Atmos. Chem. Phys. Discuss., 11, 4749–4806, 2011
www.atmos-chem-phys-discuss.net/11/4749/2011/
doi:10.5194/acpd-11-4749-2011

Atmospheric Chemistry and Physics Discussions

Towards inverse modeling of cloud-aerosol interactions – Part 1: A detailed response surface analysis

D. G. Partridge[1], **J. A. Vrugt**[2,3,4], **P. Tunved**[1], **A. M. L. Ekman**[5], **D. Gorea**[4], **and A. Sorooshian**[6,7]

[1]Dept. of Applied Environmental Science, Stockholm University, Sweden
[2]The Henry Samueli School of Engineering, Department of Civil and Environmental Engineering, University of California, Irvine, USA
[3]Earth and Environmental Sciences Division, Los Alamos National Laboratory, Mail Stop T003, Los Alamos, NM, 87545, USA
[4]Computational Geo-Ecology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands
[5]Dept. of Meteorology, Stockholm University, Sweden
[6]Dept. of Chemical and Environmental Engineering, The University of Arizona, Tucson, USA
[7]Dept. of Atmospheric Sciences, The University of Arizona, Tucson, USA

Received: 21 December 2010 – Accepted: 31 January 2011 – Published: 9 February 2011

Correspondence to: D. G. Partridge (daniel.partridge@itm.su.se)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

This paper explores the feasibility of inverse modeling to determine cloud-aerosol interactions using a pseudo-adiabatic cloud-parcel model. Two-dimensional plots of the objective function, containing the difference between the measured and model predicted droplet size distribution, are presented for selected pairs of cloud parcel model parameters. From these response surfaces it is shown that the "cloud-aerosol" inverse problem is particularly difficult to solve due to significant parameter interaction, presence of multiple regions of attractions, numerous local optima, and considerable parameter insensitivity. Sensitivity analysis is performed to help select an appropriate objective function that maximizes information retrieval from the measured droplet size distribution to help identify the unknown model parameters. The identifiability of the model parameters will be dependent on the choice of the objective function; including the interstitial aerosol will aid the calibration of parameters describing the smaller aerosol mode. Cloud parcel models that employ a moving-centre based calculation of the droplet size distribution require both the X and Y components of the $dN/d\log d_p$ size distribution function to be explicitly included in the objective function. Other possible improvements identified include an improved representation of the resolution of the region of the size spectrum associated with droplet activation within cloud parcel models, and further development of fixed-sectional cloud models that minimize numerical diffusion. Despite these developments, powerful search algorithms remain necessary to efficiently explore the parameter space and successfully solve the cloud-aerosol inverse problem.

1 Introduction

A major challenge currently facing the cloud-aerosol research community includes untangling the relative importance of size and composition for the activation of aerosol particles into cloud droplets (McFiggans et al., 2006). The difficulties in accurately

ACPD
11, 4749–4806, 2011

representing the development of a cloud droplet number concentration (CDNC) population can be partly attributed to the current state of knowledge regarding which parameters describing the properties of an aerosol distribution are most important for the cloud nucleating ability of aerosol particles (Dusek et al., 2006). This capability is a function of the size of the particle, its composition and mixing state, and the supersaturation in the cloud (Fitzgerald, 1974; Hegg and Larson, 1990; Laaksonen et al., 1998; Feingold, 2003; Conant et al., 2004; Kanakidou et al., 2005; Quinn et al., 2007; Andreae and Rosenfeld, 2008).

Numerous sensitivity studies have been conducted to determine the influence of aerosol properties on the CDNC population (e.g. Nenes et al., 2001; Feingold, 2003; Anttila and Kerminen, 2007 and Reutter et al., 2009). However, most sensitivity studies are local, i.e. investigating parameter sensitivity in the vicinity of their actual values. Global sensitivity analysis considers parameter changes over the entire multidimensional parameter domain (e.g. Pérez et al., 2006; Anttila and Kerminen, 2007). This generally leads to different, but more reliable results because parameter sensitivities in nonlinear models of complex systems typically vary considerably over the feasible space of solutions.

The aerosol indirect effect remains the largest single source of uncertainty in current estimates of the total anthropogenic radiative forcing in climate models (IPCC, 2007). This effect is attributed to an increase in cloud condensation nuclei (CCN) causing an increase in cloud albedo (at fixed cloud liquid water path) resulting in radiative cooling. To constrain the uncertainty of the indirect effect it is necessary to better understand cloud-aerosol interactions; hence it is crucial to improve the understanding of both the physiochemical properties of the aerosol, and the meteorological parameters relevant for the formation and development of clouds over their entire parameter ranges.

During the last decade the representation of the evolution of an aerosol size distribution within cloud parcel models has become increasingly complex, resulting in a significant number of model parameters. In order to improve the representation of cloud-aerosol interactions in global climate models (GCM's), while maintaining computational

ACPD
11, 4749–4806, 2011

efficiency, it is necessary to develop parameterisations that are simple yet accurate. One route by which cloud-aerosol relationships can be probed in detail is to embrace inverse modeling techniques, to scrutinize and evaluate model parameter interactions over a wide range of input and output conditions. In inverse analysis, a given model is calibrated by iteratively changing input values (calibration parameters herein) until the simulated output values match the observed data (i.e., experimental data) as closely and consistently as possible.

Inverse modeling has many practical advantages, but a key advantage is that, when properly implemented, it allows the investigation of parameter sensitivity and correlation on real world measurements. Parameter estimation by inverse modeling also provides a useful approach to diagnose structural inaccuracies in a model, which will appear as a mismatch between optimised parameter values and their directly-observed values. This potential deviation between the model predictions and observations can then be used to understand the size of the model error as a function of the prevailing aerosol/meteorological conditions. In essence, improved interpretation of parameter uncertainty can yield valuable information to enable a better judgement of the limits of our theoretical understanding of droplet activation.

In addition, the development of an inverse approach for cloud-aerosol interactions will highlight the most important model parameters to focus on in future field campaigns. Hence, parameters that appear to be insensitive are deemed unimportant. This helps to prioritise the development of instruments to allow more precise measurements towards certain atmospheric properties, thereby allowing a greater understanding of process-level mechanisms. This will then provide the means to repeat the inverse modeling approach with better measurements. Improving models and measurements in relation to one another in such an iterative manner will help to close the gaps in knowledge of cloud-aerosol interactions. If successful this will allow for an improvement in the validation of cloud-aerosol models, and subsequently the simplification of the representation of cloud droplet development in models, thereby minimising input and computational time.

ACPD
11, 4749–4806, 2011

1.1 An introduction to inverse modeling

Parameter estimation or model calibration is a common problem in many areas of process modeling, both in on-line applications, such as real time optimization, and in off-line applications, such as the modeling of reaction kinetics and phase equilibrium. The goal is to determine values of model parameters that provide the best fit to measured data, generally based on some type of least squares (sum of squared error used herein) or maximum likelihood criterion (Vrugt et al., 2006). Usually, this requires the solution of a nonlinear and frequently non-convex optimization problem.

To further illustrate inverse modeling consider Fig. 1, which presents a schematic representation of the resulting model calibration problem. In this figure, ⊕ represents measurements of the model input (calibration parameters) and response that are subject to measurement errors and uncertainty, and therefore may be different from the true values. Similarly, Φ represents the cloud parcel model, which is at best only an approximation of the underlying system. The label "output" on the y-axis of the plot on the right hand side can represent any response data; in this schematic example we assume this is the number of particles within each bin size, covering a size range that includes both the interstitial aerosol particles and activated cloud droplets.

Using a priori values of the calibration parameters, the predictions of the calibration data by the model shown in this figure (indicated with solid-red line) are behaviourally consistent with the observations (dotted line), but demonstrate a significant bias in the smaller, unactivated interstitial particles part of the size distribution. The common approach is to ascribe this mismatch between model and observations to parameter uncertainty. The goal of model calibration then becomes one of finding those values of the calibration parameters that provide the "best" possible fit to the observed behaviour (for instance solid-green line) (Vrugt et al., 2008).

If once calibrated towards a specific measurement data set, an adiabatic cloud parcel model represents both this and the associated input calibration parameters accurately, it can generally be used for the simulation or prediction of the data set used for model

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

ACPD

11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

calibration, provided that it can be reasonably assumed that the physical characteristics of the climate conditions remain similar.

Mathematically, the model calibration problem depicted in Fig. 1 can be formulated as follows. Let $\tilde{Y} = \Phi(X,\theta)$ denote predictions of, for instance the droplet size distribution, $\tilde{Y} = \{\tilde{y}_1,...,\tilde{y}_n\}$ of the model Φ with observed input variables X and model parameters θ. Let $Y = \{y_1,...,y_n\}$ represent n observations of the droplet size distribution. The difference between the model-predicted and measured droplet size distribution can be represented by the residual vector $\boldsymbol{E}$ as:

$$\boldsymbol{E}(\theta) = G(\tilde{Y}) - G(Y) = \{G(\tilde{y}_1) - G(y_1),...,G(\tilde{y}_n) - G(y_n)\} = \{e_1(\theta),...,e_n(\theta)\} \quad (1)$$

where $G(.)$ allows for various monotonic (such as logarithmic) transformations of the model outputs.

The inverse modeling approach now relies on the estimation of the set of parameters θ such that the measure $\boldsymbol{E}$, commonly called the objective function (herein denoted as "OF", is in some sense forced to be as close to zero as possible. To minimize the OF, the calibration parameters needed to perform the numerical simulation are optimised using an inverse analysis algorithm. If the model fit is not "optimal" the procedure is repeated until the model is optimized.

During the last two decades a great deal of research has been devoted to the successful application of inverse modeling for model calibration in many different areas of scientific research (Vrugt et al., 2004, 2008; Voutilainen and Kaipo, 2005; San Martini et al., 2006; Tomassini et al., 2007; Laine and Tamminen, 2008; Wraith et al., 2009 and Bikowski et al., 2010; Järvinen et al., 2010; Loridan et al., 2010).

Inverse modeling is particularly useful for model parameters whose values cannot be measured directly at the (application) scale of interest. Crump and Seinfeld (1982) and Twomey (1975) are among the first to apply inverse modeling for calibration of aerosol size distribution properties from instruments (Kandlikar and Ramachandran, 1999).

ACPD
11, 4749–4806, 2011

1.2 Goals

Due to numerous reasons, primarily the non-linear nature of cloud-aerosol interactions and the fact that cloud models are "highly parameterised" (Andreae and Rosenfeld, 2008), it is expected to be challenging for any algorithm to successfully locate the "best" calibration parameters for a single cloud case. Before we apply any automatic parameter estimation algorithm, a valuable first step would therefore be to first study the behaviour of the OF in the multi-dimensional parameter space.

The main goal of this paper is to explore the feasibility of inverse modeling for the calibration of a pseudo-adiabatic cloud parcel model using artificial measurements obtained from this model. The use of synthetic data has important practical advantages as the true parameter values are known. To achieve this goal response surfaces will be generated that plot the value of the OF for selected pairs of the calibration parameters. These surfaces will be presented for a number of different parameter combinations. Response surface analysis has been used in many fields to study the behaviour of the OF (Sorooshian and Arfi, 1982; Toorman et al., 1992; Šimúnek et al., 1998 and Vrugt et al., 2001).

Response surfaces will contain valuable information about the variation of the OF across the parameter space. To the authors' knowledge inverse modeling has never been performed for an adiabatic cloud parcel model, therefore, the response surfaces will be used to demonstrate why it is important to first analyse the posedness of the inverse problem, to ascertain the feasibility of linking such a model to an automatic search algorithm. Therefore, the aim of this first paper is to calculate response surfaces for the input aerosol, meteorological, and chemical parameters for an adiabatic cloud parcel model.

This will hopefully facilitate the selection of the best possible setup for subsequent analysis in which the aim is to automatically identify the model parameters using an automatic calibration algorithm, and also the uncertainty associated with each parameter.

ACPD
11, 4749–4806, 2011

The paper is organized as follows: first the model and generation of a set of synthetic measurements will be described. Secondly the selection of an appropriate OF will be discussed in terms of the identifiability of model calibration parameters and information content of the calibration data. Implications of the standard output of particle size distributions from cloud parcel models will then be highlighted, followed by the main analysis of the posedness of the inverse problem through response surfaces. Calibration parameter combinations that could be easily identified by automatic search algorithms, illustrating clear optimal solutions to the inverse problem will first be presented. Conversely properties of a response that would severely hamper a search algorithm's ability to locate a single true solution to the inverse problem will be shown. This will be further investigated by a sensitivity analysis of the calibration data to changes in calibration parameters, and concurrently the choice of an "optimal" OF will be discussed. Finally an example of the need for robust search algorithms will be presented by overlaying the results of a simulation from an automatic search algorithm versus a standard Latin Hypercube Monte Carlo (MC) simulation onto response surfaces.

2 Materials and methods

2.1 Adiabatic cloud parcel model

In the present study, we use a pseudo-adiabatic cloud parcel model. Such a cloud parcel model offers the possibility to study the relationship between key input parameters with respect to different output variables in a computationally efficient way. Adiabatic cloud parcel models have been used successfully as inter-comparison tools with field measurements to estimate the impact of aerosol size/composition for liquid clouds (Ayers and Larson, 1990; Nenes et al., 2002; Hsieh et al., 2009). We posit that an adiabatic cloud parcel model is a sensible trade-off between processes accounted for, and computational speed necessary to perform the thousands of simulations required for the Markov Chain Monte Carlo simulation (MCMC) of a single cloud case.

