We thank the reviewer for taking the time to review this paper. We have addressed their comments, and believe that they have led to a substantial improvement of the manuscript. Below, the reviewers comments are reproduced in green, our answers are in black.

My main concern is that the authors describe their approach as a "model". To my understanding a model is something that uses a theoretical approach with known physical relationships that is used to predict experimental results. These predictions are then compared to the experimental results and, based on this, model parameters can be adjusted. What has been done here is a parameterization. Different variables that are known (or assumed) to have an influence on CDNC are correlated to the observed CDNC, and an empirical parameterization is built from this. For example, in panel a) of Figure 5, the "N_act predicted" is simply the number of CCN (particles larger than 90 nm) that are linearly correlated to the observed CDNC (here called N_act). In section 5.1 the authors seem to realize it and use the wording "statistical relationships". No physical explanations (e.g. from Kohler theory) for the factors in equation (2) can be given. Thus, I would strongly suggest replacing "model" by "parameterization" in the whole manuscript.

As similar potential for confusion exists with the word "parameterisation", we have instead added "statistical" in many places throughout the abstract and text, so that it is always clear that this is not a physics based model. "statistical model" is a common and correct term for equations such as those used here. To maintain the readability of the paper, we have not preceded "model" with "statistical" in cases where there no potential for misunderstandings exists.

Specific comments

Page 15477, lines 22-25:
Removal of anthropogenically influenced data: How were these data identified? Particle number concentrations? CO levels? Please specify.
We have now added the following text to this part of the manuscript:
"In recent years, outdoor tourism activities around the JFJ have increased, resulting in more frequent local pollution events. Data that is likely affected by construction activities, snow groomer operation and other local anthropogenic influences (mainly cigarette smoke) have been removed from the data sets. As the JFJ is characterised as a background site, sudden, short-lived fluctuations in the size distribution can be interpreted as local pollution (Herrmann et al., 2015). Therefore the affected data were identified by visual inspection of the aerosol size distribution spectra."

Page 15479, lines 3 – 12:
Why are different definitions for cloud conditions used? The difference is justified with different SMPS operation conditions (simultaneously scanning total and interstitial vs sequentially scanning both inlets), but the cloud conditions do not depend on the SMPS scanning time.
Essentially, we had to apply a stricter definition of "cloudy" to the campaigns where the SMPS scanning time was longer (sequentially scanning), to ensure that the JFJ was not within patchy cloud, which may have led to cloud-free periods during the SMPS scanning time. We now state this explicitly in the text:
"This more stringent criterion was used to avoid the inclusion of cloud-free periods in the longer (12 minute) SMPS scanning time. On the other hand, using the criterion of Hammer et al. (2014), which was found to be adequate for excluding cloud-free periods during the 6 minute scan time, allowed the inclusion of more data from the 2010 and 2011 campaigns."

Page 15480, Line 15:
"The height of the JFJ above the cloud base (calculated from the total water content

This part now reads:

"The height of the JFJ above the cloud base is calculated by using the total water content and temperature measured at the JFJ, calculating, under the assumption of a moist adiabatic temperature lapse rate (6 K/km), the temperature (and therefore the distance below the JFJ) at which the partial pressure of water in the air mass decreases below the saturation vapour pressure. This approach is described in detail in Hammer et al. (2014), and implicitly assumes that a minimal amount of water is lost from the air mass via precipitation between the cloud base and the JFJ."

We have included this analysis and plot as suggested by the referee, in section 5.1, showing that the skill of the model to predict the observed number of droplets is not highly dependent on the size used as cut off criteria for the number of potential CCN in the range 70-100nm. However, as the model performs slightly better with a potential CCN cut off of 80nm, we adopt this value, and have re-calculated the model fit and adjusted the plots accordingly, throughout the manuscript.

References

Hammer, E., Bukowiecki, N., Gysel, M., Jurányi, Z., Hoyle, C. R., Vogt, R., Baltensperger, U., and Weingartner, E.: Investigation of the effective peak supersaturation for liquid-phase clouds at the high-alpine site Jungfraujoch, Switzerland (3580ma.s.l.), Atmos. Chem. Phys., 14, 1123–1139, doi:10.5194/acp-14-1123-2014, 2014.

Herrmann, E., Weingartner, E., Henne, S., Vuilleumier, L., Bukowiecki, N., Steinbacher, M., Conen, F., Collaud Coen, M., Hammer, E., Jurányi, Z., Baltensperger, U., and Gysel, M.: Analysis of long-term aerosol size distribution data from Jungfraujoch with emphasis on free tropospheric conditions, cloud influence, and air mass transport, Journal of Geophysical Research: Atmospheres, 120, 9459–9480, doi:10.1002/2015JD023660, http://dx.doi.org/10.1002/2015JD023660, 2015JD023660, 2015.

We thank the reviewer for taking the time to review this paper. We have addressed their comments, and believe that they have led to a substantial improvement of the manuscript. Below, the reviewers comments are reproduced in green, our answers are in black.

However, the authors present no evidence to justify these claims. To address this, I would recommend that the authors compare, if possible, their results to previous cloud droplet observations on free troposphere in order to see how the derived parameterization compares against other data sets. Second, a comparison should be made with numerical or semi-empirical, established CDNC parameterizations such as those developed by Abdul-Razzak and Ghan (2000), Nenes et al. (2003) and Kivekäs et al.(2008) by initializing these parameterizations with "typical" free tropospheric aerosol. This would give insight into the applicability and limitations of the derived parameterization. Also, such comparison is needed as one of the real advantages of the parameterization over some of these numerical schemes derived here is computational efficiency. In order to further strengthen the manuscript, the authors should also discuss what new their results bring anything else new on the table compared to previous cloud parameterizations.

Ongoing work with the JFJ data has shown that the air at the JFJ during summer is influenced by the boundary layer up to around 80% of the time. We have therefore removed the statements about the model being applicable to the free troposphere, and state rather that it is applicable to the remote troposphere, with boundary layer influence.

The reviewer has made an excellent suggestion regarding comparing our statistical model with a numerical model. To this end, we have taken the time to collaborate with Prof. Nenes, who has kindly simulated the number of cloud droplets in our data set with his parameterisation. A comparison is now presented in the results section, as well as in the discussions and conclusions. In summary, the parameterisation of Prof. Nenes performs similarly to the statistical model, which is a very good result when one considers that the statistical model has been fitted to the observational data set used. We intend to continue this collaboration by applying both models to other data sets, representative of different atmospheric regions. This will however be the topic of a future study.

2. The authors should discuss more extensively about some of the limitations of their approach. First, does the topography of the site limit the applicability of the derived parameterization? For example, is there any information on the formation mechanisms of the sampled clouds – are they formed orographically or via convective activity? If they are formed through former mechanism, how it would impact the applicability of the derived parameterization (in large scale models, for example, clouds are often assumed to be formed via convection)? Related to this, the derived parameterization contains a parameter that is specific to the measurement site: vertical distance of the site from the cloud base, H (see equation 2). To me, it seems that this contradicts with the authors claim on the general applicability of their statistical model, and therefore this issue should be discussed.

We have extended the section describing the measurement site (see below), pointing out that previous studies have found the JFJ to be representative of both convective/orographic and stratiform cloud situations; depending on conditions and wind direction. Further, the injection of boundary layer air into the free troposphere during summer does indeed appear to be enhanced by the topography. However, these injections of boundary layer air mostly

influence the number of aerosol, and the statistical model represents the number of cloud droplet residuals satisfactorily over a wide range of residual numbers.

In addition, we have performed a sensitivity analysis with the Nenes parameterisation, giving information on the range of conditions under which our statistical model should be valid. This analysis is included in the discussion section of the manuscript, and shows that the droplet activation at the JFJ is in an aerosol limited regime, which supports our findings that the number of droplets can be accurately modelled based on only the number of aerosol particles larger than 80nm.

"While measurements at JFJ often sample the free troposphere, in summer the air masses are mostly influenced by injections of boundary layer air, due to convective events (Lugauer et al., 1998; Nyeki et al., 1998) and frontal systems (Zellweger et al., 2003). On average during summer, a boundary layer influence is detected at the JFJ around 80% of the time, dropping to around 60% in spring or autumn, or lower than 40% in January (Herrmann et al., 2015). The latter study also showed that the large degree of boundary layer influence is partly due to the effect of the alpine topography on air flow."

The vertical distance from the cloud base is however not specific to the JFJ, as it is calculated for each data point, based on the water content of the air mass and the temperature at the JFJ. Assuming a moist adiabatic lapse rate within the cloud, the cloud base is then the altitude below the JFJ where the saturation vapour pressure of water is equal to the partial pressure of water. This approach is now clarified in the manuscript:

"The height of the JFJ above the cloud base was calculated by using the total water content and temperature measured at the JFJ, assuming a moist adiabatic temperature lapse rate (6 K km^{-1}) and thus calculating the temperature (and therefore the distance below the JFJ) at which the partial pressure of water in the air mass decreases below the saturation vapour pressure. This approach is described in more detail in Hammer et al. (2014), and implicitly assumes that a minimal amount of water is lost from the air mass via precipitation between the cloud base and the JFJ. The height of the JFJ above the cloud base was included as a predictor variable as it determines the amount of condensed water at the altitude of the measurements, and it is also related to the age of the cloud, during which scavenging or coagulation processes may occur."

Second, eq. 2 implies linear dependence of CDNC on the CCN concentration. Is it expected that such relationship holds on warm clouds formed on free tropospheric aerosols in general?
We acknowledge that there are restrictions on the linearity of the relationship between the number of CCN and the number of cloud droplets. However these limitations mostly apply to situations with either very high particle numbers, or very low updraft velocities. We have tested if such a dependence is expected in our measurements, using the Nenes parameterisation. This aspect of the work is now discussed thoroughly in the manuscript:

"A linear dependence of the number of cloud droplets on nCCN implies that there is not a strong competition for water vapour during most of the activation phase of cloud droplet formation. Whether or not this occurs depends on the CCN number, the slope of the CCN spectrum, vertical velocity, the degree of external mixing, the presence of giant CCN (sea-salt, dust) and temperature (e.g. Rissman et al., 2004; Reutter et al., 2009; Ghan et al., 1997; Morales Betancourt and Nenes, 2014). A good indicator of linearity is expressed by the partial sensitivity of the droplet number to the number of aerosol, $\partial Nd / \partial Na$ (also known as the Aerosol-Cloud Index, ACI) for a given set of aerosol and cloud formation conditions. The closer the ACI is to unity, the less competition effects are present, linearity applies and vice versa. The ACI can be calculated either numerically with a parcel model (Reutter et al., 2009) or with a parametrization adjoint (Rissman et al., 2004; Moore et al., 2013; Morales

Betancourt and Nenes, 2014). The latter is used here to establish the degree to which linearity holds for the conditions at the JFJ. The results of this calculation are shown in Fig. 15. In panel a, it can be seen that the ACI increases from near zero at low updraught velocities, to around 0.4 at updraught velocities of approximately 1 ms^{-1} and higher (note that the updraught velocities shown in Fig. 15 have been corrected by a factor of 0.25, as described in section 5.2). This suggests that the form of the relationship between the number of droplets and nCCN does not change at updraught velocities higher than approximately 1 ms^{-1}. Therefore while the updraught has only a small influence on the number of cloud droplets under these conditions, it does slightly influence the relationship between the number concentration of aerosol and the number of droplets. Panel b of Fig. 15 shows the sensitivity of the droplet number to nCCN as a function of nCCN. Here it can be seen that the sensitivity does not display any obvious trend with increasing nCCN, supporting our choice of a linear relationship between the number of droplets and nCCN.

These results correspond with previous studies. For example, Reutter et al. (2009) found the number of cloud droplets to be directly proportional to the particle number concentration when the ratio of updraught velocity to particle number concentration was high, but found under low ratios, the number of cloud droplets formed was only dependent on the updraught velocity. The lower limit of the regime where the number of cloud droplets depends on the number of particles was found to be an updraught to particle number concentration ratio of 10^{-3} ms^{-1} cm^3), which, for a CCN concentration of 800 cm^{-3}, requires a vertical wind speed of only 0.8 ms^{-1}. Examining Figs. 1 to 4, it can be seen that almost all of the north western wind cases, and most of the south eastern wind cases have vertical wind speeds higher than 1 ms^{-1} (if the wind speeds in Figs. 1 to 4 were corrected by a factor of 0.25, as was done for the microphysical modelling, 67% would still be above 1 ms^{-1}). Therefore, based on the study of Reutter et al. (2009), a direct dependence of the number of droplets on the number of potential CCN would be expected. The study of Partridge et al. (2012) showed that under relatively clean conditions, the details of the aerosol number size distribution determined the number of cloud droplets, however when the accumulation mode particle concentrations were above approximately 1000 cm^{-3}, the chemical composition of the particles played the major role in determining the number of cloud droplets. Partridge et al. (2012) also found that the importance of the particle chemistry increases relatively to that of the particle sizes at lower updraught velocities. Under conditions where the aerosol population is externally mixed, the number of cloud droplets formed may also not be directly dependent on the number of CCN, as changes in the relative abundance of particles with differing hygroscopicities will influence the formation of cloud droplets. Nevertheless Dusek et al. (2006) found that there was little change in the activation diameter of particles (less than 20nm) when comparing polluted and background air masses at a non-urban site. These studies support the idea that for cloud formation at remote sites such as the JFJ, with updraught velocities above approximately 1.0 ms^{-1}, and relatively low aerosol number concentrations, the number of cloud droplets formed should be dependent on the number and size of the aerosol present."
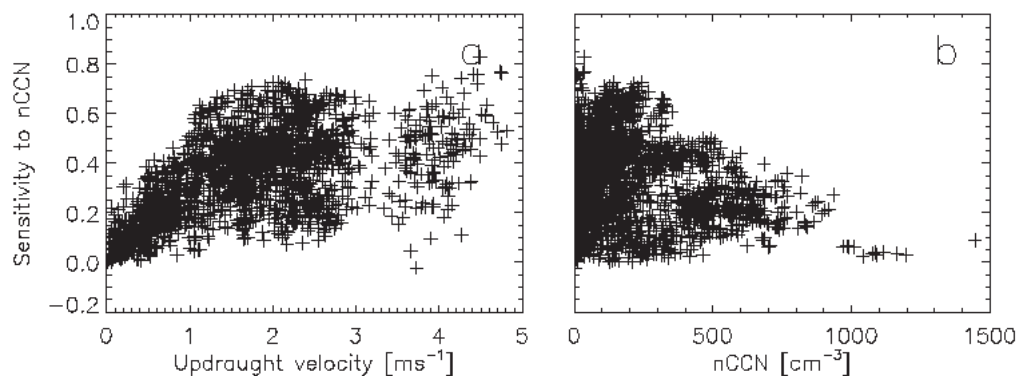
Figure 15 The sensitivity of the modelled droplet number to the updraught velocity (corrected by a factor of 0.25, following Hammer et al. 2015, panel a) and to the number of particles larger than 80nm (panel b).

Third, as stated in Section 3: "The present study utilises data collected during four summer campaigns, in 2002, 2004, 2010, and 2011". The considered data set is thus limited to the summer season. Do the characteristics of the cloud droplet activation observed at the site depend strongly on the season? And does this limit applicability of the derived parameterization?

We thank the referee for this comment.The number of potential CCN depends on the season, as boundary layer influence is much more frequent during the summer. However, the analysis shows that the presented statistical model performs similarly for low numbers of potential CCN (below 100 per cubic cm), as well as numbers up to around 800 per cubic cm. In this study we only include summer campaigns, and we filter out data where the plateau value of the activated fraction at higher aerosol diameters indicates that glaciation may have taken place. In partially glaciated clouds, the number of droplets is also influenced by the Wegener–Bergeron–Findeisen process, with some droplets evaporating over time. This is not accounted for in the statistical model. We now include a statement to this effect in the conclusions (and in the introduction):

"It should be noted that the statistical model is based only on data collected during summer campaigns, and that periods with partially or fully glaciated clouds have been excluded from the data set (as described in Sect. 4.1). During such periods the number of activated aerosol is also influenced by water uptake by ice particles, changing the relationship between the number of CCN and the number of cloud droplets. The statistical model is thus considered valid only for liquid clouds."

