Automatic Detection of Seafloor Bedforms Using Bathymetric Data

Larissa M. Freguete[1], Luiz A. Pinto[1], Alex C. Bastos[2], Gustavo M. de Almeida[1]

[1] *Control and Automation Engineering Master Program*
Instituto Federal do Espírito Santo, Serra, Brazil
lalafreguete@hotmail.com, pintoluizalberto@gmail.com, gmaia@ifes.edu.br
[2] *Oceanography and Ecology Department*
Universidade Federal do Espírito Santo, Vitória, Brazil
alex.bastos@ufes.br

Abstract. Seafloor bedforms are sedimentary structures that can reveal local hydrodynamics provide auxiliary information for the mapping of benthic habitats and can present a risk to navigation and marine structures. Therefore, it is beneficial to have a method that can provide accurate bedform classification with repeatability and shorter processing time. That is the purpose of this work, to compare the performance of Random Forest (RF) and Support Vector Machines (SVM) models for obtaining automatic bedform classification for bathymetry point clouds. The models were trained with a dataset of 4000 sampling from the Doce River and Recifes Esquecidos areas. The RF model presented an accuracy of 64% and the SVM model, 49%. A higher accuracy was observed for the none class (related to reef structures). The areas presented very differing physiography which led to a poor fitting of the models, suggesting the need for a larger dataset with different combinations of predictor variables.

Keywords: automatic detection and classification, bedforms, bathymetry, points cloud, machine learning.

1 Introduction

In the realm of classification of bathymetric data regarding the type of morphological features, it can be observed in the literature two lacking or problematic topics: a precise and repeatable methodology for the classification of these features and the use of such methodology on 3D bathymetry data, point clouds. Since the last decade, authors have been suggesting and asking for a complex approach for bedform classification as semantic segmentation or object-oriented approaches to increase classification accuracy. Until 2020, the most common approaches relied on pixel or object-based methods. However, these methods were shown to be insufficient to follow seafloor complexity given that the first method has low performance on non-uniform surfaces and the second method is sensitive to morphological discontinuity (Dekavalla and Argialas [1]; Li et al. [2]). Although research on automatic benthic mapping has been carried out since 2008, until today there is no decision on the best technique for that. Even in recent years, there are remaining problems as algorithms are specific to local scale and context, dependence on parameter adjustment, and some might require manual steps (Menandro et al. [3]; Leon et al. [4]; Summers et al. [5]). To overcome these issues, the newest works have resorted to Machine Learning (ML) techniques (Li et al. [2]; Summers et al. [5]; Le Deunf et al. [6]; Janowski et al. [7]). The use of point cloud for the marine realm is still very scarce, in general, even in publications from last year, authors indicate of need for research using this type of data. It is acknowledged there are many algorithms for point clouds generated for continental scenarios but for marine environments, there is still a lack of robust techniques.

Therefore, the purpose of this study is to show the results of the initial phase of a more comprehensive exploratory study for Machine Learning (ML) algorithms applied to the automation of the bathymetry data interpretation regarding the occurrence of seafloor bedforms. The idea is to investigate the performance of simpler algorithms such as Random Forest (RF), Support Vector Machines (SVM), Partial Least Square (PLS), and Ensemble algorithms, then move to more complex algorithms based on neural networks, such as YOLO 8 and U-Net, for image classification to ultimately projecting the classification to the bathymetry point cloud. The final stage is proposed to be the use of Deep Neural Networks (DNN) optimized for 3D data as PointNet, 3DMASC, and PointCNN for a direct automated point cloud classification. This work will present only the results of the test with RF and SVM models.

2 Methodology

This study was carried out in a gradual fashion where the first step was collecting the data and post-processing them in the laboratory. After acquiring the processed bathymetry, it was necessary to create the derivatives as the explanatory variables for the chosen models, generate the reference (called ground truth) dataset, and then implement the models and their performance, Figure 1. These steps are explained in more detail in the following subsections of this chapter.

