

## Response to Reviewer #1

We thank the anonymous Referee #1 for taking his/her valuable time to review our manuscript and provide us some very thoughtful and constructive comments. Here are the point-by-point  
5 responses to the comments.

### 1. The language problem.

“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  
10 language. The manuscript would profit considerably if the authors could find a sympathetic native English speaker to read it over.”

**Response:** Considering that we are non-native English speakers, our manuscript was thoroughly edited by a manuscript service company named American Journal Experts before submission (Certificate Verification Key: 53B9-681E-0C9B-BC77-FF4C). However, the exposition can be  
15 improved further. Overall, we tried to introduce our ideas in the scale problem with some classic theorem, so some expressions may become a little complicate to understand. We ensure that we will provide a modified version of this manuscript with higher quality English expression in the future.

### 20 2. The references problems.

“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.”

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

**Response:** The literatures written by Billingsley, Bartle and Royden are all classic works of measure theory (Maybe the texts you mentioned are *The Elements of Integration and Lebesgue Measure* by Bartle R. G. and *Real analysis* by Royden H. L.), and they were all highly cited according to the search  
30 results of Google Scholar. In addition, I also got many help from the volume titled “Stochastic

Processes and Filtering Theory” by Jazwinski A. H. during my study, which will be listed in the ‘References’ of revised manuscript. And we also thanks for the recommended literature of the Bayesian Data Assimilation (DA) in terms of the stochastic calculus. They introduced the latest advances in this research field, and the related papers will be cited in our present and future works.

5

### 3. Why the measure in manuscript was defined as a vector valued set function.

“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 correspondingly 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.”

10

15

**Response:** It is true that in the classic literatures on measure theory, measure is a scalar valued set function. However, when it comes to spatial scale, more information is necessary instead of one scalar value. For example, suppose that rectangle A is 4 meters long and 1 meter wide, and rectangle B is 2 meters long and 2 meters wide. Then if we define the area of rectangle is measure, their measures are equal but the shape of spatial scale is missed. So in our study both the length and width of rectangle composed the new measure, and they are also scalar values. However, it’s our mistake that we use the notation of vector to present the new measure, and it has made confusion that the new measure is a mismatch with its basic definition. Therefore, in the revised manuscript, we will replace the old expression with the notation  $\{a, b\}$ .

20

25

Some texts will be updated correspondingly:

- 1) In line 3 of page 6, we will explain why we use the new measure. The sentence “In this case, the measure correspondingly turns into  $\mu(A) = (a, b)^T, a, b \in [0, \infty)$ , which should also obey the countable additivity.” will be modified as “In this case, the subset of  $\Omega$  evolves across two directions because  $\Omega$  is two-dimensional. Therefore the measure should be of double scalar values so that the sufficient information of the subset can be presented. Correspondingly, we

30

define the measure as  $\mu(A) = \{a, b\}, a, b \in [0, \infty)$ , which should also obey the countable additivity.”

2) In line 16 of page 6, we also used the new form of measure to define the rectangle measure, not Lebesgue measure. The equation  $\mu_{iii}(A) = b - a = (b_1 - a_1, b_2 - a_2)^T$  goes to

5  $\mu_{iii}(A) = b - a = \{b_1 - a_1, b_2 - a_2\}$ .

3) Other changes of notations:

Position	Original Text	Revised Text
line 20 of page 6	$\frac{\sqrt{\pi}}{2}(b_1 - a_1, b_2 - a_2)^T$	$\left\{ \frac{\sqrt{\pi}}{2}(b_1 - a_1), \frac{\sqrt{\pi}}{2}(b_2 - a_2) \right\}$
line 8 and 9 of page 7	$(1,1)^T$	$\{1,1\}$
line 10 of page 7	$\frac{2}{\sqrt{\pi}}(b_1 - a_1, b_2 - a_2)^T$	$\left\{ \frac{2}{\sqrt{\pi}}(b_1 - a_1), \frac{2}{\sqrt{\pi}}[b_2 - a_2] \right\}$
line 10 of page 7	$(b_1 - a_1, b_2 - a_2)^T$	$\{b_1 - a_1, b_2 - a_2\}$

#### 4. The definition of scale.

“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  $\mu$  is the  $\sigma$ -field F, as noted in the beginning of section 2.1. The Lebesgue measure is not a point function”

**Response:** There are 2 elements should be concerned about in the definition of scale. The first one is the rectangular referential element  $A_0$ , which represents the unit of the subset and  $\Omega$ . For example, we can define the unit length and unit area as  $A_0$  for one- and two-dimensional space, respectively. And we also introduced the other elementary concept of representative region  $A$ , which is the cumulative amount of  $A_0$ . And of course,  $A \in \mathcal{F}$ . Then the measure function  $\mu(\cdot)$  means to calculate a specific feature of its domain, such as the area, the perimeter or others. So the scale  $\mu(A)$  is a description of representative region  $A$ , and the unit of scale depends on the

referential element  $A_0$ .

In line 16, page 7, we used the integral expression of scale. The reasons were that, as stated above,  $A$  is the cumulative amount of  $A_0$ , and the measure function is with the countable additivity (the second condition of measure, line 21, page 5). So the measure with a domain  $A$  can be calculated  
5 by the cumulative measures with  $A_0$  in  $A$ , which confirms the first integral (it could be more clear if  $A \gg A_0$ ). And if we want to get the reduction formula, then it's natural to replace the surface integral with a double integral, like  $\iint_A \mu(A_0)dA_0 = \iint_A f(x,y)dxdy$ . If  $\mu(\cdot)$  is the Lebesgue measure, then  $f(x,y) = 1$ . Because here the measure corresponds to the area, its output is one-dimensional scalar. However, we let  $f(x,y) = \mu(\cdot)$ , which make the second integral cannot be  
10 deduced – partly because that generally  $\mu(\cdot)$  is not equal to  $f(x,y)$ , and partly because the measure function, as you mentioned, is definitely not a point function, so  $\mu(\cdot)$  is invalid in the second integral.

