_{M,k}) - J].$$

(11)

Note that the metric in (11) sums over the whole spectrum of wavenumbers from 1 to K. However, it is often the case that the system is modeled by a coarse-grained variable that only involves first K’ Fourier modes with K’ ≤ K. For the coarse-grained variable, the uncertainty is given by the same Eq. (11) with K substituted by K’. Note that each summand in (11) is positive and thus by increasing resolution, one increases the uncertainty. A natural question arises: how does total uncertainty as well as the individual contributions in the signal and dispersion parts depend on coarse graining? This is an important issue regarding the skill of climate models for global and regional long range predictions. We will address this type of question in sections 4 and 5 with the examples of spatially extended systems.

Below, we use these formulas both to measure and improve climate fidelity of imperfect models and, more importantly, for assessing improvements in forecast skill of imperfect models. We assume that either the perfect system or the model system or both are perturbed in a fashion so that, π M,u (u) the unknown perfect distribution, π M,u (u) the measured distribution, and π M,u (u) the model distribution, all vary smoothly with parameter δ. The interest is in long-range forecasting skill and a necessary condition is climate consistency (Majda and Gershgorin 2011a,b). Climate consistency arises when Pr(π M,u) is minimized within a class of models. On the other hand simple yet instructive examples show (Majda and Gershgorin 2011a,b) that consistency is not sufficient for sensitivity—an approximate model may have perfect climate consistency Pr(π M,u) = 0, while there may be an intrinsic barrier to skill in predicting model sensitivity. The only way to overcome such a barrier is by extending the class of models by introducing more degrees of freedom.

3. A time-periodic Gaussian scalar model

In many physical situations, complex phenomena are modeled by linear Gaussian equations, when the system mean and variance are available. In fact, as we discussed in section 2, the least-biased distribution with given mean and variance is Gaussian. Another important property of many nature systems is time periodicity. The examples include the diurnal cycle and the seasonal cycle in the climate models. The role of time periodicity is extremely interesting and practically important in evaluating the skill of various predictive strategies. In particular, one may wonder if there is more uncertainty in the forecast of the winter climate or of the summer climate. Therefore, our first test model is a linear Gaussian system with time-periodic climate statistics mimicking seasonal variation of a mean jet.

a. Model description

As the simplest example of a time-periodic model, we consider a SDE (Gardiner 2010)

$$dU = -\gamma(t) Ud t + f(t)dt + \sigma(t)dW,$$  

(12)

where γ(t) is damping, f(t) is deterministic forcing, and σ(t) is the strength of the white noise forcing. Here, U(t) can be interpreted as a mean east–west zonal jet, although other interpretations are certainly possible since the model is very general. Given initial Gaussian condition U(t0), the solution U(t) stays Gaussian for all times t > t0. We assume that there is a common period T o for all three functions

$$\gamma(t) = \gamma(t + T_0),$$

$$f(t) = f(t + T_0), \text{ and}$$

$$\sigma(t) = \sigma(t + T_0).$$

(13)

If the dissipation coefficient integrated over the period is positive, that is,

$$\int_{T_0}^{T_0} \gamma(t) dt > 0,$$

(14)

then the system in (12) is stable and there is a time periodic Gaussian statistical attractor that can be described by its mean and variance. We use the following intuition about the zonal jet in the atmosphere in finding appropriate physically relevant regimes:

- the jet is stronger (has larger mean) in the winter than in the summer;
- the jet has more variability in the winter when it is stronger than in the summer when it is weaker;
- the jet has a predominantly eastward direction (positive mean);
- the variance is chosen such that the jet stays eastward (positive) within at least one standard deviation from its mean; and
- the jet decorrelates faster in the summer than in the winter, and the typical decorrelation time is on the order of 1–2 months.
The system in (12) has the following Gaussian solution for any Gaussian initial value \( U_0 \) at \( t = t_0 \)

\[
U(t) = G_U(t_0, t)U_0 + \int_{t_0}^{t} G_U(s, t)f(s) \, ds + \int_{t_0}^{t} G_U(s, t)\sigma(s) \, dW(s),
\]

where the Green’s function has the form

\[
G_U(s, t) = e^{-\int_{s}^{t} \gamma(s') \, ds'}.
\]

From (15), we find the mean and variance of \( U(t) \)

\[
\langle U(t) \rangle = G_U(t_0, t)\langle U_0 \rangle + \int_{t_0}^{t} G_U(s, t)f(s) \, ds,
\]

and

\[
\text{Var}[U(t)] = G_U(t_0, t)^2\text{Var}(U_0) + \int_{t_0}^{t} G_U(s, t)^2\sigma^2(s) \, ds.
\]

Equations (17) and (18) have time-periodic solutions that describe the attractor of the system (12) when the condition in (14) is satisfied. The periodicity conditions given by

\[
\langle U(t_0 + T_0) \rangle = \langle U_0 \rangle, \quad \text{and}
\]

\[
\text{Var}[U(t_0 + T_0)] = \text{Var}(U_0) \quad \text{leads to}
\]

\[
\langle U_0 \rangle\text{attr} = \frac{\int_{0}^{T_0} G_U(s, T_0)f(s) \, ds}{1 - G_U(0, T_0)}, \quad \text{and}
\]

\[
\text{Var}(U_0)\text{attr} = \frac{\int_{0}^{T_0} \sigma^2(s)G_U(s, T_0)^2ds}{1 - G_U(0, T_0)},
\]

where the subscript means that the value is given on the statistically periodic attractor. One can also easily show that the correlation function on the attractor is given by

\[
\text{Corr}[U(t), \tau]_{\text{attr}} = G_U(t, t + \tau)
\]

\[
= \frac{1}{\text{Var}[U(t)]}[U(t + \tau) - \langle U(t + \tau) \rangle + \langle U(t) \rangle]\cdot \text{Var}[U(t) - \langle U(t) \rangle].
\]

With Gaussian initial data, the solutions of the scalar equation in (12) are Gaussian. Since Gaussian distributions are determined by their mean and variance, the

<table>
<thead>
<tr>
<th>Regime</th>
<th>( \gamma )</th>
<th>( A_{\gamma} )</th>
<th>( \phi_{\gamma} )</th>
<th>( \sigma )</th>
<th>( A_{\sigma} )</th>
<th>( \phi_{\sigma} )</th>
<th>( A_0 )</th>
<th>( \phi_0 )</th>
<th>( A_1 )</th>
<th>( \phi_1 )</th>
<th>( A_2 )</th>
<th>( \phi_2 )</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>0.8</td>
<td>-0.4</td>
<td>-0.5</td>
<td>3.0</td>
<td>0.6</td>
<td>1.4</td>
<td>1.5</td>
<td>1.2</td>
<td>0.0</td>
<td>1.0</td>
<td>0.7</td>
<td>0.0</td>
</tr>
<tr>
<td>B</td>
<td>0.2</td>
<td>-0.1</td>
<td>-0.5</td>
<td>1.0</td>
<td>0.2</td>
<td>1.3</td>
<td>0.3</td>
<td>1.0</td>
<td>0.7</td>
<td>0.7</td>
<td>0.7</td>
<td>0.0</td>
</tr>
</tbody>
</table>

