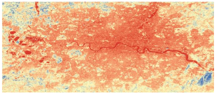
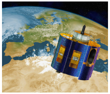
Data assimilation:
Addressing large problems with big data



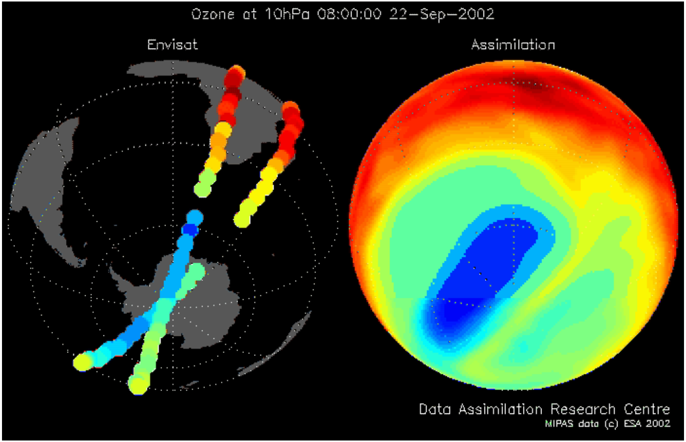


Sarah L. Dance

With thanks to Sue Ballard, Alan Geer, Graeme Kelly, Amos Lawless, Nancy Nichols, David Simonin, Jo Waller

NERC University of Reading EPSRC

What is data assimilation?



Courtesy Alan Geer

Applications of data assimilation

- Forecasting: Using recent observations to improve initial conditions for short-term predictions (e.g. numerical weather forecasting)

Applications of data assimilation

- Forecasting: Using recent observations to improve initial conditions for short-term predictions (e.g. numerical weather forecasting)
- Diagnosis, including parameter estimation: Testing and improving models by comparing predictions to observations (e.g. flood inundation forecasting)

Applications of data assimilation

- Forecasting: Using recent observations to improve initial conditions for short-term predictions (e.g. numerical weather forecasting)
- Diagnosis, including parameter estimation: Testing and improving models by comparing predictions to observations (e.g. flood inundation forecasting)
- Retrospective analysis: Learning more about how the Earth works, by using models to interpret/extend different types of data (e.g. climate studies)

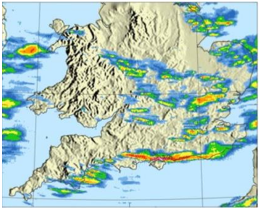
Applications of data assimilation

- Forecasting: Using recent observations to improve initial conditions for short-term predictions (e.g. numerical weather forecasting)
- Diagnosis, including parameter estimation: Testing and improving models by comparing predictions to observations (e.g. flood inundation forecasting)
- Retrospective analysis: Learning more about how the Earth works, by using models to interpret/extend different types of data (e.g. climate studies)
- Real-time Control: Use continually changing estimates of system state to determine control actions (e.g. intelligent sewers)

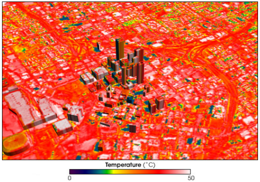
Outline

Context - Forecasting High Impact Weather

- High impact weather:
  - Intense rainfall → pluvial floods
  - Windstorms
  - Fog
  - Heat-stress



Radar data, July 11, 2012 ©Met Office 2012



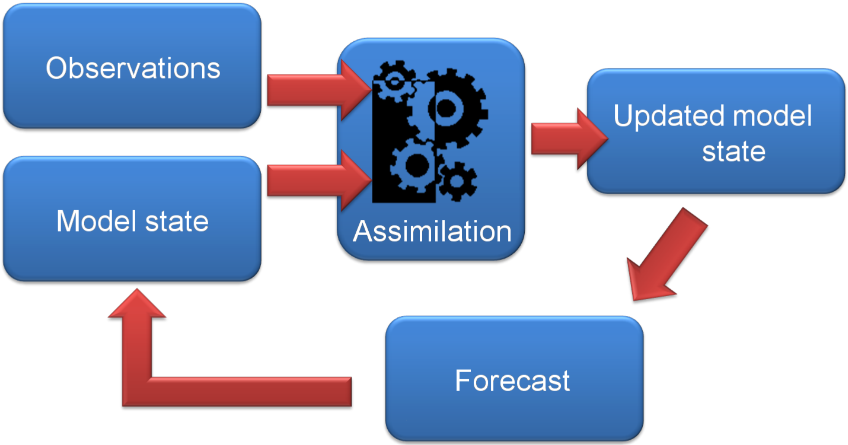
May 11-12 1997, Atlanta Georgia (NASA image). Daytime air temperatures were only 26.7° C but some of its surface temperatures soared to 47.8° C.

