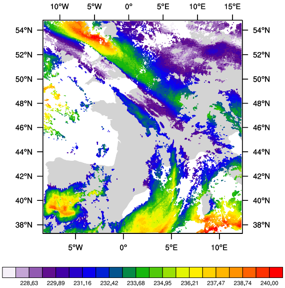
Modelling spatial correlations for observation errors with Lanczos : application to SEVIRI and RADAR data

Y. Michel[1] and S. Guedj[2]
[1] CNRM UMR 3589, Météo-France & CNRS, Toulouse, France
[2] JCSDA, College Park, MD, USA.
ISDA 2016, Reading, UK – July 20th, 2016

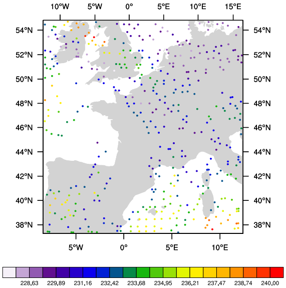
Introduction

- Horizontal correlations are not represented into the obs. error covariance matrix **R**.
- This certainly contributes to the undersampling of observations in our assimilation schemes.

Example with MSG/SEVIRI data in AROME



(a) Non-cloudy obs. (3km)



(b) Assimilated obs. (70km)

Introduction

Context

- We need $\mathbf{R}^{-1}$ when writting the gradient of the cost-function.
- $\mathbf{R}$ (and $\mathbf{R}^{-1}$) can probably be represented in block-diagonal form ; each block corresponding to an independent instrument.
- Even though, dimensions are big :

SEVIRI one channel is $p \sim 3712^2 \sim 10^7$ over the globe ; $p \sim 4 \cdot 10^5$ for AROME.

RADAR One elevation from a single radar gives $p \sim 512^2 = 2 \cdot 10^5$.

Introduction

Context

- We need $\mathbf{R}^{-1}$ when writting the gradient of the cost-function.
- $\mathbf{R}$ (and $\mathbf{R}^{-1}$) can probably be represented in block-diagonal form ; each block corresponding to an independent instrument.
- Even though, dimensions are big :

SEVIRI one channel is $p \sim 3712^2 \sim 10^7$ over the globe ; $p \sim 4 \cdot 10^5$ for AROME.

RADAR One elevation from a single radar gives $p \sim 512^2 = 2 \cdot 10^5$.

Possible approaches

- Finite elements version of the diffusion equation [Lindgren et. al 2011].
- Modelling of $\mathbf{R}$ with interpolations and estimating a truncated inverse with Lanczos [Fisher 2014].

Introduction

Context – II

- Modelling interchannel correlations (with direct methods) for infrared sounders has proven useful [Weston et.al 2014];
- Estimating **R** may be possible with the diagnostic from [Desroziers et. al 2005].
- Progress has been made for SEVIRI radiances and radial wind from Doppler RADAR [Waller et. al 2016].

Introduction

Context – II

- Modelling interchannel correlations (with direct methods) for infrared sounders has proven useful [Weston et.al 2014];
- Estimating **R** may be possible with the diagnostic from [Desroziers et. al 2005].
- Progress has been made for SEVIRI radiances and radial wind from Doppler RADAR [Waller et. al 2016].

Goal of this study

Evaluate the approach proposed by [Fisher 2014]:

- in a limited area context (for AROME) ;
- for SEVIRI and RADAR data.

Outline

1. Estimating spatial correlations in **R**
2. Modelling **R** with interpolations
3. Lanczos-based truncated inverse

Estimating spatial correlations in **R** : SEVIRI

Following [Desroziers et. al 2005]:

$$\mathbf{R} \approx \mathbb{E}(\mathbf{d}_a \mathbf{d}_b^\top)$$

Correlations in **R** are estimated through a time average/median.

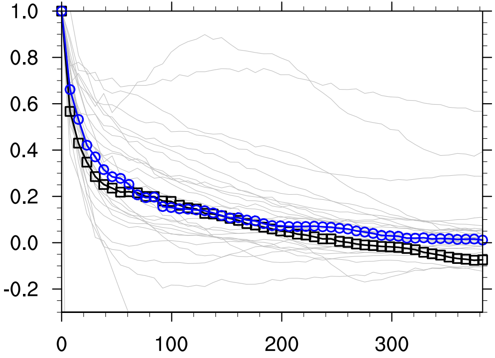


Figure : Application of the Desroziers diagnostic to the 6.2μm WV channel from SEVIRI, as in [Guedj et.al 2014].

