Ensemble Oscillation Correction (EnOC): Leveraging Oscillatory Modes to Improve Forecasts of Chaotic Systems

Eviatar Bach,[a,b,c] Safa Mote,[a,b] V. Krishnamurthy,[d] A. Surjalal Sharma,[e] Michael Ghil,[c,f]
and Eugenia Kalnay[a,b]

[a] *Department of Atmospheric and Oceanic Science, University of Maryland, College Park, College Park, Maryland*
[b] *Institute for Physical Science and Technology, University of Maryland, College Park, College Park, Maryland*
[c] *Geosciences Department and Laboratoire de Météorologie Dynamique (CNRS and IPSL), École Normale Supérieure and PSL University, Paris, France*
[d] *Center for Ocean–Land–Atmosphere Studies, George Mason University, Fairfax, Virginia*
[e] *Department of Astronomy, University of Maryland, College Park, College Park, Maryland*
[f] *Department of Atmospheric and Oceanic Science, University of California, Los Angeles, Los Angeles, California*

ABSTRACT: Oscillatory modes of the climate system are among its most predictable features, especially at intraseasonal time scales. These oscillations can be predicted well with data-driven methods, often with better skill than dynamical models. However, since the oscillations only represent a portion of the total variance, a method for beneficially combining oscillation forecasts with dynamical forecasts of the full system was not previously known. We introduce Ensemble Oscillation Correction (EnOC), a general method to correct oscillatory modes in ensemble forecasts from dynamical models. We compute the ensemble mean—or the ensemble probability distribution—with only the best ensemble members, as determined by their discrepancy from a data-driven forecast of the oscillatory modes. We also present an alternate method that uses ensemble data assimilation to combine the oscillation forecasts with an ensemble of dynamical forecasts of the system (EnOC-DA). The oscillatory modes are extracted with a time series analysis method called multichannel singular spectrum analysis (M-SSA), and forecast using an analog method. We test these two methods using chaotic toy models with significant oscillatory components and show that they robustly reduce error compared to the uncorrected ensemble. We discuss the applications of this method to improve prediction of monsoons as well as other parts of the climate system. We also discuss possible extensions of the method to other data-driven forecasts, including machine learning.

KEYWORDS: Climate prediction; Forecasting techniques; Statistical forecasting; Data assimilation; Nonlinear dynamics; Machine learning

---

1. Introduction and motivation

Weather predictability is inherently limited by chaotic error growth (Lorenz 1963, 1965). However, there are several physical processes that provide predictability for the atmosphere beyond the weather's synoptic time scale. Charney and Shukla (1981) recognized the potential for long-range predictability from slowly varying boundary conditions, including sea surface temperatures (SSTs). This predictability is particularly relevant for the tropical atmosphere, where synoptic-scale instabilities are less prominent. Recently, Bach et al. (2019) quantified the predictability provided by SST to the atmosphere from data. Coupled atmosphere–ocean interactions are also important for predictability beyond the weather time scale (Penny et al. 2019).

On intraseasonal time scales, in addition to the above sources of predictability, there is potential predictability due to climate oscillations. Because of their near-periodicity and low frequency, oscillations are predictable beyond weather time scales. The oscillations that are thought to be important for predictability on intraseasonal time scales include the Madden–Julian oscillation (MJO), monsoon intraseasonal oscillations (MISOs), and extratropical oscillations (Ghil and Robertson 2002; Stan and Krishnamurthy 2019; Krishnamurthy 2019). El Niño–Southern Oscillation (ENSO) also possesses an oscillatory component that, at times when it is more prominent, makes ENSO more predictable (Ghil and Jiang 1998). Climate oscillations on longer time scales also present an opportunity for enhanced prediction; however, there is controversy about whether robust oscillatory modes—defined as occupying a narrow band of frequencies—exist in the climate system on decadal time scales (Ghil 2001; Mann et al. 2020).

Many studies have shown that oscillatory modes extracted from data can be effectively forecasted. We will refer to these forecasts, which rely purely on data or on data combined with a low-order model, as "data-driven" forecasts; they are also referred to as empirical or statistical forecasts in the literature. These data-driven oscillation forecasts often possess comparable or higher skill than dynamical forecasts of these oscillations. See the literature on data-driven prediction of extratropical oscillations (Keppenne and Ghil 1993; Strong et al. 1995; Vautard et al. 1996), the MJO (Kang and Kim 2010; Kondrashov et al. 2013; Chen et al. 2014), and MISOs

---

Corresponding author: Eviatar Bach, eviatarbach@protonmail.com; Safa Mote, ssm@umd.edu

(Krishnamurthy and Sharma 2017; Chen et al. 2018). As an example of the latter, Krishnamurthy and Sharma (2017) showed that the leading MISO mode can be predicted for up to about 80 days using a data-driven method, more skillfully than MISO forecasts obtained from state-of-the-art models. This demonstrates the potential for improved intraseasonal prediction of monsoon rainfall by better prediction of MISO.

However, these methods are limited to improving forecasts of the oscillation itself, not of the full signal. Although the oscillations can be predicted fairly well, the magnitude of this prediction will be small compared with that of the full field, because the oscillatory components make up only a fraction of its total variance (Strong et al. 1995; Mo 2001). Data-driven models are generally not competitive with dynamical models in predicting the part of the variance comprised of high-frequency daily weather at medium range, due to the high dimensionality of the system necessitating unfeasibly large amounts of training data (Van den Dool 1994; Palmer 2020; Rasp and Thuerey 2021).

Our aim, then, is to combine dynamical forecasts with data-driven oscillation forecasts, in order to improve the former. This idea was previously suggested in Vautard et al. (1992), Strong et al. (1995), and Ghil et al. (2004); as put by Strong et al. (1995), the idea is to form a "combination of [...] an empirical model, to predict the smoothly varying part of the flow, and a dynamical model, to predict day-to-day 'weather,' superimposed on it" (p. 2628). However, a method for doing this has never been developed heretofore, to the best of our knowledge.

Here, we present a general method, called *ensemble oscillation correction* (EnOC), for improving ensemble forecasting of a chaotic system's full state by leveraging the predictability of oscillatory modes. EnOC selects ensemble members whose oscillation states are closest to a data-driven forecast of a given oscillatory mode. We also present an alternate method, *ensemble oscillation correction with data assimilation* (EnOC-DA), which treats oscillation forecasts as "observations" and uses data assimilation to combine them with a dynamical model forecast.

The rest of the paper is laid out as follows. Section 2 provides an overview of the existing methods to extract and forecast oscillations. Section 3 presents our novel EnOC and EnOC-DA methods. Section 4 describes the experiments applying EnOC and EnOC-DA to toy models, while section 5 discusses the results of these experiments. Finally, section 6 presents conclusions and future applications. The appendix contains a derivation of a theoretical error expression for EnOC and EnOC-DA.

2. Methods

We aim to use data-driven oscillation forecasts to improve prediction of systems that possess oscillatory modes. This requires three successive methods: 1) extraction of the oscillations from data, which we do here using multichannel singular spectrum analysis (section 2a), 2) mapping from a state in the full phase space to the corresponding state in the oscillations' subspace (section 2b), and 3) forecasting in this subspace (section 2c).

a. Multichannel singular spectrum analysis

Multichannel singular spectrum analysis (M-SSA), also known as extended empirical orthogonal functions (extended EOFs;

Weare and Nasstrom 1982), is a method used to extract spatiotemporal modes from multidimensional time series (Broomhead and King 1986; Plaut and Vautard 1994; Ghil et al. 2002). M-SSA helps separate such time series into a nonlinear trend, oscillatory modes, noise, and chaotic components (Ghil and Vautard 1991; Watari 1996). M-SSA has been applied in many climatic and other contexts; for instance, it has identified MISOs in Indian rainfall (Krishnamurthy and Shukla 2007; Moron et al. 2012) and been applied to Chinese rainfall (Wang et al. 1996), the MJO (Lau 2012), space weather (Sharma et al. 1993), and macroeconomic data (Groth and Ghil 2017), among others. Vautard et al. (1996) have demonstrated that the low-frequency modes extracted by M-SSA better correspond to the predictable modes of the atmosphere than those extracted by regular spatial EOFs.

We provide here a brief introduction to M-SSA; see Ghil et al. (2002), Alessio (2016, chapter 12), and Golyandina (2020) for a more detailed exposition. M-SSA is a form of principal component analysis (PCA), a widely used statistical method (Jolliffe 2002). In M-SSA, PCA is applied to sliding windows of length M along a time series in order to identify the orthogonal modes that capture the most variance in the time series. Here, we use the Broomhead–King variant of SSA, as opposed to the Toeplitz variant (see Ghil et al. 2002).

Consider a D-dimensional time series of length N, $\mathbf{x} = \{x_d(n)|d = 1, \ldots, D; n = 1, \ldots, N\}$. After choosing a window of length M based on the time scales of interest (see Vautard et al. 1992), we create lagged copies of the time series $\mathbf{x}_d(n) = [x_d(n), \ldots, x_d(n + M - 1)]$ for each dimension d, where $n = 1, \ldots, N - M + 1$. We combine these into the $(N - M + 1) \times (DM)$ trajectory matrix $\mathbf{X} = (\mathbf{x}_1, \ldots, \mathbf{x}_D)$, and form the covariance matrix $\mathbf{C} = \mathbf{X}^T\mathbf{X}/(N - M + 1)$.

The eigendecomposition of $\mathbf{C}$ yields the eigenvalues $\{\lambda_k\}$ and eigenvectors $\{\mathbf{e}_k\}$ (Plaut and Vautard 1994; Groth and Ghil 2011).[1] By analogy with the EOFs used in meteorology and oceanography (Preisendorfer and Mobley 1988), these eigenvectors are called space–time EOFs (ST-EOFs); see Ghil et al. (2002). The trace of the matrix $\mathbf{C}$ equals the total variance in the time series. It is also equal to the sum of the eigenvalues of $\mathbf{C}$, that is, $\mathrm{tr}(\mathbf{C}) = \sum_j \lambda_j$. Thus the ratio $\lambda_k/\mathrm{tr}(\mathbf{C})$ is the fraction of the total variance captured by mode k.

Due to the orthogonality of sinusoids that are in quadrature, oscillatory modes appear as pairs of eigenvectors with nearly identical eigenvalues, generalizing the sine–cosine pair of Fourier analysis (Vautard and Ghil 1989; Ghil et al. 2002; Alessio 2016). Since M-SSA may falsely detect such oscillatory modes in noise, statistical significance tests have been developed to distinguish oscillations from colored noise (Allen and Smith 1996; Ghil et al. 2002; Groth and Ghil 2015). Groth and Ghil (2011) developed a varimax algorithm to reduce mixture effects between physically distinct modes and better isolate them.

