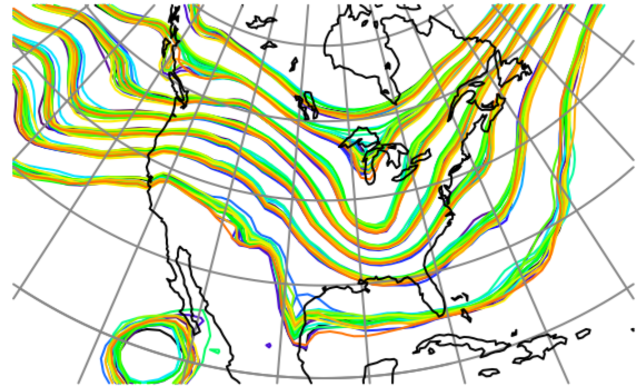


Ensemble Data Assimilation for Observations with Spatially and Temporally Correlated Errors

Jeffrey Anderson, NCAR Data Assimilation Research Section

CAHMDA VII, 21 Aug. 2017

Outline

Dealing with correlated observation error in ensemble filters.

1. Idealized correlated error.

2. Difference observations.

3. Explictly modeling instrument error.

4. Comparing the two methods.

5. Conclusions and recommendations.

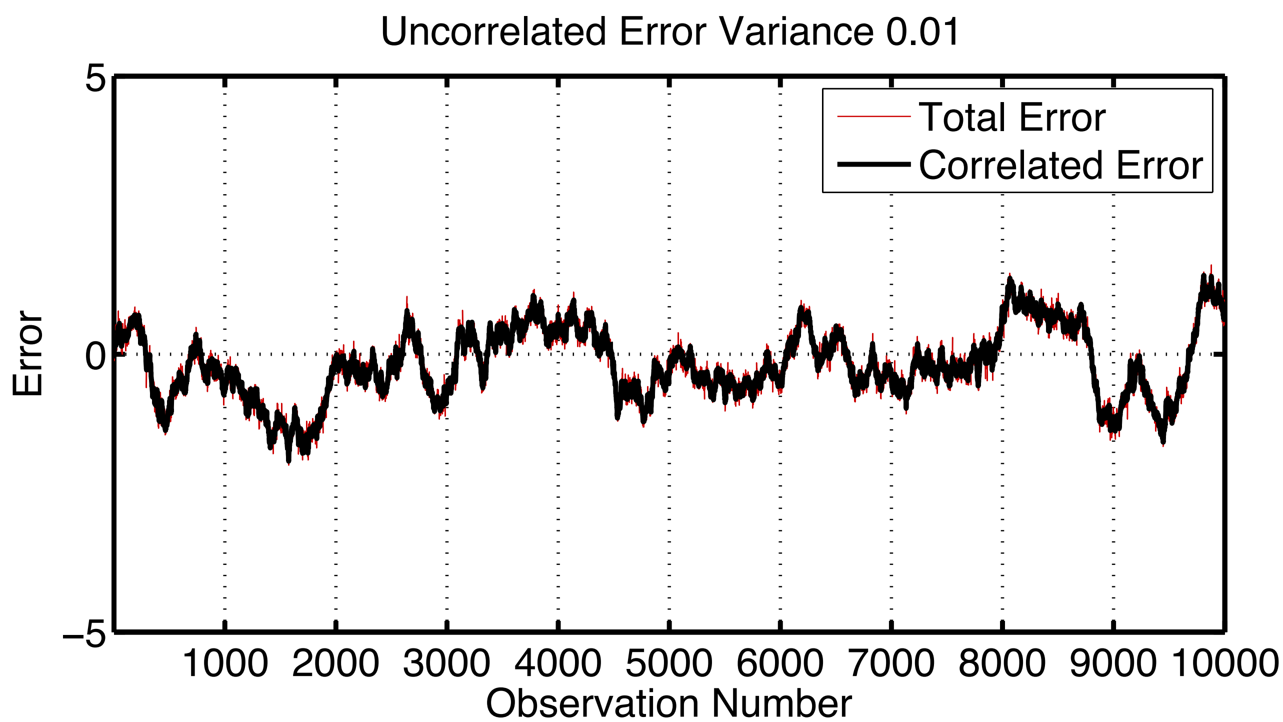
Most Observations Have Correlated Obs. Errors

Examples:

- Satellite radiances: instrument bias and aging.
- In situ soil moisture: instrument plus siting representativeness.
- Rainfall: gauge deficiencies plus siting.

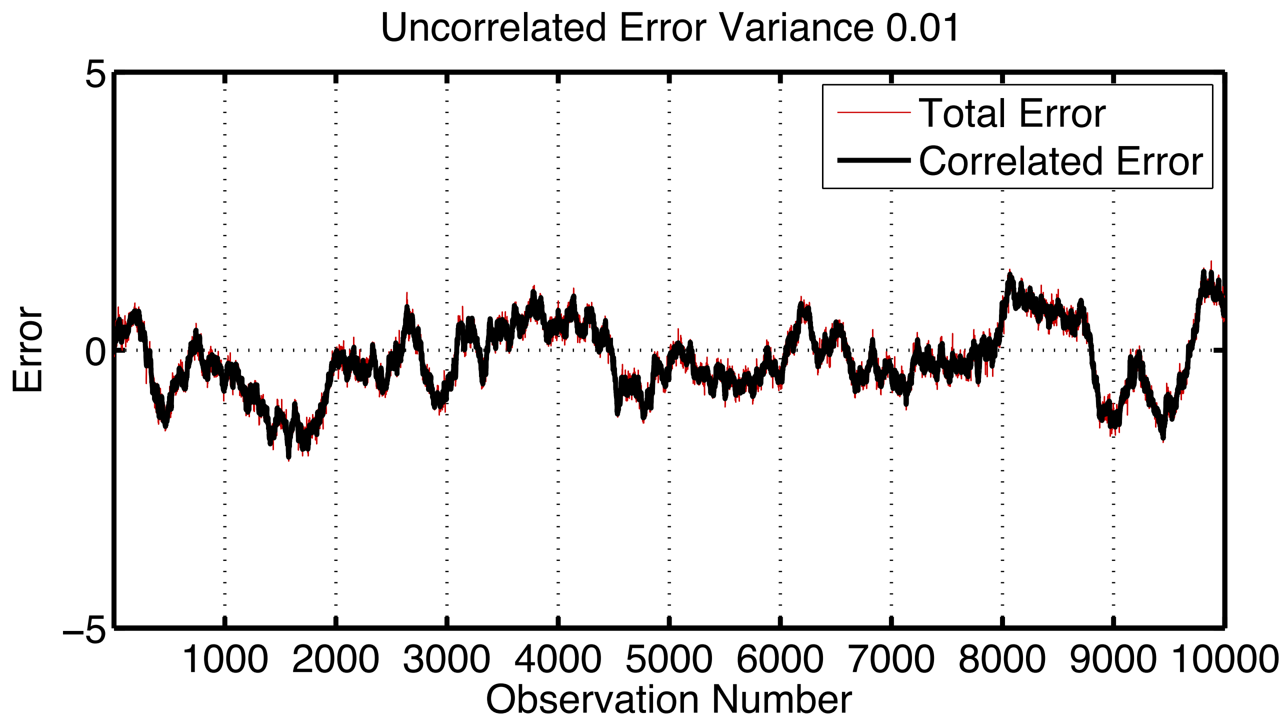
Observation Error Time Series

Example: Correlated Error AR1 with Variance 1.
Single Step Cov 0.999. Fixed for all cases.



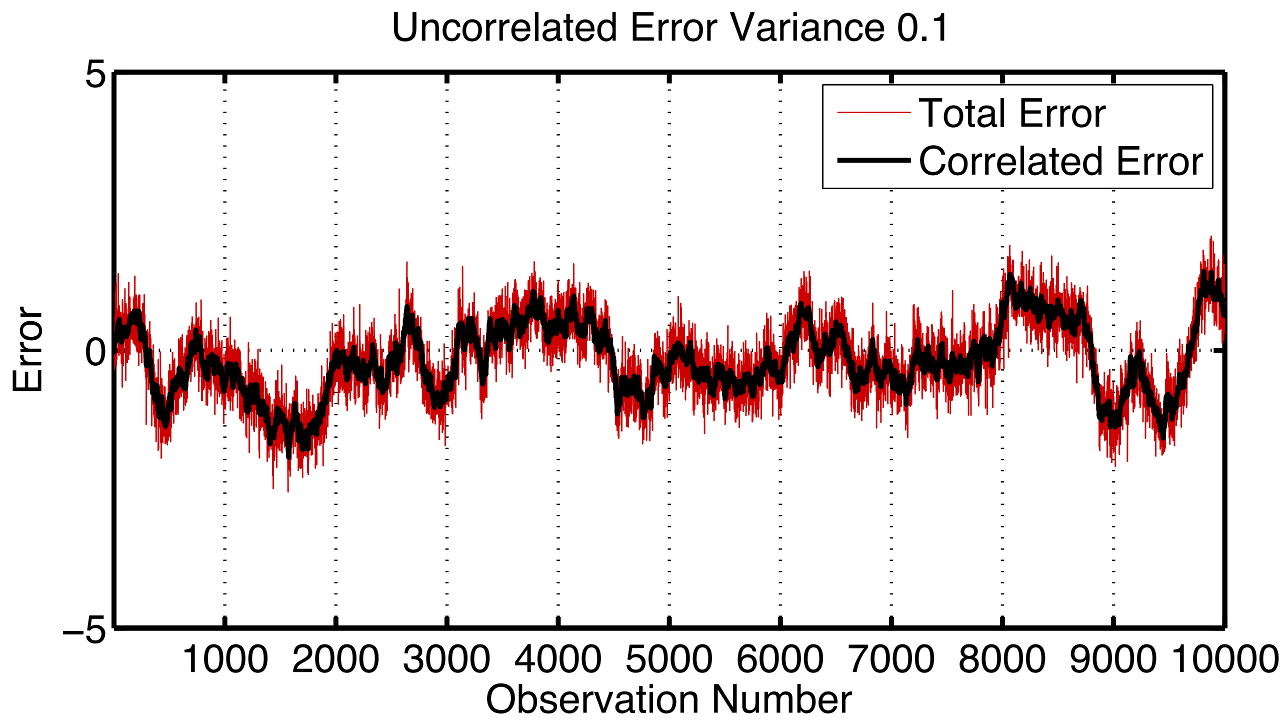
Observation Error Time Series

Example: Correlated Error AR1 with Variance 1.
Single Step Cov 0.999. Fixed for all cases.
Vary uncorrelated error variance, 0.01



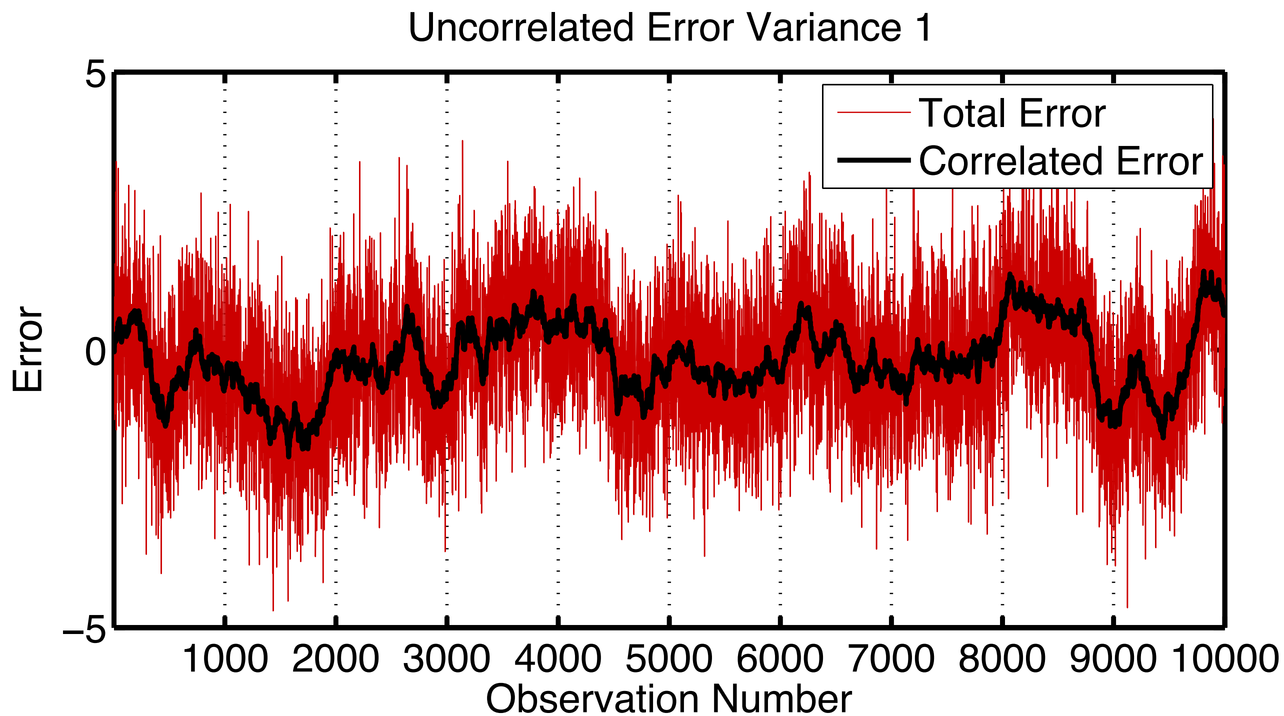
Observation Error Time Series

Example: Correlated Error AR1 with Variance 1.
Single Step Cov 0.999. Fixed for all cases.
Vary uncorrelated error variance, 0.1



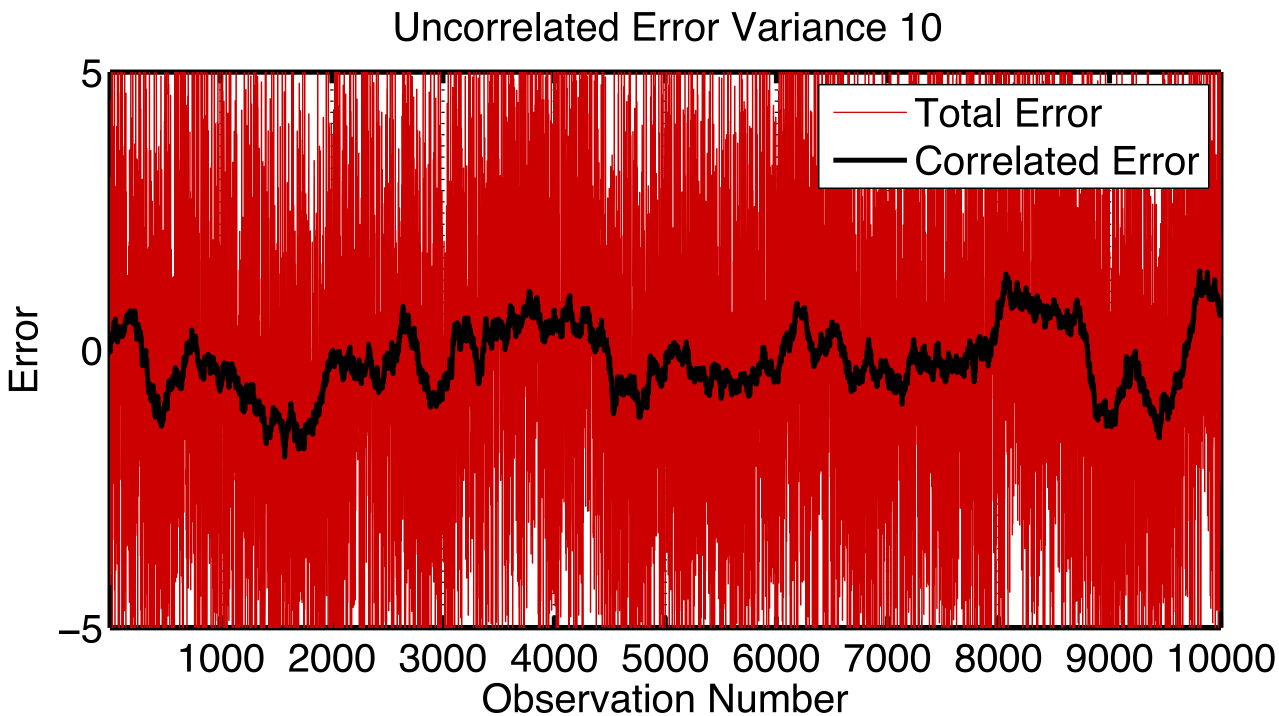
Observation Error Time Series

Example: Correlated Error AR1 with Variance 1.
Single Step Cov 0.999. Fixed for all cases.
Vary uncorrelated error variance, 1.0



Observation Error Time Series

Example: Correlated Error AR1 with Variance 1.
Single Step Cov 0.999. Fixed for all cases.
Vary uncorrelated error variance, 10.0



Possible approaches to dealing with correlated obs error

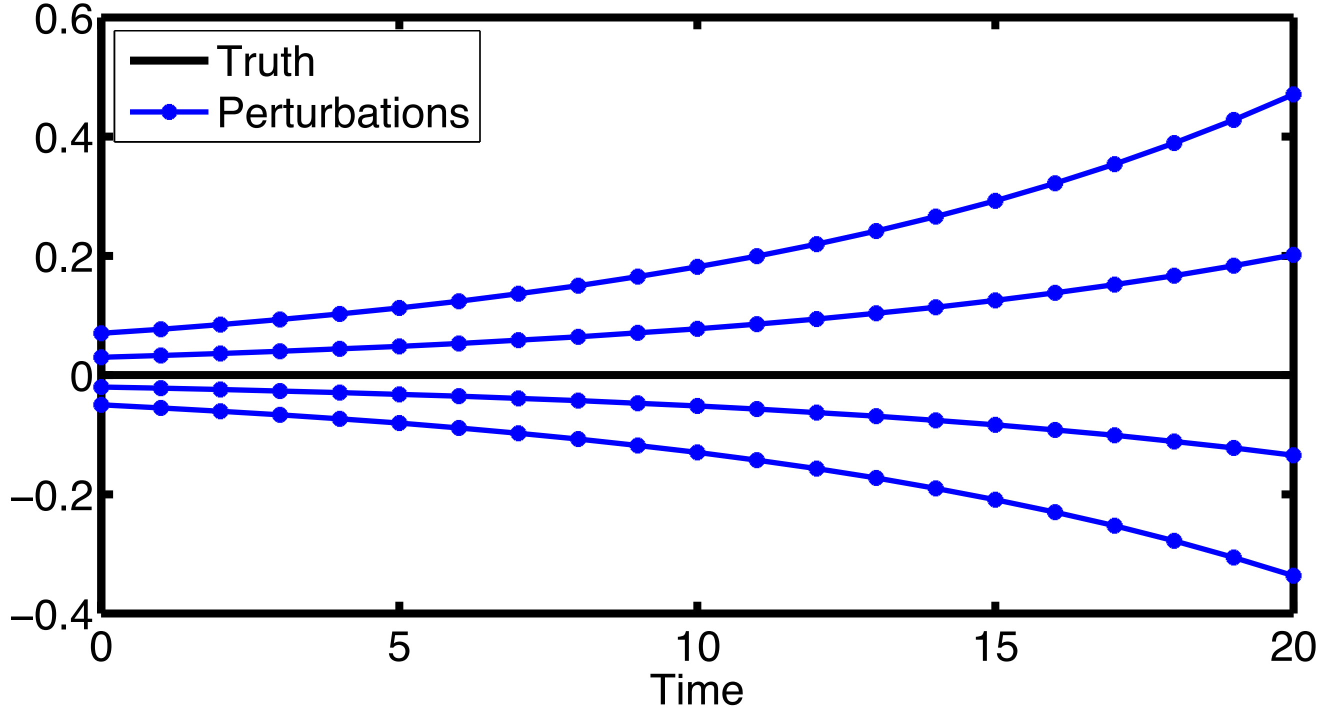
- Ignore it (common),
- Add parameters to forward operator, estimate them,
- Model it explicitly (various ways),
- Time difference observations.

1D Linear Exponential Growth Model

True trajectory is always 0.

Evolution is $x_{t+1} = 1.1 x_t$

Perturbations grow exponentially in time.



Outline

Dealing with correlated observation error in ensemble filters.

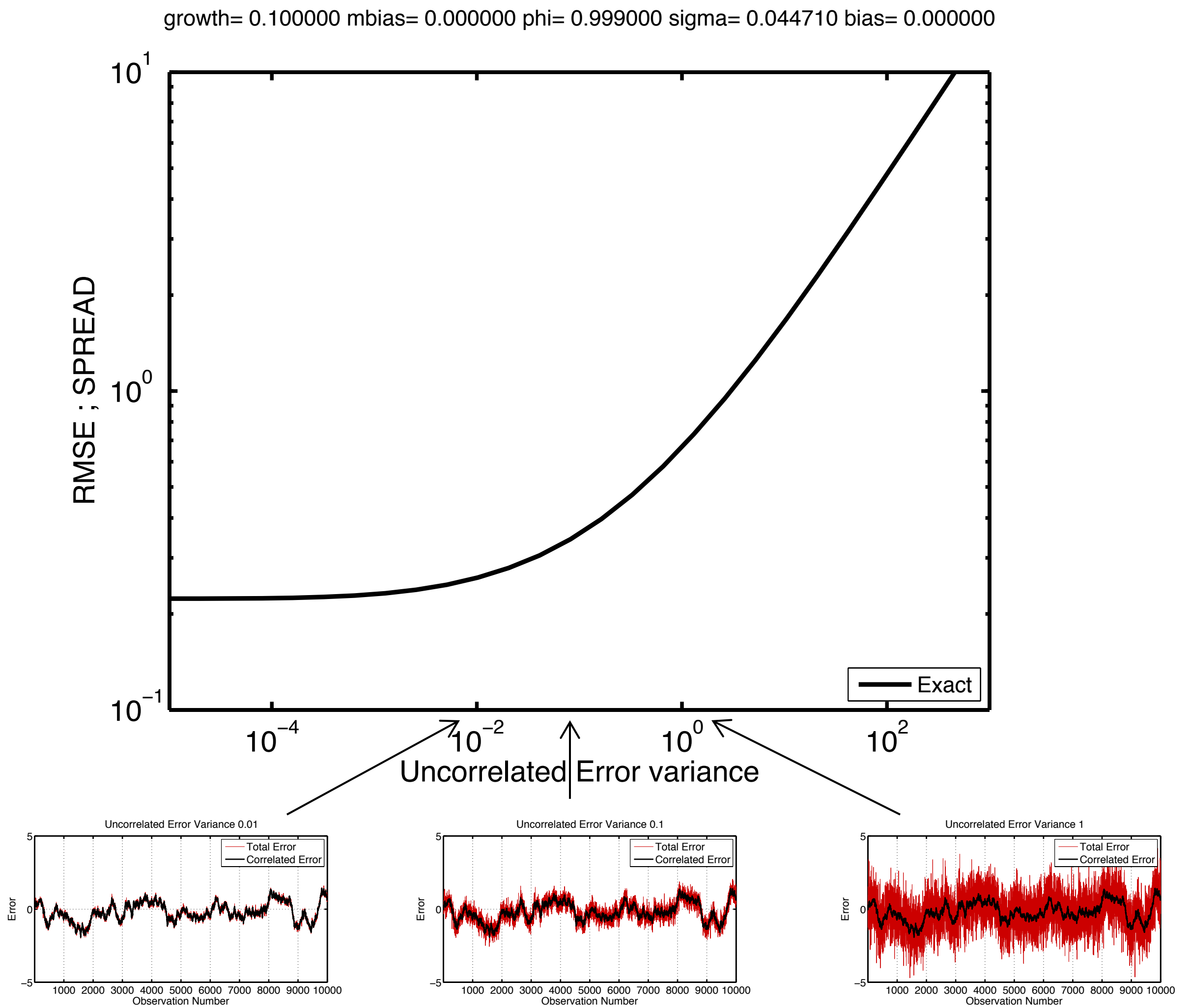
1. Idealized correlated error.
2. Difference observations.
3. Explicitly modeling instrument error.
4. Comparing the two methods.
5. Conclusions and recommendations.

Assimilating Correlated Observations

Obs1	Obs2	Obs3	Obs4	Obs5	Obs6

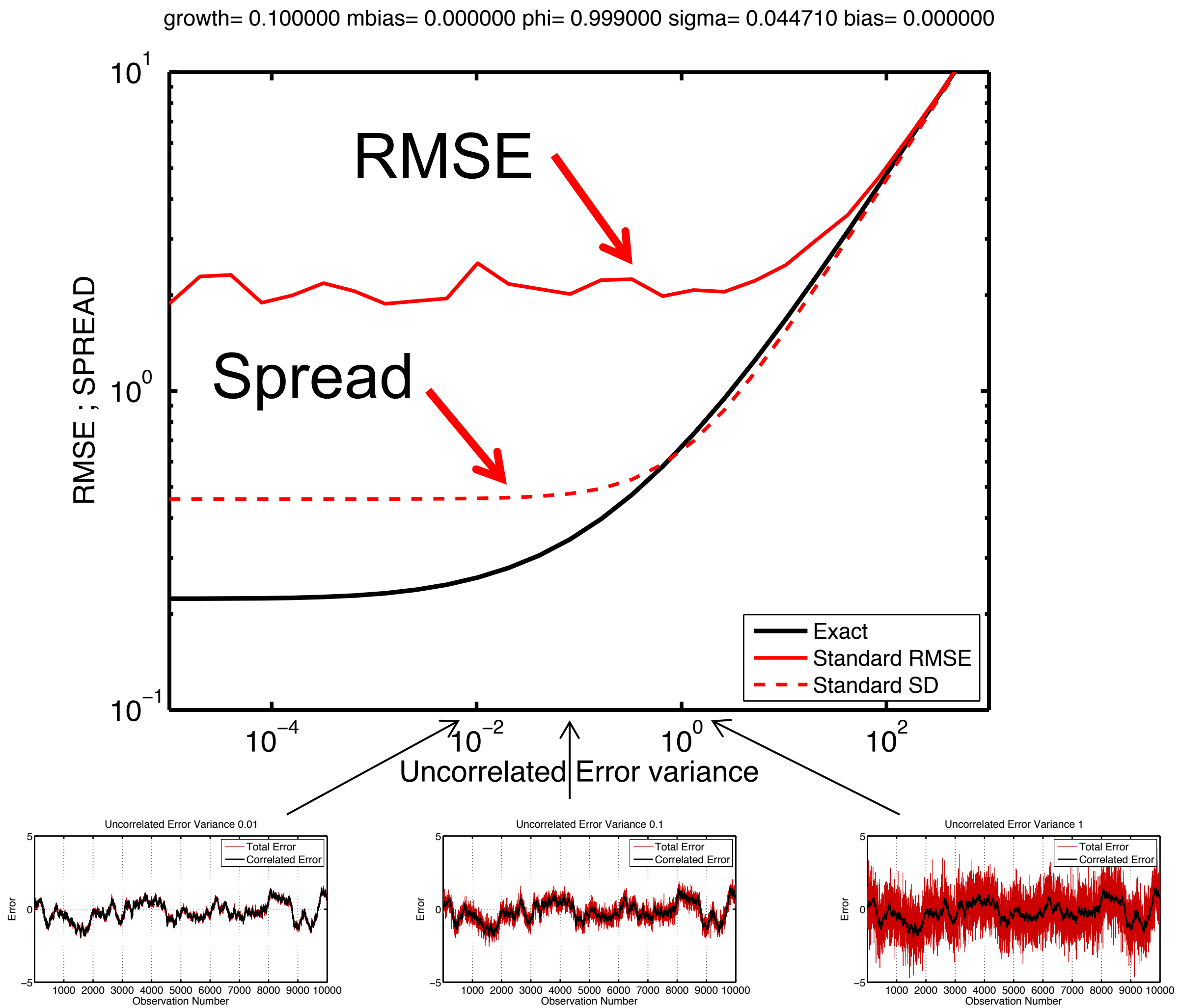
1D Exponential Growth Model Results

Exact Smoother Result. Can't do better than this.