The time-periodic formulas in (20), (21), and (22) define the statistically periodic attractor uniquely (see Majda and Wang 2010 for the theory of statistically periodic attractors). Note that in the time-periodic case, the correlation function has two arguments, the usual time lag \( \tau \) and the time \( t \), when the correlation is actually computed. Physically this means that the system in (12) can have different memory in different seasons even in this simplest Gaussian setup. The typical correlation time is computed as an integral of the correlation function over all the lags \( \tau \)

\[
T_{\text{corr}}(t) = \int_{0}^{\infty} G_U(t, t + \tau) \, d\tau.
\]

b. Parameter regimes

Next, we discuss the parameters in the model (12) keeping in mind the physical intuition discussed above. We choose a special form of the time-dependent parameters

\[
f(t) = A_0 + A_1 \cos(\omega t + \phi_1) + A_2 \cos(2\omega t + \phi_2),
\]

\[
\gamma(t) = \gamma + A_{\gamma} \cos(\omega t + \phi_{\gamma}), \quad \text{and}
\]

\[
\sigma(t) = \sigma + A_{\sigma} \cos(\omega t + \phi_{\sigma}),
\]

where \( \omega = \pi/6 \), which ensures the period \( T_0 = 12 \) that can be interpreted as a year consisting of 12 months. We note that each of these three functions has a constant part and a periodic part. Moreover, we allow the forcing \( f(t) \) to have two frequencies, \( 2\omega \) and \( 2\omega \), so that the mean jet \( \langle U \rangle \) is not a simple sine-like oscillation but has more temporal structure. Here, we distinguish between two physically relevant regimes, one that corresponds to the atmosphere-like flows with shorter correlation of 1–2 months (regime A) and another that can be attributed to ocean-like flows with longer correlation of 4–6 months (regime B). In Table 1, we give all parameters for both of these regimes. In Fig. 1, we show perfect model mean, variance, and correlation time of \( U(t) \) in the periodic, statistically steady state for both parameter regimes. We note that the parameters are chosen so that the steady state statistics satisfy the physical intuition outlined below Eq. (14). Now the question is as follows: how well can these climate statistics be approximated using an

| TABLE 1. Parameters for the scalar model describing two statistical regimes. |
imperfect model and how can such a model be improved from the information-theoretic perspective?

c. Climate fidelity and model error

After providing a formal description of this simple test model, we turn to the issue of model error in the climate forecast because of the inadequate description of the true system dynamics. Suppose that the true state of the system $U(t)$ is not known and instead an approximation $U^M(t)$ is used. As always, the superscript “$M$” corresponds to the variables and parameters of the imperfect model with model error. The approximation is governed by an imperfect model, which in this case has the same functional form (12) as the original perfect model but with one or more parameters different from those of the perfect model

$$dU^M = -\gamma^M(t)U^M dt + f^M(t)dt + \sigma^M(t)dW. \quad (25)$$

For example, suppose that the dissipation $\gamma^M(t) \neq \gamma(t)$ [current computer models tend to overestimate dissipation so that $\gamma^M(t) > \gamma(t)$ (Palmer 2001)] while the rest of the parameters take the same values as in the perfect model, and we would like to know the implications of this model error on the climate statistics. We are also interested in how much we can improve the imperfect model by increasing noise (Palmer 2001; Majda and Gershgorin 2011a). As motivated in section 2, one convenient way to assess model error is through empirical information theory. For Gaussian random fields such as $U(t)$ and $U^M(t)$, the lack of information in the imperfect model $\mathcal{P}(\pi, \pi^M)$ is given by (10), which consists of the signal and dispersion parts.

First, we consider the situation, when the model error is in the constant part of the dissipation $\gamma^M = \gamma \pm r\gamma$, where $r\gamma$ characterizes the strength of the model error. In Fig. 1, along with the perfect model statistics discussed above, we show the statistics of the imperfect model (25) for the model error in $\gamma$. In particular, in regime A, we have $\gamma^A = 1.4\gamma = 1.12$, and in regime B, we have $\gamma^B = 1.2\gamma = 0.24$. Here, the imperfect dissipation is uniformly stronger than its perfect value and hence both imperfect mean and variance decrease. In Fig. 2, we show the lack of information in the imperfect model as a function of time of year. We note that the total lack of information is dominated by the signal part in both regimes with the ratio of the signal part to the total lack of information staying at the level of 0.9 and only dropping to 0.8 in the fall season. Moreover, both signal and dispersion parts are larger in the winter season, when the mean jet and its
variance are stronger, than in the summer season when the mean statistics are weaker. Now, we follow the ideas from Majda and Gershgorin (2011a) and change noise in the Eq. (25) to compensate partially for the effects of the model error in the dissipation. In particular, we set $\sigma^M = r_\sigma \sigma$, systematically vary the parameter $r_\sigma$, and monitor the corresponding lack of information $P(\pi_{\mathrm{attr}}, \pi_{\mathrm{attr}}^M)$.

In Fig. 3, we demonstrate the corresponding lack of information as a function of $r_\sigma$ for both regimes A and B and find that there is a particular value $\sigma_{\sigma}^M$ for which the lack of information in minimized, that is,

$$P(\pi_{\mathrm{attr}}, \pi_{\mathrm{attr}}^M) = \min_{\sigma^M} P(\pi_{\mathrm{attr}}, \pi_{\mathrm{attr}}^M). \quad (26)$$

FIG. 2. Solid line shows relative entropy of the climate for the imperfect model with increased dissipation $\sigma^M > \sigma$: (top) signal part, (middle) dispersion part, (bottom) total, for regimes (left) A and (right) B given in Table 1. Dashed line shows relative entropy of the optimized imperfect model with inflated noise $\sigma^M > \sigma$. Note significant improvement (decrease of the total relative entropy) in the optimal model compared to the original imperfect model.

FIG. 3. Solid line shows the relative entropy of the climate averaged over one period as a function of the noise-tuning parameter $r_\sigma$, where $\sigma^M = r_\sigma \sigma$ for regimes (left) A and (right) B. Circles show the optimal value for $r_\sigma$ that minimizes the relative entropy. Signal and dispersion components are shown with the dashed and dash-dotted lines, respectively.
Here the line over expressions means the time average over the period \([0, T_0]\). Note that the criterion in (26) is an average over the entire year so climate fidelity as imposed here with (26) and the Gaussian Eq. (15) is a much weaker tuning requirement for a Gaussian model than separately tuning the mean and covariance as the time of year varies; in contrast, (26) requires varying a single parameter which is the noise here; nevertheless, we show below that improving (26) leads to much improved forecast skill. According to the general optimization principle from Eq. (6), \(\sigma_M^*\) is the optimal noise for the model error in dissipation. For regime A, the optimal noise strength becomes \(\sigma_M^* = 1.6\sigma = 4.8\), and for regime B we find \(\sigma_M^* = 1.35\sigma = 1.35\). Note that since the time average was taken over the whole period, these optimal values correspond to the whole year. On the other hand, we could have taken averages over distinct seasons instead of the whole year and this procedure would yield the optimal noise for a particular season. Now, we go back to Fig. 1, where we show the mean, variance, and correlation time of the optimal model as well as the original and imperfect statistics as discussed above. We note that the mean and correlation time are unaffected by the optimization procedure, while the variance is increased. However, both signal and dispersion parts of the lack of information change because of the change in the variance of the imperfect model. In Fig. 2, we show how the lack of information in the optimally tuned model compares with the lack of information in the imperfect model without extra noise. We note a significant improvement of the model with optimal noise from the information-theoretic perspective. Moreover, we note that the total uncertainty decreased for almost all the times throughout the period although we only optimized for the year-averaged lack of information. On the other hand, the dispersion part has increased in some parts of the period in regime A and through the whole period in regime B. However, since the dispersion part only takes a small fraction of the total lack of information, the decrease in the signal part compensates this minor increase in the dispersion part and makes the tuned model optimal from the information-theoretic point of view. Moreover, below we will see (Fig. 7) that this model, optimized here for climate consistency, will also have high skill for sensitivity to external perturbations, and the dispersion part of the uncertainty with external forcing in the optimized model actually decreases compared to the original imperfect model.

