ACCORD

A Consortium for COnvection-scale modelling Research and Development

Estimation and modelling of observation errors

25th EWGLAM 30th SRNWP meeting, 25-28 September 2023, Reykjavik, Iceland

Jana Sánchez-Arriola and Magnus Lindskog

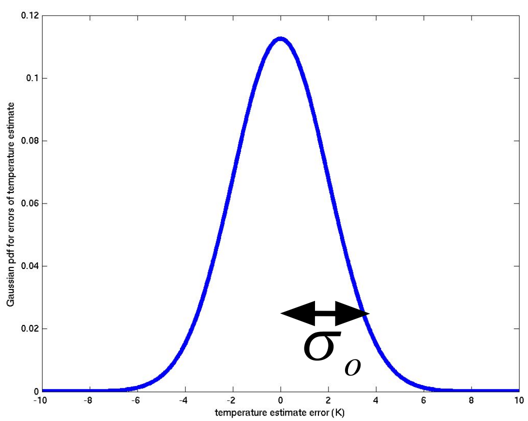
Outline

- Observation errors in data assimilation
- Estimation of observation errors
- Modelling/representation of errors
- Discussion points

Observation errors in data assimilation

- In data assimilation we usually assume Gaussian observation error distributions and characterized by their standard deviations and correlations.
- Observation systematic errors (biases) needs to be handled.
- Observations affected by Gross errors can degrade the analysis and subsequent forecast nd needs to be removed prior to data assimilation. But we do not want to reject accurate observations.
- Scale- and time differences between observations and model usually included in the observation error part (representativity and persistence error.

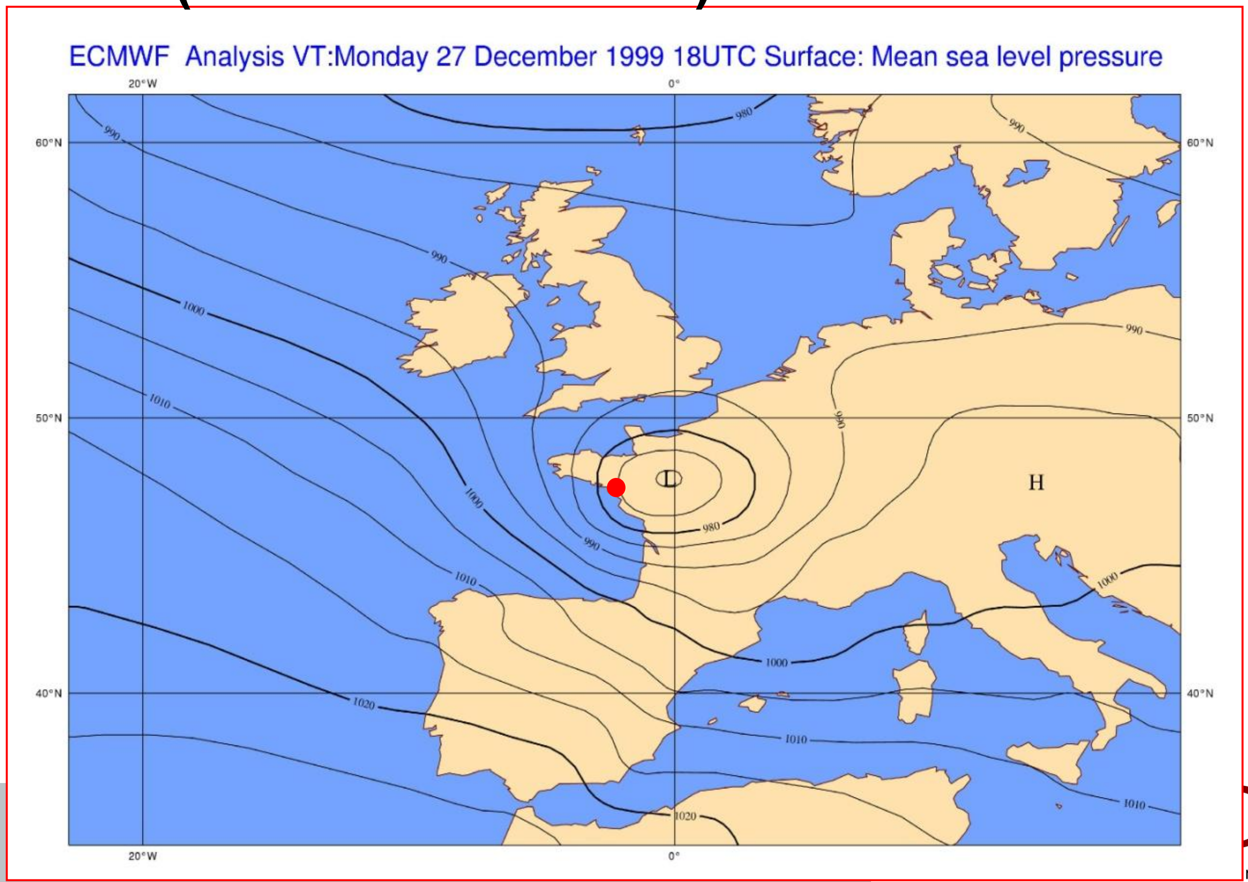
Gaussian probability distribution function



Observation errors in data assimilation

Effect of QC: Red-marked observation rejected
27.12.1999 – French storm 18UTC

- **ECMWF Era interim analysis produced a low with min 970 hPa**
- **Lowest pressure observation (SYNOP: red circle)**
  - –963.5 hPa (supported by neighbouring stations)
  - –At this station the analysis shows 977 hPa
  - –Obs – An = 16.5 hPa!

