Development of the Local Ensemble Transform Kalman Filter

Istvan Szunyogh

Institute for Physical Science and Technology

& Department of Atmospheric and Oceanic Science

AOSC Special Seminar September 27, 2007

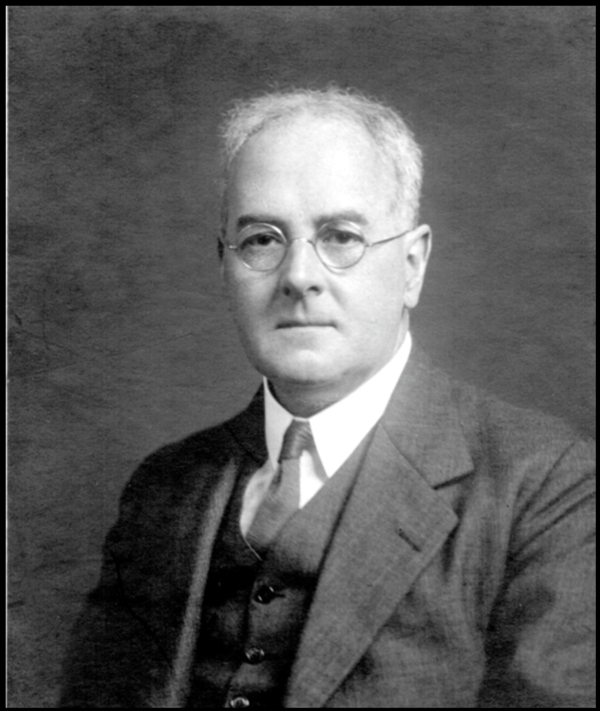
Outline

- The data assimilation problem
  - Illustrated with historical examples
- The LETKF story
  - How does it work?
- The future of the project

The First Man Who Faced the Challenge

was the first man who attempted to make a numerical weather forecast

Lewis Fry Richardson (1881-1953)



Weather Prediction by Numerical Process, 1922

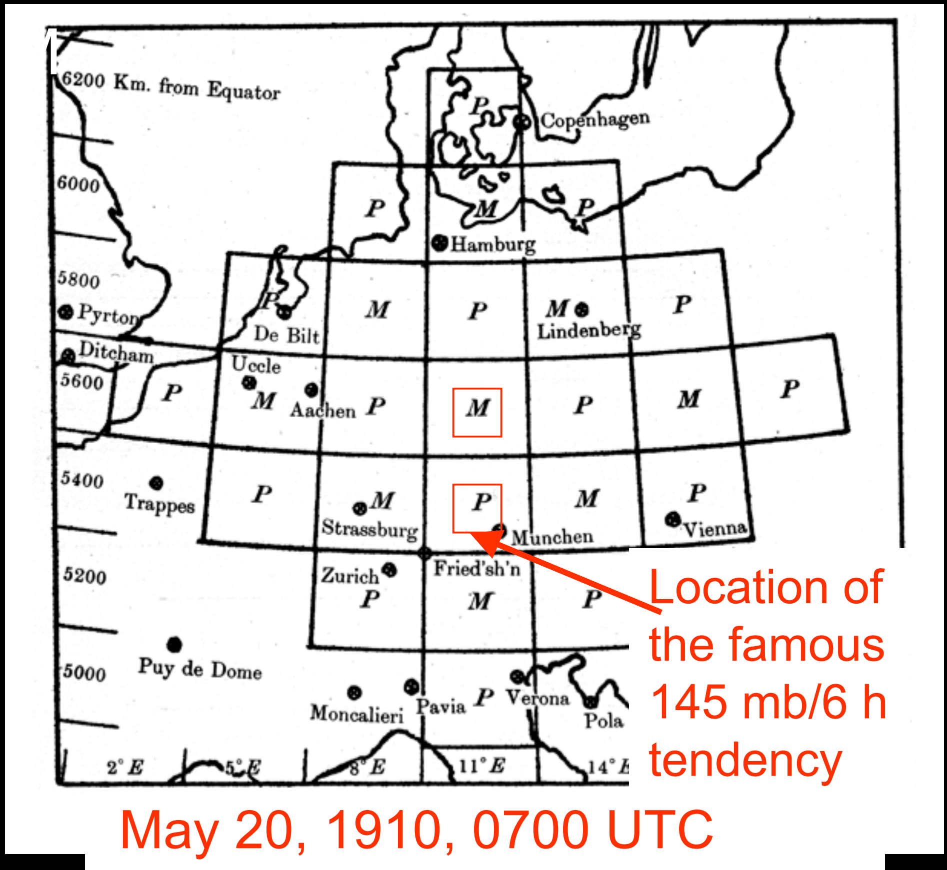
second edition was published in August 2007 by Cambridge

"Perhaps some day in the dim future it will be possible to advance the computation faster than the weather advances and at a cost less than the saving to mankind due to the information gained. But that is a dream"*-Richardson*

Did you know?