**d. The role of initial conditions**

We study the role of initial conditions in making forecasts of short, medium, and long range. To mimic realistic situations, we construct families of statistical initial conditions with initial mean sampled from the climate distribution. We assume that the initial conditions are known with high certainty, that is, the initial variances (the “error bars”) are significantly smaller than the climate variances. Then, for each member of the ensemble of initial conditions, we use the perfect model to make the prediction of the future state. In Fig. 4, we show an example of such ensemble together with the climate statistics. Note that here we only show 3 members from the ensemble of 10 trajectories. In total, we have 72 initial times taken every \(\Delta t = 12/72 \approx 1.67\) time units and 10 ensemble members at each of these times. The initial variance for each the ensemble member is equal to one tenth of the climate variance at that particular time, that is, \(\text{Var}_{\pi_0}(U) = 0.1\text{Var}_{\text{clim}}[U(t_0)]\). This is a manifestation of the fact that, initially, we assume knowledge of the state of the system with much better precision that it would be given by climate statistics. Next, we computed the forecasts for all ensemble members with a lead time up to 6 time units for regime A and 15 time units for regime B. The difference in the maximal lead times is due to the difference in the correlation time of the system in the respective regime—regime B has a significantly longer correlation time and hence longer memory of initial conditions than regime A. In Fig. 5, we show the gain of information in the forecast with initial conditions beyond the climate statistics, which is characterized as \(P[\pi_{\pi_0}(t), \pi_{\text{clim}}(t)]\), where \(\pi_{\pi_0}(t)\) is a distribution of the forecast at time \(t\) using the initial condition at time \(t_0\), and \(\pi_{\text{clim}}(t)\) is the climate distribution at time \(t\). Usually one would be interested in the average forecast skill with the initial conditions from a particular season, rather than in the forecast made with an initial condition given on a particular date. Therefore, we average the lack of information over various sets of initial conditions, the whole period (year), time units 1, 2, and 12 (winter season), and time units 6, 7, and 8 (summer season). Moreover, we average this quantity over the ensemble of all trajectories with a given initial condition, that is, in Fig. 5, we show \(\langle P[\pi_{\pi_0}(t), \pi_{\text{clim}}(t)]\rangle\), where the overline means time average over initial conditions, and the angular brackets mean ensemble average. We note that the gain of information in the forecast with initial conditions approaches zero as the lead time increases, which indicates the loss of memory of initial conditions with time compared to the climatology (Kleeman 2002). On the other hand, we note the difference in the speed of approaching zero in different seasons, for example, in winter, when the jet is stronger, the additional information decays slower (in both signal and dispersion parts) than in the summer. We note that the model has the longest memory in the winter and the shortest in the summer, while the year-averaged values are in between these two extremes. Also
note that the system in regime A indeed has shorter memory of initial conditions because of much stronger dissipation than in regime B where the dissipation is weaker (see Table 1 to compare parameters).

e. Sensitivity to external perturbations with uncertainty

Now, we study the sensitivity of the system in (12) to the perturbations of external forcing. The main points we would like to address are the following:

- How does the system respond to external perturbations?
- How well can this response be predicted using imperfect models?
- How is climate fidelity linked to the sensitivity to external perturbations?

In particular, we are interested in the mean and variance statistics of the following perturbed system:

$$\frac{dU(t)}{dt} = -\gamma(t)U(t) + f(t) + \delta f(t) + \sigma(t)\tilde{W},$$  

(27)

where $\delta f(t)$ is a perturbation of the external forcing $f(t)$, and $W$ is white noise. To imitate a climate change scenario, we assume that initially the system is in the statistically steady state without any external perturbations, that is, $\delta f(t) = 0$ for $t < t_0$ for some $t_0$. Then, the external forcing starts changing with $\delta f(t) \neq 0$ as $t \geq t_0$. We assume that the perturbation $\delta f(t)$ in (27) has both deterministic and stochastic components that represent the trend and the uncertainty in the forcing.

$$\delta f(t) = [\delta \hat{f}(t) + b(t)]\phi_{t_0}(t),$$  

(28)

where $\delta$ is small. The uncertainty in the forcing can arise from external effects like the solar cycle or anthropogenic effects. Here, $\delta \hat{f}(t)$ is either equal to $f(t)$ or to some part of it, that is, the constant part of $f(t)$, or its time-dependent part $b(t)$, is a stochastic process that models the uncertainty in $\delta f$, and $\phi_{t_0}(t)$ is a ramp function, which is equal to 0 for $t < t_0$, that grows linearly from 0 to 1 over some time $\tau_0$ that mimics the time of climate change, and then stays equal to 1 for $t > t_0 + \tau_0$.

$$\phi_{t_0}(t) = \begin{cases} 0, & t < t_0 \\ \frac{t - t_0}{\tau_0}, & t_0 \leq t \leq t_0 + \tau_0 \\ 1, & \text{otherwise}. \end{cases}$$  

(29)

Here, we model the uncertainty $b(t)$ by another independent Gaussian random process.
\[
\frac{db(t)}{dt} = -\gamma_b b(t) + \sigma_b W, \tag{30}
\]

where the parameters \(\gamma_b\) and \(\sigma_b\) characterize the correlation time and the typical size (standard deviation) of the uncertainty. The solution for (27) is given by

\[
U(t) = G_U(t_0, t)U_0 + \int_{t_0}^t G_U(s, t)[f(s) + \delta f(s)] ds \\
+ \int_{t_0}^t G_U(s, t)\sigma(s) dW(s), \tag{31}
\]

where \(G_U(s, t)\) is defined in (16). It is reasonable to assume that the stochastic process \(b(t)\) that characterizes the uncertainty in (28) is in a statistical steady state so that

\[
\langle b(t) \rangle = 0, \\
\text{Var}[b(t)] = \frac{\sigma_b^2}{2\gamma_b}, \quad \text{and} \\
\text{Cov}[b(r), b(s)] = \frac{\sigma_b^2 e^{-\gamma_b|s-r|}}{2\gamma_b}. \tag{32}
\]

We are interested in the mean and variance response to the external perturbation, which can be computed as the difference between (31) and (15)

\[
\delta\langle U(t) \rangle = \int_{t_0}^t \delta f(s)\phi_{t_0}(s)G_U(s, t) ds, \quad \text{and} \tag{33}
\]

\[
\delta\text{Var}[U(t)] = \int_{t_0}^t \int_{t_0}^t \phi_{t_0}(s)\phi_{t_0}(r)G_U(s, t)G_U(r, t)e^{-\gamma_b|r-s|} ds dr. \tag{34}
\]

It is important to note here that, because of the linear Gaussian structure, the mean response only depends on the deterministic part of the perturbation while the variance response only depends on the random part of the perturbation. The introduction of the uncertainty in the perturbation actually allows for the nontrivial variance response in a Gaussian model, which otherwise would be zero. Next, we study the form of the mean and variance response for the parameter cases discussed above in both perfect and imperfect model scenarios. The latter addresses the sensitivity issue in the model error context.
1) STATISTICAL RESPONSE FOR THE PERFECT MODEL

