Dear Veli-Matti Kerminen,

We would like to thank both yourself and the referees for the time you have spent considering our study and also for your patience with regard to the generous extension we received. Owing to the constructive feedback provided by the referees, we believe the revisions and additions have added a lot of value to the manuscript. Please find the author responses, list of manuscript changes and the annotated, revised manuscript contained below in this document.

Response to referees:

Below are the authors' point-by-point responses to each of the referees' comments. General comments are presented before major comments which precede technical corrections/minor comments. Our responses follow (bold italic text) the referee comments (standard text).

Anonymous referee 1:

General comments:

The authors present a new approach to assess the sensitivities of various parameters to CCN numbers. As opposed to numerous previous studies that used a 'one-at-a-time' approach, i.e. varying only one parameter at once, their use of response surfaces can reveal sensitivities over much wider parameter spaces. The focus on surface partitioning and surface tension and it is concluded that careful parameterization of these parameters is needed in order to successfully predict CCN under some conditions, in agreement with previous studies. The current study represents a model framework and the input data are artificially created so that perfect agreement can be achieved. So, therefore it is quite simplistic and does not lead to entirely new results. The sensitivity study for the selected parameters is more comprehensive and the approach might be promising to be applied to other problems in the future. However, not all parameters that have been identified as being important for determining the CCN numbers in previous studies have been discussed. Therefore, I think the paper's content and conclusions are somewhat weak and preliminary since only the suitability of the model framework, but not many new results are discussed.

The authors would like to take this opportunity to thank the reviewer for their constructive comments that we believe have added to the quality of the manuscript as a whole. The purpose of this study was to present a development of a new framework for probing sensitivity of aerosol activation to processes which have been studied in isolation over a number of years. The referee is absolutely right that there are a number of additional factors affecting CCN activation potential, but we focus on surfactant representations since not only is it a useful proof of concept demonstration of this framework, but there is still no clear message as to the wider relevance of surfactant behaviour. Using bulk to surface partitioning models, such as those presented here, previous studies have shown the impact of extreme behaviour, which one might consider as partitioning on or off, on the global scale, but suggested further studies would be better constrained by a wider parameter space such as provided here. In addition, as the referee points out, there are still studies promoting the potential for very low effective surface tension values, even if the concentration of surfactant material at activation might never be realised in the atmosphere. We feel demonstrating the use of a new multi-parameter sensitivity approach in helping to resolve the wider relevance of such issues is important, but requires incremental demonstrations of its use. We fully agree future studies need to tackle the issue of inter-instrument variability and process descriptions, but would also warrant much more data on systems for which we know the pure

component and mixture properties more accurately than ambient systems. We hope our response to the detailed points below make this clear, as do recommendations for future work.

Our responses are presented in bold text following the reviewer's comments. Any referencing of sections, pages or line numbers given in the response pertain to those of the revised manuscript.

Major comments:

1. Mixing state In several previous studies, it has been emphasized that the mixing state of aerosol particles might be one of the most important parameters that determines CCN number in fresh air masses. While it has been addressed briefly in the manuscript, it should be discussed more thoroughly. Could a measure of mixing state be included in the model framework?

This a good point, CCN activity may indeed be influenced by the mixing state close to aerosol sources, and as such additional text and references have been added to section 3.1. The framework developed here could in principle be used for a treatment of externally mixed aerosols, however, the choice of mechanistic aerosol-cloud model would be important. In theory, frameworks such as PartMC-MOSAIC (Tian et al 2014) could be coupled with our MCMC approach. Furthermore, the methodology developed here could be coupled with a with multi-modal cloud parcel model in the future. One of the virtues of the framework built in the present study is that many different aerosol processes and characteristics can be included in the future for more specific case studies.

2. Previous study on CCN sensitivities In a previous study, Lee et al. (2013) have performed a sensitivity study on a global scale of many parameters using a MonteCarlo-based approach. This study should be discussed in the light of the results in the current study.

The Lee et al. (2013) is indeed an interesting study. However, due to numerous differences between our study and that performed by Lee et al. we feel that the study is not directly of relevance.

Firstly, the studies have different goals. The focus of this study is to construct a framework, based on inverse modelling methods, for which model input parameters non-measurable at the scale of interest (e.g surface tension) can be calibrated against measurements of CCN spectra, and ultimately that parametric uncertainty can be evaluated and constrained using MCMC analysis (part2). In contrast, the study performed by Lee et al. (2013) is concerned only with sensitivity evaluation. In this part 1 study, by testing the validity of the inverse approach through response surfaces, we also get a visualisation of the model's parametric sensitivity in terms of Objective Function (OF) response surfaces.

It is well understood that application of inverse modelling methodologies, both for model calibration and parametric uncertainty analysis, can result in (near-)identical model predictions for non-unique parameter values if the system studied contains parameters that are non-identifiable for the definition of calibration data used. In such cases it can be challenging for the applied algorithm to converge on optimal parameter values in an attempt to constrain uncertainties. Typically, such difficulties are approached by increasing the information content in the calibration data (synthetic or real-world measurements) (Vrugt et al 2001). Therefore, we choose CCN spectra as calibration data in the present study to maximise the available information content with respect to currently known observations - as compared with the single N50 value used in Lee et al. (2013). Appropriate definition of calibration data and the importance of information content is now discussed at some length in section 5.3 of the revised manuscript as an appropriate definition of calibration data is essential for successful application of automatic search algorithms for parameter calibration and parametric uncertainty analysis. In section 5.3 we have explored the

implications of the information content contained within the calibration data for the identifiability of parameters investigated depending on the resolution of the calibration data, both with and without corruption of the synthetic measurements by a randomly generated natural variability. We conclude in the case of uncorrupted calibration data that a typical 5-band CCNC spectrum would contain sufficient information content for the presented methodology, but that it is unlikely that the use of a single value would facilitate parameter optimisation. When also accounting for natural variability, only the high-resolution calibration data, as used in the present study, would suffice for correct minimisation of the OF.

The dissimilarity in the definition of the calibration data between the present study and Lee et al (2013) is a natural consequence of another difference between the two studies – the scale. Our study is focused on developing a process model framework suitable for further uncertainty analysis in closure studies using algorithmic approaches, and we have referenced to similar studies accordingly Partridge et al. (2011, 2012). Such a study provides us with the opportunity to scrutinise the necessity for complexity, at a level of Köhler theory, offline whilst maintaining computational efficiency. Consideration of the impact of bulk atmospheric parameters (e.g. emission rates), as seen in Lee et al. (2013), does not fall under the remit of the present study and instead we choose to highlight the importance of Köhler model complexity in GCMs for the most accurate predictions of CCN.

To summarise, any sensitivity methodology relies heavily on the choice of calibration data. The focus of the present study is the construction of a framework for an inverse modelling approach to parametric uncertainty analysis and model calibration for entire CCN spectra at a process level rather than a single value on a global scale. Therefore, it is of our opinion that in some sense the present study is both more complex and robust in relation to potential model-observation evaluation and depth of process treatment. By performing this kind of analysis over a range of supersaturations the global climate modelling community can gain a lot from the end results.

Therefore, owing to the dissimilarities in methodology and purpose of the two studies, and that Monte-Carlo methods have not been employed here in part 1, direct comparison of results with respect to sensitivity is not possible and thus we have chosen to cite appropriately and not discuss Lee et al. (2013) at great length within the manuscript.

3. Uncertainty in cloud formation The authors state correctly that for data sets other than their calibration data set as used here, a perfect agreement cannot be expected. Some discussion on how accurately CCN numbers should be predicted should be added. For example, all measurements are associated with some measurement error. In addition, other factors influence CCN number (e.g. Lee et al. (2013). Given all uncertainties in the current abilities to predict cloud formation (meteorology, updrafts, emissions, etc), what is the recommendation for a tolerable uncertainty in CCN predictions?

The reviewer makes an excellent point regarding the treatment of errors and acceptable uncertainty. Following a similar request from reviewer 2, additional material has been added to the manuscript in section 5.3. Therein we have discussed at length the importance of information content for various definitions of the calibration that may arise due to different instrumentation. We also account for the natural variability in these considerations. It was concluded that, when accounting for natural variability in the analysis, the typical supersaturation resolution of CCN counters used in-situ would be challenging to correctly minimise the objective function as done when using the high-resolution spectra as seen in the present study. As natural variability in CCN measurements is typically on the order of tens of percent this effect will dominate over any instrumentation errors introduced, counting errors as deduced from Poisson statistics, for example. Thus we choose to focus on natural variability in this study to illustrate this point.

4. Solubility In previous studies, the solubility of organics has been identified as an influential parameter (Riipinen et al., 2015). However, this is not even mentioned in the current paper since all organics are assumed to be completely dissolved. How would consideration of a range of solubilities change the conclusions?

We agree that solubility may play a role, as will mixing state, gas to particle partitioning, condensed phase reactions and mass transfer limitations according to an amorphous condensed phase, depending on the RH history of the particle distribution in question. In this framework, the solubility would change the single particle hygroscopicity that would need to be constrained by an equivalent parameter space to the one presented here. This would need to be coupled to an appropriate gas phase mechanism since the solubility spectrum, such as that presented by Riipinen et al (2015), is driven by the specific abundance of compounds within different volatility ranges. We would argue then that any treatment of solubility should also be coupled to a model that can treat gas to particle partitioning since the most thermodynamically stable state of a framework that does not allow partitioning might never be met, affecting the derived parameter sensitivity. Such a model framework is beyond the scope of this study, but for sure warrants future investigation. Again, in theory this framework could be wrapped around any mechanistic or semi-empirical model, the results of this stepwise study demonstrating its use.

5. Surface-active fraction The discussion of previous studies on the effect of surface tension to CCN activation is somewhat confusing. For example, studies by Noziere et al are discussed that reveal surface tension of ~30 nM on aerosols. This number, however, is not relevant for CCN activation unless sufficient material is available to cover the complete particle/'droplet near activation. I suggest discussing the cited references more carefully. What fraction of surface-active material is needed to cause the effects as seen in the current study? Are these fractions realistic?

The reviewer has raised a very good point here. This issue is relevant for Köhler frameworks 1 (traditional Köhler theory) and 2 (redistribution of surfactant concentration and a concentration-independent fixed surface tension).

One of the key results of this paper is that the complete partitioning framework (4) produces CCN spectra and response surfaces similar to traditional Köhler theory as a consequence that surface tension is often very close to that of water at the point of activation and largely insensitive to the partitioning parameters used in the Szyszkowski equation. While this result is perhaps well known in the communities of the references in question, it is arguably less well known in the cloud physics community. Text relating to the references has been reformulated and revised, and the result has been discussed more carefully.

6. Figures Several figures appear blurry and hard to read due to very small font, e.g., Fig. 1 and 3: the indices in the legend are hard to distinguish Fig. 4: Avoid putting the legend box across the lines in the figure

We apologise for this – these figures have now been remedied.

Technical corrections/minor comments:

All minor comments have been addressed, as advised, in the revised manuscript. Thank you for taking the time to bring these to our attention.

References:

Tian, J., Riemer, N., West, M., Pfaffenberger, L., Schlager, H., and Petzold, A.: Modeling the evolution of aerosol particles in a ship plume using PartMC-MOSAIC, Atmos. Chem. Phys., 14, 5327-5347, doi:10.5194/acp-14-5327-2014, 2014.

Lee, L. A., Pringle, K. J., Reddington, C. L., Mann, G. W., Stier, P., Spracklen, D. V., Pierce, J. R., and Carslaw, K. S.: The magnitude and causes of uncertainty in global model simulations of cloud condensation nuclei, Atmos. Chem. Phys., 13, 17,8879- 8914, 10.5194/acp-13-8879-2013, 2013.

Vrugt, J.A., W. Bouten, and A.H. Weerts. 2001. Information Content of Data for Identifying Soil Hydraulic Parameters from Outflow Experiments. Soil Sci. Soc. Am. J. 65:19-27. doi:10.2136/sssaj2001.65119x

Partridge, D. G., Vrugt, J. A., Tunved, P., Ekman, A. M. L., Gorea, D., and Sorooshian, A.: Inverse modeling of cloud-aerosol interactions – Part 1: Detailed response surface analysis, Atmos. Chem. Phys., 11, 7269-7287, doi:10.5194/acp-11-7269-2011, 2011.

Partridge, D. G., Vrugt, J. A., Tunved, P., Ekman, A. M. L., Struthers, H., and Sorooshian, A.: Inverse modelling of cloud-aerosol interactions – Part 2: Sensitivity tests on liquid phase clouds using a Markov chain Monte Carlo based simulation approach, Atmos. Chem. Phys., 12, 2823-2847, doi:10.5194/acp-12-2823-2012, 2012.

Riipinen, I., Rastak, N., and Pandis, S. N.: Connecting the solubility and CCN activation of complex organic aerosols: a theoretical study using solubility distributions, Atmos. Chem. Phys., 15, 11,6305-6322, 10.5194/acp-15-6305-2015, 2015.

Anonymous referee 2:

General comments:

This study presents a methodology to investigate the sensitivity of CCN spectra to different parameters using different approaches. The results shown are not entirely new and future work is necessary in order to relate them to "real world" characterized by simultaneous measurements of aerosol chemical composition and size distribution, supersaturation, CCN spectra, etc.

The authors would like to take this opportunity to thank the reviewer for their constructive comments; we believe that they have added to the quality of the manuscript as a whole. The purpose of this study was to present the development of a new framework for probing sensitivity of aerosol activation to processes which have been studied in isolation over a number of years. By embracing an inverse modelling approach to aerosol-CCN closure, we not only build a framework for sensitivity analysis, but also a method of diagnosing both structural and parametric uncertainties in model CCN predictions by simultaneously matching input parameters and model output. In section 5.3 of the revised manuscript, additional material is presented to highlight the importance of calibration data resolution and natural variability in CCN measurements as a natural first step towards future work with observational data.Our response is presented in bold and italic text following the reviewer's comments. Any referencing of sections, pages or line numbers given in the response pertain to those of the revised manuscript.

Major comments:

1. The theoretical investigations shown here were performed at a fixed temperature, using literature data for three aerosol types: marine, polluted continental and rural. The organic aerosol and surface tension were assumed to be the same for all aerosol types. The possible effects of this data are not discussed in the manuscript.

Preliminary testing of temperature was performed and negligible sensitivity was found. We therefore chose to exclude it from our analysis. The reviewer raises a good point with regard to the

organic aerosol data. Unfortunately, as the organic aerosol fraction exhibits such high levels of complexity and spatial and temporal variability, in order to be more precise than we have been is difficult. By providing generous ranges for our model organic we hope to encompass all possibilities that may be realised in the real atmosphere. Indeed, once coupled with the MCMC algorithm (part 2) the framework developed in the present study provides a useful tool to constrain such parametric uncertainties by exploring all possibilities in the complete parameter space.

2. The manuscript looks like a report, most of the section 4.2 may be moved in a Supplementary Material. The results shown in it have to be presented in a more synthetic and comparative manner. It is not evident how much the sensitivity to parameters depends on the approach used in a quantitative manner. This is also not clear in Table 3 which is not actually discussed. Also, it has to be moved from Conclusions to Results.

We agree that the manuscript could be more concise, therefore the text has been substantially reduced and formulated in a more synthetic and concise manner. A supplementary document that contains all response surface analysis has now been attached with the manuscript. However, as inverse modelling of CCN spectra has not been performed before, we believe that the response surface analysis contained in section 5.2 (previously 4.2) is essential to the understanding of how to correctly couple the framework to automatic search algorithms, such as MCMC, which will form the focus of a part 2 study. Therefore we choose to keep the section in the main text.

3. Such a study may be very interesting if it succeeds in showing the limits/differences due to the different approaches used in modelling aerosol activation and quantify the acceptable/relevant uncertainties of measured parameters of fundamental importance in these approaches. The latter information may be important for planning future field campaigns and for development of instruments.

The reviewer raises a very good point here and we have chosen to add additional material to section 5.3 to address the issue at hand. Therein we have discussed at length the importance of information content for various definitions of the calibration that may arise due to different instrumentation. We also account for the natural variability in these considerations. It was concluded that, when accounting for natural variability in the analysis, it would be challenging to correctly minimise the OF based on the information content of a calibration data set measured by a typical CCN counter. However, correct minimisation of the OF was still achieved when corrupting the high resolution calibration data set with randomly generated natural variability. As such, this result should serve as a recommendation for the development of instrumentation for high resolution measurements of CCN spectra in-situ.

Technical corrections/minor comments:

The term “posedness” is not commonly used, replace or explain it better

The term posedness has been replaced throughout the manuscript with ‘well-posed’, ‘ill-posed’ or similar.

