Marine Debris Analysis

A workflow for identification of submerged debris objects

v. 1.0 (September 2015)

Contents

Executive Summary

The extreme variability of marine debris represents the gist of what has driven our approach in developing the general adopted workflow: after having found robust methods for reliable detections, we combine the outputs together.

The initial common scenario is represented by the availability of multiple layers of information (e.g., information collected through depth measurements or backscatter) and some kind of a-priori information (as the outcomes of a predictive model, or something simpler as a binary land/sea mask derived from the current ENC) that are able to meaningful constrain the general problem.

The developed approach combines the information content of these layers into a probability map that defines the likelihood of there being objects in a particular area. The workflow can be further customized to filter specified areas based on contingent evaluations: for instance, to focus on the marine debris present only in depths shallower that a defined threshold.

The result of such a workflow is a hot spot map of area where the presence of objects if more likely. It is then straightforward to reduce it to a list of point objects of inclusion in the database to be communicated to the outside world.

The structure of the described workflow can be divided in three stages:

- A predictive model, as a way to structuring the problem, to control some of the complexity, and to obtain a general idea on where debris are liked to be
- The detection model which takes care of the core detection aspects of this problem robustly, and
- The data management and exchange, that is relevant to make the outcomes useful.

1. Introduction

The main aim of the "Marine Debris Analysis" research theme is to facilitate the identification of marine debris by developing an approach (and identifying the required tools) that can be used to assess such a presence in the event of a storm.

The achievement of this aim is made more difficult by the extreme variability of marine debris. There has been a lot of work done on marine object detection, mostly from the mine counter measures and pipeline inspection fields, thus several of these of techniques developed in such fields have been modified to been adopted in this research theme [1-3]. However, a common complication that arises in such an operation is that the use cases of these techniques are usually very specific [4]. The specificity of the targets makes the solution to the detection problem relatively easier than in the wider case of marine debris, where targets are much less constrained in shape, material and size. In other words, if you are looking for mine-like objects, the target can be modeled quite well since you roughly know what is that you are looking for. Similarly for a pipeline or for unexploded ordinances. With marine debris, there is not a very good definition of what you are looking for [5].

Marine debris are commonly vaguely defined as being any man-made object discarded, disposed of, or abandoned that enters the coastal or marine environment [6]. They can be made of a variety of materials, directly related to those commonly used by our society: plastics, from industrial products (strapping bands, resin pellets, plastic sheeting) to common domestic material (bags, bottles), as well as other materials (metal, styrofoam, rubber, glass) that, like plastic, have a wide range of uses [7]. They tend both to break down into smaller fragments and to be worn away, but they do not biodegrade entirely. Marine debris span in size from the millimetric size of resin pellets to entire sunken vessels [6]. Thousands of abandoned and derelict crafts are presently close to the shoreline, in a multiplicity of states (e.g., semi-submerged in the intertidal zone, stranding on reefs or in marshes) [8]. When present in protected areas (e.g., lagoons), these shipwrecks may persist for years; while exposed environments can force their disintegration, and generated litter may be distributed widely through multiple habitats [7, 9]. Marine debris are commonly classified by source type into ocean-based and land-based [6]. The primary ocean-based sources are vessels (e.g., derelict fishing gears such as nets, traps, buoys, lines lost by commercial fishing vessels and recreational boats, entire containers from cargo ships in rough seas), but stationary platforms also play an important role since all items lost from these structures become litter (e.g., hard hats and gloves, storage drums). However, the largest part of debris along the shoreline comes from land-based sources [6].

Debris may be blown, swept, and washed out to sea after their accidental or intentional disposal as domestic or industrial wastes on land or in streams. In normal conditions, rain and snow-melting waters are the typical means by which these materials are carried to a nearby river or canal, or even directly to the ocean. Natural disasters such as hurricanes,

tsunamis, mudslides and floods are usually coupled with devastating effects such as heavy rains, flooding, strong winds, high waves, and storm surges [10, 11]. The direct result on marine debris is a peak rate of new depositions (in both intertidal and subtidal areas), an additional problem for regions already impacted by a natural disaster. In fact, this abnormal amount of marine debris may create threats to navigation, fishing activities, recreation, sensitive ecosystems, and generally to the environment and human safety [6, 11]. Effectively and quickly processing large amounts of hydrographic data collected using commercial systems for detection of marine debris would be highly advantageous to the necessary remediation operations [12].

A coherent framework for such a challenging task must take into account a series of different aspects and as much available information as possible, from the spatial relationship between the marine debris distribution and explanatory variables to an effective way to merge the results of different detection algorithms, required as a consequence of the vague definition in size, shape and material of the investigated targets. At the same time, some degree of modeling and approximation to make the analysis computationally attractive and sufficiently effective in practice is also required [13, 14]. Our workflow begins by analyzing the marine debris distributions in recent available data sets, from which a possible predictive modeling was outlined. The predictive step provides the initial state for a Bayesian spatial hierarchical model, but can also be used as a post-hurricane survey planning tool. For the detection, a target model was built postulating a simplified description of the debris properties, and a set of detection algorithms were specifically targeted to different possible characteristics of marine debris, detecting discrete objects which differ (e.g., protrude) from the surrounding seafloor, being close or connected to the bottom [12]. The scope of these algorithms was constrained to analyze products commonly available in existing post-processing software (mainly, bathymetric digital terrain models and backscatter mosaics with several associated data sets, such as statistics derived from the core data, or during construction) so that the technique may be quickly inserted into existing workflows, which eases resource management in a response situation. We provides an example of detector outcomes, based on survey data collected after Super Storm Sandy, to evaluate the effectiveness of the proposed technique. In concluding, we discuss a possible approach to marine debris data management and exchange.

2. Case Studies and Data Sources

Super Storm Sandy was a natural disaster with unusual characteristics [15]. Started as a classical late-season hurricane in the Caribbean Sea, a complex evolution made Sandy grow considerably in size, moving parallel to the coast of the United States (Figure 1). After turning northwestward over much cooler waters, Sandy weakened and started to lose its tropical characteristics about 45 nautical miles off Atlantic City, becoming an extra-tropical cyclone (that is, relying mainly on baroclinic processes), to make landfall near Brigantine, NJ, around 7:30 p.m. on October 29, 2012. Because of Sandy's unusually large size, the New Jersey and New York coastlines were hit by a catastrophic storm surge, accompanied by powerful damaging waves and enhanced by the fall full moon period. The impacts of Sandy were widespread, with at least 650,000 houses damaged or destroyed, cars tossed about, boats pushed well inland from the coast [15]. Sandy represented a massive source of marine debris deposition for impacted coastal areas. However, being relatively close in time, Hurricane Irene most likely also influenced the debris distribution [16].

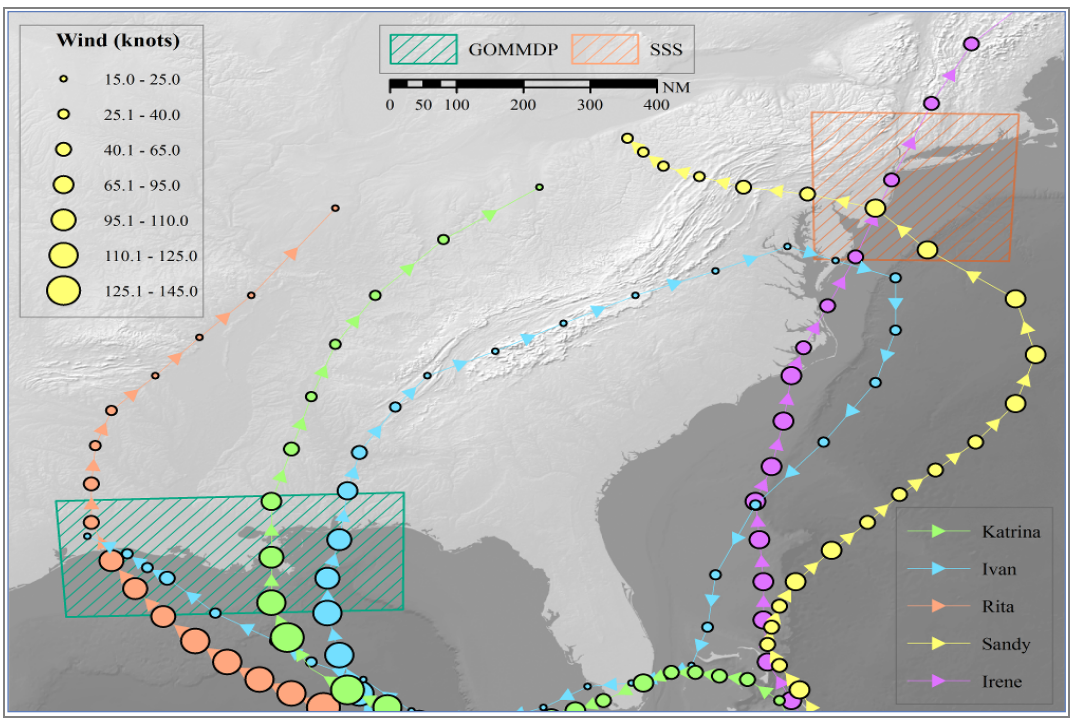

Figure 1 – Study areas, and the best tracks (associated with peak winds) of the five hurricanes of interest.

Two marine debris data sets available in the Sandy area are used in the present work:

- The NOAA Marine Debris Program (MDP) dataset, mainly focused on intertidal coastal areas (hereafter, SSS-ID), which was based on NOAA NGS imagery acquired during post-storm overflights, and follow-up shoreline survey [17]. The data set is made up of almost 70,000 debris records: 52% identified via automated Object Based Image

Analysis (OBIA) [18], and the remaining targets marked via manual heads-up digitization by imagery analysts.

- NOAA's Office of Coastal Survey (OCS) dataset, which is a growing collection of subtidal marine debris (SSS-SD), mainly based on several surveys using a variety of acoustic sensors (e.g., Multibeam Echo Sounders, Side Scan Sonars) performed by contractors. This preliminary data set (some processing is still ongoing) was retrieved directly from NOAA OCS.

The third analyzed data set is the one collected by the Gulf of Mexico Marine Debris Project (GOMMDP), a large collection of marine debris items related to Hurricanes Katrina, Rita and Ivan, and identified via side scan sonar during several surveys conducted in 2006 [19]. There are several relevant differences between the GOMMDP and the Sandy areas. The former project area is quite flat, while in comparison the latter is characterized by moderate bathymetric and topographic relief (Figure 1). The complex geography of New York Harbor and Long Island Sound generated quite distinct patterns for the maximum elevation of storm surge and peaks of strong winds; while they are almost coincident in the area impacted by Katrina. Another relevant difference is that Sandy affected one of the most densely populated areas along the East Coast, while much of Katrina effects were on rural or low-density suburbs [15, 20].

These three data sets were used for the predictive step, while survey data collected after Sandy, in the Redbird Reef area, DE and Jamaica Bay, NY, were used to test the fusion approach for the detector.

3. Predictive Model

The predictive model was developed to provide a mean by which we can spatially constrain the marine debris detection problem. Although problems of such a type have notoriously multiple possible solutions, the main idea here is to just identify a likely solution that is used as starting point for the following Bayesian inference. In other words, we are looking for obtaining a rough idea on where we think that is likely to might be marine debris.

Based on the data availability of previous hurricane events, we studied the correlation between several possible explanatory layers collected in the wake of recent Gulf of Mexico's and East Coast storms (like Katrina and Sandy) and the likelihood of generation and deposition of marine debris. As possible informative layers we used data related to the track and the intensity of the storm, to the human environment in the area (as sources of many marine debris), and to the size of the effects (significant wave, storm surge, and so on) [21]. The combination of these layers provided a mean to obtain a rough prediction, for a given storm, on where there is a larger likelihood of being significant marine debris creation [21].

The outcomes of the predictive model become very useful when we have detection algorithms that throw up a number of false positives. In such a case, we use the model results to tune back such a particular algorithm, and thus applying better constraint to the overall solution to the problem.

3.1. Spatial Analysis Tools

Spatial data are often characterized by a phenomenon, known as spatial auto-correlation (SAC), that occurs when the values of variables sampled at nearby locations are not independent from each other [22]. In marine debris data, SAC may arise from a multitude of possible causes, both from intrinsic processes, such as debris size and target-seafloor interactions, and in response to unknown (or partially known) environmental drivers, e.g. non-linear relationships between predictors and dependent variables modelled erroneously as linear, or failure in accounting for an important environmental determinant that is itself spatially structured [23]. SAC often poses serious shortcomings for hypothesis testing and prediction by violating the assumption of independently and identically distributed errors required by most commonly-used statistical procedures [24]. In the absence of a perfectly correctly specified model, SAC cannot be accounted for by non-spatial models [25], and some kind of correction, such as the one introduced by auto-covariate regression, is required [26]. However, SAC may also be seen as an opportunity since it provides useful information for inference of a process from patterns [27].

