



Article Oil Spills or Look-Alikes? Classification Rank of Surface Ocean Slick Signatures in Satellite Data

Gustavo de Araújo Carvalho ¹,*¹, Peter J. Minnett ², Nelson F. F. Ebecken ³ and Luiz Landau ¹

- ¹ Laboratório de Sensoriamento Remoto por Radar Aplicado à Indústria do Petróleo (LabSAR), Laboratório de Métodos Computacionais em Engenharia (LAMCE), Programa de Engenharia Civil (PEC), Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia (COPPE), Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro 21941-859, RJ, Brazil; landau@lamce.coppe.ufrj.br
- ² Department of Ocean Sciences (OCE), Rosenstiel School of Marine and Atmospheric Science (RSMAS), University of Miami (UM), Miami, FL 33149, USA; pminnett@rsmas.miami.edu
- ³ Núcleo de Transferência de Tecnologia (NTT), Programa de Engenharia Civil (PEC), Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa de Engenharia (COPPE), Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro 21941-901, RJ, Brazil; nelson@ntt.ufrj.br
- * Correspondence: ggus.ocn@gmail.com

Abstract: Linear discriminant analysis (LDA) is a mathematically robust multivariate data analysis approach that is sometimes used for surface oil slick signature classification. Our goal is to rank the effectiveness of LDAs to differentiate oil spills from look-alike slicks. We explored multiple combinations of (i) variables (size information, Meteorological-Oceanographic (metoc), geo-location parameters) and (ii) data transformations (non-transformed, cube root, log_{10}). Active and passive satellite-based measurements of RADARSAT, QuikSCAT, AVHRR, SeaWiFS, and MODIS were used. Results from two experiments are reported and discussed: (i) an investigation of 60 combinations of several attributes subjected to the same data transformation and (ii) a survey of 54 other data combinations of three selected variables subjected to different data transformations. In Experiment 1, the best discrimination was reached using ten cube-transformed attributes: ~85% overall accuracy using six pieces of size information, three metoc variables, and one geo-location parameter. In Experiment 2, two combinations of three variables tied as the most effective: ~81% of overall accuracy using area (log transformed), length-to-width ratio (log- or cube-transformed), and number of feature parts (non-transformed). After verifying the classification accuracy of 114 algorithms by comparing with expert interpretations, we concluded that applying different data transformations and accounting for metoc and geo-location attributes optimizes the accuracies of binary classifiers (oil spill vs. look-alike slicks) using the simple LDA technique.

Keywords: remote sensing; synthetic aperture radar (SAR); microwave sensors; optical sensors; image processing; linear discriminant analysis (LDA); oil slicks; oil spills; oil seeps; look-alike features

1. Introduction

The sea-surface signature of mineral oil contamination ("oil slicks") can be the result of natural causes seeping out of the sea floor ("oil seeps") or being spilled through human intervention ("oil spills"). Petroleum pollution in both coastal and open-ocean waters is of great ecological concern [1,2]. Oil-related incidents usually draw media attention and public awareness, leading the oil and gas industry to enforce rigorous safety protocols and invest in contingency plans, as well as causing political conflicts, economic issues, ecological problems, and scientific concerns [3,4]. A recent catastrophic oil spillage, unprecedented in the last decades, occurred at the end of 2019 when an unknown source caused a myriad of massive oil slicks along Brazil's shoreline [5].

Remote sensing can help the detection of severe events, including the recent Brazilian case [6], or in the relatively frequent minor oil slicks observed at the ocean surface; satellite



Citation: Carvalho, G.d.A.; Minnett, P.J.; Ebecken, N.F.F.; Landau, L. Oil Spills or Look-Alikes? Classification Rank of Surface Ocean Slick Signatures in Satellite Data. *Remote Sens.* 2021, *13*, 3466. https:// doi.org/10.3390/rs13173466

Academic Editor: Konstantinos Topouzelis

Received: 5 June 2021 Accepted: 2 August 2021 Published: 1 September 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Copyright: © 2021 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (https:// creativecommons.org/licenses/by/ 4.0/). data can be useful in hindcast oil-slick models [7]. Satellite-based oil pollution monitoring has been extensively employed in recent times, and among the space-borne sensors widely used to study mineral oil floating on the surface of the ocean are the Advanced Very High-Resolution Radiometer (AVHRR [8]), Sea-Viewing Wide Field-of-View Sensor (SeaWiFS [9]), Moderate Resolution Imaging Spectroradiometer (MODIS [10]), and Synthetic Aperture Radar (SAR [11]). Although SAR is considered the most well-suited tool for oil surveillance, partly because of immunity to the surface being hidden by clouds and not needing solar illumination to provide the signal, there is a crucial issue: the ambiguity of oil signatures in the radar backscatter [12]. Other environmental phenomena also produce the same signal as oil in SAR imagery—biogenic films, algal blooms, upwelling, low winds, rain cells, and others [13]. These oil-free false targets are often called "look-alike slicks".

The remote sensing community has long invested effort to improve understanding of the oil signature in SAR measurements, a process often referred to as image segmentation [14,15]. This mostly consists of identifying smooth sea surface regions with reduced radar backscattering signals, thus delineating the shape of potential oil features [16]. Following SAR image segmentation, another major task is in developing algorithms for the discrimination between possible causes of the signals—oil slicks vs. look-alike slicks [17]. Some researchers have focused on obtaining information about automatic [18] or semiautomatic approaches [19], while others rely on human interpretation [20] to identify oil in SAR imagery. Most of these discrimination algorithms involve complex machine learning techniques, e.g., the Mahalanobis classifier [21], artificial neural networks [22], fuzzy logic [23], decision trees [24], among others; Al-Ruzouq et al. [25] reviews the most frequently used machine learning techniques used for oil slick detection. These methods also use many complicated attributes; Espedal and Johannessen [26] and Stathakis et al. [27] provide an extensive compilation of frequently used attributes. Polarimetric SAR attributes (scattering matrices) have also been investigated [28]. A series of review papers have described the processes for the detection of marine oil slicks, e.g., [29–32].

Linear discriminant analysis (LDA) is a simple supervised classification technique that can be applied to satellite measurements for classifying oil slicks [33]. Even though LDA is a mathematically robust multivariate data analysis approach, it is seldom published in the scientific literature for oil slick classification [33]. Mattson et al. [34] used LDAs to classify six different infrared spectral patterns of 194 petroleum pollutant samples. The main conclusion of their analysis was that LDAs did not reach a satisfactory classification success rate; however, the LDA performance was positive after postulating it with decision trees. Xu et al. [35] compared the use of penalized LDAs with six other techniques for the classification of 198 targets (spills and look-alikes) identified in 93 satellite images. These authors confirmed that LDAs were effective with 81% to 87% success rates depending on the choice of accuracy metric; but three other methods were more effective. In an attempt to use LDAs, among three other techniques for detecting oil spills, Liu et al. [36] explored three different marine radar images to build a semi-automatic adaptive thresholding detection method. Their LDAs were capable of flagging about 80% of the spills visually identified by human interpretation, with LDAs being the second-best technique. Exploiting 267 targets (spills and look-alikes) observed in 198 SAR images, Cao et al. [37] compared four techniques, including LDA, to train active learning methods to use fewer samples to accomplish effective oil slick classification. They found LDAs as the third best technique reducing the number of utilized samples. The conclusions presented in the few papers using LDAs in classification problems are that there is room for improvement. In the Methods section below, we give more details about LDAs.

A recent research topic is the use of LDAs to differentiate between oil categories: oil seeps vs. oil spills [33]. An aspect of the seep-spill LDA investigation is that easily identified variables (e.g., area and perimeter) resulted in successful classification rates of ~70% [38–40]. The positive results of the seep-spill LDA studies, combined with the simplicity and power of the linear analyses to classify oil slicks identified in satellite imagery, form the justifications to retain this linear classification technique in the research reported

here, where we study the classification between oil spills and look-alike slicks. While LDAs were applied to remotely sensed features obtained with the Canadian RADARSAT-2 to classify seeps and spills in Gulf of Mexico waters [41], here LDAs are applied to features retrieved in images of the Canadian RADARSAT-1 to distinguish the presence of mineral oil on the sea-surface from other petroleum-free features off the Brazilian coast [42].

Our overall objective here is to rank algorithms applied to many satellite-derived parameters in various data combinations with simple data transformations, according to their success in oil-slick classification. Two experiments to assess the classification of oil spills from look-alike slicks were designed to fulfill our two objectives to rank several combinations of (i) variables and (ii) data transformations using satellite-derived measurements (microwave, infrared, and visible):

- Exclusion or inclusion of specific types of data (Experiment 1); and
- Data transformations applied to the attributes (Experiment 2).

Besides ranking the algorithms and to find the best binary classifiers, our research also seeks to provide improved baseline information for future analyses to discriminate sea-surface features identifiable in SAR imagery. The research reported here introduces five innovations (referred to as "developments"):

- 1. Implementation of stringent knowledge-driven filters;
- 2. Use of simple morphological characteristics (or simply "size information");
- Exploration of several combinations of Meteorological-Oceanographic parameters (collectively referred to as "metoc variables");
- 4. Assess the value of the including geo-location parameters ("geo-loc");
- 5. Application of different data transformations to the attributes in the same analysis.

Following the introduction and statement of objectives given in Section 1, information about the study area and the satellite-based datasets are found in Section 2; the methods are given in Section 3; results are presented in Section 4; important remarks are reported in Section 5 in the discussion of the major findings; and the paper concludes with a summary of our results and some recommendations for future work in Section 6.

2. Study Area and Data

2.1. Study Region

Our area of interest is the Campos Basin offshore of the southeast coast of Brazil (Figure 1). The relevance of this region to the Brazilian economy is due to its numerous offshore oil and gas exploration and production facilities—in 2020, 38 operational fields represented ~25% of the country's fossil fuel supply with 989,949 barrels of oil equivalent [43].



Figure 1. Area of interest offshore from the southeastern Brazilian coast: Campos Basin. The dashed square shows the region of the observed features: oil spills and look-alike slicks. Guanabara Bay (1), Cabo Frio (2), Cabo de São Tomé (3), and isobaths (50 m, 100 m, 200 m, 1000 m, 2000 m, and 3000 m) are shown. See also Section 2.1.

The Campos Basin has very dynamic meteorological and oceanographic conditions throughout the year: during the austral summer, constant northeasterly winds support upwelling events that drop the surface water temperature and increase the primary biological production, but in the winter months, strong southwesterly winds tend to roughen the sea and primary biological production is reduced [44,45]. These phenomena are not confined to the offshore region between the Cabo de São Tomé and Cabo Frio, near Guanabara Bay, but that is where they are most frequently observed (Figure 1).

2.2. Database

A tabular remote sensing dataset, including microwave, infrared, and visible satellite measurements, was exploited here. This dataset was first utilized by Bentz [46], and later explored by Moutinho [47] and Carvalho et al. [42]. An important characteristic of this dataset for our study is the classification of oil spills vs. look-alikes based on expert interpretation. We use these interpretations as the basis for assessing the LDA accuracies.

The original dataset contained 779 individual polygons that were identified in 402 scenes of the Canadian RADARSAT-1 taken between July of 2001 and June of 2003. These 8-bit, HH polarized, C-band SAR images are from different beam modes [48,49]: ScanSAR Narrow (incident angles: 20 to 46) and Extended Low (incident angles: 10 to 23). Their data were re-sampled to ground resolutions of 100 m [46]. The borders of all observed features with low-backscatter radar signal, i.e., oil and non-oil, were delimited using a multiple resolution segmentation approach [50]. 358 spills are associated with oil samples from identified exploration or production facilities and ship-spills; confirmed spills but from unknown origins are referred to as orphan-spills. 421 look-alike slicks are sea-surface expressions of five different environmental phenomena: biogenic films, algal blooms, upwelling, low wind conditions, and convective rain cells.

Each polygon was described using 34 main descriptive characteristics divided into six attribute types:

- 1. Two textural (i.e., contrast and entropy of the pixels within the features);
- 2. Four related to SAR-signatures (e.g., standard deviation and mean ratios between the pixel values inside and outside of the targets);
- 3. Three scene-related (e.g., quantity of identified features pre-SAR image);
- 4. Nine pieces of size information (e.g., area and perimeter);
- 5. Four metoc variables—cloud cover information, wind speed (WND), sea-surface temperature (SST), and chlorophyll-a concentration (CHL)); and
- 6. Twelve geo-loc parameters (e.g., bathymetry (BAT) and distance to coastline (CST) calculated to the feature centroid).

The textural and SAR-signature attributes were calculated from uncalibrated SAR measurements, i.e., digital numbers (DNs [51]). Metoc measurements were retrieved from auxiliary environmental Earth-Observation System (EOS) satellites: WND from SeaWinds scatterometer onboard the Quick Scatterometer (QuikSCAT [52]), SST from AVHRR on the National Oceanic and Atmospheric Administration (NOAA) satellites [53,54], and CHL from SeaWiFS on the OrbView-2 satellite [55] or MODIS on the Terra satellite [56]. Additionally, ancillary WND, SST, and CHL maps, derived from measurements from these sensors, were also utilized by the experts to assist their binary classifications.

All algorithms evaluated here use part of the data records and some of the attributes contained in the "original dataset" [46]. The subset of the database analyzed here is defined after the discussion of our research strategy and data mining.

3. Methods

A pair of methodological steps was performed: research strategy and data mining exercises (Figure 2). These evolved from prior analyses using LDAs to: (i) differentiate oil spills from oil seeps in RADARSAT-2 images off the Gulf of Mexico coast (Campeche Bay, Mexico) proposed by Carvalho [33] and further developed by Carvalho et al. [38–40]; and



(ii) distinguish oil spills from look-alike slicks observed in RADARSAT-1 scenes off the coast of Brazil (Campos Basin) [42].

Figure 2. Methodological steps: research strategy and data mining exercises. Experiment 1 and Experiment 2 are aligned with our objectives.