Changes to manuscript:

In order to increase the quality of the manuscript and satisfy the referees considerable revisions and additions have been made to the manuscript. In particular, a lot of text has been rephrased and some methodology sections retitled and reordered to improve the readability of the study. In most sections content has remained unchanged, however, there have been considerable additions to section 5.3 which we feel have added a new perspective to the paper. As such, results contained therein have also been carried over to the abstract and conclusions. Below are a list of changes made to the figures and text, and in the marked up manuscript that follows we have highlighted new text and text that has been reworded considerably in yellow. References that have been added during the revision period, some at the request of the referees and others independently, have been highlighted in green. We hope that the changes the manuscript has undergone are clear from this.

Figures:

1. All figures have been updated for aesthetic purposes and as a minor bug was found in the setup of the size distributions. Fortunately, the bug had no impact on the results or conclusions reached by the study.

2. The environment presented in figure 5 (previously 4) has been changed from rural continental to marine average, and mode size has also been included (orange), to link with new findings in section 5.3 (previously 4.3).

3. Two additional figures have been added to section 5.3 (previously 4.3), figures 12 and 13.

Text:

1. New results reached in section 5.3 have been outlined in the first paragraph of the abstract. Results detailed previously in the abstract have been made more concise.

2. Contents of the introduction have remained unchanged. However, to improve the readability of the introduction, the ordering has changed substantially.

3. The goals of the study (section 1.1) have been rephrased

4. A considerable amount of text has been added to section 5.3 (previously 4.3) and pre-existing text reformulated.

5. Conclusions pertaining to results in section 5.3 have now have been added to the conclusions.

Inverse modelling of Köhler theory - Part 1: A response surface analysis of CCN spectra with respect to surface-active organic species

Samuel Lowe[1, 2, 3], Daniel Partridge[3, 1, 2], David Topping[4, 5], and Philip Stier[3]

[1]Department of Environmental Science and Analytical Chemistry, Stockholm University, Stockholm, Sweden
[2]Bert Bolin Centre for Climate Research, Stockholm University, Stockholm, Sweden
[3]Atmospheric, Oceanic and Planetary Physics, Department of Physics, University of Oxford, Oxford, UK
[4]School of Earth Atmospheric and Environmental Science, University of Manchester, Manchester, UK
[5]National Centre for Atmospheric Science (NCAS), University of Manchester, Manchester, UK

Correspondence to: dan.partridge@aces.su.se

Abstract.

In this study a novel framework for inverse modelling of CCN spectra is developed using Köhler theory. The framework is established by using model-generated synthetic measurements as calibration data so as to facilitate a parametric sensitivity analysis. Assessment of the relative importance of aerosol physicochemical parameters while accounting for bulk-surface partitioning of surface-active organic species is carried out over a range of atmospherically relevant supersaturations. By introducing an Objective Function (OF) that provides a scalar diagnostic metric for deviation of modelled CCN concentrations from synthetic observations, OF response surfaces are presented as a function of input parameters. Crucially, for the chosen calibration data, aerosol-CCN spectrum closure is confirmed as a well-posed inverse modelling exercise for a subset of the parameters explored herein. The response surface analysis indicates that the appointment of appropriate calibration data is particularly important. To achieve inverse aerosol-CCN closure and constrain parametric uncertainties, it is shown that a high resolution CCN spectrum definition of the calibration data is required where single values may be expected to fail.

Using Köhler theory to model CCN concentrations requires knowledge of many physicochemical parameters some of which are difficult to measure in-situ at the scale of interest and introduce a considerable amount of parametric uncertainty to model predictions. For all partitioning schemes and environments modelled, model output showed significant sensitivity to perturbations in accumulation mode parameters, surface tension, organic-inorganic mass ratio, insoluble fraction and solution ideality. Many response surfaces pertaining to these parameters contain well-defined minima and are therefore good candidates for calibration using a Monte Carlo Markov Chain (MCMC) approach to constraining parametric uncertainties.

A complete treatment of bulk-surface partitioning is shown to model CCN spectra closely to those calculated using classical Köhler theory with the surface tension of a pure water drop, as found in previous studies. In addition, model sensitivity to perturbations in the partitioning parameters was found to be negligible. As a result, this study supports previously held recommendations that complex surfactant effects might be neglected and continued use of classical Köhler theory in global climate models (GCMs) is recommended to avoid additional computational burden. The framework developed is suitable for application to many additional composition-dependent processes that might impact CCN activation potential. However, the focus of this study is to demonstrate the efficacy of the applied sensitivity analysis to identify important parameters in those processes and will be extended to facilitate a global sensitivity analysis and inverse aerosol-CCN closure.

1 Introduction

Atmospheric aerosols have an influence on the earth's radiation balance, and thus the climate and its evolution, through many feedback effects and processes. Aerosols can act to absorb and scatter solar radiation, the direct effect (McCormick and Ludwig, 1967). In addition, aerosols larger than some critical size, referred to as the activation size, may also act as Cloud Condensation Nuclei (CCN) - viable sites for condensational growth of cloud droplets at a given supersaturation. Changes in CCN concentrations have the potential to have a influence on both cloud micro- and macro-physics, and consequently global radiative forcing (IPCC, 2013). At fixed liquid water path, an increase in aerosol concentration serves to increase CCN and cloud droplet number concentrations (CDNC), thus reducing average droplet size and increasing cloud albedo which is known as the first (Twomey) indirect effect (Twomey, 1974). Consequently, the reduced average effective droplet radius restricts the formation of droplets large enough to precipitate and is hypothesised to increase cloud lifetime, known as the second (Albrecht) indirect effect (Albrecht, 1989).

Cloud-aerosol interactions represent the largest uncertainty in current global radiative forcing estimates (IPCC, 2013). To constrain such uncertainties it is necessary to improve our understanding of the dependence of CCN concentrations on the physicochemical properties of aerosols. With an increased understanding of the cloud nucleating potential of aerosol particles, more accurate aerosol representations and droplet activation parametrizations can be implemented in global climate models (GCM) (Abdul-Razzak et al., 1998; Fountoukis and Nenes, 2005; Quinn et al., 2008). The likelihood of a given aerosol to act as a CCN is a highly non-linear function of many parameters, including the size and chemistry of the aerosol as well as the prevailing meteorological conditions (McFiggans et al., 2006). Accurate computation of CCN concentrations with Köhler theory is a problem of high dimensionality that is not currently possible online in GCMs due to computational limitations. It is therefore of critical importance that the community ascertains which parameters and processes CCN concentrations are most sensitive to. In doing so, processes that are unimportant for determining CCN concentrations can be removed to reduce computational burden and parameters currently fixed can be held at correct values.

The importance of some physical and chemical properties are certainly expected to be greater than others (Nenes et al., 2002; Quinn et al., 2008) and at a given supersaturation, a considerable spatial variability in CCN concentrations is observed in nature (Yum and Hudson, 2002). It may therefore be suspected that the relative importance of various physicochemical parameters may be environmentally dependent as both the aerosol size distribution and kinetic description of condensational growth are important in determining CCN concentrations (Lance et al., 2004). For that reason, it is important to consider multiple environments when carrying out sensitivity analyses such as the one that is conducted in the present study. As such, input parameters used in the present study are chosen to reflect aerosol populations of marine average, rural continental and polluted continental environments.

The cloud nucleating potential of aerosols is typically modelled using Köhler theory (Köhler, 1936) in which the equilibrium saturation vapour pressure ratio at the particle surface is related to the wet particle size. An aerosol is deemed to be a CCN if the peak of the growth curve, the critical supersaturation, is lower than the atmospheric saturation, thus allowing unstable growth in the presence of sufficient water vapour. The original formulation was derived to describe the growth of a binary mixture of an inorganic salt with condensing water vapour but has since been expanded upon for numerous applications to account for increasing levels of complexity to better represent real-world aerosol systems. These additions include, but are not limited to: multicomponent aerosols with concentration-dependent organic acid solubility and surface tension (Shulman et al., 1996); addition of hygroscopic material via condensational growth from trace gases (Laaksonen et al., 1997); theoretical derivation of an analytical solution for the point of activation in the presence of an insoluble core (Kokkola et al., 2008); inclusion of the bulk to surface partitioning of surface-active organics (Sorjamaa et al., 2004; Topping, 2010); and the co-condensation of Semi-Volatile Oxygenated Organic Aerosol (SV-OOA) material (Topping et al., 2013).

Over the last couple of decades the importance of complex organic aerosols in determining the activation point has been widely acknowledged (Lohmann et al., 2000; Jacobson et al., 2000; Chung and Seinfeld, 2002; Kanakidou et al., 2005). The organic fraction consists of thousands of different carbonaceous compounds of varying chemical and physical properties (Saxena and Hildemann, 1996) and constitutes 20-90% of atmospheric aerosol mass depending on the environment (Saxena and Hildemann, 1996; Jacobson et al., 2000; Putaud et al., 2004; Kanakidou et al., 2005; Zhang et al., 2007; Jimenez et al., 2009). In addition, atmospheric loadings of both primary and secondary organic aerosol have changed since the pre-industrial period (Tsigaridis et al., 2006), and thus the associated parameter ranges should be explored to investigate the influence this has had on our climate. Mircea et al. (2002) found that the presence of a water soluble organic carbon (WSOC) fraction could increase the number of CCN available in polluted regions by as much as 110%. Given this high degree of spatial variability, large range in observed organic mass fractions and its chemical complexity, an importance must not only be placed the sensitivity of modelled CCN concentrations to organic aerosol physicochemical parameters, but also on how that sensitivity interacts with size distribution parameters. This calls for a more robust sensitivity analysis than typical individual parameter analyses, particularly when considering complex organics that require additional parameters that increase the dimensionality of the problem. In particular, sensitivity and modelling studies have reported that the presence of slightly soluble and surface-active organic species can alter the point of activation for atmospheric aerosols (Shulman et al., 1996; Li et al., 1998; Sorjamaa et al., 2004; Henning et al., 2005; Topping, 2010; Topping and McFiggans, 2012). Ekström et al. (2010) concluded that bio-surfactants have the capacity to possess a greater cloud-nucleating ability than even inorganic salts on account of measured surface tension values below 30 mNm^{-1}. In addition, humic like substances (HULIS) have been acknowledged as surface-active (Li et al., 1998; Facchini et al., 1999, 2000); however, more recently it has been accepted that the concentration gradient manifested by the bulk to surface partitioning of surfactants must also be accounted for when computing the point of activation (Sorjamaa et al., 2004; Topping, 2010).

In reality, the transition between aerosol gas and liquid phases is not stepwise, i.e. the density profile is continuous rather than step-like. To calculate the influence of bulk-surface partitioning organics, Sorjamaa et al. (2004) modified traditional Köhler theory to recalculate equilibrium curves in terms of bulk and surface quantities for binary and ternary mixtures. From their results, they deduced that surfactants may enhance growth of large droplets in the atmosphere thus decreasing cloud density. Topping (2010) derived an alternative theoretical basis that is able to model the effects of an unlimited number of surface active species and concluded that in order to have a comprehensive understanding of this phenomenon, model predictions must be verified with CCN observations. Prisle et al. (2012) investigated the implications of bulk-surface partitioning for cloud droplet activation on a global scale using the ECHAM5.5-HAM2 global circulation model and recommended that an approach considering surface tension effects alone, and neglecting changes in bulk properties, is erroneous and should not be used. A full treatment of bulk-surface partitioning was found to predict similar global CDNC results to a treatment neglecting bulk-surface partitioning altogether. Nevertheless, the topic still requires observational verification on all scales, and a global sensitivity analysis (GSA) to probe the entirety of the relevant multi-dimensional parameter space, in order to confidently arrive at correct conclusions. In this study the effects of bulk-surface partitioning on CCN concentrations are analysed using the Köhler model developed by Topping (2010).

One method of evaluating predictions made by Köhler theory is to carry out an aerosol-CCN closure study. Closure is achieved when predicted CCN concentrations are within the uncertainty bounds of observations typically collected from CCN counters (CCNC) at a given supersaturation. Numerous aerosol-CCN closure studies have been performed to varying degrees of success (Bigg, 1986; Cantrell et al., 2001; Zhou et al., 2001; Broekhuizen et al., 2006; Bougiatioti et al., 2009; Martin et al., 2011). Broekhuizen et al. (2006) found that aerosol-CCN closure is often difficult to achieve and that such difficulty can be attributed to various sources of error including measurement biases or spatial and temporal variability during airborne measurements. They also indicate that studies unable to achieve closure were often those in which organic carbon (OC) was more prevalent in the particle phase and the airmass of interest was close to sources of anthropogenic aerosol. Furthermore, Facchini et al. (2000) suggested that the inability to achieve closure could be attributed to enhanced CCN activity due to accumulation of atmospheric polycarboxylic acids at the particle surface, thus depressing the surface tension as their molecular structure resembles that of HULIS.

Numerous studies have been conducted to examine the sensitivity of both the activation size and CCN concentrations with respect to relevant physicochemical parameters of the aerosol population (Fitzgerald, 1973; Roberts et al., 2002; Wex et al., 2008; Ervens et al., 2010). Such studies are instructive but are often restricted to perturbations in a single parameter, a one at a time analysis (OAT), also known as a local sensitivity analysis (LSA), thus failing to probe the entirety of the full multi-dimensional parameter space. Furthermore, aerosol size distributions often possess steep gradients so a slight change in activation diameter can have a significant impact on CCN concentrations. Sensitivity studies conducted solely on the point of activation are therefore insufficient; to confidently arrive at robust conclusions, a sensitivity analysis of CCN concentrations across all relevant supersaturations - a CCN spectrum - is preferred. One way to improve on existing sensitivity studies is to embrace an inverse modelling methodology. The benefit of such a framework is twofold. Firstly, the inverse modelling framework facilitates a GSA that is able to probe the entire multi-dimensional parameter space. A GSA captures any parameter interactions that can affect sensitivity estimates (Quinn et al., 2008; Partridge et al., 2011, 2012). Secondly, by introducing an Objective Function (OF), a sensitivity analysis be carried out across all atmospherically relevant supersaturations simultaneously.

One of the first applications of inverse modelling to assess the effects of parametric uncertainty in aerosol-cloud interactions was undertaken by Partridge et al. (2012). They carried out Markov Chain Monte Carlo (MCMC) simulations for inference of the posterior parameter distribution in a Bayesian framework. In this study, an inverse modelling framework is developed for the analysis of CCN spectra to facilitate further study using MCMC methods.

An inverse modelling framework not only enables the conditioning of parameter sensitivities on measurements, but also provides a method of diagnosing structural inaccuracies within models and simultaneous matching of model input and output. Such inaccuracies present themselves as statistically significant discrepancies between optimised parameter values and their corresponding real-world observed values. In addition, the technique also provides a method of parameter estimation for parameters which are not easily measured in-situ at the scale of interest, surface tension for example. These advantages have led to the use of inverse modelling as a method of model calibration across a broad range of research subjects (Vrugt et al., 2004; Tomassini et al., 2007; Garg and Chaubey, 2010; Partridge et al., 2012; Wikle et al., 2013).

In this study, to the best of the authors' knowledge, an inverse modelling framework for CCN spectra is developed for the first time. To diagnose the sensitivity of an entire CCN spectrum to parameter perturbations in a tangible way, an OF is introduced. The OF provides a scalar metric by which the sensitivity of CCN spectra can be quantified with respect to both individual and multiple parameter perturbations.

Before performing a GSA and parameter optimisation procedure using an automated search algorithm, it is deemed judicious to first confirm that the study is a well-posed (Pollacco and Angulo-Jaramilo, 2009; Cressie et al., 2009). That is to say that the information content of the measurement data, further referred to as calibration data, is sufficient to constrain input parameters and thus can be deemed 'identifiable' by minimisation of the OF. In this study, to calibrate the sensitivity analysis, literature-obtained best estimate parameter values are used to generate a synthetic calibration data set from the model. Should parameters be non-identifiable it may certainly be expected that algorithms employed for model calibration and GSA may fail to converge. To confirm that inverse of modelling CCN spectra is suitable for the application application of automatic search algorithms, response surfaces of the OF are invoked in this study as done by Toorman et al. (1992); Šimůnek et al. (1998); Vrugt et al. (2001); Partridge et al. (2011). Response surfaces are a graphical tool that enable the investigation of the identifiability of parameters when considering susceptibility of CCN spectra to perturbations in 2D planar subsets of the entire parameter space. In particular, surfaces containing single well-defined minimum are preferred as the gradient of minimisation points to the same point regardless of where in the parameter domain the algorithm is, thus efficient convergence can be expected.

