Comments by Gabor Vali on "Using critical area analysis to deconvolute internal and external particle variability in heterogeneous ice nucleation" by Hassan Baydoun and Ryan C. Sullivan.

General comments:

This paper presents an extensive set of equations to represent expected freezing frequencies in drop-freezing experiments with given types of ice nucleating particles (INPs). The potential for the existence of nucleating sites of various potential activity is represented via the proxy of contact angle variations. This is accomplished with the introduction of the distribution $g(\theta)$, not by assuming that the nucleation rate J can take on a broad enough range of values to represent the observed experimental spreads in freezing temperatures (for more detail on this, see Section 4.8.3 of Vali et al. 2015; V15 for short). Combining the distribution of nucleating potential of sites, $g(\theta)$ or equivalent, with a nucleation rate J associated with each site is a good representation of how nucleation can be understood; this description (model) incorporates both the stochastic element and the site-specific factor; the same idea is also presented in Section 4.7.3 of V15, with references.

Eq. 7 is the general expression incorporating the ideas described above. It is re-written for a constant cooling rate experiment as Eq. 9 which is then used in the rest of the paper. Formally, the equations follows well-established principles, but need considerable additional constraints on what is meant by the various functions. On that point there are significant debates.

First application of the equations is to data on freezing temperatures with a single INP over many cycles. It is shown that it those data are better represented using a best-fit $g(\theta)$ than using a single value of θ. This conclusion is in accord with several other papers already referenced in the paper. No new insight is gained from this exercise.

The next step in the paper is to define "critical contact angles" and "critical area". The latter is intended as a tool to identify cases where internal versus external variabilities, newly defined in the paper, need to be considered for correct interpretation of empirical results. To this reviewer, the meaning of the external variability remains obscure (cf. questions below about lines 6-8 on page 7). These definitions lead up to the major part of the paper, namely the explanation of results obtained by Broadley et al. (2012) using different concentrations of illite particles. These results are examined in the form of temperature functions of the fraction of frozen drops, $F(T)$. Three experiments are chosen for close examination, one of which was performed with a lower particle loading. The pattern apparent with these experiments was also indicated (to lesser degree) by a somewhat larger set of experiments. The $F(T)$ curves for two of the selected runs are similar in shape but shifted in temperature, while the third one, for the lower particle load, also changes shape. Explanation of this pattern is the main focus of the paper.

The explanation (scheme or model) proposed in the paper, and applied to the case in question, has three major areas of shortcomings. First, several points in the scheme are poorly defined, are counterintuitive, and/or are inadequately explained. This reviewer was unable to form a clear view of the reasoning in many places; his doubts are detailed in the list of comments that follow. Second, the focus of the paper on just one set of experiments is very limiting, specially since doubts are expressed even in the source paper about possible artifacts causing the unexpected results[1]. Third, the proposed model is restricted to interpreting only one specific type of laboratory experiment[2].

It would be beneficial for the authors to first look at a wider set of data to see if similar patterns can be

1 Page 297 of Broadley et al. (2012): "It may be possible these high weight % droplets were not stable; as the concentration of clay-in-water suspensions is increased, flocculation and settling out of material can occur; hence, results from concentrated clay-in-water suspensions should be treated with caution."

2 Laboratory experiments with suspensions of different concentrations of INPs from the same source, cooled at a steady rate, are examined and modeled in this paper. As argued in Vali (2014), such experiments with dispersed samples (drops) are effective for characterizing the INP sources (clay, etc.) but represent only one of many types of experiments that are needed to understand ice nucleation. Only combinations of several different experimental approaches constitute critical tests of interpretations, theories, or models.

identified. Also, they would be well advised to consider alternative interpretations of the data in more detail than is evident from the paper. The theory proposed in the paper is not intellectually so attractive, in the form presented, as to make it of interest without clear explanations of what is meant by various new terms introduced, without showing success in quantitative interpretations of a variety of different types of data and without demonstrating improvements over other ways of examining the data.

The paper is well written, as far as style and language are concerned. However, it is excessively long and contains a number of unnecessary repetitions.

Even though it appears that this paper was written before the publication of Vali et al. (2015; Atmos. Chem. Phys., 15, 10263–10270), for the sake of easier communication the comments below employ some of the terminology introduced there.