The distribution of marine debris density over a seafloor area is a spatially-distributed stochastic phenomenon. The density values represent a set of variates, and we want to decide whether there is any evidence that these variates are spatially correlated. However, real scenarios such as the case studies usually present a quite complex hierarchical structure

that cannot be simply modeled as regular or clustered point processes, and they may exhibit different spatial pattern characteristics at different scales [28]. The verification of the hypothesis that the debris locations tend to cluster rather than following a complete spatial randomness (CSR) process, under which their patterns are realizations of a Poisson point process, is thus an important preliminary step. Dedicated methods can be classified as global (throughout the whole study region) or local. As global methods, there are several established procedures (global Moran's I plots, Geary's C correlograms and semi-variograms), in which a measure of similarity or variance of data points is plotted as a function of the distance between them, to check whether spatial correlation is likely to impact the data analysis [29]. Moran's I varies between *–1* and *+1*, and a value close to *0* indicates a random pattern or absence of spatial autocorrelation [30]. Calculation of the *z-score* provides a means to evaluate the intensity of spatial clustering, looking for clustering lags with statistically significant peak *z-scores* (based on a randomization null hypothesis). Given the kind of pattern of this study, the focus is on a quite local level of clustering, and this method allows the correct size of the analysis to be determined from the data. Local methods such as Local Indicator or Spatial Association (LISA) and hot-spot analysis, when used in conjunction with global Moran's I, deepen the knowledge of the processes that give rise to spatial association, enabling the detection of local pockets of dependence that may not show up when using global methods [31]. The Local Indicator of Spatial Association (LISA) is a local Moran index proposed by Anselin [32]. This method highlights clusters as well as possible outliers, a low value surrounded by high values (low–high) or a high value surrounded by low values (high–low). The hot-spot analysis identifies local spatial clusters of statistically significant high (hot spot) and low (cold spot) number of debris for a grid cell by calculation of the Getis-Ord Gi* statistic [33].

The popular K-means algorithm was also adopted as an exploratory tool, applying a cluster analysis to identify possible structure between the available data in the absence of category information [34]. The selection of the number of groups was based on the pseudo F-statistic, a ratio reflecting within-group similarity and between-group differences [35].

3.2. Model Implementation

SAC occurs at all spatial scales (from meters to dozens of kilometers) for a whole suite of reasons. Since these reasons are mostly unknown, one cannot readily derive a spatial correlation structure for an entirely new and unobserved area, although it is possible to derive predictions by interpolation for missing data within the study area (i.e. by using a Gibbs sampler) [36, 37]. When models are projected into different areas the handling of spatial auto-correlation is quite problematic, sometimes even impossible. Extrapolation in space can only be based on the coefficient estimates, not on the spatial component of the model [13, 30]. Extrapolation is further complicated by model complexity: the use of non-linear predictors and interactions between environmental variables usually increases model fit, but with the price of compromising model transferability in time and space. Given such

intrinsic limitations, a predictive model was created with the main aim of providing a reliable initial state to the Bayesian detector; however, the evaluation of areas where the presence of marine debris is more probable, coupled with considerations related to the economic relevance of specific waterways, may also help to effectively prioritize the areas to survey and the data to process, speeding up identification of detected possible targets and their eventual removal. A limited number of predictive debris models have been recently developed: the web-based HurDET model [38]; USACE HAZUS-MH models [39, 40]; and, the Marine Debris Distribution model of the Gulf of Mexico, from the NOAA Marine Debris Program (MDP) [11]. While the first two are focused on terrestrial debris, some of the findings and the results from the MDP model have been adapted and incorporated in the present model, attempting both to generalize the approach (to increase flexibility) and to cast the modeling results to feed a probabilistic algorithm for marine debris detection.

A list of possible marine debris predictors (Table 1) was created following an intuition criterion, coupled with comparison with existing similar work and data availability. This latter criterion is for model flexibility, to avoid a model based on peculiar predictors that, although highly explanatory of debris presence, will not be likely available in case of natural disasters in different areas, as well as in the immediate proximity of the event. The selected predictors can be clustered in two groups: related to the storm energy (wind, storm surge, bathymetric profile, etc.) and capturing the spatial distribution of debris sources (concentrations of highly populated human areas, waterways, etc.). The intuition behind such a choice is that coastal urban areas impacted by high storm energy should have larger amounts of anthropogenic debris, due to higher potential for debris creation and mobilization.

Table 1 – Proxy Source for Model Predictors

Data Product	Predictor	Source
H*Wind Surface Wind Analysis	Wind	NOAA
Experimental Extratropical Surge and Tide Operational Forecast System (ESTOFS)	Storm surge	NOAA
NGDC 3-Arc Second Coastal Relief Model	DEM	NOAA
International Best Track Archive for Climate Stewardship (IBTrACS)	Hurricane best-track	NOAA
Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG)	Shoreline	NOAA
Natural Waterway Network	Waterway	USACE
Topologically Integrated Geographic Encoding and Referencing (TIGER)	Population	CENSUS

The wind-based predictor was created by summarizing into a single layer, only representing the peak intensity, the surface wind analysis of tropical cyclones produced by the NOAA Hurricane Research Division as part of the H*Wind Project [41], a project that merges a variety of coastal and inland data from land, space, and marine platforms. Data from the Extratropical Surge and Tide Operational Forecast System [42], a new generation hydrodynamic modeling system using the ADvanced CIRCulation model, was used as storm surge predictor. The NGDC 3-Arc Second Coastal Relief Model [43], integrating bathymetric

and topographic information from a variety of data sources, was used to explore the relationship between debris and depth. The best tracks for the hurricanes in the study have been retrieved from NOAA's International Best Track Archive for Climate Stewardship (IBTrACS) project [44], collecting the historical tropical cyclone best-track data from all available Regional Specialized Meteorological Centers (RSMCs) and other agencies [45]. The distance of each cell in the lattice grid has been calculated having as reference the World Vector Shorelines (WVS), present in the Global Self-consistent, Hierarchical, High-resolution Geography Database (GSHHG) [46], while for the waterways the USACE Natural Waterway Network was adopted [47]. Finally, a population index was derived from the 2010 census Topologically Integrated Geographic Encoding and Referencing (TIGER) product [48].

A classic Ordinary Least Squares (OLS) [30] was adopted to create an equation relating marine debris density (dependent variable) to the selected set of explanatory variables. Exploratory regressions were used to find a good specified model by evaluation of all the different possible combination of explanatory variables, balancing statistical significance, redundancy and multi-collinearity [13, 14, 30]. Such a balance was mainly evaluated by comparison of Adjusted R-squared, Akaike information criterion (AIC) and Variance Inflation Factor (VIF) values [13, 14].

3.3. Exploratory Analysis of Available Data

The global Moran's I statistic, which is used to test the null hypothesis that the spatial autocorrelation of a variable is zero, was applied to the study case data sets. The results are presented in Figure 2, with the first peak of *z-score* plotted as a red dot.

The results from LISA and hot-spot analysis are presented as Pane A and B in Figure 3 for the GOMMDP data, in Figure 4 for the SSS-ID data, and in Figure 5 for the SSS-SD data; while Pane C was used to spatially identify the areas resulting by the grouping analysis, performed using K-mean algorithms and based on the pseudo F-statistic to identify the parameter for the number of groups: two for GOMMDP and SSS-ID data sets, three for SSS-SD data set. For this last, a plot with the resulting inter-group relationships is presented in Figure 6.

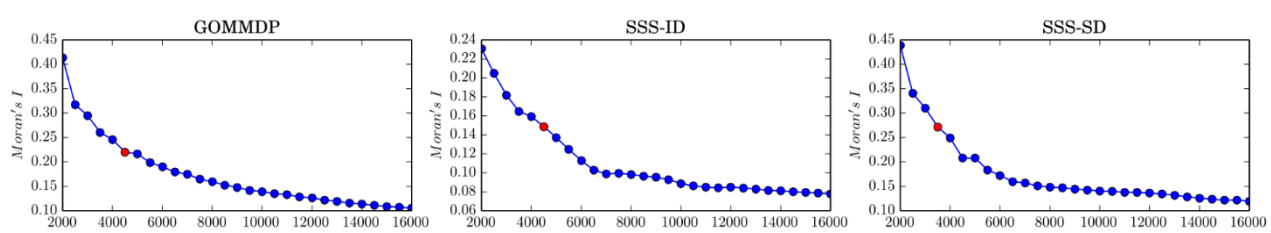

Figure 2 – Global Moran's I statistics (with, in red, the first peak of the associated z-scores) calculated for the case study data sets.

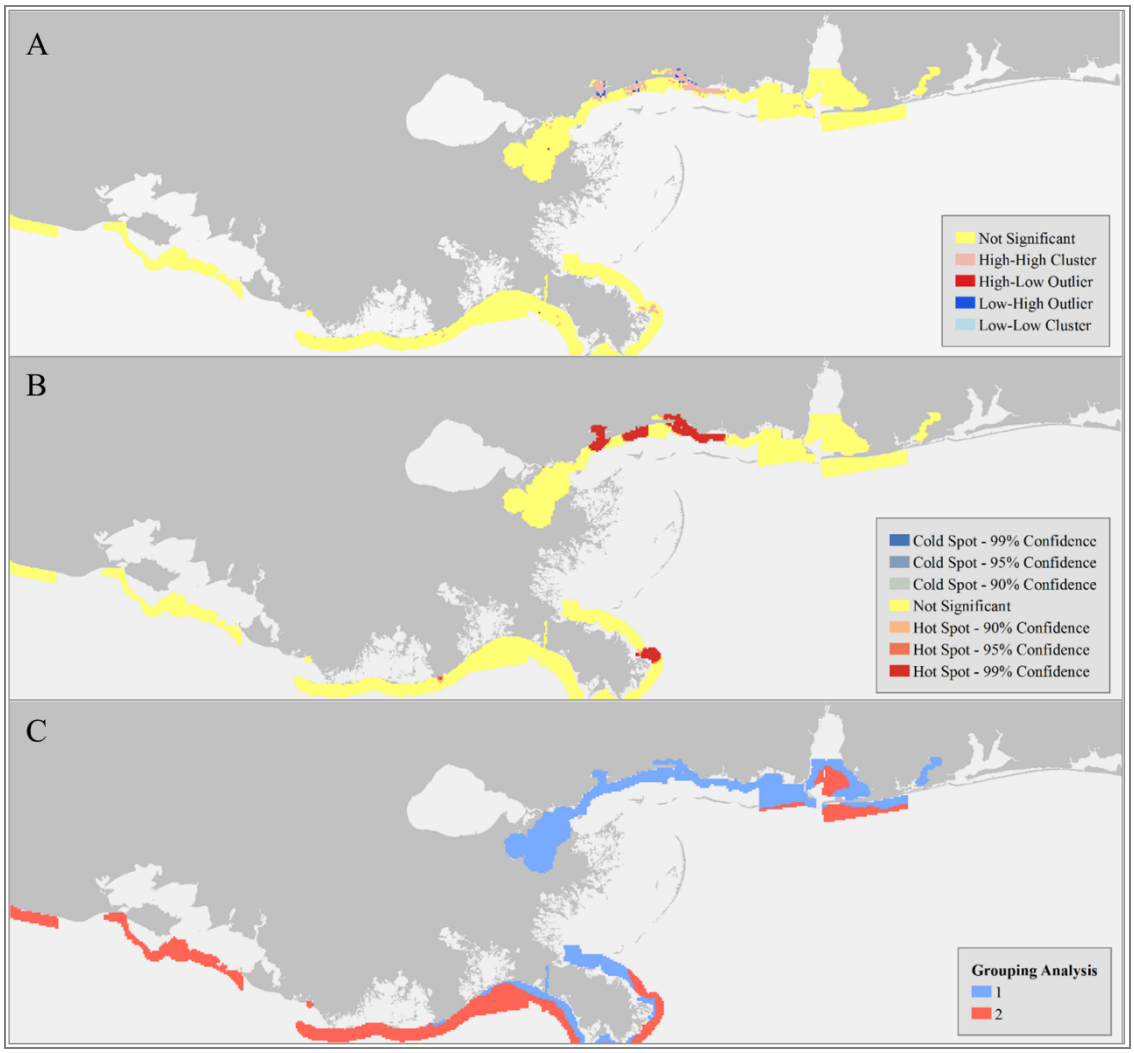

Figure 3 – Local Moran's I statistics (pane A), hot spot analysis (pane B), and grouping analysis (pane C) for GOMMDP data set.

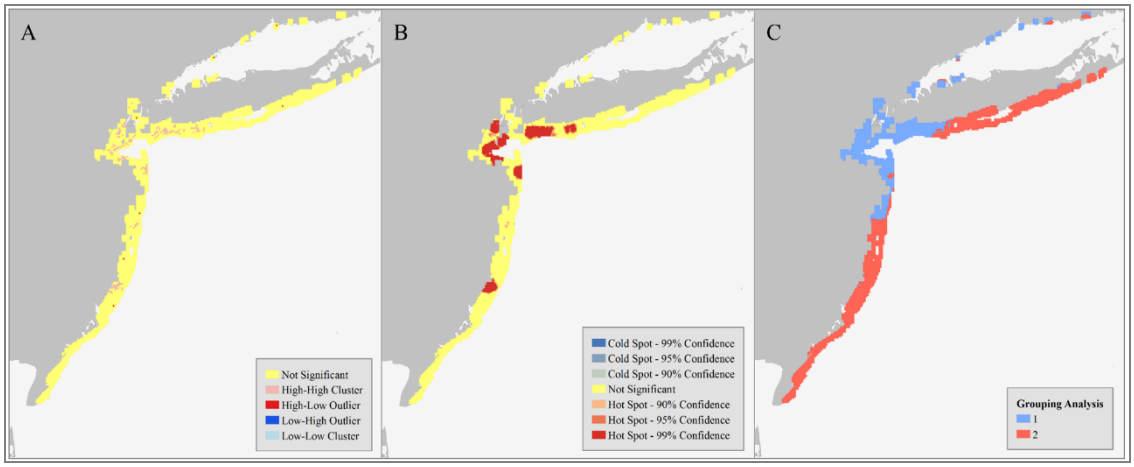

Figure 4 – Local Moran's I statistics (pane A), hot spot analysis (pane B), and grouping analysis (pane C) for SSS-ID data set.