1.1 Goals

The primary goal of this study is to build a framework for inverse modelling of CCN spectra using Köhler theory and to test the suitability of automatic search algorithms as a tool for model calibration and GSA. In constructing the framework, qualitative sensitivity information is presented in the form of OF response surfaces for simultaneous perturbations in two parameters. In addition to considerations of environmental dependent parameter sensitivities, the role of surface-active organic compounds is also explored. The specific questions to be investigated in this study are the following:

1. Is it possible to simultaneously match CCN spectra and correctly calibrate input parameters using an inverse modelling methodology?
2. Is inverse modelling of CCN spectra for a GSA using an MCMC algorithm feasible?
3. Qualitatively, how susceptible are CCN concentrations, across a range of atmospherically relevant supersaturations, to simultaneous perturbations in aerosol size distribution and physicochemical parameters?
4. Does the bulk-surface partitioning of surface-active organics play an important role in CCN activity over an atmospherically relevant range of supersaturations, and how sensitive are the associated parameters?

2 Theoretical basis and materials

2.1 Multicomponent Köhler theory

The Köhler equation describes the equilibrium saturation vapour pressure ratio s_{eq} of a condensible vapour at the surface of a wetted particle radius r_p,

$$s_{eq} = a_w exp\Big(\frac{2M_w\sigma}{RT\rho_w r_p}\Big) \tag{1}$$

where M_w is the molecular weight of water, σ is the surface tension of the wetted particle, R is the universal gas constant, T is temperature and r_p is the particle radius (Köhler, 1936; Seinfeld and Pandis, 2012). The supersaturation S_{eq}, as a percentage, is given by $S_{eq} = (s_{eq} - 1) \times 100\%$. The peak of the Köhler curve, the critical supersaturation S_c, defines the ambient supersaturation required for CCN activation.

The water activity term a_w in Eq. (1) can be written in terms of an effective mole fraction x_w^{eff},

$$a_w = x_w^{eff} = \frac{n_w}{n_w + n_s^{eff}} \tag{2}$$

where n_w is the number of moles of water, and n_s^{eff} is the effective number of moles of solute. n_s^{eff} can be calculated from the internally mixed Water Soluble Organic Carbon (WSOC) and inorganic components, n_i and n_j respectively, and their van't Hoff factors i_i and i_j,

$$n_s^{eff} = \sum_{i=1}^{p} i_i n_i \chi_i + \sum_{j=1}^{q} i_j n_j \tag{3}$$

where the indices i and j span the number of organic (p) and inorganic (q) species and χ_i is the effective soluble fraction of the organic species (Shulman et al., 1996; Sorjamaa et al., 2004). In this study, organics are assumed completely soluble, $\chi_i = 1$. Theoretically, the mathematical framework can treat a multi-component organic fraction, here however we choose to study just one, drop the summation and replace index i with org for clarity. Assuming a dilute solution, van't Hoff factors can be approximated by stoichiometric dissociation factors, ν_{org} and ν_j, and the solution's osmotic coefficient Φ (Kreidenweis et al., 2005). n_s can therefore be reformulated as

$$n_s^{eff} = \Phi\left[\nu_{org}n_{org} + \sum_{j=1}^{q}\nu_j n_j\right] \tag{4}$$

For an aerosol of dry radius r_d and insoluble fraction f_{insol}, the total number of moles of soluble substance can be re-expressed in terms of the organic fraction f, individual inorganic component sub-fractions ϵ_j and each components molecular weight M_{org} and M_j and density ρ_{org} and ρ_j

$$n_s^{eff} = \frac{4}{3}\Phi\pi r_d^3(1-f_{insol})\left[f\frac{\nu_{org}\rho_{org}}{M_{org}} + (1-f)\sum_{j=1}^{q}\frac{\epsilon_j\nu_j\rho_j}{M_j}\right] \tag{5}$$

As a final adjustment f can be expressed in terms of the organic to inorganic ratio α

$$n_s^{eff} = \frac{4}{3}\Phi\pi r_d^3(1-f_{insol})\left[\frac{\alpha}{1+\alpha}\frac{\nu_{org}\rho_{org}}{M_{org}} + \frac{1}{1+\alpha}\sum_{j=1}^{q}\frac{\epsilon_j\nu_j\rho_j}{M_j}\right] \tag{6}$$

The description of n_s given in Eq. (6) is used for this study.

2.2 Bulk-surface partitioning

A brief overview of the theory behind the bulk-surface partitioning Köhler model developed by Topping (2010) and used here is given in this section. The reader is referred to Topping (2010) for a more detailed description.

The interface between bulk liquid and gas phases is not infinitely thin as Gibbs' surface thermodynamics would suggest (Sorjamaa et al., 2004); this surface phase of finite thickness is the region in which surface active organic compounds accumulate. In order to calculate the influence of the increased surface excess on the Kelvin (surface tension) and water activity terms simultaneously, a surface excess correction to the total molar quantity is required

$$n_{org}^s = n_{org}^t - n_{org}^b \tag{7}$$

where n_{org}^s, n_{org}^t and n_{org}^b are the surface excess, total and bulk molar quantities respectively. To separate the total quantity into bulk and surface quantities a solution of the Gibbs adsorption isotherm is required

$$n_{org}^s d\mu_{org} + Ad\sigma = 0 \tag{8}$$

where A is the droplet surface area , μ_{org} the chemical potential of the organic and σ is the surface tension. The semi-empirical form of the Szyszkowski equation (Szyszkowksi, 1908) derived by Li and Lu (2001) is used

$$\sigma = \sigma_w - RT\Gamma^{wo}ln(1+Ka_{org}) \tag{9}$$

where Γ^{wo}, K and a_{org} are the saturated surface excess, adsorption equilibrium constant and activity of the organic compound, respectively, and σ_w is the surface tension of pure water. Γ^{wo} is defined to be the molar excess of the surfactant in a unit surface area of the surface region over that in the bulk liquid region assuming the same number of moles of water in the two regions (Li and Lu, 2001). The superscript wo refers to the position of the dividing interface being defined such that $n^s_w = 0$ or, equivalently, $\Gamma_w = 0$, and is dropped for notational convenience.

Solving Eq. (9) and (8), and assuming all activities can be represented by their effective bulk mole fractions, Eq. (2), the bulk mole fraction of the organic compound can be calculated from the roots of the quadratic equation,

$$(x^b_{org})^2(A\Gamma K - n^t_w K - n^t_{org}K) + x^b_{org}(n^t_{org}K - n^t_{org} - n^t_w - A\Gamma K) + n^t_{org} = 0 \tag{10}$$

the solution of which is found using the standard quadratic formula and taking the negative root such that $0 < x^b_{org} < 1$ for a physical solution. Here the eff superscript has also been dropped for notational convenience and all subsequent references to such variables are to the effective values unless stated otherwise. Assuming the surface excess of water to be zero, the number of moles of surfactant in the bulk can be calculated as (Topping, 2010)

$$n^b_{org} = n^t_w \frac{x^b_{org}}{1 - x^b_{org}} \tag{11}$$

following this the water and surfactant activities can be calculated in terms of the bulk mole fraction of the surfactant

$$a_w = \frac{n_w}{n_w + n_{inorg} + n_w\left(\frac{x^b_{org}}{1-x^b_{org}}\right)} \tag{12}$$

$$a_{org} = \frac{n_w\left(\frac{x^b_{org}}{1-x^b_{org}}\right)}{n_w + n_{inorg} + n_w\left(\frac{x^b_{org}}{1-x^b_{org}}\right)} \tag{13}$$

Equations (12) and (13) can be substituted into Eq. (1) and (9) to account for partitioning of the surfactant in Köhler growth curve.

With the partitioning described by Eq. (9), (12) and (13), there are four possible partitioning schemes for consideration. Application of the surface tension model, indicated by σ^{nf}, allows surface tension to be modelled as a function of organic activity and the empirically derived partitioning parameters, K and Γ. This is in contrast to using a fixed value indicated by σ^f. Superscripts nf and f indicate that the surface tension is not fixed and fixed, respectively. Accounting for the partitioning of the surfactant concentration to the surface phase is indicated by a^p_w while assuming that the concentration remains solely in the bulk phase is indicated by a^{np}_w, where superscripts p and np indicate partitioning and no partitioning, respectively. The resulting four schemes are:

1. $a^{np}_w \sigma^f$

2. $a_w^p \sigma^f$
3. $a_w^{np} \sigma^{nf}$
4. $a_w^p \sigma^{nf}$

Here 1. and 4. refer to simple Köhler theory and a complete treatment of bulk-surface partitioning, respectively. 2. and 3. refer to schemes accounting for the partitioning of surfactant concentration to the surface phase and surface tension depression (eq.9) independently, respectively.

Modelling of CCN with Köhler theory involves many currently uncertain parameters, especially with respect to the organic aerosol fraction. In this study, the Köhler parameters probed in the sensitivity analysis are M_{org}, ρ_{org}, Φ, σ, Γ and K and compositional parameters α and f_{insol}. In addition, the log-normal parameters of the second (accumulation) mode, N_2, σ_2 and $\bar{r}_2$, number concentration, geometric standard deviation and mean radius respectively, will also be probed as, in contrast to the first (Aitken) mode, a large fraction of accumulation mode particles are expected to be CCN active based on typical activation diameters. Thus the maximum dimensionality of the sensitivity analysis is 11, depending on the partitioning scheme used. In addition to the probed parameters, the following parameters are held fixed: $T = 285$K, $\nu_{org} = 1$, $\chi_{org} = 1$, $\nu_{\mathrm{NaCl}} = 2$, $\nu_{(\mathrm{NH_4})_2\mathrm{SO_4}} = 3$ and $\nu_{\mathrm{NH_4NO_3}} = 2$.

To illustrate the impact of the different partitioning schemes on the CCN activation point, Fig. 1 shows an example of critical supersaturation as a function of aerosol dry size for a mixture of NaCl and Suwannee River Fulvic Acid (SRFA) mass ratio of $\alpha = 2$. SRFA has been prescribed values for its molecular mass, density and surface tension in solution of 610 gmol^{-1}, 1570 kgm^{-3} and 55 mNm^{-1}, respectively, based on measurements and results obtained by Dinar et al. (2006) and Taraniuk et al. (2007). The activation point predicted by classical Köhler theory $a_w^{np} \sigma^f$ (blue) using the surface tension of a pure water droplet is well replicated by the full partitioning scheme, $a_w^p \sigma^{nf}$ (magenta), for these particular parameters. The point of activation by classical Köhler theory $a_w^{np} \sigma^f$ using a fixed surface tension of 55 mNm^{-1} (black) is well replicated by a partitioning scheme that accounts for a depleted bulk concentration of SRFA using the same fixed surface tension value $a_w^p \sigma^f$ (green). The partitioning scheme treating surface tension as concentration dependent Eq. (9), $a_w^{np} \sigma^{nf}$ (red) shows a more complex relationship for the activation points. For smaller sized dry particles it more closely matches schemes using a depressed fixed surface tension value. At larger sizes however, it approaches the classical Köhler scheme using fixed surface tension of water; this regime change is attributed to a decreased surface:volume ratio for larger particles, thus reducing the influence of surface phenomena. In the next section coupling of the above models (1-4) to a size distribution, for the computation of CCN spectra, and the model input parameters are discussed.

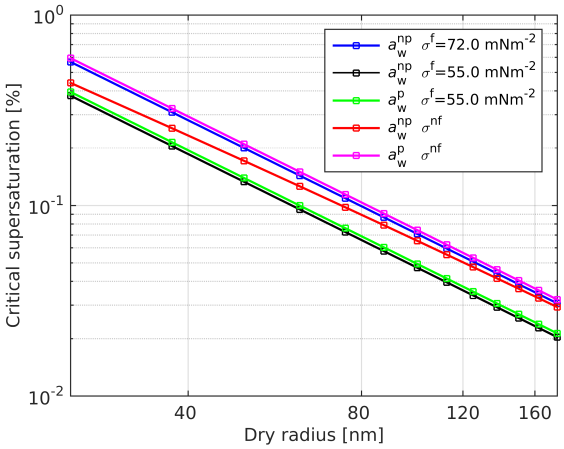

Figure 1. Critical supersaturation S_c as a function of dry radius r_d for all partitioning schemes. a_w^p and a_w^{np} labels indicate whether the partitioning effects are or are not accounted for in the water activity term a_w respectively; σ^f and σ^{nf} indicate whether the surface tension is prescribed a fixed value or modelled using equation 9. Particles are internal mixtures of NaCl and SRFA with organic to inorganic ratio $\alpha = 2$. Complete solubility and ideality are assumed. The partitioning parameters for SRFA are Γ=0.0025 and K=35942.03.

3 Model input and applicat

3.1 Physicochemical input parameters

To predict CCN spectra, Köhler theory must be coupled with an aerosol size distribution. Aerosol size distributions are well represented by a superposition of log-normal distributions (Seinfeld and Pandis, 2012)

$$n(r) = \sum_{i=1}^{m} \frac{N_i}{\sqrt{2\pi}\log\sigma_i} exp\left[-\frac{(\log r - \log\bar{r}_i)^2}{2\log^2\sigma_i}\right] \qquad (14)$$

where N_i, $\bar{r}_i$ and σ_i are the log-normal parameters for the i-th mode, number concentration, mean radius and standard deviation, respectively. In this study, only bi-modal distributions are studied, $m = 2$.

In this study, in order to analyse parameter sensitivity with respect to environmental aerosol characteristics, three distinct size distributions are taken from existing literature:

1. Marine average: Average global marine measurements from Heintzenberg et al. (2000).
2. Polluted continental: summertime air mass measurement from the Melpitz station, Germany (Birmili et al., 2001).
3. Rural continental: SMEAR II station, Hyytiälä, Finland (Tunved et al., 2005).

Average distribution parameters, and their uncertainty ranges, used for input in Eq. (14) are taken from the above references and are included in table 2. Distributions are generated over 400 logarithmically spaced dry radius bins ranging from 1nm to 1μm and are shown in Fig. 2.

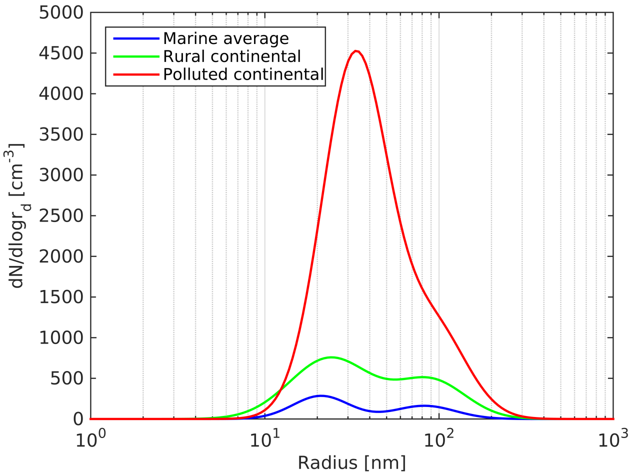

Figure 2. The marine average (blue), rural continental (green) and polluted continental (red) dry size distributions calculated using the true log-normal parameters given in table 2.

The mixing state of aerosol particles can play an important role in CCN activation and their optical properties, particularly close to sources of fresh emissions (Lesins et al., 2002; Broekhuizen et al., 2006). Here a size-independent, internally-mixed aerosol composition is prescribed for each environment for simplicity; however, the framework is suitable for application to externally mixed aerosol systems with size-dependent chemistry in future studies. The prescribed composition consists of inorganic salts, a model organic (MO) surfactant and insoluble black carbon. The composition of the average marine environment is taken as an approximate average of measurements recorded during periods of high and low biological activity at the Mace Head atmospheric research station (O'Dowd et al., 2004). Polluted continental composition is taken from the Melpitz station, Germany (Poulain et al., 2011). Rural continental from non-volatile mass measurements recorded in Hyytiälä, Finland (Häkkinen et al., 2012). The relative mass contributions are included in table 1.

The chemical properties of the MO (table 1) used in this study are based on averages calculated from organic acids documented in Topping (2010). As stated in section 2.1, the surfactant is assumed to be completely soluble. While the solubility of organic compounds has been identified to potentially influence CCN activation in some cases, the inclusion of complex conceptual frameworks such as that developed by Riipinen et al. (2015) is beyond the scope of the present study. However, the effect of solubility is explored more generally in terms of the insoluble aerosol fraction f_{insol} (table 1). The partitioning parameters for the MO were taken as the average of two strong surfactants, cis-Pinonic and Suwannee River fulvic acids, $K = 31071$ and $\Gamma = 0.00255$ mmolm^{-2}, to best capture the bulk-surface partitioning phenomena. In partitioning schemes that use a fixed depressed surface tension a value of 55 mNm^{-1} is prescribed based on results found by Taraniuk et al. (2007) for humic-like substances. The insoluble black carbon component of the aerosol is modelled as elemental carbon with a density of 2000 kgm^{-3} and molecular mass of 12.0 gmol^{-1}. The inorganic fraction is modelled as a mixture of salts, including ammonium sulphate $(\mathrm{NH_4})_2\mathrm{SO_4}$, sodium chloride NaCl and ammonium nitrate $\mathrm{NH_4NO_3}$ for each environment, the molecular masses and densities of which can be found in table 1.

