Atmos. Chem. Phys. Discuss., 11, 20051–20105, 2011
www.atmos-chem-phys-discuss.net/11/20051/2011/
doi:10.5194/acpd-11-20051-2011

Atmospheric Chemistry and Physics Discussions

Inverse modeling of cloud-aerosol interactions – Part 2: Sensitivity tests on liquid phase clouds using a Markov Chain Monte Carlo based simulation approach

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

[1]Dept. of Applied Environmental Science, Stockholm University, 10691 Stockholm, Sweden
[2]Bert Bolin Centre for Climate Research, Stockholm University, 10691 Stockholm, Sweden
[3]The Henry Samueli School of Engineering, Department of Civil and Environmental Engineering, University of California, Irvine, USA
[4]Computational Geo-Ecology, Institute for Biodiversity and Ecosystem Dynamics, University of Amsterdam, Amsterdam, The Netherlands
[5]Department of Meteorology, Stockholm University, Sweden
[6]Dept. of Chemical and Environmental Engineering, The University of Arizona, Tucson, USA
[7]Department of Atmospheric Sciences, The University of Arizona, Tucson, USA

Received: 17 June 2011 – Accepted: 6 July 2011 – Published: 15 July 2011

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

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

ACPD

11, 20051–20105, 2011

Abstract

This paper presents a novel approach to investigate cloud-aerosol interactions by coupling a Markov Chain Monte Carlo (MCMC) algorithm to a pseudo-adiabatic cloud parcel model. Despite the number of numerical cloud-aerosol sensitivity studies previously conducted few have used statistical analysis tools to investigate the sensitivity of a cloud model to input aerosol physiochemical parameters. Using synthetic data as observed values of cloud droplet number concentration (CDNC) distribution, this inverse modelling framework is shown to successfully converge to the correct calibration parameters.

The employed analysis method provides a new, integrative framework to evaluate the sensitivity of the derived CDNC distribution to the input parameters describing the lognormal properties of the accumulation mode and the particle chemistry. To a large extent, results from prior studies are confirmed, but the present study also provides some additional insightful findings. There is a clear transition from very clean marine Arctic conditions where the aerosol parameters representing the mean radius and geometric standard deviation of the accumulation mode are found to be most important for determining the CDNC distribution to very polluted continental environments (aerosol concentration in the accumulation mode $>1000\,\mathrm{cm}^{-3}$) where particle chemistry is more important than both number concentration and size of the accumulation mode.

The competition and compensation between the cloud model input parameters illustrate that if the soluble mass fraction is reduced, both the number of particles and geometric standard deviation must increase and the mean radius of the accumulation mode must increase in order to achieve the same CDNC distribution.

For more polluted aerosol conditions, with a reduction in soluble mass fraction the parameter correlation becomes weaker and more non-linear over the range of possible solutions (indicative of the sensitivity). This indicates that for the cloud parcel model used herein, the relative importance of the soluble mass fraction appears to decrease if the number or geometric standard deviation of the accumulation mode is increased.

ACPD
11, 20051–20105, 2011

This study demonstrates that inverse modelling provides a flexible, transparent and integrative method for efficiently exploring cloud-aerosol interactions efficiently with respect to parameter sensitivity and correlation.

1 Introduction

Clouds are recognised as one of the most important modulators of radiative processes in the atmosphere (Platnick and Twomey, 1994). Cloud reflectance is partially dependent on droplet size, which in turn is linked to the concentration of cloud condensation nuclei (CCN). The net effect of an increase in CCN is to increase cloud albedo (at fixed cloud liquid water path) generally resulting in a radiative cooling of the surface. In order to assess the impact of aerosols on clouds in the climate system, it is crucial to understand the underlying physical processes governing cloud-aerosol interactions. The ability of a particle to act as a CCN is a function of the size of the particle, its composition and mixing state, and the supersaturation of the air (Fitzgerald, 1974; Hegg and Larson, 1990; Laaksonen et al., 1998; Feingold, 2003; Conant et al., 2004; Kanakidou et al., 2005; Quinn et al., 2007). Untangling the relative importance of size and composition for the cloud nucleating ability of aerosol particles is at present a major challenge facing the cloud-aerosol modelling community, and this topic is at the core of the aerosol indirect effect (Dusek et al., 2006; McFiggans et al., 2006; Andreae and Rosenfeld, 2008; Stevens and Feingold, 2009).

Dusek et al. (2006) showed that particle size accounts for 84 to 96 % of observed variability in CCN concentrations. They hypothesised that aerosol-CCN relationships could be simplified by parameterising the effects of chemical composition on CCN activation for certain aerosol types. Modelling studies by Feingold (2003) and Ervens et al. (2005) also showed that for an internally-mixed aerosol, composition has a relatively small effect on droplet activation, except perhaps under conditions of both high pollution levels and small updraft velocities. However, Hudson (2007) presented a more extensive set of measurements that showed significantly more variability in the relationship

ACPD
11, 20051–20105, 2011

between dry particle size and critical supersaturation by including cleaner air masses in the analysis. Other studies have also shown that under certain combinations of meteorological/aerosol conditions the effect of chemistry may be relatively more important (e.g. Lance et al., 2004; Rissman et al., 2004; Twohy et al., 2008). In light of this, it is necessary to scrutinize and evaluate model parameters over a wide range of input and output conditions by efficiently searching the entire parameter space of relevant properties governing aerosol activation and growth.

The difficulty in untangling relationships among aerosols, clouds and precipitation has been attributed to the inadequacy of existing tools and methodologies (Stevens and Feingold, 2009). Numerous cloud-aerosol modeling sensitivity studies have been conducted (e.g. Feingold, 2003; Rissman et al., 2004 and references therein; Chuang, 2006), however, few have used statistical analysis tools to investigate the sensitivity of a cloud model to input aerosol parameters. There are two kinds of sensitivity analysis: local and global. The former studies input parameter variations across ranges that are believed to contain the appropriate values, while global sensitivity analysis considers input parameter changes over the entire multi-dimensional parameter domain (Pérez et al., 2006). When the local sensitivity to a set of model input parameters is tested, models are often run iteratively, perturbing one set of selected parameters at a time thus testing the sensitivity to these parameters individually. This approach requires prior knowledge as to how best to perturb each input parameter as the number of possible model permutations performed is usually limited. The selection of these values becomes more difficult if a parameter is non-measurable or if only limited or unreliable measurements exist.

Methods which explore the whole parameter space on the other hand have distinct advantages. Global sensitivity analysis generally leads to different, but more reliable results because parameter sensitivities in non-linear models of complex systems typically vary considerably over the feasible space of solutions. Secondly, if a model exhibits highly non-linear parameter interactions it is possible to account for parameter compensation by simultaneously varying parameters.

ACPD
11, 20051–20105, 2011

Few studies have used global sensitivity analysis to study cloud-aerosol interactions. One example is the study of Anttila and Kerminen (2007), which used the probabilistic collocation method (PCM) to test the sensitivity of cloud microphysics to Aitken mode particles (50–100 nm diameters). One of the main conclusions of their work is that parameters describing the aerosol number size distribution are generally more important than those describing chemical composition corroborating the results of e.g. Dusek et al. (2006), unless the particle surface tension or mass accommodation coefficient of water is strongly reduced due to the presence of surface-active organics. Despite the progress made, a polynomial approximation can never perfectly replace the original cloud-parcel model. Moreover, the parameters used in the polynomial function do not represent system properties, but are just fitting coefficients.

An alternative approach to global sensitivity analysis of cloud-aerosol interactions is to embrace an inverse modelling approach and invoke posterior probability density functions of model parameters using Markov Chain Monte Carlo simulation (MCMC). Such methods not only provide an estimate of the best parameter values, but also a sample set of the underlying (posterior) uncertainty. This distribution contains important information about parameter sensitivity, and correlation (interaction), and can be used to produce confidence intervals on the model predictions. The parameter sensitivity determined for the full dimensional parameter set is complimentary to the sensitivity derived from 2-D response surface analyses (Partridge et al., 2011, herein denoted P11).

MCMC approaches have found widespread application and use in a range of different disciplines to estimate posterior parameter distributions (Voutilainen and Kaipo, 2005; San Martini et al., 2006; Tomassini et al., 2007; Laine and Tamminen, 2008; Vrugt et al., 2008a; Wraith et al., 2009; Bikowski et al., 2010; Järvinen et al., 2010; Loridan et al., 2010; Vuollekoski et al., 2010).

Unfortunately, MCMC simulation requires significant computational resources and in addition, standard MCMC approaches are not particularly efficient, and typically require many thousands of model evaluations to find the posterior parameter distribution, even

ACPD
11, 20051–20105, 2011

for relatively simple problems. Therefore, it is paramount to test the performance and applicability of sophisticated state of the art MCMC algorithms for investigating cloud-aerosol parameter interactions.

P11 introduced an automatic parameter estimation framework to solve the cloud-aerosol inverse problem using the shuffled Complex Evolution (SCE-UA) global optimisation algorithm (Duan et al., 1992) in conjunction with a pseudo-adiabatic cloud parcel model (Roelofs and Jongen, 2004). Synthetic data was used to illustrate the methodology, and conclusive convergence to the appropriate parameters used to generate the synthetic data was demonstrated because we used artificially created data their true values were known a-priori. In P11 it was shown that without holding the lognormal parameters describing the Aitken mode, surface tension and updraft fixed at their true values it would be difficult to find the minimum of the objective function (OF). In particular, it was illustrated that the cloud-aerosol inverse problem is particularly difficult to solve because it is highly nonlinear, and may contain numerous local minima both within the immediate vicinity of the true solution, and far away. Although the SCE-UA algorithm was shown to successfully locate the optimum parameter values for the soluble mass fraction and lognormal aerosol parameters describing the accumulation mode, it does not provide an estimate of the underlying parameter uncertainty, associated with model nonlinearity, measurement and model error.

Explicit treatment of parameter uncertainty is possible if we adopt a Bayesian framework. In this study, we therefore pose the model calibration problem in a Bayesian framework, and use the DREAM adaptive MCMC sampling scheme (Vrugt et al., 2008b, 2009a) to approximate the posterior parameter distribution. This distribution contains the best parameter values found with SCE-UA, but also summarizes the associated parameter uncertainty. The method is used to compare the sensitivity of the pseudo adiabatic cloud parcel model to different input key parameters. The specific aims are as follows:

ACPD
11, 20051–20105, 2011

– Demonstrate that that DREAM, a current state of-the-art MCMC method, successfully solves the cloud-aerosol inverse problem, while simultaneously also providing estimates of parameter uncertainty and correlation.

– To demonstrate the applicability and power of MCMC to investigate cloud-aerosol interactions. We are particularly concerned with a global sensitivity analysis of the parameters describing the aerosol physiochemical properties.

– Pinpoint which are the dominant parameters controlling the activation of cloud droplets in different aerosol environments; from clean marine Arctic conditions to polluted continental conditions.

To the authors' knowledge this study is the first to use an MCMC framework with a pseudo-adiabatic cloud parcel model to summarize cloud-aerosol parameter and model uncertainty, and infer probability distributions of the determining factors that control the growth of droplets for different atmospheric conditions.

This paper will be presented in the following manner. First we will provide a brief introduction to inverse modelling using Bayesian inference. This will also include a detailed description of MCMC simulations using the DREAM algorithm, and a discussion about the choice of the OF.

This is followed by a short overview of the most important cloud – aerosol sensitivity tests that will be performed, followed by stepwise summary of the results. This will highlight the sensitivity of the cloud droplet number concentration (CDNC) distribution to the different calibration parameters followed by a section with the main findings and conclusions of the work considered herein.

