National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

NERC SCIENCE OF THE ENVIRONMENT

University of Reading

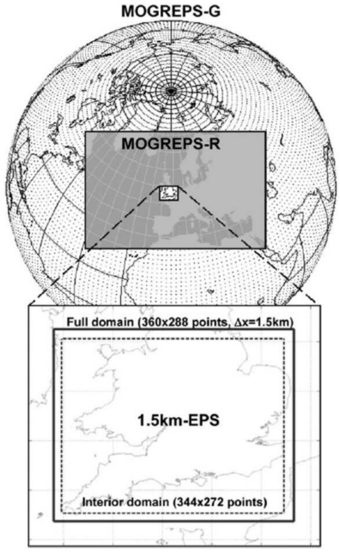
Model errors for MetUM winds – can we estimate these well with dense observations?

Ross Bannister, NCEO Reading, r.n.bannister@reading.ac.uk

- The Met Office has a high resolution MetUM over the UK.
- Little is known about the statistics of its errors.
- Diagnosing model error statistics is difficult!

There are different aspects to errors in general:

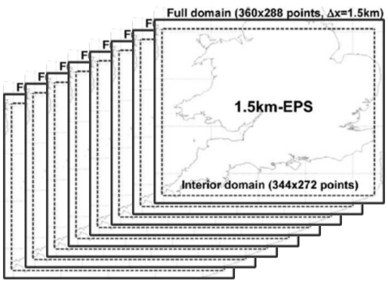
- They can have random and systematic components.
- They may be additive or multiplicative
- They may be Gaussian or non-Gaussian.
- They may be homogeneous or inhomogeneous.
- They may be isotropic or anisotropic.
- They may be correlated or uncorrelated.

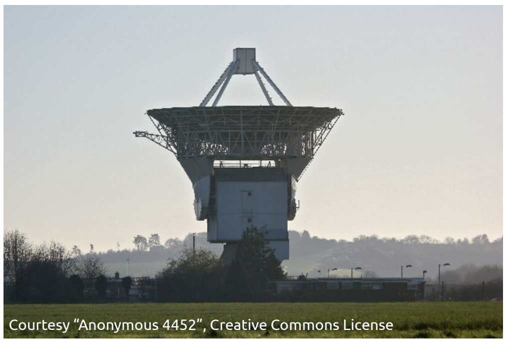


Courtesy Caron (2013), ©Crown Copyright 2013, Met Office

We have for one day (20/09/11):

- 1.8 million Doppler radar observations (Chilbolton Observatory).
- High-resolution forecasts (1.5km grid-length) over the Southern UK.
- A 93-member ensemble of forecasts.





Courtesy "Anonymous 4452", Creative Commons License

Assumptions:

	Random only?	Additive?	Gaussian?	Homo-geneous?	Isotropic?	Mutually uncorrel-ated?	Cross correl-ated?
Model errors	✓	✓	✓	✓	✓	✗	✗
Predictability errors	✓	✓	✓	✓	✓	✗	✗
Observation errors	✓	✓	✓	✓	✓	✓	✗

… plus that the ensemble is correctly spread, and that observations can be treated as 'point-like'. These assumptions are not always justified.

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

NERC SCIENCE OF THE ENVIRONMENT

University of Reading

General definitions

Forecast, analysis, or observation = 'Truth' + Error realisation

Innovation = Observation - Forecast 'observation'

Covariance = < (Quantity 1 – mean of Q1) × (Quantity 2 – mean of Q2) >

Specific definitions

Forecast (at time t)

$$\mathbf{x}_t^{\mathrm{f}} = \mathbf{x}_t^{\mathrm{t}} + \boldsymbol{\epsilon}_t^{\mathrm{f}}$$

Analysis (at $t = 0$)

$$\mathbf{x}_0^{\mathrm{a}} = \mathbf{x}_0^{\mathrm{t}} + \boldsymbol{\epsilon}_0^{\mathrm{a}}$$

Observations

$$\mathbf{y}_t = \mathcal{H}_t(\mathbf{x}_t^{\mathrm{t}}) + \boldsymbol{\epsilon}_t^{\mathrm{ob}}$$

Observation operator applied to forecast

$$\begin{aligned} \mathcal{H}_t(\mathbf{x}_t^{\mathrm{f}}) &= \mathcal{H}_t(\mathbf{x}_t^{\mathrm{t}} + \boldsymbol{\epsilon}_t^{\mathrm{f}}) \\ &\approx \mathcal{H}_t(\mathbf{x}_t^{\mathrm{t}}) + \mathbf{H}_t\boldsymbol{\epsilon}_t^{\mathrm{f}} \end{aligned}$$

