Some theoretical considerations on predictability of linear stochastic dynamics*

Michael K. Tippett†
International Research Institute for Climate Prediction, Palisades, New York

Ping Chang
Department of Oceanography, Texas A& M University, Texas

May 31, 2002

*International Research Institute for Climate Prediction Contribution Number IRI-PP/02/??.

†Corresponding author:IRI/LDEO 223 Monell, PO Box 1000 / 61 Route 9W, Palisades, NY 10964-8000. Phone: 845-680-4420. Fax: 845-680-4865. Email: tippett@iri.columbia.edu

Abstract

Predictability is a measure of prediction error relative to observed variability and so depends on both the physical and prediction systems. Here predictability is investigated for climate phenomena described by linear stochastic dynamics and prediction systems with perfect initial conditions and perfect linear prediction dynamics. Predictability is quantified using the *predictive information matrix* constructed from the prediction error and climatological covariances. Predictability measures defined using the eigenvalues of the predictive information matrix are invariant under linear state-variable transformations and for univariate systems reduce to functions of the ratio of prediction error and climatological variances. The predictability of linear stochastic dynamics is shown to be minimized for stochastic forcing that is uncorrelated in normal-mode space. This minimum predictability depends only on the eigenvalues of the dynamics and is a lower bound for the predictability of the system with arbitrary stochastic forcing. Issues related to upper bounds for predictability are explored in a simple theoretical example.

1. Introduction

The original work of Hasselmann (1976) on stochastic climate theory pioneered the use of linear stochastic dynamics for modeling and predicting various modes of climate variability. Since then, evidence has been presented that many climate phenomena are described, at least approximately, by linear stochastic dynamics (Penland and Matrosova, 1994; Whitaker and Sardeshmukh, 1998). Consequently linear stochastic models are routinely used for operational climate forecasts (Penland and Matrosova, 1998; Winkler et al., 2001). A general understanding of the predictability of phenomena described by linear stochastic dynamics is of theoretical interest and practical value.

A system is predictable on those time-scales where prediction errors do not exceed some predetermined fraction of climate variability (Lorenz, 1969). The ratio of prediction error variance to observed climatological variance is a measure of predictability in univariate systems with Gaussian distributions. This notion of predictability based on relative error can be extended to multivariate systems using the *predictive information matrix* constructed from the prediction error and climatological covariances (Schneider and Griffies, 1999). Eigenvalues of the predictive information matrix are invariant under linear transformations of the state-variable and can be used to define predictability measures that are independent of the choice of inner product or vector-norm, unlike the commonly used predictive error variance (Schneider and Griffies, 1999; Kleeman, 2002).

Prediction error statistics are a function of both the physical and prediction systems. Here we consider physical systems described by linear stochastic dynamics and prediction systems with linear deterministic dynamics. In this setting, prediction error is due to inaccurate initial conditions, misspecification of the dynamics, and the stochastic forcing present in the description of the observed system (Penland and Matrosova, 2001). In this study, we limit our investigation to systems where the only source of prediction error is the stochastic forcing. Prediction error in similar systems has been studied by many authors. Some of the issues considered have included analysis of the prediction error response to homogeneous, spatially uncorrelated stochastic forcing and characterization of the stochastic forcing structures that maximize prediction error growth (stochastic optimals) (Ioannou, 1995; Farrell and Ioannou, 1996; Kleeman and Moore, 1997). The focus of previous studies has been prediction error variance or in some cases the ratio of prediction error and climatological variances (Penland, 1989; Chang et al., 2002). However, a limitation of such results is that measures of prediction error such as variance are not invariant with respect to linear transformations of the state-variable; equivalently, they depend on the choice of vector-norm or inner product. Classification of dynamics as normal or nonnormal depending on the orthogonality of its eigenvectors, an important theme in such studies, is also not invariant with respect to linear transformations of the state-variable.

Here we examine predictability of linear stochastic dynamics using norm-independent predictability measures based on the predictive information matrix. We find that minimum predictability is achieved for stochastic forcing that is uncorrelated in normal-mode space. The minimum predictability is expressed simply in terms of the eigenvalues of the dynamics and is a lower bound for predictability in systems where the stochastic forcing has arbitrary spatial structure.

We begin our discussion in Section 2 with an introduction to a linear stochastic system and its associated prediction error; in Section 3 we analyze the predictability of the linear stochastic system under a set of general predictability measures; in Section 4 we use a theoretical example to explore open issues related to upper bounds for predictability; in Section 5 we summarize our findings and discuss their implications.

2. Linear stochastic dynamics

We assume the observed phenomenon of interest is represented by a real n-dimensional state-vector $\mathbf{w}^{\text{obs}}$ whose evolution is governed by linear stochastic dynamics. That is to say, we assume the observed state $\mathbf{w}^{\text{obs}}$ satisfies

$$\frac{d\mathbf{w}^{\text{obs}}}{dt} = \mathsf{A}\mathbf{w}^{\text{obs}} + \mathsf{F}\xi\,, \quad \mathbf{w}^{\text{obs}}(t=0) = \mathbf{w}_0^{\text{obs}}\,, \tag{1}$$

where the dynamics matrix A and the forcing matrix F are constant, real $n \times n$ matrices; ξ is n-dimensional, spatially uncorrelated Gaussian white-noise, $\langle \xi(t_1)\xi(t_2)^T \rangle = \delta(t_1 - t_2)\mathsf{I}$ where the notation $\langle \cdot \rangle$ denotes ensemble average and $()^T$ denotes matrix transpose. Properties of (1) have by studied in a geophysical context by Penland (1989), DelSole and Farrell (1995), Penland and Sardeshmukh (1995) and others.

A general deterministic prediction system has the form

$$\frac{d\mathbf{w}^{\text{pred}}}{dt} = \mathsf{A}^{\text{pred}}\mathbf{w}^{\text{pred}}\,, \quad \mathbf{w}^{\text{pred}}(t=0) = \mathbf{w}_0^{\text{pred}}\,. \tag{2}$$

Differences between the observed state $\mathbf{w}^{\text{obs}}$ and the predicted state $\mathbf{w}^{\text{pred}}$ are due to (*i*) differences between the observed initial condition $\mathbf{w}_0^{\text{obs}}$ and the prediction initial condition $\mathbf{w}_0^{\text{pred}}$, (*ii*) deficiencies in the prediction dynamics A^{pred} and, (*iii*) the presence of stochastic processes in the observations (Penland and Matrosova, 2001). This dependence is explicit in the equation for the evolution of the prediction error $\mathbf{w} \equiv \mathbf{w}^{\text{obs}} - \mathbf{w}^{\text{pred}}$

$$\frac{d\mathbf{w}}{dt} = \mathsf{A}^{\text{pred}}\mathbf{w} + (\mathsf{A} - \mathsf{A}^{\text{pred}})\mathbf{w}^{\text{obs}} + \mathsf{F}\xi\,, \quad \mathbf{w}(t=0) = \mathbf{w}_0^{\text{obs}} - \mathbf{w}_0^{\text{pred}}\,. \tag{3}$$

