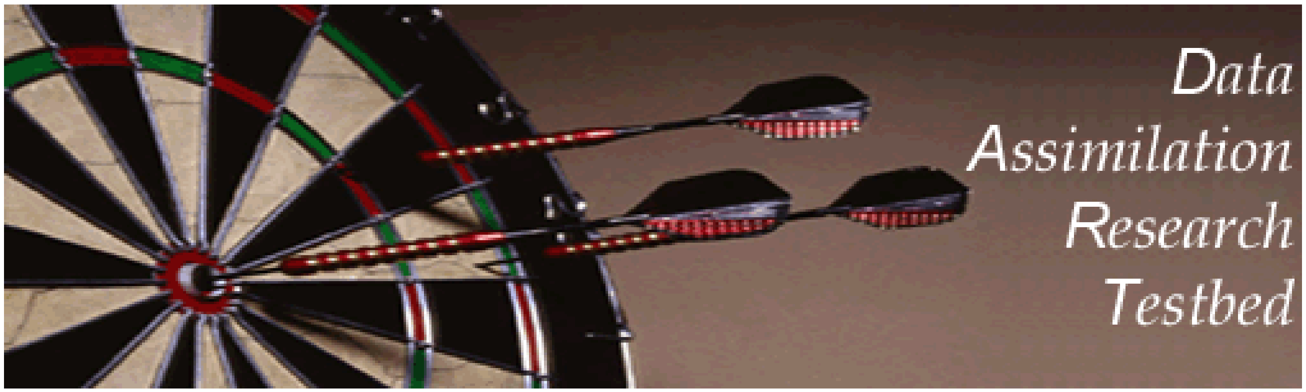
Using Small Ensembles in High Dimensions: Hierarchical Bayesian Approaches to Adaptive Ensemble Filters



Jeffrey Anderson
IMAGe Data Assimilation Research Section (DAReS)

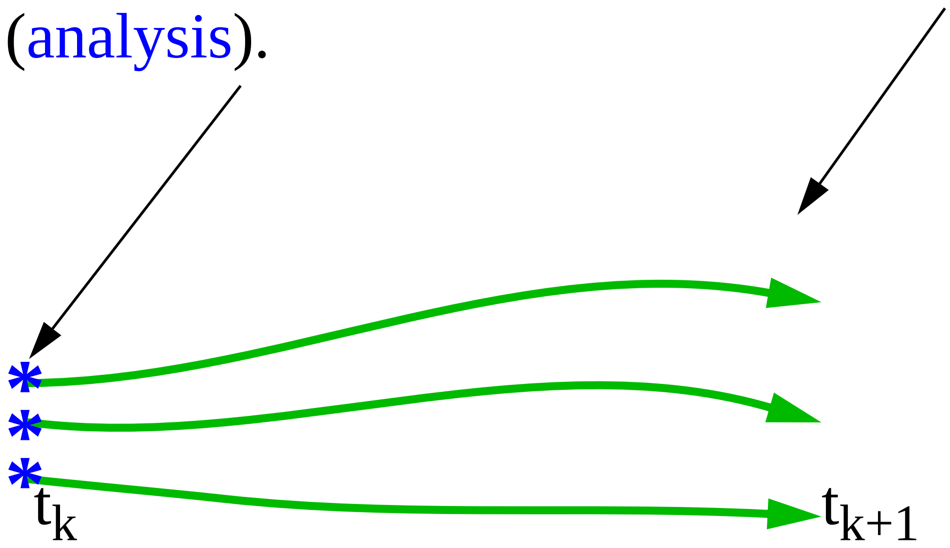
Thanks to Nancy Collins, Tim Hoar, Hui Liu, Kevin Raeder

How an Ensemble Filter Works for Geophysical Data Assimilation

1. Use model to advance ensemble (3 members here) to time at which next observation becomes available.

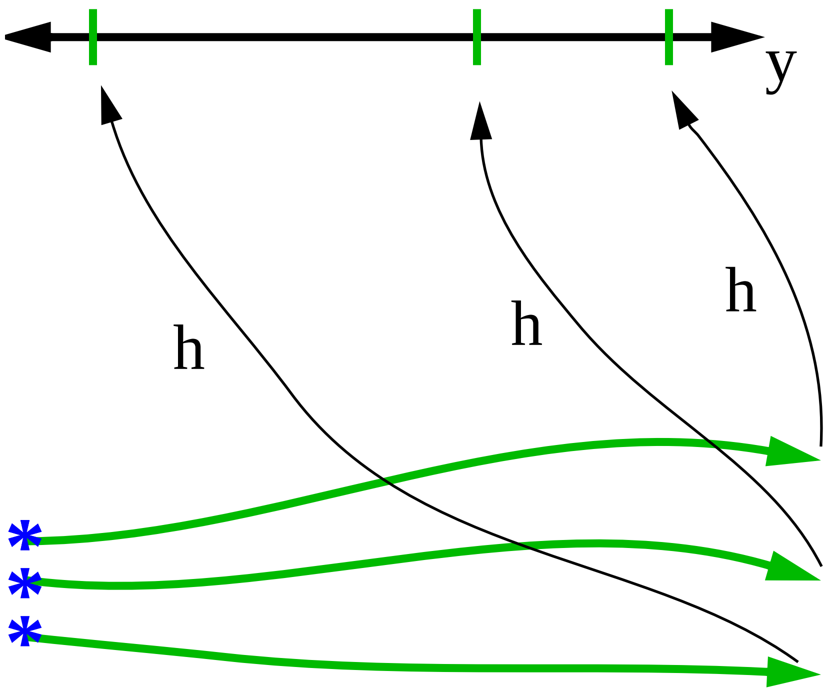
Ensemble state estimate after using previous observation (analysis).

Ensemble state at time of next observation (prior).



How an Ensemble Filter Works for Geophysical Data Assimilation

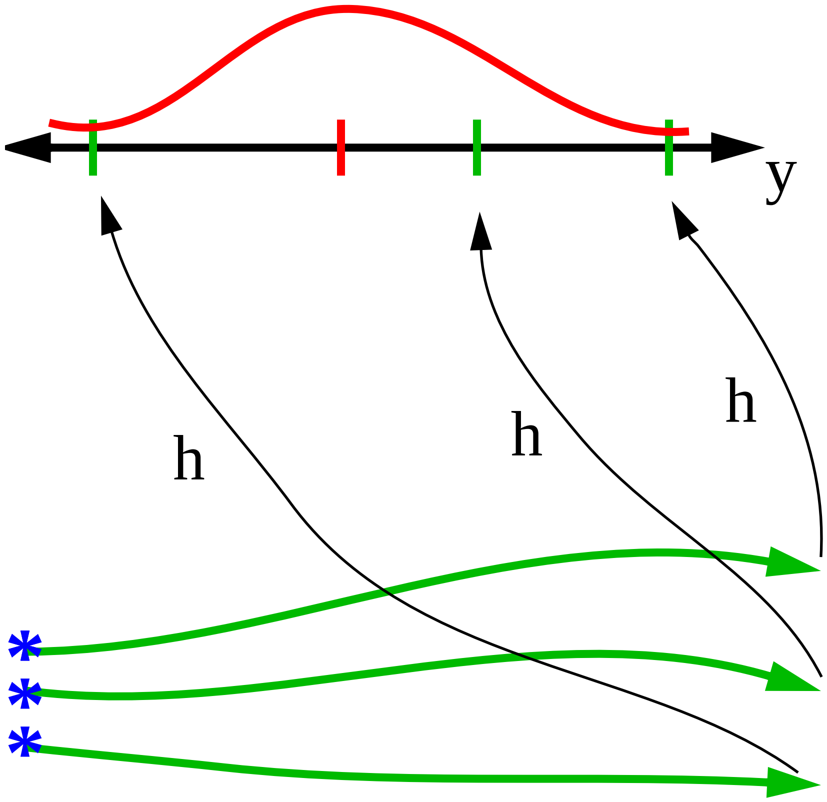
2. Get prior ensemble sample of observation, y=h(x), by applying forward operator h to each ensemble member.



Theory: observations from instruments with uncorrelated errors can be done sequentially.

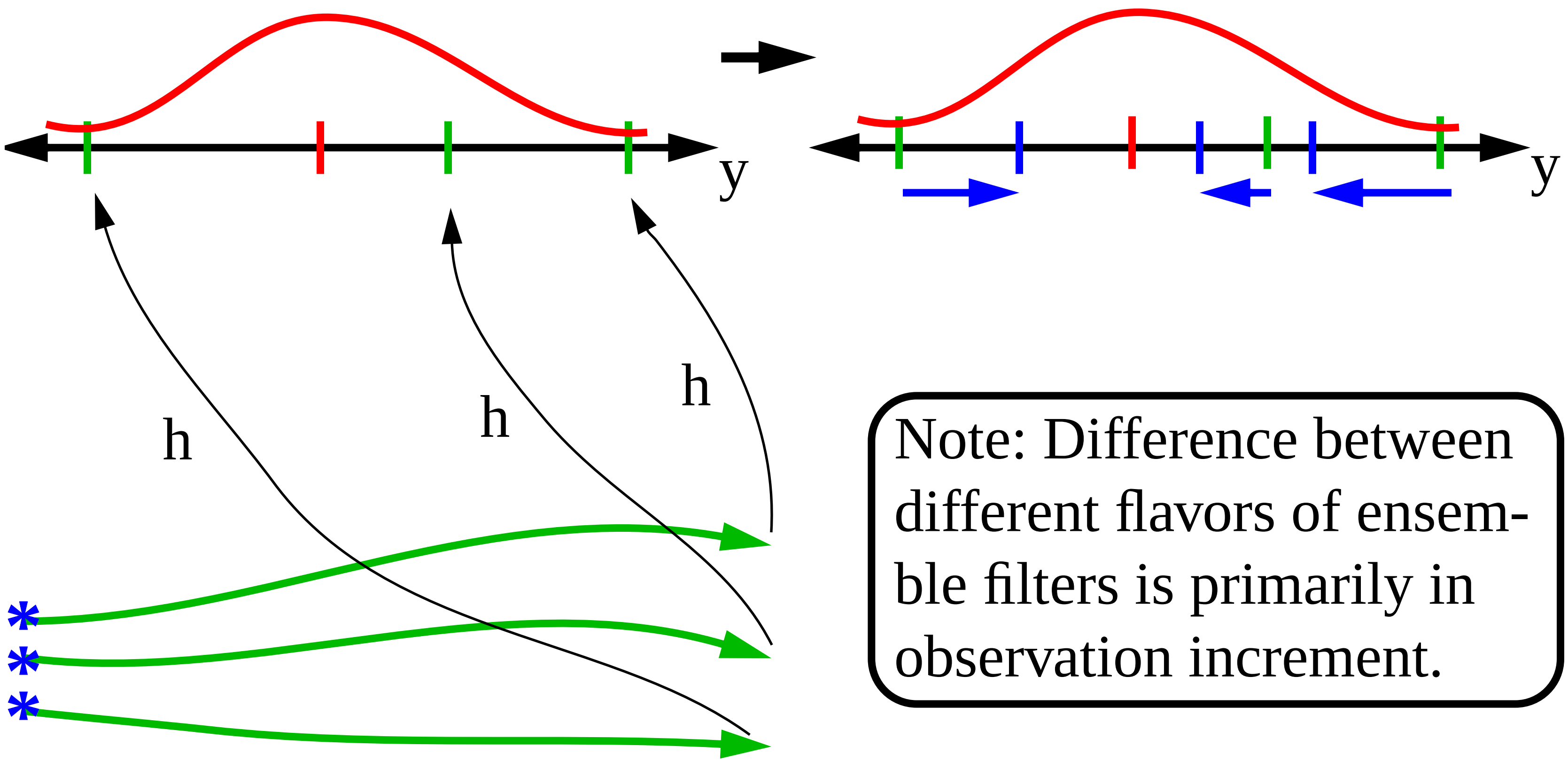
How an Ensemble Filter Works for Geophysical Data Assimilation

3. Get observed value and observational error distribution from observing system.



How an Ensemble Filter Works for Geophysical Data Assimilation

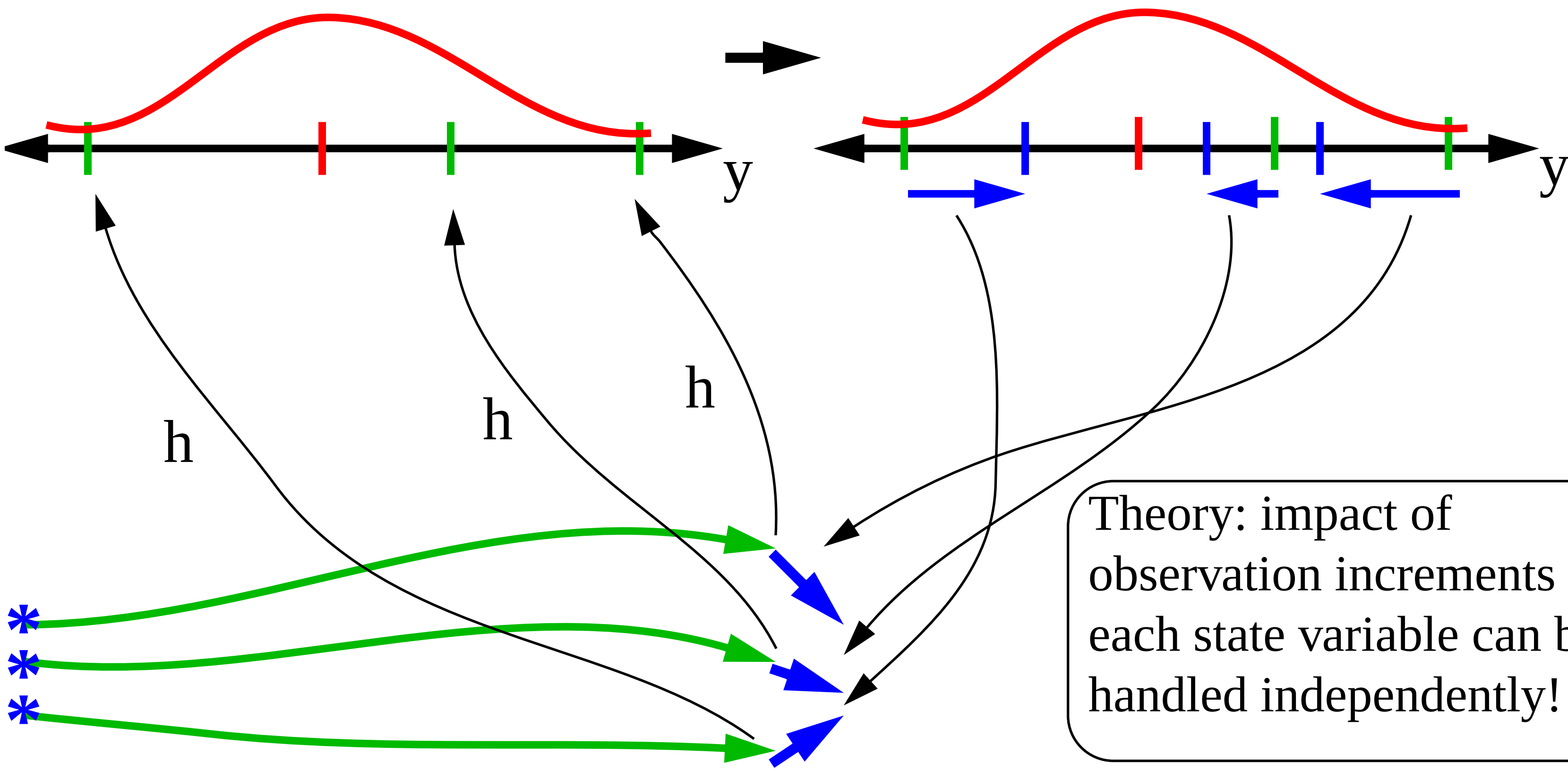
4. Find increment for each prior observation ensemble
(this is a scalar problem for uncorrelated observation errors).



Note: Difference between different flavors of ensemble filters is primarily in observation increment.

How an Ensemble Filter Works for Geophysical Data Assimilation

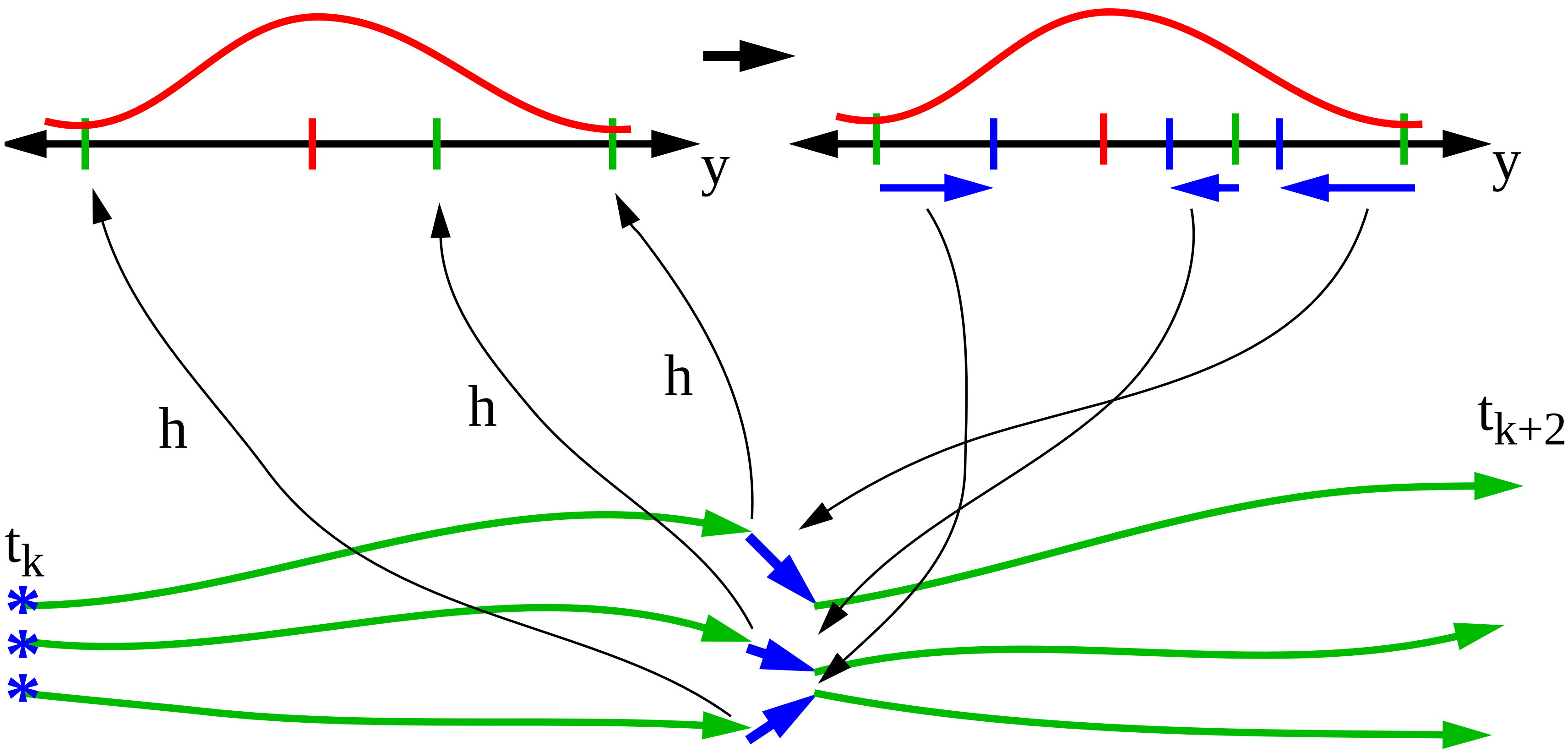
5. Use ensemble samples of y and each state variable to linearly regress observation increments onto state variable increments.



Theory: impact of observation increments on each state variable can be handled independently!

How an Ensemble Filter Works for Geophysical Data Assimilation

6. When all ensemble members for each state variable are updated, have a new analysis. Integrate to time of next observation...

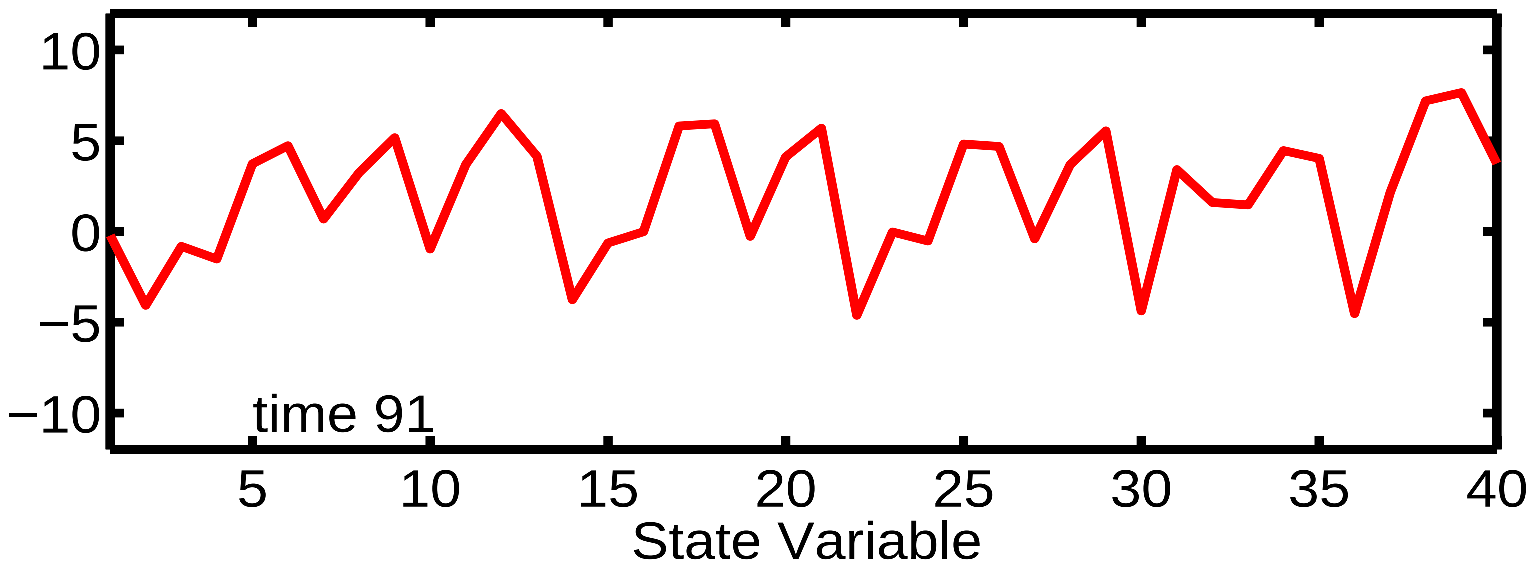


Ensemble Filter for Lorenz-96 40-Variable Model

40 state variables: X_1, X_2,..., $X_{40.}$

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

Acts ‘something’ like synoptic weather around a latitude band.

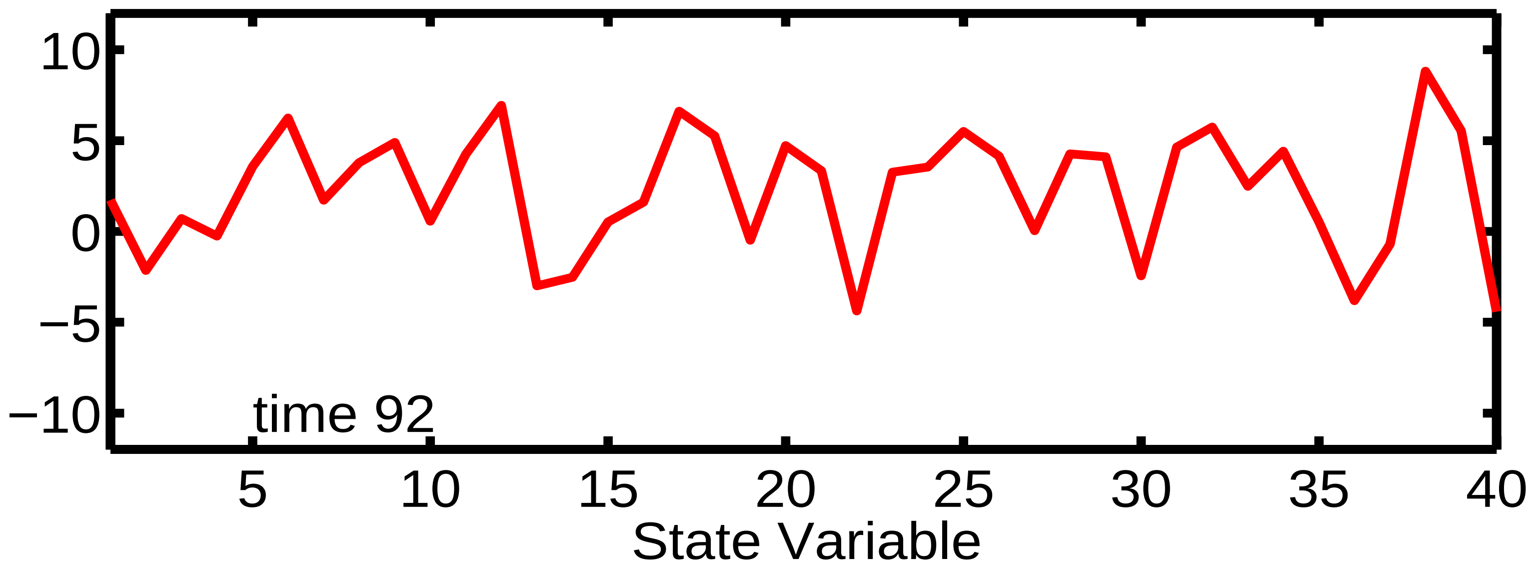


Ensemble Filter for Lorenz-96 40-Variable Model

40 state variables: $X_1, X_2, \ldots, X_{40.}$

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

Acts 'something' like synoptic weather around a latitude band.

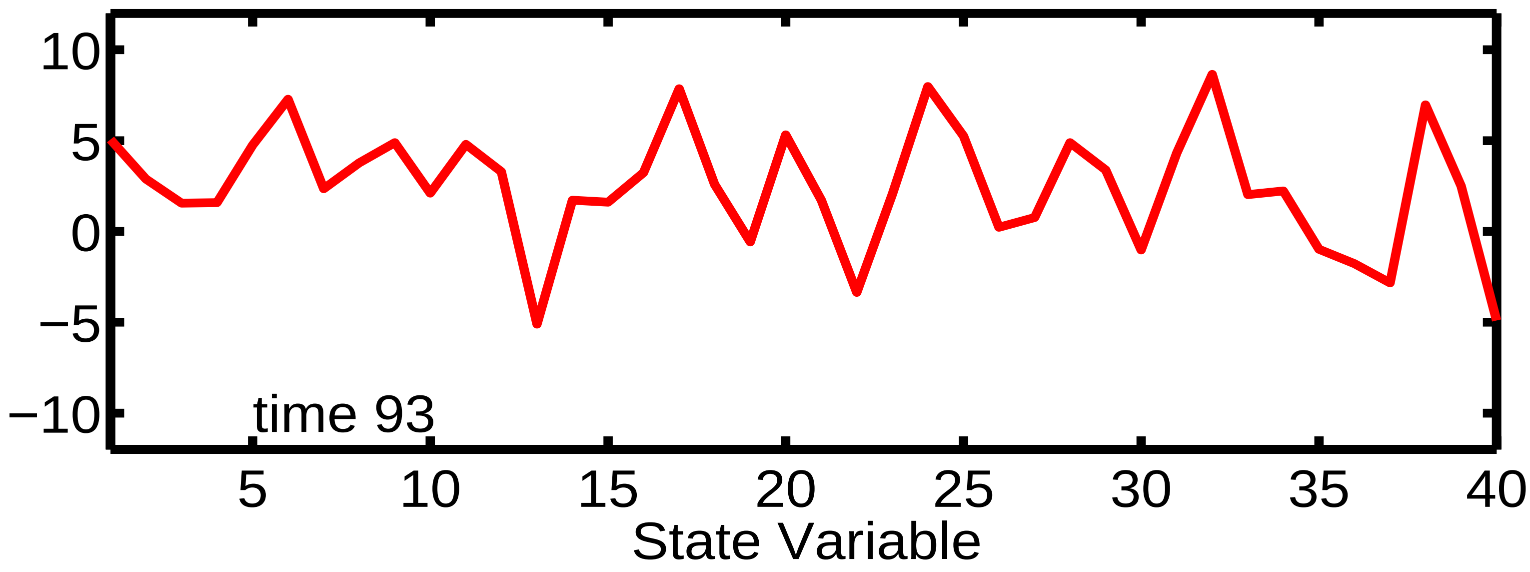


Ensemble Filter for Lorenz-96 40-Variable Model

40 state variables: $X_1, X_2, \ldots, X_{40}$.

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F$.

Acts ‘something’ like synoptic weather around a latitude band.

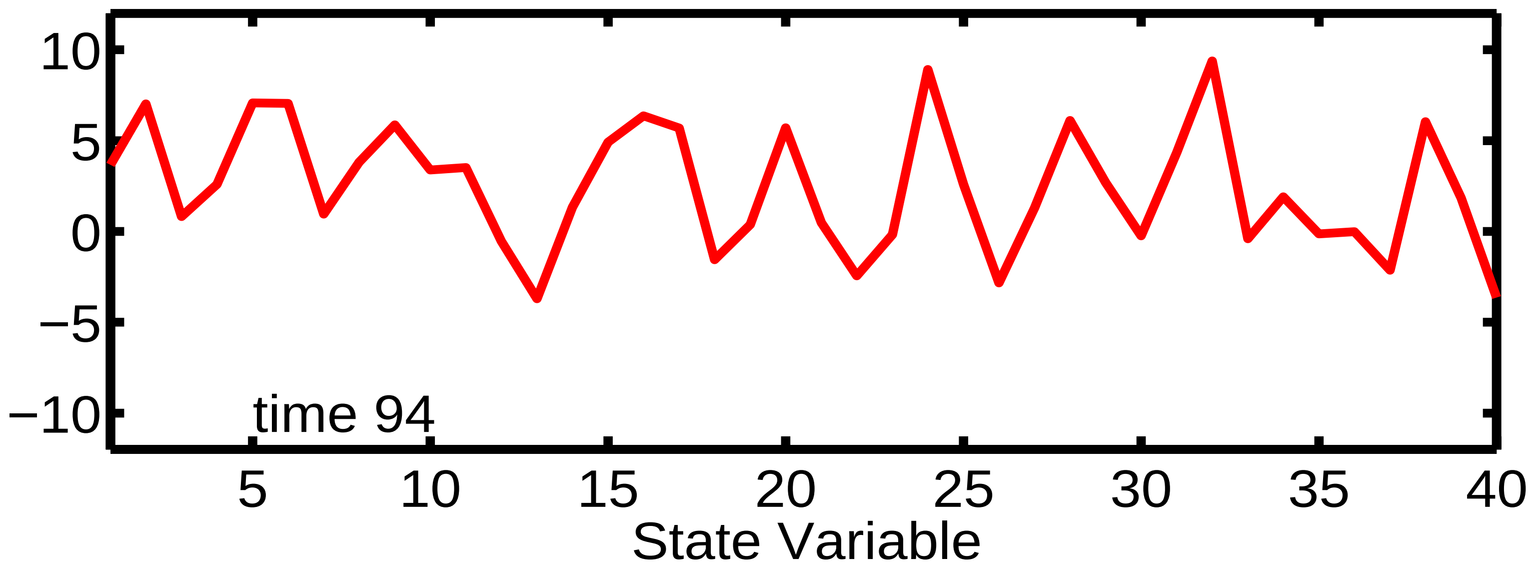


Ensemble Filter for Lorenz-96 40-Variable Model

40 state variables: X_1, X_2,..., $X_{40.}$

dX_i / dt = $(X_{i+1}$ - $X_{i-2})X_{i-1}$ - X_i + F.

Acts ‘something’ like synoptic weather around a latitude band.

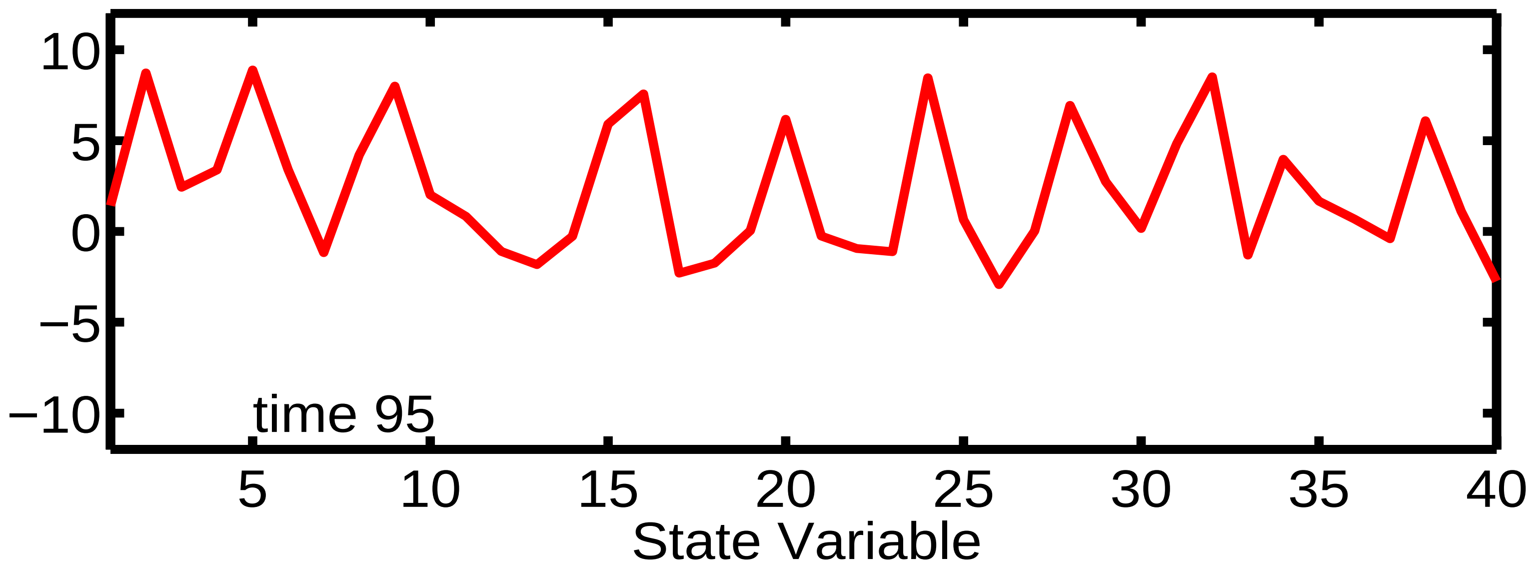


Ensemble Filter for Lorenz-96 40-Variable Model

40 state variables: X_1, X_2,..., $X_{40.}$

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

Acts ‘something’ like synoptic weather around a latitude band.

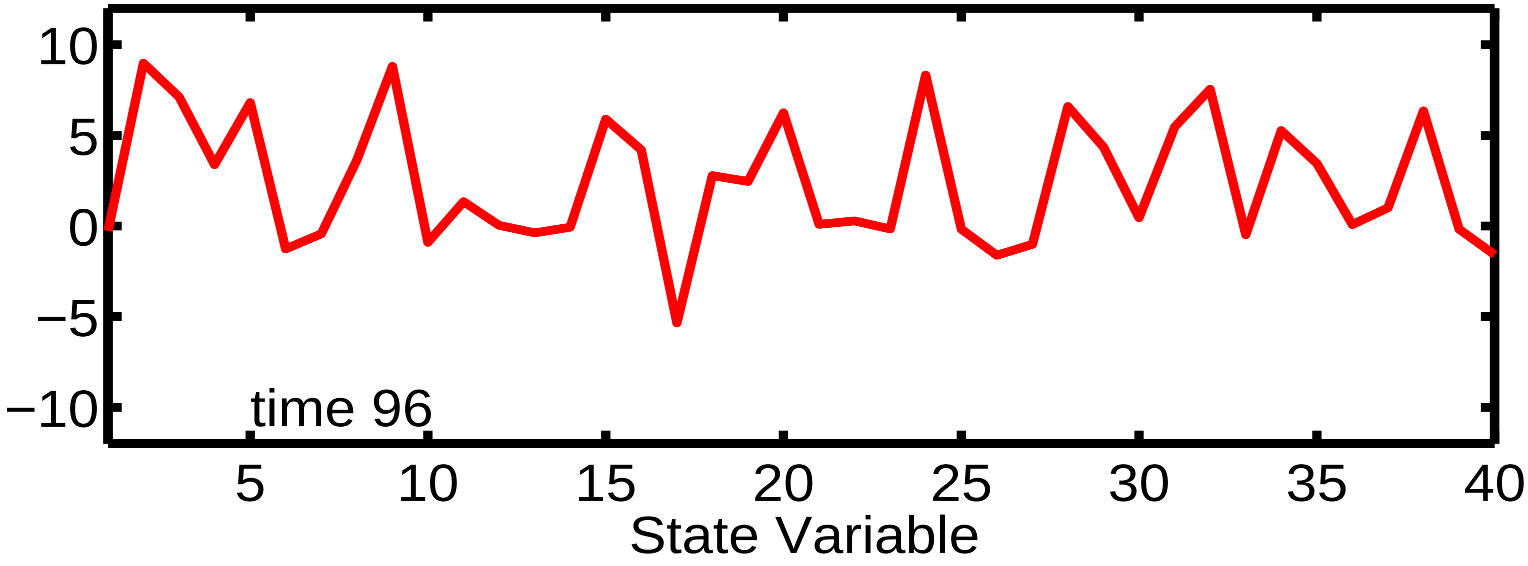


Ensemble Filter for Lorenz-96 40-Variable Model

40 state variables: X_1, X_2,..., $X_{40.}$

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

Acts ‘something’ like synoptic weather around a latitude band.