3. Usage of English in the manuscript should be improved considerably. Some of the issues with language are pointed out below, but I'd suggest that somebody with an excellent command of the language would review the usage of English in the manuscript.

The two native English speakers on the author list have carefully revised the manuscript. We thank the reviewer for the detailed review and numerous corrections provided in the technical comments, which have been implemented as suggested.

Minor and technical comments
1. Introduction, line numbers 33-34. The studies of Abdul-Razzak and Ghan (2000) and Kivekäs et al. (2008) should be cited here. Also, Petters and Kreidenweis (2007) did notpresent a parameterization for CDNC but for the impact of the particle hygroscopicity to CCN activation. Finally, I'd move the content of lines 26-34, i.e. description various CDNC parameterizations, to the end of the Introduction following the discussion on the

We now also cite the two studies suggested by the referee. We note that we did not mean to imply that Petters and Kreidenweis presented a CDNC parameterisation, rather that they describe a parameterisation of the chemical composition and mixing state on the water uptake. We have removed the citation from this paragraph to avoid confusion (the paper is still cited elsewhere). The text of lines 26-34 has been shifted as suggested.

The sentence containing these citations now reads "Previous ground-based studies have investigated statistical relationships between cloud droplet or CCN number concentration, aerosol properties and environmental variables", rather than just referring to "CCN concentration, aerosol properties and environmental variables…"

The paragraphs have been re-arranged accordingly

Done

This now reads "..surface active compounds may influence surface tension and thus the activation of aerosol particles to form cloud droplets"

Done

See answer to point 8.

The definition has been improved and now reads:
"Without entrainment, in theory all particles above a particular size will be activated during cloud formation, if the aerosol is internally mixed (as is generally the case at remote sites such as the JFJ). This size is known as the activation diameter, and depends on the peak supersaturation reached within the air parcel. The activation diameter of the aerosol was calculated for each measurement time, following Hammer et al. (2014). In atmospheric measurements, the fraction of activated particles increases between approximately 0 and 1 over a small range of diameters, rather than making a sharp transition at a particular diameter. Therefore the activation diameter is defined as that at which half the particles are activated and half are unactivated."

This second definition has been edited and shifted together with the text in the previous comment (8).

We have added the following clarification:
"..here, a potential CCN is considered to be an aerosol particle that may act as a CCN when subjected to supersaturations with respect to liquid water "

We have split this sentence into three sentences.

A detailed description has been added. The question mark was a typographical error which has been removed accordingly.

We believe that we need to motivate why we performed these additional calculations, but have now changed the end of the sentence to say "…to identify any features in the data which may be particular to these subsets."

This now reads "…during which not all particles larger than 80nm are activated…."

This now reads "….the fraction of particles that are activated appears to be lower….."

Done.

The contribution of [CO] to the explained variance is low, nevertheless piecewise analyses indicates [CO] as explanatory variable at the 99% level.
CO is frequently used as a tracer of anthropogenic emissions, therefore the negative correlation of [CO] with CNDC may indicate that more recent anthropogenic influence on an air mass implies lower particle hygroscopicity, and thus lower numbers of activated aerosol. Conversely higher levels of O_3 may indicate more photochemically processed air, and thus more hygroscopic aerosol.

Although slight differences in performance occur, the statistical model does not perform significantly worse in the SE case than in the NW case. In the revised version of the figures (using an 80nm cut off as the CCN definition; this change was made based on an analysis suggested by the other reviewer), the performance in the NW and SE cases is even more similar . The slightly poorer performance (error values) in the SE case stems mainly from SE

data from 2002. In these data, the number of cloud residuals is actually often equal to, or slightly higher than, the number of aerosol larger than 80nm, forcing and underestimate of the number of activated aerosol. This is because compared to other years, the aerosol size distribution is shifted towards smaller sizes. The main source of the larger errors is from times when the number of activated aerosol was high (above about 600 cm^{-3}). As can be seen from figure 4, most of the SE data in 2002 with high cloud residual numbers was collected over a few days (day 3 to 6), during a period with unusually high CO levels. This, combined with the number of aerosol, may imply an unusually high influence of the BL, possibly an injection of aerosol which has not yet undergone as much secondary growth as during other time periods.

19. Line number 340. ": : :in order to see: : :", instead of ": : :to see: : :".
We did not add "in order" at this point.

20. Line number 348. Remove comma from ": : :cloud droplets, to a: : :"
Done

21. Line numbers 351-352, sentence "Therefore there appears to be no systematic bias introduced by considering both wind directions in the model together.". However, according to Fig. 5C the parameterization tends to under predict observed CDNC when CDNC >500 /cm3. Please explain.
In the revised version of the figures, very little difference in model performance appears between NW and SE wind cases. Also the drop in model skill at CDNC above 500 has improved.

22. Line number 356. What does "compact correlation" mean?
We now write that the modelled and measured values are well correlated and we quantify this statement by providing the R values.

23. Line number 360. ": : : to which: : :" instead of ": : :by which: : :".
Done
24. Line numbers 369. "(Fig. 8)."
Fixed

25. Line numbers 370. See comment 22 above.
Changed to "good correlation"

26. Line numbers 370-373, sentence starting with "In this case: : :". I do not understand how good performance of the tuned parameterization against data from 2002 campaign implies that the underestimation of CDNC is not due to saturation effect. Please clarify.
In the new figures, no significant underestimation at high droplet numbers occurs; the discussion has been adapted accordingly.

27. Line number 378. What does "without replacement" mean in this context?
Sampling without replacement means that within a set of size N (in our case N=100 x 4 = 400) randomly sampled data points each data point is allowed to be drawn only once. Sampling without replacement has the advantage of avoiding sampling biases as sampling is truly random by construction. A potential disadvantage of sampling without replacement is that samples drawn might not be representative of the entire population (a disadvantage also emerging for samples with replacement), thus to overcome this problem in practice relatively large sample sizes are drawn and statistical investigations utilize multiple samples.
The text in the paper has been changed to the following:

"As a further way to assess the general applicability of the proposed linear model, we sample 100 data points at random (without replacement; i.e. individual data points are

allowed to be drawn only once to avoid a sampling bias e.g. Friedman 2015) from each year of data, and the R and error values were calculated with (i) the general model and (ii) the models fitted to each sampled set of 400 observations (i.e., 100 observations from each year) separately. To ensure for statistically robust results this analysis was performed for a set of 1000 random samples, and the results are summarized in Fig. 13."

28. Line numbers 400-404. The sentence starting with "Further: : :" should be split into two or more sentences.
We have split this sentence into two sentences.

29. Line numbers 405-408. Here it is concluded that ": : :temperature dependent influences of surface active compounds do not play a significant role in cloud droplet activation.". This is based on the lack of correlation between CDNC and cloud base temperature. I would argue that such effects (or their absence) cannot be inferred from ambient data series as the key parameters, such as exact chemical composition of CCN, that determine the cloud droplet activation process cannot be controlled in atmospheric measurements. Also, such effects could also be masked by changes in CDNC due to other temperature-related factors.
The text has been changed to "The cloud base temperature was not found to be significantly correlated with the cloud droplet number over the combined data set, therefore we find no evidence that temperature dependent influences of surface active compounds play a significant role in cloud droplet activation"

30. Line number 444. See comment no. 22 above.
Changed to "good correlation"

31. Table 2, page 21. The description of the table was hard to understand, please clarify.
The table has now been removed from the manuscript.

References

Friedman A. (2015): Statistics for Library and Information Services : A Primer for Using Open Source R Software for Accessibility and Visualization, Rowman and Littlefield, ISBN 978-1-4422-4993-6.
Hammer, E., Bukowiecki, N., Gysel, M., Jurányi, Z., Hoyle, C. R., Vogt, R., Baltensperger, U., and Weingartner, E.: Investigation of the effective peak supersaturation for liquid-phase clouds at the high-alpine site Jungfraujoch, Switzerland (3580ma.s.l.), Atmos. Chem. Phys., 14, 1123–1139, doi:10.5194/acp-14-1123-2014, 2014.

Lugauer, M., Baltensperger, U., Furger, M., Gaggeler, H., Jost, D., Schwikowski, M., and Wanner, H.: Aerosol transport to the high Alpine sites Jungfraujoch (3454masl) and Colle Gnifetti (4452masl), Tellus B, 50, 76–92, doi:10.1034/j.1600-0889.1998.00006.x, 1998.

Nyeki, S., Li, F., Weingartner, E., Streit, N., Colbeck, I., Gaggeler, H. W., and Baltensperger, U.: The background aerosol size distribution in the free troposphere: an analysis of the annual cycle at a high-alpine site, J. Geophys. Res.-Atmos., 103, 31749–31761, doi:10.1029/1998jd200029, 1998.

Zellweger, C., Hüglin, C., Klausen, J., Steinbacher, M., Vollmer, M., and Buchmann, B.: Inter-comparison of four different carbon monoxide measurement techniques and evaluation of the long-term carbon monoxide time series of Jungfraujoch, Atmos. Chem. Phys., 9, 3491–3503, doi:10.5194/acp-9-3491-2009, 2009.

Summary
11.03.2016 10:52:47

Differences exist between documents.

New Document:
acp-2015-337-discussions-typeset_source-version4
38 pages (1.99 MB)
11.03.2016 10:43:04
Used to display results.

Old Document:
acp-2015-337-discussions-typeset_source-version4_old_version
25 pages (2.86 MB)
11.03.2016 10:43:04

Get started: first change is on page 1.

No pages were deleted

How to read this report

Highlight indicates a change.
~~Deleted~~ indicates deleted content.
▲ indicates pages were changed.
⟷ indicates pages were moved.

Chemical and physical influences on aerosol activation in liquid clouds: a study based on observations from the Jungfraujoch, Switzerland

C. R. Hoyle[1,2], C. S. Webster[1,2,+], H. E. Rieder[3,4], A. Nenes[5,6,7,8], E. Hammer[1,*], E. Herrmann[1], M. Gysel[1], N. Bukowiecki[1], E. Weingartner[1], M. Steinbacher[9], and U. Baltensperger[1]

[1]Laboratory of Atmospheric Chemistry, Paul Scherrer Institute, Villigen, Switzerland
[2]WSL Institute for Snow and Avalanche Research SLF Davos, Switzerland
[3]Wegener Center for Climate and Global Change and IGAM/Department of Physics University of Graz, Austria
[4]Austrian Polar Research Institute, Vienna, Austria
[5]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA 30332
[6]School of Chemical and Biomolecular Engineering, Georgia Institute of Technology, Atlanta, GA 30332
[7]Institute of Chemical Engineering Sciences, Foundation for Research and Technology- Hellas, Patras, GR-26504, Greece
[8]Institute for Environmental Research and Sustainable Development, National Observatory of Athens, Palea Penteli, GR-15236, Greece
[9]Laboratory for Air Pollution/Environmental Technology, Empa – Swiss Federal Laboratories for Materials Science and Technology, Ueberlandstrasse 129,8600 Duebendorf, Switzerland
[*]Now at: Grolimund + Partner – Environmental Engineering, Thunstrasse 101a, 3006 Bern, Switzerland
[+]Now at: Department of Geography, Faculty of Engineering and Environment, Northumbria University, Newcastle Upon Tyne, UK

Correspondence to: C. R. Hoyle (christopher.hoyle@psi.ch)

Abstract. A simple statistical model to predict the number of aerosols which activate to form cloud droplets in warm clouds has been established, based on regression analysis of data from four summertime Cloud and Aerosol Characterisation Experiments (CLACE) at the high-altitude site Jungfraujoch (JFJ). It is shown that 79 % of the observed variance in droplet numbers can be represented by a model accounting only for the number of potential CCN (defined as number of particles larger than

5   80 nm in diameter), while the mean errors in the model representation may be reduced by the addition of further explanatory variables, such as the mixing ratios of O_3, CO and the height of the measurements above cloud base. The statistical model has a similar ability to represent the observed droplet numbers in each of the individual years, as well as for the two predominant local wind directions at the JFJ (northwest and southeast). Given the central European location of the JFJ, with air masses in summer being representative of the free troposphere with regular boundary layer in-mixing via convection, we expect that this

10   statistical model is applicable to warm clouds under conditions where droplet formation is aerosol-limited (i.e. at relatively high updraught velocities and/or relatively low aerosol number concentrations). A comparison between the statistical model and an established microphysical parametrization shows good agreement between the two, and supports the conclusion that cloud droplet formation at the JFJ is predominantly controlled by the number concentration of aerosol particles.

1 Introduction

Aerosols have a well-documented and pronounced influence on the microphysical and therefore radiative properties of clouds (e.g., Twomey, 1974, 1977; Albrecht, 1989; Hu and Stamnes, 1993). The properties of atmospheric aerosol particles thus have a strong potential to affect local and regional climates. However, the influence of aerosols on clouds remains the single largest uncertainty hampering the calculation of future climate scenarios (Boucher et al., 2013). To reduce this uncertainty, an improved understanding of the aerosol properties and environmental conditions that allow parts of the aerosol population to act as cloud condensation nuclei (CCN) and form cloud droplets is required.

Previous ground-based studies have investigated statistical relationships between cloud droplet or CCN number concentration, aerosol properties and environmental variables (e.g., Henning et al., 2002; Dusek et al., 2006; Verheggen et al., 2007; Jurányi et al., 2010, 2011; Anttila et al., 2012). Based on around 22 days of data from the Taunus Observatory in central Germany, Dusek et al. (2006) determined that the concentration of CCN (as measured at different supersaturations in a CCN counter) is largely dependent on the measured particle size distribution, with the CCN concentration increasing with increasing particle diameter, and chemical composition of the aerosol playing a secondary role.

Various studies have investigated the mechanisms through which the chemical composition of aerosol influences its water uptake and activation, and how this can be accounted for (e.g. Köhler, 1936; McFiggans et al., 2006; Petters and Kreidenweis, 2007). In addition, surface active compounds may influence surface tension and thus the activation of aerosol particles to form cloud droplets (Shulman et al., 1996; Shilling et al., 2007; King et al., 2009). Recently, it has been suggested that this may lead to a temperature influence on aerosol activation (Nenes et al., 2002; Christensen and Petters, 2012). Nevertheless, the works of, for example, Dusek et al. (2006) and Jurányi et al. (2010, 2011) suggest that the relatively small variations in chemical composition of aerosol in areas away from sources may play a smaller role in determining CCN activity of the aerosol than variations in the size distribution.

Examining one month of data from a remote site in northern Finland, Anttila et al. (2012) determined that the highest correlations with activated aerosol number occur with the number of available CCN, which was defined as the total number of particles greater than 100 nm in diameter, and that the number of droplets formed did not strongly depend on updraught velocity. A set of regimes where the number of cloud droplets formed depends on updraught velocities (at low ratios of updraught to aerosol number), and where the number of cloud droplets depends more on the number of aerosol (at high ratios of updraught to aerosol number) were described by Reutter et al. (2009), based on cloud parcel model studies. At the Jungfraujoch site, Henning et al. (2002) determined that in clouds with liquid water contents above $0.15\,\mathrm{g\,m^{-3}}$, aerosol particles larger than 100 nm in diameter were typically activated to form cloud droplets. Verheggen et al. (2007) investigated relationships between environmental variables and activated fraction, defined as the fraction of total particles, larger than 100 nm in diameter, that have been activated to form cloud droplets. The latter study based its analysis on one summer and two winter campaigns, and found that the activated fraction increased with increasing liquid water content and decreased with decreasing temperature below $0\,^{\circ}\mathrm{C}$, as clouds began to glaciate. Also using data from the Jungfraujoch site, Jurányi et al. (2010, 2011) found that with knowledge of the average chemical composition of aerosol, a very high degree of correlation could be found between the

number of activated aerosol predicted by the κ-Köhler approach (Petters and Kreidenweis, 2007) and the observed number of activated particles measured at different supersaturations in a CCN counter.