The prediction error dynamics forcing consists of two components. The first component $(\mathsf{A} - \mathsf{A}^{\text{pred}})\mathbf{w}^{\text{obs}}$ represents prediction error due to imperfect deterministic dynamics of the prediction model and depends on the observed state; the second component $\mathsf{F}\xi$ represents the unpredictable stochastic processes in the observations and is state-independent. We take the forcing to be state-independent and assume that the prediction system has perfect dynamics $\mathsf{A}^{\text{pred}} = \mathsf{A}$. Additionally we assume perfect initial conditions $\mathbf{w}_0^{\text{pred}} = \mathbf{w}_0^{\text{obs}}$. Therefore, the source of the prediction error is entirely due to the stochastic processes. Chang et al. (2002) refer to this situation as the *perfect* *initial condition* scenario and give a more detailed discussion. In this scenario, the prediction error $\mathbf{w}$ evolves according to

$$\frac{d\mathbf{w}}{dt} = \mathsf{A}\mathbf{w} + \mathsf{F}\xi\,, \quad \mathbf{w}(t=0) = 0\,. \tag{4}$$

We take the dynamics matrix A to be stable, i.e., all its eigenvalues $\lambda_k(\mathsf{A})$ have negative real part. We use the convention that the eigenvalues of A are ordered least damped to most damped so that $0 > \operatorname{Re}\lambda_1(\mathsf{A}) \geq \operatorname{Re}\lambda_2(\mathsf{A}) \cdots \geq \operatorname{Re}\lambda_n(\mathsf{A})$.

The prediction error covariance at lead-time τ is defined as $\mathsf{C}_\tau \equiv \langle \mathbf{w}(\tau)\mathbf{w}(\tau)^T \rangle$ and is well-described by its eigenvectors and eigenvalues. The stability of the dynamics means that predictions are identically zero in the limit of large lead-time τ, and consequently the infinite lead-time prediction error covariance C_∞ is the climatological covariance. The eigenvectors or EOFs of the prediction error covariance are orthogonal and order state-space according to the amount of variance they explain; $\lambda_k(\mathsf{C}_\tau)$ is the variance explained by the k-th eigenvector of C_τ; $\lambda_1(\mathsf{C}_\tau) \geq \lambda_2(\mathsf{C}_\tau) \geq \cdots \geq \lambda_n(\mathsf{C}_\tau) \geq 0$. Since orthogonality depends on the choice of norm, or equivalently on the choice of state-variable, the eigenvalue decomposition of the prediction error covariance is not invariant under linear transformations of the state-variable. If we define a new state-variable $\hat{\mathbf{w}} = \mathsf{L}\mathbf{w}$ where L is a linear transformation, the transformed prediction error covariance $\hat{\mathsf{C}}_\tau \equiv \langle \hat{\mathbf{w}}(\tau)\hat{\mathbf{w}}(\tau)^T \rangle$, is given by $\hat{\mathsf{C}}_\tau = \mathsf{L}\mathsf{C}_\tau\mathsf{L}^T$. The prediction error covariance matrices C_τ and $\hat{\mathsf{C}}_\tau$ have the same eigenvalues only when L is an orthogonal transformation, in which case the transformation L relates eigenvectors of C_τ to eigenvectors of $\hat{\mathsf{C}}_\tau$. Measures of prediction error growth that depend on the eigenvalues of C_τ, such as the total variance $\operatorname{tr}\mathsf{C}_\tau$, are invariant only under orthogonal transformations of the state-variable.

For the prediction error dynamics in (4), the prediction error covariance is

$$\mathsf{C}_\tau = \int_0^\tau e^{t\mathsf{A}}\mathsf{F}\mathsf{F}^T e^{t\mathsf{A}^T}\, dt\,. \tag{5}$$

Suppose the dynamics matrix A is diagonalizable with eigendecomposition $\mathsf{A} = \mathsf{Z}\Lambda\mathsf{Z}^{-1}$; the matrix Λ of eigenvalues is an $n \times n$ diagonal matrix whose k-th diagonal entry is $\lambda_k(\mathsf{A})$; the k-th column of the $n \times n$ matrix Z is the eigenvector $\mathbf{z}_k$ and satisfies $\mathsf{A}\mathbf{z}_k = \lambda_k(\mathsf{A})\mathbf{z}_k$. Eigenmodes of the dynamics are also called principal oscillation patterns (Hasselmann, 1988; Penland, 1989). The matrix Y of adjoint eigenvectors of A is defined by $\mathsf{Y} \equiv (\mathsf{Z}^{-1})^\dagger$ where the notation $()^\dagger$ denotes conjugate transpose. We assume without loss of generality that the columns $\mathbf{y}_k$ of Y are unit vectors with $\mathbf{y}_k^\dagger \mathbf{y}_k = 1$. Now the prediction error covariance C_τ can be expressed in the basis of the eigenvectors of the dynamics as $\mathsf{C}_\tau = \mathsf{Z}\tilde{\mathsf{C}}_\tau \mathsf{Z}^\dagger$. Using the relation $\mathsf{Y}^\dagger \mathsf{Z} = \mathsf{Z}^\dagger \mathsf{Y} = \mathsf{I}$, the matrix $\tilde{\mathsf{C}}_\tau$ is determined by

$$\begin{aligned}\tilde{\mathsf{C}}_\tau = \mathsf{Y}^\dagger \mathsf{C}_\tau \mathsf{Y} &= \int_0^\tau e^{\Lambda t} \mathsf{Y}^\dagger \mathsf{F}\mathsf{F}^T \mathsf{Y} e^{\Lambda^\dagger t}\, dt \\ &= \left(\mathsf{Y}^\dagger \mathsf{F}\mathsf{F}^T \mathsf{Y}\right) \circ \mathsf{E}_\tau\,,\end{aligned} \tag{6}$$

where the notation $\circ$ denotes Hadamard product[1], and the entries of the positive semidefinite matrix E_τ are

$$[\mathsf{E}_\tau]_{kl} = \int_0^\tau e^{(\lambda_k(\mathsf{A}) + \bar{\lambda}_l(\mathsf{A}))t}\, dt = \frac{e^{(\lambda_k(\mathsf{A}) + \bar{\lambda}_l(\mathsf{A}))\tau} - 1}{\lambda_k(\mathsf{A}) + \bar{\lambda}_l(\mathsf{A})}\,. \tag{7}$$

The matrix E_τ depends only on the eigenvalues of the dynamics and the lead-time τ. The projection P of the forcing matrix F onto the adjoint eigenvectors Y is defined by $\mathsf{P} \equiv \mathsf{F}^T \mathsf{Y}$. The representation $\mathsf{C}_\tau = \mathsf{Z}\left(\mathsf{P}^\dagger \mathsf{P} \circ \mathsf{E}_\tau\right)\mathsf{Z}^\dagger$ has the value of showing explicitly the dependence of the prediction error covariance on the eigendecomposition of the dynamics, as well as on the projection P of the forcing matrix onto the adjoint eigenvectors.

