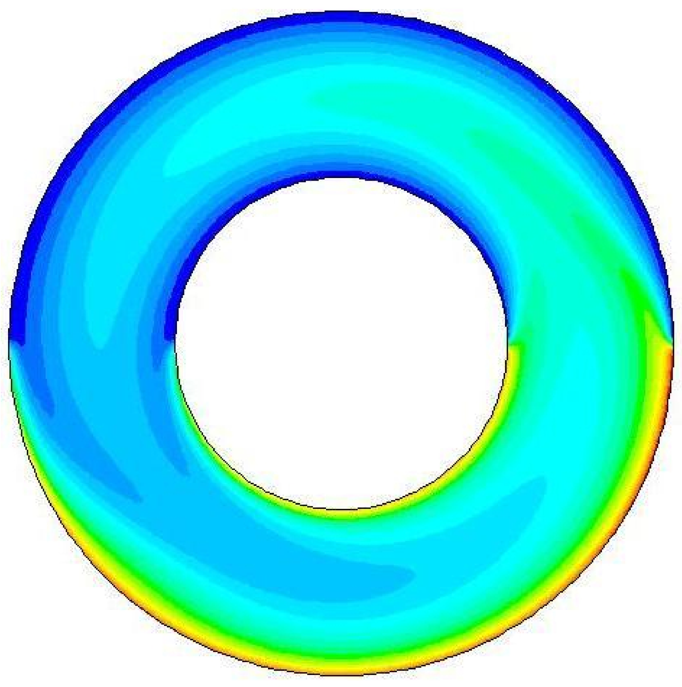
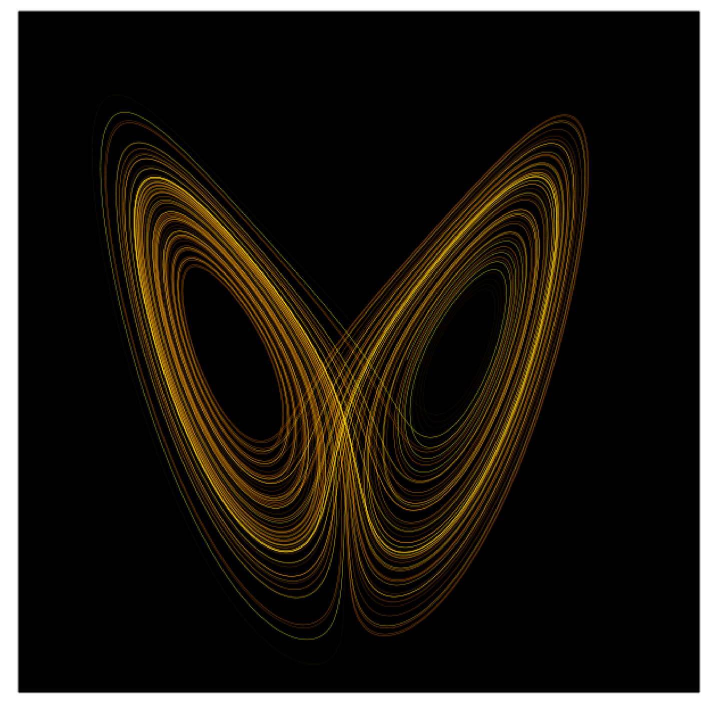
Empirical Correction of Models of Chaotic Phenomena

Nicholas A. Allgaier, Kameron D. Harris, and Christopher M. Danforth
College of Engineering and Mathematical Sciences, University of Vermont

A 3-Dimensional Model of a "Toy Climate" Convection Loop



Natural *convection* in the atmosphere is fundamental to the weather patterns we observe every day. The process can be isolated for study by filling a hula-hoop shaped loop with fluid, and then heating (cooling) the bottom (top) half to create a temperature inversion **(Left)** [1]. A high-dimensional CFD simulation of such a loop was generated to represent a "toy climate", for which we attempt to make forecasts with a 3-dimensional model.