After choosing a mode of interest k from the set of eigenmodes distinguished by M-SSA, the original time series is projected onto the eigenvector $\mathbf{e}_k$ to obtain a principal

---

[1] When $DM > N - M + 1$, using the PCA "transpose trick" will be more computationally efficient; see Ghil et al. (2002, section A2).

Full phase space   Oscillation subspace
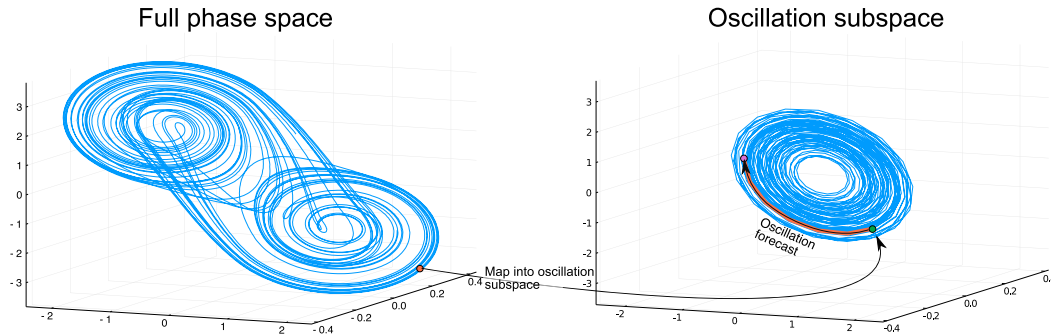

FIG. 1. Schematic diagram of projection to and forecasting in the oscillation subspace. We first project from the full phase space onto the oscillation subspace, as described in section 2b. Next, we forecast in the oscillation subspace, as described in section 2c. Note that for a pure sinusoidal oscillation—which would result, for example, from sampling a sinusoidal plane wave regularly in time at discrete points in space—the trajectories in the oscillation subspace will lie on an ellipse. We see that, in this example, the shape traced out by trajectories in the oscillation subspace deviates from an exact ellipse due to the oscillation's nonlinearity.

component (PC). However, the PC combines the behavior of all the D variables and does not possess the same phase as the original time series. The reconstruction procedure then extracts the portion of the time series corresponding to mode k for each variable and with correct phase. The resulting time series are called the reconstructed components (RCs); see Ghil and Vautard (1991), Ghil et al. (2002), and Groth and Ghil (2011).

The reconstruction at time n for the single mode k and dimension d can be written as the following operation on the original time series:[2]

$$r_{dk}(n) = \frac{1}{M_n} \sum_{m=L_n}^{U_n} \left[\sum_{d'=1}^{D} \sum_{m'=1}^{M} x_{d'}(n - m + m') e_{d'k}(m') \right] e_{dk}(m); \tag{1}$$

here $M_n = M$, $L_n = 1$, and $U_n = M$, except near the endpoints of the time series, for which the appropriate expressions are given in Vautard et al. (1992). Summing over all $\mathbf{r}_k$ yields the time series itself; $\mathbf{x}(n) = \sum_{k=1}^{DM} \mathbf{r}_k(n)$. Summing over a partial set $\mathcal{K} = \{k'_1, \ldots, k'_K\}$, with $K < DM$, yields a partial reconstruction; for example, if $\lambda_{k*} \simeq \lambda_{k*+1}$ and the other statistical significance tests are satisfied, $\mathbf{r}_{\mathcal{K}}(n)$ is the reconstruction of an oscillatory pair, with $\mathcal{K} = \{k*, \ k* + 1\}$ (Ghil et al. 2002). The RCs can be considered as the result of forward and reverse filtering of the original data (Harris and Yuan 2010).

We refer to the D-dimensional Euclidean space $\mathbb{X} \equiv \mathbb{R}^D$, with $\mathbf{x} \in \mathbb{X}$, as the *full phase space*. The DM-dimensional space $\mathbb{R}^{DM}$, in which the $\mathbf{x}_d(n)$ reside, is the *embedding space*. Any partial reconstruction with K components, $K < DM$, corresponds to a reduced subspace $\mathbb{X}_K$ spanned by the K vectors $\{\mathbf{e}_k | k \in \mathcal{K}\}$, which we refer to as the *reconstructed subspace*. When the set $\mathcal{K}$ corresponds to one or several oscillatory pairs (i.e., $\mathcal{K} = \mathcal{O}$), we refer to such a subspace as the *oscillation subspace*.

It can be seen from Eq. (1) that the reconstructed time series at time n is computed using the time n in the original time series, as well as $M - 1$ points from the immediate future and $M - 1$ points from the immediate past, except near the endpoints where increasingly fewer points are available.

These successively fewer points to average over present a problem near the endpoints, where reconstruction becomes less accurate. Methods for more accurate reconstruction near the endpoints have been developed, including the recent SSA with conditional predictions (Ogrosky et al. 2019; other methods for reconstruction near the endpoints are referenced therein).

Note that Eq. (1) is a linear transformation, which can be written as a $(2M - 1)D \times D$ matrix except near the endpoints. Since this operation is linear and rank-deficient, there is no unique sequence of $(2M - 1)$ D-dimensional full phase space states corresponding to each D-dimensional reconstructed phase space state.[3] However, the dynamics generating the time series will put additional constraints on possible sequences of full phase space states that are not captured by the linear reconstruction.

Because the reconstruction operation is not invertible, we cannot go directly from a point in the reconstructed phase space to the full space. Therefore, in order to use the forecast of the oscillatory modes to inform the forecast of the full state, we must do so indirectly. This is achieved using the EnOC and EnOC-DA algorithms described in section 3.

b. Projecting onto the oscillation subspace

For the EnOC algorithm that we introduce in section 3, we need a method to approximately map points from the full phase space to the reconstructed subspace; this projection is shown as the curved arrow from the full phase space to the oscillation subspace in Fig. 1.

---

[2] Note that this formula is not a computationally efficient way to implement reconstruction, as work can be saved by precomputing the PCs.

[3] A linear transformation is one-to-one if and only if it is of full rank.

Assuming the time series to be produced by a deterministic, autonomous dynamical system, it can be seen that any point in the full phase space corresponds to a single point in the reconstructed space. This is because the operation which maps from the full space to the reconstructed space is a function of the present, as well as $M - 1$ past and $M - 1$ future points. Since trajectories cannot cross in phase space due to uniqueness, a full phase space state will always have the same future and past, and thus the same corresponding reconstructed subspace state. Thus, if we were able to integrate the system forward and backward in time for $M - 1$ steps in each direction, we could exactly map the full phase space to the reconstructed subspace. Since an exact mapping is often not feasible, we construct an approximate mapping instead. In particular, we need an approximation of the mapping that does not require future information, since this is not available in the real-time forecasting context.

One such method relies on using the endpoint version of Eq. (1), which makes use of the present and the immediate $M - 1$ past or future points (Vautard et al. 1992). This was used in mapping from the full to the reconstructed phase space in Lynch (2019). Ogrosky et al. (2019) provided a more accurate expression for the endpoints.

Alternatively, one can use an analog method, which we use in the present paper. This method requires only the system state at the current time. We assume that we have access to a historical time series $\{\mathbf{x}(t_i)\}$ of the system, as well as the corresponding time series $\{\mathbf{r}(t_i)\}$ of the RCs. Then, to project a full state $\mathbf{x}(t)$ into the reconstructed subspace, we first find its k_f closest analogs, also referred to as nearest neighbors (Abarbanel et al. 1993), in the historical record of the full phase space, which occur at times that we denote by $t_1^*, \ldots, t_{k_f}^*$.

After finding the closest analogs in the full phase space $\mathbf{x}(t_1^*), \ldots, \mathbf{x}(t_{k_f}^*)$, we average over their corresponding points in the historical record of the RCs, $\mathbf{r}(t_1^*), \ldots, \mathbf{r}(t_{k_f}^*)$, weighted by the inverse distance of the analogs in the full phase space. In other words, we estimate the state in the reconstructed subspace as

$$\tilde{\mathbf{r}}[\mathbf{x}(t)] = \frac{\displaystyle\sum_{i=1}^{k_f} \frac{\mathbf{r}(t_i^*)}{\|\mathbf{x}(t) - \mathbf{x}(t_i^*)\|}}{\displaystyle\sum_{i=1}^{k_f} \|\mathbf{x}(t) - \mathbf{x}(t_i^*)\|^{-1}}. \qquad (2)$$

Although here we only use the present state $\mathbf{x}(t)$ to search for analogs, it may be advantageous to find analogs using a window that also includes the immediate past. Such an approach is taken in Farmer and Sidorowich (1987, 1988), Ukhorskiy et al. (2004), Chen and Sharma (2006), and Xavier and Goswami (2007).

Although the analog-based projection method works well for the low-dimensional systems considered here, it may be inaccurate for high-dimensional systems. We have found that lasso regression, where the predictors are states at multiple consecutive times, works better in high dimensions.

c. Forecasting in the oscillation subspace

We want to take a state $\mathbf{r}$ in the reconstructed subspace and forecast it from time t_i into a future time t_{i+1}; such a forecast is shown by the red arrow labeled "Oscillation forecast" on the right side of Fig. 1. We denote this operation by $\mathbf{R}_{t_i \to t_{i+1}}[\mathbf{r}(t_i)]$.

Several forecast methods based on M-SSA exist; they include fitting an autoregressive model to the RCs (Ghil and Jiang 1998; Ghil et al. 2002), SSA with conditional prediction to yield an $(M - 1)$-step forecast (Ogrosky et al. 2019), and neural networks (Lisi et al. 1995). We use an analog method here as in Krishnamurthy and Sharma (2017): we search for similar states to the present one in the historical record, and forecast based on the trajectory of these past states.

We again assume that we have a historical record $\{\mathbf{r}(t_i)\}$ of the RCs. Given the state $\mathbf{r}$ from which we want to start forecasting, we find its k_r closest analogs in the record. The trajectories of nearby points are expected to remain close before starting to diverge (Farmer and Sidorowich 1987; Krishnamurthy and Sharma 2017). Therefore, we follow the trajectory of these analogs until the end of the forecast window and average over their final states.

This method is similar to the analog method of Lorenz (1969), applied here to forecasting in the reconstructed subspace instead of the full phase space; see also the references in Farmer and Sidorowich (1988) for other early work on analog forecasting.