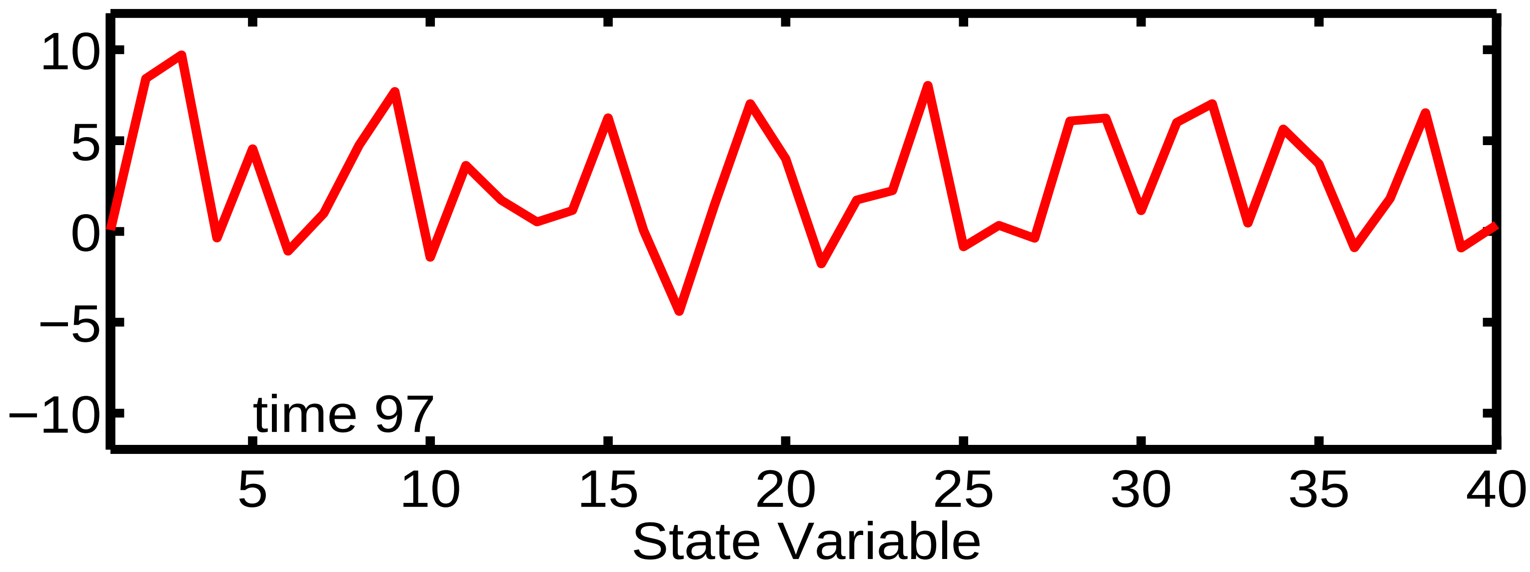


Ensemble Filter for Lorenz-96 40-Variable Model

40 state variables: $X_1, X_2, \ldots, X_{40.}$

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

Acts ‘something’ like synoptic weather around a latitude band.

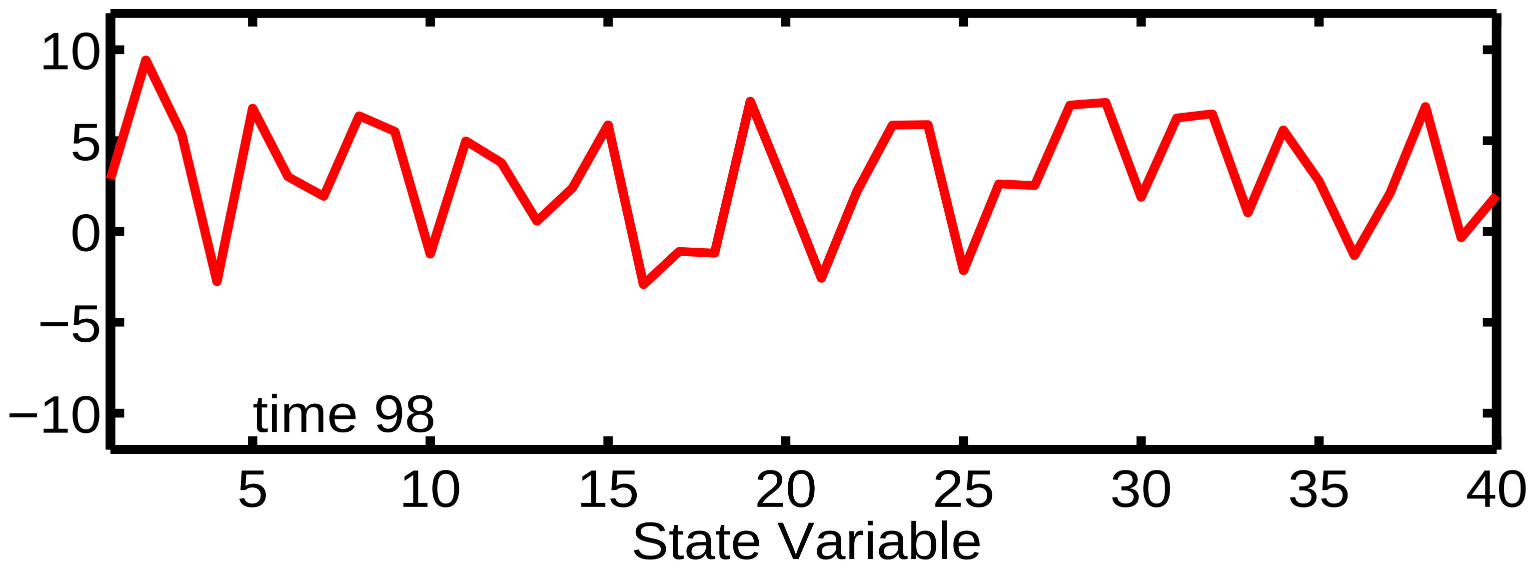


Ensemble Filter for Lorenz-96 40-Variable Model

40 state variables: X_1, X_2,..., $X_{40.}$

dX_i / dt = $(X_{i+1} - X_{i-2})X_{i-1} - X_i + F$.

Acts ‘something’ like synoptic weather around a latitude band.

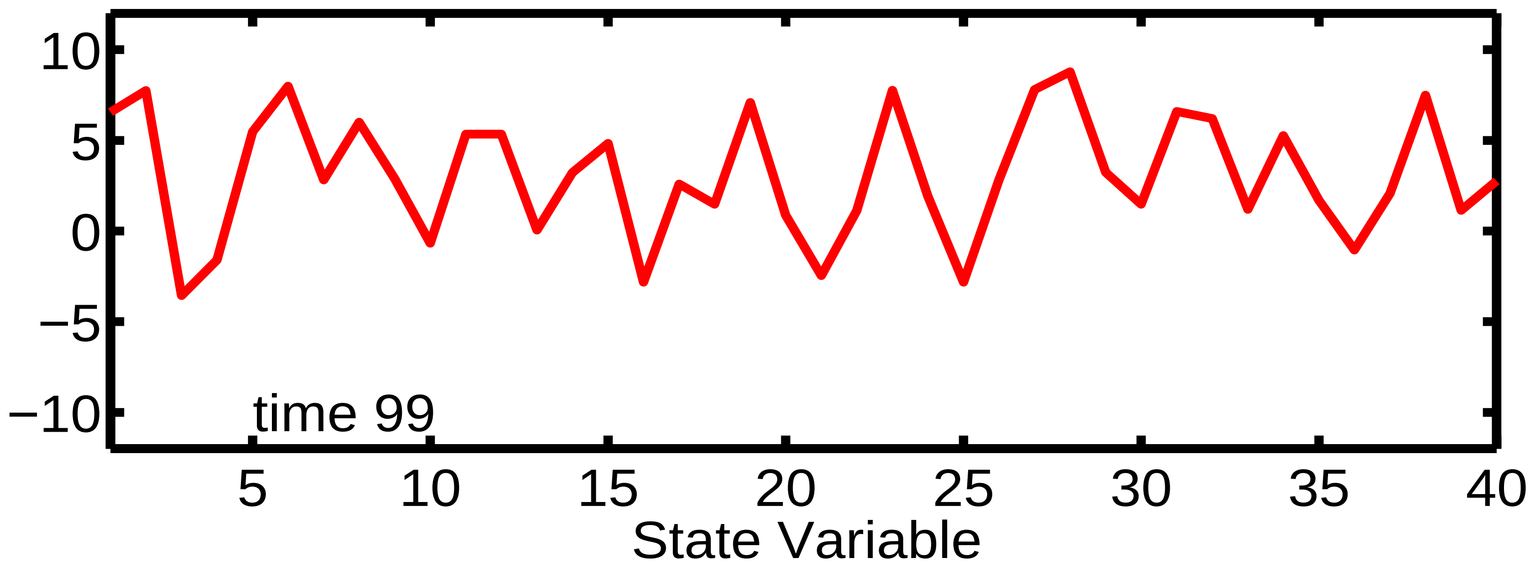


Ensemble Filter for Lorenz-96 40-Variable Model

40 state variables: $X_1, X_2, \ldots, X_{40.}$

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

Acts 'something' like synoptic weather around a latitude band.

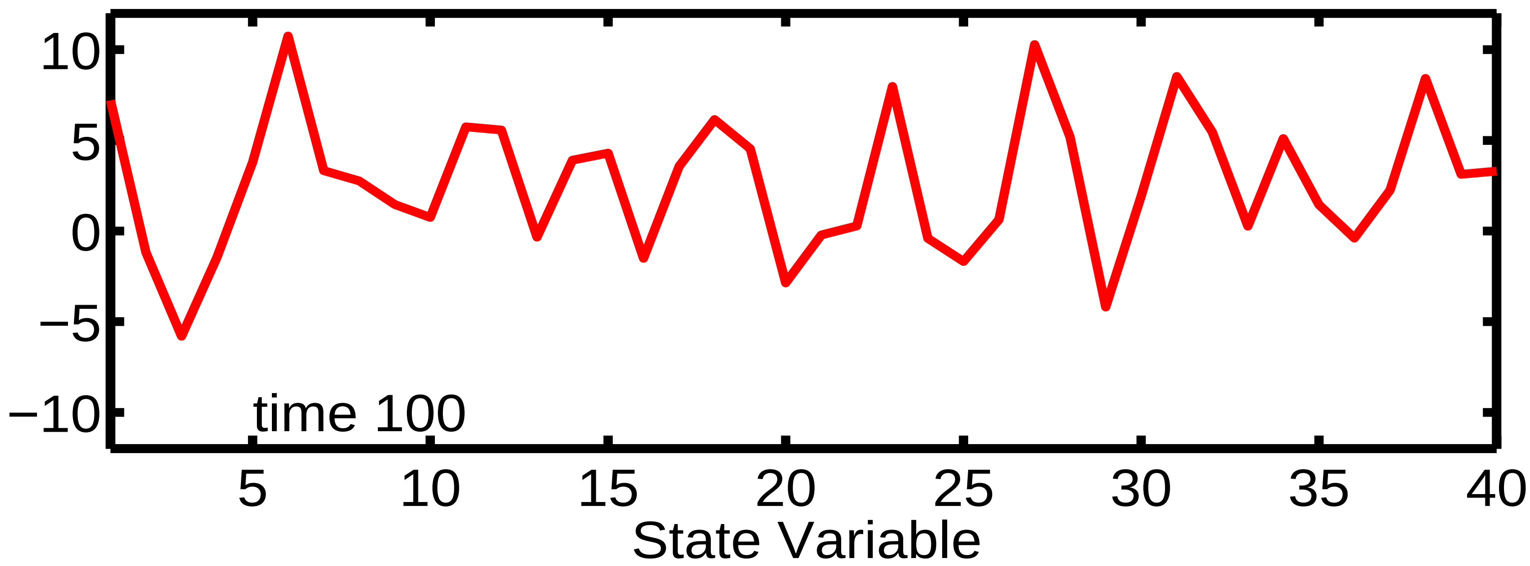


Ensemble Filter for Lorenz-96 40-Variable Model

40 state variables: X_1, X_2,..., $X_{40.}$

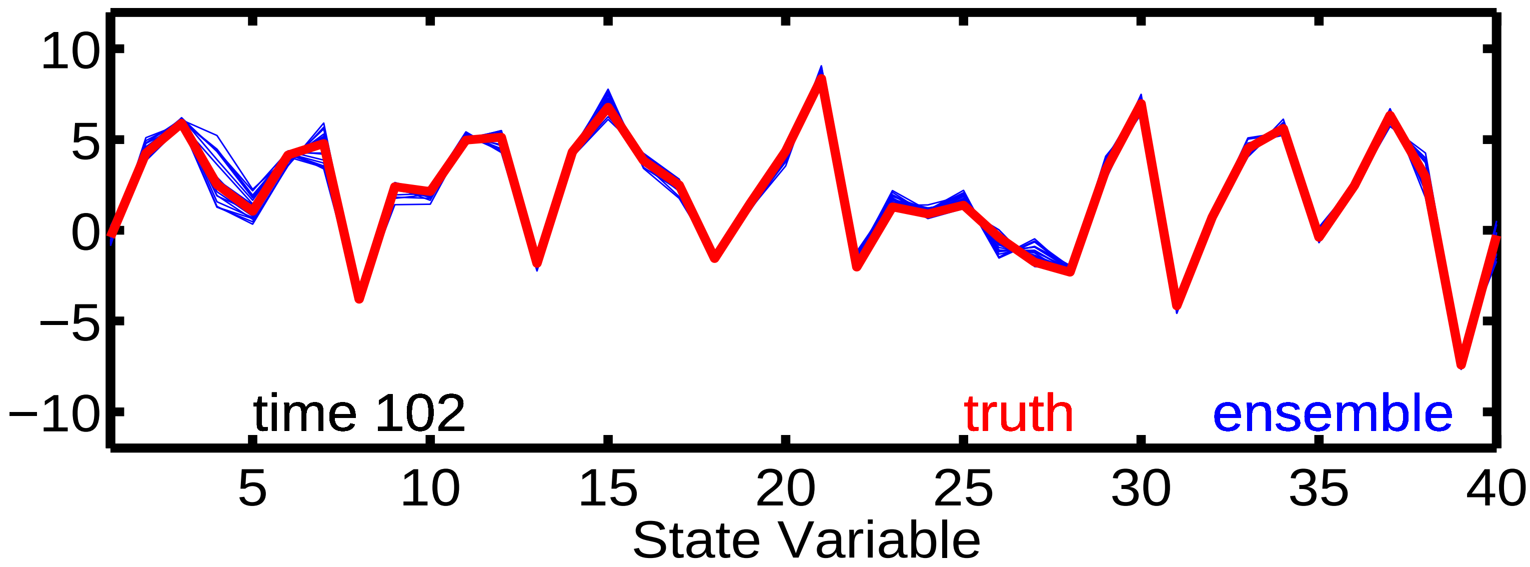
$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

Acts 'something' like synoptic weather around a latitude band.



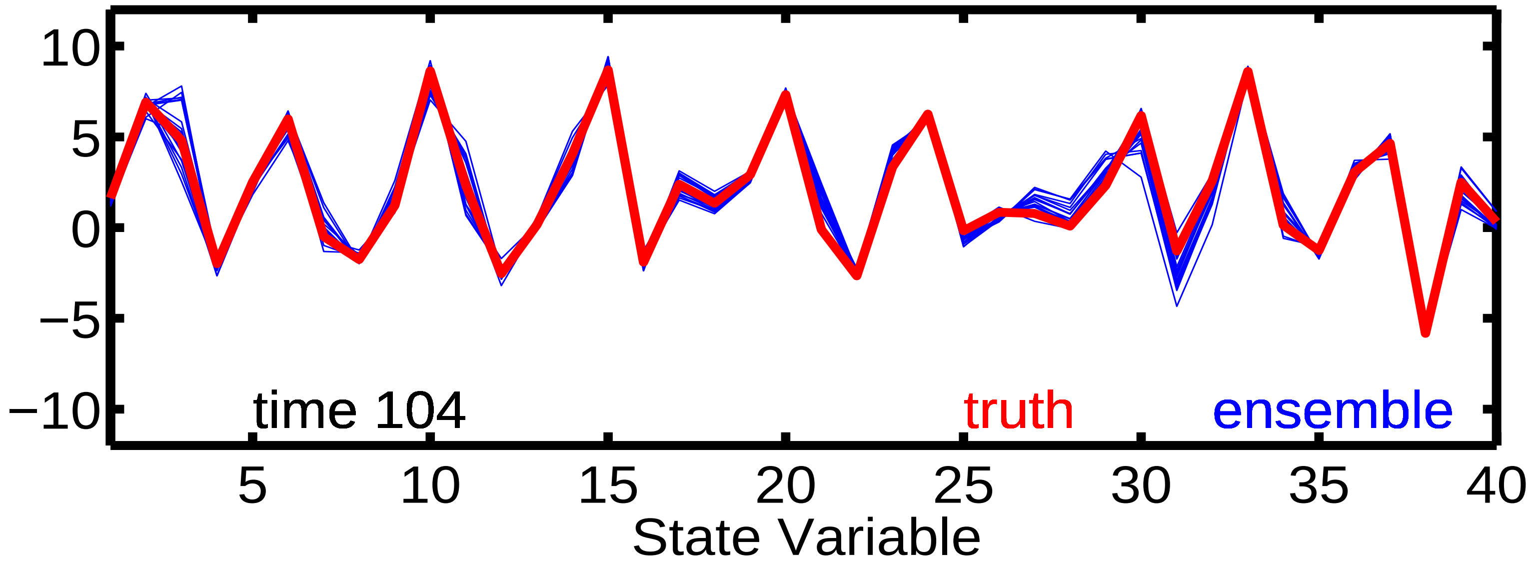
Lorenz-96 is sensitive to small perturbations

Introduce 20 ‘ensemble’ state estimates.
Each is slightly perturbed for each Xi at time 100.
Refer to unperturbed control integration as ‘truth’.



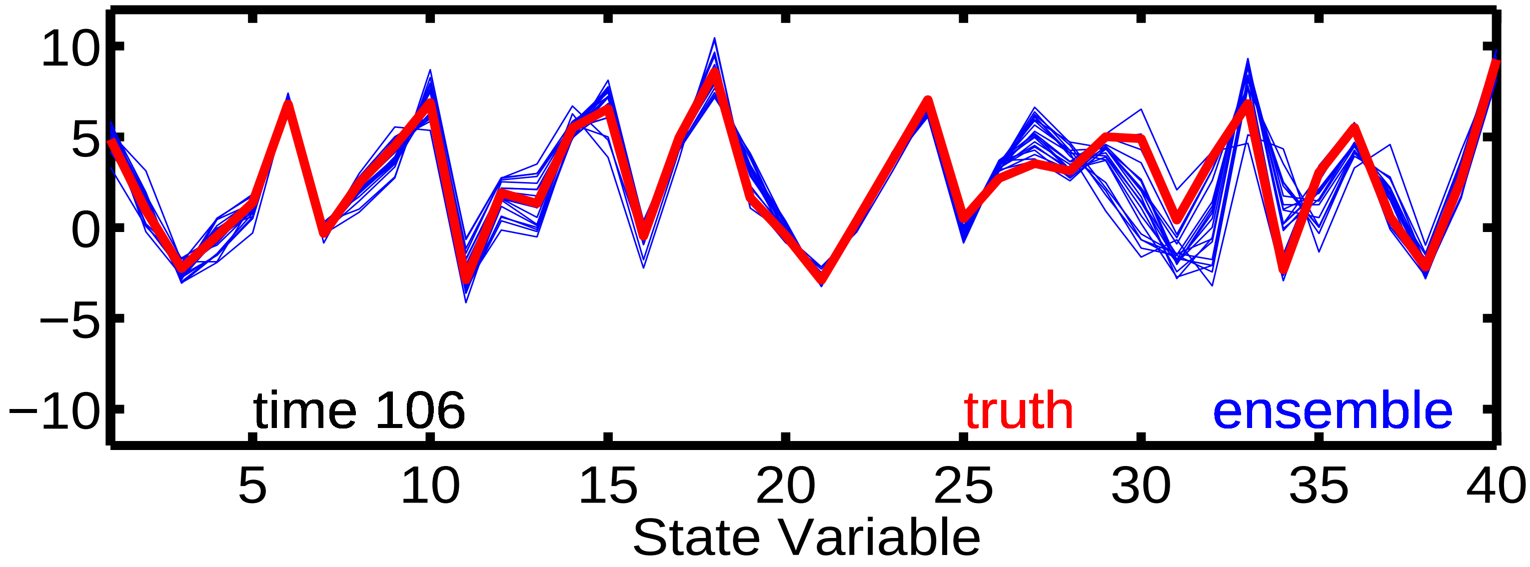
Lorenz-96 is sensitive to small perturbations

Introduce 20 ‘ensemble’ state estimates.
Each is slightly perturbed for each Xi at time 100.
Refer to unperturbed control integration as ‘truth’.



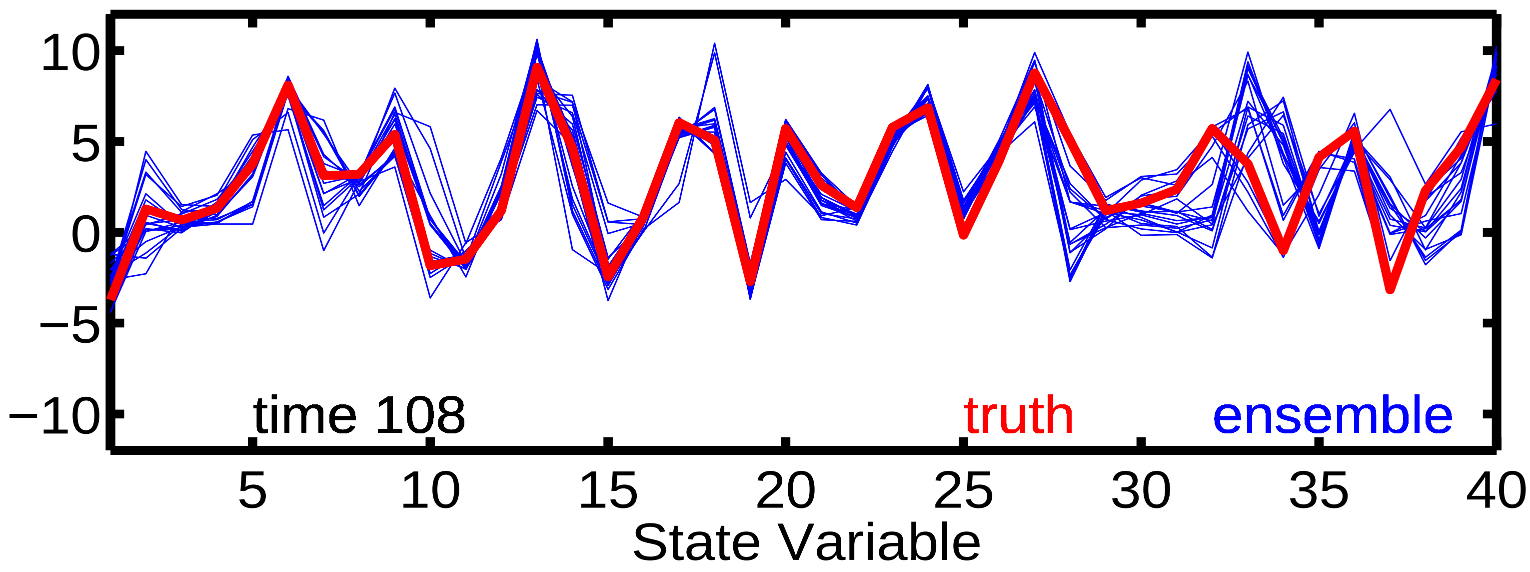
Lorenz-96 is sensitive to small perturbations

Introduce 20 ‘ensemble’ state estimates.
Each is slightly perturbed for each Xi at time 100.
Refer to unperturbed control integration as ‘truth’.



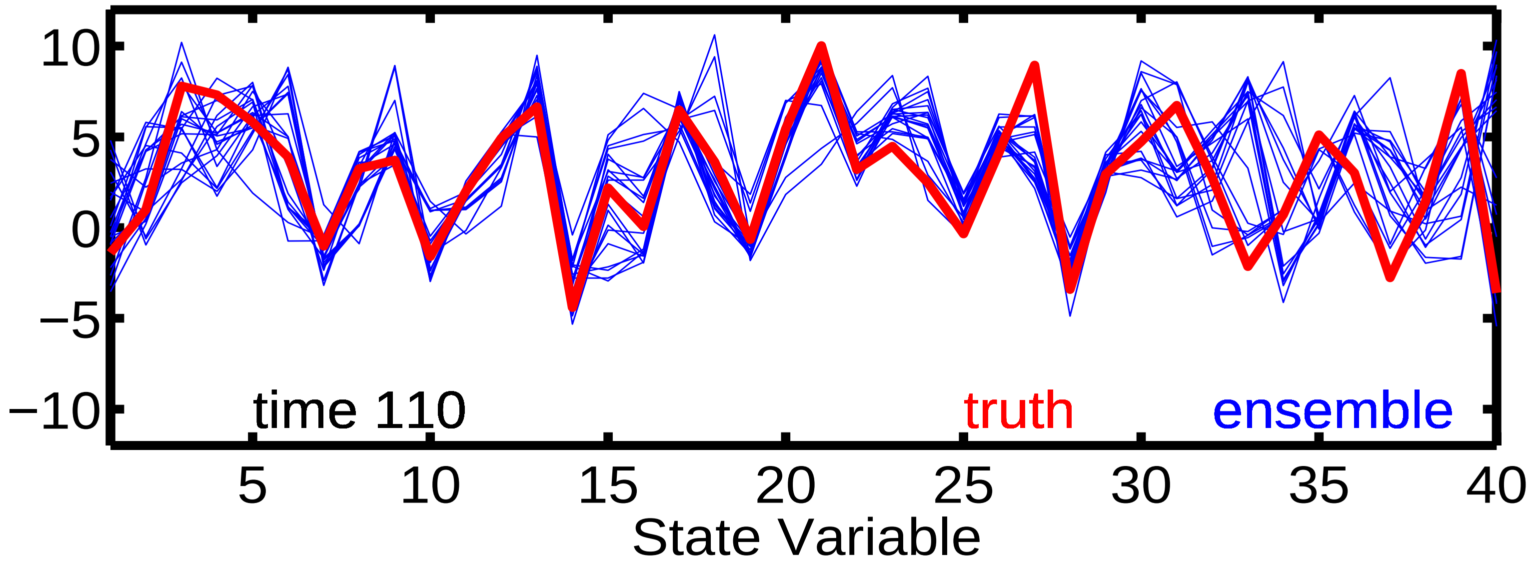
Lorenz-96 is sensitive to small perturbations

Introduce 20 ‘ensemble’ state estimates.
Each is slightly perturbed for each Xi at time 100.
Refer to unperturbed control integration as ‘truth’.



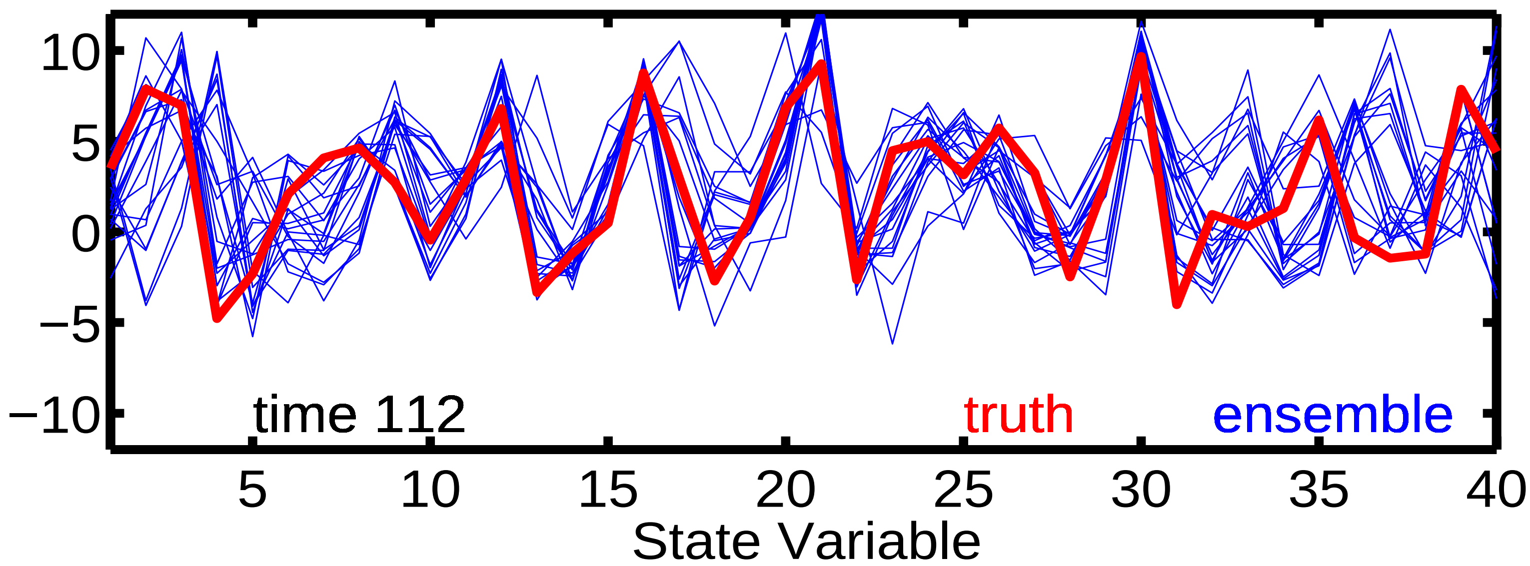
Lorenz-96 is sensitive to small perturbations

Introduce 20 'ensemble' state estimates.
Each is slightly perturbed for each Xi at time 100.
Refer to unperturbed control integration as 'truth'.



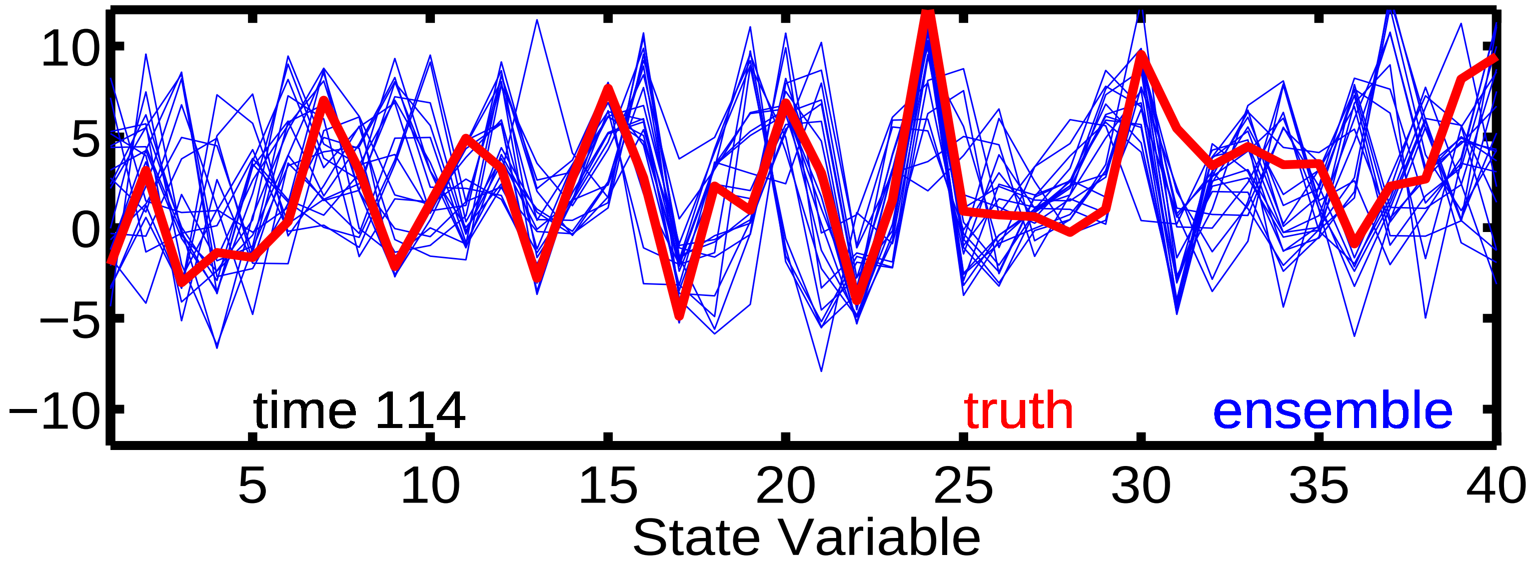
Lorenz-96 is sensitive to small perturbations

Introduce 20 ‘ensemble’ state estimates.
Each is slightly perturbed for each Xi at time 100.
Refer to unperturbed control integration as ‘truth’.



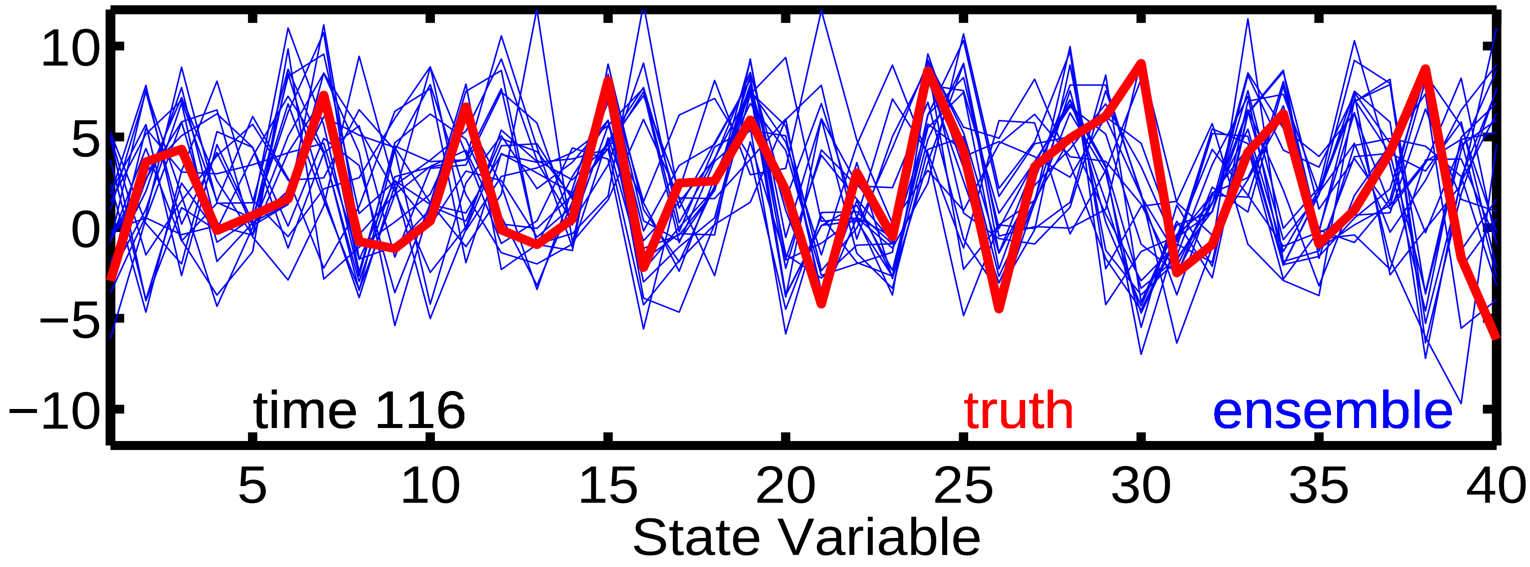
Lorenz-96 is sensitive to small perturbations

Introduce 20 ‘ensemble’ state estimates.
Each is slightly perturbed for each Xi at time 100.
Refer to unperturbed control integration as ‘truth’.



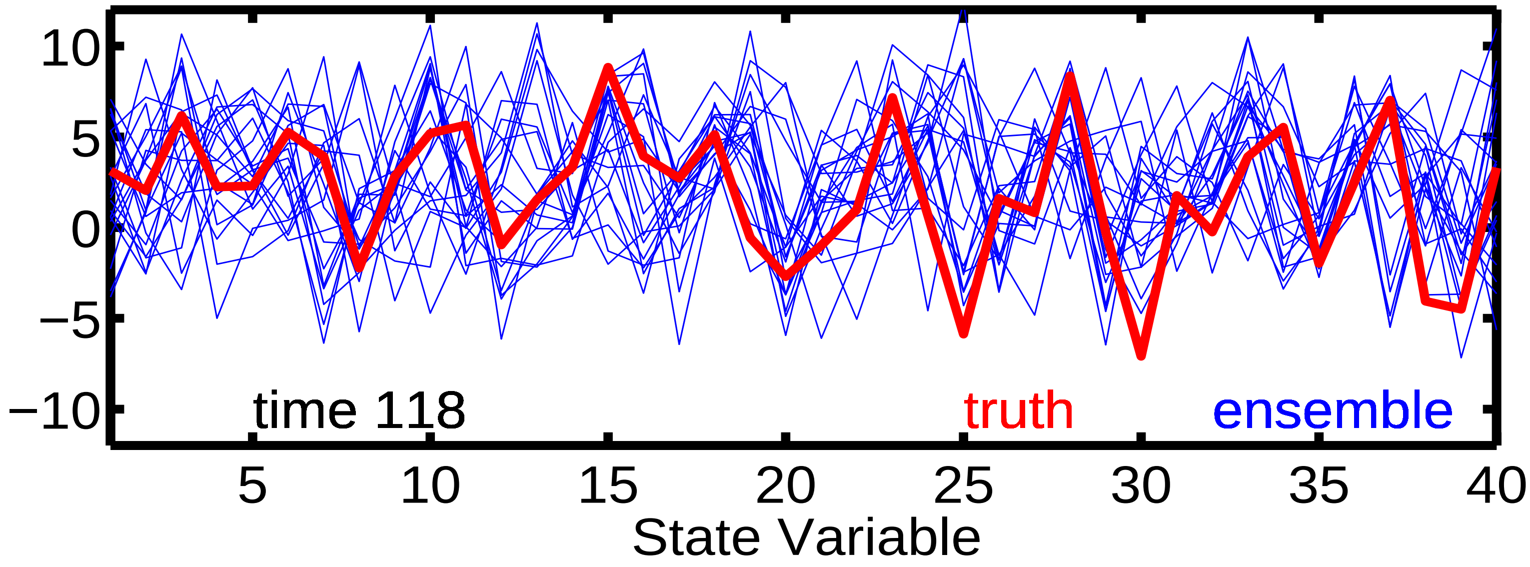
Lorenz-96 is sensitive to small perturbations

Introduce 20 ‘ensemble’ state estimates.
Each is slightly perturbed for each Xi at time 100.
Refer to unperturbed control integration as ‘truth’.



