The Ensemble of Data Assimilations

Massimo Bonavita

massimo.bonavita@ecmwf.int

Contributions from:
Lars Isaksen, Elias Holm, Mike Fisher,
Laure Raynaud

ECMWF

Outline

- KF, EKF, Reduced-rank KF, EnKF
- Hybrid methods
- The Ensemble of Data Assimilations (EDA) method
- Applications of the EDA

The EDA method

For a linear system the data assimilation update is:

$$\mathbf{x}_a^k = \mathbf{x}_b^k + \mathbf{K}_k\left(\mathbf{y}^k - \mathbf{H}_k\mathbf{x}_b^k\right)$$
$$\mathbf{x}_b^{k+1} = \mathbf{M}_k\mathbf{x}_a^k$$

Under the assumptions of statistically independent background (**P**^{b}), observation (**R**) and model errors (**Q**), the evolution of the system error covariances is given by:

$$\mathbf{P}_k^a = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\mathbf{P}_k^b\left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T$$
$$\mathbf{P}_{k+1}^b = \mathbf{M}_k\mathbf{P}_k^a\mathbf{M}_k^T + \mathbf{Q}_k$$

The EDA method

Consider now the evolution of the same system where we perturb the observations and the forecast model with random noise drawn from the respective error covariances:

$$\tilde{\mathbf{x}}_a^k = \tilde{\mathbf{x}}_b^k + \mathbf{K}_k\left(\mathbf{y}^k + \boldsymbol{\eta}_k - \mathbf{H}_k\tilde{\mathbf{x}}_b^k\right)$$
$$\tilde{\mathbf{x}}_b^{k+1} = \mathbf{M}_k\tilde{\mathbf{x}}_a^k + \boldsymbol{\zeta}_k$$

where $\boldsymbol{\eta}_k \sim \mathcal{N}(\mathbf{0},\mathbf{R})$, $\boldsymbol{\zeta}_k \sim \mathcal{N}(\mathbf{0},\mathbf{Q})$.

If we define the differences between the perturbed and unperturbed state $\boldsymbol{\varepsilon}_a \equiv \tilde{\mathbf{x}}_a - \mathbf{x}_a$ and $\boldsymbol{\varepsilon}_b \equiv \tilde{\mathbf{x}}_b - \mathbf{x}_b$, their evolution is obtained by subtracting the unperturbed state evolution equations from the perturbed ones:

$$\boldsymbol{\varepsilon}_a^k = \boldsymbol{\varepsilon}_b^k + \mathbf{K}_k\left(\boldsymbol{\eta}_k - \mathbf{H}_k\boldsymbol{\varepsilon}_b^k\right)$$
$$\boldsymbol{\varepsilon}_b^{k+1} = \mathbf{M}_k\boldsymbol{\varepsilon}_a^k + \boldsymbol{\zeta}_k$$

ECMWF

The EDA method

$$\boldsymbol{\varepsilon}_a^k = \boldsymbol{\varepsilon}_b^k + \mathbf{K}_k\left(\boldsymbol{\eta}_k - \mathbf{H}_k\boldsymbol{\varepsilon}_b^k\right)$$

$$\boldsymbol{\varepsilon}_b^{k+1} = \mathbf{M}_k\boldsymbol{\varepsilon}_a^k + \boldsymbol{\zeta}_k$$

- i.e., the perturbations evolve with the same update equations of the state (Kalman gain and model operator).

What about the errors?

If we take the statistical expectation of the outer product of the perturbations:

$$\left\langle \boldsymbol{\varepsilon}_k^a\left(\boldsymbol{\varepsilon}_k^a\right)^T\right\rangle = \left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)\left\langle \boldsymbol{\varepsilon}_k^b\left(\boldsymbol{\varepsilon}_k^b\right)^T\right\rangle\left(\mathbf{I} - \mathbf{K}_k\mathbf{H}_k\right)^T + \mathbf{K}_k\mathbf{R}_k\mathbf{K}_k^T$$

$$\left\langle \boldsymbol{\varepsilon}_{k+1}^b\left(\boldsymbol{\varepsilon}_{k+1}^b\right)^T\right\rangle = \mathbf{M}_k\left\langle \boldsymbol{\varepsilon}_k^a\left(\boldsymbol{\varepsilon}_k^a\right)^T\right\rangle\mathbf{M}_k^T + \mathbf{Q}_k$$

The EDA method

$$\left\langle \boldsymbol{\varepsilon}_k^a \left(\boldsymbol{\varepsilon}_k^a\right)^T \right\rangle = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right) \left\langle \boldsymbol{\varepsilon}_k^b \left(\boldsymbol{\varepsilon}_k^b\right)^T \right\rangle \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T$$

$$\left\langle \boldsymbol{\varepsilon}_{k+1}^b \left(\boldsymbol{\varepsilon}_{k+1}^b\right)^T \right\rangle = \mathbf{M}_k \left\langle \boldsymbol{\varepsilon}_k^a \left(\boldsymbol{\varepsilon}_k^a\right)^T \right\rangle \mathbf{M}_k^T + \mathbf{Q}_k$$

- These are the same equations for the evolution of the system error covariances:

$$\mathbf{P}_k^a = \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right) \mathbf{P}_k^b \left(\mathbf{I} - \mathbf{K}_k \mathbf{H}_k\right)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T$$

$$\mathbf{P}_{k+1}^b = \mathbf{M}_k \mathbf{P}_k^a \mathbf{M}_k^T + \mathbf{Q}_k$$

provided that:

1. The applied perturbations $\boldsymbol{\eta}_k$, $\boldsymbol{\zeta}_k$ have the right covariances (**R**, **Q**);
2. At some stage in time $\left\langle \boldsymbol{\varepsilon}_k^b \left(\boldsymbol{\varepsilon}_k^b\right)^T \right\rangle = \mathbf{P}_k^b$

The EDA method

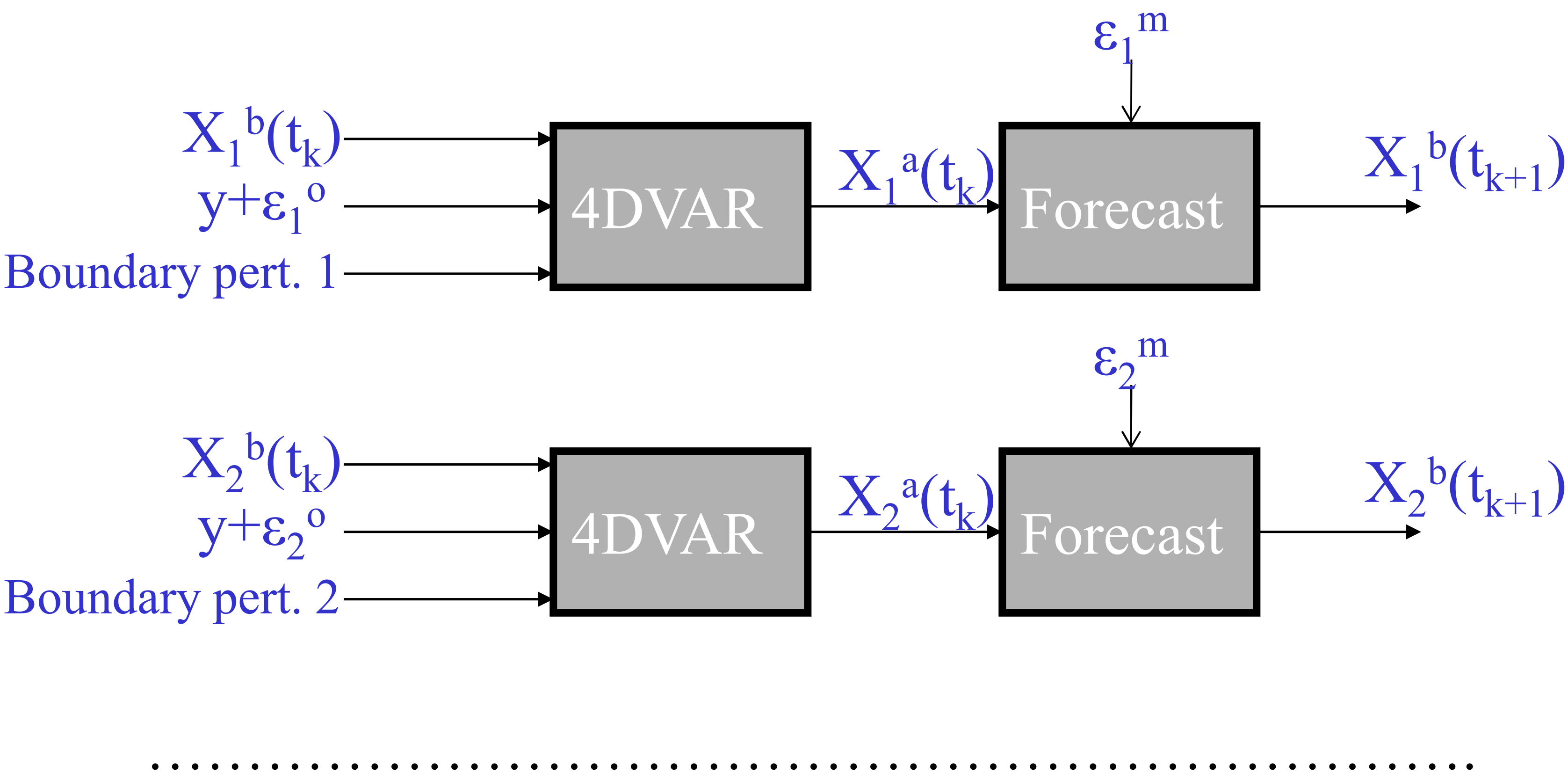
What does all this mean in practice?

- We can use an ensemble of perturbed assimilation cycles to simulate the errors of our reference assimilation cycle;
- The ensemble of perturbed DAs should be as similar as possible to the reference DA (i.e., same or similar **K** matrix)
- The applied perturbations $\boldsymbol{\eta}_k$, $\boldsymbol{\zeta}_k$ must have the required error covariances (**R**, **Q**);
- There is no need to explicitly perturb the background $\mathbf{x}_b$

The EDA method

- 25 ensemble members using 4D-Var assimilations at reduced resolution
- **T399** outer loop, T95/T159 inner loops. (Reference DA: T1279 outer loop, T159/T255/T255 inner loops)
- Observations randomly perturbed according to their estimated errors
- SST perturbed with climatological error structures
- Model error represented by stochastic methods (**SPPT,** Leutbecher, 2009)

The EDA method



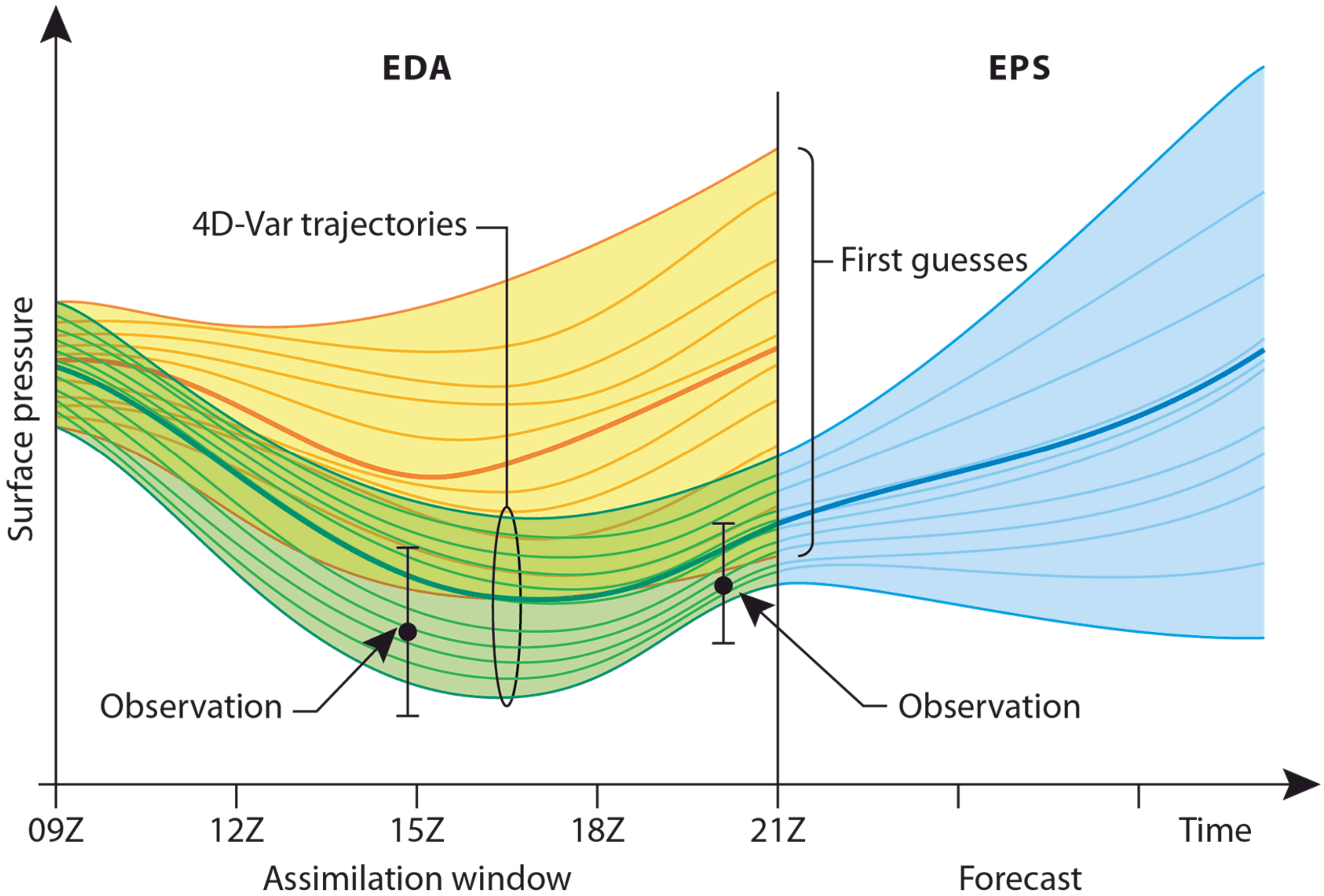
Applications of the EDA

- **The EDA simulates the error evolution of the 4DVar analysis cycle**. As such it has two main applications:

    1. Provide a flow-dependent estimate of analysis errors to initialize the ensemble prediction system (EPS)

    2. Provide a flow-dependent estimate of background errors for use in 4D-Var assimilation

Applications of the EDA



Applications of the EDA

- **Note that the EDA impacts the EPS in two distinct ways**:
  1. Directly, by providing a flow-dependent estimate of the initial perturbations of the ensemble prediction system (EPS)
  2. Indirectly, through the provision of a flow-dependent estimate of background errors for use in the 4D-Var deterministic (HRES) assimilation cycle: this improves the deterministic analysis around which the EPS initial perturbations are re-centered

ECMWF

Applications of the EDA

1. Directly, by providing a flow-dependent estimate of the initial perturbations of the ensemble prediction system (EPS)
2. Indirectly, through the provision of a flow-dependent estimate of background errors for use in the 4D-Var deterministic (HRES) assimilation cycle: this improves the deterministic analysis around which the EPS initial perturbations are re-centered

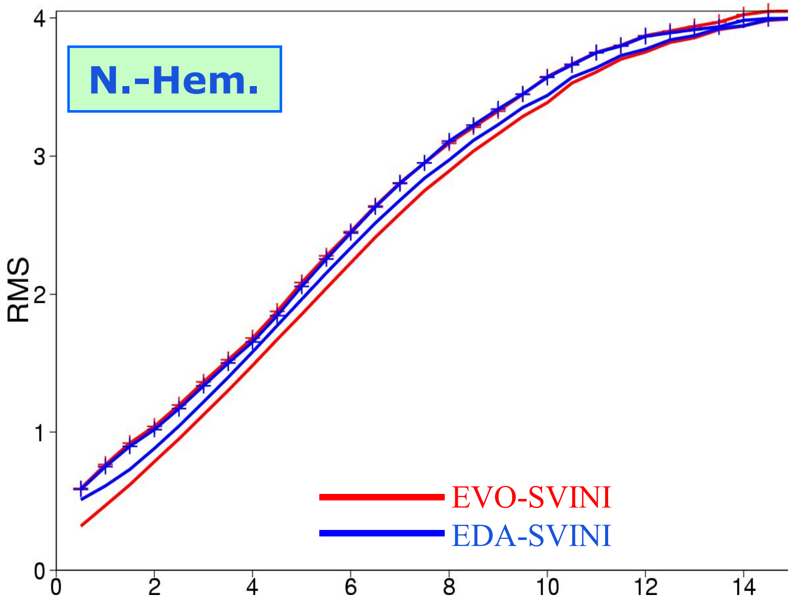
The second effect is arguably the more important:

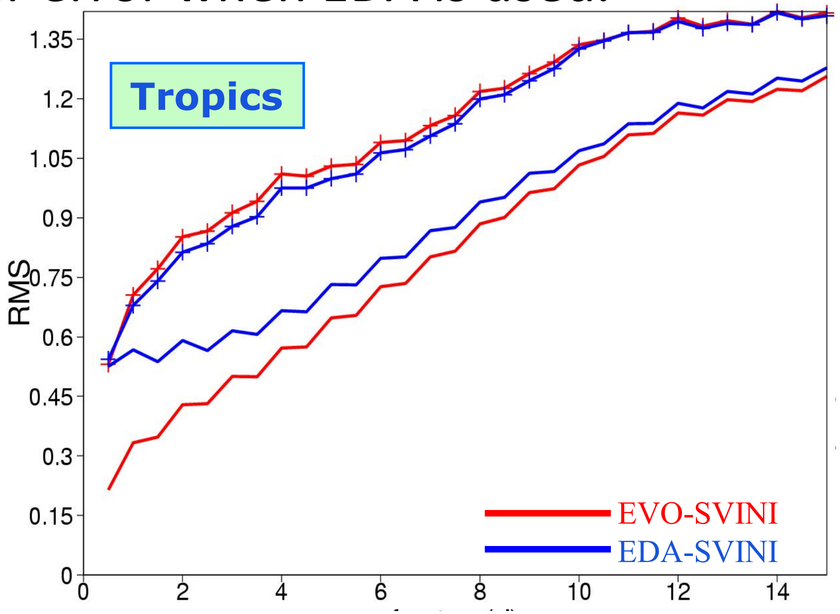
"… the main reason for the better performance of the EC-EPS in terms of the RMS, PAC, ROC and Brier Skill Score measures is the superiority of the ECMWF data assimilation (and perhaps numerical forecast modelling) system, and not necessarily the strategy used to simulate initial value and model related uncertainties in the EC-EPS." Buizza *et.al.*, 2005

Applications of the EDA

Improving Ensemble Prediction System by including EDA perturbations for initial uncertainty (implemented June 2010)

The Ensemble Prediction System (EPS) benefits from using EDA based perturbations. Replacing evolved singular vector perturbations by EDA based perturbations improve EPS spread, especially in the tropics.
The Ensemble Mean has slightly lower error when EDA is used.





