Solid Earth

Discussions

Open Access

EGU

Interactive comment on "Uncertainty in regional estimates of capacity for carbon capture and storage" *by* Mark Wilkinson and Debbie Polson

Anonymous Referee #2

Received and published: 4 June 2019

This paper deals with the uncertainty in estimation of storage capacity for CO_2 in deep saline aquifers. As a reviewer, I have a few concerns about this paper. The Editor may wish to consider these comments in deciding whether to accept the paper for publication.

Major comments:

(1) When I review papers, by far my most frequent criticism is "the novel contribution of this manuscript is not clear." That criticism pertains here. It is clear that the authors are concerned with uncertainty in estimating CO_2 storage capacity, and in coming up with some assessment over whether the uncertainty is large or small. Nevertheless, I am not able to determine the specific knowledge gap or research question being addressed here. The Introduction to the paper does not contain a hypothesis to be tested,

nor a clear statement of a central objective, nor a clear statement of a knowledge gap to be addressed. I do note that the authors state that "an estimate of the accuracy of single-value storage capacities is of practical use," and also that "For this study, an assessment of the accuracy of storage capacity estimates was conducted as part of a study of an area of the UK territorial waters." However, I dispute that the authors have assessed the "accuracy" of any estimates – that would require comparing an estimate to a known or trusted value, which is not done in this paper. Hence, in the end, I find it unclear what important contribution is represented by this manuscript that would warrant its publication.

(2) Closely related to the comment immediately above, it is not clear what we can conclude or take away from the exercise performed here by the authors. Towards the end of the discussion, the authors state that their analysis "is probably realistic for a regional study, where a potentially large number of candidate aquifers are assessed for first-order suitability for storage," but that "it is probably not applicable to a detailed study of a single aquifer, where every effort is made to reduce key uncertainties and where confidential data may be available." I think the implication here is that the estimate of uncertainty made herein would apply to "rough" estimates of storage capacity. Thus the main conclusion seems to be that initial or "rough" estimates of storage capacity carry a high degree of uncertainty. This is not surprising; it is to be expected. Hence, once again, I find it unclear what important new insight is represented by this paper. It is possible that there is an important insight or contribution here, and that I am missing it – that is possible – but if that is the case, the authors must do a better job of clarifying the importance of their work and what it offers to the community.

(3) I question part of the discussion given by the authors. In this analysis, the estimate of storage capacity is made by multiplying together six factors, four of which (A, h, NG, phi) must be estimated independently. The authors correctly note that the maximum possible estimate of storage capacity would be made by multiplying together the highest estimated values of A, h, NG, and phi. The authors also note that none of their

team of experts ever made such an estimate, i.e., in no case did one expert ever make the highest estimates of all four parameters simultaneously. The authors claim that "real" uncertainty may therefore be even greater than the range spanned by their team of experts, because "all possible combinations must be assumed to have the same probability," and "hence the storage capacity estimated using all minimum or maximum values for all variables are equally likely as any other individual combination." I do not think this is correct. First, I think it is unclear that the probability distribution of each parameter individually should be considered to be a uniform pdf. If twelve experts each make a prediction of a variable, and that variable actually has a "true" or "correct" value, is it the case that each of the 12 estimates is equally "good" as the others? Or might we expect that the values closer to the middle of the range are more likely to be "good" estimates than the highest and lowest estimates? If the 12 experts are, in fact, experts, then I think the mean/median values should be "better" estimates (i.e., lie closer to the true or correct value) than the highest and lowest estimates. This would mean that the product of the four highest parameter estimates is *not* just as likely as other combinations. And furthermore, suppose I am wrong on this point – suppose that the probability distributions of the individual parameters can be considered to be uniform pdfs, such that any value between min_value and max_value is equally likely. What does this mean about the probability distribution of the *product* of the values? Is the product of uniform-distributed variables also uniform-distributed? I would guess not! If I roll a standard die, the chances of getting 1, 2, 3, 4, 5, or 6 pips is equally likely. But suppose I roll two dice, and I multiply the values together. The chances of getting a product of 1 or 36 is *not* the same as getting a product of 12. And if I roll four dice, the probability of the product being a 1 is definitely lower than the probability of getting, say, 144 for a product. So for two reasons, I think the authors' contention is mistaken. They may want to consult with an expert in probability or statistics. I would bet that the product of uniform-distributed variables approaches a log-normal distribution or maybe a gamma distribution as the number of variables gets large. (This is just a hunch, I do not know if it is correct.)

C3

(4) It is an interesting question whether 12 experts is enough to represent the range of uncertainty of the individual parameters. Suppose we want to estimate the parameter h and we want to have some quantification of the uncertainty of the estimate. How many expert estimates of h must we obtain before we can conclude that the standard deviation of the estimates is a meaningful quantification of uncertainty? I think 12 sounds like it might be barely enough, but I do not know enough statistics to know the answer to this question – I am just going on intuition. Again the authors may want to consult with an expert in probability or statistics.

Minor comments:

(5) The text on lines 88-98 of the Introduction probably does not belong in the Introduction. Most of this text is either "site description" or "methods". I do not think any of it establishes the main idea of the paper, and therefore I would not put it in the Introduction.

(6) As far as I can tell, the paper does not indicate how many experts were used in this analysis. In my comments above, I assumed 12 experts, but I do not know if this is correct. I am assuming 12 based on the data that I see in Figures 2 and 3. It looks like maybe there were 12 experts involved. But I do not know if this is correct. The actual value should be indicated clearly in the text and possibly in the figure captions too.

Technical corrections:

(7) The grammar needs a little clean-up. The first sentence of the abstract is a run-on sentence because of the peculiarities of the conjunction "however". Later in the abstract the authors state "due a combination of using different published values", i.e., missing the word "to" after "due". Later in the abstract there is another run-on sentence where a comma should be a semi-colon. That is just the abstract. Other grammar issues are found throughout the paper. Maybe hire somebody to read the paper thoroughly and clean up any such errors. They do not impede comprehension, but they distract.

C4

(8) A couple cited papers are not in the reference list. I put this under "minor comments", but it MUST be corrected. Examples are the citations of Calvo et al. (2019) and Medina et al. (2011).

(9) Equation 1 uses h, but the following text uses H.

(10) The word "data" is plural but is incorrectly used as singular throughout the paper.