Observation errors

Niels Bormann
(n.bormann@ecmwf.int)

ECMWF

Outline of lecture

1. What are observation errors?
2. Diagnosing observation errors
3. Specification of observation errors in practice
4. Observation error correlations
5. Summary

Outline of lecture

1. **What are observation errors?**
2. Diagnosing observation errors
3. Specification of observation errors in practice
4. Observation error correlations
5. Summary

Observation error and the cost function

- **Every observation has an error vs the truth:**
  - **Systematic error**
    - **Needs to be removed through bias correction (see separate lecture)**
  - **Random error**
    - **Mostly assumed Gaussian; described by observation error covariance "R" in the observation cost function:**

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x}-\mathbf{x}_b)^T \mathbf{B}^{-1}(\mathbf{x}-\mathbf{x}_b) + \frac{1}{2}(\mathbf{y}-\mathbf{H}[\mathbf{x}])^T \mathbf{R}^{-1}(\mathbf{y}-\mathbf{H}[\mathbf{x}])$$

- **R is a matrix, often specified through the square root of the diagonals ("σ_O") and a correlation matrix (which can be the identity matrix).**

ECMWF

Role of the observation error

- R and B together determine the weight of an observation in the assimilation.
- In the linear case, the minimum of the cost function can be found at $\mathbf{x}_a$:

$$\underbrace{(\mathbf{x}_a - \mathbf{x}_b)}_{\text{Increment}} = \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1}\underbrace{(\mathbf{y} - \mathbf{H}\mathbf{x}_b)}_{\text{Departure, innovation, “o-b”}}$$

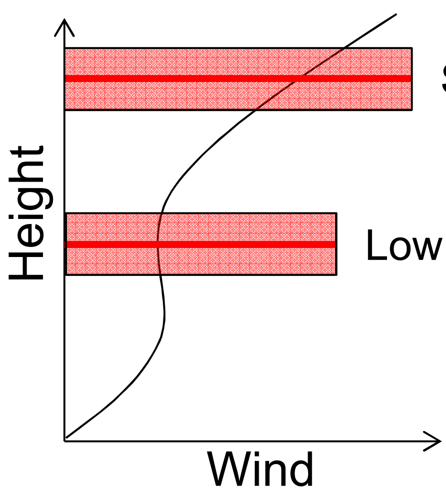
  - “Large” observation error → smaller increment, analysis draws less closely to the observations
  - “Small” observation error → larger increment, analysis draws more closely to the observations

Contributions to observation error

- Measurement error
  - E.g., instrument noise for satellite radiances
- Forward model (observation operator) error
  - E.g., radiative transfer error
- Representativeness error
  - E.g., point measurement vs model representation
- Quality control error
  - E.g., error due to the cloud detection scheme missing some clouds in clear-sky radiance assimilation

Situation-dependence of observation error

- Observation errors can be situation-dependent, especially through situation-dependence of the forward model error.
- Examples:
  - Cloud/rain-affected radiances: Representativeness error is much larger in cloudy/rainy regions than in clear-sky regions
  - Effect of height assignment error for Atmospheric Motion Vectors:



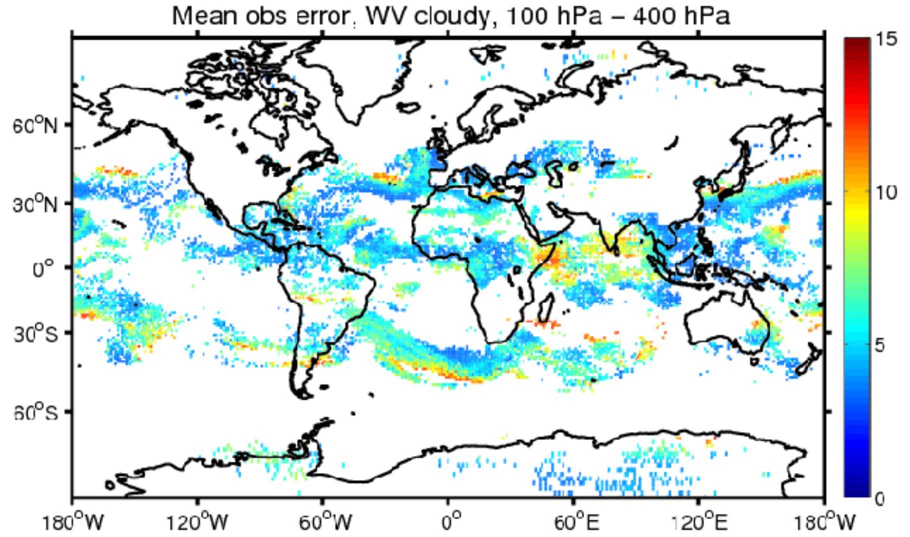
Strong shear – larger wind error due to height assignment error

Low shear – small wind error due to height assignment error

ECMWF

Situation-dependence of observation error

- Observation errors can be situation-dependent, especially through situation-dependence of the forward model error.
- Examples:
  - Cloud/rain-affected radiances: Representativeness error is much larger in cloudy/rainy regions than in clear-sky regions
  - Effect of height assignment error for Atmospheric Motion Vectors:



Current observation error specification for satellite data in the ECMWF system

- Globally constant, dependent on channel only:
  - AMSU-A, MHS, ATMS, HIRS, AIRS, IASI
- Globally constant fraction, dependent on impact parameter:
  - GPS-RO
- Situation dependent:
  - MW imagers: dependent on channel and cloud amount
  - AMVs: dependent on level and shear (and satellite, channel, height assignment method)

- Error correlations are neglected.

Outline of lecture

1. What are observation errors?
2. **Diagnosing observation errors**
3. Specification of observation errors in practice
4. Observation error correlations
5. Summary

How can we estimate observation errors?

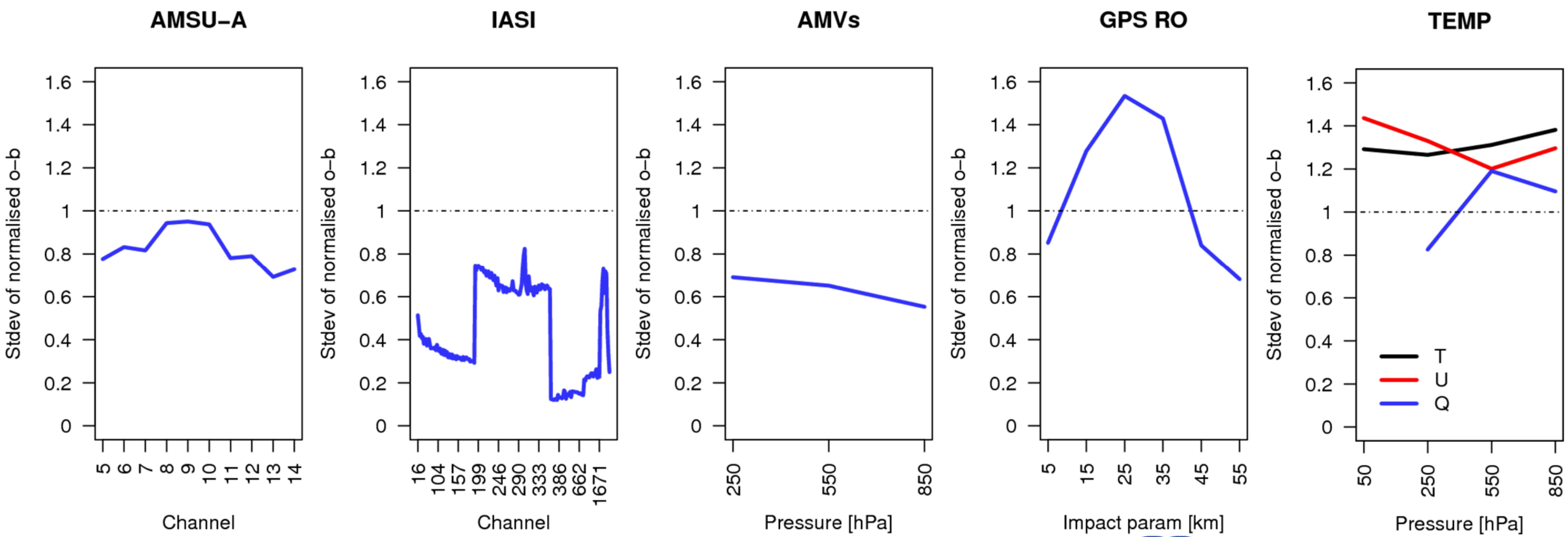
- **Several methods exist, broadly categorised as:**
  - <u>Error inventory:</u>
    - Based on considering all contributions to the error/uncertainty
  - <u>Diagnostics with collocated observations, e.g.:</u>
    - Hollingsworth/Lönnberg on collocated observations
    - Triple-collocations
  - <u>Diagnostics based on output from DA systems, e.g.:</u>
    - O-b statistics
    - Hollingsworth/Lönnberg
    - Desroziers et al 2006
    - Methods that rely on an explicit estimate of B
  - <u>Adjoint-based methods</u>

Basic departure-based diagnostics

- **If observation errors and background errors are uncorrelated then:**

$$Cov[(\mathbf{y}-\mathbf{H}[\mathbf{x}_b]),(\mathbf{y}-\mathbf{H}[\mathbf{x}_b])] = \mathbf{H}\mathbf{B}_{true}\mathbf{H}^T + \mathbf{R}_{true}$$

- **Statistics of background departures give an upper bound for the *true* observation error.**
  - Standard deviations of background departures normalised by *assumed* observation error for the ECMWF system:



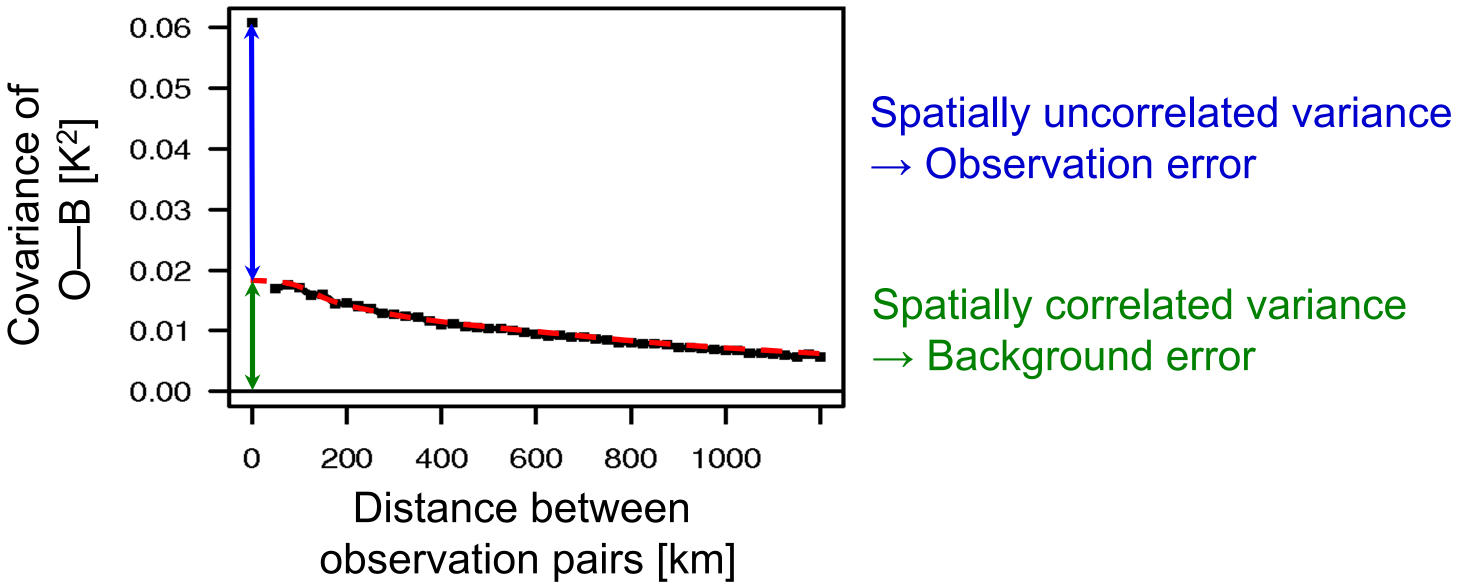
ECMWF

Departure-based diagnostics

- Standard deviations of o-b give information on observation ***and*** background error combined.
- Departure-based diagnostics try to separate contributions from background and observation errors by making assumptions (which may or may not be true).
  - Assume we know the background error → subtract background error
  - Assume a certain structure of the errors → Hollingsworth/Lönnberg
  - Assume weights used in the assimilation system are accurate → Desroziers diagnostic

Observation error diagnostics: Hollingsworth/Loennberg method (I)

- Based on a large database of pairs of departures.
- Basic assumption:
  - Background errors are spatially correlated, whereas observation errors are not.
  - This allows to separate the two contributions to the variances of background departures:

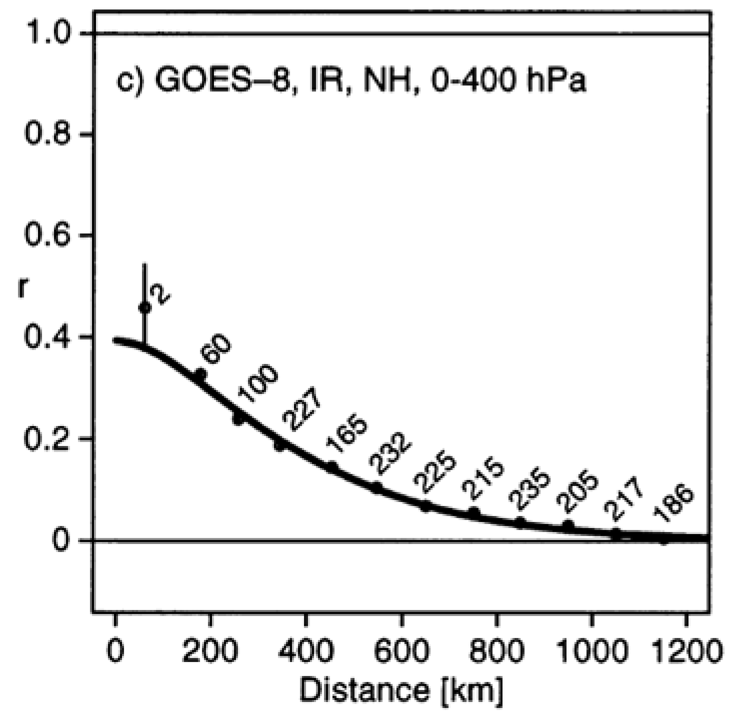


ECMWF

Observation error diagnostics: Hollingsworth/Loennberg method (II)

- Drawback: Not reliable when observation errors are spatially correlated.

- Similar methods have been used with differences between two sets of collocated observations:
  - Example: AMVs collocated with radiosondes (Bormann et al 2003).
    - Radiosonde error assumed spatially uncorrelated.



