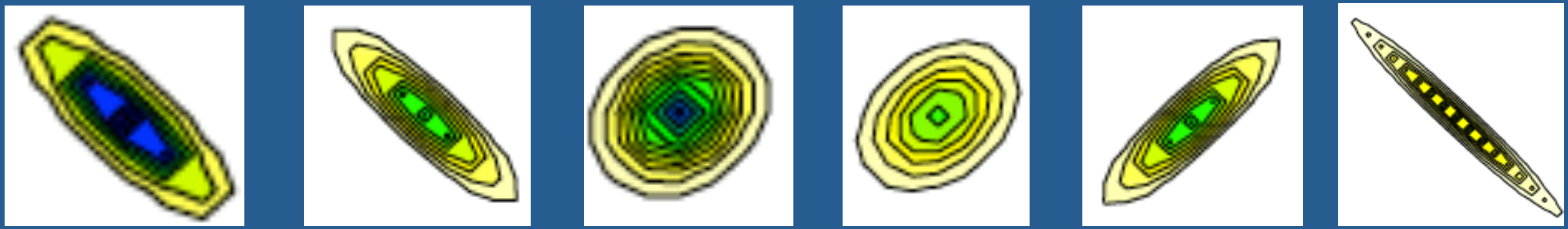


On the interaction of observation and prior error correlations in data assimilation

Alison Fowler, Sarah Dance, Jo Waller

Data Assimilation Research Centre, University of Reading, UK.

Introduction

- Until recently observation error correlations (OECs) have been neglected in NWP.
  - Difficult to estimate
  - Complicate DA algorithms
- However, OECS are known to be non-negligible for many observation types- arise due to representivity errors, observation operator errors, pre-processing…
- However, progress is being made in estimating OECs (for example using observation minus model statistics) and accounting for them in the assimilation (e.g. Campbell et al. [2017], Bormann et al. [2016], Weston et al. 2014).
- This paves the way for the assimilation of denser observations
  - could be crucial for high impact small scale weather

Introduction

- Question:
  – How does accounting explicitly for OECs impact the assimilation of the observations?
  – What are the implications for observation network design?

Disclaimer:

– Assuming observation and prior uncertainties are Gaussian, known and accounted for perfectly.
– Only considering spatial error correlations.

Examples of spatial observation error correlations

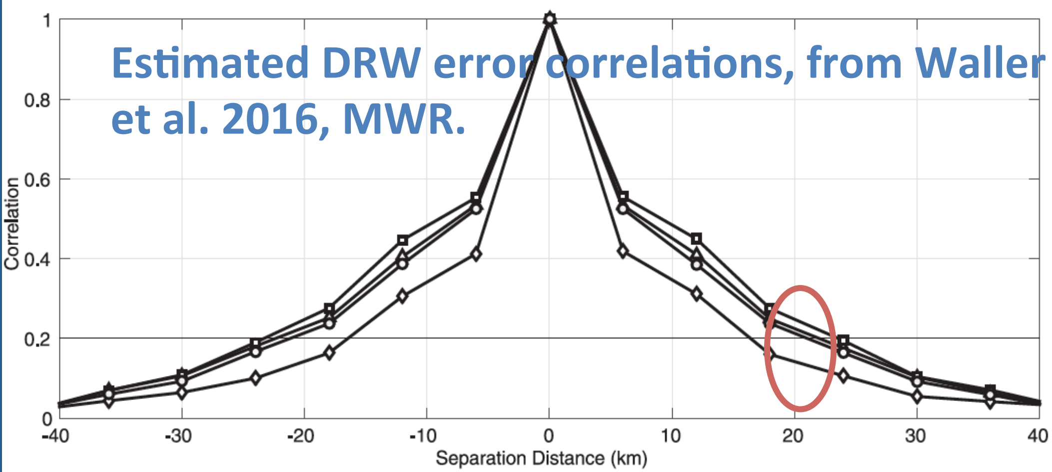


FIG. 3. All elevation horizontal observation error correlations for case 1 (control; squares), case 2 (alternate background error statistics; diamonds), case 3 (thinned raw data; triangles), and case 4 (new observation operator; circles). Error correlations are deemed to be insignificant below the horizontal line at 0.2.

In the UKV background error correlation length-scales for winds are approximately 100km. DRW are currently thinned to a distance of 6km and AMVs to 20km.