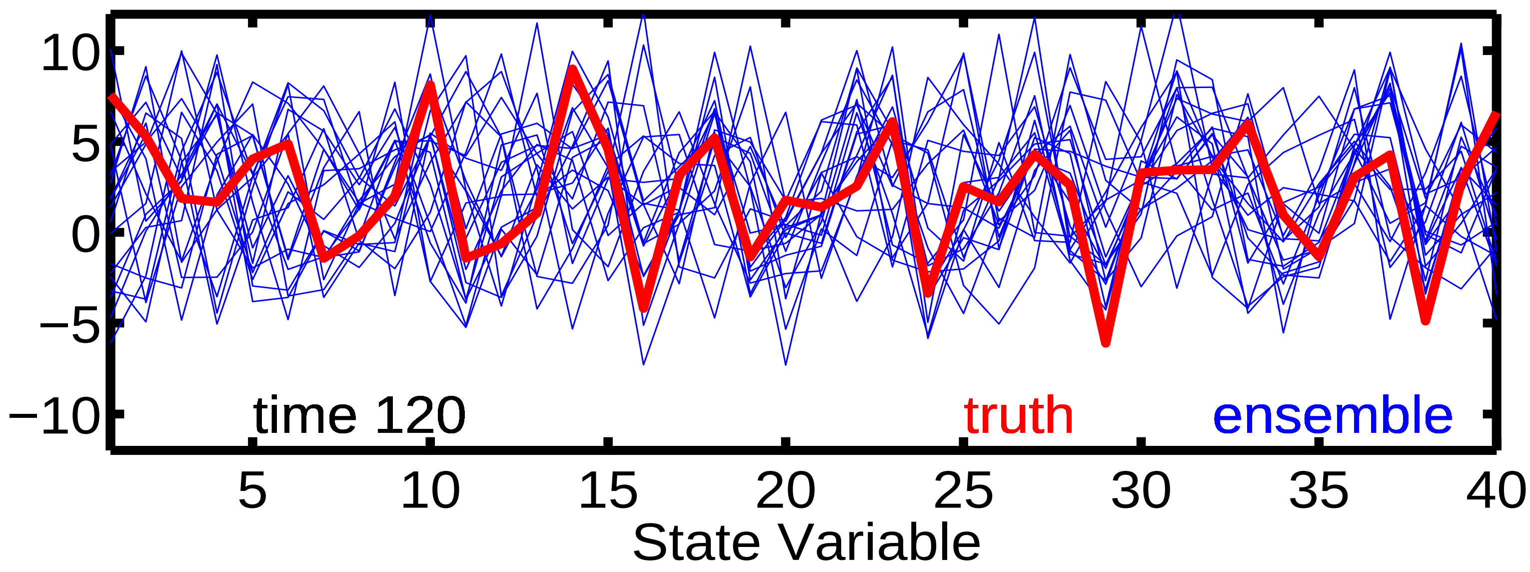
Lorenz-96 is sensitive to small perturbations

Introduce 20 ‘ensemble’ state estimates.
Each is slightly perturbed for each Xi at time 100.
Refer to unperturbed control integration as ‘truth’.



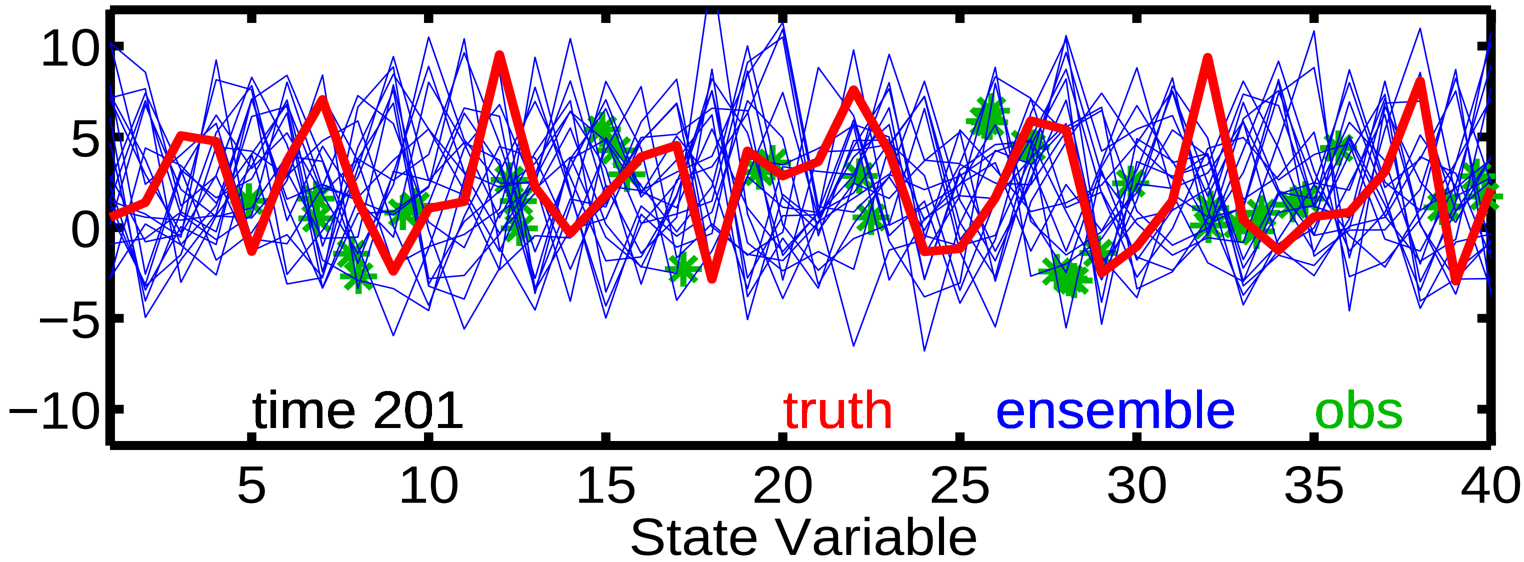
Lorenz-96 is sensitive to small perturbations

Introduce 20 ‘ensemble’ state estimates.
Each is slightly perturbed for each Xi at time 100.
Refer to unperturbed control integration as ‘truth’.



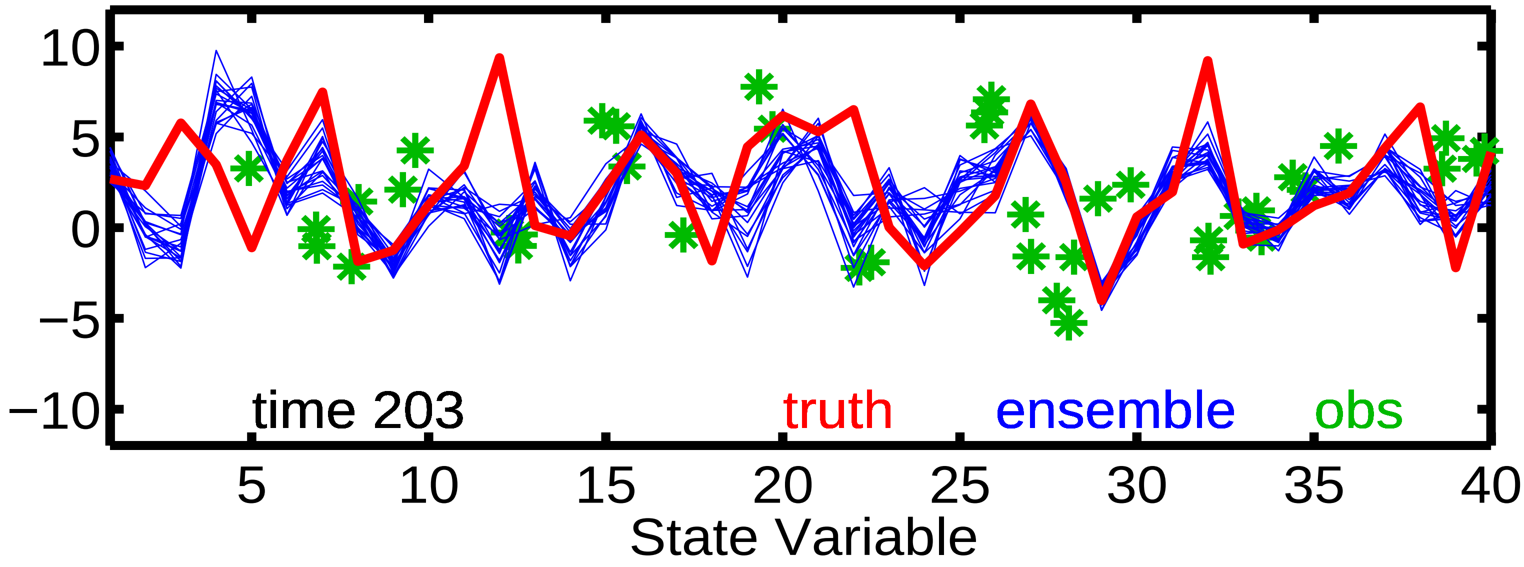
Assimilate ‘observations’ from 40 randomly located stations each step.

Observations generated by interpolating truth to station location.
Simulate observational error: Add random draw from N(0, 1) to each.
Start from ‘climatological’ 20-member ensemble.



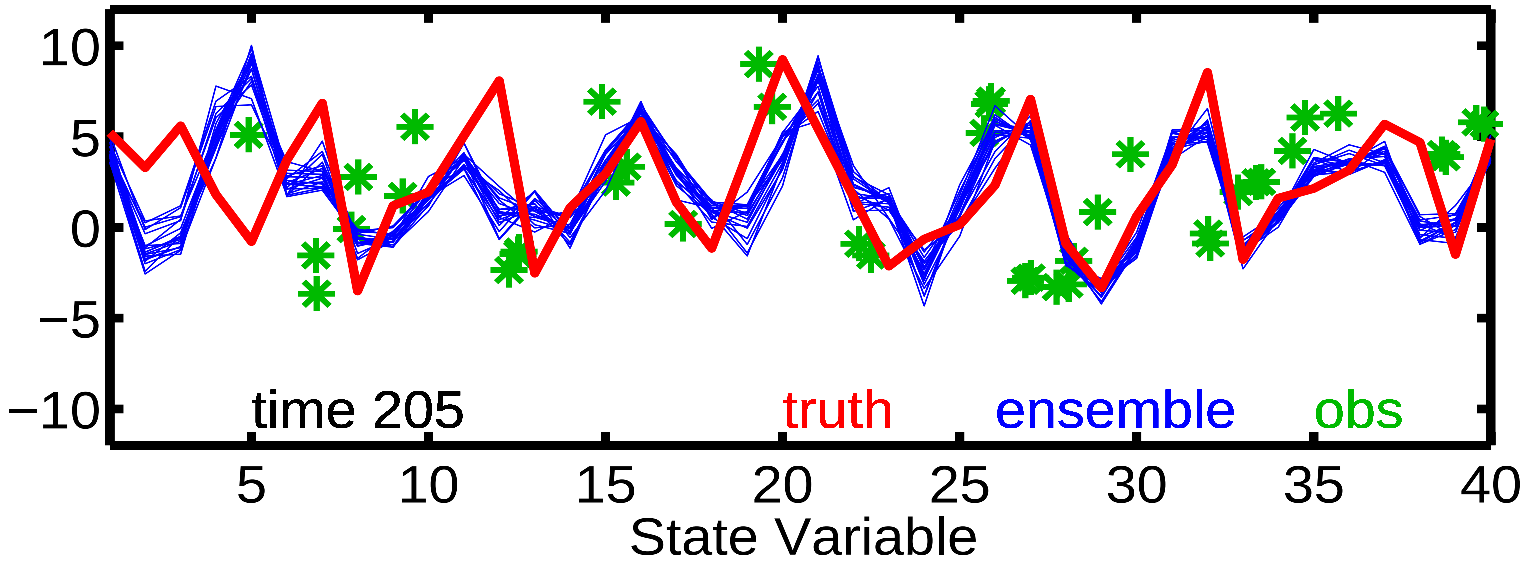
<u>Assimilate 'observations' from 40 randomly located stations each step.</u>

Observations generated by interpolating truth to station location.
Simulate observational error: Add random draw from N(0, 1) to each.
Start from 'climatological' 20-member ensemble.



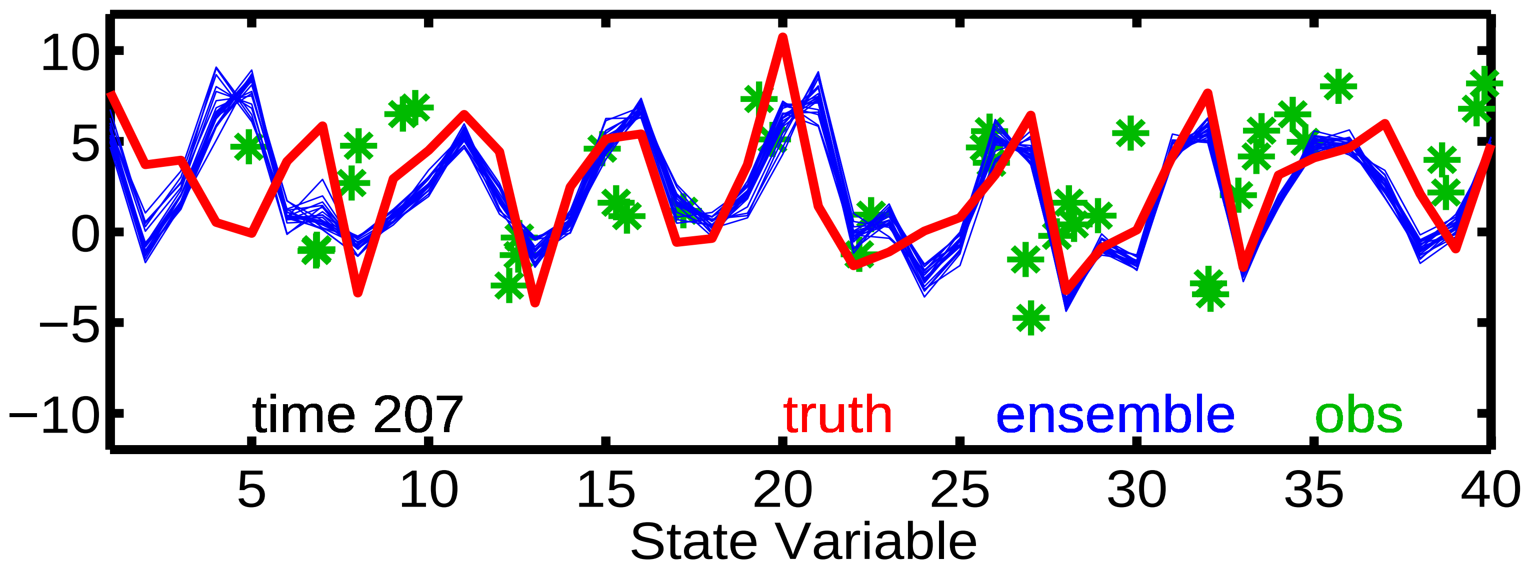
Assimilate 'observations' from 40 randomly located stations each step.

Observations generated by interpolating truth to station location.
Simulate observational error: Add random draw from N(0, 1) to each.
Start from 'climatological' 20-member ensemble.



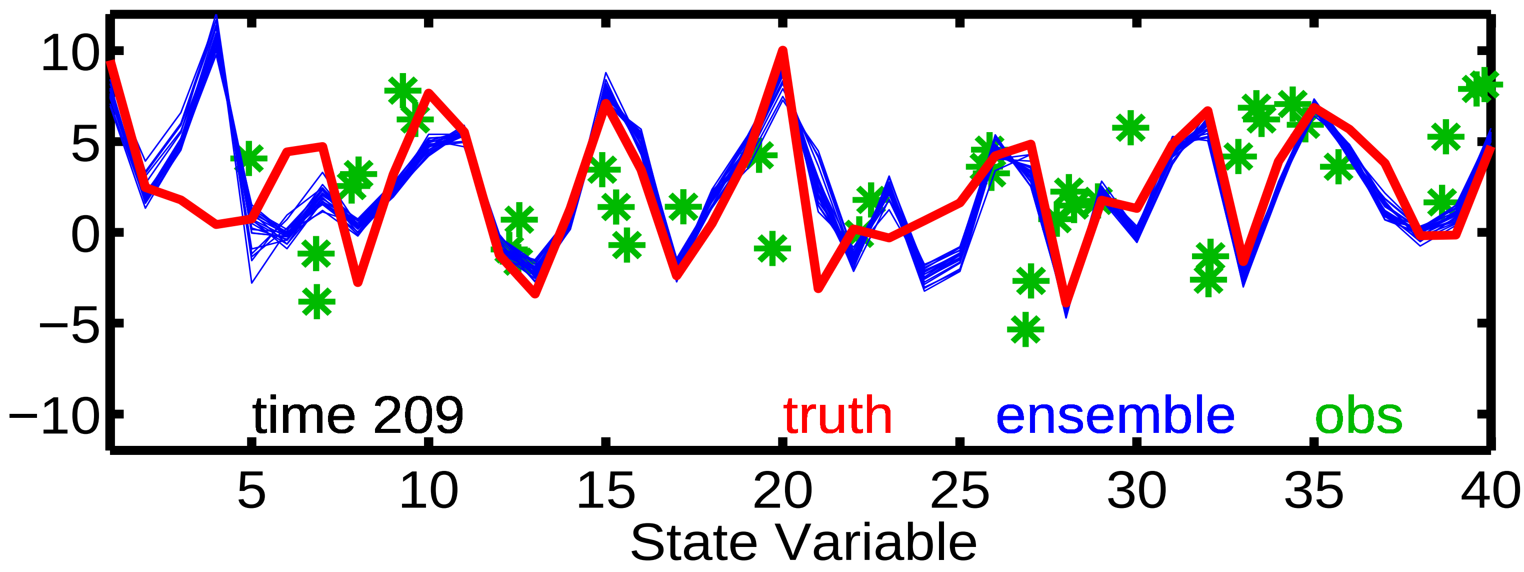
Assimilate 'observations' from 40 randomly located stations each step.

Observations generated by interpolating truth to station location.
Simulate observational error: Add random draw from N(0, 1) to each.
Start from 'climatological' 20-member ensemble.



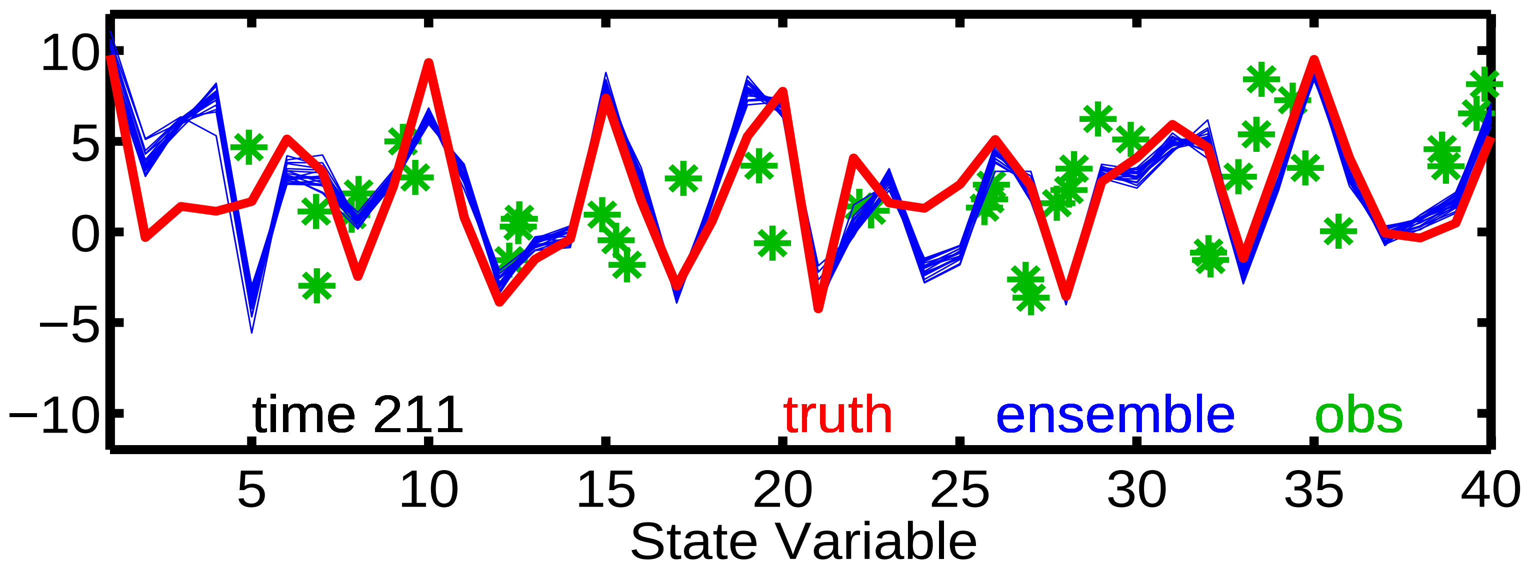
Assimilate ‘observations’ from 40 randomly located stations each step.

Observations generated by interpolating truth to station location.
Simulate observational error: Add random draw from N(0, 1) to each.
Start from ‘climatological’ 20-member ensemble.



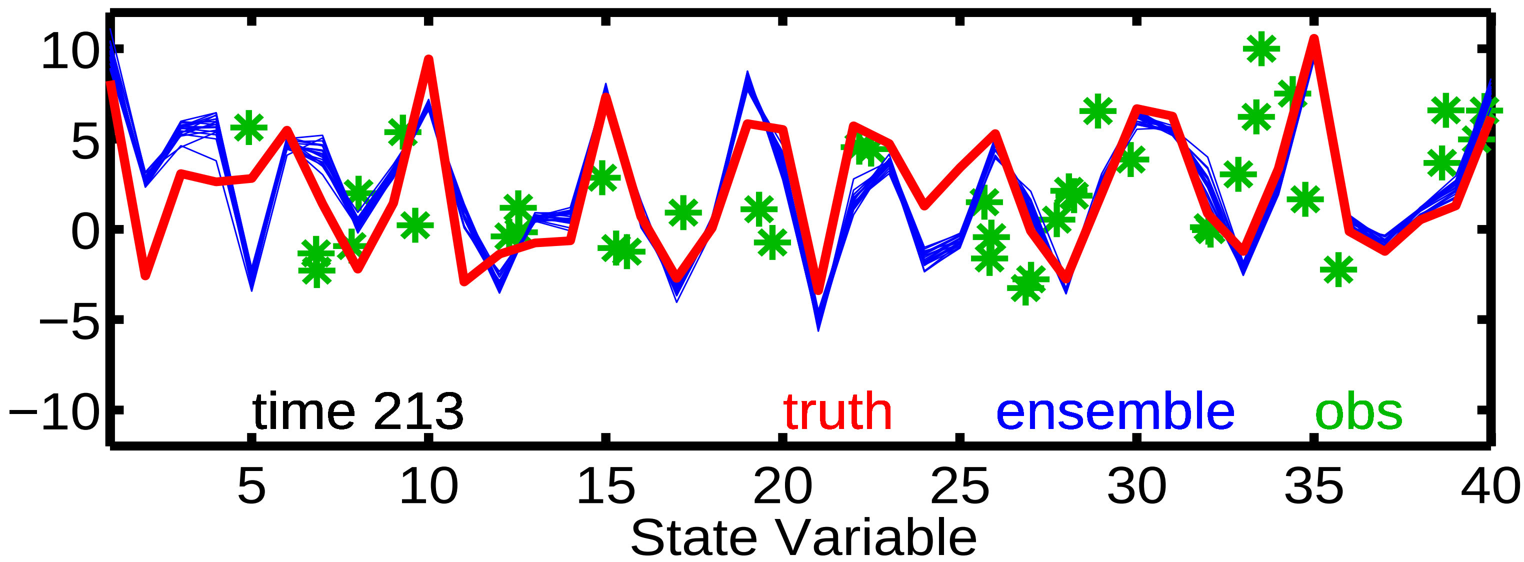
Assimilate 'observations' from 40 randomly located stations each step.

Observations generated by interpolating truth to station location.
Simulate observational error: Add random draw from N(0, 1) to each.
Start from 'climatological' 20-member ensemble.



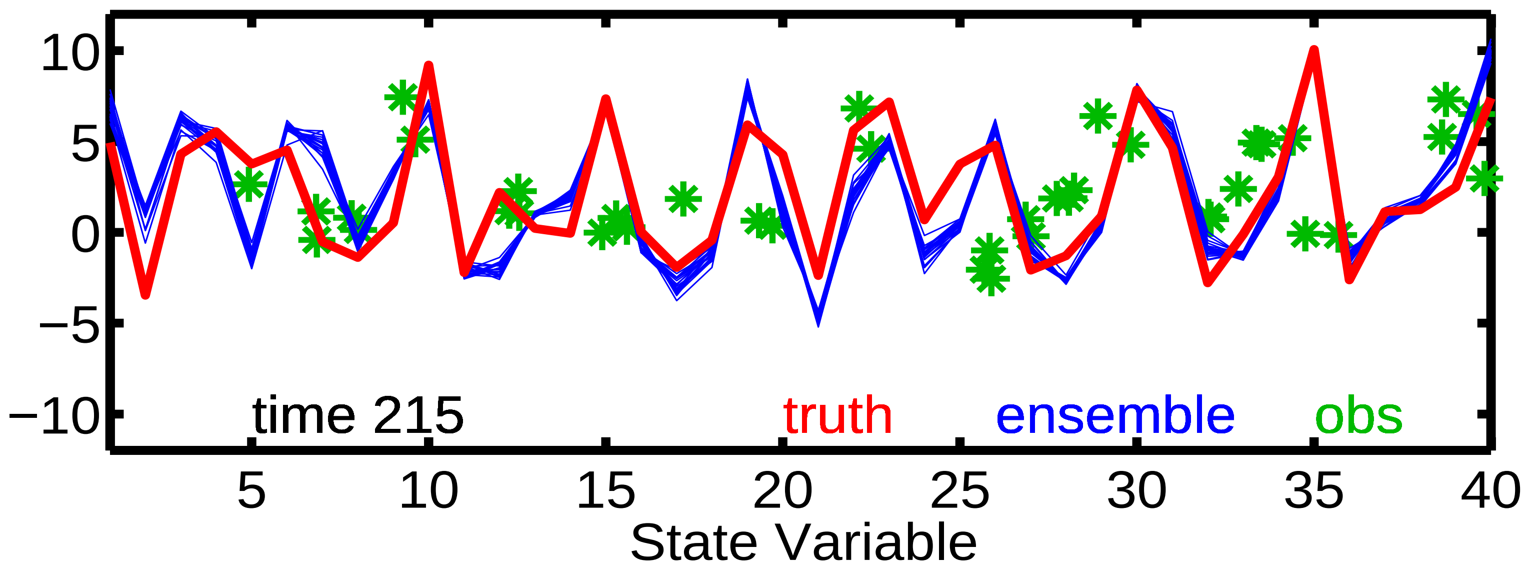
Assimilate ‘observations’ from 40 randomly located stations each step.

Observations generated by interpolating truth to station location.
Simulate observational error: Add random draw from N(0, 1) to each.
Start from ‘climatological’ 20-member ensemble.



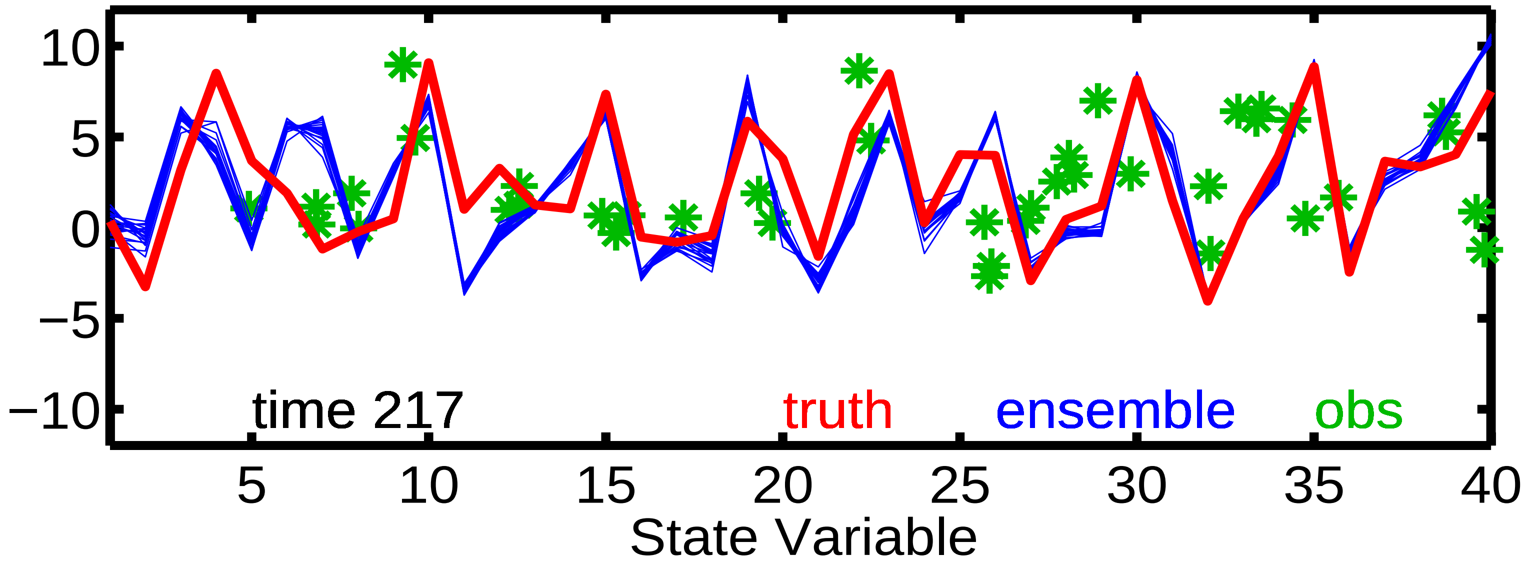
Assimilate 'observations' from 40 randomly located stations each step.

Observations generated by interpolating truth to station location.
Simulate observational error: Add random draw from N(0, 1) to each.
Start from 'climatological' 20-member ensemble.



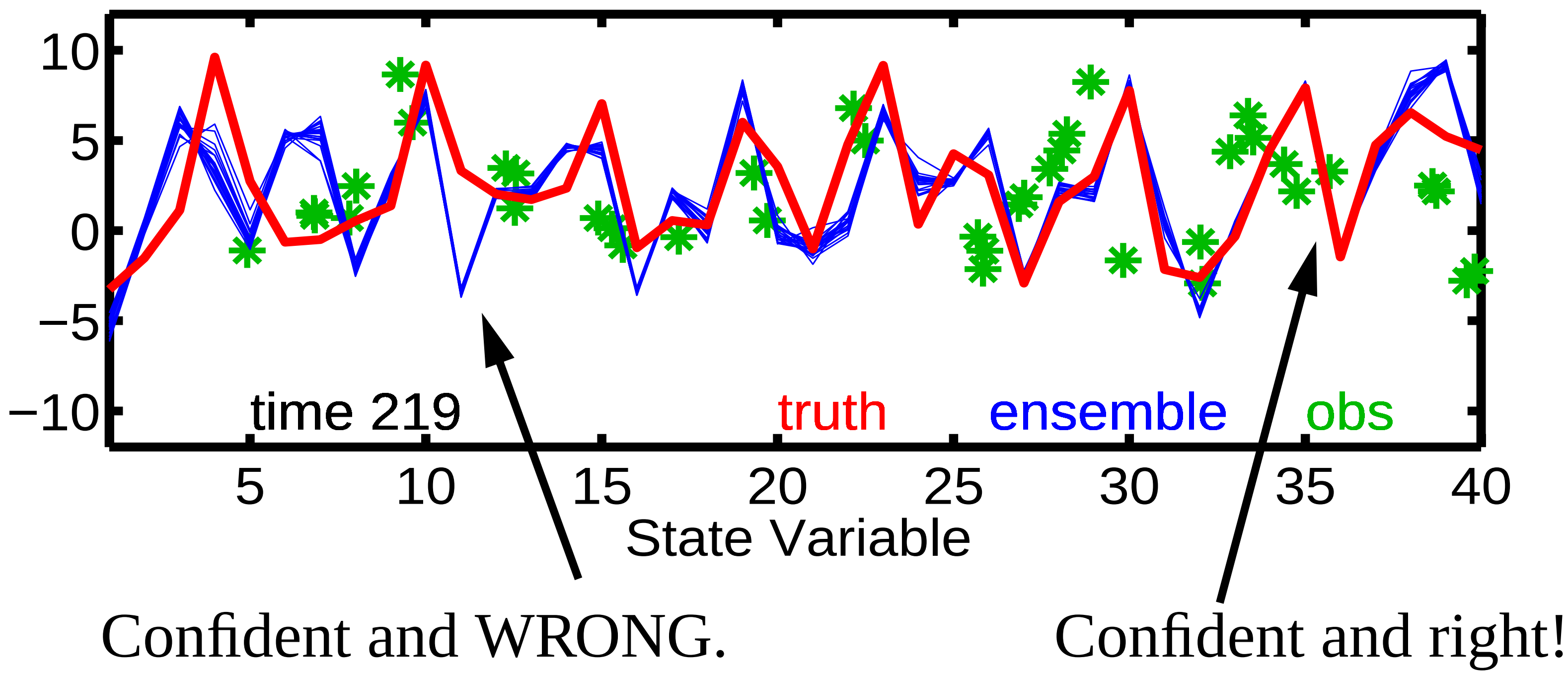
Assimilate 'observations' from 40 randomly located stations each step.

Observations generated by interpolating truth to station location.
Simulate observational error: Add random draw from N(0, 1) to each.
Start from 'climatological' 20-member ensemble.