First, we consider the response of the perfect model when all the parameters in (12) are assumed to be known exactly. This is an ideal situation, of course, but it will serve as a benchmark for testing more realistic strategies when imperfect models are used. Here, we use regime A from Table 1. Note that now we have extra parameters that characterize the uncertainty \( b(t) \), which we need to specify. The correlation time of the uncertainty is given by \( 1/\gamma_b \) and, therefore, by choosing the appropriate correlation time, we obtain the value for \( \gamma_b \). Also, the error of the uncertainty is given by the standard deviation \( \sigma[b(t)] = \sqrt{\sigma_b^2/2\gamma_b} \). We control this error by setting \( \sigma[b(t)] \) to be equal to 50% of the typical size of the deterministic part of \( f(t) \), that is,

\[
\sqrt{\frac{\sigma_b^2}{2\gamma_b}} \sim 0.5\bar{f}(t),
\]

which yields

\[
\sigma_b \sim 0.5\bar{f}\sqrt{2\gamma_b}.
\]

In our experiment, we set \( \gamma_b \approx 0.0833 \), which corresponds to 12 time units (1 year) correlation time of the uncertainty, and \( \sigma_b \approx 0.1633 \), which satisfies (36), where \( \bar{f}(t) = A_0 \). We apply the perturbation \( \delta f(t) \) for 240 times units (20 years) at a rate of 1% growth per 12 time units (1 year). This period of changing forcing is followed by another 120 time units (10 years) of the new forcing when the new climate is established. We study the sensitivity of the model to external perturbation by monitoring the gain of information in the perturbed system compared to the unperturbed one. To quantify the sensitivity of the system to external perturbation \( \delta f(t) \), we use \( P(\pi_\delta(t), \pi_{\text{attr}}(t)) \), where \( \pi_\delta \) is the probability distribution of the perturbed system. Note that both distributions, \( \pi_\delta \) and \( \pi_{\text{attr}} \), are Gaussian, and the lack of information decomposes into the signal and dispersion parts according to (10).

In Fig. 6, we show the signal and dispersion parts of \( P(\pi_\delta, \pi_{\text{attr}}) \), the additional information in the forced response beyond the climatology, as functions of time of year as well as the fraction of the signal part in the total lack of information. Note that this forced statistical response has a nonperiodic transient phase where the
information-theoretic sensitivity is modulated by the seasonal cycle. We note that the signal part of the response is larger than the dispersion part and both have strong time dependence because of the seasonal cycle in the model. We note that the signal part dominates the total lack of information just after the perturbation is applied. However, as the perturbation starts increasing, the dispersion part also grows, and as the new climate is established, the dispersion part takes 10%–15% of the total lack of information. This experiment shows, that in the climate change scenario, the role of signal and dispersion parts varies depending on the size and duration of the perturbation. Now, after we have described the response of the perfect model to the external perturbations, we are interested in how well this response can be predicted by imperfect models.

2) Statistical Response and Model Error

Next, we turn to the issue of model error. As we saw above, the system in (27) responds to the perturbations by changing both mean and variance. Now, we are interested in how well we can predict this response using imperfect models with one or more parameters different from their values in the perfect model. We focus on the same model error scenario for regime A as discussed in section 3c. In particular, the imperfect model has incorrectly specified dissipation parameters \( \bar{\gamma}^M = 1.4 \bar{\gamma} \). As a result of this error, the climate statistics are not reproduced correctly as we demonstrated in Fig. 1. Moreover, the response of the system to external perturbations also has errors because of model error. In Fig. 7, we show the signal and dispersion parts of the uncertainty in predicting the response with imperfect model \( P(\pi_s, \pi_d^M) \) for the same type of perturbation \( \delta f(t) \) as discussed above in section 3e(1). We note that the model error leads to inaccurate prediction of the system response, which is characterized by positive values of \( P(\pi_s, \pi_d^M) \). We expect to have improved prediction skill in the imperfect model \( M^\delta \) that was optimized for climate consistency according to the general principle in (6) \( (26) \). Next, we illustrate this point using the scalar model for \( U(t) \). We use the imperfect model with noise that is optimized for climate fidelity as discussed above and that is equal to \( \pi_d^M = 1.6 \bar{\gamma} \). With this inflated noise, the imperfect model reproduces the climate of the system better than the original imperfect model in the information-theoretic sense. Now, in the case of changing climate there is also a dramatic improvement of the prediction of the system response as can be seen in Fig. 7. In Fig. 7, we show the “envelope” of the time-dependent values of the uncertainty, of its signal and dispersion components, and of the fraction of the signal part in the total uncertainty. Moreover, we show segments of the actual uncertainty as it oscillates because of the seasonal cycle in the model. The improvement in the prediction skill because of the optimization of the noise is mainly attributed to the decrease of the signal part of the total uncertainty given by \( P(\pi_s, \pi_d^M) \), where \( M^\delta \) corresponds to the optimal model as discussed above. However, we note that the dispersion part also shows significant decrease as the external perturbation starts acting. This behavior is even more surprising if we recall that in the statistical steady state, while the optimized model decreased the total uncertainty, it did not decrease the dispersion part according to Fig. 2.

4. Linear stochastic PDE: Coarse graining, model error, and sensitivity

Many complex systems in nature have significant behavior over many spatiotemporal scales with subtle sensitivity in the response to changes in forcing (Majda et al. 2010a; Gritsun and Branstator 2007; Gritsun et al. 2008); the ability of an imperfect model to reproduce this multiscale behavior is a central issue as well as the capability of the imperfect model to mimic the sensitivity of the natural system. The simplest model to study...
these issues involving many degrees of freedom in a spatially extended system is through deterministically and stochastically forced constant coefficient PDEs (Gershgorin and Majda 2011; Majda and Gershgorin 2012; Majda et al. 2010b). In this setting, both the perfect and imperfect models are described by independent families of constant coefficient complex stochastic scalar equations for each Fourier mode \(v_k\), \(k = 1, \ldots, K\), where the number \(K\) represents the reciprocal of the truncation scale in space for observation of the natural extended system. Here, the perfect model has the form for spatial wavenumbers \(1 \leq k \leq K\)

\[
\frac{dv_k}{dt} = i\omega_k v_k - \gamma_k(t)v_k + \sigma_k(t)\tilde{W}_k + f_k(t), \quad (37)
\]

while the imperfect model is assumed to be given by

\[
\frac{dv_{M,k}}{dt} = i\omega_{M,k} v_{M,k} - \gamma_{M,k}(t)v_{M,k} + \sigma_{M,k}(t)\tilde{W}_{M,k} + f_{M,k}(t), \quad (38)
\]

for \(1 \leq k \leq N\). In (37) and (38), \(\tilde{W}_k\) and \(W_{M,k}\) are independent complex white noises for each \(k\), while \(f_k(t)\) and \(f_{M,k}(t)\) are the constant mean forcing. The known functions \(\omega_k, \omega_{M,k}\) represent perfect and imperfect model dispersion relations while the factors \(\gamma_k(t), \gamma_{M,k}(t)\) both represent turbulent and ordinary dissipation at a given spatial wavenumber. Here, we assume the same conditions for periodicity and climate stability as in (13) and (14) for each wavenumber \(k\). The equilibrium statistical state for both the perfect and imperfect models is a product of Gaussians for each wavenumber with mean and variance given by separate equations in the same form as in (17) and (18)