Although both Dusek et al. (2006) and Jurányi et al. (2010, 2011) have found that with a known aerosol size distribution, one can obtain good correlations between the predicted and observed number of droplets at a particular supersaturation in a CCN counter, the peak supersaturation reached in an air parcel is not generally a known quantity. It is also not possible to say how well the number of droplets predicted in this way corresponds with the number of droplets in a cloud which has formed some time ago. Although several studies exist where a good degree of closure was achieved between predicted and observed cloud droplet numbers (of the order of 20% difference betweeen calculated and observed droplet numbers, e.g. Fountoukis et al., 2007; Meskhidze et al., 2005; Conant et al., 2004), a simple method of predicting cloud droplet numbers based on easily quantifiable parameters would be useful.

It has long been recognised that the number and the size of aerosol particles strongly influences the number of CCN, and that at higher aerosol number concentrations, clouds will be composed of a greater number of droplets (Köhler, 1936; Fitzgerald and Spyers-Duran, 1973; Twomey, 1974, 1977). Several simple parametrisations of the number of cloud droplets as a function of the aerosol diameter and total aerosol number have been suggested for both continental and maritime locations, (Köhler, 1936; Raga and Jonas, 1993; Jones, 1994; Martin et al., 1994), mainly for stratus and stratocumulus clouds.

Subsequently, more advanced parametrizations were developed, allowing for the influence of the aerosol size distribution, updraught velocity, and the chemical composition and mixing state of the aerosol to be accounted for when calculating aerosol water uptake and activation to form cloud droplets (e.g., Abdul-Razzak and Ghan, 2000; Barahona and Nenes, 2007; Fountoukis and Nenes, 2005; Kumar et al., 2009; Nenes and Seinfeld, 2003). A parametrisation was also developed by Kivekäs et al. (2008), which predicts the number of cloud droplets using four parameters: the total sub-micron aerosol volume concentration, the number-to-volume aerosol concentration ratio, the soluble fraction of the particle volume and the air updraught velocity. Good agreement was found between the number of droplets predicted by this parametrization and observed droplet numbers in northern Finland.

In this study, data from four summer measurement campaigns carried out at the Jungfraujoch between 2002 and 2011 are used to develop simple statistical models of the relationship between the number of observed cloud droplets and various environmental factors, as well as the aerosol number size distribution, in liquid clouds. Using such an extensive data set collected over a period of nearly ten years allows the construction of relationships which are applicable to a wide range of conditions, although the statistical model developed here is only valid for liquid clouds. The results from the statistical models are compared to simulations using an established cloud droplet formation parametrization for use in climate model simulations of the aerosol indirect effect.

2 Measurement site

The Jungfraujoch (JFJ) high alpine measurement site is located at 3580 m a.s.l., atop an exposed crest in the Bernese Alps, Switzerland, and is accessible by train throughout the year. The site is engulfed in cloud approximately 40 % of the time

(Baltensperger et al., 1998; Nyeki et al., 1998) and local emissions are minimal with the exception of occasional construction activities. Aerosol measurements have been carried out at the JFJ since the early 1970s (Bukowiecki et al., 2016), with continuous measurements since 1986 (Baltensperger et al., 1991, 1997) and the site has been part of the Global Atmosphere Watch (GAW) programme since 1995. A review of the aerosol observations at the JFJ is provided by Bukowiecki et al. (2016).

The location of the station makes it suitable for continuous monitoring of the remote continental troposphere. The topography around the measurement site defines two predominant local wind directions, southeast or northwest. To the southeast, the Aletsch Glacier gradually slopes away from the JFJ at an approximate angle of 15°. In contrast, the northwestern side drops steeply at an average slope of approximately 46°. This difference in topography causes updraught velocities to be higher in air masses approaching the station from the northwest than from the southeast, with median peak supersaturations of around 0.41 % (representative of cumulus or orographic clouds) and 0.22 % (representative of shallow layer or stratiform clouds) being reached for the respective wind directions (Hammer et al., 2014; Lugauer et al., 1998). Therefore depending on conditions and wind direction, data gathered at the JFJ can be representative of convective or of stratiform type clouds.

The unique topography surrounding the JFJ site and the long-term measurements performed there provide substantial opportunity for not only investigating how relationships between environmental variables change between years, but also what effect the differing topography to the north and south has, through its influence on the vertical wind velocity. Furthermore, the composition of aerosols in air coming from the south is influenced by different source regions than air coming from the north. Peak supersaturation values, updraught velocity, aerosol hygroscopicity and cloud droplet number concentration were studied by Hammer et al. (2014), who found that all these quantities showed statistically significant differences between the two wind sectors. This work was extended by Hammer et al. (2015), who quantified the influence of updraught velocity and particle composition and concentration on peak supersaturation.

While measurements made at the JFJ often sample the free troposphere, in summer the air masses are mostly influenced by injections of boundary layer air, due to convective events (Lugauer et al., 1998; Nyeki et al., 1998) and frontal systems (Zellweger et al., 2003). On average during summer, a boundary layer influence is detected at the JFJ around 80% of the time, dropping to around 60% in spring or autumn, or lower than 40% in January (Herrmann et al., 2015). The latter study also showed that the large degree of boundary layer influence is partly due to the effect of the alpine topography on air flow.

The JFJ observatory is also one of 16 stations of the Swiss National Air Pollution Monitoring Network. As part of this operation, continuous in-situ observations of about 70 different trace gases are performed by Empa, the Swiss Federal Laboratories for Materials Science and Technology.

3   Data collection

Data used in this study were collected as part of the Cloud and Aerosol Characterisation Experiments (CLACE). The CLACE measurements have been conducted at the JFJ since 2000. They are a series of intensive winter and summer campaigns designed to investigate the chemical, physical and optical properties of aerosols as well as their interaction with clouds (Henning et al.,

2002; Verheggen et al., 2007; Sjogren et al., 2008; Kammermann et al., 2010; Jurányi et al., 2010, 2011; Hammer et al., 2014). The present study utilises data collected during four summer campaigns, in 2002, 2004, 2010, and 2011 (Table 1).

The following description refers to the basic experimental setup during all CLACE campaigns. The particles and hydrometeors were sampled via a total and an interstitial inlet which were installed through the roof of the laboratory (Hammer et al., 2014). The total inlet sampled all the particles that had a diameter of less than $40\,\mu m$, including the hydrometeors, at wind speeds up to $20\,m\,s^{-1}$ (Weingartner et al., 1999). The condensed water of the particles and hydrometeors was evaporated by heating up the top part of the inlet to approximately $25\,°C$ so that all particles were dried (and therefore residual aerosol particles contained in cloud droplets were set free) while reaching the instruments in the laboratory. The interstitial inlet only sampled particles smaller than 1 and $2\,\mu m$ diameter using a size discriminator of PM_1 (during CLACE2002) and PM_2 (during CLACE2004, CLACE2010 and CLACE2011) respectively. Thus, only non-activated particles (i.e. particles that did not act as CCN, and were thus not contained in cloud droplets) passed this inlet. The transition to laboratory temperatures (typically 20 to $30\,°C$) resulted in the drying of the particles at a relative humidity less than $10\,\%$. The difference between the number of aerosol sampled through the total inlet and the number sampled through the interstitial inlet gives the number of aerosol which were activated to form cloud droplets, n_{act}. It has been shown by Henning et al. (2002), in a comparison with Forward Scattering Spectrometer Probe (FSSP) droplet measurements, that this value can be used as a proxy for the number of cloud droplets. Therefore this is the approach that we adopt in the present study.

Downstream of the inlets, a Scanning Mobility Particle Sizer (SMPS) was used to measure the total and interstitial aerosol size distribution, respectively. The SMPS measured particles in the size range of 16 to $600\,nm$. One scan required $6\,min$. During CLACE2002 and CLACE2004, the SMPS was installed behind a pinch valve to switch between the two inlets after each scan (i.e. $6\,min$). The data in 2002 and 2004 is therefore at $12\,min$ resolution. For CLACE2010 and CLACE2011, two SMPS measured simultaneously behind each inlet so that a higher time resolution (approximately $6\,min$) could be achieved. Each SMPS consisted of a differential mobility analyser (DMA), a bipolar charger to obtain charge equilibrium (krypton source, ^{85}Kr) and a condensation particle counter (CPC) (Wiedensohler et al., 2012). During cloud-free periods, the interstitial and the total SMPS should measure the same aerosol number size distribution. For the campaigns where two SMPS measured simultaneously, the out-of-cloud particle size distribution showed differences of up to $10\,\%$ for particles with diameters between 20 and $600\,nm$ (Hammer et al., 2014). This is within the typical uncertainty for this type of measurements (Wiedensohler et al., 2012). To account for these differences between the two units, the interstitial number size distributions (for each campaign specific instrument) were corrected towards the total aerosol size distribution. A size and time dependent correction factor was determined by comparing the total and interstitial number size distributions during all cloud-free periods (Verheggen et al., 2007).

To monitor the cloud presence, the liquid water content (LWC) was measured using a particle volume monitor (PVM-100; Gerber, 1991), which measures the LWC by forward light scattering.

A measurement of the horizontal wind speed and direction was provided by the Rosemount Pitot tube anemometer, which is mounted on a $10\,m$ mast as part of the SwissMetNet network of MeteoSwiss. Likewise, temperature measured at the site as part of the SwissMetNet network was used.

In recent years, outdoor tourism activities around the JFJ have increased, resulting in more frequent local pollution events. Data that is likely affected by construction activities, snow groomer operation and other local anthropogenic influences (mainly cigarette smoke, Fröhlich et al., 2015) have been removed from the data sets. As the JFJ is characterised as a background site, sudden, short-lived fluctuations in the aerosol size distribution can be interpreted as local pollution (Herrmann et al., 2015).
5   Therefore the affected data were identified by visual inspection of the aerosol size distribution spectra.

In-situ trace gas measurements of O_3 and CO were conducted as part of the Swiss National Air Pollution Monitoring Network (NABEL). Measurements were recorded at 10 min intervals throughout all study periods, using a UV absorption technique for O_3 (Thermo Environmental Instrument, TEI49C), and non-dispersive IR absorption photometry (NDIR) for CO (Horiba APMA360, APMA370) (Gilge et al., 2010; Zellweger et al., 2009).

10   ## 4   Data analysis

4.1   Data processing

For years where two SMPS were operating simultaneously (CLACE2010, CLACE2011), n_{act}, as a function of dry particle diameter, could be calculated directly from the difference between the total and the interstitial particle number size distributions. For the remaining two years (CLACE2002 and CLACE2004), the SMPS was switched between the total and the interstitial
15   inlet. For these two years, the total measurement was taken to be the first measurement, with the interstitial measurement immediately following it used to calculate n_{act}. The two scans inside this 12 min period were assumed to represent the same atmospheric conditions.

In order to exclude cloud periods that were influenced by the entrainment of dry air, as well as to exclude mixed phase clouds, the fraction of activated particles was analysed as a function of particle diameter. Without entrainment, in theory all particles
20   above a particular size will be activated during cloud formation, if the aerosol is internally mixed (as is generally the case at remote sites such as the JFJ). This size is know as the activation diameter, and depends on the peak supersaturation reached within the air parcel. The activation diameter of the aerosol was calculated for each measurement time, following Hammer et al. (2014). In atmospheric measurements, the fraction of activated particles increases between approximately 0 and 1 over a small range of diameters, rather than making a sharp transition at a particular diameter. Therefore the activation diameter is
25   defined as that at which half the particles are activated and half are unactivated.

As described below, for the aged aerosol found at the JFJ, the critical diameter lies around 80–100 nm. Entrainment and mixing of air into the cloud will lead to non-activated particles larger than the activation diameter co-existing with activated particles and therefore the maximum activated fraction above the activation diameter will be less than one. Similarly, the lower water vapour pressure over ice particles in mixed phase clouds will lead to evaporation of droplets and deactivation of
30   aerosol, reducing the activated fraction above the activation diameter. A threshold of 0.9 was defined, and all measurements with maximum activated fractions of less than this threshold were assumed to be influenced by entrainment or partial glaciation of the cloud and thus excluded from the analysis.

6

The data were also filtered to remove any data points that were measured outside of clouds, in patchy cloud, or on the edges of clouds. This was achieved based on the measured LWC. For the campaigns that had two SMPS scanners operating simultaneously (CLACE2010 and CLACE2011), the criterion follows Hammer et al. (2014), where cloud was defined to be present when the 30th-percentile of the 10 s LWC values' distribution during one 6 min scan period was higher than 5 mg m^{-3}. For the other campaigns which had only one SMPS system operating (CLACE2002 and CLACE2004), creating a 12 min resolution data set, the criteria used was that of Henning et al. (2002) and Cozic et al. (2008), which defined cloudy conditions if the LWC was higher than 20 mg m^{-3} for more than 85 % of an hourly period. This more stringent criteria was used to avoid the inclusion of cloud-free periods in the longer (12 minute) SMPS scanning time. On the other hand, using the criteria of Hammer et al. (2014), which was found to be adequate for excluding cloud-free periods during the 6 minute scan time, allowed the inclusion of more data from the 2010 and 2011 campaigns.

Total water content (TWC) was calculated by adding measured LWC to calculated gas phase water (GPW), except during CLACE2010 where it could be determined directly from a dew point measurement in air sampled through the total inlet. In campaigns other than CLACE2010, such dew point measurements were not available and the GPW was calculated, using the ambient temperature, under the assumption that the in-cloud relative humidity was 100 %.

Data was classified according to wind direction (north and south), in order to determine if different factors influence the CCN quality depending on the origin of the aerosol particles.

For the purposes of this study, an estimate of the updraught velocity (w_{act}) at cloud base was calculated, similarly to Hammer et al. (2014), from the local topography and the horizontal wind speed measured at the JFJ (v_{JFJ}^h) using:

$$w_{act} = \tan(\alpha)v_{JFJ}^h, \tag{1}$$

where α is the inclination angle of the flow lines at the cloud base. These values were $\alpha = 46°$ for the northern terrain and $\alpha = 15°$ for the southern terrain (for further details see Hammer et al., 2014). This equation is based on the assumptions that the flow lines of the updraught strictly follow the terrain on either side of the JFJ research station, and that there is neither sideways convergence nor divergence of the flow lines between the cloud base and the JFJ.

4.2 Selection of predictor variables

Six different predictor variables either measured at the JFJ, or calculated for the cloud base, were included in the statistical analysis. These were height of the JFJ above cloud base, updraught velocity, number of available potential CCN particles (hereafter referred to as n_{CCN}, see definition below), air temperature at the cloud base, CO and O$_3$.

The height of the JFJ above the cloud base was calculated by using the total water content and temperature measured at the JFJ, assuming a moist adiabatic temperature lapse rate (6 K km^{-1}) and thus calculating the temperature (and therefore the distance below the JFJ) at which the partial pressure of water in the air mass decreased below the saturation vapour pressure. This approach is described in detail in Hammer et al. (2014), and implicitly assumes that a minimal amount of water is lost from the air mass via precipitation between the cloud base and the JFJ. The height of the JFJ above the cloud base was included

7

as a predictor variable as it determines the amount of condensed water at the altitude of the measurements, and it is also related to the age of the cloud, during which scavenging or coagulation processes may occur.

The updraught velocity, estimated as described in Sect. 4.1, was chosen as it is known to influence the peak supersaturation achieved during cloud formation, and therefore the activation diameter of the aerosol and the activated fraction of a particular aerosol size distribution.

The n_{CCN} is estimated from the measured aerosol size distributions. As described in Sect. 1, the aerosol number size distribution is known to play an important role in defining the number of cloud droplets formed, with larger particles more likely to be activated, and the smallest particles rarely playing a role in cloud formation. Therefore, it is necessary to choose a minimum diameter, above which a particle can be considered a potential CCN (here, a potential CCN is considered to be an aerosol particle that may act as a CCN when subjected to supersaturation with respect to liquid water). As described above, at aerosol number concentrations larger than approximately $100\,\mathrm{cm}^{-3}$, Henning et al. (2002) found that the activation diameter at the JFJ is around $100\,\mathrm{nm}$. Further, Hammer et al. (2014) reported that there is a systematic difference in the activation diameter for aerosol in air masses approaching the JFJ from the north ($87\,\mathrm{nm}$) and from the south ($106\,\mathrm{nm}$). Here we have chosen a diameter of $80\,\mathrm{nm}$ as the lower size bound defining potential CCN. The relatively low value was chosen so as not to exclude potentially important sizes of aerosols.

