Developing An Ensemble Kalman Filter for Data Assimilation in CTIPE

Stefan Codrescu
CU/CIRES Boulder
NOAA/NCEI Boulder
stefan.codrescu@noaa.gov

Mihail Codrescu
NOAA/SWPC Boulder

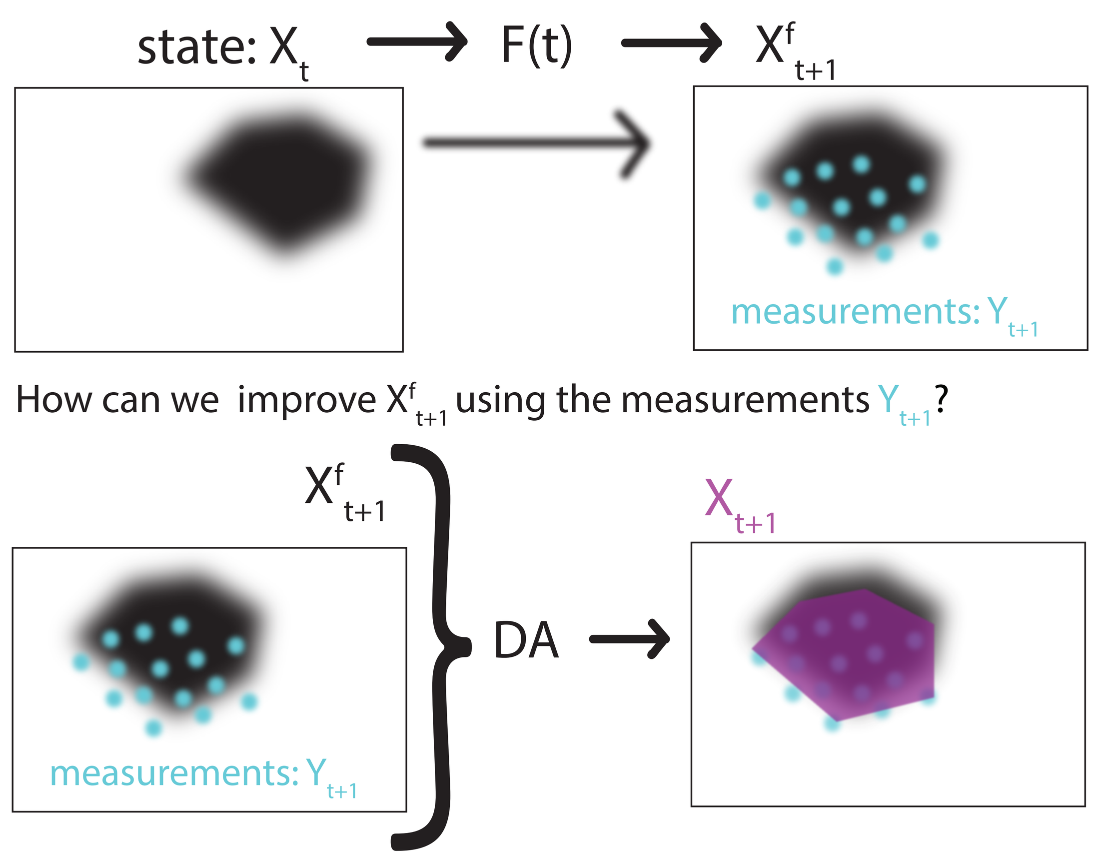
Mariangel Fedrizzi
CU/CIRES Boulder
SWPC/NOAA Boulder

Abstract

We are developing an ensemble kalman filter to assimilate data into the Coupled Thermosphere Ionosphere Plasmasphere and Electrodynamics (CTIPE) model. The Ensemble Kalman Filter (EnKF) approach is useful for approximating the non stationary covariance of the state, especially in high dimensional state spaces like that of CTIPE. Challenges include creating a representative distribution of the state uncertainty in the ensemble, limiting the computational complexity of the scheme, and managing the effects of measurement biases. Finally, we present preliminary results in assimilating simulated measurements.

Introduction

Data assimilation is a process in which measurements are incorporated into the state of a model to improve specification and forecasting. Our scheme implements an Ensemble Kalman Filter (EnKF).



Conceptually, a data assimilation (DA) scheme combines the model forecast X^f_{t+1} with measurements Y_{t+1} to estimate the state X_{t+1} with a smaller uncertainty.

