TRAINING
COURSE

EUMETSAT/ ECMWF NWP-SAF satellite data assimilation

NWP SAF
Numerical Weather Prediction

ECMWF

Background Errors for Satellite Data assimilation

Data Assimilation …combining background information with observations

- Models give a complete description of the atmospheric, but **errors grow rapidly** in time

- Observations provide an **incomplete description** of the atmospheric state, but bring up to date information

- Data assimilation **combines** these two sources of information to produce an optimal (best) estimate of the atmospheric state

- This state (the *analysis*) is used as **initial conditions** for extended forecasts.

The cost function J (X)

model state

background error covariance

$$J(x) = (x - x_b)^T \mathbf{B}^{-1} (x - x_b) + (y - \mathrm{H}[x])^T \mathbf{R}^{-1} (y - \mathrm{H}[x])$$

observations

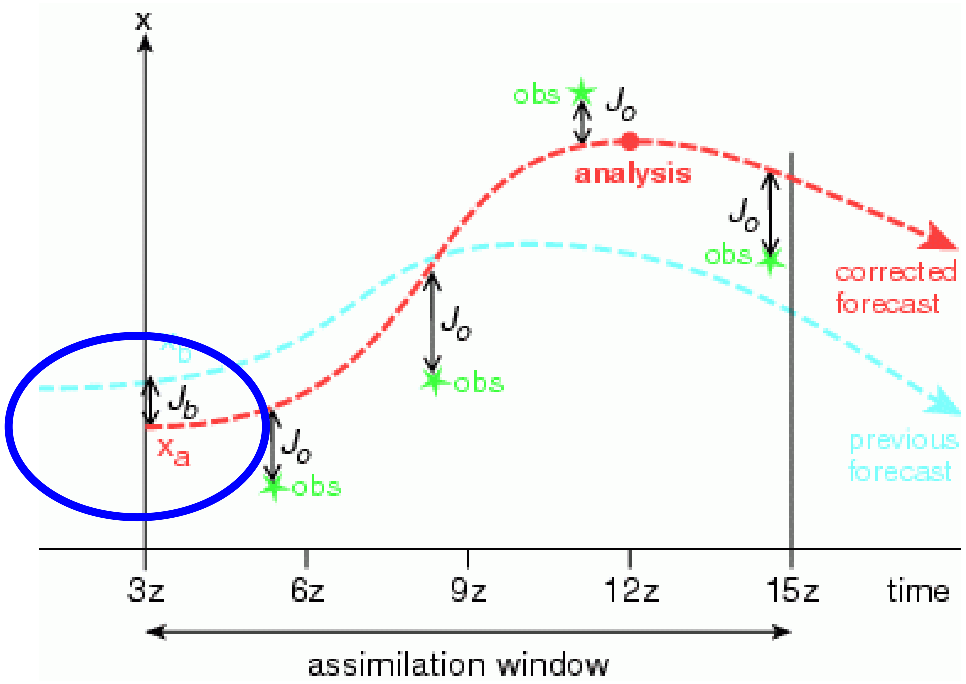
observation* error covariance

observation operator (maps the model state to the observation space)

The 4D-Var Algorithm J_b

$$J(x) = \boxed{(x - x_b)^T \mathbf{B}^{-1} (x - x_b)} +$$

$$(y - \mathrm{H}[x])^T \mathbf{R}^{-1} (y - \mathrm{H}[x])$$



What do we want our background errors to do ?

What do we want our background errors to do ?

- Describe our <u>confidence in the background</u> estimate of the atmosphere X_b

- Describe how background errors are <u>correlated</u> with each other:
  - vertically (between different model levels)
  - spatially (between different grid points)
  - between variables (T / Q / O3 / wind)
  - impose balance (e.g. geostrophic)

- **They should be data and flow dependent!**

How do we determine background errors?

Simply compare X_b to
the true state of the atmosphere ?

…but we don’t have the truth!

How do we <u>estimate</u> background errors?

How do we estimate background errors?

• Innovation departure statistics – i.e. comparison of X_b with **radiosondes** (best estimate of truth but limited coverage)

•comparison of forecasts differences e.g. 48hr and 24hr (so called ***NMC method***)

•comparison of **ensembles** of analyses made using perturbed observations

None of these approaches are perfect!

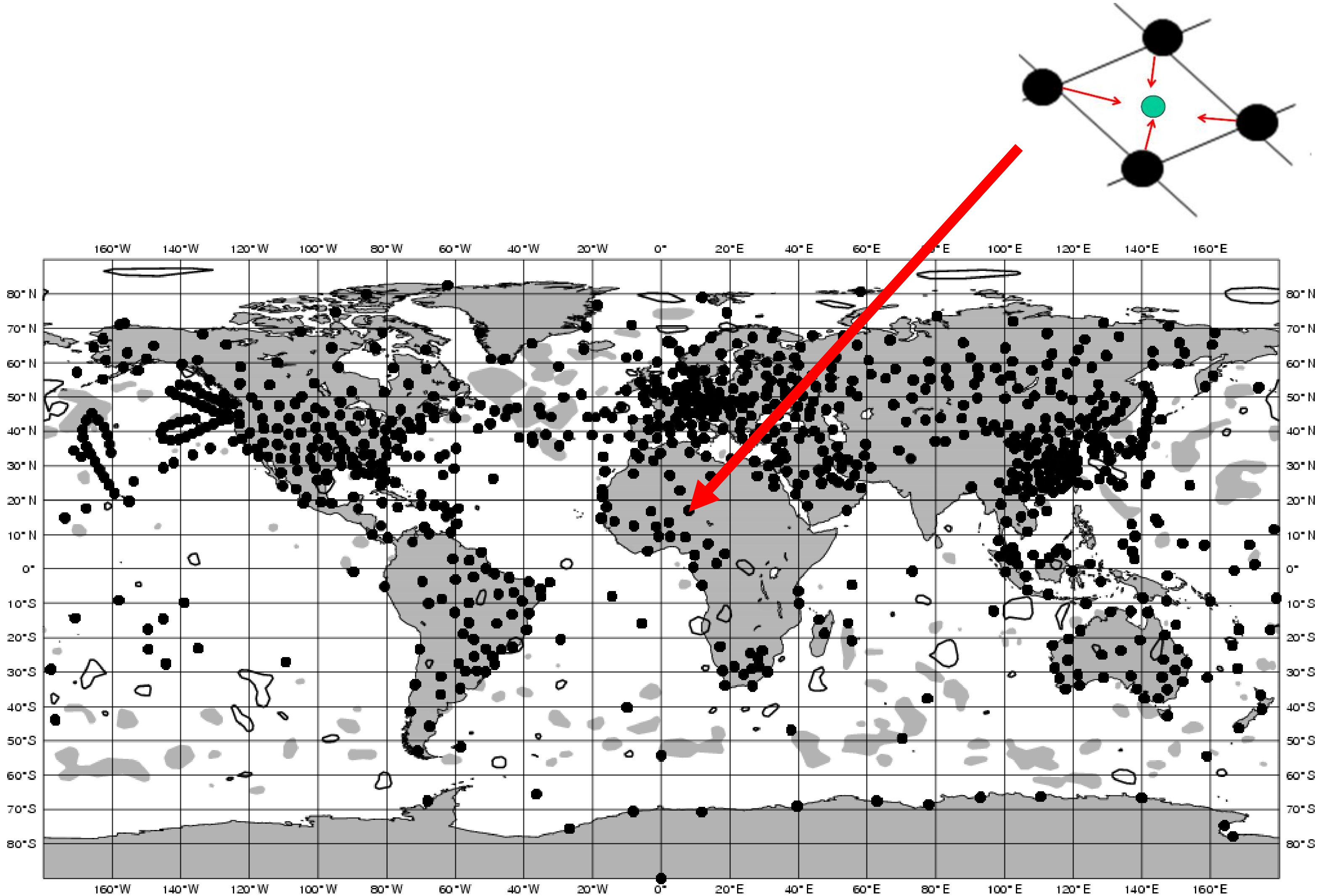
How do we estimate background errors?

- **Innovation statistics:**
  - ☺ The only direct method for diagnosing background error statistics.
  - ☹ Provides statistics of background error in observation space.
  - ☹ Statistics are not global, and do not cover all model levels.
  - ☹ Requires a good uniform observing network.
  - ☹ Statistics are biased towards data-dense areas.
- **Forecast Differences:**
  - ☺ Generates global statistics of model variables at all levels.
  - ☺ Inexpensive.
  - ☹ Statistics are a mixture of analysis and background error.
  - ☹ Not good in data-sparse regions.
- **Ensembles of Analyses:**
  - ☹ Assumes statistics of observation error (and SST, etc.) are well known.
  - ☺ Diagnoses the statistics of the actual analysis system.
  - ☹ Danger of feedback. (Noisy analysis system => noisy stats => noisier system.)

How do we estimate background errors?

- **Innovation statistics:**
  - ☺ The only direct method for diagnosing background error statistics.
  - ☹ Provides statistics of background error in observation space.
  - ☹ Statistics are not global, and do not cover all model levels.
  - ☹ Requires a good uniform observing network.
  - ☹ Statistics are biased towards data-dense areas.
- **Forecast Differences:**
  - ☺ Generates global statistics of model variables at all levels.
  - ☺ Inexpensive.
  - ☹ Statistics are a mixture of analysis and background error.
  - ☹ Not good in data-sparse regions.
- **Ensembles of Analyses:**
  - ☹ Assumes statistics of observation error (and SST, etc.) are well known.
  - ☺ Diagnoses the statistics of the actual analysis system.
  - ☹ Danger of feedback. (Noisy analysis system => noisy stats => noisier system.)

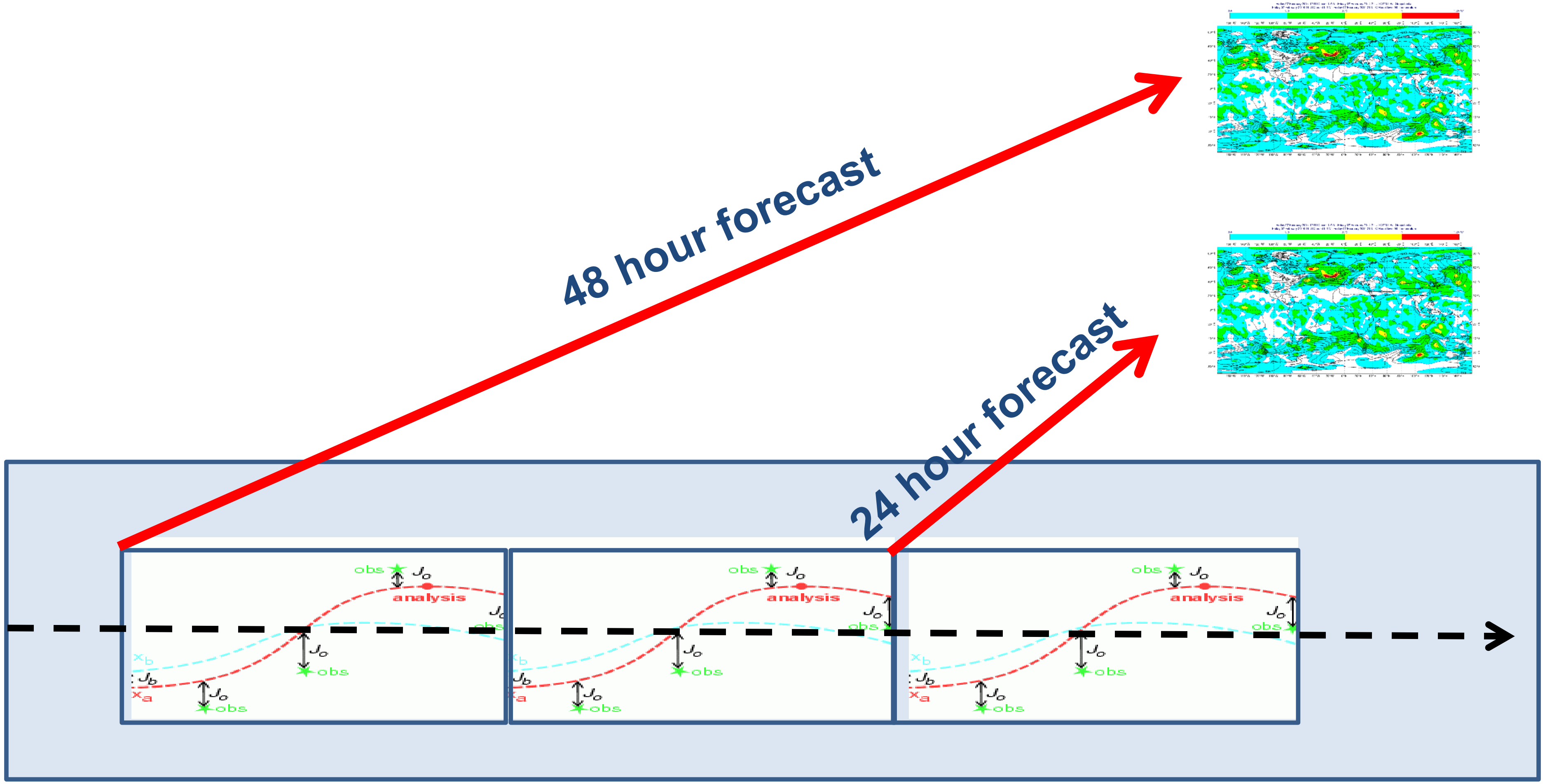
Innovation statistics …
e.g. compare X_b to radiosondes



How do we estimate background errors?

- **Innovation statistics:**
  - ☺ The only direct method for diagnosing background error statistics.
  - ☹ Provides statistics of background error in observation space.
  - ☹ Statistics are not global, and do not cover all model levels.
  - ☹ Requires a good uniform observing network.
  - ☹ Statistics are biased towards data-dense areas.
- **Forecast Differences:**
  - ☺ Generates global statistics of model variables at all levels.
  - ☺ Inexpensive.
  - ☹ Statistics are a mixture of analysis and background error.
  - ☹ Not good in data-sparse regions.
- **Ensembles of Analyses:**
  - ☹ Assumes statistics of observation error (and SST, etc.) are well known.
  - ☺ Diagnoses the statistics of the actual analysis system.
  - ☹ Danger of feedback. (Noisy analysis system => noisy stats => noisier system.)

