A comparison between mixed and transform data assimilation schemes on short-, medium- and long-term forecasts

Erin A. Kashawlic[1,2], Steven J. Fletcher[3], John M. Forsythe[3], Andrew S. Jones[3] and Thomas H. Vonder Haar[3,4]
[1]Center for Multi-Scale Modeling of Atmospheric Processes, Colorado State University
[2]Department of Atmospheric, Oceanic and Space Sciences, University of Michigan
[3]Center for the Geosciences/Atmospheric Research, Cooperative Institute for Research in the Atmosphere, Colorado State University
[4]Department of Atmospheric Science, Colorado State University

Introduction

Operational forecast centers are starting to rely heavily upon data assimilation in order to produce more accurate forecasts that are based on current observations. As of present, the transform scheme is used most widely. It minimizes the cost function with respect to ln(**x**), as opposed to **x**, and then changes it back to the **x** space to complete the new forward run. It also takes the observations and considers them to be in the ln(**x**) space.

Fletcher and Zupanski (2006,2007) have developed the mixed DA system in which the minimization as well as the observations are kept in the **x** space. With this, there is no need to convert into a different space, thus retaining more information which is used to produce a, in theory, more accurate forecast.

We will compare the two schemes at varying lengths of forecasts to determine if one produces more correct forecasts than the other and if so, by what magnitude.

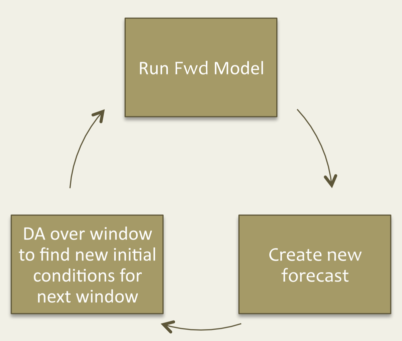
Objective and Methodology

This study was conducted to evaluate which data assimilation scheme is more accurate when producing short-, medium- and long-term forecasts.

- Run Fletcher's 4DVAR data assimilation code for MatLab, on Apple desktop during summer 2010 internship at CMMAP
  - Uses Lorenz '63 model – represents convection
    - Uses second-order Runge-Kutta scheme to solve ODEs
  - Z variable is lognormally distributed forecast while x,y are Gaussian distributed
  - Creates 'true' solution, the desired output
  - Randomly generates observations to use during assimilation of each window
  - Creates mixed and transform forecasts
  - Minimizes cost function, $J(\mathbf{x})$
  - Updates background error covariance matrix for next forecast
- Compare x,z variable differences
  - Y variable is not discussed – similar results to x
  - Each cycle of forecast has a short, medium and long forecast
  - Repeat with different observational error

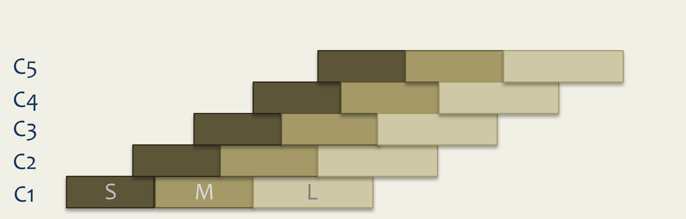
Lorenz '63 model equations:

$$\frac{dx}{dt} = -\sigma(x+y),$$
$$\frac{dy}{dt} = -xz + \rho x - y,$$
$$\frac{dz}{dt} = xy - \beta z,$$
$$\sigma = 10, \rho = 28, \beta = \frac{8}{3}$$



Chosen variables:
- Forecast length: 1500 timesteps
- Number of windows: 5
- Window length: 300 timesteps
- Number of observations per window: 20
- Observational errors (σ_o): 0.25, 1.5

- Short-term forecast: from 0-400 timesteps
- Medium-term forecast: from 400-900 timesteps
- Long-term forecast: from 900-1500 timesteps