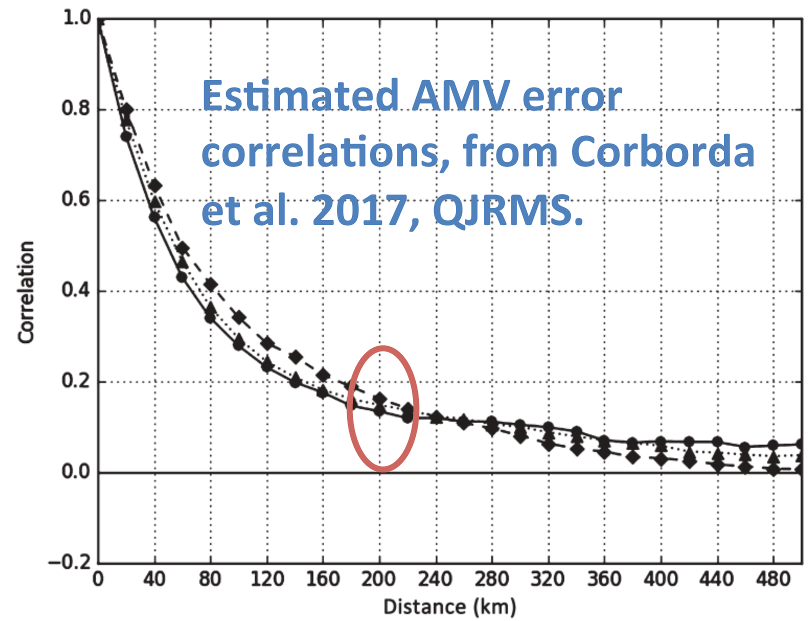
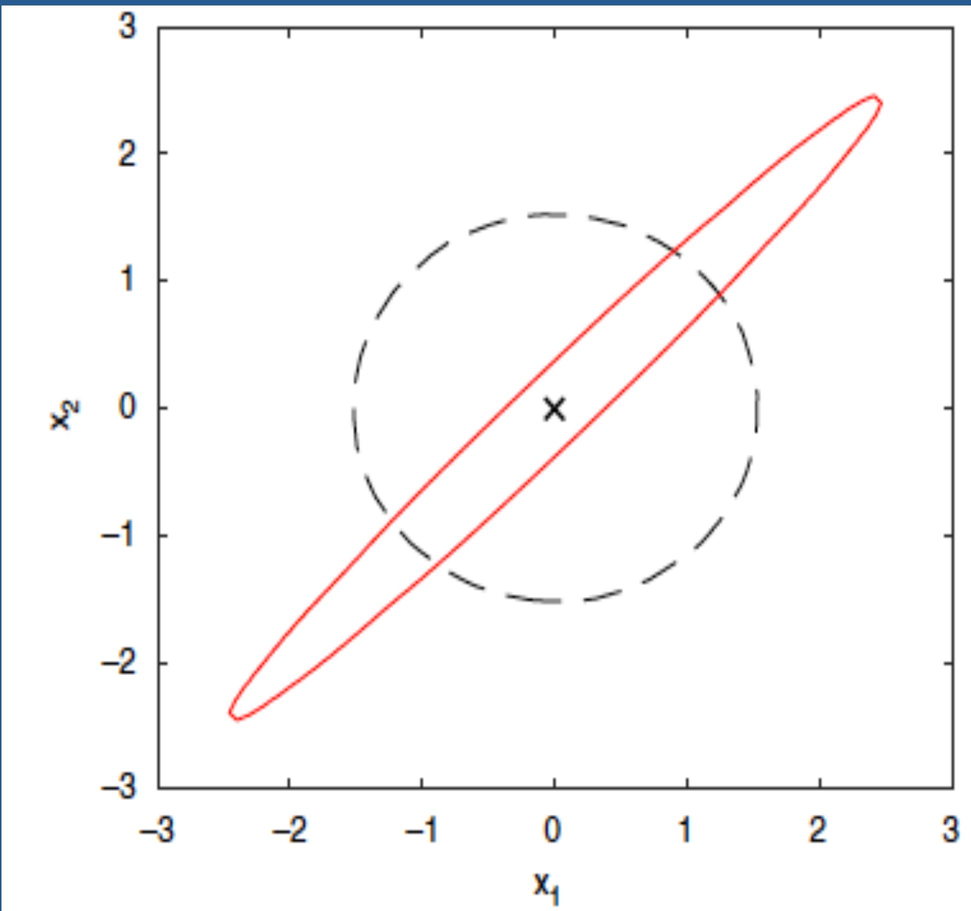


Figure 5. AMV wind-speed observation horizontal error correlation function estimate (20 km bin size) for the three SEVIRI channels, IR108 (solid with circles), WV062 (dashed with diamonds) and WV073 (dotted with triangles), at high level.

