

Forecasting regime changes in a chaotic toy climate

Kameron Decker Harris,^{1,2} El Hassan Ridouane,³

Darren L. Hitt,³ and Christopher M. Danforth⁴

¹*Department of Mathematics and Statistics*

²*Vermont Advanced Computing Center*

³*Mechanical Engineering Program*

⁴*Complex Systems Center*

University of Vermont, Burlington, VT 05405 USA

(Dated: June 1, 2009)

Abstract

A low-dimensional model of natural convection, similar to the classic Lorenz system, is derived and compared to flow simulations in order to (1) compare data assimilation methods, (2) better understand the physics of the thermosyphon, and (3) predict changes between different flow regimes. A thermosyphon is a physical apparatus which provides a simple platform for the study of natural convection. The flow dynamics of the thermosyphon and those of the atmosphere have similar forecasting difficulties, and both exhibit distinct behavioral regimes, so we consider the thermosyphon a useful toy model for climate. The thermosyphon is represented by a long time direct numerical simulation of the flow, which serves as a reference “truth”. Forecasts are then made using low-dimensional Lorenz-like models and synchronized to noisy and limited observations of the truth using 3D-Var, the extended Kalman filter (EKF), the ensemble square root filter (EnSRF), and the ensemble transform Kalman filter (ETKF). Regime changes in the toy climate are forecast up to 5 minutes in advance and up to 97% of the time. Residency times are predicted with a probabilistic method, yielding $\approx 50\%$ improvement over climatology. Regime changes have been successfully forecast in the Lorenz system under the perfect model scenario but never in the presence of model error or unobserved variables. Finally, we provide new details concerning the physical processes present in the thermosyphon during these regime changes.

Keywords: Bred vectors, chaos, data assimilation, ensemble, forecast, Kalman filter, Lorenz system, thermosyphon

I. INTRODUCTION

Many of the problems associated with weather and climate forecasting do not result from our lack of knowledge of the geophysical governing equations. Even though such equations are well known, nonlinearities tend to amplify uncertainty in our estimate of the current atmospheric state, known as the initial condition (IC). In fact, even for a perfect IC, uncertainty in the parameters used to represent sub-grid scale phenomena (e.g., clouds) and other sources of model error lead to forecast divergence. Weather models define temperature, pressure, and other relevant quantities at regularly-spaced gridpoints throughout the atmosphere, but meteorologists do not have observations at every one of these locations, so knowledge of the IC is inherently incomplete. To estimate the initial state of the atmosphere, statistically informed guesses are made for unmeasured variables. Over time, these guesses are validated by new measurements in a forecast-observe-analyze cycle called data assimilation (DA). Improving upon filters from control theory, modern DA techniques are used to estimate the best ICs for the short-term forecasts made by operational numerical weather prediction (NWP) models. However, DA is applicable to any modeling endeavor where only partial knowledge of the “truth” is available.

The *ensemble* forecasting methodology approximates uncertainty in the initial state with a finite set of perturbed ICs. The ensemble of ICs is integrated forward in time using the best available model; the resulting ensemble of forecasts estimates the probability distribution of potential outcomes. The ensemble spread quantifies forecast uncertainty, and the ensemble mean can give a better guess for the true state than any one member. However, the ensemble members must be chosen in a clever manner, because if the deviations of multiple ensemble members from the mean trajectory are linearly dependent, then the information they carry is redundant [3]. Furthermore, the number of degrees of freedom of modern NWP models is $\mathcal{O}(10^9)$, so the numerical cost of integrating these models limits ensembles to $< \mathcal{O}(10^2)$ members in practical applications.

In this paper, we use a toy model for atmospheric convection to compare various methods of DA for prediction of nonlinear phenomena. The toy climate investigated is a thermosyphon, also known as a natural convection loop or non-mechanical heat pump. Thermosyphons are used in solar water heaters [4], cooling systems for computers [5], roads and railways that cross permafrost [6], nuclear power plants [7, 8], and other industrial

applications. In the system, buoyant forces move fluid through a closed-loop with circular geometry (see Fig. 1). As the thermal forcing increases, the state of the thermosyphon undergoes bifurcations from an initial purely conducting state to a steady convecting state. Under further forcing, steady convection becomes unstable and the flow reverses directions chaotically. As first suggested by Lorenz [9], this system is illustrative of the unpredictable behavior governing weather and climate dynamics.

Following previous experiments [10–15], we consider a circular thermosyphon geometry, analogous to a vertically-oriented hula hoop. An imposed wall temperature T_h on the lower half of the loop ($-\frac{\pi}{2} < \phi < \frac{\pi}{2}$) heats the fluid contained in this section. Similarly, a wall temperature $T_c < T_h$ is imposed on the upper half ($\frac{\pi}{2} < \phi < \frac{3\pi}{2}$) to cool the upper section (Fig. 1). The forcing is constant, i.e., we examine the case of developed flow. This creates a temperature inversion. As fluid in the bottom of the loop heats past the point of simple conduction, buoyant forces overcome those of friction and viscosity, and the warm fluid rises (cool fluid sinks). The cyclical flow of hot and cold fluid forms a circular structure called a convection cell. Unlike Rayleigh-Bénard convection, fluid in the thermosyphon is restricted to a single cell by the geometry and must initially rotate clockwise or counter-clockwise. Each direction of rotation is said to constitute a *regime* of behavior; a flow reversal is referred to as a *regime change*. The initial regime is determined randomly, since the symmetry of the tube’s geometry and uniform heating do not favor any particular direction. The fluid will accelerate until the buoyant force is balanced by friction and gravity, and the flow stabilizes. With sufficient heating, i.e., large enough $\Delta T_w = T_h - T_c$, the flow aperiodically reverses direction in a chaotic manner while still laminar. This means that slightly perturbed states diverge exponentially in time, or that there is at least one positive Lyapunov exponent. Among the goals of this research is to develop an accurate method for predicting the duration of the current and next regime, given a time series of measurements of the mass flow rate averaged over the loop volume.

This paper is structured in the following way: In Section II, we describe the low-dimensional system of ODEs used for forecasting and the direct numerical simulation used to generate a synthetic true state or “nature run” of the toy climate. In Section III, we describe the various nonlinear DA and ensemble methods used in this experiment. In Section IV, we present methods and results, and we conclude with implications for NWP in Section V.

FIG. 1. The thermosyphon has a simple circular geometry. The bottom wall is heated to a constant hot temperature T_h while the top wall is maintained at the temperature T_c , creating an unstable temperature inversion of hot fluid below cold fluid. Energy in the bottom half must find a way to the top half. If conduction alone cannot stabilize this temperature inversion, then the fluid will begin to move and convection becomes the dominant process of heat transfer. The most important model state variables are proportional to the fluid velocity u and the temperature difference across the loop ΔT_{3-9} . The radius ratio $R/r = 24$ is shown.

II. MODEL DESCRIPTION

Several studies have examined the periodic [10] and chaotic [11–21] behavior of toroidal thermosyphons, many exhibiting behavior closely mimicking that of the Lorenz system [9]. Gorman et al. [14] developed a dynamical model for constant heat flux through the bottom half. A typical example of such a flux condition would be observed experimentally when the forcing is provided by a heating tape. Ehrhard and Müller [15] derived similar equations of motion with fixed wall temperature for the top and bottom halves of the loop. We replicate these Dirichlet boundary conditions. One recent study [16] used control theory to suppress chaotic behavior in both numerical and physical experiments. Jiang and Shoji [17] used a multiscale analysis of the field equations to derive low-dimensional models for arbitrary boundary conditions. Burroughs et al. [18] compared a partial differential equation model to three-dimensional (3D) direct numerical simulations; their model was able to determine the location of the various bifurcation points more accurately than the Lorenz equations, although they did not thoroughly investigate the chaotic regime due to high computational cost. To our knowledge, Desrayaud et al. [19] implemented the first two-dimensional (2D) direct numerical simulation of a thermosyphon. They were able to capture spatiotemporal details of the flow and found stable, periodic, and chaotic behavior. Ridouane et al. [21] found similar results with a 2D numerical simulation and characterized the time-dependent structure of the flow immediately prior to a regime change. In the present study, we utilize this same flow simulation capability to establish the true state of the toy climate; computational details are presented in Section II B.