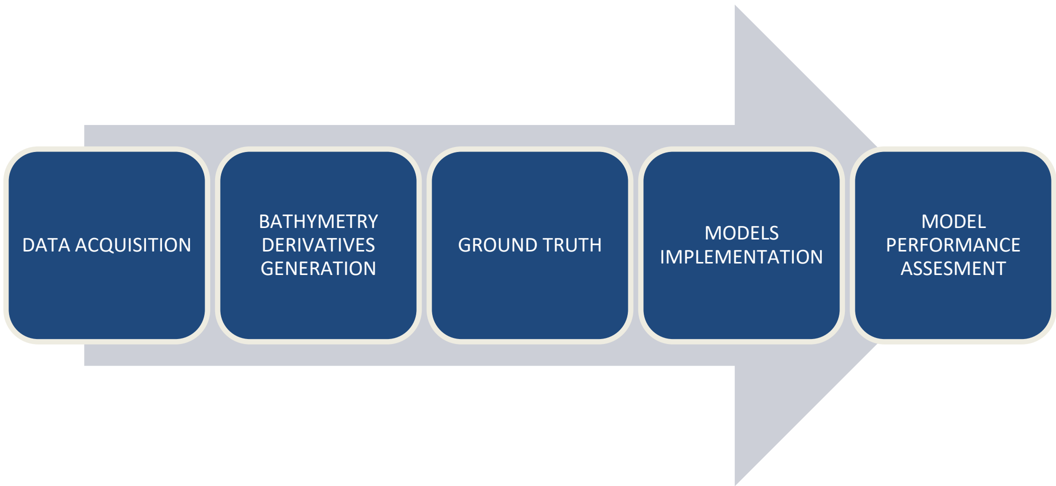

Figure 1. Overview of the methodology steps.

2.1 Dataset Acquisition

The point clouds and Digital Bathymetry Models (DBMs) were obtained from hydrography surveys over the Espírito Santo Continental Shelf (southeastern Brazil). The data was collected with an R2Sonic 2024 multibeam echosounder (MBES). The surveyed sites are in front of the Doce River (RD) mouth and within an area richly covered with reef structures, the Recifes Esquecidos (RE), at the north of the river mouth (Vieira et al. [8]).

The predictor datasets were built with ESRI ArcGIS Pro software, where 9 types of rasters were created containing the following variables: bathymetry, slope, aspect, planar curvature, profile curvature, curvature, geomorphons, landform units, and the Terrain Ruggedness Index (TRI) as a proxy of rugosity. The predictor dataset for the RD mouth area was generated with a spatial resolution of 10 x 10 m, and RE area, of 1 x 1 m.

Seafloor bedform classification can be a cumbersome task as there are diverse manners of doing so as classification based on morphology, sedimentology, or even genetics. For a precise classification, it is necessary to know the local hydrodynamics, the sediment granulometry, and the feature's real dimensions, as highlighted by Sánchez [9]. As one way of overcoming this issue, the classification used in this work considered the features' arrangement to the flow direction, so the bedforms were categorized as only 3 types: transversal bedforms (structures oriented perpendicularly to the flow), longitudinal (structures are the ones with spacing parallel to the flow), and the transitional (features which arrangement orientation couldn't be well defined).

Apart from these three classes (transversal, parallel, and transitional), there were also included other 3 classes: artifacts (related to the corrugation-like features caused by the surveying method inherent issues), none (for this specific case, they are related to reefs, which are rigid structures and don't represent the local hydrodynamics), and the ground (chosen as the background class), Figure 2. The ground truth dataset totalizes in 4000 samplings. The distribution of classes within the ground truth dataset is shown in Table 1.

The 4000-sampling dataset was split into training and testing with a ratio of 80% (3198 samplings) for training and 20% (800 samplings) for testing. The splitting was achieved in a random manner using Scikit Learn's train_test_split tool. To avoid model overfitting, a k-fold cross-validation was performed with the training dataset using the cross-validate tool from the same Scikit module.

Table 1. Distribution of ground truth for each class.