1D Exponential Growth Model Results

EAKF Poor Unless Uncorrelated Error Dominates



CAHMDA VII, 21 Aug. 2017

Two Types of Difference Observations

Obs1	Obs2	Obs3	Obs4	Obs5	Obs6

Unlinked Difference Observations

Unlinked Diff 1

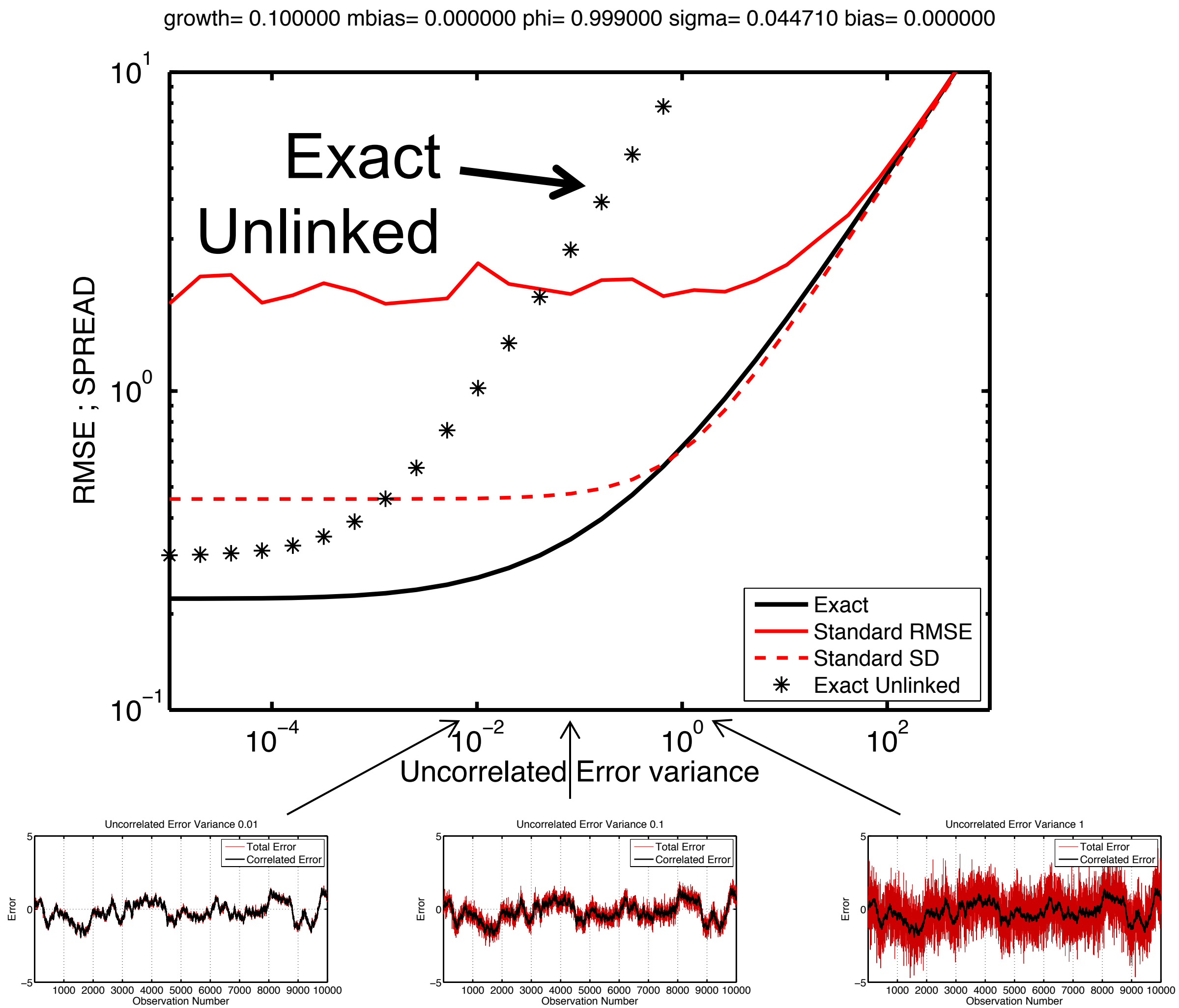
Unlinked Diff 3

Unlinked Diff 5

Obs1 Obs2 Obs3 Obs4 Obs5 Obs6

1D Exponential Growth Model Results

Exact Unlinked Difference Obs Much worse.



CAHMDA VII, 21 Aug. 2017

Uninked Difference Observations

$$Var \approx 2\sigma_u^2$$

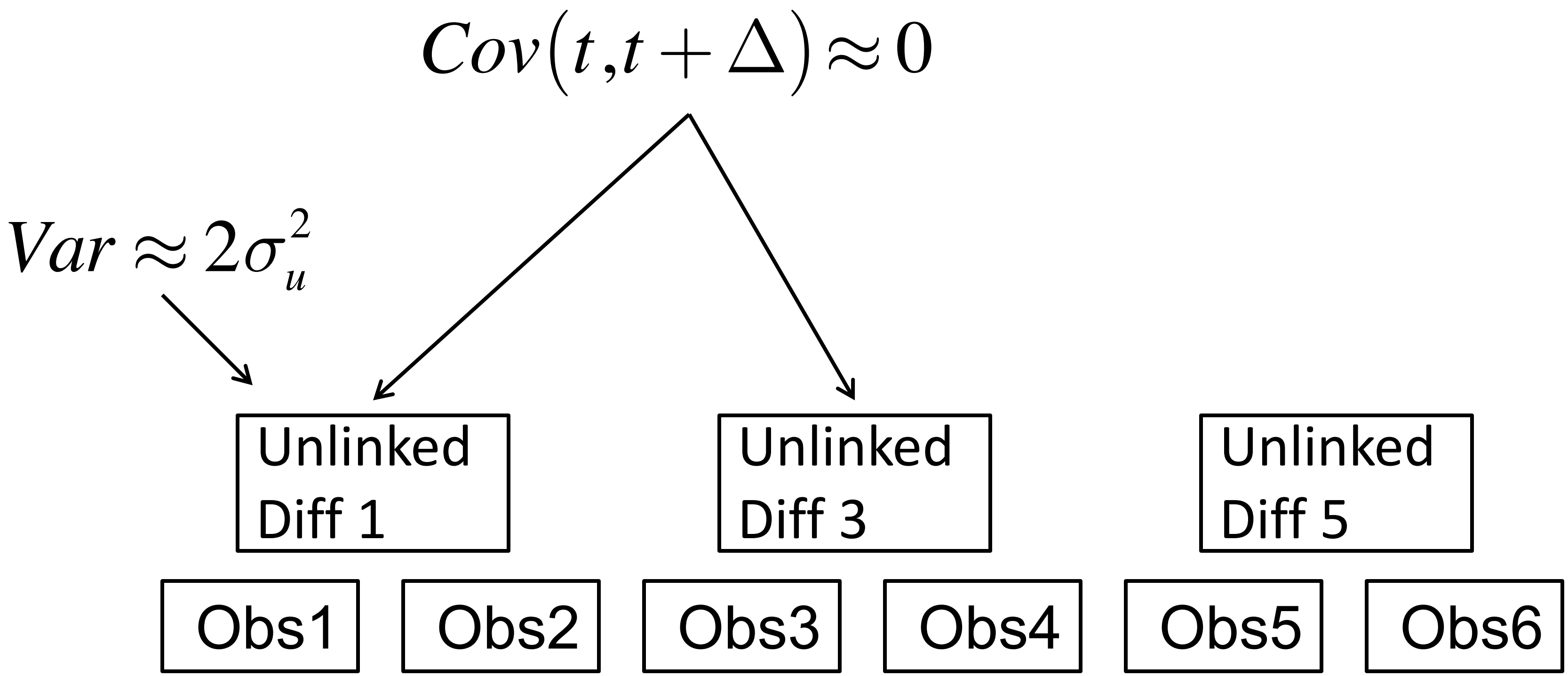
Unlinked Diff 1

Unlinked Diff 3

Unlinked Diff 5

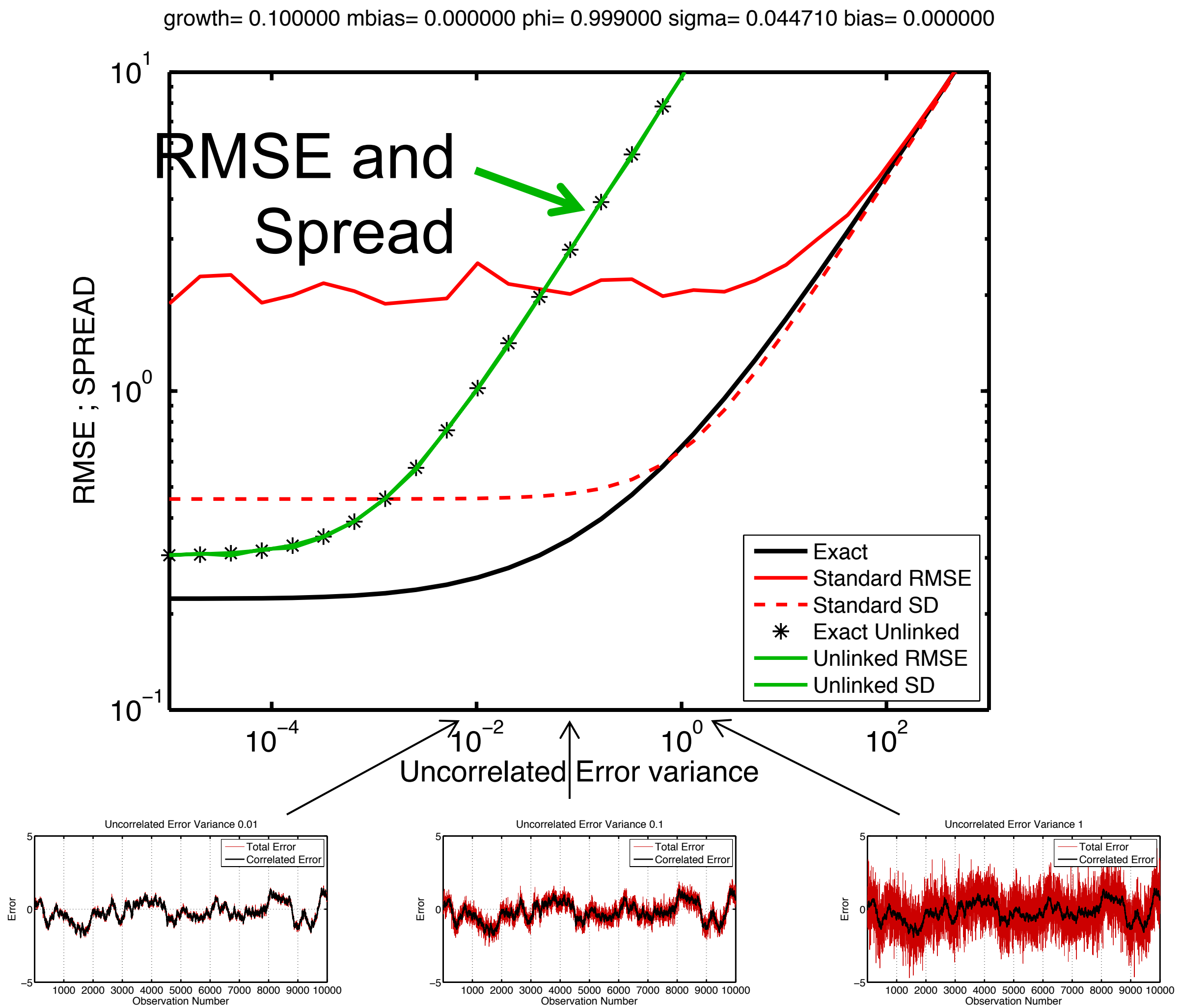
Obs1 Obs2 Obs3 Obs4 Obs5 Obs6

Uninked Difference Observations



1D Exponential Growth Model Results

EAKF is nearly exact for Unlinked Difference Obs.



Linked Difference Observations

Linked Diff 1

Linked Diff 3

Linked Diff 5

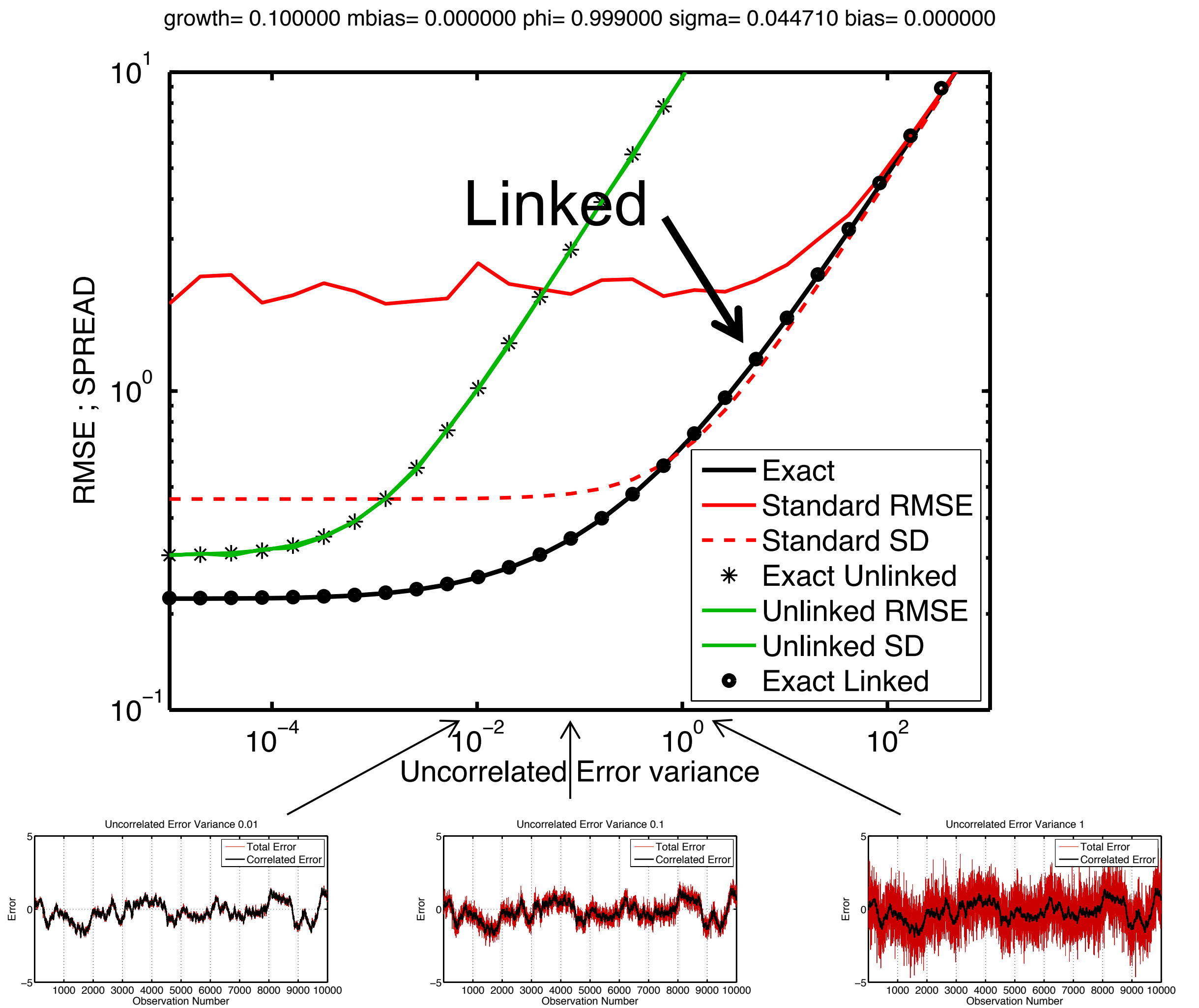
Obs1 Obs2 Obs3 Obs4 Obs5 Obs6

Linked Diff 2

Linked Diff 4

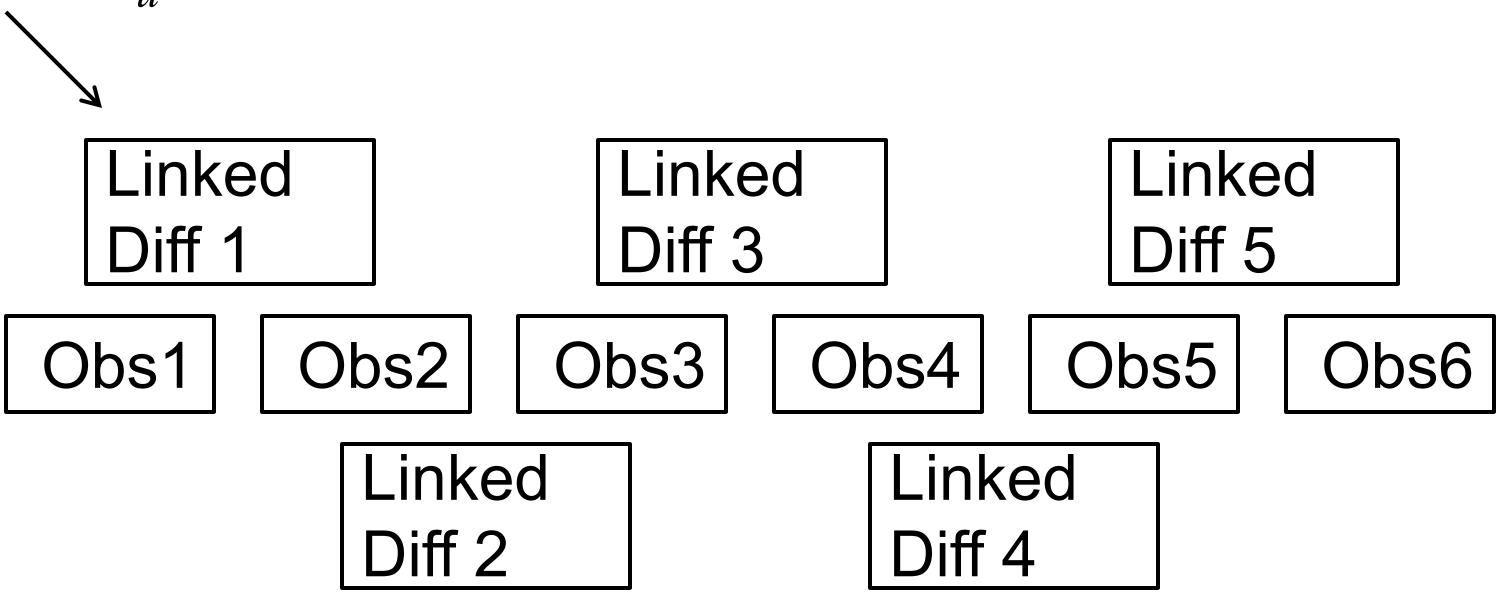
1D Exponential Growth Model Results

Exact linked Difference Obs Nearly Identical to Analytic.



Linked Difference Observations

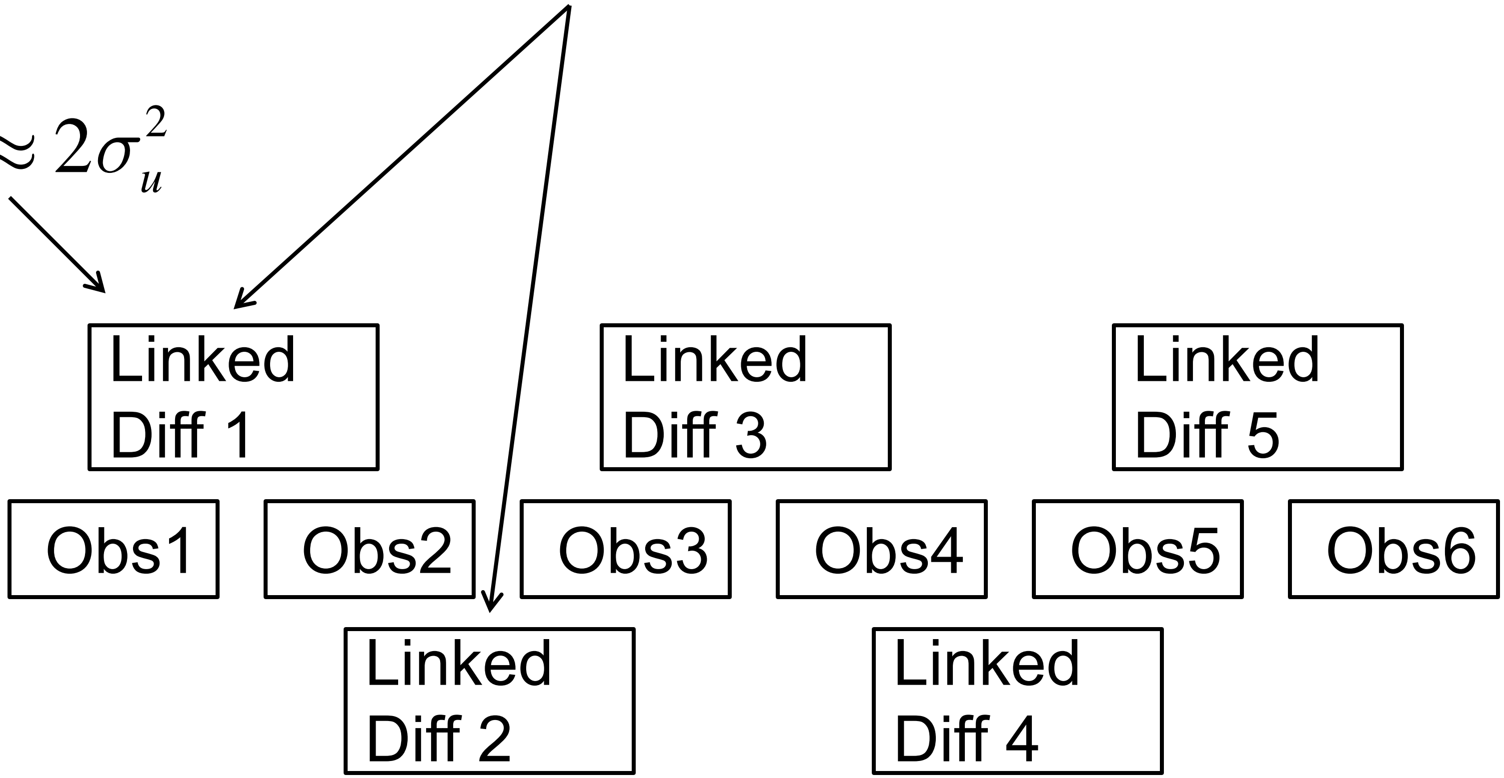
$$Var \approx 2\sigma_u^2$$



Linked Difference Observations

$$Cov(t,t+1) \approx -\sigma_u^2$$

$$Var \approx 2\sigma_u^2$$

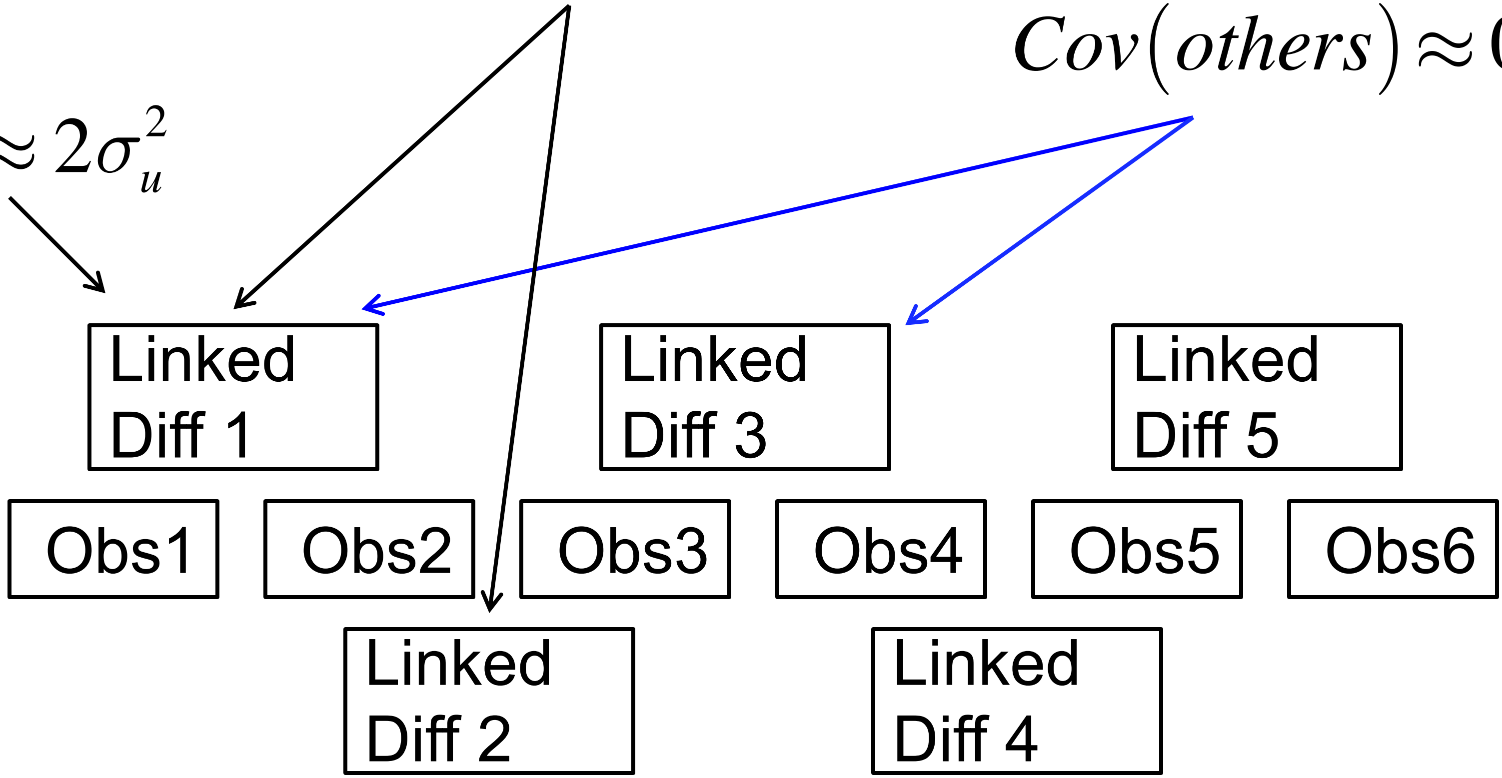


Linked Difference Observations

$$Cov(t, t+1) \approx -\sigma_u^2$$

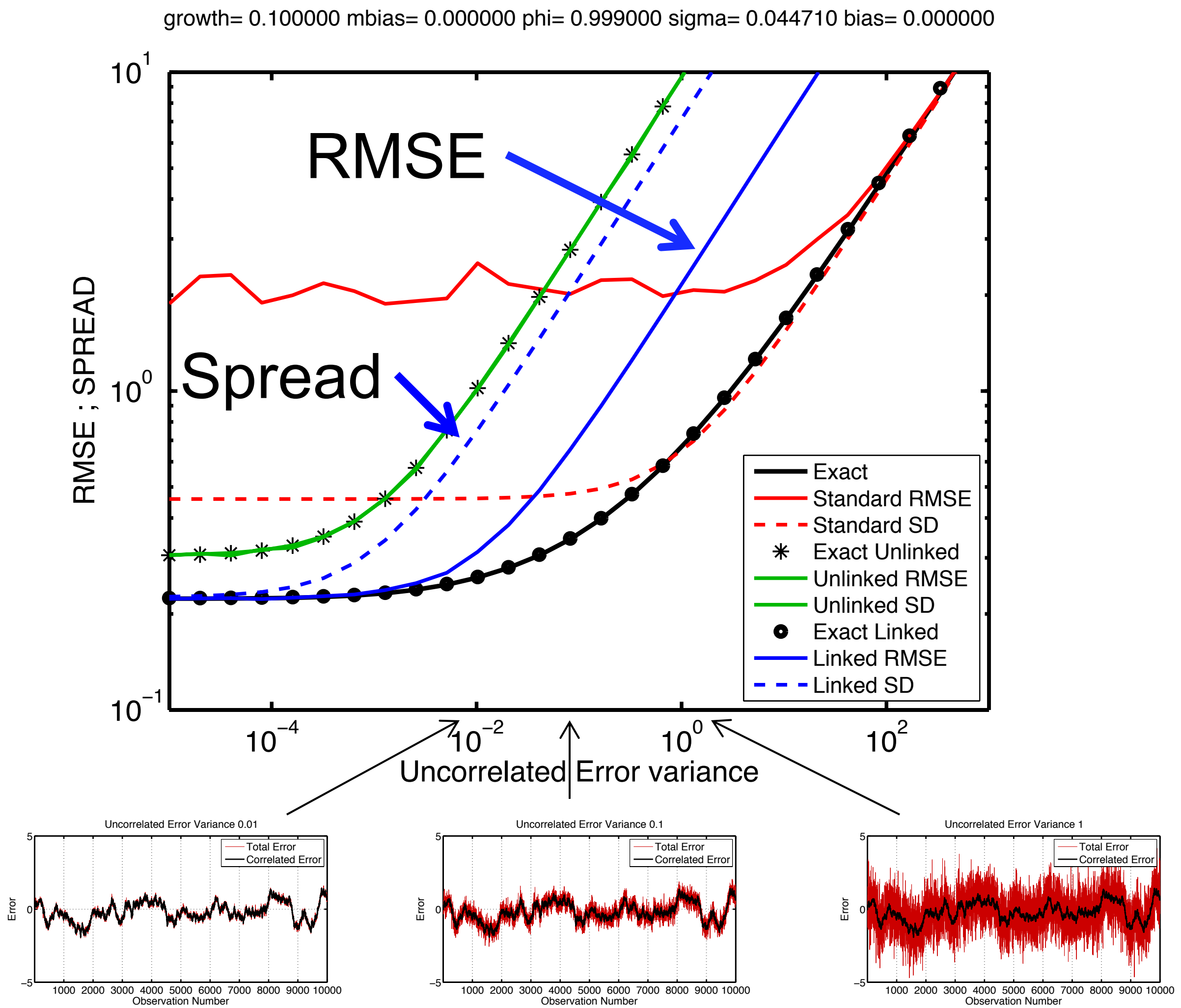
$$Cov(others) \approx 0$$

$$Var \approx 2\sigma_u^2$$



1D Exponential Growth Model Results

EAKF Linked Diff. Obs. Good when correlated error dominates.