- That Richardson designed a decent model (except for proposing a time step that did not satisfy the CFL condition and a need for diffusion), which is capable to provide a decent 1-day forecast
- But, he failed (so famously!) even before calculating the first time step, because he did not have a decent analysis !
- For a modern interpretation of Richardson's work see Peter Lynch: The Emergence of Numerical Weather Prediction: Richardson's Dream, 2006, Cambridge

Challenge #1: The model variables are not observed directly



City names within grid: Observations

M: Momentum (wind) variables

P: Mass variables

"It makes one wish that pilot balloons stations could be arranged in rectangular order, alternating with stations for registering balloons…"-*Richardson*

A simple interpolation problem for Richardson, but remotely sensed observations took the difficulties to a whole new level

Challenge #2: There are many more model variables than assimilated observations

- **Richardson's calculation**
  - Figure shows the region of highest observational density!
  - Unusually large number of observations to detect the effects of the passing Haley comet on the atmosphere (regular upper air soundings started only after World War II)
- **Current global circulation model of NCEP***
  - Number of model variables: about 385 million
  - Number of assimilated observations: 7-8 million observations per day (about two orders of magnitude less then the # of variables)
  - Number of observations received: 1.43 billion observations per day not all assimilated due to (i) time constraint (total time available for data processing and analysis is 35 minutes), (ii) quality problems, (iii) lack of observation operator, (iv) redundancy

*Source: J. Derber's presentation at UMD workshop on satellite DA

Challenge #3: Internal consistency of the state estimate-”Balance”

- **Richardson**
  - He did not know about the importance of this issue
  - Although he thought that the root of his problem was the unrealistically large divergence in the analysis near the surface
- **Current NWP**
  - The issue is much broader than controlling gravity waves (e.g., spin-up in atmospheric water cycle)
  - The models are robust to many types of inconsistencies in the initial conditions (they survive, but forecast accuracy suffers)
  - Initialization of the analysis field (external filters) can also help
  - But, in principle, the data assimilation scheme is expected to do a good job (e.g., by making initialization part of the analysis process-internal filters)

The Background (First Guess)

- The analysis $\boldsymbol{x^a}$ is obtained by updating a background $\boldsymbol{x^b}$ based on the observational information:
  - $\mathbf{x^a=x^b+f(y^o)}$,
  - The components of the state vector, $\boldsymbol{x}$, are the model variables at the grid points and **f** is a function of the observations $\mathbf{y}^o$
  - This approach provides an estimate of all state variables
  - The background can be constructed to be well balanced and can propagate information from the past
  - For the first time in history, Bergthorsson and Döös in 1955 (Tellus) obtained $\boldsymbol{x^b}$ by linearly combining a short term model forecast with climatology and called it "preliminary"

The State-Of-The-Art Background: A Short-term Forecast

- In a modern data assimilation system x^b is a short-term model forecast from the analysis at the previous time.
  - It reflects the combined effect of all past observations, filling up gaps in the observing network
  - Model dynamics do the filtering and build realistic dynamical “balance” between the observed and unobserved variables

The Least-Square Problem

- Cost function: $J(\mathbf{x})=[\mathbf{x}-\mathbf{x}^b]^T(P^b)^{-1}[\mathbf{x}-\mathbf{x}^b]+[\mathbf{y}-h(\mathbf{x})]^T R^{-1}[\mathbf{y}-h(x)]$
  - $P^{b:}$ Background error covariance matrix
  - R: Observation error covariance matrix
  - h: Observation operator
- Essentially all data assimilation schemes are based on minimizing J(**x**), (variational schemes often have an extra penalty term)
- Observations **y** without the associated **R** and **h** are useless (typically, R is most problematic for retrievals, while h is most problematic for radiances)

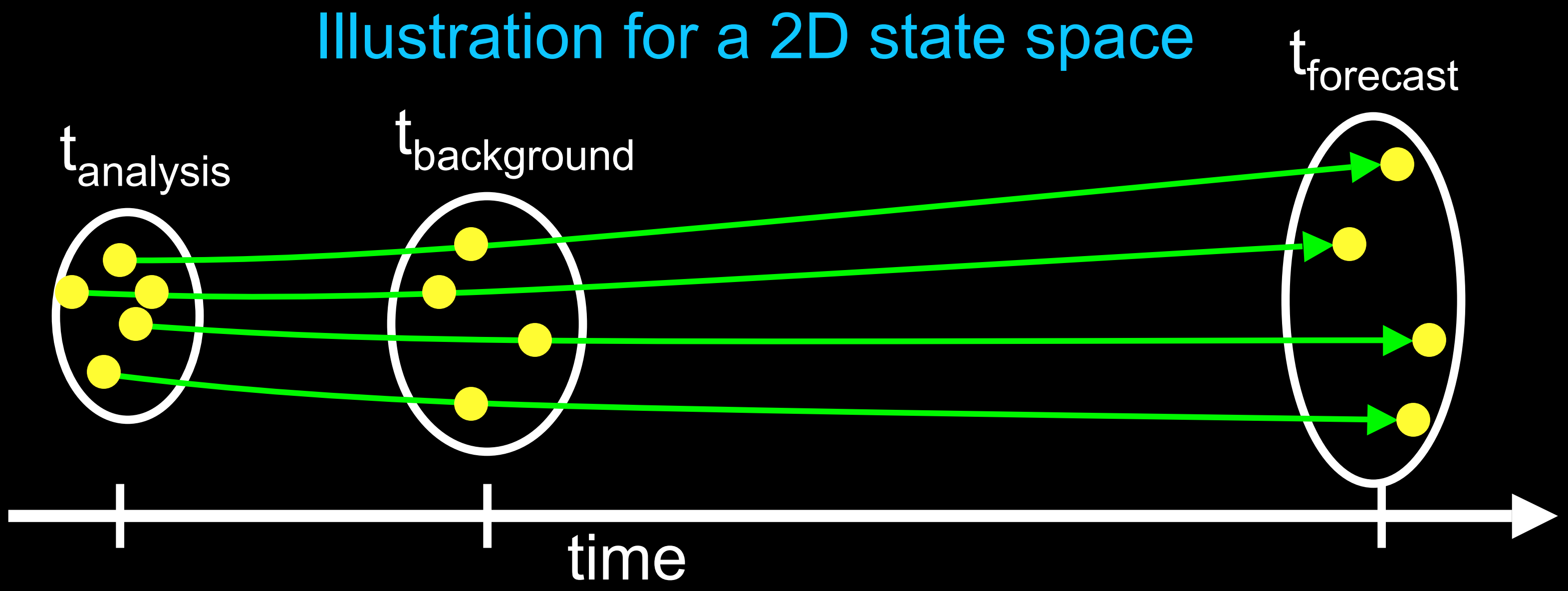
Extended Kalman Filter:

the four main components at time t_n

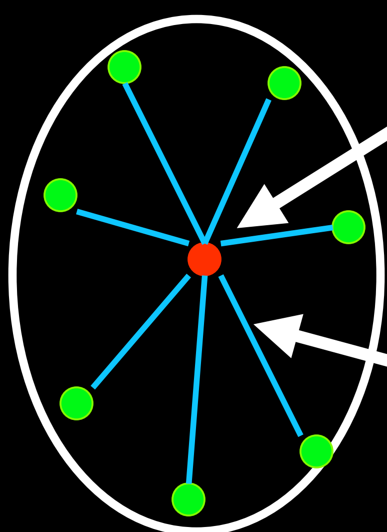
1. $\mathbf{x}^b = M\mathbf{x}^a(t_{n-1})$: Obtaining the backgound
   - $\mathcal{M}$: Nonlinear model from time t_{n-1} to t_n
2. $\mathbf{P}^b = M\mathbf{P}^a(t_{n-1})M^T$: Obtaining the background error covariance matrix
   - M: Linearization of $\mathcal{M}$ around $\mathbf{x}^a(t_{n-1})$
   - Prohibitively expensive computationally
   - Issues of linearization
3. $\mathbf{x^a} = \mathbf{x}^b + \mathbf{K}[\mathbf{h}(\mathbf{x^b}) - \mathbf{y}]$: Update Equation
   - $\mathbf{K} = \mathbf{P}^b\mathbf{H}^T(\mathbf{H}\mathbf{P}^b\mathbf{H}^T + \mathbf{R})^{-1}$: Kalman Gain Matrix
   - H: $\mathbf{h}(\mathbf{x_b})$ linearized around $\mathbf{x_b}$
4. $\mathbf{P}^a = (\mathbf{I} - \mathbf{KH})\mathbf{P}^b$: Analysis Error Covariance Matrix

Ensembles

- The model state is considered to be a probabilistic variable: The probability distribution is evolved by a representative ensemble of model states



Ensemble Representation of the Background



The ensemble mean is the background

The background error covariance matrix is defined by the ensemble of background perturbations

