Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

How much information do extinction and backscattering measurements contain about the chemical composition of atmospheric aerosol?

Michael Kahnert[1,2] and Emma Andersson[2]

[1]Research Department, Swedish Meteorological and Hydrological Institute, Folkborgsvägen 17, SE-601 76 Norrköping, Sweden
[2]Department of Earth and Space Science, Chalmers University of Technology, SE-412 96 Gothenburg, Sweden

Correspondence to: Michael Kahnert (michael.kahnert@smhi.se)

Abstract.

We theoretically and numerically investigate the problem of assimilating lidar observations of extinction and backscattering coefficients of aerosols into a chemical transport model. More specifically, we consider the inverse problem of determining the chemical composition of aerosols from these observations. The main questions are how much information the observations contain to constrain the particles' chemical composition, and how one can optimise a chemical data assimilation system to make maximum use of the available information. We first quantify the information content of the measurements by computing the singular values of the observation operator. From the singular values we can compute the number of signal degrees of freedom and the reduction in Shannon entropy. For an observation standard deviation of 10 %, it is found that simultaneous measurements of extinction and backscattering allows us to constrain twice as many model variables as extinction measurements alone. The same holds for measurements at two wavelengths compared to measurements at a single wavelength. However, when we extend the set of measurements from two to three wavelengths then we observe only a small increase in the number of signal degrees of freedom, and a minor change in the Shannon entropy. The information content is strongly sensitive to the observation error; both the number of signal degrees of freedom and the reduction in Shannon entropy steeply decrease as the observation standard deviation increases in the range between 1 and 100 %. The right singular vectors of the observation operator can be employed to transform the model variables into a new basis in which the components of the state vector can be divided into signal-related and noise-related components. We incorporate these results in a chemical data assimilation algorithm by introducing weak constraints that restrict the assimilation algorithm to acting on the signal-related model variables only. This ensures that the information contained in the measurements is fully exploited, but not over-used. Numerical experiments confirm that the constrained data assimilation algorithm solves the inverse problem in a way that automatises the choice of control variables, and that restricts the minimisation of the costfunction to the signal-related model variables.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

1 Introduction

Atmospheric aerosols have a substantial, yet highly uncertain impact on climate, they can cause respiratory health problems, degrade visibility, and even compromise air-traffic safety. The physical and chemical properties of aerosols play a key role in understanding these effects. The aerosol properties are determined by a complex interplay of different chemical, microphysical, and meteorological processes. These processes are investigated in environmental modelling by use of chemical transport models (CTMs). However, modelling aerosol processes is plagued by substantial biases and errors (McKeen et al., 2007). It is, therefore, fundamentally important to evaluate and constrain CTMs by use of measurements.

Measurements from satellite instruments provide consistent long-term data sets with global coverage. However, it is notoriously difficult to compare measured radiances to modelled aerosol concentrations. An alternative to using radiances is to make use of satellite retrieval products. For instance, one of the products of the CALIPSO lidar instrument (Cloud-Aerosol Lidar and Infrared Pathfinder Satellite Observations) is a rough classification of the aerosol types (i.e. dust, smoke, clean/polluted continental, and clean/polluted marine). This retrieval product is based on lidar depolarisation measurements (Omar et al., 2009). For the evaluation of aerosol transport models this provides us with a qualitative check for the chemical composition of aerosols. However, this is of limited practical use, since what we really need is quantitative information on the particles' chemical composition (which can be size-dependent). The most popular approach in evaluating and constraining aerosol transport models is the use of retrieved optical properties, such as aerosol optical depth, or extinction and backscattering coefficients. Yet another idea is to provide the particles' refractive index as a retrieval product (e.g. Mishchenko et al. (2007)). However, the use of such retrieval products still leaves us with the challenge of solving an ill-posed inverse problem, namely, of determining the particles chemical composition from their retrieved optical or dielectric properties.

A systematic class of statistical methods for solving this inverse problem is known as data assimilation. Recent studies have applied data assimilation to aerosol models with varying degrees of sophistication, ranging from simple dust models (Khade et al., 2013) and mass transport models (Zhang et al., 2014) to microphysical aerosol models based on modal (Rubin and Collins, 2014) or sectional descriptions (Sandu et al., 2005; Saide et al., 2013) of the aerosol size distribution. The assimilation techniques that have been used comprise variational methods, such as 2D (Zhang et al., 2014), 3D (Kahnert, 2008; Liu et al., 2011), and 4D variational methods (Benedetti et al., 2009), as well as ensemble approaches (Sekiyama et al., 2010). Assimilation of satellite products for trace gases is relatively straightforward, since observed and modelled trace gas concentrations are almost directly comparable. However, aerosol optical properties observed from satellites are not directly comparable to the modelled size distribution and chemical composition of the aerosols. Solving this problem amounts to regularising a severely under-constrained inverse problem. Previous aerosol assimilation attempts have been mainly based on educated guesses about the information content of the observations. For instance, there have been studies on the assimilation of aerosol optical depth (AOD) in which all chemical aerosol components in all size classes and at all model layers were used as independent control variables (Liu et al., 2011). This is a rather bold approach that largely disregards the problems involved in inverse modelling. By contrast, it has been proposed to only allow for the total aerosol mass concentration to be corrected by data assimilation of AOD (Benedetti et al., 2009). This is a more prudent approach based on the plausible assumption that a

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

single optical variable only contains enough information to control a single model variable. There have also been intermediate approaches in which the total aerosol mass per size bin have been used as control variables (Saide et al., 2013).

In all such approaches the choice of control variables is based on ad hoc assumptions. Numerical assimilation experiments by Kahnert (2009) suggest that observations of several aerosol optical properties at multiple wavelengths may allow us to constrain more than just the total mass concentration, but certainly not *all* aerosol parameters. However, it is still an unsolved mystery how much information a given set of observations actually contains about the size distribution and chemical composition of aerosols, and exactly which model variables are related to the observed signals, and which ones are related to noise. Thus a prerequisite for assimilating remote sensing observations into aerosol transport models is to thoroughly understand the information content of the observations as well as the relation between the model variables and the signal degrees of freedom.

In a recent study, Burton et al. (2016) have investigated the information content of "$3\beta+2\alpha$"lidar measurements, i.e., observations of backscattering at three wavelengths and extinction at two wavelengths, where the information content was analysed with regard to the refractive index and number distribution of the aerosol particles. As mentioned earlier, the refractive index is a very useful retrieval product of remote sensing observations. However, from the point of view of chemical transport modelling, the main quantities of interest are the concentrations of the different chemical species of which the aerosol particles are composed. Although the chemical composition determines the refractive index, the inversion of this relationship is still under-determined, hence an ill-posed problem. In this paper, we want to investigate the inverse problem that goes all the way from optical properties to the chemical composition of particles.

Thus the two main goals of this paper are (i) to apply a systematic method for analysing the information content of aerosol optical properties with regard to the particles' chemical composition, and (ii) to test an algorithm for making an automatic choice of control variables in chemical data assimilation that relate to the signal degrees of freedom, while all other model variables remain unchanged by the assimilation procedure. The focus will be on spectral observations of extinction and backscattering coefficients, which can be retrieved from lidar observations.[1] We will not restrict this analysis to any fixed choice of wavelengths, such as $3\beta+2\alpha$. Instead, we will investigate the information content for varying combinations of the three main Nd:YAG wavelengths. However, it should be mentioned that extinction measurements at the lowest harmonic of 1064 nm can be difficult and plagued by high errors; in practice, this will affect the observation error, resulting in a low information content of this particular measurement.

The paper is organised as follows. Section 2 gives a rather concise introduction of the modelling tools and of the numerical approach employed to studying the information content of extinction and backscattering observations. Section 3 presents the main results of this study, and Sect. 4 offers concluding remarks. To make this paper self-contained, we included an appendix that gives a brief introduction to some essential concepts of data assimilation, and a detailed explanation of the methods we used for quantifying the information content of aerosol optical observables. It is advisable to read the appendix first before

[1] In addition to lidar measurements from ground-based and aircraft-carried instruments, there are currently two space-borne lidar instruments in orbit. The CALIOP instrument on-board the CALIPSO satellite has been launched in April 2006; it has three receiver channels, one at 1064 nm, and two channels at 532 nm to measure orthogonally polarised components. The CATS instrument on-board the International Space Station has been operational since January 2015; It measures backscattering at 355 nm, 532 nm, and 1064 nm, were the latter two have two orthogonal polarisation channels. It is also capable of performing high spectral resolution measurements at 532 nm. A third instrument is planned to be launched in 2018 (ATLID on-board EarthCARE).

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

reading the body of the paper. Readers that are not interested in the theory may leave out the appendix at the risk of missing some of the discussions in the main body of the paper.

2 Methods

This study consists of two parts. In the first part we quantify the information content of extinction and backscattering coefficients at multiple wavelengths. In the second part we perform a numerical experiment to investigate to what extent the concentrations of different chemical aerosol components can be constrained by observations of extinction and backscattering coefficients. The modelling tools required for this study are (i) a chemical transport model; (ii) an aerosol optics model; and (iii) a data assimilation system.

2.1 Multiple scale Atmospheric Transport and CHemistry modelling system (MATCH)

We employ the chemical transport model MATCH, which is an off-line Eulerian CTM with flexible model domain. It has been previously used from regional to hemispheric scales. Here we use a model version that contains a photochemistry module with 64 chemical species, among them four secondary inorganic aerosol (SIA), namely, ammonium sulphate, ammonium nitrate, other sulphates, and other nitrates. It also contains a module with 16 primary aerosol variables, namely, seasalt, elemental carbon (EC), organic carbon (OC), and dust particles, each emitted in four different size bins. Thus, the total number of aerosol model variables is 20.

The model reads in emission data, meteorological data, and land use data and computes transport processes, chemical transformation, and dry and wet deposition of the various trace gases and aerosols. As output, it provides concentration fields of gases and aerosols, the deposition of these chemical species to land and water-covered areas, as well as the temporal evolution of these variables.

We mention that there exists another model version that includes aerosol microphysical processes, such as nucleation, condensational growth, and coagulation. In that model version the aerosol size distribution evolves dynamically. The model has 20 size bins and seven chemical species (EC, OC, dust, seasalt, particulate sulphate (PSOX), particulate nitrate (PNOX), and particulate ammonium (PNHX)), although not all species are encountered in all size bins. The total number of model variables in the present setup is 82.

More complete information about the mass transport model can be found in Andersson et al. (2007). The seasalt module is discussed in Foltescu et al. (2005). The aerosol microphysics module is described in Andersson et al. (2015).

For the sake of simplicity we here use the mass transport model without aerosol microphysical processes (see next section). The model is set up over Europe covering 33° in the longitudinal and 42° in the latitudinal direction in a rotated lat-long grid with $0.4° \times 0.4°$ horizontal resolution. In the vertical direction the model domain extends up to 13 hPa, using 40 η-layers with variable thickness depending on the underlying topography. The meteorological input data are taken from the numerical weather prediction model HIRLAM (Undén et al., 2002). For the emissions we used EMEP data for the year 2007, where EC

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

and OC emissions were computed from total primary particle emissions based on the data in Kupiainen and Klimont (2004, 2007).