ACPD
11, 20051–20105, 2011

2 Method

2.1 Bayesian inference

To start we provide a short summary of Bayesian inference. For a comprehensive review see e.g. Tamminen and Kyrölä, 2001; Jackson et al., 2004; Villagran et al., 2008. Bayesian inference represents a mathematically rigorous approach to parameter estimation. This statistical method treats the model parameters as random variables with a joint (but yet unknown) posterior probability distribution. This distribution is the product of the prior distribution and the likelihood function and conveys all desired information about the current knowledge of the parameters, and implicitly carries information about their best values (also called maximum likelihood), underlying uncertainty, and possible multi-dimensional correlation. The posterior probability density function of the parameters, hereafter referred to as $P(\theta|\boldsymbol{Y})$ can be written as follows using Bayes law:

$$p(\theta|\boldsymbol{Y}) = p(\theta) \times L(\theta|\boldsymbol{Y}) \quad (1)$$

where $P(\theta|\boldsymbol{Y})$ denotes the prior distribution of the parameters, and $L(\theta|\boldsymbol{Y})$ signifies the likelihood (objective) function. This function essentially measures the distance between the model predictions and corresponding observations. Many different formulations of this function are available in the (Bayesian) literature. Schoups and Vrugt (2009) recently introduced a generalized likelihood function that encapsulates most of these different formulations, and amongst others is especially developed to explicitly treat autocorrelation, heteroscedasticity, and non-Gaussianity of the residuals.

Once the posterior parameter distribution is known, model predictive uncertainty can be assessed by running samples of the posterior parameter distribution through the respective model, and inspecting the resulting range of the model predictions.

The prior distribution defines the knowledge about the parameters that is available before any data is collected or processed. This distribution typically constitutes information about the system of interest, and ensures that the parameter estimates at least partially adhere to prior knowledge. The likelihood function provides a diagnostic

ACPD
11, 20051–20105, 2011

measure of how well the model fits the data. The highest likelihood is generally found for those parameter values that provide the least squares fit to the experimental data. Additional observations (new evidence) are easily processed in this framework and will result in changes in the posterior parameter distribution. Hence, when confronted with new data, the likelihood function (and prior distribution) will likely change and alter parameter and predictive uncertainty.

In the past decade, much progress has been made in the development of efficient sampling methods that approximate the posterior distribution within a limited number of model evaluations. The Markov Chain Monte Carlo (MCMC) scheme was introduced by Metropolis et al. (1953), the basis of which is a Markov chain, which generates a random walk through the search space and successively visits solutions stemming from a fixed probability distribution (Vrugt et al., 2009b). This sampling procedure operates in two steps: (1) the proposal step: A candidate value is sampled from a "proposal distribution". (2) The acceptance/rejectance step: the candidate value is either accepted or rejected using the Metropolis acceptance probability (Järvinen et al., 2010). MCMC is especially designed to generate samples from the posterior distribution. This distribution contains important information about parameter sensitivity (width of the posterior distribution for each parameter), and correlation.

The original Metropolis MCMC scheme was extended for posterior inference in a Bayesian framework by Gelfand and Smith (1990), and has subsequently enjoyed widespread use in many fields of study (Vrugt et al., 2009a and references therein). MCMC algorithms are typically used to summarize parameter and model output uncertainty, without recourse to studying parameter sensitivities. A few studies exist that have used MCMC simulation to study "global" parameter sensitivities (Benke et al., 2008; Kanso et al., 2006; Vrugt et al., 2006, 2008b), yet such contributions are rather novel. This is rather remarkable as the posterior distribution directly conveys information about parameter sensitivity.

Existing theory and experiments prove the convergence of well constructed MCMC schemes to the appropriate limiting distribution under a variety of different conditions.

ACPD
11, 20051–20105, 2011

However, in practice this convergence is observed to be frustratingly slow, the efficiency being limited by the scale/orientation of the proposal distribution (Vrugt et al., 2009a). Slow convergence towards the correct target distribution is frequently caused by an inappropriate selection of the proposal distribution used to generate trial moves in the Markov Chain. This indicates the need for preliminary test runs or arduous hand tuning of the proposal distribution. Naturally this is a particular hindrance for the successful application of Bayesian inference for models that are CPU intensive, necessitating the use of more sophisticated and efficient MCMC methods which improve on the efficiency of older methods by employing adaptive techniques that "learn" during the sampling process. This allows the continuous adaptation of the shape/size of the proposal distribution such that the sampler more rapidly evolves towards the appropriate limiting distribution (Vrugt et al., 2009a). Convergence can also be hindered for inverse problems that contain numerous local minima in the posterior parameter space when using single chain MCMC methods. Gelman and Rubin (1992) advocate the use of MCMC algorithms that run multiple different Markov chains (trajectories) in parallel. This not only reduces the chance of getting stuck in local solutions, it also enables the use of a powerful array of statistical measures to diagnose convergence to a limiting distribution. For instance, a simple comparison of the within and in-between variances of the different chains will help judge whether the same distribution is being sampled by the different parallel chains.

Therefore, for the efficient investigation the cloud-aerosol inverse problem we employ a state of the art self adaptive DiffeRential Evolution Adaptive Metropolis algorithm (DREAM) (Vrugt et al., 2009a) in this study.

2.2 DiffeRential Evolution Adaptive Metropolis algorithm: DREAM

The DREAM sampling scheme is an adaptation of the Shuffled Complex Evolution Metropolis (SCEM-UA) global optimisation algorithm (Vrugt et al., 2003) but maintains detailed balance and ergodicity. The DREAM algorithm uses differential evolution as a genetic algorithm for population evolution with a Metropolis selection rule to

ACPD
11, 20051–20105, 2011

decide whether to accept the candidate points (offspring) or not. In DREAM, N different Markov chains are run in parallel, and jumps in each chain are generated using a fixed multiple of the difference of the states of one or more randomly chosen pairs of chains. The scale and orientation of this discrete proposal distribution is continuously changing en route to the posterior target distribution. The samples generated after convergence can be used to summarize the posterior distribution, and communicate parameter and model predictive uncertainty. The number of steps in each chain required to reach stationarity (convergence) is commonly called "burn-in", and these samples are removed from the analysis (Dekker et al., 2011).

Synthetic and real-world case studies have shown that this new approach elicits good efficiencies for complex, highly nonlinear, and multimodal target distributions (Vrugt et al., 2009a) typical for the parameters involved in cloud-aerosol interactions (P11). It is therefore well suited to the purpose of this investigation.

2.3 Pseudo-adiabatic cloud parcel model

Adiabatic cloud parcel models have been used successfully 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). To complete an MCMC simulation for a single cloud case with just a few calibration parameters, many thousands of cloud model evaluations are required to explore the posterior distribution. The computational requirements of MCMC could therefore hinder the use of CPU intensive models. In this paper, we utilize a computationally efficient pseudo-adiabatic cloud parcel model that provides a reasonable trade-off between processes accounted for, and computational speed. This provides us with flexibility to run different MCMC trials with different data sets, and calibration parameters. The chosen cloud parcel model (Roelofs and Jongen, 2004) 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. The reader is referred to P11 for a description of the model setup and to Roelofs and Jongen, 2004 for more information on the parcel model.

ACPD
11, 20051–20105, 2011

2.4 Calibration parameters

To test a wide range of input aerosol size distributions, data from four distinctively different aerosol environments were used as outlined in P11. These are:

1. Marine Arctic: summertime measurements performed at Ny-Ålesund, Svalbard (P. Tunved, personal communication, 2011).
2. Marine general: global measurements (Heintzenberg et al., 2000).
3. Rural continental: measurements from the well-established SMEAR II station at Hyytiälä (Tunved et al., 2005).
4. Polluted continental: summer continental air mass measurements from Melpitz station (Birmili et al., 2001).

The base value for all 10 input parameters of the pseudo-adiabatic cloud parcel model and the associated lower and upper limits for the four parameters to be optimised can be found in Table 1 for marine Arctic and marine general conditions; in Table 2 for rural continental and polluted continental environments. For each aerosol environment the base value, and lower/upper bounds for the lognormal parameters describing the accumulation mode were obtained using the statistics from P. Tunved, personal communication, 2011; Heintzenberg et al., 2000; Tunved et al., 2005; and Birmili et al., 2001 as a guide. The aerosol size distributions for marine Arctic, and marine general environments used to generate the synthetic CDNC distribution data can be found in P11. The base soluble mass fraction and upper and lower limits are selected based on values from the literature (P11). The only difference to P11 is that we constrain the prior limits for each environment so that they are more physically realistic. It is important that the prior limits are representative of the real atmosphere or else subsequently derived sensitivity may be misleading.

ACPD
11, 20051–20105, 2011

Synthetic calibration data

To benchmark our MCMC algorithm, it is useful to start the inverse modelling analysis with numerically generated cloud data (i.e. "synthetic" calibration data) simulated using known values of the model parameters. This is important to ensure that the subsequent sensitivity analysis is not contaminated by model error or parameter non-identifiability

The choice of the calibration data set essentially determines the posterior distribution of the parameters. More information available in the calibration data allows for more parameters to be constrained. On the contrary, noisy data with poor sensitivity to the individual parameters will result in uncertainty in the posterior distribution. Hence, in such situations it will be difficult to reduce parameter uncertainty, and appropriately calibrate the pseudo-adiabatic cloud parcel model. Thus, the information content of the calibration data directly determines the identifiability, uncertainty, and correlation of the pseudo-adiabatic cloud parcel parameters (P11).

The main thrust of this paper is to assess the impact of the calibration parameters on the number of activated cloud droplets and we therefore remove the interstitial aerosols from our calibration data set. This droplet size distribution is output at 100 m above cloud base as the calibration target.

To investigate the influence of environmental conditions on the posterior distribution and sensitivity of the governing pseudo-adiabatic cloud parcel model parameters we synthetically generate CDNC distributions using input from four different aerosol environments (cf. Sect. 2.4). The resulting CDNC distributions are depicted in Fig. 1.

2.5 Coupling pseudo-adiabatic cloud parcel model to MCMC algorithm

Figure 2 provides a schematic overview of the cloud-parcel parameter estimation problem using MCMC simulation with DREAM. The plot is essentially divided in two main parts. The top part corresponds to "the real-world" (in our case represented by synthetically generated data). The environmental conditions (denoted with "true input") act on the "real cloud" to produce a certain particle size distribution (dotted blue line, Fig. 2).

ACPD
11, 20051–20105, 2011

The terminology "true" and "observed" response is used to differentiate between reality and respective observations of reality that are prone to measurement error and uncertainty. Our framework thus explicitly recognizes the role of measurement error. The DREAM algorithm is now used to find those values of the pseudo-adiabatic cloud parcel parameters that provide the best possible fit to the measured droplet size distribution. This results in an ensemble of parameter values that define the posterior distribution.

Mathematically, the model calibration problem can be formulated as follows: Let $\hat{\boldsymbol{Y}}\phi(x,\theta) = \{\tilde{y}_1,\ldots,\tilde{y}_n\}$ denote predictions of the model Φ with observed input variables X and model parameters θ. Let, $\boldsymbol{Y} = \{y,\ldots,y_n\}$ 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:

$$\boldsymbol{E}(\theta) = G(\tilde{Y}) - G(Y) = \{G(\tilde{y}_1) - G(y_1),\ldots,G(\tilde{y}_n) - G(y_n)\} = \{e_1(\theta),\ldots,e_n(\theta)\} \quad (2)$$

where $G(.)$ allows for various monotonic (such as logarithmic) transformations of the output. The inverse modeling approach now relies on the estimation of the set of input parameters θ such that the measure $\boldsymbol{E}$, commonly called the objective function (OF), is in some sense forced to be as close to zero as possible.

We run the DREAM algorithm with the parameter bounds listed in Table 1 and with 10 different Markov chains and 75 000 cloud parcel model evaluations. Our experience with other parameter estimation problems of similar dimension suggests that these settings are appropriate. Such a setup completes an MCMC simulation in approximately two days using a standard desktop computer.

ACPD
11, 20051–20105, 2011

2.6 Defining the Objective Function (OF)

The most popular OF is the simple least squares (SLS) or maximum likelihood estimator. We follow this assumption and use the following definition of the OF:

$$\mathrm{OF} = \sum_{i=1}^{n} w_i [y_i - \Phi(X_i, \theta)]^2 = \sum_{i=1}^{n} w_i e_i(\theta)^2 \quad (3)$$

The θ vector consists of calibration parameters, for the cloud-aerosol inverse problem these are the input lognormal parameters describing the accumulation mode and soluble mass fraction, whereas the w_i denote weights associated with a particular measurement point. In this study in which we use a synthetically generated calibration dataset we assume the weights of the individual data points in Eq. (3) are similar and equal to one.

The identifiability of calibration parameters is highly dependent on the definition of the OF. Adiabatic cloud parcel models that employ a moving centre (MvCr) framework are particularly problematic for inverse modelling techniques as both the droplet radius and number are simultaneously changing in each run (P11).