Detailed comments:

page/line

6/6 -->	The wording "discrete ice active surface site" needs to be explained more fully. Are the sites surface features that are assumed to be unchanged with time, or are they formations that develop randomly on the surface due to chance? I have the impression that the authors mean the former. If so, it should be clearly stated.
6/12	It is incorrect to refer to sites as being infinitesimally small. For the sake of allowing an integration to be indicated instead of a summation, it is sufficient for dA to be a small fraction of the particle surface area.
7/6 - 7/8	Why would there be "differences in the *g* distributions" among particles of the same type? If it is because of their size differences, than they can differ because of the chance allocation of sites drawn from the same *g* distribution. Apparently you mean something different. Can you cite some reasons for why to expect that?
8/24 - 9/6	If the drop is kept at a constant temperature of 255.5 K, how is a distribution of freezing temperatures obtained, as shown in Fig. 1 with the dashed-line curve,? This plot extends over ~5 degrees in temperature? Please explain.
9/5 - 9/10	This is a prediction, with no empirical support. Right?
9/7 - 9/14	Why not test the calculations against the observed shifts in freezing temperatures with changes in the rate of cooling? Results from such experiments are described in Section 3.2.2 of V14.
9/24	The wording "ice active site activity" is not a fortunate description. Suggest changing to something else.
9/27 - 9/28	"distributions" here refer to the $g(\theta)$ function?
9/27 - 9/29	To which side of the Gaussian curve does this comment apply? Please rephrase this sentence.
10/4 - 10/8	Representing the distribution of sites of different potential activity as one site with a continuum of activity is very puzzling. I see no reason for doing this. Neither does it follow from the arguments presented about exponential dependence of freezing probability on J and exponential dependence of J

on temperature. Please elaborate both on why this is useful to understanding the model and why it is justified.

Fig. 2	The upper right inset and the second line of the caption are misleading and need to be corrected. The caption mentions " ... a representative effective ice active surface site " and the inset appears to indicate that th value of θ changes in concentric circles around a specific site. The histogram in the main part of Fig. 2 is a better representation of the information to be conveyed.
11/6	In Eq. 10 the right-most expression is approximately equal to the preceding expression with substantial differences for narrow range of integral limits. Thus, Eq. 10 cannot be "satisfied" - this sentence should be omitted.
11/6	What case is being depicted here? The red curve is not the same as that in Fig. 1. What J-function is assumed?
Fig. 3	Please indicate that $\theta_{c1} = 0$ is assumed for this diagram. Also the value of μ and σ that were used.
11/15 - 11/21	Is this example for case described in Section 3.1?
11/17	The estimate of site area is dependent on temperature and contact angle. The numerical value quoted should be referenced to the assumed values.
12/10	Reference (2012) is incomplete.
12/10 -->	The discussion appears to proceed as if particle count per unit volume of water was a single number. In fact different size particles exist in most cases, even when attempts are made to produce nearly monodisperse powders for laboratory tests. Thus, for the authors' argument to make sense, the mono-disperse assumption has to be stated, or saturation of external variability need to be achieved for all sizes (probably impossible in reality). Also, it is implied that all particles have identical chemical and mean surface properties. Thus, the treatment here given applies only to laboratory experiments in which particles of a given substance are added to the water. These assumptions should be spelled out.
12/25	What does 'one system' mean?
12/27	What is system i ? One particle?
13/16	This critical area notion is in contradiction with the monotonic decrease of $g(\theta)$ as θ approaches 0, i.e. in principle this critical area can only be reached with nucleation at the melting point (e.g. 273 K). If the lower limit $\theta_{c1} \neq 0$, the definition may make sense but remains of questionable practical meaning.
13/25 - 14/6	Again, experiments are mentioned without stating that a specific type of experiment is being discussed. This has not been clearly established in the foregoing. This is a serious constraint on the applicability of the scheme developed in the paper and need to be fully explained at least at the beginning of Section 3.3, specially since a different type of experiment in discussed in Section 3.1.
14/4	Can the authors spell out what they consider significant divergence?
14/15 - 14/16	This sentence is crucial to the view represented in the paper: " ... variability of active sites remains constrained within droplets." The authors view is focussed on the distribution of contact angles (as a proxy for real factors). This is expressed by talking about variability remaining constrained in the drops, i.e. an attempt to separate what they call external and internal variability. Diluting any sample containing suspended INPs and thereby the reducing the particle content per drop volume used in an