2.2 Aerosol optics model

We have two different optics models coupled to MATCH, one to the mass transport module, and another to the aerosol microphysics module. The former assumes that all aerosol species are homogeneous spheres, and that each chemical species is contained in separate particles. Under these assumptions the optics model is linear, i.e., the optical properties are linear functions of the concentrations of the chemical aerosol species. The latter model accounts for the fact that in reality different chemical species can be internally mixed, i.e., they can be contained in one and the same particle. That model also accounts for the inhomogeneous internal structure of black carbon mixed with other aerosol components, and for the irregular fractal aggregate morphology of bare black carbon particles (Kahnert et al., 2012a, 2013). Under these assumptions the optics model becomes nonlinear, which introduces additional complications in the inverse-modelling problem. This is the main reason why we chose to use the simpler mass transport optics model in this study. Much of the theory explained in the appendix relies on the assumption that the optics model is either linear, or that it is only mildly nonlinear, so that it can be linearised — see Eq. (B6). More information about the aerosol optics models implemented in MATCH can be found in Andersson and Kahnert (2016).

2.3 3-dimensional variational data assimilation (3DVAR)

Data assimilation is a class of statistical methods for combining model results and observations. The algorithm weighs these two pieces of information according to their respective error variances an covariances. As output the assimilation returns a result in model space of which the error variances are smaller than those of the original model estimate. Data assimilation is commonly employed for constraining model results by use of observations. However, one can also employ data assimilation as an inverse-modelling tool, i.e. for retrieving a model state from measurements. A brief summary of the theoretical basis of variational data assimilation is given in the appendix.[2]

The MATCH model contains a 3DVAR data assimilation module. This model uses a spectral method, i.e., the model state vector is Fourier-transformed in the two horizontal coordinates. All error correlations in the horizontal direction are assumed to be homogeneous and isotropic. Error correlations in the vertical direction and among different chemical species are *not* assumed to be separable. The background error covariance matrix of the model a priori is modelled with the NMC method (Parrish and Derber, 1992). A more complete description of our 3DVAR program can be found in Kahnert (2008).

[2]One should actually distinguish between data *assimilation* and data *analysis*. The latter refers to post-processing the model output by statistically weighing model results and observations. The former refers to a process in which the data analysis is incorporated into the time-integration of the CTM. Our 3DVAR code can be used in either analysis or assimilation mode. However, in this study we only perform numerical experiments at a fixed point in time. Thus we use the 3DVAR code as a data analysis tool.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

2.4 Analysis of the information content of aerosol optical parameters

The questions we ask are these.

1. Given m observations of, e.g., m_1 different parameters at m_2 different wavelengths, so that $m_1 \cdot m_2 = m$, how many independent model variables ℓ can we constrain to better than observation error? Obviously, the best we can achieve would be $\ell = m$; but in general, we will have $\ell \leq m$.

2. Which are the ℓ model variables (or linear combinations of model variables) that can be constrained by the measurements?

It turns out the the answer to these questions are found by performing a singular value decomposition of the Jacobian of the observation operator — see Eq. (C6) in the appendix. The right singular vectors can be used to construct a transformation of the model state vector — see Eq. (D16) in the appendix. The transformed vector components fall into two categories, namely, the signal-related components, which can be constrained by the measurements, and the noise-related components, which cannot be constrained by the measurements. From the singular values we can compute the number of signal degrees of freedom, i.e., the number of model variables that can be constrained to better than observation error. We can further compute how much a set of measurements reduces the Shannon-entropy of the model state. This is a quantitative measure for the information content of the measurements.

Readers who are unfamiliar with these concepts are urged to read the brief introduction in the appendix. A more complete discussion of information aspects and inverse methods for atmospheric sounding can be found in Rodgers (2000).

2.5 Numerical assimilation experiments

We use the results of this analysis to modify our 3DVAR program. More specifically, we implement weak constraints into the 3DVAR program such that only the signal-related (transformed) model variables are allowed to be adjusted in the data-analysis procedure, while the noise-related components are not altered. We study the performance of the 3DVAR system by performing a numerical experiment. To this end, we first perform a reference run by driving the MATCH model with analysed meteorological data. These reference results are taken as the "true" chemical state of the atmosphere. We apply the optics model to the model output to generate synthetic "observations". Next we run the MATCH model again, this time driven with 48 hour-forecast meteorological data. The results are taken as a proxy for a background model-estimate that is impaired by uncertainties. Finally, we perform a 3DVAR-analysis of the "observations" and the background estimate in an attempt to restore the reference results. In this numerical experiment we have perfect knowledge of the true state, and we assume that our optics model is nearly perfect, thus providing nearly perfect observations. (We assumed an observation error standard deviation of 10 %). The only factor that may prevent us from fully restoring the reference state is a lack of information in the observed parameters on the chemical composition of the aerosol particles. Thus, comparison of the retrieval and reference results gives us an indication of how strongly different model variables can be controlled by the information contained in the observations.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

3 Results

3.1 Analysis of the information content of aerosol optical parameters

To be specific, we consider the set of parameters $\{k_{\rm ext}(\lambda_1), k_{\rm ext}(\lambda_2), k_{\rm ext}(\lambda_3), \beta_{\rm sca}(\lambda_1), \beta_{\rm sca}(\lambda_2), \beta_{\rm sca}(\lambda_3)\}$, where $k_{\rm ext}$ and $\beta_{\rm sca}$ denote the extinction and backscattering coefficients, respectively, and the wavelengths $\lambda_1 = 1064$ nm, $\lambda_2 = 532$ nm, and $\lambda_3 = 355$ nm denote the first three Nd:YAG harmonics. Out of this six-parameter set we pick different subsets and analyse the singular values of the corresponding observation operators. From those we compute the number of signal degrees of freedom as well as the change in Shannon-entropy for each subset of measurements. The results are listed in Table 1, which shows a number of interesting findings:

1. When we increase the number of wavelengths from one to two, then the number of signal degrees of freedom N_s shows a corresponding increase from 1 to around 1.95–2.00. The change in entropy H indicates a similar trend; it increases from around 7 for a single wavelength to up to 12 for two wavelengths. (Compare, e.g., cases 1., 2., and 3. to cases 4., 5., and 6.) Hence we almost double the information contained in the measurements.

2. When we increase the number of wavelengths further from two to three, then N_s only increases from around 2.0 to around 2.7 (compare cases 4., 5., and 6. to case 7, or case 9. to case 10.) This is also reflected in H; it only increases from around 12 to 13. This indicates that in our particular case there is little extra information to be gained by extending the number of spectral measurements beyond 2–3 wavelengths.

3. Supplementing extinction with backscattering measurements results in a significant increase in N_s and H. This can be seen by comparing, e.g., cases 5. and 11. By adding $\beta_{\rm sca}$ observations to $k_{\rm ext}$ observations the number of signal degrees of freedom increases from 2 to 3.9, so it almost doubles, while H increases from 12 to 18.

4. Case 12 clearly reveals the limitations of extending the set of observed parameters; N_s is only 4.6, significantly lower than the number of observed parameters, m=6.

5. An inspection of w_i shows that the singular values often display quite a dramatic decrease from the largest to the smallest value. By contrast, the contribution N_s^i to the total number of signal degrees of freedom decreases more gently. This fact is interesting in relation to the choice of the covariance matrix of the weak constrains — see Eqs. (D18), (D19) and (D21) in the appendix. In view of our findings here, we conclude that (D18) would yield a very sharp transition from unconstrained to constrained model variables, Eq. (D19) would give a smooth transition, and Eq. (D21) would give a moderately sharp transition.

We performed a sensitivity study on how the observation error affects the information content in the analysis. It is important to understand that the observation error is not the same as the measurement error. The latter contributes to the former, but the observation error contains also other sources of error. For instance, if we deal with morphologically complex particles, but our lack of knowledge forces us to make assumptions and invoke approximations about the particle shapes, then this source

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

Table 1. Signal degrees of freedom N_s and change in entropy H for the lowest model layer (closest to the surface) and different subsets of $\{k_{\rm ext}(\lambda_1), k_{\rm ext}(\lambda_2), k_{\rm ext}(\lambda_3), \beta_{\rm sca}(\lambda_1), \beta_{\rm sca}(\lambda_2), \beta_{\rm sca}(\lambda_3)\}$, where $k_{\rm ext}$ denotes the extinction coefficient, $\beta_{\rm sca}$ represents the backscattering coefficient, and the wavelengths $\lambda_1 = 1064$ nm, $\lambda_2 = 532$ nm, and $\lambda_3 = 355$ nm denote the first three Nd:YAG harmonics. Also shown are the singular values w_i and their contributions N_s^i and H_i to N_s and H, respectively. The results have been obtained by assuming an observation standard deviation of 10 %.

No.	Parameters	Wavelengths	i	w_i	N_s^i	H_i	N_s	H
1.	$k_{\rm ext}$	λ_3	1	163	1.00	7.35	1.00	7.35
2.	$k_{\rm ext}$	λ_2	1	172	1.00	7.43	1.00	7.43
3.	$k_{\rm ext}$	λ_1	1	194	1.00	7.52	1.00	7.52
4.	$k_{\rm ext}$	λ_2, λ_3	1	237	1.00	7.89	1.95	10.1
			2	4.57	0.95	2.23		
5.	$k_{\rm ext}$	λ_1, λ_3	1	253	1.00	7.98	2.00	11.9
			2	15.0	1.00	3.91		
6.	$k_{\rm ext}$	λ_1, λ_2	1	259	1.00	8.02	1.99	11.5
			2	10.9	0.99	3.45		
7.	$k_{\rm ext}$	λ_1, λ_2, λ_3	1	305	1.00	8.25	2.72	13.2
			2	15.8	1.00	3.98		
			3	1.64	0.73	0.94		
8.	$\beta_{\rm sca}$	λ_1	1	226	1.00	7.82	1.00	7.82
9.	$\beta_{\rm sca}$	λ_1, λ_3	1	292	1.00	8.19	2.00	12.3
			2	17.6	1.00	4.14		
10.	$\beta_{\rm sca}$	λ_1, λ_2, λ_3	1	354	1.00	8.47	2.65	13.4
			2	18.2	1.00	4.18		
			3	1.36	0.65	0.76		
11.	$k_{\rm ext}$, $\beta_{\rm sca}$	λ_1, λ_3	1	386	1.00	8.59	3.89	17.9
			2	32.8	1.00	5.04		
			3	4.80	0.96	2.29		
			4	3.66	0.93	1.92		
12.	$k_{\rm ext}$, $\beta_{\rm sca}$	λ_1, λ_2, λ_3	1	467	1.00	8.87	4.61	19.4
			2	37.8	1.00	5.24		
			3	5.54	0.97	2.49		
			4	4.18	0.95	2.10		
			5	0.95	0.48	0.47		
			6	0.53	0.22	0.18		

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

Table 2. Signal degrees of freedom N_s and change in entropy H as a function of observation standard deviation, taken from the first model layer and case 12 in table 1.

Obs. std. dev. (%)	N_s	H
1	5.96	37.6
5	5.29	24.3
10	4.61	19.3
50	3.00	10.5
100	2.33	7.8

of error contributes to the observation error. The same is the case if we lack information about the particles' size distribution. Such assumptions also enter into our relatively simple optics model, so our previous assumption of an observation standard deviation of 10 % represents, most likely, a highly idealised case. [3]

To get an idea about the significance of the observation error on the amount of information we can extract from measurements, we consider case 12 in table 1, and we varied the observation standard deviation from 1% to 100%. Table 2 shows how the total entropy and signal degrees of freedom vary with the observation standard deviation. The larger the standard deviation, the less information can be obtained from the observations. Both the total entropy H and the signal degrees of freedom N_s decrease with increasing standard deviation. For a standard deviation of 100 %, we only have two signal degrees of freedom contained in the six observed optical parameters. This demonstrates two important things.

1. It is essential to develop accurate and realistic aerosol optics models. The most accurate measurements may intrinsically contain a wealth of information on aerosol properties. But we can only make use of this information to the extent that our observation operator is able to accurately describe the relation between the physical and chemical particle characteristics and their optical properties.