\[
\langle v_k(t) \rangle = G_{v_k}(t_0, t)\langle v_k(t_0) \rangle + \int_{t_0}^t G_{v_k}(s, t)f_k(s)\,ds, \quad \text{and}
\]

\[
\text{Var}[v_k(t)] = G_{v_k}(t_0, t)^2\text{Var}[v_k(t_0)] + \int_{t_0}^t G_{v_k}(s, t)^2\sigma_k^2(s)\,ds, \quad (39)\]

where \(G_{v_k}(s, t) = e^{-\int_t^s \gamma_k(s')\,ds' + \omega_k(t-s)}\). Equations (39) and (40) have time-periodic solutions that describe the attractor of the system [(37)] when the condition in (14) is satisfied for \(\gamma_k\). The periodicity implies that

\[
\langle v_k(0) \rangle_{\text{attr}} = \int_0^{T_0} G_{v_k}(s, T_0)f_k(s)\,ds \quad \text{and}
\]

\[
\text{Var}_{\text{attr}}[v_k(0)] = \frac{\int_0^{T_0} \sigma_k^2(s)G_{v_k}(s, T_0)^2\,ds}{1 - G_{v_k}(0, T_0)}, \quad (42)\]

with similar expressions for the imperfect model.

**Model error and coarse graining**

For spatially extended systems, the issue of coarse graining becomes important in the context of model climate fidelity and sensitivity. Here, we discuss how coarse graining is reflected in the quantification of model error using information metrics. For simplicity, we assume here, that \(\gamma_k, \sigma_k\), and \(f_k\) are all constants for both perfect and imperfect models, which means that there is no temporal dependence in the statistically steady state. Then, the expressions for the mean and variance on the attractor simplify

\[
\langle v_k \rangle_{\text{attr}} = \frac{f_k}{\gamma_k - i\omega_k},
\]

\[
\text{Var}_{\text{attr}}(v_k) = \frac{\sigma_k^2}{2\gamma_k}, \quad \text{and}
\]

\[
\text{Var}_{\text{attr}}(v_{M,k}) = \frac{\sigma_{M,k}^2}{2\gamma_{M,k}}.
\]

Note that each Fourier mode \(v_k\) has Gaussian probability distribution, and therefore the relative entropy has the form in (10). In particular, we use \(P(\pi, \pi^M)\) to quantify the uncertainty in the imperfect model for various coarse-grainings. Note that since each Fourier mode in (37) and (38) is given by a complex equation, we treat each Fourier mode as a pair of real variables, that is, real and imaginary parts of a complex variable, respectively. In equilibrium, there is no correlation between these two components because the noises \(\tilde{W}_k\) and \(W_{M,k}\) have independent real and imaginary parts. Therefore, the covariance matrix for the system of \(K\) complex Fourier modes is diagonal and has the size \(2K \times 2K\).

The total uncertainty in the system is computed as a sum of uncertainties for each individual Fourier mode \(v_k\), which is given by Eq. (11) with \(J = 2\).

Next, we present the results of a particular simulation for (37). We consider a model with \(K = 100\) Fourier modes, each mimicking the dynamics of the \(k\)th atmospheric Rossby wave with \(\gamma_k = \nu(F_k + k^2), \omega_k = \beta k(F_k + k^2)^{-1}, \nu\) the viscosity, and the power spectrum \(E_k = \text{Var}(v_k)\) of a particular form.
The value $u = 3$ is appropriate for large-scale turbulence in the atmosphere. In Fig. 8, we show the statistics of the model in (37) and (38) in equilibrium. We used the following parameters for the perfect model: $\nu = 10^{-3}$, $\beta = 8.91$, $F_s = 16$, $K_0 = 10$, and $\theta = 3$. Also, we picked arbitrary forcing parameters $f_k$ so that the magnitude of the equilibrium mean $\langle v_k \rangle$ is roughly twice the magnitude of the standard deviation of $v_k$, that is,

$$|f_k| \sim 2 \frac{\sigma_k}{\sqrt{2} \gamma_k} \sqrt{\gamma_k^2 + \omega_k^2}. \quad (44)$$

Most models for turbulence in spatially extended systems have too much dissipation (Palmer 2001) because of inadequate resolution and deterministic parameterization of unresolved features. This is due to wave–mean flow interaction and stochastic backscatter, which is not captured in the simple models in (37) and (38). Nevertheless, we mimic this overdamping in these simple models. Thus, for simplicity it is assumed that the energy spectra satisfy

$$E_k > E_{M_k} \quad \text{for} \quad k = 1, \ldots, K. \quad (45)$$

For the purpose of comparison, we also consider the opposite situation, when the imperfect spectrum is larger than the perfect one, that is, $E_k < E_{M_k}$. So, for the imperfect model, we specify two modeled power spectra: $E_{M_k}$

$$E_k = \begin{cases} 1, & k \leq K_0, \\ \left(\frac{k}{K_0}\right)^{-\theta}, & \text{otherwise}. \end{cases} \quad (43)$$

$$E_{M_k} = \begin{cases} 0.8, & k \leq K_0, \\ 0.8 \left(\frac{k}{K_0}\right)^{-\theta_{M_k}}, & \text{otherwise}, \end{cases} \quad (46)$$

with $\theta_{M_1} = 3.5$ for the overdamped imperfect model and $E_{M_k}$

$$E_{M_k} = \begin{cases} 1.2, & k \leq K_0, \\ 1.2 \left(\frac{k}{K_0}\right)^{-\theta_{M_2}}, & \text{otherwise}, \end{cases} \quad (47)$$

with $\theta_{M_2} = 2.5$ for the underdamped imperfect model. We assume that the noise strength of the imperfect model, the forcing, and the dispersion are the same as in the perfect model, that is, $\sigma_{M,k} = \sigma_k$, $f_{M,k} = f_k$, and $\omega_{M,k} = \omega_k$ for both models. The dissipation of the imperfect model becomes

$$\gamma_{M,k} = \frac{\sigma_k^2}{2E_{M_k}}. \quad (48)$$

In Fig. 9, we show the lack of information in the imperfect models as a function of coarse graining for both overdamped and underdamped cases. We note that as we take more modes into consideration, that is, refine the resolution [and hence include more terms in the summation in (11)], the uncertainty of the imperfect model grows. In our setup, for the coarse graining of up to 15 modes, the uncertainty is mostly concentrated in the dispersion; however, as more modes are added, there is a qualitative jump in the distribution of the uncertainty between the signal and dispersion parts, and most
of the uncertainty is contained in the signal part, as the number of modes in coarse graining exceeds 20. Interestingly, we note that the fraction of the signal part in the total lack of information is practically constant as the number of modes in the coarse graining is increased beyond 20 for both model error scenarios. We also note that in the case of the overdamped imperfect model the fraction of the signal part is around 60% while in the case of the underdamped imperfect model, this value rises up to 80%. Furthermore, (Majda and Gershgorin 2011a) the total lack of information is smaller for the underdamped model and there is less demand in improving the model in this case as the spatial resolution increases than in the case of overdamped imperfect model. Note that in Fig. 9, we specifically show the area that corresponds to the first 15 modes separately to demonstrate the behavior of the lack of information at large scale coarse-graining.

