Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2018-353-AC1, 2018

Biogeosciences Discussions
Open Access

Interactive comment on "Quantitative mapping and predictive modelling of Mn-nodules' distribution from hydroacoustic and optical AUV data linked by Random Forests machine learning" *by* Iason-Zois Gazis et al.

Iason-Zois Gazis et al.

igazis@geomar.de

Received and published: 5 October 2018

Reviewer #1: General comments

This is a well-written paper on the combined use of AUV imagery and acoustic surveys for the assessment of manganese nodules, which shows clear scientific and industrial relevance. However, it shows some similarities to Alevizos et al 2018 (similar approaches, but different locations). Both the size of the area covered and the number of images, highlight the use of AUVs and the importance of automated approaches for

Printer-friendly version

Discussion paper

environmental assessment. The authors have made a point of being very transparent about their approach, and the statistical details provided in the appendices provide extra confidence in the results presented (more of the statistical explanations of the results could be moved to the appendix, e.g. assessment of normality in 4.1).

Authors comment:

We welcome all comments of Reviewer #1 and we appreciate the time and effort put to review this manuscript. Below we present our reply for each of the Reviewer's points:

We believe that the differences in this paper, compared with the paper from Alevizos et al, (2018), are not limited only to different locations. Alevizos et al, (2018) applied three different techniques in order to estimate the distribution of the Mn-nodules inside their study area. Two of them (Bayesian probability on beam backscatter and ISODATA classification) classify the bottom in areas with higher and lower number of Mn-nodules (based on backscatter values), while the third (RandomForests machine learning) predicts the Mn-nodule abundance in each location based on a number of predictor variables (MBES data) and training data (optic data). In our study, we focus only on the Random Forests machine learning prediction performance by applying and tuning the algorithm. Different predictor variables and in different scales (compared with Alevizos et al, 2018) were used; by doing so we supported the investigation of the role of predictor variables in different areas. It is worth to mention that in our study the absence of backscatter information as a predictor variable, showed that topographic factors alone can achieve relatively accurate predictions. Differently to Alevizos et al., an extensive statistical analysis (e.g. assessment of normality, spatial clustering) was performed in order to further investigate the distribution characteristics of the Mn-nodules. This analysis combined with the correlation analysis between the number of Mn-nodules/m2 and the derivatives highlighted the value of the Random Forests algorithm as a tool for complex spatial predictions. Furthermore, this study examined the distribution of the median size and its correlation with the number of Mn-nodules. The idea was to go one step further than Alevizos et al 2018, by introducing and applying a relatively simple operational workflow, highlighting at the same time the advantages and disadvantages of the existing sampling procedures.

The assessment of normality in 4.1 was moved to Appendix B in order to strengthen the important results. Finally, the paper by Alevizos et al. was redrawn and will not be published in the way it has been discussed in Biogeoscience. âĂČ Reviewer #1: Specific comments

(1) There is reference to a particular MBES depth data processing that guarantees removal of artifacts and improvements of georeferencing, but no reference or descriptions are given. Particularly in deep waters, inaccurate AUV positioning will be an issue, especially when trying to related photographs to 3m resolution bathymetric grids.

Authors comment:

We added some more explanation in the text.

(2) As mentioned in line 420-425, choice of scale is important in deriving terrain metrics and a quantitative justification for the choice of chosen scales should be provided.

Authors comment:

Lines 420-423 state that the arbitrary choice of scale limits the value of the terrain metrics as explanatory variables exactly because of the scale dependency in environmental modelling. Thus using derivatives of different scales (e.g.fine or broad scale BPI) contributions critically to environmental modelling results. Due to the lack of relevant literature for AUV scale data sets, the Concavity and Terrain Ruggedness indexes were created with the default scale of SAGA GIS v.6.3.0 (radius of 10 cells as stated in Table 1). The three different values for the Topographic Position Index were selected based on the minimum possible correlation among them (surface correlation tool, in SAGA).

(3) The calibration of the model section is not as clear as it could be. Lines 223-225 need to clearly state that the default values were used for the assessment of

training/testing sample size only. Lines 238 says that after training sample size was determined were mtry and ntree examined, but Line 241 mentions that for each case of different training sample size, ntree and mtry, the model was run ten times. The latter sentence should be split to clarify that each training sample size was not also tested for each different numbers of ntree and mtry. Similarly, Line 243, if I understood correctly, Appendix A only presents the averages for the 10 training size runs, and not the ntree and mtry runs. The wording here also needs to be clarified.

Authors comment:

Lines 223-225, 239-240 (in the submitted manuscript) were changed accordingly to the recommendations, stating clearly that the default RF values (for regression) were used only during the investigation of the optimum training size. Line 241 (in the submitted manuscript) was changed, now stating that ten different mtry and ntree values were applied for ten times each, only in the optimum selected training size. All tables regarding the statistical characteristics of the performance after 10 runs are presented in Appendix B.

(4) I am not convinced that the approach taken can be used to determine the optimal training sample size proportion. More data is likely to yield better models, but by decreasing number of testing data points, one can also expect MSR to keep decreasing (as was shown here). A much more interesting question would be how many samples are needed to obtain accurate predictions.

Authors comment:

Indeed, the less testing data points you have the more likely it is to achieve a lower error only because your model fits relatively well. In lines 356 – 361, we justify our choice to use the model with 80% because of the higher number of validation data compared to the 90% model.

(5) RF models also provide a measure of uncertainty, it would be interesting to provide

uncertainty maps for the associated predictions and discuss potential spatial trends if any.

Authors comment:

Indeed, the use of uncertainty prediction maps is a useful tool in spatial predictive mapping as they can reveal spatial trends (e.g. areas with higher uncertainty in predicted value) which might lead to additional sampling in order to advance the model and support the interpretation. The RF models inside the randomForests R package (Liaw and Wiener, 2002), can give an accurate prediction of the conditional mean of the response variable. The uncertainty (conditional quantiles) around this mean can be estimated by the use of the Quantile Regression Forests (Meinshausen, 2006), as they keep all the values in each node in each tree (not only the mean value), allowing the construction of prediction intervals. Quantile Regression Forests (QRF) models can be developed using the quantregForest R package (Meinshausen, 2012). The used MGET toolbox (Roberts et al, 2010) includes only the randomForests R package (Liaw and Wiener, 2002) and the party R package (Hothorn et al., 2006; Strobl et al., 2007; Strobl et al., 2008). MGET was selected as tool to keep the proposed workflow simple and, in a graphic environment familiar to many geoscientists. Recent comparative studies showed that the accuracy of the quantregForest R package against standard RF does not differ considerably, while it increased the computational time (Tung et al, 2014), without adding any other information regarding the variable importance. The use of other recently proposed methodologies as the Jackknife method (Wager et al, 2014), the Monte Carlo approach (Coulston et al, 2016) and U-statistics approach (Mentch and Hooker, 2016)) are far beyond of the aim and purposes of this paper.

1.Liaw, A. and Wiener, M.: Classification and regression by randomForest. R News, 2/3:18–22, 2002. http://CRAN.R-project.org/doc/Rnews/

2.Breiman, L.: Random forests. Machine Learning, 45, 5–32, 2001. https://doi.org/10.1023/A:101093340

3.Meinshausen, N.: 2 Quantile Regression Forests. Journal of Machine Learning Research 7, 983–999, 2006.

4.Hothorn T, Hornik K, Zeileis A.: "Unbiased Recursive Partitioning: A Conditional Inference Framework." Journal of Computational and Graphical Statistics, 15(3), 651–674, 2006. https://doi.org/10.1198/106186006X133933

5.Strobl, C. Boulesteix, A.L., Zeileis, A. and Hothorn, T.: Bias in random forest variable importance measures: Illustrations, sources, and a solution. BMC Bioinformatics, 8:25, 2007. https://doi.org/10.1186/1471-2105-8-25

6.Strobl, C., Boulesteix, A.L., Kneib, T., Augustin, T. and Zeileis, A.: Conditional variable importance for random forests. BMC Bioinformatics, 9:307, 2008. https://doi.org/10.1186/1471-2105-9-307

7.Tung N.T., Huang J.Z., Khan I., Li M.J., Williams G.: Extensions to Quantile Regression Forests for Very High-Dimensional Data. In: Tseng V.S., Ho T.B., Zhou ZH., Chen A.L.P., Kao HY. (eds) Advances in Knowledge Discovery and Data Mining. PAKDD 2014. Lecture Notes in Computer Science, vol 8444. Springer, Cham https://doi.org/10.1007/978-3-319-06605-9_21

8.Wager, S., Hastie, T., and Efron, B.: Confidence intervals for random forests: the jackknife and the infinitesimal jackknife. Journal of Machine Learning Research, 15(1):1625–1651, 2014.

9.Coulston, J. W., Blinn, C. E., Thomas, V. A., and Wynne, R. H.: Approximating prediction uncertainty for random forest regression models. Photogrammetric Engineering & Remote Sensing, 807 82(3):189 – 197, 2016.

10.Mentch, L. and Hooker, G. (2016). Quantifying uncertainty in random forests via confidence intervals and hypothesis tests. Journal of Machine Learning Research, 17(1):841–881.

(6) Autocorrelation in Mn nodule distribution was discussed, but whether model resid-

uals showed any spatial autocorrelation was nor assessed, nor were the effects of this autocorrelation on model assessment discussed.

Authors comment:

The spatial autocorrelation analysis of the residuals using the Global Moran's Index (same settings as Appendix A), showed low, but significant spatial autocorrelation ($I=0.112112$ $p<0.01$ and Z-score>2.58). The index number of the residuals is relatively low compared with the high initial values of the original data ($I=0.69890$ and $I=0.697747$ for the entire dataset and the 80% training dataset, respectively). The 5% trimmed residuals (see Appendix B-Table B8) showed that their spatial autocorrelation is only 0.093832. According to similar studies (i.e. regression RF), the presence of spatial autocorrelation in the residuals of the model can result in underestimation of the true prediction error (Ruß und Kruse, 2010). The presence of low spatial autocorrelation values in the residuals of regression RF has been reported also by other authors (e.g. Mascaro et al, 2014; Xu et al, 2016); it is a common problem in all the well-established machine learning methods (e.g. RandomForests, Neural Network, Gradient Boosting Machine, and Support Vector Machines) when dealing with regression predictions of spatial variables (Gilardi and Bengio, 2009; Ruß und Kruse, 2010; Santibanez et al, 2015 a,b). The spatial plotting and visual examination of the residuals (Figure 1) showed that this spatial clustering exists mainly in the small sub-area b, and especially in the areas which are associated with an increased slope (>3°), where the AUV is forced to vary its altitude between the ascending and descending phase (Figure 7b) and consequently affects the image quality and the later modelling results.

1.Ruß, G., and Kruse, R.: Regression Models for Spatial Data: An Example from Precision Agriculture. CDM 2010. Lecture Notes in Computer Science, vol 6171. Springer, Berlin, Heidelberg, 2010. https://doi.org/10.1007/978-3-642-14400-4_35

2.Mascaro, J., Asner, GP., Knapp, DE., Kennedy-Bowdoin, T., Martin, RE., Anderson, C., Higgins, M., and Chadwick, D.: A Tale of Two "Forests": Random Forest Ma-

chine Learning Aids Tropical Forest Carbon Mapping. PLoS ONE 9(1): e85993, 2014. https://doi.org/10.1371/journal.pone.0085993

3.Xu, L., Saatchi, SS., Yang, Y., Yu, Y., and White, L.: Performance of nonâĂŚparametric algorithms for spatial mapping of tropical forest structure. Carbon Balance Manage, 11:18, 2016. https://doi.org/10.1186/s13021-016-0062-9

4.Gilardi, N., and Bengio, S.: Comparison of four machine learning algorithms for spatial data analysis. Conf. Signals Syst. Comput., 17, 160–167, 2009.

5.Santibanez, S., Lakes, T., and Kloft, M.,: Performance Analysis of Some Machine Learning Algorithms for Regression Under Varying Spatial Autocorrelation. The 18th AGILE International Conference on Geographic Information Science, Lisboa (Portugal), 9-12 June, 2015a.

6.Santibanez, Sebastian F., Marius Kloft and Tobia Lakes. "Performance Analysis of Machine Learning Algorithms for Regression of Spatial Variables. A Case Study in the Real Estate Industry." the 13th International Conference of GeoComputation, Dallas (USA), May 20 – 23, 2015b.

(7) The discussion is very much focused on the model and although the exploratory nature of machine learning algorithm is mentioned, a little more discussion of the causation mechanisms (or potential hypothesis) would be valuable.

Authors comment:

Classic studies have shown that the bathymetry and the variation of the topographic characteristics of the seafloor affects the sediment deposition environment, bottom currents and thus also geochemical processes in the sediment. All these factors determine Mn-nodule growth and thus affect the distribution of Mn-nodules on regional scales (e.g. Craig, 1979; Sharma and Kodagali, 1993;). It is unknown how these properties influence the Mn-nodule distribution on meter to tens of meters scales as seen in our AUV data. The non-linear relationship between Mn-nodules and bathymetry on such high-resolution scales only started very recently (Peukert et al, 2018 and also th withdrawn submission by Alevizos et al). To elaborate more on the hydrodynamic and geochemical reasons behind the observed distribution pattern, we would need more investigations at and in the sediment on the same scale. Without such data, any elaboration on the reasons for the distribution would be purely speculative, without additional 'ideas' than the known and published influencing parameters. 1.Craig, J. D.: The relationship between bathymetry and ferromanganese deposits in the north equatorial Pacific, Marine Geology, 29, 165–186, 1979. https://doi.org/10.1016/0025-3227(79)90107-5

2.Sharma, R. and Kodagali, V.: Influence of seabed topography on the distribution of manganese nodules and associated features in the Central Indian Basin: A study based on photographic observations, Marine Geology, 110, 153–162, 1993. https://doi.org/10.1016/0025-3227(93)90111-8 3.Peukert, A., Schoening, T., Alevizos, E., Köser, K., Kwasnitschka, T., and Greinert, J.: Understanding Mn-nodule distribution and evaluation of related deep-sea mining impacts using AUV-based hydroacoustic and optical data. Biogeosciences, 15, 2525-2549, 2018. https://doi.org/10.1007/978-3-319-57852-1_24

4.Alevizos et al, Schoening T., Koeser K., Snellen M. and Greinert J.: Quantification of the fine-scale distribution 1 of Mn-nodules: insights from AUV multi-beam and optical imagery data fusion. Biogeosciences Discussions. pp. 1-29, 2018. https://doi.org/10.5194/bg-2018-60

Reviewer #1: Technical corrections - Authors comments:

Line 38 I would suggest changing sea bottom for seafloor - Done

Lines 39-45 I would suggest specifically introducing the term backscatter, as I believe that to be one of the main data product used for to show Mn trends in regional surveys - Done

Line 61 Reference style - Done

Lines 81-83 Awkward, please rephrase - Done

Line114 In the marine environment, - Done

Line 131 remove scale) -Done

Line 133 is deeper and has less relief - Done

Line 189, while - Done

Line 222 a threshold of 0.95 for correlation of variable seems very high – Done (we changed the term highly with perfectly correlated. In similar studies even higher thresholds have been used during the selection of predictor variables (Che Hasan et al, 2014; Li et al, 2016; Li et al, 2017)).

1.Che Hasan, R., Ierodiaconou, D., Laurenson, L., Schimel, A.: Integrating Multibeam Backscatter Angular Response, Mosaic and Bathymetry Data for Benthic Habitat Mapping. PLoS ONE 9 (5), e97339, 2014. https://doi.org/10.1371/journal.pone.0097339

2.Li J, Tran, M, Siwabessy, J: Selecting Optimal Random Forest Predictive Models: A Case Study on Predicting the Spatial Distribution of Seabed Hardness. PLoS ONE 11 (2): e0149089, 2016. https://doi.org/10.1371/journal.pone.0149089

3.Li, J., Alvarez, B., Siwabessy, J., Tran, M., Huang, Z., Przeslawski, L., Radke, L., Howard, F. and Nichol, S.: Application of random forest, generalised linear model and their hybrid methods with geostatistical techniques to count data: Predicting sponge species richness. Environmental Modelling & Software, 97, 112-129, 2017. https://doi.org/10.1016/j.envsoft.2017.07.016

Line 256 in the study area, - Done

Line 259 First sentence seems repetitive - Done

Line 260 I do not think that the word ‘alternation’ here is the right one – Done (now:

change)

Line 263 to 0.18% - Done

Line 265 Awkward, please rephrase - Done

Line 266 change approved by to supported - Done

Line 270 measurements - Done

Line 274 like Kriging - Done

Line 275 area, and it is an important step - Done

Line 276 and the produced bathymetric derivatives - Done

Line 290 after a distance of 400m - Done

Line 345 I would suggest to avoid finishing a sentence with too - Done

Line 356 For our data, - Done

Figure 12 b) for which mtry and c) for which ntree? - Done

Line 385 Table 5 MAE, MSE and RMSE were not introduced previously, only MSR was mentioned in to method section - Done

Line 413 The analysis of RF variable importance - Done

Line 414 specific depth ranges - Done

Line 415-417 I would suggest removing this sentence as it is not necessary - Done (removed from here and added to the discussion part as we consider that is important to refer that other authors have found such relationships in the marine and terrestrial environment).

Line 418 All of them also contribute in a nonlinear way - Done

Line 444 the study area, equal to - Done

Lines 458-459 Awkward, please rephrase - Done

Lines 468-471 Awkward, please rephrase - Done

Line 471 Conversely, - Done

Line 475 clues as to why - Done

Line 491 Along these lines, several authors have included - Done

Line 513 ‘a priori’- Done

Line 516 as well as their size - Done

Line 518 interest). Finally, - Done

Line 561 the remaining derivatives - Done

Lines 565-566 training and testing records - Done

Line 570 Should be 4.1 - Done Throughout, ground truth vs ground-truth, hydro-acoustic vs hydroacoustic, circa vs ca., space vs no-space between value and unit: - Done (ground-truth, hydroacoustic, ca. and space between value and unit, were selected)

Please also note the supplement to this comment:
https://www.biogeosciences-discuss.net/bg-2018-353/bg-2018-353-AC1-supplement.pdf

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2018-353, 2018.

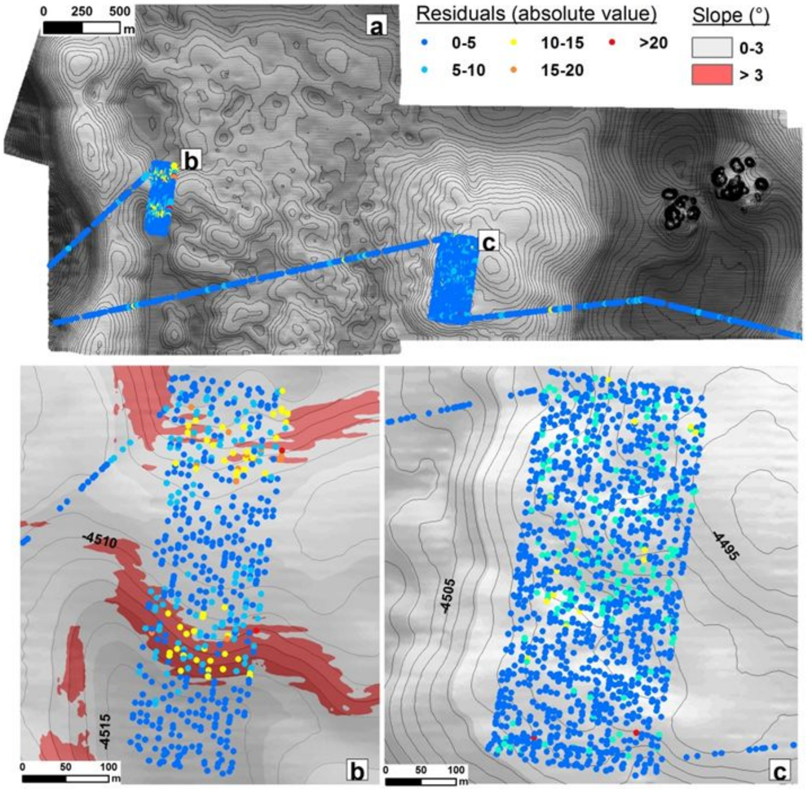

Figure 1. Spatial plotting of the RF residuals (absolute values). The intervals of their range are in accordance with the Table B9 (Appendix B) in the submitted manuscript.

Fig. 1.

Printer-friendly version

Discussion paper

Biogeosciences Discuss.,
https://doi.org/10.5194/bg-2018-353-AC2, 2018

Biogeosciences Discussions

BGD

Interactive comment

Interactive comment on "Quantitative mapping and predictive modelling of Mn-nodules' distribution from hydroacoustic and optical AUV data linked by Random Forests machine learning" *by* Iason-Zois Gazis et al.

Iason-Zois Gazis et al.

igazis@geomar.de

Received and published: 5 October 2018

Reviewer #2 General comments:

This paper was very interesting to read and clearly demonstrates how combining several state of the art scientific tools can achieve results that were, until recently, difficult to produce. The main idea in the manuscript; using Machine Learning to derive abundance of nodules from predictor variables remotely sensed with an AUV, has been applied by these authors and others datasets and this paper combines data and methodologies that have both been featured in other publications (cited in the paper). However, it presents a thorough protocol to make use of these tools, combine and optimize them, account for known caveats in the procedure and demonstrates the applicability of this protocol in a practical situation. The scientific approach is complex but transparently detailed throughout the method, results and appendixes. Thus, this paper is a useful case study and a method that should be applicable to other similar datasets and, as such, is a valuable contribution to the exploration of the Manganese nodules fields in the CCZ. It is well written but could be streamlined and made easier to read. The important findings could be further highlighted in the results section by moving some of the subsections in the appendix (as highlighted by reviewer 1). In addition, I found that several sentences or groups of sentences in the discussion either were confused in their formulation or didn't make a clear point. Furthermore, the discussion could be structured into several paragraphs to help readers perceive the different points made by the authors.

Authors comments:

We welcome all comments of Reviewer #2 and we appreciate the time and effort put to review this manuscript. Below we present our reply for each of the reviewer's points:

Similar to Reviewer #1, Reviewer #2 highlights the transparent, thorough and well-written workflow, and notices also that this methodology has been applied in the past, as well as the need for a different structure of some parts in the manuscript. Both comments have been already answered to Reviewer #1 and are considered in the revised version. In addition, we followed the recommendation of the Reviewer #2 to divide the discussion part into several paragraphs in an effort to state clearly the points of our study.

