Quantitative Network Design

Peter Rayner
Thomas Kaminski

Kaminski and Rayner "Assimilation and Network Design" in "Observing the Greenhouse Gas Balance" Dolman et al., Springer, 2008

A scientist uses statistics like a drunk uses a lamp-post, for support rather than illumination

Outline

- Designing a network for flux determination
  - Basic theory
  - Some examples
- Network design in an assimilation context
- Lessons and perspectives

Basic theory

$$\chi^2 = \frac{1}{2}\left[(\vec{D} - \mathbf{J}\vec{S})^T \mathbf{C}(\vec{D})^{-1}(\vec{D} - \mathbf{J}\vec{S}) + (\vec{S} - \vec{S}_0)^T \mathbf{C}(\vec{S}_0)^{-1}(\vec{S} - \vec{S}_0)\right]$$

where $\vec{S}$ is sources, $\vec{D}$ data, $\mathbf{J}$ transport Jacobian, and the subscript 0 indicates the prior and $\mathbf{C}$ covariance.

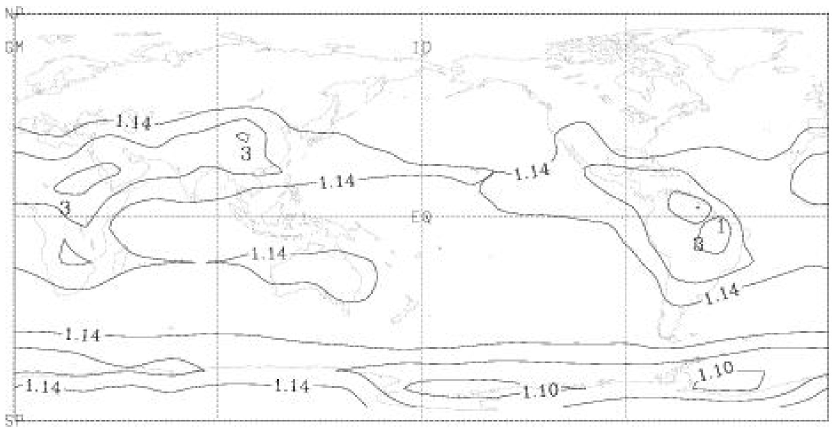
$$\vec{S} = \vec{S}_0 + \mathbf{C}(\vec{S})\mathbf{J}^T\mathbf{C}(\vec{D})^{-1}(\vec{D} - \mathbf{J}\vec{S}_0)$$

$$\mathbf{C}(\vec{S})^{-1} = \mathbf{C}(\vec{S}_0)^{-1} + \mathbf{J}^T\mathbf{C}(\vec{D})^{-1}\mathbf{J}$$

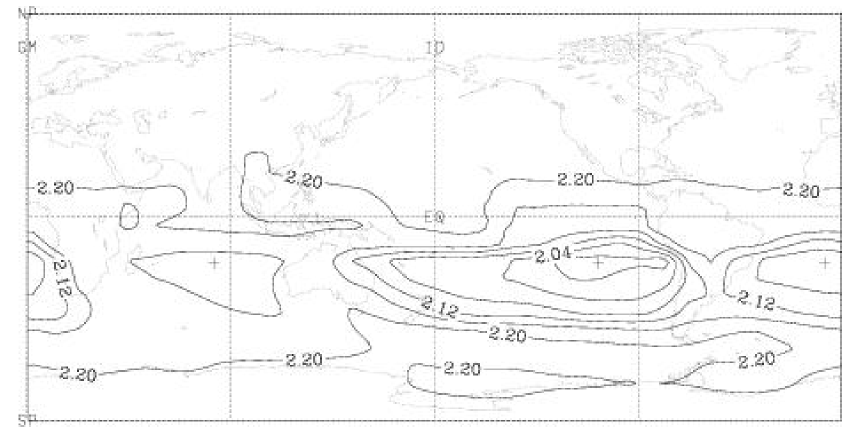
Recipe for network optimization

- Borrowed from geophysics
- Choose some property of $\mathbf{C}(\vec{S})$
- Manipulate $\mathbf{J}$ e.g. choosing sampling locations
- Use nonlinear minimization to optimize