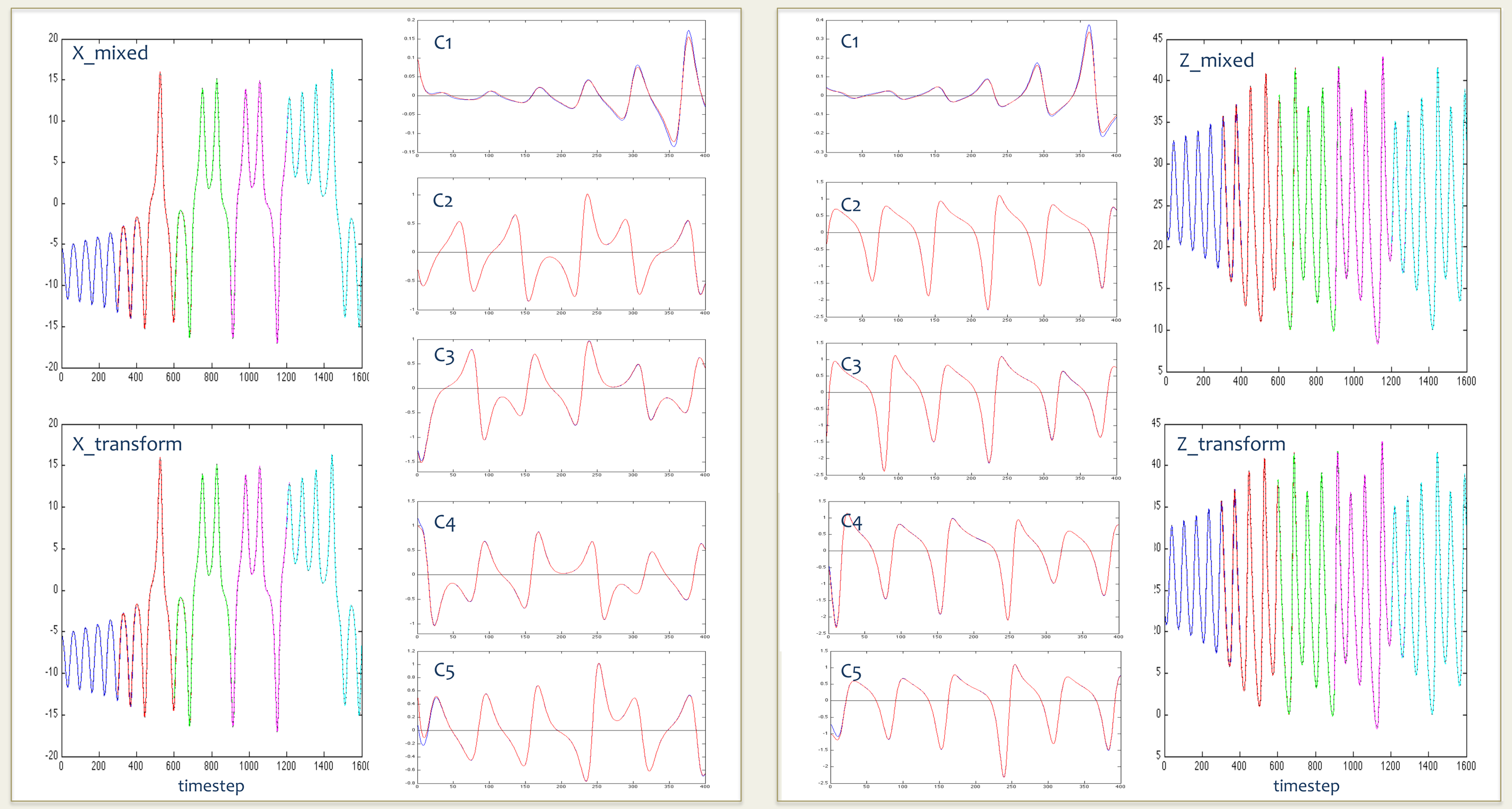


Results

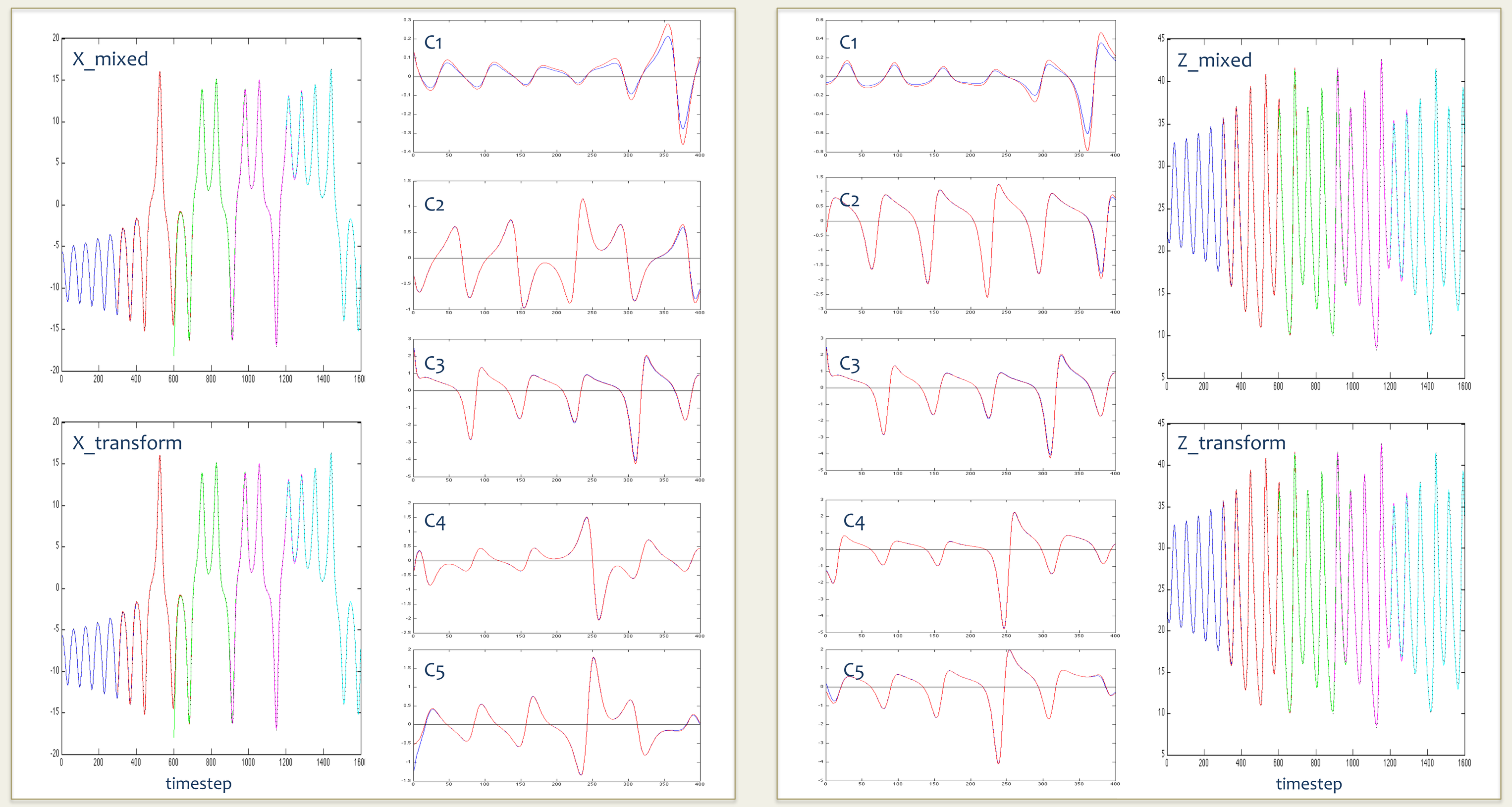
σ_o=0.25

σ_o=1.5

Short-term