Background Ensemble

Ensemble-based Kalman Filter
data assimilation schemes

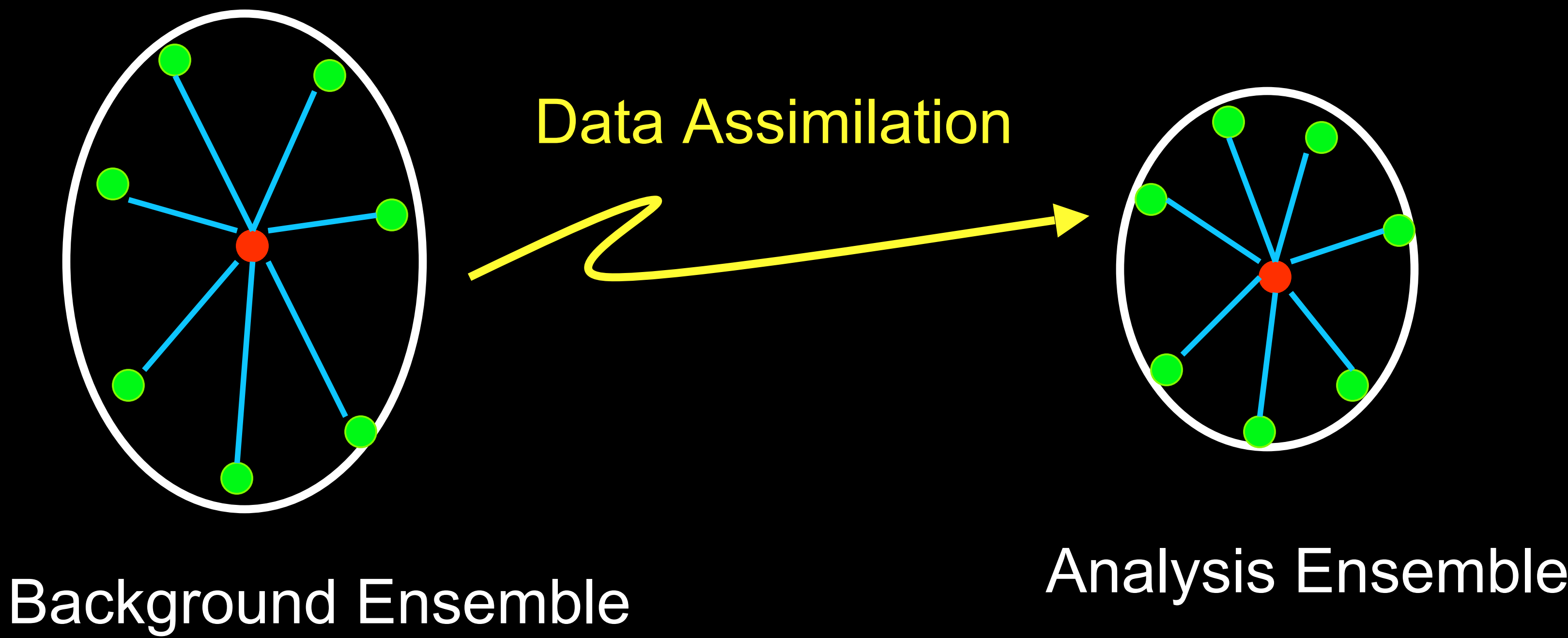
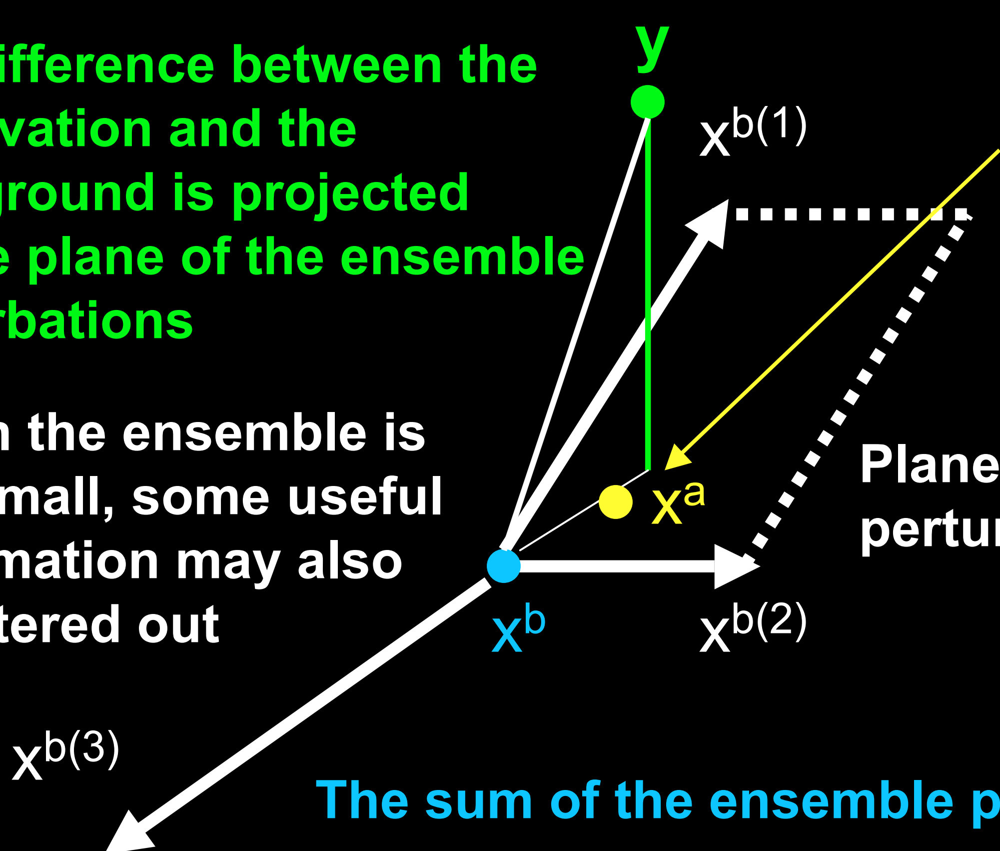


Illustration in State Space

3d state space, 3-member ensemble on a plane

The difference between the observation and the background is projected on the plane of the ensemble perturbations



x^b-x^a is obtained in the plane of the ensemble perturbations: potentially an efficient filter of observational noise