For comparisons between different simulations to be meaningful, it is essential to construct a calibration data set that is constant with respect to the droplet size grid regardless of the prescribed calibration input parameters. If the OF is defined using only the raw MvCr output of the $dN/d\log D_\mathrm{p}$ function, without any radius information, then it is in theory possible to achieve exactly the same function shape for different parameter combinations, i.e. the calibration parameters are non-identifiable.

To avoid this, a direct interpolation of the droplet size distributions (Fig. 1) is performed so that the corresponding model predictions of the $dN/d\log D_\mathrm{p}$ size distribution function, $\hat{\boldsymbol{Y}} = \{\tilde{y}_1, \ldots, y_n\}$ are interpolated to the size grid of the calibration data, $\boldsymbol{Y} = \{y_1, \ldots, y_n\}$. Unfortunately this interpolation (termed "interpolation method" herein) results in poorly defined and chaotic response surfaces (P11) which result in non-identifiability problems. Tests show using the SCE-UA algorithm (P11) that for the

ACPD
11, 20051–20105, 2011

current definition of the calibration data we are restricted to studying only four parameters or the algorithm struggles to locate the calibration parameter values used to generate our calibration data for the perfect case (no measurement error).

In reality neither the adiabatic cloud model nor the measurements are perfect. Therefore, in order to investigate parameter sensitivity when using a synthetically generated calibration data set, it is necessary to corrupt the calibration data with a "measurement error", (Koda and Seinfeld, 1978). This is done in the following manner: First an error fraction is defined as 10 % of the calibration data, and thus a sigma vector, $\boldsymbol{\sigma}$, representing a synthetic variability is calculated as:

$$\boldsymbol{\sigma} = 0.10 \cdot \boldsymbol{Y}, \tag{4}$$

where the error is then defined from this σ using the Matlab *normrnd* function as:

$$\boldsymbol{Y}_{\text{error}} = normrnd(\mu, \sigma, N_Y) \tag{5}$$

where N_Y denotes the number of observations (cloud model size bin resolution) used to calibrate the cloud-parcel model. The *normrnd* function generates random numbers from the normal distribution with mean parameter $\mu = 0$ and standard deviation σ. To obtain our corrupted calibration data, the calibration data vector is corrupted with this measurement error vector by.

$$\boldsymbol{Y} = \boldsymbol{Y} + \boldsymbol{Y}_{\text{error}} \tag{6}$$

The simulation is then re-run to obtain the parameter sensitivity from the posterior distribution. This is calculated by allowing an 80 % burn in (cf. Sect. 2.2) of the MCMC simulation (i.e. we only take the last 20 % of the simulations at which point the algorithm has reached a stationary posterior distribution).

The algorithm was run first with no measurement error to be sure of convergence and then again with the calibration data corrupted with a synthetic measurement error (Eq. 4–6) to obtain parameter uncertainty.

ACPD
11, 20051–20105, 2011

3 Results

3.1 Performed sensitivity simulations and analysis

In this first study using DREAM we limit ourselves to investigating four parameters when using the "Interpolation method" for the OF definition. Thus, simulations and analysis will be presented for the calibration parameters deemed to be of most interest. Those are the number concentration, mean radius, and geometric standard deviation of the accumulation mode as well as the soluble mass fraction (cf. Tables 1–2). The analysis is performed for four aerosol environments (Sect. 2.4)

In the following, we will:

1. Show that MCMC simulation with DREAM converges nicely to the known parameter values (in the case of synthetic data) used to create the synthetic (artificial) data for marine general and rural continental aerosol environments.

2. Perform an initial sensitivity analysis of the calibration input parameters for marine general and rural continental environments.

3. Examine the posterior parameter distributions for all four aerosol environments in order to present a more detailed sensitivity analysis whilst concurrently revealing the effects of parameter compensation within the adiabatic cloud parcel model.

4. Repeat step 3 for a "low" and "high" updraft velocity to study the effect of updraft velocity on the derived sensitivity.

3.2 Performance of MCMC algorithm

In order to make sure that the MCMC algorithm can successfully find the true optimal solution (RMSE = 0) for every calibration parameter for the "perfect case" (no measurement error) simulations were initially performed for marine general and rural continental conditions using Eq. (3) for the calculation of the OF. The results are shown in Fig. 3

ACPD
11, 20051–20105, 2011

and Fig. 4 for marine general and rural continental data, respectively. The success of the MCMC algorithm in reaching the true synthetic calibration input parameter values is illustrated by the convergence of the Markov Chains (red lines) during the length of the simulation towards the true values (green dotted line). It is clear that when the size distribution measurements are not corrupted with a synthetic measurement error the single true optimal solution for every parameter is successfully located in fewer than 40 000 function evaluations.

The results from subsequent simulations in which we corrupted our calibration data with a 10 % synthetic measurement error were overlaid for each aerosol environment onto Figs. 3 and 4 respectively with blue dots so that the sensitivity bounds with respect to the true optimal solution are visible. The range on the Y-axis of each subplot in Figs. 3 and 4 corresponds to the prior range defined in Tables 1, 2 for marine general and rural continental conditions within which the algorithm is allowed to search. This means that the range of the posterior distribution (last 15 000 simulations) for a specific parameter in relation to the prior distribution (seen at iteration = 0) provides key information as to how sensitive the particle size distribution is to changes in a parameter. Since all input parameters are simultaneously optimised within this framework, a parameter whose posterior distribution has a small spread about the true solution is of high importance; as there are few combinations for which it can be defined in the model input and still get a measurement output which is close to the calibration data.

In order to visualise the convergence of the individual Markov Chains to the optimal solution we illustrate the evolution of the OF during the MCMC simulation in Fig. 5 for one parameter, the soluble mass fraction, for marine general conditions. This parameter creates some difficulties for the search algorithm as the convergence is not rapid. This can be explained by the fact that it is possible to achieve the same closeness of fit between the generated data set and the calibration data set (RMSE) for different values of the soluble mass fraction. It is hypothesised that this is attributed to parameter compensation within the cloud parcel model which will be discussed further in the following subsections.

ACPD
11, 20051–20105, 2011

3.3 Sensitivity analysis

3.3.1 Initial results from optimisation procedure

The parameter sensitivity is explored by corrupting our calibration data set with a synthetic measurement error (Eqs. 3–5).

Based on the width of the posterior distribution (cf. Sect. 3.2) it is clear from Figs. 3a and 4a that for both aerosol environments the key calibration parameter for describing the CDNC distribution is the number of particles in the accumulation mode, as its posterior range is the narrowest out of all calibration parameters relative to its prior range. Conversely for marine general conditions these figures indicate that the least important calibration parameter for our pseudo-adiabatic cloud parcel model is the soluble mass fraction. For rural continental conditions the difference between the widths of the posterior distributions is less clear.

The CDNC distribution associated with this posterior distribution is shown in Fig. 6 for all four aerosol environments. It is clear that the solutions stored within the posterior distribution bound the calibration data set for all aerosol conditions.

In order to confirm these preliminary indications and see the true relative sensitivity between different calibration parameters we normalise the posterior ranges by the prior ranges for each individual parameter. The last 20 % of the samples generated with DREAM are considered, thus a burn-in of 80 %. Burn-in is required to give the MCMC sampler time to converge to the posterior distribution. For our simulations this results in the last 15 000 parameter combinations stored in the individual chains. These parameter values correspond to the stationary distribution of the calibration parameters and can be used to define parameter and predictive sensitivity.

3.3.2 Parameter sensitivity

We will now explore the relative sensitivity between the parameters by investigating the posterior distribution for each of the four calibration parameters for each of the four

ACPD
11, 20051–20105, 2011

aerosol environments (Fig. 7). A larger normalised posterior range represents smaller sensitivity to a calibration parameter. It should be noted here that our normalised ranges used to infer parameter sensitivity are dependent on the prior range. It is for this reason that the prior ranges have to represent physically reasonable lower and upper limits for each parameter (cf. Sect. 2.4).

The results for marine general aerosol conditions (Fig. 7b) confirm those displayed in Fig. 3, i.e., for the pseudo-adiabatic cloud parcel model used in this study the particle concentration of the accumulation mode is the most important parameter for the activation of cloud droplets. The geometric standard deviation of the accumulation mode and soluble mass fraction are least important. For marine Arctic conditions (Fig. 7a) the results are similar; however, the geometric standard deviation is of larger importance. This low sensitivity to chemistry in clean CCN limited environments is intuitive; it does not matter how soluble a particle is if it does not exist. Thus, the number of particles must be, up to a certain threshold the limiting factor in any environment for the cloud droplet nucleating ability of an aerosol population. This will be especially true for environments in which the number of available cloud condensation nuclei (CCN) is limited (P11). This is also consistent with current observations and theory for cleaner environments (e.g. Dusek et al., 2006).

For rural continental conditions, the overall picture is the same, the number of aerosol particles in the accumulation mode is still the key parameter and the soluble mass fraction is the least important calibration parameter (Fig. 7c). However, now the soluble mass fraction is relatively more important. The geometric standard deviation is of equal importance as the mean radius, and there is a dramatic increase in the accumulation mode number normalised posterior range compared to marine general conditions. Moving to a yet further polluted environment (Fig. 7d) we see a shift in the dominant parameter for describing droplet activation to the parameter representing particle chemistry, with the difference between sensitivity of the lognormal aerosol parameters describing the accumulation mode decreasing further. The results are consistent with current theory for conditions in which the environment is polluted and the updraft

ACPD
11, 20051–20105, 2011

is relatively low, ($0.3\,m\,s^{-1}$). For more polluted aerosol conditions with a low updraft velocity the higher concentration of larger particles results in the activation of larger droplets, followed by a suppression of peak supersaturation which tends to reduce the total number of droplets activated. This allows for the soluble mass fraction to be relatively more important, in agreement with previous studies (Feingold, 2003; Lance et al., 2004; Ervens et al., 2005; Quinn et al., 2008). It is expected at higher updraft velocities a greater fraction of the larger aerosol would be able to achieve the critical supersaturation required for activation (regardless of composition), thereby decreasing the relative sensitivity of the aerosol composition compared to aerosol size (Antilla and Kerminen, 2007).

The evolution of the calibration parameter sensitivity from very clean (marine Arctic) to more polluted conditions is in keeping with our two dimensional response surface analysis of the sensitivity between number and chemistry (P11). This is caused by the shift from CCN limited to CCN saturated environments and the associated competition for water vapour. Also similar to P11 for which the updraft was relatively low, there is a clear tipping point in the calibration parameter sensitivity between marine general and rural continental conditions (Fig. 7).

For low updraft velocities ($0.3\,m\,s^{-1}$) the chemistry appears to be of similar importance to the accumulation mode mean radius for rural continental conditions, and more important than the accumulation mode mean radius for polluted continental conditions. This result highlights the importance of accurately representing the chemical composition of aerosols. For rural continental conditions the geometric standard deviation of the accumulation mode is only slightly less important for the cloud nucleating ability of particles than the mean radius. This is in agreement with the study of Antilla and Kerminen (2007).

ACPD
11, 20051–20105, 2011

3.4 Distribution of parameter values

Table 3 lists values of the derived posterior mean, minimum, maximum, coefficient of variation (CV) and maximum likelihood (ML) value of the four input parameters under investigation for all four aerosol environments. The ML value is the value associated with a calibration parameter that gave the best fit to the CDNC distribution stored in the calibration data. For all aerosol environments the soluble mass fraction has the highest coefficient of variation, showing the parameter to have the highest uncertainty within the posterior parameter distribution. Overall the minimum and maximum ranges after optimisation are generally more constrained for the two clean environments compared to the two more polluted conditions. In addition the ML of the soluble mass fraction for polluted continental conditions is 0.41, considerably lower than the true value of 0.60. This indicates that for polluted regions, more variability in the parameters describing the activation of cloud droplets is possible, whilst still achieving approximately the same CDNC distribution.

