Prioris, posteriors and uncertainty

Johanna Tamminen

`johanna.tamminen@fmi.fi`

ESA Summer School on
Earth System Monitoring and Modeling
July 30 – Aug 11, 2012, Frascati

August 1, 2012

Topics of todays lecture

- Example what a good prior could do
- Uncertainty and error characterization:
    - random errors
    - systematic errors
    - uncertainties in model selection
- Analysing retrievals: residuals and χ^2
- Validating operational retrievals using MCMC
- Examples in atmospheric remote sensing

Vertical inversion

The *vertical inversion* is the problem

$$N_\ell^{\text{gas}} = \int\limits_\ell \rho^{\text{gas}}(z(s)) ds, \qquad \ell = \ell_1, \ldots, \ell_M.$$

for the gas profiles $\rho^{\text{gas}}(z)$. By discretizing the atmosphere into layers the problem can be solved separately for each gas as a linear inversion problem

$$N^{\text{gas}} = A\rho^{\text{gas}}, \qquad N^{\text{gas}} = (N_\ell^{\text{gas}}), \ \ell = \ell_1, \ldots, \ell_M$$

In so called Onion peeling formulation local spherical symmetry of the atmosphere is assumed.

The matrix *A* contains the lengths of the line of sight in the layers and depends on the discretization. In the operational retrieval the discretization is fixed so that the number of layers is the same as the number of measurement lines in each occultation.

$$A = \begin{bmatrix} a_{11} & & & \\ 2a_{21} & a_{22} & & \\ 2a_{31} & 2a_{32} & a_{33} & \\ \cdots & & & \end{bmatrix}$$
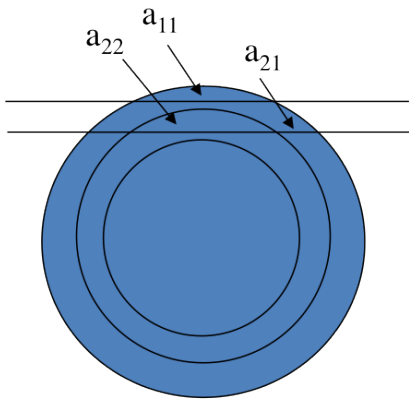
Tikhonov regularization - smoothness prior

In the case of GOMOS, the direct inversion produces vertical profiles accompanied with unphysical oscillations. In order to remove these, the inversion is performed using the Tikhonov regularization method. The classical Tikhonov regularized solution of the problem is a minimizer of the functional:

$$F(\mu) = \|A\rho - N\|^2 + \mu\|H\rho\|^2 \tag{1}$$

Here μ is the regularization parameter.

A modified version of Tikhonov regularization with smooth second derivatives is applied in GOMOS retrieval. The regularisation matrix H is

$$H = \mathrm{diag}\left[\frac{1}{h_i^2}\right]\begin{bmatrix} 0 & 0 & 0 & ... & 0 \\ 1 & -2 & 1 & ... & 0 \\ ... & .. & ... & ... & ... \\ 0 & ... & 1 & -2 & 1 \\ 0 & 0 & ... & 0 & 0 \end{bmatrix} \tag{2}$$

where h_i denotes local altitude differences.

Note: Tikhonov regularization can be seen as Bayesian solution resulting the MAP estimate given the data and a prior described by μ and H.

Vertical inversion cont.

The Tikhonov regularized solution of is given by

$$\widehat{\rho} = (A^T A + \mu H^T H)^{-1} A^T N \qquad (3)$$

Since the problem is linear and the noise is assumed to be Gaussian the posterior covariance matrix is obtained by standard matrix computation:

$$C_\rho = L C_N L^T$$

where L denotes the retrieval matrix $\widehat{\rho} = LN$.

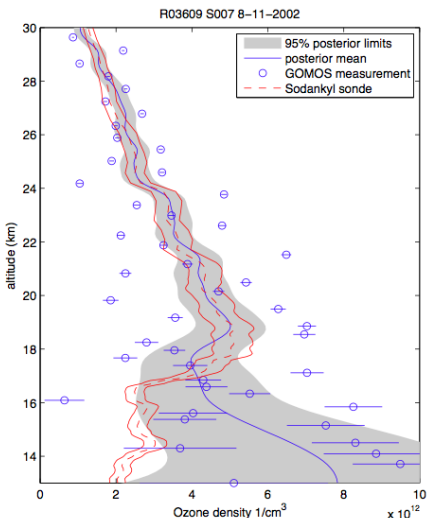# Example: GOMOS vertical inversion – prior information

- PRIOR 1: Flat (non-informative prior). Retrieval - blue circles.
- PRIOR 2: Based on earlier measurements it has been found that the vertical profiles ozone themself vary, but they are rather smooth. Retrieval - blue solid line and gray area around it.
- Retrieval using PRIOR 2 results in ozone profile which corresponds well to independent ozone sounding measurements (red dashed line)