2. It is equally essential to accurately estimate the contribution of the uncertainties in the aerosol optics model to the observation error. If we underestimate this error, we will rely too much on the measurements than we should, thus assimilating noise. If we overestimate this error, we will waste information contained in the observations.

3.2 Numerical inverse-modelling experiment

We integrated the findings of 3.1 into our 3DVAR program by constraining the algorithm to varying only the signal-related model variables. We employed the weak-constraint approach described in the appendix. More specifically, the constraints are formulated by use of the ansatz given in Eq. (D21), where we set the constant $c'' = \min\{w_k, 0.1\}$, and where w_k is the smallest singular value of the scaled observation operator — see Eq. (C6).

[3] A more realistic optics model, such as the one investigated in Andersson and Kahnert (2016) would help to reduce the observation standard deviation. For future studies, such a model should be linearised and investigated in a similar way.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

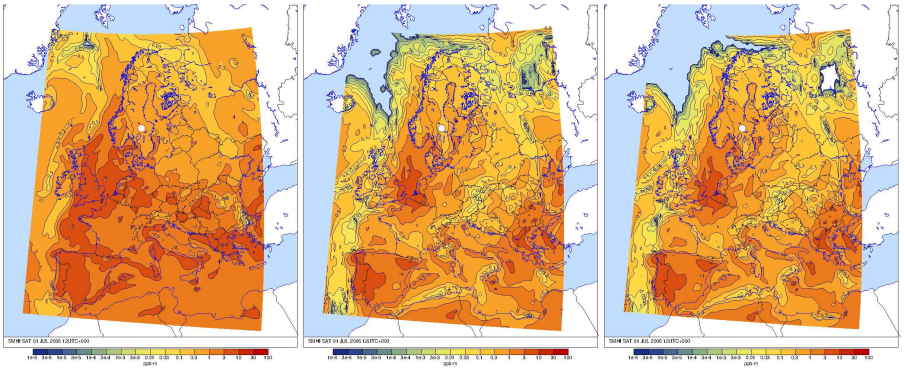

Figure 1. Ammonium sulphate mixing ratio over Europe. Left: reference field, centre: background field, right: 3DVAR analysis. The observation site is indicated by the white circle. Note the nonlinear colour scale!

To illustrate the method we conduct a numerical experiment as described in Sect. 2.5. We perform a 3DVAR analysis of the background field by assimilating four different vertical profiles of optical properties, namely, the backscattering coefficient, β_{bak}, and the extinction coefficient, k_{ext}, each at a wavelength of 355 nm and 1064 nm. Thus in our case the number of singular values in each vertical layer is $k = 4$. We assume an idealised situation in which the observation standard deviation is only 10 %. As we see in Table 1 (row 11), the number of singular values larger than unity is $\ell = k = 4$, and the number of signal degrees of freedom is N_s =3.9. So we have as many signal degrees of freedom as we have measurements.

As an example, Fig. 1 shows the ammonium sulphate mixing ratio in the lowest model layer (closest to the surface) computed for the reference run (left), the background estimate (centre), and the 3DVAR analysis (right)[4]. Clearly, the background field underestimates the reference field in most areas. We picked a location in Central Sweden (60°N, 15°E) to compute backscattering and extinction profiles from the reference results, which were then 3DVAR-analysed in conjunction with the background field. The analysis (right) restores the reference mixing ratios at and near the observation site. So, at least for ammonium sulphate mixing ratios in the lowest model layer we seem to obtain a satisfying solution to the inverse modelling problem.

A closer inspection of the analysis performance is given in Fig. 2. Each panel shows vertical profiles of mixing ratios at the observation site. We compare the analysis results (red solid line) to both the background estimate (blue dashed line) and the reference results (black solid line). The reference results of the secondary inorganic aerosol species (SIA, which is the sum of all sulphate, nitrate, and ammonium mixing ratios) are almost completely restored by the 3DVAR analysis at all altitudes. For elemental carbon (EC), organic carbon (OC), dust, and, even more so, sodium chloride (NaCl) the analysis overestimates the reference results at altitudes between 0–2 km, while above 2 km the reference results are at least partially restored by the analysis. When we compare the different scales on the x-axes, we see that SIA makes the dominant contribution to the aerosol mixing ratio. Accordingly, the total aerosol mass mixing ratio (PM_{10}) is almost equally well restored by the analysis as the SIA mixing ratio.

[4]The approach for generating a reference and background model-run has been explained in Sect. 2.5

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

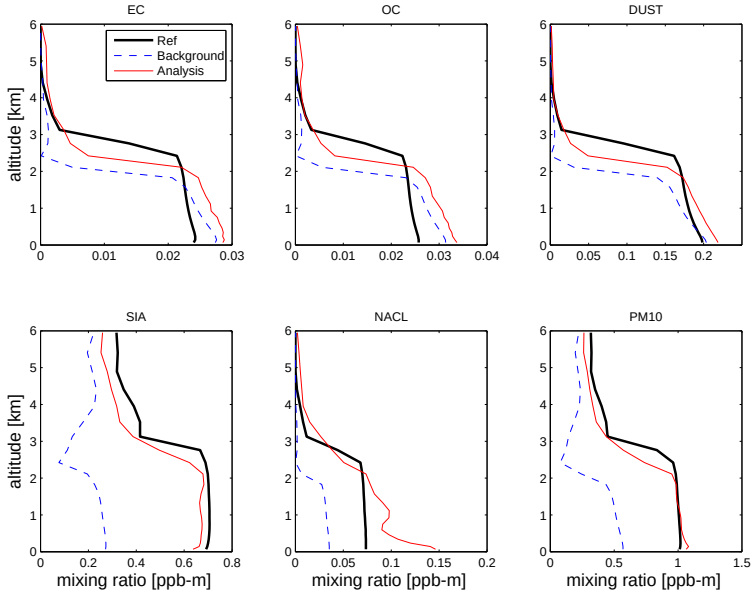

Figure 2. Vertical profiles of elemental carbon (EC), organic carbon (OC), dust, secondary inorganic aerosols (SIA), sea salt, and total aerosol mass mixing ratio (PM_{10}). Each panel shows the reference results (black solid line), background estimate (dashed blue line), and the 3DVAR analysis (red solid line).

Figure 3 shows the observations (black solid line) as well as the observation-equivalents of the background estimate (blue dashed line) and the 3DVAR analysis (red solid line) for all four observations, namely, β_{bak} at 355 nm wavelength (top left), β_{bak} at 1064 nm (top right), k_{ext} at 355 nm (bottom left), and k_{ext} at 1064 nm (bottom right). We learn from this figure that the analysis follows the observations faithfully. The reason for this is that we assumed that the observations were highly accurate with an error standard deviation of only 10 %.

We have seen that the analysis provides a reasonable, but, as expected, not a perfect answer to the inverse problem. We have further seen that at (and near) the observation site it relies more on the observations than on the background estimate. However, the previous figures tell us little about the effect of the constraints we introduced. To learn about that we need to inspect the analysis in the abstract space of the transformed model variables $\delta\boldsymbol{x}'$ given in Eq. (C16). Figure 4 shows vertical profiles of all 20 variables $\delta x'_i$. The error variance within which each of these variables is allowed to vary in the analysis is given by the diagonal elements of the matrix $\mathbf{B}_G$ in Eq. (D21). The first four of these are the singular values given in Table 1 (row 11). The remaining 16 variances are set to 0.1. Thus the first element $\delta x'_1$ has by far the largest freedom to be adjusted by the 3DVAR algorithm. The error variance of the second element, $\delta x'_2$, is smaller by roughly one order of magnitude. The error variances of $\delta x'_3$ and $\delta x'_4$ are of comparable magnitude, and each one is about one order of magnitude smaller than that of $\delta x'_2$. Finally, the error variances of the remaining 16 elements are about one order of magnitude smaller than that of $\delta x'_4$. Accordingly, the first element, $\delta x'_1$, is the one that deviates most strongly from zero. The elements $\delta x'_2$, $\delta x'_3$, and $\delta x'_4$ are varied

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

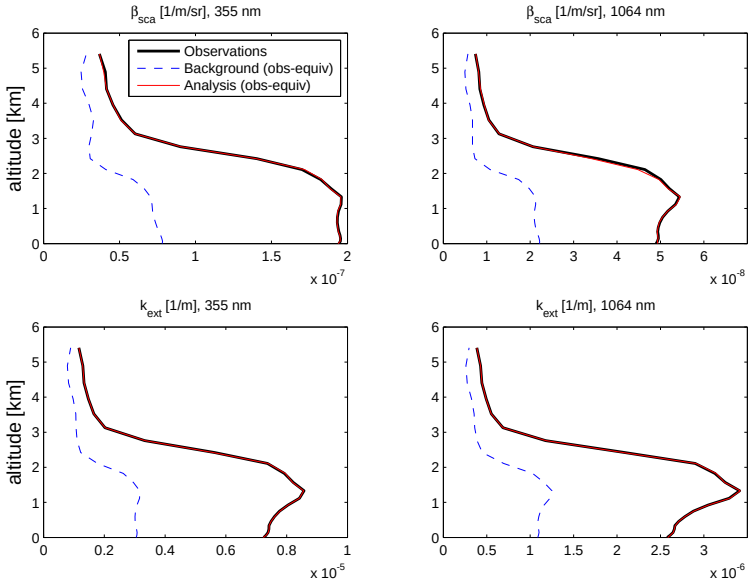

Figure 3. Observations (black solid line), and observation-equivalents of the background estimate (dashed blue line) and of the 3DVAR analysis (red solid line). The optical parameters and wavelengths are indicated above each panel.

much less in comparison. If we had imposed strong constraints, then the remaining elements would be exactly zero. However, our weak-constraint formulation allows even the other elements to deviate from zero within relatively tight limits. But several of them are, in fact, very close to zero, notably the elements $\delta x_i'$ for $i = 7$–9, 11–14, 16, and 19.

4 Summary and conclusions

We have quantified the information content of extinction and backscattering measurements with regard to the chemical composition of aerosol particles. This has been done by determining the singular values of the observation operator, by computing the number of signal degrees of freedom, and by calculating the change in Shannon-entropy caused by taking measurements. We first assumed a relatively low observation standard deviation of 10 %. In that case, when adding measurements of β_{bak} to measurements of k_{ext}, the information content nearly doubles. The same is true when we increase the number of wavelengths from a single wavelength to two wavelengths. However, when we further increase the number of wavelengths from two to three, then the gain in information is rather modest; there appeared to be little use in increasing the number of optical wavelengths beyond three. When the full set of six observations (both optical parameters at three wavelengths) is considered, then the number of signal degrees of freedom is 4.6. Thus we can constrain, at most, 4–5 model variables with this set of observations.

