Review of "Heterogeneous ice nucleation on dust particles sourced from 9 deserts worldwide - Part 1: Immersion freezing" by *Boose et al*.

**Summary**

The authors investigated the relationship between the immersion freezing behavior of diverse natural dusts from nine desert regions around the globe (4 airborne and 11 sieved/milled surface samples) and their mineralogical compositions (that are based on bulk XRD). The immersion freezing measurements were conducted in the wide range of sub-zero temperatures (235-263 K) using two independent ice nucleation (IN) measurement techniques (i.e., mainly by IMCA-ZINC-IODE for $T$ below ~250 K and partially by FRIDGE for selected samples at temperatures above ~250 K). The authors evaluated the immersion freezing behavior in two metrics, frozen fraction ($FF$) and ice active surface-site density, $n_s(T)$.

The authors suggest that their dust samples exhibit somewhat different freezing behavior potentially due to variations in mineralogy (quartz and K-feldspar seem highly efficient immersion freezing components of dust) and sampling preparation procedures prior to the aerosolization (e.g., sieving and milling). More interestingly, the authors suggest that the surface-collected samples contain more efficient INPs, inferred by $FF$ and $n_s(T)$, as compared to the airborne samples. The authors imply that mineralogy may play a significant role to explain the observed difference (P12 L18-20; P12 L25-27). Yet, they point out that further rigorous composition-resolved (i.e., both bio- and mineral-resolved) IN study of airborne dusts are necessary (P16 L8-13).

**General comments**

The experimental approach presented in this work is rigorous, and the authors undoubtedly put an enormous amount of effort to examine a variety of samples with careful sample preparation methods. However, the authors need quite a bit more work to improve the clarity regarding their analysis as the current manuscript contains some misleading statements and over-interpreted results. Critical comments are listed below. Six major comments are listed first, followed by the section-based specific and technical revisions. Additionally, I encourage the authors of the manuscript to thoroughly proof-read and language-check their manuscript as the list of specific/technical comments is too long.

Having said that, the research questions raised in this study (i.e., airborne dust vs. surface dust, mineralogy-resolved IN study) are interesting, and the research topic is certainly an important addition to ACP. If the authors improve the clarity of the paper by addressing the comments/suggestions below, the manuscript would be suitable for publication in ACP.

**Major comments**

1. *Kaufmann et al*. vs. *Boose et al*.: One of the major conclusions from this study (that is, mineralogy matters for dust IN) goes in the opposite direction of that from the study done by *Kaufmann et al*. (2016, AMTD; K16 hereafter), which points out the similar freezing behavior amongst multiple surface dust samples despite the difference/variation in mineralogy. Does this difference appear due to the technical issue/limitation (i.e., IMCA-ZINC-IODE vs. DSC) or totally something different? I am missing such discussion in this manuscript. Please elaborate to the extent of the community-wide scale. Providing a solid answer with respect to this point would be potentially important for modelers to consider if they require a computationally expensive mineralogy-resolved parameterization for their future simulation works or not.

Along similar lines, K16 suggests that microcline-containing particles are not abundant in the atmosphere and, hence, may have overall small contributions to atmospheric ice nucleation (IN) and glaciation (e.g., P20 L22-25 of K16). Do the authors agree with that statement? The reviewer thinks that microcline is in general an important dust composition that can contribute to efficient IN of material composites. Anyhow, it is somehow puzzling for the reviewer to see that two papers from the same institute report opposite things…

2. $n_{s,\,geo}$ vs. $n_{s,\,BET}$: As the authors may be aware, there are two sub-metrics in $n_s$, namely geometric-based $n_s$ and BET based $n_s$ (*Hiranuma et al*., 2015, ACP; H15 hereafter). It appears to me that both $n_s$ metrics co-exist in Fig. 5 - Atokinson, Murray and Broadley seem BET-based, while Niemand and all of your data are geometric-based. This is an apples-to-oranges comparison. Note that the difference in these two $n_s$ metrics could be more than an order magnitude (e.g., for illite NX in H15), which would have a substantial effect on

your stats presented in Table 5 & P13 L5-33. I strongly urge the authors to select either one metric (the conversion method is introduced in H15), re-evaluate your stats and re-visit your statement in P1 L21-23. Your conclusion may change.

