High Resolution Satellite Bathymetry Mapping: Regression and Machine Learning Based Approaches

Francisco Eugenio, *Senior Member, IEEE,* Javier Marcello, *Senior Member, IEEE,* Antonio Mederos-Barrera, Ferran Marqués, *Senior Member, IEEE*

***Abstract*—Remote spectral imaging of coastal areas can provide valuable information for their sustainable management and conservation of their biodiversity. Unfortunately, such areas are very sensitive to changes due to human activity, natural phenomenona, introduction of non-native species and climate change. Thus, the main objective of this research is the implementation of a robust image processing methodology to produce accurate bathymetry maps in shallow coastal waters using high resolution multispectral WorldView-2/3 satellite imagery for the monitoring at the maximum spatial and spectral resolutions. Two different island ecosystems have been selected for the assessment, since they stand out for their richness in endemic species and they are more vulnerable to climate change: Cabrera National Park and Maspalomas Natural Protected area, located in the Balearic and Canary Islands, Spain, respectively. In addition, a third example to show the applicability of the mapping methodology to monitor the construction of a new port in Granadilla (Canary Islands) is presented. Contributions of this work focus on improving the preprocessing methodology and, mainly, on the proposal and assessment of new satellite derived regression and machine learning bathymetric models, which have been validated and compared with respect to measured reference bathymetry. After a thorough analysis of 9 techniques, using visual and quantitative statistical parameters, ensemble learning approaches have demonstrated excellent performance, even in challenging scenarios up to 35 m depth, with mean RMSE values around 2 m.**

***Index Terms*— Satellite-Derived Bathymetry (SDB), multispectral WorldView-2/3, shallow coastal water, regression and machine learning based techniques.**

I. Introduction

Currently, remote sensing of coastal areas can provide valuable information for optically characterizing and monitoring shallow coastal waters. The International Hydrographic Organization (IHO) indicates that, at least, half of the world's shallow coastal waters remain unmapped or poorly understood [1]. About 15% (20 million km^2) of all land on the planet is under some form of protection. Specifically, in Spain, protected areas cover 27.21% of its total surface and the coastline length is about 7,880 km. Preserved marine regions exceed the 8% but, nevertheless, only 25% of the littoral areas are being monitored in detail [2]. These coastal areas are essential for the conservation of biodiversity and the provision of basic services for the society.

Specifically, extracting bathymetric information and mapping the seafloor to monitor coral reefs or benthic habitats, are crucial for a wide range of near-shore activities and applications such as climate change monitoring, environmental management, engineering and maintenance of infrastructures, hydrographic applications, navigation or aquaculture, among others.

Bathymetric studies in shallow waters can be accurately performed with single or multibeam echosounders; however, such technologies are laborious and extremely expensive when covering wide areas [3]. In this context, for the monitoring and protection of coastal areas, new remote sensing sensors and techniques are suitable [4-6]. Specifically, high resolution satellite imagery can be an excellent solution to provide continuous and updated information. However, the accurate derivation of satellite products is very complex and challenging in littoral areas due to atmospheric disturbances, sunglint effects on the sea surface and the water column absorption and backscattering caused by the water inherent optical properties and the dissolved and particulate constituents [7, 8]. Consequently, selection of the appropriate satellite imagery (e.g. suitable dates, sensing geometry, calm sea areas to avoid sunglint effects, water transparency to achieve the maximum solar radiation penetration, etc.) is critical, as well as the application of advanced atmospheric and deglinting correction techniques [9-11]. Also, estimating bathymetry using optical remote data is challenging as it requires knowledge of the water-leaving radiation as a function of depth for different water composition and seafloor materials [12].

Different passive and active space and airborne platforms have been effectively used to derive bathymetry using optical, LIDAR and radar sensors [13, 14]. Passive optical remote sensing is a popular approach that takes advantage of the significant penetration capabilities of shortwave radiation, mainly around the blue and green channels. Based on passive systems, the techniques for obtaining high-resolution satellite-derived bathymetry, implemented in this research, have been grouped into: Regression-based Bathymetry Models (RBMs), and Machine Learning Bathymetry Models (MLBMs) [15].

Among the Physical-based Bathymetry Models (PBMs), we have studied the approach implemented by Lyzenga et al. [16, 17], based on the linear relationship between the depth and the neperian logarithm of the reflectance extracting the reflectance in the near-infrared (NIR) in deep waters, and the bathymetric model based on the resolution of the Radiative Transfer Equation, detailed in Eugenio et al. [8]. This model is based on the multispectral adaptation of the Hyperspectral Optimization Process Exemplar Model (HOPES) [18]. Physical-based

models have, in general, the limitation that they depend on a multitude of input parameters and training data.

To solve the problem of parameter dependence, empirical models were developed, which apply conditions and approximations to the radiometric equations, based on the use of training data. Regarding the empirical models based on regression, it is important to highlight the Stumpf model [19], widely used in Satellite-Derived Bathymetry (SDB) and considered a benchmark algorithm. It is based on the linear relationship of depth with the logarithms of different spectral bands. However, the linearity of the model is not always fulfilled in the transformed Stumpf space and the model shows a dependency on the type of seafloor. To solve the non-linearity, a multitude of approaches have been proposed. For example, the use of a quadratic regression in the Stumpf transform space [20] or the work proposed by Kanno et al. [21], applying a semi-parametric regression. In summary, regression techniques have limitations in the maximum depth at which they can obtain reliable bathymetric maps, in addition to the strong dependence with the substrate albedo.

A different paradigm to derive bathymetric information is based on Machine Learning (ML) techniques, which overcome the preceding limitations, and allow exploiting all the spectral information of the images [22]. Some works obtaining accurate bathymetric maps using Support Vector Machine (SVM) models are presented in [23-25]. Random Forest (RF) algorithms have also been used to derive bathymetry [25-27]. In particular, [26] and [27] describe the use of the K-Nearest Neighbor (KNN) and Random Forest (RF) algorithms, respectively, to achieve lower errors than Stumpf et al. [19]. In [28], Artificial Neural Networks (ANN) are used to obtain bathymetric maps, reporting minor errors. In [29] and [30] ensemble models have been applied with success in very shallow waters. In summary, better results could be obtained using MLBMs but with a high dependence on data. Therefore, it is essential to select the most suitable algorithm for an effective estimation of bathymetry [31].

In this work, contributions are provided in the overall methodology proposed to estimate bathymetry, as shown in Figure 3.

First, a combined model-based atmospheric correction and sunglint technique is presented and statistical assessed with respect to in-situ measurements. Specifically, the reflectivity obtained using the adapted 6S (Second Simulation of a Satellite Signal in the Solar Spectrum) atmospheric correction plus a deglinting method, which includes a noise reduction algorithm, is evaluated using the reference signatures measured with a field spectroradiometer, acquired at the time of satellite overflight.

Second, based on the optimal results obtained in [8] for high resolution SDB, both the intercomparison of RBMs and MLBMs is carried out. The RBMs chosen for the analysis are those implemented by Stumpf: linear and quadratic. In addition, a novel regression model is presented, termed Sigmoid-based Bathymetry Model (SBM), based on the use of the sigmoid function for modeling the non-linearity of the Stumpf relationship. On the other hand, the implemented and tested MLBMs are the linear SVM, SVM with Gaussian kernel, KNN, Decision Trees [32], and additional models that apply Ensemble Learning [33], specifically the Bagging Tree (BT) [34] and Subspace KNN (S-KNN) [35].

To our knowledge, this is the first work that provides a comprehensive review of such a large number of techniques to extract bathymetric maps. It also should be noted that a new model has been developed and others, not previously applied in this context, have been incorporated. In addition, it is important to highlight that many works only address studies at very shallow depths (typically less than 15 m) and in optimal areas with transparent and calm waters, while in this work, a complete analysis has been carried out in different areas and reaching up to considerable depths (25 m and 35 m) to study the performance in such challenging scenarios. Finally, a real case study is presented to demonstrate its applicability during the construction of a new port infrastructure.

This paper is organized as follows: Section II presents the studied areas, as well as the high resolution satellite and in-situ data used. Section III discusses the applied data pre-processing techniques. Section IV focuses on the processing methodology for bathymetry mapping of shallow-water environments, describing the regression and machine learning remote sensing-based approaches. The main results are presented and discussed in Section V. Finally, conclusions are included in Section VI.

II. Study Areas and Datasets

Two coastal areas have been selected for the study with very different characteristics: Cabrera and Maspalomas, located in the Balearic Islands (Mediterrranean Sea) and the Canary Islands (Northwest African coast, Atlantic Ocean), respectively, as shown in Fig. 1. The Cabrera Archipelago is made up of 19 islands or islets and it is the best example of undisturbed island ecosystems in the Spanish Mediterranean Sea. Since 1991, the Cabrera Archipelago has been declared National Park to preserve its rich biodiversity and, specially, the *Posidonia oceanica* seagrass beds that provide habitat for fish and crustaceans. On the other hand, the Natural Reserve of Maspalomas, south of Gran Canaria Island, includes a beach of 6 km long with a mobile dune system of white sand and a small lagoon. For several decades, the dune system has suffered increasingly obvious degradation with a constant loss of sand, mainly due to urban touristic development processes and human impact, which has altered the dynamics of the wind and the dunes. It is estimated that around 45,000 m^3 of sand are lost each year that ends up at the bottom of the sea. For this reason, a bathymetric analysis is very important as local authorities are running an ambitious project based on the extraction of sand from a specific intertidal zone and its reintroduction in the system.

Finally, a case example is introduced to show the applicability of the bathymetric mapping in a specific task: to monitor the construction of a new port in Granadilla (Tenerife Island, Canary Islands). The Granadilla coast (Fig. 1), located in the South-East of the island of Tenerife, is an environment of high ecological value that was compromised due to the construction of the port, whose works began in 2009. To guarantee environmental conservation, the Granadilla Environmental Observatory Public Foundation (OAG) was created, which is in charge of monitoring this environment, not only through traditional in-situ methods, but also through the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3135462, IEEE Transactions on Geoscience and Remote Sensing

processing of satellite imagery, within the framework of the Port of Granadilla Environmental European Monitoring Program, established in 2010, in order to guarantee sustainable environmental quality, inside and outside the port, during and after its construction.

