Met Office

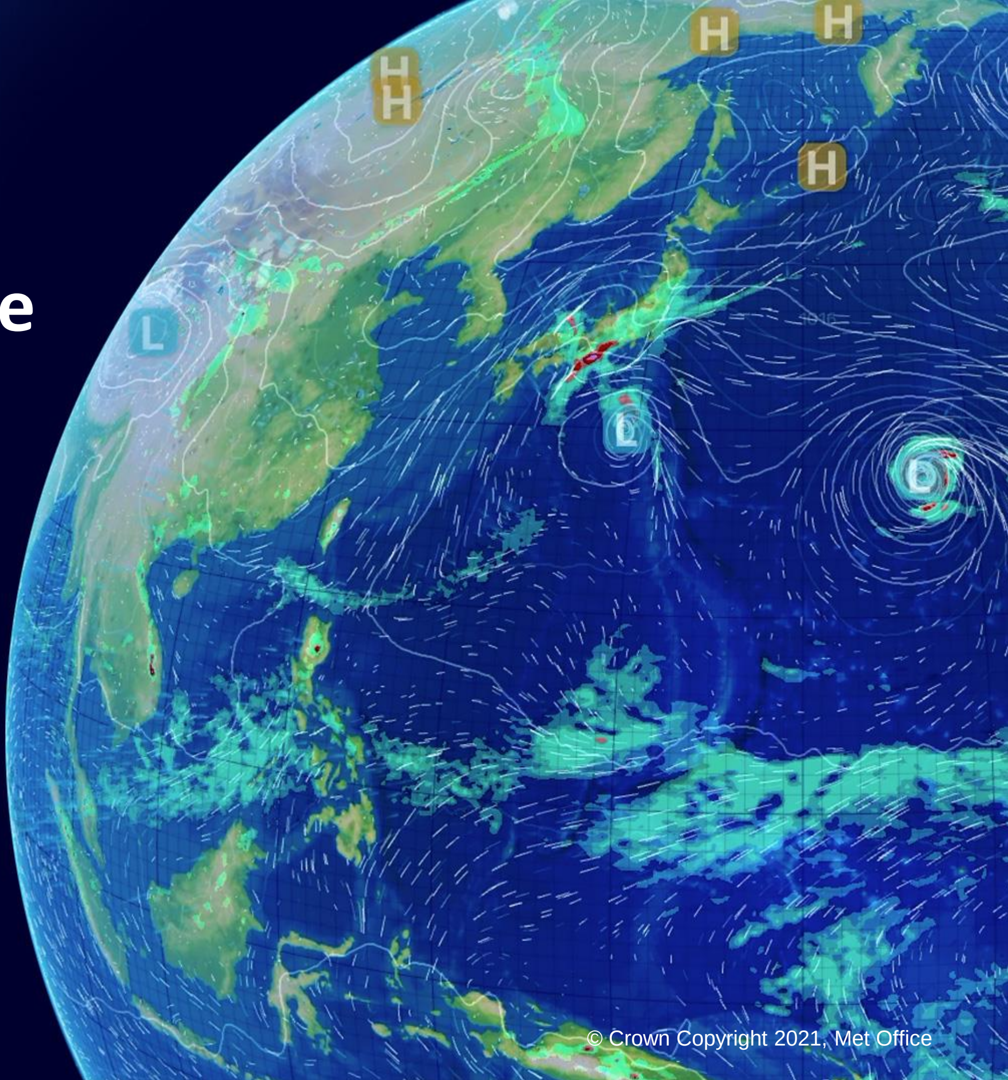
Observation error covariances, can we handle spatial observation error correlations?

ECMWF Annual Seminar

13th – 17th September 2021

Joanne A. Waller

With thanks to D. Simonin, S. L. Dance, N. K. Nichols, S. P. Ballard, C. Charlton-Perez and many other colleagues



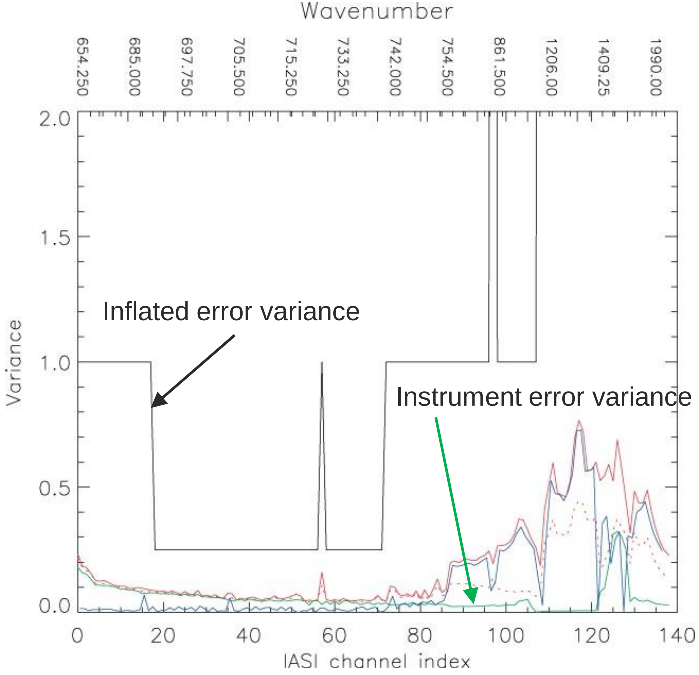
Overview

- Introduction
- Benefits of accounting for spatial error correlations
- Methodology for using spatial correlations
- Illustration with Doppler radar winds
- Conclusions

Introduction

Motivation

- For many years the observation error correlation matrices, $\boldsymbol{R}$, used in assimilation systems had been diagonal:
  - For computational speed and efficiency
  - As observation error statistics were poorly known
- To account for any correlations that were neglected:
  - the observations are thinned
  - the error variances are inflated



Error variance for IASI Channels

Figure from Stewart et al. 2013

Challenging assumptions

Original focus mainly on inter-channel correlations since technically simpler to implement:

- Finite number of channels => Small fixed observation error covariance matrix
- Single position for all observations with correlated error

Still a need to tune $\boldsymbol{R}$ to ameliorate issues with slow convergence or filter divergence (e.g. Tabeart et al. 2020).