Observation error diagnostics: Desroziers diagnostic (I)

- Basic assumptions:
  - Assimilation process can be adequately described through linear estimation theory.
  - Weights used in the assimilation system are consistent with true observation and background errors.
- Then the following relationship can be derived:

$$\tilde{\mathbf{R}} = Cov[\mathbf{d}_a, \mathbf{d}_b]$$

with $\mathbf{d}_a = (\mathbf{y} - \mathbf{H}[\mathbf{x}_a])$ (analysis departure)

$\mathbf{d}_b = (\mathbf{y} - \mathbf{H}[\mathbf{x}_b])$ (background departure)

(see Desroziers et al. 2005, QJRMS)

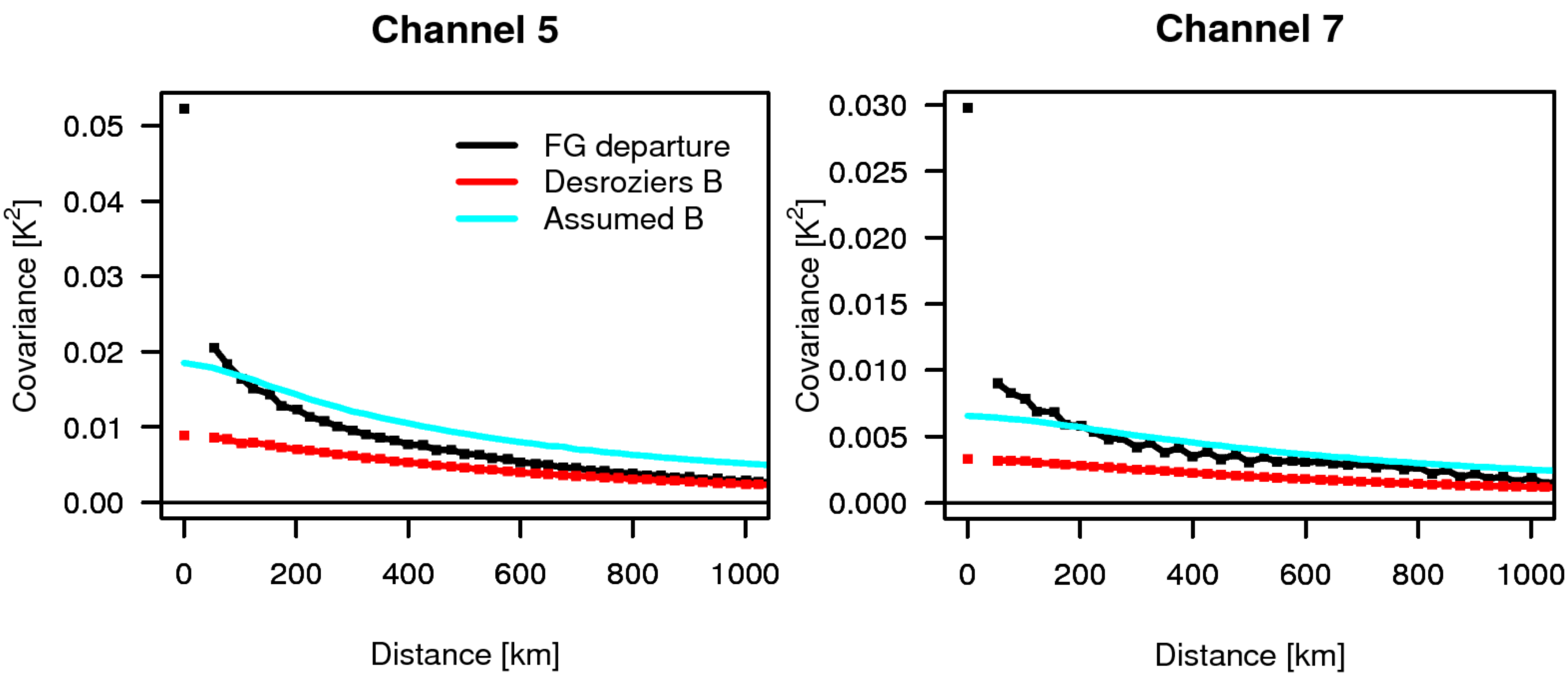
- Consistency diagnostic for the specification of R.

Observation error diagnostics: Desroziers diagnostic (II)

- Desroziers diagnostic can be applied iteratively.
- Simulations in toy-assimilation systems:
  - Good convergence if the correlation length-scales for observation errors and background errors are sufficiently different.
  - Mis-leading results if correlation length-scales for background and observation errors are too similar.
- For real assimilation systems, the applicability of the diagnostic for estimating observation errors is still subject of research.

Examples of observation error diagnostics: AMSU-A

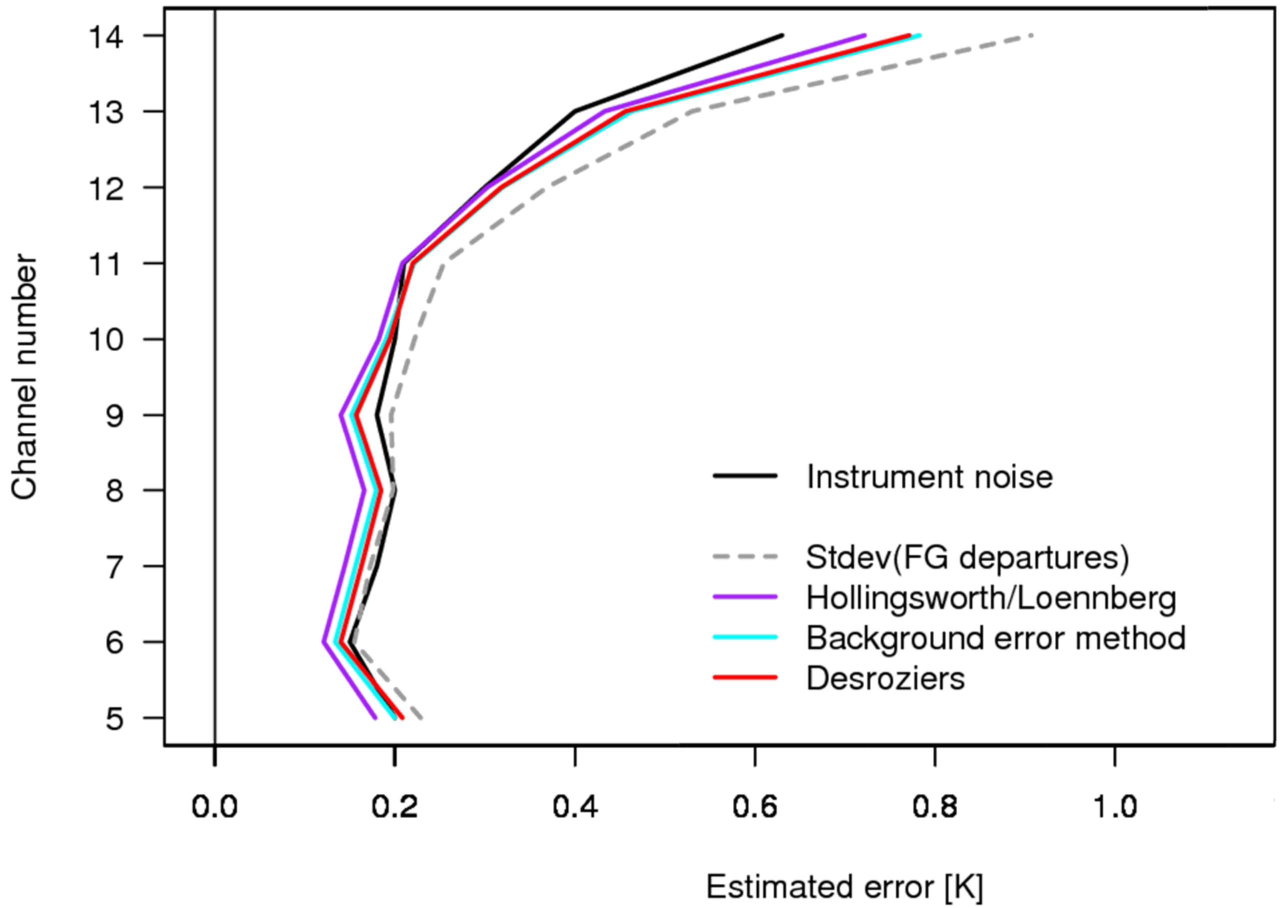
Spatial covariances of background departures:



(See also Bormann and Bauer 2010)

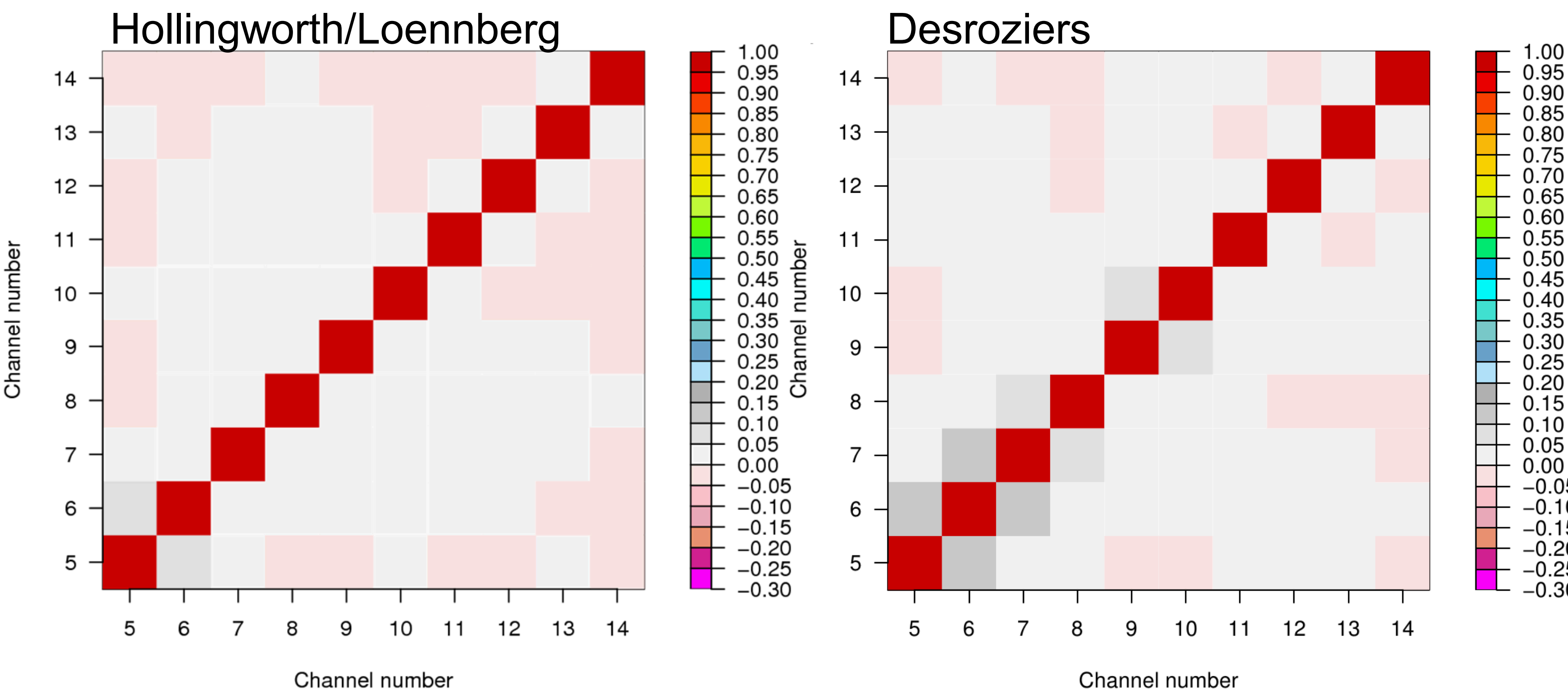
Examples of observation error diagnostics: AMSU-A

Diagnostics for σ_o



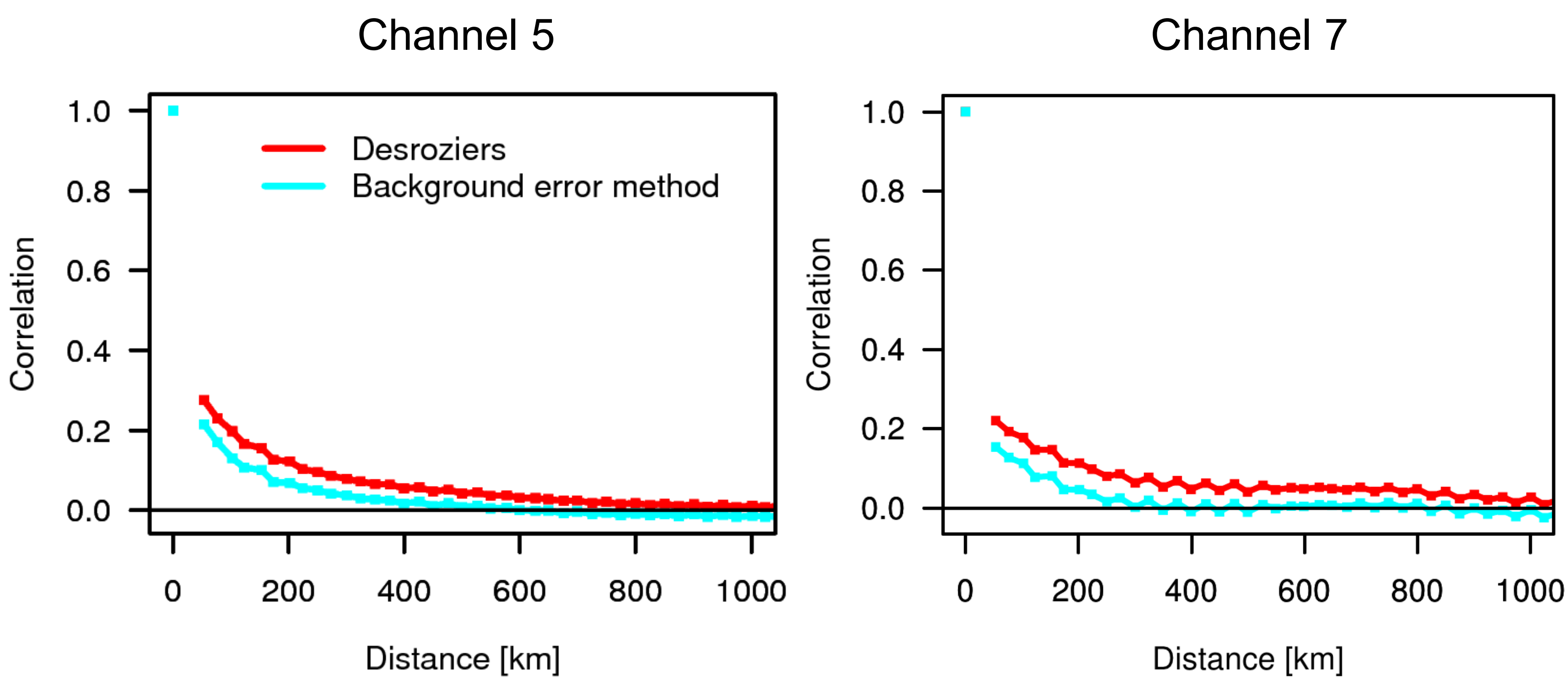
Examples of observation error diagnostics: AMSU-A

Inter-channel error correlations:



Examples of observation error diagnostics: AMSU-A

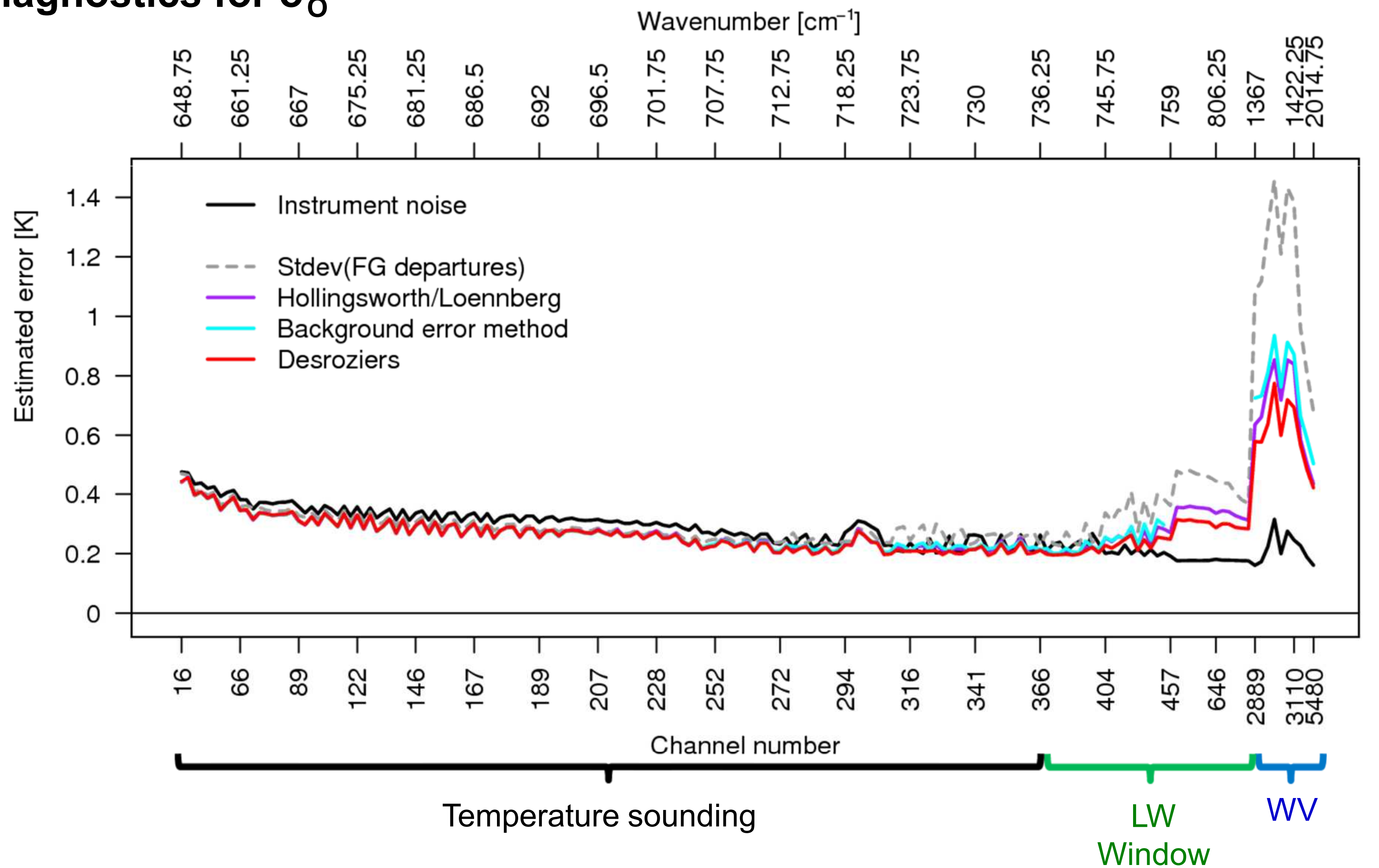
Spatial error correlations:



ECMWF

Examples of observation error diagnostics: IASI

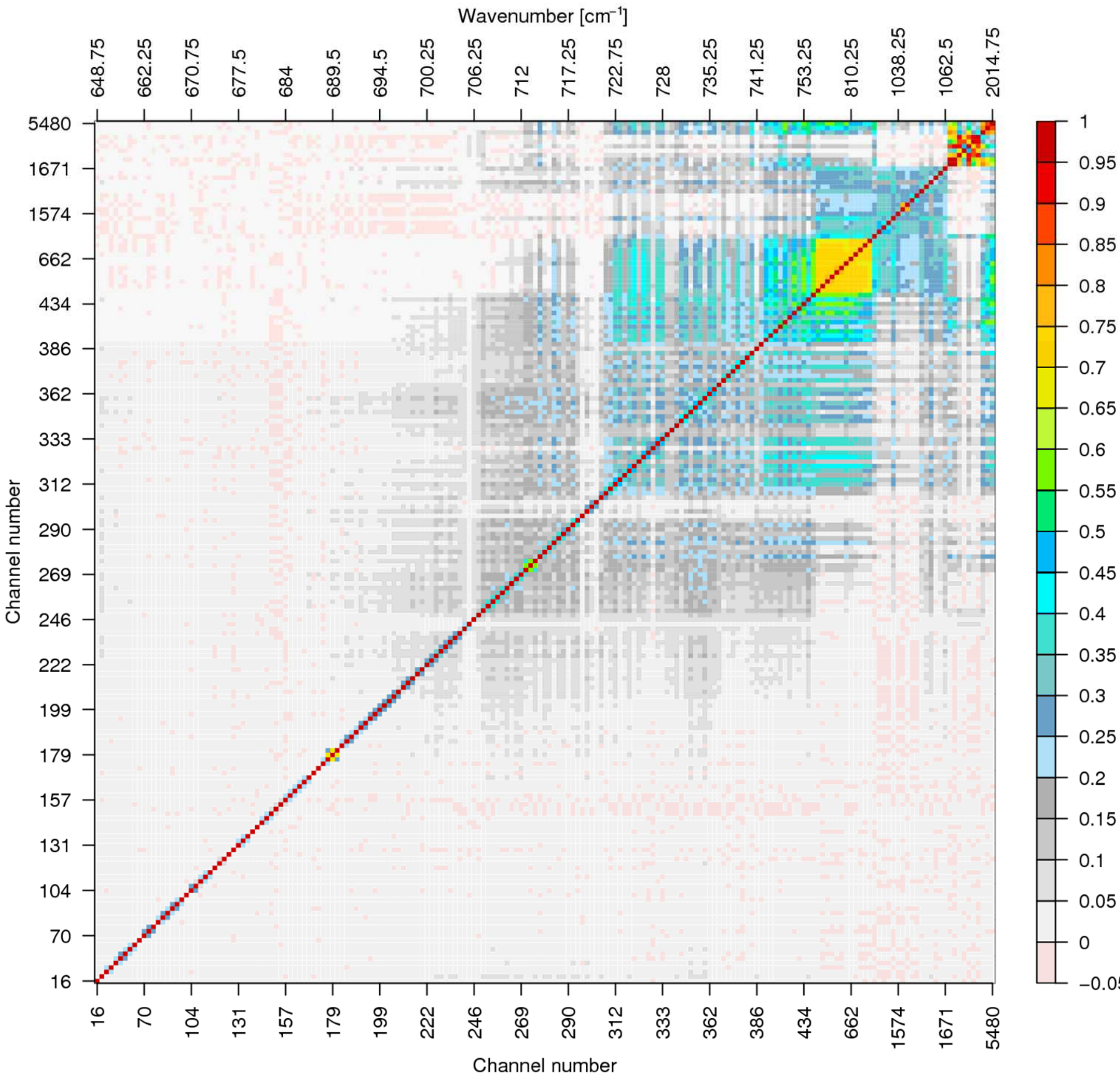
Diagnostics for σ_o



ECMWF

Examples: IASI

Inter-channel error correlations:

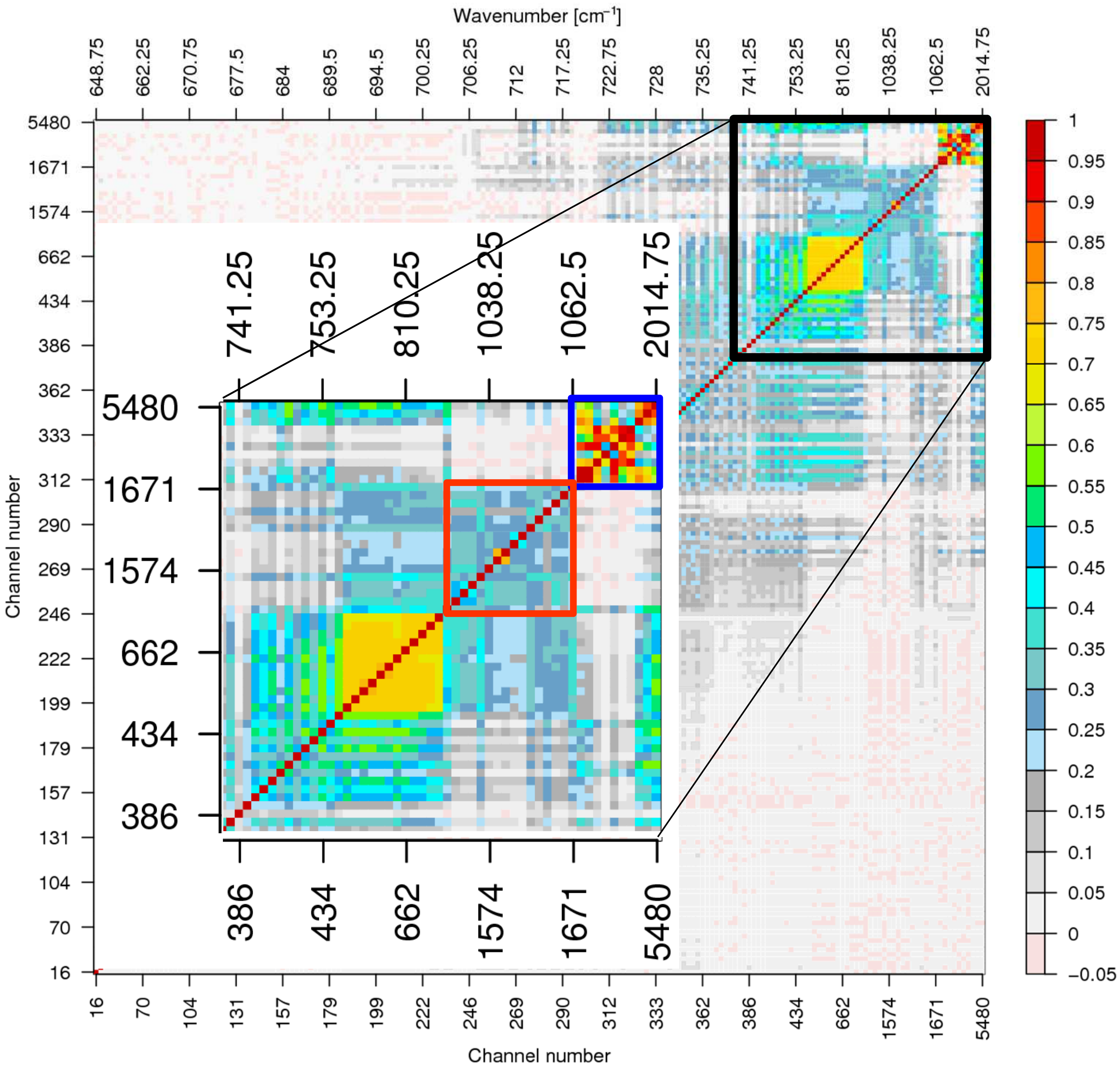


Examples: IASI

Inter-channel error correlations:

Humidity

Ozone



Outline of lecture

1. What are observation errors?
2. Diagnosing observation errors
3. **Specification of observation errors in practice**
4. Observation error correlations
5. Summary

How to specify observation errors in practice?

- Diagnostics can provide guidance for observation error specification, including:
  - Relative size of observation and background errors:
    - For most satellite data, the errors in the observations are larger than the errors in the background.
  - Presence of observation error correlations.

- BUT: Observation errors specified in assimilation systems are often simplified:
  - Observation error covariance is mostly assumed to be diagonal.
    - Observations with “complicated” observation errors can be more difficult to assimilate.
    - *Assumed* observation errors may need adjustments.

Too large assumed observation errors tend to be safer than too small ones. Why?