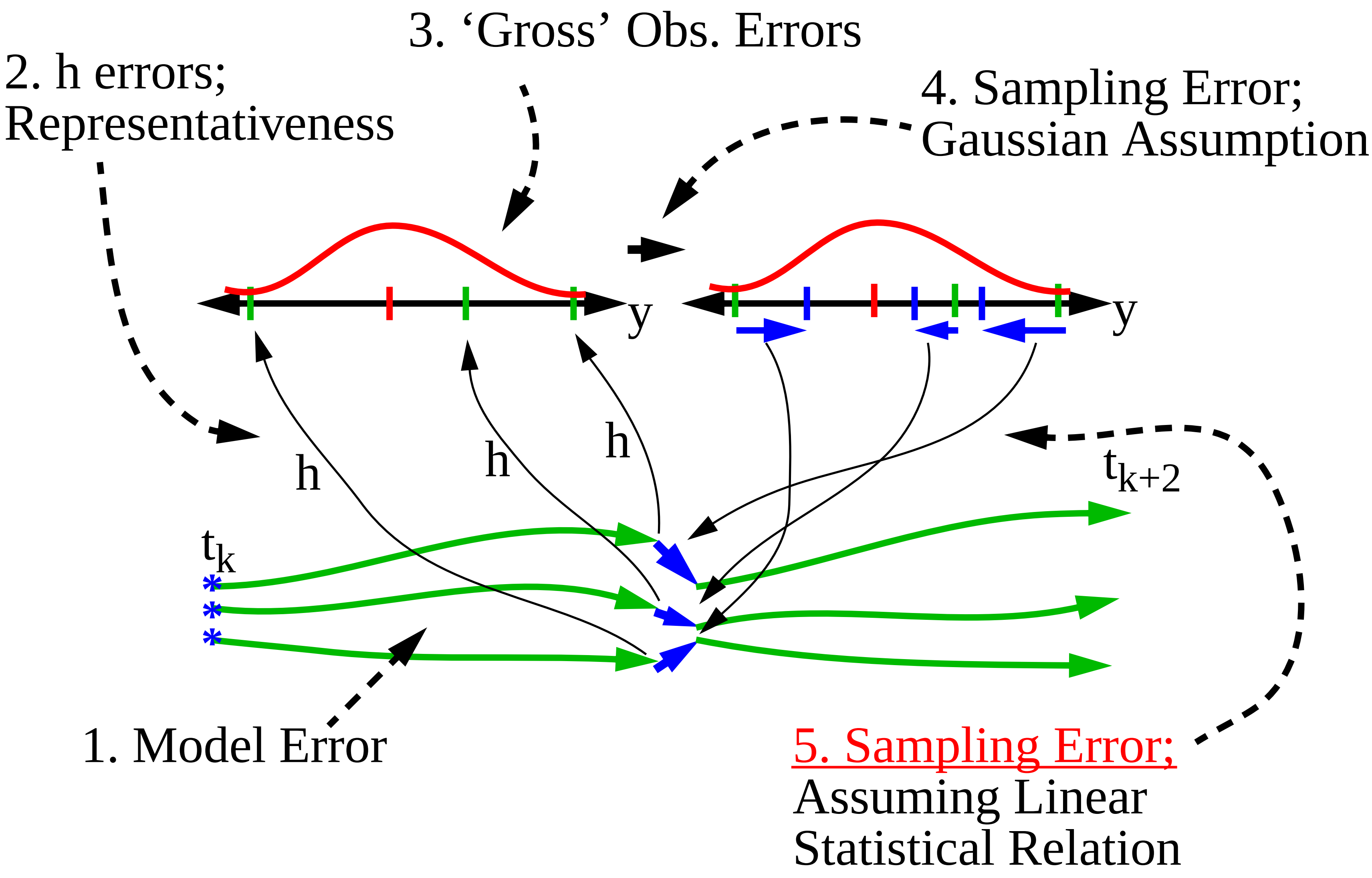


<u>Assimilate 'observations' from 40 randomly located stations each step.</u>

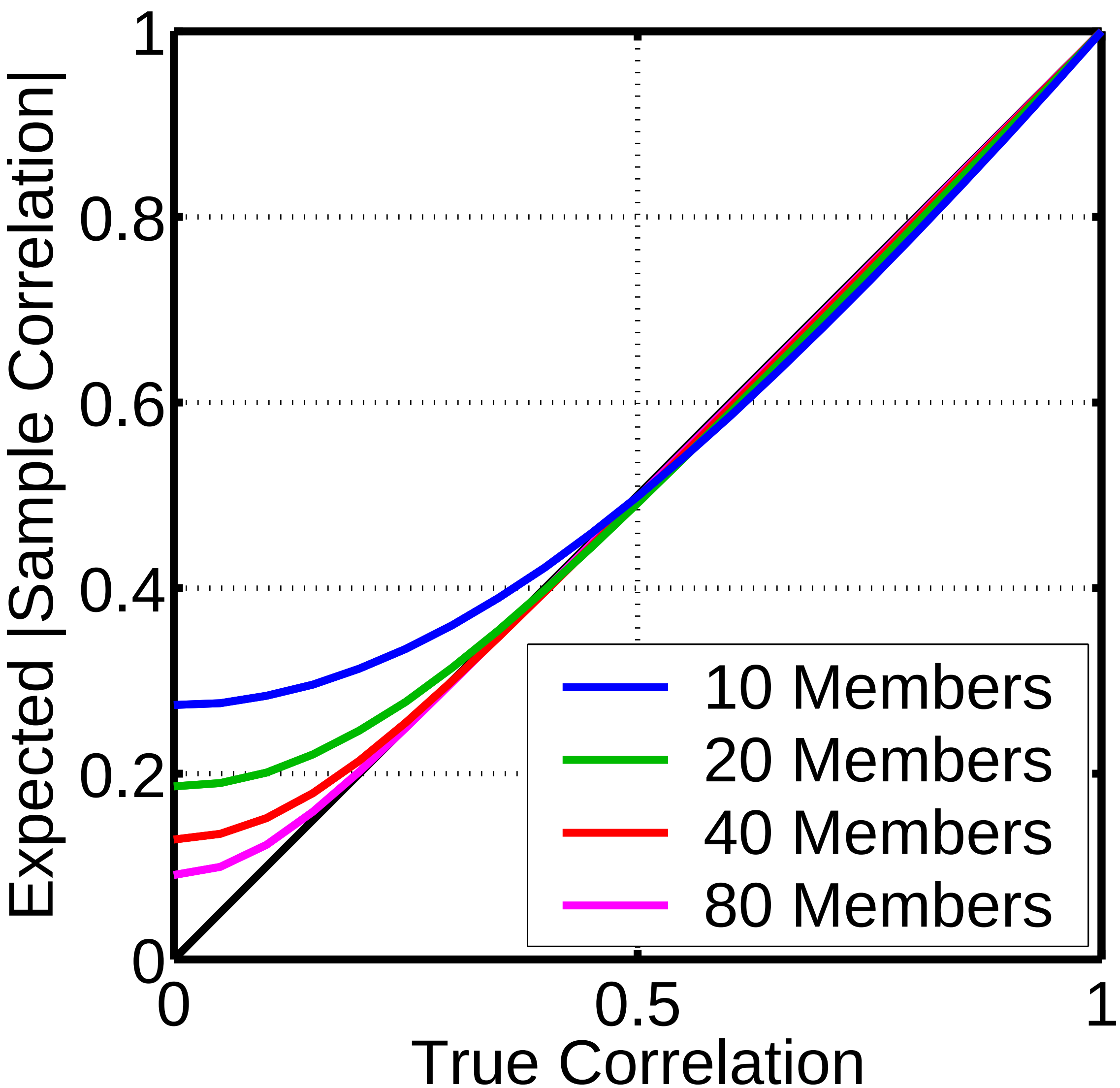
This isn't working very well.
Ensemble spread is reduced, but...,
Ensemble is inconsistent with truth most places.



Some Error Sources in Ensemble Filters



Observations impact unrelated state variables through sampling error.



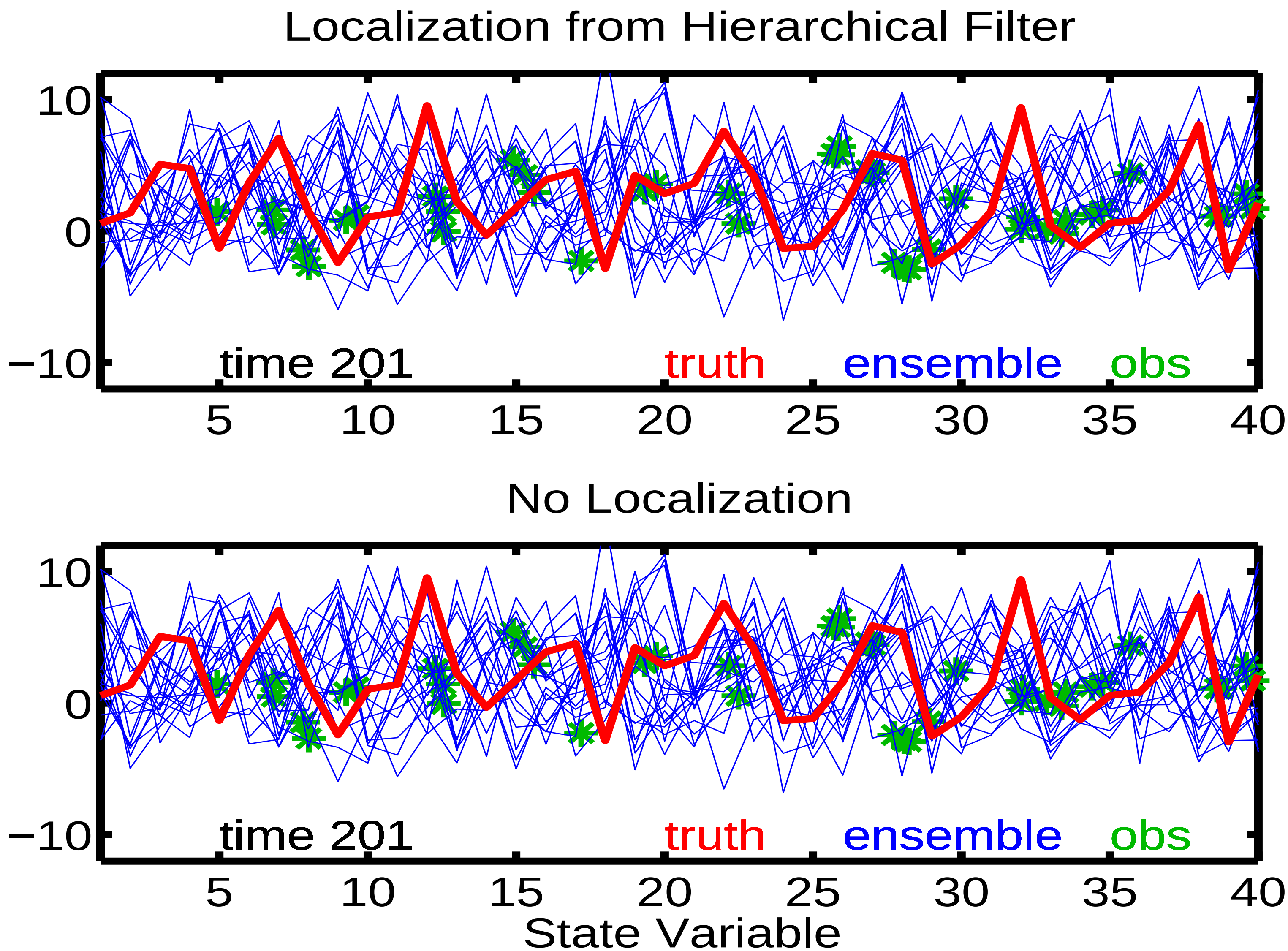
Plot shows expected absolute value of sample correlation vs. true correlation.

Unrelated obs. reduce spread, increase error.

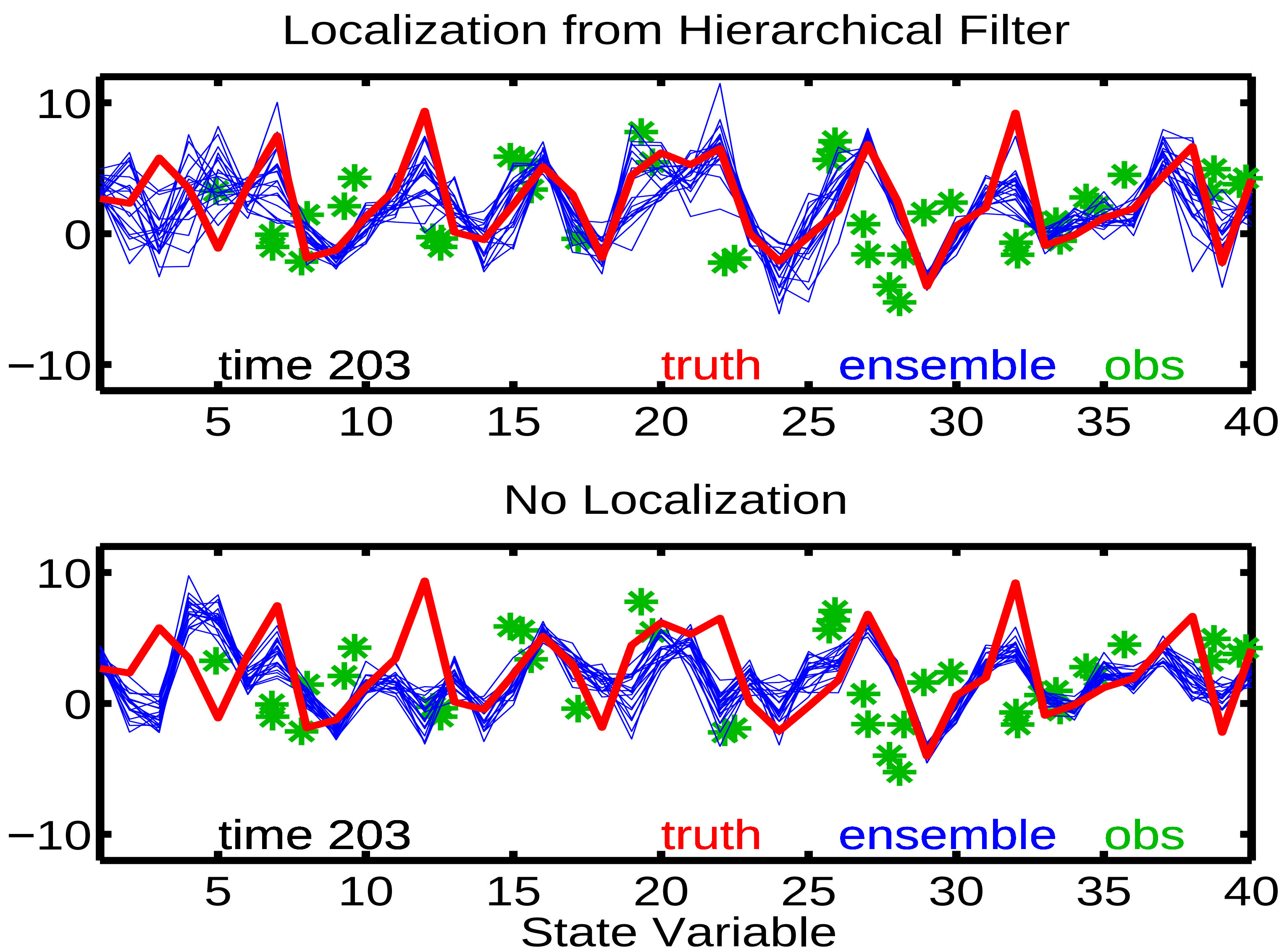
Attack with localization.

Don't let obs. impact unrelated state.

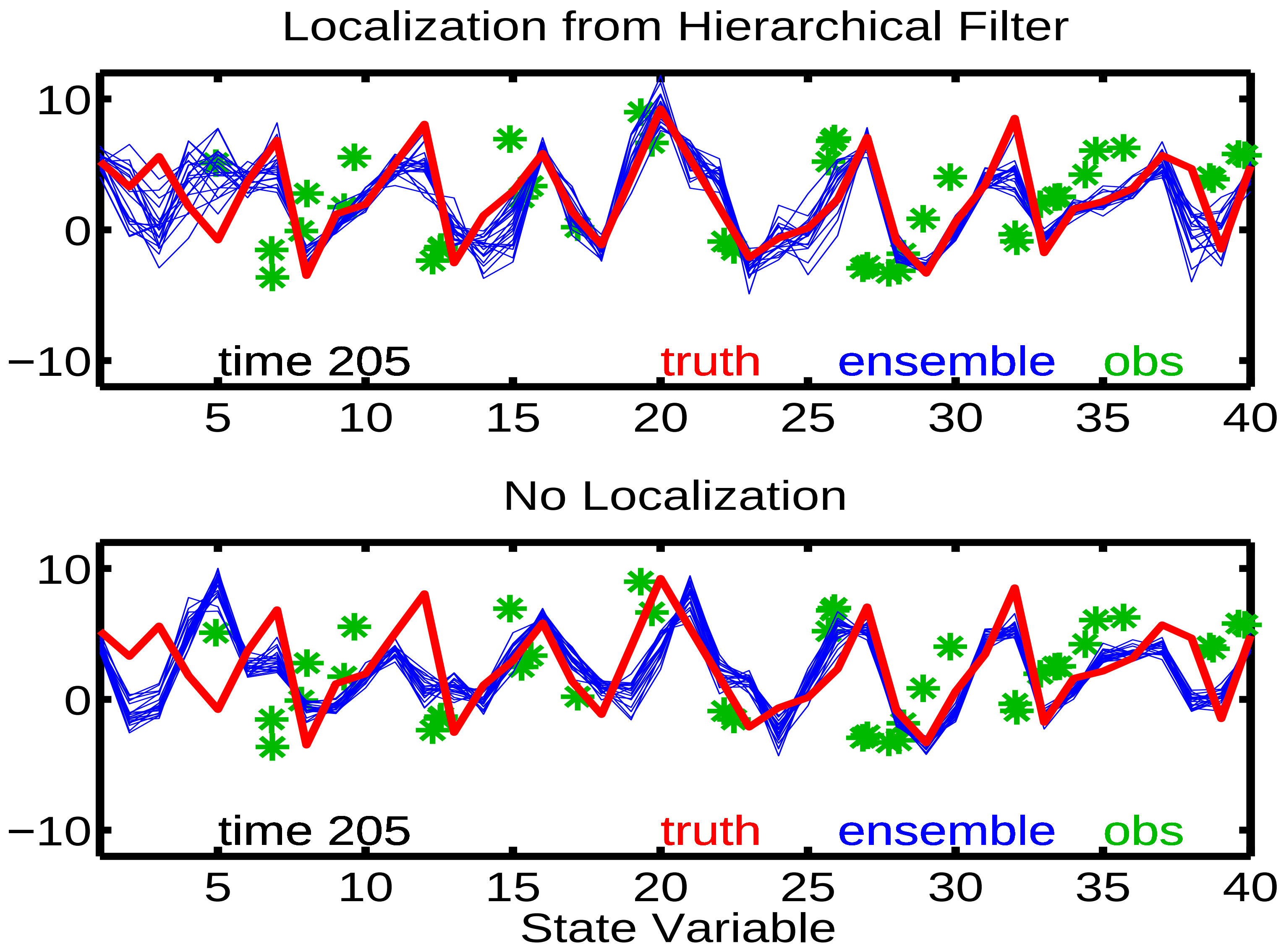
Lorenz-96 Assimilation with localization of observation impact.



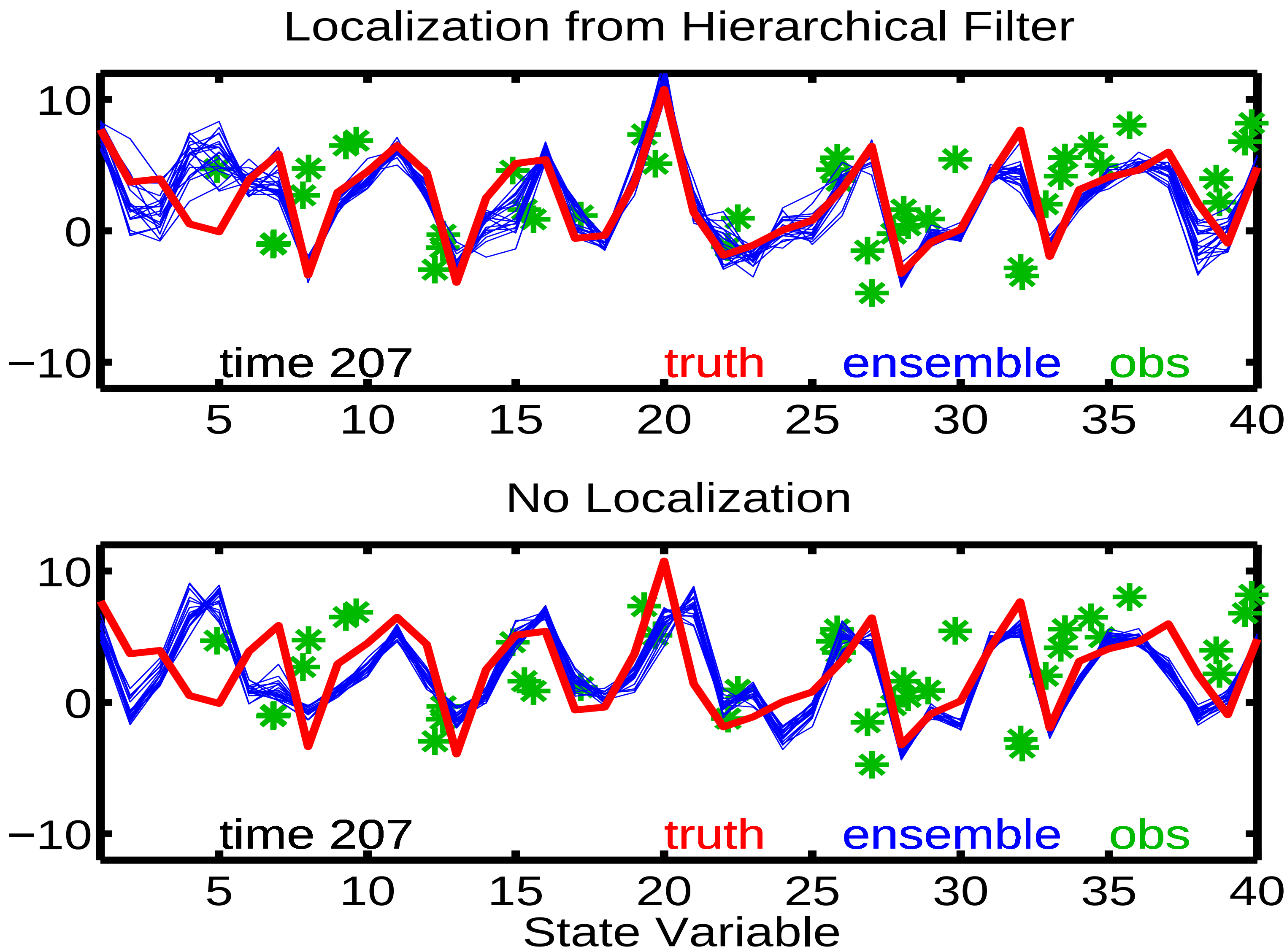
Lorenz-96 Assimilation with localization of observation impact.



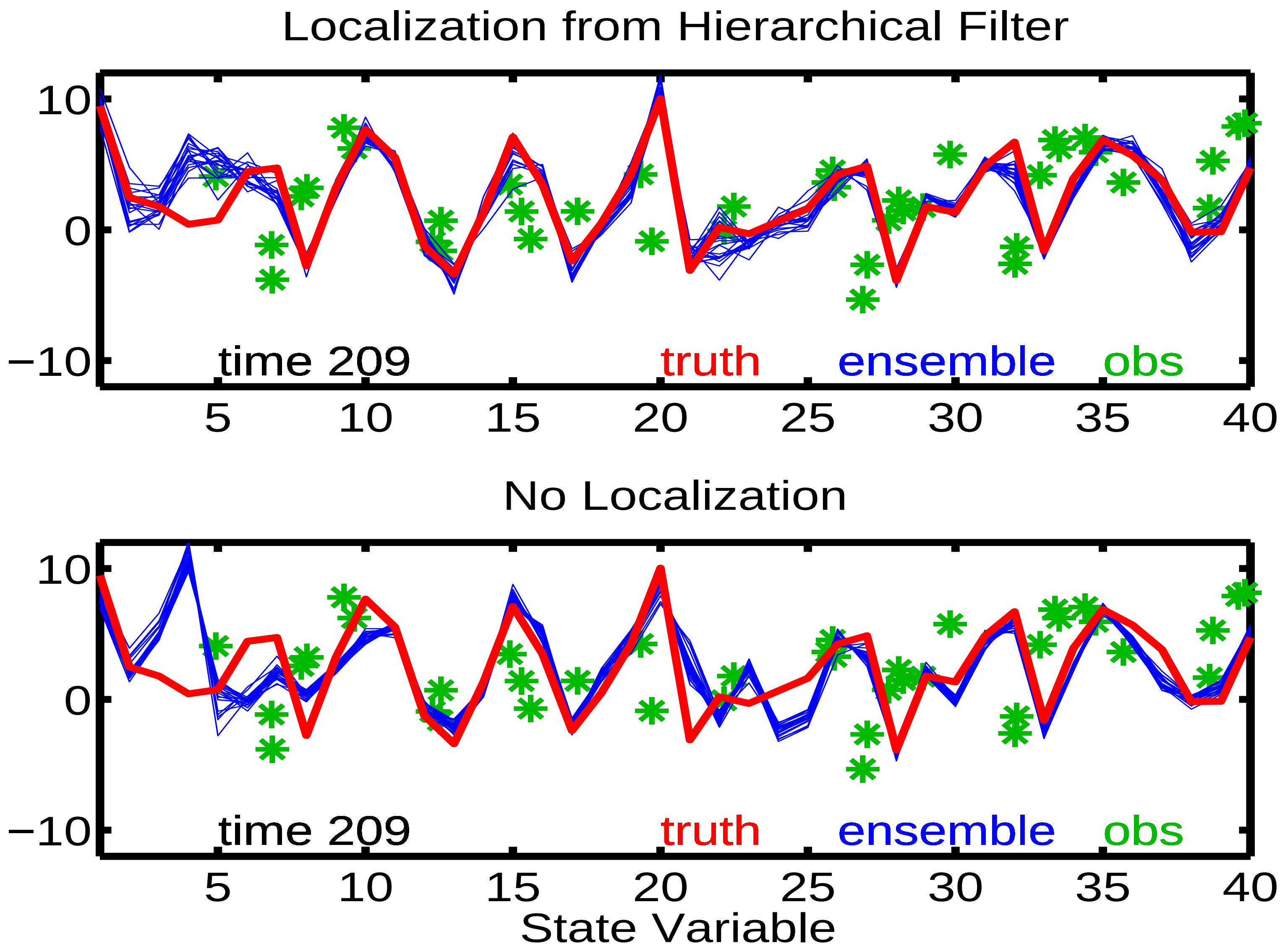
Lorenz-96 Assimilation with localization of observation impact.



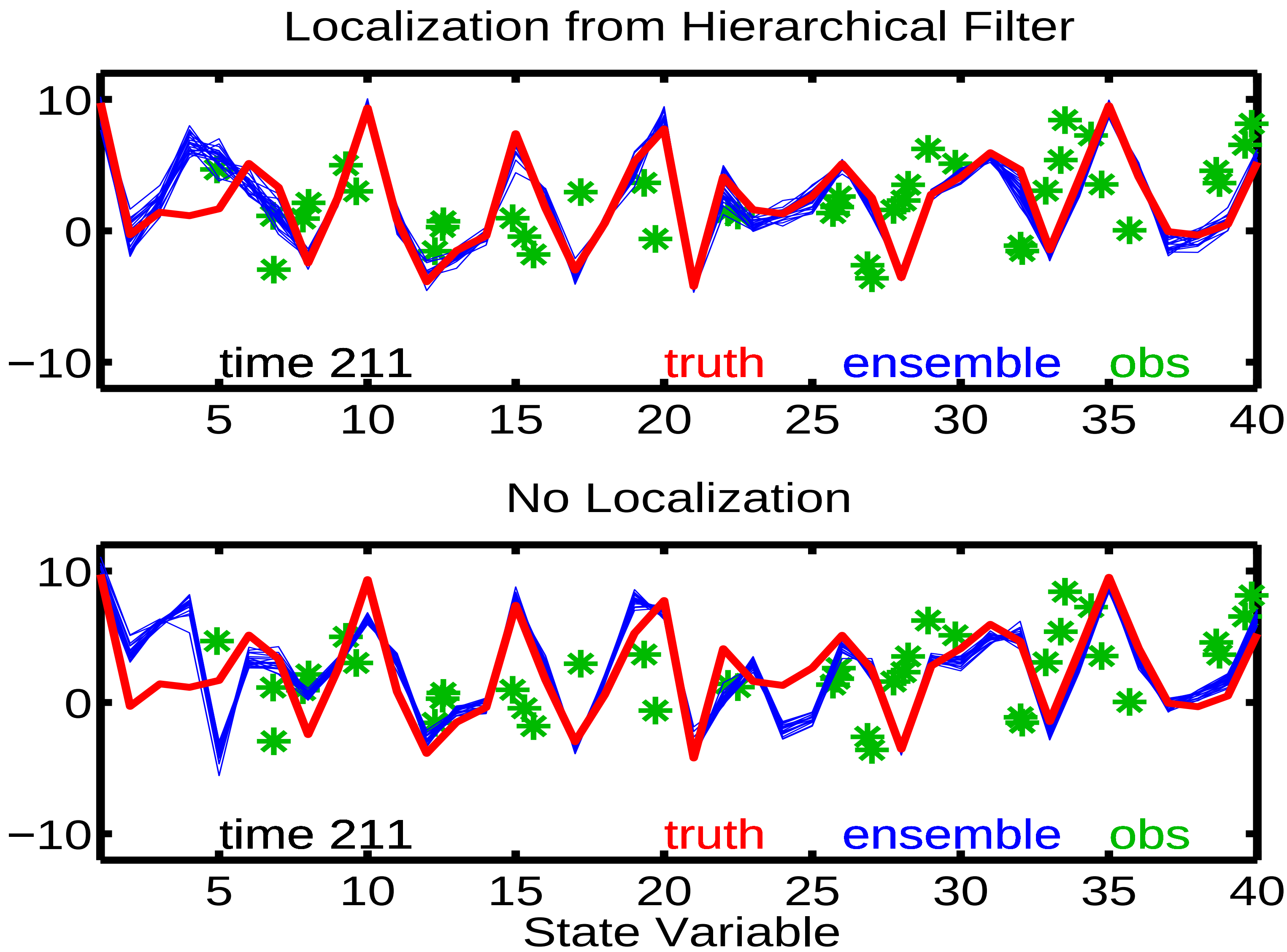
Lorenz-96 Assimilation with localization of observation impact.



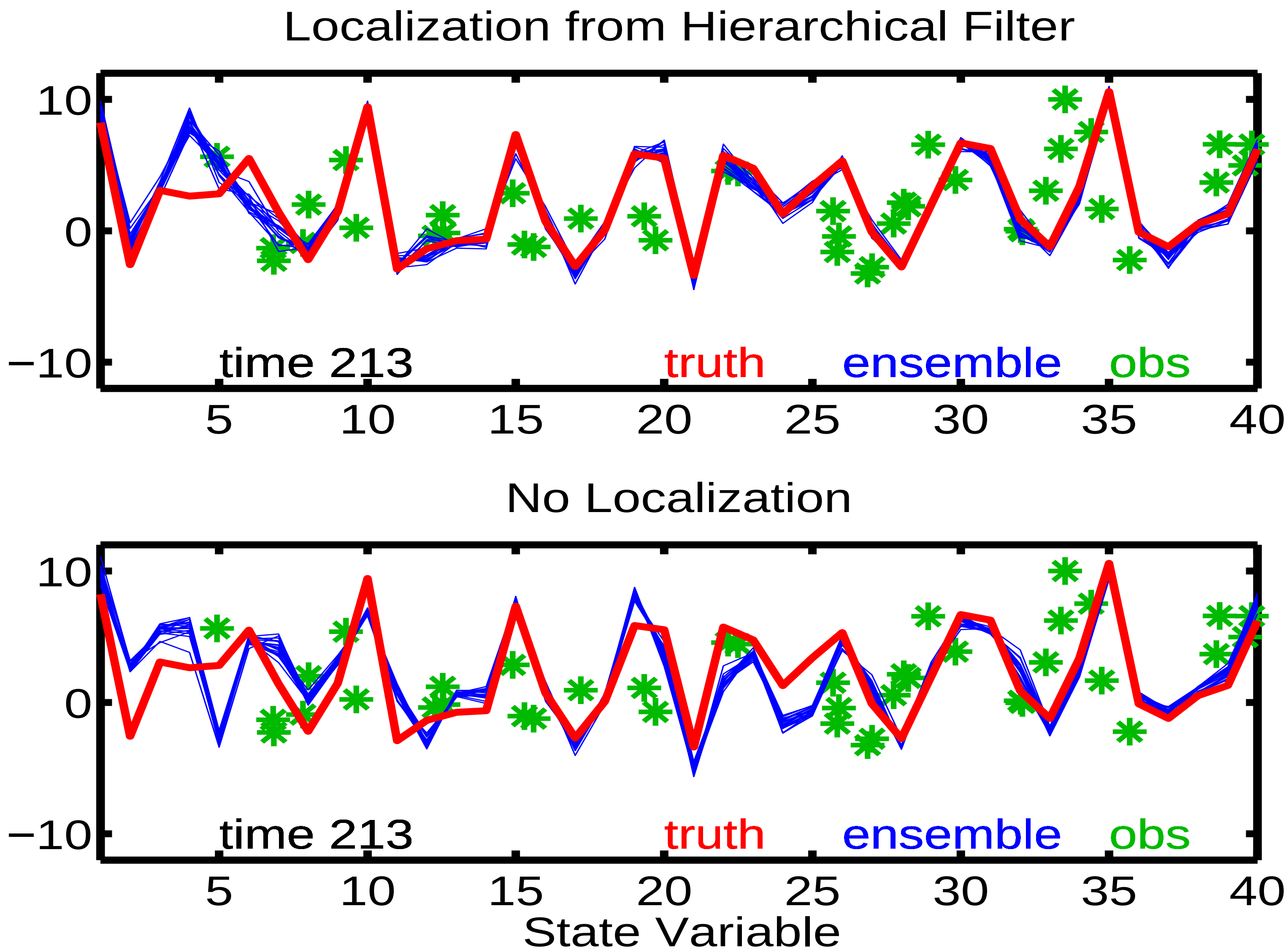
Lorenz-96 Assimilation with localization of observation impact.



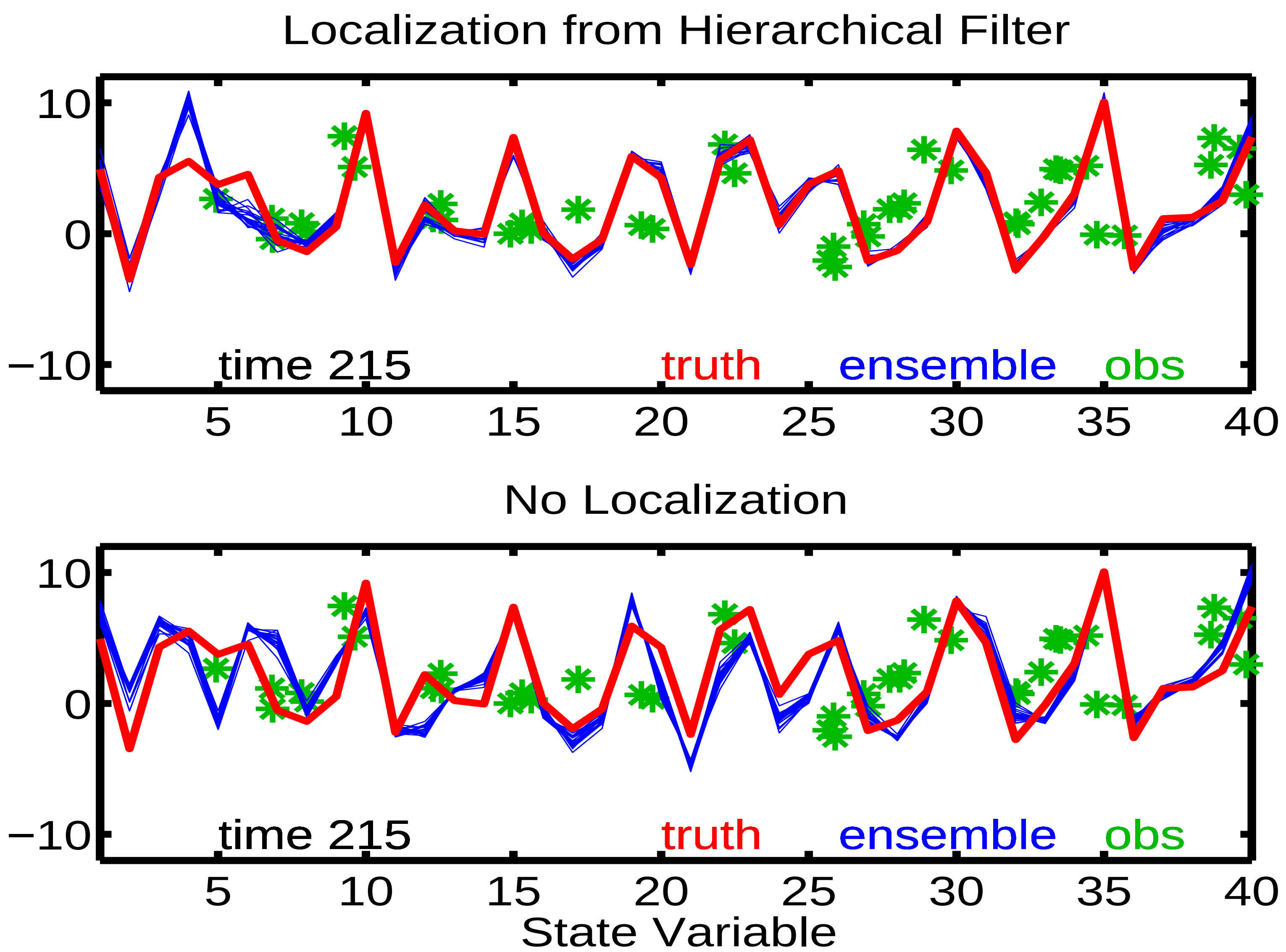
Lorenz-96 Assimilation with localization of observation impact.



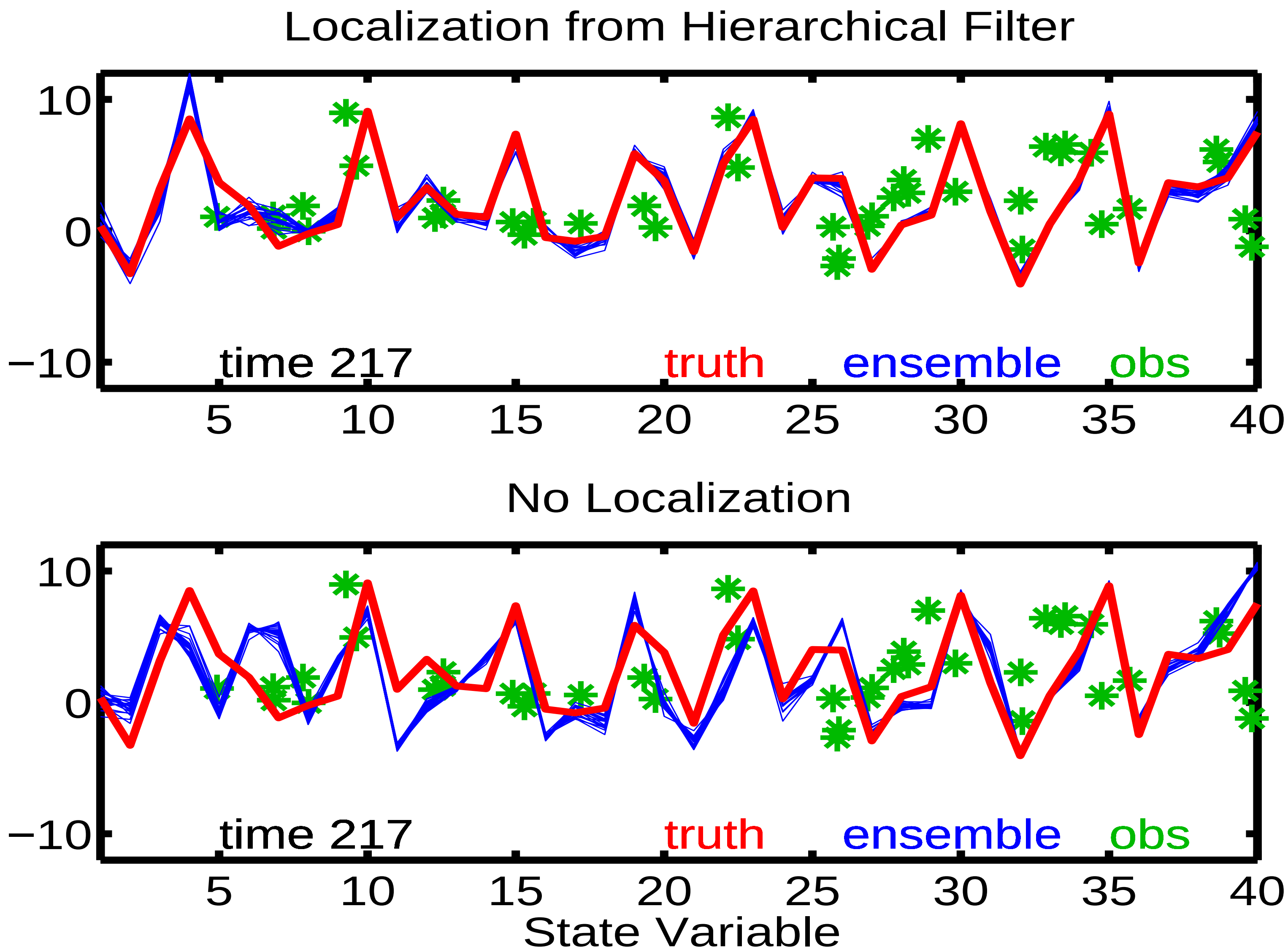
Lorenz-96 Assimilation with localization of observation impact.