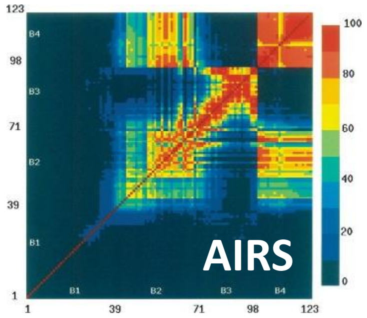


Figure from Garand et al 2007

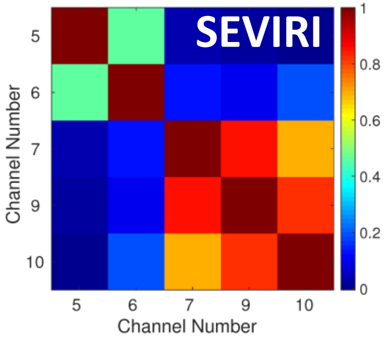


Figure from Waller et al 2016a

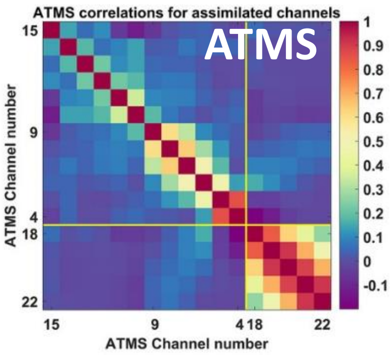


Figure from Campbell et al. 2017

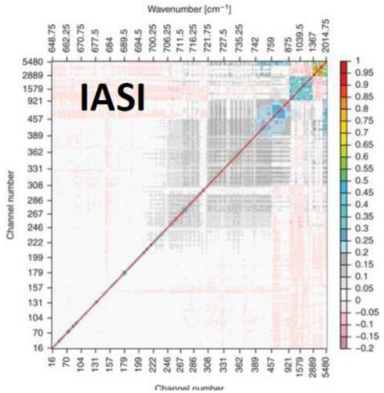


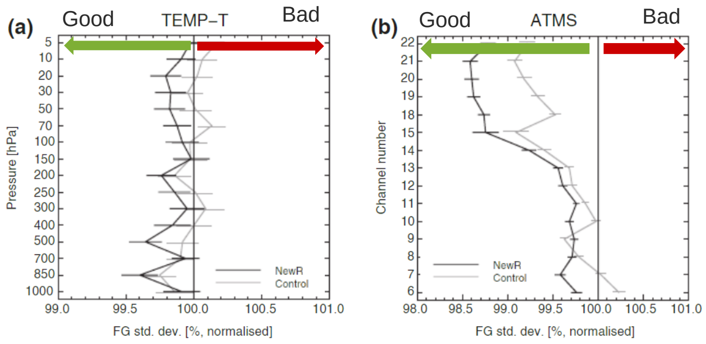
Figure from Bormann et al. 2016

Met Office

Motivation

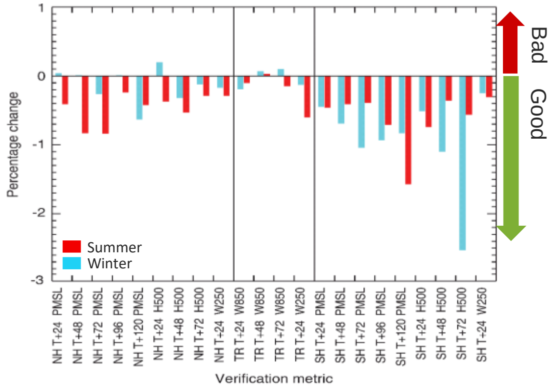
In global numerical weather prediction for a number of different satellite instruments it has been shown that inter-channel correlated errors leads to:

- Increase in the analysis accuracy.
- Improved fit of background to observations.
- Improvement in the forecast skill score.



Global standard deviations of background departures for (a) radiosonde temperature observations; (b) ATMS brightness temperatures.

Figure from Bormann et al. 2016

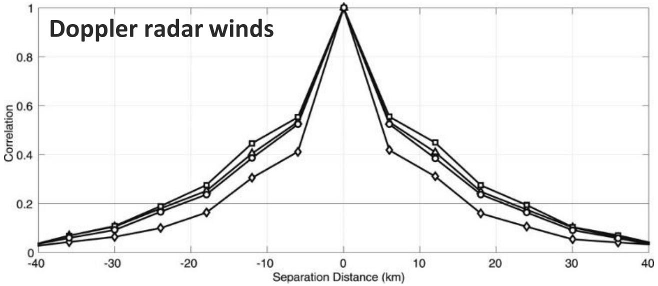


Change in RMSE and weighted skill against observations when accounting for correlated errors for IASI.

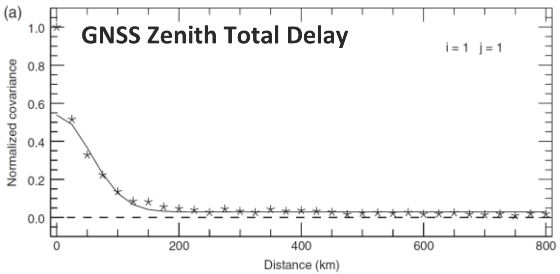
Figure from Weston et al. 2014

Spatial error correlations

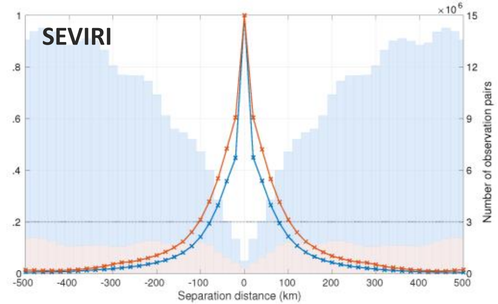
Inter-channel error correlation success sparked interest in spatial error correlations.



