

Article

Design of Real—Time Sampling Strategies for Submerged Oil Based on Probabilistic Model Predictions

Chao Ji ^{1,*} , James D. Englehardt ¹ and Cynthia Juyné Beegle-Krause ² ¹ College of Engineering, University of Miami, Coral Gables, FL 33146, USA; jenglehardt@miami.edu² Environment and New Resources Department, SINTEF Ocean, NO 7465 Trondheim, Norway; CJ.Beege-Krause@sintef.no

* Correspondence: cxj357@miami.edu

Received: 5 October 2020; Accepted: 26 November 2020; Published: 3 December 2020



Abstract: Locating and tracking submerged oil in the mid depths of the ocean is challenging during an oil spill response, due to the deep, wide-spread and long-lasting distributions of submerged oil. Due to the limited area that a ship or AUV can visit, efficient sampling methods are needed to reveal the real distributions of submerged oil. In this paper, several sampling plans are developed for collecting submerged oil samples using different sampling methods combined with forecasts by a submerged oil model, SOSim (Subsurface Oil Simulator). SOSim is a Bayesian probabilistic model that uses real time field oil concentration data as input to locate and forecast the movement of submerged oil. Sampling plans comprise two phases: the first phase for initial field data collection prior to SOSim assessments, and the second phase based on the SOSim assessments. Several environmental sampling techniques including the systematic random, modified station plans as well zig-zag patterns are evaluated for the first phase. The data using the first phase sampling plan are then input to SOSim to produce submerged oil distributions in time. The second phase sampling methods (systematic random combined with the kriging-based sampling method and naive zig-zag sampling method) are applied to design the sampling plans within the submerged oil area predicted by SOSim. The sampled data obtained using the second phase sampling methods are input to SOSim to update the model's assessments. The performance of the sampling methods is evaluated by comparing SOSim predictions using the sampled data from the proposed sampling methods with simulated submerged oil distributions during the Deepwater Horizon spill by the OSCAR (oil spill contingency and response) oil spill model. The proposed sampling methods, coupled with the use of the SOSim model, are shown to provide an efficient approach to guide oil spill response efforts.

Keywords: submerged oil; environmental sampling; geospatial sampling; probabilistic model

1. Introduction

During the Deepwater Horizon (DWH) oil spill, ~2 million barrels of oil submerged [1] were found in the water column at depths of ~1000–1300 m. An important task was to understand where the oil was in the water column, which required defining the subsurface well blowout peeling layers [2]. The submerged oil droplets released from the plume were neutrally buoyant in the isopycnal layers with a density range of 1027.62–1027.70 kg/m³ [3]. These layers are represented by the area within the red dashed lines in Figure 1 which mark specific density contours (isopycnals). The areas with submerged dissolved oil are marked in light brown, and the oil droplets are plotted as dark brown dots. The depths of submerged oil are below the main thermocline which is ~1000 m. The oil transport is governed by geostrophic balance due to the Earth's rotation, which makes the isopycnal increasingly

tilted in locations where the horizontal velocity increases. The submerged oil posed significant environmental and economic risks, potentially causing hypoxic or anoxic zones [4] from oil droplet biodegradation, impacting offshore ecosystems such as coral reefs [5] or other sensitive ecosystems. In order to determine and monitor the environmental effects of submerged oil, field observations are needed to detect the oil, characterize different oil components and determine their concentration distributions [6].