Estimating spatial correlations in **R** : RADAR

[Waller et. al 2016]

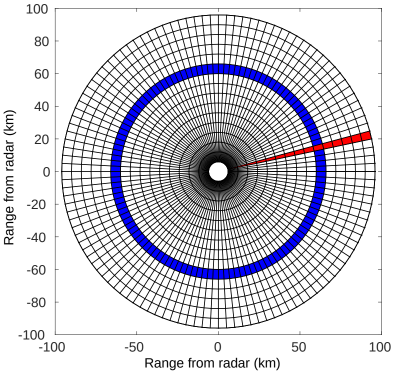


Figure : Geometry

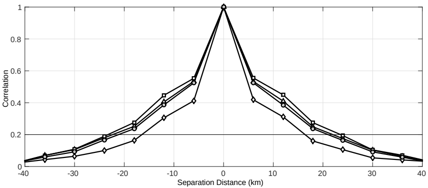


Figure : Azimuthal direction

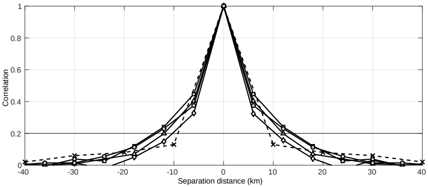


Figure : Radial direction

Outline

1. Estimating spatial correlations in **R**
2. Modelling **R** with interpolations
3. Lanczos-based truncated inverse

Modelling $\mathbf{R}$ with interpolations

[Fisher 2014] has proposed the following square-root model :

$$\mathbf{R} = \mathbf{U}\mathbf{U}^{\mathsf{T}}$$

where $\mathbf{U}$ is a sequence of operators :

$$\mathbf{U} = \mathbf{\Sigma_o}\mathbf{P}\mathbf{S}^{-1}\mathbf{D}\mathbf{S}$$

where :

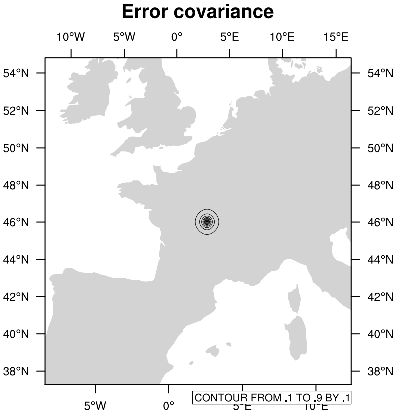
- $\mathbf{S}^{-1}\mathbf{D}\mathbf{S}$ is a spectral based correlation model ;
- $\mathbf{P}$ is the interpolation from a regular grid to the observation space ;
- $\mathbf{\Sigma_o}$ is the multiplication by the standard deviations.

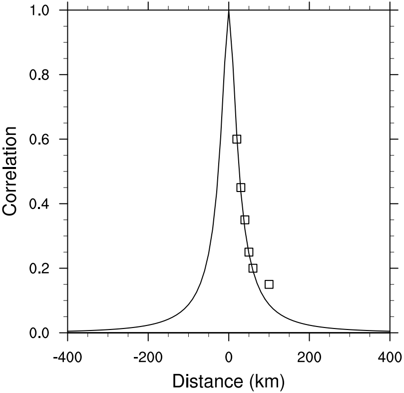
Note :

- the regular grid may be of coarse resolution ;
- we do not need to cover the whole analysis domain ;
- we can use alternative (gridpoint) correlation models.

Modelling **R** with interpolations : SEVIRI

- Define a grid covering the AROME domain, at 10 km spatial resolution (240 $\times$ 256 points).
- Use non-periodic hyperGaussians recursive filters ($\sigma = 60$ km, $\gamma = 5$, Purser et. al 2003b).