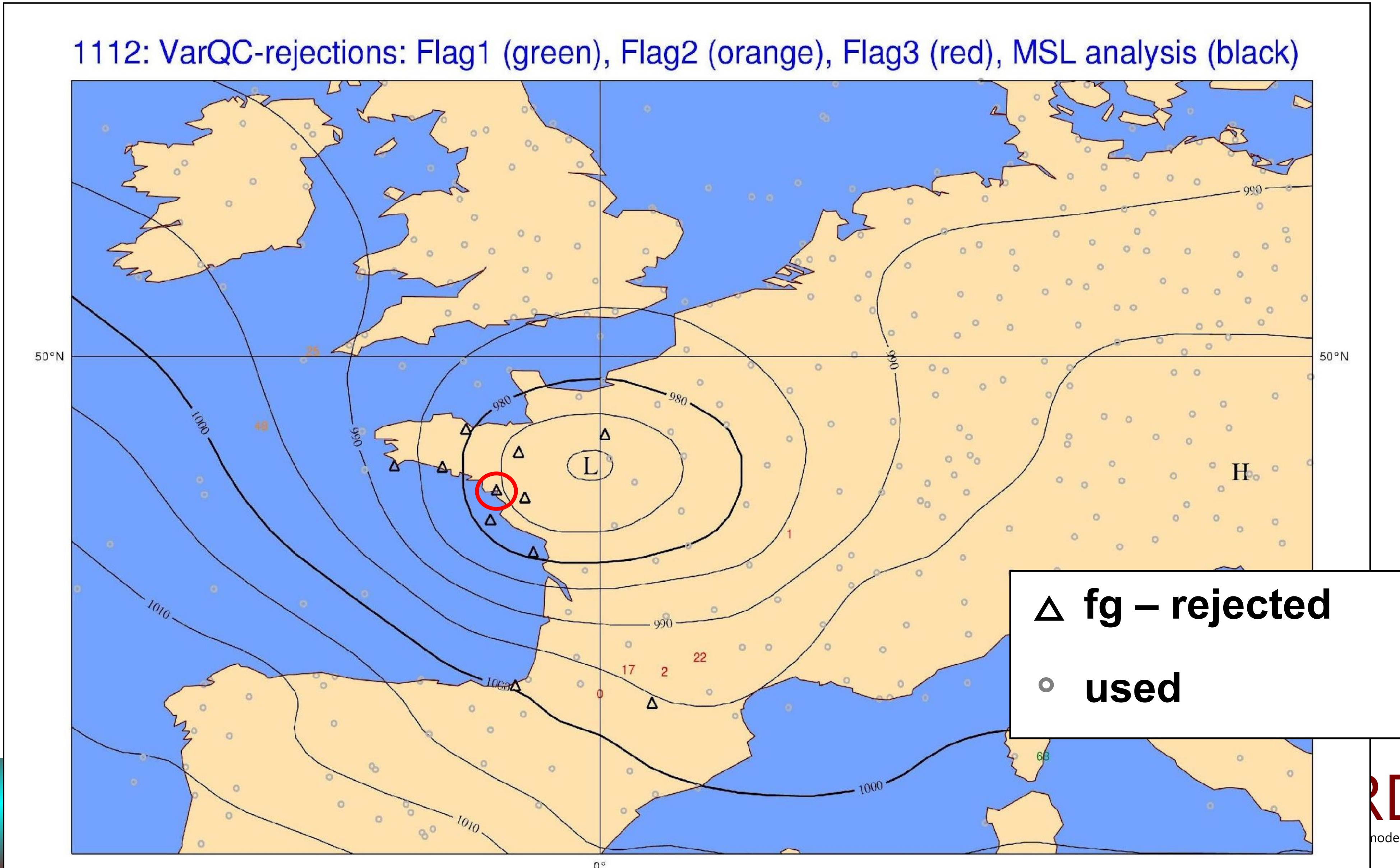


- High density of good quality surface data

(from Lars Isaksen och Christina Tavolato, ECMWF)