OECs and information content



<- the region of 95% probability of a Gaussian PDF with covariance matrix given by

I. $\mathbf{I}$ (black dashed line)
   - entropy = 2.8379

II. $[(1, 0.99)^T, (0.99, 1)^T]$ (solid line).
   - entropy = 0.8794.

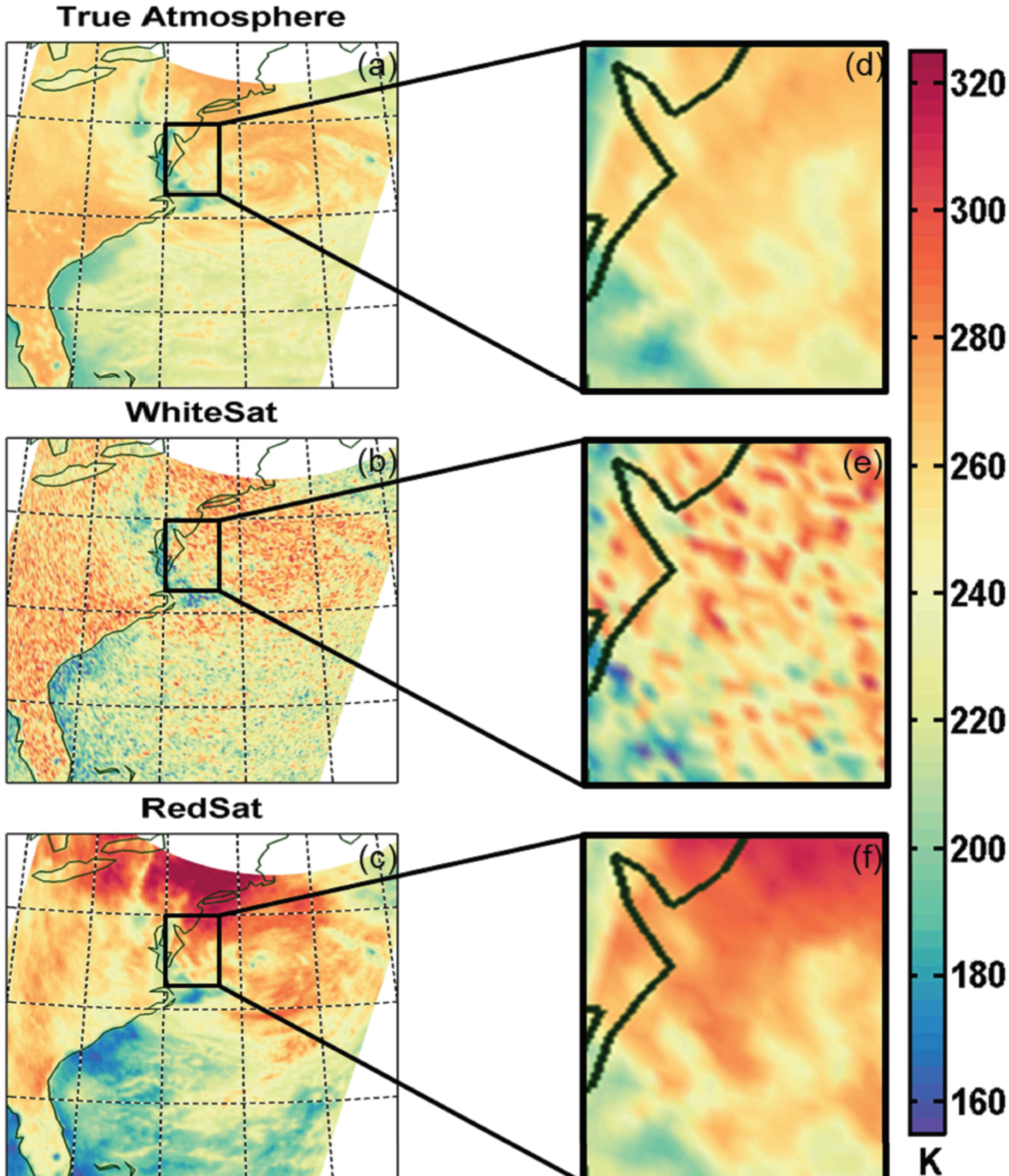


Figure 1. (a) A microwave satellite image of Hurricane *Sandy* on 24 October 2012, which is treated as truth. (b) Panel (a) plus white (uncorrelated) noise; (c) panel (a) plus red (spatially correlated) noise. (d)–(f) Detail views of (a)–(c), respectively. The colour scale for all panels is brightness temperature in Kelvin. This figure is for illustrative purposes.