Modelling **R** with interpolations : SEVIRI

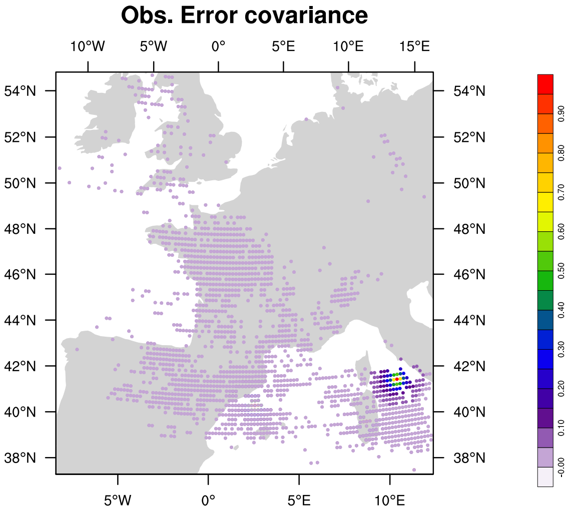


Fig: Spatial correlations modelled in **R**.

Modelling **R** with interpolations : SEVIRI

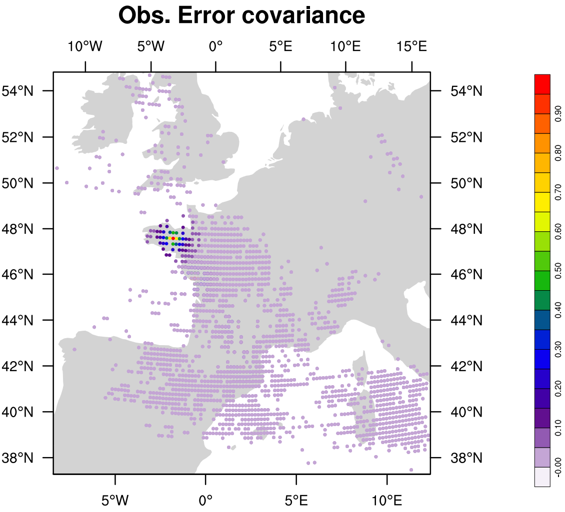


Fig: Spatial correlations modelled in **R**.

Modelling **R** with interpolations : SEVIRI

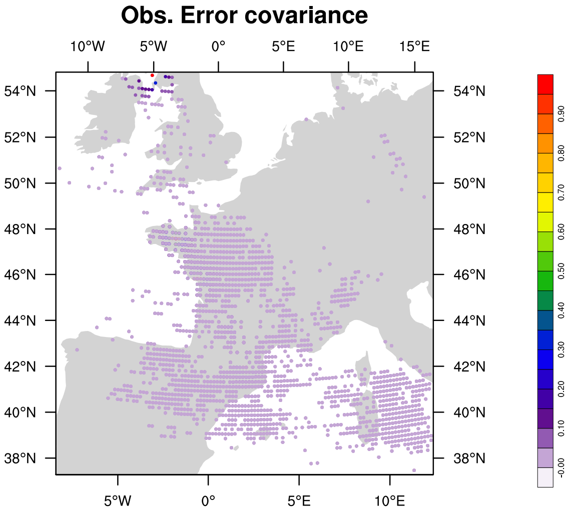
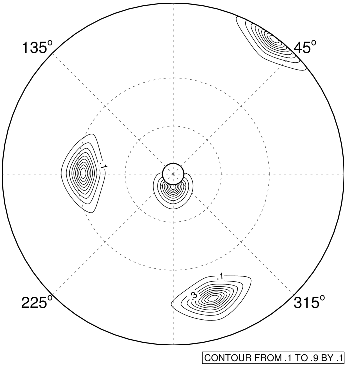


Fig: Spatial correlations modelled in **R**.

Modelling **R** with interpolations : RADAR

- Grid covering the radar only.
- Convolution with 1D recursive filters ;
- Periodic in the azimuth, non-periodic in the radial direction.



