Interactive comment on "The Stochastic Calculus Reformulation of Data Assimilation: on Scale" by Feng Liu and Xin Li

Anonymous Referee #1

Received and published: 12 July 2016

We all carry with us the notion of "scale," and we have been performing scale transformations on mathematical formulations of problems since we were beginning students. Upon reflection, few would dispute the proposition that quantitative evaluation of uncertainty is fundamental to data assimilation, and that our evaluation of uncertainty is highly scale dependent. I don't recall seeing a rigorous treatment of the role of scale in uncertainty quantification as applied to data assimilation, so this article has the potential of introducing a useful new idea to the community. Setting down a rigorous definition of scale and scale transformation, and showing how the definition of scale can be applied to quantification of uncertainty would be a worthwhile contribution to the field of data assimilation. In addition, I welcome the explicit introduction of measure theory and stochastic calculus to the data assimilation community, which sees very little of either, especially the former.

C1

Unfortunately, the exposition is extremely hard to follow, and, after reading the article through carefully, I still don't understand what the authors are doing. Part of the problem is probably the language. The manuscript would profit considerably if the authors could find a sympathetic native English speaker to read it over. More importantly, I find much of the exposition puzzling. I don't know the book by Billingsley. In my day students in the USA learned measure theory from the texts by Bartle and Royden, and, relative to my background, much of the material is written very unconventionally. I found this manuscript very hard to read, as much of it seemed to conflict with my most basic intuitions.

The usual intuition for the concept of measure is that measure is a generalization of the concepts of length, area and volume, and is thus a scalar valued set function. The authors' response at the end of section 2.1 to Prof. Talagrand's comment is inadequate. There is nothing intrinsically wrong with choosing a vector valued measure, but that choice requires more explanation than simply "the measure correspondingy turns to . . ." The authors should explain why they want to define measure as a vector valued set function, rather than simply defining the measure of a rectangle in Euclidean space as its area. Again, maybe Billingsley defines it differently, but the Lebesgue measure of a rectangle is its area, not a vector whose components are the lengths of its sides as the authors assert on line 16 of page 6.

I do not understand the definition of scale. First, measure is a function whose domain is the sigma field F, as noted at the very beginning of section 2.1. The integrals in the definition of scale, line 16, page 7, don't make sense to me. A_0 is a specific set. (Pardon my TeX, I don't know how to make superscripts, subscripts or special characters). In the analysis texts I learned from, $\mu(A_0)$ is the area of the set A_0. I don't understand the expression $\mu(A_0)dA_0$. I cannot make sense of the second integral. The domain of the measure μ is the σ-field F, as noted in the beginning of section 2.1. The Lebesgue measure is not a point function.

It would help a great deal if there were more explanation of figure 1. In particular, after

C2

reading and rereading the last paragraph on page 8, I can't understand how C_2 can have the same measure as C_1 and C_3, and D_1 has the same measure as D_2, though they have the same "scale." The problem may be the terminology: As I recall my long-ago analysis classes, the common intuition for measure was that measure corresponds to area, and, in particular, the Lebesgue measure of a geometrical figure in the plane is its area.

Finally, the manuscript seems inconsistent with itself. As examples, consider the abstract. "...measure theory was used to propose [a definition of] spatial scale ... [and the] Jacobian matrix [was used] to describe the change of scale. The Jacobian matrix is introduced on page 7 in the well known change of variables formula, change of scale by the Jacobian matrix is defined on page 8, and the Jacobian is not mentioned again until the summary. No further discussion of the effects of change of scale appears. Again, in the abstract, "...the variation range of this type of error is proportional to the scale gap, ..." I'm sure I'm not the only reader for whom the phrase "scale gap" conjures up ideas of inertial range from turbulence theory and similar notions. The term "scale gap" is never mentioned in the body of the article.

The Bayesian expression of DA in terms of the stochastic calculus appears in many places. The authors should consult the volume by Jazwinski and the recent work of P. J. van Leeuwen and M. Bocquet.