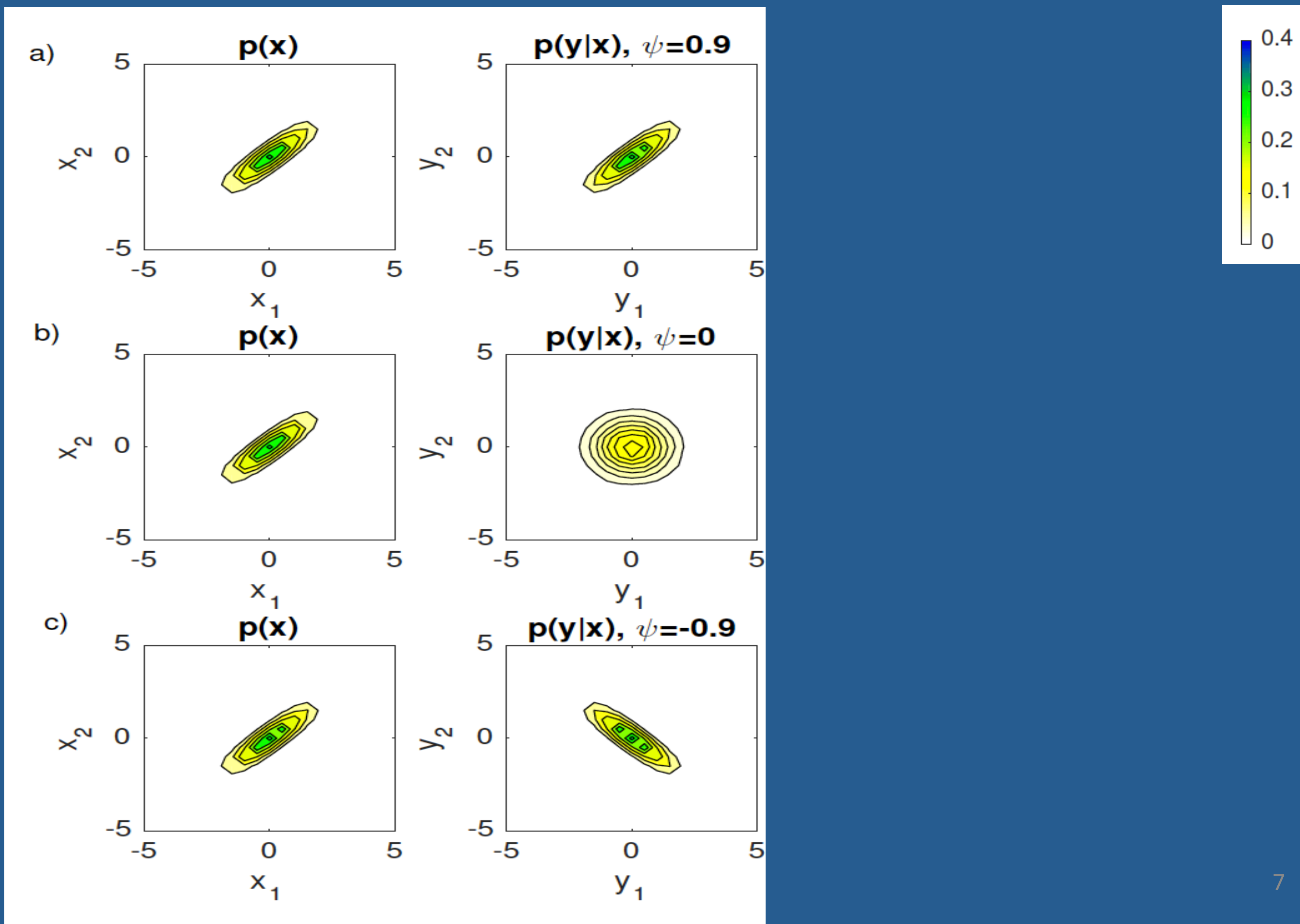
Illustration of the effect of positive spatially correlated observation errors, from Rainwater et al. 2015, QJRMS.