- **Consider linear combination of two estimates x_b and y:**

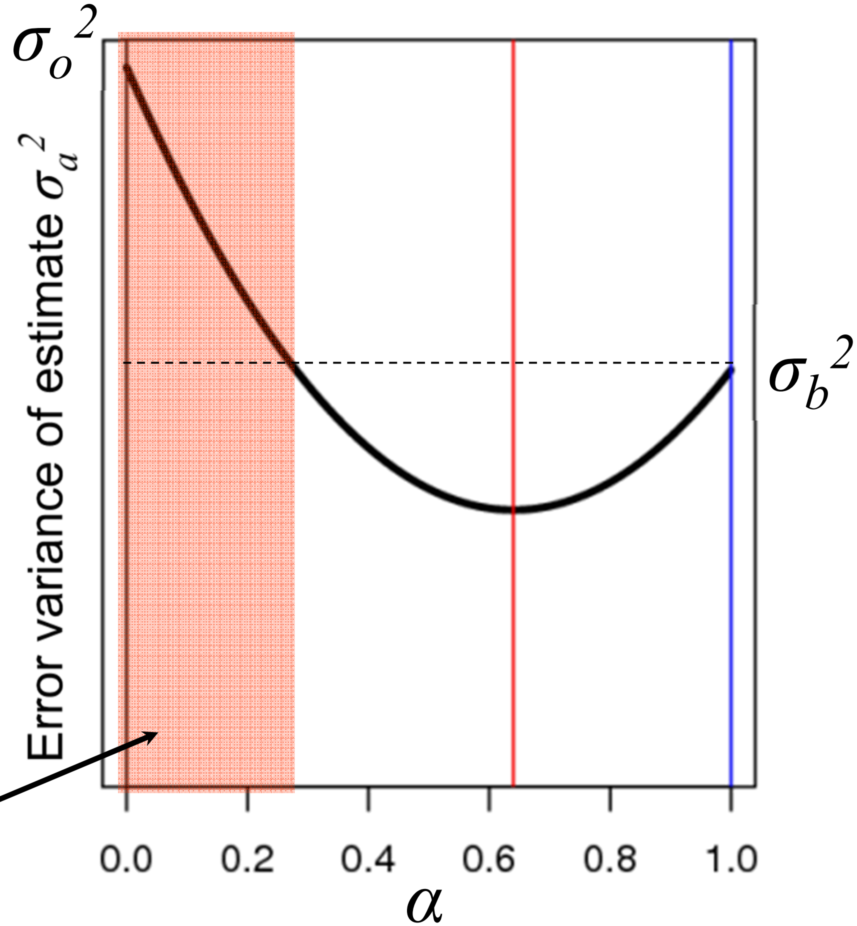
$$x_a = \alpha x_b + (1-\alpha) y$$

- **The error variance of the linear combination is:**

$$\sigma_a^2 = \alpha^2 \sigma_b^2 + (1-\alpha)^2 \sigma_o^2$$

- **Optimal weighting:**

$$\alpha = \frac{\sigma_o^2}{\sigma_b^2 + \sigma_o^2}$$

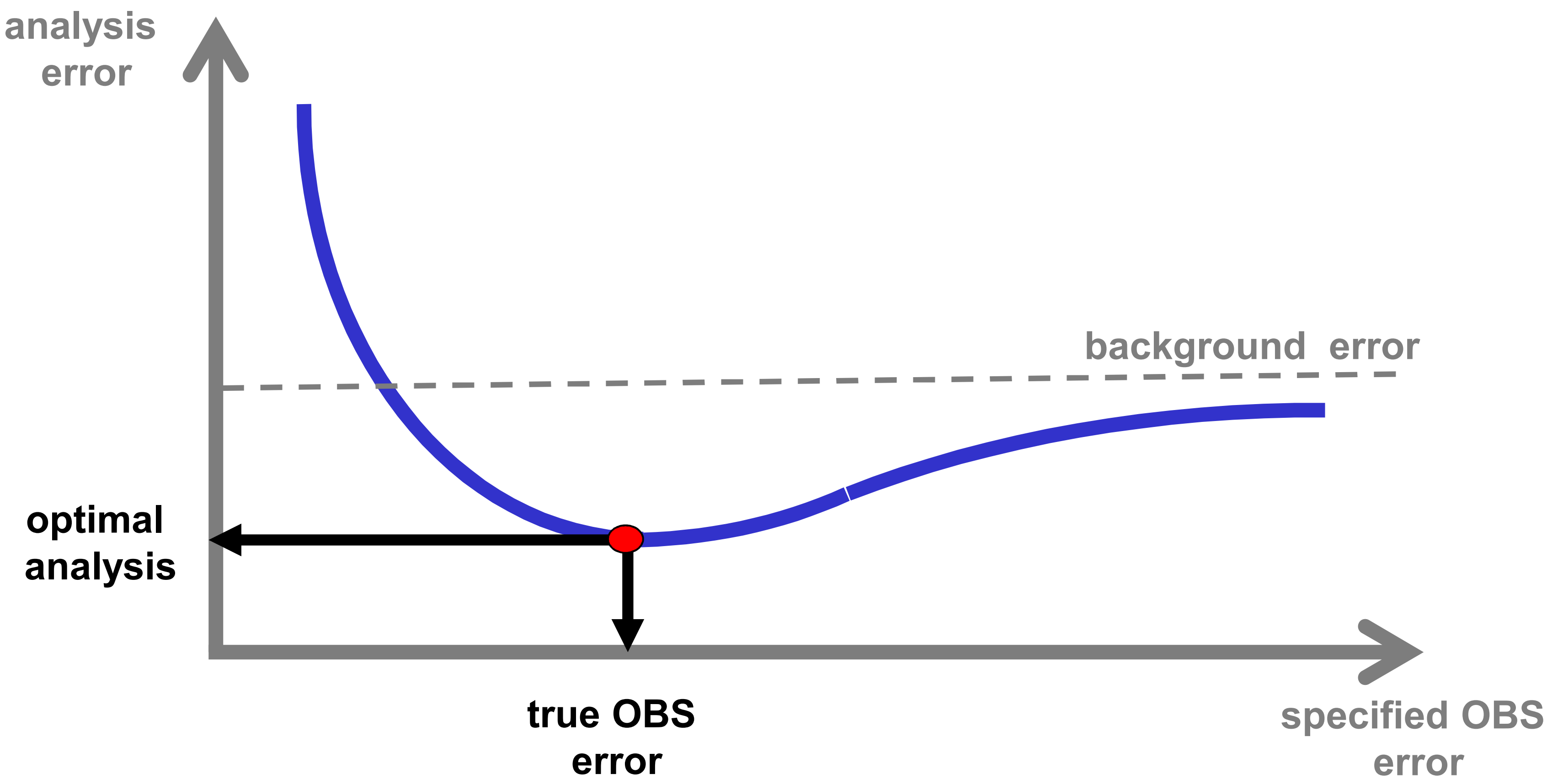


Danger zone: Too small *assumed* σ_o will lead to an analysis worse than the background when the (true) $\sigma_o > \sigma_b$.

Assuming an inflated σ_o will never result in deterioration.

Observation errors:

- Specifying the correct observation error produces an optimal analysis with minimum error.



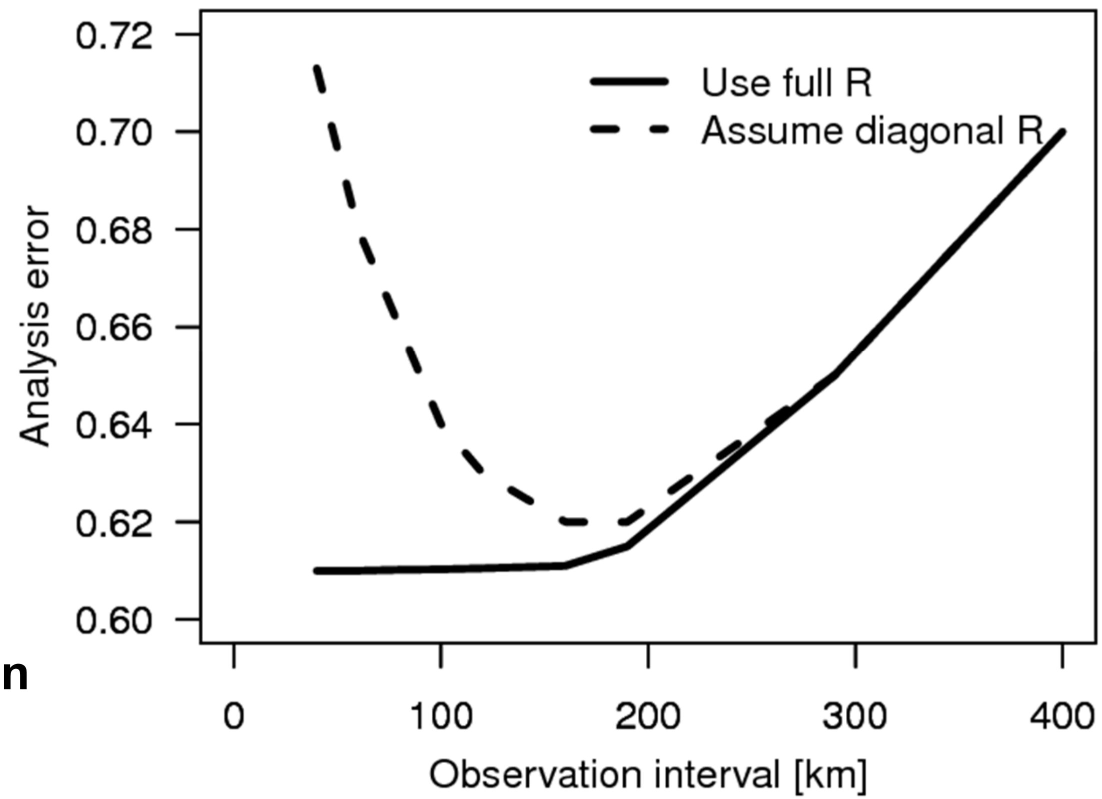
What to do when there are error correlations?

- Thinning
  - Ie, reduce observation density so that error correlations are not relevant.
- Error inflation
  - Ie, use diagonal R with larger σ_O than diagnostics suggest.
- Take error correlations into account in the assimilation

Spatial error correlations and thinning

- If the observations have spatial error correlations, but these are neglected in the assimilation system, assimilating these observations too densely can have a negative effect.

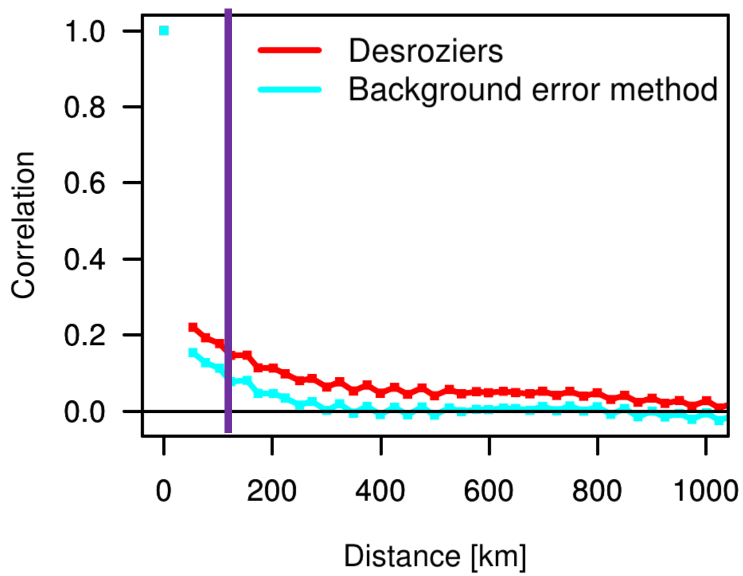
- Practical solution: Thinning, ie select one observation within a "thinning box".
- Using fewer observations gives better results!
- See Liu and Rabier (2002), QJRMS: "Optimal" thinning when r ≈ 0.15-0.2



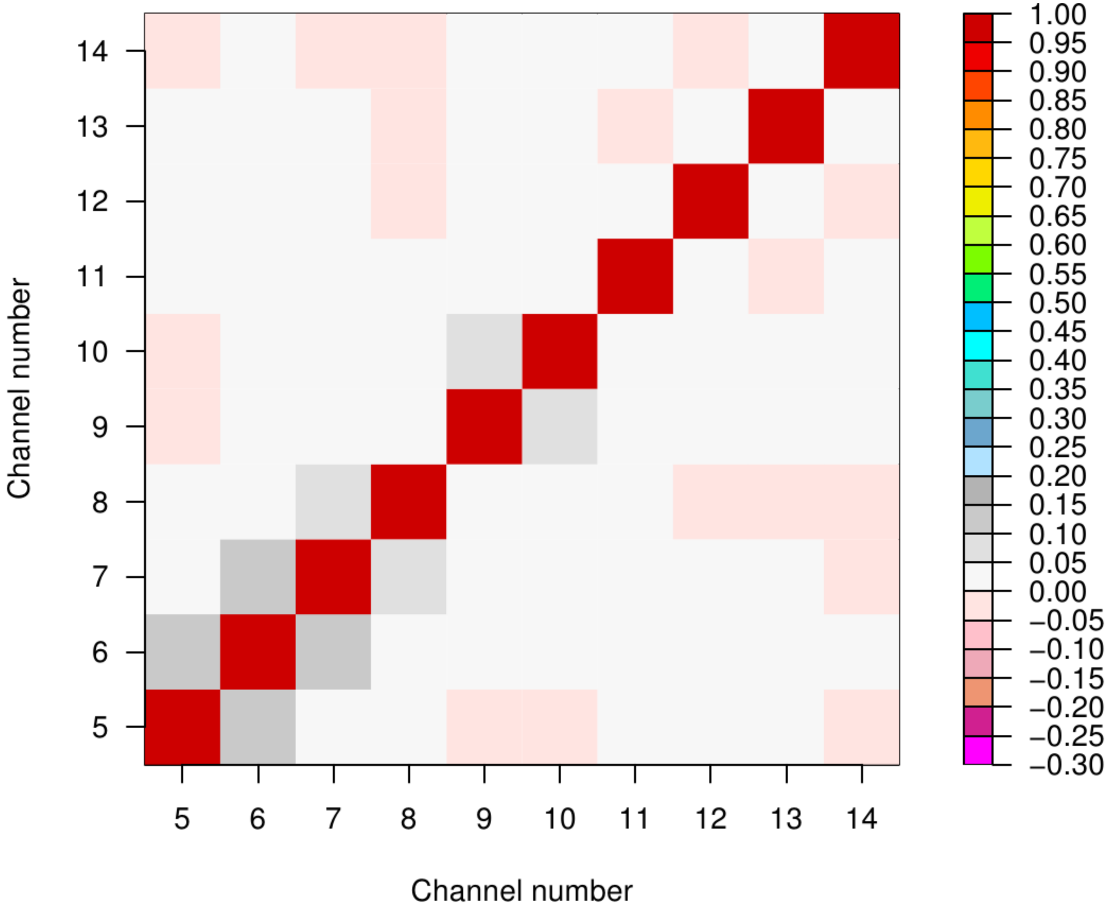
Example: AMSU-A

- **After thinning to 120 km, error diagnostics suggest little correlations…**

Spatial error correlations
for channel 7:

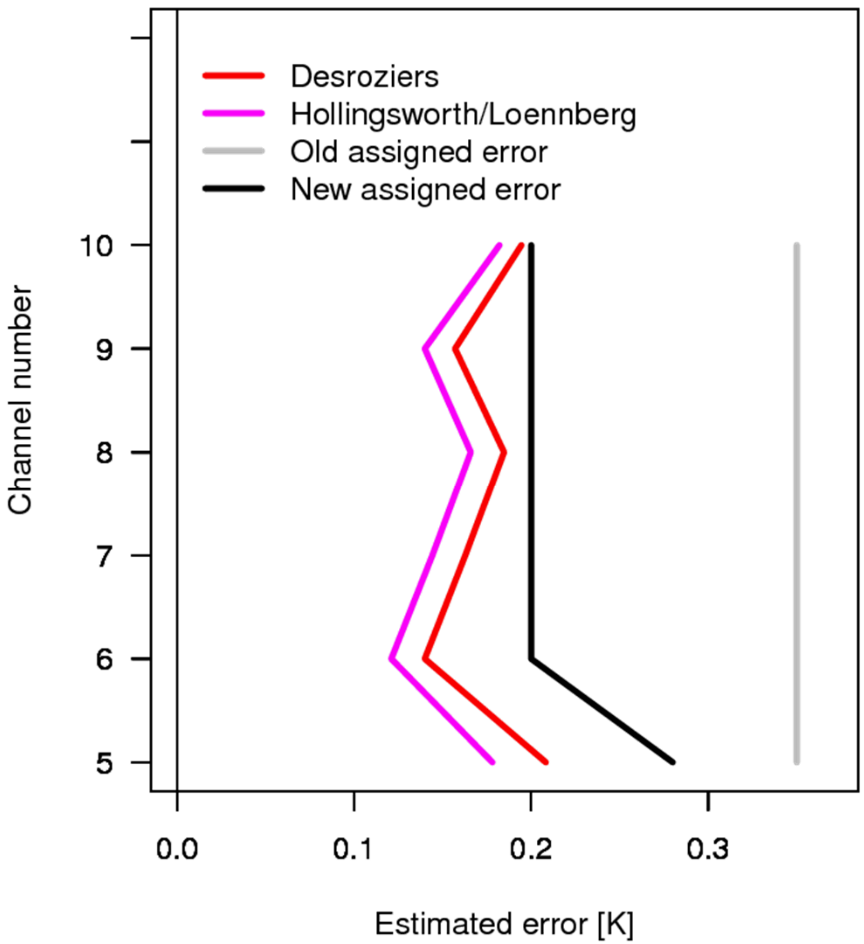


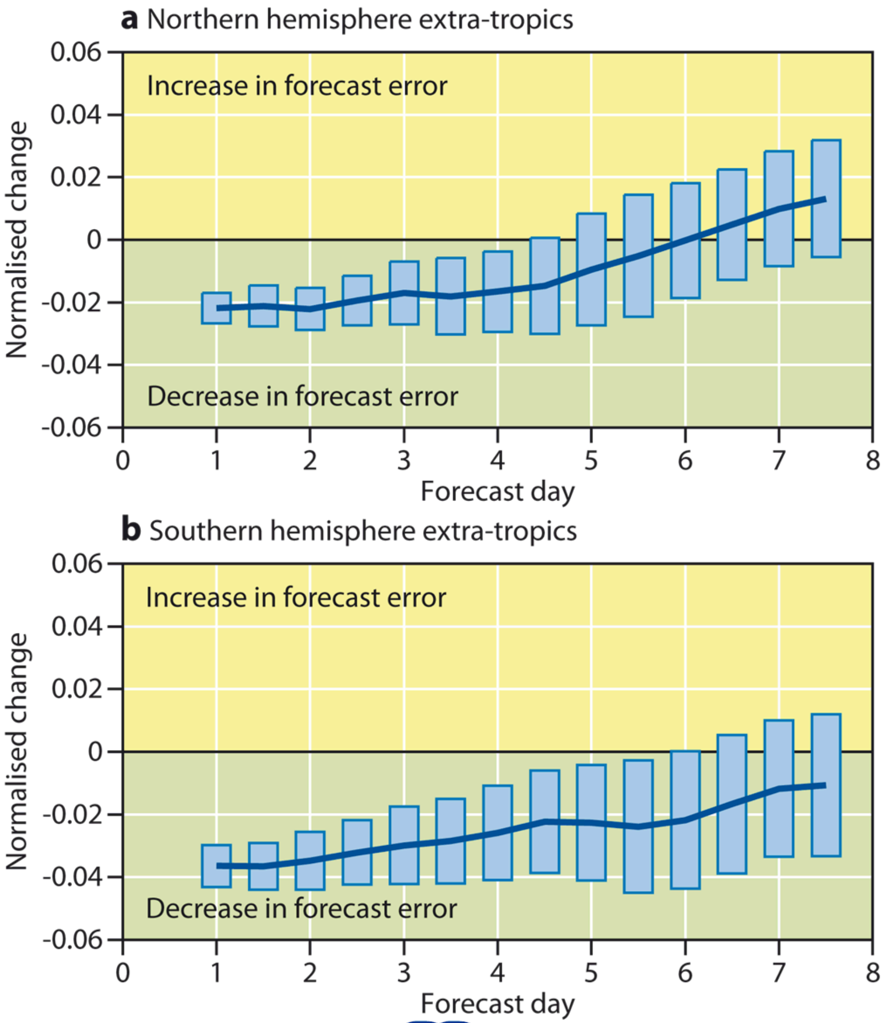
Inter-channel error correlations:



Example: AMSU-A

- … diagonal R a good approximation.

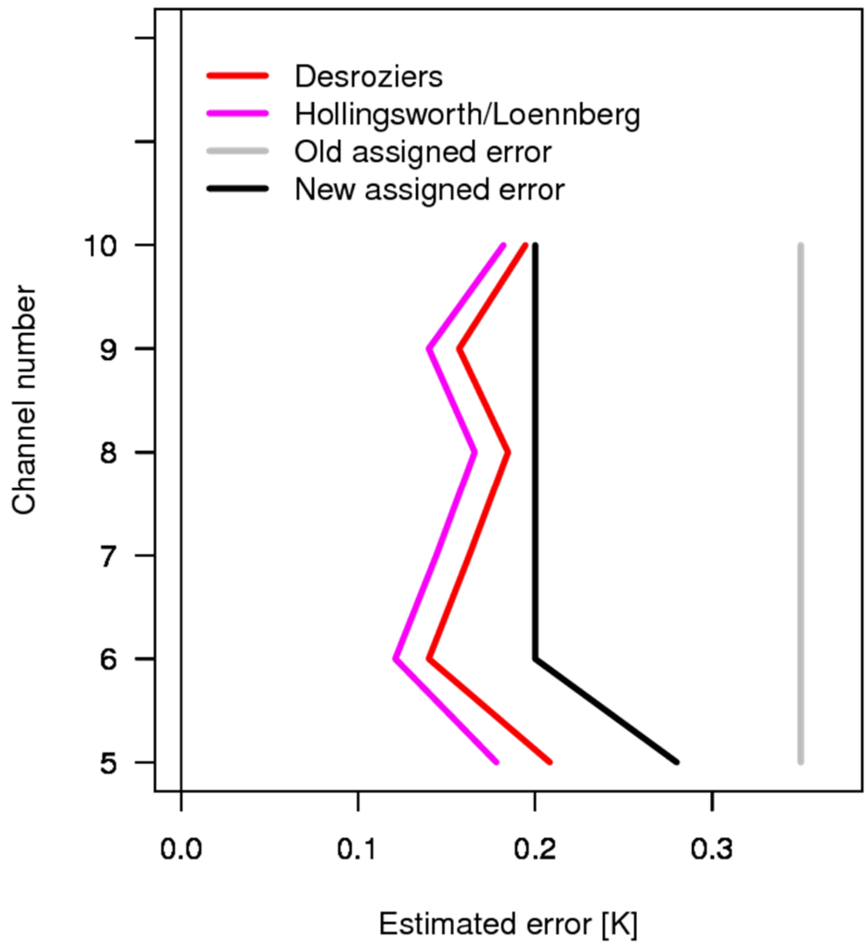


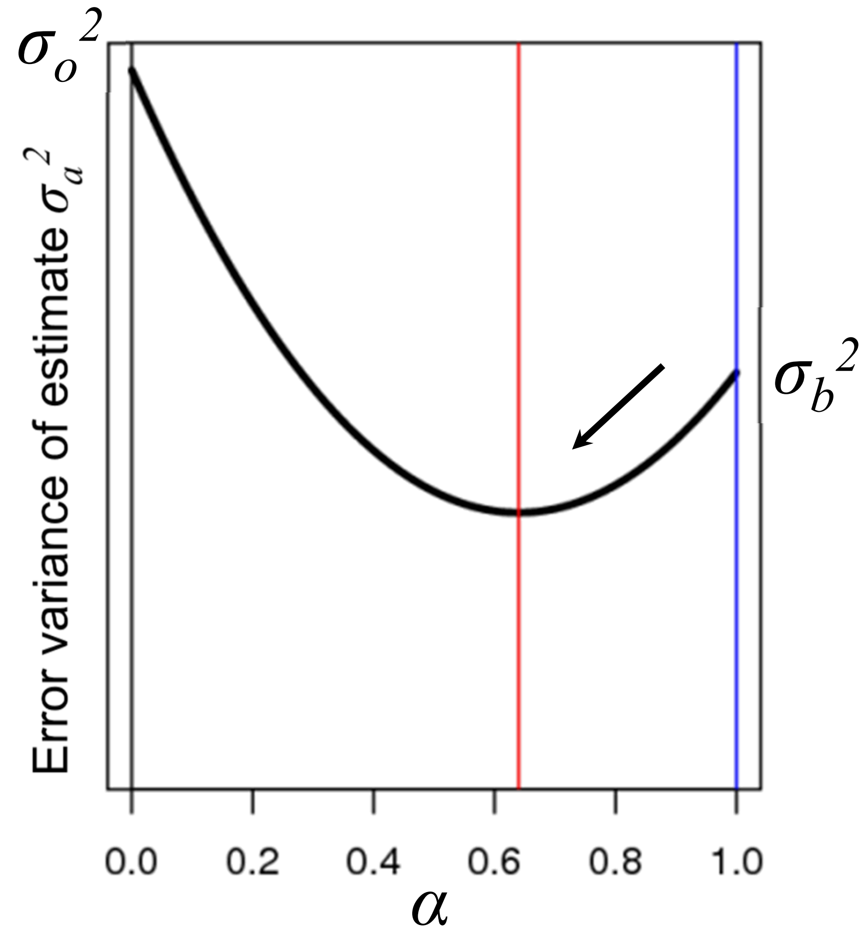


ECMWF

Example: AMSU-A

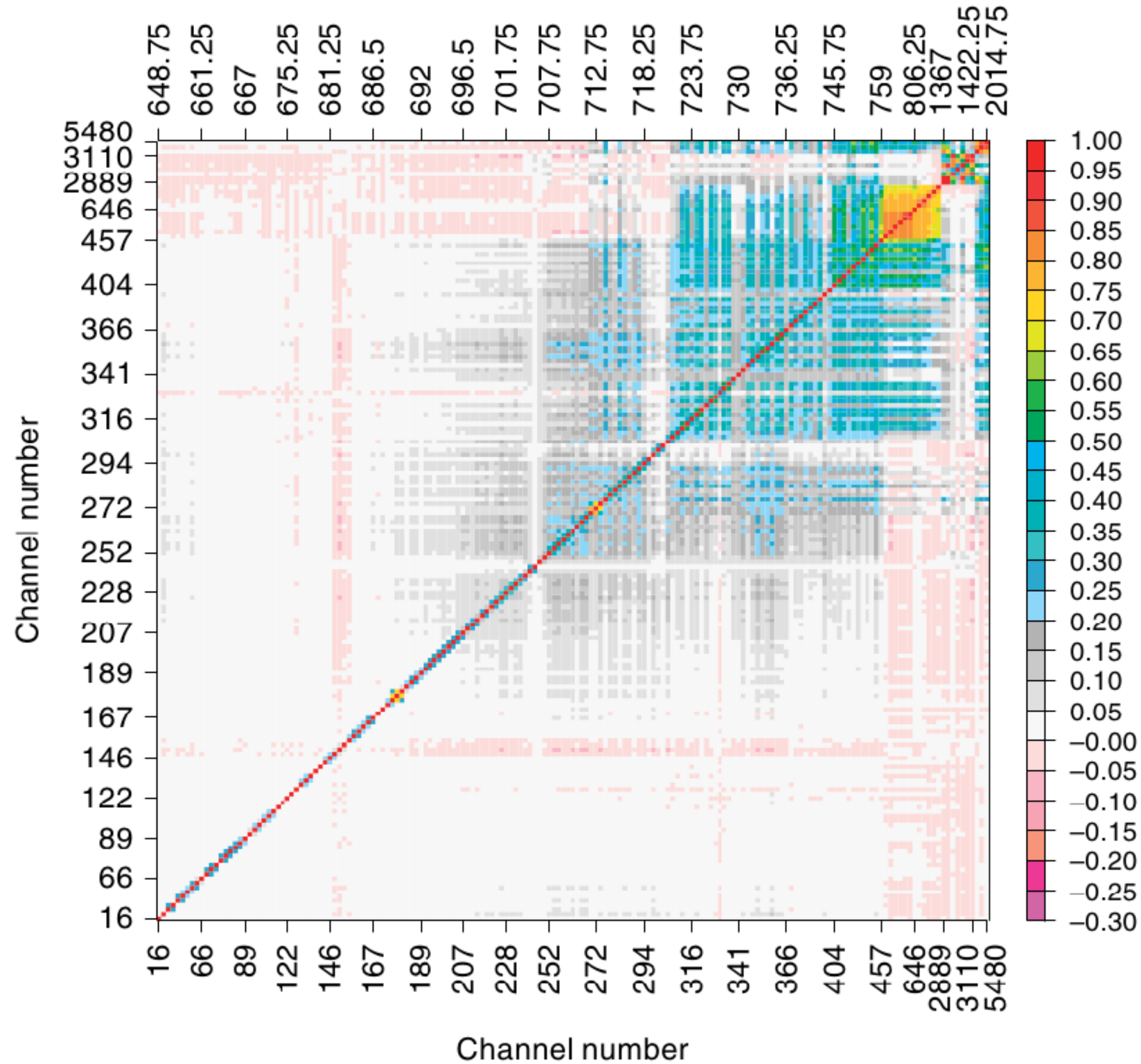
- … diagonal R a good approximation.





Examples of observation error diagnostics: IASI

Inter-channel error correlations:

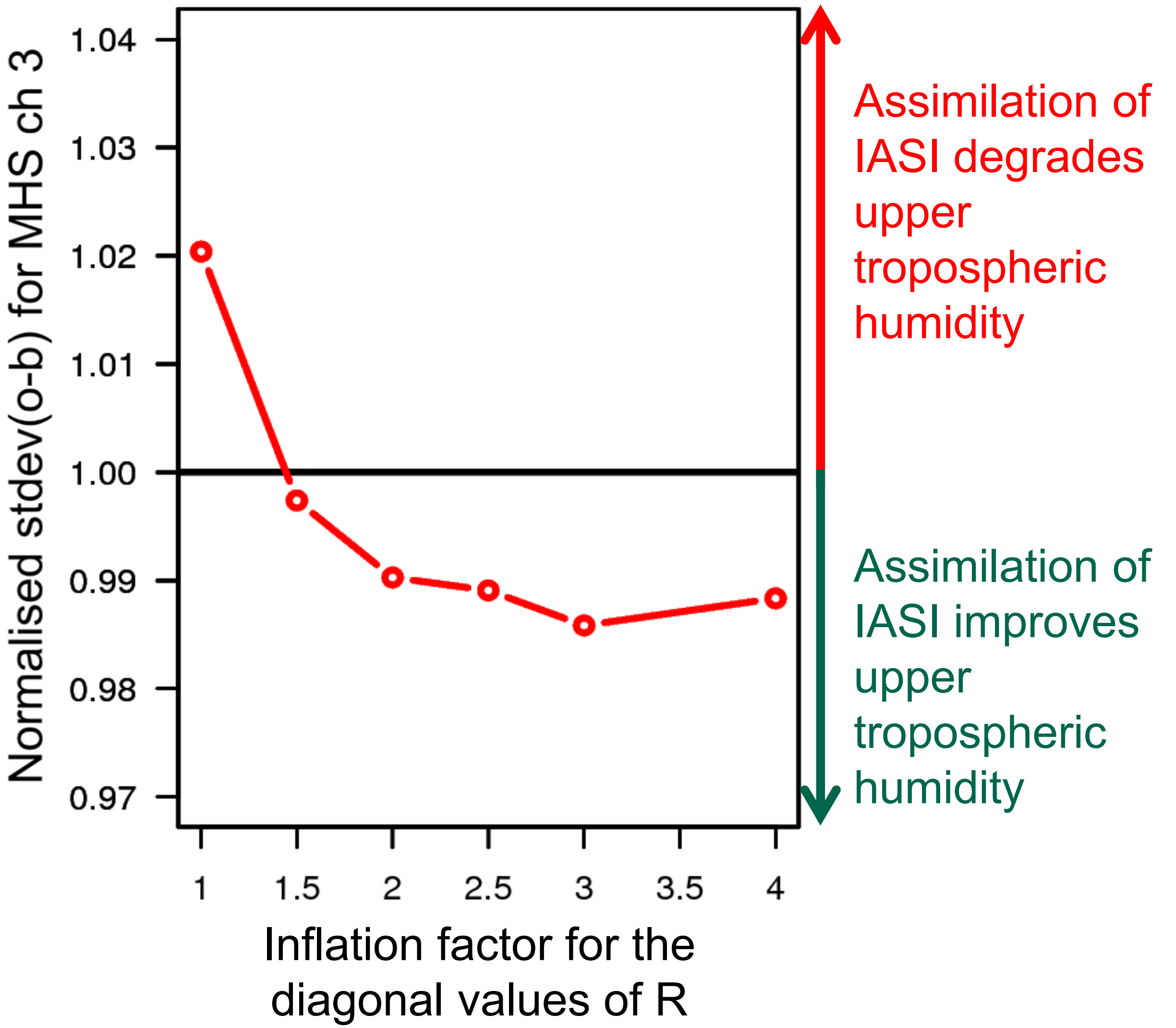


ECMWF

Example: IASI

- **<u>Very common approach</u>: Assume diagonal R, but with larger σ_O than diagnostics suggest ("Error inflation").**

