A review of the evolution of setting observation errors in satellite DA

Niels Bormann

with material from many people

Errors in observations

- **Every observation has an error vs the truth:**
  - **Systematic error**
    - Needs to be removed through bias correction (see Dick Dee's talk)
  - **Random error**
    - Mostly assumed Gaussian in DA.
    - Denoted by the observation error covariance matrix "**R**" in the observation cost function:

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

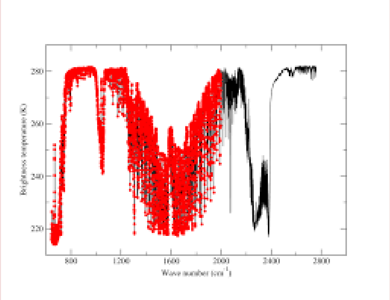
    - Often specified through the square root of the diagonals ("σ_o") and a correlation matrix (which can be the identity matrix).

Contributions to observation error

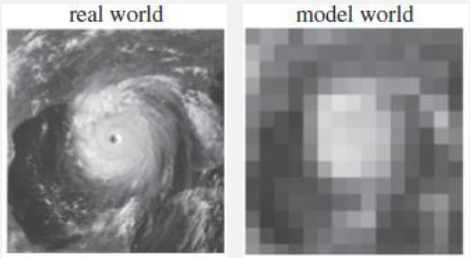
Representation error
(e.g., Janjić et al 2017)

Measurement error
E.g., instrument noise for satellite radiances

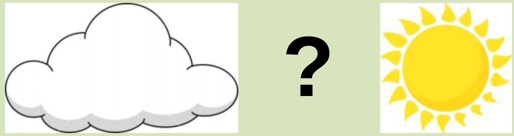
Forward model (observation operator) error
E.g., radiative transfer error



Representativeness error
E.g., point measurement vs model representation



Quality control/pre-processing error
E.g., error due to the cloud detection scheme missing some clouds in clear-sky radiance assimilation



Contributions to observation error

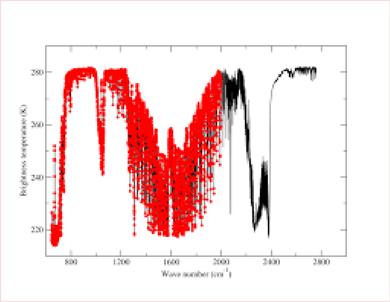
Representation error
(e.g., Janjić et al 2017)

Measurement error

E.g., instrument noise for satellite radiances

Forward model (observation operator) error

E.g., radiative transfer error

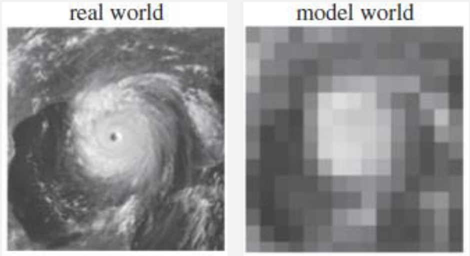


Observation errors can be:

- situation-dependent
- correlated between observations (spatially, temporally, between channels)

Representativeness error

E.g., point measurement vs model representation



Quality control/pre-processing error

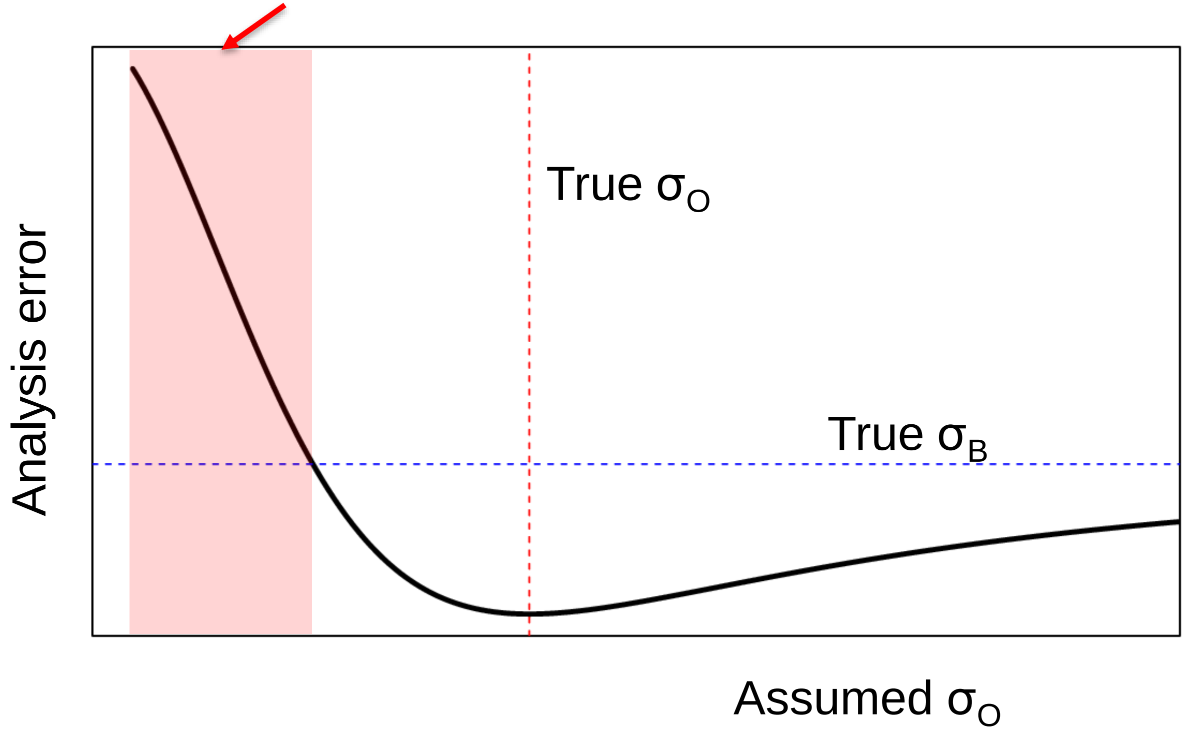
E.g., error due to the cloud detection scheme missing some clouds in clear-sky radiance assimilation

?

Observation error specification 20 years ago

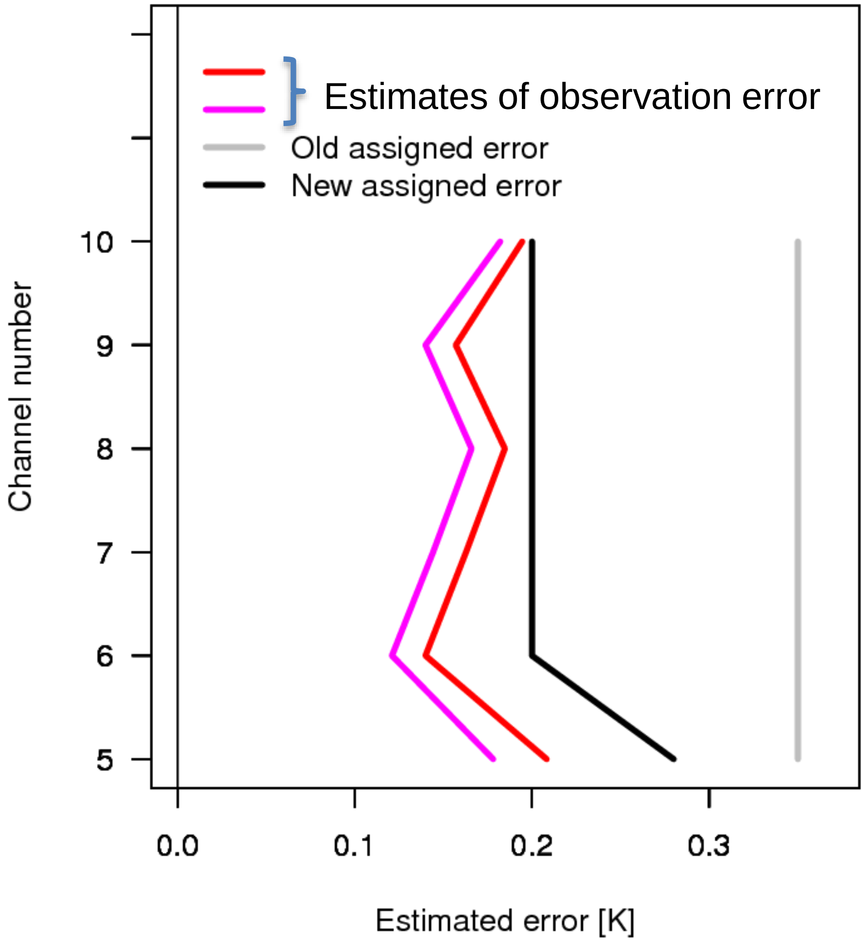
- **R diagonal**, one constant number per channel/level
- **Thin data**, to avoid spatial error correlations

- **Prevailing wisdom: Make σ_O large**
  - To counter-act remaining error correlations
  - To stay away from the **danger zone**

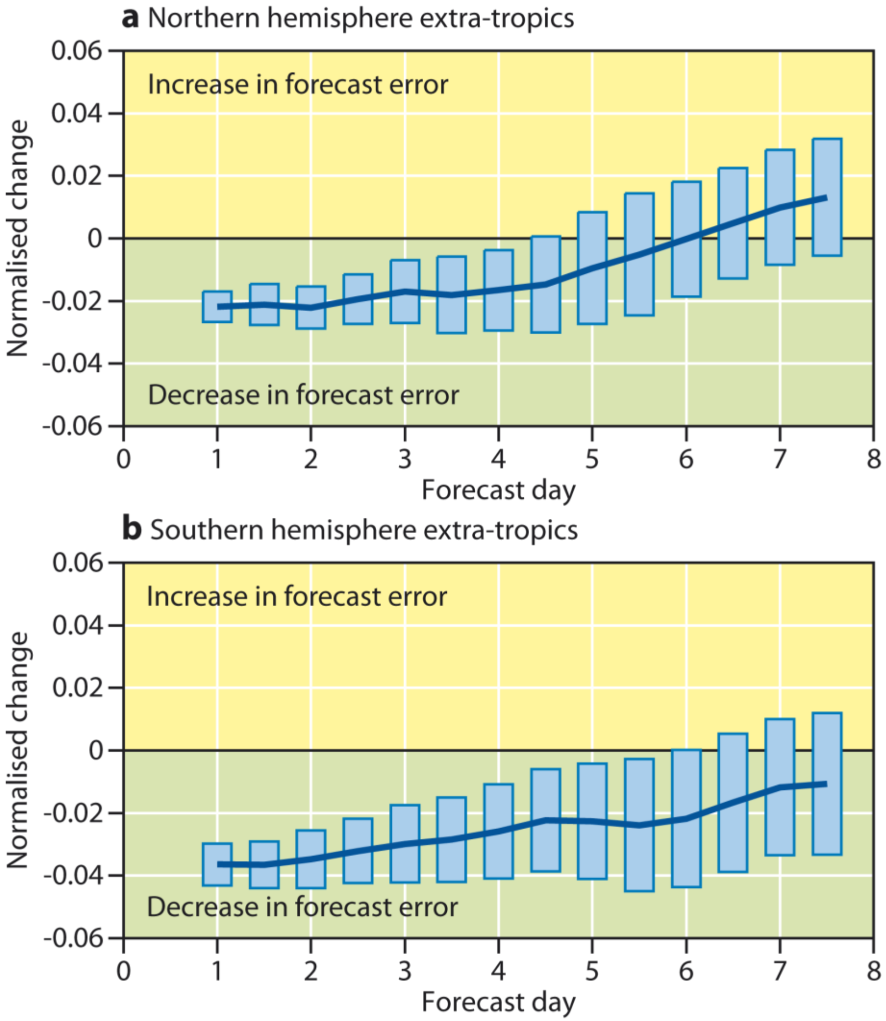


Assigning observation errors matters

AMSU-A observation error revision at ECMWF, 37r2, 2011



Impact on Z500 RMSE



Increased sophistication of observation error assignment

- **Current observation error setting** at ECMWF reflect **two main strands** of recent development in observation error modelling:

Situation-dependent observation errors:

- **AMSU-A:** dependent on satellite, channel, cloudiness, surface emissivity error
- **All-sky error model for MW imagers, MW humidity sounders:** dependent on channel and cloud amount
- **AMVs:** dependent on level and wind shear (and satellite, channel, height assignment method)
- **Aeolus:** based on physically estimated error for each derived wind

Observation errors with inter-channel error correlations taken into account (globally constant):

- **IASI, CrIS**
- **ATMS**
- WV channels from **geostationary imagers**

Outline

Outline

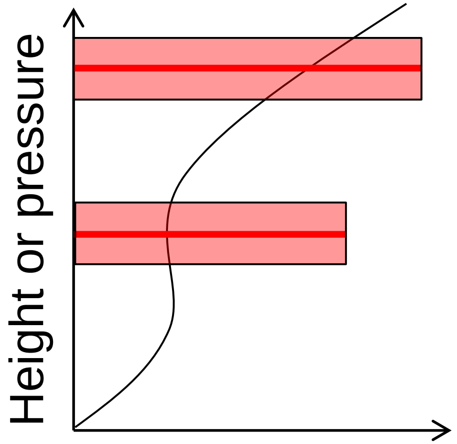
1. Introduction
2. **Situation-dependent observation errors**
3. Correlated observation errors
4. Error inventories and closure studies
5. Summary

Situation-dependence of observation errors

- Observation errors can be situation-dependent, particularly the contributions from representation error.
- To account for this, observation errors are modelled as a function of situation-dependent parameters.

