Mathematical foundations of hybrid data assimilation from a synchronization perspective

Stephen G. Penny

View Online   **Export Citation**   **CrossMark**

ARTICLES YOU MAY BE INTERESTED IN

A unifying view of synchronization for data assimilation in complex nonlinear networks
Chaos: An Interdisciplinary Journal of Nonlinear Science **27**, 126802 (2017); https://doi.org/10.1063/1.5001816

"FORCE" learning in recurrent neural networks as data assimilation
Chaos: An Interdisciplinary Journal of Nonlinear Science **27**, 126804 (2017); https://doi.org/10.1063/1.4990730

Using machine learning to replicate chaotic attractors and calculate Lyapunov exponents from data
Chaos: An Interdisciplinary Journal of Nonlinear Science **27**, 121102 (2017); https://doi.org/10.1063/1.5010300
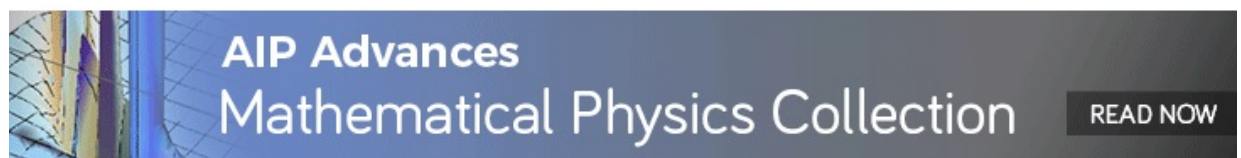
AIP Advances
Mathematical Physics Collection
READ NOW

AIP Publishing

Mathematical foundations of hybrid data assimilation from a synchronization perspective

Stephen G. Penny

Department of Atmospheric and Oceanic Science, University of Maryland, College Park, Maryland 20742, USA; National Centers for Environmental Prediction (NCEP), College Park, Maryland 20740, USA; and RIKEN Advanced Institute for Computational Science, Kobe, Hyogo 650-0047, Japan

The state-of-the-art data assimilation methods used today in operational weather prediction centers around the world can be classified as generalized one-way coupled impulsive synchronization. This classification permits the investigation of hybrid data assimilation methods, which combine dynamic error estimates of the system state with long time-averaged (climatological) error estimates, from a synchronization perspective. Illustrative results show how dynamically informed formulations of the coupling matrix (via an Ensemble Kalman Filter, EnKF) can lead to synchronization when observing networks are sparse and how hybrid methods can lead to synchronization when those dynamic formulations are inadequate (due to small ensemble sizes). A large-scale application with a global ocean general circulation model is also presented. Results indicate that the hybrid methods also have useful applications in generalized synchronization, in particular, for correcting systematic model errors. *Published by AIP Publishing.* [http://dx.doi.org/10.1063/1.5001819]

Data assimilation is a mathematical discipline that arose out of the development of numerical weather prediction (NWP), which required initialization of numerical models based on a set of relatively sparse observations. In recent years, popular solution approaches for this state estimation problem have evolved into two main categories: variational methods and ensemble-based Kalman filters. As each approach has its own unique benefits, researchers began creating hybrid combinations of these methods to take advantage of the strengths of both. We describe how data assimilation can be interpreted as a type of synchronization problem in which a modeled system is driven by observations of a natural system and extend this formalism to include the aforementioned hybrid data assimilation techniques.

INTRODUCTION

Data assimilation can broadly be defined as the mathematical discipline dedicated to the optimal reconciliation of a theoretical model with observed data—a primary application being state estimation of an imperfectly known dynamical system. It is often desired to make predictions using a theoretical model that must be initialized using information about the current state of a dynamical system. We assume that the dynamical system under investigation may be queried via measurement; however, measurements may be sparse, due to cost and technological limitations, and are subject to error. To make accurate predictions, one must utilize all available information from current and past measurements, incorporate understanding of the dynamical system via a theoretical model, and account for all errors present in the forecasting process. At major operational weather forecasting centers around the world, hybrid data assimilation methods that combine a climatological (time-averaged) estimate of forecast errors with a dynamic (time-varying) estimate have been adopted as the primary approach for initializing numerical weather prediction (NWP) models.

It has been suggested (Duane *et al.*, 2006; Yang *et al.*, 2006; Carrassi *et al.*, 2008; and Abarbanel *et al.*, 2010) that synchronization may be a useful mechanism to explore for applications to data assimilation. NWP represents a relatively general class of synchronization problem, and we suggest that studying its successes in the prediction of geophysical systems may benefit the general nonlinear dynamics community. In turn, we hope that the nonlinear dynamics community may help to elucidate the behaviors of operational data assimilation systems used for operational NWP.

To help bridge these two communities, we attempt to form a rigorous definition of data assimilation with respect to synchronization. We interpret sequential data assimilation as a form of generalized (Rulkov *et al.*, 1995 and Hunt *et al.*, 1997) one-way coupled impulsive synchronization in which the following statements hold:

(1) The driver (nature) system $f^{(N)}$ and its state are unknown. The driver system is typically a natural system (e.g., for example, a geophysical system such as the ocean, atmosphere, or complete Earth system).

(2) The response (model) system $f^{(M)}$ is a numerical model that is capable of resolving and forecasting a selection of spatiotemporal scales of the nature system.

(3) The response (model) state $\mathbf{x}^{(M)}$ is related to the driver (nature) state $\mathbf{x}^{(N)}$ by an unknown function, $\mathbf{x}^{(M)} = \phi(\mathbf{x}^{(N)})$. In NWP, estimating this relationship is the critical role of the forecaster and is related to "post-processing" methodologies such as Model Output Statistics (MOS) (Glahn and Lowry, 1972).

(4) There is a measurement operator $H^{(N)}$ that maps the nature state to the observation space $\mathbf{y}^o = H^{(N)}(\mathbf{x}^{(N)}, \eta^o)$, which has possible errors and biases represented by parameter η^o. A

physical sensor device typically performs this operation and may be satellite-based or *in situ*.

(5) There is an observation operator $H^{(M)}$ that maps the model state $\mathbf{x}^{(M)}$ to the observation space $\mathbf{y}^b = H^{(M)}(\mathbf{x}^{(M)}, \eta^\beta)$, which has possible errors and biases represented by parameter η^β.

(6) Using a Gaussian approximation, there are statistical covariance estimates for the error characteristics of the forecast (**P**), model (**Q**), and observations (**R**), while the true error characteristics are unknown.

The last point is perhaps the largest departure from traditional synchronization studies. The nonlinear error dynamics are encoded into the forecast error covariance matrix to enable coupling of a potentially sparsely observed driver system with a numerical model as the response system. In NWP, this forecast error covariance information is either estimated from a long time-averaged history of the system's forecast errors (i.e., a climatology) typically denoted as **B**, produced adaptively to estimate the instantaneous "errors of the day" (Kalnay, 2003) typically denoted as $\mathbf{P}^b$, or some combination of the two (Hamill and Snyder, 2000; Wang *et al.*, 2007a, 2007b, 2008a, 2008b, 2010, 2013; Kleist 2012; Penny, 2014; Penny *et al.*, 2015; Hamrud *et al.*, 2014; and Bonavita *et al.*, 2015). Such methods that combine static and dynamic error representations are typically referred to as hybrid methods and have recently been reviewed by Asch *et al.* (2017) and Bannister (2017). The nonlinear dynamics of the spatially extended response system are assumed to be reasonably well known but are typically under-resolved, while the sub-gridscale physics is parameterized (i.e., approximated with simple parametric models usually dependent on the resolved scales) in order to have some representation of processes that cannot be resolved explicitly.

The measurement and observation operators are possibly non-invertible and nonlinear. For example, the measurement operator $H^{(N)}$ may represent a satellite-based passive sensor, while the observation operator $H^{(M)}$ is realized by a numerical model (e.g., the Community Radiative Transfer Model; Han *et al.*, 2006) designed to emulate the measurement operator $H^{(N)}$. The data assimilation system is synchronized on the observation space—an idealized space in which both the measurement operator $H^{(N)}$ and observation operator $H^{(M)}$ attempt to form equivalent representations. The characteristics of the observation space are essential for the stability of the system; they determine the "observability" of the system. For example, Whartenby *et al.* (2013) began to address this issue for shallow water equations, Bocquet *et al.* (2017) provided the analytic connection between observability in the unstable-neutral subspace and the filter stability for deterministic linear dynamics, while the nonlinear deterministic case is investigated numerically by Carrassi *et al.* (2008). Ni and Zhang (2016) described the relationship between observability and filter stability for a deterministic linear system.

For a number of practical reasons, sequential data assimilation methods used for operational NWP typically apply an impulse at constant time intervals (called an analysis cycle). These intervals are determined primarily by the delay in collecting and processing observations at the operational center and by computational limitations on the forecast frequency,

typically performed at regular intervals throughout the day (e.g., every 6 h). Incremental analysis updates (Bloom *et al.*, 1996) are a common technique used to spread the impact of the analysis update into smaller constituent parts at each model timestep. Such incremental updates are quite similar to those typically applied in synchronization studies; however, they are held constant over multiple model timesteps until recomputed by a new analysis. Recently, there have been efforts to extend this method to provide some time dependence to the increments.

A straightforward synchronization application might be to couple a particular set of gridpoints of the driver and response systems (e.g., Kostuk, 2012). With imperfect models, noisy observations, and potentially sparse observing networks, direct insertion of observations is not a viable option for NWP in practice (Kalnay, 2003). We must include as much dynamical information as possible from the model to ensure that all information from the observations is utilized, while still maintaining a state that is consistent with the numerical model. Such was the motivation for iterative techniques such as the "Running in Place" method (Yang *et al.*, 2012 and Penny *et al.*, 2013).

