DIAGNOSING OBSERVATION ERROR STATISTICS FOR NUMERICAL WEATHER PREDICTION

J. Waller, S. Dance, N. Nichols (University of Reading)
D. Simonin , S. Ballard, G. Kelly (Met Office)

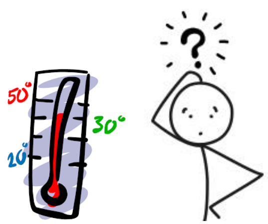
AIMS

OBSERVATION ERRORS

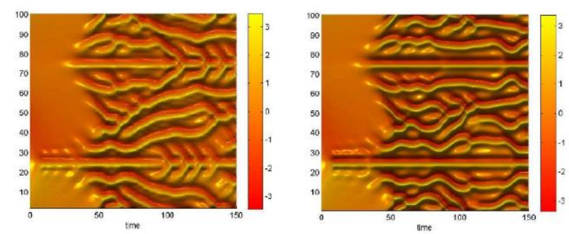
- Only 20% of observations are utilised in atmospheric data assimilation.
- This is in part due to the unknown observation error statistics.

Observation errors arise from four main sources:

- Instrument error

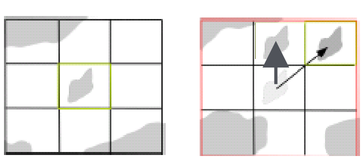


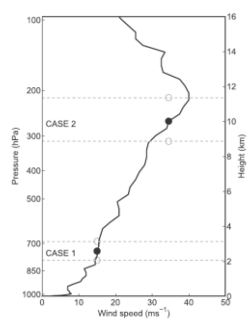
- Representativity error



Low resolution High resolution

- Observation pre-processing





- Observation operator error

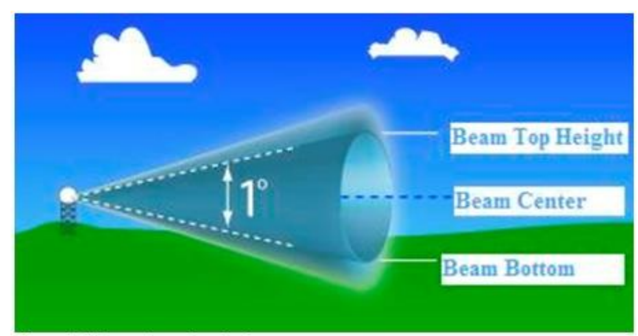
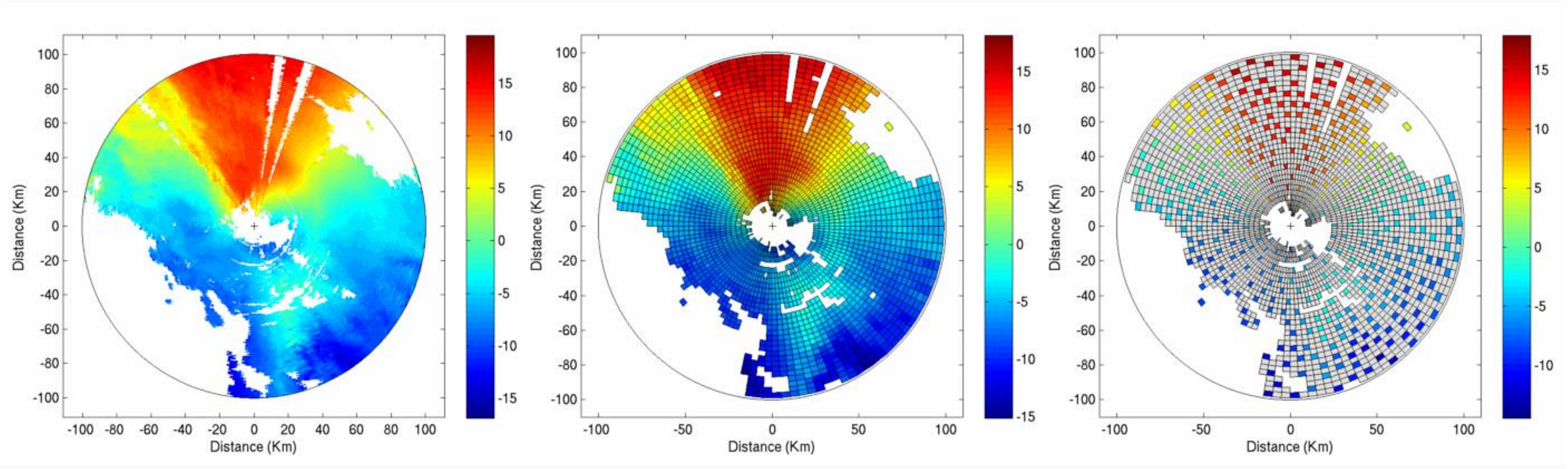


Figure 1 Illustration of radar beam

CURRENT TREATMENT

Currently errors are assumed uncorrelated. This is achieved by 'superobbing' and thinning.



Processed raw observations

Super-observations

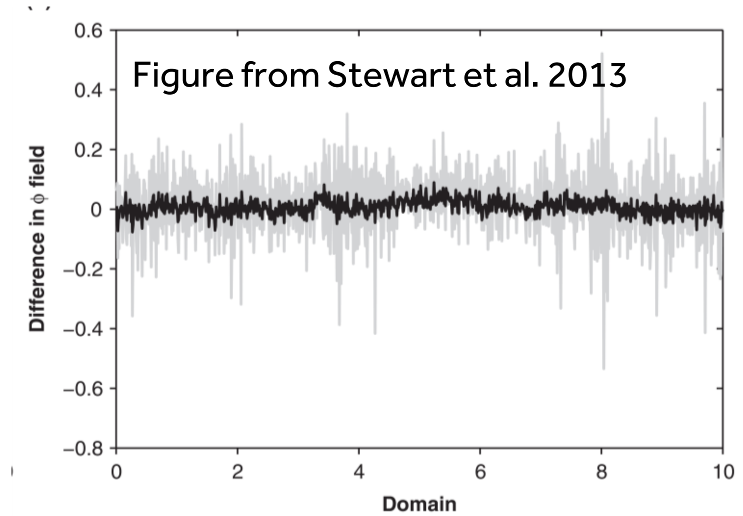
Thinned super-observations

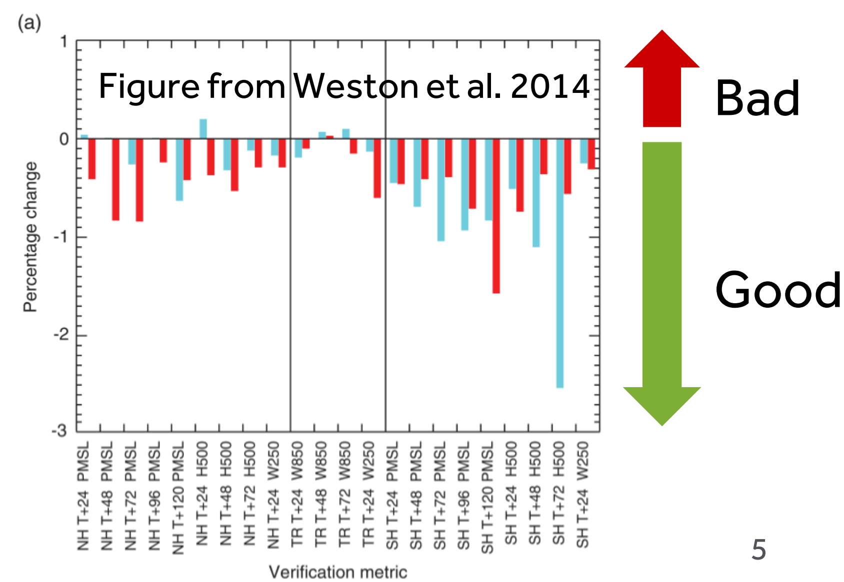
Figures from D. Simonin

WHY USE CORRELATED ERRORS?

Using correlated errors:

- Leads to an increase in the analysis accuracy (Stewart et al. 2013).
- Leads to an increase in the NWP skill score (Weston et al. 2013).
- Makes more use of the available data.





CHALLENGES

- The observation errors are unknown and may only be calculated statistically. They will also be dependent on the operational configuration.

- Once the errors are calculated we cannot use them - the computational framework to implement correlated errors is non-existent.

- The computational framework will be complex and is potentially very costly.

DIAGNOSING OBSERVATION ERROR STATISTICS

Estimate observation error statistics for SEVIRI and Doppler radar radial winds using the Desroziers Diagnostic (Desroziers et al., 2005),

Background residual: $d_b^o = y - H(x^b)$

Analysis residual: $d_a^o = y - H(x^a)$

$$\boldsymbol{R} \approx \mathrm{E}[d_a^o {d_b^o}^T]$$

- The behaviour of the diagnostic is unknown when error statistics used in the assimilation are in error.

- Has produced good estimates of ***R*** even when the ***R*** and ***B*** are not correctly specified (Mènard et al., 2009, Desroziers et al., 2009).

$$\mathrm{R}^e = \tilde{\mathrm{R}}(\mathrm{H}\tilde{\mathrm{B}}\mathrm{H}^\mathrm{T} + \tilde{\mathrm{R}})^{-1}(\mathrm{H}\mathrm{B}\mathrm{H}^T + \mathrm{R})$$

THEORETICAL ANALYSIS OF THE DIAGNOSTIC

Theoretical results provide insight into the behaviour of the diagnostic (Waller et al. 2016a):

1. Show that the diagnostic can provide good estimated when background and observation error length scales are similar.

2. Determine the behaviour of the diagnostic given information about error statistics that are used in the assimilation.

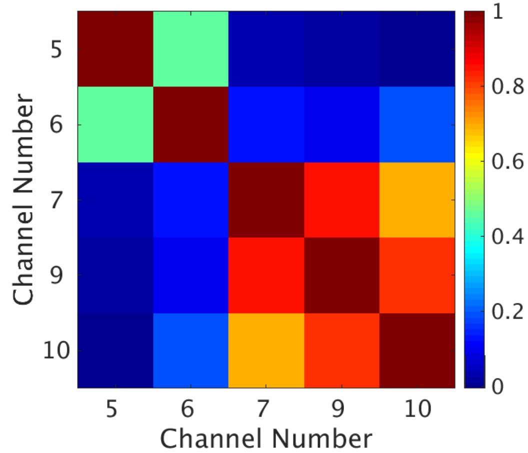
3. If correlations are neglected in the assimilation then the diagnostic will underestimate the correlation length scale.

4. If the background correlations are over estimated then the diagnostic will underestimate the correlation length scale.