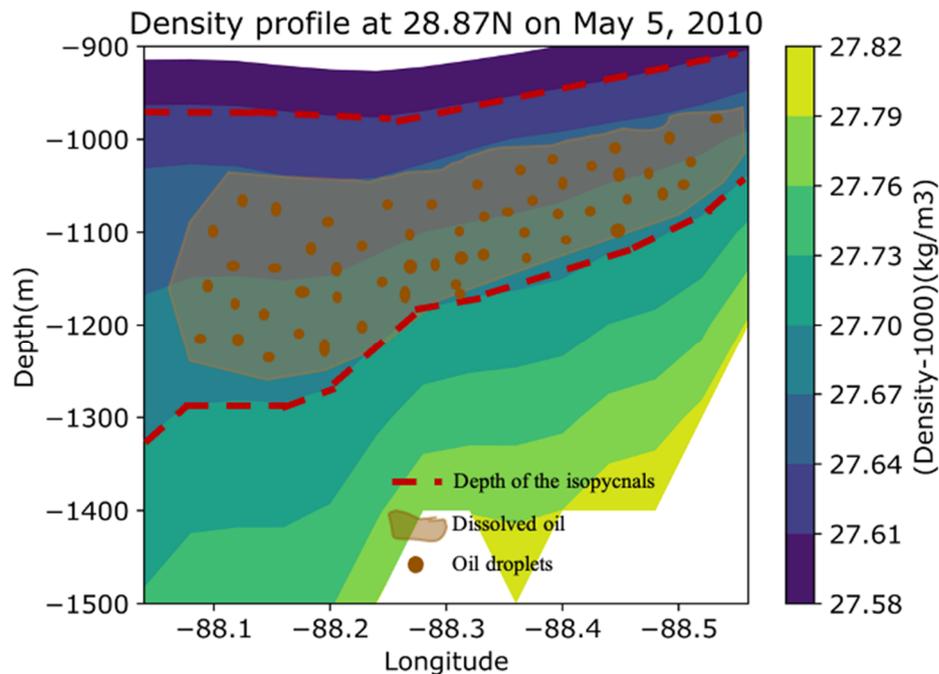


Figure 1. The cross section of water column density distribution at 28.87° N around the Deepwater Horizon (DWH) well blowout location in Gulf of Mexico on 5 May 2010. The submerged oil has been found within a ~1000–1300 m depth range, corresponding to a constant density layer of 1027.62–1027.70 kg/m³.

Over the past decades, oil spill field samples in the water column have been collected by ships. However, ship-based methods for detecting and sampling dynamic submerged oil features are often laborious and difficult. More recently, AUVs (autonomous underwater vehicles) have become available to autonomously track and detect oil. In both methods, because submerged oil from a well blowout is likely distributed over a large volume of the ocean over time, any vessel and sampling method is limited in coverage. Hence, an efficient sampling plan is required to identify representative portions of the submerged oil features present in the real distribution. This need was highlighted during the DWH oil spill emergency response period, when oil was detected by ship in only a few collected samples each day, with each sample taking ~2–3 h [7] to collect. Although oil spill models may help guide the sampling procedure, oil spill model predictions varied in their accuracy [8], before the required temperature and salinity profiles were available to calibrate the velocity field of ocean models [9,10].

The objective of this work is to design plans for efficient sampling of submerged oil during the emergency response, by applying environmental and geospatial sampling methods together with the output of the probabilistic model Subsurface Oil Simulator (SOSim) [11]. The sampling procedures include a first sampling stage, before SOSim predictions are available, and a second sampling stage, after the first sampled data set of submerged oil becomes available as input to SOSim [12].

A sampling plan can be designed using environmental and geospatial sampling methods, including random, systematic, and stratified sampling methods [13–16]. Using random sampling, sample points are selected randomly in a region of interest. In systematic sampling, the region is evenly discretized with a regular grid and the same number of data points are sampled within each grid cell.

Such systematic sampling is further classified as systematic random, centric systematic, or systematic unaligned sampling, based on the way samples are selected from each grid. If the sample is selected randomly in each grid, the sampling plan is called systematic random sampling. If the sample is selected in the center of each grid, the sampling plan is called centric systematic sampling. If the sample is selected from different relative locations, the sampling plan is called systematic unaligned sampling. In contrast with systematic sampling in which the sampling region is evenly divided into small grids, stratified sampling divides the population into nonoverlapping strata, usually based on expected density, and samples are collected within each stratum.

Geospatial sampling methods such as adaptive sampling and the kriging-based sampling take the spatial distribution of the population into consideration. The methods can be used after initial sampling, to collect additional samples if the number of samples is insufficient or the interpolation results are unsatisfactory [17]. For example, adaptive cluster sampling (ACS) begins with a coarse sampling grid and then proceeds to grids of higher density around these data points when a characteristic of interest is present in an initial unit, or when the initial unit has an observed value meeting some pre-specified conditions (e.g., when a critical threshold value is exceeded) [18–21]. ACS is useful for sparse but clustered data [21]. The kriging-based sampling is a method based on minimizing prediction error (kriging variance) of new samples based on the result of previous sampling. The kriging-based sampling has been used in soil contamination [22–24] and aquatic organism population sampling, etc. [25,26].

During the response to the DWH oil spill, submerged oil sampling plans were generally either pre-determined or adaptive [27–29]. For the pre-determined sampling, sampling locations are determined before sampling. Such a sampling plan is generally designed for a broad search around the spill location. For example, the NOAA FSV Gunter cruise [30] used a fixed sawtooth path around the well site to detect the vertical submerged oil distribution between 900 and 1200 m from 27 May to 4 June 2010. Alternatively, using an adaptive sampling plan, sampling locations are changed adaptively based on prior observations or knowledge. For example, in the DWH spill, Payne and Driskell [31] developed an adaptive sampling procedure that followed the previously observed south-southwestern movement of oil, to determine subsequent sampling locations. The NIUST (National Institute for Undersea Science and Technology) oil response team monitored the deep submerged oil plume using a modified station procedure based on in-situ sensor data collected by the R/V Pelican cruise in May 2010 [32]. The modified station procedure was that the ship followed the direction in which oil was continuously detected and then changed directions when no oil was detected.