experiment has been found to lead to lowering of freezing temperatures in numerous experiments. This results in retrieving a different segment of the $n_s(T)$ or $k(T)$ spectra (Fig. 4 in Vali 1971 and many later examples). The data plotted as the fraction frozen versus temperature may or may not show a change in shape, depending on whether the slope of the $n_s(T)$ or $k(T)$ spectra happens to change over the observed range of freezing temperatures.

14/24 ".. green curve diverges ..." is an incorrect interpretation of the experiments discussed. It is not plausible for a well controlled experiment with a stable suspension of INPs to produce higher freezing temperatures (higher fraction frozen of higher n_s values) with a reduced particle content per drop.

15/1 - 15/5 These data should be presented.

15/8 - 15/15 This description is difficult to understand. How does the particle surface area influence the result from Eq. (16)? The total surface area of the particles within each drop is the parameter that is modeled, yet it does not appear in the description.What do you mean by optimizing the choice of n_{draw}?

16/7 - 16/8 This statement cannot be supported because of the limited scope of the evaluations made in this paper. It may refer to some apparent problems in the Broadley et al. (2012) paper to see how the data can be reconciled with the description based on surface site density. Shifts in the $F(T)$ curves with no change in shape is not a requirement at all for the applicability of the interpretation of observations in terms of $n_s(T)$ or $k(T)$ spectra. The note above for 14/15 - 14/16 explains this.

16/12 - 16/14 Is the g-bar distribution determined using Eq. (9)? If so, is the integral over contact angle applied as indicated (0 to π) or some smaller range? It would appear illogical, as it is also argued on page 10, to consider both the ascending and descending parts of the normal distribution. The details of this fit should be clearly described in the text for the process to be comprehensible to readers. The fit being determined for experiment (6a) is used for (6b) which has approximately factor 3.7 higher particle surface area. Thus, the frequency values extracted from $g(\theta)$ are reduced by about the same factor. While this is a fairly small factor compared to overall range of values needed to reproduce the freezing frequencies, it is important to know what part of the Gaussian curve comes into play.

16/29 - 16/30 Following the questions raised in the preceding two comments, is the random draw taken from the entire *g*-bar function, i.e. for $0 < \theta < \pi$?

17/10 - 17/26 Understanding of this paragraph is hindered by the use of expressions like "very active" when the model is constructed around the idea of a continuum of activities, albeit of different frequencies of occurrence. Similarly, 'leftover" drops goes counter to the model. There is no surprise in the fact that lower concentration of INPs lead to lower freezing temperatures. That there are a small numbers of freezing events at similar temperatures than for the higher surface area drops is due only to the relatively small change in the total surface area per drop. For any given temperature at the warm tail of the distribution the frequencies of these events can be expected to scale with surface area of INP per drop.

19/6 - 19/8 The criticism of pervious works for not having distinguished above and below "critical threshold" conditions sounds hollow, since the idea of critical threshold is introduced only in this paper. The real test is whether those previous treatments were successful, or not, in representing all aspects of the empirical data.

19/14 --> Again, contrast is drawn with previous work in a way that only focuses on differences in procedure not on the success of the interpretation. In any case, no theory can be considered of general validity when it applies only to laboratory preparations with a series of suspensions from the same source of INPs and in one specific manner of testing.

20/17 - 20/18 This has been a limitation of this paper from the beginning. The fraction frozen curves are incomplete representation of the information content of the data.

21/2 What is meant by 'freezing behavior'? If it refers to the breadth of the F(T) curves, that represents a narrow view of what the empirical data indicates.

20/2 - 23/10 Most of the four pages of Section 3.5 is an unnecessary repeat of the features of the proposed model.

24/1 -24/3 This is highly arguable. The case cited is just one of many other studies of time-dependence.

Sections 4 and 5 remain of doubtful value until the preceding material is improved.