However, through a series of simplifying assumptions, we can recover a traditional (Pecora and Carroll, 1990 and Pecora *et al.*, 1997) one-way coupled synchronization:

(1) The driver and response systems are identical so that $f = f^{(N)} = f^{(M)}$ (i.e., a "perfect model" scenario),
(2) The measurement operator $H^{(N)}$ and observation operator $H^{(M)}$ are identical (i.e., $H = H^{(N)} = H^{(M)}$),
(3) The observational noise is zero,
(4) The estimated statistical error covariance matrices are diagonal (i.e., errors are spatially uncorrelated).

The history of operational NWP has seen an evolution from direct nudging to observations, to optimal interpolation (OI) and the 3D-variational (3DVar) method, to the 4D-variational (4DVar) and ensemble Kalman filter (EnKF) methods, and most recently to various hybrid methods combining elements of both variational and ensemble methods. Kalnay (2003) and Ghil and Malanotte-Rizzoli (1991) provided thorough reviews of this history, supplemented by Asch *et al.* (2017) and Bannister (2017) for more recent developments. We give a synopsis here for context.

An early data assimilation approach called "nudging" (e.g., see Kistler, 1974 and Hoke and Anthes, 1976) is very similar to the direct insertion of observations used in many synchronization studies. Nudging adds a forcing term to the prognostic equations, which relaxes the dynamics toward the observed value with a timescale specified by a tuning parameter. Statistical estimation theory inspired later developments in DA, leading to a collection of multivariate least squares methods that remain the most commonly used today. These are generally aimed at minimizing a cost functional at each time step t_i derived from a multivariate Gaussian distribution, such as

$$J(\mathbf{x}_i) = \frac{1}{2}(\mathbf{x}_i - \mathbf{x}_i^f)^T \mathbf{P}_i^{-1}(\mathbf{x}_i - \mathbf{x}_i^f)$$
$$+ \frac{1}{2}(\mathbf{y}_i^o - H_i(\mathbf{x}_i))^T \mathbf{R}_i^{-1}(\mathbf{y}_i^o - H(\mathbf{x}_i)). \quad (1)$$

The techniques for minimizing (1) can be broadly categorized into "sequential" and "variational" methods (Lorenc, 1986), with the former achieved by a direct algebraic solution (e.g., Optimal Interpolation (OI) and variants of the Kalman Filter; Kalman,1960) and the latter achieved by a numerical minimization of the cost functional (Sasaki, 1970). For a number of years, these two alternative branches developed somewhat independently (and competitively). The OI and 3D-Variational (3DVar) methods in their basic formulation are algebraically equivalent (see Kalnay, 2003 for proof); however, implementation details introduce meaningful differences in practice. All approaches analyze the forecast and observations at a particular time or in a time range to determine an optimal combination of the two sets of information given their respective assumed known (but in practice estimated) errors.

The variational methods (Le Dimet and Talagrand, 1986) were improved by allowing observations to be analyzed not only at one time per analysis but instead over a longer time window. Thus, synoptic weather oscillations observed over 12–24 h could be fitted to a specific model trajectory that would ideally minimize the disagreement with all available observations within that window. This 4D-Variational (4DVar; e.g., Talagrand and Coutier, 1987 and Thepaut and Courtier, 1991) method has been very successful, with the caveats that finding the trajectory that minimizes Eq. (1) over a time window requires (a) the gradient of the cost functional and thus the development of a Tangent Linear Model (TLM; a numerical approximation of the linear evolution operator) of the entire NWP model and (b) multiple integrations of this model must be repeated over the same time period as part of the minimization algorithm. While the European Centre for Medium-Range Weather Forecasts (ECMWF) has dedicated resources to maintain a TLM for their operational NWP model, thus enabling the use of 4DVar for operational forecasts, it is not common practice for developers of large-scale geophysical models to maintain a TLM.

The sequential methods progressed from OI, which uses a fixed forecast error covariance estimate, to the Kalman filter, which uses an evolving forecast error covariance estimate. This essentially allows for an adaptive weighting of the synchronization strength that depends on anticipated forecast accuracy. Because the nonlinear adaptation of the Kalman filter, the extended Kalman filter (EKF), was too expensive to apply with the state-of-the-art numerical weather models, a Monte-Carlo approach was introduced called the ensemble Kalman filter (EnKF; Evensen, 1994). Early applications of the EnKF used multiple concurrent but independent DA cycles with each one assimilating randomly perturbed observations. Later, deterministic methods were developed that formed the new initial states of the ensemble using a linear transform applied to the forecast states (Bishop *et al.*, 2001 and Whitaker and Hamill, 2002).

Ensemble methods like the EnKF have been popular due to their ease of use, particularly in that they automatically generate the forecast error covariance information relating unobserved variables to observed variables. There are two major caveats to the use of EnKFs, however. The first is the computational cost—they typically require the integration of a minimum O(10)-O(100) simultaneous instances of the model. As the model dimension of operational NWP systems can exceed O(10^9), this point in nontrivial. The second caveat is the susceptibility to model deficiencies. Systematic errors in the numerical model can lead to biases in long-term error statistics and thus are often called "model biases." The EnKF uses an ensemble of model states to represent the first two statistical moments of the forecast error. If the model has a systematic error, then all members may closely cluster around an inaccurate state, thereby underestimating the model error and ignoring meaningful corrections that should be made due to observations.

Given the strengths and weaknesses of the EnKF and 4DVar methods, there were a number of years of debate over which one would prove to be the superior approach (Lorenc, 2003 and Kalnay *et al.*, 2007). Through that debate, it became clear that neither approach on its own would be sufficient, but rather the strengths of both should be leveraged in hybrid methods (Lorenc 2003). Such strengths included the dynamically varying estimate of the EnKF forecast error covariance and the carefully tuned climatological forecast error covariance estimate used for the variational methods.

Here, we focus on a widely used deterministic EnKF variant, the Local Ensemble Transform Kalman Filter (LETKF; Hunt *et al.*, 2007), and a hybridization method called Hybrid-Gain (Penny 2014, 2015; Hamrud *et al.*, 2014; and Bonavita *et al.*, 2015). We will examine the role the components of the Hybrid-Gain method play in stabilizing the filter. Trevisan and Palatella (2011) conjectured that the minimum required ensemble size to prevent filter divergence is equal to the number of unstable and neutral Lyapunov exponents. Ng *et al.* (2011) extended the proposition by Trevisan and Palatella (2011) to account for the dynamics of the system under investigation, indicating that if exposed to noise, an EnKF must account for not only the unstable and neutral Lyapunov modes but also a subset of stable modes that have transient unstable dynamics that could project onto the unstable and neutral modes as a result of nonlinear interactions. Palatella and Trevisan (2015) further extended the work by Trevisan and Palatella (2011) to consider the effects of nonlinear deviations from the linear assumption of the EKF, related to the results from the study by Ng *et al.* (2011). Recently, Bocquet and Carrassi (2017) indicated that the ETKF belongs to the class of deterministic EnKFs for which the relationship between error convergence and the unstable-subspace holds, and provided a proof of the relationship between the required minimum ensemble size and the number of nonnegative Lyapunov modes for linear systems. It was shown that the analysis error drops to its minimum when a full geometrical alignment is achieved between the EnKF perturbations and the unstable subspace.

Samoilenko and Perestyuk (1995) presented stability results in the context of nonlinear impulsive differential equations. Given a nonlinear impulsive dynamical system, simple constraints (essentially a global Lipschitz condition) on the nonlinear part are sufficient to allow the linearized dynamics to determine the asymptotic stability of the system (Thm. 37, Samoilenko and Perestyuk, 1995). We explore

this first in the context of a simplified model and then consider the applicability of these results to a more realistic high-dimensional geophysical system as used in an operational environment.

We make one final distinction. Although the traditional definition of synchronization typically assumes that the driver and response systems will remain synchronized beyond the coupling period, we only require synchronization during a finite forecast period. In geophysical data assimilation applications for example, it is common to examine the forecasts from one analysis cycle (impulse interval) to the next. This is due to the sensitive dependence on initial conditions (Lorenz 1963 and Lorenz 1969), which cannot be known with complete accuracy.

MATHEMATICAL FOUNDATIONS

We commence with a coupled nonlinear dynamical system that includes the driver/nature (N) dynamics

$$\frac{d\mathbf{x}}{dt} = f^{(N)}(t, \mathbf{x}(t)), \tag{2}$$

and the response/model (M) dynamics

$$\begin{cases} \frac{d\mathbf{x}}{dt} = f^{(M)}(t, \mathbf{x}(t)) & t \neq \tau_i \\ \Delta\mathbf{x}|_{t=\tau_i} = g_i(\mathbf{x}) & t = \tau_i. \end{cases} \tag{3}$$

The function g is an impulsive control that will be applied to the response system at regular intervals to produce synchronization with the driver system. Without loss of generality, we will study the error system, which is equivalent to the response system with the zero vector as its solution (where $\mathbf{x}$ is defined to be the state in the error phase space, i.e., $\mathbf{x} = \mathbf{x}^{(N)} - \mathbf{x}^{(M)}$). This may be formed from the nonlinear system by the change in coordinates (p. 120, Samoilenko and Perestyuk, 1995).