We explored many data combinations: 60 combinations of variables (Experiment 1) and 54 combinations of data transformations (Experiment 2). In practice, each combination was considered as an individual "LDA algorithm". The data combinations in our algorithms are different to those explored in earlier studies, but are similar in number of combinations in three other papers, i.e., 32 [39] + 61 [40] + 39 [42] = 132. Of the combinations analyzed here (60 + 54 = 114), only nine have been previously investigated, but were modified as discussed below.

3.1. Research Strategy

This section has three parts describing the data filtering, the removal or inclusion of data (Experiment 1), and the consideration of various data transformations in the same analysis (Experiment 2).

3.1.1. Data-Filtering Scheme

The first development is that we removed samples based on the likelihood of them being outliers. Because of a common issue faced in data classification problems, i.e., to define a good collection of instances with representative characteristics of each class [57,58], the proposed filtering was based on local, historical, and empirical knowledge. As a result, we designed quality control tests to remove samples that include values of any variables that are unlikely to contribute to the oil spill vs. look-alike classification. The number of instances in the experiments was determined by this filtering.

3.1.2. Data Information: Removal or Inclusion

This section presents the different ways the attributes were combined to verify the consequences of removal or inclusion of data. These actions assisted in the ranking of the different combinations of variables, which is our first objective.

Of the six attribute types in the original dataset, three were not considered: textural, SAR-signature, and scene-related information (Section 2.2). In the original dataset, texture and SAR-signature had not been converted to backscatter coefficients (sigma-, beta-, or gamma-naught [59]) making it impossible to compare time series of images, but instead, they had been registered in uncalibrated DN values, therefore permitting only relative comparisons within individual scenes. Scene parameters (i.e., number of identified features per scene, sum of the areas of all features within each SAR image, etc.) cannot contribute to a classification scheme, as these are functions of the SAR swath width and not of the slicks. We thus utilized variables from the remaining three attribute types: size information, metoc variables, and geo-loc parameters (Section 2.2). Within these attribute types, we explored

three subdivisions: "Size Plus Metoc Set" (Figure 3A: blue panel); "Size Set" (Figure 3B: green panel); and "Metoc Set" (Figure 3C: gray panel). These subdivisions, which were analyzed in conjunction with geo-loc parameters (white circles in Figure 3), are further discussed below.



Figure 3. Data combinations explored to evaluate the linear discriminant analysis (LDA) algorithms during the data-information experiment fulfilling our first objective, i.e., determine the best combination of variables for linearly discriminating oil spills from look-alike slicks. Color-coded circles represent attribute types. Yellow: size information—area, compact index (CMP: $(4.\pi.area)/(\text{perimeter}^2))$, aspect ratio (length-to-width ratio: LtoW), perimeter-to-area ratio (PtoA), fractal index (FRA: 2.ln(perimeter/4)/ln(area)), and number of parts of each feature (NUM). Black: Meteorological-Oceanographic (metoc) variables—wind speed (WND), sea-surface temperature (SST), and chlorophyll-a concentration (CHL). White: geo-location (geo-loc) parameters—bathymetry (BAT) and distance to coastline (CST). Colored panels correspond to attribute-type subdivisions: (A) blue ("Size Plus Metoc Set"—9 data combinations); (B) green ("Size Set"—3 data combinations); and (C) gray ("Metoc Set"—8 data combinations). Each of these 20 combinations had all variables subjected to the same data transformation (i.e., non-transformed, cube root, or log₁₀), thus forming 60 combinations. Combinations previously explored in Carvalho et al. [42] are indicated (#). See also Section 3.1.2.

3.1.2.1. Size Information

The second development here is the independent use of simple size information. Besides the nine geometry, shape, and dimensions characteristics—area, perimeter, shape index (SHP = (perimeter/4).(area^{1/2})), compact index (CMP = $(4.\pi.area)/(perimeter^2)$), asymmetry (ASY = 1 – (ratio between feature's length and width)), aspect ratio (LtoW = length/width), density (DEN), curvature (CUR), and number of parts of each feature (NUM)—we also explored two other morphologic variables: perimeter-to-area ratio (PtoA), and fractal index (FRA = 2.ln(perimeter/4)/ln(area)). However, several of these eleven attributes are correlated: area with perimeter, CMP with SHP and DEN, LtoW with ASY, and PtoA with CUR [38–40,42]. The FRA and NUM variables did not correlate with any other attribute. The choice of uncorrelated attributes is given below (Section 3.2.2). Because the five correlated characteristics (i.e., perimeter, SHP, DEN, ASY, and CUR) led to no LDA classification improvements [42], they are not pursued here. Thus, we use the six uncorrelated variables to define the Size Set; in Figure 3 they are represented by yellow circles:

- area;
- CMP;
- LtoW;
- PtoA;
- FRA; and
- NUM.

3.1.2.2. Metoc Variables

Of the four metoc variables (clouds, WND, SST, and CHL), only cloud cover information was discretely registered as the absence (0) or presence (1) of clouds within the polygons, and is not explored further here due to its binary character. The third development explored three different combinations of metoc variables to quantify their influence (individual and combined) on the algorithm's accuracy. In Figure 3, black circles correspond to the three combinations defining the Metoc Set:

- WND, SST, and CHL;
- WND; and
- SST and CHL.

3.1.2.3. Geo-Location Parameters

The fourth development is the use of geo-loc parameters. Because most geo-location attributes are site-specific (e.g., distance to petroleum platforms or to underwater pipelines) we only considered two of them:

- bathymetry (BAT); and
- distance to coastline (CST).

In Figure 3, these parameters are shown by white circles. One should note that even though they are considered independently, they are always analyzed together with size information and/or metoc variables.

3.1.2.4. Data Transformations

The application of data transformations to the attributes prior to using them in the machine learning methods is, in principle, capable of improving algorithm classification accuracy [35]. Carvalho et al. [39] tested the LDA performance with data from eight non-linear transformations, and based on their results, we analyzed the data without any transformation (i.e., "non-transformed set") and with two data transformations:

- cube root; and
- logarithm base 10 (\log_{10}).

It should be noted that the FRA variable contains negative values and cannot be subjected to logarithmic transformation.

3.1.2.5. Data Combinations

Eleven variables were carried forward in our study: six pieces of size information (Section 3.1.2.1), three metoc variables (Section 3.1.2.2), and two geo-loc parameters (Section 3.1.2.3). These resulted in nine data combinations of the Size Plus Metoc Set subdivision with and without geo-loc (Figure 3A), three Size Set combinations with and without geo-loc (Figure 3B), and eight Metoc Set combinations with and without geo-loc (Figure 3C). The three attribute-type subdivisions when analyzed with or without geo-loc parameters formed 20 different data combinations. Each of these combinations was analyzed three times—in which all variables were subjected to the same data transformation: non-transformed, cube root, or log₁₀ (Section 3.1.2.4). In the first experiment (denoted as "Data-Information Experiment") we compared the performance of as many as 60 LDAs (20×3). This collection of LDAs was implemented to reach our first objective (Experiment 1) and differ from those proposed in the section to follow to attain our second objective (Experiment 2).

Three of the 39 combinations investigated by Carvalho et al. [42], indicated in Figure 3 by the # symbol, are also evaluated here: (i) all-size information plus all-metoc variables; (ii) all-size information; and (iii) all-metoc variables. However, Carvalho et al. [42] did not include any geo-location data, but all variables were also subjected to the same data transformations as those used in this experiment. This resulted in nine combinations (3×3) in common with their study, but here, these combinations are treated differently due to two of the five developments: the elimination of some samples and the analysis including geo-loc parameters.

3.1.3. Combined Use of Several Data Transformations in the Same Analysis

The fifth development of this research in relation to any other published binary classification studies (to our knowledge), is that we verified the influence of applying different data transformations to the attributes in the same analysis, i.e., our second objective. Three selected variables were each subjected to different transformations. Table 1 depicts the 27 possible combinations and its pool of 27 different LDAs—3 variables vs. 3 transformations. LDAs were implemented in two distinct assemblages of variables:

1. "Metoc Assemblage": WND, SST, and CHL; and

Table 1. The 27 possible data combinations of three variables (Var.) each of which are subjected to three data transformations in the same analysis: none, cube root, or log₁₀. Two distinct assemblages were used in the "Data-Transformation Experiment" to address the second objective—establish the best combination of data transformation for the discrimination of oil spills from look-alike slicks. Baseline combinations with the same transformation are given in the first row. "Metoc Assemblage": wind speed (WND), sea-surface temperature (SST), and chlorophyll-a concentration (CHL). "Size Assemblage": area, aspect ratio (length-to-width ratio: LtoW), and number of parts of each feature (NUM)—see also Figure 4 in Carvalho et al. [42]. See also Section 3.1.3.

Var. 1	Var. 2	Var. 3	Var. 1	Var. 2	Var. 3	Var. 1	Var. 2	Var. 3
None	None	None	Cube	Cube	Cube	log ₁₀	log ₁₀	log ₁₀
None	None	Cube	Cube	Cube	None	log ₁₀	log ₁₀	None
None	Cube	Cube	Cube	None	None	log ₁₀	None	None
None	Cube	None	Cube	None	Cube	log ₁₀	None	log ₁₀
None	None	log ₁₀	Cube	Cube	log ₁₀	log ₁₀	log ₁₀	Cube
None	log ₁₀	log ₁₀	Cube	log ₁₀	log ₁₀	log ₁₀	Cube	Cube
None	log ₁₀	None	Cube	log ₁₀	Cube	log ₁₀	Cube	log ₁₀
None	Cube	log ₁₀	Cube	None	log ₁₀	log ₁₀	None	Cube
None	log ₁₀	Cube	Cube	log ₁₀	None	log ₁₀	Cube	None

^{2. &}quot;Size Assemblage": area, LtoW, and NUM.

These two assemblages resulted in another series of 54 LDAs (27×2) that are used in the second experiment, referred to as the "Data-Transformation Experiment". Regarding the "Assemblage" nomenclature, the reader should not get confused with the terms using "Set" previously defined in Section 3.1.2: Size Set and Metoc Set.

While the Size Assemblage was chosen based on inspection of the dendrograms identifying uncorrelated variables (see Figure 4 in Carvalho et al. [42]), the Metoc Assemblage verifies if we can exclude the use of SAR data and solely use measurements from environmental EOSs sensors. One should note that even though the Metoc Assemblage has the same metoc variables as those from the first Metoc Set, the attributes of this assemblage are subjected to different transformations instead of the same transformation as in the set.

3.2. Data Mining Exercises

This section has three parts describing the selection of attributes, the LDA algorithms, and the evaluation of the algorithm accuracy. An open-access software package was used: Paleontological Statistics (PAST [60]).

3.2.1. Attribute-Selection Approach

Rooted tree dendrograms (Unweighted Pair Group Method with Arithmetic mean: UPGMA [61]) were used to assess the level of correlation among variables. The threshold for uncorrelated attributes using dendrograms is user-defined, and two of the most common approaches have been separately applied here:

- In Experiment 1, an across-dendrogram numeric threshold (phenon line [62]) was used to identify groups of correlated variables from which one attribute is selected per group. This used a fixed Pearson's r correlation coefficient (0.3 > r > -0.3 [63]); and
- In Experiment 2, a visual identification of correlated groups of variables, from which
 one attribute is manually selected for each group.

3.2.2. Linear Discriminant Analysis (LDA)

In addition to being used to reduce the dimensionality of data classification analyses, LDAs can be used as a classification technique [64]. In our analyses we explore conventional LDAs, but many other LDA variants exist: global-local LDA [65], probabilistic LDA [66], dual-space LDA [67], null-space LDA [68], penalized LDA [69], among others. While Tharwat et al. [70] and Legendre and Legendre [71] discuss these linear analyses in a wider context, a summary of the main benefits and weaknesses of conventional LDAs is given below:

- Advantages: LDA is a supervised classification method that uses the observed values (attribute magnitudes) of the data (samples) to determine the location of a specific boundary (a linear discriminant axis) between each group (in our case, oil and look-alikes). The LDA general concept is to use the data according to two criteria: (i) maximization of the distance between the average value of each group; and (ii) minimization of the scatter within each group. The ratio of these two criteria, mean squared differences to sum of the variances, is projected onto a line (the linear discriminant axis), providing the ability to linearly separate the groups of samples. This projected lower-dimensional space inherently preserves the group discriminat tory information, if one exists. A covariance matrix is calculated for each group along with a within-group scatter matrix to create what is called a discriminant function [72]. Numerically, this function, which corresponds to the dependent variable (DF(X)), is the sum of the product of the independent variables' values (X_n) with a calculated independent variables' weight (W_n); a constant offset may apply (C): DF(X) = (X₁W₁ + X₂W₂ + ... + X_nW_n) C [73].
- **Disadvantages:** LDA outcomes tend to support good classification decisions, but there are limitations. The number of variables must not exceed the number of samples [74]. LDAs are restricted to linearly separable groups. In addition, the variables used should have as small a correlation as possible [75]. This was accomplished through

10 of 30

the pre-selection of attributes. Another aspect to consider is that the dataset must include a binary labeling that can be used to assess the LDA performance [76]: the accuracies of our supervised learning method were verified against the baseline of the experts' classifications.

3.2.3. Classification-Accuracy Assessment

The outcomes of the LDA algorithms ("predicted classes") were assessed by comparison with the baseline interpretation of experts ("true classes") with all samples used as the training-set. We choose to work with five straightforward evaluators obtained from 2-by-2 confusion matrices [77] (Figure 4: Panel 1). Because the standalone use of the common performance metric, i.e., overall accuracy (ratio of all correct decisions to all possible outcomes), can be misleading, four additional metrics were used: sensitivity, specificity, positive- and negative-predictive values [78]. Different nomenclatures are found in the literature for these metrics, for instance: "recall" rather than sensitivity, "precision" instead of positive-predictive value, etc. [79]. These four performance metrics play equally important roles alongside the overall accuracy in measuring the success of binary classification algorithms. While sensitivity and specificity indicate the amount of previously known features correctly identified by the LDAs (the predicted classes), the positive- and negative-predictive values report how many of the features predicted by the LDA match the a priori knowledge (the true classes). Figure 4 illustrates the domains of these metrics:

- Panel 1: Diagonal analysis produces the overall accuracy;
- **Panel 2:** Horizontal analysis provides the sensitivities and specificities (the producer's accuracy), and their complements (false negatives and false positives: Type I error or omission error); and
- **Panel 3:** Vertical analysis gives the positive- and negative-predictive values (the user's accuracy) and their counterparts (inverse of the positive- and inverse of negative-predictive values: Type II error or commission error).