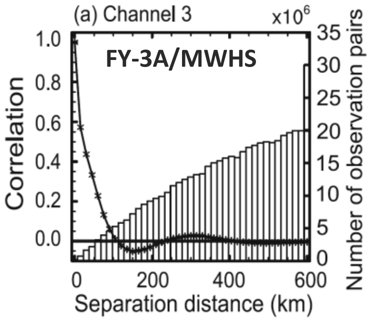
[Figure from Waller et al. 2016b](#)



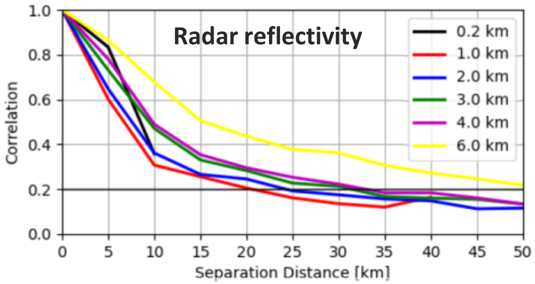
[Figure from Macpherson & Laroche 2019](#)



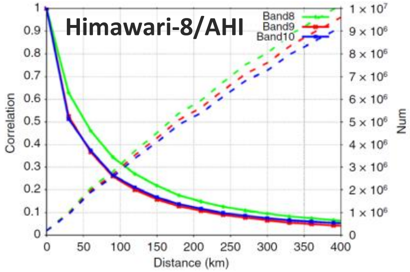
[Figure from Waller et al 2016a](#)



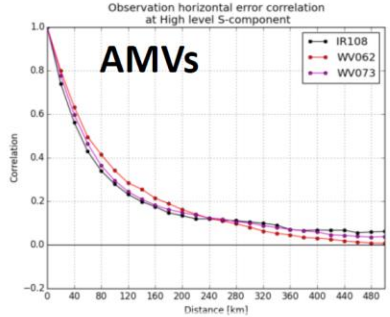
[Figure from Wang et al. 2018](#)



[Figure from Zheng et al. 2021](#)



[Figure from Okamoto et al. 2019](#)



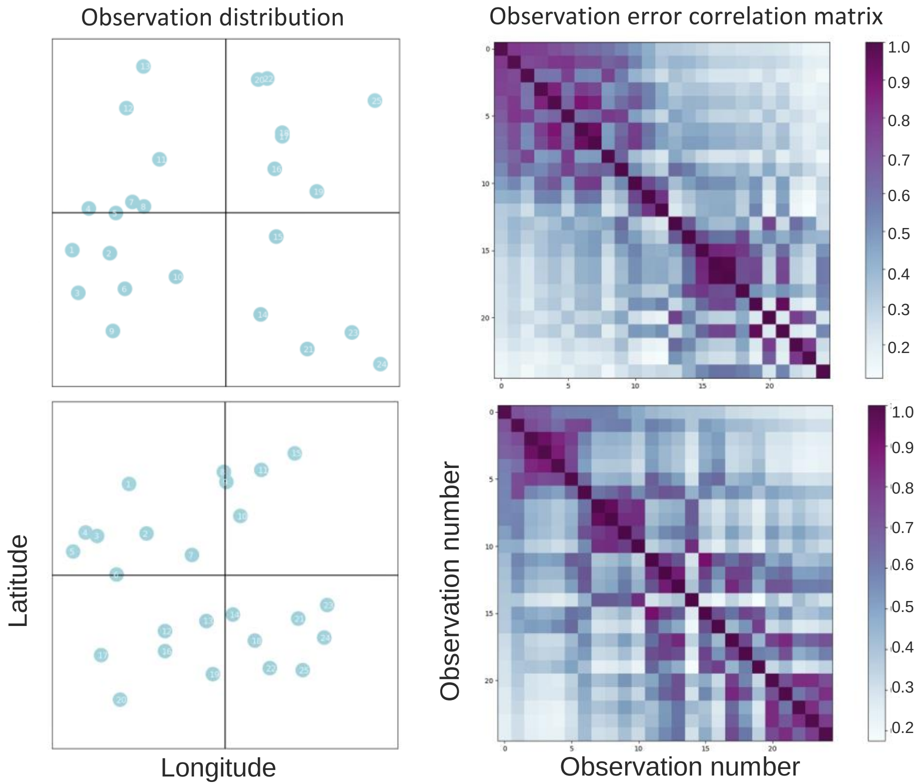
[Figure from Cordoba et al. 2017](#)

Can we efficiently account for them in assimilation?

Met Office

The challenge

- Challenge is in technical implementation because a spatial observation error correlation matrix:
  - May be very large
  - Likely to change every assimilation cycle
- Observations spread over spatial domain which is more complex for parallel computation.
- May need to combine with existing inter-channel error correlations.



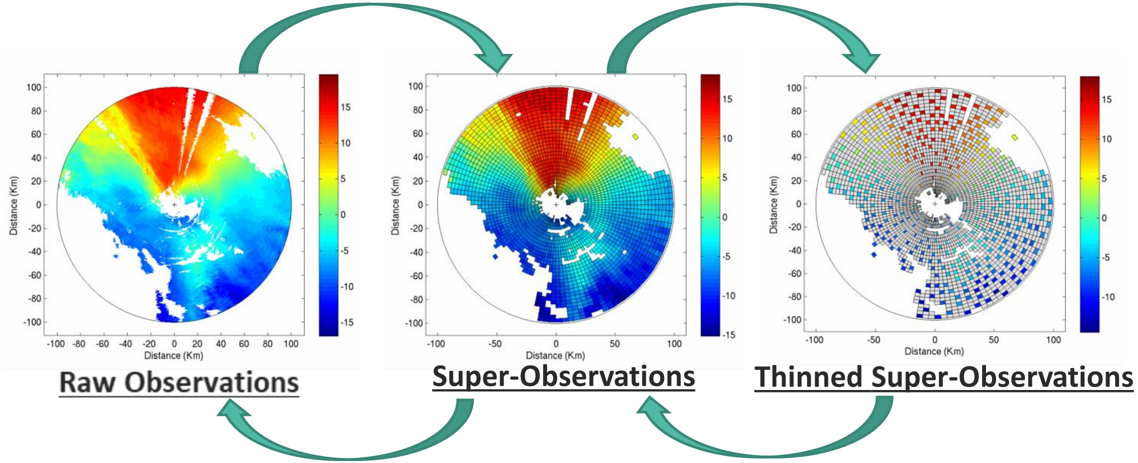
Benefits

Benefits of accounting for spatial correlation

Improved use of available observations

Reduce the need to thin data as no need to satisfy uncorrelated error assumption. This allows:

- Better use of available information – better value for money!
- Higher resolution information to be ingested into the analysis.
- Will be critical for sub-kilometre NWP



Benefits of accounting for spatial correlation

Met Office

Improved analyses and forecast

- Given the success of inter-channel correlations it is expected that accounting for spatial error correlations will improve analysis and forecast accuracy.
- Work in idealised systems has demonstrated that not accounting for spatial error correlations can lead to analyses that are grossly suboptimal (Rainwater et al. 2015).