Ensemble spread and Ensemble mean RMSE for 850hPa T

Applications of the EDA

- The EDA simulates **the error evolution** of the 4DVar analysis cycle. As such it has two main applications:
  1. Provide a flow-dependent estimate of analysis errors to initialize the ensemble prediction system (EPS)
  2. Provide a flow-dependent estimate of background errors for use in 4D-Var assimilation

Hybrids: EDA

In the ECMWF 4D-Var, the $\mathbf{B}$ matrix is defined implicitly in terms of a transformation from the background departure ($\mathbf{x}-\mathbf{x}_b$) to a control variable χ:

$$(\mathbf{x}-\mathbf{x}_b) = \mathbf{L}\chi$$

So that the implied $\mathbf{B}=\mathbf{L}\mathbf{L}^T$.

In the current wavelet formulation (Fisher, 2003), the variable transform can be written as:

$$(\mathbf{x}-\mathbf{x}_b) = \mathbf{T}^{-1}\mathbf{\Sigma}_b^{1/2}\sum_j \psi_j \otimes \left[\mathbf{C}_j^{1/2}(\lambda,\phi)\chi_j\right]$$

$\mathbf{T}$ is the balance operator, i.e. the operator that links the control variables to the model variables

$\mathbf{\Sigma}_b$ is the gridpoint variance of background errors

$\mathbf{C}_j(\lambda,\varphi)$ is the vertical covariance matrix for wavelet index j

ψ_j are the set of radial basis function that define the wavelet transform.

Hybrids: EDA

$$\left(\mathbf{x}-\mathbf{x}_b\right)=\mathbf{T}^{-1}\boldsymbol{\Sigma}_b^{1/2}\sum_j \psi_j \otimes\left[\mathbf{C}_j^{1/2}(\lambda,\phi)\chi_j\right]$$

$\mathbf{C}_j(\lambda,\varphi)$ are full vertical covariance matrices, function of (λ,φ). They determine both the horizontal and vertical background error *correlation structures*;

In standard 4D-Var $\mathbf{T}$ and $\mathbf{C}_j$ are computed off-line using a climatology of EDA perturbations.

$\boldsymbol{\Sigma}_b$ is computed by random sampling of the static $\mathbf{B}$ matrix (randomization procedure, Fisher and Courtier, 1995)

How do we make this error covariance model flow-dependent?

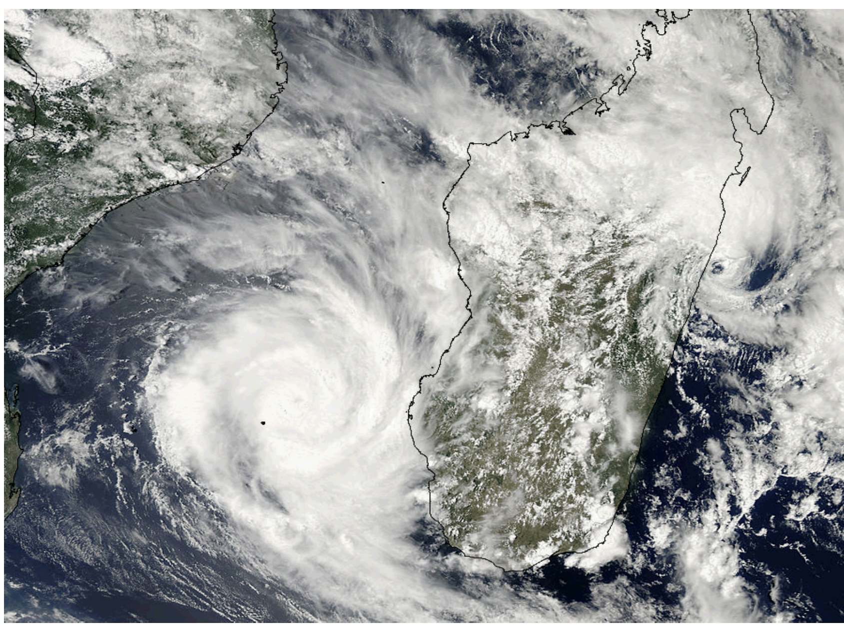
We look for flow-dependent EDA estimates of $\boldsymbol{\Sigma}_b$ and $\mathbf{C}_j(\lambda,\varphi)$

ECMWF

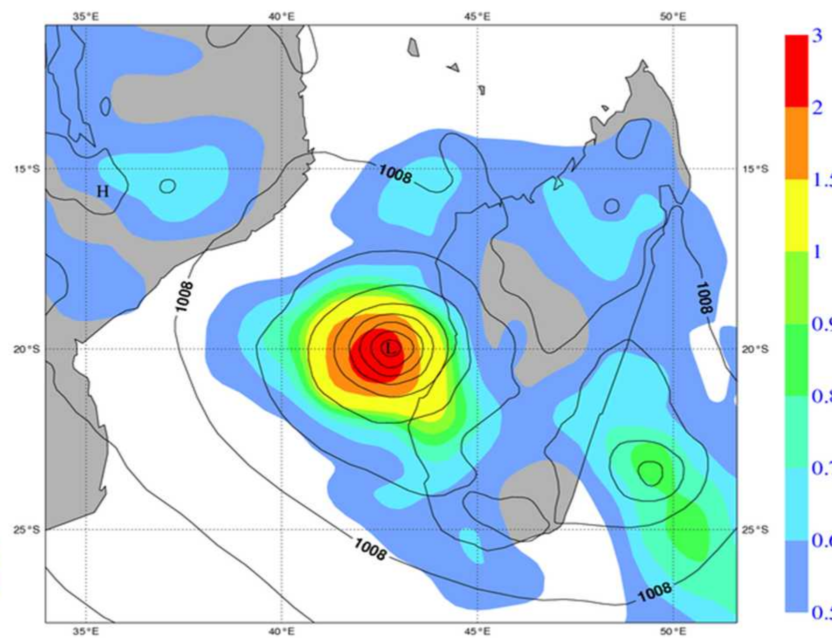
Applications of the EDA

- We want to use EDA perturbations to simulate 4DVar flow-dependent background error covariance evolution
- We start with the EDA flow-dependent estimates of background error variances (diagonal of the B matrix, $\boldsymbol{\Sigma}_b$)

Hurricane Fanele, 20 January 2009



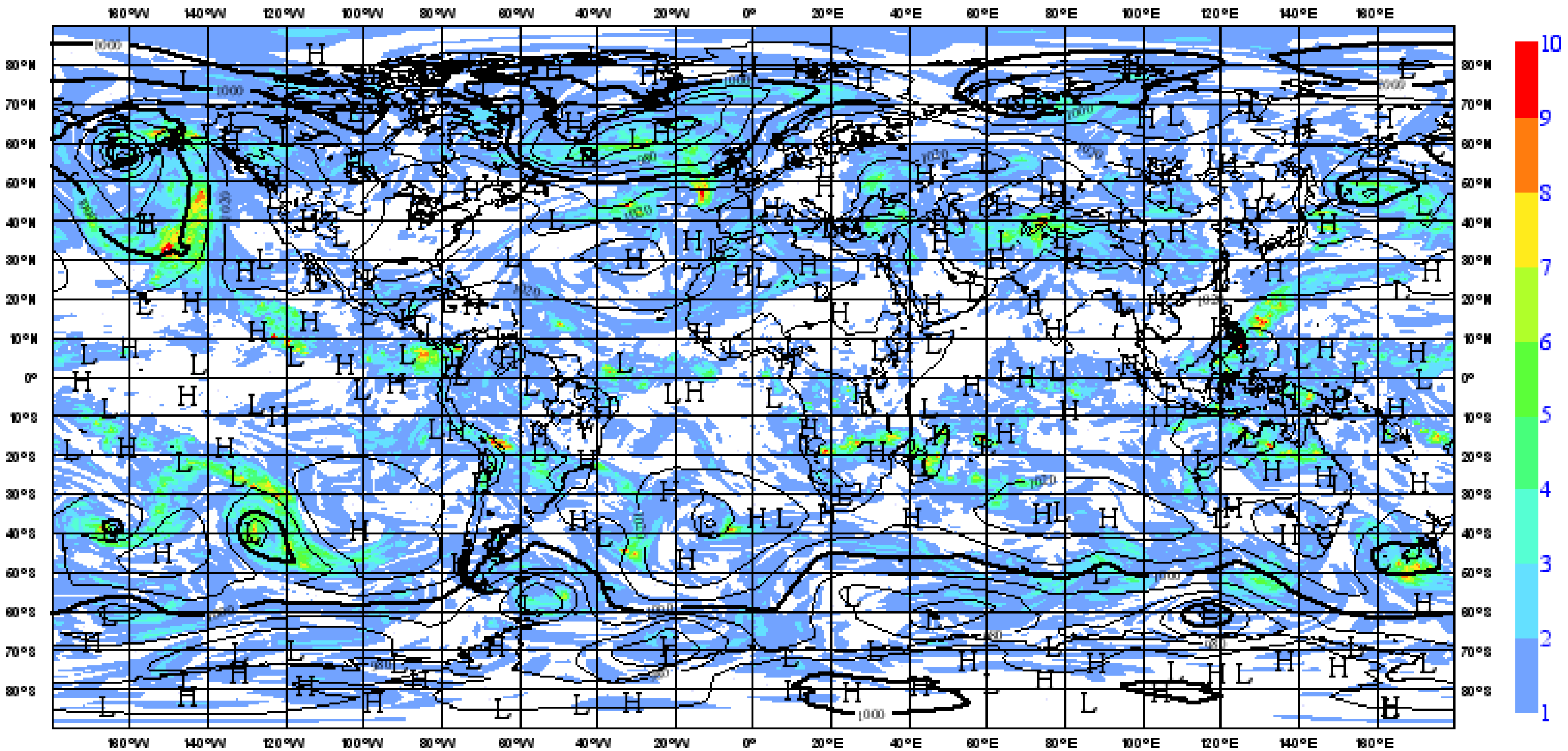
EDA based background error variance for surface pressure



Applications of the EDA

What do raw ensemble variances look like?

Spread of Vorticity t+9h 500hPa



Applications of the EDA

- Noise level is due to sampling errors: 25 member ensemble
- EDA is a stochastic system: error variance of variance estimator $\sim 1/N_{ens}$
- We need a system to effectively filter out noise from first guess ensemble forecast variances: Reduce the random component of the estimation error

Applications of the EDA

Mallat *et al.*: 1998, Annals of Statistics, 26,1-47

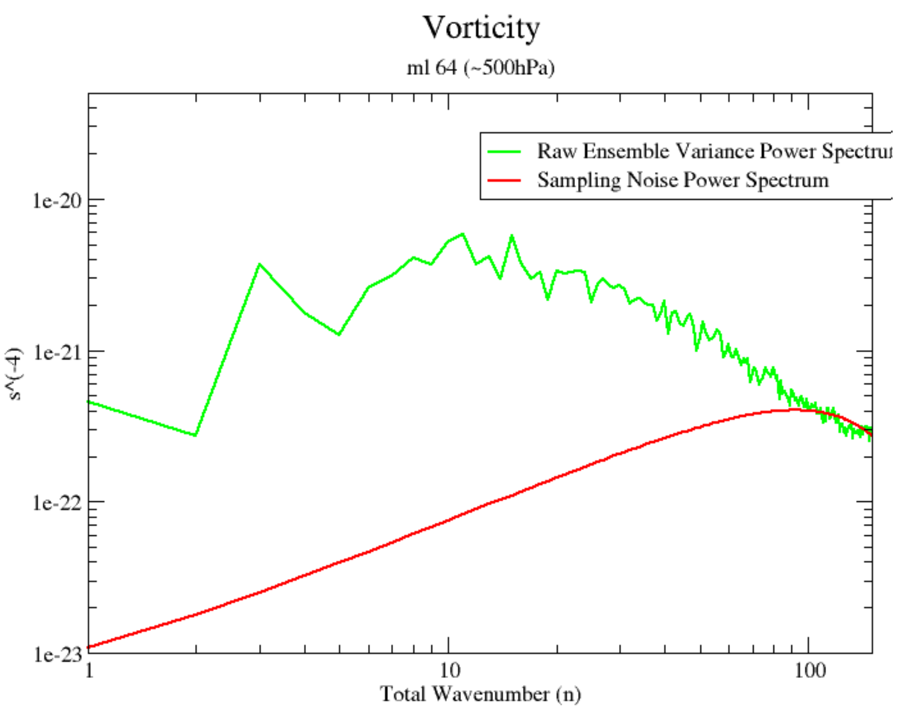
Define $G^e(i)$ as the random component of the sampling error in the estimated ensemble variance at gridpoint *i*:

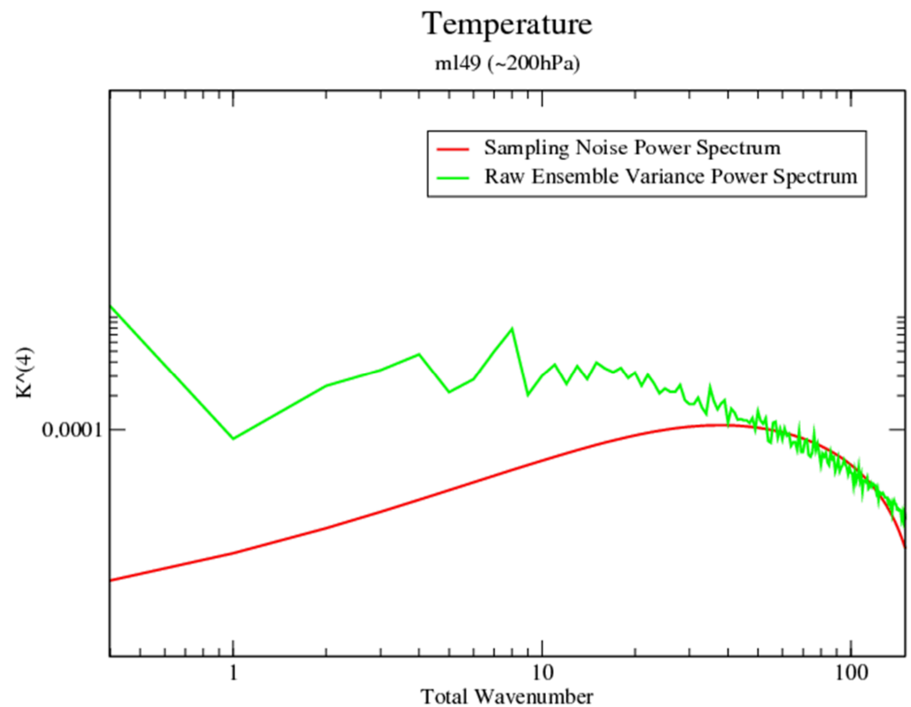
$$G^e(i) \equiv \widetilde{B}_{ii} - E\left[\widetilde{B}_{ii}\right]$$

Then the covariance of the sampling noise can be shown to be a simple function of the expectation of the ensemble-based covariance matrix:

$$E\left[G^e(i)G^e(j)\right] = \frac{2}{N-1}\left(E\left[\widetilde{B}_{ij}\right]\right)^2 \qquad (1)$$

