Uncertainty Quantification in Geophysical Systems

Craig H. Bishop

Naval Research Laboratory, Monterey, CA 93940

Acknowledgements: Elizabeth Satterfield, David Kuhl, Daniel Hodyss all from Naval Research Laboratory, Monterey and Tom Rosmond from Forks Washington via Science Apllications International Corporation.

BIRS workshop, July 2011

Overview

- Introduction
- Hidden volatility (variance) in chaotic systems
  - Illustrated using a simple chaotic system and an even simpler univariate example
- Using knowledge of errors in error variance estimates to improve DA via Hybrid error covariance models
  - Illustrated using Navy's obs space 4D-VAR scheme NAVDAS-AR
- Conclusions

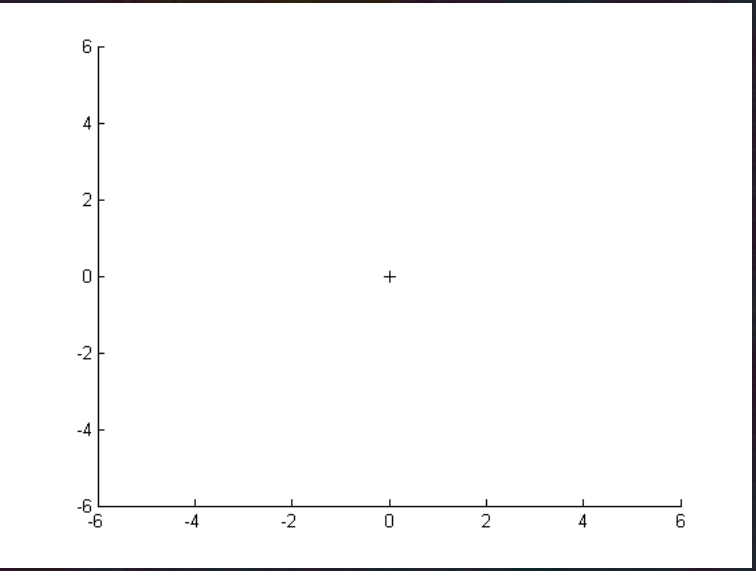
Donald Rumsfeld

“There are known knowns; there are things we know we know. We also know there are known unknowns; that is to say we know there are some things we do not know. But there are also unknown unknowns ...”— Donald Rumsfeld, 2/12/2002, then U.S. Secretary of Defense.

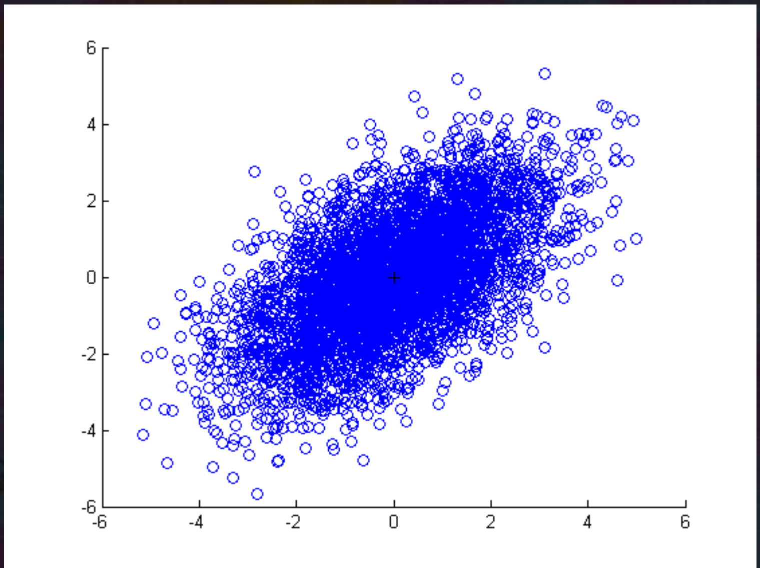
Grimweld the groundhog interprets Rumsfeld

- Vertical position of dropped ground-hog is a *known known*

Belief, Uncertainty and repeatable stochastic phenomena



Belief, Uncertainty and repeatable stochastic phenomena



Empirical quantification of uncertainty is straightforward for (easily) repeatable stochastic phenomena like ball throwing

Not so easy for chaotic systems like atmosphere, stock-market, politics, ocean, etc.

Grimweld the Ground-Hog

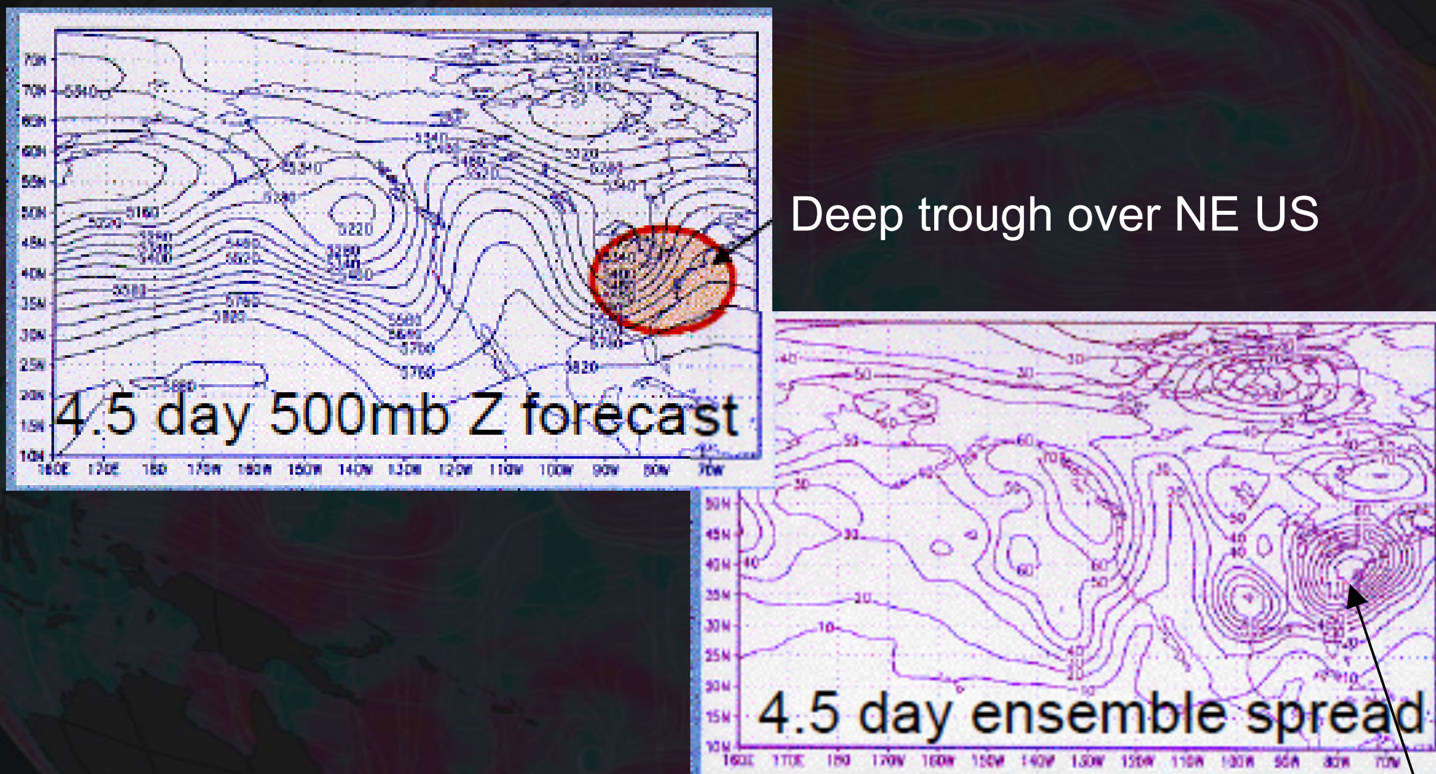
• Vertical position of dropped ground-hog is a *known known*

• Impact location of ground-hog thrown at X by Craig is a *known unknown* provided we have previously *empirically* defined the distribution of ground-hog locations through repeated Craig throws – (assuming that Craig's throwing accuracy never improves). In addition, the error covariance **P** of the throws is precisely defined.

• Impact location of ground-hog thrown at X on wall by randomly selected audience member is an *unknown unknown* (no empirical prior distribution or **P** is available)

• Similarly, forecast error covariance associated with today's weather is an unknown unknown

Meteorological example of uncertainty that is difficult to quantify



Ensemble based uncertainty prediction

Hidden volatility: When the error variance (volatility) depends on a flow pattern that is unlikely to repeat itself, it is *hidden* because one or two realizations of error do not enable an estimation of variance.

Replicate Earths for ultimate uncertainty quantification

(Slartibartfast – Magrathean designer of planets, D. Adams, Hitchhikers …)

- Imagine an unimaginably large number of quasi-identical Earths.

Each Earth has one true state and one prediction but these differ from one Earth to another.

Collect all Earths having the same historical observations $\mathbf{y}$ but differing true atmospheric states $\mathbf{x}^t$ to define $\rho\left(\mathbf{x}^t \mid \mathbf{y}_{i-1}, \mathbf{y}_{i-2}, \ldots\right)$. The differences between individual truths and the mean truth within this set define the fixed-obs forecast error distribution.

Collect all Earths having the same true state $\mathbf{x}^t$ but differing forecasts $\mathbf{x}^f$ of this state to define $\rho\left(\mathbf{x}^f \mid \mathbf{x}^t\right)$. The fixed-truth error distribution is the distribution of differences between individual forecasts and single truth within this set.

Replicate Earths with simple model

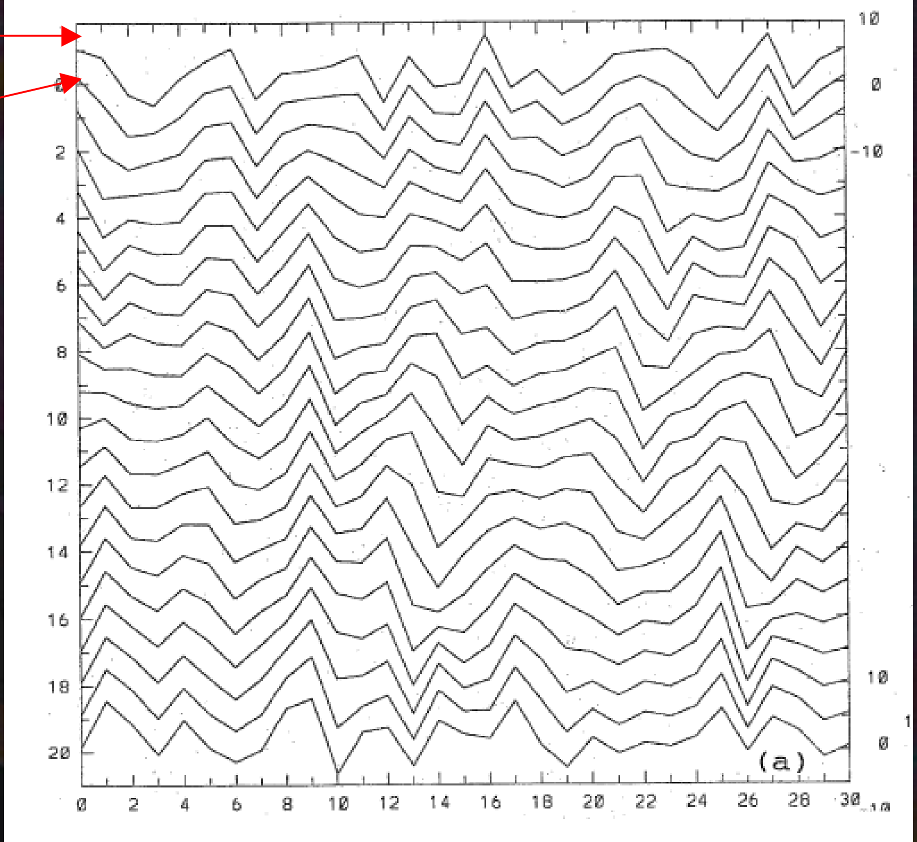
Suppose Earth evolution governed by the following chaotic model

Lorenz (2005, J. Atmos. Sci), model 1

$$dX_n/dt = -X_{n-2}X_{n-1} + X_{n-1}X_{n+1} - X_n + F. \quad (1)$$

Initial state

State after 1 time step



Time =>

("Earth" only has 30 varibles in this example)

Model grid point

Replicate Earths with simple model (continued)

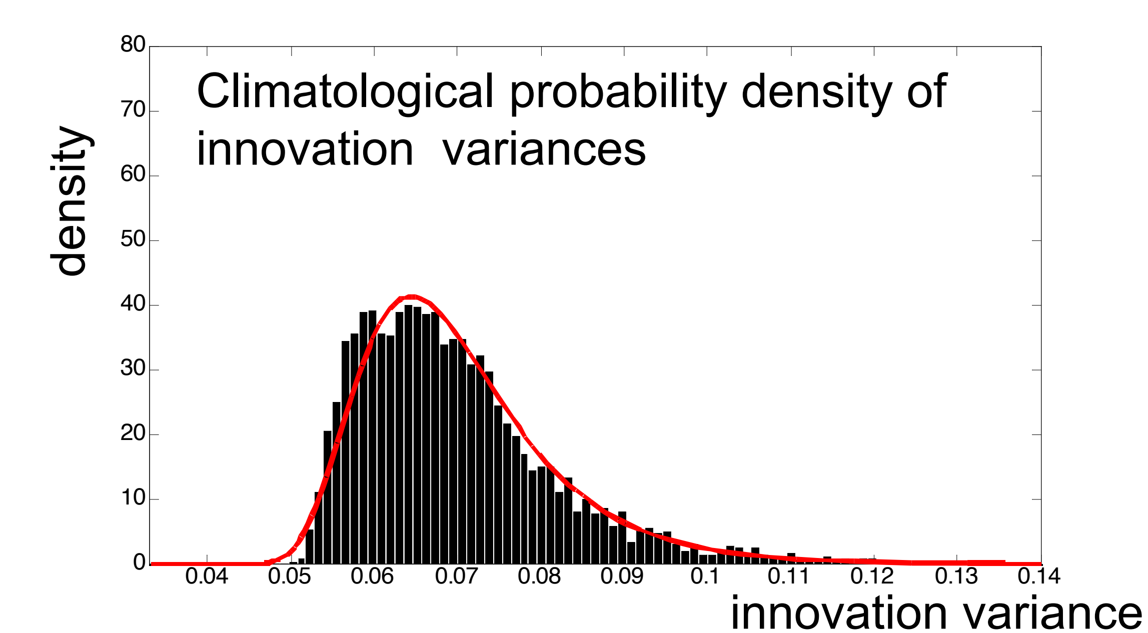
1. To reduce computational expense, only use 10 horizontal grid points.
2. Perform very long integration and call this the truth.
3. Create pseudo-observations at each time step and at every second grid point by adding a random Gaussian number $\zeta \sim N(0,R)$ to the truth at the observation sites.
4. Create a large number of replicate Earths having the same truth but differing time sequences of observations by repeating step 3 using differing random number seeds.
5. At time zero, create an *initial* analysis x^a by adding a random ζ to the truth at each grid point so that initially $x^a \sim N(\mathbf{x}^t, \mathbf{R})$ so initially $\mathbf{P}^a = \mathbf{R}$.
6. Create replicate Earths having the same truth but differing *initial* analyses by repeating 5 using differing random number seeds.
7. Create replicate Earths having the same truth but differing time sequences of analyses and forecasts by running a *deterministic* 10 member ensemble Kalman filter on each of the replicate Earths.
8. Choose R to be quite small so that error dynamics are approximately linear.