What is numerical weather prediction?

Picture from Met Office

- Numerical solution of a set of coupled PDES in complex geometry
- Initial-boundary value problem: we need an initial condition plus forcing at the top of the atmosphere and earth's surface

The forecast-assimilation cycle - provision of initial conditions



Practical and Numerical aspects

- Large problem
  - State vector $x \sim O(10^9)$, observations $y \sim O(10^7)$

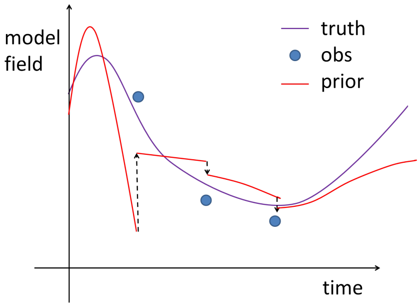
Practical and Numerical aspects

- Large problem
  - State vector $x \sim O(10^9)$, observations $y \sim O(10^7)$
- Time constrained
  - Need to wait for latest observations to arrive
  - A late forecast is useless!

Practical and Numerical aspects

- Large problem
  - State vector $x \sim O(10^9)$, observations $y \sim O(10^7)$
- Time constrained
  - Need to wait for latest observations to arrive
  - A late forecast is useless!
- Robustness
  - Critical operations rely on forecasts (Civil aviation, emergency response, wind power generation etc.)
  - Robustness more important than accuracy.

3D-Var



Minimize

$$J(\mathbf{x}) = \left(\mathbf{x} - \mathbf{x}^b\right)^T \mathbf{B}^{-1} \left(\mathbf{x} - \mathbf{x}^b\right) + (\mathbf{y} - \mathbf{H}\mathbf{x})^T \mathbf{R}^{-1} (\mathbf{y} - \mathbf{H}\mathbf{x})\,.$$

Statistical viewpoint of background errors

$$\mathbf{x}^t = \mathbf{x}^b + \varepsilon^b; \quad \mathbf{x}^t, \mathbf{x}^b, \varepsilon^b \in \mathbb{R}^n$$

Statistical viewpoint of background errors

$$\mathbf{x}^t = \mathbf{x}^b + \varepsilon^b; \quad \mathbf{x}^t, \mathbf{x}^b, \varepsilon^b \in \mathbb{R}^n$$

The background errors are assumed

- unbiased, $\mathbb{E}(\varepsilon^b) = 0$

Statistical viewpoint of background errors

$$\mathbf{x}^t = \mathbf{x}^b + \varepsilon^b; \quad \mathbf{x}^t, \mathbf{x}^b, \varepsilon^b \in \mathbb{R}^n$$

The background errors are assumed

- unbiased, $\mathbb{E}(\varepsilon^b) = 0$
- with covariance $\mathbb{E}\left[\varepsilon^b(\varepsilon^b)^T\right] = \mathbf{B}$

Statistical viewpoint of background errors

$$\mathbf{x}^t = \mathbf{x}^b + \varepsilon^b; \quad \mathbf{x}^t, \mathbf{x}^b, \varepsilon^b \in \mathbb{R}^n$$

The background errors are assumed

- unbiased, $\mathbb{E}(\varepsilon^b) = 0$
- with covariance $\mathbb{E}\left[\varepsilon^b (\varepsilon^b)^T\right] = \mathbf{B}$
- and Gaussian distributed $\varepsilon^b \sim N(\mathbf{0}, \mathbf{B})$

Tikhonov regularization viewpoint

- The minimization problem,

$$\min J(\mathbf{x}) = (\mathbf{y} - H(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{y} - H(\mathbf{x})) .$$

  is ill-posed, since there are not enough observations to define **x** uniquely.