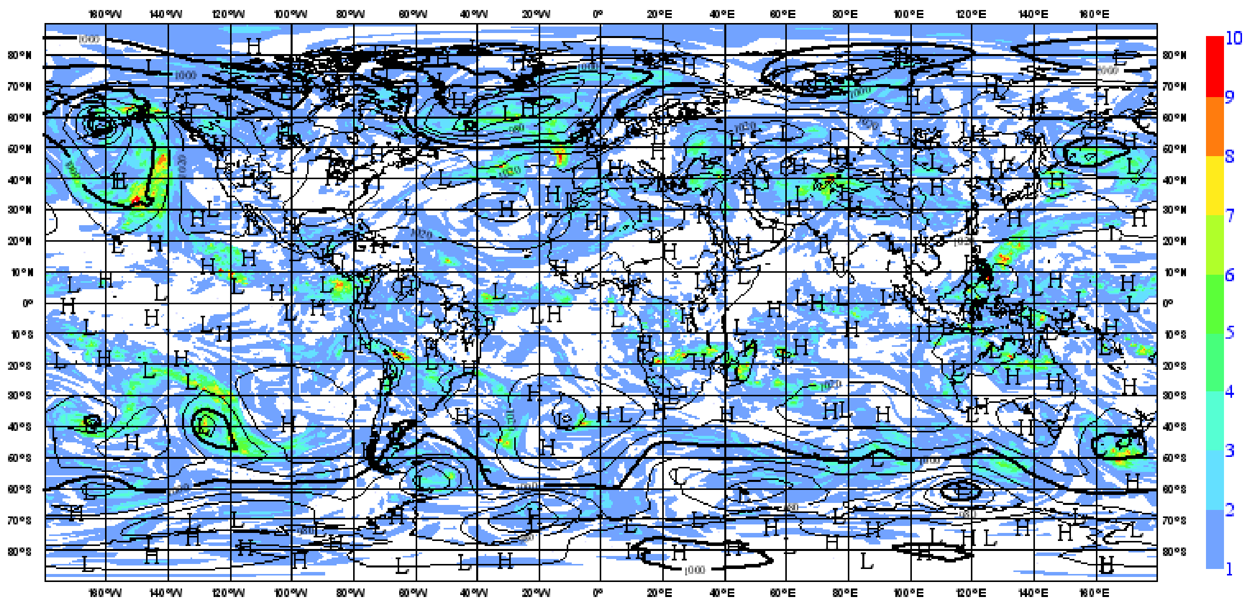
Applications of the EDA



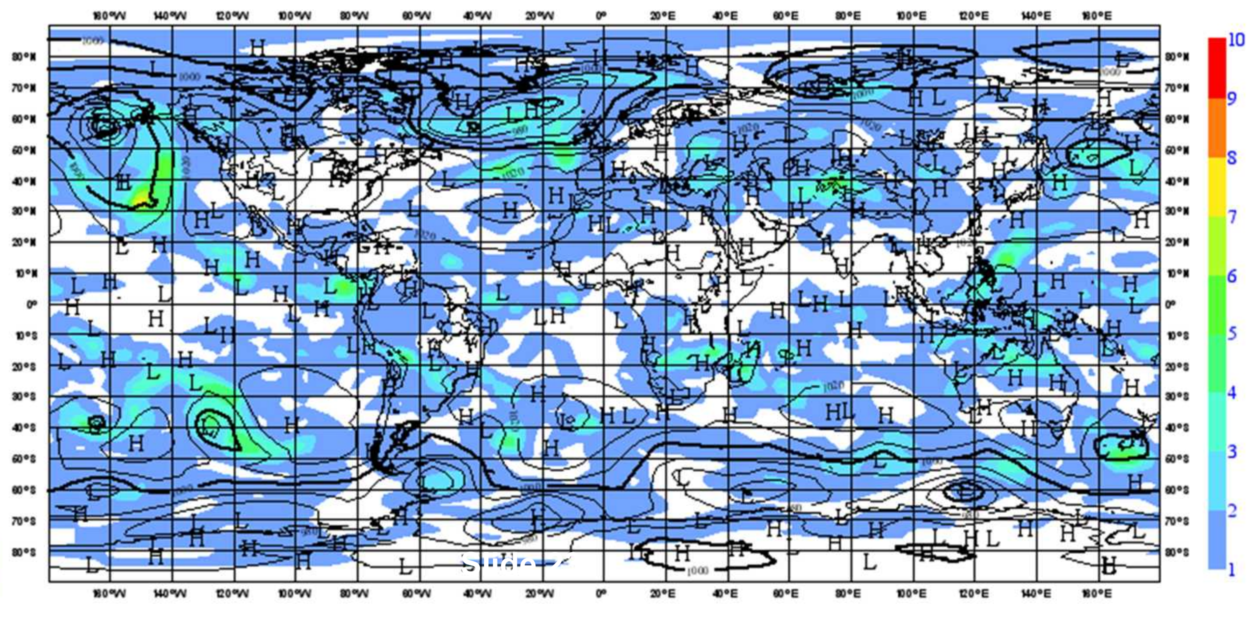


- We can use a spectral filter to disentangle noise from signal
- Truncation wavenumber is determined by **maximizing signal-to-noise** ratio of filtered variances (Raynaud *et al.*, 2009; Bonavita *et al.*, 2011)

Applications of the EDA

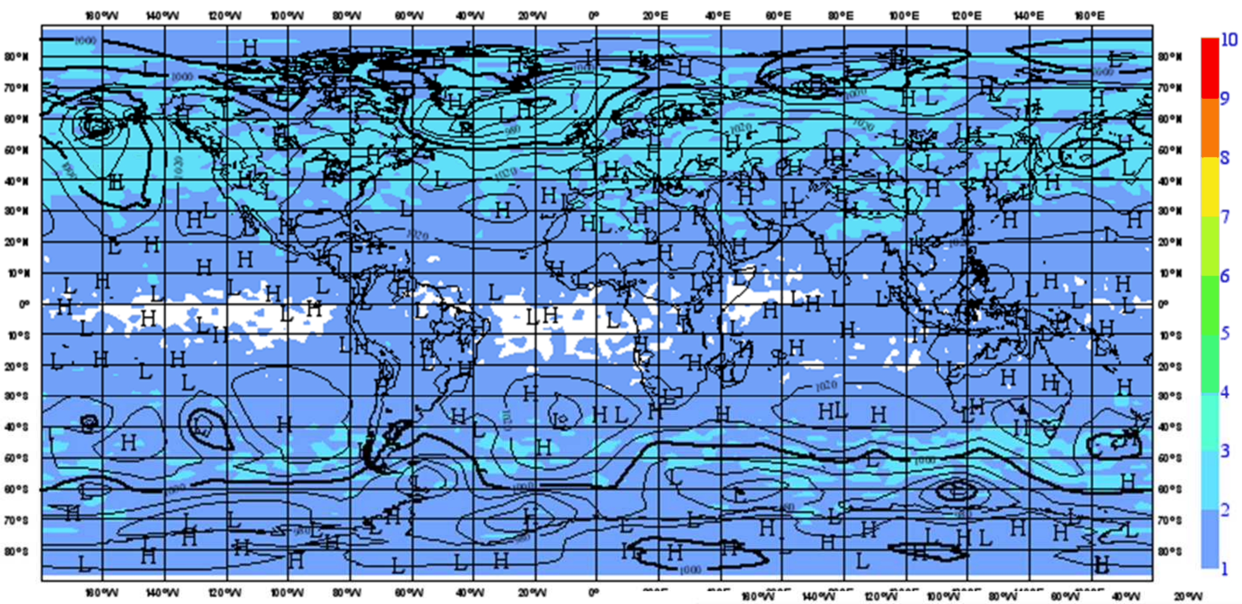


Raw EDA StDev
Vorticity 500 hPa

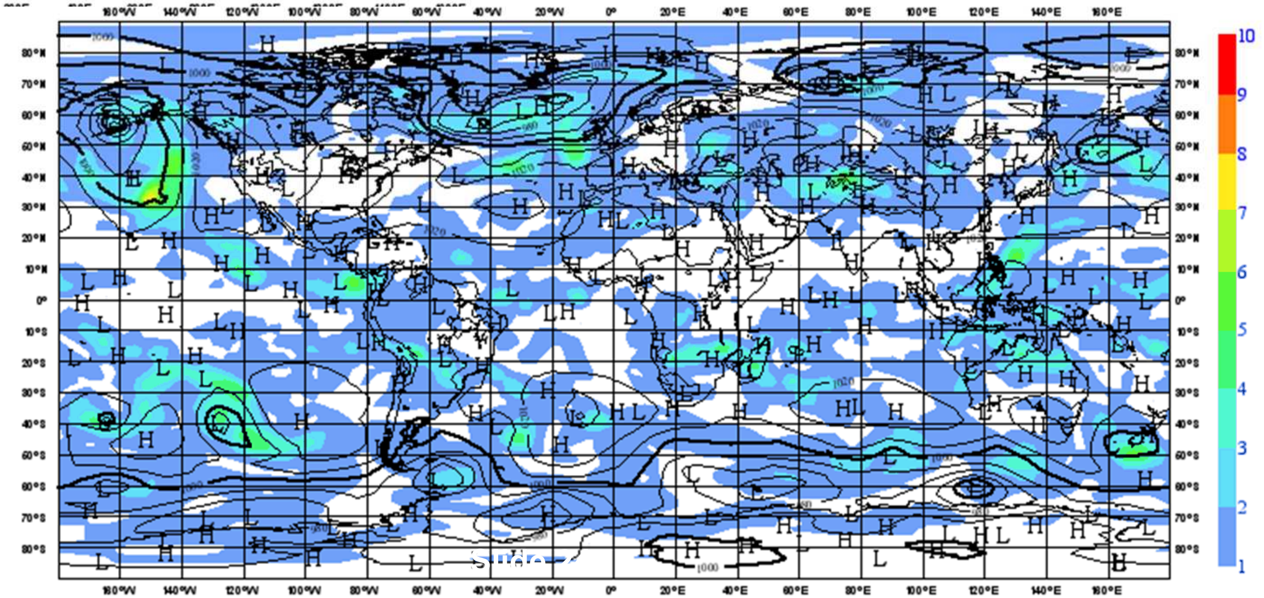


Filtered EDA StDev
Vorticity 500 hPa

Applications of the EDA



StDev of Vorticity at 500 hPa
estimated from **static B**
Random. Method
(Fisher & Courtier, 1995)



Filtered EDA estimate of
StDev of Vorticity 500 hPa

Applications of the EDA

Is there also a systematic error in our EDA sampled variances?

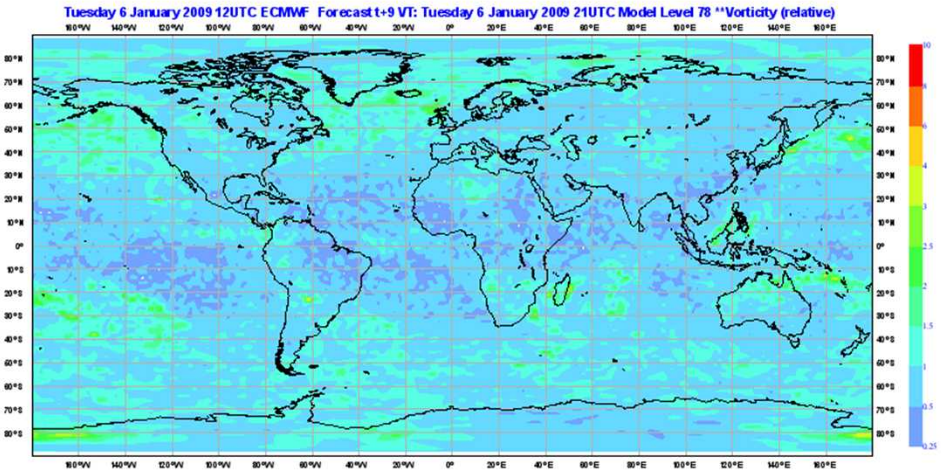
A statistically consistent ensemble satisfies:

$$(1-1/N_{ens})^{-1}\langle Ens\_Var \rangle = (1+1/N_{ens})^{-1}\langle Ens\_Mean\_Square\_Error \rangle$$

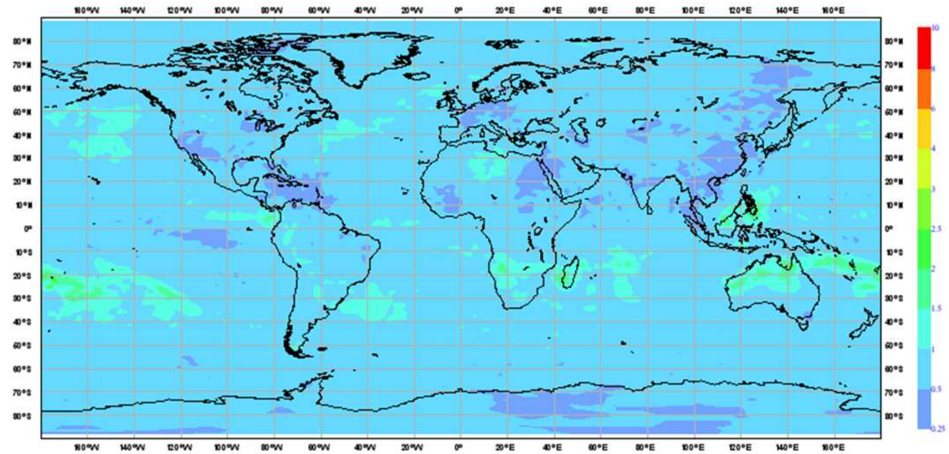
Applications of the EDA

Vorticity ml 78 (~850hPa)