Climatological probability density of innovation variance

9. Compute squared innovation $(y\text{-}Hx^f)^2$ at each observation site at each forecast time

10. Take the average value of squared innovation across all replicate Earths. This is the *true* innovation variance at the observation site of interest at a given point in time *given the current true state*. The corresponding true forecast-error-variance=innovation-variance – *R*

11. Use values from step 10 to estimate climatological distribution of innovation variances

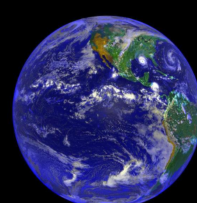


Black bars give frequency divided by bin size of occurrences of innovation variance

Red line gives inverse-gamma distribution having same mean and variance as observed innovation variances

Monte-Carlo approximation to replicate Earths

- Use models and observations to produce your single best-guess prediction
- Use your knowledge of unknowns in model and observations to produce an *ensemble* of plausible perturbed forecasts.
- Assume that the ensemble distribution of forecasts either gives $\rho\left(\mathbf{x}^t \mid \mathbf{y}_{i-1}, \mathbf{y}_{i-2}, \ldots\right)$ or $\rho\left(\mathbf{x}^f \mid \mathbf{x}^t = \textit{best guess}\right)$
- Compute the required measures of uncertainty from these distributions
- If the ensemble variance is smaller than the error variance of the ensemble mean, do something *ad-hoc* to account for the unknown sources of forecast error; e.g. amplify perturbations by some factor or add some sort of random noise to the perturbations.

Example of Monte-Carlo ensemble from Houtekamer et al. (2005, Mon. Wea. Rev.)

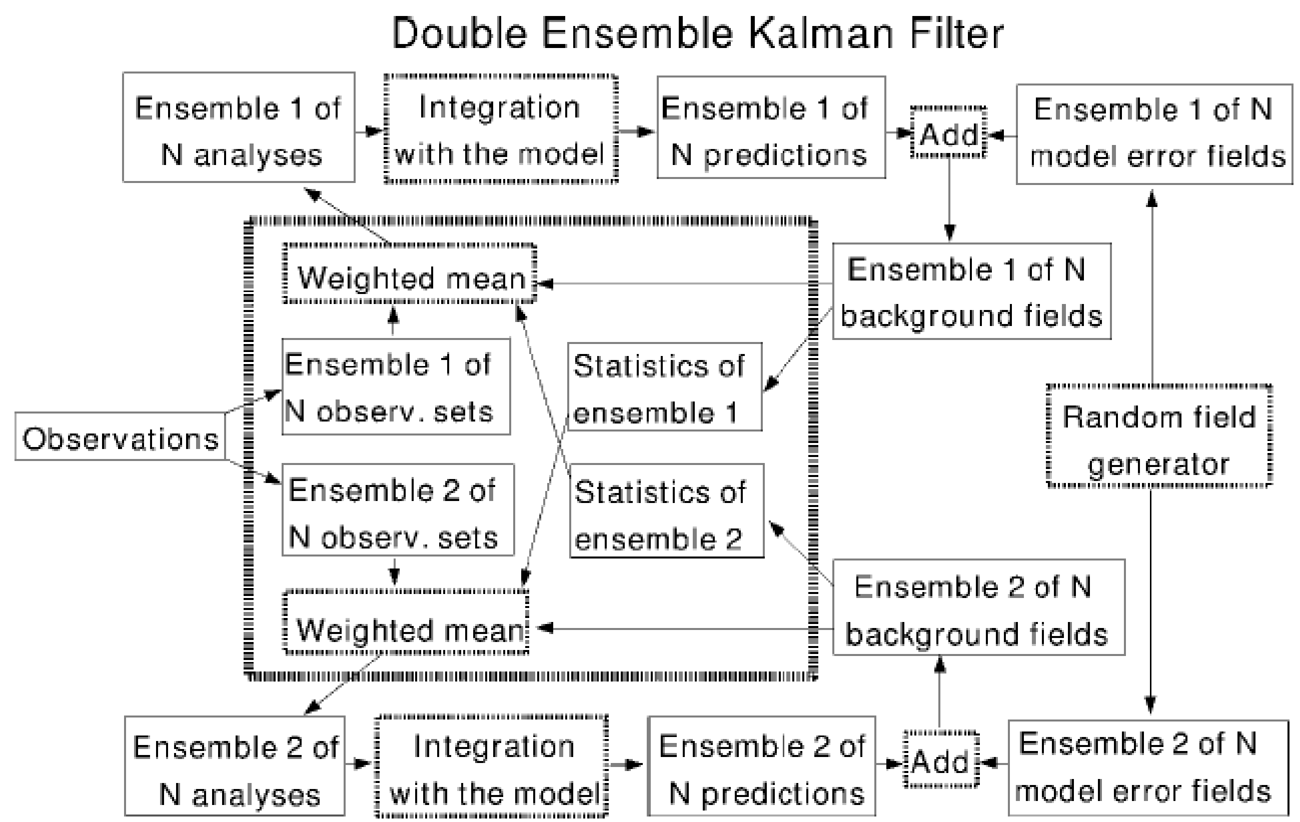
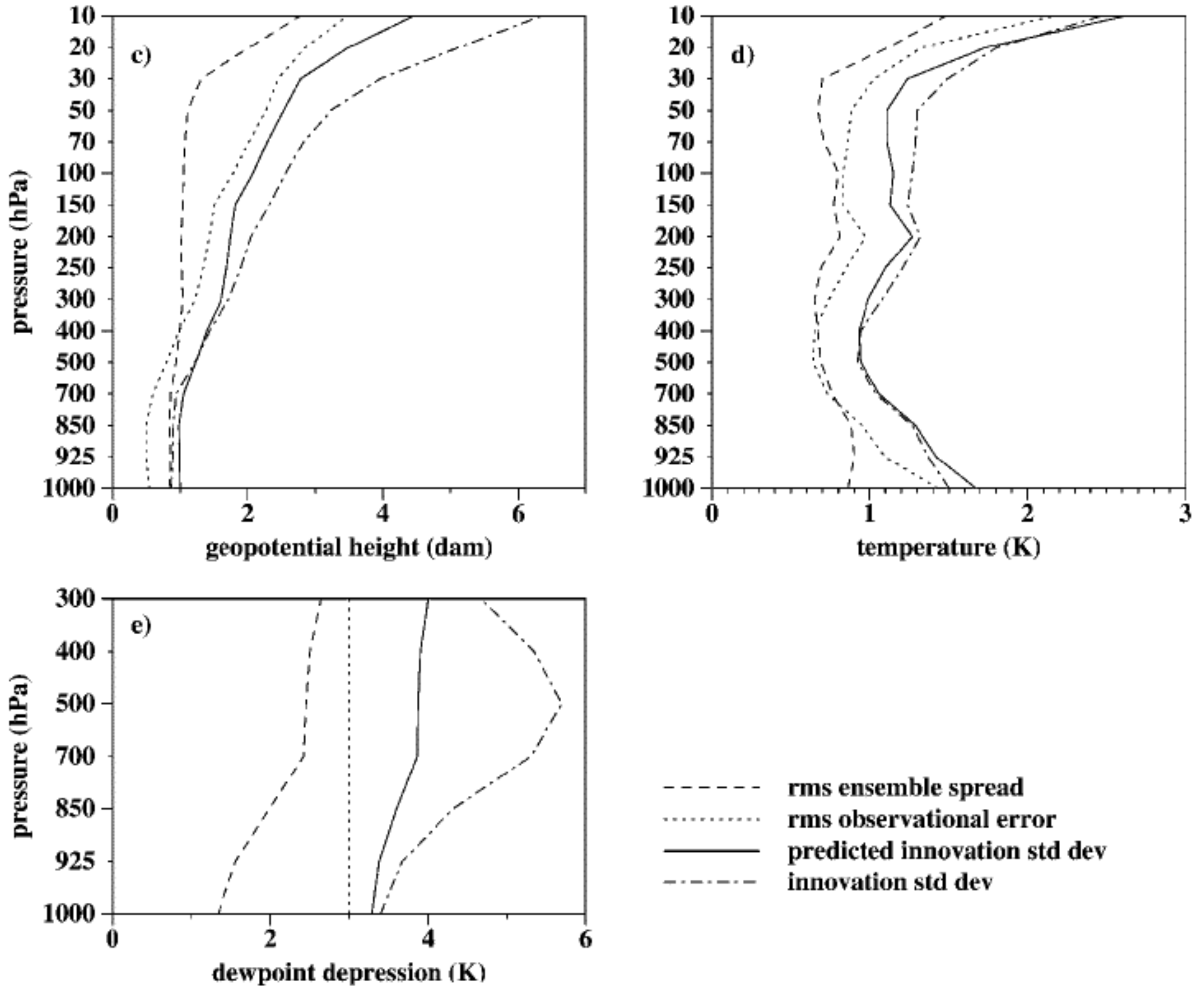


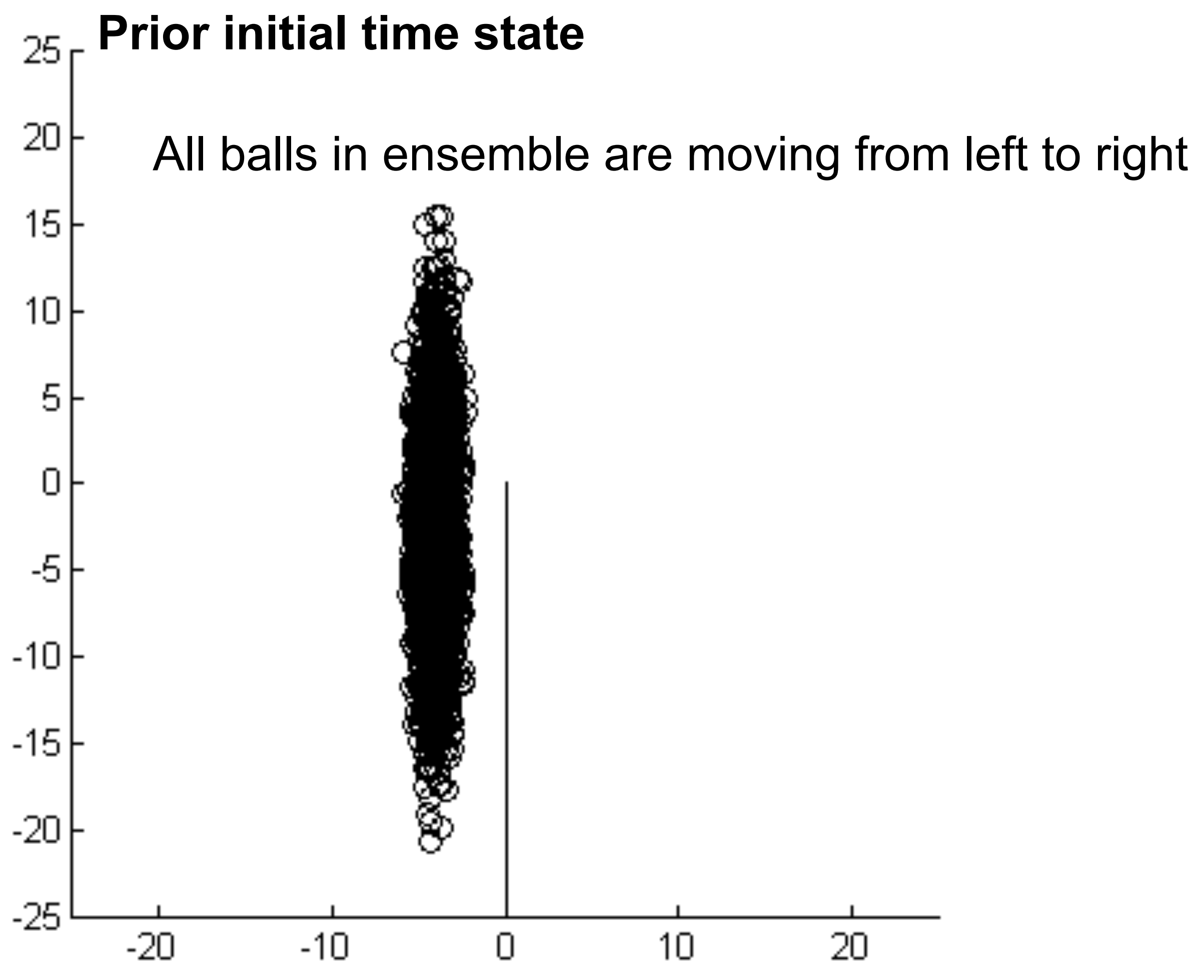
FIG. 1. The procedure used to perform a data assimilation cycle with the double EnKF.

Example of Monte-Carlo ensemble from Houtekamer et al. (2005, Mon. Wea. Rev.)