Situation-dependence of observation errors: Example: height-assignment error in AMVs

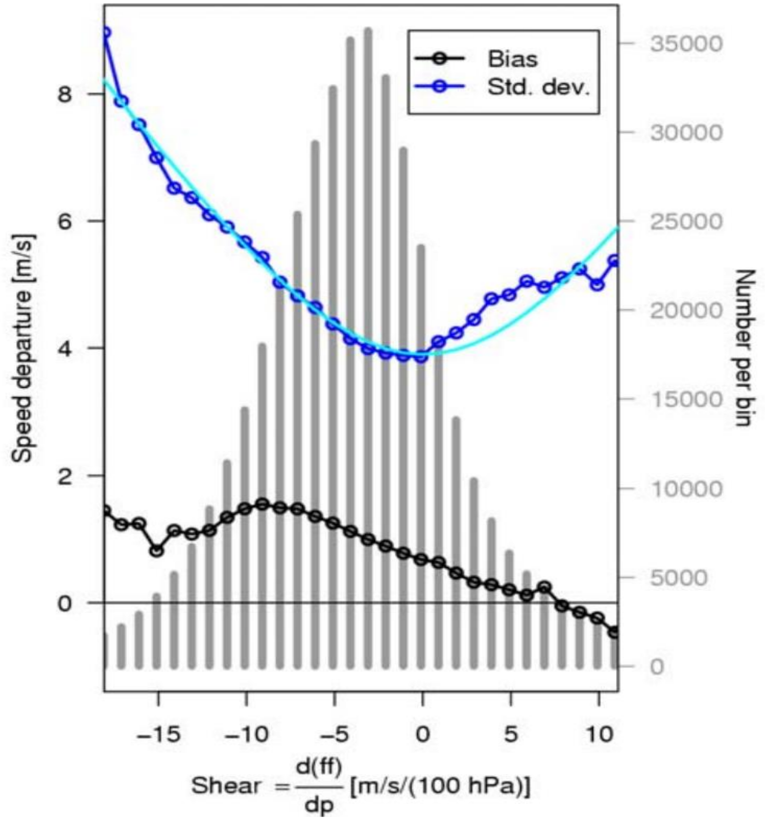
(e.g., Forsythe and Saunders 2008; Bormann et al 2008; Salonen and Bormann 2012)

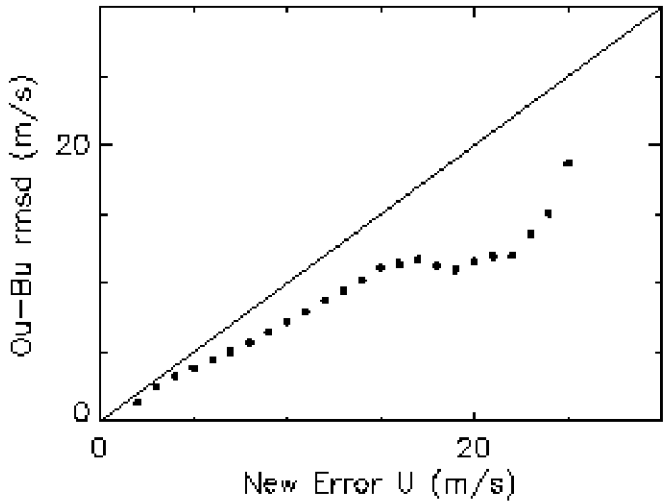


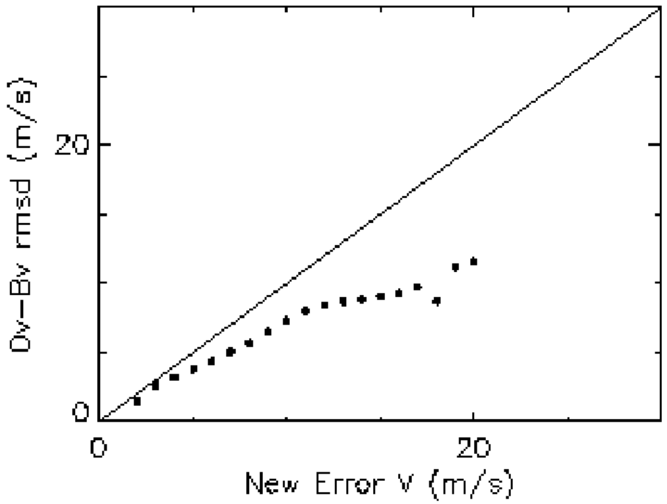
Large **shear** – larger error due to **height assignment error**

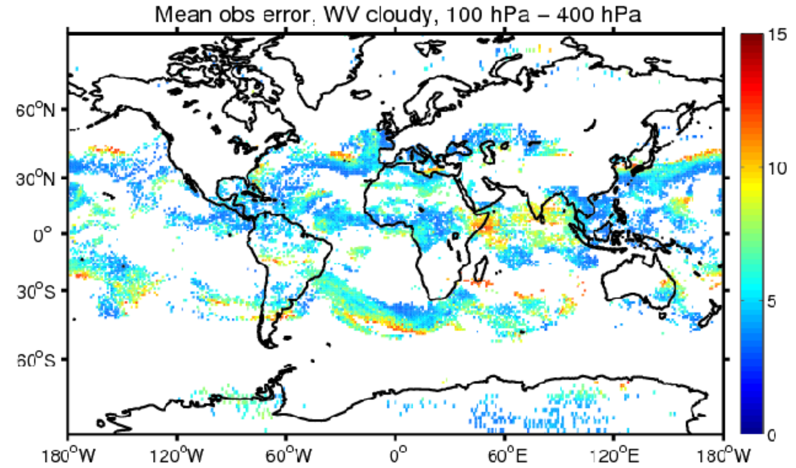
Low **shear** – smaller error due to **height assignment error**

$$(\sigma_O)^2 \approx (\sigma_{O,\,\mathrm{Tracking}})^2 + \left(\sigma_p \frac{dv}{dp}\right)^2$$









→ Mary Forsythe's talk on AMV errors

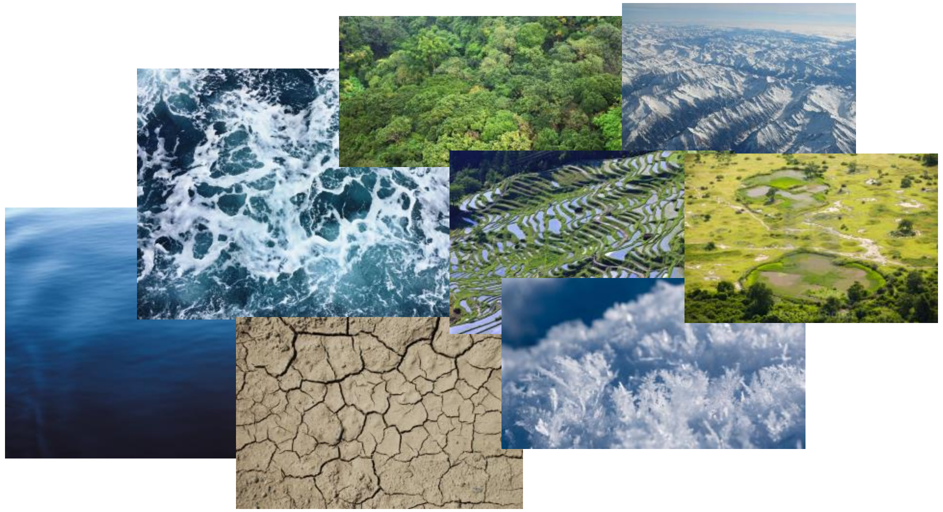
Situation-dependence of observation errors: Example: surface-related errors

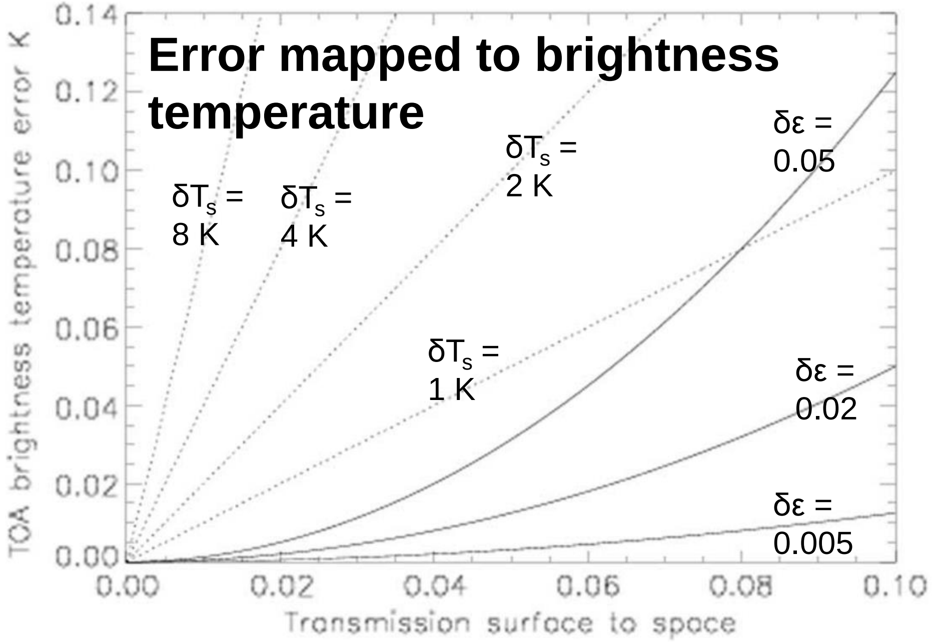
(e.g., English et al 2008;
Lawrence et al 2015; etc)

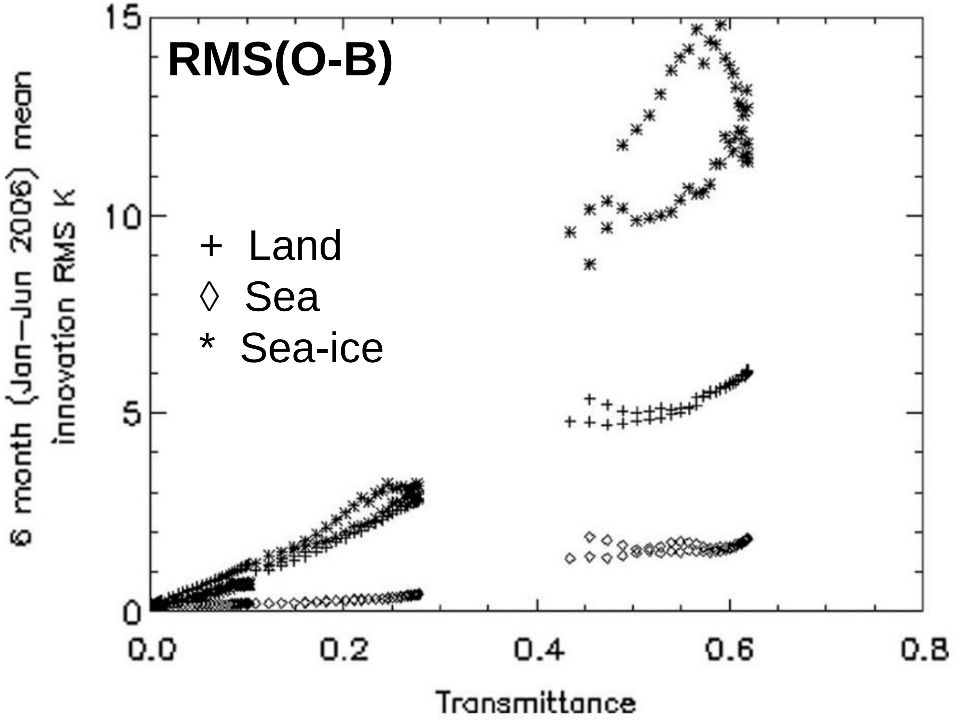
Contributions from emissivity and skin-temperature errors to forward-modelling for surface-sensitive radiances:

$$dI = \epsilon\tau\delta T_s + ((T_s - T)\tau + (T - T_c)\tau^2)\delta\epsilon$$

$$\sigma_O{}^2 = (\sigma_{O\ NeDT})^2 + (dI)2$$



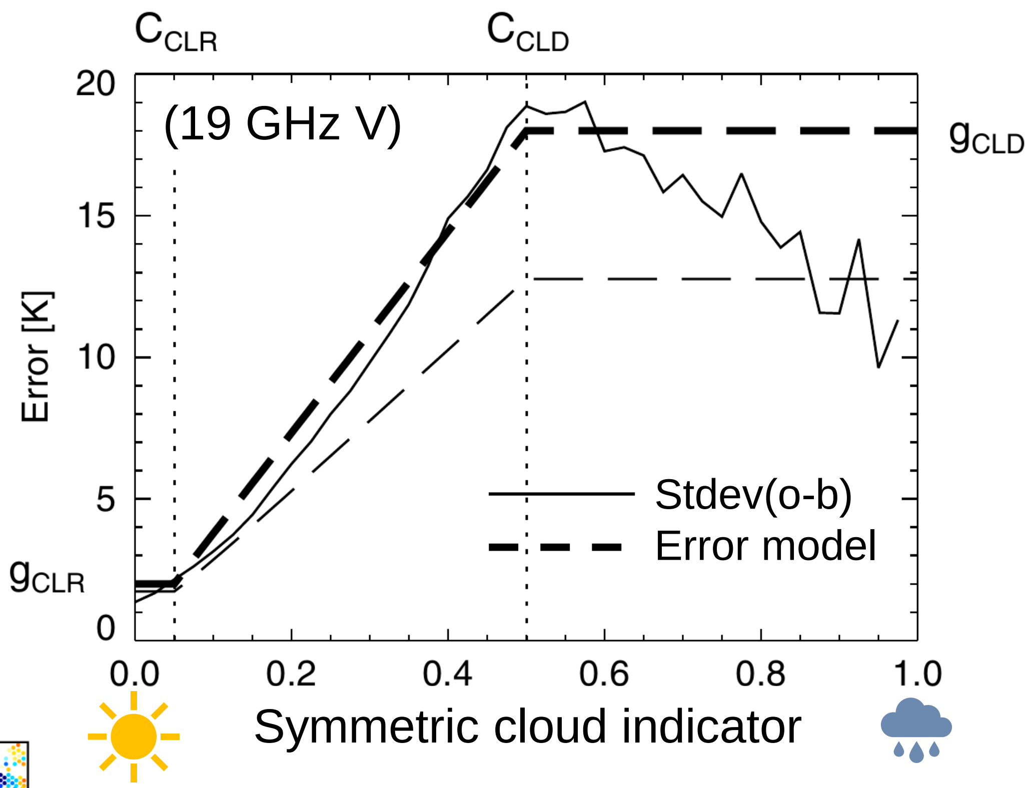


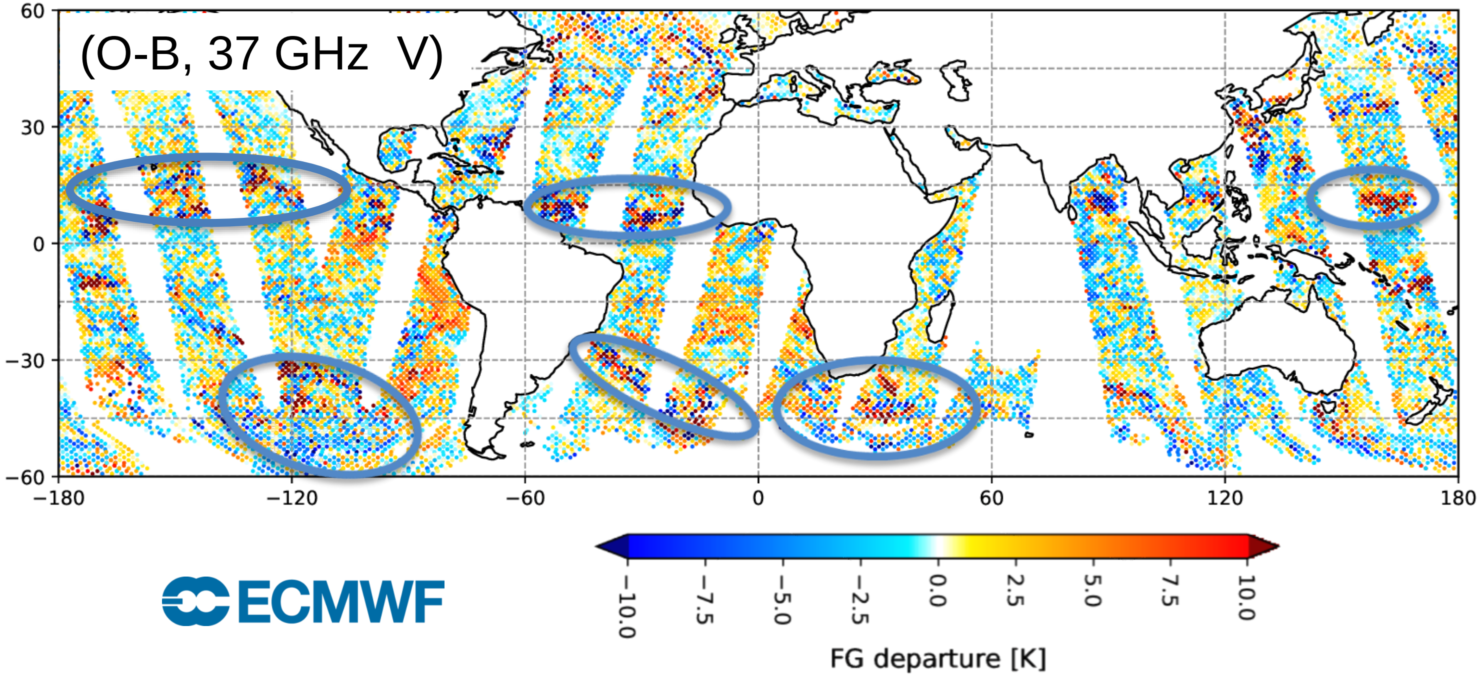


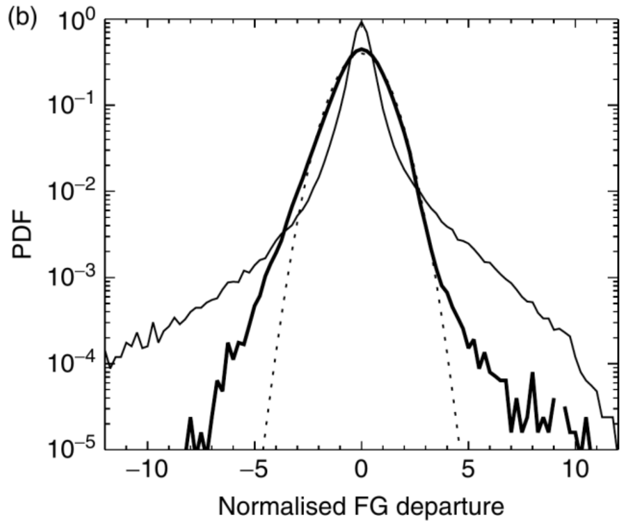
Situation-dependence of observation errors: Example: All-sky assimilation