If we assume that $\mathbf{x}$ is a small perturbation from the true state determined by the driver system and apply a Taylor series decomposition evaluated at $\mathbf{0}$, then we have

$$f(t, \mathbf{x}(t)) = \mathbf{A}(t)\mathbf{x} + \hat{f}(t, \mathbf{x}(t)), \tag{4}$$

where $\mathbf{A} = \mathrm{D}f$ is the Jacobian matrix of f. We then separate the response system into terms representing its linear and nonlinear parts

$$\begin{cases} \frac{d\mathbf{x}}{dt} = \mathbf{A}(t)\mathbf{x} + \hat{f}(t, \mathbf{x}(t)) & t \neq \tau_i \\ \Delta\mathbf{x}|_{t=\tau_i} = \tilde{\mathbf{K}}_i\mathbf{x} + \hat{g}_i(\mathbf{x}) & t = \tau_i. \end{cases} \tag{5}$$

Next, we consider the first order approximation of the system (Samoilenko and Perestyuk, 1995)

$$\begin{cases} \frac{d\mathbf{x}}{dt} = \mathbf{A}(t)\mathbf{x} & t \neq \tau_i \\ \Delta\mathbf{x}|_{t=\tau_i} = \tilde{\mathbf{K}}_i\mathbf{x} & t = \tau_i. \end{cases} \tag{6}$$

Let $\mathbf{X}$ be the fundamental matrix solution of the linearized driver system

$$\frac{d\mathbf{X}}{dt} = \mathbf{A}(t)\mathbf{X}. \tag{7}$$

The evolution operator Φ is a mapping of solutions from one time to the next

$$\mathbf{X}(t) = \Phi(t, t_0)\mathbf{X}(t_0). \tag{8}$$

The nonlinear system is asymptotically stable as long as the linearized system is asymptotically stable, and the nonlinearities in the system are not too large. For example, Samoilenko and Perestyuk (1995, Thm. 37, p. 120) gave the condition that, if

$$\|\Phi(t, t_i)\| \leq Le^{-\gamma(t-t_i)}, \quad \text{with } L \geq 1 \quad \text{and} \quad \gamma > 0, \tag{9}$$

for all t and t_i, $t_0 \leq t_i \leq t$, and the nonlinear parts are bound so that

$$\|\hat{f}(t, \mathbf{x})\| \leq a\|\mathbf{x}\|, \quad \|\hat{g}_i(\mathbf{x})\| \leq a\|\mathbf{x}\|, \tag{10}$$

for all $t \geq t_0$ and $i = 1, 2, \ldots$, for $\|\mathbf{x}\| \leq h$ and with $h > 0$, then for sufficiently small constants a, the zero solution of the linearized system is asymptotically stable.

The linear operator $\tilde{\mathbf{K}}$ is the gain applied to the impulsive correction. When $\tilde{\mathbf{K}}$ is a constant diagonal matrix (e.g., Abarbanel *et al.* 1993; Pecora *et al.*, 1997; and Chen *et al.*, 2013), this system describes the traditional one-way coupled synchronization formulations. For example, a basic synchronization formulation was given by Pecora *et al.* (1997), with notation adjusted to match that here as

$$\frac{d\mathbf{x}^{(N)}}{dt} = f(\mathbf{x}^{(N)}), \tag{11}$$

$$\frac{d\mathbf{x}^{(M)}}{dt} = f(\mathbf{x}^{(M)}) + \alpha\tilde{\mathbf{K}}(\mathbf{x}^{(N)} - \mathbf{x}^{(M)}). \tag{12}$$

Let $\mathbf{H}_{l \times m}$ act as an indicator function that identifies which model grid points are observed and then maps those points into the reduced observation dimension. Further, let $\mathbf{C}_{l \times l}$ be a diagonal matrix indicating the coupling strength in the observation space (with subscripts indicating the matrix dimension). If we let $\mathbf{K}_{m \times l} = \mathbf{H}^T\mathbf{C}$, then the matrix $\tilde{\mathbf{K}} = \mathbf{KH}$ is the coupling matrix as described by Pecora *et al.* (1997).

The simplest data assimilation methods used in operational NWP centers, such as 3DVar and OI, use $\tilde{\mathbf{K}}$ that may be idealized as constant but not diagonal. They can be defined using the Kalman gain matrix

$$\mathbf{K} = \mathbf{BH}^T(\mathbf{HBH}^T + \mathbf{R})^{-1}, \tag{13}$$

where $\mathbf{B}$ is a climatological (i.e., time-averaged) forecast (also called the "background") error covariance matrix. Matrix $\mathbf{R}$ is the observation error covariance, and $\mathbf{H}$ is the linearized observation operator. In this case, we have assumed a static observing network.

For a more accurate time-varying estimate, EnKFs use an adaptive scheme where $\mathbf{K}_i$ is the time-dependent Kalman gain and $\mathbf{H}_i$ is the time-dependent linearized observation operator. The Kalman gain of the EnKF at a particular time-step is then defined

$$\mathbf{K}_i = \mathbf{P}_i^b \mathbf{H}_i^T \left(\mathbf{H}_i \mathbf{P}_i^b \mathbf{H}_i^T + \mathbf{R}_i \right)^{-1}. \tag{14}$$

Here, the true forecast error covariance matrix is represented by a reduced-rank estimate (i.e., the ensemble size $k \ll m$ the model dimension)

$$\mathbf{P}_i^b = \frac{1}{(k-1)} \mathbf{X}_i \mathbf{X}_i^T, \tag{15}$$

where $\mathbf{X}_i$ is an $m \times k$ matrix whose columns are perturbations from the ensemble mean state $\bar{\mathbf{x}}$. The ensemble mean is assumed close to the true state so that small perturbations from the mean are representative of the error dynamics within the tangent linear subspace. Matrix $\mathbf{R}_i$ is an estimate of the true observation error covariance matrix, and $\mathbf{H}_i$ is the linearized observation operator for an evolving observing network, all specified at time t_i.

The impulse matrix $\tilde{\mathbf{K}}$ applied to achieve synchronization using the EnKF can thus be given as

$$\tilde{\mathbf{K}}_i = -\mathbf{K}_i \mathbf{H}_i. \tag{16}$$

We note that at each analysis time, the dimensions of the impulse matrices resulting from the EnKF are determined by the model dimension (m), observation count (l), and ensemble size (k), labeled in subscripts

$$\mathbf{K}_{m \times l} \mathbf{H}_{l \times m} = \frac{1}{k-1} \mathbf{X}_{m \times k}^b \mathbf{X}_{k \times m}^{bT} \mathbf{H}_{m \times l}^T$$
$$\times \left(\frac{1}{k-1} \mathbf{H}_{l \times m} \mathbf{X}_{m \times k}^b \mathbf{X}_{k \times m}^{bT} \mathbf{H}_{m \times l}^T + \mathbf{R}_{l \times l} \right)^{-1} \mathbf{H}_{l \times m}. \tag{17}$$

Thus, the rank of $\mathbf{K}_i \mathbf{H}_i$ is at most $\min\{m,l,k\}$, which in practice is typically the ensemble size k. In general, $k \ll l \ll m$, although in conditions with sparse observing networks, it is possible that $l < k$. Further, since $\mathbf{X}^b$ is composed of perturbations from the ensemble mean state, the matrix $\mathbf{X}^b$ (and thus $\mathbf{K}_i \mathbf{H}_i$) can have rank only as large as $(k-1)$.

The most computationally intensive part of the data assimilation process is typically the integration of the forward model, putting a high cost on increasing the rank of $\mathbf{K}_i \mathbf{H}_i$. This has resulted in two successful techniques that have been widely adopted: (1) localization and (2) hybridization. Localization essentially reduces the problem size to a series of manageable sub-domains so that a given ensemble size is sufficient to achieve synchronization. For geophysical problems, the physical justification for this is based on the decorrelation length scales of the fluid. Hybridization essentially increases the effective rank of $\mathbf{K}_i \mathbf{H}_i$ to enable synchronization, even when the ensemble size k is smaller than the number of non-negative Lyapunov exponents.

THE TANGENT LINEAR MODEL

The TLM is used in NWP as a numerical representation of the evolution operator (8), also known as the linear propagator (Kupstov and Parlitz, 2012), state transition matrix

(Balci *et al.*, 2012), or matricient (Samoilenko and Perestyuk, 1995). This evolution operator can be defined as

$$\Phi(t, t_0) = \mathbf{X}(t)\mathbf{X}(t_0)^{-1}, \tag{18}$$

where $\mathbf{X}$ is a fundamental matrix solution, as in (7), and has the following properties:

$$\Phi(t, t) = \mathbf{I}, \forall t, \tag{19}$$

$$\Phi(t_2, t_0) = \Phi(t_2, t_1)\Phi(t_1, t_0), \tag{20}$$

$$\Phi(t_1, t_0) = \Phi^{-1}(t_0, t_1), \tag{21}$$

$$\frac{d\Phi(t, t_0)}{dt} = \mathbf{A}(t)\Phi(t, t_0). \tag{22}$$

The singular values of the matrix Φ determine the time averaged expansion and contraction rates over the interval $[t_0, t]$ along the local Lyapunov vectors (LLVs) defined in that interval.

Numerical approximations to the matrix Φ can be made via series truncation or other means. For example, Balci *et al.* (2012) approximate Φ for some finite number n of intermediate time steps δt using the exponential Euler method. Assuming that $\mathbf{A}_i = \mathbf{A}(t_i)$ is fixed for some small finite time dt

$$\Phi^{(n)}(t, t_0) \approx e^{\mathbf{A}_{n-1}\delta t} e^{\mathbf{A}_{n-2}\delta t} \cdots e^{\mathbf{A}_1 \delta t} e^{\mathbf{A}_0 \delta t}. \tag{23}$$

Because the matrix $\mathbf{A}(t)$ will in general be non-commutative, the exponentiation rules for scalar quantities do not apply, i.e.,

$$e^{\mathbf{A}} e^{\mathbf{B}} \neq e^{\mathbf{A}+\mathbf{B}}. \tag{24}$$

Thus even in the limit,

$$\lim_{n \to \infty} e^{\mathbf{A}_{n-1}\delta t} e^{\mathbf{A}_{n-2}\delta t} \cdots e^{\mathbf{A}_1 \delta t} e^{\mathbf{A}_0 \delta t} \neq e^{\int_{t_0}^t \mathbf{A}(t)dt}. \tag{25}$$

However, Φ can be defined by the generalized Peano-Baker series and may be calculated with reasonable accuracy using geometric numerical integration methods. For example, the Magnus expansion (Magnus, 1954) permits the familiar exponential solution form

$$\Phi(t, t_0) = \exp(\Omega(t)), \tag{26}$$

$$\Omega(t) = \sum_{k=1}^{\infty} \Omega_k(t). \tag{27}$$

Each term of the Magnus series is defined using the matrix commutator $[\mathbf{A}, \mathbf{B}] = \mathbf{AB} - \mathbf{BA}$. The first three terms of the series are

$$\Omega_1 = \int_0^t \mathbf{A}(s_1)ds_1,$$

$$\Omega_2 = \frac{1}{2}\int_0^t ds_1 \int_0^{s_1} ds_2 \left[\mathbf{A}(s_1), \mathbf{A}(s_2)\right], \tag{28}$$

$$\Omega_3 = \frac{1}{6}\int_0^t ds_1 \int_0^{s_1} ds_2 \int_0^{s_2} ds_3 ([\mathbf{A}(s_1), [\mathbf{A}(s_2), \mathbf{A}(s_3)]]$$
$$+ [\mathbf{A}(s_3), [\mathbf{A}(s_2), \mathbf{A}(s_1)]]).$$

Blanes *et al.* (2009) provided a generic recursive definition for all terms Ω_μ of the Magnus expansion.