Observation errors in data assimilation

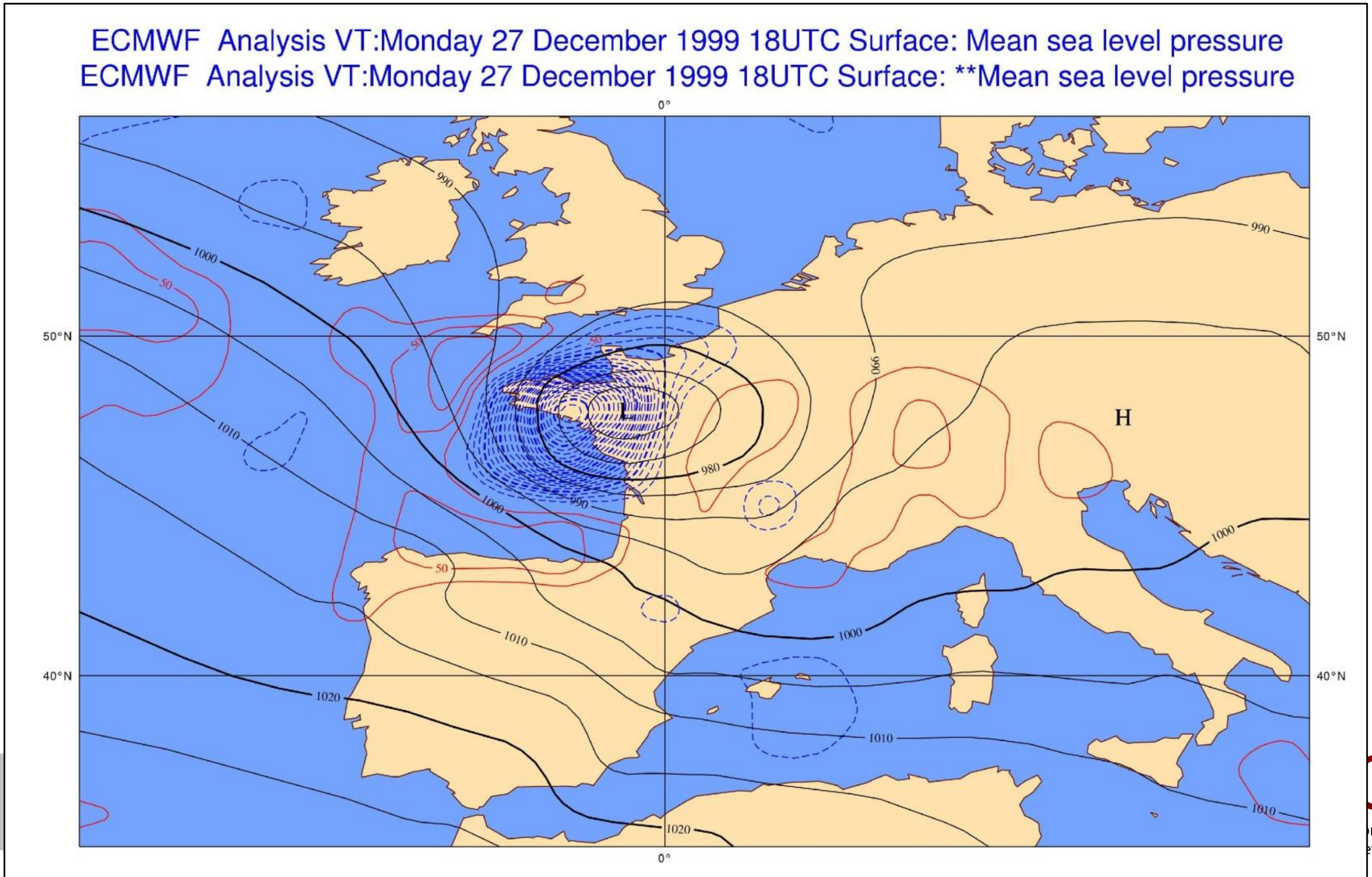
Data rejection and VarQC weights – ECMWF Era interim



Observation errors in data assimilation

Effect of revised quality control

- New min 968 hPa
- Low correctly shifted towards west and intensified in better agreement with surface pressure observations



r COnvection-scale modelling
evelopment

Estimation of observation errors

Standard deviations and correlations

- Hollingswoth-Lönnberg (Tellus, 1986)
- Desroziers method (QJR, 2005)
- Bucanek/Bormann and Bauer (QJR, 2010)
- …..

Biases

- Harris-Kelly (QJR, 2001)
- Variational bias correction (Dee, 2005)
- Bias correction simply based on averaged deviation from model (Poli et al, JGR 2007), or on a 10-day running mean (McPherson et al, MWR, 2008). Applied for GNSS, assuming model unbiased.
- Agusti-Panareda et al, QJRMS, 2009: bias-correction assuming the night-time RS-92 is bias-free, using the model as an intermediate.
- ….

Identification of Gross errors

- Järvinen-Andersson (QJR, 1999)
- Variational quality control
- Background check
- Buddy check (Benjamin et al, MWR, 2004)
- Adaptive QC (Dee et al, QJR, 2001)
- ….

Modelling/representation of errors

Standard deviations and correlations

- Thinning of observations
- Inflation of observation error standard deviations
- Inter-channel observation error correlation representation
- Spatial observation error correlation representation
- …..

Biases

- Static bias correction methods
- Adaptive
- ….