Fig. 1. Cabrera archipelago, Maspalomas and Granadilla geographic locations (Google©).

In this work, very high resolution WorldView-2/3 satellite imagery has been used. Data have a radiometric resolution of 11 bits and a spatial resolution of 1.8/1.3 m, at the nadir, for the 8 multispectral bands (0.400-1.040 µm) [36]. Specifically, the level 2 ortho-ready product has been selected. Fig. 2 shows the true color composite images of Cabrera, Maspalomas and Granadilla for a small subset of the scene. The Cabrera image was recorded by WV-2 on September 1, 2016 (10:29 hr. UTC) with an off-nadir view angle of 24.6º, while the Maspalomas image by WV-3 on May 22, 2018 (12:06 hr. UTC) with an off-nadir angle of 10.6º. For Granadilla, 2 images were used: WV-2 of December 1, 2011 (12:11 hr. UTC) with an off nadir angle of 3.6º, and WV-3 recorded on December 21, 2019 (12:02 hr UTC) with an off nadir angle of 16.5º (Fig. 2(c)).

For the systematic use of the WV-2/3 multispectral data, a thorough assessment of the combined atmospheric and sun glint correction procedures has been performed in the Natural Reserve of Maspalomas. The goal was to select the best methodology to improve the water leaving reflectance estimation in near-shore environments. Moreover, a complete multi-algorithm study to map bathymetry has been carried out in the Cabrera Island through the use of the multispectral high-resolution satellite data. In addition, a summary of the bathymetric maps derived by the best algorithms is also provided for the Maspalomas area. Finally, the optimal overall methodology has been applied to monitor the construction of the port of Granadilla.

In-situ field data, for the Maspalomas shallow coastal waters, was collected simultaneously to the WorldView-3 satellite over flight [11]. Specifically, 4 points were sampled (Fig. 2(b)) and their reflectance in the 350–2500 nm spectral range was recorded with the spectroradiometer ADS Fieldspec 3.

III. Remote Sensing Data Preprocessing

New remote sensing platforms with improved sensors can contribute to the generation of accurate information for the management of coastal areas and vulnerable ecosystems [11, 37, 38]. However, the radiation acquired by space sensors is affected by perturbations produced by the atmospheric scattering and absorption and by the sunglint effect on the sea surface. These unwanted contributions affect to the accurate extraction of information and, consequently, have to be properly addressed.

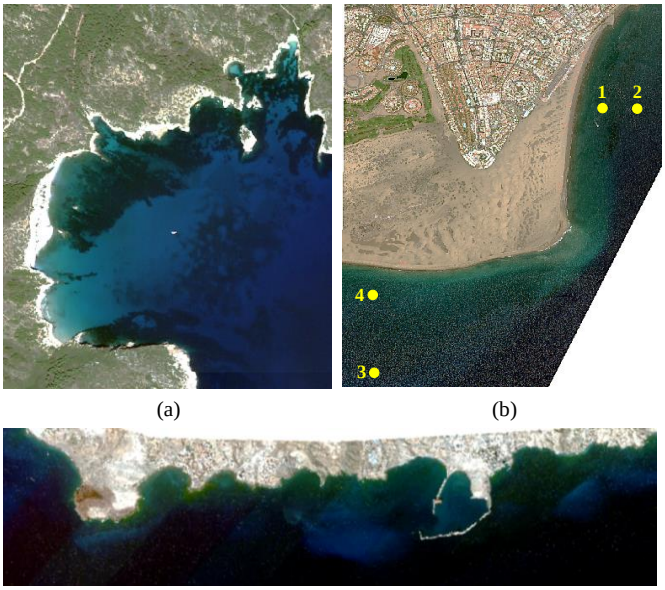


Fig. 2. Worldview-2/3 scenes used in the analysis: (a) Cabrera, 2016. (b) Maspalomas, 2018. (c) Granadilla, 2019.

Regarding the atmospheric correction, different approaches have been implemented. Some works considering WorldView-2 imagery have demonstrated that the Second Simulation of a Satellite Signal in the Solar Spectrum (6S) [39] has an excellent performance achieving low RMSE values in littoral zones [8-11].

The 6S is a radiative transfer model to predict the reflectance at the top of atmosphere (ρ_{TOA}). For an accurate atmospheric correction, some input parameters have to be properly set. Specifically, the type of satellite sensor, geographical coordinates and altitude of the area, the sun angles, the time and date of image acquisition, the atmosphere and aerosol models and the aerosol optical thickness have to be introduced. The Mid-Latitude Summer is the most suitable atmosphere model for the climate of the Balearic and Canary Islands for the sensing dates. The Maritime aerosol model was selected as island areas are considered in the study. The aerosol optical thickness parameter was properly adjusted using in-situ or satellite sensor information (i.e. MODIS at 550 nm), because major errors in their estimation can significantly affect the surface reflectivity computed. The contribution of adjacent pixels was considered to take into account the spatial mixing of radiance among nearby pixels caused by atmospheric scattering.

Next, the removal of sunglint is essential to retrieve bathymetry and the seabed mapping in shallow waters. Sunglint correction techniques have been proposed for open ocean and coastal applications [9, 40]. In this work, to remove the sunglint on the sea surface from the visible channels of the high-resolution imagery, a method combining physical principles and image processing techniques has been applied. According to the physical model proposed by Kay et al. [40], for a selected deep water area, WV-2/3 visible bands 1 to 5 are included in a

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3135462, IEEE Transactions on Geoscience and Remote Sensing

linear regression with respect to the WV-2/3 NIR bands 7 and 8 to calculate the slope b_λ. Pixels of the image scene are deglinted applying the following expression,

$$\rho_{VIS}^{deg}(\lambda) = \rho_{VIS}(\lambda) - b_\lambda(\rho_{NIR} - MIN\rho_{NIR}) \quad (1)$$

where $\rho_{VIS}^{deg}(\lambda)$ is the deglinted water surface reflectance of the visible channels, b_λ is the slope of the linear regression and ρ_{VIS} and ρ_{NIR} are the water leaving reflectance of the visible and near-infrared channel, respectively.

Apart from the physical deglinting approach, a histogram matching technique is applied to statistically equalize each channel of the deglinted image to match the water reflectivity prior to the deglinting process, as the presence of whitecaps or foam affects to the spectral content of the image altering the reflectance intensity.

After the glint is removed, depending on the state of the sea surface state, certain amount of noisy pixels can be expected. Therefore, a median filter is used for noise removal and to facilitate, for example, the classification of subsurface features.

After the evaluation and selection of the combined atmospheric-deglinting preprocessing tasks, to improve the water reflectance in near-shore environments, it is possible to obtain accurate bathymetric information using the reflectivity information from the multispectral corrected channels with greater penetration.

IV. Satellite Derived Bathymetry Mapping

With the increase of coastal population densities, marine ecosystems are under significant stress due to overfishing, pollution, and coastal development. Moreover, global environmental changes, such as sea-level rise and increases in ocean temperature and acidification, are applying additional pressure to littoral water ecosystems [7, 41]. New satellite sensors and techniques for coastal monitoring and protection are critically needed. In this context, the monitoring of multitemporal bathymetric information will enhance the understanding of some of the main climate change effects as, littoral erosion, benthic habitats resilience, sea-level variation, etc.

The systematic framework to derive an accurate model for high-resolution bathymetry mapping of shallow littoral waters is included in Fig. 3. The multi-algorithm analysis developed in this work exploits the spectral capabilities of the WorldView-2/3 multispectral imagery. Different Regression and Machine Learning Bathymetry Techniques are analyzed.

Two different analysis were performed to assess the accuracy and robustness of the different techniques when estimating bathymetry down to 2 challenging depths (25 m and 35 m). During the training and test of the bathymetric techniques, the bathymetry derived by a multibeam echosounder was used as a reference. Specifically, master isobaths every meter were chosen to extract a total of 67.660 and 82.640 samples for the 25 m and 35 m depth studies, respectively. For each meter of depth, 80% of the samples were used to train the models and the remaining 20% for their evaluation. To avoid geolocation errors, flat seafloor areas were chosen.

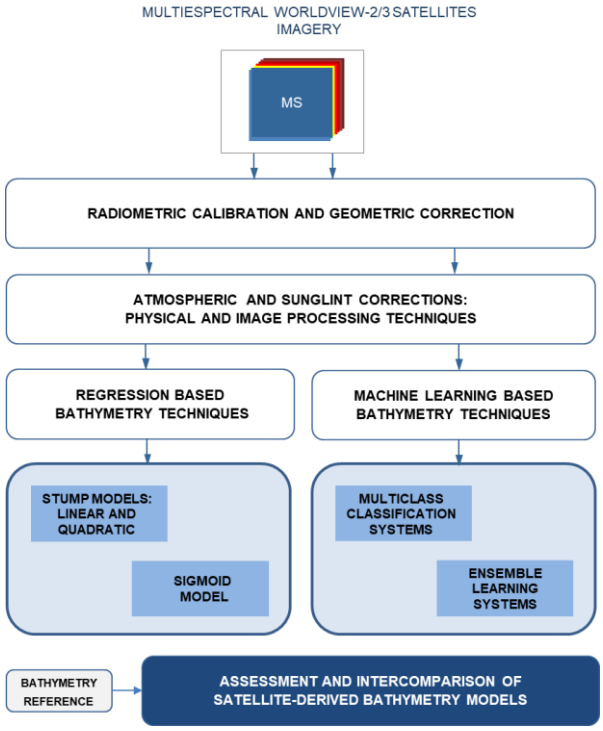


Fig. 3. Multi-algorithm diagram for the bathymetry-derived mapping of littoral areas with multispectral high-resolution satellite remote sensing imagery.

