Reply to Referee #3 (Anonymous):

We appreciate your constructive feedback. After each comment, our response is given in italics, with revised text in red and unchanged text for context in blue.

Pg3 L73-74: "However, due to the Stefan-Boltzmann law infrared thermometry at low temperatures is limited to large droplets." I agree with this point but wonder if a limit be calculated as to the size of droplet / temperature range accessible? This is an interesting point, and this type of calculation or some sort of limiting value would add useful context if possible.

An exact minimum size is very equipment dependent. Our statement was based on the example of IR-NIPI developed by Harrison et al., which uses 50 µl aliquots. However, TINA developed by Kunert et al. can detect droplets as small as 0.1 µL freezing. In the case of TINA they do not detect the temperature via IR thermometry, only the latent heat released.

We modified the sentence at line 73:

However, due to the Stefan-Boltzmann law infrared thermometry at low temperatures is usually limited to large droplets (Harrison 2018) although the latent heat released by droplets as small as 0.1 µL freezing has been reported (Kunert 2018).

Harrison, A. D. et al., (2018), Atmos. Meas. Tech, 11, 5629–5641. https://doi.org/10.5194/amt-11-5629-2018

Kunert, A. T. et al., (2018), Atmos. Meas. Tech, 11, 6327–6337. https://doi.org/10.5194/amt-11-6327-2018

Pg5 L159: What would be the smallest droplet measurable by this technique? How sensitive are the films to the latent heat released? Do you anticipate a signal / noise issue at smaller droplet sizes?

The smallest droplet size detected during testing was 0.2 µl. However, for this droplet size we came up against the non-uniform sensitivity of the PVDF we used in our experiments

We have added the following at line 170:

The spatial variation in pyroelectric coefficient also means that although droplets smaller than 1 µl could be detected in places, in order to guarantee detection across the whole surface the minimum size was set at 1 µl.

Pg6 L174: Please expand on this method. What was the temperature that the sample was heated to? For how long? How was it rapidly quenched? How was it milled? As it stands, it would not be possible to repeat this process based on the information given.

We have replaced the first paragraph at line 172 as follows:

The crystalline K-feldspar comes from the Bureau of Analysed Samples (BCS-CRM No. 376/1), as used by Atkinson et al. 2013. ~~No further processing of this sample was done~~. The sample was crushed in a ball mill with agate balls before being sieved using a fine mesh (aperture size 20 µm).

The glassy K-feldspar sample was made from the crystalline sample melted in a platinum crucible. It was held at 1250°C overnight to remove moisture, before being heated to 1600°C for two hours. After this, the sample was removed from the furnace and allowed to quench in air. A few sections of the glass formed were examined under a polarizing microscope and no birefringent regions were observed. The glassy sample was then crushed and sieved using the same method described for the crystalline sample.

Pg6 L179: Feldspar is susceptible to acid ageing. Did you measure the pH to ensure that the acid was removed by the Milli-Q rinses? How much rinsing was done?

We did not measure the pH after rinsing, however each sample container was thoroughly rinsed at least 6 times. The following paragraph has been added to the end of the sample preparation section:

Feldspar materials are susceptible to surface changes in aqueous solutions (Lee and Parsons, 1995) and when exposed to extreme pH (Kumar et al., 2018), which could lead to a change in their ice nucleating ability. Peckhaus et al. (2016) measured a 2°C decrease in freezing temperatures of K-feldspar stored in aqueous solution for five months. However, Kumar et al. (2018) recorded no change in the ice nucleating ability of crystalline K-feldspar after one week in water suspension and Harrison et al. (2016) noted no significant changes in freezing temperatures of crystalline K-feldspar due to time spent in water suspension. We assume that any aging of K-feldspar in aqueous solution is sufficiently slow to not have an effect on our results. Due to the identical chemical composition of the glassy sample we assume that any aging effects are similarly slow.

Pg6 L183-185: "The values were 5.0±0.7 m2g-1 and 1.8±0.4 m2g-1 for crystalline and glassy K-feldspar respectively. The percentage errors associated with the surface area per unit mass dominate the error in calculating surface area present in each droplet, but are comparable to other experiments." I think this statement needs to be evidenced with citations. In some articles, the uncertainties in surface area per drop are not reported separately, so it is difficult to assess their magnitude. This is particularly true when the Poisson error is said to be the dominant source of uncertainty. The evidence you provide here demonstrates that this is important to account for, particularly when considering the sensitivity of ice nucleation measurements to surface area uncertainties (Alpert and Knopf, 2016).