Lorenz-96 Assimilation with localization of observation impact.

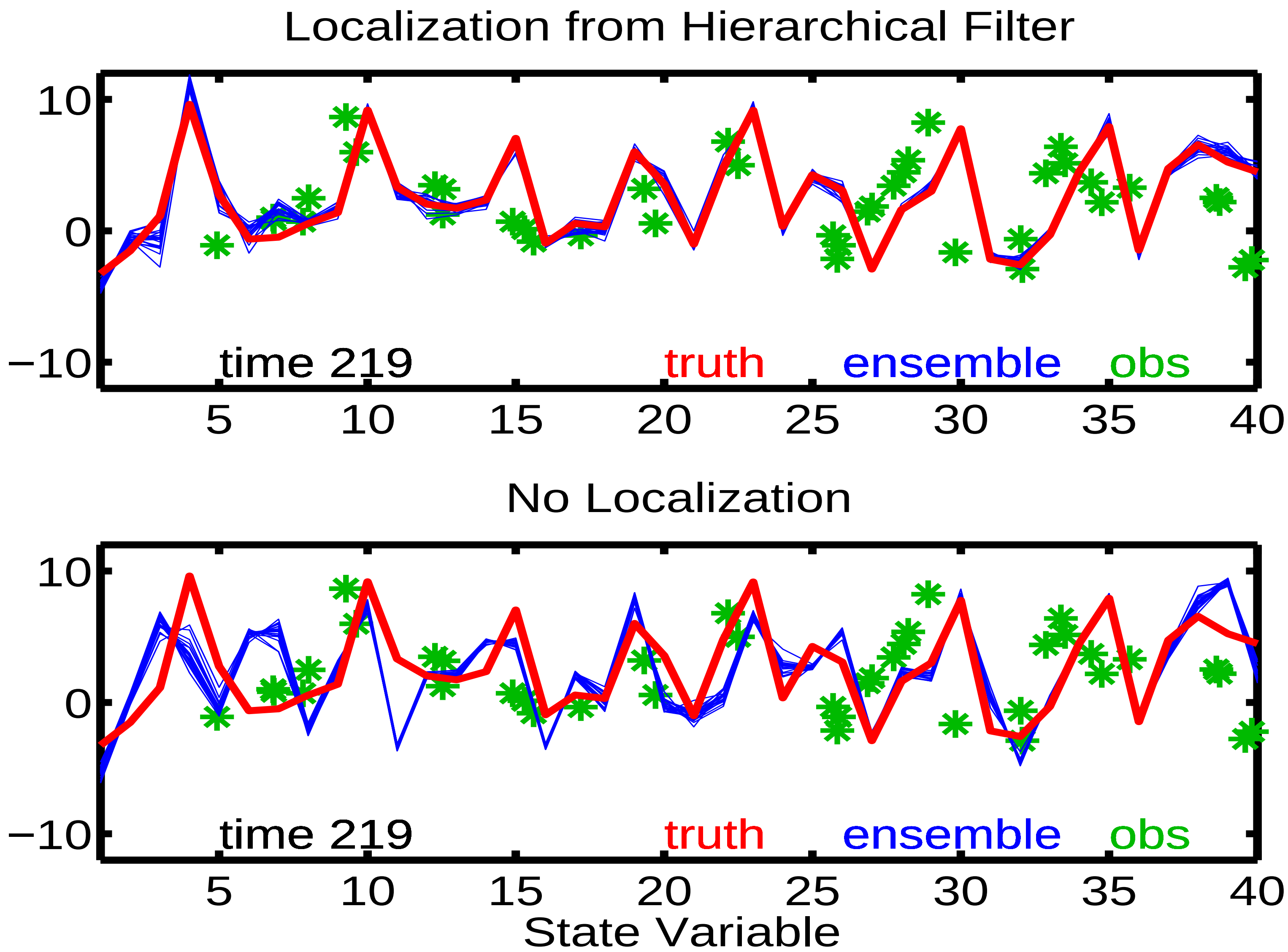


Lorenz-96 Assimilation with localization of observation impact.



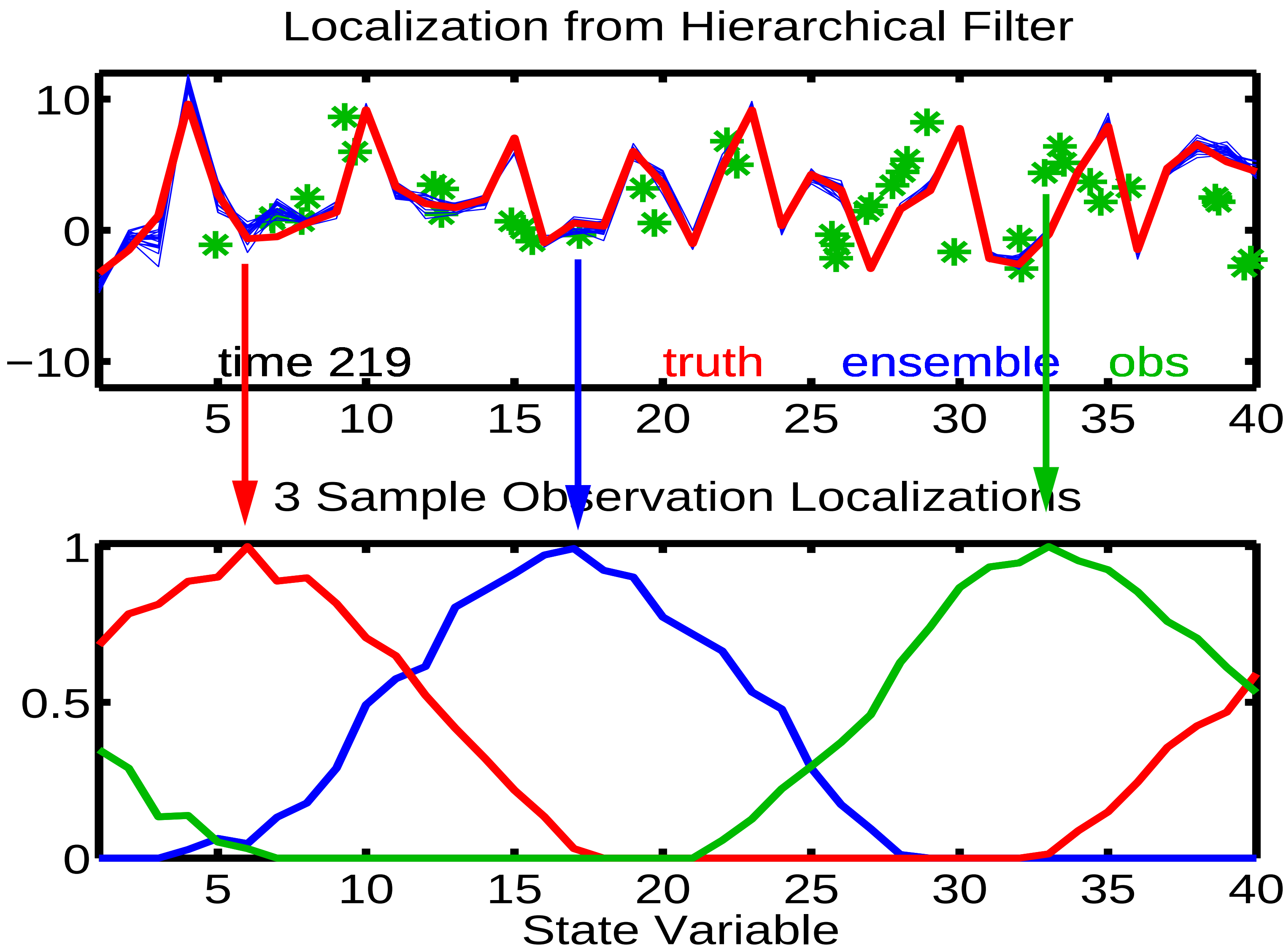
Lorenz-96 Assimilation with localization of observation impact.

Ensemble is much more consistent with truth.

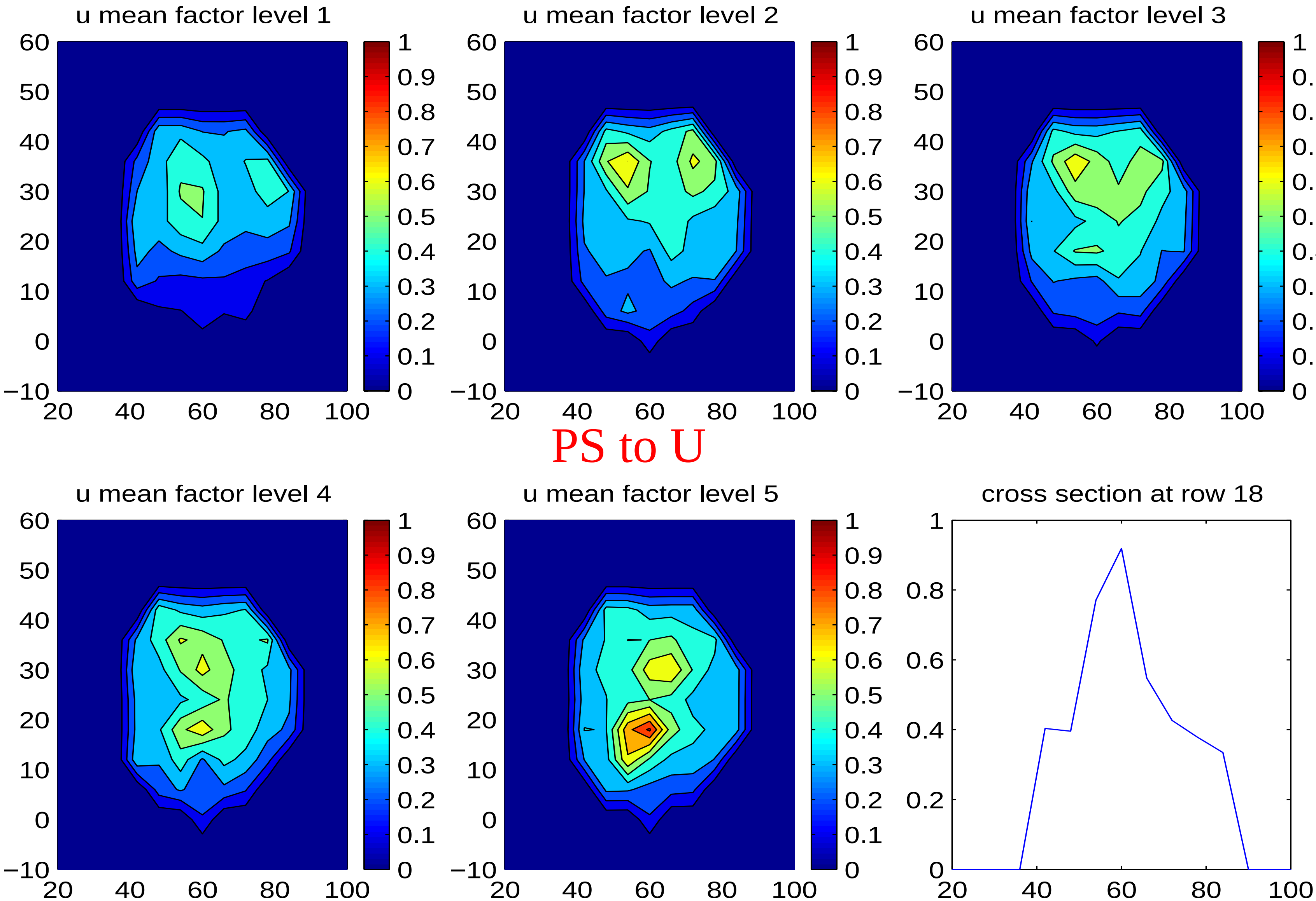


<u>Localization computed by adaptive hierarchical filter.</u>

A tuning run of 4, 20-member ensembles maximizes signal.

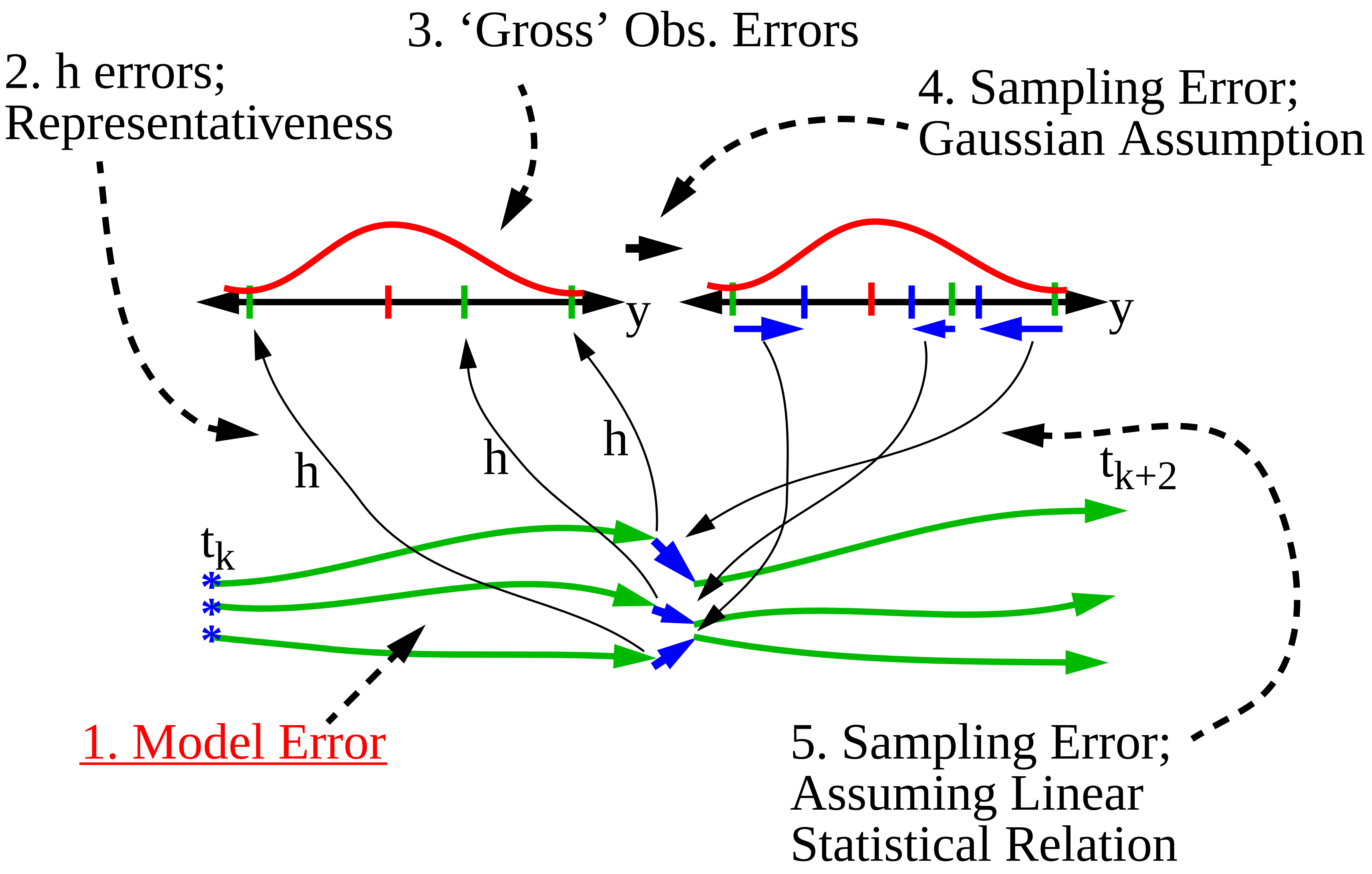


Localization in GCM can be very complex. Surface Pressure Obs. at 20N, 60E



MUST HAVE ADAPTIVE HELP FOR NON-EXPERT USERS.

Some Error Sources in Ensemble Filters

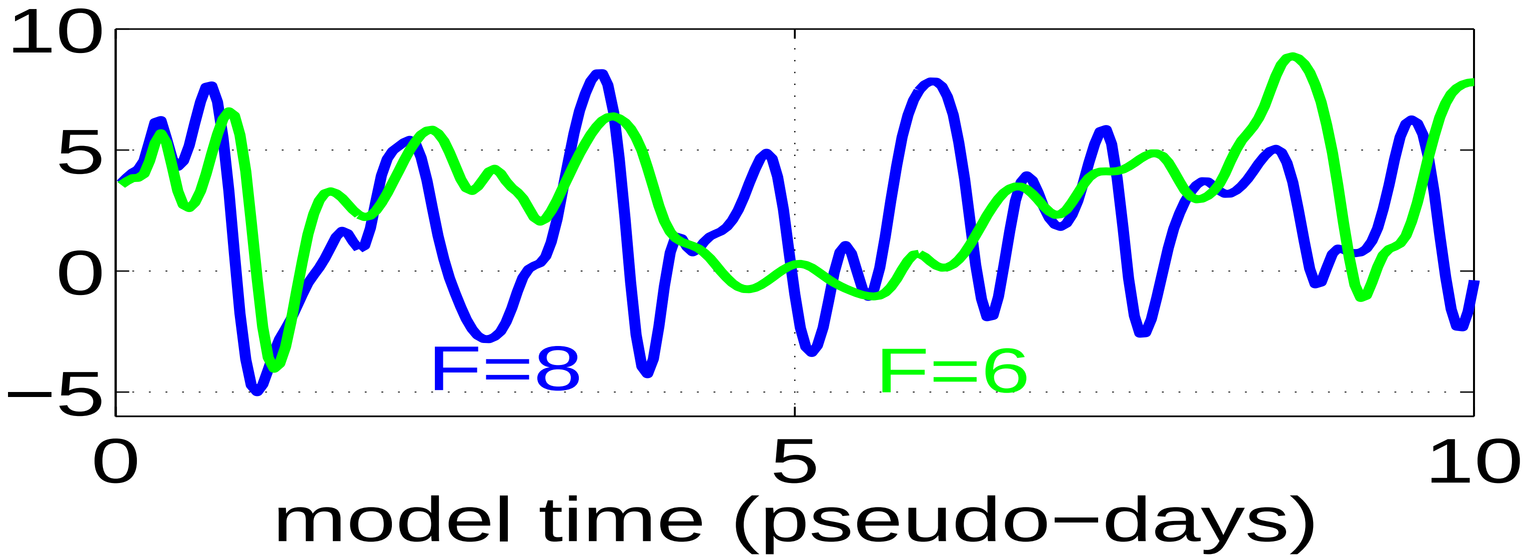


Assimilating in the presence of simulated model error.

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

For truth, use F = 8.

In assimilating model, use F = 6.



Time evolution for X1 shown.

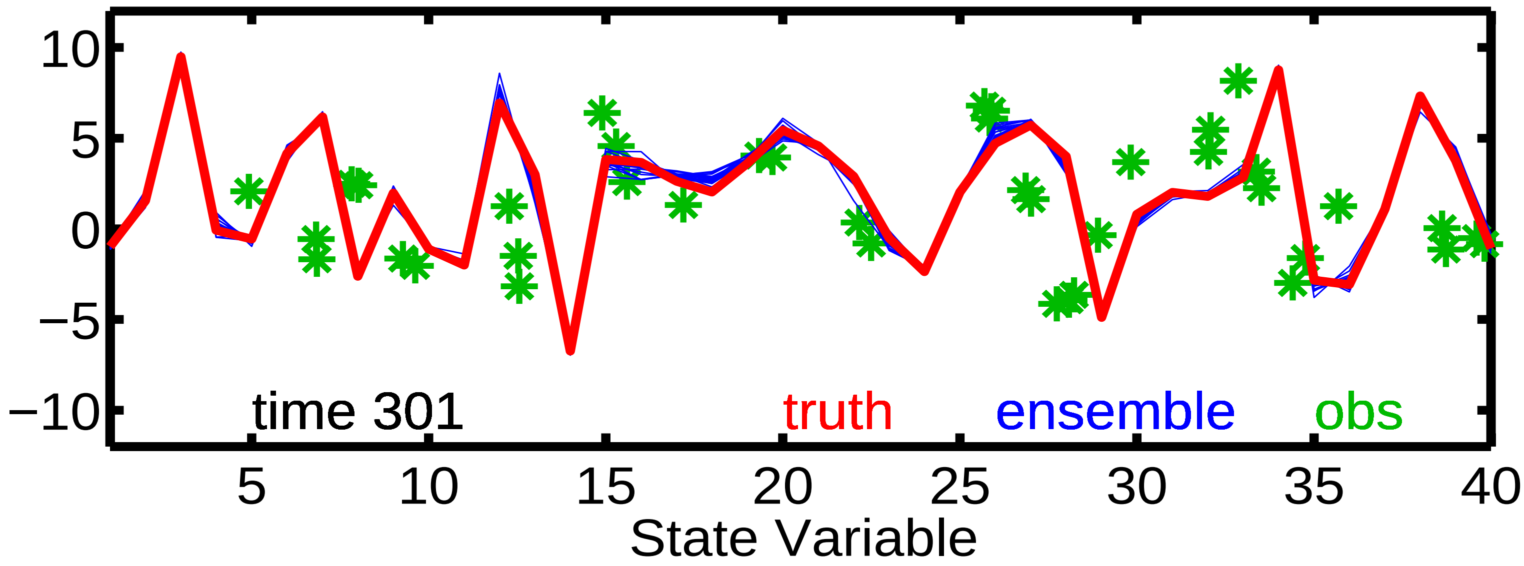
Assimilating model quickly diverges from ‘true’ model.

Assimilating in the presence of simulated model error

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

For truth, use F = 8.

In assimilating model, use F = 6.

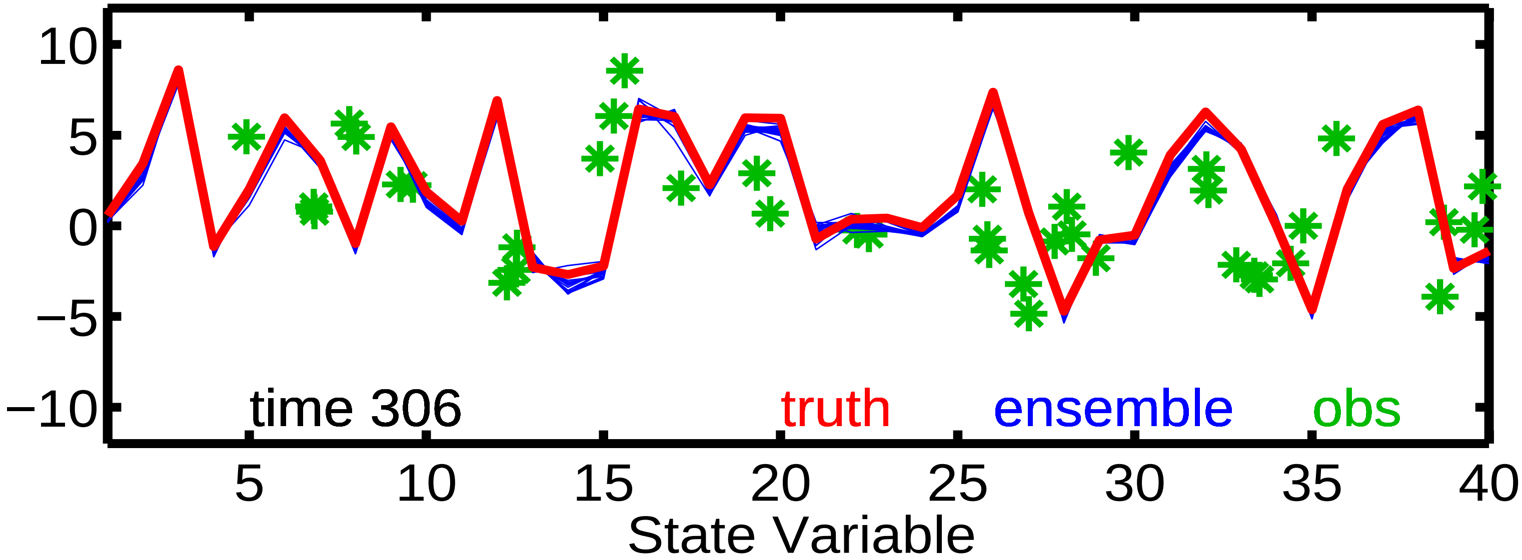


Assimilating in the presence of simulated model error

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

For truth, use F = 8.

In assimilating model, use F = 6.

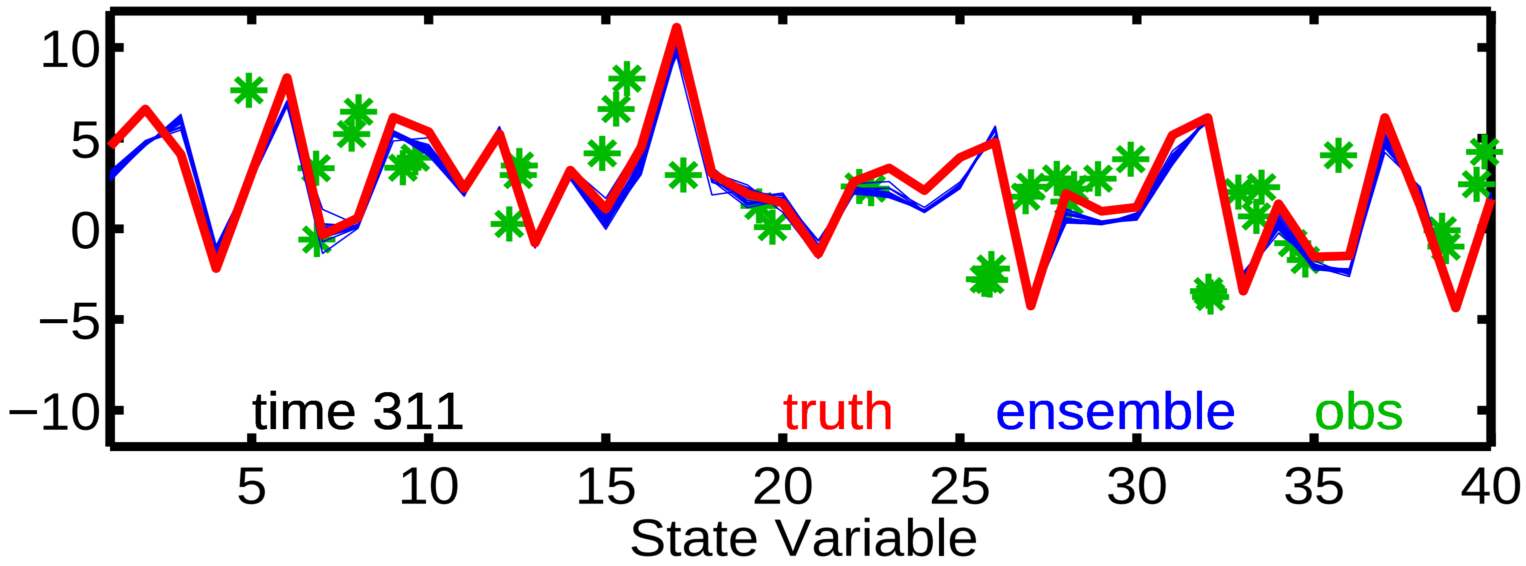


Assimilating in the presence of simulated model error

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

For truth, use F = 8.

In assimilating model, use F = 6.

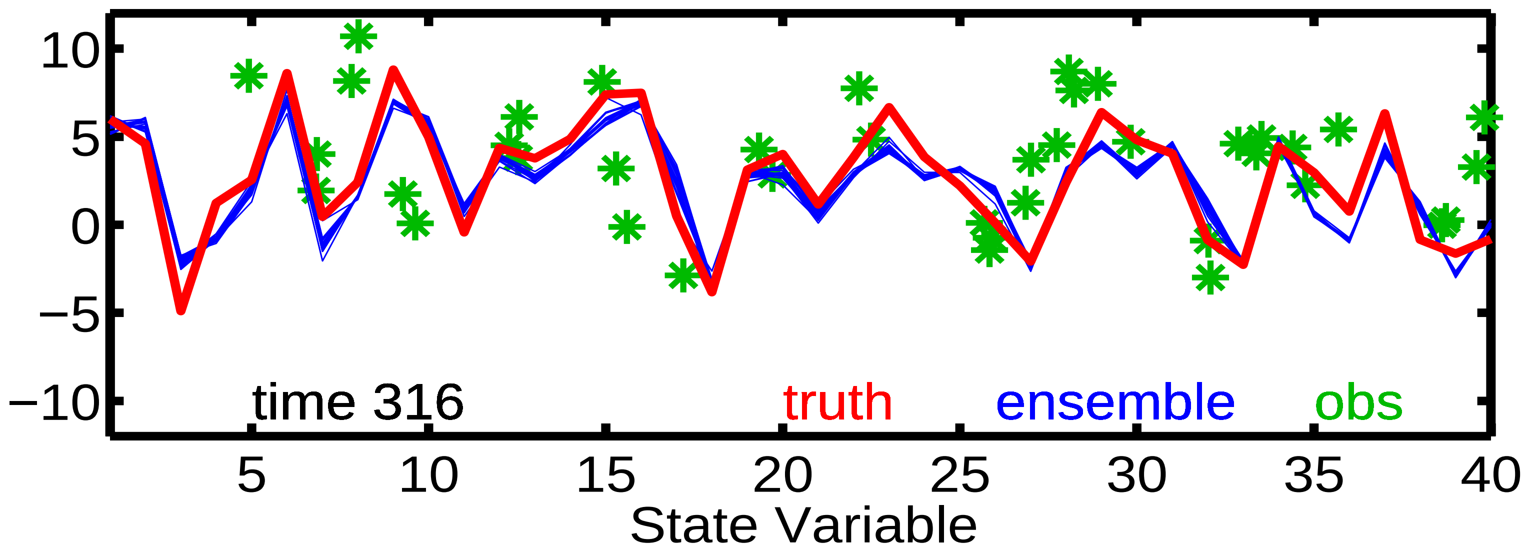


Assimilating in the presence of simulated model error

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

For truth, use F = 8.

In assimilating model, use F = 6.

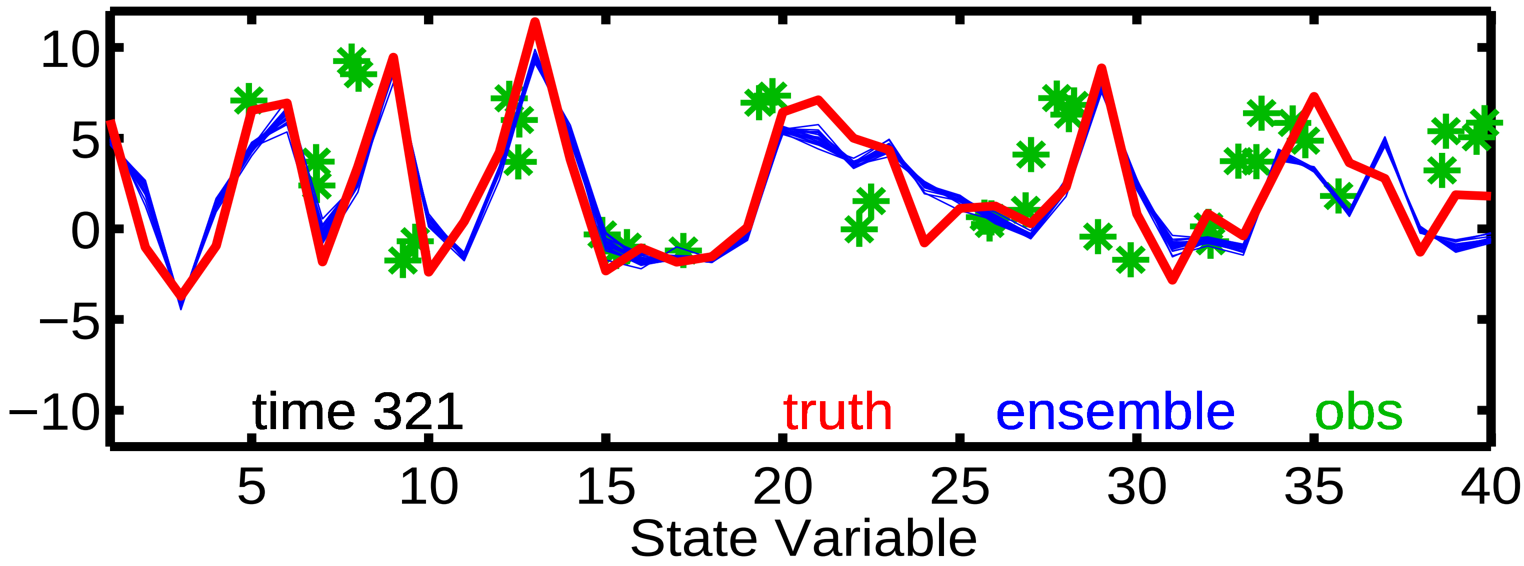


Assimilating in the presence of simulated model error

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

For truth, use F = 8.

In assimilating model, use F = 6.

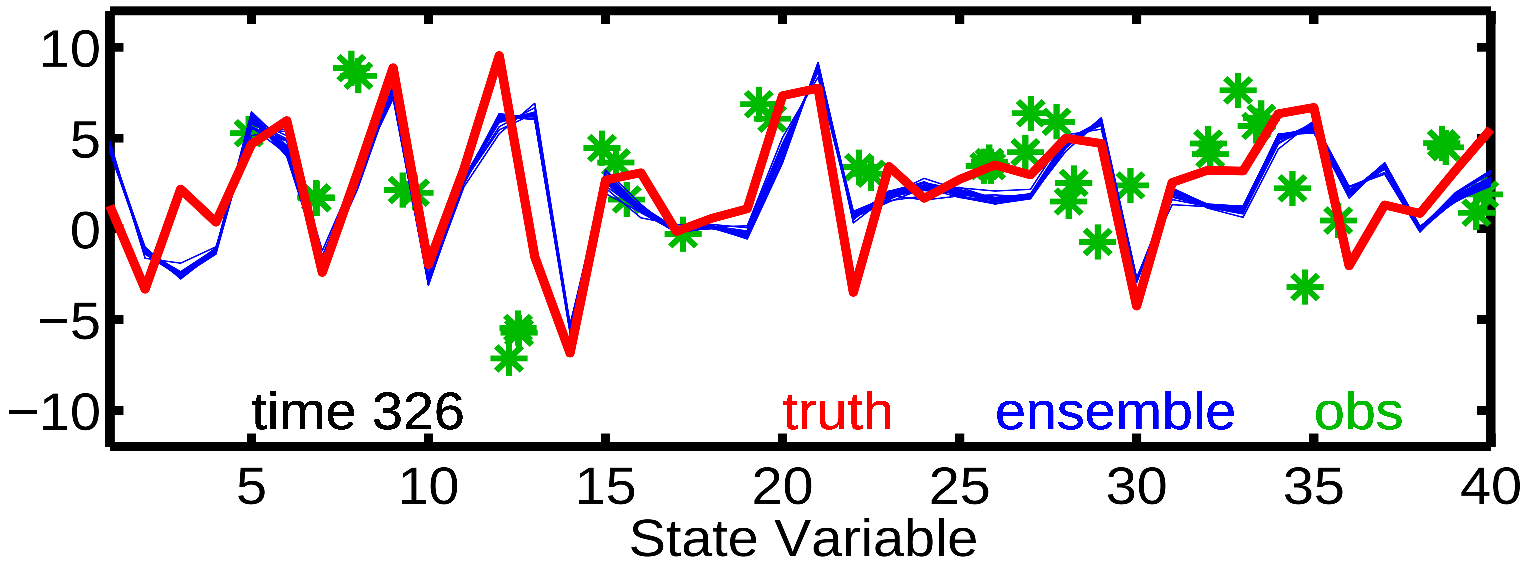


Assimilating in the presence of simulated model error

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

For truth, use F = 8.

In assimilating model, use F = 6.

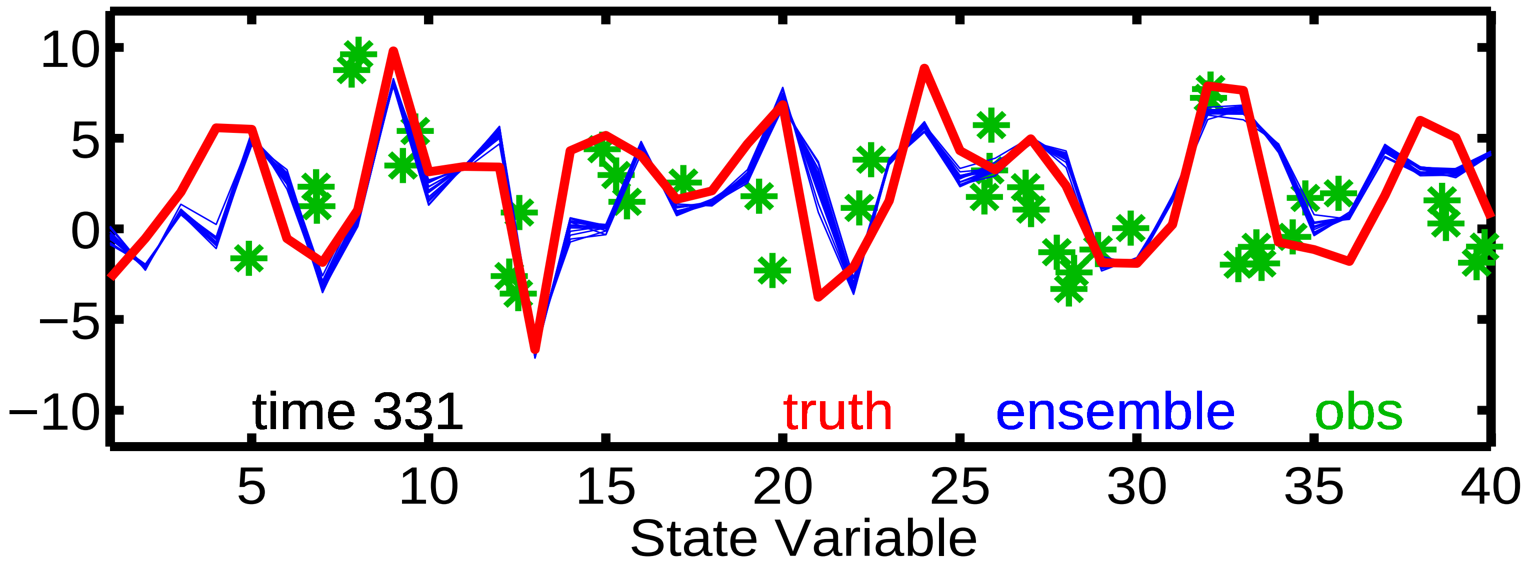


Assimilating in the presence of simulated model error

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

For truth, use F = 8.