The air temperature at cloud base (calculated from the temperature at the JFJ) was chosen to account for any temperature dependent effects on water uptake to the aerosols which may influence activation. However, the cloud base temperature was found not to contribute significantly in the linear regression models for the years 2010 and 2011 (i.e., the years with most observational data). It was thus excluded by backward-elimination of explanatory variables for final model selection. Likewise, no significant relationship between air pressure and n_{act} was found.

Finally, the two chemical tracers CO and O_3 were included in the analysis to account for the history of the air parcels. While CO is a primary pollutant and O_3 is produced photochemically as a secondary pollutant from precursors such as volatile organic compounds and nitrogen oxides, both of these can act as tracers of anthropogenic emissions or of biomass burning events (e.g. Staudt et al., 2001; Liang et al., 2004; Yashiro et al., 2009; Zhang et al., 2006, 2009; Gilge et al., 2010), and therefore in this study they are used as indicators of the degree of influence of polluted air masses, in an attempt to determine if this has an important effect on particle activation at the JFJ. Ozone at the JFJ may be influenced by stratospheric intrusions, however a modelling study (Cui et al., 2009) has suggested that this is the case for less than $20\,\%$ of the year, making such events relatively rare.

4.3 Statistical analysis

In order to determine if and how environmental and chemical factors can be related to the number of cloud droplets (i.e. the number of activated aerosol, n_{act}), we chose a simple multiple linear regression model for the analysis. Multiple linear regression is a commonly used statistical method for explanatory and theory-testing purposes, thus it is appropriate to use in assessing how the environmental and chemical variables contribute to the prediction of n_{act} (Johnson et al., 2004; Tonidandel and LeBreton, 2011). It is likely that several of the predictor variables selected for this analysis will be cross-correlated, thus

traditional regression indices (p value, regression coefficients) will fail to appropriately partition the predictor variables into respective contribution to the overall R^2 of the model (Tonidandel and LeBreton, 2011). Nevertheless, active research in the statistical sciences has led to a set of tools for the assessment of the relative importance of individual covariates in linear regression models in the presence of correlated explanatory variables. A widely used approach, first proposed by Lindeman et al. (1980), hence referred to as LMG, but better known in the sequential additive version proposed by Kruskal (1987), allows assigning shares of "relative importance" to a set of regressors in a linear model (Grömping, 2007). Here we use the LMG method, in its implementation in the *relaimpo* package, developed by Grömping (2006) available for the scientific computing language R (R Core Team, 2014), to assess the relative importance of individual explanatory variables in a simple linear regression model for the cloud droplet numbers in warm tropospheric clouds.

Below we detail the LMG method and its application to our statistical model following Grömping (2006). Once the set of explanatory variables/regressors $(x_{i1}, ..., x_{ip})$ is defined, as in our analysis in Eq. (5), the multiple linear regression model is fitted and the regression coefficients for each explanatory variable $(\beta_k, k = 0, ..., p)$ included in the model are estimated by minimising the sum of squared unexplained parts. The coefficient of determination (R^2) can then be expressed using the fitted response values $(\hat{y}_i)$ and estimated coefficients $(\hat{\beta}_k)$ as the ratio between the model and total sum of squares (MSS and TSS, respectively) i.e., $R^2 = \frac{MSS}{TSS} = \frac{\sum_{i=1}^{n}(\widehat{y_i - \bar{y}})^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2}$. The LMG method decomposes the coefficient of determination into non-negative contributions that sum to the total R^2. First sequential (i.e., regressors are used in listed order e.g. as given in our model in Eq. 5) sums of squares (SSS) are derived via analysis of variance (ANOVA). These sequential sums of squares, for each regressor, sum to the MSS of the TSS. Next sequential R^2 contributions are derived by dividing SSS by TSS. These sequential R^2 contributions are then utilized in the LMG method. As the order of the explanatory variables in any regression models is a permutation of the available regressors $x_1, ..., x_p$, it can be denoted by the tuple of indices $r = (r_1, ..., r_p)$. The set of regressors entered in the model before regressor x_k in the order of r can then be denoted as $S_k(r)$. Thus the portion of R^2 allocated to explanatory variable x_k in the order r can be written as

$$seq R^2(\{x_k\} | S_k(r)) = R^2(\{x_k\} \cup S_k(r)) - R^2(S_k(r)). \tag{2}$$

Using equation 2 the metric LMG can be written as

$$LMG(x_k) = \frac{1}{p!} \sum_{permutation} seq R^2(\{x_k\} | r), \tag{3}$$

which can be further simplified to

$$LMG(x_k) = \frac{1}{p!} \sum_{S \subseteq \{x_1,...,x_p\}/\{x_k\}} n(S)! (p - n(S) - 1)! \, seq R^2(\{x_k\} | S) \tag{4}$$

as orders with the same $S_k(r)$ can be summarized into one summand (Grömping, 2006).

In the following we propose simple linear regression models developed based on 4 years of observations from the JFJ, Switzerland. Additionally the best performing regression model was run for subsets of the data corresponding to the different years, and wind directions, to identify any features in the data which were particular to these subsets. The aim of this analysis

was to determine whether a single statistical model can be constructed which will be generally applicable for the prediction of the number of cloud droplets for all years and wind directions.

5 Results

In total, 2399 data points were included in the analysis, with the majority being from 2010 (1087) and 2011 (896). Data was limited in 2002 (206 points) and 2004 (210 points) compared to that in 2010 and 2011 since there were more episodes of entrainment or partially glaciated clouds, where data was excluded from this analysis. The 2002 campaign was relatively short and the time resolution of the measurement data set was lower in 2004 and 2002 than in later years, as described above, yielding fewer data points. In Figs. 1 to 4, time series of the predictor variables are shown for each campaign. In these plots, it can be seen that the data sets include a wide range of conditions with respect to meteorology and air parcel composition. In the upper panels of the plots, n_{CCN} is plotted together with n_{act}. In 2011 and 2010 (Figs. 1 and 2) there are episodes of relatively high n_{CCN}, during which not all particles larger than 80 nm are activated, as shown by the lower n_{act} numbers. Additionally, the fraction of particles that are activated appears to be lower when the wind is from the southeast (red symbols in the bottom panel of the plots). In 2004 however (Fig. 3), n_{CCN} is generally fairly low, with, in a few cases, larger n_{act} than CCN, indicating that also particles below the chosen cut off diameter for potential CCN are being activated. In 2002 (Fig. 4), there is a broad range of n_{CCN} values, and activation appears to be high in almost all cases, regardless of wind direction or updraught velocity. In all years, the concentrations of CO and O_3 (second panel) appear to be fairly well correlated with each other, except around day 12 of the 2002 campaign (overall $R = 0.65$). There does not appear to be an appreciable link between wind direction and CO or O_3 concentration. The temperature range is similar for all the data sets, with temperatures generally between 270 and 280 K. An episode of warmer temperatures in the first half of the 2010 campaign corresponds with relatively high CO and O_3 values, and also higher aerosol number concentrations. The cooling after day 20 is accompanied by a marked reduction in n_{CCN}, and also an increase in the fraction of aerosol which are activated to form cloud droplets. As can be seen in the bottom panel of each plot, the updraught velocities are generally lower when the wind is from the southeast than when it is from the northwest, consistent with the findings of Hammer et al. (2014).

5.1 Statistical relationships for combined data

The modelled number of cloud droplets is plotted against the observed number (n_{act}), for a variety of statistical model formulations, in Fig. (5). In panel a, only n_{CCN} is used to predict the number of cloud droplets. Already here a good relationship is found, with a correlation (R) of 0.89, however the intercept in the model leads to an unphysical cut off at low modelled numbers. Including the updraught velocity improves the model slightly, while the R value remains the same, the root mean squared error (RMSE) reduces slightly from 59.7 to 58.1. Further improvements are found by including all five selected explanatory variables (panel c), and, in panel d, by using all variables as well as the log of the updraught velocity rather than the updraught velocity itself. The latter statistical model was found to provide the best representation of the observed number of droplets, with an R value of 0.91, RMSE of 54.2 and a mean error (ME) of 38.1.

The statistical model presented in panel d of Fig. 5 provides a simple and reasonably accurate way of predicting the number of cloud droplets formed based on only a few explanatory variables. The number of activated aerosol (considered equivalent to the number of droplets) predicted by this model is given by:

$$n_{\text{act}} = 0.57 n_{\text{CCN}} + 2.58 O_3 + 0.03 H - 1.0 \text{CO} + 28.48 \log(\omega) - 41.28, \tag{5}$$

where ω is the estimated updraught velocity at cloud base in m s^{-1}, CO and O_3 are mixing ratios in ppb, and H is the height of the JFJ above the cloud base in meters (H must be greater than zero).

The model considering only the number of CCN, as shown in panel a of Fig. 5, is:

$$n_{\text{act}} = 0.57 n_{\text{CCN}} + 43.27 \tag{6}$$

The same analysis was performed with reductions in the minimum size of aerosol considered to be CCN to 70, 90 and 100 nm (Fig. 6), however this did not improve the model skill in relation to the results obtained when counting only aerosol larger than 80 nm to determine n_{CCN}. In fact, there was little variation in the model skill when these different size criteria were used in the definition of potential CCN.

It should be noted that at very low n_{CCN}, the statistical model may return negative values for the number of droplets, which is obviously un-physical. However, this only applies to a very small number of points (16 of the 2399 points presented here), and thus does not compromise the general applicability of the proposed model.

5.2 Comparison with physically based parametrization

To put the results presented in Fig. 5 and 6 into the context of previous work, a state of the art cloud droplet formation parametrization was used to calculate the cloud droplet number for the same data points. Here we apply the sectional form of the cloud droplet formation model of Nenes and Seinfeld (2003) and Fountoukis and Nenes (2005), with the giant CCN correction as described by Barahona et al. (2010). In applying this parametrization, input data is required, describing the chemical composition, aerosol size distribution, updraught velocity, pressure and temperature. For the aerosol, the size distributions obtained by the SMPS are used (in original bin form), while an average aerosol hygroscopicity of 0.25 (corresponding to an aerosol mixture of roughly 42 % ammonium sulfate and 48 % insoluble aerosol) is assumed, which is similar to the hygroscopicity value found from 17 months of measurements at the JFJ by Jurányi et al. (2011), for particles with a critical dry diameter of around 80-100nm. The parametrization was also run for the overall median hygroscopicity value given by Jurányi et al. (2011) of 0.2, and a value of 0.3, to test the sensitivity of the results to small changes in assumed hygroscopicity, within the bounds of that which has been measured at the JFJ. Vertical velocity for the parametrization input was calculated using the method of Hammer et al. (2014), multiplied by an estimated correction factor of 0.25, following the suggestions of Hammer et al. (2015). Pressure and temperature at cloud base are also used, calculated in the same way as for the statistical model. A comparison of the predicted number of cloud droplets and the number of observed cloud residuals is shown in Fig. 7. The agreement between the modelled and observed data is excellent, with an R value of 0.86, RMSE of 67.2 and an ME of 42.8. The errors for Eq. (5), in panel d of Fig. 5 are only slightly lower than this. The R and error values for the microphysical

parametrization run for the three different hygroscopicity parameters are shown in table 2. There it can be seen that within the range of likely hygroscopicity values for the JFJ, there is little variation in the R values or errors from the model calculations. A slight decrease in the RMSE and ME is found when the hygroscopicity value is increased from 0.2 to 0.25 and 0.3.

5.3 Difference between wind directions

It was observed by Hammer et al. (2014) that the number and properties of aerosol in air parcels approaching the JFJ from the southeast was different from those in air approaching from the northwest. Further, they found that the activation diameter of particles differed considerably between the two wind directions. Therefore the total data set used here was divided according to wind direction, and the statistical model given by Eq. (5) was applied, to see if its ability to reproduce the observed number of droplets differed between the two wind directions. This comparison is shown in Fig. 8. The R values for the northwestern wind direction and the southeastern wind direction are the same (0.9), but the RMSE and ME are both substantially lower for the northwestern wind direction (RMSE of 49.3 vs. 67.4 and ME of 34.5 vs. 49.4). In the northwesterly case it can be seen that the model shifts from a slight overestimation of the observed number of cloud droplets to a slight underestimation, with the crossover occurring at about $150 \, \mathrm{drops \, cm^{-3}}$. The data in the southeastern case appears to closely follow the $1:1$ line. Therefore there appears to be no systematic bias introduced by considering both wind directions in the model together.

The results of the microphysical parametrization simulations, separated by wind direction, are shown in Fig. 9. Here it is seen that the microphysical parametrization is better able to represent the number of droplets in the northwestern wind case (R of 0.91), while in the southeastern case, the RMSE increases to 107, and the model underestimates the number of cloud droplets, particularly for numbers of residuals above about $300 \, \mathrm{cm^{-3}}$. This may be due to differences in turbulence and vertical wind velocity, between the northwestern and southeastern wind cases, which are not resolved by our vertical wind velocity estimation.

5.4 Difference between years

To determine how representative the model in Eq. (5) is for data from different years, the results were broken up into data for each year, shown in Fig. 10. For 2002, 2010 and 2011, the modelled data is well correlated with the observed number of droplets (R of between 0.89 and 0.95), however the slope varies between different years. While the data from 2011 lies along the $1:1$ line, the 2010 data seems to be composed of two different groups of points with different slopes, below and above approximately $300 \, \mathrm{drops \, cm^{-3}}$. It is not surprising that the R and error values are better for 2010 and 2011, as these years provide by far the most data points to which the model was fitted. The R for 2002 (0.95) was the highest of all years, however many of the data points are below the $1:1$ line and the RMSE was higher than for the other years (82.7). The data collected during 2004 is less well fit by the model (R of 0.76, RMSE of 47.2). However, as there were so few data points in 2004, and these were mostly at low droplet numbers, it is difficult to say if this is due to the data sampled, or if the conditions were fundamentally different during 2004.

Again, the results of the microphysical parametrization are shown, this time separated by year, in Fig. 11. The RMSE for the 2002 data is higher than for the statistical model (124 vs. 82.7), and the microphysical parametrization was found

to generally underestimate the number of cloud droplets in cases where there were more than approximately 200 residuals cm^{-3}. It is interesting to note that the statistical model also generally underestimates the observed values for 2002. For 2004, the microphysical parametrization represents the observational data better than the statistical model, with an R value of 0.82 compared with the 0.76 of the statistical model and an RMSE of 42.9 compared to 47.2 for the statistical model. For 2010 and 2011, both the statistical model and the microphysical parametrization represent the observed data well.

The differences between the years were also investigated by re-fitting the statistical model to each individual year of data (Fig. 12). Naturally, this results in higher values of R, and smaller errors. For example, in 2002 a good correlation is seen, with R of 0.96, and an RMSE of only 53.5. In 2002, it can also be seen that the model underestimation of points above $500\,\mathrm{drops}\,cm^{-3}$ seen in previous plots is not due to a saturation effect, as the observed droplet number can be predicted over the whole range of n_{CCN} with one set of parameters. The model representation of 2004 is improved when the model is fitted to only 2004 data, however the R value is still only 0.83, lower than for the other years. This appears to be related to the overall low range of n_{CCN} observed in 2004. Both 2010 and 2011 are well represented by models fitted specifically to these data.

As a further way to assess the general applicability of the proposed linear model, we sampled 100 data points at random (without replacement; i.e. individual data points are allowed to be drawn only once to avoid a sampling bias e.g. Friedman, 2015) from each year of data, and the R and error values were calculated with (i) the general model and (ii) the models fitted to each sampled set of 400 observations (i.e., 100 observations from each year) separately. To ensure for statistically robust results this analysis was performed for a set of 1000 random samples, and the results are summarized in Fig. 13. Due to the small number of data points in 2004 (210) and 2002 (206), the samples for these years did not differ greatly. In Fig. 13, it is apparent that the individually fitted models for the 1000 subsets perform slightly better than the simultaneously applied general model (as expected), however given the small differences in both R and error values between the individual and general models, illustrated by the overlap of the inner quartile ranges in both R and error values, the general model can be considered to be robust for the data set and applicable over a wide range of observed conditions.