We apologize because the line comparing to other experiments was intended for the values themselves, not the errors. We have corrected the text as follows to make this clear:

The surface area of both samples was measured via Brunauer-Emmett-Teller (BET) nitrogen gas absorption. Three repeats were taken, with the mean to extreme range used as the error. The values were 5.0 ± 0.7 m^2g^{-1} and 1.8±0.4 m^2g^{-1} for crystalline and glassy K-feldspar respectively, which are comparable to other experiments. The percentage errors associated with the surface area per unit mass dominate the error in calculating surface area present in each droplet~~, but are comparable to other experiments~~.

Pg6 L185: The Poisson error is mentioned in the SI, but it is not propagated into the uncertainties (which instead use the surface area uncertainties). Can the authors comment on the calculation that led to the "exceeds 10%" comment in the SI and why these errors are not included?

The analysis of the data has been changed significantly. In order to include the score confidence interval, suggested by referee no. 1, the data has been binned into 0.5°C bins. An example histogram with stochastic uncertainty is shown in the updated supplementary information. The stochastic/binomial errors were then propagated with the surface area errors. We find that the stochastic error dominates at higher temperature, while the surface area error dominates at lower temperatures. Also, since this analysis was done we have started using an improved method to fit the liquid proportion curves which is also detailed in the supplementary information.

Figure 5 has been updated to reflect these changes.

The results and discussion section has been modified to reflect these changes (line 200):

~~The dashed lines are generalised logistic functions (see supplementary information). These have no basis in theory, but provide good, monotonically decreasing, lines of best fit which can be differentiated analytically.~~ The solid lines are taken from a fit assuming the liquid proportion curves follow a non-homogeneous Poisson process, referring to the fact that the rate constant is changing as a function of temperature. A full derivation of the fit can be found in the supplementary information.

A brief error section has been added to the supplementary information as follows:

Stochastic errors were estimated using the Wilson score confidence interval on each temperature bin with more than one nucleation event. An example is shown in Figure S1. The errors in the liquid proportion were calculated based on how the minimum and maximum number of freezing events would affect the liquid proportion at that temperature bin, assuming the mean number of events were seen in all higher temperature bins. These errors were then combined with the errors in INP area in the calculation of n_s and J using standard propagation techniques.

Pg6 L185: Should this list also include weighing uncertainties?

We have added them for completeness

There are also errors associated with the masses of K-feldspar and water when making suspensions, the volume of each droplet pipetted, ~~and~~ amount of material which settled out of suspension during pipetting (Tarn et al. 2018)

Pg 6 L165: If the nucleant has a much steeper gradient in fraction frozen (for example pollen, fungal or bacterial INPs), would there be an overlap of signal, and what would be the limit of the experiment? In other words, how many droplets freezing per second could this method distinguish. This limit may also be interesting to determine for potential application if it were to be used for many smaller droplets, for example in a microfluidics experiment.

For the data presented here the voltage was sampled at 1kHz, and this was sufficient to detect all droplet freezing events. However, the sampling rate could be increased by several orders of magnitude by using a higher bandwidth analogue to digital converter, so this is unlikely to prove a limiting factor.

We have added the following at line 142,

The output from the charge amplifier was monitored using an analogue to digital converter (NI USB-6002), sampled at 1 kHz, which is fast enough to detect all droplets freezing, without creating unnecessarily large data files. For INPs that freeze over a very narrow temperature range, the sampling rate for this analogue to digital converter could be increased to 50 kHz to reduce the chance of near simultaneous freezes not being detected as separate events.

Pg6 L193: Did you perform experiments at cooling rates other than 1°*C* /min? What would be the uncertainty based on literature thermal conductivity for different cooling rates?

Unfortunately, we do not have any data on different cooling rates. At 1°C/min we see a small negative offset in the voltage level, as the PVDF continually discharges through the feedback resistor. This offset would increase with cooling rate, however, as long as the rate of release of latent heat from a freezing droplet was greater we should still expect to see a spike in the voltage.

Pg7 L196: Given that the background showed some heterogeneous nucleation, did you consider accounting for the background from your data using the differential spectra? (Vali, 2019)

We investigated the background effect using a slightly different method. This has been described in the supplementary information as follows:

Each liquid proportion curve $P(T)$ is the probability that a droplet is still liquid at temperature T. The measured curve, $P_{meas}(T)$ is the product of the probability that there has been no ice nucleation event caused by the ice nucleating agent of interest $P_{INP}(T)$, and the probability that there has been no background event $P_{back}(T)$. Hence the background can be removed by dividing $P_{meas}(T)$ by $P_{back}(T)$ The influence of the background can be seen in Figure S2, where only the 1% glassy sample shows a noticeable change, although it is a maximum of 0.2°C.

and the following has been added at line 200:

The influence of background freezing events on the liquid proportion curve of each experiment was calculated (more details in the supplementary information), but in all cases the corrected curve lay within the temperature errors.