(e.g., Geer and Bauer 2011; Okamoto et al 2014; Harnisch et al 2016)

Representation error larger in cloudy regions: observation error modelled as function of cloud indicator; observation error model derived from stdev(o-b)







ECMWF

→ Alan Geer's talk on R for all-sky IR

Some remarks on modelling situation-dependent observation errors

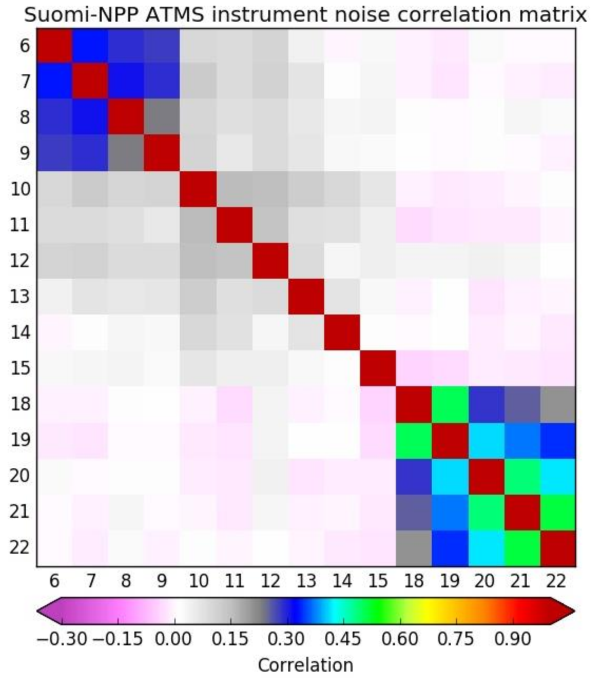
- Current approaches aim to identify and model the main situation-dependent contributions, based on physical considerations
  - Models are mostly specified based on observation departure statistics (stdev(o-b)), with ad-hoc assumptions on the behaviour of background errors.
    - How valid are the underlying assumptions on background errors?
    - Scope for more independent specification of error sources?
- What situation-dependent variations are we currently missing?
  - E.g., convective vs stratiform clouds in all-sky; larger errors in H for obs at the end of 4D-Var window?
  - What level of sophistication is useful and desirable for situation-dependent observation errors? What can we model reliably?

Outline

1. Introduction
2. Situation-dependent observation errors
3. **Correlated observation errors**
4. Error inventories and closure studies
5. Summary

Observation error correlations

- **Representation error** is likely to be **correlated** between different observations, e.g.:
  - An error in cloud detection is likely be similar for other channels with similar cloud-sensitivity in clear-sky assimilation.
  - A radiative transfer error is likely to be similar for spectrally-similar channels.
  - A height-assignment error for AMVs is likely to be similar for neighbouring AMVs derived from a similar cloud.
- And even instrument noise can be correlated between channels:



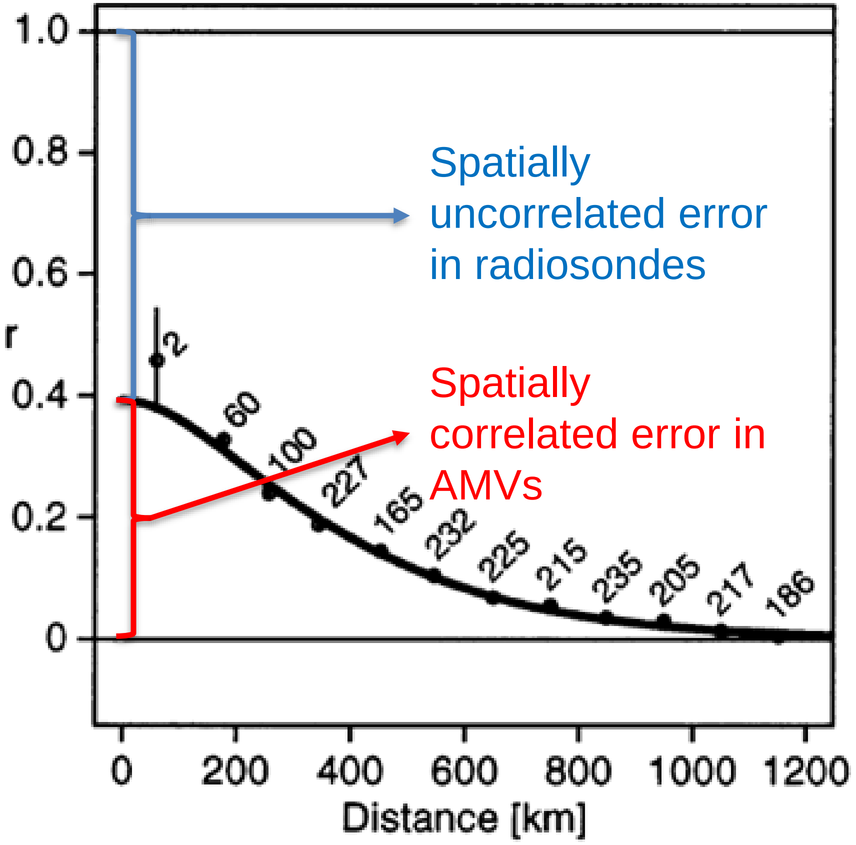
ATMS instrument noise correlation, from independent instrument characterisation.

Estimating spatial error correlations for AMVs

(e.g., Bormann et al 2003)

- Estimated using a Hollingsworth/Lönnberg approach:
  - Use pairs of collocated AMVs & radiosondes.
  - Assume errors in radiosondes uncorrelated.

Correlations between AMV/radiosonde differences



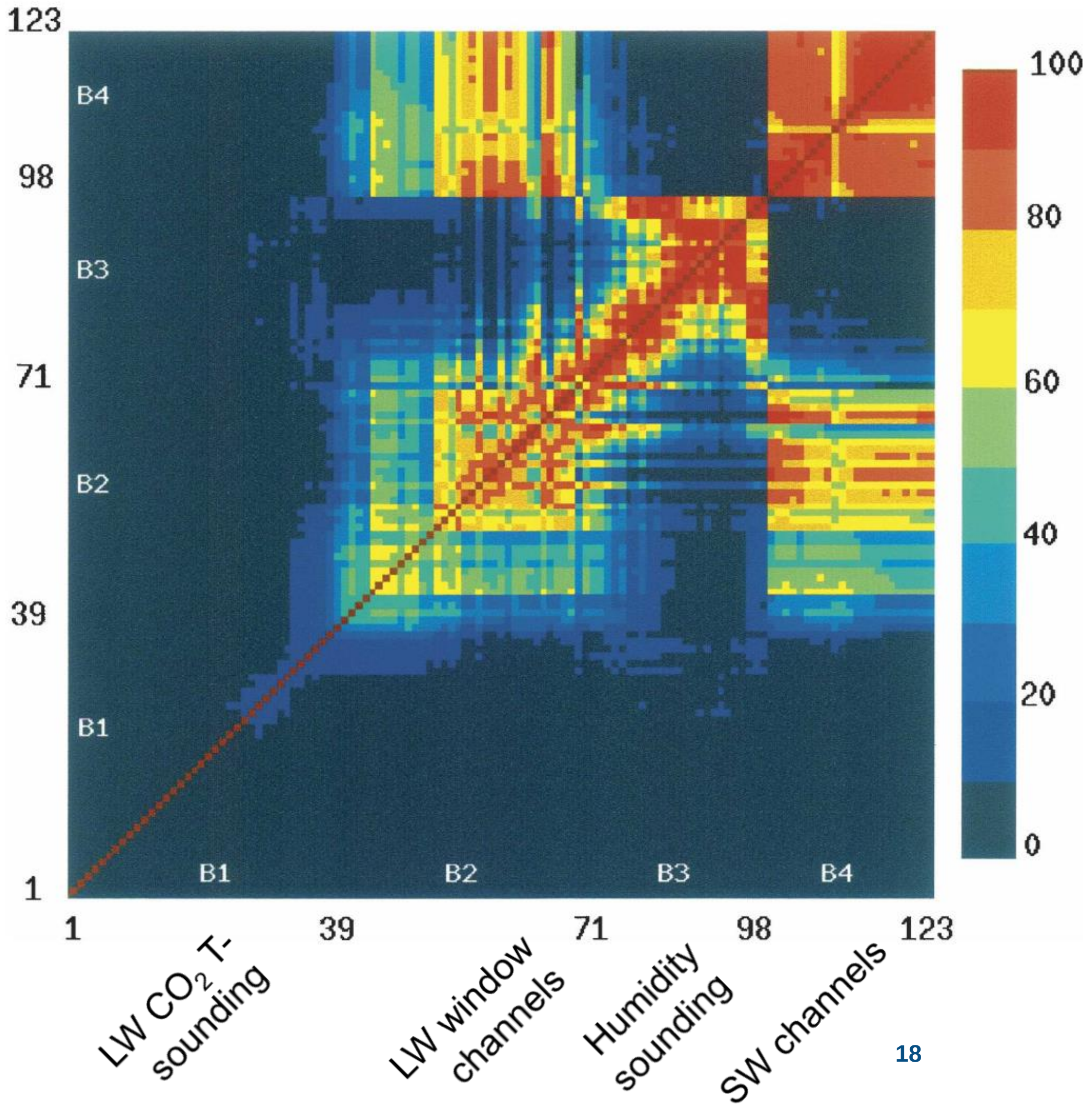
Estimating inter-channel error correlations for hyper-spectral IR observations

(e.g., Garand et al 2007)

- Estimated using a Hollingsworth/Lönnberg approach:
    - Use pairs of o-b for AIRS.
    - Assume AIRS observation errors are spatially uncorrelated.

- Possible source of error correlation:
    - Cloud detection
    - Spatial representativeness
    - Radiative transfer

Diagnosed error correlations for AIRS [%]



Estimating inter-channel error correlations for hyper-spectral IR observations and the Desroziers diagnostic

(e.g., Desroziers et al 2005)

- **Basic assumptions:**
  - Linear estimation theory; errors in observation and background uncorrelated.
  - Weights used in the assimilation system are consistent with true observation and background errors.
- Then the following relationship can be derived:

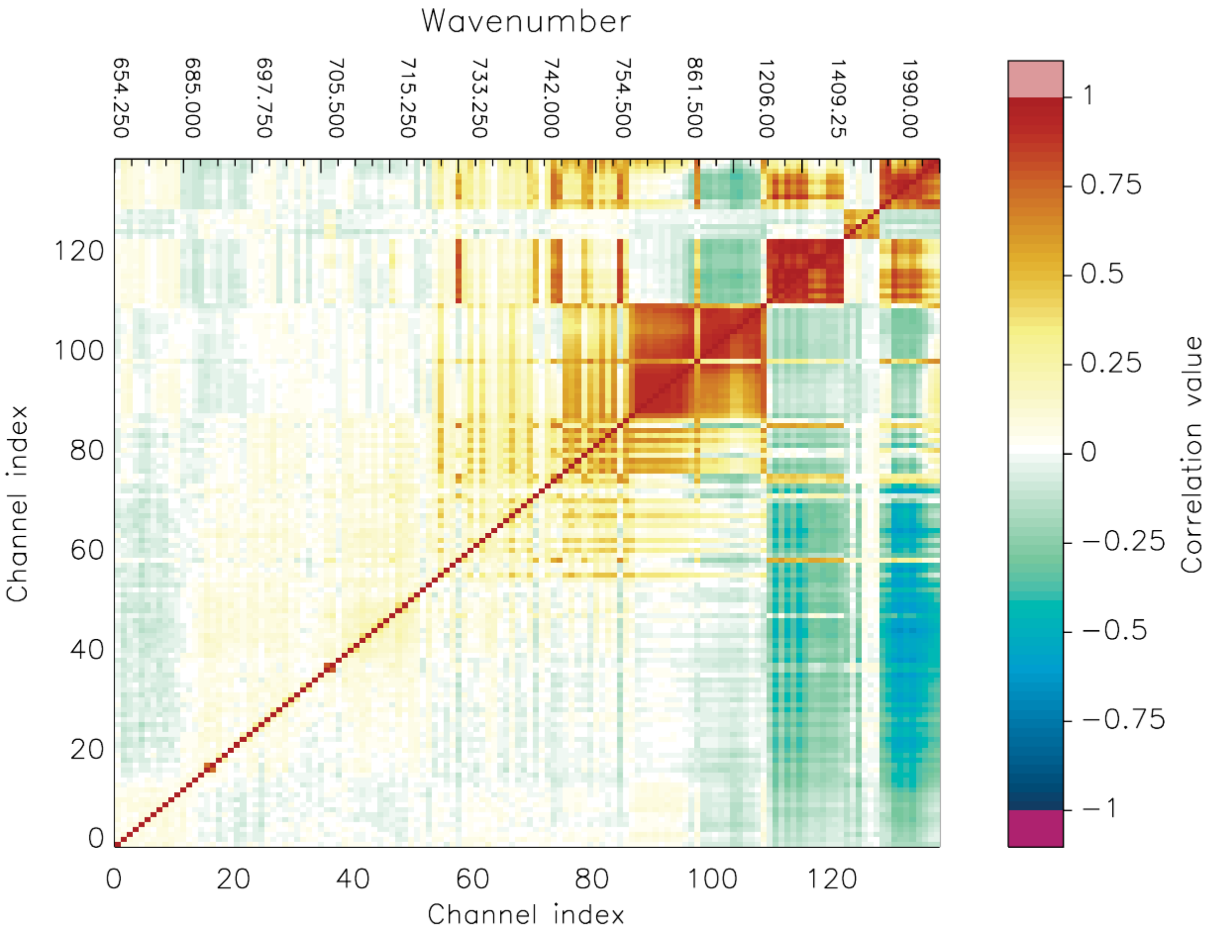
$$\mathbf{R} \quad = \quad Cov[\mathbf{d}_a, \mathbf{d}_b]$$

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)