Oil spill models are useful in estimating submerged oil distributions and thus assist in designing sampling plans. However, most models are based on results obtained from hydrodynamic models, which may not reflect conditions during the emergency response period reliably [10]. For example, during the DWH spill, there was no particle tracking model that was used by U.S. Coast Guard for the submerged oil response [9]. Additionally, until there were measured temperature and salinity profiles for use with circulation models, such models produced disparate results. Moreover, large-scale ocean circulation models are generally more diffusive than in the real ocean, to maintain model stability [5]. For example, the simulated concentrations using the Hybrid Coordinate Ocean Model (HYCOM) and the Navy Coupled Ocean Data Assimilation (NCDA) are lower than field observations in a tracer study in the Gulf of Mexico, and the locations are “smeared” by the diffusion into a larger area [33].

Alternatively, SOSim is a statistical model and is designed for real-time predictions of the location of the droplet layer using field observations of oil locations and concentrations as input, together with the output of another oil spill model if available, to infer the spatial distribution of submerged oil within the isopycnal as a function of time [34–37] using Bayesian methods [10,38]. Instead of using velocity field from hydrodynamic model, SOSim infers the velocity field by assimilating the field observations on submerged oil concentration and locations. The SOSim model is based on two assumptions: (1) the movement of submerged oil in the open ocean can be modelled by the superposition of up to four distinct forecast patches of submerged oil, each represented by dimensionally-correlated dispersive

and time-varying advective components of transport; and (2) that the effect of fate processes can be sufficiently accounted for by internal Bayesian estimation of model parameters based on the input field data, reflecting current environmental and oil conditions, and the output of a selected hydrodynamic/fate/transport model. The output of SOSim is a probability map that presents the relative probability of finding submerged oil across the area of interest, and the assessment can be updated as new observations become available. Thus, SOSim is designed to help find the submerged oil in the ocean and is suitable for oil spill response to guide sampling collections.

In this paper, sampling plans are designed for two phases. For the first phase, methods include the systematic random, adaptive systematic, modified station, and adaptive tracking path methods are proposed. For the second phase, systematic sampling combined with the kriging-based method is used, as well as the naive zig-zag method. The difference in the two phases lies in that the SOSim forecasts are available and can be used to assist sampling in the second phase. The effectiveness of different sampling plans is evaluated by comparing corresponding SOSim predictions using these sampled data with real distributions. The OSCAR simulation of the DWH spill [39] is regarded as the real distribution, because the field observations of submerged oil concentrations collected following the DWH oil spill were sparse in space and time. In this case study, the outputs of other oil spill models were not used as prior information and only field observations were provided as input for SOSim.

This paper is organized as follows. In the next section, the sampling methods are presented for both the first and second phases. The proposed sampling plans are then demonstrated using OSCAR-simulated data of the DWH spill, and the results are described and discussed in Section 3. Discussion of different proposed sampling plans and conclusions are given in Section 4.

2. Methods

2.1. Data Source

The submerged oil distribution simulations at a depth of ~1000–1300 m from 5 May to 9 May, 2010 of the DWH oil spill by OSCAR [39] were taken as the real distributions. OSCAR is a comprehensive 3-D subsea oil spill model [40]. The resolution used in this case study is 0.005° in longitude, 0.004° in latitude, and the time step is 1 day. The DWH oil spill began on 20 April 2010 and stopped on 15 July 2010. The submerged oil was found in early May. The submerged oil was simulated by OSCAR beginning on 5 May. The simulation results on 5 May, 6 May, and 9 May are plotted in Figure S1. In this study, the first phase sampling plan is used to sample from the simulation on 5 May. The second phase sampling plan is used to sample from simulations on 6 May and 9 May.

2.2. Sampling Procedures

The proposed sampling plans for the two phases of sample collections during the oil response are shown in Figure 2. The first phase sampling plan is made before a SOSim prediction is available, and the second phase sampling plan is guided by SOSim predictions. First, the density of the submerged oil isopycnal layer is determined, which indicates the depths at which to sample. After that, the first phase sampling methods (systematic random, adaptive systematic, modified station combined with random and systematic sampling, and three zig-zag paths methods) are used to collect the initial samples. The collected samples in the first phase are then input to SOSim to make 1–3 days' forecasts. Such a time gap for forecast is chosen due to the decrease in accuracy of input information for longer future predictions [41]. An area with high probability of oil presence is determined by SOSim and the second phase sampling methods (systematic random combined with the kriging-based sampling method and a naive zig-zag method) are used to determine where to sample. The samples are input to SOSim to make further real time predictions, which are then used for subsequent samplings.