Figure 4. Confusion matrix, i.e., 2-by-2 table (panel 1): "Predicted classes": algorithm outcome. "True classes": expert interpretation. (**A**) Correctly classified oil spills. (**B**) Misidentified oil spills. (**C**) Misidentified look-alike slicks. (**D**) Correctly classified look-alike slicks. Number of correctly classified features: **A** + **D**. A priori known oil spills (**A** + **B**) and look-alikes (**C** + **D**)—these are fixed values established in the data-filtering scheme. Numbers of classified oil spills (**A** + **C**) and lookalikes (**B** + **D**) differ for each algorithm. Performance metrics: overall accuracy (panel 1), sensitivity and specificity (panel 2: horizontal analysis), and positive- and negative-predictive values (panel 3: vertical analysis). Compact confusion matrix form (panel 4) used to facilitate the comparison of the many explored classifiers: 114—i.e., 60 (Figure 3) plus 54 (Table 1). See also Section 3.2.3.

The classification-accuracy assessment using these three 2-by-2 matrix domains (diagonal, horizontal, and vertical) differs from other published investigations exploring oil-slick LDA classifiers, which do not report their accuracies in such a succinct manner as we do here. Some papers ignore the vertical-analysis metrics (e.g., [35]) or even both, horizontal and vertical (e.g., [34,36]).

Algorithms were deemed "void" if an evaluator was below 60%. Another reason to void the algorithms was due to unbalanced classification rates, i.e., algorithms correctly identifying 30% or more of one class than another; see Section 4.1 for the balance sampling percentages of the database analyzed here.

Because of the generation of multiple 2-by-2-tables (60 + 54 = 114), the five performance metrics are given in a compact confusion matrix form. This compact structure is shown in Figure 4 (Panel 4) and displays: the five metrics; the number of correctly identified oil spills and look-alikes (A and D, respectively); and the quantity of all correct classifications (A + D). This simple configuration enables us to construct a single table accounting for all 60 combinations (Experiment 1), and two other tables with 27 combinations each (Experiment 2).

4. Results

This section follows the research strategy (Figure 2). Throughout this section we list 15 important "remarks" that are revisited in the discussion section.

4.1. Data-Filtering Scheme

In the first part of our research (Figure 2) we indicated the number of instances utilized in the 114 LDA algorithms. The outcomes of the knowledge-driven filters are summarized in Table 2. Ten samples (eight spills and two look-alikes) were identified as having transcription errors, thus removing 1.3% of the original dataset (Table 2). Apart from these, only the WND and SST variables presented unexpected values, and their removal is summarized below:

Table 2. Summary of the data-filtering scheme showing the number of eliminated records. Wind speed (WND) filter: <3 m/s and >6 m/s. Sea surface temperature (SST) filter: <11 °C. Transcription errors (typo) filter. The statistics of all removed samples, of the original dataset instances [46], and of the analyzed database are also given. See also Section 4.1.

Class/Catagory	Oı	ginal	V	VND Filte	r	SST	Typo	All	An	alvzed
Class/Category	Da	ataset	<3 m/s	>6 m/s	Both	Filter	Filter	Filters	Da	tabase
Formation Tests	65	(8.3%)	0	-10	-10	0	-3	-13	52	(9.3%)
Accidental Discards	149	(19.1%)	$^{-2}$	-19	-21	0	-3	-24	125	(22.3%)
Ship-Spills	76	(9.9%)	-1	-13	-14	0	0	-14	62	(11.1%)
Orphan-Spills	68	(8.7%)	-4	-20	-24	0	-2	-26	42	(7.5%)
Oil Spills	358	(46.0%)	-7	-62	-69	0	-8	-77	281	(50.2%)
Biogenic Films	203	(26.1%)	-40	-1	-41	-4	0	-45	158	(28.2%)
Algal Blooms	61	(7.8%)	-18	0	-18	0	0	-18	43	(7.7%)
Upwelling	27	(3.5%)	$^{-2}$	-5	-7	0	-1	-8	19	(3.4%)
Low Wind	51	(6.5%)	-38	0	-38	0	-1	-39	12	(2.1%)
Rain Cells	79	(10.1%)	0	-26	-26	-6	0	-32	47	(8.4%)
Slick-Alikes	421	(54.0%)	-98	-32	-130	-10	-2	-142	279	(49.8%)
	Or	oinal	I	WND filter	r	SST	Typo	A11	An	alvzed
Class/Category	Da	ataset	<3 m/s	>6 m/s	Both	filter	filter	Filters	Da	tabase
		779	-105	-94	-199	-10	-10	-219		560
All reatures	10	0.0%	-13.5%	-12.0%	-25.5%	-1.3%	-1.3%	-28.1%	7	1.9%

• WND Filter: The SAR-detection ability to identify sea-surface features relies on reduced radar backscatter from the sea-surface, which is dependent on the local wind field [80]. However, the wind limits (lower and upper) to identify sea-surface features in SAR images are not agreed upon by the remote sensing community [81–83]. Weak wind conditions (<3 m/s) may prevent correct classification of features as the ambient water around them is also smooth [81]. Even though some authors have pointed out that oil slicks can be observed in ~10 m/s or higher winds (e.g., [82]), others have found the upper wind limit is ~6 m/s (e.g., [83]). To eliminate unwanted wind influence on our classifiers, samples having wind speed <3 m/s and >6 m/s were not considered. WND filtering removed 199 features (69 spills and 130 look-alikes) that represent 25.5% of the original dataset (Table 2). A primary concern about the WND variable is the ground resolution disparity between the QuikSCAT wind data and the SAR pixel: ~25 km vs. ~100 m. Although we used the wind information already included in the original dataset [46], finer wind measurements could produce different outcomes. The reader is referred to Remark 5 below, where we discuss the WND variable impact on the LDA classification decision.

SST Filter: The upwelled cold water that usually surfaces in the Campos Basin region comes from the South Atlantic Central Water and has temperatures between 6 °C and 20 °C [84]. However, an analysis of all AVHRR images from the year 2001 in this basin, 176 cloud-free scenes, did not indicate SSTs <11 °C even in the coldest core of the upwelling between Cabo de São Tomé and Cabo Frio [45]. Thus, all samples with SSTs <11 °C were removed prior to the analysis. This SST filtering did not remove any spill samples but eliminated 10 look-alike slicks amounting to 1.3% of the original dataset (Table 2). The ground resolution discrepancy between the AVHRR SSTs and SAR measurements is not as marked as that with the wind, but may also be a matter to bear in mind: ~1 km vs. ~100 m. As this filter only removed 10 look-alikes (Table 2), it is most likely that it did not exert as much influence as the WND filter on the analysis. Even though our choice of 11 °C was based on an earlier analysis, other SST thresholds could influence the LDA outcomes.</p>

These filters removed 21.5% of the oil spills (77) and 33.7% of the look-alike slicks (142) from our analyses (Table 2), resulting in ~28% fewer instances (219) being analyzed in relation to the 779 samples in the original dataset [46]. Consequently, the database analyzed here has 560 records. Since all LDAs were evaluated using the same collection of samples, the discretization resolution of our analyses is 0.18%, i.e., one misclassified feature (1/560).

While the original dataset had a somewhat unbalanced sampling percentage, 46% (358 spills) and 54% (421 look-alikes), the filtered database used here has fortuitously a well-balanced sampling: 50.2% (281 spills) and 49.8% (279 look-alikes); Table 2. This balance increased the chances of reaching good predictability levels among the five performance metrics, thus enabling a more meaningful comparison of the performance of the LDA algorithms.

The data-filtering scheme determined the most effective combination of samples by considering the magnitudes of all selected variables, thus adequately accomplishing its goal of establishing a collection of samples using a conservative approach to reduce the chances of incorrect classification in the two experiments presented below. Other factors influencing the delineation of oil-slick features in the SAR signal include oil type (light or heavy oil), slick age (time in the sea-surface since its release), acquisition geometry (incident angle), among others [85,86]. However, these were not stored as separate attributes in the dataset to allow their implementation as filters.

4.2. Experiment 1: Data Information

Sixty data combinations were analyzed in the second part of our research (Figure 2). The UPGMA dendrograms revealed that the bulk of these combinations had variables correlated at levels practically within the similarity threshold of 0.3 > r > -0.3. The LDA outcomes are presented in Table 3A,B.

Table 3. Classification accuracy from testing 60 LDA algorithms to determine the best combination of variables—first objective, i.e., "Data-Information Experiment". Inclusive hierarchy runs from 1 to 60 and is divided in three color-coded blocks: (A) Size Plus Metoc Set (blue: 1–29) (B) Size Set (green: 25–36) and Metoc Set (gray: 37–60), all of which were analyzed with or without at least one geo-location parameter and were subjected to the same data transformation (Transf.): none, cube root, or log₁₀. A ranking within attribute-type subdivisions is also provided between parentheses: 1–27 (Size Plus Metoc Set: blue), 1–9 (Size Set: green), and 1–24 (Metoc Set: gray). Blocks match the three attribute-type subdivisions (Figure 3). Size information: area, compact index (CMP: $(4.\pi.area)/(\text{perimeter}^2)$), aspect ratio (length-to-width ratio: LtoW), perimeter-to-area ratio (PtoA), fractal index (FRA: 2.ln(perimeter/4)/ln(area)), and number of parts of each feature (NUM). Meteorological-Oceanographic (metoc) variables: wind speed (WND), sea surface temperature (SST), and chlorophyll-a concentration (CHL). Geo-location (geo-loc) parameters: bathymetry (BAT) and distance to coastline (CST). Variables not used are indicated with a dot. # indicates combinations previously investigated [42]. \$ indicates hierarchies out of order. * indicates void algorithms: at least one performance metric below 60%, i.e., specificity. For the interpretation of thick table

	lines see Section 4.2.	Detailed statistica	l information	is found	in Figure 4.
--	------------------------	---------------------	---------------	----------	--------------

							(A)							
Hierarchy	(Rank)	Size		Metoc		Geo	o-Loc	Transf.	Oil	Spills	Slick	-Alikes	All F	eatures
1	(1)	Sizo		сст	СШ	рат		Cuba	251	89.3%	222	79.9%	474	84 6 9/
1	(1)	5120	WIND	331	CHL	DAI		Cube	201	81.8%	223	88.1%	4/4	04.0 /0
2	(2)	Size	WND	SST	СНІ	ват		10910	251	89.3%	221	79.2%	472	84.3%
	()	JIZC		001		DAI		10,510	201	81.2%	221	88.0%	1/2	04.070
3	(3)	Size	WND	SST	CHL		CST	Cube	250	89.0%	222	79.6%	472	84.3%
	(0)	5120	ning.	001	CIIL		COT	Cube	200	81.4%		87.7%	172	01.070
4	(4)	Size	WND	SST	CHL		CST	None	245	87.2%	226	81.0%	471	84.1%
		0120		001						82.2%		86.3%		• • • • •
5	(5)	Size	WND	SST	CHL		CST	log ₁₀	250	89.0%	221	79.2%	471	84.1%
								010		81.2%		87.7%		
# 6	(6)	Size	WND	SST	CHL			None	244	86.8%	226	81.0%	470	83.9%
										82.2%		85.9%		
# 7	(7)	Size	WND	SST	CHL			Cube	250	89.0%	220	78.9%	470	83.9%
										80.9%		87.6%		
8	(8)	Size	WND			BAT		log ₁₀	247	87.9%	223	79.9%	470	83.9%
										81.5%		86.8%		
9	(9)	Size	WND	SST	CHL	BAT		None	243	86.5%	226	81.0%	469	83.8%
TT'	(D = = 1)	<u> </u>		Mataa		0	T	Turnel	0:1	82.1%	C1'-1	85.6%	A 11 T	
Hierarchy	(Kank)	Size		Metoc		Geo	o-Loc	Iransi.	Uil	5pills	SIICK	-Alikes	All F	eatures
10	(10)	Size	WND		•		CST	Cube	247	87.9%	220	78.9%	467	83.4%
										80.7%		86.6%		
11	(11)	Size	WND		•		CST	None	239	80.1%	227	01.4%	466	83.2%
										02.1 /0		04.4 /0		
12	(12)	Size	WND	•	•	•	CST	log ₁₀	247	07.9 /0 80 5%	219	<u>70.3 /0</u>	466	83.2%
										86.1%		70.0%		
13	(13)	Size	WND	•	•	•	•	Cube	242	81.2%	223	85.1%	465	83.0%
										86.5%		79.6%		
14	(14)	Size	WND	•	•	BAT	•	Cube	243	81.0%	222	85.4%	465	83.0%
										89.0%		77.1%		
15	(15)	Size	•	SST	CHL	BAT	•	Cube	250	79.6%	215	87.4%	465	83.0%
										84.3%		81.0%		
16	(16)	Size	WND	•	•	•	•	None	237	81.7%	226	83.7%	463	82.7%
										84.3%		81.0%		
17	(17)	Size	WND	•	•	BAT	•	None	237	81.7%	226	83.7%	463	82.7%
										01.7 /0		00.1 /0		