In this paper we use the Euclidean distance to define the closest analogs. Other metrics can also be used, such as distance in a reduced-dimensional space. For instance, Krishnamurthy and Sharma (2017) reduced the dimension D by projecting onto the leading spatial EOFs of the RCs. For a better algorithmic time complexity than a sequential search, the record can be stored in a k-dimensional (k-d) tree (Bentley 1975).

3. EnOC forecasting using oscillatory modes

a. EnOC algorithm

Here we introduce EnOC as a method for improving the ensemble forecasts by averaging solely over the ensemble members whose oscillation state is close to the data-driven oscillation forecast. We assume that we have access to a historical time series of noisy observations or an analysis[4] of the real system. This time series is used to perform the M-SSA, to map the full phase space $\mathbb{X}$ into the reconstructed subspace $\mathbb{X}_K$, and to forecast the oscillation, as in Krishnamurthy and Sharma (2017).

EnOC relies on using an ensemble of dynamical model forecasts of the full state $\{\mathbf{x}_i\}$, such as from an operational weather forecast model. Denote the projection of $\mathbb{X}$ into $\mathbb{X}_K$ by $\tilde{\mathbf{r}}(\mathbf{x})$, as per section 2b above, and the operation that forecasts the reconstructed subspace state $\mathbf{r}(t_i)$ to time t_{i+1} as $\mathbf{R}_{t_i \to t_{i+1}}[\mathbf{r}(t_i)]$, as per section 2c above. The EnOC algorithm is given by the following steps:

1) Using the best current estimate $\hat{\mathbf{x}}(t_0)$ of the full system state at time t_0—based, for instance, on an analysis combining the ensemble forecasts with observations—find the corresponding point $\tilde{\mathbf{r}}[\hat{\mathbf{x}}(t_0)] = \tilde{\mathbf{r}}(t_0)$ in the reconstructed subspace. Carry out an oscillatory-subspace forecast into the future from time t_0 to t_1, and call it $\bar{\mathbf{r}}(t_1) = \mathbf{R}_{t_0 \to t_1}[\tilde{\mathbf{r}}(t_0)]$.

---

[4] Objective analysis, or just analysis for short, refers to model estimates of the system that are corrected by observations using data assimilation (Bengtsson et al. 1981; Kalnay 2002).
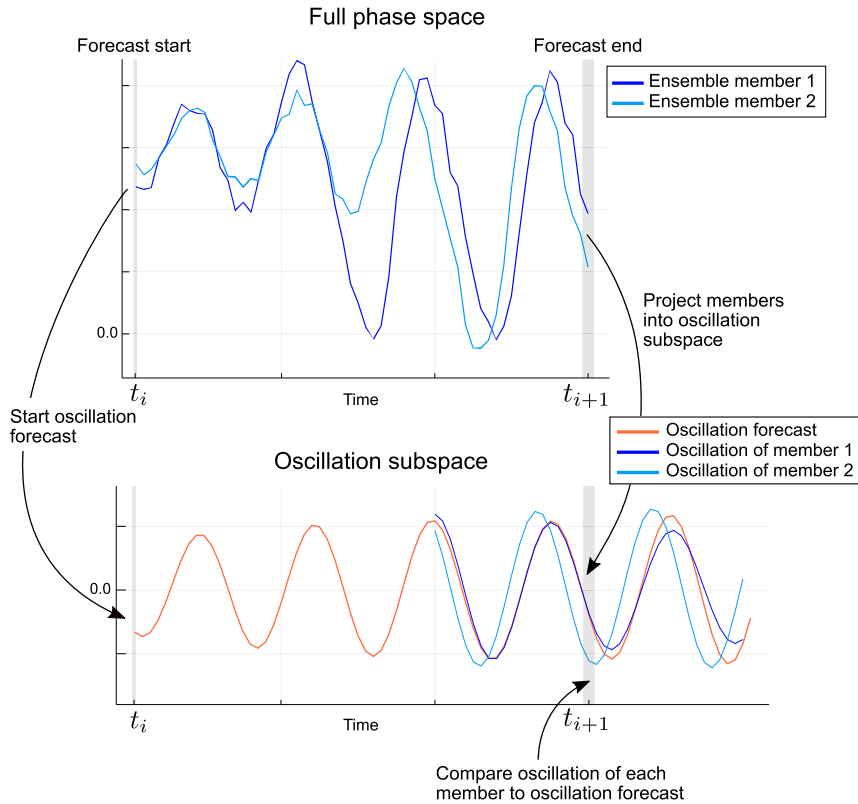
Full phase space

FIG. 2. Schematic diagram of EnOC for two ensemble members. (top) The dynamical model is integrated in time for each member. At the beginning of the forecast window (thin gray bar), the best estimate of the system state is mapped into the oscillation subspace, and an oscillation forecast is started (red curve). At the end of the forecast window (thick gray bar), the model forecasts are mapped into the oscillation subspace, and compared to the oscillation forecast. In this case, we see that the oscillation of ensemble member 1 is closer to the oscillation forecast than ensemble member 2. (bottom) A time window is shown for the oscillation of each ensemble member for visualization purposes, but in reality the method uses only the single state at the end of the forecast window.

2) Evolve each ensemble member $\{\mathbf{x}_i | i = 1, \ldots, m\}$ forward from time t_0 until the next forecast time t_1, using the dynamical model.

3) Find the points $\tilde{\mathbf{r}}_i(t_1) = \tilde{\mathbf{r}}[\mathbf{x}_i(t_1)]$ in the oscillatory subspace that correspond to each ensemble member at time t_1. Then, select the $m' \leq m$ ensemble members with the smallest distances $d_i = \|\tilde{\mathbf{r}}_i(t_1) - \overline{\mathbf{r}}(t_1)\|$ at time t_1, and compute the new ensemble mean $\overline{\mathbf{x}}^{(m')}(t_1)$ using only these m' members.

4) Repeat steps 1–3 for the next forecast time t_1.

These steps are illustrated schematically in Fig. 2.

Note that we may be interested in the oscillations in only some of the variables, like rainfall in the case of the monsoon. If so, only the variables of interest need be used in the M-SSA, and subsequently in steps 1 and 3 above.

b. Optimal number m' of ensemble members to average over

In step 3 above, we did not specify the number m' of "best" ensemble members to average over. This will differ depending on the system, the forecast lead time, the total number of

ensemble members m, and the ensemble spread, among other factors. Generally, the choice of m' will be a trade-off between 1) the improvement due to using a larger m' in the ensemble mean (Christiansen 2018; Wilks 2019), in the smaller error variance around the mean (Kalnay 2019), and in the ability to characterize the forecast uncertainty (Wilks 2019), versus 2) the inclusion of ensemble members that are far from the oscillation forecast. Here, we only optimize m' with respect to the root-mean-square error (RMSE) in the ensemble mean.

For a given setup—and assuming we have access to a historical record of ensembles, oscillation forecasts, and a good estimate of the system state $\hat{\mathbf{x}}(t)$ at those times—we can estimate m' by checking which one *would have* resulted in the smallest error. In other words,

$$m' = \arg\min_{\mu} \sum_t \|\overline{\mathbf{x}}^{(\mu)}(t) - \hat{\mathbf{x}}(t)\|^2, \tag{3}$$

where t is the time in the historical record, $\hat{\mathbf{x}}(t)$ is the estimate of the true state at that time, and $\overline{\mathbf{x}}^{(\mu)}(t)$ is the average over the μ ensemble members with the smallest distances d_i from

the oscillation forecast. The norm can be chosen to optimize the desired features of the forecast.

Note also that in this paper we pick a single m' for all the variables. However, in cases where an oscillation comprises significantly more variance in some variables than others—or in some regions in the case of a spatiotemporal oscillation—assigning different values of m' could improve the performance.

c. Data assimilation of oscillatory modes (EnOC-DA)

We present an alternate EnOC method, which we call ensemble oscillation correction with data assimilation (EnOC-DA). Instead of selecting ensemble members that have better oscillation forecasts, we use an ensemble data assimilation system to assimilate oscillation forecasts as pseudo-observations of the system. We are thereby combining the model forecast—treated as the first guess or background in the data assimilation (DA)—with the data-driven forecast of the oscillation, treated as an observation in DA.

EnOC-DA does not provide a significant advantage over regular EnOC in our toy model experiments in section 5, but it may do so for operational weather and climate prediction. For introductions to DA, see any number of books and review papers (e.g., Bengtsson et al. 1981; Ghil and Malanotte-Rizzoli 1991; Kalnay 2002; Asch et al. 2016; Carrassi et al. 2018).

In EnOC-DA, we assimilate the oscillation forecasts using the same forecast window that we would normally apply in EnOC. This choice differs from conventional DA in that we perform the assimilation at a point in time in the future, prior to having real observations at that time.

We use the ensemble transform Kalman filter (ETKF: Bishop et al. 2001), which allows for a nonlinear observation operator, which in this case is the projection of the full phase space onto the reconstructed subspace $\tilde{\mathbf{r}}[\mathbf{x}(t)]$, as discussed in section 2b. We estimate the observation error covariance matrix $\mathbf{R}$ empirically from the error statistics of the oscillation forecast. Thus, given oscillation forecasts $\bar{\mathbf{r}}(t_i)$ at points $\{t_i | i = 1, \ldots, N\}$ in time,

$$\hat{\mathbf{R}} = \frac{1}{N-1} \sum_{i=1}^{N} [\hat{\mathbf{r}}(t_i) - \bar{\mathbf{r}}(t_i)][\hat{\mathbf{r}}(t_i) - \bar{\mathbf{r}}(t_i)]^{\mathrm{T}}, \quad (4)$$

where $\hat{\mathbf{r}}(t_i)$ is the true value of the oscillation at time t_i.

d. Notes

If the system contains anticorrelated large-amplitude modes that are not included in the reconstruction then the error reduction may be small, since an improvement in the included modes will be partially offset by the missing ones that make up a large portion of the variance. Thus, the correlations between the RCs need to be inspected and all correlated and anticorrelated modes should be included.

This is related to the problem of separability in the SSA literature, and the correlations between the RCs—slightly modified to account for the endpoint effects—are referred to as $\mathbf{w}$ correlations (Golyandina et al. 2001). Since our EnOC algorithm reduces the ensemble spread in the oscillatory modes, having them be uncorrelated from the rest of the modes also helps the ensemble maintain a sufficient spread in these other modes.