Evaluating the TLM is useful because it acts as a linear map from one discrete time to the next, emulating the numerical modeling procedure used in data assimilation. Also, it is required by the minimization algorithms used by the 4DVar methods. The data assimilation analysis update (impulse) applied to this system can be expressed by

$$\Phi^+(t_i, t_{i-1}) = (\mathbf{I} - \mathbf{K}_i \mathbf{H}_i)\Phi^-(t_i, t_{i-1}). \tag{29}$$

Trevisan and Uboldi (2004), Uboldi *et al.* (2005), and Carrassi *et al.* (2008) have identified that the stability properties of this type of DA system is dependent on both the Kalman gain matrix K_i and the observing network as defined by H_i. As demonstrated by Ng *et al.* (2011) and later expanded by Bocquet and Carrassi (2017), the stability of the EnKF is affected by whether there are sufficient ensemble members to represent all the unstable and neutral Lyapunov modes (as determined by the covariant Lyapunov vectors with nonnegative Lyapunov exponents). When there is an insufficient number of members, the EnKF experiences "catastrophic filter divergence." It has been found that hybrid methods, which compensate ensemble methods with a climatological component, are able to account for the "missing" representation of the unstable modes by including a static component of full rank (m).

HYBRID-GAIN DATA ASSIMILATION

One can consider a further decomposition of the linearized system into static and time varying component

$$\begin{cases} \dfrac{d\mathbf{x}}{dt} = \bar{\mathbf{A}}\mathbf{x} + \hat{\mathbf{A}}(t)\mathbf{x} & t \neq \tau_i \\ \Delta\mathbf{x}|_{t=\tau_i} = \tilde{\mathbf{K}}^B\mathbf{x} + \tilde{\mathbf{K}}_i^P\mathbf{x} & t = \tau_i, \end{cases} \tag{30}$$

where $\hat{\mathbf{A}}(t)$ is a piecewise continuous matrix. Samoilenko and Perestyuk (1995, Thm. 35, p. 117) showed that solutions of the linearized system are exponentially stable as long as $\|\hat{\mathbf{A}}(t)\| \leq c, \|\tilde{\mathbf{K}}_i^P\| \leq c$, for sufficiently large t and i and sufficiently small c. Thus, if the mean "climatological" dynamics can be constrained by a static gain matrix (e.g., via 3DVar), then the time-varying gain is only required to constrain the climatological anomalies.

In order to improve the stability of the data assimilation system, traditional hybrid methods attempt to modify the gain matrix $\mathbf{K}_i$ via the constituent background error covariance matrix $\mathbf{P}^b$ (Hamill and Snyder, 2000; Wang *et al.*, 2007a, 2007b, 2008a, 2008b, and 2013). A simpler and more direct approach is to apply additional contractions to the evolution operator (Penny 2014; Penny *et al.*, 2015; Hamrud *et al.*, 2014; and Bonavita *et al.*, 2015). For example, after the contraction is determined by the reduced-rank dynamic ensemble estimate of the forecast error covariance $\mathbf{P}^b$, a second contraction can be applied based on a full-rank static climatological estimate $\mathbf{B}$,

$$\Phi^+(t_{i+1}, t_i) = (\mathbf{I} - \mathbf{K}_i^B \mathbf{H}_i)(\mathbf{I} - \mathbf{K}_i^P \mathbf{H}_i)\Phi^-(t_{i+1}, t_i). \tag{31}$$

Through a simple algebraic reformulation, this process can be depicted in a form that allows the consecutive contractions to be represented by a single gain matrix

$$(\mathbf{I} - \mathbf{K}_i^B \mathbf{H}_i)(\mathbf{I} - \mathbf{K}_i^P \mathbf{H}_i) = (\mathbf{I} - (\mathbf{K}_i^P + \mathbf{K}_i^B - \mathbf{K}_i^B \mathbf{H}_i \mathbf{K}_i^P)\mathbf{H}_i). \tag{32}$$

Thus, we have

$$\mathbf{K}_i^{BP} = (\mathbf{K}_i^P + \mathbf{K}_i^B - \mathbf{K}_i^B \mathbf{H}_i \mathbf{K}_i^P). \tag{33}$$

Because each of these component gain matrices are approximations, Penny (2014) allowed a constant scaling to adjust the impact of each term in the assimilation

$$\mathbf{K}_i = (\beta_1 \mathbf{K}_i^P + \beta_2 \mathbf{K}_i^B - \beta_3 \mathbf{K}_i^B \mathbf{H}_i \mathbf{K}_i^P). \tag{34}$$

The stability of the linearized system can be diagnosed by examining the linear propagator applied over the full time horizon of the data assimilation experiment $[t_0, t_N]$, as noted by Trevisan and Uboldi (2004) and Carrassi *et al.* (2008),

$$\Phi(t_N, t_0) = \prod_{i=1}^{N}(\mathbf{I} - \mathbf{K}_i \mathbf{H}_i)\Phi^-(t_i, t_{i-1}). \tag{35}$$

We reiterate that when the filter is divergent for a deterministic EnKF, this indicates that the unstable and neutral modes are not adequately represented in the corresponding reduced-rank gain matrix $\mathbf{K}^P$. Sakov and Oke (2008) and Carrassi *et al.* (2009) provided distinctions of the properties of the stochastic versus deterministic versions of the EnKF. The hybrid gain matrix $\mathbf{K}$ provides a convenient mechanism for representing these missing modes, thus allowing error reduction in these directions. Further, the climatological gain matrix $\mathbf{K}^B$ can be constructed using knowledge of the model errors, thus providing a mechanism for making corrections to systematic errors in the model forecasts where it is expected that the ensemble will not provide a sufficient representation of uncertainty.

EXPERIMENTAL SETUP

We examine results using the Hybrid-Gain method for synchronization in two systems. First, we examine the simple Lorenz (1996) model (L96) with forcing term $F = 8.0$ in reduced dimension ($m = 6$) as explored by Ng *et al.* (2011). We conduct a series of experiments showing the impact of applying impulsive synchronization with a diagonal coupling term, with a coupling term determined via a variational method using static forecast error covariances, with a coupling term informed by correlations determined through the ensemble-derived error covariance matrix, and with a hybridization of static and dynamically varying coupling terms.

All observations are made at the analysis time. Thus, for simplicity, we chose 3DVar as the variation method. If observations are taken over an interval between analysis times, the higher-dimensional extension 4DVar has generally been shown to be more accurate (Lorenc and Rawlins, 2005 and Yang *et al.*, 2009). Bonavita *et al.* (2015) showed
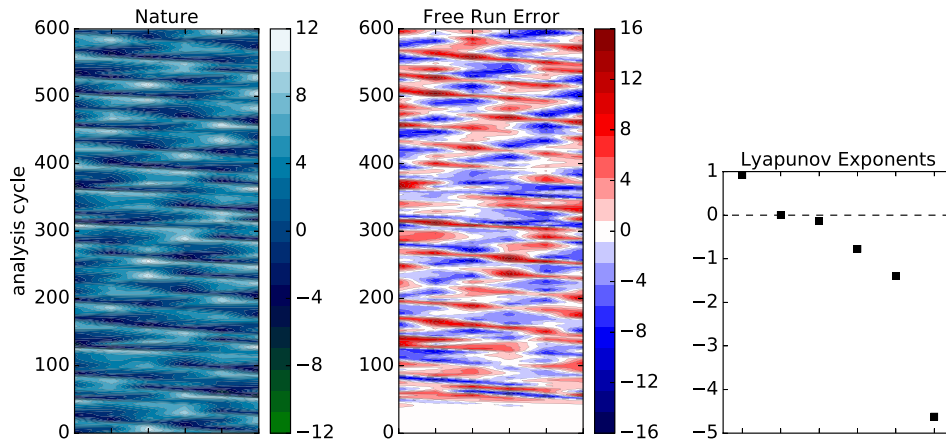
FIG. 1. (a) Nature run for the L96 system with parameters $F = 8.0$ and $m = 6$, from which observations were sampled, (b) error from a free run started with a small perturbation from the true state, and (c) numerically estimated Lyapunov exponents for the L96 system over 1000 time units. The y-axes in (a) and (b) indicate that the analysis cycles 0 through 600 (covering the first 30 time units).

performance gains with the Hybrid-Gain method combining 4DVar and LETKF compared to using either alone when applied to the operational forecast system of the European Centre for Medium-Range Weather Forecasts (ECMWF).

The 3DVar static forecast error covariance **B** is assigned as in Penny (2014) as a unit diagonal matrix with exponential decay on the off-diagonals with a radius of two grid points. We note that for stability in the conjugate gradient minimization used by 3DVar, the initial guess for the solution should be the forecast state. For the traditional synchronization approach using a diagonal coupling matrix, we use as a constant multiplier the ratio of the forecast error variance to the observation error variance. Due to the small system size, we use LETKF without localization and designate this ETKF. For the hybrid method, we use $\alpha = 0.5$.