The ML value is very close to the true values of the calibration parameter for marine general and rural continental conditions (Table 3; Fig. 7b). For marine Arctic and rural conditions the ML value departs considerably from the true values for the soluble mass fraction (0.36 compared to 0.60). For polluted conditions the ML for the number concentration in the accumulation mode is 1457 cm^{-3}, ~250 cm^{-3} higher than the true value. The reason for this departure from the true value can be partially ascribed to the magnitude of the corrupted calibration data. The sigma values calculated using a 10 % error in Sect. 2.6 were generally positive, meaning that the corrupted droplet size distribution on average had a higher peak droplet number than the calibration data set. Therefore, it is logical that the MCMC algorithm tends towards a ML accumulation mode number concentration that is higher than the true value for this parameter. This is more noticeable for CCN saturated conditions for which there is a reduced sensitivity (higher uncertainty) to the particle concentration as more particles remain unactivated, staying within the interstitial size regime.

ACPD
11, 20051–20105, 2011

3.5 Parameter compensation and correlation

To explore these statistics further we derive the marginal distributions for each aerosol environment and present the results in Fig. 8. These histograms are derived by plotting the DREAM generated samples of each individual parameter. A marginal distribution that extends over the entire prior ranges is indicative for poor parameter sensitivity. On the contrary, if the histogram is well defined with narrow ranges, then this parameter is well defined, and sensitive to the calibration data. The marginal density is the probability distribution of the variables contained in our four dimensional inverse problem and provides us with counts of the calibration parameters values over their posterior distribution range, thus providing the shape of the posterior distribution. The scale and orientation of the inferred parameter distributions provide important diagnostic information about the structure of the adiabatic cloud parcel model under investigation.

For polluted continental aerosol conditions (subplots M-P) the histograms shows significant parameter variation across the posterior range indicating that there is a great range of possible aerosol physiochemical properties that can be considered optimal for the given environmental conditions. This results in a decrease in the relative importance of the aerosol parameters describing the accumulation mode distribution. The spread of the posterior distribution around the "correct" modal value for each calibration parameter is generally more constrained for cleaner aerosol conditions, especially for the lognormal aerosol parameters describing the accumulation mode. This indicates that for clean environments these parameters are particularly important for the accurate prediction of the droplet size distribution.

The shape of the marginal density distribution for each aerosol environment except marine Arctic indicates the presence of strong correlations between the four calibration parameters under investigation. For each of these three environments, many of the four calibration parameters depart from a normal distribution. The probability density is forced to accumulate at the parameter bounds to the left of the true optimal solution, causing the marginal distributions to be skewed.

ACPD
11, 20051–20105, 2011

This indicates that aerosol physiochemical properties within the pseudo-adiabatic cloud parcel model compensate each other to achieve the same CDNC distribution. To examine this in more detail consider Table 4 that presents correlation coefficients of the samples of the posterior parameter distribution for all environments. For each aerosol environment there are three calibration parameters that show significant co-variation (correlation coefficient $|r| > 0.6$) which have been highlighted in bold. For all except the most polluted conditions the parameters show significant correlation. For instance, consider the relationship between the soluble mass fraction with both number concentration and geometric standard deviation of the accumulation mode, and the correlation between the number and geometric standard deviation. For the polluted continental environment the correlation between the soluble mass fraction and geometric standard deviation is slightly lower (0.49), there is also a stronger relationship between the mean radius and geometric standard deviation of the accumulation mode. As three of the four aerosol environments share common correlations between different calibration parameters we present these three in the form of scatter plots for all conditions (Fig. 9). These scatter plots can potentially be used to gauge at which point within the parameter space a specific parameter used to describe the activation of cloud droplets becomes important for a certain atmospheric environment. All parameter combinations present in the posterior distribution shown in Fig. 9 give approximately the same cloud droplet size distribution for each aerosol environment respectively (Fig. 6).

The parameter combinations in the posterior distribution for the geometric standard deviation versus the number of particles in the accumulation mode show clear positive correlation for all four environments. Thus, in order to reach the same CDNC distribution it is necessary for both the number and geometric standard deviation to increase simultaneously. This is in agreement with other studies. For instance Quinn et al. (2008), reported that for a given mean particle diameter and total number concentration, increases in the geometric standard deviation lead to a decrease in the total droplet concentration because a broader mode suppresses the supersaturation due to the presence of more larger particles. The shape of the correlation across the

ACPD
11, 20051–20105, 2011

parameter space indicates the variation in sensitivity between the parameters. For the two cleaner environments, the relative importance of number increases as the geometric standard deviation increases; the inverse being true for the two more polluted environments. This is clearly shown by the increase in scatter for larger values of the geometric standard deviation for polluted continental conditions, and decrease for marine general conditions. This analysis highlights the importance of a proper representation of the geometric standard deviation for estimating the cloud nucleating ability of particles (cf. Sect. 3.3).

There is a strong negative relationship between the soluble mass fraction and the number of aerosol particles as well as the geometric standard deviation of the accumulation mode for all aerosol environments. There is a clear shift in the linearity of the correlation as we move into polluted environments which can be attributed to the increased sensitivity of the soluble mass fraction relative to the lognormal aerosol properties describing the accumulation mode. From the shape of the correlation for polluted continental conditions we can also see that the relative importance of the soluble mass fraction decreases if the number or the geometric standard deviation of the accumulation mode is increased. This is in agreement with current theory that for more polluted environments the effect of a decrease in supersaturation with a larger geometric standard deviation is larger in the presence of more large particles. Thus the ability for the soluble mass fraction to compensate in such conditions is reduced, evident from the increase in scatter in the posterior distribution. In Fig. 10 the relationship between all four calibration parameters is presented, clearly illustrating that in order to achieve the same CDNC distribution within the parameter thresholds provided by the posterior distribution these three parameters must compensate each other so that if the soluble mass fraction is reduced both the number of particles and geometric standard deviation must increase, and the mean radius of the accumulation mode must increase.

The correlations are less clear for marine Arctic conditions (Fig. 10a) and it is likely this can be partly attributed to the very narrow CDNC distribution and the loss of information caused by an interpolation of this function to a fixed size grid (P11).

ACPD
11, 20051–20105, 2011

The scatter plots presented in Fig. 10 illustrate that a wide range of aerosol physiochemical properties exists that result in very similar cloud microphysical properties. Therefore, for inverse modelling of cloud-aerosol interactions detailed measurements of cloud properties are required in order for the different clouds to be "unique". For instance height resolved measurements, size resolved chemistry, and interstitial aerosol measurements are all crucial.

In summary, the sensitivity analysis presented in Sect. 3 highlights that the size of the aerosol particle is only "sometimes" more important than its chemical composition. This must be considered in the future development of parameterisations used to calculate droplet number with respect to subsequent calculations of the aerosol indirect effect, thus it is paramount to estimate the importance of chemical effects for a variety of environments and meteorological conditions globally.

4 Effect of updraft velocity

As we cannot infer more than four parameters simultaneously given the limited information content of the data without non-identifiability contaminating our sensitivity analysis (P11), we now investigate what happens with the posterior parameter distributions if the updraft is changed to $0.15\,\mathrm{m\,s^{-1}}$ and to $0.60\,\mathrm{m\,s^{-1}}$, respectively. It is important to ascertain the effect of updraft on the sensitivity of the parameters describing the aerosol physiochemical characteristics as it has a strong influence on the number and size of cloud droplets formed (Rissman et al., 2004; Brenguier and Wood, 2009). We also showed from our initial response surface analysis (P11) that the CDNC distribution was most sensitive updraft perturbations.

To present the results from all updraft simulations simultaneously we calculate our relative sensitivity as 1-normalised posterior ranges for every parameter, and plot these against the accumulation mode number concentration for each aerosol environment (Fig. 11) it is clear that the relative importance of the chemistry compared to the accumulation mode radius increases for all aerosol environments when the updraft is

ACPD
11, 20051–20105, 2011

halved (Fig. 11a). When the updraft is doubled (Fig. 11c) the parameters become somewhat more identifiable with lesser dispersion of the posterior distribution. This is especially true for the number concentration of the accumulation mode. The soluble mass fraction increases in sensitivity when the updraft velocity is increased, although for the more polluted environments its importance relative to the remaining parameters is higher than for cleaner environments. This result can be explained by a combination of parameter compensation and CCN saturation of the aerosol environment. For clean aerosol conditions an increase in the value of the base updraft velocity (keeping all other values fixed) results in almost all of the available CCN becoming activated. Therefore, as the updraft isn't optimised during the MCMC simulation it cannot act as a limiting factor, and smaller perturbations in the remaining parameters will be amplified causing clean environments to exhibit higher sensitivity to changes in aerosol physiochemical properties.

To check this hypothesis a simple sensitivity analysis to the input parameters was performed as in P11 (figures not shown). For the higher updraft base case a small perturbation in these parameters resulted in a larger change on the CDNC distribution. This effect becomes weaker as the environment becomes more CCN saturated due to the effect of parameter compensation (cf. Sect. 3.5). The smaller relative change from halving the updraft compared to doubling it with respect to the base case (Fig. 11b) can be easily explained by the non-linear relationship between updraft and the accumulation mode concentration and soluble mass fraction as shown by our response surface analysis (P11), so that below a certain updraft value only small changes in the sensitivity will be observed. In summary, for low updraft the critical saturation is the limiting factor, whereas for high updraft conditions the non-linear physiochemical effects relating to the aerosol are limiting.

ACPD
11, 20051–20105, 2011

5 Discussion

The sensitivity analysis presented in Sects. 3 and 4 shows that the importance of the chemistry for the cloud nucleating ability of aerosol particles varies substantially as a function of both the aerosol concentration and the updraft velocity. We have probed an idealised cloud using synthetically generated CDNC distribution measurements with respect to four of the key calibration parameters of a pseudo-adiabatic cloud parcel model. The restricted number of input parameters was caused by the current "Interpolation method" definition of the OF (cf. P11).Improvements to the information content of the calibration data set are required in order to constrain more input parameters, for instance the updraft velocity, in order to allow a better understanding of cloud-aerosol interactions using MCMC. This is important parameter to include in our MCMC analysis since updraft is highly uncertain; it is both difficult to measure and highly variable (Lance et al., 2004). It has also been shown that for clean aerosol conditions the fraction of aerosols activated to droplets is a weak function of vertical velocity and a much stronger function of vertical velocity when aerosol concentrations are typical of polluted continental conditions (Snider and Brenguier, 2000). In relation to the parameter compensation (cf. Sect. 3.5) initial tests show that if the updraft is included, then for CDNC distributions generated with a high updraft the soluble mass fraction will be compensated by this parameter, the effect being stronger the more polluted the environment. This has implications with respect to future climate change as it can be envisaged that an increase in updraft velocities in clean aerosol environments could increase the susceptibility of the CDNC distribution to changes in aerosol physiochemical properties.

The strong correlation between three of the four parameters investigated in this synthetic study provides hope for simplification of parameterisations describing droplet activation (Kivekäs et al., 2007), and this motivates applying the MCMC to real world observations of cloud-aerosol properties. The strong parameter correlation and compensation for all aerosol environments also highlights the need for detailed measurements of cloud properties if we wish to constrain the cloud-aerosol inverse problem

ACPD
11, 20051–20105, 2011

using physically based cloud models. Future measurement campaigns should direct efforts towards measuring cloud microphysical properties at multiple height levels and include the interstitial aerosol and size resolved chemistry. The parameter compensation and correlation, in particular for polluted environments also highlights the difficulty in ascertaining the true parameter sensitivity using synthetic studies. Parameter compensation can mask sensitivity and in this respect the methodology presented in this study is complemented well by 2-D response surface analysis (P11).

In the future, improvements to the method to keep the size grid fixed without loss of information should be developed to in order for to allow more calibration parameters to be investigated, more efficiently.

6 Conclusions

In this study, we have coupled a state of-the-art MCMC algorithm, to a pseudo-adiabatic cloud parcel model. By using synthetically generated observations for marine Arctic, marine general, rural continental and polluted continental conditions, we have shown that the MCMC algorithm is able to efficiently locate the true optimal solution and the associated sensitivity of four of the most important input parameters for describing the development of a CDNC population in a pseudo-adiabatic cloud parcel model.

The most important advantages of the approach adopted are:

- MCMC algorithms can successfully be coupled with adiabatic cloud parcel models. This framework opens up new ways forward to investigate cloud-aerosol interactions.
- It is possible to simultaneously quantify both parameter sensitivity and cloud parcel model performance/investigate model structure. This framework results in a high level of transparency with respect to statistical inference of parameter uncertainty and correlation, and assessment model prediction uncertainty ranges.