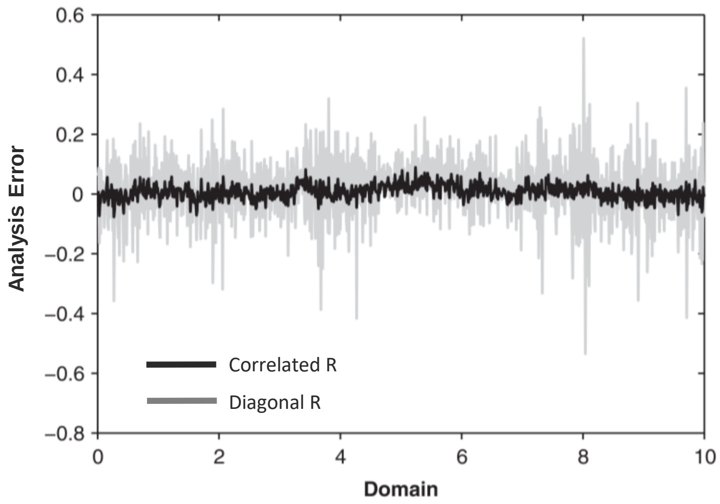
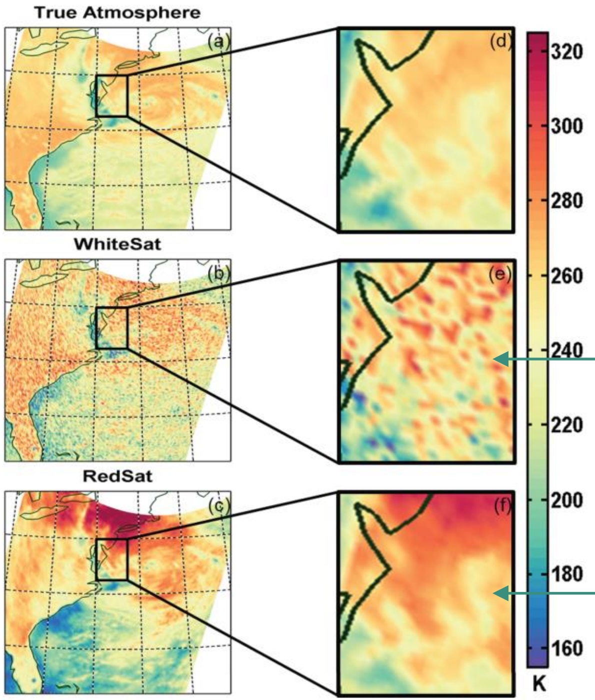


Figure from Stewart et al. 2013

Benefits of accounting for spatial correlation

Scales

- High-density observations with correlated error contain more information about small resolvable scales.
- Including spatial error correlations in the assimilation can reduce significantly reduce error in the small scale.
- Hence correctly accounting for spatial error correlations will be particularly important for high resolution NWP and prediction of high-impact weather e.g. tropical cyclones, thunderstorms.



Truth with uncorrelated noise added
- Larger scale information is close to the truth
- Detailed information is lost

Truth with correlated noise added
- Large scale information wrong magnitude.
- Detailed structure retained.

Figure from Rainwater et al. 2015

Potential Methods

Methodology

How can we handle spatial error correlations?

Account for correlation information but keep diagonal $\boldsymbol{R}$:

- Use variables or transform to spaces where errors are uncorrelated (e.g. Chabot et al. 2020)
- Assimilate observations and 'spatial difference' observations (Bédard and Buehner, 2020)

Explicitly use correlated observation error matrix:

- Make use of observation families – *Family method*
- Consider a diffusion-based approach – *Diffusion method*
- Make use of small $\boldsymbol{R}$ matrices in localized ensemble assimilation schemes (e.g. See WCRP-WWRP symposium posters from Terasaki et al. and Yeh et al.).

Methodology

Family method

- A family is a group of observations that have correlated error (the specific observations in each family may change each assimilation cycle).
- Traditional approach: send observations to processors based on location or equal distribution
- New approach: keep families on single processors.

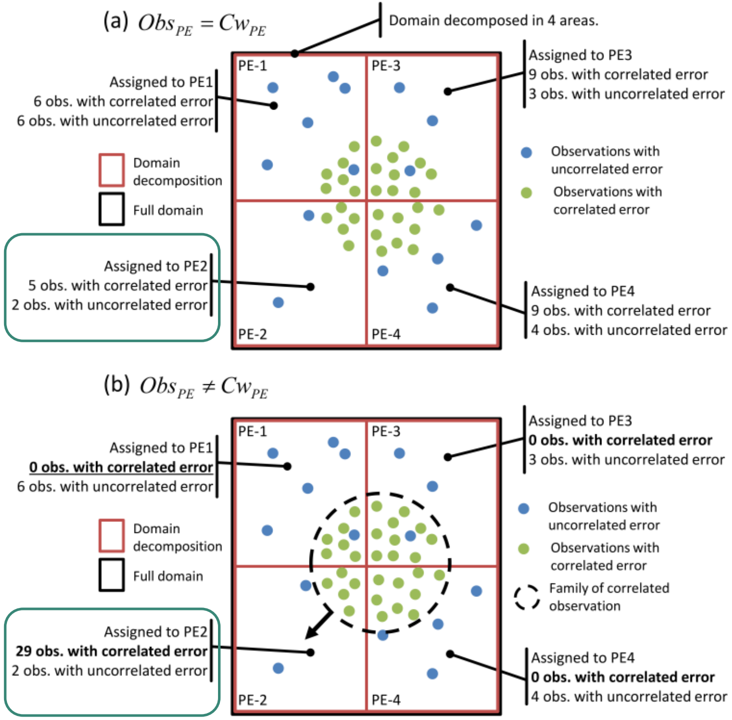


Figure from Simonin et al. 2019

Methodology

Family method - Error covariance matrix

- For each family the observation error covariance matrix $\boldsymbol{R} = \boldsymbol{DCD}^{\boldsymbol{T}}$ is defined on the fly for each assimilation cycle.
- The correlation matrix, $\boldsymbol{C}$, is determined by calculating the distance between each pair of observations in the family and using the separation distance, $\Delta y_{i,j}$, in a correlation function e.g.

$$C_{i,j} = exp\left(\frac{-|\Delta y_{i,j}|}{L_r}\right),$$

where parameters such as length scale, L_r, must be predetermined.