Tikhonov regularization viewpoint

- The minimization problem,

$$\min J(\mathbf{x}) = (\mathbf{y} - H(\mathbf{x}))^T \mathbf{R}^{-1} (\mathbf{y} - H(\mathbf{x})) \,.$$

  is ill-posed, since there are not enough observations to define $\mathbf{x}$ uniquely.
- Addition of background term can be viewed as regularization.
- $\mathbf{B}$ is chosen to give a balance between fitting the observations closely and making the problem easier to solve.

Role of $\mathbf{B}$ in the analysis

Consider the solution for the analysis in the following form,

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-T}(\mathbf{y} - \mathbf{H}\mathbf{x}^b).$$

Role of **B** in the analysis

Consider the solution for the analysis in the following form,

$$\mathbf{x}^a = \mathbf{x}^b + \mathbf{B}\mathbf{H}^T(\mathbf{H}\mathbf{B}\mathbf{H}^T + \mathbf{R})^{-T}(\mathbf{y} - \mathbf{H}\mathbf{x}^b).$$

Analysis increments are linear combinations of the columns of **B.**

Spreading and smoothing

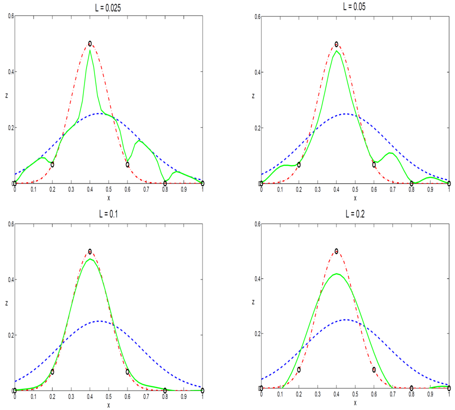


Figure 5.5: Effect of varying the correlation length scale L in the Gaussian state background error covariance matrix $\mathbf{B_{zz}}$ (5.22). The red dot-dash line represents the true bathymetry $\mathbf{z}^t$, observations $\mathbf{y}$ are given by circles, the background $\mathbf{z}^b$ by the blue dashed line and the analysis $\mathbf{z}^a$ by the solid green line.

Smith (2010)

Multivariate effects

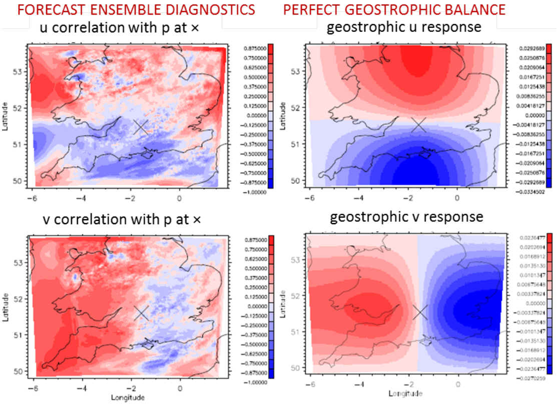


Figure from Ross Bannister

Background errors summary

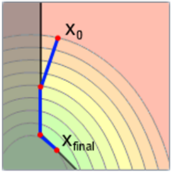
- Background errors play a large role in
  - Spreading and smoothing the increments
  - Setting the multivariate balance properties of the analysis

Background errors summary

- Background errors play a large role in
  - Spreading and smoothing the increments
  - Setting the multivariate balance properties of the analysis
- **B** cannot be estimated directly so it is often estimated using forecast differences, or ensembles (of forecasts or analyses) as a proxy.
- For more information see Bannister (2008a,b); Buehner (2010)

Open question - Constrained DA

- Some variables need to be constrained for physical reasons e.g.
  - humidity
  - rainfall rate
- Can resolve this by using constraints in minimization BUT reduction in minimization rate
  - Brute force enforcement
  - Karush-Kuhn-Tucker
  - Barrier methods
- How can we enforce positivity without increasing wall-clock time?