Reviewer #2 Specific comments:

I also have a couple of specific remarks and suggestion to add to those of reviewer 1: R400: If RF is not good at predicting outside the ranges of the training set, could it affect the projected map of nodule abundance? Other studies projecting RF models (of species distribution) in space (or time) have used multivariate environmental similarity surfaces (MESS) maps (Elith et al. 2010). This procedure is mapping how dissimilar to known data points the predictors are across the projection area. This could potentially highlight that predictions in deeper and shallower areas than where nodule abundance samples are should be considered with care. This could also help target areas for future sampling. See Elith J, Kearney M, Phillips S (2010) The art of modelling range-shifting species. Methods in Ecology and Evolution 1:330-342

Authors comments:

Doubtlessly, the 'weakness' of the RF method is to predict outside of the range of the training set, this can influence the accuracy of the final abundance map. The need for extrapolation is always given in deep ocean studies by the limited numbers of actual samples. The problem of having not 'entirely' representative samples can only be solved by collecting a great number of sample points (like images in our case) that are well-distributed inside the study area (i.e. data that will include the entire range of the number of Mn-nodules/m2 and they are come from all the different sub-terrains). The comparative use of different machine learning algorithms (Support Vector Machines and Artificial Neural Networks) for the same dataset, which are able to extrapolate beyond the training range (e.g. Balabin and Lomakina, 2011; Martious and Lambert, 2017), can reveal the size of this 'weakness' in RF predictions. Such extrapolated predictions should be treated carefully regarding their accuracy and should always been validated with samples from the outer parts (lower and upper) of the training range. The main difficulty of our approach, is the need for different representative large training data in every different study area. The use of multivariate environmental similarity surfaces (MESS) can contribute to Mn-nodule exploration, by indicating other similar Mn-nodules fields in the wider area, based on the similarity of morphological characteristics of the already studied areas. To our knowledge the combined use of RF and MESS has not been applied yet as. Elith et al. (2010) used Boosted Regression Tree

Printer-friendly version

Discussion paper

(BRT) and Maximum Entropy (MaxEnt) machine learnings approaches; an approach interesting for future studies. Another promising, although complex would be the use of the Transfer Learning Approach. This approach can overcome the drawback of traditional machine learning, in which the training predictive algorithms should be trained each time based on previously collected (labelled or unlabeled) data from the study area. By using Transfer Learning, one can take an already trained model and transfer the part of the model that contains the necessary built relationships into a new model (usually smaller) that has to learn only the extra relationships/patterns that may exist in the new study area (e.g. Pan and Yang, 2010; Lu et al, 2015). Thus, the non-linear relationship between the number of Mn-Nodules/m2 and the topographic factors can be transferred and applied to other potential areas, where there is a lack of labelled optic data, and may include slightly different bathymetric range and topographic characteristics.

1. Balabin, R.M. and Romakina, E.I.: Support vector machine regression (LS-SVM) - an alternative to artificial neural networks (ANNs) for the analysis of quantum chemistry data? Phys. Chem. Chem. Phys., 13, 11710–11718, 2011. https://doi.org/10.1039/c1cp00051a

2. Martius, G., Lampert, C.H.: Extrapolation and learning equations. CoRR abs/1610.02995, 2016. http://arxiv.org/abs/1610.02995

3. Elith, J., Kearney, M. and Phillips, S.: The art of modelling range-shifting species. Methods in Ecology and Evolution, 1, 330–342, 2010. https://doi.org/10.1111/j.2041-210X.2010.00036.x

4. Pan, S. J., and Yang, Q.: A Survey on Transfer Learning. IEEE Transactions on knowledge and data engineering, Vol.22, No. 10, 1345-1359, 2010.

5. Lu, J., Behbood, V., Hao, P., Zuo, H., Xue, S., Zhang, G.: Transfer learning using computational intelligence: A survey. Knowledge-Based Systems, 80, 14–23, 2015. http://dx.doi.org/10.1016/j.knosys.2015.01.010

R415: The relevance of the depth as the most important predictor could be discussed further. Is there a geological reason why depth is the main driver of nodules distribution (as it looks unintuitive as to why such small changes in depth could drive nodule distribution)? Is it likely to be a proxy for another driver?

Authors comments:

This question is also asked in the specific comment (7) from Reviewer #1, and it is answered there.

R499: Minor point but Judging by figure 12, the relation between MSR and the different tuning parameters, particularly the number of training samples is not linear and thus, could either increase asymptotically towards a maximum or might continue increase logarithmically. Either way, It is unclear if more data would be a major improvement. Thus, collection of new data should focus on better-distributed data

Authors comments:

Indeed, it is not clear if more data would be a major improvement. The availability of more data and especially if they were better distributed, would most likely reinforce the model to build better and wider relationships between the predictor and response variables. This would allow keeping a larger number of validation data points. The need for more and better-distributed data has been stated in lines 498 – 504, especially when considering the spatial clustering inside the study area. The influence of the number of training data for model performance still remains a discussion point between studies showing an improvement by adding more data (e.g. Bishop, 2006), and other studies presenting stable performance of the model even if more data are added (e.g. Zhu et al, 2012).

1. Bishop, C.M.: Pattern Recognition and Machine Learning. Information Science and Statistics. Springer, Heidelberg, 2006.

2. Zhu, X., Vondrick, C., Ramanan, D., and Fowlkes, C. C.: Do We Need More Training

Data or Better Models for Object Detection? In BMVC, 3, 5, 2012.

R510: Given the rarity of corers data compared to photo data, would it not be better to take all cores where there is photos to strengthen the comparison between the two nodule counting methods? The photos of areas where some of the cores have been taken can still be excluded from the RF model and externally validated afterwards in order to make the best use of available ship time and data.

Authors comments:

In an ideal scenario with a sufficient number of available box-corers (many box corers, in which we are referring in the manuscript), both scenarios should be applied. The greatest number of them should be deployed in areas with photos in order to calculate better the factor between counted Mn-nodules in photos and in box-corers, and the rest in areas without photos in order to estimate the accuracy of the model in areas far away from the optic data but still inside the study area. However, in a realistic scenario the amount box-core samples will always be limited and thus they should be deployed in areas with photos to establish a better relationship between these two quantitative methods.

And a few technical corrections and suggestions -Authors comments:

R56: "data points"? "Data sets"? – Done we use data sets

R180: could you specify what the correction would be? - Done. This correction can be a simple factor that describes the ratio between the number of Mn-nodules seen in the photo and the number of nodules counted in box-corers (considering for the different spatial scales). Kuhn and Rathke (2014) used this approach, but also considered two different nodule size spectra.

1. Kuhn, T., Rathke, M.: Report on visual data acquisition in the field and interpretation for SMnN. Deliverable D1.31 of the EU-Project Blue Mining. BGR Hannover, 34 pp. 2017. www.bluemining.eu/downloads

R474: "resulting in biased results where Mn-nodules are bigger"? – Done (This phrase is in lines 472-473). The sentence has been changed.

R480: This is true, but is it necessary to state that here? Maybe it could be moved to the introduction

R476 - 485: It is hard to follow the authors point here. Do you mean that the observed influence of bathymetric factors on the nodule distribution cannot necessarily be explained? This observation is an interesting fact in itself and may lead to a better understanding of an underlying Mn-nodule formation process?

R490: "as it ignores"? – Done

R490: "To this end, several authors, have included the values of latitude/longitude and even LMI as predictor variables"? – Done (based on Referee's #1 suggestion)

516: "thus, high priority areas (e.g. these with highest commercial interest) can be targeted for sampling based on the results of optic data and RF modelling"? – Done

Please also note the supplement to this comment:
https://www.biogeosciences-discuss.net/bg-2018-353/bg-2018-353-AC2-supplement.pdf

Interactive comment on Biogeosciences Discuss., https://doi.org/10.5194/bg-2018-353, 2018.

Quantitative mapping and predictive modelling of Mn-nodules' distribution from hydroacoustic and optical AUV data linked by Random Forests machine learning.

Iason-Zois Gazis[1], Timm Schoening[1], Evangelos Alevizos[1], Jens Greinert[1,2]

[1]GEOMAR Helmholtz Centre for Ocean Research Kiel, Wischhofstrasse 1-3, 24148 Kiel, Germany
[2]Christian-Albrechts University Kiel, Institute of Geosciences, Ludewig-Meyn-Str. 10-12, 24098 Kiel, Germany

Correspondence to: Iason-Zois Gazis (igazis@geomar.de)

Abstract. In this study, high-resolution bathymetric multibeam and optical image data, both obtained within the Belgian manganese (Mn) nodule mining license area by the autonomous underwater vehicle (AUV) Abyss, were combined in order to create a predictive Random Forests (RF) machine learning model. AUV bathymetry reveals small-scale terrain variations, allowing slope estimations and calculation of bathymetric derivatives such as slope, curvature, and ruggedness. Optical AUV imagery provides quantitative information regarding the distribution (number and median size) of Mn-nodules. Within the area considered in this study, Mn-nodules show a heterogeneous and spatially clustered pattern and their number per square meter is negatively correlated with their median size. A prediction of the number of Mn-nodules was achieved by combining information derived from the acoustic and optical data using a RF model. This model was tuned by examining the influence of the training set size, the number of growing trees (*ntree*) and the number of predictor variables to be randomly selected at each RF node (*mtry*) on the RF prediction accuracy. The use of larger training data sets with higher *ntree* and *mtry* values increases the accuracy. To estimate the Mn-nodule abundance, these predictions were linked to ground-truth data acquired by box coring. Linking optical and hydroacoustic data revealed a non-linear relationship between the Mn-nodule distribution and topographic characteristics. This highlights the importance of a detailed terrain reconstruction for a predictive modelling of Mn-nodule abundance. In addition, this study underlines the necessity of a sufficient spatial distribution of the optical data to provide reliable modelling input for the RF.

1. Introduction

High-resolution quantitative predictive mapping of the distribution and abundance of manganese nodules (Mn-nodules) is of interest for both the deep-sea mining industry and scientific fields as marine geology, geochemistry, and ecology. The distribution and abundance of Mn-nodules are affected by several factors such as local bathymetry (Craig 1979; Kodagali, 1988; Kodagali and Sudhakarand, 1993; Sharma and Kodagali, 1993), sedimentation rate (Glasby, 1976; Frazer and Fisk, 1981; von Stackelberg and Beiersdorf 1991; Skornyakova and Murdmaa, 1992), availability of nucleus material (Glasby, 1973), and bottom current strength (Frazer and Fisk, 1981; Skornyakova and Murdmaa, 1992). As a consequence, the distribution and abundance of Mn-nodules is heterogeneous (Craig, 1979; Frazer and Fisk, 1981; Kodagali, 1988; Kodagali and Sudhakar, 1993; Kodagali and Chakraborty, 1999; Kuhn et al., 2011), even on fine scales of 10 to 1,000 m (Peukert et al., 2018a; Alevizos et al., 2018). This increases the difficulty for quantitative predictive mapping using remote sensing methods. Vast areas of the seafloor can be mapped by ship-mounted, multibeam echo-sounder systems (MBES). State-of-the-art MBES systems feature a low frequency (12 kHz) and can map ca. 300 km^2 of seafloor in 4,500 m water depth per hour. Hence, low-resolution regional maps can be created at a grid cell size of 50 to 100 m within which the main Mn-nodule occurrence can become apparent, based on the backscatter intensity (Kuhn et al., 2011; Rühlemann et. al., 2011; Jung et al., 2001). A general separation in areas of high and low abundance (kg/m²) of Mn-nodules seems possible, especially in flat areas where sedimentological changes and physical influences on the footprint size and incidence angle of the transmitted acoustic pressure wave can be corrected accurately (De Moustier, 1986; Kodagali and Chakraborty, 1999; Chakraborty and Kodagali, 2004; Kuhn et al.al., 2010 and 2011, Rühlemann et al., 2011 and 2013). However, the patchy distribution of Mn-nodules in size and number at meter-scale cannot be resolved with ship-mounted MBES data (Petersen, 2017). For an operational resource assessment, a higher resolution of few meters grid cell size is needed to supply accurate depth, slope, and Mn-nodule distribution variability (Kuhn et al., 2011). Supplementary to the spatial mapping by acoustic sensors, point-based measurements from box-corer samples are used as ground-truth data for training and validation of geostatistical techniques (e.g. kriging) in order to create quantitative maps of Mn-nodule abundance (Mucha et al., 2013; Rahn, 2017). However, the generally low number of ground-truth samples during surveys (usually below 10), their limited sampling area (typically 0.25 m^2) and the relatively large distance between them (> 1 nmi) prevent an accurate correlation with the ship-based MBES data and thus a good prediction of the total Mn-nodules' mass and distribution in large areas (Petersen, 2017). Importantly, the sparse sampling with box corers affects the performance of interpolation and geostatistical techniques, which are typically applied during data analysis (Li and Heap, 2011 & 2014; Kuhn et al., 2016). In this article, we address this challenge by combining high-resolution hydroacoustic and optical data sets acquired with an Autonomous Underwater Vehicle (AUV) and connecting those data with a Machine Learning (ML) algorithm (here Random Forests), in order to predict the spatial distribution of the number of Mn-nodules per square meter. Unlike geostatistical methods, ML can be used to incorporate information from different bathymetric derivative layers and to detect complex relationships among predictor variables without making any prior assumptions about the type of their relationship or value distribution (Garzón et al., 2006;

Lary et al., 2016). To this direction, first predictions have already been achieved (Knobloch et al. 2017; Vishnu et al., 2017; Alevizos et al., 2018). Here, we present a complete data analysis workflow for potential mining operations (Figure 1).

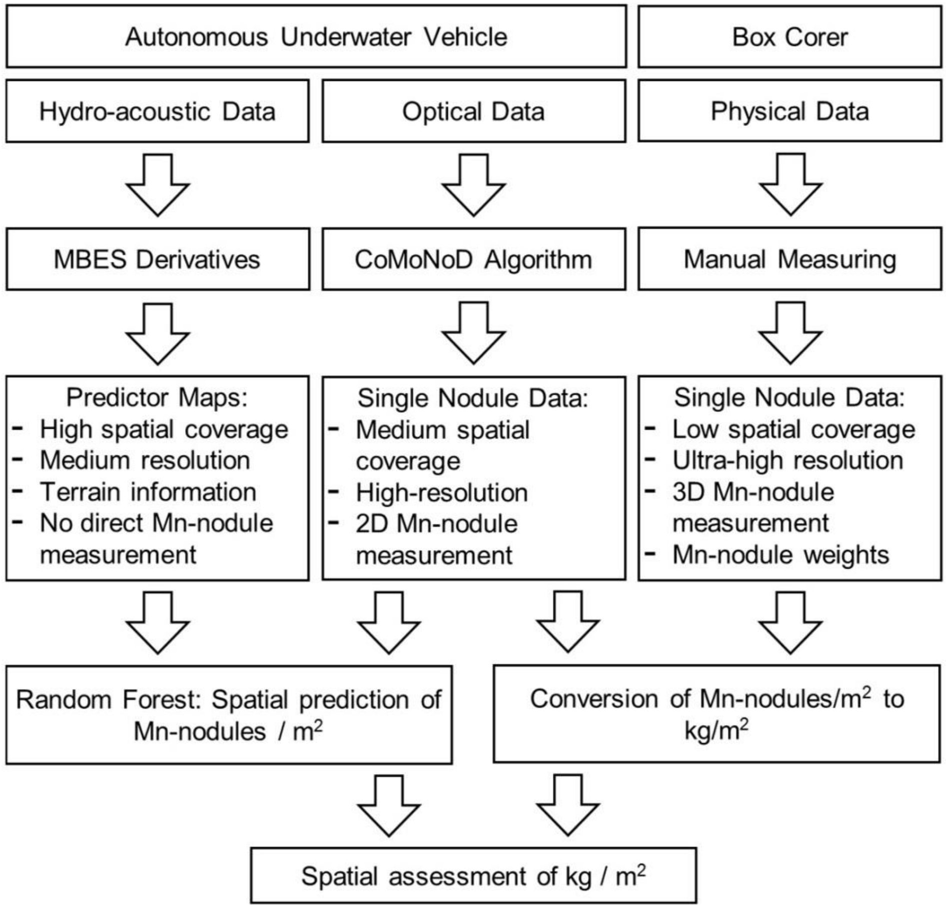

Figure 1. Schematic workflow of the data sets used in this study to enable the spatial assessment of Mn-nodules inside the study area. The medium resolution of AUV MBES (m scale) is referring to the comparison of the optical and physical data (cm scale).

1.1 AUV hydroacoustic mapping

AUVs have proven their usefulness for multibeam data acquisition in the deep-sea environment (Grasmueck et al., 2006; Deschamps et al., 2007; Haase et al., 2009; Wynn et al., 2014; Clague et al., 2014 and 2018; Pierdomenico et al., 2015; Peukert et al., 2018a). They achieve higher spatial and vertical resolution compared to ship-mounted MBESs. This is due to their operation close to the seafloor which results in a smaller footprint at a given beam angle and enables the use of higher frequencies (Henthorn et al., 2006; Mayer, 2006; Caress et al., 2008; Paduan et al., 2009). Additionally, AUVs avoid problems like near-surface turbulences, bubbles, ship-noise and strong sound velocity changes (Kleinrock et al., 1992a and 1992b; Jakobson et al., 2016; Paul et al., 2016). They work independently from the surface vessel and operate at a stable altitude. AUVs can efficiently conduct a dive pattern of dense survey lines and thus reduce survey effort and costs (Chance et al., 2000; Bellingham, 2001; Bingham et al., 2002; Danson, 2003; Roman and Mather, 2010). High-resolution bathymetry enables computing bathymetric derivatives like slope and rugosity with a similarly high resolution. These derivatives play an important role in predicting Mn-nodules' distribution and abundance (Craig, 1979; Kodagali, 1988; Skornyakova and Murdmaa, 1992; Kodagali and Sudhakar, 1993, Sharma & Kodagali, 1993; Ko et al., 2006). However, a small number of recent studies have investigated this role in an AUV scale (Okazaki and Tsune, 2013; Peukert et al., 2018a; Alevizos et al., 2018).

1.2 Underwater optical data

Underwater optical data have generally played an important role in the qualitative analysis of the seafloor features and for the specific task of assessing Mn-nodules' distribution explicitly (Glasby, 1973; Rogers, 1987; Skornyakova and Murdmaa, 1992; Sharma et al., 1993). The development of automated detection algorithms enabled quantitative optical image data analysis and subsequent statistical interpretation of Mn-nodule densities. The spatial coverage of optical imaging is much higher than for box core sampling. The data resolution remains high enough to reveal the high variance in the spatial distribution of nodules at meter scale. Thus optical data can fill the investigation gap between ground-truth sampling and hydroacoustic remote sensing (Sharma et al., 2010 and 2013; Schoening et al., 2012a, 2014, 2015, 2016 and 2017a; Kuhn and Rathke, 2017). Moreover, mosaicking of optical data could reveal mining obstacles such as outcropping basements or volcanic pillow lava flows. In addition, seafloor photos are the source for evaluating benthic fauna occurrences and related habitats on a wider area (Schoening et al., 2012b; Durden et al., 2016).

1.3 Box corer sampling

Box coring is common to obtain physical samples of Mn-nodules and sediments for resource assessments and biological studies. While optical data reveal only the exposed and semi-buried Mn-nodules, box corers collect the top 30-50 cm of the seafloor with minimum disturbance, allowing an accurate measure of the Mn-nodules' abundance (kg/m^2). Box coring data are used for training and validation in geostatistical methods for quantitative and spatial predictions of Mn-nodules (e.g. Mucha et al., 2013; Knobloch et al., 2017). The representativeness of box coring data is disputable as few deployments can be made due to time constraints (ca. 4h per core) and as the spatial coverage of one sample is rather low (ca. 0.25 m^2).

1.4 Random Forests

Random Forests (RF) is an ensemble machine learning (ML) method composed of multiple weaker learners, namely classification or regression trees (Breiman, 2001a). Within RF an ensemble of distinct tree models is trained using a random subsample of the training data for each tree until a maximum tree size is reached. In each tree, each node is split using the best among a subset of predictors randomly chosen at that node instead of using the best split among all variables (Liaw & Wiener, 2002). Thus, the process is double-randomized which further reduces the correlation between trees. About two thirds of the training data are used to tune the RF while the remaining 'out-of-bag' (OOB) samples are used for an internal validation. By aggregating the predictions of all trees (majority votes for classification, the average for regression) new values can be predicted. This aggregation keeps the bias low while it reduces the variance, resulting in a more powerful and accurate model. RFs have the ability to estimate the importance of each predictor variable which enables data mining of the high-dimensional prediction data. Terrestrial studies use RFs in prospectivity mapping of mineral deposits (Carranza and Laborte, 2015a; 2015b; 2016; Rodriguez-Galiano et al., 2014 and 2015). In the marine environment, RFs have been used to combine MBES bathymetry, backscatter, their derivatives, sediment sampling, and optical data for various seabed classification and regression tasks (e.g. Li et al., 2010; Li et al., 2011a; Che Hasan et al., 2014; Huang et al., 2014). Further studies showed the robustness of RFs for selected data sets compared to other ML algorithms (Che Hasan et al., 2012; Stephens and Diesing, 2014; Diesing and Stephens, 2015; Herkul et al. 2017), as well as to geostatistical and deterministic interpolation methods (Li et al., 2010, 2011b and 2011b; Diesing et al., 2014).

2. Study Area

The study area lies in the Clarion–Clipperton Zone (CCZ; ca. $4x10^6$ km^2) in the Eastern Central Pacific Ocean. The CCZ triggered scientific and industrial interest for several decades due to its high resource potential in Mn-nodules deposits (Hein et al., 2013; Petersen et. al., 2016) with an average nodule abundance of 15 kg/m^2 (SPC, 2013). At the time of writing, the International Seabed Authority (ISA) has granted 17 exploration licences inside the CCZ (Figure 2a). The study area described here is part of the Belgian GSR license area (Figure 2b) and will be referred to as Block G77 (Figure 2c). Overall, this part of the Belgian license area has high bathymetric range, and complex morphology, due to the presence of submarine volcanoes, solitary seamounts and seamount chains. Block G77 is characterized by a low bathymetric range (77 m) and mostly gentle slopes (95% of the area below 5°). An exception is located in the eastern part, where sub-recent small-scale volcanic activity created 15 cone-shaped morphological features of up to 55m height and 150m width that are clustered in an area of ca. 700 m x 380 m. Despite the gentle slopes, bock G77 is characterized by an uneven micro-relief (according to Dikau (1990)) especially in the western part, where small (2-4 m) depressions coexist next to short (2-4 m) protrusions. In the central part, a 30 m high elevation acts as a natural barrier between the western part of the study area that features more relief and the eastern part that is deeper and has less relief (Figure 2c).