Forecast model (no stochastic scheme)

$$\mathcal{M}(\mathbf{x}_{t-1}^{\mathrm{t}}) = \mathbf{x}_t^{\mathrm{t}} + \boldsymbol{\eta}_t$$

Using the definitions

Applying the forecast model

$$
\begin{aligned}
\mathbf{x}_t^{\mathrm{f}} &= \mathcal{M}_t(\mathbf{x}_{t-1}^{\mathrm{a}}) \\
&= \mathcal{M}_t(\mathbf{x}_{t-1}^{\mathrm{t}} + \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}}) \\
&\approx \mathcal{M}_t(\mathbf{x}_{t-1}^{\mathrm{t}}) + \mathbf{M}_t\boldsymbol{\epsilon}_{t-1}^{\mathrm{a}} \\
\mathbf{x}_t^{\mathrm{t}} + \boldsymbol{\epsilon}_t^{\mathrm{f}} &= \mathbf{x}_t^{\mathrm{t}} + \underbrace{\boldsymbol{\eta}_t}_{\text{model error}} + \underbrace{\mathbf{M}_t\boldsymbol{\epsilon}_{t-1}^{\mathrm{a}}}_{\text{predictability error}}
\end{aligned}
$$

Innovations

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

Ensemble perturbations

$$
\begin{aligned}
\boldsymbol{\delta}_t^{\mathrm{f}} &\approx \mathbf{x}_t^{\mathrm{f}} - \langle \mathbf{x}_t^{\mathrm{f}} \rangle \\
&= \mathcal{M}_t(\mathbf{x}_{t-1}^{\mathrm{a}}) - \langle \mathcal{M}_t(\mathbf{x}_{t-1}^{\mathrm{a}}) \rangle \\
&= \mathcal{M}_t(\mathbf{x}_{t-1}^{\mathrm{t}} + \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}}) - \langle \mathcal{M}_t(\mathbf{x}_{t-1}^{\mathrm{t}} + \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}}) \rangle \\
&\approx \mathcal{M}_t(\mathbf{x}_{t-1}^{\mathrm{t}}) + \mathbf{M}_t\boldsymbol{\epsilon}_{t-1}^{\mathrm{a}} - \langle \mathcal{M}_{t-1\to t}(\mathbf{x}_{t-1}^{\mathrm{t}}) \rangle \\
&= \mathbf{M}_t\boldsymbol{\epsilon}_{t-1}^{\mathrm{a}}.
\end{aligned}
$$

National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

NERC SCIENCE OF THE ENVIRONMENT

University of Reading

Procedure – interpretation of Daley (1992)

Daley, Roger (1992) The effect of serially correlated observation and model error on atmospheric data assimilation, Monthly weather review 120 (1), 164 – 177.

1. Analysis of innovations

$$
\begin{aligned}
\left\langle \mathbf{d}_t^{\mathrm{o,b}} \mathbf{d}_t^{\mathrm{o,b}\,\mathrm{T}} \right\rangle &= \left\langle \left(\mathbf{y}_t - \mathcal{H}_t(\mathbf{x}_t^{\mathrm{f}})\right) \left(\mathbf{y}_t - \mathcal{H}_t(\mathbf{x}_t^{\mathrm{f}})\right)^{\mathrm{T}} \right\rangle \\
&= \left\langle \left(\boldsymbol{\epsilon}_t^{\mathrm{ob}} - \mathbf{H}_t \boldsymbol{\epsilon}_t^{\mathrm{f}}\right) \left(\boldsymbol{\epsilon}_t^{\mathrm{ob}} - \mathbf{H}_t \boldsymbol{\epsilon}_t^{\mathrm{f}}\right)^{\mathrm{T}} \right\rangle \\
&= \left\langle \left(\boldsymbol{\epsilon}_t^{\mathrm{ob}} - \mathbf{H}_t \boldsymbol{\eta}_t - \mathbf{H}_t \mathbf{M}_t \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}}\right) \left(\boldsymbol{\epsilon}_t^{\mathrm{ob}} - \mathbf{H}_t \boldsymbol{\eta}_t - \mathbf{H}_t \mathbf{M}_t \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}}\right)^{\mathrm{T}} \right\rangle \\
&\approx \underbrace{\underbrace{\left\langle \boldsymbol{\epsilon}_t^{\mathrm{ob}} \boldsymbol{\epsilon}_t^{\mathrm{ob}\,\mathrm{T}} \right\rangle}_{\mathbf{R}_t}}_{\text{observation error part}} + \underbrace{\mathbf{H}_t \underbrace{\left\langle \boldsymbol{\eta}_t \boldsymbol{\eta}_t^{\mathrm{T}} \right\rangle}_{\mathbf{Q}_t} \mathbf{H}_t^{\mathrm{T}}}_{\text{model error part}} + \underbrace{\mathbf{H}_t \mathbf{M}_t \underbrace{\left\langle \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}} \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}\,\mathrm{T}} \right\rangle}_{\mathbf{A}_{t-1}} \mathbf{M}_t^{\mathrm{T}} \mathbf{H}_t^{\mathrm{T}}}_{\text{predictability error part}}
\end{aligned}
$$