Comparing Quality Measures

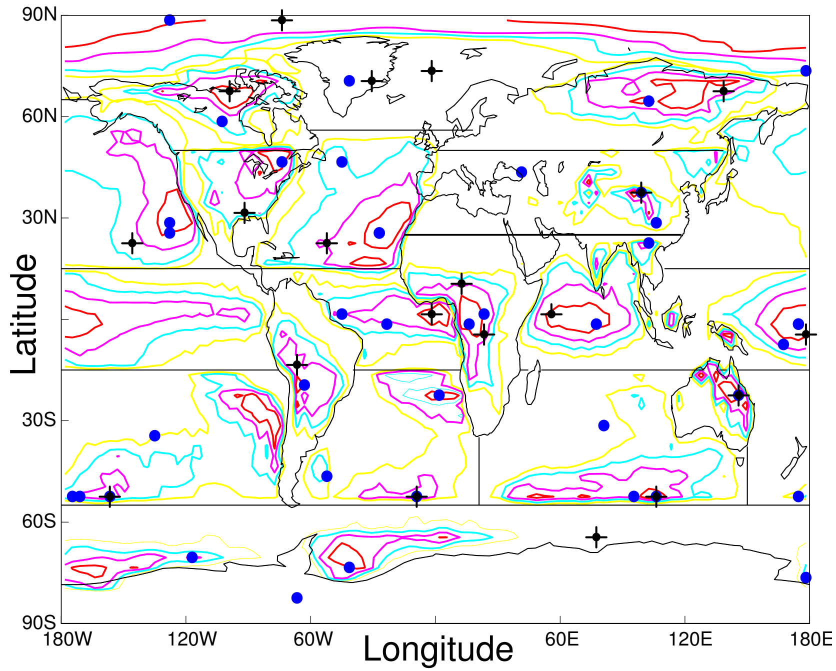


$$\mathrm{var}\left(\sum_1^N f_{O,i}\right)$$

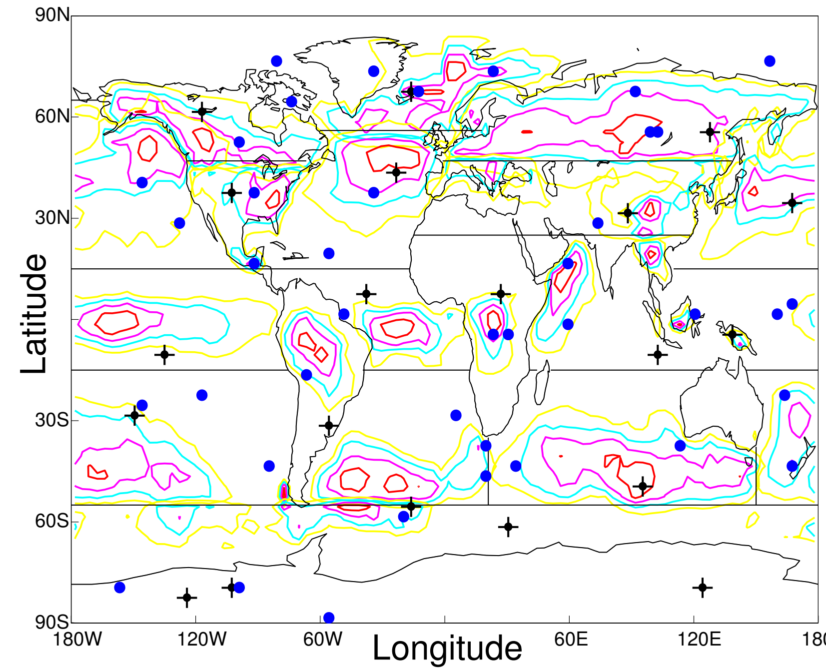


$$\sum_1^N \mathrm{var}(f_{O,i})$$

Comparing Error Models



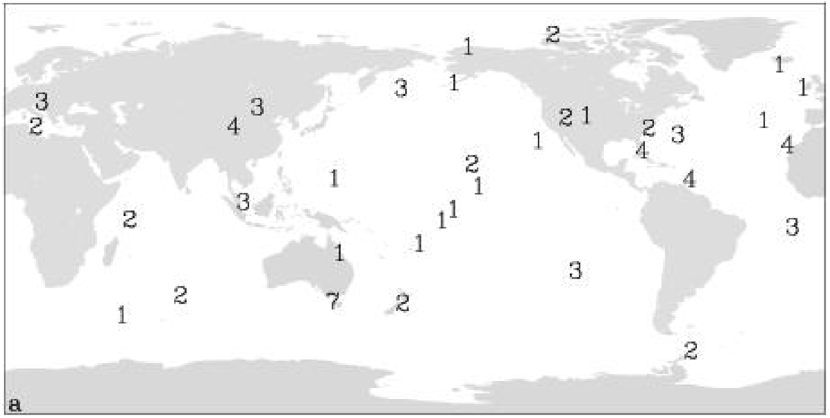
Optimal network in the case where all data uncertainties are equal