To correct this mistake, there should be some following changes:

1) From line 16 to line 18, page 7, the new content is “the scale is  $s = \mu(A) = \iint_A \mu(A_0)dA_0$ .

15 From a geometric perspective, the measure refers to the shape of the observation region, and the scale further indicates the size of the region; therefore, the scale increases with increases in the value of the measure. Specifically we further define that the measure is the area of its domain of integration, then  $s = \mu(A) = \iint_A \mu(A_0)dA_0 = \iint_A dxdy$ . This equation simplifies the measure by replacing the surface integral with a double integral, and will be applied to  
20 the following studies.”

2) Related equations should be also changed. The equations in line 20, page 7 will be  $s_1 = \mu_1(A_1) = \iint_{A_1} \mu_1(A_0)dA_0 = \iint_{A_1} dxdy$  and  $s_2 = \mu_2(A_2) = \iint_{A_2} \mu_2(A_0)dA_0 = \iint_{A_2} dudv$ ,  
respectively. The equation in line 1, page 8 will be  $s_1 = \iint_{A_1} dxdy = \iint_{A_2} |J(u,v)|dudv$ . And the Eq. (1) in line 4, page 8 will be  $s_1 = \xi^2 \iint_{A_2} dudv = \xi^2 \mu_2(A_2) = \xi^2 s_2$ . The equation in

25 line 2, page 9 will be  $s_0 = \mu_0(A_0) = \iint_{A_0} dxdy = 1$ .

## 5. About figure 1.

“It would help a great deal if there were more explanation of figure 1. In particular, after 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  
5 for measure was that measure corresponds to area, and, in particular, the Lebesgue measure of a geometrical figure in the plane is its area.”

**Response:** I feel very sorry because the instruction of this figure is not enough to make you understand, and also I think there may be some inconsistency problems of terminology between us. In my manuscript, measure refers to the function  $\mu(\cdot)$ , and its output  $\mu(A)$  is scale. So  
10 measure is abstract, it becomes a real value when its argument (or domain) is confirmed. As noted in line 13 and 14, page 8,  $C_1, C_2$  and  $C_3$  have the same measures because they all calculate the area of the inscribed circle in a square region. And the outputs of measures cannot be obtained until the square regions are confirmed. Therefore, the output, which was defined as scale in our manuscript, is related to both the measure function and the function argument. The scale of  $C_2$   
15 is larger than the scales of  $C_1$  and  $C_3$  because the square region of  $C_2$  is bigger.

However, as you stated, more explanation of figure 1 can help to make our manuscript more clear, so we decide to add more information in here. In addition, in order not to cause confusion, we replace the “measure” by “measure function” in the update version of explanation.

Based on the text from line 10 to line 16, page 8, the new explanation is as follows:

20 “The measure space  $\Omega = [\alpha, \beta] = [x, y: 0 \leq x \leq 4, 0 \leq y \leq 4]$  is regularly divided by a referential element defined with unit area. Let  $\mu_{C_1}, \mu_{C_2}$  and  $\mu_{C_3}$  be the measure functions of the disc measurements  $C_1, C_2$  and  $C_3$ , which present the calculation function of the area of the inscribed circle in a square region; and  $\mu_{D_1}, \mu_{D_2}$  be the measure functions of the diamond measurements  $D_1$  and  $D_2$ , which are also to calculate the area of the inscribed diamond in a  
25 square region, as shown in Figure 1. Then  $\mu_{C_1} = \mu_{C_2} = \mu_{C_3}$  because they are the same functions. And based on the definition, scale is related to both the measure function and the size of representative region. Therefore, the scale of  $C_2$  is not equal to the two other scales because of their representative regions are different. However, their scales are in a one-dimensional law because their measure functions are identical and the Jacobian matrices are diagonal. Similarly, we  
30 have  $\mu_{D_1} = \mu_{D_2}$ ; their scales are also different but are in a one-dimensional law. In addition, the

value of scale is also based on the referential element, which is defined by the unit area. So if the referential element is changed, such as to increase to twice of unit area, the scale is also increased proportionally. However, scales which are in a one-dimensional law will still keep their relationship intact, regardless of whether the referential element changed or not.”

5

## 6. Inconsistency problems.

“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.”

15

**Response:** There are mainly two inconsistency problems you mentioned. The first one is why the Jacobian matrix and the change of scale disappeared after the Sect. 2. Actually, they were not omitted, they were simplified by the one-dimensional law (defined in line 5, page 8) to suit stochastic calculus. And then we used this simplified version of scale change to investigate the uncertainties in DA (we also mentioned this problem in the second paragraph of Sect. 5). Although it maybe the simplest case that how the change of scale can influence the evolution of uncertainties in DA, the results were still complicated and some new components of uncertainty were discovered (Eq. 21~27). However, it is comprehensive and universal to study the change of scale by Jacobian matrix, and it will be launched in our following work, but not in this study. The second one is the phrase “scale gap” in the abstract. I’m sorry for the trouble that this word did. Here the term “scale gap” stands for the quadratic variation between  $s_X$  and  $s_Y$  (see Eq. 25~27). However, it is hard to fully explicate them in the abstract, so we had to use the term “scale gap”. Maybe it’s better to replace it with “the difference between scales” in next time.

20

25