
Bayesian error covariance estimates in variational retrieval algorithms.

Johannes K. Nielsen and Stig Syndergaard

ROM SAF DMI

Melbourne April 2015

The well known basic problem of variational methods:

$$J(\mathbf{x}) = \frac{1}{2}(\mathbf{x} - \mathbf{x}^b)^T \mathbf{B}_x^{-1}(\mathbf{x} - \mathbf{x}^b) + \frac{1}{2}(\mathbf{y}^o - \mathbf{H}(\mathbf{x}))^T \mathbf{O}_y^{-1}(\mathbf{y}^o - \mathbf{H}(\mathbf{x}))$$

The problem is: What are **O** and **B** actually? - And what about biases?

Desroziers relations puts some restriction on errors

Notation in refractivity space:

$$\mathbf{o} = \mathbf{y}_o, \mathbf{b} = H(\mathbf{x}_b), \mathbf{O} = \mathbf{O}_y \text{ and } \mathbf{B} = \mathbf{H}\mathbf{B}_x\mathbf{H}^T$$

Desroziers et al. (2005): $\langle (\mathbf{o} - \mathbf{b})(\mathbf{o} - \mathbf{b})^T \rangle = \mathbf{B} + \mathbf{O}$, and 3 more (redundant) equations.

- ▶ Only constrains the sum of the error covariances. One must assume that either background noise or observation noise are known, to fix \mathbf{O} and \mathbf{B} .
- ▶ Implies, of course, that IF one has either the \mathbf{O} or \mathbf{B} matrix, both matrices can be found.

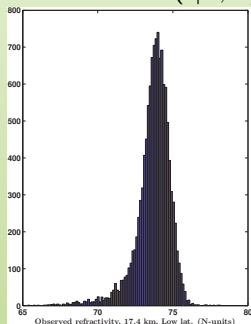
The error covariances are hidden in the data

The model

$$\mathbf{o} = \mathbf{t} + \omega + \varepsilon_o, \varepsilon_o \sim N(\varepsilon_o | \mathbf{0}, \mathbf{O})$$

$$\mathbf{b} = \mathbf{t} + \beta + \varepsilon_b, \varepsilon_b \sim N(\varepsilon_b | \mathbf{0}, \mathbf{B})$$

Where $\mathbf{t} \sim N(\mathbf{t} | \mathbf{0}, \mathbf{T})$, and $\omega, \beta, \varepsilon_o$ and ε_b are vectors.



Example of refractivity histogram

We shall not assume anything

The likelihood, of observations \mathbf{o} and \mathbf{b} , given the true values:

$$L(\mathbf{o}, \mathbf{b} | \mathbf{t}, \omega, \beta, \mathbf{O}, \mathbf{B}) = \frac{1}{(2\pi |\mathbf{O}|^{\frac{1}{2}} |\mathbf{B}|^{\frac{1}{2}})^N} \times \\ \exp\left\{-\frac{1}{2} \sum_{n=1}^N [(\mathbf{b}_n - \mathbf{t}_n)^T \mathbf{B}^{-1} (\mathbf{b}_n - \mathbf{t}_n) + (\mathbf{o}_n - \mathbf{t}_n)^T \mathbf{O}^{-1} (\mathbf{o}_n - \mathbf{t}_n)]\right\}$$

Posterior:

$$P(\mathbf{O}, \mathbf{B} | \mathbf{o}, \mathbf{b}, \mathbf{t}, \omega, \beta) \propto \\ L(\mathbf{o}, \mathbf{b} | \mathbf{t}, \omega, \beta, \mathbf{O}, \mathbf{B}) p(\omega) p(\beta) p(\mathbf{t} | \mathbf{T}) p(\mathbf{T}) p(\mathbf{O}) p(\mathbf{B}),$$

where

$$p(\mathbf{t} | \mathbf{T}) = \frac{1}{(2\pi |\mathbf{T}|)^N} \exp\left\{-\frac{1}{2} \sum_{n=1}^N (\mathbf{t}_n)^T \mathbf{T}^{-1} (\mathbf{t}_n)\right\}$$

and $p(\mathbf{O})$, $p(\mathbf{B})$, $p(\mathbf{T})$, $p(\omega)$ and $p(\beta)$ are chosen to be flat priors.

Example of sampling, $O_{30,30}$ and $B_{30,30}$

Rather than performing the integrals

$$P(\mathbf{O}) = \int dt d\mathbf{B} d\omega d\beta$$

$P(\mathbf{O}, \mathbf{B} | \mathbf{o}, \mathbf{b}, \mathbf{t}, \omega, \beta)$, we find the marginal distribution of the parameters with a Markov Chain Monte Carlo algorithm, that samples $\mathbf{t}, \omega, \beta, \mathbf{O}$ and \mathbf{B} .