A. Ehrhard-Müller (EM) System

The dynamics of the thermosyphon are dominated by convection, resulting in Lorenz-like flow. By deriving the dynamical equations governing the thermosyphon from basic principles based on its geometry and the imposed boundary conditions, one is left with parameters in terms of physically meaningful constants and the scalings needed to make the state variables dimensionless. In principle, the physical properties of the fluid and loop define all parameters of the model. In practice, this is more complicated, because some parameters must be measured empirically (see Appendix A).

1. Derivation

Following the derivations of Gorman et al. [14] and Ehrhard and Müller [15], we consider the forces acting upon a control volume of incompressible fluid in the loop. All fluid properties are cross-sectionally averaged, and the radial components of velocity and heat conduction within the fluid are neglected. The fluid velocity $u = u(t)$ is assumed to be constant at all points. Applying Newton's second law, the sum of all forces on the control volume must equal its change in momentum [14, 15]:

$$F_p + F_f + F_g = \rho\pi r^2 R d\phi \frac{du}{dt} \quad (1a)$$

where

$$F_p = -\pi r^2 R d\phi \nabla p = -\pi r^2 d\phi \frac{\partial p}{\partial \phi} \quad (1b)$$

$$F_f = -\rho\pi r^2 R d\phi f_w \quad (1c)$$

$$F_g = -\rho\pi r^2 R d\phi g \sin \phi. \quad (1d)$$

The angular coordinate ϕ and loop dimensions r and R are defined in Fig. 1; g is the acceleration of gravity, ρ is the fluid density, u is velocity, and p is pressure. The total force in Eqn. (1a) is comprised of the net pressure (F_p), friction from shear within the fluid (F_f), and the force of gravity (F_g). The pressure term, Eqn. (1b), is the volume times the pressure gradient. The friction term, Eqn. (1c), is written in this form in order to simplify the analysis; all frictional effects are contained in f_w which will depend on fluid velocity, to be discussed later.

Before we write the momentum equation, it is convenient to apply the Boussinesq approximation, which assumes that variations in fluid density are linear with temperature. In other words, $\rho = \rho(T) \approx \rho_0(1 - \gamma(T - T_0))$ where ρ_0 is the reference density, γ is the coefficient of volumetric thermal expansion, and $T_0 = \frac{1}{2}(T_h + T_c)$ is the reference temperature. The Boussinesq approximation also states that the density variation is insignificant except in terms multiplied by g . Thus, the density ρ is replaced by ρ_0 in all terms of Eqn. (1) except gravity, Eqn. (1d). Using the Boussinesq approximation, gathering terms, and dividing out common factors gives the momentum equation

$$\rho_0 \frac{du}{dt} d\phi = -d\phi \left(\frac{1}{R} \frac{\partial p}{\partial \phi} + \rho_0 (1 - \gamma(T - T_0)) g \sin \phi + f_w \right). \quad (2)$$

Integrating about the loop, the momentum equation is simplified because u and f_w are independent of ϕ and other terms drop out due to periodicity.

$$\rho_0 \frac{du}{dt} = \frac{\rho_0 \gamma g}{2\pi} \int_0^{2\pi} d\phi T \sin \phi - f_w \quad (3)$$

We now must account for the transfer of energy within the fluid, and between the fluid and the wall. All modes of heat transfer are neglected except convection, which is a valid approximation when $r \ll R$ (see [11, 15]). The energy rate of change (D/Dt is the material derivative with respect to time) in the control volume is

$$\rho_0 \pi r^2 R d\phi c_p \frac{DT}{Dt} \equiv \rho_0 \pi r^2 R d\phi c_p \left(\frac{\partial T}{\partial t} + \frac{u}{R} \frac{\partial T}{\partial \phi} \right) \quad (4)$$

which must be equal to the heat transfer through the wall

$$\Delta Q = -\pi r^2 R d\phi h_w (T - T_w), \quad (5)$$

where c_p is the specific heat of the fluid, h_w is the heat transfer coefficient, which depends on velocity, and T_w is the temperature at the wall. Combining Eqns. (4) and (5) gives the energy equation

$$\left(\frac{\partial T}{\partial t} + \frac{u}{R} \frac{\partial T}{\partial \phi} \right) = -\frac{h_w}{\rho_0 c_p} (T - T_w). \quad (6)$$

Together, Eqns. (3) and (6) represent a simple model of the flow in the loop.

The transport coefficients f_w and h_w characterize the interaction between the fluid and the wall. They are defined by the constitutive relations [15]

$$h_w = h_{w0} [1 + KH(|x_1|)] \quad (7)$$

$$f_w = \frac{1}{2} \rho_0 f_{w0} u, \quad (8)$$

where $x_1 \propto u$ is the dimensionless velocity. The function $H(x) = \Theta(1-x)p(x) + \Theta(x-1)x^{1/3}$ in Eqn. (7) determines the velocity dependence of the heat transfer coefficient, which varies as $u^{1/3}$ for moderate u [15]. We introduce the fitting polynomial $p(x) = 44/9 x^2 - 55/9 x^3 + 20/9 x^4$ to ensure that h_w is analytic at $x_1 = 0$. The Heaviside step function $\Theta(x)$ causes $H(x)$ to vary as $p(x)$ for $x \leq 1$ and $x^{1/3}$ for $x > 1$. Eqn. (8) gives the deceleration of the fluid when $|u| > 0$, and the $\rho_0/2$ term is retained to simplify the final solution. Dimensionally, f_w is an acceleration (m/s^2) and h_w is power per unit volume per unit temperature ($\text{W/m}^3\text{K}$). These coefficients h_{w0} , f_{w0} , and K must be estimated from experiments (e.g., [11, 14, 15]) or from other empirical means. In Appendix A, we describe the empirical methods used for parameter estimation.

Ehrhard and Müller [15] solved the system of two coupled, partial differential equations (Eqns. (3) and (6)) by introducing an infinite Fourier series for T . The essential dynamics can be captured by the lowest modes, i.e.,

$$T(\phi, t) = C_0(t) + S(t) \sin \phi + C(t) \cos \phi. \quad (9)$$

Because this form of T separates the variables ϕ and t , the problem is transformed into a set of ordinary differential equations. Substituting Eqn. (9) into Eqn. (3) and integrating gives the equation of motion for u . Similarly, Eqn. (6) is integrated by $\oint d\phi \sin \phi$ and $\oint d\phi \cos \phi$ to separate the two temperature modes S and C . The system is written in dimensionless form

$$\frac{dx_1}{dt'} = \alpha (x_2 - x_1) \quad (10a)$$

$$\frac{dx_2}{dt'} = \beta x_1 - x_2 (1 + KH(|x_1|)) - x_1 x_3 \quad (10b)$$

$$\frac{dx_3}{dt'} = x_1 x_2 - x_3 (1 + KH(|x_1|)) \quad (10c)$$

where the following linear transformations have been made to create dimensionless variables

$$\left. \begin{aligned} t' &= \frac{h_{w0}}{\rho_0 c_p} t \\ x_1 &= \frac{\rho_0 c_p}{R h_{w0}} u \\ x_2 &= \frac{1}{2} \frac{\rho_0 c_p \gamma g}{R h_{w0} f_{w0}} \Delta T_{3-9} \\ x_3 &= \frac{1}{2} \frac{\rho_0 c_p \gamma g}{R h_{w0} f_{w0}} \left(\frac{4}{\pi} \Delta T_w - \Delta T_{6-12} \right) \end{aligned} \right\} \quad (11)$$

Physically, x_1 is proportional to the mean fluid velocity, x_2 to the temperature difference across the convection cell or ΔT_{3-9} (between 3 o'clock and 9 o'clock), and x_3 is proportional to the deviation of the vertical temperature profile (characterized by the temperature difference between 6 o'clock and 12 o'clock, ΔT_{6-12}) from the value it takes during conduction.

The parameter $\alpha = \frac{1}{2}\rho_0 c_p f_{w_0}/h_{w_0}$ is comparable to the Prandtl number, the ratio of momentum diffusivity and thermal diffusivity. Similar to the Rayleigh number, the heating parameter

$$\beta = \frac{2}{\pi} \frac{\rho_0 c_p \gamma g}{R h_{w_0} f_{w_0}} \Delta T_w \quad (12)$$

determines the onset of convection as well as the transition to the chaotic regime. The nonlinear system of Eqn. (10) can be written as a vector function $\dot{\mathbf{x}} = F(\mathbf{x})$, where \mathbf{x} is the state vector $(x, y, z)^T$ and the dot is Newtonian notation for the time derivative; we write the parameters as a vector $\mathbf{p} = (\alpha, \beta, K)^T$. The numerical integration of this autonomous ODE was performed, as is customary for the Lorenz system, with a fourth-order Runge-Kutta method and timestep 0.01 in MATLAB R2008b [22].