These conclusions depend, to be sure, on the assumed observation standard deviation. We therefore performed a sensitivity study were we investigated how the observation standard deviation affects the information content. We observed a rather

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

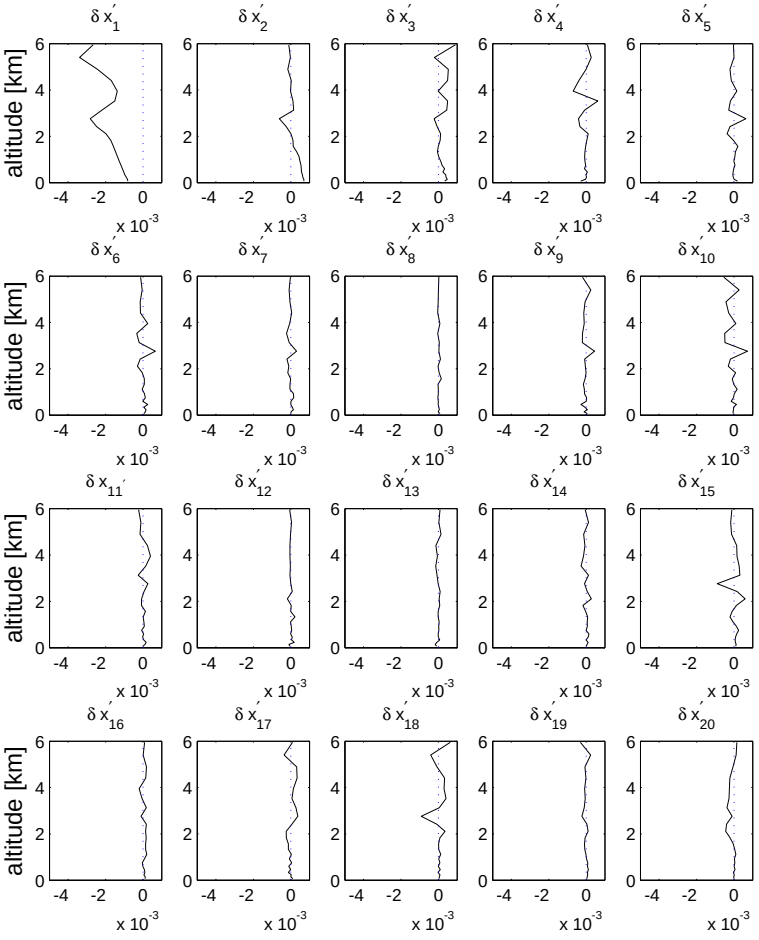

Figure 4. Vertical profiles of the transformed model variables $\delta \boldsymbol{x}'$.

dramatic decrease in both the entropy and signal degrees of freedom with increasing observation standard deviation. Note that not only the measurement error, but also the uncertainties in the aerosol optics model contribute to the observation error. This highlights the importance of developing accurate aerosol optics models and of obtaining an accurate estimate of the observation error, especially of the uncertainty in the aerosol optics model. This is a prerequisite for extracting as much information as possible from the measurements, while avoiding to extract noise rather than signal. More often than not, computational limitations and lack of knowledge force us to introduce simplifying assumptions about the particles' morphologies. However, we know that aerosol optical properties can be highly sensitive to the shape (Mishchenko et al. (1997); Kahnert (2004)), small-scale surface roughness (Kahnert et al., 2012b), inhomogeneity (Mishchenko et al., 2014; Kahnert, 2015), ag-

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

gregation (Fuller and Mackowski, 2000; Liu and Mishchenko, 2007; Kahnert and Devasthale, 2011), irregularity (Muinonen, 2000; Bi et al., 2010), porosity (Vilaplana et al., 2006; Lindqvist et al., 2011; Kylling et al., 2014), and combinations thereof (Lindqvist et al., 2009; Kahnert et al., 2013; Lindqvist et al., 2014). We need to know how much these sources of uncertainty contribute to the observation standard deviation. One way of estimating this is to compare aerosol optical properties computed with simple shape models to either measurements or to computations based on more realistic particle shape models — see Kahnert et al. (2016) for a recent review and a more detailed discussion.

We exploited our analysis of the information content and the number of signal degrees of freedom by formulating weak constraints in a 3DVAR algorithm. More specifically, we transformed the model variables into a new basis in which the components of the state vector can be divided into signal-related and noise-related components. We then added weak constraints to the assimilation algorithm in such a way that only the signal-related transformed model variables are varied by the 3DVAR analysis. Numerical experiments showed that the 3DVAR algorithm provided a reasonable solution to the inverse problem; when mapped into observation space, the analysis result closely reproduces the measurements. It also appeared that among the original model variables, secondary inorganic aerosol components were most faithfully retrieved by the inverse modelling solution. Most importantly, it was demonstrated that the 3DVAR analysis follows, indeed, the imposed constraints; the transformed model variables are adjusted within certain limits according to how strongly they relate to the signal degrees of freedom.

The results presented here suggest further questions that should be addressed in future studies. We have performed this investigation with a mass transport model, thus focusing on the information content of optical measurements on the chemical composition of aerosols. When we include aerosol microphysical processes, then the model delivers the aerosols' size distribution, as well as their size-resolved chemical composition. This makes the problem quite different from the one we investigated here. First, the dimension of the model space is considerably larger for an aerosol microphysics transport model. Constraining such a model with limited information from measurements becomes even more challenging than in the case of a mass transport model. On the other hand, an aerosol microphysics model delivers information on the particles size distribution and mixing state. Therefore, this would require us to make fewer assumptions in the aerosol optics model, which may reduce the observation error. The present study should be extended to investigate the information contained in extinction and backscattering measurements for simultaneously constraining the chemical composition and the size of aerosol particles.

Another important, and often highly underrated issue concerns the choice of the aerosol optics model. In the present study we employed a simple homogeneous-sphere model in which all chemical components were assumed to be externally mixed. There is little one can put forward in defence of this model other than pure convenience. In this model the observation operator is linear, which is a prerequisite for much of the theoretical foundations of this study — see the appendix for details. However, it has been demonstrated that drastically simplifying assumptions, such as the external-mixture approximation, can give model results for aerosol optical properties that differ substantially from those obtained with more realistic nonlinear optics models (Andersson and Kahnert, 2016). It would therefore be important to extend the present study to include more accurate and realistic optics models. A first step could be to analyse the degree of nonlinearity of optics models that account for internal mixing of different aerosol species. If they turn out to be only mildly nonlinear, then one can linearise them and work with

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

the Jacobian of the nonlinear observation operator. Otherwise the theoretical methods employed in this paper would have to be extended in order to accommodate nonlinear observation operators.

Appendix A: Inverse problems

Suppose we have a system described by a set of variables $x_1, \ldots, x_n$, summarised in a vector $\boldsymbol{x}$. Suppose also that we have an operator $\hat{H} : \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{x} \mapsto \boldsymbol{y} = \hat{H}(\boldsymbol{x})$ that allows us to compute a set of variables $y_1, \ldots, y_m$, summarised in a vector $\boldsymbol{y}$. To take a specific example, we may think of $\boldsymbol{x}$ as a vector of mixing ratios of chemical aerosol species, $\boldsymbol{y}$ as a set of aerosol optical properties, and $\hat{H}$ as an aerosol optics model. We consider the following two problems:

1. **Direct problem:** Given $\boldsymbol{x}$ and $\hat{H}$, calculate $\boldsymbol{y} = \hat{H}(\boldsymbol{x})$.
2. **Inverse problem:** Given $\boldsymbol{y}$ and $\hat{H}$, solve $\boldsymbol{y} = \hat{H}(\boldsymbol{x})$ for $\boldsymbol{x}$.

A pair of such problems is inverse *to each other*; it is, therefore, somewhat arbitrary which problem we choose to call the direct problem, and which one we call the inverse problem. However, one of the problems is usually *well-posed*, while the other one is *ill-posed*. Such is also the case in aerosol optics modelling. It is customary to call the well-posed problem the *direct problem*, and the ill-posed one the *inverse problem*.

An equation $\boldsymbol{y} = \hat{H}(\boldsymbol{x})$ is called *well-posed* if it has the following properties:

1. **Existence:** For every $\boldsymbol{y} \in \mathbb{R}^m$, there is at least one $\boldsymbol{x} \in \mathbb{R}^n$ for which $\boldsymbol{y} = \hat{H}(\boldsymbol{x})$.
2. **Uniqueness:** For every $\boldsymbol{y} \in \mathbb{R}^m$, there is at most one $\boldsymbol{x} \in \mathbb{R}^n$ for which $\boldsymbol{y} = \hat{H}(\boldsymbol{x})$.
3. **Stability:** The solution $\boldsymbol{x}$ depends continuously on $\boldsymbol{y}$.

If any of these properties is not fulfilled, then the problem is called *ill-posed*.

Appendix B: 3-dimensional variational data assimilation

Data assimilation is usually employed for constraining models by use of measurements, but it can also be used to solve inverse problems. Here we focus on one specific data assimilation method known as 3-dimensional variational data assimilation, or 3DVAR.

In a CTM we discretise the geographic domain of interest into a 3-dimensional grid. In each grid cell, the aerosol particles are characterised by the mass concentrations of each chemical component in the aerosol phase, such as sulphate, nitrate, ammonium, mineral dust, black carbon, organic carbon, and sea salt. Suppose we summarise all these mass concentrations from all grid cells into one large vector $\boldsymbol{x} \in \mathbb{R}^n$. The model provides us with a first guess of the atmospheric aerosol state, known as a *background estimate*[5] $\boldsymbol{x}_b$. Suppose also that we have m observations, which we summarise in a vector $\boldsymbol{y} \in \mathbb{R}^m$. We

[5] In the remote sensing and inverse modelling community, the background estimate is more commonly referred to as the *a priori* estimate.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

further have an observation operator $\hat{H} : \mathbb{R}^n \to \mathbb{R}^m$, $\boldsymbol{x} \mapsto \hat{H}(\boldsymbol{x})$ that maps the state vector $\boldsymbol{x}$ from model space to observation space[6]. We further denote by $\boldsymbol{x}_t$ the true state of the atmosphere, by $\boldsymbol{\epsilon}_b = \boldsymbol{x}_t - \boldsymbol{x}_b$ the error of the background estimate, and by $\boldsymbol{\epsilon}_o = \hat{H}(\boldsymbol{x}_t) - \boldsymbol{y}$ the observation error.[7] If the background errors are not correlated with the observation errors, then their joint probability distribution becomes separable, i.e.

$$P(\boldsymbol{\epsilon}_b, \boldsymbol{\epsilon}_o) = P_b(\boldsymbol{\epsilon}_b) P_o(\boldsymbol{\epsilon}_o). \tag{B1}$$

The true state of the atmosphere is, of course, unknown. Therefore, our definition of the errors and their probability distribution is only of conceptual use, but not of any practical value. However, we can reinterpret the probability distributions by replacing $\boldsymbol{\epsilon}_b$ in the argument of P_b with $\boldsymbol{x} - \boldsymbol{x}_b$, and by replacing $\boldsymbol{\epsilon}_o$ in the argument of P_o with $\hat{H}(\boldsymbol{x}) - \boldsymbol{y}$. We further assume that both the background and the observation errors are normally distributed. Thus we may write

$$P_b(\boldsymbol{x}) = (2\pi \mid \mathbf{B} \mid)^{-1/2} \exp\left(-\frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_b)^T \cdot \mathbf{B}^{-1} \cdot (\boldsymbol{x} - \boldsymbol{x}_b)\right) \tag{B2}$$

$$P_o(\boldsymbol{x}) = (2\pi \mid \mathbf{R} \mid)^{-1/2} \exp\left(-\frac{1}{2}(\hat{H}(\boldsymbol{x}) - \boldsymbol{y})^T \cdot \mathbf{R}^{-1} \cdot (\hat{H}(\boldsymbol{x}) - \boldsymbol{y})\right). \tag{B3}$$

Here $\mathbf{B}$ and $\mathbf{R}$ denote the covariance matrices of the background and observation errors, respectively, and $|\cdot|$ denotes the matrix determinant. In this form, $P_b(\boldsymbol{x})$ represents the probability that the atmospheric aerosol particles are found in state $\boldsymbol{x}$, given a background estimate $\boldsymbol{x}_b$ with error covariance matrix $\mathbf{B}$. Similarly, $P_o(\boldsymbol{x})$ is the probability that the system is found in state $\boldsymbol{x}$, given measurements $\boldsymbol{y}$ with error covariances $\mathbf{R}$.[8]

Equations (B1)–(B3) can be summarised in the form

$$P(\boldsymbol{x}) = \frac{1}{2\pi(\mid \mathbf{B} \mid \cdot \mid \mathbf{R} \mid)^{-1/2}} \exp(-J(\boldsymbol{x})) \tag{B4}$$

$$J(\boldsymbol{x}) = \frac{1}{2}\left[(\boldsymbol{x} - \boldsymbol{x}_b)^T \cdot \mathbf{B}^{-1} \cdot (\boldsymbol{x} - \boldsymbol{x}_b) + (\hat{H}(\boldsymbol{x}) - \boldsymbol{y})^T \cdot \mathbf{R}^{-1} \cdot (\hat{H}(\boldsymbol{x}) - \boldsymbol{y})\right], \tag{B5}$$

where J is suggestively called the costfunction, since it can be interpreted as a measure for how "costly" it is for a state $\boldsymbol{x}$ to simultaneously deviate from the background estimate and the measurements within the permitted error bounds. The deviations are weighted with the inverse error covariance matrices. For instance, this means that for measurements with a small error variance, a deviation $\hat{H}(\boldsymbol{x}) - \boldsymbol{y}$ becomes "more costly".