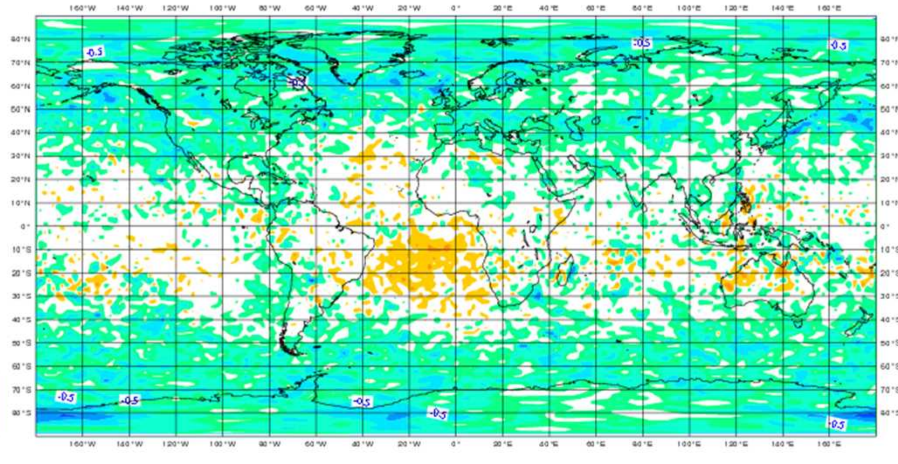
Ensemble Error



Ensemble Spread

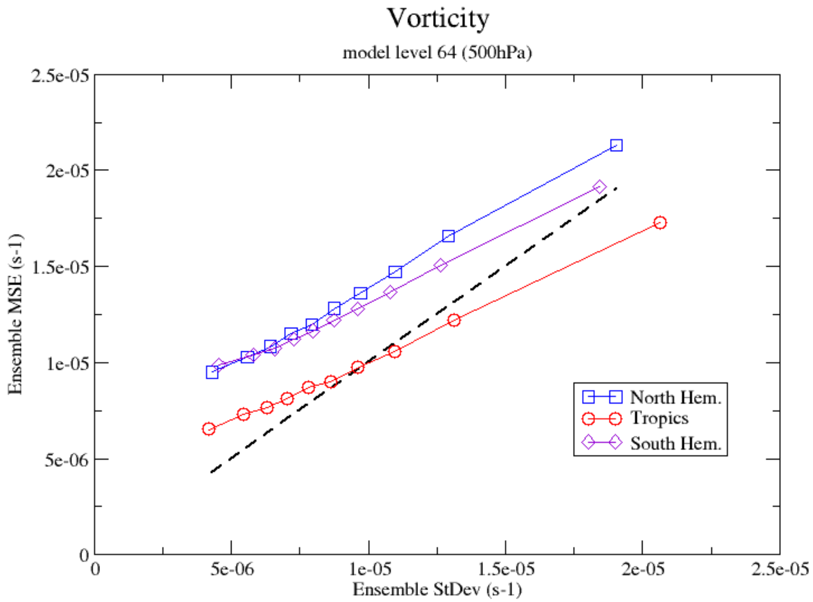


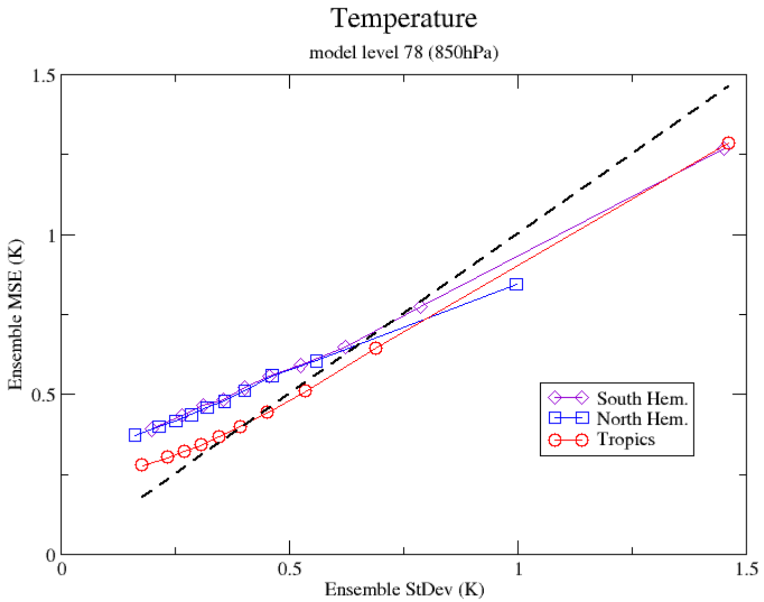
Spread - Error

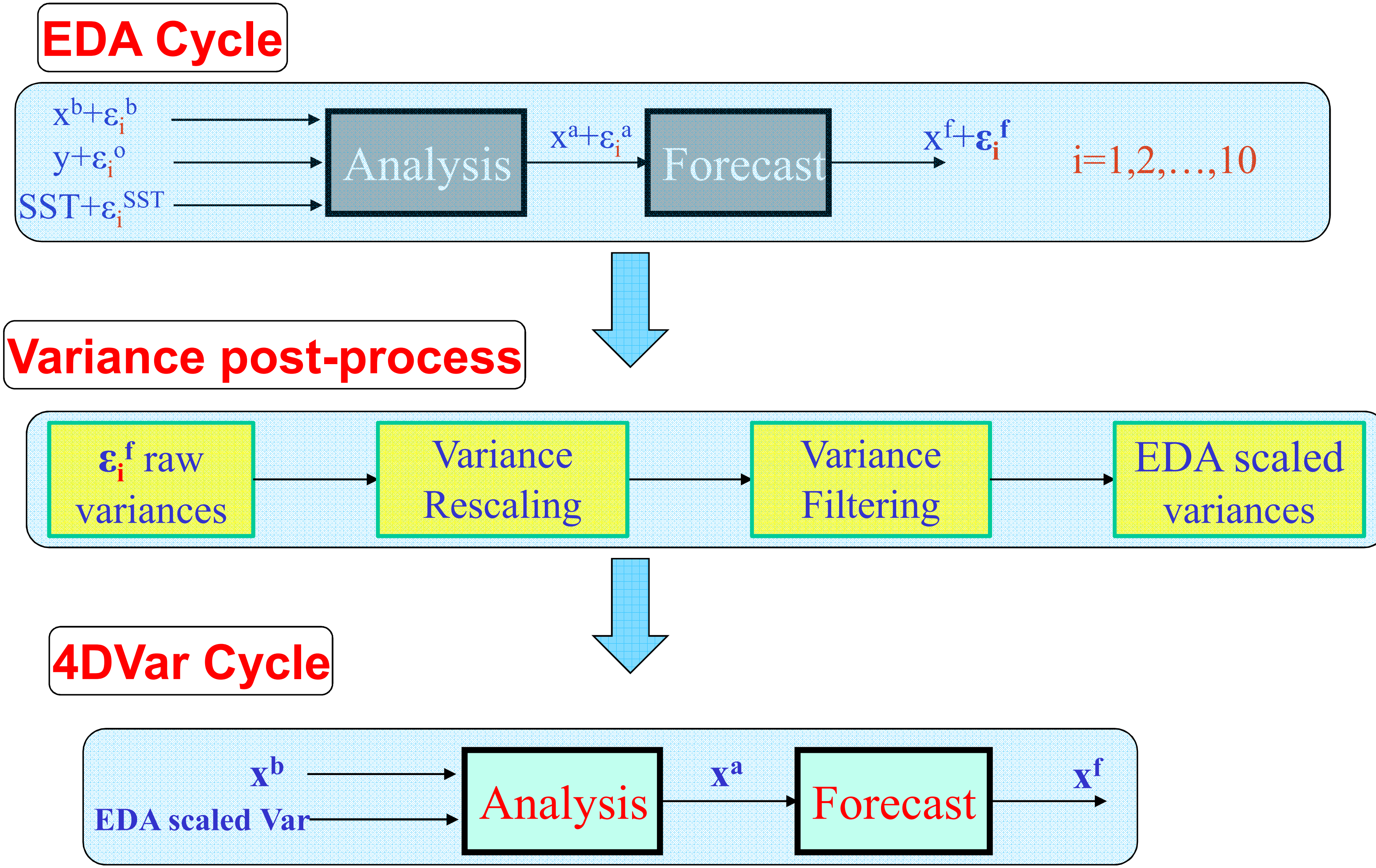


Applications of the EDA

- To get statistically consistent EDA variances we need to perform an online calibration (Ensemble Variance Calibration; Kolczynsky et al., 2009, 2011; Bonavita et al., 2011)
- Calibration factors are also state-dependent, i.e. depend on the size of the expected error
- Need to perform calibration of variances reflects underlying problem in **Q** and **R** models, system non-linearities, ensemble size






EDA Cycle
x^b+ε_i^b
y+ε_i^o
SST+ε_i^SST
Analysis
x^a+ε_i^a
Forecast
x^f+ε_i^f
i=1,2,…,10
Variance post-process
ε_i^f raw variances
Variance Rescaling
Variance Filtering
EDA scaled variances
4DVar Cycle
x^b
EDA scaled Var
Analysis
x^a
Forecast
x^f

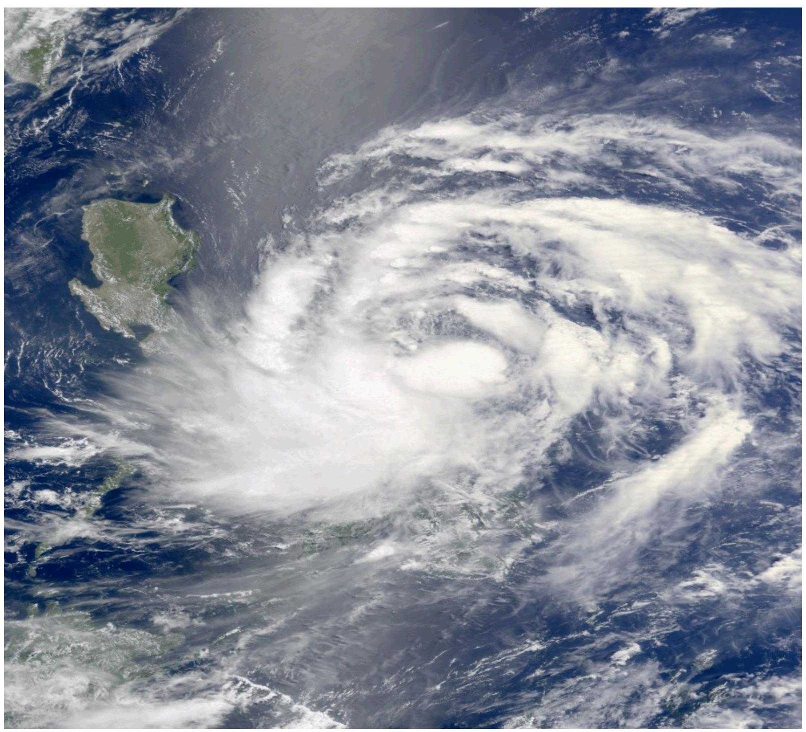
Use of EDA variances in 4DVar

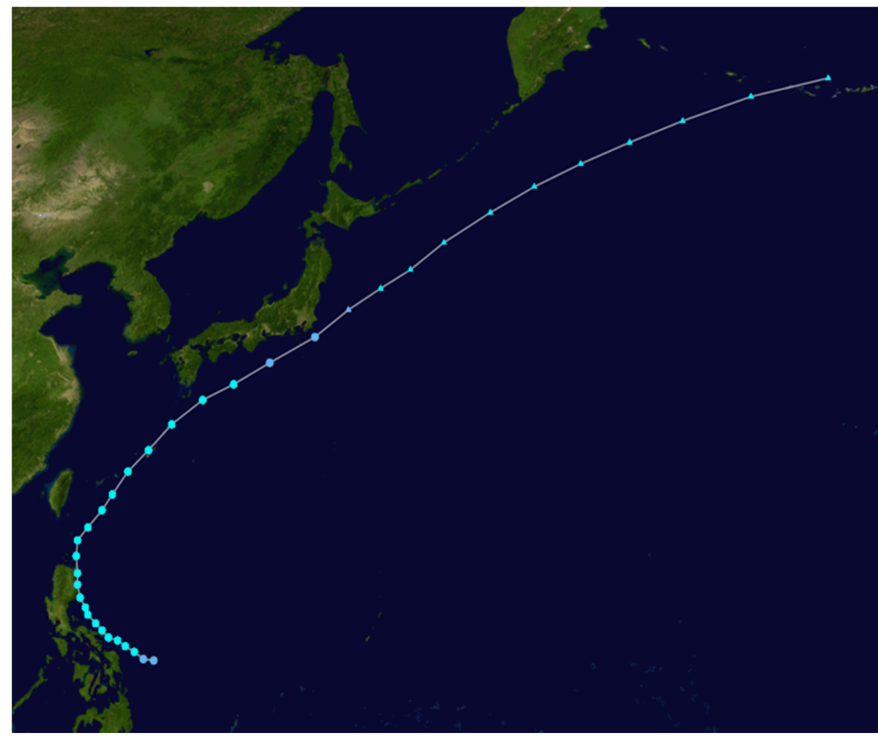
- How does the flow-dependent structure of the EDA variances affect the 4DVar analysis?
- Two mechanisms:
  1. Inside 4DVar EDA variances change the shape and size of analysis increments
  2. Before 4DVar they affect the observation quality control decisions

Use of EDA variances in 4DVar

1. Inside 4DVar EDA variances change the shape and size of analysis increments

- Tropical Cyclone Aere, Philippines 8-9 May 2011.





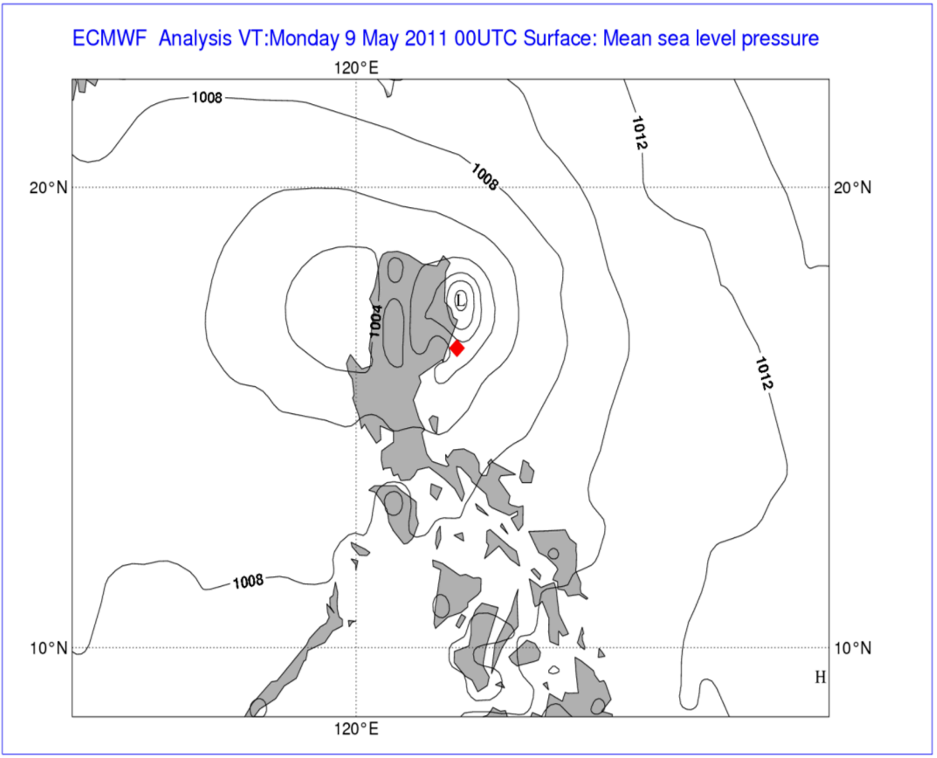
ECMWF

Use of EDA variances in 4DVar

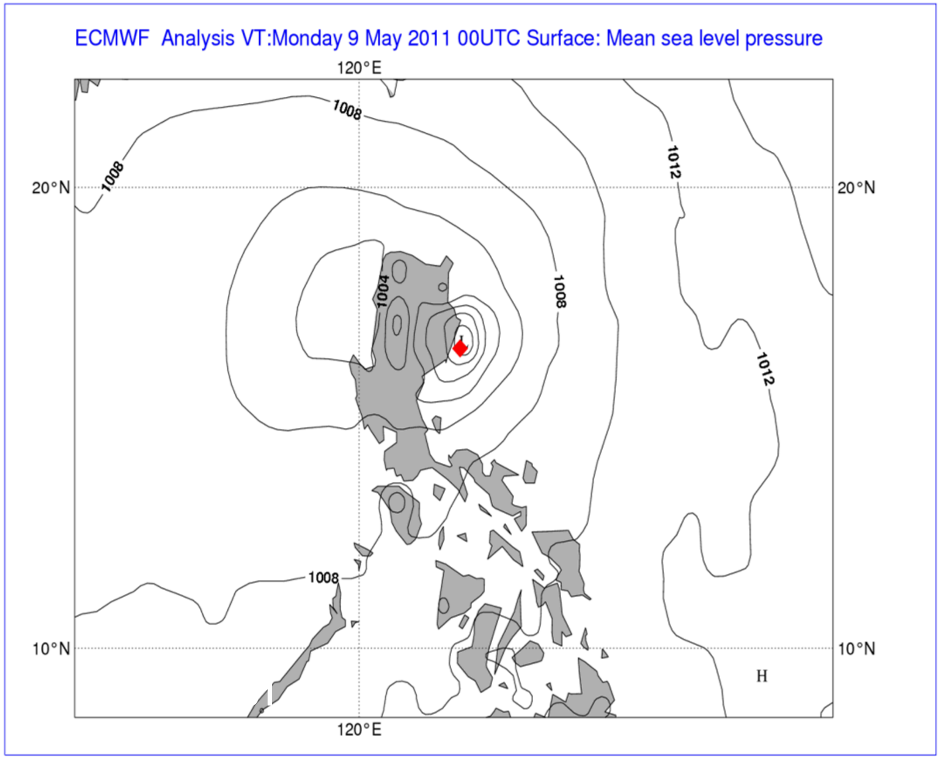
1. Inside 4DVar EDA variances change the shape and size of analysis increments

