Authors reply to Referee 1

We are grateful to Anonymous Referee 1 for a careful and in depth review. The comprehensive feedback on the methodology for statistical evaluation of upscaling uncertainties is especially useful and appreciated. We feel confident that we will be able to submit an improved revised manuscript with the support of these comments.

Major/method comments

Below we provide our view on some of the major comments of Anonymous Referee 1 which largely relate to the statistical approach we have taken when estimating upscaling uncertainty.

Inconsistency in use of confidence intervals (CI): We agree that the use of several different metrics of uncertainty makes the paper cumbersome to read and interpret. We will revise the manuscript to only report 95% CI in the main text and figures/tables and only report the 99% CI intervals in supplementary materials for those data-users that may wish to look at these wider uncertainty ranges.

Incorrect use of the formulas for error propagation: We find the ideas and concepts presented by the referee very thoughtful and indeed useful. In general the idea that SOC stocks are correlated within soil orders but uncorrelated across soil orders is sound and we will apply this reasoning in the revised manuscript.

However, it is possible that the suggestions for changes in the calculations are based on a misunderstanding of how the uncertainties were calculated? If there was a misunderstanding, that is a clear sign that our method section needs to better explainined. We constructed a figure to clarify how uncertainties have been calculated and added together (fig. 1). This figure illustrates how uncertainties were calculated for the initially submitted manuscript. A figure showing how we propose to calculate uncertainties in the revised manuscript is presented below.

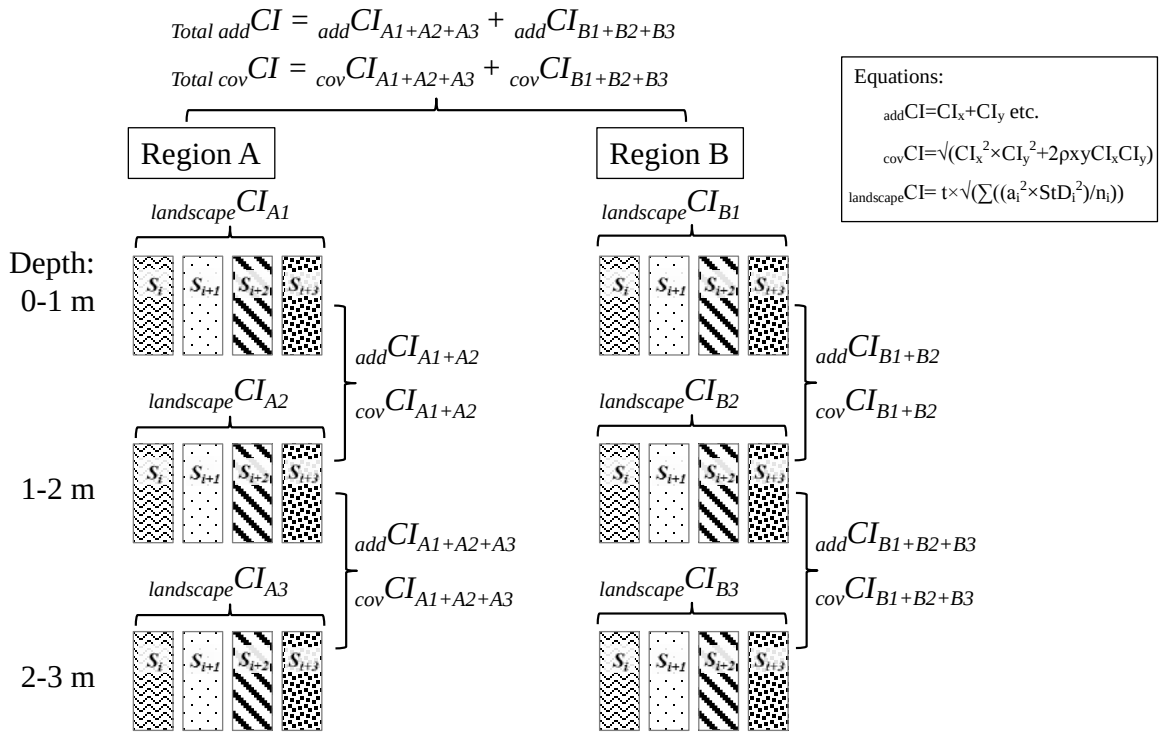

Figure 1. Schematic diagram illustrating how the different CI components were calculated and summarised. The diagram illustrates two separate regions (A and B) with the occurrence of four soil classes (S_{i+x}) over three depth ranges (0-1 m, 1-2 m and 2-3 m). Italic text represents the different applications of formulas. The complete formulas are provided in the box to the upper right.