Although the previous derivation assumes a 3D geometry, the direct numerical flow simulations described in Section II B were performed for the 2D case. A 2D geometry corresponds to infinite concentric cylinders as opposed to the quasi-1D torus. Due to cross-sectional averaging, the EM equations of motion (10) are the same in 2D or 3D; the change may be realized by letting $\pi r^2 \rightarrow 2r$ in Eqns. (1), (4), and (5). The only differences arise in the non-dimensional transformations and parameters, which were found empirically (see Appendix A).

2. Equilibrium Solutions and Flow Behavior

The EM system exhibits two kinds of equilibrium solutions: a conductive steady state, located at the origin in state space and corresponding to no motion; and two convective steady states, corresponding to constant velocity in the clockwise or counterclockwise directions. Performing a linear stability analysis, these equilibrium states are found from the positive real root ξ^* of the equation

$$(\beta - 1) + (\beta - 2)KH(\sqrt{\xi}) - K^2 H(\sqrt{\xi})^2 - \xi = 0 \quad (13)$$

which gives the convecting equilibria

$$[\pm\sqrt{\xi^*}, \pm\sqrt{\xi^*}, \xi^*/(1 + KH(\sqrt{\xi^*}))^T. \quad (14)$$

These equilibrium points were used in our regime change forecasts, for example when calculating the phase space distance from either equilibrium.

As mentioned previously, the thermosyphon undergoes bifurcations from conduction to convection [11–14]. For low β , stable conduction with a time-independent temperature field is the observed behavior. At a critical value of the heating parameter ($\beta = 1$, as for the Lorenz system), the two convective steady states bifurcate from the conducting equilibrium at the origin and become the attracting states for the system. These convecting equilibria move apart as β is increased, until they lose stability through a Hopf bifurcation and the system becomes globally chaotic. For heating parameters immediately below the transition to global chaos, the flow exhibits transient chaos which decays to stable oscillations or steady flow in a single regime. At slightly higher β , the system is subcritical, meaning there is the possibility of both chaotic and stable flows, depending on the IC [13, 14]. Finally, for even higher values of ΔT_w , the laminar assumption breaks down and convection is again stable [12–14].

3. Relationship to Lorenz System

Written in dimensionless form, the Lorenz [9] equations read

$$\frac{dx}{dt} = s(y - x) \quad (15a)$$

$$\frac{dy}{dt} = rx - y - xz \quad (15b)$$

$$\frac{dz}{dt} = xy - bz. \quad (15c)$$

Lorenz’s model describes the convection cells that develop in fluid placed between two plates. It was later discovered that the equations are also applicable to circular thermosyphons, although in Lorenz’s derivation the connection of the parameters to the thermosyphon design is not transparent. When $K = 0$, the EM system (Eqn. (10)) is analogous to the Lorenz system (Eqn. (15)) with $b = 1$. The lack of a geometric factor b in the EM system is due to the circular geometry, which forces a circular convection cell. The Lorenz state variables x ,

FIG. 2. Flow simulation results showing the temperature profile, units of K, of steady counter-clockwise convecting flow. The loop parameters are $Ra = 1.2 \times 10^4$ and $R/r = 3$ (for visualization). In the chaotic case, opposite anomalous regions of warmer and cooler fluid are superimposed on this temperature profile. As these pass through the loop, the “tongues” of warm and cool fluid extending into the top and bottom halves of the loop will grow and shrink simultaneously until the hot tongue visible near 2 o’clock reaches the opposite side of the loop. The flow then stalls and reverses direction.

FIG. 3. Visualization of the temperature profile (units are K) near the forcing discontinuity at the right end of the heating section is shown during a regime change. Initially, the flow is clockwise. As the fluid slows down, cool fluid mixes with warm fluid in the region due to a Kelvin-Helmholtz instability [21]. Heat is always passing from the wall to the fluid, the end result being an excess of hot fluid (e) where a “tongue” of cold (a) resided in the clockwise regime. This drives the flow in the counterclockwise direction. If the system pauses between regimes for an extended period, the fluid in the bottom becomes very hot. When it starts to rotate, it does so quickly and with large deviation from the convecting equilibrium. $R/r = 24$.

y , and z are comparable to x_1 , x_2 , and x_3 in the EM system. The Lorenz equations have been widely used in nonlinear dynamics to study chaos and in NWP as a model system for testing DA [16, 20, 23–26].

B. Direct Numerical Simulations

In order to establish an independent reference of the thermosyphon state to which the EM model may be compared, 2D direct numerical simulations of the natural convective flow have been performed using computational fluid dynamics. The computational model used here has been described in detail in a previous study [21]; however, for completeness we summarize here its essential elements.

Consistent with the preceding modeling assumptions, temperature-dependent variations of material properties are regarded as negligible, save for the density ρ . The standard Boussinesq approximation is invoked and all fluid properties are assumed to be constant and evaluated at the reference temperature T_0 . The flow is assumed to be laminar, two-dimensional,

with negligible viscous dissipation due to low velocities. Under these circumstances, the governing dimensionless equations are the unsteady, 2D laminar Navier-Stokes equations along with the energy equation. When cast in dimensionless form, the governing equations become:

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0 \quad (16)$$

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} = -\nabla p + \text{RaPr} T + \text{Pr} \nabla^2 \mathbf{u} \quad (17)$$

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T = \nabla^2 T \quad (18)$$

In the above equations \mathbf{u} is the dimensionless velocity field and T is the dimensionless temperature. The length scale used in the non-dimensionalization is based on the loop thickness $2r$ and a diffusion time is used as a basis for temporal scaling. The quantities Ra and Pr are the Rayleigh and Prandtl numbers, respectively, introduced earlier in this article. Specifically, the Rayleigh number is defined here as

$$\text{Ra} = \frac{g\gamma\Delta T_w d^3}{\nu\kappa} \quad (19)$$

where g is the gravitational acceleration, γ is the thermal expansion coefficient, ν is the kinematic viscosity, κ is the thermal diffusivity, and ΔT_w is the applied temperature differential applied to the walls. For consistency with previous study [21], the dimensions of the loop are chosen with $R = 36$ cm and $r = 1.5$ cm to yield a radius ratio of 24 (see Fig. 1). A computational mesh for this geometry consisting of $\approx 10^4$ elements has been previously shown to yield adequate spatial resolution for the simulations.

As in the classic Rayleigh-Bénard problem, the Rayleigh number determines the onset of convection in the thermosyphon. For the numerical simulations on this fixed geometry, a range of Rayleigh numbers can be imposed by varying the value of the gravitational acceleration. As the Rayleigh number is increased from zero, the flow behavior transitions from a stationary, conduction state to a steady, unidirectional state of convection. At still higher values of Ra , chaotic flow oscillations can be observed. Unless otherwise indicated, the simulation results presented in this paper correspond to a value of $\text{Ra} = 1.5 \times 10^5$.

The numerical simulations were performed using a finite-volume-based flow modeling software package (FLUENT 6.3 [27]). During the course of the simulations, flow monitors are used to record the time-varying mass flow rate within the system as well as the mean

cross-sectional temperatures at the $\phi = \pm\pi/2$ locations at 10 second intervals. In doing so, a synthetic time series of toy climate observations is recorded that can be used with the EM model in a forecasting scheme.

III. FORECASTING METHODS

A. Data Assimilation

DA is the process by which observations of a dynamical system are combined with forecasts from a model to estimate error covariances and calculate a “best guess” for the current state of the system, typically referred to as the *analysis*. In practice, this problem is difficult, because the forecaster uses an inexact forecasting model and never knows the true state of the dynamical system. The number of state variables in a NWP model is typically $\mathcal{O}(10^3)$ times the number of observations. Nevertheless, this best guess is used to produce a forecast, which is then used in the next assimilation cycle to fill in the blanks when new observations become available. Observations and forecasts are weighted depending on the confidence in each, given their error covariance matrices.

A variety of filters are capable of solving the basic DA problem for reasonably good models. The Kalman filter (KF) is the optimal algorithm for state estimation problems in a linear system. One of its first applications was to trajectory estimation and correction of missiles and rockets [28]. A number of nonlinear DA schemes are implemented in this study. In 3D variational DA (3D-Var), the background error covariance is estimated a single time, offline, prior to the data assimilation procedure. In the extended Kalman filter (EKF), it is evolved according to the linear tangent model, which approximates the evolution of small perturbations about the trajectory. Ensemble Kalman filters (EnKFs, of which the EnSRF and ETKF are examples) use ensembles of forecasts to estimate the background error and can also be used for parameter estimation. Detailed descriptions of each method, including the state update equations, are included in Appendix B.