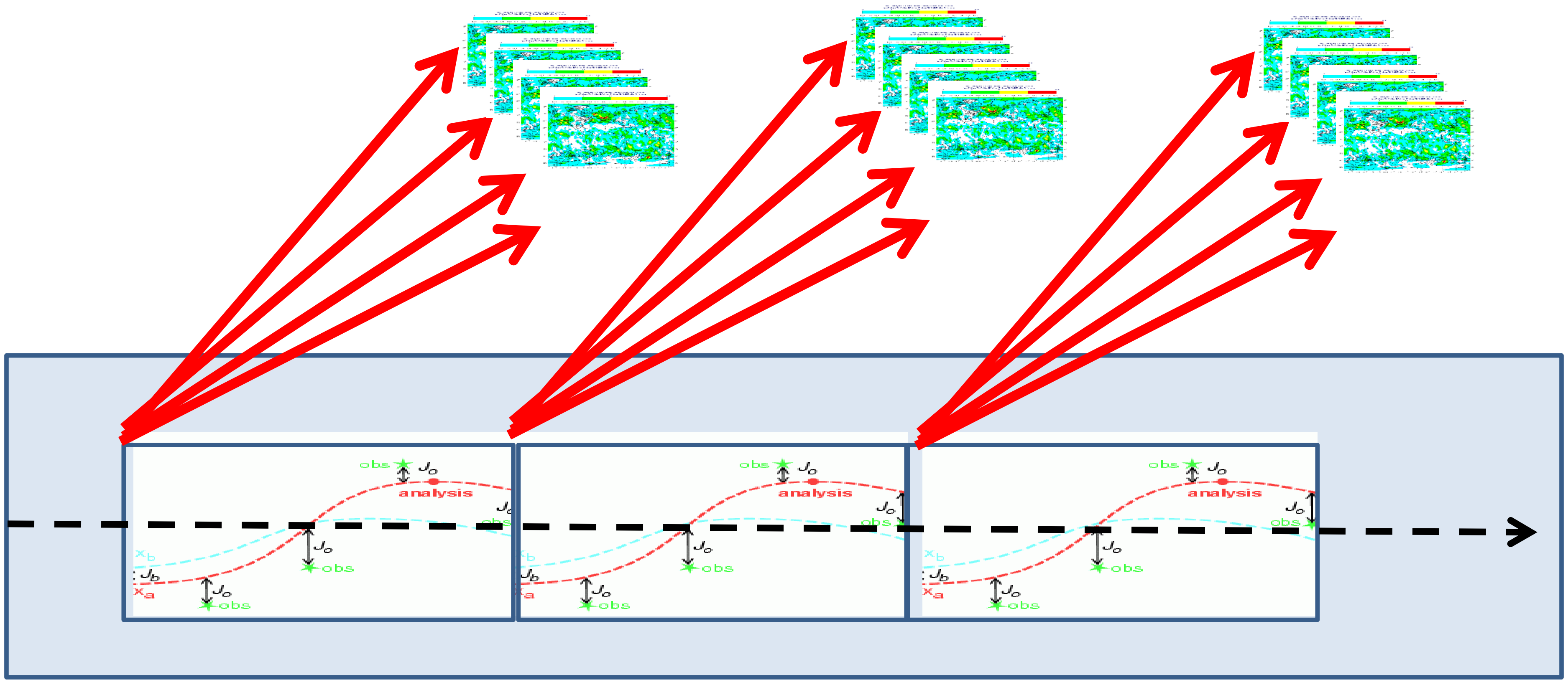
Compare forecasts of the same state from different ranges



How do we estimate background errors?

- **Innovation statistics:**
  - ☺ The only direct method for diagnosing background error statistics.
  - ☹ Provides statistics of background error in observation space.
  - ☹ Statistics are not global, and do not cover all model levels.
  - ☹ Requires a good uniform observing network.
  - ☹ Statistics are biased towards data-dense areas.
- **Forecast Differences:**
  - ☺ Generates global statistics of model variables at all levels.
  - ☺ Inexpensive.
  - ☹ Statistics are a mixture of analysis and background error.
  - ☹ Not good in data-sparse regions.
- **Ensembles of Analyses:**
  - ☹ Assumes statistics of observation error (and SST, etc.) are well known.
  - ☺ Diagnoses the statistics of the actual analysis system.
  - ☹ Danger of feedback. (Noisy analysis system => noisy stats => noisier system.)

Perturb observations and physics to produce an ensemble of analyses



How do we estimate background errors?

- Innovation statistics:
  - ☺ The only direct method for dia
  - ☹ Provides statistics of
  - ☹ Statistics
  - ☹

ound error.

- E yses:
  - ☹ statistics of observation error (and SST, etc.) are well known.
  - ☺ Diagnoses the statistics of the actual analysis system.
  - ☹ Danger of feedback. (Noisy analysis system => noisy stats => noisier system.)

But! It is generally the case that where accurate background errors are most important we have the least information to estimate them!!

What do background errors look like?

…or at least what do we <u>think</u> they look like ?

What do background errors look like?

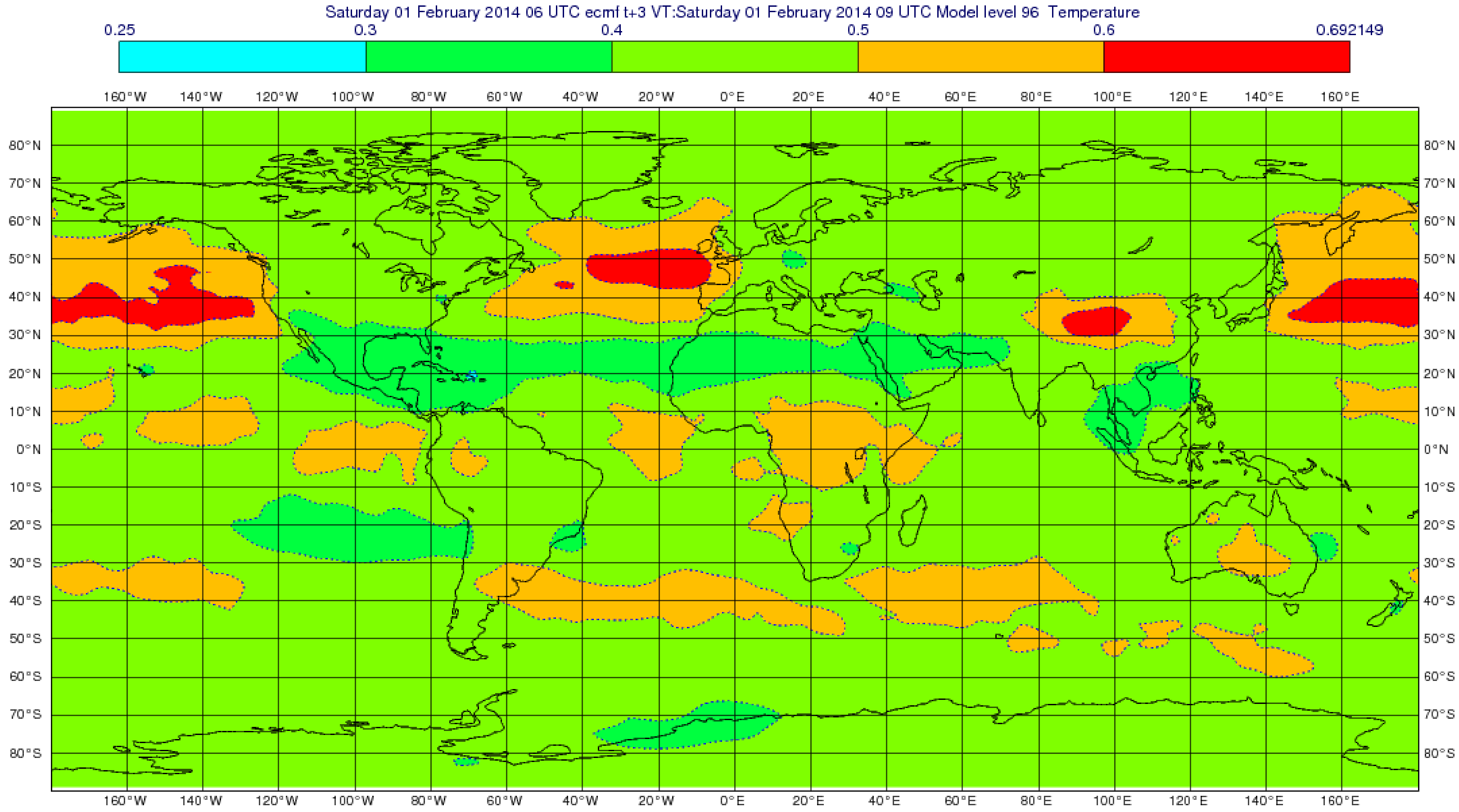
- Error magnitude (variances)
- Spatial error correlations
- Vertical error correlations

What do background errors look like?

- **Error magnitude (variances)**
- Spatial error correlations
- Vertical error correlations

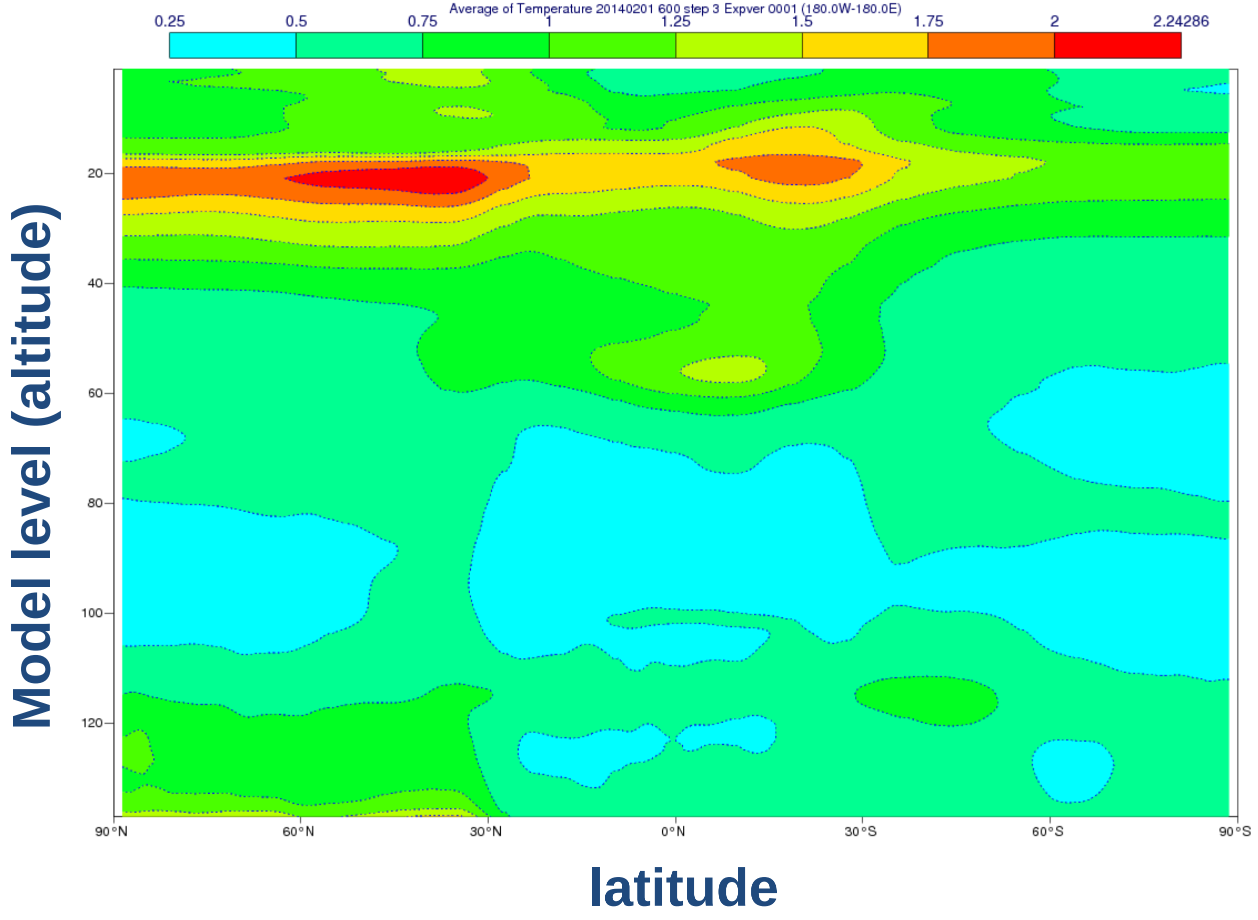
Magnitude of Background Errors

Monthly average of standard deviation of temperature error at 500hPa



Magnitude of Background Errors

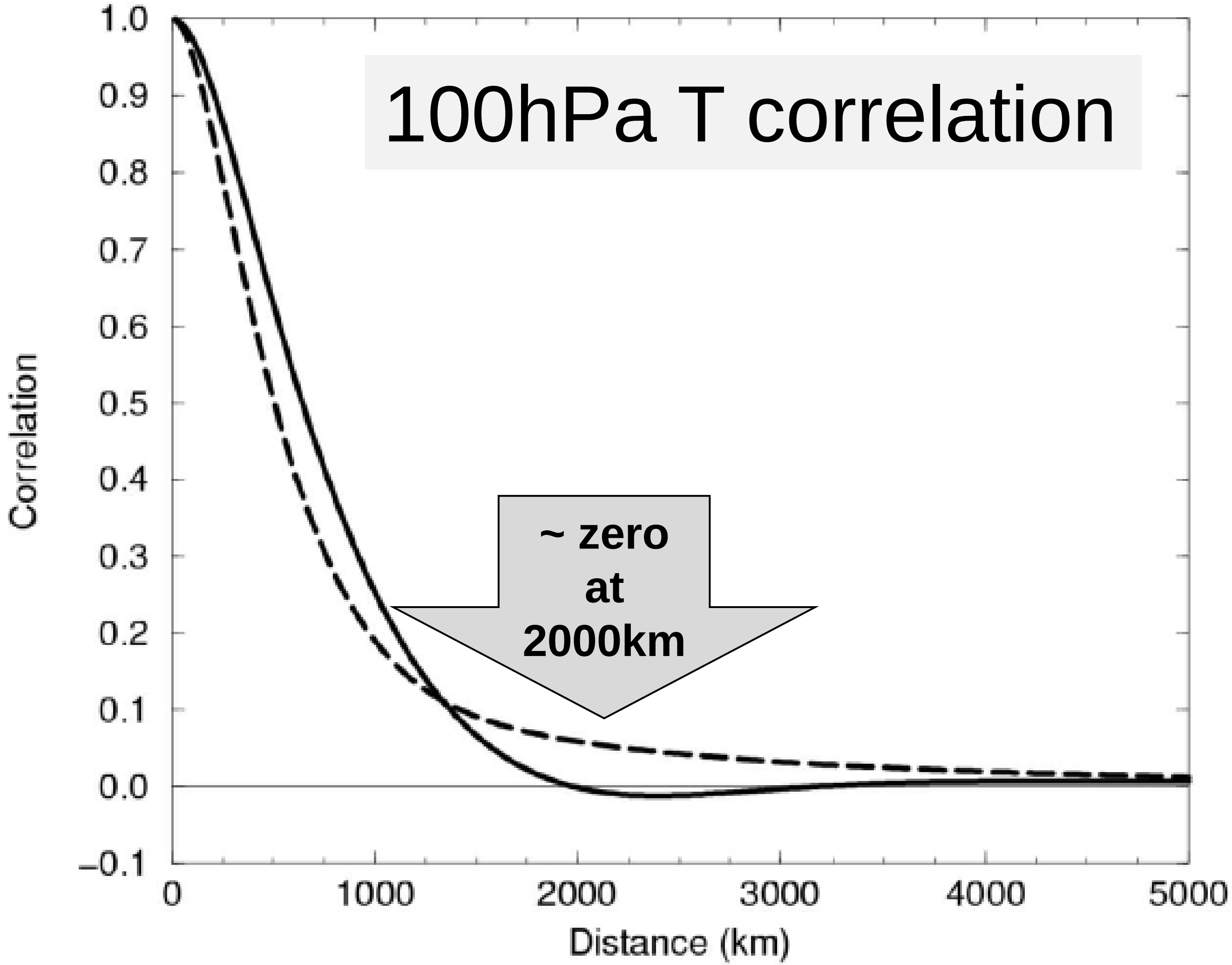
Monthly zonal average of standard deviation of temperature error at 500hPa