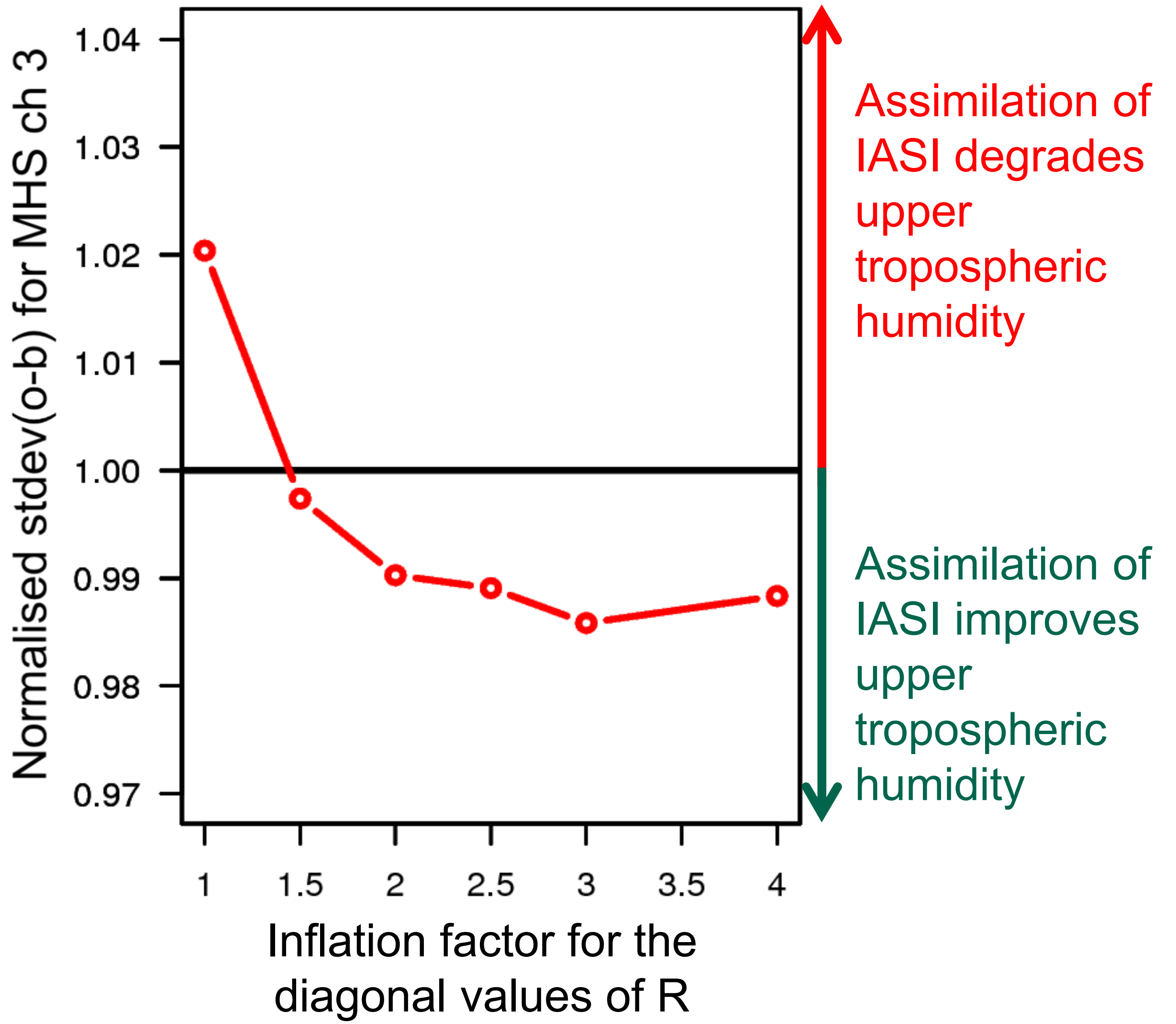
- **Neglecting error correlation with no inflation can result in an analysis that is *<u>worse</u>* than the background!**



ECMWF

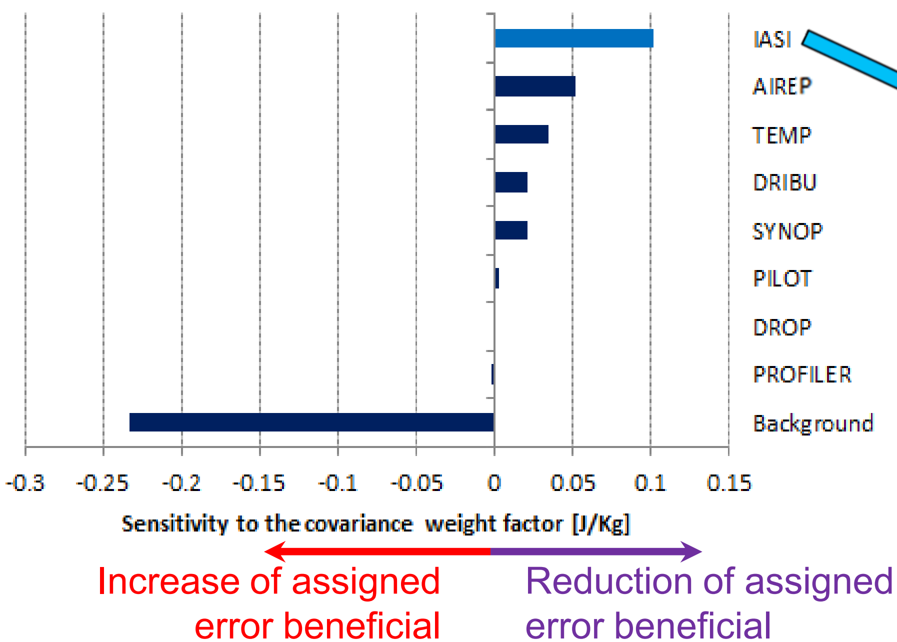
Example: IASI

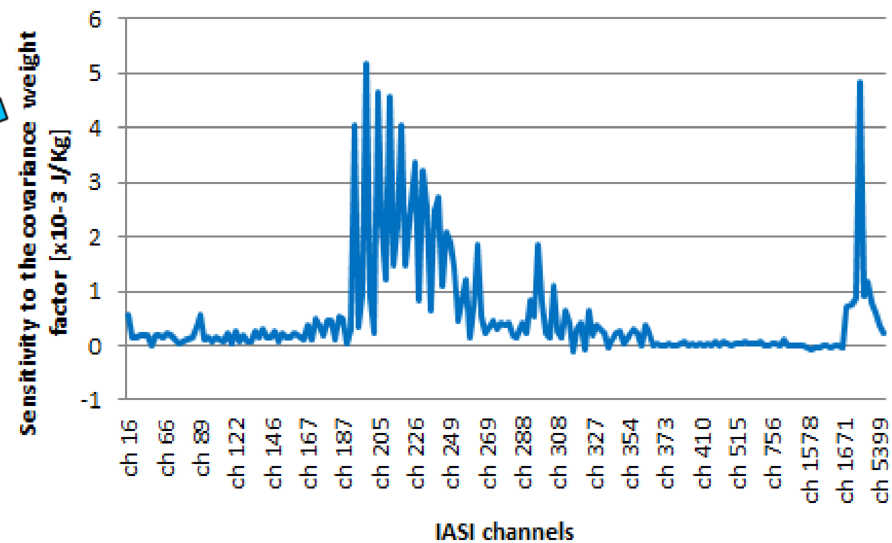
- **Background departure statistics for other observations are a useful indicator to tune observation errors.**



Adjoint diagnostics for observation errors

- Adjoint diagnostics can be used to assess the sensitivity of forecast error reduction to the observation error specification (e.g., Daescu and Todling 2010).
- Example: Assessment of IASI in a depleted observing system with only conventional and IASI data.



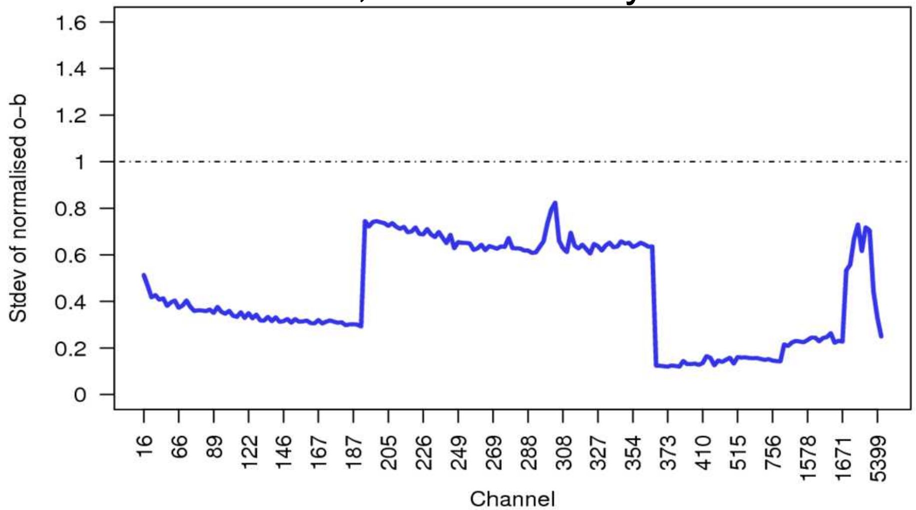


(Cristina Lupu, Carla Cardinali)

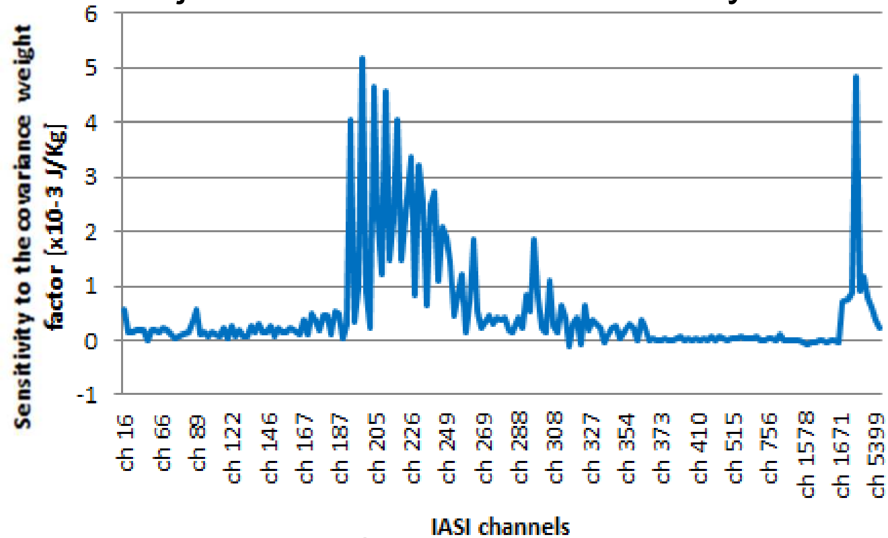
Adjoint diagnostics for observation errors

- **Adjoint diagnostics can be used to assess the sensitivity of forecast error reduction to the observation error specification.**
- **Example: Assessment of IASI in a depleted observing system with only conventional and IASI data.**

Stdev of o-b, normalised by assumed R:



Adjoint-based forecast sensitivity to R:



- **Further work required regarding the applicability of this diagnostic (consistency of results with estimates of true observation errors).**

Outline of lecture

1. What are observation errors?
2. Diagnosing observation errors
3. Specification of observation errors in practice
4. **Accounting for observation error correlations**
5. Summary

Accounting for error correlations

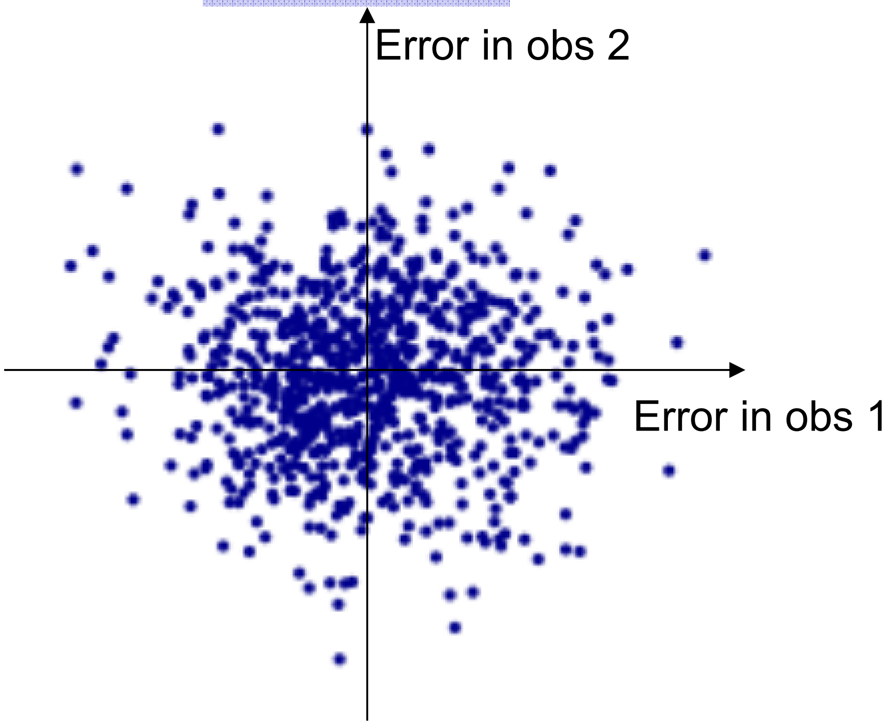
- Accounting for observation error correlations is an area of active research.
- Efficient methods exist if the error correlations are restricted to small groups of observations (e.g., inter-channel error correlations).
  - E.g., calculate $R^{-1} (y - H(x))$ without explicit inversion of R, by using Cholesky decomposition (algorithm for solving equations of the form Az = b).
  - Used operationally for IASI at the Met Office.

- Accounting for spatial error correlations is technically more difficult.

What is the effect of error correlations?