We essentially calculated separate uncertainties for each depth interval of across one upscaling region. This was done with the formula for calculating spatial $_{landscape}$CI ranges presented by Thompson (1992). This formula combines uncertainties across soil orders assuming no correlation between SOC stocks of the different soil orders (which is also following the recommendations made by the referee). After this, the different depths were summed within regions using either $_{add}$CI or $_{cov}$CI formulas. The CI ranges for different regions were then added up to provide the total CI ranges for the permafrost zone.

The referee argues that the application of $_{cov}$CI formulas when summing up the different depths is erroneous because they should be completely correlated. In the application of $_{cov}$CI formulas we reasoned that 0-1 m SOC stocks are only partly auto-correlated to 1-2 m and 2-3 m SOC stocks. SOC at depths below 1 m is often several thousands of years older than surface SOC and has been physically disconnected from the surface SOC-turnover by either anoxia or freezing. This condition is quite unique to permafrost environments where the vast bulk of SOC is disconnected from the actively cycling surface soils. Correlation analyses showed that the correlation coefficients between surface and deep SOC stocks are ca 0.4-0.5 (supplementary material of the initial manuscript).

However, we acknowledge that this process-oriented reasoning falls outside the realm of mathematical applications of Gaussian error propagation. As Referee 1 suggested we will present only the $_{add}$CI ranges. In the end, this actually corresponds to the recommendation that we had already made in the paper, but with a clearer methodological reasoning behind it. We will apply these corrections in an updated manuscript. For consistency and clarity, we will do all calculations

with the $_{cov}$CI formula but adapting the correlation coefficient (ρxy) to either 1.0 (within soil orders) or 0.0 (across soil orders).

Representation error: This is a very good point raised by the referee. Indeed, we are not confident that our current samples for some soil orders in the thin sediment regions (or the High Arctic) are truly representative. To some extent, the $_{landscape}$CI formula already accounts for small sample sizes. If the sample size deviates far below what is a mathematically robust sample size to approximate a t-distribution with the given mean and StD, the derived CI ranges are relatively wide. The referee states that thick and thin regions have similar CI ranges in table 2. We do not agree with this statement. In fact, the relative range of the 95% CI in thin sediment regions is three to five times wider than for the thick sediment regions (Table 2). This reflects the small sample sizes for that region. However, as the referee points out, our current calculations do not explicitly account for the degree of representativeness of the smaller samples. We would thus propose to calculate the representation error CI ($_{rep_err}$CI) by simulating subsampling of the larger dataset for thick sediment areas as the referee suggest. The $_{rep_err}$CI will then be combined with the $_{landscape}$CI (using the quadrature formula of uncorrelated errors).

We created a figure showing the intended changes we will make to the calculations of CI ranges (fig. 2). We propose to add this figure to the method section of the supplemental material in a revised manuscript. The referee suggests putting together an error covariance matrix for summing up the errors regionally which is also a very constructive suggestion. We will add a covariance matrix to the supplemental materials.