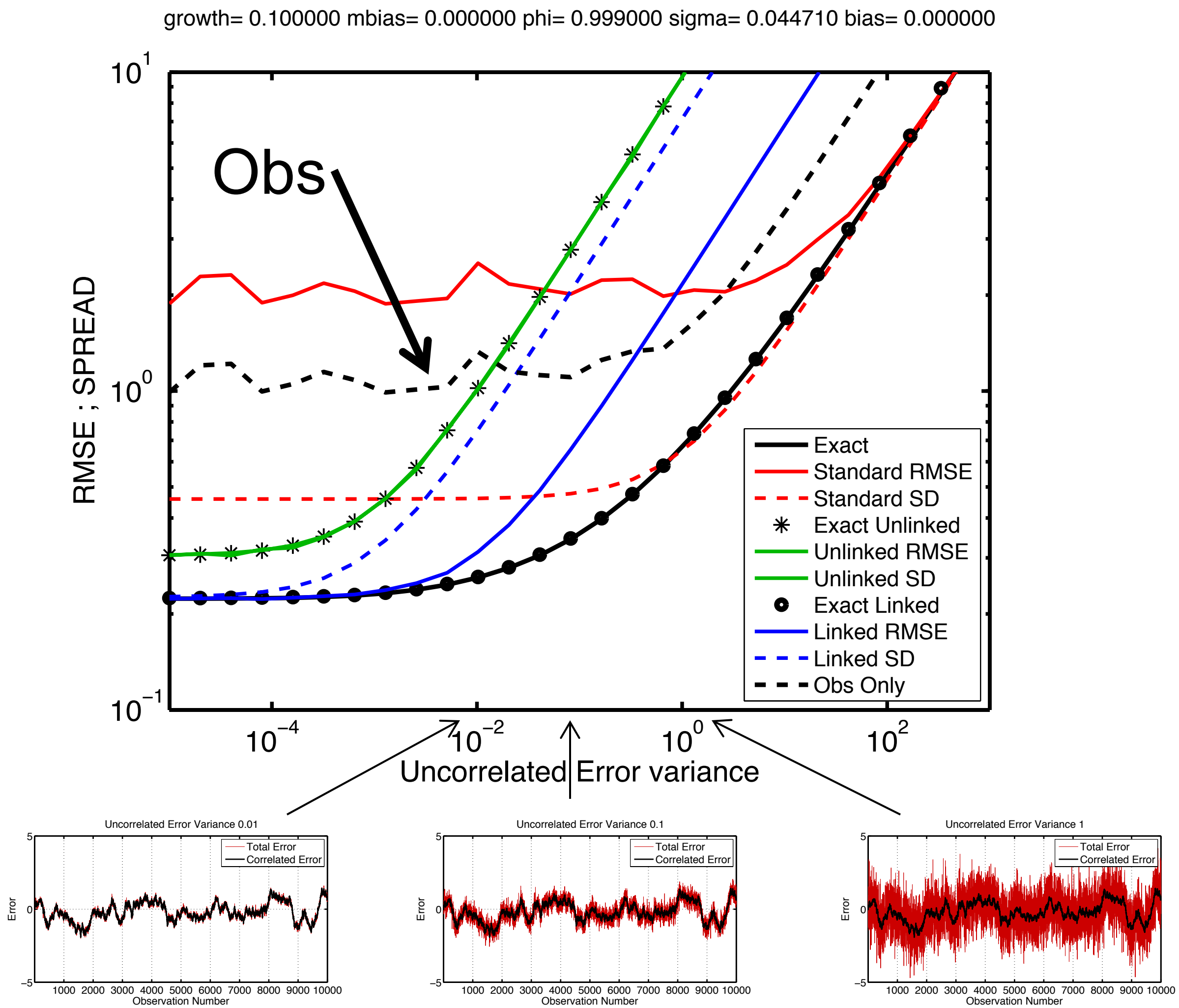


CAHMDA VII, 21 Aug. 2017

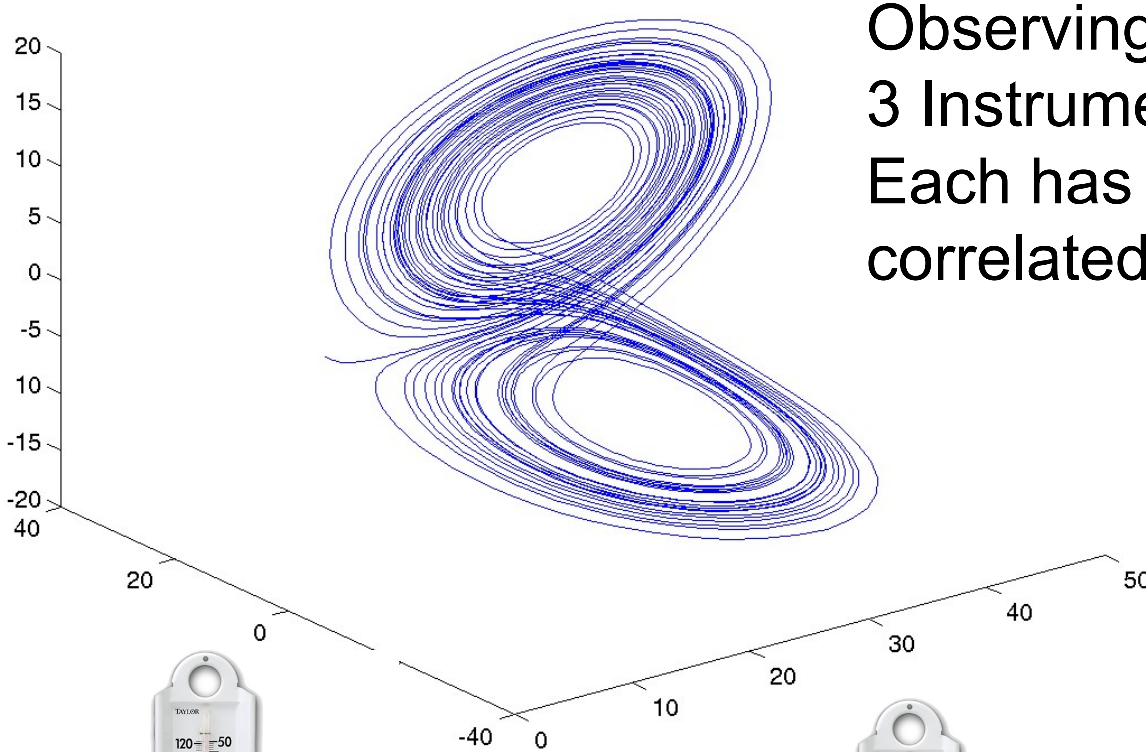
Data Assimilation Research Testbed

1D Exponential Growth Model Results

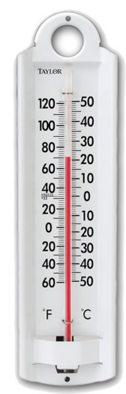
Comparison to Just Using Raw Observations



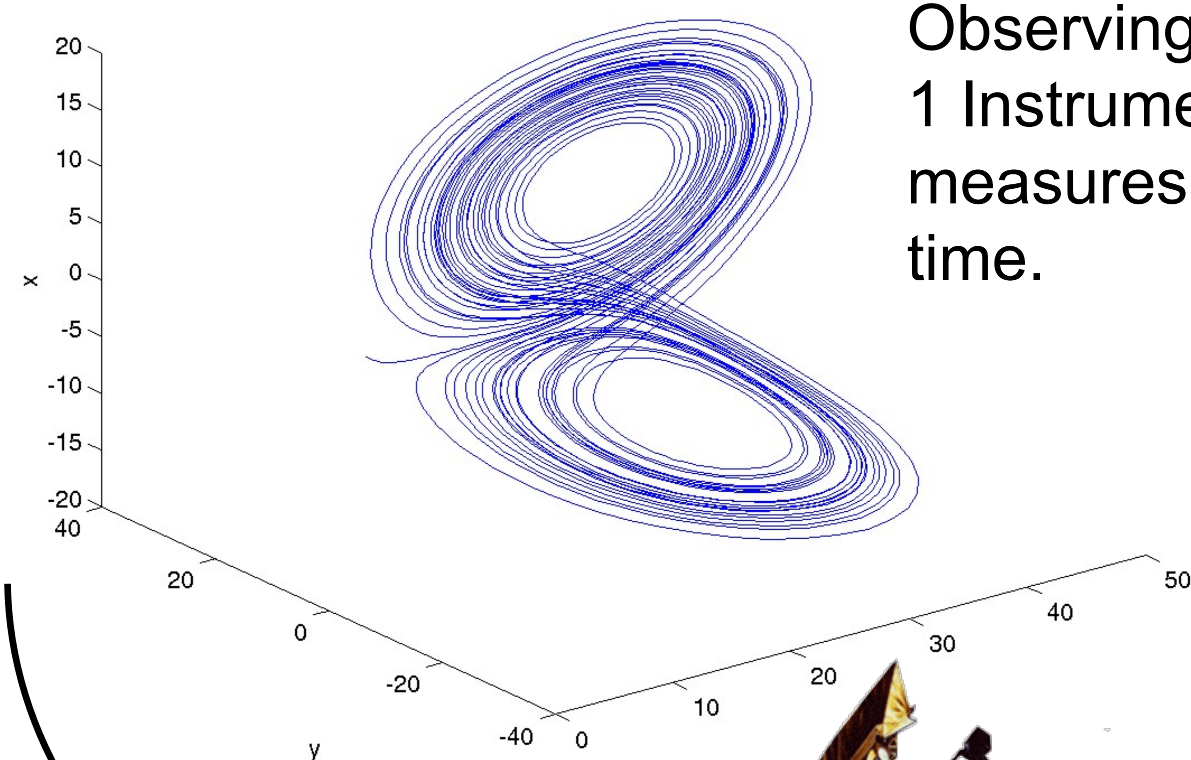
Lorenz 63 Model



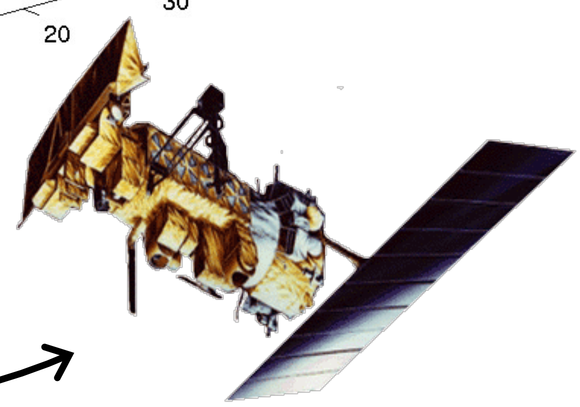
Observing System 1
3 Instruments.
Each has own correlated error.



Lorenz 63 Model

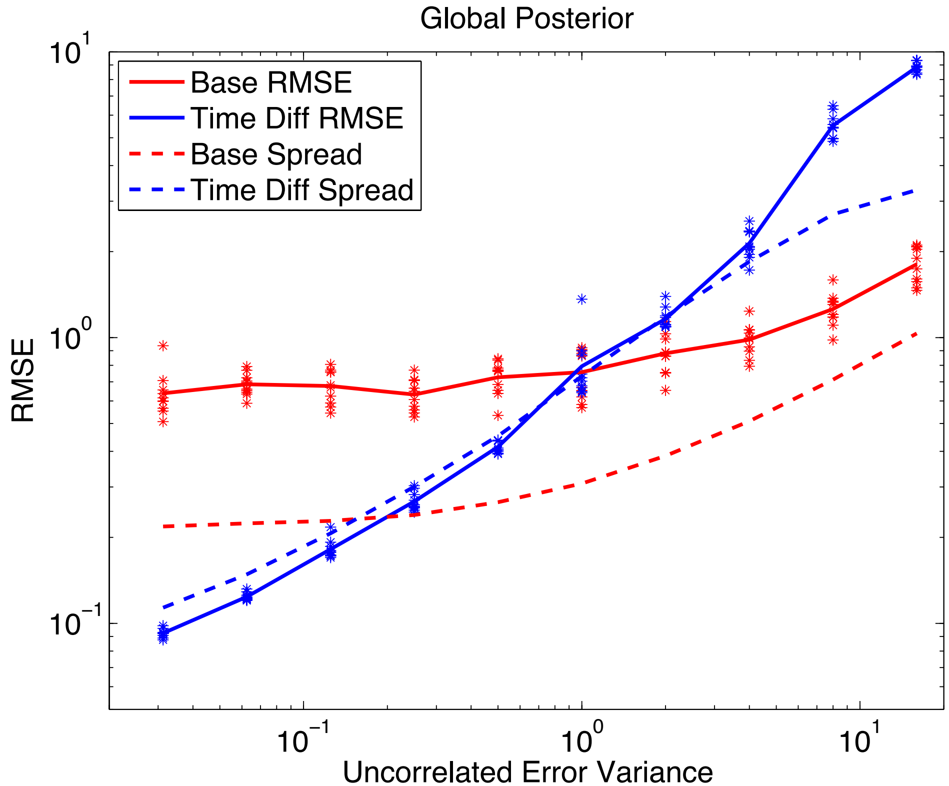


Observing System 2
1 Instrument
measures x,y,z each
time.



L63 Results, Linked Difference Obs

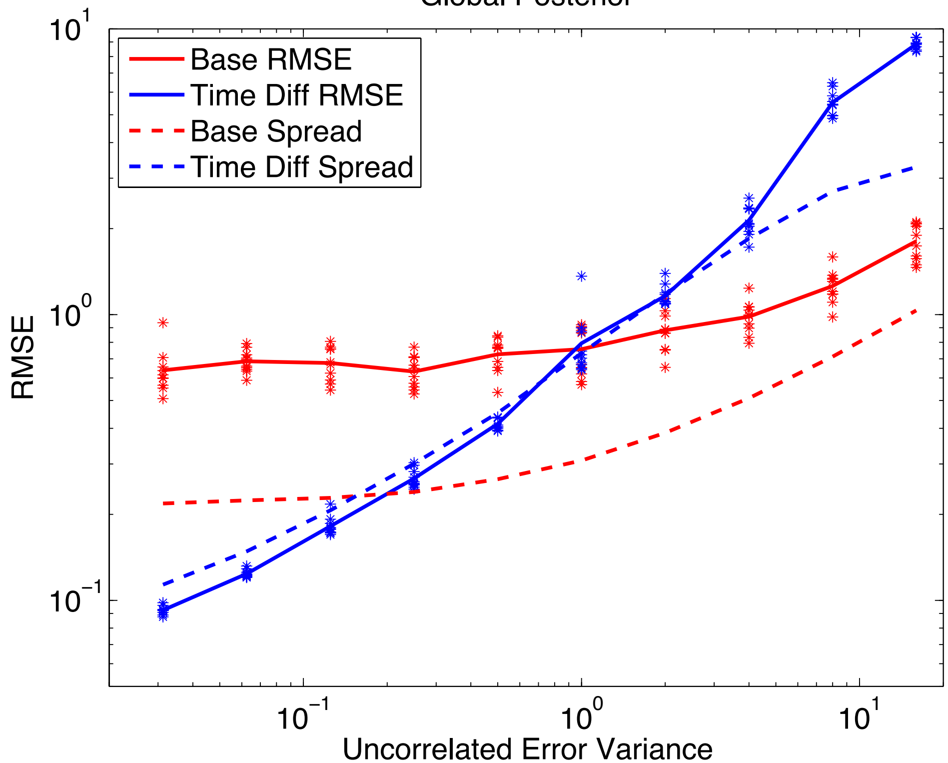
3 Instruments



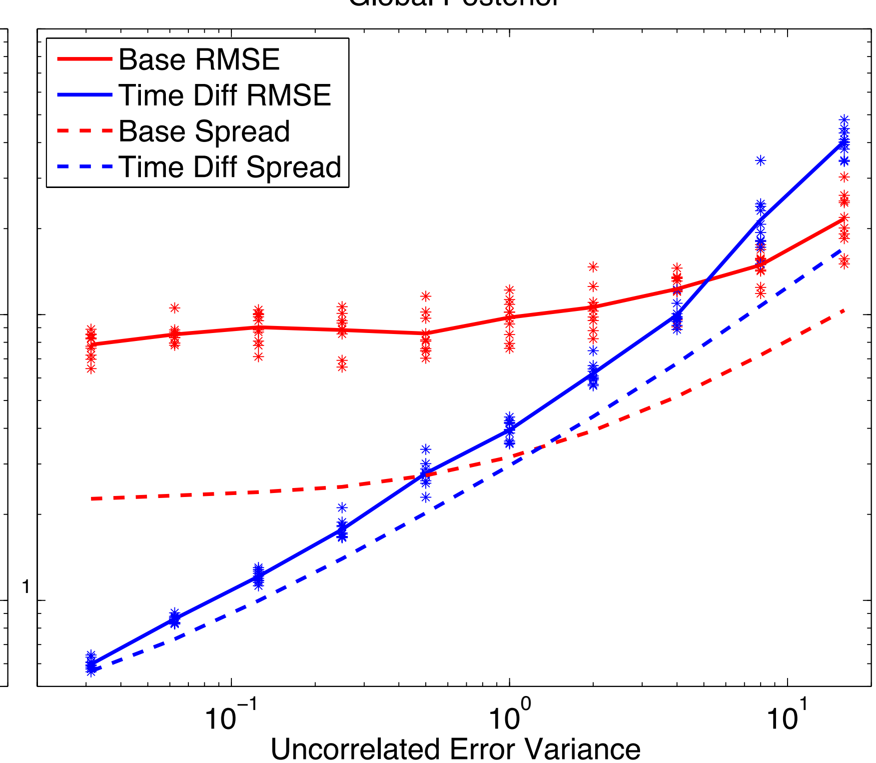
5 ensemble members.
Adaptive inflation.
Observations every 6 model timesteps.

L63 Results, Linked Difference Obs

3 Instruments



1 Instrument

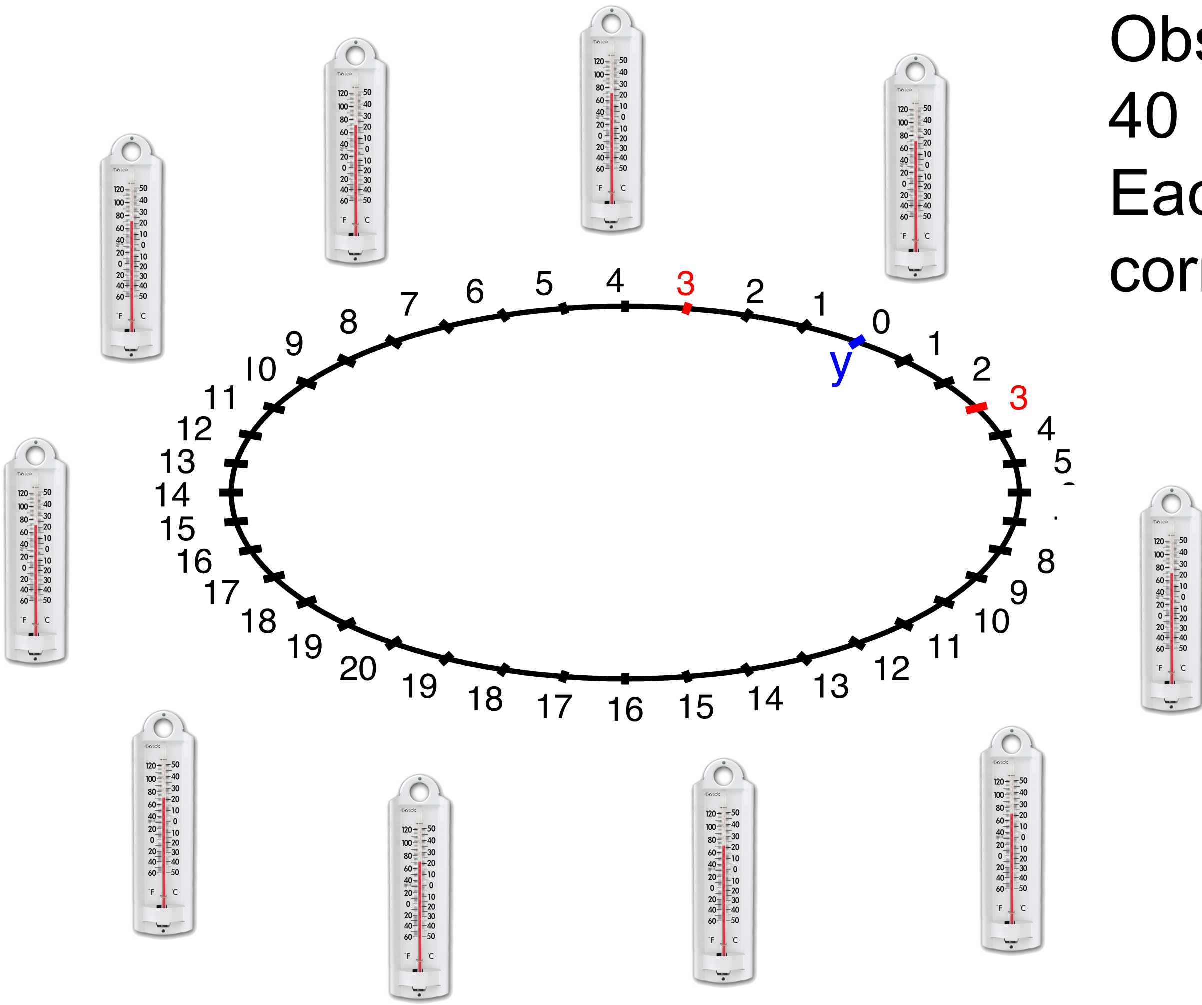


5 ensemble members.
Adaptive inflation.
Observations every 6 model timesteps.

L63 Summary

- Difference obs better unless uncorrelated error variance dominates.
- Improvement greater for single instrument.
- Ensembles often under-dispersive (what a surprise!).

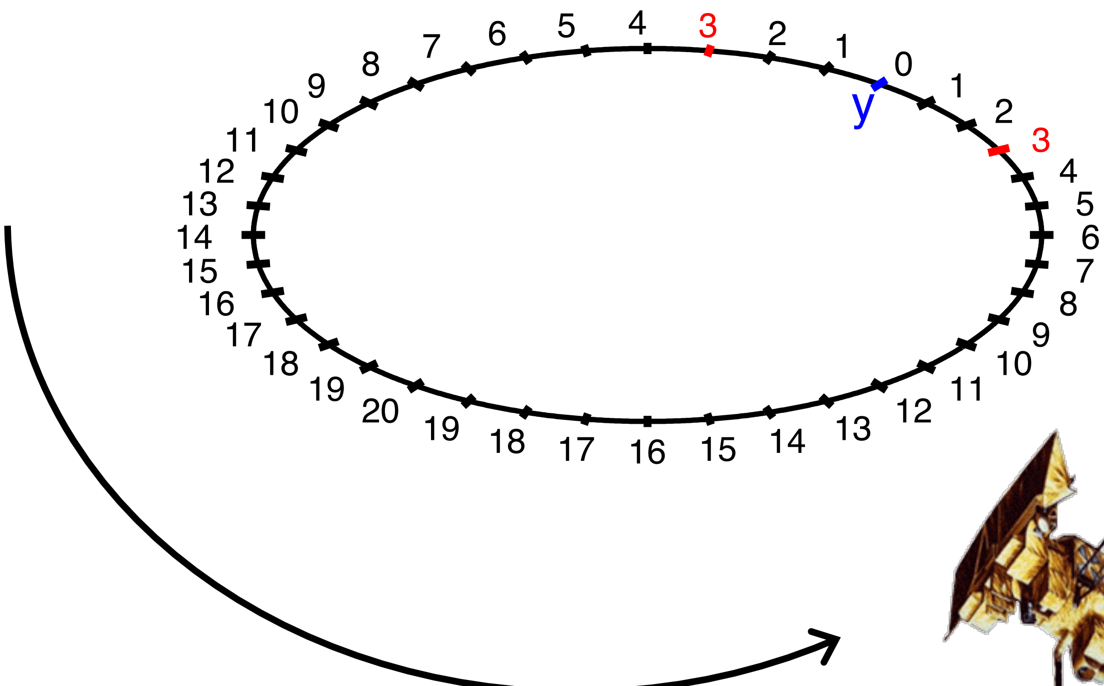
Lorenz 96 Model, 40-variables

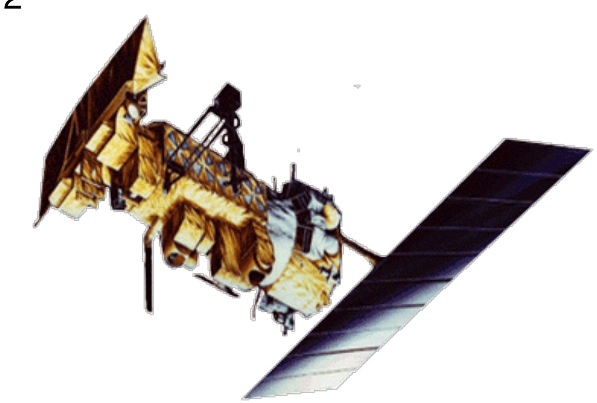


Observing System 1
40 Instruments.
Each has own
correlated error.