B. Bred Vectors

To address IC uncertainty growth, an ensemble of perturbed states forming a small ball around the analysis can be used to represent uncertainty in the IC. A nonlinear system will

dynamically stretch and shrink such a ball around its trajectory as it moves through the attractor [29]. Small perturbations to points on a trajectory are integrated forward in time, and the differences between perturbed and unperturbed solutions are called *bred vectors* (BVs).

A set of $k = 3$ orthogonal vectors of length $r = 0.001$ were integrated forward m time steps, corresponding to one assimilation window. The resulting BVs populate the columns of a matrix \mathbf{Z} which is analyzed using the singular value decomposition (SVD), $\mathbf{Z} = \mathbf{U}\mathbf{S}\mathbf{V}^T$ [30]. The unitary matrix $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, \dots, \mathbf{u}_k]$ forms an orthonormal basis for a subspace of the state space which preserves the direction of maximum growth, and the rescaled vectors $r\mathbf{u}_i$ are used as the initial perturbations in the next breeding cycle. The singular values s_i , entries in the diagonal matrix \mathbf{S} , measure the exponential growth rate g_i in the direction \mathbf{u}_i according to

$$g_i \equiv \frac{1}{m} \ln \left(\frac{s_i}{r} \right), \quad (20)$$

while the average growth rate is denoted \bar{g} .

The average BV growth rate is a useful measure of local instabilities [31]. A fast BV growth rate implies a higher forecast error, which can be used to improve forecasts in DA by adaptively inflating the forecast error estimate. BVs and SVD can also be used in ensemble initialization to ensure that ensemble members are linearly independent and span the subspace of highest uncertainty [29]. Evans et al. [25], studying perfect-model forecasting of the Lorenz system, set a BV growth rate threshold which accounted for 91.4% of the observed regime changes. In Sec. IV C, we show that the same method, and other predictors, can be used to predict the regime changes of the simulated thermosyphon.

IV. METHODS AND RESULTS

A. Data Assimilation

A perfect model scenario, in which the Lorenz equations were used to forecast a synthetic truth created by that same system, was implemented in order to compare our forecast performance to previous results. We found analysis errors similar to what was reported by Yang et al. [20] (3D-Var and EKF) and Kalnay et al. [26] (ETKF), using the same model and tuning parameters.

(a)

time

de-

lay

(b)

as-

sim-

i-

lated

FIG. 4. Two views of the numerically simulated thermosyphon attractor. A 60 s time-delay reconstruction, using the monitored mass flow rate, is shown in (a). In (b), data points show x_1 and x_2 of the EM analyses generated by EKF with an assimilation window of 30 s. Each is colored by the bred vector growth rate over the prior window. Note how trajectories that move through the far edge of either lobe create distinctive loops near the center of the opposite lobe after a regime change.

As described in Section II B, the mass flow rate q and cross-sectionally averaged temperature at $\phi = \pm\pi/2$ were reported during the course of numerical simulations. The mass flow rate gave the fluid velocity $u = q/(2r\rho_0) \propto x_1$, since $2r$ is the 2D cross-sectional area. The temperature measurements were used to calculate $\Delta T_{3-9} \propto x_2$. The temperature data in our time series, however, were noisy. This is due to the neglect of higher order temperature modes and the presence of a Kelvin-Helmholtz instability [21] at the boundary between heat source and sink (see Fig. 3). These effects are not captured by the EM model, which produces smooth trajectories. Assimilation of mass flux alone proved sufficient. Gaussian noise with standard deviation equal to 6×10^{-4} kg/s, approximately 0.8% of the mass flux climatological mean $\sqrt{\langle q^2 \rangle}$, was added to the synthetic truth to create observations.

When creating forecasts for the simulated thermosyphon with the EM model, the observations y of state variable q provide the only validation, so $\text{RMSE} = \sqrt{\langle \delta q^2 \rangle}$ is calculated in observation space. The residual at a specific assimilation cycle is given by $\delta q = q - \mathbf{H}\mathbf{x}^b$. Analysis error in observation space, which can be small even for large assimilation windows,

(a)

as-

sim-

i-

la-

tion

win-

dow

=

120

s

(b)

as-

sim-

i-

la-

tion

win-

dow

=

300

s

(c)

as-

sim-

i-

la-

tion

win-

dow

15=

600

is not an appropriate metric for assessing model performance. For example, 3D-Var with a 10 minute time window yields analysis and background errors of 0.000506 and 0.0703 kg/s, respectively. The analysis error is less than the observational noise, but the background error is on the order of the climatological mean (RMS) mass flow rate $\sqrt{\langle q^2 \rangle} \approx 0.076$ kg/s. In other words, background forecasts are practically meaningless, but the filter knows this and weights the observations heavily over the background forecasts. The unobserved EM state variables are still poor predictions.

Model errors further complicate the implementation of DA in realistic forecasting scenarios. Kalnay et al. [26] found that a Lorenz forecasting model with a slightly different forcing parameter ($r = 26$ versus $r = 28$) required a 10-fold increase in the multiplicative inflation factor Δ when using a 3 member EnKF. Model error is likely even more significant for our forecasts, since the EM model is a weaker approximation of the numerically simulated thermosyphon than the near-perfect model used above. Within the KF update equations, model error appears as the \mathbf{Q} covariance matrix. We set $\mathbf{Q} = 0$ and relied upon additive and multiplicative background covariance inflation to capture model error; better results might have been obtained with an optimized \mathbf{Q} .

The performance metric used instead was the background error normalized by the climatological mean of the mass flow rate. Following the historical S1 score convention [32], relative error above 70% is considered a “useless” forecast, while under 20% the forecast is “perfect” (these numbers are somewhat arbitrary). Perfect forecasts for 3D-Var were found with up to a 3.5 minute assimilation window, while the limit was 4 minutes for the EnKFs and 4.5 minutes for EKF (see Fig. 6).

FIG. 6. Relative background RMSE over 3000 assimilation cycles with respect to $\sqrt{\langle q^2 \rangle}$ plotted for different DA algorithms against assimilation window. As the window becomes larger, the relative error increases until saturation. The lower dashed line in the main figure shows the limit of a “perfect” forecast while the upper demarcates a “useless” forecast [32]. Because 3D-Var outperforms the other methods for larger windows, a hybrid method would likely yield further improvements.

B. Regime Change Mechanism

The first explanation of regime changes was presented by Welander [11] and repeated by Creveling et al. [12]. Welander, who was also the first to discover that thermosyphons exhibit aperiodic oscillatory behavior, explained the instability of steady convecting flow by considering a thermal anomaly or “warm pocket” of fluid. For low heating rates, the convecting equilibrium is stable because viscous and thermal dissipations are in phase, thus an increase (decrease) in flow rate leads to an increase (decrease) in friction and a decrease (increase) in buoyancy, and such perturbations are damped out. At higher heating rates, the warm pocket is amplified with each cycle through the loop due to out of phase viscous and thermal dissipations. Welander explained that when the warm pocket emerges from the heating section and enters the cooling section, it feels a greater buoyant force than the surrounding fluid and accelerates, exiting the cooling section quickly, giving it less time to radiate away its energy. As the pocket moves into the heater, the buoyant force it experiences is again higher than normal, so now the pocket decelerates and passes slowly through the heating section, gaining more energy.

This positive feedback effect causes the pocket to grow hotter and larger. In the thermosyphon, it manifests as growing oscillations in the mass flow rate and temperature field. In the Lorenz system, it repels trajectories from the unstable foci at the center of each butterfly wing in the Lorenz attractor. But these oscillations do not increase forever; eventually, the pocket becomes large and hot enough that its descent towards the heating section is stopped entirely by its own buoyancy. Without movement, the pocket dissipates, but its remnant heat biases new rotation in the opposite direction. This regime change is the transition from one wing of the Lorenz attractor to the other.