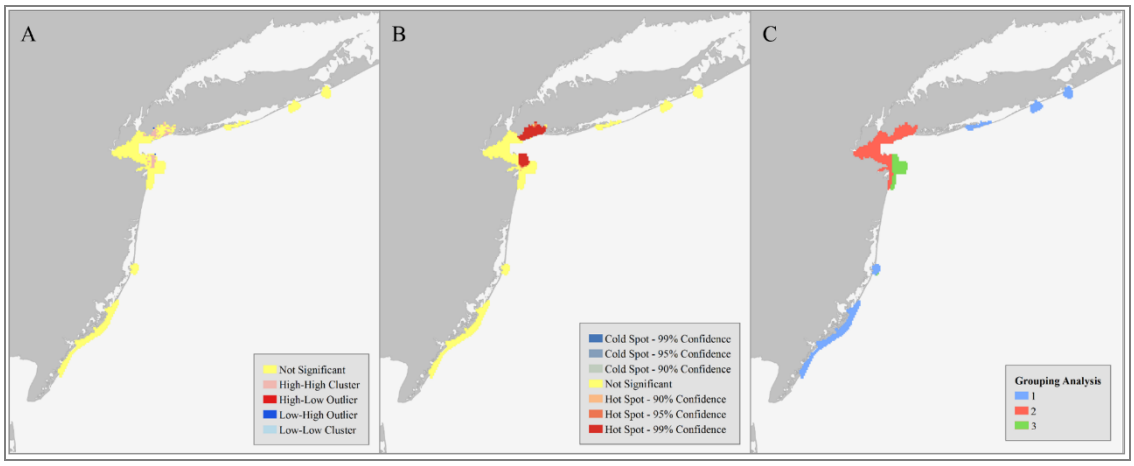

Figure 5 – Local Moran's I statistics (pane A), hot spot analysis (pane B), and grouping analysis (pane C) for SSS-SD data set.

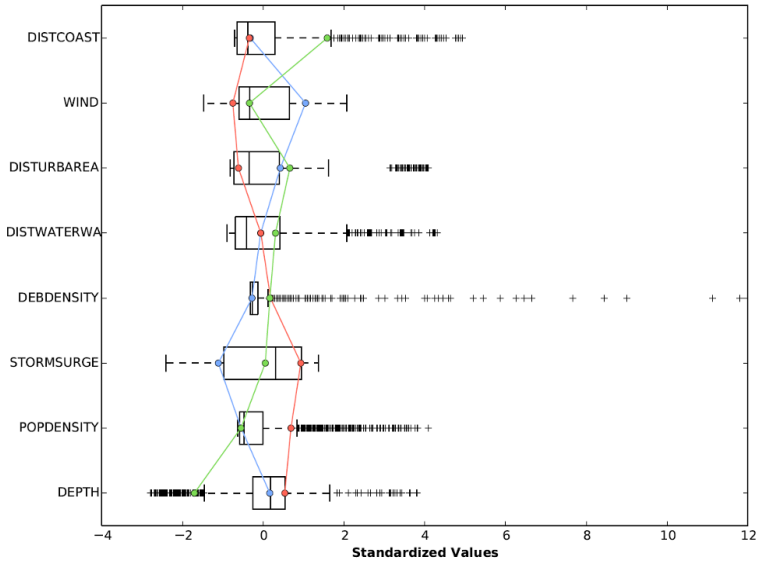

Figure 6 – SSS-SD grouping analysis outcomes between debris density ("DEBDENSITY") and predictors: distance from the shoreline ("DISTCOAST"), maximum wind ("WIND"), distance from the urban area ("DISTURBAREA"), distance from waterway ("DISTWATERWAY"), maximum storm surge height ("STORMSURGE"), population density index ("POPDENSITY"), and average depth ("DEPTH").

The resulting global Moran's I values show a similar trend for the GOMMDP and the SSS-SD data sets, with peaks at 3.5 and 4.5 km respectively, but lower values in the SSS-ID (although with a consistent peak at 4.5 km). Such a difference could be justified by the fact that restoration efforts are more likely to have been applied on the intertidal zones due to their accessibility. The Getis-Ord Gi* test identified several clusters of points that have higher values than expected by chance. For the GOMMDP dataset, the analysis results suggest a relationship between the hurricane Katrina track, the urban areas and hot-spots of marine debris density (Figure 3). Several hot- and cold-spots are also present for the Sandy datasets (intertidal debris in Figure 4, and subtidal debris in Figure 5), although relationships with

tracks and urban areas are less visually evident. The results from the hot-spot analysis are compatible with the LISA outcomes, and only a very limited number of possible outliers are present. An interesting result from the grouping analysis is the different number of groups (three rather than two) for the SSS-SD data set. As is clear in Pane C of Figure 5, the anomaly is spatially localized (the results of the grouping analysis are expanded in Figure 6) and almost exactly matches with an area assigned to a specific contractor. Thus, it is possible to speculate that this is most likely due to different criteria followed in the target detection analysis (most strict). Consequently, the debris distribution in this area was clipped from the remaining analysis (although it is worthwhile of additional future investigation). The evaluation of whether a debris distribution exhibits a spatial clustering pattern implies the question if elevated debris rates arise simply by chance and thus have no predictive utility. Since the case studies show statistically significant patterns, marine debris density warrants study since it can be used, e.g., to identify hot spots that can be used by post-hurricane survey planners to prioritize and target the data collection on limited areas.

3.4. Predicted Debris Distribution

After an exploratory regression based on the seven predictors listed in Table 1, both study areas obtained similar results indicating that storm surge, population density index, and distance from urban areas are quite good predictors of marine debris presence (R-squared higher that 0.5). The addition of other available predictors does not provide significant contributions to the R-squared values, adding decrements in terms of AIC and VIF values. Using the resulting model parameter coefficients, a prediction of debris distribution for the areas covering is presented in Figure 7, for the SSS-SD data set, and in Figure 8, for the GOMMDP data set.

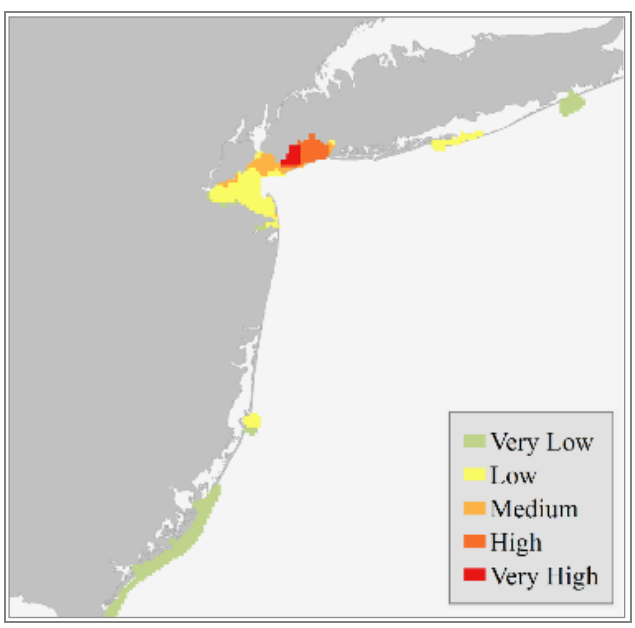

Figure 7 – Predicted distribution density of marine debris in the SSS study area.

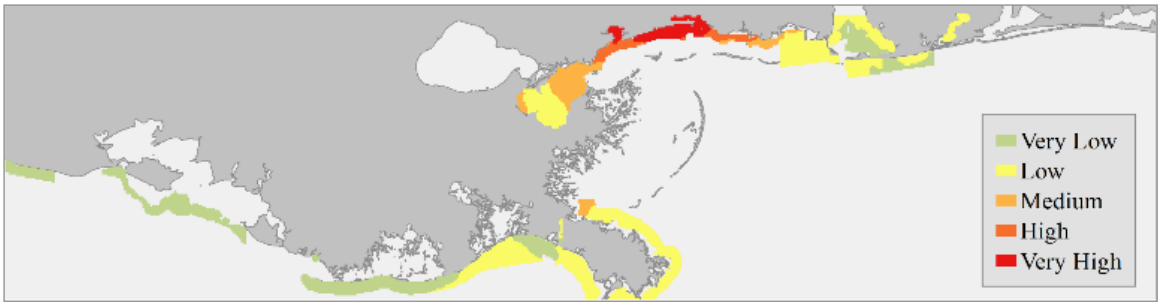

Figure 8 – Predicted distribution density of marine debris in the GOMMDP study area.

The model outcomes highlighted higher likelihood of debris presence in areas that received during the hurricane both higher wind energy and storm surge water elevations, and that are proximal to more developed and populated urban areas. The predictive model was developed thanks to the availability of several data sources, variously related to marine debris, collected and made publicly available after hurricanes Ivan, Katrina, Rita, Irene and Sandy. The analyzed debris data sets represent two case studies, with distinct peculiarities. The future addition of similar rich databases will surely help to obtain a clearer picture of how individual characteristics of hurricanes interact with human land use to lead to various types and degrees of marine debris deposition. In fact, this deposition may or may not occur, or occur to varying degrees, depending upon individual hurricane characteristics (e.g., category, breakpoint, maximum wind speed, height of storm surge, and path after landfall). Landfall in a populous area, a post-landfall trajectory upriver toward a headwater region, or a relative slow speed of hurricane passage are just some of the many variables that, leading to increased damages, can carry to a massive creation of marine debris. The different intensities and tracks for the hurricanes affecting the study areas are clear examples of how each storm-like event has its own peculiarities. A practical difficulty in common to these and similar data sets is the complexity involved in distinguishing between anthropogenic and natural, storm-generated and pre-existing targets. In order to make the analysis possible, it was assumed that each of them provides a representative picture of the distribution of storm-related and anthropogenic marine debris. Furthermore, the model also relies on the assumption that the distribution of marine debris objects is essentially static, even months after the event occurred, in order to use data sets collected after a variable amount of time since the event. This is increasingly likely to be faulty in function of the time passed since the event or, for instance, in case of another storm-like event occurred. Furthermore, the seven adopted predictors do not directly capture all the possible causes of concentrations of marine debris: for instance, areas with particular activities (e.g., recreational marinas), specific land use in the neighborhood such as dumping areas, and strong energy impacts of wave run-up effects. Additional predictors capturing such causes might be evaluated in future studies, with the challenging criterion to maintain model's spatial portability.

4. Detection Model

Given the complexity of marine debris objects does not have a very good definition, a slightly more sophisticated approach than the ones usually applied in related fields of research was required (Figure 9). The overall idea here is that we use a selection of detection algorithms, each one not necessarily reliable when taken stand-alone, but that together provide a more reliable output: a quite robust indication on where marine debris should be located.

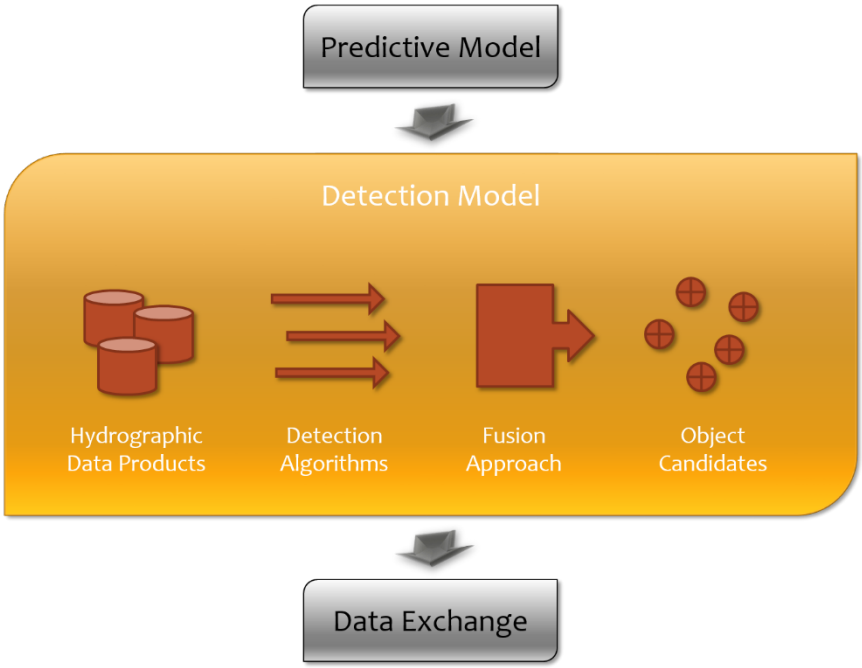

Figure 9 – The three connected sub-problems that have been investigated for the marine debris analysis theme. The core topic of the research is represented by the detection model (in orange). It takes hydrographic data products as input and provides a list of marine debris candidates based on the fusion of the outcomes of several detection algorithms.