R03609 S007 8-11-2002

95% posterior limits
posterior mean
GOMOS measurement
Sodankyl sonde

altitude (km)

Ozone density 1/cm³   x 10¹²

Uncertainties in practice

- Random errors
  - caused by measurement noise through error propagation
  - caused by random features not modeled correctly
  - typically considered as 'easy' errors as they cancel out when a lot of measurements are averaged
- Systematic errors
  - assumptions
  - approximations
  - simplifications in modeling
  - model uncertainty
  - Typically condered as 'difficult' errors since they can cause bias in the data.
  - However, if only trends are studied, (small) stable systematic errors may not be crucial
- Mixtures of random and systematic errors
- Both random and systematic errors can depend on geographical conditions, measurement geometry, solar angles, ...

Uncertainties in practice (cont.)

- Time evolution in errors

    - Aging of the instrument: measurement noise typically increasing

    - Particularly important in remote sensing since calibration more complicated

    - Drifting orbit may cause changes in errors.

Aging of the instrument - example
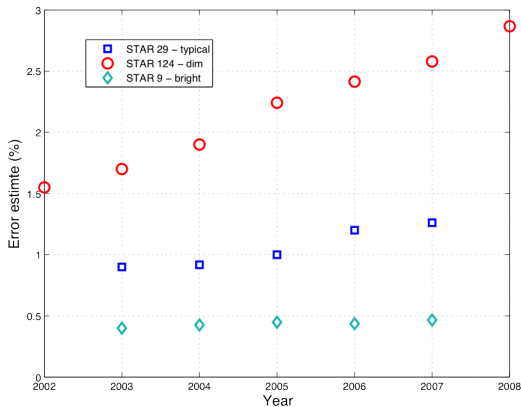
Due to GOMOS instrument aging the measurement noise is increasing. The aging is seen in the error estimates

Example: uncertainties in GOMOS data

- Random errors:
  - Measurement noise
  - Uncorrected scintillations
- Systematic errors:
  - Uncertainties in aerosol modeling
  - Uncertainties in cross sections
  - Uncertainty in temperature and neutral density
  - Uncertainties in ray tracing

Methods of studying the uncertainties and validating the error estimates

- Error propagation
  - Analysing residuals
  - Analysing χ^2 values
  - Non-linearity can be studied using MCMC

- Off line sensitivity studies: simulations using varying assumptions or initial conditions.

- Off line processing with detailed forward model.

- Off line processing with extendid inversion: characterizing modeling uncertainties in the retrievals

- Geophysical validation - comparison with independent measurements

Useful diagnostics: analysing residuals and χ^2

- Residual: $(f_i(x) - y_i)$ should be of the same size as the noise. It should not show systematic behaviour if noncorrelated noise is assumed.

- $\chi^2 = (f(x) - y)C_y^{-1}(f(x) - y)^T$ should be ~ 1 in close to Gaussian case.
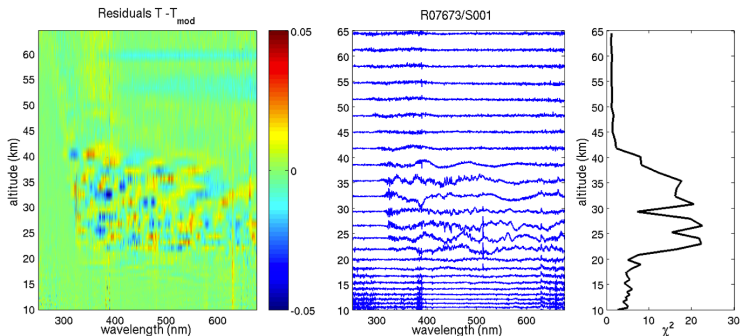
Example of GOMOS residuals and χ^2 in old Version 5 processing. Note high values of χ^2 and systematic features in residuals.
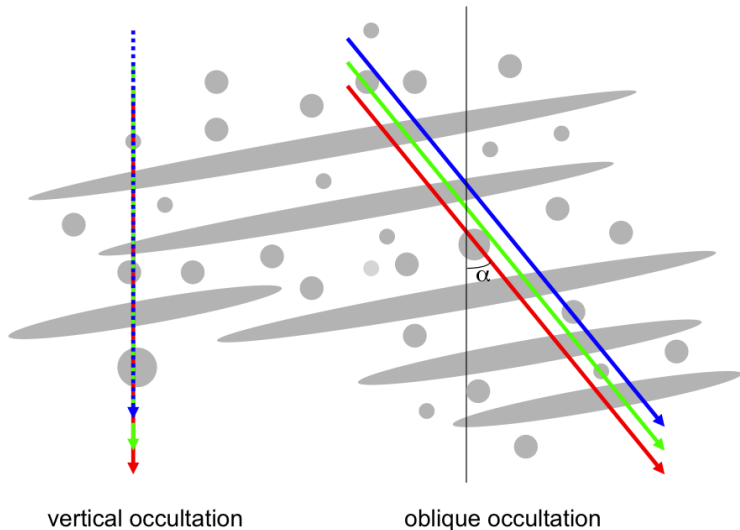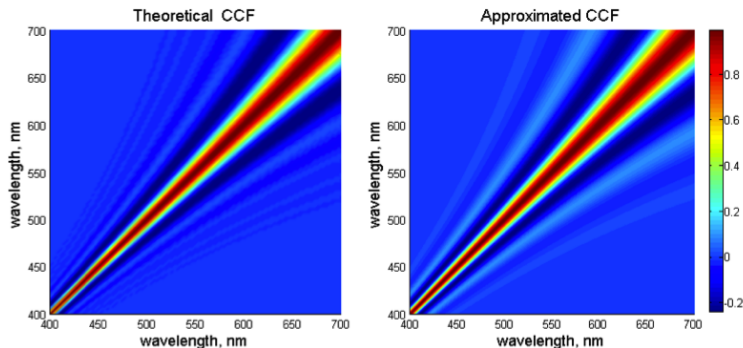
Interesting finding: Small scale structures in the atmosphere