ACPD
11, 4749–4806, 2011

The cloud parcel model used in this study simulates the pseudo-adiabatic ascent of an air parcel, condensation and evaporation of water vapor on aerosols, particle activation, condensational growth, collision and coalescence between droplets, and aqueous phase sulfur chemistry. It can be initialized with aerosol populations consisting of one or more internal and/or external mixtures of $(NH_4)_mH_{2-m}SO_4$ ($m = 0, 1$ or 2), organic carbon (OC), black carbon (BC), mineral dust and sea salt. The model is currently set up so that the aerosol is represented as an internal mixture of compounds. This should be a reasonable assumption for aged aerosol in the marine environment. For rural continental conditions this assumption may be less realistic, but we keep the mixing state consistent to allow a more straight-forward analysis. The size distribution of each external aerosol type in this study are described by two lognormal modes, which are defined by the geometric standard deviation of the mode, total particle concentration, and mean diameter. This is represented by 400 size bins; for which the lower and upper limits of the dry aerosol diameters are set to 8 nm and 2000 nm respectively. Each aerosol bin is characterized by a dry and a wet particle radius, the latter of which employs a moving centre approach so that the wet radii are continuously modified by condensation or evaporation of water.

The ascending parcel equations are from Pruppacher and Klett (1997). Aerosol activation and condensation/evaporation of water are calculated according to the Köhler equation (Köhler, 1936) and parameterized according to Hänel (1987). The Köhler equation was reformulated in terms of the solute concentrations (Roelofs, 1992) to allow for modifications of the Raoult term by chemical processes, e.g., dissolution of gaseous HNO_3 or partial dissolution of aerosol organic matter. Organic matter is represented by dicarboxylic acids and its effect on the surface tension is accounted for. For more information on the model and references the reader is referred to Roelofs and Jongen (2004).

For this study the model is run with the processes of collision-coalescence and entrainment turned off, and the updraft velocity kept constant. This maintains computational speed and is deemed reasonable for this study, in which the focus is on CCN

ACPD
11, 4749–4806, 2011

activation and condensational growth above the cloud base. In reality clouds are much more complex in nature, including processes such as mixing, variable updrafts and entrainment. The approach of this study will therefore only probe an idealised system; nevertheless, it addresses the most critical cloud/droplet forming mechanisms.

2.2 Artificial measurements

To benchmark our inverse modeling approach, it is useful to start our analysis with numerically generated data sets (termed "synthetic" data herein). This is beneficial, since the exact values of the aerosol/chemistry/meteorological calibration parameters are known precisely a priori (Toorman et al., 1992; Šimúnek et al., 1998), which enables us to analyze possible discrepancies between the estimated and true parameters and thus test our methodology. Using real-world data, parameters will diverge from their true values because of model and measurement error. It is critical that first the optimal solution can be accurately and efficiently found using synthetic data, or else solutions found when calibrated against real world measurements could be misleading. Therefore, we first work with a synthetic droplet size distribution simulated with an adiabatic cloud parcel model. This artificially generated data set serves as the "truth", and was created using known input parameter values from the literature.

2.2.1 Calibration input parameters

To test a wide range of input aerosol size distributions, data from two distinctively different environments were used. For marine conditions we utilized the aerosol size distribution measurements compiled by Heintzenberg et al. (2000). Measurements from the well-established SMEAR II station at Hyytiälä, as documented by Tunved et al. (2005) are used to represent a rural continental environment. The mean values of these observations are presented in Table 1. For each environment the upper and lower parameter limits for the aerosol size distribution were selected using the statistics

ACPD
11, 4749–4806, 2011

available in Heintzenberg et al. (2000) and Tunved et al. (2005) respectively. These ranges were rounded or extended slightly (see Table 1) to make sure that our upper and lower bounds appropriately encompass the large variety of conditions observed in the aerosol/climatic system. The aerosol size distributions for the marine and rural conditions used to generate the synthetic data are illustrated in Fig. 2. For the updraft velocity the base value was chosen to be 0.5 $\mathrm{m\,s^{-1}}$, with a lower limit of 0.1 $\mathrm{m\,s^{-1}}$ and upper limit of 3 $\mathrm{m\,s^{-1}}$, to represent a wide range of meteorological conditions. In all the calculations reported herein, the cloud depth is fixed to 200 m. In reality however, summertime marine shallow cumulus or stratocumulus clouds can extend beyond 200 m (Lu et al., 2007); yet our approach ensures computational tractability, as each model simulation can be performed within a reasonable time frame. The base updraft and cloud depth were kept fixed for the marine and rural continental environment to allow easier comparison of the behaviour of the less easily measured calibration parameters (i.e. the parameters describing the lognormal aerosol size distribution). In the pseudo-adiabatic cloud parcel model it is of course necessary to define an ambient temperature profile. In the absence of detailed prior information, we use a temperature profile measured from MASE II measurement campaign (A. Sorooshian, personal communication, 2010).

Ideally the chemistry would be represented by the mass fractions of multiple different aerosol components. This however, is not conducive for two-dimensional parameter analysis, since in order to vary parameter values independently from one another it is not possible to include mass fractions for multiple components, whilst simultaneously keeping all other components fixed and achieving unity for the sum of all mass fractions (MF). The chemistry therefore, was defined as a two-component scheme consisting of either a soluble component, ammonium bisulphate (NH_4HSO_4), or an insoluble component, black carbon (BC). In the calculation of the response surfaces (cf. Sect. 2.4 onwards) we only allow one of the chemistry components to vary. The soluble component is allowed to vary between 0.05 and 1, and the mass fraction (MF) of the insoluble component is calculated directly from the equation, $\mathrm{MF_{INSOLUBLE}} = 1 - \mathrm{MF_{SOLUBLE}}$. The

ACPD
11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

true values of the MF were selected rather arbitrarily, with a higher $MF_{SOLUBLE}$ for the marine environment compared to rural continental as would be expected in reality (Lance et al., 2004).

A log-transformation was applied to the number of aerosol particles in mode one and two (calibration parameters 36; see Table 1), as the ranges of these vary by several orders of magnitude. Such transformation generally improves parameter search efficiency. Hence, these calibration parameters are sampled in the transformed space and then back transformed before running the cloud parcel model.

2.2.2 Synthetic calibration data

Within an inverse modeling framework the choice of calibration data is an important factor, as this information is directly translated into calibration parameter values. Therefore, some discussion is required regarding the most appropriate form of output from an adiabatic cloud parcel model to be used for parameter estimation. To examine aerosol-cloud interactions, in particular the influence of aerosols on the cloud microphysical properties, the change in cloud model output properties that are of most interest are the liquid water content (LWC), cloud droplet effective radius (R_e), CDNC, cloud fraction, and cloud phase. In this first paper, we are focusing on liquid-phase clouds using an adiabatic cloud parcel model; hence we are limited to investigating microphysical properties from the model output in the form of droplet number distributions or bulk parameters describing this distribution. Cloud bulk properties such as LWC and R_e are not used as this could potentially result in an ill-posed problem, in which many input parameter combinations could possibly result in the same “bulk” value as will be discussed further in Sect. 3.2 in relation to the calibration dataset that was chosen for this study. Therefore, we constrain the setup, such that the droplet size distribution is used as the model output for this investigation.

When defining a suitable OF from a cloud parcel model a primary difficulty arises. Adiabatic cloud parcel models typically employ a moving centre approach for the numerical representation of the particle size distribution (Jacobson, 1997; Korhonen et al., 2005). Yet, from an inverse modeling perspective a moving centre approach has distinct disadvantages. As both the x (radius) and y (number of droplets) are simultaneously changing in each run, it is not trivial to compare measured with modelled size distributions. Ideally, the model outputs y-values at values of x where we have measurements. In that case, the difference between measured and simulated y-values provides an accurate diagnostic for the distance between the model-predicted droplet size distribution and the respective observational data. Unfortunately, this is impossible with a moving centre approach. This begs the question of why we are using a moving centre approach during inverse modeling.

Adiabatic cloud parcel models generally use the moving centre approach to calculate the evolution of a droplet size distribution as this framework eliminates the numerical diffusion upon condensation and evaporation attributed to fixed sectional methods, by letting the particles grow or shrink to their exact sizes, making it well suited for cloud models. Solvers from air quality models were evaluated by Zhang et al. (1999), with the conclusion that a moving centre based structure most accurately reproduces the qualitative features of the size distribution. The fixed sectional approach is clearly inferior for cloud models; therefore, in order for the comparisons between different simulations to be meaningful in a moving centre framework, it is essential to construct a calibration data set that is constant with respect to the size grid for which the droplet numbers are used, regardless of the prescribed calibration parameters (as in a moving centre framework a simulation with a different updraft velocity for instance, will give the raw output on a new size grid). The simplest way to achieve this would be to perform a post-processing procedure in the form of a standard re-binning of the size distribution onto a fixed grid. This can potentially introduce problems, as the shape of the output size distribution will change significantly due to a loss of resolution (it is not possible to re-bin to the same or close resolution as the raw output resolution, as the size distribution would then become “patchy” owing to bins that contain no particles). The loss of resolution will lead to a very “spiky” CDNC, with the majority of the information content stored within the peak over a limited number of bins. This has the potential to create

ACPD
11, 4749–4806, 2011

an ill-posed or "non-unique" inverse problem, the extent to which will be dependent on the original shape of the CDNC.

The common approach when measuring size distributions is to describe the size distribution function in the form $dN/d\log d_p$ (the number is normalised by the bin width to account for the fact that this can vary in a moving centre framework). Therefore, to maintain a structurally consistent calibration data set with respect to the bin sizes, and also information content through resolution, the natural solution is to interpolate this size distribution function for each different parameter combination onto the size grid of the calibration data set. Unfortunately, a direct interpolation of the size distributions onto the grid of the calibration data is not an optimal solution, as will now be discussed in relation to the calculation of the OF.

2.3 Selection of Objective Function (OF)

After running our adiabatic model with a set of parameters and input variables we are left with a predicted size distribution that we need to compare against its measured counterpart. We could simply take the difference of both vectors. This provides a measure of distance between the model and data. In practice however, it is particularly difficult to work with an n-element vector of error residuals, $E(\theta)$, and find the best parameter values. It has therefore become common practice to aggregate the vector of error residuals, $E(\theta) = \mathrm{e}_1(\theta), \ldots, e_n(\theta)$ into a single measure of distance between the model predictions and observations. This measure is also referred to as the OF. The development of such measure that mathematically measures the "size" of $E(\theta)$ is typically based on assumptions regarding the distributions of the measurement errors presented in the data. By far the most popular OF is the simple least squares (SLS) or maximum likelihood estimator, appropriate when the measurement errors are believed to be homoscedastic and uncorrelated. We follow this assumption and use the following definition of the OF:

ACPD
11, 4749–4806, 2011

$$OF = \sum_{i=1}^{n} w_i [y_i - \Phi(X_i, \theta)]^2 = \sum_{i=1}^{n} w_i e_i(\theta)^2 \quad (2)$$

This function is similar to the SSE function used in many other fields of study to minimise the quadratic simulation error during model calibration. Note that the minimum value of the OF for a numerically generated data set with no error is zero. The θ vector consists of calibration parameters, for instance those describing the input lognormal aerosol size distribution over different modes, updraft, and mass fractions of chemical compounds, whereas the w_i denote weights associated with a particular measurement point. In the absence of compelling prior information about the measurement error it is common practice to weight the observations by their respective measurement variance (Šimùnek et al., 1998). If the errors are believed to be heteroscedastic the weights should be adjusted to reflect this deviation from a fixed measurement error variance. Yet, in this paper we are using synthetically generated observations, and the weights of the individual data points are assumed similar and equal to one.

It is important to realize that projection of an n-dimensional observation space onto a single dimension, OF, results in a colossal loss of information from the data. This frustrates the search for the right parameter values. An appropriate OF attempts to consolidate the information contained in the individual measurements. For this cloud model which contains 9 varying input parameters in this study, the inverse problem can be said to have relatively high dimensionality, as the aim is to inversely obtain values for each (hence 9 dimensions) from information stored in one dimension.

It should be evident from our discussion that the definition of the OF exerts a strong influence on the identifiability of the different cloud-parcel model parameters. Yet, the information content of the calibration data stored within the OF is also important; in relation to calibration data from particle size distribution this information content can be viewed as resolution and/or particle size range. Data that exhibit significant (spatial and temporal) variation are preferred as such observations are likely to excite different model processes (and thus parameters). It is therefore important to construct an OF

ACPD
11, 4749–4806, 2011

that contains independent (orthogonal) information about each of the individual calibration parameters. Such an OF is sensitive to each individual parameter and therefore should contain a well-defined global optimum within the parameter space.

For a cloud experiment, a natural choice for the calibration data $Y = \{y_1, ..., y_n\}$ would be the droplet size distribution in the form $dN/d\log d_p$. This selection of measurement data poses a main problem, that is that the corresponding model predictions $\tilde{Y} = \{\tilde{y}_1,...,\tilde{y}_n\}$ of the adiabatic model, $\tilde{Y} = \{\tilde{y}_1,...,\tilde{y}_n\}$ would need to be interpolated onto the size grid of $Y = \{y_1,...,y_n\}$. Otherwise the OF would be rather meaningless, because we are then comparing simulated and measured values of $dN/d\log d_p$ at different x-values (the radius).

Computational results suggest that a simple interpolation of the size distribution is not an ideal approach for cloud parcel models. In essence, parameter identifiability is deteriorated with this approach as widely varying parameter values resulted in very similar values of the OF. This really frustrates the parameter search, and makes it very complicated to find the appropriate parameter values. Our experience suggests that a more productive approach is to explicitly include both the number and associated radius components of the size distribution in the formulation of the OF. We compute both the x- and y-components of the droplet size distribution and include both these vector of values in one OF. This was found to work better than interpolating the y-values of the simulated curve to the respective x-values of the measured curve, as discussed previously. A question that remains is whether we need to weight the individual data types because they each have different units. Initial numerical tests have demonstrated that such weighting is not necessary for our particular case study. Thus, our vector of model predictions, $Y = \{y1,...,yn;x1,...,xn\}$ consists of a joint vector of $\log(dN/d\log d_p)$ and $\log(\mathrm{d}_p)$ values. This data set can be computed at either one, or at multiple altitudes within the cloud.