We are interested in the most probable aerosol state of the atmosphere, i.e., in that state $\boldsymbol{x}_a$ for which the probability distribution attains its maximum. This is obviously the case when the argument of the exponential in Eq. (B4) assumes a

[6]The optics model $\hat{H}$ usually has to invoke assumptions about physical aerosol properties that are relevant for the optical properties, but not provided by the CTM output, e.g. assumptions about the morphology of the particles. If the CTM is a simple mass-transport model without aerosol microphysics, then it is also necessary to invoke assumptions about the size distribution of the aerosols.

[7]The observation error must not be confused with the measurement error. The latter contributes to the former, but the observation error contains also other sources of error. For instance, if we deal with morphologically complex particles, but our lack of knowledge forces us to make assumptions and invoke approximations about the particle shapes, then this source of error contributes to the observation error. The same is the case if we lack information about the particles' size distribution.

[8]The observation errors are often uncorrelated. In such case the matrix $\mathbf{R}$ is diagonal, where the diagonal elements are the observation error variances.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

minimum. Thus we seek to minimise the costfunction J. The variational method is based on computing the gradient of the costfunction, ∇J, and to use this in a descent algorithm to iteratively search for the minimum of J.

In practice it is common to introduce the variable $\delta\boldsymbol{x} = \boldsymbol{x} - \boldsymbol{x}_b$, and use the first-order Taylor expansion of the observation operator,

$$\hat{H}(\boldsymbol{x}) = \hat{H}(\boldsymbol{x}_b) + \mathbf{H} \cdot \delta\boldsymbol{x}, \tag{B6}$$

where the $(m \times n)$-matrix $\mathbf{H}$ denotes the Jacobian of $\hat{H}$ at $\boldsymbol{x} = \boldsymbol{x}_b$. If $\hat{H}$ is only mildly non-linear, and if the components of $\delta\boldsymbol{x}$ are sufficiently small, then we can substitute this first-order approximation into Eq. (B5), which yields

$$J = J_b + J_o \tag{B7}$$

$$J_b(\delta\boldsymbol{x}) = \frac{1}{2}\delta\boldsymbol{x}^T \cdot \mathbf{B}^{-1} \cdot \delta\boldsymbol{x} \tag{B8}$$

$$J_o(\delta\boldsymbol{x}) = \frac{1}{2}\left(\hat{H}(\boldsymbol{x}_b) + \mathbf{H} \cdot \delta\boldsymbol{x} - \boldsymbol{y}\right)^T \cdot \mathbf{R}^{-1} \cdot \left(\hat{H}(\boldsymbol{x}_b) + \mathbf{H} \cdot \delta\boldsymbol{x} - \boldsymbol{y}\right) \tag{B9}$$

The components of the vector $\delta\boldsymbol{x}$ are the *control variables* that are iteratively varied by the algorithm until the minimum of the costfunction is found.

The solution to the equation $\nabla J = \mathbf{0}$ is a solution to the inverse problem; we input the observations $\boldsymbol{y}$ into the algorithm, and as output we obtain a result in model space that is consistent with the measurements (within the given error bounds). What if the measurements contain insufficient information about the state $\boldsymbol{x}$? The algorithm will still provide an answer to the inverse problem, but the missing information will be supplemented by the background estimate $\boldsymbol{x}_b$. The weighting of the two pieces of information, $\boldsymbol{x}_b$ and $\boldsymbol{y}$, is controlled by the respective error covariance matrices. Thus data assimilation is a statistical approach, which can be expected to give good results *on average*, but not in every single time-step of the model run. This can become highly problematic if we only have very few observations, i.e., $m \ll n$. If we allow all model variables to be freely adjusted by the assimilation algorithm in such a severely under-constrained case, then the algorithm may just assimilate noise rather than signal, resulting in unreasonable solutions to the inverse problem (e.g. Kahnert (2009)). To avoid such problems, one needs to systematically analyse the information content of the observations and constrain the assimilation algorithm to only operate on the signal degrees of freedom.

Appendix C: Information content of measurements

Our ultimate goal is to formulate the data assimilation problem in such a way that the information contained in the measurements is fully exploited, but not over-used. To this end, we first need to know how many independent quantities can be determined from a specific set of measurements. We investigate this question by borrowing ideas from retrieval and information theory — see Rodgers (2000) for more detailed explanations.

The main idea is to compare the variances of the model variables to those of the observations. Only those model variables whose variance is larger than those of the observations can be constrained by measurements. However, to actually make such

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

a comparison is rather tricky. The first problem is that one cannot readily compare error covariance *matrices*. The second problem is that model variables and measurements are in different spaces. We first address the second problem.

When we account for observation errors $\boldsymbol{\epsilon}_o$, then the basic relation between model variables and observations is, to first order

$$\boldsymbol{y} = \hat{H}(\boldsymbol{x}_b) + \mathbf{H} \cdot \delta\boldsymbol{x} + \boldsymbol{\epsilon}_o. \tag{C1}$$

The error covariance matrices are given by the expectation values $\mathbf{B} = \langle \delta\boldsymbol{x} \cdot \delta\boldsymbol{x}^T \rangle$, and $\mathbf{R} = \langle \boldsymbol{\epsilon}_o \cdot \boldsymbol{\epsilon}_o^T \rangle$, where the dot denotes a dyadic product. The covariance matrix of $\delta\boldsymbol{y} = \boldsymbol{y} - \hat{H}(\boldsymbol{x}_b)$ is given by $\langle \delta\boldsymbol{y} \cdot \delta\boldsymbol{y}^T \rangle = \mathbf{H} \cdot \mathbf{B} \cdot \mathbf{H}^T + \mathbf{R}$, where we assumed that background and observation errors are uncorrelated. This last equation suggests that we can compare model and observation errors in the same space by transforming the background error covariance matrix from the space of $(n \times n)$ matrices to the space of $(m \times m)$ matrices viz. $\mathbf{H} \cdot \mathbf{B} \cdot \mathbf{H}^T$.

To address the first problem, we diagonalise the covariance matrices by making the following change of variables

$$\delta\tilde{\boldsymbol{x}} = \mathbf{B}^{-1/2} \cdot \delta\boldsymbol{x} \tag{C2}$$

$$\delta\tilde{\boldsymbol{y}} = \mathbf{R}^{-1/2} \cdot (\boldsymbol{y} - \hat{H}(\boldsymbol{x}_b)) \tag{C3}$$

$$\tilde{\mathbf{H}} = \mathbf{R}^{-1/2} \cdot \mathbf{H} \cdot \mathbf{B}^{1/2}. \tag{C4}$$

Here $\mathbf{B}^{1/2}$ denotes the positive square root[9] of the matrix $\mathbf{B}$, and $\mathbf{B}^{-1/2}$ denotes its inverse. In the new basis, the costfunction in (B7)–(B9) becomes

$$J = \frac{1}{2}\delta\tilde{\boldsymbol{x}}^T \cdot \delta\tilde{\boldsymbol{x}} + \frac{1}{2}\left(\tilde{\mathbf{H}} \cdot \delta\tilde{\boldsymbol{x}} - \delta\tilde{\boldsymbol{y}}\right)^T \cdot \left(\tilde{\mathbf{H}} \cdot \delta\tilde{\boldsymbol{x}} - \delta\tilde{\boldsymbol{y}}\right). \tag{C5}$$

The covariance matrices are now unit matrices. This can also be seen by considering the transformed errors, e.g. $\tilde{\boldsymbol{\epsilon}}_o = \mathbf{R}^{-1/2} \cdot \boldsymbol{\epsilon}_o$ and computing $\langle \tilde{\boldsymbol{\epsilon}}_o \cdot \tilde{\boldsymbol{\epsilon}}_o^T \rangle = \mathbf{R}^{-1/2} \cdot \langle \boldsymbol{\epsilon}_o \cdot \boldsymbol{\epsilon}_o^T \rangle \cdot \mathbf{R}^{-1/2} = \mathbf{1}$, since $\langle \boldsymbol{\epsilon}_o \cdot \boldsymbol{\epsilon}_o^T \rangle = \mathbf{R}$. Similarly, we find $\langle \delta\tilde{\boldsymbol{x}} \cdot \delta\tilde{\boldsymbol{x}}^T \rangle = \mathbf{1}$. The covariance matrix of the transformed measurement vector $\delta\tilde{\boldsymbol{y}}$ is given by $\langle \delta\tilde{\boldsymbol{y}} \cdot \delta\tilde{\boldsymbol{y}}^T \rangle = \tilde{\mathbf{H}} \cdot \tilde{\mathbf{H}}^T + \mathbf{1}$. The first term is the model error covariance term transformed into observation space, while the second term (the unit matrix) is the diagonalised observation error covariance matrix.

We are still not in a position to make a meaningful comparison of model and observation errors, since the first term, $\tilde{\mathbf{H}} \cdot \tilde{\mathbf{H}}^T$, is still not diagonal. To make it so we need to perform one more transformation. To this end, we consider the singular value decomposition of the matrix $\tilde{\mathbf{H}}$,

$$\tilde{\mathbf{H}} = \mathbf{R}^{-1/2} \cdot \mathbf{H} \cdot \mathbf{B}^{1/2} = \mathbf{V}_L \cdot \mathbf{W} \cdot \mathbf{V}_R^T. \tag{C6}$$

Here $\tilde{\mathbf{H}}$ is a $(m \times n)$-matrix, the matrix of the left-singular vectors $\mathbf{V}_L$ is a $(m \times m)$-matrix, the matrix $\mathbf{V}_R$ containing the right-singular vectors is a $(n \times n)$-matrix, and the $(m \times n)$-matrix $\mathbf{W}$ consists of two blocks. If $m < n$, then the left block of $\mathbf{W}$ is a $(m \times m)$-diagonal matrix containing the m singular values $w_1, \ldots, w_m$ on the diagonal; the right block is a $(m \times (n-m))$-nullmatrix. Similarly, if $m > n$, then the upper block of $\mathbf{W}$ is a $(n \times n)$-diagonal matrix containing the n singular values on the diagonal, while the lower block is a $((m-n) \times n)$-nullmatrix.