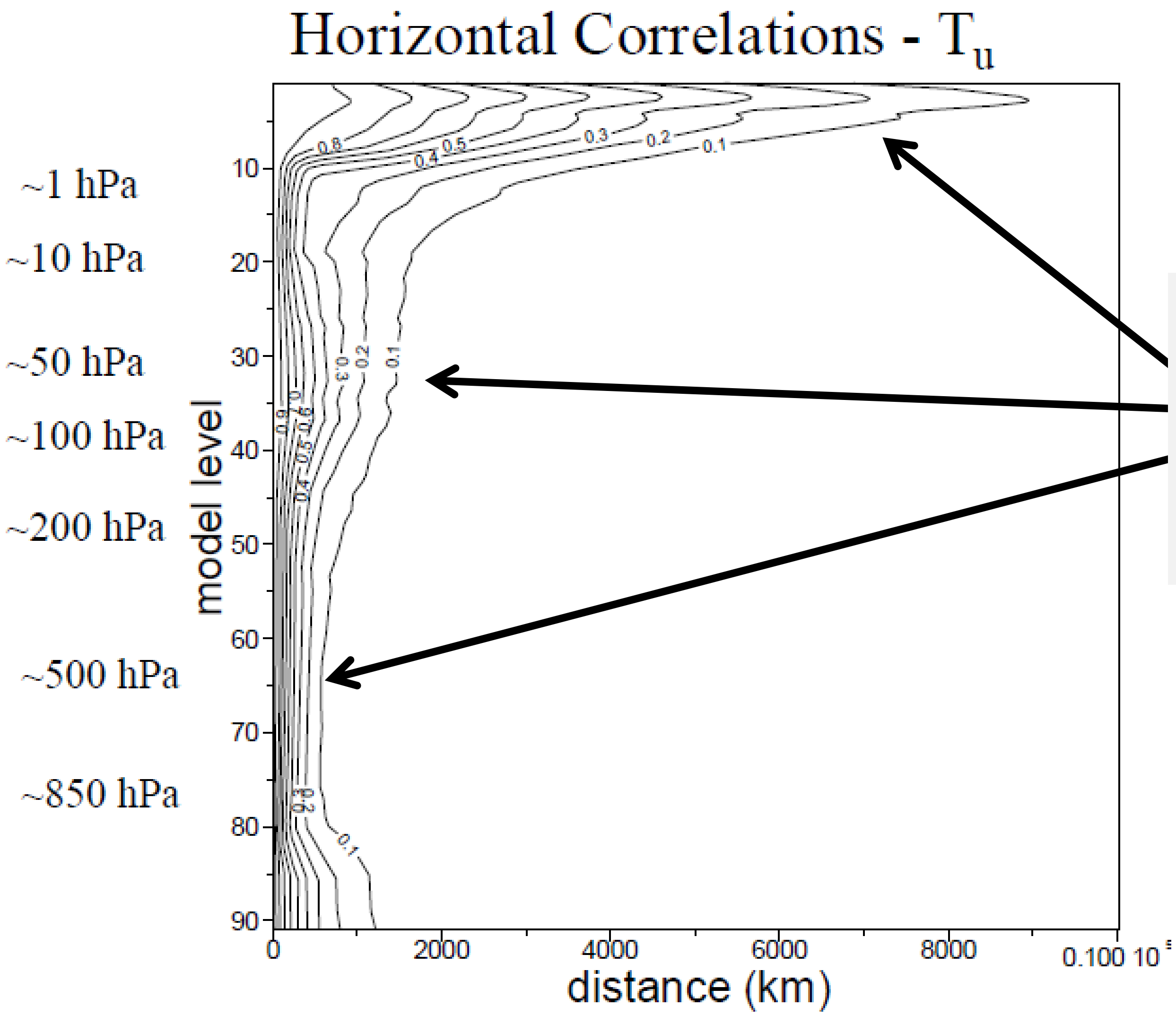
What do background errors look like?

- Error magnitude (variances)
- **Spatial error correlations**
- Vertical error correlations

Spatial Error Correlations



Spatial Error Correlations



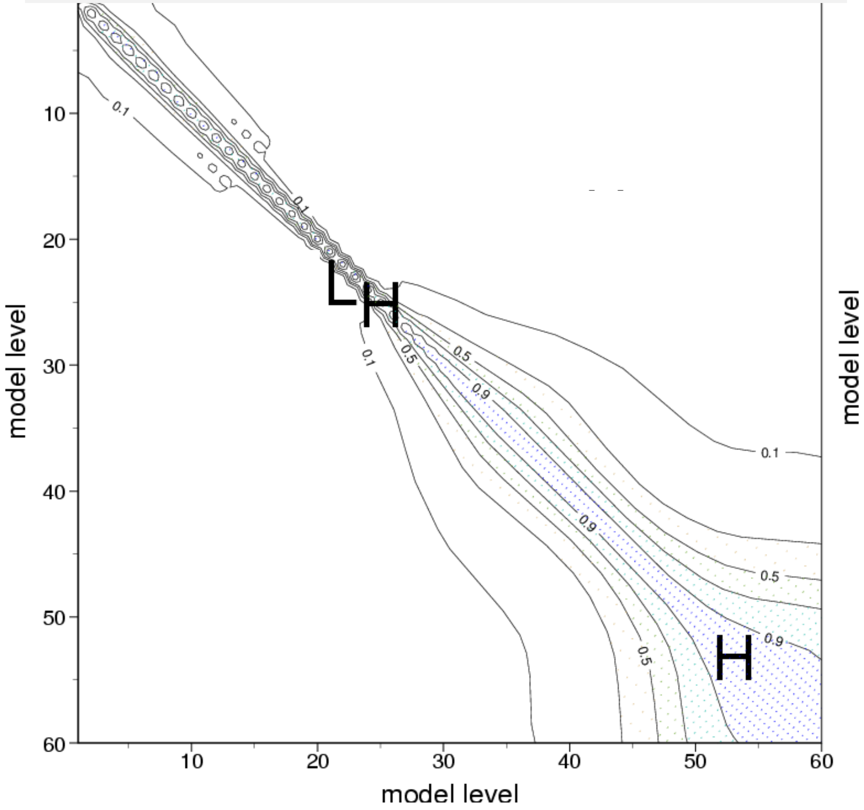
Spatial correlation lengths vary significantly with altitude

What do background errors look like?

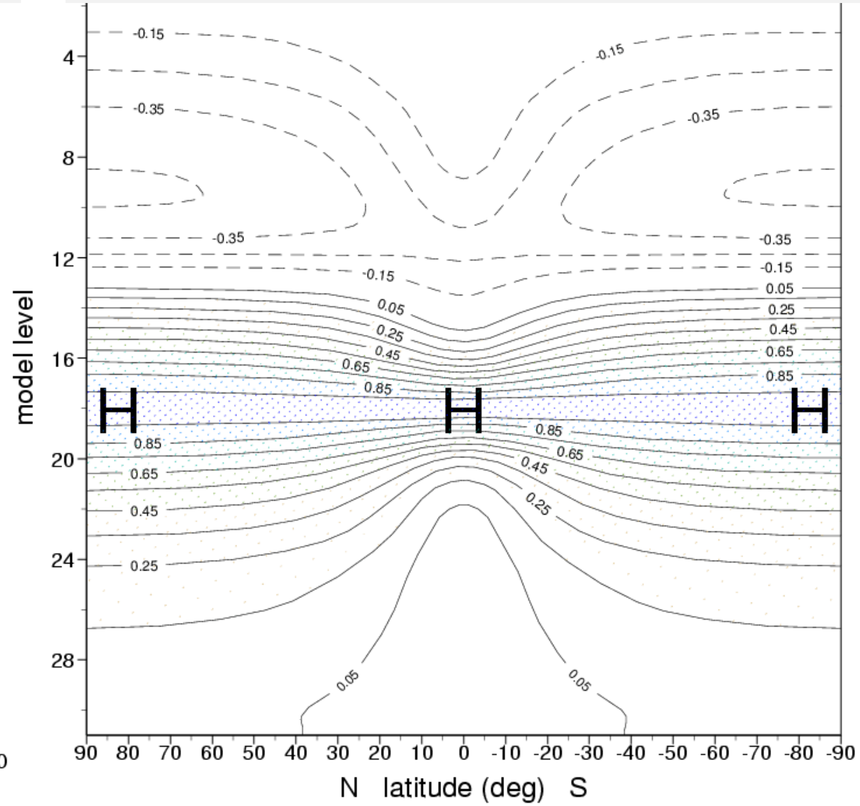
- Error magnitude (variances)
- Spatial error correlations
- **Vertical error correlations**

Vertical (inter-level) error correlations

Inter-level error vorticity correlations



Cross section of T error correlations



Background error dependencies

Background error dependencies

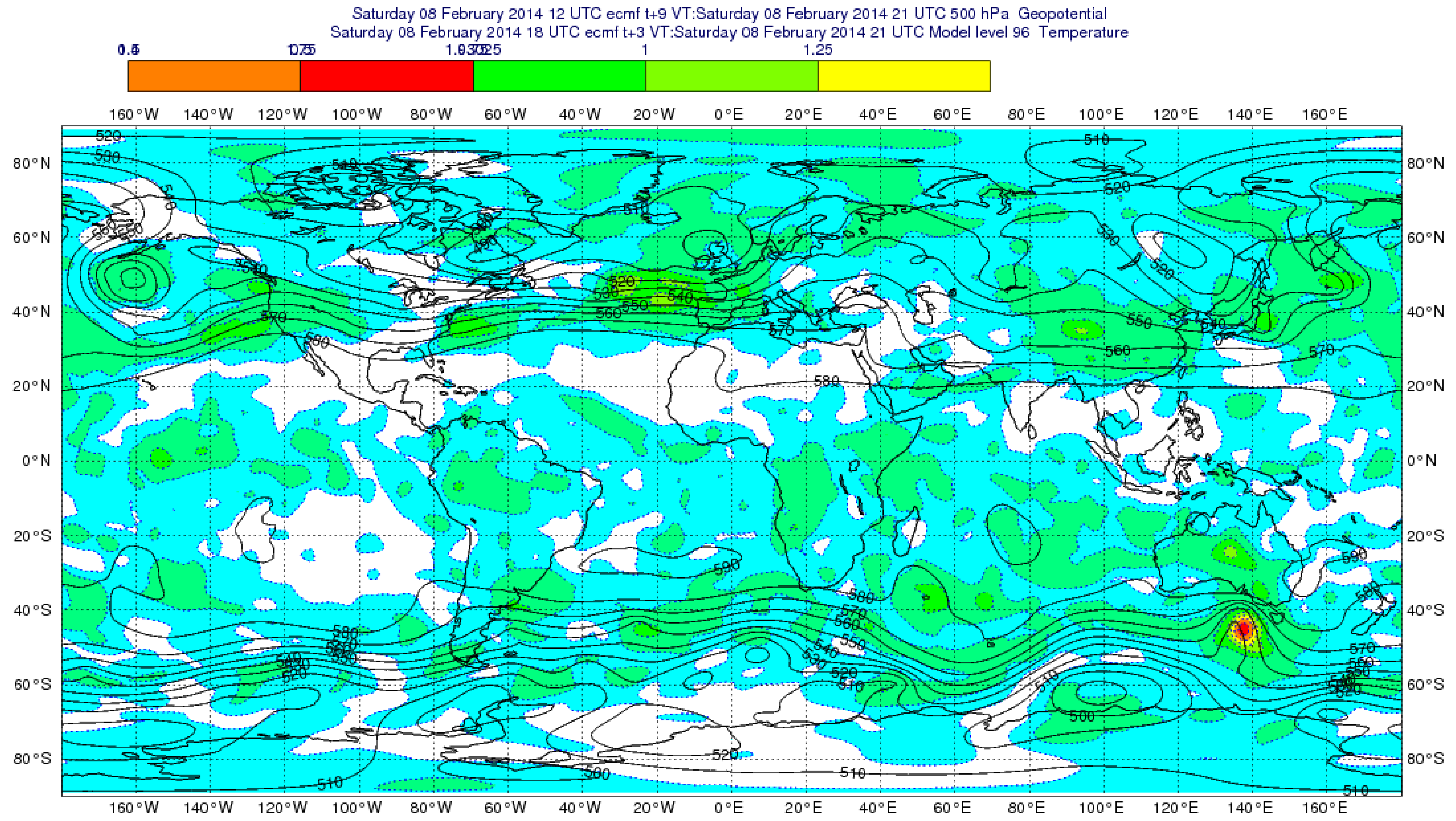
- **Flow / regime dependence**
- **Observation dependence**
- **Method dependence**

Background error dependencies

- **Flow / regime dependence**
- **Observation dependence**
- **Method dependence**