3. Your $n_s$ vs. Niemand's $n_s$: The parameterization described in *Niemand et al*. (2012; N12 hereafter) is essentially based on the observed ice number normalized to the measured "total" surface area (see Eqn. 1-4 in N12), while your parameterization introduces another approximated parameter, the weighted mean aerosol surface area ($A_{ve,w}$; P6 L11). Are these two $n_s$ values apples-to-apples? The performance comparison of IMCA-ZINC-IODE to the AIDA chamber, where the N12 study was conducted, with a reference IN material seems available in H15. The authors may extend the discussion given in H15 a bit more to justify the use of $A_{ve,w}$ in its first appearance.

In addition, the general discussion of the $n_s$ concept seems scattered over the multiple sections and somehow cumbersome (i.e., P11 L3-6, P9 L 4-7 and P10 L12-16). For clarity, I suggest the authors to briefly describe the following two things within one section (any):

> 1) Three assumptions of the $n_s$ parameterization, which is relevant to the immersion freezing characterization, include i) the probability of ice nucleation is proportional to the available surface area of immersed aerosols, ii) ice nucleation active sites are uniformly distributed over individual particle surfaces and iii) ice nucleation occurs at specific site in a deterministic manner (predominantly *T* dependent).
>
> 2) The use of the weighted mean aerosol surface area ($A_{ve,w}$) is something unique in this study (and different from N12).

Accordingly, I suggest revising the conclusion (P 16 L15-17).

4. The effect of milling: The discussion regarding the effect of milling is misleading. Specifically, the reviewer's concerns are as follows:

P1 L18-20: This statement is misleading as it sounds like the milling process generally deteriorates IN efficiency of any composite materials. In fact, this statement contradicts to the IN results of the Atacama samples ($n_{s,milled} > n_{s,sieved}$) presented in Sect. 3.4. (i.e., P14 L29-30). I suggest rephrasing this sentence to be more specific (adding the "may" word would not help).

P15 L11-13: Without any evidence of the alternation in defect densities on the Isralel dust surface or calcite surface per milling, this statement (morphology vs. mineralogy) seems too ambitious/strong. It is likely the micro pores on the surface (*Subramanyam et al*., 2016, Appl. Mater. Interfaces, doi:10.1021/acsami.6b01133) that enhance the IN activity of particles due to the inverse Kelvin effect (*Marcolli et al*., 2014, ACP). If calcite etc. is breaking up but maintaining smooth surface, there is no reason for its fragments to enhance IN activity. The authors may be aware, but the nature of active sites is still uncertain and under investigation. Concerning these points, I suggest the authors to soften the tone of this statement.

5. Airborne vs. Surface: To me, the essence of this paper is summarized in P10 L9-10. This single statement and the method to reach this point would be worth a paper. The authors have already provided a great conclusion (P16 L8-10 & P1 L9-10). Focus your story along with this line. I have some questions regarding airborne dust vs. surface-collected dust as follows:

P11 L9-10: Interesting. The authors are right - according to Table 4, airborne samples in general seem having larger $a$ ($0.48 \pm 0.07$ $K^{-1}$) as compared to the average of the rest ($0.38 \pm 0.16$ $K^{-1}$), suggesting high $T$ dependency of the airborne samples. But, this is not all about $T$ dependency. This may be rather an indication that the airborne particles miss certain active sites that can be activated at high $T$ (owing to the difference in mineralogy???). The authors may discuss and clarify this point. This observation seems important.

P12 L26-27: This is another important statement. Elaborate a bit further by discussing the atmospheric relevancy of the minerals uniquely found in the airborne samples. Put clear emphasis if your bulk measurements at least suggest the atmospheric relevance of quartz/K-feldspar. According to P7 L14-15 and P7 L20-22, I have a feeling that there would be reduced amount of quartz/K-feldspar in air. Any comments? This argument seems the core research question of this manuscript.

Table 3: It is really bothering me that the Tenerife sample (6% microcline!) is not showing any superb IN behavior as compared to other airborne dusts. The authors said that the mineralogy inferred by XRD is representative of the aerosolized Tenerife sample (P7 L28-30). The authors disregard the contribution of atmospheric processes (P14 L17-19). Then, what is limiting the IN of this particular dust?