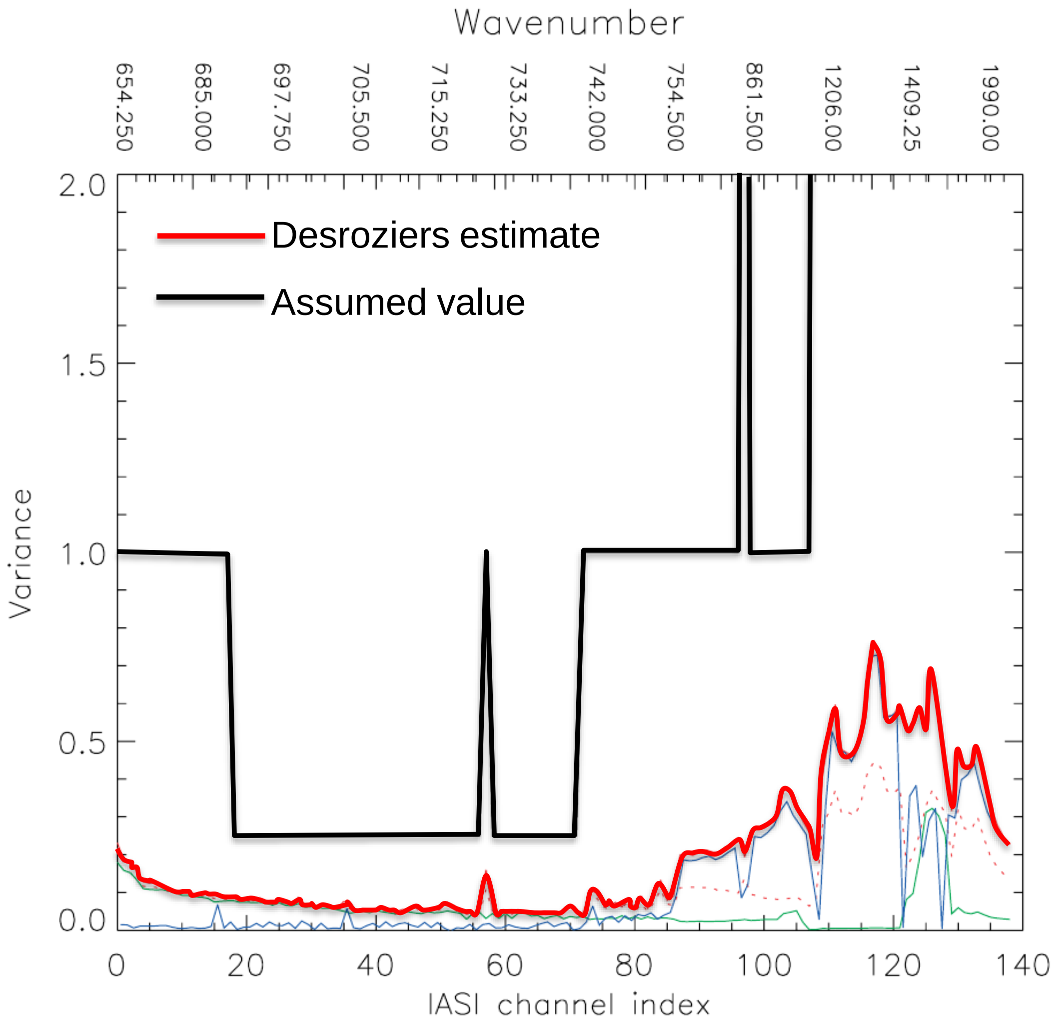
- ***Consistency diagnostic*** for the specification of **R**. Increasingly used to estimate **R**.

Diagnosed error correlations for IASI (Stewart et al 2009, 2014)

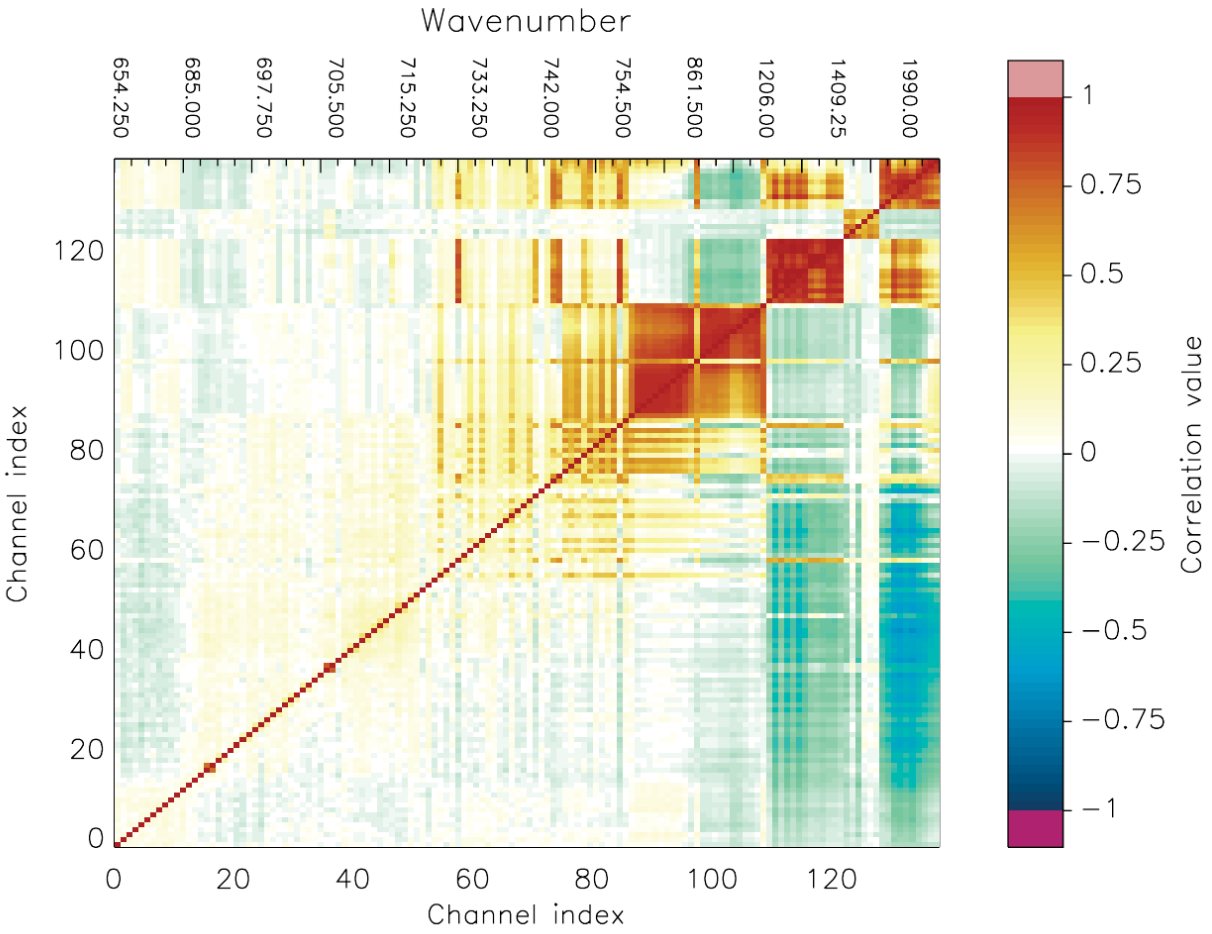


→ *Sarah Dance's talk on error diagnostics*

Estimating inter-channel error correlations for hyper-spectral IR observations and the Desroziers diagnostic



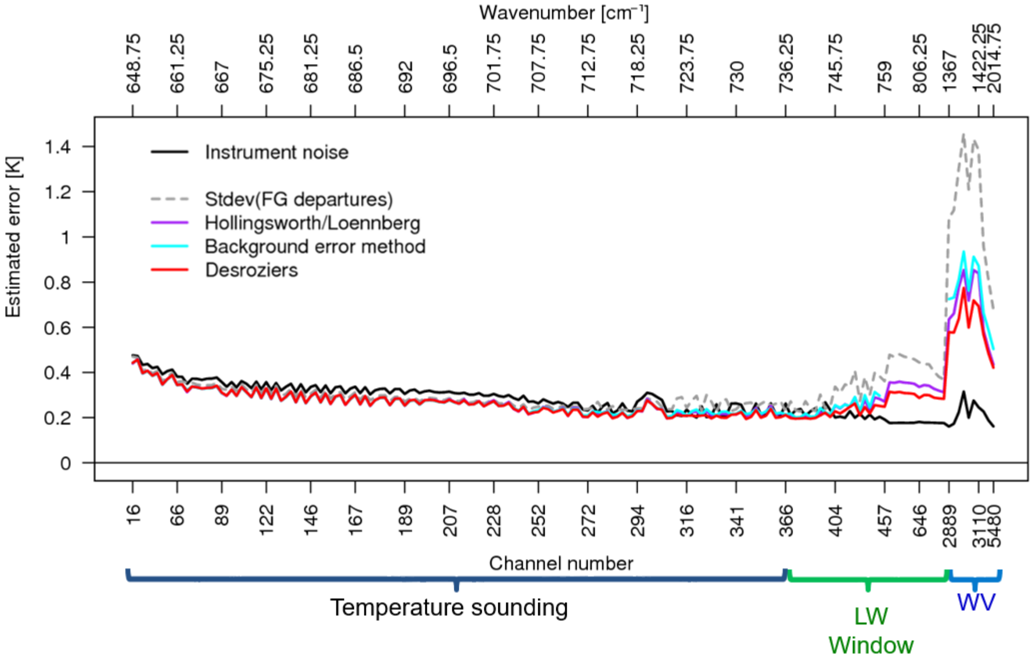
Diagnosed error correlations for IASI
(Stewart et al 2009, 2014)

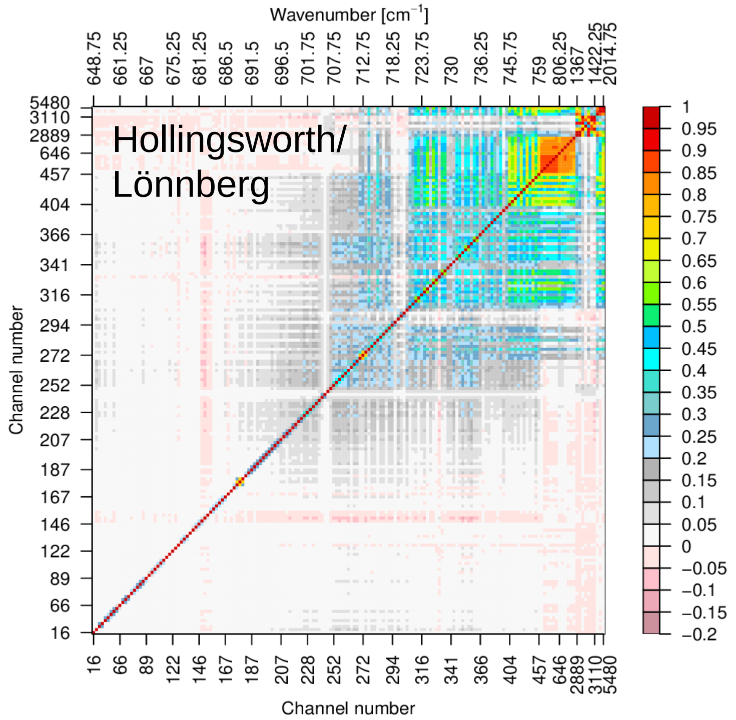


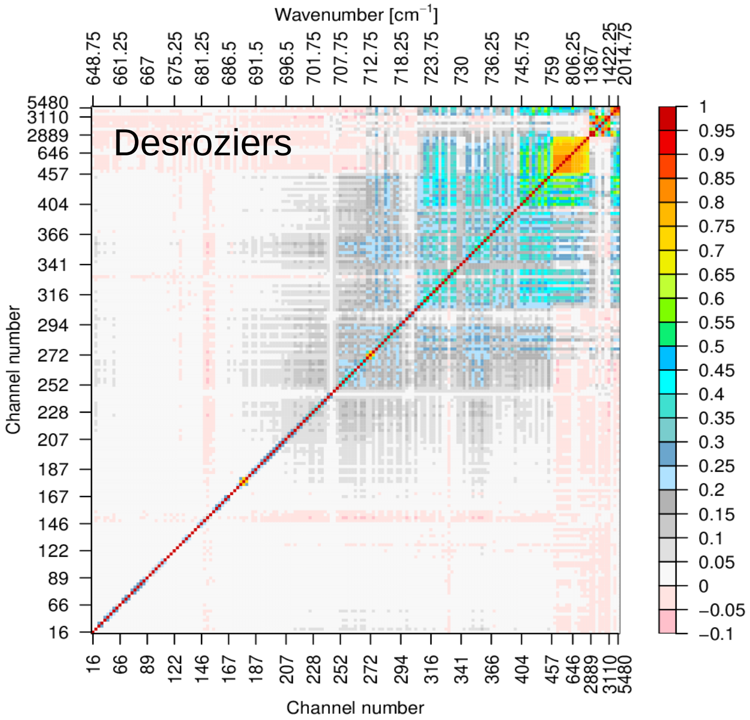
→ *Sarah Dance's talk on error diagnostics*

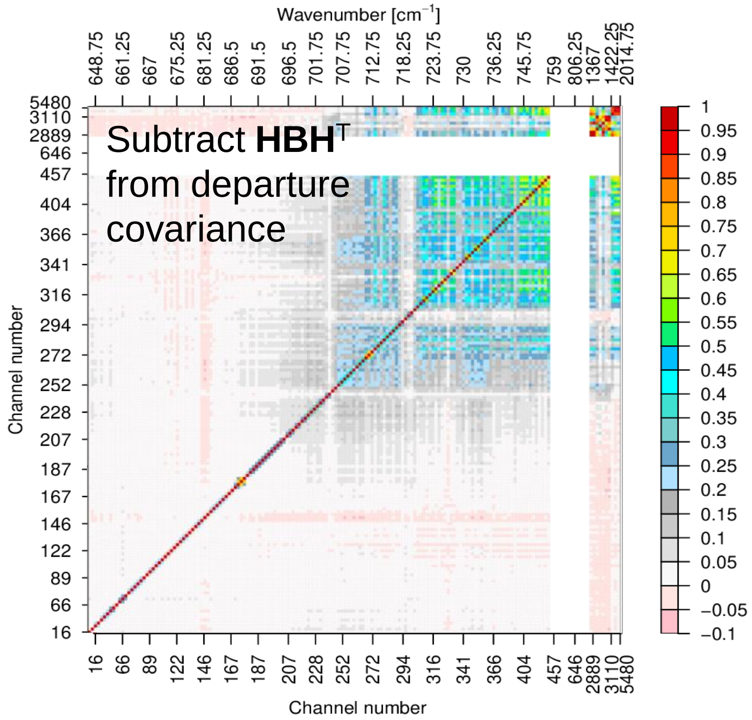
Estimating inter-channel error correlations for hyperspectral IR: Different diagnostics, similar results

(Bormann et al 2010)