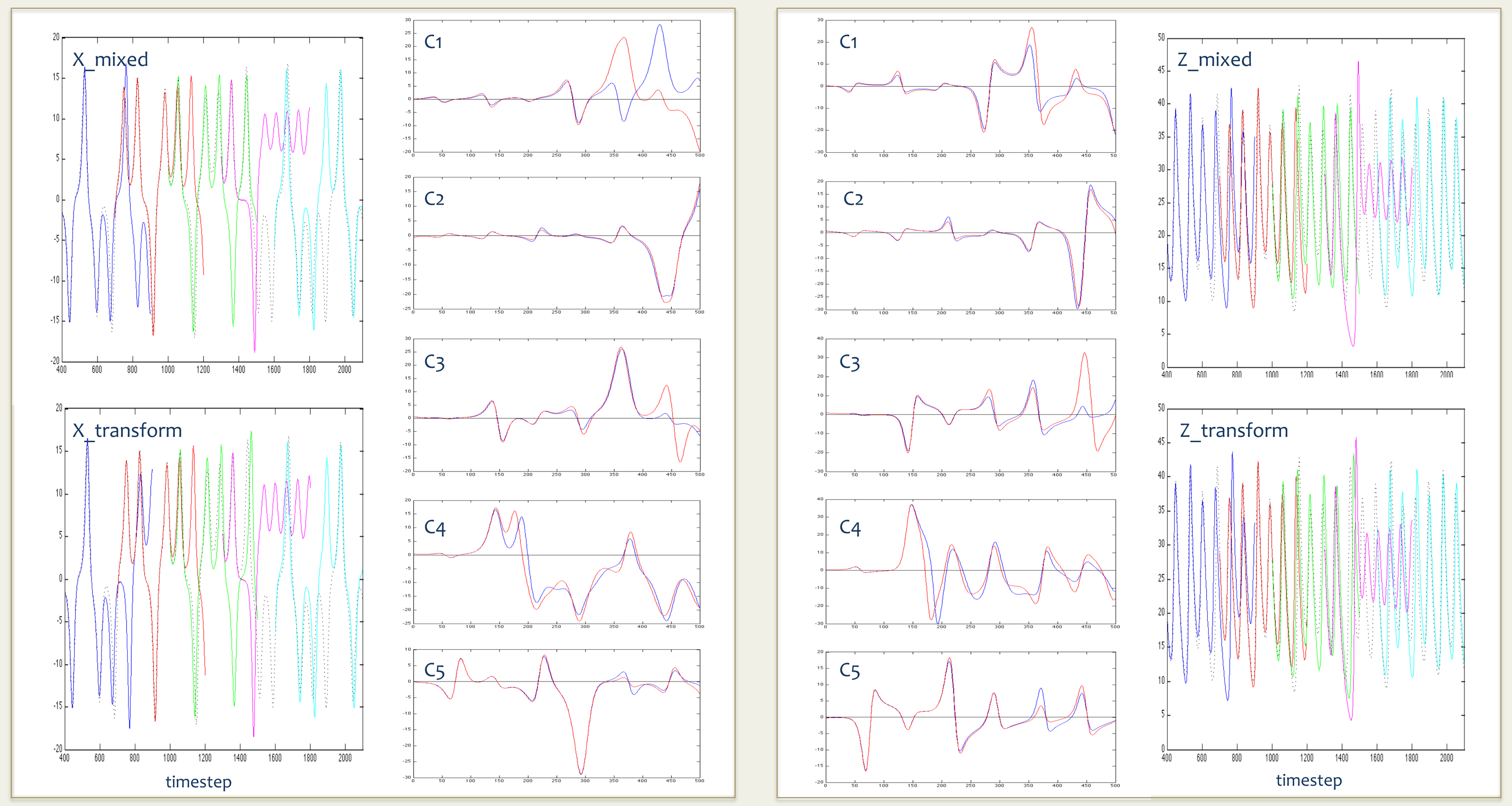
Full assimilation forecast plots: cycle 1-blue, cycle 2-red, cycle 3-green, cycle 4-magenta, cycle 5-cyan, true-black dashed; Difference plots ($\mathbf{x}_{true} - \mathbf{x}_{forecast}$): mixed scheme-blue, transform scheme-red