[9] A matrix $\mathbf{A}$ is called a square root of a matrix $\mathbf{B}$ if $\mathbf{A} \cdot \mathbf{A} = \mathbf{B}$. The *positive* square root of $\mathbf{B}$, which is denoted by $\mathbf{B}^{1/2}$, has the property $\boldsymbol{x}^T \cdot \mathbf{B}^{1/2} \cdot \boldsymbol{x} \geq 0$ for all $\boldsymbol{x}$. If $\mathbf{B}$ is itself positive and symmetric, as is the case for covariance matrices, then the positive square root exists and is unique.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

We now make another change of variables:

$$\delta \boldsymbol{x}' = \mathbf{V}_R^T \cdot \delta \tilde{\boldsymbol{x}} \tag{C7}$$

$$\delta \boldsymbol{y}' = \mathbf{V}_L^T \cdot \delta \tilde{\boldsymbol{y}} \tag{C8}$$

$$\mathbf{H}' = \mathbf{V}_L^T \cdot \tilde{\mathbf{H}} \cdot \mathbf{V}_R. \tag{C9}$$

The matrices $\mathbf{V}_L$ and $\mathbf{V}_R$ are orthogonal, i.e., $\mathbf{V}_L^T \cdot \mathbf{V}_L = \mathbf{1}$, and similarly for $\mathbf{V}_R$. Thus, substitution of (C7)–(C9) into (C5) yields

$$J = \frac{1}{2} \delta \boldsymbol{x}'^T \cdot \delta \boldsymbol{x}' + \frac{1}{2} \left(\mathbf{H}' \cdot \delta \boldsymbol{x}' - \delta \boldsymbol{y}' \right)^T \cdot \left(\mathbf{H}' \cdot \delta \boldsymbol{x}' - \delta \boldsymbol{y}' \right). \tag{C10}$$

Evidently, the transformation given in (C7)–(C9) preserves the diagonality of the background and observation error covariance matrices. What about the covariance matrix $\langle \delta \boldsymbol{y}' \cdot \delta \boldsymbol{y}'^T \rangle$ in the new basis? Using $\boldsymbol{\epsilon}_o' = \mathbf{V}_L^T \cdot \tilde{\boldsymbol{\epsilon}}_o = \mathbf{V}_L^T \cdot \mathbf{R}^{-1/2} \cdot \boldsymbol{\epsilon}_o$, as well as Eqs. (C1), (C2)–(C4), and (C7)–(C9), we obtain $\langle \delta \boldsymbol{y}' \cdot \delta \boldsymbol{y}'^T \rangle = \mathbf{H}' \cdot \mathbf{H}'^T + \mathbf{1}$. The contribution of the background error covariances in this coordinate system is $\mathbf{H}' \cdot \mathbf{H}'^T$, which is a diagonal matrix. This becomes clear from Eqs. (C6) and (C9), which yields

$$\mathbf{H}' \cdot \mathbf{H}'^T = \mathbf{W} \cdot \mathbf{W}^T, \tag{C11}$$

which is a $(m \times m)$ diagonal matrix. Thus in this coordinate system we can readily compare the diagonal elements of the transformed background error covariance matrix $\mathbf{H}' \cdot \mathbf{H}'^T$ to the diagonal (unit) elements of the observation error covariance matrix $\mathbf{1}$. Roughly, those singular values w_i on the diagonal of $\mathbf{W}$ that are larger than unity correspond to model variables $\delta x_i'$ that can be controlled by the measurements. Those singular values smaller than unity correspond to model variables that are only related to noise.

In the above discussion we relied on plausibility arguments. We mention that there are more systematic ways of approaching the problem. Here we merely state some key results without going into details. The interested reader is referred to Rodgers (2000). However, in all approaches the main quantities of interest are always the singular values of the matrix $\mathbf{R}^{-1/2} \cdot \mathbf{H} \cdot \mathbf{B}^{1/2}$.

One can compute the number of signal degrees of freedom N_s from the expectation value of J_b in Eq. (B8). The result can be expressed in terms of the singular values w_i of the transformed observation operator in Eq. (C6):

$$N_s = \sum_i w_i^2 / (1 + w_i^2). \tag{C12}$$

Another approach is based on information theory. Given a system described by a probability distribution function $P(x)$, one defines the Shannon-entropy

$$S(P) = -\int P(x) \ln \left(\frac{P(x)}{P_0(x)} \right) \mathrm{d}x, \tag{C13}$$

where P_0 is a normalisation factor needed to make the argument of the logarithm dimensionless. A decrease in entropy expresses an increase in our knowledge of the system. For instance, if we initially describe the system by $P_i(x)$, and, after taking measurements, by $P_f(x)$, then the measurement process has changed the entropy by an amount

$$H = S(P_i) - S(P_f). \tag{C14}$$

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

In our case, we assume that all errors are normally distributed. In that case, one can show that

$$H = \frac{1}{2}\sum_i \ln(1+w_i^2). \tag{C15}$$

H can be interpreted as a measure for the information content of a set of measurements.

Our findings so far suggest a general strategy for how to optimise the amount of information that we can extract from measurements. First, we need to compute the singular value decomposition in Eq. (C6), as well as the transformation given in (C2) and (C7), which we can summarise as

$$\delta\boldsymbol{x}' = \mathbf{V}_R^T \cdot \mathbf{B}^{-1/2} \cdot \delta\boldsymbol{x}. \tag{C16}$$

Then we want to formulate the minimisation of the costfunction in such a way that only those components of $\delta\boldsymbol{x}'$ are adjusted by the assimilation algorithm that correspond to the largest singular values of the matrix $\mathbf{W}$ in (C6). All other elements of $\delta\boldsymbol{x}'$ should be left alone. In other words, we want to constrain the minimisation of the costfunction to the subspace of the signal degrees of freedom of the state vector. Thus, in order to implement this idea, we first need to discuss how to incorporate constraints into the theory.

Appendix D: Minimisation of the costfunction with constraints

In the minimisation of the costfunction all elements of the control vector $\delta\boldsymbol{x}$ are independently adjusted until the minimum of J is found. This may not be a prudent approach if the information contained in the observations is insufficient to constrain all model variables. In such case one should introduce constraints that reduce the number of independent control variables. However, this needs to be done in a clever way; the goal is to neither under-use the measurements (thus wasting available information), nor to over-use them (thus assimilating noise).

For reasons we will explain later we formulate the constraints as weak conditions. However, for didactic reasons as well as for the sake of completeness, we will also mention how to formulate constraints as strong conditions.

D1 Minimisation of the costfunction with strong constraints

Given k constraints in the form $g_i(\delta\boldsymbol{x})$=0, $i = 1,\ldots,k$, the most general way of finding the minimum of $J(\delta\boldsymbol{x})$ under the constraints g_i is the method of Lagrange multipliers. More specifically, one introduces k Lagrange multipliers $\lambda_1,\ldots,\lambda_k$ and defines the function

$$L(\delta x_1,\ldots,\delta x_n,\lambda_1,\ldots,\lambda_k) = J(\delta x_1,\ldots,\delta x_n) + \sum_{i=1}^{k} \lambda_i g_i(\delta x_1,\ldots,\delta x_n); \tag{D1}$$

then one solves the minimisation problem

$$\nabla L(\delta x_1,\ldots,\delta x_n,\lambda_1,\ldots,\lambda_k) = \mathbf{0}, \tag{D2}$$

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

where $\nabla = \nabla_{\delta x_1,\ldots,\delta x_n,\lambda_1,\ldots,\lambda_k}$ is now a $(n+k)$-dimensional gradient operator. Note that in this general formulation of the problem the constraints can even be nonlinear. We are specifically interested in linear constraints, which can be expressed in the form $\mathbf{G}\cdot\delta\boldsymbol{x} = \mathbf{0}$. Then the constrained minimisation problem becomes

$$L(\delta\boldsymbol{x},\boldsymbol{\lambda}) = J(\delta\boldsymbol{x}) + \boldsymbol{\lambda}^T\cdot\mathbf{G}\cdot\delta\boldsymbol{x} \tag{D3}$$

$$\nabla_{\delta\boldsymbol{x},\boldsymbol{\lambda}} L(\delta\boldsymbol{x},\boldsymbol{\lambda}) = \begin{pmatrix} \nabla_{\delta\boldsymbol{x}} J(\delta\boldsymbol{x}) + \boldsymbol{\lambda}^T\cdot\mathbf{G} \\ \mathbf{G}\cdot\delta\boldsymbol{x} \end{pmatrix} = \mathbf{0}. \tag{D4}$$

Compared to the unconstrained minimisation problem, the introduction of k constraints has increased the dimension of the problem from n to $n+k$. Naively, one may have expected that the dimension would, on the contrary, be reduced to $n-k$. This is indeed the case if the constraints are linear, and if the function J is quadratic, as is the case in Eqs. (B7)–(B9). To see this, let us first write those equations more concisely in the form

$$J = \frac{1}{2}\left(\delta\boldsymbol{x}^T\cdot\mathbf{Q}_1\cdot\delta\boldsymbol{x} + \mathbf{Q}_2^T\cdot\delta\boldsymbol{x} + \delta\boldsymbol{x}^T\cdot\mathbf{Q}_2 + \mathbf{Q}_3\right) \tag{D5}$$

$$\mathbf{Q}_1 = \mathbf{B}^{-1} + \mathbf{H}^T\cdot\mathbf{R}^{-1}\cdot\mathbf{H} \tag{D6}$$

$$\mathbf{Q}_2 = \mathbf{H}^T\cdot\mathbf{R}^{-1}\cdot(\hat{H}(\boldsymbol{x}_b) - \boldsymbol{y}) \tag{D7}$$

$$\mathbf{Q}_3 = (\hat{H}(\boldsymbol{x}_b) - \boldsymbol{y})^T\cdot\mathbf{R}^{-1}\cdot(\hat{H}(\boldsymbol{x}_b) - \boldsymbol{y}). \tag{D8}$$

(Note that the covariance matrices and their inverses are symmetric, i.e., $\mathbf{R}^T = \mathbf{R}$, etc.) The unconstrained minimisation problem requires us to solve the equation $\nabla J = \mathbf{Q}_1\cdot\delta\boldsymbol{x} + \mathbf{Q}_2 = \mathbf{0}$. Now we want to minimise the costfunction subject to the the linear constraints

$$\mathbf{G}\cdot\delta\boldsymbol{x} = \mathbf{0}, \tag{D9}$$

where $\mathbf{G}$ is a $(k\times n)$-matrix, $\delta\boldsymbol{x}$ is an n-vector, and $\mathbf{0}$ is the null-vector in $\mathbb{R}^k$. Let us denote the kernel[10] of $\mathbf{G}$ by ker($\mathbf{G}$). Let further $\boldsymbol{z}_1,\ldots,\boldsymbol{z}_{n-k}$ denote a basis of ker($\mathbf{G}$). We define the $(n\times(n-k))$-matrix

$$\mathbf{Z} = \begin{pmatrix} \boldsymbol{z}_1 & \cdots & \boldsymbol{z}_{n-k} \end{pmatrix} \tag{D10}$$

the column vectors of which are just the basis vectors of ker($\mathbf{G}$). Obviously, $\mathbf{G}\cdot\mathbf{Z} = \mathbf{0}$, where $\mathbf{0}$ denotes the $((k\times(n-k))$-nullmatrix. If $\delta\boldsymbol{x}$ is a vector in $\mathbb{R}^n$ for which there exists a vector $\boldsymbol{\xi}\in\mathbb{R}^{n-k}$ such that $\mathbf{Z}\cdot\boldsymbol{\xi} = \delta\boldsymbol{x}$, then we automatically have $\mathbf{G}\cdot\delta\boldsymbol{x} = \mathbf{0}$, i.e., $\delta\boldsymbol{x}$ satisfies the linear constraints. Thus we can formulate the constrained minimisation problem by substitution of $\delta\boldsymbol{x} = \mathbf{Z}\cdot\boldsymbol{\xi}$ into Eq. (D5), which yields

$$J = \frac{1}{2}\left(\boldsymbol{\xi}^T\cdot\mathbf{Z}^T\cdot\mathbf{Q}_1\cdot\mathbf{Z}\cdot\boldsymbol{\xi} + \mathbf{Q}_2^T\cdot\mathbf{Z}\cdot\boldsymbol{\xi} + \boldsymbol{\xi}^T\cdot\mathbf{Z}^T\cdot\mathbf{Q}_2 + \mathbf{Q}_3\right) \tag{D11}$$

$$\mathbf{0} = \nabla J = \mathbf{Z}^T\cdot\mathbf{Q}_1\cdot\mathbf{Z}\cdot\boldsymbol{\xi} + \mathbf{Z}^T\cdot\mathbf{Q}_2. \tag{D12}$$

Thus we have reduced the $(n+k)$-dimensional constrained minimisation problem given in Eq. (D4) to a problem consisting of the following two steps.