We first assume that the observation network is fixed so that $\mathbf{H}_i = \mathbf{H}$ and $\mathbf{R}_i = \mathbf{R}$. Thus, the only source of the time dependence that we investigate here is that of the dynamics, indicated by the background error covariance matrix $\mathbf{P}^b_i$. In later experiments, we relax that constraint and allow observation locations to vary randomly so that $\mathbf{H}_i$ is time-dependent. Observational error is assigned from a Gaussian distribution using standard deviation $\sigma_r = 0.01$ to minimize the impacts of observational noise as in Ng et al. (2011). We examine three cases: (1) using an observing network with full spatial coverage, (2a) a case using a sparse stationary observing network, and (2b) a case with a sparse but non-stationary observing network. The Hybrid-Gain method has been demonstrated with L96 using larger observational error ($\sigma_r = 0.5$), larger system dimensions ($m = 40$), and larger forcing ($F = 20.0$) by Penny (2014).

Next, we examine the bias-correction effects in a realistic global ocean model, namely, the Geophysical Fluid Dynamics Laboratory (GFDL) Modular Ocean Model (MOM4p1) that is used by NCEP for climate forecasts (Saha et al., 2006, 2010, 2014). Briefly, it is a hybrid combination of the LETKF-based gain and the NCEP operational 3DVar-based gain applied to a $1/2°$ global ocean model. For the results presented, the system assimilates observed vertical profiles of temperature and salinity, primarily available from the Argo float program (Roemmich and Owens, 2000). The LETKF component of the hybrid global ocean data

assimilation system uses 56 ensemble members, each forced with atmospheric surface forcing perturbations from the 20th-Century Reanalysis (Compo et al., 2011), which are re-centered at NCEP/DOE (Department of Energy) Reanalysis 2 (R2; Kanamitsu et al., 2002). The stand-alone 3DVar system uses the same model forced by the R2. Further details of the system configuration are provided by Penny et al. (2015).

RESULTS

Lorenz-96

We examine the L96 model with forcing term $F = 8.0$ in its reduced dimension ($m = 6$). The system has 1 unstable and 1 neutral Lyapunov mode, with the remaining 4 modes being asymptotically stable (Fig. 1). In a given analysis cycle, there may be any number of transient finite-time Lyapunov exponents that are nonnegative (unstable or neutral).

We list the characteristics of the six-dimensional (6D) L96 system over a fixed time period in Table I and indicate our expectations for the behavior of the system when we apply data assimilation. The starting point on the attractor was determined by integrating the L96 model using the 4th-order Runge-Kutta method started from a random initial condition and iterated 28 000 timesteps with $dt = 0.001$.

TABLE I. System statistics for the Lorenz 96 experiments.

|	
System dimensions (m)	6
Experiment duration	600 analysis cycles each 0.05 time units in length ($\sim$6 h equivalent per cycle)
Starting point on the attractor	$[-4.362, 0.171, 5.566, 7.483, 2.772, -0.91]$
Average $x^{(N)}$ (nature)	$[1.87\ 2.26\ 2.93\ 2.61\ 2.64\ 2.70]$
Maximum $x^{(N)}$	$[11.88\ 9.58\ 11.31\ 10.58\ 10.73\ 11.80]$
Minimum $x^{(N)}$	$[-7.42\ -6.40\ -5.76\ -6.06\ -7.24\ -7.42]$
Number of nonnegative Lyapunov exponents	2
Predicted required ensemble size for EnKF without localization	3 (so that $k - 1 = 2$)
Predicted required ensemble size for Hybrid-ETKF	2 (smallest possible)
Observation error	0.01
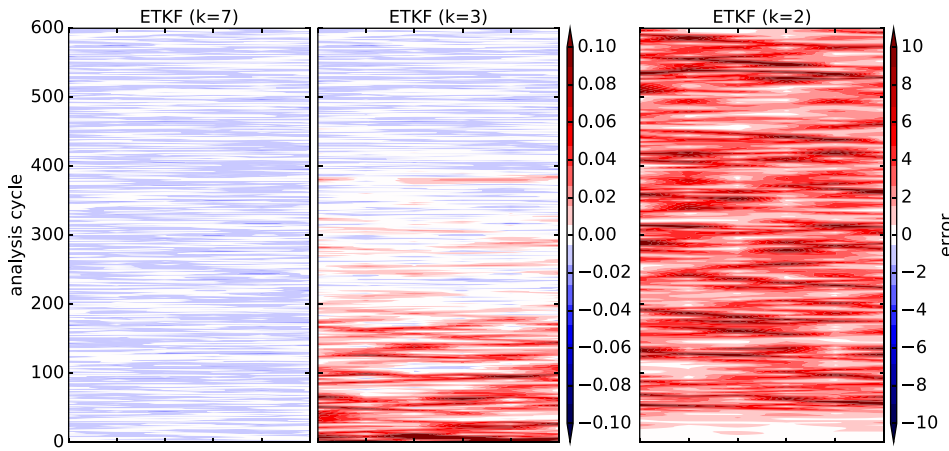
FIG. 2. Absolute analysis error for the (a) 7-member, (b) 3-member, and (c) 2-member ETKF minus absolute analysis error for 3DVar. The standard deviation of Gaussian observational error is 0.01. Negative values (dashed contours, blue shading) indicate that the ETKF produces lower errors. Note the change in the scale in (c).

Case 1: Full-coverage observing network

The first case we examine uses 6-h updates with observations covering the full domain, similar to the scenario investigated by Ng *et al.* (2011). The ETKF has slightly better accuracy than the 3DVar. The 3DVar and traditional impulsive synchronization (essentially equivalent to 3DVar with the off-diagonal elements set to zero in the forecast error covariance matrix **B**) are nearly identical, and thus in this case, the synchronization results are not shown. We show results from the ETKF and Hybrid compared to the 3DVar as a baseline in Fig. 2. The primary advantage provided by the ETKF is that it adaptively estimates the magnitude of the forecast error covariance by sampling the growth of instabilities along the system trajectory. However, as the ensemble size decreases, the ETKF can no longer represent all the long-term unstable and neutral modes. As predicted above, stability is still achieved with three members, but the filter quickly diverges when the ensemble size is reduced to two members. The hybrid filter, while partially reducing the accuracy of the ETKF for the "large" ensemble case ($k = 7$), shows similar accuracy when reducing the ensemble size to $k = 3$ and $k = 2$, preventing catastrophic filter divergence in the latter case (Fig. 3).

Case 2: Sparse observing network

We next consider a sparse observing network, which is more typical in geophysical applications. Using a sparse

observing network, the off-diagonal terms in the estimated forecast error covariance matrix become much more important. As case (2a), we consider a fixed observing network (e.g., emulating moored buoys in the ocean), and then as case (2b), we consider a random network (e.g., emulating drifter observations in the ocean).

In case (2a), observations are made at alternating nodes 0, 2, and 4. We see that the 3DVar case suffers from inadequately specified off-diagonal error covariance estimates that degrade the analysis compared to the traditional impulsive synchronization approach (Fig. 4). A large ensemble ($k = 7$) case using the ETFK however is capable of adequately representing the dynamical relationship between the observed nodes so that the unobserved nodes can be synchronized. As the ensemble size is reduced to two members (one fewer than that required by the analysis of the Lyapunov exponents), the ETKF quickly diverges.

Somewhat surprisingly, with the application of the static correction derived from the 3DVar to the 2-member ETKF, both of which are independently divergent, the hybrid filter is stabilized. Computing the Lyapunov exponents using Eq. (35), we see (Fig. 5) that the leading Lyapunov exponent of the ETKF system with $k = 2$ is positive, leading us to expect a divergent filter. The 3DVar system has a negative leading Lyapunov exponent, but its magnitude is small, approximately -8.88×10^{-2}. The small magnitude of the leading Lyapunov exponent renders the 3DVar susceptible to noise in the observations of a similar magnitude, and thus, the
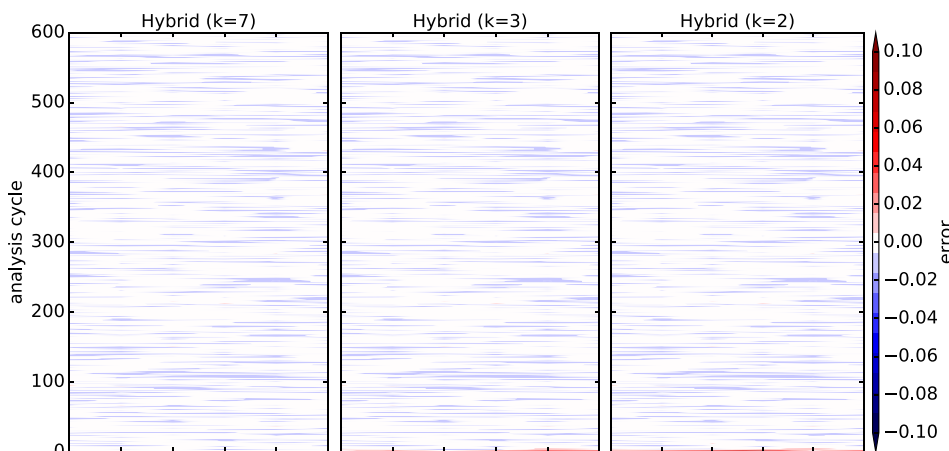