- Use a Cholesky decomposition to calculate the sensitivity: $\boldsymbol{Q} = \boldsymbol{R}_s^{-1}(\boldsymbol{y} - H(\boldsymbol{x}))$.

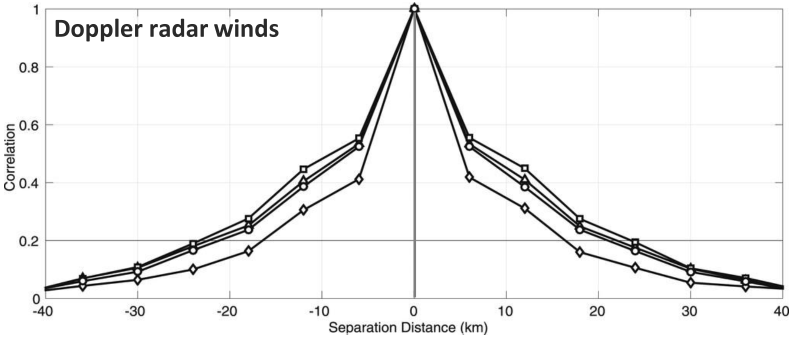


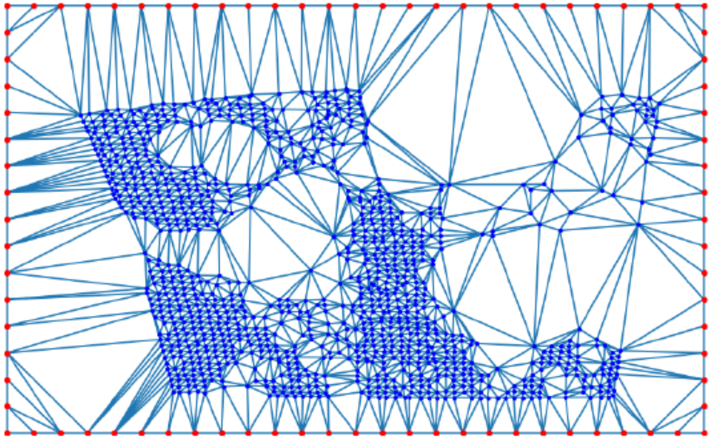
Figure from Waller et al. 2016b

Family method - pros and cons

Pros	Cons
Simple to apply - application of ***R*** similar to inter-channel approach	Sets of observations with correlated errors must be sufficiently small.
R used in assimilation is not further approximated.	Large correlation length scales may be costly to handle

Diffusion Method

- The method is based on the numerical solution of a diffusion equation.
- Has been commonly used for modelling $\boldsymbol{B}$ in ocean applications with structured meshes (e.g. Weaver and Courtier, 2001).
- Expanded by Guillet et al (2019) to be applied on unstructured mesh which is necessary when dealing with observations.



(a) SEVIRI observations.

Unstructured finite element mesh for SEVIRI observations.

jo.waller@metoffice.gov.uk

Met Office

Diffusion Method - Overview

- Obtain a diffusion operator by integrating a diffusion equation over a finite number of steps using a backward Euler scheme.
- The solution of the resulting equation can be interpreted as a correlation operator whose kernel is a correlation function from the Matérn family.
- Using the finite element method to solve the equation results in mass, $\boldsymbol{M}$, and stiffness, $\boldsymbol{K}$, matrices that can approximate the inverse of the correlation matrix :

$$\boldsymbol{R}^{-1} = \boldsymbol{D}^{-1}\boldsymbol{C}^{-1}\boldsymbol{D}^{-1} = \boldsymbol{D}^{-1}\boldsymbol{\Gamma}^{-1}[\boldsymbol{M}^{-1}(\boldsymbol{M}+\boldsymbol{K})]^{m}\boldsymbol{M}\boldsymbol{\Gamma}^{-1}\,\boldsymbol{D}^{-1}$$

Note $\boldsymbol{\Gamma}$ is a normalisation matrix.

Examples of Matérn function

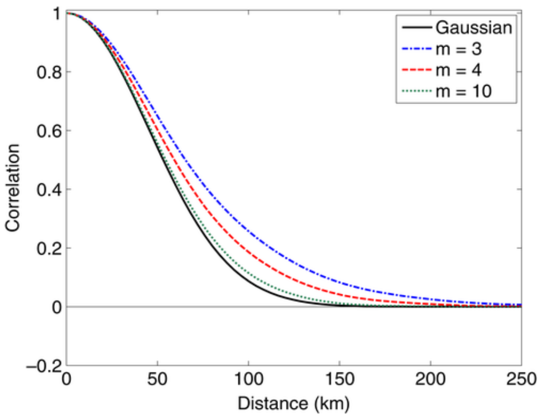
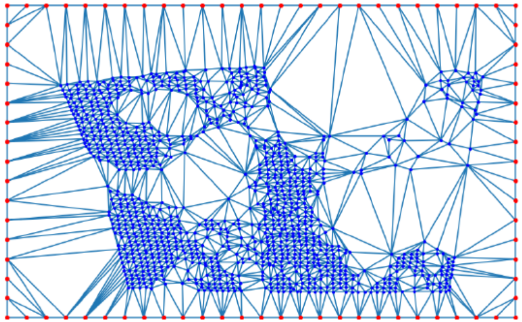


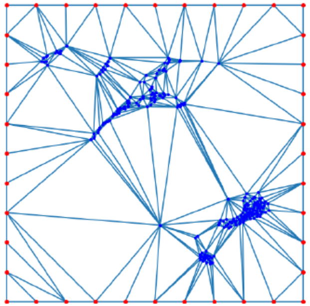
Figure from Guillet et al. 2019

Approximation errors

- Quality of mesh impacts quality of approximated correlation matrix.