The analysis of the prediction error covariance is particularly clear when the stochastic forcing is uncorrelated in normal-mode space. In this case, the forcing covariance can be written $\mathsf{F}\mathsf{F}^T = \mathsf{Z}\mathsf{D}\mathsf{Z}^\dagger$ where the matrix $\mathsf{D} = \mathsf{P}^\dagger \mathsf{P}$ is diagonal and gives the amplitudes of the forcing. This characterization of the forcing is invariant with respect to linear transformations of the state-variable. When the stochastic forcing is uncorrelated in normal-mode space, the adjoint eigenvec-

[1]The Hadamard product of two matrices X and Y with entries X_{kl} and Y_{kl}, respectively, is the matrix whose entries are $\mathsf{X}_{kl}\mathsf{Y}_{kl}$.

tors of the dynamics diagonalize the prediction error covariance at all lead times, i.e., eigenmodes are uncorrelated, and $\mathbf{C}_\tau = \mathbf{Z}(\mathbf{D} \circ \mathbf{E}_\tau)\mathbf{Z}^\dagger$; $\mathbf{D} \circ \mathbf{E}_\tau$ is a diagonal matrix whose k-th diagonal entry is $\mathbf{D}_{kk}[\mathbf{E}_\tau]_{kk}$. Although this representation separates temporal and spatial structures of the prediction error covariance, it is not the eigenvalue decomposition since the matrix of eigenvectors $\mathbf{Z}$ is only orthogonal when the dynamics matrix $\mathbf{A}$ is normal. A lower bound for the total prediction error variance in the case of normal-mode uncorrelated stochastic forcing is

$$\begin{aligned} \operatorname{tr} \mathbf{C}_\tau &= \sum_{k=1}^{n} \frac{e^{2\operatorname{Re}\lambda_k(\mathbf{A})\tau} - 1}{2\operatorname{Re}\lambda_k(\mathbf{A})} \mathbf{D}_{kk} \mathbf{z}_k^\dagger \mathbf{z}_k \\ &= \sum_{k=1}^{n} \frac{e^{2\operatorname{Re}\lambda_k(\mathbf{A})\tau} - 1}{2\operatorname{Re}\lambda_k(\mathbf{A})} \mathbf{D}_{kk} \sec^2\theta_k \\ &\geq \sum_{k=1}^{n} \frac{e^{2\operatorname{Re}\lambda_k(\mathbf{A})\tau} - 1}{2\operatorname{Re}\lambda_k(\mathbf{A})} \mathbf{D}_{kk} = \operatorname{tr}(\mathbf{E}_\tau \circ \mathbf{D}) \, , \end{aligned} \tag{8}$$

where θ_k is the angle between the eigenvector $\mathbf{z}_k$ and the adjoint eigenvector $\mathbf{y}_k$.[2] The prediction error variance is controlled by the proximity of the eigenvalues of the dynamics in the complex plane to the imaginary axis, the angle between eigenvectors and adjoint eigenvectors, and the size of the stochastic forcing.

Although the expression in (8) for the prediction error variance applies only to normal-mode uncorrelated stochastic forcing, it provides some useful insights. First, when the dynamics matrix $\mathbf{A}$ is normal and the stochastic forcing is uncorrelated in normal-mode space, the eigenvalues of the prediction error covariance are the elements of the diagonal matrix $\mathbf{D} \circ \mathbf{E}_\tau$, and the inequality in (8) is an equality. Therefore (8) extends the result that nonnormality increases variance, proven by Ioannou (1995) for unitary stochastic forcing, to normal-mode uncorrelated stochastic forcing. Second, the expression in (8) illustrates how a linear change of state-variable affects prediction error variance. A linear change of state-variable does not change the eigenvalues of the dynamics or the forcing amplitudes $\mathbf{D}_{kk}$ but being equivalent to a change of inner product can change the

[2] Recall that $\sec\theta_k = \|\mathbf{z}_k\|\|\mathbf{y}_k\|/\mathbf{y}_k^\dagger\mathbf{z}_k$ and $\|\mathbf{y}_k\| = \mathbf{y}_k^\dagger\mathbf{z}_k = 1$.

angles between eigenvectors and adjoint eigenvectors and consequently change the prediction error variance.

The example of normal-mode uncorrelated stochastic forcing highlights the need to normalize or re-scale the prediction error; prediction error variance can be increased by simply increasing the stochastic forcing amplitude. Scalar normalization have proved useful in other contexts, such as the identification of forcing structures that most efficiently excite error growth (Farrell and Ioannou, 1996; Kleeman and Moore, 1997; Tippett and Cohn, 2001). Here, for the special case of normal-mode uncorrelated stochastic forcing, evolution of the prediction error of each eigenmode is independent, and the univariate approach of measuring predictability by the ratio of prediction error and climatological variances can be applied to each eigenmode. Normalizing the prediction error variance of each eigenmode by its climatological variance removes both the forcing-amplitude and nonnormal factors. The resulting predictability measure can also be used to compare the predictability of different eigenmodes. This approach to measuring predictability is extended to general stochastic forcing in the following section.

3. Predictability

A prediction is useful on time-scales where the prediction error is, in some sense, less than the climatological variability. Therefore, the utility of a prediction at lead-time τ depends on both the prediction error covariance $\mathbf{C}_\tau$ and the climatological error covariance $\mathbf{C}_\infty$. This notion of predictability is the basis for predictability measures defined using the *predictive information matrix* $\mathbf{G}_\tau \equiv \mathbf{C}_\tau \mathbf{C}_\infty^{-1}$ (Schneider and Griffies, 1999).[3] The eigenvalues of the predictive information matrix measure the relative error variance of a set of state-space patterns chosen to optimize relative error

[3]Here, the climatological covariance matrix is invertible, i.e., there are no perfectly predictable components, if the pair $(\mathbf{A}, \mathbf{F})$ is controllable; a sufficient condition for controllability is that forcing covariance $\mathbf{F}\mathbf{F}^T$ be invertible. When the climatological covariance matrix is computed from data assumed to be Gaussian distributed, high dimensional state-space and short data records make it singular. Regularization methods, such as projecting the climatological covariance matrix onto truncated set of EOFs, are required and can significantly limit the number of predictable components that can robustly estimated (Schneider and Griffies, 1999; Schneider and Held, 2001).

variance and so can provide a multivariate generalization of the univariate relative error variance s_τ^2/s_∞^2 and the associated univariate predictability measure, $1 - s_\tau^2/s_\infty^2$; s_τ^2 is the prediction error variance at lead-time τ and s_∞^2 is the climatological variance.