FIG. 3. Absolute analysis error for the (a) 7-member, (b) 3-member, and (c) 2-member Hybrid minus absolute analysis error for 3DVar. The standard deviation of Gaussian observational error is 0.01. Negative values (dashed contours, blue shading) indicate that the Hybrid produces slightly lower errors. It is evident that although for the larger ensemble size, the Hybrid lessens the advantage of the ETKF over 3DVar, for smaller ensemble sizes, the Hybrid stabilizes the ETKF and provides a faster spin-up to a synchronized status.
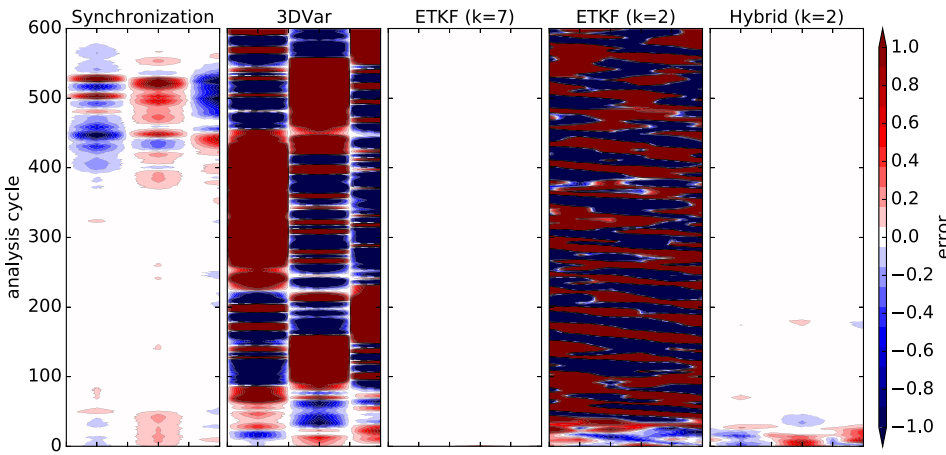
FIG. 4. Absolute analysis error using a sparse observing network alternating observations at model grid points 0, 2, and 4 for (a) direct coupling synchronization method with eventual instabilities at the unobserved nodes, (b) 3DVar with large errors at the unobserved nodes, (c) 7-member ETKF with negligible error, (d) 2-member ETKF that diverges, and (e) 2-member hybrid that is stable despite the fact that neither of its constituent 2-member ETKF and 3DVar components are stable.
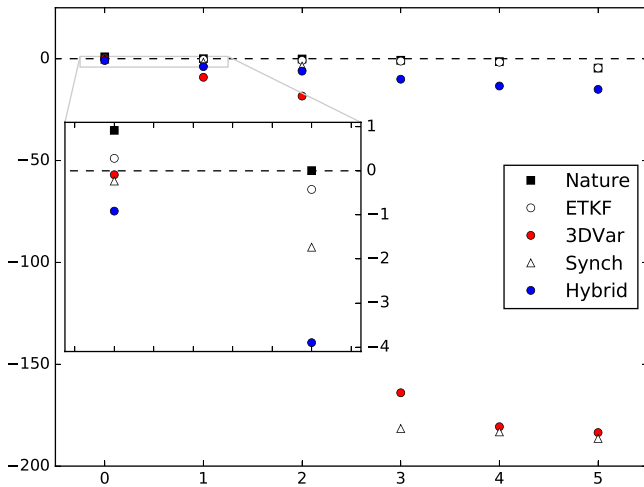

FIG. 5. Lyapunov exponents for the methods in case (2a) (corresponding to Fig. 4). The ETKF and Hybrid-Gain are shown only for an ensemble size of $k = 2$. The 3DVar data assimilation system has a negative leading Lyapunov exponent close to 0 but diverges due to noise in the observations. The hybrid method (with an ensemble size of $k = 2$) is stable.

system is unstable. The hybrid method combining both ETKF and 3DVar reduces the leading Lyapunov exponent and achieves stability.

In case (2b), we observe 3 random nodes per cycle. As we allow the observations to migrate in time, we see improved filter performance for all but the small ensemble case ($k = 2$) for ETKF (Fig. 6). This is the result of the dynamical error

growth in L96 being relatively linear over multiple passes of the 0.05 timestep analysis cycle, thus providing a relatively high probability that the entire domain will be sampled within the linear regime of error growth.

OCEAN GENERAL CIRCULATION MODEL

In realistic applications, numerical forecast models contain errors that result in systematic biases. We now show results using a global ocean model and real observations as used by NCEP for ocean monitoring and seasonal climate forecasts. We compare the LETKF and the Hybrid-Gain data assimilation methods applied to the GFDL MOM4p1 ocean model (Penny *et al.*, 2015). We use *in situ* observations consisting of temperature and salinity vertical profiles, focusing particularly on the equatorial Atlantic. In the first case (a), LETKF is used in its standard form, while in the second case (b), the Hybrid-Gain method applies LETKF as the initial state estimate with a mean correction applied to the ensemble at each analysis cycle based on the NCEP operational 3DVar. This implementation of 3DVar uses a background error covariance matrix that is static in the horizontal direction and slightly varying in the vertical direction depending on the seasonal variation of the thermocline depth. Saha *et al.* (2010) give further details of the NCEP operational 3DVar.

Brandt *et al.* (2011) reported observations in the equatorial Atlantic that indicated stacked zonal jets alternating with the depth. However, these stacked currents are not produced in the model. The absence of these features in the reanalysis
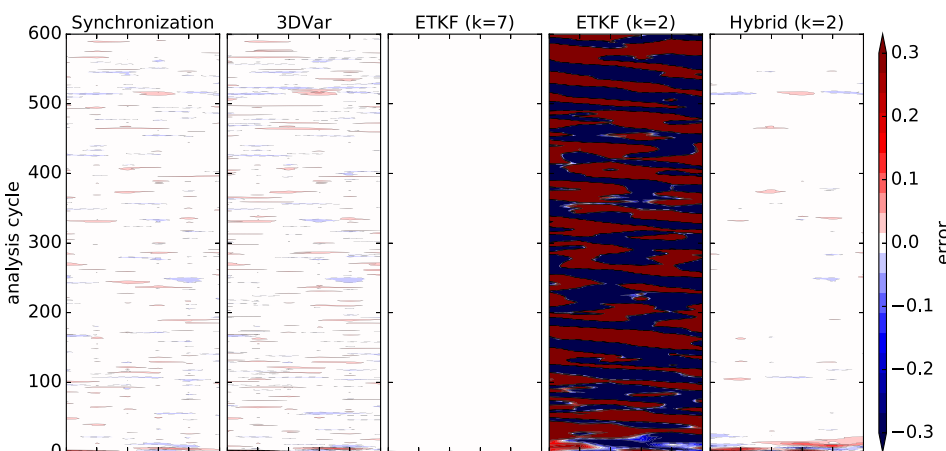

FIG. 6. Absolute analysis error using a sparse observing network consisting of 3 randomly observed nodes per cycle for (a) direct coupling synchronization method, (b) 3DVar, (c) 7-member ETKF, (d) 2-member ETKF that diverges, and (e) 2-member hybrid that regains stability.
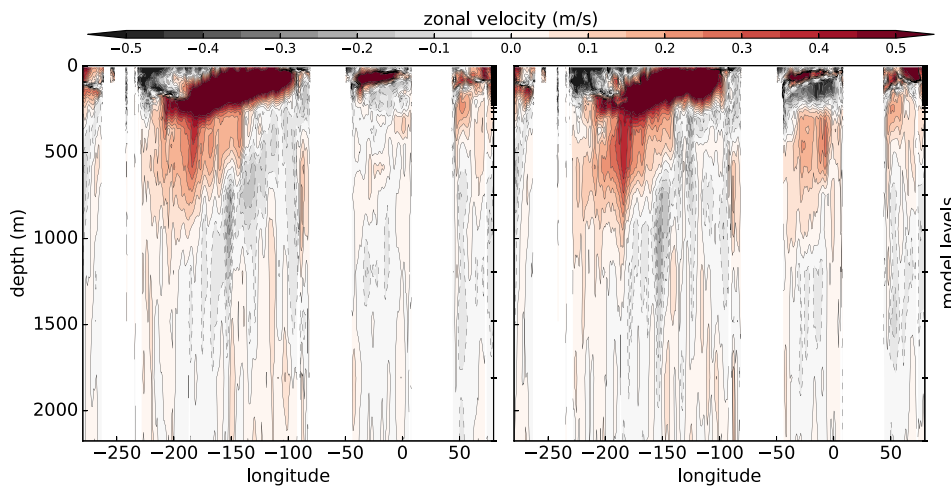
FIG. 7. Zonal currents from depths of 0 to 2000 m at the equator for (a) LETKF and (b) Hybrid-Gain LETKF. The latter shows a recovery of the observed stacked alternating zonal jets in the equatorial Atlantic (between longitudes 50W and 10E) due to the bias-correction effects of the hybrid approach. Model levels are shown with ticks on the right hand side, indicating the higher vertical resolution near the surface.

may be due to either errors in the surface wind stress provided by NCEP Reanalysis 2 (Kanamitsu *et al.*, 2002) or due to errors in the ocean model response.

Compared to the observed features in zonal velocity at the equator, we see in Fig. 7 that the hybrid approach (b) is able to recover the stacked deep ocean jets that are not adequately reproduced by LETKF alone (a). This fault occurs for LETKF because each member of the ensemble exhibits the same model bias, and thus, the ensemble does not adequately represent the uncertainty in the model forecast. Thus, important observations are largely ignored in the region of the known model deficiency. By supplementing the ensemble with a static error covariance derived from long time-averaged error information (which may include information about known systematic model errors), the hybrid is able to correct the state estimate accordingly. At the deeper levels, the jets are still weak due to the reduced vertical resolution in the model.