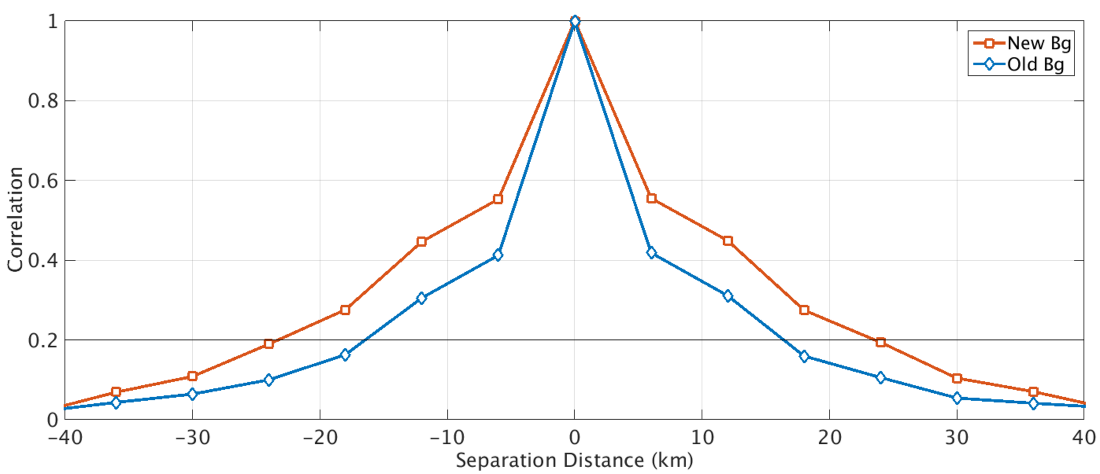
DIAGNOSING ERRORS FOR NWP OBSERVATIONS

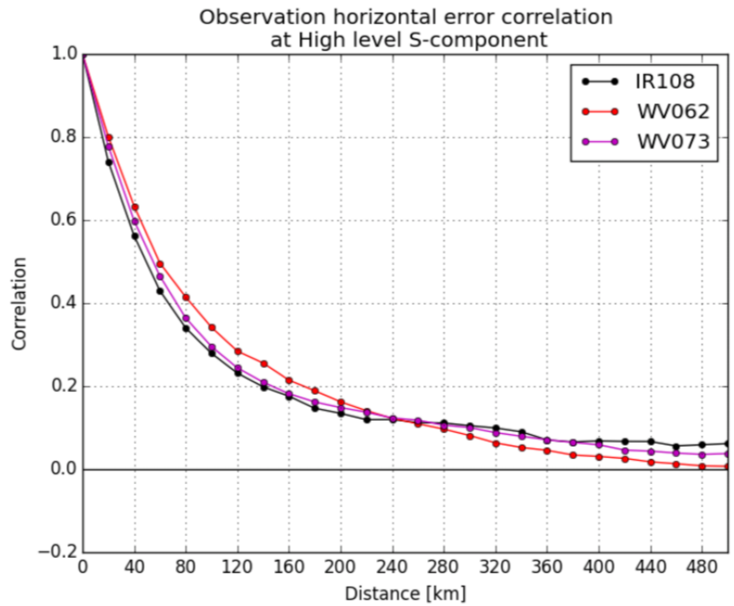
Met Office
NERC SCIENCE OF THE ENVIRONMENT
University of Reading

We have used the diagnostic to estimate errors for:



- Doppler radar radial winds (Waller et al., 2016b)
- SEVIRI observations (Waller et al. 2016c)
- Atmospheric motion vectors (Cordoba et al. 2107)



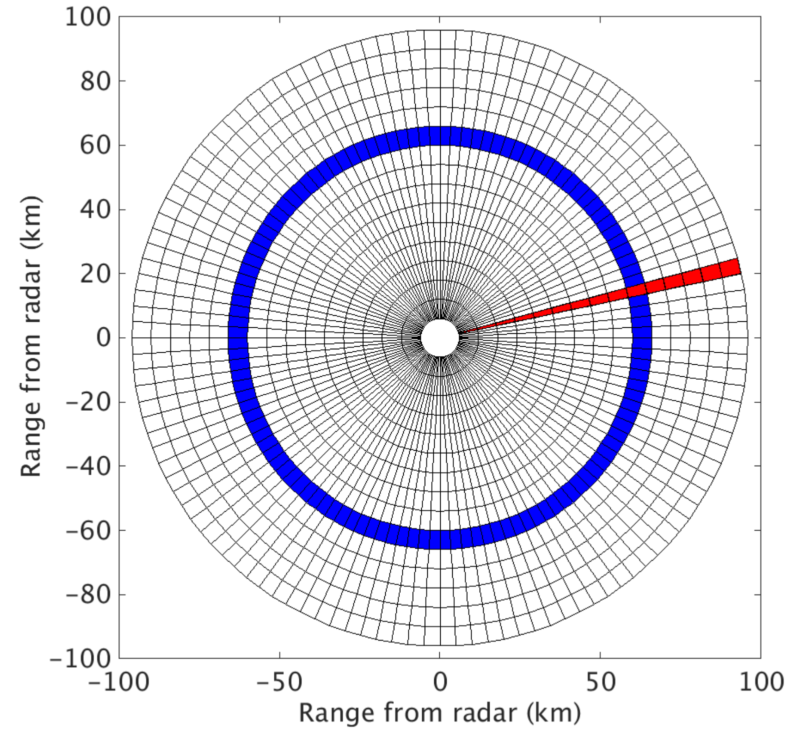


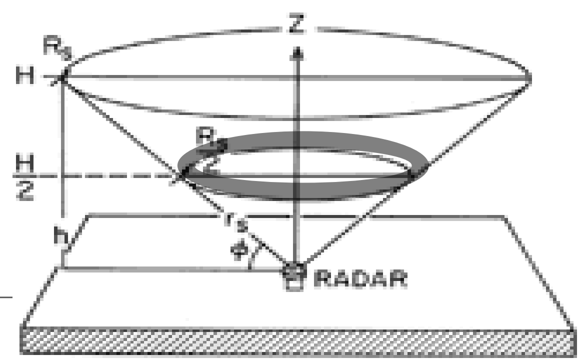
DOPPLER RADAR RADIAL WIND (DRW) ERROR STATISTICS

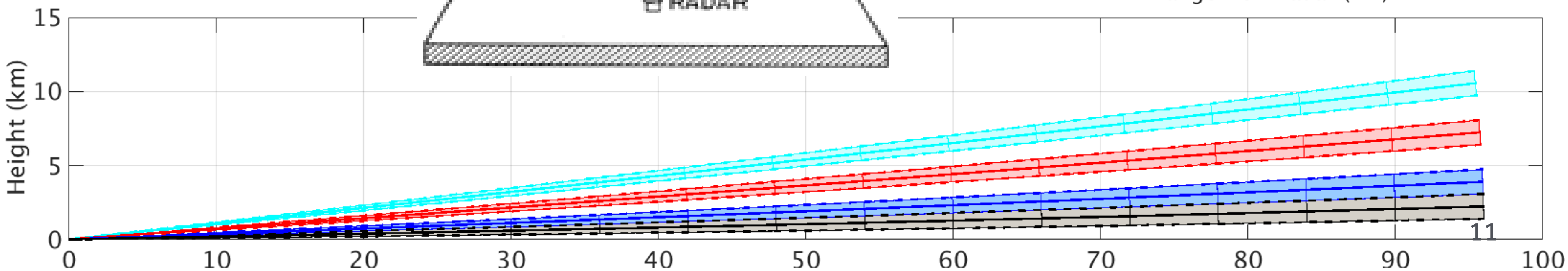
DRW OBSERVATIONS FOR THE UKV

Each radar beam produces observations of radial velocity out to a range of 100km with measurements taken:

- Every 75m along the beam.
- Every degree.
- At five different elevation angles.
- Superobbed to 3° by 3km.
- Thinned to 6km.

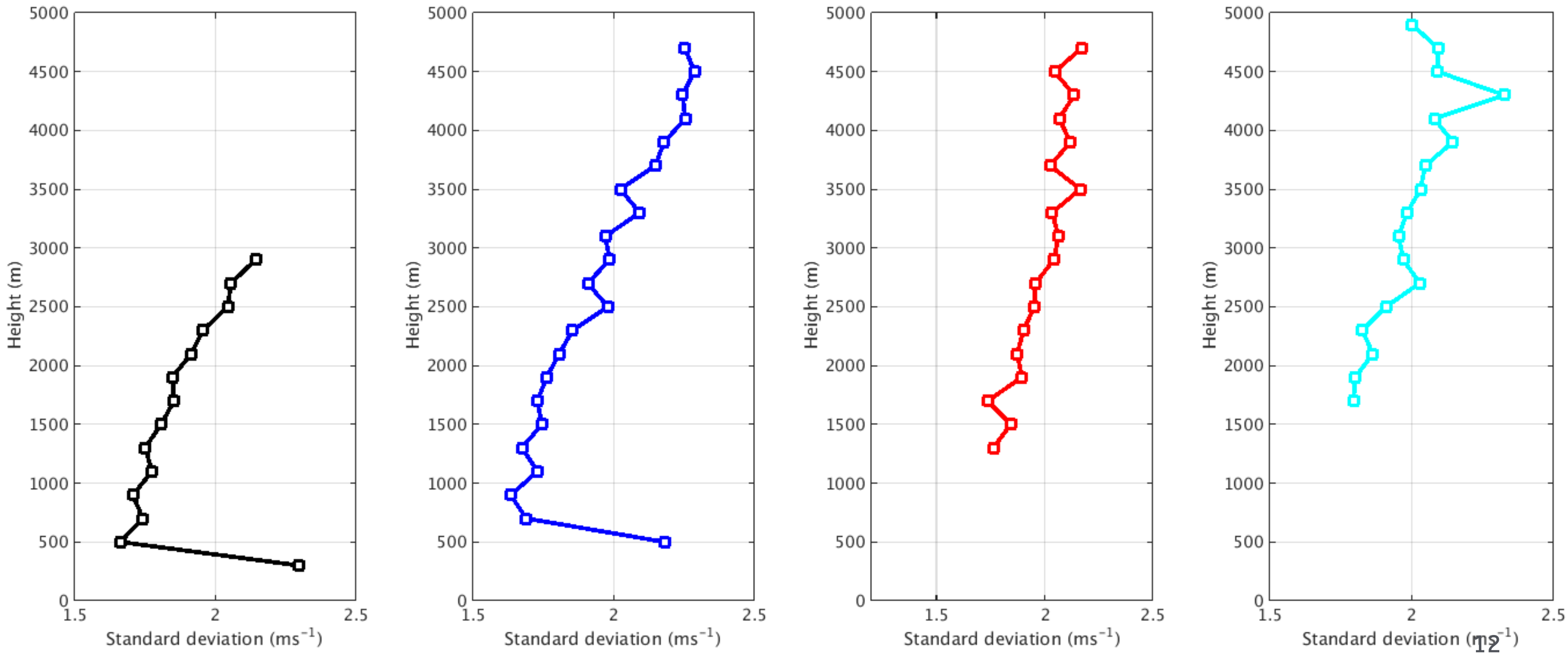




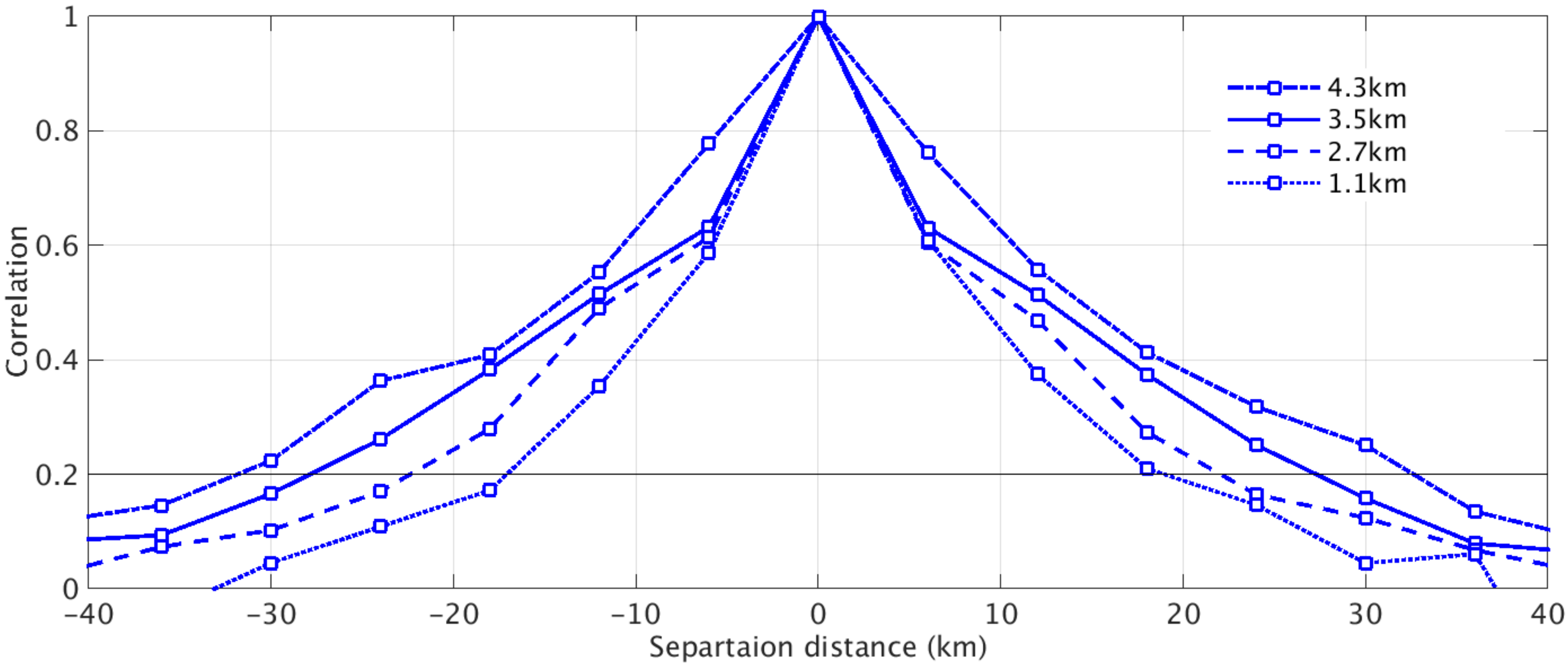


HORIZONTAL DRW ERROR VARIANCES

- Standard deviation increases with height due to the increasing measurement volume with height.
- The exception, larger errors at the lowest height, are likely to be a result of representativity errors.

