Data Assimilation and Genetic Algorithms for the Parameter Estimation Problem in Simple Climate Models

Morgan R. Frank[1], Andrew Reagan[2]

Computational Story Lab, Department of Mathematics and Statistics, Vermont Complex Systems Center, Vermont Advanced Computing Core,

University of Vermont, Burlington, Vermont, United States of America

[1]mrfrank@uvm.edu, [2]andrew.reagan@uvm.edu

December 10, 2013

Abstract

Given observations of an atmospheric phenomenon and a well-principled model of that phenomenon, the parameters for the model must be properly tuned if the model is to mimic the data. We investigate the use of genetic algorithm in comparison to data assimilation as a means of performing parameter estimation when tuning models to data. We compare results while tuning chaotic dynamics, observation noise and frequency, and system dimensionality while performing parameter estimation for the Lorenz '63 and Lorenz '96 systems.

1 Introduction

Weather forecasts have become an expected part of everyday life in the modern society. Things like air-travel, disaster preparation, and daily planning rely on accurate predictions [13]. However, predicting future states of the atmosphere proves to be difficult, as chaotic systems exhibit sensitive dependence of initial conditions, and the underlying processes in weather are known to be chaotic [14–16]. This hurdle is overcome by utilizing computationally expensive global forecasting systems (GFS) for prediction and advanced methods for initial condition determination, but scientists working to improve weather forecasting often lack the time or computational power to execute many high resolution GFS experiments. Instead, climate scientists often use simple models that account for particular aspects of the weather forecasting problem.

Edward Lorenz has made major contributions to the fields of dynamical systems and atmospheric prediction [17–19]. Two such contributions are the wildly popular Lorenz '63 system [20] and the Lorenz '96 system [21]. The Lorenz '63 system (L63), which yields the widely known Lorenz Attractor, is a simple three-variable model with highly tunable dynamics, allowing researchers a computationally tractable means to experiment in the predictability of chaotic systems. The Lorenz '96 system (L96) exhibits tunable chaotic dynamics as well, while additionally providing a computationally tractable way to change the system dimensions and tune the accuracy of data observations. Both systems provide interesting and computationally manageable test beds for the parameter estimation problem across several different types of systems. Figure 1 shows example trajectories for each system.

2 Methods

2.1 The Lorenz '63 Model

In 1962, Barry Saltzmann attempted to model convection in a Rayleigh-Bérnard cell by reducing the equations of motion into their core processes [12]. Then in 1963 Edward Lorenz reduced this system ever further to 3 equations, leading to his landmark discovery of deterministic non-periodic flow [11]. This system, which we will call the Lorenz 63 system, exhibits sensitive dependence on initial conditions, meaning that small errors in an approximation will lead to exponential error growth. These equations have since been the subject of intense study and have changed the way we view prediction and determinism, remaining the simple system of choice for examining

1

nonlinear behaivor today [10]. The three equations are:

$$\frac{dx}{dt} = \sigma(y-x)$$

$$\frac{dy}{dt} = \rho x - y - xz$$

$$\frac{dz}{dt} = xy - \beta z.$$

The cannonical choice of $\sigma = 10, \beta = 8/3$ and $\rho = 28$ produce the well known butterfly attractor, and to adjust the strength of nonlinearity (chaos) we tune the ρ parameter.
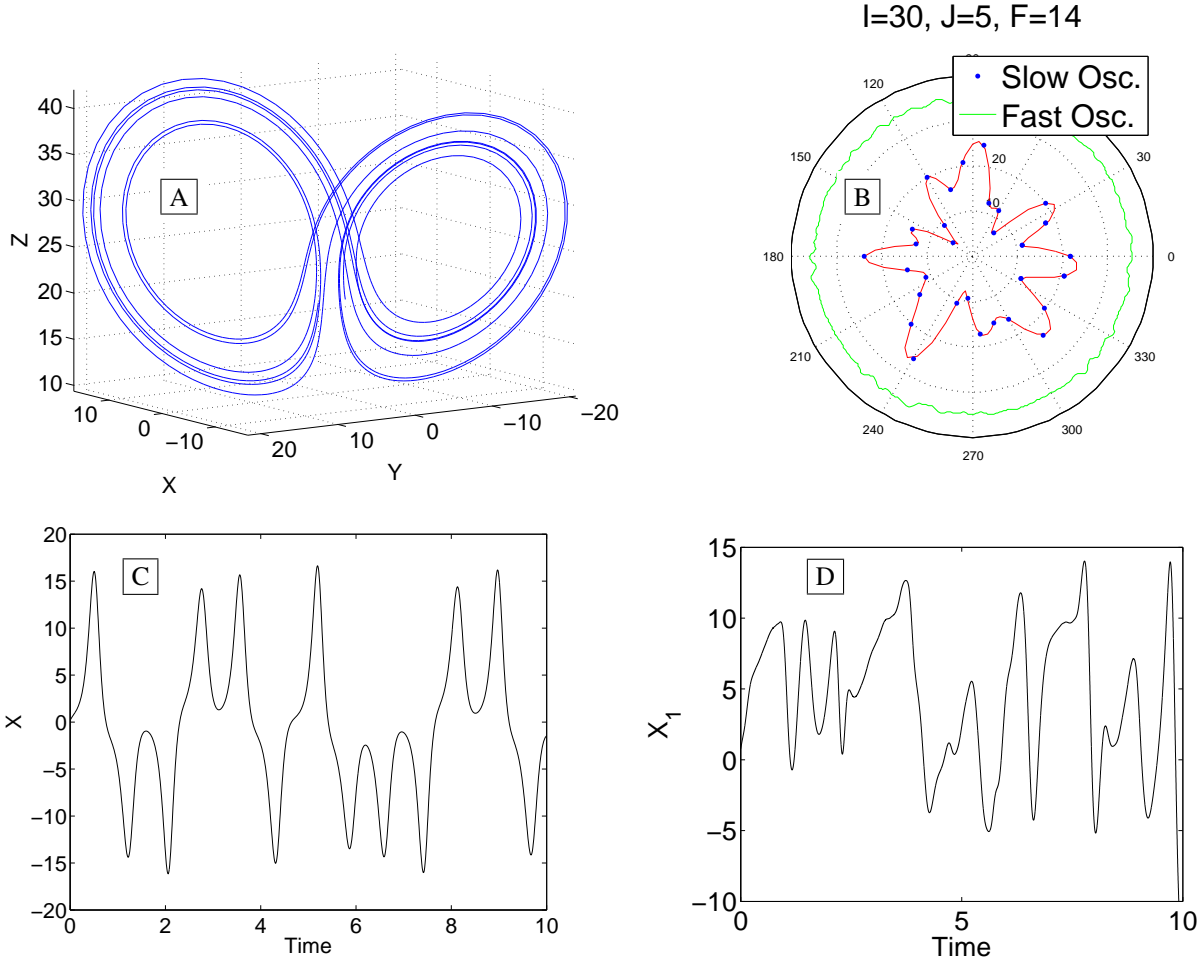

Figure 1. (A) The popular "Lorenz Attractor" produced with the Lorenz '63 system. This three-variable system produces a "butterfly"-like chaotic attractor that is well-known among fractal and chaos enthusiasts. (B) An snapshot of a trajectory of the Lorenz '96 system. Each blue point is a slow oscillator, and the adjacent sections of green represent the fast oscillators coupled with the corresponding slow oscillator. The origin represents the lowest value achieved by any of the slow oscillators on this trajectory. The red line is a cubic spline interpolation of the blue data points. (C) An example trajectory of the X variable from the Lorenz '63 system. (D) An example trajectory for a slow oscillator of the Lorenz '96 system.

Parameter	Values Explored	Interpretation, if any
Observed Variables (63 Only)	$[x_1,\text{all}]$	Limited observations
Observational Noise	Normal in $[0,.01,.05,.1,.25,.5,1,2]$ Uniform in $[0,.5,2,4,6,8,10]$	Measurement and representativeness errors
Nonlinearity	ρ in $[22,28,35]$ or I in $[4,8,10,15]$	Chaotic behavior
Subsampled observations	$[1,5,25,50]$	Infrequent observations

Table 1: Experimental parameter choices on which we test the performance of Data Assimilation and a Genetic Algorithm for fitting model parameters.

2.2 The Lorenz '96 Model

In 1995, Edward Lorenz introduced the following I-dimensional model [17, 19]. The key characteristics of this model include tunable chaotic behavior when subject to enough forcing, and tunable dimensionality. The predecessor to the current model is given by

$$\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - x_i + F \tag{1}$$

where $i = 1, 2, \ldots, I$ and F is the forcing parameter. Each x_i represents observations of some atmospheric quantity, like temperature, evenly distributed about a given latitude of the globe. This implies a modularity in the indexing that is described by $x_{i+I} = x_{i-I} = x_i$.