<u>6. Table 5 (P13 L5-P14 L24):</u> This part includes a number of flaws (i.e., $n_{s,\,geo}$ vs. $n_{s,\,BET}$) and needs substantial improvements. In fact, with given limitations/assumptions in P14 L6-24, I am not convinced that these 'relative' correlations add much meaning to the manuscript. Do the authors really require this part to draw their conclusion? This whole statistics part of the manuscript could be deleted?

Again, I generally agree that quartz and microcline are IN active and may have potential importance in the atmospheric IN. What would be more valuable to see is if there is any 'absolute' relation between mineralogy and IN. For example, the authors may explore if the natural dust $n_s$ (or *FF*) can be optimized/predicted by its composition and associated individual $n_s$ scaled to the surface fraction (*SF*) of each component ($x_1$,$x_2$ … to $x_i$) [i.e., $n_{s,dust}$ = ($n_{s,x1}$ x $SF_{x1}$) + ($n_{s,x2}$ x $SF_{x2}$) + … + ($n_{s,xi}$ x $SF_{xi}$)]. Note that H15 attempted, but no success. Give more in-depth thoughts regarding the role of particular mineral IN.

**Specific comments/suggestions**

<u>Introduction</u>

P2 L22: Briefly explain what the authors mean for 'contradicting results'.

P2 L24: Briefly describe 'certain minerals'.

P2 L25: "For the implementation…" - I suggest starting a new paragraph here regarding the IN parameterizations. This way, the previous paragraph (L18-25) reads more like a general introduction of atmospheric INPs and their diversity (biological and non-biological).

P3 L19: The authors may mention that the abundance of quartz in atmospheric dusts is consistently high (i.e., ~10% in volume) in the size range of ~1 to 35 µm geometric diameter (see Table 1 of *Kandler et al*., 2009, Tellus; cited in this paper), which would add the atmospheric relevance and general importance of quartz. Such information could also fit in P4 L1-3.

<u>Methods</u>

P5 L26-27: Does the size distribution of particles in the chamber change over 3.5 hours (i.e., the filter sampling period; P8 L15-16)? Large particles settle down faster than the smaller ones, and the authors infer that certain mineral compositions are large in their sizes (e.g., P15 L33). Please clarify and discuss potential consequences, if any.

P5 L28: This background (~10%) seems high for the IN research. Any justification or measurements that this much background particles have no substantial contributions to heterogeneous freezing?

P5 L30: Briefly describe "coincidence effects" or cite proper papers.

P7 L2: According to Table 3 of K16 (*Kaufmann et al*.), the Israel sample contains some sanidine. Moreover, sanidine in the natural surface samples seems non-negligible as 6 out of 12 natural surface samples examined in K16 shows the presence of sanidine. Was your sample totally different from K16? Or the the sample may not be completely homogeneous in terms of mineralogical distribution even within a same batch? Was the XRD data interpretation method somewhat different from K16? Please clarify.

P7 L20-22: Doesn't this just mean that a majority of large particles (up to 32 µm) break up by the RBG milling? I mean that RBG may do more than just aerosol dispersion, correct? Long story short, is it really fair to assume that those aerosolized particles are identical to the sieved bulk used for XRD as the authors mention in P13 L20-23? Further clarification seems necessary. Accordingly, the authors may consider rephrasing the relevant text in the conclusion (i.e., P15 L29-32). 2.5 µm sounds like a magic number as it is right now.

P7 L28-30: For the reason given above, I am not sure if the authors can asset like this.

Results & Discussion

P10 L3: The homogeneous freezing regime presented in this manuscript is based on CNT? Or anything different? I suggest adding proper reference(s) here at least.

P10 L9-10: I encourage the authors to clearly state that the same trend holds true for another metric, $n_s(T)$.

P11 L2-3: It looks to me that the Niemand parameterization falls in the middle of your 15 $n_s(T)$ spectra. Please clarify what "rather at the lower end" means quantitatively.

P11 L13-14: The word "overpredicts" implies that the Niemand parameterization (N12) is wrong. The authors may want to soften the tone and rephrase. Your assumption vs. assumption made in N12 would be discussed here.