When the ensemble is too small, some useful information may also be filtered out

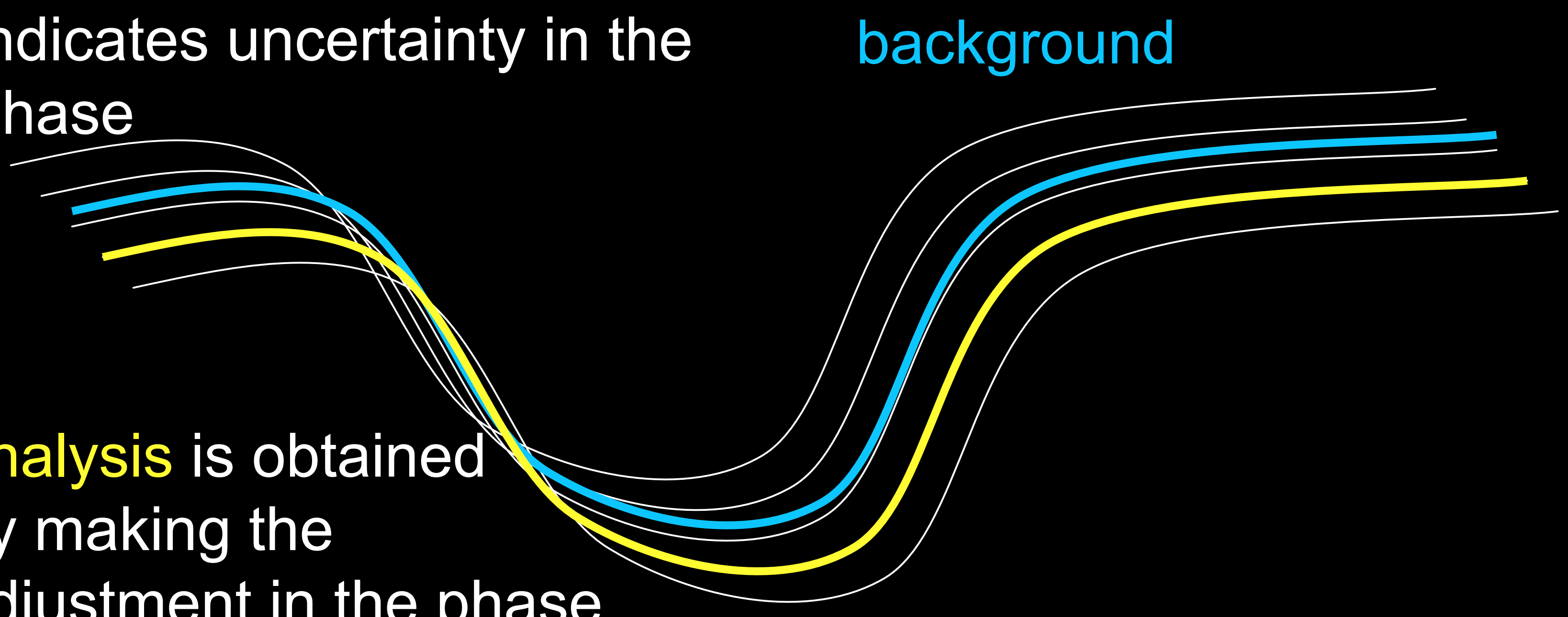
Plane of the ensemble perturbations

The sum of the ensemble perturbations is zero

Illustration in physical space

uncertainty in the phasing of a wave

background ensemble
indicates uncertainty in the
phase



analysis is obtained
by making the
adjustment in the phase
based on the observations

Generation of the Analysis Ensemble Perturbations

- Perturbed-observations method:
  - First proposed by Houtekamer and Mitchell (1998), Burgers et al. (1998)
  - Each of the k ensemble members is updated assimilating a set of randomly perturbed observations
  - It provides an analysis ensemble with the right P^a, when k goes to infinity
- Square-root filters:
  - Schemes proposed by Bishop et al. (2001), Anderson (2001), Whitaker and Hamill (2002), Ott et al. (2002), a nice paper on the subject is Tippett et al. (2003), LETKF
  - First calculates P^a, then generates a set of analysis perturbations that exactly satisfy that P^a
  - More accurate for smaller ensembles (better representation of R)

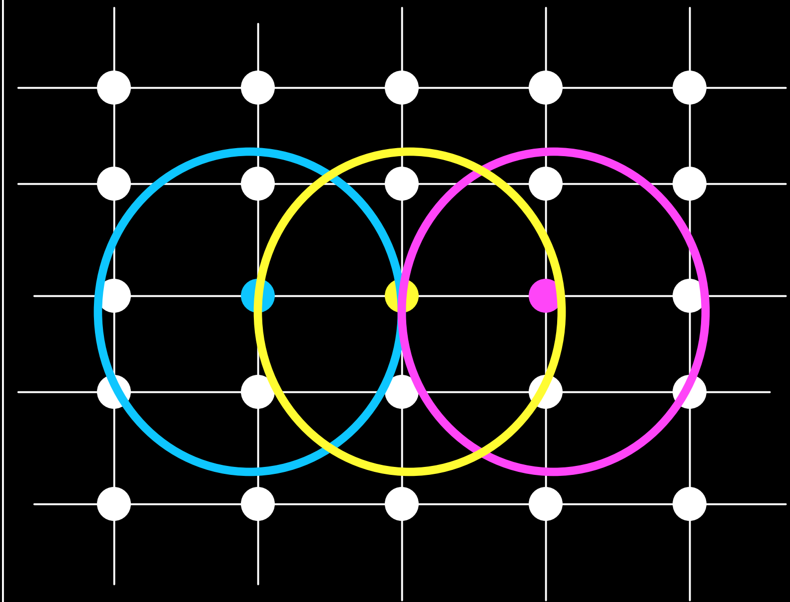
Optimal order of calculations?

- The analysis for the different state vector components can be processed independently
  - Sounds trivial, but nobody considered doing it in the context of an ensemble-based DA system before us
  - In part, because it is assumed to be computationally suboptimal (even in such recent books as Evensen, 2006: Data assimilation: The Ensemble Kalman Filter, Springer)
  - In reality, for a high-resolution model and a large number of observations, this is the most efficient approach on a parallel computer
- Observations can be assimilated serially or simultaneously
  - In a serial scheme, he observations are assimilated one by one iteratively updating the backgound and the background error matrix
  - When the number of observations is large, the serial approach is computationally more expensive (Whitaker 2007).