vertical occultation   oblique occultation

Figure by V. Sofieva (FMI)

Correlated modeling error

- In GOMOS version 6 Spectral inversion correlated modeling error C_m was included:

$$C_y := C_y + C_m$$

where C_y was diagonal and C_m band matrix describing the modeling error and its correlation in wavelength.

Useful diagnostics: analysing residuals and χ^2

- The retrieval become much slower but error characterization was improved.

Sirius, obliquity >30°
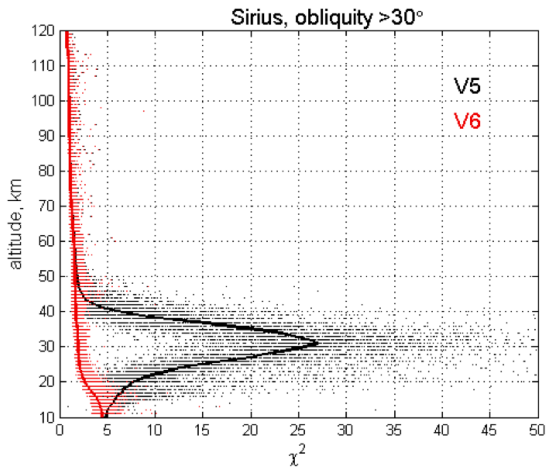
GOMOS validation: operational vs MCMC

Motivation

- Validate the assumption of solution to be close to Gaussian
- Different prior and noise structures can be easily included
    - Studying different a priori information (positivity priors: nonsymmetric posteriors)
    - Example with noisy data, outliers: estimation with robust ℓ_1 norm may yield to a more stable results (example yesterday).
- Modelling errors:
    - Include in the retrieval as nuisance parameters: wider posteriors
    - Model selection

Examples of sampled posterior distribution

- MCMC sampling in ∼5 dimensional space in GOMOS spectral inversion. Unknowns: horizontal densities of O3, NO2, NO3, aerosols, (neutral density)

- Visually we can analyse 2-dimensional marginal distributions and their correlations.
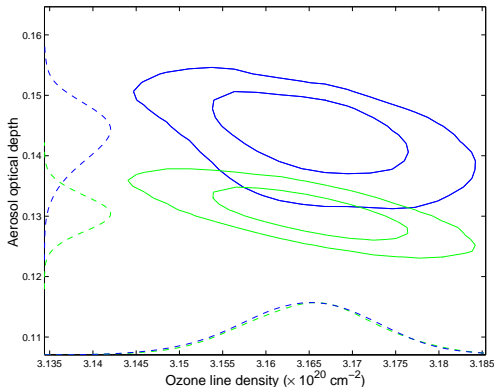
How Gaussian is the posterior distribution?

- The posterior distribution of GOMOS Spectral inversion computed with MCMC and the Gaussian estimate (point estimate and covariance matrix) computed using operational Levenberg-Marquard algorithm practically coincide if no additional information is used.
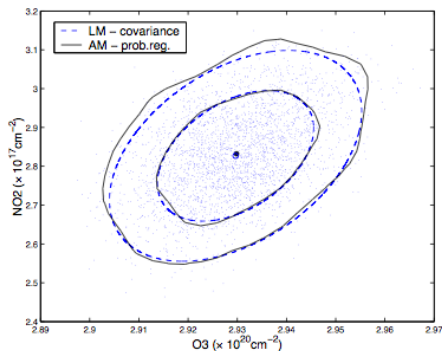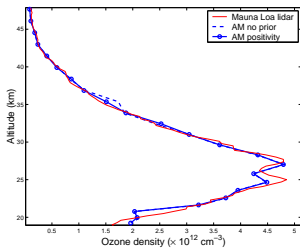
Positivity prior
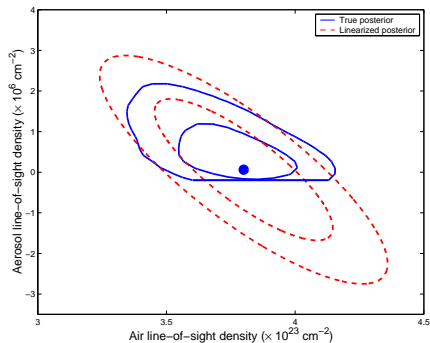
- Assuming positivity:

$$p(N \mid y) \propto \begin{cases} \exp(-\frac{1}{2} \sum_{\lambda=1}^{\Lambda} \frac{(T_\lambda(N) - y_\lambda)^2}{s_\lambda^2}) & \text{if } N_j > 0 \\ 0 & \text{otherwise} \end{cases}.$$

- Without constraints some gases may get too small (negative), others too large values.
- True nonsymmetric posteriors easily by MCMC.

Positivity prior cont.

- Improvement in ozone retrieval when positivity prior introduced.
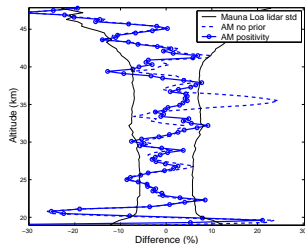
Positivity prior cont.

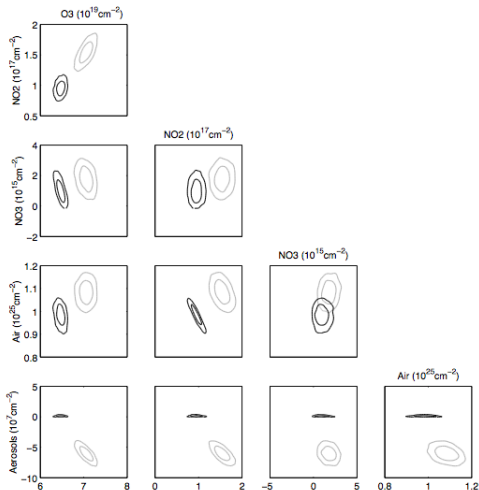- Unrealistic aerosol density causes bias for other constituents.

Figure 6. Two-dimensional marginal posterior distributions, from left to right, line-of-sight densities of: O_3, NO_2, NO_3 and air and from top to bottom: NO_2, NO_3, air and aerosols. The contour curves denote 68.3 and 95 % probability regions. Gray contours shows the result without any prior knowledge and black contours assuming positivity. GOMOS measurement at 35 km, dim star.

Modelling error: uncertainty in temperature
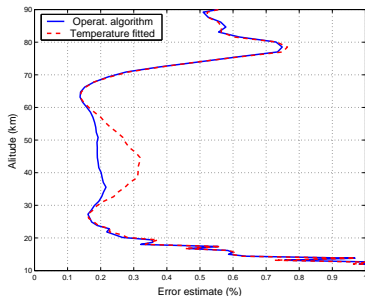
- Uncertainties in temperature: use tempereture as a nuisance parameter b, sample $p(x, b \mid y)$ to obtain:

$$p(x \mid y) = \int p(x, b \mid y) p(b \mid y) b.$$

- The error estimates of ozone are clearly increased by assuming uncertainty in temperature (Gaussian with 2 K deviation)
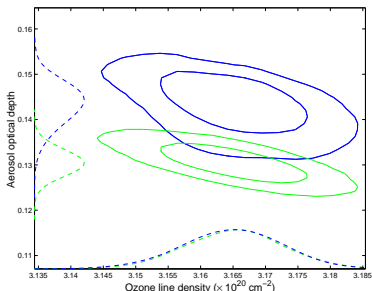
Modelling errors: aerosol size distribution

Simple aerosol model:

$$\tau_{\mathrm{aero}}(\lambda) = N_{\mathrm{aero}}\alpha_{\mathrm{aero}}(\lambda).$$

where $\alpha_{aero}(\lambda) = \alpha_0/\lambda^\beta$ and a choice $\alpha_0 = 3 \times 10^{-7} \mathrm{cm}^2$, $\beta = 1$ is applied

Including uncertainty about Angstrom coefficient to transmission model:

$$T_\lambda(N, \beta) = \exp\left(-N_{\mathrm{aero}}\frac{\alpha_0}{\lambda_\lambda^\beta} - \sum_{j=1}^{\mathrm{n\ of\ gases}} N_j\alpha_{j,\lambda}\right).$$

- Uncertainty modelled again using nuisance parameter.
- Uncertainty of aerosol cross sections increases in particular the uncertainty of aerosol optical depth.