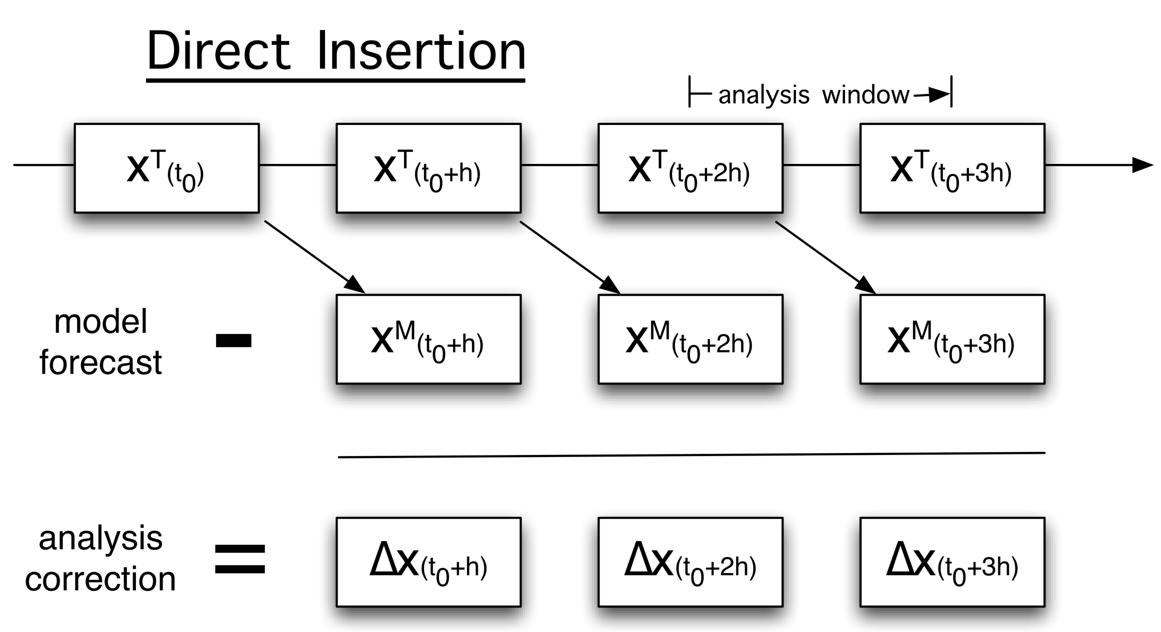


(Right) System of ODEs developed by Lorenz to model convection [2], and **(Left)** a visualization of the *chaotic attractor* of the system. Unstable convective equilibria are located at the centers of the two lobes, representing clockwise and counter-clockwise flow. A qualitatively identical system was used as the forecast model for this study.

$$\dot{x}_1 = \sigma(x_2 - x_1)$$
$$\dot{x}_2 = \rho x_1 - x_2 - x_1 x_3$$
$$\dot{x}_3 = x_1 x_2 - \beta x_3$$