Parameter ranges explored in the sensitivity analysis of this study are taken from literature where possible. Ranges for compositional, α and f_{insol}, and modal distribution parameters are deduced from statistics and measurements contained within the references discussed above in relation to each environment. The density and molecular mass of the surfactant are perturbed between minimum and maximum values of the five compounds studied in Topping (2010), while the partitioning parameters K and Γ are perturbed between the values of two strong surfactants, cis-Pinonic and Suwannee River fulvic acid. Surface tension was allowed to vary between 30 mNm^{-1} and 72.8 mNm^{-1} (pure water) to account for particularly strong surfactants such as bio-surfactants (Ekström et al., 2010). The effect of non-ideal solutions is also explored by analysing spectra sensitivity to perturbations in Φ between 0.75 and 1.0.

3.2 Interpolation methods for CCN spectra modelling

In this section the practicalities of coupling the Köhler model to the size distribution, in order to calculate the number concentration of CCN, N_{CCN}, as a function of the ambient supersaturation, are discussed. For a given supersaturation, the activation radius r_{act} is defined as the dry radius of the aerosol such that

$$S_c(r_d = r_{act}) = S_a \tag{15}$$

for a given internally mixed composition. Köhler curves are generated for each dry size class of the size distributions discussed in section 3.1. In practice, owing to the discrete nature of the size classes, S_a will be between two critical supersaturations, S_c^i and S_c^{i+1}, corresponding to a smaller and larger dry size r_d^i and r_d^{i+1}, between which r_{act} lies. A linear interpolation is employed to calculate unique values of r_{act} for each supersaturation. With r_{act} determined, N_{CCN} can be calculated by integrating the size distribution

$$N_{\mathrm{CCN}} = \int\limits_{r_{act}}^{\infty} n(r)dr \tag{16}$$

where $n(r)$ is the number concentration size distribution function. It must be noted that r_{act} will lie between the lower and upper bounds of the activated size class q. The practical difficulty this causes is twofold: Firstly, when evaluating equation 16 as a summation, one must either discount the first bin number concentration or take its total number concentration. Secondly, should two or more r_{act} values fall into the same size class, then non-unique calculations of N_{CCN} will occur for different S_a, producing a step-like curve for the CCN spectrum. To circumvent this, fractional interpolation within the first activated size bin is employed between the upper and lower bounds, r_q^u and r_q^l respectively. Thus, in practice N_{CCN} is calculated as follows

$$N_{\mathrm{CCN}} = N^q \frac{(r_q^u - r_{act})}{(r_q^u - r_q^l)} + \sum_{i=q+1}^{400} N^i \tag{17}$$

The vector of N_{CCN} values together with their corresponding S_a values form the CCN spectrum. The importance of intra-bin interpolation, in relation to the resolution of the size distribution, for successful application of inverse modelling procedure, is explored fully in section 5.3 below.

4 Inverse modelling materials

Inverse modelling is a methodology often used for finding a set model input parameter values that produce model outputs that best represent measurement data. The optimisation procedure is usually performed using a least squares or maximum likelihood criterion with respect to some objective function (Vrugt et al., 2006). Mathematically, it is formulated as follows. Let $\tilde{\boldsymbol{C}} = \psi(X,\theta)$ denote the vector of n model predictions, say CCN concentrations $\tilde{\boldsymbol{C}} = (\tilde{c}_1,\ldots,\tilde{c}_n)$, where ψ denotes the model and X and θ are the fixed input variables and parameters for optimisation, respectively. Given a vectorial set of observations $\boldsymbol{C} = (c_1,\ldots,c_n)$, say observed CCN concentrations, then the deviation of model predictions, for a given set of θ, can be calculated as a vector of residual concentrations $r(\theta)$

$$\boldsymbol{R} = \tilde{\boldsymbol{C}}(\theta) - \boldsymbol{C} = [(\tilde{c}_1 - c_1),\ldots,(\tilde{c}_n - c_n)] = [r_1(\theta),\ldots,r_n(\theta)] \tag{18}$$

Thus, inverse modelling seeks to minimise $\boldsymbol{R}$ with respect to θ. In practice minimising a vector quantity can be challenging. This challenge can be overcome by introducing an OF - a scalar aggregate of the residuals. The aim now is to minimise this model-measurement discrepancy metric with respect to input parameter values. A parameter set that returns a zero valued OF corresponds to a perfect match between observations and model predictions. Producing a zero valued OF function with real-world observations is unlikely; however, synthetic modelling studies using model-generated measurements, such as this one, will result a zero-valued OF for parameter values used to generate the synthetic measurements.

Successful application of an inverse modelling approach to any given problem is reliant on an appropriate definition of both the calibration data, $\boldsymbol{C} = (c_1,\ldots,c_n)$, and OF. In subsections 4.1 and 4.2 below definitions of these objects are presented.

4.1 Synthetic calibration data: CCN spectra

Real-world measurement data is normally used as calibration data in model calibration and sensitivity studies. Here, however, synthetic measurements are numerically-generated from the model by using best estimate parameter values to represent real-world atmospheric conditions, henceforth referred to as the 'true' parameter values θ^{true} or calibration parameters, for input parameters. These true parameters are documented collectively in table 2. The calibration data is thus denoted $\tilde{\boldsymbol{C}}(\theta^{true})$ and is a vector of CCN concentrations where each element corresponds to each point on a prescribed supersaturation grid spanning 0.1-1.5% in increments of 0.1%. In choosing a CCN spectrum as the calibration data, the sensitivity analysis that follows includes information regarding a range of prevalent meteorological conditions that define various clouds types. The additional information content a CCN spectrum contains versus, say, a single CCN concentration at a fixed supersaturation is considerable, the importance of which is discussed at length in section 5.3.

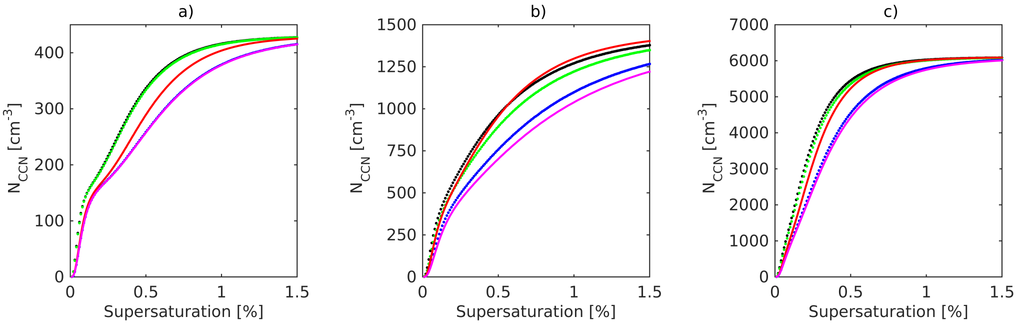

Figure 3. CCN spectra calculated from true parameter values (table 2) for: a) Marine average, b) rural continental and c) polluted continental environments. Partitioning schemes are coloured as in Fig.1. Models using a fixed surface tension value are shown by dotted lines while models using Eq. 9 are presented as line-plots.

All 12 sets of calibration data generated from true parameter values for each partitioning scheme and environment are presented in Fig. 3. Also included in Fig. 3, for reference, are CCN spectra generated from classical Köhler theory using the surface tension value of pure water. The differences between calibration data sets for different partitioning schemes arise for the same reasons as the changes in activation points shown in Fig. 1 and discussed in section 2.2.

4.2 The objective function (OF)

Care should be taken when choosing the functional form of the OF. The functional description should reflect the characteristics of measurement errors seen in the relevant observation data set. Common definitions for the OF include the simple least squares (SLS) or some maximum likelihood estimator. Definitions such as SLS or root mean square error (RMSE) are valid when the measurement errors are believed to be equal throughout the data set (homoscedastic) and uncorrelated. More generally, a weighted RMSE definition can be applied

$$OF = \left[\frac{1}{n}\sum_{i=1}^{n} w_i[\tilde{c}_i - c_i]^2\right]^{1/2} = \left[\frac{1}{n}\sum_{i=1}^{n} w_i r_i(\theta)^2\right]^{1/2} \tag{19}$$

where w_i is the weighting of the ith element. In the present study, the main sensitivity analysis carried out in section 5.2 assumes homoscedsaticity and therefore weightings are set to unity, $w_i = 1$ for all i. In section 5.3 the implications of supersaturation dependent natural variability (heteroscedasticity) in CCN measurements is explored with appropriately defined weightings.

To illustrate how the OF behaves in relation to perturbations in a single parameter, Fig. 4 shows how the rural continental CCN spectrum varies according to a perturbation in N_2 (upper panel) and the corresponding change in the OF (lower panel).

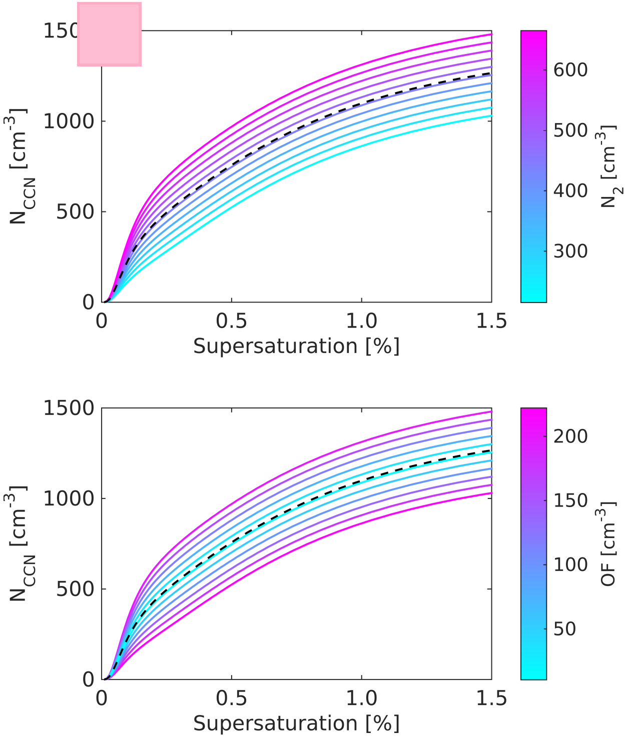

Figure 4. Rural continental CCN spectra for partitioning scheme $a_w^{np}\sigma^f$. In the top panel, the colour mapping indicates modelled CCN spectra as a function of N_2 within uncertainty ranges specified in table 2. In the bottom panel, the colour mapping indicates variation in the OF between the modelled spectra and the calibration data (dashed black line) for the corresponding calculations with respect to N_2.

5 Results and discussion

5.1 One at a time (OAT) parametric sensitivities

Typically, studies provide one at a time (OAT) sensitivity analyses of model outputs, e.g Wex et al. (2008). Although this methodology can be instructive it is not ideal. By performing an OAT analysis, large volumes of the full multi-dimensional parameter space remain unexplored and as a consequence the analysis may miss important parameter interactions that could result in suppressed or increased sensitivity. A brief OAT analysis of a subset of the parameters considered in this study is included here as an instructive step in developing the response surface methodology. In addition, the OAT analysis facilitates the identification of supersaturations at which individual parametric sensitivities are greatest.

Consider a fractional perturbation P to the true value of parameter i. The sensitivity $\delta\boldsymbol{C}$ to the ith parameter is thus calculated as follows

$$\theta_i^{perturbed} = (1+P)\theta_i^{true} \quad P = 0.1 \tag{20}$$

$$\delta\boldsymbol{C} = \tilde{\boldsymbol{C}}(\theta^{true}) - \tilde{\boldsymbol{C}}(\theta_i^{perturbed}, \theta_{j\neq i}^{true}) = (r_1,\ldots,r_n) \tag{21}$$

where $\theta_i^{perturbed}$ is the perturbed parameter for OAT analysis and the index j runs over all other parameters. By applying the same fractional perturbation to each parameter the relative parametric sensitivities can be assessed. Figure 5 shows the calculated sensitivities for perturbations in N_2, $\bar{r}_2$, α, ρ_{org}, σ, K and Γ as a function of supersaturation. This is repeated for all partitioning schemes for the average marine environment.

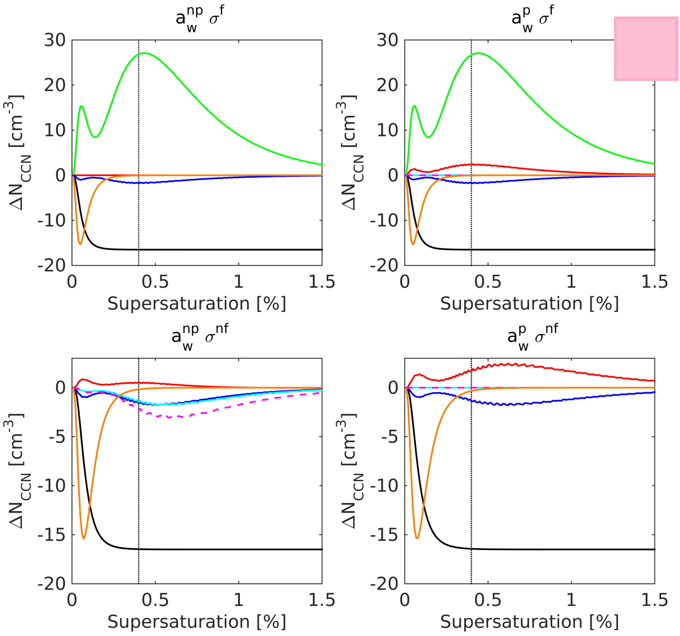

Figure 5. Sensitivity curves for marine average CCN concentrations as a function of supersaturation. Selected parameters are perturbed individually by 10% for all partitioning schemes; N_2 (black), $\bar{r}_2$ (orange), α (red), ρ_{org} (blue), σ (green), K (cyan) and Γ (magenta). A vertical, grey dashed line is included at $S_a = 0.4\%$ to indicate a regime change between stratiform and convective cloud types.

Global variability in updraft velocities has considerable importance for the aerosol indirect effect as it leads to the development of different cloud types and a range of supersaturations (West et al., 2014). In Fig. 5 the distinction between stratiform and convective cloud types is illustrated by a grey, vertical dashed line at $S_a = 0.4\%$ corresponding to an updraft of approximately 0.5 ms^{-1} in marine environments (Chuang, 2006). In Fig. 5 there is a local sensitivity maximum for most parameters, for all partitioning schemes, around 0.1-0.2% corresponding to stratiform cloud types. Furthermore, sensitivity is at a maximum for many parameters at humidities close to the convective threshold and higher. Higher humidities are characteristic of deep convective systems typically seen in the tropics giving rise to Hadley cell circulation. Therefore, when simulating CCN concentrations in convective models, such as the CRM-ORG developed by Murphy et al. (2015), with complex organic representations, special consideration may be required when choosing the physicochemical aerosol input parameters.

It is clear that the surface tension σ is the most sensitive Köhler parameter when considering $a_w^{np}\sigma^f$ and $a_w^p\sigma^f$ bulk-surface partitioning schemes. This suggests that constraining uncertainties in σ is important for CCN activation. As σ shows a high degree of sensitivity and is difficult to measure in-situ at the scale of interest, it is likely a good candidate for optimisation using an inverse methods. For the simple Köhler case, $a_w^{np}\sigma^f$, this result is in agreement with results obtained by Wex et al. (2008) which show a strong sensitivity of the critical supersaturation to σ perturbations. CCN concentrations are also highly sensitive to the number concentration of the accumulation mode particles N_2 for all partitioning schemes. Large sensitivity to N_2 is expected as a large fraction of the accumulation mode particles have radii greater than that of typical activation radii, even at lower supersaturations. CCN concentrations are also sensitive to chemical and compositional parameters ρ_{org} and α but less so. There is some symmetry in the sensitivity above and below the $\delta\boldsymbol{C} = 0$ line for ρ_{org} and α, therefore, when considering simultaneous perturbations to these parameters, it is likely there will be a non-unique set of parameter pairs returning minimal deviation from the calibration data due to parameter interactions.

For $a_w^{np}\sigma^{nf}$ and $a_w^p\sigma^{nf}$ partitioning schemes, σ is replaced by the partitioning parameters Γ and K that are used to model the surface tension using Eq. (9). For $a_w^{np}\sigma^{nf}$ there is moderate sensitivity to both partitioning parameters, suggesting that if these empirically derived parameters are to be used in Köhler modelling they must be known to some degree of accuracy for meaningful conclusions to be reached. For the full partitioning scheme $a_w^p\sigma^{nf}$ however, there is negligible sensitivity. The lack of sensitivity for the complete scheme is likely due to the competing effects of reduced surface tension and increased bulk water activity on the point of activation. It is necessary to analyse these two parameters simultaneously across their uncertainty ranges to provide a clearer picture, an instructive tool for such analysis is the response surface.