Modelling **R** with interpolations : RADAR

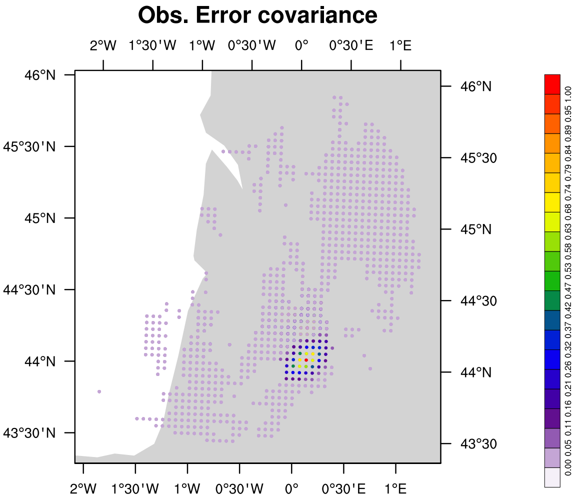


Fig: Spatial correlations modelled in **R**.

Modelling **R** with interpolations : RADAR

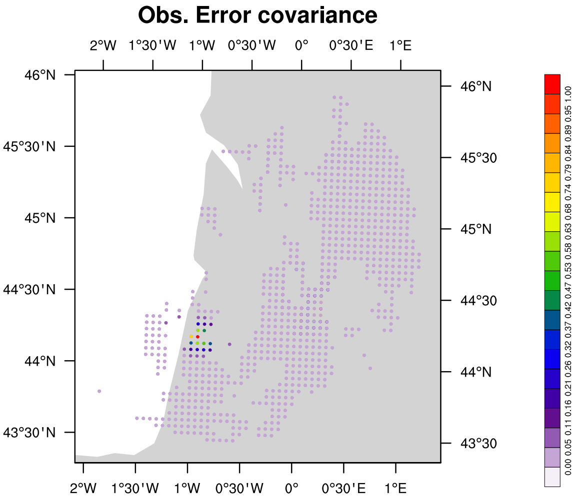


Fig: Spatial correlations modelled in **R**.

Modelling **R** with interpolations : RADAR

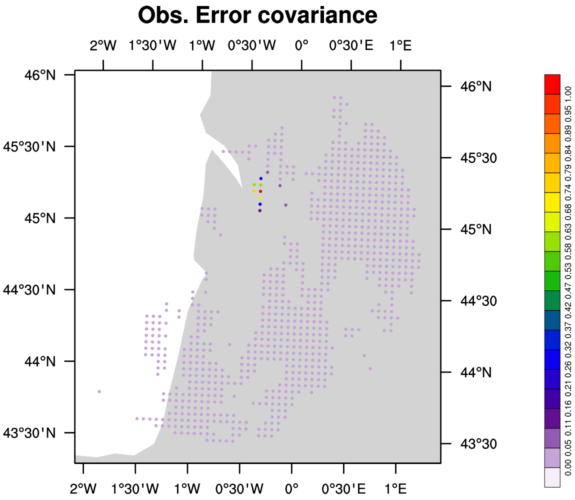


Fig: Spatial correlations modelled in **R**.

Outline

Lanczos-based truncated inverse

A reduced rank approximation of **R** may be obtained from the Lanczos algorithm :

$$\mathbf{R} \approx \mathbf{\Sigma}_o \left(\sum_{k=1}^{K} \lambda_k \mathbf{v}_k \mathbf{v}_k^{\mathrm{T}} \right) \mathbf{\Sigma}_o$$

When $K < p$, this approximation can be regularized :

$$\mathbf{R} \approx \widehat{\mathbf{R}} \equiv \mathbf{\Sigma}_o \left(\alpha \mathbf{I} + \sum_{k=1}^{K} (\lambda_k - \alpha) \mathbf{v}_k \mathbf{v}_k^{\mathrm{T}} \right) \mathbf{\Sigma}_o$$

allowing the explicit inversion formula :

$$\widehat{\mathbf{R}}^{-1} = \mathbf{\Sigma}_o^{-1} \left(\alpha^{-1} \mathbf{I} + \sum_{k=1}^{K} (\lambda_k^{-1} - \alpha^{-1}) \mathbf{v}_k \mathbf{v}_k^{\mathrm{T}} \right) \mathbf{\Sigma}_o^{-1}$$

Lanczos-based truncated inverse

The regularisation parameter α is chosen to conserve the total trace :

$$\mathrm{Tr}[\mathbf{R}] = \mathrm{Tr}[\widehat{\mathbf{R}}] \Longrightarrow \alpha = \frac{p - \sum_{k=1}^{K} \lambda_k}{p - K}$$

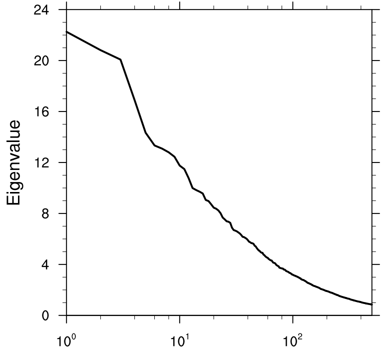
α is one minus the fraction of variance explained by the first K eigenvectors.