Warning
Analysis of innovations will be susceptible to biases in obs and in model

2. Ensemble analysis (no localisation)

$$
\begin{aligned}
\left\langle \left(\mathbf{H}_t \left[\mathbf{x}_t^{\mathrm{f}} - \langle \mathbf{x}_t^{\mathrm{f}} \rangle\right]\right) \left(\mathbf{H}_t \left[\mathbf{x}_t^{\mathrm{f}} - \langle \mathbf{x}_t^{\mathrm{f}} \rangle\right]\right)^{\mathrm{T}} \right\rangle &\approx \Big\langle \left(\mathbf{H}_t \left[\mathbf{x}_t^{\mathrm{t}} + \boldsymbol{\eta}_t + \mathbf{M}_t \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}} - \langle \mathbf{x}_t^{\mathrm{t}} + \boldsymbol{\eta}_t + \mathbf{M}_t \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}} \rangle\right]\right) \times \\
&\qquad \left(\mathbf{H}_t \left[\mathbf{x}_t^{\mathrm{t}} + \boldsymbol{\eta}_t + \mathbf{M}_t \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}} - \langle \mathbf{x}_t^{\mathrm{t}} + \boldsymbol{\eta}_t + \mathbf{M}_t \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}} \rangle\right]\right)^{\mathrm{T}} \Big\rangle \\
&= \Big\langle \left(\mathbf{H}_t \left[\mathbf{M}_t \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}} - \langle \mathbf{M}_t \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}} \rangle\right]\right) \times \\
&\qquad \left(\mathbf{H}_t \left[\mathbf{M}_t \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}} - \langle \mathbf{M}_t \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}} \rangle\right]\right)^{\mathrm{T}} \Big\rangle \\
&\approx \left\langle \left(\mathbf{H}_t \mathbf{M}_t \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}}\right) \left(\mathbf{H}_t \mathbf{M}_t \boldsymbol{\epsilon}_{t-1}^{\mathrm{a}}\right)^{\mathrm{T}} \right\rangle \\
&= \underbrace{\mathbf{H}_t \mathbf{M}_t \mathbf{A}_{t-1} \mathbf{M}_t^{\mathrm{T}} \mathbf{H}_t^{\mathrm{T}}}_{\text{predictability error part}}
\end{aligned}
$$

3. Estimate the model error contribution

$$
\mathbf{H}_t \left\langle \boldsymbol{\eta}_t \boldsymbol{\eta}_t^{\mathrm{T}} \right\rangle \mathbf{H}_t^{\mathrm{T}} \approx \left\langle \mathbf{d}_t^{\mathrm{o,b}} \mathbf{d}_t^{\mathrm{o,b}\,\mathrm{T}} \right\rangle - \left\langle \left(\mathbf{H}_t \left[\mathbf{x}_t^{\mathrm{f}} - \langle \mathbf{x}_t^{\mathrm{f}} \rangle\right]\right) \left(\mathbf{H}_t \left[\mathbf{x}_t^{\mathrm{f}} - \langle \mathbf{x}_t^{\mathrm{f}} \rangle\right]\right)^{\mathrm{T}} \right\rangle - \mathbf{R}_t
$$