Hier	archy	(Rank)	Size		Metoc		Geo	-Loc	Transf.	Oil	Spills	Slick	-Alikes	All Fe	eatures
#	18	(18)	Size	WND	SST	CHL	•	•	log ₁₀	244	86.8% 80.0%	218	78.1% 85.5%	462	82.5%
	19	(19)	Size	•	SST	CHL	•	•	None	246	87.5% 79.6%	216	77.4% 86.1%	462	82.5%
	20	(20)	Size	•	SST	CHL	•	CST	Cube	250	89.0% 78.9%	212	76.0% 87.2%	462	82.5%
	21	(21)	Size	•	SST	CHL	•		log ₁₀	246	87.5% 79.1%	214	76.7% 85.9%	460	82.1%
	22	(22)	Size	•	SST	CHL	•	•	Cube	245	87.2% 78.8%	213	76.3% 85.5%	458	81.8%
	23	(23)	Size	•	SST	CHL	•	CST	None	244	86.8%	214	76.7% 85.3%	458	81.8%
	24	(24)	Size	•	SST	CHL	BAT	•	None	243	86.5% 79.2%	215	77.1% 85.0%	458	81.8%
\$	26	(25)	Size	•	SST	CHL	BAT	•	log ₁₀	247	87.9% 77.9%	209	74.9% 86.0%	456	81.4%
\$	27	(26)	Size	WND	•	•	•	•	log ₁₀	240	85.4% 79.2%	216	77.4% 84.0%	456	81.4%
\$	29	(27)	Size	•	SST	CHL	•	CST	log ₁₀	248	88.3% 77.3%	206	73.8%	454	81.1%
								(B)							
Hier	archy	(Rank)	Size		Metoc		Geo	-Loc	Transf.	Oil	Spills	Slick	-Alikes	All Fe	eatures
\$	25	(1)	Size	•	•	•	BAT	•	Cube	245	87.2% 78.3%	211	75.6% 85.4%	456	81.4%
\$	28	(2)	Size			•	BAT		log ₁₀	245	87.2% 78.0%	210	75.3% 85.4%	455	81.3%
	30	(3)	Size		•	•	•	CST	Cube	248	88.3% 77.0%	205	73.5% 86.1%	453	80.9%
#	31	(4)	Size	•	•	•	•	•	log ₁₀	237	84.3% 78.7%	215	77.1% 83.0%	452	80.7%
	32	(5)	Size	•		•	•	CST	log ₁₀	245	87.2% 76.8%	205	73.5% 85.1%	450	80.4%
Hier	archy	(Rank)	Size		Metoc		Geo	-Loc	Transf.	Oil	Spills	Slick	-Alikes	All Fe	eatures
	33	(6)	Size		•		BAT		None	240	85.4% 76.9%	207	74.2% 83.5%	447	79.8 %
#	34	(7)	Size	•	•	•	•	•	None	233	82.9% 77.9%	213	76.3% 81.6%	446	79.6 %
#	35	(8)	Size						Cube	233	82.9% 77.9%	213	76.3% 81.6%	446	79.6 %
	36	(9)	Size	•	•	•	•	CST	None	241	85.8% 76.0%	203	72.8% 83.5%	444	79.3 %
Hier	archy	(Rank)	Size		Metoc		Geo	-Loc	Transf.	Oil	Spills	Slick	-Alikes	All Fe	eatures
	37	(1)	٠	WND	SST	CHL	BAT	٠	log ₁₀	220	78.3% 73.3%	199	71.3% 76.5%	419	74.8%
	38	(2)	•	WND	SST	CHL	٠	CST	log ₁₀	219	77.9% 73.2%	199	71.3% 76.2%	418	74.6%
#	39	(3)		WND	SST	CHL	٠	٠	Cube	217	77.2% 73.3%	200	71.7% 75.8%	417	74.5%

Table 3. Cont.

Hie	rarchy	(Rank)	Size		Metoc		Geo	-Loc	Transf.	Oil	Spills	Slick	Alikes	All Fe	eatures
	40	(4)		WND	SST	CHL	٠	CST	Cube	217	77.2% 73.3%	200	71.7% 75.8%	417	74.5%
#	41	(5)		WND	SST	CHL	٠	٠	log ₁₀	217	77.2%	199	71.3%	416	74.3%
	42	(6)		WND	SST	CHL	BAT	٠	Cube	216	76.9% 72.7%	198	71.0%	414	73.9%
Hier	rarchy	(Rank)	Size		Metoc		Geo	-Loc	Transf.	Oil	Spills	Slick	Alikes	All Fe	eatures
	43	(7)	٠	WND	٠	٠	٠	CST	Cube	215	76.5% 72.6%	198	71.0%	413	73.8%
	44	(8)	٠	WND	SST	CHL	BAT	٠	None	209	74.4% 73.6%	204	73.1%	413	73.8%
	45	(9)	٠	WND	٠	٠	BAT	٠	log ₁₀	214	76.2% 72.5%	198	71.0%	412	73.6%
	46	(10)	•	WND	SST	CHL		CST	None	210	74.7% 73.2%	202	72.4% 74.0%	412	73.6%
#	47	(11)	٥	WND	SST	CHL	÷	٠	None	208	74.0% 73.2%	203	72.8%	411	73.4%
	48	(12)	٠	WND	٠	٠	٠	CST	log ₁₀	217	77.2%	193	69.2% 75.1%	410	73.2%
	49	(13)	•	WND	•	•	BAT	٠	Cube	211	75.1%	197	70.6%	408	72.9%
	50	(14)		WND	•		•	CST	None	208	74.0%	198	71.0%	406	72.5%
	51	(15)	٠	WND	•	٠	BAT	٠	None	204	72.6%	197	70.6%	401	71.6%
Hie	rarchy	(Rank)	Size		Metoc		Geo	-Loc	Transf.	Oil	Spills	Slick	Alikes	All Fe	eatures
*!	52	(16)			SST	CHL	BAT	٠	Cube	223	79.4% 63.7%	152	54.5% 72.4%	375	67.0%
*!	53	(17)	•	٠	SST	CHL	٠	٠	Cube	221	78.6% 63.7%	153	54.8%	374	66.8%
*!	54	(18)	٠	٠	SST	CHL	BAT	٠	log ₁₀	209	74.4% 64.1%	162	58.1% 69.2%	371	66.3%
*!	55	(19)	٠	÷	SST	CHL	٠	CST	log ₁₀	210	74.7% 63.6%	159	57.0% 69.1%	369	65.9%
*!	56	(20)	٠	¢	SST	CHL	٠	CST	Cube	216	76.9% 63.2%	153	54.8% 70.2%	369	65.9%
*!	57	(21)	٠	٠	SST	CHL	۰	٠	None	212	75.4%	151	54.1% 68.6%	363	64.8%
*!	58	(22)	٠	٠	SST	CHL	۰	CST	None	211	75.1% 61.7%	148	53.0% 67.9%	359	64.1%
*!	59	(23)	٠	•	SST	CHL	٠	٠	log ₁₀	197	70.1%	158	56.6%	355	63.4%
*!	60	(24)	٠	•	SST	CHL	BAT	•	None	206	73.3%	145	52.0% 65.9%	351	62.7%

Table 3. Cont.

The best classification accuracy had an overall accuracy of 84.6%, in which 474 samples were correctly identified: 251 oil spills and 223 look-alike slicks. This was achieved

with good levels of other performance metrics (sensitivity (89.3%), specificity (79.9%), positive- (81.8%), and negative-predictive (88.1%) values), and was identified from a combination of ten cube-transformed attributes: six pieces of size information (area, CMP, LtoW, PtoA, FRA, and NUM) plus the three metoc variables (WND, SST, and CHL), with one geo-location parameter (BAT). In contrast, the poorest classification accuracy resulted from a combination of three non-transformed variables (SST, CHL, and BAT): 62.7%—351 successful predictions (206 oil spills and 145 look-alikes). The classification accuracy difference between the poorest and the best classifier is ~22% (123 samples). However, the result of the poorest classifier was considered void because its specificity was below 60% and it had an unbalanced identification rate (Section 3.2.3). Consequently, the lowest "valid" classification accuracy was reached with only two non-transformed variables: WND and BAT. The performance metrics were ~70%—sensitivity (72.6%), specificity (70.6%), positive-(71.3%), and negative-predictive (71.9%) values. Its overall accuracy (71.6%—401 good decisions: 204 oil spills and 197 look-alikes) is 13% (73 samples) lower than the best of all classifications.

An inclusive hierarchy based on the classifier's overall accuracies is provided in Table 3A,B: running from 1 to 60. These are assembled into "hierarchy blocks", color-coded as in Figure 3. All combinations are grouped in three major blocks corresponding to the three proposed attribute-type subdivisions with and without one of the two geo-loc parameters: size information plus metoc variables (1 to 29: blue), Size Set (25 to 36: green), and Metoc Set (37 to 60: gray)—in Table 3A,B. See Remark 1 below. The averaged values per block are presented in Table 4. Each of these color-coded blocks was also ranked within attribute-type subdivisions. These define the "subdivision ranks" which are given in parentheses in Table 3A,B: 1–27 (Size Plus Metoc Set: blue), 1–9 (Size Set: green), and 1–24 (Metoc Set: gray). Each major block was further divided in "subgroups" (Table 3A,B), based on the characteristics of the variables. The averaged subgroup information is also given in Table 4. See Remark 2 below.

Table 4. Averaged overall accuracies of Experiment 1 (Data Information). Three hierarchy blocks and their respective subgroups (as color-coded in Table 3A,B): size information plus Meteorological-Oceanographic (metoc) variables (blue: 1–29), "Size Set" (green: 25–36), and "Metoc Set" (gray: 37–60), all of which were analyzed with or without at least one geo-location (geo-loc) parameter and were subjected to the same data transformations. Averaged number of correctly classified samples is provided in parentheses. Blocks match the proposed attribute-type subdivisions (Figure 3). + indicates the range of accuracies (and samples) in these blocks. * indicates unbalanced identification rate: algorithms correctly identifying 30% or more oil spills than look-alike slicks. ! indicates void algorithms: at least one performance metric below 60%, i.e., specificity. See also Section 4.2.

Blocks	Subdivisions	Percentages	(Samples)	Subg	groups	Percentages	(Samples)
		83.0%	(465)	Top Group	WND, SST, and CHL	84.1%	(471)
Top- Blue (1–29)	Size Plus Metoc Set	2 (9/	(20)	Middle Group	WND	83.0%	(465)
(/		3.0%	(20) +	Bottom Group	SST and CHL	81.9%	(459)
Middle-	Size Set	80.3%	(450)	First Group	log_{10} or cube root	80.9%	(453)
(25–36)	Size Set	2.1%	(12) +	Second Group	Original set	79.6%	(446)
		70.5%	(395)	Top Group	WND, SST, and CHL	74.4%	(417)
Bottom- Gray (37–60)	Metoc Set	10 10/	((0))	Middle Group	WND	73.1%	(410)
		12.1%	(60) +	Bottom Group	SST and CHL	65.2%	(365) *!

The thick lines in the table show a subtle characteristic displayed in Table 3A,B that link combinations with equal overall accuracies. See Remark 3 below. Even though the hierarchy blocks and subdivision ranks are interchangeably used when we refer to blocks, hierarchies run from 1 through 60, whereas references to ranks match the attribute-type subdivision count given above. A series of findings apparent in Table 3A,B and Table 4 is discussed by subdivision rank below.

4.2.1. Size Plus Metoc Set, with or without Geo-Location (Blue: 1–27)

Within this top hierarchy block, three subgroups are identified. The top ranked nine combinations are primarily formed by the combinations of the Size Plus Metoc Set. As stated above, the best accuracy is 84.6%. The middle group has eight combinations predominantly based on size plus WND combinations. The lowest subgroup has ten combinations mostly formed by size plus SST and CHL combinations. More details are given in Remark 3 below.

Although the difference between the best and worst classification rate is 3.6% (20 samples; Table 4), there is a demonstrable synergy in combining different attributes: firstly, the six pieces of size information plus the three metoc variables (size + WND, SST, and CHL) out-performed size with only one metoc (size + WND), and secondly, size + WND surpassed size plus the other two metoc (size + SST and CHL). Regarding the use of geo-location parameters, when either of them was included, there was a gain in accuracy. In this hierarchy block, there was no improvement of the data-transformed combinations over the non-transformed set.

4.2.2. Size Set, with or without Geo-Location (Green: 1-9)

There are two subgroups in the middle hierarchy block. The first has five combinations, all of which were transformed: cube root or log_{10} . The best combination was the six size plus BAT cube transformed: 81.4% (456 samples correctly classified; Table 3B). The second subgroup has four combinations (ranks 6–9) and most of them are non-transformed combinations: size with and without geo-loc. The exception was a cube-transformed combination (rank 8: size without geo-loc) that was not within the first subgroup, but in the second.

While the averaged overall accuracy of the first group was ~81% (453 samples), the second group average was ~80% (446 samples); Table 4. The inclusion of geo-loc parameters promoted an improvement of the classification accuracies. The difference between the most and least accurate classification in this block is 2.1% (12 samples; Table 4), but data-transformed combinations have better outcomes than those without transformation—indeed this is the basis for the formation of groups in this block.

4.2.3. Metoc Set, with or without Geo-Location (Gray: 1–24)

The lowest hierarchy block has three subgroups. The top subgroup has six combinations using all three metoc variables (with and without one geo-loc) that have been transformed: cube root or log₁₀. The most successful combination in this block has three metoc variables with log transformed BAT: 74.8% (419 samples). The middle subgroup has nine combinations (ranks 7–15) that include the three non-transformed combinations of three metoc variables (with and without one geo-loc), and the six combinations only using WND plus either of the geo-loc parameters. The lowest subgroup has nine combinations (ranks 16–24) using SST and CHL, with or without geo-loc. However, they were all considered void for the two reasons given in Section 3.2.3: (i) their specificity was below 60% (Table 3B); and (ii) they had unbalanced classification rates.