$p(\mathbf{x}|\mathbf{y}) \propto p(\mathbf{x})p(\mathbf{y}|\mathbf{x})$

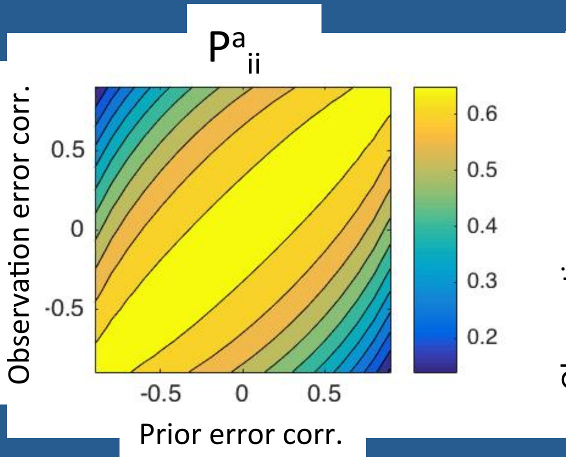
Direct observations

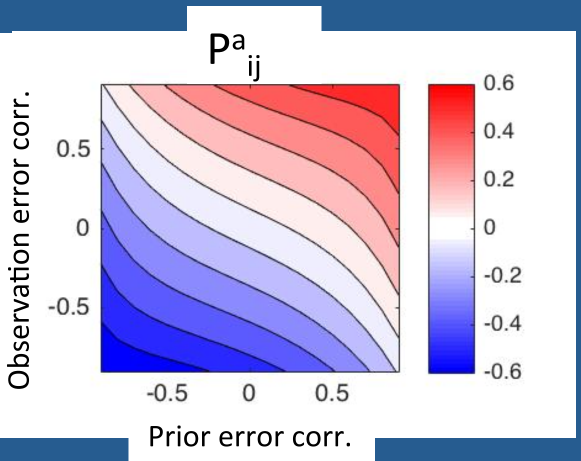
Fowler et al. 2018, QJRMS

Bayes' illustration



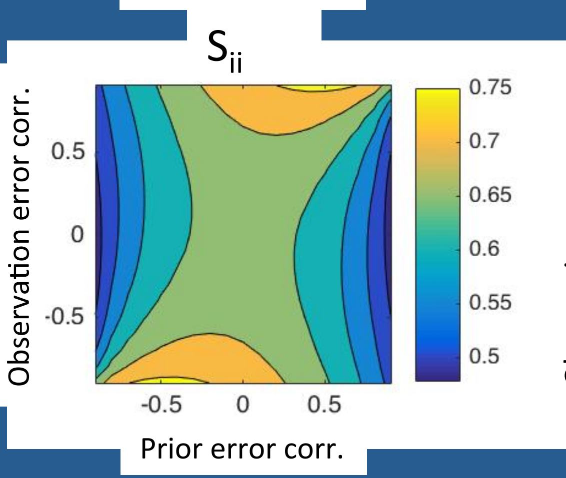
2-variable illustrations continued…

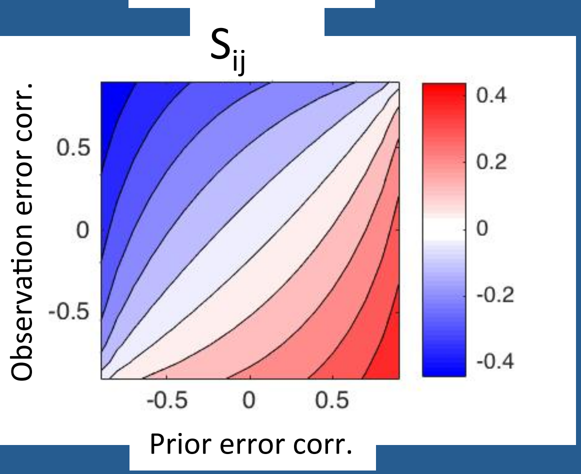




Analysis more accurate at small scales

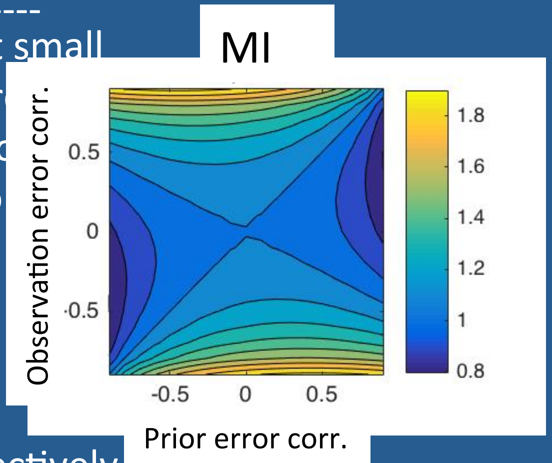
Analysis more accurate at large scales





Analysis at large scales more sensitive to observations

Analysis at small scales mor sensitive t observatio



$$\mathbf{P}^a = (\mathbf{B}^{-1} + \mathbf{R}^{-1})^{-1}$$

$$\mathbf{S} = d\mathbf{x}^a / d\mathbf{y}$$

$$MI = \ln(\det(\mathbf{B}(\mathbf{P}^a)^{-1}))$$
$$= -0.5\ln(\det(\mathbf{I}-\mathbf{S}))$$

