

Review #2:

Comment: Summary:

This work presents a thorough, although perhaps at times exhaustive, discussion on the definitions and distinctions of the terms ‘error’ and ‘uncertainty’ and how their viewpoint, namely that the two are the same, differs from that put forth in “Evaluation of measurement data – Guide to the expression of uncertainty in measurement”, issued by the Joint Committee for Guides in Metrology, colloquially referred to as GUM. On the whole, the topic of this paper is somewhat subjective and more a matter of the definition of terminology. While I would, at first, question its appropriateness for this journal, I think the topic is useful if not necessary and find that this journal is likely the most appropriate option for its target audience. I mostly agree with the authors’ viewpoint and rationale throughout but would argue that it likely should be reorganized and made more concise.

Reply: The authors thank the reviewer for their open- mindedness with respect to a topic that is somewhat off the mainstream of the usual AMT articles.

Comment: General Comments:

I find this paper unique amongst all those I have reviewed in the past in the sense that its premise borders almost on the philosophical rather than the purely scientific even though most of its arguments are rooted soundly in established mathematics and statistics. Much of the paper hinges on the authors’ interpretation of GUM and addressing what they believe to be logical fallacies and inconsistencies. However, it is clear from reading both this work and the review of Referee #1 and his colleagues, which demonstrates an extensive familiarity with GUM, that there are differing interpretations of GUM. I must admit that I am only somewhat aware of the content of GUM and have not read it in its entirety, which unfortunately impedes my ability to thoroughly evaluate this paper. For this reason, I must defer much of the debate on the meaning / intention of GUM to what will be subsequent interactions between the authors, Referee #1, and the editor, but overall this supports an underlying premise of the paper that further clarification of the nomenclature is necessary.

Reply: We do agree that some of our interpretations are subjective, but since GUM is in some places somewhat ambiguous and leaves room for interpretation, this subjectivity seems unavoidable to us.

Comment: The authors go into detail in Sections 3 and 4 of what they interpret GUM is saying on the meanings of error and uncertainty. However, at the end of Section 4 (paragraph starting at Line 338) the authors more clearly state what error and uncertainty are,

though it is not clear if this is the authors' interpretation of what they believe GUM is actually saying (as opposed to what the authors interpret GUM intends to say) or if it is what the authors believe. If the former, it should be better clarified. If the latter, this seems rather subjective (as opposed to definitive) because to attribute any meaning to either of these words will always be a matter of opinion, even if a strict definition is likely necessary. If the authors wish to firmly establish their own definitions of error and uncertainty, then they must make this a clearer priority of this paper. Another alternative is that might be a good topic for a companion paper.

Reply: The paragraph starting with line 338 (in the original manuscript) describes how the terms 'error' and 'uncertainty' were understood in the pre-GUM language. Some confusion is generated by bringing this point not until here. In the restructured manuscript, we redistribute our arguments into a terminological section and a conceptual section. We start the subsection on terminology with the connotation of these terms in common language. Restructuring and rewriting hopefully has made this issue clearer in the revised version.

Action: The manuscript has been restructured and partly rewritten as described above.

Comment: While Sections 5.2 and 5.3 are interesting, they do not appear to be necessary to make the authors' main arguments pointing out the inconsistencies in GUM. However, as mentioned previously, if the authors wish to firmly establish their own definitions of error and uncertainty, particularly as they apply to remote sensing and inversion schemes, then these sections become more useful.

Reply: Section 5.2: In the restructured version of the manuscript, we make more clearly that we conceive measurement (along with its evaluation) as an inverse process, and we put all the relevant arguments, which were distributed over the paper in the original manuscript, in one subsection. We trust that this action has made our point clearer, and that it is now more obvious, why the base-rate-fallacy issue is relevant.

Section 5.3: If the relation between the value of the measurand and the measured signal is nonlinear, the error in the estimate of the true value depends on the true value. We still think that this is a serious challenge for the claim that uncertainties can be estimated without referring to the true value. Our rewriting of Section 5.2 should have paved the way towards a better understanding of the relevance of (old) Section 5.3.

Action: The manuscript was restructured and partly rewritten.

Comment: Normally I would avoid addressing the content of another referee's review, but given the extensive nature of the review

it seems appropriate to avoid being repetitive and ease the subsequent burden on the authors to respond. Additionally, while I have some opinions on the philosophical debates on frequentists versus Bayesians and other topics within this work, it does not appear that my opinions on the matter are as strongly held as written in either the paper or the review so I will abstain so as not to add complication to the review process. However, I would like to chime in on the issue of “systematic” versus “random” uncertainties. I understand the Referee’s point-of-view regarding the nomenclature (i.e., preferring “persistent” and “volatile”), and while technically more accurate it might be simpler and more relatable to the general reader to stick with the more familiar terms. Additionally, I would like to second the authors’ desire to try to classify uncertainties in this fashion (i.e., the “classical” approach according to Colle as discussed by Referee #1) as it is important for the user to know the nature of the associated uncertainty.

Reply: We agree that the terms “systematic error” and “random error” should not be superseded by a new terminology.

Action: We stick to the established terminology but we add a descriptive statement that systematic errors are persistent and that random errors are volatile.

Comment: I suppose this is somewhat expected given the nature of the topic of this work, but in general I find the content to be filled with jargon and philosophical tangents. I would imagine that the average reader that may not have an extensive knowledge of statistics might find this paper to be unapproachable. It would be in the authors’ best interests to cut unnecessary content (e.g., many of the tangents regarding the comparisons between frequentists and Bayesians outside of Section 6.4) to the extent possible to both shorten the paper and appeal to as wide an audience as possible, particularly if they choose to further discuss their own preferred definitions of error and uncertainty.

Reply: We agree that the original version of the paper contained unnecessary material that made it hard to read.

Action: We have restructured and shortened the paper and have removed unnecessary formalism.