The averaged overall accuracies of these groups are ~74%, ~73%, and ~65%, respectively with the number of samples correctly identified per group being 417, 410, and 365 (Table 4). The highest and lowest classification rate had a difference of 12.1% (68 samples). There was an evident synergy in using all metoc variables together, as they improved the ability of the classifier to discriminate oil spills from look-alike slicks. Likewise, the sole use of WND (with any geo-loc) produced better classifiers than those using the other two metoc variables, i.e., SST and CHL (with or without geo-loc). The use of geo-loc parameters improved the classification accuracy. There was a clear dependence on the use of data transformations in the top and middle groups, with the absence of transformations producing the least accurate classifications.

4.2.4. Comparative Classification Accuracy

In this section we compare the results of nine data combinations that have been analyzed by Carvalho et al. [42] that are indicated in Figure 3. Table 5 shows the main classification accuracy differences extracted from Table 3A,B and Table 7 in Carvalho et al. [42]; see Remark 4 below. Two differences in percentages (Diff.) are reported in Table 5, comparing (i) our results with those of Carvalho et al. [42] and (ii) the inclusion of geo-loc parameters. These are described below:

- Comparisons with Earlier Results of Carvalho et al. [42]: Although the classification accuracy is improved compared with earlier results by using subdivision of the Size Plus Metoc Set in nearly all combinations, there was one exception: log_{10} without geo-loc (82.5% - 83.0% = -0.5%). Likewise, all accuracies of the Size Set increased $(\log_{10} \text{ transformation without geo-loc: } 80.7\% - 78.0\% = 2.7\%)$. On the contrary, all combinations of the Metoc Set showed decreased accuracy, and this was independent of the inclusion of geo-loc (no transformation and no geo-loc: 73.4% - 76.9% = -3.5%). See Remark 5 below. Table 5 contains a local ordering of the three data transformations of each attribute-type subdivision. This ordering confirmed that there was no clear consistency to show which data transformation was best; in Table 5, asterisks indicate best accuracies per subdivision. An example of the lack of consistency is seen in the subdivisions of the Size Set that indicated different best transformations in each study: the overall accuracy without any transformation (79.1%) reported by Carvalho et al. [42] surpassed the application of transformations, while here, the most successful transformation without geo-loc was log_{10} (80.7%), but the best outcome including a geo-loc parameter (BAT) was with the cube-transformed combination (81.4%). See Remark 6 below.
- Including Geo-Location: In nearly all cases, combinations including at least one geo-location parameter had better performance than those without; the exception being the Metoc Set cube-transformed that remained the same with or without geo-loc: 74.5%. The largest overall accuracy increases when geo-loc parameters were considered was ~2%: the Size Set combination with cube root transformation (from 79.6% to 81.4%) and the Size Plus Metoc Set combination with log₁₀ transformation (from 82.5% to 84.3%). See Remark 7 below. In the combinations including geo-loc, BAT was preferable to CST. In only two of nine cases CST achieved superior accuracy. Indeed, among the combinations, the best classifier (cube transformed Size Plus Metoc Set) was improved by ~1% with the use of BAT: from 83.9% to 84.6% (Tables 3 and 5). See Remark 8 below.

4.3. Experiment 2: Data Transformation

Fifty-four data combinations were considered in the third part of our research (Figure 2). The analyses of the UPGMA dendrograms showed that the correlation of these combinations of variables were within the recommended similarity threshold: 0.3 > r > -0.3. Tables 6 and 7 condense the classifications of the two distinct assemblages: Metoc Assemblage and Size Assemblage; each of which having 3 variables subjected to 3 transformations—27 LDAs each. See Remark 9 below. These Results are presented below.

Table 5. Classification accuracy comparisons between our results (see the # symbol in Figure 3 and Table 3A,B) and those in Carvalho et al. [42]—see their Table 7. Attribute-type subdivision (Section 3.1.2): size information plus Meteorological-Oceanographic (metoc) variables, "Size Set", and "Metoc Set". In both studies, variables have been subjected to the same data transformation (Transf.). Herein, combinations were analyzed with or without at least one geo-location (geo-loc) parameter: bathymetry (BAT) or distance to coastline (CST). Overall accuracies are shown in bold font. A pair of differences in percentages (Diff.) are reported: (i) this study compared to Carvalho et al. [42]; and (ii) present study: with minus without geo-loc. A local order is provided per subdivision. The hierarchy (shown in parentheses) has been taken from Table 3A,B and Table 7 in Carvalho et al. [42]. * indicates the best accuracy within subdivisions. See also Section 4.2.4.

		Carvalho	et al. [42]	This Pa	per (without G	eo-Loc)	This Pa	aper (with Geo-	Loc)	
Sub Division	Transf.	Overall Accuracy	Order (Hierarchy)	Overall Accuracy	Order (Hierarchy)	Diff. i	Overall Accuracy	Order (Hierarchy)	Diff. ii	Geo- Loc
Cine Dlue	None	83.1%	2 (5)	83.9%	* 1 (6)	0.8%	84.1%	3 (4)	0.2%	CST
Size Plus	Cube Root	83.7%	* 1 (2)	83.9%	2 (7)	0.2%	84.6%	* 1 (1)	0.7%	BAT
Metoc Set	\log_{10}	83.0%	3 (7)	82.5%	3 (18)	-0.5%	84.3%	2 (2)	1.8%	BAT
	None	79.1%	* 1 (19)	79.6%	2 (34)	0.5%	79.8%	3 (33)	0.2%	BAT
Size Set	Cube Root	78.9%	2 (21)	79.6%	3 (35)	0.7%	81.4%	* 1 (25)	1.8%	BAT
	log_{10}	78.0%	3 (24)	80.7%	* 1 (31)	2.7%	81.3%	2 (28)	0.6%	BAT
	None	76.9%	2 (27)	73.4%	3 (47)	-3.5%	73.8%	3 (44)	0.4%	BAT
Metoc Set	Cube Root	77.1%	* 1 (26)	74.5%	* 1 (39)	-2.6%	74.5%	2 (40)	0.0%	CST
	log ₁₀	76.7%	3 (29)	74.3%	2 (41)	-2.4%	74.8%	* 1 (37)	0.5%	BAT

4.3.1. Metoc Assemblage (WND, SST, and CHL) with Different Data Transformations

Unlike the 27 combinations of the Size Assemblage (see below: Section 4.3.2), those with variables from the Metoc Assemblage did not form identifiable blocks (Table 6). Additionally, there was no combination being deemed void in the Metoc Assemblage (Tables 3 and 7). See Remark 10 below.

Table 6. Classification accuracy hierarchy of the 27 algorithms using three variables subjected to different data transformations in the same analysis—Meteorological-Oceanographic data (metoc: "Metoc Assemblage"—wind speed (WND), sea surface temperature (SST), and chlorophyll-a concentration (CHL)). Bold font indicates baseline combinations with the same transformation. For the interpretation of thick table lines see Section 4.3.1. Detailed statistical information is found in Figure 4. See also Tables 1 and 7, and Section 4.3: Experiment 2 (Data-Transformation).

Hierarchy	WND	SST	CHL	Oil	Spills	Slick-	Alikes	All F	eatures
1	NT	Cult	log	014	76.2%	205	73.5%	410	5 4.09/
1	None	Cube	10g ₁₀	214	74.3%	205	75.4%	419	74.8%
2	Nterre	NT	log	015	76.5%	202	72.8%	410	E 4.69/
2	None	None	10g ₁₀	215	73.9%	203	75.5%	418	74.6%
2	Nterre	log	log	010	75.8%	205	73.5%	410	E 4.69/
3	None	10g ₁₀	10g ₁₀	213	74.2%	205	75.1%	418	74.6%
	log	NT	log	010	77.6%	200	71.7%	410	E 4.69/
4	10g ₁₀	None	10g ₁₀	218	73.4%	200	76.0%	418	74.6%
_	log	NT	<u> </u>	210	77.9%	100	71.3%	410	E 4.69/
5	10g ₁₀	None	Cube	219	73.2%	199	76.2%	418	74.6%
6	0.1	<u> </u>	100	010	77.6%	200	71.7%	410	E 4.69/
6	Cube	Cube	10g ₁₀	218	73.4%	200	76.0%	418	74.6%
	C. I.	C. I.	C. L.	017	77.2%	200	71.7%	417	E4 E 0/
7	Cube	Cube	Cube	217	73.3%	200	75.8%	417	74.5%

Hierarchy	WND	SST	CHL	Oil	Spills	Slick-	Alikes	All F	eatures
		001	0112	011	77.2%	onen	71.7%		
8	Cube	log ₁₀	log_{10}	217	73.3%	200	75.8%	417	74.5%
					77.2%		71.7%		
9	Cube	log ₁₀	Cube	217	73.3%	200	75.8%	417	74.5%
	,		1		77.6%		71.3%		
10	\log_{10}	Cube	\log_{10}	218	73.2%	199	76.0%	417	74.5%
	1	1	1	015	77.2%	100	71.3%	11.6	=1.00/
11	10g ₁₀	10g ₁₀	10g ₁₀	217	73.1%	199	75.7%	416	74.3%
10	log	Nama	Norra	21(76.9%	200	71.7%	416	E4 29 /
12	\log_{10}	None	None	216	73.2%	200	75.5%	416	74.3%
10	logu	Cult	C. I.	3 10	77.6%	100	71.0%	416	E4 29 /
15	10g10	Cube	Cube	210	72.9%	190	75.9%	410	74.3 /0
14	Cuba	None	logia	216	76.9%	200	71.7%	116	74.29/
	Cube	None	10910	210	73.2%	200	75.5%	410	74.3 /0
15	Cuba	Nono	Cuba	216	76.9%	100	71.3%	115	74 19/
15	Cube	None	Cube	210	73.0%	199	75.4%	415	/4.1 /0
16	10910	Cubo	None	214	76.2%	201	72.0%	115	74 1%
10	10510	Cube	INOTIC	214	73.3%	201	75.0%	415	74.1 /0
17	None	10910	Cube	214	76.2%	201	72.0%	115	74 1%
	None	10810	Cube	214	73.3%	201	75.0%	415	74.170
18	Cube	None	None	213	75.8%	201	72.0%	414	73.9%
	Cube	ivone	ivone	210	73.2%	201	74.7%	111	75.570
19	None	None	Cube	213	75.8%	201	72.0%	414	73.9%
		ittorite	cube	210	73.2%	201	74.7%		
20	None	Cube	Cube	214	76.2%	200	71.7%	414	73.9%
					73.0%		74.9%		
21	\log_{10}	log ₁₀	Cube	218		196	70.3%	414	73.9%
		0			72.4%		75.7%		
22	Cube	Cube	None	212	75.4%	201	72.0%	413	73.8%
					73.1%		74.4%		
23	Cube	\log_{10}	None	211	75.1%	201	72.0%	412	73.6%
					73.0%		74.2%		
24	None	None	None	208	74.0%	203	72.8%	411	73.4%
					73.2%		73.6%		
25	None	Cube	None	209	74.4%	202	72.4%	411	73.4%
					73.1%		73.7%		
26	None	log ₁₀	None	209	/4.4%	202	72.7%	411	73.4%
					75.0%		73.7%		
27	log ₁₀	log ₁₀	None	213	73.8%	198	71.0%	411	73.4%
					12.470		/ 4.4 %		

Table 6. Cont.

The minimum (73.4%) and maximum (74.8%) overall accuracy rate difference was only 1.4% (8 samples; Table 6). Within this classification range, there were six values that had many combinations with similar performance, 73.4% to 74.6%, these are delineated in Table 6 by thick lines. A characteristic of most of them is that they did not correctly identify the same samples; this is apparent in Table 6 in the number of correctly classified oil spills and look-alike slicks—for instance, hierarchies 19, 20, and 21 all identified 414 samples (73.9%) but their classifications per class varied: spill (213, 214, and 218 samples, respectively) and look-alikes (201, 200, and 196, respectively). See Remark 11 below.

If we consider the baseline combinations with the three variables subjected to the same transformation (bold font in Tables 1 and 6), the cube root (74.5%) surpassed the log-transformed (74.3%), as well as the non-transformed (73.4%). These are hierarchies 7, 11, and 24, in Table 6. Note that the non-transformed version was worse by ~1% compared to the two with a transformation.

4.3.2. Size Assemblage (Area, LtoW, and NUM) with Different Data Transformations

The key outcomes of the 27 LDAs of the Size Assemblage (Table 7) are as follows: (i) two combinations reached best classification accuracy (80.9%): area (\log_{10}) and NUM (non-transformed) with LtoW (either log- or cube-transformed) (see Remark 12 below); (ii) the poorest combination was area (without transformation), LtoW (without transformation), and NUM (\log_{10} transformed)—67.0%. Nonetheless, this was considered void because specificity <60% and unbalanced identification rate differing by >30% (see Remark 13 below); and (iii) the lowest valid classification accuracy was reached with area (cube transformed), LtoW (\log_{10} transformed), and NUM (\log_{10} transformed), and NUM (\log_{10} transformed).

A remarkable accuracy improvement was observed from worst to best classifiers with different data transformations: 13.9% (78 samples; Table 7). Considering the baseline combinations with the three variables subjected to the same transformation (bold font in Tables 1 and 7), the log-transformed (78.6%) surpassed the cube-transformed (77.3%), and the non-transformed (70.2%; void). These are hierarchies 10, 14, and 20, in Table 7. The non-transformed version was poorer by >7% and voided. See Remark 14 below.

The 27 combinations within the Size Assemblage were divided into three major blocks mostly guided by a specific attribute: area (Table 7). A secondary group, apparent in these blocks, is controlled by another variable: NUM—these are shown in Table 7 by thick lines. In the blocks guided by the area variable, the application of the log transformation forms the top block, followed by combinations subjected to the cube transformation, and lastly by non-transformed versions. On the other hand, the groups controlled by the NUM variable had the non-transformed assemblage being more accurate than those with the application of cube root and log transformations. See Remark 15 below.