(a) SEVIRI observations.



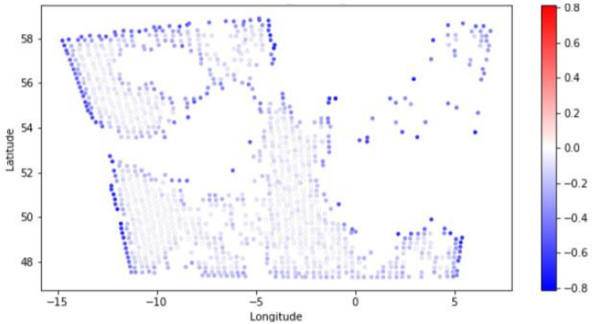
(b) Doppler radial wind observations.

Mesh for different observation types.

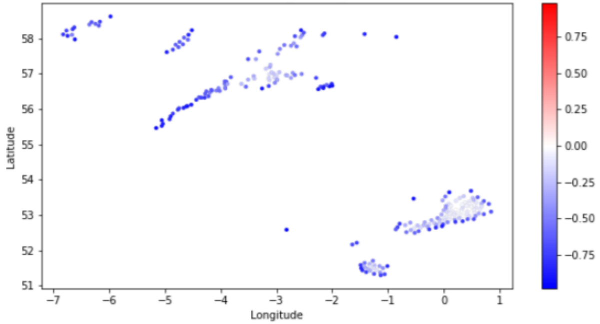
Approximation errors

- Quality of mesh impacts quality of approximated correlation matrix.
- Correlations well estimated in dense data areas
- Statistics poorly estimated for observations:
  - Near domain boundaries
  - On cluster boundaries
  - For sparse observations

- Method best suited to data in dense clusters.



(a) SEVIRI observations.



(b) Radar observations.

Error in approximated error variance.

Diffusion Method - pros and cons

Pros	Cons
Do not need to calculate full ***R,*** have immediate access to $\boldsymbol{R}^{-1}$	Restricted to Matérn functions
R and sensitivity calculation easily parallelizable as only sparse matrix computation is required	Mesh will need to be calculated every cycle and this may not be easily parallelizable
Applicable to large observations subsets	Quality of matrix approximation dependent on quality of mesh – method may not suit observations in small dense clusters
Mesh refinement could be linked to observation selection/thinning	Need determine how to use spatial and inter-channel error correlations simultaneously

Doppler radial wind illustration

Met Office

Radar observations

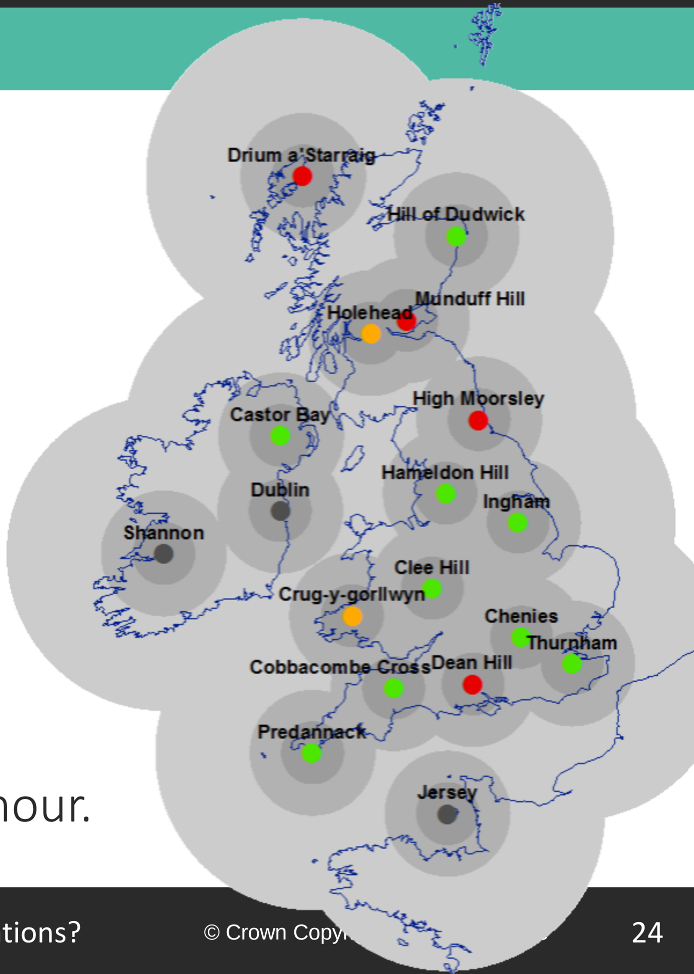
UK Radar Network

- 18 radars currently in the network
- PPI Scan elevations between $1°$and $9°$
- $1°$ beamwidth, $600/300/75$ m gate size
- Volume scan available every 5 minutes

Provides observations of:

- Reflectivity
- Refractivity
- Doppler wind

Resulting in approximately 30 million observations per hour.