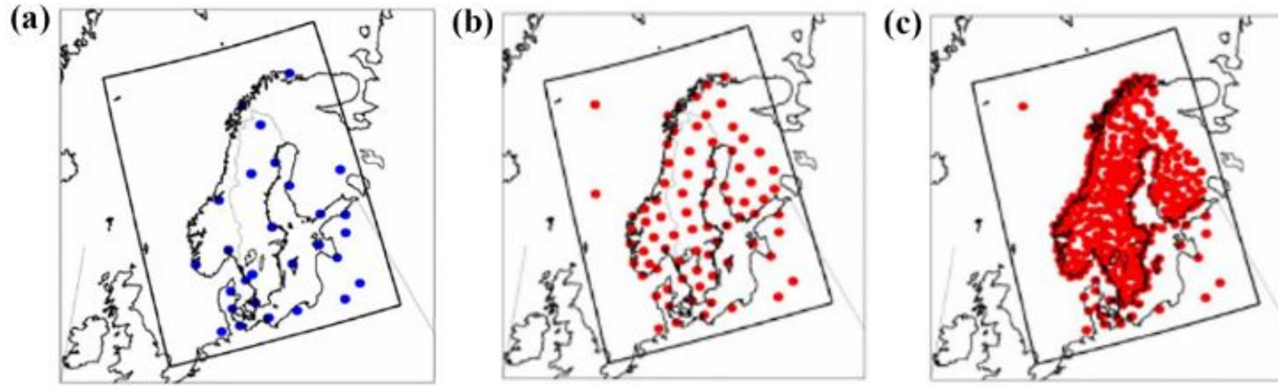
Identification of Gross errors

- Blacklist
- First guess check
- Buddy check
- Variational quality control
- Adaptive methods
- ….

Modelling/representation of errors

Thinning

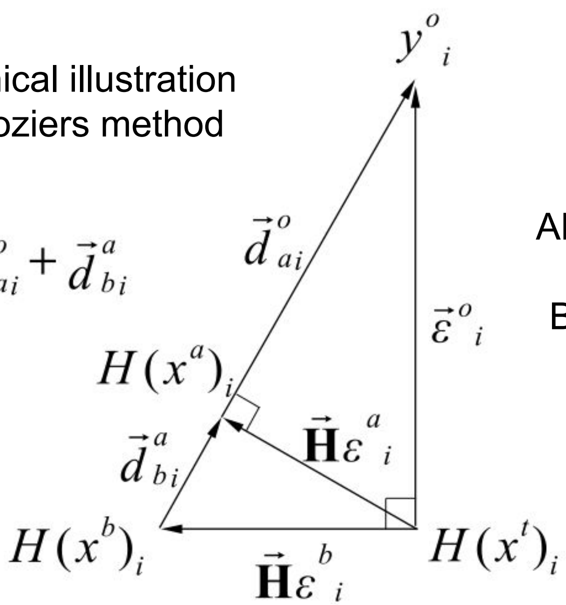
Illustration GNSS thinning a) 90 km, b) 40 km, c) original raw (from Lindskog et al., 2017)



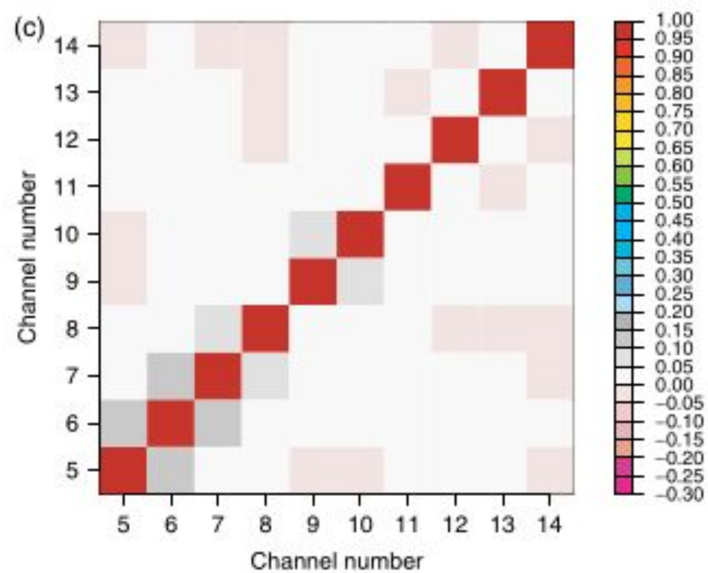
Standard deviations and correlations

Graphical illustration Desroziers method

$$\vec{d}^{\,o}_{bi} = \vec{d}^{\,o}_{ai} + \vec{d}^{\,a}_{bi}$$



AMSU-A inter-channel correlations (from Bormann and Bauer, 2010)



Modelling/representation of errors

Identification of Gross errors

First guess check

$$([H(\boldsymbol{x}_{\mathrm{b}})]_i - y_i)^2 / \sigma^2_{\mathrm{b},i} > \mathrm{FgLim} \times \lambda$$

Buddy check

$$p = p\left(\frac{|x_i - x_i^\star|^2}{S_{ii}^\star} > \tau^2\right) = \sqrt{\frac{2}{\pi}} \int_\tau^\infty e^{-t^2/2} dt.$$

We partition the residual vector $\mathbf{v}$ as

$$\mathbf{v} = \begin{bmatrix} \mathbf{x} \\ \mathbf{y} \end{bmatrix}, \qquad (3)$$

where $\mathbf{x}$ contains the residuals associated with suspect observations, and $\mathbf{y}$ those associated with buddies. We then define corresponding blocks of the residual covariance,

$$\mathbf{S} = \begin{bmatrix} \mathbf{S}_x & \mathbf{S}_{xy} \\ \mathbf{S}_{xy}^T & \mathbf{S}_y \end{bmatrix}.$$

Variational QC

$$p_i = P(G_i)F_i + (1 - P(G_i))N_i$$

FIGURE 1: probability density function P(Gi)=0.05 l=5 sigma=1.