Next, we study the improvement of the climate prediction by using the optimal noise in the imperfect model. The tuning procedure is exactly the same as described in section 3 except that here we have $K$ different Fourier modes instead of 1. For simplicity, we add the same noise to all the modes so there is only a single model parameter and find the optimal noise according to the general principle in (6), for which the uncertainty in the imperfect model is minimized. We note that the optimal value of the noise strongly depends on the coarse graining because here the uncertainty $P(p_{\text{attr}}, p_{\text{M}})$ in (6) is computed for a certain coarse graining $1 \leq k \leq K'$. In Fig. 10, we compare the model error for the climate change prediction using two imperfect models with various coarse grainings, where both models have overdamped energy spectrum as in (46) and one model is not optimized for climate consistency and the other one is optimized. We note drastic improvements in the predictive skill for the whole range of coarse grainings in the optimally tuned model. The improvement is significant in both signal and dispersion components of the uncertainty. Moreover, it is worth noting that the fraction of the signal part becomes larger in the optimized model as the resolution increases in the considered
model error scenario. Also in Fig. 10, we show the value of the optimal additional noise parameter $D_s$, which is obtained using the optimization procedure based on (6). Note that this parameter is a function of coarse-graining $K_9$ and for a given coarse graining, the additional noise is the same for all the modes $k \leq K_9$ according to the setup of our experiment. It turns out that the additional noise increases as the number of modes in the coarse graining increases.

5. Gaussian model for turbulent tracers

In this section, we utilize an instructive model (Majda and Gershgorin 2010; Gershgorin and Majda 2011; Majda and Gershgorin 2011a,b, 2012) with nontrivial eddy-diffusivity, variance spectrum and intermittent non-Gaussian statistics like tracers in the atmosphere (Neelin et al. 2010). Here, we consider a special case of the tracer model, when the tracer has Gaussian statistics, as our next example of Gaussian test models for uncertainty quantification. Note that in this example we incorporate all the features that we discovered for simple models for scalar Gaussian fields in section 3 and for a linear PDE model for Rossby waves discussed in section 4. Here, new effects due to nontrivial dependence of the tracer on the velocity field become important.

a. Model description

This model is a special case of a general model for advection–diffusion of a passive tracer in a turbulent velocity field

$$\frac{\partial T}{\partial t} + \nu \cdot \nabla T = \kappa \Delta T - d_T T, \quad (49)$$

where $T$ is a tracer, $\nu$ is velocity field, and $\kappa$ is molecular diffusion. In the case of a two-dimensional velocity field $\nu = [\mathcal{U}(t), \mathcal{V}(x, t)]^T$ with a tracer mean gradient
the model becomes one-dimensional for the perturbations of the tracer around the mean gradient if they depend on the $x$ coordinate alone. We denote the perturbations around the mean gradient as $T$ from now on, resulting in

$$\frac{\partial T}{\partial t} + U(t)\frac{\partial T}{\partial x} = \kappa \frac{\partial^2 T}{\partial x^2} - d_T T - \alpha v,$$  

where the uniform damping term $-d_T T$ was added to damp the zeroth mode. Here, we further simplify the model by assuming that $U(t)$ is deterministic and satisfies the linear equation

$$\frac{dU}{dt} = -\gamma U + f(t),$$  

where $\gamma$ is constant, and $f(t)$ is given in (24). For physical relevance, we demand that $U$ stays positive (eastward direction of the zonal jet) as discussed in section 3 and

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$A_0$</th>
<th>$A_1$</th>
<th>$A_2$</th>
<th>$A_3$</th>
<th>$K$</th>
<th>$K_0$</th>
<th>$v$</th>
<th>$F_t$</th>
<th>$\beta$</th>
<th>$E_{1-K}$</th>
<th>$E_{K+1-K}$</th>
<th>$d_T$</th>
<th>$\kappa$</th>
<th>$\alpha$</th>
</tr>
</thead>
<tbody>
<tr>
<td>2/3</td>
<td>4</td>
<td>1.5</td>
<td>1.5</td>
<td>1.1</td>
<td>3</td>
<td>20</td>
<td>5</td>
<td>$10^{-3}$</td>
<td>16</td>
<td>8.91</td>
<td>0.8</td>
<td>$-k^{-3}$</td>
<td>0.1</td>
<td>$10^{-3}$</td>
</tr>
</tbody>
</table>

**FIG. 11.** Statistics of the Gaussian tracer model for perfect (solid line) and imperfect (dashed line) parameters as a function of time. (top) Deterministic mean jet, (next row) mean tracer, (next row) variance of the tracer, and (bottom) correlation of the tracer and velocity field. (left to right) Modes 1, 6, and 13. Note that higher dissipation in the imperfect model leads to damped statistics.
\( y(x, t) \) is linear Gaussian given by (37). Equation (50) can be solved in Fourier space for each mode, \( T_k(t) \)

\[
T_k(t) = G_{T_k}(t, t)T_k(t_0) - \alpha \int_{t_0}^{t} G_{T_k}(s, t)v_k(s) \, ds,
\]

where \( G_{T_k}(s, t) \) is the Green’s function for the tracer

\[
G_{T_k}(s, t) = e^{-\left(\sigma_k + ik^2\right)(t-s) - ik \int_{s}^{t} U(s') \, ds'}.
\]

In the study of climatology and climate response to external perturbations, we are interested in the steady state statistics of the solution in (57). To mimic the seasonal cycle in the atmosphere, we assume that the velocity field is statistically periodic with a period \( T_0 \), which means that \( U(t + T_0) = U(t) \) and the mean and covariance of \( v_k(t) \) are also periodic with the same period \( T_0 \). Then, we can find the periodic steady state statistical solution for 57. In particular, for the mean, \( \langle T_k \rangle \), we have

\[
\langle T_k(0) \rangle = -\alpha \int_{0}^{T_0} G_{T_k}(s, T_0)v_k(s) \, ds - \frac{1}{1 - G_{T_k}(0, T_0)}.
\]

Similar explicit formulas as in [[20], (21) for \( \text{Var}[T_k(t)] \) and \( \text{Cov}[T_k(t), v_k(t)] \) for the climate with \( t \) varying through the year are recorded in the appendix.

Next, we give a statistical description of the steady state of this linear tracer model. The fact that the jet is deterministic here is reflected in the Gaussianity of the tracer, however, time periodicity of the jet leads to time periodicity of the mean and variance of the tracer as well as the cross covariance of the tracer and the waves. In Table 2, we provide parameters for all three components of the tracer model, that is, \( U \), \( v_k \), and \( T_k \). In Fig. 11, we show the statistics of the tracer as function of time in the periodic statistically steady state. Each Fourier mode is independent from others; however, \( v_k \) and \( T_k \) are correlated for the same wavenumber \( k \). We note that the variance of the tracer increases as the jet decreases and vice versa. We also note that the cross correlation of the waves and the tracer (which we compute as cross-covariance normalized by the product of standard deviations) is significant and should not be ignored. Next, in Fig. 12, we show the spectrum of the waves that is given by \( \text{Var}(v_k) \) and the spectrum of the tracer \( \text{Var}(T_k) \). The variance of the waves is constant in time while the variance of the spectrum depends on time as we have shown in Fig. 11; therefore, in Fig. 12, we show the time-averaged spectrum of the tracer. The spectrum of the waves is flat for the first 5 modes and then decays as the power law \( \sim k^{-3} \) for the remaining 15 modes. The corresponding spectrum of the tracer has an approximate power law decay \( k \sim k^{-1.8} \) for the first 5 modes that correspond to the flat part of the spectrum of the waves. For the remaining 15 modes, the spectrum of the tracer decays as \( \sim k^{-5} \).