Full assimilation forecast plots: cycle 1-blue, cycle 2-red, cycle 3-green, cycle 4-magenta, cycle 5-cyan, true-black dashed; Difference plots ($\mathbf{x}_{true} - \mathbf{x}_{forecast}$): mixed scheme-blue, transform scheme-red

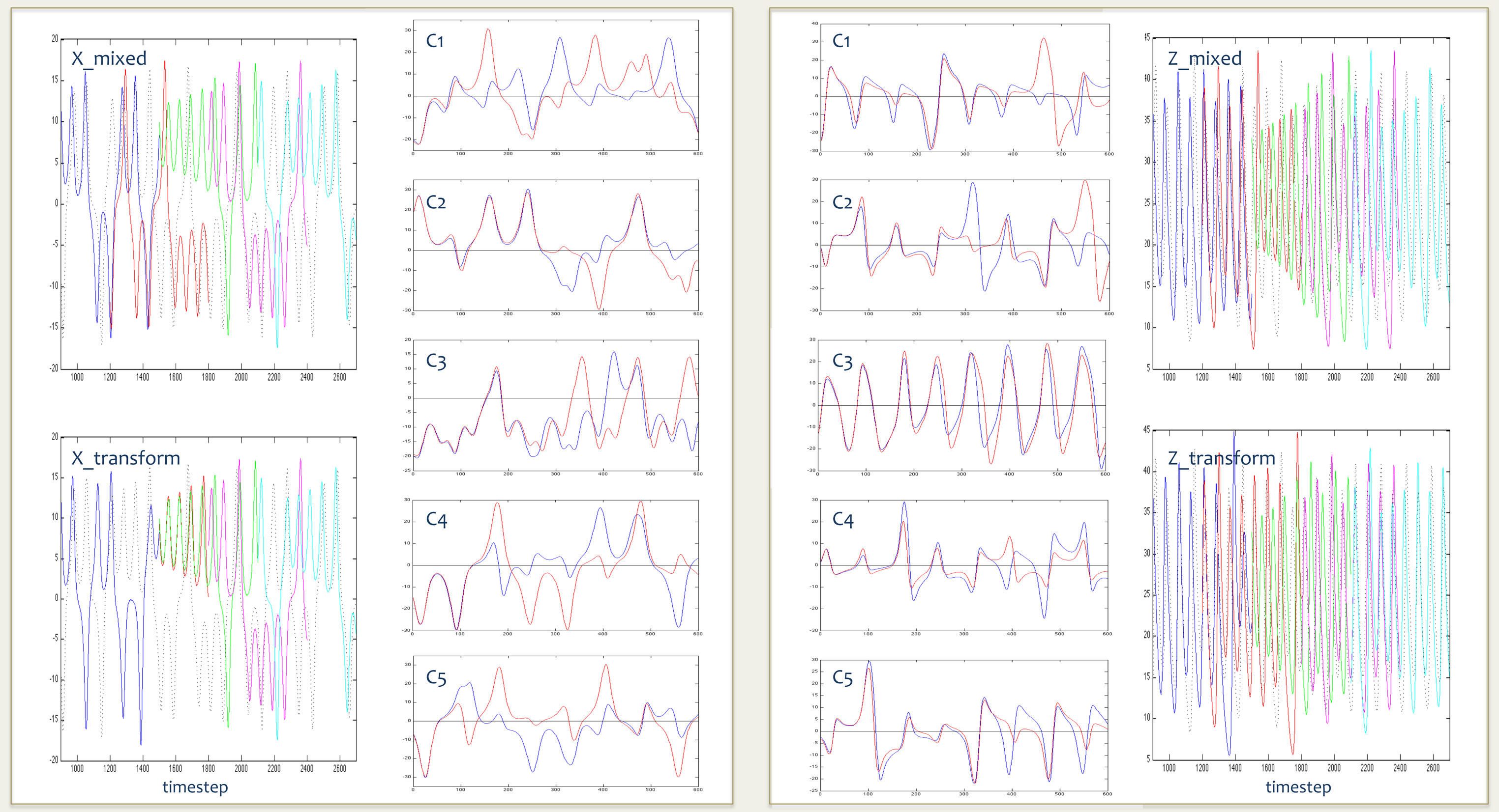
Medium-term

Full assimilation forecast plots: cycle 1-blue, cycle 2-red, cycle 3-green, cycle 4-magenta, cycle 5-cyan, true-black dashed; Difference plots ($\mathbf{x}_{true} - \mathbf{x}_{forecast}$): mixed scheme-blue, transform scheme-red