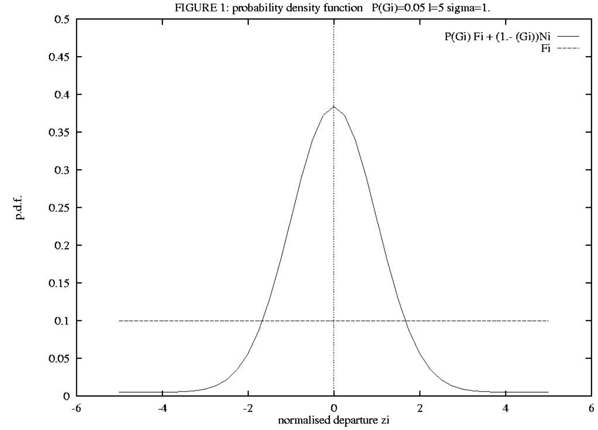
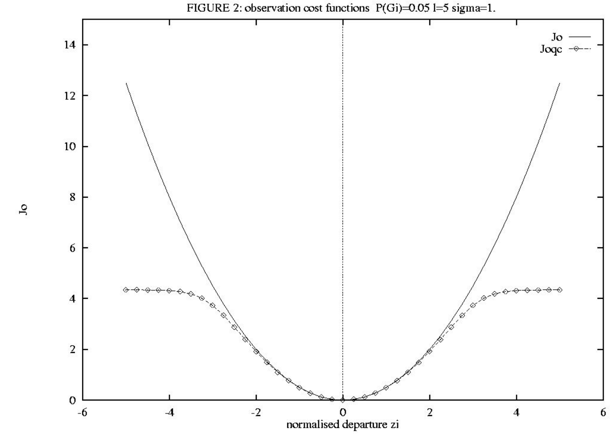
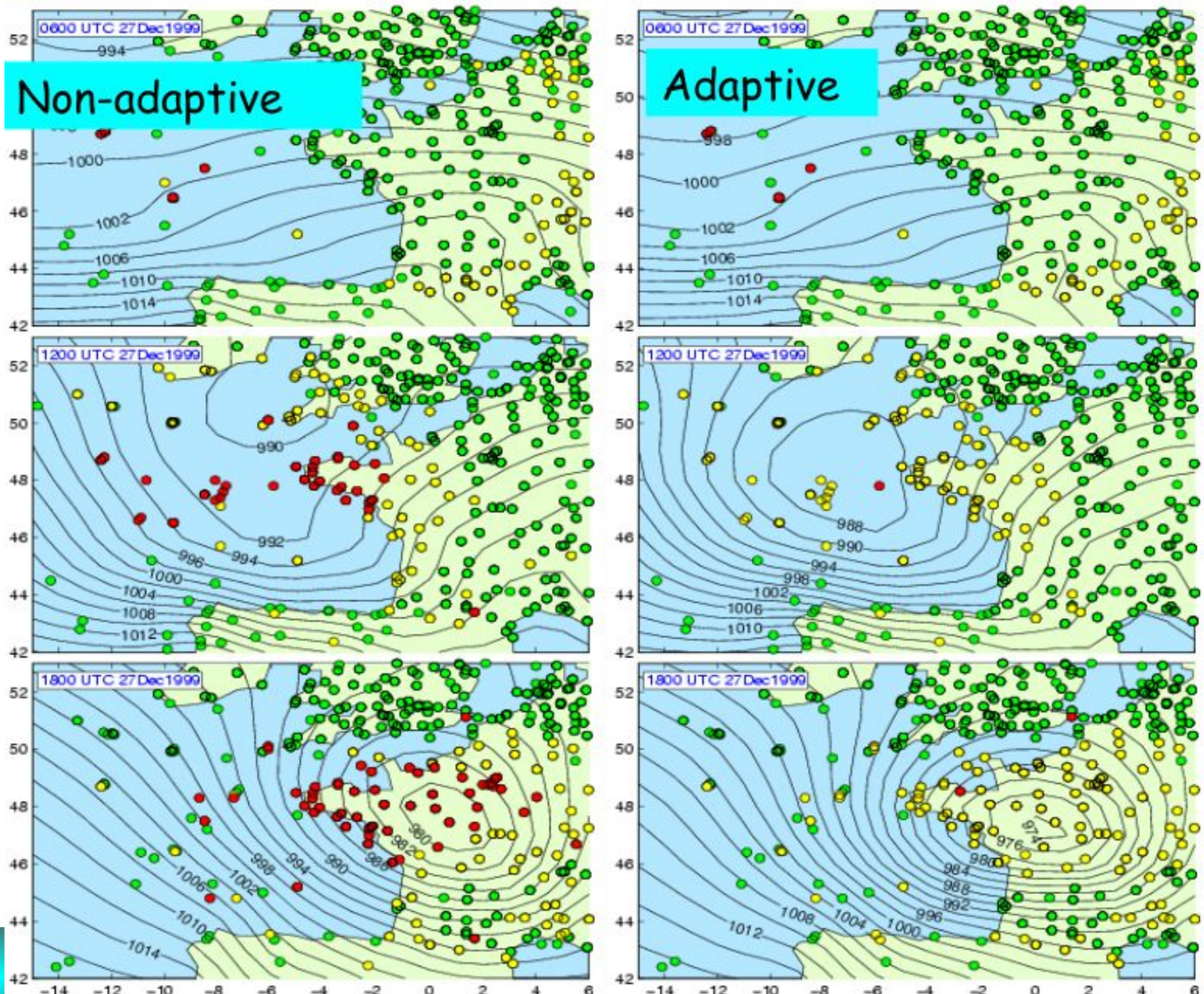


FIGURE 2: observation cost functions P(Gi)=0.05 l=5 sigma=1.