To illustrate our approach and data set, please consider Fig. 3 that presents an artificially created particle size distribution for marine and rural continental conditions. The associated X and Y components of this distribution are depicted in log space in Fig. 3b

ACPD
11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

and Fig. 3c respectively. Note that we present the entire particle size distribution, but in practice we are only interested in droplets greater than 1 µm in size. The bin number corresponding to this radius size is illustrated by the dotted lines for marine (blue) and rural continental (pink) within the figures. Therefore, for the first response surfaces presented in this study in which we focus on the activated particles (droplets), the OF is only calculated from the bins to the right of the dotted lines.

2.4 Response surfaces as a graphical tool to analyse the inverse problem

Results of a response surface analysis match closely with global sensitivity analysis, with the purpose of providing a qualitative graphical illustration of the outcome of global sensitivity analysis. Thus response surfaces convey important information about the posedness of the inverse problem. For instance, surfaces that are smooth (deterministic) and contain a single well-defined minimum are preferred, as the gradient (slope of the response surface) points to the same minimum anywhere in the search space, irrespective of the location in the feasible parameter domain. Such surfaces are easy to traverse and find the optimum parameter values. Therefore, a well defined global solution is desirable, as it will help identify the values of the model parameters which are essential for predictive purposes. Calibration parameters that are very sensitive and relatively easy to identify can be efficiently solved using any type of search algorithm (Luo et al., 2009). As a visualization aid, a response surface for a parameter combination which has a well defined optimal solution is illustrated in Fig. 4. The response surface has a clear smooth convex gradient towards the optimal solution, which consists of a relatively small basin of attraction

On the contrary, response surfaces that are flat (OF is insensitive to changes in parameter values) are rather difficult to traverse as the direction of improvement is hard to find, leading to considerable uncertainty in the optimal values of the parameters. Response surfaces that exhibit an erratic (chaotic) pattern, and/or contain multiple basins of attraction, are also particularly difficult to solve, as most search algorithms will be prone to local convergence and tend to get stuck in pursuit of finding the global

ACPD
11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

optimum parameter values. Therefore, optimization problems with many local solutions and flat response surfaces are more difficult to solve, and typically require global search methods that launch multiple concurrent searches from different starting points in the parameter space to reduce the risk of becoming trapped in a local minima en pursuit of the global optimal solution.

If the various response surfaces already exhibit considerable peculiarity in two-dimensions, it is to be expected that the behavior of the OF in the full-dimensional parameter space will only be more erratic. In summary, if the response surfaces do not display a well-defined global minimum in the two-dimensional parameter planes, the conventional inverse parameter estimation technique may certainly be expected to be unsuccessful in a multidimensional plane. It is stressed that the behaviour of the OF in these parameter planes can only suggest how the OF might behave in our 9-dimensional continuum. For example, local minima of the OF could exist and not show up in the cross-sectional planes (Šimúnek et al., 1998; Vrugt et al., 2001). Nevertheless, the response surfaces provide a useful approximate view of the behaviour of the OF in the entire parameter space, allowing the identification of calibration parameters that may need fixing prior to optimisation for the successful coupling to an automatic parameter search algorithm.

In summary such parameters are important to locate, and it is equally important to understand the reason behind the non-ideal shape of their respective response surfaces to be able to effectively apply an optimisation algorithm to investigate cloud-aerosol interactions. Response surface analysis also allows a qualitative evaluation of whether data measured conventionally, for instance during an airborne measurement campaign of cloud and aerosol properties, provides enough information to enable the identification of the input aerosol chemical and meteorological properties from the inverse problem with an adiabatic cloud parcel model.

The focus of the following sections will involve the analysis of a set of response surfaces and this will be approached in a way similar to that done previously by Toorman et al. (1992), Šimúnek et al. (1998), and Vrugt et al. (2001). The response surfaces

ACPD
11, 4749–4806, 2011

are calculated by solving the OF for many possible combinations of a selected pair of parameters from Table 1 on a rectangular grid. A response surface is then obtained by changing these two selected parameters around their true values, whilst keeping other parameters constant at their true values.

Therefore, an inverse problem containing 9 input parameters will result in a total of 36 paired parameter combinations, and thus 36 possible response surfaces. Each parameter domain was separated into 30×30 discrete points, resulting in 900 grid points for each response surface simulation.

3 Results and discussion

3.1 Response surface analysis

From the 36 calibration parameter combinations a base selection of six response surfaces have been chosen to represent a wide range of possible shapes for the optimal solution observed in the two-dimensional plane, and illustrate the nature of the cloud-aerosol inverse problem. These were calculated for both marine (Fig. 5) and rural continental conditions (Fig. 6). The remaining response surface parameter combinations were also investigated (figures not shown).

Figures 5d and 6d illustrate response surfaces for the input aerosol number concentration in the largest aerosol mode (mode two) versus the mean radius in mode two for marine and rural continental aerosol input conditions respectively. For these parameter combinations there is a clear, well-defined optimal solution to the inverse problem as indicated by the OF values smoothly decreasing in a generally convex manner towards the location of the optimal solution denoted by the blue cross. The smooth gradient of the OF towards the optimal parameter values could be easily and efficiently located using any number of optimisation algorithms currently available.

The same can be said for both environments, where parameter one is now the mean radius in mode two and parameter two on the x-axis is the updraft velocity (Fig. 5f, 6f). There is a clear optimal solution to the problem; the only difference being that the minimum is located in the outer range of the parameter space. From a detailed analysis of response surfaces for other parameter combinations (figures not shown), it was apparent that the parameter combinations that lead to the most identifiable response surfaces were the updraft velocity, mean radius of mode two, and number concentration (Na) of mode two. This is understandable as the physical structure of adiabatic cloud parcel model running with constant updraft velocity is such that in the absence of entrainment/mixing processes a decrease in the required critical supersaturation (S^*) by increasing the updraft velocity (v) will monotonically increase the total number of droplets (Nd) until all of the particles become activated; $Nd = f\left(S^*\right) = f\left(v, dN_a/d\log d_p\right)$.

Parameters also exist that illustrate non-identifiable characteristics. A model that contains non-identifiable calibration parameters can lead to non-uniqueness. Non-uniqueness indicates the presence of more than one set of calibration parameters, each yielding minimum values for the OF (Sorooshian and Gupta, 1985; Duan et al., 1992). Thereby the information content available to define an inverse modeling problem does not allow a single or unambiguous mathematical solution to the identification problem (Vrugt et al., 2001, 2005; Beven, 2006). Equifinality is also used to describe the attribute of either different models or dissimilar parameter values of the same model corresponding with data equally well, without the ability to distinguish which models or parameter values are better than others. For a more detailed discussion of these terms the reader is directed to Luo et al. (2009) and Beven (2006). To solve the problem of non-uniqueness it is first paramount to understand which parameters are non-identifiable, and aside from response surface analysis there are other methods and analytical tools to help the user ascertain these (Pollacco and Angulo-Jaramilo, 2009 and Cressie et al., 2009). Once these parameters are identified, if they are not deemed significant for the study it is possible to fix them during inverse modeling so that they do not impede the optimisation algorithm locating the remaining parameters that are easily identifiable. If, however they are important parameters the next step is to either change the model structure or formulation of the OF.

ACPD
11, 4749–4806, 2011

Calibration parameter combinations that are non-identifiable are illustrated for both marine (Fig. 5a) and rural continental (Fig. 6a) aerosol environments for the aerosol concentration in mode 1 versus the updraft, and also in Figs. 5b, 6b for when the updraft is replaced by the soluble mass fraction. Changes in the aerosol concentration over the limits defined in Table 1 result in insignificant changes in the OF as seen in the Y plane of the response surface. Such a non-identifiable calibration parameter cannot be estimated using current available observations unless it is clearly constrained by some other calibration parameter combinations. The measurements simply do not contain the appropriate and required information to warrant its estimation. This informs us that relatively this is not an important calibration parameter within the model structure under investigation compared to others to cause significant changes in the OF. Non-identifiable input parameters can result in a distribution of parameter values existing (the posterior parameter distribution results from an optimisation simulation) that exhibit virtually identical OF values. This would likely translate into significant model predictive uncertainty. In other words, if many or all of the input calibration parameters are non-identifiable in nature it will be impossible for an algorithm to converge to a single true optimal solution, making it difficult to ascertain the contributions of this non-identifiability to parameter uncertainty. Undesirable side effects to this include that the uncertainty of the calibration parameter(s) might increase considerably outside the calibration period when they might become active and sensitive. A single optimised vector of calibration parameter values does not communicate this uncertainty, illustrating the need for sensitivity analysis in the form of response surfaces.

Response surfaces for certain calibration parameter combinations are not necessarily completely non-identifiable; some were also found to exhibit partially non-identifiable characteristics. Such a parameter combination can be classified as poorly constrained, and an example of this can be seen in Figs. 5e, 6e for the aerosol concentration in mode 2 versus the soluble mass fraction. For this calibration parameter combination the OF is relatively insensitive to changes in the soluble mass fraction up to a certain point in the parameter space; it is not as well confined compared to when the soluble

ACPD
11, 4749–4806, 2011

mass fraction is replaced by the updraft velocity (figures not shown). Therefore, the threshold change in soluble mass fraction that results in a significant change in the OF is larger; i.e. the droplet size distribution within an adiabatic parcel model is less sensitive to the composition of the aerosol than, for instance, the updraft velocity, as would be expected since the updraft controls the critical supersaturation required for activation into droplets.

Generally the response surfaces using a marine aerosol size distribution (Fig. 5a–f) have the same general shape as when the synthetic measurements were generated using a rural continental distribution (Fig. 6a–f). However, for marine aerosol conditions the OF surface is neither as smooth nor continuous; the gradient varying in a more unpredictable manner through the parameter space. This hampers the progress of optimization algorithms (Duan et al., 1993). The meteorological conditions remain the same between both simulations, so it is obviously the nature of the aerosol size distribution that creates a slightly more chaotic route to the optimal solution. This difference between the shape and smoothness was found to be more noticeable the larger the difference between the number of particles in the aerosol environments. With respect to the smoothness of the response surfaces it is possible that this is due to the location of information content within the OF, since for the marine case the interstitial and cloud droplet distribution is narrower than for the rural continental environment (i.e. the information is stored over a fewer number of bins). This narrower calibration data set does affect the smoothness of the response surfaces since the droplet size distribution has a steeper gradient. Small changes in the input parameters can lead to bigger "jumps" in the error residual vector from one bin to the next, hence the OF appears more discontinuous ("jagged") when represented by response surfaces. Even though the response surfaces are less smooth for marine conditions, a search or optimization algorithm should not exhibit too much difficulty locating the optimal solution for these two parameter combinations as the OF is well defined.

Since a significant number of parameters were found to be non-identifiable, one or more may need fixing at their true values prior to coupling an adiabatic cloud parcel

ACPD
11, 4749–4806, 2011

model with a MCMC algorithm. The parameters of most interest are those for which there is greatest uncertainty in the literature, the debate concerning the relative importance of size compared to chemistry for the cloud nucleating ability of particles being of specific interest. The question of fixing non-identifiable parameters also has to be balanced with respect to the dimension of the problem. With 9 calibration parameters and a limited calibration data set, it is particularly difficult to identify and constrain each individual parameter. Indeed, at this stage it is envisaged that due to their non-identifiable characteristics in multiple response surfaces, one or more of the calibration parameters describing the smaller aerosol mode (mode one) will need fixing for a successful optimisation of the remaining calibration parameters.

3.2 Optimisation difficulties

Non-identifiable calibration parameter combinations aside, there are also other conditions which can impede inverse modeling, and these should also be identified. In this section we will demonstrate why other qualitative features of the response surfaces indicate that inverse modeling of adiabatic cloud parcel model calibration parameters is expected to be difficult.

A search algorithm may encounter difficulties if several major regions of attraction exist into which it may converge (i.e. it is non-convex), and if the parameters exhibit varying degrees of sensitivity and a great deal of highly non-linear interaction. It is also possible for a region of attraction within the parameter space to contain numerous (possible uncountable) local “minima” or “optima” (location of low OF values, herein denoted as local minima). These may occur both close to and at various distances from the best solution (Duan et al., 1993). This characteristic can make it difficult for a search algorithm to find which of the local minima is the “true” minima, and hence the true solution to the inverse problem; single chain MCMC algorithms are especially vulnerable. This feature can be seen in the response surfaces for the updraft and surface tension for marine and rural continental conditions. This is seen for both aerosol environments and is demonstrated for marine conditions in Fig. 7. The OF for this response

ACPD
11, 4749–4806, 2011

surface is represented now in the Z co-ordinate, so the difficulty of a search algorithm becoming trapped within a local minima can be more easily visualised. For the updraft and surface tension many local minima, both near to and far from the true solution, can be identified, resulting in the response surface having a multi-modal nature. The resulting shape and multi-modal nature of the location of possible "close to" optimal solutions over the parameter space would be challenging for most optimisation algorithms. Such a parameter combination ideally needs to be probed using MCMC algorithms that simultaneously launch multiple chains that share information, to avoid becoming trapped within local minima seen by the basins in the response surface. Nevertheless, the general shape of these response surfaces are logical, based on our knowledge of the interactions of the parameters in question in relation to the OF. The updraft is expected to compensate changes in the surface tension so that the higher the updraft a higher surface tension is required to activate the same number of particles. However, the question remains why there is not just a single broad minimum, i.e. a solution that is poorly constrained. It can be hypothesised that this can be attributed to the bimodal nature of the entire model output before a limit is imposed with respect to bin number in relation to particle sizes smaller than 1 µm in the calibration data. A calibration data set containing both interstitial aerosols and droplets on a moving grid means that potentially the change in gradient between simulations is not smooth. This is likely to be attributed to the shifting location of the region of the size distribution spectrum at which particles become activated into droplets. For instance, if an input parameter has a sensitivity per bin of the smallest activated droplet that is far greater or far smaller than the sensitivity of the smallest unactivated droplet, it is possible that there is a tipping point in the subsequently calculated value for the OF associated with the model resolution within this size range. This translates into chaotic changes in the error residual over adjoining bins and thus the observed response surface contains a series of minima regions associated with different calibration parameter values. A more physical way of interpreting this in relation to the model structure of a cloud parcel model is that there is a break between the interstitial aerosol and activated droplets (Fig. 3a). This will result

ACPD
11, 4749–4806, 2011

in the highest variability in the size distribution with respect to the bin sizes occurring at the "split" between the unactivated interstitial particles and activated droplets. This "split" is determined essentially by the updraft via the critical supersaturation, therefore, it is not surprising that response surfaces demonstrating multiple modes contain the updraft as one of the parameter pairs.