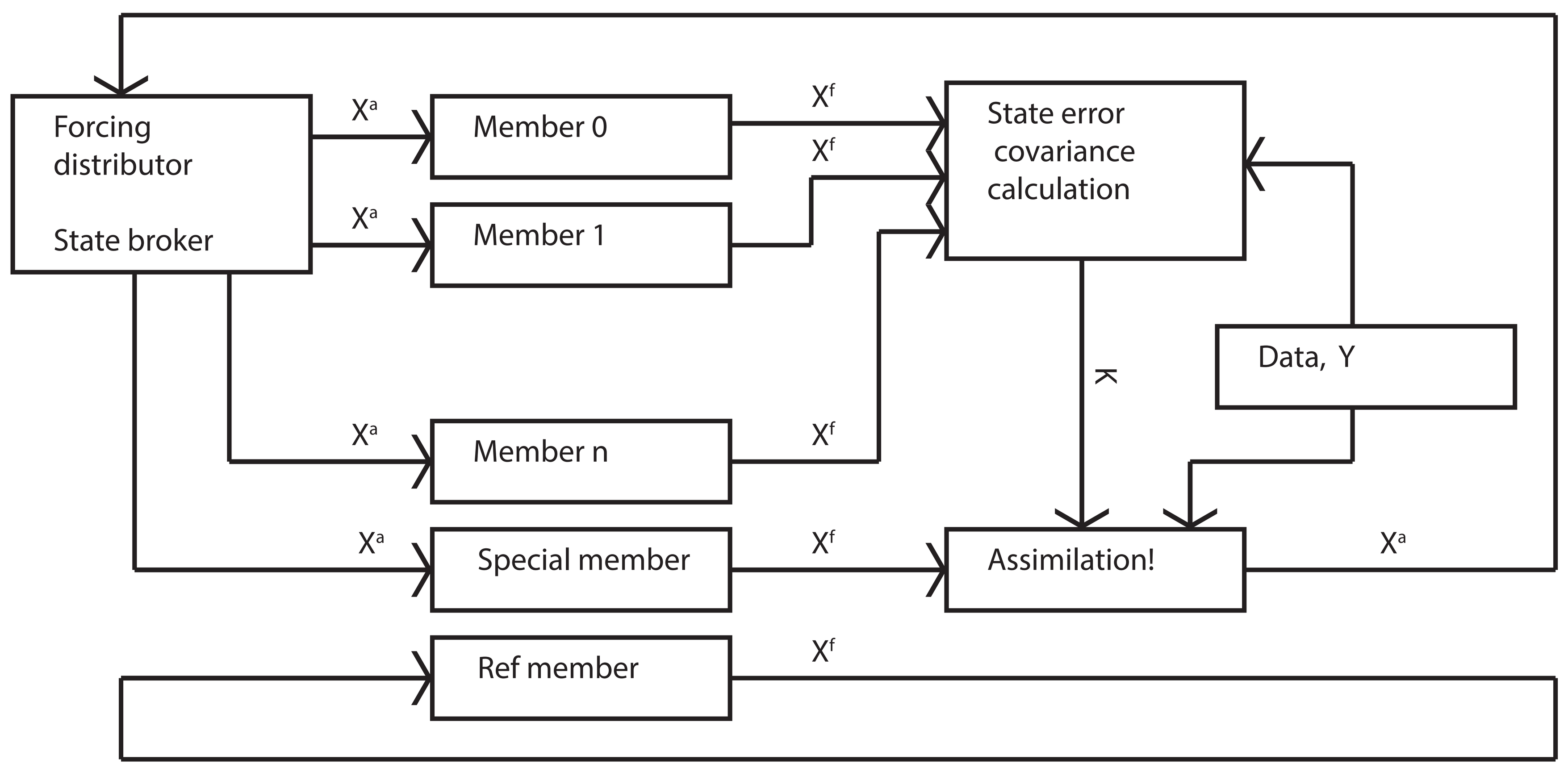
Here, we represent uncertainty roughly with blur.

Many techniques for data assimilation exist. Some techniques such as the extended Kalman Filter or 4DVAR linearize the dynamics which may be more epensive and underestimate the correlations compared to the Ensemble Kalman Filter.