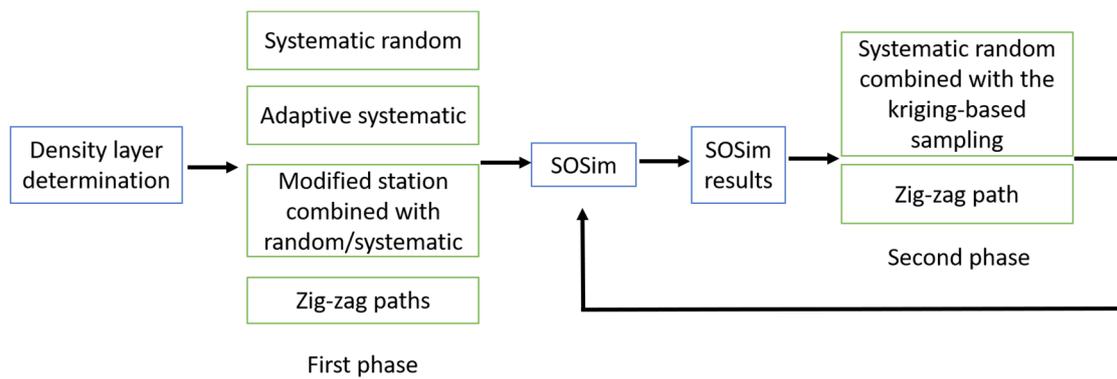


Figure 2. The proposed sampling plans for the two phases.

2.2.1. Determination of the Density Layer

Below the mixed layer in the ocean, the effects of wind are not significant and stratification dominates. The submerged oil thus travels along the isopycnal layer, which may be tilted in the deep ocean due to the geostrophic balance. The potential density ranges where the submerged oil is located determine the potential depths of finding submerged oil. The density ranges can be determined by salinity, temperature observations or be predicted by plume models. After that, we follow the isopycnal to find the potential depths of submerged oil. For the DWH case study here, the density of the isopycnal layer is known as 1027.62–1027.70 kg/m³ based on observations, corresponding to the depth range of 1000–1300 m [42].

2.2.2. First Phase Sampling Plan

This section describes the sampling plan to guide the initial collection of field data. In total, 36 data points are collected on 5 May 2010, across the area of interest, which is in the range of −88.57° W–88.17° W and 28.54° N–28.94° N. The number of data points are determined by the sampling ability of devices [32] and the number of samples SOSim needs for an accurate prediction. These data are used to reveal the distribution of submerged oil and for informing the subsequent sampling plan design. For this first phase of sampling, the systematic random, adaptive systematic, modified station combined with random sampling, modified station combined with systematic, and three zig-zag path methods are applied and compared. The systematic random and adaptive systematic methods sample across the area of interest. The modified station sampling method applies the algorithms to find the oil plume boundary, and then the systematic random and random sampling methods are applied to sample within the boundary. The zig-zag paths are designed within the search boundary found by the modified station plan method. In this work, the path to visit all sampling locations is determined by solving the ‘traveling salesman problem’ using the well-known nearest neighbor algorithm. That is, the path starts from an initial location, connects to the closest node, then extends by iterating until return to the initial sampling location. The method quickly yields a short tour, although the tour may not be optimal. In field implementation, if supported by onboard or remote computing, a more optimal solution can be generated to save significant time when collecting samples [43].

(1) Systematic random sampling plan

The area of interest is evenly divided into a 6 × 6 grid. One sample is randomly selected within each grid cell. The sampled data are plotted as red dots in Figure S2a and the designed sampling path is plotted in Figure S2b.

(2) Adaptive systematic sampling plan

The area of interest is evenly divided into a 3×3 grid. At each grid cell, one point is randomly chosen. The sampled points are plotted as red dots in Figure S3. Further sampling is randomly made 0.03° in latitude and longitude around the locations where oil is detected.

(3) Modified station sampling plan

A modified station sampling plan was used for sample collection during the DWH oil response [32], and a similar concept is adopted here. The ship or AUV begins sampling in a certain location and moves around to find the boundary of submerged oil. The moving directions of the ship or AUV are based on empirical considerations and field detected oil concentrations. Here, we take the directions to be from -10° to 30° (polar coordinate with respect to the east direction), which correspond to the two outer search boundary lines in Figure S4. The predetermined stations are taken at the cross points of the concentric rings, which are 5 km apart from 9 to 48 km away from the spill site, and the half lines, which are 10° apart from -10° to 30° . The vessel starts sampling along the southernmost half line 9 km away from the well location, and moves along the same half line until oil is not detected. Then, the vessel moves northward along the concentric ring to the next half line. If oil is detected, the vessel keeps moving outward along the same half line until oil is not detected, otherwise the vessel moves ~ 5 km inward at the same radial direction until oil is detected again. Subsequently, the vessel keeps moving northward to the next half line, until the other side of the submerged oil boundary is reached. The boundary of submerged oil layer is thus determined.