During the development of such a detection model we had to provide answers to three main issues:

- What products use as data input (and their collection requirements)
- How doing the initial detection (i.e., selection of detection algorithms)
- How fusing together the outcomes of different algorithms

4.1. Data Input and Related Collection Requirements

About the product inputs, we took the approach to only use hydrographic standard products. This means that our approach will not require the acquisition of particularly 'exotic' data. This is mainly based on the consideration that the approach is designed to work in an event of a storm-like scenario, when a disaster response is in act, and the users of such an approach will not have time and resources to do nothing particularly fancy. In a

stress situation where the priorities are to provide supplies to people and to repair damaged infrastructures, we cannot require the acquisition of, for instance, magnetic data. Most likely, we have to deal with uncomplete and far from be perfect acoustic data acquired with quick-to-deploy devices. We thus focused on use of backscatter and bathymetry, products that we get conventionally out of a survey anyway, as source of our useful information.

However, analysis of the requirements for the techniques adopted by the workflow to be successful has led to some caveats on the data being used, and the processing being applied [12]. In particular, the acoustic system used should be fully understood, with particular attention to the internal backscatter processing, otherwise the appropriate corrections cannot readily be made. In addition, the system used to collect the data should be calibrated, and the resulting calibration parameters correctly applied (in real time or post processing), or the results may be misleading. Finally, the environment should be properly characterized (e.g., absence of issues with the sound speed profiles, correct absorption coefficients, etc.) for the corrections being done to be effective. Missing one or more of the above requirements may affect the developed algorithms, at least with respect to performance, emphasizing the need for careful survey planning and management.

System calibration is mainly required by the fact that elements in the receive array do not usually have absolutely identical characteristics or mounting position [49]. The resulting differences in magnitude and phase must be taken into account to relate the received data to absolute values of backscattering strength with the level of accuracy required by adopted algorithms based on a physical model. Similarly, any signal distortion on the backscatter time series collected around the seafloor detection point should be reduced / avoided.

At the same time, in case of availability of pre- and post-disaster datasets, relevant insights may come from a comparative approach. However, proof that the seabed changes (e.g., presence of marine debris) are not related to instrumental and integration artifacts requires confidence in the absolute accuracy of both the bathymetric and backscatter output of the integrated sonar system [50, 51]. The relevance of the accuracy tends to increase with the reduction of marine debris size (e.g., spatial scale of decimeters), and it often lies at the limit of many acoustic systems used to collect disaster-driven datasets. A possible consequence of the described situation is the appearance of features in a dataset not present in the pre-disaster products due to a better 'focusing' of the used instrument (e.g., higher operational frequency), in case different systems are in use, or to different settings (e.g., operational modes) where the same system is in use.

More generally, important information that should be properly evaluated both in the selection of the system for data collection and in the assessment of the workflow results is the achievable resolution of a specific system, and its variability as a function of the settings selected in the field. In fact, even if specific performance is theoretically achievable based on the manufacturer product specifications, all systems have fundamental constraints and

trade-offs that are a function of operational frequency, resolution, and range of transmission [51].

The resolution directly influences the size of detectable marine debris. In case of a MBES, the resolution is strongly related to the beam footprint which is characterized by the transmit (along-track resolution) and receive beam-width (across-track resolution), and the equivalent length of the transmitted pulse, properly projected on the seabed [52]. In case of detection based on phase measurement, features can be discerned at lower grazing angles at a scale significantly finer than the beam footprint dimension [50]. Another factor that influences the minimum size of detectable marine debris is represented by the beam forming approach. The approaches most commonly in use are equiangular, equidistance, and high definition modes. The last of these, although not implemented by all sonar manufacturers, usually represents the best trade-off among the beam spacing in the nadir region and the required additional detection solutions at lower grazing angles (where multiple soundings are defined within a single beam) [53, 54]. Alternatively, a similar result can be obtained by increasing the number of beams in equidistance mode; another common solution to the same issue is having multiple pings in water at the same time (with slightly different frequencies) [50].

In general, it is highly desirable that the data density within all the acquired dataset is uniform, both in the along- and in the across-track direction. Irregular sounding density coming, for instance, from not properly compensated pitch and yaw may result in undetected features. A 'full sea floor search' of the survey area should be not simply based on the assumption that everything within the bounds of the edge of the swath is 'covered' [55]. In fact, a lack of local data density can drastically reduce the reliable detection of small targets that is based on the assumption that the seafloor is sampled at a scale significantly finer than the target dimension to be resolved [56].

The common target of maintaining three swaths on any given target is a useful rule of thumb, and more swaths per target should be maintained where possible. Assuming roll and pitch stabilization (offered by almost all manufacturers) a yaw-stabilized MBES system may be advantageous where available. In fact, given a 1-meter cube as the assumed lowest bound for debris size, the along track spacing among swaths in shallow waters can require particularly low speed for small boats, usually characterized by higher yaw rates than larger vessels at low engine regime [56]. Simply increasing the data density does not necessary imply better data quality, but it often provides a wider margin for data filtering and statistic tools application.

The data density along-track for single-ping MBES system is mainly controlled by the two way travel time required by the outermost area of each swath to be received. The main implication is that any attempt to improve the swath coverage reduces the ping rate (and then the along-track distance between each ping increases).

MBES along-track beam-width is usually much wider than that used by conventional sidescan sonars (SSS), so that SSS imagery tends to be better quality [57]. However, unless the SSS is hull-mounted (which has its own difficulties) variable distortions are usually introduced due to the uncertainty in the towing fish position and weakness of the flat-seafloor assumption. A better solution, when feasible, is to integrate accurate MBES bathymetry and high resolution SSS imagery. In such a case, the SSS-based mosaic can also take advantage of being properly geometrically corrected by using the MBES-based DTM.

Together with the resolution, it is also important to reduce all the possible sources of Total Propagated Uncertainty (TPU) and absolute accuracy. Since a large part of the adopted algorithms are based on products that combine different survey lines (e.g., mosaic, DTM), the areas of overlap (that is, each node whose final value is based on the integration of data coming from more than a single survey line) will be variously affected by any introduced 'corruption' (e.g., ray tracing with incorrect sound speed profiles, inaccurate tide reduction, loss of GPS differential corrections, time delays between the different sensors in use) with the double risk to mask the presence of marine debris (defocusing) and to create false detections driven by artifacts. The adoption of commonly used patch test procedures before and after the survey (as well as after any variation in the vessel configuration) usually helps to reduce and track many of the possible issues [58, 59]. However, there could be residual misalignment or mistiming of sensors relative to each other that may produce both static biases and dynamic residuals, called wobbles [60]. This latter can be confused with or mask the presence of marine debris.

Similar issues arise comparing products built with the same identical parameters (e.g., grid spacing for DTM), but with different uncertainties and accuracies. In such a case, only scales of seabed change larger than the combination of the accuracies characterizing the compared surveys will become detectable [50].

The characteristics of the water column are continuously changing both in time and in space [61]. As a consequence, there is not a simple direct relationship between the time since, and the distance from, the sound speed measurement in use. The measurements of sound speed must be taken often enough to capture both the actual spatial and temporal variability [62]. If an underway profiler is available, an adequate sampling interval should be adopted [63].

As an additional consideration, it is of overall importance to know and/or have experience with the adopted system so that the best settings will be adopted for target detection. In fact, many manufacturers have specific bottom detection algorithms and operation modes for this type of survey where the requirements are different than a standard bathymetric survey [64].

The workflow will take advantage, when available, of well-collected and calibrated hydrographic data. Nonetheless, thanks to a variable system of weights for the available algorithms (in fact, some of them are less affected than others by improper data

acquisition), marine debris detection will still be possible, with expected increased false alarm rates, even in the event of lack of some of the described best practices.

4.2. Detection Algorithms

4.2.1. Selection Criteria

A comparison of the SSS-SD data set and the original survey data was used to estimate the criteria used by the analyst for defining the presence of a possible marine debris. From the analysis of the targets selected so far within the SSS-SD data set, several common selection patterns emerged [12]. For instance, a first group containing a rounded shape and/or a jump in seafloor reflectivity was common to many of the several hundred targets examined. A second group was based solely on bathymetry evaluation. A third group comes from an integrated analysis of the Digital Terrain Model (DTM) and the acoustic backscatter. Finally, although such data can now be readily collected on many systems, there are no examples of debris selection based on water column data, although the extent of availability of this data (and appropriate tools) to the observers is unknown. It appears that operator debris detection was mainly based on the bathymetry and the reflectivity of the seafloor, assuming any deviation from the 'natural average background' as hints of possible debris. From that consideration, and given the intrinsic complexity of the targets, it is likely that a single algorithm will not be successful for robust marine debris detection. The proposed solution is therefore based on multiple algorithms to process different sources (mainly bathymetry and backscatter from acoustic systems, but easily extendable to water column data, as well as lidar data), fused together so as to be adaptive to the environment, the context, and *a priori* knowledge (if available) of the possible targets. The goal was to use a collection of algorithms working at different levels (e.g., through per beam, single swath, snippet and pixel level operators), which are then fused by the core engine. One of the primary advantages of this approach is operating over different data with independent algorithms can reduce inter-algorithm cross-correlation and therefore the probability of false alarm [12].

Another consideration that has driven the development is that the problem of the marine debris detection is usually made more difficult from the large areas without any debris that can drive the creation of large number of false positive. In other words, the scenario could be described as the famous problem to a find a (very small) needle in a (very big) haystack. What we have done in our approach is not to necessarily inventing new algorithms to do the object detection (since there is a lot out there that we can pull from), but what we are actually doing is to provide extra structure. We have adopted an algorithm to segment the area under analysis into smaller (but significantly related) seafloor patches, and then to calculate for each of these small areas a vector of meaningful metrics that captures various aspects of the backscatter. In essence, we are building a series of haystacks, smaller than the big haystack, so that we can look for debris in smaller areas. This provides a mechanism to adapt the context (the local conditions of the data that we are attempting to deal with)

at each particular seafloor patch. This makes the problem a lot simpler because we do not have to deal with the global problem, dealing instead with a much more localize problem.

4.2.2. Backscatter-based Algorithms

For the backscatter mosaic, as in many existing algorithms, target detection is based on the observation that denser material (often anthropogenic) makes debris returns much stronger than the surrounding background. However, the often used approach to object detection through simple thresholding (e.g., based on the premise that on a mosaic the object return is brighter than the background) was modified since it tends to fail when the background is textured (i.e., detectors are not aware of image correlation). This algorithm is based on an acoustic backscatter mosaic, and takes advantages of previous NOAA-sponsored work at the Joint Hydrographic Center to properly geometrically and radiometrically correct the collected data [12, 65]. The resulting mosaic is segmented into areas with similar reflectivity values through a clustering analysis, and a histogram of backscatter values as a function of angle of incidence is then computed for each clustered area (effectively forming a bivariate histogram). A simple Bayesian classifier is subsequently used to identify areas in each segment where the statistics of a small window do not match that of the overall background distribution (as characterized by the appropriate marginalization of the histogram constructed previously). Areas of low probability of background membership are identified as potential marine debris. Subsequent edge detection and hierarchical filtering are applied to remove misdetections along the mosaic boundaries (Figure 10) [12].

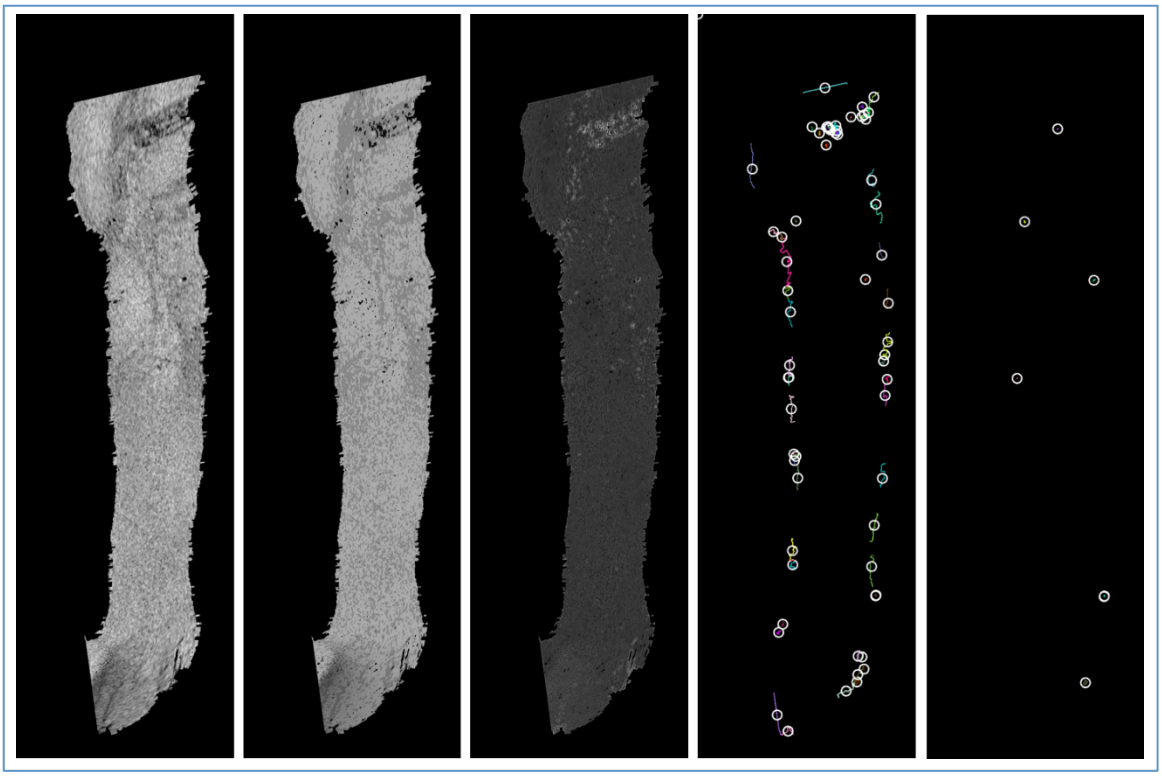

Figure 10 – Stages in the Bayesian analysis of backscatter anomalies. The sub-images show, left to right: the geometrically and radiometrically corrected backscatter mosaic; the clustered mosaic (clustering is based on simple backscatter values in the mosaic); the probability map for membership of each analysis window in its surrounding background (high values indicate lack of membership); edge detected segments indicating potential objects; and hierarchically filtered objects showing those likely not associated with edge effects in the mosaic. The limited number of detections is promising (from the point of view of limiting false alarms), and corresponds well to operator inspection of the mosaic.