Not all regime changes occur when the system reaches the same oscillation amplitude, nor do all regimes last the same amount of time. During a regime change, the fluid motion stalls after hot fluid extends across the entire heating section into the cooling section (see Figs. 2 and 7). The magnitude of this hot “tongue” and, likewise, the opposite cold tongue affects the stability of the system as it transitions between regimes. If the oscillation is small, it will mostly dissipate before the new regime is entered, bringing the temperature profile close to that of conduction. This is a highly unstable equilibrium, since the vertical temperature gradient builds until the fluid in the bottom is much hotter than the fluid

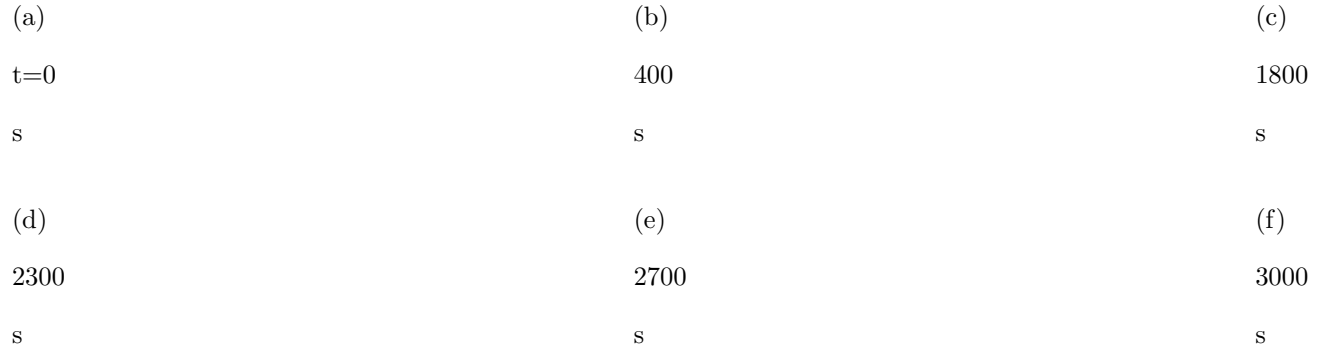


FIG. 7. Temperature profiles during two consecutive regime changes for the chaotic case; $Ra = 1.8 \times 10^4$ and $R/r = 3$. The color bar is the same as in Fig. 2. The strength of the oscillation (deviation from steady convection) shown in (a) is relatively weak; the temperature profile quickly approaches that of conduction before the flow reverses direction (c). The extreme instability of the conducting state produces a large oscillation in the clockwise direction (d), immediately causing another flow reversal back to counterclockwise. As the system enters the new regime, the fluid is well mixed and remnant heat stabilizes the flow (f), necessitating a longer residency (5 oscillations) in the new regime while the instability grows.

above (illustrated in Fig. 9). When the fluid begins to rotate, it accelerates rapidly. The large amount of heat carried by the fluid brings the system state far from the convecting equilibrium. If the oscillation is large (corresponding to a large deviation from convecting equilibrium in temperature and velocity), remnant warm and cool areas will be present in the top and bottom sections of the loop, respectively. These stabilize the new regime near its convective equilibrium. The resulting regime lasts longer since the instability requires more time to grow before causing the next reversal. These two situations are illustrated in Fig. 7. Animations of the simulated temperature field during regime reversal are consistent with this explanation [33].

C. Early Warning Signs of Regime Changes

Three separate tests were developed to predict, at each assimilation step, whether a regime change would occur within the next oscillation period ($\tau \approx 11$ minute [34]). The simplest test forecasts a regime change whenever the background forecast x_1 crosses through zero. Note that the forecast is unable to predict regime changes that occur beyond the lead time; we call this the lead forecast test. A BV threshold test was also implemented, since the BV growth rate spikes as the system transitions between regimes, as shown by Evans et al. [25]. The third test uses the fact that regime changes are suspected to be caused by out of phase viscous and thermal dissipations. Since the friction term grows with x_1 and the thermal dissipation grows with the size of the temperature anomaly $\sim x_2$, we examined the correlation between those two variables over a tunable number of previous analysis cycles. Specifically, when the slope of the least-squares linear fit of I previous analysis points $[x_2^{(i)}, x_1^{(i)}]^T$ for $i = 1, 2, \dots, I$ exceeds a threshold, we can predict a regime change with high probability. See Fig. 8. Interestingly, increasing autocorrelation of the state seems to be a universal property of many systems in advance of critical transitions [35].

The results of the three tests are shown in Table I as two-by-two contingency tables. Also shown are the threat score (TS), false alarm ratio (FAR), and hit rate (HR). Given a nonprobabilistic yes/no forecast with a hits, b false alarms, c misses, and d correct negatives for a total of n events, these are defined as $TS=a/(a+b+c)$, $FAR=b/(a+b)$, $HR=a/(a+c)$ (see [36]). Because regime changes are relatively rare events, the FARs in each case are all quite low. TS is an appropriate overall performance metric since it disregards these frequent

FIG. 8. Correlation test: Whenever the correlation (green) exceeds the threshold (red), a regime change is predicted. Successes are shown as filled circles and false positives as open circles. The starred point corresponds to the inset, which shows how correlation is computed as the slope of the least squares fit (green line) of previous analysis points. Since these points trace out the assimilated thermosyphon attractor, the correlation is larger when the system state is further from the convecting equilibrium foci. For this figure, we use a threshold of correlation > 1.42 on the previous $I=18$ points. The timeseries shown corresponds to that in Fig 5

negative events and takes into account both false alarms and misses.

Considering TS alone, the lead forecast performed best, followed by the correlation test, with the BV test performing poorest. There are also discrepancies in average regime change detection times. The lead forecast, BV, and correlation tests had average detection times of 1.3, 3.4, and 5.1 minutes respectively. We note that the correlation test performed satisfactorily in all areas which could indicate that out of phase dissipations are indeed the cause of regime changes. A heuristic that combines all three predictors would almost certainly yield improvements in accuracy across the board.

FIG. 9. The EKF-assimilated trajectory during the largest-oscillation regime change, corresponding to the bottom right-most point of Fig. 10. The system starts in the negative (counterclockwise) regime and stalls near the convecting state for a long time, causing fluid in the bottom to heat up, manifesting in a highly negative x_3 . Note that the filter is not observing this variable, but it has learned that when fluid in the thermosyphon is stalled, x_3 should increase in magnitude. The fluid turns over one more time, the wide swing in the trajectory to the left of the figure, then begins to rotate clockwise before quickly returning to the counterclockwise regime. Color indicates the 30 s assimilation window BV growth rate.

D. Predicting New Regime Duration

We found that the amplitude of the analysis x_1 -extremum preceding regime change is correlated to the duration of the following regime (shown in Fig. 10). The durations come at discrete steps corresponding to an integer number of oscillations. Furthermore, the average

TABLE I. Contingency tables for the three regime change forecasting rules. The tuning parameters found to maximize TS were: lead forecast cutoff = 3 cycles; BV threshold = 0.6991; correlation threshold = 1.42, I = 18 cycles. Note that, in optimizing the TS, the other skill scores are necessarily made suboptimal. These results are for the tuned EKF with 30 s between observations/analyses.

(a) Lead forecast: TS= 84%, FAR=13%, HR=97%, n=160542

		Observed	
		Yes	No
Fcast	Yes	4115	639
	No	135	155653

(b) BV test: TS= 63%, FAR=36%, HR=97%, n=124118

		Observed	
		Yes	No
Fcast	Yes	4131	2323
	No	119	122522

(c) Correlation test: TS= 77%, FAR=7.1%, HR=82%, n=157978

		Observed	
		Yes	No
Fcast	Yes	3473	266
	No	777	153426

BV growth rate measured over the assimilation window preceding that extremum follows a clear gradient as the growth rate increases with oscillation amplitude. The BV growth rate gradient implies that more unstable system states precede longer duration regimes. These correlations make the x_1 oscillation amplitude and BV growth rates plausible predictors for the new regime's duration. Outliers with $|x_1| \gtrsim 15$ result in shorter regime durations than expected. A possible explanation is that this is due to mixing that occurs when the temperature or mass flux oscillation is very large, effectively setting an upper limit on the number of oscillations in a regime; this phenomenon does not occur in the Lorenz or EM systems.