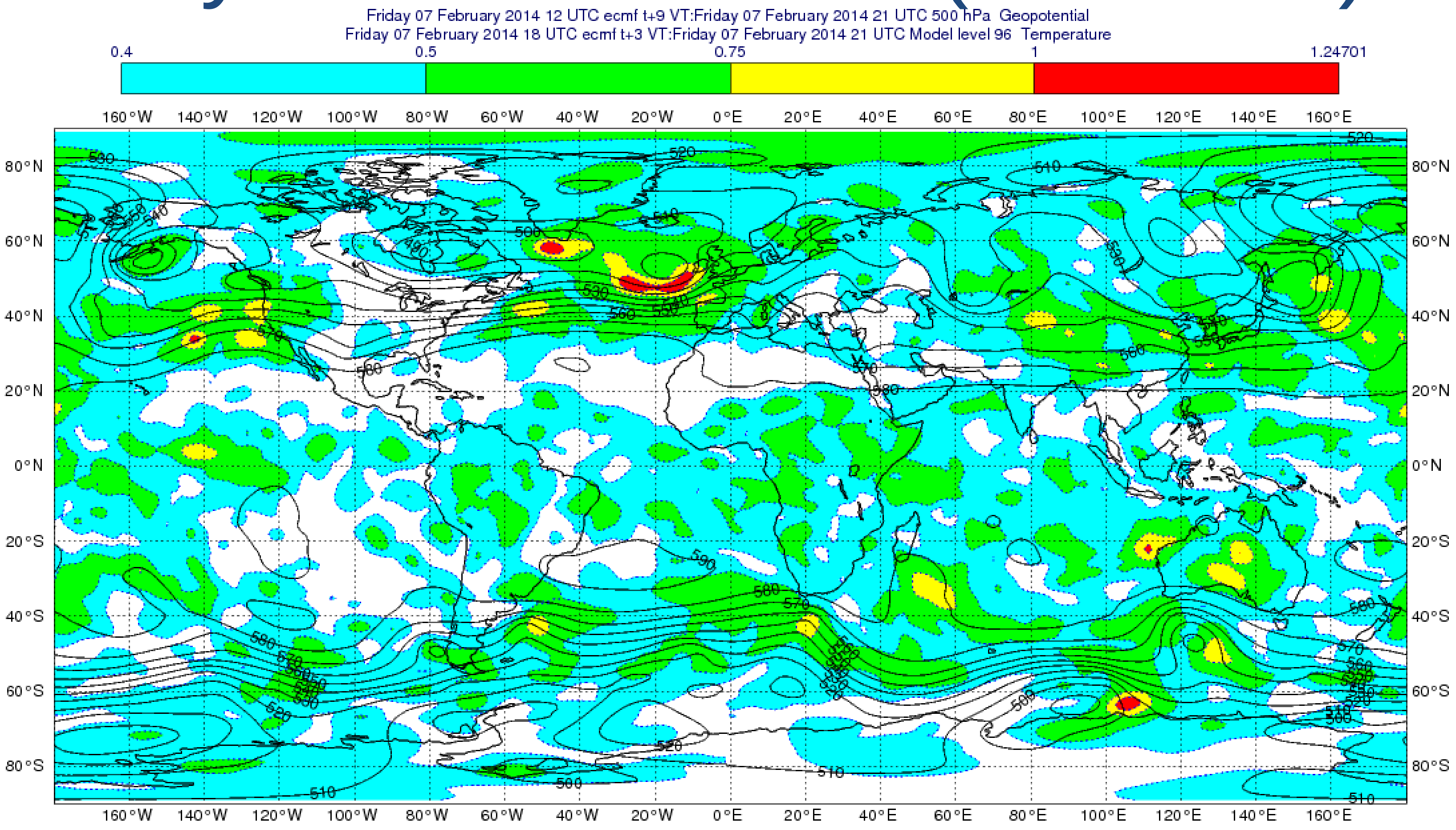
Flow / regime dependent background errors

Daily 500hPa T error (8/2/2014)



Flow / regime dependent background errors

Daily 500hPa T error (7/2/2014)



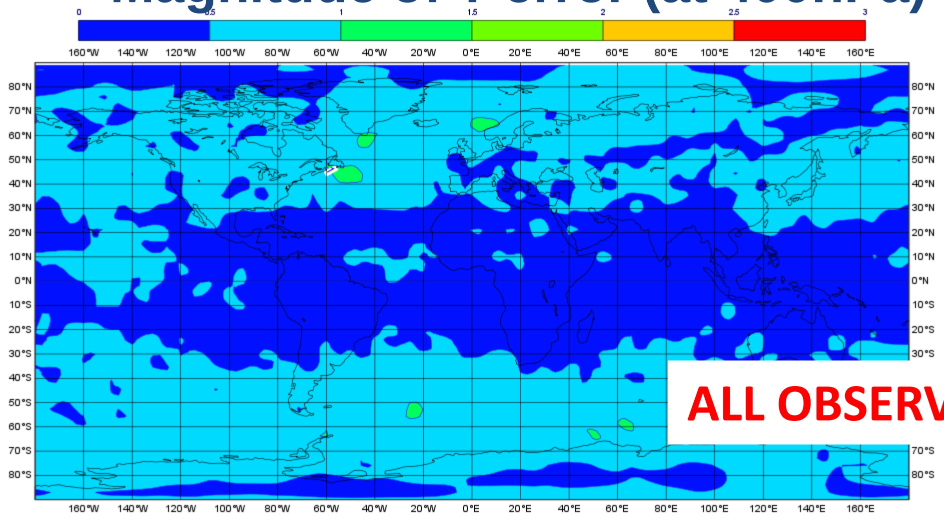
Background error dependencies

- Flow / regime dependence
- **Observation dependence**
- Method dependence

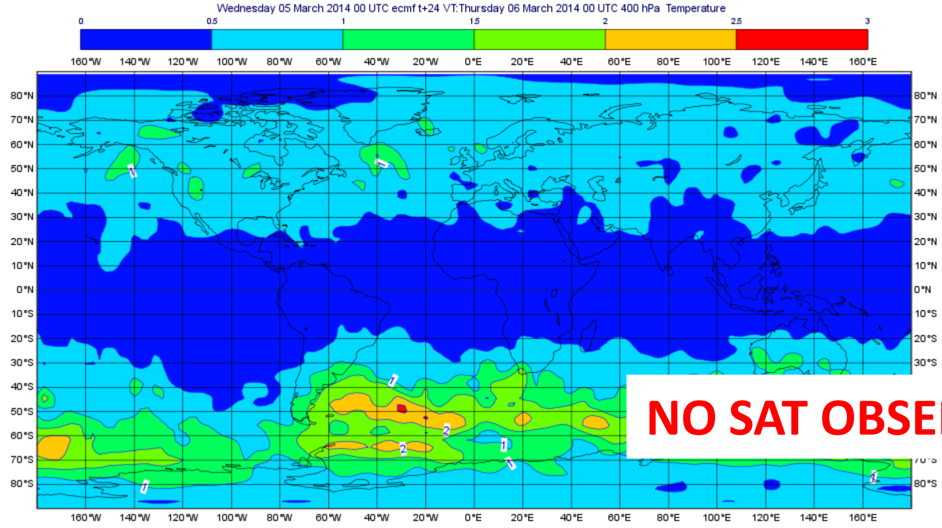
Background errors depend on which observations are assimilated in the system

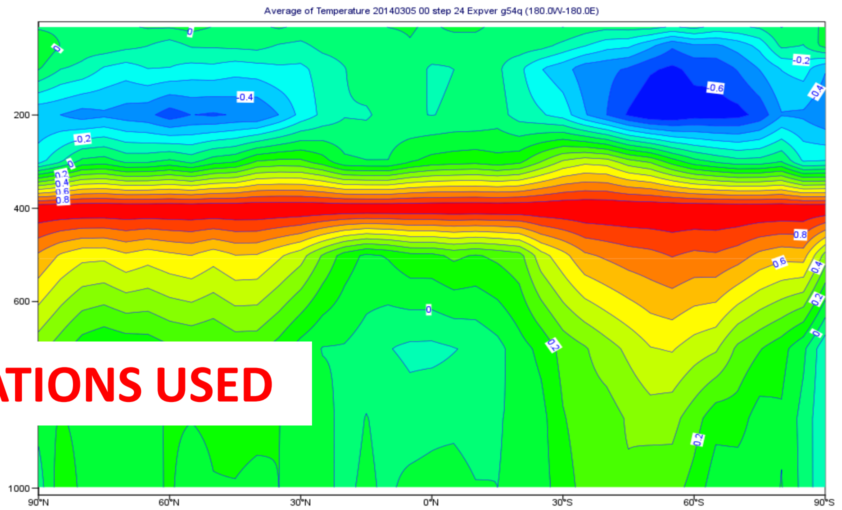
Magnitude of T error (at 400hPa)

Correlation of T error (at 400hPa)



ALL OBSERVATIONS USED





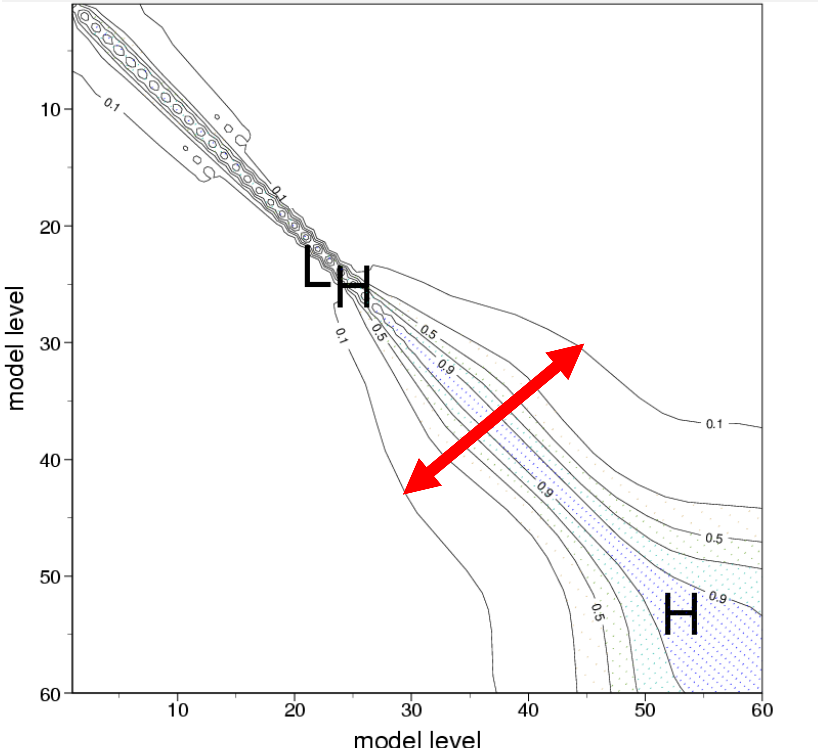
NO SAT OBSERVATIONS USED

Background error dependencies

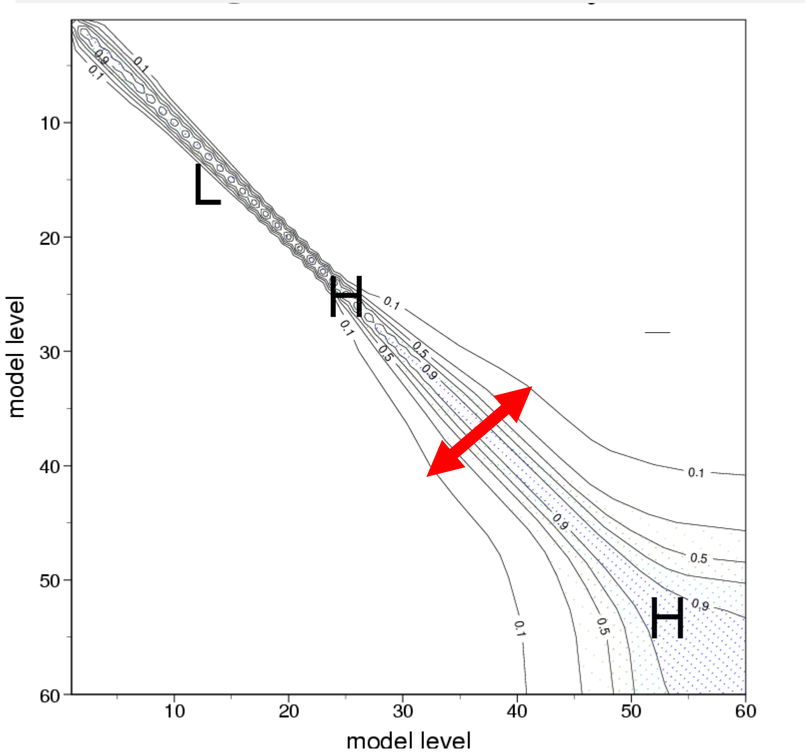
- Flow / regime dependence
- Observation dependence
- **Method dependence**

The background error estimates depends on the method you use!

Forecast Differences (NMC method)



Analysis Ensembles (EDA method)



Background errors and radiance assimilation

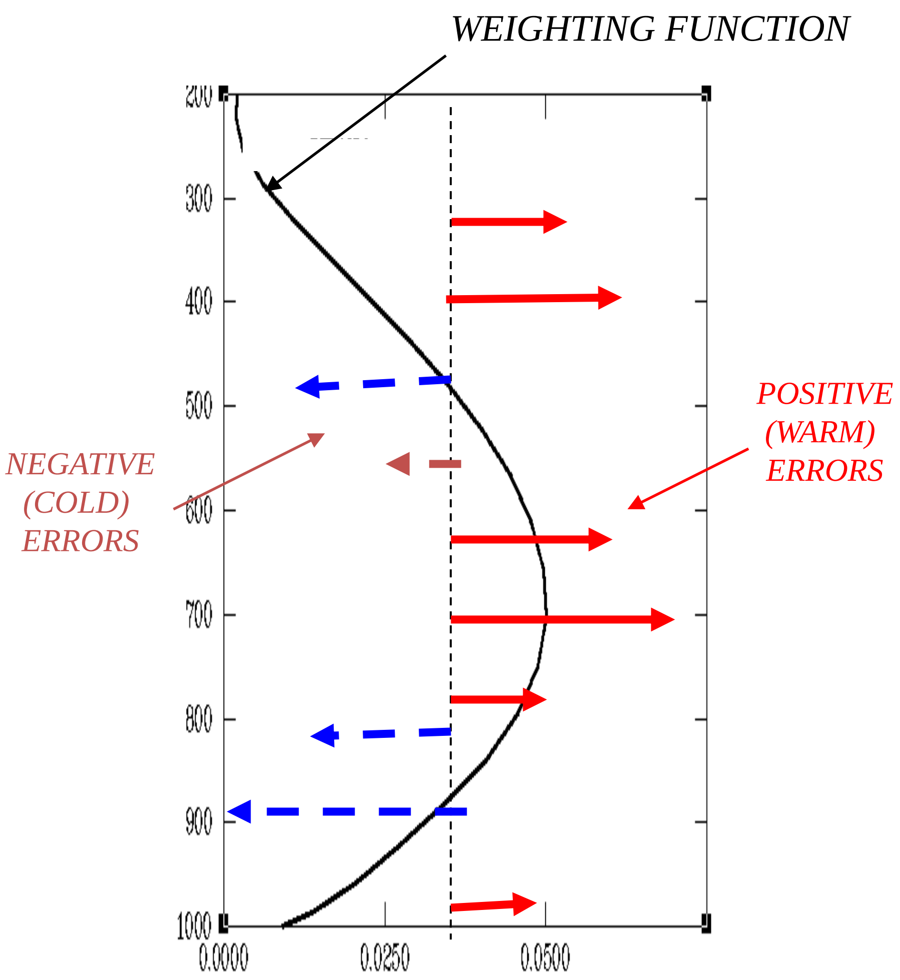
Background errors and radiance assimilation

- The physics of radiative transfer mean that radiances measured by downward looking satellite sounders have very **poor vertical resolution** (they are broad vertical averages)

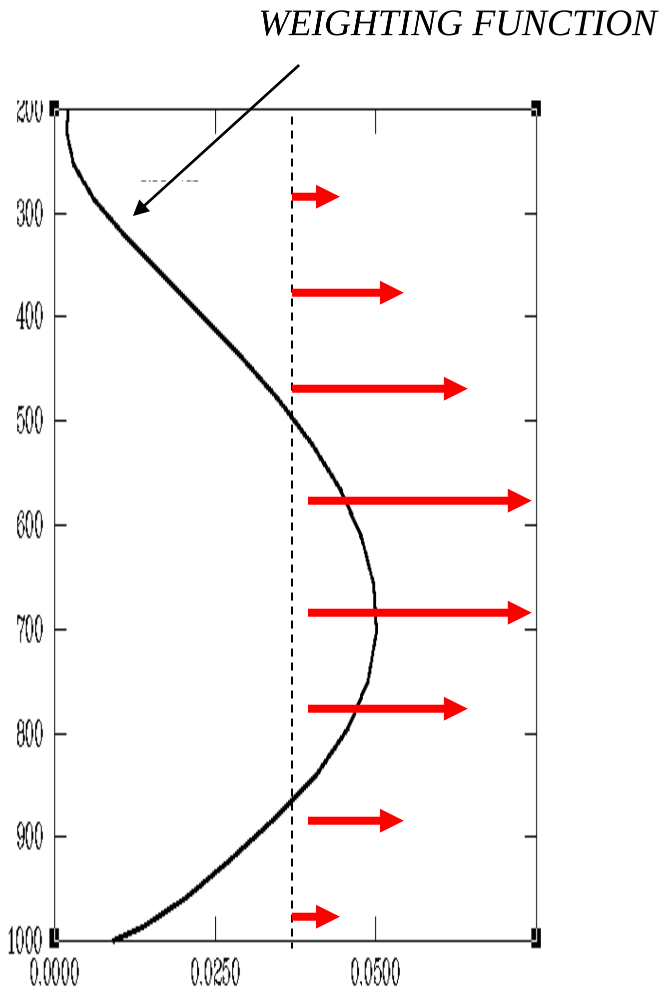
- If we wish to **correct errors in the background** with radiance observations (in DA) the vertical structure of these errors is very important.