5. Discussion

Other than the oil-slick classification studies described in Carvalho et al. [38–40,42], involving LDA algorithms to discriminate surface ocean slicks detected in RADARSAT measurements, there are few publications in the literature (to our knowledge) classifying satellite-detected features using LDAs in a similar fashion as reported here. Most papers using LDAs to classify oil slicks differ from our research in that: (i) they were only successful once LDAs were postulated with another machine learning technique (e.g., [34]), while we reached successful discriminations solely based on the use of conventional LDAs; (ii) they fail to report essential accuracy metrics (e.g., [35]), thus ignoring the importance of reporting a full algorithm's accuracy assessment in a more efficient and effective manner; (iii) they explored marine radar images (e.g., [36]), rather than SAR satellite imagery. A pair of other characteristics set our study apart from these earlier investigations: the pre-selection of specific data (Experiment 1) and combination of attributes subjected to several data transformations in the same algorithm form (Experiment 2). In addition, of the 114 LDA algorithms tested here, only nine have been previously examined, with these

being modified here. The remainder of this section discusses the 15 remarks previously introduced in the results section.

Table 7. Classification accuracy hierarchy of the 27 algorithms using three attributes subjected to different data transformations in the same analysis—morphological characteristics ("Size Assemblage": area, aspect ratio (length-to-width ratio: LtoW), and number of parts of each feature (NUM)). The explanation of the hierarchy blocks: 1–6, 7–18, and 19–27 is given in the text. * indicates unbalanced identification rate: algorithms correctly identifying 30% or more oil spills than look-alike slicks. ! indicates void algorithms: at least one performance metric below 60%, i.e., specificity. Bold font indicates baseline combinations with the same transformation. For the interpretation of thick table lines see Section 4.3.2. Detailed statistical information is found in Figure 4. See also Tables 1 and 6, and Section 4.3: Experiment 2 (Data-Transformation).

Hier	archy	Area	LtoW	NUM	Oil S	Spills	Slick	Alikes	All F	eatures
	1	log	log	NT	250	89.0%	202	72.8%	450	00.00/
	1	10g ₁₀	10g ₁₀	None	250	76.7%	203	86.8%	453	80.9%
	•	lag	C. L.	NT	051	89.3%	202	72.4%	450	00.00/
	2	10g ₁₀	Cube	None	251	76.5%	202	87.1%	453	80.9%
	•	1	NT	NT	250	89.0%	001	72.0%	451	00 - 0/
	3	10g ₁₀	None	None	250	76.2%	201	86.6%	451	80.5%
		1	1	0.1	2.47	87.9%	200	71.7%	4.47	50 00/
	4	\log_{10}	\log_{10}	Cube	247	75.8%	200	85.5%	447	79.8%
	_	1	Ъ.т.	<u> </u>	244	87.5%	100	71.3%		
	5	\log_{10}	None	Cube	246	75.5%	199	85.0%	445	79.5%
	-	1	<u> </u>	<u> </u>		87.5%	100	71.3%		
	6	\log_{10}	Cube	Cube	246	75.5%	199	85.0%	445	79.5%
Hier	archy	Area	LtoW	NUM	Oil S	Spills	Slick	Alikes	All F	eatures
	_	<u> </u>			• (0	95.7%	4	62.7%		
*	7	Cube	None	None	269	72.1%	175	93.6%	444	79.3%
	-	<u> </u>			• • • •	94.7%		62.7%		T 2 20/
*	8	Cube	\log_{10}	None	266	71.9%	175	92.1%	441	78.8%
					2/7	95.0%		62.4%		
*	9	Cube	Cube	None	267	71.8%	174	92.6%	441	78.8%
		1	1	1		85.1%	• • •	72.0%		
	10	log ₁₀	log ₁₀	\log_{10}	239	75.4%	201	82.7%	440	78.6%
		1	Ът	1	240	85.4%	100	71.0%	100	
	11	\log_{10}	None	\log_{10}	240	74.8%	198	82.8%	438	78.2%
		1	6.1	1	220	85.1%	100	71.3%	100	- 0.00/
	12	\log_{10}	Cube	\log_{10}	239	74.9%	199	82.6%	438	78.2%
	10	<u> </u>	1	<u> </u>	051	89.3%	100	65.6%	40.4	0/
*	13	Cube	\log_{10}	Cube	251	72.3%	183	85.9%	434	77.5%
*	44	6.1	<u> </u>	0.1	051	89.3%	100	65.2%	400	77 20/
*	14	Cube	Cube	Cube	251	72.1%	182	85.8%	433	77.3%
*	15	<u> </u>	NT	0.1	250	89.0%	101	64.9%	401	77 00/
*	15	Cube	None	Cube	250	71.8%	181	85.4%	431	77.0%
	4.5	<u> </u>	ŊŢ	1	2.42	86.5%	100	67.4%	101	00/
	16	Cube	None	\log_{10}	243	72.8%	188	83.2%	431	77.0%
	4 -	<u> </u>	<u> </u>	lac	0.41	85.8%	105	66.3%	107	E C 40/
	17	Cube	Cube	\log_{10}	241	71.9%	185	82.2%	426	76.1%
	10	<u> </u>	lac	lac	0.10	86.1%	104	65.9%	101	B (10/
	18	Cube	\log_{10}	\log_{10}	242	71.8%	184	82.5%	426	76.1%

Hier	archy	Area	LtoW	NUM	Oil	Spills	Slick	-Alikes	All F	eatures
*1	10	NT	100	NT	246	87.5%	140	53.4%	205	70 50/
.1	19	None	10g ₁₀	None	246	65.4%	149	81.0%	395	70.5%
*1	20	None	None	None	247	87.9%	146	52.3%	202	70.2%
	20	None	None	none	247	65.0%	140	81.1%	393	70.2%
*1	01	Mana	Culta	Nama	247	87.9%	140	52.3%	202	70.09/
	21	None	Cube	Inone	247	65.0%	140	81.1%	393	70.2%
*1	22	None	logu	Cuba	220	81.9%	150	56.6%	200	60.29/
	22	None	10g ₁₀	Cube	230	65.5%	158	75.6%	388	69.3%
*1	22	Mana	Nama	Culta	224	83.3%	152	54.8%	207	(0.19/
	25	None	None	Cube	234	65.0%	155	76.5%	387	69.1%
*1	24	None	Cuba	Cuba	220	81.9%	157	56.3%	207	60.19/
	24	None	Cube	Cube	230	65.3%	157	75.5%	367	09.1 /0
*1	25	None	logra	logia	221	78.6%	159	56.6%	270	67 79/
	25	None	10g10	10g10	221	64.6%	136	72.5%	379	07.7 /0
*1	26	None	Cube		210	77.9%	160	57.3%	270	67 79/
.1	26	none	Cube	10g10	219	64.8%	100	72.1%	379	0/.7%
*1	27	None	Nana	logu	210	77.9%	156	55.9%	275	(7.09/
*!	27	inone	inone	10g ₁₀	219	64.0%	136	71.6%	3/5	67.0%

Table 7. Cont.

5.1. Data-Information Experiment

- **Remark 1:** Considering the hierarchy blocks, when variables from Size Plus Metoc Set were combined, the algorithms were more accurate than those using variables from one type alone. Additionally, when comparing the sole use of size information, the classification accuracies were superior to those using only the metoc variables. A corresponding hierarchical pattern was also observed among the 61 data combinations reported in Carvalho et al. [40]. The hierarchy block formation was only disrupted by two combinations of the Size Set (hierarchies 25 and 28: green group) that were more accurate than a few combinations of the Size Plus Metoc Set (hierarchies 26, 27, and 29: blue group).
- **Remark 2:** Regarding the subgroups, it is noteworthy that some data combinations achieve classifications better than others (Table 3A,B). Table 4 shows the top-blue (Size Plus Metoc Set) and middle-green (Size Set) blocks have an average difference of ~1% between each of their groups: ~84% to ~80%. The differences between the middle-green and lowest-gray (Metoc Set) blocks were greater, as were those within the groups in the last block.
- **Remark 3:** Of the many combinations that had the same overall accuracies (to the number of decimal places indicated), most of them did not correctly identify the same samples—this is seen in Table 3A,B: the number of correctly classified oil spills and look-alike slicks. Only hierarchies 34 and 35 (79.6%—Size Set without geo-loc: non-transformed and cube root, respectively) and hierarchies 39 and 40 (74.5%—Metoc Set with and without CST, both cube-transformed) identified the same samples.

5.1.1. Comparative Classification Accuracy

• **Remark 4:** Although nearly all accuracies were improved in the Size Plus Metoc Set and Size Set subdivisions described by Carvalho et al. [42], the same did not hold for the Metoc Set subdivision that had its overall accuracies reduced (Table 5). While the largest improvements were ~3% in two log-transformed Size Set combinations: without geo-loc (from 78.0% to 80.7%) and with geo-loc (from 78.0% to 81.3%), the best of all combinations (cube transformed Size Plus Metoc Set) had its accuracy increased by ~1% by the inclusion of one geo-loc parameter (BAT): from 83.7% to 84.6% (Table 5).

These improvements demonstrate the success of the removal of samples that are unlikely to contribute to the classification and the addition of geo-loc attributes.

5.1.2. Comparisons with Earlier Results

- **Remark 5:** The Metoc Set combinations did not produce high-ranking accuracies in comparison with the earlier results of Carvalho et al. [42] (Table 5). This may be due to many records having been removed based on the WND thresholds: lower (<3 m/s: 105 samples) and upper (>6 m/s: 94 samples)—i.e., 25.5% of the original dataset (Table 2), even though the exclusion of these cases was based on physical reasoning.
- **Remark 6:** There was not a clear pattern to indicate which data transformation was best. The non-transformed set and log₁₀ had only two cases each as the best combination among the nine compared, and the cube-transformed combinations were more accurate in five cases (Table 5).

5.1.3. Geo-Location Inclusion

- **Remark 7:** Two geo-loc parameters available in the original dataset were studied here, but they were not considered together because they are highly correlated. The inclusion of geo-loc parameters results in improved accuracies (Table 5).
- Remark 8: Combinations using Bathymetry (BAT, ranging from 5 m to ~4 km) tended to have improved accuracies compared to those using the distance to coastline (CST, 186 m to ~435 km); Table 5.
- 5.2. Data-Transformation Experiment
- **Remark 9:** The investigation of two assemblages of only three variables subjected to three data transformations indicated that the Metoc Assemblage did not show an advantage of using the different transformations, however, the results of using different data transformations within the variables of the Size Assemblage were promising; see below Remarks 12 and 15, and Future Work Recommendations.

5.2.1. Metoc Assemblage: WND, SST, and CHL

- Remark 10: There was a lack of hierarchy blocks in the Metoc Assemblage subjected to different transformations. This may be due to the relatively small range among the analyzed features (WND (3 to 6 m/s), SST (11.44 to 29.43 °C), and CHL (0.003 and 9.7 mg/m³)).
- **Remark 11:** Even though there was a span of 1.4% (8 samples) between the best and worst accuracy among the 27 combinations of the Metoc Assemblage, if we compare the baseline combinations of three pieces of metoc variables with the same transformation (shown in bold in Table 6), we notice that subjecting variables to different data transformations in the same analysis slightly improved the accuracies of the LDA algorithms.

5.2.2. Size Assemblage: Area, LtoW, and NUM

- **Remark 12:** The use of three pieces of size information subjected to different transformations (i.e., the two combinations that tied with 80.9%—area (log₁₀), LtoW (logor cube-transformed), and NUM (non-transformed); hierarchies 1 and 2 in Table 7) reached an equivalent accuracy to the best combination of six pieces of size information log-transformed without geo-loc or metoc (80.7%; hierarchy 31 in Table 3B). Clearly, the combination of various attributes subjected to several data transformations in the same analysis, can lead to improving the LDA algorithm accuracy.
- **Remark 13:** The combinations using non-transformed areas were void—hierarchies 19 to 27 in Table 7. The lack of data transformation may also be negatively influencing other combinations of variables using the non-transformed area, for example, those among the 60 depicted in Figure 3, and presented in Table 3A,B. As such, other

variables may also be suffering from using non-transformed areas, and this should be further investigated. See also Remark 15 below.

- **Remark 14:** The best of the three baseline combinations of three pieces of size information with the same transformation (shown in bold font in Tables 1 and 7) was subjected to log₁₀—78.6% (hierarchy 10 in Table 7). However, nine other combinations were better, the best being 80.9% (hierarchies 1 and 2 in Table 7). This improvement of 2.3% is another indication that the combined use of attributes subjected to different data transformations improves the LDA classification accuracy.
- **Remark 15:** Considering the major hierarchy blocks and secondary groups, among the 27 combinations that use three pieces of size information with three data transformations (Table 7), one reason is given for this ranking: among the 560 analyzed features, areas have a large range of continuous values (from oil spills with 0.45 km² to look-alikes with 8177.24 km² cause by upwelling events), whereas the NUM variable with its discrete values had features with only 1 part up to look-alike slicks with 24 different parts caused by biogenic films.