A variation of the modified station plan discussed above is made so that samples are also collected within the boundary. Inside the boundary, systematic random and random sampling are applied to collect data within the mesh grids in Figure S4.

(4) Zig-zag paths

In addition, two types of zig-zag paths are designed to continuously track the submerged oil within the areas bounded by the two the half lines with polar angle -10° to 30° : naive zig-zag and adaptive zig-zag patterns.

Naive zig-zag pattern. Using the naive zig-zag pattern, the vessel follows pre-defined track lines, which are represented by the green lines in Figure S5.

Adaptive zig-zag pattern. The adaptive zig-zag pattern is a modification of the naive zig-zag pattern. The path is changed adaptively based on real time oil detections. We propose two adaptive algorithms: I and II. Both algorithms are shown in Figure S6a,b. Overall, compared with the naive zig-zag pattern in Figure S5, the device travels a shorter distance for the adaptive plans. To get these data for SOSim input, we uniformly selected the samples within each straight track line segment along the sampling path with oil. The sampled points are the red dots in Figure S7.

2.2.3. Second Phase Sampling Plan

The locations and concentrations of the initial data collected in the first phase are input to SOSim for predictions. The predictions provide a region with higher probability of finding submerged oil than other places. In the second phase, sampling is conducted within this region. Two approaches are used for the second phase sampling, the systematic random sampling combined with the kriging-based sampling method and the naive zig-zag sampling method. The sampling results are then input to SOSim to update the prediction, either once or iteratively as more observations are available.

(1) Systematic random combined with the kriging-based sampling

First, the systematic random sampling method is applied to sample within the SOSim predicted region of submerged oil. The locations of these sampled data and their concentrations predicted by SOSim are input to a kriging model to generate more sampling locations. That is, the kriging method is evaluated here as one approach to select points in the ocean at which a responder would collect

additional water samples to be assessed for submerged oil concentration. The measured concentrations and coordinates of these samples would then be used as input to SOSim, to update the SOSim output. SOSim would use the newly enlarged dataset, to compute new estimates of the parameters with which SOSim would generate a new forecast. The detailed explanations of kriging method are in the Supplementary Materials.

In detail, based on the systematic random sampled data within the area predicted by SOSim $\{x_i, i = 1, \dots, M\}$, kriging interpolations can be done to generate an estimation $\hat{C}(x)$ of the submerged oil concentration over the area of interest and we get a kriging variance Var_1 . To add new samples, an optimization algorithm is used to find the sampling locations $\{x_i, i = M + 1, \dots, M'\}$ that reduce the kriging variance most. That is, with the new determined samples, the kriging variance Var_2 is the smallest compared to the variance using other samples of the same size. Here, we use the R library 'gstat' [44,45] to assimilate the universal kriging variance and the function 'ssaOptim' of the library 'intamapInterative' [46] to find the optimal sampling points that minimize the kriging variance [17]. The method is applied to design a sampling plan on 6 May. The sampling locations and the sampling path are given in Figure S7. Total distance is 184 km.

(2) Naive zig-zag sampling

As an alternative, a naive zig-zag path, shown in Figure S5, to design the sampling path within the SOSim-forecasted region is used. Notice here the path is not determined by two radial boundaries but by the boundary of the SOSim predicted region. The designed path for 6 May, and the resulting pattern of sample collection, are shown in Figure S8. The distance of the tracking path is 442 km.

2.3. Sampling Plan Effectiveness Evaluation

To evaluate the sampling effectiveness of the proposed sampling methods, we compare the SOSim predicted submerged oil distributions with the OSCAR simulations. In this section, we define several quantities to compare SOSim predictions with the OSCAR simulations: (a) the relative prediction error, (b) the extent distance, and (c) the extent area of the submerged oil distribution and the success rate. The relative error compares the relative concentrations assessed by SOSim with OSCAR data. The extent distance and area compare the locations of submerged oil. The success rate, defined as the percentage of grid points containing oil that are enclosed within the SOSim confidence bound agreeing with OSCAR data, measures the accuracy of SOSim predictions. These quantities are combined to produce a comprehensive comparison of SOSim predictions and OSCAR simulations. In addition to the above quantities, we also define the sampling efficiency as the percentage of sampling locations at which submerged oil is found.