From the response surface analysis we have established that certain calibration parameter combinations with the current OF definition have characteristics that could pose problems when we couple the adiabatic parcel model to an optimisation algorithm. It has been shown that the "smoothness" of the response surfaces for parameters expected to be key for the activation of particles is dependent on the initial input aerosol size distribution for which we define the calibration data. Subsequently a response surface that can appear smooth and well defined using polluted aerosol size distribution as the "truth" can appear slightly discontinuous and vary in a more unpredictable manner across the parameter space for cleaner conditions.

It is worth noting that these results only represent response surfaces for the current choice of calibration data used in the calculation of the OF. A different setup in which the OF is calculated at a different height in the cloud (for instance below the peak supersaturation), or includes the interstitial aerosol, could potentially result in an altogether different shape of the response surface. This has implications depending on whether the new response surfaces are as identifiable or not. It is advantageous to have the most identifiable response surfaces with the fewest model output variables, as it is simpler to perform measurements of cloud droplet number concentrations over smaller size range and (up to a point) lower resolution. On the other hand, including more information terms of the range of the size distribution does not necessarily mean a more identifiable response surface. For instance, if the location of highest information content (i.e. highest sensitivity) within the CDNC for a particular model input parameter becomes masked by a less sensitive region of the output in the calculation of the OF. In other words, the information content of data is not determined by the length of the data set, but by variability of the data set.

ACPD
11, 4749–4806, 2011

Suffice to say, it has been shown that a simple adiabatic cloud parcel model satisfies all of the conditions necessary to make the inverse problem extremely difficult to solve with an optimisation algorithm. In order to understand these inherent differences between marine and rural continental conditions, sensitivity tests are needed with respect to the location of information content within the OF which will be the focus of the next section.

3.3 Objective function: importance of location of information content-clean clouds vs. polluted clouds

Experience gained from other fields where inverse modeling is frequently utilised has shown that it is particularly useful to know how individual measurements constrain parameters, and to figure out the information content of the various observations. The spread of this information is of particular importance for inverse modeling as it is lumped into a single OF value. It would be expected that different regions of the interstitial aerosol and droplet size distribution will provide different degrees of information for the calibration parameters; that the location of the highest information content will vary between parameters and also be dependent on the initial aerosol regime. The simplest way to identify the variation of the location of the information content of calibration data is by a performing a very simple sensitivity analysis on the $dN/d\log d_p$ size distribution function over the entire particle size range available. As an optimisation algorithm has not previously been applied to a cloud parcel model there is not currently a consensus in the literature as to what model output provides the optimal information for applying inverse modeling to study cloud-aerosol interactions.

This serves two purposes; firstly it allows us to better understand the results from our response surface analysis which portray the variation of the OF over the parameter space, in turn allowing for a possible improvement in the OF. Secondly it shows the potential of inverse modeling for cloud parcel models when linking the model to actual measurements instead of using synthetic data. We can then better understand over what size ranges it is most important to measure the cloud particle/droplet distribution.

ACPD
11, 4749–4806, 2011

There are numerous ways to perform such a sensitivity analysis to compare the sensitivities of the calibration data to different calibration parameters. The same approach as Šimúnek et al. (1998) is chosen, i.e. an increase of the calibration parameter values by a fixed percentage.

The sensitivity is calculated for each bin as:

$$\frac{S_d N}{d\log d_{p,i}} = \frac{dN}{d\log d_{p,i}}(P+\Delta P) - d\log d_{p,i}(P) \tag{3}$$

$$\Delta P = 0.05P \tag{4}$$

Where $S_{dN}/d\log d_p$ in Eq. (3) is the change in the size distribution corresponding to a 5% change in parameter P by ΔP. We used a one-sided approach, in which parameter sensitivities are determined by perturbing the reference set with 5%. A centred sensitivity analysis approach, in which the outcome of a positive and negative parameter perturbation are averaged, yielded very similar results.

As a reference point the sensitivity is first provided when the size distribution is plotted against the actual radius for each sensitivity simulation for marine conditions (Fig. 8). For the subsequent analysis the information content will be interpreted with respect to bin number. In the definition of the OF the Y component of the residuals are calculated between data points corresponding to bin numbers for the $dN/d\log d_p$ size distribution function. Therefore in the following plots, to allow a consistent analysis of the information content at each bin point (each having equal weighting), the sensitivity figures are plotted against bin number instead of radius. This also allows for a clearer interpretation of the location of the information content with respect to the OF, as an algorithm would view the size distribution in this format. Both components of the OF were analysed, and in this section results from the Y component will be presented, as these alone provide adequate information for the purpose of this section.

From the nine possible calibration parameters, a sensitivity analysis will be provided for a selection of five calibration parameters to address the issues raised in the previous section. These are the aerosol concentration in mode one and two, the aerosol radius

ACPD
11, 4749–4806, 2011

in mode two, the updraft and the soluble mass fraction. The updraft and soluble mass fraction are plotted in a separate figure for clearer viewing.

The results for marine conditions are presented in Fig. 9 and the rural continental environment in Fig. 10. Firstly it can be seen that the interstitial region of the size distribution is more sensitive to a perturbation of the aerosol concentration in mode one. This is expected as mode one is located in the lower size end, thus here its sensitivity towards the selection of modal number concentration is largest, as these smallest particles are less easily activated into cloud droplets. For the rural continental environment the sensitivity spans a greater size range due to the greater number and wider base particle size distribution (Fig. 10a, b) compared to the marine case (Fig. 9a, b). From this it can be deduced that the identifiability of this calibration parameter could be vastly improved by including the interstitial aerosol in the OF.

A perturbation of the mean radius of particles in the larger mode 2 results in a shift in the size distribution towards larger sizes. This can be expected for a cloud parcel model in which cloud droplet activation is described using the Köhler equation. An increase in the mean radius of mode 2 results in the model being initiated with a greater number of larger sized aerosol particles. The associated Kelvin effect results in a lower critical supersaturation required for activation into droplets, meaning that after the peak supersaturation is reached these particles grow faster at the expense of smaller unactivated particles. No sensitivity to the interstitial aerosol is observed for this calibration parameter, whereas for the aerosol concentration in mode 2 there is a very small sensitivity to be seen for the largest interstitial particles. Therefore, for these two parameters not much information will be added to the OF by including the interstitial aerosol in its calculation.

When the updraft is perturbed by 5% the impact for the marine environment (Fig. 9.c, d) is most visible in the droplet size range greater than 1 µm. This is undesirable since the updraft controls the supersaturation profile, and hence the moisture available to condense on a particle at a certain height. Therefore an increase in the updraft velocity leads to an increase in number of particles able to activate. The change in soluble mass

ACPD
11, 4749–4806, 2011

fraction impacts particle activation via the Raoult term and due to particles growing we see an increase in the number of smaller activated droplets and reduction in the number of larger droplets. The same shape of the sensitivity to updraft and soluble mass fraction is observed for the rural continental case (Fig. 10c, d), however, as the base case contains a greater number of larger particles increasing the updraft and soluble mass fraction has less impact in this size range, resulting in a broader sensitivity distribution for more polluted environments.

Sensitivity tests were also performed to examine the effect of the initial selection of updraft velocity as it is known to influence droplet activation efficiency strongly. Therefore, additional simulations for both rural continental and marine conditions were completed where the "true" updraft was increased from $0.5\,\mathrm{m\,s^{-1}}$ to $2\,\mathrm{m\,s^{-1}}$ (figures not shown). For the high updraft case, when the updraft or soluble mass fraction is perturbed the effect is to reduce the overall sensitivity for each bin, and more noticeably alter the difference in magnitude between the sensitivity for the smallest droplets compared to the larger droplets. A 5% increase in updraft or soluble mass fraction makes a much smaller difference on whether or not a particle can grow to the bin size required for activation into a droplet, since almost all possible particles reach the critical radius, i.e. the supersaturation profile for the high updraft case allows the growth and activation of a greater percentage of the original aerosol. The effect of increasing the base updraft value on the sensitivity was much stronger for the cleaner marine environment.

The chaotic nature of the sensitivity to updraft, seen in particular over a few number of data points associated with the transition between interstitial aerosol and activated droplets helps explain the multiple minima seen in the response surfaces containing the updraft. This can be seen in Fig. 9d by the large negative and positive changes for a perturbation in updraft velocity at single data points at the dotted line illustrating the 1 µm bin number for the calibration data set in Fig. 3b. This effect is more prevalent for the marine environment due to the narrower droplet mode. In essence the strength of this artefact as seen in the response surfaces will be dependent on the initial aerosol distribution used in relation to the sensitivity of a particle reaching its critical radius with

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

ACPD

11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

respect to a perturbation of a calibration parameter; particularly the updraft velocity, and to a lesser extent the soluble mass fraction and surface tension.

This break in size range between interstitial particles and droplets is a feature of the moving centre output and creates undesirable jumps from one size bin to another in the sensitivity at this activation region of the particle spectrum. The smoothness of the transition from interstitial aerosol to activated droplets could be improved by developing the approach to determine the fraction of activated particles (Korhonen et al., 2005; Takeda and Kuba, 1982) or by employing an adaptive spectrum refinement procedure such as the one developed by Arabas and Pawlowska (2011). Another proposal to improve the definition of the OF would be to fill the gap between the interstitial aerosols and droplets with some smooth filling procedure for the sizes containing zero particles (i.e. synthetically fill these bins). This should help smooth out the gradient, and also the response surfaces. One further potential improvement to smooth out the numerous local minima experienced in the response surfaces would be to weight the model output in an appropriate manner. This was shown to have a big impact on wavelength experiments by Cochran and Horne (1977). There are numerous possible weighting methods, and a weighting technique has been used successfully in the determination of narrow aerosol size distribution measurements (Voutilainen et al., 2000). Unfortunately this is no simple task when using synthetic calibration data for a simple adiabatic parcel model, as there is no variation in time with which to apply the weighting, and if an arbitrary weighting is applied to obtain a smoothing it becomes very easy to bias the results (also such a weighting may not be successful for all aerosol environments).

The reason the response surfaces appear less well defined in cleaner conditions seems to lie in the fact that a lot of important information is contained over a smaller number of bin numbers as hypothesised in Sect. 3.2. This potentially has a strong impact for this synthetically generated calibration data in which we give equal weighting to every bin, i.e. the number and radius of particles are not weighted by a measurement error as they would be in the case of using real measurements.

ACPD
11, 4749–4806, 2011

Title Page
Abstract Introduction
Conclusions References
Tables Figures
◀◀ ▶▶
◀ ▶
Back Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

More polluted clouds will contain more particles in the interstitial size region compared to clouds generated from cleaner aerosol environments. Also clouds generated from cleaner conditions will have more variable but narrower sensitivity distribution shape across the size bins. Experiments that yield a droplet size distribution, for which the geometric standard deviation is large, either due to a high number of aerosol particles or high updraft/soluble conditions, are beneficial for parameter estimation studies, since the measurements then contain independent information about most of the parameters. This increases the identifiability of the parameters and enhances the likelihood of uniqueness of the final parameter estimates. In summary the feasibility of optimisation will be highly dependent on the shape of the input aerosol size distribution. In essence a trade-off exists between the shape of the aerosol size distribution and the distribution of information over the bins; since in turn these characteristics are related to the smoothness and identifiability of a response surface.

3.4 Improving the information content of the objective function

It is clear from previous analysis that the successful solution to the inverse problem for cloud-aerosol interactions will be sensitive to the size limits imposed on the model output in the calculation of the OF. In the past the consensus was that the best solution to the problem of non-uniqueness was to include additional and more accurate measurements. Recent research into data requirements reveals that actually the information content of the data used for calibration is far more important than the quantity (resolution) of the data (Kuczera, 1982; Sorooshian et al., 1983; Sorooshian and Gupta, 1985; Yapo et al., 1996; Gupta et al., 1998; Vrugt et al., 2001). In this respect we can conclude from Sect. 3.2 that the identifiability of the aerosol concentration in mode one could be improved by including the interstitial aerosol in the calculation of the OF. In light of this a subset of three response surfaces were recalculated for both marine and rural continental environments, this time including the entire size range in the calculation of the change in X and Y-values of the size distribution function (i.e. using the entire functions illustrated in Fig. 3b, c). It is clear that for both environments, including particles less than 1 µm in size, the solution for response surfaces containing updraft and soluble mass fraction versus the aerosol concentration in mode one become more identifiable (Fig. 11a–d).

Conversely, including the interstitial aerosol contributes to the presence of local minima for certain calibration parameters most important for droplet activation as seen in Fig. 11e–f for the aerosol concentration in mode two versus the radius of mode two. This is believed to be caused by an amplification of the reasons determined for the presence of local minima as discussed in Sect. 3.2. By including the interstitial aerosol the size break between the interstitial particles and activated droplets is present in its entirety, allowing more instances for large changes in error residual values over a finite number of bins for a small change in updraft velocity. A similar problem has been experienced in the inversion of aerosol measurements in the 1980's for which algorithms were found to work well for unimodal distributions; however, for bimodal distributions they tended to find local minima (Helsper et al., 1982). Unfortunately within a moving centre framework, if only the activated particles are used in the OF then local optima will continue to exist for certain parameter combinations as previously discussed.