CLASS	COUNTS	%
Artifact	959	23.98
Ground	1194	29.86
None	467	11.68
Parallel	705	17.63
Transitional	407	10.18
Transversal	266	6.65

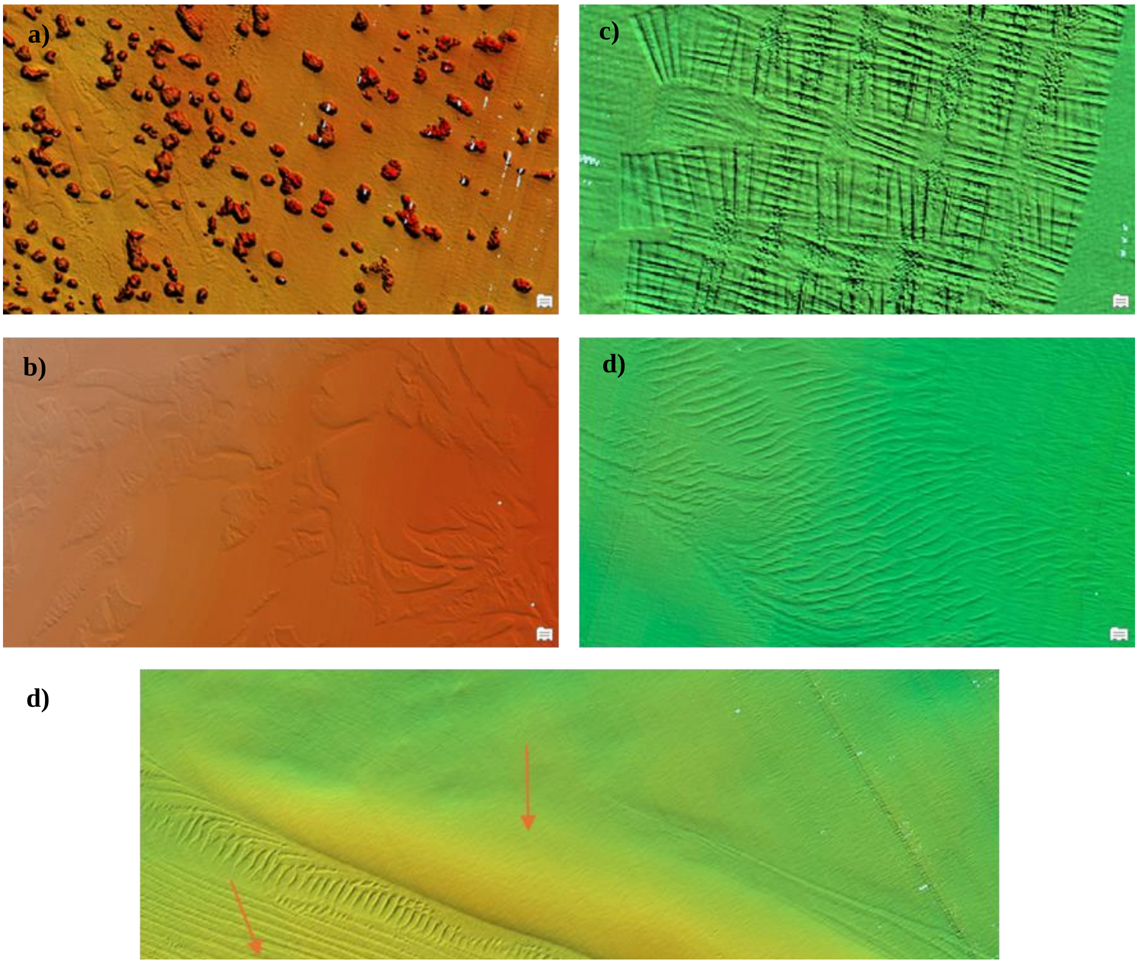

Figure 2. Exemplification of classes. a) None (reefs). b) Artifact (corrugation/noise). c) Transitional. d) Transversal. e) Parallel.