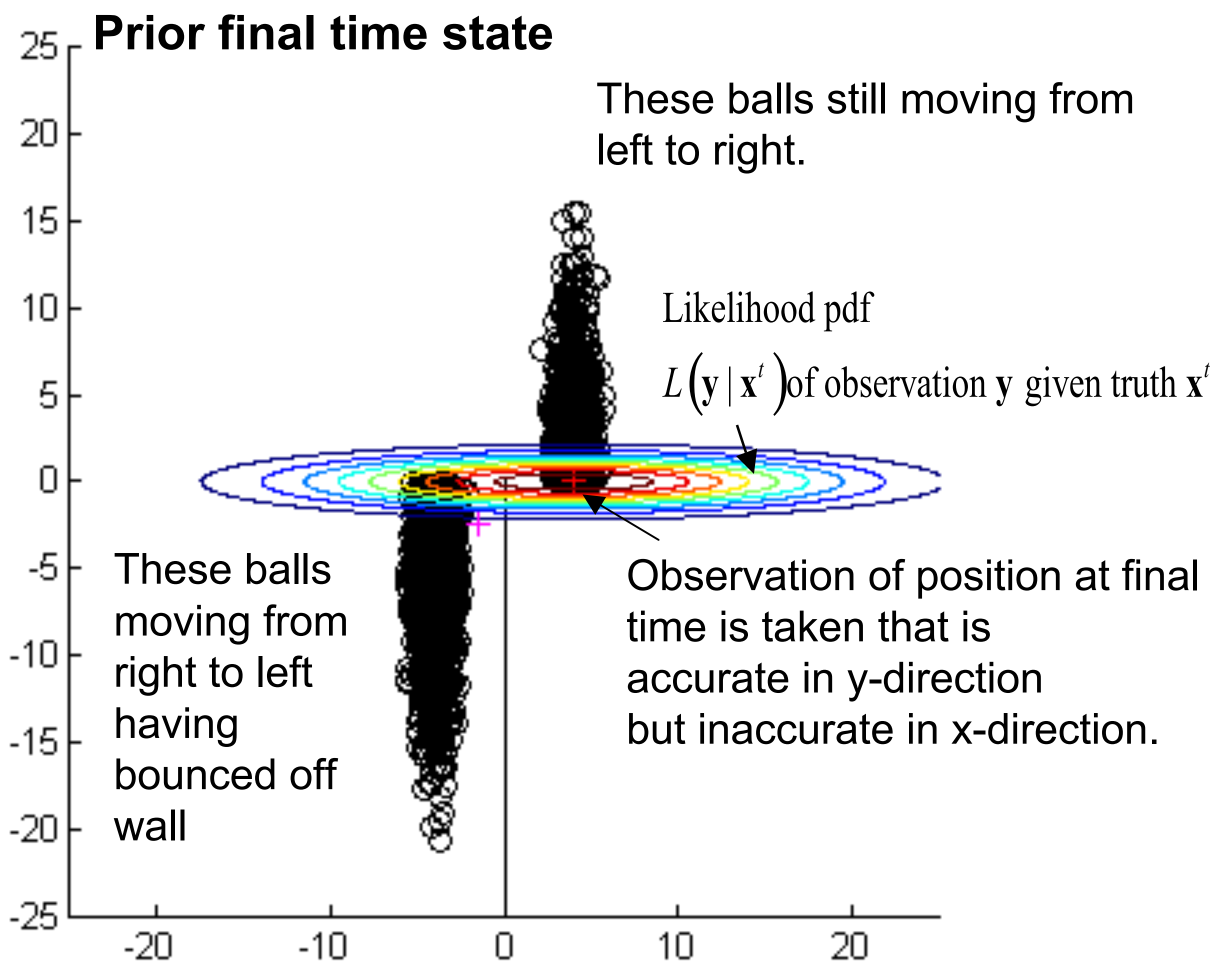


Comparison of predicted std of difference of forecast and observation (solid line) and ensemble prediction of same thing (.-line) indicates that, in an average sense, the Monte-Carlo approach worked.

Prior and posterior state estimates of poorly observed billiard balls



Prior and posterior state estimates of poorly observed billiard balls (particle filter).



Particle filter

1. Compute $L(\mathbf{y} \mid \mathbf{x}_i)$ for observation $\mathbf{y}$ and ith ball $\mathbf{x}_i$
2. Define weights $w_i \propto L(\mathbf{y} \mid \mathbf{x}_i)$ such that $\max(w_i) = 1$
3. Draw a random uniform number r between 0 and 1.

   if $r \leq w_i$

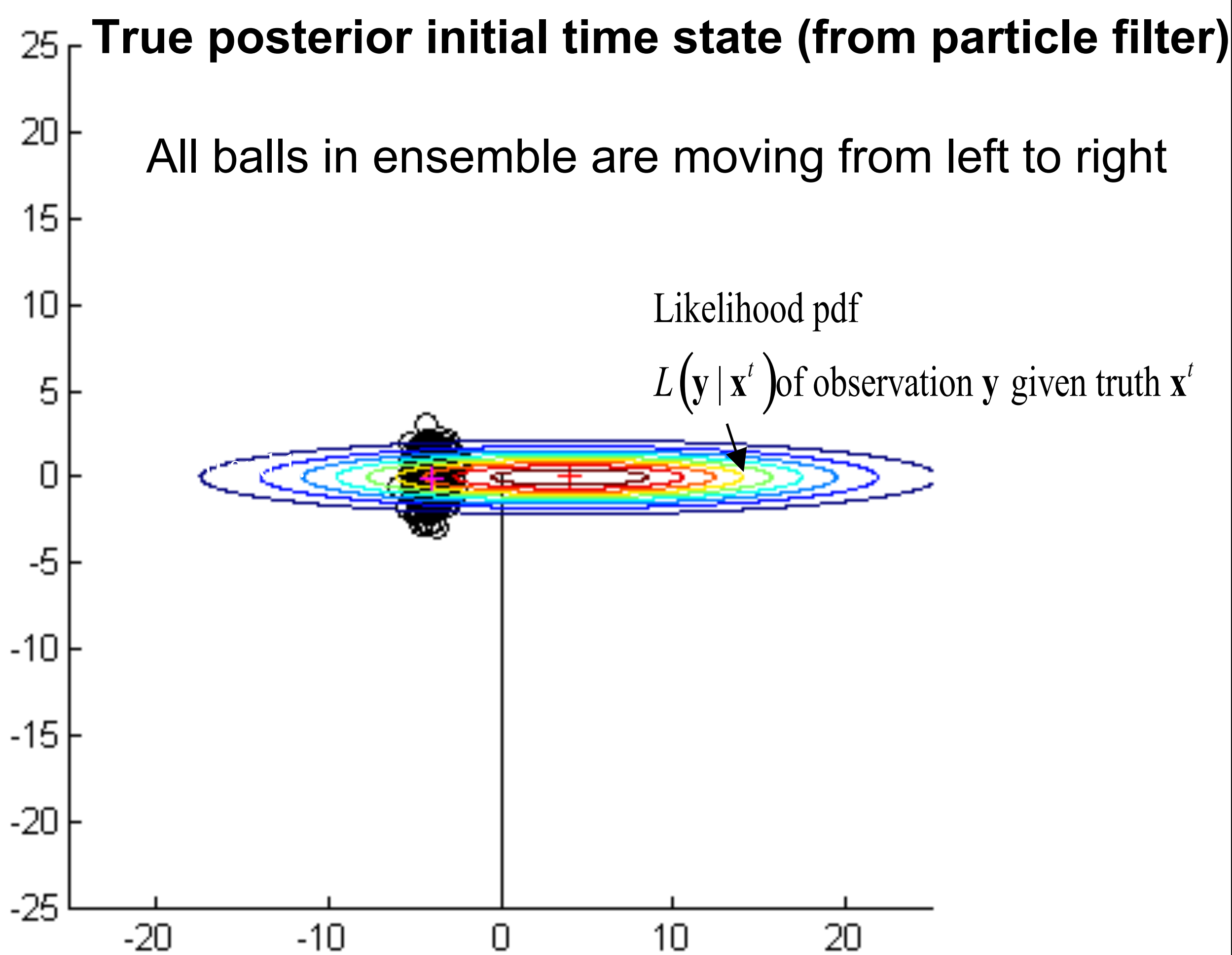
       select ith ball

     else

       reject ith ball

   end
4. Repeat step 3 for all balls.

Prior and posterior state estimates of poorly observed billiard balls (particle filter).



Particle filter

1. Compute $L(\mathbf{y}\,|\,\mathbf{x}_i)$ for observation $\mathbf{y}$ and ith ball $\mathbf{x}_i$
2. Define weights $w_i \propto L(\mathbf{y}\,|\,\mathbf{x}_i)$ such that $\max(w_i) = 1$
3. Draw a random uniform number r between 0 and 1.

   if $r \le w_i$

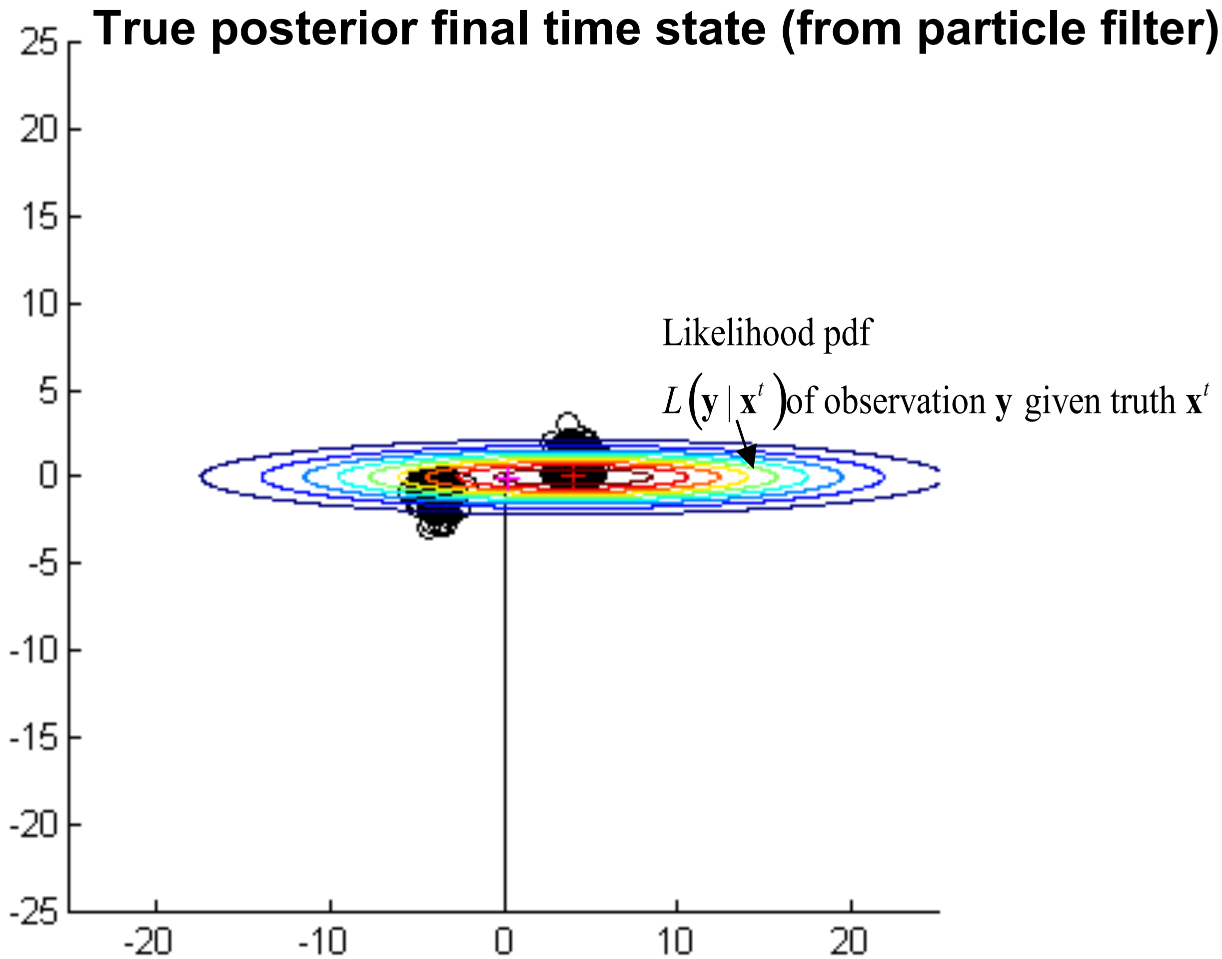
   select ith ball

   else

   reject ith ball

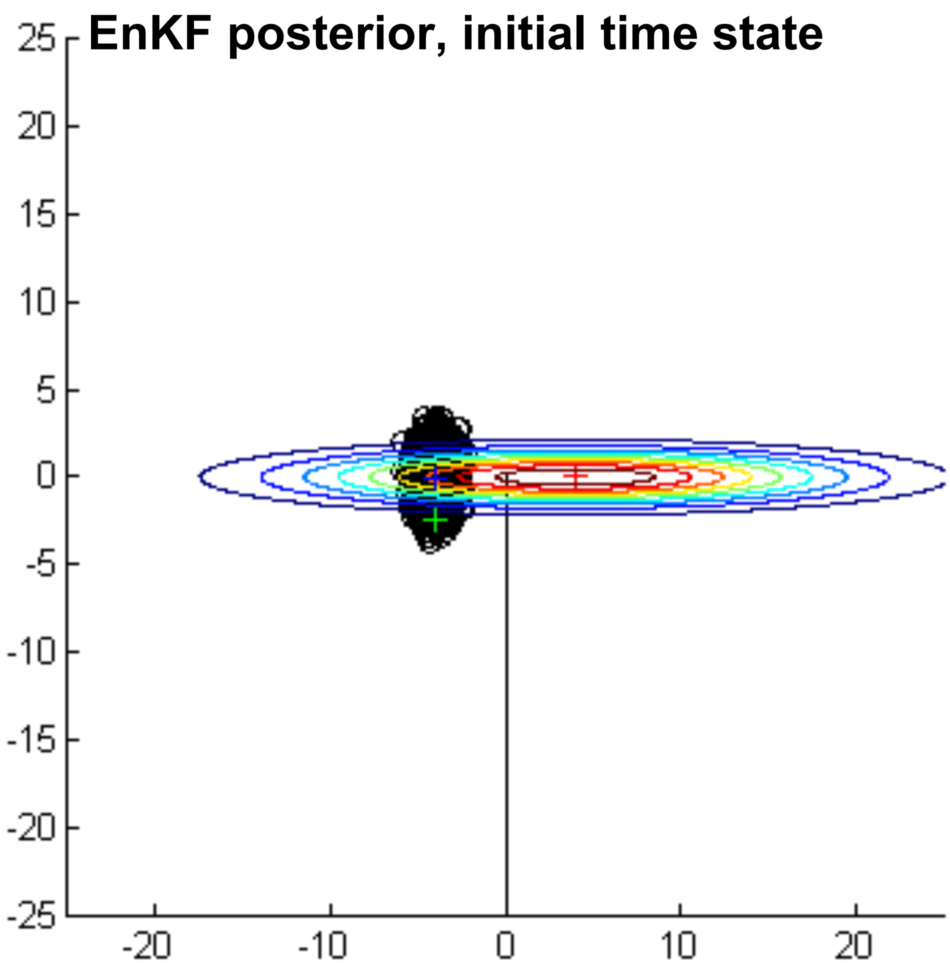
   end
4. Repeat step 3 for all balls.