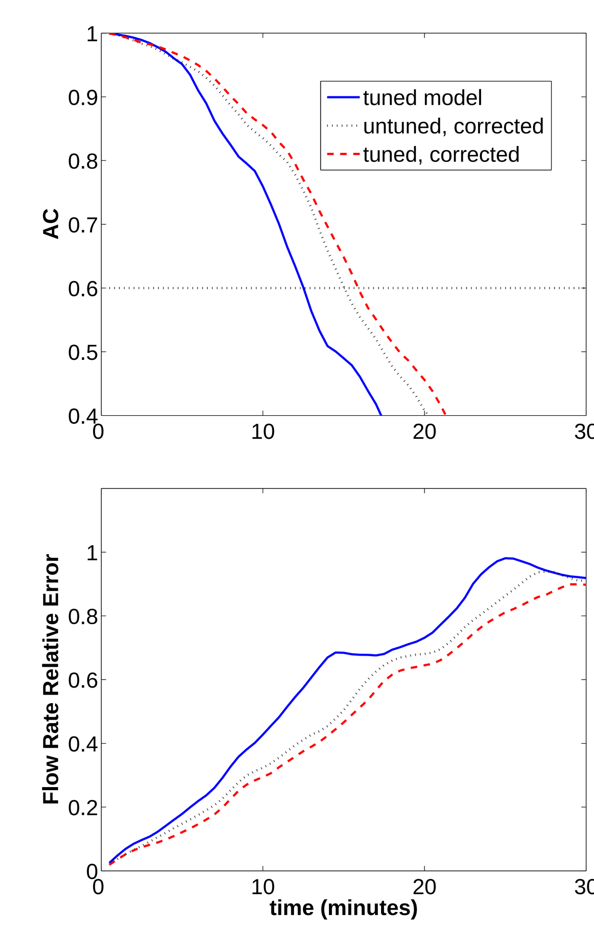
- $x_1 \propto$ flow velocity
- $x_2 \propto$ temperature difference across tube
- $x_3 \propto$ deviation of the convecting temperature profile from that of conduction (linear)
- $\beta \backsim$ tube geometry
- $\sigma =$ Prandtl number
- $\rho =$ Rayleigh number

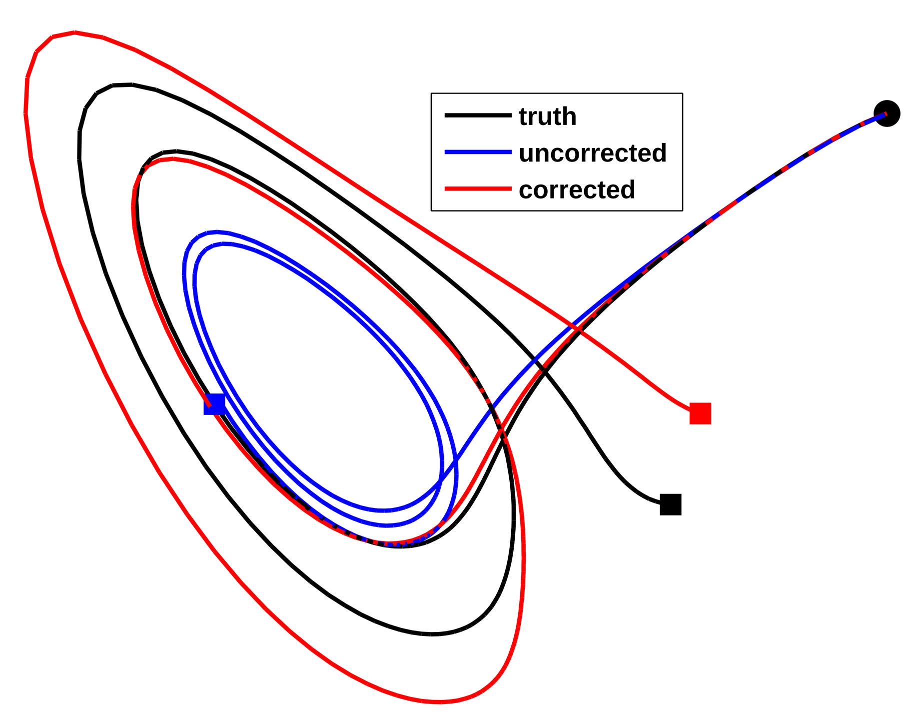
Learning from the Past: Empirical Correction of the Forecast Model