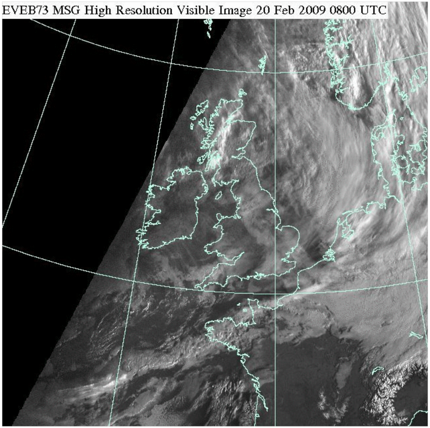
What are observation errors?

In data assimilation, we consider the observation equation

$$\mathbf{y} = H(\mathbf{x}) + \varepsilon.$$

We assume ε is unbiased, $\mathbb{E}(\varepsilon) = 0$, and has covariance $\mathbf{R}$ such that

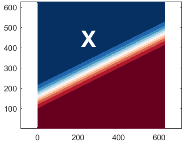
$$\mathbf{R}_{ij} = \mathbb{E}(\varepsilon_i \varepsilon_j).$$



Where do observation errors come from?

The error vector, ε, contains errors from four main sources:

- Instrument noise
- Observation processing
- Forward model error
- Representativity error
  - contrasting model and observation resolutions
  - observations resolve spatial scales or features that the model cannot
  - Problem: we know little about natural variability in cities



http://www.met.reading.ac.uk/micromet/ir-images/

Problems dealing with observation error correlations

- Magnitude and character of observation weighting matrices largely unknown - can only be estimated in a statistical sense, not observed directly

Problems dealing with observation error correlations

- Magnitude and character of observation weighting matrices largely unknown - can only be estimated in a statistical sense, not observed directly
- In the early days of data assimilation, there were fewer observations from remote sensing, so historically the most effort has been put into modelling and preconditioning **B**.

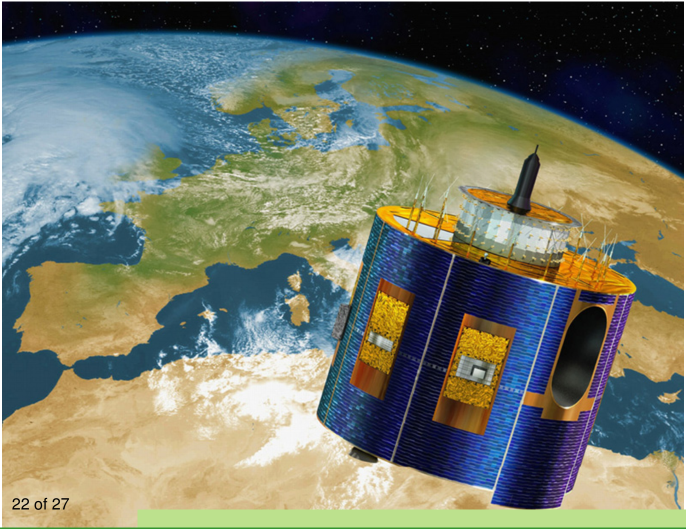
Problems dealing with observation error correlations

- Magnitude and character of observation weighting matrices largely unknown - can only be estimated in a statistical sense, not observed directly
- In the early days of data assimilation, there were fewer observations from remote sensing, so historically the most effort has been put into modelling and preconditioning **B**.
- Observations thinned spatially, and the **R** matrix treated as diagonal (about 80% loss)

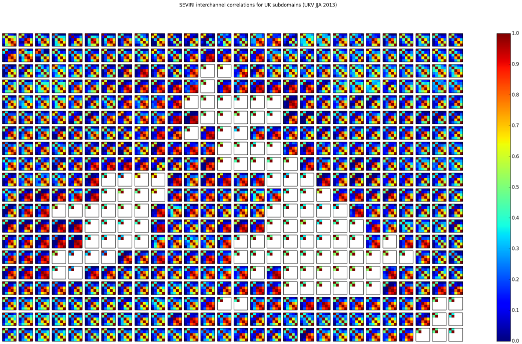
Problems dealing with observation error correlations

- Magnitude and character of observation weighting matrices largely unknown - can only be estimated in a statistical sense, not observed directly
- In the early days of data assimilation, there were fewer observations from remote sensing, so historically the most effort has been put into modelling and preconditioning **B**.
- Observations thinned spatially, and the **R** matrix treated as diagonal (about 80% loss)
- However with higher resolution forecasting we need to retain the observed information on finer scales.