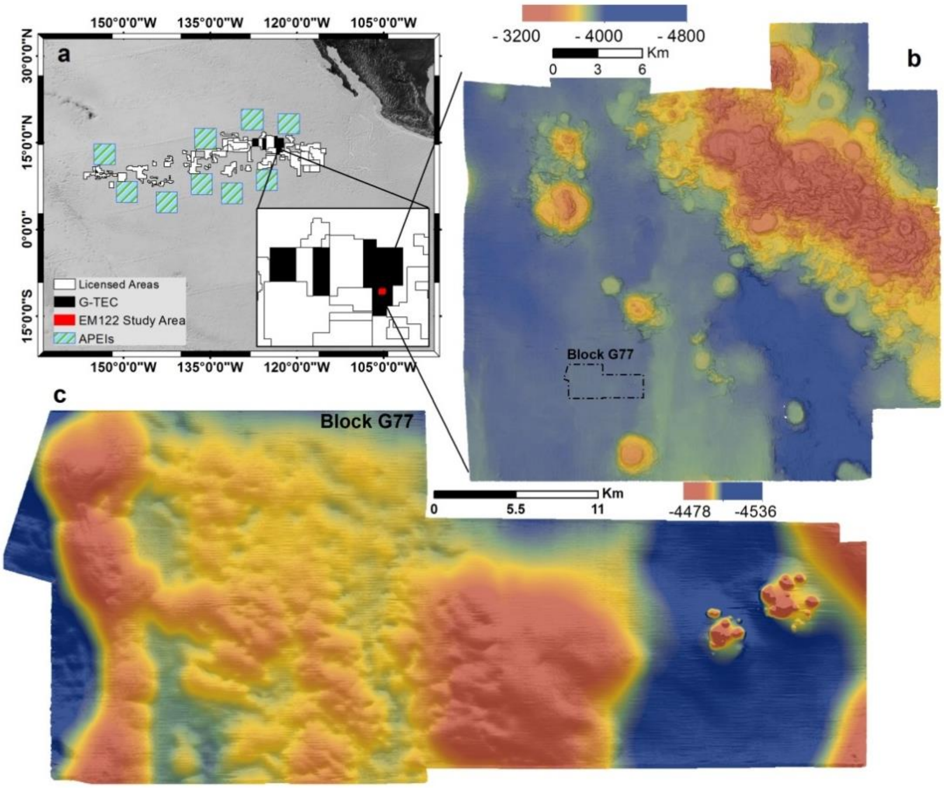

Figure 2. a) Areas of Particular Environmental Interest (APEIs), licensed areas (white) and the Belgium / GSR licenses area (black) within the CCZ. b) Regional bathymetric map of the study area, created by the EM 122 MBES on R/V SONNE (cruise SO239). c) Block G77, mapped by AUV Abyss with a Teledyne Reson Seabat 7125 MBES.

3. Methodology

3.1 Hydroacoustic Data Acquisition & Post Processing

The data (Greinert, 2016) were collected in March 2015 during cruise SO239 EcoResponse (Martínez Arbizu & Haeckel, 2015) with the German Research Vessel Sonne. Ship-based mapping was conducted with a hull-mounted Kongsberg EM 122 MBES (12 kHz, 0.5° along- and 1° across-track beam angle, 432 beams with 120° swath angle). High-resolution MBES data were acquired with AUV Abyss (GEOMAR, 2016) inside Block G77 equipped with a Teledyne Reson Seabat 7125 MBES (200 kHz, 2° along- and 1° across-track beam angle, 256 beams with 130° swath angle). The data (60 km of survey lines) were acquired from 50m altitude and with 100% swath overlap resulting in an insonification of 9.5 km^2. Post-processing of the AUV data was conducted with the Teledyne PDS2000 software for data conversion of the raw data into s7k and GSF format. Further multibeam processing (sound velocity calibration, pitch/roll/yaw/latency artifacts correction) was performed using the Qimera™ software. The largest uncertainties during AUV operations result from inaccurate navigation and localization in the deep-sea environment (Paull et al., 2014). AUV Abyss has a combination of five different systems for navigation and positioning: Global Positioning System (GPS) when at the sea surface, Doppler Velocity Log (DVL) when 100 m or less from the ground, Inertial Navigation System (INS), Long Baseline Acoustic Navigation (LBL) and dead reckoning (GEOMAR, 2016). Each system has its own limitations that contribute to the total navigation error (Sibenac et al., 2004; Chen et al., 2013) that generally results in positioning drifts over time. Consequently, this affects the position accuracy of the MBES and optical data. Our AUV MBES data processing and an absolute geo-referencing of the resulting AUV-bathymetry grid with the EM122 ship data, supplemented with the use of MBnavadjust in MB-Systems, (Caress et al, 2017) resulted in a well calibrated AUV bathymetric dataset. The position of the AUV image data 'only' relies on the above mentioned sensors with a not quantifiable 'small' position error. Backscatter data were excluded from the modelling procedure due to artifacts and a general poor quality. The output grid cell size for the analyses was set to 3 m x 3 m. The depth raster was exported as ASCII format for further analysis in SAGA GIS v.6.3.0. SAGA includes numerous tools that focus on DEM and Terrain Analysis (Conrad, 2015). Eight bathymetric derivatives were computed (Table 1) with the SAGA algorithms (Appendix A).

Table 1. The bathymetric derivatives computed in SAGA GIS and used as predictor variables.

Derivative	Description
Slope (S)	The first derivative of the bathymetry and describes the steepness of a surface.
Plan Curvature (Pl.C)	The second derivative of the bathymetry and perpendicular to the direction of the maximum slope (Zevenbergen and Thorne, 1987).
Profile Curvature (Pr.C)	The second derivative of the bathymetry and parallel to the direction of the maximum slope (Zevenbergen and Thorne, 1987).
Topographic Position Index (TPI)	Compares the elevation of a single pixel to the average of multiple cells surrounding it in a defined distance (Weiss, 2001).
Broad-scale (TPI_B)	Distance: 150-400 m
Medium-scale (TPI_M)	Distance: 50-150 m
Fine-scale (TPI_F)	Distance: 0-50 m
Concavity (C)	In each cell its value is defined as the percentage of concave downward cells within a constant radius (Iwahashi & Pike, 2007). Here, a 10 cell radius was used.
Terrain Ruggedness Index (TRI)	A quantitative measure of surface heterogeneity and can be explained as the sum change in elevation between a central pixel and its neighborhood (Riley et al, 1999). Here, a 10 cell radius was used.

3.2 Optical Data Acquisition & Post Processing

High-resolution optical data (20.2 Megapixels) was acquired by the DeepSurveyCamera system on board AUV Abyss (Kwasnitschka et al., 2016). During image acquisition the altitude above ground was 5 to 11 m, resulting in an overlap between the images of ca. 60% in each direction. In total, 11,276 photos acquired in block G77 (Greinert, 2017) and analysed with the automated image analysis algorithm CoMoNoD (Schoening et al., 2017a, 2017b and 2017c). For each image this algorithm delineates each individual Mn-nodule and provides quantitative information on each nodule (size in cm^2, alignment of main axis, geographical coordinate of the nodule). This information is further aggregated per image to provide the average number of Mn-nodules per square meter (Mn-nodules/m^2), the nodule coverage of the seafloor in percent and the nodule size distribution in cm^2 size quantiles. The algorithm has successfully been applied for quantitative assessment and predictive modelling of Mn-nodules (Peukert et al., 2018a, Alevizos et al., 2018). Nevertheless, the derived number of Mn-nodules/m^2 is subject to uncertainties due to the limitations of the CoMoNoD algorithm and the non-constant altitude of the AUV, especially in areas with slopes. The CoMoNoD algorithm cannot detect sediment-covered Mn-nodules due to the low or non-existent contrast. It may count two or more adjacent small Mn-nodules as one big nodule or misinterpret benthic fauna or rock fragments with similar visual features as Mn-nodules. The CoMoNoD algorithm fits an ellipsoid around each detected Mn-nodule, which limits the first two disadvantages as it splits huge Mn-nodules and accounts for potentially buried parts (see discussions in Schoening et al., 2017a). In general, the first two disadvantages lead to underestimations while the third one results in an overestimation of the number of Mn-nodules per m^2. These limitations are common and the need for corrections (e.g. a factor that describes the ratio between the number of Mn-nodules seen in the photo and the number of nodules counted in box-corers, considering for the different spatial scales), has been acknowledged (Sharma and Kodagali, 1993; Sharma et al., 2010 and 2013; Tsune and Okazaki, 2014; Kuhn and Rathke, 2017). Recent studies show that the difference between image estimates and the abundance in box corer data (due to sediment covered Mn-nodules) can be two to four times higher (Kuhn and Rathke, 2017). In this study, none of the box-corers was obtained exactly at a location for which optical data exists, thus no direct comparison and verification exist. Taking box corer samples for verification requires Ultra Short Baseline (USBL) navigation and imaging of the seafloor prior to the physical sampling. The effects of the non-constant flying altitude on the detection of Mn-nodules per m^2 are explained in detail below. For each photo location, the depth and the bathymetric derivative values were extracted from the hydroacoustic data. As no absolute geo-referencing could be performed for the AUV-based photo surveys, drifting sensor data will have an effect on the alignment between bathymetric and photo information, which was considered, while interpreting the results.

3.3 Data Exploration and Spatial Analysis

The data exploration, spatial plotting and analysis was performed with ArcMap™ 10.1, PAST v3.19 (Hammer et al., 2001), and R (R, 2008). All data were projected as UTM Zone 10N coordinate system (to enable spatial analysis). The existence of spatial autocorrelation in the distribution of Mn-nodules/m^2 was examined by the Global Moran's Index (GMI) and Anselin Local Moran's Index (LMI). Both, GMI (Moran, 1948 and 1950) and LMI (Anselin, 1995) are well-established for examining the overall (global) and local spatial autocorrelation, respectively (e.g. Goodchild, 1986; Fu et al., 2014). GMI attains values between -1 and 1 with high positive values indicating strong spatial autocorrelation. High positive LMI index values indicate a local cluster. This cluster could be a group of observations with high-high (H-H) or low-low (L-L) values regarding the examined variable. A high negative index value implies local outliers, like high-low (H-L) or low-high (L-H) clusters, in which an observation has a higher or lower value in comparison to its adjacent observations. Both Moran's Index analyses were performed in ArcMap™ 10.1 (for parameter settings see Appendix A). One decimal was retained in the presentation of the results from statistical analysis and RF modelling.

3.4 Box corer Data

A total of five box-corers (0.5 m x 0.5 m surface area) were obtained close to the study area (coordinates not given due to confidentiality). However, one is located within Block G77 (Figure 3a); this is the result of independent sampling schemes and purposes during the cruise. Nevertheless, all box core samples (maximum distance <1.5 km), were analyzed and used for further analyses. In each box-corer, the number, size, and weight of nodules were measured and the abundance (kg/m^2) was estimated (mean value: 26.5 kg/m^2). The total number of Mn-nodules within each box corer was compared with the number of Mn-nodules on the surface resulting in an average ratio of 1.32 (Table 2). This means that ≈ 25% of the nodules are not seen on the surface but are completely buried within the sediment (down to a depth of about 15 cm).

Table 2. The number of Mn-nodules on the sediment surface, the total number of Mn-nodules per box core, the ratio of those two values, and the distance of the box corer deployments from the study area in block G77.

box corer station	total number of Mn-nodules	number of Mn-nodules at the surface	ratio	abundance (kg/m^2)	distance from G77 area (km)
BC20	40	27	1.5	-	0
BC21	67	58	1.1	27.1	1.4
BC22	29	21	1.4	27.1	0.6
BC23	32	20	1.6	25.2	0.1
BC24	17	16	1.0	-	1
Average	**37**	**28**	**1.32**	**26.5**	**0.6**

3.5 RF Predictive modelling

The RF modelling was performed with the Marine Geospatial Ecology Tools (MGET) toolbox in ArcMap™ 10.1. MGET (Roberts et al., 2010) uses the *randomForests* R package for classification and regression (Liaw and Wiener, 2002). Our target variable (number of Mn-nodules/m^2) is continuous, so regression was applied. We followed the three main steps to establish a good model by selecting predictor variables, calibration/training of the model and finally validating the model results.

Selection of Predictor Variables: The depth (D) and its derivatives (Table 1) were used as predictor variables. Although RFs can handle a high number of predictor variables with similar information, the exclusion of highly correlated variables can improve the RF performance and decrease computation time (Che Hasan, 2014; Li et al., 2016). Thus, the correlation between derivatives was investigated using the Spearman's rank correlation coefficient. None of the variable pairs was perfectly correlated ($\rho \geq 95$) and consequently, all of them were used for RF modelling (Appendix A).

Calibration of the model: During the calibration process, the RF parameters were adjusted as follows. The number of predictor variables to be randomly selected at each node (*mtry*), the minimum size of the terminal nodes (*nodesize*) and the number of trees to grow (*ntree*) were set to the default values, in order to investigate the optimum training size. For regression RF the default *mtry* value is 1/3 of the number of predictor variables (rounded down), *nodesize* is 5 and *ntree* is 500 (Liaw and Wiener, 2002). RF has demonstrated to be robust regarding these parameters and the default values have given trustworthy results (e.g. Liaw and Wiener, 2002; Diaz-Uriarte and de Andres, 2006; Cutler et al., 2007, Okun and Priisalu, 2007; Li et al., 2016 & 2017). With regards to the subsampling method (*replace*), the subsampling without replacement was selected. Although the initial implementations of the RF algorithm use subsampling with replacement (Breiman, 2001a), later studies showed that this process might cause biased selection of predictor variables that vary in their scale and/or in their number of categories, resulting in a biased variable importance measurement (Strobl et al., 2007, 2009; Mitchell, 2011). Based on recent findings, the raw variable importance was preferred (*unscaled*) as the final parameter (Diaz-Uriarte and de Andres, 2006; Strobl et al., 2008a, 2008b, 2009). Using these settings, the influence of the training sample size was examined (10 to 90% of the total sample in steps of 10%) and compared based on the Mean of Squared Residuals (MSR) using the respective equation provided in the *randomForests* R package (Liaw and Wiener, 2002). The different training groups need to be considered as representative of the total sample, in order to capture the heterogeneity of the Mn-nodules' spatial distribution. The spatially random selection of subsamples by MGET ensured similar statistical characteristics in each group (Appendix A). For each case of different training sample size, the model was run ten times and the results are presented as the average value of these ten runs (Appendix B). Since the optimal training sample size was defined, the influence of the number of growing trees (*ntree*) and the influence of the number of predictor variables to be randomly selected at each node (*mtry*) was examined. Only for the already defined optimum training size ten different *ntree* values (100 to 1000 in steps of 100) and seven different *mtry* values (1 to 7 in steps of 1) were tested and compared based on the MSR values. In each case of different *ntree* and *mtry* parameter, the model was run ten times and the results are presented as the average value of these ten runs (Appendix B). Selection and external validation of the optimal model: Based on the above-mentioned results and considering the sampling and computational cost, the optimal model was selected, run for 30 iterations and applied to the entire study area. Its predicted values were validated with the observed values from the remaining dataset that was not used. Several validation measures were used including the Mean Absolute Error (MAE), the Mean Squared Error (MSE) and the Root Mean Squared Error (RMSE). The combined use of MAE and RMSE is a well-established procedure as the MAE can evaluate better the overall performance of a model (all individual differences have equal weight), while the RMSE gives disproportionate weight to large errors showing an increased sensitivity to the presence of outliers. Due to this characteristic RMSE is suitable for outlier detection analysis but should not be used solely as an index for the model performance (Willmott & Matsuura, 2005). Both, MEA and RMSE are measured in the same unit as the data. In addition, the R^2, Pearson (r) and Spearman's rank correlation coefficients were used to identify the correlation between predicted and initial values. Finally, the descriptive statistics of predicted and initial values were compare and a residual analysis was performed.

3.6 Resource Assessment

As the optimal RF model was applied to the entire Block G77, an estimate of the abundance (kg/m^2) was computed, based on the analogy between the corresponding abundance measured from the average number of Mn-nodules in the box corer data and the number of Mn-nodules/m^2 in each cell of the final result of the RF model. Considering that the collector can recover buried Mn-nodules from a maximum depth of 10-15 cm (Sharma, 1993 and 2010), the ratio of 1.32 was applied to account for Mn-nodules not detected in the images, and areas with a slope of >3° were excluded, assuming that a potential mining vehicle is limited to less steep slopes (UNOET, 1987).

4. Results

4.1 Data Exploration

The analysis of AUV photos with the CoMoNoD algorithm (Schoening et al., 2017a) revealed a rather heterogeneous pattern of Mn-nodules/m^2 in the study area, showing adjacent areas with high and low Mn-nodules number (Figure 3a). The number of Mn-nodules/m^2 changes within less than 100m in the overall study area and in the two main sub-areas b and c (Figures 3a-c). In half of the photos (48%), the number of Mn-nodules/m^2 varies from 30 – 43 with the mean value being 36.6 Mn-nodules/m^2. The very small change of 5% trimmed mean value indicates the absence of extreme outliers, which is confirmed by box-plot analysis (Appendix B). Further analysis of their descriptive and distribution characteristics was performed in order to assess the presence of normality in the data, resulting that the number of Mn-nodules/m^2 is approximately normal distributed (Appendix B).. Although the presence of normality in data is not a prerequisite assumption in order to perform the RF (Breiman, 2001a); as it is with geostatistical interpolation techniques like kriging (e.g. Kuhn et al., 2016), this examination can give us a better understanding of the Mn-nodules' distribution inside the study area, and it is an important step in order to examine potential extreme observations which may be derived from wrong measurements and could artificially change the training range during RF predictive modelling. Moreover, absence of linear correlation was observed between Mn-nodules/m^2 and the produced bathymetric derivatives, indicating the complexity of the phenomenon (Appendix B).

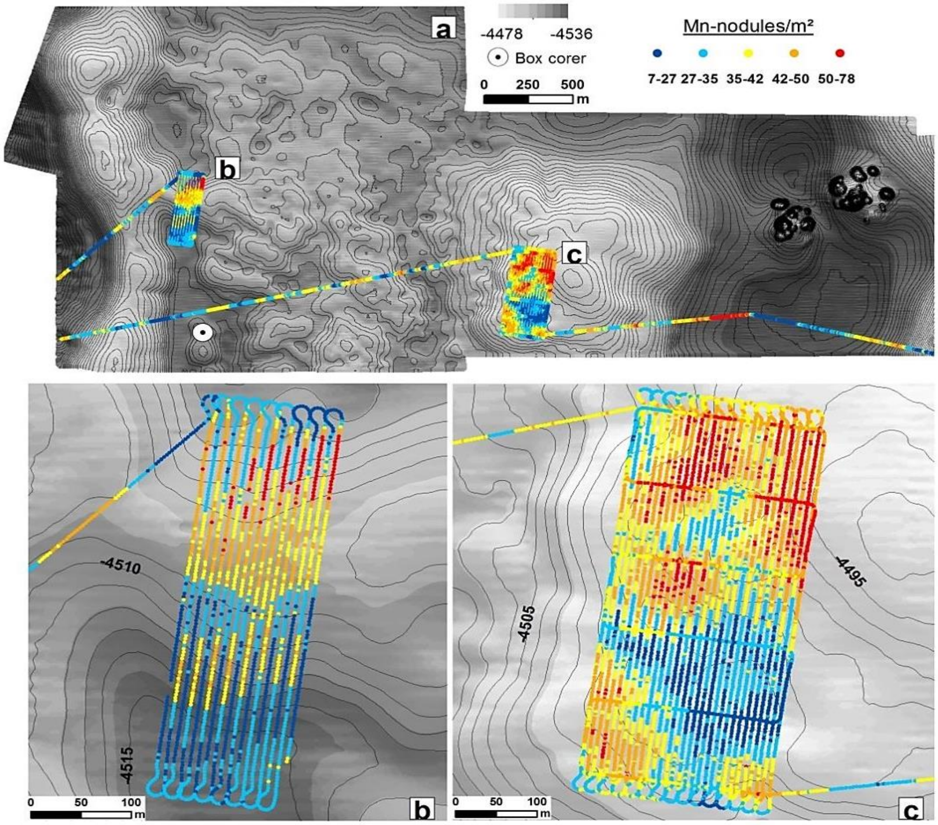

Figure 3. a) The spatial distribution of Mn-nodules/m^2 inside block G77 and the box corer position. b) The spatial distribution of Mn-nodules/m^2 inside the sub-area b. c) The spatial distribution of Mn-nodules/m^2 inside the sub-area c.

4.2 Spatial Analyses

Spatial analyses revealed the presence of a spatial autocorrelation in the distribution of Mn-nodules/m^2. The GMI, with I=0.6989, $p<0.01$ and Z-score>2.58 indicates a positive spatial autocorrelation. According to the incremental analysis, the index takes its highest value in the first 50 m with a gradual decrease, approaching 0 values after a distance of 400 m (Figure 5a). Similarly, the results from the LMI show that the main size of the spatial clusters does not exceed 400m in either direction (Figure 6a). The main types of these clusters are H-H and L-L groups (Table 4 & Figure 6a). A distinct 'buffer/transitional zone' with Mn-nodules was found between these two clusters, which does not show a significant autocorrelation (Figure 6b & 6c). Approximately one-third of the data does not have a significant clustering (NS). In the sub-area c, in the outer parts of these zones without significant spatial clustering, the few local H-L and L-H groups are located.

Both H-L and L-H (from the entire study area) only account for 2.1% of the data (Table 4). The comparison of the number of Mn-nodules/m^2 between the groups shows a clear discrimination between H-H and L-L clusters (Figure 5b). The H-H clusters are in areas with 37.9-78.2 Mn-nodules/m^2 whilst the L-L clusters are in areas with 6.8-35.2 Mn-nodules/m^2.

