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.

1. The authors state in abstract (line numbers 12-13) that ": : :we expect that this model is applicable to warm, free tropospheric clouds over the European continent.". Similarly, at the end of the Introduction (line numbers 80-81) it is stated that ": : :allows the construction of relationships which are applicable to a wide range of conditions." and at the end of Conclusions section (line numbers 477-478): ": : : we expect Equation 2 and Equation 3 [derived parameterizations for CDNC] to be broadly applicable to the European free troposphere."
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."

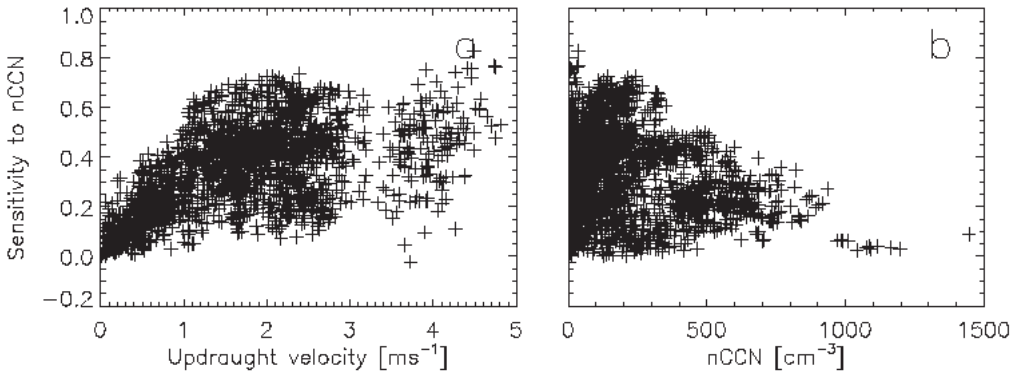

Figure 15 The sensitivity of the modelled droplet number to the updraught velocity (corrected by a factor of 0.25, following Hammer et al. 2015, panel a) and to the number of particles larger than 80nm (panel b).

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

experimental CCN and cloud formation studies.
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.

2. Introduction, lines 36-37. The studies of the Henning et al. (2002) and Verheggen et al. (2007) cited here did not consider CCN but only cloud droplet formation.
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…”

3. Introduction, fourth, fifth and sixth paragraphs. Here previous studies on CCN activation and cloud droplets formation are discussed in mixed order. I’d reorganize the discussion so that the CCN studies are treated first, followed by the cloud formation studies.
The paragraphs have been re-arranged accordingly

4. Line number 61. I’d write “: : :the mechanisms through which: : :” instead of “: : :the way which: : :”, for example.
Done

5. Line number 64. “: : :particles and thus their activation into cloud droplets...” instead of “: : : aerosol droplets and thus activation to form droplets: : :”, for example.
This now reads “..surface active compounds may influence surface tension and thus the activation of aerosol particles to form cloud droplets”

6. Line number 101. “: : :is influenced: : :”, instead of “: : :will be influenced: : :”.
Done

7. Line number 179. “. Without entrainment, all particles above a particular size will be activated.”. This is strictly true only for internally mixed aerosols. Otherwise, particle population may contain a mixture of large, non-hygroscopic and hygroscopic particles so that the former particles do not activate, leading to a non-monotonic activation curve.
See answer to point 8.

8. Line numbers 179-181. The definition of the activation diameter given here is extremely confusing. To me, it does not even sound right – critical diameter cannot be defined solely based on “critical diameter of the least hygroscopic particle” as the particle dry size plays a role too. Please clarify.
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.”

9. Line numbers 216-218. Another definition is given here for activation diameter.

Please define the concept in consistent and precise manner.
This second definition has been edited and shifted together with the text in the previous comment (8).

10. Line 235. What is “potential CCN”? Please clarify.
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 ”

11. Lines 243-247. This sentence should be split into two or more sentences.
We have split this sentence into three sentences.

12. Section 4.3. Please give a short description of the applied statistical method as it would benefit understanding the results of this study. Also, in line no. 274, there is a question mark in parenthesis following “: : :developed by: : :”.
A detailed description has been added. The question mark was a typographical error which has been removed accordingly.

13. Line numbers 278-279. The end of the sentence (“: : :to investigate its general applicability”) can be omitted.
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.”

14. Line number 293. The expression “: : :is not complete,” is not fitting in this context.
This now reads “…during which not all particles larger than 80nm are activated….”

15. Line numbers 293-294. “: : :activation appears to be lower: : :”. See the previous comment.
This now reads “….the fraction of particles that are activated appears to be lower…..”

16. Line number 307. Parenthesis should be removed from the end of the line.
Done.

17. Line number 323, equation 2. According to this equation increasing ozone levels lead to increases in CDNC, while increasing [CO] decreases CDNC. Is there any physical explanation for these trends? Also, the derived parametrization does not contain any variable that would directly reflect the particle chemical composition. Does this limit applicability of eq. 2 as cloud droplet concentrations depend, at least to some extent, on the chemical composition of CCN?
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.

18. Section 5.2. Is there any explanation as for the worse performance of the parameterization against data from SE wind sector?
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.