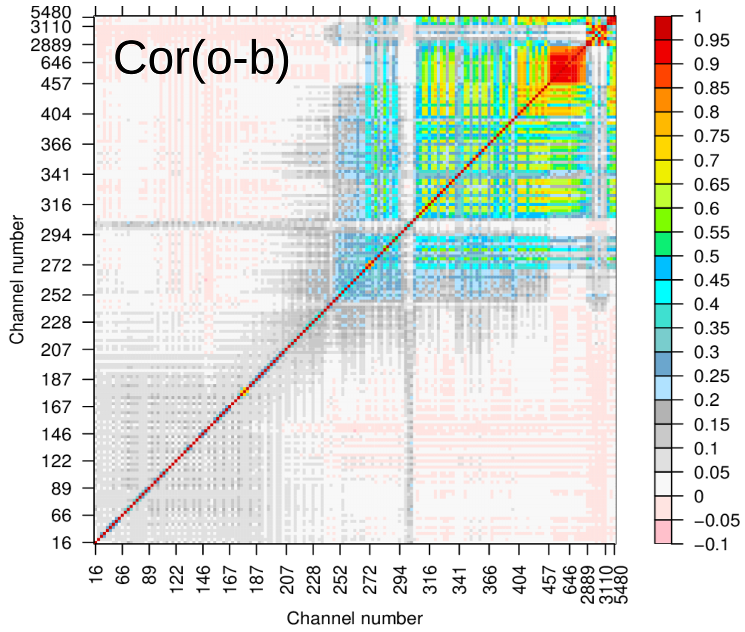


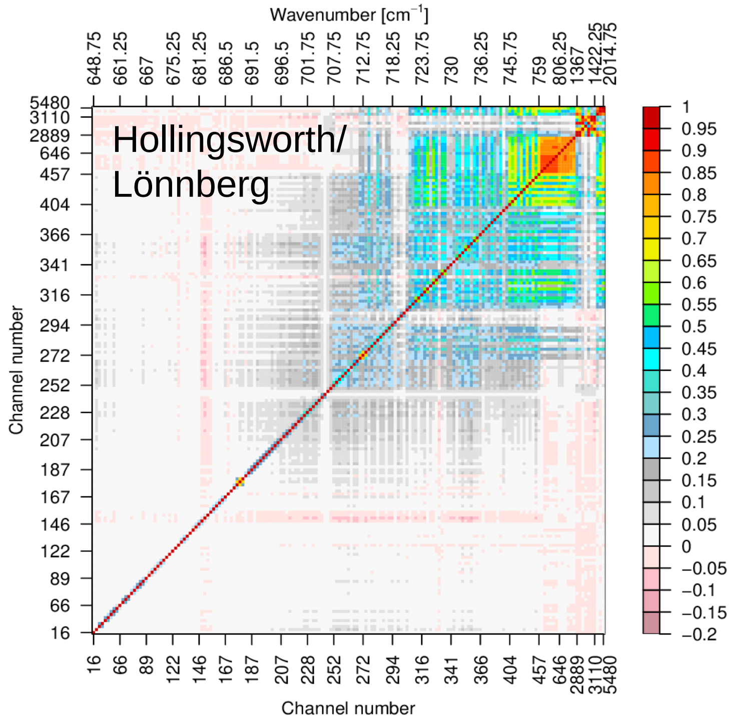


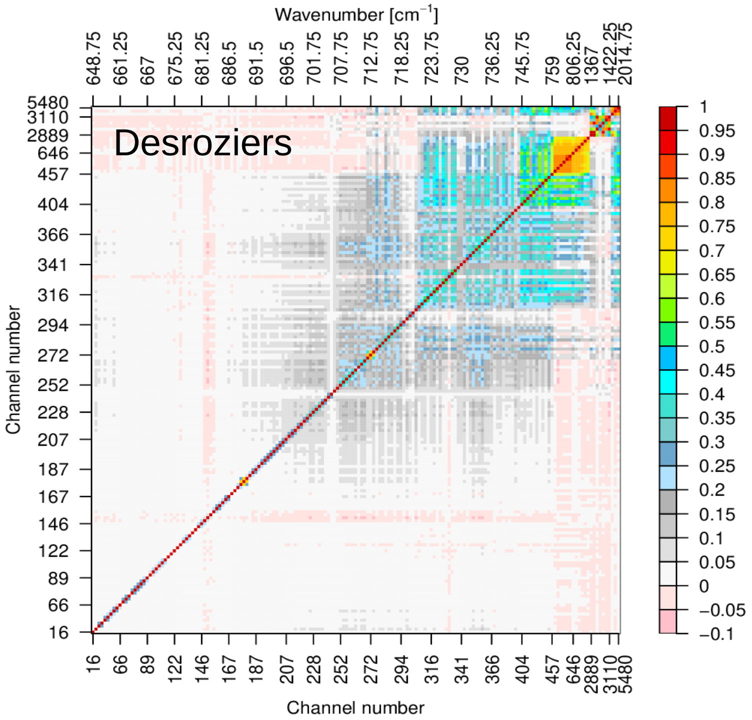


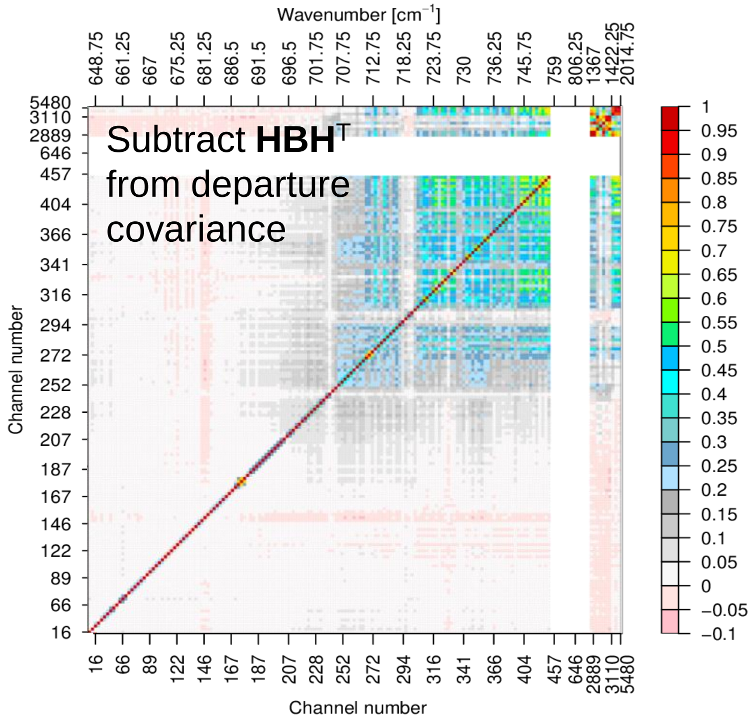
Estimating inter-channel error correlations for hyperspectral IR: Different diagnostics, similar results

(Bormann et al 2010)





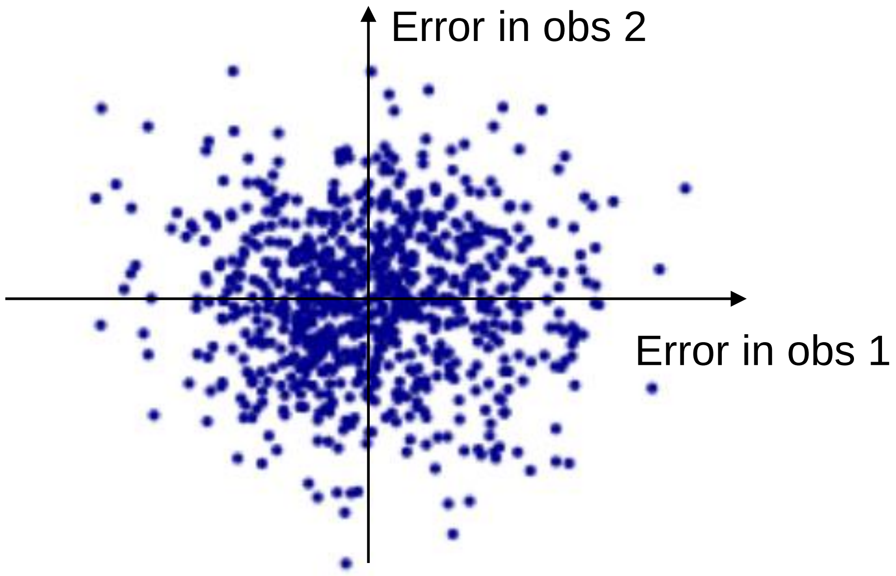




What is the effect of error correlations?

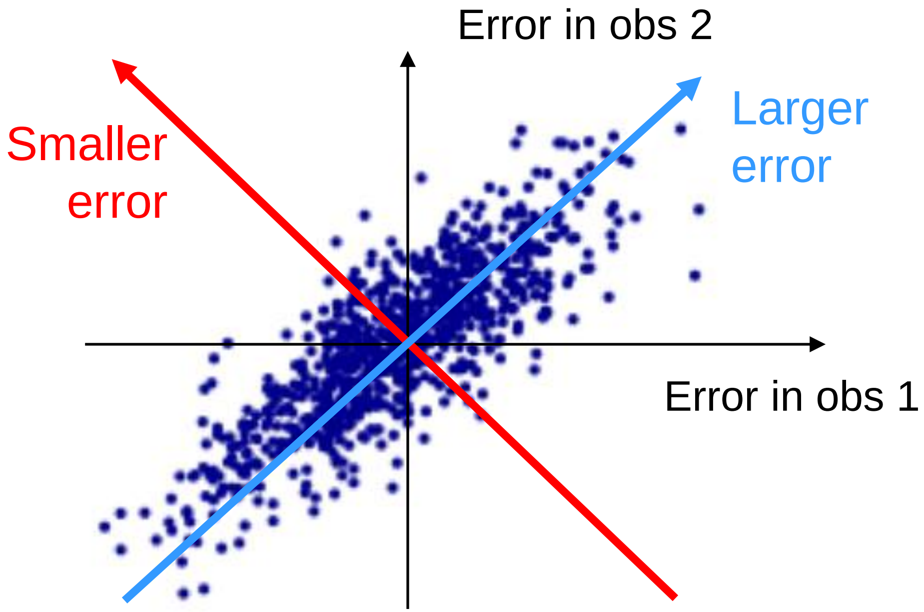
Uncorrelated error

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



Correlated error

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

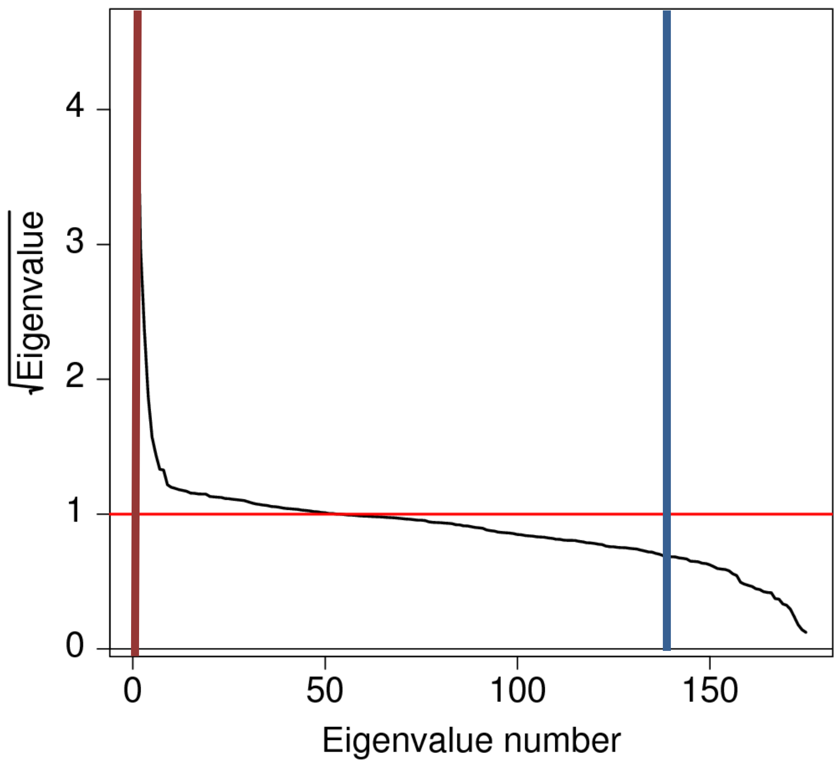


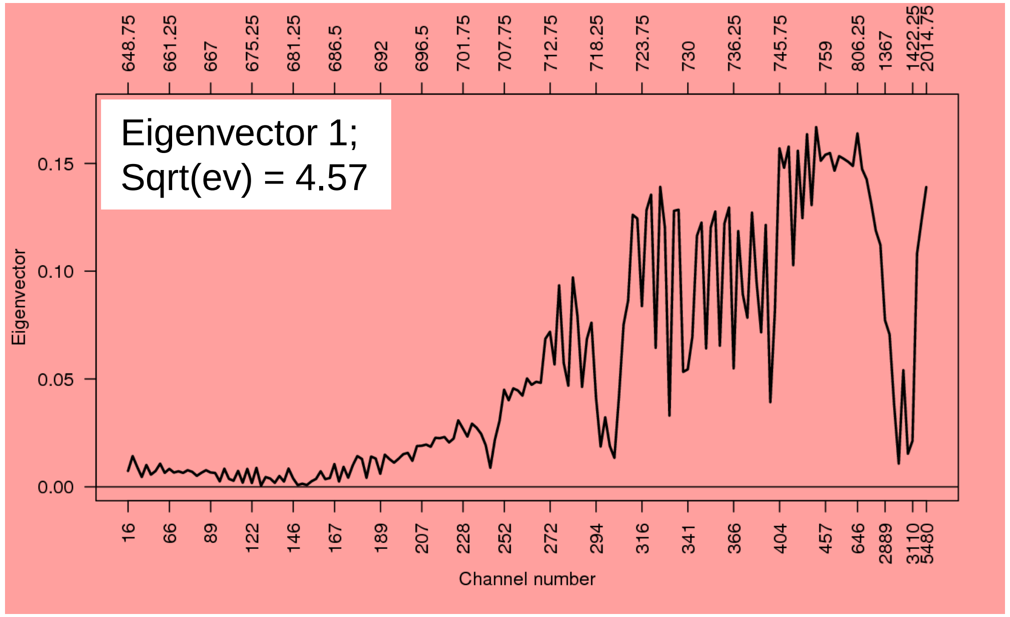
Compared to diagonal errors, ***positive error correlations imply***…

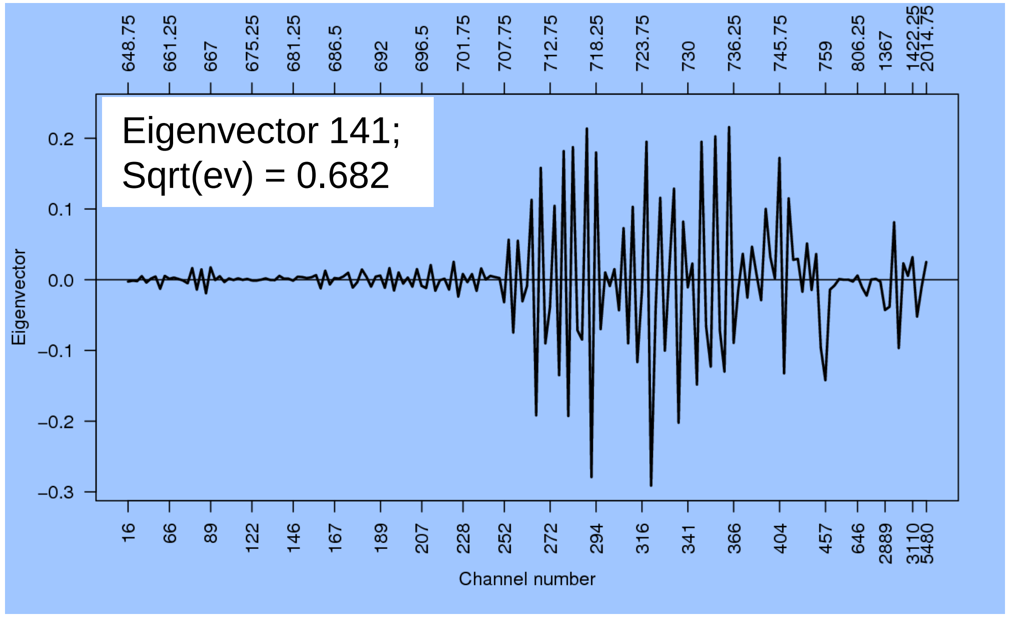
- … ***larger errors*** for features along the blue direction (mean-like features).
- … ***smaller errors*** for features along the red direction (differencee-type features).