5.2 Response Surface Analysis

A GSA is preferred over traditional OAT analyses as it provides a comprehensive analysis that spans the entirety of the parameter space (Pérez et al., 2006) thus arriving at a more extensive and reliable set of results. This is particularly pertinent when applied to highly non-linear systems such as those found in cloud-aerosol interactions as multidimensional parameter interactions can significantly affect individual parameter sensitivities when the entire parameter space is explored (Partridge et al., 2011, 2012). With the application of an automatic search algorithm an inverse modelling framework can be used to simultaneously facilitate a GSA and parameter optimisation (Partridge et al., 2012). In this section a graphical tool that provides qualitative CCN spectra sensitivity information for simultaneous perturbations in two parameters is employed - response surfaces. These surfaces also provide some indication of the viability of a GSA using, for instance, an MCMC-based automatic search algorithm.

Traditionally, in 2D sensitivity analyses the surface illustrates the response in a single model output variable, for example Quinn et al. (2008) investigated the response in CCN concentration, at fixed superaturations, to perturbations in insoluble fraction and mean diameter. Here, however, response surfaces are used as a graphical tool used to illustrate the response of the

OF as a function of the perturbed aerosol physicochemical parameters so as to capture CCN sensitivity information across a range of supersaturations and cloud types.

Consideration of the behaviour of the OF in 2D planes of the full parameter space is also instructive for testing that aerosol-CCN closure is an appropriate problem for investigation using inverse methods. While the response surfaces only suggest how the OF may evolve when traversing the full parameter space, if the surfaces do not show a single well defined minimum then it may certainly be expected that inverse parameter optimisation may be unsuccessful (Partridge et al., 2011). Parameters that have a wide range of values while maintaining minimal deviation in CCN spectra from the calibration data are deemed 'non-identifiable' and will be difficult to calibrate based on the current information content of the calibration data. Having many such parameters may reduce algorithm efficiency and hamper calibration of more important parameters. Response surfaces provide a way of visually discerning such parameters to be removed from the optimisation procedure. Surfaces possessing a well defined minimum are preferred as algorithms tend to iterate more efficiently if the gradient of improvement points toward a single attractor within search space. Response surfaces containing single attractors and steep gradients suggest that the associated parameters are both sensitive and viable candidates for calibration. A high degree of sensitivity also implies that it is important to represent such parameters well in GCMs for accurate predictions of climate evolution. Using response surfaces to visualise the evolution of an OF across 2D parameter planes has been used effectively in similar highly non-linear atmospheric inverse problems (Partridge et al., 2011).

Figure 6a) illustrates how the critical supersaturation S_c of a dry aerosol with a 75 nm radius and marine average composition, as modelled using traditional Köhler theory, evolves through the parameter space when subject to simultaneous perturbations in α and σ. It is clear that the sensitivity of the activation point to perturbations in σ is greater than that of α. In addition, non-unique values of α can result in the same S_c value, this result is similar to results obtained by Wex et al. (2007) for a bulk parameter of chemical properties. Figure 6b) shows the OF response in relation to the CCN spectrum for the same composition and parameter perturbations. Blue crosses indicate the coordinates of the true parameter values and thus falls on a point where the OF is zero.

In what follows, parameter sensitivities for all four partitioning schemes in the marine average environment are analysed in sections 5.2.1 - 5.2.4 before considering environmental dependencies in section 5.2.5. Attention is given to the marine environment as the extensive spatial coverage, high contrasting albedo relative to the surface (Warren et al., 1986, 1988) and long synoptic lifetime (Brenguier and Wood, 2009) of marine stratocumulus result in a greater climate sensitivity to changes in CCN concentrations than other environments. All response surfaces calculated can be found in the supplementary information.

Parameters of interest are perturbed across ranges of values that reflect uncertainties found in existing observations that include both laboratory and in-situ measurements. These ranges are documented in table 2 and discussed in section 3.1. Blue crosses indicate the true parameter values, and collectively correspond to the full true parameter set θ^{true}, and therefore lie on a point where the OF is zero. Constant value OF contours and a colour mapping are used to visualise the deviation of CCN spectra from the calibration data. Dark (hot) regions of the parameter plane indicate a high value OF and thus large deviations of modelled CCN spectra from the calibration data; light (cold) regions indicate low values of the OF and thus small deviations from calibration data.

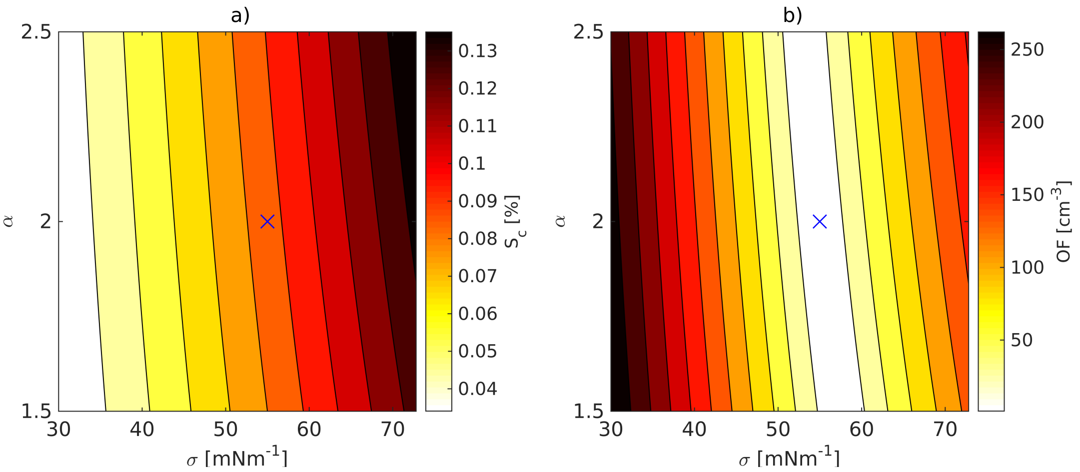

Figure 6. a) Response surface for the critical supersaturation S_c of a $r_d = 75$nm marine average aerosol for perturbations in α and σ across ranges given in table 2. b) Response surface of the OF for marine average CCN spectra with respect to the same parameter perturbations as in a). Blue cross indicate the true values of the two parameters.

5.2.1 Classical Köhler theory: $a_w^{np}\sigma^f$

Figure 7a-d) shows response surfaces for four parameter combinations for classical Köhler theory $a_w^{np}\sigma_{ws}^f$ in the marine average environment. This particular formulation of Köhler theory presents 9 parameters for analysis, leading to the calculation of 36 response surfaces, 14 of which possess a well defined minimum. The response surfaces presented in Fig. 7 are chosen to illustrate the relative sensitivities of size distribution and Köhler theory parameters. The complete set of response surfaces (documented in the supplementary information) indicates that σ, α, f_{insol}, Φ, N_2 and $\bar{r}_2$ are the most sensitive parameters. In Fig. 7d) the response surface for perturbations in α and σ does not contain a well defined minimum as seen in Fig. 7a-c) for other parameters. Interactions between σ and α allow α to take any value across its uncertainty range and return a zero OF for a narrow band of σ values close to its true value.

5.2.2 Redistribution of surfactant concentration: $a_w^p\sigma^f$

Response surfaces were recalculated for the inclusion of bulk-surface partitioning effects in the Raoult term $a_w^p\sigma_{ws}^f$ which accounts for the reduced bulk concentration of surfactant when calculating the water activity. This allows the effect of bulk-surface partitioning on the bulk water activity to be isolated from the effects of a concentration dependent model of surface tension, Eq. (9). To model the effects on the bulk activity the partitioning parameters, Γ and K, must be introduced. Thus all 11 parameters are analysed in this setup, totalling 55 response surfaces 11 of which contain a well defined minimum. Figure 8 shows response surfaces for this setup. Response surfaces for parameter pairs common to this scheme and classical Köhler theory showed negligible changes, indicating that the effect concentration partitioning is relatively unimportant for determining CCN spectra. This result is further highlighted in Fig. 8d), while parameter interactions similar to those seen in

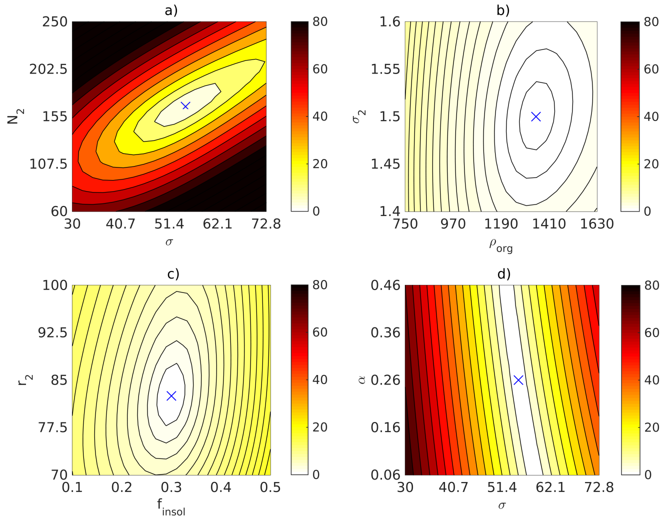

Figure 7. Response surfaces for $a_w^{np}\sigma^f$ in the marine average environment. Blue crosses indicate the true parameter values θ^{true} used to calculate the calibration data. The colour scale represents the value of the OF calculated for the modelled spectra against the calibration data for parameter values across the uncertainty ranges (table 2).

Fig. 7d) are present, the OF takes a much lower value indicating low sensitivity within the specified uncertainty range. The lack of sensitivity to the partitioning scheme as well as the partitioning parameters is attributed to its action solely through the water activity term. The water activity, in its mole fraction form, Eq. (2), is typically close to unity at the point of activation as $n_w \gg n_s$ and therefore any changes to bulk concentrations of the solute moles n_s may certainly be expected to have negligible influence on the mole fraction. This is further reinforced by the black ($a_w^{np}\sigma_{ws}^f$) and green ($a_w^p\sigma_{ws}^f$) curves in figures 1 and 3, wherein only small changes to the critical supersaturation S_c and the CCN spectrum, respectively, are seen. This scheme should not be considered as an accurate representation of what occurs in nature; here a concentration-dependent bulk water activity has been used whilst a fixed concentration-independent surface tension has been applied. Nevertheless, it remains instructive to isolate and ascertain the magnitude of the effect of such a phenomenon on CCN activation so that it can be disregarded in future studies and model and parametrization developments.

The sensitivity to perturbations in solution ideality is shown in Fig. 8b) and is found to have a similar sensitivity to the modal radius $\bar{r}_2$. Information on ideality of atmospheric particles is challenging to measure in-situ and given the relatively high sensitivity of the parameter shown here, it is a particularly good candidate for optimisation using inverse methods if there is adequate constraint on additional processes. The chemical properties of the surfactant, M_{org} and ρ_{org} (see supplementary information), were found to be relatively insensitive when compared with other parameters for this partitioning scheme and classical theory discussed in section 5.2.1. The relative unimportance of these chemical properties is in-line with the general conclusion reached by Dusek et al. (2006) that chemistry is less important than size. However, in contrast, the sensitivities of compositional parameters, α and f_{insol}, and the ideality of the solution Φ exhibit similar sensitivities when compared to

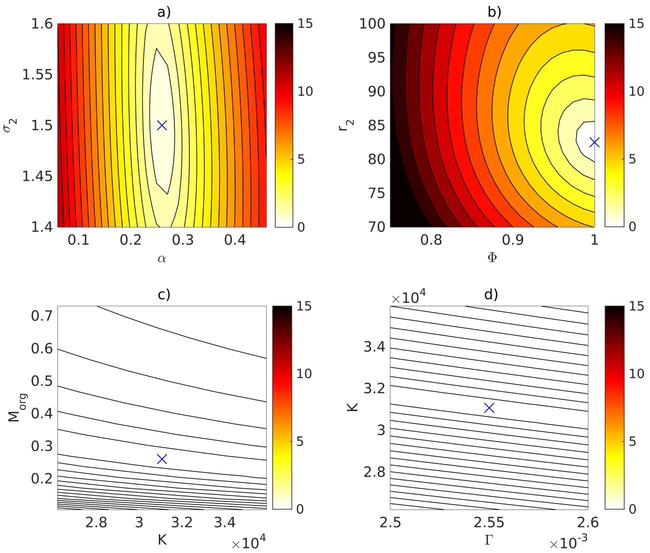

Figure 8. Response surfaces for $a_w^p\sigma^f$ in the marine average environment. Blue crosses indicate the true parameter values θ^{true} used to calculate the calibration data. The colour scale represents the value of the OF calculated for the modelled spectra against the calibration data for parameter values across the uncertainty ranges (table 2).

size distribution parameters σ_2 and $\bar{r}_2$. Dusek et al. (2006) carried out their sensitivity analysis on CCN size distributions at individual supersaturations of 0.25%, 0.4% and 1.0% thereby missing sensitivity information for all stratiform cloud types below 0.25% and convective cloud types above 1.0%. Here the analysis has been carried out over a highly resolved range of atmospheric supersaturations. This difference in methodology is likely the cause of contrasting results as the discrete nature of their analysis may miss peaks in individual parameter sensitivities such as those seen in Fig. 5, the effect of which are quantified and accounted for in the OF methodology developed here.

5.2.3 Surface tension considerations: $a_w^{np}\sigma^{nf}$

Here the effects of a concentration-dependent surface tension, Eq. (9), are accounted for while concentration partitioning is not accounted for in evaluating the bulk water activity - $a_w^{np}\sigma^{nf}$. In this partitioning scheme the partitioning parameters, Γ and K, replace surface tension so the analysis covers 10 parameters and thus 45 parameters planes, 21 of which possess a well defined minimum.

Response surfaces for this partitioning scheme are shown in Fig. 9. Figure 9d) illustrates that the CCN spectrum shows a higher degree of sensitivity to changes in Γ and K than in the $a_w^p\sigma^f$ case, confirming results from the OAT analysis. This result is expected on account of their action through the surface tension - the Köhler parameter commonly found to be one of the most sensitive in determining S_c (Wex et al., 2008). This effect manifests itself as an increased CCN activity and can be readily seen in Fig. 3 (red) for all environments when compared with simple Köhler theory using surface tension for water

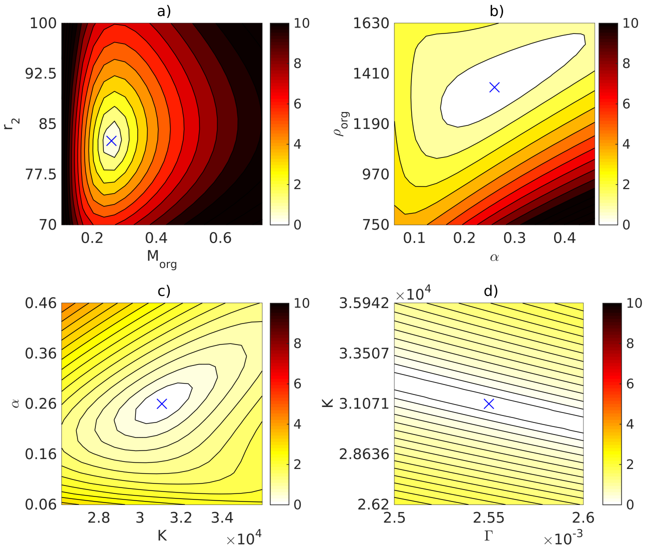

Figure 9. Response surfaces for $a_w^{np}\sigma^{nf}$ in the marine average environment. Blue crosses indicate the true parameter values θ^{true} used to calculate the calibration data. The colour scale represents the value of the OF calculated for the modelled spectra against the calibration data for parameter values across the uncertainty ranges (table 2).

(blue). In addition, the parameter plane for the partitioning parameter Fig. 9d) also shows a strong interaction between the two parameters that can result in non-unique optimised parameter values for a zero OF. This suggests that the chosen calibration data may not contain the necessary information to correctly calibrate K and Γ for this partitioning scheme.

5.2.4 The complete partitioning scheme: $a_w^p\sigma^{nf}$

Here the full partitioning framework is considered. The surface tension is calculated using the partitioning parameters Γ and K as in section 5.2.3, therefore there are the same number of parameters and surfaces for consideration. 10 response surfaces show a well defined minimum.

Figure 10 contains response surfaces for this comprehensive partitioning scheme. There are clearly strong similarities between Fig. 10c-d) and Fig. 8c-d) and the partitioning parameters seem to be relatively insensitive when compared to size distribution parameters. This is explained by the dependence of surface tension on the organic activities as well the partitioning parameters. For the $a_w^{np}\sigma^{nf}$ scheme there is substantial depression of surface tension through Γ and K and therefore significant sensitivity, here however the dependence on the organic activity pushes the value of surface tension back towards that of water at the point of activation. This is also clear from the calibration data plotted in Fig. 3 (pink and blue) for a single point in the parameter space.