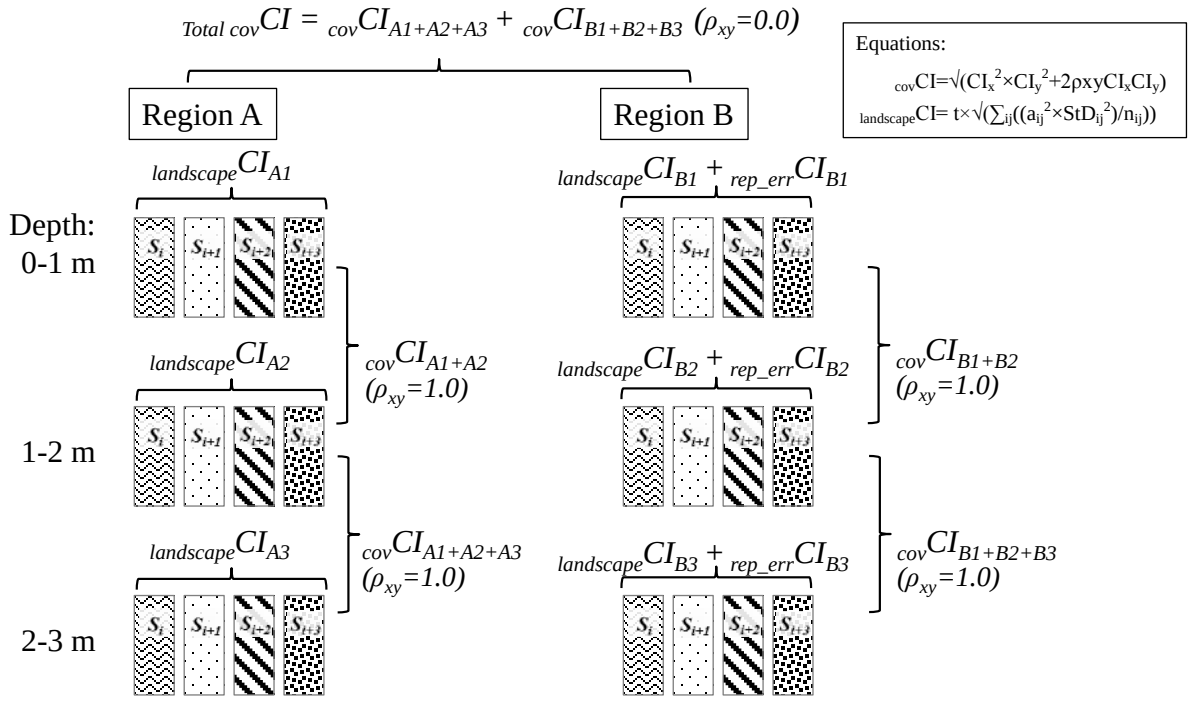

Figure 2. Schematic diagram illustrating how we propose to calculate and summarise the different CI components in the updated manuscript. The diagram illustrates two separate regions (A and B) with the occurrence of four soil classes (S_{i+x}) over three depth ranges (0-1 m, 1-2 m and 2-3 m). Italic text represents the different applications of formulas and applied correlation coefficients are given in brackets. The complete formulas are provided in the box to the upper right. In this example, region B is represented by very small

sample sizes. Therefore, for region B the $_{cov}$CI for each depth range includes both the $_{landscape}$CI and the $_{rep_err}$CI (see text for further explanation).

Minor/editorial comments

Referee 1 provides extensive editorial feedback and many minor comments. We find these comments useful and constructive. We will strive to incorporate them in an updated manuscript. Below we provide explicit response to some of the minor comments by referee 1. Any comments that are not explicitly addressed we will strive to accommodate.

We agree that the terms uncertainty and variance are sometimes used interchanged, we will correct this throughout the manuscript.

The referee asks that we remove algebraic symbols (such as < and >) entirely from the text. We feel that this makes the text unnecessarily long. We will refer this question to the editorial staff of the journal and follow their recommendations.

The referee suggests adapting figures 2 and 3 to contain only the data used to calculate stocks and move the comparisons between the two datasets to supplementary materials. This is a good suggestion that will help us make the manuscript more streamlined and readable.

We will add the SOC storage numbers for 0-30 cm depth in the figures.

The referee suggests that much of the results from the running text be transferred to a table. We will consider this advice. This may make the article even longer, but we may save space with some of the other suggested changes.

The referee suggests that we add maps of spatially distributed uncertainties. This would entail significant additional work beyond what has already been undertaken. Such analyses would also mean significant additions to methods, results and discussion of this manuscript and we feel that this manuscript is already very long and intricate. We are currently pursuing such GIS-analyses but will present this in a separate manuscript.