Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2018-60-RC1, 2018

Interactive comment on "Quantification of the fine-scale distribution of Mn-nodules: insights from AUV multi-beam and optical imagery data fusion" *by* Evangelos Alevizos et al.

Anonymous Referee #1

Received and published: 6 March 2018

Review of BG-2018-60: Quantification of the fine-scale distribution of Mn-nodules: insights from AUV multi-beam and optical imagery data fusion

The manuscript submitted by Alevizos et al. presents a technical study, covering the development of optimal methods to map out the spatial distribution of Mn nodules in the Peru Basin. The authors work at a relatively fine scale (maps with m-scale pixels), finer than most studies have been able to achieve until today, because they have acquired high-resolution multibeam echosounder (MBES) data and photographs of their study area using an Autonomous Underwater Vehicle (AUV). The team present three different methods to map out nodule density from the MBES acoustic data in a (semi-

)automated way: Bayesian classification of MBES backscatter values, ISODATA classification of the backscatter mosaic, and Random Forest predictive modelling based on MBES bathymetry, backscatter, and their derived variables. To ground-truth their unsupervised classification results, and to create their training dataset for the RF modelling, the authors have applied the automated nodule identification and measurement routines developed and described in Schoening et al. (2017).

Overall, the team behind this manuscript hold a lot of technical knowledge in all the fields covered by this paper (backscatter processing, optical image classification, machine learning, the use of AUVs for deep-sea mapping). The technical quality of the work they have carried out is beyond doubt. However, it is my feeling that the manuscript is not presented in the best possible way. In addition I would suggest that the authors reflect more on the aims of the work they have done and the techniques they have chosen to apply, that they express this clearer in the paper, and discuss the outcomes in more depth in the discussion. As a result of these issues, the paper will need a major revision before its publication. I will explain myself in more detail below:

- With regard to the presentation of the work, although the figures are of good quality (bar a few issues) and the writing style is good, there are some issues with the structure of the paper. I have made comments in the attached pdf document, but in summary I had the feeling that in several cases descriptions of the methods that were used were only provided in the 'Results' section, while the 'Discussion' section contained a number of results that should have been presented in the 'Results' section. Furthermore, I was also missing essential information about the datasets used in this study (see comments in the pdf). This should all not be too difficult to revise.

- My second main comment relates to the aim of the work presented here. Unfortunately, I found this not well expressed in the introduction. Section 1.3. provides a short literature review of studies that have looked at the use of acoustics to map out Mn nodules, but the authors fail to draw from this list of previous work the essential state of the art in the field, and more importantly the current knowledge gaps, which would

Printer-friendly version

Discussion paper

then lead them into which of those gaps they are tackling with their paper.

- The comment above has a further implication: the authors state that their aim is the "quantitative mapping of Mn nodule densities in the Peru Basin" (Line 138). Hence, I understand that their aim is to map out the spatial distribution of a continuous variable. The RF modelling they have applied indeed provides very convincing results, including a clear evaluation of the importance of each of the explanatory variables (in addition the RF technique can also map out the prediction uncertainty over the study area, which would be good to add – see e.g. Robert et al. (2015)). It then surprises me that the authors choose two techniques which turn what is in essence a continuous variable (backscatter strength) into a categorical one (classification results), only to then comment on the fact that the categorical results are in fact ordinal (line 375) and to demonstrate in the final figures that there is a near-linear correlation between the backscatter strength and the Mn nodule density. I would advise that the authors reconsider their choice of methods, and balance techniques which create distinct classes against the potential use of e.g. regression techniques to relate the Mn density to the backscatter strength directly. If they choose to still use the classification methods, there should be a very strong justification why this would be a good method to use and what they want to achieve with it. I also note that neither for the Bayesian, nor for the ISODATA classification, the classification results are actually interpreted or groundtruthed (i.e. there is no description of what class 1,2,3,. . . represents). This would then need to be added.

- I do note that the ISODATA algorithm took in more than just the backscatter strength as input variable. However, the four input variables that the authors have chosen (mean BS, mode of the BS, 10% percentile and 90% percentile of the BS) are all very highly correlated (between 50 and 90%, Line 390), which indicates that even with the ISODATA classification method, the algorithm basically works on the BS strength. If the authors decide that unsupervised classification is still a technique that they would like to use, I would suggest that instead they look into the use of (a) the MBES bathymetry

& derived variables as further input into the ISODATA algorithm (also given the importance of variables such as BPI10_100 and bathymetry, as illustrated by Fig. 6c) and (b) the use of other, higher order MBES backscatter variables (e.g. variance, skewness etc. calculated over the 10x10m neighbourhood).

- With regard to the Discussion section, once the further results are moved to the Results section (Table 5 and Fig. 8 and their description), the discussion may feel a little light. The authors compare the three maps they have made with the Mn nodule density results from the photographs (Fig. 7) – this would be the place to come back to the gaps in the current state-of-the-art and discuss what advances have been made. The other topic that would merit some more discussion, to my feeling, is the actual spatial pattern that has been revealed by the authors. The fine-scale spatial distribution of Mn nodules and its relation to subtle variations in the seabed bathymetry is an important result that has not been published very often, and it would be interesting to see more discussion on what the underlying mechanisms would be, even if there are still a lot of questions to be answered.

- Finally, I have one more comment about the methods used, with regard to the way the authors have dealt with potential spatial autocorrelation in their RF training and validation datasets. They correctly indicate that one should be aware of this issue and should try to correct for it, but from the way the manuscript is currently written, it feels as if they have only taken this into account while choosing the validation points. It is important to make sure also the training data does not show spatial autocorrelation. In addition, the authors have estimated that the spatial scale of the autocorrelation is six metres. They have not provided a justification for their estimate. A robust estimate could be obtained by a range of methods (e.g. a semivariogram, or using Moran's I). It is necessary to include such a justification to make sure the applied method and results are robust. I have also made a number of smaller comments in the attached pdf, which should all be addressed during the review, but should be relatively easy to address. My apologies that some of it may look a little scattered, I had problems with the highlight tool in Acrobat, but all the comments should be there.

Given that the work behind this manuscript is of good quality, and that the issues are mainly related to the presentation of it, or to choices of methods that either need reconsidering or better justification, I am confident that the authors will be able to revise the manuscript and create a strong paper. I am looking forward to seeing it published in Biogeosciences.

Best wishes,

Veerle Huvenne

Please also note the supplement to this comment:
https://www.biogeosciences-discuss.net/bg-2018-60/bg-2018-60-RC1-supplement.pdf

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2018-60, 2018.