Assuming that the RCs are uncorrelated, we derive in the appendix a rough estimate for the "best-case" ratio of the RMSE of the EnOC-corrected forecast to the RMSE of the uncorrected forecast:

$$\frac{\mathrm{RMSE}'}{\mathrm{RMSE}} = \left(1 - \frac{\sum_{j \in \mathscr{O}} \lambda_j}{\sum_{j \in \mathscr{P}} \lambda_j} \right)^{1/2}; \quad (5)$$

here $\mathscr{O}$ is the set of indices of the oscillatory modes, and $\mathscr{P}$ is the set of indices of all modes excluding a mean mode, if it exists; see the appendix for further details. The two major MISOs make up about 14% of the variance in the daily rainfall anomalies (Krishnamurthy and Shukla 2007). Other oscillations in the large-scale atmospheric flow are estimated to comprise 20%–30% of the variance (Ghil and Robertson 2002; Stan and Krishnamurthy 2019).

The derivation assumes that EnOC results in a perfect prediction of the oscillation, so the real error improvement will generally be smaller than suggested by this equation. Unsurprisingly, we see that the larger the portion of the variance comprised by the oscillatory modes, the more the potential error reduction.

Note that EnOC is not restricted to oscillatory modes and is potentially useful given any modes that are predicted well using a data-driven forecasting method. However, we consider oscillatory modes the most likely to have useful data-driven forecasts for high-dimensional geophysical systems, as already demonstrated by several authors and reviewed here in section 1. We discuss the possible extension of the method to nonoscillatory modes in section 6a.

For the method to work, it is important to have ensemble members that yield oscillations approximating the correct one. This condition may not be satisfied for models that miss important physics (Ghil and Robertson 2000), or if the ensemble spread is too small. Such "underdispersion" occurs when the ensemble underestimates the uncertainty in the forecast, and it may then be beneficial to increase the ensemble spread.

Both EnOC and EnOC-DA are a correction performed after the ensemble forecast, so they can be considered a form of ensemble postprocessing (Vannitsem et al. 2018). EnOC is similar to the ensemble subsetting methods in Dong and Zhang (2016) and Ancell (2016) in that it picks a subset of ensemble members whose mean outperforms the ensemble mean. However, these methods fundamentally differ from EnOC in that they verify against real-time observations instead of a data-driven forecast, and in that they pick a subset and uses it for future lead times instead of a possibly different subset at every lead time.

e. Software implementation

In section 4, EnOC and EnOC-DA are applied to three nonlinear, chaotic systems of ordinary differential equations (ODEs). The experiments were performed using open-source Julia code that is available on GitHub.[5] To efficiently search for analogs, the k-d tree data structure was used from the NearestNeighbors.jl library (Carlsson 2020). The Lyapunov exponents of the systems

---

[5] https://github.com/eviatarbach/ensemble_oscillation_correction.jl

TABLE 1. Parameters, properties, and configuration for the ODE systems used in the experiments; nondimensional time units.

	Colpitts, Eq. (6)	Chua, Eq. (7)	Forced Lorenz (1963), Eq. (8)
Parameters	$p_1 = 5, p_2 = 0.0797, p_3^{(1)} = 9,$ $p_3^{(2)} = 10.5,$ $p_4 = 0.6898, c_{(2,1)} = 0.05$	$\alpha = 15.6, \beta = 25.58,$ $m_0 = -8/7,$ $m_1 = -5/7$	$\sigma = 10, \beta = 8/3, \rho = 28, c = 5,$ $\Omega = 0.3, u(0) = 0, v(0) = 3$
Maximum Lyapunov exponent (Lyapunov time)	0.097 (10.3)	0.48 (2.1)	0.36 (2.8)
Peak oscillation frequency (period)	0.18 (5.6)	0.63 (1.6)	0.048 (20.9)
Perturbed parameters	$p_1' = 5.1, p_2' = 0.0897$	$\alpha' = 15.7, \beta' = 24.58$	$\Omega' = 0.32, c' = 5.1$
Embedding dimension used	30	60	100
Sampling time used	0.4	0.1	0.5
Oscillation eigenvalue indices	2, 3	3, 4	1, 2
Portion of variance due to oscillations	29%	18%	51%

were computed using the DynamicalSystems.jl library (Datseris 2018). The continuous ranked probability score (CRPS) was computed with the properscoring library (Climate Corporation 2015). The time integrations used the fourth-order Runge–Kutta scheme. We used the parasweep library for Python of Bach (2021) to facilitate the running in parallel of multiple experiments at different lead times and with different parameters.

4. Experiments applying EnOC to toy models

We test EnOC and EnOC-DA on three ODE systems: two coupled Colpitts oscillators (e.g., Kennedy 1994), a Chua oscillator (Chua et al. 1986), and a periodically forced Lorenz (1963) model. These three systems exhibit oscillatory behavior and are all chaotic, with maximal Lyapunov exponents $\lambda_{\max}$ greater than zero. Lyapunov exponents characterize the rate of exponential error growth (Ott 2002). The Lyapunov time $\lambda_{\max}^{-1}$ is the characteristic time scale of chaotic error growth.

We apply parametric model errors to each model in order to test a scenario as in Krishnamurthy and Sharma (2017), in which the oscillatory modes are poorly predicted by the dynamical model. For each system, we list its parameters, Lyapunov exponents, peak oscillation frequency, experiment configuration, and oscillation properties in Table 1.

Our system of coupled Colpitts oscillators follows Rey et al. (2014):

$$\dot{x}_1^{(i)}(t) = p_1 x_2^{(i)}(t) + c_{(i+1,i)}[x_1^{(i+1)}(t) - x_1^{(i)}(t)], \quad (6a)$$

$$\dot{x}_2^{(i)}(t) = -p_2[x_1^{(i)}(t) + x_3^{(i)}(t)] - p_3^{(i)} x_2^{(i)}(t), \quad (6b)$$

$$\dot{x}_3^{(i)}(t) = p_4\{x_2^{(i)}(t) + 1 - \exp[-x_1^{(i)}(t)]\}. \quad (6c)$$

Here $\dot{x} = dx/dt$, the indices are cyclical, and we use two oscillators. For the values of the parameters in Table 1, the two oscillators do not synchronize.

The dimensionless Chua oscillator describes the so-called double-scroll family, and it is governed by the following equations (Chua et al. 1986):

$$\dot{x} = \alpha[y - x - f(x)], \quad (7a)$$

$$\dot{y} = x - y + z, \quad (7b)$$

$$\dot{z} = -\beta y, \quad (7c)$$

$$f(x) = m_1 x + \frac{1}{2}(m_0 - m_1)(|x + 1| - |x - 1|). \quad (7d)$$

For this system, due to mixture of low-frequency variability in the oscillatory modes, we follow Groth and Ghil (2011) and perform varimax rotation.

Mani et al. (2009) estimated the largest Lyapunov exponent for the Indian monsoon rainfall in the years 1979–2004 to be $\lambda_{\max} \simeq 0.36$ day^{-1}, corresponding to a Lyapunov time $\lambda_{\max}^{-1} \simeq 2.8$ days. The Lyapunov time is thus significantly smaller than the leading MISO period of about 45 days (Krishnamurthy and Shukla 2007). To study a toy system with a Lyapunov time much smaller than the oscillation period, we use the following modified Lorenz (1963) system, with the addition of sinusoidal forcing in the x component:

$$\dot{x} = \sigma(y - x) + cu, \quad (8a)$$

$$\dot{y} = x(\rho - z) - y, \quad (8b)$$

$$\dot{z} = xy - \beta z, \quad (8c)$$

$$\dot{u} = v, \quad \dot{v} = -\Omega^2 u. \quad (8d)$$

Ghil and Jiang (1998) used a similar system to illustrate the effectiveness of single-channel SSA in real-time ENSO prediction. Here, the periodic forcing given by Eq. (8d) represents the MISO cycle rather than the previous authors' seasonal cycle. It should be noted that the maximum Lyapunov exponent for this sinusoidally driven Lorenz system is less than the value of 0.906 for the same parameters but with no driving. For this Lorenz system we use only the x and y variables in EnOC and EnOC-DA, since the oscillation is most prominent in these variables of the forced subsystem.

To create the "historical data" used in the EnOC algorithm of section 3a in order to carry out the M-SSA analysis of the dataset, project from the full phase space to the oscillation subspace, and forecast the oscillatory modes, we run a transient of $3000\Delta t$, followed by a record of $22\,000\Delta t$, where Δt is the sampling time. A Gaussian error with variance of 10% of the standard deviation of each variable is then added to that variable.
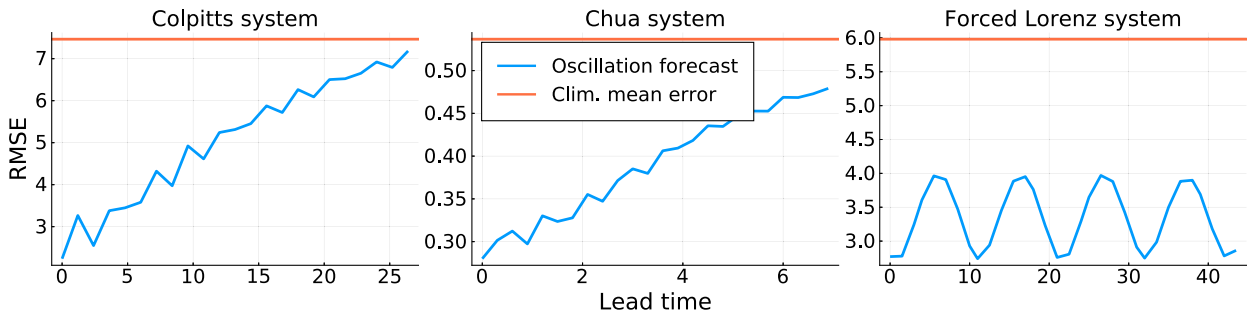
FIG. 3. The root-mean-square error (RMSE) error of the oscillation forecasts at a given lead time (blue), along with the error obtained by predicting the climatological mean value of the oscillation (red). This error is computed on a test set of $2200\Delta t$.

To generate the ensembles, we draw from a Gaussian distribution around the true state, with standard deviations of 20% of the standard deviation of each variable. For the Chua system, randomly perturbed states are often outside the strange attractor's basin of attraction and the trajectory becomes unbounded; we thus verify that each perturbed state is on a bounded trajectory, and discard it if it is not. We use $m = 20$ ensemble members.