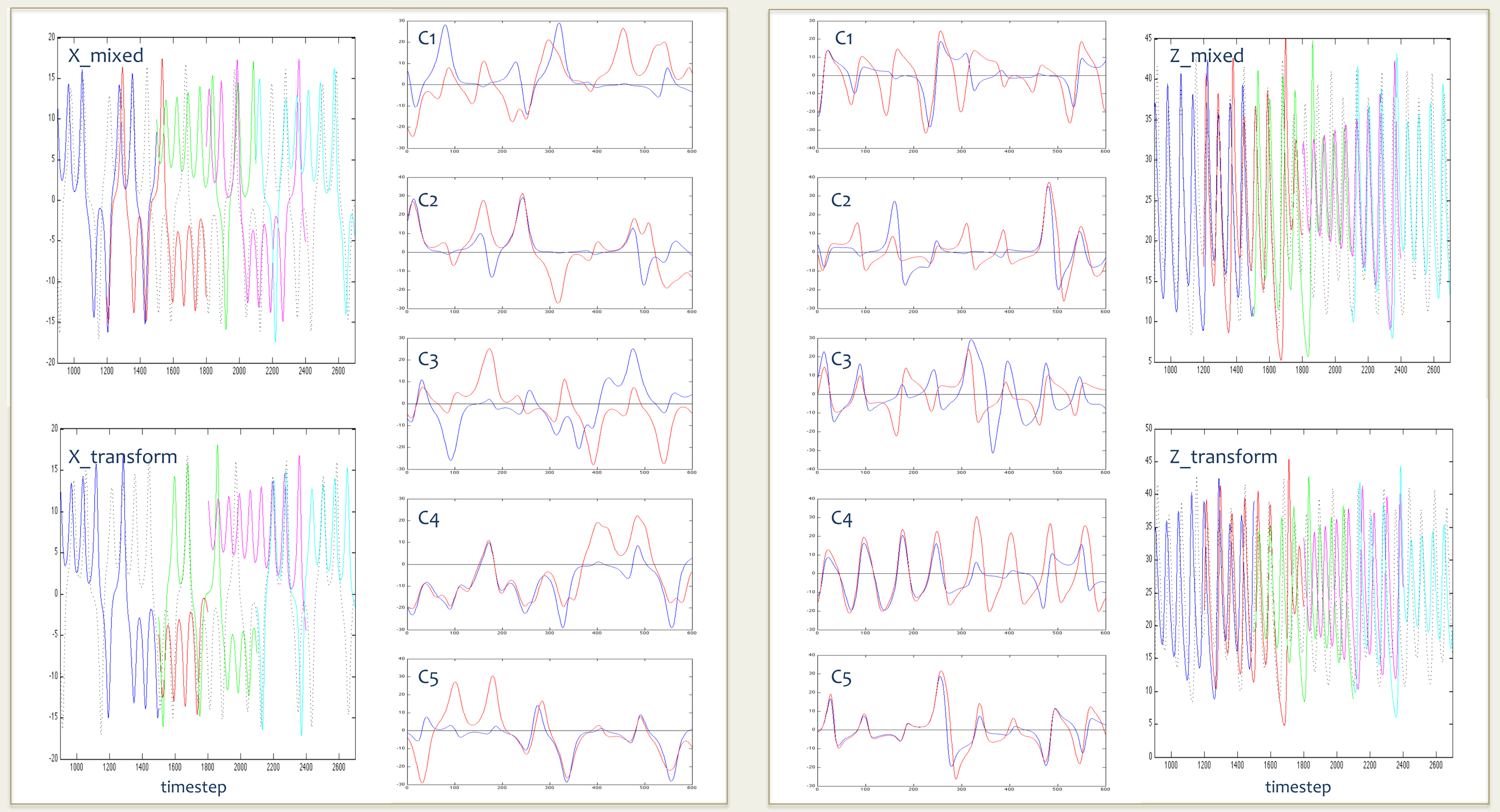


Full assimilation forecast plots: cycle 1-blue, cycle 2-red, cycle 3-green, cycle 4-magenta, cycle 5-cyan, true-black dashed; Difference plots ($\mathbf{x}_{true} - \mathbf{x}_{forecast}$): mixed scheme-blue, transform scheme-red

Long-term



Full assimilation forecast plots: cycle 1-blue, cycle 2-red, cycle 3-green, cycle 4-magenta, cycle 5-cyan, true-black dashed; Difference plots ($\mathbf{x}_{true} - \mathbf{x}_{forecast}$): mixed scheme-blue, transform scheme-red



Full assimilation forecast plots: cycle 1-blue, cycle 2-red, cycle 3-green, cycle 4-magenta, cycle 5-cyan, true-black dashed; Difference plots ($\mathbf{x}_{true} - \mathbf{x}_{forecast}$): mixed scheme-blue, transform scheme-red

Conclusions

Short
- Either scheme shows similar results
- Better initial conditions with mixed with small error obs
- Amplitude not as large with mixed with large error obs

Medium
- Either scheme shows similar results
- Mostly, mixed shows smaller peak amplitude of ~1-3
- As cycles progress, errors decrease by ~60% during peaks in general

Long
- Either scheme is very chaotic
- Mixed does better in X, transform does better in Z with small obs
- Mixed does better in Z with large obs

- For less accurate observations, trends show mixed scheme is a better
- As more cycles progress, mixed becomes more accurate
- Changing attractors in X prove hard to get back to forecast in medium- and long-term

Future Work

Plans for further work include rerunning this model with each assimilation scheme but with different parameters. We would like to see if fewer, but more accurate, observations can produce more correct forecasts rather than incorporating more, but less accurate, observations. Also, we would like to extend as well as shorten the window lengths and see how changing the window lengths affect the projected forecast.