- This structure is described by the **vertical correlations in the background error covariance**

Background errors (and vertical resolution)



"Difficult" to correct



"Easy" to correct

Can we quantify the impact of vertical background error correlations on analysis accuracy ?

…a helpful linear analogue …

model state

background error covariance

$$J(x) = (x - x_b)^T \mathbf{B}^{-1} (x - x_b) +$$

$$(y - \mathrm{H}[x])^T \mathbf{R}^{-1} (y - \mathrm{H}[x])$$

observations

observation* error covariance

observation operator
(maps the model state to the observation space)

…when we minimise J(x) …

...we correct background errors

It can be shown that the state that minimizes the cost function is equivalent to a linear **correction** of the background using the observations:

$$x_a = x_b + [\mathbf{HB}]^T[\mathbf{HBH}^T + \mathbf{R}]^{-1}(y - \mathbf{H}x_b)$$

…where the **correction** is the **Kalman Gain Matrix** multiplied by the **innovation vector** (observation minus radiances simulated from the background)

correction term = $\underbrace{\mathbf{HB}]^T[\mathbf{HBH}^T + \mathbf{R}]^{-1}}_{\text{Kalman gain}} \times \underbrace{(y - \mathbf{H}x_b)}_{\text{innovation}}$

…and reduce the error …

Furthermore when we apply this **correction** we produce a state (the analysis) that is more accurate than the background. We can compute the improvement as an **error reduction** of the analysis error (**A**) compared to the error in the background (**B**) …

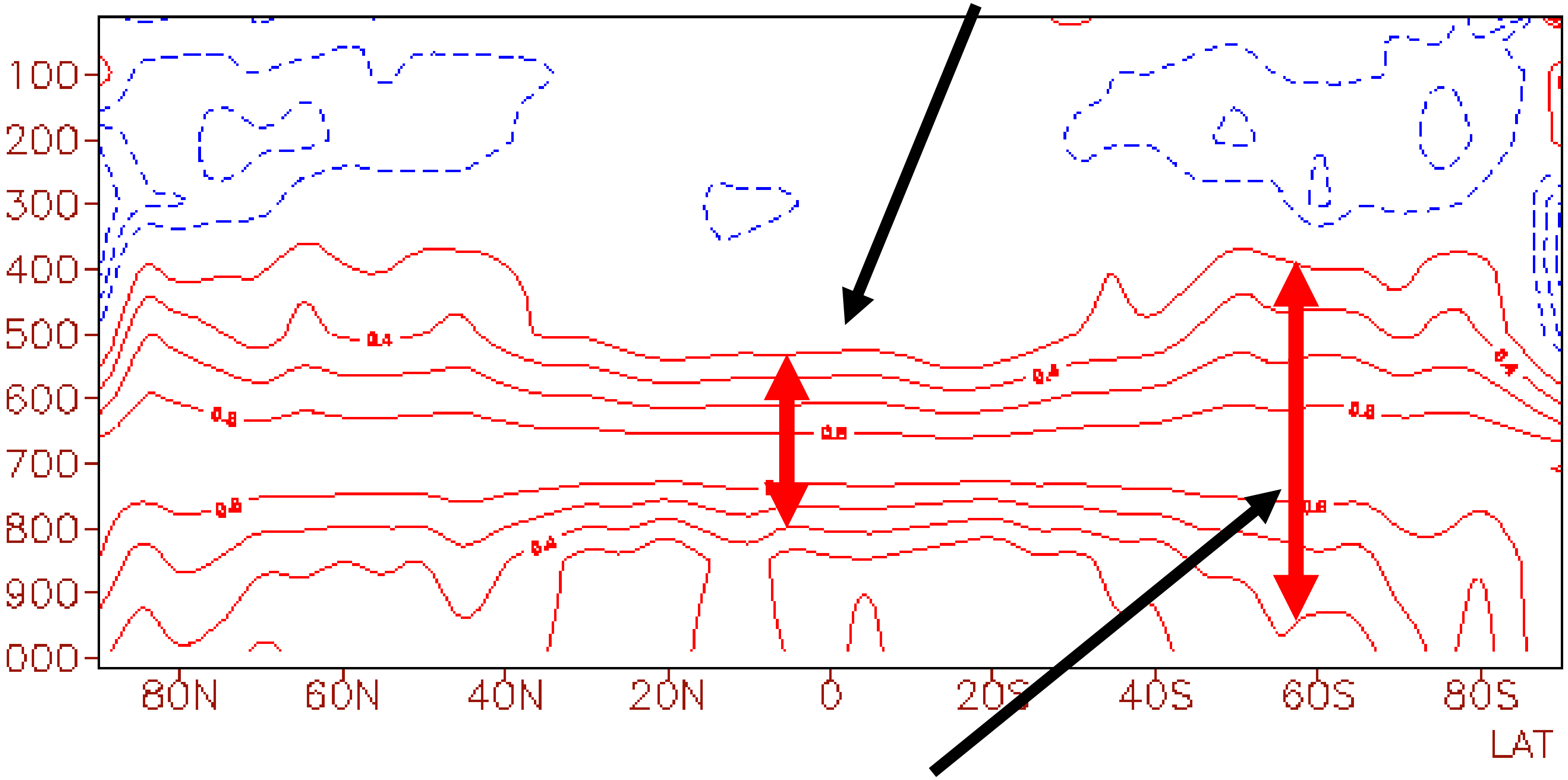
$$\mathbf{A} = \mathbf{B} - \underbrace{[\mathbf{HB}]^T[\mathbf{HBH}^T + \mathbf{R}]^{-1}\,\mathbf{HB}}_{\text{error reduction}}$$

So we can look at how **A** performs for two different types of vertical correlation present in **B** (for example using **8461 IASI channels** correct background errors in the analysis)

Sharp and broad vertical correlation

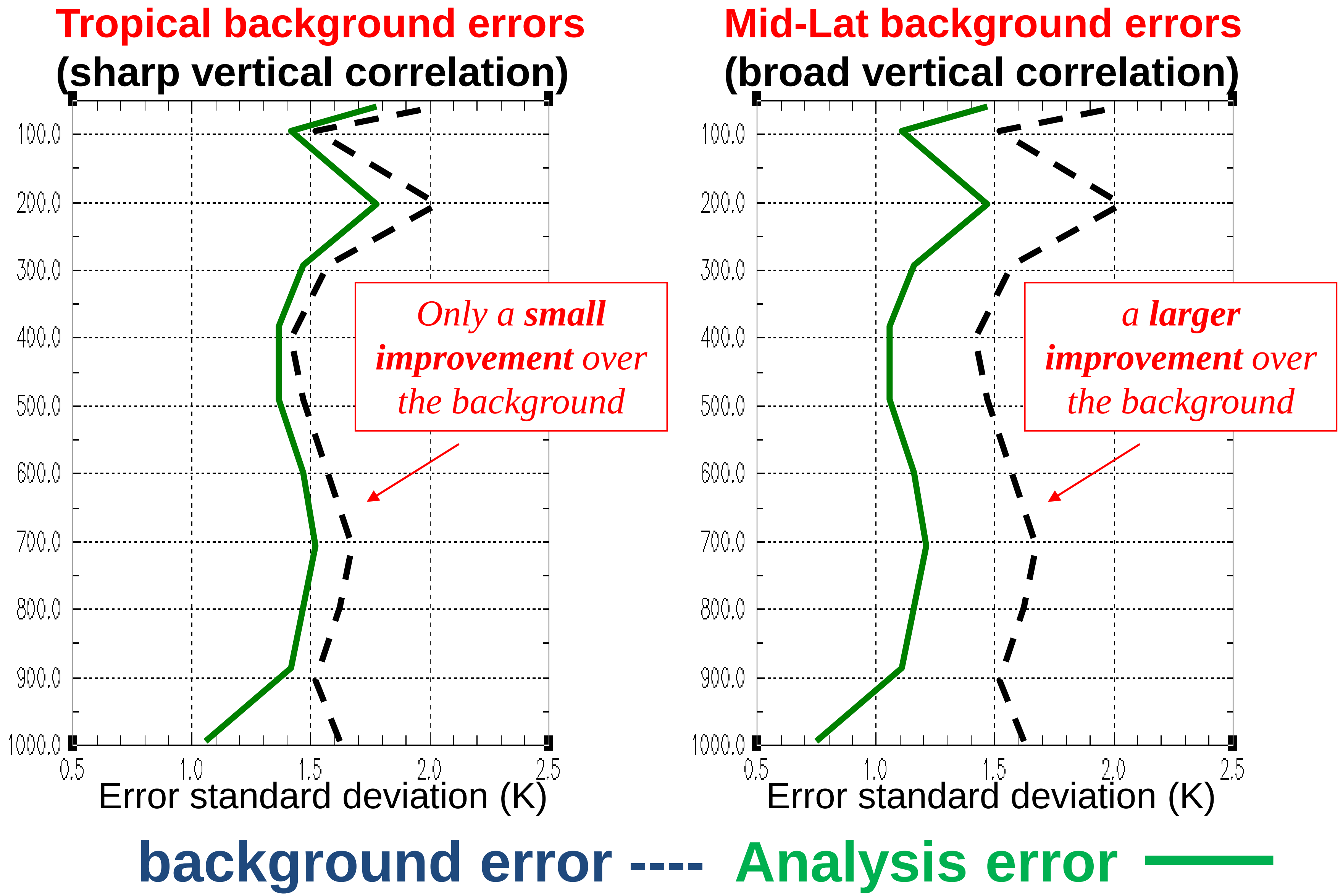
Sharper error correlations
in the Tropics

700hPa T error



Broader vertical correlations
in the mid-latitudes

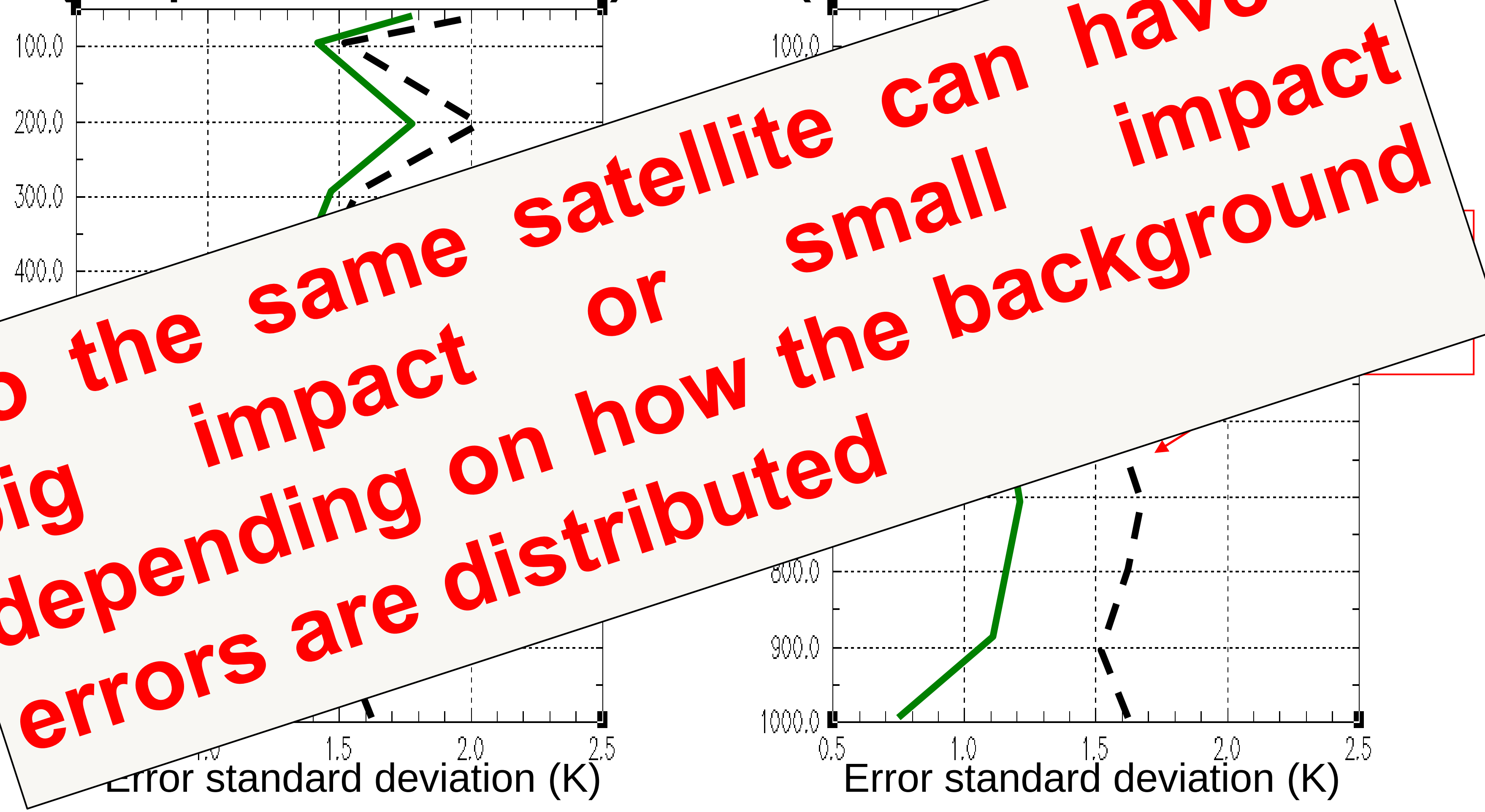
Sharp and broad vertical correlation



Sharp and broad vertical correlation

Tropical background errors
(sharp vertical correlation)

Mid-Lat background
(broad verti



So the same satellite can have a big impact or small impact depending on how the background errors are distributed

background error ---- Analysis error

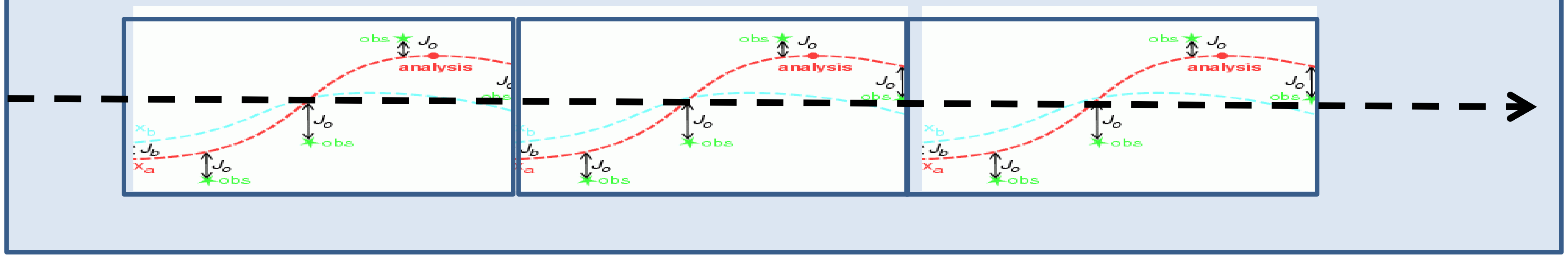
Background error specification in Observing System Experiments (OSE)

For example…..