# Sensitivity study related to GOMOS aerosol model

- Uncertainty in modeling wavelength dependence of aerosols (related to aerosol type)

- Sensitivity stydy performed for a test data set of about 1000 occultations made at different geographical locations using seven different aerosol models.

- Averaged ozone profiles show that there is a large systematic difference in the ozone profiles depending on the aerosol model used (up to 10-30%) below 20 km.
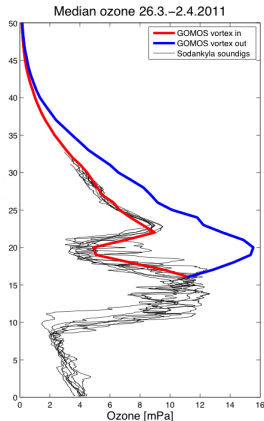

O3 relative difference
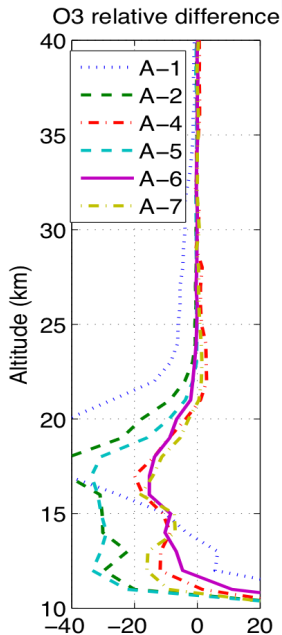
Arctic ozone depleation in 2011: ozone measured by OMI/Aura and GOMOS/Envisat

Example: improving ozone at low altitudes

- Introducing smoothnes prior to aerosols. Application of 'one-step' algorithm (together spectral and vertical inversion).
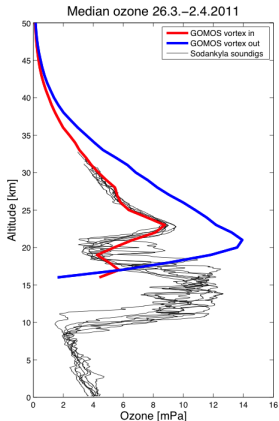
Operative algorithm   Improved UTLS-algorithm

Variants of Adaptive Metropolis MCMC algorithm

Variants of AM

- SCAM: Single component AM algorithm (Haario et. al 2004)
  - High dimensions!
- DRAM: Delayed rejection AM algorithm (Haario et al. 2006)
  - Robust adaptation!
- AARJ: Adaptive Automatic Reversible Jump MCMC (Laine and Tamminen, 2009)
  - Practical model selection!

SCAM algorithm

- Single component AM algorithm

The i:th coordinate X_t^i ($i = 1, \ldots, d$) of the t:th state X_t is obtained by 1-dimensional Metropolis step:

Step 1: Sample y^i from 1-dimensional normally distributed proposal distribution q_t centered at X_{t-1}^i with variance C_t^i. After an initial period,

$$C_t^i = s_1 \operatorname{Var}(X_0^i, \ldots, X_{t-1}^i) + s_1 \varepsilon.$$

Step 2: Accept the candidate point y^i with probability

$$\min\left(1, \frac{\pi(X_t^1, \ldots, X_t^{i-1}, z^i, X_{t-1}^{i+1}, \ldots, X_{t-1}^d)}{\pi(X_t^1, \ldots, X_t^{i-1}, X_{t-1}^i, \ldots, X_{t-1}^d)}\right),$$

in which case set $X_t^i = y^i$, otherwise $X_t^i = X_{t-1}^i$.

High-dimensional one-step inversion with MCMC: SCAM

- Free of restrictive assumptions on model or data
- A direct solution for the gas profiles leads to a problem with 70000–140000 measurements and 250–500 unknowns.
- AM converges slowly in high dimensional problems, but Single component adaptive Metropolis algorithm (SCAM) works well.
- SCAM combines the ideas of componentwise sampling (Gibbs sampling) and AM algorithm

SCAM applied to GOMOS 'one-step' problem (simultaneus spectral and vertical inversion

- Test case: dimension of the problem 90.