Naturally the choice of including both the interstitial aerosol and activated particles or only the activated particles within the OF will depend on the calibration parameters of interest. As the parameters clearly exhibit varying degrees of sensitivity and a great deal of interaction and compensation a setup of just including the CDNC at the cloud top may give the smoothest, best-defined optimal solution for a one parameter combination, compared to including the interstitial aerosol in the calibration data. However, if the main aim of the study was to inversely ascertain the aerosol calibration parameters describing the smallest aerosol mode, which are less important for droplet activation, the best chance for an optimisation algorithm would be achieved by including the interstitial aerosol.

Another possible method of adding more information would be to add extra cloud height levels to the calibration data. This was performed for marine conditions by

ACPD
11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

additionally including the output at 20 m and 100 m above cloud base (figures not shown), and the general change across response surfaces for each parameter combination was that the optimal solution became better defined, but this improvement was not particularly spectacular for any of the considered parameter combinations. This is not really surprising since for this model set up of our pseudo-adiabatic cloud parcel model the processes of entrainment and collision-coalescence were deactivated, therefore, the number of activated droplets change as a step function with respect to height. Also processes such as mixing are not explicitly accounted for in the model and these are necessary to accurately reproduce the observed CDNC values and/or dynamics at the cloud top when comparing against real world measurements. Therefore, it must be noted that our model remains a considerable simplification of real clouds. The use of data from multiple heights will potentially improve our model parameterization, yet will pose difficulties during calibration. The increased dimensionality of the parameter space will likely increase parameter uncertainty and correlation. Nonetheless, data from multiple cloud heights will prove instrumental for the evaluation of our cloud parcel model against real world measurements, and will help identify structural inadequacies that are not evident from a single level.

Therefore, on balance it is clear that it is the inclusion of the right type of information, not necessarily more information that is most important. It has also been ascertained that when the interstitial aerosol and droplets are included in the calibration data set for the moving centre based model used in this study problems could arise as indicated by the response surfaces being not perfectly smooth and exhibiting numerous minima solutions for certain parameter combinations.

An alternative approach would be to include more moments of the size distribution function in the calculation of the OF. It is common to use moment based methods to estimate parameters for the droplet size distribution, although the location of information content can become biased depending on the choice of moment, which can in turn hamper their effectiveness (Smith et al., 2009). It is not within the scope of this study to re-calculate all response surfaces using moments of the CDNC in detail, however

ACPD
11, 4749–4806, 2011

initial results indicate (figures not shown) that this has the potential to add desirable information for the largest particle sizes, but does not circumvent the problem associated with the particle not being on a fixed size grid.

Another possible solution is to use multiple-objective functions by including bulk variables in the OF to constrain the solution to the inverse problem. This could potentially be advantageous for the benchmarking of parameterisations designed to calculate bulk variables of total droplet number for instance against real world measurements.

The OF used in this study provides smooth and generally well defined response surfaces for the calibration parameters known to be most important (aerosol concentration/mean radius of mode two and updraft). This will facilitate the efficient solution to the inverse problem when a pseudo-adiabatic cloud parcel model that employs a moving centre calculation of the droplet size distribution is coupled to an MCMC algorithm. This does not circumvent the possible bias that may be introduced to parameter sensitivity results from an MCMC analysis that uses an OF that contains both the X and Y-components of a size distribution. Therefore, this section can be concluded by stating that we advocate the development of fixed sectional cloud parcel models which avoid the drawbacks of numerical diffusion (Lehtinen and Kulmala, 2003).

4 Monte Carlo (MC) parameter sampling and automatic model calibration

In this section the results of a standard Latin hypercube sampling MC simulation will be compared with a deterministic optimisation algorithm: the Shuffled Complex Evolution global optimisation algorithm (SCE-UA) (Duan et al., 1992).

A simple MC based approach can only give an estimate of the global optimum solution, whereas a search algorithm such as SCE-UA is especially designed to find the minimum of the response surface (OF) in a minimum number of function evaluations. The success of these two different search methods can be visually demonstrated by marking onto the response surfaces which solutions have been visited by each individual method.

ACPD
11, 4749–4806, 2011

Figure 12 plots the various solutions that have been created with the MC algorithm onto a selected set of response surfaces for rural continental and marine conditions. It has been previously demonstrated that the cloud-aerosol parameter estimation problem is relatively difficult to solve. The determination of the minimum of the OF will become increasingly difficult with increasing dimensionality of the parameter space. The more calibration parameters in the cloud-parcel model the more difficult it will become to estimate them accurately. With the knowledge gained from our response surfaces analysis, we decided to fix the surface tension, known to be highly multi-modal, to its known (true) value. The non-identifiable calibration parameters describing mode one were included in this test to ascertain the impact on the performance of the search algorithm. Altogether, this leaves us with eight "unknown" parameters for optimization against the measured size distribution. The green dots represent the lowest 10% of the OF values from 3 000 different MC realizations. Each dot corresponds to a different parameter combination. A brute force MC approach is not only inefficient, but potentially also misleading. If we draw inferences based on this set of 3 000 different solutions, then our ensemble of best solutions is still sufficiently removed from the actual optimum solution (blue cross). This highlights the need for robust search algorithms such as SCE-UA which can more efficiently find the optimal solution, especially when estimating many calibration parameters at the same time. This is indicated by the best 10% of the SCE-UA simulations as shown by the red dots being located densely over the true optimal solution for each response surface. About 2 000 model evaluations were required to locate their exact optimum values for the calibration parameters illustrated in these surfaces. Fortunately, the adiabatic cloud model used herein is computationally efficient. Imagine that each model simulation for a different parameter combination would take at least 30 minutes. High performance (parallel) computing would then be necessary to solve the parameter estimation problem. Nevertheless, our findings illustrate the need for a sophisticated search algorithm that within a small number of function evaluations can efficiently and effectively explore the parameter space. It appears evident from our results that MC sampling is rather inefficient. The solutions

ACPD
11, 4749–4806, 2011

exhibit significant scatter, and even after 3 000 trials not a single parameter set can be found that spots the optimum perfectly. In a subsequent paper (part 2) we will use a state-of-the-art algorithm called DiffeRential Evolution Adaptive Metropolis (DREAM: Vrugt et al., 2009) to efficiently search the parameter space. This method belongs to the class of MCMC methods (Bayesian statistics) that not only provide an estimate of the best parameter values, but also a sample set of the underlying (posterior) uncertainty. This distribution contains all the desired information about parameter sensitivity, interaction and correlation, and can be used to produce confidence intervals on the model predictions.

5 Conclusions

In this paper we have explored the feasibility of inverse modeling for estimating parameters in cloud adiabatic parcel models. For a successful and efficient calibration of the cloud-aerosol inverse problem, cloud parcel models that employ a moving-centre based calculation of the droplet size distribution require both the X and Y-components of the $dN/d\log d_p$ size distribution function to be explicitly included in the OF.

A response surface analysis was used to gain a visual understanding of the behaviour of the OF in the parameter space. This analysis reveals important information about the identifiability of the pseudo-adiabatic cloud parcel model parameters and the information content of the calibration data. Response surface analysis is very useful, not only to help guide finding the appropriate parameter values when coupling a model to an automatic search algorithm, but also to provide experimentalists with better insight as to where to focus development of instruments to enable cloud modellers to evaluate, or calibrate their models to a higher level of accuracy.

We have shown the usefulness of applying a qualitative response surface analysis to cloud-aerosol interactions before moving onto coupling a cloud model to an optimisation algorithm. Cloud models are increasingly becoming more complex (high

ACPD
11, 4749–4806, 2011

dimensional problem due to large number of calibration parameters), are highly non-linear in nature due to the evolving nature of an aerosol size distribution into a cloud droplet distribution, combined with being highly parameterised. From our analysis it has been shown that these qualities mean that the cloud-aerosol inverse problem will be particularly difficult to solve. The calibration parameters describing the aerosol mode one have been found to be non-identifiable for the cloud model used in this study and these may need fixing to their true values for the successful application of an optimisation algorithm, unless the interstitial aerosol is included in the calculation of the OF. However, including this extra information in the current form of the OF does not overcome the bimodal nature of the particle size distribution which will likely introduce more local optima into other calibration parameter combinations that are actually of primary interest, potentially with detrimental affects to the overall success of the calibration. Therefore it is recommended that the mode one aerosol concentration, mean radius and geometric standard deviation be fixed to their true values, rather than including the interstitial aerosol, especially since measurements of interstitial aerosol inherently are rarer and contain high uncertainty.

The region of the size distribution associated with the activation of particles can also introduce local minima into the response surfaces due to the split between unactivated interstitial aerosol and activated droplets resulting in discontinuous error residuals. This will create difficulties for any optimisation algorithm, and this impact is stronger for cleaner clouds than for polluted clouds. The effect was observed to be strong for the surface tension; therefore, this calibration parameter may also need fixing to its true value before the successful coupling of an automatic search algorithm to a cloud parcel model.

The scope of our future research will be to test MCMC methods for inference of the posterior probability density function of the parameters. This distribution contains all desired information about parameter sensitivity, correlation and interaction, and can be used to determine how many parameters are warranted by the calibration data. In particular, we will use the recently developed DREAM algorithm (Vrugt et al., 2009) with

ACPD
11, 4749–4806, 2011

synthetic data, and test the effect of measurement errors of the size distribution on the final parameter estimates of the adiabatic cloud model and their nonlinear uncertainty, and correlation.

Acknowledgements. We gratefully acknowledge the financial support of the Bert Bolin Centre for Climate research. The calculations presented herein have been made possible using the LISA cluster from the SARA centre for parallel computing at the University of Amsterdam, the Netherlands. Technical programming assistance from Hamish Struthers during model development to a different platform is appreciated. AS acknowledges support from an Office of Naval Research YIP award.

References

Andreae, M. O. and Rosenfeld, D.: Aerosol-cloud-precipitation interactions, Part 1, The nature and sources of cloud-active aerosols, Earth Sci. Rev., 89, 13–41, 2008.

Anttila, T. and Kerminen, V.-M.: On the contribution of Aitken mode particles to cloud droplet populations at continental background areas - a parametric sensitivity study, Atmos. Chem. Phys., 7, 4625–4637, doi:10.5194/acp-7-4625-2007, 2007.

Ayers, G. P. and Larson, T. V.: Numerical study of droplet size dependent chemistry in oceanic, wintertime, stratus cloud at southern midlattitudes, J. Atmos. Chem., 11(1–2), 143–167, 1990.

Arabas, S. and Pawlowska, H.: Adaptive method of lines for multi-component aerosol condensational growth and CCN activation, Geosci. Model Dev., 4, 15–31, doi:10.5194/gmd-4-15-2011, 2011.

Beven, K.: A manifesto for the equifinality thesis, J. Hydrol., 320(1–2), 18–36, 2006.

Bikowski, J., van der Kruk, J., Huisman, J. A., Vereecken, H., and Vrugt, J.A.: Inversion and sensitivity analysis of GPR data with waveguide dispersion using Markov Chain Monte Carlo simulation, Ground Penetrating Radar (GPR), 13th International Conference, 2010.

Cochran, R. N. and Horne, F. H.: Statistically weighted principle component analysis of rapid scanning wavelength kinetics experiments, Anal. Chem., 49(6), 846–853, 1977.

Conant, W. C., VanReken, T. M., Rissman, T. A., Varutbangkul, V., Jonsson, H. H., Nenes, A., Jimenez, J. L., Delia, A. E., Bahreini, R., Roberts, G. C., Flagan, R. C., and Seinfeld, J. H.:

ACPD
11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

Aerosol, cloud drop concentration closure in warm cumulus, J. Geophys. Res., 109(D13), D13204, doi:10.1029/2003JD004324, 2004.
Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., and Wikle, C. K.: Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling, Ecol. Appl., 19(3), 553–570, 2009.
Crump, J. G. and Seinfeld, J. H.: A New Algorithm for Inversion of Aerosol Size Distribution Data, Aerosol. Sci. Technol., 1(1), 15–34, 1982.
Duan, Q. Y., Sorooshian, S., and Gupta, V.: Effective and Efficient Global Optimization for Conceptual Rainfall-Runoff Models, Water Resour. Res., 28(4), 1015–1031, 1992.
Duan, Q. Y., Gupta, V. K., and Sorooshian, S.: Shuffled Complex Evolution Approach for Effective and Efficient Global Minimization, J. Optimiz. Theory App., 76(3), 501–521, 1993.
Dusek, U., Frank, G. P., Hildebrandt, L., Curtius, J., Schneider, J., Walter, S., Chand, D., Drewnick, F., Hings, S., Jung, D., Borrmann, S., and Andreae, M. O.: Size matters more than chemistry for cloud-nucleating ability of aerosol particles, Science, 312(5778), 1375–1378, 2006.
Feingold, G.: Modeling of the first indirect effect: Analysis of measurement requirements, Geophys. Res. Lett., 130(19), ASC7.1–ASC7.4, 2003.
Fitzgerald, J. W.: Effect of aerosol composition on cloud droplet size distribution – numerical study, J. Atmos. Sci., 31, 1358–1367, 1974.
Gupta, H. V., Sorooshian, S., and Yapo, P. O.: Toward improved calibration of hydrologic models: Multiple and noncommensurable measures of information, Water Resour. Res., 34(4), 751–763, 1998.
Hänel, G.: The role of aerosol properties during the condensational stage of cloud: a reinvestigation of numerics and microphysics, Beitr. Phys. Atmosph., 60, 321–339, 1987.
Hegg, D. A. and Larson, T. V.: The effects of microphysical parameterisation on model predictions of sulfate production in clouds, Tellus B, 42, 272–284, 1990
Heintzenberg, J., Covert, D. C., and van Dingenen, R.: Size distribution and chemical composition of marine aerosols: a compilation and review, Tellus B, 52(4), 1104–1122, 2000.
Helsper, C., Fissan, H., Kapadia, A., and Liu, B. Y. H.: Data Inversion by simplex minimization for the electrical aerosol analyzer, Aerosol Sci. Technol., 1, 135–146, 1982.
Hsieh, W. C., Nenes, A., Flagan, R. C., Seinfeld, J. H., Buzorius, G., and H., Jonsson.: Parameterization of cloud droplet size distributions: Comparison with parcel models and observations, J. Geophys. Res., 114, D11205, doi:10.1029/2008JD011387, 2009.