When a regime change is predicted by one of the above methods, we then examine

FIG. 10. Regime duration plotted against the amplitude of x_1 (proportional to the mass flow rate) at the last extremum before regime change. This amplitude is calculated from the EM model EKF analysis of thermosyphon observations, using a 30 s assimilation window. Inset, upper left: time series corresponding to a single point in the scatter plot, marked with a black cross. This figure contains almost 93 days of simulated flow and 4252 regime changes. Points are colored by the average BV growth rate over the preceding assimilation window, showing a BV growth rate gradient that increases in the positive direction along both axes. Inset, lower right: histogram showing the likelihood of different regime durations, given an x_1 amplitude between 12 and 13. The most probable duration is about 33 minutes, the middle “step” for that amplitude range.

a x_1^{max} -slice through Fig. 10, where x_1^{max} is the amplitude found from analysis and lead forecast data available at the time when the regime change is predicted. From the relative abundance of points in each step of the x_1^{max} -slice through Fig. 10, we assign a probability to each step. This generates a probabilistic forecast for the number of oscillations in the following regime. Obviously, we only do this when the regime change algorithm makes a correct prediction. The skill of this method is compared to climatology (the distribution of durations disregarding amplitude) with a ranked probability score (RPS) [36]. The lead forecast, BV, and correlation tests had RPS’s of 51%, 63%, and 48%, respectively. In each case, forecasts were made using the parameters from Table I.

V. CONCLUSIONS

DA was shown to be an effective way of coupling a low-dimensional model to simulations of the toy climate. Although background errors were always larger than observational noise, the filters were still able to produce analyses within observational noise of the truth for reasonable assimilation windows. Proper tuning of multiplicative and additive inflation factors was essential for avoiding filter divergence and achieving low forecast error. Three different predictors of regime changes were proposed and tested with reasonable success. Finally, the amplitude of the final oscillation in the current regime is correlated with the duration of the following regime (Fig. 10), and we give a physical explanation for this phenomenon. Analysis oscillation amplitudes were then used to create reliable probabilistic forecasts of regime durations.

A laboratory thermosyphon device is in construction. The next stage of this research will apply similar methods to forecasting the system state, regime changes, and regime durations. We would also like to carry out 3D numerical flow simulations. Spatial DA techniques, such as the LETKF [26, 37], could be applied to finite-volume or finite-element models. These imperfect model experiments could be used to compare the relative performance of other DA algorithms (4D-Var [26]), synchronization approaches (adaptive nudging, see [20]), and empirical correction techniques [32, 38].

Although the thermosyphon is far from representing anything as complex and vast as Earth’s weather and climate, there are characteristics our toy climate shares with global atmospheric models. Sophisticated atmospheric models are, at best, only an approximate representation of the numerous processes that govern the Earth’s climate. Global weather models and the EM model both parameterize fine-scale processes that interact nonlinearly to determine large-scale behavior. Clouds and precipitation are sub-grid-scale processes in a global weather model, and the correlations for the heat transfer and friction coefficients are parameterizations of fluid behavior on a finer scale than can be dealt with in the derivation.

The methods we use to forecast the toy model are also similar to the methods used for global geophysical systems. Both require state estimation to find the IC from which to generate forecasts. Also, when forecasts are made in either system, the climatology is often more important than specific behavior: the occurrence and duration of regime changes for the thermosyphon; periodic regime shifts like the El Niño Southern Oscillation, and statistics such as globally and regionally-averaged temperatures and their effects on rainfall, ice cover, etc. for climate. Each of these is a statistic that must be post-processed from the model output. To meet these global challenges, many techniques are needed in the modeling toolbox. In this way, toy models can provide us with insights that are applicable to the most important prediction problems of today.

ACKNOWLEDGMENTS

We would like thank Dennis Clougherty, Peter Dodds, Nicholas Allgaier, and Ross Lieb-Lappen for comments and discussion. We also wish to acknowledge financial support from the Vermont Space Grant Consortium, NASA EPSCoR, and the Vermont Advanced Computing Center.

Appendix A: EMF Parameter Estimation

MULTIPLE SHOOTING STUFF HERE

Before any forecasts could be generated, the parameters matching the EM model to the thermosyphon simulation needed to be determined. Ehrhard and Müller used experimental measurements to determine the correlation coefficients for friction, f_{w_0} , and heat transfer, h_{w_0} and K . They achieved this by opening the loop at $\phi = \pi/2$ and providing a developed flow with adjustable velocity. By measuring the pressure loss ($\propto f_w$) and heat transfer for a range of velocities, they were able to find the correlation coefficients using regression. We were unable to accomplish this with a direct numerical simulation of an open-loop geometry.

The appropriate parameters were found by a simple search of the parameter space for $\alpha \in \{1, 2, \dots, 20\}$, $\beta \in \{20, 21, \dots, 55\}$, and $K \in \{0, 0.01, \dots, 0.5\}$. This required integrating the EM system 36720 times for 2500 time units. Because of the size of the search space, this was performed in parallel using the Vermont Advanced Computing Center (VACC) cluster. Note that for more complicated models, this brute force search of the parameter space becomes computationally prohibitive. The optimal EM parameters were chosen to be those that produce solutions which most closely match the climatology of the thermosyphon system, as measured by three statistics. The first two of these statistics characterize regime change behavior: the average number of oscillations per regime and the maximum number of oscillations before a regime change. The third statistic measures the relative height of the velocity oscillations. The relative oscillation height was measured by computing of RMS distance from the convecting equilibrium and dividing by the magnitude of the convecting equilibrium itself. The three statistics were calculated first for the simulation-reported q time series, then for the x_1 time series generated for each triple in the parameter search space. Only those parameter triples that produced a maximum of 7 oscillations before regime change, the same as was found from the q data, were considered. Minimizing the 2-norm difference, between model runs and simulation data, of the remaining two statistics gave the model parameters $\mathbf{p} = (7, 33, 0.07)^T$.

With this set of parameters, the dimensionless time scale and observation operator were estimated empirically. The time scale constant of proportionality (t'/t) was given by $\tau_{\text{EM}}/\tau_{\text{sim}}$, where the τ 's are the average oscillation periods in steps of the corresponding time series (EM model and numerically simulated thermosyphon). The one nonzero entry

in the observation operator connects dimensionless velocity to mass flow rate. The velocity scale ($[\mathbf{H}]_{11} = x_1/q$) was calculated with the ratio of average velocity magnitudes $\langle |x_1| \rangle / \langle |q| \rangle$. Although this empirical method was successful, scalable methods of parameter estimation are needed for higher-dimensional models.

Appendix B: Data Assimilation Algorithms

1. Kalman Filter

The KF is well-known and widely used in linear DA and control problems. Although the thermosyphon is highly nonlinear, the linear update equations are similar to those of the nonlinear algorithms used for this experiment. The KF attempts to assimilate observations and forecasts for a process of the form

$$\mathbf{x}_k^t = \mathbf{W}\mathbf{x}_{k-1}^t. \quad (\text{B1})$$

In this case, \mathbf{x}^t is the true state, which advances in time according to the linear process \mathbf{W} , which is unknown but approximated by the model \mathbf{M} . Subscripts index the time step. Using the model, the analysis from the previous time step is integrated to generate the background forecast for the current time step

$$\mathbf{x}_k^b = \mathbf{M}\mathbf{x}_{k-1}^a \quad (\text{B2})$$

where \mathbf{M} is the linear model, \mathbf{x}^a is the old analysis, and \mathbf{x}^b is the background. Because \mathbf{M} is only an approximation of \mathbf{W} , a perfect initial condition will not lead to a perfect forecast, so

$$\mathbf{x}_k^t = \mathbf{M}\mathbf{x}_{k-1}^t + \epsilon_k^q \quad (\text{B3})$$

where the model errors ϵ^q have covariance \mathbf{Q} (usually assumed to be constant in time) and are written on the right hand side for convenience. When deemed unnecessary, time subscripts are left out.

Given an observation \mathbf{y} and background forecast \mathbf{x}^b , the KF finds the optimal way to combine them into the *analysis* \mathbf{x}^a , the best guess of the current state. This becomes the IC when forecasting with the model, Eqn. (B2). In an operational context, we usually cannot observe every state variable. If $\mathbf{x} \in \mathbb{R}^N$ and $\mathbf{y} \in \mathbb{R}^M$, then $M < N$ (in NWP $M \ll N$), so we define the observation operator $\mathcal{H} : \mathbb{R}^N \rightarrow \mathbb{R}^M$ that takes the background forecast

from the model state space into the observation space. This serves two purposes: first, it avoids extrapolation of observations to gridpoints in state space; and second, it enables us to interpret our forecasts by comparing them directly to observations. For the thermosyphon, \mathcal{H} is linear, so we write it as \mathbf{H} , but this is usually not the case for the observations in NWP, e.g., satellite radiances and radar reflectivities.