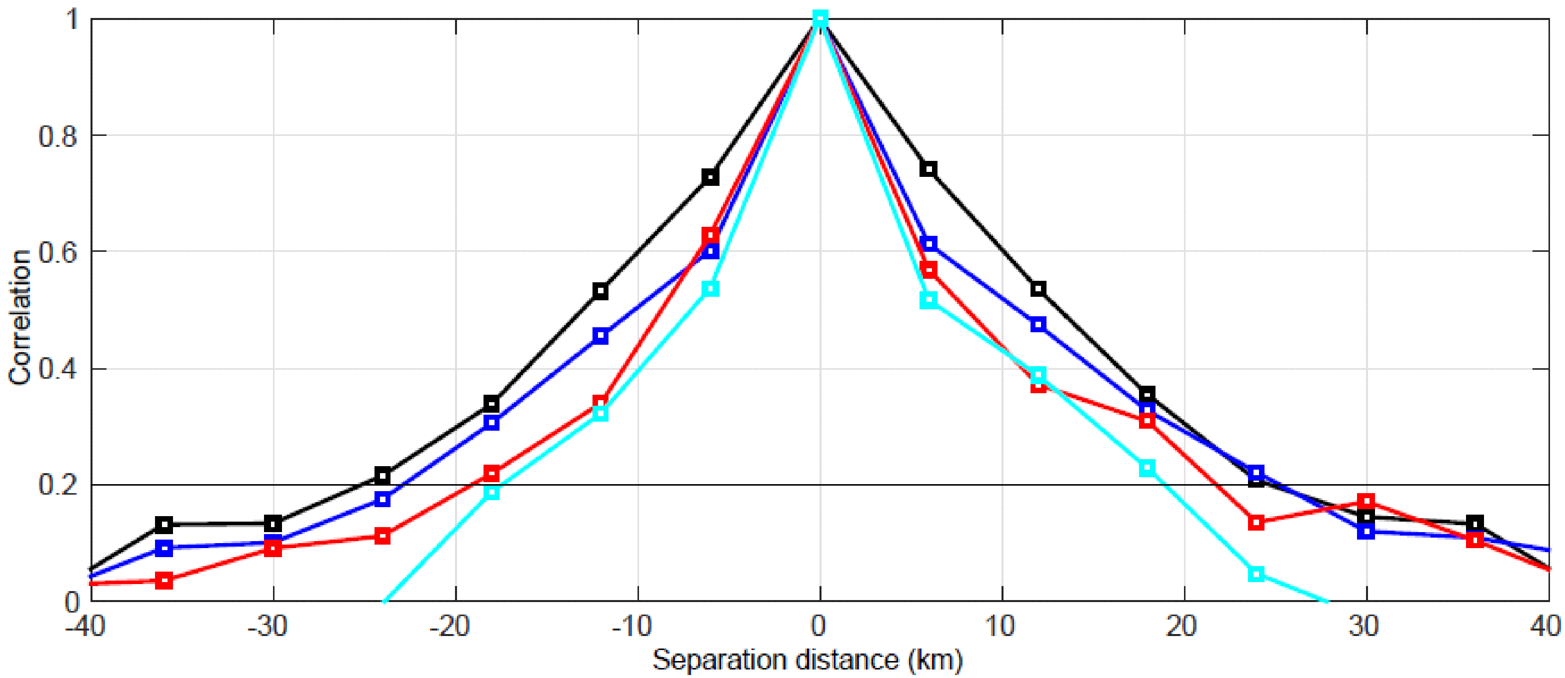


HORIZONTAL DRW ERROR CORRELATION

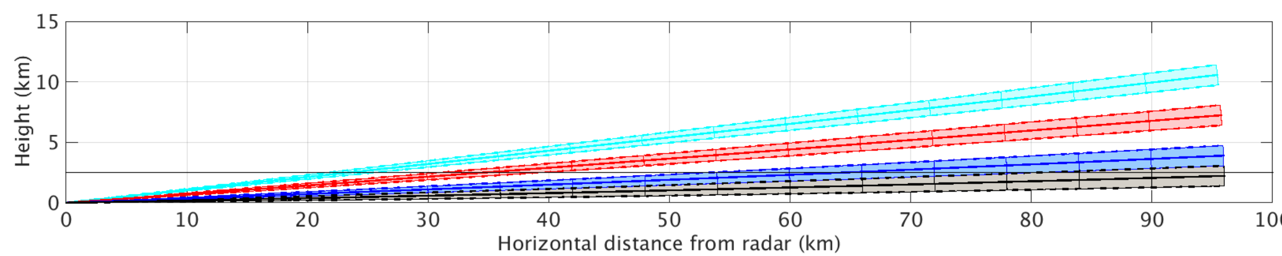


- Correlation length scale increases with height.
- Greater heights have larger errors in the observation operator.

HORIZONTAL DRW ERROR CORRELATION



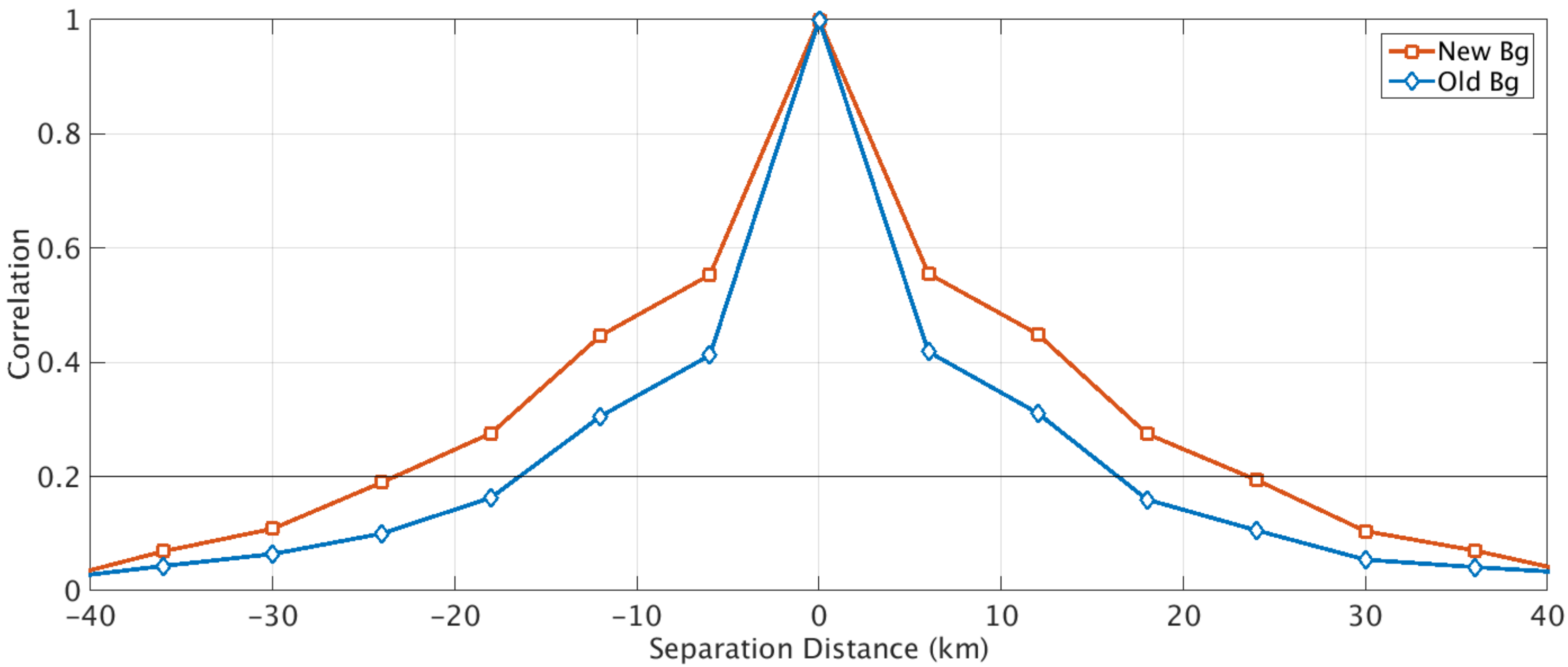
- Correlation length scale at a given height (2.5km) shorter for lower elevations.



- Lower elevations have smaller measurement volumes

SENSITIVITY TO ASSIMILATED B

Case	B statistics	Superobs	Observation operator	Standard deviation (m/s)
New Bg	New	Yes	Old	1.97
Old Bg	Old	Yes	Old	1.57