IPCC 2007: Climate Change: Summary for Policymakers, in: The Physical Science Basis, Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change, edited by: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K. B., Tignor M., and Miller, H. L., New York: Cambridge Univ. Press, 2007.

Jacobson, M. Z.: Development and application of a new air pollution modeling system-II, Aerosol model structure and design, Atmos. Environ., 31, 131–144, 1997.

Järvinen, H., Räisänen, P., Laine, M., Tamminen, J., Ilin, A., Oja, E., Solonen, A., and Haario, H.: Estimation of ECHAM5 climate model closure parameters with adaptive MCMC, Atmos. Chem. Phys., 10, 9993–10002, doi:10.5194/acp-10-9993-2010, 2010.

Kanakidou, M., Seinfeld, J. H., Pandis, S. N., Barnes, I., Dentener, F. J., Facchini, M. C., Van Dingenen, R., Ervens, B., Nenes, A., Nielsen, C. J., Swietlicki, E., Putaud, J. P., Balkanski, Y., Fuzzi, S., Horth, J., Moortgat, G. K., Winterhalter, R., Myhre, C. E. L., Tsigaridis, K., Vignati, E., Stephanou, E. G., and Wilson, J.: Organic aerosol and global climate modelling: a review, Atmos. Chem. Phys., 5, 1053–1123, doi:10.5194/acp-5-1053-2005, 2005.

Kandlikar, M. and Ramachandran, G.: Inverse methods for analysing aerosol spectrometer measurements: A critical review, J. Aerosol. Sci., 30(4), 413–437, 1999.

Köhler, H.: The nucleus in and the growth of hygroscopic droplets, Trans. Faraday Soc., 32, 1152–1161, 1936.

Korhonen, H., Kerminen, V.-M., Lehtinen, K. E. J., and Kulmala, M.: CCN activation and cloud processing in sectional aerosol models with low size resolution, Atmos. Chem. Phys., 5, 2561–2570, doi:10.5194/acp-5-2561-2005, 2005.

Kuczera, G.: On the Relationship between the Reliability of Parameter Estimates and Hydrologic Time-Series Data Used in Calibration, Water Resour. Res., 18(1), 146–154, 1982.

Laaksonen, A., Korhonen, P., Kulmala, M., and Charlson, R. J.: Modification of the Köhler equation to include soluble trace gases and slightly soluble substances, J. Atmos. Sci., 55(5), 853–862, 1998.

Laine, M. and Tamminen, J.: Aerosol model selection and uncertainty modelling by adaptive MCMC technique, Atmos. Chem. Phys., 8, 7697–7707, doi:10.5194/acp-8-7697-2008, 2008.

Lance, S., Nenes, A., and Rissman, T. A.: Chemical and dynamical effects on cloud droplet number: Implications for estimates of the aerosol indirect effect, J. Geophys. Res., 109, D22208, doi:10.1029/2004JD004596, 2004.

Lehtinen, K. E. J. and Kulmala, M.: A model for particle formation and growth in the atmosphere with molecular resolution in size, Atmos. Chem. Phys., 3, 251–257, doi:10.5194/acp-3-251-2003, 2003.

Loridan, T., Grimmond, C. S. B., Grossman-Clarke, S., Chen, F., Tewari, M., Manning, K., Martilli, A., Kusaka, H., and Best, M.: Trade-offs and responsiveness of the single-layer urban canopy parameterization in WRF: an offline evaluation using the MOSCEM optimization algorithm and field observations, Q. J. Roy. Meteorol. Soc., 136, 997–1019, doi:10.1002/qj.614, 2010.

Lu, M. L., Conant, W. C., Jonsson, H. H., Varutbangkul, V., Flagan, R. C., and Seinfeld, J. H.: The Marine Stratus/Stratocumulus Experiment (MASE): Aerosol-cloud relationships in marine stratocumulus, J. Geophys. Res., 112, D10209, doi:10.1029/2006JD007985, 2007.

Luo, Y. Weng, Q., E. S., Wu, X. W., Gao, C., Zhou, X. H., and Zhang, L.: Parameter identifiability, constraint, and equifinality in data assimilation with ecosystem models, Ecol. Appl., 19(3), 571–574, 2009.

McFiggans, G., Artaxo, P., Baltensperger, U., Coe, H., Facchini, M. C., Feingold, G., Fuzzi, S., Gysel, M., Laaksonen, A., Lohmann, U., Mentel, T. F., Murphy, D. M., O'Dowd, C. D., Snider, J. R., and Weingartner, E.: The effect of physical and chemical aerosol properties on warm cloud droplet activation, Atmos. Chem. Phys., 6, 2593–2649, doi:10.5194/acp-6-2593-2006, 2006.

Nenes, A., Ghan, S., Abdul-Razzak, H., Chuang, P. Y., and Seinfeld, J. H.: Kinetic limitations on cloud droplet formation and impact on cloud albedo, Tellus, 53, 133–149, 2001.

Nenes, A., Charlson, R. J., Facchini, M. C., Kulmala, M., Laaksonen, A., and Seinfeld, J. H.: Can chemical effects on cloud droplet number rival the first indirect effect?, Geophys. Res. Lett., 29(17), 1848, doi:10.1029/2002GL015295, 2002.

Pérez, C. J., Martín, J., and Rufo, M. F.: Sensitivity estimations for Bayesian inference models solved by MCMC methods, Reliab. Eng. Sys. Safe., 91, 1310–1314, 2006.

Pollacco, J. A. P. and Angulo-Jaramilo, R.: A Linking Test that investigates the feasibility of inverse modeling: application to a simple rainfall interception model for Mt Gambier, southeast South Australia, Hydrol. Process., 23(14), 2023–2032, 2009.

Pruppacher, H. R. and Klett, J. D.: Microphysics of Clouds and Precipitation (Atmospheric and Oceanographic Sciences Library), 976 pp., Springer, 1996.

Quinn, P. K., Bates, T. S., Coffman, D. J., and Covert, D. S.: Influence of particle size and chemistry on the cloud nucleating properties of aerosols, Atmos. Chem. Phys., 8, 1029–1042, doi:10.5194/acp-8-1029-2008, 2008.

ACPD
11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

Reutter, P., Su, H., Trentmann, J., Simmel, M., Rose, D., Gunthe, S. S., Wernli, H., Andreae, M. O., and Pöschl, U.: Aerosol- and updraft-limited regimes of cloud droplet formation: influence of particle number, size and hygroscopicity on the activation of cloud condensation nuclei (CCN), Atmos. Chem. Phys., 9, 7067–7080, doi:10.5194/acp-9-7067-2009, 2009.

Roelofs, G. J.: On the drop and aerosol size dependence of aqueous sulfate formation in a continental cumulus cloud, Atmos. Environ., 26A, 2309–2321, 1992.

Roelofs, G. J. and Jongen, S.: A model study of the influence of aerosol size and chemical properties on precipitation formation in warm clouds, J. Geophys. Res., 109, D22201, doi:0.1029/2004JD004779, 2004.

San Martini, F. M., Dunlea, E. J., Grutter, M., Onasch, T. B., Jayne, J. T., Canagaratna, M. R., Worsnop, D. R., Kolb, C. E., Shorter, J. H., Herndon, S. C., Zahniser, M. S., Ortega, J. M., McRae, G. J., Molina, L. T., and Molina, M. J.: Implementation of a Markov Chain Monte Carlo method to inorganic aerosol modeling of observations from the MCMA-2003 campaign - Part I: Model description and application to the La Merced site, Atmos. Chem. Phys., 6, 4867–4888, doi:10.5194/acp-6-4867-2006, 2006.

Šimúnek, J., Wendroth, O., and van Genuchten, M. T.: Parameter estimation analysis of the evaporation method for determining soil hydraulic properties, Soil Sci. Soc. Am. J., 62(4), 894–905, 1998.

Smith, P. L., Kliche, D. V., and Johnson, R. W.: The Bias and Error in Moment Estimators for Parameters of Drop Size Distribution Functions: Sampling from Gamma Distributions, J. Appl. Meteorol. Clim., 48(10), 2118–2126, 2009.

Sorooshian, S. and Gupta, V. K.: The Analysis of Structural Identifiability – Theory and Application to Conceptual Rainfall-Runoff Models, Water Resour. Res., 21(4), 487–495, 1985.

Sorooshian, S. and Arfi, F.: Response-Surface Parameter Sensitivity Analysis-Methods for Post-Calibration Studies, Water Resour. Res., 18(5), 1531–1538, 1982.

Sorooshian, S., Gupta, V. K., and Fulton, J. L.: Evaluation of Maximum-Likelihood Parameter-Estimation Techniques for Conceptual Rainfall-Runoff Models – Influence of Calibration Data Variability and Length on Model Credibility, Water Resour. Res., 19(1), 251–259, 1983.

Takeda, T. and Kuba, N.: Numerical Study of the Effect of CCN on the Size Distribution of Cloud Droplets .1. Cloud Droplets in the Stage of Condensation Growth, J. Meteorol. Soc. Jpn., 60(4), 978–993, 1982.

Tao, T., Golovitchev, V. I., and Chomiak, J.: A phenomenological model for the prediction of soot formation in diesel spray combustion, Combust. Flame, 136(3), 270– 282, 2004.

ACPD
11, 4749–4806, 2011

Tomassini, L., Reichert, P., Knutti, R., Stocker, T. F., and Borsuk, M. E.: Robust Bayesian uncertainty analysis of climate system properties using Markov chain Monte Carlo methods, J. Climate, 20(7), 1239–1254, 2007.
Toorman, A. F., Wierenga, P. J., and Hills, R. G.: Parameter-Estimation of Hydraulic-Properties from One-Step Outflow Data, Water Resour. Res., 28(11), 3021–3028, 1992.
Tunved, P., Nilsson, E. D., Hansson, H. C., Strom, J., Kulmala, M., Aalto, P., and Viisanen, Y.: Aerosol characteristics of air masses in northern Europe: Influences of location, transport, sinks, and sources, J. Geophys. Res.-Atmos., 110(D7), D07201, doi:10.1029/2004JD005085, 2005.
Twomey, S.: Comparison of Constrained Linear Inversion and an Iterative Nonlinear Algorithm Applied to Indirect Estimation of Particle-Size Distributions, J. Comput. Phys., 18(2), 188–200, 1975.
Voutilainen, A. and Kaipio, J. P.: Sequential Monte Carlo estimation of aerosol size distributions, Comput. Stat. Data An., 48(4), 887–908, 2005.
Voutilainen, A., Stratmann, F., and Kaipio, J. P.: A non-homogeneous regularization method for the estimation of narrow aerosol size distributions, J. Aerosol. Sci., 31(12), 1433–1445, 2000.
Vrugt, J. A., Bouten, W., and Weerts, A. H.: Information content of data for identifying soil hydraulic parameters from outflow experiments, Soil Sci. Soc. Am. J., 65(1), 19–27, 2001.
Vrugt, J. A., Schoups, G., Hopmans, J. W., Young, C., Wallender, W. W., Harter, T., and Bouten, W.: Inverse modeling of large-scale spatially distributed vadose zone properties using global optimization, Water Resour. Res., 40(6), W06503, doi:10.1029/2003WR002706, 2004.
Vrugt, J. A., Diks, C. G. H., Gupta, H. V., Bouten, W., and Verstraten, J. M.: Improved treatment of uncertainty in hydrologic modeling: Combining the strengths of global optimization and data assimilation, Water Resour. Res., 41(1), W01017, doi:10.1029/2004WR003059, 2005.
Vrugt, J. A., Nualláin, B., Ó. B., Robinson, A., Bouten, W., Dekker, S. C., and Sloot, P. M. A.: Application of parallel computing to stochastic parameter estimation in environmental models, Comput.Geosci., 32(8), 1139–1155, 2006.
Vrugt, J. A., Stauffer, P. H., Wohling, T., Robinson, B. A., and Vesselinov, V. V.: Inverse modeling of subsurface flow and transport properties: A review with new developments, Vadose Zone J., 7(2), 843–864, 2008.
Vrugt, J. A., Braak Ter., C. J. F., Diks, C. G. H., Robinson, B. A., Hyman, J. M., and Higdon, D.: Accelerating Markov chain Monte Carlo simulation by differential evolution with self-adaptive

ACPD
11, 4749–4806, 2011

randomized subspace sampling, Int. J. Nonlin. Sci. Num., 10, 273–290, 2009.
Wraith, D., Alston, C., Mengersen, K., and Hussein, T.: Bayesian mixture model estimation of aerosol particle size distributions, Environmetrics, 22, 23-34, doi:10.1002/env.1020, 2010.
Yapo, P. O., Gupta, H. V., and Sorooshian, S.: Automatic calibration of conceptual rainfall-runoff models: Sensitivity to calibration data, J. Hydrol., 181(1–4), 23–48, 2006.
Zhang, Y., Seigneur, C., Seinfeld, J. H., Jacobson, M. Z., and Binkowski, F. S.: Simulation of aerosol dynamics: A comparative review of algorithms used in air quality models, Aerosol Sci. Technol., 31, 487–514, 1999.

ACPD
11, 4749–4806, 2011

Table 1. Model parameter values used to generate synthetic data for marine and rural continental aerosol environments (bold), as well as lower and upper parameter bounds used to generate the grid over which the response surfaces were calculated.