The backscatter information content, as variable angular response return, is also used in a model-based detection algorithm (Figure 11). The angular response for each acoustically clustered area (i.e., detecting anomalies from the average angular response, which is quite different from the mosaic response where angular differences are removed), and evaluation of the half-swath patch (i.e., stacking a certain number of successive pings to stabilize the statistics and reduce the noise), again looking for anomalies in the angular response of the immediate area [12].

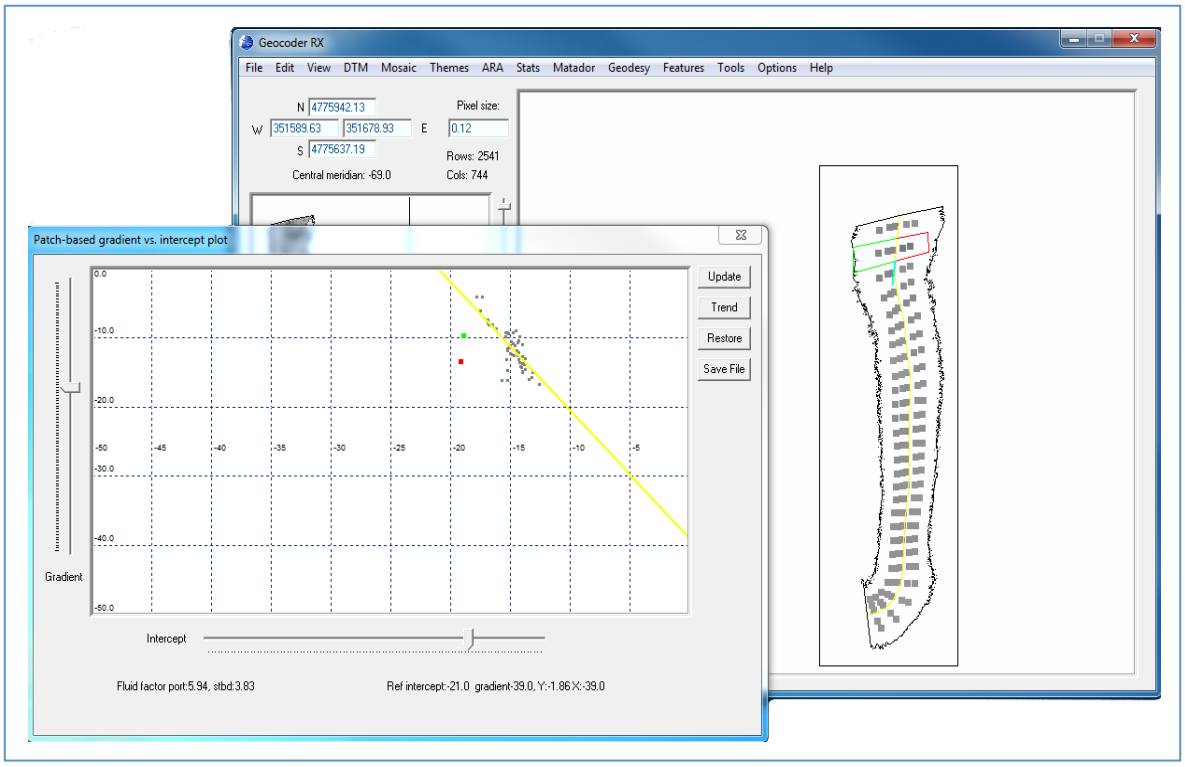

Figure 11 – Bivariate plot of acoustic backscatter-derived features computed from a half-swath patch of MBES data. The red/green indicated half-swaths have distinctly different behavior (as measured by the slope and intercept of a line fitted to the acoustic backscatter angular response in the patches) from the other patches, an indication of anomalous behavior. Use of multivariate combinations of features can help to clarify detections and reduce false alarm rates.

Finally, we have also created algorithms that adapt well-known general-use image processing techniques to marine debris detection.

4.2.3. Bathymetric-based Algorithms

For the bathymetric DTM, a few spatial indices were used as proxies for any discontinuity and, thus, a possible target.

In particular, an algorithm is based on the Combined Uncertainty and Bathymetry Estimator (CUBE), a weighted surface-construction algorithm for automatically that processes big and dense bathymetric data set, and that addresses issues as efficiency, objectivity, robustness and accuracy [66]. Based on available soundings, the nodes independently assimilate propagated information to form depth hypotheses which are then tracked and updated on-line as more data is gathered. CUBE manages groups of soundings that are mutually inconsistent, but internally consistent, by segregating them in sacrificial alternate hypothesis, avoiding cross-contamination of estimates [66]. The state of knowledge about the data is summarized for each estimation node by a list of depth hypotheses (Multiple Hypothesis Tracking). This work explored the use of CUBE's auxiliary products for marine debris detection (Figure 12). The map of hypothesis counts clearly shows areas of difficulty in the gridding process, and, together with CUBE's estimate of the correctness of the

hypothesis selected by the disambiguation engine (hypothesis strength) [67], may be used as a proxy for debris presence. Low values of hypothesis strength are used to assess a node depth reconstruction that is sufficiently robust to be reliable [68].

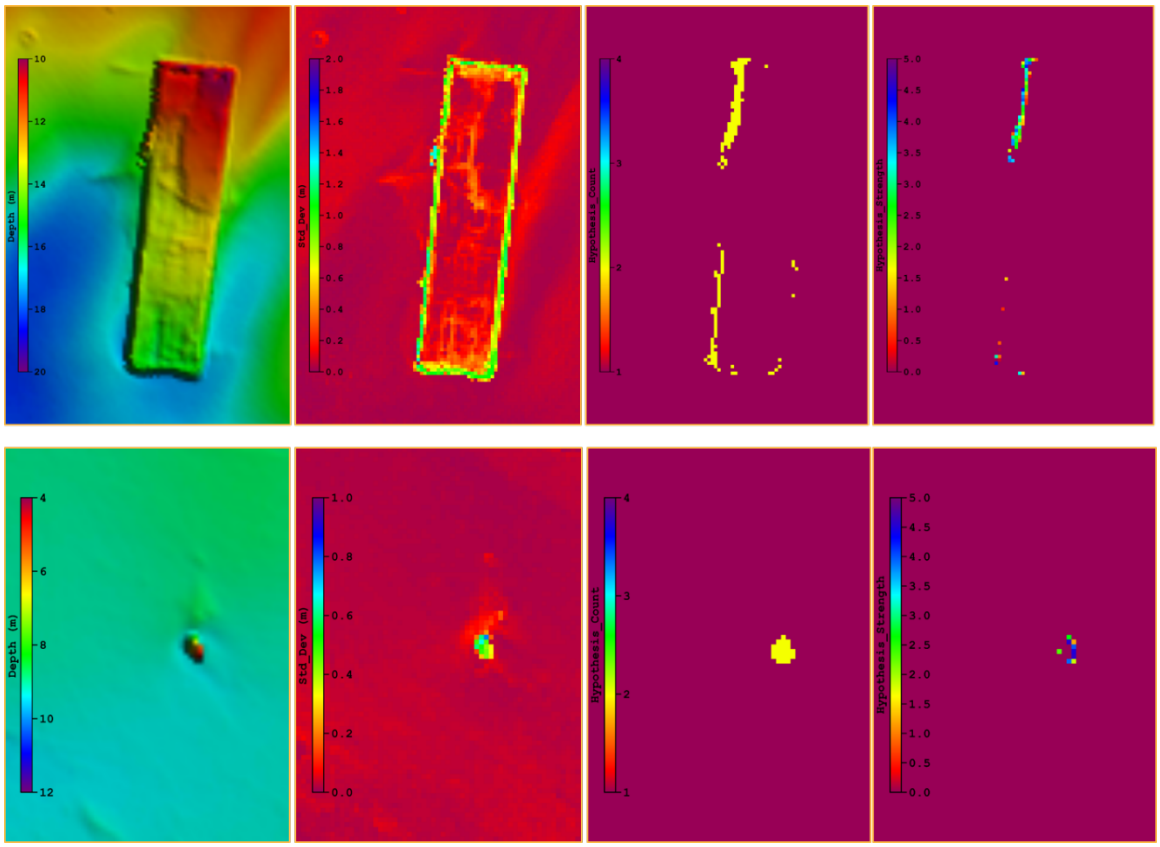

Figure 12 – Examples of additional useful metrics related to the statistical bathymetric representation of two objects (from left to right, bathymetry, standard deviation, hypothesis count, and hypothesis strength).

4.2.4. Additional Remarks

For both backscatter and bathymetry, the adoption of classical estimation techniques usually generates point estimate or a confidence interval, which becomes important when fusion of target information coming from different products is attempted. In order to provide appropriate distributions for exploitation, Bayesian methods were adopted since they permit use of multiple-source asymmetric and discontinuous posterior distributions that may be carried into further analysis. A hierarchical scheme is proposed where a series of modeling tasks are implemented through a probabilistic model, casting the debris detection problem as one of estimating properties of the posterior distribution describing the probability of objects occurring given the observed data products.

4.3. Fusion Approach

The described predictive model provides the areas of higher likelihood of there being an object (debris prediction). The information obtained from the various detectors in use

represents an estimation of how likely will that there will be a detection given that there is an object there in a given node. Unfortunately, nothing of them is what we want, since what we want is how likely that there is an object given the detections that we saw. To solve this issue, we picked one of the available statistical techniques, structuring it as a Bayesian problem. What we are doing is combining all these information together in order to predict what we want. The Bayesian hierarchical model provides us with a probabilistic structure that we use to combine together the information that we are gathering from the detectors (building a model on how likely is that the model is working correctly given such parameters) and the extra prior information that we have from the predictive model (or from ENC or from other sources). Wrapped up it into a Bayesian estimation problem we adopted computation tools to calculate the probability of there being an object in any particular area.

Apart from the strong mathematical framework and the very well understood mathematical structure that comes along with it, one of the big advantages of this technique is that allows us to build an extra knowledge tool. This comes back to the idea that this is a very complex problem and that we are attempting to put as much as possible structure here to constrain the solution as more as we can to get reliable solutions. One of the things that we can try to take advantage of is the spatial autocorrelation. When you have a system like the one we developed, you tend to have multiple detections all flashing at the same spatial location consistently from the various detectors. Assuming that this is an indication that there is a valid detection, we can build it in as an additional variable, and the mathematical methodology (i.e., auto-logistic model) provides a way to add such an information. The result is that we take advantage of that spatial context to obtain a more robust estimate of what is going on in our data.

4.3.1. Adopted Hierarchical Spatial Modeling

The adopted approach embraces hierarchical spatial modeling in the conviction that it is the sharp statistical tool needed to ascend the knowledge pyramid in marine debris detection and recognition. Its real power becomes apparent with complicated dependencies, when each component distribution can be both further decomposed and simplified by introducing modeling assumptions. The proposed fusion approach provides a means to statistically merge together the output of a set of detection algorithms, each of them consuming hydrographic data products (e.g., bathymetric surfaces, backscatter mosaics, seafloor characterization analyses). Since the detector has been developed with the task of detecting the presence of marine debris over a surveyed area, this task can be basically reduced to a classification problem based on only two classes. After having discretized the surveyed area using a regular lattice, the classifier maps each cell of the lattice to one element of the set $\{t, n\}$, respectively representing target and not-target labels. Similar to other existing classification models, the classifier produces an intermediate step that estimates membership probability to the 'target' class for each cell. The application of a different

threshold to such an estimate can be used to tune the classifier behavior, directly influencing its balance between hit rates and false alarm rates. For notational clarity, the label set $\{Y, N\}$ is specifically introduced for the detector output, to distinguish predictions from the actual class. The detector may reach four possible states, as summarized using a two-by-two matrix, called contingency table, in Figure 13.

