Response to referee 1

We thank the referee for their helpful comments. We have attempted to improve the paper as recommended, the reviewer comments are italic, our response is in normal type and major changes to the manuscript are reproduced in bold type.

Major comment

'However, the idea of cold stage experiments is quite old. Cold stages have been used for decades (e.g., there is a very old paper from Roberts and Hallett, 1967). And even if the cold stages differ from each other in several details, they are always based on the same fundamental idea. What I missed here is a discussion what's really new. Furthermore, experimental data measured with this instrument are already presented in previous papers (as also stated in the text there is a very high ranking level paper presenting K-feldspar data). Except the detailed discussion of the setup (and maybe several Snowmax and AgI data) there is nothing new in the manuscript. I was wondering what was the reason to write this separate paper. The authors should motivate this much more.'

The major comment from the reviewer is that the cold stage instrument is an old idea and that it may therefore not be worth a separate paper to describe the technique. While we agree that cold stages (and the droplet freezing technique) have been used in the past, we think that it is important to properly describe the details of the specific experiment used. The µL-NIPI has already been used extensively by our group and will continue to be used in the future, but we have not yet described the technique in any detail, nor have we discussed the potential pitfalls of the technique. Importantly, it took us a lot of time and effort to iron out problems and artifacts, and this paper will help others in this rapidly expanding field with design of their own similar experiments in the future.

The reviewer highlights exactly why this technique needs to be described carefully. Not all cold stage experiments are the same. In fact, the paper by Roberts and Hallett the reviewer mentions describes a rather different technique in which ice crystals are grown directly from the vapour phase to yield ice-onset relative humidities.

In order to stress the importance of this paper, we have modified the final paragraph of the introduction:

'In this paper, we present a method of conducting a droplet freezing experiment using microlitre scale droplets. While the principle of the technique is not new, the specific application of the technique is and the aim of this paper is to document the equipment, methods and analysis associated with the technique. This instrument, the microlitre Nucleation by Immersed Particle Instrument (µl-NIPI), is based around a Stirling cryocooler which cools a hydrophobic surface supporting microlitre volume droplets and the freezing of the droplets is monitored using a digital camera. We present new data for the nucleation efficiency of K-feldspar (microcline), kaolinite, chlorite, Snomax®, and silver iodide. This data is used to illustrate several potential artefacts and how to avoid them.'

Specific comments and suggestions

abstract*: I would recommend to shorten at least the first part of the abstract. I think it would be enough to give most of the information (radiative properties of clouds, ...) in the introduction section.*

The abstract has been abridged, the first sentence has been removed and the second reduced.

introduction: *The authors try to highly motivate their experimental concept. Generally, this is totally okay. However, my feeling is that they overdid it. By reading the introduction one can get the impression that experiments with cloud sized droplets are completely senseless and not needed. But this is absolutely not true. Droplet freezing experiments with microlitre sized droplets are a very good complement to those with cloud sized droplets, but cannot replace them. On the one hand, using bigger droplets allows the determination of much smaller ns values. But on the other hand, (overlapping) experiments with cloud sized droplets are still needed to show the connection to atmospherically relevant droplet sizes (i.e. relevant amounts of dispersed material). This has to be done because in terms of nucleant surface area the experiments with big droplets (as done at mL-NIPI) are not atmosperically relevant.*

It was certainly not our intention to give the impression that cloud sized droplet experiments are completely senseless and we state in the summary that different approaches are complementary. We have changed the relevant section to make this clearer. It now reads:

'In general, instrumentation employing single aerosol particles suspended in gas and droplet experiments working with cloud sized droplets (10's of micrometers) report values of n_s down to about 10^3 cm^{-2} (e.g. see Fig. 18 of Murray et al., 2012 for a compilation). These measurements are clearly valuable and applicable to the atmosphere, but even smaller n_s values are also relevant (Murray et al., 2012). For example, if we consider a dust influenced atmosphere with 5 dust particles per cubic centimetre with a mean radius of 500 nm, then in order to generate 10 ice crystals per cubic metre an n_s of only 60 cm^{-2} would be required. Hence, it is important that we have the capacity to measure n_s smaller than 10^3 cm^{-2} in addition to the capacity to measure larger values.'

p. 9513, l. 19: *The authors claim that experiments "can be performed with excellent reproducibility". Hovever, this has not been proven in the paper. It is shown that there is a spread of 2-3K for the pure water samples. Furthermore, there is the information included that melting points are average values of 5 measurements. But that's it. I guess, the data shown in the figs. are single experiments. If not, standard deviations should be included. The authors should discuss the topic of reproducibility a bit more.*

We overstated this perhaps. We have removed the statement. We do however think reproducibility is strong. Figure 5 shows that freezing temperatures for an identical sample are almost identical on repeated runs. We have also produced repeated data on ice nucleation by NX illite and multiple experiments with K-feldspar here and in Atkinson et al consistently fall on the same line in n_s. We have inserted the below into the test experiments and analysis section.