b. Improving climate consistency by optimizing noise

Imperfect models are usually used for predicting the future states of climate models. The reasons for the
deficiencies in the current computer climate models, include the lack of physical understanding of the system dynamics or inappropriate parameterization of small-scale processes because of the limitations in computing power. One very common type of model error comes from overdissipation (Palmer 2001) that often results in an eddy diffusivity approximation. Here, we consider this type of model error in the context of the Gaussian tracer model. In particular, we assume that the spectrum of the imperfect model’s velocity field $v_k$ is overdamped as in section 4

$$E_{M,k} = \begin{cases} 0.7, & k \leq K_0, \\ 0.7 \left( \frac{k}{K_0} \right)^{-4}, & \text{otherwise}, \end{cases}$$ \hspace{1cm} (55)$$

and the diffusion increases to $\kappa_M = 0.6$ instead of the original $\kappa = 0.001$. The rest of the parameters of the imperfect model are equal to their counterparts in the perfect model (see Table 2). The climate statistics of this imperfect model are shown in Fig. 11 with dashed line. We note that the statistics of the imperfect model are overdamped especially for higher wavenumbers. In Fig. 13, we show the uncertainty in the climate state due to model error measured as $P(\pi, \pi^M)$ for various coarse grainings. We also show the fraction of the signal in the total uncertainty for the coarse graining of $k \leq 13$ modes in linear scale along the $y$ axis that emphasizes the temporal structure over the period. The peak at time $t = 8$ corresponds to weaker zonal jet while the peak at time $t = 0$ corresponds to overdamped variance of the imperfect model (see Fig. 11).

Fig. 13. (top) Relative entropy of climate as a function of time in period for different coarse grainings: $k \leq 1$ (solid), $k \leq 3$ (dashed), $k \leq 6$ (dashed-dotted), $k \leq 9$ (circles), and $k \leq 13$ (dotted). Note the log scale of the $y$ axis for easier comparison. (left) For both waves and tracer and (right) only for the tracer. Relative entropy grows significantly as more modes are included in the coarse graining. (middle and bottom) Relative entropy of climate for the coarse graining of $k \leq 13$ modes in linear scale along the $y$ axis that emphasizes the temporal structure over the period. The peak at time $t = 8$ corresponds to weaker zonal jet while the peak at time $t = 0$ corresponds to overdamped variance of the imperfect model (see Fig. 11).
only the tracer is used in quantifying uncertainty (real and imaginary parts of the tracer). It is important to realize that in the former case, each covariance matrix $R_k$ is not diagonal because of the nontrivial covariance $\text{Cov}(y_k, T_k)$. We note that the uncertainty of the climate state due to model error increases rapidly as more modes are included into the coarse graining. We also note higher uncertainty during the time when the zonal jet $U(t)$ has smaller values (around time units 7, 8, and 9 in Fig. 13). The uncertainty measured for the tracer only is much lower than for the whole system of the waves and the tracer. Another interesting effect that is demonstrated in Fig. 13 is that for the coarse graining that includes higher wave numbers, $1 \leq k \leq 13$, there is a rise of uncertainty around the yearly time $t = 0$. We attribute this to the fact that, according to Fig. 11, the variance of the imperfect model practically vanished at $t = 0$ for high wavenumbers. Therefore, in the system with high resolution there are two different sources of uncertainty, one comes from the weaker zonal jet and the other comes from the particular type of model error, when the variance is overdamped at certain times. In Fig. 13, we also compare the case when both the velocity $v_k$ and the tracer $T_k$ are taken into account with the case, when the uncertainty is computed for the tracer alone. Note that, in the former case, the uncertainty is much higher than in the case with tracer alone because of high values of uncertainty in the velocity as a result of incorrect spectrum specification. From Fig. 13, we see that, at the system coarse grained to 13 modes, the fraction of the signal part is approximately 0.5 in the total uncertainty, with higher values around yearly times 7, 8, and 9 when the zonal jet is weaker and lower values otherwise, when the zonal jet is stronger and the dispersion part slightly dominates.

As we have shown in sections 3 and 4, one natural way to improve the model error due to overdamping is to add a noise term to the dynamic equation.

FIG. 14. (top) Relative entropy of the imperfect model $P(\pi, \pi^M)$, as a function of the noise-tuning parameter $\sigma_T$ for different coarse grainings. Circles show the minima of the relative entropy that yield the optimal values for $\sigma_T$. (bottom left) Optimal noise $\sigma_T$ as a function of coarse graining. (bottom right) Model improvement as a ratio of the relative entropy after and before adding optimal noise.
\[
\frac{\partial T_M}{\partial t} + U(t) \frac{\partial T_M}{\partial \alpha} = \kappa \frac{\partial^2 T_M}{\partial \alpha^2} - d_T T_M - \alpha v_M + \sigma_T \dot{W}.
\]

(56)

This noise term increases the variance of the tracer, which in the statistically steady case becomes

\[
\text{Var}(T_M)_{\sigma_T} = \text{Var}(T_M)_{\sigma_T=0} + \frac{\sigma_T^2}{2(\kappa k^2 + d_T)}
\]

(57)

where \(\text{Var}(T_M)_{\sigma_T=0}\) is computed via (A1). This noise term does not have impact on the mean of the tracer and the covariance of the tracer and the velocity. To find the best value for the noise \(\sigma_T\), we optimize it with respect to the lack of information following the general principle in (6) and (26), that is, we find \(\sigma_T^*\) that satisfies

\[
P[\pi, \pi_M(\sigma_T^*)] = \min_{\sigma_T} P[\pi, \pi_M(\sigma_T)],
\]

(58)

where \(M^*\) is the model with an optimal value for \(\sigma_T^*\), and \(M\) is a general model with some noise \(\sigma_T\). Recall that the bar denotes time average. We note that there are a few important aspects that need to be emphasized here:

- optimal noise depends on coarse graining;
- optimal noise depends on the metric we use, it can be relative entropy for the full system that includes the velocity field and the tracer or just for the tracer; and
- optimal noise depends on the time interval over which the time averaging is taken.

In Fig. 14 (upper panel), we show how the relative entropy depends on the noise parameter \(\sigma_T\) for different coarse grainings. Note that at a given coarse graining, we use the same noise parameter \(\sigma_T\) for every Fourier mode as we did in section 4 for the system of linear Rossby waves. We show that for any coarse graining, there is an optimal value of the noise, for which the relative entropy is minimal. Moreover, we note a rather sharp drop in the uncertainty, when the noise parameter \(\sigma_T\) starts growing from zero. After reaching its minimum value, the relative entropy starts growing again. Also in Fig. 14 (lower left panel), we show the optimal noise parameter \(\sigma_T\) as a function of coarse graining. We note that as the number of modes in the coarse graining grows from 1 to 5, the optimal noise also grows from \(\sigma_T \approx 0.14\) to \(\sigma_T \approx 0.26\). Then, the optimal noise stays at approximately the same level up until the coarse graining includes 12 modes, and then the optimal noise slightly declines to \(\sigma_T \approx 0.23\) at the coarse graining of 20 modes, which is the whole system in our case. In Fig. 14 (lower right panel), we show the ratio of the relative entropies after and before adding optimal noise. Naturally, we expect this ratio to be smaller than one, otherwise, no actual improvement was made according to (7). Here, this ratio is very small and decreases as the number of modes in the coarse graining increases. Quantitatively, the actual improvement in the information theoretic measure is more than 97%.