6 Discussion

The analysis above shows that the number of cloud droplets can be reasonably well predicted by a single statistical model, containing the n_{CCN}, the log of the updraught velocity, the height above cloud base and the mixing ratios of CO and O_3. The contribution of each variable to the variance explained by Eq. (5) is shown in Fig. 14, along with error bars, denoting the range of the contributions of each variable in the random sampling analysis described in the previous section. The range of the parameters included in Fig. 14 is relatively small, indicating that the contribution to the explained variance is similar regardless of the sample taken from the data set.

By far the greatest contribution to the explained variance is from n_{CCN}, however including additional explanatory variables does improve the model with respect to absolute biases. The O_3 and CO mixing ratios contributed around 10 and $4\,\%$ respectively of predictive ability to the model, suggesting that for sites such as the JFJ, which are located relatively far from direct emissions sources, the chemical history, or source region of the air mass is not greatly relevant in predicting the activation

of aerosol to cloud droplets. Previously Jurányi et al. (2011) and Hammer et al. (2014) have found that the hygroscopicity parameter of aerosols observed at the JFJ is not highly variable. The results presented here also indicate that changes in aerosol properties, which would generally be correlated with CO or O_3 concentrations, are not large enough to substantially influence aerosol activation. The height above the cloud base, H, contributed a small amount (around 7 %) to the explained variance. This is likely due to the height above cloud base being a measure of the total amount of condensible water in the cloud, with greater condensible water generally leading to more droplets. The cloud base temperature was not found to be significantly correlated with the cloud droplet number over the combined data set, therefore we find no evidence that temperature dependent influences of surface active compounds play a significant role in cloud droplet activation. A previous study carried out at the JFJ, by Henning et al. (2002), found that when the number of potential CCN with diameter greater than 100 nm reduced below 100 cm^{-3}, the activation diameter shifted to smaller sizes, so that significant numbers of aerosol smaller than 100 nm began to activate. However, the ability of Eq. (5) to predict n_{act} does not deteriorate at low particle numbers, possibly because in this work, particles larger than 80 nm are considered potential CCN.

A linear dependence of the number of cloud droplets on n_{CCN} implies that there is not a strong competition for water vapour during most of the activation phase of cloud droplet formation. Whether or not this occurs depends on the CCN number, the slope of the CCN spectrum, vertical velocity, the degree of external mixing, the presence of giant CCN (sea-salt, dust) and temperature (e.g. Rissman et al., 2004; Reutter et al., 2009; Ghan et al., 1997; Morales Betancourt and Nenes, 2014). A good indicator of linearity is expressed by the partial sensitivity of the droplet number to the number of aerosol, ∂N_d / ∂N_a (also known as the Aerosol-Cloud Index, ACI) for a given set of aerosol and cloud formation conditions. The closer the ACI is to unity, the less competition effects are present, linearity applies and vice versa. The ACI can be calculated either numerically with a parcel model (Reutter et al., 2009) or with a parametrization adjoint (Rissman et al., 2004; Moore et al., 2013; Morales Betancourt and Nenes, 2014). The latter is used here to establish the degree to which linearity holds for the conditions at the JFJ. The results of this calculation are shown in Fig. 15. In panel a, it can be seen that the ACI increases from near zero at low updraught velocities, to around 0.4 at updraught velocities of approximately 1 ms^{-1} and higher (note that the updraught velocities shown in Fig. 15 have been corrected by a factor of 0.25, as described in section 5.2). This suggests that the form of the relationship between the number of droplets and n_{CCN} does not change at updraught velocities higher than approximately 1 ms^{-1}. Therefore while the updraught has only a small influence on the number of cloud droplets under these conditions, it does slightly influence the relationship between the number concentration of aerosol and the number of droplets. Panel b of Fig. 15 shows the sensitivity of the droplet number to n_{CCN} as a function of n_{CCN}. Here it can be seen that the sensitivity does not display any obvious trend with increasing n_{CCN}, supporting our choice of a linear relationship between the number of droplets and n_{CCN}.

These results correspond with previous studies. For example, Reutter et al. (2009) found the number of cloud droplets to be directly proportional to the particle number concentration when the ratio of updraught velocity to particle number concentration was high, but found under low ratios, the number of cloud droplets formed was only dependent on the updraught velocity. The lower limit of the regime where the number of cloud droplets depends on the number of particles was found to be an updraught to particle number concentration ratio of 10^{-3} ms^{-1} cm^3), which, for a CCN concentration of 800 cm^{-3}, requires a vertical

wind speed of only 0.8 ms^{-1}. Examining Figs. 1 to 4, it can be seen that almost all of the north western wind cases, and most of the south eastern wind cases have vertical wind speeds higher than 1 ms^{-1} (if the wind speeds in Figs. 1 to 4 were corrected by a factor of 0.25, as was done for the microphysical modelling, 67% would still be above 1 ms^{-1}). Therefore, based on the study of Reutter et al. (2009), a direct dependence of the number of droplets on the number of potential CCN would be

5   expected. The study of Partridge et al. (2012) showed that under relatively clean conditions, the details of the aerosol number size distribution determined the number of cloud droplets, however when the accumulation mode particle concentrations were above approximately 1000 cm^{-3}, the chemical composition of the particles played the major role in determining the number of cloud droplets. Partridge et al. (2012) also found that the importance of the particle chemistry increases relatively to that of the particle sizes at lower updraught velocities. Under conditions where the aerosol population is externally mixed, the number

10 of cloud droplets formed may also not be directly dependent on the number of CCN, as changes in the relative abundance of particles with differing hygroscopicities will influence the formation of cloud droplets. Nevertheless Dusek et al. (2006) found that there was little change in the activation diameter of particles (less than 20nm) when comparing polluted and background air masses at a non-urban site. These studies support the idea that for cloud formation at remote sites such as the JFJ, with updraught velocities above approximately 1.0 ms^{-1} and relatively low aerosol number concentrations, the number of cloud

15 droplets formed should be dependent on the number and size of the aerosol present.

Finally, the statistical models and the microphysical parametrization presented in this study are compared with two existing parametrisations, those of Jones (1994) and Martin et al. (1994), both of which used n_{CCN} to predict the number of cloud droplets which would be formed. The Martin et al. (1994) parametrisation is given by:

$$N_{\text{droplets}} = -2.10 \times 10^{-4} A^2 + 0.568A - 27.9 \tag{7}$$

20 where A is the number of aerosol in the size range 100 nm–3.0 μm in diameter. We use the version suggested for use in maritime air masses (their Eq. 12), as the version for continental air masses (their Eq. 13) produces a very poor representation of the number of observed droplets at the JFJ (not shown). This is possibly because the maritime parametrisation is more representative for air masses with relatively low aerosol number concentrations, as encountered at the JFJ. The maritime parametrisation is described as being valid over the range of aerosol number concentrations of 36 to 280 cm^{-3}.

25 The Jones (1994) parametrisation is derived from a combination of the continental and maritime parametrisations of Martin et al. (1994) and should therefore be valid over the range of aerosol number concentrations of 36 to 1500 cm^{-3}. It is given by:

$$N_{\text{droplets}} = 375(1 - \exp[-2.5 \times 10^{-3} \, A]) \tag{8}$$

The modelled cloud droplet number concentration is plotted against the measured values for Eqs. (5) and (6) as well as against the models of Martin et al. (1994) and Jones (1994) and the microphysical parametrization, in Fig. 16. Comparison of

30 Eqs. (5) and (6) with the other models considered shows that, although all five models provide a similar degree of explained variance (between 74% and 83%), error values are higher for the Jones (1994) and Martin et al. (1994) models. The micro-physical parametrization has a slightly lower R value than the other models, but has better error values than the Jones (1994) and Martin et al. (1994) models. While all five models show a good correlation between modelled and measured cloud droplet

numbers, the model of Martin et al. (1994) has a too shallow slope, resulting in a general underestimation of the observed values. Both the Jones (1994) and Martin et al. (1994) models have included a saturation effect at higher n_{CCN} which limits the number of cloud droplets formed, similarly to the effect described by Reutter et al. (2009). No such saturation effect is observed at the JFJ, however it cannot be ruled out that such an effect may occur at higher aerosol number concentrations than those presented here.

7   Conclusions

Using data from four summertime CLACE campaigns performed at the high altitude research station at the Jungfraujoch, we have shown that the number of cloud droplets formed in warm clouds can be rather accurately represented by a simple statistical model (Eq. 5), producing a similar degree of accuracy to that achieved with a microphysical parametrization. The majority of the variance in the observed droplet numbers is explained by the number of potential CCN, which is defined in this study as the total number of particles with a dry diameter greater than 80 nm. Using the number of potential CCN alone, 79 % of the observed variance is explained (Eq. 6). With the addition of further explanatory variables, such as CO and O_3 mixing ratios, and the height above cloud base, the RMSE and ME errors can be slightly reduced.

Although tuning the statistical model to each year of data separately produces slightly improved results, Eq. (5) represents the observed droplet numbers from the individual years quite adequately. Likewise, the model is applicable to data from both of the predominant wind directions at the JFJ, and although there is more variability in the models ability to predict the number of droplets formed during southeasterly wind conditions, there appears to be no substantial bias.

In contrast to previous studies where such models were constructed (e.g. Martin et al., 1994; Jones, 1994), no evidence for a saturation effect of high CCN numbers was observed, rather the number of droplets formed increased continually with n_{CCN}. Such a saturation effect is expected to occur at higher aerosol number concentrations, for example closer to aerosol sources or in more polluted environments.

It should be noted that the statistical model is based only on data collected during summer campaigns, and that periods with partially or fully glaciated clouds have been excluded from the data set (as described in Sect. 4.1). During such periods the number of activated aerosol is also influenced by water uptake by ice particles, changing the relationship between the number of CCN and the number of cloud droplets. The statistical model is thus considered valid only for liquid clouds.

Due to the location of the JFJ station on the alpine divide, with air masses approaching from both the north and the south, we expect Eqs. (5) and (6) to be broadly applicable to the remote European continental troposphere, however with a boundary layer influence. Indeed, these equations should be generally applicable to conditions where droplet activation occurs in the aerosol limited regime. While such empirically derived relationships have their limitations, and may not remain valid under substantially perturbed atmospheric conditions, they provide a simple and computationally efficient way to calculate the number of cloud droplets in warm clouds, when appropriately applied.

5

17

References

Abdul-Razzak, H. and Ghan, S. J.: A parameterization of aerosol activation: 2. Multiple aerosol types, Journal of Geophysical Research: Atmospheres, 105, 6837–6844, doi:10.1029/1999JD901161, http://dx.doi.org/10.1029/1999JD901161, 2000.

Albrecht, B. A.: Aerosols, cloud microphysics, and fractional cloudiness, Science, 245, 1227–1230, doi:10.1126/science.245.4923.1227, 1989.

Anttila, T., Brus, D., Jaatinen, A., Hyvärinen, A.-P., Kivekäs, N., Romakkaniemi, S., Komppula, M., and Lihavainen, H.: Relationships between particles, cloud condensation nuclei and cloud droplet activation during the third Pallas Cloud Experiment, Atmos. Chem. Phys., 12, 11 435–11 450, 2012.

Baltensperger, U., Gaggeler, H. W., Jost, D. T., Lugauer, M., Schwikowski, M., Weingartner, E., and Seibert, P.: Aerosol climatology at the high-alpine site Jungfraujoch, Switzerland, J. Geophys. Res.–Atmos., 102, 19 707–19 715, doi:10.1029/97jd00928, 1997.

Baltensperger, U., Schwikowski, M., Jost, D. T., Nyeki, S., Gaggeler, H. W., and Poulida, O.: Scavenging of atmospheric constituents in mixed phase clouds at the high-alpine site Jungfraujoch part I: Basic concept and aerosol scavenging by clouds, Atmos. Environ., 32, 3975–3983, doi:10.1016/S1352-2310(98)00051-X, 1998.

Baltensperger, U., Gaggeler, H., Jost, D., Emmenegger, M., and Nageli, W.: Continuous background aerosol monitoring wiht the epiphaniometer, Atmospheric Environment Part A-General Topics, 25, 629–634, doi:10.1016/0960-1686(91)90060-K, 1991.

Barahona, D. and Nenes, A.: Parameterization of cloud droplet formation in large-scale models: Including effects of entrainment, J. Geophys. Res.–Atmos., 112, doi:10.1029/2007JD008473, 2007.

Barahona, D., West, R. E. L., Stier, P., Romakkaniemi, S., Kokkola, H., and Nenes, A.: Comprehensively accounting for the effect of giant CCN in cloud activation parameterizations, Atmospheric Chemistry and Physics, 10, 2467–2473, doi:10.5194/acp-10-2467-2010, http://www.atmos-chem-phys.net/10/2467/2010/, 2010.

Boucher, O., Randall, D., Artaxo, P., Bretherton, C., Feingold, G., Forster, P., Kerminen, V.-M., Kondo, Y., Liao, H., Lohmann, U., Rasch, P., Satheesh, S., Sherwood, S., Stevens, B., and Zhang, X.: Cloud and Aerosols. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)], Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, pp. 571–657, 2013.

Bukowiecki, N., Weingartner, E., Martin, G., Collaud Coen, M., Zieger, P., Herrmann, E., Steinbacher, M., Gäggeler, H. W., and Baltensperger, U.: A Review of More Than 20 Years of Aerosol Observation at the High Altitude Research Station Jungfraujoch, Switzerland (3580 m asl), Aerosol and Air Quality Research, doi:10.4209/aaqr.2015.05.0305, 2016.

Christensen, S. I. and Petters, M. D.: The role of temperature in cloud droplet activation, J. Phys. Chem. A, 116, 9706–9717, doi:10.1021/jp3064454, 2012.

Conant, W. C., VanReken, T. M., Rissman, T. A., Varutbangkul, V., Jonsson, H. H., Nenes, A., Jimenez, J. L., Delia, A. E., Bahreini, R., Roberts, G. C., Flagan, R. C., and Seinfeld, J. H.: Aerosol–cloud drop concentration closure in warm cumulus, J. Geophys. Res.–Atmos., 109, n/a–n/a, doi:10.1029/2003JD004324, http://dx.doi.org/10.1029/2003JD004324, d13204, 2004.

Cozic, J., Verheggen, B., Mertes, S., Connolly, P., Bower, K., Petzold, A., Baltensperger, U., and Weingartner, E.: Scavenging of black carbon in mixed phase clouds at the high alpine site Jungfraujoch, Atmos. Chem. Phys., 7, 1797–1807, 2007.

Cozic, J., Verheggen, B., Weingartner, E., Crosier, J., Bower, K. N., Flynn, M., Coe, H., Henning, S., Steinbacher, M., Henne, S., Coen, M. C., Petzold, A., and Baltensperger, U.: Chemical composition of free tropospheric aerosol for PM1 and coarse mode at the high alpine site Jungfraujoch, Atmos. Chem. Phys., 8, 407–423, 2008.

Cui, J., Sprenger, M., Staehelin, J., Siegrist, A., Kunz, M., Henne, S., and Steinbacher, M.: Impact of stratospheric intrusions and interconti-nental transport on ozone at Jungfraujoch in 2005: comparison and validation of two Lagrangian approaches, Atmospheric Chemistry and Physics, 9, 3371–3383, doi:10.5194/acp-9-3371-2009, http://www.atmos-chem-phys.net/9/3371/2009/, 2009.

Dusek, U., Frank, G. P., Hildebrandt, L., Curtius, J., Schneider, J., Walter, S., Chand, D., Drewnick, F., Hings, S., Jung, D., Borrmann, S., and Andreae, M. O.: Size Mmatters more than chemistry for cloud-nucleating ability of aerosol particles, Science, 312, 1375–1378, doi:10.1126/science.1125261, http://www.sciencemag.org/content/312/5778/1375.abstract, 2006.

Fitzgerald, J. W. and Spyers-Duran, P. A.: Changes in cloud nucleus concentration and cloud droplet size distribution associated with pollution from St. Louis, J. Appl. Meteor., 12, 511–516, doi:10.1175/1520-0450(1973)012<0511:CICNCA>2.0.CO;2, 1973.

Fountoukis, C. and Nenes, A.: Continued development of a cloud droplet formation parameterization for global climate models, J. Geophys. Res.–Atmos., 110, D11 212, doi:10.1029/2004JD005591, 2005.