[10] The *kernel* or *nullspace* of a matrix is the set of all vectors $\boldsymbol{z}$ such that $\mathbf{G}\cdot\boldsymbol{z} = \mathbf{0}$. The kernel is a subspace of the full vector space $\mathbb{R}^n$ with dim ker($\mathbf{G}$)$= n-k$.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

1. Determine a basis of the nullspace ker($\mathbf{G}$); this yields the matrix $\mathbf{Z}$.

2. Solve the unconstrained $(n-k)$-dimensional optimisation problem given in Eq. (D12). From the $(n-k)$-vector $\boldsymbol{\xi}$ that minimises the costfunction in (D11), we then obtain the solution $\delta\boldsymbol{x} = \mathbf{Z}\cdot\boldsymbol{\xi}$ that minimises the costfunction in (D5) subject to the constraint (D9).

D2 Minimisation of the costfunction with weak constraints

In the approach described in the previous section the solution satisfies the constraints exactly. Therefore, this approach is known as the minimisation of the costfunction with *strong constraints*. In the *weak-constraint* approach the constraints only need to be satisfied within specified error bounds.

The formulation of the weak-constraint approach is conceptually quite simple. One incorporates the constraints by adding an extra term to the costfunction (B7), i.e.

$$J = J_b + J_o + J_G \tag{D13}$$

$$J_G = \frac{1}{2}\delta\boldsymbol{x}^T\cdot\mathbf{G}^T\cdot\mathbf{B}_G^{-1}\cdot\mathbf{G}\cdot\delta\boldsymbol{x}, \tag{D14}$$

which also gives an extra term in the gradient of the costfunction,

$$\nabla J_G = \mathbf{G}^T\cdot\mathbf{B}_G^{-1}\cdot\mathbf{G}\cdot\delta\boldsymbol{x}. \tag{D15}$$

We will assume that the matrix $\mathbf{B}_G$=diag$(\sigma_1^G,\ldots,\sigma_k^G)$ is diagonal, where k is the number of constraints. The "error variances" σ_i^G along the diagonal of $\mathbf{B}_G$ allow us to fine-tune the influence of each constraint on the solution. If σ_i^G is small, then the ith constraint is relatively strong, and vice versa. The choice of these variances is a matter of experimenting and tuning. Typically, if the σ_i^G are made too large, then there is a risk that the minimisation algorithm ignores the constraints all together. In that case the solution will be very similar to the unconstrained solution. On the other hand, if the σ_i^G are made too small, then J_G can make the dominant contribution to J. In that case, there is a risk that the minimisation routine largely ignores the observations and returns a solution that lies quite close to the background estimate.

D3 Constraints designed for making optimum use of the information contained in the observations

We now want to incorporate the results of Section C into the variational data assimilation method. More specifically, we want to formulate weak constraints, Eq. (D14), based on the singular values of the observation operator in Eq. (C6). To this end, we make the change of variables given in Eq. (C16). We assume, without loss of generality, that the first ℓ singular values are greater than unity. Thus we only want to use the corresponding components $\delta x_1',\ldots,\delta x_\ell'$ as independent control variables in the 3DVAR algorithm, while the remaining components remain unchanged, at least approximately within specified error bounds.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

If we were to formulate this requirement as a strong constraint, as in Eq. (D9), then it would take the form

$$\delta \boldsymbol{x}' = \mathbf{V}_R^T \cdot \mathbf{B}^{-1/2} \cdot \delta \boldsymbol{x} = \begin{pmatrix} \delta x'_1 \\ \vdots \\ \delta x'_\ell \\ 0 \\ \vdots \\ 0 \end{pmatrix}. \tag{D16}$$

Thus the matrix expressing the constraints is given by $\mathbf{G} = \mathbf{V}_R^T \cdot \mathbf{B}^{-1/2}$, which is a $(n \times n)$ matrix.

The weak constraint approach is, arguably, more suitable in our case. We have, in the preceding text, frequently used the terms *signal degrees of freedom* and *noise degrees of freedom*. Although it was conceptually useful to make this distinction, it is important to stress that there is no sharp boundary between the two. Rather, there is a smooth transition from singular values $w_1 > w_2 > \cdots > w_\ell \geq 1$ to singular values $1 > w_{\ell+1} > w_{\ell+2} > \cdots > w_k$ ($k = \min\{n, m\}$). For this reason we choose to formulate the constraints as weak constraints. This allows us to make a smooth transition from free to constrained control variables, where the transition from one regime to the other can be controlled by the singular values.

In order to apply the weak-constraint approach, we need to substitute the constraint-matrix $\mathbf{G} = \mathbf{V}_R^T \cdot \mathbf{B}^{-1/2}$ into Eq. (D14), which yields

$$J_G = \frac{1}{2} \delta \boldsymbol{x}^T \cdot \mathbf{B}^{-1/2} \cdot \mathbf{V}_R \cdot \mathbf{B}_G^{-1} \cdot \mathbf{V}_R^T \cdot \mathbf{B}^{-1/2} \cdot \delta \boldsymbol{x}, \tag{D17}$$

where $\mathbf{B}_G$ is a $(n \times n)$ matrix. We want to set up this matrix in such a way that we obtain a smooth transition from freely adaptable control variables $\delta x'_1, \ldots \delta x'_\ell$ to increasingly constrained variables $\delta x'_{\ell+1}, \ldots \delta x'_k, \ldots, \delta x_n$. One possible choice of the matrix $\mathbf{B}_G$, which is suggested by Eq. (C11), would be

$$\mathbf{B}_G = \sigma_G \, \mathrm{diag}(w_1^2, w_2^2, \ldots, w_\ell^2, \ldots, w_k^2, c, \ldots, c), \tag{D18}$$

where σ_G is a free scaling factor, and where the last $n - k$ diagonal elements are equal to a constant c chosen to be much smaller than w_k^2. Another possible choice would be

$$\mathbf{B}_G = \sigma_G \cdot \mathrm{diag}(\lambda_1, \lambda_2, \ldots, \lambda_\ell, \ldots, \lambda_k, c', \ldots, c'), \tag{D19}$$

$$\lambda_i = w_i^2/(1 + w_i^2), \tag{D20}$$

where $c' \ll \lambda_k$. This ansatz is suggested by Eq. (C12), i.e., each of the elements $\delta x'_1, \ldots \delta x'_k$ is weighted with its corresponding contribution to the number of signal degrees of freedom. It turns out that Eq. (D18) gives a relatively sharp transition from unconstrained to constrained model variables, while Eq. (D19) gives a very gentle transition. Another ansatz that lies in between these two extremes would be

$$\mathbf{B}_G = \sigma_G \, \mathrm{diag}(w_1, w_2, \ldots, w_\ell, \ldots, w_k, c'', \ldots, c''), \tag{D21}$$

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

where $c'' \ll w_k$.

Despite the mathematical foundation of this approach we are left with some room for experimentation in the formulation of the matrix $\mathbf{B}_G$. It is a matter of experience to test different approaches and select the one that proves to be most suited.

Appendix E: Practical aspects of the implementation

One of the main practical problems is the dimension n of the model space. The grid-size is typically on the order $N_x \times N_y \times N_z \sim 100 \times 100 \times 10$, and the number of aerosol components is on the order of $N_c \sim$10–100. Hence the dimension of the model space is $n \sim 10^6$–10^7. In our case, the matrix $\tilde{\mathbf{H}}$ in (C6) is a $(m \times n)$ matrix. To numerically perform a singular value decomposition of such a large matrix would be a formidable task.

In variational data assimilation we encounter a similar problem in the inversion of the matrix $\mathbf{B}$. In our 3DVAR code this problem is alleviated by using a so-called spectral formulation. The idea is to make a Fourier-transformation in the horizontal coordinates and to assume that all horizontal error correlations are homogeneous and isotropic. Under these assumptions one obtains one background error covariance matrix for each horizontal wavenumber; each of these matrices has dimension $N_z \times N_c \sim 10^3$–10^4. This can further be reduced to about 10^2 by making a reduced eigenvalue diagonalisation. The details are explained in Kahnert (2008).

In our case we are primarily interested in constraining the chemical components. Therefore, the best solution for us is to simply restrict ourselves to the chemical subspace. To this end, we select a grid-point[11] (i,j,l) and consider the reduced background error covariance matrix $\mathbf{B}^0$ with components $B^0_{k;k'} = B_{i,j,l,k;i,j,l,k'}$, where $k,k' = 1,\ldots,N_c$ Similarly, we consider reduced matrices $\mathbf{R}^0$, $\mathbf{H}^0$, and $\tilde{\mathbf{H}}^0 = (\mathbf{R}^0)^{-1/2} \cdot \mathbf{H}^0 \cdot \mathbf{B}^0$, and we numerically compute the singular value decomposition of the latter. From this we obtain, for each level l, a constraint matrix $(\mathbf{V}^0_R)^T \cdot (\mathbf{B}^0)^{-1/2}$ analogous to the one in Eq. (D16), but reduced to the chemical subspace. We then simply apply these chemical constraints throughout the horizontal domain (again, assuming horizontal homogeneity).

Another aspect concerns the positive square root of the background error covariance matrix, which appears in essential parts of the theory, namely, in Eqs. (C6) and (D16). In theoretical developments it is, arguably, didactically expedient to work with the matrix $\mathbf{B}^{1/2}$. But in practice there are numerically more efficient formulations. One such approach is discussed in Kahnert (2008) in the context of a spectral formulation of the variational method. In our present problem we employ the Cholesky decomposition[12] of the B-matrix,

$$\mathbf{B} = \mathbf{C}_u^T \cdot \mathbf{C}_u, \tag{E1}$$

where $\mathbf{C}_u$ is an upper triangular matrix. Thus the actual algorithm we used for formulating the constrained minimisation of the costfunction is obtained by replacing in the preceding formulas all incidences of the matrix $\mathbf{B}^{1/2}$ with the matrix $\mathbf{C}_u^T$ (and, similarly, by replacing the inverse matrix $\mathbf{B}^{-1/2}$ by the inverse of the Cholesky factor, $\mathbf{C}_u^{-T}$).

[11] Since the error correlations are assumed to be homogeneous in space any point in the horizontal direction will do.

[12] The Cholesky decomposition is, essentially, a special case of a LU-decomposition, which applies to symmetric real (or Hermitian complex), positive definite matrices.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

Author contributions. MK worked with the theoretical developments and and numerical implementation, EA performed the testing of the method.