A. Regression-Based Techniques

Considering the optical properties of water and underwater reflectance and using the best WorldView-2/3 bands, empiric algorithms have been analyzed to extract bathymetry at different depth ranges.

The most relevant regression model in literature was proposed by Stumpf et al. [19]. This model is based on the exponential relationship between the attenuation coefficients of different bands and depth. Depth can be expressed by,

$$z = m_0 + m_1 \frac{\ln(n\,R(\lambda_i))}{\ln(n\,R(\lambda_j))} \quad (2)$$

where, m_0 is the offset, m_1 is the slope, $R(\lambda_{i,j})$ is the reflectivity in band$_{i,j}$ and n is an arbitrary multiplicative value so that values in the logarithms are positive. Different values of "n" are commonly used, such as 100 or 1000. In our case we choose 100 and results were satisfactory. The model (2), presented by Stumpf, indicates that there is independence with the type of seafloor. However, in practice, for large areas, there are different slopes and offsets in different types of bottoms, depending on the Field of View (FoV) of the satellite. In order to model this behavior, an improved Stumpf model is implemented through a quadratic regression, given by [42]:

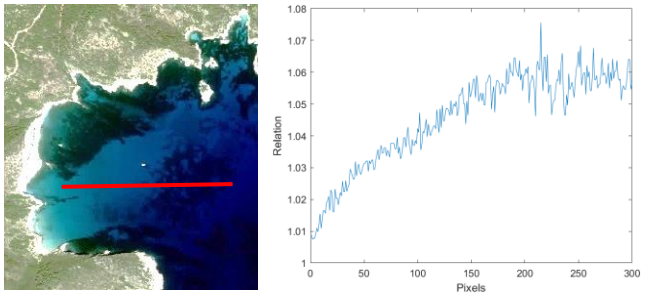
$$z = m_0 + m_1 \frac{\ln(n\,R(\lambda_i))}{\ln(n\,R(\lambda_j))} + m_2 \left(\frac{\ln(n\,R(\lambda_i))}{\ln(n\,R(\lambda_j))}\right)^2 \quad (3)$$

This quadratic function can be approximated to a linear function for low depth areas leading to an identical behavior to the linear Stumpf model.

Fig. 4 shows an example of the evolution of the Stumpf ratio with respect to bathymetry up to 25 m depth. The red line represents the trajectory sampled in the WV-2 image of Cabrera

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3135462, IEEE Transactions on Geoscience and Remote Sensing

island at increasing depths and the corresponding curve is plotted. Values of the ordinate axis correspond to the relationship $\frac{\ln(n\,R(\lambda_i))}{\ln(n\,R(\lambda_j))}$. In this example, the blue and green bands have been chosen. A saturation of the curve can be appreciated that will remain for deeper waters. Consequently, the linear and quadratic regressions do not properly model the Stumpf ratio in deeper areas. To solve this problem and to improve the robustness and accuracy, a new regression approach, called the Sigmoid-based Model, is proposed in this work.



(a) (b)

Fig. 4. Stumpf ratio variation with depth: (a) sampled transect (red line). (b) Stumpf ratio at the sampled points.

- *Sigmoid-based Model: A New Regression Approach*

As previously indicated, pixel saturation of deep areas produces a curve of values in relation to depth. A saturation in the curve is clearly observed for deep pixels. Taking advantage of the s-shaped curve of the Sigmoid function and its ability to model non-linearities, the following expression is proposed to estimate the bathymetry,

$$f = m_1\left(\frac{1}{1+e^{-m_0 \cdot z}} - \frac{1}{2}\right) + m_2 \tag{4}$$

where the constant $\left(-\frac{1}{2}\right)$ is chosen to center the sigmoid at zero, m_0 is the scale of the inputs, m_1 corresponds to the sigmoid scale, m_2 is the continuous component, z corresponds to the depth, and f is the Stumpf relation $\left(f = \frac{\ln(n\,R(\lambda_i))}{\ln(n\,R(\lambda_j))}\right)$. Therefore, depth can be formulated by:

$$z = -\frac{1}{m_0} ln\left(\frac{2\,m_1}{2\,f - 2\,m_2 + m_1} - 1\right) \tag{5}$$

This novel model using a sigmoid function does not present the problem of unwanted behavior in deeper areas or outside the training range, as occurs with polynomial regressions, since it properly models the saturation of deep pixels.

Although variable bottom-reflectance and variation in the optical properties of the water column impact the accurate depth extraction, using several optical bands increases the robustness of the depth estimation.

However, SDB based on multi-spectral sensors still presents problems for practical use. Previously used methods basically require supervised training depth data for each image, and bathymetry cannot be extracted only using the satellite image.

B. *Machine Learning-Based Techniques*

At present, Machine Learning techniques [12, 13, 43] represent a new paradigm to obtain high resolution satellite bathymetry. These methods allow, from some training data, obtaining a model that maps inputs to desired outputs. To accomplish this objective, a metric is iteratively minimized based on statistical optimization. Even though it is not simple to develop a model that relates water depth and the multispectral bands, under variable sensing conditions, machine learning can offer a good solution. In consequence, we have explored the effectiveness of machine learning algorithms for Satellite-Derived Bathymetry (SDB).

Machine Learning systems not only model the non-linear relationships of the transformed Stumpf space, but also allow obtaining bathymetric maps from the information included in all the bands of the original image. The latter makes it possible to improve the exploitation of the spectral information, at the cost of increasing the computational borden. In this work two families of Machine Learning techniques have been analyzed: Multiclass Classification and Ensemble Learning approaches.

- *Multiclass classification systems*

Multiclass systems classify the entries into a finite set of classes [44], which in this case corresponds to the estimation of the bathymetry map levels. The number of classes, according to the maximum depth (Max_{depth}), minimun depth (Min_{depth}), and bathymetry resolution ($Resolution_{depth}$) is:

$$\boldsymbol{Classes} = \left\lfloor \frac{\boldsymbol{Max_{depth} - Min_{depth}}}{\boldsymbol{Resolution_{depth}}} \right\rfloor \tag{6}$$

The selected multiclass classification systems implemented in this work are:

(i) Support Vector Machines [45] are based on searching the division of the set of inputs by means of a hyperplane, finding relationships between bands of the image. To accomplish this, the cost function is minimized in relation to the W parameters:

$$J(W) = -\frac{C}{m}\sum_{i=0}^{m} \left(z^{(i)}\, cost_0\left(w_0 + w_1 x_1^{(i)} + \cdots + w_N x_N^{(i)}\right) + \left(1 - z^{(i)}\right) cost_1\left(w_0 + w_1 x_1^{(i)} + \cdots + w_N x_N^{(i)}\right)\right) \tag{7}$$

where C is a scaling parameter, m is the number of samples, $x_N^{(i)}$ correspond to the input image channels and $z^{(i)}$ is the expected bathymetry.

SVM systems perform a binary classification. However, to address a multiclass classification and obtain bathymetry maps, SVM algorithms are applied iteratively until the target resolution is achieved. In addition, to perform a non-linear classification, a kernel can be applied on the data set to transform a non-linear feature space into a linear space, depending on the nature of the data and the kernel used. The Gaussian kernel produces good results [46]:

$$\boldsymbol{k(x)} = \boldsymbol{e}^{-\frac{\left\|x - r_i\right\|^2}{2\sigma^2}} \tag{8}$$

where $k(x)$ is the linearized value for each point x, r_i are the reference points and σ^2 is the variance of the Gaussian.

(ii) K-Nearest Neighbors [47] are based on the assignment of a point in the feature space to a class (depth) using the Euclidean distance:

$$d(x, x_{ref}) = \sqrt{\sum_{i=1}^{p}(x_i - r_i)^2} \quad (9)$$

where x is the analyzed point, r_i is the reference point and p is the dimension of the feature space.

In estimation, the input is classified as the class, or depth, of greatest presence between the nearest K points.

(iii) Decision Trees [32] are based on establishing successive conditionals to classify an entry. Each conditional corresponds to a node in the tree, where the data set is divided into two subsets. The resulting tree must have the necessary bifurcations to have as many final classes as depth levels on the bathymetric map. Decision Tree training is based on the homogeneity of the resulting group at each branch. To quantify homogeneity, the Gini Index [48] is applied, which is given by the expression:

$$Gini = 1 - \sum_{i=1}^{n} p_i^2 \quad (10)$$

where n refers to the number of classes (depths) and p_i indicates the probability of a class i in the data subset at a given node.

In this research, the following Multiclass Classification algorithms have been implemented and assessed: linear SVM, SVM with Gaussian kernel, KNN and Decision Tree.

- *Ensemble Learning systems*

Systems that apply Ensemble Learning techniques [33] are based on the idea that a joint election produces better results than individual elections, obtaining, as demonstrated in Section V, bathymetric maps of greater accuracy.

Starting from a data set, N data subsets are generated by sampling, which are fed during the training phase to N classification models with the same initial configuration. Once the N models have been trained, N possible outputs are estimated when faced with a new entry. The output of the overall system is obtained based on a decision metric taking into account the individual outputs. The final estimated depth is, therefore, given by:

$$\hat{z} = \max_c(\bar{z}) \quad (11)$$

where $\bar{z}$ is the set of depth estimates for each classification model and $\max_c$ is the function that obtains the class with the highest presence among the outputs of each model.

The Machine Learning models implemented in this work, which apply Ensemble Learning techniques to improve water depth estimation, are: Bagging Tree and Subspace KNN.

(i) Bagging Tree [34] applies the Bootstrap Aggregating (Bagging) technique [49] to Decision Tree classification systems. In Bagging models, a uniform sampling with replacement is performed, so there will be repeated elements between subsets.