The eigenvalue decomposition of the predictive information $\mathbf{G}_\tau$ decomposes phase-space into uncorrelated patterns ordered by their relative prediction error variance (Schneider and Griffies, 1999). When λ is an eigenvalue of $\mathbf{G}_\tau$, its adjoint eigenvector $\mathbf{q}$ satisfies $\mathbf{q}^T\mathbf{G}_\tau = \lambda\mathbf{q}^T$. The eigenvalues of $\mathbf{G}_\tau$ are between zero and unity since

$$\lambda = \frac{\mathbf{q}^T\mathbf{C}_\tau\mathbf{q}}{\mathbf{q}^T\mathbf{C}_\infty\mathbf{q}}\,, \tag{9}$$

and $\mathbf{q}^T\mathbf{C}_\infty\mathbf{q} \geq \mathbf{q}^T\mathbf{C}_\tau\mathbf{q} \geq 0$; $\mathbf{q}^T(\mathbf{C}_\infty - \mathbf{C}_\tau)\mathbf{q} \geq 0$ follows from (5). The eigenvalues of $\mathbf{G}_\tau$ behave like relative error, initially zero because of the perfect initial condition assumption, and increasing with lead-time until they reach unity in the limit of large lead-time. For any direction $\mathbf{q}$ in state-space, the projection of the state-variable $\mathbf{w}$ in the direction $\mathbf{q}$ is $(\mathbf{q}^T\mathbf{w})$ and has relative error given by the quantity $(\mathbf{q}^T\mathbf{C}_\tau\mathbf{q})/(\mathbf{q}^T\mathbf{C}_\infty\mathbf{q})$. The direction $\mathbf{q}_n$ that minimizes the relative error in (9) is the n-th adjoint eigenvector of the predictive information matrix and defines the first *predictable component* $(\mathbf{q}_n^T\mathbf{w})$ and its relative error variance $\lambda_n(\mathbf{G}_\tau)$ (Schneider and Griffies, 1999). The eigenvector $\mathbf{p}_n$ corresponding to the adjoint eigenvector $\mathbf{q}_n$ is the first *predictable pattern* (Schneider and Griffies, 1999). The second predictable component is determined by the next smallest eigenvalue of $\mathbf{G}_\tau$ and is temporally uncorrelated with the first predictable component.

The eigenvalues of the predictive information matrix are invariant with respect to linear transformations of the state-variable. If a new state-variable $\hat{\mathbf{w}} \equiv \mathbf{L}\mathbf{w}$ and its prediction error covariance $\hat{\mathbf{C}}_\tau$ are defined, the new predictive information matrix $\hat{\mathbf{G}}_\tau$ is related to $\mathbf{G}_\tau$ by a similarity transformation

$$\hat{\mathbf{G}}_\tau = \hat{\mathbf{C}}_\tau\hat{\mathbf{C}}_\infty^{-1} = \mathbf{L}\mathbf{C}_\tau\mathbf{C}_\infty^{-1}\mathbf{L}^{-1} = \mathbf{L}\mathbf{G}_\tau\mathbf{L}^{-1}\,, \tag{10}$$

so that $\mathbf{G}_\tau$ and $\hat{\mathbf{G}}_\tau$ have the same eigenvalues. Therefore, predictability measures defined by eigenvalues of the predictive information matrix are invariant with respect to linear transformations of the state-variable and are norm-independent. The eigenvectors of the predictive information matrix transform in the same manner as the state-variable; if $\mathbf{p}$ is an eigenvector of $\mathbf{G}_\tau$ then $\mathbf{L}\mathbf{p}$ is an eigenvector of $\hat{\mathbf{G}}_\tau$.

The predictable patterns and their relative error variances are simply related to the dynamics when the stochastic forcing is uncorrelated in normal-mode space, i.e., when the forcing covariance can be represented $\mathbf{F}\mathbf{F}^T = \mathbf{Z}\mathbf{D}\mathbf{Z}^\dagger$ and $\mathbf{D}$ is diagonal. We require that $\mathbf{D}$ be invertible to ensure that the climatological covariance $\mathbf{C}_\infty$ is also invertible. In this case, the predictive information matrix $\mathbf{G}_\tau$ has the simple form

$$\begin{aligned}\mathbf{G}_\tau &= \mathbf{Z}(\mathbf{D}\circ\mathbf{E}_\tau)\mathbf{Z}^\dagger\left(\mathbf{Z}^\dagger\right)^{-1}(\mathbf{D}\circ\mathbf{E}_\infty)^{-1}\mathbf{Z}^{-1} \\ &= \mathbf{Z}\,\mathrm{diag}(\mathbf{E}_\tau)\,\mathrm{diag}(\mathbf{E}_\infty)^{-1}\mathbf{Z}^{-1}\,,\end{aligned} \tag{11}$$

and is remarkably independent of the forcing coefficient $\mathbf{D}$; $\mathrm{diag}(\mathbf{E}_\tau)$ is the diagonal matrix whose entries are the diagonal elements of $\mathbf{E}_\tau$. Since (11) is the eigendecomposition of $\mathbf{G}_\tau$, the eigenvalues of $\mathbf{G}_\tau$ are

$$\lambda_k(\mathbf{G}_\tau) = 1 - e^{2\,\mathrm{Re}\,\lambda_{n-k+1}(\mathbf{A})\tau}\,, \tag{12}$$

and depend only on the real part of the eigenvalues of the dynamics; the eigenvectors of $\mathbf{G}_\tau$ are the eigenvectors of the dynamics. Therefore, the leading eigenvector $\mathbf{z}_1$ of the dynamics is the first predictable pattern at all lead-times of a system with normal-mode uncorrelated stochastic forcing; the relative error of the associated first predictable component is $1 - e^{2\,\mathrm{Re}\,\lambda_1(\mathbf{A})}\tau$. The predictability of the first predictable component decreases exponentially with decay-rate determined by the least damped eigenmode of the dynamics.