Prior and posterior state estimates of poorly observed billiard balls (particle filter).



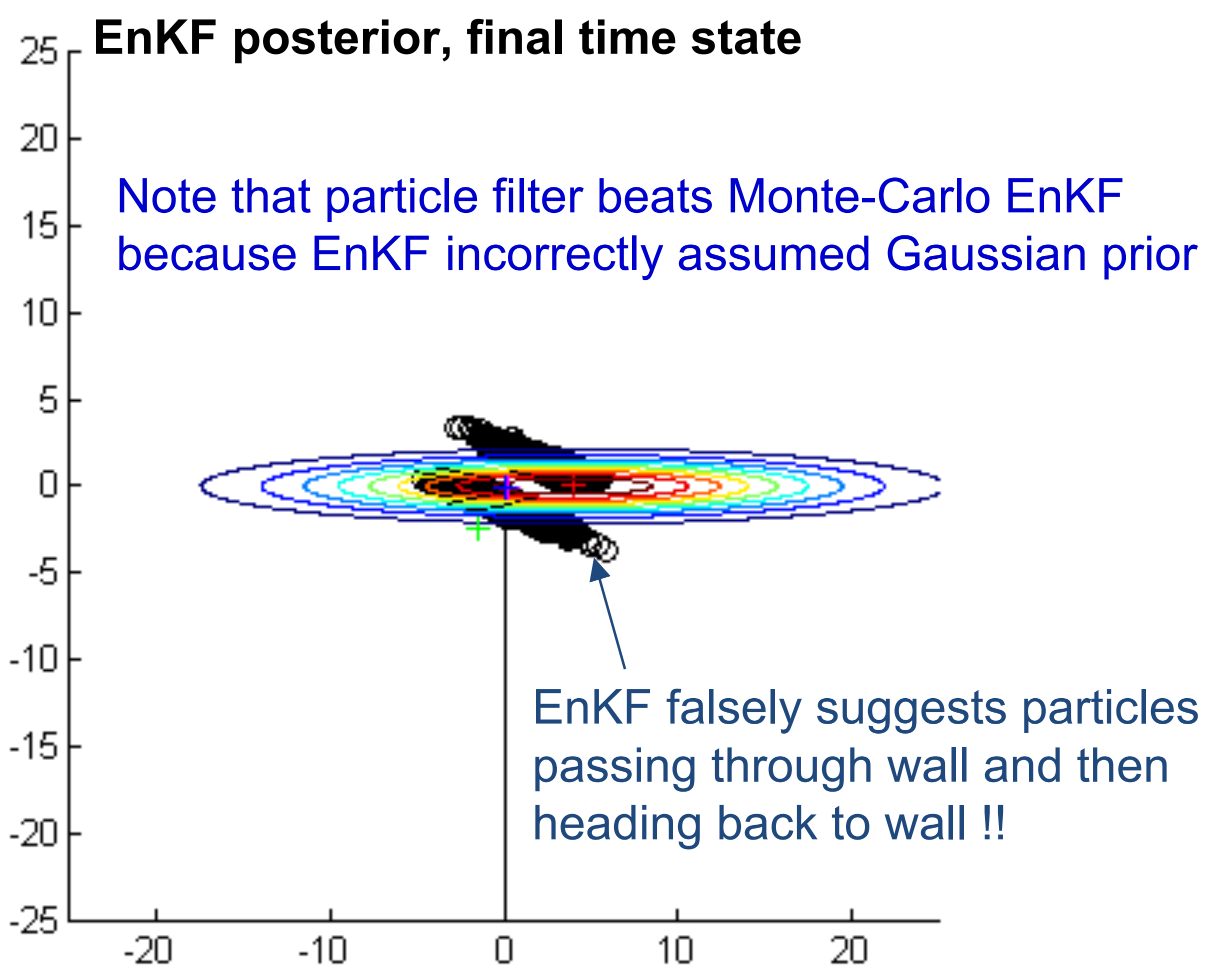
It is this final time posterior distribution of states that should be used to initialize the next ensemble forecast. The resulting ensemble forecast will be used to assimilate observations in the next data assimilation window.

Prior and posterior state estimates of poorly observed billiard balls (EnKF).



This initial time posterior state from the EnKF is very accurate.

Prior and posterior state estimates of poorly observed billiard balls (EnKF).



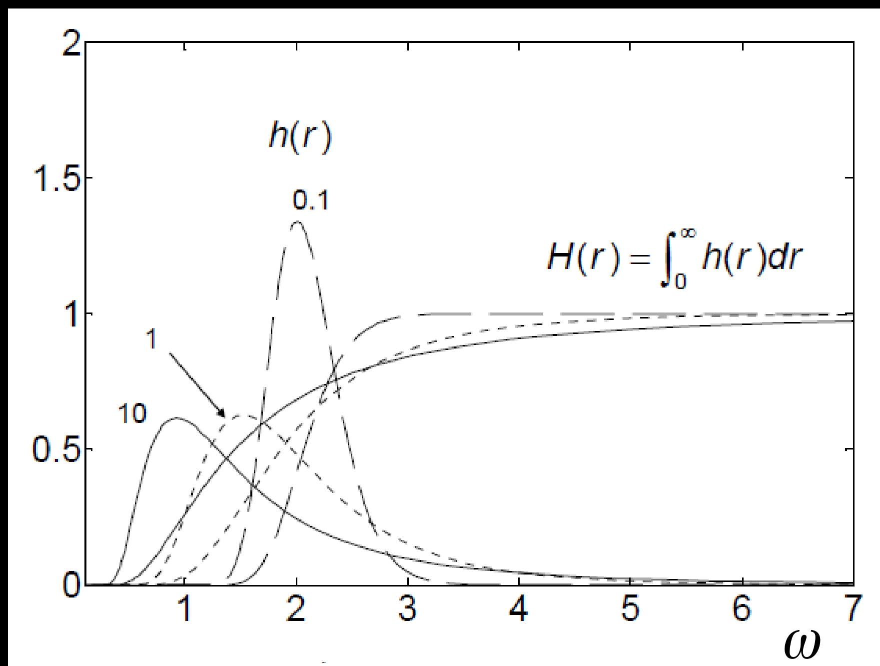
This initial time posterior state from the EnKF is very **inaccurate.** Initializing next ensemble with this set of particles would give a very inaccurate mean and covariance for the next DA window.

Recapitulation

- Uncertainty quantification not so easy for chaotic systems in which it is difficult to find “repeat” or “near repeat” events
- Replicate Earth thought experiment represents “ideal” quantification of certainty and uncertainty
- In Monte-Carlo approach all known unknowns of a prediction system are varied across an ensemble of predictions. The resulting ensemble of predictions may be viewed as a proxy for an ensemble of replicate Earths.
- However, if the system fails to accurately account for the effect of observations and/or model error on uncertainty, the measures of uncertainty derivable from the ensemble will be inaccurate.
- The desired *conditioned* forecast error covariance is *hidden* from observations because in aperiodic chaotic systems the conditions do not repeat.

Quantifying uncertainty in presence of Hidd Volatility (1D idealization)

1. Assume distribution of true climatological innovation variances ω is an inverse-gamma distribution. (innovation $= y - H\left(x^f\right)$)



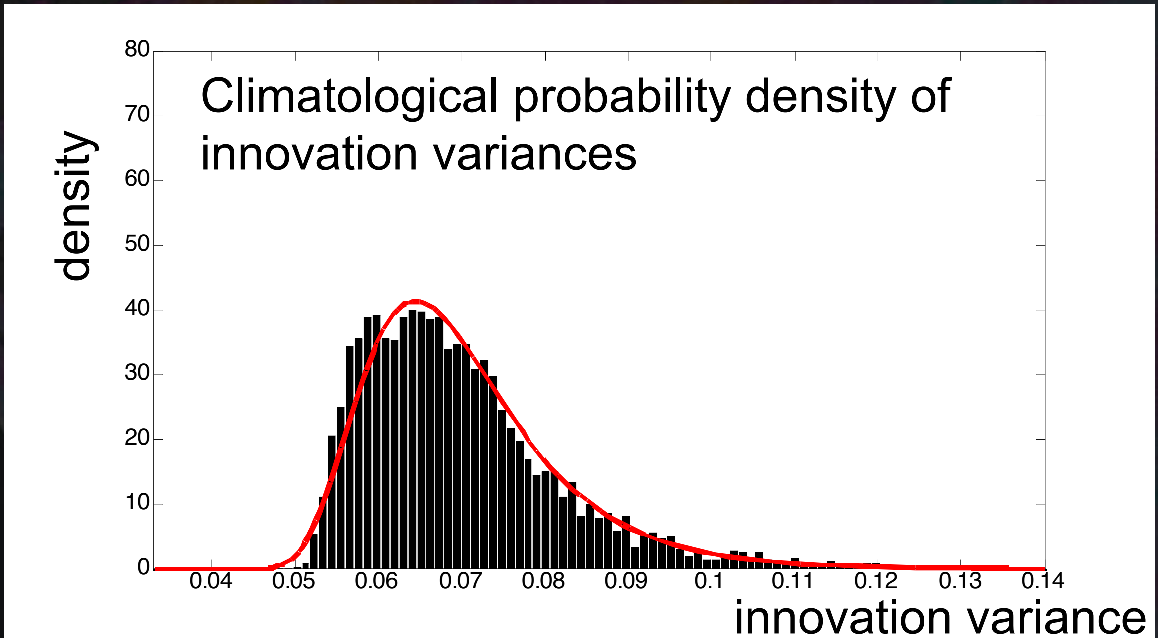
Note that inverse-gamma ensures that innovation variance is never equal to zero.

$t = \omega - R$ is the forecast error covariance and

$\overline{x^f} = x^t + \varepsilon^f$, where $\left\langle \varepsilon^f \right\rangle = 0$ and $\left\langle \left(\varepsilon^f\right)^2 \right\rangle = t$

- Assumption 1 is supported by replicate Earth experiment using Lorenz model

Prior climatological pdf $\rho(\omega)$



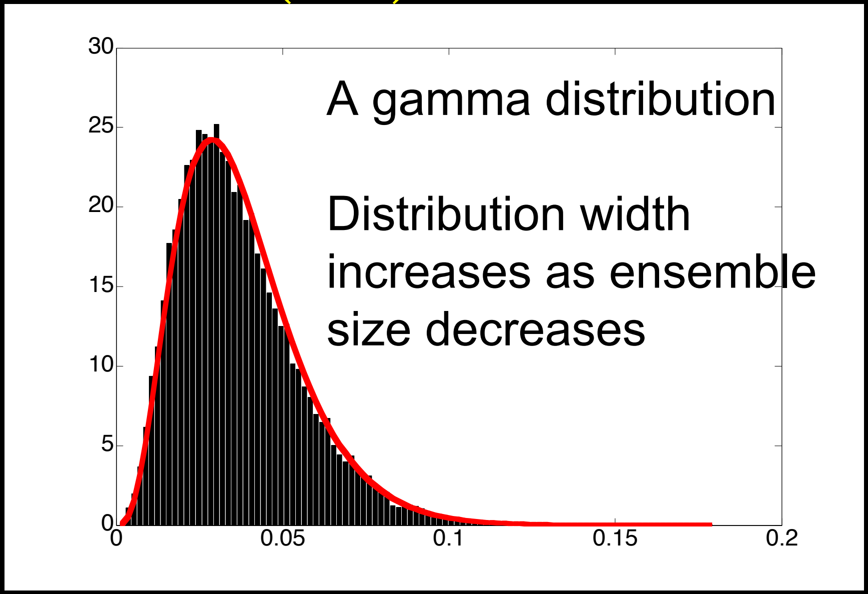
Black bars give frequency divided by bin size of occurrences of innovation variance

Red line gives inverse-gamma distribution having same mean and variance as observed innovation variances

Quantifying uncertainty in presence of Hidden Volatility (1D idealization)

2. Assume that ensemble sample variance is a gamma distribution with mean $s^2 = s^2_{\min} + a\left(\omega^2 - R\right)$ and variance determined by an effective ensemble size M. R is the observation error variance and a is a constant. (See Figure)

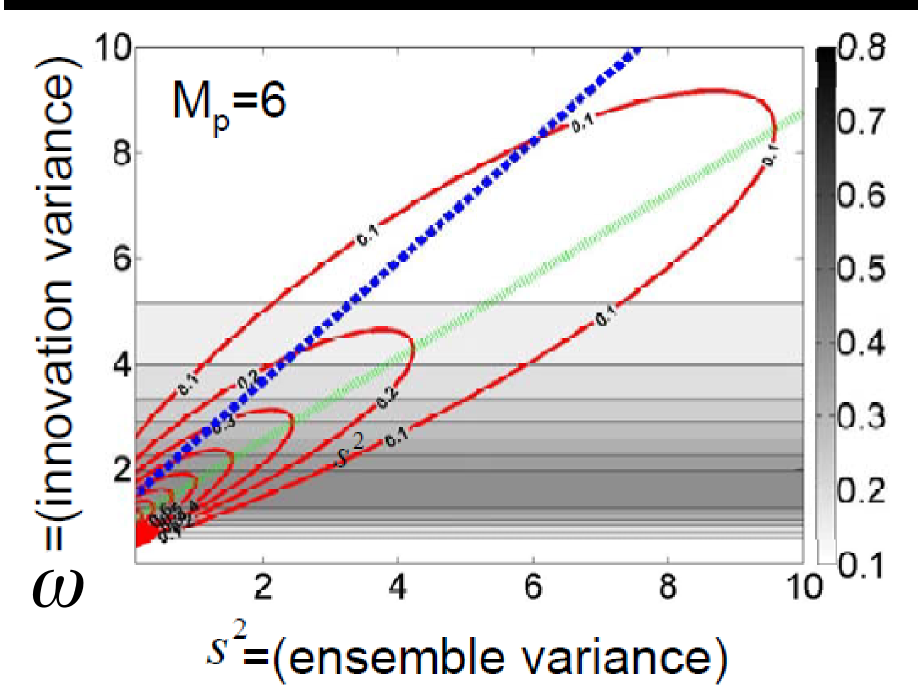
Likelihood density $\rho\left(s^2 \mid \omega\right)$