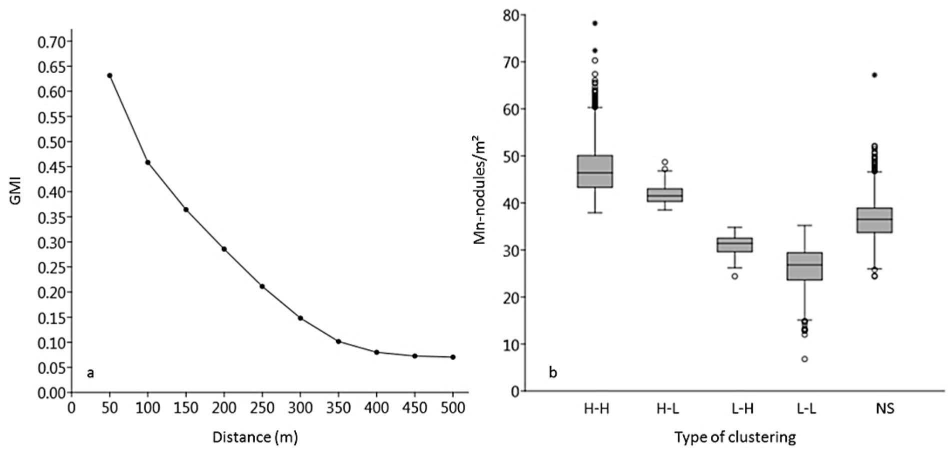

Figure 5. a) The GMI decrement due to increasing distance, after the first 50m. b) The range of Mn-nodules/m^2 in each clustered group.

Table 4. Number and % percentage of samples in each type of spatial clustering.

Cluster Type	H-H	H-L	L-H	L-L	NS
Counts (n)	3472	121	113	3523	4047
Counts (%)	30.8	1.1	1.0	31.2	35.9

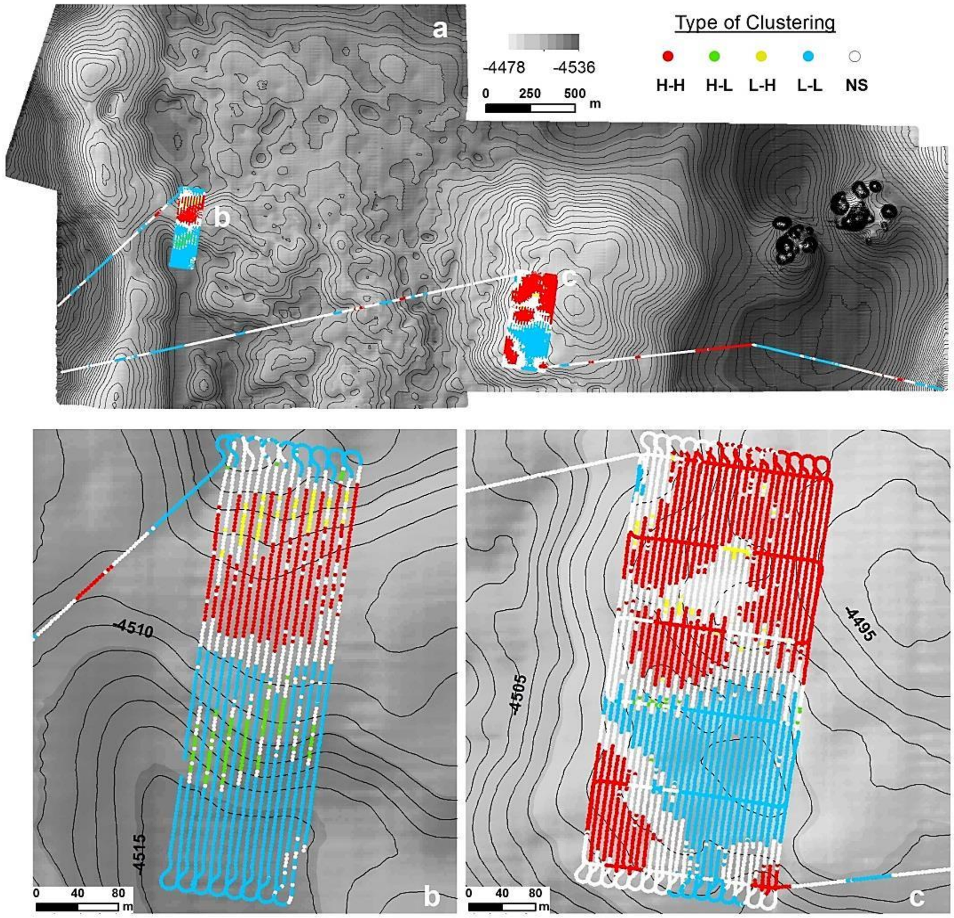

Figure 6. a) The spatial distribution of the significant cluster types inside the block G77. b) The spatial clusters inside the sub-area b. c) The spatial clusters inside the sub-area c

The application of the LMI reveals a bias that exists in the data due to the sampling procedure, especially in the sub-area b (Figure 6b). Here, the presence of the slope around 2.8° forced the AUV to vary its altitude between the ascending and descending phase (Figure 7b). This variation seems to affect the image quality resulting in counting fewer nodules for higher altitudes of the AUV (Figure 8 & 9). This is also confirmed by the distribution map of the Mn-nodules/m^2 (Figure 3b). It is important to emphasize that this difference clearly shows up in the LMI results (Figure 6b) and not in the distribution map (Figure 3b); here the arbitrary choice of color scale can hide this bias during plotting. The comparison of the detected Mn-nodules/m^2 in these adjacent lines, inside the small sub-area b, gives a ratio ≈1.4 between photos that have been acquired in 7-9 m altitude and those in 9-11 m altitude. The ratio is higher (≈1.8) between photos from 5-7 m and 9-11 m altitude. In contrast, the ratio between photos from 5-7 m altitude and those in 7-9 m altitude is ≈1.25 indicating that the problem mainly exists in upper and lower flying altitudes. Despite their different ratio, none of these groups contain extreme high or low values of Mn-nodules/m². Moreover, in several parts of the block, the photos from higher altitude are the only source of information without the ability for further comparison and consequently, they cannot be excluded from the modelling procedure.

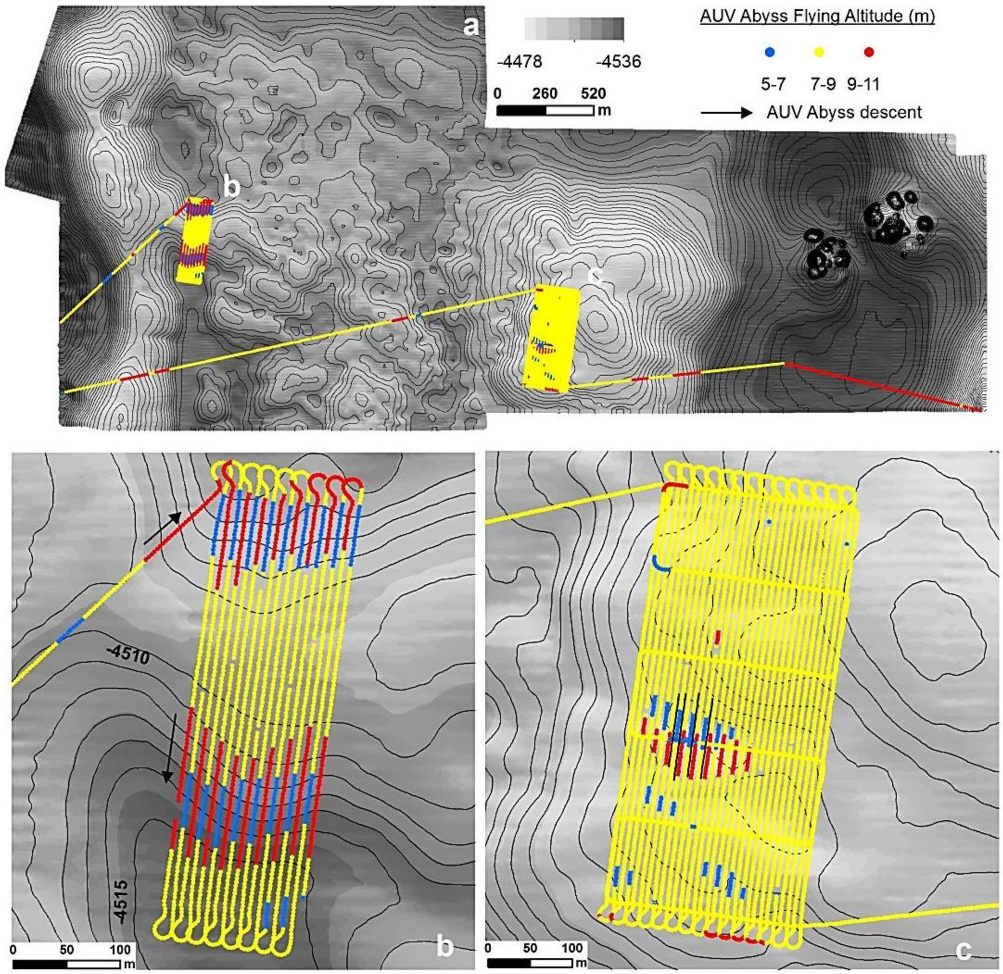

Figure 7. a) The altitude of AUV Abyss inside block G77. b) The altitude inside the small sub-area b, where the presence of the slope forces the AUV to modify its altitude, flying closer to the seafloor in the ascending phase (blue lines) and farther from the seafloor in the descending phase (red lines). c) In the big sub-area c, the AUV flying altitude is mainly constant between 7-9 m for the entire part.

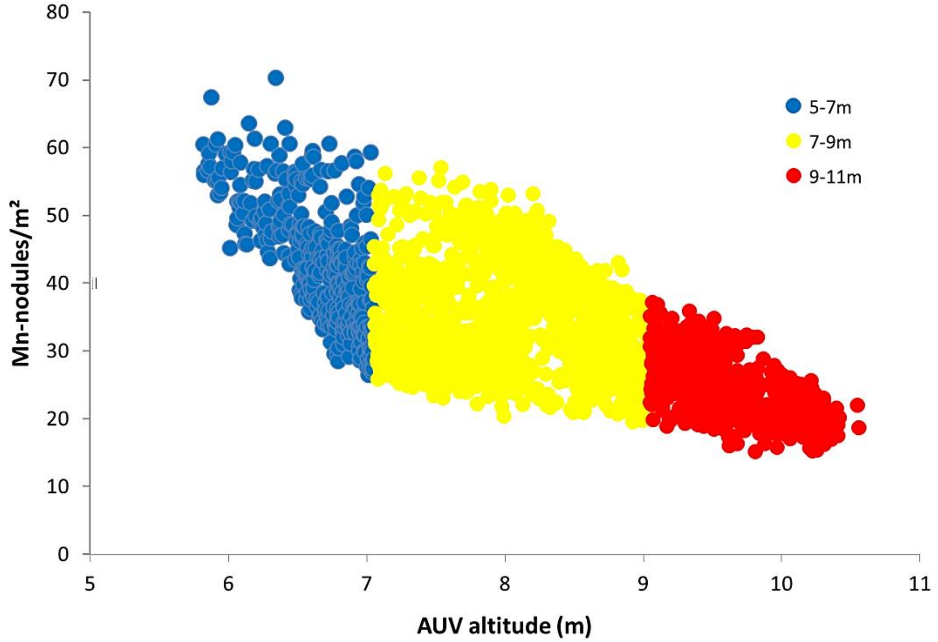

Figure 8. Scatterplot of the AUV altitude (m) and the estimated number of Mn-nodules/m² inside sub-area b. The colours correspond to the colour scale in figure 7.

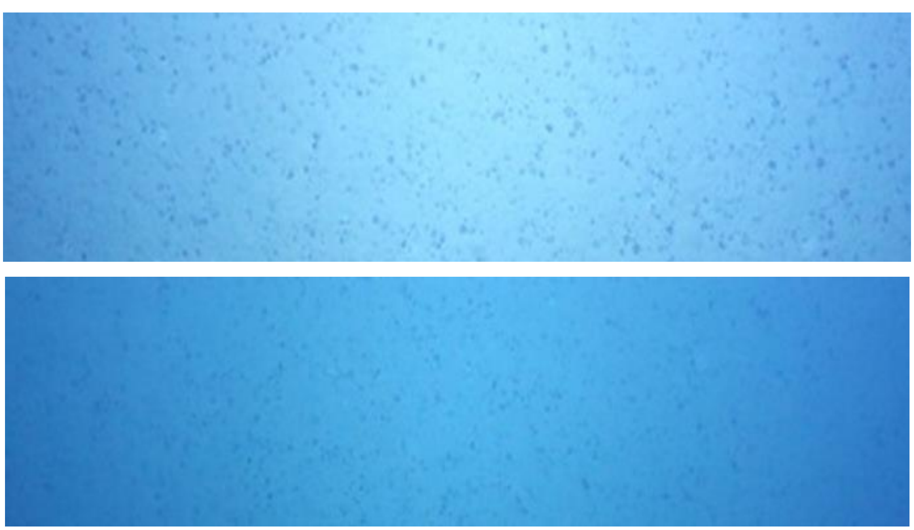

Figure 9. Adjacent AUV photos from consecutive dive tracks that have been obtained inside sub-area b, from: a) lower (5-7 m) and b) higher (9-11 m) altitude. Notice the decrement in the image brightness. (The area of the photos represent the central part of the photo, ca.1/4 of the original photo size).

Spatial distribution of median size: Plotting of the median size in cm^2 (Figure 10) showed that the number of Mn-nodules/m^2 is anti-correlated to the median Mn-nodule size. The Spearman's rank correlation coefficient and R^2 between these two variables are -0.50 and 0.25 respectively, supporting this observation (Figure 11a); other studies found similar results (Okazaki and Tsune, 2013; Kuhn and Rathke, 2017; Peukert et al., 2018a). The box plot analysis of the median size values between the H-H and L-L clustered groups showed that although the L-L group contains the entire range of median size values (2.8 to 15.9 cm^2), the H-H group does not contain values above 10 cm^2 (2.7-10 cm^2). This means in consequence that in areas with significant clustering of higher numbers of Mn-nodules/m^2 the size of Mn-nodules tends to be smaller (Figure 11b).

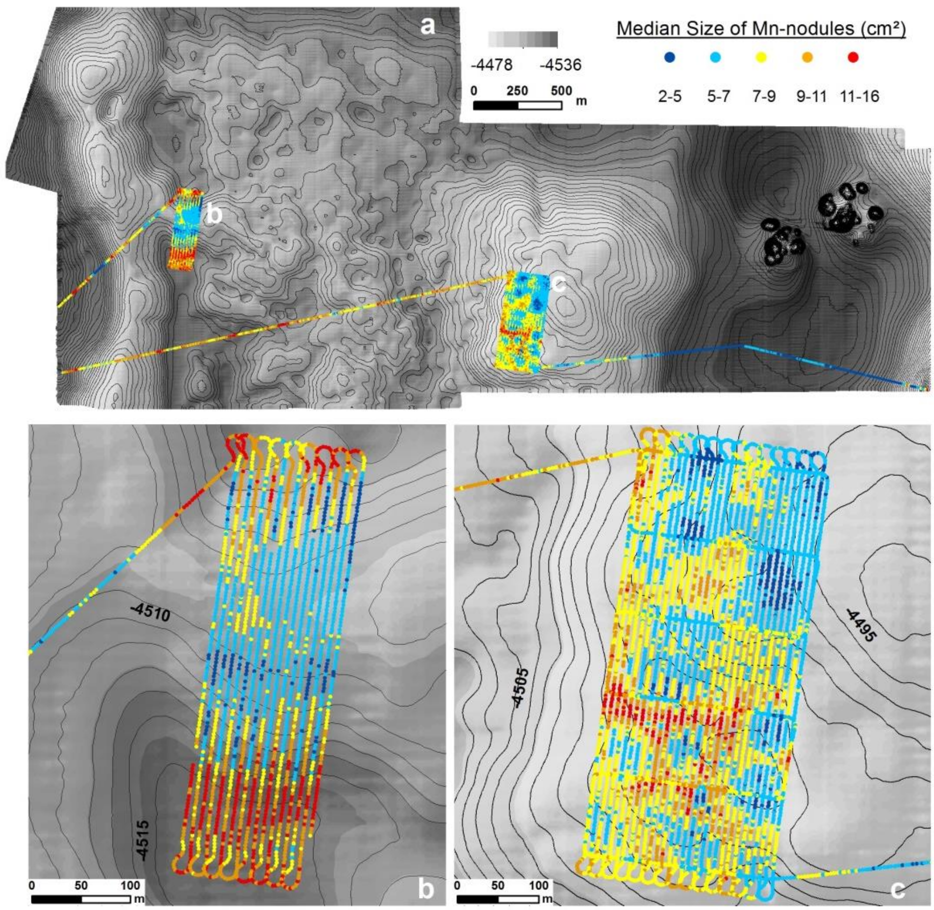

Figure 10. a) The spatial distribution of median Mn-nodule size (in cm^2). b) The estimation of median Mn-nodule size in sub-area b and mainly in its southern part has been probably affected by the non-constant altitude of the AUV. c) The distribution of the median size inside sub-area c shows also a clumped pattern.

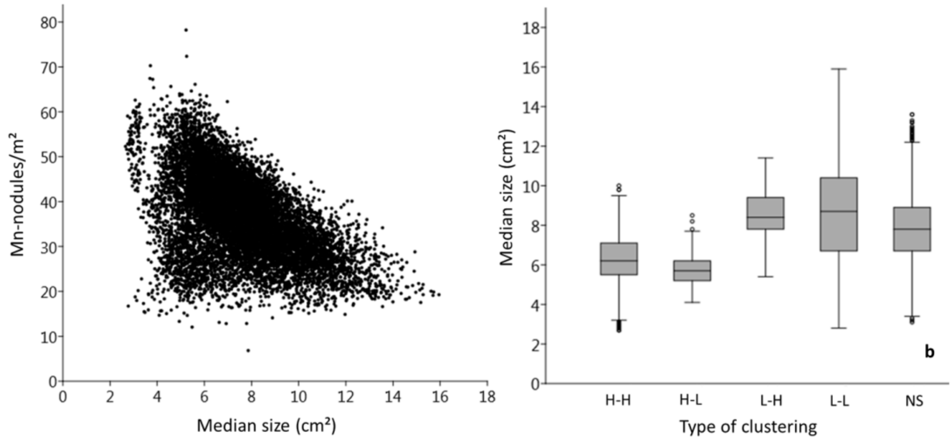

Figure 11. a) The plot of median size (cm^2) and number of Mn-nodules/m^2. e) The range of median size (cm^2) in type of cluster. Notice the distinct difference in the range between the H-H and L-L cluster type.

4.3 RF Predictive modelling

4.3.1 Effect of training sample size, *ntree* and *mtry* parameter

The results of the modelling procedure demonstrate that the RF algorithm is influenced by the size of the training sample (Figure 12a). This finding is in accordance with other studies, in which larger training samples tended to increase the performance of RF (Li et al., 2010 and 2011b; Millard and Richardson, 2015). The inclusion of a more representative range of the observed values and consequently larger spectrum of the causal underlying relationships, assist the RF to build a better model for the prediction of the value distribution inside the study area. For our data, the decrement becomes smaller when the size of the training sample increases further; it reaches a minimum value of 0.2 between 80% and 90%, showing that these additional 10% do not notably benefit the RF model. However, the absence of stabilization of the error to a minimum value indicates that more optical data are needed from this block. The small decrement in error between 80% and 90% was the decisive factor to select 80% of the data as training samples (also considering the larger number of remaining validation data and the reduced computational effort). Based on this dataset, the examination of different numbers of trees showed that the RF error remains constant after 600 trees (Figure 12b). Less trees result in a larger error; this particularly becomes evident with less than 300 trees. With more than 300 trees the range of the error is reduced (Appendix B). A higher number of trees enables higher *mtry* values as there are more opportunities for each variable to occur in several trees (Strobl et al., 2009). Similarly to the *ntree* parameter, a larger number of *mtry* values results in a reduced error (Figure 12c). The error reaches a minimum and cannot be reduced further for *mtry* = 6; with values below 3 the error increases significantly. The different numbers of *ntree* reduced the error by only 0.6 in the MSR (from 18.8 to 18.2), in contrast different *mtry* values reduced the error by 5.8 in the MSR (23.4 to 17.6), highlighting its importance for the prediction accuracy. In general a higher number of *mtry* values is suggested for RF studies with correlated variables to result in a less biased result regarding the importance of each variable; this is because the higher number increases the competition between highly correlated variables, giving more chances for different selections (Strobl et al., 2008a). The finally selected *mtry* value of 6 coincides with the recommended approach for *mtry* (default, half of the default, and twice the default) suggested by Breiman (2001a). Albeit the importance of this analysis, within the model with 80% of the data as training sample, the decrease in error by the use of RF tuned values instead of RF default values was only 0.7 in the MSR values, whilst the greatest reduction in error (16.5 in the MSR values) came from the increase in training data set size. This highlights the increased sensitivity of the method with respect to training data and that the recommended settings in the R *randomforest* package (Lia and Wiener, 2002) give trustworthy results, increasing its simplicity and operational character.

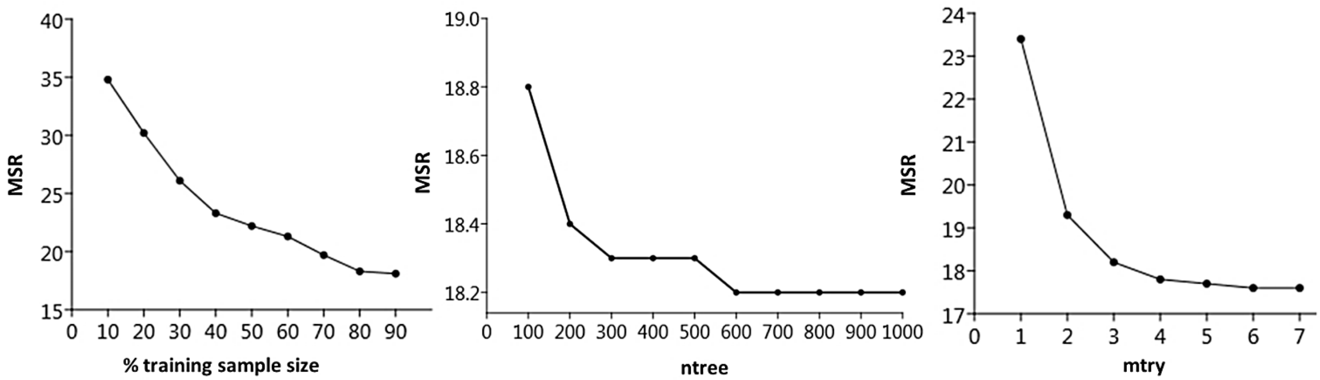

Figure 12. a) The effect of training sample size in RF error (in MSR). b) The effect of *ntree* parameter in RF error (in MSR) for the 80% training size. c) The effect of *mtry* parameter in RF error (in MSR) for the 80% training size.