For general stochastic forcing, it is useful to write the prediction error covariance $\mathbf{C}_\tau$ as

$$\begin{aligned}\mathbf{C}_\tau &= \int_0^\infty e^{t\mathbf{A}}\mathbf{F}\mathbf{F}^T e^{t\mathbf{A}^T}\,dt - \int_\tau^\infty e^{t\mathbf{A}}\mathbf{F}\mathbf{F}^T e^{t\mathbf{A}^T}\,dt\,, \\ &= \mathbf{C}_\infty - e^{\tau\mathbf{A}}\mathbf{C}_\infty e^{\tau\mathbf{A}^T}\,,\end{aligned} \tag{13}$$

so that

$$\mathbf{G}_\tau = \mathbf{I} - e^{\mathbf{A}\tau}\mathbf{C}_\infty e^{\mathbf{A}^T\tau}\mathbf{C}_\infty^{-1}\,. \tag{14}$$

The eigenvalues of $\mathbf{G}_\tau$ are then

$$\begin{aligned}\lambda_k(\mathbf{G}_\tau) &= \lambda_k(\mathbf{I} - \mathbf{W}_\tau\mathbf{W}_\tau^T)\\ &= 1 - \sigma^2_{n-k+1}(\mathbf{W}_\tau)\,,\end{aligned} \tag{15}$$

where we define $\mathbf{W}_\tau \equiv e^{\hat{\mathbf{A}}\tau}$ and $\hat{\mathbf{A}} \equiv \mathbf{C}_\infty^{-1/2}\mathbf{A}\mathbf{C}_\infty^{1/2}$; the notation $\sigma_k(\mathbf{X})$ denotes the k-th singular value of a matrix $\mathbf{X}$ ordered so that $\sigma_1(\mathbf{X}) \geq \sigma_2(\mathbf{X}) \geq \ldots \sigma_n(\mathbf{X}) \geq 0$. The matrix $\hat{\mathbf{A}}$ is the dynamics matrix of the whitened state-variable $\hat{\mathrm{w}} \equiv \mathbf{C}_\infty^{-1/2}\mathrm{w}$. The climatological covariance $\hat{\mathbf{C}}_\infty$ is the identity matrix in the whitened state-variable; $\mathbf{W}_\tau$ is the state-propagator of the whitened state-variable. The eigendecomposition of $\mathbf{G}_\tau$ is determined by the singular value decomposition of $\mathbf{W}_\tau$. Another way of arriving at (15) is to recall that the eigenvalues of $\mathbf{G}_\tau$ are invariant under linear transformations of the state-variable and note that in the whitened state-variable

$$\hat{\mathbf{G}}_\tau = \hat{\mathbf{C}}_\tau = \mathbf{I} - \mathbf{W}_\tau\mathbf{W}_\tau^T\,. \tag{16}$$

Predictability analysis is equivalent to principle component analysis of the whitened state-vector $\hat{\mathrm{w}}$ (Schneider and Griffies, 1999). The appropriate choice of vector-norm or state-variable for questions of predictability is one that makes the climatological covariance the identity matrix.

The whitened dynamics $\hat{\mathbf{A}}$ is normal when the stochastic forcing is uncorrelated in normal-mode space. In this case, the whitened-state propagator $\mathbf{W}_\tau$ is also normal, $\sigma_k^2(\mathbf{W}_\tau) = |\lambda_k(\mathbf{W}_\tau)|^2 = e^{2\,\mathrm{Re}\,\lambda_k(\mathbf{A})\tau}$, and the expression in (12) for the eigenvalues of the predictive information matrix is a consequence. When the stochastic forcing is correlated in normal-mode space, $\mathbf{W}_\tau$ is nonnormal and its singular values are not determined by its eigenvalues. However, the eigenvalues and singular values of any matrix must satisfy certain inequalities. For instance, the largest singular value must be larger than the modulus of the largest eigenvalue. That is to say, $\sigma_1^2(\mathbf{W}_\tau) \geq |\lambda_1(\mathbf{W}_\tau)|^2$, which combined with (15) gives the upper bound

$$\lambda_n(\mathbf{G}_\tau) \leq 1 - e^{2\,\mathrm{Re}\,\lambda_1(\mathbf{A})\tau}\,, \tag{17}$$

for the smallest eigenvalue of the predictive information matrix. The left-hand side of (17) is the relative error of the first predictable component of a system with general stochastic forcing, and the right-hand side is the relative error of the first predictable component of a system with normal-mode uncorrelated stochastic forcing. Therefore the inequality in (17) means that the relative error of the first predictable component is maximized when the stochastic forcing is uncorrelated in normal-mode space. In terms of predictability, (17) says that the predictability of the first predictable component of a system is minimized when the stochastic forcing is uncorrelated in normal-mode space.

A more physical interpretation of this result is that when the stochastic forcing is correlated in normal-mode space, normal modes are correlated, and linear combinations of normal modes can be constructed whose predictability is greater than that of any single normal mode. This potential for constructive combination of normal modes is the same mechanism that allows systems with nonnormal dynamics and unitary stochastic forcing to present *smaller* prediction error than equivalent normal systems. Equation (7) of Ioannou (1995) shows that the smallest singular value of $\mathbf{W}_\tau = e^{\hat{\mathbf{A}}\tau}$ is maximized when $\mathbf{W}_\tau$ is normal. In the present context, this fact means that the error of the first predictable component is maximized for stochastic forcing that is uncorrelated in normal-mode space.

The same conclusion that normal-mode uncorrelated stochastic forcing minimizes predictability is true for general predictability measures defined abstractly by

$$\text{predictability} \equiv \sum_{k=1}^{n} h(1 - \lambda_k(\mathbf{G}_\tau))\,, \tag{18}$$

where the function h has the property that $h(e^x)$ is a convex and increasing function of x. Theorem

1 of the Appendix and (15) give that

$$\text{predictability} = \sum_{k=1}^{n} h(\sigma_k^2(\mathbf{W}_\tau)) \geq \sum_{k=1}^{n} h(|\lambda_k(\mathbf{W}_\tau)|^2) = \sum_{k=1}^{n} h\left(e^{2\,\mathrm{Re}\,\lambda_k(\mathbf{A})}\right) . \tag{19}$$

The right-hand side of (19) is the predictability of the system with normal-mode uncorrelated stochastic forcing and depends only on the real part of the eigenvalues of the dynamics. General predictability measures are minimized when the stochastic forcing is uncorrelated in normal-mode space.