2.3.1. Relative Error

SOSim predicts the relative concentration (probability) (C_{mi}) of the submerged oil on the grid points of OSCAR (x_i). OSCAR concentration data are in parts per billion (ppb). To compare these two values, OSCAR data are normalized such that all the concentrations of all these data sums to 1. In order to evaluate the accuracy of SOSim predictions, we define a relative error given by

$$\varepsilon = \sum_{i=1}^I \text{abs}(C_{rr_i} - C_{m_i}). \tag{1}$$

Here, the relative concentration (C_{rr_i}) of OSCAR at the point (x_i) is defined as the real concentration (C_{r_i}) divided by the sum of the concentrations at all data coordinates ($\sum_{i=1}^I C_{r_i}$). The error (ε) is the summation of the error at all coordinates. SOSim provides two predictions using parameters given by the maximum likelihood estimation and by the 95% confidence bound of the maximum likelihood estimator, so errors are computed for both predictions. We use ε_b to denote the error of SOSim prediction using the optimal parameter, and ε_m for the error of SOSim prediction using the 95% confidence bound.

2.3.2. Extent Distance and Area

The extent distance of the submerged oil is defined as the furthest distance of submerged oil from the release location. Here, the locations where the relative concentration is smaller than 1% in the SOSim predictions and the locations where the concentration is smaller than 1 ppb in OSCAR data [47] are regarded to be free of oil. The Hausdorff distance [48] is used to compute the furthest distance from the boundary of the submerged oil area (Y represents the predicted boundary) to the spill location (x_0).

$$d(x_0, Y) = \sup_{y \in Y} d(x_0, y), \tag{2}$$

in which \sup represents supremum. The real extent distances of OSCAR (d_r) and SOSim prediction (d_m) are calculated by Equation (2).

The relative error of the extent distance is defined as:

$$\varepsilon_d = \frac{d_r - d_m}{d_r} \tag{3}$$

The area of the submerged oil is defined as the submerged oil extent area. The relative error of the extent area is defined as:

$$\varepsilon_s = \frac{s_r - s_m}{s_r}, \tag{4}$$

in which s_r is the real extent area of OSCAR data, and s_m is the extent area of SOSim prediction. The extent distance and area are also computed for both SOSim predictions with the maximum likelihood parameter (ε_{db} and ε_{sb}) and the 95% confidence bound (ε_{dm} and ε_{sm}).

2.3.3. Success Rate

The success rate is computed using the Peirce skill score (PSS) by [49,50]:

$$PSS = \frac{(ad - bc)}{(a + c)(b + d)}, \tag{5}$$

in which a is the number of data points predicted with oil by both OSCAR and SOSim, b is the number of data points with oil predicted by SOSim and without oil predicted by OSCAR, c is the number of data points with oil predicted by OSCAR and without oil predicted by SOSim, and d is the number of data points without oil predicted by OSCAR and SOSim. The success rate represents the proportion of the overlaid area of submerged oil between the SOSim prediction and the real oil distribution. When ε_o equals 100%, SOSim's prediction covers the entire real oil distribution. The success rate ε_o is computed for predictions using both maximum likelihood (denoted by ε_{ob}) and 95% confidence bound parameters (denoted by ε_{om})

2.3.4. Sampling Efficiency

The sampling efficiency is defined here as the proportion of data points with oil detected across the entire sample. We use this quantity because those data points without oil detected give less information on the submerged oil distribution, in some sense. However, the efficiency of the sampling plan does not only depend on the sampling efficiency we defined here, but also depends on the distance of sampling routes, the concentrations of sampled data, and also the representativeness of sampled data.

3. Results and Discussion

3.1. Initial Sampling Plan

3.1.1. Systematic Random Sampling

We repeated the sampling 30 times and the average sampling efficiency of these 30 repeated tests using the systematic random method for 5 May is $4.91 \pm 3.40\%$. SOSim infers the submerged oil distribution for 5 May by inputting the systematic random sampled data in Figure S2 and the result is plotted in Figure 3. In Figure S2b, the black dots represent OSCAR data, the filled contours represent the relative oil concentration predicted by SOSim using the maximum likelihood parameter and the green line represents the area with relative oil concentration greater than 1% at the 95% confidence level. As can be seen from Figure S2b, this area covers the real distribution but is broader. This is partially because the predicted submerged oil distribution is dispersed in two separate directions, which does not agree with the real distribution. The contours show that the submerged oil exists at ~1000–1300 m depth in the area of interest.