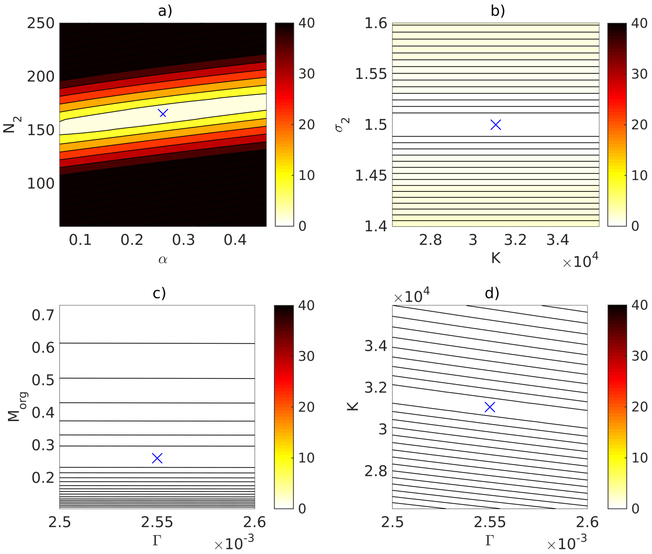

Figure 10. Response surfaces for $a_w^p\sigma^{nf}$ in the marine average environment. Blue crosses indicate the true parameter values θ^{true} used to calculate the calibration data. The colour scale represents the value of the OF calculated for the modelled spectra against the calibration data for parameter values across the uncertainty ranges (table 2).

The ability of simple Köhler theory, when the surface tension of water is used, to approximately replicate the CCN concentrations generated from the full partitioning treatment is in agreement with existing literature (Prisle et al., 2012, 2010; Sorjamaa et al., 2004).

For the full partitioning scheme considered here, the relative sensitivity of each parameter, and both their linear and non-linear interactions, are summarised in table 3. Parameters that are indicated to have high or very high sensitivities are good candidates for a future study using automated search algorithms to provide a quantitative GSA and parameter optimisation with respect to appropriate definition of calibration data (section 5.3).

5.2.5 Environmental considerations

Sensitivities of the organic chemical parameters were not found to vary a significant amount between environments and therefore we have not included response surfaces for all environments in sections 5.2.1-5.2.4. In Fig. 11a-c) the response surfaces for N_2 and α perturbations are shown for all three environments. In panels A (marine average) and B (rural continental) very similar parameter interactions are evident - a zero value OF is returned across the entire range of uncertainty in α, suggesting α is insensitive and thus cannot be calibrated to a unique value based on the information content of the CCN spectrum alone for these environments. In panel C (polluted continental), a higher degree of sensitivity to α is clear from a steeper gradient parallel to its axis. The increased sensitivity allows the uncertainty to be constrained subject to interactions with N_2 in more polluted environments. However, this constrained uncertainty in α has come at the cost of a less well defined N_2. It should be noted however, that despite this challenging outlook, response surfaces provide only a glimpse of the full parameter space and a well defined minimum may exist where a third, or several parameters, push CCN spectra into a different sensitivity regime. To perform a rigorous analysis, automatic search algorithms must be employed using a selection of parameters believed to be identifiable from a thorough response surface analysis as presented here.

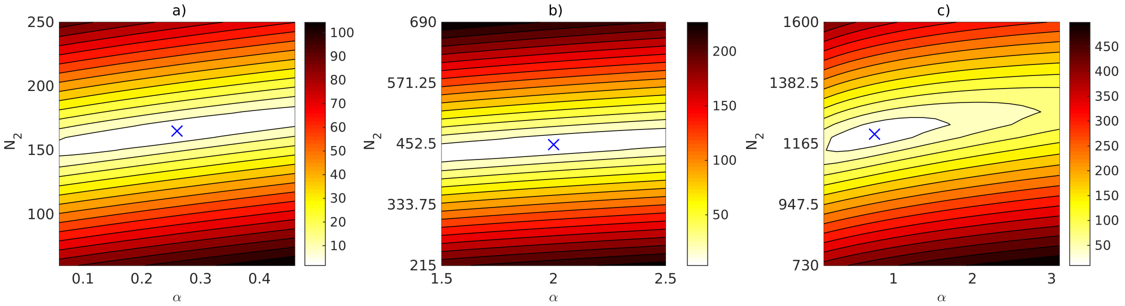

Figure 11. N_2 vs. α response surfaces of the OF for the $a_w^p\sigma^{nf}$ partitioning scheme in the a) marine average, b) rural continental and c) polluted continental environments. Blue crosses indicate true parameter values.

5.3 Information content and CCN observations as calibration data

For a single parameter to be considered identifiable, only one response surface showing a well-defined minimum is required. However, if the parameter does not exhibit well defined minima in several parameter pairs, and in particular if these surfaces are relatively flat, then automatic search algorithms will likely struggle to converge on unique parameters values. Response surfaces that are flat with respect to perturbations in a particular parameter indicate that such a parameter is insensitive and thus accurate calibration is unnecessary for the model under consideration. Insensitive parameters can be removed from the optimisation procedure and replaced with a fixed value. In GCMs parameters such as surface tension and the hygroscopicity parameter κ (Petters and Kreidenweis, 2007) are often implemented as fixed values. However, CCN and cloud droplet concentrations are know to be sensitive to these parameters. For example, Ervens et al. (2010) showed that different assumptions regarding the value of κ can result in upto a twofold difference in CCN concentrations. Studies seeking to calibrate parameters which are not measurable in-situ at the scale of interest must take care when defining calibration data so as to include as much information content as possible. Response surfaces in sections 5.2.1-5.2.4 and the supplementary information show that using CCN spectra alone as calibration data will likely not be sufficient for calibration of all parameters considered in the present study. Therefore, it is recommended that further studies are conducted to identify the appropriate in-situ measurements required and thus inform the experimentalist community accordingly. One way to help abate this issue is to define the calibration data set such that it contains a greater information content. For example, additional information content can be introduced through further temporal and spatial measurements or higher resolution calibration data.

Figure 12a) shows that without interpolation CCN spectra exhibit a ’stepping’ in CCN concentrations as multiple activation sizes, corresponding to continuous intervals on the supersaturation axis, fall between the limits of the same size class. Thus a considerable amount of information content has been added to the calibration data by providing unique values of CCN concentrations at different supersaturations. Stepped CCN spectra result in multiple local minima in response surfaces Fig. 12b and d). Convergence of automatic search algorithms in parameter spaces containing multiple local minima proves challenging, therefore smoothing of the calibration data is crucial to the methodology presented.

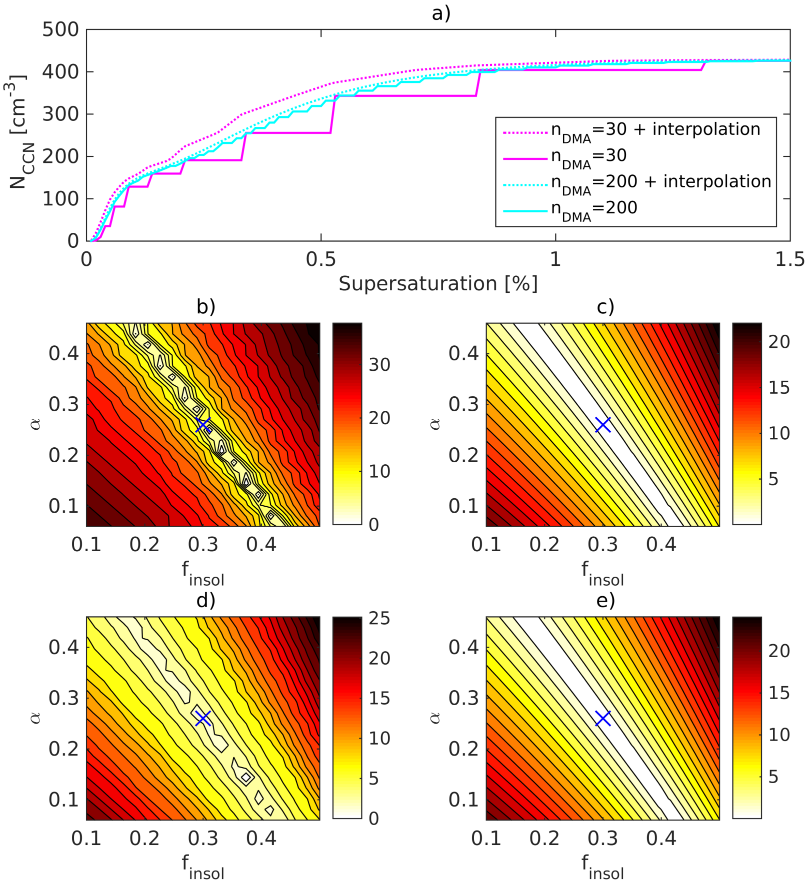

Figure 12. a) marine average calibration data generated from an aerosol size distribution of 30 sizes bins (magenta) and 200 sizes bins (cyan) with (dotted line) and without (solid line) intra-bin interpolation. The corresponding OF response surface for b) 30 bins and no interpolation; c) 30 bins with interpolation; d) 200 bins and no interpolation; e) 200 bins with interpolation.

In the absence of direct measurements of CCN spectra for real world calibration data sets, model predictions of the activation point could be used to derive pseudo-synthetic CCN spectra from aerosol size distributions measured by differential mobility analysers (DMA). DMA instrumentation can vary substantially in size resolution. Figures 12a,b and d) show that decreasing the number of size classes n_{DMA} from 200 to 30 reduces the information content in the calibration data considerably and produces rougher OF response surfaces as result. However, it is encouraging that the topography of the OF response surfaces for these calibration data sets do not depend on n_{DMA} when interpolating, Fig. 12c) and e). That is to say, derived parametric sensitivities are independent of n_{DMA} despite increased the CCN concentrations at lower resolutions, Fig. 12a), that arise due to increased bin-width.

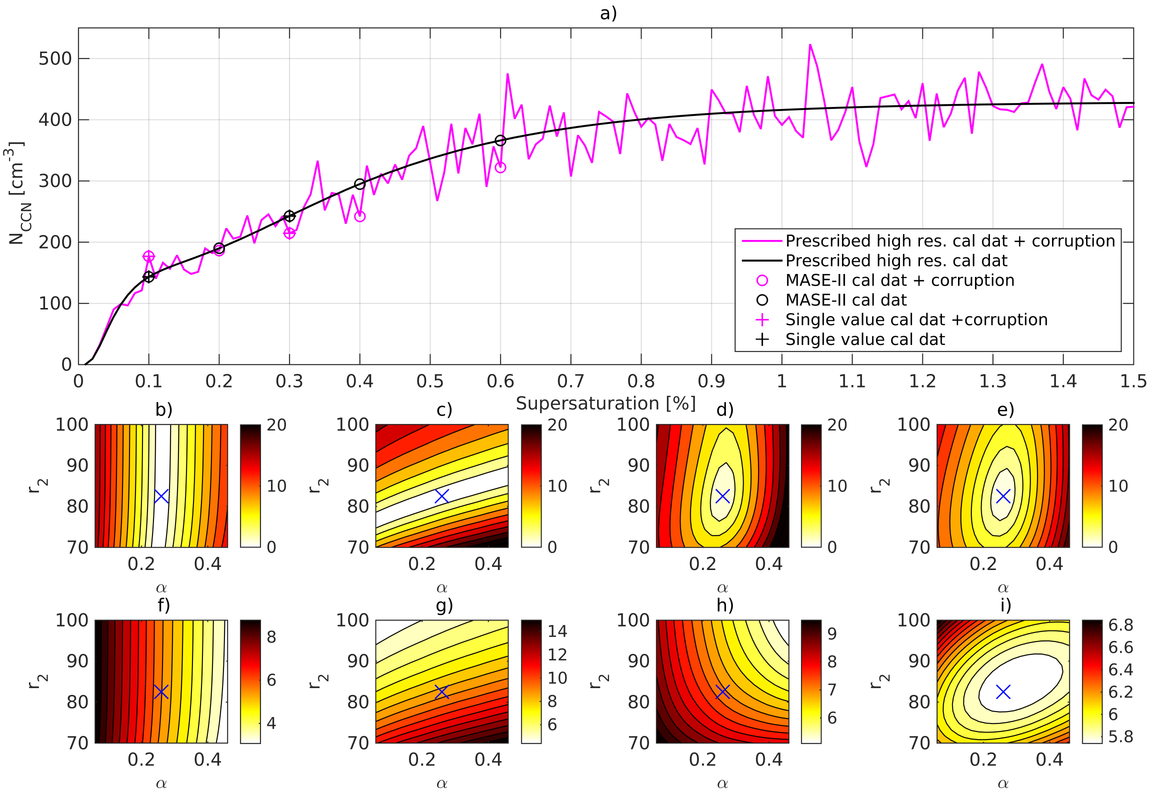

Figure 13. a) Multiple definitions of calibration data (cal dat). Curves and data points in black correspond to uncorrupted definitions; high resolution synthetic CCN spectrum (solid line), supersaturation bands corresponding to CCNC measurements from the MASE-II campaign (circles) and two single value definitions at 0.1 and 0.3% (crosses). Magenta: calibration data corrupted with randomly generated error assuming a standard deviation of 10% of the uncorrupted values. b)-e) OF response surfaces corresponding to the uncorrupted single values 0.3% and 0.1%, MASE-II supersaturation bands and synthetic CCN spectrum definitions of calibration data.f)-i) response surfaces as in b-e) but with the corruption of calibration data.

The importance of information content is particularly evident when considering the resolution and range supersaturations spanned by the calibration data. Multiple definitions of calibration data are shown in Fig. 13a) both with and without corruption by randomly generated synthetic natural variability: single-values at 0.1 and 0.3% corresponding to supersaturations typically used in analyses such as Aerosol Comparisons between Observations and Models (AEROCOM) (Ghan et al., 2016); a typical five-band CCN spectrum representative of data collected in the second marine stratus/stratocumulus experiment (MASE-II) campaign (Lu et al., 2009) and a high resolution CCN spectrum such as that used in the present study. Figure 13b-e) show response surfaces that highlight the relevance of resolution for calibration of the mode size and organic to inorganic mass ratio. Panel b) indicates that at a supersaturation of 0.3% sensitivity to mode size is negligible, as seen in the OAT analysis Fig. 5, the parameter is non-identifiable and attempts to constrain its parametric uncertainty will fail. In panel c) single valued calibration data at a supersaturation of 0.1% shows increased sensitivity to the mode size and that it may be possible to reduce parametric uncertainty through MCMC analysis. However, the parameter remains non-identifiable within a narrower range. Therefore, when using single valued calibration data such as seen in Lee et al. (2013), care must be taken as parametric sensitivities are expected to be dependent on the cloud base supersaturations in GCMs which, in reality, are subject to considerable sub-grid scale variability (West et al., 2014). By increasing the range and resolution of the calibration data Fig. 13d-e) additional information content is introduced that allows the mode size to be identified on the response surface. Surfaces corresponding to calibration data defined from a CCNC measured spectrum shows little difference when compared with the high-resolution case as sufficient information content for identifying the mode radius is contained at lower supersaturations. The dependence of parameter identifiability on calibration data range and resolution indicates that constraint of parametric uncertainty using algorithmic approaches may be expected to fail when using single valued definitions opposed to spectral definitions. Such a result provides motivation for and highlights the virtue of a response surface pre-analysis to arrive at an appropriate definition of calibration data and selection of parameters to optimise.

In real world CCN measurements substantial natural variability is present and dominates instrumentation errors. To represent natural variability in the calibration data the i-th data point is corrupted by a normally distributed error with a standard deviation $\sigma_i = 0.1 \times \tilde{c}_i(\theta^{true})$. For a robust treatment of heteroscedastic errors associated with CCN spectra, residual weightings w_i in the OF (eq.19) can not be taken to be unity, instead they are calculated as $w_i = 1/\sigma_i$ in a similar approach to that presented by Partridge et al. (2012). Such a treatment provides preliminary insights into the feasibility of achieving aerosol-CCN closure when confronting the model with real world CCN observations using the inverse framework. Response surfaces given in Fig. 13f-i) repeat those in calculated in panels b-e) but with the inclusion of natural variability and appropriate weightings. Here the OF can no longer be minimised to zero as the model can not be considered 'perfect' in relation to the corrupted calibration data. Panels f-h) that correspond to single value and CCNC-like definitions of calibration data indicate that simultaneous minimisation of the OF and parameter calibration is not possible. That is to say that inverse closure is not possible. If a response surface analysis were to be bypassed, the inability to achieve inverse closure using MCMC or similar may lead one to conclude that structural uncertainties may exist in the model. However, for the high resolution case, panel i), inverse closure is possible for the parameters considered. Therefore, failure to attain inverse closure with respect to panels f-h) should not be attributed to model structural uncertainties but rather insufficient information content at such resolutions. Thus continued study using MCMC and calibration data defined by corrupted synthetic measurements or a functional fitting of CCNC-obtained spectra is recommended. The reliance of the success of inverse closure on calibration data resolution should also serve as a recommendation for the development of instrumentation, for in-situ deployment, which has a higher resolution than existing CCNCs.