Model accuracy can be improved by comparing short model forecasts to the true (or best estimate) system state over a training period. Effectively, by analyzing the "past mistakes" of the model, the error in future forecasts can be reduced. **(Above)** The direct insertion procedure for comparing short model forecasts to the truth. $\mathbf{x}^T$ represents a vector time series of the reference truth, and the analysis window h is the number of timesteps between estimations of the true system state. $\mathbf{x}^M$ represents a vector time series of forecasts with duration equal to the analysis window, each of which is started from the previous true state. The time-average of the analysis corrections $\langle \Delta \mathbf{x} \rangle$ divided by the analysis window h approximates the *model bias* $\mathbf{b}$, which is independent of system state. A model correction $\mathbf{Lx}$ that depends on the system state $\mathbf{x}$ can further improve accuracy. The correction operator $\mathbf{L}$ is computed by $\mathbf{L} \equiv \mathbf{C}_{\Delta x x^T} \mathbf{C}^{-1}_{x^T x^T}$, the product of the cross-covariance $\mathbf{C}_{\Delta x x^T}$ between the analysis corrections and the reference truth, and the inverse of the true state covariance $\mathbf{C}^{-1}_{x^T x^T}$ [3, 4]. The model is adjusted by adding the bias term $\mathbf{b}$ and the state-dependent correction $\mathbf{Lx}$ to the model prediction at every timestep.

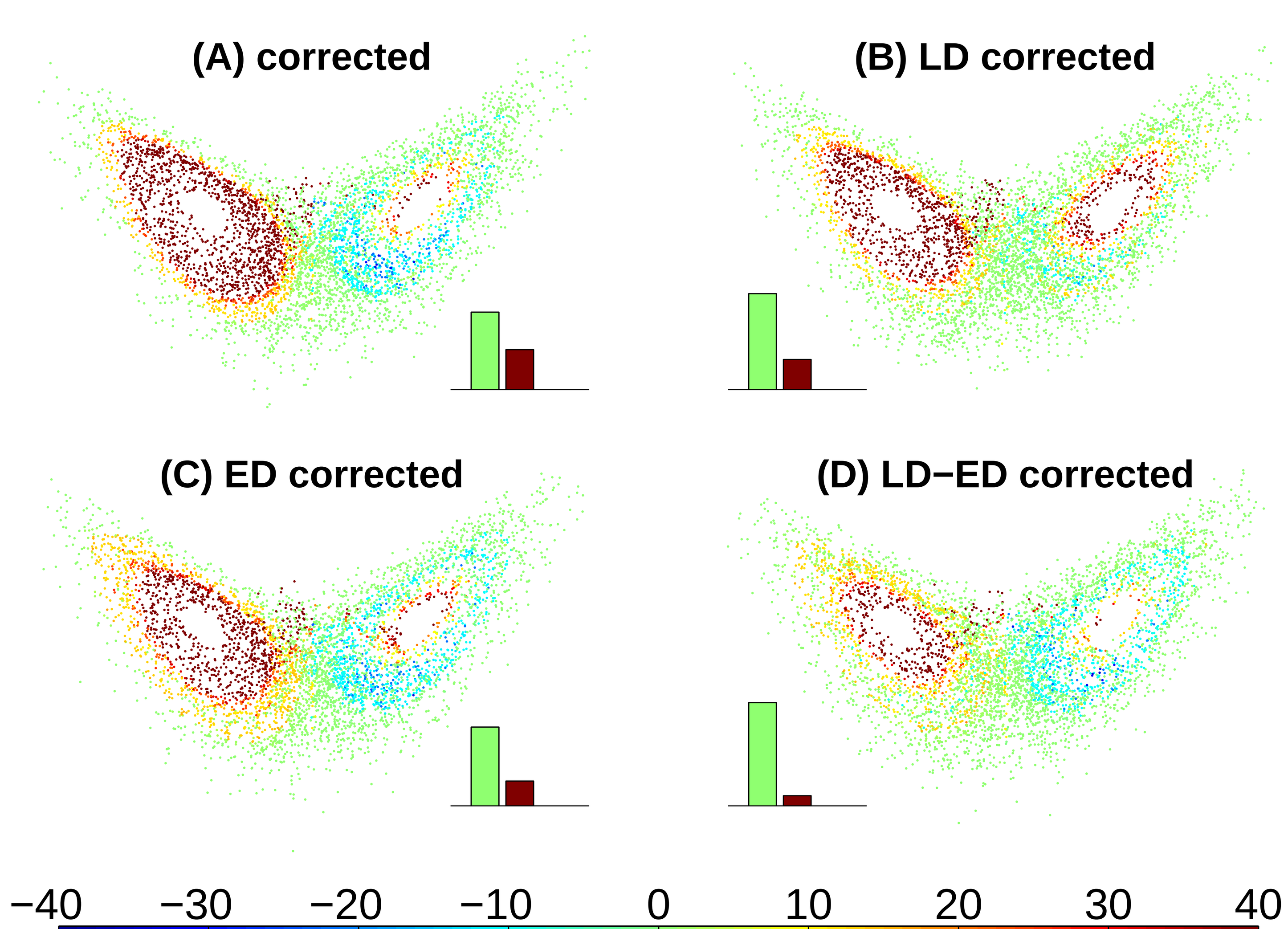
To test the effectiveness of empirical correction, 1000 trial forecasts were performed starting from randomly selected, independent initial states after the training period. **(Below)** One particularly positive outcome. Trajectories of the true system (black), uncorrected model (blue) and corrected model (red) all starting from the same initial state (black circle). The corrected model remains close to the truth for far longer than the uncorrected model. **(Right)** Average error statistics over the 1000 trials. (top) A forecast is generally considered useful for as long as its anomaly correlation (AC) remains above 0.6 [5]. Empirical correction increases the average duration of usefulness by approximately 30%. (bottom) Average error in predicted fluid velocity x_1, taken relative to the natural variability of the system.