Example: error correlations for IASI

Eigenvalues of the
error correlation matrix:







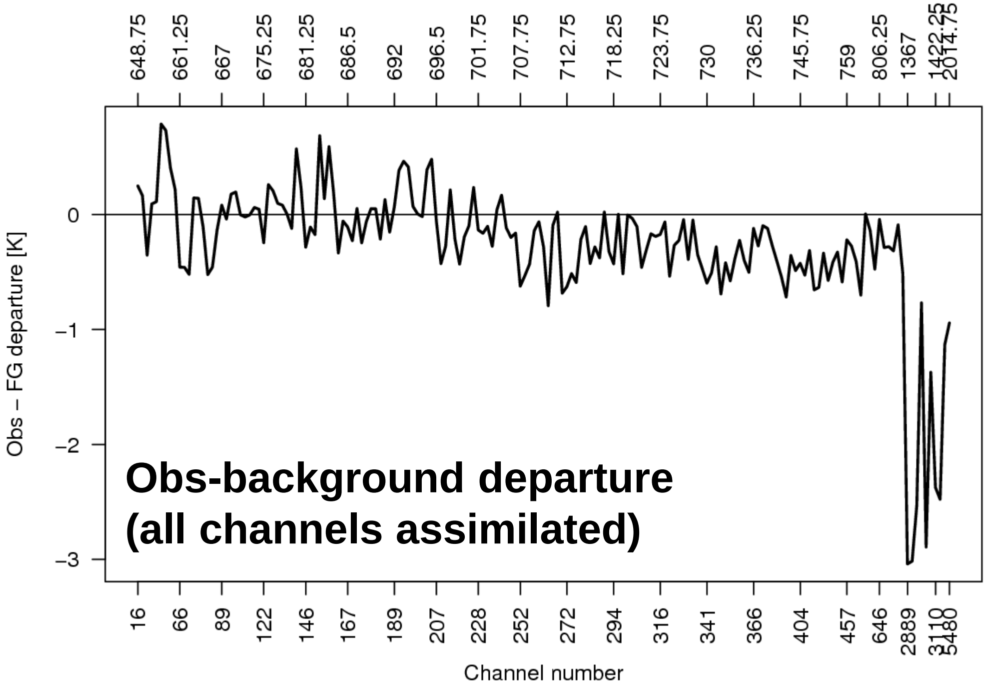
Example: Assimilation of a single IASI spectrum (I)

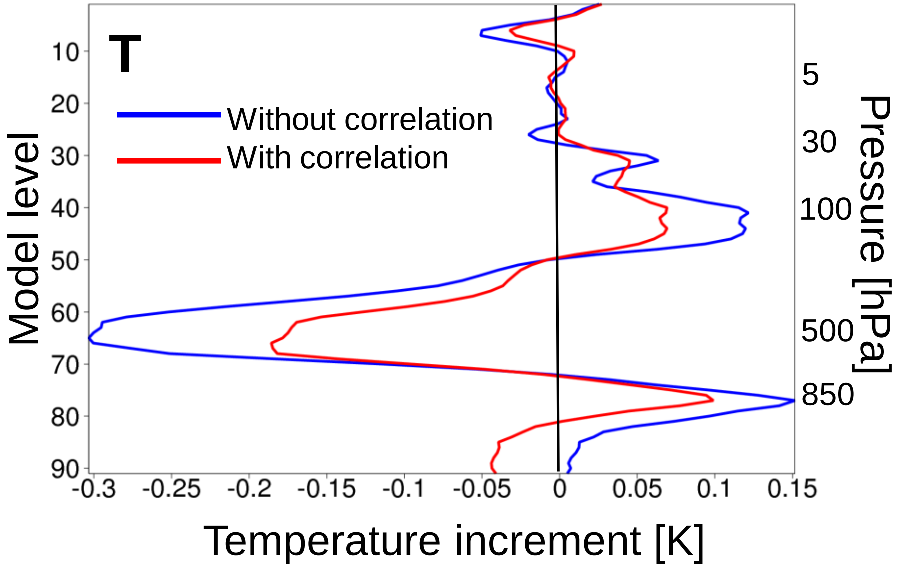
“Similar” departures → increments reduced with error correlations taken into account

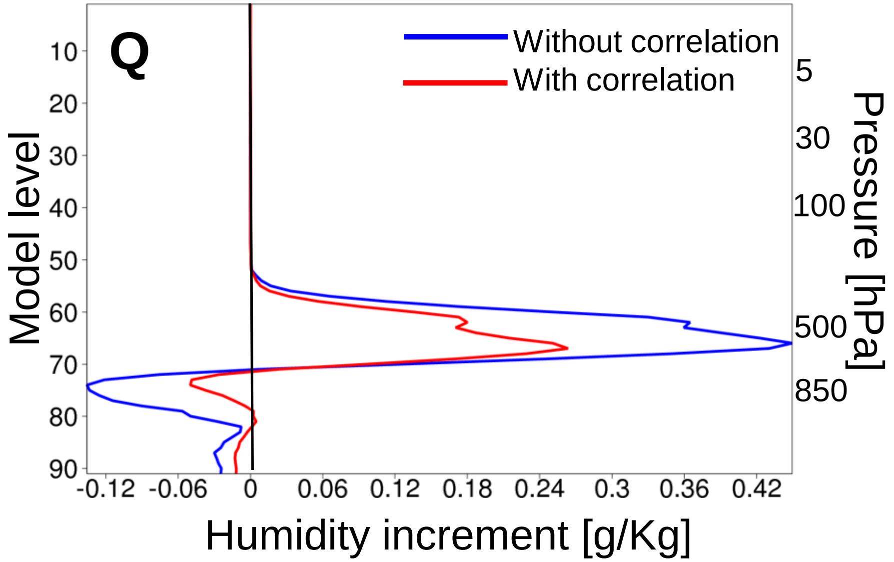
Assimilate a single IASI spectrum,

- assuming **no error correlations**,
- assuming diagnosed error correlations

(σ_o unchanged in both cases).





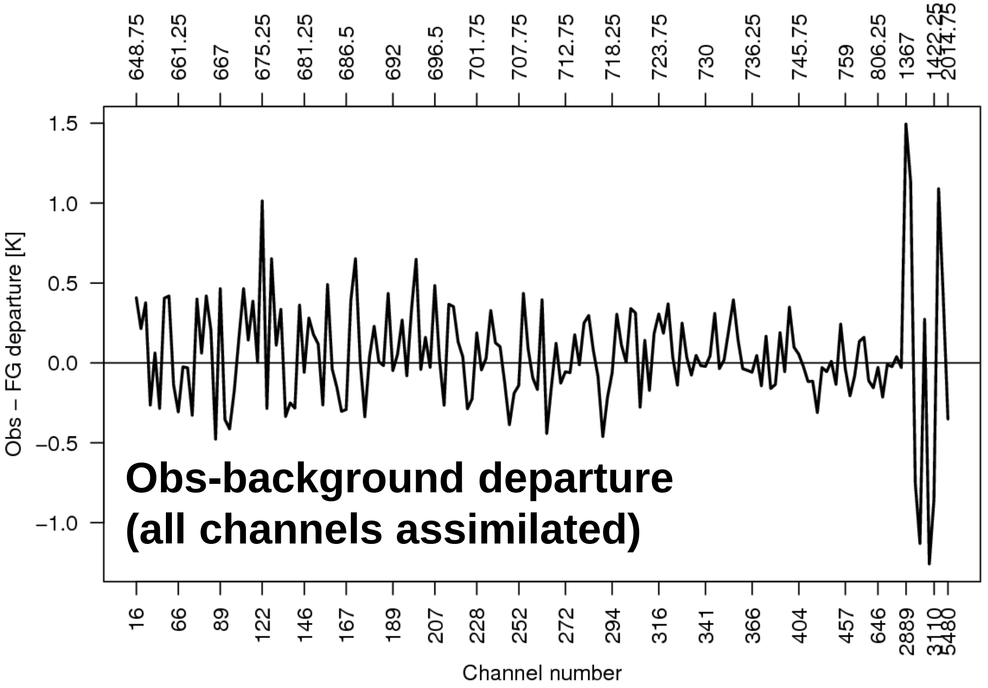


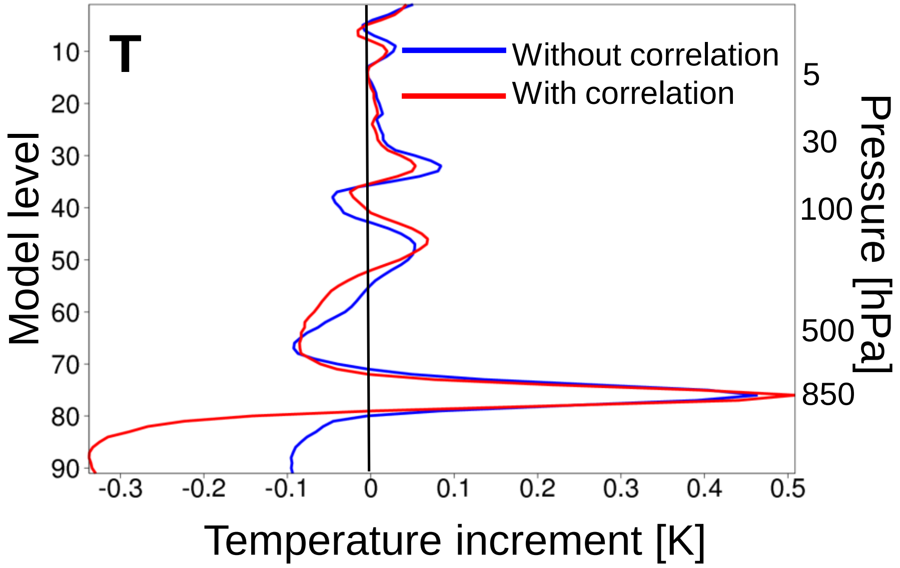
Example: Assimilation of a single IASI spectrum (II)

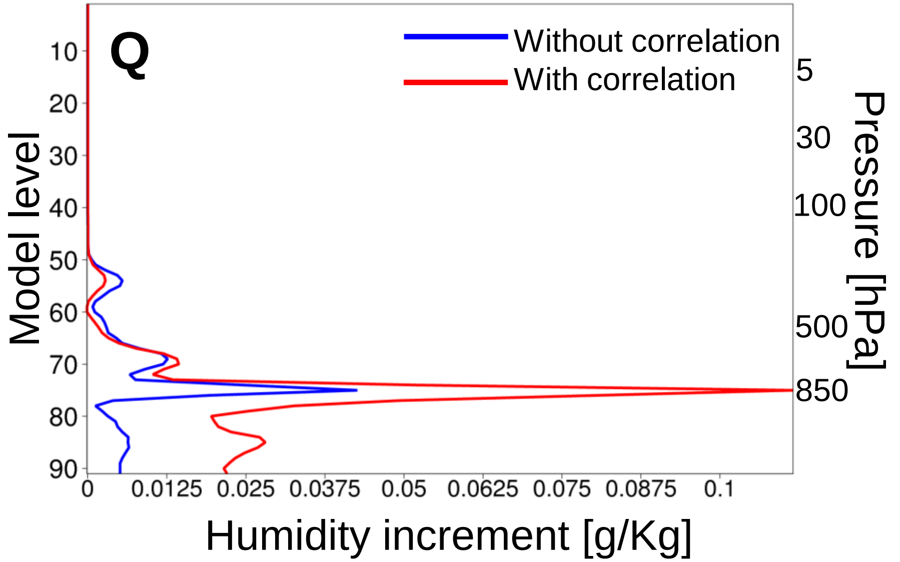
"Different" departures → increments ***increased*** with error correlations taken into account

Assimilate a single IASI spectrum,
- assuming **no error correlations**,
- assuming diagnosed error correlations

(σ_o unchanged in both cases).





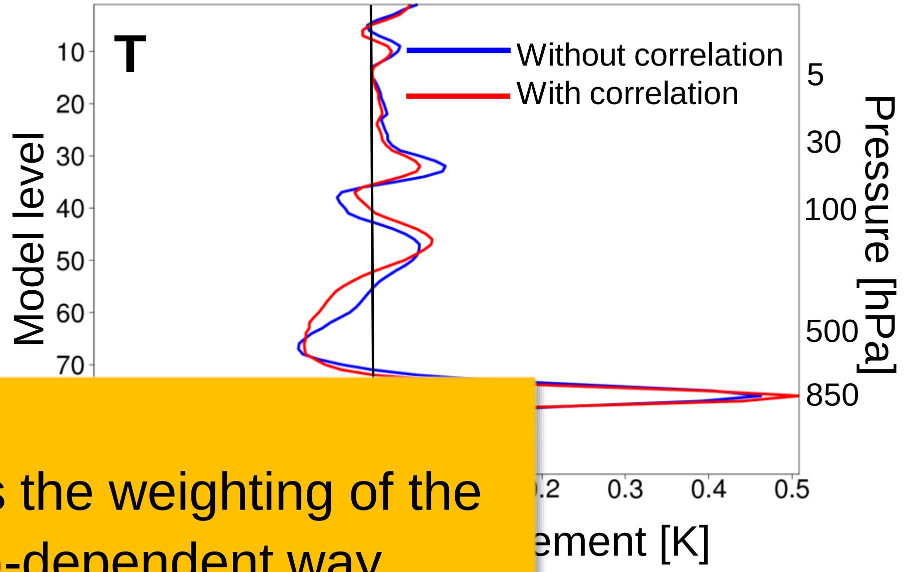


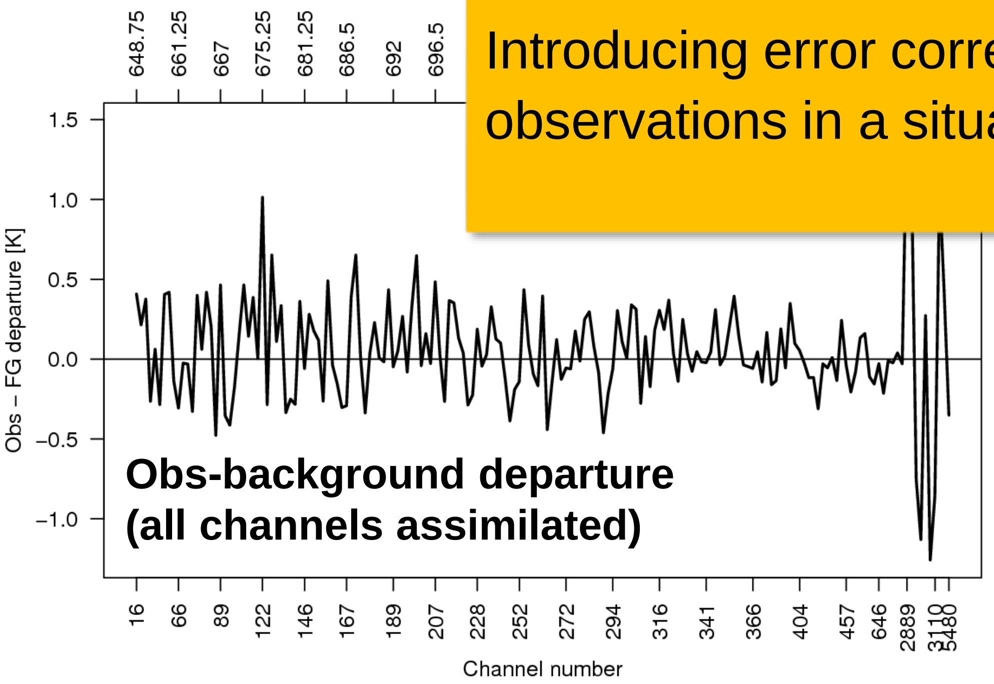
Example: Assimilation of a single IASI spectrum (II)