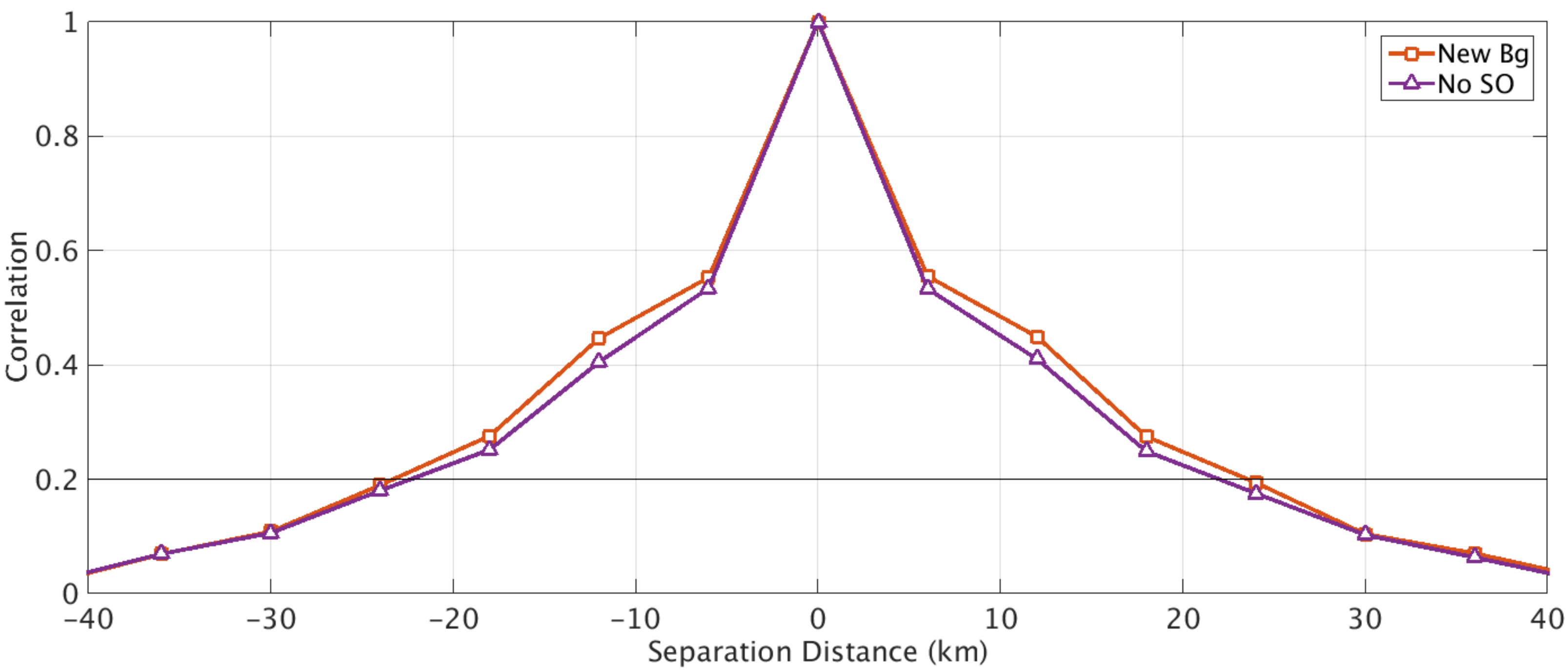


Increasing variance and length scale in **B** reduces variance and length scale in estimated **R**.

SENSITIVITY OF RESULTS TO SUPEROBSERVATIONS

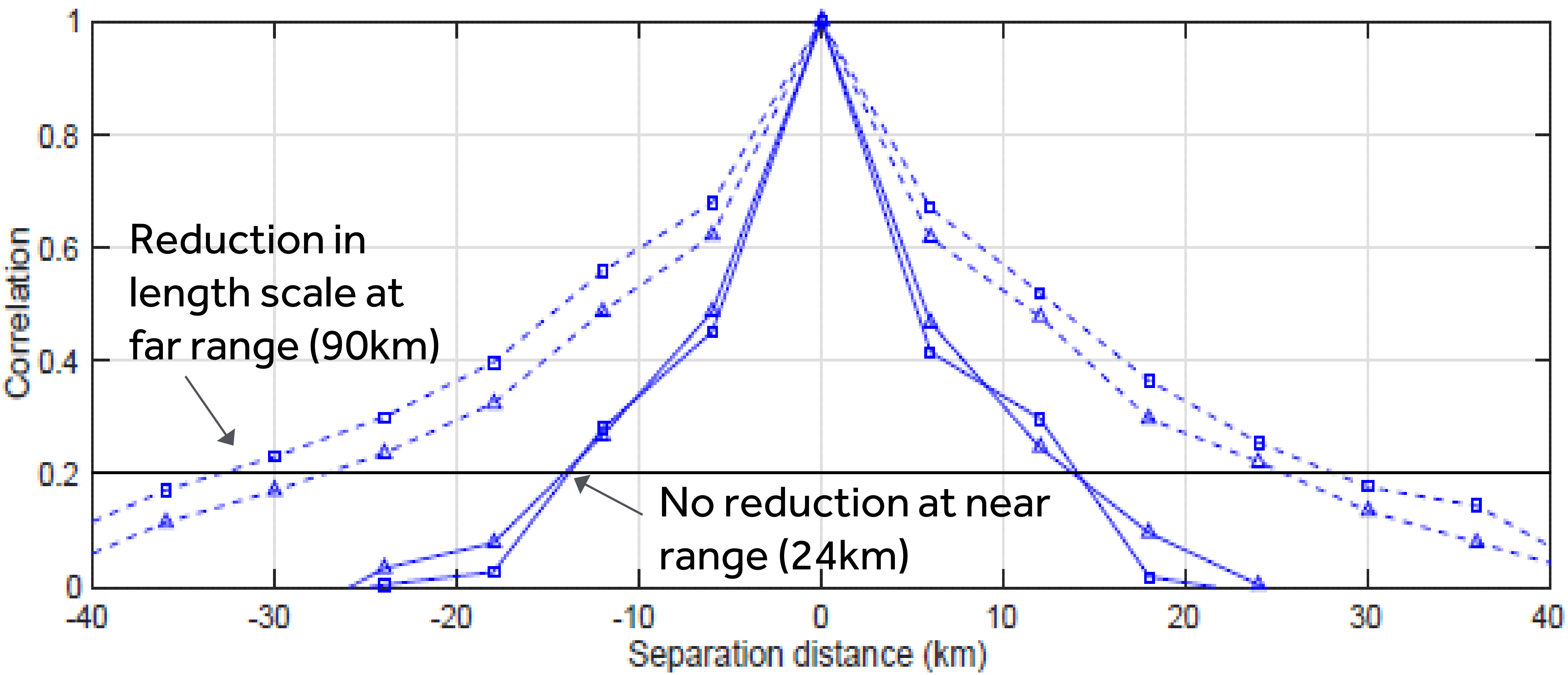
Met Office | NERC SCIENCE OF THE ENVIRONMENT | University of Reading

Case	B statistics	Superobs	Observation operator	Standard deviation (m/s)
New Bg	New	Yes	Old	1.97
No SO	New	No	Old	1.96



- Using thinned data slightly reduces correlation length scale.
- Larger reduction at far range where superobservations larger.

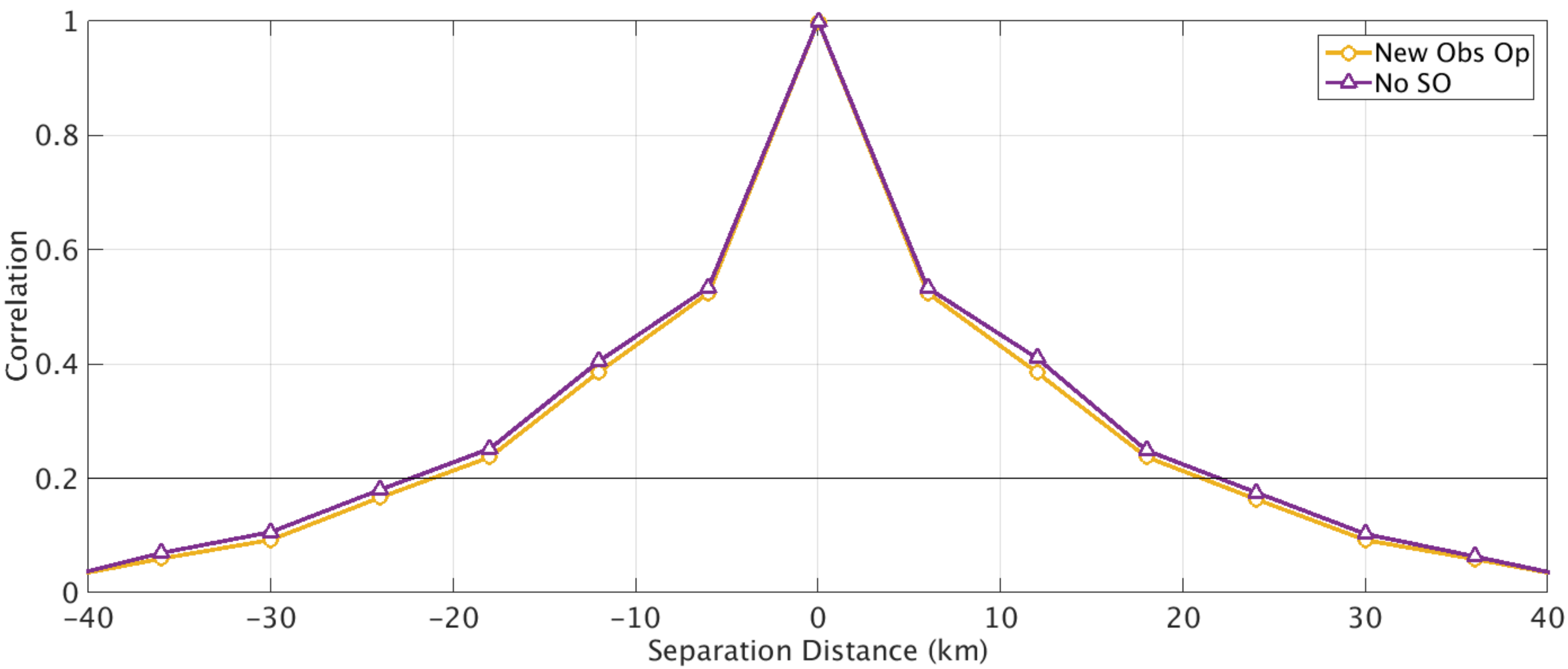
SENSITIVITY OF RESULTS TO SUPEROBSERVATIONS



N.b. control (squares) and without superobs (triangles).

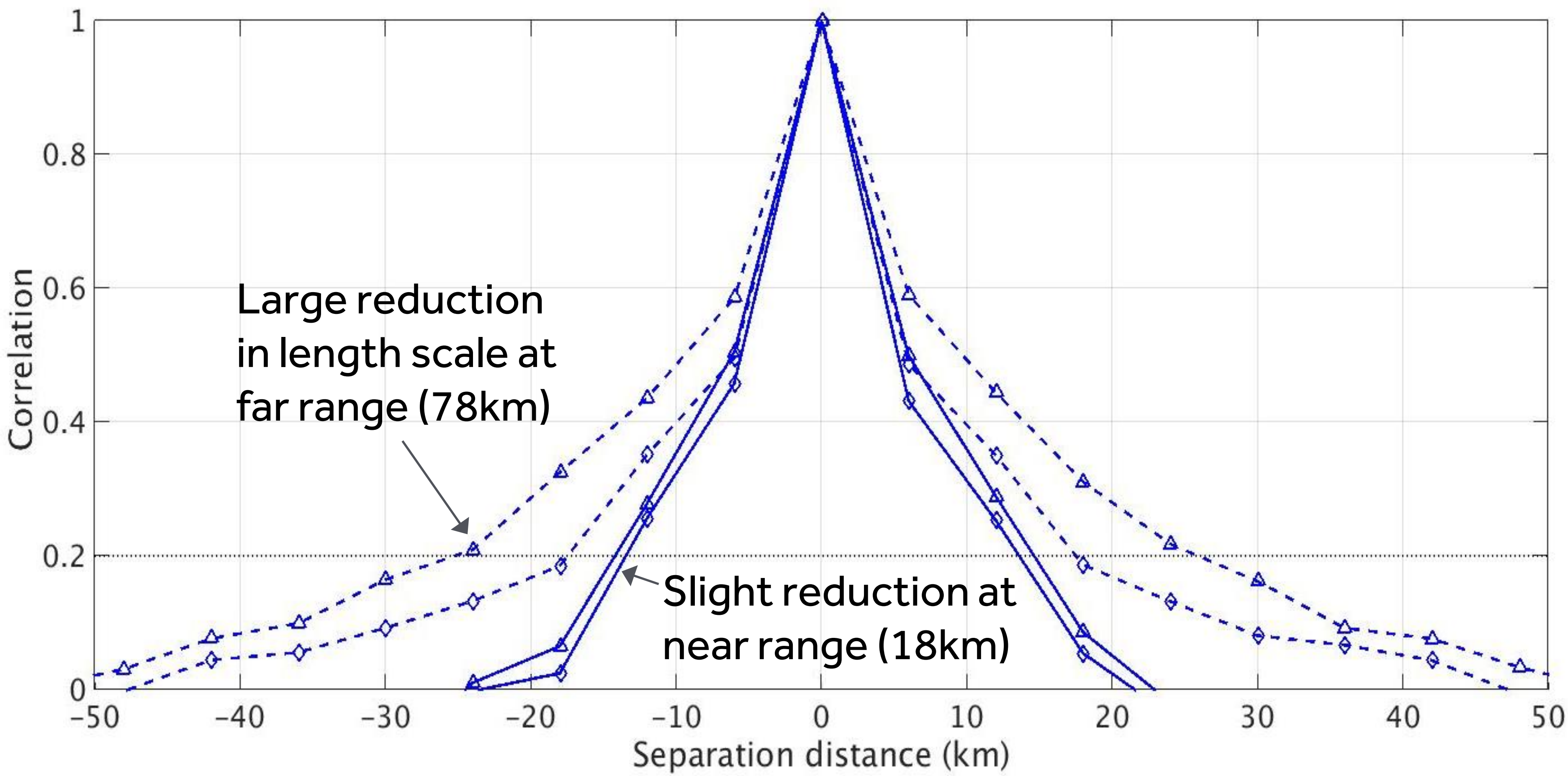
SENSITIVITY OF RESULTS TO OBSERVATION OPERATOR

Case	B statistics	Superobs	Observation operator	Standard deviation (m/s)
No SO	New	No	Old	1.96
New Obs Op	New	No	New	1.82



- Improving **H** slightly reduces variance and correlation length scale.
- Larger reduction at far range where improved **H** has most impact .