When adding information content in a synthetic study such as the one performed here, it is important to be mindful that such information content could be retrieved from field observations as the end goal is to compare with an observational data set rather than synthetic measurements. In this study the information content is solely composed of the deviation of model predictions for a single calibration data set. If, in future studies, parameter search algorithms do not converge in an efficient manner additional information content can be supplied from in-flight temporal CCN measurements taken by counters at fixed supersaturations, or from the interstitial fraction of the the aerosol size distribution.

6 Conclusions

A methodology that is able to scrutinize the sensitivity of Köhler theory to perturbations in physicochemical parameters across a range of atmospherically relevant supersaturations has been constructed. The response surface analysis provides a visualisation of pairwise parameter sensitivity while simultaneously confirming aerosol-CCN spectrum closure as a well-posed inverse problem for appropriately defined calibration data. Across all partitioning schemes and environments a total of 543 response surfaces were calculated.

In agreement with Djikaev and Ruckenstein (2014), the response surface analysis here confirms that the density and molecular weight of the surfactant have only a small effect on aerosol activation. In further agreement, Wex et al. (2007) parametrised their Köhler model with a bulk chemical parameter that contained the density, molecular weight and effective dissociation of the organic substance, and found it to take a constant value over a range of values for those properties, resulting in a constant activation point.

For all partitioning cases, model sensitivity to surface tension, solution ideality and compositional fractions is on the order of that of the size distribution parameters. This is in contradiction to conclusions reached by Dusek et al. (2006) and certainly warrants further investigation. The ability of this novel framework to probe sensitivity over a range of atmospherically relevant supersaturation is likely the source of this difference as considerations in Dusek et al. (2006) are at fixed supersaturations. As response surfaces have shown relative similarities in model sensitivity to these parameters, an inverse modelling study using MCMC with high resolution calibration data that accounts for natural variability appears to be feasible, and therefore will form the focus of a second study to better quantify parametric sensitivities and constrain parametric uncertainties in a full GSA.

Nozière et al. (2014) used state of the art extraction techniques (Baduel et al., 2012) that allowed them to observe surface tension values as low as 30 mNm^{-1} in aerosols when allowing for equilibration times. Here, however, it has been deduced that due competing partitioning effects - surface tension depression and redistribution of surfactant concentration - result in values of surface tension close to that of water at the point of CCN activation. This result supports conclusions reached by Prisle et al. (2012) for global simulations CDNC. Therefore, the full treatment of bulk-surface partitioning returns CCN spectra almost identical to those calculated using simple Köhler theory with a surface tension value of water. Furthermore, the measurements recorded by Nozière et al. (2014) suggest the bulk-surface framework employed here may not be correct model for the behaviour of all surfactants and concentrations such as those they reported. This warrants further theoretical and/or experimental investigations. It should be noted that care should be taken when choosing values for surface tension in future aerosol-CCN closure studies based on the framework built here; CCN counter residence times, for example, are not long enough to facilitate surfactant equilibration times such as that reported by Nozière et al. (2014), therefore a closure study will likely prove unsuccessful or, at best, successful for the wrong reasons. Similarly, it has been suggested that lack of control of the saturation ratio of all semi-volatiles in such instruments might influence retrieved single particle properties (Topping and McFiggans, 2012). With respect to future use of the four partitioning schemes employed here, while treatment of surface tension depression or water activity alone lacks a physically justified basis, we should not necessarily continue to use a simple Köhler theory with the surface tension of water or, at additional computational demand, the full treatment blindly, as neither formulations have seen adequate verification from experimental data relating to surfactants. This is complicated by the possibility of additional composition dependent processes not only related to surfactant behaviour. Nonetheless, in order to increase current understanding of the role that bulk-surface partitioning plays in cloud nucleation, development of more sophisticated instrumentation, such as that of Baduel et al. (2012), must be a priority. In particular, CCN counters that operate on a highly resolved range of atmospheric supersaturations and have residence times to sufficiently capture equilibrium of surfactants are required. In addition, the use of single particle levitation techniques such as those used by Lienhard et al. (2015) might provide additional insights into the role of surface tension should they be able to access complex mixed aerosol.

Ervens et al. (2005) examined several chemical and compositional effects simultaneously and found compensating parameters resulted in a decreased sensitivity of total cloud droplet numbers when compared to studies treating the effects individual. Therefore, to have a good understanding of these effects global sensitivity analyses (GSA) are required. Response surfaces have here shown that several parameters may be identifiable in the complete parameter space and also that there are many interacting parameter pairs. Interacting parameters indicate the model under consideration can be simplified by reducing interacting parameters in single parameter as performed for κ by Kreidenweis et al. (2005). Therefore the inverse modelling framework developed here will be revisited in a future study. By implementing a Monte Carlo Markov Chain (MCMC) algorithm in a similar manner as performed by Partridge et al. (2012), a statistically conditioned parameter optimisation and GSA can be conducted. The applicability of the algorithm will be first benchmarked against synthetic measurement data, i.e. calibration data used in this study, before being applied to real world measurements taken from the European Integrated project on Aerosol Cloud Climate and Air Quality interactions (EUCAARI) (Paramonov et al., 2015). While the response surface analysis here suggests that a proper treatment of bulk-surface partitioning produces CCN concentrations similar to those of the classic Köhler theory, thus questioning its use in already computationally demanding global modelling, it only provides insight into 2D planes of the full parameter space. Using an MCMC simulation this preliminary conclusion can be readdressed using a more rigorous approach that also provides a greater understanding of the entire parametric landscape.

At this stage, results show that there are many parameter interactions present in CCN modelling. In addition, it is also clear that log-normal distribution parameters, compositional fractions, surface tension and solution ideality are all parameters that exhibit high sensitivity and as a community we must seek to reduce uncertainties in these parameters for effective global climate modelling. Herein it has been demonstrated that inverse modelling of CCN spectra may indeed be an effective methodology for constraining these uncertainties under an appropriate definition for the calibration data. Both the resolution and range of the calibration data are important not only for diagnosing parametric sensitivities, but also for simultaneous minimisation of the OF and correct parameter calibration, i.e. ascertaining the feasibility of inverse closure. In particular, a future study using an MCMC approach to GSA and inverse modelling seems promising when employing uncorrupted CCNC-like synthetic calibration data. However, the end goal is to confront the model with real world observations. By corrupting the synthetic calibration data to represent natural variability, it is shown that a high resolution definition of the calibration data is required for successful aerosol-CCN inverse closure with respect to some parameters. As such, this result should serve as a recommendation for the development of instrumentation that can be used in-situ to measure CCN spectra at higher resolution. In the absence of such instrumentation, functional fitting of CCN spectra obtained from current instrumentation will be required.

Acknowledgements. This work was supported by the UK Natural Environment Research Council grants NE/I020148/1 (Aerosol-Cloud Interactions - A Directed Programme to Reduce Uncertainty in Forcing) and NE/J024252/1 (Global Aerosol Synthesis And Science Project). P. Stier would like to acknowledge funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) ERC project ACCLAIM (grant agreement no. FP7-280025).

References

Abdul-Razzak, H., Ghan, S. J., and Rivera-Carpio, C.: A parameterization of aerosol activation: 1. Single aerosol type, J. Geophys. Res., 103, 6123–6131, 1998.

Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, Science, 245, 1227–1230, 1989.

Baduel, C., Nozière, B., and Jaffrezo, J.-L.: Summer/winter variability of the surfactants in aerosols from Grenoble, France, Atmos. Environ., 47, 413–420, doi:10.1016/j.atmosenv.2011.10.040, 2012.

Bigg, E.: Discrepancy between observation and prediction of concentrations of cloud condensation nuclei, Atmospheric Research, 20, 81–86, 1986.

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, 32 005–32 018, 2001.

Bougiatioti, A., Fountoukis, C., Kalivitis, N., Pandis, S. N., Nenes, A., and Mihalopoulos, N.: Cloud condensation nuclei measurements in the marine boundary layer of the Eastern Mediterranean: CCN closure and droplet growth kinetics, Atmos. Chem. Phys, 9, 7053–7066, 2009.

Brenguier, J.-L. and Wood, R.: Observational strategies from the micro to meso scale, Clouds in the Perturbed Climate System: Their Relationship to Energy Balance, Atmospheric Dynamics, and Precipitation, pp. 487–510, 2009.

Broekhuizen, K., Chang, R.-W., Leaitch, W., Li, S.-M., and Abbatt, J.: Closure between measured and modeled cloud condensation nuclei (CCN) using size-resolved aerosol compositions in downtown Toronto, Atmos. Chem. Phys., 6, 2513–2524, 2006.

Cantrell, W., Shaw, G., Cass, G. R., Chowdhury, Z., Hughes, L. S., Prather, K. A., Guazzotti, S. A., and Coffee, K. R.: Closure between aerosol particles and cloud condensation nuclei at Kaashidhoo Climate Observatory, J. Geophys. Res., 106, 28 711–28 718, 2001.

Chuang, P.: Sensitivity of cloud condensation nuclei activation processes to kinetic parameters, J. Geophys. Res., 111, D09 201, doi:10.1029/2005JD006529, 2006.

Chung, S. H. and Seinfeld, J. H.: Global distribution and climate forcing of carbonaceous aerosols, J. Geophys. Res., 107, AAC–14, 2002.

Cressie, N., Calder, C. A., Clark, J. S., Hoef, J. M. V., and Wikle, C. K.: Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modeling, Ecological Applications, 19, 553–570, 2009.

Dinar, E., Mentel, T., and Rudich, Y.: The density of humic acids and humic like substances (HULIS) from fresh and aged wood burning and pollution aerosol particles, Atmos. Chem. Phys., 6, 5213–5224, 2006.

Djikaev, Y. S. and Ruckenstein, E.: Thermodynamics of Water Condensation on a Primary Marine Aerosol Coated by Surfactant Organic Molecules, The Journal of Physical Chemistry A, 118, 9879–9889, 2014.

Dusek, U., Frank, G., Hildebrandt, L., Curtius, J., Schneider, J., Walter, S., Chand, D., Drewnick, F., Hings, S., Jung, D., et al.: Size matters more than chemistry for cloud-nucleating ability of aerosol particles, Science, 312, 1375–1378, 2006.

Ekström, S., Nozière, B., Hultberg, M., Alsberg, T., Magnér, J., Nilsson, E. D., and Artaxo, P.: A possible role of ground-based microorganisms on cloud formation in the atmosphere, Biogeosciences, 7, 387–394, doi:10.5194/bg-7-387-2010, 2010.

Ervens, B., Feingold, G., and Kreidenweis, S. M.: Influence of water-soluble organic carbon on cloud drop number concentration, J. Geophys. Res., 110, doi:10.1029/2004JD005634, 2005.

Ervens, B., Cubison, M., Andrews, E., Feingold, G., Ogren, J., Jimenez, J., Quinn, P., Bates, T., Wang, J., Zhang, Q., et al.: CCN predictions using simplified assumptions of organic aerosol composition and mixing state: a synthesis from six different locations, Atmos. Chem. Phys., 10, 4795–4807, 2010.

Facchini, M. C., Mircea, M., Fuzzi, S., and Charlson, R. J.: Cloud albedo enhancement by surface-active organic solutes in growing droplets, Nature, 401, 257–259, 1999.

Facchini, M. C., Decesari, S., Mircea, M., Fuzzi, S., and Loglio, G.: Surface tension of atmospheric wet aerosol and cloud/fog droplets in relation to their organic carbon content and chemical composition, Atmos. Environ., 34, 4853–4857, 2000.

Fitzgerald, J. W.: Dependence of the supersaturation spectrum of CCN on aerosol size distribution and composition, Journal of the atmospheric sciences, 30, 628–634, 1973.

Fountoukis, C. and Nenes, A.: Continued development of a cloud droplet formation parameterization for global climate models, J. Geophys. Res., doi:10.1029/2004JD005591, 2005.

Garg, V. and Chaubey, I.: A computationally efficient inverse modelling approach of inherent optical properties for a remote sensing model, International Journal of Remote Sensing, 31, 4349–4371, 2010.

Ghan, S., Wang, M., Zhang, S., Ferrachat, S., Gettelman, A., Griesfeller, J., Kipling, Z., Lohmann, U., Morrison, H., Neubauer, D., et al.: Challenges in constraining anthropogenic aerosol effects on cloud radiative forcing using present-day spatiotemporal variability, Proceedings of the National Academy of Sciences, p. 201514036, 2016.

Häkkinen, S., Äijälä, M., Lehtipalo, K., Junninen, H., Backman, J., Virkkula, A., Nieminen, T., Vestenius, M., Hakola, H., Ehn, M., et al.: Long-term volatility measurements of submicron atmospheric aerosol in Hyytiälä, Finland, Atmos. Chem. Phys., 12, 10 771–10 786, 2012.

Heintzenberg, J., Covert, D., and Van Dingenen, R.: Size distribution and chemical composition of marine aerosols: a compilation and review, Tellus B, 52, 1104–1122, doi:10.1034/j.1600-0889.2000.00136.x, 2000.

Henning, S., Rosenørn, T., D'Anna, B., Gola, A., Svenningsson, B., and Bilde, M.: Cloud droplet activation and surface tension of mixtures of slightly soluble organics and inorganic salt, Atmos. Chem. Phys., 5, 575–582, 2005.

IPCC: Summary for Policymakers, book section SPM, p. 1–30, Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, doi:10.1017/CBO9781107415324.004, www.climatechange2013.org, 2013.

Jacobson, M., Hansson, H., Noone, K., and Charlson, R.: Organic atmospheric aerosols: Review and state of the science, Rev. Geophys., 38, 267–294, 2000.

Jimenez, J. L., Canagaratna, M. R., Donahue, N. M., Prevot, A. S. H., Zhang, Q., Kroll, J. H., DeCarlo, P. F., Allan, J. D., Coe, H., Ng, N. L., Aiken, A. C., Docherty, K. S., Ulbrich, I. M., Grieshop, A. P., Robinson, A. L., Duplissy, J., Smith, J. D., Wilson, K. R., Lanz, V. A., Hueglin, C., Sun, Y. L., Tian, J., Laaksonen, A., Raatikainen, T., Rautiainen, J., Vaattovaara, P., Ehn, M., Kulmala, M., Tomlinson, J. M., Collins, D. R., Cubison, M. J., E., Dunlea, J., Huffman, J. A., Onasch, T. B., Alfarra, M. R., Williams, P. I., Bower, K., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Salcedo, D., Cottrell, L., Griffin, R., Takami, A., Miyoshi, T., Hatakeyama, S., Shimono, A., Sun, J. Y., Zhang, Y. M., Dzepina, K., Kimmel, J. R., Sueper, D., Jayne, J. T., Herndon, S. C., Trimborn, A. M., Williams, L. R., Wood, E. C., Middlebrook, A. M., Kolb, C. E., Baltensperger, U., and Worsnop, D. R.: Evolution of organic aerosols in the atmosphere, Science, doi:10.1126/science.1180353, 2009.

Kanakidou, M., Seinfeld, J., Pandis, S., Barnes, I., Dentener, F., Facchini, M., Dingenen, R. V., Ervens, B., Nenes, A., Nielsen, C., et al.: Organic aerosol and global climate modelling: a review, Atmos. Chem. Phys., doi:10.5194/acp-5-1053-2005, 2005.

Köhler, H.: The nucleus in and the growth of hygroscopic droplets, Trans. Faraday Soc., doi:10.1039/TF9363201152, 1936.

Kokkola, H., Vesterinen, M., Anttila, T., Laaksonen, A., and Lehtinen, K.: Technical note: Analytical formulae for the critical supersaturations and droplet diameters of CCN containing insoluble material, Atmos. Chem. Phys., 8, 1985–1988, 2008.

Kreidenweis, S. M., Koehler, K., DeMott, P. J., Prenni, A. J., Carrico, C., and Ervens, B.: Water activity and activation diameters from hygroscopicity data - Part I: Theory and application to inorganic salts, Atmos. Chem. Phys., 5, 1357–1370, doi:10.5194/acp-5-1357-2005, http://www.atmos-chem-phys.net/5/1357/2005/, 2005.