Predictability measures that can be written in the form of (18) are: the quantity

$$\begin{aligned} 1 - \mathrm{tr}\,\mathbf{G}_\tau/n &= \frac{1}{n}\sum_{k=1}^{n} \sigma_k^2\left(\mathbf{W}_\tau\right) \\ &\geq \frac{1}{n}\sum_{k=1}^{n} e^{2\,\mathrm{Re}\,\lambda_k(\mathbf{A})} , \end{aligned} \tag{20}$$

the *predictive information* R_τ, defined by (Schneider and Griffies, 1999)

$$\begin{aligned} R_\tau \equiv -\frac{n}{2}\log\det\mathbf{G}_\tau &= -\frac{n}{2}\sum_{k=1}^{n}\log\left(1-\sigma_k^2(\mathbf{W}_\tau)\right) \\ &\geq -\frac{n}{2}\sum_{k=1}^{n}\log\left(1-e^{2\,\mathrm{Re}\,\lambda_k(\mathbf{A})}\right) , \end{aligned} \tag{21}$$

and *relative entropy* r_τ defined by (Kleeman, 2002)

$$\begin{aligned} r_\tau &\equiv \frac{1}{2}\left[-\log\left(\det\mathbf{G}_\tau\right) + \mathrm{tr}\left(\mathbf{G}_\tau\right) - n\right] \\ &= -\frac{1}{2}\sum_{k=1}^{n}\log\left(1-\sigma_k^2(\mathbf{W}_\tau)\right) + |\sigma_k(\mathbf{W}_\tau)|^2 \\ &\geq -\frac{1}{2}\sum_{k=1}^{n}\log\left(1-e^{2\,\mathrm{Re}\,\lambda_k(\mathbf{A})}\right) + e^{2\,\mathrm{Re}\,\lambda_k(\mathbf{A})} . \end{aligned} \tag{22}$$

Normal-mode uncorrelated stochastic forcing minimizes all these measures of predictability and gives lower bounds for predictability that depend only on the eigenvalues of the dynamics (see Lemmas 1 and 2 of the Appendix). These inequalities show that normal-mode analysis of the predictability problem gives lower bounds for the system with general forcing structure. The inequalities in (20) – (22) are also valid when the sums are truncated, and predictability is measured using the trailing eigenvalues of the predictive information matrix.

These results have an interpretation in the framework where the forcing is fixed as unitary, i.e., when $\mathbf{F}\mathbf{F}^T = \mathbf{I}$. In this case, the stochastic forcing is uncorrelated in normal-mode space when $\mathbf{Y}^\dagger\mathbf{Y}$ is diagonal. When $\mathbf{Y}^\dagger\mathbf{Y}$ is diagonal, $\mathbf{Y}^\dagger\mathbf{Y} = \mathbf{I}$ and the dynamics is normal since the columns of $\mathbf{Y}$ are unit vectors. Therefore, for unitary stochastic forcing, predictability is minimized when the dynamics is normal. Nonnormal dynamics is more predictable than normal dynamics with the same eigenvalues for unitary stochastic forcing. Ioannou (1995) demonstrated in this setting that nonnormality increases prediction error, and here we have shown that it increases predictability as well.

4. Upper bounds for predictability

Our analysis shows that the predictability of a system described by linear stochastic dynamics is minimized when the stochastic forcing is uncorrelated in normal-mode space. However, some questions remain unanswered. For instance, what stochastic forcing produces maximum predictability, and when is this maximum predictability strictly larger than that given by normal-mode uncorrelated forcing? Such questions are difficult to answer generally because predictability is a nonlinear function of the stochastic forcing. We explore these issues in a simple 2-D linear stochastic system. The relevance of this type of simple model to real climate systems is discussed in Chang et al. (2002) where a 2-D stochastically driven damped inertial oscillator is used as a prototype coupled system to elucidate the importance of nonnormal growth in enhancing the predictability of climate systems.

Consider the 2×2 dynamics matrix given by

$$\mathbf{A} = \begin{bmatrix} \lambda_1 + i\phi & 0 \\ 0 & \lambda_2 - i\phi \end{bmatrix} , \tag{23}$$

where λ_1 and λ_2 are negative real constants; ϕ is a real constant and $i = \sqrt{-1}$. We assume that the two eigenvalues of $\mathbf{A}$ are either both real or complex conjugates. There is no loss of generality

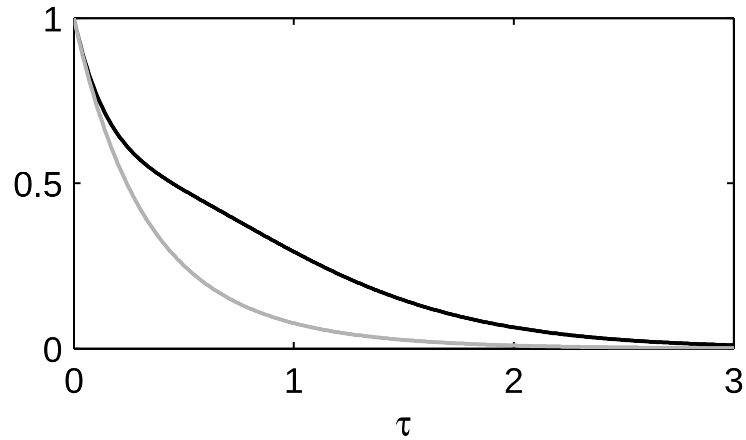

Figure 1. The predictability measure $1 - \mathrm{tr}\,\mathbf{G}_\tau/2$ for normal-mode uncorrelated forcing (gray line) and rank-1 forcing (black line) with $\lambda_1 = -1$ and $\lambda_2 = -2$.

in taking the dynamics to be diagonal since our predictability measures are invariant under linear transformation of the state variable. Normal-mode uncorrelated stochastic forcing for diagonal dynamics corresponds to $\mathbf{F}\mathbf{F}^T$ being diagonal. In this case, the predictive information matrix $\mathbf{G}_\tau$ is diagonal and given by

$$\mathbf{G}_\tau = \begin{bmatrix} 1 - e^{2\lambda_1\tau} & 0 \\ 0 & 1 - e^{2\lambda_2\tau} \end{bmatrix}\,, \tag{24}$$

and can be used to compute the minimum predictability of the system. When the eigenvalues of the dynamics are identical, $\lambda \equiv \lambda_1 = \lambda_2$ and $\phi = 0$, the dynamics is essentially scalar and

$$\mathbf{C}_\tau = \frac{1 - e^{2\lambda\tau}}{2\lambda}\mathbf{F}\mathbf{F}^T\,, \tag{25}$$

for general stochastic forcing. In this case, the climatological covariance $\mathbf{C}_\infty = \mathbf{F}\mathbf{F}^T/(2\lambda)$ is invertible when $\mathbf{F}\mathbf{F}^T$ is, and the predictive information matrix is independent of the forcing and given by (24) with $\lambda_1 = \lambda_2$. Therefore no forcing structure increases predictability when the eigenvalues of the dynamics are identical, a result that is true in general.

