Supplementary Information

S1 Supplementary figures

Figure S1 shows the individual unaged and water aged Fuego (FUE) ash ice nucleation (IN) spectra. Figure S2 shows the result of applying the same type of analysis of the different methods of calculating empirically generated confidence bands for the water aged FUE spectrum, as opposed to the unaged FUE spectrum shown in the main text. Figure S3a and S3b show the effects of changing the number of droplets in the original experiment and the number of simulations used the generate confidence bands for the water aged FUE spectrum respectively.

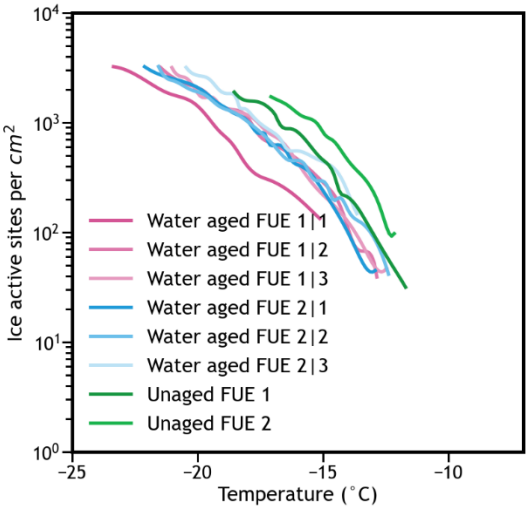

Figure S1. Individually interpolated cumulative ice nucleation surface active density spectra of unaged and water aged volcanic ash using the smoothedPCHIP algorithm

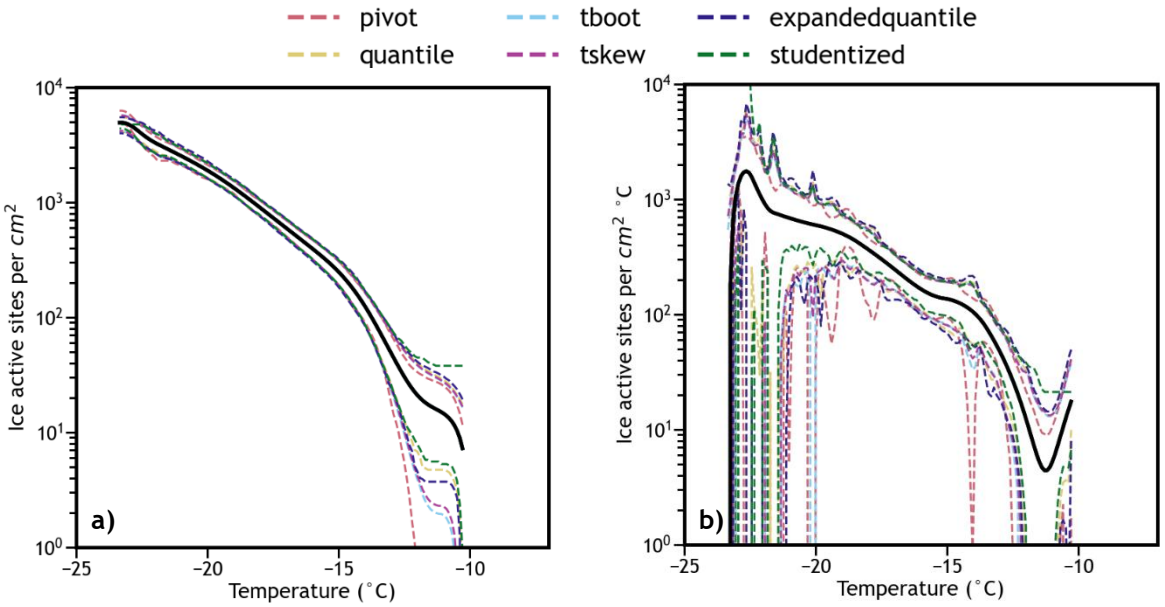

Figure S2. Comparison of methods to calculate CIs on the a) cumulative and b) differential combined water aged ice active site density spectra representing 286 droplets. The interpolated experimental spectra are shown in black.