(ii) Subspace KNN [35] applies the Random Subspace technique [50] to KNN classification systems. In the Random Subspace, uniform sampling is performed without replacement, so the intersections of the generated subsets are the empty set, that is, there are no repeated data.

It is worth mentioning that bathymetric maps must be continuous, as in regression models, since depth evolves in nature in this way. However, applying Machine Learning techniques, discrete maps are obtained, with as many classes as defined according to Equation (6). In this work, the $Resolution_{depth}$ selected was 1 m. Therefore, to transform the bathymetry map obtained into a continuous map, a Gaussian filter with a standard deviation of 0.5 was applied.

V. Results

This section includes the assessment of the pre-processing corrections and the detailed analysis of the bathymetric mapping applying the previous 9 techniques on very high resolution multispectral imagery.

A. Assessment of Combined Atmospheric-Deglinting Preprocessing

A comprehensive assessment of the preprocessing steps was performed to the Maspalomas shallow coastal waters using the Worldview-2/3 imagery. As indicated in Section II, this work implements a combined 6S atmospheric correction model with deglinting techniques. A preliminary analysis of the 6S parameterization was performed and, basically, changes in the different input parameters only led to slight variations in the estimated reflectance.

Next, the estimated reflectance of the atmospherically corrected image was compared with the real reflectance measured by the ADS Fieldspec 3 spectroradiometer in representative coastal sites [51]. The result of combining the 6S atmospheric correction model, properly parameterized, a simple correction of the solar reflection, based on the physical algorithm of Kay et al. [40], and a subsequent histogram adjustment, according to the Moorea model [52], provides optimal results, as shown in Fig. 5, for the 4 strategically selected coastal water points (Fig. 2). As presented in Table I, the appropriate combination of the 6S model and the improved deglinting technique achieves reflectivity values very close to those measured in-situ, with approximate RMSE of 0.5%.

The outstanding results of this important stage of the WV-2/3 data preprocessing, are as well applicable to the coastal areas of Cabrera and Granadilla, where imagery with suitable sea state conditions was selected.

TABLE I
STATISTICS RESULTS FOR THE PREPROCESSED IMAGES WITH RESPECT TO THE IN-SITU MEASUREMENTS AND CORRECTED SUNGLINT (HEDLEY METHOD).

Atmospheric Model	Deglinting	RMSE ↓	BIAS ↓
6S	No correction	0,02580	-0,016859
	Hedley + Post-processing	**0.00508**	**0.000638**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3135462, IEEE Transactions on Geoscience and Remote Sensing

Some results of the complete preprocessing methodology applied to the Maspalomas and Cabrera littoral zones are shown in Fig.6. The removal of the sunglint effects are clearly seen in both areas and, specially, for Maspalomas where pronounced glint is present on the sea surface.

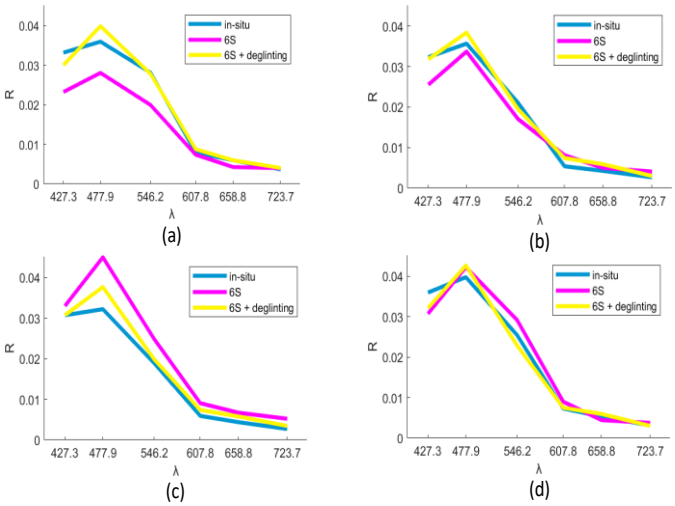


Fig. 5. Spectral reflectivity signatures with the appropriate combination of the 6S model and the improved deglinting technique in the 4 coastal shallow water sites sampled (see Fig. 2 (b)): (a) Point 1. (b) Point 2. (c) Point 3. (d) Point 4.

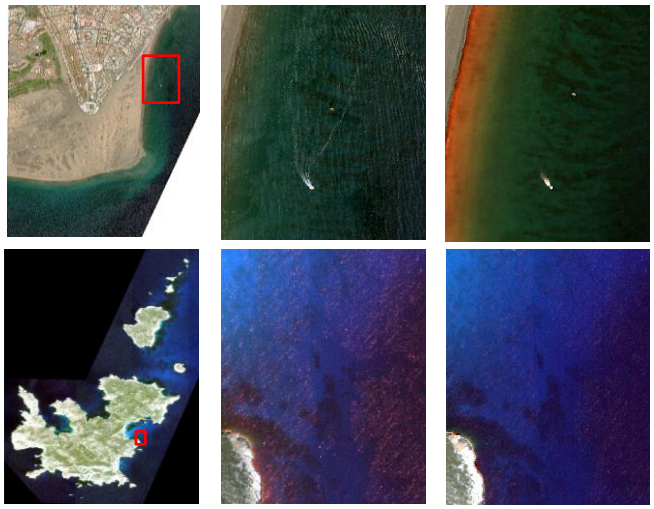


Fig. 6. WorldView-2/3 images of Maspalomas (up) and Cabrera (down): (a) true color composite. (b) Images after the 6S correction. (c) Images after the deglinting process.

B. *Assessment of Satellite-Derived Bathymetry Models*

In this section, the detailed results of the selected algorithms for coastal bathymetry mapping in Cabrera Island are presented. In addition, a summary of the main results for the Maspalomas area is included, allowing the robust mapping of water depth in littoral waters.

The main goal of this work is to assess the performance of 9 techniques to map bathymetry in complex shallow water areas. For this reason, 2 case studies have been selected. In the first case, a maximum depth of 25 m has been used to train and test the different approaches, while in the second study, depths up to 35 m were considered to assess the robustness in such a challenging scenario.

This multi-algorithm analysis is divided into, on the one hand, the study of regression-based models and machine learning-based models, and, on the other hand, the comparative study of both approaches.

In each case, for the qualitative analysis, the bathymetry obtained by each model is shown and can be compared to the reference bathymetry. For the quantitative analysis, the values of the RMSE and R^2 are presented. In addition, the RMSE value is analyzed for each meter of depth, and, finally, the curve that relates the estimated versus the expected depth is presented for a specific transect.

In particular, the results for the subscene of Fig. 7(a) are presented. It corresponds to an atmospherically corrected and deglinted WV-2 multispectral image from a selected beach of Cabrera Island.

- *Evaluation of Regression-based Bathymetry Models*

Two different experiments were carried out using training data limited to values up to 25 and 35 meters. The aim was to evaluate the precision and robustness of the existing and proposed models in complex scenarios reaching the limits of radiation penetration capability: (i) typical coastal areas (25 m) and (ii) areas with calm and transparent waters (35 m).

After a preliminary analysis, blue and green bands were selected for the regression models. Fig. 7 shows the bathymetry obtained by each technique using the same training data. Left and right columns show the SDB maps obtained up to 25 m and 35 m, respectively. As indicated, each model has been trained in 2 different ways, with depth data up to 25 m and up to 35 m, and then bathymetric maps have been generated for the whole scene. To facilitate the visual comparison in both situations, a single colormap has been chosen.

In Fig. 7 (a), the RGB composite is presented while in Fig. 7 (b) the reference bathymetry is provided. Figs. 7 (c-f) show the SDB results obtained by Stumpf, Quadratic Stumpf~~, RTE~~ and Sigmoid models, respectively. The visual analysis indicates that the different models provide similar results.

On the other hand, Fig. 8 shows the fitting of the different regression models with respect to the training data for both depths (25 and 35 m). A considerable dispersion of the training data is observed, especially at shallow depths. That is, in the relation of 2 band reflectivities for each meter of considered depth. Fig. 8 (a) shows the Stumpf linear model, which correctly fits the data but only in the linear regime. In Fig. 8 (b), the Stumpf model with quadratic regression is presented. In this case, the model, likewise, is not appropriate at water depths over 25 m, where the curve saturates. Finally, Fig. 8 (c) shows the results for the proposed Sigmoid model, where the curve perfectly fits the data in both cases, properly modeling the non-linearity, even at high depths. Therefore, when a more challenging scenario is considered, only the Sigmoid model is able to properly characterize the saturated values.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3135462, IEEE Transactions on Geoscience and Remote Sensing