In each case, we present the results after finding the optimal m' based on 1000 forecast cycles. We then run 10 000 forecast cycles to compare the error of the corrected versus uncorrected ensemble using this m'.

For the nearest-neighbor analog mapping from the full phase space to the oscillation subspace, as well as the oscillation forecast, we use $k_f = k_r = 30$ neighbors. In applying EnOC-DA, for each system and lead time, we run experiments with multiplicative inflation factor λ (corresponding to inflating the background covariance matrix by λ^2), between 1.0 and 1.45 at increments of 0.05, and choose the one that results in the lowest error.

5. Results

a. Oscillation forecasts

Before showing the results of EnOC and EnOC-DA, we first consider the error growth with lead time of the oscillation forecasts themselves. This is shown in Fig. 3. The RMSE is computed as $\|\bar{\mathbf{r}}(t) - \hat{\mathbf{r}}(t)\|_2/\sqrt{D'}$, where $\bar{\mathbf{r}}(t)$ is the oscillation forecast, $\hat{\mathbf{r}}(t)$ is the true state of the oscillation, and D' is the number of variables used in the oscillation forecasts. The error obtained by predicting the climatological mean value of the oscillation is also included.

When the error of the oscillation forecast reaches the error obtained by predicting the climatological mean of the oscillation, the oscillation forecast no longer offers any useful skill. Whereas for the Colpitts and Chua systems the error grows with time, we see that for the forced Lorenz system the error instead oscillates without growing. This is because the oscillation in x and y for the forced Lorenz system is close to being purely harmonic, rather than broad-peaked, and it can thus be predicted well at long lead times; see Ghil and Childress (1987, section 12.6).

b. EnOC and EnOC-DA applied to ensemble mean

Figure 4 shows the RMSE as a function of lead time in forecasting the Colpitts system given by Eq. (6), comparing the ensembles corrected with EnOC and EnOC-DA to the uncorrected forecasts. We also plot the error as theoretically estimated from Eq. (5). Note that the RMSE here refers to $\|\bar{\mathbf{x}}(t) - \hat{\mathbf{x}}(t)\|_2/\sqrt{D}$—where $\bar{\mathbf{x}}(t)$ is the forecast and $\hat{\mathbf{x}}(t)$ is the true state—averaged over all the forecast cycles.

Up until a lead time of about 7, the error reductions of EnOC are close to the predictions of Eq. (5). However, there are significant improvements at even the largest lead times past 25. At longer lead times, the oscillation forecast loses skill and EnOC can contribute less, but is still providing predictability. We see that EnOC-DA performs worse than EnOC except at the longest lead times.

Figure 5 displays the results of applying EnOC and EnOC-DA to the Chua system of Eq. (7). The error reductions closely track the estimates of Eq. (5), and EnOC and EnOC-DA perform similarly, with EnOC slightly better. The error reductions continue to the longest lead times shown on the plot, when the errors reach the level attained by predicting the climatological mean.
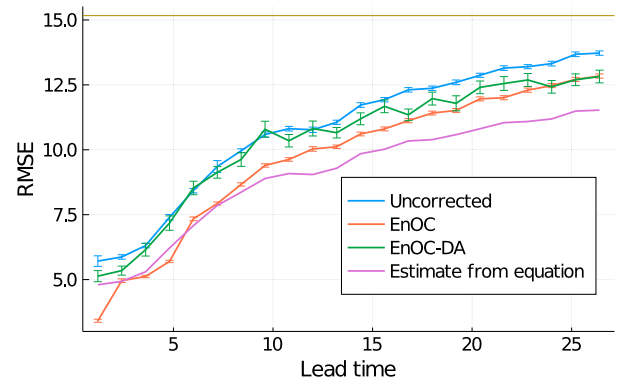

FIG. 4. The RMSE for the forecasts of the Colpitts system of Eq. (6) at a given lead time, comparing the uncorrected forecast, the EnOC forecast, the EnOC-DA forecast, and the estimate from Eq. (5). The error bars are the standard error in the mean over 10 000 forecast cycles for EnOC and 1000 cycles for EnOC-DA. We include, for reference, the error obtained by predicting the climatological mean as the dark yellow line.
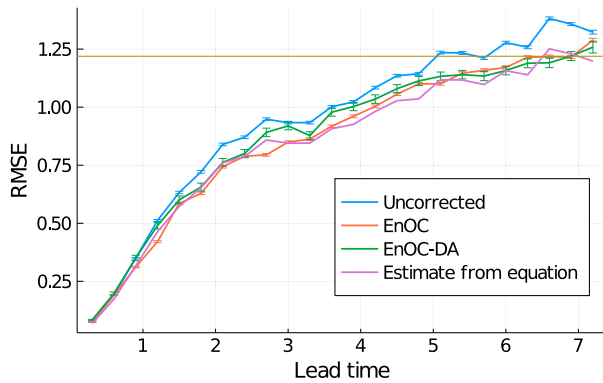
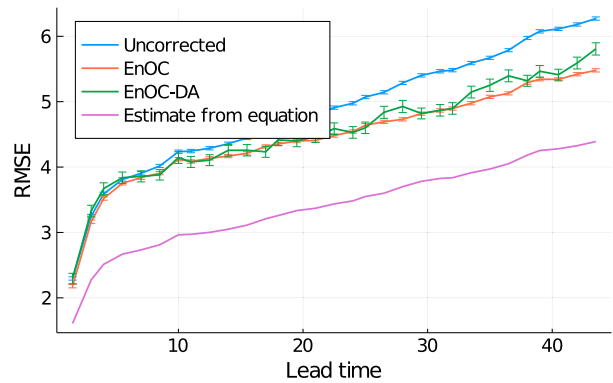FIG. 5. As in Fig. 4, but for the Chua system of Eq. (7).

FIG. 6. As in Fig. 4, but for the periodically forced Lorenz (1963) system of Eq. (8). The error obtained by predicting the climatological mean is not shown since it is outside of the range of the y axis, at around 15.8.

Finally, Fig. 6 shows the error for the periodically forced Lorenz system. The results here are qualitatively quite different from those of the Colpitts and Chua systems in Figs. 4 and 5, since the oscillation period is significantly longer than the Lyapunov time. We can see only insignificant error reductions until around time 1, and the reductions keep increasing in magnitude with increasing lead time.

The larger reductions at longer lead times can be explained by the long oscillation period: at small lead times, the oscillation has only completed a small fraction of its period, and therefore does not provide significant predictability for the system. Moreover, as discussed in section 5a, the oscillation of this system can be predicted at long lead times without growing error. Here, the error reductions are much smaller than the "best case" predicted by Eq. (5) at all lead times.

For the above EnOC results of Figs. 4–6, we have picked the optimal number m' of ensemble members to average over as described in section 3b. Note that this is not done for EnOC-DA, which uses all the ensemble members.

Figure 7 shows an example of an RMSE vs. m' curve for the Chua system. We can clearly see the trade-off between two factors: too small an ensemble leads to a larger RMSE, while including too many ensemble members that are not close enough to the oscillation forecast also leads to a larger RMSE. This convex shape of the RMSE vs. m' curve was common in our experiments.

In the case at hand, the optimal m' is 11. In cases where the oscillation forecast is very informative about the system state, the optimal m' will be small. If the oscillation forecast is useless, the error will generally be monotonically decreasing with increasing m', and $m' = m$ (the total number of ensemble members) will be optimal.

For comparison, we also try picking random subsets of ensemble members. As we see in Fig. 7, the error generally decreases monotonically with increasing ensemble size, demonstrating that the subsets of ensemble members chosen by EnOC are better than chance.

c. EnOC for probabilistic forecasts

Thus far, we have only considered the error in the ensemble mean. Often in ensemble forecasting, however, the distribution of the ensemble members is also of interest. Here we evaluate the performance of EnOC in terms of its impact on the forecast probability distribution.

A common error score for probabilistic forecasts is the continuous ranked probability score (CRPS: Hersbach 2000). Although defined for continuous probability distributions, a version called the ensemble CRPS can be computed by assuming that each ensemble member is a sample from a forecast probability distribution. A lower score of CRPS indicates a better forecast (Wilks 2019).

Here, instead of using EnOC to compute the ensemble mean, we instead use it to pick the best m' ensemble members to characterize the forecast distribution. We then compare the CRPS computed by considering only those members to the CRPS computed by considering all the members.

Figure 8 displays the CRPS for the Chua, periodically forced Lorenz (1963), and Colpitts systems, for both the uncorrected and EnOC forecasts. We see that EnOC results in a significant reduction in the CRPS for the Chua and forced Lorenz systems. For the Colpitts system the reduction in CRPS is very small, and stops at lead times beyond 15.
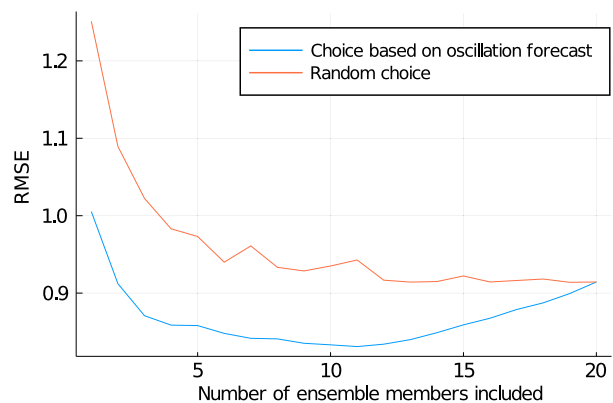

FIG. 7. The RMSE of the ensemble mean $\bar{\mathbf{x}}^{(m')}$ vs m' for the Chua system at a lead time of 3.0 time units (blue), along with the RMSE when choosing a random subset of ensemble members of length m' (red).
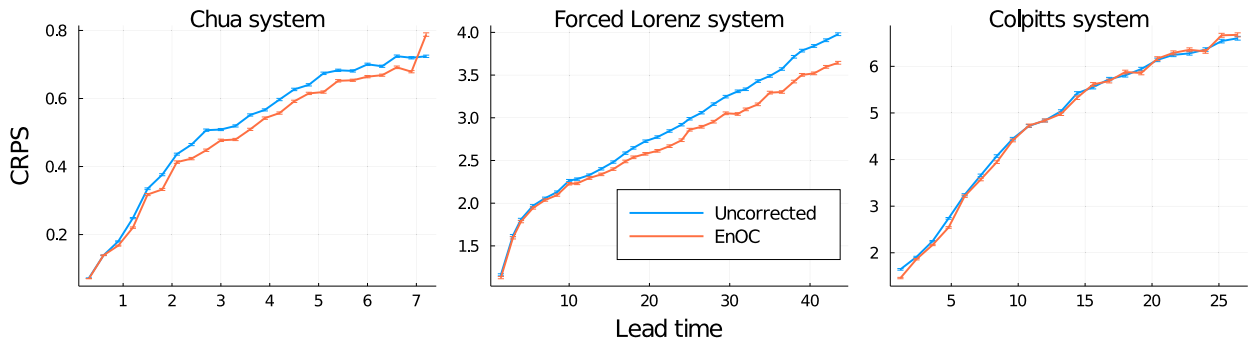
FIG. 8. CRPS for the three systems at a given lead time, comparing the uncorrected and EnOC forecasts.