UMCP Weather & Chaos Project
http://weatherchaos.umd.edu

- Started in 2000 by J. Yorke and E. Kalnay with the aim
  - To develop a data assimilation system for spatio-temporally chaotic systems
  - To study predictability in spatiotemporally chaotic systems
  - I was hired to lead the project in 2001
- Main achievement: The project produced specific science problems that led to
  - 10 Ph.D. thesis in four different programs (AOSC, AMSC, Physics, EE)
  - 2 more are expected by the end of the calendar year
  - There are several others in progress
- Unique feature of our approach: Local in grid space

Illustration of the Local Approach
for a 2D model grid



- A local region is associated with each grid point
- Properties assigned to a grid point are calculated using information from the associated local region
- For instance, the analysis for a given grid point is calculated using $\boldsymbol{x}^a$, $\boldsymbol{x}^b$, $\boldsymbol{K}$, $\boldsymbol{y}$, $\boldsymbol{B}$, and $\boldsymbol{R}$ defined for the local region

Motivations for our approach of the development were

- In 2001, it was yet to be seen whether an ensemble-based Kalman filter coupled with a state-of-the-art forecast model can be used to assimilate observations of the real atmosphere. The major concerns were
  - An estimate of the background error covariance matrix based on a reasonably small ensemble would be hopelessly rank-deficient
  - An ensemble-based Kalman filter would be computationally hopelessly expensive
  - Some scientists also argued that model errors were hopelessly large for an indefinitely long cycling of an ensemble base Kalman filter
- Our goal was to design a scheme to address these concerns and a series of experiments to separate real challenges from assumed difficulties
- We wanted to design a scheme for parallel computers

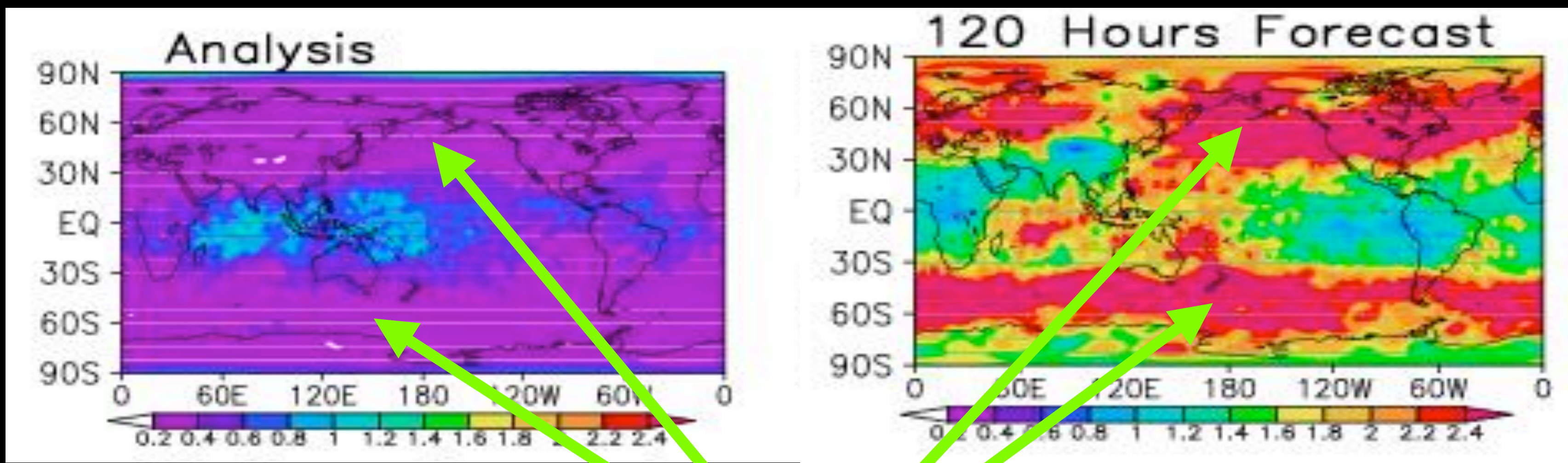
Local Ensemble Kalman Filter (LEKF)

- First Formulation: Ott, Hunt, Szunyogh et al. 2004, *Tellus* **A**
  - Introduced the idea of localization in grid space
  - Introduced the idea of preparing the analysis independently for the different grid points
  - Investigated the conditions under which the local approach provided a smooth global analysis
  - Scheme was tested on the Lorenz-96 model (40-120 variables)
- First experiments with the NCEP GFS were designed to address the following issue
  - Is it possible to track the state of the model with a small (40-80-member) ensemble under the perfect model scenario?
  - Results were reported in Szunyogh, Kostelich, Gyarmati et al., 2005, ***Tellus A***

Experimental design

- Observations: Noisy observations of a time series of true states (generated by a long model integration), full vertical soundings are located at randomly selected model grid point location (10% coverage)
- Data Assimilation: LETKF with 40 ensemble members
- Model: NCEP GFS at resolution T62 (about 150 km) and 28-levels
- Error Statistic collected for 45 days (January-February)

Geographical Distribution of Errors



The analysis errors are the smallest where the forecast errors grow fastest (For a detailed investigation of the analysis errors see IS et al., 2005, *Tellus;* of the forecast errors see Kuhl et al., 2007, *JAS.*)

Main Conclusions of the Study

- The key is to find a good balance between the number of ensemble members and the size of the local region (larger region requires a larger ensemble)
  - A 40-member ensemble with 5x5x3 grid points is about as accurate as an 8-member ensemble with 7x7x3 grid points, but computationally more efficient
  - 3x3x3 local cubes are always suboptimal (too few observations in local cubes)
- Where the 6-hour error growth is fast (storm track regions) the analysis is extremely accurate, because the background ensemble is very efficient in capturing the space of uncertainties