4.3.2 Selection, application and external validation of the optimal model

Based on the above-mentioned findings, the optimal RF regression model which uses 80% of training data, 600 trees and 6 predictor variables to be randomly selected at each node, was selected and applied to the entire block G77. The comparison of the predicted values with the observed values from the remaining 20% (2,255 observations) of validation data showed a good predictive performance (Table 5). Analytically, MAE and RMSE have very low values, R^2 has a high value and both Pearson’s and Spearman’s correlation coefficients show a strong positive correlation between the predicted and observed values. The small deviation between MAE and RMSE and the same good correlation of the Pearson and Spearman factor point towards the absence of extremely high or low predicted values (outliers). Moreover, the performance is rather stable among all the iterations (Appendix B).

Table 5: The values of validation measures between predicted and observed data.

MAE	MSE	RMSE	R^2	Pearson	Spearman
3.1	19.0	4.4	0.8	0.9	0.9

The scatterplot and box plot (Figure 13a and 13b) illustrate this good match between predicted and observed values, as confirmed also by the descriptive statistics, which have almost equal mean, median, skewness and kurtosis values (Table 6). The residual analysis confirmed further the robustness of the model (Appendix B).

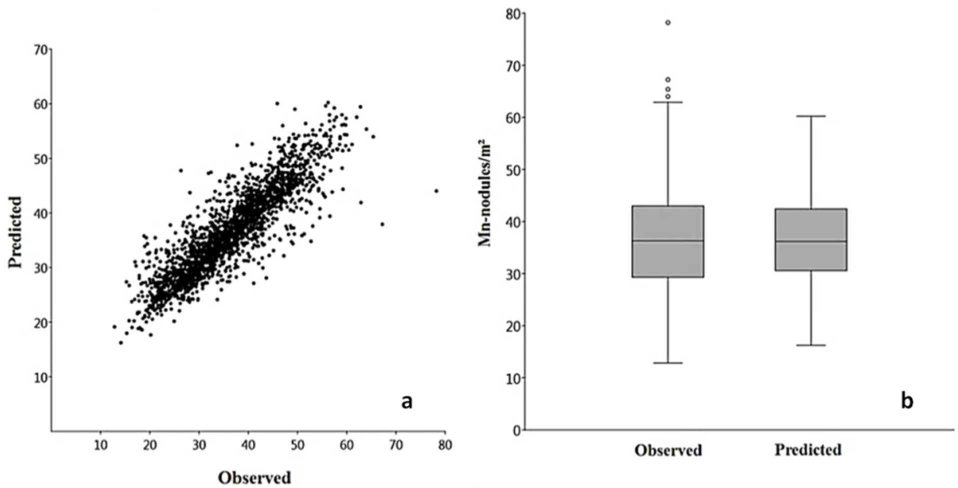

Figure 13. Comparison between observed and predicted values: scatterplot (left) and box-plots (right)

Table 6. Descriptive statistics of observed and predicted values

	Mean	Std. Error	5% Trim. Mean	Median	Mode	SD	Min.	Max.	C.L (95%)
Observed	36.5	0.2	36.3	36.3	40.8	9.4	12.8	78.2	0.4
Predicted	36.7	0.2	36.5	36.2	33.9	7.8	16.2	60.2	0.3

The statistical analysis also reveals the limitations of the RF model which cannot predict beyond the range of training values. It underestimates the maximum predicted values and overestimates the minimum values (Figure 13b & Table 6), a limitation also mentioned by other authors (e.g. Horning, 2010). This happens, because in regression RF the result is the average value of all the predictions (Breiman, 2001a).

4.3.3 RF predicted distribution of Mn-nodules/m²

The final application of the RF model for the entire block G77 predicts that the majority of the area is covered by 30-45 Mn-nodules/m² (Figure 14). In the central-western part the distribution is quite uniform (at this scale) with few small areas of lower numbers. In the western part, there are two extended areas along the base of the hill with the lowest number of Mn-nodules/m^2. Both of these areas have a linear shape in N-S direction and follow the seafloor topography with increased slope (>3°). The third main patch with minimum Mn-nodules/m^2 occurs in the eastern depression part. In contrast, areas of higher number of Mn-nodules/m^2 are located mainly in the central upper part of the hill and eastward facing slope of eastern depression and south of the sub-recent hydrothermally active area.

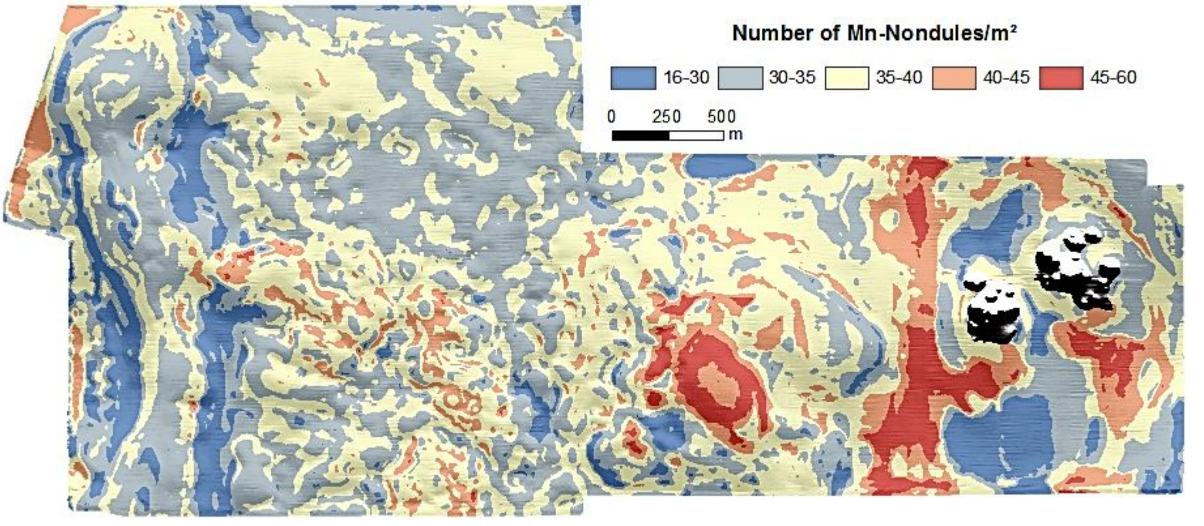

Figure 14. The RF predicted distribution of Mn-nodules/m² inside block G77.

4.3.4 RF importance

The analysis of the RF variable importance showed that the best explanatory variable for the distribution of Mn-nodules/m² is depth (Figure 15a). The partial dependence plot of depth shows that there are specific, which promote higher numbers of Mn-nodules/m² aggregated in a nonlinear way (Figure 15b). The following two most important variables are the TPI_B and TPI_M. TRI, TPI_F, C, and S follow in importance (Figures 15a). All of them also contribute in a nonlinear way. (Appendix B). Pl.C and Pr.C do not contribute significantly as explanatory variables in the performance of the RF model (Figure 15a and Apendix B). Although the RF demonstrates good overall performance, the small study area and the arbitrary choice of the spatial scales for the TPI and other derivatives, limit the potential of these variables as indicative explanatory variables on a broader scale. It is well established that surface derivatives are scaled-depended with different analysis scales to create alterations in results. Thus the combined use of different scales (here TPI) in the analysis and modelling procedure can produce models that do capture the natural variability and scale dependence (Wilson et al., 2007; Miller et al., 2014; Ismail et al., 2015; Leempoe et al., 2015). Due to the lack of relevant literature for AUV scale data sets, the C and TRI were created with the default scales of SAGA GIS v.6.3.0, while the three different TPI values were selected based on the minimum possible correlation among them.

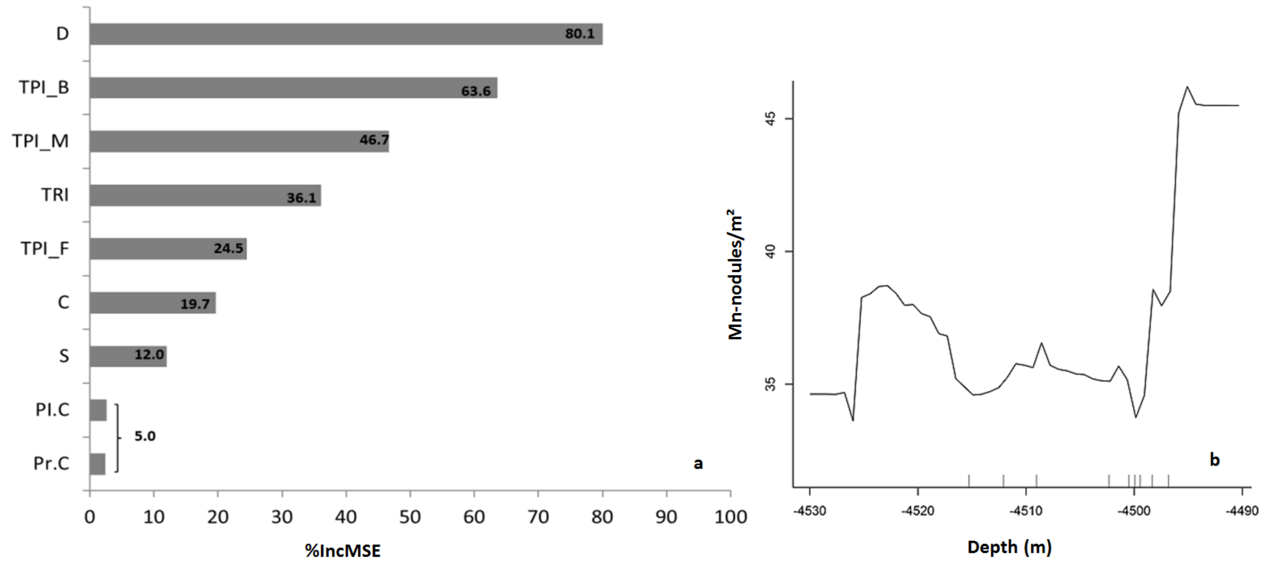

Figure 15: a) The variable importance of each predictor in the RF model. b) The partial dependence plot of the Depth (right). The ticks inside the graphs indicate the deciles of the data.

4.3.5 Estimation of abundance (kg/m²) of Mn-nodules

The predicted Mn-nodule distribution was combined with the abundance from box corer data (and corrected with the ratio between buried/unburied Mn-nodules, in order to include the top ~15 cm of the sediment), resulting in the Mn-nodules' abundance map shown in Figure 16. According to this map, block G77 is a promising area for mining operations. The entire block is above the cut-off abundance of 5 kg/m2 (UNOET, 1987), with a mean value of 33.8 kg/m2. We calculated that 84% of block G77 has slopes below 3°, steeper slopes are located mainly at the outer parts of the block, a fact that would ease establishing an ideal mining path. In this respect, the AUV-scale mapping provides vital information for a potential mining path by decreasing the possibility of machine failure due to poorly mapped steep slopes not detected e.g. by ship-based bathymetry (Peukert et al., 2018b). Mn-nodule distribution maps with this resolution increase the mining efficiency because local deposit variations can significantly affect the performance of the pick-up rate, which is likely determined by technical parameters of the mining vehicle as well as the size, burial depth and abundance of Mn-nodules in the seafloor (Chung, 1996). The exclusion of areas with slopes > 3° resulted in 8 km^2 mineable seafloor surface. Assuming a constant 80% collection efficiency (Volkmann & Lehnen, 2018) and a 30% reduction of the Mn-nodule weight by removal of water (Das & Anand, 2017), the dry mass of Mn-nodules that can be extracted from the surface and the first 15 cm of the sediment column amounts to ca. 190,000 t. In a back-of-the-envelope calculation this quantity, assuming constant metal content inside the study area, equal to the average metal concentrations inside the CCZ (Table 7) (Volkmann, 2015), and 90% metal recovery efficiency; could result in an estimated resource haul of 45,450 t Mn, 2,232 t Ni, 1891 t Cu, 374 t Co, and 102 t Mo (Table 7).

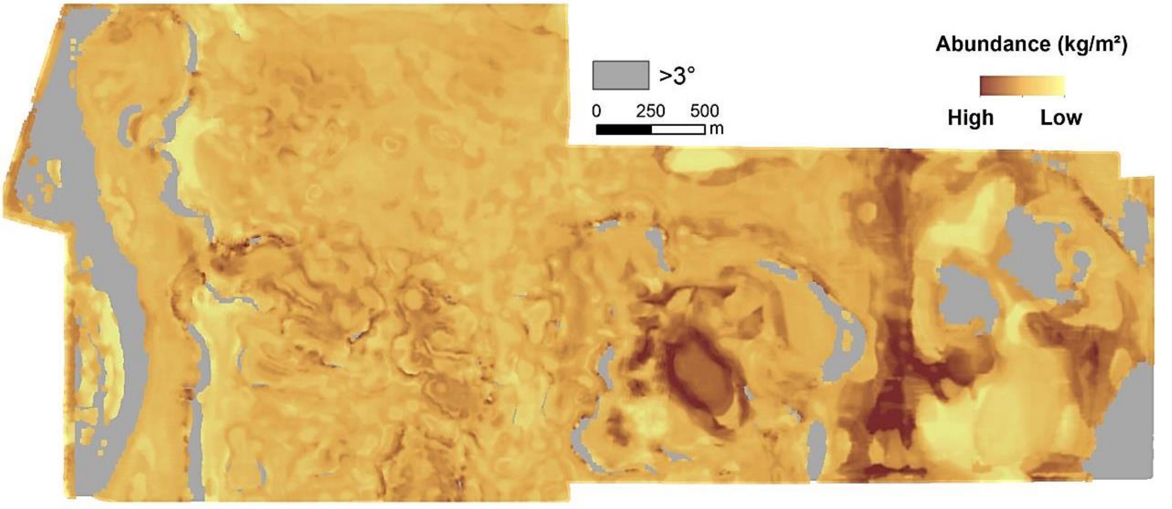

Figure 16: The total abundance of Mn-nodules from the surface and embedded in the sediment (max. 15 cm), in areas with slope ≤ 3° inside block G77 (continuous values of abundance are not given due to confidentiality).

Table 7. The estimated amount of metal mass for 5 metals, based on the average values of metal content inside CCZ and a 5metal HCl-leach recovery method (Volkmann, 2015).

Total Wet Mass (t):	270,400				
Total Dry Mass (t):	189,280				
Metal Content	Mn	Ni	Cu	Co	Mo
wt%:	26.68	1.31	1.11	0.22	0.06
Equal to (t):	50,500	2,480	2,101	416	113
90% metal recovery (t)	45,450	2,232	1,891	374	102

5. Discussion

We present a case study that highlights the applicability of the combination of AUV bathymetric and optical data for Mn-nodules resource modelling using RF machine learning. The use of AUVs for collecting hydroacoustic and optical data in areas of scientific and commercial interest can provide more precise bathymetric and Mn-nodules distribution maps. Regarding the bathymetric maps, the accurate and detailed reconstruction of the seafloor bathymetry in meter-scale resolution enables to use bathymetry and its derivatives as source data layers within a high-resolution RF model. These data should have high-quality characteristics, as the presence of acquisition artefacts may affect the robustness of the modelling procedure (Preston, 2009; Herkül et al., 2017). The combined use of cameras as the DeepSurveyCamera (Kwansnitschka et al., 2016) for acquiring high-resolution photographs, and an automated analysis with a state-of-the-art algorithm (Schoening et al., 2017a) provides essential quantitative information about the distribution of Mn-nodules. Image analysis results are more robust for constant AUV altitudes (7-9m) above flat areas (<3°), while the alternation of the flying altitude and camera orientation during the ascending & descending phases limit the quality of the obtained images and can affect the derived number of Mn-nodules/m².

Inside block G77, the number of Mn-nodules/m^2 seems to follow a normal distribution without extreme outliers and without being linearly correlated with the used predictor variables. Spatially, a clumped autocorrelated pattern is demonstrated, mainly with clustered areas of H-H and L-L values. It is still unclear if this heterogeneity is caused by external processes (e.g. topographic characteristics, geochemical conditions, availability of nucleus material etc.) or it is resulted from the interaction of neighboring Mn-nodules. The areas with higher number of Mn-nodules could provide more fragments as potential nucleus material. However, the less available space in these areas may make more difficult the individual Mn-nodule growth, resulting in smaller median sizes. Conversely, a recent study from Kuhn and Rathke (2017) showed that the blanketing of the Mn-nodules by sediments is higher for larger Mn-nodules and, as a result, fewer large nodules will be counted; resulting in biased results in areas, where the Mn-nodules are bigger. Probably, all of these effects can happen at the same time (with different degrees of influence) promoting a given, scale-dependent spatial structure.

Although the exact reasons for the patchy distribution are not fully understood, the knowledge of the distribution pattern is essential for planning box corer sampling. A random spatial sampling reduces the possibility of dependence among observations in a homogeneously distributed population (Cochran 1977), but it is not appropriate for clustered populations as it cannot eliminate autocorrelation between neighboring sample locations that are inside the spatial influence of the underlying phenomenon (Legendre and Legendre, 1998). In other words, if two or more box-corers are obtained from the same patch, the results will still not be representative for the entire study area. Thus, the deployment of box corers should be executed only after the acquisition, processing and spatial analysis of bathymetric and optical data. The number of box corers should be the maximum feasible, with at least one within each main patch and preferable in locations with available photos in order to calculate better the factor between counted Mn-nodules in photos and in box-corers. A smaller number of box-corers should be deployed and in areas without photos in order to estimate the accuracy of the RF model in areas far away from the optical data but still inside the study area. In other words, the optical data acquisition should be guided by the bathymetric and backscatter seafloor characteristics, and be followed by box core sampling that targets all defined 'seafloor-classes' by considering direct correlations with the previously gathered optical and hydroacoustic data.

This study did not consider geochemical properties of the sediments as input data in the modelling process, which might give additional clues as to why Mn-nodules are distributed as they are. However, RF importance and partial dependence plots show that bathymetric and topographic factors tend to affect this distribution in a non-linear way and with the bulk of data plotting in specific ranges of the bathymetric derivatives. Classic studies have shown that the bathymetry and the variation of the topographic characteristics of the seafloor affects the sediment deposition environment, bottom currents and thus also geochemical processes in the sediment. All these factors determine Mn-nodule growth and thus affect the distribution of Mn-nodules on regional scales (e.g. Craig, 1979; Sharma and Kodagali, 1993). It is still unknown how these properties influence the Mn-nodule distribution on meter to tens of meters scales as seen in our AUV data. The non-linear relationship between Mn-nodules and bathymetry on such high resolution scales only started investigating very recently (e.g. Peukert et al, 2018, Alevizos et al, 2018). To elaborate more on the hydrodynamic and geochemical reasons behind the observed distribution pattern, we would need more investigations at and in the sediment on the same scale.

It should be acknowledged that the aim of any ML predictive model is to derive accurate predictions based on an existing (large) number of measurements, to capture a complex underlying relationship (e.g. non-linear and multi-variate) between different types of data, for which our theoretical knowledge or conceptual understanding is still under development (Schmueli, 2010; Lary et al 2016). Especially due to the constantly increasing size of scientific multivariate data in marine sciences, and the existence of such non-linear relationships between predictor and response variables (e.g. Zhi et al., 2014; Li et al., 2017), ML and RF are considered important analytic tools that can objectively reveal patterns of a (unknown) phenomenon (Genuer et al, 2017; Kavenski et al, 2009; Lary et al 2016). Such predictions may be used to derive causalities or may drive the creation of new hypotheses. In other words, for a predictive model, the 'unguided' data analyses come first and the interpretation follows (Breiman, 2001b; Schmueli, 2010; Obermeyer and Emanuel, 2016). This 'a priori' knowledge of the distribution of the Mn-nodules number and size in such scale can contribute to the biological data survey planning, too. Recent studies showed that the abundance and species richness of nodule fauna inside the CCZ is affected by the abundance of Mn-nodules (Amon et al., 2016; Vanreusel et al., 2016) as well as their size (Veillette et al., 2007). Thus, high priority areas (e.g. these with highest commercial interest) can be targeted for sampling based on the results of optic data and RF modelling. The RF modelling takes advantage of the multi-layer information (here: hydroacoustic and optical data) and handling effectively their complex relationships while being resistant to overfitting (Breiman, 2001a). Moreover, the randomization of the input training points in each tree in each run, resulting in a completely different training dataset each time with mixed points from the entire study area. This random selection and mixing of points, is appropriate for clustered data, as it ignores their spatial locations and consequently limits the influence of spatial autocorrelation (Appendix B). Along these lines, several authors have included the values of latitude/longitude and even the LMI values as predictor variables in order to increase the model performance (e.g. Li, 2013; Li et al., 2011b; Li et al., 2013). RF has a high operational character due to its relatively simple calibration, which does not request extensive data preparation/transformation or need for geostatistical assumptions (e.g. stationarity). The selection of the MGET toolbox (Roberts et al., 2010) increased further the simplicity of the workflow, as the RF modelling was performed entirely inside a graphic environment familiar to many geoscientists. As RF model runs can be implemented inside various software packages in future implementations of this workflow, it would be interesting to include the uncertainty for the associated predictions e.g. with the use of the Quantile Regression Forests (Meinshausen, 2006) from the *quantregForest R package* (Meinshausen, 2012). However, this will increase the computational time (Tung et al, 2014) and the simplicity of the procedure, especially if used other recently proposed methodologies of estimating the uncertainty: the Jackknife method (Wager et al, 2014), the Monte Carlo approach (Coulston et al, 2016) and U-statistics approach (Mentch and Hooker, 2016).

Similarly to other studies (e.g. Cutler et al., 2007; Millard and Richardson, 2015), RF showed increased stability in its performance, allowing a small number of iterations to compute sufficient results. The examination of the main two tuning parameters (*ntree* and *mtry*) showed that the model performance can be increased compared to default values. However, the largest improvement results from using more training data. In this respect, more photos would potentially improve the RF performance as no clear threshold was observed. Although the number of 11,276 photos seems to represent a large data set, the heterogeneity of the distribution and the occurrence of spatial clusters (patches) in different sizes and the inherent need of RF and ML in general for big training datasets (van der Ploeg et al, 2014; Obermeyer and Emanuel, 2016), stresses the need for collecting more and well distributed data. The influence of the number of training data for model performance still remains a discussion point between studies showing an improvement by adding more data (e.g. Bishop, 2006), and other studies presenting stable performance of the model even if more data are added (e.g. Zhu et al, 2012). The availability of more data and especially if they are better distributed (i.e. data that will include the entire range of the number of Mn-nodules/m^2 and they are come from all the different sub-terrains), would most likely reinforce the model to build better and wider relationships between the predictor and response variables; keeping also a larger number of validation data points.