		Actual Labels: t	Actual Labels: n
Detector Labels	Y	***TP*** *True Positives*	***FP*** *False Positives*
	N	***FN*** *False Negatives*	***TN*** *True Negatives*
		P *Total Positives*	***N*** *Total Negatives*

Figure 13 – Contingency table and notations adopted for the marine debris detector.

The classifier states can be combined in various classification metrics, such as:

- The true positive rate (TPR), or sensitivity, evaluated as

$$TPR = TP / P \qquad (1)$$

  where P represents the 'real' total positives.
- The specificity, directly related to the false positive (or false alarm) rate (FPR),

$$spec = TN / (FP + TN) = 1 - FPR. \qquad (2)$$

The performance of the proposed detector can be depicted using two-dimensional Receiver Operating Characteristics (ROC) graphs [69]. On these graphs, a diagonal line ($y = x$) is added to represent the strategy of randomly guessing a class. A single scalar representing the portion of the area of the unit square, called the Area Under the Curve (AUC), is commonly used to compare classifiers. Since random guessing produces an area of 0.5, no realistic classifier should provide a value less than 0.5.

The debris data are necessarily categorical: either binary (presence/absence) or abundance (number of objects at a given location). The developed workflow focuses on binary data, seen as the result of the fusion approach, that may be easily fitted using logistic regression, which is straightforward to implement using standard maximum likelihood [70]. However, like all linear models, logistic regression is incompatible with SAC observations since it assumes independence of errors. Thus, among several possible approaches for modeling autocorrelation in binary data, the auto-logistic approach [23, 71], that extends the logistic

model to allow dependence between nearby observations [23], was selected. Common past objections against the use of auto-models seem to be due to incorrect model implementations, providing a general conclusion of validity for auto-model analysis of SAC data [72].

The full likelihood for the auto-logistic model is known analytically only within a normalization constant [73], an analytically intractable function of the regression parameters. This issue has driven the development of an approximation scheme known as the maximum pseudo-likelihood estimate method (MPLE). Although auto-covariate term are included in the regression, MPLE assumes that the sites are independent for the purpose of constructing the likelihood function (that simplifies to a simple product form). Several researchers have observed inaccuracy (e.g.,. overestimation of intrinsic auto-correlation in data with strong intrinsic SAC) resulting from MPLE application [74]. Thus, a simple Bayesian implementation of the auto-logistic model was adopted, avoiding MPLE but adding a computation burden. This solution can be simultaneously used for model fitting, based on covariates, and for making predictions about unsurveyed parts of the study area. In several recent examples, the auto-logistic model has been implemented in a Bayesian framework adopting the MCMC methodology first described by Geman and Geman [75]. This methodology constructs a non-normalized posterior distribution for the unknown parameters and any missing observations (which are treated as unknown parameters). The resulting distribution, conditional on the observations, is sampled according to one of the possible MCMC procedures (e.g., the Gibbs sampler) that guarantee that certain sequences of dependent samples (Markov chain) will converge to the target distribution. This Bayesian approach allows for a more flexible incorporation of possible complications (observer bias, missing data, and different error distributions) at the expense of more computational-intensive requirements.

4.3.2. Case study

In order to evaluate the described approach, Jamaica Bay, an area in the East Coast of US affected by Super Storm Sandy, was selected. On this area we have imagery, bathymetry, and sidescan, as well as we have ground truth object detections done by hand (Figure 14). So that we have a way to compare the results against the detections that the algorithm comes up with. And of course we have also the ENC information that allows us to give extra constraint to the context, and we are using the predictive probability from our prediction model too.

Figure 14 – From top to bottom, bathymetry, backscatter mosaic, and analyst-based detection (with land/sea mask from ENC).

Figure 15 focuses down on a particular area, in the southwest region of Jamaica Bay, showing the marine debris prediction as hot spot map. The map shows strong debris prediction on particular zones. Part of this prediction is related on there being an anthropogenic structures in the area that can feed into the water there (or extra effects for storm surge, as example).

Since the system is naturally scalable, it is possible to zoom in a particular sub-area (Figure 16), so that is possible to increase the map resolution without being too computationally expensive. If we zoom in, we can look in more details. The figure shows a good correlation between the high peaks in the hot spot probability map and where are the analyst-selected objects.

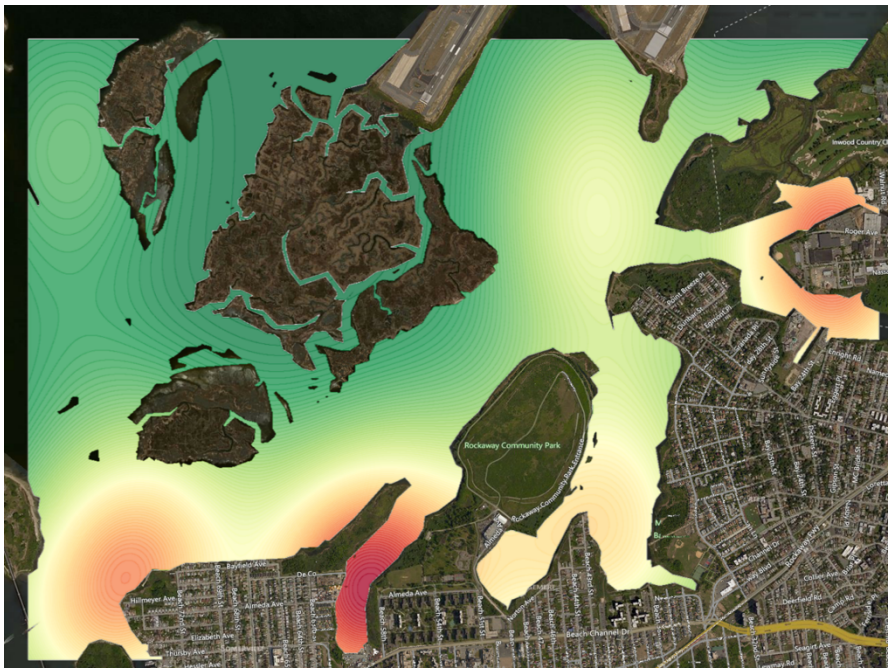

Figure 15 – Hot spot map generated by the workflow for a particular area of Jamaica Bay (NYC).

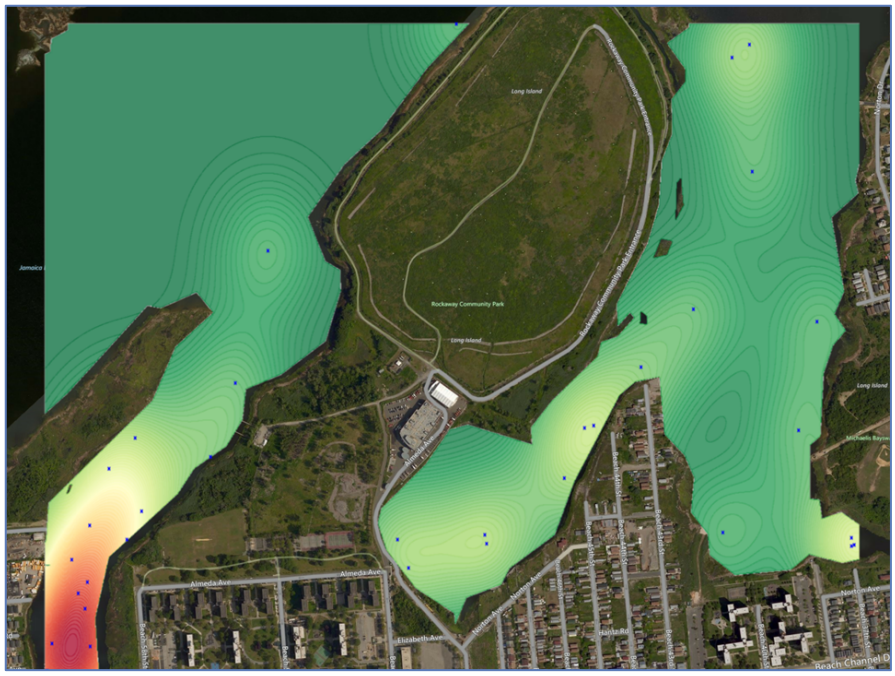

Figure 16 - Resulting of the hot spot analysis in Jamaica Bay (NYC), with the ground-truth positions of the marine debris designed by human analysts showed as blue dots.

The detection outcomes based on auto-logistic and logistic regressions have been compared. The comparison has been performed separately on two data sets: a first data set was artificially created, injecting SAC as described in [72], and a second set of outcomes was obtained from the application of the detection algorithms to real hydrographic products (the cited Jamaica Bay data set collected after the Super Storm Sandy event). The resulting ROC graphs for both data sets are presented in Figure 17. The graphs for the artificial data set present AUC values of 0.923 (logistic regression) and 0.940 (auto-logistic regression); while, for the real survey data set, AUCs of 0.862 and 0.880 characterize the logistic and the auto-logistic regressions, respectively.

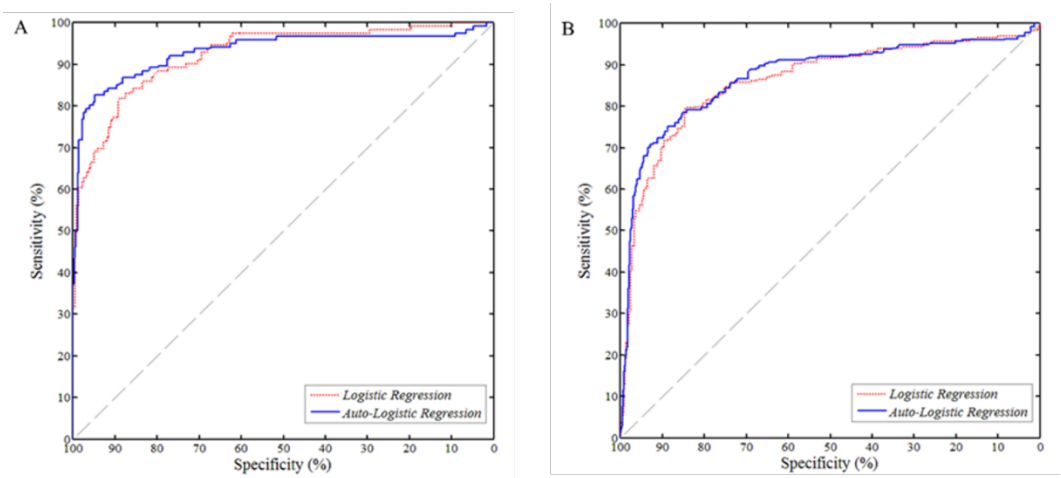

Figure 17 – ROC curves for artificial data (pane A) and real survey data (pane B).

The auto-logistic implementation outperformed the basic logistic model in removing residual auto-correlation. Thus, the auto-logistic model provides a better description of the observed clustering of objects, since the logistic model cannot represent clustering at all unless it is present in the covariates. Furthermore, ROC plots (Figure 17) indicates better overall predictive performance by the auto-logistic model due to much higher true positive rates at small false positive rates, although the logistic model slightly outperforms at low (artificial data set) and intermediate (real data set) specificity. Classifiers appearing on the left-hand side of an ROC graph (near the *Y* axis), are usually evaluated as more "conservative", since they make positive classifications only with strong evidence (so they make few false positive errors) [69]. Since the marine debris detection domain is usually dominated by large numbers of negative instances, performance like the auto-logistic regression becomes more interesting. Similar tests with data sets collected over different areas and using various sensors will be applied to verify the robustness of the approach.

5. Target Management and Data Exchange

5.1. *Marine Object Manager*

The combination of information present in bathymetric and imagery-based products is a key requirement for any modern feature-detection approach that aims to be adopted in coastal areas whereas the seafloor is deep enough that optic means are not reliable. If the data sources and the processing involved are correctly weighted in a fusion algorithm, the detection task can be extended beyond a simple binary (presence/absence) decision to provide a meaningful metric that evaluates confidence in the presence of new features. In combination with other existing information (such as that present in ENCs), this metric can become a proxy for areas with high probability of change (for features to be either added or removed) with respect to the baseline knowledge of the area. The dual, and partially contradictory, goals of such a system are to highlight areas with high probability of change, and to use the existing nautical documentation as a spatial filter to resource consumption on known features. Determining an appropriate balance between these is an interesting challenge. However, we recognize that the algorithm we are never going to be perfect (very few algorithms actually are).

Based on such considerations, this work describes an approach for how to effectively assist data analysts in combining the results of different target detection algorithms, as well as in comparing such results with existing features present on ENCs and geographic databases (e.g., spatial DBMS). The main goal is to help the analyst in focusing on specific areas (with higher likelihood of new features), prioritizing them on safety-of-navigation criteria and reducing the common pitfall of subjectivity in the processing workflow. Although mainly aimed at rapid response to the short-term increase in marine debris deposition related to major events like hurricanes and floods, the approach is also well suited for different scenarios such as reducing the "ping-to-chart" time.