6. Summary and Conclusions

We report on successful differentiation of oil spills from look-alike slicks using simple, linear discriminant analyses (LDAs) of satellite-based information (RADARSAT-1, QuikSCAT, AVHRR, SeaWiFS, and MODIS) from the Campos Basin, Brazil (Figure 1). A series of effective classification algorithms was produced based on the combination of characteristics of three attribute types: (i) morphological characteristics (size information: area, compact index (CMP), aspect ratio (length-to-width ratio: LtoW), perimeter-to-area ratio (PtoA), fractal index (FRA), and number of feature's parts (NUM)); (ii) Meteorological-Oceanographic (metoc) variables (wind speed (WND), sea-surface temperature (SST), and chlorophyll-a concentration (CHL)); and (iii) geo-location (geo-loc) parameters (bathymetry (BAT) and distance to coastline (CST)). Two data transformations were considered in addition to non-transformed: cube root and log₁₀. The quantitative accuracy of 114 LDA algorithms was evaluated and ranked with five performance metrics: overall accuracy, sensitivity, specificity, positive-, and negative-predictive values (Figure 4). This study was built upon the ability to distinguish sea-surface features in SAR images using LDAs—oil spills vs. look-alike slicks [42], as well as oil spills vs. oil seeps [38–40]—and included developments beyond past research [33–37]. Our two objectives have been achieved through two separate experiments (Figure 2):

6.1. Objective 1

The "Data-Information Experiment" sought the most effective combination of variables among 60 combinations (Figure 3). Three proposed attribute-type subdivisions were hierarchized in major blocks: "Size Plus Metoc Set", "Size Set", and "Metoc Set" (Table 3A,B). These were considered with or without at least one geo-loc parameter and all variables were subjected to the same data transformations. The best accuracies were reached with all variables from each subdivision. Each block was further stratified in subgroups related to the variable's characteristics (Table 4). Bathymetry (BAT) was generally better than distance to coastline (CST). The main developments used here—sample removal (data filter) and inclusion of geo-loc information—improved classification accuracy (Table 5). The main results regarding the LDA accuracies (Table 3A,B) are summarized as:

- if all variables are available, the best accuracy is 84.6% (hierarchy 1; cube-transformed);
- without geo-loc parameters, the best accuracy is 83.9% (hierarchy 6; non-transformed);
- if Oceanographic data are not available, the best accuracy is 83.9% (hierarchy 8; log-transformed);
- if Meteorological data are unavailable, the best accuracy is 83.0% (hierarchy 15; cube-transformed);
- if only size information is given, the best accuracy is 80.7% (hierarchy 31; log-transformed);
- without size information, the best accuracy is 74.8% (hierarchy 37; log-transformed);

- if only Meteorological data and geo-loc are used, the best accuracy is 73.8% (hierarchy 43; cube-transformed); and
- if only Oceanographic data are accounted for (with or without geo-loc), the results are considered void (hierarchies 52–60).

6.2. Objective 2

The "Data-Transformation Experiment" sought the most effective combination of data transformations to improve accuracy. This experiment is a development over published binary classification papers, as here we combined variables undergoing different data transformations in the same analysis. Two distinct assemblages of 27 data combinations each with three variables were tested with three data transformations (Table 1). In the first assemblage ("Metoc Assemblage": WND, SST, and CHL-Table 6), there was no noteworthy classification improvement as revealed by the small range of $\sim 1.5\%$ (8 samples) from its best (74.8%) to worst (73.4%) overall accuracies. On the contrary, the second assemblage ("Size Assemblage": area, LtoW, and NUM-Table 7) showed accuracy improvements from different transformations—the best (80.9%) to worst (67.0%) accuracy had a remarkable difference of ~14.0% (78 samples). Two combinations subjected to three transformations tied as the most effective LDA—80.9% (453 samples): area (log₁₀), LtoW (log- or cube-transformed), and NUM (non-transformed). These two best combinations of three variables vs. three transformations were superior to the best baseline combination with the same transformation applied to all variables—78.6% (440 samples): area (log_{10}), LtoW (log_{10}), and NUM (log_{10}); Tables 1 and 7. Moreover, they achieved a comparable outcome to the best combination using the six pieces of size information (without metoc or geo-loc) all being subjected to \log_{10} transformation (80.7%; 452 samples). The framework of combining different data transformations in the same classification algorithm simplifies and optimizes the LDA classification as fewer attributes were used to reach the same result.

6.3. Future Work Recommendations

Future work could apply other linear and non-linear methods (e.g., decision tree, random forest, support vector machine, artificial neural network) to guide the development of improved classifiers. A continuation of this research could include a larger collection of variables being subjected to different data transformations in the same classification algorithm, as it would be interesting to investigate if the behavior observed in the Data-Transformation Experiment also occurs with other attributes, i.e., testing different data transformations on a greater number of variables. For instance, what would happen if in the best Size Set combinations that accounts for six variables (without metoc or geo-loc) all of which were subjected to log₁₀ (i.e., 80.7%; 452 samples; hierarchy 31 in Table 3A,B) had been subjected to different transformations?

Author Contributions: G.A.C.: Data interpretation and analyses, experiment design, draft preparation, writing, funding acquisition. P.J.M.: Project supervision, experiment design, draft preparation, paper revision. N.F.F.E.: Project supervision, experiment design, draft preparation. L.L.: Project supervision, funding acquisition. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Brazilian National Postdoctoral Program (*Programa* Nacional de Pós Doutorado: PNPD) of the Coordination for the Improvement of Higher Education Personnel (*Coordenação de Aperfeiçoamento de Pessoal de Nível Superior*: CAPES).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Not applicable.

Acknowledgments: We thank Roberta Santana for constructive discussions, and gratefully acknowledge Cristina Bentz for advice on the characteristics of the dataset, as well as Lindzai Taylor, Patricia McCoy, and Lucas Williams for suggestions to clarify the text. We thank the Canadian Space Agency (CSA) and the National Aeronautics and Space Administration (NASA) for data from their Earth observation satellites, the developers of the open-access PAleontological STatistics (PAST) software, and the reviewers for their comments that led to an improved paper.

Conflicts of Interest: The authors declare no conflict of interest.

References

- 1. RCC (National Research Council Committee). *Oil in the Sea: Inputs, Fates, and Effects;* The National Academies Press: Washington, DC, USA, 1985. Available online: https://www.nap.edu/read/314/chapter/1 (accessed on 30 July 2021).
- NRCC (National Research Council Committee). Oil in the Sea III: Inputs, Fates, and Effects; The National Academies Press: Washington, DC, USA, 2003; ISBN 9780309084383. Available online: https://www.nap.edu/read/10388/chapter/1 (accessed on 30 July 2021).
- 3. Leifer, I.; Lehr, W.J.; Simecek-Beatty, D.; Bradley, E.; Clark, R.; Dennison, P.; Hu, Y.; Matheson, S.; Jones, C.E.; Holt, B.; et al. State of the art satellite and airborne marine oil spill remote sensing: Application to the BP Deepwater Horizon oil spill. *Remote Sens. Environ.* **2012**, *124*, 185–209. [CrossRef]
- 4. Neuparth, T.; Moreira, S.; Santos, M.; Henriques, M.A.R. Review of oil and HNS accidental spills in Europe: Identifying major environmental monitoring gaps and drawing priorities. *Mar. Pollut. Bull.* **2012**, *64*, 1085–1095. [CrossRef] [PubMed]
- Soares, M.D.O.; Teixeira, C.; Bezerra, L.E.A.; Paiva, S.; Tavares, T.C.L.; Garcia, T.M.; de Araújo, J.T.; Campos, C.C.; Ferreira, S.M.C.; Matthews-Cascon, H.; et al. Oil spill in South Atlantic (Brazil): Environmental and governmental disaster. *Mar. Policy* 2020, 115, 103879. [CrossRef]
- 6. Soares, M.O.; Teixeira, C.; Bezerra, L.E.; Rossi, S.; Tavares, T.; Cavalcante, R. Brazil oil spill response: Time for coordination. *Science* 2020, *367*, 155. [CrossRef] [PubMed]
- Coppini, G.; De Dominicis, M.; Zodiatis, G.; Lardner, R.; Pinardi, N.; Santoleri, R.; Colella, S.; Bignami, F.; Hayes, D.R.; Soloviev, D.; et al. Hindcast of oil-spill pollution during the Lebanon crisis in the Eastern Mediterranean, July–August 2006. *Mar. Pollut. Bull.* 2011, 62, 140–153. [CrossRef] [PubMed]
- 8. Stringer, W.J.; Ahlnäs, K.; Royer, T.C.; Dean, K.E.; Groves, J.E. Oil spill shows on satellite image. *EOS Trans.* **1989**, *70*, 564. [CrossRef]
- 9. Banks, S. SeaWiFS satellite monitoring of oil spill impact on primary production in the Galápagos Marine Reserve. *Mar. Pollut. Bull.* **2003**, 47, 325–330. [CrossRef]
- Pisano, A.; Bignami, F.; Santoleri, R. Oil Spill Detection in Glint-Contaminated Near-Infrared MODIS Imagery. *Remote Sens.* 2015, 7, 1112–1134. [CrossRef]
- Jackson, C.R.; Apel, J.R. Synthetic Aperture Radar Marine User's Manual; NOAA/NESDIS; Office of Research and Applications: Washington, DC, USA, 2004. Available online: http://www.sarusersmanual.com (accessed on 30 July 2021).
- 12. Gens, R. Oceanographic Applications of SAR Remote Sensing. GIScience Remote Sens. 2008, 45, 275–305. [CrossRef]
- 13. Espedal, H.A.; Johannessen, O.M.; Knulst, J. Satellite detection of natural films on the ocean surface. *Geophys. Res. Lett.* **1996**, *23*, 3151–3154. [CrossRef]
- 14. Garcia-Pineda, O.; Zimmer, B.; Howard, M.; Pichel, W.G.; Li, X.; MacDonald, I.R. Using SAR images to delineate ocean oil slicks with a texture-classifying neural network algorithm (TCNNA). *Can. J. Remote Sens.* **2009**, *35*, 411–421. [CrossRef]
- 15. Yekeen, S.T.; Balogun, A.; Yusof, K.B.W. A novel deep learning instance segmentation model for automated marine oil spill detection. *ISPRS J. Photogramm. Remote Sens.* 2020, 167, 190–200. [CrossRef]
- Ayed, I.B.; Mitiche, A.; Belhadj, Z. Multiregion level-set partitioning of synthetic aperture radar images. *IEEE Trans. Pattern Anal. Mach. Intell.* 2005, 27, 793–800. [CrossRef]
- 17. Topouzelis, K.; Karathanassi, V.; Pavlakis, P.; Rokos, D. Detection and discrimination between oil spills and look-alike phenomena through neural networks. *ISPRS J. Photogramm. Remote. Sens.* **2007**, *62*, 264–270. [CrossRef]
- 18. Marghany, M. RADARSAT automatic algorithms for detecting coastal oil spill pollution. *Int. J. Appl. Earth Obs. Geoinf.* **2001**, *3*, 191–196. [CrossRef]
- Calabresi, G.; Del Frate, F.; Lichtenegger, I.; Petrocchi, A.; Trivero, P. Neural networks for the oil spill detection using ERS–SAR data. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS '99), Hamburg, Germany, 28 June–2 July 1999; pp. 215–217. [CrossRef]
- 20. Jones, B. A comparison of visual observations of surface oil with Synthetic Aperture Radar imagery of the Sea Empress oil spill. *Int. J. Remote Sens.* 2001, 22, 1619–1638. [CrossRef]
- Fiscella, B.; Giancaspro, A.; Nirchio, F.; Pavese, P.; Trivero, P. Oil spill monitoring in the Mediterranean Sea using ERS SAR data. In Proceedings of the Envisat Symposium, ESA, Göteborg, Sweden, 16–20 October 1998; 9p.
- 22. Del Frate, F.; Petrocchi, A.; Lichtenegger, J.; Calabresi, G. Neural networks for oil spill detection using ERS-SAR data. *IEEE Trans. Geosci. Remote Sens.* 2000, *38*, 2282–2287. [CrossRef]
- Keramitsoglou, I.; Cartalis, C.; Kiranoudis, C.T. Automatic identification of oil spills on satellite images. *Environ. Model. Softw.* 2006, 21, 640–652. [CrossRef]
- 24. Topouzelis, K.; Psyllos, A. Oil spill feature selection and classification using decision tree forest on SAR image data. *ISPRS J. Photogramm. Remote Sens.* **2012**, *68*, 135–143. [CrossRef]