c. Improving predictive skill of imperfect model

Here, we discuss the predictive skill of imperfect models when the external perturbations change the climate statistics. The external perturbation \(\delta f(t)\) is applied to the forcing of the deterministic mean jet similarly to the perturbation considered in section 3. However, unlike in that case, here we only consider deterministic perturbations of the form

\[
\delta f(t) = \delta A_0 \phi_n(t),
\]

(59)

where the ramp function \(\phi_n(t)\) is given by Eq. (29) and increases the jet strength. The response to this type of perturbation is the decrease of both mean and variance of the tracer for each wavenumber as shown in Fig. 15. Note that since the external perturbation starts acting at \(t > t_0 = 0\) both mean and variance start decreasing compared to their climate values until the forcing perturbation stops changing and then the new climate is established. The question we address next is how well can we capture the true system response using imperfect models and how can we improve the predictive skill of a given model? Based on the previous experience with the simple models from sections 3 and 4, we anticipate that the models that are tuned to capture the climate
state better in the information-theoretic sense in (26) also have better prediction skill in the context of climate change. To demonstrate this behavior on this particular tracer model, we monitor the measure of model error $P(\pi_\delta, \pi^M_\delta)$ as external perturbations is applied to the forcing of $U(t)$. For the imperfect model, we choose the same imperfect model as in the climate consistency study in section 5b when the spectrum of the velocity field is overdamped and the diffusivity is replaced by rather strong eddy diffusivity. In Figs. 16 and 17 (left panels), we demonstrate the uncertainty in the climate change prediction measured as $P(\pi_\delta, \pi^M_\delta)$ as a function of time for various coarse grainings from $K = 1$ to $K = 20$. Moreover, we also separate two stages of climate change: in the first stage, the external perturbation, $\delta f(t)$, grows linearly, and in the second stage, it stays constant, which leads to the new climate. We note that in the first stage of climate change, the uncertainty is generally higher than in the second stage. Also, as more modes are included in the coarse graining, the uncertainty increases dramatically as we already demonstrated in Fig. 13 in the climate consistency study. Next, we compare the prediction skill of the imperfect model with the predictive skill of the optimized imperfect model. The optimization is performed for climate consistency and is discussed above. The measure of the predictive skill $P(\pi_\delta, \pi^M_\delta)$ is shown in Figs. 16 and 17 (right panels) next to the corresponding values of uncertainty for the original imperfect model. For the coarse grainings of $K' = 1, 3,$ and $5$ modes, the improvement is roughly 4 times, while for higher coarse grainings of up to $K' = 20$ modes, the improvement reaches the values of 20 times.

6. Conclusions

Here, the recently developed information theoretic framework for improving model fidelity and sensitivity for complex systems (Majda and Gershgorin 2010, 2011a,b)
is applied to a suite of mathematically tractable physically relevant test models. In particular, we utilized three linear Gaussian test models that mimic certain features of the earth’s climate system, to demonstrate how two important aspects of model error, namely climate consistency and sensitivity to external and internal perturbations, can be improved simultaneously within this framework. The models mimic the behavior of the zonal jet in the atmosphere with the seasonal cycle, the dynamics of linear Rossby waves with turbulent spectrum, and the evolution of a passive tracer in the turbulent velocity field with deterministic cross sweep. This last model mimics the global and regional behavior of turbulent passive tracers under various climate change scenarios. We consider a typical model error scenario, when the system’s dissipation is overestimated in the imperfect model, which is a common deficiency of contemporary computer climate models. We address whether model deficiencies are systematically improved by suitable additional stochastic noise Palmer (2001). We successfully apply the general principle of information theory that minimizes the lack of information in the imperfect model to optimize such a model for climate consistency via additional stochastic noise. Moreover, for all three models, we observed significant improvement of the predictive skill in the climate change scenarios for imperfect models that are optimized with stochastic noise in the unperturbed climate. Furthermore, we studied the role of initial conditions in making short-, medium-, and long-range forecasts in different seasons. We also focused on spatial coarse graining that is an important aspect of uncertainty quantification in spatially extended systems for local and regional forecasting skill. We find that the optimal noise value nontrivially depends on the level of coarse graining.

While adding noise to improve complex climate models has been suggested long ago (Hasselman 1976) all the results here indicate that an information-theoretic framework for tuning the climate of imperfect models is useful for designing the optimal noise to enhance predictive skill. An important future research direction that naturally follows this work is to use nonlinear and non-Gaussian models to demonstrate various facets of the information theoretic approach to model error. This is the subject of ongoing research that emphasizes a much longer memory of the initial conditions compared to the Gaussian case, the role of coarse-graining, and the link between
climate fidelity and sensitivity through the fluctuation–
dissipation theorem (Majda and Gershgorin 2011b).
While we have emphasized the use of additional noise to
improve overdamped imperfect models, different strat-

gies need to be developed for improving underdamped
imperfect models.

Acknowledgments. This research of A.J.M. is par-
tially supported by National Science Foundation Grant
DMS-0456713 and the office of Naval Research Grants
25-74200-F6607 and N00014-11-1-0306. B.G. is supported
as a postdoctoral fellow on the first grant.

APPENDIX

Second-Order Statistics of the Tracer

Here, we show the formulas for the variance of the tracer

\[
\text{Var}[T_k(0)] = \alpha^2 \frac{\int_0^{T_0} \int_0^{T_0} G_{T_k}(s, T_0) G_{v_k}(r, T_0) \exp[-\gamma_0 |s-r| + i\omega_0 (r-s)] \, ds \, dr}{1 - |G_{T_k}(0, T_0)|^2} - \alpha \frac{\int_0^{T_0} G_{T_k}(0, T_0) \text{Cov}[T_k(0), v_k(0)] \int_0^{T_0} G_{T_k}(s, T_0) G_{v_k}(0, s) \, ds + c.c.}{1 - |G_{T_k}(0, T_0)|^2}.
\]

and for the cross-covariance of the tracer and velocity
field

\[
\text{Cov}[T_k(0), v_k(0)] = \alpha \frac{\int_0^{T_0} G_{T_k}(s, T_0) G_{v_k}(s, T_0) \, ds}{1 - |G_{T_k}(0, T_0)|^2}.
\]

REFERENCES


Branstator, G., and H. Teng, 2010: Two limits of initial-value de-
cadal predictability in a CGCM. J. Climate, 23, 6292–6311.

Burnham, K., and D. Anderson, 2002: Model Selection and Multi-


——, 2005: Predictability and information theory. Part II: Imper-


Gardiner, C., 2010: Stochastic Methods: A Handbook for the Nat-


Giannakis, D., and A. J. Majda, 2012a: Quantifying the predictive skill in long-range forecasting. Part I: Coarse-grained predic-
tions in a simple ocean model. J. Climate, 25, 1793–1813.

——, ——, and ——, 2012b: Quantifying the predictive skill in long-

Gritsun, A., and G. Branstator, 2007: Climate response using a three-
dimensional operator based on the fluctuation–dissipation theo-

——, ——, and A. Majda, 2008: Climate response of linear and
quadratic functionals using the fluctuation–dissipation theo-


Kleeman, R., 2002: Measuring dynamical prediction utility using