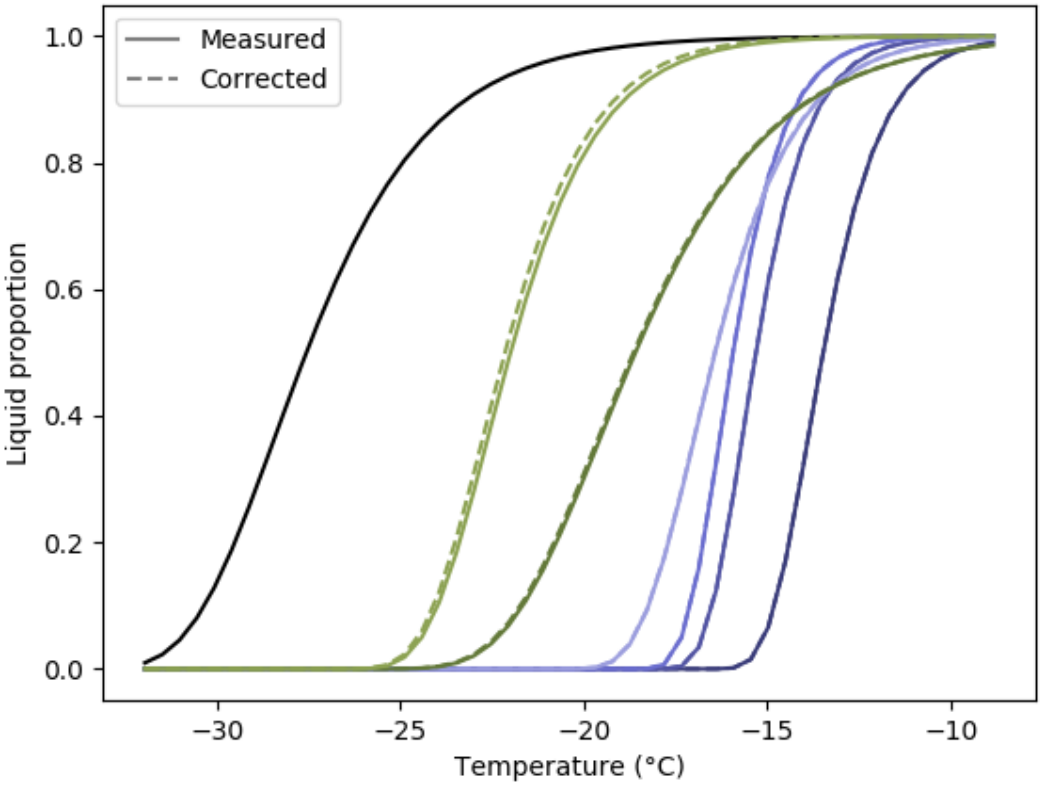

Figure S2. The background adjusted liquid proportion curve fits for pure water, crystalline K-feldspar and glassy K-feldspar. The colours are the same as Figure 5a in the manuscript.

Pg7 L208: Since you describe site specific nucleation, is it more appropriate to use freezing rate rather than nucleation rate as discussed by Vali? (Vali, 2014)

We have addressed this in combination with the next comment.

Pg7 L222: "The value of jhet(T) found for glassy crystalline K-feldspar here represents the average particle." Can you please clarify what is meant by the average particle here? Would the

heterogeneity of the sample not bias the jhet(T) measurement? (Herbert et al., 2014;Holden et al., 2019)

Line 222 has been modified and extended as follows:

The value of j_{het} found for glassy and crystalline K-feldspar here represents the freezing rate (Vali 2014) divided by the surface area measured by BET ~~the average particle~~.

Figure 3: It is made clear in the figure caption that positive spikes represent freezing events. What is the source of the negative spikes? If this is an artefact, is the same artefact possible with positive values (i.e. recording false positive signal)?

We are unsure what causes the negative spikes, possibly thermal contraction of the PTFE clamp holding the PVDF to the cooling stage. Since the number of freezing events detected always matched the number of droplets in any given experiment, we do not believe that false positives are a problem..

Figure 5: Whilst I agree that not displaying all error bars helps with clarity, I think it would be helpful to add more than just the first data points for ease of interpretation(perhaps at 25%, 50%, 75% and 100%?)

As the data has now been binned into 0.5°C bins we have included error bars on every data point.