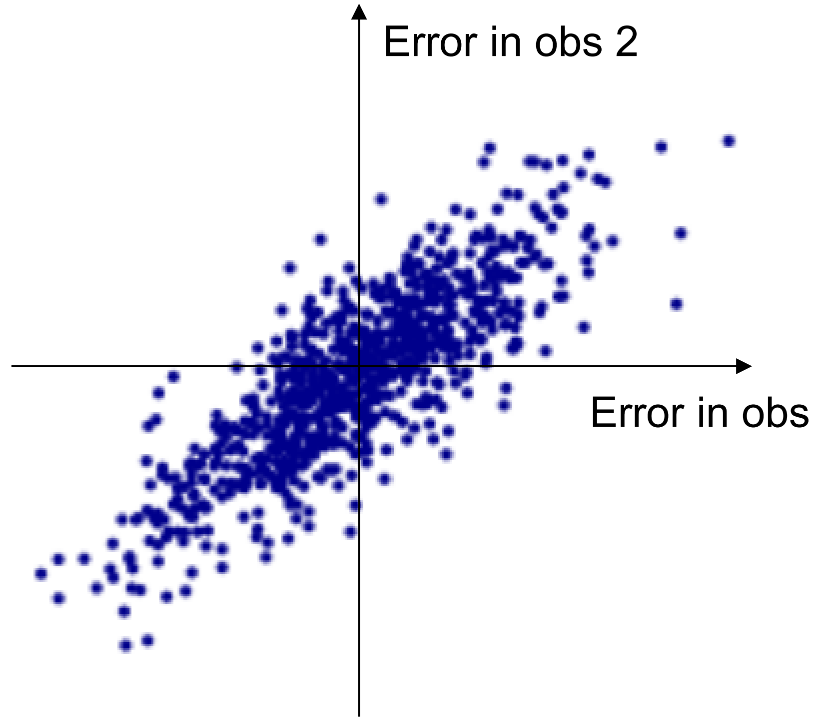
Uncorrelated error

$$\mathbf{R} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$



Correlated error

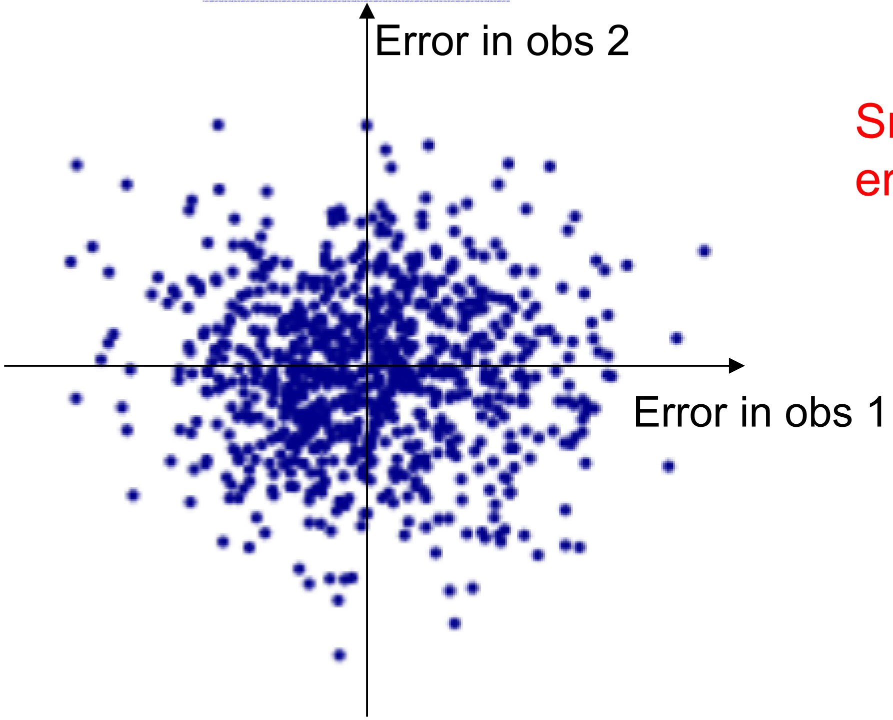
$$\mathbf{R} = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$$



What is the effect of error correlations?

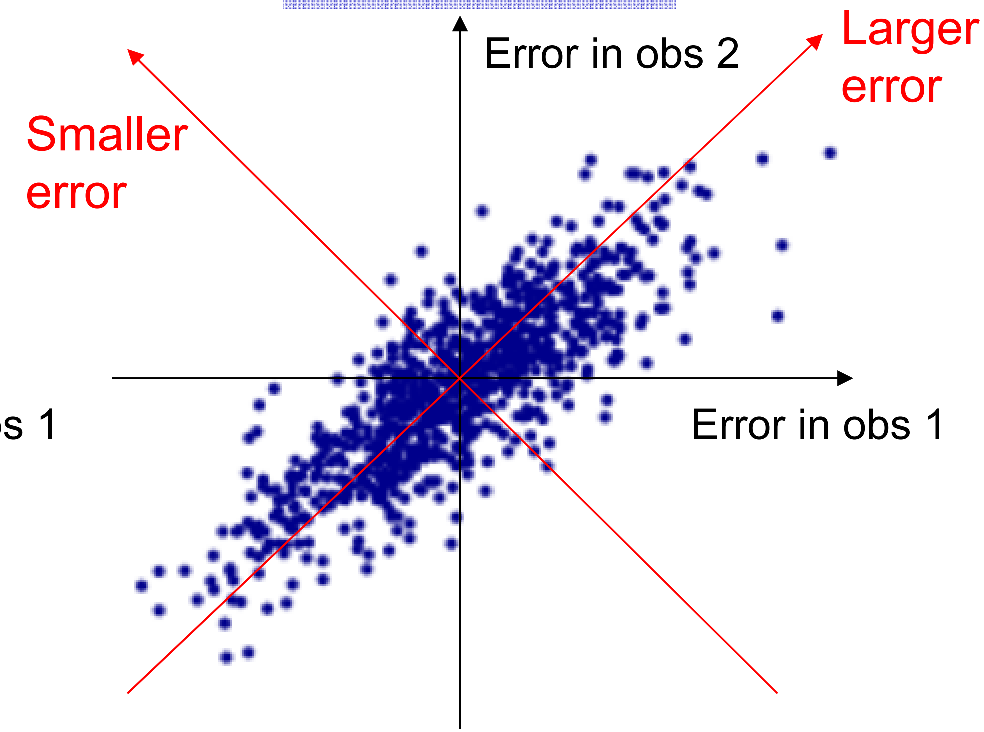
Uncorrelated error

$$\mathbf{R} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

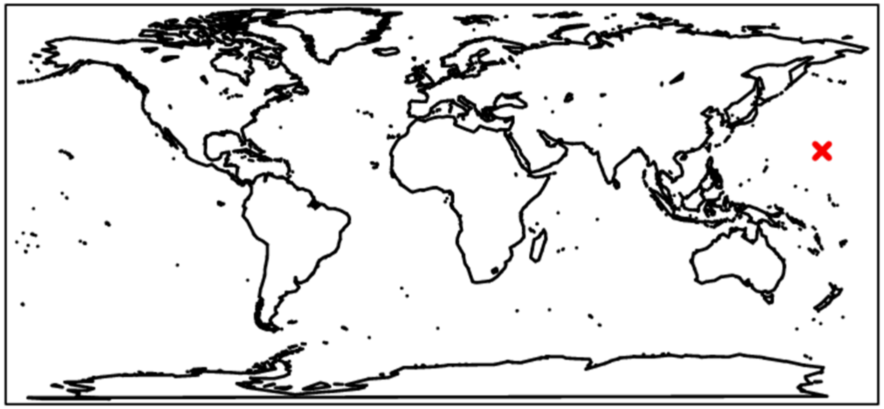


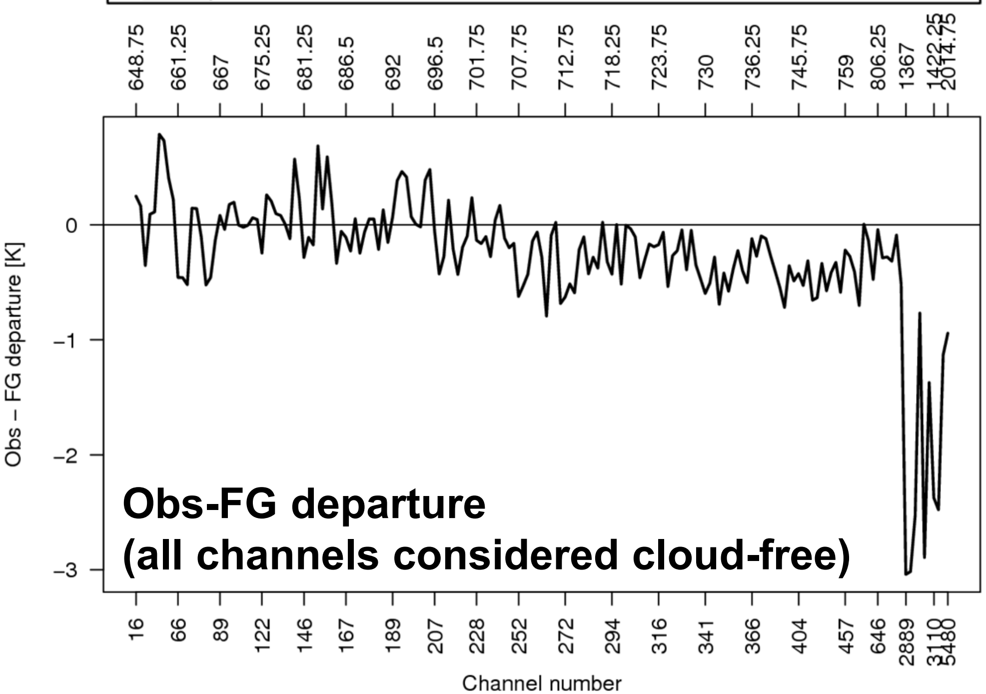
Correlated error

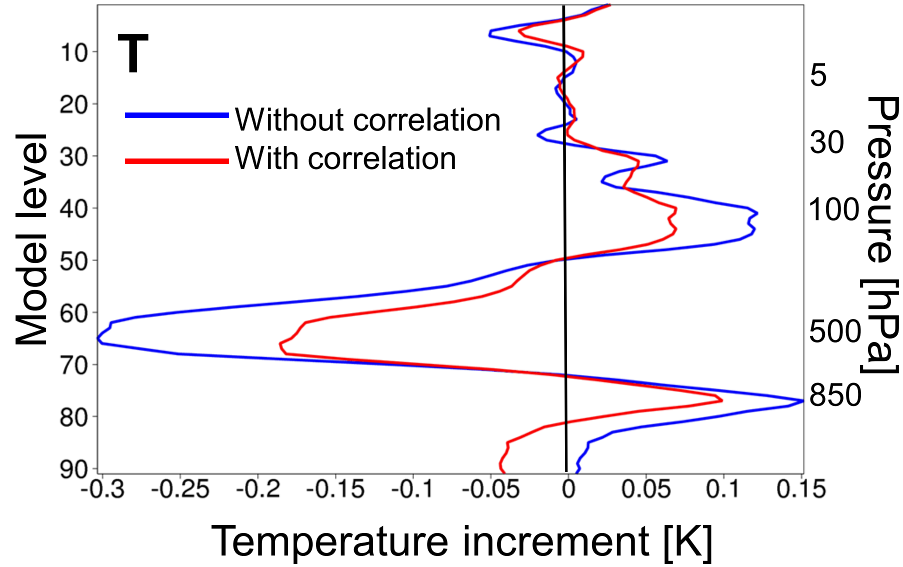
$$\mathbf{R} = \begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$$

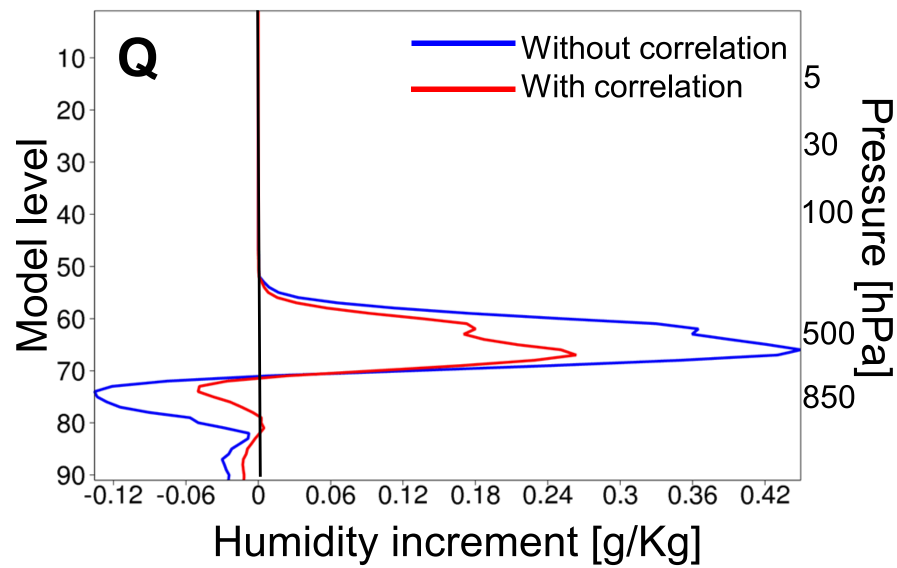


Single IASI spectrum assimilation experiments (I)