Evaluating the impact of Satellite Observations with denial experiments

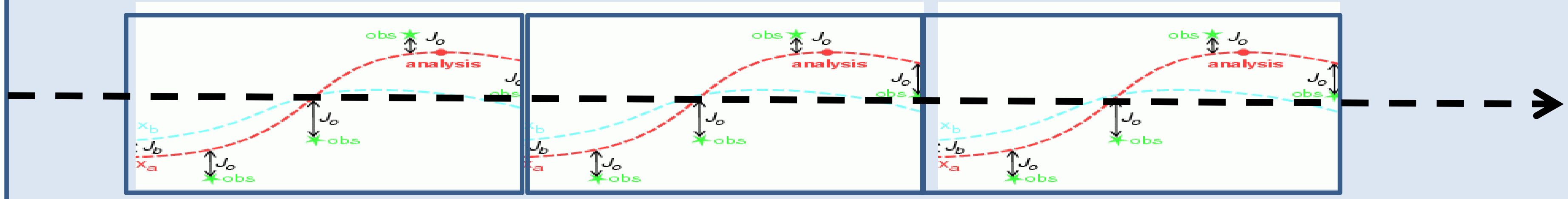
Control assimilation

Control assimilation system with all observations

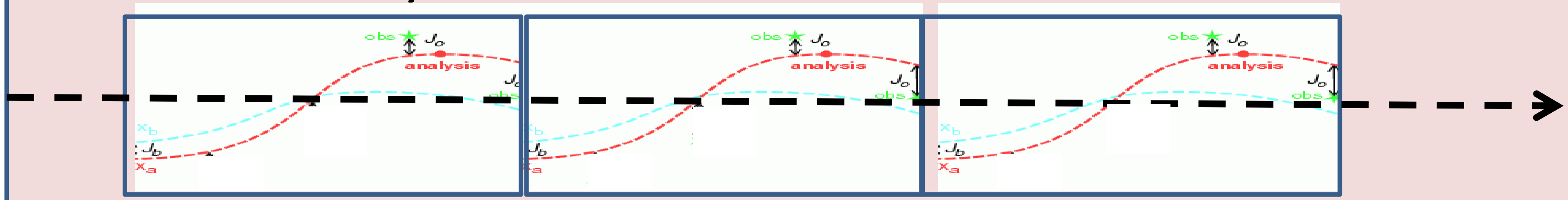


Assimilation with no satellite data

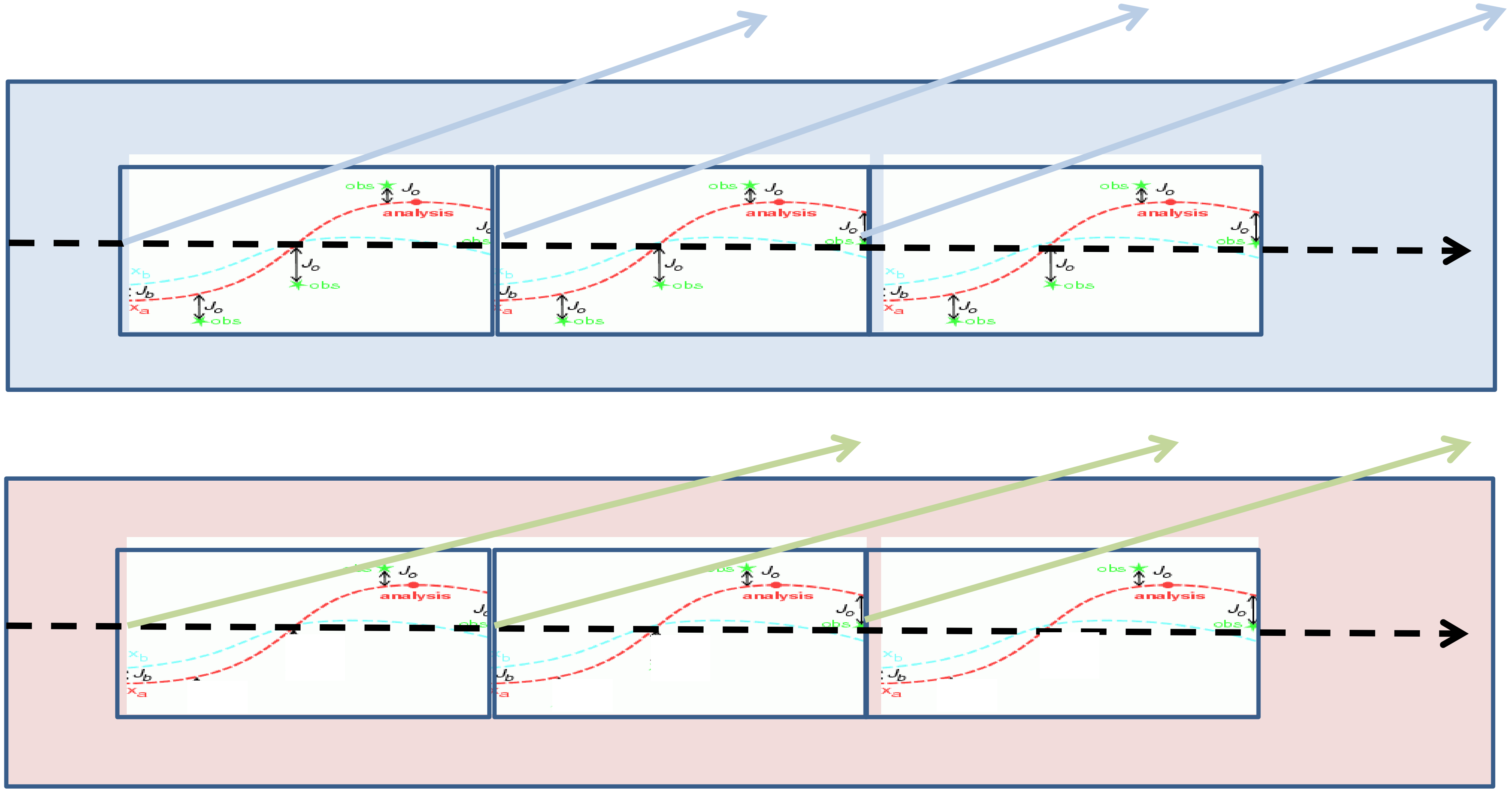
Control assimilation system with all observations used



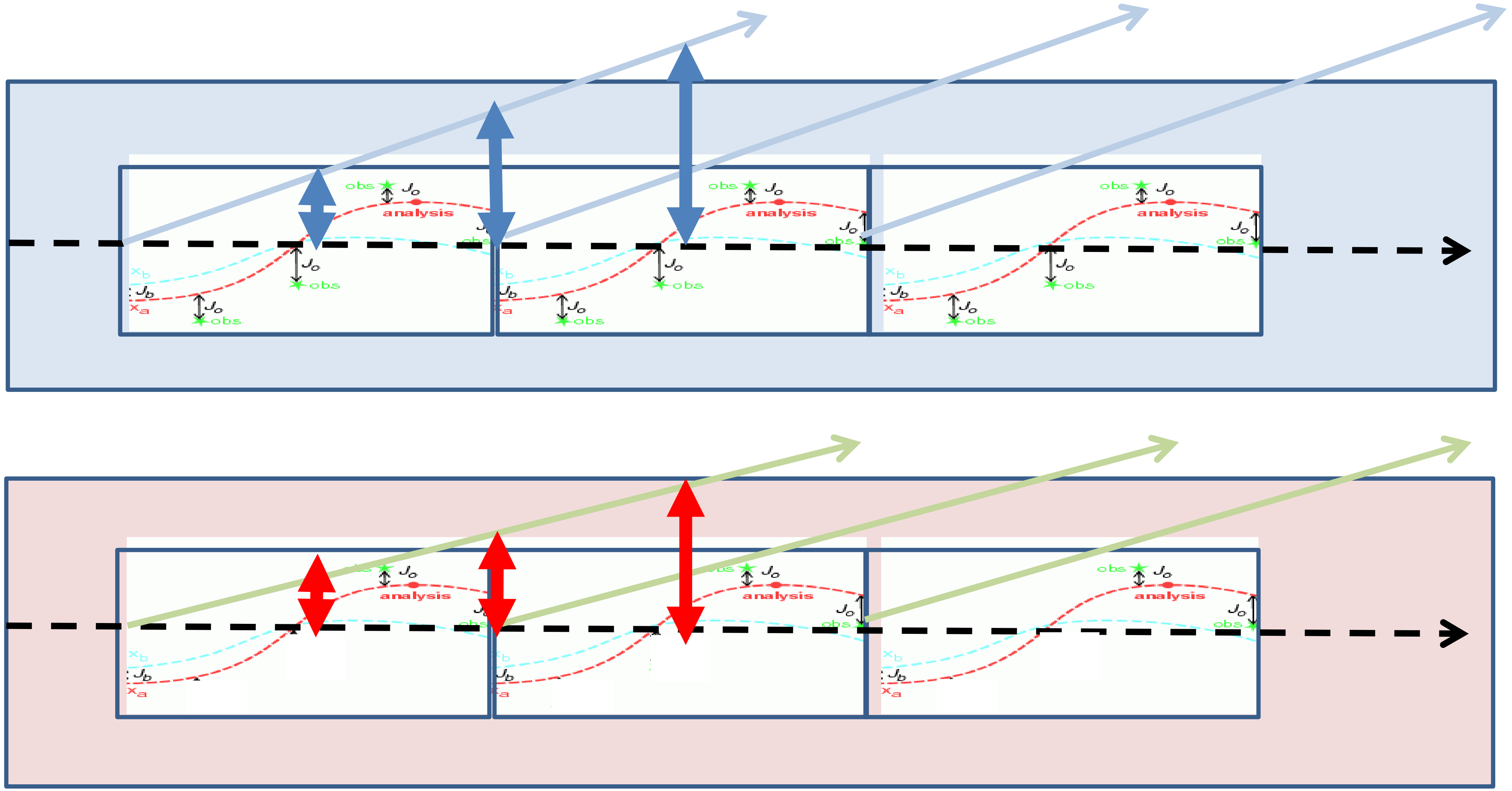
Assimilation system with all SATELLITE observations denied