2.2 Models

The models were generated and executed with Scikit-Learn's RandomForestClassifier and SVM tools inside a Python environment, 3.11 version, in a machine with the following settings: Intel Core i7-8700 3.20 GHz (12 CPUs) processor, 64 GB RAM, and two Nvidia GeForce GTX 1050 Ti graphic cards.

The RF model was built with the following hyperparameters: Gini as the function for measuring the quality of data split, maximum tree depth as 10, minimum number of samples to split node ass 2, minimum number of samples leaf node as 1, and number of trees as 100. For the SVC model, the chosen hyperparameter values were regularization equal to 1, and radial basis function as the kernel.

2.3 Model Assessment Metrics

The models` performance was evaluated based on the precision, recall, and F1-score metrics. Precision is a metric that helps to evaluate the model's ability to classify correctly the pixels and it is obtained by

$$\text{Precision}=\frac{\text{TP}}{\text{TP+FP}}\ . \tag{1}$$

Recall, or sensitivity, measures the model's ability to correctly predict the background pixels and is calculated by

$$\text{Recall}=\frac{\text{TP}}{\text{TP+FN}}\ . \tag{2}$$

And the last metric, which was considered as the metric for accuracy, is the F1-score. F1-score is a harmonic mean of the recall and precision and is calculated by

$$\text{F1 Score}=\frac{2xPrecisioxRecall}{Precision+Recall}\ . \tag{3}$$

3 Results and Discussion

3.1 Bedform Classification

The RF and SVM were chosen as models based on results found in the work of Summers et al. [5] and Janowski et al. [7] where they found accuracies of 94% for the RF model and 73% for the SVM model in applications similar to this study.

The training and validation phase results show that for this specific dataset, the RF had a better overall accuracy of 64.25% while the SVM could only distinguish the classes with 44% accuracy. Regarding their inter-class performance, both could classify the reef structures (class "none" or 1) with high accuracy, F1-scores of 96% and 88%, RF, and SVM, respectively. The reefs are distinguishable in all the predictors, enabling a better model fitting for this class. Apart from this class, their behavior was different for each class.

The SVM had the second highest F1-score for the artifact class (class 3), with 52%. This class represents the corrugation-like effect in the DBM and creates a powerful and distinguishable signal in the slope map, for example. This can lead to a false positive, as other features can also have strong slope values and be mistakenly classified as artifacts. The assumption of this effect can be justified by the low precision value for this class, 42%, Table 2. For the remaining classes, the SVM had a poor performance with accuracy less than or equal to 50%.

The RF model had accuracy values for the ground (2), artifacts (3), and transversal (5) classes that were close to the overall accuracy, 67%, 64%, and 62% respectively, Table 2. After the none class, the ground has the highest Recall score, indicating the model missed a few ground pixels. This was not expected, as the ground class represents pixels with smooth/mild signals for all the predictor rasters comparatively to artifacts. However, when looking at the Precision values, the RF model did have lower values for the ground class, indicating misclassification, as in the case of the top of sandbars (class parallel) being designated as ground in the RE section, Figure 3. All the other classes, parallel (5) and transitional (6), had accuracy lower than 50%.

We can assume that the models' poor performance can be linked to the dataset used for model training. The selected areas are geomorphically different from each other, the RD sector is located in front of a river mouth which is considered to have a higher sediment apport and a high incidence of incised valley and hard ground. While the RE sector shows a relatively dense reef coverage and a more diverse occurrence of parallel, transversal, transitional forms. This creates a wider range of values of predictor variables for each class and makes necessary a larger dataset that will allow a better-fitted model.

In a qualitative analysis, we can see the issues shown by the assessment metric. In the output classification maps, Figure 3a, the reefs were rightly classified as class none, the ground class was well represented but often misclassified as parallel. The furrow marks caused by the flow-driven erosion were correctly detected as a parallel class, however, the model also assigned a parallel class to the pixel closer to the artifacts that are not related to true parallel bedforms. Artifact class was assigned to the corrugation-like pixel but also to pixels where especially the slope has a higher value, for example, the border of furrow marks.