Fountoukis, C., Nenes, A., Meskhidze, N., Bahreini, R., Conant, W. C., Jonsson, H., Murphy, S., Sorooshian, A., Varutbangkul, V., Brechtel, F., Flagan, R. C., and Seinfeld, J. H.: Aerosol–cloud drop concentration closure for clouds sampled during the Interna-tional Consortium for Atmospheric Research on Transport and Transformation 2004 campaign, J. Geophys. Res.–Atmos., 112, n/a–n/a, doi:10.1029/2006JD007272, http://dx.doi.org/10.1029/2006JD007272, d10S30, 2007.

Friedman, A.: Statistics for library and information services: A primer for using open source R software for accessibility and visualization, Rowman & Littlefield, 2015.

Fröhlich, R., Cubison, M. J., Slowik, J. G., Bukowiecki, N., Canonaco, F., Croteau, P. L., Gysel, M., Henne, S., Herrmann, E., Jayne, J. T., Steinbacher, M., Worsnop, D. R., Baltensperger, U., and Prévôt, A. S. H.: Fourteen months of on-line measurements of the non-refractory submicron aerosol at the Jungfraujoch (3580 m a.s.l.) - chemical composition, origins and organic aerosol sources, Atmospheric Chemistry and Physics, 15, 11 373–11 398, doi:10.5194/acp-15-11373-2015, http://www.atmos-chem-phys.net/15/11373/2015/, 2015.

Gerber, H.: Supersaturation and droplet spectral evolution in fog, J. Atmos. Sci., 48, 2569–2588, doi:10.1175/1520-0469(1991)048<2569:SADSEI>2.0.CO;2, 1991.

Ghan, S. J., Leung, L. R., Easter, R. C., and Abdul-Razzak, H.: Prediction of cloud droplet number in a general circulation model, Journal of Geophysical Research: Atmospheres, 102, 21 777–21 794, doi:10.1029/97JD01810, http://dx.doi.org/10.1029/97JD01810, 1997.

Gilge, S., Plass-Duelmer, C., Fricke, W., Kaiser, A., Ries, L., Buchmann, B., and Steinbacher, M.: Ozone, carbon monoxide and nitrogen oxides time series at four alpine GAW mountain stations in central Europe, Atmos. Chem. Phys., 10, 12 295–12 316, doi:10.5194/acp-10-12295-2010, 2010.

Grömping, U.: Relative importance for linear regression in R: the package relaimpo, Journal of Statistical Software, 17, 1–27, 2006.

Grömping, U.: Estimators of relative importance in linear regression based on variance decomposition, American Statistican, 61, 139–147, doi:10.1198/000313007X188252, 2007.

Hammer, E., Bukowiecki, N., Gysel, M., Jurányi, Z., Hoyle, C. R., Vogt, R., Baltensperger, U., and Weingartner, E.: Investigation of the effective peak supersaturation for liquid-phase clouds at the high-alpine site Jungfraujoch, Switzerland (3580 m a.s.l.), Atmos. Chem. Phys., 14, 1123–1139, doi:10.5194/acp-14-1123-2014, 2014.

Hammer, E., Bukowiecki, N., Luo, B. P., Lohmann, U., Marcolli, C., Weingartner, E., Baltensperger, U., and Hoyle, C. R.: Sensitivity estimations for cloud droplet formation in the vicinity of the high-alpine research station Jungfraujoch (3580 m a.s.l.), Atmospheric Chemistry and Physics, 15, 10 309–10 323, doi:10.5194/acp-15-10309-2015, http://www.atmos-chem-phys.net/15/10309/2015/, 2015.

Henning, S., Weingartner, E., Schmidt, S., Wendisch, M., Gaggeler, H. W., and Baltensperger, U.: Size-dependent aerosol activation at the high-alpine site Jungfraujoch (3580 m asl), Tellus Ser. B, 54, 82–95, doi:10.1034/j.1600-0889.2002.00299.x, 2002.

Herrmann, E., Weingartner, E., Henne, S., Vuilleumier, L., Bukowiecki, N., Steinbacher, M., Conen, F., Collaud Coen, M., Hammer, E., Jurányi, Z., Baltensperger, U., and Gysel, M.: Analysis of long-term aerosol size distribution data from Jungfraujoch with emphasis on free tropospheric conditions, cloud influence, and air mass transport, Journal of Geophysical Research: Atmospheres, 120, 9459–9480, doi:10.1002/2015JD023660, http://dx.doi.org/10.1002/2015JD023660, 2015JD023660, 2015.

Hu, Y. X. and Stamnes, K.: An accurate parameterization of the radiative properties of water clouds suitable for use in climate models, J. Climate, 6, 728–742, doi:10.1175/1520-0442(1993)006<0728:AAPOTR>2.0.CO;2, 1993.

Johnson, B. T., Shine, K. P., and Forster, P. M.: The semi-direct aerosol effect: Impact of absorbing aerosols on marine stratocumulus, Quart. J. Roy. Meteor. Soc., 130, 1407–1422, doi:10.1256/qj.03.61, 2004.

Jones, A.: A climate model study of indirect radiative forcing by anthropogenic sulphate aerosols, Nature, 370, 450–453, 1994.

Jurányi, Z., Gysel, M., Weingartner, E., DeCarlo, P. F., Kammermann, L., and Baltensperger, U.: Measured and modelled cloud condensation nuclei number concentration at the high alpine site Jungfraujoch, Atmos. Chem. Phys., 10, 7891–7906, doi:10.5194/acp-10-7891-2010, 2010.

Jurányi, Z., Gysel, M., Weingartner, E., Bukowiecki, N., Kammermann, L., and Baltensperger, U.: A 17 month climatology of the cloud condensation nuclei number concentration at the high alpine site Jungfraujoch, J. Geophys. Res.–Atmos., 116, 2156–2202, doi:10.1029/2010JD015199, 2011.

Kammermann, L., Gysel, M., Weingartner, E., and Baltensperger, U.: 13-month climatology of the aerosol hygroscopicity at the free tropospheric site Jungfraujoch (3580 m a.s.l.), Atmos. Chem. Phys., 10, 10 717–10 732, doi:10.5194/acp-10-10717-2010, 2010.

King, S. M., Rosenoern, T., Shilling, J. E., Chen, Q., and Martin, S. T.: Increased cloud activation potential of secondary organic aerosol for atmospheric mass loadings, Atmos. Chem. Phys., 9, 2959–2971, 2009.

Kivekäs, N., Kerminen, V.-M., Anttila, T., Korhonen, H., Lihavainen, H., Komppula, M., and Kulmala, M.: Parameterization of cloud droplet activation using a simplified treatment of the aerosol number size distribution, Journal of Geophysical Research: Atmospheres, 113, n/a–n/a, doi:10.1029/2007JD009485, http://dx.doi.org/10.1029/2007JD009485, d15207, 2008.

Köhler, H.: The nucleus in and the growth of hygroscopic droplets, Trans. Faraday Soc., 32, 1152–1161, doi:10.1039/TF9363201152, 1936.

Kruskal, W.: Relative importance by averaging over orderings, American Statistican, 41, 6–10, doi:10.2307/2684310, 1987.

Kumar, P., Sokolik, I. N., and Nenes, A.: Parameterization of cloud droplet formation for global and regional models: including adsorption activation from insoluble CCN, Atmos. Chem. Phys., 9, 2517–2532, doi:10.5194/acp-9-2517-2009, 2009.

Liang, Q., Jaeglé, L., Jaffe, D. A., Weiss-Penzias, P., Heckman, A., and Snow, J. A.: Long-range transport of Asian pollution to the northeast Pacific: Seasonal variations and transport pathways of carbon monoxide, Journal of Geophysical Research: Atmospheres, 109, doi:10.1029/2003JD004402, http://dx.doi.org/10.1029/2003JD004402, 2004.

Lindeman, R., Merenda, P., and Gold, R.: Introduction to Bivariate and Multivariate Analysis, Scott Foresman, Glenview IL., 1980.

Lugauer, M., Baltensperger, U., Furger, M., Gaggeler, H., Jost, D., Schwikowski, M., and Wanner, H.: Aerosol transport to the high Alpine sites Jungfraujoch (3454 m asl) and Colle Gnifetti (4452 m asl), Tellus Series B-Chemical and Physical Meteorology, 50, 76–92, doi:10.1034/j.1600-0889.1998.00006.x, 1998.

Martin, G. M., Johnson, D. W., and Spice, A.: The measurement and parameterization of effective radius of droplets in warm stratocumulus clouds, J. Atmos. Sci., 51, 1823–1842, doi:10.1175/1520-0469(1994)051<1823:TMAPOE>2.0.CO;2, 1994.

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

Meskhidze, N., Nenes, A., Conant, W. C., and Seinfeld, J. H.: Evaluation of a new cloud droplet activation parameterization with in situ data from CRYSTAL-FACE and CSTRIPE, J. Geophys. Res.–Atmos., 110, n/a–n/a, doi:10.1029/2004JD005703, http://dx.doi.org/10.1029/2004JD005703, d16202, 2005.

Moore, R. H., Karydis, V. A., Capps, S. L., Lathem, T. L., and Nenes, A.: Droplet number uncertainties associated with CCN: an assessment using observations and a global model adjoint, Atmospheric Chemistry and Physics, 13, 4235–4251, doi:10.5194/acp-13-4235-2013, http://www.atmos-chem-phys.net/13/4235/2013/, 2013.

Morales Betancourt, R. and Nenes, A.: Understanding the contributions of aerosol properties and parameterization discrepancies to droplet number variability in a global climate model, Atmospheric Chemistry and Physics, 14, 4809–4826, doi:10.5194/acp-14-4809-2014, http://www.atmos-chem-phys.net/14/4809/2014/, 2014.

Nenes, A. and Seinfeld, J. H.: Parameterization of cloud droplet formation in global climate models, J. Geophys. Res.–Atmos., 108, 4415, doi:10.1029/2002JD002911, 2003.

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?, Geophysical Research Letters, 29, 29–1–29–4, doi:10.1029/2002GL015295, http://dx.doi.org/10.1029/2002GL015295, 1848, 2002.

Nyeki, S., Baltensperger, U., Colbeck, I., Jost, D., Weingartner, E., and Gaggeler, H.: The Jungfraujoch high-Alpine research station (3454m) as a background clean continental site for the measurement of aerosol parameters, J. Geophys. Res.–Atmos., 103, 6097–6107, doi:10.1029/97JD03123, 1998.

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, Atmospheric Chemistry and Physics, 12, 2823–2847, doi:10.5194/acp-12-2823-2012, http://www.atmos-chem-phys.net/12/2823/2012/, 2012.

Petters, M. D. and Kreidenweis, S. M.: A single parameter representation of hygroscopic growth and cloud condensation nucleus activity, Atmos. Chem. Phys., 7, 1961–1971, doi:10.5194/acp-7-1961-2007, 2007.

R Core Team: R: A Language and Environment for Statistical Computing, R Foundation for Statistical Computing, Vienna, Austria, http://www.R-project.org, ISBN 3-900051-07-0, 2014.

Raga, G. B. and Jonas, P. R.: On the link between cloud-top radiative properties and sub-cloud aerosol concentrations, Quart. J. Roy. Meteor. Soc., 119, 1419–1425, doi:10.1002/qj.49711951410, 1993.

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

Rissman, T., Nenes, A., and Seinfeld, J.: Chemical amplification (or dampening) of the Twomey effect: Conditions derived from droplet activation theory, Journal of the Atmospheric Sciences, 61, 919–930, doi:10.1175/1520-0469(2004)061<0919:CAODOT>2.0.CO;2, 2004.

Shilling, J. E., King, S. M., Mochida, M., and Martin, S. T.: Mass spectral evidence that small changes in composition caused by oxidative aging processes alter aerosol CCN properties, J. Phys. Chem. A, 111, 3358–3368, doi:10.1021/jp068822r, 2007.

Shulman, M., Jacobson, M., Carlson, R., Synovec, R., and Young, T.: Dissolution behavior and surface tension effects of organic compounds in nucleating cloud droplets, Geophys. Res. Lett., 23, 277–280, doi:10.1029/95GL03810, 1996.

Sjogren, S., Gysel, M., Weingartner, E., Alfarra, M. R., Duplissy, J., Cozic, J., Crosier, J., Coe, H., and Baltensperger, U.: Hygroscopicity of the submicrometer aerosol at the high-alpine site Jungfraujoch, 3580 m a.s.l., Switzerland, Atmospheric Chemistry and Physics, 8, 5715–5729, doi:10.5194/acp-8-5715-2008, 2008.

Spiegel, J. K., Zieger, P., Bukowiecki, N., Hammer, E., Weingartner, E., and Eugster, W.: Evaluating the capabilities and uncertainties of droplet measurements for the fog droplet spectrometer (FM-100), Atmos. Meas. Tech., 5, 2237–2260, doi:10.5194/amt-5-2237-2012, 2012.

Staudt, A., Jacob, D., Logan, J., Bachiochi, D., Krishnamurti, T., and Sachse, G.: Continental sources, transoceanic transport, and inter-hemispheric exchange of carbon monoxide over the Pacific, J. Geophys. Res.–Atmos., 106, 32 571–32 589, doi:10.1029/2001JD900078, 2001.

Tonidandel, S. and LeBreton, J. M.: Relative importance analysis: A useful supplement to regression analysis, Journal of Business and Psychology, 26, 1–9, doi:10.1007/s10869-010-9204-3, 2011.

Twomey, S.: Pollution and the planetary albedo, Atmos. Environ., 8, 1251 – 1256, doi:doi:10.1016/0004-6981(74)90004-3, 1974.

Twomey, S. A.: The influence of pollution on the shortwave albedo of clouds, J. Atmos. Sci., 34, 1149–1152, doi:10.1175/1520-0469(1977)034<1149:TIOPOT>2.0.CO;2, 1977.

Verheggen, B., Cozic, J., Weingartner, E., Bower, K., Mertes, S., Connolly, P., Gallagher, M., Flynn, M., Choularton, T., and Baltensperger, U.: Aerosol partitioning between the interstitial and the condensed phase in mixed-phase clouds, J. Geophys. Res.–Atmos., 112, 2156–2202, doi:10.1029/2007JD008714, 2007.

Weingartner, E., Nyeki, S., and Baltensperger, U.: Seasonal and diurnal variation of aerosol size distributions (10 < D < 750 nm) at a high-alpine site (Jungfraujoch 3580 m asl), J. Geophys. Res.–Atmos., 104, 26 809–26 820, 1999.

Wiedensohler, A., Birmili, W., Nowak, A., Sonntag, A., Weinhold, K., Merkel, M., Wehner, B., Tuch, T., Pfeifer, S., Fiebig, M., Fjäraa, A. M., Asmi, E., Sellegri, K., Depuy, R., Venzac, H., Villani, P., Laj, P., Aalto, P., Ogren, J. A., Swietlicki, E., Williams, P., Roldin, P., Quincey, P., Hüglin, C., Fierz-Schmidhauser, R., Gysel, M., Weingartner, E., Riccobono, F., Santos, S., Grüning, C., Faloon, K., Beddows, D., Harrison, R., Monahan, C., Jennings, S. G., O'Dowd, C. D., Marinoni, A., Horn, H.-G., Keck, L., Jiang, J., Scheckman, J., McMurry, P. H., Deng, Z., Zhao, C. S., Moerman, M., Henzing, B., de Leeuw, G., Löschau, G., and Bastian, S.: Mobility particle size spectrometers: harmonization of technical standards and data structure to facilitate high quality long-term observations of atmospheric particle number size distributions, Atmos. Meas. Tech., 5, 657–685, doi:10.5194/amt-5-657-2012, 2012.

Yashiro, H., Sugawara, S., Sudo, K., Aoki, S., and Nakazawa, T.: Temporal and spatial variations of carbon monoxide over the western part of the Pacific Ocean, J. Geophys. Res.–Atmos., 114, doi:10.1029/2008JD010876, 2009.

Zellweger, C., Forrer, J., Hofer, P., Nyeki, S., Schwarzenbach, B., Weingartner, E., Ammann, M., and Baltensperger, U.: Partitioning of reactive nitrogen (NO_y) and dependence on meteorological conditions in the lower free troposphere, Atmospheric Chemistry and Physics, 3, 779–796, doi:10.5194/acp-3-779-2003, 2003.