Finally, the resource assessment showed that block G77 is a potential mining area with high average Mn-nodules density and gentle slopes. While here the threshold of 3° (UNOET, 1987) was used, newer plans for mining machines seem to enable operations on steeper slopes (Atmanand and Ramadass, 2017) increasing the total amount of collected Mn-nodules within the herein considered area.

6. Conclusions

The results of this study show that the acquisition and analysis of optical seafloor data can provide quantitative information on the distribution of Mn-nodules. This information can be combined with AUV-based MBES data using RF machine learning to compute predictions of Mn-nodule occurrence on small operational scales. Linking such spatial predictions with sampling based physical Mn-nodule data provides an efficient and effective tool for mapping Mn-nodule abundance.

Competing interests: The authors declare that they have no conflict of interest.

Special issue statement: This article is part of the special issue “Assessing environmental impacts of deep-sea mining – revisiting decade-old benthic disturbances in Pacific nodule areas”. It is not associated with a conference.

Acknowledgments: We thank the captain and crew of RV SONNE for their contribution to a successful cruise. We express our gratefulness to the GEOMAR AUV team for their support during the cruise. We thank Anja Steinführer for pre-processing of the AUV MBES data and Mareike Kampmeier for advice during the post-processing analysis of the MBES data. We thank Inken Preuss for proof-reading the manuscript. Finally, we thank the GEOMAR Library team for its support in gathering the necessary bibliography. All data were acquired within the framework of the JPIO Project "Ecological Aspects of Deep-Sea Mining", funded through BMBF-Grant 03F0707A. This is publication 35 of the DeepSea Monitoring Group at GEOMAR Helmholtz Centre for Ocean Research Kiel.

Edited by:

Reviewed by:

Appendix A: Methodology

3.1 Hydroacoustic Data Acquisition & Post Processing

The calculation of the bathymetric derivatives was performed with the SAGA GIS v6.3.0 Morphometry library (http://www.saga-gis.org/saga_tool_doc/6.3.0/ta_morphometry.html).

3.4 Spatial Statistics

Global Moran's I and Local Moran's I were performed with the ArcMap™ 10.1 software, using the Spatial Statistic Toolbox, according to its provided equations. As a null hypothesis, it is assumed that the examined attribute is randomly distributed among the features in the study area. For the optimal conceptualization of spatial relationships, the Inverse Euclidian Distance Method was selected, as it is appropriate for modelling processes with continuous data in which the closer two samples are in space, the more likely they are to interact/influence each other or have been influenced from the same reasons. The distance threshold was set at 50m and the increment analysis was performed with a step of 50m. Moreover, the spatial weights were standardized in order to minimize any bias that exists due to sampling design (uneven number of neighbors). Apart from the index value, the p-value and z-score are also provided. The Local Moran's I indicates statistically significant clusters and outliers for a 95% confidence level. The high number of observations (>>30) that was used ensures the robustness of the indexes.

3.5 RF Predictive modelling (Selection of Predictor Variables)

Correlation among the derivatives was checked by Spearman's correlation coefficient (ρ). This coefficient was preferred due to the skewed distribution of the values in the derivatives. The majority of the possible pairs is uncorrelated or weakly correlated. Only C vs. TPI_F and TRI vs. S have a strong correlation. However, they should not be excluded as they express different topographic characteristics and they are not correlated with the remainder of derivatives. . It should be mentioned that in similar studies even higher thresholds have been used during the selection of predictor variables (Che Hasan et al, 2014; Li et al, 2016; Li et al, 2017)).

Table A1. Spearman's correlation coefficient for each pair of predictor variables

	D	**S**	**Pl.C**	**Pr.C**	**TPI_B**	**TPI_M**	**TPI_F**	**C**	**TRI**
D									
S	-0.07								
Pl.C	0.06	-0.02							
Pr.C	0.08	-0.01	0.37						
TPI_B	0.76	-0.09	0.13	0.16					
TPI_M	0.36	-0.06	0.20	0.27	0.72				
TPI_F	0.23	-0.05	0.33	0.41	0.47	0.77			
C	-0.30	0.05	-0.25	-0.34	-0.54	-0.79	-0.90		
TRI	-0.10	0.91	-0.02	-0.03	-0.12	-0.06	0.04	0.05	

The 9 training samples with different size were created by the MGET tool: Randomly Split Table into training and testing records. The spatial randomness of the procedure, combined with the many available data resulted in training samples with similar descriptive statistics.

Table A2. Descriptive Statistics of different training samples

% Training Sample:	**10%**	**20%**	**30%**	**40%**	**50%**	**60%**	**70%**	**80%**	**90%**
Training set size	1127	2255	3383	4511	5638	6766	7894	9021	10148
Mean	36.5	36.3	36.6	36.6	36.6	36.7	36.6	36.7	36.6
Std. Error	0.3	0.2	0.2	0.1	0.1	0.1	0.1	0.1	0.1
Std. Deviation	9.3	9.2	9.4	9.2	9.2	9.3	9.3	9.2	9.3
Minimum	7	13	12	13	12	14	7	7	7
Maximum	63	70	72	66	78	78	78	72	78

Appendix B: Results

3.5 RF Predictive modelling (Calibration of the model):

The descriptive statistics of the performance of each model were used as decision factors for the number of iterations. In all cases, the mean value with very low standard error, the very low standard deviation, range and the 95% confidence interval indicate a rather stable performance, without the need for further iterations.

Table B2. Descriptive statistics of MSR from different training set sizes, after 10 iterations with default settings.

% Training Sample:	10%	20%	30%	40%	50%	60%	70%	80%	90%
Mean	34.8	30.2	26.1	23.3	22.2	21.3	19.7	18.3	18.1
Std. Error	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Median	34.8	30.3	26.1	23.2	22.2	21.3	19.7	18.3	18.1
Mode	34.7	30.3	26.1	23.2	22.2	21.3	19.7	18.3	18.1
Std. Deviation	0.2	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.1
Minimum	34.5	30.1	25.9	23.2	22.1	21.2	19.6	18.2	18.1
Maximum	35.1	30.4	26.3	23.5	22.3	21.3	19.7	18.3	18.1
C.I. (95.0%)	0.1	0.1	0.1	0.1	0.1	0.0	0.0.	0.0	0.0

Table B3. Descriptive Statistics of MSR from a different number of *ntree* parameter, after 10 iterations with 80% of the sample as training data and *mtry* = 3.

ntree:	100	200	300	400	500	600	700	800	900	1000
Mean	18.8	18.4	18.3	18.3	18.3	18.2	18.2	18.2	18.2	18.2
Std. Error	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Median	18.8	18.4	18.3	18.3	18.3	18.2	18.2	18.2	18.2	18.2
Mode	18.8	18.4	18.3	18.3	18.3	18.2	18.2	18.2	18.2	18.2
Std. Deviation	0.1	0.1	0.1	0.1	0.0	0.1	0.1	0.1	0.0	0.0
Minimum	18.5	18.4	18.2	18.2	18.2	18.1	18.1	18.1	18.1	18.1
Maximum	18.9	18.5	18.5	18.4	18.3	18.3	18.3	18.3	18.2	18.2
C.I. (95.0%)	0.1	0.0	0.1	0.0	0.0	0.0	0.0	0.0	0.0	0.0

Table B4. Descriptive Statistics of MSR from different number of *mtry* parameter, after 10 iterations with 80% of the sample as training data and *ntree* = 600.

mtry:	1	2	3	4	5	6	7
Mean	23.4	19.3	18.2	17.9	17.7	17.6	17.6
Std. Error	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Median	23.4	19.3	18.2	17.9	17.7	17.6	17.6
Mode	23.4	19.3	18.2	17.9	17.7	17.6	17.6

Std. Deviation	0.0	0.1	0.1	0.1	0.0	0.0	0.0
Minimum	23.3	19.1	18.1	17.8	17.6	17.5	17.6
Maximum	23.5	19.4	18.3	17.9	17.7	17.7	17.7
C.I. (95.0%)	0.0	0.1	0.0	0.0	0.0	0.0	0.0

Table B5. Descriptive Statistics of MSR for the optimum selected RF model, after 30 iterations with 80% of the sample as training data, *ntree* = 600, and *mtry* = 6.

	Mean	Std. Error	Median	Mode	Std. Deviation	Minimum	Maximum	C.I. (95%)
Optimum RF	17.6	0.0	17.6	17.6	0.0	17.5	17.7	0.0

Table B6. Descriptive statistics of validation measures for the optimum RF model, after 30 iterations with 80% of the sample as training data, *ntree* = 600, and *mtry* = 6 .

RF Importance:	**Depth**	**TPI_B**	**TPI_M**	**TRI**	**TPI_F**	**C**	**S**	**Pl.C**	**Pr.C**
Mean	80.1	63.6	46.7	36.1	24.5	19.7	12.0	2.6	2.4
Std. Error	0.1	0.1	0.1	0.0	0.0	0.0	0.0	0.0	0.0
Median	80.1	63.5	46.7	36.1	19.7	19.7	11.9	2.6	2.4
Mode	80.1	63.3	46.9	36.1	19.8	19.8	11.9	2.6	2.4
Std. Deviation	0.4	0.6	0.6	0.2	0.2	0.2	0.2	0.0	0.0
Minimum	79.1	62.6	45.0	35.7	19.2	19.2	11.7	2.5	2.3
Maximum	80.8	64.9	47.7	36.4	20.1	20.1	12.4	2.6	2.5
C.I. (95.0%)	0.1	0.2	0.2	0.1	0.1	0.1	0.1	0.0	0.0

4.1 Data Exploration:

The histogram of Mn-nodules/m^2 (Figure B1) shows a good fit with the superimposed theoretical normal curve, with the shape of the distribution being rather symmetrical. This fact is supported by the equal mean and median and the slightly different mode (Table B6). Similarly, the visual inspection of the probability plot (Figure B1) shows a good match as a linear pattern is observed for the greatest part, with slight deviation existing only in the outer parts of the curve. According to the boxplot, there are only 21 mild outliers (according to Hoaglin et al., 1986; Dawson, 2011), which correspond to 0.18% of the total observation. This percentage is smaller than the 0.8% threshold that has been suggested for normal disturbed data (Dawson, 2011).The small values for skewness and kurtosis combined with the large sample size further support the normal distributed pattern of the data (Table B6). Especially for large data samples, measurements of skewness and kurtosis combined with the visual inspection of histogram and probability plot are recommended ways of examining normality of data (D' Agostino et al., 1990; Yaziki and Yolacan, 2007; Field, 2009; Ghasemi and Zahediasl, 2012; Kim, 2013).

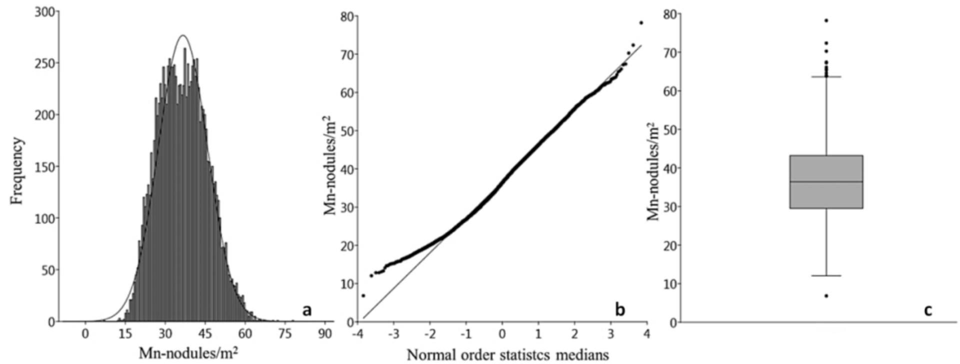

Figure B1 a) Histogram of Mn-nodules/m^2 with the superimposed normal curve. b) The normal probability plot of Mn-nodules/m^2. c) The box plot of Mn-nodules/m^2.

Table B6. The descriptive statistics of the number of Mn-nodules/m².

	Mean	5% Trim. Mean	Median	Mode	SD	Min.	Max.	Skew.	Kurtosis
Mn-nodules/m²	36.6	36.4	36.4	39	9.2	6.8	78.2	0.1	-0.4

A potential linear correlation between depth, bathymetric derivatives, and number of Mn-nodules/m^2 was investigated using the Spearman's rank correlation coefficient (ρ) because of the skewed distribution and presence of extreme values in the depth and bathymetric derivative values (Mukaka, 2012).

Table B7. The Spearman's rank correlation coefficient between Mn-nodules/m^{2,} depth, and bathymetric derivatives

Depth	Slope	TRI	Pl.C	Pr.C	TPI_B	TPI_M	TPI_F	Con.
0.38	0.08	0.07	0.03	0.04	0.29	0.24	0.05	-0.14

4.3.2 Selection, application and external validation of the optimal model

Despite the fact that RF is a full non-parametric technique and there is no need for the residuals to follow specific assumptions (Breiman, 2001a), the examination of them can provide an in-depth look into its performance characteristics. The scatterplot of residuals against predicted values shows a random pattern, which is also confirmed by the low values of Pearson, Spearman, and R² coefficients between predicted values and residuals (Figure B1a and Table B6). Moreover, the residuals tend to cluster towards the middle of the plot without being systematically high or low, and having zero mean value (Figure B1 and Table B7). Their constant variance (homoscedasticity) implies that the distribution of error has the same range for almost all fitted values. Indeed, 99.3% of the residuals are inside the range ±15 and mainly the 81.2% inside the range ±5 (Table B8). The presence of outliers is very limited without affecting the main statistical characteristics of residuals (Table B7) indicating that the model adequately fits the overwhelming majority of the observations (>2165) and only random variation (that exists in any real natural phenomenon) or noise can occur.

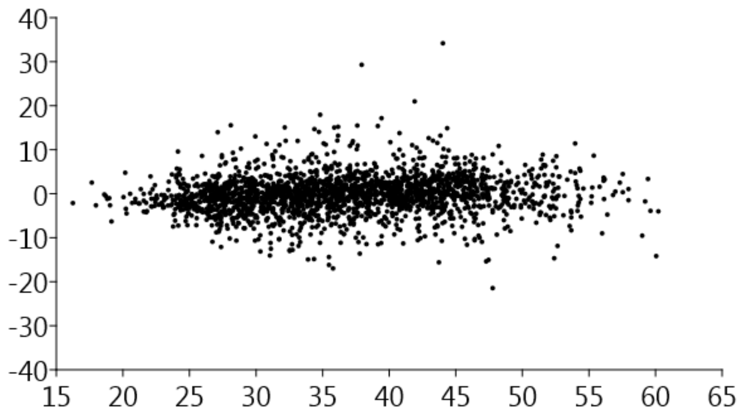

Figure B2. Scatterplot between residuals and predicted values.

Table B8. Pearson, Spearman, and R^2 correlation coefficients between residuals and predicted values.

	Pearson	Spearman	R^2
Correlation of residuals and predicted values	0.1	0.2	0.0

Table B9. Main descriptive statistics of residuals and 5% trimmed residuals

	Mean	Std. Error	Median	Mode	Std. Deviation
Residuals	-0.2	0.1	-0.2	0.6	4.4
5% Trimmed Residuals	-0.2	0.1	-0.2	0.6	2.9

Table B10. Residuals range

Residuals Range	±20	±15	±10	±5
% of Residuals	99.8	99.3	96.1	81.2

The spatial autocorrelation analysis of the residuals using the Global Moran’s Index (same settings as Appendix A), showed low spatial autocorrelation (I=0.112112 p<0.01 and Z-score>2.58). The index number of the residuals is relatively low compared with the high initial values of the original data (I=0.69890 and I=0.697747 for the entire dataset and the 80% training dataset, respectively). The 5% trimmed residuals (see Appendix B-Table B8) showed that their spatial autocorrelation is only 0.093832. According to similar studies (i.e. regression RF), the presence of spatial autocorrelation in the residuals of the model can result in underestimation of the true prediction error (Ruß und Kruse, 2010). The presence of low spatial autocorrelation values in the residuals of regression RF has been reported also by other authors (e.g. Mascaro et al, 2014; Xu et al, 2016); and it is a common problem in all the well-established machine learning methods (e.g. RandomForests, Neural Network, Gradient Boosting Machine, and Support Vector Machines) when dealing with regression predictions of spatial variables (Gilardi and Bengio, 2009; Ruß und Kruse, 2010; Santibanez et al, 2015a,b). The spatial plotting and visual examination of the residuals (Figure B3) showed that this spatial clustering exists mainly in the small sub-area b, and especially in the areas which are associated with an increased slope (>3°), where the AUV is forced to vary its altitude between the ascending and descending phase and consequently affects the image quality and the later modelling results.

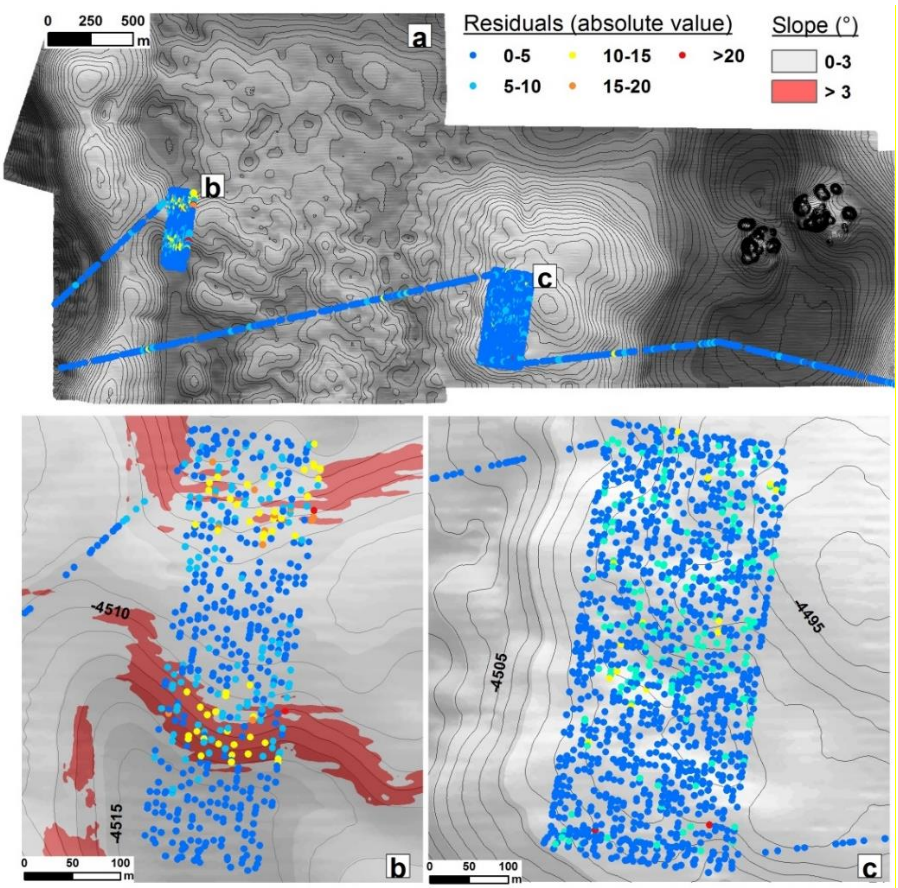

Figure B3. Spatial plotting of the RF residuals (absolute values). The intervals of their range are in accordance with the Table B9.

4.3.4 RF importance

The production of the RF partial depedence plots, show the non-linear character between the Mn-nodules/m² and the bathymetric derivatives.

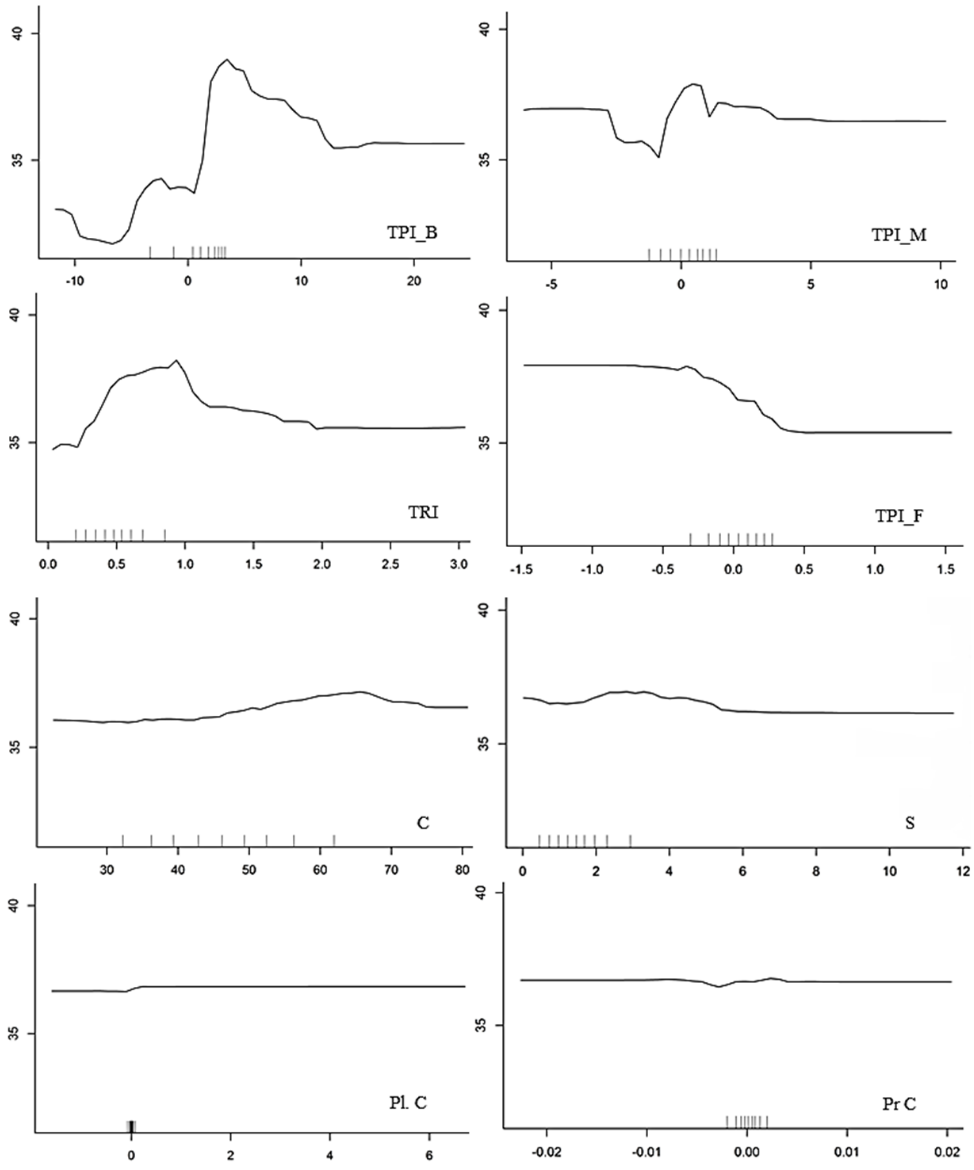

Figure B2. Partial dependence plots for each of the predictor variables. The y axis represent the number of Mn-nodules/m² and the x axis the values of each predictor variable (depth derivatives). The ticks inside the graphs indicate the deciles of the data.