These results demonstrate that EnOC can improve probabilistic forecasts. However, the results for the Colpitts system show that it is possible for EnOC to improve the ensemble mean forecast without significantly improving the probabilistic forecast. These results may be improved if CRPS is chosen to optimize m' instead of RMSE (see section 3b). Moreover, we have found that EnOC-DA can reduce the CRPS more than EnOC in some cases.

6. Concluding remarks

a. Summary

Oscillatory modes are prevalent in the climate system, and tend to be more predictable than the overall signal (Ghil and Childress 1987; Ghil and Robertson 2002; Ghil et al. 2019; Krishnamurthy 2019). Previous studies have focused on the prediction of these modes, but did not use this information systematically to improve the prediction of the overall system.

Here, we presented a general method, which we term ensemble oscillation correction (EnOC), to improve ensemble forecasts of systems with oscillatory modes. We demonstrate a robust error reduction over uncorrected forecasts. Equation (5) provides a rough "best case" error reduction for this method based only on the percentage of the variance comprised by the oscillation.

We also introduced an alternate method that uses data assimilation, EnOC with Data Assimilation (EnOC-DA). EnOC-DA generally results in similar but slightly smaller error reductions compared to EnOC for the chaotic toy models we tested. This may be because spinup is important for the performance of ensemble Kalman filters (Kalnay and Yang 2010), and the filter was not spun up here. A hybrid method that incorporates a climatological background error covariance matrix could mitigate this problem (Penny 2017).

In this paper we used M-SSA due to the extensive literature on its use for extraction and prediction of climate oscillations, as described in section 2a. As an extension of this work, the EnOC ideas could be tested using a number of other methods of extracting oscillatory modes from multivariate time series data; see von Storch and Zwiers (1999), Jolliffe (2002), and Ghil et al. (2002) for an overview. These include multitaper frequency domain-singular value decomposition (MTM-SVD: Mann and Park 1999), Hilbert EOFs (Rasmusson et al. 1981), principal

oscillation patterns (Hasselmann 1988), predictive oscillation patterns (Kooperberg and O'Sullivan 1996), or the recent data-adaptive harmonic (DAH)–multilayer Stuart–Landau model (MSLM) methodology (Kondrashov et al. 2018).

Even more broadly, the idea of EnOC and EnOC-DA of optimally combining dynamical and data-driven forecasts could be extended to use machine learning forecasts, or other low-order, robust forecasts, instead of the oscillatory-mode forecasts on which the full-space forecast builds herein. As one example of many, the reduced-subspace forecast could rely on empirical model reduction (EMR: Kravtsov et al. 2010) that has been widely applied in the climate sciences and beyond (Kondrashov et al. 2015). EMR forecasts have participated for over a decade in the ENSO forecast plume of the International Research Institute for Climate and Society and have been found still in 2012 to be highly competitive with ENSO forecasts of high-end climate models (Barnston et al. 2012).

b. Discussion: Application to real-time predictions

Motivated by the poor forecasts of the monsoon intraseasonal oscillations (MISOs) with state-of-the-art climate models (e.g., Krishnamurthy and Sharma 2017), we intend to apply this method to the South Asian monsoon region (see Pentakota et al. 2020).

The error reduction will vary in space according to the percentage of the variance represented by MISOs at that location. Rainfall reflects latent heating and thus the patterns of potential vorticity, which in turn are a key determinant in the atmospheric flow. Therefore, if EnOC ensemble members are chosen with EnOC based on their predicted rainfall, it is likely that their error in forecasts of other variables could also be improved.

This idea is supported by Lien et al. (2013), who found that data assimilation of observed precipitation reduced the forecast error in other fields. They attributed this improvement to the increased weight placed on ensemble members with more correct potential vorticity, which is done implicitly with the ensemble transform Kalman filter.

APPENDIX

Derivation of an A Priori Error Reduction Estimate

We derive a rough estimate of a "best case" expected error reduction of EnOC given some reasonable assumptions.

We start, for simplicity, with the squared prediction error of the forecast $\hat{y}(t)$ of the one-dimensional time series $y(t)$ at a particular time step t; the same results hold for multiple dimensions and multiple time steps. We assume that both the real and predicted time series have the same set of SSA eigenvectors, such that both can be decomposed into the same set of modes $\mathscr{J}$, that is,

$$y = \sum_{j \in \mathscr{J}} r_j \quad \text{and} \quad \hat{y} = \sum_{j \in \mathscr{J}} \hat{r}_j,$$

where we drop the dependence on t. This assumption will hold if and only if the two corresponding covariance matrices commute with each other.

We regard each j as a single mode, although the derivation also works if each j is a subset of modes. In particular, each oscillatory mode will generally correspond to a pair of eigenvectors in quadrature; see section 2a. We then obtain

$$(\hat{y} - y)^2 = \left(\sum_j \hat{r}_j - \sum_j r_j \right)^2 = \left[\sum_j (\hat{r}_j - r_j) \right]^2$$
$$= \sum_j (\hat{r}_j - r_j)^2 + \sum_{k \neq \ell} (\hat{r}_k - r_k)(\hat{r}_\ell - r_\ell)$$
$$= \sum_j (\hat{r}_j - r_j)^2$$
$$\quad + \sum_{k \neq \ell} [(\hat{r}_k - \bar{r}_k) - (r_k - \bar{r}_k)][(\hat{r}_\ell - \bar{r}_\ell) - (r_\ell - \bar{r}_\ell)],$$

where $\bar{r}_i$ is the mean of the ith RC and we assume that the mean is the same for the real and predicted RCs.

Next, we assume that the different groups of RCs are uncorrelated, as discussed in section 2d, so that the second term above vanishes and

$$(\hat{y} - y)^2 = \sum_j (\hat{r}_j - r_j)^2.$$

Assume further that EnOC corrects mode j' "perfectly," so that $(\hat{r}_{j'} - r_{j'})^2 = 0$. Then, the ratio of the uncorrected mean-square error MSE to the corrected mean-square error MSE$'$ will be

$$\frac{\text{MSE}}{\text{MSE}'} = \frac{\sum_{j \neq j'} (\hat{r}_j - r_j)^2 + (\hat{r}_{j'} - r_{j'})^2}{\sum_{j \neq j'} (\hat{r}_j - r_j)^2}$$
$$= 1 + \frac{(\hat{r}_{j'} - r_{j'})^2}{\sum_{j \neq j'} (\hat{r}_j - r_j)^2}.$$

Now, we assume that the prediction error of each mode is proportional to the variance captured by that mode [i.e., $(\hat{r}_j - r_j)^2 = c_j \text{Var}_j$]. Then,

$$\frac{\text{MSE}}{\text{MSE}'} = 1 + \frac{c_{j'} \text{Var}_{j'}}{\sum_{j \neq j'} c_j \text{Var}_j}.$$

Often the highest-variance mode corresponds simply to the long-term mean of the time series; unless, that is, $y(t)$ is already an anomaly time series with mean zero. Assuming that the model estimates this mean well, c_j will be small for this mode, and we assume that its $c_j = 0$.

We expect that the noise modes will have large c_j, due to their unpredictability, and modes that are well predicted by the model will have small c_j. Here we make the simplifying assumption that, except for the mode corresponding to the mean, all the other modes are predicted equally well by the model and thus their c_j terms are the same. One thus obtains

$$\frac{\text{MSE}}{\text{MSE}'} = 1 + \frac{\text{Var}_{j'}}{\text{Var}_{\text{tot}} - \text{Var}_{j'}},$$

where Var_{tot} is the total variance of all the modes, excluding the mean mode.

Expressing now the ratio of the corrected to uncorrected error in terms of the RMSE, we get

$$\frac{\text{RMSE}'}{\text{RMSE}} = \left(1 - \frac{\text{Var}_{j'}}{\text{Var}_{\text{tot}}} \right)^{1/2}$$
$$= \left(1 - \frac{\sum_{j \in \mathscr{O}} \lambda_j}{\sum_{j \in \mathscr{P}} \lambda_j} \right)^{1/2},$$

where λ_j is the eigenvalue corresponding to mode j, $\mathscr{O}$ is the set of indices of the oscillatory modes, and $\mathscr{P}$ is the set of indices of all modes excluding a mean mode. The last equality follows from the discussion in section 2a.

REFERENCES

Abarbanel, H. D. I., R. Brown, J. J. Sidorowich, and L. S. Tsimring, 1993: The analysis of observed chaotic data in physical systems. *Rev. Mod. Phys.*, **65**, 1331–1392, https://doi.org/10.1103/RevModPhys.65.1331.

Alessio, S. M., 2016: Singular spectrum analysis (SSA). *Digital Signal Processing and Spectral Analysis for Scientists: Concepts and Applications*, Springer, 537–571, https://doi.org/10.1007/978-3-319-25468-5.

Allen, M. R., and L. A. Smith, 1996: Monte Carlo SSA: Detecting irregular oscillations in the presence of colored noise. *J. Climate*, **9**, 3373–3404, https://doi.org/10.1175/1520-0442(1996)009<3373:MCSDIO>2.0.CO;2.

Ancell, B. C., 2016: Improving high-impact forecasts through sensitivity-based ensemble subsets: Demonstration and initial tests. *Wea. Forecasting*, **31**, 1019–1036, https://doi.org/10.1175/WAF-D-15-0121.1.

Asch, M., M. Bocquet, and M. Nodet, 2016: *Data Assimilation: Methods, Algorithms, and Applications*. Society for Industrial and Applied Mathematics, 306 pp., https://doi.org/10.1137/1.9781611974546.

Bach, E., 2021: parasweep: A template-based utility for generating, dispatching, and post-processing of parameter sweeps. *SoftwareX*, **13**, 100631, https://doi.org/10.1016/j.softx.2020.100631.