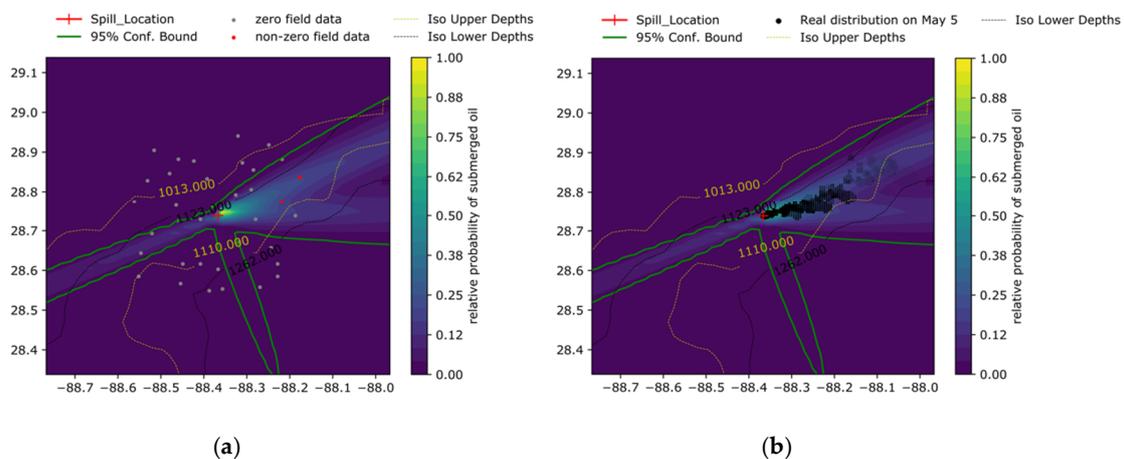


Figure 3. SOSim prediction for 5 May using the systematic random sampling (a); and SOSim prediction overlaid with the real distribution on 5 May (b).

3.1.2. Adaptive Systematic Sampling

The average sampling efficiency of 30 repeated tests using the adaptive sampling method is $40.46 \pm 17.37\%$. The SOSim prediction for 5 May using the data shown in Figure S3 is plotted in Figure 4. These sampled data are clustered around the spill location, which leads to the inferred distribution narrower than the real distribution. The submerged oil area of the OSCAR data is not fully covered by the SOSim prediction. These sampled points using the adaptive sampling plan may be concentrated in a small region and thus lack spatial variability.

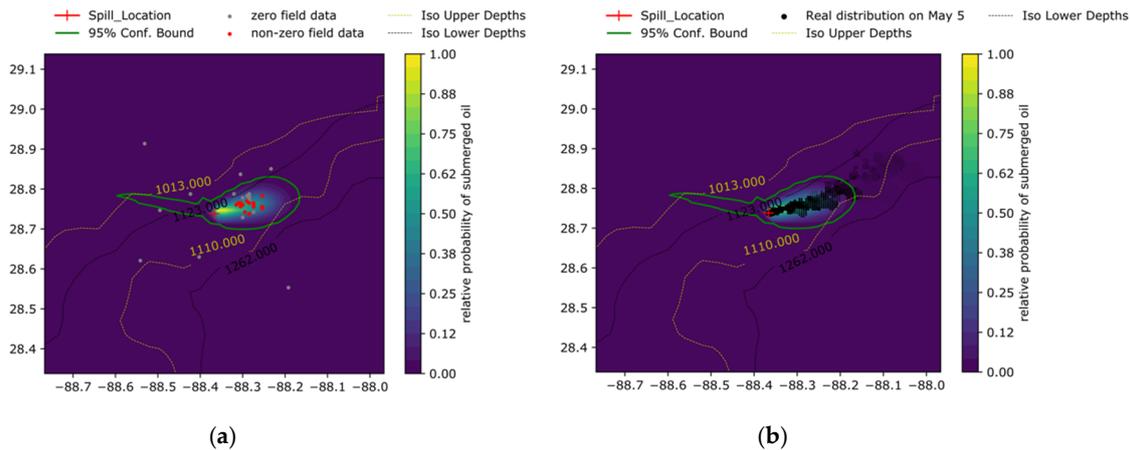


Figure 4. Subsurface Oil Simulator (SOSim) prediction for 5 May using the adaptive sampling method (a); and SOSim prediction overlaid with the real distribution on 5 May (b).

3.1.3. Modified Station Plan Combined with Systematic Random Sampling

The average sampling efficiency of 30 repeated tests using the modified station plan combined with the systematic random sampling is $44.54 \pm 4.92\%$. The SOSim prediction on 5 May using data in Figure S4a is plotted in Figure 5. The prediction covers most of the submerged oil area of all the OSCAR output and shows a similar moving tendency as the real distribution.