Ensemble DA Comparison Project
funded by NOAA THORPEX, 2003-2007

- 4 groups were asked to develop ensemble-based DA systems for the NCEP GFS model
- Was the motivation to develop the LETKF (Hunt, Kostelich, Szunyogh, 2007: Physica D) from the LEKF
- Two groups succeeded: UMD and ESRL/NOAA (Jeff Whitaker and Tom Hamill), UMD team has a paper in press in Tellus, ESRL team has a paper in press in MWR
- As we hoped, LETKF is the computationally most efficient scheme
- A consensus system is being implemented at NCEP, for further testing, based on the LETKF by Jeff Whitaker

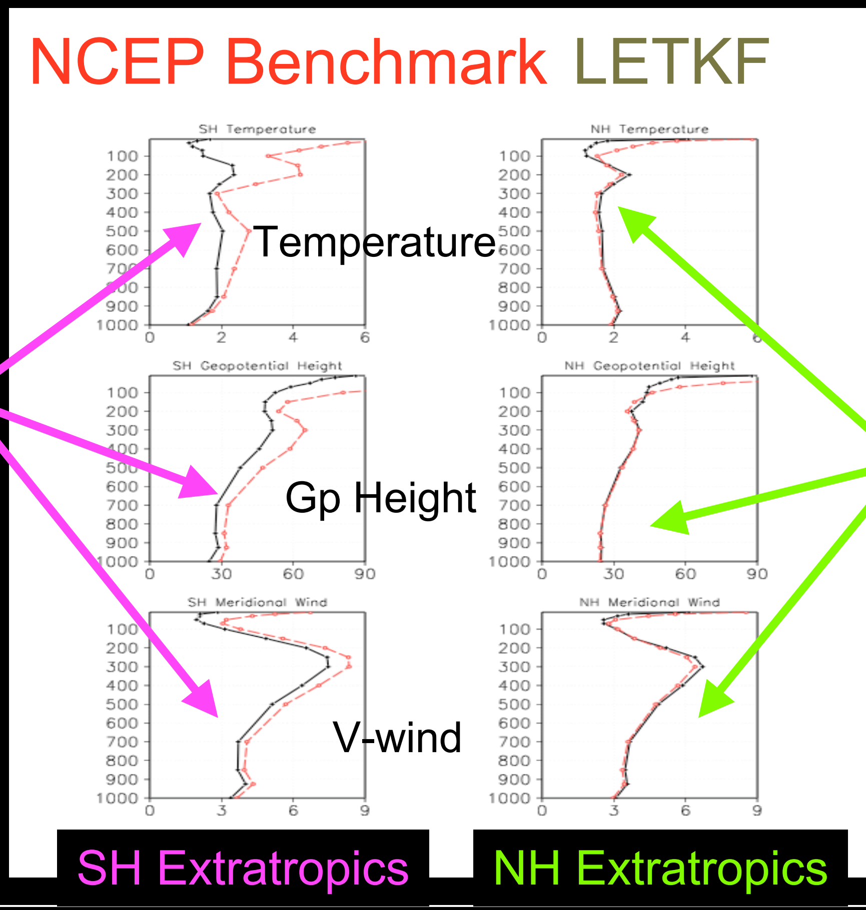
Validation Experiments with the NCEP GFS at resolution T62L28-reanalysis resolution

- Observations of the real atmosphere, except for radiances (Szunyogh, Kostelich, Gyarmati et al. 2007, Tellus, in press)
  - The LETKF and the Benchmark SSI system use different **H** operators; the one used with the LETKF is less sophisticated. This may affect the results near the surface and in areas of high observational density
  - Benchmark SSI data are provided by NCEP (Y. Song and Z. Toth)
  - 60-member ensemble

Comparison of the LETKF and the SSI

48-hour forecasts with real observations (no radiances)