——, S. Motesharrei, E. Kalnay, and A. Ruiz-Barradas, 2019: Local atmosphere–ocean predictability: Dynamical origins, lead times, and seasonality. *J. Climate*, **32**, 7507–7519, https://doi.org/10.1175/JCLI-D-18-0817.1.

Barnston, A. G., M. K. Tippett, M. L. L'Heureux, S. Li, and D. G. DeWitt, 2012: Skill of real-time seasonal ENSO model predictions during 2002–11: Is our capability increasing? *Bull. Amer. Meteor. Soc.*, **93**, 631–651, https://doi.org/10.1175/BAMS-D-11-00111.1.

Bengtsson, L., M. Ghil, and E. Källen, Eds., 1981: *Dynamic Meteorology: Data Assimilation Methods*. Applied Mathematical Sciences Series, Vol. 36, Springer-Verlag, 330 pp., https://doi.org/10.1007/978-1-4612-5970-1.

Bentley, J. L., 1975: Multidimensional binary search trees used for associative searching. *Commun. ACM*, **18**, 509–517, https://doi.org/10.1145/361002.361007.

Bishop, C. H., B. J. Etherton, and S. J. Majumdar, 2001: Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects. *Mon. Wea. Rev.*, **129**, 420–436, https://doi.org/10.1175/1520-0493(2001)129<0420:ASWTET>2.0.CO;2.

Broomhead, D. S., and G. P. King, 1986: Extracting qualitative dynamics from experimental data. *Physica D*, **20**, 217–236, https://doi.org/10.1016/0167-2789(86)90031-X.

Carlsson, K., 2020: NearestNeighbors.jl: High performance nearest neighbor data structures and algorithms for Julia. GitHub, accessed 14 May 2021, https://github.com/KristofferC/NearestNeighbors.jl.

Carrassi, A., M. Bocquet, L. Bertino, and G. Evensen, 2018: Data assimilation in the geosciences: An overview of methods, issues, and perspectives. *Wiley Interdiscip. Rev.: Climate Change*, **9**, e535, https://doi.org/10.1002/wcc.535.

Charney, J. G., and J. Shukla, 1981: Predictability of monsoons. *Monsoon Dynamics*, J. Lighthill and R. P. Pearce, Eds., Cambridge University Press, 99–110, https://doi.org/10.1017/CBO9780511897580.009.

Chen, J., and A. S. Sharma, 2006: Modeling and prediction of the magnetospheric dynamics during intense geospace storms. *J. Geophys. Res. Space Phys.*, **111**, A04209, https://doi.org/10.1029/2005JA011359.

Chen, N., A. J. Majda, and D. Giannakis, 2014: Predicting the cloud patterns of the Madden–Julian oscillation through a low-order nonlinear stochastic model. *Geophys. Res. Lett.*, **41**, 5612–5619, https://doi.org/10.1002/2014GL060876.

——, ——, C. T. Sabeerali, and R. S. Ajayamohan, 2018: Predicting monsoon intraseasonal precipitation using a low-order nonlinear stochastic model. *J. Climate*, **31**, 4403–4427, https://doi.org/10.1175/JCLI-D-17-0411.1.

Christiansen, B., 2018: Ensemble averaging and the curse of dimensionality. *J. Climate*, **31**, 1587–1596, https://doi.org/10.1175/JCLI-D-17-0197.1.

Chua, L., M. Komuro, and T. Matsumoto, 1986: The double scroll family. *IEEE Trans. Circuits Syst.*, **33**, 1072–1118, https://doi.org/10.1109/TCS.1986.1085869.

Climate Corporation, 2015: properscoring. Accessed 14 May 2021, https://github.com/TheClimateCorporation/properscoring.

Datseris, G., 2018: DynamicalSystems.jl: A Julia software library for chaos and nonlinear dynamics. *J. Open Source Softw.*, **3**, 598, https://doi.org/10.21105/joss.00598.

Dong, L., and F. Zhang, 2016: OBEST: An observation-based ensemble subsetting technique for tropical cyclone track prediction. *Wea. Forecasting*, **31**, 57–70, https://doi.org/10.1175/WAF-D-15-0056.1.

Farmer, J. D., and J. J. Sidorowich, 1987: Predicting chaotic time series. *Phys. Rev. Lett.*, **59**, 845–848, https://doi.org/10.1103/PhysRevLett.59.845.

——, and ——, 1988: Exploiting chaos to predict the future and reduce noise. *Evolution, Learning and Cognition*, Y. C. Lee, Ed., World Scientific, 277–330, https://doi.org/10.1142/9789814434102_0011.

Ghil, M., 2001: Hilbert problems for the geosciences in the 21st century. *Nonlinear Processes Geophys.*, **8**, 211, https://doi.org/10.5194/npg-8-211-2001.

——, and S. Childress, 1987: *Topics in Geophysical Fluid Dynamics: Atmospheric Dynamics, Dynamo Theory, and Climate Dynamics*. Springer, 512 pp., https://doi.org/10.1007/978-1-4612-1052-8.

——, and P. Malanotte-Rizzoli, 1991: Data assimilation in meteorology and oceanography. *Advances in Geophysics*, Vol. 33, Elsevier, 141–266, https://doi.org/10.1016/S0065-2687(08)60442-2..

——, and R. Vautard, 1991: Interdecadal oscillations and the warming trend in global temperature time series. *Nature*, **350**, 324–327, https://doi.org/10.1038/350324a0.

——, and N. Jiang, 1998: Recent forecast skill for the El Niño/Southern Oscillation. *Geophys. Res. Lett.*, **25**, 171–174, https://doi.org/10.1029/97GL03635.

——, and A. W. Robertson, 2000: Solving problems with GCMs: General circulation models and their role in the climate modeling hierarchy. *General Circulation Model Development: Past, Present, and Future*, D. A. Randall, Ed., Academic Press, 285–325.

——, and ——, 2002: "Waves" vs. "particles" in the atmosphere's phase space: A pathway to long-range forecasting? *Proc. Natl. Acad. Sci. USA*, **99**, 2493–2500, https://doi.org/10.1073/pnas.012582899.

——, and Coauthors, 2002: Advanced spectral methods for climatic time series. *Rev. Geophys.*, **40**, 1003, https://doi.org/10.1029/2000RG000092.

——, D. Kondrashov, F. Lott, and A. W. Robertson, 2004: Intraseasonal oscillations in the mid-latitudes: Observations, theory and GCM results. *ECMWF/CLIVAR Workshop on Simulation and Prediction of Intra-Seasonal Variability with Emphasis on the MJO*, Reading, United Kingdom, ECMWF, 35–54.

——, A. Groth, D. Kondrashov, and A. W. Robertson, 2019: Extratropical sub-seasonal to seasonal oscillations and multiple regimes: The dynamical systems view. *Sub-Seasonal to*

Seasonal Prediction: The Gap between Weather and Climate Forecasting, A. W. Robertson, and F. Vitart, Eds., Elsevier, 119–142, https://doi.org/10.1016/B978-0-12-811714-9.00006-1.

Golyandina, N., 2020: Particularities and commonalities of singular spectrum analysis as a method of time series analysis and signal processing. *Wiley Interdiscip. Rev. Comput. Stat.*, **12**, e1487, https://doi.org/10.1002/wics.1487.

——, V. Nekrutkin, and A. A. Zhigljavsky, 2001: *Analysis of Time Series Structure: SSA and Related Techniques*. 1st ed., Chapman and Hall/CRC, 320 pp.

Groth, A., and M. Ghil, 2011: Multivariate singular spectrum analysis and the road to phase synchronization. *Phys. Rev. E*, **84**, 036206, https://doi.org/10.1103/PhysRevE.84.036206.

——, and ——, 2015: Monte Carlo singular spectrum analysis (SSA) revisited: Detecting oscillator clusters in multivariate datasets. *J. Climate*, **28**, 7873–7893, https://doi.org/10.1175/JCLI-D-15-0100.1.

——, and ——, 2017: Synchronization of world economic activity. *Chaos*, **27**, 127002, https://doi.org/10.1063/1.5001820.

Harris, T. J., and H. Yuan, 2010: Filtering and frequency interpretations of singular spectrum analysis. *Physica D*, **239**, 1958–1967, https://doi.org/10.1016/j.physd.2010.07.005.

Hasselmann, K., 1988: PIPs and POPs: The reduction of complex dynamical systems using principal interaction and oscillation patterns. *J. Geophys. Res.*, **93**, 11 015–11 021, https://doi.org/10.1029/JD093iD09p11015.

Hersbach, H., 2000: Decomposition of the continuous ranked probability score for ensemble prediction systems. *Wea. Forecasting*, **15**, 559–570, https://doi.org/10.1175/1520-0434(2000)015<0559:DOTCRP>2.0.CO;2.

Jolliffe, I. T., 2002: *Principal Component Analysis*. 2nd ed. Springer-Verlag, 488 pp., https://doi.org/10.1007/b98835.

Kalnay, E., 2002: *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, 341 pp.

——, 2019: Historical perspective: Earlier ensembles and forecasting forecast skill. *Quart. J. Roy. Meteor. Soc.*, **145**, 25–34, https://doi.org/10.1002/qj.3595.

——, and S.-C. Yang, 2010: Accelerating the spin-up of ensemble Kalman filtering. *Quart. J. Roy. Meteor. Soc.*, **136**, 1644–1651, https://doi.org/10.1002/qj.652.

Kang, I.-S., and H.-M. Kim, 2010: Assessment of MJO predictability for boreal winter with various statistical and dynamical models. *J. Climate*, **23**, 2368–2378, https://doi.org/10.1175/2010JCLI3288.1.

Kennedy, M., 1994: Chaos in the Colpitts oscillator. *IEEE Trans. Circuits Syst.*, **41**, 771–774, https://doi.org/10.1109/81.331536.

Keppenne, C. L., and M. Ghil, 1993: Adaptive filtering and prediction of noisy multivariate signals: An application to subannual variability in atmospheric angular momentum. *Int. J. Bifurcation Chaos*, **03**, 625–634, https://doi.org/10.1142/S0218127493000520.

Kondrashov, D., M. D. Chekroun, A. W. Robertson, and M. Ghil, 2013: Low-order stochastic model and "past-noise forecasting" of the Madden–Julian oscillation. *Geophys. Res. Lett.*, **40**, 5305–5310, https://doi.org/10.1002/grl.50991.

——, ——, and M. Ghil, 2015: Data-driven non-Markovian closure models. *Physica D*, **297**, 33–55, https://doi.org/10.1016/j.physd.2014.12.005.