“Different” departures → increments ***increased*** with error correlations taken into account

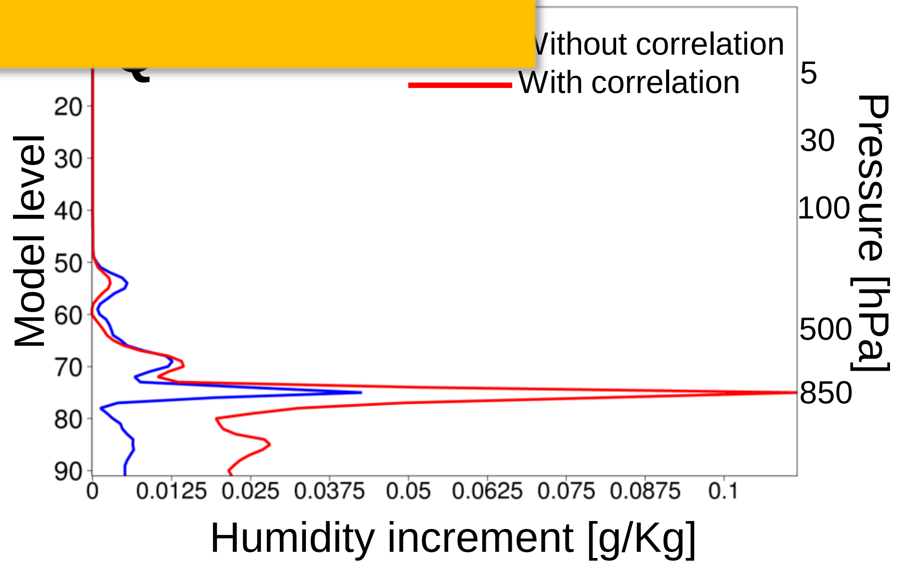
Assimilate a single IASI spectrum,
- assuming **no error correlations**,
- assuming diagnosed error correlations

(σ_o unchanged in both cases).

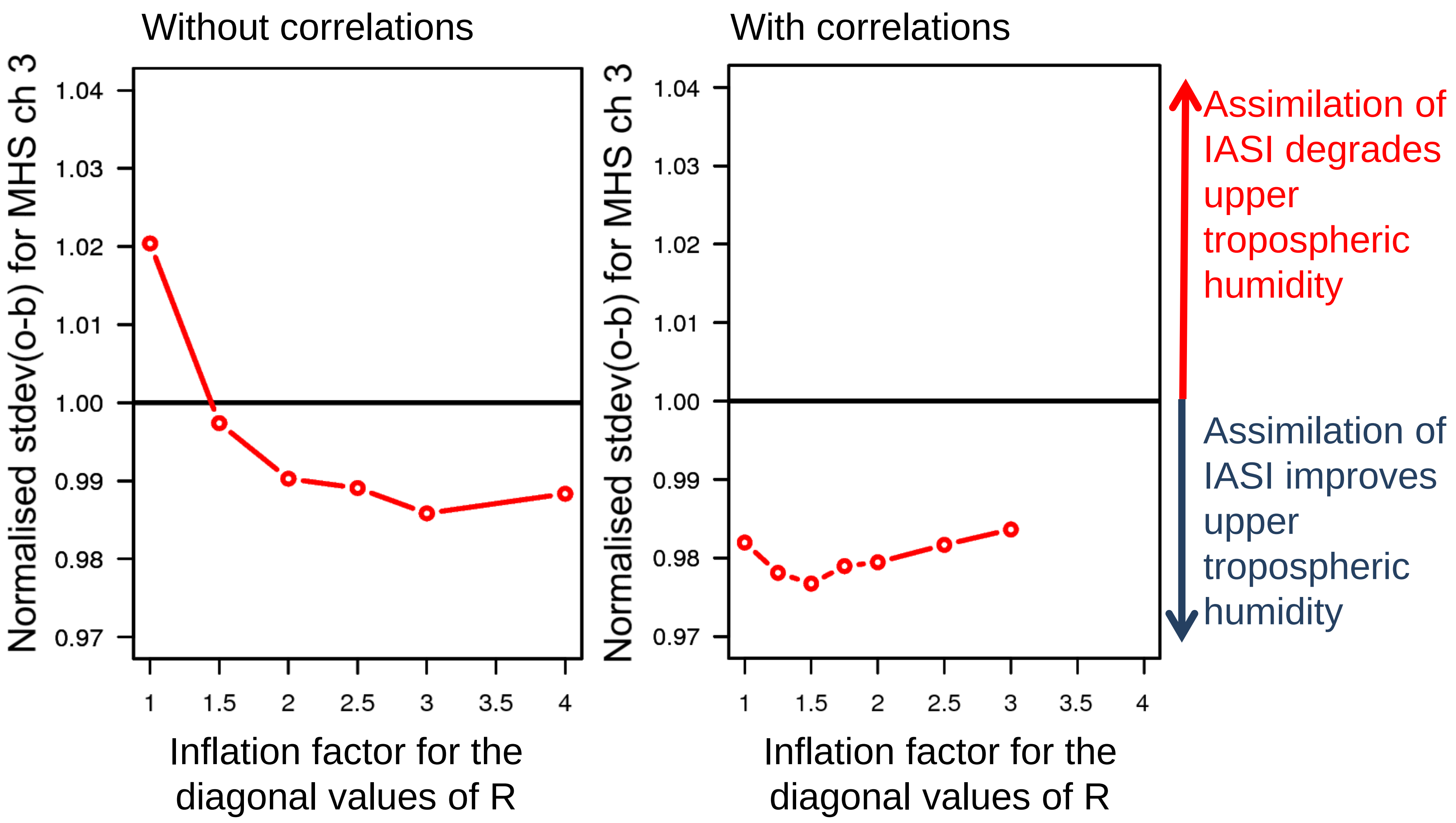




Introducing error correlations changes the weighting of the observations in a situation(/departure)-dependent way.



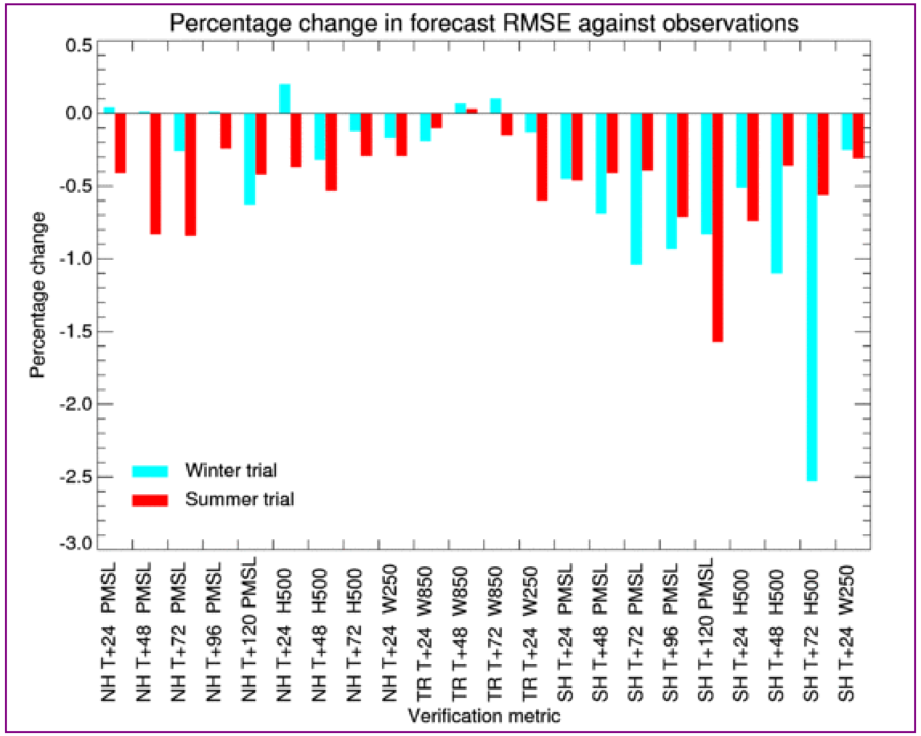
Effect of accounting for inter-channel error correlations in the assimilation of IASI



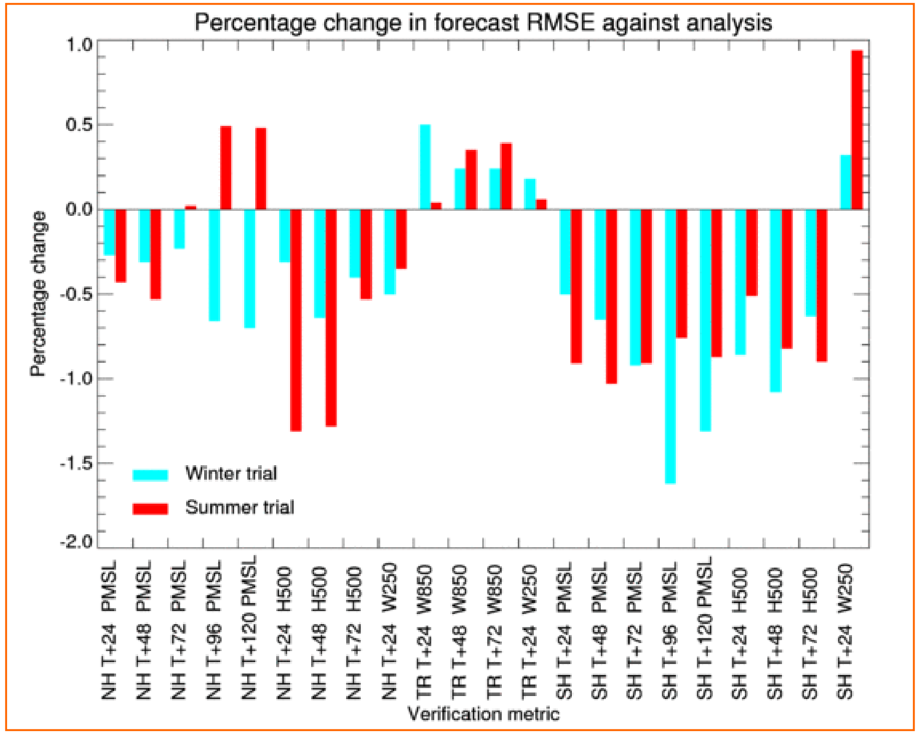
Accounting for inter-channel error correlations in the assimilation

- Now widely used at operational centres, for hyperspectral IR, geostationary imager radiances, ATMS, etc.
- E.g., Weston et al (2014), Bormann et al (2016), Campbell et al (2017), Weston and Bormann (2018), Burrows (2018), Bathmann and Collard (2020), …
- → *Fiona Smith's talk on the status of R for hyperspectral IR*

Verification v Observations



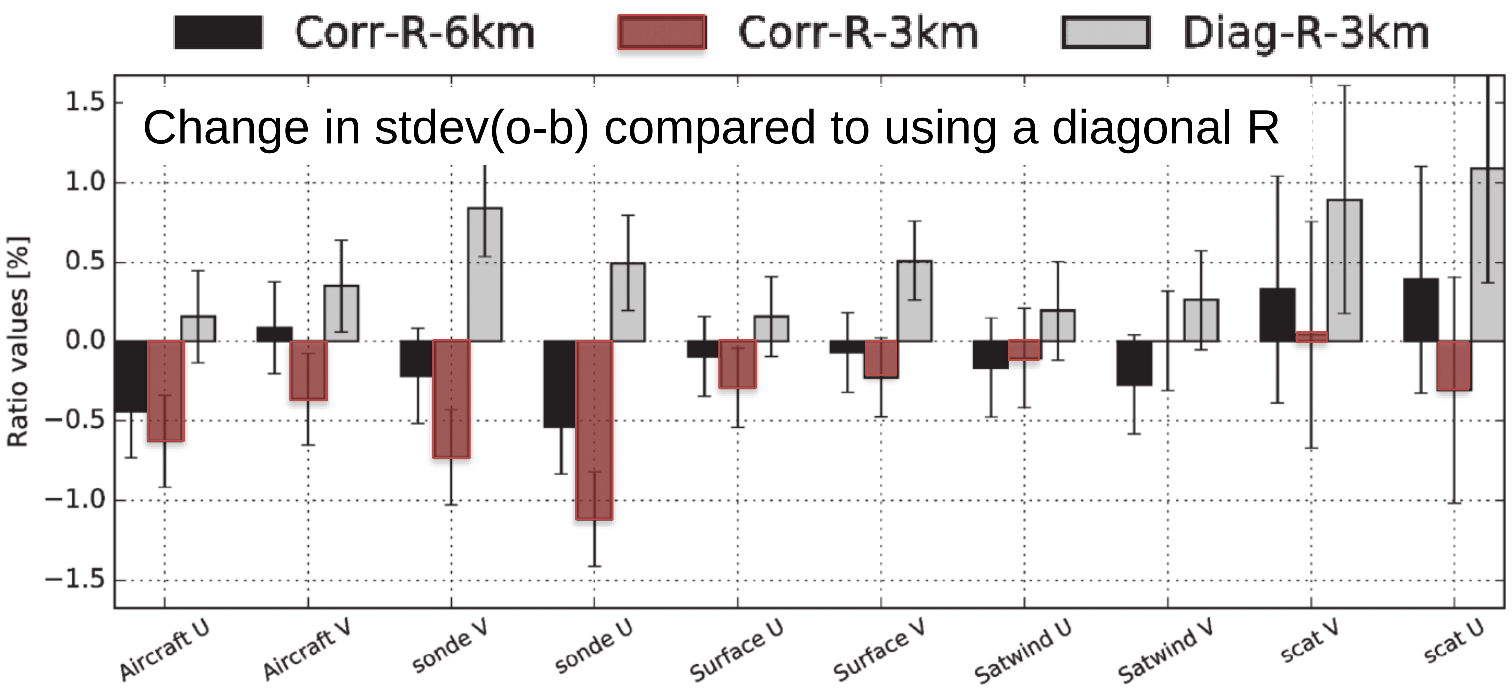
Verification v Analyses



Weston et al (2014)

Accounting for spatial error correlations

- Less work has been done on accounting for spatial error correlations in NWP, partly as it is ***technically more difficult in variational frameworks***.
- But recent activity in several areas (→ *talks by Koji Terasaki, Oliver Guillet, Joël Bédard*)
  - First ***operational application*** in Met Office UKV system for radial winds from Doppler radar (Simonin et al 2019):



Accounting for spatial error correlations allows beneficial assimilation of radar winds with less thinning.

  - Particular interest for regional models, to ***improve small-scale representation***.