P11 L11-20: I do not find the scientific significance of having FRIDGE data included in this manuscript. The authors briefly discuss about the FRDIGE results in this part and only a bit more afterward. The reviewer does not find that the FRIDGE results are complementing IMCA-ZINC-IODE, vice versa. In addition, no proper justification for why FRIDGE was conducted for a subset of samples is provided. Does the authors' conclusion change without the FRIDGE data?

Fig. 5: Adding quartz reference spectrum (*Atkinson et al*., 2013) would be nice.

Fig. 5 Cont'd: Having another panel depicting the highest-median-lowest spectra of airborne $n_s$ vs. those of surface-collected $n_s$ would add some clarity and strengthen the paper.

P12 L4-9: The discussion given here makes the review think something other than minerals (e.g., P5 L15-17) are competing for IN at given *T* range. The reviewer is aware that the focus of the current work is on mineralogy vs. IN (and no biological INP perspective at all). That said, the authors should extend the discussion regarding the potential bioaccessibility of dust surface (*Augustin-Bauditz et al*., 2016, ACP; *O'Sullivan et al*., 2016, ACP) and other IN species that might be present in soil (*O'Sullivan et al*., 2014, ACP; *Tobo et al*., 2014, ACP; *Pummer et al*., 2015, ACP). Such information would strengthen the manuscript.

Conclusion

P15 L18-19: Which data infers the *FF* and $n_s(T)$ results at temperatures above 250 K? I do not see them in Fig. 4. FRIDGE?

P16 L2-4: Your data presented in Tables 2 and 3 (i.e., Atacama milled vs. sieved; Israle milled vs. sieved) seem contradicting to this statement. For instance, I do not see any increase in the quartz fraction.

P16 L6-8: This part seems contradicting to the previous statement (that is, P15 L33).

**Technical comments/suggestions**

P1 L1-2: "Traditionally, clay minerals were assumed to determine…" → "Since natural dusts are composite in nature, clay minerals were typically used as a proxy to determine…"

P1 L14: "…activity in a given sample above 253 K that can be attributed to…"

P1 L17 and hereafter: Use the Italic font for *T* throughout the manuscript.

P2 L3: "determine" → "influence" - Besides primary ice nucleation, secondary ice processes can also contribute to the lifetime of clouds.

P2 L20: *Pummer et al*. missing as INM references

P2 L21: *DeMott et al*. missing as soot IN references

P2 L32: ", respectively"

P2 L32-P3 L1: Too many things packed in a single sentence. Split into two sentences.

P2 L33: between → amongst

P3 L13: "K-feldspars (microcline, orthoclase and sanidine) were…"

P3 L20-22: Redundant (P3 L4-6)

P3 L27: Delete "hence". Size and composition are inherently related. Is that what the authors want to say? If so, state so.

P3 L31: No "hence"

P4 L29: "The GPS coordinates of our collection sites…"

P5 L15-16: "The composition of all natural samples are presumably heterogeneous…"

P6 L10: "… sample (see …"

P6 L15: "… (sieving/milling)"

P8 L5: “The cloud droplets” → “The simulated cloud droplets” or “The activated droplets”
P8 L15: “subsequent” → “independent”
Eqn. 4: Missing negative sign on the RHS of Eqn.4?
P9 L4-5: Awkward sentence. I suggest rephrasing.
P9 L14-15: Awkward sentence. Simply say something like our measurements are valid with $\chi$ in the range of 1.1 to 1.6.
P9 L21-22: Awkward sentence. Simply say …because of the presence of predominantly large particles… or something similar.
P 10 L12-13: “Due to the heterogeneous and possibly size-dependent particle compositions, a partial step function like activation spectrum could be…”
P10 L27-28: two consecutive “overalls” are bothering.
P11 L2: “The comparison between the Niemand parameterization and our parameterizations based on…”
P11 L5: “polydisperse nature” → “heterogeneous properties”
P11 L23-24: This sentence does not fit in here. I suggest deleting.
P14 L34: The authors may want to remind the reader that 2.5 µm is $D_{50}$ of your cyclone.
P15 L10-13: “However… Thus…” - Awkward transition. Rephraze.
P15 L22 above or below?
P15 L27-28: I do not understand this sentence. Rephrase.
P15 L29-32: Awkward sentence. Rephrase.
P16 L 12: “can not” → “cannot” (to be consistent with the rest of the manuscript)
P16 L 21-22: Redundant (P4 L22)