Suppose that the forcing matrix is rank-1 so that the stochastic forcing covariance can be written $\mathbf{F}\mathbf{F}^T = \mathbf{f}\mathbf{f}^T$ where $\mathbf{f}$ is a vector with nonzero elements. A general result that follows directly from

(6) is that

$$\mathbf{C}_\tau = \mathbf{f}\mathbf{f}^T \circ \mathbf{E}_\tau = \mathrm{Diag}(\mathbf{f})\ \mathbf{E}_\tau\, \mathrm{Diag}(\mathbf{f})\,, \tag{26}$$

when the forcing matrix $\mathbf{F}$ is rank-1 and the dynamics are diagonal; the notation $\mathrm{Diag}(\mathbf{f})$ denotes the diagonal matrix whose diagonal entries are the elements of the vector $\mathbf{f}$. In general, the invertibility of the climatological covariance $\mathbf{C}_\infty$ is not guaranteed for rank-1 forcing. The climatological covariance $\mathbf{C}_\infty$ is invertible for 2×2 diagonal dynamics when the entries of $\mathbf{f}$ are nonzero and the eigenvalues of the dynamics are distinct. For rank-1 forcing, the eigenvalues of $\mathbf{G}_\tau$ are given by $\lambda_k(\mathbf{G}_\tau) = \lambda_k(\mathbf{E}_\tau \mathbf{E}_\infty^{-1})$ and remarkably, are independent of the forcing $\mathbf{f}$. A direct calculation gives that the predictability measure (20) of the system with rank-1 forcing is

$$1 - \frac{1}{2}\,\mathrm{tr}\,\mathbf{G}_\tau = \frac{1}{2}\left(e^{2\,\lambda_1\,t} + e^{2\,\lambda_2\,t}\right) + \frac{2\lambda_1\lambda_2\left(\left(e^{\lambda_1\tau} - e^{\lambda_2\tau}\right)^2 + 2e^{(\lambda_1+\lambda_2)\tau}\left(1 - \cos 2\phi\right)\right)}{4\phi^2 + (\lambda_1 - \lambda_2)^2}\,. \tag{27}$$

The result in (27) is valid when the eigenvalues of the dynamics are distinct. The first term on the right-hand side of (27) is the minimum predictability of the system, and the second term is strictly positive. Therefore, rank-1 forcing gives predictability that is *strictly* larger than normal-mode uncorrelated forcing. Figure 1 compares the quantity $1 - \mathrm{tr}\,\mathbf{G}_\tau/2$ for normal-mode uncorrelated and for rank-1 forcing as function of lead-time τ with $\lambda_1 = -1$ and $\lambda_2 = -2$. In this example, rank-1 stochastic forcing is seen to enhance predictability on time-scales comparable to the system e-folding time. Rank-1 forcing also increases predictability as measured by the predictive information R_τ and the relative entropy r_τ. In fact, there is no rank-2 forcing that gives more predictability. The optimization problems of maximizing $1 - \mathrm{tr}\,\mathbf{G}_\tau/2$, R_τ and r_τ can be solved in closed form for the 2×2 system and show that rank-1 forcing maximizes predictability. Whether rank-1, or perhaps approximately rank-1, forcing gives maximum predictability in general remains an open question. However, limited numerical experiments do not contradict this conjecture.

5. Discussion

In this work we have considered prediction error growth and predictability in linear stochastic systems with perfect initial conditions and state-independent stochastic forcing. We have focused our study on the impact of stochastic forcing spatial-structure on prediction error growth and predictability. Two classes of forcing were considered: forcing that is uncorrelated in normal-mode space and forcing with arbitrary spatial-structure. The analysis of systems with normal-mode uncorrelated stochastic forcing is simple and complete since normal modes in such systems are uncorrelated and evolve independently. Prediction error growth depends on details of the forcing, even for the special case of normal-mode uncorrelated forcing where prediction error depends on the forcing amplitudes. Appropriately chosen stochastic forcing with general spatial structures can produce prediction error growth more efficiently or less efficiently than normal-mode uncorrelated stochastic forcing (Ioannou, 1995). The norm-dependence of prediction error growth further complicates its analysis.

Predictability can be defined in a norm-invariant manner using eigenvalues of the *predictive information matrix* (Schneider and Griffies, 1999). We consider a family of predictability measures that includes *predictive information* and *relative entropy* (Schneider and Griffies, 1999; Kleeman, 2002). Remarkably, the predictability of a system with normal-mode uncorrelated stochastic forcing is independent of the forcing amplitudes and is expressed simply in terms of the eigenvalues of the dynamics. Moreover, the predictable patterns are simply the eigenmodes of the dynamics. Predictability is more difficult to analyze in systems where the stochastic forcing has general spatial structure. However, in contrast to the results of prediction error growth analysis where general stochastic forcing can excite error growth more or less efficiently then normal-mode uncorrelated forcing, general stochastic forcing always increases predictability; normal-mode uncorrelated forcing gives minimum predictability. Therefore normal-mode analysis gives lower bounds for the predictability of systems with general stochastic forcing, and these bounds are simple functions of the eigenvalues of the dynamics.

These results can also be interpreted in the framework where eigenvectors of the dynamics are allowed to vary and the stochastic forcing is fixed as unitary. Prediction error growth has been studied extensively in this framework (Farrell and Ioannou, 1996). Unitary stochastic forcing is uncorrelated in normal-mode space when the dynamics is normal. Therefore, for unitary stochastic forcing and fixed eigenvalues, a system with nonnormal dynamics is more predictable than a system with normal dynamics.

Our results are in the form of lower bounds for predictability. Since the lower bound for predictability depends only the real part of the eigenvalues of the dynamics, predictability does not depend on the oscillatory behavior of the system when the stochastic forcing is uncorrelated in normal-mode space. A topic for future research is the formulation of *upper* bounds for predictability. We expect such upper bounds to depend only on the eigenvalues of the dynamics, as in the simple example. Another issue is the characterization of forcing structures that give maximum predictability as in Chang et al. (2002) where the forcing producing maximum predictability for norm-dependent error variance predictability measures was found.

6. Acknowledgement

This manuscript benefited from the insightful comments and suggestions of two anonymous reviewers. The authors heartily thank Tim DelSole (COLA), Steve Cohn (NASA/DAO) and Tapio Schneider (NYU/CIMS) for their comments, suggestions and corrections. IRI is supported by its sponsors and NOAA Office of Global Programs Grant number NA07GP0213. PC is grateful to IRI for its hospitality and generous support during his sabbatical leave when this research was initiated. PC also acknowledges the supports by the NSF grant ATM-99007625 and the NOAA grant NA16GP1572.