	Environment	Marine			Rural continental		
Param. Number	Parameter	Lower Param. Value	True Value	Upper Param. Value	Lower Param. Value	True Value	Upper Param. Value
1	Surface Tension dyn cm^{-1}	20	70	100	20	70	100
2	Updraft (ms^{-1})	0.1	0.5	3	0.1	0.5	3
3	Particle Conc. M1(cm^{-3})	100.00	288.00	600.00	400.00	1010.00	2000.00
4	Radius: M1(nm)	10.00	20.00	50.00	10.00	23.70	50.00
5	GMD: Mode1	1.30	1.46	1.70	1.40	1.71	1.90
6	Particle Conc. M2(cm^{-3})	50.00	159.00	500.00	100.00	451.00	800.00
7	Radius: M2(nm)	50.00	82.50	100.00	50.00	89.80	100.00
8	GMD: M2	1.30	1.49	1.70	1.30	1.58	1.70
9	Soluble Mass Fraction	0.05	0.9	1	0.05	0.7	1

ACPD
11, 4749–4806, 2011

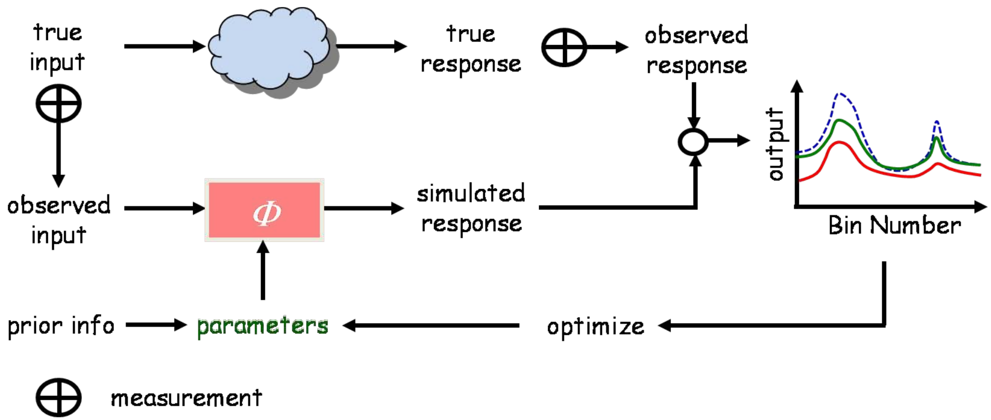

Fig. 1. A schematic representation of inverse modeling. The model parameters are iteratively adjusted so that the predictions of the model, Φ (represented with the solid line) approximate as closely and consistently as possible the observed response (interstitial-droplet number distribution over 400 bins represented with dotted line).

ACPD
11, 4749–4806, 2011

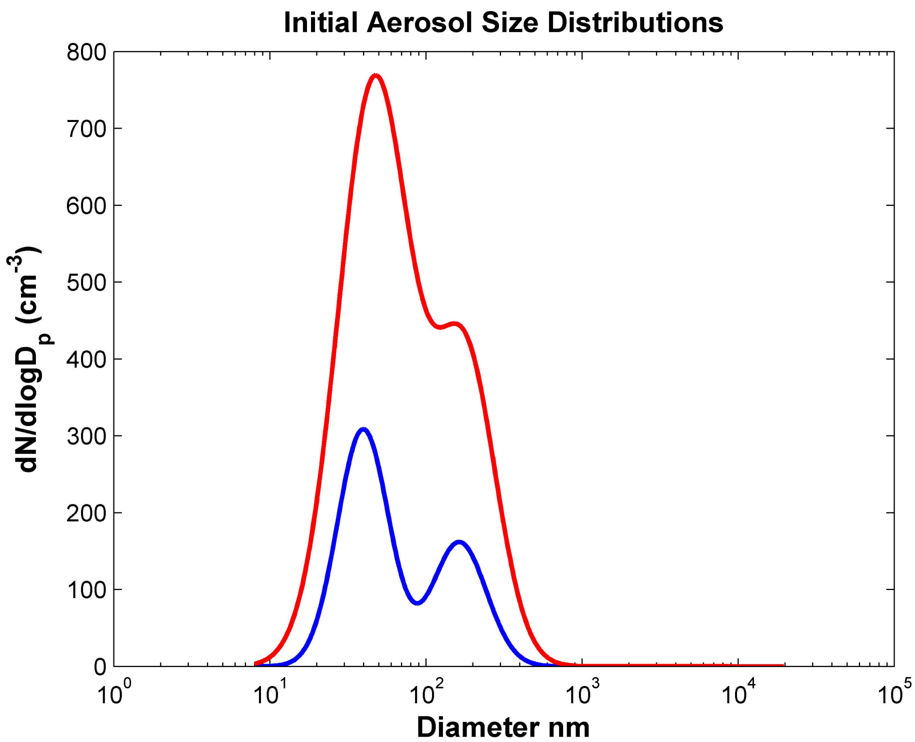

Fig. 2. The initial aerosol size distributions for marine (blue) and rural environments (red) used to calculate the respective calibration data sets.

ACPD
11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

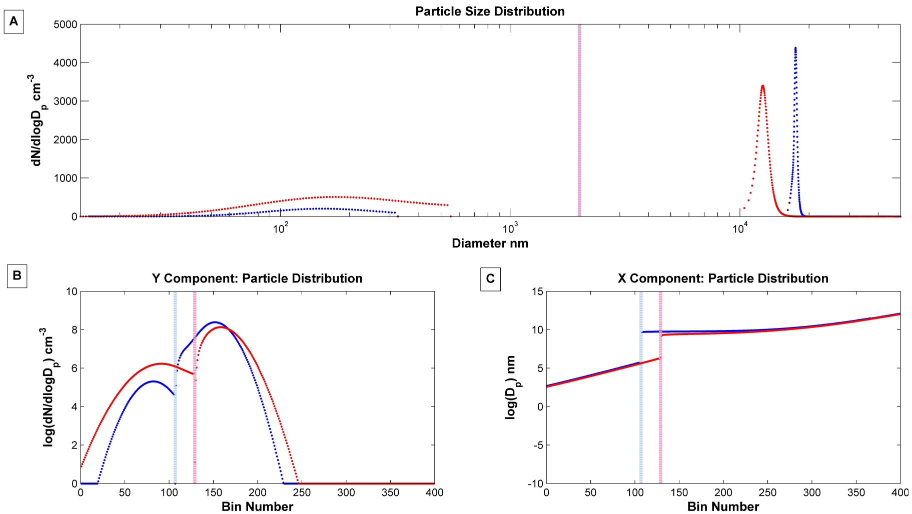

Fig. 3. (A–C): **(A)**: The $dN/d\log d_p$ size distribution for particles for both marine (blue curves) and rural continental conditions (red curves). Pink/Cyan dotted line represents location of 1 µm radius; the criteria used to select only activated droplets. **(B)**: The Y component in log space of the size distribution function. Cyan dotted line represents the bin number for which the cut-off point is determined in the calculation of the OF associated with particles greater than 1 µm. Pink dotted line represents this bin number for rural continental conditions. **(C)**: The X component in log space of the size distribution function. Dotted lines at same bin number for each environment as in panel **(B)**.

ACPD
11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

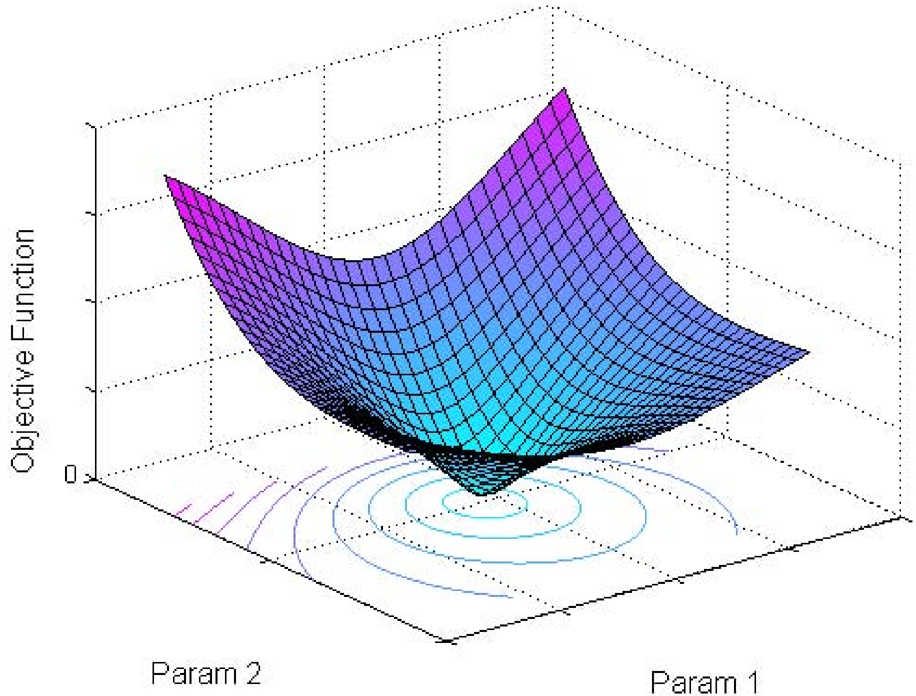

Fig. 4. A schematic illustrating a response surface which has a clear optimal solution to the inverse problem.

ACPD
11, 4749–4806, 2011

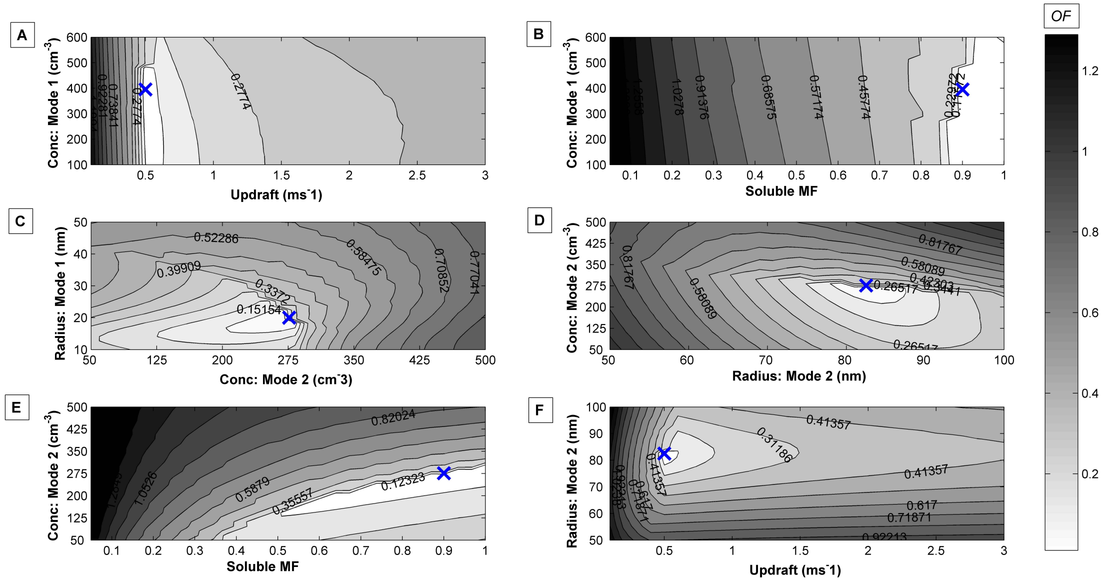

Fig. 5. **(A–F)**: 2-D response surface planes for a selection of parameter combination pairs: Marine Aerosol environment, cloud top. Blue cross in each denotes the true solution for each parameter that was used to generate the synthetic measurements.

ACPD
11, 4749–4806, 2011

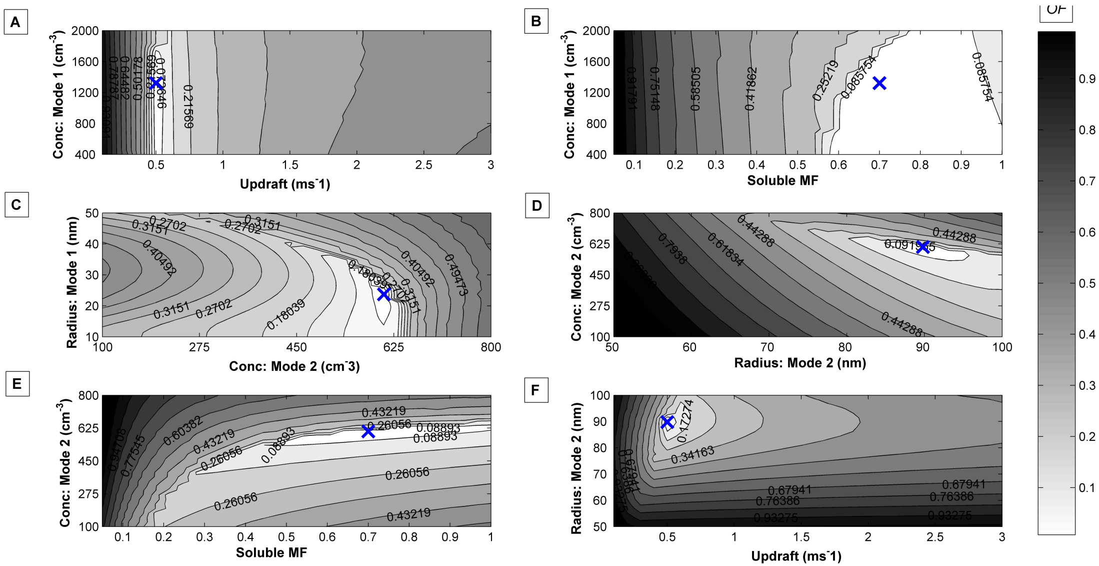

Fig. 6. (A–F): 2-D response surface planes for a selection of parameter combination pairs: Rural Continental Aerosol environment, cloud top. Blue cross in each denotes the true solution for each parameter that was used to generate the synthetic measurements.