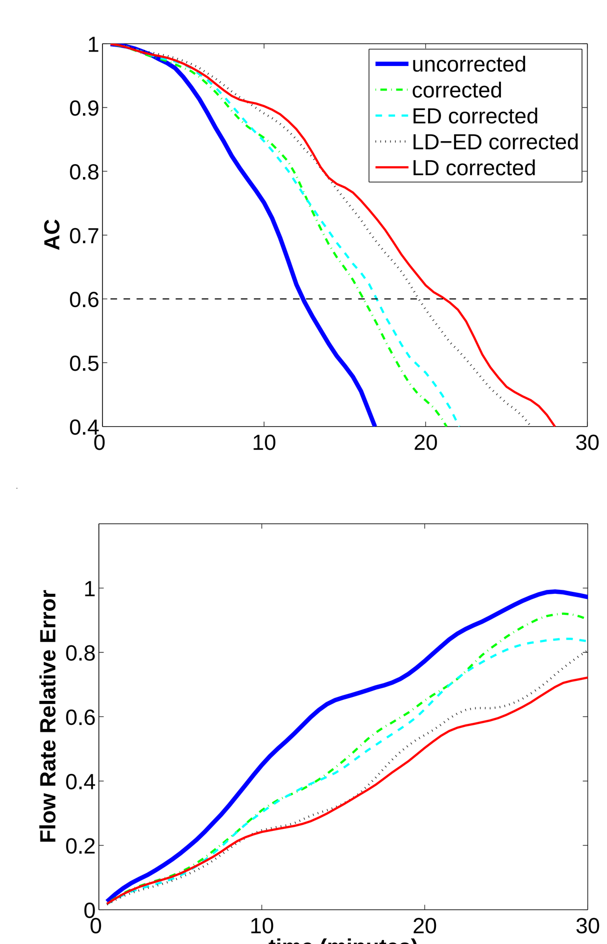
Anomaly correlation **(Top)** and flow-rate relative error **(Bottom)** over 1000 independent trials, for the uncorrected model (solid blue), and corrected models with parameters untuned (black) and optimally tuned (red). Corrected models show reduced error. Additionally, the small difference in performance between corrected models with tuned vs. untuned parameters suggests that empirical correction can overcome parameter inaccuracy.