The complete application of the KF consists of a forecast step

$$\mathbf{x}_k^b = \mathbf{M}\mathbf{x}_{k-1}^a \quad (\text{B4a})$$

$$\mathbf{B}_k = \mathbf{M}\mathbf{A}_{k-1}\mathbf{M}^T + \mathbf{Q} \quad (\text{B4b})$$

and an analysis step

$$\mathbf{x}_k^a = \mathbf{x}_k^b + \mathbf{K}_k(\mathbf{y}_k - \mathbf{H}\mathbf{x}_k^b) \quad (\text{B5a})$$

$$\mathbf{A}_k = (1 - \mathbf{K}_k\mathbf{H})\mathbf{B}_k \quad (\text{B5b})$$

with the Kalman gain \mathbf{K}_k given by

$$\mathbf{K}_k = \mathbf{B}_k\mathbf{H}^T(\mathbf{H}\mathbf{B}_k\mathbf{H}^T + \mathbf{R})^{-1}. \quad (\text{B6})$$

The forecast equations create the background forecast and update the background error covariance. The new background error covariance is the old analysis error integrated forward plus the model error \mathbf{Q} . In the analysis step, this background forecast is incremented by the gain times the innovation $(\mathbf{y} - \mathbf{H}\mathbf{x}^b)$ to produce the analysis. The difference between the analysis and the background is referred to as the *analysis increment*; statistical properties of these increments can be used to reduce model error [39–41]. The new analysis error is equal to the background error reduced by a factor of $(1 - \mathbf{K}\mathbf{H})$. By finding the analysis, the filter has revealed the best possible starting point for the next background forecast. In fact, if the system is linear, the KF is the optimal algorithm for state-estimation.

2. 3D-Var

Rather than minimize the analysis error variance, the analysis equations can also be derived by finding the analysis state \mathbf{x}^a that minimizes the quadratic scalar cost function $2C(\mathbf{x}) = (\mathbf{x} - \mathbf{x}^b)^T\mathbf{B}^{-1}(\mathbf{x} - \mathbf{x}^b) + (\mathbf{y} - \mathcal{H}\mathbf{x})^T\mathbf{R}^{-1}(\mathbf{y} - \mathcal{H}\mathbf{x})$. The cost $C(\mathbf{x})$ has its minimum at $\mathbf{x} = \mathbf{x}^a$, where \mathbf{x}^a is given by Eqn. (B5). This is called the 3D variational (3D-Var)

method since the minimization for NWP is with respect to a state vector embedded in a three-dimensional field (latitude, longitude, and height).

Formally, both 3D-Var and the KF yield the same solution [32]. However, in this case the control variable is the analysis, while in the KF the control variable is the weight matrix itself. In operational NWP, where the dimension of the state space N is of $\mathcal{O}(10^9)$, the numerical implementations of 3D-Var and the nonlinear KF are drastically different. Because 3D-Var assumes the background error \mathbf{B} is fixed in time, the Kalman gain \mathbf{K} needs to be calculated only once. The calculation of \mathbf{K} is the most computationally prohibitive part of DA because it requires solving a linear system in N variables. A constant \mathbf{K} thus makes the algorithm computationally simple; the most difficult part of implementing 3D-Var is finding the optimal \mathbf{B} .

However, a static \mathbf{B} is not realistic. From a dynamical systems standpoint, uncertainty is closely related to stability, which is clearly dependent on the system state. In the thermosphere, the true background error is typically smaller when the system state is near the unstable convecting equilibria than when the state is near the more unstable conducting equilibrium. Because 3D-Var is computationally cheap, the National Centers for Environmental Prediction (NCEP) employ it to estimate ICs for the National Weather Service 14-day global forecasts. However, it cannot detect so-called “errors of the day”, state-dependent forecast errors which grow quickly but are not represented in the 3D-Var background error covariance matrix [32, 38].

3. EKF

The EKF is essentially the KF applied to a nonlinear model. Given a nonlinear model \mathcal{M} , the error covariances are updated by the *linear tangent model* $\mathbf{M} \equiv \partial\mathcal{M}/\partial\mathbf{x}|_{\mathbf{x}=\mathbf{x}^b}$ which takes the place of \mathbf{M} in Eqn. (B4b). This model propagates small perturbations around the trajectory \mathbf{x}^b forward in time. To operate on the matrix \mathbf{A} with the linear tangent model, first take the Jacobian of F (the right hand side of the nonlinear differential equation $\dot{\mathbf{x}} = F(\mathbf{x})$ which describes the model \mathcal{M}) and evaluate it at the background point \mathbf{x}^b ; call this matrix \mathbf{J} . Each column \mathbf{a}_i of \mathbf{A} , which can be thought of as an error perturbation to the analysis state, is then integrated forward in time according to the linear ODE $\dot{\mathbf{a}}_i = \mathbf{J} \mathbf{a}_i$.

Also note that if the observation operator \mathcal{H} is nonlinear, it is replaced by a similar linear

tangent model \mathbf{H} in the matrix equations (B5) and (B6). The transpose of these matrix functions are called *adjoint models*, which are used in sensitivity analysis of the state to perturbations.

Some modifications to the EKF algorithm are necessary to prevent the covariance matrices from becoming unbalanced. Because of its inherent approximations, the EKF tends to systematically underestimate the background error. As the background error shrinks to zero, the filter stops trusting the observations and follows the background trajectory. This is known as *filter divergence*. To combat this, a multiplicative inflation of $(1 + \Delta)$ was first applied to the background covariance matrix \mathbf{B} . Random numbers uniformly distributed between 0 and μ were added to the diagonal elements of \mathbf{A} after performing the analysis, the same type of additive inflation used by Yang et al. [20].

To propagate the background covariance without the explicit adjoint model, as Eqn. (B4b) would require, \mathbf{B} was first decomposed with the Cholesky factorization [30] into the product of a lower and upper diagonal matrix before its columns were integrated forward with the linear tangent model \mathbf{M} .

$$\mathbf{B}_{k-1} = \mathbf{L}_{k-1} \mathbf{L}_{k-1}^T \quad (\text{B7})$$

$$\mathbf{T}_k = \mathbf{M}_{k-1} \mathbf{L}_{k-1} \quad (\text{B8})$$

$$\mathbf{A}_k = \mathbf{T}_k \mathbf{T}_k^T + \mathbf{Q} \quad (\text{B9})$$

This guarantees symmetry for the new analysis error covariance \mathbf{A} .

4. Ensemble Filters (EnKF)

The EnKF is a method that replaces a single forecast state with an ensemble of states. The spread of the ensemble about its mean gives an approximation of the background error covariance and forecast uncertainty, while the ensemble average gives the best guess of the forecast. The EnKF was first introduced by Evensen [42], and he gives a comprehensive overview of ensemble filters in [43]. It was shown that if the observation, which has random error with covariance \mathbf{R} , is perturbed with P random errors (again with covariance \mathbf{R}), to make an P -member ensemble of independent observations $\{\mathbf{y}_i\}$, then the background error

covariance can be written [43]

$$\mathbf{B} \approx \frac{1}{P-1} \sum_{i=1}^P (\mathbf{x}_i^b - \overline{\mathbf{x}^b})(\mathbf{x}_i^b - \overline{\mathbf{x}^b})^T = \frac{1}{P-1} \mathbf{X}^b \mathbf{X}^{bT} \quad (\text{B10})$$

which is simply the unbiased average outer product of background perturbations $\mathbf{X}^b = [\mathbf{x}_1^b, \dots, \mathbf{x}_P^b]$. The background forecast of ensemble member i is denoted \mathbf{x}_i^b , $\overline{\mathbf{x}^b}$ is the background forecast ensemble average, and $\mathbf{x}_i'^b = \mathbf{x}_i^b - \overline{\mathbf{x}^b}$ is the i th member's deviation from the mean. In this case, each ensemble member is updated according to the KF equations for their associated observation.

As with the EKF, ensemble filters tend to underestimate the background error, resulting in an ensemble spread which is typically less than the RMSE found when verifying. We used multiplicative inflation of \mathbf{B} by $(1 + \Delta)$, a common method shown to be successful in [20, 24, 26, 43, 44].

The advantages of the EnKF are many: there is no linear tangent model to compute, the number of ensemble members can be small ($\mathcal{O}(10^2)$ for NWP) relative to the dimensionality of the state space, and prior knowledge about the structure of the forecast errors is not necessary. Currently, 4D-Var (like 3D-Var but also taking into account older observations) and ensemble filters are the most promising candidates being considered to replace 3D-Var in operational NWP.

