

The manuscript titled “A universally applicable method of calculating confidence bands for ice nucleation spectra derived from droplet freezing experiments” by Fahey et al., demonstrates a method to derive ice nucleation (IN) spectrum and confidence bands for ice nucleation studies. They use this to determine an IN spectrum normalized to particles surface area, a quantity referred to as $ns(T)$, the number of ice nucleation events normalized to particle surface area as a function of temperature. Bootstrapping is used in this method to simulate freezing temperatures multiple times by random sampling and calculating $ns(T)$ from these simulated data. Variability in simulated freezing and $ns(T)$ is used to derive confidence bands. Finally, the authors are then able to deduce statistical similarity between datasets. Overall, this paper stresses in the importance of deriving uncertainties in $ns(T)$ and should be of great benefit to ice nucleation community.

Unfortunately, there are many instances when clarity and precision are lacking in describing the methods, as well as when the authors mix what can be stated as options or as claims with evidence. I cannot recommend publication until the comments below are addressed. The first major comment relates to poor writing, due to poorly explaining (or not explaining at all) complex topics/methods that are the main subjects of this manuscript. Another major comment is about assuming variability on the droplet by droplet basis in Eq 1 and 2. A final major comment is about the method for assessing the similarity of IN spectra is not explained at all. Although, the paper is quite novel, addresses relevant scientific questions, and has important conclusions. They demonstrate sufficient evidence through their simulations to support their conclusion.

Major Comments:

1. The article and especially the introduction has many vague terms, descriptions and opinions. I have outlined most in the minor comments below, however, a few major comments are described. Most notably the introduction is poorly written, with insufficient details about what the author will present and how it relates to previous work. The authors have not reviewed previous literature and what advances were made in ice nucleation simulation to lead them to this work. They are not the first to model or simulate ice nucleation. Instead, their intro is a narrative about the failure in the community. This is unacceptable and must be rewritten.
 - a. “ice nucleation processes” The authors must be more specific about what this process is. As a reader, the phrase “it is simply a fact of the ice nucleation process” is meaningless because these facts are not outlined or given in the manuscript up to this point. I read only that ice nucleation is a process, and it does not help understanding. Please be specific, what is the process.
 - b. “IN spectra”, “ $ns(T)$ ”, “ k ”, “ K ”, These are defined in many places throughout the first pages of the manuscript. Would the authors please consolidate a common and non-redundant definitions of these.
 - i. p. 3, l. 75: k and K are ice active site density spectra
 - ii. p. 3, l. 76: k is a differential spectra and K is a cumulative spectra
 - iii. p. 4, l. 81: K is the number of ice active sites at temperature T per unit of suspension volume. K can also be nm. K can also be ns.
 - iv. p. 4, l. 6-7: K are the cumulative ice nucleation active site density curves
 - c. p. 2, l. 44-46, This sentence insults the ice nucleation community. Yes, ice nucleation analysis is complex, however, the statistical tests, approximations, and methods of previous work are not inaccurate. They are peer-reviewed and explained. Please rewrite this sentence, emphasize previous work that has been done on which your manuscript is based, and how the authors fill a missing gap.
 - d. p. 4, l. 81: “...number of ice active sites...” This is a vague term. What is a site and how does Eq 1 and 2 give sites? In fact, eq1 and eq2 are using counted freezing events and

- there are no ice active sites in them. There are many assumptions made between Eq 1 and 2, and then making a claim about quantitative measures of ice active sites.
- e. The word “complex” is overused. On p. 5, l. 23, what make a curve complex and what makes a curve not complex? On p. 2, l. 55, what makes a dataset complex and not complex? On p. 2, l. 45, what makes ice nucleation complex and what makes it not complex? On p. 11, l. 80, math is complex. It appears that when the authors do not want to take the time to explain ice nucleation, IN spectra/curves on graphs, and bootstrapping, they call it “complex” and move on. This shows lack of care and work put in to write a comprehensive manuscript, as these complex things are the main subjects of the manuscript. Please search for the word complex in the manuscript and try to replace it with specific details and explanation. These should be briefly explained.
 - f. The words/phrases “superior”, “best”, “most powerful” are used through the manuscript. In all context, these are opinions of the authors. I suggest removing all.
2. Equation 1 and 2 has an important and unclaimed assumption. Why are the authors so sure that the normalization constant, X, is identical for each droplet? On l. 35-37, it states many differences that could cause variability from droplet to droplet. If so, these equations are not valid. If X were to be different for each droplet, then it would need to be accounted for. However, this equation inherently assumes some identical normalization. In other words, one constant for all liquid droplets. In other derivations besides Vali, such as in the Pruppacher and Klett textbook, eqs 7-64 and 9-56, the natural log appears due to the assumption that all drop volumes are the same, or surface area in drops is the same. The authors must claim this assumption in the manuscript, and that their analysis only holds if this assumption is valid.
 3. p. 15, l. 92 - p. 18, l. 97: This method descriptions is far too short. The authors claim taking a ratio or subtracting two different $n_s(T)$ curves is supposed to be between 0 and 1? That is the test of similarity. When taking a ratio or difference, the authors are doing this on a log10 scale? When they take the ratio, does the larger spectra need to be in the denominator? When taking the difference, the smaller values need to be subtracted from the larger values to get only positive numbers? The step-by-step procedure here is not so clear and so it is difficult to review Fig. 5 and any discussion around it. It would help to guide the reader a bit more slowly here. Finally, can the authors give a name or equations for “these metrics” stated p. 15 l. 91?