Some further points on accounting for observation error correlations

- Accounting for error correlations puts ***more weight on differences between observations.***
  - Are these differences reliable? How reliable are ***inter-channel calibration/bias correction***?
  - Are the ***estimates of error correlations reliable***?
- Accounting for observation error correlations can affect the ***conditioning*** of the assimilation and lead to slower convergence.
- The importance of accounting for error correlations may additionally depend on the structure of the ***background error***.

Outline

1. Introduction
2. Situation-dependent observation errors
3. Correlated observation errors
4. **Error inventories and closure studies**
5. Summary

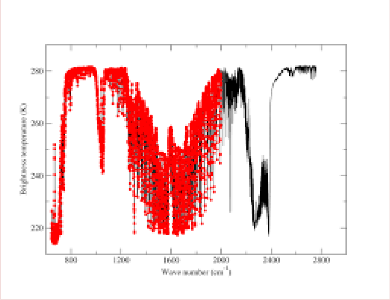
Contributions to observation error

Measurement error

E.g., instrument noise for satellite radiances

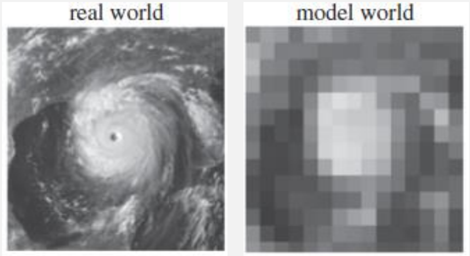
Forward model (observation operator) error

E.g., radiative transfer error



Representativeness error

E.g., point measurement vs model representation



Quality control/pre-processing error

E.g., error due to the cloud detection scheme missing some clouds in clear-sky radiance assimilation

?

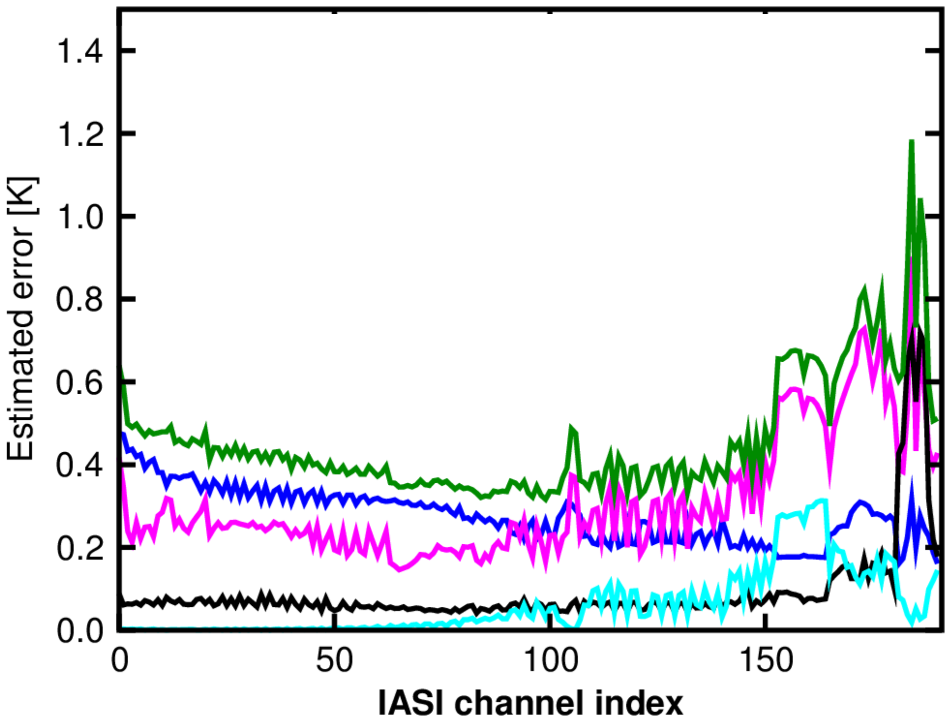
→ talks in session 2

Error inventory

(e.g., Chun et al 2015)

• Idea: Estimate the observation error from ***estimates of all uncertainty*** contributions.

• Example: error inventory for IASI



Instrument noise (information from data providers)

Radiative transfer error (estimated through reference obs)

Spatial representativeness error (e.g, through high vs low-resolution simulations)

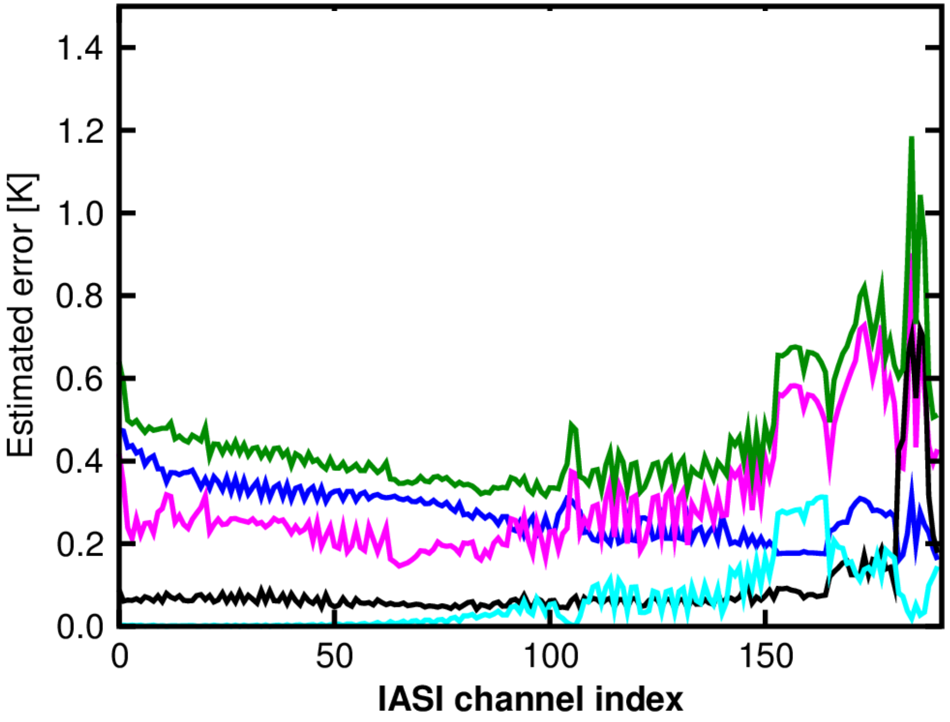
Cloud detection error (e.g., using simulations of cloudy radiances)

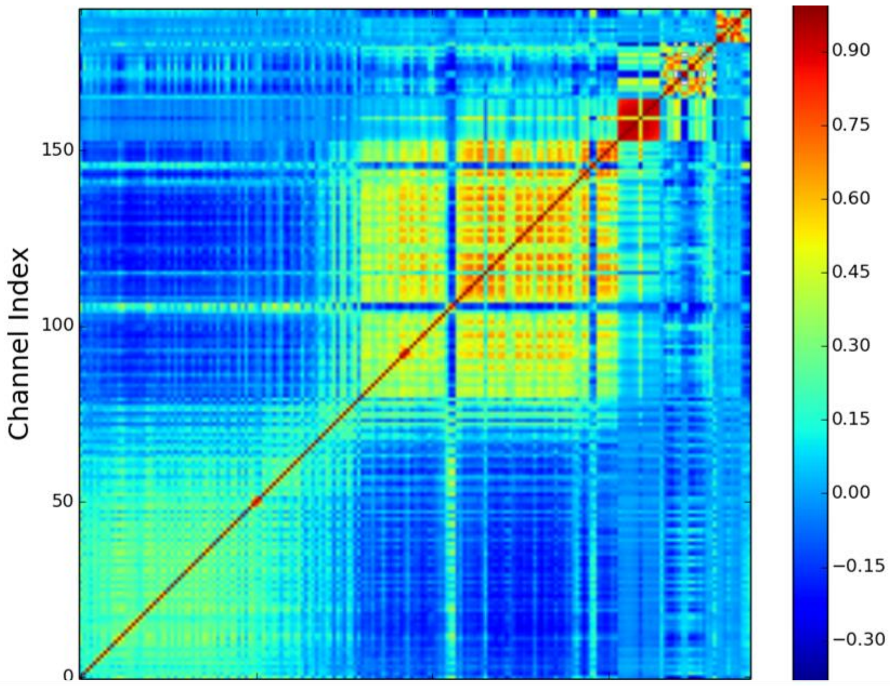
Combined observation error

Error inventory

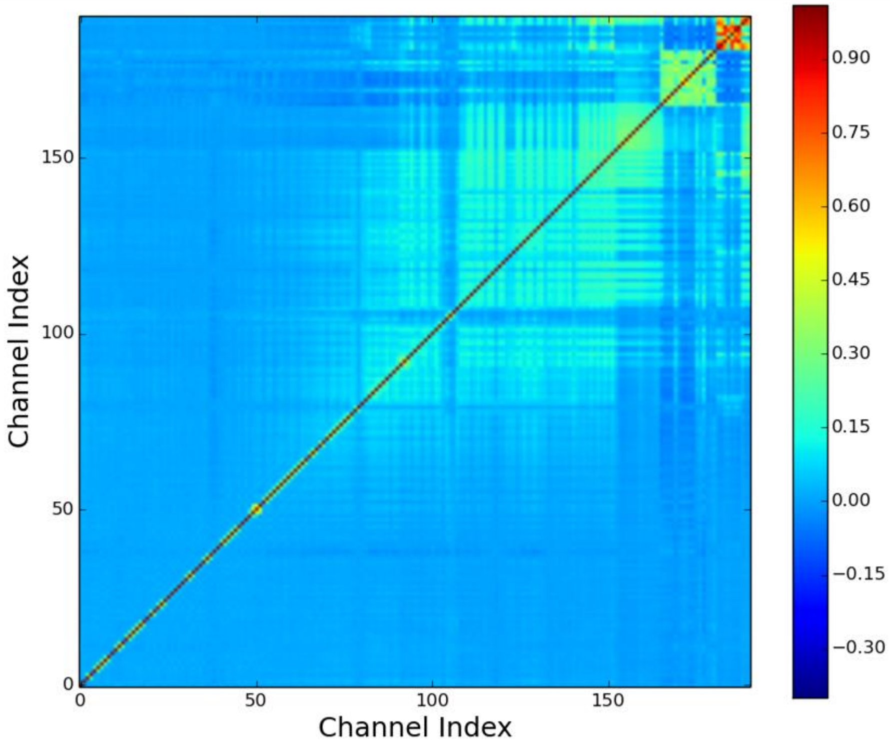
(e.g., Chun et al 2015)

- Idea: Estimate the observation error from ***estimates of <u>all</u> uncertainty*** contributions.
- Example: error inventory for IASI





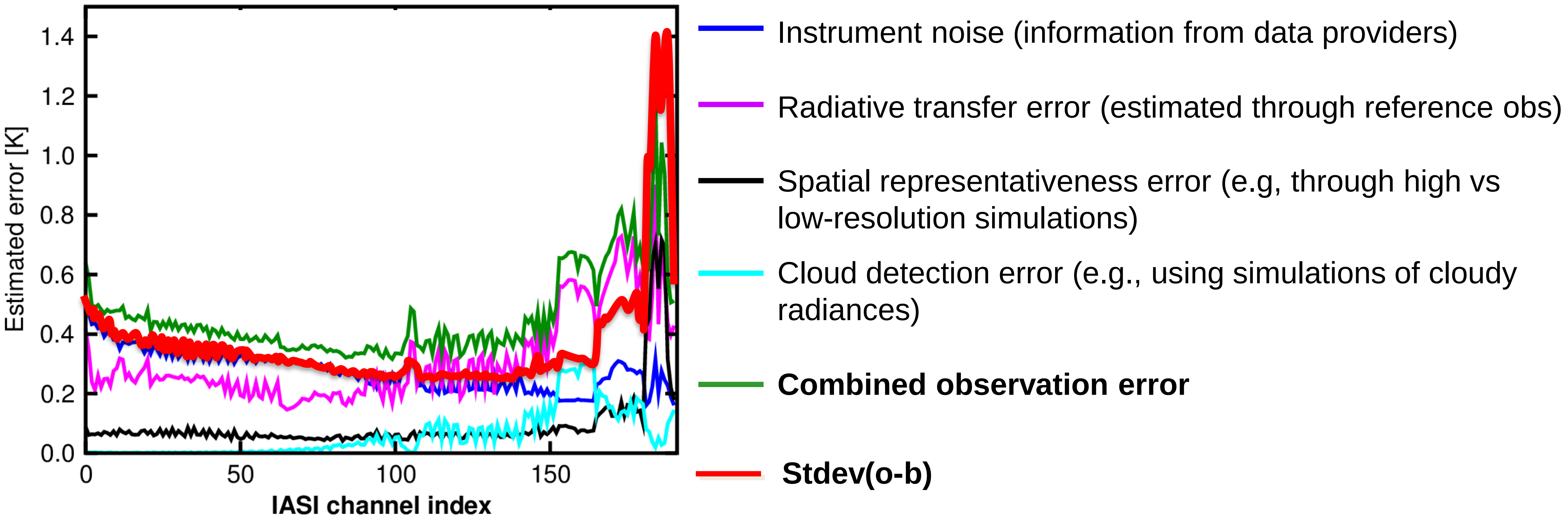
Inventory approach



Diagnosed using Desroziers

Error inventory … and closure studies

- How do the separate error estimates compare to the total (observation + background) error estimate from observation departures?
- Here: **Combined observation error estimate alone is (mostly) larger than stdev(o-b).**
  - Overestimation of error contributions?
  - Correlations between background and observation errors (e.g., cloud detection error)?



Outline

1. Introduction
2. Situation-dependent observation errors
3. Correlated observation errors
4. Error inventories and closure studies
5. **Summary**

Summary

- A lot of progress in specifying observation errors in recent years; ***more aspects of observation error are being taken into account***.
  - ***Situation-dependence*** of observation errors increasingly taken into account, based on ***physical considerations paired with departure statistics***.
  - ***Inter-channel error correlations*** are now widely accounted for, using results of ***departure-based diagnostics*** with some adjustments to specify R.
  - Accounting for ***horizontal error correlations*** is emerging.
  - Continue to see ***significant benefit*** for forecast skill from better specifications of observation errors.
- Most sophistications of observation error modelling are based on departure statistics in one way or another.
  - Stdev(o-b), Hollingsworth/Lönnberg, Desroziers, Cov(o-b) – $\mathbf{HBH}^T$; collocated observations/triple collocations
  - All rely on a ***range of assumptions***, which may or may not be true.
  - Sometimes ***adjustments*** are necessary (inflation/reconditioning), sometimes they aren't.
  - ***Error inventories*** can instead shed light on the dominant sources of error, and they can bring ***further independent information*** to error modelling.

Some thoughts for the working groups

- A lot of progress in specifying observation errors in recent years with increased sophistication.
  - But what level of (further) sophistication is useful/desirable?
    - Limitations in the available estimates for specifying R
    - Maintainability of error modelling - responding to changes in the size of error contributions
- What tools do we have to estimate observation errors and how well do they cover our needs?
  - How can we make more use of uncertainty characterisation beyond departure-based diagnostics (e.g., instrument characterisation, metrological approaches, etc)?
- When do observation error correlations matter?
  - Compare, for instance, success of taking inter-channel error correlations into account for hyper-spectral IR vs the diagonal observation error modelling in successful MW all-sky assimilation (where representation error is huge and correlated).
- What aspects of observation error modelling may become more important in the future?
  - E.g., due to higher-resolution analyses; observations with higher temporal resolution; Earth system approaches