In assimilating model, use F = 6.

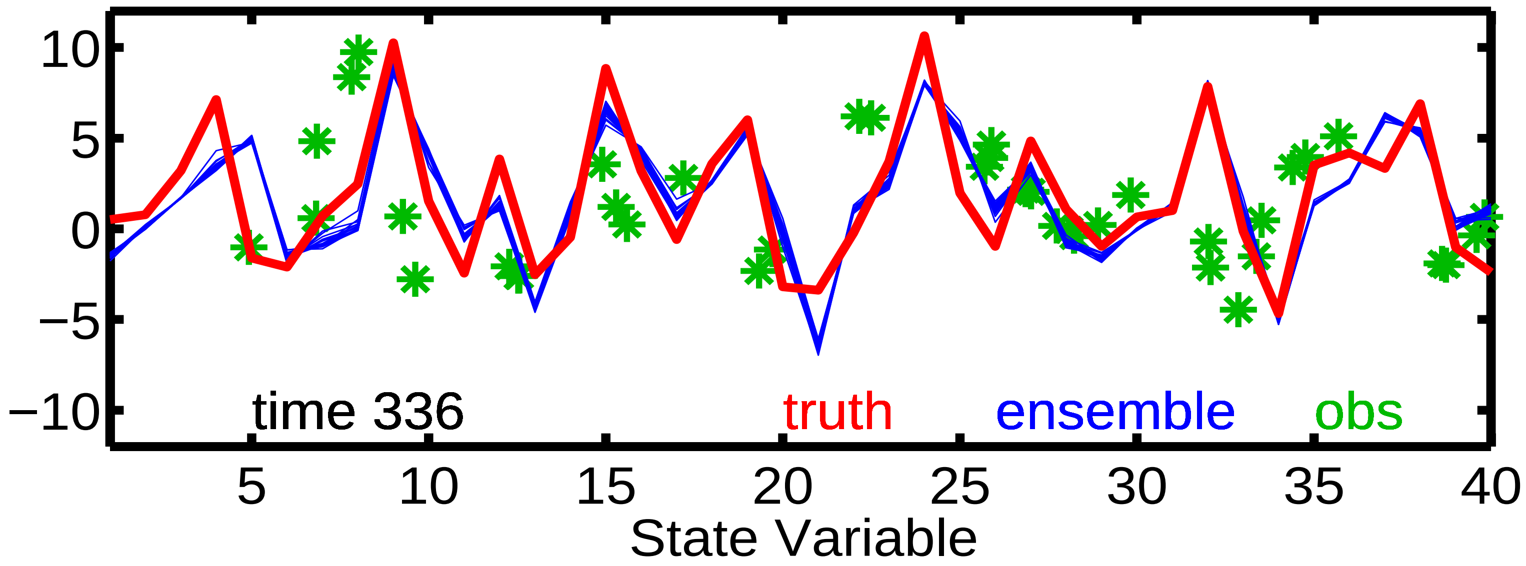


Assimilating in the presence of simulated model error

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

For truth, use F = 8.

In assimilating model, use F = 6.

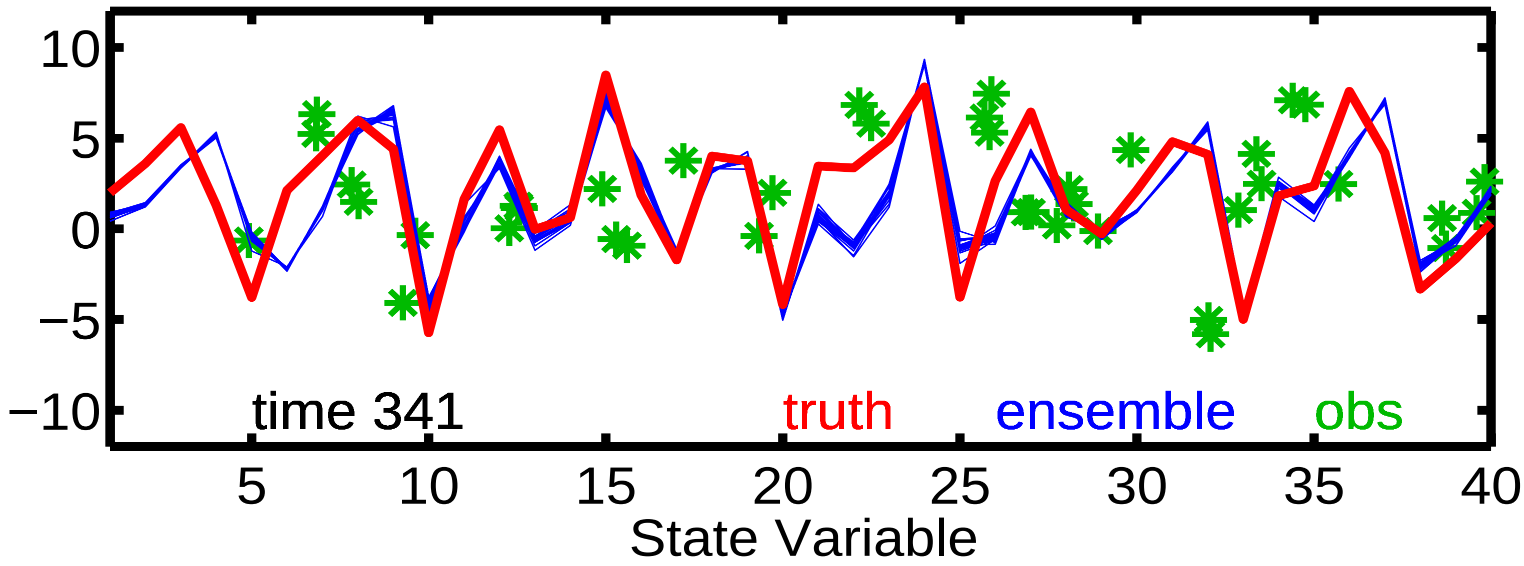


Assimilating in the presence of simulated model error

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

For truth, use F = 8.

In assimilating model, use F = 6.

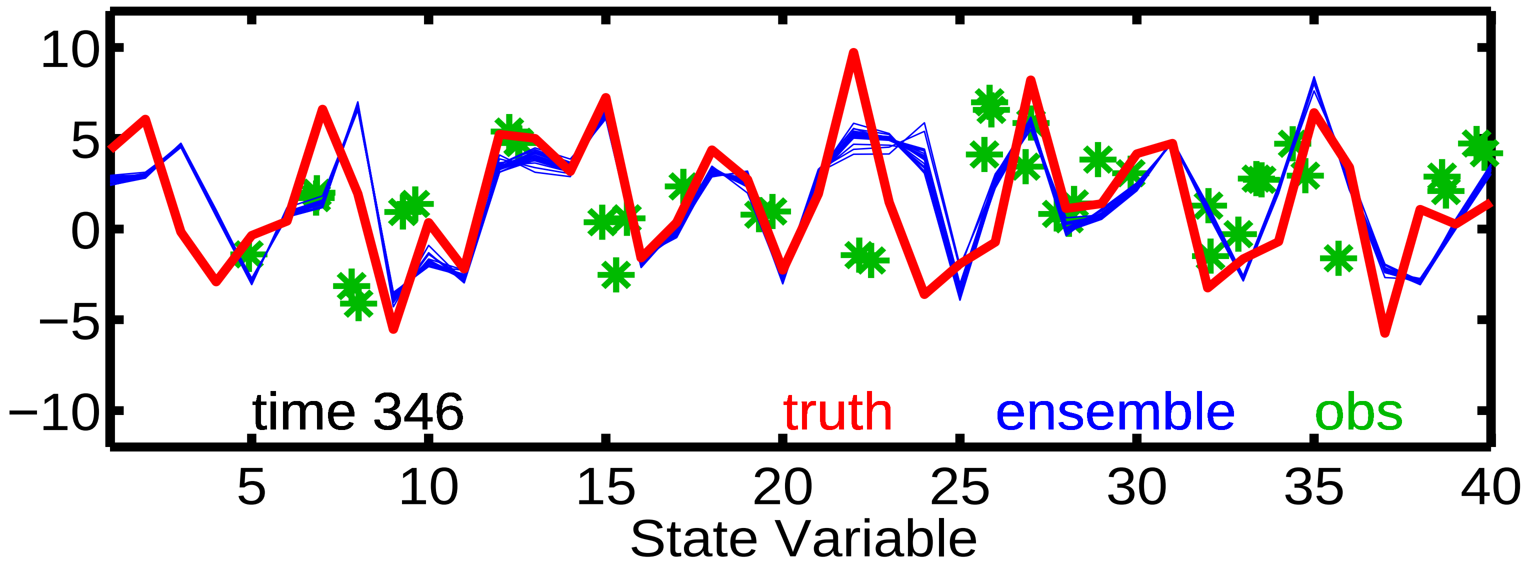


Assimilating in the presence of simulated model error

$dX_i / dt = (X_{i+1} - X_{i-2})X_{i-1} - X_i + F.$

For truth, use F = 8.

In assimilating model, use F = 6.

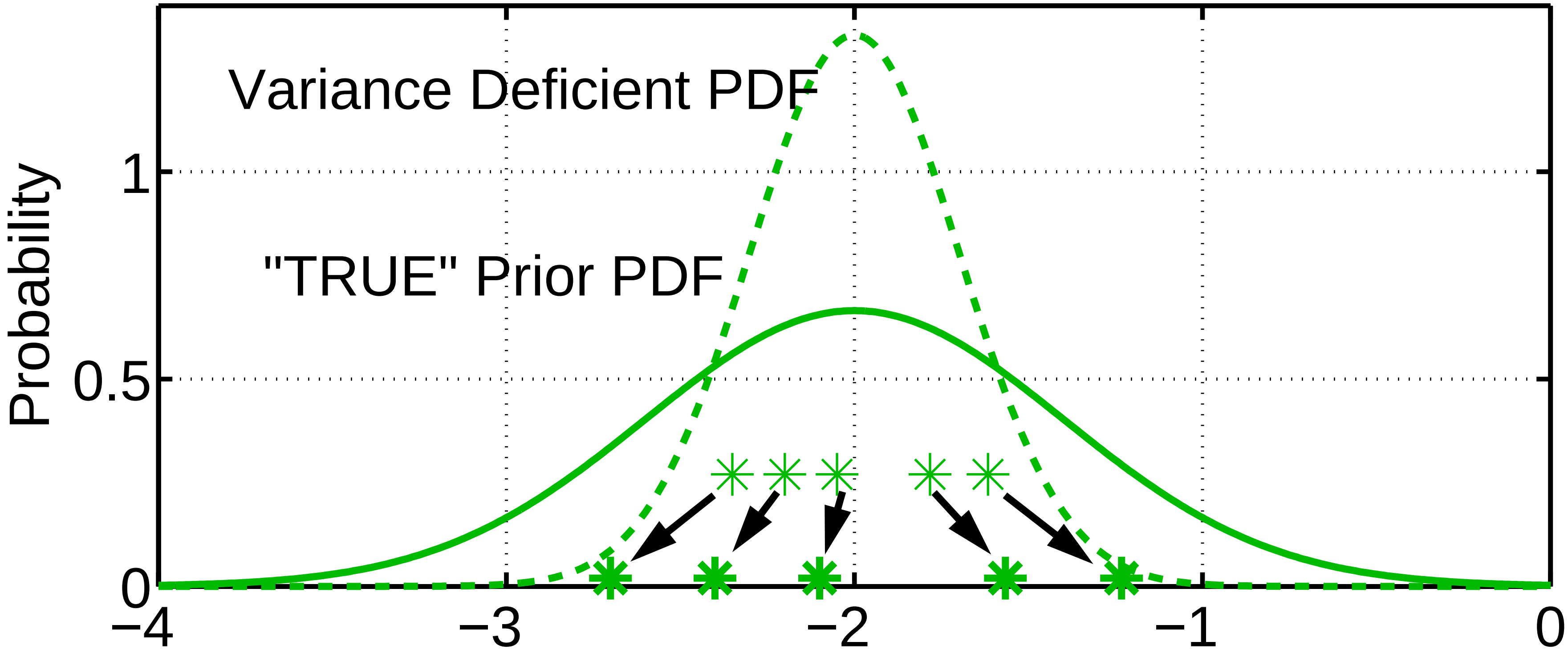


This isn't working again!

It will just keep getting worse.

Model/Filter Error; Filter Divergence and Variance Inflation

1. History of observations and physical system => 'true' distribution.
2. Sampling error, some model errors lead to insufficient prior variance.

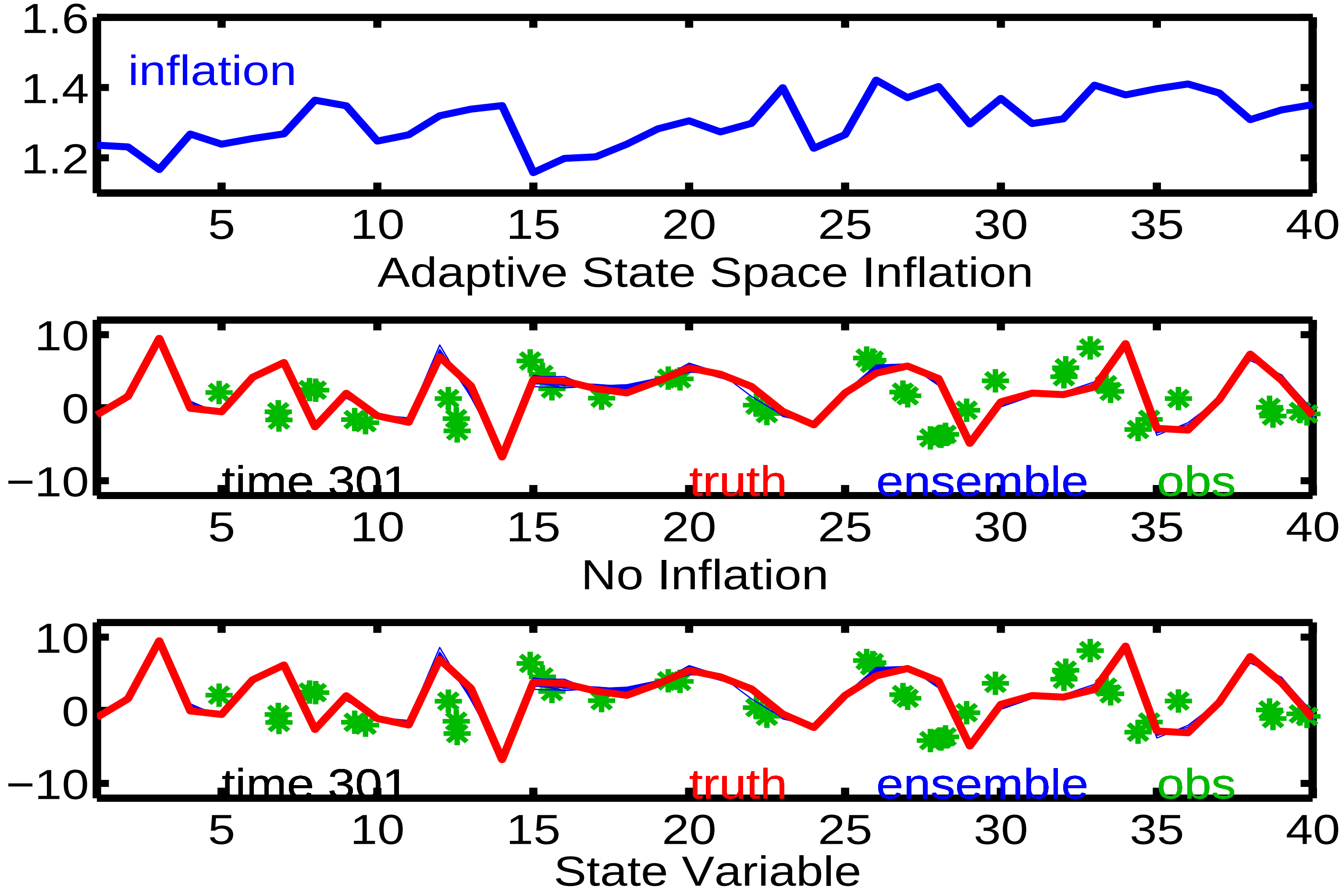


3. Naive solution is Variance inflation: just increase spread of prior
4. For ensemble member i, $inflate(x_i) = \sqrt{\lambda}(x_i - \bar{x}) + \bar{x}$.

Assimilating with Inflation in presence of model error

Inflation is a function of state variable and time.

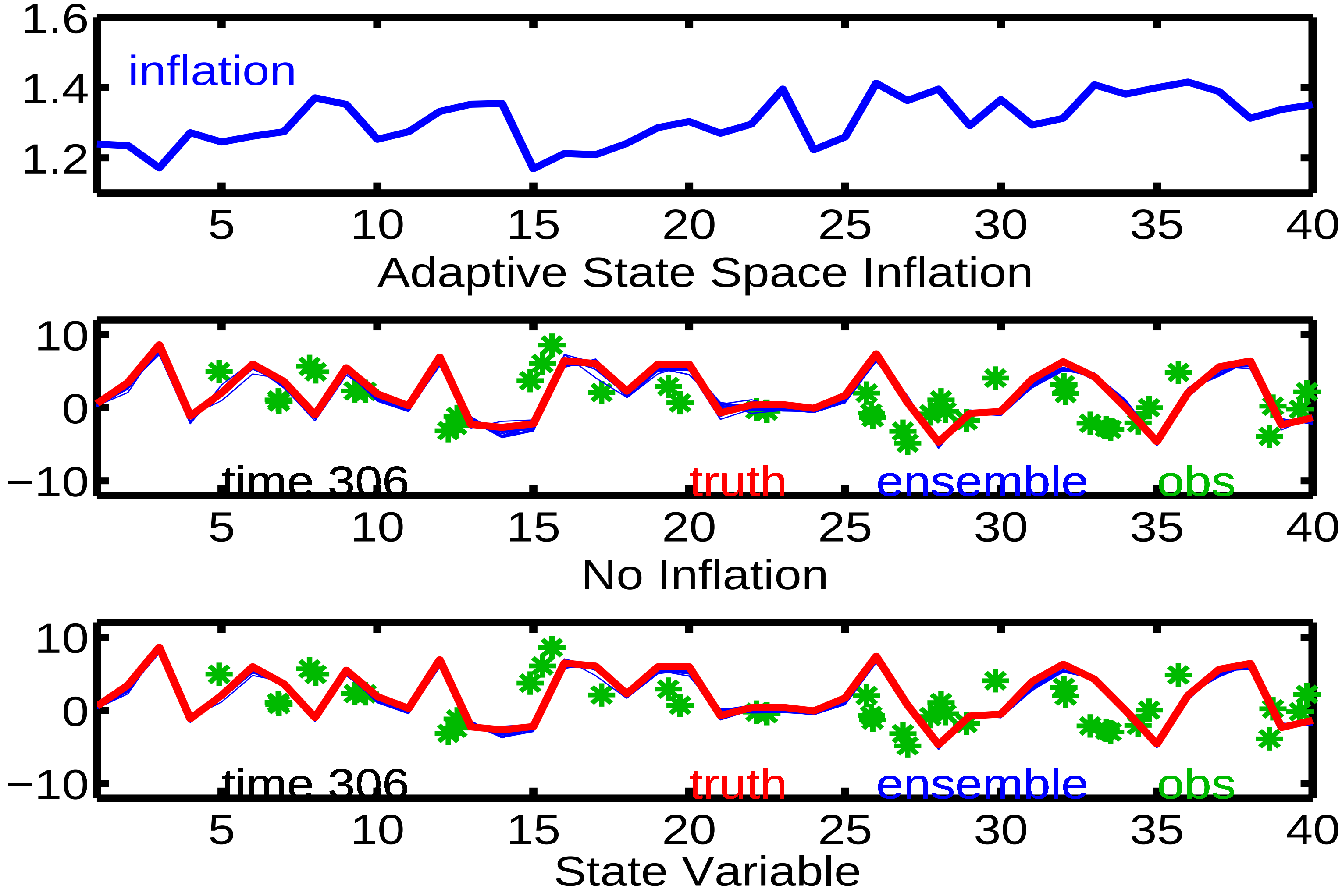
Automatically selected by adaptive inflation algorithm.



Assimilating with Inflation in presence of model error

Inflation is a function of state variable and time.

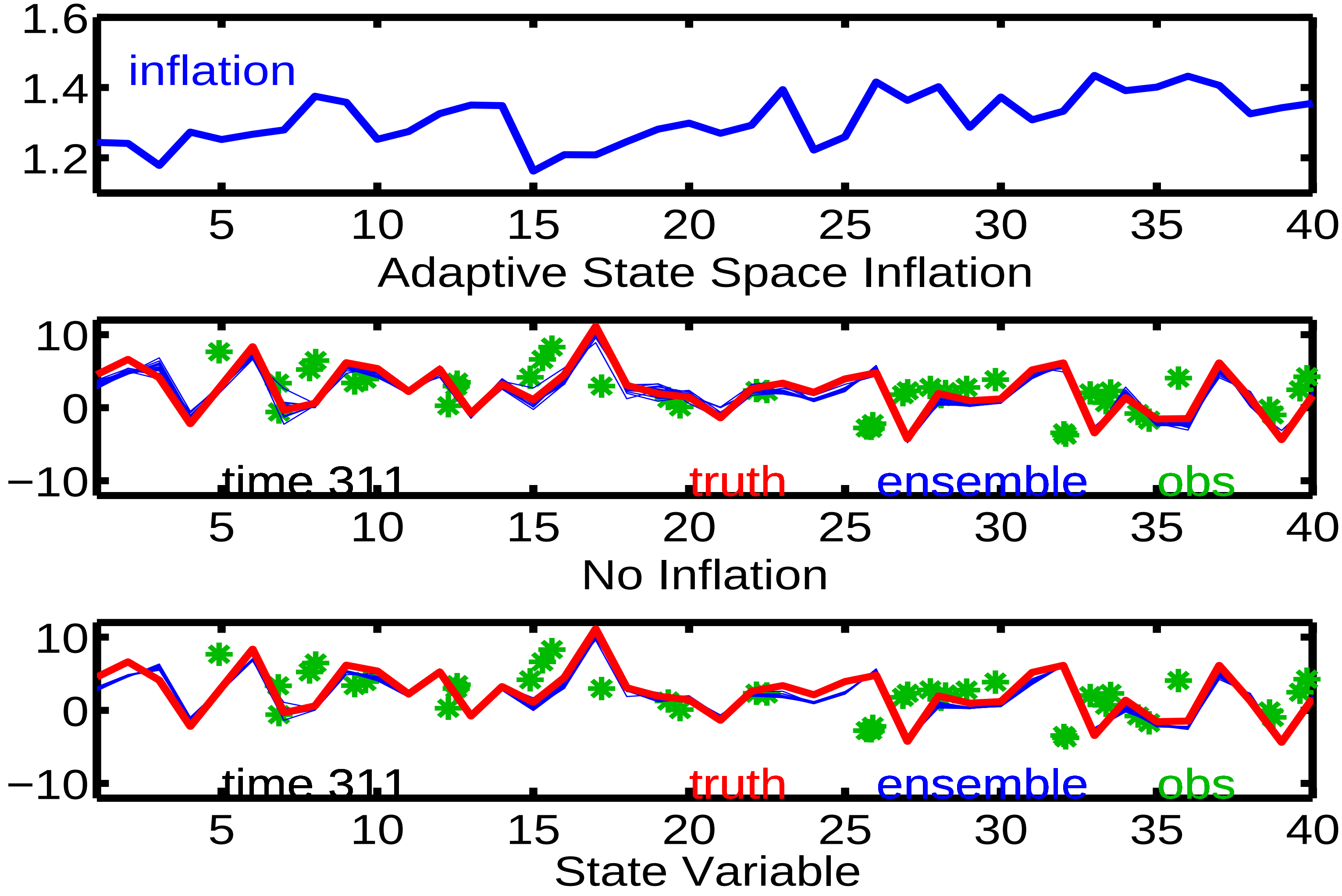
Automatically selected by adaptive inflation algorithm.



Assimilating with Inflation in presence of model error

Inflation is a function of state variable and time.

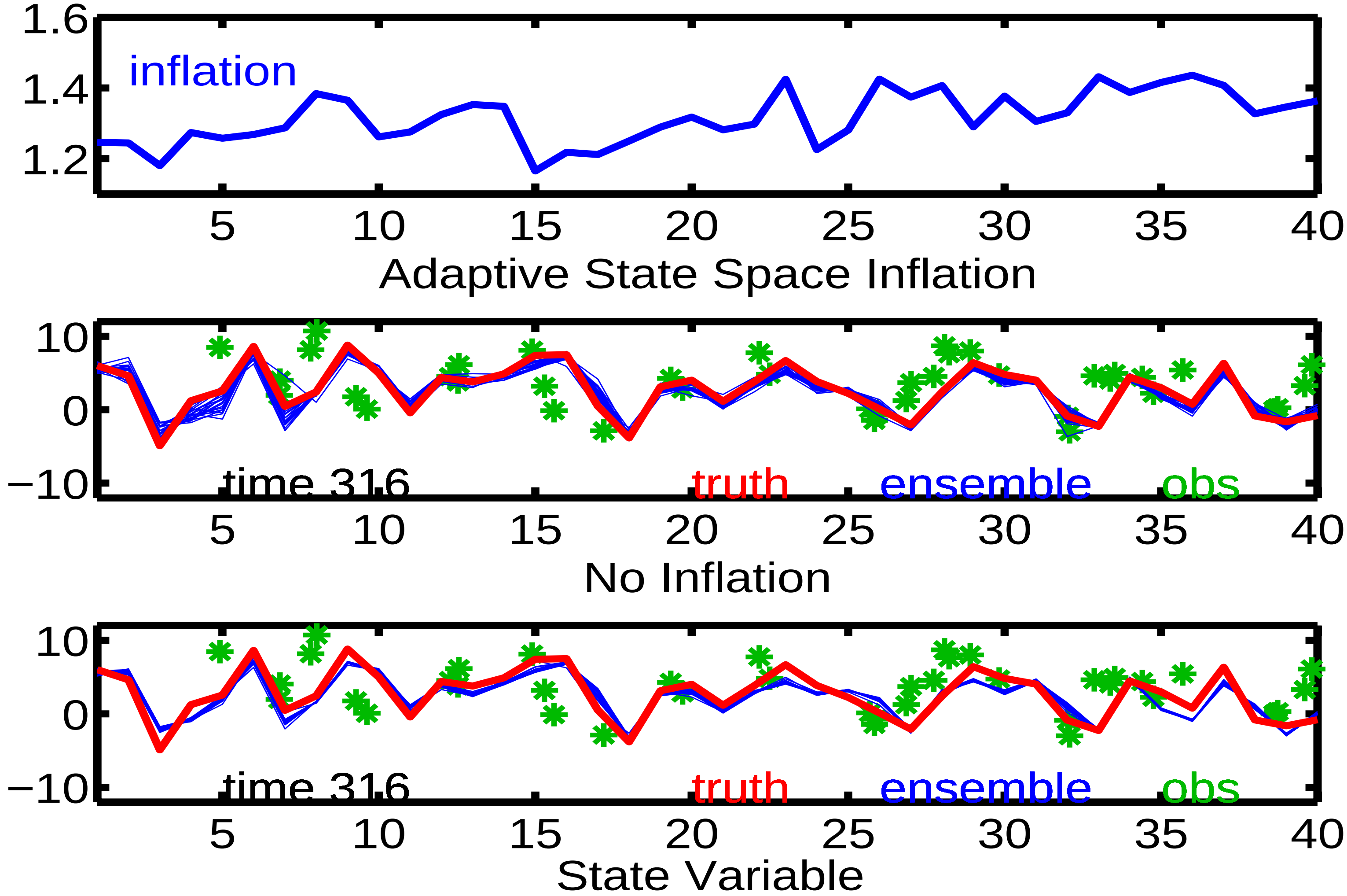
Automatically selected by adaptive inflation algorithm.



Assimilating with Inflation in presence of model error

Inflation is a function of state variable and time.

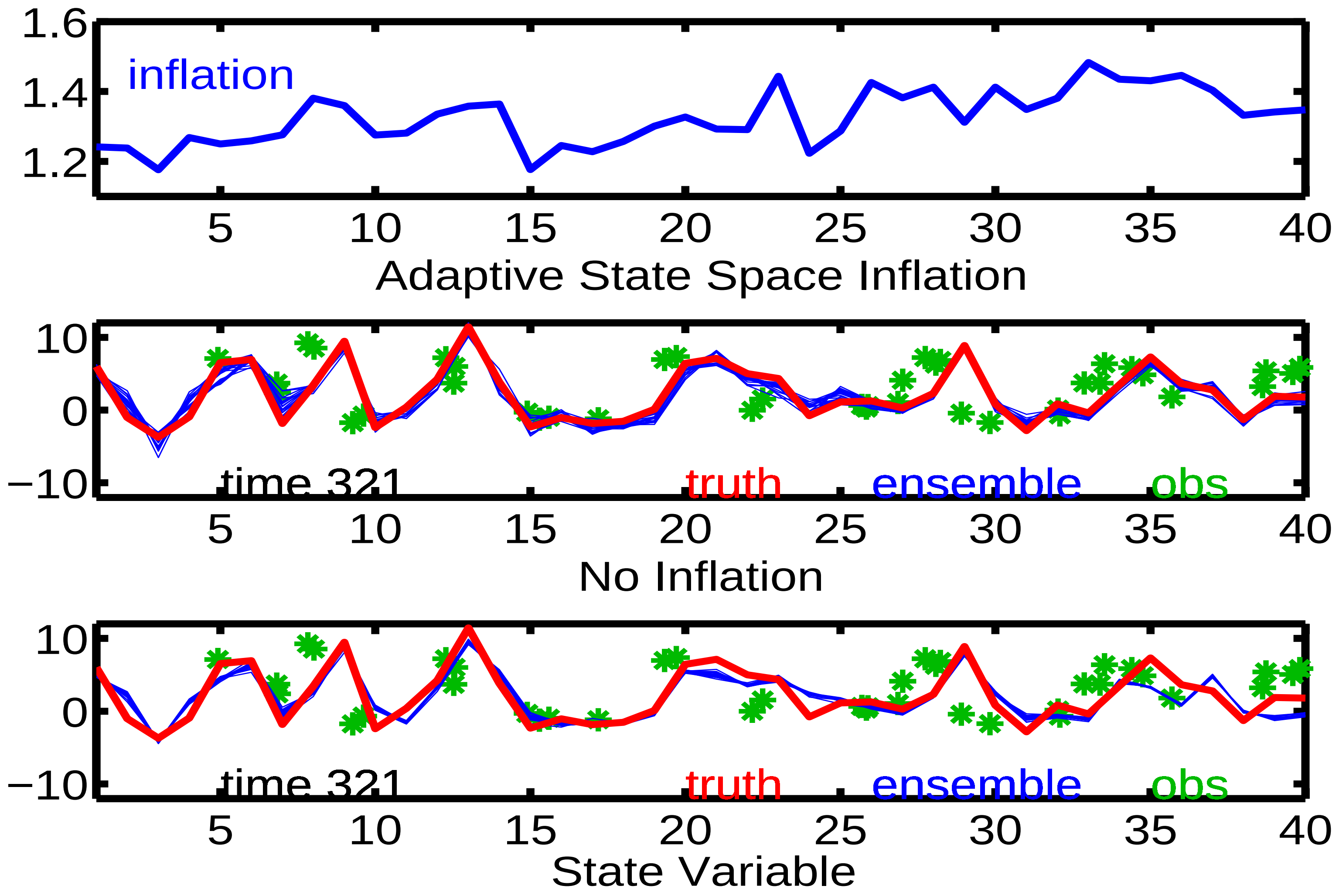
Automatically selected by adaptive inflation algorithm.



Assimilating with Inflation in presence of model error

Inflation is a function of state variable and time.

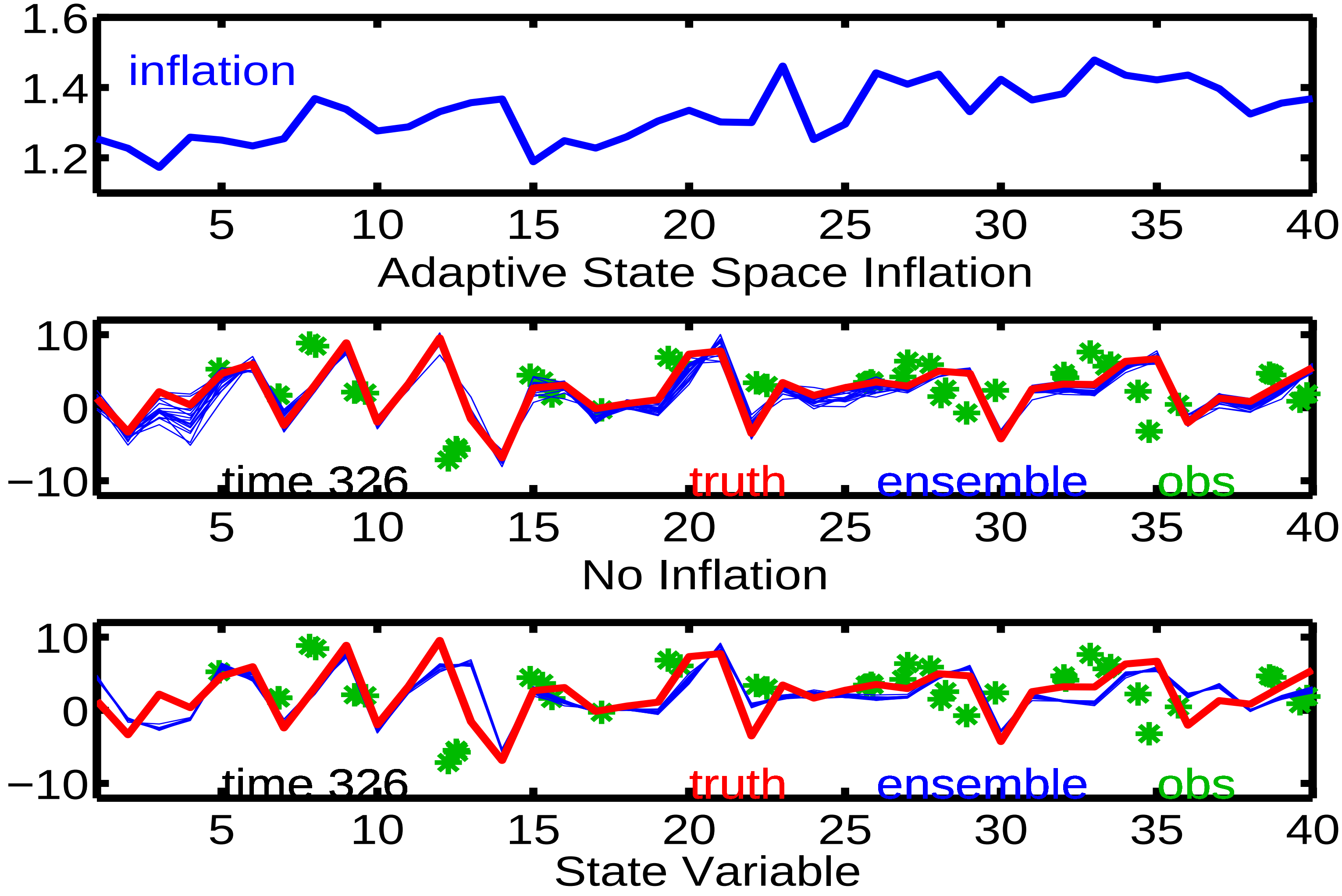
Automatically selected by adaptive inflation algorithm.



Assimilating with Inflation in presence of model error

Inflation is a function of state variable and time.