Acknowledgements. This work was funded by the Swedish National Space Board through the OSCES project (no. 101/13).

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

Atmospheric Chemistry and Physics Discussions

References

Andersson, C., Langner, J., and Bergström, R.: Interannual variation and trends in air pollution over Europe due to climate variability during 1958-2001 simulated with a regional CTM coupled to the ERA40 reanalysis, Tellus, 59B, 77–98, 2007.

Andersson, C., Bergström, R., Bennet, C., Robertson, L., Thomas, M., Korhonen, H., Lehtinen, K. H. J., and Kokkola, H.: MATCH-SALSA — Multi-scale Atmospheric Transport and CHemistry model coupled to the SALSA aerosol microphysics model — Part 1: Model description and evaluation, Geosci. Model Dev., 8, 171–189, 2015.

Andersson, E. and Kahnert, M.: Coupling aerosol optics to the MATCH (v5. 5.0) chemical transport model and the SALSA (v1) aerosol microphysics module, Geosci. Model Dev., 9, 1803–1826, 2016.

Benedetti, A., Morcrette, M. J.-J., Boucher, O., Dethof, A., Engelen, R. J., Huneeus, M. F. H. F. N., Jones, L., andS. Kinne, J. W. K., Mangold, A., Razinger, M., Simmons, A. J., and Suttie, M.: Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: 2. Data assimilation, J. Geophys. Res., 114, D13 205, 2009.

Bi, L., Yang, P., Kattawar, G., and Kahn, R.: Modeling optical properties of mineral aerosol particles by using nonsymmetric hexahedra, Appl. Opt., 49, 334–342, 2010.

Burton, S. P., Chemyakin, E., Liu, X., Knobelspiesse, K., Stamnes, S., Sawamura, P., Moore, R. H., Hostetler, C. A., and Ferrare, R. A.: Information Content and Sensitivity of the $3\alpha + 2\beta$ Lidar Measurement System for Aerosol Microphysical Retrievals, Atmos. Meas. Tech. Discuss., 2016, doi:10.5194/amt-2016-240, 2016.

Foltescu, V., Pryor, S. C., and Bennet, C.: Sea salt generation, dispersion and removal on the regional scale, Atmos. Environ., 39, 2123–2133, 2005.

Fuller, K. A. and Mackowski, D. W.: Electromagnetic scattering by compounded spherical particles, in: Light scattering by nonspherical particles, edited by Mishchenko, M. I., Hovenier, J. W., and Travis, L. D., pp. 226–273, Academic Press, San Diego, 2000.

Kahnert, F. M.: Reproducing the optical properties of fine desert dust aerosols using ensembles of simple model particles, J. Quant. Spectrosc. Radiat. Transfer, 85, 231–249, 2004.

Kahnert, M.: Variational data analysis of aerosol species in a regional CTM: background error covariance constraint and aerosol optical observation operators, Tellus, 60B, 753–770, 2008.

Kahnert, M.: On the observability of chemical and physical aerosol properties by optical observations: Inverse modelling with variational data assimilation, Tellus, 61B, 747–755, 2009.

Kahnert, M.: Modelling radiometric properties of inhomogeneous mineral dust particles: Applicability and limitations of effective medium theories, J. Quant. Spectrosc. Radiat. Transfer, 152, 16–27, 2015.

Kahnert, M. and Devasthale, A.: Black carbon fractal morphology and short-wave radiative impact: a modelling study, Atmos. Chem. Phys., 11, 11 745–11 759, 2011.

Kahnert, M., Nousiainen, T., Lindqvist, H., and Ebert, M.: Optical properties of light absorbing carbon aggregates mixed with sulfate: assessment of different model geometries for climate forcing calculations, Opt. Express, 20, 10 042–10 058, 2012a.

Kahnert, M., Nousiainen, T., Thomas, M. A., and Tyynelä, J.: Light scattering by particles with small-scale surface roughness: comparison of four classes of model geometries, J. Quant. Spectrosc. Radiat. Transfer, 113, 2356–2367, 2012b.

Kahnert, M., Nousiainen, T., and Lindqvist, H.: Models for integrated and differential scattering optical properties of encapsulated light absorbing carbon aggregates, Opt. Express, 21, 7974–7992, 2013.

639 Kahnert, M., Nousiainen, T., and Markkanen, J.: Morphological models for inhomogeneous particles: light scattering by aerosols, cometary
640 dust, and living cells, in: Light Scattering Reviews 11, edited by Kokhanovsky, A., Springer, Berlin, p. 299–339, 2016.
641 Khade, V. M., Hansen, J. A., Reid, J. S., and Westphal, D. L.: Ensemble filter based estimation of spatially distributed parameters in a
642 mesoscale dust model: experiments with simulated and real data, Atmos. Chem. and Phys., 13, 3481–3500, 2013.
643 Kupiainen, K. and Klimont, Z.: Primary emissions of submicron and carbonaceaous particles in Europe and the potential for their control,
644 Tech. Rep. IR-04-079, IIASA, Laxenburg, Austria, 2004.
645 Kupiainen, K. and Klimont, Z.: Primary emissions of fine carbonaceous particles in Europe, Atmos. Environ., 41, 2156–2170, 2007.
646 Kylling, A., Kahnert, M., Lindqvist, H., and Nousiainen, T.: Volcanic ash infrared signature: porous non-spherical ash particle shapes com-
647 pared to homogeneous spherical ash particles, Atmos. Meas. Tech., 7, 919–929, 2014.
648 Lindqvist, H., Muinonen, K., and Nousiainen, T.: Light scattering by coated Gaussian and aggregate particles, J. Quant. Spectrosc. Radiat.
649 Transfer, 110, 1398–1410, doi:10.1016/j.jqsrt.2009.01.015, 2009.
650 Lindqvist, H., Nousiainen, T., Zubko, E., and Muñoz, O.: Optical modeling of vesicular volcanic ash particles, J. Quant. Spectrosc. Radiat.
651 Transfer, 112, 1871–1880, 2011.
652 Lindqvist, H., Jokinen, O., Kandler, K., Scheuvens, D., and Nousiainen, T.: Single scattering by realistic, inhomogeneous mineral dust
653 particles with stereogrammetric shapes, Atmos. Chem. Phys., 14, 143–157, doi:10.5194/acp-14-143-2014, 2014.
654 Liu, L. and Mishchenko, M. I.: Scattering and radiative properties of complex soot and soot-containing aggregate particles, J. Quant. Spec-
655 trosc. Radiat. Transfer, 106, 262–273, 2007.
656 Liu, Z., Liu, Q., Lin, H.-C., Schwartz, C. S., Lee, Y.-H., and Wang, T.: Three-dimensional variational assimilation of MODIS aerosol optical
657 depth: Implementation and application to a dust storm over East Asia, J. Geophys. Res., 116, D23 206, 2011.
658 McKeen, S., Chung, S. H., Wilczak, J., Grell, G., Djalalova, I., Peckham, S., Gong, W., Bouchet, V., Moffet, R., Tang, Y., Carmichael, G. R.,
659 Mathur, R., and Yu, S.: Evaluation of several PM2.5 forecast models using data collected during the ICARTT/NEAQS 2004 field study, J.
660 of Geophys. Res., 112, d10S20, 2007.
661 Mishchenko, M. I., Travis, L. D., Kahn, R. A., and West, R. A.: Modeling phase functions for dustlike tropospheric aerosols using a shape
662 mixture of randomly oriented polydisperse spheroids, J. Geophys. Res., 102, 16,831–16,847, 1997.
663 Mishchenko, M. I., Cairns, B., Kopp, G., Schueler, C. F., Fafaul, B. A., Hansen, J. E., Hooker, R. J., Itchkawich, T., Maring, H. B., and Travis,
664 L. D.: Accurate Monitoring of Terrestrial Aerosols and Total Solar Irradiance: Introducing the Glory Mission, Bull. Am. Met. Soc., 88,
665 677–691, 2007.
666 Mishchenko, M. I., Dlugach, Z. M., and Zakharova, N. T.: Direct demonstration of the concept of unrestricted effective-medium approxima-
667 tion, Opt. Lett., 39, 3935–3938, 2014.
668 Muinonen, K.: Light scattering by stochastically shaped particles, in: Light scattering by nonspherical particles, edited by Mishchenko, M. I.,
669 Hovenier, J. W., and Travis, L. D., pp. 323–354, Academic Press, San Diego, 2000.
670 Omar, A. H., Winker, D. M., Vaughan, M. A., Hu, Y., Trepte, C. R., Ferrare, R. A., Lee, K.-P., Hostetler, C. A., Kittaka, C., Rogers, R. R.,
671 Kuehn, R. E., and Liu, Z.: The CALIPSO automated aerosol classification and lidar ratio selection algorithm, J. Atmos. and Ocean.
672 Technol., 26, 1994–2014, 2009.
673 Parrish, D. F. and Derber, J. C.: The National Meteorological Centre's spectral statistical interpolation analysis system, Mon. Wea. Rev., 120,
674 1747–1763, 1992.
675 Rodgers, C. D.: Inverse methods for atmospheric sounding, World Scientific, Singapore, 2000.

Atmos. Chem. Phys. Discuss., doi:10.5194/acp-2016-914, 2016
Manuscript under review for journal Atmos. Chem. Phys.
Published: 10 November 2016

676 Rubin, J. I. and Collins, W. D.: Global simulations of aerosol amount and size using MODIS observations assimilated with an Ensemble
677 Kalman Filter, J. Geophys. Res., 119, 12,780–12,806, 2014.
678 Saide, P. E., Charmichael, G. R., Liu, Z., Schwartz, C. S., Lin, H. C., da Silva, A. M., and Hyer, E.: Aerosol optical depth assimilation
679 for a size-resolved sectional model: impacts of observationally constrained, multi-wavelength and fine mode retrievals on regional scale
680 analysis and forecasts, Atmos. Chem. Phys., 13, 10 425–10 444, 2013.
681 Sandu, A., Liao, W., adn D. K. Henze, G. R. C., and Seinfeld, J. H.: Inverse modeling of aerosol dynamics using adjoints: Theoretical and
682 numerical considerations, Aerosol Sci. Technol., 39, 677–694, 2005.
683 Sekiyama, T. T., Tanaka, T. Y., Shimizu, A., and Miyoshi, T.: Data assimilation of CALIPSO aerosol observations, Atmos. Chem. Phys., 10,
684 39–49, 2010.
685 Undén, P., Rontu, L., Järvinen, H., Lynch, P., Calvo, J., Cats, G., Cuxart, J., Eerola, K., Fortelius, C., Garcia-Moya, J. A., Jones, C., Lender-
686 link, G., McDonald, A., McGrath, R., Navascues, B., Nielsen, N. W., Ødegaard, V., Rodriguez, E., Rummukainen, M., Rõõm, R., Sattler,
687 K., Sass, B. H., Savijärvi, H., Schreur, B. W., Sigg, R., The, H., and Tijm, A.: HIRLAM-5 Scientic Documentation, http://www.hirlam.org,
688 2002.
689 Vilaplana, R., Moreno, F., and Molina, A.: Study of the sensitivity of size-averaged scattering matrix elements of nonspherical particles
690 to changes in shape, porosity and refractive index, J. Quant. Spectrosc. Radiat. Transfer, 100, 415–428, doi:10.1016/j.jqsrt.2005.11.068,
691 2006.
692 Zhang, J., Campbell, J. R., Hyer, E. J., Reid, J. S., Westphal, D. L., and Johnson, R. S.: Evaluating the impact of multisensor data assimilation
693 on a global aerosol particle transport model, J. Geophys. Res., 119, 4674–4689, 2014.