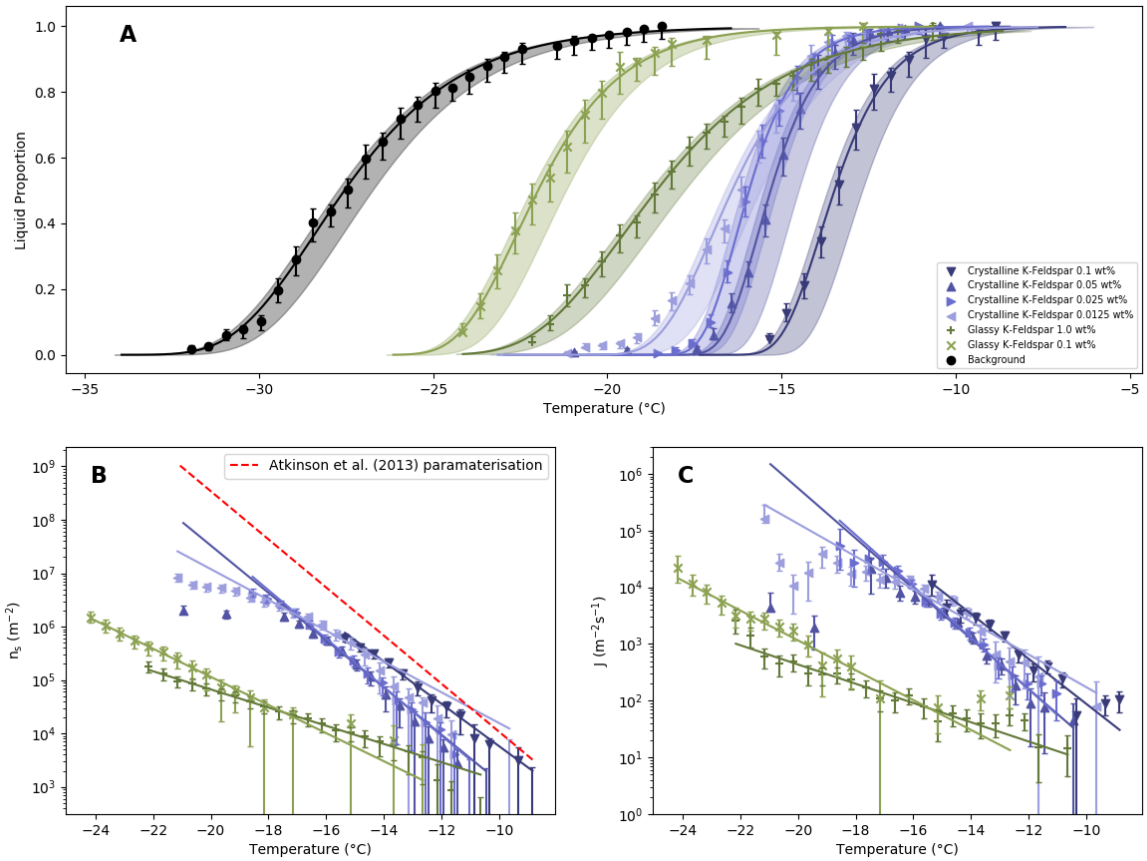

Figure 5. **A.** Liquid proportion ~~Frozen fraction~~ curves for 1 µl droplets of water containing different fractions of glassy and crystalline K-Feldspar. The background freezing rate of the instrument is also shown~~ as the pure water frozen fraction~~. Temperature errors are shown by the shading. Details on the lines of best fit can be found in the supplementary information. **B.** Ice nucleation active site density, normalised by the surface area present in each droplet. The red dashed line is the parameterisation from Atkinson et al. (2013) which is partly based on microlitre sized droplets with similar concentrations to those used here. **C.** ~~Nucleation~~ Freezing rates normalised by the surface area present in each droplet. ~~The nucleation rate calculated from classical nucleation theory. Only the error bars of the first datapoint of each sample are shown for clarity.~~

Figure 5b: As the sample used is the same feldspar as Atkinson (2013), perhaps it would be helpful to display the parameterisation from this paper, so that the performance of the new stage can be compared to that of the optical cold stage they used

This has been included in the updated Figure 5b.

Figure 5b: Can the authors comment on the variation in ns for different wt% suspensions? In particular, there seems to be an offset between 0.5 wt% / 0.25 wt% and 0.1wt% / 0.0125 wt% for crystalline feldspar. For example, at ns = 104 the uncertainties displayed do not explain the differences in the data. Is this expected based on the uncertainties in ns? Or could this relate to the length of time suspensions were kept for before experiments?

We suspect the discrepancy is due to the inherent randomness of these types of experiment. The crystalline suspensions were made simultaneously, with experiments conducted from the highest wt% to the lowest in the timespan of a week, which would discount correlations due to aging in

suspension. The difference is smallest towards the centre of each dataset, where the errors are the smallest.

SI: Is [64] a reference? If so, please correct to AMT format

The supplementary information has been largely re-written and no longer contains this reference.

General comment: The spaces between sentences and references are inconsistent (sometimes there is a space and sometimes there isn't).

This will be corrected to consistently have a space.