Modelling/representation of errors



Adaptive buddy check

flow- dependent tolerances for outlier observations

(Dee et al, QJRMS, 2001)

Dec 1999 storm

- rejected by buddy check
- passed buddy check
- passed FG check

Discussion points

- Alternative methods for estimation of observation errors.
- Plans/experiences for modelling of observation error correlations (spatial and inter-channel correlations).
- Alternative ways of handling of observation biases?
- Methods and ideas for rejected observations? Adaptive, ensemble based, machine learning?
- Exchange of blacklist information?
- Inter-consortia comparisons of error estimation and handling?

Extra slides

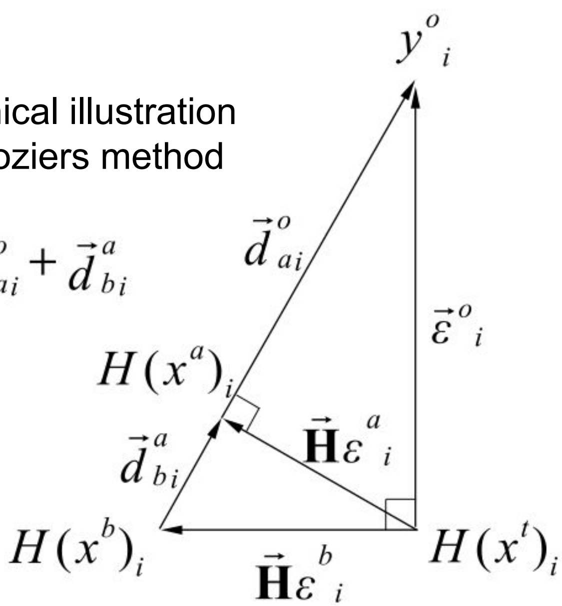
Desroziers method Q. J. R. Meteorol. Soc. (2005), 999, pp. 1–999

AIM: To compare used background- and observation-errors with theoretical ones calculated by Desroziers method and exploit if revisions of error standard deviation specifications needed.

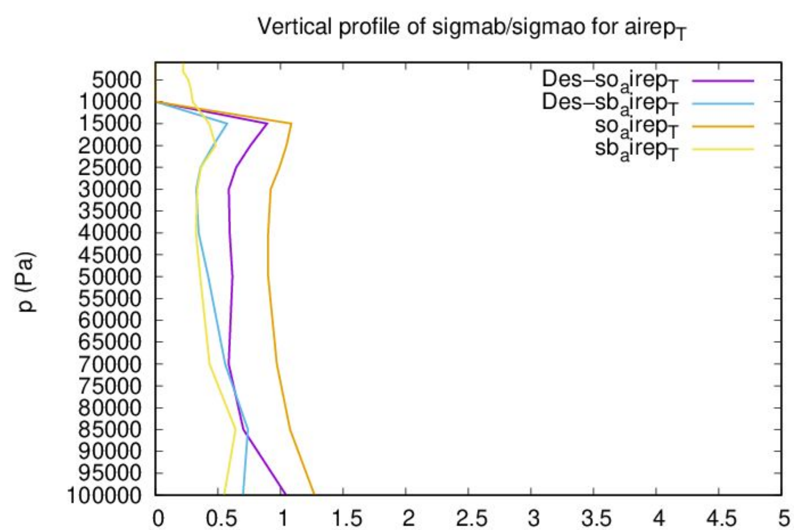
HOW: Use DA feedback statistics of residuals and innovations from parallel cy46 evaluation experiments. Investigate plots of the current prescribe and the by desroziers method suggested observation and background error values.

Graphical illustration Desroziers method

$$\vec{d}^{\,o}_{bi} = \vec{d}^{\,o}_{ai} + \vec{d}^{\,a}_{bi}$$



Example for ABO temperatures



background- and observation-errors (K)

Extra slides

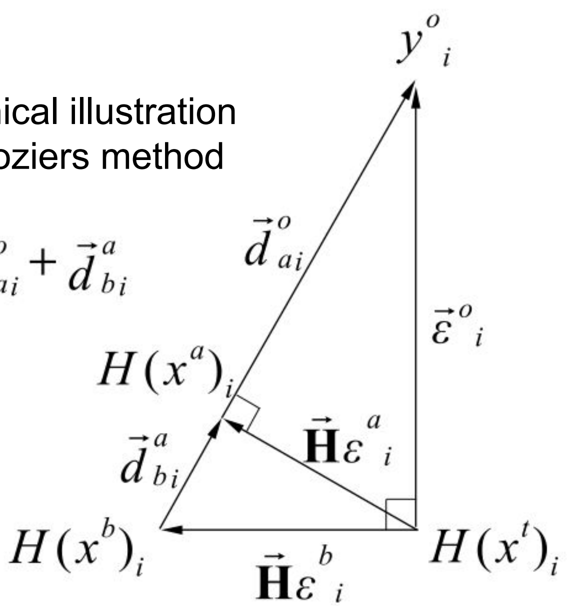
Desroziers method Q. J. R. Meteorol. Soc. (2005), 999, pp. 1–999