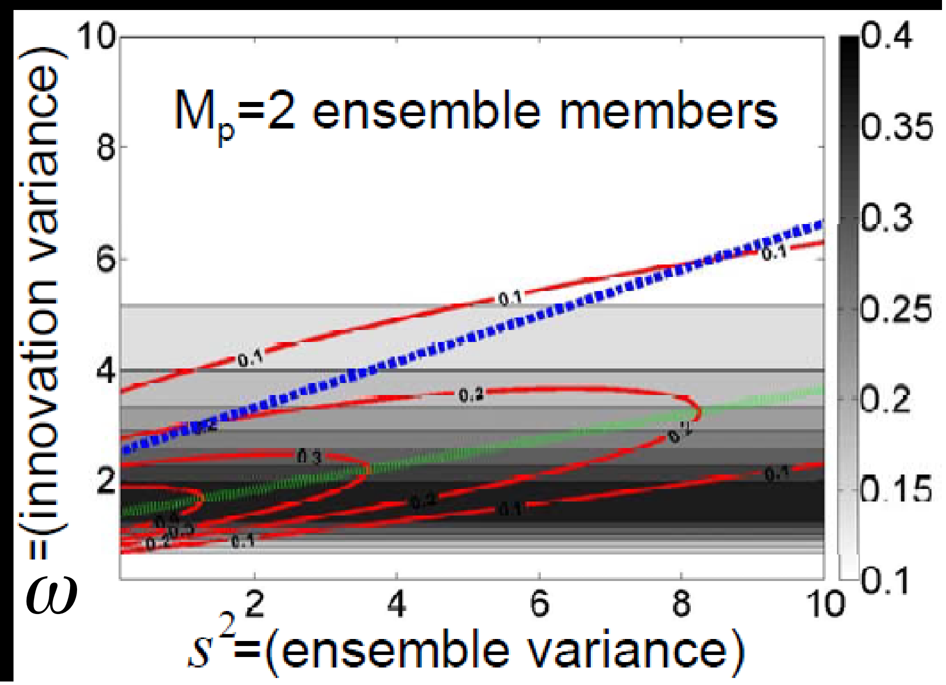


Quantifying uncertainty in presence of Hidden Volatility (1D idealization)

3. Use Bayes' theorem to combine prior $\rho(\omega)$ and likelihood $\rho(s^2 \mid \omega)$ to obtain posterior pdf $\rho(\omega \mid s^2)$ of true variances given the ensemble sample variance s^2.

$\rho(\omega \mid s^2)$ $\quad a = 0.5$ and $s^2_{\min} = 0$ $\quad \rho(\omega \mid s^2)$





Quantifying uncertainty in presence of Hidd Volatility (1D idealization)

The equation $s^2 = s_{\min}^2 + a\left(\omega - R\right)$ implies a naive estimate of the true innovation variance given by $\omega_n = R + \dfrac{s^2 - s_{\min}^2}{a}$.

Can show that the posterior mean innovation variance has the form

$\bar{\omega}_{post} = w\omega_n + \left(1 - w\right)\bar{\omega}_{prior}$ where w is a weight between 0 and 1.

$w \sim \dfrac{\left(1/r_n\right)}{\left(1/r_n\right) + \left(1/r_{prior}\right)}$, where r_n and r_{prior} are relative error variances of ω_n and $\bar{\omega}_{prior}$

When ensemble variance prediction is imperfect, the "best" error variance estimate is a linear combination of a flow dependent estimate and a static estimate.

Amounts to a 1st principles motivation for Hybrid covariance model that is a linear combination of a flow dependent estimate and a static estimate.

Quantifying uncertainty in presence of Hidd Volatility (1D idealization)

- It may be shown that all of the parameters defining this simple model of error variance prediction in the presence of hidden volatility may be obtained using appropriate regression formula from a large number of independent (ensemble-variance, innovation) pairs.
- This enables
  - Weights for Hybrid error covariance models
  - Measurement of hidden volatility prediction accuracy in terms of "effective-ensemble size".
  - Ensemble post-processing that accounts for hidden volatility prediction inaccuracy.

Effective ensemble size of 32 member ET ensemble T49L42 only conventional obs

- *Effective number of ensemble members:*
  Northern Hemisphere: 6.3
  Tropics: 5.1
  Southern Hemisphere: 6.0
- i.e. the 32 member ensemble is behaving as if it were comprised of 6 random draws from a distribution having the true innovation variance.
- Suggests that a Hybrid forecast error covariance model that mixes ensemble covariances with static covariances would be superior to one based on ensemble covariances alone.

Hybrid in Navy 4D-VAR

- Terminology
  - Navy 4D-VAR is in observation space and is called NAVDAS-AR
  - Have created a new form of NAVDAS-AR that allows for the incorporation of localized ensemble covariances. We will call this Hybrid form Ensemble-AR

How is ensemble-AR different from NAVDAS-AR?

NAVDAS-AR has

$$\underline{\mathbf{P}}_{NAVDAS-AR} = \begin{bmatrix} \mathbf{P}_0^b & \mathbf{P}_0^b \mathbf{M}^T \\ \mathbf{M}\mathbf{P}_0^b & \mathbf{M}\mathbf{P}_0^b \mathbf{M}^T + \mathbf{Q} \end{bmatrix}$$

Ensemble-AR currently enables the following option for the initial covariance

$$\mathbf{P}_{0\_Hybrid}^b = \alpha \mathbf{P}_{0\_Ensemble}^b + (1-\alpha)\mathbf{P}_0^b$$

The code also allows for variations in α across 6 geographical regions. We can also run the code with the TLM and adjoint turned off by using the 4D localized ensemble covariance

$$\underline{\mathbf{P}}_{ENSEMBLE-AR} = \mathbf{N}\left\{\left[\frac{1}{K-1}\sum_{i=1}^{K}(\underline{\mathbf{x}}_i - \underline{\overline{\mathbf{x}}})(\underline{\mathbf{x}}_i - \underline{\overline{\mathbf{x}}})^T\right] \circ \mathbf{C}\right\}\mathbf{N}^T$$

$$= \mathbf{N}\left\{\left[\frac{1}{K-1}\sum_{i=1}^{K}(\underline{\mathbf{x}}_i - \underline{\overline{\mathbf{x}}})(\underline{\mathbf{x}}_i - \underline{\overline{\mathbf{x}}})^T\right] \circ \mathbf{W}\mathbf{W}^T\right\}\mathbf{N}^T$$

where $(\underline{\mathbf{x}}_i - \underline{\overline{\mathbf{x}}})$ is ensemble perturbation, $\mathbf{C}$ is localization matrix and $\mathbf{N}$ is TLM of balance operator ($\mathbf{N}$ is optional). $\mathbf{P}_{0\_Ensemble}^b$ is obtained by using the above equation but only computing the part of it that pertains to the initial time.

Description of Experiment

- Cycling analysis from Nov. 20, 2008 to Dec. 31, 2008
- Assimilating only conventional observations (no radiances)
- Background error covariance matrix at beginning of DA time window (Pb0) is a combination of 75% Pb0_static and 25% Pb0_ensemble

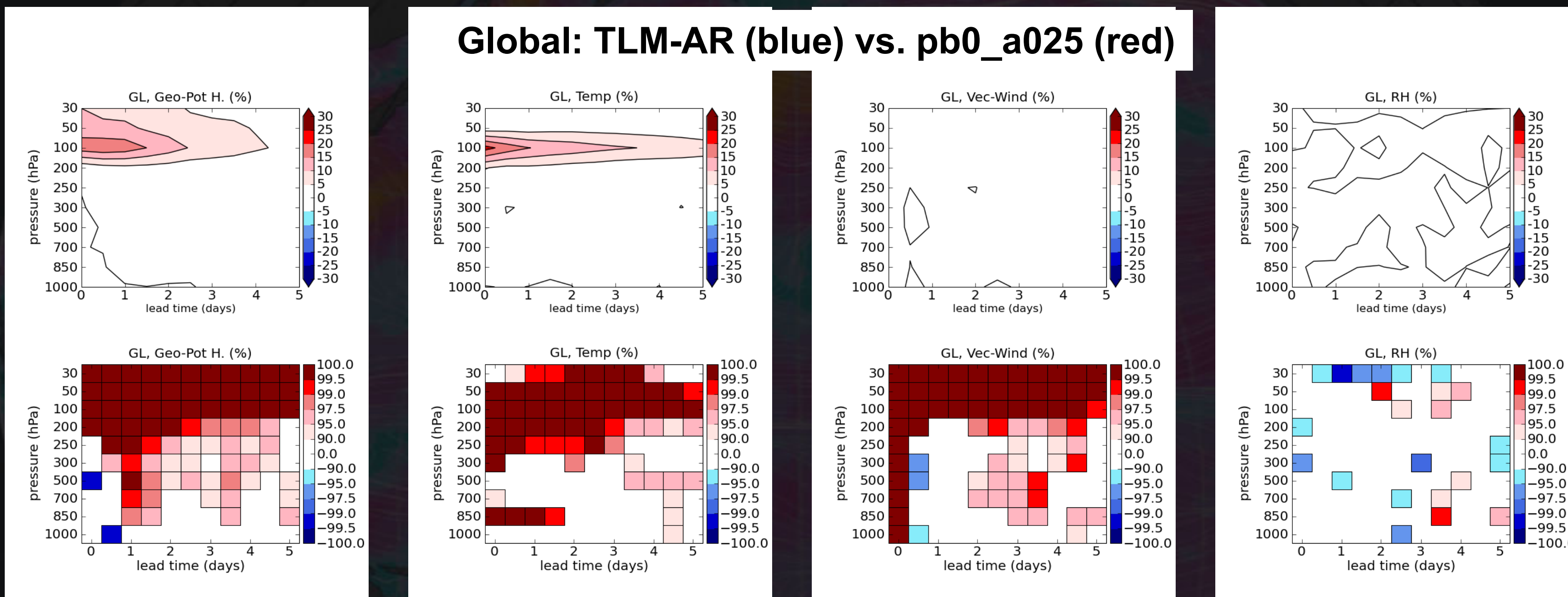
Pb0_hybrid=0.75*Pb0_static+0.25*Pb0_ensemble

- 32 member ET ensemble (Bishop&Toth, 99,McLay et al 08)
- Model resolution: T119L42 Outer, T47L42 Inner
- Skip first 10 days of analysis for spin-up
- 5-day forecasts from each analysis
- Verification of forecasts with Radiosondes
- Comparison with current NAVDAS-AR at same resolution using static covariance model

The red in the panels below show the improvement of the DA system when the initial Hybrid error covariance model

$\mathbf{P}^{b}_{0\_Hybrid} = 0.25\mathbf{P}^{b}_{0\_Ensemble} + 0.75\mathbf{P}^{b}_{0\_Static}$ is used instead of just $\mathbf{P}^{b}_{0\_Static}$

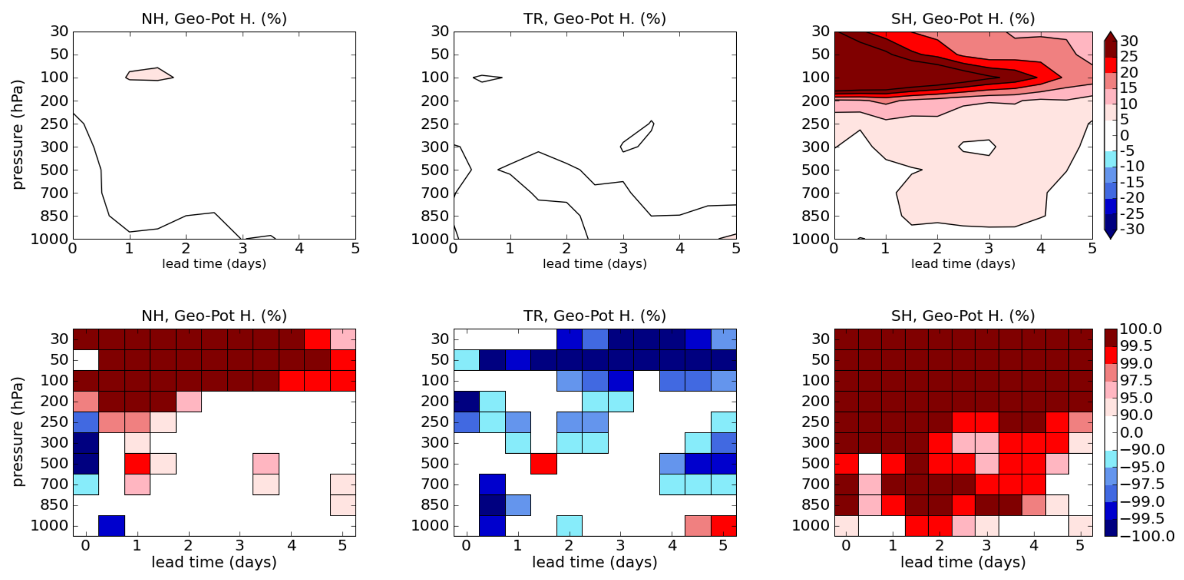
These results are for a relatively small ensemble (32 members).



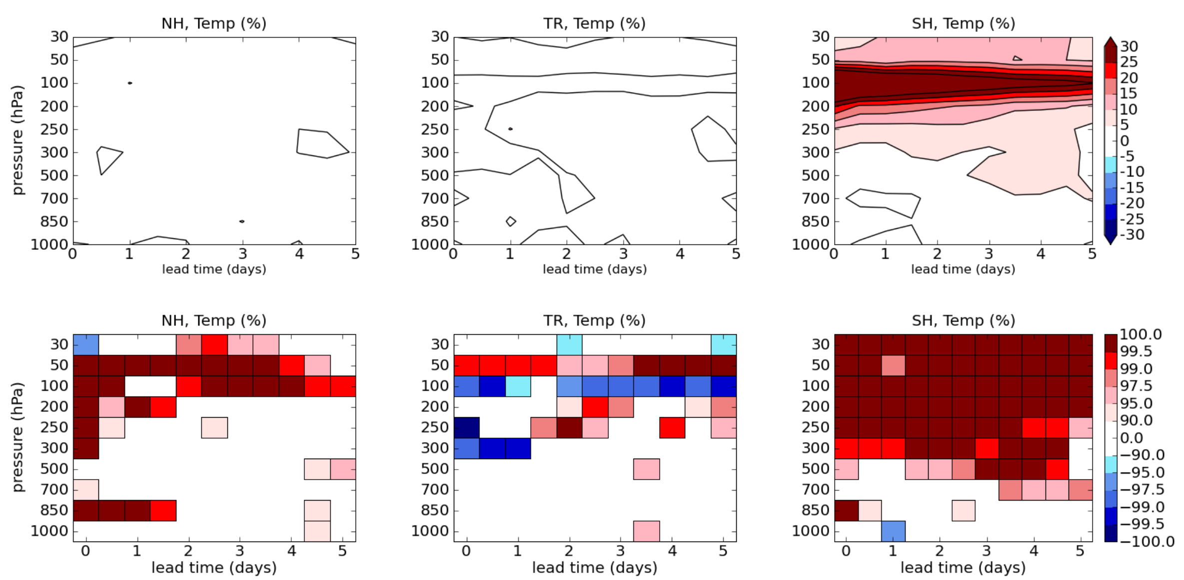
Contours filled with red indicate regions where Pb0_hybrid reduced rms error (as measured by radiosondes) by more than 5% relative to Pb0_static. The stronger the tone of red the greater the percentage reduction. Colored blocks indicate the significance of difference with red again signifying superiority of Pb0_hybrid.