——, ——, X. Yuan, and M. Ghil, 2018: Data-adaptive harmonic decomposition and stochastic modeling of Arctic sea ice. *Advances in Nonlinear Geosciences*, A. A. Tsonis, Ed., Springer, 179–205, https://doi.org/10.1007/978-3-319-58895-7_10.

Kooperberg, C., and F. O'Sullivan, 1996: Predictive oscillation patterns: A synthesis of methods for spatial-temporal de-

composition of random fields. *J. Amer. Stat. Assoc.*, **91**, 1485–1496, https://doi.org/10.1080/01621459.1996.10476716.

Kravtsov, S., D. Kondrashov, and M. Ghil, 2010: Empirical model reduction and the modeling hierarchy in climate dynamics and the geosciences. *Stochastic Physics and Climate Modelling*, T. N. Palmer and P. Williams, Eds., Cambridge University Press, 35–72.

Krishnamurthy, V., 2019: Predictability of weather and climate. *Earth Space Sci.*, **6**, 1043–1056, https://doi.org/10.1029/2019EA000586.

——, and J. Shukla, 2007: Intraseasonal and seasonally persisting patterns of Indian monsoon rainfall. *J. Climate*, **20**, 3–20, https://doi.org/10.1175/JCLI3981.1.

——, and A. S. Sharma, 2017: Predictability at intraseasonal time scale. *Geophys. Res. Lett.*, **44**, 8530–8537, https://doi.org/10.1002/2017GL074984.

Lau, W. K. M., 2012: El Niño Southern Oscillation connection. *Intraseasonal Variability in the Atmosphere–Ocean Climate System*, 2nd ed. W. K. M. Lau and D. E. Waliser, Eds., Springer-Verlag, 297–334, https://doi.org/10.1007/978-3-642-13914-7.

Lien, G.-Y., E. Kalnay, and T. Miyoshi, 2013: Effective assimilation of global precipitation: Simulation experiments. *Tellus*, **65A**, 19915, https://doi.org/10.3402/tellusa.v65i0.19915.

Lisi, F., O. Nicolis, and M. Sandri, 1995: Combining singular-spectrum analysis and neural networks for time series forecasting. *Neural Process. Lett.*, **2**, 6–10, https://doi.org/10.1007/BF02279931.

Lorenz, E. N., 1963: Deterministic nonperiodic flow. *J. Atmos. Sci.*, **20**, 130–141, https://doi.org/10.1175/1520-0469(1963)020<0130:DNF>2.0.CO;2.

——, 1965: A study of the predictability of a 28-variable atmospheric model. *Tellus*, **17**, 321–333, https://doi.org/10.3402/tellusa.v17i3.9076.

——, 1969: Atmospheric predictability as revealed by naturally occurring analogues. *J. Atmos. Sci.*, **26**, 636–646, https://doi.org/10.1175/1520-0469(1969)26<636:APARBN>2.0.CO;2.

Lynch, E. M., 2019: Data driven prediction without a model. Doctoral thesis, University of Maryland, College Park, 107 pp., https://doi.org/10.13016/quty-dayf.

Mani, N. J., E. Suhas, and B. N. Goswami, 2009: Can global warming make Indian monsoon weather less predictable? *Geophys. Res. Lett.*, **36**, L08811, https://doi.org/10.1029/2009GL037989.

Mann, M. E., and J. Park, 1999: Oscillatory spatiotemporal signal detection in climate studies: A multiple-taper spectral domain approach. *Advances in Geophysics*, Vol. 41, Elsevier, 1–131, https://doi.org/10.1016/S0065-2687(08)60026-6.

——, B. A. Steinman, and S. K. Miller, 2020: Absence of internal multidecadal and interdecadal oscillations in climate model simulations. *Nat. Commun.*, **11**, 49, https://doi.org/10.1038/s41467-019-13823-w.

Mo, K. C., 2001: Adaptive filtering and prediction of intraseasonal oscillations. *Mon. Wea. Rev.*, **129**, 802–817, https://doi.org/10.1175/1520-0493(2001)129<0802:AFAPOI>2.0.CO;2.

Moron, V., A. W. Robertson, and M. Ghil, 2012: Impact of the modulated annual cycle and intraseasonal oscillation on daily-to-interannual rainfall variability across monsoonal India. *Climate Dyn.*, **38**, 2409–2435, https://doi.org/10.1007/s00382-011-1253-4.

Ogrosky, H. R., S. N. Stechmann, N. Chen, and A. J. Majda, 2019: Singular spectrum analysis with conditional predictions for real-time state estimation and forecasting. *Geophys. Res. Lett.*, **46**, 1851–1860, https://doi.org/10.1029/2018GL081100.

Ott, E., 2002: *Chaos in Dynamical Systems*. 2nd ed. Cambridge University Press, 487 pp.

Palmer, T., 2020: A vision for numerical weather prediction in 2030. 11 pp., https://arxiv.org/abs/2007.04830.

Penny, S. G., 2017: Mathematical foundations of hybrid data assimilation from a synchronization perspective. *Chaos*, **27**, 126801, https://doi.org/10.1063/1.5001819.

——, E. Bach, K. Bhargava, C.-C. Chang, C. Da, L. Sun, and T. Yoshida, 2019: Strongly coupled data assimilation in multiscale media: Experiments using a quasi-geostrophic coupled model. *J. Adv. Model. Earth Syst.*, **11**, 1803–1829, https://doi.org/10.1029/2019MS001652.

Pentakota, S., and Coauthors, 2020: Advances in coupled data assimilation, ensemble forecasting, and assimilation of altimeter observations. *CLIVAR Exchanges*, No. 79, International CLIVAR Project Office, Southampton, United Kingdom, 27–30.

Plaut, G., and R. Vautard, 1994: Spells of low-frequency oscillations and weather regimes in the Northern Hemisphere. *J. Atmos. Sci.*, **51**, 210–236, https://doi.org/10.1175/1520-0469(1994)051<0210:SOLFOA>2.0.CO;2.

Preisendorfer, R. W., and C. D. Mobley, 1988: *Principal Component Analysis in Meteorology and Oceanography*. Elsevier, 425 pp.

Rasmusson, E. M., P. A. Arkin, W.-Y. Chen, and J. B. Jalickee, 1981: Biennial variations in surface temperature over the United States as revealed by singular decomposition. *Mon. Wea. Rev.*, **109**, 587–598, https://doi.org/10.1175/1520-0493(1981)109<0587:BVISTO>2.0.CO;2.

Rasp, S., and N. Thuerey, 2021: Data-driven medium-range weather prediction with a Resnet pretrained on climate simulations: A new model for WeatherBench. *J. Adv. Model. Earth Syst.*, **13**, e2020MS002405, https://doi.org/10.1029/2020MS002405.

Rey, D., M. Eldridge, U. Morone, H. D. I. Abarbanel, U. Parlitz, and J. Schumann-Bischoff, 2014: Using waveform information in nonlinear data assimilation. *Phys. Rev. E*, **90**, 062916, https://doi.org/10.1103/PhysRevE.90.062916.

Sharma, A. S., D. Vassiliadis, and K. Papadopoulos, 1993: Reconstruction of low-dimensional magnetospheric dynamics by singular spectrum analysis. *Geophys. Res. Lett.*, **20**, 335–338, https://doi.org/10.1029/93GL00242.

Stan, C., and V. Krishnamurthy, 2019: Intra-seasonal and seasonal variability of the Northern Hemisphere extra-tropics. *Climate Dyn.*, **53**, 4821–4839, https://doi.org/10.1007/s00382-019-04827-9.

Strong, C., F. Jin, and M. Ghil, 1995: Intraseasonal oscillations in a barotropic model with annual cycle, and their predictability.

J. Atmos. Sci., **52**, 2627–2642, https://doi.org/10.1175/1520-0469(1995)052<2627:IOIABM>2.0.CO;2.

Ukhorskiy, A. Y., M. I. Sitnov, A. S. Sharma, and K. Papadopoulos, 2004: Global and multi-scale features of solar wind–magnetosphere coupling: From modeling to forecasting. *Geophys. Res. Lett.*, **31**, L08802, https://doi.org/10.1029/2003GL018932.

Van den Dool, H. M., 1994: Searching for analogues, how long must we wait? *Tellus*, **46A**, 314–324, https://doi.org/10.3402/tellusa.v46i3.15481.

Vannitsem, S., D. S. Wilks, and J. W. Messner, Eds., 2018: *Statistical Postprocessing of Ensemble Forecasts*. Elsevier, 362 pp., https://doi.org/10.1016/C2016-0-03244-8.

Vautard, R., and M. Ghil, 1989: Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series. *Physica D*, **35**, 395–424, https://doi.org/10.1016/0167-2789(89)90077-8.

——, P. Yiou, and M. Ghil, 1992: Singular-spectrum analysis: A toolkit for short, noisy chaotic signals. *Physica D*, **58**, 95–126, https://doi.org/10.1016/0167-2789(92)90103-T.

——, C. Pires, and G. Plaut, 1996: Long-range atmospheric predictability using space–time principal components. *Mon. Wea. Rev.*, **124**, 288–307, https://doi.org/10.1175/1520-0493(1996)124<0288:LRAPUS>2.0.CO;2.

von Storch, H., and F. W. Zwiers, 1999: *Statistical Analysis in Climate Research*. Cambridge University Press, 484 pp., https://doi.org/10.1017/CBO9780511612336.

Wang, X. L., J. Corte-Real, and X. Zhang, 1996: Intraseasonal oscillations and associated spatial-temporal structures of precipitation over China. *J. Geophys. Res.*, **101**, 19 035–19 042, https://doi.org/10.1029/96JD01225.

Watari, S., 1996: Separation of periodic, chaotic, and random components in solar activity. *Sol. Phys.*, **168**, 413–422, https://doi.org/10.1007/BF00148065.

Weare, B. C., and J. S. Nasstrom, 1982: Examples of extended empirical orthogonal function analyses. *Mon. Wea. Rev.*, **110**, 481–485, https://doi.org/10.1175/1520-0493(1982)110<0481:EOEEOF>2.0.CO;2.

Wilks, D. S., 2019: *Statistical Methods in the Atmospheric Sciences*. 4th ed. Elsevier, 840 pp., https://doi.org/10.1016/C2017-0-03921-6.

Xavier, P. K., and B. N. Goswami, 2007: An analog method for real-time forecasting of summer monsoon subseasonal variability. *Mon. Wea. Rev.*, **135**, 4149–4160, https://doi.org/10.1175/2007MWR1854.1.