In the subset area shown in Figure 3b, the model had a performance increase for artifact class, this area subset presented a high occurrence of "noise" due to the surveying issues (corrugation texture) and the pixels were correctly classified, except for a small quantity of pixels around the parallel classes (sandbars). There was good coverage of ground pixels, however, with a significant number of pixels misclassified as transitional.

Figure 3c is a classification map for the RD sector, where it was observed to be a more challenging area. This subset is covered with not well-defined transitional/parallel bedforms and, additionally, is hollowed by many incised valleys. The valleys were wrongly assigned to artifact pixels and there was confusion between the transitional and parallel classes.

Table 2. Validation metrics results for RF and SVM models.

	RF			SVM		
CLASS	PRECISION	RECALL	F1-SCORE	PRECISION	RECALL	F1-SCORE
1	0.93	1.00	0.96	0.83	0.95	0.88
2	0.59	0.77	0.67	0.42	0.62	0.50
3	0.62	0.66	0.64	0.49	0.56	0.52
4	0.52	0.40	0.45	0.42	0.15	0.22
5	0.67	0.58	0.62	0.37	0.56	0.45
6	0.64	0.30	0.41	0.00	0.00	0.00
WEIGHTED AVERAGE	0.64	0.64	0.63	0.43	0.49	0.44

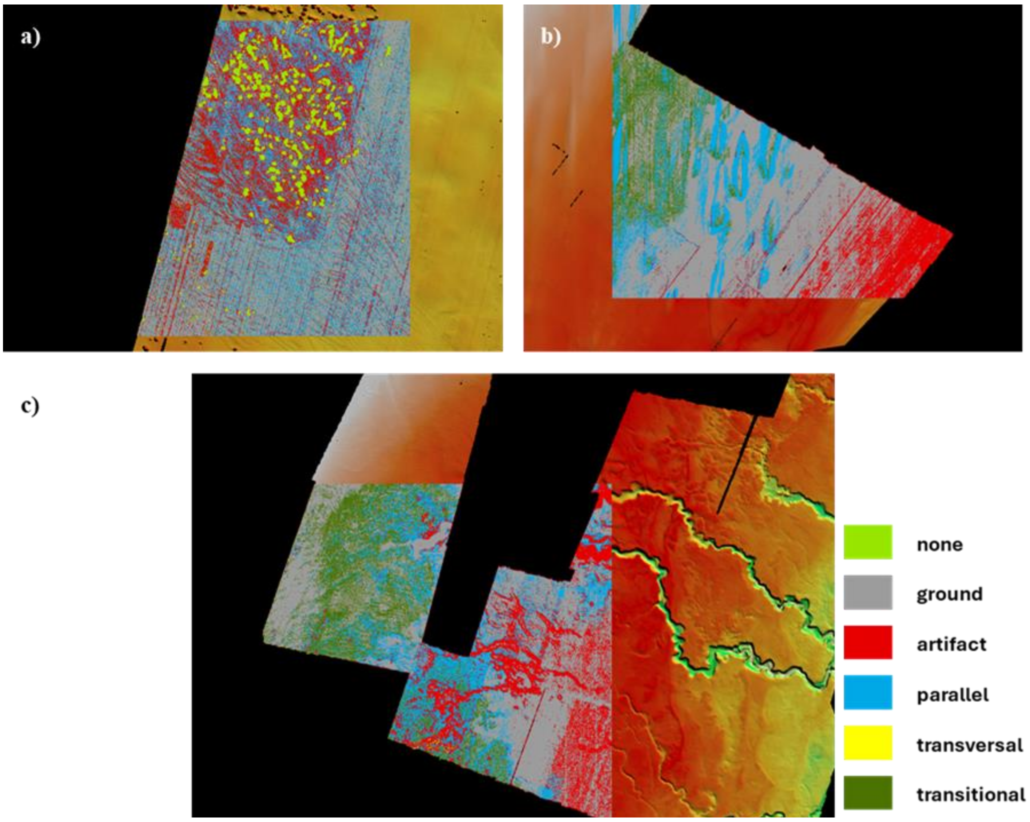

Figure 3. Output classification maps for the RE and RD sectors. Classification results are overlaying the bathymetry rasters.