5. EnSRF

The original EnKF adds noise to create linearly independent observations and is classified as a *perturbed observations* method [26]. This necessarily introduces additional sampling error into the forecast. For this reason, Whitaker and Hamill [44] introduced the ensemble square root filter (EnSRF) as an improved EnKF. In the EnSRF, the ensemble mean is updated with the traditional Kalman gain (Eqn. (B6))

$$\overline{\mathbf{x}^a} = \overline{\mathbf{x}^b} + \mathbf{K}(\mathbf{y} - \mathcal{H}\overline{\mathbf{x}^b}) \quad (\text{B11})$$

and deviations from the mean are updated by

$$\mathbf{X}^a = (1 - \tilde{\mathbf{K}}\mathcal{H})\mathbf{X}^b \quad (\text{B12})$$

where

$$\tilde{\mathbf{K}} = \mathbf{B}\mathbf{H}^T \left[\left(\sqrt{\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}} \right)^{-1} \right]^T \times \left[\sqrt{\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}} + \sqrt{\mathbf{R}} \right]^{-1}. \quad (\text{B13})$$

When the observation is a scalar, it can be shown that

$$\tilde{\mathbf{K}} = \left(1 + \sqrt{\frac{\mathbf{R}}{\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}}} \right)^{-1} \mathbf{K}. \quad (\text{B14})$$

If observation errors are uncorrelated (\mathbf{R} is diagonal), then Eqn. (B14) can be used to process observations one at a time [44]. The updated analysis ensemble is then $\{\mathbf{x}_i^a\}$, where $\mathbf{x}_i^a = \overline{\mathbf{x}}^a + \mathbf{x}_i'^a$. Square root filters have better numerical stability and speed than their standard KF counterparts. For these reasons, the Potter square root filter was employed for navigation in the Lunar Module of the Apollo program [28].

6. ETKF

The ETKF is another type of deterministic square root filter. In this variant, the analysis perturbations are assumed to be equal to the background perturbations postmultiplied by a transformation matrix \mathbf{T} so that the analysis error covariance satisfies Eqn. (B5b). The analysis covariance is written

$$\mathbf{A} = \frac{1}{P-1} \mathbf{X}^a \mathbf{X}^{aT} = \mathbf{X}^b \hat{\mathbf{A}} \mathbf{X}^{bT}$$

where $\hat{\mathbf{A}} = [(P-1)\mathbf{I} + (\mathbf{H}\mathbf{X}^b)^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{X}^b)]^{-1}$. The analysis perturbations are $\mathbf{X}^a = \mathbf{X}^b \mathbf{T}$, where $\mathbf{T} = [(P-1)\hat{\mathbf{A}}]^{1/2}$. See Kalnay et al. [26] for further details.

The local ensemble transform filter (LETKF) is a variant that computes the analysis at a given gridpoint using only local observations. This allows for efficient parallelization. Localization removes spurious long-distance correlations from \mathbf{B} and allows greater flexibility in the global analysis by allowing different linear combinations of ensemble members at different spatial locations [26, 37].

[1]

[2] 08 (1).

[3] D. J. Patil, E. Ott, B. R. Hunt, E. Kalnay, and J. A. Yorke, Physical Review Letters **86**, 5878 (2001).

- [4] V. Belessiotis and E. Mathioulakis, *Solar Energy* **72**, 307 (2002).
- [5] M. H. Beitelmal and C. D. Patel, Tech. Rep., HP Labs (2002).
- [6] A. Lustgarten, *FORTUNE Magazine* **153** (2006).
- [7] B. Beine, V. Kaminski, and W. Von Lensa, *Nuclear Engineering and Design* **136**, 135 (1992).
- [8] W. Kwant and C. E. Boardman, *Nuclear Engineering and Design* **136**, 111 (1992).
- [9] E. N. Lorenz, *Journal of the Atmospheric Sciences* **20**, 130 (1963).
- [10] J. B. Keller, *Journal of Fluid Mechanics* **26**, 599 (1966).
- [11] P. Welander, *Journal of Fluid Mechanics* **29**, 17 (1967).
- [12] H. F. Creveling, J. F. De Paz, J. Y. Baladi, and R. J. Schoenhals, *Journal of Fluid Mechanics* **67**, 65 (1975).
- [13] M. Gorman and P. J. Widmann, *Physical Review Letters* **52**, 2241 (1984).
- [14] M. Gorman, P. J. Widmann, and K. A. Robbins, *Physica D* **19**, 255 (1986).
- [15] P. Ehrhard and U. Müller, *Journal of Fluid Mechanics* **217**, 487 (1990).
- [16] P. K. Yuen and H. H. Bau, *Physics of Fluids* **11**, 1435 (1999).
- [17] Y. Y. Jiang and M. Shoji, *Journal of Heat Transfer* **125**, 612 (2003).
- [18] E. A. Burroughs, E. A. Coutias, and L. A. Romero, *Journal of Fluid Mechanics* **543**, 203 (2005).
- [19] G. Desrayaud, A. Fichera, and M. Marcoux, *International Journal of Heat and Fluid Flow* **27**, 154 (2006).
- [20] S.-C. Yang, D. Baker, H. Li, K. Cordes, M. Huff, G. Nagpal, E. Okereke, J. Villafañe, E. Kalnay, and G. S. Duane, *Journal of the Atmospheric Sciences* **63**, 2340 (2006).
- [21] E. H. Ridouane, C. M. Danforth, and D. L. Hitt, *Int. J. of Heat Mass Transfer* (2009).
- [22] *MATLAB Manual*, 3 Apple Hill Drive, Natick, MA 01760-2098 USA (2008).
- [23] R. N. Miller and M. Ghil, *Journal of the Atmospheric Sciences* **51**, 1037 (1994).
- [24] J. D. Annan and J. C. Hargreaves, *Tellus* **56A**, 520 (2004).
- [25] E. Evans, N. Bhatti, J. Kinney, L. Pann, M. Peña, S.-C. Yang, E. Kalnay, and J. Hansen, *Bulletin of the American Meteorological Society* pp. 520–524 (2004).
- [26] E. Kalnay, H. Li, T. Miyoshi, S.-C. Yang, and J. Ballabrera-Poy, *Tellus* **59A**, 758 (2007).
- [27] *FLUENT Manual*, Centerra Resource Park, 10 Cavendish Court, Lebanon, NH 03766 USA (2006).
- [28] R. T. Savely, B. F. Cockrell, and S. Pines, Tech. Rep., NASA (1972).

- [29] C. M. Danforth and J. A. Yorke, Physical Review Letters **96**, 144102 (2006).
- [30] G. H. Golub and C. F. V. Loan, *Matrix Computations* (Johns Hopkins University Press, 1996).
- [31] M. J. Hoffman, E. Kalnay, J. A. Carton, and S.-C. Yang, Geophysical Research Letters **36** (2009).
- [32] E. Kalnay, *Atmospheric Modeling, Data Assimilation and Predictability* (Cambridge University Press, 2002), ISBN 0521796296.
- [33] Note1, a movie of the case shown in Fig. 7 is provided online at <http://www.uvm.edu/~kharris/thermosyphon/T-Ra-18000-new.mp4>.
- [34] Note2, in laboratory experiments, the period of oscillations is on the order of 1 minute. The 11 minute period we observed is a result of 2D simulations.
- [35] M. Scheffer, J. Bascompte, W. A. Brock, V. Brovkin, S. R. Carpenter, V. Dakos, H. Held, E. H. Van Nes, M. Rietkerk, and G. Sugihara, Nature **461**, 53 (2009), ISSN 1476-4687.
- [36] D. S. Wilks, *Statistical Methods in the Atmospheric Sciences*, vol. 59 of *International Geophysics Series* (Academic Press, 1995).
- [37] B. R. Hunt, E. J. Kostelich, and I. Szunyogh, Physica D **230**, 112 (2007).
- [38] H. Li, E. Kalnay, T. Miyoshi, and C. M. Danforth, Monthly Weather Review **137**, 3407 (2009).
- [39] C. M. Danforth, E. Kalnay, and T. Miyoshi, Monthly Weather Review **135**, 281 (2007).
- [40] C. M. Danforth and E. Kalnay, Journal of the Atmospheric Sciences **65**, 1467 (2008).
- [41] C. M. Danforth and E. Kalnay, Geophysical Research Letters **35**, L24805 (2008).
- [42] G. Evensen, Journal of Geophysical Research **99**, 10143 (1994).
- [43] G. Evensen, Ocean Dynamics **53**, 343 (2003).
- [44] J. S. Whitaker and T. M. Hamill, Monthly Weather Review **130**, 1913 (2002).