Lorenz 96 Model, 40-variables

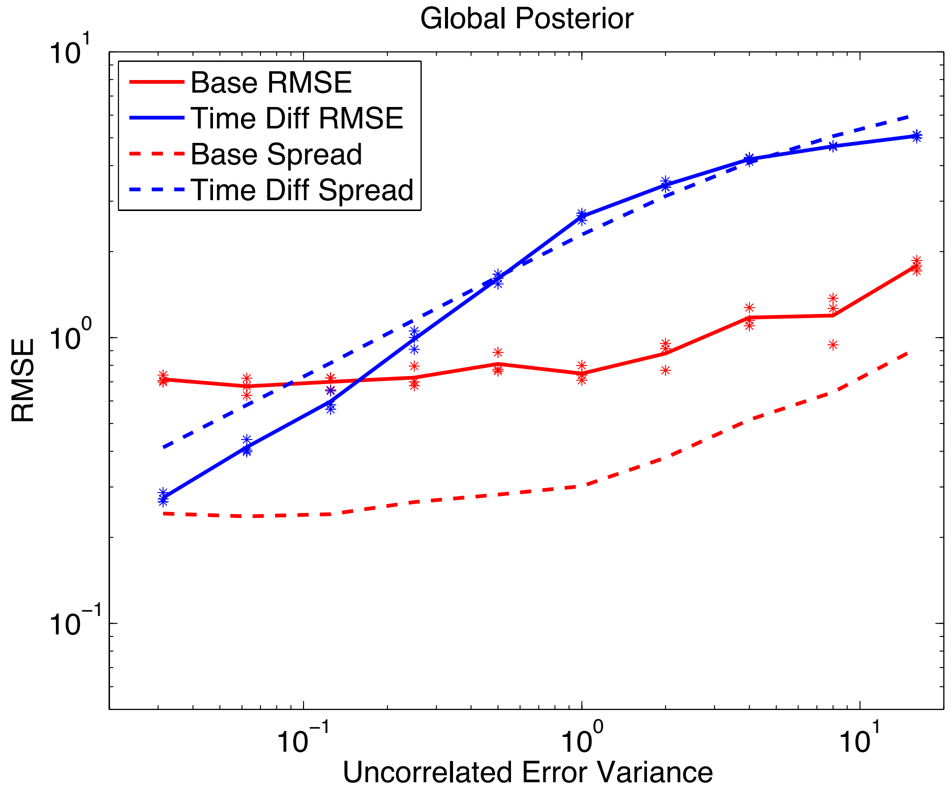
Observing System 2
1 instrument
measures all 40
variables each time.





L96 Results, Linked Difference Obs

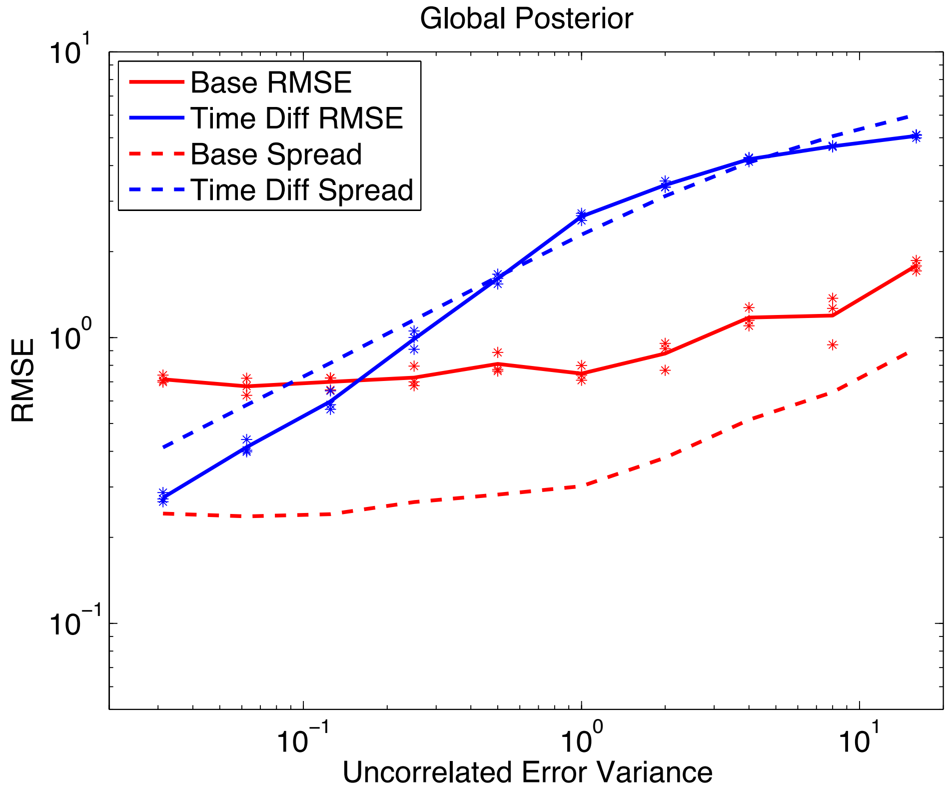
40 Instruments



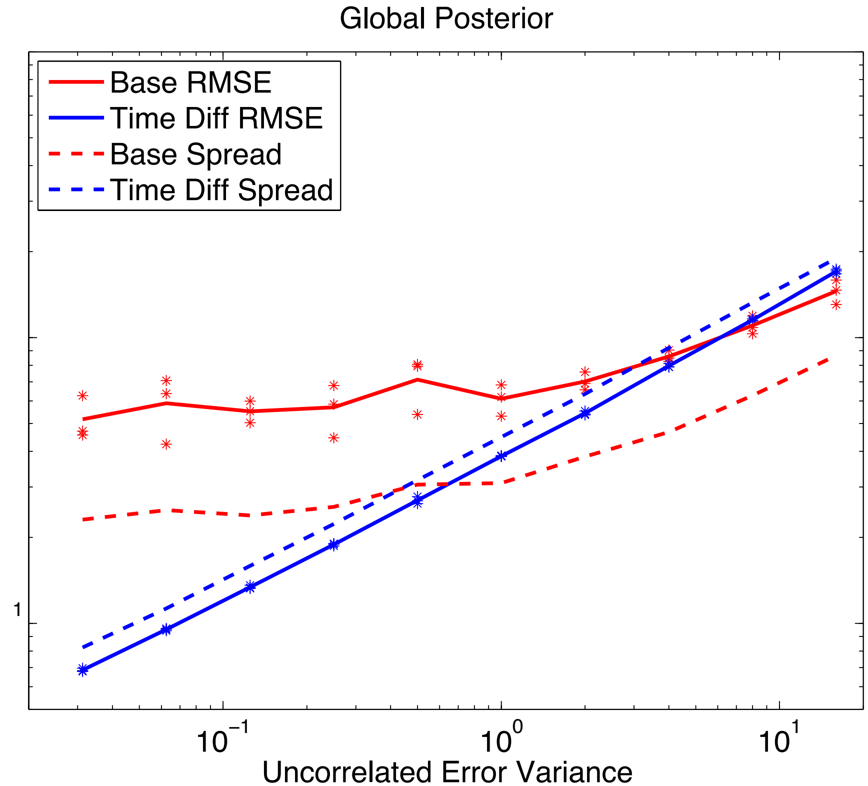
10 ensemble members.
Adaptive inflation, 0.2 halfwidth localization.
Observations every model timestep.

L96 Results, Linked Difference Obs

40 Instruments



1 Instrument

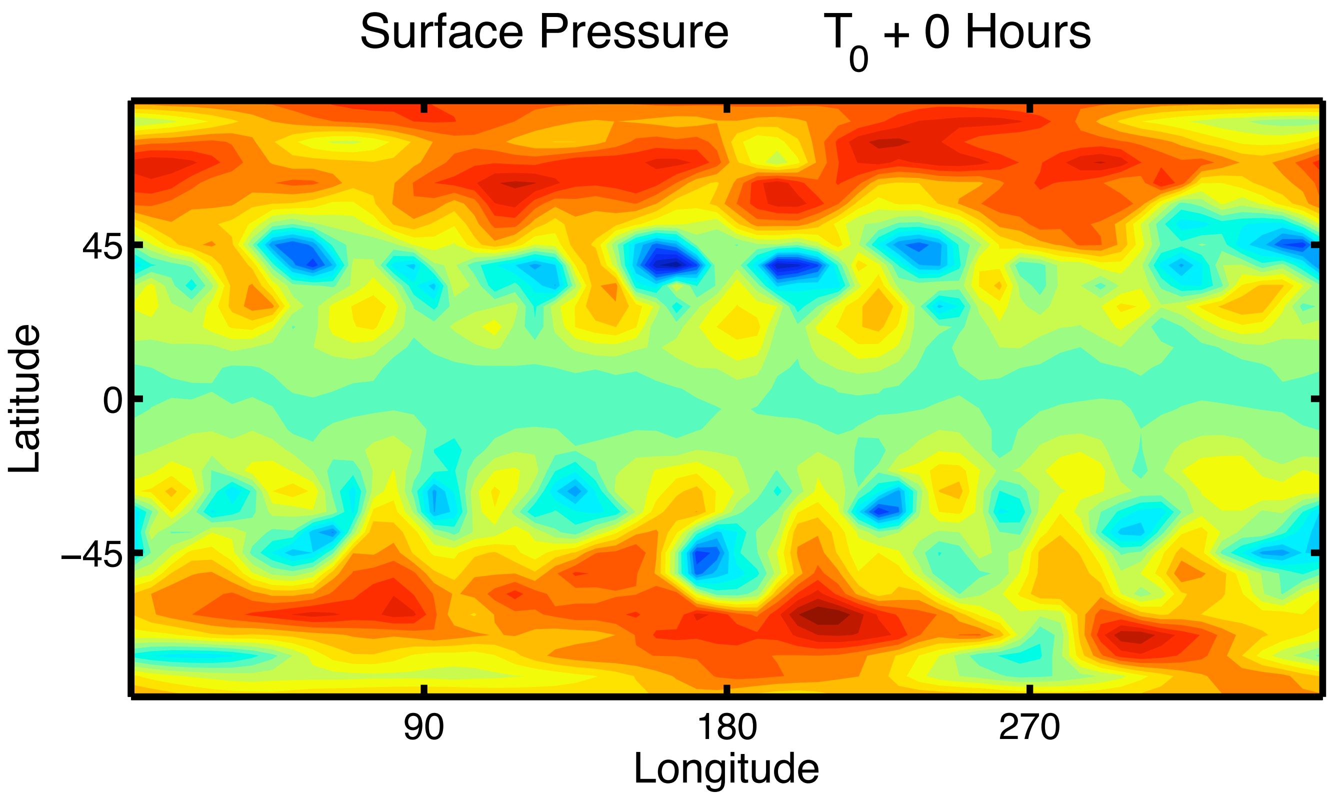


10 ensemble members.
Adaptive inflation, 0.2 halfwidth localization.
Observations every model timestep.

L96 Results, Linked Difference Obs

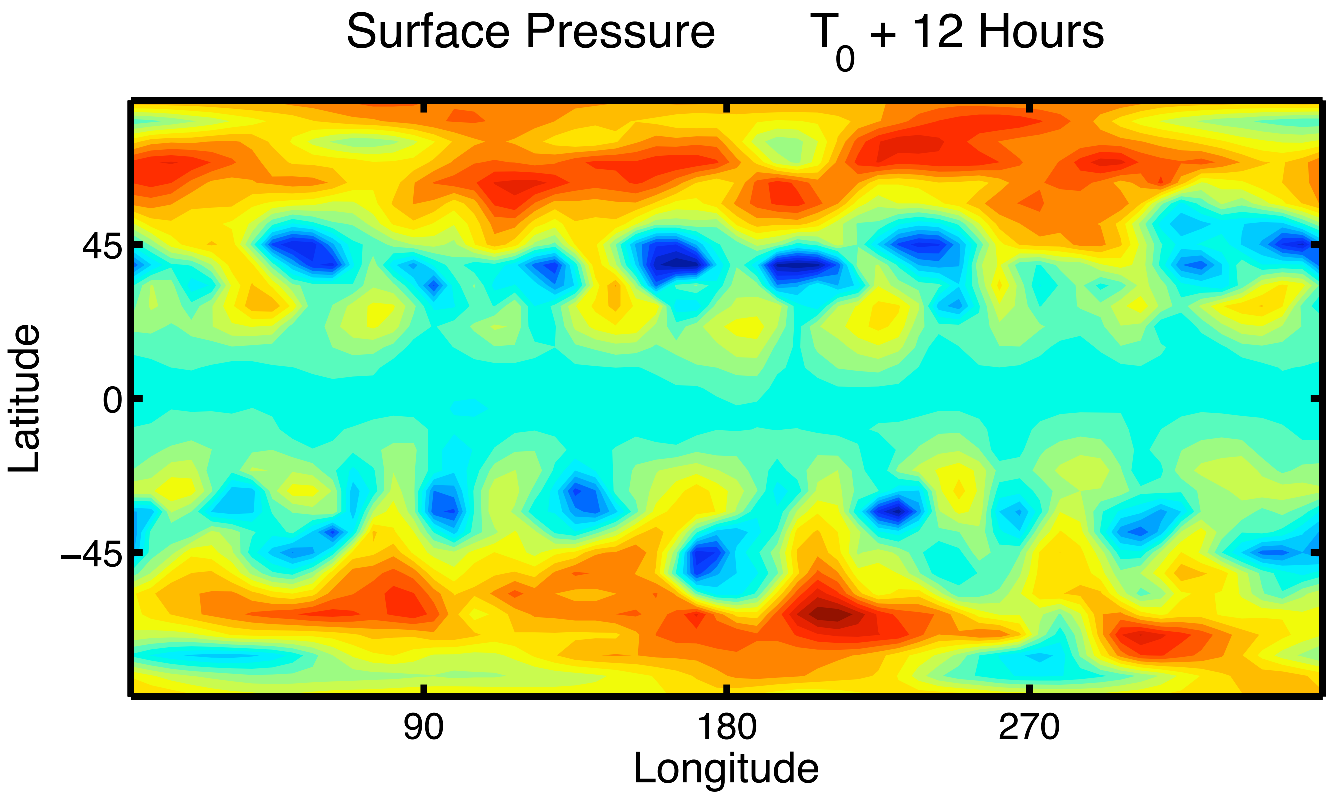
- Difference obs better unless uncorrelated error variance dominates.
- Improvement much greater for single instrument.
- Ensembles often over-dispersive.
- Dealing with time correlation harder than space correlation.

Low-Order Dry Dynamical Core



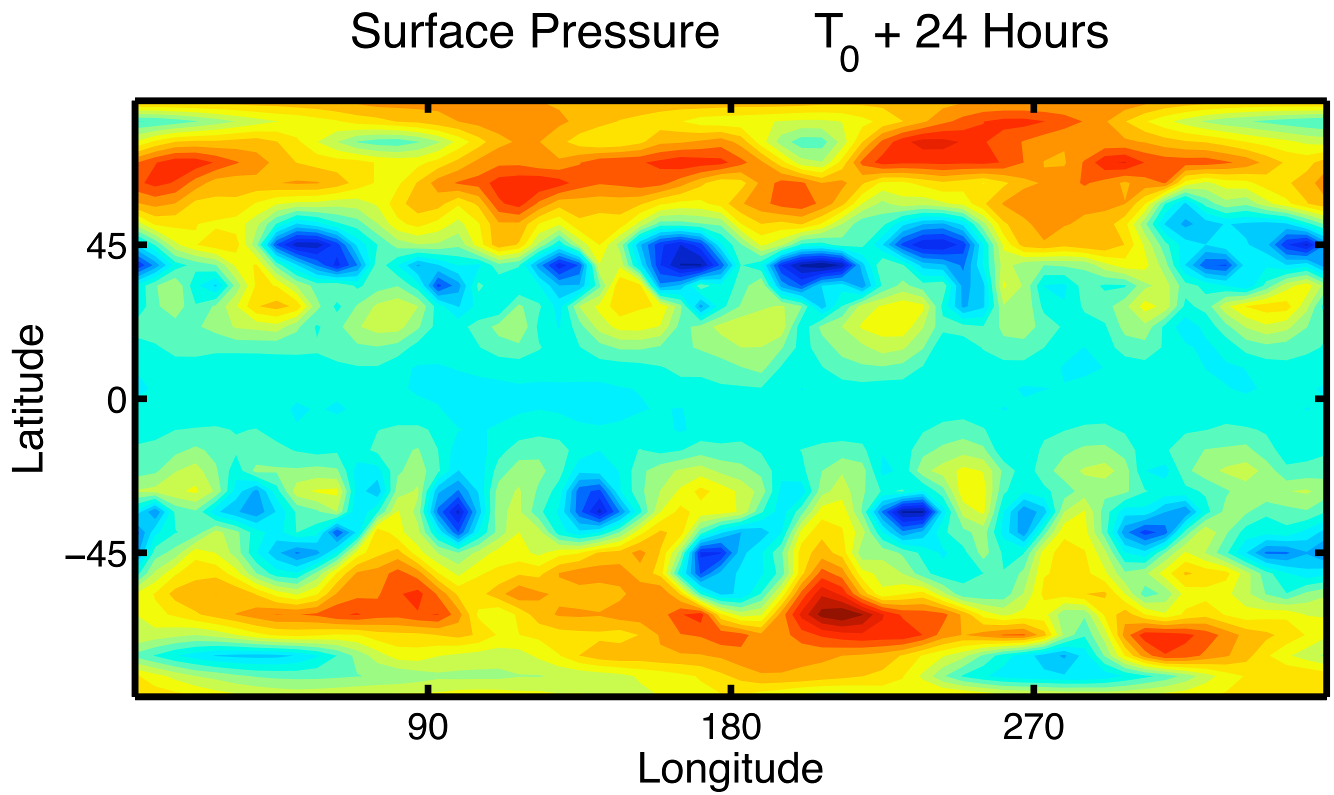
Evolution of surface pressure field every 12 hours.
Has baroclinic instability: storms move east in midlatitudes.

Low-Order Dry Dynamical Core



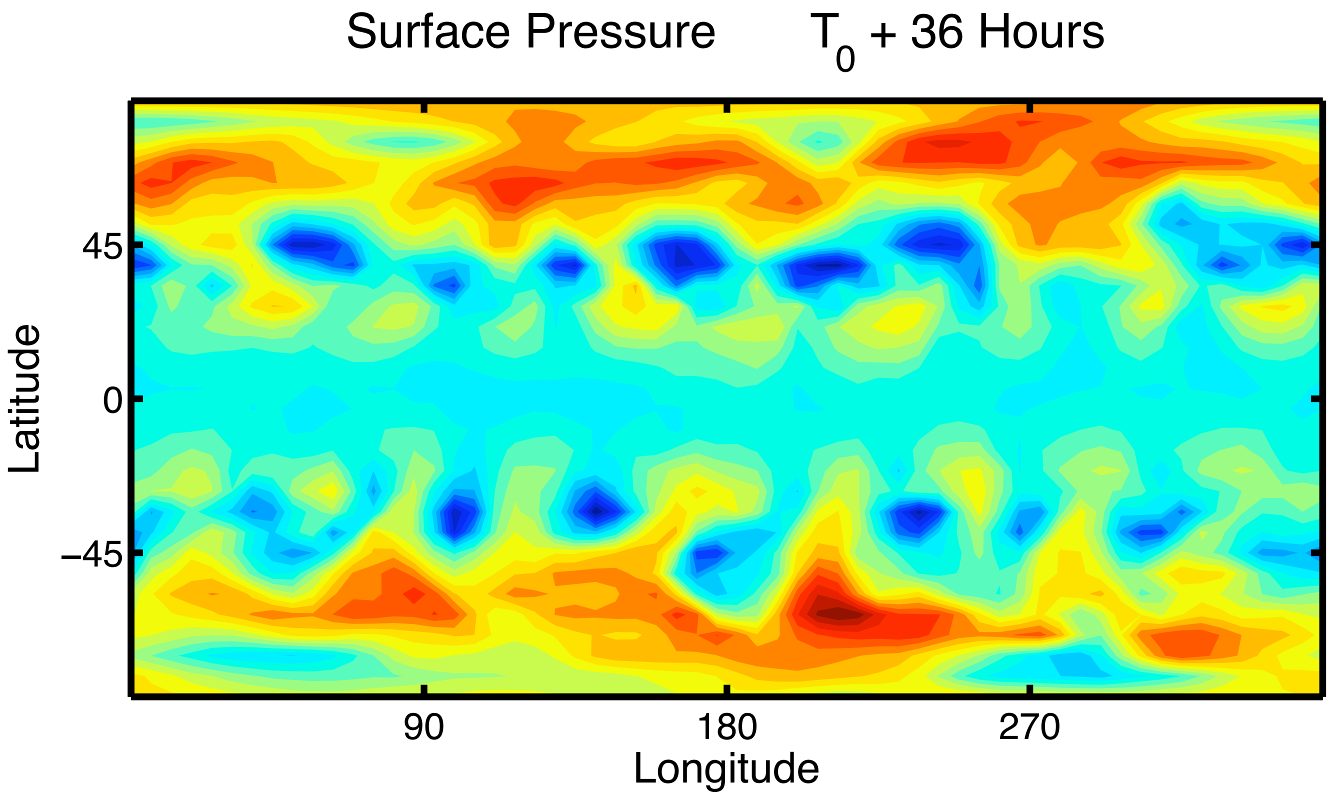
Evolution of surface pressure field every 12 hours.
Has baroclinic instability: storms move east in midlatitudes.

Low-Order Dry Dynamical Core



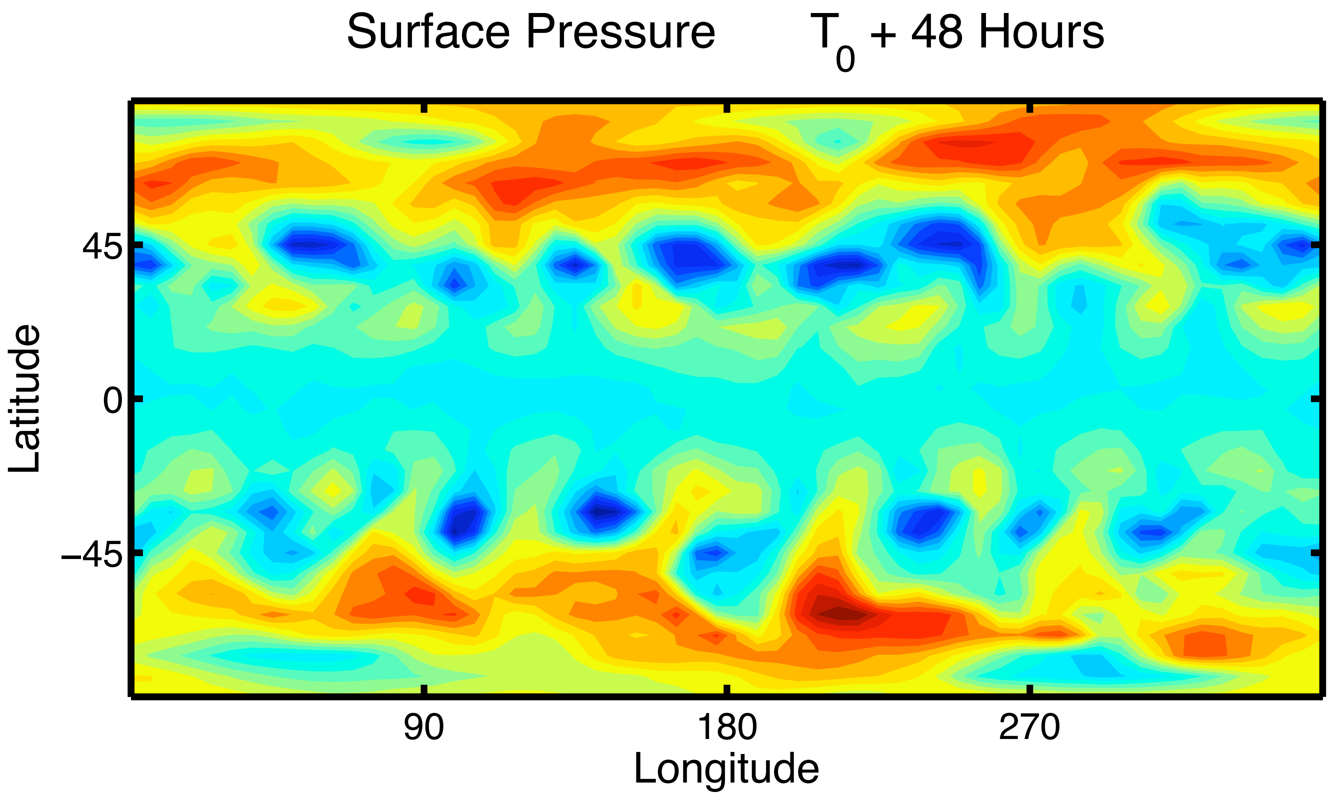
Evolution of surface pressure field every 12 hours.
Has baroclinic instability: storms move east in midlatitudes.