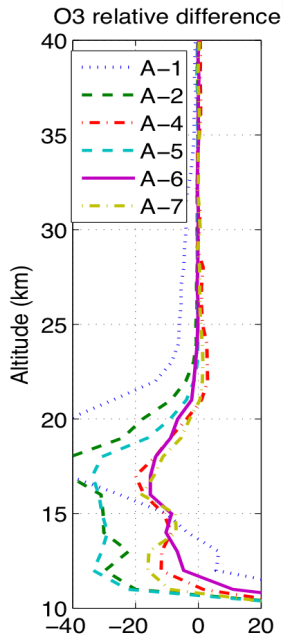- Full view of the correlations of btw gases and altitudes.
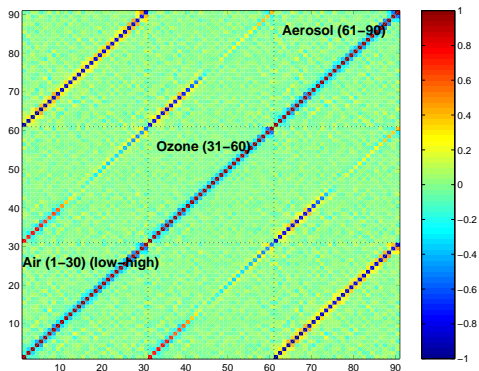
Fig. Correlation coefficients of the error structure by 'one-step' inversion and SCAM.

...Back to uncertainty of GOMOS aerosol model

O3 relative difference

Model selection

- It is also possible that we do not know which model is correct/best
- In such a case we need to perform model selection
- It would be good if the model selection could be done simultaneusly with solving the inverse problem.

• Bayesian model selection can be applied using eg. MCMC algorithms called Reversible jump MCMC (Ref. Green 1995) which jumps between different models

MCMC method for model selection I

Model selection using automatic adaptive Reversible Jump MCMC (AARJ)
(ref. Laine and Tamminen 2009).

- AARJ = Automatic RJMCMC by
  Green (2003, 1995) with AM
  style adaptation.

- Uses Gaussian approximations
  of the target distributions to
  perform model to model
  parameter transformations.

- Both the target approximations
  and proposal covariances are
  adapted.

- First approximations by initial
  runs.
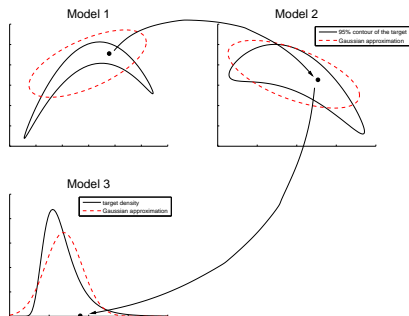
Suitable for model selection and model averaging problems with 2-10
competing models.

Aerosol model selection in GOMOS retrieval

- Four different aerosol models considered:

1. $\alpha(\lambda) = a_0 \lambda_{\mathrm{ref}}/\lambda$
2. $\alpha(\lambda) = a_0 + \tilde{a}_1(\lambda - \lambda_{\mathrm{ref}}) + \tilde{a}_2(\lambda - \lambda_{\mathrm{ref}})^2$
3. $\alpha(\lambda) = a_0 \lambda_{\mathrm{ref}}^2/\lambda^2$
4. $\alpha(\lambda) = a_0 + \tilde{a}_1(1/\lambda - 1/\lambda_{\mathrm{ref}}) + \tilde{a}_2(1/\lambda - 1/\lambda_{\mathrm{ref}})^2$

The model is parameterized so that constant term a_0 gives the optical extinction when $\lambda_{\mathrm{ref}} = 500$ nm and parameters a_1 and a_2 correspond to wavelengths $\lambda_1 = 300$ nm and $\lambda_2 = 600$ nm.
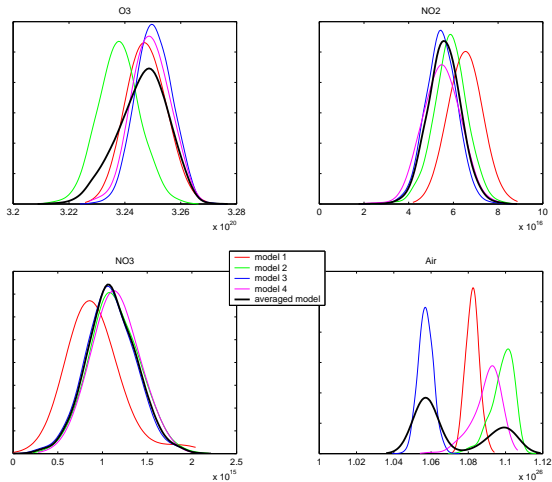
- Simultaneusly as the horizontal has densities are fitted also different aerosol models are sampled.

- MCMC sampling now done in 4 different state spaces (with different dimensons).
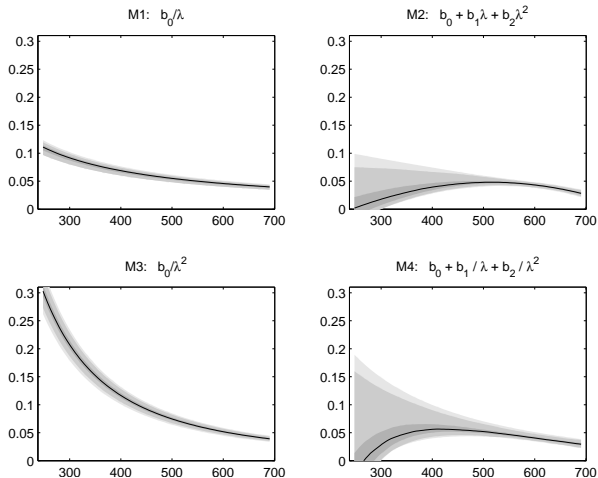
RJMCMC chains and model probabilities