This is a realistic ocean application in which applying the hybrid system reduces biases in the analysis versus an EnKF baseline. Penny *et al.* (2015) demonstrated a similar bias-correcting effect in a perfect-model experiment that applied an imposed bias to the ensemble surface fluxes, using the same Hybrid-Gain ocean data assimilation system that is used for this analysis. Next, we extend the reanalysis of Penny *et al.* (2015) using real historical observations to include an additional decade of results, extending into the modern Argo era. As the costs associated with a full-scale stability analysis would be prohibitive on an operational system, we use the root mean square (RMS) deviations and global biases in the observation-minus-forecast statistics as a proxy to evaluate the performance of the Hybrid-Gain method in comparison to 3Dvar. We see in Fig. 8 that compared to the NCEP operational 3DVar, the hybrid produces consistently reduced RMS deviations compared to observations and also a reduction in the global upper (0–700 m) ocean temperature and salinity biases. As the number of salinity observations increases dramatically from the first half to the second half of the experiment period with the
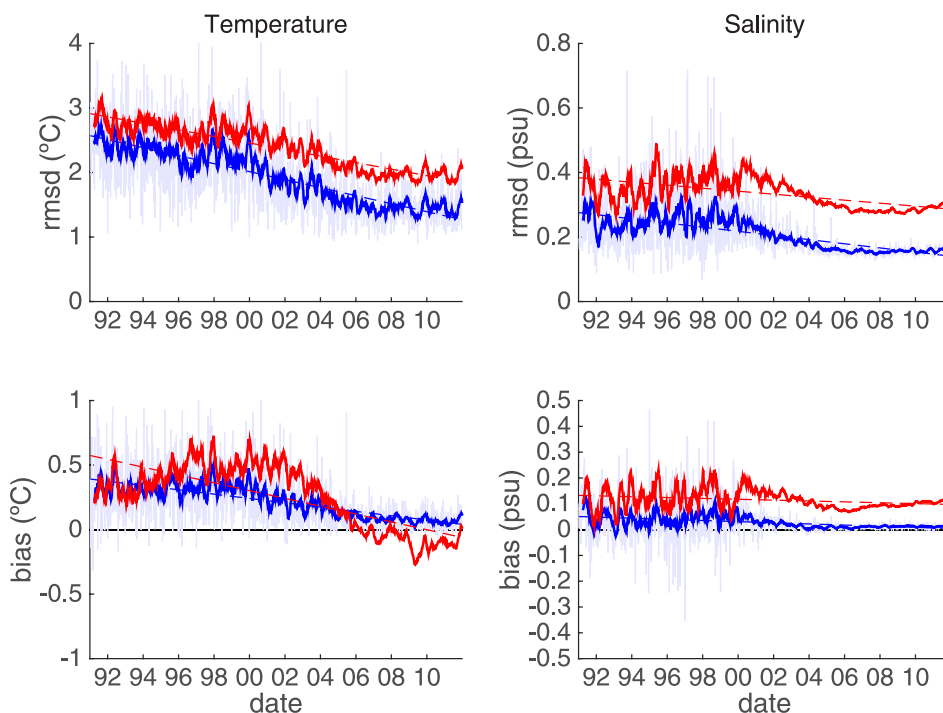

FIG. 8. Three-month moving average of global upper ocean (0–700 m) (a) temperature RMSD (°C), (b) salinity RMSD (psu), (c) temperature bias, and (d) salinity bias. The NCEP operational 3DVar is shown in red, while the future replacement hybrid system is shown in blue. Dashed lines indicate the respective linear trend. The hybrid reduces RMS deviations compared to observations and reduces systematic biases.

introduction of the Argo float program, 3DVar suffers from assuming a static climatological relationship between temperature and salinity. The hybrid method is able to adapt to the changing observing network while producing a corresponding reduction in global bias.

CONCLUSION

We have described common operational data assimilation methods in the context of synchronization as generalized one-way coupled impulsive synchronization. The stability properties of the resulting impulsive differential equation can be investigated through the analysis of the associated Kalman gain matrix and observation operator.

Hybrid data assimilation methods are able to recover lost stability in ensemble methods that have an insufficient number of members to represent the unstable modes that correspond to the positive and neutral Lyapunov exponents. The hybrid methods can also make up for potential inadequacies in static error covariance estimates used in variational methods.

The benefit of utilizing forecast error covariance information becomes clear when the observation coverage is sparse, while the benefit of the hybridization becomes clear when the ensemble sampling is sparse. When considering model error (i.e., generalized synchronization), the hybrid method is able to correct model biases that the pure ensemble Kalman filter cannot.

In the future, we will examine the stability properties of the Hybrid-Gain approach in greater detail. For the application to large-scale systems, candidate methods for this investigation are the breeding on the data assimilation system (BDAS) of Trevisan and Uboldi (2004) and Carrassi et al. (2007) and the use of the Local Ensemble Tangent Linear Model (LETLM) introduced by Bishop et al. (2017) and Frolov and Bishop (2016) to evaluate the linearized dynamics. More recent investigations of time-delay methods (Pazò et al., 2016 and An et al., 2017) may also provide useful supplements to the current state-of-the-art data assimilation methodologies, particularly for sparsely observed systems.

ACKNOWLEDGMENTS

Abarbanel, H. D. I., Brown, R., Sidorowich, J. J., and Tsimring, L. S., "The analysis of observed chaotic data in physical systems," Rev. Mod. Phys. **65**(4), 1331–1392 (1993).

Abarbanel, H. D. I., Kostuk, M., and Whartenby, W., "Data assimilation with regularized nonlinear instabilities," Q. J. R. Meteorol. Soc. **136**(A), 769–783 (2010).

An, Z., Rey, D., Ye, J. X., and Abarbanel, H. D. I., "Estimating the state of a geophysical system with sparse observations: Time delay methods to achieve accurate initial states for prediction," Nonlinear Processes Geophys. **24**, 9–22 (2017).

Asch, M., Bocquet, M., and Nodet, N., *Data Assimilation: Methods, Algorithms, and Applications* (SIAM, 2017), ISBN 978-1-611974-53-9.

Balci, N., Mazzucato, A. L., Restrepo, J. M., and Sell, G. R., "Ensemble dynamics and bred vectors," Mon. Weather Rev. **140**, 2308–2334 (2012).

Bannister, R. N., "A review of operational methods of variational and ensemble-variational data assimilation," Q. J. R Meteor. Soc. **143**(703), 607–633 (2017).

Bishop, C. H., Etherton, B. J., and Majumdar, S. J., "Adaptive sampling with the ensemble transform Kalman filter. Part I: Theoretical aspects," Mon. Weather Rev. **129**, 420–436 (2001).

Bishop, C. H., Frolov, S., Allen, D. R., Kuhl, D. D., and Hoppel, K., "The local ensemble tangent linear model: An enabler for coupled model 4D-Var," Q. J. R. Meteorol. Soc. **143**, 1009–1020 (2017).

Blanes, S., Casas, F., Oteo, J. A., and Ros, J., "The Magnus expansion and some of its applications," Phys. Rep. **470**, 151–238 (2009).

Bloom, S. C., Takacs, L. L., da Silva, A. M., and Ledvina, D., "Data assimilation using incremental analysis updates," Mon. Weather Rev. **124**, 1256–1271 (1996).

Bocquet, M. and Carrassi, A., "Four-dimensional ensemble variational data assimilation and the unstable subspace," Tellus A **69**, 1304504 (2017).

Bocquet, M., Gurumoorthy, K. S., Apte, A., Carrassi, A., Grudzien, C., and Jones, C. K. R. T., "Degenerate Kalman filter error covariances and their convergence onto the unstable subspace," preprint arXiv:1604.02578 (2017).

Bonavita, M., Hamrud, M., and Isaksen, L., "EnKF and hybrid gain ensemble data assimilation part II: EnKF and hybrid gain results," Mon. Weather Rev. **143**, 4865–4882 (2015).

Brandt, P., Funk, A., Hormann, V., Dengler, M., Greatbatch, R. J., and Toole, J. M., "Interannual atmospheric variability forced by the deep equatorial Atlantic ocean," Nature **473**, 497–501 (2011).

Carrassi, A., Ghil, M., Trevisan, A., and Uboldi, F., "Data assimilation as a nonlinear dynamical system problem: Stability and convergence of the prediction-assimilation system," Chaos **18**, 023112 (2008).

Carrassi, A., Trevisan, A., and Uboldi, F., "Adaptive observations and assimilation in the unstable subspace by breeding on the data assimilation system," Tellus A **59**, 101–113 (2007).

Carrassi, A., Vannitsem, S., Zupanski, D., and Zupanski, M., "The maximum likelihood ensemble filter performances in chaotic systems," Tellus A **61**, 587–600 (2009).

Chen, S., Xi F., and Liu, Z., "Supreme local Lyapunov exponents and chaotic impulsive synchronization," Int. J. Bifurcation Chaos **23**(10), 1350169 (2013).

Duane, G., Tribbia, J. J., and Weiss, J. B., "Synchronicity in predictive modeling: A new view of data assimilation," Nonlinear Processes Geophys. **13**, 601–612 (2006).

Evensen, G., "Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics," J. Geophys. Res. **99**(10), 10143–10162, doi:10.1029/94JC00572 (1994).

Frolov, S. and Bishop, C. H., "Localized ensemble-based tangent linear models and their use in propagating hybrid error covariance models," Mon. Weather Rev. **144**, 1383–1405 (2016).

Ghil, M. and Malanotte-Rizzoli, P., "Data assimilation in meteorology and oceanography," Adv. Geophys. **33**, 141–266 (1991).

Glahn, H. R. and Lowry, D. A., "The use of model output statistics (MOS) in objective weather forecasting," J. Appl. Meteorol. **11**, 1203–1211 (1972).

Hamill, T. M. and Snyder, C., "A hybrid ensemble Kalman filter-3D variational analysis scheme," Mon. Weather Rev. **128**, 2905–2919 (2000).

Hamrud, M., Bonavita, M., and Isaksen, L., "EnKF and hybrid gain ensemble data assimilation," ECMWF Tech. Rep. **733**, 1–34 (2014), available at https://www.ecmwf.int/sites/default/files/elibrary/2014/9766-enkf-and-hybrid-gain-ensemble-data-assimilation.pdf.

Han, Y., van Delst, P., Liu, Q., Weng, F., Yan, B., Treadon, R., and Derber, J., "Community radiative transfer model (CRTM): Version 1," NOAA Technical Report No. 1-122, 2006.