So we are building together with this workflow some prototype tools for a human investigation part that we expected will be always the case. We can identify object very easily, but identify whether they are natural or marine debris is not necessarily straightforward for an algorithm (Figure 18), without a level of sophistication of the algorithm that makes it very slow. Detecting natural features like this and artificial like that sometimes is not so hard, sometimes is very difficult. We are leaving for now to a human operator. What we are doing however is to provide tools to accelerate such a mechanism. So we want to provide a visualization more efficient as we can. The mechanism is driven by a higher probability of being an object. We also evaluated tools to extract shape both automatically and by hand (Figure 19).

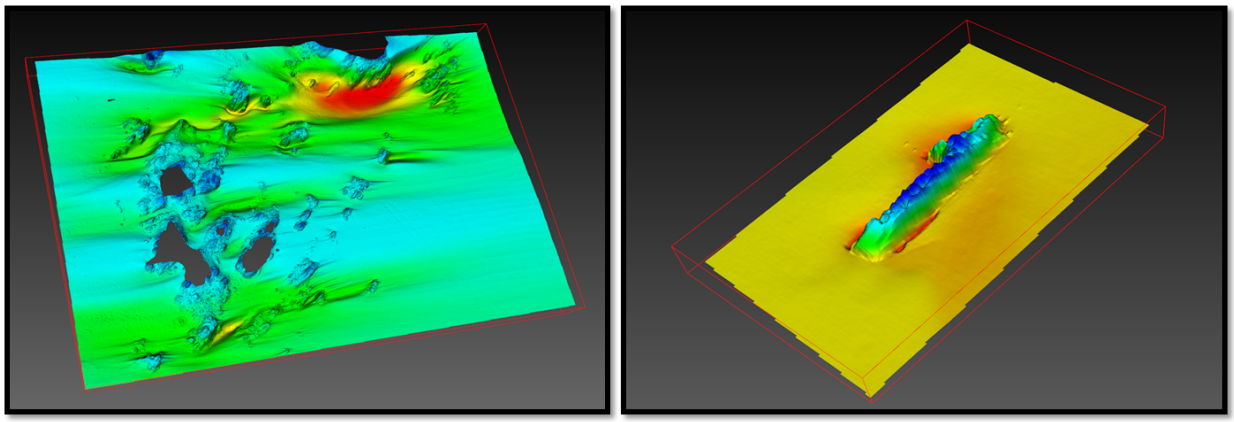

Figure 18 – Examples of 'natural' (left) and 'marine debris' (right) candidates provides as output of the detection model.

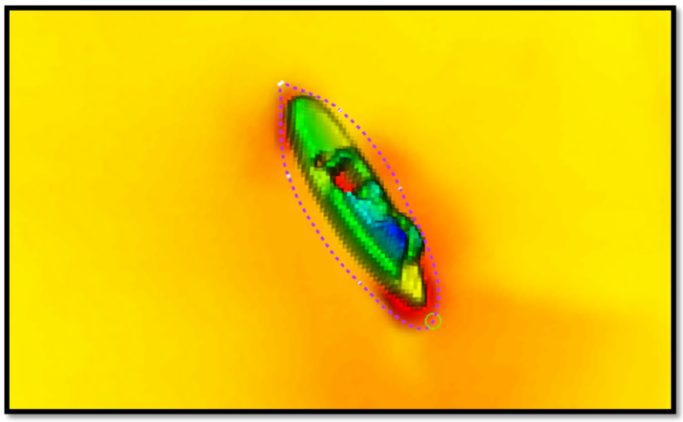

Figure 19 – Example of auxiliary tool provided to the analyst to define the shape of the target.

5.2. Marine Debris Markup Language

A driven consideration for the data exchange part was that having an algorithm that produces nice results but these do not go anywhere, the overall approach is much less useful. As a direct consequence, the workflow pays some attention to data interoperability, a key factor that is often neglected. After a marine disaster many agencies conduct surveys collecting information that may be related in different ways to the presence of marine debris. The data acquired in these surveys can be used for planning purposes, operational support, and to manage risks associated with marine debris. Most surveys adhere to standards and best practices in creating outputs, but there is currently a lack of standardization for debris detections, which are typically delivered in unstructured CAD and GIS files. Although flexible, this lack of structure makes integration and interoperability with these files difficult. This make sharing marine debris data difficult.

A collaboration with NOAA's Marine Debris Program has therefore been undertaken to jointly evaluate how this data can be better managed; the result is a Marine Debris Markup Language (MDML).

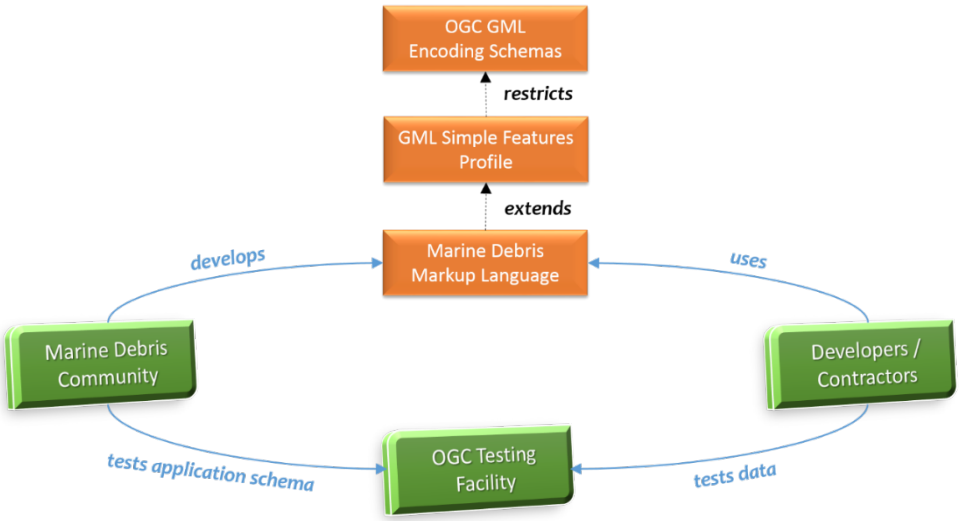

Figure 20 – Interactions between the community interested to marine debris data and users of the Marine Debris Markup Language (MDML). The MDML uses several of the primitives defined in GML Core Schemas as restricted by GML Simple Feature Profile (Level Zero).

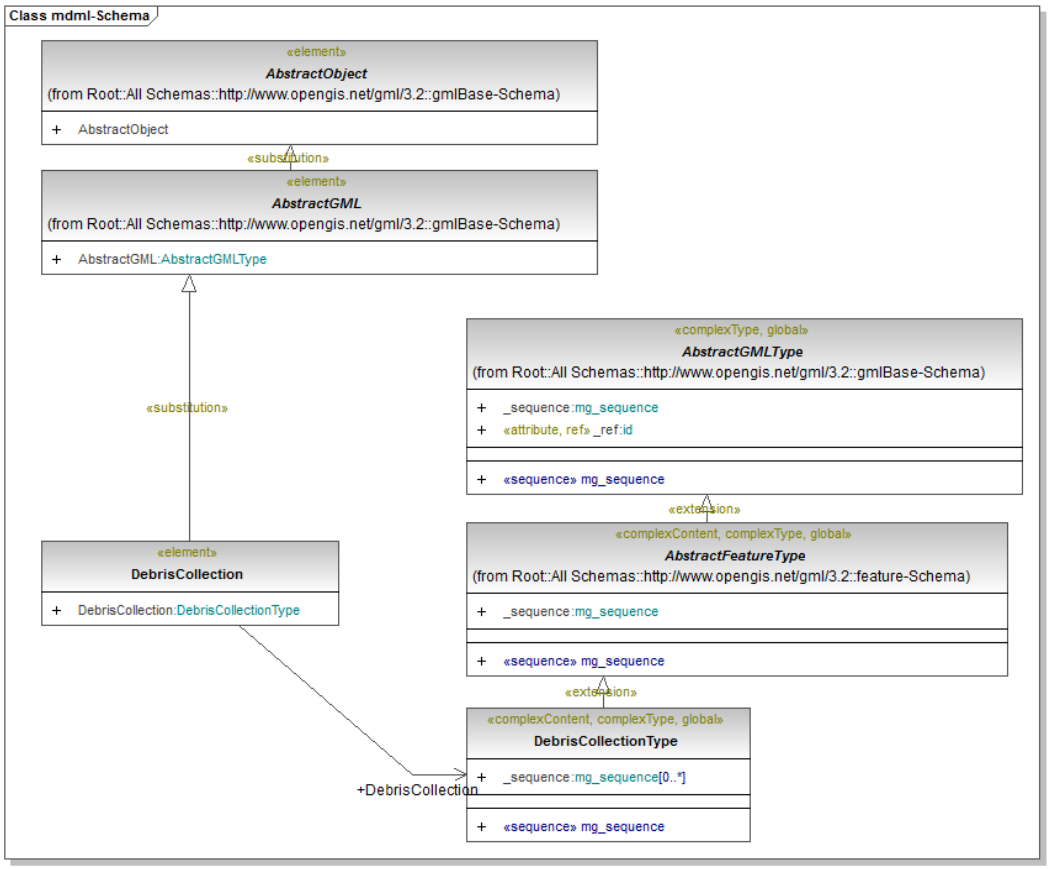

Figure 21 – Details of the MDML implementation: each class, as the DebrisCollection class in the diagram, is derived from GML abstract elements that have proper counterparts within other spatial object models (e.g., ESRI Data Model).

This model provides a common vocabulary to describe marine debris based on sound geo-information management principles and practice, which could be used to provide guidelines for the acquisition of marine debris data in future events. The MDML template is implemented using a GML-aware XML schema that is the *lingua franca* for geographic data exchange (Figure 20). However, in principle the data model can be implemented on any GIS and transferred via an open GIS data exchange format (Figure 21). Because the schema uses the Open Geospatial Consortium GML Encoding Standard and the GML Simple Feature Profile (Level Zero), many applications are already able to process MDML data files, thus supporting intra- and inter-community interoperability.

Furthermore, by building onto a standard feature profile, this gives us something that is easily reconstructed in standard OGC tools, GDAL and so on. The advantage is that the output by following the Marine Debris Markup Language can be readily imported in other tools, without any extra effort.

6. Final Remarks

The problem of marine debris identification is complex, particularly because what we are attempting to find is not particularly well defined (There is not real good definition of what a marine debris really is). So this workflow assumes to have to deal with this variable situation. For this the workflow defines a selection of algorithms that, although singularly not quite reliable, together permits to fuse a solution that is reliable and that we can use. Given such a consideration, the workflow is easy to be improved by introducing better predictive model and detection algorithms, since the more structure we can put to the problem, the better it is.

However, as described in the previous chapters, the current workflow provides a way to structure a complex problem by merging the output of different algorithm in a hot spot map that is easy to understand. A partially open question is whether having new detection algorithms actually improves the detection. Our current impression based on some testing is that is not always the case. We explain these results from the fact that the more metrics you generate from the algorithms and then use in the fusion you could cause more confusion in the detection that you gain from having the extra algorithm there. So more algorithms are not necessarily better. What you want if a limited number of independent algorithms that gives you independent views to the data.

The adopted Bayesian approach allows for a more flexible incorporation of possible complications (observer bias, missing data, and different error distributions) and prior beliefs at the expenses of higher computational-intensive requirements. Furthermore, a good understanding of the influence of prior distributions and convergence assessment of Markov chains is crucial to properly evaluate the methodology results. The scalability of the fusion technique permits its straightforward extension to additional detection algorithms for *ad-hoc* created data products, with expected improvements both in robustness against outliers and in detection performance. Given such a framework flexibility, it is possible to extend the detector with a set of *ad hoc* hydrographic products. Although the probability of false alarm based on a combined analysis of multiple data sources is expected to be generally lower than when a single source is used, there are particular cases where a particular object might only be observable within a single data source. For example, a semi-buried target, or one with a flattened shape, might only be visible through acoustic backscatter. A careful analysis of the benefits of different algorithms and different data sources is therefore indicated in case of future research along the direction showed by this approach.

From a more general point of view, we believe that outcomes of this project can be used to improve current hydrographic data processing for target detection. Since the current processing is usually subjective and time consuming, increasing the automation of the process would be largely beneficial toward subjectivity reduction.

References

[1] D. Li, M. R. Azimi-Sadjadi, and G. J. Dobeck, "Comparison of different neural network classification paradigms for underwater target discrimination," 2000, pp. 334-345.

[2] J. Stack, "Automation for underwater mine recognition: current trends and future strategy," in *SPIE Defense, Security, and Sensing*, 2011, pp. 80170K-80170K-21.

[3] B. A. Telfer, H. H. Szu, G. J. Dobeck, J. P. Garcia, H. Ko, A. C. Dubey, *et al.*, "Adaptive wavelet classification of acoustic backscatter and imagery," *Optical Engineering,* vol. 33, pp. 2192-2203, 1994.

[4] G. Masetti and B. Calder, "A Marine Object Manager as an Information Fusion Tool for Detected and Database-stored Shipwrecks," in *Wrecks of the World III: Shipwreck Risk Assessment*, Gothenburg, Sweden, 2015.

[5] G. Masetti, B. R. Calder, and M. Wilson, "A Marine Object Manager for Detected and Database-stored Features," in *Shallow Survey 2015*, Plymouth, UK, 2015.

[6] C. A. Ribic, T. R. Dixon, and I. Vining, "Marine debris survey manual," *NOAA/NMFS Technical Report 108,* 1992.

