Maurel at al. present a comparison of two averaging methods used to determine seismic velocities in anisotropic media, outlining the derivation of velocity-averaging and elasticity-tensor averaging for single grain and polycrystal ice exhibiting typical anisotropic fabric. The authors demonstrate the shortcomings of the velocity-averaging method which result in erroneous P-wave velocity estimates and unphysical S-wave velocities with subsequent erroneous average S-wave velocities. By contrast, the elasticity-tensor averaging method is shown to be robust in the cases presented. The authors go on to outline the shortcomings of the Bennett (1968) method. This paper is an extension of Maurel et al. (2015), with reproduction here of a number of derivations and equations reported previously. Maurel et al. (2015) includes a section 5(b) Comparison with previous work which is essentially a digest of, or prelude to, this paper, commenting that the differences will be discussed in detail in future work. At the end of this section they essentially report the findings presented here, but arguably in a more succinct manner: The agreement is excellent [referring to Bennett], although less good for the S-wave than for the P-wave, and this will be analysed in more detail in forthcoming work. The resulting agreement on the velocities is 0.07% for the P-wave and 0.7% for the S- wave, without any adjustment (figure 10) (red and black curves). For completeness, we also report in figure 10 the results obtained from slowness averaging as used in [24] [referring to Gusmeroli or slowness averaging as a method], omitting the T- average (equations (5.5)) (green curves). This latter case leads to two different S- wave velocities, which is unphysical as the wave propagates along the symmetry axis. Incidentally, there is a slightly more notable disagreement with Bennett [21], with 0.9% for both the P- and the S-wave velocities (when compared with the highest S-velocity).

General comments:

comments from Referees, *Although the result presented here are important, in that they highlight potential flaws in methodology, the significance in terms of the scale of errors introduced is not obvious: 1. The errors introduced using the velocity-averaging method for P-waves are very different to those of S-waves. As such, the significance to different experiments will vary.*

author's response Yes, you are right, for ice, the significance to different experiments will vary (if the experiments involve P- or S- wave propagation).

comments from Referees, *2. In general, seismic anisotropy is presented as a percentage of velocity which is more tangible and indicative of its significance when applied to data. This is not the case here. Presentation of results as percentages would allow readers to determine their significance much more readily and put the errors in context. At present, it is not possible to ascertain the scale of these errors when compared to observational errors, which are currently significant in seismic studies of in-situ ice. Percentages were used in Maurel et al (2015), as highlighted above. 3. It would appear from the plots presented that the*

errors introduced by the velocity- averaging method would be of the order of 1% for P-wave velocities and 1% for S-waves (again, as highlighted above). Errors at this level may be acceptable when comparing to field observations, although future studies will of course need to include this uncertainty.

author's response Indeed, in Maurel et al (2015), percentage were used. Nevertheless, the objective in this paper was to compare several models, and not to inspect their validity. This is the objective of the present paper, and as such, the focus is different. Entering in a discussion on the consequences (in percentage) of using a model which leads to wrong results would be confusing. What would be the message then ? Inspecting the validity of the velocity average method reveals that it conduces to non acceptable results (two different S- velocities). This would be different if two models, with different hypothesis, led to acceptable results; in this case, it would be necessary indeed to inspect whether or not the differences between the two results overcome the uncertainties in the measurements.

comments from Referees, *4. The inclusion of the example of zinc is particularly confusing and of no interest to the general glaciological community. The authors include this example to emphasise the potential errors introduced but its inclusion over-complicates what is already a fairly inaccessible piece of work. More useful would be to present the errors introduced for all the likely anisotropic fabrics of ice (of which two are already presented, and all of which are already outlined in Maurel et at (2015)).*

author's response and changes, This point is related to the previous ones. We have removed the example of the zinc. We also added the velocities coming from the Bennett's calculations. We have changed the text in order to stress that the relative agreement between the observed values does not support the idea that erroneous models can be used. It only explains why the error in using such models for simple textures (cluster and girdles) has not been detected. Next, inspecting other fabrics would be useless since one cannot cover all the possible textures, thus we cannot guaranty that the error in using an erroneous model will be always not too important. Thus, we choose a texture which allows to demonstrate that the models are erroneous, with no need for comparison with experiments or with a reference model (a reference model would require to demonstrate that it is the best one, which is not possible in general).

comments from Referees, *5. There is no glaciological context with regards the stress regime responsible for the fabrics presented, or why anisotropic fabric in ice is of interest etc. Again, this reduces the target audience. The manuscript is poorly written, difficult to follow, poorly structured and with an un- scientific style in places. As such, the main findings of the work are not clear and will be overlooked by the vast majority of readers. The structure needs attention and section headings need to be more specific and descriptive to improve the flow of the manuscript. Method and application should be in separate sections. The style, gram- mar and vocabulary also need attention: the paper is currently below*

the standard where a reviewer can be expected to correct all of the grammar and style issues.