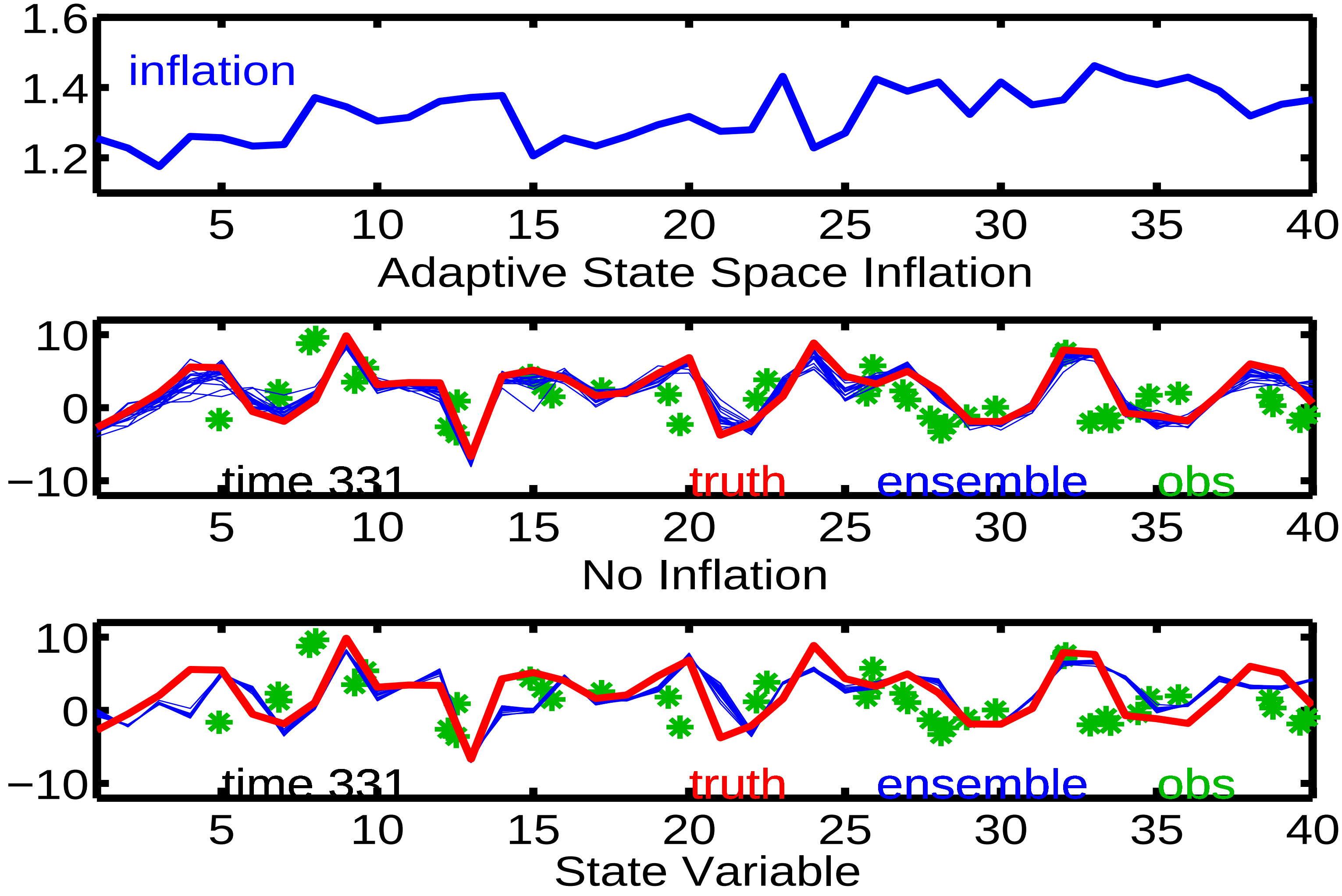
Automatically selected by adaptive inflation algorithm.



Assimilating with Inflation in presence of model error

Inflation is a function of state variable and time.

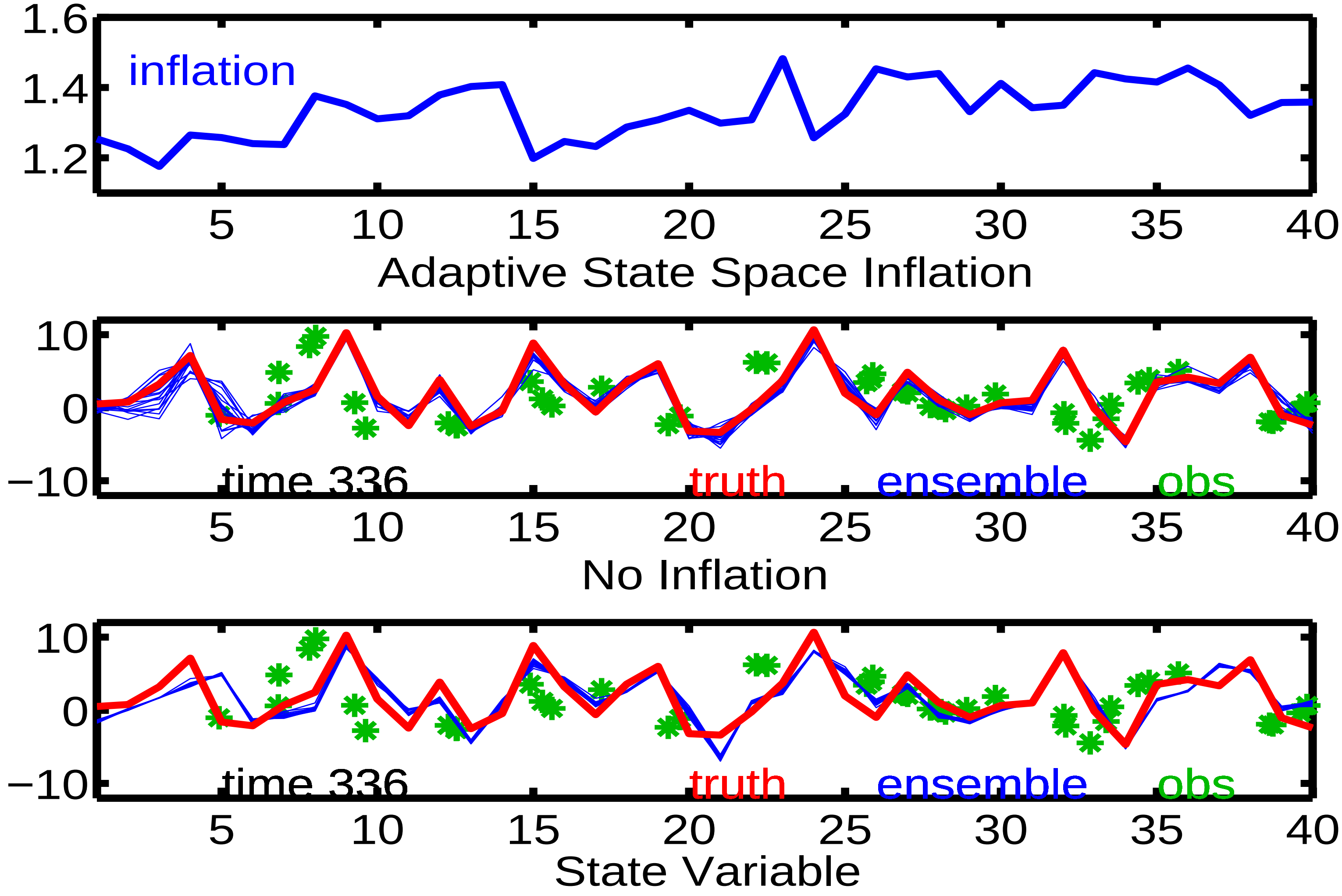
Automatically selected by adaptive inflation algorithm.



Assimilating with Inflation in presence of model error

Inflation is a function of state variable and time.

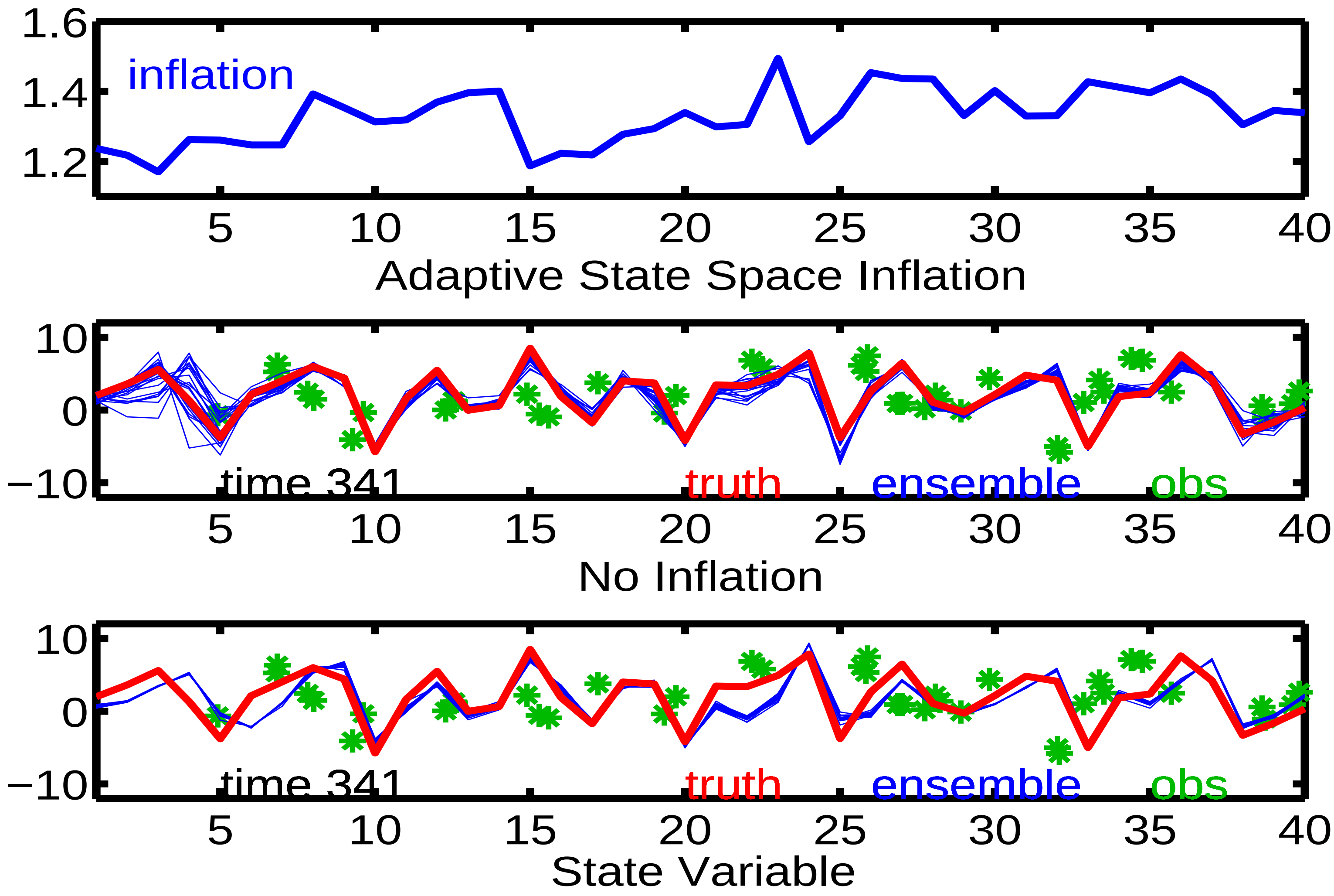
Automatically selected by adaptive inflation algorithm.



Assimilating with Inflation in presence of model error

Inflation is a function of state variable and time.

Automatically selected by adaptive inflation algorithm.

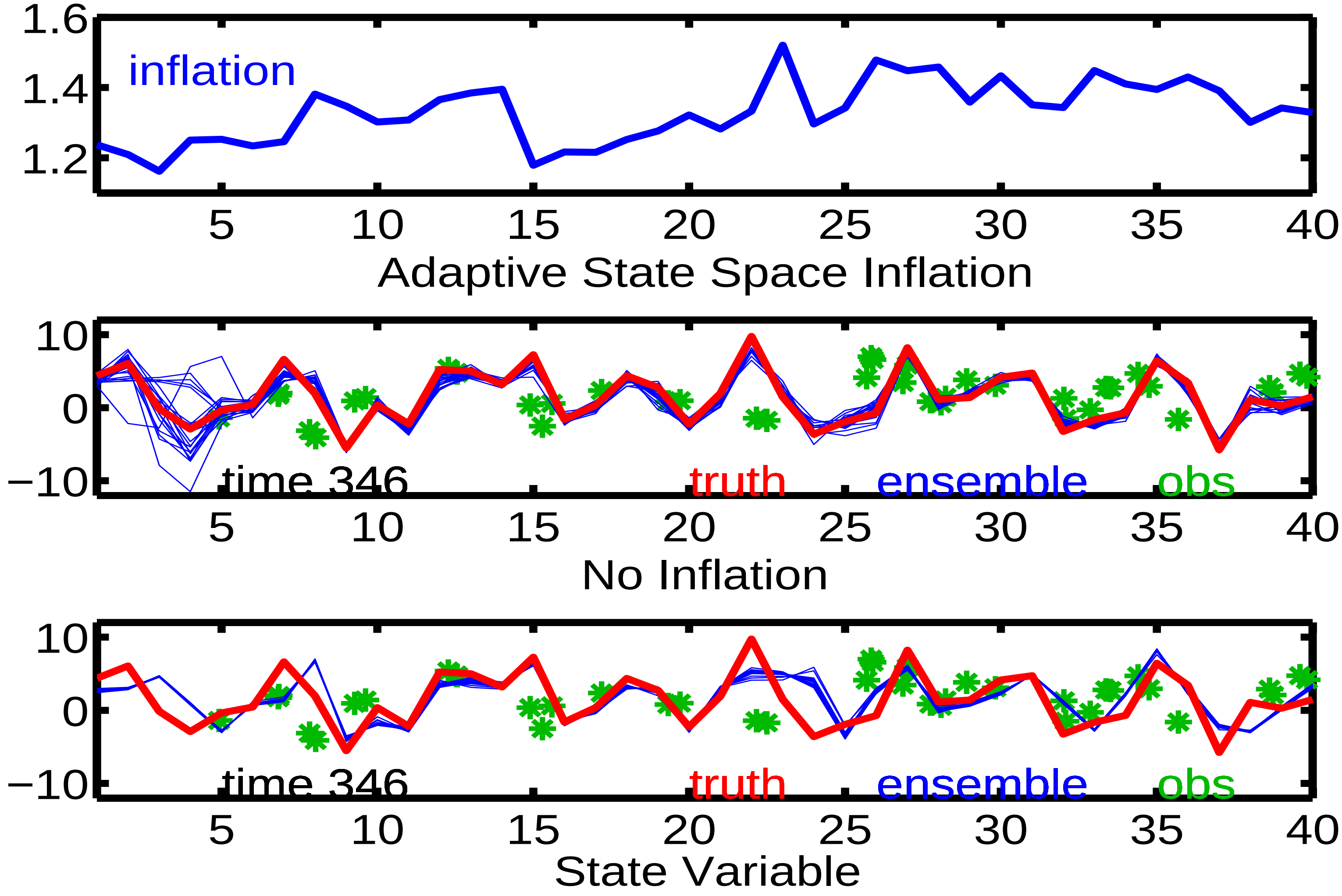


Assimilating with Inflation in presence of model error

Inflation is a function of state variable and time.

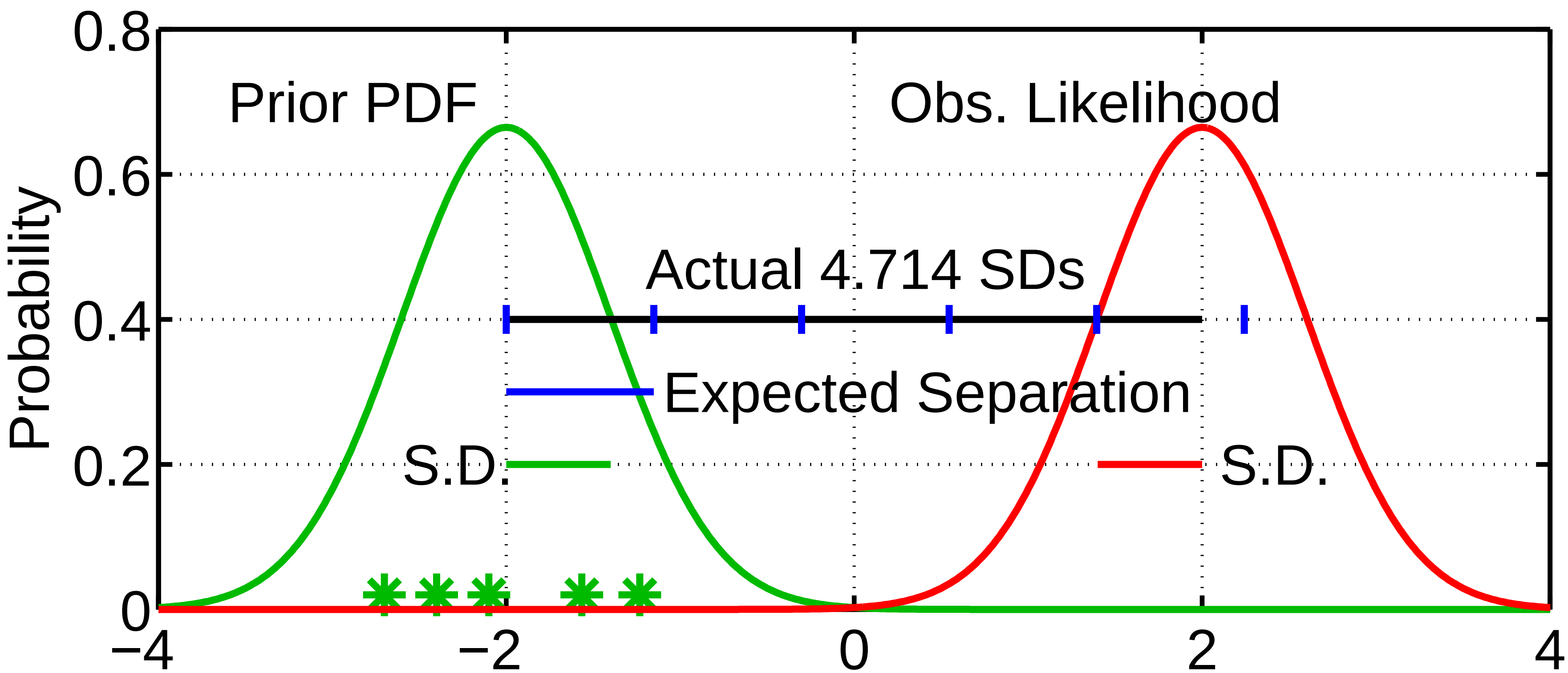
Automatically selected by adaptive inflation algorithm.

It can work, even in presence of severe model error.



Adaptive Inflation for Ensemble Filtering

1. For observed variable, have estimate of prior-observed inconsistency.



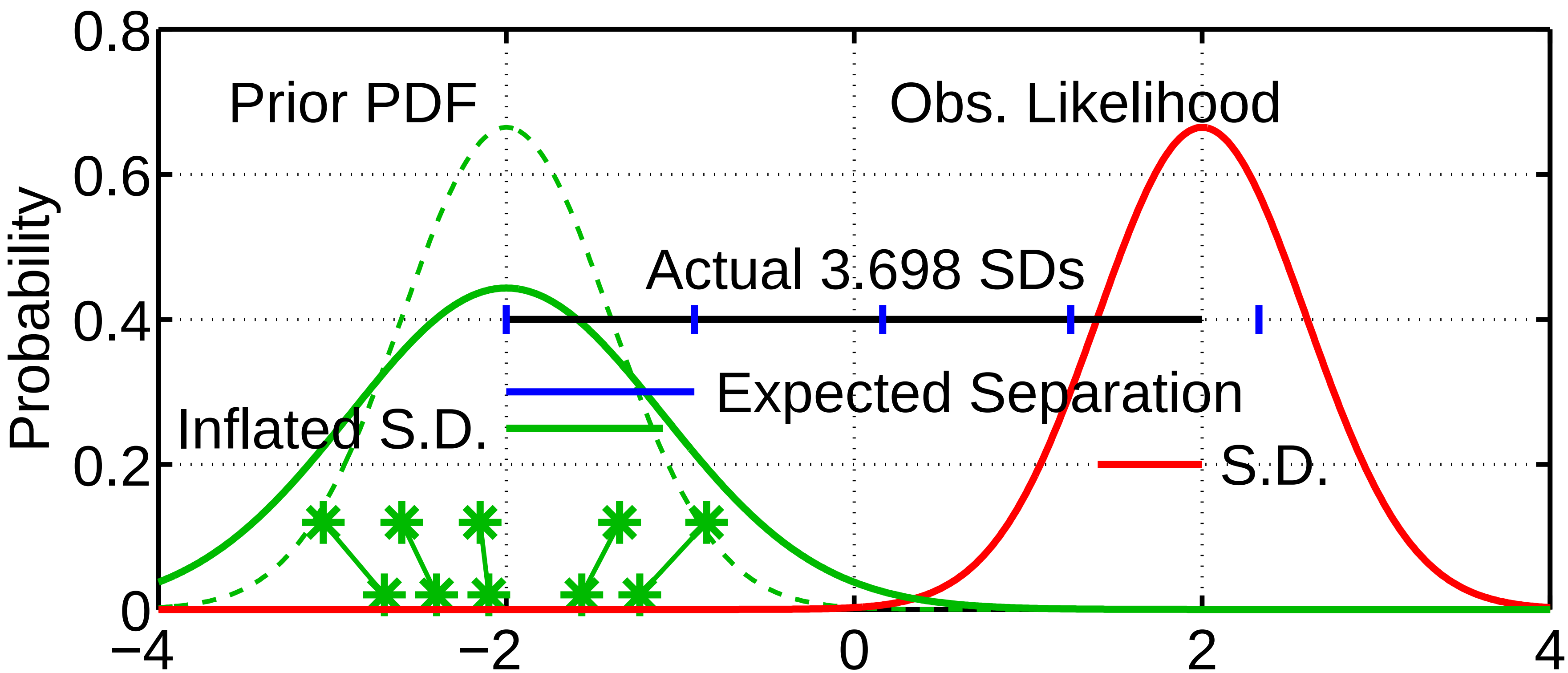
2. Expected(prior mean - observation) = $\sqrt{\sigma^2_{prior} + \sigma^2_{obs}}$.

Assumes that prior and observation are supposed to be unbiased.

Is it model error or random chance?

Adaptive Inflation for Ensemble Filtering

1. For observed variable, have estimate of prior-observed inconsistency.



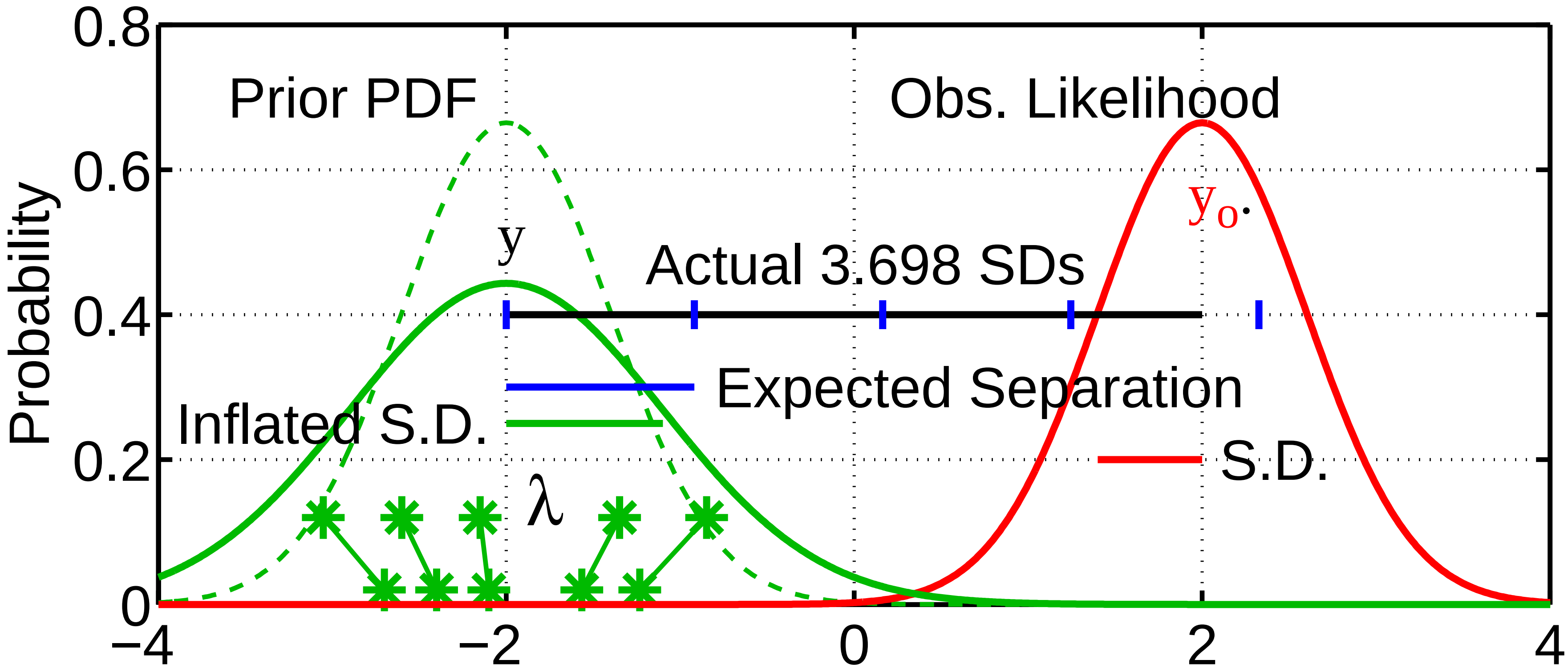
2. Expected(prior mean - observation) = $\sqrt{\sigma^2_{prior} + \sigma^2_{obs}}$.

3. Inflating increases expected separation.

Increases 'apparent' consistency between prior and observation.

Adaptive Inflation for Ensemble Filtering

1. For observed variable, have estimate of prior-observed inconsistency.

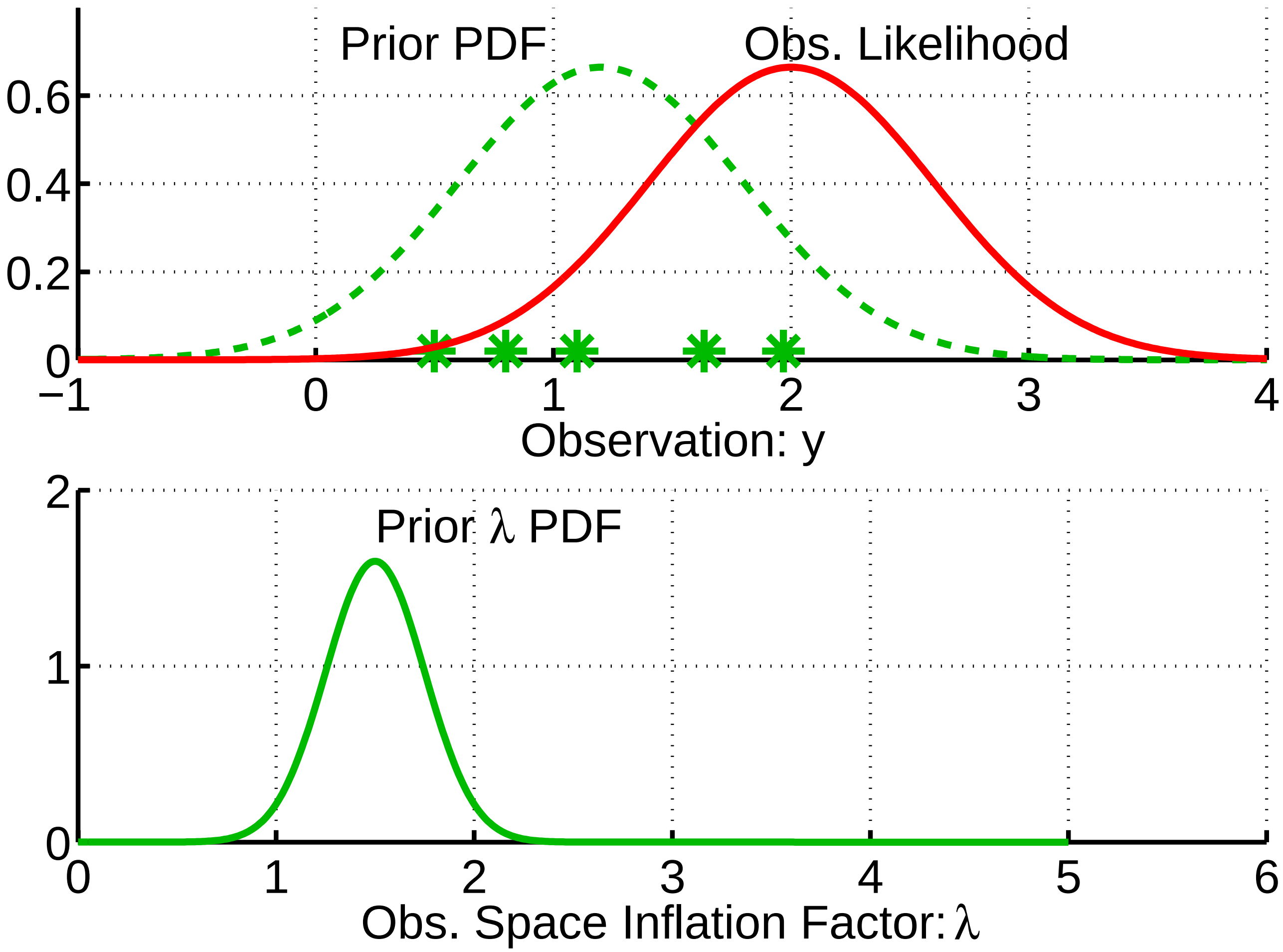


Distance, D, from prior mean y to obs. is $N\left(0, \sqrt{\lambda\sigma^2_{prior} + \sigma^2_{obs}}\right) = N(0, \theta)$

Prob. y_o is observed given λ: $p(y_o|\lambda) = (2\Pi\theta^2)^{-1/2}\exp(-D^2/2\theta^2)$

Adaptive Inflation for Ensemble Filtering

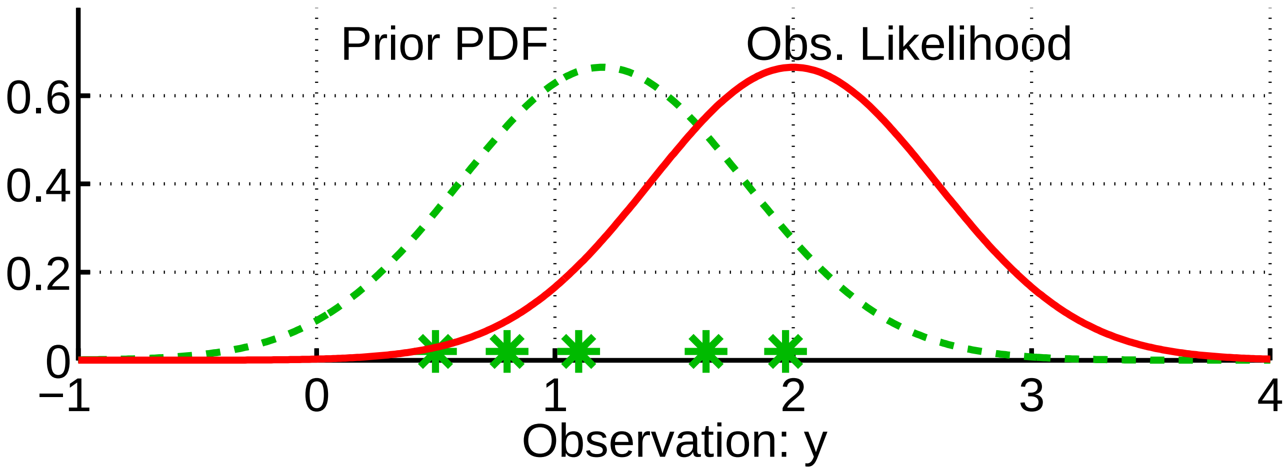
Use Bayesian statistics to get estimate of inflation factor, λ.



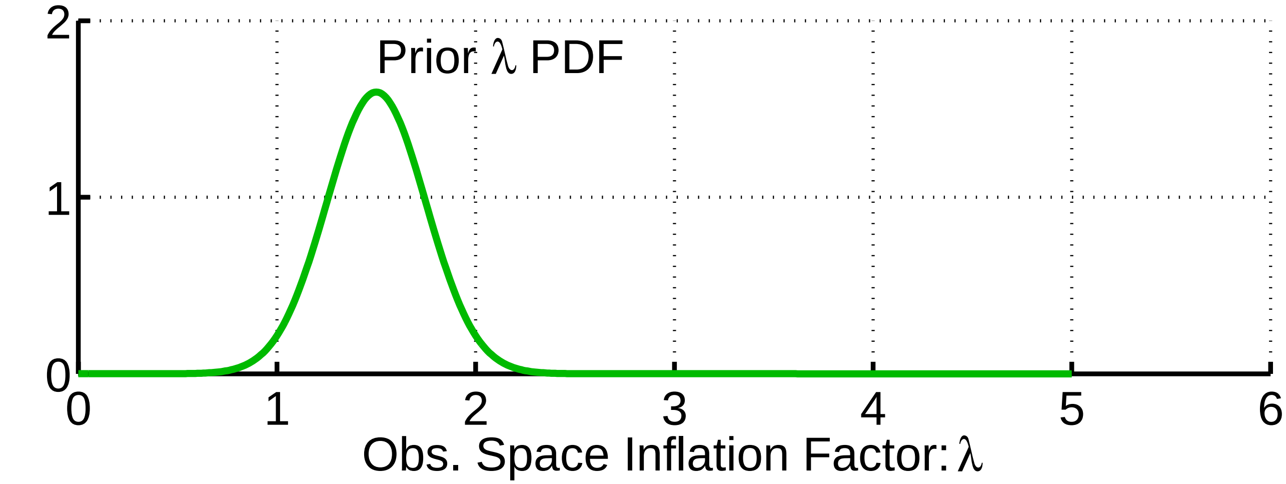
Assume prior is gaussian; $p(\lambda | Y_{prev}) = N(\bar{\lambda}_p, \sigma^2_{\lambda, p})$.

Adaptive Inflation for Ensemble Filtering

Use Bayesian statistics to get estimate of inflation factor, λ.



We've assumed a gaussian for prior $p(\lambda|Y_{prev})$.

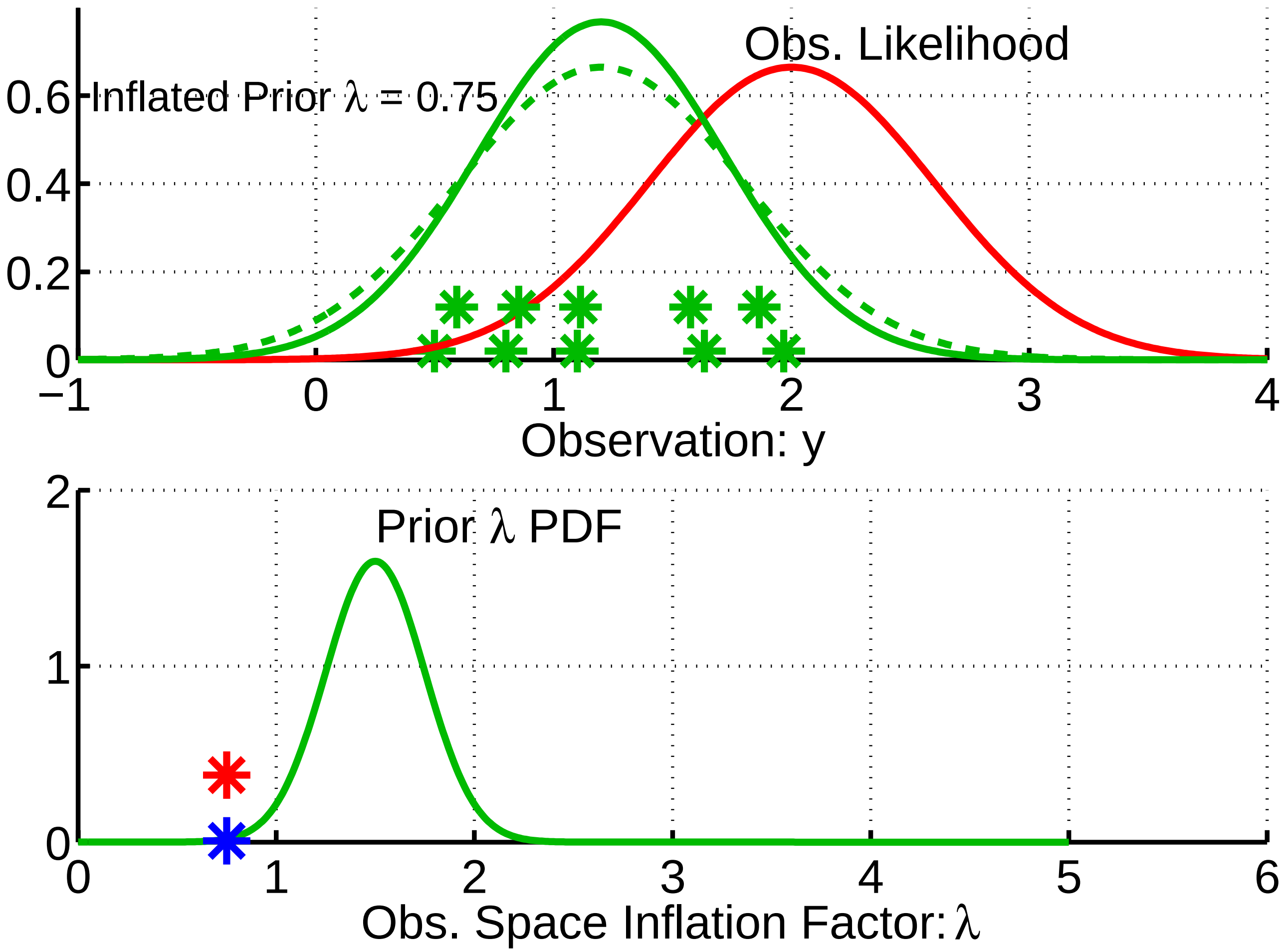


Recall that $p(y_o|\lambda)$ can be evaluated from normal PDF.

$$p(\lambda|Y_{prev}, y_o) = p(y_o|\lambda)p(\lambda|Y_{prev})/normalization.$$

Adaptive Inflation for Ensemble Filtering

Use Bayesian statistics to get estimate of inflation factor, λ.



Get $p(y_o|\lambda = 0.75)$ from normal PDF.