Architecture

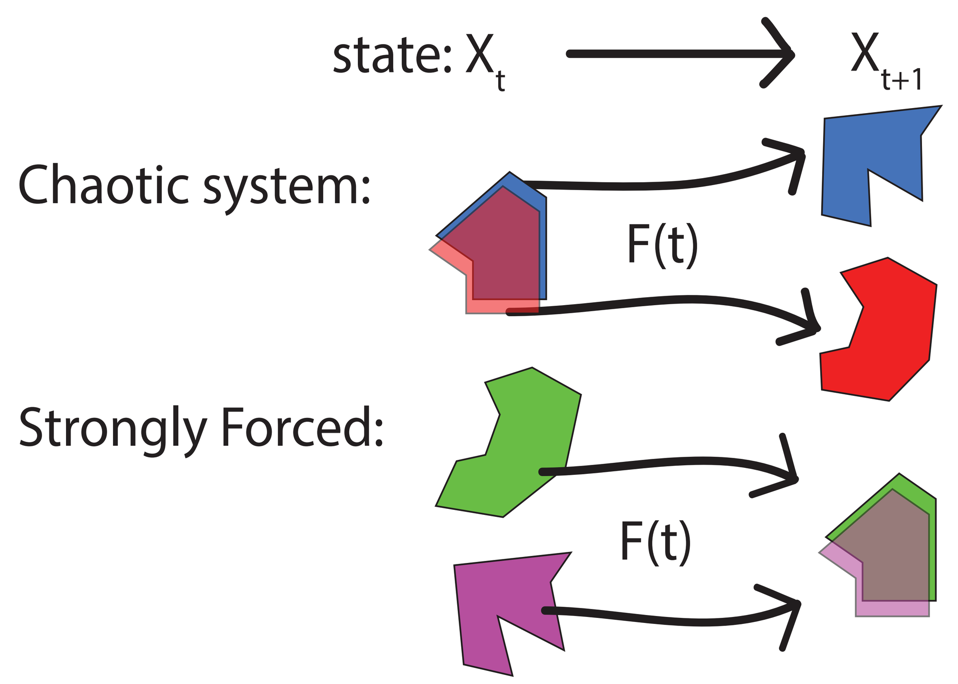
CTIPE is a non-linear, coupled thermosphere-ionosphere-plasmasphere physically based numerical code that includes a self-consistent electrodynamics scheme for the computation of dynamo electric fields. The model consists of four distinct components which run concurrently and are fully coupled.

The assimilation scheme is written from scratch in C++11. The program manages the ensemble, gathering measurements, and assimilation calculations. The Eigen library is used for matrix computations. Additionally, the GNU Multi-precision Library is needed.



Ensemble

Performance of the filter depends critically on the covariance estimate obtained from the ensemble. The goal is to distribute the ensemble over the state space according to the error of the current state and the possible evolutions in time through state space.



Note the difference in behavior between a strongly forced system and a chaotic system due to some forcing F(t).

Chaotic systems are characterized by a strong sensitivity to initial condition, thus it makes sense to perturb the initial condition across the ensemble to distribute members over the state space.

CTIPE in contrast, has less dependency on the initial condition and more dependency on the forcing applied. We call this a strongly forced system. To distribute the ensemble over the state space, we perturb the forcing.

The forcing in the ensemble is perturbed according to independent draws from normal distributions described to the right.

These are values that seemed to have acceptable performance. Investigating optimal distribution parameters is future work.

$$B_n \sim \mathcal{N}(0, 10)$$
$$B_\theta \sim \mathcal{N}(0, 15)$$
$$V_{sw} \sim \mathcal{N}(0, 50)$$
$$\rho_{sw} \sim \mathcal{N}(0, 5)$$
$$F_{10.7} \sim \mathcal{N}(0, 50)$$

Challenges

- The filter can cause the state to become inconsistent and cause model instance to crash.

- Density data from CHAMP and GRACE are more sparse than our simulated data and appears to have bias with respect to each other. Assimilating real density makes the filter diverge.

- How many members required to accurately approximate the covariance?

- How do we scale to more measurements? The most expensive calculation is the pseudo-inverse required for the Kalman gain. Can we avoid this by assimilating one measurement at a time?

Results

We show results from assimilating simulated neutral density measurements from March 20, 2007 (quiet day near solar minimum). Neutral density measurements are simulated by sampling a run of CTIPE forced with an artificial F10.7 value. Although the test is weak, it indicates that we are on the right track.