This early model failed to produce realistic growth rate of the large-scale errors along with lacking tenability in observation reliability. Lorenz went on to introduce a more flexible model in 1996 by coupling two systems similar to the model in equation (1), but differing in time scales. The equations for the Lorenz '96 model [21] are given as

$$\frac{dx_i}{dt} = x_{i-1}(x_{i+1} - x_{i-2}) - x_i + F - \frac{hc}{b}\sum_{j=1}^{J} y_{(j,i)} \tag{2}$$
$$\frac{dy_{(j,i)}}{dt} = cby_{(j+1,i)}(y_{(j-1,i)} - y_{(j+2,i)}) - cy_{(j,i)} + \frac{hc}{b}x_i \tag{3}$$

where $i = 1, 2, \ldots, I$ and $j = 1, 2, \ldots, J$. The parameters b and c indicate the time scale of solutions to equation (3) relative to solutions of equation (2), and h is the coupling parameter. The coupling term can be thought of as a parameterization of dynamics occurring at a spatial and temporal scale unresolved by the x variables. Again, each x_i represents an atmospheric observation about a latitude that oscillates in slow time, and the set of $y_{(j,i)}$ are a set of J fast time oscillators that act as a damping force on x_i. The y's exhibit a similar modularity described by $y_{(j+IJ,i)} = y_{(j-IJ,i)} = y_{(j,i)}$.

2.3 Data Assimilation

Areas as disparate as quadcopter stabilization [1] to the tracking of ballistic missile re-entry [2] use data assimi-

lation. The purpose of data assimilation in weather prediction is defined by Talagrand as "using all the available information, to determine as accurately as possible the state of the atmospheric (or oceanic) flow." [3] The data assimilation algorithm that we use here, the Kalman filter, was originally implemented in the navigation system of Apollo program [4, 5].

Data assimilation algorithms consist of a 3-part cycle: predict, observe, and assimilate. Formally, the data assimilation problem is solved by minimizing the initial condition error in the presence of specific constraints. The prediction step involves making a prediction of the future state of the system, as well as the error of the model, in some capacity. Observing systems come in many flavors: rawindsomes and satellite irradiance for the atmosphere, temperature and velocity reconstruction from sensors in experiments, and sampling the market in finance. Assimilation is the combination of these observations and the predictive model in such a way that minimizes the error of the initial condition state, which we denote the analysis.

In addition to determining the initial conditions, we can extend the Extended Kalman Filter (EKF) to determine the model parameters. This is accomplished by considering the model parameters as variables of the model itself, with their differential equation being equal to 0, since they do not change with the solution. The value of this consideration is that the covariance of the model variables and model parameters is now included in the Tangent Linear Model (the Jacobian of the extended analytical system) and hence is updated by the Kalman gain matrix.

The formulation of the filter we employ is the standard formulation, since the incorporation of parameters into the estimation is independent of the filter itself. Using the notation of Kalnay [6], this amounts to making a forecast with the nonlinear model M (either Lorenz 63 or Lorenz 96 in this study), and updating the error covariance matrix $\mathbf{P}$ with the TLM L, and adjoint model L^T

$$\mathbf{x}^f(t_i) = M_{i-1}[\mathbf{x}^a(t_{i-1})]$$
$$\mathbf{P}^f(t_i) = L_{i-1}\mathbf{P}^a(t_{i-1})L_{i-1}^T + \mathbf{Q}(t_{i-1})$$

where $\mathbf{Q}$ is the noise covariance matrix (model error). In the experiments here, $\mathbf{Q} = 0$ since our model is perfect. In NWP, $\mathbf{Q}$ must be approximated, e.g. using statistical moments on the analysis increments [7, 8]. The analysis step is then written as (for H the observation operator):

$$\mathbf{x}^a(t_i) = \mathbf{x}^f(t_i) + \mathbf{K}_i \mathbf{d}_i \qquad (4)$$

$$\mathbf{P}^a(t_i) = (\mathbf{I} - \mathbf{K}_i \mathbf{H}_i)\mathbf{P}^f(t_i) \qquad (5)$$

where

$$\mathbf{d}_i = \mathbf{y}_i^o - \mathbf{H}[x^f(t_i)]$$

is the innovation. The Kalman gain matrix is computed to minimize the analysis error covariance P_i^a as

$$\mathbf{K}_i = \mathbf{P}^f(t_i)\mathbf{H}_i^T[\mathbf{R}_i + \mathbf{H}_i \mathbf{P}^f(t_i)\mathbf{H}^T]^{-1}$$

where $\mathbf{R}_i$ is the observation error covariance. Since we are making observations of the truth with known standard deviation ε, the observational error covariance matrix $\mathbf{R}$ is a diagonal matrix with the standard deviatoin ε along the diagonal. This information is an additional assumption, we could however not use this information and simply sample ε as a part of the experiment. The most difficult, and most computationally expensive, part of the EKF is deriving and integrating the TLM. Here we use a differentiated Runge-Kutta scheme of 2-nd order to accurately integrate the TLM. For more details on this implemenation, see Reagan [9].