Low-Order Dry Dynamical Core



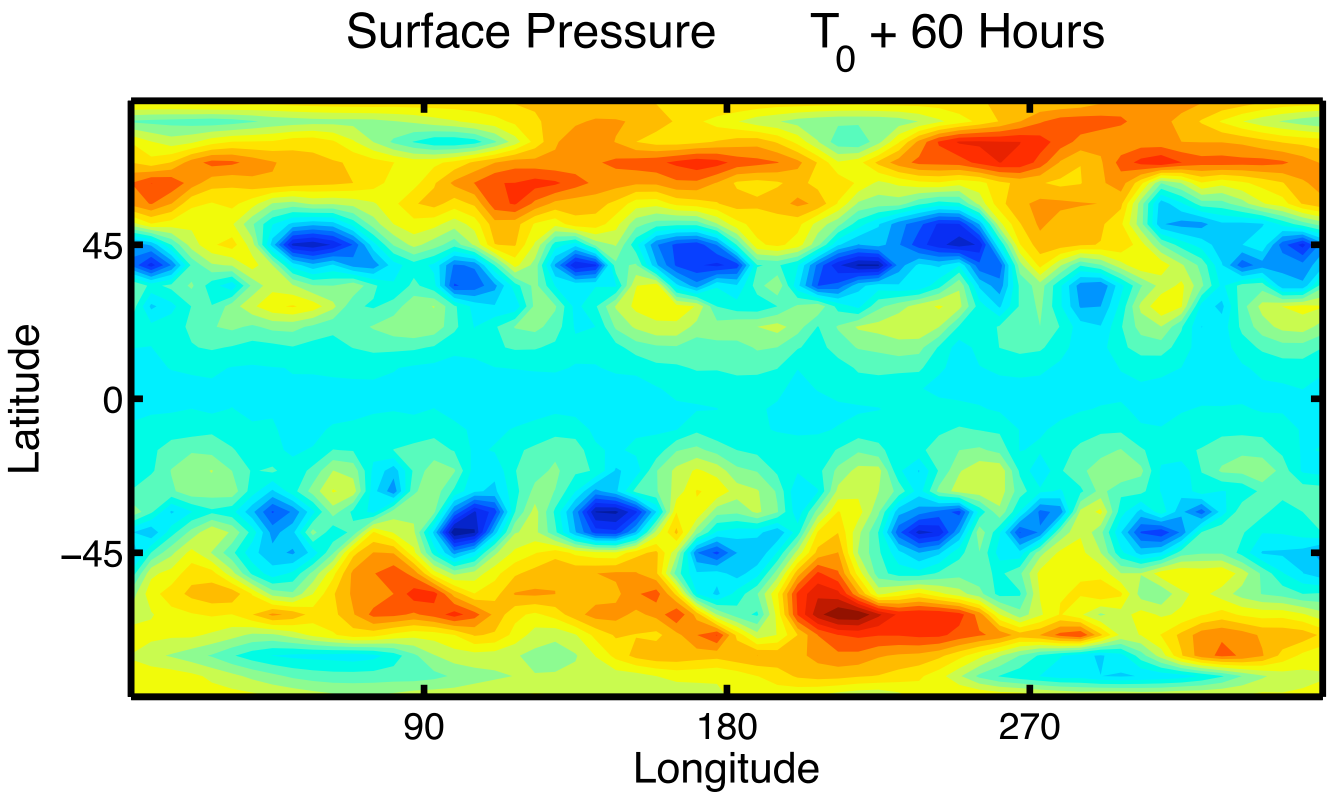
Evolution of surface pressure field every 12 hours.
Has baroclinic instability: storms move east in midlatitudes.

Low-Order Dry Dynamical Core



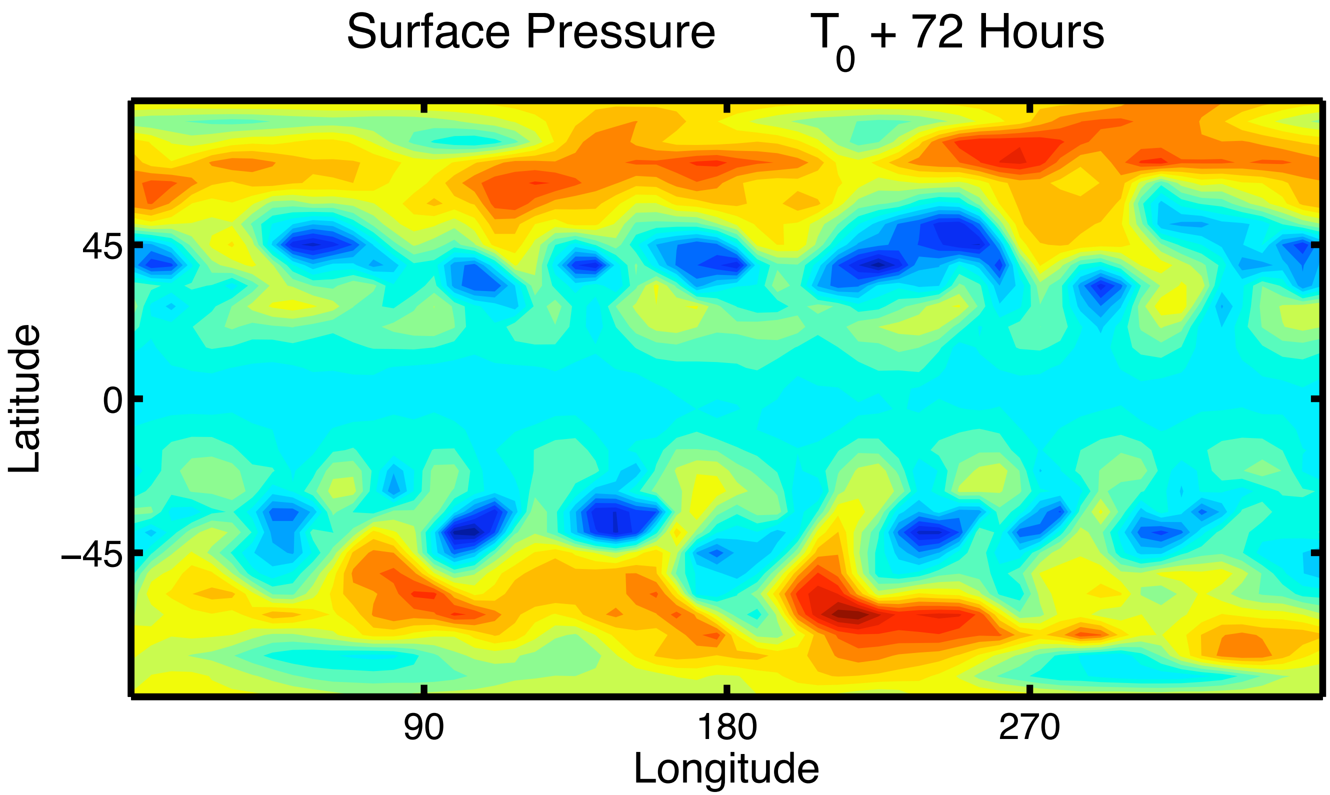
Evolution of surface pressure field every 12 hours.
Has baroclinic instability: storms move east in midlatitudes.

Low-Order Dry Dynamical Core



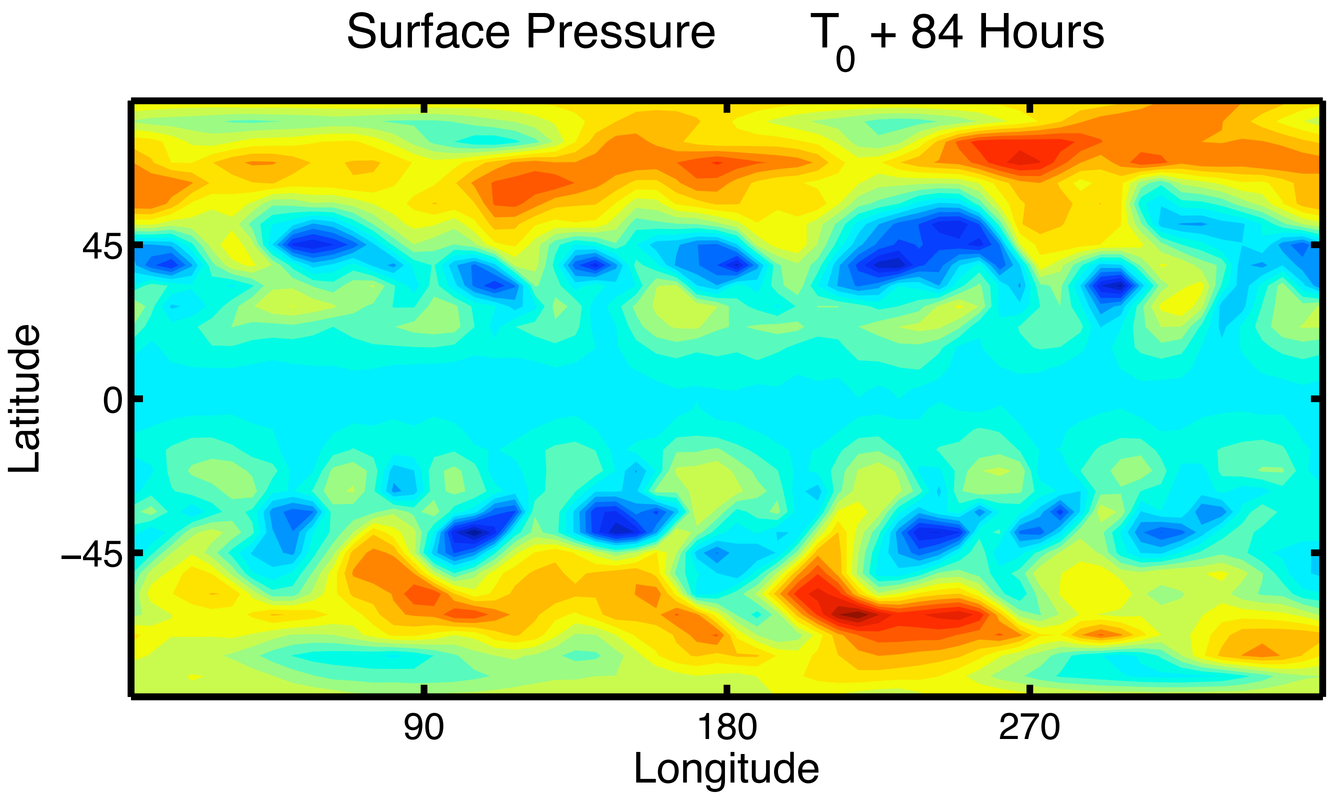
Evolution of surface pressure field every 12 hours.
Has baroclinic instability: storms move east in midlatitudes.

Low-Order Dry Dynamical Core



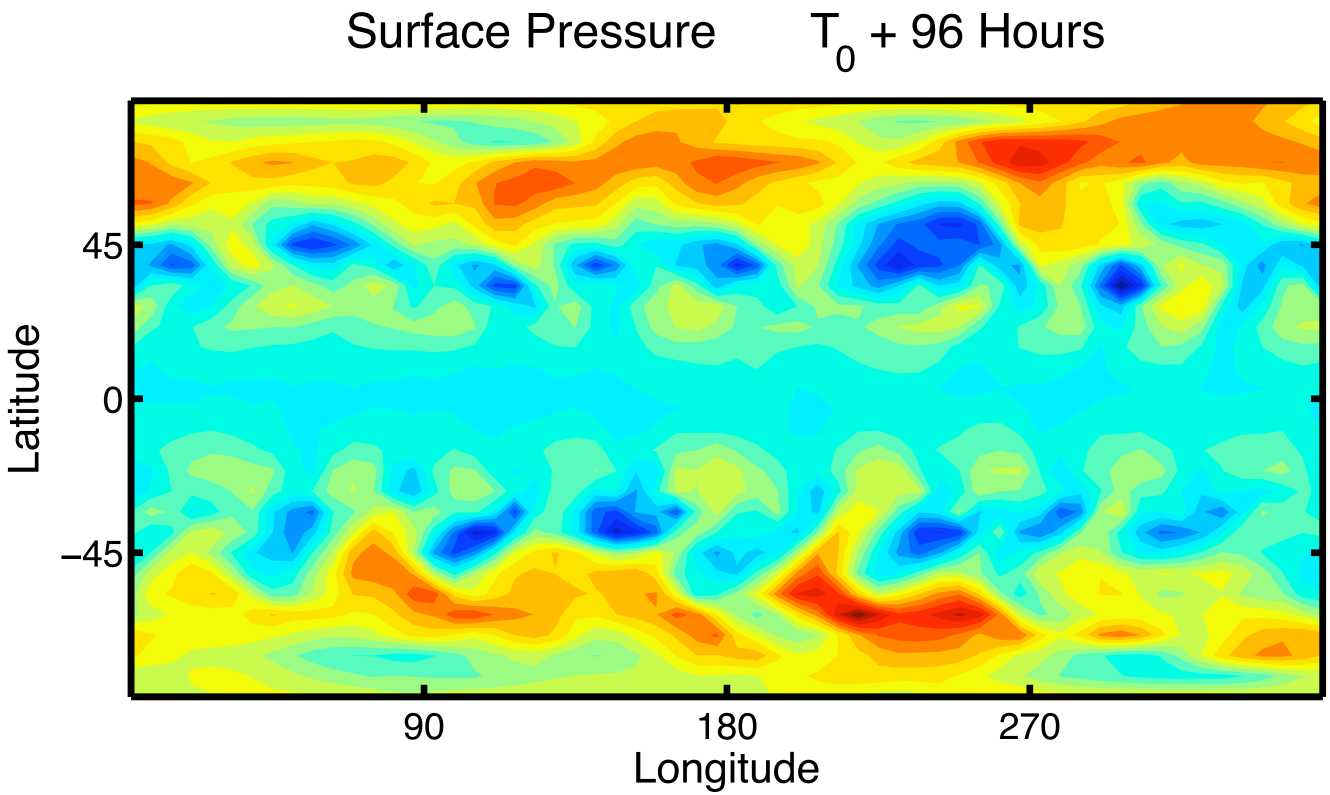
Evolution of surface pressure field every 12 hours.
Has baroclinic instability: storms move east in midlatitudes.

Low-Order Dry Dynamical Core



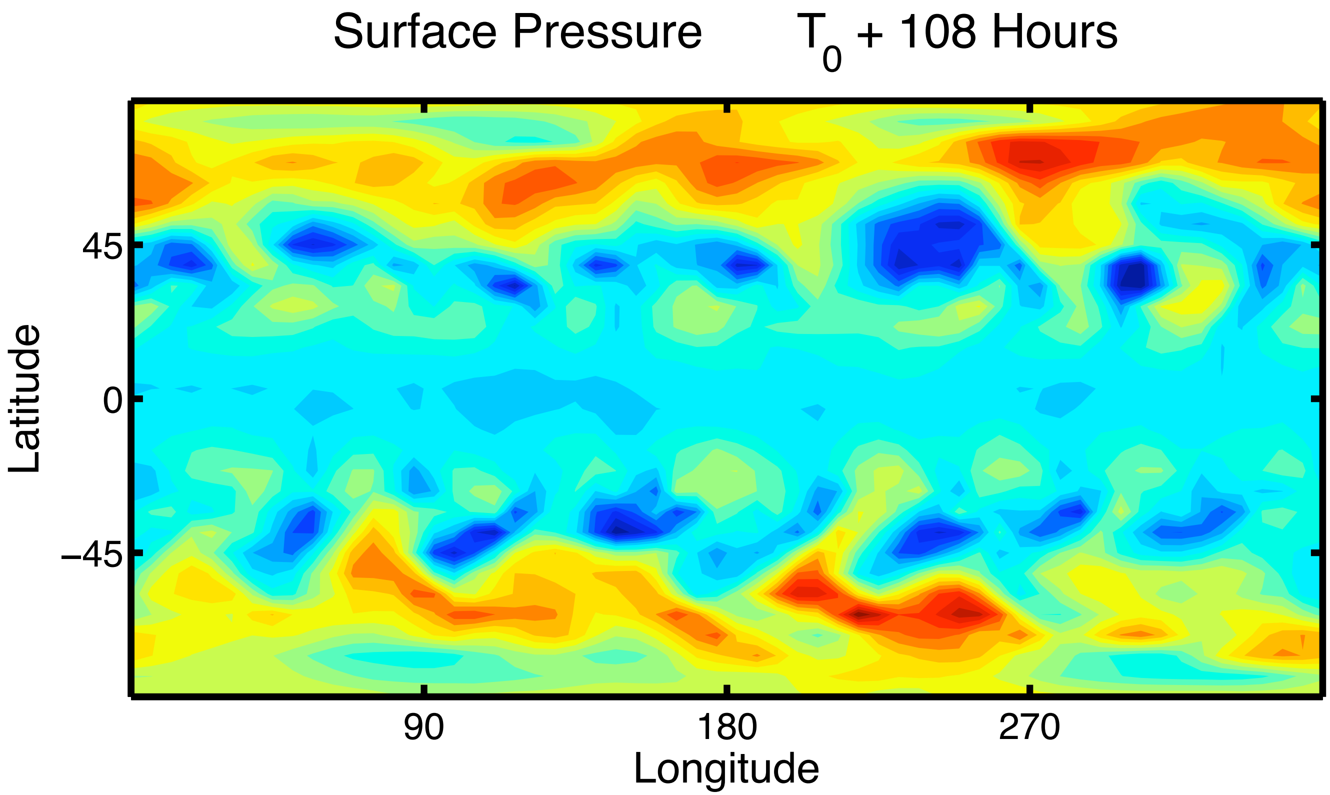
Evolution of surface pressure field every 12 hours.
Has baroclinic instability: storms move east in midlatitudes.

Low-Order Dry Dynamical Core



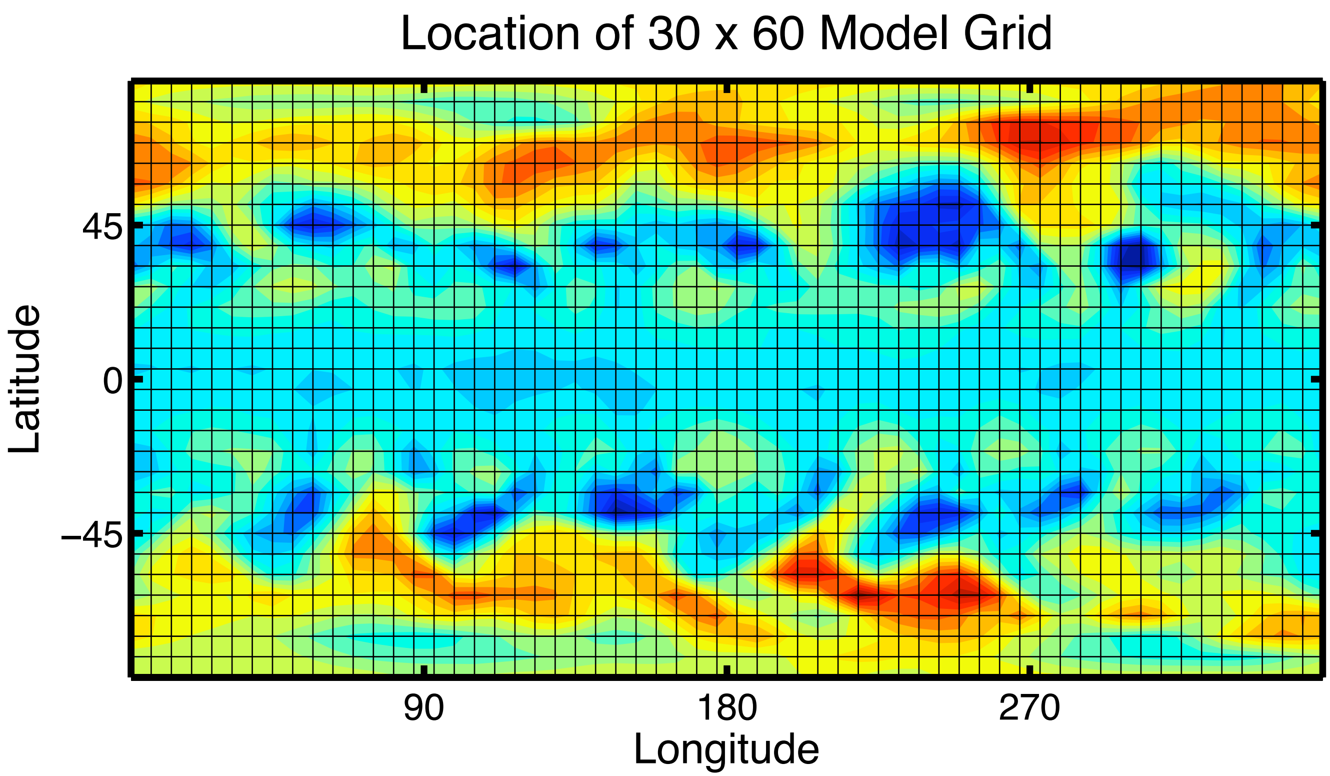
Evolution of surface pressure field every 12 hours.
Has baroclinic instability: storms move east in midlatitudes.

Low-Order Dry Dynamical Core



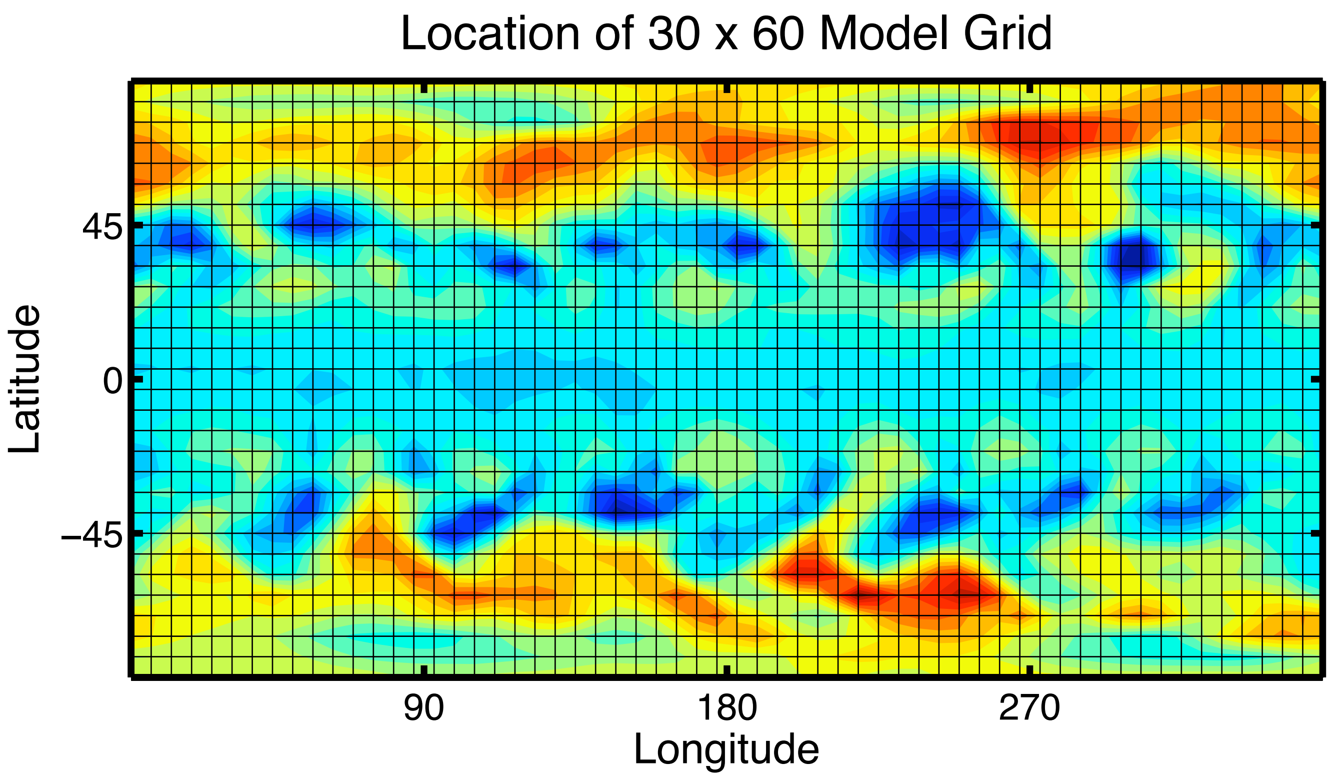
Evolution of surface pressure field every 12 hours.
Has baroclinic instability: storms move east in midlatitudes.

Low-Order Dry Dynamical Core: Grid



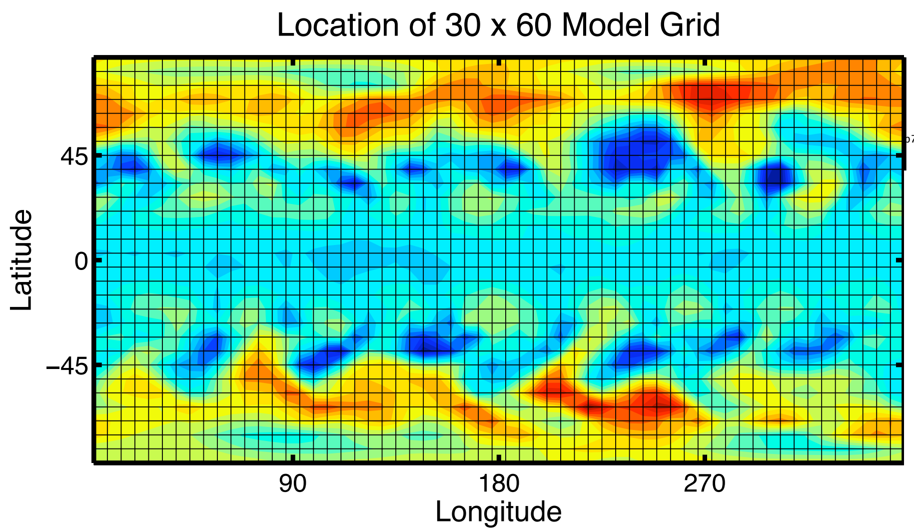
30x60 horizontal grid, 5 levels.
Surface pressure, temperature, wind components.
28,800 variables.

Low-Order Dry Dynamical Core: Observations

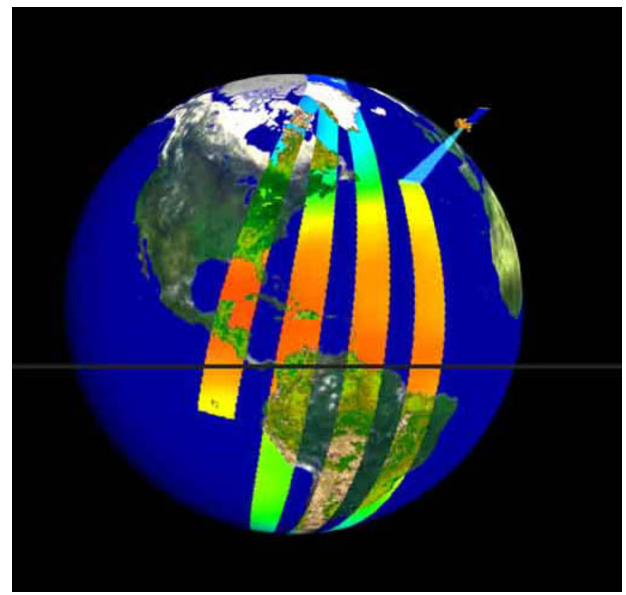


Assimilate once per day. 0.2 radian localization.
Observe each surface pressure grid point.
Uncorrelated obs error variance 100 Pa.