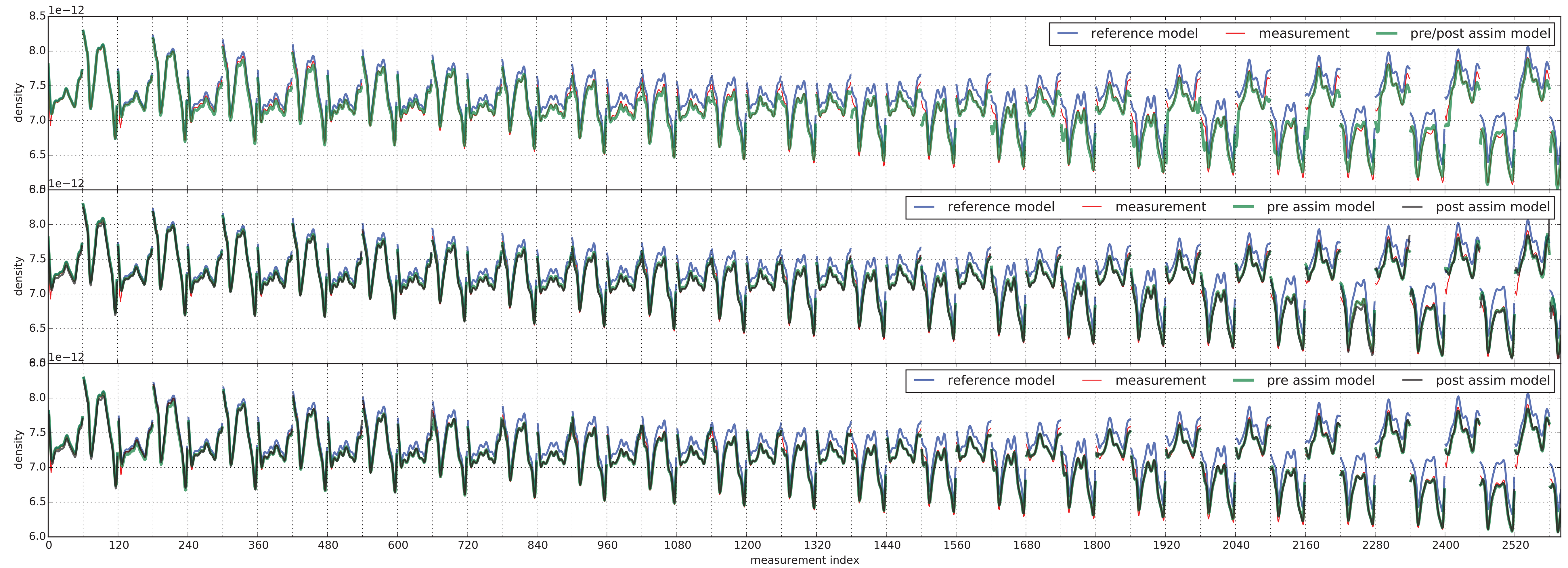
The combination of Panel 1 highlight the strongly forced nature of the system, the filter is able to reproduce the simulated measurements only by modifying the forcing Panel 2 illustrates that fixing the state under a certain forcing will maintain the trajectory. Panel 3 shows the behvavior when both forcing and state are modified by the assimilation.

Three modes

- Panel 1: only allowing the filter to modify the forcing.

- Panel 2: the filter is only allowed to modify the Kalman state elements.

- Panel 3: both state and forcing are modified by the filter.



Codrescu, M. V., T. J. Fuller-Rowell, and C. F. Minter. "An ensemble-type Kalman filter for neutral thermospheric composition during geomagnetic storms." Space Weather 2.11 (2004).

Evensen, G. (1994), Sequential data assimilation with a nonlinear quasigeostrophic model using Monte Carlo methods to forecast error statistics, J. Geophys. Res., 99, 10,143--10,162.

Gillijns, Steven, et al. "What is the ensemble Kalman filter and how well does it work?." American Control Conference, 2006. IEEE, 2006.

Houtekamer, P. L., and H. L. Mitchell (1998), Data assimilation using an ensemble Kalman filter technique, Mon. Weather Rev., 126, 796-- 811.

Kalman, R. E. (1960), A new approach to linear filtering and prediction problems, Trans. ASME J. Basic Eng., 82, 35--45.