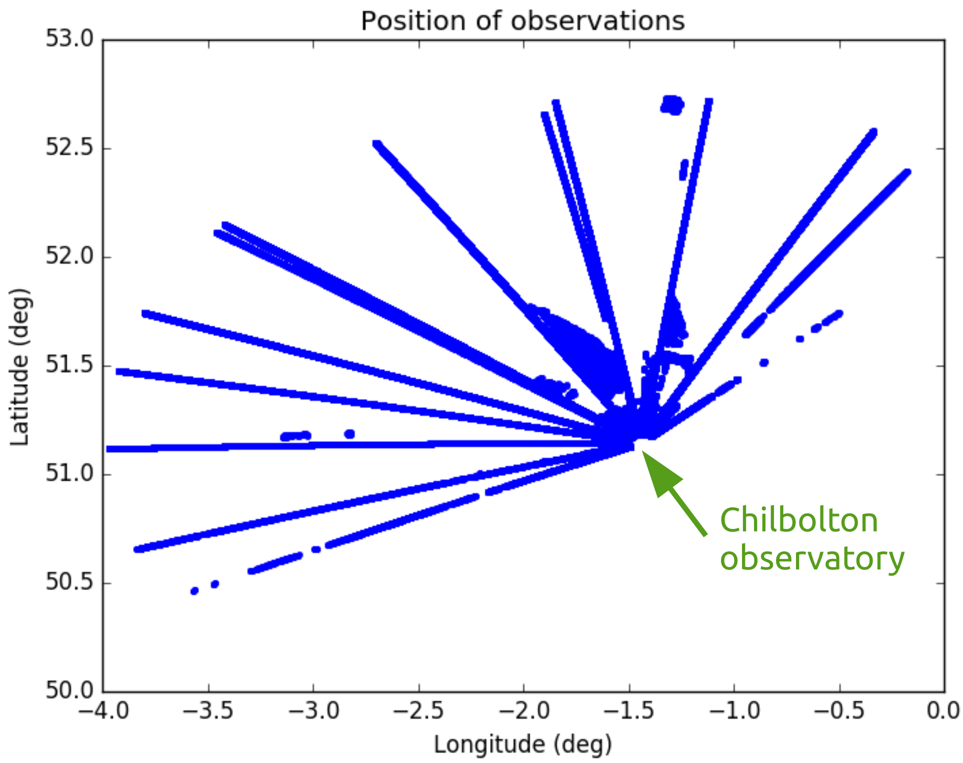
Warning
Model error is estimated as the residual of imperfectly known contributions

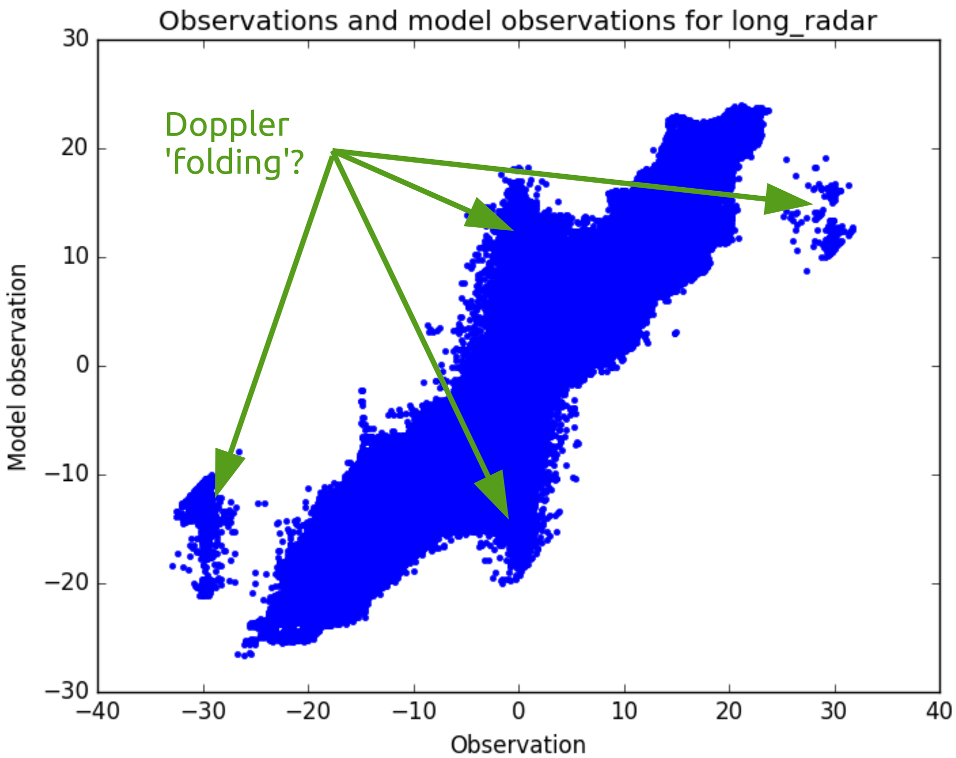
National Centre for Earth Observation
NATURAL ENVIRONMENT RESEARCH COUNCIL