SENSITIVITY OF RESULTS TO OBSERVATION OPERATOR



N.b. operational **H** (triangles) and improved **H** (diamonds).

DRW SUMMARY

- Estimated variances for operational case are similar to the operational values.
- Variances increase with height, with the exception of the lowest levels.
- Correlation length scale increases with height and with elevation angle and is much larger than thinning distance.
- Some correlation caused by use of superobservations and simplified observation operator.

SEVIRI OBSERVATION ERROR STATISTICS

SEVIRI OBSERVATIONS

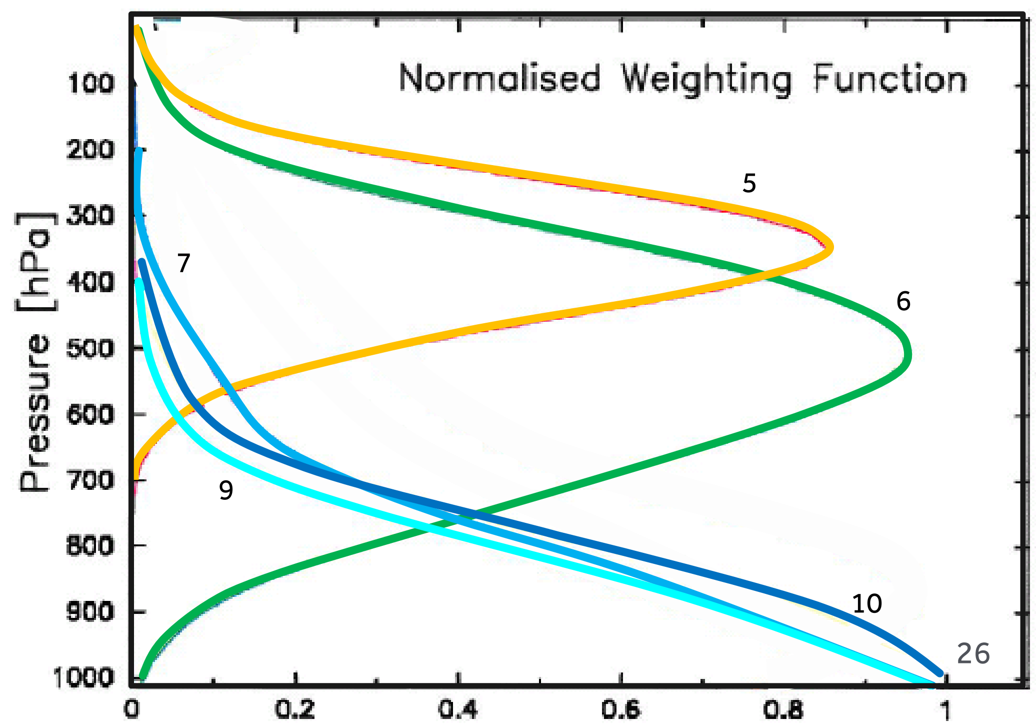
- The SEVIRI instrument on board the Meteosat Second Generation satellite produces observations of top of atmosphere radiances from 12 different spectral channels every 15 min at a 3km spatial resolution.

- Observations thinned to 24km assimilated from 5 channels.

- Simulated brightness temperatures from RTTOV model.

- Humidity channels over clear sky (channel 5 low cloud).

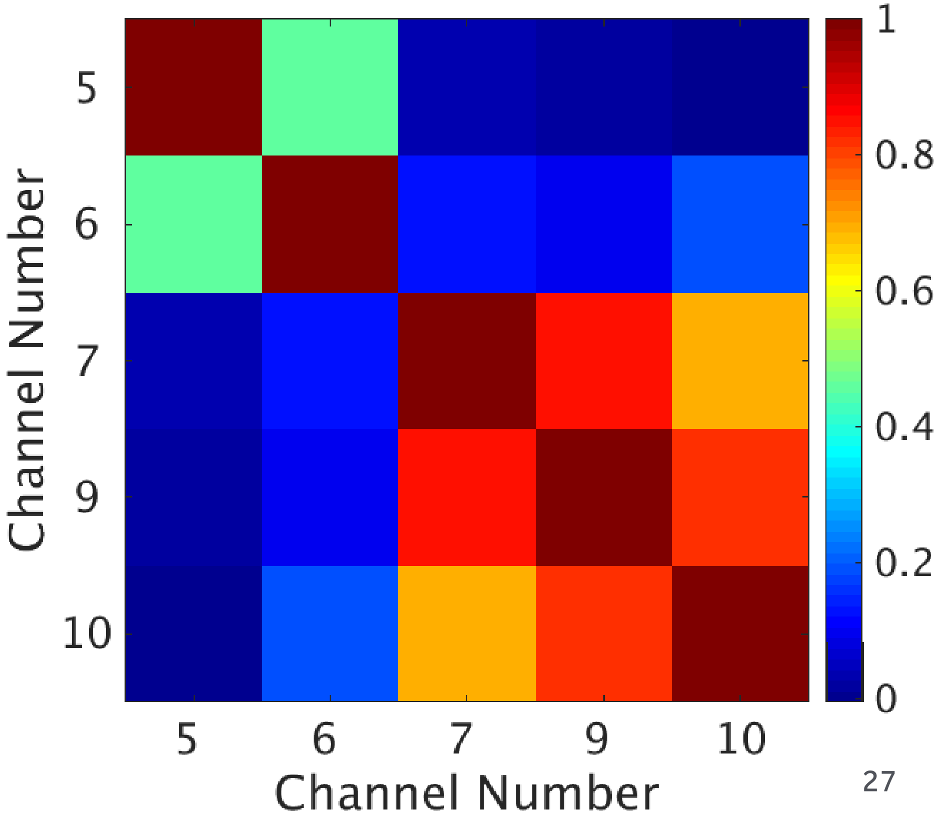
- Surface channels only over clear sky and ocean.



SEVIRI INTER-CHANNEL CORRELATIONS

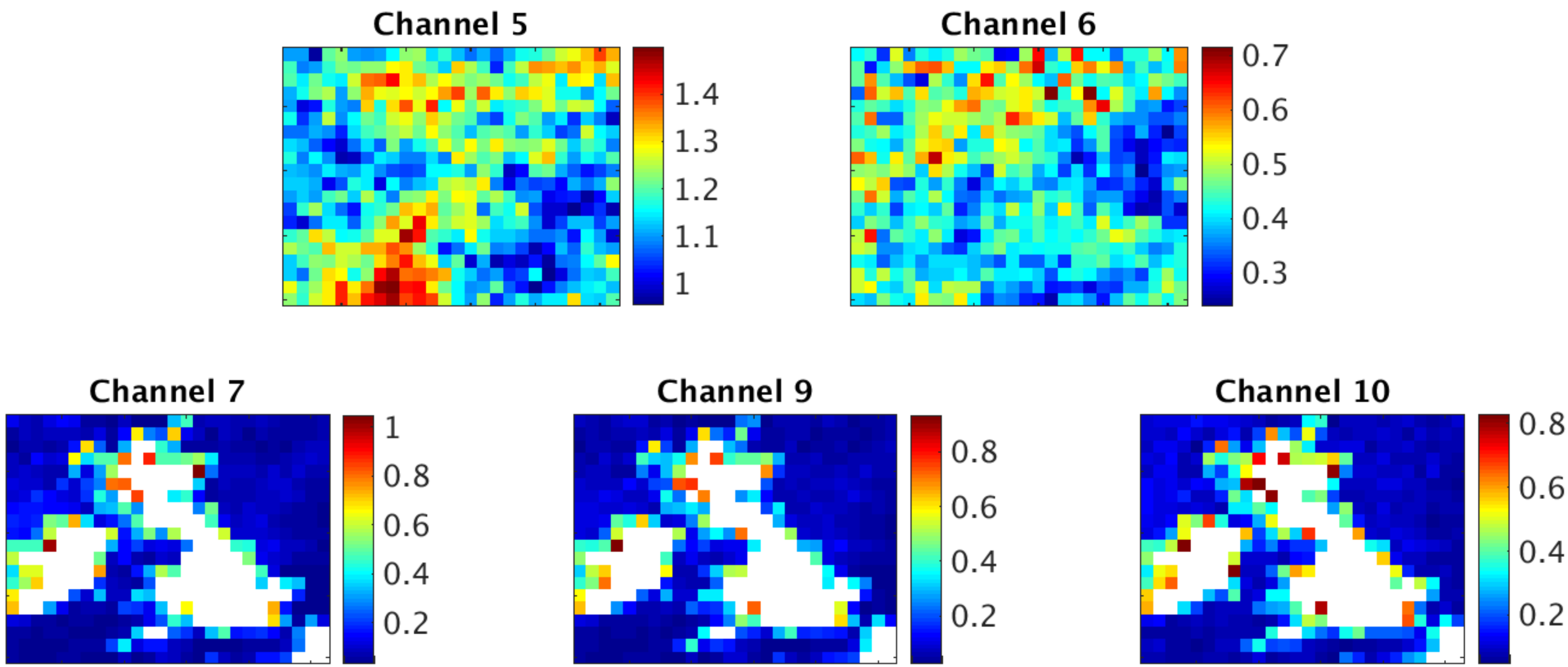
Channel	5	6	7	9	10
Operational variance	4.0	4.0	1.0	1.0	1.0
Estimated variance	1.2	0.4	0.2	0.2	0.2

- Error variances much smaller than those used operationally.
- Significant inter-channel correlation due to overlapping weighting functions.