Low-Order Dry Dynamical Core: Observations

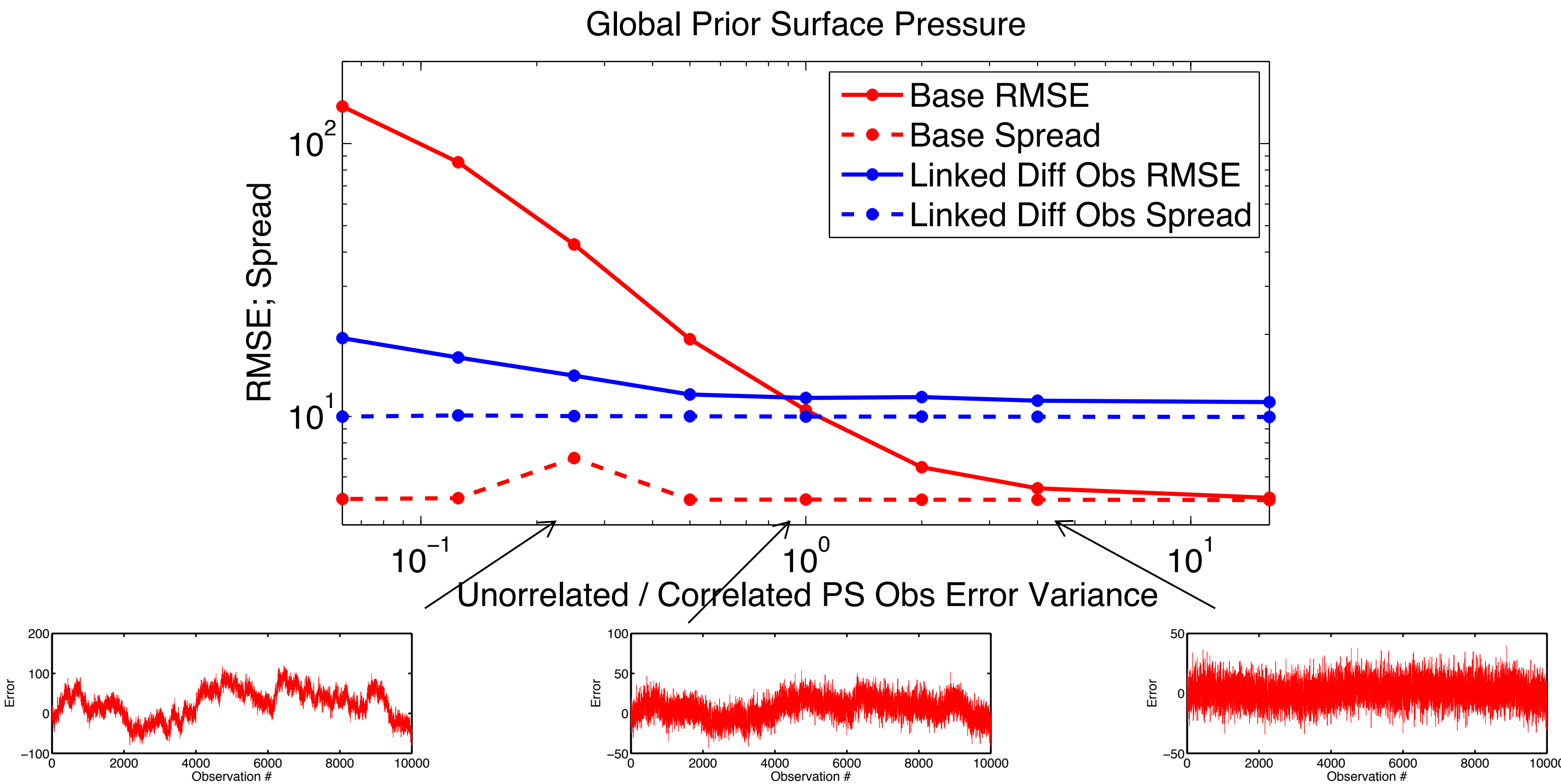


76.jpg 510×490 pixels
1/5/16, 10:53 AM



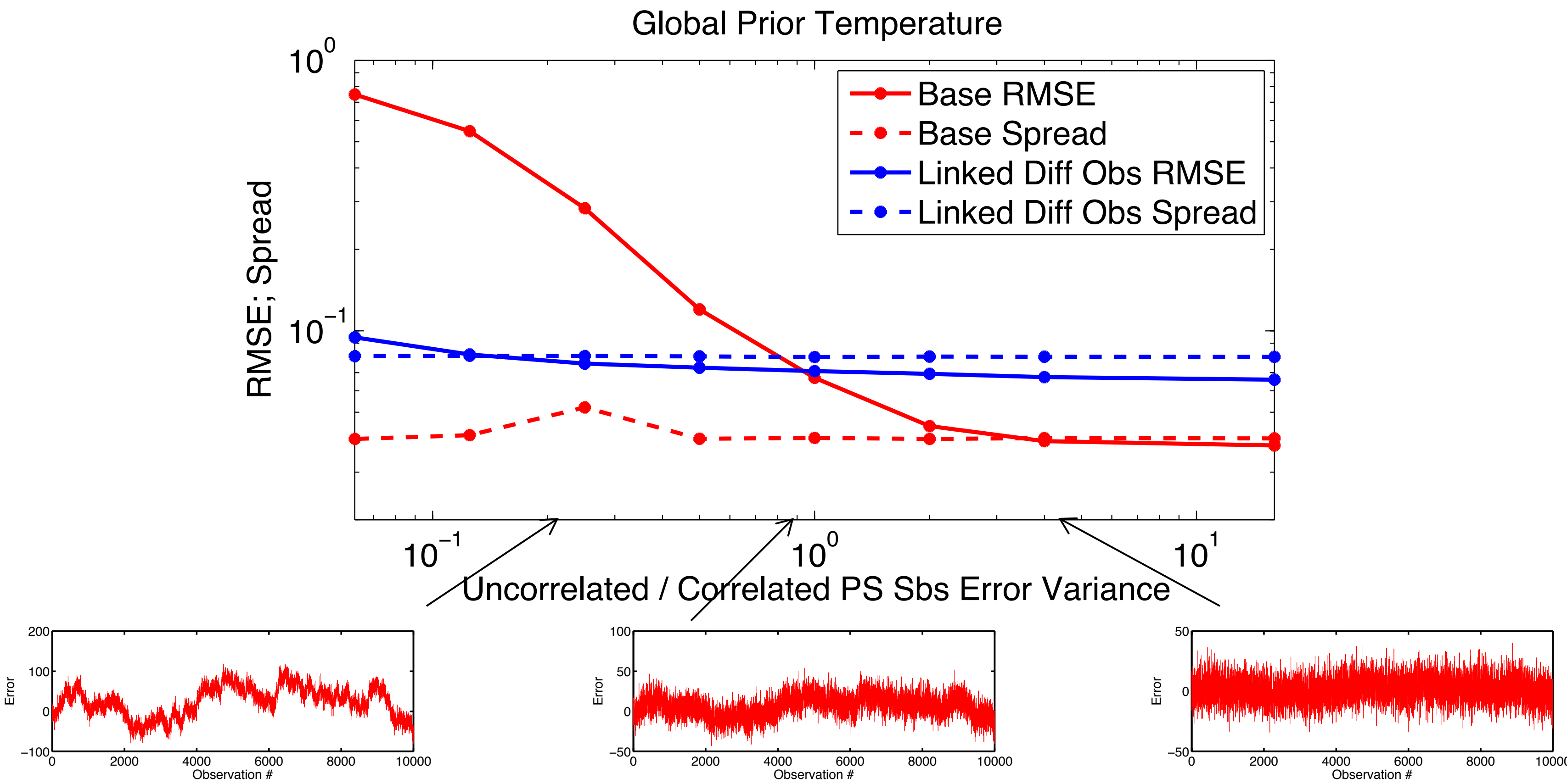
Uncorrelated obs error variance 100 Pa.
Correlated obs error along ‘simulated polar orbiter track’.
Vary ratio of correlated to uncorrelated obs error variance.

Low-Order Dry Dynamical Core: PS Results



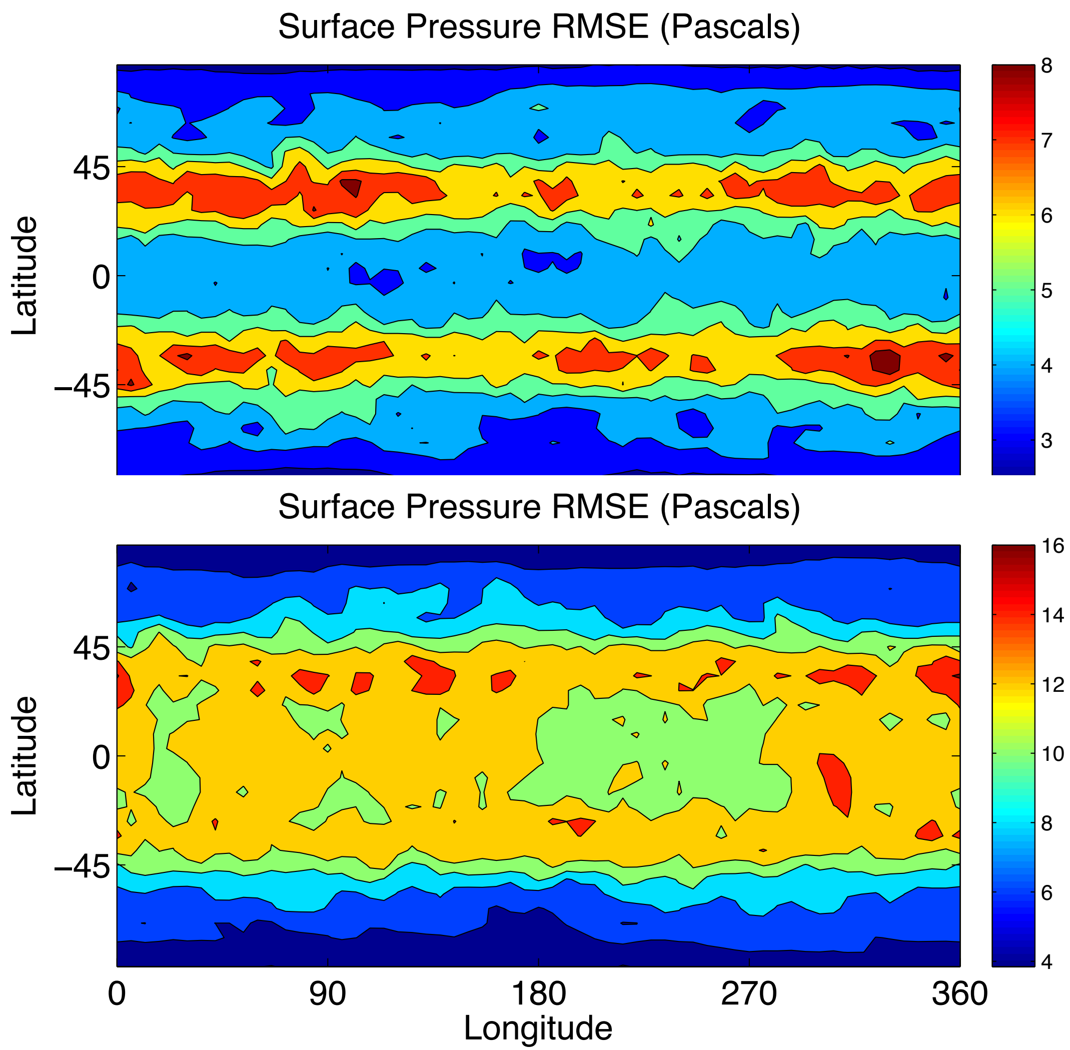
Linked difference better for large correlated error.
Standard better for small correlated error.

Low-Order Dry Dynamical Core: T Results



Linked difference better for large correlated error.
Standard better for small correlated error.

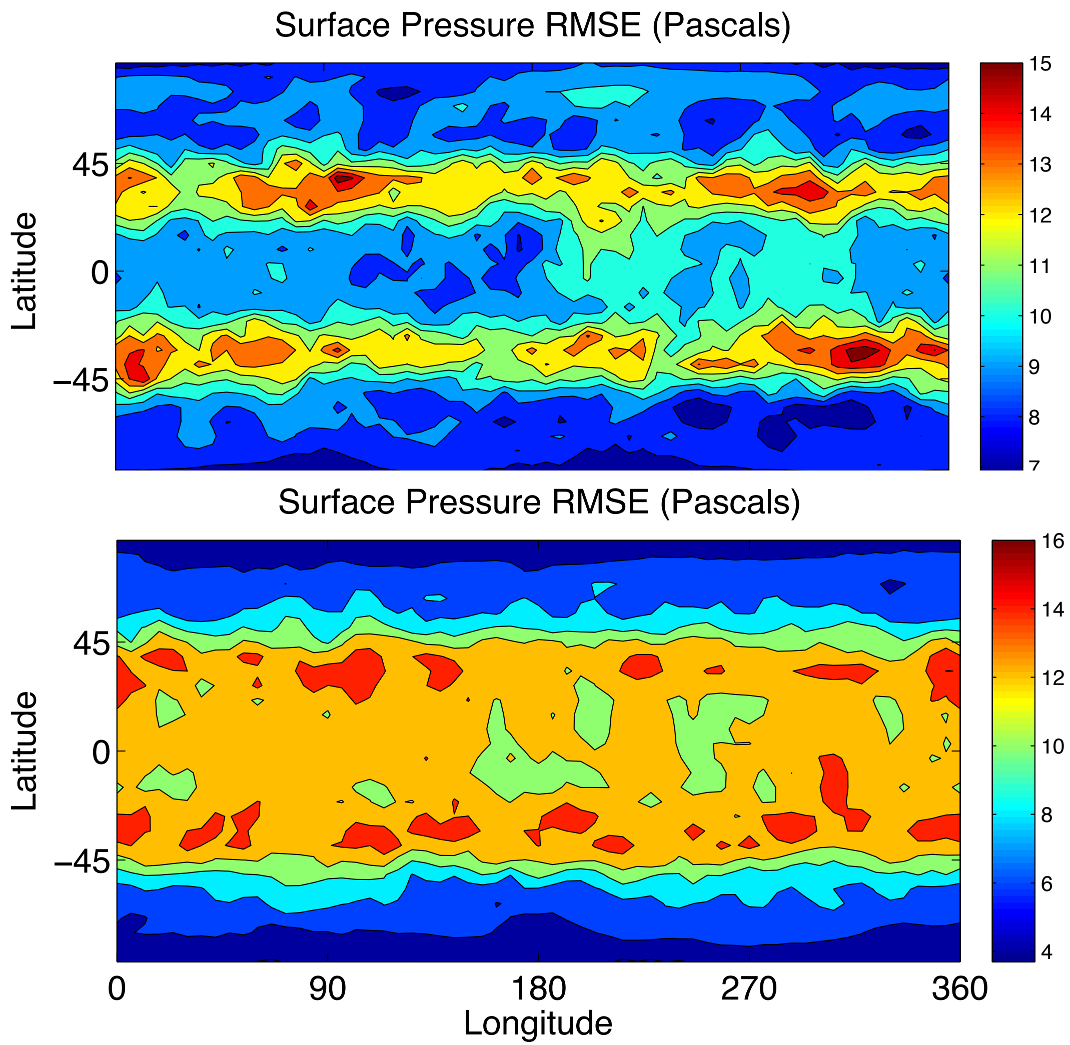
PS RMSE Structure: Large Uncorrelated Error, Ratio 4



Base errors largest in storm tracks.

Linked difference errors largest in broad tropical band.

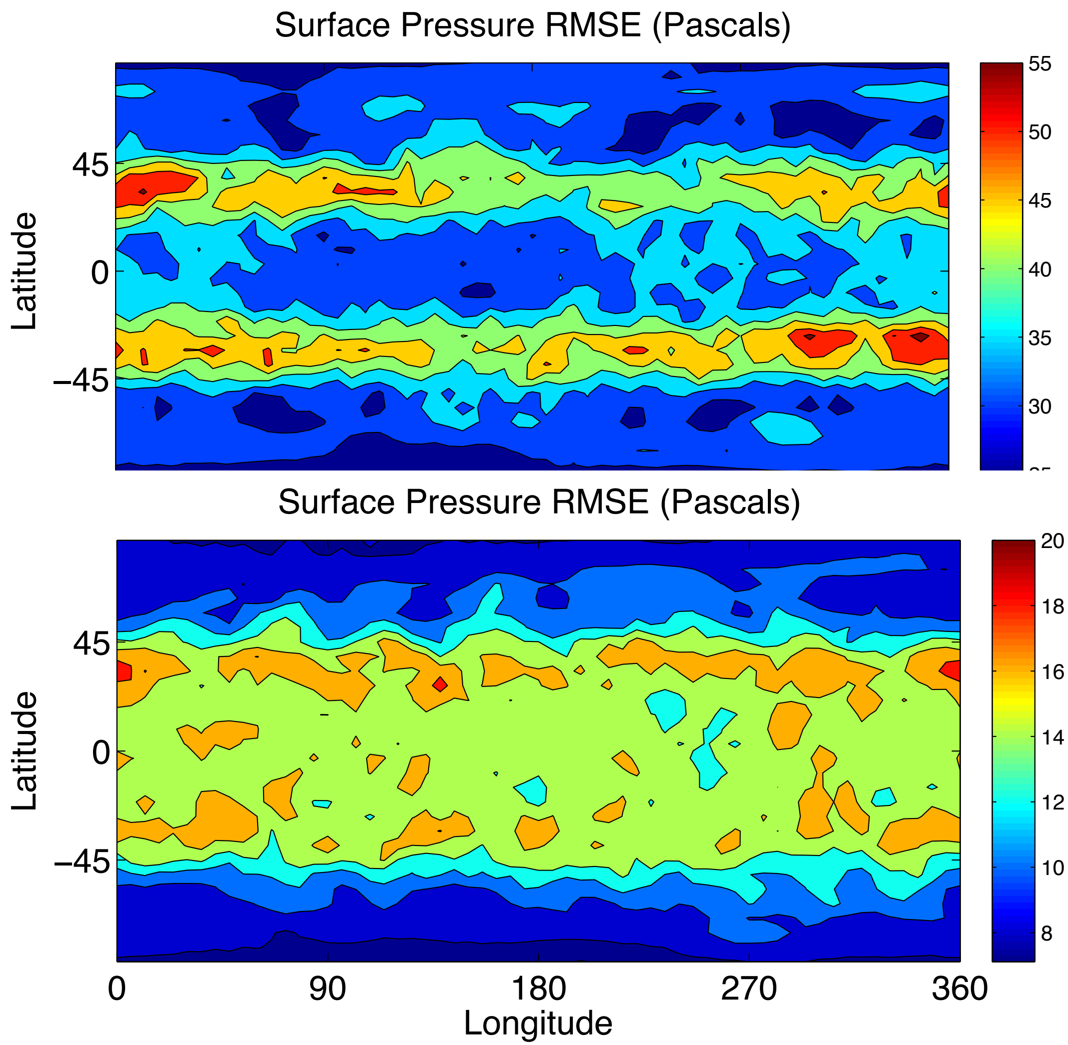
PS RMSE Structure: Moderate Uncorrelated Error, Ratio 1



Base errors largest in storm tracks.

Linked difference errors largest in broad tropical band.

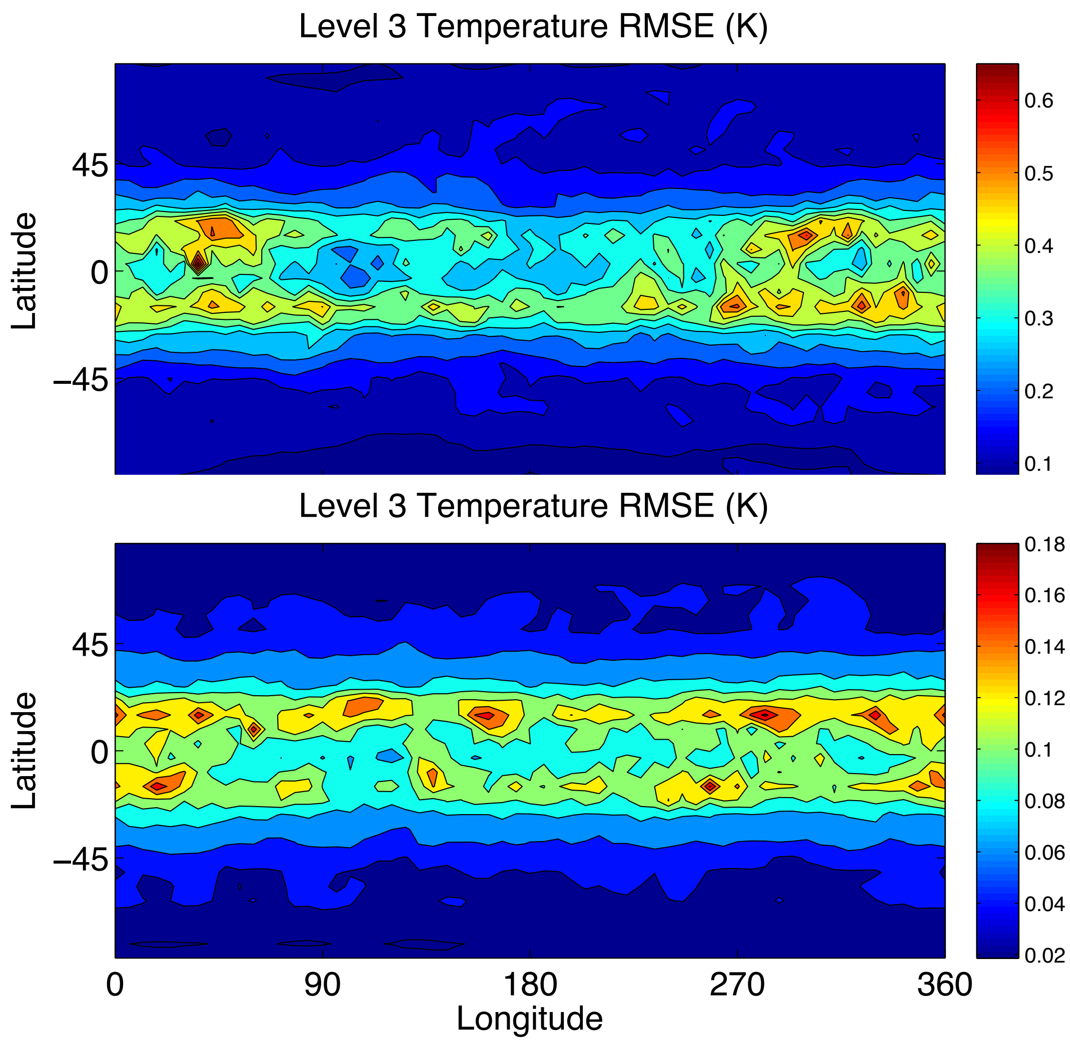
PS RMSE Structure: Small Uncorrelated Error, Ratio 1/4



Base errors largest in storm tracks.

Linked difference errors largest in broad tropical band.

T RMSE Structure: Small Uncorrelated Error, Ratio 1/4



Base errors largest in tropics.

Linked difference errors have similar pattern.

Low-Order Dry Dynamical Core Summary

- Linked difference obs better for large correlated error.
- Linked difference not sensitive to correlated error size.
- Adaptive inflation struggles with large correlated error.
- Could use base approach for uncorrelated obs, difference for correlated error obs.
- For example, base for sondes, difference for radiances.
- Difference obs allows assimilating before knowing correlated error characteristics.

Outline

Dealing with correlated observation error in ensemble filters.

1. Idealized correlated error.
2. Difference observations.
3. Explicitly modeling instrument error.
4. Comparing the two methods.
5. Conclusions and recommendations.

Modeling Correlated Observation Error

Error in examples is AR1:

(other types may need other methods).

Given correlated error now, can predict it at later time.

Have ensemble of model state.

Also ensemble of correlated error for each instrument.

Modeling Correlated Observation Error

1. Forecast: Advance model & correlated error ensembles.
2. Forward operator (for each ensemble member):

   Apply standard forward operator to state, H(x),

   Add correlated error.
3. Observation Increments: Compute normally.
4. State variable update:

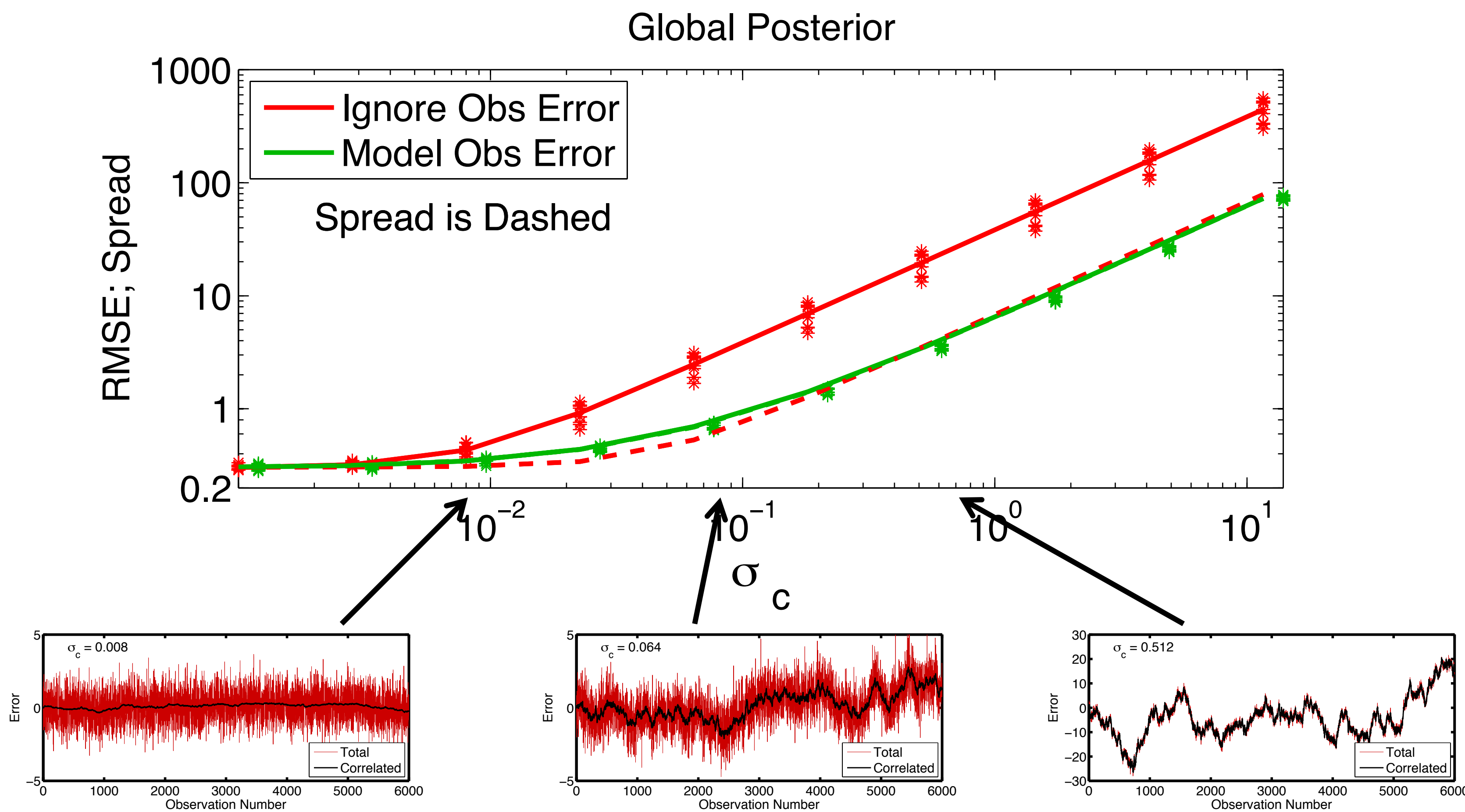
   Use regression (ensemble Kalman gain) to update:

   Model state variables,

   Correlated observation variables.