Lanczos-based truncated inverse

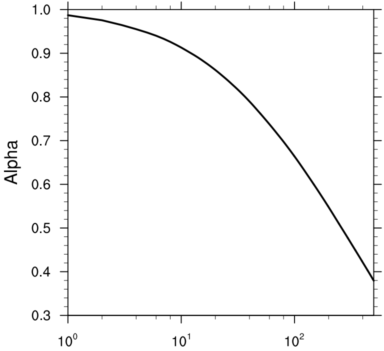
The regularisation parameter α is chosen to conserve the total trace :

$$\mathrm{Tr}[\mathbf{R}] = \mathrm{Tr}[\widehat{\mathbf{R}}] \implies \alpha = \frac{p - \sum_{k=1}^{K} \lambda_k}{p - K}$$

α is one minus the fraction of variance explained by the first K eigenvectors.



Eigenspectrum λ_k – SEVIRI



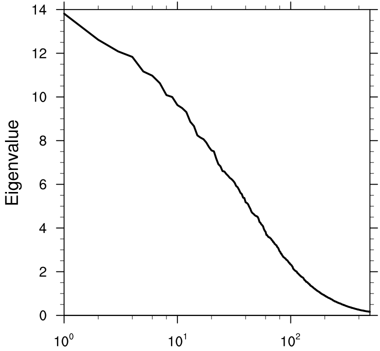
$\alpha(K)$ – SEVIRI

Lanczos-based truncated inverse

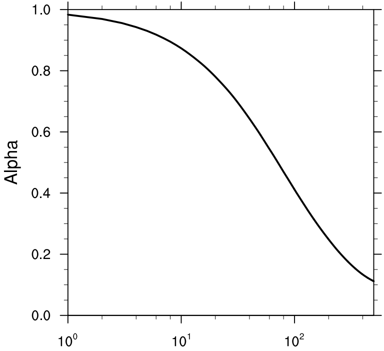
The regularisation parameter α is chosen to conserve the total trace :

$$\mathrm{Tr}[\mathbf{R}] = \mathrm{Tr}[\widehat{\mathbf{R}}] \implies \alpha = \frac{p - \sum_{k=1}^{K} \lambda_k}{p - K}$$

α is one minus the fraction of variance explained by the first K eigenvectors.



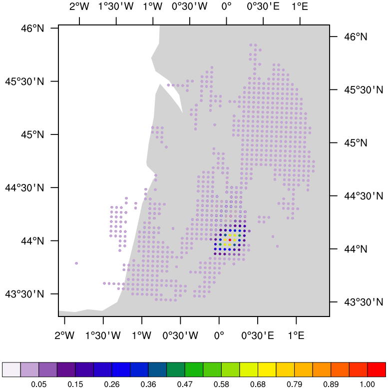
Eigenspectrum λ_k – RADAR



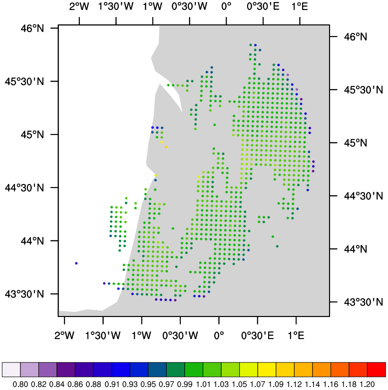
$\alpha(K)$ – RADAR

Effects of the truncation – RADAR

Truncation $K = 500$



Spatial correlations



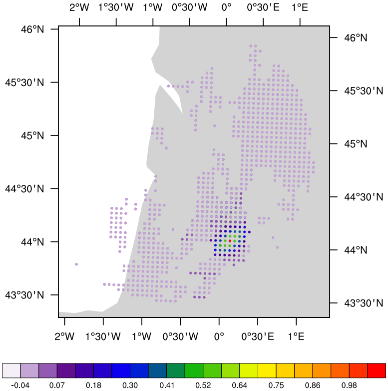
Variances

Truncating the spectrum introduces :