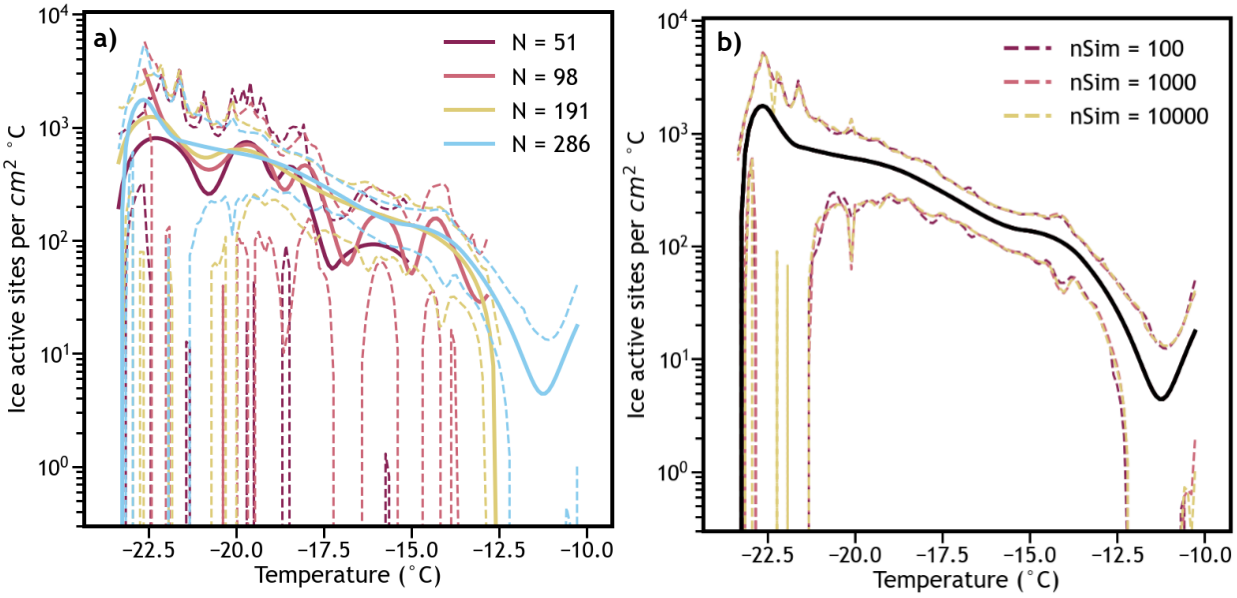

Figure S3. Freezing spectra with tskew CIs calculated with varying numbers of a) datapoints with 1000 bootstrapping simulations, and b) bootstrapping simulations, with 50 datapoints with the experimental spectra shown in black.

S2 Bootstrapping algorithms and confidence interval calculations

In this section, the algorithms and equations used to generate simulated experiments, boostrapped statistics, and confidence intervals/bands are described in plain text. When discussing ice active site density spectra, any cumulative or differential spectrum can be used (and in fact, are all used in the provided implementation of these methods), however, for generality and simplicity, 'ice active site density spectrum' or 'IN spectrum' will be used here.

To generate empirically simulated ice nucleation experiments, the following steps were taken:

1. The observed freezing experiment (whether combined or not) was reduced into a list of freezing temperatures T_{obs} = [T_0, $T_{1, \ldots}$ T_i] where i is the number of droplets that froze in that experiment. Temperatures are repeated if multiple droplets froze at a single temperature.
2. A new list of temperatures, $T_{sim,}$ is generated by randomly choosing an element of T_{obs} i times with replacement and adding each one to the new list.
3. Step 2 is repeated j times, where j is the number of desired simulated experiments, usually with j >= 1000.
4. Each new list of temperatures, $T_{sim,j}$ is separately sorted, converted into a frozen fraction, inverted to calculate ice active site density spectra, and interpolated using the desired approach to produce $f_{sim,j,}$ a function which takes a temperature and outputs the corresponding ice active site density at that temperature.

To calculate statistics using these new simulated spectra:

1. The temperature range of the observed freezing experiment is divided into a grid of x temperatures. In this paper x was chosen to be 10 times the difference between the highest and lowest freezing temperatures observed.
2. For temperature x, a list $\hat{G}_x$ consisting of $f_{sim,j}(x)$ for all j is generated – the list of possible simulated ice active site densities at that temperature.
3. Step 2 is repeated for all x
4. For all x, the desired statistic is calculated on the list $\hat{G}_x$ using elementary statistical equations (common statistics include mean, standard error or standard deviation of the mean, and quantiles).
5. The statistic is then interpolated over all x to provide a continuous bootstrapped estimation of the statistic.

A variety of methods for calculating confidence intervals of the mean are provided in the main text of this paper. Equations S1-S8 provide mathematical formulations for these confidence interval calculations based on bootstrapped statistics calculated as above. Confidence intervals are calculated like statistics on the simulated spectra and are similarly interpolated to provide full confidence bands. Notation and formulations are taken from Hesterberg (2015). In some cases, due to the nature of ice nucleation experiments, an experimental standard error or mean is not available (see Section 4.1 in the main text). When these statistics would normally be used in calculating bootstrapped confidence intervals, we have substituted the bootstrapped versions as the bootstrapped version of any given statistic can be interpreted as an estimation of the sampled version of that

statistic. When a confidence interval is symmetric, it is represented using the ± symbol, and where it is not the form (*a*, *b*) is used, where *a* is the lower confidence limit and *b* is the upper confidence limit.