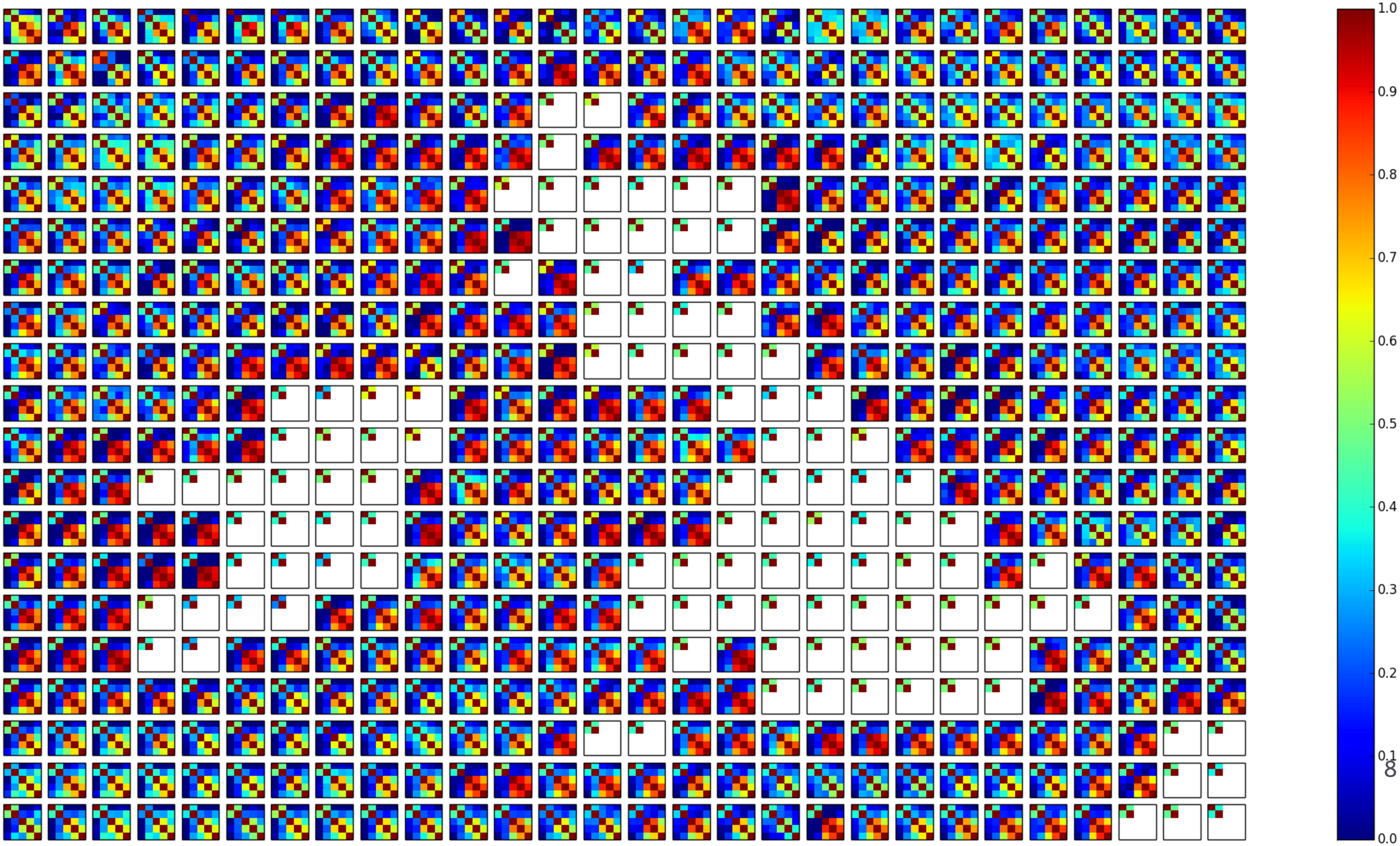
SPATIAL DEPENDENCE VARIANCE

- Variance varies across the domain particularly for surface channels.



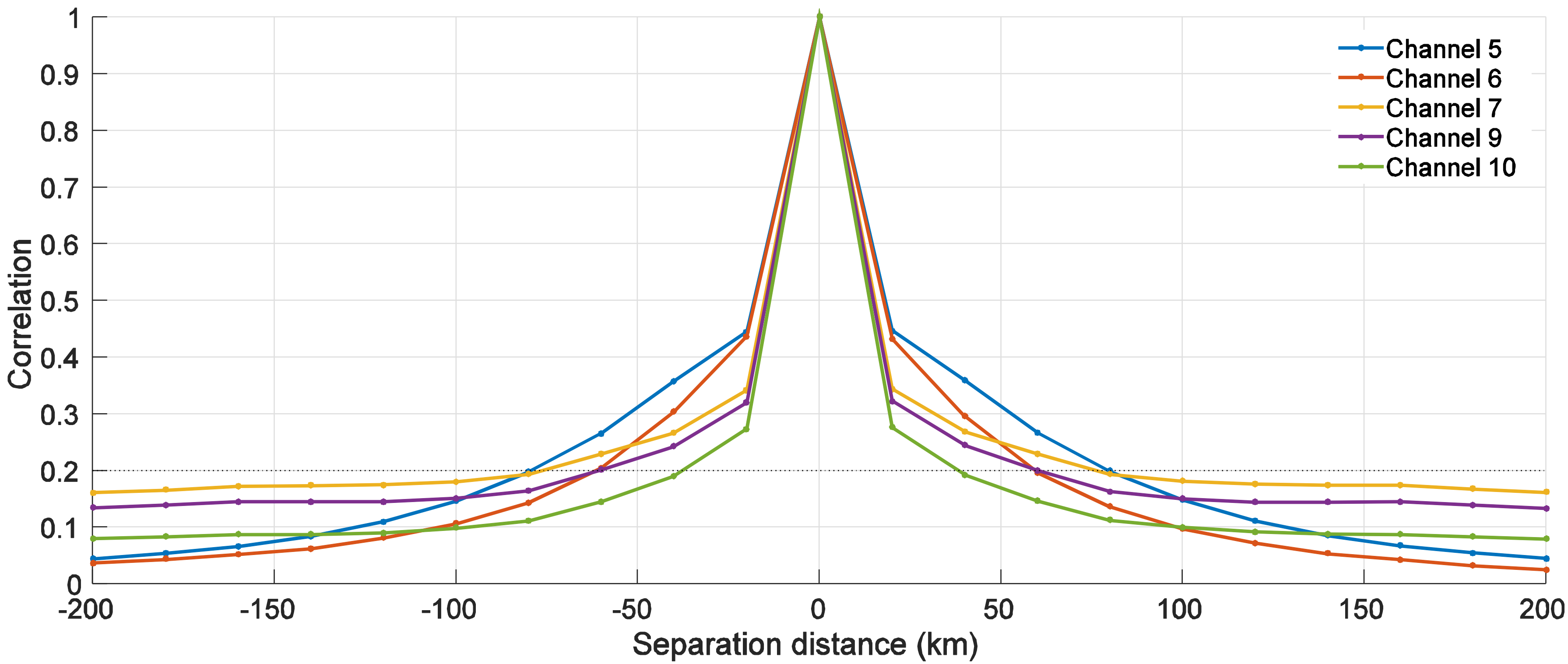
SPATIAL DEPENDENCE CORRELATION

Correlation strength dependent on surface type.



HORIZONTAL CORRELATIONS

Horizontal correlation longer than observation thinning distance.



SEVIRI SUMMARY

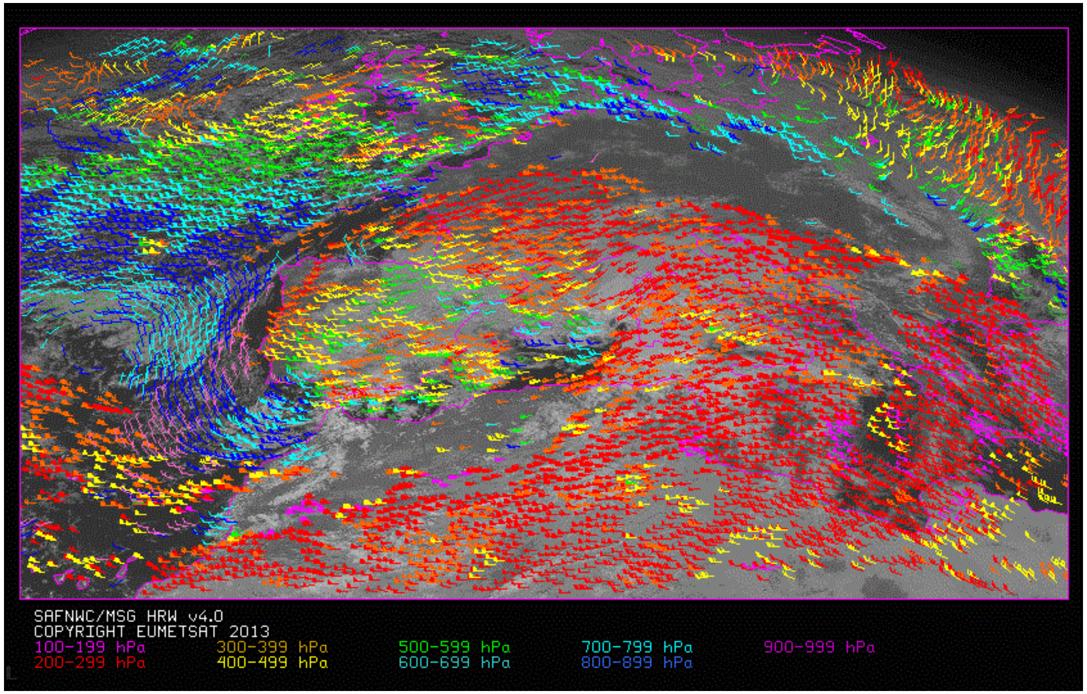
- Estimated variances are much smaller than those used operationally.
- Horizontal correlation longer than observation thinning distance.
- Inter-channel correlations are significant and vary across the domain.
- Processes diagnosed observation bias and quality control problem both of which have/are being fixed!

ATMOSPHERIC MOTION VECTOR (AMV) ERROR STATISTICS

ATMOSPHERIC MOTION VECTORS

AMVs are wind observations derived by:

- Selecting suitable features from satellite images.
- tracking the feature over consecutive images.
- Assigning a height to the tracked feature



AMV ERRORS

There are two main contributors to the total AMV error:

- tracking error
- Error in the height assignment .

The magnitude of the error is influenced by specific atmospheric situations including:

- wind shear,
- temperature inversions,
- the jet stream.

Mid-level features are hard to track as they are difficult to distinguish from features above and below.