ACPD
11, 20051–20105, 2011

– The ability of DREAM to search the entire parameter space significantly reduces the chance of getting stuck in local optima. Yet, population based search and optimization algorithms pose computational problems, particularly when the model requires significant time to run and produce the desired output.

– The DREAM algorithm is can be run in parallel, and distributed computing opens up new possibilities for solving complex, and computationally demanding parameter estimation problems.

The most important limitations are:

– Care must be taken that the parameter sensitivity results presented herein are dependent on the choice of the calibration data set, and likelihood (objective) function used.

– Sensitivity of the model to input parameters can be potentially masked by parameter correlation, thus care must be taken in the sensitivity analysis.

– To inspire confidence in the MCMC inverse modeling approach, a successful demonstration using real rather than synthetic measurements is required. This is a prerequisite to accurately predict cloud-aerosol interactions across a range of spatial scales.

We found strong correlations between certain input parameters, most notably the solubility versus the number and geometric standard deviation of the accumulation mode aerosol. In light of this it is crucial to improve our knowledge of the physical upper and lower limits of aerosol physio-chemical properties in the real atmosphere by performing more detailed measurements. This will ensure a better confidence in subsequently derived parameter sensitivity using MCMC methods.

The applied algorithm shows that for marine Arctic and marine general aerosol conditions the aerosol particle size and mean radius of the accumulation mode are the most important parameters when simulating the cloud droplet number concentrations,

ACPD
11, 20051–20105, 2011

whereas the chemical composition is the least important. However, for the present updraft applied ($0.3\,m\,s^{-1}$) in more polluted environments (aerosol concentration of the accumulation mode $> 400\,cm^{-3}$) the relative importance of the soluble mass fraction increases considerably. In CCN saturated conditions (aerosol concentration of the accumulation mode $> 1000\,cm^{-3}$) chemistry dominates the lognormal aerosol parameters describing the accumulation mode.

Whilst these main conclusions mostly confirm those obtained by previous studies, the method presented considers and displays a number of important findings in an integrative way, providing a visually clear way to deconstruct complex cloud-aerosol interactions into a visually simple form.

The results presented here are not derived using real-world cloud data, the findings so far being limited to synthetic cases only. In a related future study we will investigate cloud-aerosol interactions in an inverse framework using real measurements, the objective being to show whether the model input parameters match their measurement estimates when the model output is successfully optimised to the associated observations. Showing this to be the case improves our confidence in the adiabatic cloud parcel model and the parameters most important for describing the model output.

Acknowledgements. We gratefully acknowledge the financial support of the Bert Bolin Centre for Climate research. We gratefully appreciate G. J. Roelofs, IMAU, Utrecht, The Netherlands, for providing us with the pseudo-adiabatic cloud parcel model used in this study. Discussions with Thomas Loridan during the early stages of this work are greatly appreciated. AS acknowledges support from an Office of Naval Research YIP Award (N00014-10-1-0811). The authors acknowledge the Swedish Environmental Monitoring Program and Naturvårdsverket (Swedish environmental protection agency). The DREAM algorithm used herein can be obtained from the second author upon request.

ACPD

11, 20051–20105, 2011

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.

Benke, K. K., Lowell, K. E., Hamilton, A. J.: Parameter uncertainty, sensitivity analysis and prediction error in a water-balance hydrological model, Math. Comput. Modell., 47, 1134–1149, 2008.

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.

Birmili, W., Wiedensohler, A., Heintzenberg, J., and Lehmann, K.: Atmospheric particle number size distribution in Central Europe: statistical relations to air masses and meteorology, J. Geophys. Res., 106, 32005–32018, doi:10.1029/2000JD000220, 2001.

Brenguier, J. L. and Wood, R.: Observational strategies from the micro-to mesoscale, in: Clouds in the Perturbed Climate System: There relationship to Energy Balance, Atmospheric Dynamics and Precipitation, edited by: Heintzenberg. J. and Charlson, R. J., 487–510, MIT Press, Cambridge, Mass, 2009.

Chuang, P. Y.: Sensitivity of cloud condensation nuclei activation processes to kinetic parameters, J. Geophys. Res., 111(D9), D09201, doi:10.1029/2005JD006529, 2006.

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.: Aerosol, cloud drop concentration closure in warm cumulus, J. Geophys. Res., 109(D13), D13204, doi:10.1029/2003JD004324, 2004.

Dekker, S. C., Vrugt, J. A., and Elkington, R. J.: Significant variation in vegetation characteristics and dynamics from ecohydrological optimality of net carbon profit, Ecohydrology, doi:10.1002/eco.177, in press, 2011.

Duan, Q. Y., Sorooshian, S., and Gupta, V.: Effective and Efficient Global Optimization for

ACPD
11, 20051–20105, 2011

Conceptual Rainfall-Runoff Models, Water Resour. Res., 28(4), 1015–1031, 1992.
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.
Ervens, B., Feingold, G., and Kreidenweis, S. M.: The influence of water-soluble organic carbon on cloud drop number concentration, J. Geophys. Res., 110, D18211, doi:10.1029/2004JD005634, 2005.
Feingold, G.: Modeling of the first indirect effect: Analysis of measurement requirements, Geophys. Res. Let., Vol. 30, doi:10.1029/2003GL017967, 2003.
Fitzgerald, J. W.: Effect of aerosol composition on cloud droplet size distribution – numerical study, J. Atmos. Sci., 31, 1358–1367, 1974.
Gelman, A. and Rubin, D. B.: Inference from iterative simulation using multiple sequences, Stat. Sci., 7, 457–511, 1992.
Gelfand, A. E. and Smith A. F.: Sampling based approaches to calculating marginal densities, J. Am. Stat. Assoc., 85, 398–409, 1990.
Hegg D. A. and Larson T. V.: The effects of microphysical parameterisation on model predictions of sulfate production in clouds, Tellus, 42B, 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.
Hsieh, W. C., Nenes, A., Flagan, R. C., Seinfeld, J. H., Buzorius, G., and Jonsson, H.: Parameterization of cloud droplet size distributions: Comparison with parcel models and observations, J. Geophys. Res., 114, D11205, doi:10.1029/2008JD011387, 2009.
Hudson, J. G.: Variability of the relationship between particle size and cloud nucleating ability, Geophys. Res. Lett., 34, L08801, doi:10.1029/2006GL028850, 2007.
Jackson, C., Sen, M. K., and Stoffa, P. L.: An efficient stochastic Bayesian approach to optimal parameter and uncertainty estimation for climate model predictions, J. Climate, 17(14), 2828–2841, 2004.
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,

ACPD
11, 20051–20105, 2011

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.

Kanso, A., Ghebbo, G., and Tassin, B.: Application of MCMC-GSA model calibration method to urban runoff quality modeling, Reliability Engineering and System Safety, 91, 1398–1405, 2006.

Kivekäs, N., Kerminen, V.-M., Anttila, T., Hakola, H., Komppula, M., and Lihavainen, H.: Using Aerosol Number to Volume Ratio in Predicting Cloud Droplet Number Concentration, in: Nucleation and Atmospheric Aerosols, edited by: O'Dowd, C. D. and Wagner, P. E., 551–555, Springer, The Netherlands, 2007.

Koda, M. and Seinfeld, J. H.: Estimation of Urban Air-Pollution, Automatica, 14(6), 583–595, 1978.

Laaksonen, A., P. Korhonen, 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.

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.

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.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., and Teller, E.: Equation of state calculations by fast computing machines, J. Chem. Phys., 21, 1087–1092, 1953.

ACPD
11, 20051–20105, 2011

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.

Partridge, D. G., Vrugt, J. A., Tunved, P., Ekman, A. M. L., Gorea, D., and Sorooshian, A.: Towards inverse modeling of cloud-aerosol interactions – Part 1: A detailed response surface analysis, Atmos. Chem. Phys. Discuss., 11, 4749–4806, doi:10.5194/acpd-11-4749-2011, 2011.

Pérez, C. J., Martín, J., and Rufo, M. F.: Sensitivity estimations for Bayesian inference models solved by MCMC methods, Reliability Engineering and System Safety, 91, 1310–1314, 2006.

Platnick, S. and Twomey, S.: Determining the susceptibility of cloud albedo to changes in droplet concentration with the Advanced Very High Resolution Radiometer, J. Appl. Meteor., 33, 334–347, 1994.

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.

Rissman, T., Nenes, A., and Seinfeld, J. H.: Chemical amplification (or dampening) of the Twomey effect: Conditions derived from droplet activation theory, J. Atmos. Sci., 61(8), 919–930, 2004.

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:10.1029/2004JD004779, 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.

Snider, J. R. and Brenguier, J.-L.: Cloud condensation nuclei and cloud droplet measurements during ACE-2, Tellus B, 53, 828–842, 2000.

Stevens, B. and G. Feingold.: Untangling aerosol effects on clouds and precipitation in a buffered system, Nature, 461(7264), 607–613, 2009.

Tamminen, J. and E. Kyrölä.: Bayesian solution for nonlinear and non-Gaussian inverse

ACPD

11, 20051–20105, 2011

problems by Markov chain Monte Carlo method, J. Geophys. Res., 106, 14377–14390, doi:10.1029/2001JD900007, 2001.
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.
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.
Twohy, C. H. and Anderson, J. R.: Droplet nuclei in non-precipitating clouds: composition and size matter, Environ. Res. Lett., 3(4), 045002, doi:10.1088/1748-9326/3/4/045002, 2008.
Villagran, A., Huerta, G., Jackson, C. S., and Sen, M. K.: Computational Methods for Parameter Estimation in Climate Models, Bayesian Anal, 3(4), 823–850, 2008.
Voutilainen, A. and Kaipio, J. P.: Sequential Monte Carlo estimation of aerosol size distributions, Comput. Stat Data An, 48(4), 887–908, 2005.
Vrugt, J. A., Gupta, H. V., Bouten, W., and Sorooshian, S.: A Shuffled Complex Evolution Metropolis algorithm for optimization and uncertainty assessment of hydrologic model parameters, Water Resour. Res., 39(8), 1201, doi:10.1029/2002WR001642, 2003.
Vrugt, J. A., Clark, M. P., Diks, C. G. H., Duan, Q., and Robinson, B. A.: Multi-objective calibration of forecast ensembles using Bayesian model averaging, Geophys. Res. Lett., 33, L19817, doi:10.1029/2006GL027126, 2006.
Vrugt, J. A., Braak, C. J. F. T., Clark, M. P., Hyman, J. M., and Robinson, B. A.: Treatment of input uncertainty in hydrologic modelling: Doing hydrology backward with Markov chain Monte Carlo simulation, Water Resour. Res., 44, W00B09, doi:10.1029/2007WR006720, 2008a.
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, 2008b.
Vrugt, J. A., Braak, C. J. F. Ter., 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 randomized subspace sampling, Int. J. Non. Sci. Num. Sim., 10, 273–290, 2009a.
Vrugt, J. A., Robinson, B. A., and Hyman, J. M.: Self-Adaptive Multimethod Search for Global Optimization in Real-Parameter Spaces, IEEE Transactions on Evolutionary Computation,

ACPD
11, 20051–20105, 2011

13(2), 243–259, 2009b.
Vuollekoski, H., Boy, M., Kerminen, V. M., Lehtinen, K. E. J., and Kulmala, M.: MECCO: A method to estimate concentrations of condensing organics – Description and evaluation of a Markov chain Monte Carlo application, J. Aerosol Sci., 41(12), 1080–1089, 2010.
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, 2009.

ACPD
11, 20051–20105, 2011

Table 1. Model parameter values used to generate synthetic data for marine Arctic and marine general aerosol environments (bold), as well as their respective lower and upper bounds used to create posterior distributions derived with DREAM.