Assimilating doppler radial winds

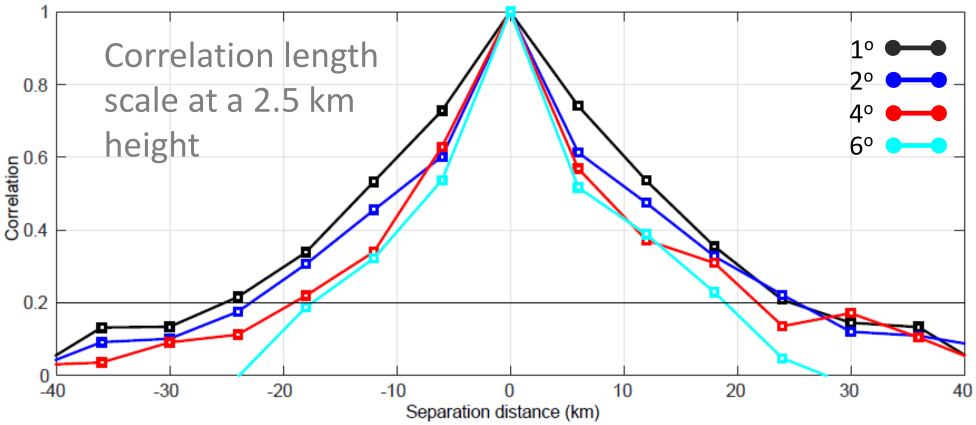
Met Office

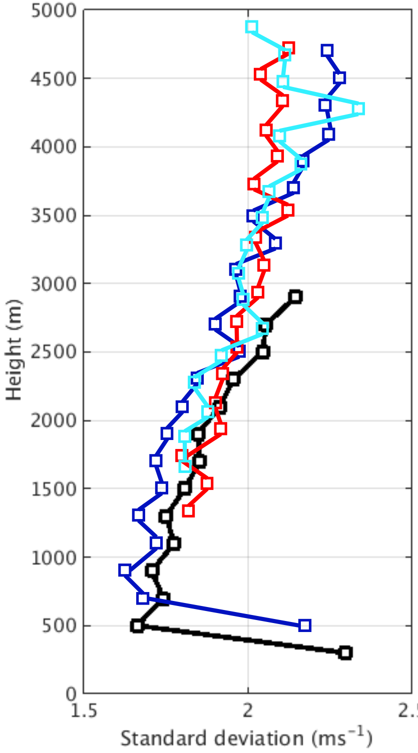
Doppler wind error statistics

Doppler Radar radial wind error statistics have been shown to be spatially correlated:

- Correlation length scale range from ~15km to ~30km.
- Correlation length scale increase with distance from radar.

Figure from Waller et al. 2016b



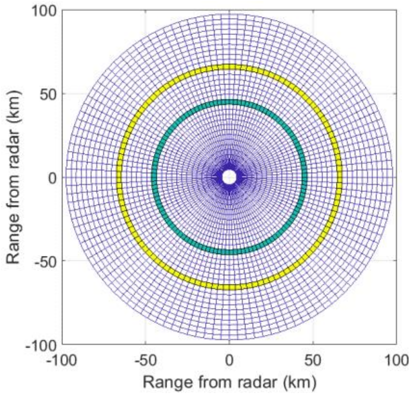


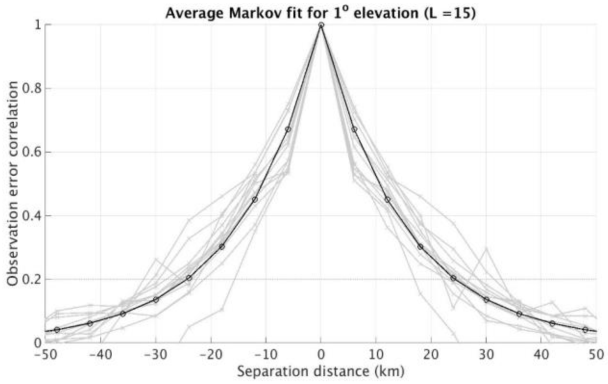
Assimilating doppler radial winds

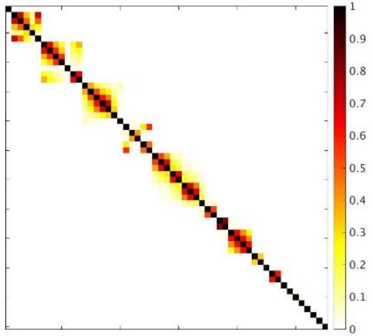
Accounting for correlated error

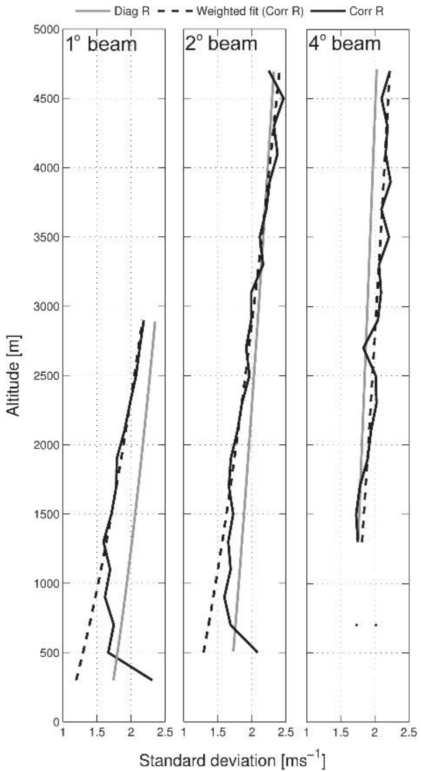
Use families methodology

- Each scan is divided in altitude bands (200m in thickness).
- Bands at a given height from a given scan elevation is assigned to a family.
- Fit functions to estimated observation error statistics.









jo.waller@metoffice.gov.uk

Met Office

Experimental Design

Simonin et al. 2019 trial setup:

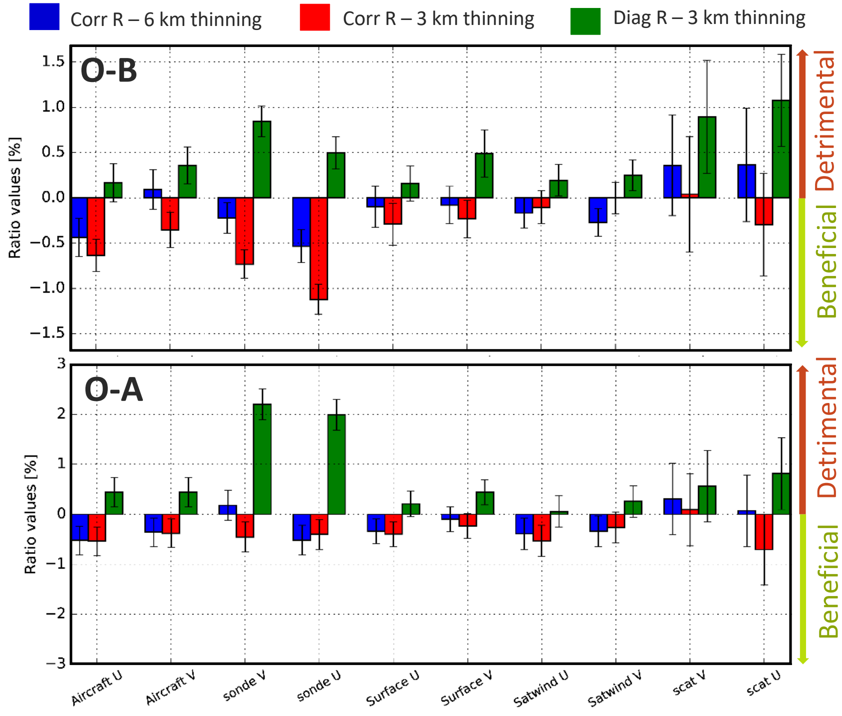
- Met Office convection permitting UKV with 3 hourly 3D Var assimilation.
- 1 scan per radar at the center of the assimilation window.
- Three weeks from the 1st April 2016
- Four experiments

Experiment	Doppler wind observation error matrix	Doppler wind super-observation thinning distance
Control	Diagonal observation error covariance matrix (Operational)	6 km (~2000 rad obs. per cycle)
Corr-R-6km	Correlated observation error covariance matrix	6 km (~2000 rad obs. per cycle)
Corr-R-3km	Correlated observation error covariance matrix	3 km (~8000 rad obs. per cycle)
Diag-R-3km	Diagonal observation error covariance matrix (Operational)	3 km (~8000 rad obs. per cycle)

Assimilating doppler radial winds

Impact on innovations and residuals

- Overall improved fit to observations when observations assimilated with correlated errors.
- Further improvement to some innovations and residuals when observation density is increased.
- Increasing observation density without accounting for error correlations has some considerable detriment to fit to observations.



Trial average standard deviation ratio between the experiments and the Control

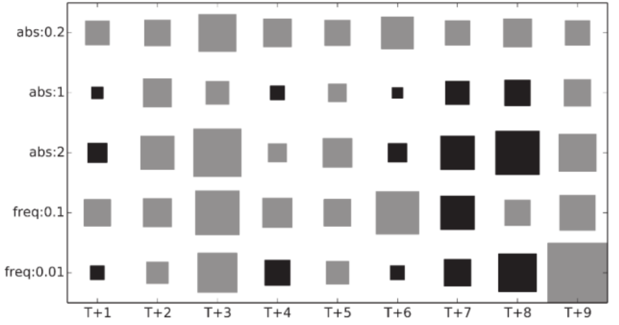
Figures from Simonin et al. 2019

jo.waller@metoffice.gov.uk

Impact on precipitation forecast

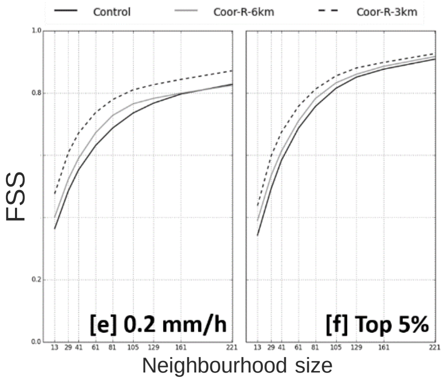
When accounting for error correlations but not increasing observation resolution there is no real impact on the precipitation forecast.

When increasing observation density whilst accounting for error correlations there are small improvements.

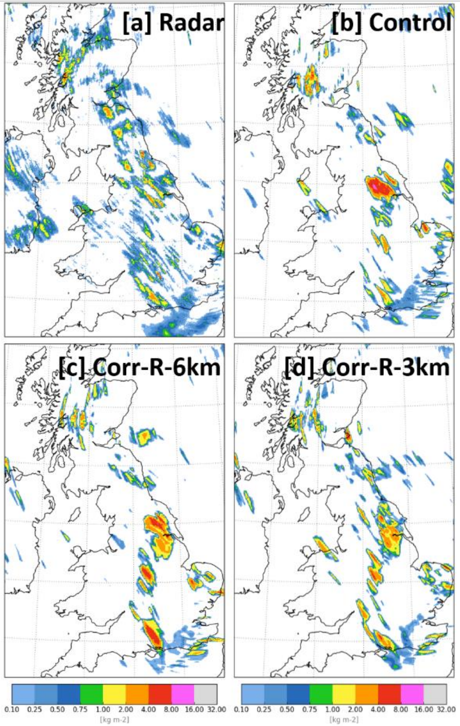


Trial average FSS differences between the Corr-R-3km and the Control experiment

Figures from Simonin et al. 2019



Equivalent fraction skill score as a function of neighbourhood size (1500 UTC on 7 April 2016)



Hourly accumulated precipitation forecasts for 1500 UTC on 7 April 2016 at at T+3.

Assimilating doppler radial winds

Met Office

Impact on the performance of the assimilation system

Using a full observation error covariance matrix is expected to have some impact on the cost of the assimilation system since:

- There is additional communication between processors.
- There is additional cost related to the computation and application of $\boldsymbol{R}$.

Experiment	Trial average iterations	Trial average time (s)
Control	27.4	272
Corr-R-6km	27.7	293
Corr-R-3km	28.2	288

However, the overall cost is manageable and not really significant compared to the overall cost of a forecast.

Met Office

Conclusions

Conclusions

Summary

- Methodologies emerging that can be used to handle spatial observation error correlations. Though different methods appropriate in different situations.
- Beginning to see benefit of using spatial error correlations. These include:
  - Improvement to analysis and forecasts.
  - Improved use of high resolution observations.
  - Injection of small scale information.
- May be critical for sub-kilometre NWP
- Technical challenges still remain:
  - Use of very large datasets with correlated errors.
  - Use of correlated errors in muti-dimensions (e.g. horizonal and inter-channel, temporal).
  - Getting good estimates of the observation error statistics! (not discussed here)

Thank you.

Any questions?

jo.waller@metoffice.gov.uk