Rather than providing a mathematical prove:

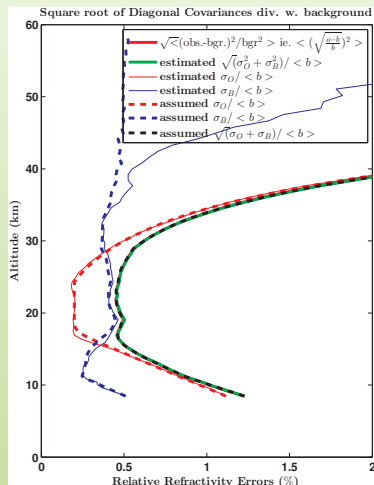
You give me two data series

$$o = t + \omega + \varepsilon_o, \varepsilon_o \sim N(\varepsilon_o | 0, \sigma_o^2)$$

$$b = t + \beta + \varepsilon_b, \varepsilon_b \sim N(\varepsilon_b | 0, \sigma_b^2)$$

Where $t \sim N(t | 0, \sigma_t^2)$ Then I give you σ_o^2, σ_b^2 and $\omega - \beta$

Covariance diagonals (surrogate data with bias)

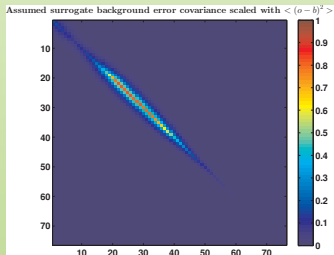
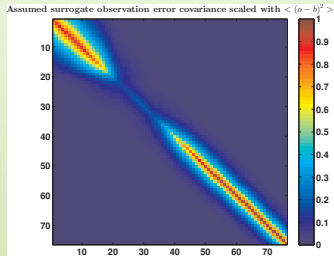


Two artificial data sets “observation” and “background” with known error characteristics are produced from 2000 true refractivity profiles (Metop A and B, latitude < 30 deg.):

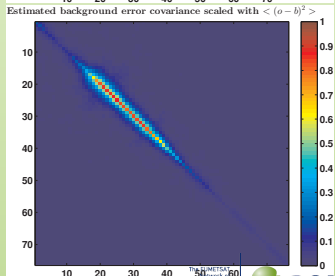
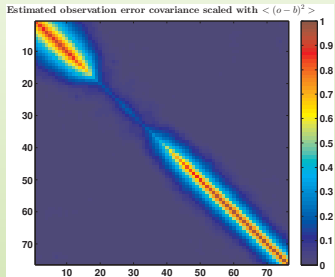
$o = t + \varepsilon_o$ and $b = t + \varepsilon_b$,
where ε_o and ε_b are noise terms. A Bayesian estimate of the two error covariance matrices is obtained by a MCMC method

Finding **O** and **B** matrices

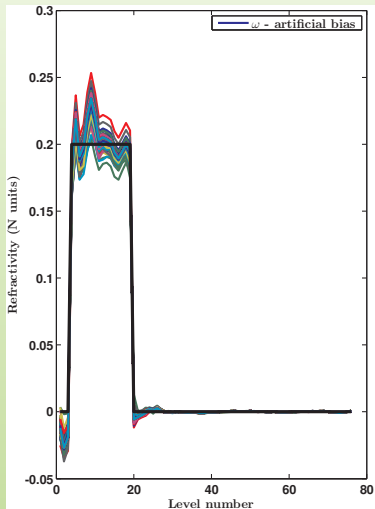
Surrogate



Estimated with MCMC

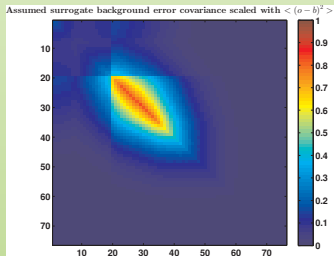
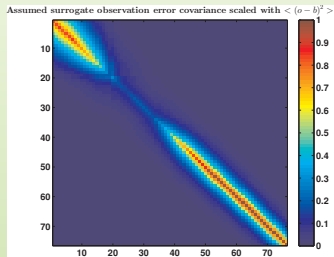


Finding bias

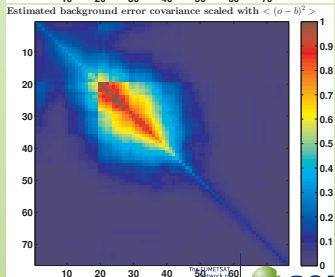
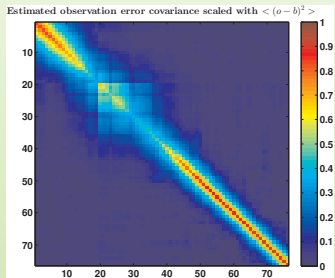


Finding **O** and **B** matrices from biased data

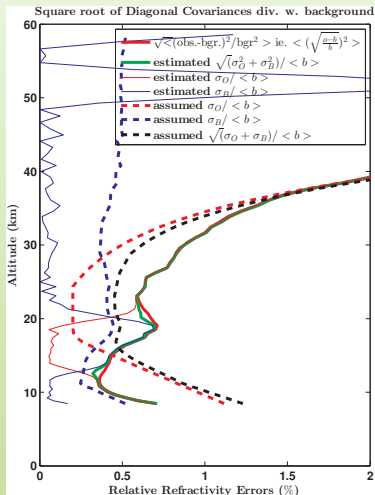
Surrogate



Estimated with MCMC



Finding real covariances



Conclusions

- ▶ Observation error covariances can be obtained from two data series measuring the same property.
- ▶ Alternative method for error estimates.
- ▶ Alternative method for validation.
- ▶ Problems with (inter) correlations and too large correlations.
- ▶ It is certainly not “plug and play”.