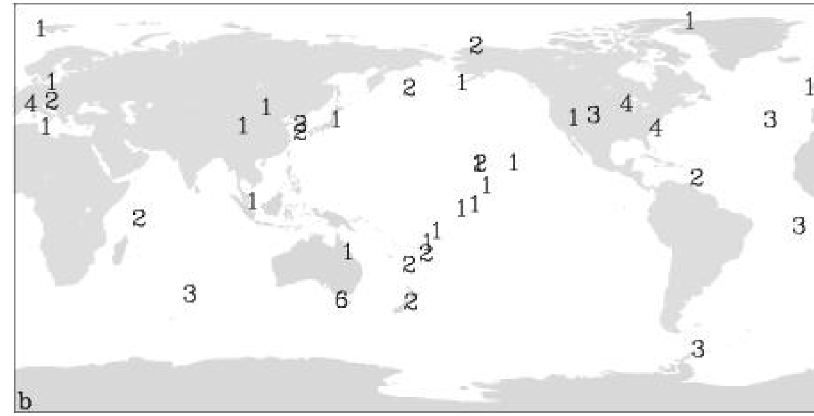
Optimal network in the case where data uncertainties vary according to simulated variability

Impact of Model Error

- Rayner, ACP 2004
- Transcom repeats inversions with different models
- Minimize uncertainty plus model spread
- Need real data for this
- Genetic algorithm
- Choose 76 stations of possible 110



Optimized networks considering uncertainty alone. Numbers indicate the number of times a site was included.

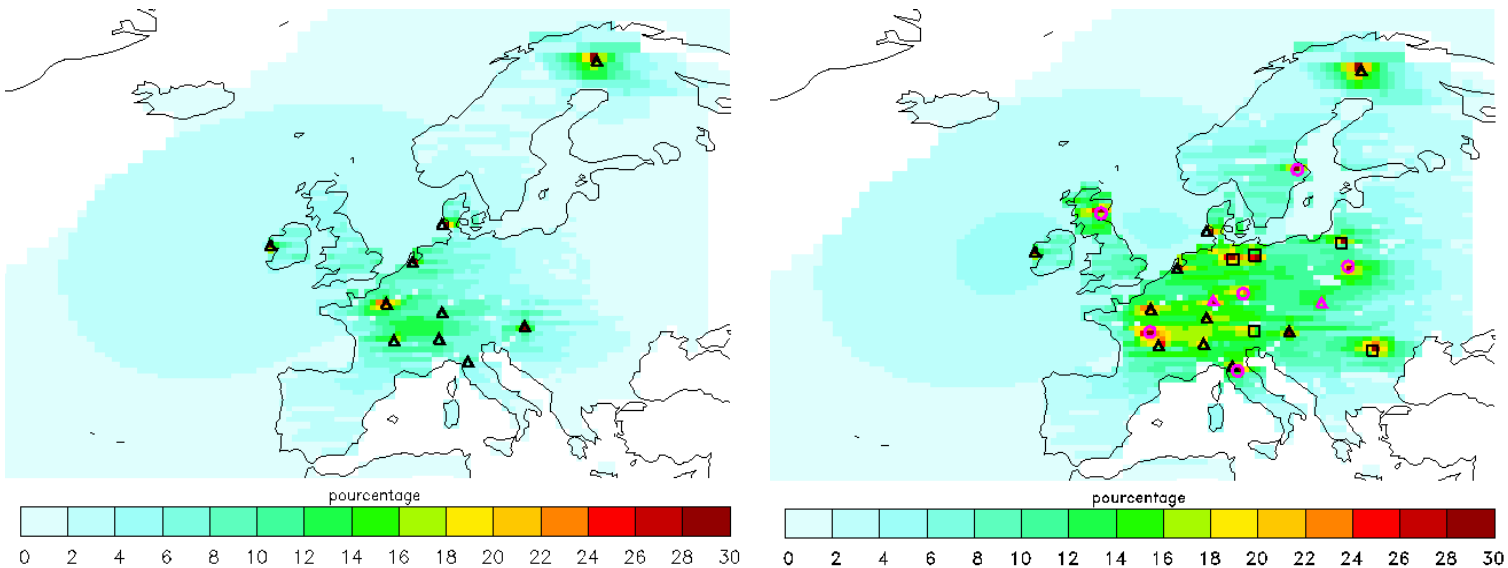


Optimized networks considering uncertainty plus model spread. Numbers indicate the number of times a site was included.