- Significant operational analysis error, corrected by 4DVar with EDA variances

4DVar with Static errors

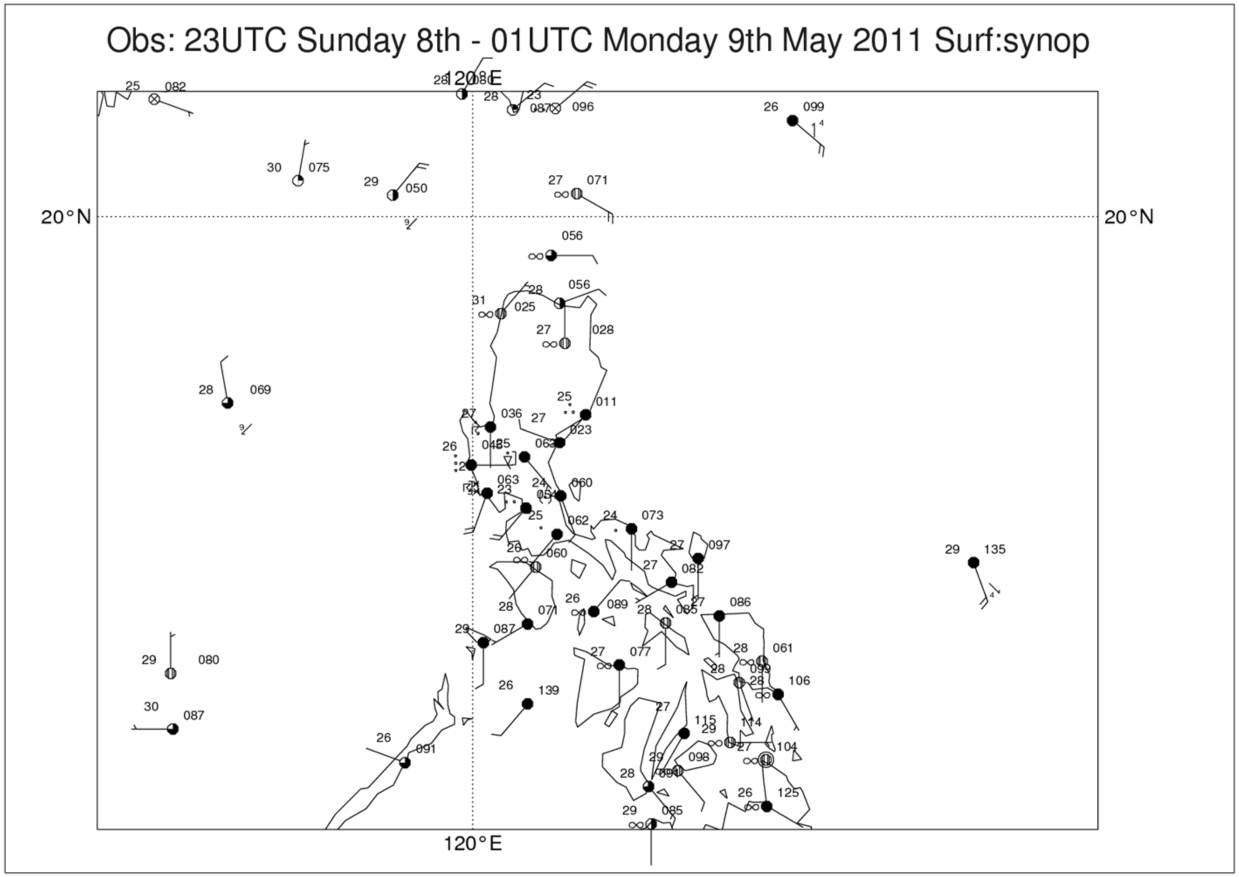


4DVar with EDA errors



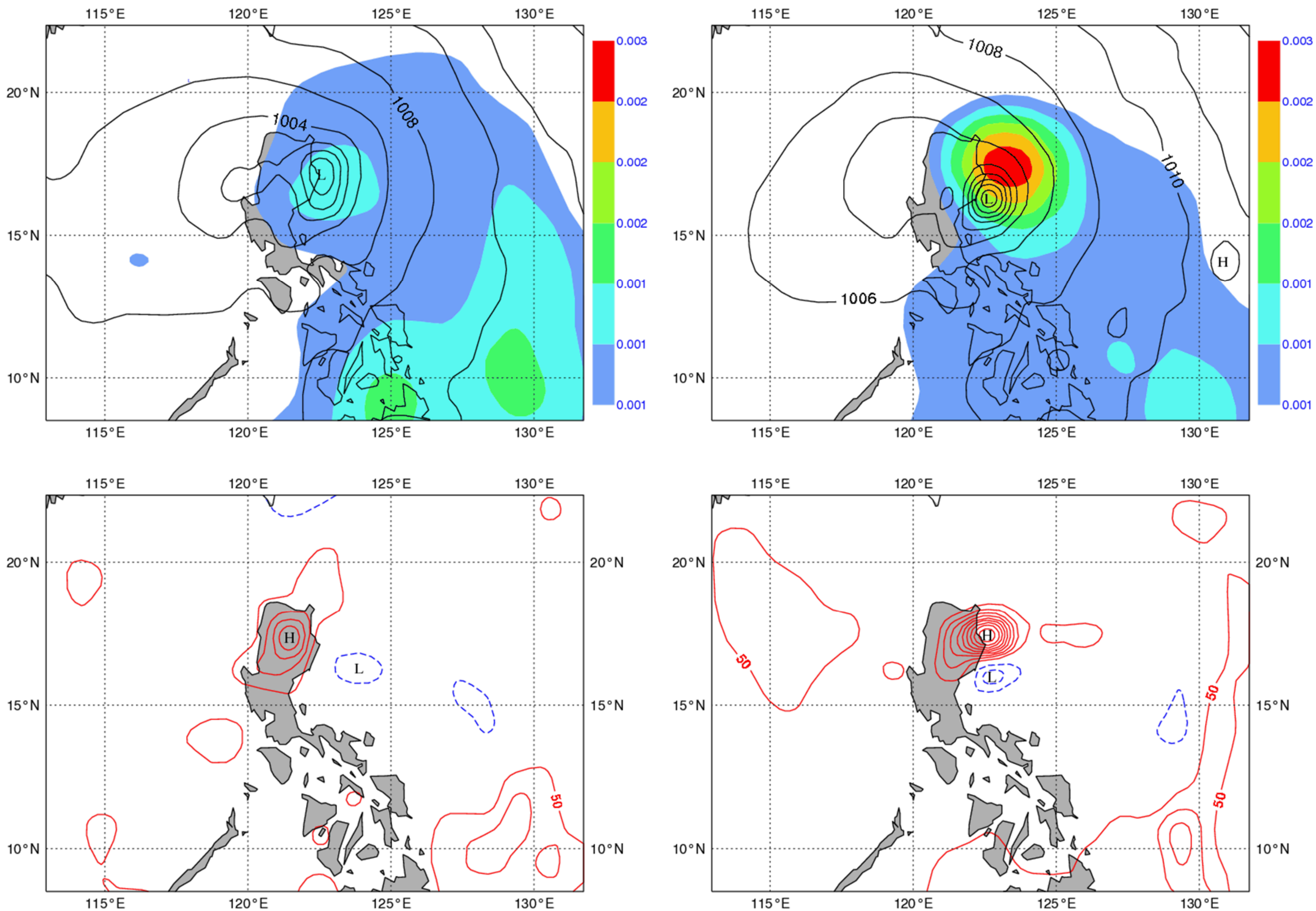
Use of EDA variances in 4DVar

1. Inside 4DVar EDA variances change the shape and size of analysis increments



$\log(P_s)$ Static errors

$\log(P_s)$ EDA errors



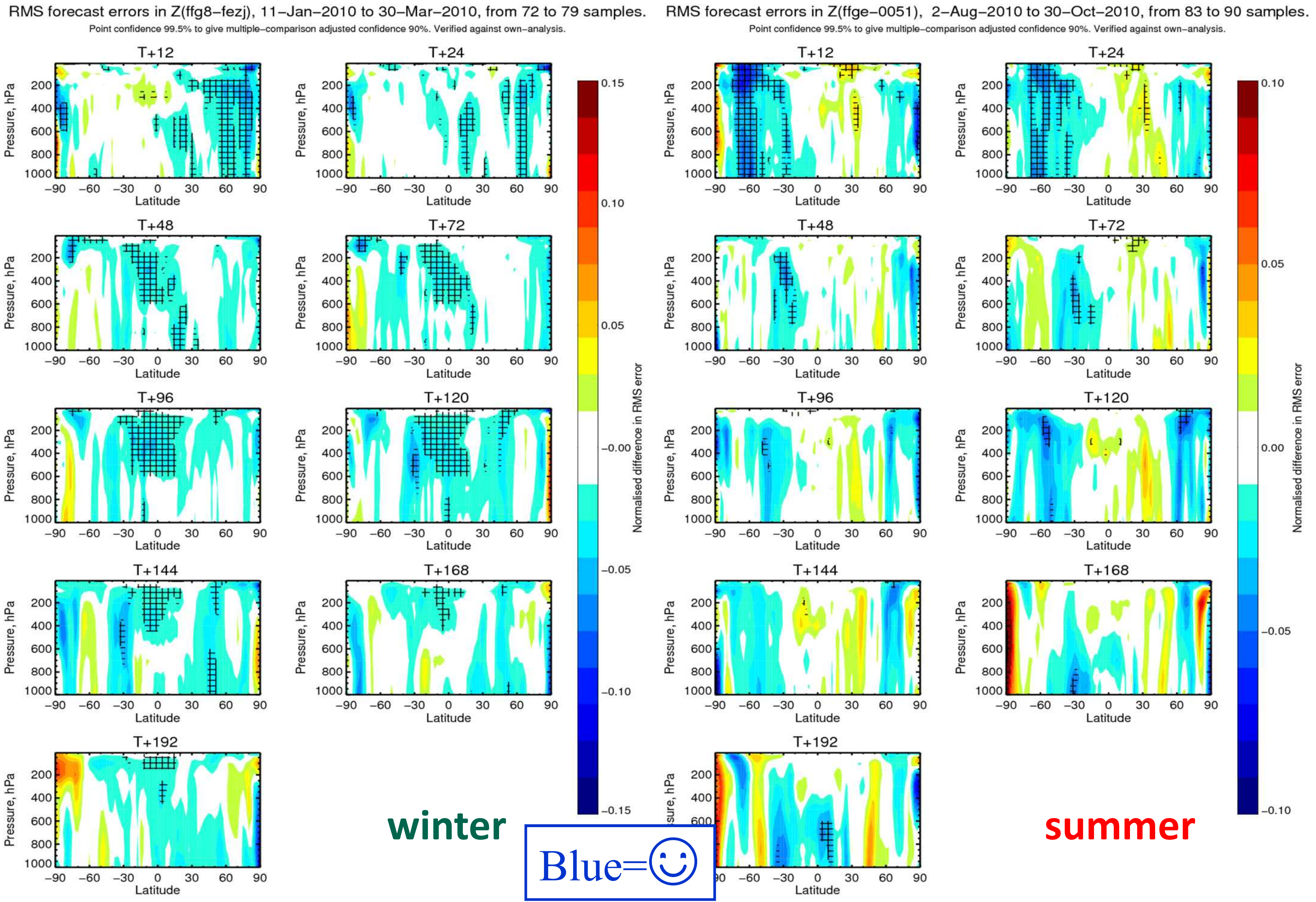
Static mslp ana incr.

EDA mslp ana incr.

Use of EDA variances in 4DVar

- Flow-dependent EDA errors have been used operationally since May 2012 (CY37R2)
- The effect of using flow-dependent EDA estimated errors is large on average skill scores

Geopotential RMSE reduction



Use of EDA covariances in 4DVar

$$(\mathbf{x}-\mathbf{x}_b)=\mathbf{T}^{-1}\mathbf{\Sigma}_b^{1/2}\sum_j \psi_j \otimes \left[\mathbf{C}_j^{1/2}(\lambda,\phi)\chi_j\right]$$

$\mathbf{T}$ is the balance operator

$\mathbf{\Sigma}_\mathrm{b}$ is the gridpoint variance of background errors

$\mathbf{C}_\mathbf{j}(\lambda,\varphi)$ is the vertical covariance matrix for wavelet index *j*

ψ_j are the set of radial basis function that define the wavelet transform

$\mathbf{C}_j(\lambda,\varphi)$ are fields of full vertical covariance matrices, defined for each wavelet band. They determine both the horizontal and vertical background error correlation structures.

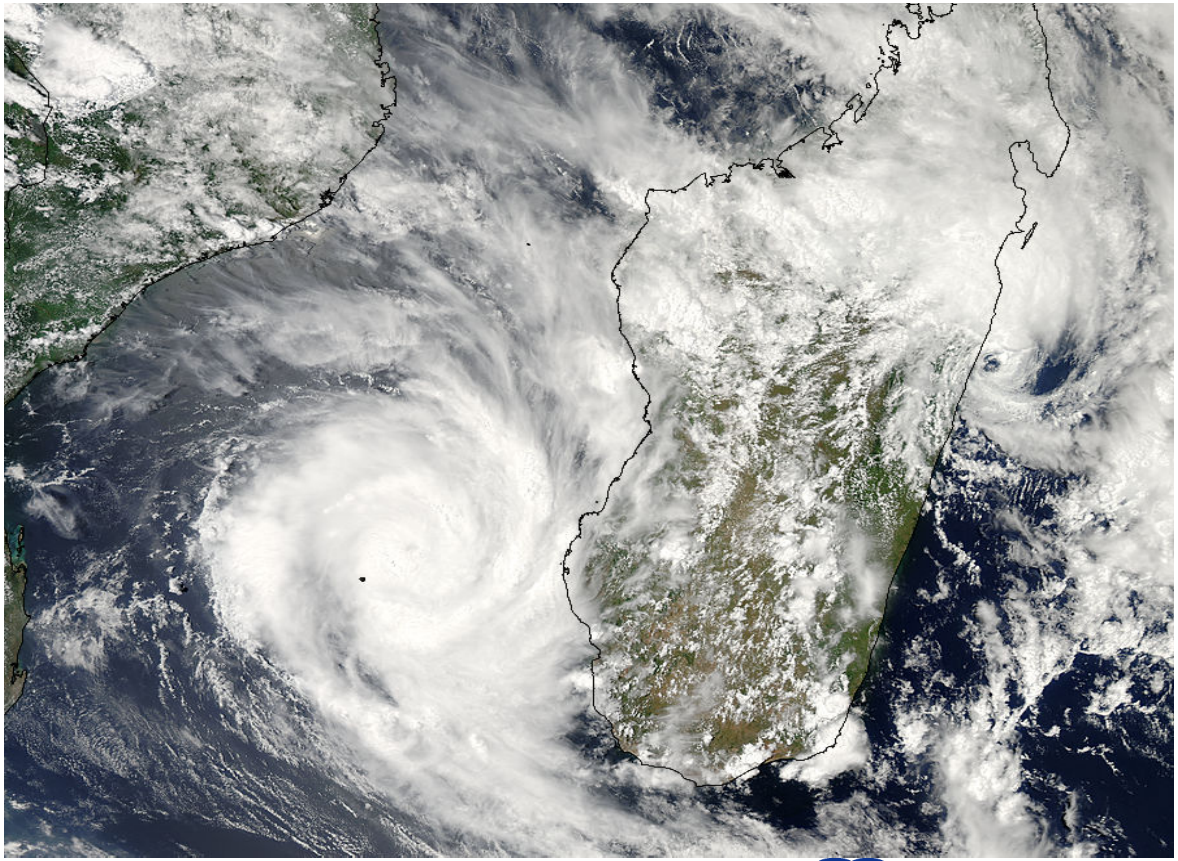
In order to get flow-dependent error covariances we need flow-dependent estimates of $\mathbf{C}_j(\lambda,\varphi)$.