Minor Comments:

1. Abstract, p. 1, l. 15: “...and if used properly...” Would you please rephrase this? It reads as if there is a way to also improperly use your methods.
2. p. 1, l. 19: “By improving the statistical tools available, this work will improve...” Can you rephrase this? It reads like the authors still need to improve their methods. I do not think that is their intension.
3. p. 1, l. 19: “...improve the quality and accuracy...” Accuracy is improved with instrumentation that is more accuracy. Statistical analysis does not make instruments more or less accurate. Would you please state exactly what is more accurate here?
4. p. 1, l. 39: The authors write the word “this”, but do they refer to all of the uncertainties and variabilities in experimental investigation mentioned above? Or are they still talking about a perfect experimental setup?
5. Intro, p. 2, l. 33: Would the authors please precisely define “IN activity”? There should be a distinction between what is measured and what is derived.
6. p. 2, l. 42: “If we cannot eliminate experimental error, it must...” The word “we” is usually referred to the authors. I suggest to replace with “Experimental error is always present and must be ...”
7. p. 2, l. 43-44: It is the authors opinion that there is no widely accepted approach. Please remove this sentence, or write this is an opinion.

8. p. 2, l. 50: "...remaining questions..." What are the questions the author is referring? What are those questions that remain? Please be precise in formulating your scientific questions.
9. p. 2, l. 50: "...experimental parameters..." What are the specific experimental parameters the authors are referring to. Please be specific.
10. p. 2, l. 51: "...these approaches ..." What are the approaches the authors are referring to. Please be specific.
11. p. 2, l. 55: What is the difference between a toy IN dataset and an IN dataset. Maybe the authors would like to change this to read, "example of an IN dataset"?
12. p. 2, l. 61: please state the diameter
13. p. 3, l. 68-73: These sentences are redundant. Please rewrite.
14. p. 4, l. 99: Should "...temperature density of freezing events..." be "...number of freezing events depending on temperature..."?
15. p. 5, l. 49-40: This is a strange way to end. This is a strange way to end. Please tell the reader exactly what you are looking for, instead of telling them you are looking "elsewhere".
16. p. 6, l. 44 and l. 49: What is the difference between the terms "accurate interpolation", "faithful interpolation", and simply "interpolation". Can an interpolation be "unfaithful" or "inaccurate"? I can find no quantitative meaning of these in the paper. I suggest to remove "accurate" and "faithful". Please search through the manuscript for these.
17. p. 8, l. 83: "...are unreliable." should be changed to "...may be unreliable."
18. p. 8, l. 88-89: I tried looking up the reference Kaufmann et al. 2017, but it is not in the references list at the end of the manuscript. Would you please check that all references are actually included?
19. p. 10, l. 45: "...originally proposed..." What exactly was originally proposed? There are many things proposed up until this point reading the manuscript.
20. p. 10, l. 47: "...with replacement..." Would the authors please be specific about what this means? I thought it was a typo at first, but appears like it means something important to them.
21. p. 10, l. 61: "bye" is a typo.
22. p. 10, l. 62-65: This sentence is unnecessary, negative and offensive. Please remove it. It is affronting members of the ice nucleation community as unfamiliar with statistics and as unable to access information. This is not true.
23. Through the manuscript there is use of single quoted words and phrases. It is not clear to the reader why these have to have single quotes. Would the authors please elucidate the reason for this, or just remove the quotes and explain things clearly.
24. p. 11, l. 78: What is a "re-interpolation", and how can this be accurate or inaccurate?
25. p. 12, l. 4: Please write out what alpha is.
26. p. 15, l. 82: Please add commas so that the words "...method such as empirical bootstrapping rooted..." is changed to "...method, such as empirical bootstrapping, rooted..."
27. p. 17, l. 22-24: This is plenty of evidence that droplets on a substrate surrounded by oil or air can be used to measure homogeneous freezing. Yes, other studies have a background of heterogeneous ice nucleation occurring. Generalizing this to all substrate based approaches is not correct. Please remove this sentence.