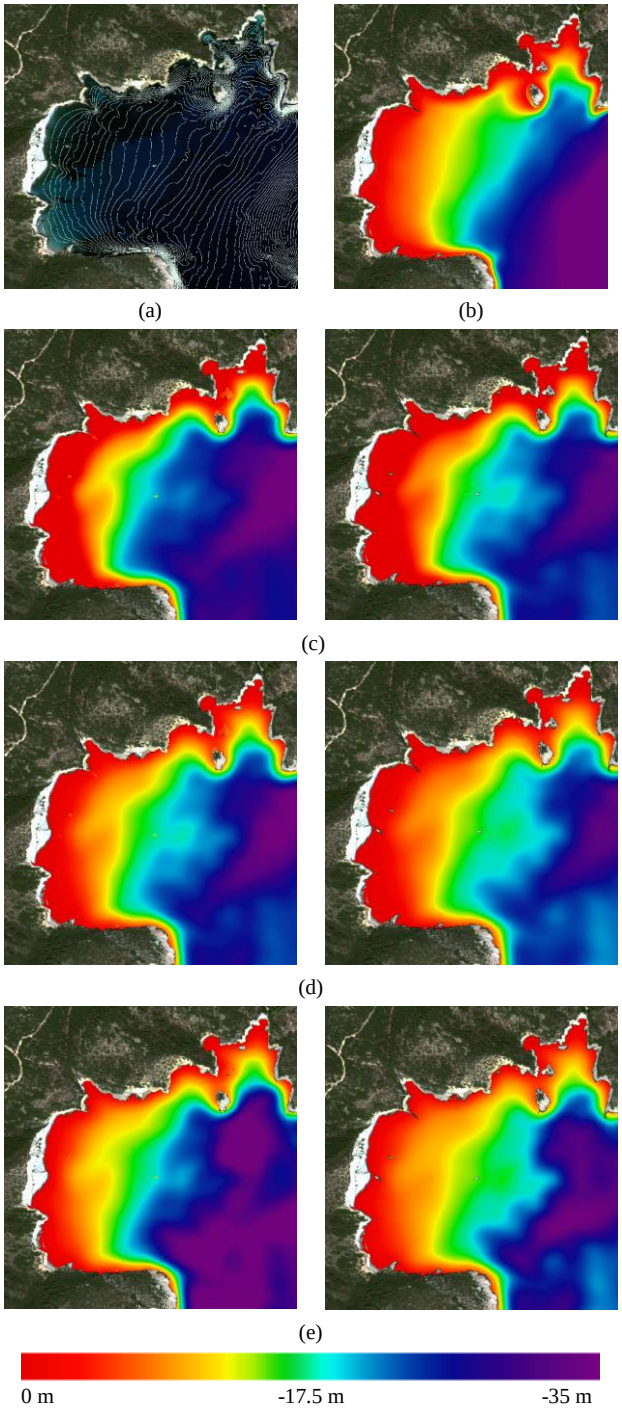


Fig. 7. SDB of the regression models for 25 m (left) and 35 m (right) depth: (a) RGB WorldView-2 preprocesed image. (b) Reference bathymetry. (c) Stumpf. (d) Stumpf quadratic. (e) Sigmoid.

Quantitatively, Table II shows the RMSE and R^2 values in both cases. Note that the study has been carried out up to high depths and that the dispersion of the training data is large, which results in higher values of the measured errors with respect to the reference bathymetry. Despite the differences, all the models present acceptable metric values. In particular, Sigmoid model have the lowest RMSE and both approaches achieve the highest R^2 value, obtaining superior bathymetric maps. The quadratic Stumpf model obtains the worse statistics and, therefore, the linear model would be preferably.

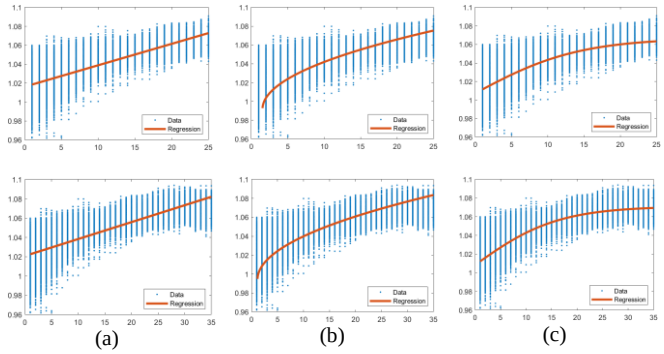


Fig. 8. Regression models fit using training data with maximum depth of 25 m (upper row) or 35 m (lower row): (a) Stumpf (linear adjustment). (b) Quadratic (order 2 polynomial fit). (c) Sigmoid Model proposed in this work (modified sigmoid fit).

TABLE II
STATISTIC RESULTS OF THE REGRESSION-BASED BATHYMETRY MODELS RESPECT TO THE REFERENCE BATHYMETRY (BEST RESULTS IN BOLD).

Method	RMSE (m) ↓		R^2 ↑	
	25 m	35 m	25 m	35 m
Stumpf Linear	3.63	5.01	0.744	0.770
Stumpf Quadratic	4.20	5.33	0.693	0.718
Sigmoid	**3.43**	**4.36**	**0.800**	**0.823**

As a conclusion, it can be pointed out that, within the empirical-based category, the Sigmoid model is the appropriate choice as it is a simple approach that provides excellent performance. Sigmoid technique can model the non-linearity produced by deep pixels, allowing them to fit the training data without the need for depth limitation, as it is necessary in the Stumpf model with linear and quadratic regressions. However, by limiting the maximum depth to lower values, where the energy can properly reach the seabed, all models present a similar behavior, as shown in Figure 8.

- *Evaluation of Machine Learning-based Bathymetry Models*

High-resolution bathymetry maps obtained for the two groups of implemented Machine Learning techniques (Multiclass Classification and Ensemble Learning) are presented in Fig. 9 for both cases (25 and 35 m). Fig. 9 (a) shows the RGB composite of the WV-2 multispectral image atmospherically corrected and without sunglint, including the 1 meter isobaths lines. The reference bathymetry map is shown in Fig. 9 (b). Figs. 9 (c-h) show the results of the machine learning models obtained by Support Vector Machine, SVM with Gaussian kernel, Decision Tree, K-Nearest Neighbors, Subspace KNN and Bagged Tree models, respectively. Clearly, Decision Tree provides the worst map when compared to the reference in Fig. 9 (b). The rest of models properly estimate bathymetry, although it should be noted that the ensemble models provide excellent results regardless of the water depth and the seafloor topology.

Comparing both experiments, as expected, better estimations are produced for deep waters when models are trained up to 35

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3135462, IEEE Transactions on Geoscience and Remote Sensing

m. In fact, maps trained up to 25 m (left column) cannot produce accurate maps in very deep waters.

Table III presents the RMSE, and R^2 values for each Machine Learning based model implemented. Accuracies are high, even in the challenging situation of mapping up to 35 m depth. As expected, Decision Tree is the worst technique, while models implementing Ensemble Learning have the lowest RMSE errors and the highest R^2 coefficients. Both models can be chosen to accurately derive high resolution bathymetric maps.

Based on the previous results, Ensemble Learning techniques improve mapping when compared to single models, thanks to the decrease in the variance of the data. This phenomenon is observed between KNN and Subspace KNN models, where the method implemented by Ensemble Learning has a reduction in the RMSE error of 0.255 m (25 m) and 0.279 (35 m).

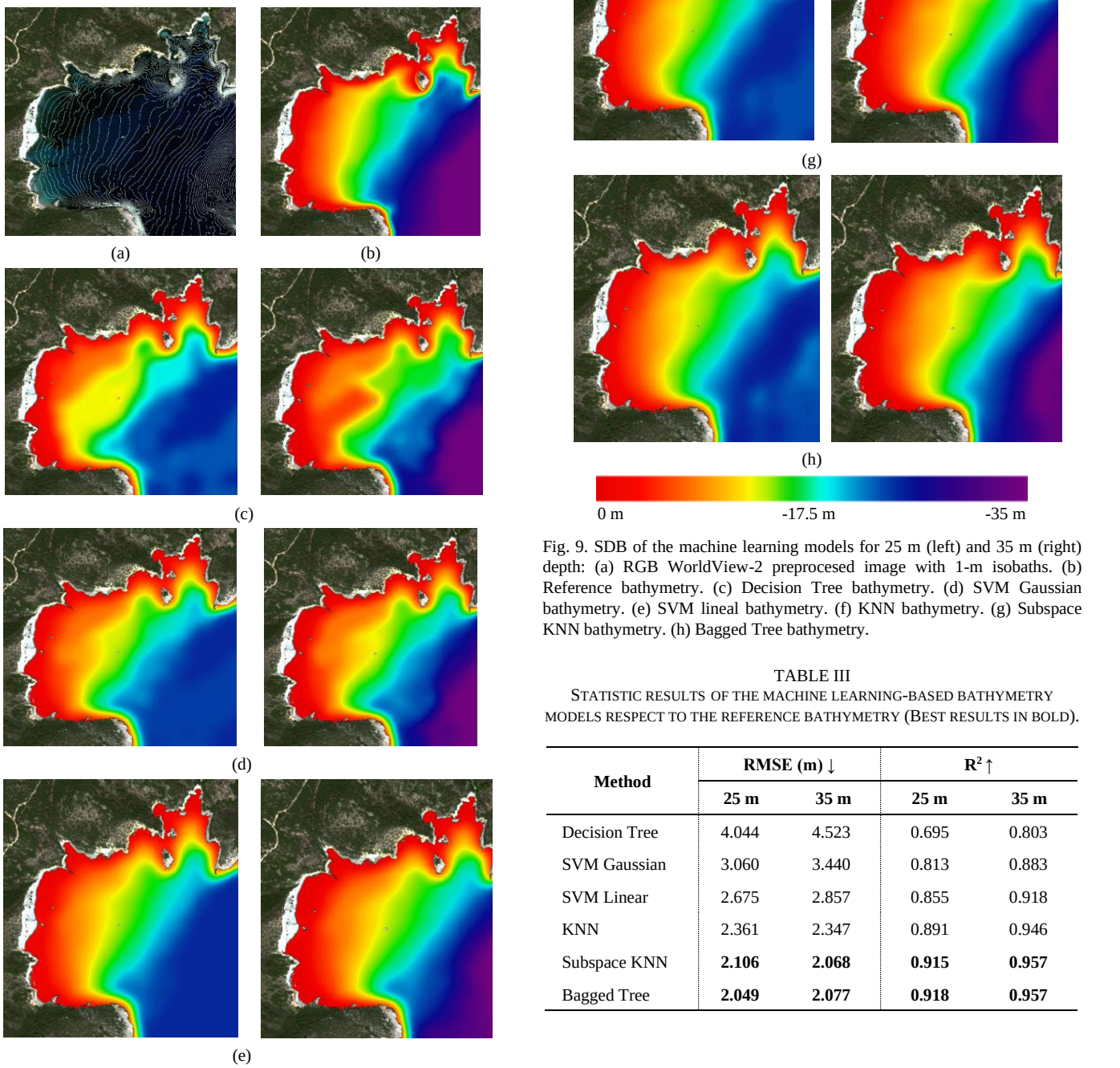


Fig. 9. SDB of the machine learning models for 25 m (left) and 35 m (right) depth: (a) RGB WorldView-2 preprocesed image with 1-m isobaths. (b) Reference bathymetry. (c) Decision Tree bathymetry. (d) SVM Gaussian bathymetry. (e) SVM lineal bathymetry. (f) KNN bathymetry. (g) Subspace KNN bathymetry. (h) Bagged Tree bathymetry.