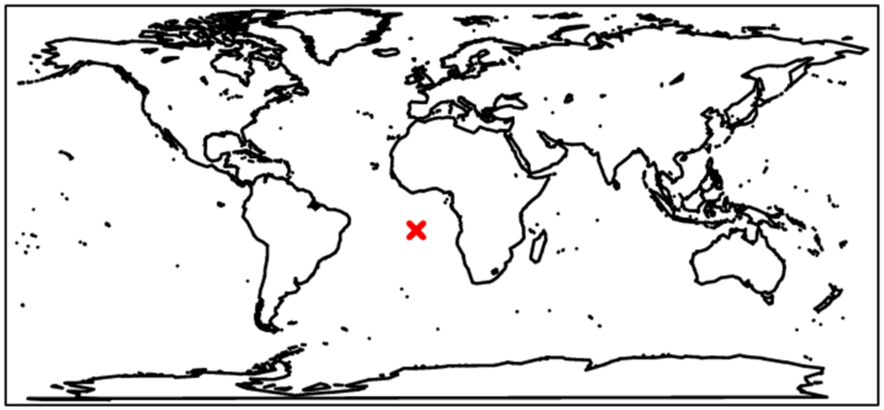


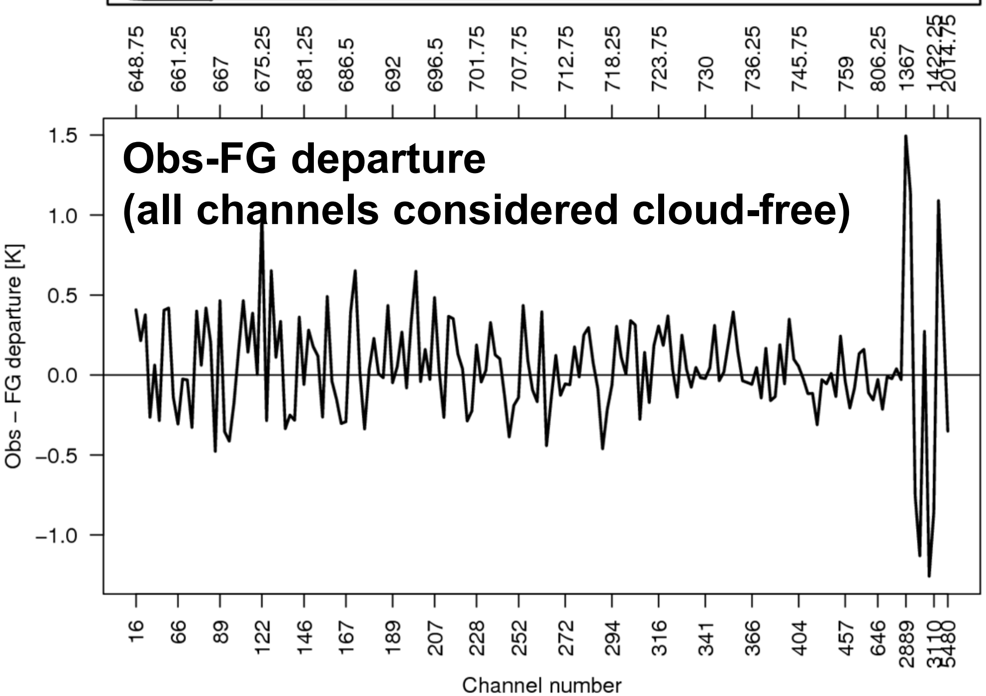


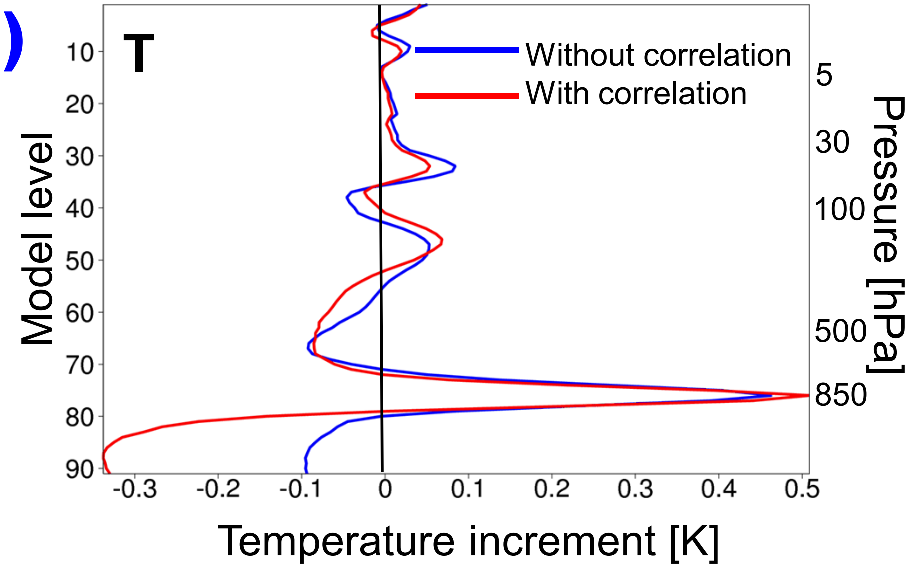


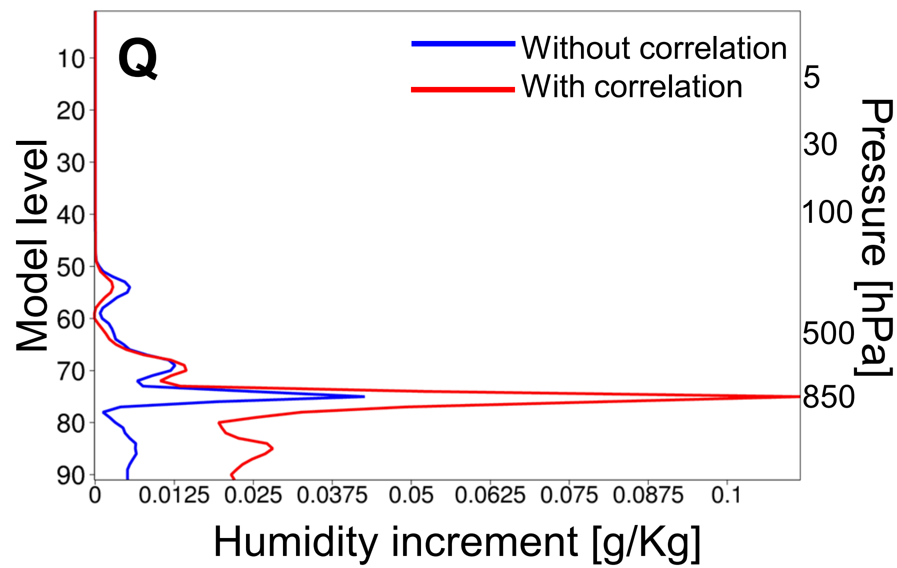
ECMWF

Single IASI spectrum assimilation experiments (II)

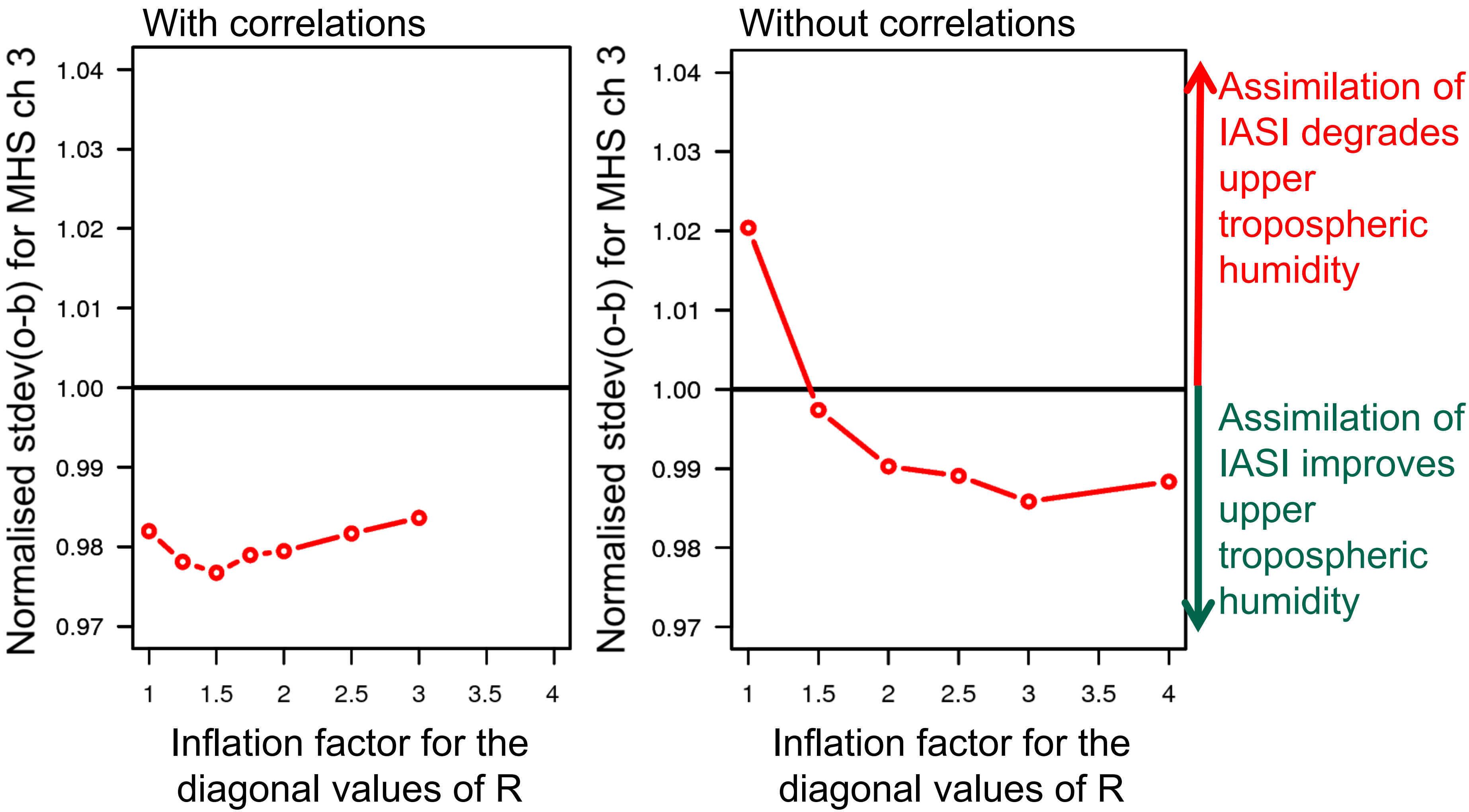








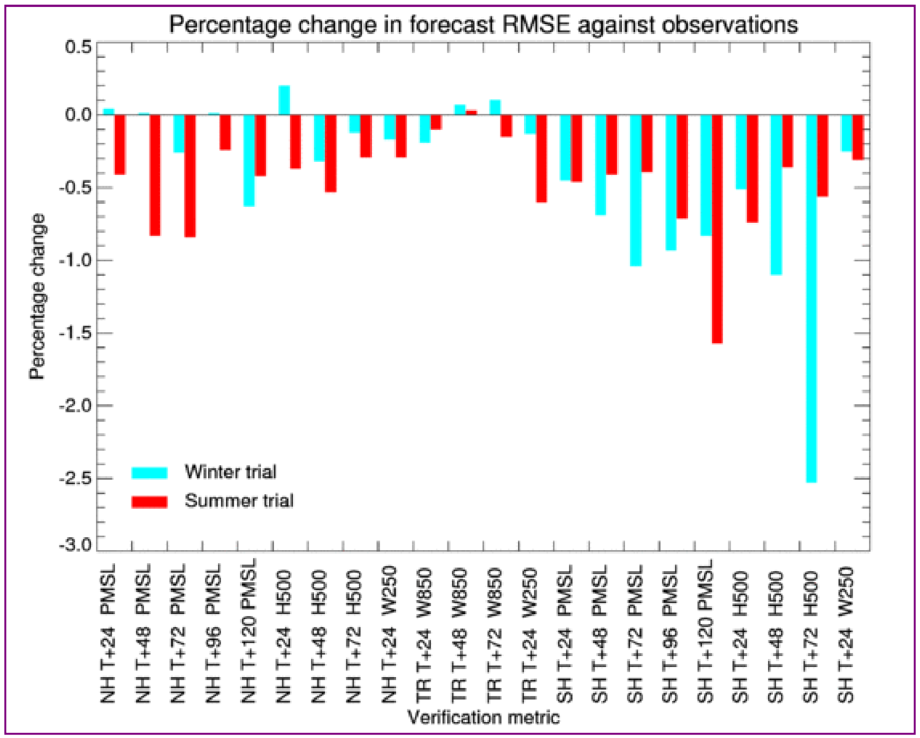
Effect of error correlations on the assimilation of AIRS and IASI



ECMWF

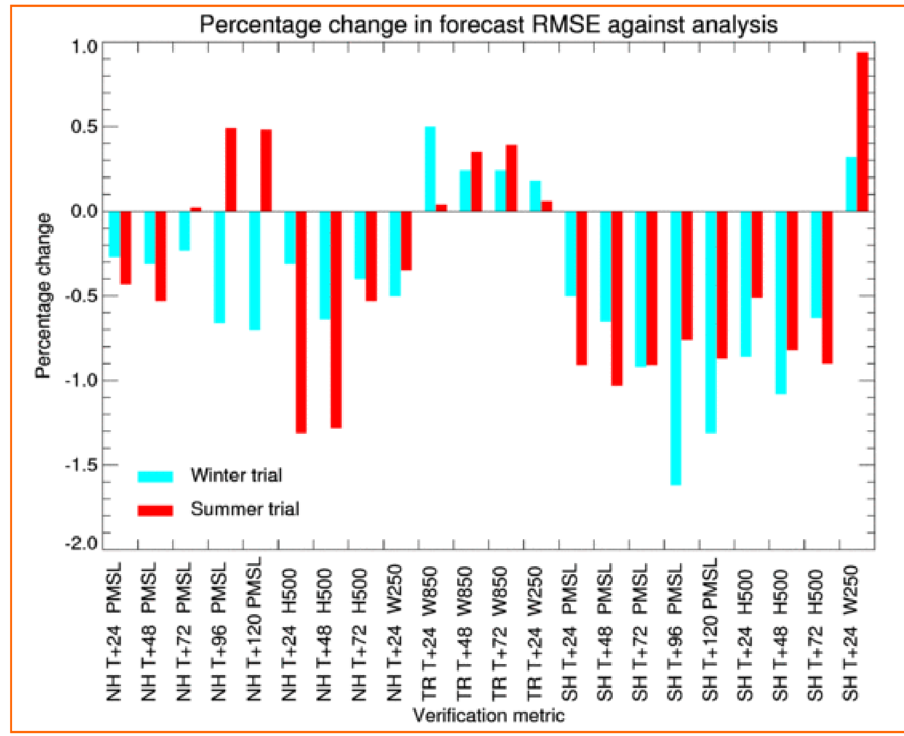
Correlated observation errors for IASI at Met Office

The use of correlated observation errors for IASI was implemented operationally at the Met Office in January 2013, Weston et al 2014.



Verification v Observations

+0.209/0.302% UKMO NWP Index



Verification v Analyses

+0.241/0.047% UKMO NWP Index

Outline of lecture

1. What are observation errors?
2. Diagnosing observation errors
3. Specification of observation errors in practice
4. **Summary**

Summary of main points

- Assigned observation and background errors determine how much weight an observation receives in the assimilation.
- For satellite data, “true” observation errors are often correlated (spatially, in time, between channels).
- Diagnostics on departure statistics from assimilation systems can be used to provide guidance on the setting of observation errors.
- Most systems assume diagonal observation errors, and thinning and error inflation are used widely to counter-act the effects of error correlations.

Further reading (I)

Bormann and Bauer (2010): Estimates of spatial and inter-channel observation error characteristics for current sounder radiances for NWP, part I: Methods and application to ATOVS data. QJRMS, 136, 1036-1050.

Bormann et al. (2010): Estimates of spatial and inter-channel observation error characteristics for current sounder radiances for NWP, part II: Application to AIRS and IASI. QJRMS, 136, 1051-1063.

Daescu, D. N. and Todling, R., 2010: Adjoint sensitivity of the model forecast to data assimilation system error covariance parameters. *Q.J.R. Meteorol. Soc*., 136, 2000–2012.

Desroziers et al. (2005): Diagnosis of observation, background and analysis error statistics in observation space. QJRMS, 131, 3385-3396.

Hollingworth and Loennberg (1986): The statistical structure of short-range forecast errors as determined from radiosonde data. Part I: The wind field. Tellus, 38A, 111-136.

ECMWF

Further reading (II)

Liu and Rabier (2003): The potential of high-density observations for numerical weather prediction: A study with simulated observations. QJRMS, 129, 3013-3035.

Weston et al (2014): Accounting for correlated error in the assimilation of high-resolution sounder data. Q.J.R. Meteorol. Soc., 140: 2420–2429. doi: 10.1002/qj.2306

Cost function diagnostics

- **Consistency diagnostic based on the minimum of the cost function:**
  - **If background and observation errors are correctly specified, it can be shown that**

$$E[J(x_a)] = n$$

with the number of observations *n* and the expectation operator $E[]$ **(see Talagrand 1999).**

  - **Can be used to check/tune assumed error characteristics.**