Use of EDA covariances in 4DVar

Diagnosing the Background Error Correlation Length-Scales

$$L \equiv \sqrt{-\frac{\rho(r)}{\partial^2 \rho(r)/\partial r^2}\bigg|_{r=0}}$$

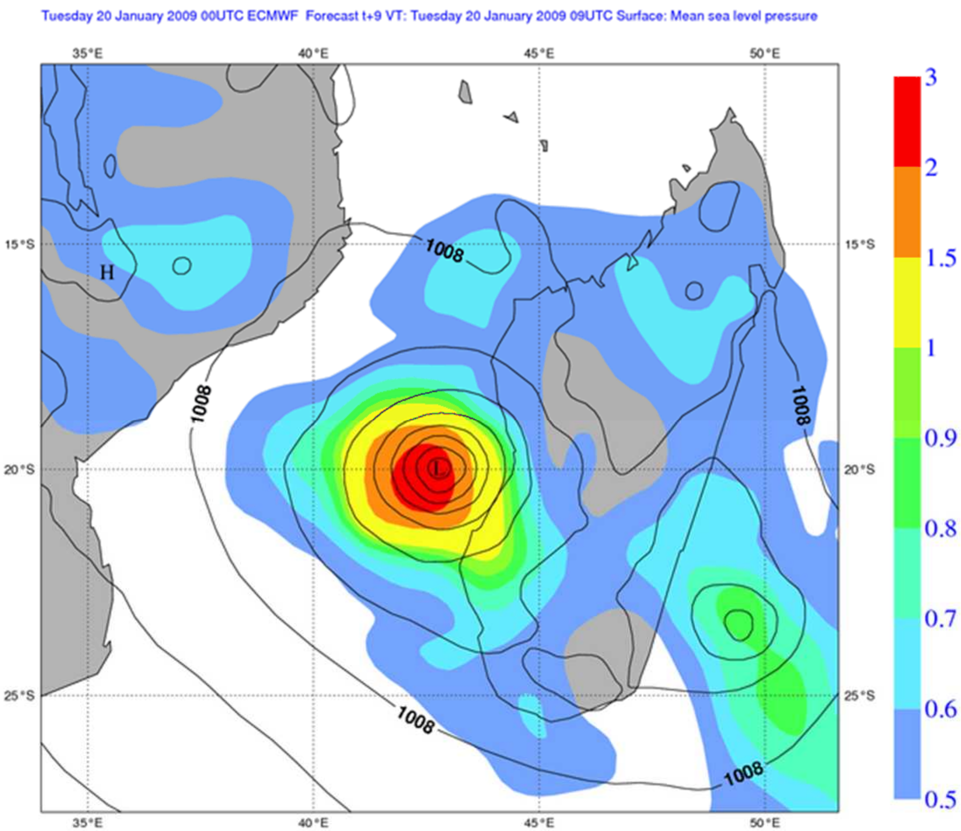
Hurricane Fanele
20 January 2009

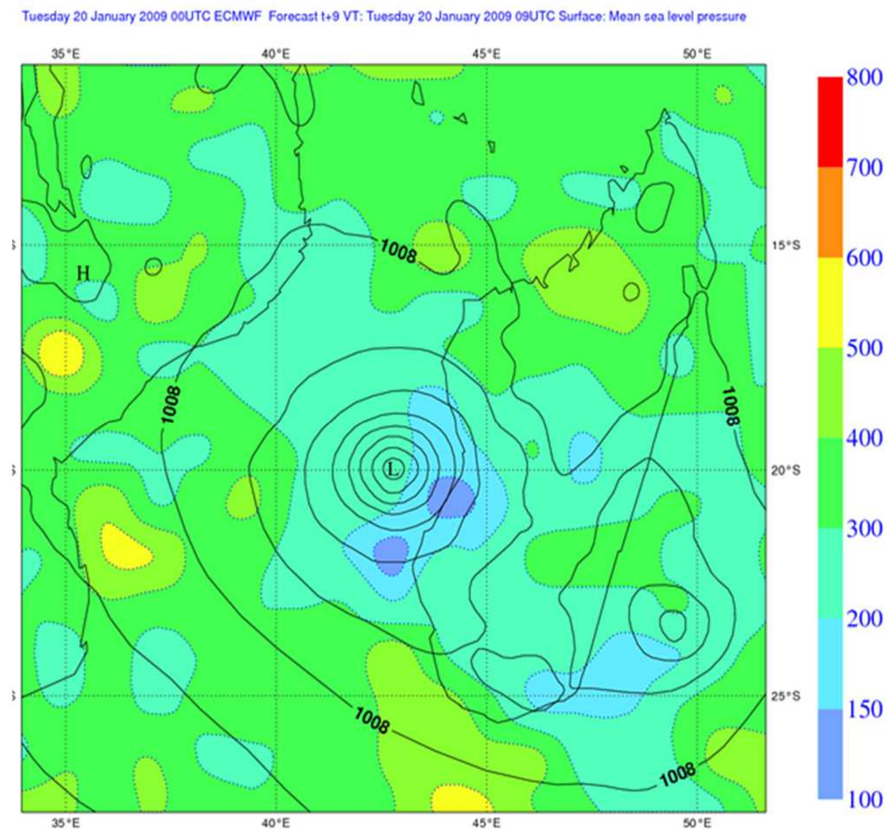


Use of EDA covariances in 4DVar

20 member EDA

Surf. Press. Background Err. St.Dev. Surf. Press. BG Err. Correlation L. Scale



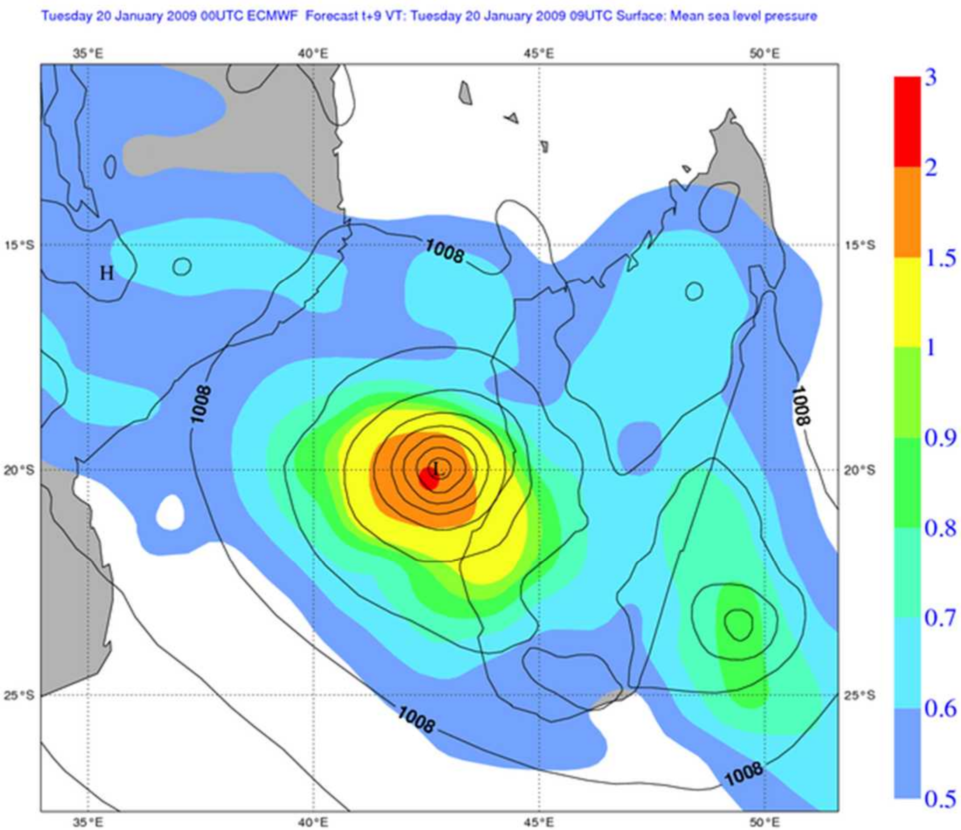


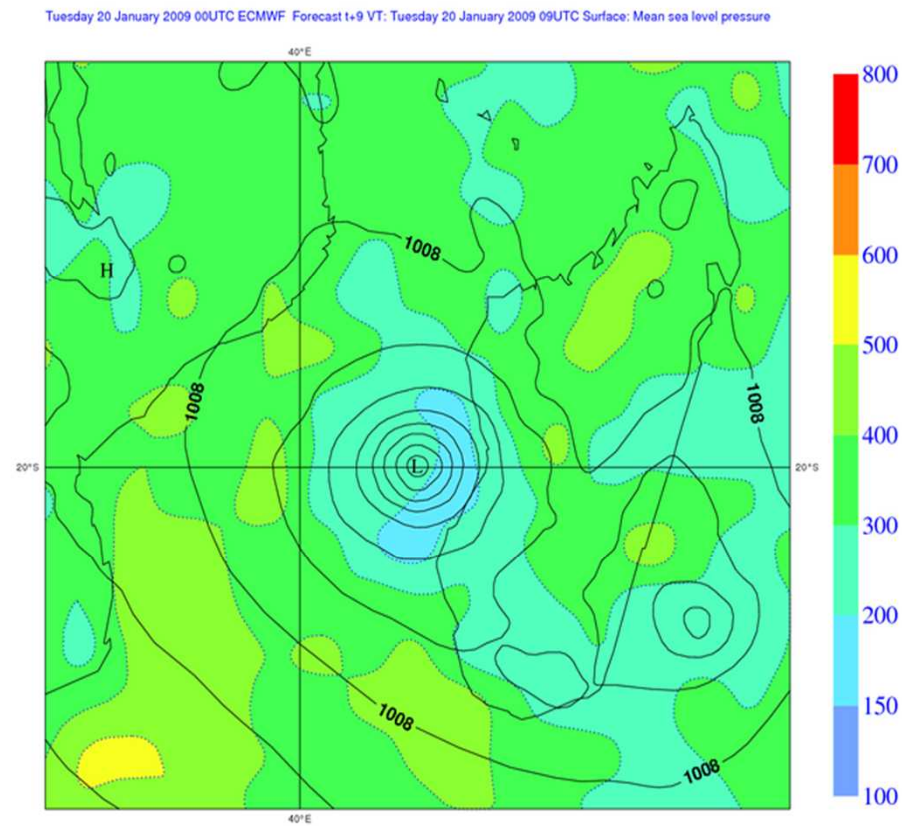
ECMWF

Use of EDA covariances in 4DVar

50 member EDA

Surf. Press. Background Err. St.Dev. | Surf. Press. BG Err. Correlation L. Scale





ECMWF

Use of EDA covariances in 4DVar

a) From November 2013 (CY40R1) background error covariances (wavelet JB) are computed online, i.e. they are updated at every assimilation time (00, 12 UTC)

b) EDA perturbations from the past 12 days are used, with an exponential decay factor (i.e., reduce noise at the cost of losing some flow-dependent detail)

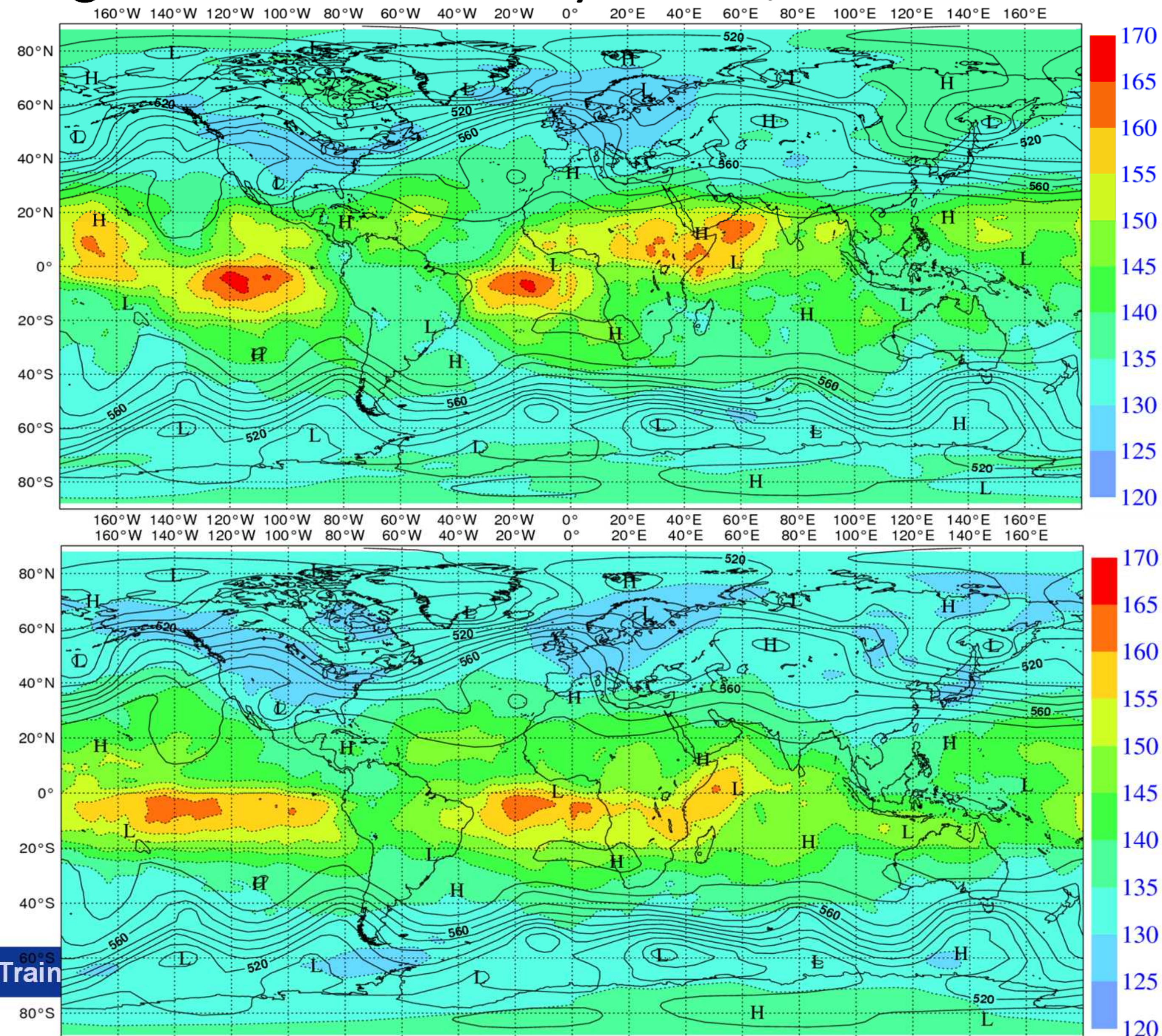
c) Continuously updated JB is used in High Resolution 4D-Var

Why is a flow-dependent JB better?

Correlation length scale of Vorticity errors, ~500 hPa

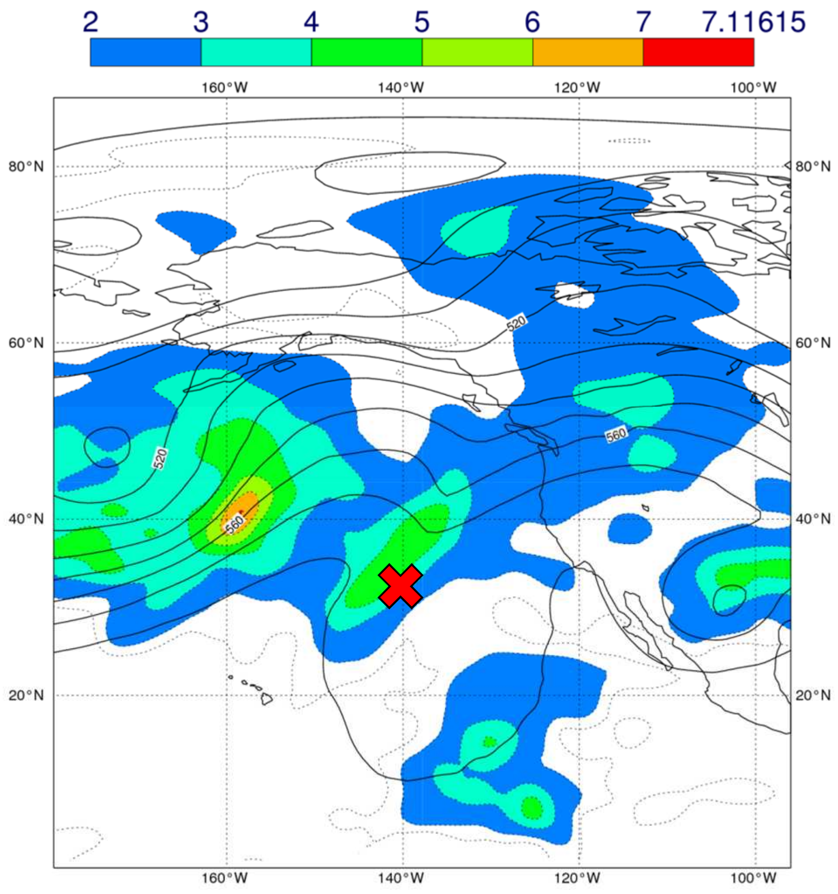
Online wavelet JB, valid 20120110 00Z

Static wavelet JB

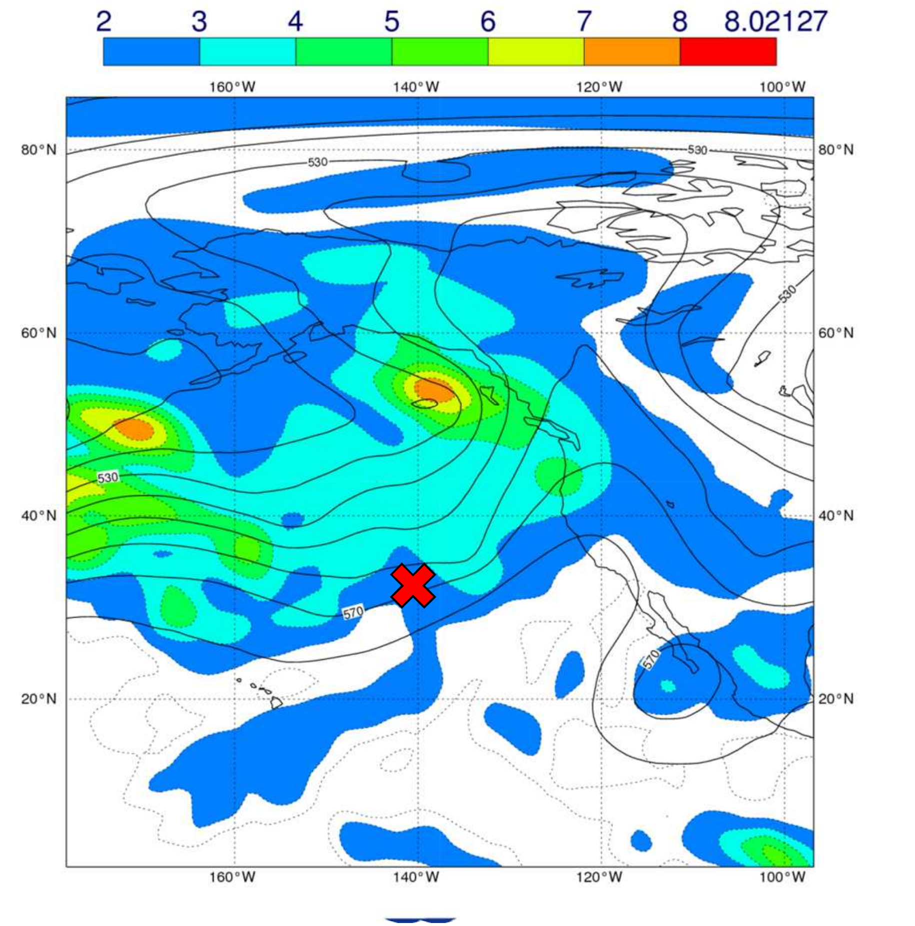


Why is a flow-dependent JB better?