Acknowledgments

Ms. Kashawlic would like to extend gratitude towards Eric Guillot for being her graduate mentor during her internship.

Dr. Fletcher and Mr. Forsythe were funded by the Department of Defence (DoD) Center for Geosciences/ Atmospheric Research at Colorado State University under Cooperative Agreement (W911NF-060200015) with the Army Research Laboratory.

This work has been supported by the National Science Foundation Science and Technology Center for Multi-Scale Modeling of Atmospheric Processes, managed by Colorado State University under Cooperative Agreement No. ATM-0425247.

References

Fletcher, S.J. and Zupanski, M 2006a. A data assimilation method for log-normally distributed observational errors. *Quart. J. Roy. Meteor. Soc.* **132**, 2505-2519.
Fletcher, S.J. and Zupanski, M 2006b. A hybrid normal and lognormal distribution for data assimilation. *Atmos. Sci. Lett.* **7**, 43-46.
Fletcher, S.J. and Zupanski, M 2007. Implications and impacts of transforming lognormal variables into normal variables in VAR. *Meteorologische Zeitschrift* **16**, 755-765.
Fletcher, S.J. 2010. Mixed Gaussian-lognormal four-dimensional data assimilation. *Tellus* **62A**, 266-287.
Kalnay, E. 2003. *Atmospheric Modeling, Data Assimilation and Predictability*. Cambridge University Press, Cambridge.
Lorenc, A.C. 1986. Analysis methods for numerical weather prediction. *Quart. J. Roy. Meteor. Soc.* **112**, 1177-1194.
Lorenz, E.N. 1963. Deterministic nonperiodic flow. *J. Atmos. Sci.* **20**, 130-141.

Contact information: Erin A. Kashawlic, eakashaw@umich.edu