Hybrid_Ensemble_AR performed significantly better than current system

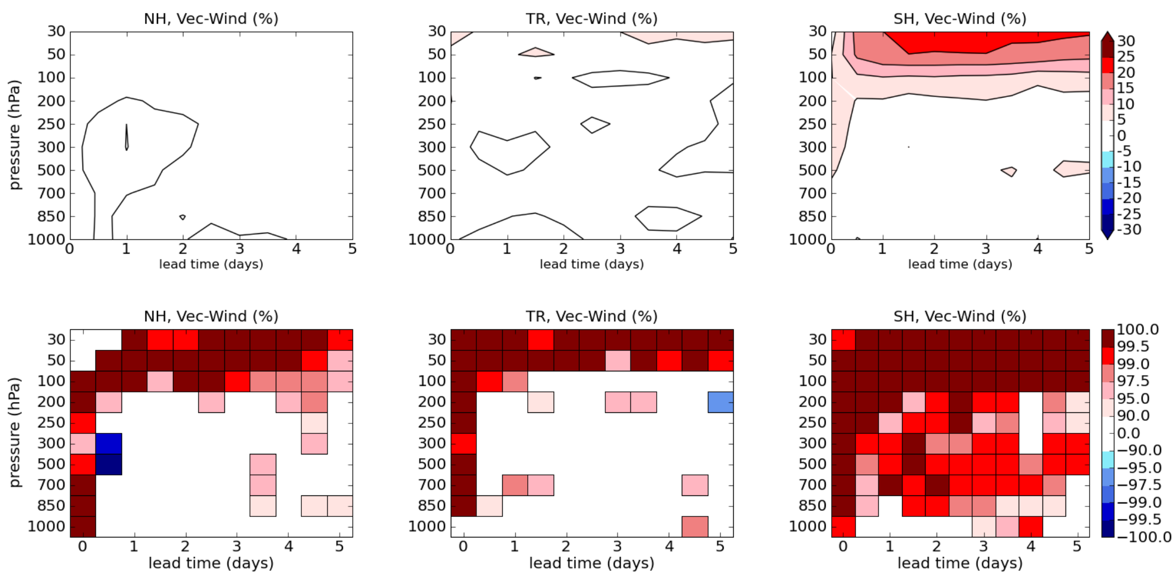
Height: TLM-AR (blue) vs. pb0_a025 (red)



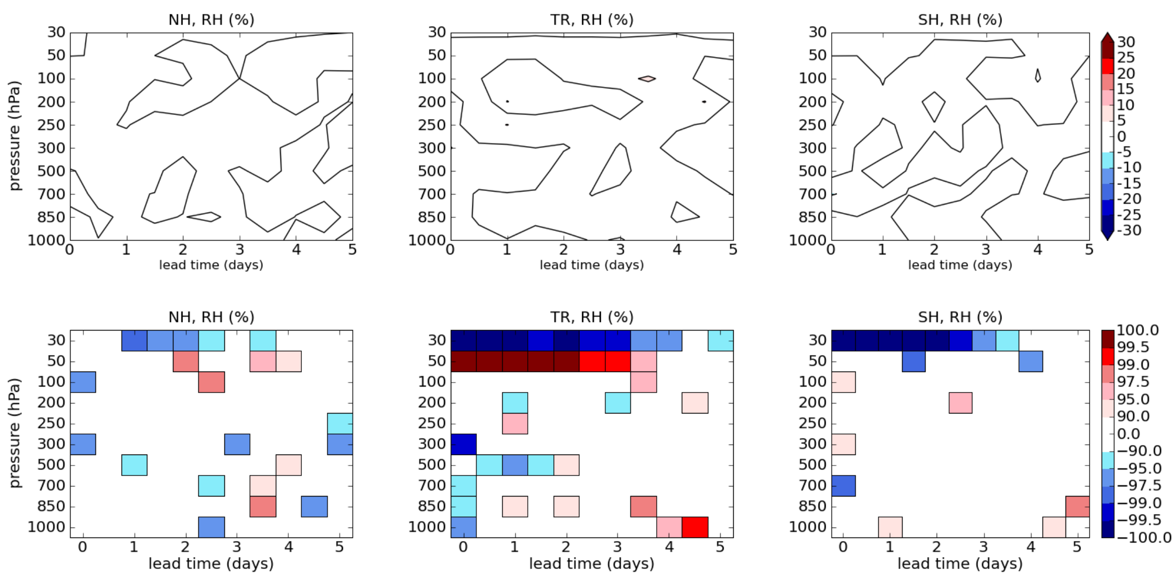
Temp: TLM-AR (blue) vs. pb0_a025 (red)



Vec-Wind: TLM-AR (blue) vs. pb0_a025 (red)



Rel. Humid.: TLM-AR (blue) vs. pb0_a025 (red)

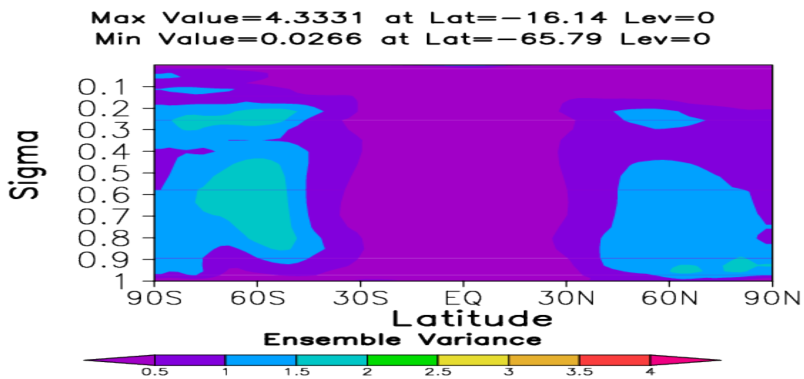


Derivation of Hybrid weights from analysis innovations

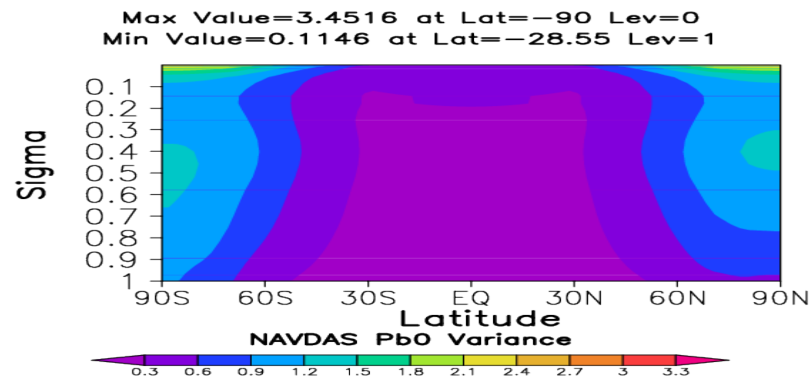
- Assuming that all of the assumptions of simple model of error variance prediction are satisfied, Hybrid weights for optimal error variance prediction may be derived using appropriate regression formula from a large number of independent (ensemble-variance, innovation) pairs.

Comparison of variances from globally tuned Hybrid weights with those from theoretically optimal weights for 6 regions (SH, Tropics, NH all below 400 hPa and then all 3 again above 400 hPa)

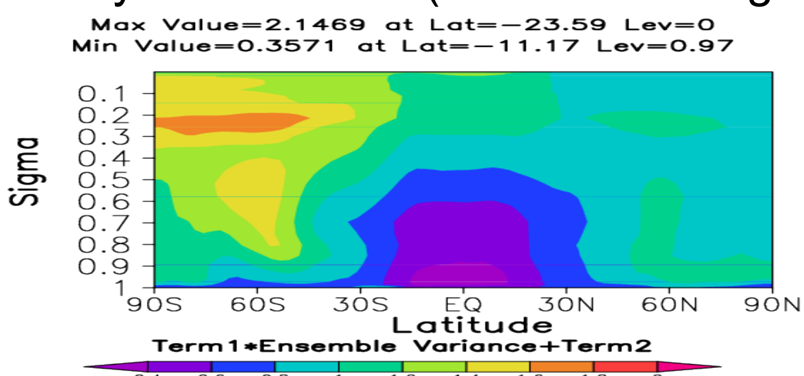
Average ensemble variance



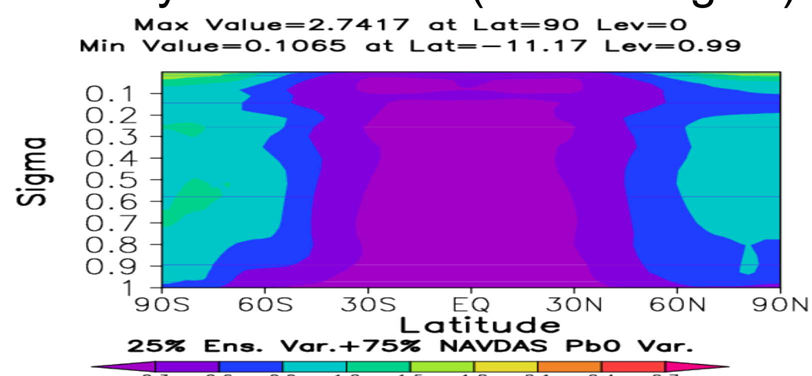
Static variance from Pb0



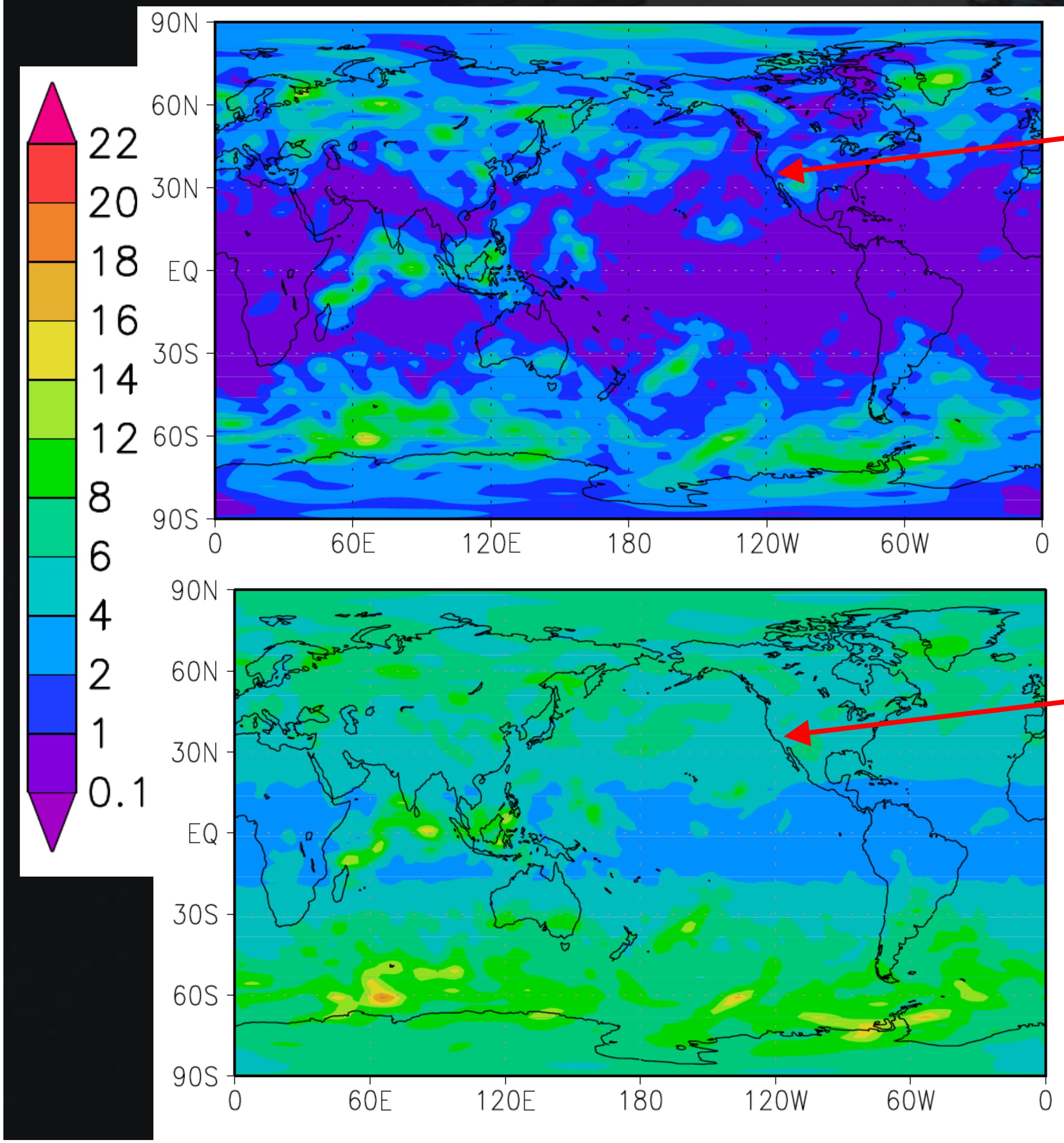
Hybrid variance (theoretical weights)



Hybrid variance (tuned weights)



U-wind 850mb



Raw 6 hr ET variance exhibits large regions of very small variance. This spot is on the West Coast where radiosondes are present but the ensemble variance is below 2.