2.4 Genetic Algorithm

2.5 Experiment

For both of the systems, we study the performance of our parameter estimation scheme under varying observational noise, observational density, observational frequency, nonlinearity and dimension. This amounts to 4 (Lorenz 63) and 5 (Lorenz 96) dimensions of the experiment, and we outline the specific choices for each experiment in Table 1. The experiments were chosen to mimic realisitic conditions under which simple models are fit to data.

3 Results

4 Discussion

References

[1] Achtelik, M., T. Zhang, K. Kuhnlenz, and M. Buss (2009). Visual tracking and control of a quadcopter using a stereo camera system and inertial sensors. In *Mechatronics and Automation, 2009. ICMA 2009. International Conference on*, pp. 2863–2869. IEEE.

[2] Siouris, G. M., G. Chen, and J. Wang (1997). Tracking an incoming ballistic missile using an extended interval kalman filter. *Aerospace and Electronic Systems, IEEE Transactions on 33*(1), 232–240.

[3] Talagrand, O. (1997). Assimilation of observations, an introduction. *JOURNAL-METEOROLOGICAL SOCIETY OF JAAN SERIES 2 75*, 81–99.

[4] Kalman, R. and R. Bucy (1961). New results in linear prediction and filtering theory. *Trans. AMSE J. Basic Eng. D 83*, 95–108.

[5] Savely, R., B. Cockrell, , and S. Pines (1972). Apollo experience report - onboard navigational and alignment software. *Technical Report*.

[6] Kalnay, E. (2003). *Atmospheric modeling, data assimilation, and predictability*. Cambridge university press.

[7] Danforth, C. M., E. Kalnay, and T. Miyoshi (2007). Estimating and correcting global weather model error. *Monthly weather review 135*(2), 281–299.

[8] Li, H., E. Kalnay, T. Miyoshi, and C. M. Danforth (2009). Accounting for model errors in ensemble data assimilation. *Monthly Weather Review 137*(10), 3407–3419.

[9] Reagan, A (2013). Predicting Flow Reversals in a Computational Fluid Dyanmics Simulated Thermosyphon Using Data Assimilatoin. University of Vermont Master's Thesis, *arXiv:1312.2142 [math.DS]*.

[10] Kalnay, E., H. Li, T. Miyoshi, S.-C. YANG, and J. BALLABRERA-POY (2007). 4-d-var or ensemble kalman filter? *Tellus A 59*(5), 758–773.

[11] Lorenz, E. N. (1963). Deterministic nonperiodic flow. *Journal of the atmospheric sciences 20*(2), 130–141.

[12] Saltzman, B. (1962). Finite amplitude free convection as an initial value problem-i. *Journal of the Atmospheric Sciences 19*(4), 329–341.

[13] R. A. Kerr. *Weather Forecasts Slowly Clearing Up*. Science 9 November 2012: Vol. 338 no. 6108 pp. 734-737 DOI: 10.1126/science.338.6108.734

[14] Farmer, J. D., and J. J. Sidorowich. *Predicting Chaotic Time Series.* Phys. Rev. Lett. 59(8) (1987): 845-848.

[15] D. Orrell, *Role of the Metric in Forecast Error Growth: How Chaotic is the Weather?*, Tellus 54A (2002) 350362.

[16] C. M. Danforth, J. A. Yorke. 2006. *Making Forecasts for Chaotic Physical Processes.* Physical Review Letters, 96, 144102.

[17] E.N. Lorenz, *Predictability A problem partly solved*, in: ECMWF Seminar Proceedings on Predictability, Reading, United Kingdom, ECMWF, 1995, pp. 118.

[18] E. N. Lorenz. *The predictability of a flow which possesses many scales of motion*. **Tellus XXI**, 289 (1968).

[19] E.N. Lorenz, K.A. Emanuel, *Optimal sites for supplementary weather observations: simulation with a small model*, J. Atmos. Sci. 55 (1998) 399414.

[20] Lorenz, Edward N., 1963: *Deterministic Nonperiodic Flow. J. Atmos. Sci.*, **20**, 130.141.

[21] E. N. Lorenz, Proc. Seminar on Predictability 1, 1 (1996).

[22] R. England. *Error estimates for Runga-Kutta type solutions to systems of ordinary differential equations. The Computer Journal* (1969) 12 (2): 166-170. doi: 10.1093/comjnl/12.2.166

[23] D. S. Wilks, *Effects of Stochastic Parametrizations in the Lorenz 96 System*, Quart. J. Roy. Meteo. Soc. 131 (2005) 389407.