	Environment	Marine Arctic			Marine general		
	Parameter	Lower Limit	True Value	Upper Limit	Lower Limit	True Value	Upper Limit
1	Mass Accom Coefficient	N/A	**1.00**	N/A	N/A	**1.00**	N/A
2	Surface Tension m N m^{-1}	N/A	**70.00**	N/A	N/A	**70.00**	N/A
3	Updraft (ms-1)	N/A	**0.30**	N/A	N/A	**0.30**	N/A
4	N1 (cm^{-3})	N/A	**80.00**	N/A	N/A	**265.00**	N/A
5	R1 (nm)	N/A	**17.40**	N/A	N/A	**21.00**	N/A
6	GSD1	N/A	**1.43**	N/A	N/A	**1.45**	N/A
7	N2 (cm^{-3})	36.50	**74.50**	150.00	60.00	**165.00**	250.00
8	R2 (nm)	35.00	**48.00**	65.00	70.00	**82.50**	100.00
9	GSD2	1.50	**1.68**	1.85	1.40	**1.50**	1.60
10	Sol MF	0.30	**0.60**	1.00	0.45	**0.90**	1.00

ACPD
11, 20051–20105, 2011

Table 2. Model parameter values used to generate synthetic data for rural continental and polluted continental aerosol environments (bold), as well as their respective lower and upper bounds used to create posterior distributions derived with DREAM.

	Environment	Rural Continental			Polluted Continental		
	Parameter	Lower Limit	True Value	Upper Limit	Lower Limit	True Value	Upper Limit
1	Mass Accom Coefficient	N/A	**1.00**	N/A	N/A	**1.00**	N/A
2	Surface Tension m N m^{-1}	N/A	**70.00**	N/A	N/A	**70.00**	N/A
3	Updraft ($m\,s^{-1}$)	N/A	**0.30**	N/A	N/A	**0.30**	N/A
4	N1 (cm^{-3})	N/A	**1010.00**	N/A	N/A	**4900.00**	N/A
5	R1 (nm)	N/A	**23.70**	N/A	N/A	**33.00**	N/A
6	GSD1	N/A	**1.71**	N/A	N/A	**1.55**	N/A
7	N2 (cm^{-3})	215.00	**451.00**	690.00	730.00	**1200.00**	1600.00
8	R2 (nm)	75.00	**89.80**	105.00	75.00	**93.50**	105.00
9	GSD2	1.40	**1.58**	1.75	1.50	**1.55**	1.62
10	Sol MF	0.25	**0.70**	1.00	0.20	**0.60**	1.00

ACPD
11, 20051–20105, 2011

Table 3. Prior ranges and true values for each environment are presented under heading "Initial Range" for marine Arctic, marine general, rural continental and polluted continental conditions. Summary statistics of the derived final (posterior) distribution are also listed for each parameter.

Environment	Initial Range			Optimised Range: DREAM			DREAM	
Parameter	Min	Truth	Max	Min	Max	Mean	CV	ML
				Arctic				
7. N2 (cm^{-3})	36.50	**74.50**	150.00	72.55	93.36	83.89	0.05	85.68
8. R2 (nm)	35.00	**48.00**	65.00	45.06	53.51	49.27	0.03	49.91
9. GSD2	1.50	**1.68**	1.85	1.65	1.79	1.72	0.01	1.72
10. Sol MF	0.30	**0.60**	1.00	0.30	0.63	0.41	0.18	0.36
RMSE				38.81	40.12	39.20	0.01	38.81
				Marine general				
7. N2 (cm^{-3})	60.00	**165.00**	250.00	161.21	187.39	169.51	0.03	165.12
8. R2 (nm)	70.00	**82.50**	100.00	74.63	89.09	81.79	0.03	81.31
9. GSD2	1.40	**1.50**	1.60	1.47	1.60	1.52	0.02	1.50
10. Sol MF	0.45	**0.90**	1.00	0.47	1.00	0.80	0.15	0.89
RMSE				45.08	47.48	45.65	0.01	45.08
				Rural Continental				
7. N2 (cm^{-3})	215.00	**451.00**	690.00	394.78	606.27	454.61	0.09	428.87
8. R2 (nm)	75.00	**89.80**	105.00	75.89	105.00	96.55	0.06	104.85
9. GSD2	1.40	**1.58**	1.75	1.43	1.75	1.56	0.05	1.50
10. Sol MF	0.25	**0.70**	1.00	0.28	1.00	0.67	0.23	0.68
RMSE				54.18	57.07	54.76	0.01	54.18
				Polluted continental				
7. N2 (cm^{-3})	730.00	**1200.00**	1600.00	1074.10	1599.90	1392.10	0.10	1457.30
8. R2 (nm)	75.00	**93.50**	105.00	80.67	104.99	95.72	0.06	100.25
9. GSD2	1.50	**1.55**	1.62	1.50	1.62	1.57	0.02	1.58
10. Sol MF	0.20	**0.60**	1.00	0.31	0.74	0.51	0.17	0.41
RMSE				40.69	43.48	41.07	0.01	40.69

ACPD
11, 20051–20105, 2011

Table 4. Correlation structure induced between the parameters of the posterior distribution derived with DREAM for four contrasting environments including marine Arctic, marine general, rural continental and polluted continental environments". We separately list the best attainable values of the likelihood (objective) function (lowest value of the RMSE). Correlation coefficients larger than $|r| > 0.6$ are highlighted in bold.

Environment	N2	R2	GSD2	Sol MF	RMSE
			Arctic		
N2	1.00				
R2	0.17	1.00			
GSD2	**0.76**	0.49	1.00		
Sol MF	**−0.95**	−0.43	**−0.76**	1.00	
RMSE	−0.17	0.07	−0.06	0.22	1.00
			Marine general		
N2	1.00				
R2	0.08	1.00			
GSD2	**0.97**	−0.09	1.00		
Sol MF	**−0.98**	−0.23	**−0.93**	1.00	
RMSE	0.53	0.01	0.50	−0.48	1.00
			Rural Continental		
N2	1.00				
R2	−0.42	1.00			
GSD2	**0.94**	−0.67	1.00		
Sol MF	**−0.90**	0.14	**−0.81**	1.00	
RMSE	0.48	−0.30	0.45	−0.34	1.00
			Polluted continental		
N2	1.00				
R2	−0.40	1.00			
GSD2	**0.94**	**−0.62**	1.00		
Sol MF	**−0.72**	−0.31	−0.53	1.00	
RMSE	−0.15	−0.06	−0.08	0.16	1.00

ACPD
11, 20051–20105, 2011

Table A1. Prior ranges and true values for each environment are presented under heading "Initial Range" for marine Arctic, marine general, rural continental and polluted continental conditions. Summary statistics of the derived final (posterior) distribution are presented for rural continental and marine aerosol conditions.

Environment	Initial Range			Optimised Range: DREAM			DREAM	
Parameter	Min	Truth	Max	Min	Max	Mean	CV	ML
	Arctic: Updraft=0.15 m s^{-1}							
7. N2 (cm^{-3})	36.50	**74.50**	150.00	76.83	107.99	91.62	0.07	88.12
8. R2 (nm)	35.00	**48.00**	65.00	39.76	51.30	45.32	0.05	43.50
9. GSD2	1.50	**1.68**	1.85	1.73	1.85	1.78	0.01	1.76
10. Sol MF	0.30	**0.60**	1.00	0.30	0.65	0.44	0.21	0.51
RMSE				26.80	27.91	27.11	0.01	26.80
	Marine general: Updraft=0.15 m s^{-1}							
7. N2 (cm^{-3})	60.00	**165.00**	250.00	159.13	193.30	178.19	0.05	178.34
8. R2 (nm)	70.00	**82.50**	100.00	70.01	99.93	82.33	0.08	76.02
9. GSD2	1.40	**1.50**	1.60	1.47	1.60	1.56	0.02	1.57
10. Sol MF	0.45	**0.90**	1.00	0.45	1.00	0.68	0.19	0.80
RMSE				30.28	32.41	30.57	0.01	30.28
	Rural Continental: Updraft=0.15 m s^{-1}							
7. N2 (cm^{-3})	215.00	**451.00**	690.00	345.01	687.22	480.03	0.17	401.70
8. R2 (nm)	75.00	**89.80**	105.00	75.02	104.98	94.48	0.08	96.85
9. GSD2	1.40	**1.58**	1.75	1.45	1.75	1.60	0.05	1.53
10. Sol MF	0.25	**0.70**	1.00	0.25	0.94	0.56	0.27	0.73
RMSE				24.02	25.63	24.27	0.01	24.02
	Polluted continental: Updraft=0.15 m s^{-1}							
7. N2 (cm^{-3})	730.00	**1200.00**	1600.00	961.04	1599.90	1290.20	0.13	1087.60
8. R2 (nm)	75.00	**93.50**	105.00	81.54	105.00	97.32	0.06	98.78
9. GSD2	1.50	**1.55**	1.62	1.50	1.62	1.55	0.02	1.51
10. Sol MF	0.20	**0.60**	1.00	0.30	0.72	0.52	0.15	0.63
RMSE				23.06	24.53	23.35	0.01	23.06

ACPD
11, 20051–20105, 2011

Table A2. Prior ranges and true values for each environment are presented under heading "Initial Range" for marine Arctic, marine general, rural continental and polluted continental conditions. Summary statistics of the derived final (posterior) distribution are presented for each.

Environment	Initial Range			Optimised Range: DREAM			DREAM	
Parameter	Min	Truth	Max	Min	Max	Mean	CV	ML
	Arctic: Updraft=0.60 m s^{-1}							
7. N2 (cm^{-3})	36.50	**74.50**	150.00	72.62	95.66	81.85	0.06	79.52
8. R2 (nm)	35.00	**48.00**	65.00	42.18	51.34	46.38	0.03	46.79
9. GSD2	1.50	**1.68**	1.85	1.66	1.82	1.72	0.02	1.71
10. Sol MF	0.30	**0.60**	1.00	0.35	0.65	0.51	0.12	0.53
RMSE				63.47	65.60	64.01	0.01	63.47
	Marine general: Updraft=0.60 m s^{-1}							
7. N2 (cm^{-3})	60.00	**165.00**	250.00	157.86	169.26	162.08	0.01	162.42
8. R2 (nm)	70.00	**82.50**	100.00	76.20	85.47	80.79	0.02	80.17
9. GSD2	1.40	**1.50**	1.60	1.46	1.56	1.51	0.01	1.51
10. Sol MF	0.45	**0.90**	1.00	0.79	1.00	0.90	0.04	0.88
RMSE				77.81	82.82	78.72	0.01	77.81
	Rural Continental: Updraft=0.60 m s^{-1}							
7. N2 (cm^{-3})	215.00	**451.00**	690.00	439.43	569.61	489.88	0.05	459.88
8. R2 (nm)	75.00	**89.80**	105.00	76.82	104.96	91.26	0.08	87.02
9. GSD2	1.40	**1.58**	1.75	1.49	1.75	1.63	0.04	1.61
10. Sol MF	0.25	**0.70**	1.00	0.36	0.79	0.58	0.14	0.67
RMSE				74.40	77.67	75.16	0.01	74.40
Environment	**Polluted continental: Updraft=0.60 m s^{-1}**							
7. N2 (cm^{-3})	730.00	**1200.00**	1600.00	1108.40	1598.90	1302.10	0.07	1203.70
8. R2 (nm)	75.00	**93.50**	105.00	85.95	105.00	96.69	0.05	93.32
9. GSD2	1.50	**1.55**	1.62	1.50	1.62	1.56	0.02	1.55
10. Sol MF	0.20	**0.60**	1.00	0.33	0.73	0.55	0.13	0.61
RMSE				65.52	70.23	66.23	0.01	65.52

ACPD
11, 20051–20105, 2011

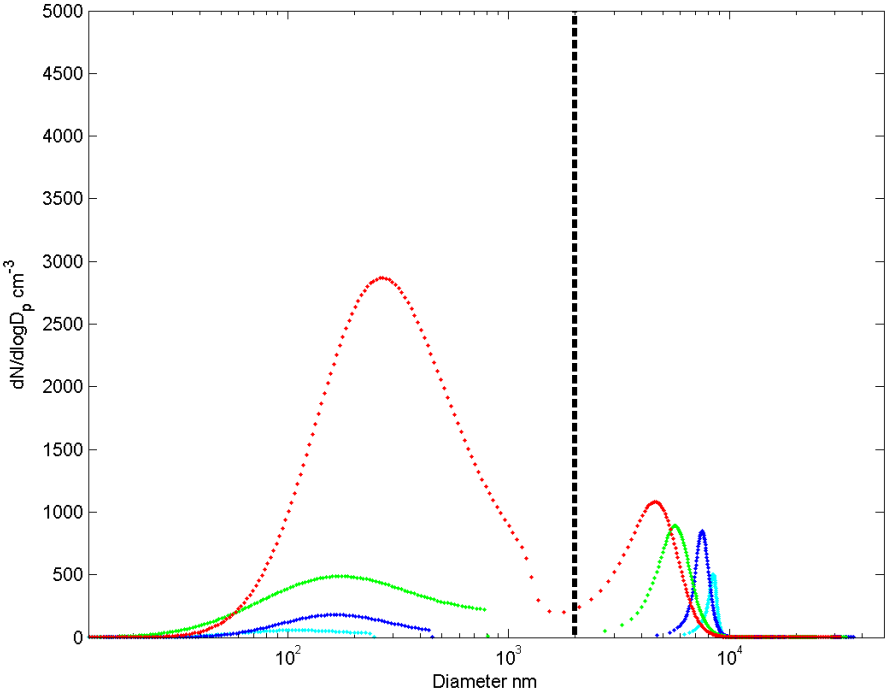

Fig. 1. The $dN/d\log D_p$ particle size distribution generated for marine Arctic (cyan), marine general (blue), rural continental (green) and polluted continental (red) aerosol environments. Black dotted line represents location of 1 µm radius.