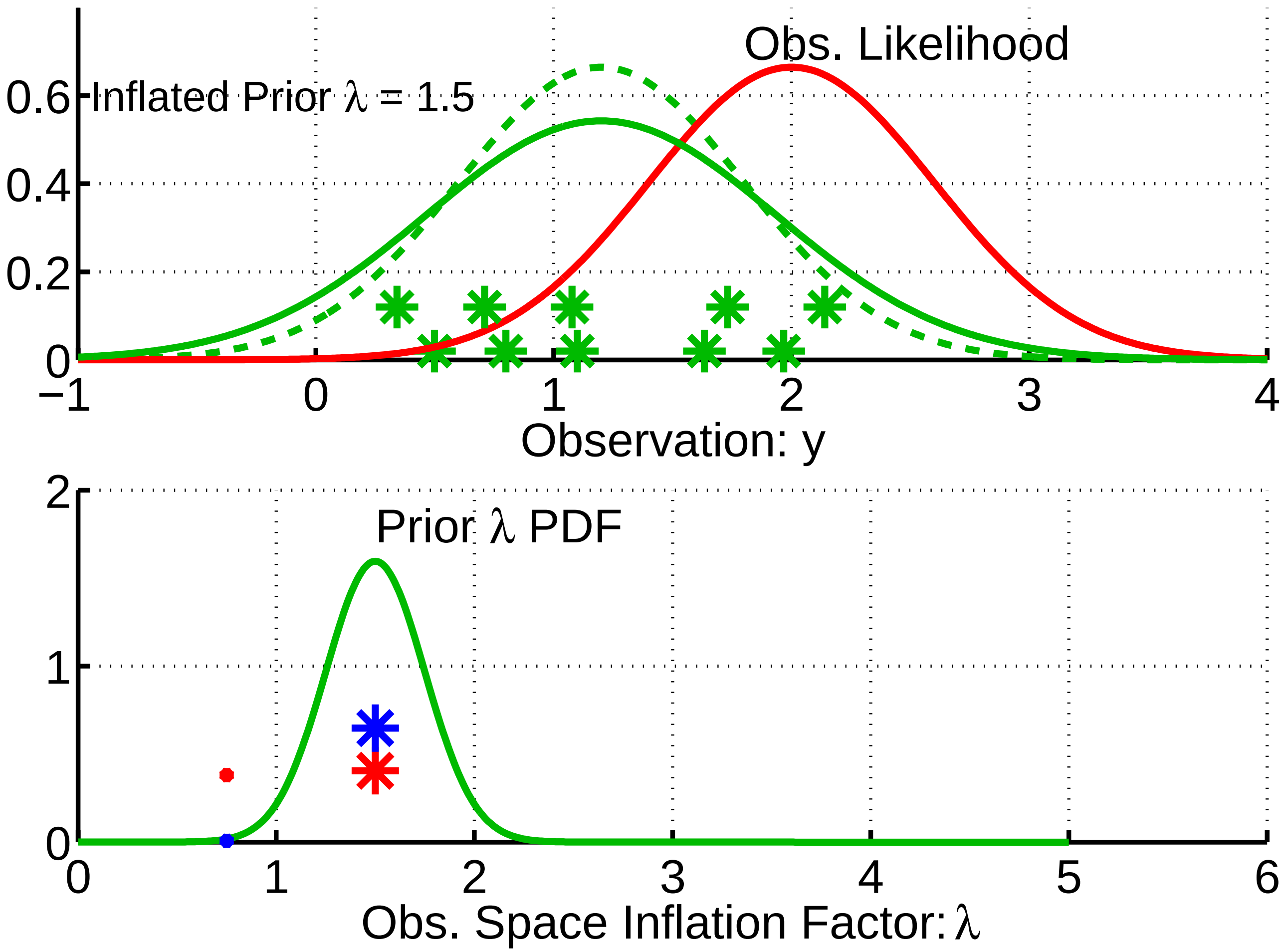
Multiply by $p(\lambda = 0.75|Y_{prev})$ to get

$p(\lambda = 0.75|Y_{prev}, y_o)$

$$p(\lambda|Y_{prev}, y_o) = p(y_o|\lambda)p(\lambda|Y_{prev})/normalization .$$

Adaptive Inflation for Ensemble Filtering

Use Bayesian statistics to get estimate of inflation factor, λ.



Get $p(y_o|\lambda = 1.5)$

from normal PDF.

Multiply by

$p(\lambda = 1.5|Y_{prev})$

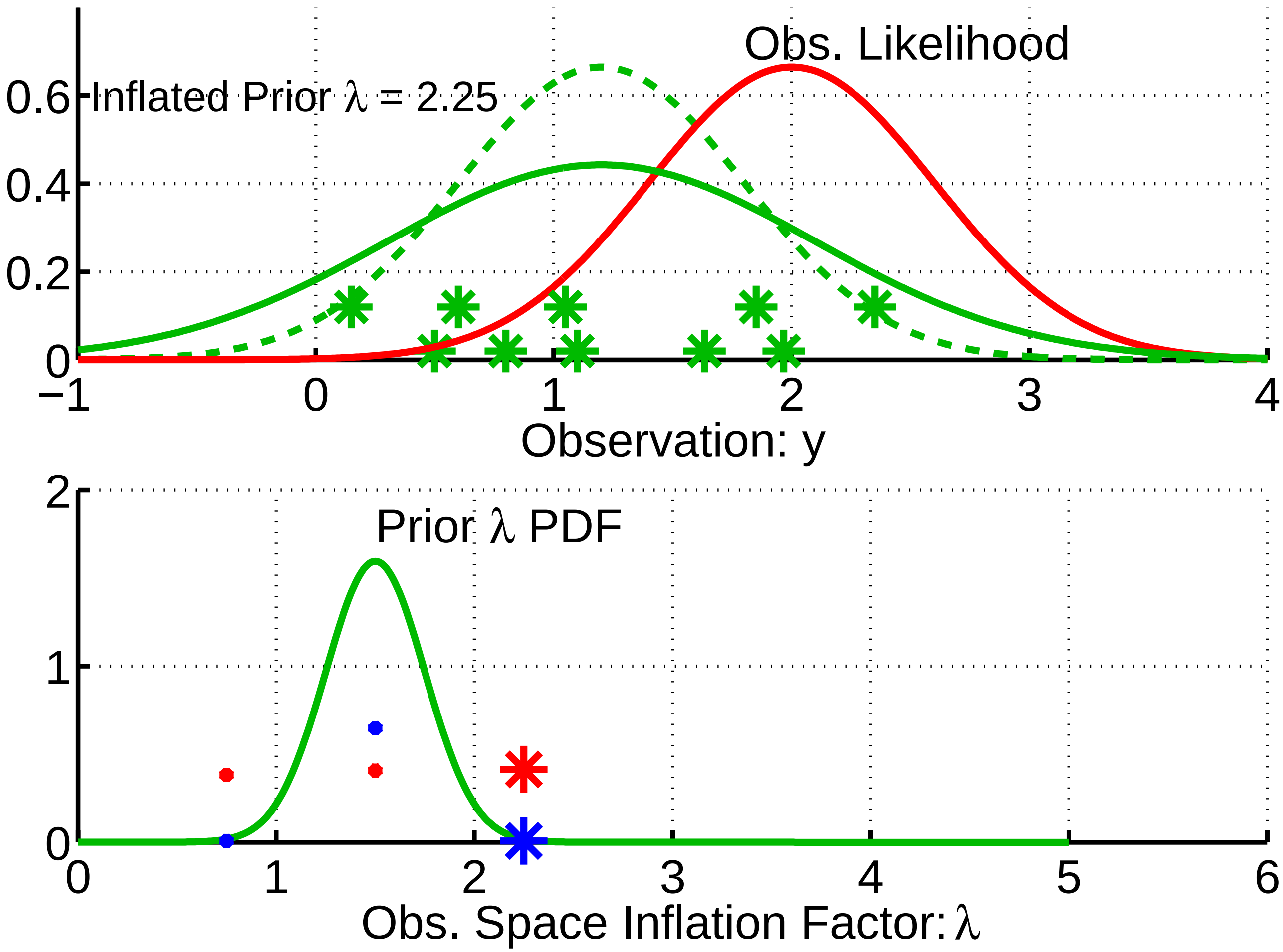
to get

$p(\lambda = 1.5|Y_{prev}, y_o)$

$$p(\lambda|Y_{prev}, y_o) = p(y_o|\lambda)p(\lambda|Y_{prev})/normalization\ .$$

Adaptive Inflation for Ensemble Filtering

Use Bayesian statistics to get estimate of inflation factor, λ.



Get $p(y_o|\lambda = 2.2)$

from normal PDF.

Multiply by

$p(\lambda = 2.2|Y_{prev})$

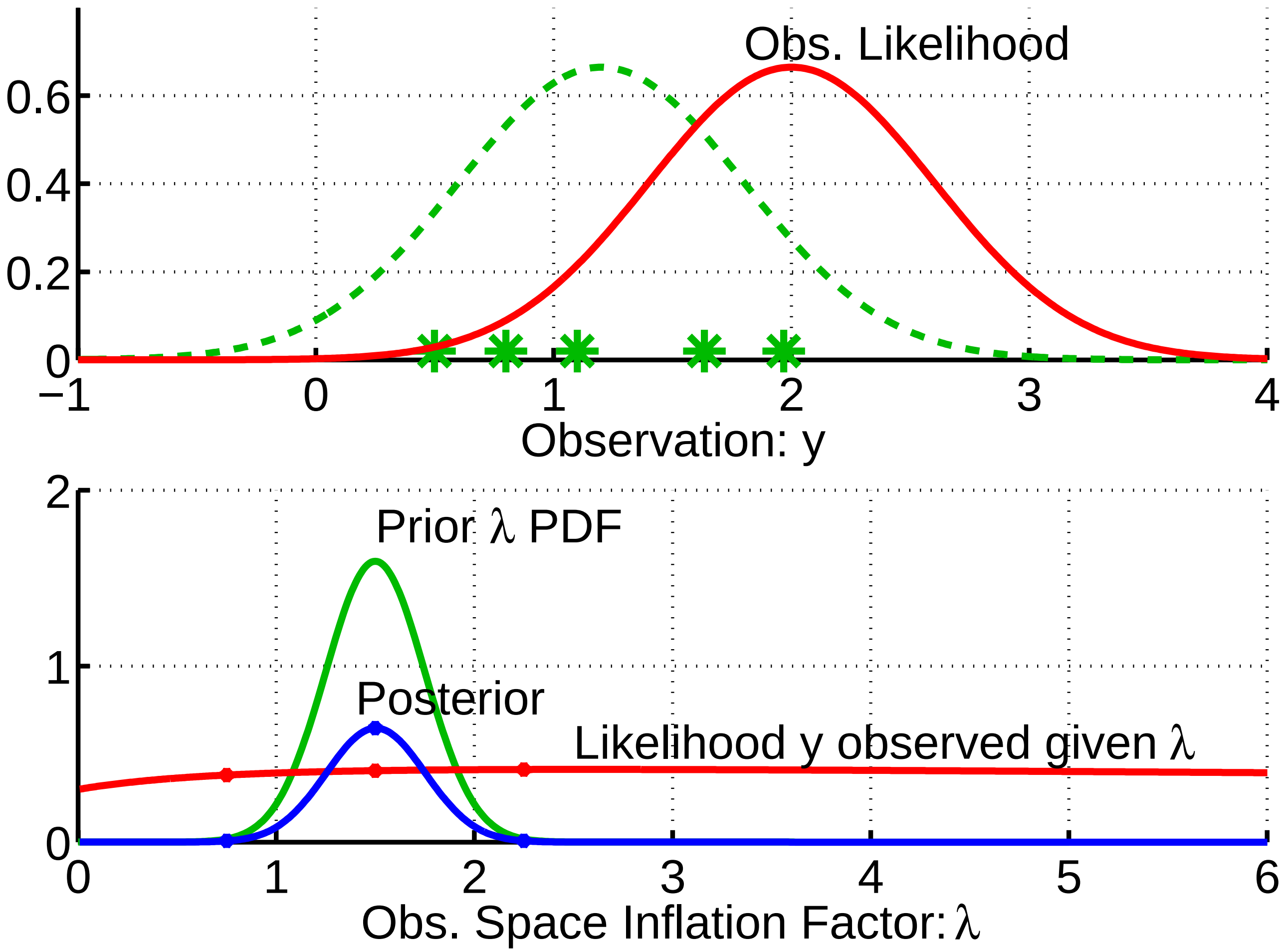
to get

$p(\lambda = 2.2|Y_{prev}, y_o)$

$$p(\lambda|Y_{prev}, y_o) = p(y_o|\lambda)p(\lambda|Y_{prev})/normalization\ .$$

Adaptive Inflation for Ensemble Filtering

Use Bayesian statistics to get estimate of inflation factor, λ.



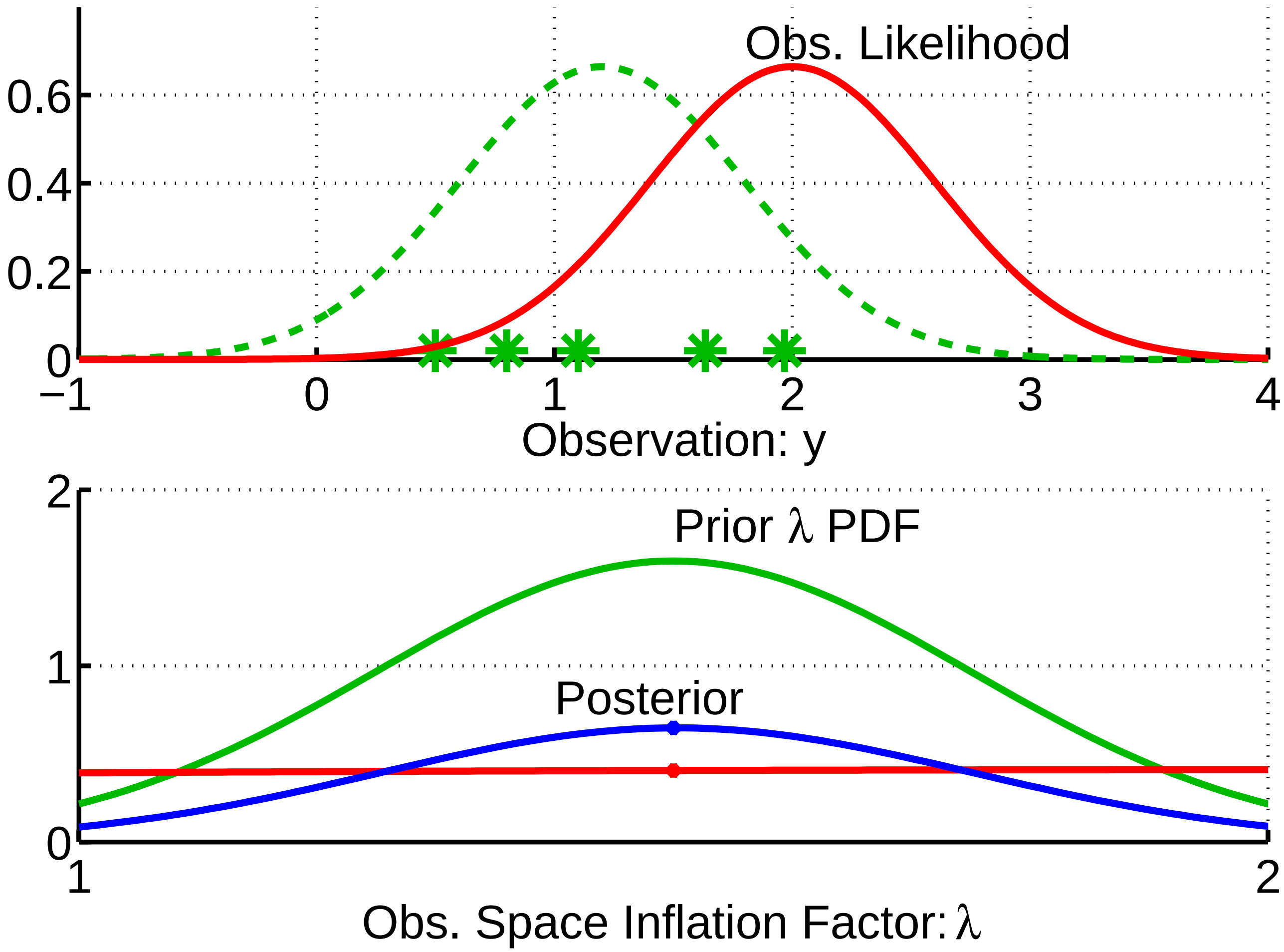
Repeat for a range of values of λ.

Now must get posterior in same form as prior (gaussian).

$$p(\lambda | Y_{prev}, y_o) = p(y_o | \lambda) p(\lambda | Y_{prev}) / normalization\ .$$

Adaptive Inflation for Ensemble Filtering

Use Bayesian statistics to get estimate of inflation factor, λ.



Very little information about λ in a single observation.

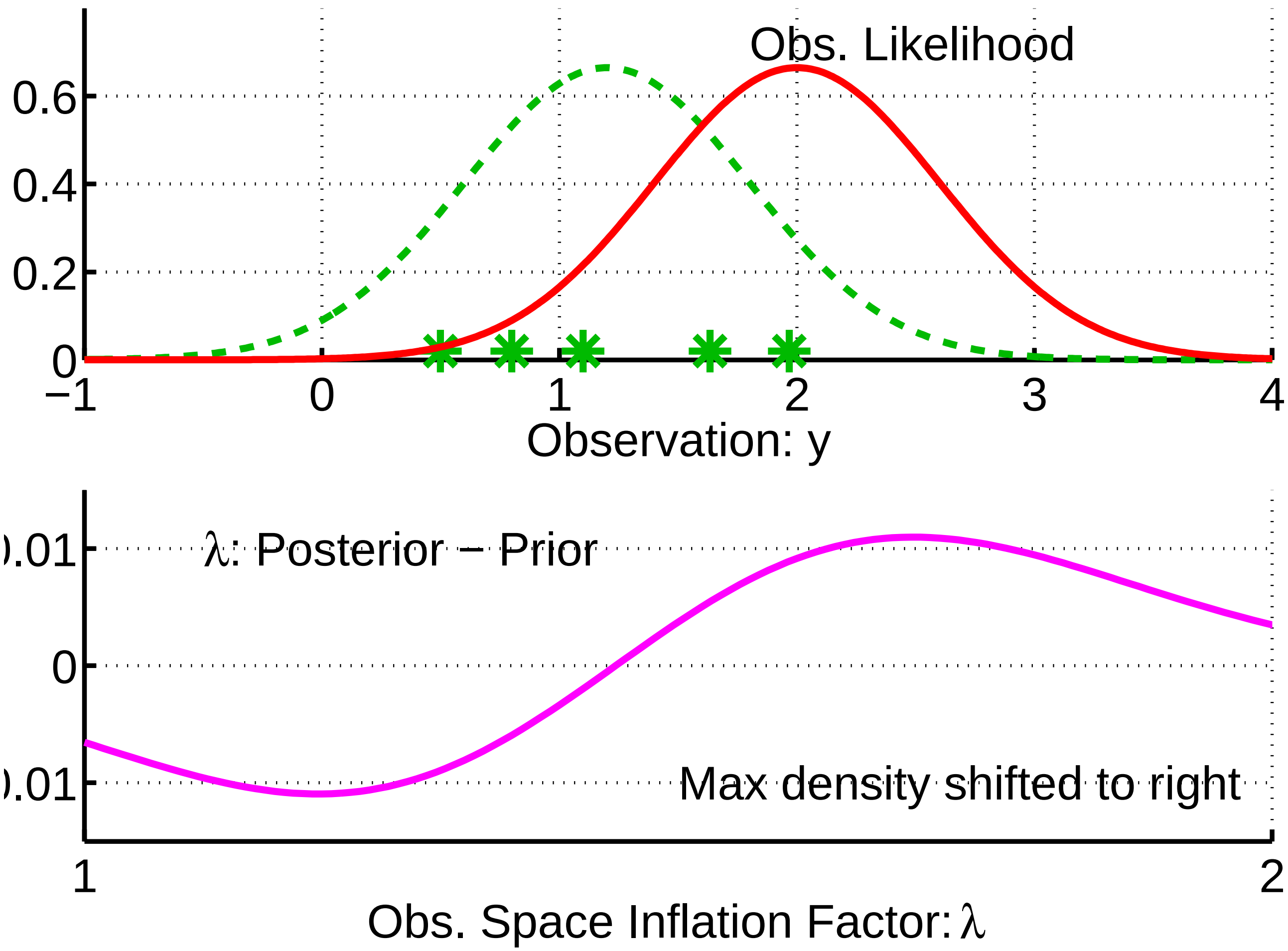
Posterior and prior are very similar.

Normalized posterior indistinguishable from prior.

$$p(\lambda|Y_{prev}, y_o) = p(y_o|\lambda)p(\lambda|Y_{prev})/normalization .$$

Adaptive Inflation for Ensemble Filtering

Use Bayesian statistics to get estimate of inflation factor, λ.



Very little information about λ in a single observation.

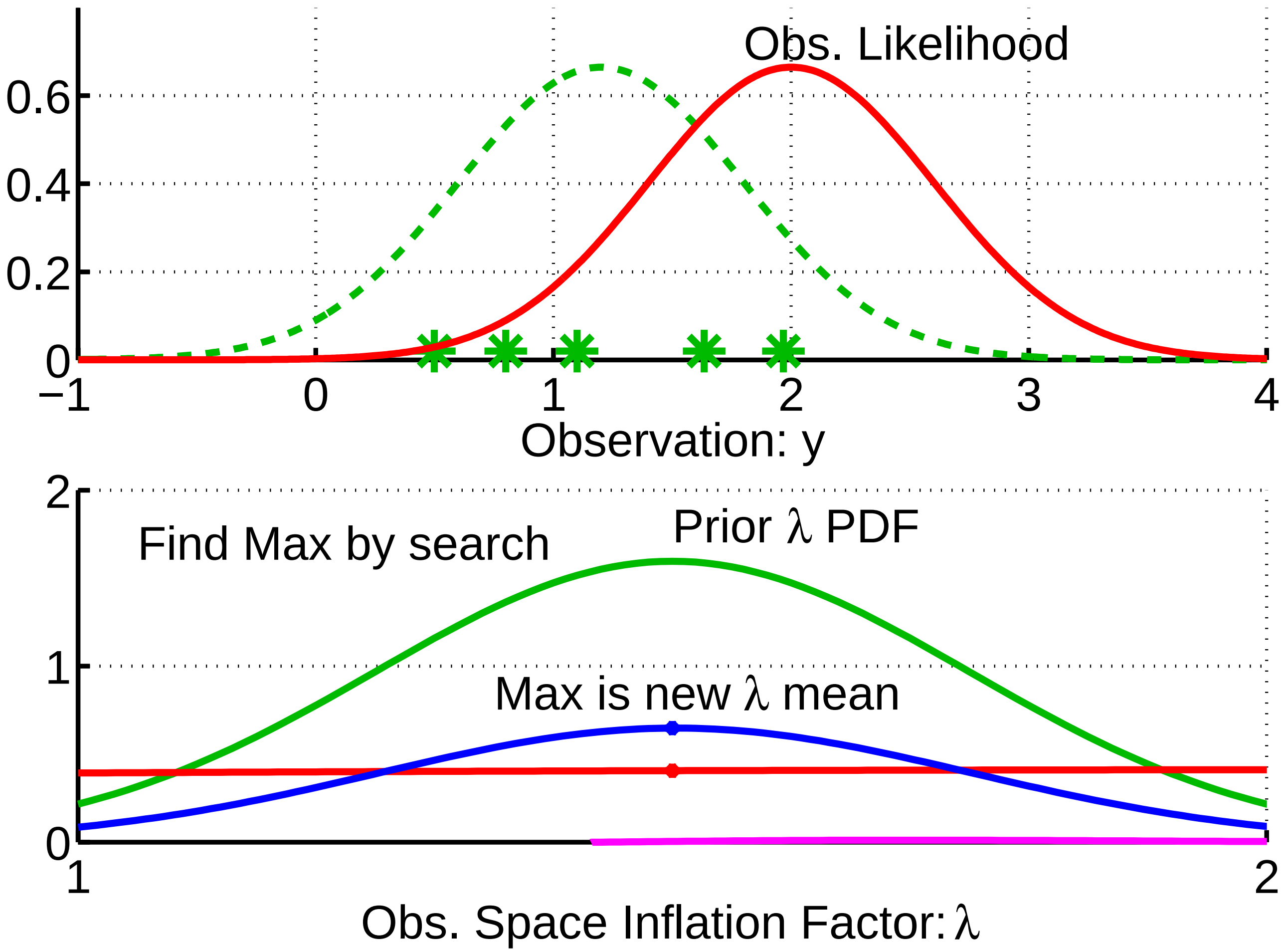
Posterior and prior are very similar.

Difference shows slight shift to larger values of λ.

$$p(\lambda | Y_{prev}, y_o) = p(y_o | \lambda) p(\lambda | Y_{prev}) / normalization .$$

Adaptive Inflation for Ensemble Filtering

Use Bayesian statistics to get estimate of inflation factor, λ.



One option is to use Gaussian prior for λ.

Select max (mode) of posterior as mean of updated Gaussian.

Do a fit for updated standard deviation.

$$p(\lambda | Y_{prev}, y_o) = p(y_o | \lambda) p(\lambda | Y_{prev}) / normalization \ .$$

Adaptive Inflation for Ensemble Filtering

A. Computing updated inflation mean, $\bar{\lambda}_u$.

Mode of $p(y_o|\lambda)p(\lambda|Y_{prev})$ can be found analytically!

Solving $\partial\left[p(y_o|\lambda)p(\lambda|Y_{prev})\right]/\partial\lambda = 0$ leads to 6th order poly in θ

This can be reduced to a cubic equation and solved to give mode.

New $\bar{\lambda}_u$ is set to the mode.

This is relatively cheap compared to computing regressions.

Adaptive Inflation for Ensemble Filtering

A. Computing updated inflation variance, $\sigma^2_{\lambda,u}$

1. Evaluate numerator at mean $\bar{\lambda}_u$ and second point, e.g. $\bar{\lambda}_u + \sigma_{\lambda,p}$.

2. Find $\sigma^2_{\lambda,u}$ so $N(\bar{\lambda}_u, \sigma^2_{\lambda,u})$ goes through $p(\bar{\lambda}_u)$ and $p(\bar{\lambda}_u + \sigma_{\lambda,p})$.

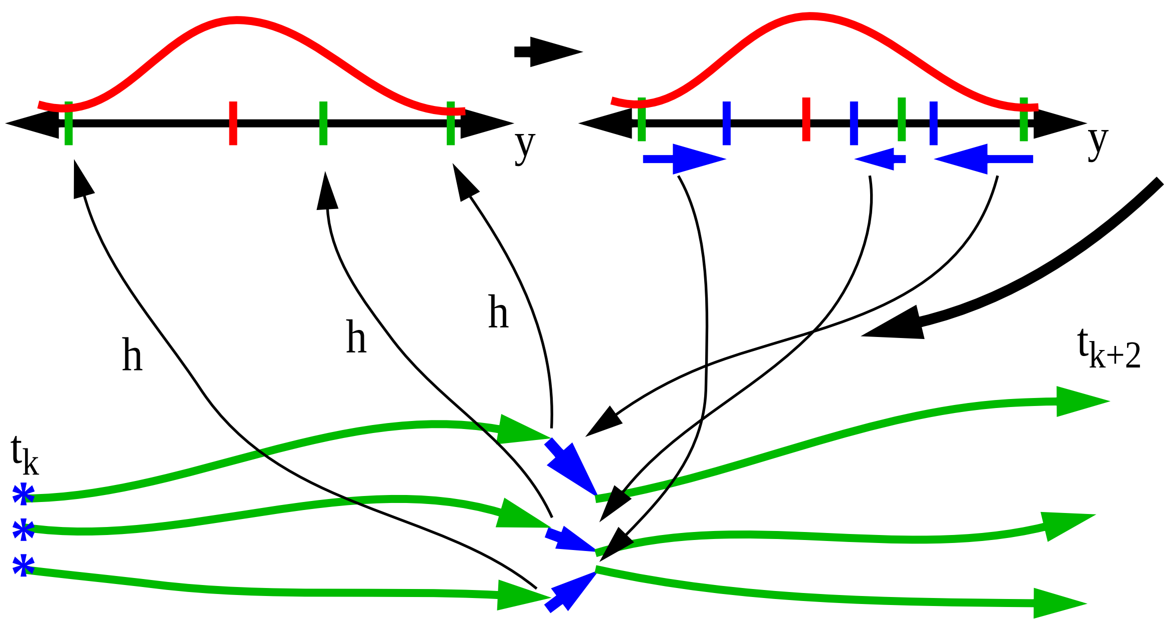
3. Compute as $\sigma^2_{\lambda,u} = -\sigma^2_{\lambda,p} / 2 \ln r$ where $r = p(\bar{\lambda}_u + \sigma_{\lambda,p}) / p(\bar{\lambda}_u)$.

State Space Adaptive Inflation

Computations so far adapt inflation for observation space.

What is relation between observation and state space inflation?

Have to use prior ensemble observation/state joint distribution.



Regress changes in inflation onto state variable inflation.

Spatially varying adaptive inflation algorithm:

Have a distribution for λ at each time for each state variable, $\lambda_{s,i}$.

Use prior correlation from ensemble to determine impact of $\lambda_{s,i}$ on prior variance for given observation.

If γ is correlation between state variable i and observation then

$$\theta = \sqrt{[1 + \gamma(\sqrt{\lambda_{s,i}} - 1)]^2 \sigma_{prior}^2 + \sigma_{obs}^2}.$$

Equation for finding mode of posterior is now full 12th order:

Analytic solution appears unlikely.

Can do Taylor expansion of θ around $\lambda_{s,i}$.

Retaining linear term is normally quite accurate.

There is an analytic solution to find mode of product in this case!

Adaptive Inflation in Global NWP

Model: CAM 3.1 T85L26.

Several million model variables: winds, temperature, moisture.

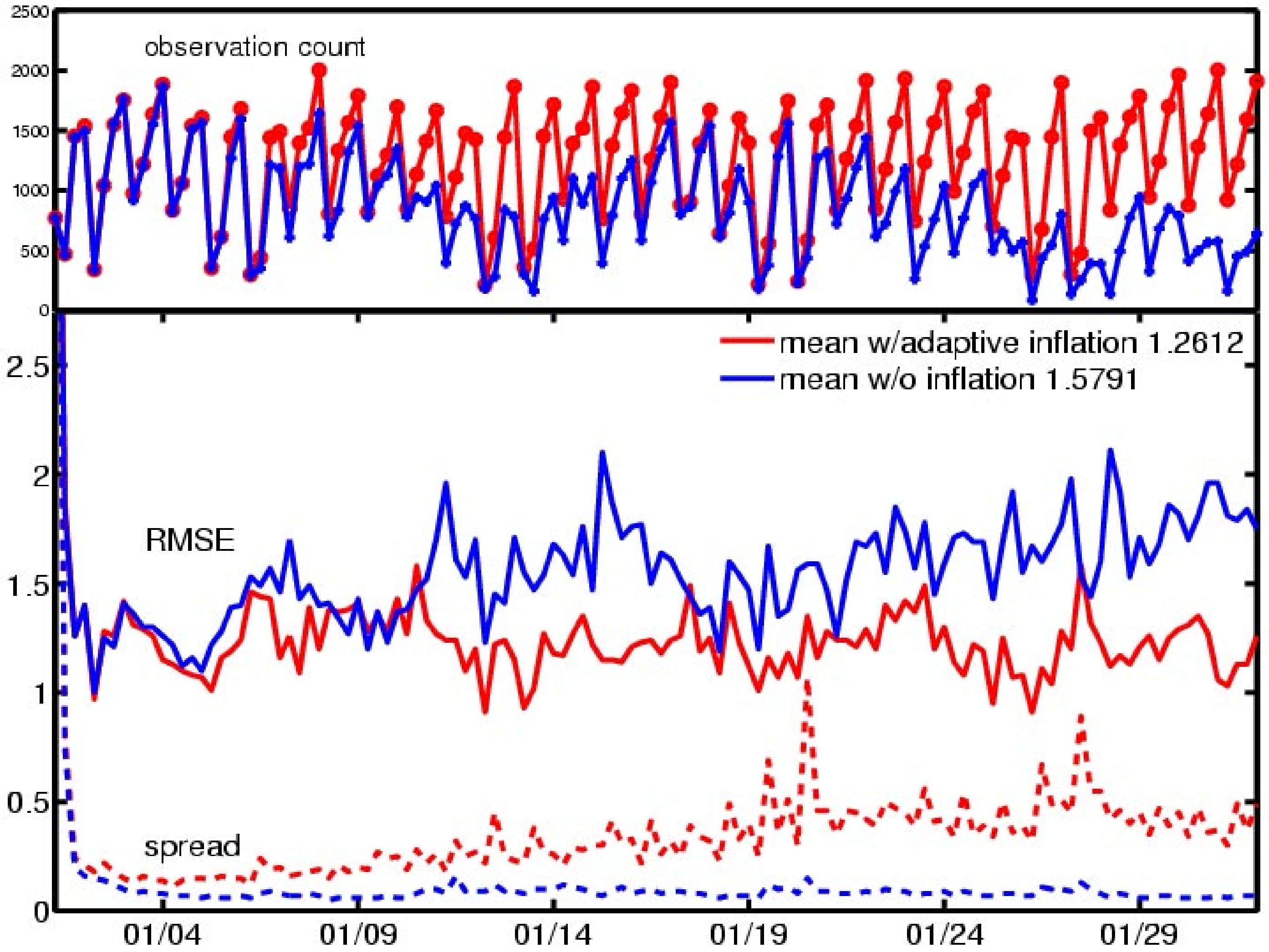
Initialized from a climatological distribution (huge spread).

Observations: Radiosondes, ACARS, Satellite Winds.

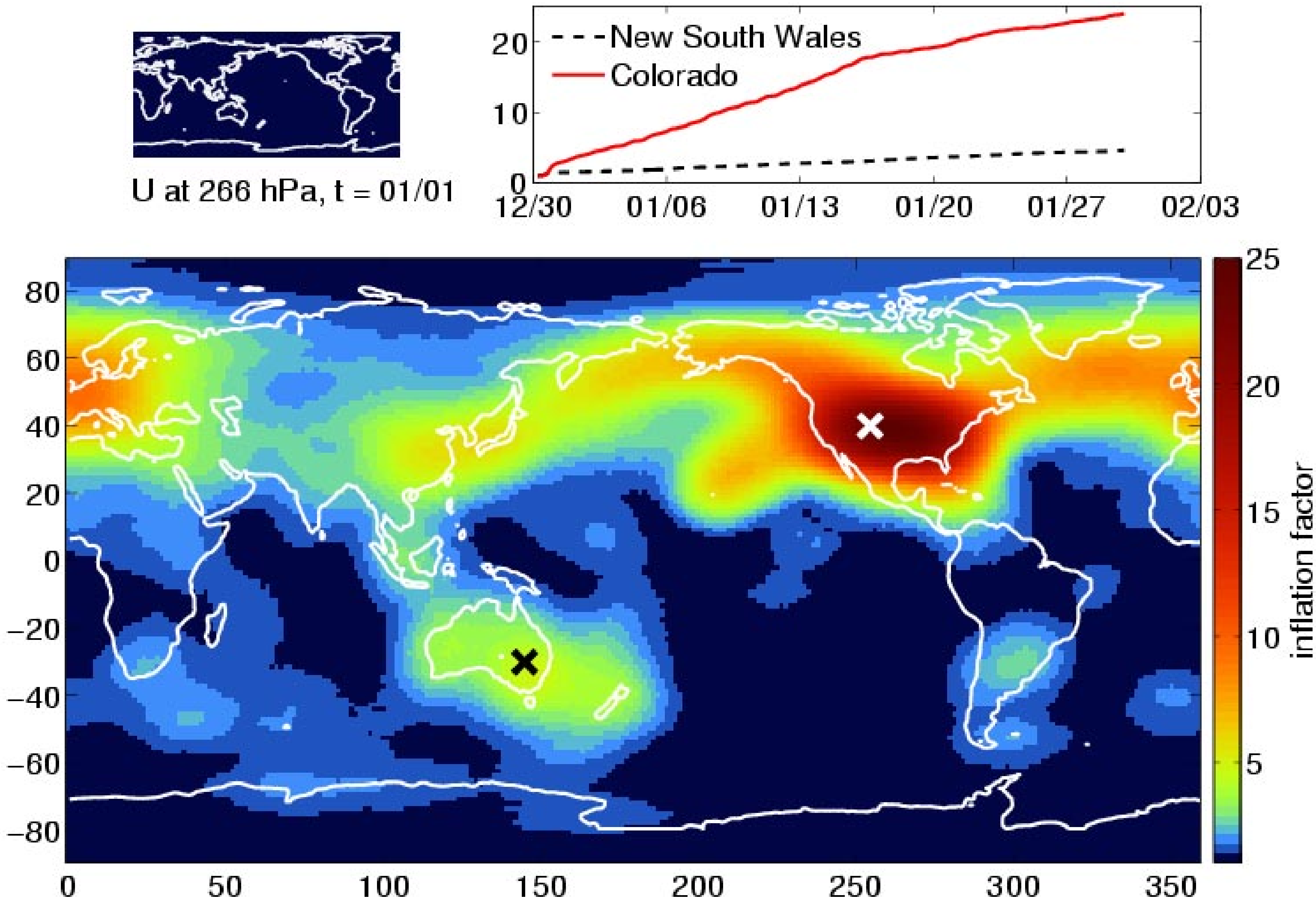
Subset of observations used in NCAR/NCEP reanalysis.

Several hundred thousand observations per day.

Adaptive Inflation in CAM; 500 hPa T Obs. Space Prior RMS, Spread



Adaptive Inflation in CAM after 1-Month; 266 hPa U



Adaptive Inflation for Numerical Weather Prediction in CAM

1. Largest inflation caused by model error in densely observed regions.

2. RMS reduced, spread increased.

3. Fewer observations rejected.

Combined with localization, allows 80 member ensemble to work well!

Hierarchical Bayesian Methods for Adaptive Filters: Summary

1. Localization:
    - Run an ensemble of ensembles.
    - Use regression coefficient signal-to-noise ratio to minimize error.

2. Inflation:
    - Use each observation twice.
        - Once to adjust parameter (inflation) of filter system.
        - Second time to adjust mean and variance of estimate.

Conclusions

Lots of cool statistics being done poorly.

Work in small models can give insight.

Applications in large models are important.

Looking for collaborations with interested statisticians.