TABLE III
STATISTIC RESULTS OF THE MACHINE LEARNING-BASED BATHYMETRY MODELS RESPECT TO THE REFERENCE BATHYMETRY (BEST RESULTS IN BOLD).

Method	RMSE (m) ↓		R^2 ↑	
	25 m	35 m	25 m	35 m
Decision Tree	4.044	4.523	0.695	0.803
SVM Gaussian	3.060	3.440	0.813	0.883
SVM Linear	2.675	2.857	0.855	0.918
KNN	2.361	2.347	0.891	0.946
Subspace KNN	**2.106**	**2.068**	**0.915**	**0.957**
Bagged Tree	**2.049**	**2.077**	**0.918**	**0.957**

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3135462, IEEE Transactions on Geoscience and Remote Sensing

- *Regression-based versus Machine Learning-based Bathymetry Models*

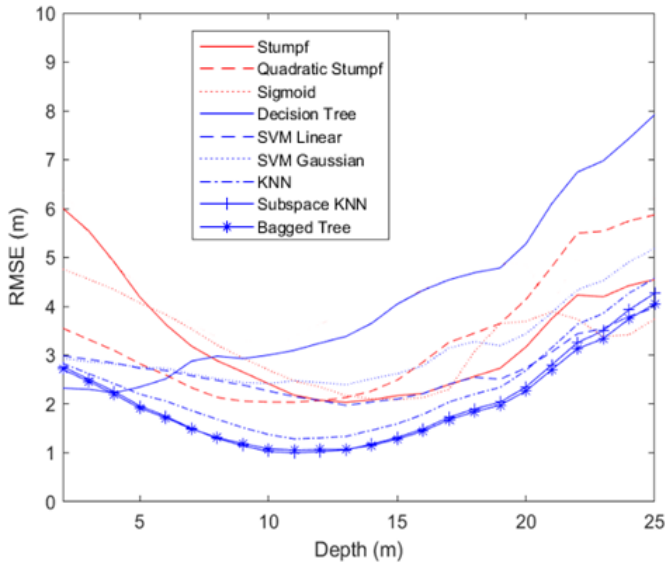
To perform a detailed joint analysis of both family of models, the RMSE value computed at each meter has been obtained and plotted in Fig. 10. Fig. 10 (a) shows the results when models are trained up to 25 m and Fig. 10 (b) up to 35 m. In general, the best accuracies are achieved for mid-depth water (10 to 15 m range). Clearly, Machine Learning models (blue) outperform regression models (red), except for the Decision Tree method. regression models cannot provide satisfactory results in very deep waters (over 28 m deep) where the penetration capability of the spectral channel is insufficient. However, the Sigmoid technique stands out for its good performance at medium-high depths, providing errors similar to some Machine Learning approaches (see Fig. 10(a)). From the regression models, the quadratic approach shows fairly low errors at low-medium depths. On the other hand, Ensemble models provide excellent results with errors below 3 m for most of the water depths, even in very deep waters. The minimum RMSE appears between 10 and 15 m and at higher depths, errors slightly increase with depth. Among the models that do not implement ensemble techniques, KNN achieves the best performance and follows the same error pattern as the ensemble methods. Finally, note the evident improvement in accuracy at high depths achieved by machine learning models when trained with samples up to 35 m. Thus, comparing the RMSE errors obtained at depths around 25 m in both figures, a clear improvement in performance is appreciated using more training information around deep waters.

Fig. 11 plots the estimated depth for a representative transect section covering different types of benthic classes and reaching 25 and 35 m deep. The figure includes the measured (reference) depth and bathymetry predicted by the 9 algorithms. We can appreciate that regression models are, in a greater manner, affected by the type of seafloor substrate, showing higher errors in sandy areas located in the middle of the transect. Machine learning techniques accurately map the water depth all along the transect pixels.

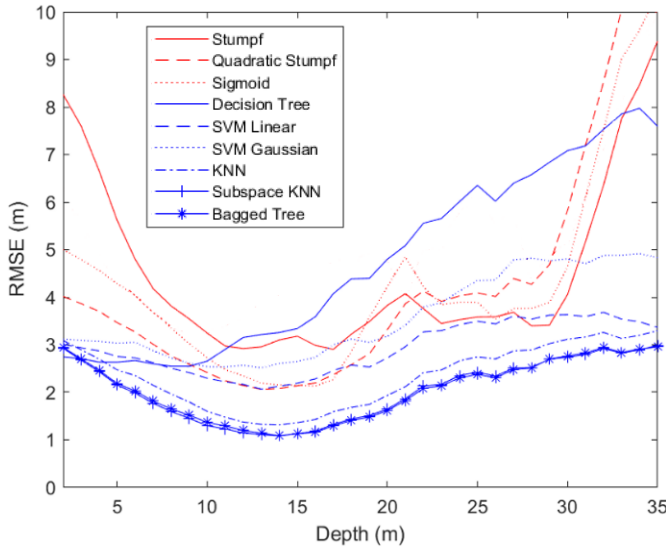
Concerning the robustness and transferability of the different algorithms is a critical issue still unresolved and that needs future research. It is necessary to take into account the great complexity of coastal zones, their heterogeneity and their great variability in the concentrations of the different parameters of the water column. This fact, together with the unwanted effects caused by the sea surface in swell conditions (sunglint), forces the choice of the appropriate image and the adjustment of the parameters of the models to each area of interest.

- *Satellite-Derived Bathymetry Maps for Maspalomas*

The methodology of Fig. 3 has been applied to the Maspalomas area. This coastal zone has a steep depth gradient in some regions, so that, a few meters from the coast, the bathymetry drops abruptly. This insular zone is characterized by the usual presence of wind and strong currents, which makes difficult to estimate the bathymetry or extract the seabed mapping using high resolution remote sensing data. In fact, as



(a)



(b)

Fig. 10. Comparative RMSE values for the analysis up to: (a) 25 and, (b) 35 m.

shown in Fig. 6, a considerable sunglint appears in the images which makes pre-processing tasks critical in this area to obtain adequate results.

Fig. 12 shows the bathymetric maps obtained by the most robust SDB methods of each family (regression and machine learning models). This particular zone has been selected because, as indicated in Section II, it is highly dynamic and local authorities are running a project to extract sand from the sink area marked in Fig. 12 (a) and to reintroduce it. For this reason, periodic monitoring is necessary to estimate the amount of sand accumulated near the shore.

After visual comparison with the reference map in Fig. 12 (b), it can be seen that both algorithms perform satisfactorily although, as expected, the Bagged Tree, as it applies ensemble learning techniques, obtains exceptional results, very close to the measured depth data.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TGRS.2021.3135462, IEEE Transactions on Geoscience and Remote Sensing

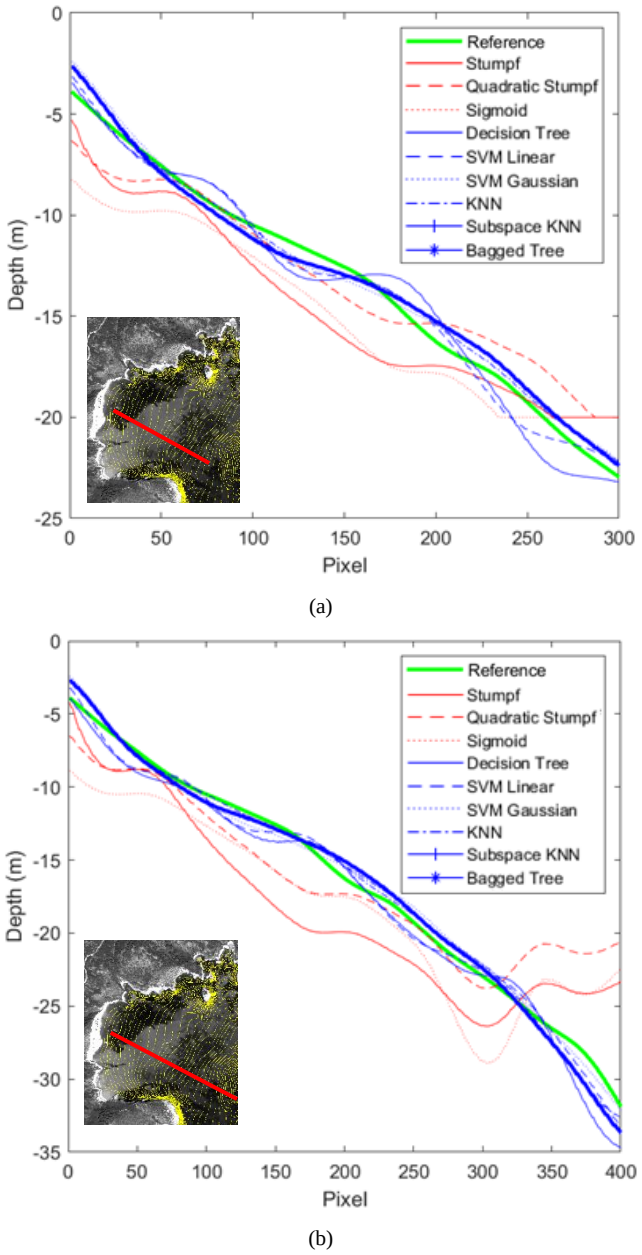


Fig. 11. Depth with respect to the distance from the shore (in pixels) for the 10 models analyzed (the reference bathymetry is included in green) trained up to: (a) 25 m and, (b) 35 m.