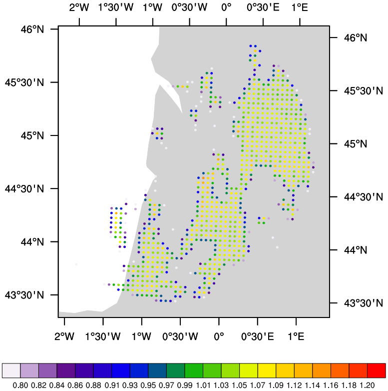
- erroneous long range correlations ;
- errors in the variances.

Effects of the truncation – RADAR

Truncation $K = 100$



Spatial correlations



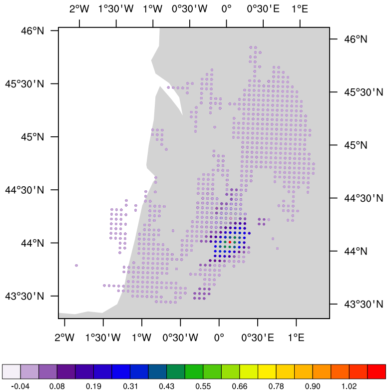
Variances

Truncating the spectrum introduces :

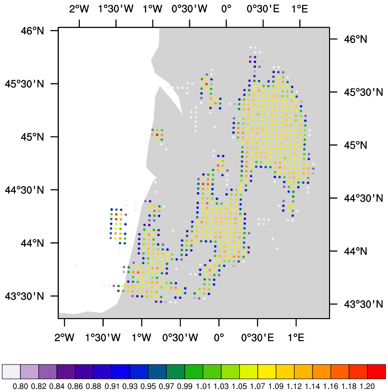
- erroneous long range correlations ;
- errors in the variances.

Effects of the truncation – RADAR

Truncation $K = 50$



Spatial correlations



Variances

Truncating the spectrum introduces :

- erroneous long range correlations ;
- errors in the variances.

Conclusions

Estimation : based on Desroziers' diagnostic, both SEVIRI and RADAR observations are significantly correlated.

Modelling : we can build a spatial correlation model to represent those correlations in observation space.

Inverse : use of the Lanczos algorithm [Fisher 2014]...

- may require a large number of eigenvectors (*e.g.*, 500).
- may introduce erroneous long range correlations / wrong variances if truncation is too severe.

but otherwise works !

References I

Desroziers, G., Berre, L., Chapnik, B. and Poli, P. (2005)
Diagnosis of observation, background and analysis-error statistics in observation space.
Q.J.R. Meteorol. Soc., **131**, 3385-3396.

M. Fisher, 2014.
Accounting for Correlated Observation Error in Variational Data Assimilation.
ESA workshop on correlated observation errors in Data Assimilation, April 24th, 2014.

S. Guedj, V. Guidard, B. Ménétrier and J.-F. Mahfouf (2014)
First estimates of observation error correlations for the future assimilation of MTG-IRS radiances.
ESA workshop on correlated observation errors in Data Assimilation, April 24th, 2014.

Lindgren, F., Rue, H. and Lindström, J. (2011)
An explicit link between Gaussian fields and Gaussian Markov random fields: the stochastic partial differential equation approach.
Journal of the Royal Statistical Society: Series B (Statistical Methodology), **73**: 423-498.

References II

J. A. Waller, D. Simonin, S. L. Dance, N. K. Nichols, and S. P. Ballard (2016)
Diagnosing observation error correlations for Doppler radar radial winds in the Met Office UKV model using observation-minus-background and observation-minus-analysis statistics.
Monthly Weather Review.

Weston, P. P., Bell, W. and Eyre, J. R. (2014)
Accounting for correlated error in the assimilation of high-resolution sounder data.
Q.J.R. Meteorol. Soc., **140**, 2420-2429