Zellweger, C., Hüglin, C., Klausen, J., Steinbacher, M., Vollmer, M., and Buchmann, B.: Inter-comparison of four different carbon monoxide measurement techniques and evaluation of the long-term carbon monoxide time series of Jungfraujoch, Atmos. Chem. Phys., 9, 3491–3503, doi:10.5194/acp-9-3491-2009, 2009.

Table 1. Dates and references for each CLACE campaign.

Campaign	Start date	End date	Reference
CLACE2002	4 Jul 2002	20 Jul 2002	=
CLACE2004	16 Jul 2004	30 Sep 2004	Cozic et al. (2007)
CLACE2010	1 Jul 2010	13 Aug 2010	Spiegel et al. (2012)
CLACE2011	1 Jul 2011	23 Aug2011	Hammer et al. (2014)

Table 2. Parameters describing the performance of the microphysical parameterisation in capturing the number of observed cloud droplets, when run for three different hydroscopicity parameters.

Hydroscopicity	R	RMSE	ME
0.2	0.85	70.8	44.7
0.25	0.86	67.2	42.8
0.3	0.86	65.7	42.2

Zhang, L., Jacob, D. J., Bowman, K. W., Logan, J. A., Turquety, S., Hudman, R. C., Li, Q., Beer, R., Worden, H. M., Worden, J. R., Rinsland, C. P., Kulawik, S. S., Lampel, M. C., Shephard, M. W., Fisher, B. M., Eldering, A., and Avery, M. A.: Ozone-CO correlations determined by the TES satellite instrument in continental outflow regions, Geophys. Res. Lett., 33, doi:10.1029/2006GL026399, 2006.

Zhang, L., Jacob, D. J., Kopacz, M., Henze, D. K., Singh, K., and Jaffe, D. A.: Intercontinental source attribution of ozone pollution at western US sites using an adjoint method, Geophys. Res. Lett., 36, doi:10.1029/2009GL037950, 2009.