Example - Spinning Enhanced Visible and Infrared Imager (SEVIRI)



Diagnosed Interchannel correlations (Waller et al., 2016)



What is happening on the coasts?

Observation Error summary Conclusions

- It is important to be able to account for observation error correlations
  - Avoid thinning (high resolution forecasting)
  - More information content
  - Better analysis accuracy
  - Improved NWP skill score

Observation Error summary Conclusions

- It is important to be able to account for observation error correlations
  - Avoid thinning (high resolution forecasting)
  - More information content
  - Better analysis accuracy
  - Improved NWP skill score
- Observation error review Janjić et al. (2016)

Open question - new computing architecture

- How can we write data assimilation algorithms that are efficient on new massively parallel computing architecture?
- Can we cope with heterogeneous, intermittent, low quality data and still get good results?
- Keep up to date with new data assimilation blog and network(Get in touch if you want to contribute!)
  `http://blogs.reading.ac.uk/dare/`

Useful references

New textbooks:

- Asch et al. (2016) is a text placing data assimilation in the broader context of inverse problems.
- Reich and Cotter (2015) take a more probabilistic Bayesian approach.

Classics:

- Daley (1991)
- Bennett (2002)
- Jazwinski (1970)

Research reviews/collections covering a range of applications

- Park and Xu (2017)
- Lahoz et al. (2010)

References

M. Asch, M. Bocquet, and M. Nodet. *Data Assimilation: Methods, Algorithms, and Applications:*. Fundamentals of Algorithms. SIAM, 2016. ISBN 9781611974539.

R. N. Bannister. A review of forecast error covariance statistics in atmospheric variational data assimilation. i: Characteristics and measurements of forecast error covariances. *Quarterly Journal of the Royal Meteorological Society*, 134(637):1951–1970, 2008a.

R. N. Bannister. A review of forecast error covariance statistics in atmospheric variational data assimilation. ii: Modelling the forecast error covariance statistics. *Quarterly Journal of the Royal Meteorological Society*, 134(637):1971–1996, 2008b.

A.F. Bennett. *Inverse Modeling of the Ocean and Atmosphere*. Cambridge University Press, 2002.

M. Buehner. Error statistics in data assimilation: Estimation and modelling. In W. Lahoz, B. Khattatov, and R. Menard, editors, *Data Assimilation: Making Sense of Observations*, pages 93–112. Springer, Berlin, Heidelberg, 2010. doi: 10.1007/978-3-540-74703-1_5.

R. Daley. *Atmospheric Data Analysis*. Cambridge University Press, Cambridge, 1991.

T. Janjić, N. Bormann, M. Bocquet, J. A. Carton, S. E. Cohn, S. L. Dance, S. N. Losa, N. K. Nichols, R. Potthast, J. A. Waller, and P. Weston. On the representation error in data assimilation. *submitted*, 2016.

A. H. Jazwinski. *Stochastic Processes and Filtering Theory*. Academic Press, 1970. 376 pp.

W. Lahoz, B. Khattatov, and R. Menard, editors. *Data Assimilation: Making Sense of Observations*. Springer, 2010.

S. K. Park and L. Xu, editors. *Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications (Vol. III)*. Springer, 2017.

S. Reich and C. Cotter. *Probabilistic Forecasting and Bayesian Data Assimilation*. Cambridge University Press, 2015. ISBN 9781107663916.

P.J. Smith. *Joint state and parameter estimation using data assimilation with application to morphodynamic modelling*. PhD thesis, Dept of Mathematics, University of Reading, 2010. URL `http://www.reading.ac.uk/maths-and-stats/research/theses/maths-phdtheses.aspx`.

J. A. Waller, S. P. Ballard, S. L. Dance, G. Kelly, N. K. Nichols, and D. Simonin. Diagnosing horizontal and inter-channel observation error correlations for seviri observations using observation-minus-background and observation-minus-analysis statistics. *Remote Sensing*, 8(7), 2016. doi: 10.3390/rs8070581.