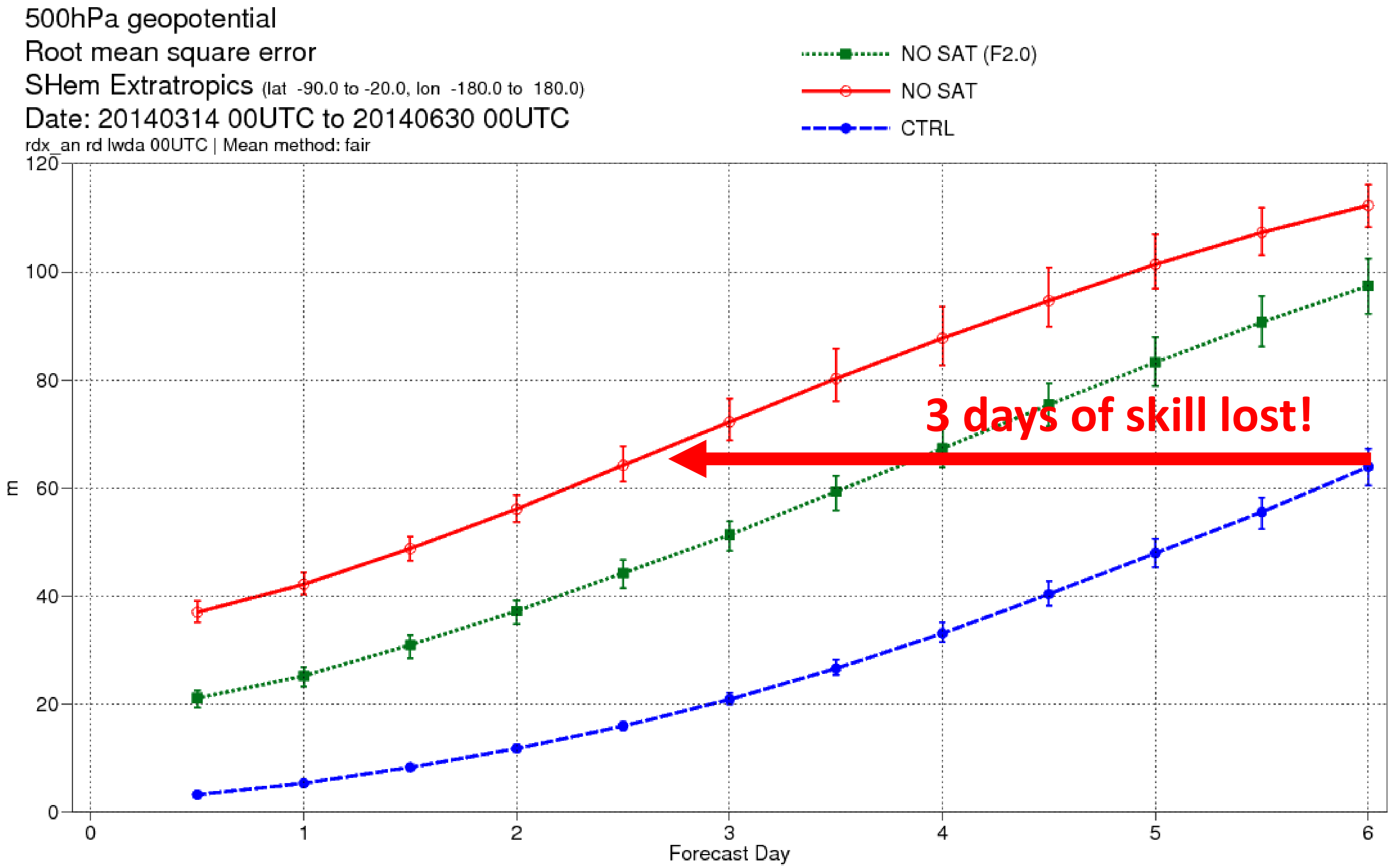
We run forecasts from each assimilation



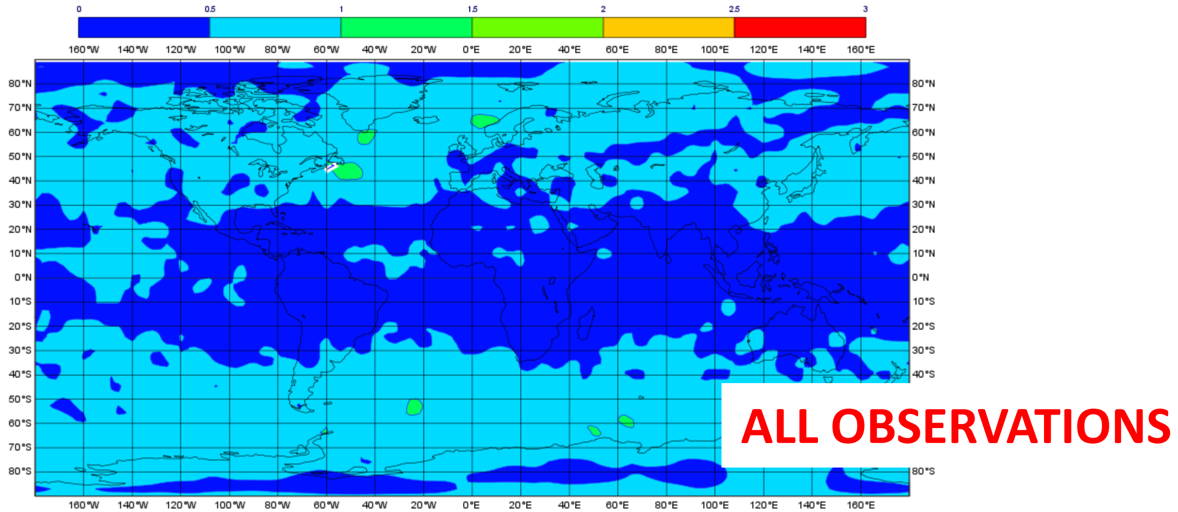
We verify the forecast of each assimilation

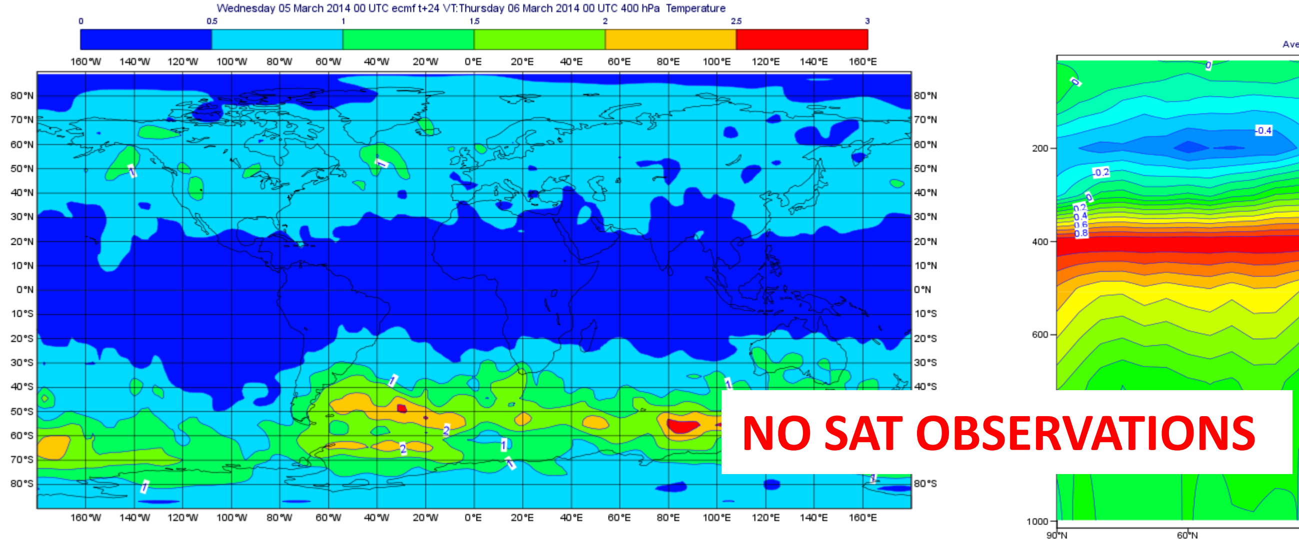


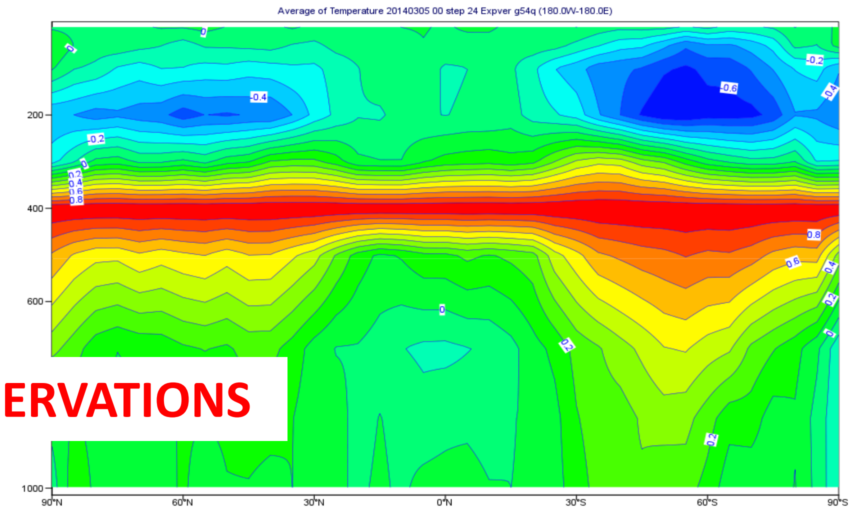
Impact of removing satellite observations



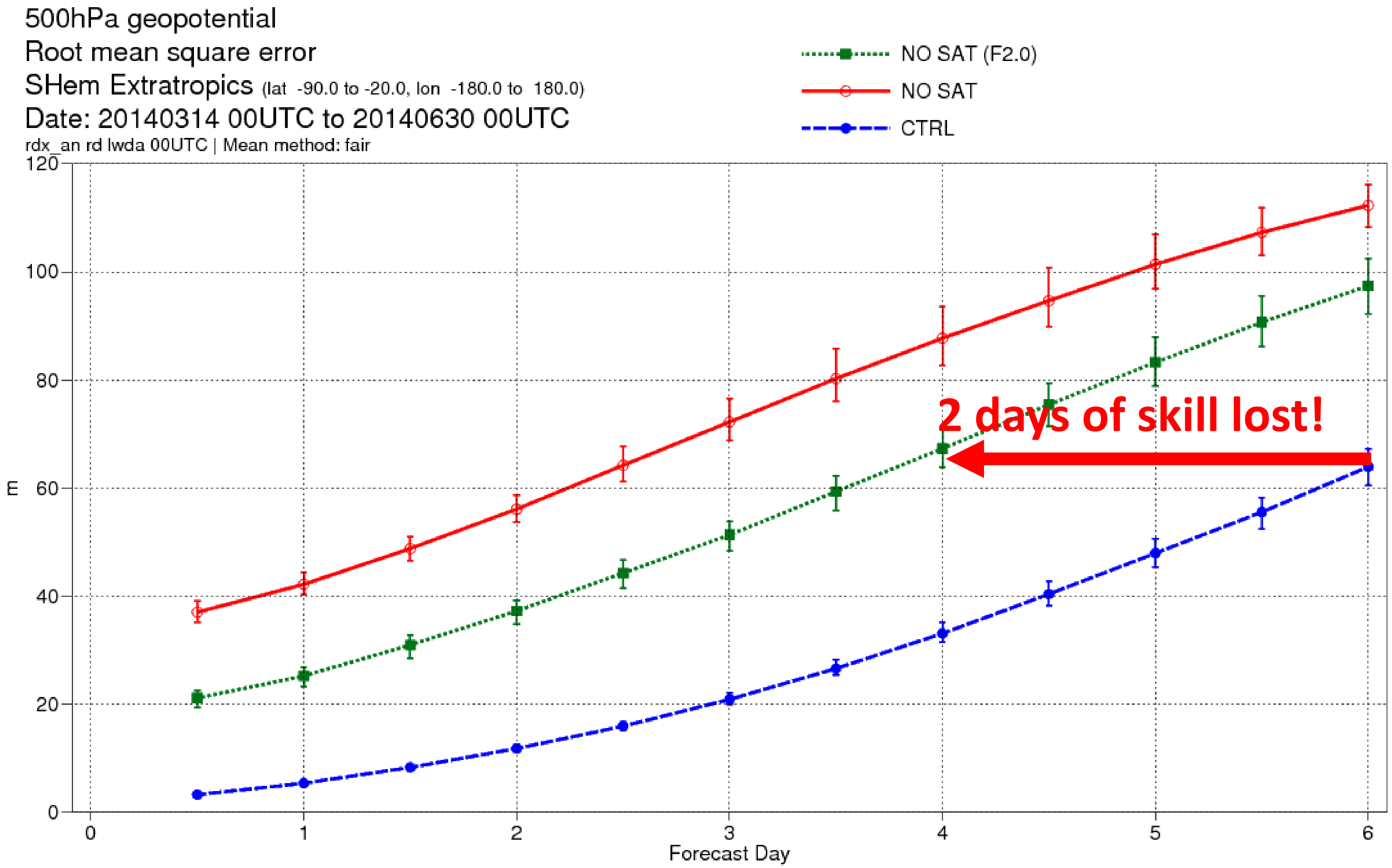
But we should retune the background errors!