VI. Conclusions

Bathymetry is considered a key information for hydrological engineering and coastal applications. Traditionally, it is obtained via costly and time-consuming shipboard echosounders campaigns. However, nowadays, remote sensing imagery can be a wide-coverage, low-cost and fast solution for the bathymetric mapping.

In this research, a new methodology is developed to estimate bathymetry using very high-resolution multispectral Worldview-2/3 imagery, but it is applicable to similar satellites. An improved atmospheric and sun-glint correction method is presented and validated using the real reflectance measured by a field spectroradiometer. Mean RMSE values about 0.5% are achieved in representative coastal sites.

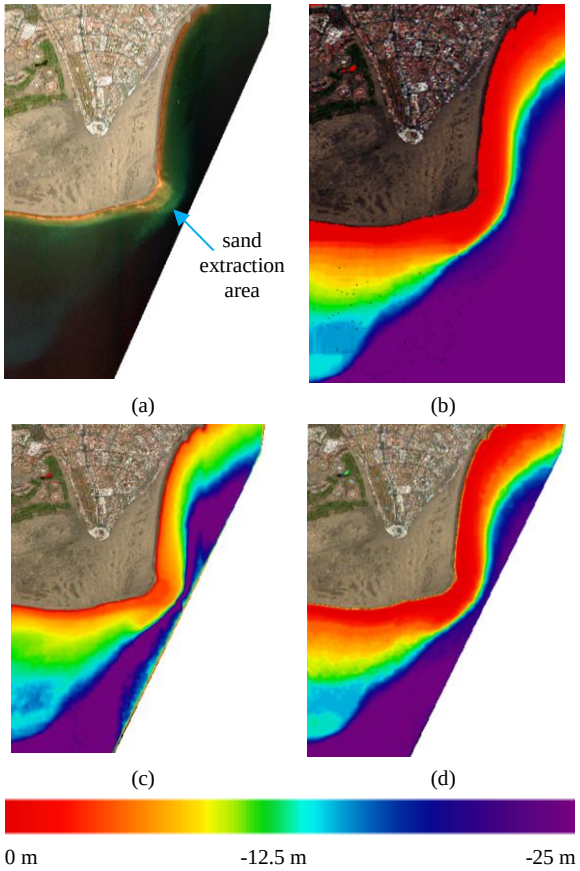


Fig. 12. Bathymetry of the Natural Reserve of Maspalomas: (a) Original image with 6S atmospheric correction and sunglint correction. (b) Reference bathymetry. (c) Bathymetry with the Sigmoid model. (d) Bathymetry with the Bagged Tree model.

A. Mapping Application: Monitoring of the Granadilla Port

Finally, the developed methodology (Fig. 3) has been used, operationally, in the context of the Multitemporal European Monitoring Program for the Port of Granadilla [53]. The research is linked to the monitoring of the bathymetry during the construction of the commercial port, and afterwards, within the framework of compliance with the evaluation of the conservation status of the species and habitats included in the European Habitat Directive (Directive 2008/105/CE of the Parliament and European Council). To that respect, Fig. 13 includes two examples of high resolution WV-2/3 bathymetric maps used for monitoring the Granadilla Port environment in different dates (2011 and 2019). Analogously to the results obtained in the other sites (Cabrera and Maspalomas), the Bagged Tree bathymetry technique obtains exceptional results.

On the other hand, to select the most suitable algorithm to derive the bathymetric maps, a thorough assessment of 9 Regression and Machine Learning models (linear Stumpf, quadratic Stumpf, Sigmoid, linear SVM, Gaussian SVM, KNN, Decision Trees, Bagging Tree and Subspace KNN) has been carried out. Note that, to our knowledge, this is the first time that a comprehensive evaluation of such a large number of techniques is presented. Additionally, a new regression model is proposed, and specific machine learning techniques are newly applied to derive bathymetry. Furthermore, challenging

scenarios have been considered, reaching water depths up to 25 m and 35 m, unlike the large number of works that only address very shallow depths [54].

Satellite-derived maps were compared to echosounder depth data, and it was demonstrated that machine learning models achieved superior robustness and performance than regression models, providing lower RMSE errors and higher R^2 coefficients. In particular, ensemble learning approaches revealed an outstanding performance in coastal areas down to 35 m depth, with mean RMSE values around 2 and R^2 of 0.957.

The excellent results provided for the proposed high resolution bathymetry mapping methodology has provided a systematic framework for the monitoring of coastal areas in Balearic and Canary Islands and, specifically, in the frame of the multitemporal European Monitoring Program for the new Port of Granadilla.

Therefore, Machine Learning models is one of the emerging areas in SDB studies and has a vast potential to develop operational SDB models but further research is required to discover solutions for sites having highly turbid areas, or variable bottom types using all contemporary available data and techniques. In summary, an assessment of the potential and limitations of these different methods over different coastal areas must be made. Also, comparison between these methods over the same study sites in a multitemporal manner is, as well, essential to obtain a concise knowledge about the robustness and difference in their performances.

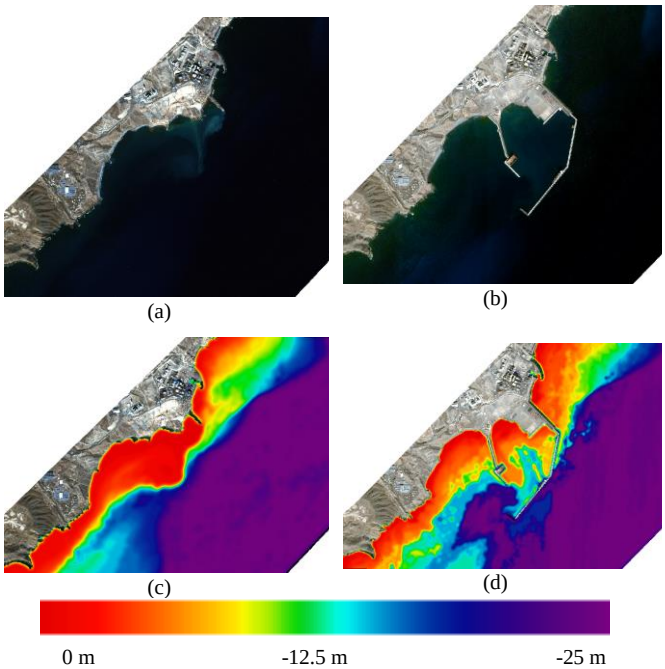


Fig. 13. High Resolution Satellite Bathymetry Mapping obtained by the machine learning developed methodology in the framework of Port of Granadilla Environmental Monitoring Programme: (a, b) WV-2/3 preprocessed imagery: December, 2011 and December 2019. (c, d) Bathymetry obtained with the Bagged Tree model (construction of a new port in the period 2010-2020).

ACKNOWLEDGMENT

This work has been supported by the European Project MAC-CLIMA and the ARTEMISAT-2 (CTM2016-77733-R) project, funded by the Spanish Agencia Estatal de Investigación (AEI), and the European Regional Development Funds (ERDF). Also, this work has been supported by the Observatorio Ambiental de Granadilla (OAG), Contract SERV-CM-1/202.

REFERENCES

[1] IHO, “International Hydrographic Organization Standards for Hydrographic Surveys S-44 Edition 6,” International Hydrographic Organization, 2020. www.iho.int (accessed February, 2021).

[2] Europarc-España. Anuario 2013 del estado de las áreas protegidas en España. Ed. Fundación Fernando González Bernáldez. Madrid, 2014.

[3] L. Rossi, I. Mammi I, and F. Pelliccia, “UAV-Derived Multispectral Bathymetry,” *Remote Sensing*, vol. 12, no. 23, pp. 3897, 2020

[4] S. Purkis, and V, Klemas, Remote Sensing and Global Environmental Change, John Wiley & Sons Ltd.: Oxford, UK, 2011.

[5] E. Horning, J. Robinson, E. Sterling, W. Turner, and S. Spector, Remote Sensing for Ecology and Conservation; Oxford University Press: New York, USA, 2010.

[6] J. A. Richards, Remote Sensing Digital Image Analysis, Springer-Verlag, Berlin, 2013.

[7] Y. Wang, Remote Sensing of Coastal Environments, Taylor and Francis Series; CRC Press: Boca Raton, FL, USA, 2010.

[8] F. Eugenio, J. Marcello, and J. Martin, “High-Resolution Maps of Bathymetry and Benthic Habitats in Shallow-Water Environments Using Multispectral Remote Sensing Imagery,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 53, no. 7, pp. 3539-354, 2015.

[9] J. Martin, F. Eugenio, J. Marcello, and A, Medina, “Automatic sunglint removal of multispectral high-resolution Worldview-2 imagery for retrieving coastal shallow water parameters,” *Remote Sensing*, vol. 8, no. 37, 2016.

[10] F. Eugenio, J. Marcello, J. Martin, and D. Rodríguez-Esparragón, “Benthic habitat mapping using multispectral high-resolution imagery: evaluation of shallow water atmospheric correction techniques,” *Sensors*, vol. 17, 2639, 2017.

[11] J. Marcello, F. Eugenio, C. Gonzalo-Martín, D. Rodríguez-Esparragón and F. Marqués, "Advanced Processing of Multiplatform Remote Sensing Imagery for the Monitoring of Coastal and Mountain Ecosystems," *IEEE Access*, vol. 9, pp. 6536-6549, 2021.

[12] S. Jawak, S. Vadlamani, and A. Luis, “A Synoptic Review on Deriving Bathymetry Information Using Remote Sensing Technologies: Models, Methods and Comparisons,” *Advances in Remote Sensing*, vol. 4, pp. 147-162, 2015.