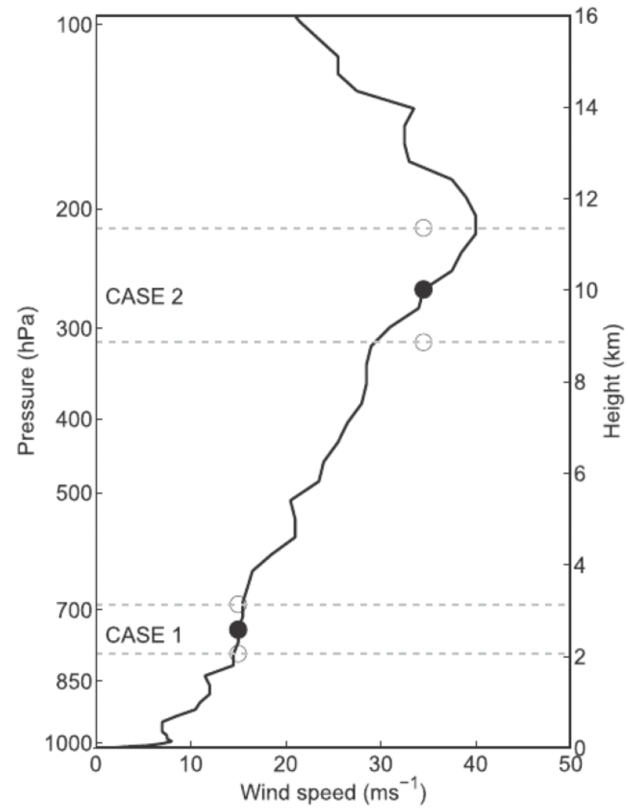


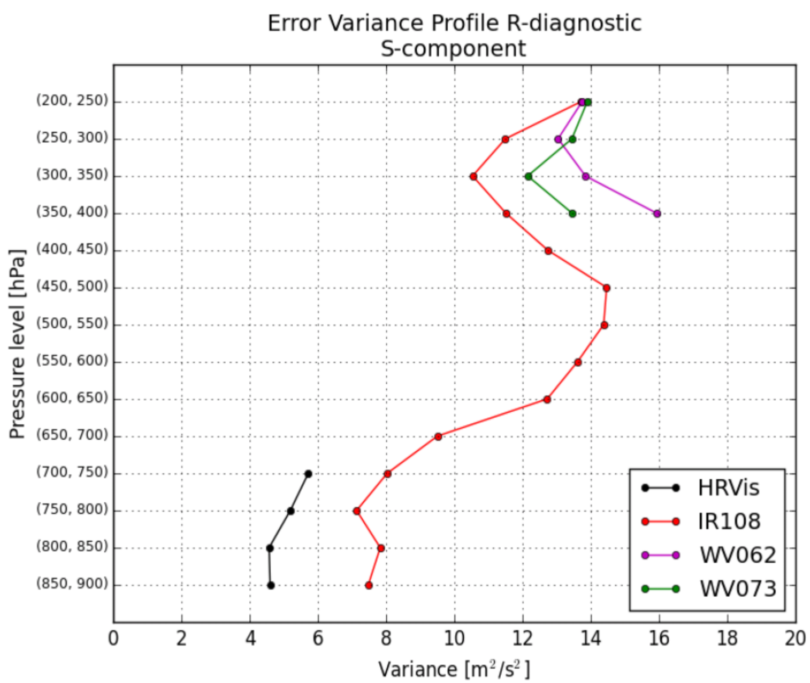
Figure from Salonen et al. 2014

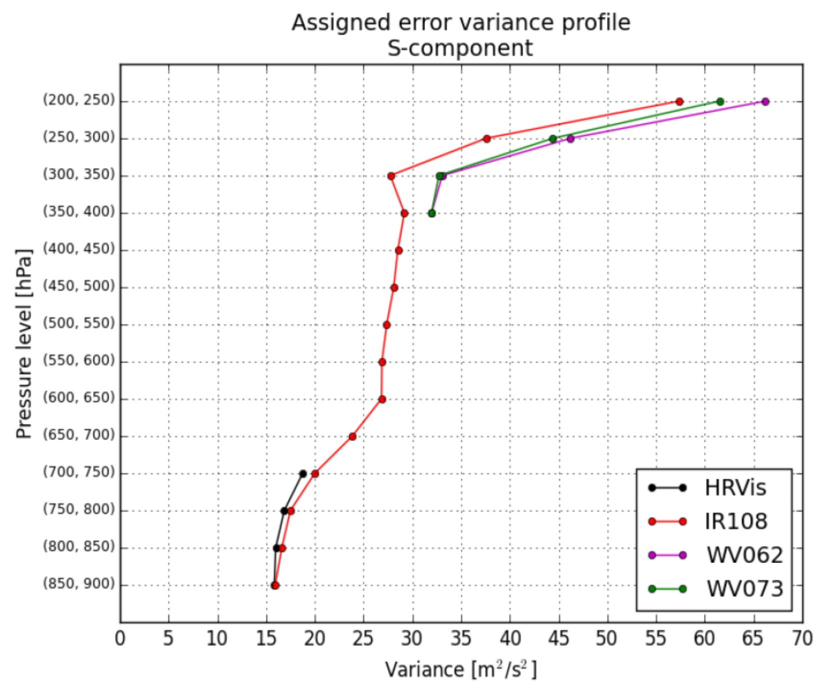
AMVS USED IN THIS STUDY

- Estimate error statistics for AMV observations assimilated in to the Met Office high resolution (1.5km) UK model.
- AMVs derived using images from 4 channels (IR108, WV062, WV073,HRVis) of the SEVIRI instrument.
- Operationally observations are thinned to 20km before being assimilated. This work assimilates a denser data set of AMVs where observations are thinned to 5km.

AMV ERROR CORRELATIONS

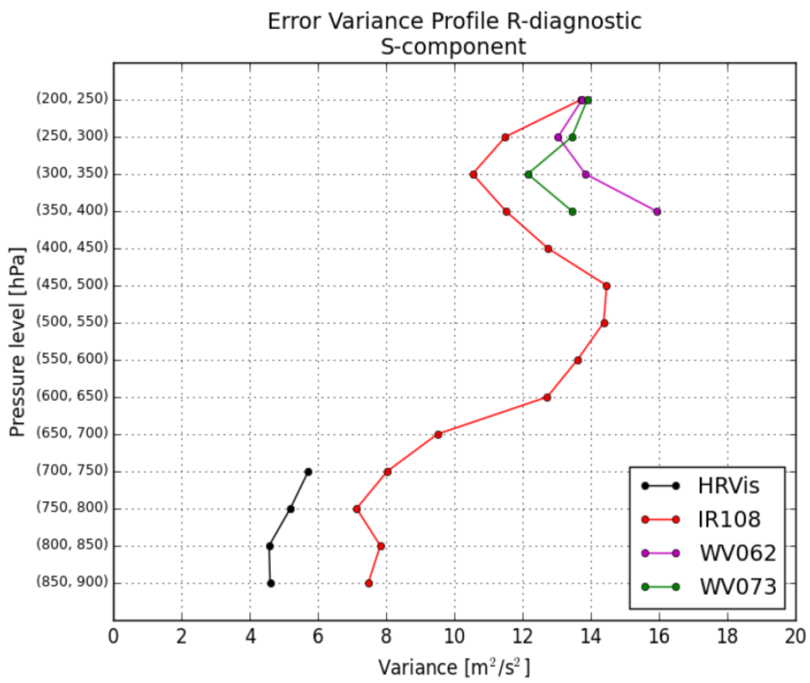
- Estimated variances
  - are smaller than those used in the assimilation
  - vary significantly with height

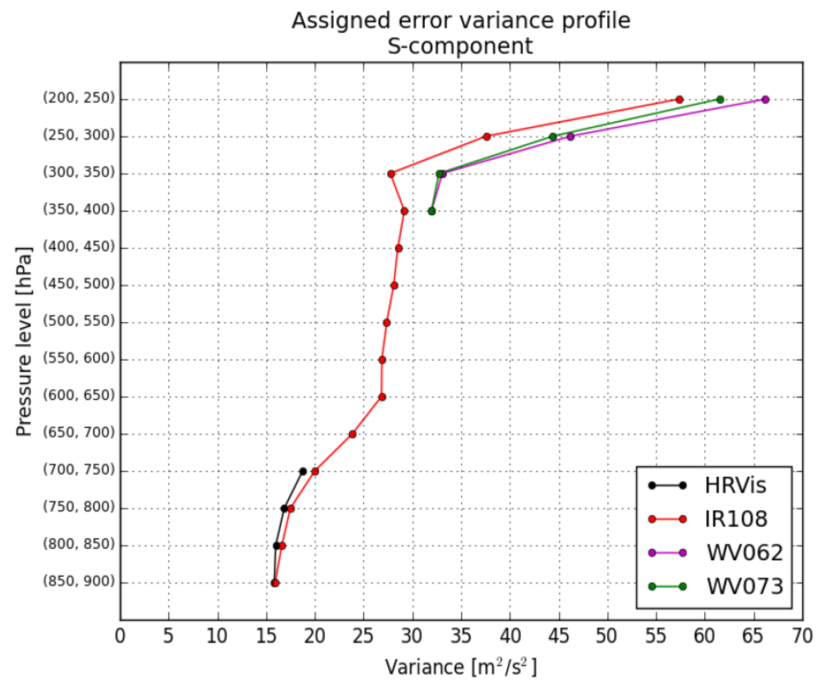




AMV ERROR CORRELATIONS

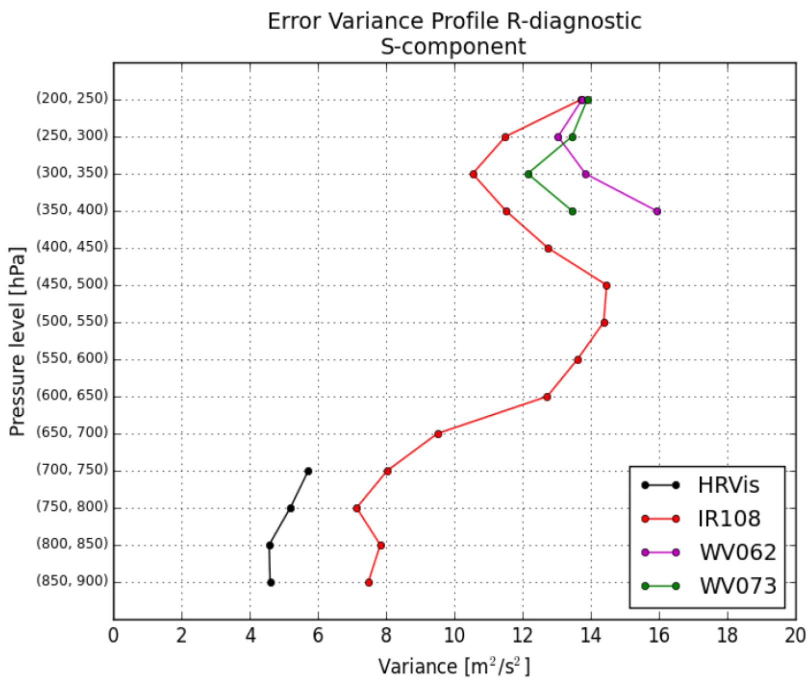
- The largest errors are found in the mid levels where
  - wind shear is large
  - tracked features are more frequently contaminated

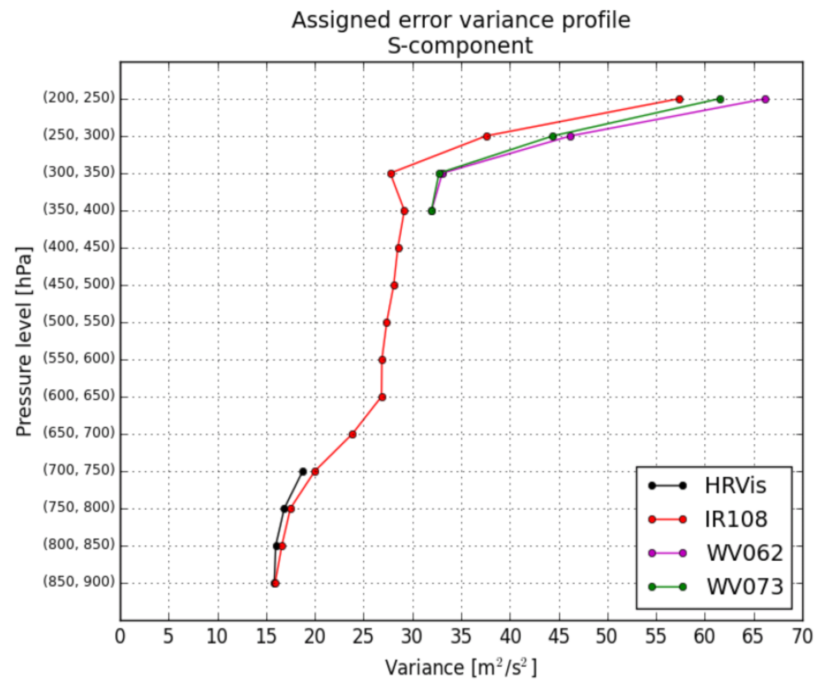




AMV ERROR CORRELATIONS

- Large error variances found at high levels may be due to:
  - very large assigned variances affecting the results of the diagnostic or
  - related to the high wind shear at these levels.