ACPD
11, 20051–20105, 2011

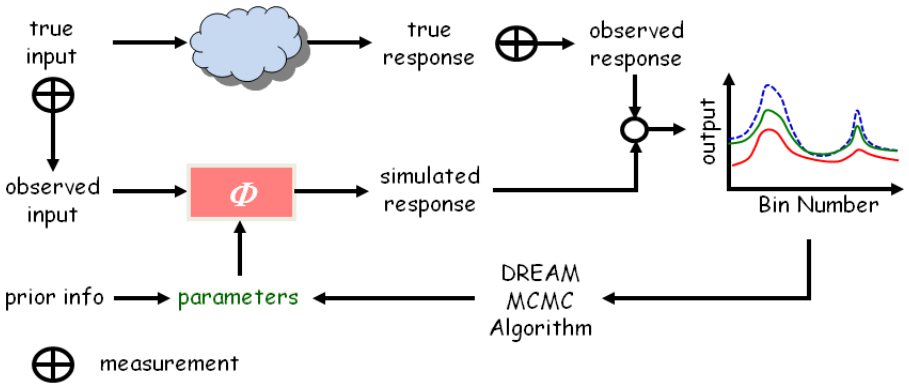

Fig. 2. A schematic representation of inverse modeling. The rectangular box in the bottom panel Φ represents the cloud – parcel model that is being used to predict the observed particle size distribution from given input data (also called forcing or boundary conditions), and some a-priori values of the model parameters. The model parameters are iteratively adjusted so that the predictions of the model,Φ (represented by the green and red solid lines) approximate as closely and consistently as possible the observed response (measured particle size distribution).

ACPD
11, 20051–20105, 2011

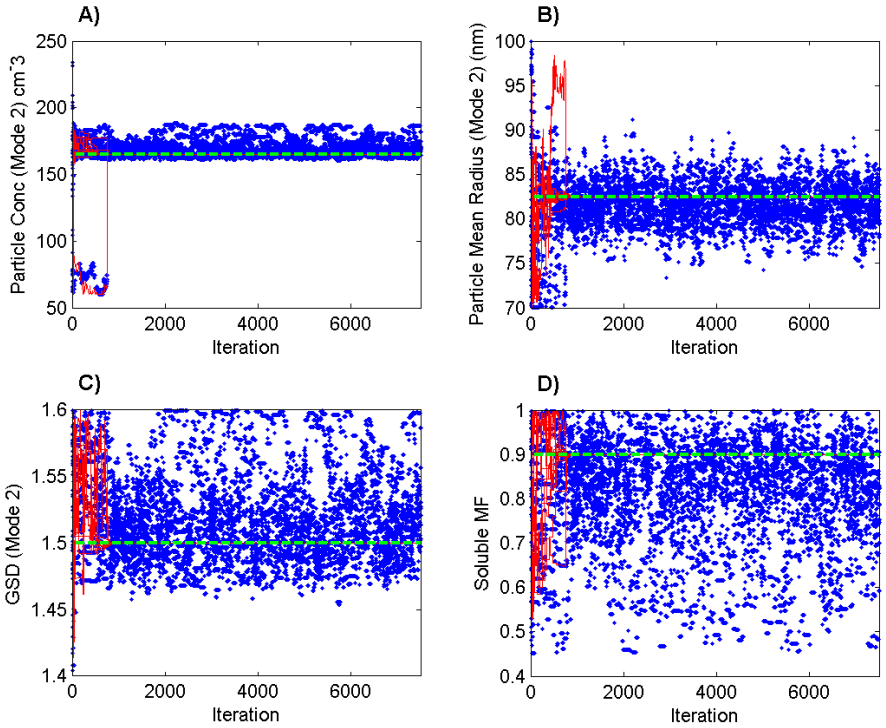

Fig. 3. Marine general aerosol environment. Evolution of the DREAM sampled Markov chains (different dots) towards the stationary posterior distribution of the lognormal parameters describing the accumulation mode and soluble mass fraction. Each panel considers a different parameter. The dashed green line represents the actual values of the calibration parameters used to generate the synthetic aerosol size distribution. The red line represents the convergence of DREAM algorithm when the calibration data set is not corrupted with a measurement error.

ACPD
11, 20051–20105, 2011

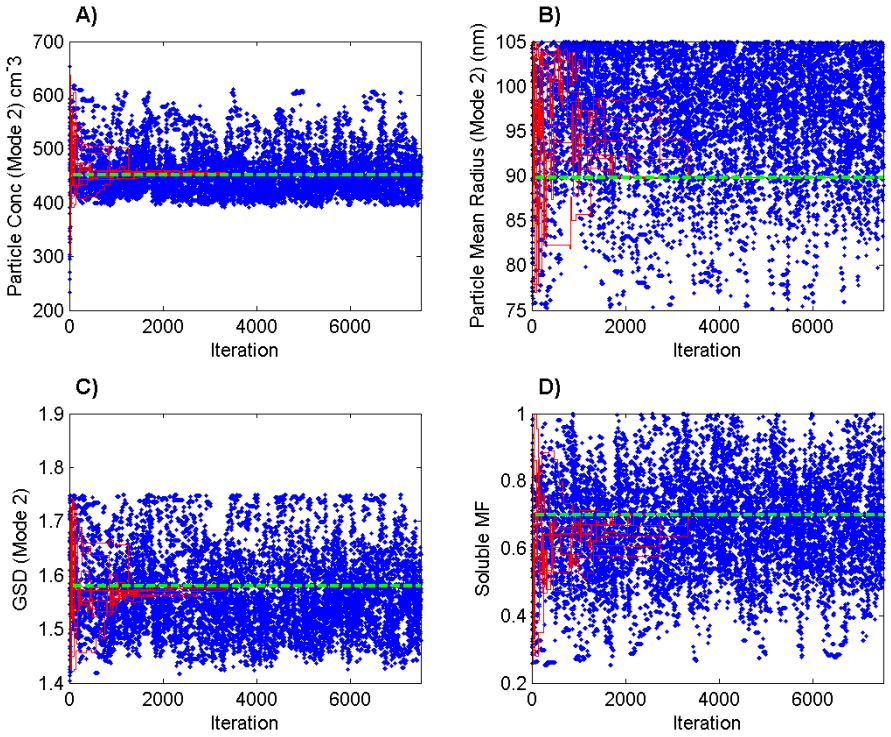

Fig. 4. Rural Continental aerosol environment. Evolution of the DREAM sampled Markov chains (different dots) towards the stationary posterior distribution of the lognormal parameters describing the accumulation mode and soluble mass fraction. Each panel considers a different parameter. The dashed green line represents the actual values of the calibration parameters used to generate the synthetic aerosol size distribution. The red line represents the convergence of DREAM algorithm when the calibration data set is not corrupted with a measurement error.

ACPD
11, 20051–20105, 2011

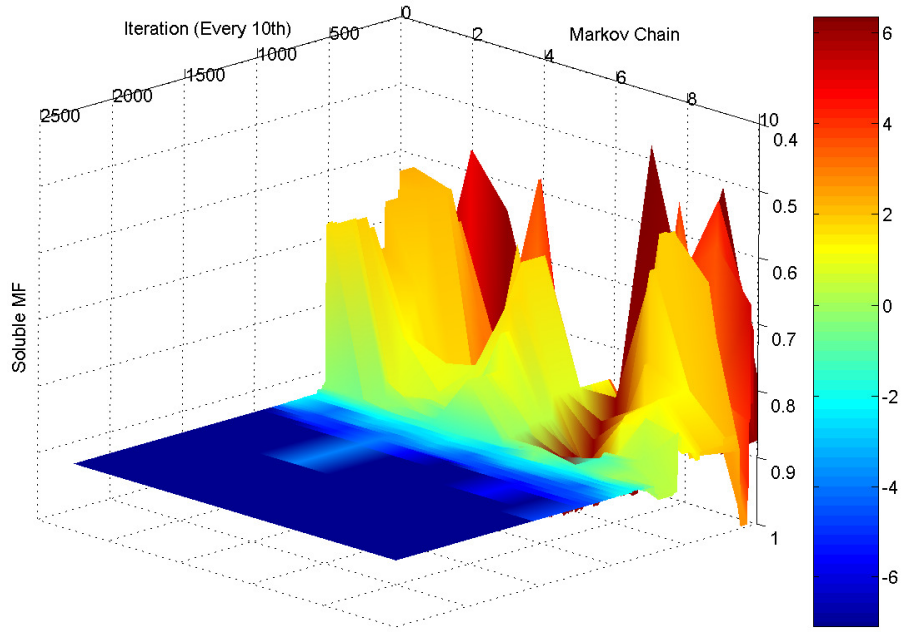

Fig. 5. Marine general aerosol environment: evolution of the DREAM generated Markov chains to the stationary posterior distribution. The y-axis (soluble mass fraction) is up-side-down, and the OF was transformed to the log-space. This made the interpretation easier.

ACPD
11, 20051–20105, 2011

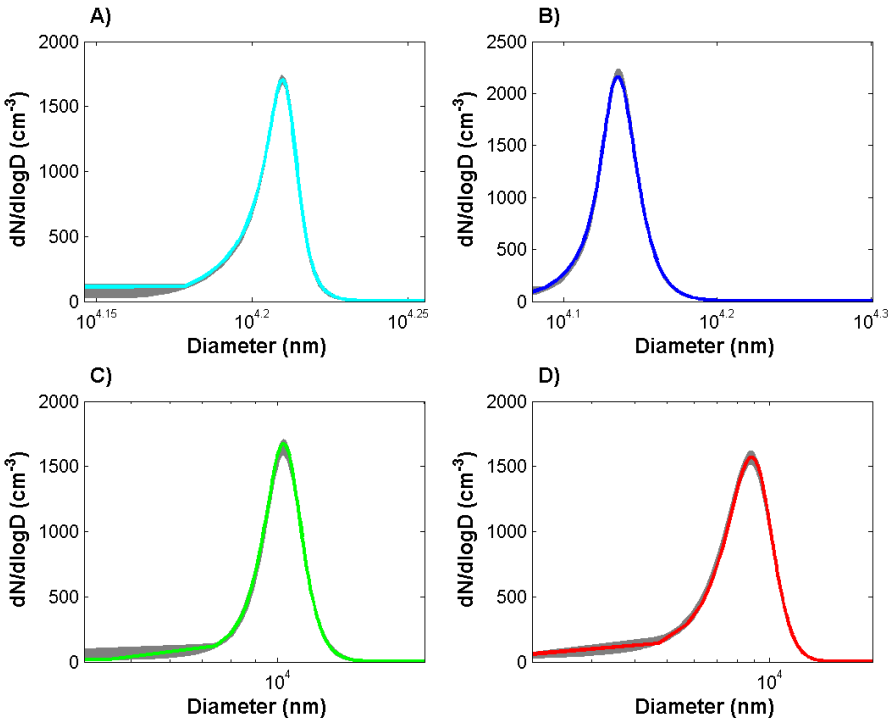

Fig. 6. The range of droplet size distributions associated with the posterior parameter distribution values (last 20 % of the posterior samples derived with DREAM). **(A)** Marine Arctic aerosol environment (cyan), **(B)** marine general aerosol environment (blue), **(C)** rural continental aerosol environment (green), **(D)** polluted continental aerosol environment (red). Coloured lines represent the calibration data set (synthetic observations). Each grey line represents one sample from the posterior distribution.