Potential of the European Network

Present error reduction

Future error reduction

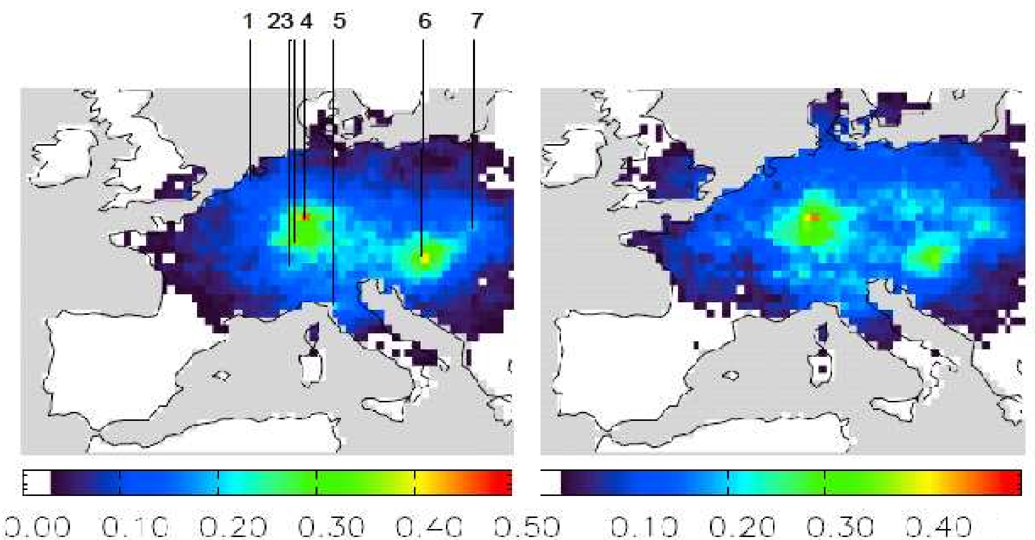


Carouge thèse

(Carouge et al., 2008) Error reduction for current (left) and enhanced (right European network. Reduction of error concentrated around stations. Requires prior correlations to achieve even this.

Using Higher Frequency Data

Nouvelle réduction d'erreur sur les flux , moyenne de l'été 2003



<u>D'Ouest en Est</u> :

1-CBW, 200m

2-PRS, 3485m

3-SCH, 1205m

4-HEI, 116m

5-CMN, 2165m

6-HUN, 363m

7-KAS, 1987m

<u>TEMOIN</u> : Erreur indexée sur la respiration du modèle ORCHIDE [Chevallier et al., 2006]

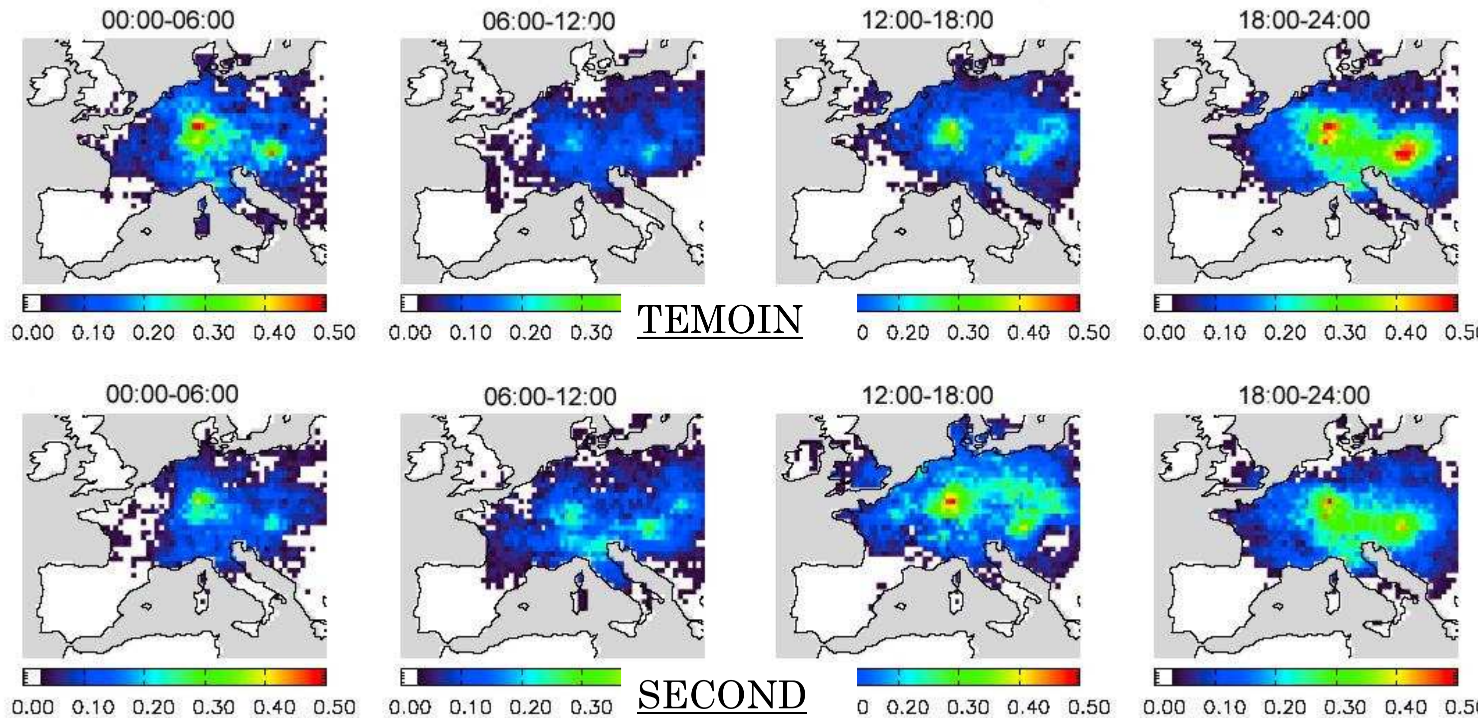
<u>SECOND</u> : Erreur indexée sur la dev. std. entre 10 modèles de flux [Rodenbeck et al., 2003]