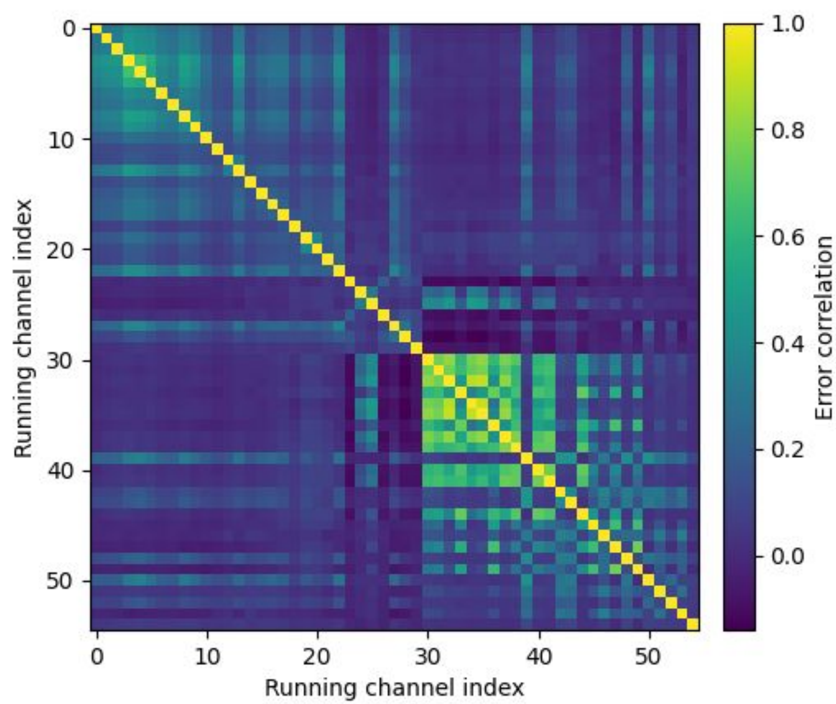
AIM: Estimate inter-channel observation error correlations

HOW: Use DA feedback statistics from innovations and residuals.

Graphical illustration
Desroziers method

$$\vec{d}^{\,o}_{bi} = \vec{d}^{\,o}_{ai} + \vec{d}^{\,a}_{bi}$$





METOP-B IASI INTERCHANNEL CORRELATION (K)

Extra slides

Hollingsworth-Lönnberg method Tellus(1986), 38A, 111-136

AIM: To estimate background and observation errors.

HOW: Extrapolate innovations to zero distance and derive error standard deviations from intersect with x-axis. Correlations from assumptions on either background or observation errors.

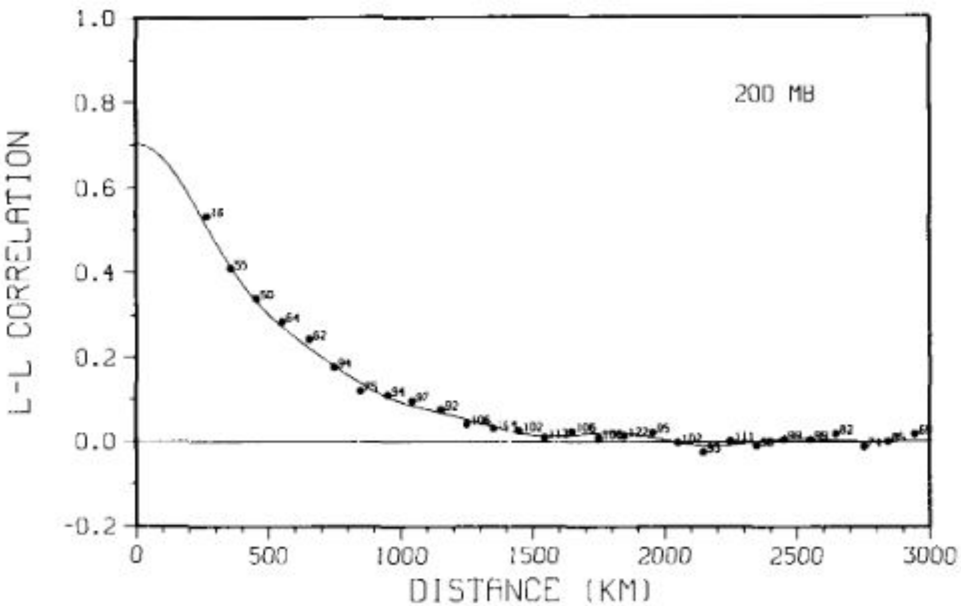


Fig. 5. The variation at 200 mb of the $\langle l, l \rangle$ or longitudinal correlation with station separation: the dots show the empirically determined average value for each 100 km "bin", together with the number of station pairs in that bin. All the data out to 3000 km was used in the least squares procedure to determine the fitting curve with a truncation of 10 synoptic scale terms.