Four different aerosol models fitted with AARJ algorithm.
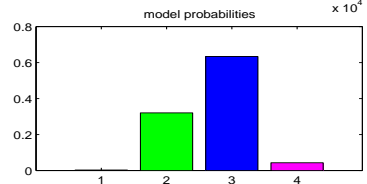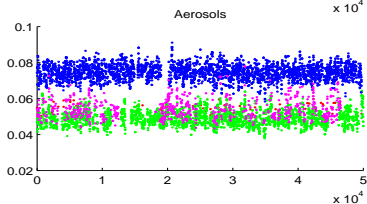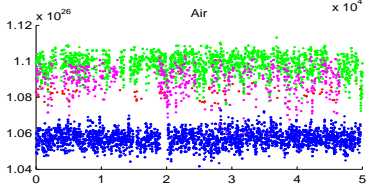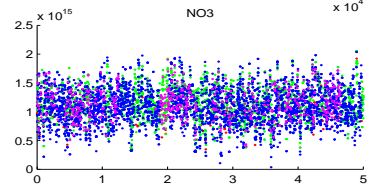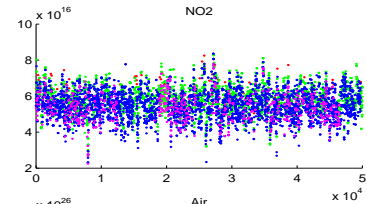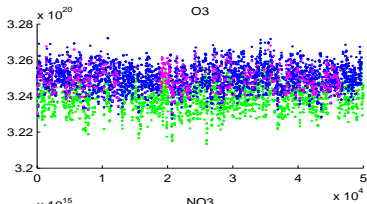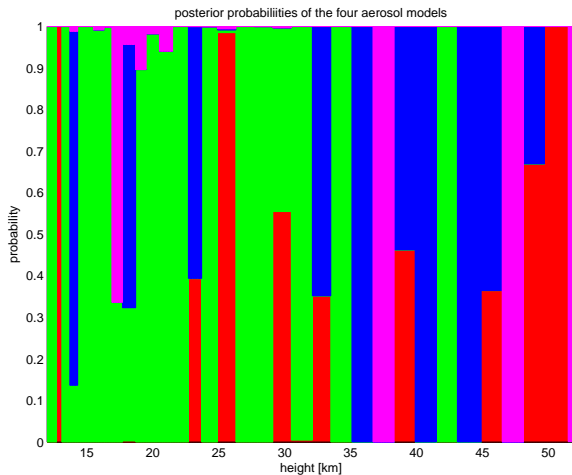
1d Posterior distributions

Predictive distributions of the wavelength dependence

Models and their posterior uncertainty.

Posterior probability at different heights

posterior probabiliities of the four aerosol models

Aerosol model selection in OMI

- OMI - Ozone Monitoring Instrument

- Dutch-Finnish Instrument onboard NASA's EOS-Aura Satellite, launched in 2004.

- ESA third party mission

- UV-VIS instrument measures: O3, NO2, SO2, aerosols, UV-radiation, ...

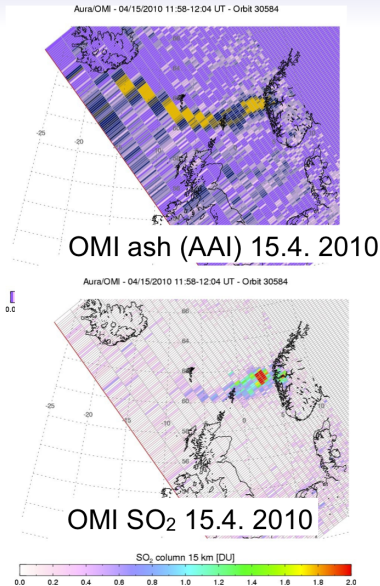- Global coverage in one-two days

- Pixel size at nadir 24×13 km.
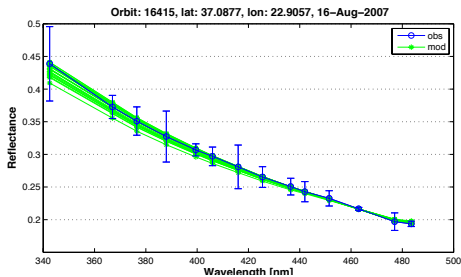

OMI ash (AAI) 15.4. 2010
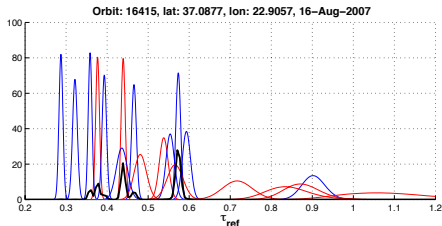

OMI SO$_2$ 15.4. 2010

Fig. by NASA

Aerosol model selection in OMI

- Different models give different optical depth values

- The mean value is multimodal.