Laaksonen, A., Korhonen, P., Kulmala, M., and Charlson, R. J.: Modification of the Köhler equation to include soluble trace gases and slightly soluble substances, Journal of the atmospheric sciences, 55, 853–862, 1997.

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, doi:10.1029/2004JD004596, 2004.

Lee, L., Pringle, K., Reddington, C., Mann, G., Stier, P., Spracklen, D., Pierce, J., and Carslaw, K.: The magnitude and causes of uncertainty in global model simulations of cloud condensation nuclei, Atmos. Chem. Phys., 13, 8879–8914, doi:10.5194/acp-13-8879-2013, 2013.

Lesins, G., Chylek, P., and Lohmann, U.: A study of internal and external mixing scenarios and its effect on aerosol optical properties and direct radiative forcing, Journal of Geophysical Research: Atmospheres, 107, 2002.

Li, Z. and Lu, B. C.-Y.: Surface tension of aqueous electrolyte solutions at high concentrations—representation and prediction, Chemical engineering science, 56, 2879–2888, 2001.

Li, Z., Williams, A. L., and Rood, M. J.: Influence of soluble surfactant properties on the activation of aerosol particles containing inorganic solute, Journal of the atmospheric sciences, 55, 1859–1866, 1998.

Lienhard, D., Huisman, A., Krieger, U., Rudich, Y., Marcolli, C., Luo, B., Bones, D., Reid, J., Lambe, A., Canagaratna, M., et al.: Viscous organic aerosol particles in the upper troposphere: diffusivity-controlled water uptake and ice nucleation?, Atmos. Chem. Phys. Discussions, 15, 24 473–24 511, 2015.

Lohmann, U., Feichter, J., Penner, J., and Leaitch, R.: Indirect effect of sulfate and carbonaceous aerosols- A mechanistic treatment, J. Geophys. Res., 105, 12 193–12 206, 2000.

Lu, M.-L., Sorooshian, A., Jonsson, H. H., Feingold, G., Flagan, R. C., and Seinfeld, J. H.: Marine stratocumulus aerosol-cloud relationships in the MASE-II experiment: Precipitation susceptibility in eastern Pacific marine stratocumulus, Journal of Geophysical Research: Atmospheres, 114, 2009.

Martin, M., Chang, R.-W., Sierau, B., Sjogren, S., Swietlicki, E., Abbatt, J., Leck, C., and Lohmann, U.: Cloud condensation nuclei closure study on summer arctic aerosol, Atmos. Chem. Phys., 11, 11 335–11 350, 2011.

McCormick, R. and Ludwig, J. H.: Climate Modification by Atmospheric Aerosols, Science, doi:10.1126/science.156.3780.1358, 1967.

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

Mircea, M., Facchini, M. C., Decesari, S., Fuzzi, S., and Charlson, R. J.: The influence of the organic aerosol component on CCN supersaturation spectra for different aerosol types, Tellus B, 54, 74–81, 2002.

Murphy, B. N., Julin, J., Riipinen, I., and Ekman, A. M. L.: Organic aerosol processing in tropical deep convective clouds: Development of a new model (CRM-ORG) and implications for sources of particle number, J. Geophys. Res., 120, 10,441–10,464, doi:10.1002/2015JD023551, http://dx.doi.org/10.1002/2015JD023551, 2015JD023551, 2015.

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, 29–1, 2002.

Nozière, B., Baduel, C., and Jaffrezo, J.-L.: The dynamic surface tension of atmospheric aerosol surfactants reveals new aspects of cloud activation, Nature communications, 5, 3335, doi:10.1038/ncomms4335, 2014.

O'Dowd, C. D., Facchini, M. C., Cavalli, F., Ceburnis, D., Mircea, M., Decesari, S., Fuzzi, S., Yoon, Y. J., and Putaud, J.-P.: Biogenically driven organic contribution to marine aerosol, Nature, 431, 676–680, 2004.

Paramonov, M., Kerminen, V.-M., Gysel, M., Aalto, P., Andreae, M., Asmi, E., Baltensperger, U., Bougiatioti, A., Brus, D., Frank, G., et al.: A synthesis of cloud condensation nuclei counter (CCNC) measurements within the EUCAARI network, Atmos. Chem. Phys. Discussions, 15, 15 039–15 086, 2015.

Partridge, D. G., Vrugt, J. A., Tunved, P., Ekman, A. M. L., Gorea, D., and Sorooshian, A.: Inverse modeling of cloud-aerosol interactions: Part 1: Detailed response surface analysis, Atmos. Chem. Phys., 11, 7269–7287, doi:10.5194/acp-11-7269-2011, 2011.

Partridge, D. G., Vrugt, J. A., Tunved, P., Ekman, A. M. L., Struthers, H., and Sorooshian, A.: Inverse modelling of cloud-aerosol interactions: Part 2: Sensitivity tests on liquid phase clouds using a Markov chain Monte Carlo based simulation approach, Atmos. Chem. Phys., 12, 2823–2847, doi:10.5194/acp-12-2823-2012, 2012.

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

Petters, M. and Kreidenweis, S.: A single parameter representation of hygroscopic growth and cloud condensation nucleus activity, Atmos. Chem. Phys., 7, 1961–1971, 2007.

Pollacco, J. and Angulo-Jaramilo, R.: A Linking Test that investigates the feasibility of inverse modelling: application to a simple rainfall interception model for Mt Gambier, southeast South Australia, Hydrological processes, 23, 2023–2032, 2009.

Poulain, L., Spindler, G., Birmili, W., Plass-Dülmer, C., Wiedensohler, A., and Herrmann, H.: Seasonal and diurnal variations of particulate nitrate and organic matter at the IfT research station Melpitz, Atmos. Chem. Phys., 11, 12 579–12 599, 2011.

Prisle, N., Asmi, A., Topping, D., Partanen, A.-I., Romakkaniemi, S., Dal Maso, M., Kulmala, M., Laaksonen, A., Lehtinen, K., McFiggans, G., et al.: Surfactant effects in global simulations of cloud droplet activation, Geophys. Res. Lett., 39, L05 802, doi:10.1029/2011GL050467, 2012.

Prisle, N. L., Raatikainen, T., Laaksonen, A., and Bilde, M.: Surfactants in cloud droplet activation: mixed organic-inorganic particles, Atmos. Chem. Phys., 10, 5663–5683, doi:10.5194/acp-10-5663-2010, 2010.

Putaud, J.-P., Dingenen, R. V., Dell'Acqua, A., Raes, F., Matta, E., Decesari, S., Facchini, M., and Fuzzi, S.: Size-segregated aerosol mass closure and chemical composition in Monte Cimone (I) during MINATROC, Atmos. Chem. Phys., 4, 889–902, doi:10.5194/acp-4-889-2004, 2004.

Quinn, P., Bates, T., Coffman, D., and Covert, D.: Influence of particle size and chemistry on the cloud nucleating properties of aerosols, Atmos. Chem. Phys., 8, 1029–1042, 2008.

Riipinen, I., Rastak, N., and Pandis, S.: Connecting the solubility and CCN activation of complex organic aerosols: a theoretical study using the Solubility Basis Set (SBS), Atmos. Chem. Phys. Discussions, 14, 28 523–28 569, 2015.

Roberts, G. C., Artaxo, P., Zhou, J., Swietlicki, E., and Andreae, M. O.: Sensitivity of CCN spectra on chemical and physical properties of aerosol: A case study from the Amazon Basin, J. Geophys. Res., doi:10.1029/2001JD000583, 2002.

Saxena, P. and Hildemann, L. M.: Water-soluble organics in atmospheric particles: A critical review of the literature and application of thermodynamics to identify candidate compounds, Journal of atmospheric chemistry, 24, 57–109, 1996.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric chemistry and physics: from air pollution to climate change, John Wiley & Sons, 2012.

Shulman, M. L., Jacobson, M. C., Carlson, R. J., Synovec, R. E., and Young, T. E.: Dissolution behavior and surface tension effects of organic compounds in nucleating cloud droplets, Geophys. Res. Lett., 23, 277–280, 1996.

Sorjamaa, R., Svenningsson, B., Raatikainen, T., Henning, S., Bilde, M., and Laaksonen, A.: The role of surfactants in Köhler theory reconsidered, Atmos. Chem. Phys., 4, doi:10.5194/acp-4-2107-2004, 2004.

Szyszkowksi, B.: Experimentelle Studien über kapillare Eigenschaften der wässrigen Lösungen von Fettsäuren, Z. Physik. Chem., 64, 385–414, 1908.

Taraniuk, I., Graber, E. R., Kostinski, A., and Rudich, Y.: Surfactant properties of atmospheric and model humic-like substances (HULIS), Geophys. Res. Lett., 34, 2007.

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, Journal of Climate, 20, 1239–1254, 2007.

Toorman, A., Wierenga, P., and Hills, R.: Parameter estimation of hydraulic properties from one-step outflow data, Water Resources Research, 28, 3021–3028, 1992.

Topping, D.: An analytical solution to calculate bulk mole fractions for any number of components in aerosol droplets after considering partitioning to a surface layer, Geoscientific Model Development, 3, 635–642, doi:10.5194/gmd-3-635-2010, 2010.

Topping, D. and McFiggans, G.: Tight coupling of particle size, number and composition in atmospheric cloud droplet activation, Atmos. Chem. Phys., 12, 3253–3260, 2012.

Topping, D., Connolly, P., and McFiggans, G.: Cloud droplet number enhanced by co-condensation of organic vapours, Nature Geoscience, 6, 443–446, 2013.

Tsigaridis, K., Krol, M., Dentener, F., Balkanski, Y., Lathiere, J., Metzger, S., Hauglustaine, D., and Kanakidou, M.: Change in global aerosol composition since preindustrial times, Atmospheric Chemistry and Physics, 6, 5143–5162, 2006.

Tunved, P., Nilsson, E., Hansson, H.-C., Ström, 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., 110, 2005.

Twomey, S.: Pollution and the planetary albedo, Atmos. Environ., 8, 1251–1256, 1974.

Vrugt, J., Bouten, W., and Weerts, A.: Information content of data for identifying soil hydraulic parameters from outflow experiments, Soil Science Society of America Journal, 65, 19–27, 2001.

Vrugt, J. A., Schoups, G., Hopmans, J. W., Young, C., Wallender, W. W., Harter, T., and Bouten, W.: Inverse modeling of large-scale spatially distributed vadose zone properties using global optimization, Water Resources Research, 40, 2004.

Vrugt, J. A., Nualláin, B. Ó., Robinson, B. A., Bouten, W., Dekker, S. C., and Sloot, P. M.: Application of parallel computing to stochastic parameter estimation in environmental models, Computers & Geosciences, 32, 1139–1155, 2006.

Šimůnek, J., van Genuchten, M. T., and Wendroth, O.: Parameter estimation analysis of the evaporation method for determining soil hydraulic properties, Soil Science Society of America Journal, 62, 894–905, 1998.

Warren, S. G., Hahn, C. J., London, J., Chervin, R. M., and Jenne, R. L.: Global distribution of total cloud cover and cloud type amounts over land, NCAR Technical Note TN-273+STR, 29 pp, 1986.

Warren, S. G., Hahn, C. J., London, J., Chervin, R. M., and Jenne, R. L.: Global distribution of total cloud cover and cloud type amounts over the ocean, NCAR Technical Note TN-317+STR, 42 pp, 1988.

West, R., Stier, P., Jones, A., Johnson, C., Mann, G., Bellouin, N., Partridge, D., and Kipling, Z.: The importance of vertical velocity variability for estimates of the indirect aerosol effects, Atmos. Chem. Phys., 14, 6369–6393, 2014.

Wex, H., Hennig, T., Salma, I., Ocskay, R., Kiselev, A., Henning, S., Massling, A., Wiedensohler, A., and Stratmann, F.: Hygroscopic growth and measured and modeled critical super-saturations of an atmospheric HULIS sample, Geophys. Res. Lett., 34, 2007.

Wex, H., Stratmann, F., Topping, D., and McFiggans, G.: The Kelvin versus the Raoult term in the Köhler equation, Journal of the Atmospheric Sciences, 65, 4004–4016, 2008.

Wikle, C. K., Milliff, R. F., Herbei, R., Leeds, W. B., et al.: Modern statistical methods in oceanography: A hierarchical perspective, Statistical Science, 28, 466–486, 2013.

Yum, S. S. and Hudson, J. G.: Maritime/continental microphysical contrasts in stratus, Tellus B, 54, 61–73, 2002.

Zhang, Q., Jimenez, J. L., Canagaratna, M. R., Allan, J. D., Coe, H., Ulbrich, I., Alfarra, M. R., Takami, A., Middlebrook, A. M., Sun, Y. L., Dzepina, K., Dunlea, E., Docherty, K., DeCarlo, P. F., Salcedo, D., Onasch, T., Jayne, J. T., Miyoshi, T., Shimono, A., Hatakeyama, S., Takegawa, N., Kondo, Y., Schneider, J., Drewnick, F., Borrmann, S., Weimer, S., Demerjian, K., Williams, P., Bower, K., Bahreini, R., Cottrell, L., Griffin, R. J., Rautiainen, J., Sun, J. Y and, Z. Y. M., and Worsnop, D. R.: Ubiquity and dominance of oxygenated species in organic aerosols in anthropogenically-influenced Northern Hemisphere midlatitudes, Geophys. Res. Lett., 34, doi:10.1029/2007GL029979, 2007.

Zhou, J., Swietlicki, E., Berg, O. H., Aalto, P. P., Hämeri, K., Nilsson, E. D., and Leck, C.: Hygroscopic properties of aerosol particles over the central Arctic Ocean during summer, J. Geophys. Res., 106, 32 111–32 123, 2001.

			Mass fractions		
Component	ρ [kgm^{-3}]	M [gmol^{-1}]	Marine average	Polluted continental	Rural continental
Model organic	1350	260	0.18	0.40	0.60
BC	2000	12	0.30	0.075	0.10
$(NH_4)_2SO_4$	1770	132	0.00	0.2625	0.15
NaCl	2160	58.44	0.52	0.00	0.00
NH_4NO_3	1720	80.55	0.00	0.2625	0.15

Table 1: Density, molecular weight and mass fraction of each aerosol component in all environments. The mass fractions included here are used to derive true parameter values for f_{insol} and α in table 2.

Environment	Marine Average			Polluted Continental			Rural Continental		
Parameter	Min	True	Max	Min	True	Max	Min	True	Max
ρ_{org} [kgm^{-3}]	750	1350	1630	750	1350	1630	750	1350	1630
M_{org} [gmol^{-1}]	105	260	730	105	260	730	105	260	730
Φ	0.75	1.0	1.0	0.75	1.0	1.0	0.75	1.0	1.0
σ [mNm^{-1}]	30.0	55.0	72.8	30.0	55.0	72.8	30.0	55.0	72.8
α	0.06	0.26	0.46	0.12	0.76	3.10	1.50	2.00	2.50
f_{insol}	0.10	0.30	0.50	0.03	0.075	0.12	0.05	0.10	0.15
K	26200	31071	35942	26200	31071	35942	26200	31071	35942
Γ [mmolm^{-2}]	0.0025	0.00255	0.0026	0.0025	0.00255	0.0026	0.0025	0.00255	0.0026
N_1[cm^{-3}]	-	265.00	-	-	4900.00	-	-	1010.00	-
σ_1	-	1.45	-	-	1.55	-	-	1.71	-
$\bar{r}_1$[nm]	-	21.00	-	-	33.00	-	-	23.70	-
N_2[cm^{-3}]	60.00	165.00	250.00	730.00	1200.00	1600.00	215.00	451.00	690.00
σ_2	1.40	1.50	1.60	1.50	1.55	1.62	1.40	1.58	1.75
$\bar{r}_2$[nm]	70.00	82.50	100.00	75.00	93.50	105.00	75.00	89.80	105.00

Table 2: True parameter values used for calibration data for all environments and their corresponding parameter ranges used for perturbations in the response surface analysis.

Parameter	Relative sensitivity	Linear interactions	Non-linear interactions
N_2	Very high	α, Φ	f_{insol}
σ_2	Medium	-	-
$\bar{r}_2$	High	-	-
α	High	N_2,Φ	ρ_{org}, f_{insol}
f_{insol}	High	-	N_2, ρ_{org}, α, Φ
K	Low	Γ	-
Γ	Low	K	-
Φ	High	N_2, α	f_{insol}
ρ_{org}	Medium	-	α, f_{insol}
M_{org}	Low	-	-
σ	Very high	σ_2	M_{org}, ρ_{org}, α, f_{insol}, Φ

Table 3: Summary of qualitative sensitivities and parameter interactions observed in response surfaces for all parameters used in the the complete partitioning scheme $a_w^p\sigma^{nf}$ for the marine environment. The surface tension σ for classical Köhler theory $a_w^{np}\sigma^f$ is also included at the bottom of the table.