'The data in Fig. 7 also illustrate the reproducibility of the derived n_s values with this technique, with repeat runs for kaolinite and chlorite being identical within uncertainty.'

p. 9514, l. 8- *"40 to 50 droplets". This seems to be not true. All experiments shown in the manuscript were done with less than 40 droplets.*

We have amended this line. Experiments presented in this paper are somewhat older experiments- since then we have improved our pipetting skills and tend to get more droplets on a slide. The new data inserted into the paper reflects this.

p. 9517-9518 *The authors discuss the formation of Frost and the resulting artefacts. They also present a solution how Frost growth can be avoided. However, I'm not really satisfied with the phrase "largely eliminated". What does this statement mean? Is there still Frost growth or not? Furthermore, can "largely" transformed into "completely" by using an even higher flow of N2? In other words, what is the motivation to use a N2 flow of 0.2 l/min (increasing droplet evaporation at higher flows?, cooling/heating by the dry flow? Is the N2 flow temperature controlled?)? I was also wondering, why there was no Frost growth problem in the experiments discussed in section 3.2. In 3.1. it was shown that Frost growth can be observed even at -5◦C. By turning the dry flow off there should be also Frost growth in the experiments shown in Fig. 5, right?*

Some small amount of frost growth is still observed in the immediate vicinity of a frozen droplet, taking the form of a very narrow ring around the droplet, but it does not extend far enough to interfere with neighbouring droplets.

The section has been modified to make this clearer:

'Our example illustrates that the use of a 0.2 l min^{-1} flow of dry gas eliminates the problems of frost haloes and frost growth. The condensation halo, consisting of very small droplets, evaporates rapidly in low humidity conditions and frost growth is reduced to the extent that it does not impact neighbouring droplets. Some small amount of frost growth is still observed in the immediate vicinity of a frozen droplet, taking the form of a very narrow ring around the droplet, but it does not extend far enough to interfere with neighbouring droplets.'

p. 9520, l. 15: *The other substances (Snomax, AgI) should be included in the discussion of Fig. 7.*

Snomax and AgI could not be included in figure 7 since we do not think that a specific surface area can be used for aqueous suspensions of this material. This is because, AgI is sparingly soluble and therefore the surface area changes in suspension. Snomax® is a material which contains some ice nucleating proteins mixed with various cell materials including soluble material hence a BET derived surface area would not be relevant for an aqueous suspension/solution.

p. 9521, l. 7-19: *Here, the authors discuss the influence of "impurities". I don't get the argument why this is particularly important for large droplet experiments. Small amounts of high efficient nucleants would also dominate experiments with small droplets. The advantage of experiments using small droplets is that they usually observe much more droplets, which makes it statistically more easier to separate the impact of several nucleants. This section should be rewritten being more precisely what is really meant here.*

The argument is that it takes far smaller levels of contamination for the contaminant to dominate freezing with drop assay experiments using large droplets. We have added the following to the mentioned section.

'**In this example it would easy to attribute the ice nucleation activity observed to kaolinite, when in fact it all comes from the feldspar. If the aim of a study is to further fundamental understanding of ice nucleation by a particular material it is important to recognise the potential impact of small impurities when working with larger droplets. While it is true that such an issue might also be observed with systems investigating smaller droplets, the smaller the amount of material in a droplet the less chance there is of a contaminating particle being present. Hence, there is a trade-off between measuring smaller n_s values and increasing the risk of contamination by minor components.**'

p. 9531-9536: *The quality of the figs. is okay but can be further improved by being a bit more consistent in terms of scales, labels, symbols and colors. For example, I suppose that the K-feldspar data in Fig. 5 and Fig. 6 are the same. Why not using equal colors and symbols?*

We have improved the figures as suggested

p. 9535, Fig. 7: *The uncertainty in the surface area is estimated by +/-15 percent. The authors should give some information where this number comes from. The authors give a number for the uncertainty of the droplet volume (p. 9514, l. 12), but there seems to be no connection to the estimated uncertainty of the surface area.*

We have inserted the following into section 2.1.

'Uncertainty in surface area per droplet has been estimated at around ±15% by propagating uncertainties in weighing of samples, the BET surface area and the volume error of the pipette.'