ACPD
11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

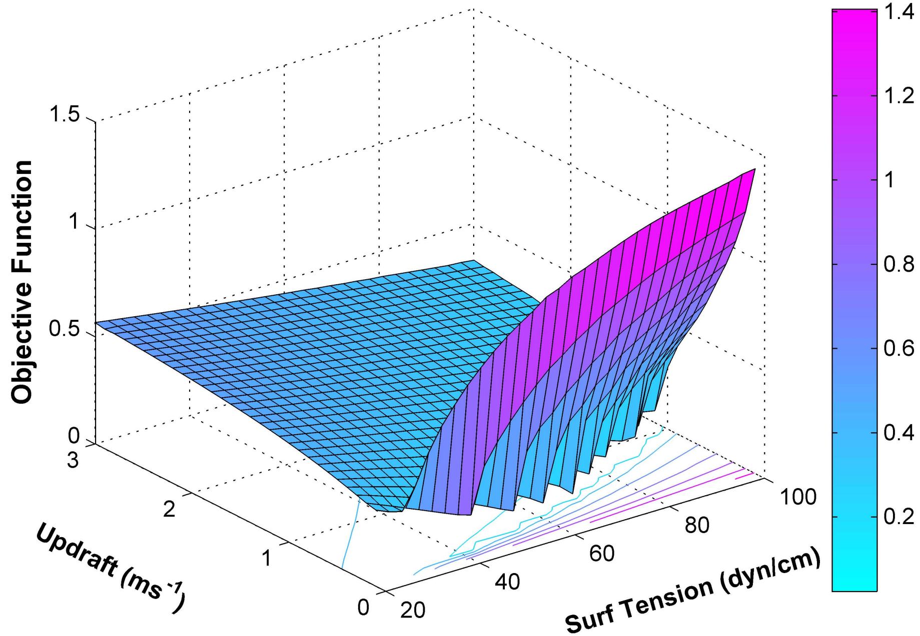

Fig. 7. Response surface of input updraft Vs surface tension: Marine Aerosol environment, cloud top.

ACPD
11, 4749–4806, 2011

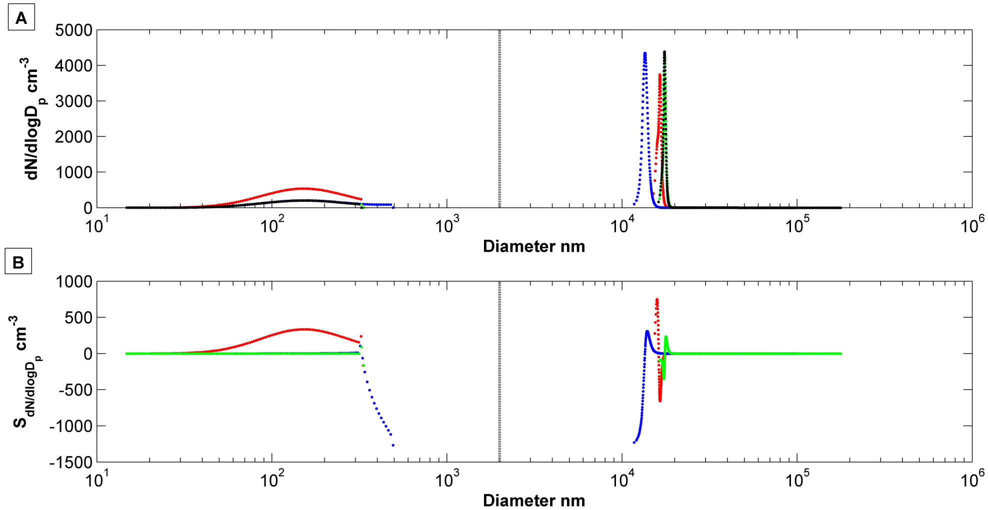

Fig. 8. (A–B): Marine Aerosol environment. **(A)**: The $dN/d\log d_p$ size distribution functions for the base case- black curve; input aerosol concentration (Mode 1) – red, aerosol concentration (Mode 2) – blue, aerosol mean radius (Mode 2) – green. **(B)**: The associated sensitivity of the $dN/d\log d_p$ size distribution functions to calibration parameters in plot **(A)**. The grey dotted lines represent 1 µm radius.

ACPD
11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

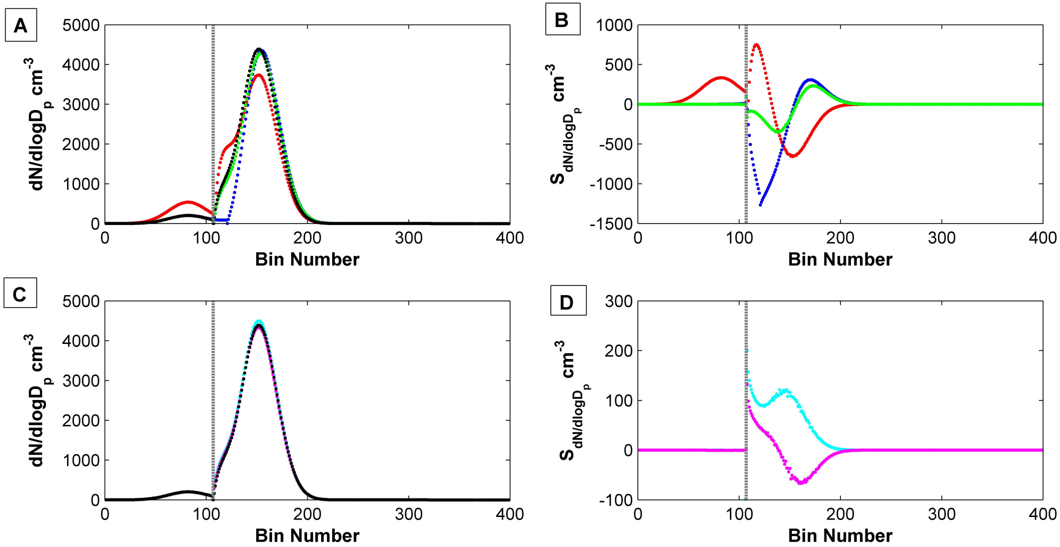

Fig. 9. (A–D): Marine Aerosol environment. **(A)**: The $dN/d\log d_p$ size distribution function for the base case- black curve; input aerosol concentration (Mode 1) – red, aerosol concentration (Mode 2) – blue, aerosol mean radius (Mode 2) – green. **(B)**: The associated sensitivity of the $dN/d\log d_p$ size distribution functions to calibration parameters in plot **(A)**. **(C)**: The $dN/d\log d_p$ size distribution functions for the base case- black curve; input updraft – cyan, aerosol soluble mass fraction– magenta. **(D)**: The associated sensitivity of the $dN/d\log d_p$ size distribution functions to calibration parameters in plot **(C)**. The grey vertical dotted line represents the 1 µm bin number for the calibration data (black line).

ACPD
11, 4749–4806, 2011

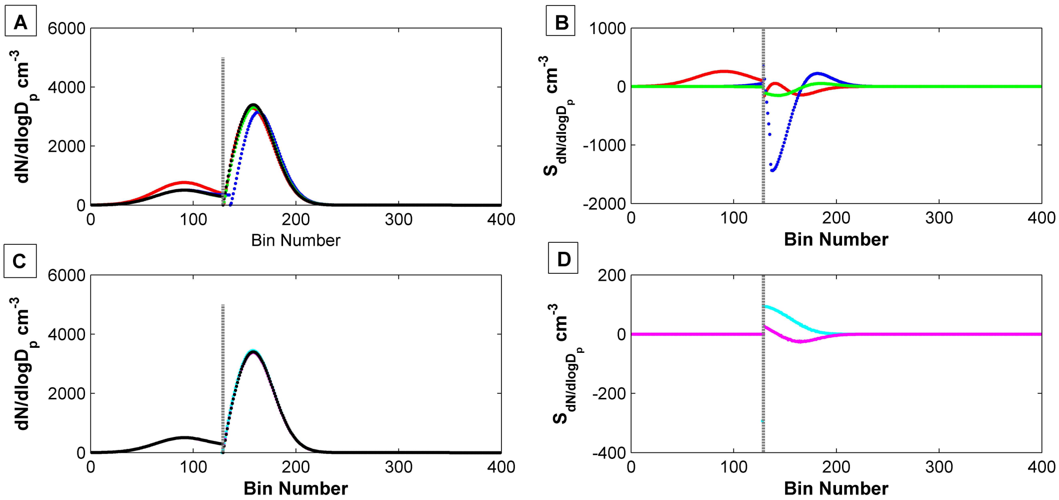

Fig. 10. (A–D):Rural Continenta Aerosol environment. **(A)**: The $dN/d\log d_p$ size distribution function for the base case- black curve; input aerosol concentration (Mode 1) – red, aerosol concentration (Mode 2) – blue, aerosol mean radius (Mode 2) – green. **(B)**: The associated sensitivity of the $dN/d\log d_p$ size distribution functions to calibration parameters in plot **(A)**: The $dN/d\log d_p$ size distribution functions for the base case- black curve; input updraft – cyan, aerosol soluble mass fraction– magenta. **(D)**: The associated sensitivity of the $dN/d\log d_p$ size distribution functions to calibration parameters in plot **(C)**. The grey vertical dotted line represents the 1 µm bin number for the calibration data (black line).

ACPD
11, 4749–4806, 2011

Towards inverse modeling of cloud-aerosol interactions – Part 1

D. G. Partridge et al.

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

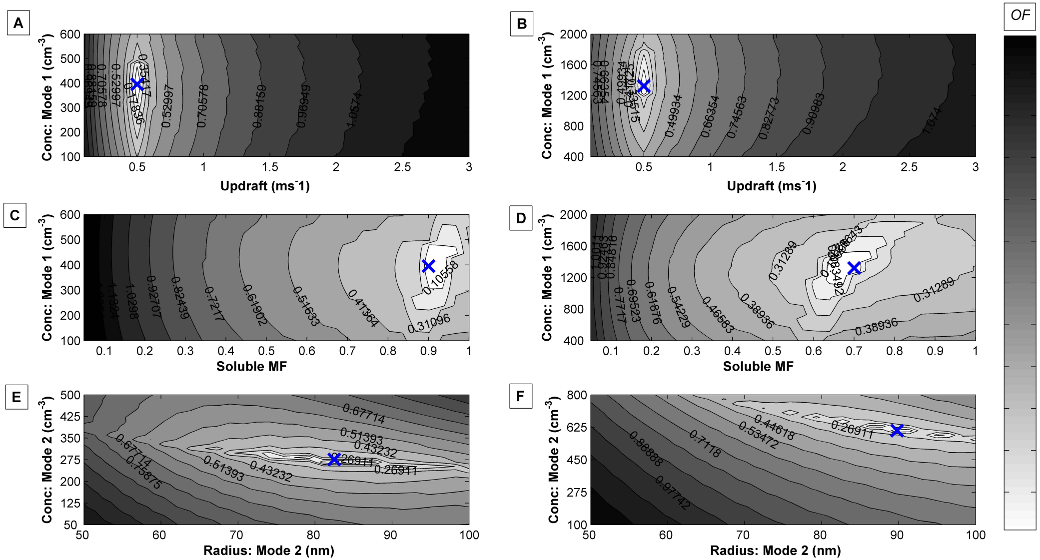

Fig. 11. (A–F): 2-D response surface planes for a selection of parameter combination pairs for when the interstitial aerosol is included in the calculation of the objective function. Marine Aerosol environment : Left panels; Rural Continental Aerosol environment: Right panels. Blue cross in each denotes the true solution for each parameter that was used to generate the synthetic measurements.

ACPD
11, 4749–4806, 2011

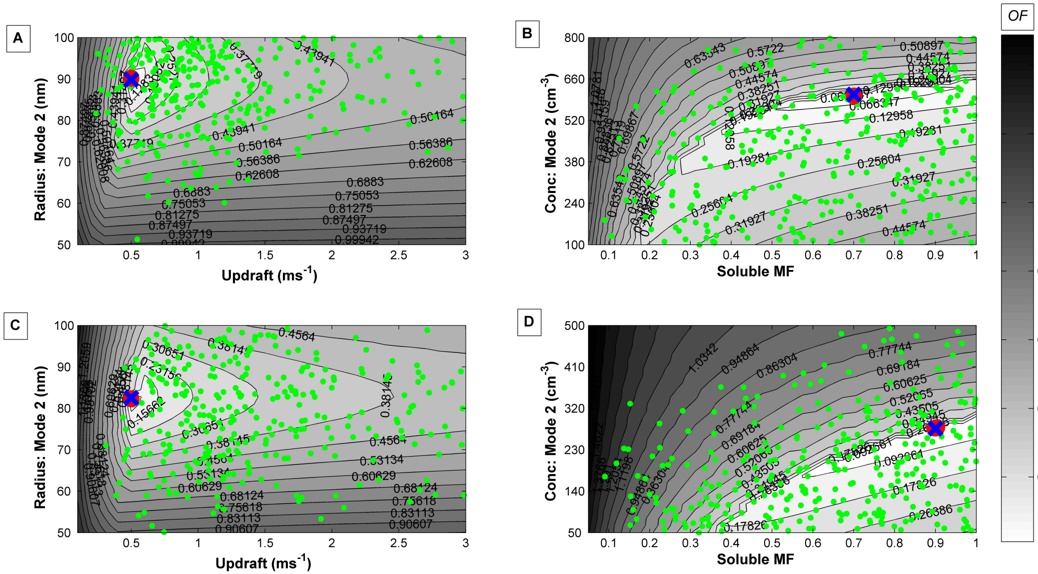

Fig. 12. (A–D): Lowest 10% of RMSE of 3000 LHS Monte Carlo simulations (green dots) overlaid on a 2-D response surface for two selected parameter pairs. Blue cross represents true solution, and red dot is the 10% lowest solutions from the SCE-UA optimisation algorithm. **(A, B)**: Rural Continental aerosol conditions; **(C, D)**: Marine aerosol conditions.

ACPD

11, 4749–4806, 2011