[7] S. Katsanevakis, "Marine debris, a growing problem: Sources, distribution, composition, and impacts," *Marine Pollution: New Research. Nova Science Publishers, New York,* pp. 53-100, 2008.

[8] G. Masetti and B. Calder, "A risk index methodology for potentially polluting marine sites," *Environment Systems and Decisions,* vol. 34, pp. 391-405, 2014/09/01 2014.

[9] G. Masetti and B. Calder, "Design of a standardized geo-database for risk monitoring of potentially polluting marine sites," *Environment Systems and Decisions,* vol. 34, pp. 138-149, 2014/03/01 2014.

[10] N. Barnea, J. Michel, B. Bray, Z. Nixon, G. Imahori, and C. Moegling, "Marine Debris Response Planning in the North-Central Gulf of Mexico," *NOAA Technical Memorandum NOS-OR&R-31,* 2009.

[11] Z. Nixon and N. Barnea, "Development of the Gulf of Mexico Marine Debris Model," *NOAA Technical Memorandum NOS-OR&R-35,* 2010.

[12] G. Masetti and B. Calder, "Development of an adaptive fusion algorithm for marine debris recognition within the post-Sandy restoration framework," in *Canadian Hydrographic Conference 2014*, St. John's, NL (Canada), 2014, p. 14.

[13] N. A. C. Cressie and C. K. Wikle, *Statistics for spatio-temporal data*. Hoboken, N.J.: Wiley, 2011.

[14] S. Banerjee, B. P. Carlin, and A. E. Gelfand, *Hierarchical modeling and analysis for spatial data*: CRC Press, 2015.

[15] K. Sullivan and L. Uccellini, "Service Assessment: Hurricane/Post-Tropical Cyclone Sandy, October 22–29, 2012," *NOAA/NWS,* 2012.

[16] J. Lubchenco and L. Furgione, "Service Assessment: Hurricane Irene, August 21–30, 2011," *NOAA/NWS,* 2012.

[17] NOAA. (2015, 21/2/2015). ERMA Atlantic. Available: https://erma.noaa.gov/atlantic/erma.html

[18] S. Lang, "Object-based image analysis for remote sensing applications: modeling reality," in *Object-Based Image Analysis*, T. Blaschke, S. Lang, and G. Hay, Eds., ed: Springer Berlin Heidelberg, 2008, pp. 3-27.

[19] NOAA. (2015, 21/2/2015). Gulf of Mexico Marine Debris Project. Available: http://www.ncddc.noaa.gov/website/GOM_MarineDebris/viewer.htm

[20] D. L. Johnson, "Service Assessment: Hurricane Katrina, August 23–31, 2005," *NOAA/NWS,* 2006.

[21] G. Masetti and B. Calder, "A Bayesian marine debris detector using existing hydrographic data products," in *IEEE Oceans Conference*, 18-21 May 2015, 2015.

[22] W. R. Tobler, "A Computer Movie Simulating Urban Growth in the Detroit Region," *Economic Geography,* vol. 46, pp. 234-240, 1970.

[23] J. Besag, "Spatial interaction and the statistical analysis of lattice systems," *Journal of the Royal Statistical Society. Series B (Methodological),* pp. 192-236, 1974.

[24] L. Anselin, "Under the hood Issues in the specification and interpretation of spatial regression models," *Agricultural Economics,* vol. 27, pp. 247-267, 2002.

[25] J. Besag, J. York, and A. Mollié, "Bayesian image restoration, with two applications in spatial statistics," *Annals of the Institute of Statistical Mathematics,* vol. 43, pp. 1-20, 1991/03/01 1991.

[26] T. H. Keitt, O. N. Bjørnstad, P. M. Dixon, and S. Citron-Pousty, "Accounting for spatial pattern when modeling organism-environment interactions," *Ecography,* vol. 25, pp. 616-625, 2002.
[27] L. Palma, P. Beja, and M. Rodrigues, "The use of sighting data to analyse Iberian lynx habitat and distribution," *Journal of Applied Ecology,* vol. 36, pp. 812-824, 1999.
[28] P. Grabarnik and S. N. Chiu, "Goodness-of-fit test for complete spatial randomness against mixtures of regular and clustered spatial point processes," *Biometrika,* vol. 89, pp. 411-421, June 1, 2002 2002.
[29] J. N. Perry, A. M. Liebhold, M. S. Rosenberg, J. Dungan, M. Miriti, A. Jakomulska*, et al.*, "Illustrations and guidelines for selecting statistical methods for quantifying spatial pattern in ecological data," *Ecography,* vol. 25, pp. 578-600, 2002.
[30] F. Wang, *Quantitative methods and applications in GIS*: CRC Press, 2006.
[31] A. Getis and J. K. Ord, "The Analysis of Spatial Association by Use of Distance Statistics," *Geographical Analysis,* vol. 24, pp. 189-206, 1992.
[32] L. Anselin, "Local indicators of spatial association—LISA," *Geographical Analysis,* vol. 27, pp. 93-115, 1995.
[33] J. K. Ord and A. Getis, "Local spatial autocorrelation statistics: distributional issues and an application," *Geographical Analysis,* vol. 27, pp. 286-306, 1995.
[34] A. K. Jain, "Data clustering: 50 years beyond K-means," *Pattern Recognition Letters,* vol. 31, pp. 651-666, 2010.
[35] T. Caliński and J. Harabasz, "A dendrite method for cluster analysis," *Communications in Statistics,* vol. 3, pp. 1-27, 1974/01/01 1974.
[36] C. F. Dormann, J. M. McPherson, M. B. Araújo, R. Bivand, J. Bolliger, G. Carl*, et al.*, "Methods to account for spatial autocorrelation in the analysis of species distributional data: a review," *Ecography,* vol. 30, pp. 609-628, 2007.
[37] N. Augustin, M. Mugglestone, and S. Buckland, "An autologistic model for the spatial distribution of wildlife," *Journal of Applied Ecology,* pp. 339-347, 1996.
[38] D. Umpierre and G. Margoles, "Broward county's web-based Hurricane debris estimation tool (HurDET)," in *2005 ESRI International User Conference Proceedings*, 2005.
[39] FEMA, "Appendix B - USACE Hurricane Debris Estimating Model," in *Public Assistance - Debris Management Guide*, FEMA-325 / July 2007 ed, 2007, p. 260.
[40] FEMA, *Hazus - MH 2.1 - Hurrican Model - User Manual*. Washington, D.C.: Department of Homeland Security, 2013.
[41] M. D. Powell, S. Murillo, P. Dodge, E. Uhlhorn, J. Gamache, V. Cardone*, et al.*, "Reconstruction of Hurricane Katrina's wind fields for storm surge and wave hindcasting," *Ocean Engineering,* vol. 37, pp. 26-36, 2010.
[42] NOAA. (2015, 03/20/2015). Extratropical Surge and Tide Operational Forecast System (ESTOFS). Available: http://www.stormsurge.noaa.gov/models_obs_modeling.html
[43] B. W. Eakins and P. R. Grothe, "Challenges in Building Coastal Digital Elevation Models," *Journal of Coastal Research,* pp. 942-953, 2014.
[44] NOAA. (2015, 22/2/2015). International Best Track Archive for Climate Stewardship (IBTrACS). *National Climatic Data Center*. Available: http://www.ncdc.noaa.gov/ibtracs/index.php?name=ibtracs-data
[45] K. R. Knapp, M. C. Kruk, D. H. Levinson, H. J. Diamond, and C. J. Neumann, "The International Best Track Archive for Climate Stewardship (IBTrACS)," *Bulletin of the American Meteorological Society,* vol. 91, pp. 363-376, 2010/03/01 2010.
[46] P. Wessel and W. H. Smith, "A global, self-consistent, hierarchical, high-resolution shoreline database," *Journal of Geophysical Research: Solid Earth (1978–2012),* vol. 101, pp. 8741-8743, 1996.
[47] USACE. (2015, 03/25/2015). U.S. Waterway Data - National Waterway Network. *US Army Corps of Engineers - Navigation Data Center*. Available: http://www.navigationdatacenter.us/data/datanwn.htm
[48] F. R. Broome and D. B. Meixler, "The TIGER Data Base Structure," *Cartography and Geographic Information Systems,* vol. 17, pp. 39-47, 1990/01/01 1990.
[49] T. Meurling and B. Volberg, "The Evolution and Future of Multibeam Echo Sounder Technology," in *Underwater Technology and Workshop on Scientific Use of Submarine Cables and Related Technologies*, Tokyo (Japan), 2007.

[50] J. E. Hughes Clarke, "Optimal use of multibeam technology in the study of shelf morphodynamics," *Sediments, Morphology and Sedimentary Processes on Continental Shelves: Advances in technologies, research and applications (Special Publication 44 of the IAS),* vol. 108, p. 3, 2012.

[51] L. A. Mayer, R. Raymond, G. Glang, M. D. Richardson, P. Traykovski, and A. C. Trembanis, "High-Resolution Mapping of Mines and Ripples at the Martha's Vineyard Coastal Observatory," *Oceanic Engineering, IEEE Journal of,* vol. 32, pp. 133-149, 2007.

[52] X. Lurton, *An introduction to underwater acoustics : principles and applications*, 2nd ed. Heidelberg ; New York: Springer, published in association with Praxis Publishing, Chichester, UK, 2010.

[53] G. Masetti and B. Calder, "Remote identification of a shipwreck site from MBES backscatter," *Journal of Environmental Management,* vol. 111, pp. 44-52, 2012.

[54] Kongsberg, "Sector coverage and beam spacing modes for multibeam echosounders," Kongsberg MaritimeDecember 2013 2013.

[55] IHO, *IHO S-44 - Standards for Hydrographic Surveys*, 5th Edition ed. Monaco: IHB, 2008.

[56] Kongsberg, "The advantages of yaw stabilization in multibeam surveying," Kongsberg MaritimeSeptember 2013 2013.

[57] F. Pohner, J. Bakke, O. Nilsen, T. Kjaer, and F. L., "Integrating Imagery from Hull Mounted Sidescan Sonars with Multibeam Bathymetry," in *U.S. Hydrographic Conference*, Norfolk, VA (USA), 2007, p. 16.

[58] G. Wheaton, "PATCH TEST, A system check for multibeam survey systems," in *Proc. Hydrographic Conf*, 1988, pp. 12-15.

[59] J. Eisenberg, M. Davidson, J. Beaudoin, and S. Brodet, "Rethinking the Patch Test for Phase Measuring Bathymetric Sonars."

[60] J. E. Hughes Clarke, "Dynamic Motion Residuals in Swatch Sonar Data: Ironing out the Creases," *International Hydrographic Review,* vol. 4, pp. 6-23, 2003.

[61] W. S. Burdic, *Underwater acoustic system analysis*: Prentice Hall New Jersey, 1991.

[62] J. Beaudoin, B. Calder, J. Hiebert, and G. Imahori, "Estimation of sounding uncertainty from measurements of water mass variability," *International Hydrographic Review,* pp. 20-38, November 2009 2009.

[63] M. J. Wilson, J. Beaudoin, and S. Smyth, "Water-column variability assessment for underway profilers to improve efficiency and accuracy of multibeam surveys," in *U.S. Hydrographic Conference*, New Orleans, LA (USA), 2013, p. 24.

[64] Kongsberg, "Detector modes in SIS for EM 2040, EM 2040C and EM 3002," Kongsberg MaritimeSeptember 2013 2013.

[65] L. Fonseca and B. Calder, "Geocoder: An Efficient Backscatter Map Constructor," in *U.S. Hydrographic Conference*, San Diego (USA), 2005.

[66] B. R. Calder and L. A. Mayer, "Automatic processing of high-rate, high-density multibeam echosounder data," *Geochemistry, Geophysics, Geosystems,* vol. 4, p. 1048, 2003.

[67] B. Calder and D. Wells, "CUBE user's manual," *CCOM/JHC, UNH, Durham, NH, USA,* 2007.

[68] B. Calder, "Automatic statistical processing of multibeam echosounder data," *International Hydrographic Review,* vol. 4, pp. 53-68, 2003.

[69] T. Fawcett, "An introduction to ROC analysis," *Pattern Recognition Letters,* vol. 27, pp. 861-874, 2006.

[70] D. Collett, *Modelling binary data*: CRC press, 2002.

[71] B. Wintle and D. Bardos, "Modeling species-habitat relationships with spatially autocorrelated observation data," *Ecological Applications,* vol. 16, pp. 1945-1958, 2006.

[72] D. C. Bardos, G. Guillera-Arroita, and B. A. Wintle, "Valid auto-models for spatially autocorrelated occupancy and abundance data," *arXiv preprint arXiv:1501.06529,* 2015.

[73] F. W. Huffer and H. Wu, "Markov Chain Monte Carlo for Autologistic Regression Models with Application to the Distribution of Plant Species," *Biometrics,* vol. 54, pp. 509-524, 1998.

[74] C. J. Geyer and E. A. Thompson, "Constrained Monte Carlo Maximum Likelihood for Dependent Data," *Journal of the Royal Statistical Society. Series B (Methodological),* vol. 54, pp. 657-699, 1992.

[75] S. Geman and D. Geman, "Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images," *Pattern Analysis and Machine Intelligence, IEEE Transactions on,* pp. 721-741, 1984.