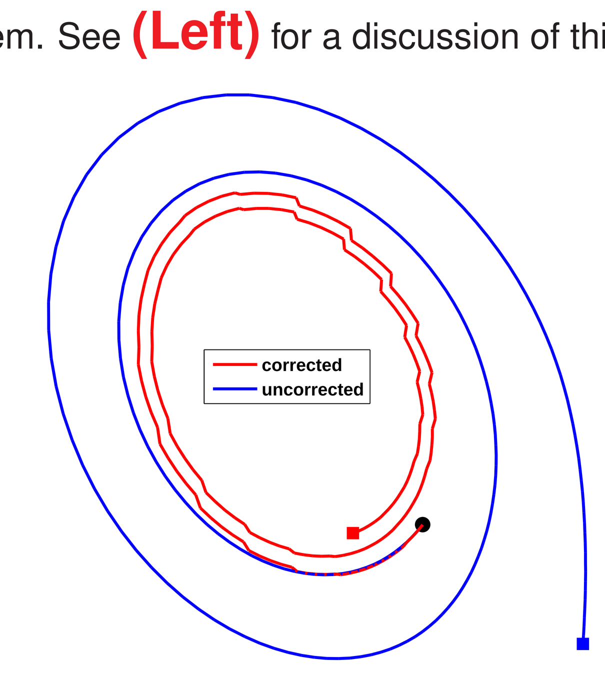
Incorporating System-Specific Dynamical Knowledge into the Correction Procedure



Improved error statistics hide some important dynamical consequences of empirical correction, including broken symmetry and altered stability of equilibrium solutions. **(Above)** Difference (in minutes) between predicted and actual time of first regime change (flow-reversal), taken $t_{\text{model}} - t_{\text{actual}}$, and plotted by initial state, for the (A) corrected, (B) lobe-dependent corrected, (C) equilibrium-dependent, and (D) lobe and equilibrium-dependent corrected models. The green dots show good predictions, while the red dots show initial states in the basins of the spuriously stable convective equilibria, (inset histograms show frequency of each). The lobe and equilibrium dependent corrected model is the best dynamical match to the truth.

The general empirical correction procedure can be tailored to take advantage of system-specific dynamical knowledge. The present system, for example, is known to have two distinct *regimes* characterized by opposite directions of flow in the tube, (corresponding to the left and right lobes in state-space). **(Right)** Computing a different bias term $\mathbf{b}$ and operator $\mathbf{L}$ for each regime, and applying them appropriately in a forecast scenario, results in further error reduction. Results are also shown for corrected models incorporating knowledge of distance from equilibrium (from the centers of the lobes), and models incorporating both types of dynamical knowledge. Another reason to adapt the correction procedure: **(Below)** A corrected model gone wrong. Empirical correction has altered the stability of the convective equilibria, which now attract nearby states in contrast with the true system. See **(Left)** for a discussion of this unfortunate consequence.





Anomaly correlation **(Top)** and flow-rate relative error **(Bottom)** over 5000 trials, for the uncorrected model (solid blue), and models corrected without dynamical knowlege, with knowlege of lobe (LD), distance from equilibrium (ED), and both (LD-ED). The best performing model incorporated flow direction (lobe) only (solid red), where forecasts are useful for about 80% longer than those made by the uncorrected model, more than doubling the improvement made by dynamically uninformed correction.

References and Acknowledgements

[1] K. D. Harris, E.-H. Ridouane, D. L. Hitt, and C. M. Danforth. Data assimilation applied to a chaotic toy climate. *In preparation*, 2011.
[2] E. Lorenz. Deterministic nonperiodic flow. *Journal of Atmospheric Sciences*, 20:130–141, 1963.
[3] C. E. Leith. Objective methods for weather prediction. *Annual Review of Fluid Mechanics*, 10:107–128, 1978.
[4] C. M. Danforth, E. Kalnay, and T. Miyoshi. Estimating and correcting global weather model error. *Monthly Weather Review*, 135(2):281–299, 2007.
[5] E. Kalnay. *Atmospheric Modelling, Data Assimilation and Predictability*. Cambridge University Press, 2003.

The authors would like to acknowledge support from NSF EPSCoR, NASA EPSCoR, the Vermont Space Grant Consortium, and an NSF grant for the Mathematics & Climate Research Network.

http://www.uvm.edu/~cdanfort
{nicholas.allgaier,kameron.harris,chris.danforth}@uvm.edu