References

1. Alevizos et al, Schoening T., Koeser K., Snellen M. and Greinert J.: Quantification of the fine-scale distribution 1 of Mn-nodules: insights from AUV multi-beam and optical imagery data fusion. Biogeosciences Discussions . pp. 1-29, 2018. https://doi.org/10.5194/bg-2018-60
2. Anselin, L.: Local Indicators of Spatial Association—LISA. Geographical Analysis 27(2): 93–115, 1995. https://doi.org/10.1111/j.1538-4632.1995.tb00338.x
3. Atmanand M.A. and Ramadass G.A.: Concepts of Deep-Sea Mining Technologies. In: Sharma R. (eds): Deep-Sea Mining. Resource Springer, Cham. 2017. Online ISBN 978-3-319-52557-0. https://doi.org/10.1007/978-3-319-52557-0_6
4. Bellingham, J.: Autonomous underwater vehicles (AUVs): In Encyclopedia of Ocean Sciences, ed.-in chief John H. Steele, Academic Press, San Diego, pp. 212-216. 2001. https://doi.org/10.1006/rwos.2001.0303
5. Bingham, D., Drake, T., Hill, A., Lott, R.: The Application of Autonomous Underwater Vehicle (AUV) Technology in the Oil Industry – Vision and Experiences. FIG XXII International Congress Washington, D.C. USA, April 19-26 1-13, 2002.
6. Breiman, L.: Random forests. Machine Learning, 45, 5–32, 2001a. https://doi.org/10.1023/A:101093340
7. Breiman, L.: Statistical Modeling: The Two Cultures. Statistical Science, 16, 3, 199-215, 2001b. https://doi.org/10.1214/ss/1009213726
8. Caress, D.W., Thomas, H., Kirkwood, W.J., McEwen, R., Henthorn, R., Clague, D.A., Paull, C.K., and Paduan, J.. High-resolution multibeam, sidescan and subbottom surveys using the MBARI AUV D. Allan B. In: Greene, H.G., In: Reynolds, J.R. (Eds.), Marine Habitat Mapping Technology for Alaska. Alaska Sea Grant College Program, University of Alaska Fairbanks, pp. 47–69, 2008. https://doi.org/10.4027/mhmta.2008.04
9. Caress, D. W., and Chayes, D. N.: MB-System: Mapping the Seafloor, 2017. https://www.mbari.org/products/research-software/mb-system
10. Carranza, E. J. M. and Laborte, A. G.: Random Forest Predictive Modelling of Mineral Prospectivity with Small Number of Prospects and Data with Missing Values in Abra (Philippines). Computers and Geosciences, 74, 60-70, 2015a. http://dx.doi.org/10.1016/j.cageo.2014.10.004
11. Carranza, E. J. M. and Laborte, A. G.: Data-driven predictive mapping of gold prospectivity, Baguio district, Philippines: Application of Random Forests algorithm. Ore Geology Reviews, 7, 777–787, 2015b. http://dx.doi.org/10.1016/j.oregeorev.2014.08.010
12. Carranza, E. J. M. and Laborte, A. G.: Data-Driven Predictive Modeling of Mineral Prospectivity Using Random Forests: A Case Study in Catanduanes Island (Philippines). Natural Resources Research, 25, 35-50, 2016. https://doi.org/10.1007/s11053-015-9268-x

13. Chance, T., Kleiner, Art and Northcutt, J.: The autonomous underwater vehicle (AUV): A cost-effective alternative to deep-towed technology. Integrated Coastal Zone Management, 2, 65-69, 2000.
14. Chakraborty, B. and Kodagali, V.: Characterizing Indian Ocean manganese nodule-bearing seafloor using multi-beam angular backscatter. Geo-Mar Lett, 24, 8–13, 2004. https://doi.org/10.1007/s00367-003-0153-y
15. Che Hasan, et al R., Ierodiakonou, D. and Monk, J.: Evaluation of Four Supervised Learning Methods for BenthicHabitat Mapping Using Backscatter from Multi-Beam Sonar. Remote Sensing 4, 3427-3443, 2012. https://doi.org/10.3390/rs4113427
16. Che Hasan, R., Ierodiaconou, D., Laurenson, L., Schimel, A.: Integrating Multibeam Backscatter Angular Response, Mosaic and Bathymetry Data for Benthic Habitat Mapping. PLoS ONE 9 (5), e97339, 2014. https://doi.org/10.1371/journal.pone.0097339
17. Chen, L., Wang, S., McDonald-Maier K., Hu, H.: Towards autonomous localization and mapping of AUVs: a survey. International Journal of Intelligent Unmanned Systems, Vol. 1 Issue: 2, 97-120, 2013. https://doi.org/10.1108/20496421311330047
18. Chung, Jin S.: Deep-Ocean Mining: Technologies for Manganese Nodules and Crusts. International Journal of Offshore and Polar Engineering, Vol. 6, No.4, 1996.
19. Clague, D. A., B. M. Dreyer, J. B. Paduan, J. F. Martin, D. W. Caress, J. B. Gill, D. S. Kelley, H. Thomas, R. A. Portner, J. R. Delaney, T. P. Guilderson, and M. L. McGann: Eruptive and tectonic history of the Endeavour Segment, Juan de Fuca Ridge, based on AUV mapping data and lava flow ages, Geochem. Geophys. Geosyst., 15, 3364–3391, 2014. https://doi.org/10.1002/2014GC005415
20. Clague, D. A., Caress, D. W., Dreyer, B. M., Lundsten, L., Paduan, J. B., Portner, R. A., Spelz-Madero, R., Bowles, J.A., Castillo P.R., Guardado-France, R., Le Saout, M., Martin, J.F., Santa Rosa-del Rio, M. A., and A. Zierenberg, R.: Geology of the Alarcon Rise, southern Gulf of California. Geochemistry, Geophysics, Geosystems, 19, 807–837, 2018. https://doi.org/10.1002/2017GC007348
21. Clements, A. J., Strong, J. A., Flanagan, C., and Service, M.: Objective stratification and sampling-effort allocation of ground-truthing in benthic-mapping surveys. ICES Journal of Marine Science, 67, 628–637, 2010.
22. Cochran, W. G.: Sampling Techniques, 3rd edn. Wiley, New York, 1977.
23. Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., and Boehner, J.: A system for Automated Geoscientific Analyses (SAGA) v. 2.1.4. Geoscientific Model Development 8, 1991-2007. https://doi.org/10.5194/gmd-8-1991-2015
24. Coulston, J. W., Blinn, C. E., Thomas, V. A., and Wynne, R. H.: Approximating prediction uncertainty for random forest regression models. Photogrammetric Engineering & Remote Sensing, 807 82(3):189 – 197, 2016.
25. Craig, J. D.: The relationship between bathymetry and ferromanganese deposits in the north equatorial Pacific, Marine Geology, 29, 165–186, 1979. https://doi.org/10.1016/0025-3227(79)90107-5

26. Cutler, D.R., Edwards, T.C., Beard Karen, H., Cutler, A., Hess, K.T., Gibson, J.C., Lawler, J.J.: Random forests for classification in ecology. Ecology, 88(11), pp. 2783–2792, 2007.
27. D' Agostino, R.B., Belanger, A. and D' Agostino, R.B. Jr.: A Suggestion for Using Powerful and Informative Tests of Normality. The American Statistician 44, no. 4, 316-21, 1990. https://doi.org/10.2307/2684359
28. Danson, E.: The Economies of Scale: Using Autonomous Underwater Vehicles (AUVs) for Wide-Area Hydrographic Survey and Ocean Data Acquisition, FIG XXII International Congress Washington, D.C. USA, April 19-26, 2002.
29. Das, R.P., and Anand, S.: Metallurgical Processing of Polymetallic Ocean Nodules. In Sharma R. (eds): Deep-Sea Mining. Resource Springer, Cham. 2017 0. https://doi.org/10.1007/978-3-319-52557-0_12
30. Dawson, R.: How Significant is a Boxplot Outlier? Journal of Statistics Education, 19:2, 2011. https://doi.org/10.1080/10691898.2011.11889610
31. De Moustier, C.: Beyond bathymetry: Mapping acoustic backscattering from the deep seafloor with Sea Beam. Journal of the Acoustical Society of America 79 (2), 316-331, 1986.
32. Deschamps, A., Maurice, T., Embley, R. W., Chadwick, W.W.: Quantitative study of the deformation at Southern Explorer Ridge using high-resolution bathymetric data. Earth and Planetary Science Letters 259, 1–17, 2007. https://doi.org/10.1016/j.epsl.2007.04.007
33. Diaz-Uriarte, R. and de Andres, A.: Gene selection and classification of microarray data using random forest. BMC Bioinformatics, 7:3, 2006.
34. Diesing, M., Green, S. L., Stephens, D., Lark, R.M., Stewart, H.A., Dove D.: Mapping seabed sediments: Comparison of manual, geostatistical, object-based image analysis and machine learning approaches. Continental Shelf Research, 84, 107-119, 2014. http://dx.doi.org/10.1016/j.csr.2014.05.004
35. Diesing, M. and Stephens, D.: A multi-model ensemble approach to seabed mapping. Journal of Sea Research, 100, 62-69, 2015. http://dx.doi.org/10.1016/j.seares.2014.10.013
36. Dikau, R.: Geomorphic landform modelling based on hierarchy theory. In: Brassel, K., Kishimoto, H. (Eds.), Proceedings of the 4th International Symposium on Spatial Data Handling. Department of Geography, University of Zürich, Zürich, Switzerland, 230–239, 1990.
37. Durden, JM, Schoening, T., Althaus, F., Friedman, A. Garcia, R., Glover, A.G., Greinert, J., Jacobsen, Stout N., Jones, DOB, Jordt, A., Kaeli, JW., Koser, K., Kuhnz, LA., Lindsay, D., Morris, KJ., Nattkemper, TW., Osterloff, J., Ruhl, HA., Singh, H., Tran, M., Bett, BJ.: Perspectives in visual imaging for marine biology and ecology: from acquisition to understanding. Oceanography and Marine Biology: An Annual Review, 54. pp. 1-72, 2016. https://doi.org/10.1201/9781315368597
38. Frazer, J. and Fisk, M. B.: Geological factors related to characteristics of sea-floor manganese nodule deposits, Deep Sea Research Part A. Oceanographic Research Papers, 28, 1533–1551, 1981. https://doi.org/10.1016/0198-0149(81)90096-0

39. Field, A. P.: Discovering statistics using SPSS: (and sex and drugs and rock 'n' roll). (OKS Print.) Los Angeles [i.e. Thousand Oaks, Calif.: SAGE Publications.

40. Fu, W. J., Jiang, P. K., Zhou, G. M. and Zhao, K. L.: Using Moran’s I and GIS to study the spatial pattern of forest litter carbon density in a subtropical region of southeastern China. Biogeosciences, 11, 2401–2409, 2014. https://doi.org/10.5194/bg-11-2401-2014

41. Garzón, M.B., Blazek, R., Neteler, M., Sánchez de Dios, R., Ollero, H.S. and Furlanello, C.: Predicting habitat suitability with machine learning models: The potential area of Pinus sylvestris L. in the Iberian Peninsula. Ecological Modelling, 197, 383-393, 2006. https://doi.org/10.1016/j.ecolmodel.2006.03.015

42. Genuer, R., Poggi J.,Tuleau-Malot, C, Villa-Vialaneix, N.: Random Forests for Big Data. Big DataResearch, 9, 28–46, 2017. http://dx.doi.org/10.1016/j.bdr.2017.07.003

43. GEOMAR Helmholtz-Zentrum für Ozeanforschung. (2016). Autonomous Underwater Vehicle „ABYSS“ Journal of large-scale research facilities, 2, A79. http://dx.doi.org/10.17815/jlsrf-2-149

44. Ghasemi, A. and Zahediasl, S.: Normality Tests for Statistical Analysis: A Guide for Non-Statisticians. Int J Endocrinol Metab, 10 (2), 486-489, 2012. https://doi.org/10.5812/ijem.3505

45. Gilardi, N., and Bengio, S.: Comparison of four machine learning algorithms for spatial data analysis. Conf. Signals Syst. Comput., 17, 160–167, 2009.

46. Glasby, G.P.: Distribution of manganese nodules and lebensspuren in underwater photographs from the Carlsberg Ridge, Indian Ocean, New Zealand. Journal of Geology and Geophysics, 16:1, 1-17, 1973. http://dx.doi.org/10.1080/00288306.1973.10425383

47. Glasby, G. P.: Manganese nodules in the South Pacific: A review. New Zealand Journal of Geology and Geophysics, 19:5 707-736, 1976. https://doi.org/10.1080/00288306.1976.10426315

48. Goodchild, M. F.: Spatial autocorrelation. Concepts and Techniques in Modem Geography, 47, l-56, 1986.

49. Grasmueck, M., Eberli, G. P., Viggiano, D. A., Correa, T., Rathwell, G., and Luo, J. : Autonomous underwater vehicle (AUV) mapping reveals coral mound distribution, morphology, and oceanography in deep water of the Straits of Florida. Geophysical Research. Letters, 33, L23616, 1-6, 2006. https://doi.org/10.1029/2006GL027734

50. Greinert, J.: Swath sonar multibeam EM122 bathymetry during SONNE cruise SO239 with links to raw data files. PANGAEA. 2016. https://doi.pangaea.de/10.1594/PANGAEA.859456

51. Greinert, J., Schoening, T., Köser K. and Rothenbeck M.: Seafloor images and raw context data along AUV tracks during SONNE cruises SO239 and SO242/1. GEOMAR - Helmholtz Centre for Ocean Research Kiel, PANGAEA, 2017.https://doi.org/10.1594/PANGAEA.882349

52. Haase, K.M., Koschinsky, A., Petersen, S., Devey, C.W., German, C., Lackschewitz, K.S., Melchert, B., Seifert, R., Borowski, C., Giere, O., Paulick, H., M64/1, M68/1 and M78/2 Scientific Parties. Diking, young volcanism and diffuse hydrothermal activity on the southern Mid-Atlantic Ridge: the Lilliput field at 9°33′S. Marine Geology 266, 52–64, 2009. https://doi.org/10.1016/j.margeo.2009.07.012

53. Hammer, Ø., Harper, D.A.T., Ryan, P.D.: PAST: Paleontological statistics software package for education and data analysis. Palaeontologia Electronica 4 (1): 9pp, 2001. http://palaeo-electronica.org/2001_1/past/issue1_01.htm
54. Herkül, K., Peterson, A. and Paekivi, S.: Applying multibeam sonar and mathematical modeling for mapping seabed substrate and biota of offshore shallows. Estuarine, Coastal and Shelf Science, 192, 57-71, 2017. http://dx.doi.org/10.1016/j.ecss.2017.04.026
55. Henthorn, R., Caress D.W., Thomas H., McEwen R., Kirkwood W.J., Paull C.K. and Keaten R.: High-resolution multibeam and subbottom surveys of submarine canyons, deep-sea fan channels, and gas seeps using the MBARI mapping AUV. In: Proceedings of the IEEE OCEANS Conference, 1-6, 2006. https://doi.org/10.1109/OCEANS.2006.307104
56. Hoaglin, D. C., Iglewicz, B. and Tukey, J. W.: Performance of Some Resistant Rules for Outlier Labeling. Journal of the American Statistical Association 81, no. 396, 991-999, 1986. http://dx.doi.org/10.2307/2289073
57. Horning, N.: RandomForests : An algorithm for image classification and generation of continuous fields data sets, International Conference on Geoinformatics for Spatial Infrastructure Development in Earth and Allied Sciences (GIS-IDEAS) , 9-11 December, Hanoi,Vietnam, 2010. http://wgrass.media.osaka cu.ac.jp/gisideas10/ viewpaper. php ?id=342
58. Huang, Z., Justy, S., Scott, L.N. Brendan, P.B.: Predictive mapping of seabed substrata using high-resolution multibeam sonar data: A case study from a shelf with complex geomorphology. Marine Geology, 357, 37–52, 2014. http://dx.doi.org/10.1016/j.margeo.2014.07.012
59. Ismail, K., Huvenne, V.A.I ., Masson, D.G.: Objective automated classification technique for marine landscape mapping in submarine canyons. Marine Geology, 362,17–32, 2015. http://dx.doi.org/10.1016/j.margeo.2015.01.006
60. Iwahashi, J. and Pike, R.J.: Automated Classifications of Topography from DEMs by an Unsupervised Nested-Means Algorithm and a Three-Part Geometric Signature. Geomorphology, 86, 409-440, 2017. http://dx.doi.org/10.1016/j.geomorph.2006.09.012
61. Jakobsson, M., Gyllencreutz, R., Mayer, L. Dowdeswell, J. A. Canals, M., Todd, B. J., Dowdeswell, E. K., Hogan K. A. and Larter, R. D.: Mapping submarine glacial landforms using acoustic methods. Geological Society, London, Memoirs, 46, 17-40, 2016, https://doi.org/10.1144/M46.182
62. Jung, H.S., Ko, Y.T., and Moon, J.W.: Characteristics of Seafloor Morphology and Ferromanganese Nodule Occurrence in the Korea Deep-sea Environmental Study (KODES) Area, NE Equatorial Pacific. Marine Georesources and Geotechnolog y, 19, 167-180, 2001. https://doi.org/10.1080/10641190109353811
63. Kavenski, M., Pozdnukhov, A. and Timonin, V.: Machine learning for spatial environmental data. Theory, applications and software. EPFL Press, First edition, 2009. https://doi.org/10.1201/9781439808085
64. Kim, H.Y.: Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restorative Dentistry & Endodontics, 38 (1), 52-54, 2013. https://doi.org/10.5395/rde.2013.38.1.52

65. Kleinrock, M.C: Chapter 2: Capabilities of some systems used to Survey the Deep-Sea Floor. In: Richard G. A. (eds), Handbook of geophysical exploration at sea. Second Edition. Hard Minerals, CRC Press, p37, 1992a.
66. Kleinrock, M.C., Hey, R.N., Theberger, A.E., Jr.: Practical geological comparison of some seafloor survey instruments. Geophysical Research Letters, VOL.19, NO.13, 1407-1410, 1992b. https://doi.org/10.1029/92GL01390
67. Knobloch, A., Kuhn, T., Rühlemann, C., Hertwig, T., Zeissler, K.O., Noack, S.: Predictive Mapping of the Nodule Abundance and Mineral Resource Estimation in the Clarion-Clipperton Zone Using Artificial Neural Networks and Classical Geostatistical Methods. In: Sharma R. (eds): Deep-Sea Mining. Resource Springer, Cham. 2017 https://doi.org/10.1007/978-3-319-52557-0_6
68. Ko, Y., Lee, S., Kim, J., Kim. K.H., and Jung M.S.: Relationship between Mn Nodule Abundance and Other Geological Factors in the Northeastern Pacific: Application of GIS and Probability Method. Ocean Science Journal, Vol. 41 No. 3, 149-161, 2006. https://doi.org/10.1007/BF03022420
69. Kodagali, V. N. and Sudhakar, M.: Manganese nodule distributionin different topographic domains of the Central Indian Basin, Marine Georesources & Geotechnology, 11:4, 293-309, 1993. http://dx.doi.org/10.1080/10641199309379925
70. Kodagali, V.: Influence of Regional and Local Topography on the Distribution of Polymetallic Nodules in Central Indian Ocean Basin, Geo-Marine Letters 8, 173-178, 1988.
71. Kodagali, V. and Chakraborty, B.: Multibeam Echosounder PseudoSidescan Images as a tool for Manganese Nodule Exploration. Proceedings of the Third Ocean Mining Symposium Goa, India, November 8-10, 97-104, 1999.
72. Kuhn, T., Wiedicke – Hombach, M., Barckhausen ,U. Schwarz-Schampera, U., Rutkowski, J. and Lehmnn, S.: New Insights of Mn Nodule Exploration from the German License Area in the Pacific Manganese Nodule Belt. Toward the Sustainable Development of Marine Minerals: Geological, Technological, and Economic Aspects, 39th Underwater Mining Institute, 4-9 October, Gelendzhik, Russia, 2010.
73. Kuhn, T., Rühlemann, C. and Wiedicke-Hombach, M.: Development of Methods and Equipment for the Exploration of Manganese Nodules in the German License Area in the Central Equatorial Pacific, Proceedings of the ninth ISOPE Ocean Mining Symposium, Maui, Hawaii, USA, June 19-24, 174-177, 2011.
74. Kuhn, T., Rühlemann C., Knobloch, A.: Classification of manganese nodule estimates: can we reach the "measured resource" level? Resource and Environmental Assessments for Seafloor Mining Development, 45th Underwater Mining Conference , Incheon, Korea, 9-13 October 2016.
75. Kuhn, T., Rathke, M.: Report on visual data acquisition in the field and interpretation for SMnN. Deliverable D1.31 of the EU-Project Blue Mining. BGR Hannover, 34 pp. 2017. www.bluemining.eu/downloads