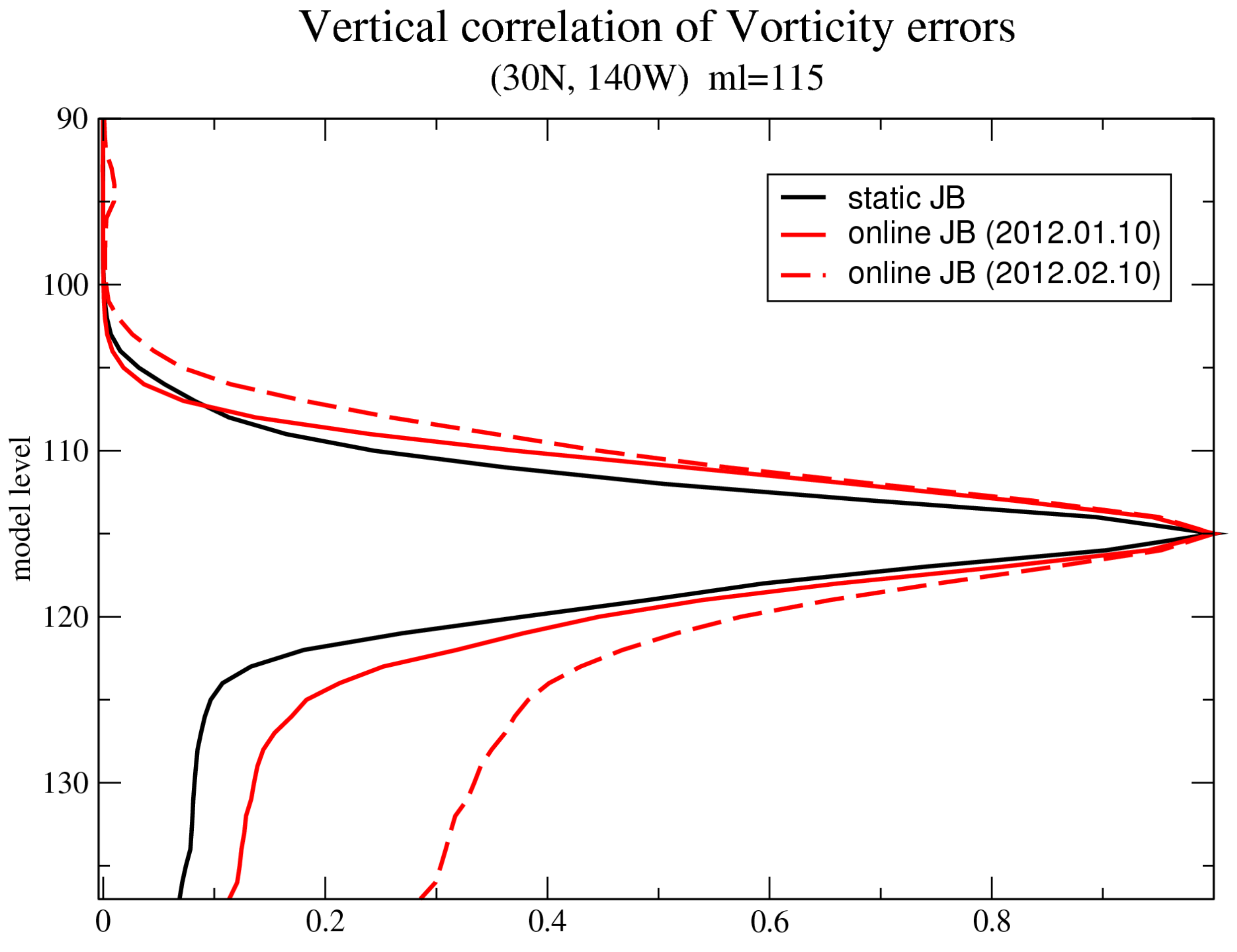
2012-01-01 00Z



2012-02-01 00Z

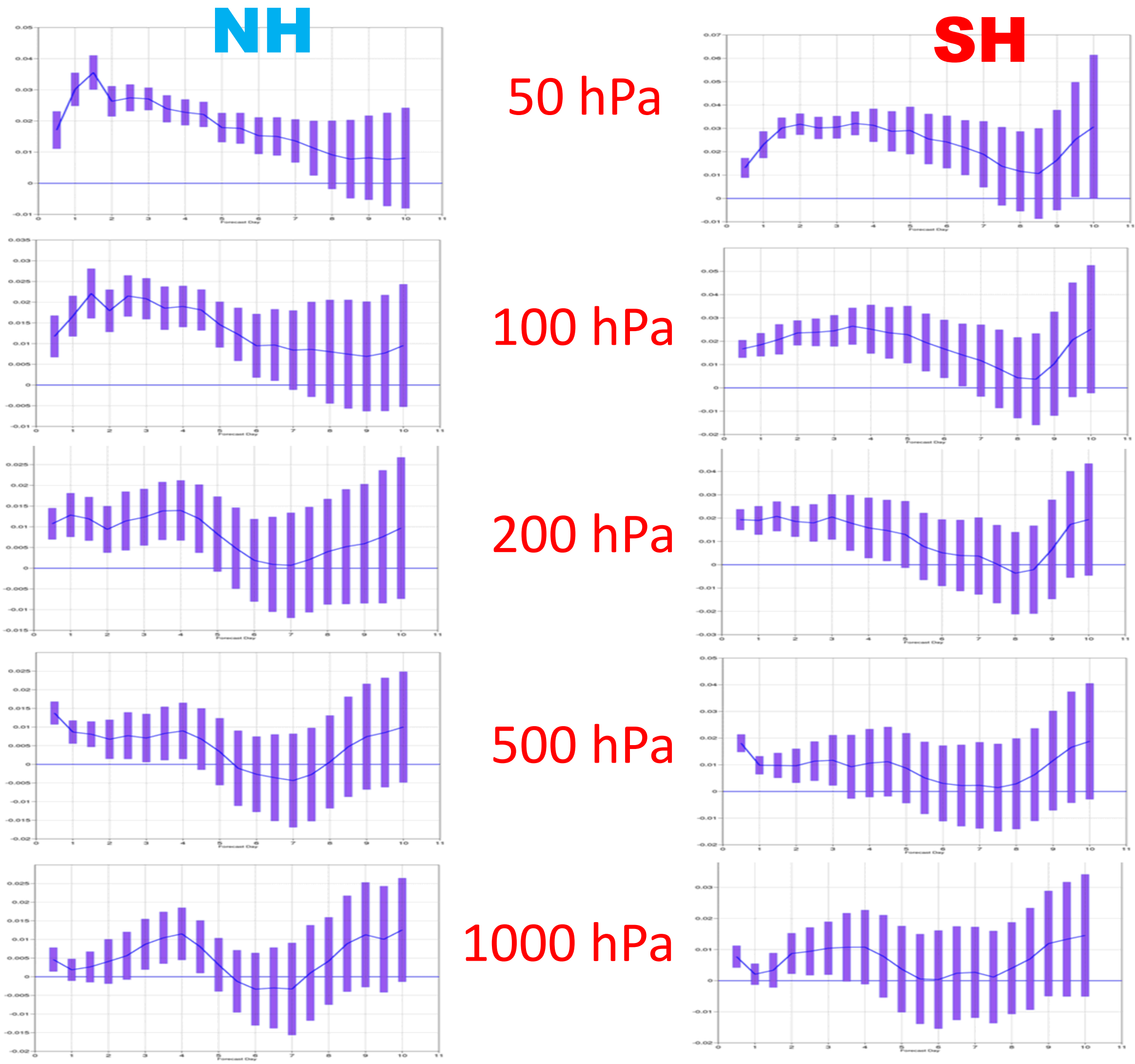


Why is a flow-dependent JB better?



Impact of online JB

Reduction in Geopotential RMSE - 95% confidence



Period: Feb - June 2012

T511L91, 3 Outer Loops (T159/T255/T255)

Verified against operational analysis

Summary

- The EDA is the best available tool at ECMWF to estimate analysis and background errors of the ECMWF analysis
- The EDA derived analysis errors contribute to the estimation of initial perturbations of the Ensemble Prediction System (EPS)
- The EDA background perturbations provide flow-dependent estimates of background errors variances and covariances for use in the High-Resolution 4DVar analysis

Summary

- The use of EDA background perturbations in a variational analysis is currently done in two ways:

  a) adding an ensemble, flow-dependent component to the static B used in 4D-Var (UK MetO, NCEP; alpha control variable method: Lorenc, 2003)

  b) using EDA perturbations to get an on-line, flow-dependent estimate of parameterised **B** (ECMWF, Meteo France; EDA approach)

Summary

- Flow-dependent background errors from EDA variances have been used in ECMWF 4D-Var since April 2011 (Cycle 37R2)

- They benefit the analysis and forecast skill by:

  a) changing the weight given to observations near dynamically active zones;

  b) introducing a level of flow-dependency in the analysis increments

  c) Allowing a state and flow-dependent QC of observations

- Increase in ensemble size will benefit the system allowing less aggressive filtering of raw variances (Bonavita *et al.*, 2011)

Summary

- Flow-dependent background error covariances (**B**) estimated from EDA perturbations has been implemented recently (CY40R1, November 2013)
- They also have an important effect on the analysis by providing flow-dependent, non homogeneous, correlation structures
- The introduction of flow-dependent background errors and covariances estimated from the EDA has been one of the largest source of improvement in recent years in the analysis and forecast skill of the IFS

References

1. Anderson, J.L., 2007: An adaptive covariance inflation error correction algorithm for ensemble filters. Tellus, 59A, 210–224.
2. Bishop, C. H., and D. Hodyss, 2007: Flow adaptive moderation of spurious ensemble correlations and its use in ensemble-based data assimilation. Quart. J. Roy. Meteor. Soc., 133, 2029–2044.
3. Bishop, C. H., and D. Hodyss, 2009a: Ensemble covariances adaptively localized with ECO-RAP, Part 1: Tests on simple error models. Tellus, 61, 84–96.
4. Bonavita M., L. Raynaud and L. Isaksen, 2010: Estimating background-error variances with the ECMWF Ensemble of Data Assimilations system: the effect of ensemble size and day-to-day variability. *Q. J. R. Meteorol. Soc., 137: 423–434*.
5. Bonavita M., L. Isaksen and E. Holm, 2012: On the use of EDA background error variances in the ECMWF 4D-Var. *Q. J. R. Meteorol. Soc..* Early Online Release, doi: 10.1002/qj.1899
6. R. Buizza, P. L. Houtekamer, Z. Toth, G. Pellerin, M. Wei & Y. Zhu, 2005: A comparison of the ECMWF, MSC and NCEP Global Ensemble Prediction Systems. Mon. Wea. Rev., 133, 5, 1076-1097.
7. Burgers, G., P. J. van Leeuwen, and G. Evensen, 1998: Analysis scheme in the ensemble Kalman filter. Mon. Wea. Rev., 126, 1719-1724.
8. Campbell, William F., Craig H. Bishop, Daniel Hodyss, 2010: Vertical Covariance Localization for Satellite Radiances in Ensemble Kalman Filters. Mon. Wea. Rev., 138, 282–290.
9. Evensen, G., 1994: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics. J. Geophys. Res., 99, 10143-10162.

References

10. Evensen, G., 2003: The Ensemble Kalman Filter: theoretical formulation and practical implementation. Ocean Dynamics 53: 343–367
11. Fisher M, Leutbecher M, Kelly GA. 2005 On the equivalence between Kalman smoothing and weak-constraint four-dimensional variational data assimilation. Q. J. R. Meteorol. Soc. 131 3235–3246.
12. Fisher, M., 2003: Background error covariance modelling. Proceedings of the ECMWF Seminar on recent developments in data assimilation for atmosphere and ocean, ECMWF, pages 45–63. (Available from: http://www.ecmwf.int/publications/)
13. Fisher, M. and Courtier, P., 1995: Estimating the covariance matrices of analysis and forecast error in variational data assimilation. ECMWF Technical Memorandum 220. (Available from: http://www.ecmwf.int/publications/)
14. Isaksen, L., M. Bonavita, R. Buizza, M. Fisher, J. Haseler, M. Leutbecher and Laure Raynaud, 2010: Ensemble of data assimilations at ECMWF. *ECMWF Technical Memorandum* No. 636. (Available from: http://www.ecmwf.int/publications/)
15. Kolczynski W. C., Jr., D. R. Stauffer, S. E. Haupt, A. Deng, 2009: Ensemble Variance Calibration for Representing Meteorological Uncertainty for Atmospheric Transport and Dispersion Modeling. J. Appl. Meteor. Climatol., 48, 2001-2021.
16. Kolczynski W. C., Jr., D. R. Stauffer, S. E. Haupt, N. S. Altman, A. Deng, 2011: Investigation of Ensemble Variance as a Measure of True Forecast Variance. Mon. Weather Rev., Early Online Release, doi: 10.1175/MWR-D-10-05081.1
17. Leutbecher, M., 2010: Diagnosis of ensemble forecasting systems. *Proceedings of the ECMWF Seminar on Diagnosis of Forecasting and Data Assimilation Systems,* 7-10 September 2009, ECMWF, pages 235-266. (Available from: http://www.ecmwf.int/publications/)

References

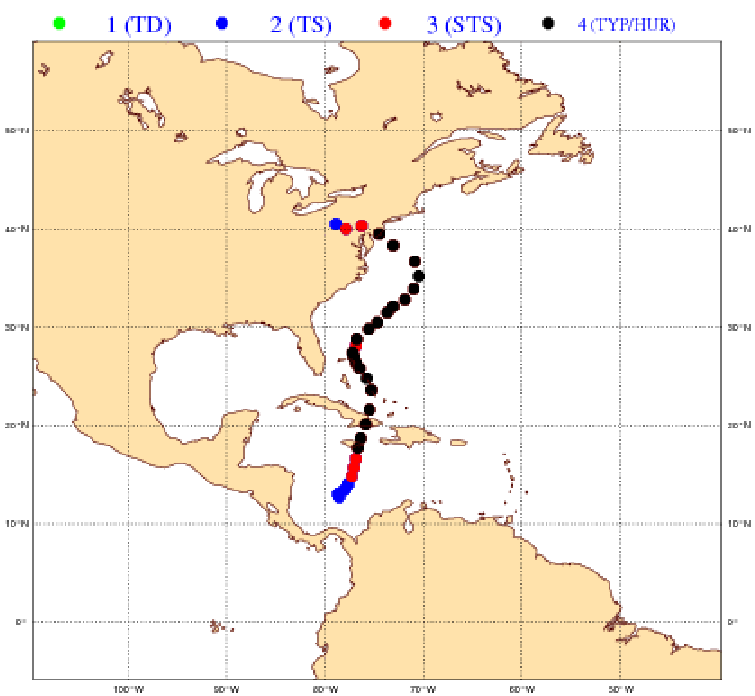
17. Liu C, Xiao Q, Wang B. 2008. An ensemble-based four-dimensional variational data assimilation scheme. part i: Technical formulation and preliminary test. Mon. Weather Rev. 136: 3363–3373.
18. Lorenc, A.C.,2003: The potential of the ensemble Kalman filter for NWP—A comparison with 4D-VAR. *Q. J. R. Meteorol. Soc.*, 129: 3183–3203.
19. Mallat, S., G. Papanicolau, and Z. Zhang, 1998: Adaptive Covariance Estimation of Locally Stationary Processes. Annals of Statistics, 26,1,1-47.
20. Ott E, Hunt BR, Szunyogh I, Zimin AV, Kostelich EJ, Corazza M, Kalnay E, Patil DJ. 2004. A local ensemble Kalman filter for atmospheric data assimilation. Tellus 56A: 415–428.
21. Raynaud, L., L. Berre, and G. Desroziers, 2008: Spatial averaging of ensemble-based background-error variances. Q. J. R. Meteorol. Soc., 134, 1003–1014.
22. Snyder, C, 2012: Particle filters, the "optimal" proposal and high-dimensional systems. *Proceedings of the ECMWF Seminar on Data assimilation for atmosphere and ocean, 6*-9 September 2011, ECMWF, pages 161-170. (Available from: http://www.ecmwf.int/publications/)
23. Thepaut, J.-N., Courtier, P., Belaud, G. and Lemaıtre, G., 1996: Dynamical structure functions in a four-dimensional variational assimilation: A case study. Q. J. R. Meteorol. Soc., 122, 535–561
24. Van Leeuwen, P.J., 2012: Nonlinear large-scale data assimilation: the potential of particle filters. *Proceedings of the ECMWF Seminar on Data assimilation for atmosphere and ocean, 6*-9 September 2011, ECMWF, pages 171-187. (Available from: http://www.ecmwf.int/publications/)
25. Wikle, C. K., & Berliner, L. M., 2007: A Bayesian tutorial for data assimilation. Physica D: Nonlinear Phenomena, 230(1), 1-16.

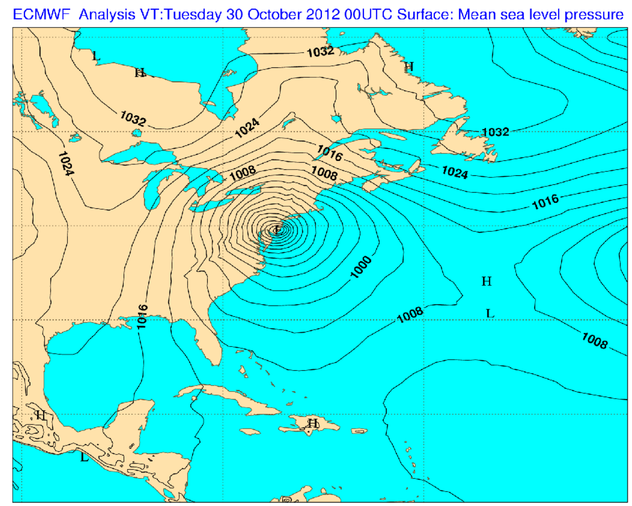
Additional Slides

Use of EDA variances in 4DVar

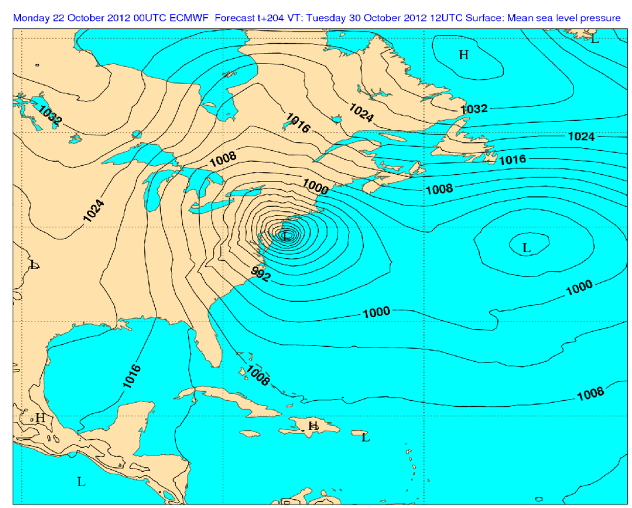
2. Before 4DVar they affect the observation quality control decisions

- Super Storm Sandy





Mslp Ana 30/10/2012 00UTC



Mslp t+204h Fcst valid at 30/10/2012 00UTC

Use of EDA variances in 4DVar

What happens if we withhold polar-orbiters observations (i.e., approx. 90% of obs. counts)?