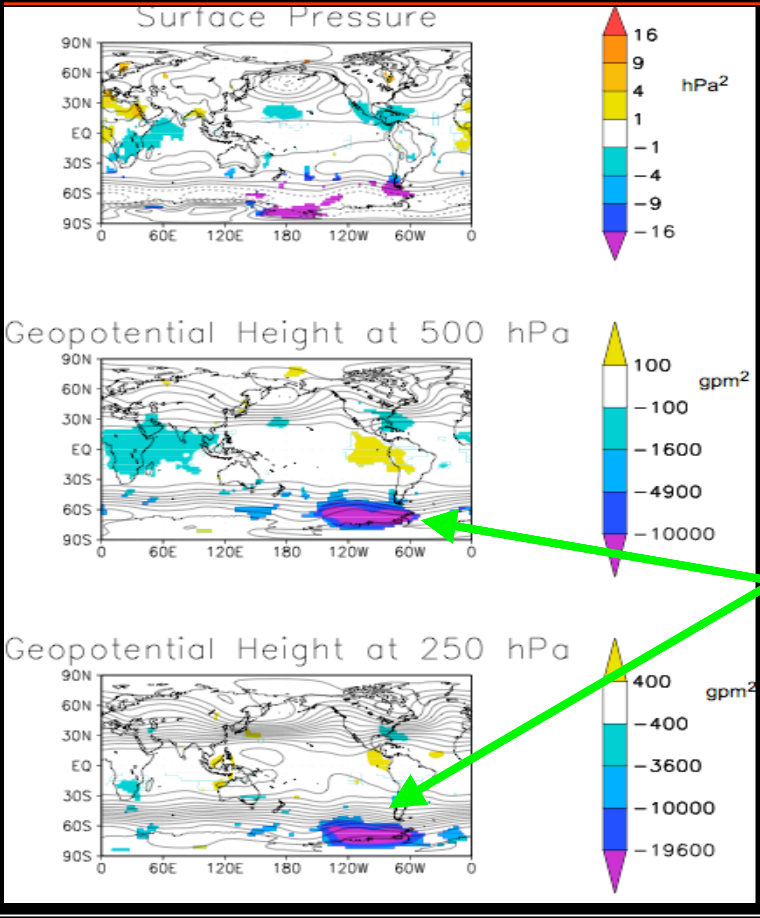
In the SH XT
The LETKF
is more
accurate



In the NH XT
the two systems
are comparable

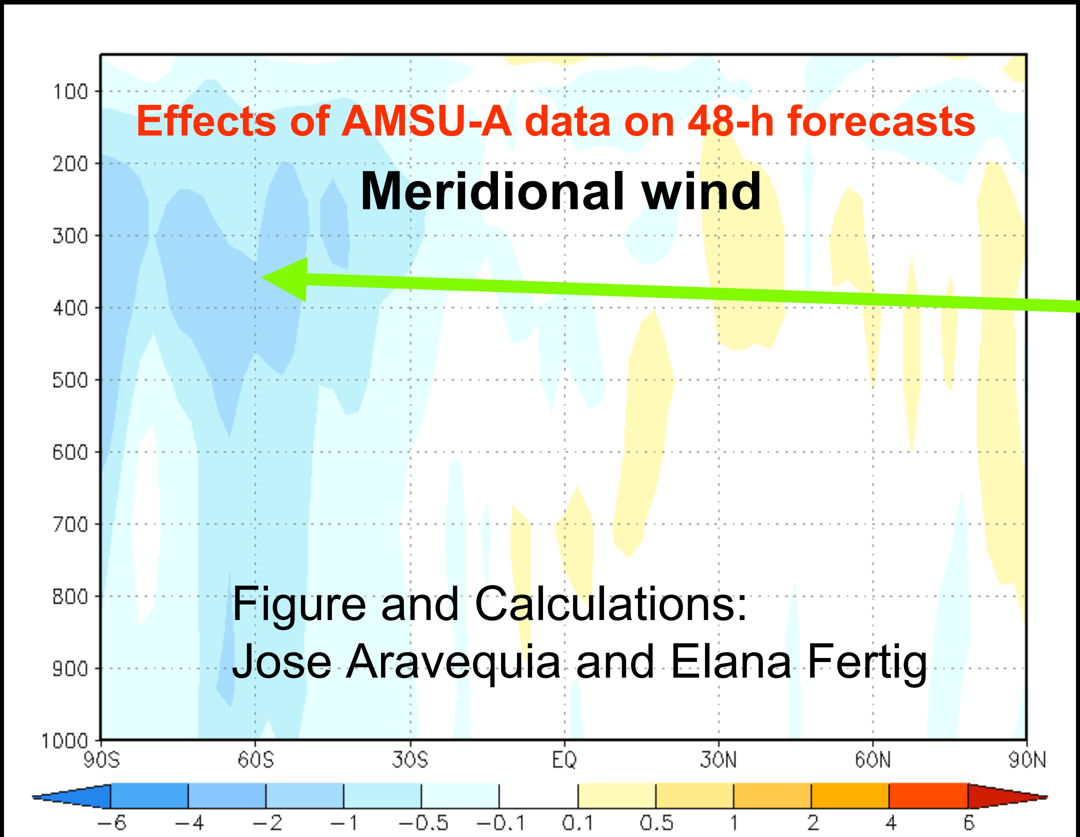
Comparison of the LETKF and the SSI

48-hour forecasts with real observations (no radiances)



The advantage of the LETKF is the largest where the observation density is the lowest

Latest results: capability to assimilate satellite radiances



- The large improvements in the SH suggests, that there is a lot of useful information in the estimated background error covariance matrix between the temperature (most closely related to the radiances) and the wind

The Goal
is to convince others that they should use our code and/or algorithm

- Those who use our code
  - CPTEC Brazil is in the process of implementing in operations
  - Atmospheric and Environmental Research Inc. (ocean DA for Navy, Phase 2 starts in October)
  - University of Massachusetts-Dartmouth (ocean)
  - ECMWF expressed interest for research-depends on availability of funding
  - UCLA/JPL proposal to couple it with ROMS
- Those who use our algorithm
  - Japan Meteorology Agency (See talk by Takemasa Miyoshi)
  - Jeff Whitaker (effort on NCEP computer)

The Future
has already started

- Further investigation of predictability with the LETKF/GFS system: 3-year NSF funded project started in August, Liz Satterfield)
- Martian Data Assimilation (2 NASA funded project will start in October--the goal is to couple the GFDL Mars model (also a community model) and the LETKF, to study predictability in the Martian atmosphere, and to carry out a reanalysis of Martian observations, at least 2 new GRAs
- Impact of wildfire emission (1 NASA funded project, Dave Kuhl)
- Carbon cyclone data assimilation (4-year DOE funded project led by Eugenia)

Reminder:

http://weatherchaos.umd.edu

- Information available through the web page
  - Papers
  - Information and presentations from the Summer Workshop on Satellite DA
  - Presentations from AOSC615
- Most complete review paper available from the web
  - Szunyogh et al., 2007: The Local Ensemble Transform Kalman Filter and its implementation on the NCEP global model at the University of Maryland. ECMWF proceedings, in press.