Hoke, J. and Anthes, R., "The initialization of numerical models by a dynamic relaxation technique," Mon. Weather Rev. **104**, 1551–1556 (1976).

Hunt, B. R., Kostelich, E. J., and Szunyogh, I., "Efficient data assimilation for spatiotemporal chaos: a local ensemble transform Kalman filter," Physica D **230**, 112–126 (2007).

Hunt, B. R., Ott, E., and Yorke, J. A., "Differentiable generalized synchronization of chaos," Phys. Rev. E **55**(4), 4029–4034 (1997).

Kalman, R. E., "A new approach to linear filtering and prediction problems," J. Basic Eng. **82**, 35–45 (1960).

Kalnay, E., *Atmospheric Modeling, Data Assimilation and Predictability* (Cambridge University Press, 2003), p. 341.

Kalnay, E., Li, H., Miyoshi, T., Yang, S.-C., and Ballabrera-Poy, J., "4-D-Var or ensemble Kalman filter?," Tellus A **59**, 758–773 (2007).

Kanamitsu, M., Ebisuzaki, W., Woollen, J., Yang, S.-K., Hnilo, J. J., Fiorino, M., and Potter, G. L., "NCEP–DOE AMIP-II reanalysis (R-2)," Bull. Am. Meteorol. Soc. **83**, 1631–1643 (2002).

Kistler, R. E., "A study of data assimilation techniques in an autobarotropic primitive equation channel model," MS thesis (Department of Meteorology, Pennsylvania State University, 1974).

Kleist, D. T., "An evaluation of hybrid variational-ensemble data assimilation for the NCEP GFS," Ph.D. dissertation (University of Maryland, 2012), p. 149.

Kostuk, M., "Synchronization and statistical methods for the data assimilation of HVc neuron models," Ph.D. dissertation (University of California, San Diego, 2012), p. 135.

Kupstov, P. V. and Parlitz, U., "Theory and computation of covariant lyapunov vectors," J. of Nonlinear Science **22**(5), 727–762 (2012), available at https://link.springer.com/article/10.1007%2Fs00332-012-9126-5.

Le Dimet, F. X. and Talagrand, O., "Variational algorithms for analysis and assimilation of meteorological observations: Theoretical aspects," Tellus **38A**, 97–110 (1986).

Lorenc, A. C., "Analysis methods for numerical weather prediction," Q. J. R. Meteorol. Soc. **112**, 1177–1194 (1986).

Lorenc, A. C., "The potential of the ensemble Kalman filter for NWP – a comparison with 4D-Var," Q. J. R. Meteorol. Soc. **129**, 3183–3203 (2003).

Lorenc, A. C. and Rawlins, F., "Why does 4D-Var beat 3D-Var?," Q. J. R. Meteorol. Soc. **131**, 3247–3257 (2005).

Lorenz, E. N., "Deterministic nonperiodic flow," J. Atmos. Sci. **20**, 130–141 (1963).

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

Lorenz, E. N., "Predictability—A problem partly solved," in *Proceedings of a Seminar Held at ECMWF on Predictability, ECMWF Seminar Proceedings* (1996), Vol. **1**, pp. 1–18.

Magnus, W., "On the exponential solution of differential equations for a linear operator," Commun. Pure Appl. Math. **7**, 649–673 (1954).

Ng, G.-H. C., Mclaughlin, D., Entekhabi, D., and Ahanin, A., "The role of model dynamics in ensemble Kalman filter performance for chaotic systems," Tellus **63A**, 958–977 (2011).

Ni, B. and Zhang, Q., "Stability of the Kalman filter for continuous time output error systems," Syst. Control Lett. **94**, 172–180 (2016).

Palatella, L. and Trevisan, A., "Interaction of Lyapunov vectors in the formulation of the nonlinear extension of the Kalman filter," Phys. Rev. E **91**, 042905 (2015).

Pazò, D., Carrassi, A., and Lòpez, J. M., "Data assimilation by delay-coordinate nudging," Q. J. R. Meteorol. Soc. **142**, 1290–1299 (2016).

Pecora, L. and Carroll, T., "Synchronization in chaotic systems," Phys. Rev. Lett. **64**, 821–824 (1990).

Pecora, L. M., Carroll, T. L., Johnson, G. A., Mar, D. J., and Heagy, J. F., "Fundamentals of synchronization in chaotic systems, concepts, and applications," Chaos **7**(4), 520–543 (1997).

Penny, S. G., "The hybrid local ensemble transform Kalman filter," Mon. Weather Rev. **142**, 2139–2149 (2014).

Penny, S. G., Behringer, D., Carton, J. A., and Kalnay, E., "A hybrid global ocean data assimilation system at NCEP," Mon. Weather Rev. **143**, 4660–4677 (2015).

Penny, S. G., Kalnay, E., Carton, J. A., Hunt, B. R., Ide, K., Miyoshi, T., and Chepurin, G. A., "The local ensemble transform Kalman filter and the running-in-place algorithm applied to a global ocean general circulation model," Nonlinear Processes Geophys. **20**, 1031–1046 (2013).

Roemmich, D. and Owens, W. B., "The Argo project: Global ocean observations for the understanding and prediction of climate variability," Oceanography **13**, 45–50 (2000).

Rulkov, N., Sushchik, M., Tsimring, L., and Abarbanel, H., "Generalized synchronization of chaos in directionally coupled chaotic systems," Phys. Rev. E **51**, 980–994 (1995).

Saha, S., *et al.*, "The NCEP climate forecast system," J. Clim. **19**, 3483–3517 (2006).

Saha, S., *et al.*, "The NCEP climate forecast system reanalysis," Bull. Am. Meteorol. Soc. **91**, 1015–1057 (2010).

Saha, S., *et al.*, "The NCEP climate forecast system version 2," J. Clim. **27**, 2185–2208 (2014).

Sakov, P. and Oke, P. R., "Implications of the form of the ensemble transformation in the ensemble square root filters," Mon. Weather Rev. **136**, 1042–1053 (2008).

Samoilenko, A. M. and Perestyuk, N. A., *Impulsive Differential Equations*, World Scientific Series on Nonlinear Science, Series A, edited by L. O. Chua (World Scientific, 1995), Vol. 14.

Sasaki, Y., "Some basic formalisms in numerical variational analysis," Mon. Weather Rev. **98**, 875–883 (1970).

Talagrand, O. and Courtier, P., "Variational assimilation of meteorological observations with the adjoint vorticity equation (I): Theory," Q. J. R. Meteorol. Soc. **113**, 1311–1328 (1987).

Thepaut, J.-N. and Courtier, P., "Four-dimensional variational data assimilation using the adjoint of a multilevel primitive-equation model," Q. J. R. Meteorol. Soc. **117**, 1225–1254 (1991).

Trevisan, A. and Palatella, L., "On the Kalman filter error covariance collapse into the unstable subspace," Nonlinear Processes Geophys. **18**, 243–250 (2011).

Trevisan, A. and Uboldi, F., "Assimilation of standard and targeted observations within the unstable subspace of the observation–analysis–forecast cycle system," J. Atmos. Sci. **61**, 103–113 (2004).

Uboldi, F., Trevisan, A., and Carrassi, A., "Developing a dynamically based assimilation method for targeted and standard observations," Nonlinear Processes Geophys. **12**, 149–156 (2005).

Wang, X., "Incorporating ensemble covariance in the gridpoint statistical interpolation variational minimization: A mathematical framework," Mon. Weather Rev. **138**, 2990–2995 (2010).

Wang, X., Barker, D. M., Snyder, C., and Hamill, T. M., "A hybrid ETKF–3DVAR data assimilation scheme for the WRF Model. Part I: Observing system simulation experiment," Mon. Weather Rev. **136**, 5116–5131 (2008a).

Wang, X., Barker, D. M., Snyder, C., and Hamill, T. M., "A hybrid ETKF–3DVAR data assimilation scheme for the WRF Model. Part II: Real observation experiments," Mon. Weather Rev. **136**, 5132–5147 (2008b).

Wang, X., Hamill, T. M., Whitaker, J. S., and Bishop, C. H., "A comparison of hybrid ensemble transform Kalman filter–optimum interpolation and ensemble square root filter analysis schemes," Mon. Weather Rev. **135**, 1055–1076 (2007a).

Wang, X., Parrish, D., Kleist, D., and Whitaker, J., "GSI 3DVar-based ensemble–variational hybrid data assimilation for NCEP Global Forecast System: Single-resolution experiments," Mon. Weather Rev. **141**, 4098–4117 (2013).

Wang, X., Snyder, C., and Hamill, T. M., "On the theoretical equivalence of differently proposed ensemble–3D-VAR hybrid analysis schemes," Mon. Weather Rev. **135**, 222–227 (2007b).

Whartenby, W. G., Quinn, J. C., and Abarbanel, J. C., "The number of required observations in data assimilation for a shallow-water flow," Mon. Wea. Rev. **141**, 2502–2518 (2013).

Whitaker, J. S. and Hamill, T. M., "Ensemble data assimilation without perturbed observations," Mon. Wea. Rev. **130**, 1913–1924 (2002).

Yang, S.-C., Baker, D., Li, H., Cordes, K., Huff, M., Nagpal, G., Okereke, E., Villafane, J., Kalnay, E., and Duane, G., "Data assimilation as synchronization of truth and model: experiments with the three-variable Lorenz system," J. Atmos. Sci. **63**, 2340–2354 (2006).

Yang, S.-C., Corazza, M., Carrassi, A., Kalnay, E., and Miyoshi, T., "Comparison of Local Ensemble Transform Kalman Filter, 3DVAR, and 4DVAR in a Quasigeostrophic Model," Mon. Weather Rev. **137**, 693–709 (2009).

Yang, S.-C., Kalnay, E., and Hunt, B., "Handling nonlinearity in an ensemble Kalman filter: experiments with the three-variable Lorenz model," Mon. Weather Rev. **140**, 2628–2646 (2012).