- It looks like the errors do not correctly reflect the uncertainty in the data.

Orbit: 16415, lat: 37.0877, lon: 22.9057, 16−Aug−2007

Orbit: 16415, lat: 37.0877, lon: 22.9057, 16−Aug−2007

Aerosol model selection in OMI

- The model discrepancy can be taken into account using Gaussian processes.

- Results look more realistic. Mean value wide and unimodal.

- The agreement with models seem to reflect the uncertainty in the data.
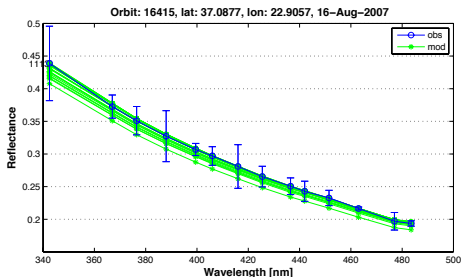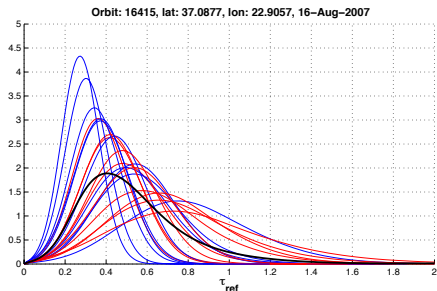
(From A. Määttä et al., in preparation 2012)

Summary: Uncertainties in remote sensing

Large amounts of data and complex models

- Important to understand and characterize uncertainties:
    - The users of the data need to know the uncertainties.
    - Possibility to improve the retrievals
    - We may learn new things
- Bayesian approach gives natural tools to characterize uncertainties.
- Tools: Analysing retrievals, sensitivity analysis, alternative retrievals, full modeling of uncertainties using off line models, geophysical validation.
- MCMC is one option for helping in uncertainty quantification.

References I

- Tarantola, Albert (2005). *Inverse Problem Theory and Methods for Model Parameter Estimation*. SIAM. ISBN 978-0-89871-572-9. (Rewritten version of Tarantola, 1987)

- Clive D. Rodgers *Inverse Methods for Atmospheric Sounding: Theory and Practice*, World Scientific, 2000.

- Kaipio and Somersalo, Statistical and Computational Inverse Problems, 2005.

- Box and Tiao: Bayesian Inference in Statistical Analysis, 1992.

- Tikhonov, A. N.; Arsenin, V. Y. (1977). Solutions of Ill-Posed Problems. New York, Winston

- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E. (1953). *Equation of State Calculations by Fast Computing Machines*, The Journal of Chemical Physics, 21(6), pp. 1087–1092.

References II

- Hastings W, *Monte Carlo sampling method using Markov chains and their applications*, Biometrika, 57, 97–109, 1970.

- Markov chain Monte Carlo in practice, editors W. R. Gilks, S. Richarrdson and D. J. Spiegelhalter, 1995.

- Haario H., Saksman E., Tamminen J.: *An adaptive Metropolis algorithm*. Bernoulli 7(2) 223–242, 2001

- Haario H., M.Laine, M. Lehtinen, E. Saksman, J. Tamminen: *Markov chain Monte Carlo methods for high dimensional inversion in remote sensing*, with discussion, J.R. Statist. Soc. B, 66, part 3 pp. 591-607, 648–649, 2004.

- M. Laine, J. Tamminen: *Aerosol model selection and uncertainty modelling by adaptive MCMC technique*. Atmos. Chem. Phys. 2009

References III

- Green, P. J.: *Reversible jump Markov chain Monte Carlo computation and Bayesian model determination*, Biometrika, 82, 711–732, 1995.

- Green, P. J.: *Trans-dimensional Markov chain Monte Carlo, in: Highly Structured Stochastic Systems*, edited by: Green, P. J., Hjort, N. L., and Richardson, S., no. 27, Oxford Statistical Sci- ence Series, pp. 179-198, Oxford University Press, 2003.

- J. L. Bertaux et. al., *Global ozone monitoring by occultation of stars: an overview of GOMOS measurements on ENVISAT*, ACP, GOMOS Special Issue, 2010

- Kyrola. et al. *Retrieval of atmospheric parameters from GOMOS data*, ACP, GOMOS Special Issue, 2010

- Tamminen et. al. *GOMOS data characterisation and error estimation*, ACP, GOMOS Special Issue, 2010