ACPD
11, 20051–20105, 2011

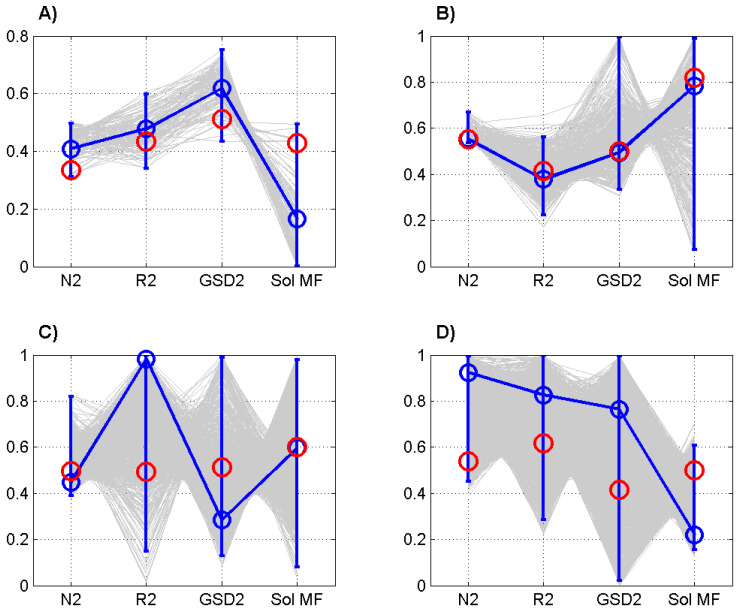

Fig. 7. Normalized posterior parameter ranges for **(a)** Marine Arctic aerosol environment, **(b)** Marine general aerosol environment, **(c)** Rural Continental aerosol environment, and **(d)** polluted continental aerosol environment. The last 20 % of the samples generated with DREAM were used to derive the results. The y-axes are scaled between 0 and 1 using the prior ranges defined in Table 1 to yield normalized ranges. The blue error-bars represent define the 1 %–99 % limits of the posterior distribution. The blue circles are used to signify the maximum likelihood values of the parameters that provide the closest fit (lowest RMSE) to the measured aerosol size distribution, whereas the red circles denote the true parameter values used to create the synthetic data. Each grey line going from left to right through each panel is a different parameter sample from the posterior distribution.

ACPD
11, 20051–20105, 2011

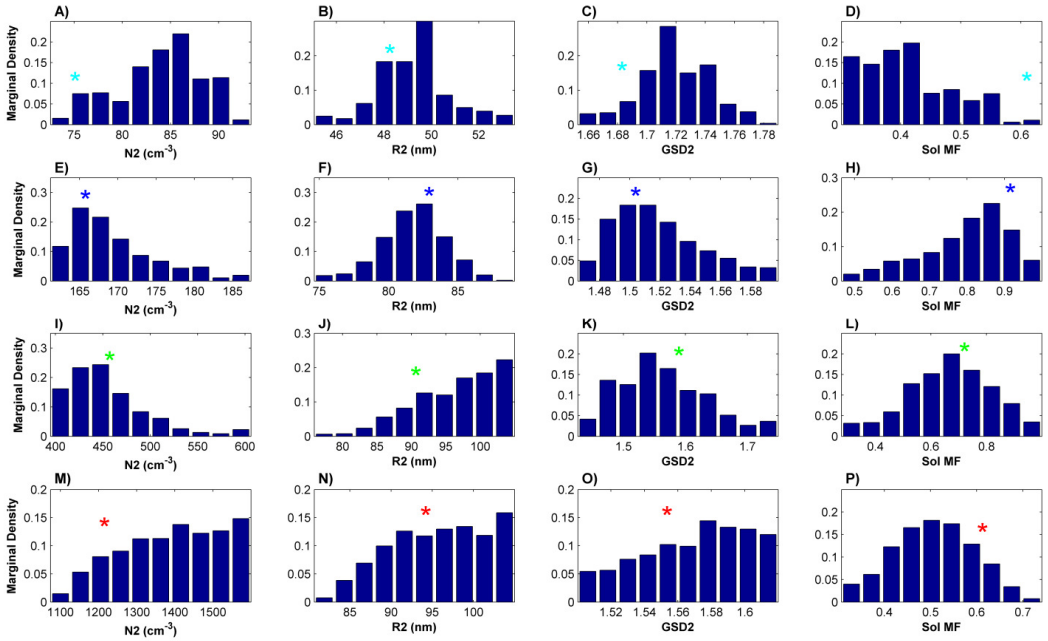

Fig. 8. Histograms of the marginal distributions of the four different adiabatic cloud parcel model parameters for **(A–D)** Marine Arctic, **(E–H)** Marine general, **(I–L)** Rural Continental, and **(M–P)** polluted continental conditions. The star in each subplot is used to separately indicate the true values of the cloud model parameters used to create the calibration data set.

ACPD
11, 20051–20105, 2011

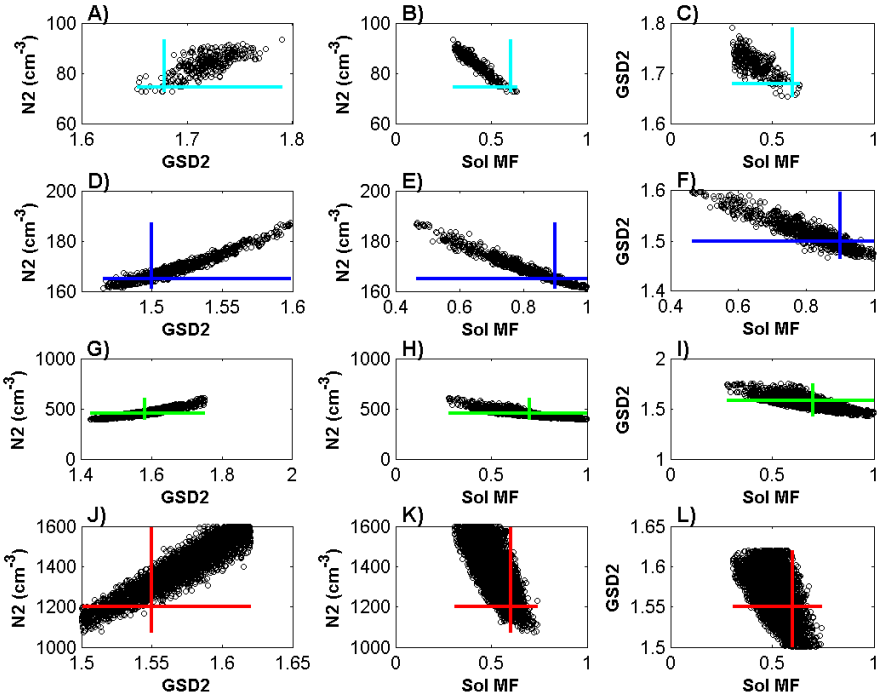

Fig. 9. Scatterplots of the posterior samples of the adiabatic cloud parcel model parameters for **(A–C)** Marine Arctic, **(D–F)** Marine general, **(G–I)** Rural Continental, and **(J–L)** polluted continental environments. The solid lines in each individual plot denote the posterior range of each individual parameter.

ACPD
11, 20051–20105, 2011

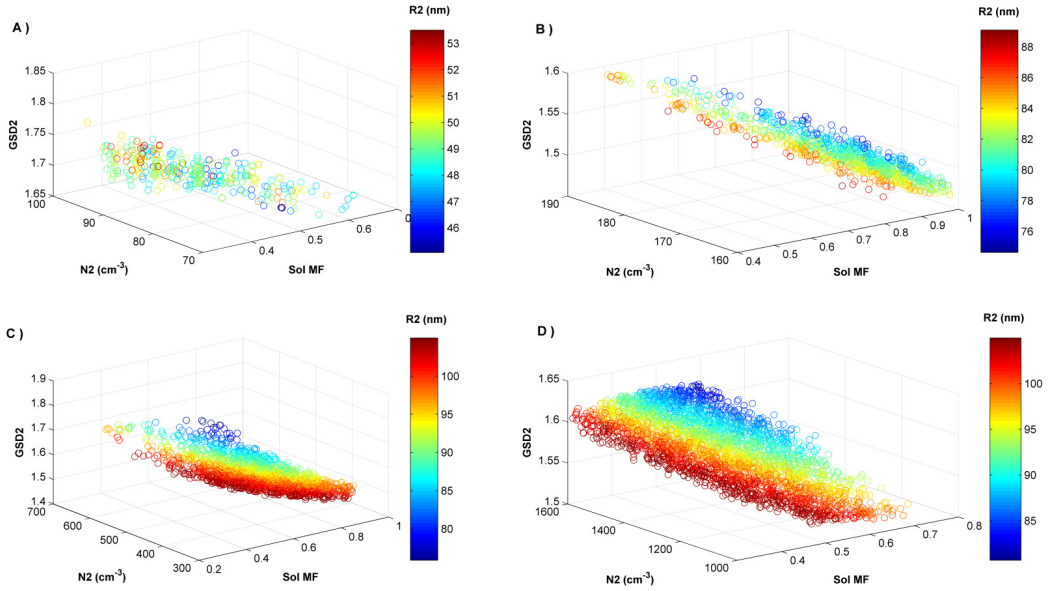

Fig. 10. Scatterplots of all four pseudo-adiabatic cloud parcel model parameters for **(a)** Marine Arctic, **(b)** Marine general, **(c)** Rural Continental, and **(d)** polluted continental environments.

ACPD
11, 20051–20105, 2011

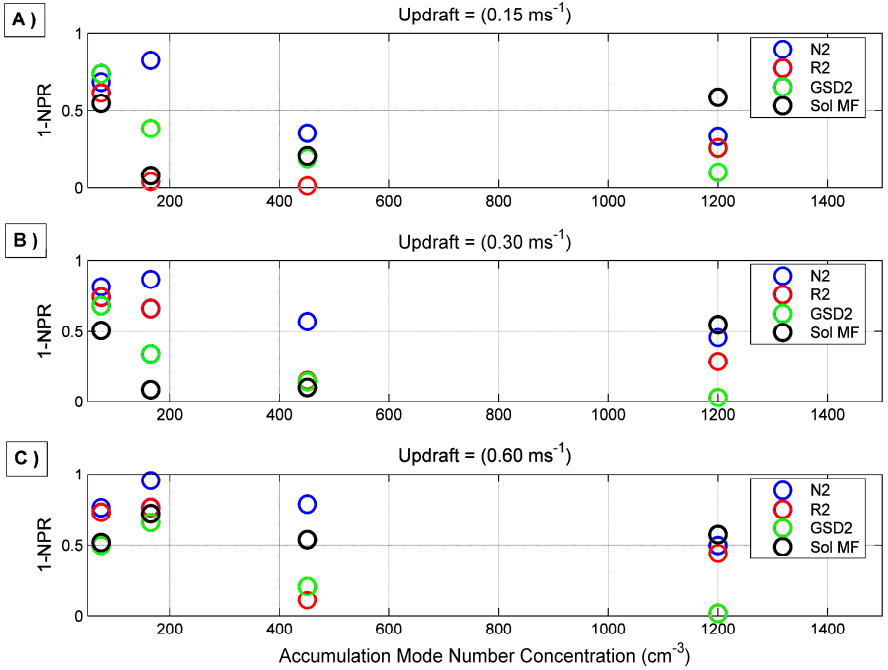

Fig. 11. Parameter relative sensitivity for **(a)** Updraft = 0.15 m s^{-1}, **(b)** updraft = 0.30 m s^{-1}, **(c)** updraft = 0.60 m s^{-1}. The last 20 % of the samples generated with DREAM were used to derive the results. The y-axes NPR labels correspond to "Normalized posterior parameter range". Thus, we present the relative sensitivity for each parameter as the aerosol environment becomes more polluted. A higher value of 1-NPR indicates a parameter having higher relative sensitivity. Going from left the light the x-axes corresponds to the accumulation mode number of marine Arctic, marine general, rural continental, and polluted continental conditions respectively.

ACPD
11, 20051–20105, 2011