NERC SCIENCE OF THE ENVIRONMENT

University of Reading

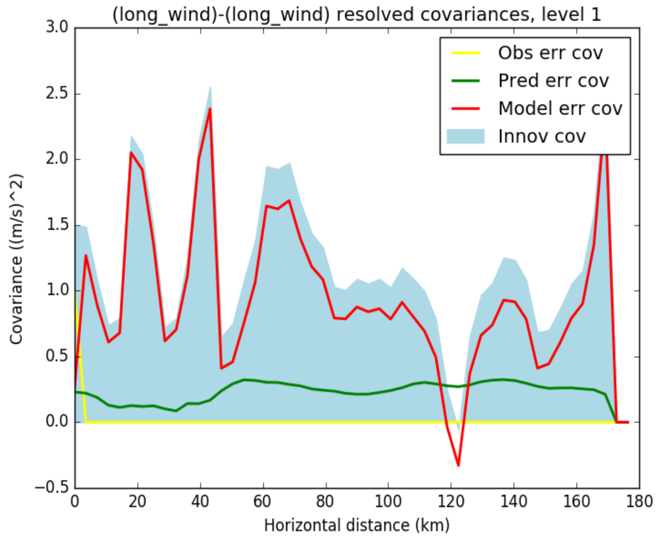
The observations and the model's equivalents



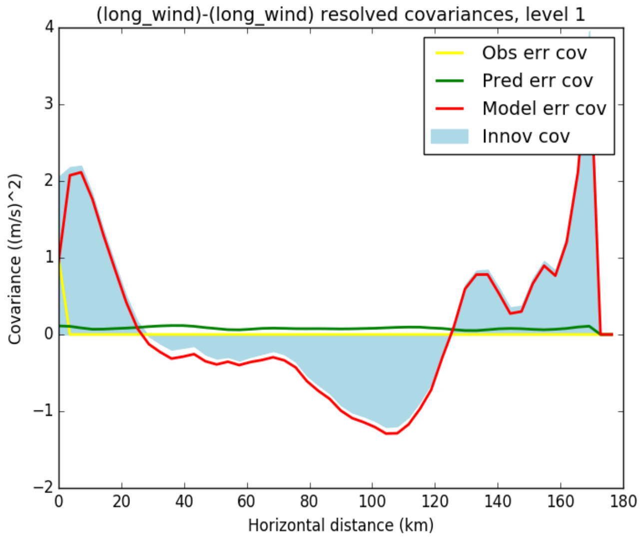


Provisional computations

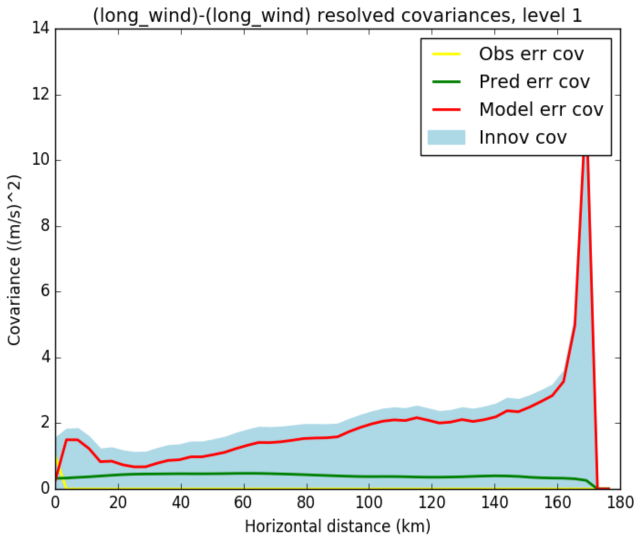
Forecast lead time 1hr ± 1hr



Forecast lead time 2hr ± 1hr



Forecast lead time 3hr ± 1hr



Comments

- No. of samples ~$O(num\ obs)^2$.
- Sampling errors are deemed negligible.
- Have assumed contribution from (i) *w* (vertical wind), and (ii) effect of atmospheric refraction, to 'model obs' are negligible.
- The above results show systematic problems with the procedure on P. 4.
- Some covariances look noisy and increase with distance.
  - Is this an artefact?
  - If not, why does this happen?
- What is the source of the large innovation covariance at ~170 km?
- Need to do quality control on observations (e.g. Doppler 'folding')?
- Need to check for systematic errors, and remove where possible.
- How can these results be used?