Appendix A

The necessary and sufficient condition relating the eigenvalues $\lambda_k(\mathbf{W})$ and singular values $\sigma_k(\mathbf{W})$ of an invertable matrix $\mathbf{W}$ is the inequality

$$\sum_{k=1}^{j} \log \sigma_k(\mathbf{W}) \geq \sum_{k=1}^{j} \log |\lambda_k(\mathbf{W})| \,, \tag{A.1}$$

for $j = 1, 2, \ldots, n$ with equality for $j = n$ (Marshall and Olkin, 1979). The sequence $\{\log \sigma_k(\mathbf{W})\}$ is said to *majorize* the sequence $\{\log |\lambda_k(\mathbf{W})|\}$. The following classical result of Weyl identifies functions that preserve majorization.

Theorem 1 (Weyl (1949); Chapter 5, A.2.a of Marshall and Olkin (1979)**).** *Suppose* $\mathbf{W}$ *is an invertible* $n \times n$ *matrix with eigenvalues* $\lambda_k(\mathbf{W})$ *and singular values* $\sigma_k(\mathbf{W})$*. If* $h(e^x)$ *is an increasing convex function then,*

$$\sum_{k=1}^{j} h(\sigma_k(\mathbf{W})) \geq \sum_{k=1}^{j} h(|\lambda_k(\mathbf{W})|) \,, \tag{A.2}$$

for $j = 1, 2, \ldots, n$*.*

The following lemmas are consequences of Theorem 1 with specific choices of the function h. Lemma 1 follows from taking $h(x) = x^s$ and from $h(e^x) = e^{sx}$ being an increasing convex function for $s > 0$.

Lemma 1. *Suppose* $\mathbf{W}$ *is an* $n \times n$ *matrix with with eigenvalues* $\lambda_k(\mathbf{W})$ *and singular values* $\sigma_k(\mathbf{W})$*. For* $s > 0$*,*

$$\sum_{k=1}^{j} \sigma_k^s(\mathbf{W}) \geq \sum_{k=1}^{j} |\lambda_k(\mathbf{W})|^s \,, \quad j = 1, 2, \ldots, n \,. \tag{A.3}$$

The function $h(e^x)$ is a convex increasing function for $-\infty \leq x \leq 0$ when h is an increasing convex function on the interval $(0, 1)$. The following lemma is a consequence of $-\log(1 - x^2)$ and $-(\log(1 - x^2) + x^2)$ being convex increasing functions on the interval $(0, 1)$.

Lemma 2. *Under the assumptions of Theorem 1, and* $\sigma_1(\mathbf{W}) < 1$

$$-\sum_{k=1}^{j} \log\left(1 - |\sigma_k(\mathbf{W})|^2\right) \geq -\sum_{k=1}^{k} \log\left(1 - \lambda_k^2(\mathbf{W})\right) \ , \tag{A.4}$$

and

$$-\sum_{k=1}^{j} \log\left(1 - |\lambda_k(\mathbf{W})|^2\right) + |\lambda_k(\mathbf{W})|^2 \geq -\sum_{k=1}^{j} \log\left(1 - \sigma_k^2(\mathbf{W})\right) + |\sigma_k(\mathbf{W})|^2 \, , \tag{A.5}$$

for $j = 1, 2, \ldots, n$.

References

Chang, P., Saravanan, R., DelSole, T., and Wang, F. 2002. Predictability of Linear Coupled Systems. Part I: Theoretical Analyses. *J. Climate*. submitted.

DelSole, T. and Farrell, B. F. 1995. A Stochastically Excited Linear System as a Model for Quasigeostrophic Turbulence: Analytic Results for One- and Two-layer Fluids. *J. Atmos. Sci.*, **52**, 2531–2547.

Farrell, B. F. and Ioannou, P. J. 1996. Generalized Stability Theory. Part I: Autonomous Operators. *J. Atmos. Sci.*, **53**, 2025–2040.

Hasselmann, K. 1976. Stochastic climate models I. Theory. *Tellus*, **28**, 473–485.

Hasselmann, K. 1988. PIPs and POPs–A general formalism for the reduction of dynamical systems is terms of Principal Interaction Patterns and Principal Oscillation Patterns. *J. Geophys. Res.*, **93**, 11 015–11 021.

Ioannou, P. J. 1995. Nonnormality increases variance. *J. Atmos. Sci*, **52**, 1155–1158.

Kleeman, R. 2002. Measuring Dynamical Prediction Utility Using Relative Entropy. *J. Atmos. Sci.*, **59**, 2057–2072.

Kleeman, R. and Moore, A. M. 1997. A Theory for the Limitation of ENSO Predictability Due to Stochastic Atmospheric Transients. *J. Atmos. Sci.*, **54**, 753–767.

Lorenz, E. N. 1969. The predictability of a flow which possesses many scales of motion. *Tellus*, **21**, 289–307.

Marshall, A. W. and Olkin, I. *Inequalities: Theory of Majorization and its applications*, volume 143 of *Mathematics in Sciences and Engineering*. Academic Press, New York, New York, 1979.

Penland, C. 1989. Random forcing and forecasting using principal oscillation pattern analysis. *Mon. Wea. Rev.*, **117**, 2165–2185.

Penland, C. and Matrosova, L. 1994. A Balance Condition for Stochastic Numerical Models with Application to the El Niño-Southern Oscillation. *J. Climate*, **7**, 1352–1372.

Penland, C. and Matrosova, L. 1998. Prediction of Tropical Atlantic Sea Surface Temperatures Using Linear Inverse Modeling. *J. Climate*, **11**, 483–496.

Penland, C. and Matrosova, L. 2001. Expected and actual errors of linear inverse model forecasts. *Mon. Wea. Rev.*, **129**, 1740–1745.

Penland, C. and Sardeshmukh, P. D. 1995. The optimal-growth of tropical sea-surface temperature anomalies. *J. Climate*, **8**, 1999–2024.

Schneider, T. and Griffies, S. 1999. A Conceptual Framework for Predictability Studies. *J. Climate*, **12**, 3133–3155.

Schneider, T. and Held, I. M. 2001. Discriminants of twentieth-century changes in earth surface temperatures. *J. Climate*, **14**, 249–254.

Tippett, M. K. and Cohn, S. E. 2001. Adjoints and low-rank covariance representation. *Nonlinear Processes in Geophysics*, **8**, 331–340.

Weyl, H. 1949. Inequalities between two kinds of eigenvalues of a linear transformation. *Proc. Nat. Acad. Sci U.S.A.*, **35**, 408–411.

Whitaker, J. S. and Sardeshmukh, P. D. 1998. A Linear Theory of Extratropical Synoptic Eddy Statistics. *J. Atmos. Sci.*, **55**, 237–258.

Winkler, C. R., Newman, M., and Sardeshmukh, P. D. 2001. A Linear Model of Wintertime Low-Frequency Variability. Part I: Formulation and Forecast Skill. *J. Climate*, **14**, 4474–4494.