76. Kwasnitschka, T., Köser, K., Sticklus, J., Rothenbeck, M., Weiß, T., Wenzlaff, Em. , Schoening, T., Triebe, L., Steinführer, An. , Devey, C. and Greinert, J. 2016 DeepSurveyCam - A Deep Ocean Optical Mapping System. Sensors, 16, 164, 2016. http://dx.doi.org/10.3390/s16020164
77. Lary, D.J., Alavi, A.H., Gandomi, A. H., Walker A.L.: Machine learning in geosciences and remote sensing. Geoscience Frontiers, 7, 3-10, 2016. http://dx.doi.org/10.1016/j.gsf.2015.07.003
78. Leempoel, K., Parisod, C., Geiser, C., Dapra, L., Vittoz, P., and Joost, S.: Very high-resolution digital elevation models: are multi-scale derived variables ecologically relevant? Methods in Ecology and Evolution 6, 1373–1383, 2015. https://doi.org/10.1111/2041-210X.12427
79. Legendre, P., and Legendre, L.: Numerical Ecology. second ed., Elsevier, Amsterdam, p. 853, 1998:
80. Li, J. Predicting the spatial distribution of seabed gravel content using random forest, spatial interpolation methods and their hybrid methods. 20th International Congress on Modelling and Simulation, Adelaide, Australia, 1–6 December, p.394-400, 2013.
81. Li, J. and Heap, A.D.: A review of comparative studies of spatial interpolation methods in environmental sciences: Performance and impact factors. Ecological Informatics, 6, 228–241, 2011. https://doi.org/10.1016/j.ecoinf.2010.12.003
82. Li, J. and Heap, A.D.: Spatial interpolation methods applied in the environmental sciences: A review. Environmental Modelling & Software, 53, 173-189, 2014. https://doi.org/10.1016/j.envsoft.2013.12.008
83. Li, J., Potter, A., Huang, Z., Daniell, J. J. and Heap, A. D.: Predicting Seabed Mud Content across the Australian Margin: Comparison of Statistical and Mathematical Techniques Using a Simulation Experiment. Geoscience Australia, Record 2010/11, 146 pp., 2010.
84. Li J., Heap, A.D., Potter, A., Daniell, J.J.: Application of machine learning methods to spatial interpolation of environmental variables. Environmental Modelling & Software, 26, 1647-1659, 2011a. https://doi.org/10.1016/j.envsoft.2011.07.004
85. Li J., Heap, A.D., Potter, A., Huang, Z., Daniell, J.J.: Can we improve the spatial predictions of seabed sediments? A case study of spatial interpolation of mud content across the southwest Australian margin. Continental Shelf Research, 31, 365–1376, 2011b. https://doi.org/10.1016/j.csr.2011.05.015
86. Li, J.,Siwabessy, P.J., Maggie, T., Zhi, H., and Andrew, D.H.: Predicting Seabed Hardness Using Random Forest in R. Data Mining Applications with R. Publisher: Elsevier, Editors: Y. Zhao and Y. Cen, pp.299-329, 2013. https://doi.org/10.1016/B978-0-12-411511-8.00011-6
87. Li J, Tran, M, Siwabessy, J: Selecting Optimal Random Forest Predictive Models: A Case Study on Predicting the Spatial Distribution of Seabed Hardness. PLoS ONE 11 (2): e0149089, 2016. https://doi.org/10.1371/journal.pone.0149089
88. Li, J., Alvarez, B., Siwabessy, J., Tran, M., Huang, Z., Przeslawski, L., Radke, L., Howard, F. and Nichol, S.: Application of random forest, generalised linear model and their hybrid methods with geostatistical techniques to

count data: Predicting sponge species richness. Environmental Modelling & Software, 97, 112-129, 2017. https://doi.org/10.1016/j.envsoft.2017.07.016

89. Liaw, A. and Wiener, M.: Classification and regression by randomForest. R News, 2/3:18–22, 2002.
90. Martínez Arbizu, P. and Haeckel, M. (eds.): RV SONNE Fahrtbericht / Cruise Report SO239: EcoResponse Assessing the Ecology, Connectivity and Resilience of Polymetallic Nodule Field Systems, Balboa (Panama) – Manzanillo (Mexico,) 11.03.-30.04.2015. GEOMAR Report, N. Ser. 025. GEOMAR Helmholtz-Zentrum für Ozeanforschung, Kiel, Germany, 204 pp, 2015. http://dx.doi.org/10.3289/GEOMAR_REP_NS_25_2015 .
91. Mascaro, J., Asner, G.P., Knapp, DE., Kennedy-Bowdoin, T., Martin, RE., Anderson, C., Higgins, M., and Chadwick, D.: A Tale of Two ''Forests'': Random Forest Machine Learning Aids Tropical Forest Carbon Mapping. PLoS ONE 9(1): e85993, 2014. https://doi.org/10.1371/journal.pone.0085993
92. Mayer, L.A. Frontiers in seafloor mapping and visualization. Marine Geophysical Researches, 27, 7-17, 2007. http://dx.doi.org/10.1007/s11001-005-0267-x
93. Meinshausen, N.: Quantile Regression Forests. Journal of Machine Learning Research 7, 983–999, 2006.
94. Meinshausen, N.: quantregForest 0.2-3. R package, 2012. https://CRAN.R-roject.org/package=quantregForest
95. Mentch, L. and Hooker, G.: Quantifying uncertainty in random forests via confidence intervals and hypothesis tests. Journal of Machine Learning Research, 17(1):841–881, 2016.
96. Millard, K. and Richardson, M.: On the Importance of Training Data Sample Selection in Random Forest Image Classification: A Case Study in Peatland Ecosystem Mapping. Remote Sensing, 7, 8489-8515, 2015. https://doi.org/10.3390/rs70708489
97. Miller, B. A., Koszinski, S., Wehrhan, M., Sommer, M.: Impact of multi-scale predictor selection for modeling soil properties. Geoderma 239–240, 97–106, 2015. http://dx.doi.org/10.1016/j.geoderma.2014.09.018
98. Mitchell, M. W.: Bias of the Random Forest Out-of-Bag (OOB) Error for Certain Input Parameters. Open Journal of Statistics, 1, 205-211, 2011. https://doi.org/10.4236/ojs.2011.13024
99. Moellering, H. and Tobler, W.: Geographical variances. Geogr. Anal. 4(1), 34–50, 1972. https://doi.org/10.1111/j.1538-4632.1972.tb00455.x
100. Moran, P.A.P.: The interpretation of statistical maps, Journal of the Royal Statistics Society, Series B (Methodological), 10, 2, 243 – 251, 1948.
101. Moran, P.A.P.: Notes on Continuous Stochastic Phenomena. Biometrika, Vol. 37, No. 1/2, 17-23, 1950. http://dx.doi.org/10.2307/2332142
102. Mucha, J. and Wasilewska – Błaszczyk, M.: Variability and Accuracy of Polymetallic Nodules Abundance Estimations in the IOM Area − Statistical and Geostatistical Approach. Proceedings of the Tenth (2013) ISOPE Ocean Mining and Gas Hydrates Symposium, Szczecin, Poland, September 22-26, 27-31, 2013.
103. Mukaka, M.M.: Statistics corner: A guide to appropriate use of correlation coefficient in medical research. Malawi Medical Journal, 24 (3), 69-71, 2012.

104. Obermeyer, Z., and Emanuel, E.J.: Predicting the Future — Big Data, Machine Learning, and Clinical Medicine. N Engl J Med., 29, 375(13): 1216–1219, 2016. http://dx.doi.org/10.1056/NEJMp1606181

105. Okazaki, M. and Tsune, A.: Exploration of Polymetallic Nodule Using AUV in the Central Equatorial Pacific. Proceedings of the Tenth ISOPE Ocean Mining and Gas Hydrates Symposium, Szczecin, Poland, September 22-26, 2013.

106. Okun, O. and Priisalu, H.: Random Forest for Gene Expression Based Cancer Classification: Overlooked Issues. Pattern Recognition and Image Analysis. 4478, 483-490, 2007. http://dx.doi.org/10.1007/978-3-540-72849-8_61

107. Paduan, B. J., Caress, D.W.; Clague D.A., Paull, C.K., Thomas H. High-Resolution Mapping of Mass Wasting, Tectonic, and Volcanic Hazards Using the MBARI Mapping AUV. Rend. online Soc. Geol. It., 7, 181-186, 2009.

108. Paul, C.K., Anderson, K., Caress, D.W., Lundsten, E., Gwiazda, R.: Chapter 15: Fine Scale Morphology of Tubenworm Slump, Monterey Canyon. In: Submarine mass movements and their consequences, 7th International Symposium, edited by Lamarche G., Mountjoy J., Bull S., Hubble T, Krastel S., Lane E., Micallef Aar., Moscardelli L., Mueller C., Pecher In., Woelz S. Advances in Natural and Technological Hazards Research 41, Springer, 2016. https://doi.org/10.1007/978-3-319-20979-1_15

109. Paull, L., Saeedi S., Seto M.and Li H.: AUV Navigation and Localization: A Review. IEEE Journal of oceanic Engineering, 39 (1), 131-149, 2014. http://dx.doi.org/10.1109/JOE.2013.2278891

110. Petersen S., Hannington, M. and Krätschell, A.: Technology developments in the exploration and evaluation of deep-sea mineral resources. Responsabilité and Environment, 85, 14-18, 2017.

111. Peukert, A., Schoening, T., Alevizos, E., Köser, K., Kwasnitschka, T., and Greinert, J.: Understanding Mn-nodule distribution and evaluation of related deep-sea mining impacts using AUV-based hydroacoustic and optical data, Biogeosciences, 15, 2525-2549, 2018a. https://doi.org/10.1007/978-3-319-57852-1_24

112. Peukert, A., Petersen, S., Greinert, J., and Charlot, F.: Seabed Mining, In: Micallef, A, Krastel, S., and Savini, A. (eds): Submarine Geomorphology, p.481-502. Springer Geology. Springer, Cham, 2018b. https://doi.org/10.1007/978-3-319-57852-1

113. Pierdomenico, M., Guida, V.G., Macelloni, L., Chiocci, F.L., Rona, P.A., Scranton, M.I, Asper, V., Diercks, A. Sedimentary facies, geomorphic features and habitat distribution at the Hudson Canyon head from AUV multibeam data. Deep-Sea Research II 121, 112–125, 2015. http://dx.doi.org/10.1016/j.dsr2.2015.04.016

114. Preston, J.: Automated acoustic seabed classification of multibeam images of Stanton Banks. Applied Acoustics, 70, 1277–1287, 2009. http://dx.doi.org/10.1016/j.apacoust.2008.07.011

115. R Development Core Team: R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2008. ISBN 3-900051-07-0. http://www.R-project.org

116. Rahn, M.: Deposit models. Deliverable D3.11 of the EU-Project Blue Mining. BGR Hannover, 43-47, 2017. www.bluemining.eu/downloads

117. Riley, S. J., DeGloria, S. D. and Elliot, R: A terrain ruggedness index that quantifies topographic heterogeneity, Intermountain Journal of Sciences, vol. 5, No. 1-4, 1999.

118. Roberts, J.J., Best, B.D., Dunn, D.C., Treml, E.A. and Halpin, P.N.: Marine Geospatial Ecology Tools: An integrated framework for ecological geoprocessing with ArcGIS, Python, R, MATLAB, and C++. Environmental Modelling & Software 25: 1197-1207. http://dx.doi.org/10.1016/j.envsoft.2010.03.029

119. Rodriguez-Galiano, V.F., Chica-Olmo, M. and Chica-Rivas, M.: Predictive modelling of gold potential with the integration of multisource information based on random forest: a case study on the Rodalquilar area, Southern Spain, International Journal of Geographical Information Science, 28:7, 1336-1354, 2014. http://dx.doi.org/10.1080/13658816.2014.885527

120. Rodriguez-Galiano, V.F., Sanchez-Castillo, M., Chica-Olmo, M. and Chica-Rivas, M.: Machine learning predictive models for mineral prospectivity: An evaluation of neural networks, random forest, regression trees and support vector machines. Ore Geology Reviews, 71, 804–818, 2015. http://dx.doi.org/10.1016/j.oregeorev.2015.01.001

121. Rogers, J.: Seismic, Bathymetric, and Photographic Evidence of widespread erosion and a manganese-nodule pavement along the continental rise of the southeast Cape Basin. Marine Geology, 78, 57-76, 1987. https://doi.org/10.1016/0025-3227(87)90068-5

122. Roman, C. and Mather, R.: Autonomous underwater vehicles as tools for deep-submergence archaeology. Proceedings of the Institution of Mechanical Engineers, Part M: Journal of Engineering for the Maritime Environment 224 (4), 327-340, 2010. http://dx.doi.org/10.1243/14750902JEME202

123. Rühlemann, C., Kuhn, T., Vink, A.,Wiedicke, M.: Methods of Manganese Nodule Exploration in the German License Area. Recent Developments in Atlantic Seabed Minerals Exploration and Other Topics, 42nd Underwater Mining Institute , 21-29 October, Rio de Janeiro and Porto de Galinhas, Brazil, 2013.

124. Rühlemann, C, Kuhn, T, Wiedicke, M.: Current Status of Manganese Nodule Exploration in the German License Area Proceedings of the Ninth ISOPE Ocean Mining Symposium Maui, Hawaii, USA, June 19-24, 168-173, 2011.

125. Ruß, G., and Kruse, R.: Regression Models for Spatial Data: An Example from Precision Agriculture. CDM 2010. Lecture Notes in Computer Science, vol 6171. Springer, Berlin, Heidelberg, 2010. https://doi.org/10.1007/978-3-642-14400-4_35

126. Santibanez, S., Lakes, T., and Kloft, M.,: Performance Analysis of Some Machine Learning Algorithms for Regression Under Varying Spatial Autocorrelation. The 18th AGILE International Conference on Geographic Information Science, Lisboa (Portugal), 9-12 June, 2015a.

127. Santibanez, S. F., Kloft, M., and Lakes T.: "Performance Analysis of Machine Learning Algorithms for Regression of Spatial Variables. A Case Study in the Real Estate Industry." the 13th International Conference of GeoComputation, Dallas (USA), May 20 – 23, 2015b.

128. Schmueli, G.: To Explain or to Predict? Statistical Science; 25, 3, 289–310, 2010. http://dx.doi.org/10.1214/10-STS330

129. Schoening, T., Kuhn, T., Nattkemper, T.W.: Estimation of poly-metallic nodule coverage in benthic images. In: Proceedings of the 41st Conference of the UnderwaterMining Institute. UMI, 2012a.

130. Schoening, T., Bergmann, M., Ontrup, J., Taylor, J., Dannheim, J. Gutt, J., Purser, A., Nattkemper, T. W.: Semi-Automated Image Analysis for the Assessment of Megafaunal Densities at the Arctic Deep-Sea Observatory HAUSGARTEN. PLoS ONE 7(6): e38179, 2012b. http://dx.doi.org/10.1371/journal.pone.0038179

131. Schoening, T., Kuhn, T., Nattkemper, T.W.: Seabed classification using a bag-of prototypes feature representation. In: 2014 ICPR Workshop on Computer Vision for Analysis of Underwater Imagery (CVAUI). IEEE, pp. 17–24, 2014. http://dx.doi.org/10.1109/CVAUI.2014.9

132. Schoening, T., Thomas, K., Bergmann, M., and Nattkemper, T.W.: DELPHI—fast and adaptive computational laser point detection and visual footprint quantification for arbitrary underwater image collections. Front.Mar.Sci.2:20.2015. http://dx.doi.org/10.3389/fmars.2015.00020

133. Schoening, T., Kuhn, T., Jones, D.O.B., Simon-Lledo, E., Nattkemper, T.W., 2016. Fully automated image segmentation for benthic resource assessment of poly-metallic nodules. Methods Oceanogr. 15–16, 78–89, 2016. http://dx.doi.org/10.1016/j.mio.2016.04.002

134. Schoening, T., Jones, D. O.B. and Greinert, J.: Compact-Morphology-based polymetallic Nodule Delineation. Scientific Reports, 7.1, 13338, 2017a. https://doi.org/10.1038/s41598-017-13335-x

135. Schoening, T.: Source code for the Compact Morphology-based Nodule Delineation (CoMoNoD) algorithm. PANGAEA, https://doi.org/10.1594/PANGAEA.875070 .Supplement to: Schoening T., Jones D. O. B., Greinert, J.: Compact-Morphology-based poly-metallic Nodule Delineation. Scientific Reports, 7 (1), 2017b. https://doi.org/10.1038/s41598-017-13335-x

136. Schoening, T.: Results of nodule detection along AUV tracks during SONNE cruises SO239 and SO242/1. PANGAEA, 2017c. https://doi.org/10.1594/PANGAEA.883838

137. Sibenac, M., Podder, T., Kirkwood, W. Thomas, H.: Autonomous Underwater Vehicles for Ocean Research: Current Needs and State of the Art Technologies. Marine Technology Society Journal, 38 (2), 63-72, 2004. https://doi.org/10.4031/002533204787522848

138. Sharma, R.: Quantitative estimation of seafloor features from photographs and their application to nodule. Marine Georesources & Geotechnology, 11:4, 311-331, 1993. http://dx.doi.org/10.1080/10641199309379926

139. Sharma, R. and Kodagali, V.: Influence of seabed topography on the distribution of manganese nodules and associated features in the Central Indian Basin: A study based on photographic observations, Marine geology, 110, 153–162, 1993. https://doi.org/10.1016/0025-3227(93)90111-8

140. Sharma, R. Sankar, S. J., Samanta, S., Sardar, A.A. and Gracious, D.: Image analysis of seafloor photographs for estimation of deep-sea minerals. Geo-Mar Lett. 30, 617–626, 2010. http://dx.doi.org/10.1007/s00367-010-0205-z

141. Sharma, R., Khade, N.H., and Sankar, S. J.: Assessing the distribution and abundance of seabed minerals from seafloor photography data in the Central Indian Ocean Basin. Int. J. Remote Sens., vol.34, 1691-1706, 2013. https://doi.org/10.1080/01431161.2012.725485

142. Skornyakova, N. and Murdmaa, I.: Local variations in distribution and composition of ferromanganese nodules in the Clarion-Clipperton Nodule Province, Marine Geology, 103, 381–405, 1992. https://doi.org/10.1016/0025-3227(92)90028-G

143. SPC: Deep Sea Minerals: Manganese Nodules, a physical, biological, environmental, and technical review. Baker, E., and Beaudoin, Y. (Eds.) Vol. 1B, Secretariat of the Pacific Community, 2013.

144. Stephens, D. and Diesing, M.: A Comparison of Supervised Classification Methods for the Prediction of Substrate Type Using Multibeam Acoustic and Legacy Grain-Size Data. PLoS ONE 9 (4), e93950, 2014. https://doi.org/10.1371/journal.pone.0093950

145. Strobl, C. Boulesteix, A.L., Zeileis, A. and Hothorn, T.: Bias in random forest variable importance measures: Illustrations, sources and a solution. BMC Bioinformatics, 8:25, 2007. https://doi.org/10.1186/1471-2105-8-25

146. Strobl, C., Boulesteix, A.L., Kneib, T., Augustin, T. and Zeileis, A.: Conditional variable importance for random forests. BMC Bioinformatics, 9:307, 2008a. https://doi.org/10.1186/1471-2105-9-307

147. Strobl, C., Zeileis, A.: Danger: High Power! – Exploring the Statistical Properties of a Test for Random Forest Variable Importance. Proceedings of the 18th International Conference on Computational Statistics, Porto, Portugal 2008b.

148. Strobl, C., Malley, J., Tutz, G.: An introduction to recursive partitioning: rationale, application, and characteristics of classification and regression trees, bagging, and random forests. Psychol Methods, 14 (4): 323-48, 2009. https://doi.org/10.1037/a0016973

149. Tsune, A. and Okazaki, M.: Some Considerations about Image Analysis of Seafloor Photographs for Better Estimation of Parameters of Polymetallic Nodule Distribution. Proceedings of the Twenty-fourth (2014) International Ocean and Polar Engineering Conference, Busan, Korea, June 15-20, 72-77, 2014.

150. Tung N.T., Huang J.Z., Khan I., Li M.J., Williams G.: Extensions to Quantile Regression Forests for Very High-Dimensional Data. In: Tseng V.S., Ho T.B., Zhou ZH., Chen A.L.P., Kao HY. (eds) Advances in Knowledge Discovery and Data Mining. PAKDD 2014. Lecture Notes in Computer Science, vol 8444. Springer, Cham https://doi.org/10.1007/978-3-319-06605-9_21

151. UNOET: Delineation of mine sites and potential in different sea areas, vol. 9. Seabed Minerals Series 4. London: Graham & Trotman, 1987.

152. van der Ploeg, T., Austin, P.C., and Steyerberg E.W.: Modern modelling techniques are data hungry: a simulation study for predicting dichotomous end points. Medical Research Methodology, 14:137, 2014. https://doi.org/10.1186/1471-2288-14-137

153. Hari, V. N., Kalyan, B., Chitre, M., and Ganesan, V.: Spatial Modeling of Deep-Sea Ferromanganese Nodules With Limited Data Using Neural Networks. IEEE Journal of Oceanic Engineering, 1-18, 2017. https://doi.org/10.1109/JOE.2017.2752757

154. Volkmann, S.E.: Concept for Sustainable Economic Evaluation. Deliverable D1.41 of the EU-Project Blue Mining. BGR Hannover, 43-47, 2017. www.bluemining.eu/downloads

155. Volkmann, S.E., and Lehnen, F.: Production key figures for planning the mining of manganese nodules. Marine Georesources & Geotechnology, 36:3, 360-375,2018. https://doi.org/10.1080/1064119X.2017.1319448

156. von Stackelberg, U. and Beiersdorf, H.: The formation of manganese nodules between the Clarion and Clipperton fracture zones southeast of Hawaii, Marine Geology, 98, 411–423, 1991. https://doi.org/10.1016/0025-3227(91)90113-I

157. Wager, S., Hastie, T., and Efron, B.: Confidence intervals for random forests: the jackknife and the infinitesimal jackknife. Journal of Machine Learning Research, 15(1):1625–1651, 2014.

158. Willmott, C.J. and Matsuura, K.: Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. Vol. 30: 79–82, 2005. https://doi.org/10.3354/cr030079

159. Wilson, M.F.J., O'Connell, B., Brown, C., Guinan J.C., Grehan, A.J.: Multiscale Terrain Analysis of Multibeam Bathymetry Data for Habitat Mapping on the Continental Slope. Marine Geodesy, 30: 3–35, 2007. https://doi.org/10.1080/01490410701295962

160. Wynn, R.B., Huvenne, V.A.I., Le Bas, T.P., Murton, B.J., Connelly, D.P., Bett, B.J., Ruhl, H.A., Morris, K.J., Peakall, J., Parsons, D.R., Sunner, S.J., Darber S.E., Dorrell, R.M., Hunt, J.E. Autonomous Underwater Vehicles (AUVs): Their past, present and future contributions to the advancement of marine geoscience. Marine Geology, 352, 451–468, 2014. http://dx.doi.org/10.1016/j.margeo.2014.03.012

161. Xu, L., Saatchi, SS., Yang, Y., Yu, Y., and White, L.: Performance of non-parametric algorithms for spatial mapping of tropical forest structure. Carbon Balance Manage, 11:18, 2016. https://doi.org/10.1186/s13021-016-0062-9

162. Yazici, B. and Yolacan, S.: A comparison of various tests of normality, Journal of Statistical Computation and Simulation, 77:2, 175-183, 2007. http://dx.doi.org/10.1080/10629360600678310

163. Zevenbergen, L. W. and Thorne, C. R.: Quantitative analysis of land surface topography. Earth Surf. Process. Landforms, 12: 47–56, 1987. https://doi.org/10.1002/esp.3290120107