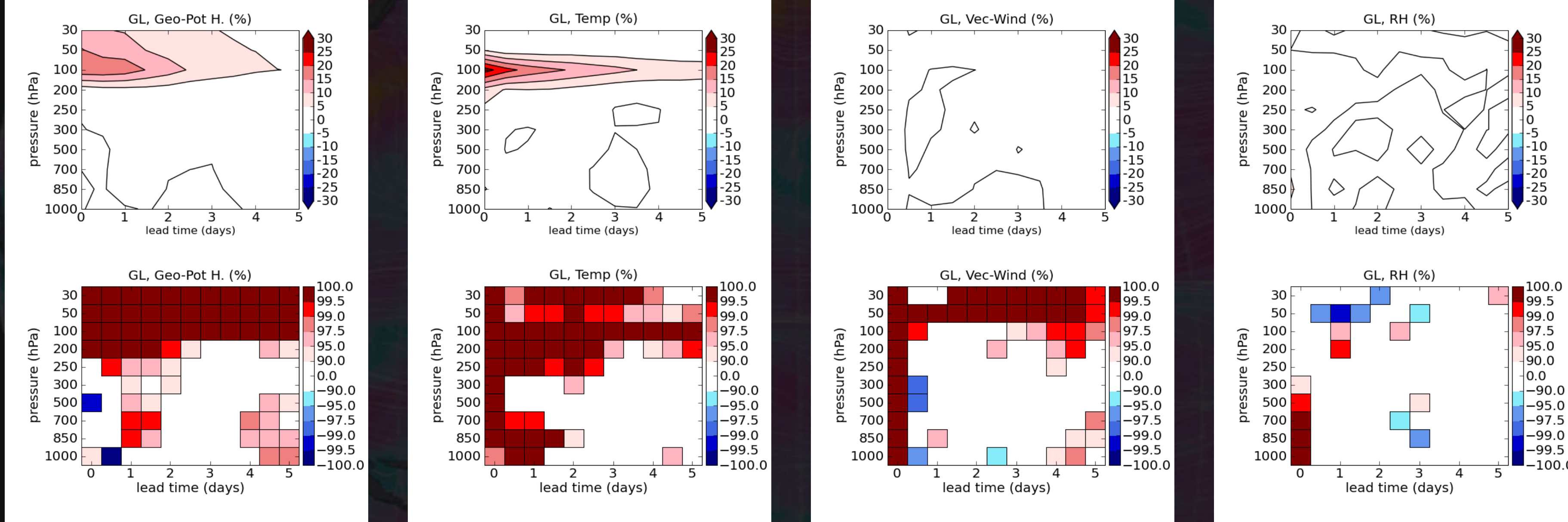
Equation hybrid variance exhibits a smaller range of variances. The analysis will be able to draw to the observations near the West Coast.

Test of ensemble-AR using eq Hybrid

- Machine: 8 processor Linux cluster
- 32 member ET ensemble
- Observations: All conventional obs
- Period: 11/20/08 to 12/31/08 (1st week removed from assessment and last 5 days from assessment of 120 hr forecasts)
- 6 hr DA cycle.

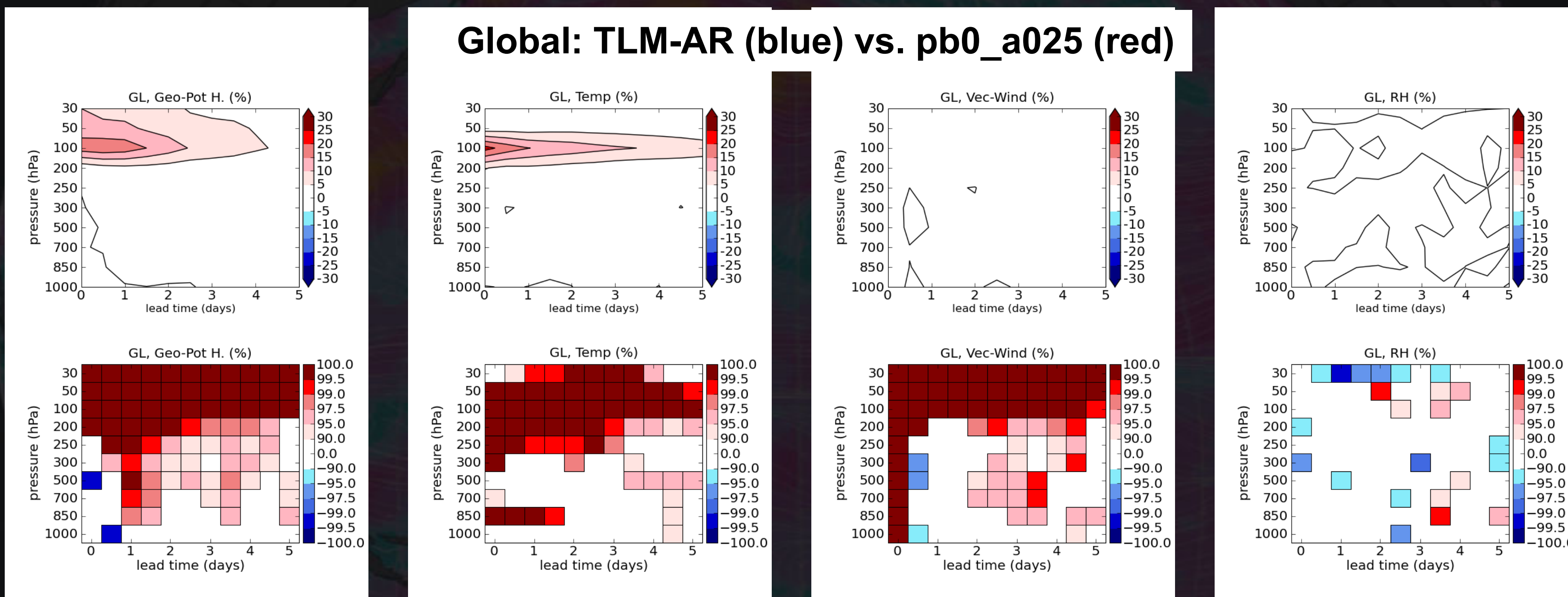
Equation Hybrid

Global: TLM-AR (blue) vs. pb0206E1 (red)



Globally, performance of Equation Hybrid and standard Hybrid is about the same – but no costly tuning experiments are required for equation Hybrid

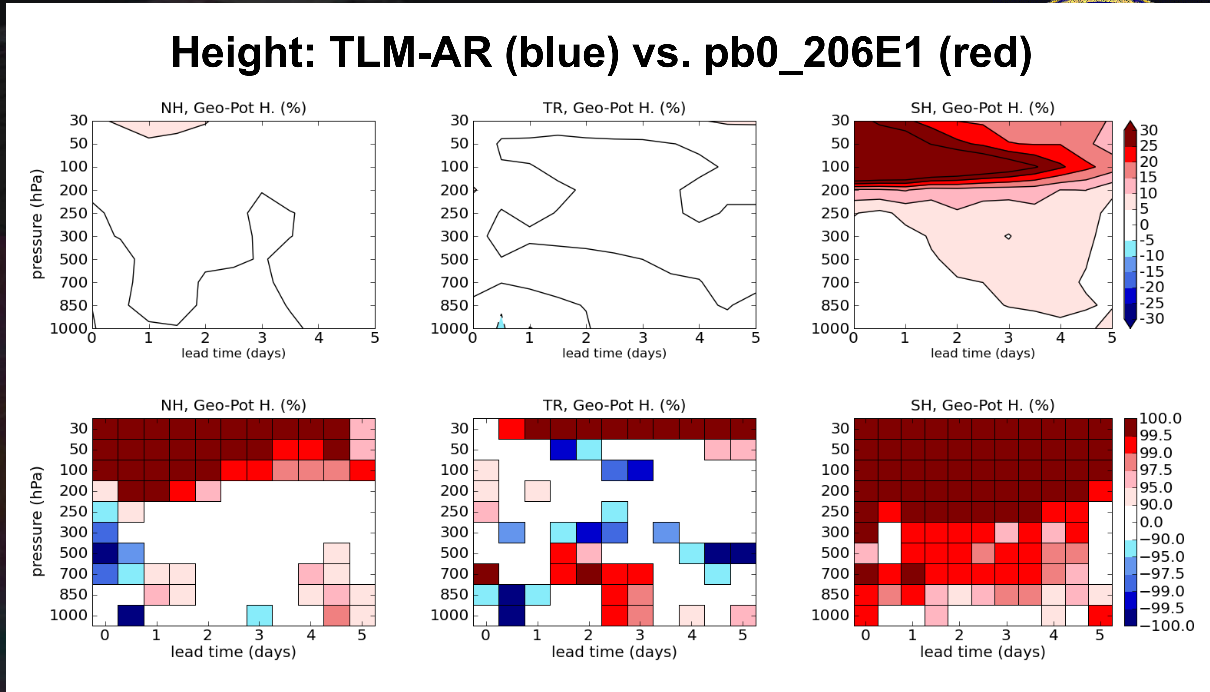
Standard Hybrid

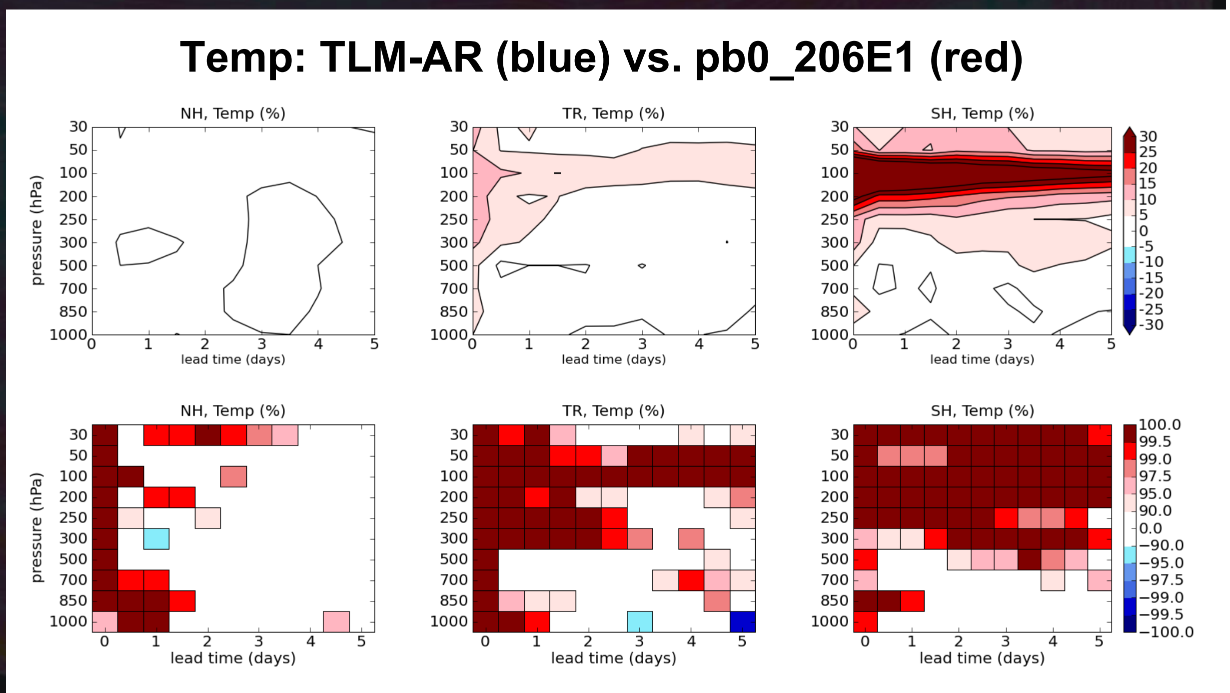


Globally, performance of Equation Hybrid and standard Hybrid is about the same – but no costly tuning experiments are required for equation Hybrid

Equation Hybrid

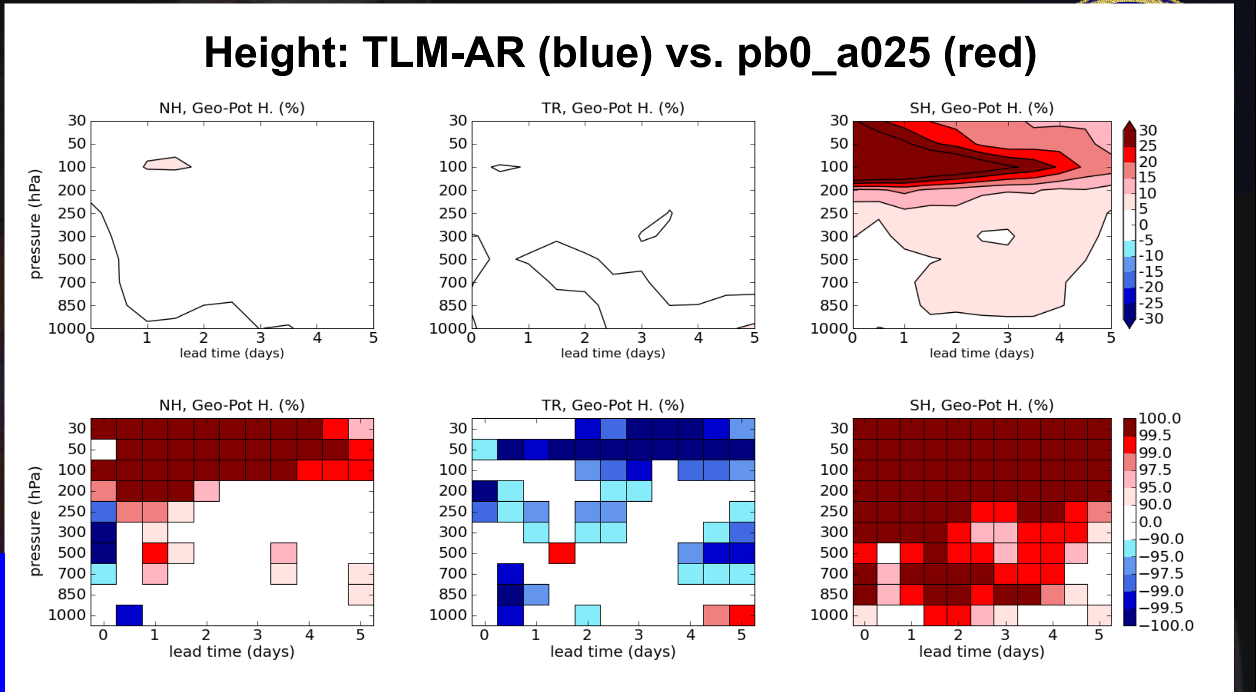
Regionally, the equation-Hybrid is better at avoiding areas of degradation than the standard-Hybrid. (tropics particularly)