- 25. Al-Ruzouq, R.; Gibril, M.; Shanableh, A.; Kais, A.; Hamed, O.; Al-Mansoori, S.; Khalil, M. Sensors, Features, and Machine Learning for Oil Spill Detection and Monitoring: A Review. *Remote Sens.* **2020**, *12*, 3338. [CrossRef]
- Espedal, H.A.; Johannessen, O.M. Cover: Detection of oil spills near offshore installations using synthetic aperture radar (SAR). Int. J. Remote Sens. 2000, 21, 2141–2144. [CrossRef]
- 27. Stathakis, D.; Topouzelis, K.; Karathanassi, V. Large-scale feature selection using evolved neural networks. *Remote Sens.* 2006, 6365, 636513. [CrossRef]
- 28. Li, G.; Li, Y.; Hou, Y.; Wang, X.; Wang, L. Marine Oil Slick Detection Using Improved Polarimetric Feature Parameters Based on Polarimetric Synthetic Aperture Radar Data. *Remote Sens.* **2021**, *13*, 1607. [CrossRef]
- 29. Alpers, W.; Holt, B.; Zeng, K. Oil spill detection by imaging radars: Challenges and pitfalls. *Remote Sens. Environ.* 2017, 201, 133–147. [CrossRef]
- 30. Fingas, M.F.; Brown, C.E. Review of oil spill remote sensing. Spill Sci. Technol. Bull. 1997, 4, 199–208. [CrossRef]
- 31. Fingas, M.; Brown, C. Review of oil spill remote sensing. Mar. Pollut. Bull. 2014, 83, 9–23. [CrossRef] [PubMed]
- 32. Fingas, M.; Brown, C.E. A Review of Oil Spill Remote Sensing. Sensors 2017, 18, 91. [CrossRef]
- 33. Carvalho, G.A. Multivariate Data Analysis of Satellite-Derived Measurements to Distinguish Natural from Man-Made Oil Slicks on the Sea Surface of Campeche Bay (Mexico). Ph.D. Thesis, COPPE, Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, Brazil, 2015; 285p. Available online: http://www.coc.ufrj.br/pt/teses-de-doutorado/390-2015/4618-gustavo-de-araujo-carvalho (accessed on 30 July 2021).
- 34. Mattson, J.S.; Mattson, C.S.; Spencer, M.J.; Spencer, F.W. Classification of petroleum pollutants by linear discriminant function analysis of infrared spectral patterns. *Anal. Chem.* **1977**, *49*, 500–502. [CrossRef] [PubMed]
- 35. Xu, L.; Li, J.; Brenning, A. A comparative study of different classification techniques for marine oil spill identification using RADARSAT-1 imagery. *Remote Sens. Environ.* 2014, 141, 14–23. [CrossRef]
- Liu, P.; Li, Y.; Liu, B.; Chen, P.; Xu, A.J. Semi-Automatic Oil Spill Detection on X-Band Marine Radar Images Using Texture Analysis, Machine Learning, and Adaptive Thresholding. *Remote Sens.* 2019, 11, 756. [CrossRef]
- 37. Cao, Y.; Xu, L.; Clausi, D. Exploring the Potential of Active Learning for Automatic Identification of Marine Oil Spills Using 10-Year (2004–2013) RADARSAT Data. *Remote Sens.* 2017, *9*, 1041. [CrossRef]
- Carvalho, G.A.; Minnett, P.J.; de Miranda, F.P.; Landau, L.; Paes, E.T. Exploratory Data Analysis of Synthetic Aperture Radar (SAR) Measurements to Distinguish the Sea Surface Expressions of Naturally-Occurring Oil Seeps from Human-Related Oil Spills in Campeche Bay (Gulf of Mexico). *ISPRS Int. J. Geo-Inf.* 2017, *6*, 379. [CrossRef]
- Carvalho, G.A.; Minnett, P.J.; Paes, E.T.; de Miranda, F.P.; Landau, L. Refined Analysis of RADARSAT-2 Measurements to Discriminate Two Petrogenic Oil-Slick Categories: Seeps versus Spills. J. Mar. Sci. Eng. 2018, 6, 153. [CrossRef]
- Carvalho, G.A.; Minnett, P.J.; Paes, E.T.; de Miranda, F.P.; Landau, L. Oil-Slick Category Discrimination (Seeps vs. Spills): A Linear Discriminant Analysis Using RADARSAT-2 Backscatter Coefficients in Campeche Bay (Gulf of Mexico). *Remote Sens.* 2019, 11, 1652. [CrossRef]
- Carvalho, G.A.; Minnett, P.J.; de Miranda, F.P.; Landau, L.; Moreira, F. The Use of a RADARSAT-derived Long-term Dataset to Investigate the Sea Surface Expressions of Human-related Oil spills and Naturally Occurring Oil Seeps in Campeche Bay, Gulf of Mexico. *Can. J. Remote Sens.* 2016, 42, 307–321. [CrossRef]
- Carvalho, G.A.; Minnett, P.J.; Ebecken, N.F.F.; Landau, L. Classification of Oil Slicks and Look-Alike Slicks: A Linear Discriminant Analysis of Microwave, Infrared, and Optical Satellite Measurements. *Remote Sens.* 2020, 12, 2078. [CrossRef]
- 43. ANP (Agência Nacional do Petróleo, Gás Natural e Biocombustíveis). Oil and Natural Gas Production Bulletin, External Circulation; n. 120; ANP (Agência Nacional do Petróleo, Gás Natural e Biocombustíveis): Brasilia, Brazil, 2020; 46p. Available online: http://www.anp.gov.br/publicacoes/boletins-anp/2395-boletim-mensal-da-producao-de-petroleo-e-gas-natural (accessed on 30 July 2021).
- 44. Campos, E.; Gonçalves, J.E.; Ikeda, Y. Water mass characteristics and geostrophic circulation in the South Brazil Bight: Summer of 1991. J. Geophys. Res. Space Phys. 1995, 100, 18537–18550. [CrossRef]
- Carvalho, G.A. Wind Influence on the Sea Surface Temperature of the Cabo Frio Upwelling (23° S/42° W—RJ/Brazil) during 2001, through the Analysis of Satellite Measurements (Seawinds-QuikScat/AVHRR-NOAA). Bachelor's Thesis, UERJ, Rio de Janeiro, Brazil, 2002; 210p. Available online: goo.gl/reqp2H (accessed on 30 July 2021).
- 46. Bentz, C.M. Reconhecimento Automático de Eventos Ambientais Costeiros e Oceânicos em Imagens de Radares Orbitais. Ph.D. Thesis, Universidade Federal do Rio de Janeiro (UFRJ), COPPE, Rio de Janeiro, Brazil, 2006; 115p. Available online: http://www.coc.ufrj.br/index.php?option=com_content&view=article&id=1048:cristina-maria-bentz (accessed on 30 July 2021).
- Moutinho, A.M. Otimização de Sistemas de Detecção de Padrões em Imagens. Ph.D. Thesis, Universidade Federal do Rio de Janeiro (UFRJ), COPPE, Rio de Janeiro, Brazil, 2011; 133p. Available online: http://www.coc.ufrj.br/index.php/teses-dedoutorado/155-2011/1258-adriano-martins-moutinho (accessed on 30 July 2021).
- Fox, P.A.; Luscombe, A.P.; Thompson, A.A. RADARSAT-2 SAR modes development and utilization. *Can. J. Remote Sens.* 2004, 30, 258–264. [CrossRef]
- 49. MDA (MacDonald, Dettwiler and Associates Ltd.). *RADARSAT-2 Product Description*; Technical Report RN-SP-52-1238, Issue/Revision: 1/13; MDA: Richmond, BC, Canada, 2016; p. 91.

- Baatz, M.; Schape, A. Multiresolution segmentation—An optimization approach for high quality multi-scale image segmentation. In Angewandte Geographische Informationsverarbeitung XI, Beiträge zum AGIT—Symposium 1999; Herbert Wichmann Verlag: Kalsruhe, Germany, 1999.
- 51. Chan, Y.K.; Koo, V.C. an introduction to synthetic aperture radar (SAR). Prog. Electromagn. Res. B 2008, 2, 27–60. [CrossRef]
- 52. Tang, W.; Liu, W.; Stiles, B. Evaluation of high-resolution ocean surface vector winds measured by QuikSCAT scatterometer in coastal regions. *IEEE Trans. Geosci. Remote Sens.* 2004, 42, 1762–1769. [CrossRef]
- 53. Kilpatrick, K.A.; Podestá, G.; Evans, R. Overview of the NOAA/NASA advanced very high resolution radiometer Pathfinder algorithm for sea surface temperature and associated matchup database. *J. Geophys. Res. Space Phys.* **2001**, *106*, 9179–9197. [CrossRef]
- 54. Kilpatrick, K.A.; Podestá, G.; Walsh, S.; Williams, E.; Halliwell, V.; Szczodrak, M.; Brown, O.B.; Minnett, P.J.; Evans, R. A decade of sea surface temperature from MODIS. *Remote Sens. Environ.* **2015**, *165*, 27–41. [CrossRef]
- 55. O'Reilly, J.E.; Maritorena, S.; Mitchell, B.G.; Siegel, D.A.; Carder, K.L.; Garver, S.A.; Kahru, M.; McClain, C. Ocean color chlorophyll algorithms for SeaWiFS. *J. Geophys. Res. Space Phys.* **1998**, 103, 24937–24953. [CrossRef]
- 56. Esaias, W.; Abbott, M.; Barton, I.; Brown, O.; Campbell, J.; Carder, K.; Clark, D.; Evans, R.; Hoge, F.; Gordon, H.; et al. An overview of MODIS capabilities for ocean science observations. *IEEE Trans. Geosci. Remote Sens.* 1998, 36, 1250–1265. [CrossRef]
- 57. Figueredo, G.P.; Ebecken, N.F.F.; Augusto, D.A.; Barbosa, H.J.C. An immune-inspired instance selection mechanism for supervised classification. *Memet. Comput.* 2012, *4*, 135–147. [CrossRef]
- Passini, M.L.C.; Estébanez, K.B.; Figueredo, G.P.; Ebecken, N.F.F. A Strategy for Training Set Selection in Text Classification Problems. Int. J. Adv. Comput. Sci. Appl. 2013, 4, 6. [CrossRef]
- 59. MDA (MacDonald, Dettwiler and Associates Ltd.). *RADARSAT-2 Product Format Definition;* Technical Report RN-RP-51–2713, Issue/Revision: 1/10; MDA: Richmond, BC, Canada, 2011; 83p.
- 60. Hammer, Ø.; Harper, D.A.T.; Ryan, P.D. PAST: Paleontological Statistics software package for education and data analysis. *Palaeontol. Electron.* **2001**, *4*, 9.
- 61. Sneath, P.H.A.; Sokal, R.R. *Numerical Taxonomy—The Principles and Practice of Numerical Classification*; W.H. Freeman and Company: San Francisco, CA, USA, 1973; 573p, ISBN 0716706970. Available online: http://www.brclasssoc.org.uk/books/Sneath/ (accessed on 30 July 2021).
- 62. Kelley, L.A.; Gardner, S.P.; Sutcliffe, M.J. An automated approach for clustering an ensemble of NMR-derived protein structures into conformationally related subfamilies. *Protein Eng. Des. Sel.* **1996**, *9*, 1063–1065. [CrossRef] [PubMed]
- 63. Zar, H.J. Biostatistical Analysis, 5th ed.; New International Edition; Pearson: Upper Saddle River, NJ, USA, 2014; ISBN 1292024046.
- 64. Rao, C.R. The use and interpretation of principal component analysis in applied research. Sankhyä Indian J. Stat. 1964, 26, 329–358.
- 65. Zhang, D.; He, J.; Zhao, Y.; Luo, Z.; Du, M. Global plus local: A complete framework for feature extraction and recognition. *Pattern Recognit.* **2014**, 47, 1433–1442. [CrossRef]
- Li, P.; Fu, Y.; Mohammed, U.; Elder, J.H.; Prince, S.J.D. Probabilistic Models for Inference about Identity. *IEEE Trans. Pattern Anal. Mach. Intell.* 2011, 34, 144–157. [CrossRef]
- 67. Wang, X.; Tang, X. 2004, Dual-Space Linear Discriminant Analysis for Face Recognition. In Proceedings of the Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'04), Washington, DC, USA, 27 June–2 July 2004; Volume 2, p. 15. [CrossRef]
- 68. Chen, L.-F.; Liao, H.-Y.M.; Ko, M.-T.; Lin, J.-C.; Yu, G.-J. A new LDA-based face recognition system which can solve the small sample size problem. *Pattern Recognit.* 2000, *33*, 1713–1726. [CrossRef]
- 69. Hastie, T.; Buja, A.; Tibshirani, R. Penalized Discriminant Analysis. Ann. Stat. 1995, 23, 73–102. [CrossRef]
- 70. Tharwat, A.; Gaber, T.; Ibrahim, A.; Hassanien, A.E. Linear discriminant analysis: A detailed tutorial. *AI Commun.* **2017**, *30*, 169–190. [CrossRef]
- 71. Legendre, P.; Legendre, L. Numerical Ecology. In *Developments in Environmental Modelling*, 3rd English ed.; Elsevier Science B.V.: Amsterdam, The Netherlands, 2012; Volume 24, 990p, ISBN 978–0444538680.
- 72. Witten, I.H.; Frank, E.; Hall, M.A.; Pal, C.J. *Data Mining: Practical Machine Learning Tools and Techniques*, 4th ed.; Morgan Kaufmann: San Francisco, CA, USA, 2016; 654p.
- 73. Lohninger, H. Teach/Me Data Analysis; Springer: Berlin, Germany; New York, NY, USA; Tokyo, Japan, 1999; ISBN 3540147438.
- Clemmensen, L.K.H. On Discriminant Analysis Techniques and Correlation Structures in High Dimensions; Technical Report-2013 No. 04; Technical University of Denmark: Lyngby, Denmark, 2013; Available online: https://backend.orbit.dtu.dk/ws/portalfiles/ portal/53413081/tr13_04_Clemmensen_L.pdf (accessed on 30 July 2021).
- 75. McLachlan, G. Discriminant Analysis and Statistical Pattern Recognition; John Wiley & Sons, Inc.: Milton, Australia, 1992; 534p, ISBN 0-471-61531-5.
- 76. Aurelien, G. Hands-On Machine Learning with Scikit-Learn and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent System; O'Reilly Media: Newton, MA, USA, 2017.
- 77. Fawcett, T. An introduction to ROC analysis. Pattern Recognit. Lett. 2006, 27, 861–874. [CrossRef]
- Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* 1991, 37, 35–46. [CrossRef]
- 79. Powers, D.M.W. Evaluation: From Precision, Recall and F-Factor to ROC, Informedness, Markedness & Correlation. *J. Mach. Learn. Technol.* **2011**, *2*, 37–63.

- 80. Christiansen, M.B.; Koch, W.; Horstmann, J.; Hasager, C.B.; Nielsen, M. Wind resource assessment from C-band SAR. *Remote Sens. Environ.* **2006**, *105*, 68–81. [CrossRef]
- 81. Bern, T.-I.; Wahl, T.; Anderssen, T.; Olsen, R. Oil Spill Detection Using Satellite Based SAR: Experience from a Field Experiment. *Photogramm. Eng. Remote Sens.* **1993**, *59*, 423–428.
- 82. Johannessen, J.A.; Digranes, G.; Espedal, H.; Johannessen, O.M.; Samuel, P.; Browne, D.; Vachon, P. SAR Ocean Feature Catalogue; ESA Publication Division: Noordwijk, The Netherlands, 1994; 106p.
- 83. Staples, G.C.; Hodgins, D.O. RADARSAT-1 emergency response for oil spill monitoring. In Proceedings of the 5th International Conference on Remote Sensing for Marine and Coastal Environments, San Diego, CA, USA, 5–7 October 1998; pp. 163–170.
- 84. Silveira, I.C.A.; Schmidt, A.C.K.; Campos, E.J.D.; Godoi, S.S.; Ikeda, Y. The Brazil Current off the Eastern Brazilian Coast. *Rev. Bras. De Oceanogr.* **2000**, *48*, 171–183. [CrossRef]
- 85. Brown, C.E.; Fingas, M. New Space-Borne Sensors for Oil Spill Response. In Proceedings of the International Oil Spill Conference, Tampa, FL, USA, 26–29 March 2001; pp. 911–916.
- 86. Brown, C.E.; Fingas, M. The Latest Developments in Remote Sensing Technology for Oil Spill Detection. In Proceedings of the Interspill Conference and Exhibition, Marseille, France, 12–14 May 2009; p. 13.