Céline Aulagnier

PhD thesis. Reduction of error on 6-hourly fluxes using hourly data with network as of 2003

More with High Frequency Data

Nouvelles réductions d'erreur sur les flux, inégalités jour / nuit

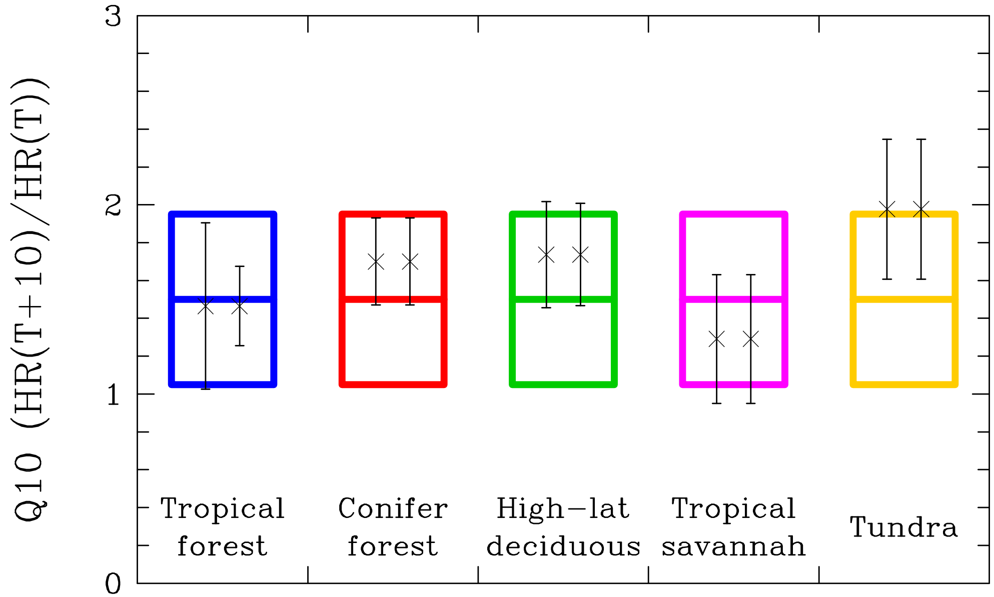


Reduction for Different periods of the day showing sensitivity to transport

Network Design in a Data Assimilation Context

- CCDAS combines different kinds of observations
- Must consider representativeness of measurements along with all previous caveats
- Network Optimization will be model dependent
- More complex optimization since we can trade off different measurement types as well as locations

Optimized Q_{10}



Predicted Q_{10} with (right bar) and without (left bar) the addition of a simulated flux station measurement in the broadleaf evergreen biome. The horizontal line shows the prior estimate, the box the 1 standard deviation of prior uncertainty, and the cross the optimized value. The vertical bar shows the one standard deviation confidence interval of the predicted value.

Lessons

1. Define your question carefully (multi-criterion optimisation helps)
2. Optimized network depends on prior knowledge
3. Optimized network depends on error model
4. Also depends on observation operator
5. Different kinds of uncertainty will generate different networks
6. Can work in a CCDAS context but model dependent