ESTIMATED AMV CORRELATION

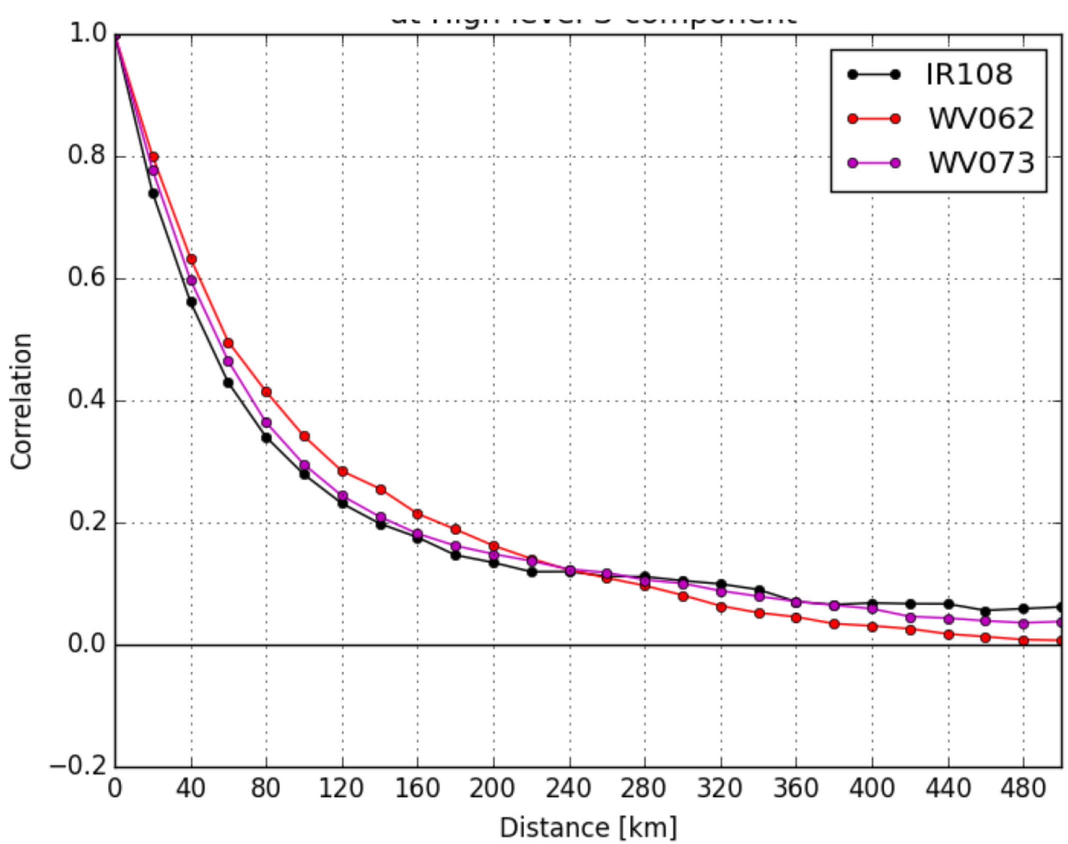
- Estimated horizontal correlation length scales range between 120km and 360km.
- Correlation length scales are significantly longer than the current thinning distance of 20km.
- Correlation length scales also vary with height, with larger length scales in the mid-levels.

	High Level			Medium Level			Low Level		
	U	V	S	U	V	S	U	V	S
IR108	120	140	140	200	200	210	140	150	140
WV062	160	200	180	-	-	-	-	-	-
WV073	150	170	160	-	-	-	-	-	-
HRVis	-	-	-	-	-	-	320	220	360

Table 1: AMV wind speed (U, V and S components) observation error horizontal correlation length scale [km] for the 4 SEVIRI channels at high, medium and low level

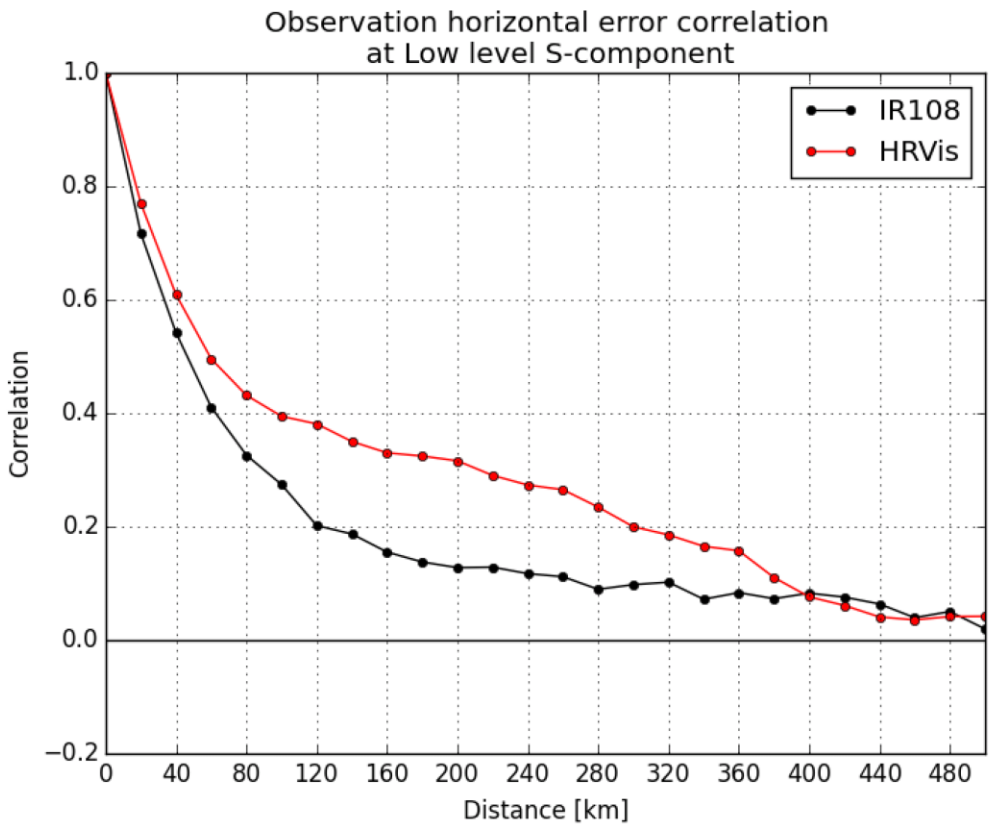
DEPENDENCE ON CHANNEL

Similar correlation length scales suggest that the error sources for each channel are similar.



DEPENDENCE ON TIME

Long correlation length scales for HRVis AMVs suggest that at least one source of error has larger length scale during daytime than at night.



AMV SUMMARY

- The estimated variances vary significantly with height and are smaller than those used in the assimilation.
- Variances are largest between 400hPa and 700hPa where wind shear is large and the tracked features are most likely to be contaminated.
- The horizontal length scales found are significantly larger than the current thinning distance of 20km.
- At least one source of error has longer correlation length scales during day than at night.

CONCLUSIONS

SUMMARY

- Accurate error statistics required for data assimilation to produce accurate analysis.
- Currently observation errors are treated as uncorrelated when, in fact, they are not.
- The errors can only be estimated statistically.
- We have proved results that describe the behaviour of a diagnostic often used to calculate observation errors.
- Using this diagnostic we have estimate observation errors for different observation types.
- Reducing correlated observation errors may be a possibility if results can provide information on the source of error.
- Some correlated error is being neglected in operational assimilation. Either observations must be thinned further or the correlated errors must be accounted for.

REFERENCES

- M. Cordoba, S. L. Dance, G. A. Kelly, N. K. Nichols and J. A. Waller. Diagnosing atmospheric motion vector observation errors for an operational high-resolution data assimilation system. Q.J.R. Meteorol. Soc., 143: 333–341. doi:10.1002/qj.2925. 2017.
- G. Desroziers, L. Berre, B. Chapnik, and P. Poli. Diagnosis of observation, background and analysis error statistics in observation space *Q.J.R. Meteorol. Soc.*, 131:3385–3396, 2005.
- G. Desroziers, L. Berre, and B. Chapnik. Objective validation of data assimilation systems: diagnosing sub-optimality. In Proceedings of ECMWF Workshop on diagnostics of data assimilation system performance, 2009.
- R. Mènard, Y. Yang, and Y. Rochon. Convergence and stability of estimated error variances derived from assimilation residuals in observation space. In Proceedings of ECMWF Workshop on diagnostics of data assimilation system performance, 2009.
- Salonen K, Cotton J, Bormann N and Forsythe M. 2015. Characterizing Atmospheric Motion Vectors height assignment error by comparing best-fit pressure statistics from the Met Office and ECMWF data assimilation systems. Journal of Applied Meteorology and Climatology. 54:225-242.
- L. M. Stewart, S. L. Dance, N. K. Nichols, J. R. Eyre, and J. Cameron. Estimating inter-channel observation error correlations for IASI radiance data in the Met Office system. *Q.J.R. Meteorol. Soc.*, 2013.
- J.A. Waller, S.L. Dance, N.K. Nichols Theoretical insight into diagnosing observation error correlations using observation-minus-background and observation-minus-analysis statistics. *Q.J.R. Meteorol. Soc.* 142: 418–431. doi:10.1002/qj.2661 2016a.
- J.A. Waller, D. Simonin S.L. Dance, N.K. Nichols and S.P. Ballard. Diagnosing observation error correlations for Doppler radar radial winds in the Met Office UKV model using observation-minus-background and observation-minus-analysis statistics. MWR. DOI: http://dx.doi.org/10.1175/MWR-D-15-0340.1. 2016b.
- J. A. Waller, S. P. Ballard, S. L. Dance, G. A. Kelly, N. K. Nichols, D. Simonin. Diagnosing Horizontal and Inter-Channel Observation Error Correlations for SEVIRI Observations Using Observation-Minus-Background and Observation-Minus-Analysis Statistics. *Remote Sens* , *8*, 581. 2016c.
- P. P. Weston, W. Bell, and J. R. Eyre. Accounting for correlated error in the assimilation of high resolution sounder data. *Q.J.R. Meteorol. Soc.* , 2013. DOI: 10.1002/qj.2306.

DIAGNOSTIC DERIVATION

$$\begin{aligned}\mathbf{d}^b = \mathbf{y} - \mathcal{H}(\mathbf{x}^b) &= \mathbf{y} - \mathcal{H}(\mathbf{x}^t) + \mathcal{H}(\mathbf{x}^t) - \mathcal{H}(\mathbf{x}^b), \\ &\approx \epsilon^o + \mathbf{H}(\mathbf{x}^t - \mathbf{x}^b), \\ &\approx \epsilon^o + \mathbf{H}\epsilon^b,\end{aligned}$$

DIAGNOSTIC DERIVATION

$$
\begin{aligned}
\mathbf{d}^a &= \mathbf{y} - \mathcal{H}\mathbf{x}^a, \\
&= \mathbf{y} - \mathcal{H}(\mathbf{x}^b + \mathbf{K}(\mathbf{y} - \mathcal{H}\mathbf{x}^b)), \\
&= \mathbf{y} - \mathcal{H}(\mathbf{x}^b + \mathbf{K}\mathbf{d}^b), \\
&\approx \mathbf{d}^b - \mathbf{H}\mathbf{K}\mathbf{d}^b, \\
&= (\mathbf{I} - \mathbf{H}\mathbf{K})\mathbf{d}^b, \\
&= \mathbf{R}(\mathbf{H}\mathbf{B}\mathbf{H}^{\mathrm{T}} + \mathbf{R})^{-1}\mathbf{d}^b.
\end{aligned}
$$

DIAGNOSTIC DERIVATION

$$
\begin{aligned}
E[\mathbf{d}^a \mathbf{d}^{b^T}] &= \mathbf{R}(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} E[\mathbf{d}^b \mathbf{d}^{b^T}], \\
&= \mathbf{R}(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-1} (\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R}), \\
&= \mathbf{R}.
\end{aligned}
$$