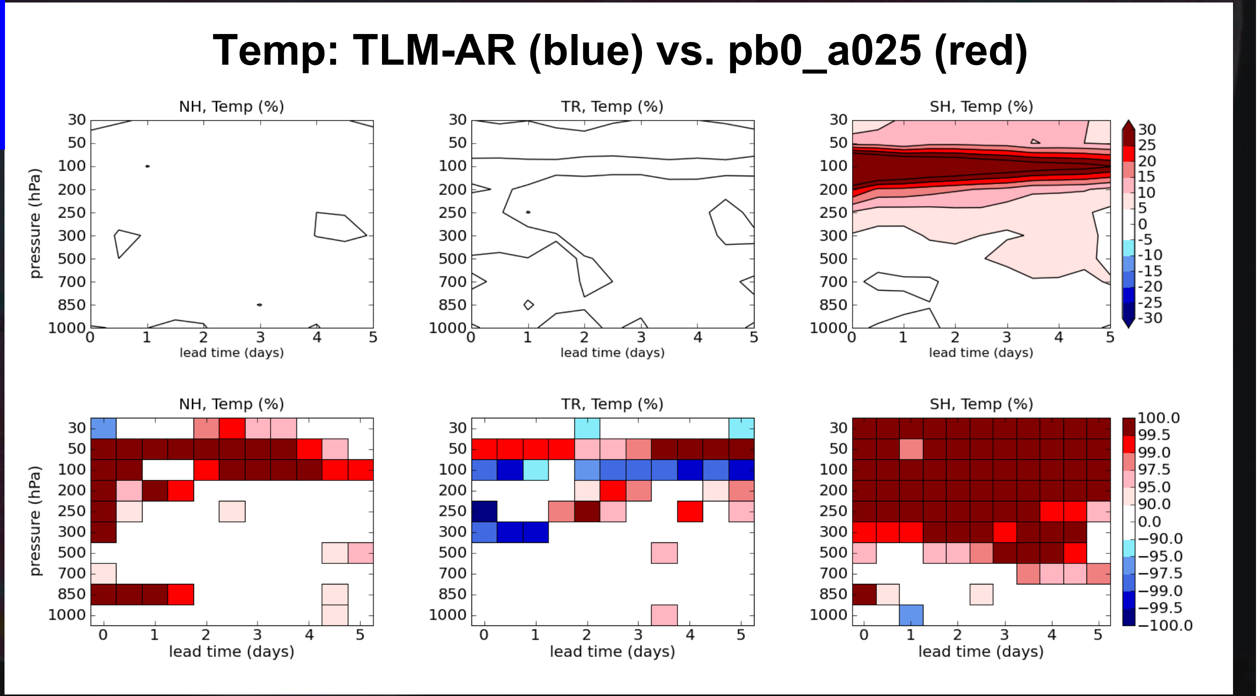




Standard Hybrid

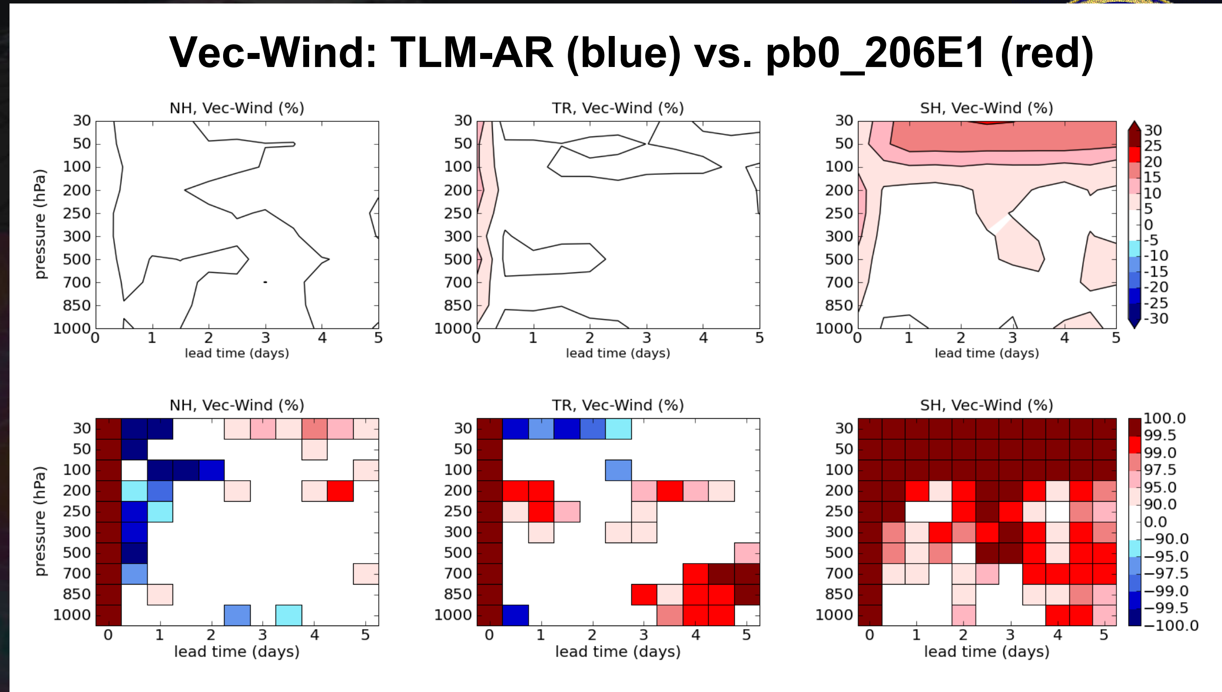
Regionally, the equation-Hybrid is better at avoiding areas of degradation than the standard-Hybrid. (tropics particularly)

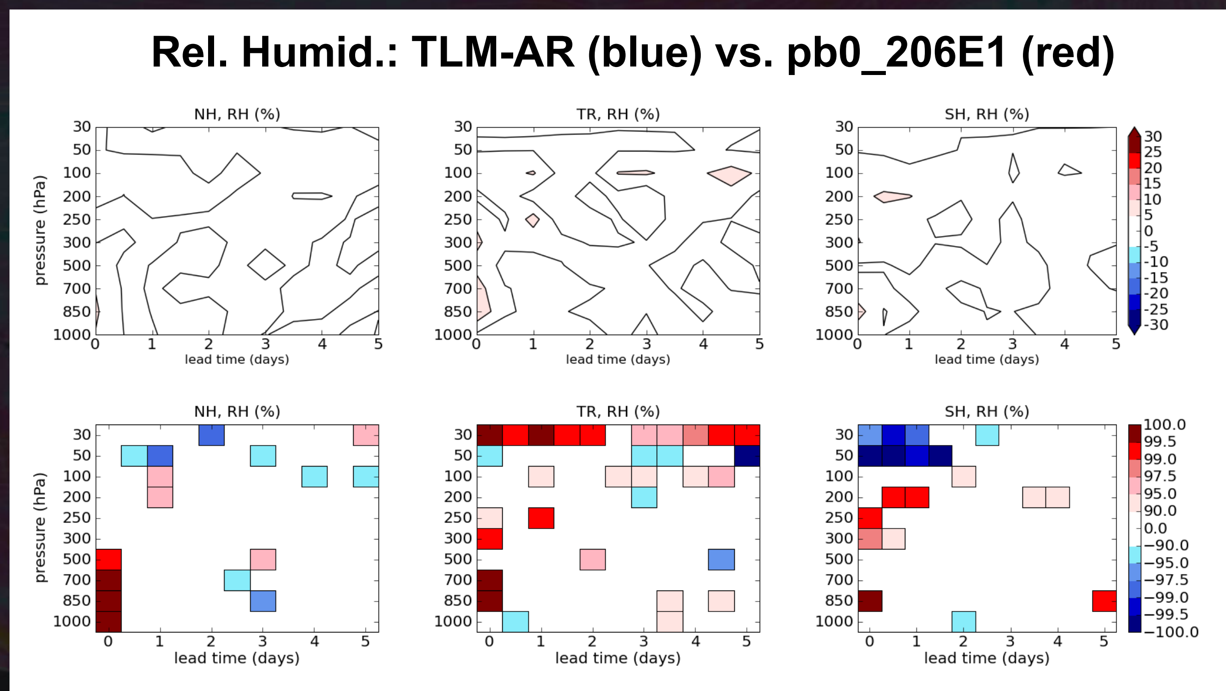




Equation Hybrid

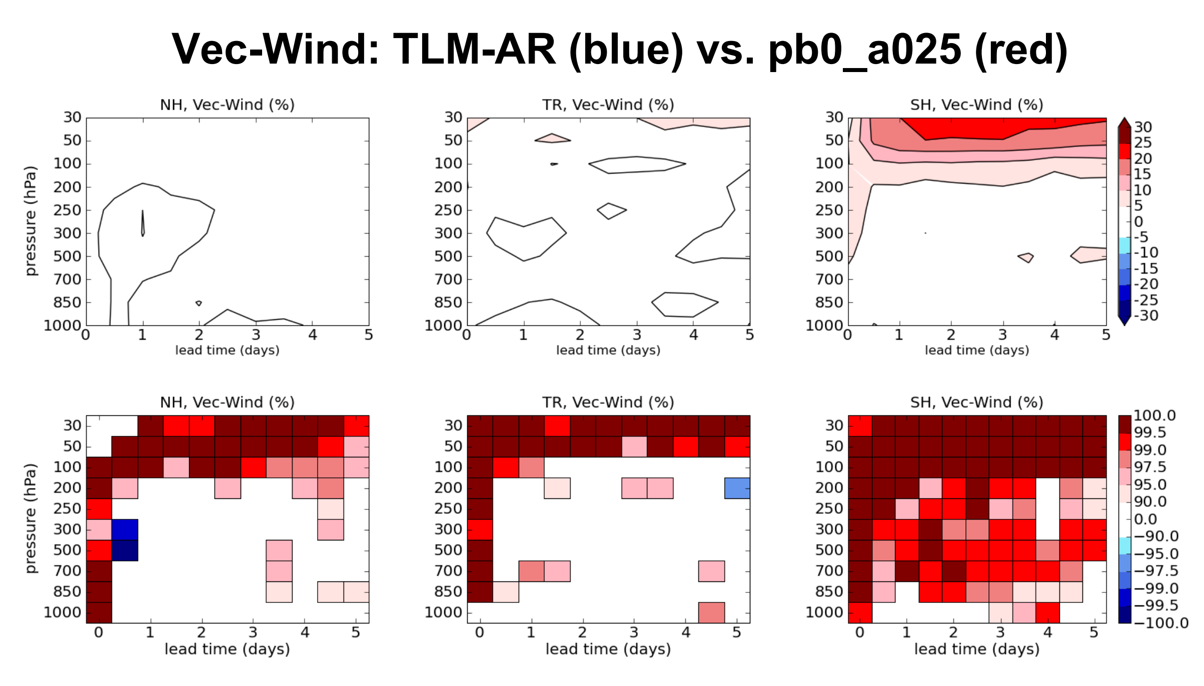
Regionally, the equation-Hybrid is better at avoiding areas of degradation than the standard-Hybrid. (tropics particularly)

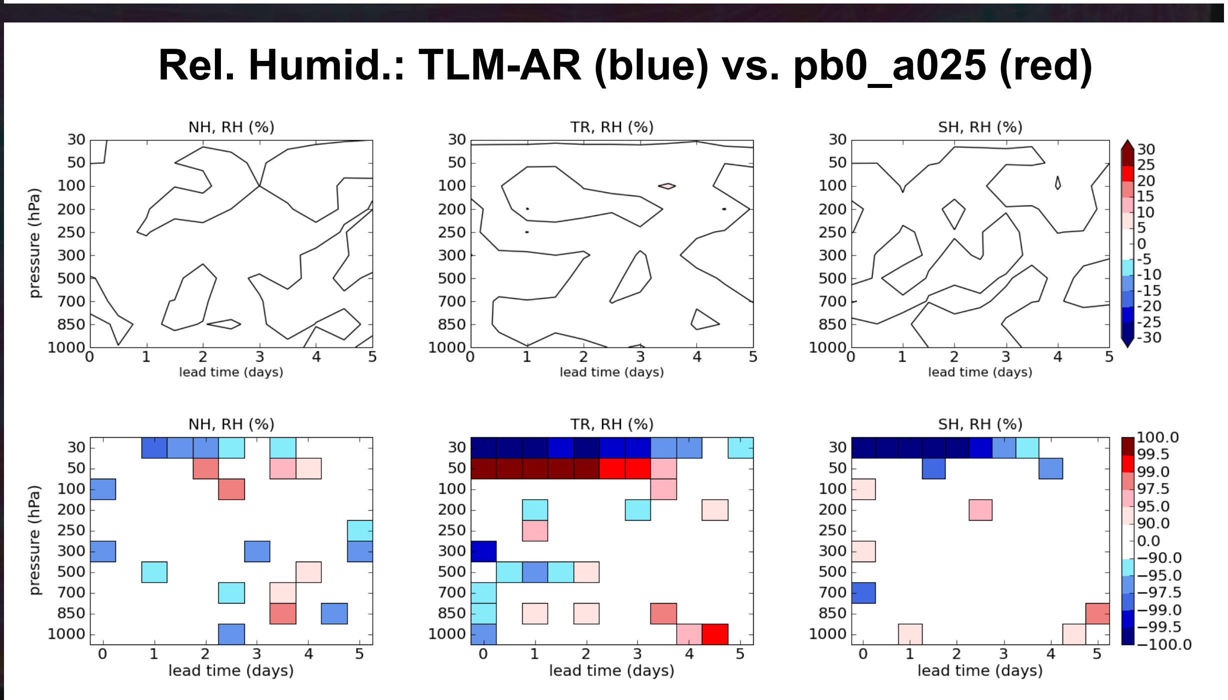




Standard Hybrid

Regionally, the equation-Hybrid is better at avoiding areas of degradation than the standard-Hybrid. (tropics particularly)





Summary

- All error variance predictions are inaccurate
- When two or more *independent* error variance predictions are available, linear combination is more accurate than either one individually – hybrid allows such combinations
- Brute force tuning of Hybrid weights is very expensive
- Analytical theoretical model for error variance prediction has been developed that allows weights to be derived directly from archive of (ensemble-variance, innovation) pairs.
- Theory estimates accuracy of static and flow-dependent error variance predictions – quantifies errors in error-variances.
- Low-resolution experiments using Navy forecast model indicate that (a) standard-Hybrid and equation-Hybrid superior to NAVDAS-AR (b) standard-Hybrid and equation-Hybrid about the same but weights for

Furthermore,

There once was a ground-hog called Grimweld
Who wanted to interpret Rumsfeld
He fell on the floor
Got shot at the screen
And now wise Grimweld lies pummeled

… thank you Grimweld