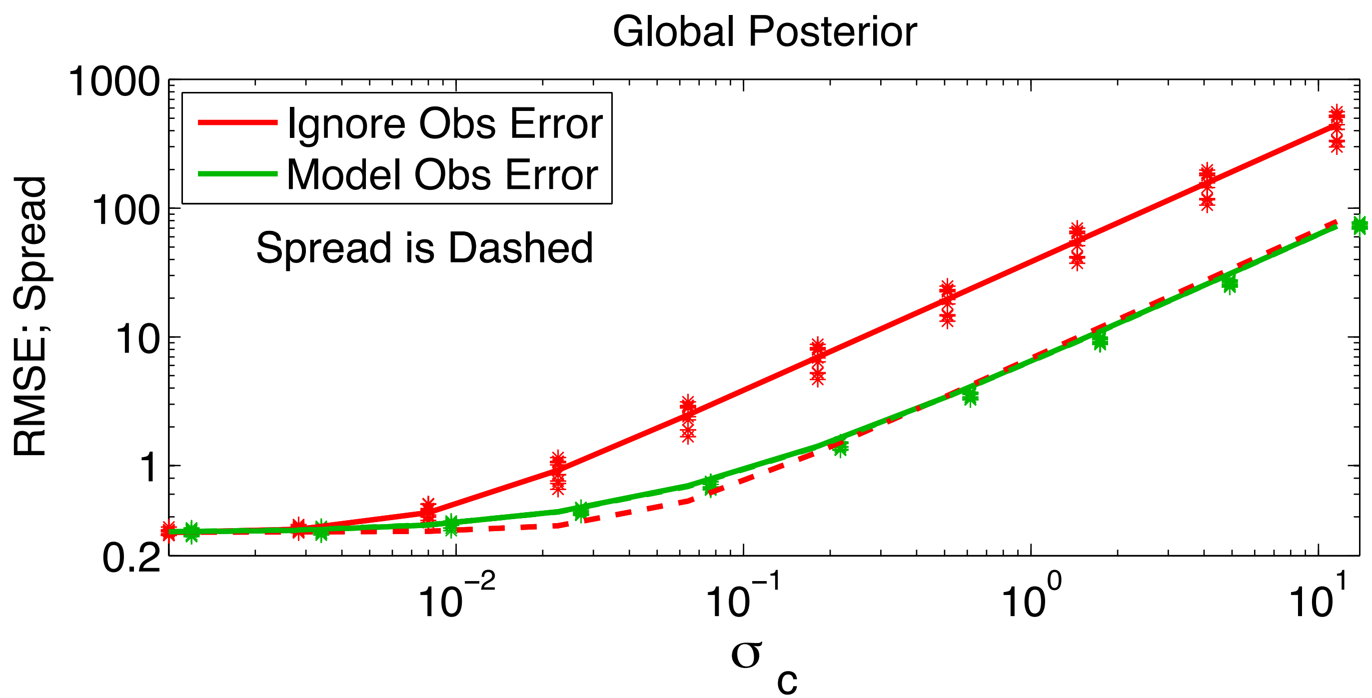
1D Exponential Growth Model Results

320 Member deterministic ensemble filter (EAKF) State



1D Exponential Growth Model Results

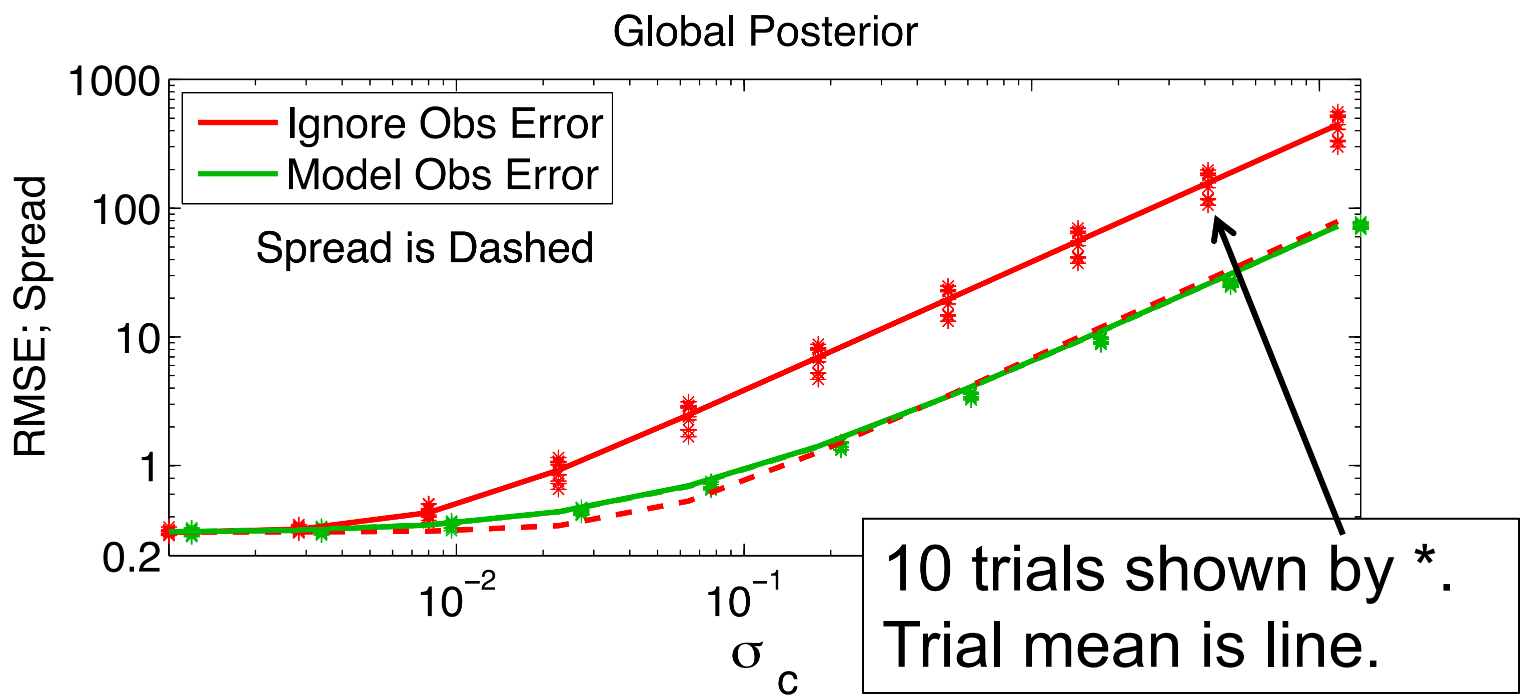
320 Member deterministic ensemble filter (EAKF) State



All results for 5000 steps after 1000 step spin-up.

1D Exponential Growth Model Results

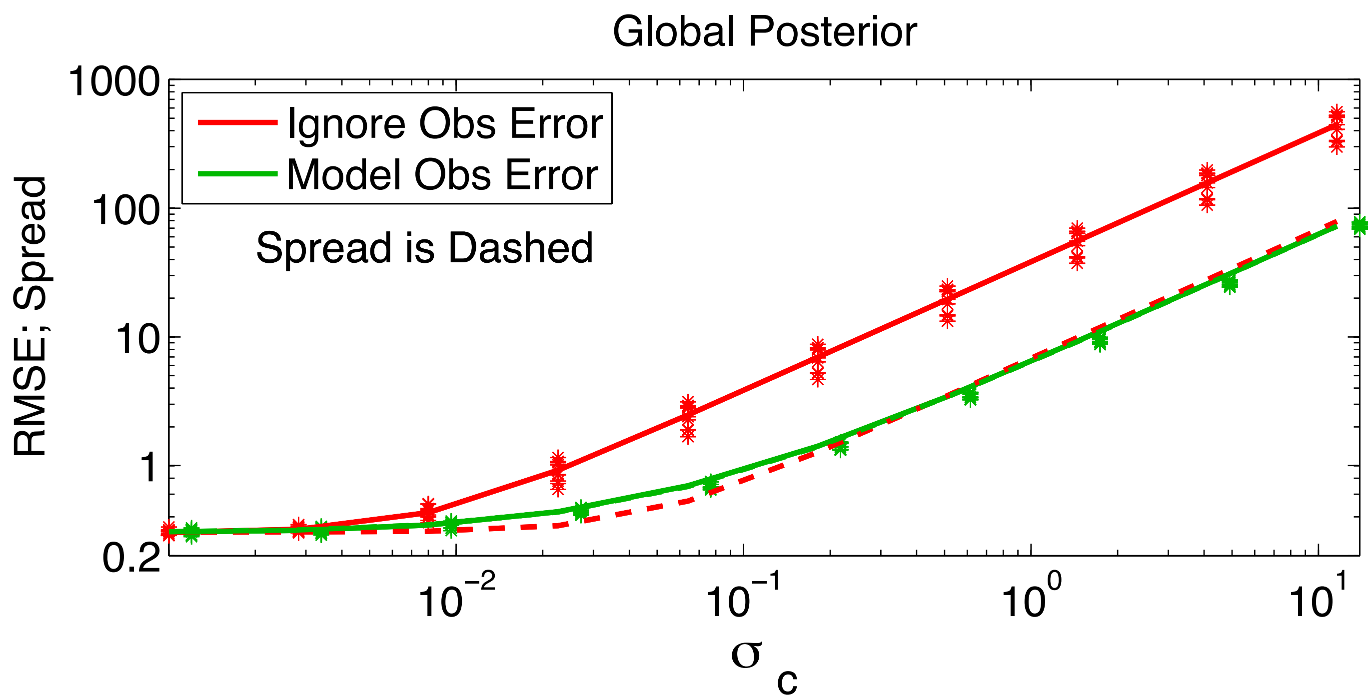
320 Member deterministic ensemble filter (EAKF) State



All results for 5000 steps after 1000 step spin-up.

1D Exponential Growth Model Results

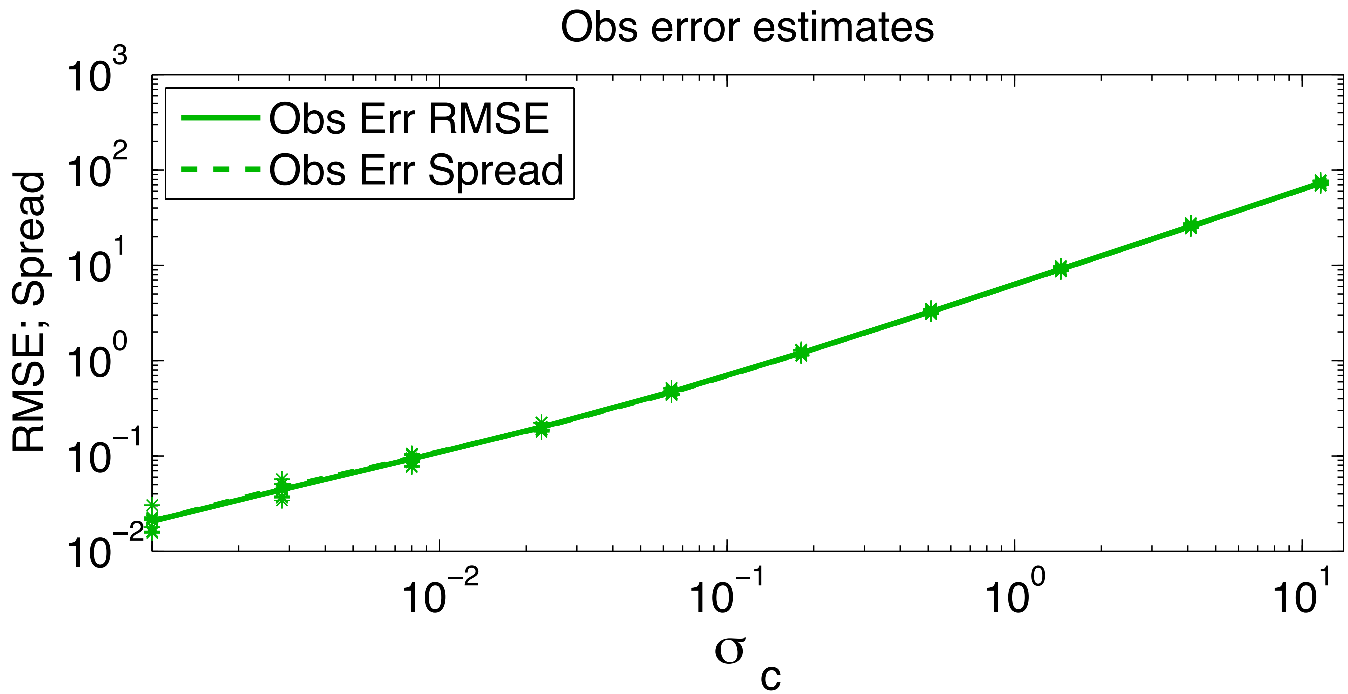
320 Member deterministic ensemble filter (EAKF) State



Exact asymptotic solution can be computed.
Indistinguishable from 320 member ensemble.

1D Exponential Growth Model Results

320 Member deterministic ensemble filter (EAKF) Obs. Error

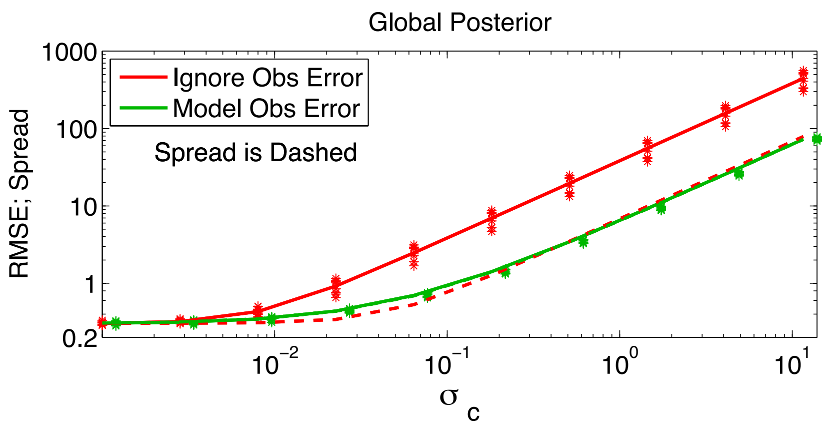


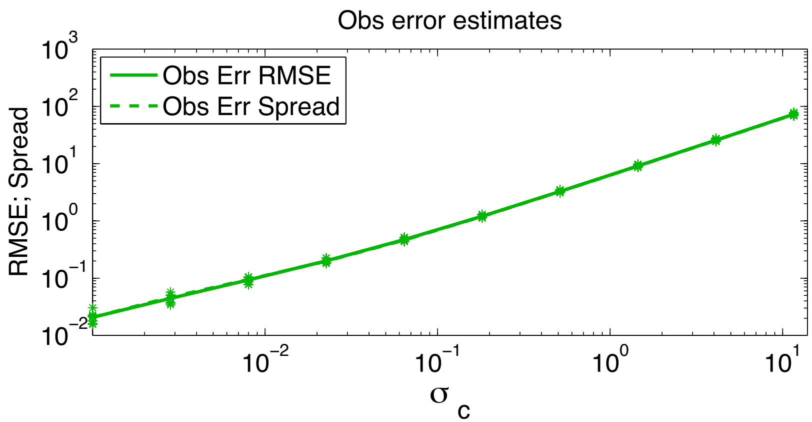
Exact asymptotic solution can be computed.
Indistinguishable from 320 member ensemble.

1D Exponential Growth Model Results

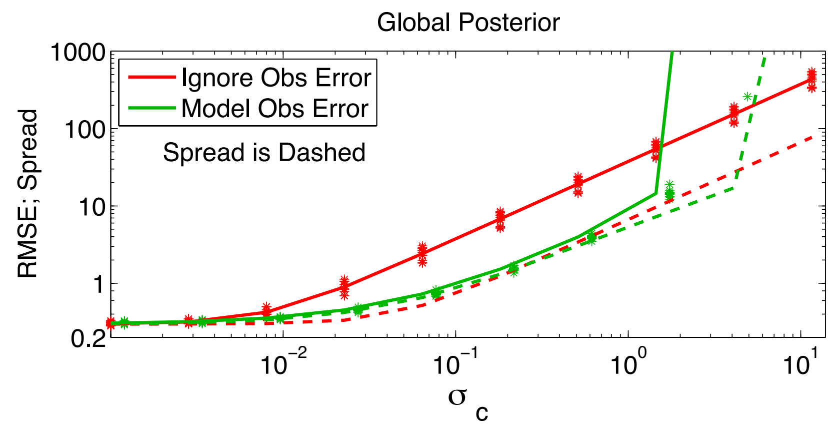
Fails for small ensembles with large correlated error.

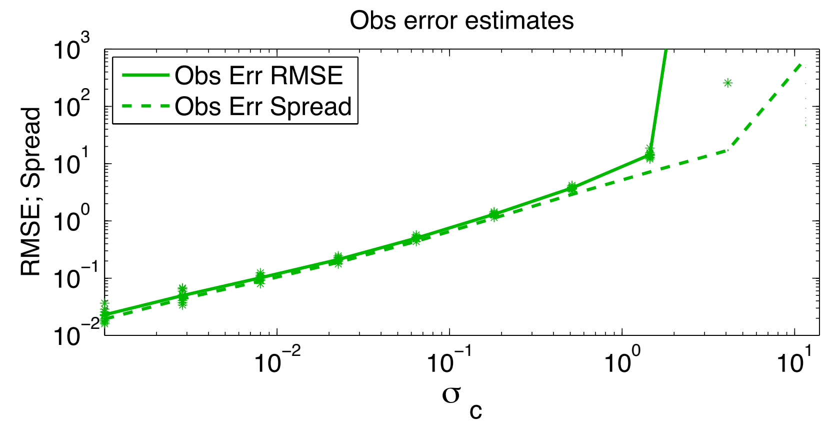
320 Member EAKF





20 Member EAKF

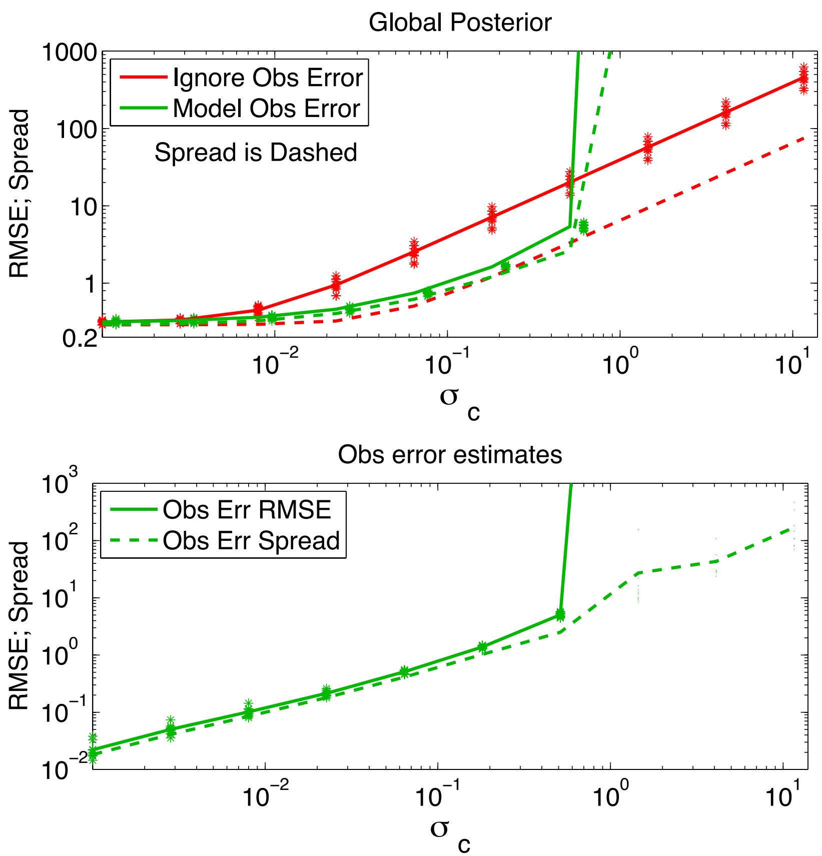




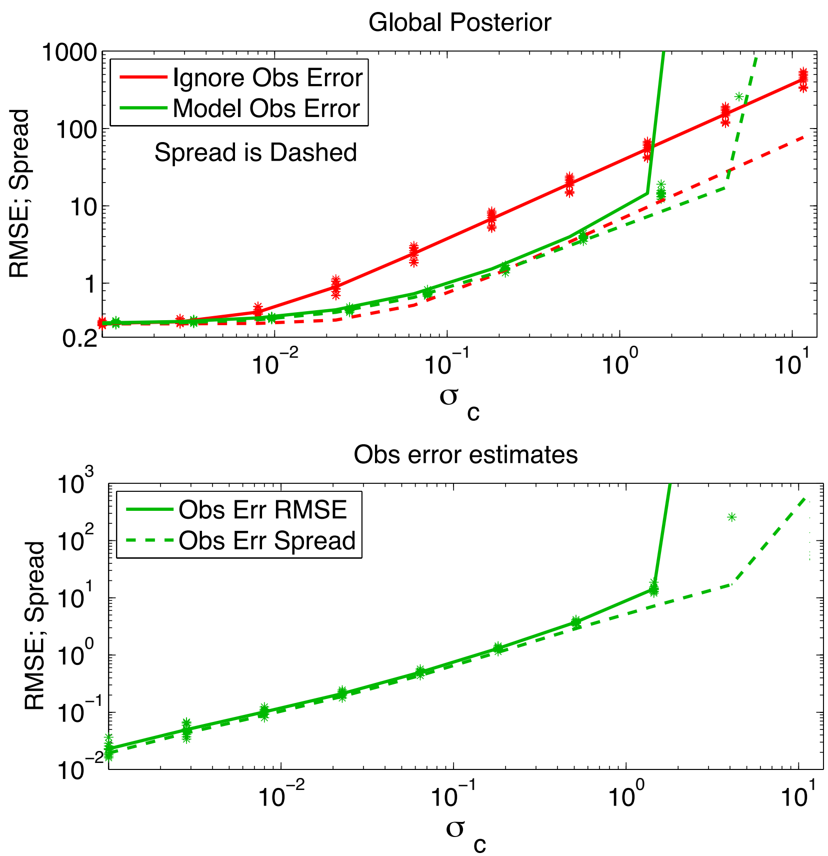
1D Exponential Growth Model Results

Fails for small ensembles with large correlated error.

10 Member EAKF



20 Member EAKF



Ensemble Filters Scale Poorly for Random Fields

Ensemble size > 1 is exact with no correlated obs error.

Random walk evolution of correlated error is a problem.

Can reduce this by reducing ‘randomness’ of ensemble.

AR1 series for observation error is: $e_t = \phi e_{t-1} + Normal\left(0, \sigma_c^2\right)$

Given a posterior ensemble estimate of *e* at previous time:

Expected prior mean at next time is: $E\left(e_p\right) = \phi E\left(e_u\right)$

Expected prior variance is: $E\left[\mathrm{var}\left(e_p\right)\right] = \phi^2 E\left[\mathrm{var}\left(e_u\right)\right] + \sigma_c^2$

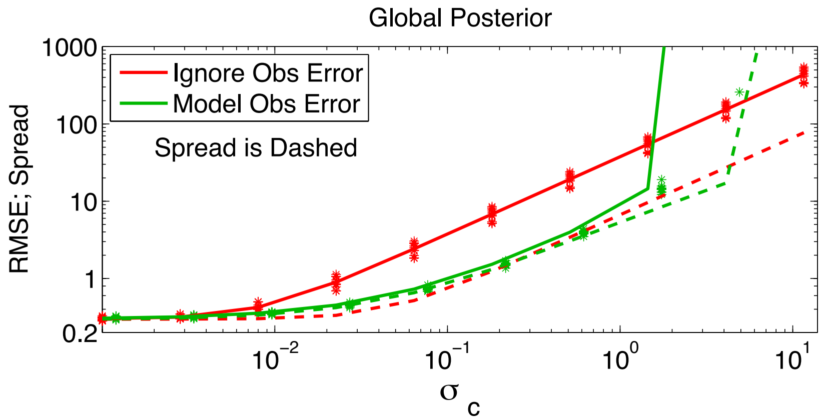
‘Deterministic’ forecast for observation error:

‘Adjust’ ensemble to have exactly these statistics.

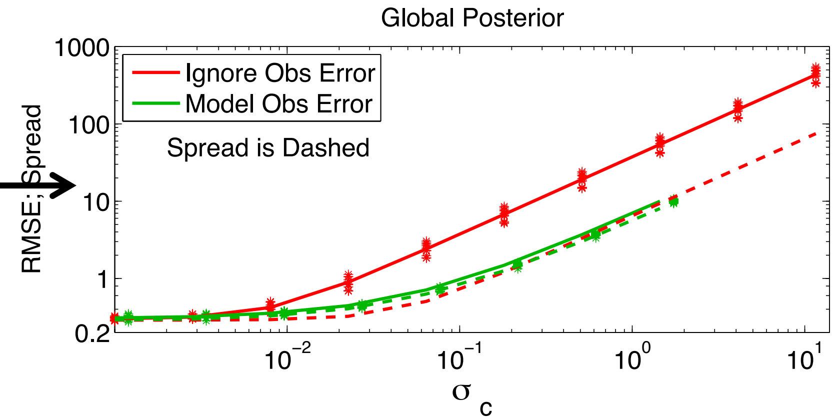
1D Exponential Growth Model Results

Deterministic works with smaller ensembles. Used hereafter.