Where B and R are the prior and observation error covariance matrices respectively

2-variable illustrations conclusions

- In general as the magnitude of the OECs increase the information in the observations increases too.
- However, the Impact of OECs on the analysis cannot be considered in isolation of prior error correlations.
  - The greater the difference in the structures of prior and likelihood
    - The greater the reduction in analysis error variances
    - The greater the spread in information form the observations

Data thinning and data compression

- Observations with positive OECs have more small scale information than observations with uncorrelated error => greater benefit to having a denser observation network
- Can reduce amount of data by compressing observations such that the maximum amount of information is retained.
- Let $\mathbf{M}=\mathbf{R}^{-1/2}\mathbf{H}\mathbf{B}^{1/2}=\mathbf{U}\mathbf{\Lambda}^{\mathrm{M}}\mathbf{V}^{\mathrm{T}}$
- Then $MI = 0.5 \ln det(\mathbf{I} + \mathbf{M}\mathbf{M}^{\mathrm{T}})$ $DFS = trace(\mathbf{M}\mathbf{M}^{\mathrm{T}}(\mathbf{I} + \mathbf{M}\mathbf{M}^{\mathrm{T}})^{-1})$
- Can compress the observations using $\mathbf{C} = \mathbf{I}^{\mathrm{c}}\mathbf{U}^{\mathrm{T}}\mathbf{R}^{-1/2}$, where $\mathbf{I}^{\mathrm{c}} \in \mathbb{R}^{p_c \times p}$
- Ordering the observations w.r.t the singular values of M allows for the first p_c observations with the maximum information to be selected for assimilation

$$MI^{\mathrm{c}} = \sum_{k=1}^{p_c} \ln(1 + \lambda_k^{M^2})^{1/2} \qquad DFS^{\mathrm{c}} = \sum_{k=1}^{p_c} \lambda_k^{M^2} / (1 + \lambda_k^{M^2})$$

Isotropic, homogenous example

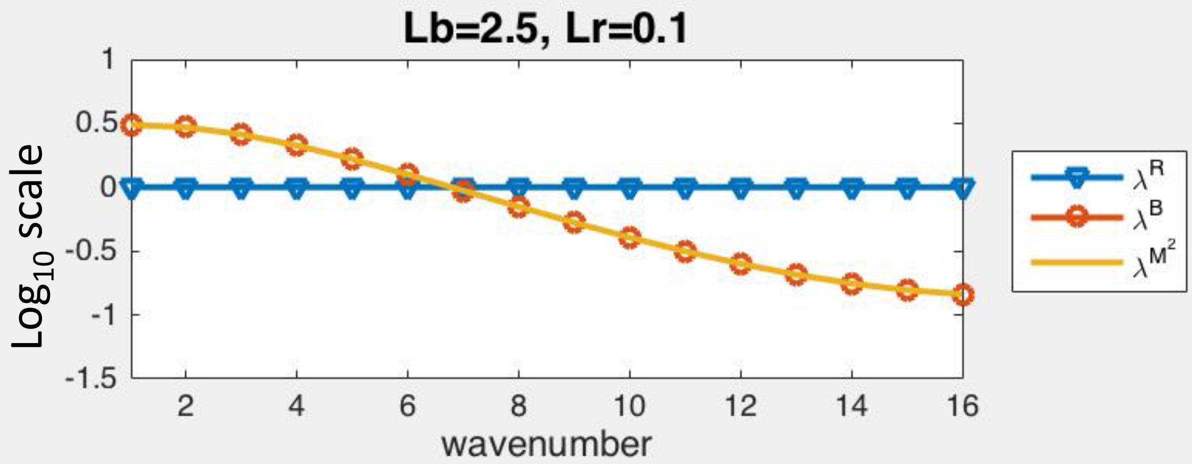
- Circulant matrices have the property that eigenvectors are given by the Fourier basis, F.
- Let $\mathbf{B}=\mathbf{F}\boldsymbol{\Gamma}\mathbf{F}^{\mathrm{T}}$, $\mathbf{R}=\mathbf{F}\boldsymbol{\Psi}\mathbf{F}^{\mathrm{T}}$ and $\mathbf{H}=\mathbf{I}$ (direct observations of the state)
- Then $\mathbf{C}=\mathbf{I}^{c}\boldsymbol{\Psi}^{-1/2}\mathbf{F}^{\mathrm{T}}$
- $\lambda_i^{M^2} = \gamma_i/\psi_i,$ where γ_i and ψ_i are the ith eigenvalue of $\mathbf{B}$ and $\mathbf{R}$ respectively.