4 Conclusions

This study had the goal of comparing the performance of RF and SVM models for bedforms classification as the initial stage of exploratory study for bathymetry point cloud automatic classification. Both models could detect the reefs (class none) with a fair high F1-score, due to their distinguishable bathymetric signature and, consequently, in all the derivatives. The RF model resulted in an F1-score of 96% and the SVM, 88%, for this class. This is the class with the best performance for both models.

The RF presented the highest overall accuracy with an F1-score of 64%, however, its performance varied between the study areas. Conversely, it was also observed ubiquitous low performance for the bedforms with smoother physiography, the transitional and parallel classes, with F1-scores lower than 50% for both models.

The qualitative and quantitative analyses show that for performance improvement, it is necessary a larger dataset to include more variation into the dataset, it is also possible the inclusion of more explanatory variables, an analysis of its effect, and variable selection. As a recommendation of a promising variable, future work can add backscatter data as the additional variable.

As another recommendation for succeeding in the driven intention of this study, which is automating bedforms detection and classification, more models can be tested and verified.

Acknowledgements. I must say this is only the initial stage of a more complex application which could only be kicked off with financial support from the Programa de Monitoramento da Biodiversidade Áquatica (PMBA) and the support from the Marine Geosciences Laboratory (LaboGeo) which provided the data, the scientific knowhow, and the physical resources for computational processing.

Authorship statement. The authors hereby confirm that they are the sole liable persons responsible for the authorship of this work and that all material that has been herein included as part of the present paper is either the property (and authorship) of the authors or has the permission of the owners to be included here.

References

[1] M. Dekavalla. and D. Argialas. “Object-based classification of global undersea topography and geomorphological features from the srtm30_plus data”. Geomorphology, Elsevier, v. 288, p. 66–82, 2017.
[2] S. Li, L. Xiong, G. Tang, J. Strobl. “Deep learning-based approach for landform classification from integrated data sources of digital elevation model and imagery”. Geomorphology, Elsevier, v. 354, p. 107045, 2020.
[3] P.S. Menandro, A. C. Bastos, G. Boni, L. C. Ferreira, F. V. Vieira, A. C. Lavagnino, R. L. Moura, M. Diesing. “Reef mapping using different seabed automatic classification tools”. Geosciences, MDPI, v. 10, n. 2, p. 72, 2020.
[4] A. Z. Leon, V. A. I. Huvene, N. Benoist, M. Fergusson. “Assessing the repeatability of automated seafloor classification algorithms, with application in marine protected area monitoring”. Remote Sensing, MDPI, v. 12, n. 10, p. 1572, 2020.
[5] G. Summers, A. Lim, A. J. A. Wheeler.” A scalable, supervised classification of seabed sediment waves using an object-based image analysis approach”. *Remote Sensing*, MDPI, v. 13, n. 12, p. 2317, 2021.
[6] J. Le Deunf, R. Mishra, Y. Pastol, R. Billot, S. Oudot.“Seabed prediction from airborne topo-bathymetric lidar point cloud using machine learning approaches”. In: IEEE. *OCEANS 2021: San Diego–Porto*. p. 1–9.
[7] L. Janowski, R. Wroblewski, M. Rucinska, A. Kubowicz-Grajewska, P. Tysiac. “Automatic classification and mapping of the seabed using airborne lidar bathymetry”. *Engineering Geology*, Elsevier, v. 301, p. 106615, 2022.
[8] F. V. Vieira, A. C. Bastos, V. da S. Quaresma. “Submerged reef and inter-reef morphology in the western south atlantic, abrolhos shelf (Brazil)”. *Geomorphology*, Elsevier, v. 442, p. 108917, 2023.
[9] Q. G. Sánchez. “Bedforms and associated sediment dynamics on the inner shelves at different spatio-temporal scales”. PhD thesis. Universitat Politècnica de Catalunya, 2019.