author's response , There are two different points in this comment. First, the glaciological context with regard the stress regimes responsible for the fabrics presented. This is clearly outside the scope of the present paper, since our conclusions hold for other polycrystals, as soon as sonic logging measurements are concerned. The second point is : Why the anisotropic fabric of ice is of interest is another question, and clearly pertinent here. In fact, it motivated the present study. Because sonic loggers start to be used in the context of glaciology, and in this context, high accuracy is required because of the weak anisotropy of single ice crystal, it is particularly important to use accurate models. At this stage, we cannot claim that the presented model is sufficiently accurate (this needs comparison with results of well controlled experiments). We stress that the average velocity model and Bennett's results are erroneous and as such, it would be better not to use them. One could make an exception for the case of the clustered textures, because it seems the error to the erroneous model falls within the uncertainties of the measurements, but it is not so helpful. Indeed, in this case, the expressions found by the effective medium theory are also easy to use.

author's changes, We have modified the text and more specifically the introduction and the conclusion to make clearer the goal in the presented study (analysis of several models used to invert the elastic velocities to get the anisotropy of ice polycrystal); also the motivation coming from the use of sonic logging measurements in the context of glaciology. We have tried to improve the english.

comments from Referees, *Recommendation This is a useful and timely piece of work and adds to the growing body of papers investigating anisotropy in ice, highlighting pitfalls of a previously-applied methodology and assumptions made therein. The discrepancies introduced by using the velocity- averaging method are an important finding which must be heeded by future workers in the field.*

As such, the findings are suitable for publication in The Cryosphere. However, I have two main concerns and suggested improvements: 1. The significance of the findings to the glaciological community is not well presented, and possibly of only minor significance when compared to observational uncertainties, and as a result they may be overlooked. The glaciological context and significance needs more discussion.

author's response and changes, As previously said, we have tried to make clearer the goal of the present study and the significance of the findings to the glaciology community. With regard to the sonic logging measurements used recently in boreholes, accurate models have to be developed, able to describe the simple or more complex textures of ice polycrystals. A first step is to avoid erroneous models.

comments from Referees, *2. One can regard this paper as an application of Maurel et al (2015). I would therefore recommend two strategies to improve the manuscript. If this paper is to be published in The Cryosphere this is critical to ensure accessibility to the likely readership: Option 1: Significantly more use could be made of referring to Maurel et al. (2015) with the removal of some of the repeated text and functions to improve the readability of this paper, e.g., parts of Section 2.1 and the Introduction. Alternatively, some of the content could be moved to an appendix or supplementary material and reported in more detail, as per Maurel et al (2015). At present, this manuscript is a poorly structured version of Maurel et al. (2015), which uses a much more coherent and well-structured presentation style. Option 2: This paper could be submitted as a Brief Communication, highlighting the results and discrepancies of the different methods and removing a significant part of the introduction and methodology by referring to Maurel et al (2015).*

author's response and changes, We judge that the option 1 is better. A brief communication removing the specific detailed calculations will not help the accessibility to the likely readership. We are not presenting the discrepancies between different methods (in which case we agree that removing the calculations would be incidental). We are explaining why two of them are erroneous and this requires to be specific in the description of these methods. With regard to the Option 1, we have significantly re-written the paper (notably by removing repeated text).

comments from Referees,

Specific Comments The manuscript is full of grammatical and linguistic errors, beyond what I regard as reasonable for a reviewer to highlight, and is in need of significant proof reading and editing prior to re-submission. Title: The title as it stands does not describe the manuscript. The paper is a comparison and evaluation of averaging methods. Abstract: The abstract does not fully describe the manuscript, only outlining the aims ??of the work and not the method, results or conclusions. Sections headings should be more concise and descriptive Section 2 As per RC1, this section could be re-structured to improve readability. Consider moving the first part of Section 3 (the two boxed sections, or perhaps all of 3.1) here to create a self-contained section of methodology followed by application only in Section 3. Section 3.2 This should become section 4 to improve readability. Section 4 The section discussing Bennett (1968) again builds on Maurel et al (2015). However, in the previous paper the errors introduced were described in terms of percentages. However, in this manuscript this section is poorly structured and difficult to follow. Figures 5 and 6 can be merged if zinc is dropped, and similar plots with percentage anisotropy included. The sections discussing previous work (Diez, Bennett, Gusmeroli) are poorly structured and lack focus or specifics (such as section or equation numbers in the previous work). The reader will therefore struggle to fully understand the issues with the previous work. Conclusions: As with the abstract, the conclusions do

not encapsulate the full body of work. Similarly, a separate discussion section is required, most likely a re-working of existing text will suffice.

author's response and changes, We have rewritten the abstract, the introduction and the conclusion.

Notably, the introduction has been changed in order to stress the difference in the focus of the present paper with respect to Maurel (2015). The section discussing Bennett has been shortened and we think that it better stresses the method used in Bennett 1968. The new Figs. (removing the example of Zinc) includes Bennett's predictions, rather than the averages which were not discussed .