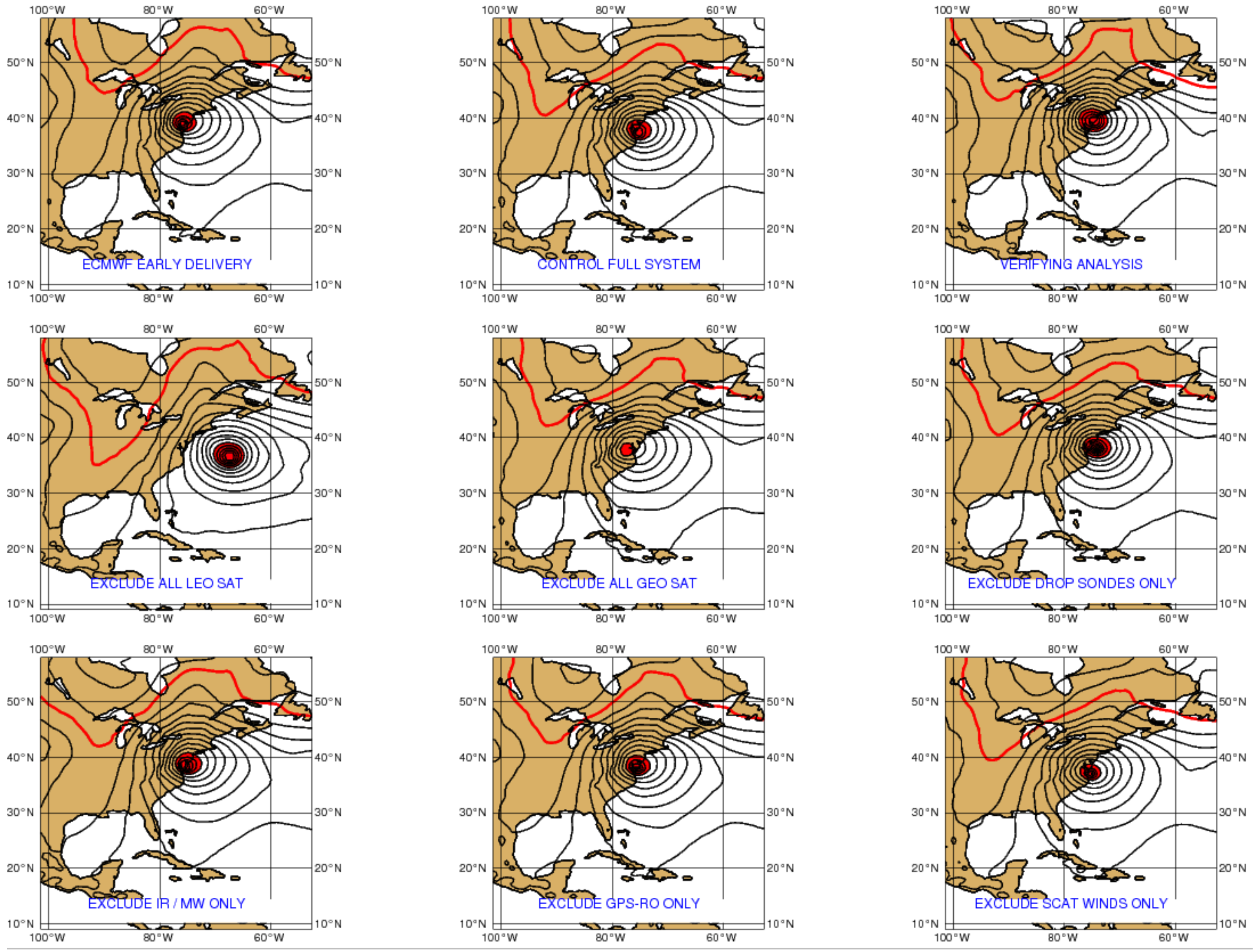


Satellite Impact with retuned (larger) background errors

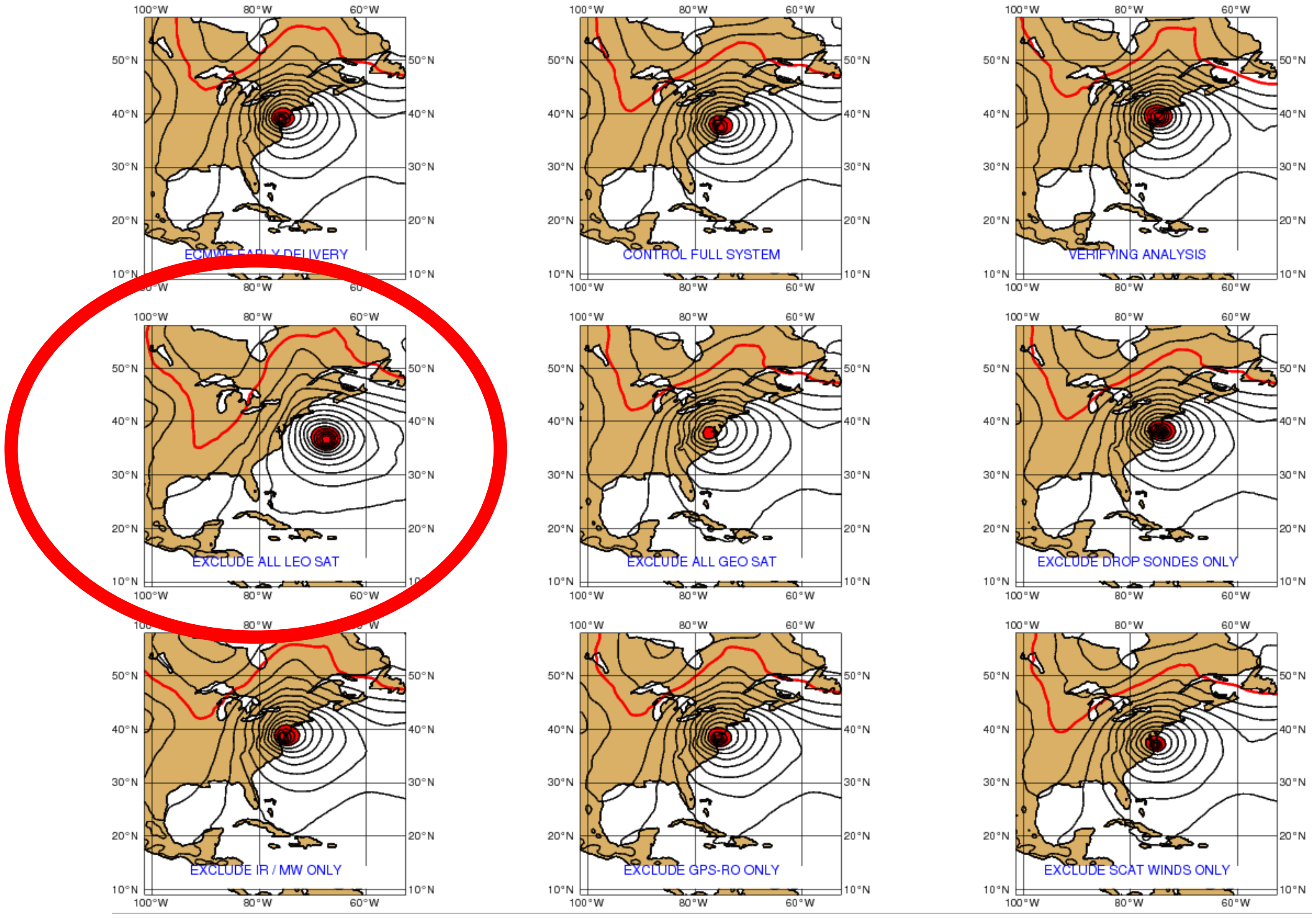


Hurricane Sandy

Data denial experiments

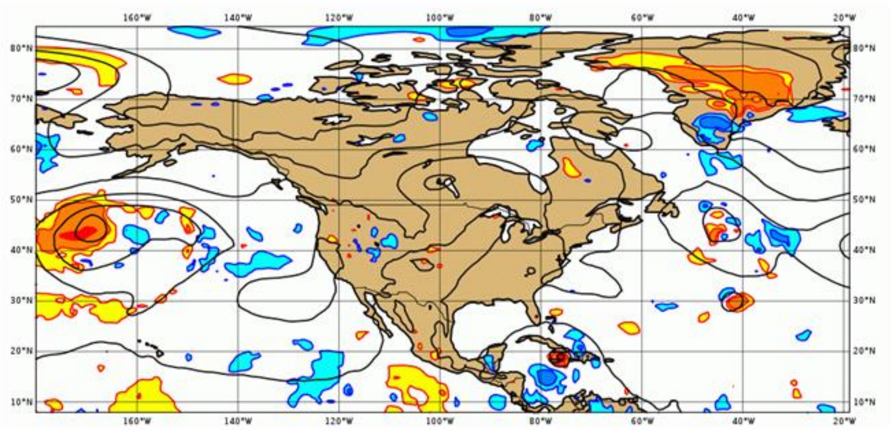


Data denial experiments

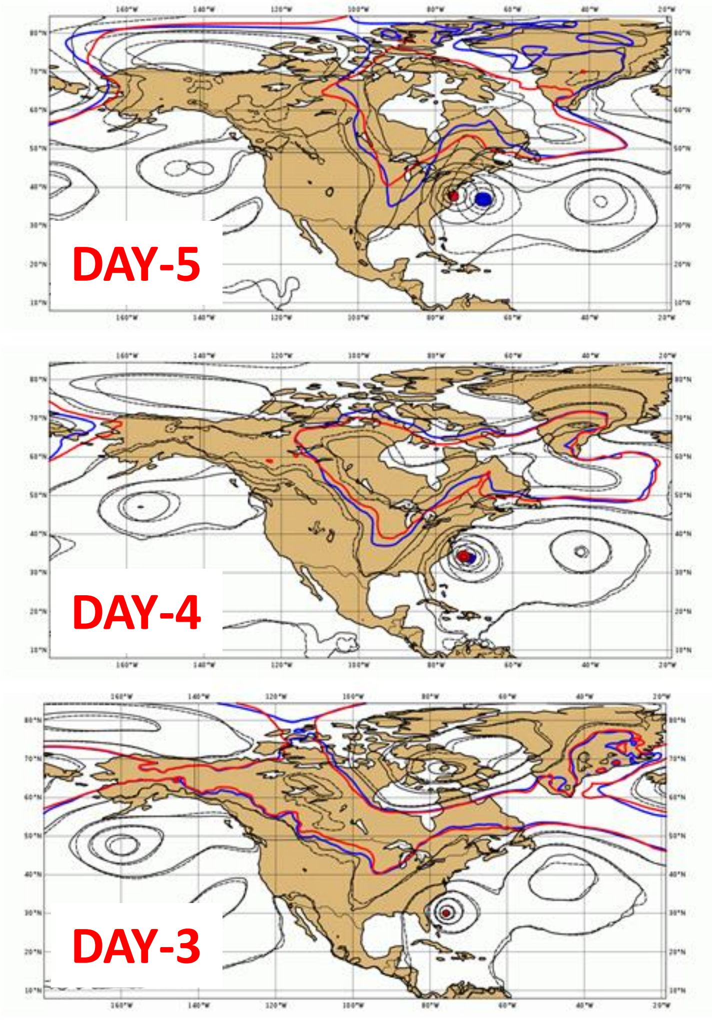


Data denial experiments (no LEO satellites)

Changes to the analysis (2012-10-25) when all LEO satellite observations are removed



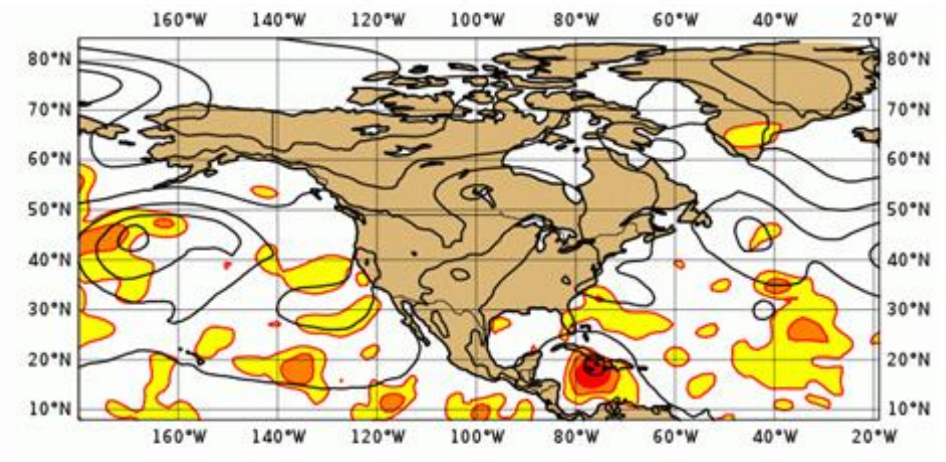
Changed the forecast



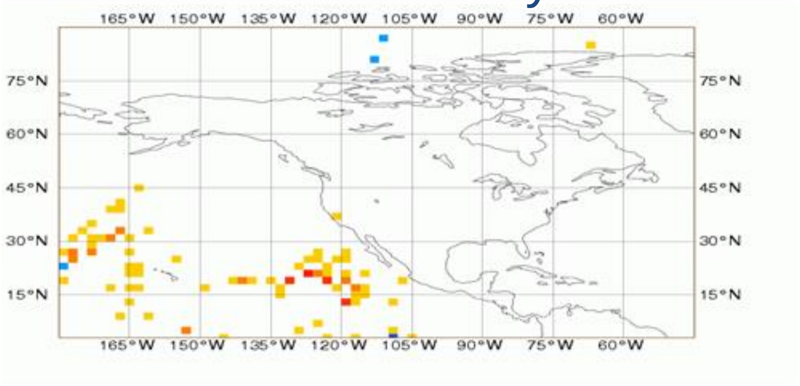
Re-calibrate (inflate) background errors

(to account for the LEO data being removed)

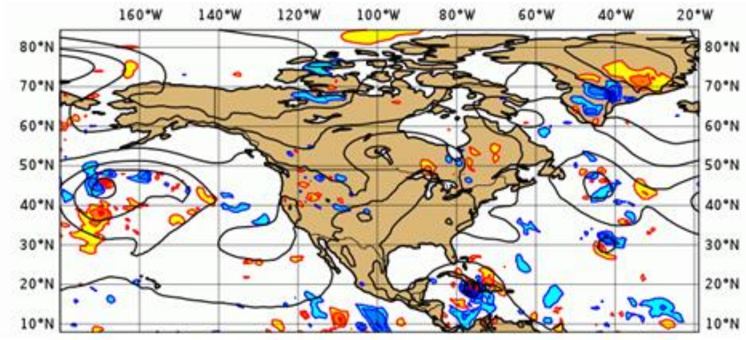
Inflated background errors



More weight to other OBS kept in the assimilation system

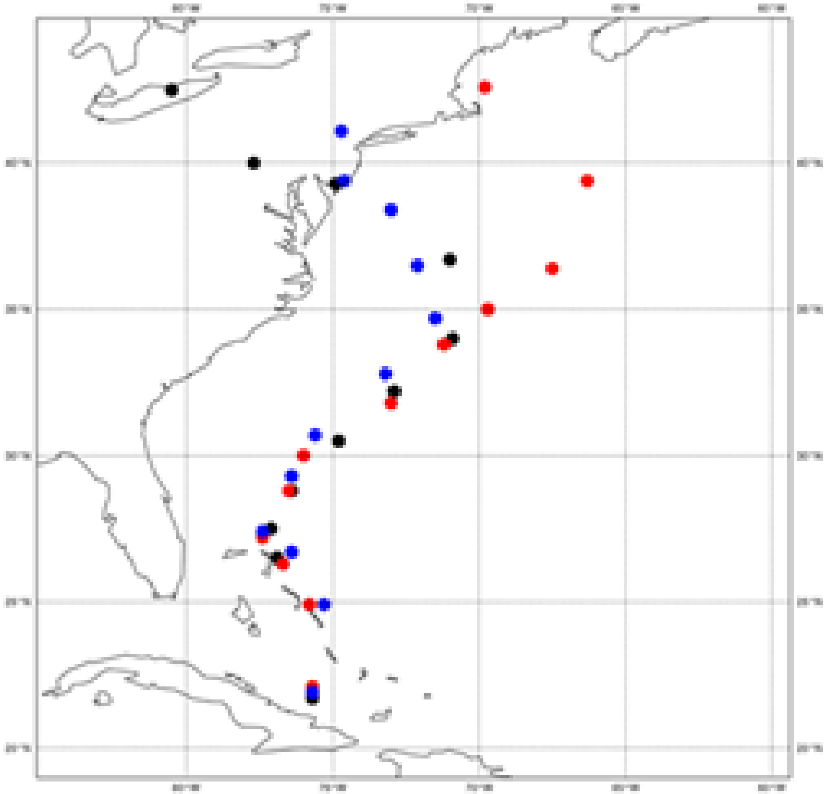


Changes to the analysis

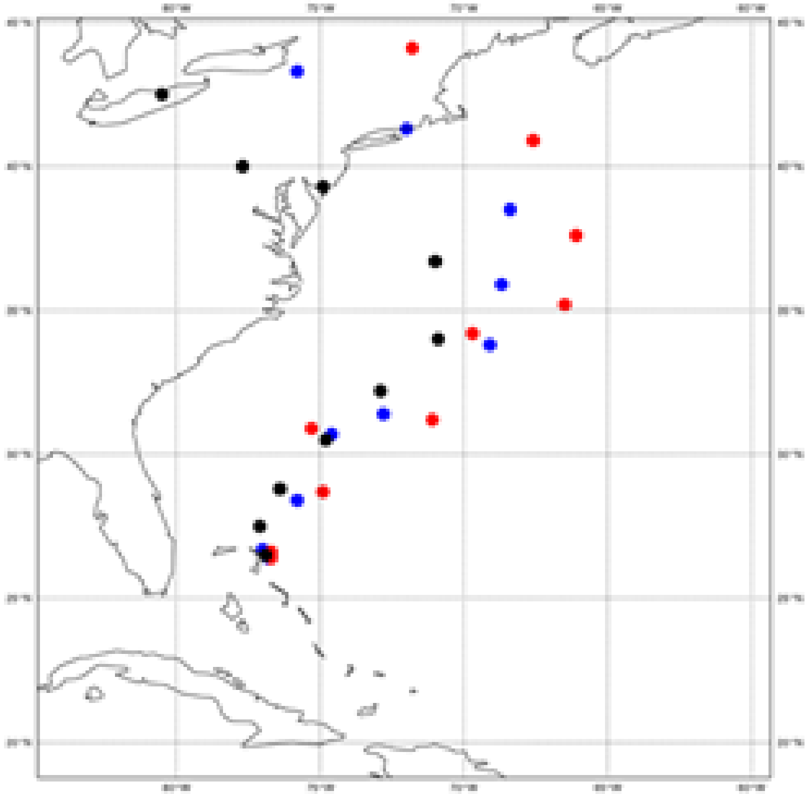


Re-calibrated background errors change the hurricane forecast!

Day 5 from 2012-10-25



Day 4 from 2012-10-26



Summary

- Background errors are a crucial element of any data assimilation system and particularly important for radiance observations (due to their poor vertical resolution)

- Background errors are flow-dependent, but also depend on the observations in the system and even the method used to evaluate them!

- The impact of satellite observations depends on the characteristics of the background errors we are attempting to correct with the data

- Great care must be taken in evaluating satellite impact with observing system experiments

Questions ?

You may want to impose additional constraints