5
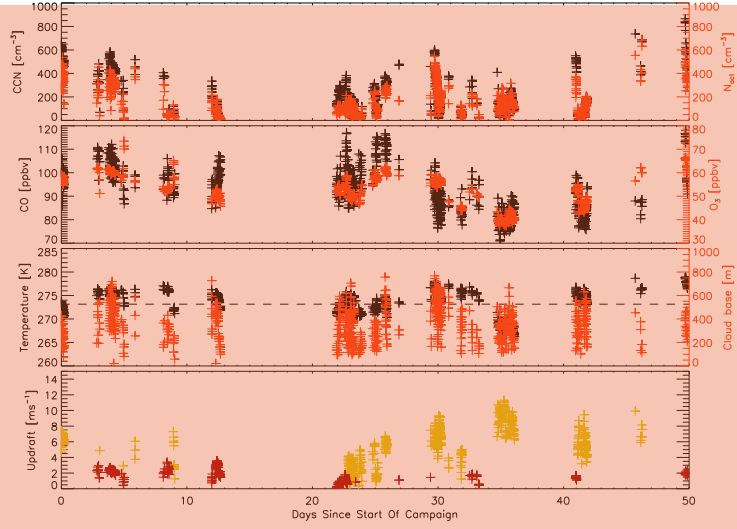
Figure 1. Time series for several quantities measured directly during the CLACE2011 campaign, or derived from other CLACE2011 data. The number of CCN, shown in the top panel, refers the number of particles larger than 80 nm in diameter, which are considered potential CCN in constructing the statistical models (see text). In the bottom panel, the colour of the points indicates the wind direction, with yellow showing wind classified as being northwest and red southeast. Data are only plotted for times when the JFJ was in-cloud.
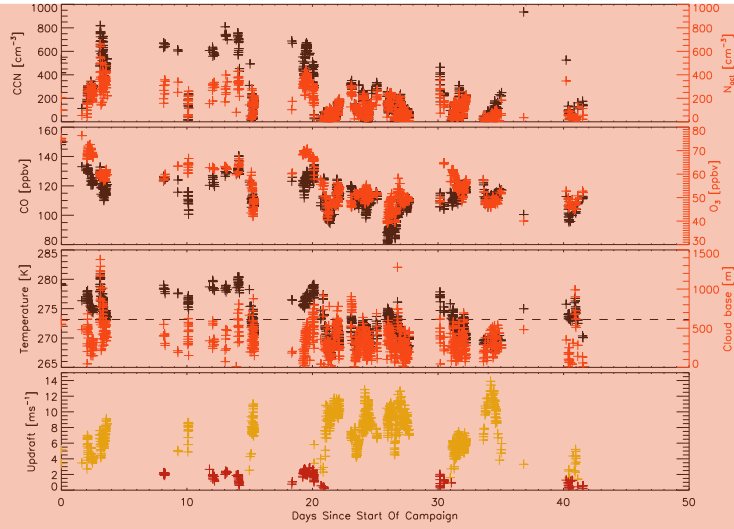
Figure 2. As for Fig. 1, but for CLACE2010. Note that the axis ranges differ from those in Fig. 1.
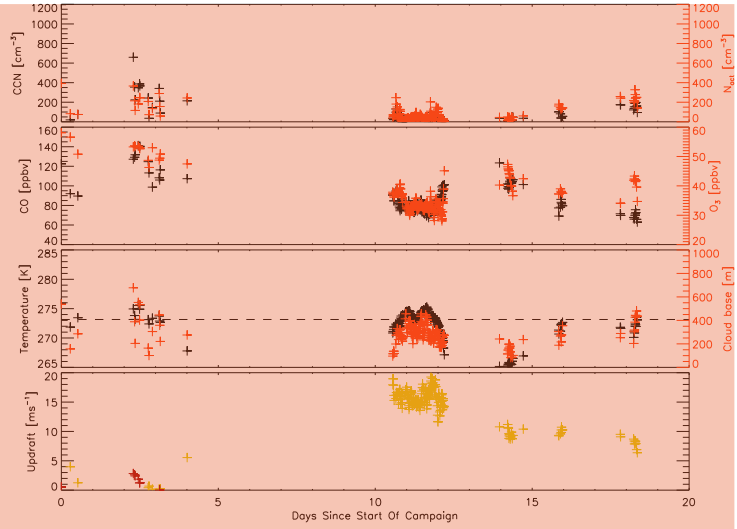
Figure 3. As for Fig. 1, but for CLACE2004. Note that the axis ranges differ from those in Fig. 1.
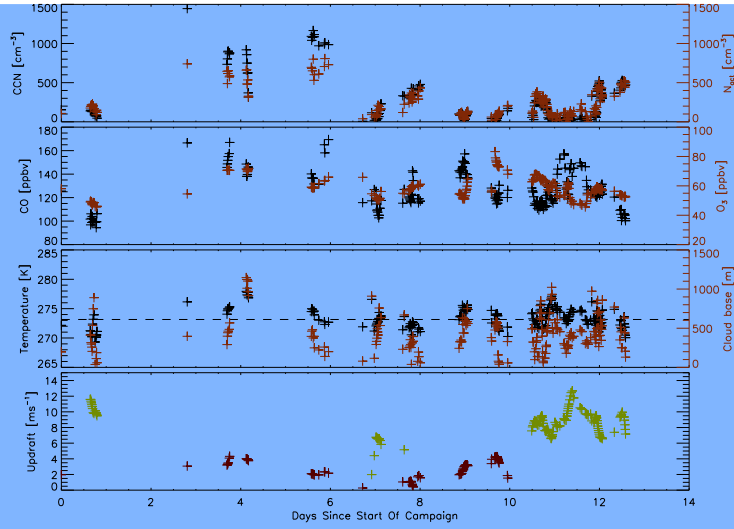
Figure 4. As for Fig. 1, but for CLACE2002. Note that the axis ranges differ from those in Fig. 1.
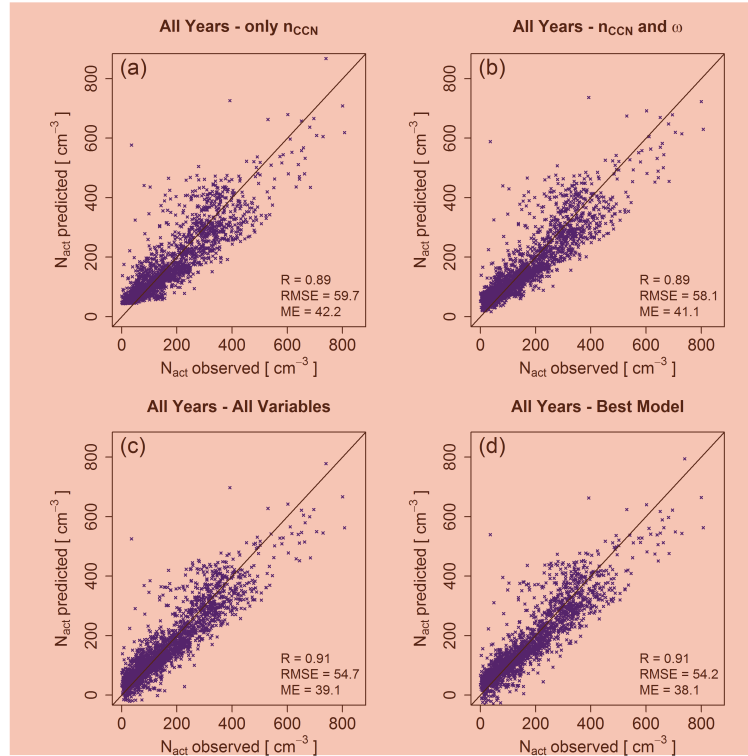
Figure 5. The number of cloud droplets calculated using different statistical models, plotted against the observed number of residuals. The model used for panel **(a)** included only the n_{CCN}, for **(b)** n_{CCN} and updraught velocity are included in the model, and in panel **(c)** all variables are included. In panel **(d)**, all variables are included, however the log of the updraught velocity is used.
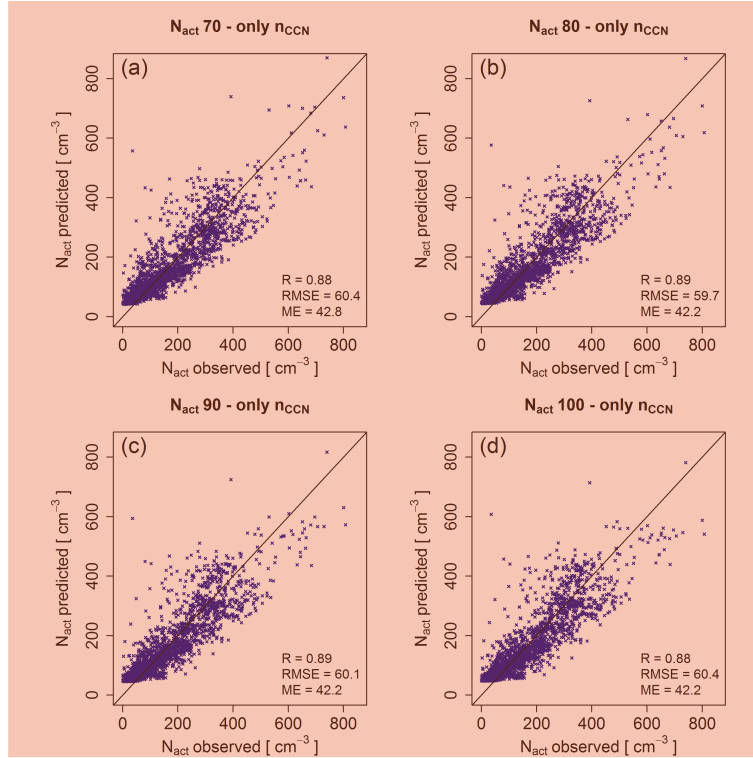
Figure 6. The number of cloud droplets calculated using a statistical model, based on a regression analysis including only the number of potential CCN, plotted against the observed number of residuals. Panel **(a)**: Potential CCN are considered to be all particles with a diameter larger than 70 nm, Panel **(b)**: 80 nm, Panel **(c)**: 90nm, Panel **(d)**: 100nm.
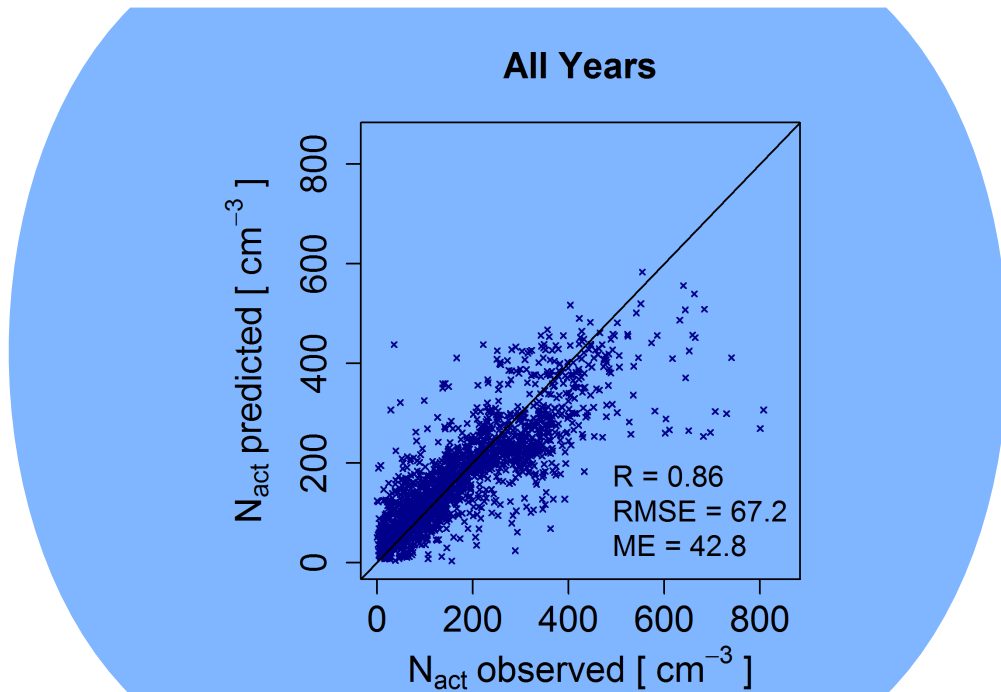
Figure 7. The number of cloud droplets calculated with the microphysical parametrization, plotted against the measured number of residuals.
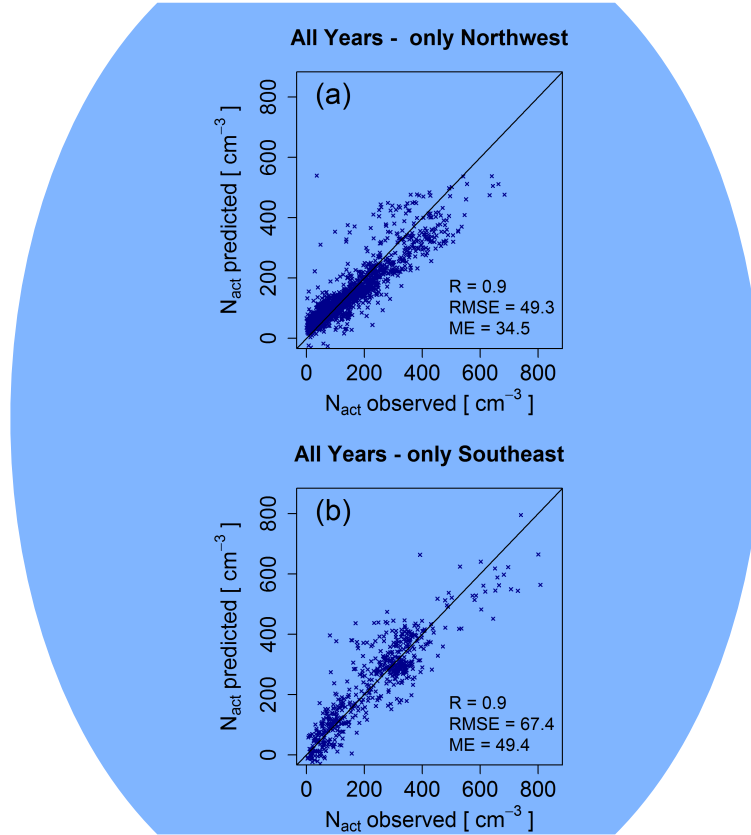
Figure 8. The modelled number of cloud droplets (Eq. 5) plotted against the observed number of residuals. Only data for northwestern wind conditions are included in panel (**a**), while only data for southeastern wind directions are included in panel (**b**).
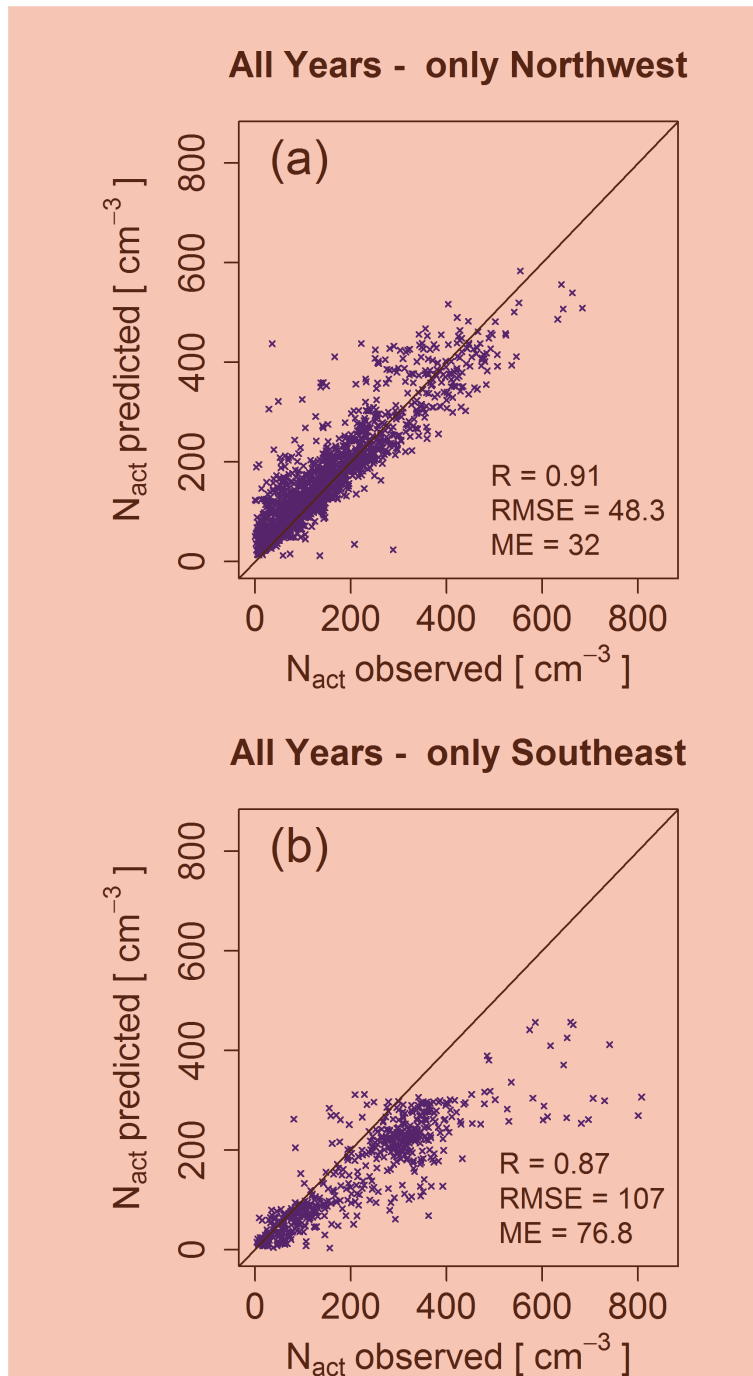
Figure 9. The number of cloud droplets calculated by the microphysical parametrization, separated by wind direction, compared to the number of observed cloud droplet residuals.
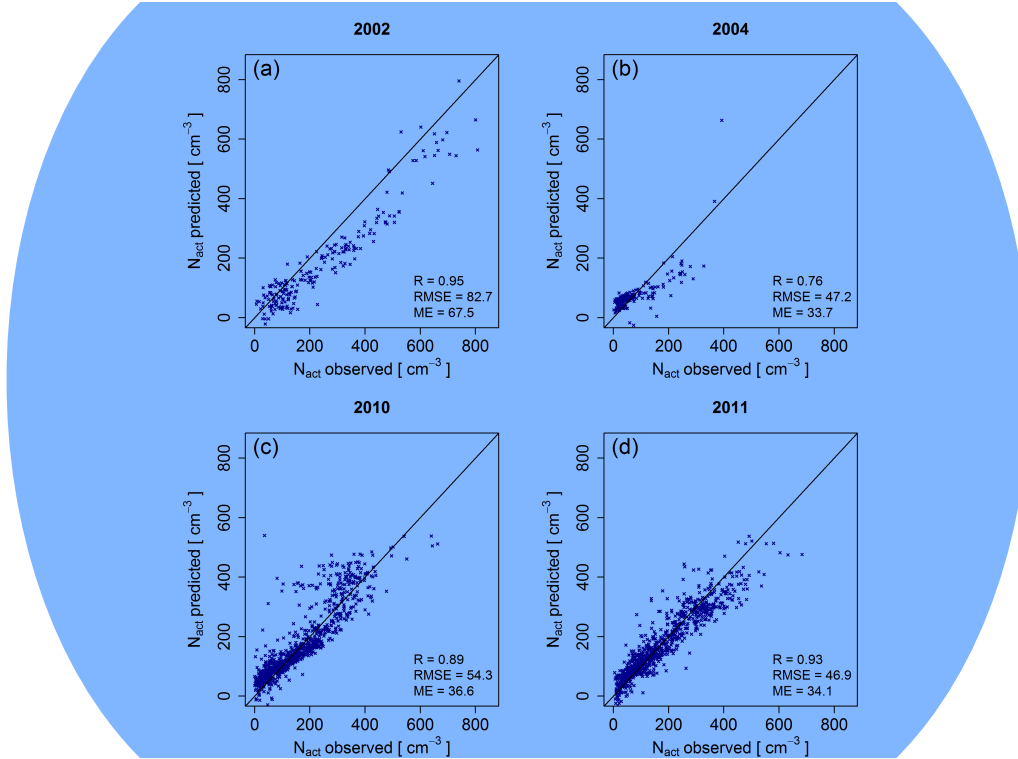
Figure 10. The modelled droplet numbers (Eq. 5) plotted against the observed number of residuals for each year separately.
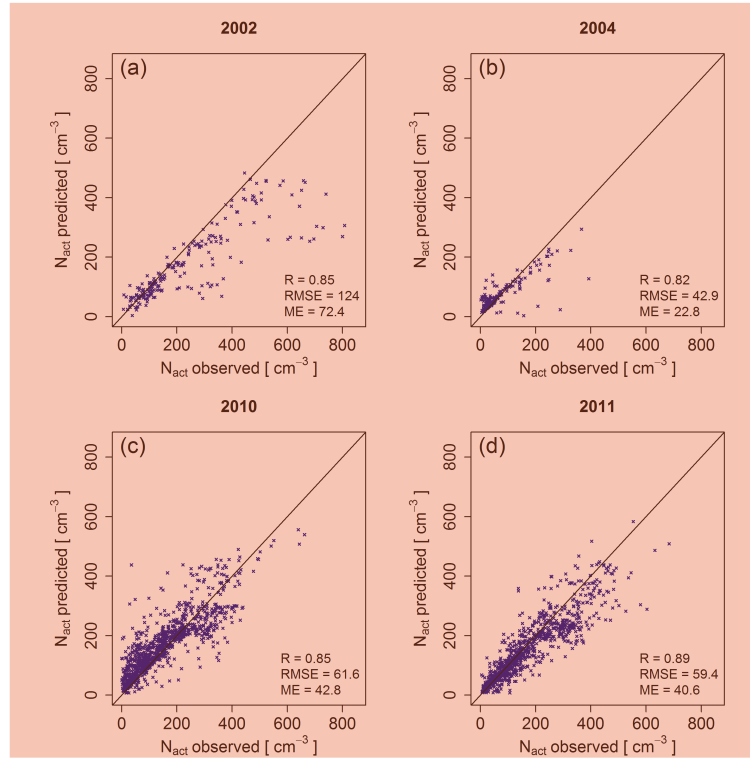
Figure 11. The number of cloud droplets calculated by the microphysical parametrization, separated by year, compared to the number of observed cloud droplet residuals.
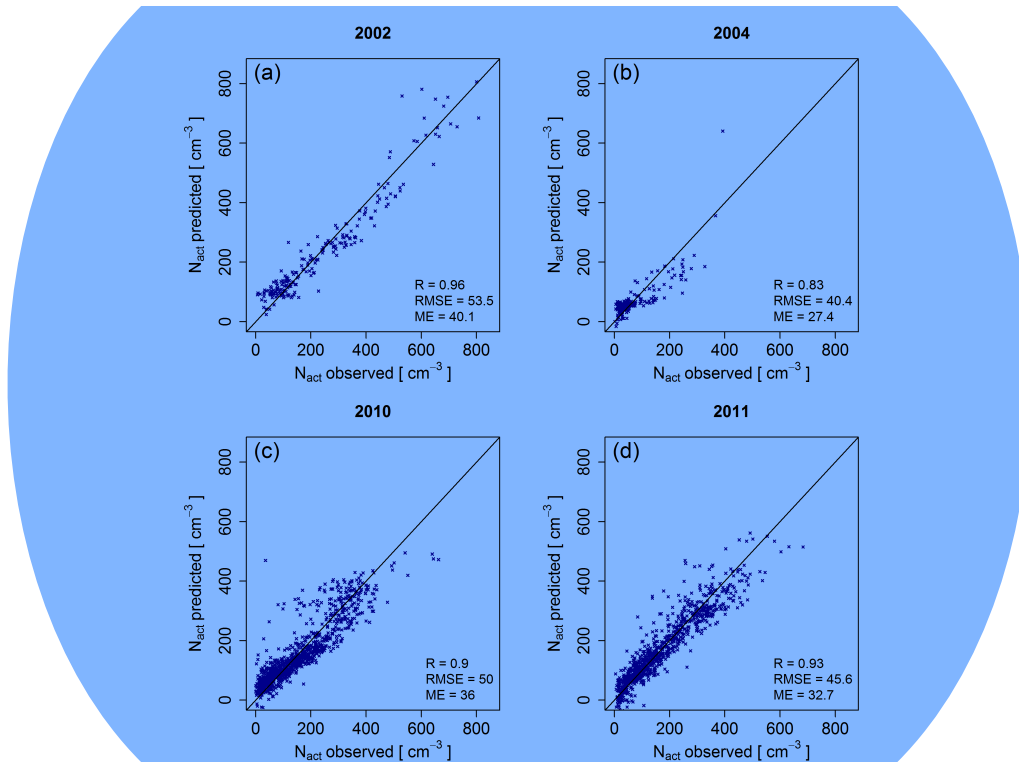
Figure 12. As for Fig. 10, however the model was re-calculated to provide the best fit for each year individually.
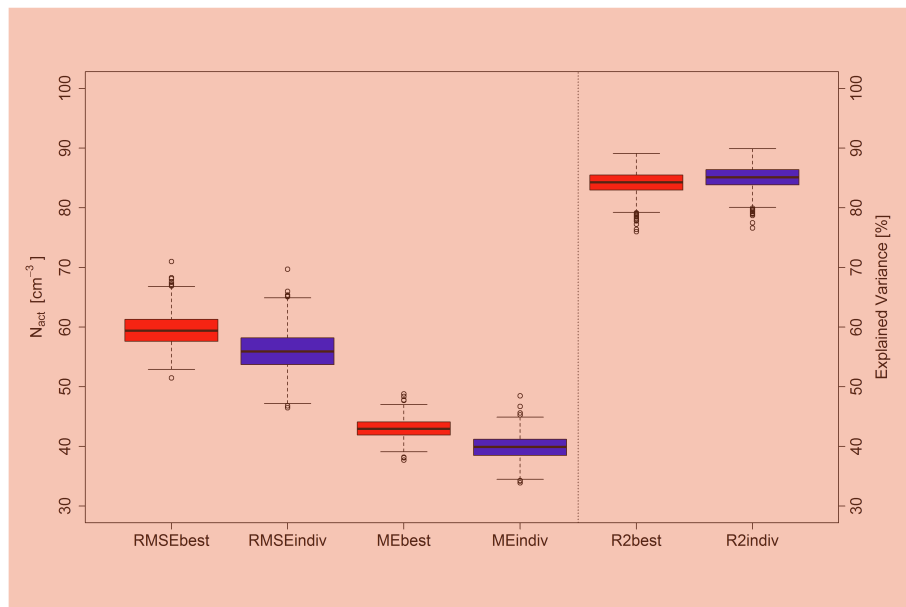
Figure 13. A box plot of the R^2, RMSE and ME values for the application of the general model (Eq. 5) to 1000 random samples of 100 data points from each year. The red boxes show the range of R^2 and error values when Eq. (5) is applied to the sampled data, while the blue boxes show the ranges when the model is refitted individually to the data sampled in each case.
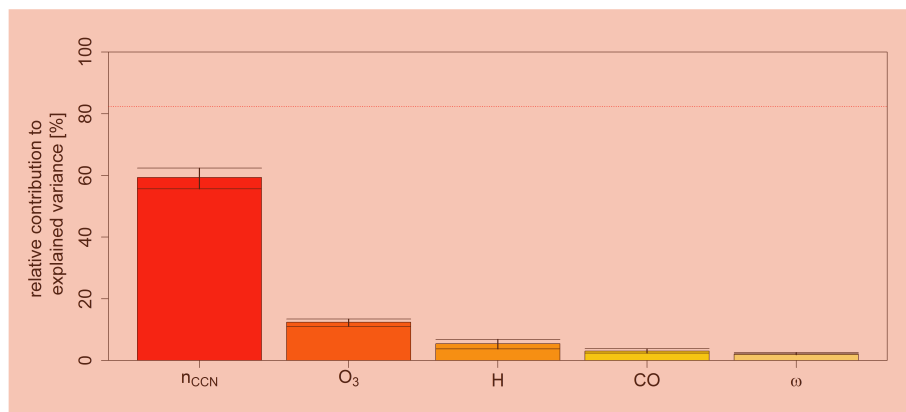

Figure 14. The contribution of each of the model variables in Eq. (5) to the explained variance. The error bars show the spread of the variation of the contribution values in the random samples from Fig. 13.
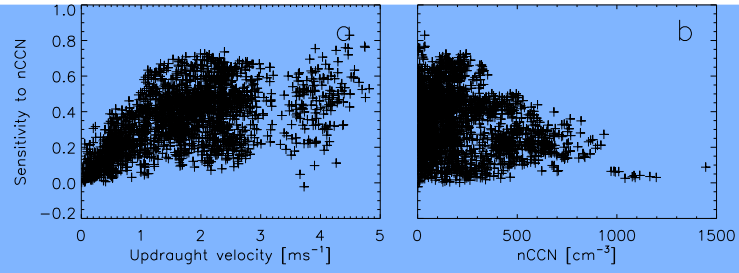
Figure 15. The sensitivity of the modelled droplet number to the updraught velocity (corrected by a factor of 0.25, following Hammer et al. 2015, panel a) and to the number of particles larger than 80nm (panel b).
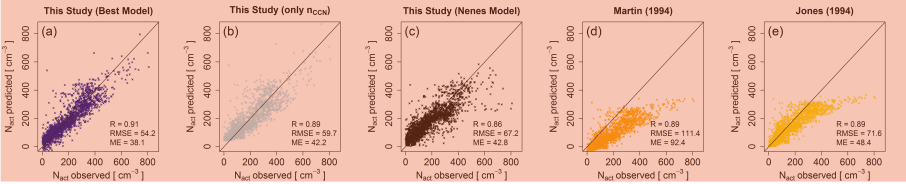
Figure 16. A comparison of the model developed in this study, and the microphysical parametrization, with the performance of two existing models, those of Martin et al. (1994) and Jones (1994), which are based only on n_{CCN}.