Extra slides

Andersson&Järvinen Q. J . R. Meteorol. SOC. (1999). 125, pp. 691-122

AIM: To select appropriate check limits, FgLim for background check. Assumption is that observations with errors outside Gaussian distribution are affected by Gross errors and should be removed prior to the data assimilation.

$$([H(\boldsymbol{x}_\mathrm{b})]_i - y_i)^2/\sigma^2_{\mathrm{b},i} > \mathrm{FgLim} \times \lambda$$

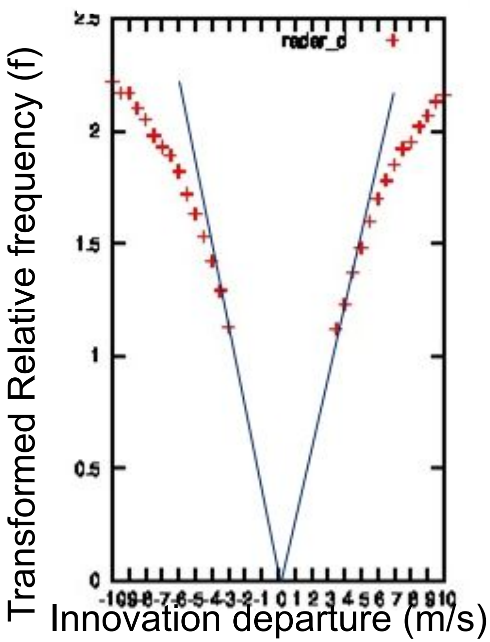
where $\lambda = 1 + \sigma^2_{\mathrm{o},i}/\sigma^2_{\mathrm{b},i}$, FgLim is the rejection limit and $[H(\boldsymbol{x}_\mathrm{b})]_i$ denotes the projection of the model state on y_i observation, where the potential observation bias has been accounted for. $\sigma_{\mathrm{o},i}$ and $\sigma_{\mathrm{b},i}$ are the standard deviation of the observation error and background error equivalent, respectively.

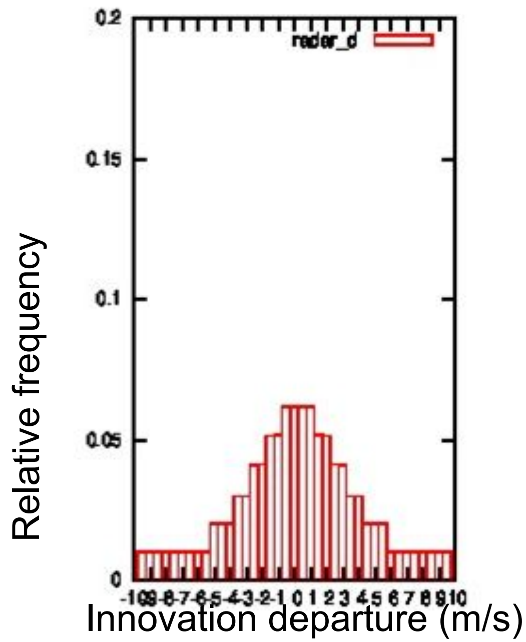
HOW: Plots histograms and transformed histograms of innovations to identify when distribution starts to deviate from Gaussian and where to put rejection limit.

Example for radial winds:

Transformation formula:

$$\tilde{f} = \sqrt{-2\ln[f/\max(f)]}$$





A Consortium for COnvection-scale modelling Research and Development

Extra slides

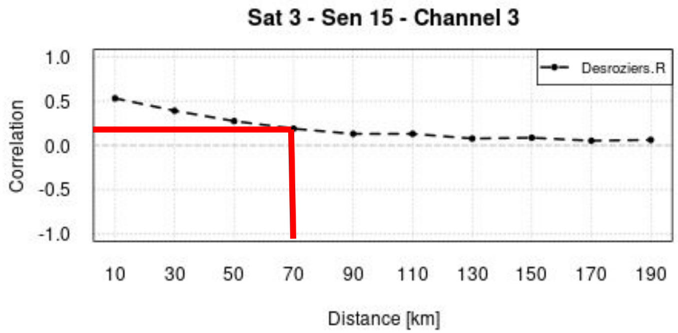
Obstool (based on P .Benachecks developments)

AIM: To set the thinning distances applied to high spatial density data in accordance with estimated observation error correlation length scales. The spatial thinning is applied both to limit data amounts and to compensate for our current lack of representation of spatial observation error correlations.

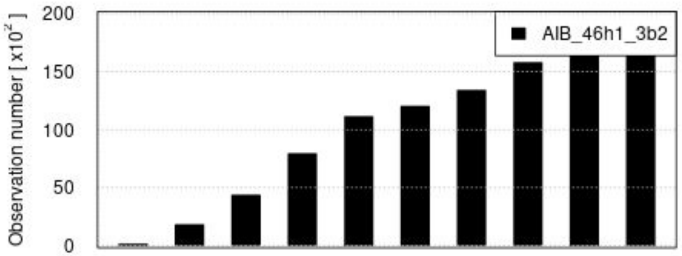
HOW: Based on DA feedback statistics files, innovations are separated into observation error correlations and background error correlations. From plots of the observation error correlation part, appropriate thinning distance is estimated with distance when the observation correlation drop to 0.2.

Example for satellite MHS channel 3 data

Derived observation error correlation as function of distance between data pairs.



Number of data in each bin as function of distance between data pairs.



ACCORD
A Consortium for COnvection-scale modelling Research and Development