- The most informative compressed observations are those associated with the scales at which the prior uncertainty is relatively large compared to the observation uncertainty.
- The reduction in the analysis error variance compared to the prior is given by

$$trace(\mathbf{B}-\mathbf{P}^{\mathrm{a}})^{\mathrm{c}} = \sum_{k=1}^{p_{\mathrm{c}}} \frac{\gamma_k \lambda_k^{M^2}}{1+\lambda_k^{M^2}}$$

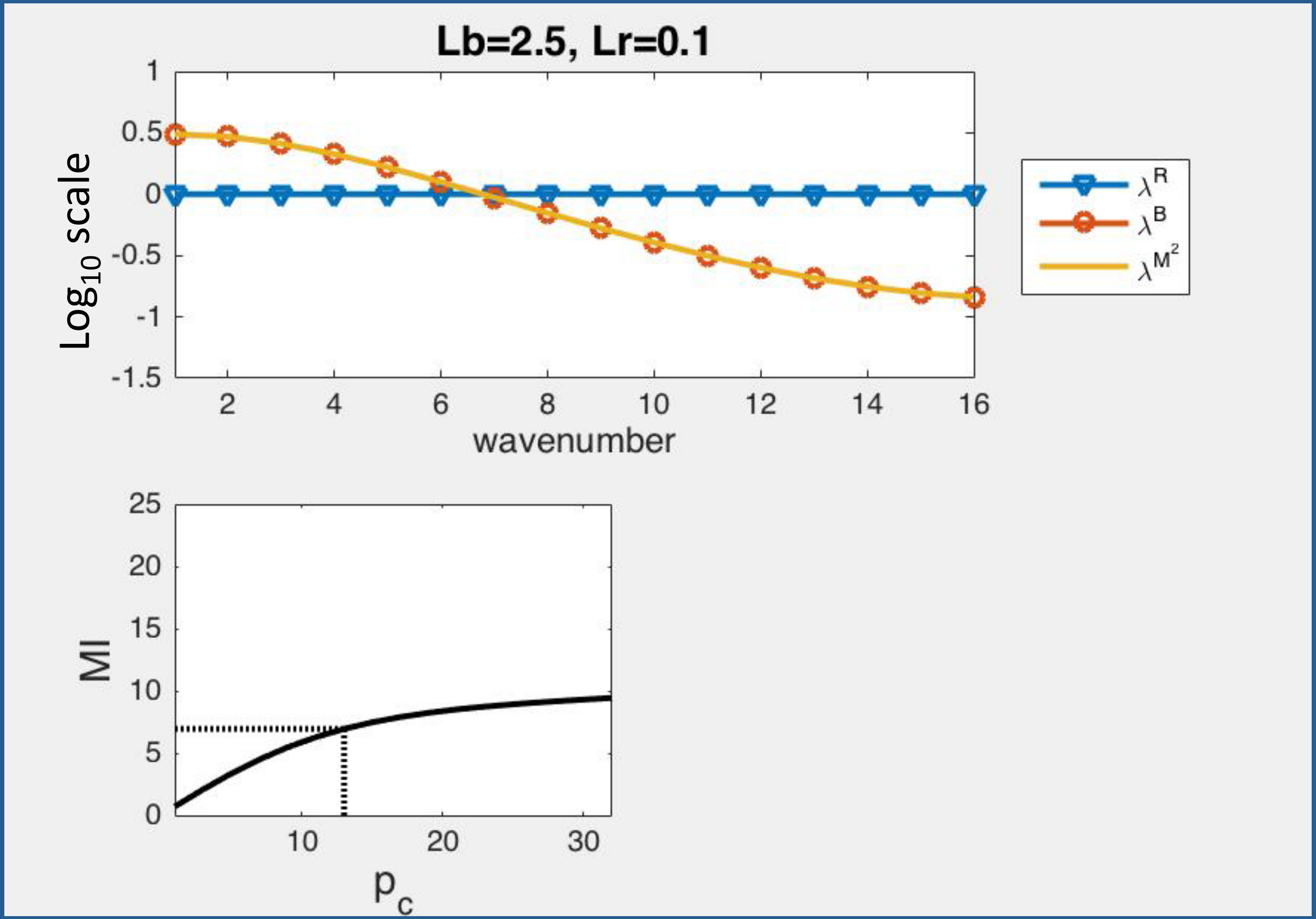
Isotropic, homogenous example…

circular grid discretised into 32 grid points. SOAR correlation structure.