[13] Salameh, E.; Frappart, F.; Almar, R.; Baptista, P.; Heygster, G.; Lubac, B.; Raucoules, D.; Almeida, L.P.; Bergsma, E.W.J.; Capo, S.; De Michele, M.; Idier, D.; Li, Z.; Marieu, V.; Poupardin, A.; Silva, P.A.; Turki, I.; Laignel, B., “Monitoring Beach Topography and Nearshore Bathymetry Using Spaceborne Remote Sensing: A Review,” *Remote Sensing*, vol. 11, 2212, 2019.

[14] Yue Ma, Nan Xu, Zhen Liu, Bisheng Yang, Fanlin Yang, Xiao Hua Wang, Song Li, “Satellite-derived bathymetry using the ICESat-2 lidar and Sentinel-2 imagery datasets,” *Remote Sensing of Environment*, vol. 250, 112047, ISSN 0034-4257, 2020.

[15] Mohammad Ashphaq, Pankaj K. Srivastava, D. Mitra, “Review of near-shore satellite derived bathymetry: classification and account of five decades of coastal bathymetry research,” *Journal of Ocean Engineering and Science*, 2021.

[16] D. Lyzenga, "Passive remote sensing techniques for mapping water depth and bottom features," *Applied Optics*, vol. 17, 379-383, 1978.

[17] D. R. Lyzenga, N. P. Malinas, and F. J. Tanis, "Multispectral bathymetry using a simple physically based algorithm," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 44, no. 8, pp. 2251-2259, Aug. 2006.

[18] Z.P. Lee, K.L. Carder, C.D. Mobley, R.G. Steward, and J.S. Patch, "Hyperspectral remote sensing for shallow waters. 2. Deriving bottom depths and water properties by optimization," *Applied Optics*, vol. 38, pp. 3831–3843, 1999.

[19] R. P. Stumpf, K. Holderied, and M. Sinclair, "Determination of water depth with high-resolution satellite imagery over variable bottom types," *Limnol. Oceanogr.*, vol. 48, no. 1 II, pp. 547–556, 2003.

[20] J. Kibele, and N. T. Shears, "Nonparametric empirical depth regression for bathymetric mapping in coastal waters," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 9, no. 11, pp. 5130 – 5138, 2016.

[21] A. Kanno, Y. Koibuchi, and M. Isobe, "Shallow Water Bathymetry from Multispectral Satellite Images: Extensions of Lyzenga's Method for Improving Accuracy," *Coastal Engineering Journal*, vol. 53, no. 4, pp. 431-450, 2011.

[22] T. M. Mitchell, Machine learning. McGraw-Hill Series in Computer Science, 1997.

[23] L. Wang, H. Liu, H. Su, and J. Wang, "Bathymetry retrieval from optical images with spatially distributed support vector machines", *GIScience & Remote Sensing*, vol. 56, no. 3, pp. 323-337, 2019.

[24] A. Misra, Z. Vojinovic, B. Ramakrishnan, A. Luijendijk, and R. Ranasinghe, "Shallow water bathymetry mapping using Support Vector Machine (SVM) technique and multispectral imagery," *International Journal of Remote Sensing*, vol. 39, no. 13, pp. 4431-4450, 2018.

[25] F. Tonion, F. Pirotti, G. Faina, and D. Paltrinieri, "A Machine Learning Approach to Multispectral Satellite Derived Bathymetry", *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.*, vol. 3, pp. 565–570, 2020.

[26] T. Sagawa, Y. Yamashita, T. Okumura, T. Yamanokuchi, "Satellite Derived Bathymetry Using Machine Learning and Multi-Temporal Satellite Images," *Remote Sensing*, vol. 11, no. 10, 1155, 2019.

[27] M. D. M. Manessa et al., "Satellite-derived bathymetry using random forest algorithm and Worldview-2 imagery," *Geoplanning: Journal of Geomatics and Planning*, vol. 3, no. 2, pp. 117-126, Oct. 2016.

[28] S. Liu, Y. Gao, W. Zheng, and X. Li, "Performance of two neural network models in bathymetry," *Remote sensing letters*, vol. 6, no. 4, pp. 321-330, 2015.

[29] H. Mohamed, A. Negm, M. Zahran, and O. C. Saavedra, "Bathymetry Determination from High Resolution Satellite Imagery Using Ensemble Learning Algorithms in Shallow Lakes: Case Study El-Burullus Lake," *International Journal of Environmental Science and Development*, vol. 7, no. 4, pp. 295-301, 2016.

[30] H. Mohamed, A. Negm, M. Salah, K. Nadaoka, and M. Zahran, "Assessment of proposed approaches for bathymetry calculations using multispectral satellite images in shallow coastal/lake areas: a comparison of five models," Arabian Journal of Geosciences, vol. 10, no. 42, 2017.

[31] A. Melet, P. Teatini, G. Le Cozannet, C. Jamet, A. Conversi, J. Benveniste, and R. Almar, "Earth Observations for Monitoring Marine Coastal Hazards and Their Drivers," *Surveys in Geophysics*, vol. 41, pp. 1489–1534, 2020.

[32] S. R. Safavian, and D. Landgrebe, "A survey of decision tree classifier methodology," *IEEE transactions on systems, man, and cybernetics*, vol. 21, no. 3, pp. 660-674, 1991.

[33] C. Zhang, and Y. Ma, Ensemble machine learning: methods and applications. Springer Science & Business Media, 2012.

[34] L. Breiman, Bagging predictors, *Machine learning*, vol. 24, no. 2, pp. 123-140, 1996.

[35] T.K. Ho, "The random subspace method for constructing decision forests," *IEEE transactions on pattern analysis and machine intelligence*, vol. 20, no. 8, pp. 832-844, 1998.

[36] Digitalglobe. Accuracy of Worldview Products. White Paper. 2016. Available online: https://dg-cms-uploadsproduction.s3.com/uploads//DG_ACCURACY_WP_V3.pdf (accessed on 2 February 2021).

[37] S.M. Adler-Golden, P.A. Acharya, A. Berk, M. Matthew and D. Gorodetzky, "Remote bathymetry of the littoral zone from AVIRIS, LASH and QuickBird imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 43, no. 2, pp. 337-347, February 2005.

[38] S.C.J. Palmer, P. Kutser, and D.Hunter, "Remote sensing of inland waters: Challenges, progress and future directions," *Remote Sensing of Environment*, vol. 157, pp. 1-8, 2015.

[39] Y. Svetlana, E. Kotchenova, F. Vermote, M. Raffaella, J. Frank and Jr. Klemm, "Validation of vector version of 6s radiative transfer code for atmospheric correction of satellite data. Part I. Parth radiance," *Applied Optics*, vol. 45, no. 26, pp. 6762-6774, 2006.

[40] S. Kay, J. Hedley and S. Lavender, "Sun Glint correction of high and low spatial resolution images of aquatic scenes: A review of methods for visible and near-infrared wavelengths," *Remote Sensing*, vol. 1, pp. 697-730, 2009.

[41] J. Gao, "Bathymetric mapping by means of remote sensing: methods, accuracy and limitations," *Progress Physical Geographical*, vol. 33, no. 1, pp. 103-116, 2009.

[42] S.M. Hamylton, J.D. Hedley, and R.J. Beaman, "Derivation of High-Resolution Bathymetry from Multispectral Satellite Imagery: A Comparison of Empirical and Optimisation Methods through Geographical Error Analysis," *Remote Sensing*, vol. 7, no. 12, pp. 16257-16273, 2015.

[43] J. Wan, and Y. Ma, "Shallow Water Bathymetry Mapping of Xinji Island Based on Multispectral Satellite Image using Deep Learning," *Journal Indian Society Remote Sensing*, 2021.

[44] J. Z. Huang, An Introduction to Statistical Learning: With Applications in R, JABES vol. 19, pp. 556–557, 2014.

[45] G. Mountrakis, J. Im, and C. Ogole, "Support vector machines in remote sensing: A review," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 66, pp. 247-259, 2011.

[46] U. Maulik, and D. Chakraborty, "Remote Sensing Image Classification: A survey of support-vector-machine-based advanced techniques," *IEEE Geoscience and Remote Sensing Magazine*, vol. 5, no. 1, pp. 33-52, 2017.

[47] O. Kramer, "K-Nearest Neighbors," in Dimensionality Reduction with Unsupervised Nearest Neighbors. Intelligent Systems Reference Library, vol 51. Springer, Berlin, Heidelberg, 2013.

[48] L.E. Raileanu, and K. Stoffel, "Theoretical Comparison between the Gini Index and Information Gain Criteria," *Annals of Mathematics and Artificial Intelligence*, vol. 41, pp. 77–93, 2004.

[49] T. Hothorn, and B. Lausen, "Double-bagging: combining classifiers by bootstrap aggregation," *Pattern Recognition*, vol. 36, no. 6, pp. 1303-1309, 2003.

[50] A. Gul, A. Perperoglou, Z. Khan, and O. Mahmoud, "Ensemble of a subset of kNN classifiers," *Advanced Data Analysis and Classification*, vol. 12, 827–840, 2016.

[51] Jiménez, M.; González, M.; Amaro, A.; Fernández, A. Field Spectroscopy Metadata System Based on ISO and OGC Standards. *ISPRS Int. J. Geo-Inform*ation, vol. 3, pp. 1003–1022, 2014

[52] D.A. Siegel, S. Maritorena, N.B. Nelson, M.J. Behrenfeld, and C.R. McClain, "Colored dissolved organic matter and its influence on the satellite-based characterization of the ocean biosphere Geophysical," *Research Letters*, vol. 32, L20605, 2005.

[53] Plan de vigilancia ambiental del puerto de Granadilla en fase operativa (2017-2021), OAG_PVAGr_8/2016, S/C de Tenerife: Observatorio Ambiental Granadilla, 2016.

[54] Nan Xu, Xin Ma, Yue Ma, Pufan Zhao, Jian Yang and Xiao Hua Wang, "Deriving Highly Accurate Shallow Water Bathymetry From Sentinel-2 and ICESat-2 Datasets by a Multitemporal Stacking Method," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 14, pp. 6677-6685, 2021.