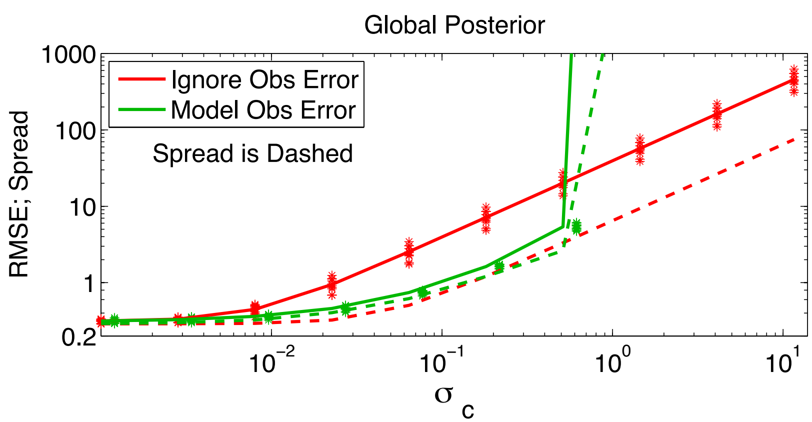
Nondeterm 20 Member



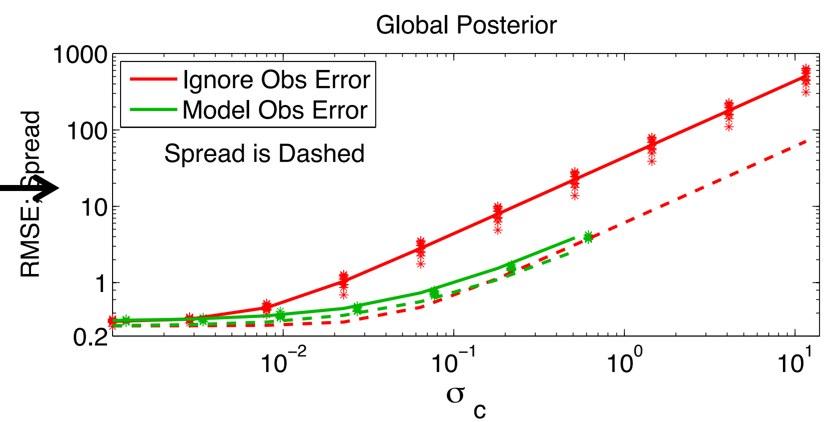
Determ 10 Member



Nondeterm 10 Member



Determ 5 Member

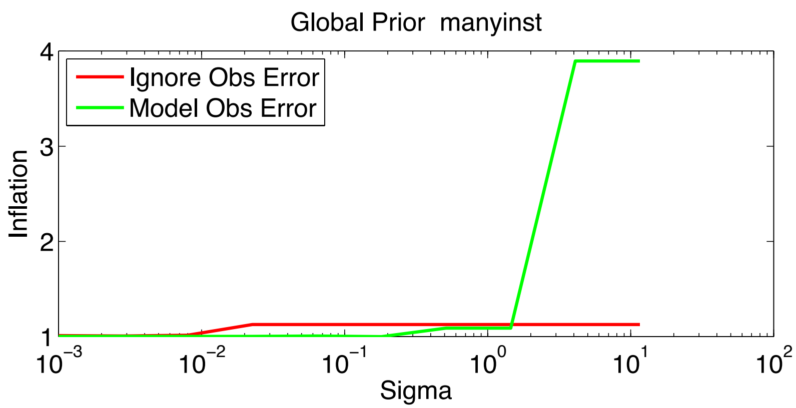


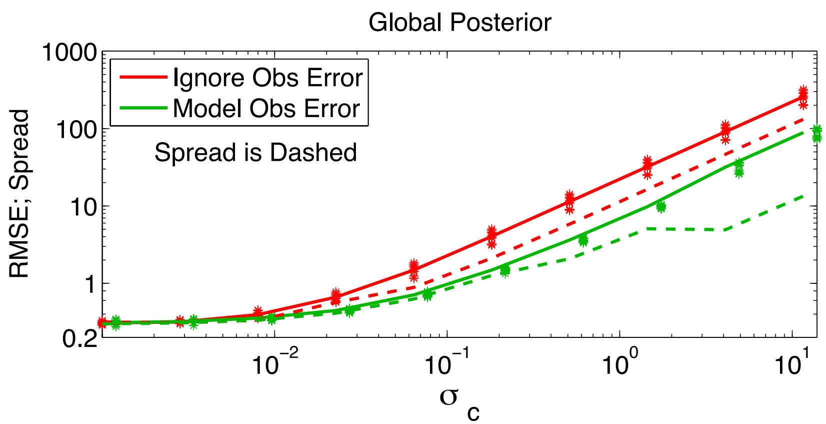
1D Exponential Growth Model Results

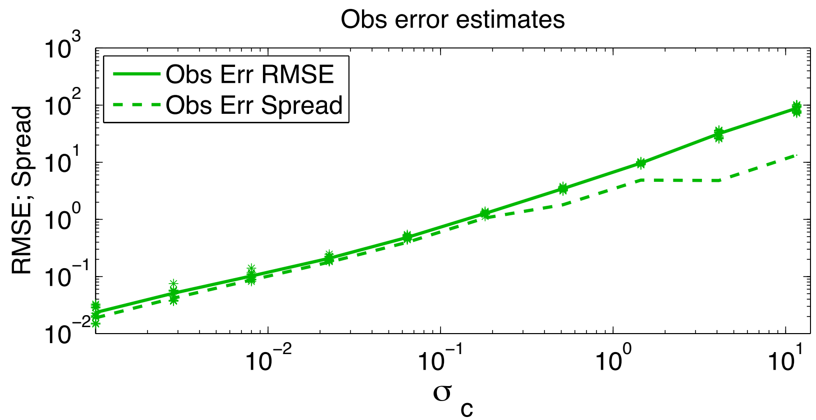
Try multiplicative inflation of state.

10 Member inflated

Optimal inflation gets large.





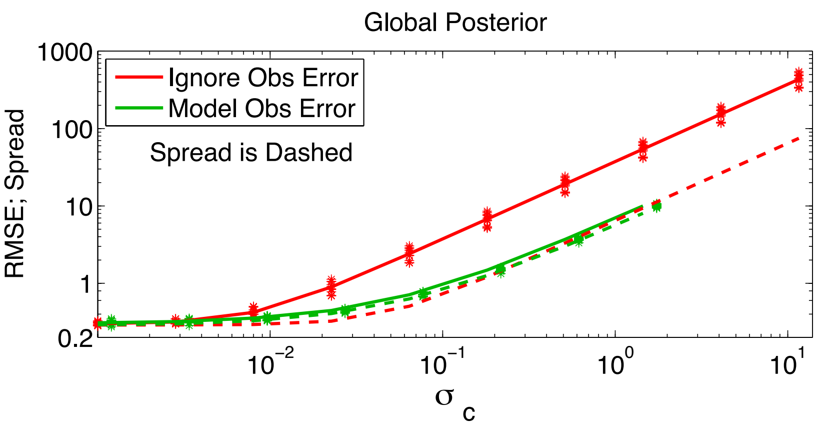


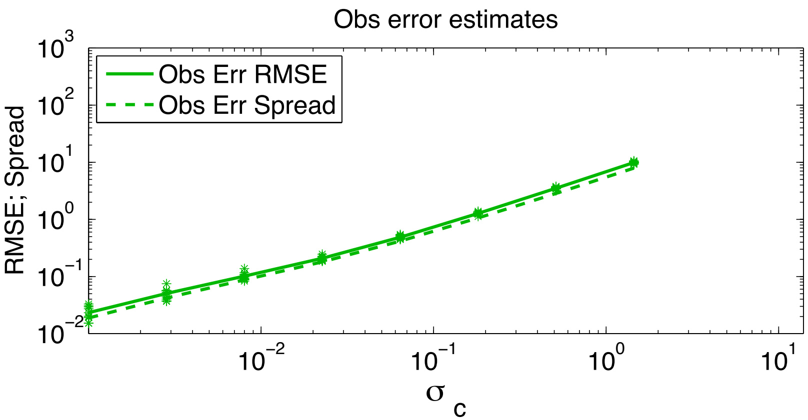
Multiplicative inflation for obs error ensemble is bad.

1D Exponential Growth Model Results

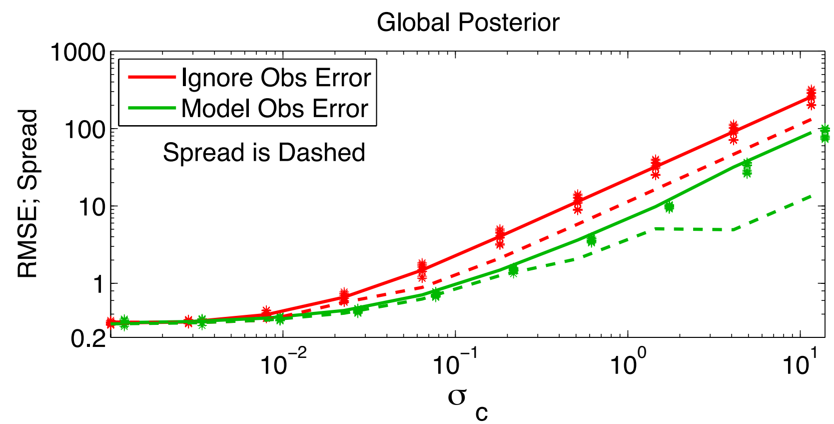
Multiplicative inflation for state improves performance.

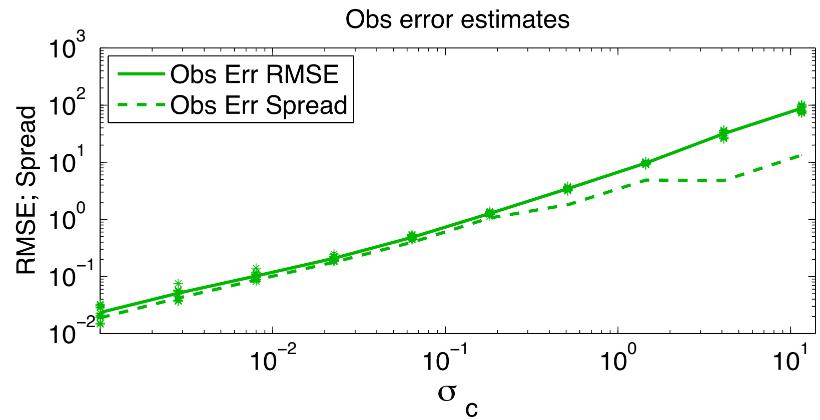
10 Member



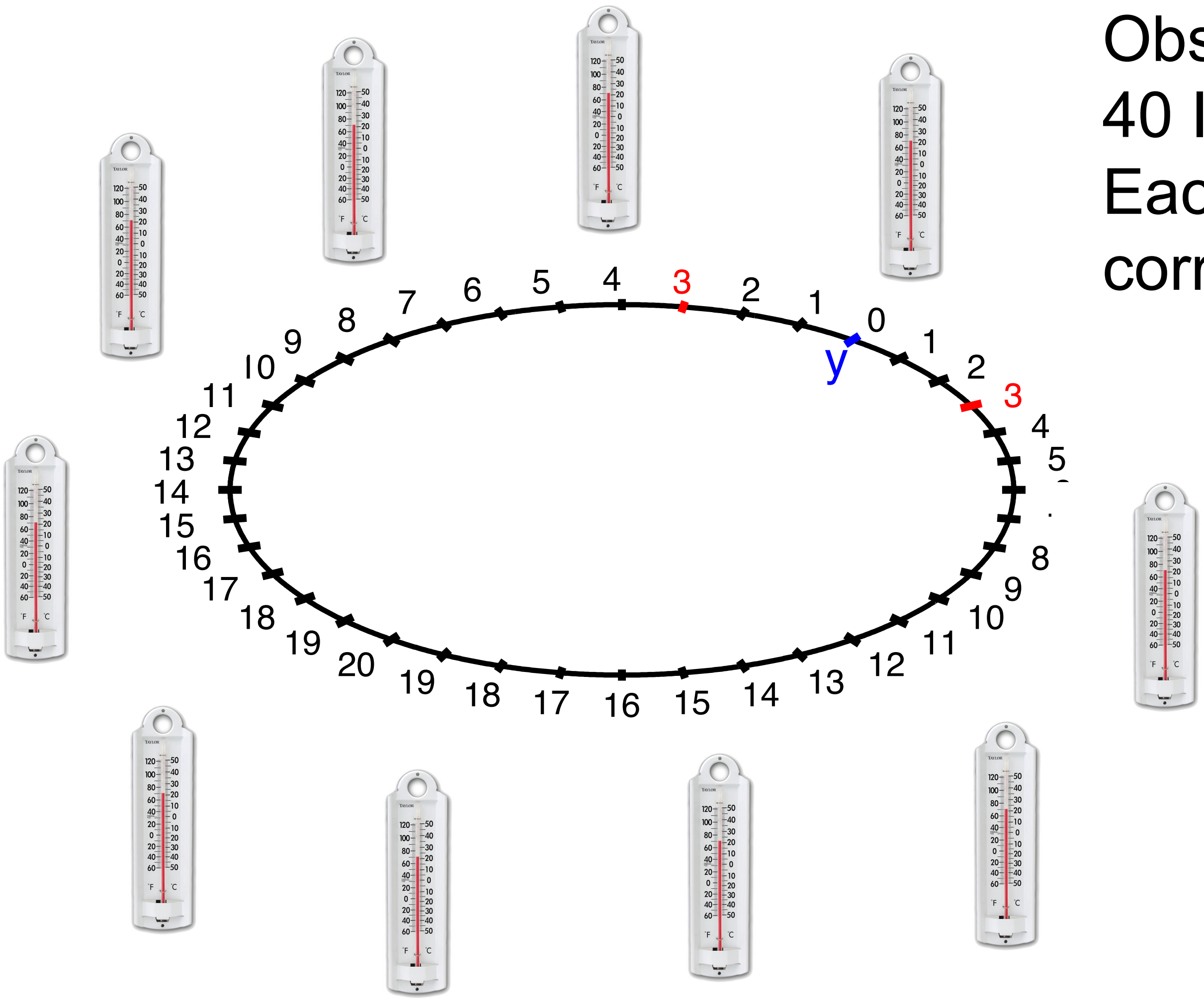


10 Member inflated



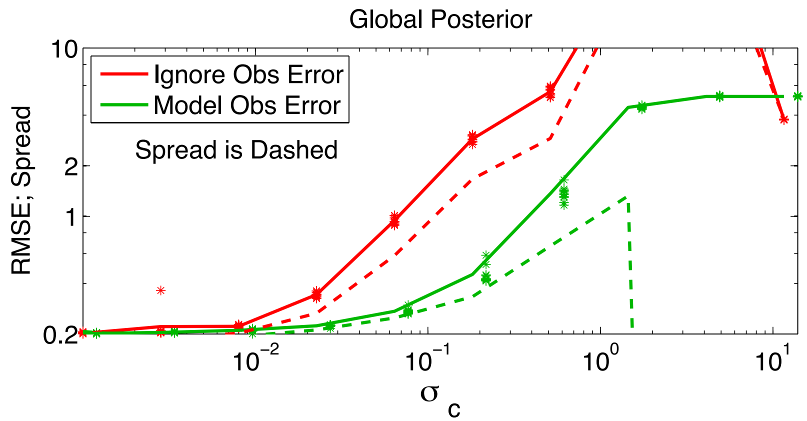


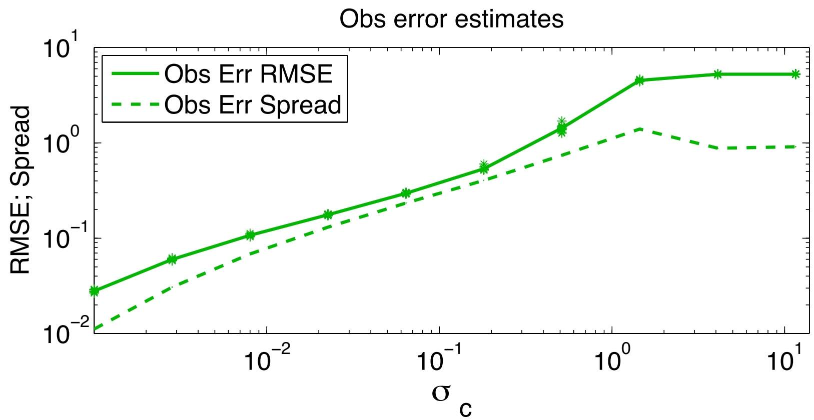
Lorenz 96 Model, 40-variables



Observing System 1:
40 Instruments.
Each has own
correlated error.

Lorenz 96 Model, 40-instruments





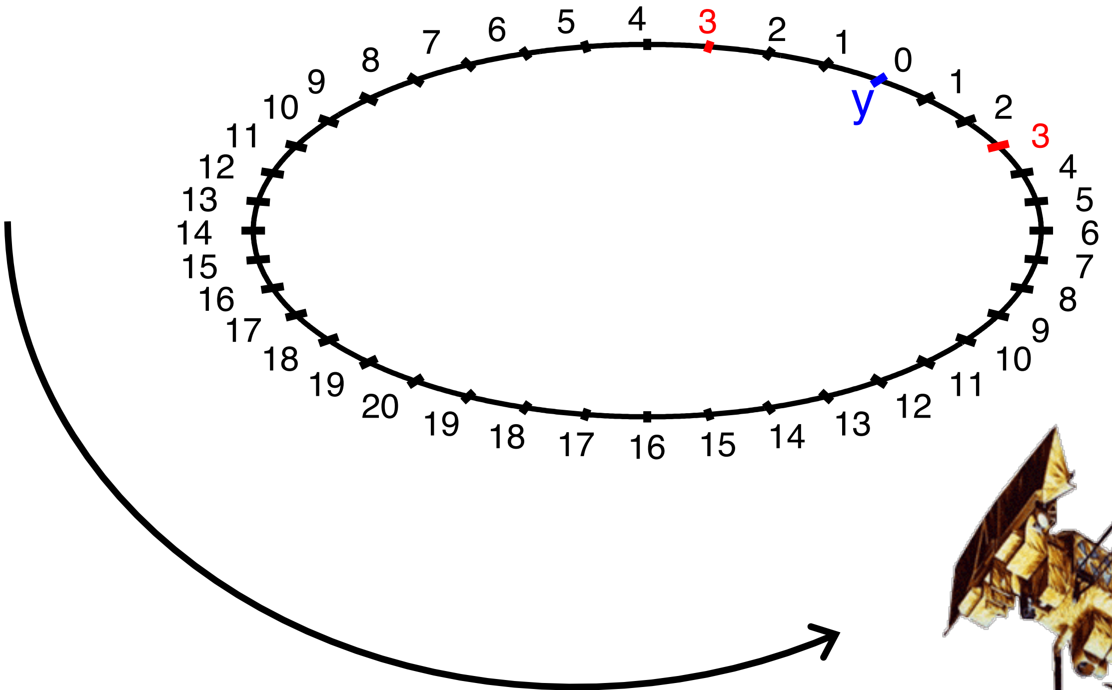
20 member EAKF.

Optimal inflation.

Localization halfwidth 0.2

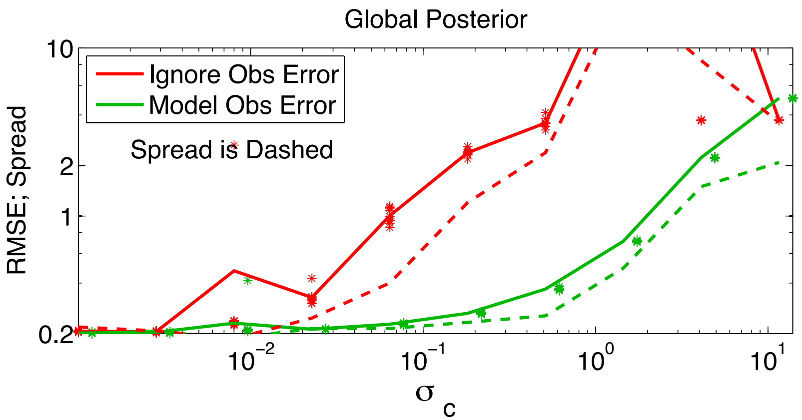
Modeling obs error helps.
Spread is deficient.

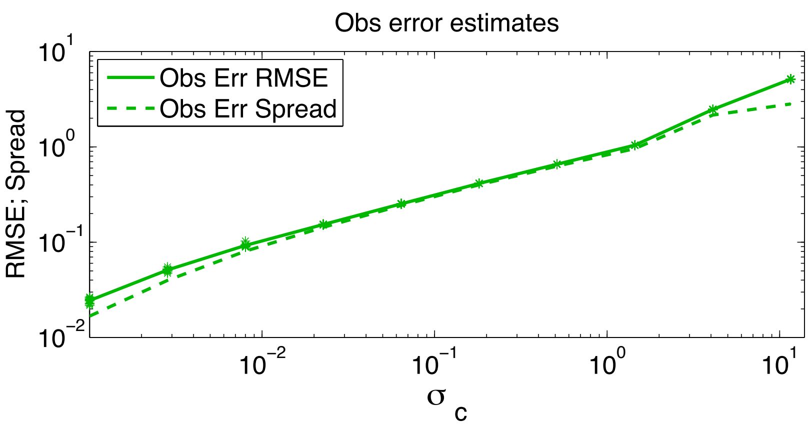
Lorenz 96 Model, 40-variables



Observing System 2:
1 instrument measures all 40 variables each time.

Lorenz 96 Model, 1-instrument





20 member EAKF.

Optimal inflation.

Localization halfwidth 0.2

Modeling obs error helps.
Spread is better than many instrument case.

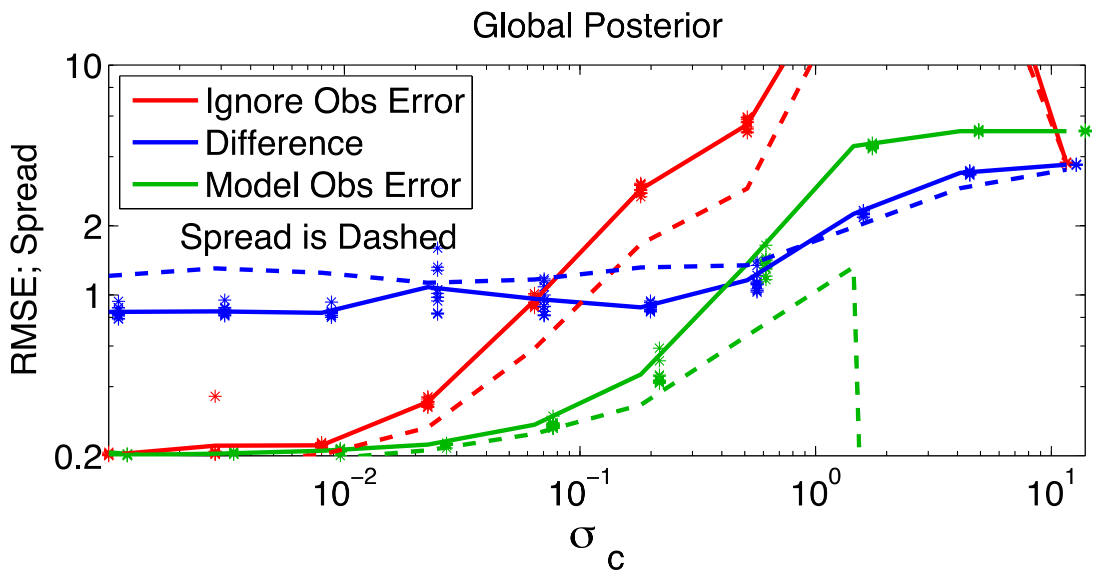
Outline

Dealing with correlated observation error in ensemble filters.

1. Idealized correlated error.
2. Difference observations.
3. Explictly modeling instrument error.
4. Comparing the two methods.
5. Conclusions and recommendations.

Lorenz 96 Model, 40-instruments

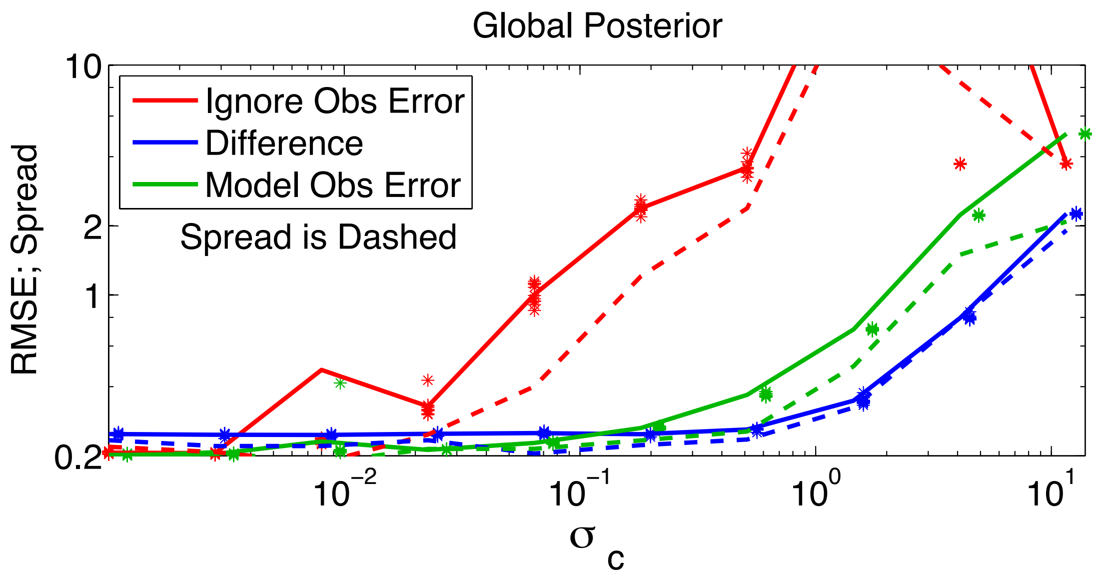
Time difference assimilation best for large correlated error.
Terrible for small correlated error.



20 member EAKF.
Optimal inflation.
Localization halfwidth 0.2

Lorenz 96 Model, 1-instrument

Time difference assimilation best for large correlated error.
Not bad for small correlated error.



20 member EAKF.
Optimal inflation.
Localization halfwidth 0.2

Outline

Dealing with correlated observation error in ensemble filters.

1. Idealized correlated error.
2. Difference observations.
3. Explictly modeling instrument error.
4. Comparing the two methods.
5. Conclusions and recommendations.

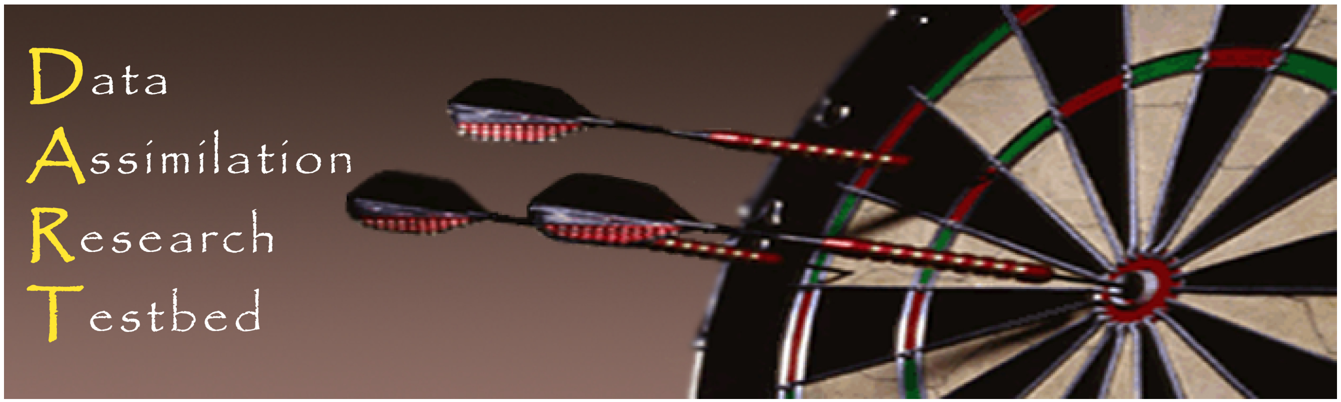
Conclusions

- Modeling correlated obs error ‘optimal’ for large ensemble.
- Sampling error is a problem for small ensembles.
- Multiplicative state inflation can reduce this problem.
- Additive inflation for obs error may help?
- Time difference obs effective for large correlated error.

General things to keep in mind:

- Details of filtering problem determine best methods.
- Making models/filters more deterministic generally helps.

Learn more about DART at:



www.image.ucar.edu/DAReS/DART

Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., Arellano, A., 2009: *The Data Assimilation Research Testbed: A community facility.* BAMS, **90**, 1283—1296, doi: 10.1175/2009BAMS2618.1