A simple bootstrap Z interval is

$$CI_{Z\ interval} = \mu \pm \hat{s} Z_{\alpha/2} \tag{S1}$$

where μ is the experimental value for the IN spectrum at *x*, $\hat{s}$ is the standard error of the mean of the bootstrap distribution $\hat{G}x$, $Z\alpha_{/2}$ is the z-score at alpha level $\alpha/2$, and α is the confidence level desired.

A quantile confidence interval is

$$CI_{quantile} = \left(q_{\alpha/2}, q_{1-\alpha/2}\right) \tag{S2}$$

where q_α is the α quantile of $\hat{G}_x$.

The pivot (or reverse quantile) confidence interval is

$$CI_{pivot} = \left(2\mu - q_{1-\alpha/2}, 2\mu - q_{\alpha/2}\right). \tag{S3}$$

The bootstrapped t interval is

$$CI_{t\ interval} = \mu \pm \hat{s} t_{\alpha/2,n-1} \tag{S4}$$

where $t_{\alpha/2,\ n-1}$ is the t-score for alpha level alpha level $\alpha/2$ and *n* is the number of droplets in the original freezing array.

The skew-corrected bootstrap t interval based on Johnson (1978) is

$$CI_{skew-corrected\ t\ interval} = \mu + \kappa(1 + 2t^2_{\alpha/2,n-1}) \pm \hat{s} t_{\alpha/2,n-1} \text{ with } \kappa = \frac{\widehat{skew}}{6\sqrt{n}} \tag{S5}$$

where $\widehat{Skew}$ is the skewness of $\hat{G}x$.

The expanded quantile confidence interval based on Efron (1987)

$$CI_{expanded\ quantile} = \left(q_{\alpha'/2}, q_{1-\alpha'/2}\right) \text{ with } \alpha'/2 = \Phi\left(-\sqrt{\frac{n}{n-1}} t_{\alpha/2,n-1}\right) \tag{S6}$$

where Φ is the normal cumulative distribution function.

Finally, the studentized confidence interval (or bootstrap t interval; (Diciccio and Efron, 1996 and references therein; Efron, 1979) is

$$CI_{studentized} = \left(\mu - \hat{s}\hat{q}_{1-\alpha/2}, \mu - \hat{s}\hat{q}_{\alpha/2}\right) \tag{S7}$$

where $\hat{q}_\alpha$ is the α quantile of a distribution $\hat{G}^*_x$, which is the distribution of the t-statistic for *each bootstrapped spectrum*, calculated as

$$\hat{G}_x^* = \frac{(\hat{\mu}_j - \mu)}{\hat{s}_j^*} \quad \text{(S8)}$$

where $\hat{\mu}_j$ is the *j*th bootstrapped value for the IN spectrum at *x*, and $\hat{s}_j^*$is an estimation of the standard error for $\hat{\mu}_j$. To estimate $\hat{s}_j^*$, another iteration of bootstrapping is required, which is performed the same way as the first iteration of bootstrapping, only using the *j*th bootstrapped simulation from the first iteration to generate the list of temperatures being sampled from for the second iteration. The sorting, analysing, and interpolating processes are the same from there, and calculating the statistics from the resulting resimulations is identical to calculating statistics for the first round of simulations.

References

Diciccio, T. J. and Efron, B.: Bootstrap Confidence Intervals, Statistical Science, 189–228 pp., 1996.

Efron, B.: Bootstrap Methods: Another Look at the Jackknife, Ann. Stat., 7, 1–26, https://doi.org/10.1214/aos/1176344552, 1979.

Efron, B.: Better bootstrap confidence intervals, J. Am. Stat. Assoc., 82, 171–185, https://doi.org/10.1080/01621459.1987.10478410, 1987.

Hesterberg, T. C.: What Teachers Should Know About the Bootstrap: Resampling in the Undergraduate Statistics Curriculum, Am. Stat., 69, 371–386, https://doi.org/10.1080/00031305.2015.1089789, 2015.

Johnson, N. J.: Modified t tests and confidence intervals for asymmetrical populations, J. Am. Stat. Assoc., 73, 536–544, https://doi.org/10.1080/01621459.1978.10480051, 1978.