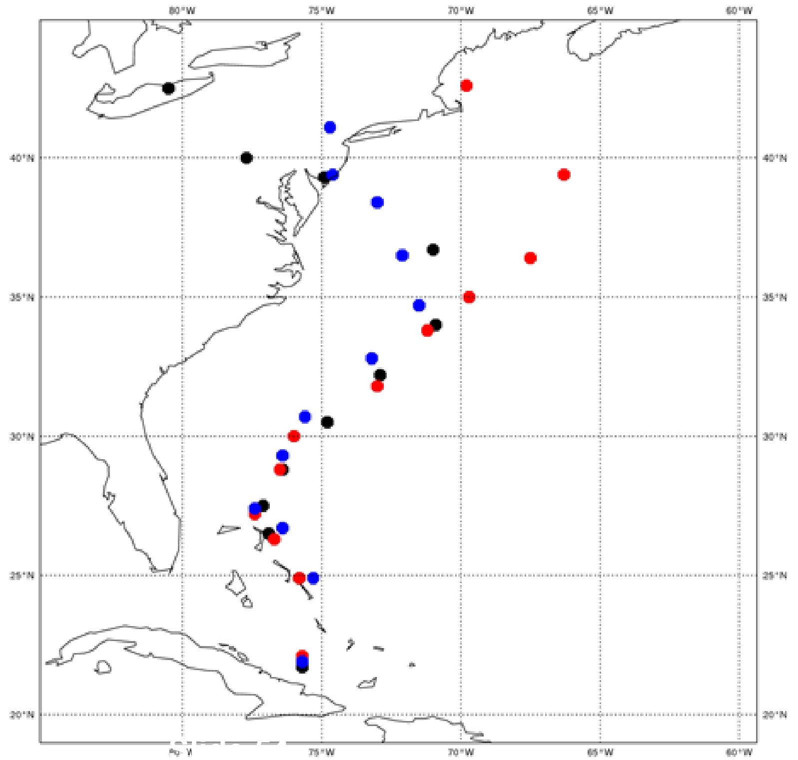
The forecast performance is obviously degraded, and only 5 days before landfall the system recovers the correct track

Sandy's forecast tracks 25 Oct 2013 00UTC

Operational forecast

Forecast from HRES assimilation cycle without polar orbiters and errors from operational EDA

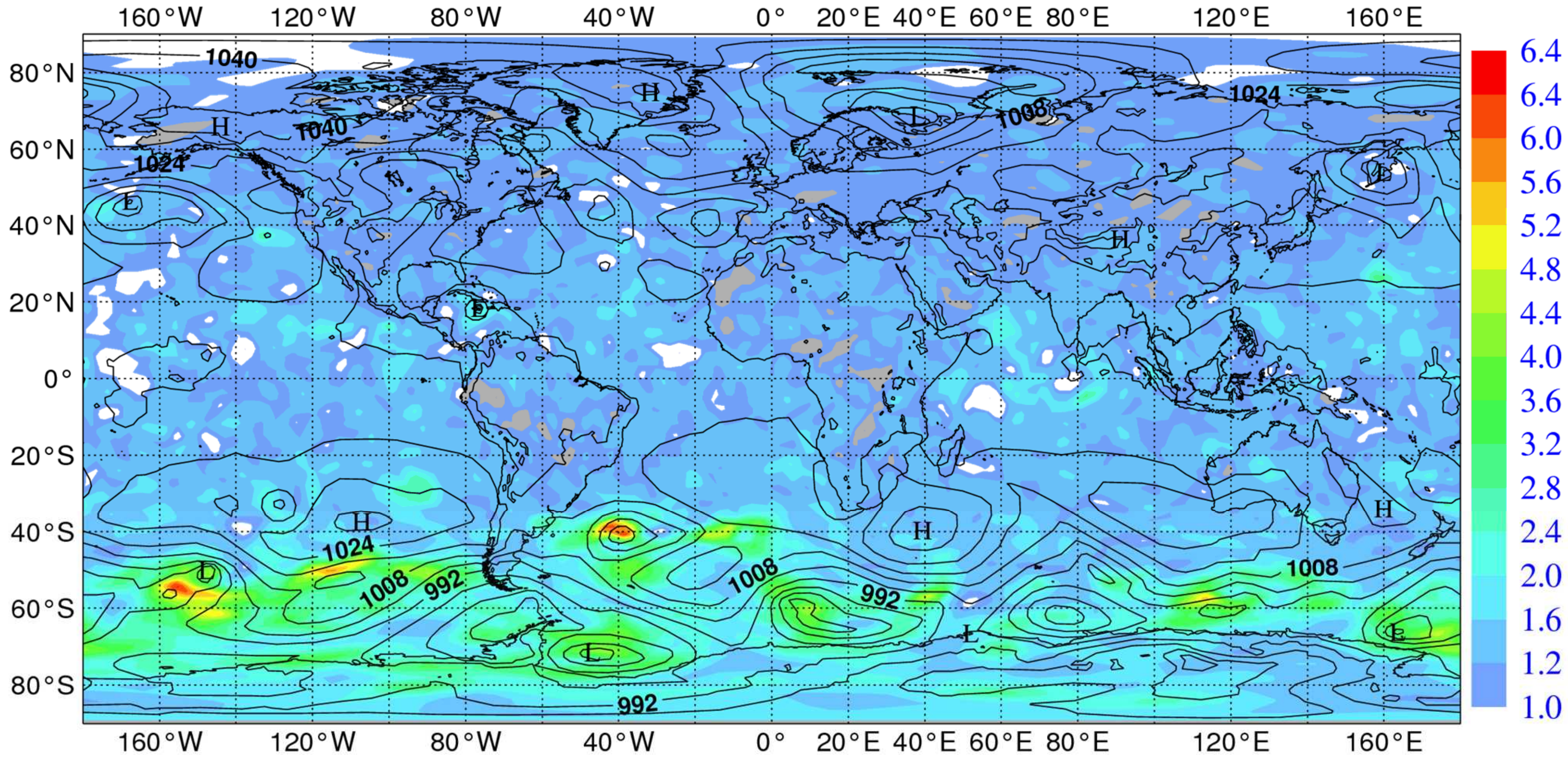
Forecast from HRES assimilation cycle and EDA both without polar orbiters data



Use of EDA variances in 4DVar

EDA without polar orbiters' data has larger spread than operational EDA

EDA spread for u-wind component at 850 hPa: No-Polar/Oper ratio

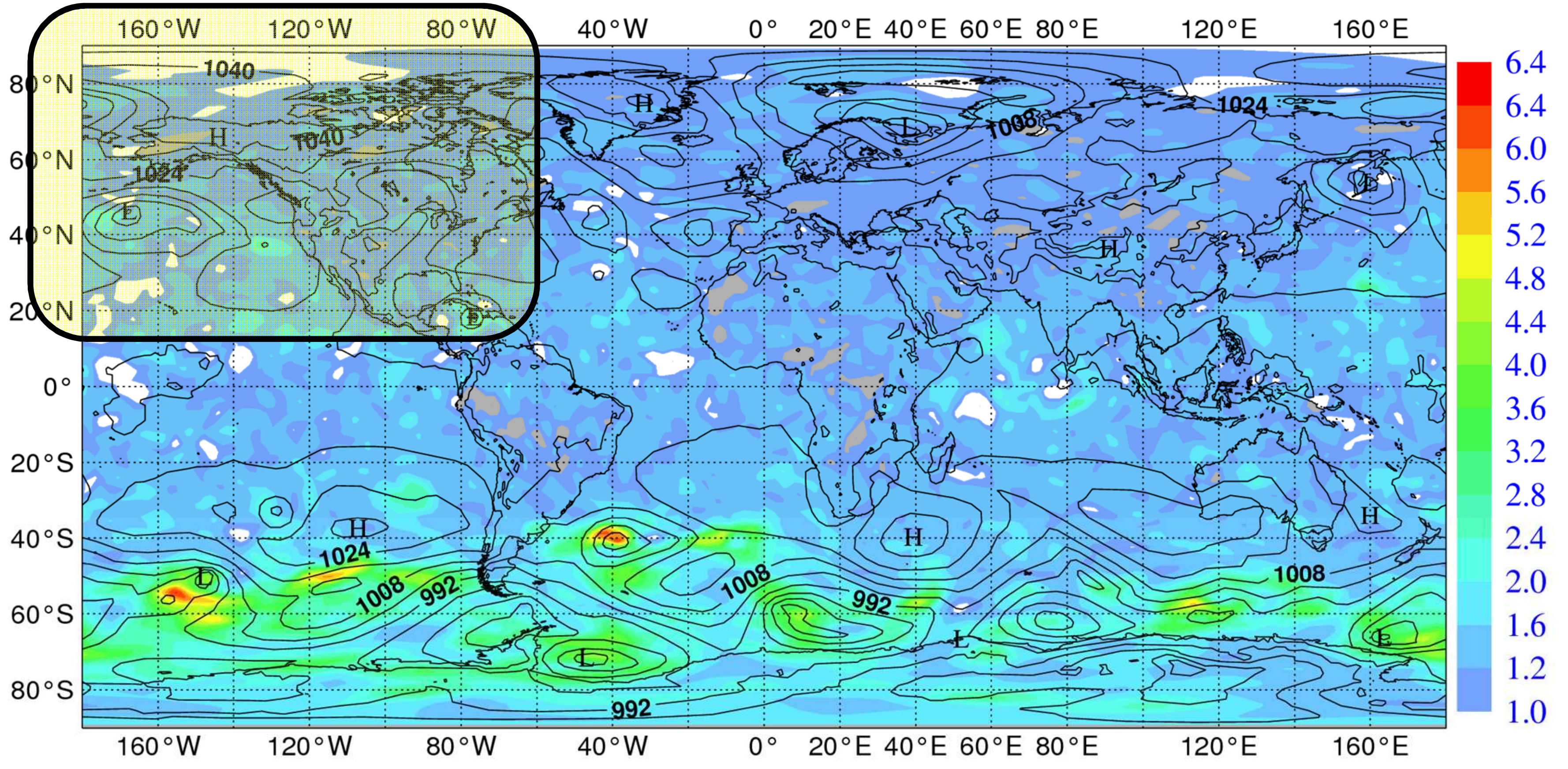


ECMWF

Use of EDA variances in 4DVar

EDA without polar orbiters' data has larger spread than operational EDA

EDA spread for u-wind component at 850 hPa: No-Polar/Oper ratio

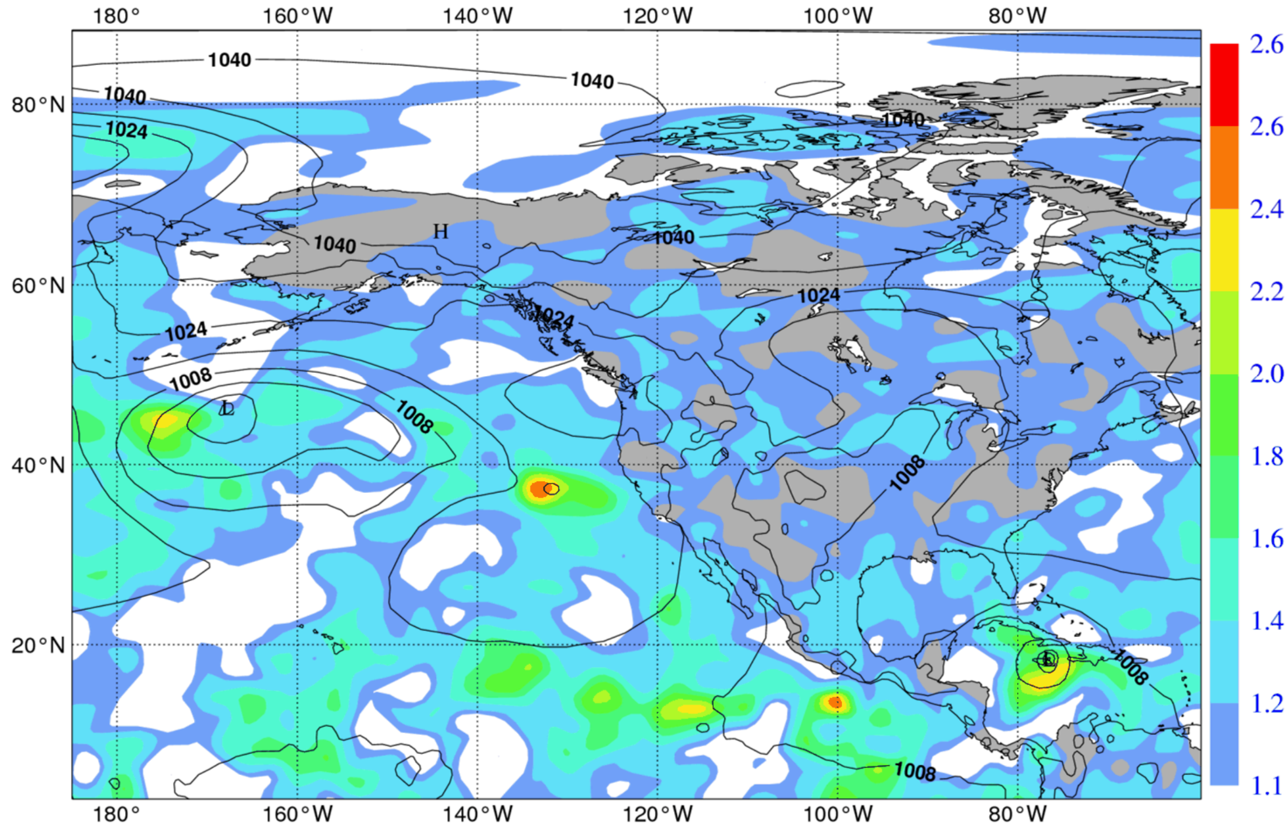


ECMWF

Use of EDA variances in 4DVar

EDA without polar orbiters' data has larger spread than operational EDA

EDA spread for u-wind component at 850 hPa: No-Polar/Oper ratio

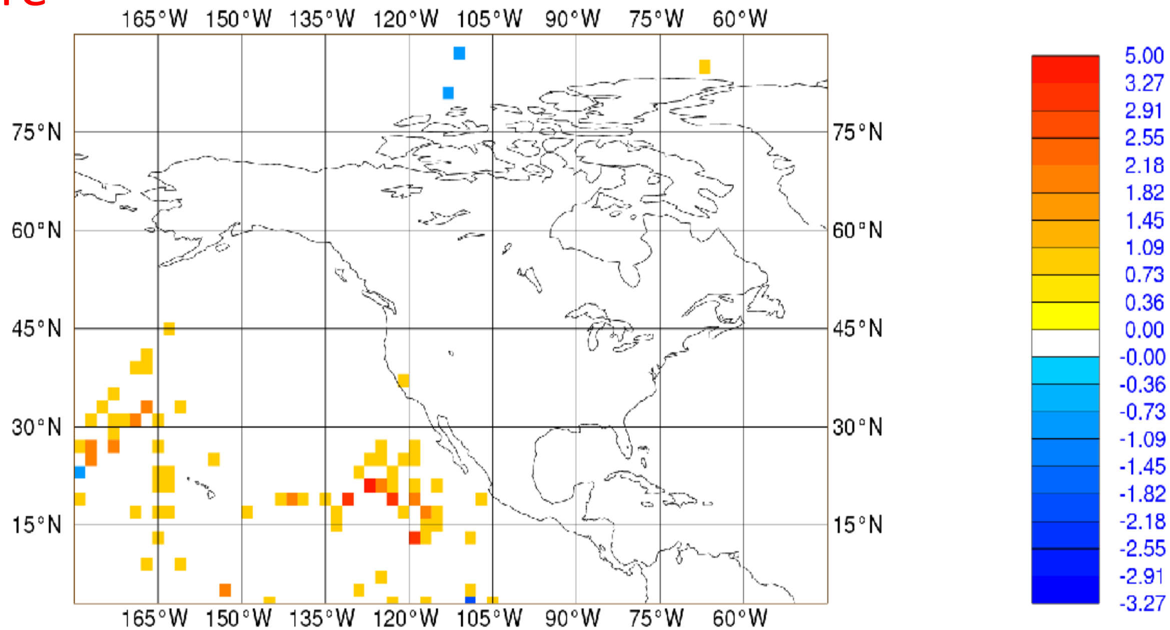


Use of EDA variances in 4DVar

EDA without polar orbiters' data has larger spread than operational EDA

This has two effects: a) Observations are more closely fit and b) More observations pass first guess quality control: $(y-\mathcal{H}(x))^2 \le \alpha(\sigma_b^2 + \sigma_o^2)$

In this case more AMVs from geostationary satellites are assimilated



In this case more AMVs from geostationary satellites are assimilated