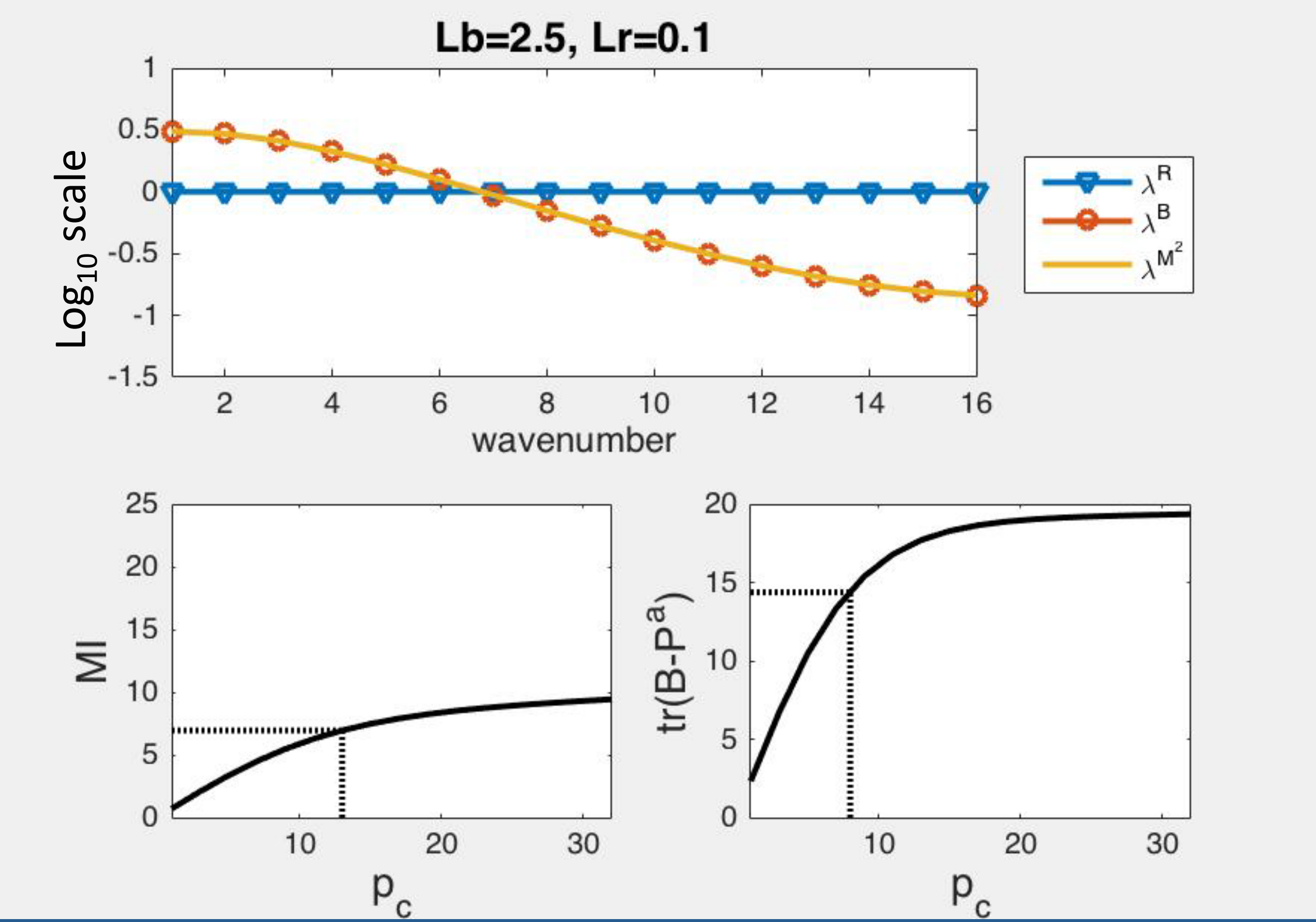
Isotropic, homogenous example…

circular grid discretised into 32 grid points. SOAR correlation structure.



Isotropic, homogenous example…

circular grid discretised into 32 grid points. SOAR correlation structure.



Observation network design Conclusions

- As the length-scales in the observation errors, Lr, increase the observations become more informative about the small scales.
- When Lr>Lb, the observations are more certain at small scale than the prior and so the benefit of denser observations increases.
  - Data compression can be used to help reduce the amount of data while retaining the small scale information (opposite to Super-obbing!)
  - Assimilating just the small-scale information may not result in the greatest reduction in analysis error
    - is this an issue for nested models?
    - use a metric which focuses on accuracy of small scales?

THANK YOU FOR LISTENING