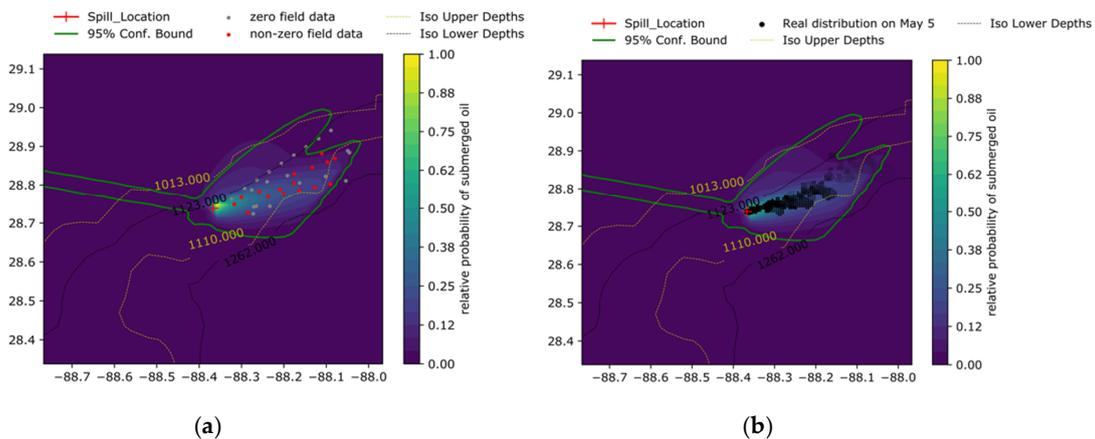


Figure 5. SOSim prediction for 5 May using the modified station plan combined with synthetic random sampling method (a); and SOSim prediction overlaid with the real distribution on 5 May (b).

3.1.4. Modified Station Plan Combined with Random Sampling

The average sampling efficiency of 30 repeated tests using the modified station plan combined with the random sampling is $41.11 \pm 6.21\%$. The SOSim prediction result for 5 May using data in Figure S4b is plotted in Figure 6. The prediction also covers the submerged oil area of all the OSCAR output. Compared to the result in Figure 5, the predicted submerged oil area here is narrower and is more accurate.

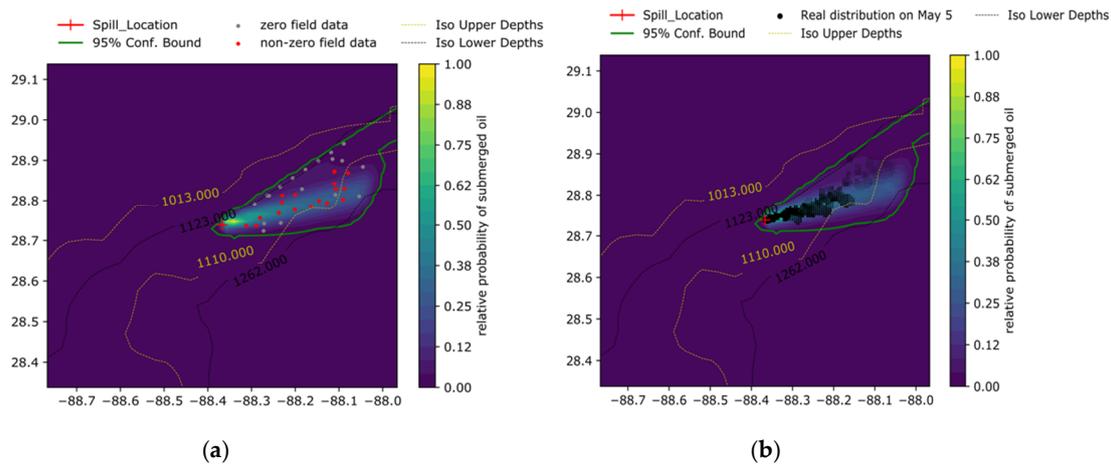


Figure 6. SOSim prediction for 5 May using the modified station plan combined with random sampling method (a); and SOSim prediction overlaid with the real distribution on 5 May (b).

3.1.5. Zig-Zag Paths

Three types of zig-zag paths described in Section 2.2.2 are applied to sample submerged oil on 5 May: the naive zig-zag pattern and the adaptive zig-zag patterns I and II. The sampling efficiencies of these sampled data using the three different paths all reach 100%. The SOSim predictions for 5 May, following the naive zig-zag pattern, the adaptive I pattern, and the adaptive II pattern are plotted in Figures 7–9, respectively. In these Figures, the real distribution is represented by the black dots and is overlaid with the predicted distribution. In general, all the results follow the moving tendency and cover the area of OSCAR data. For both the real and predicted distributions, the highly concentrated areas are both in the east and northeast areas near the spill location. The yellow and black dashed lines represent the depths of the submerged oil on 5 May, which are ~1000–1300 m.

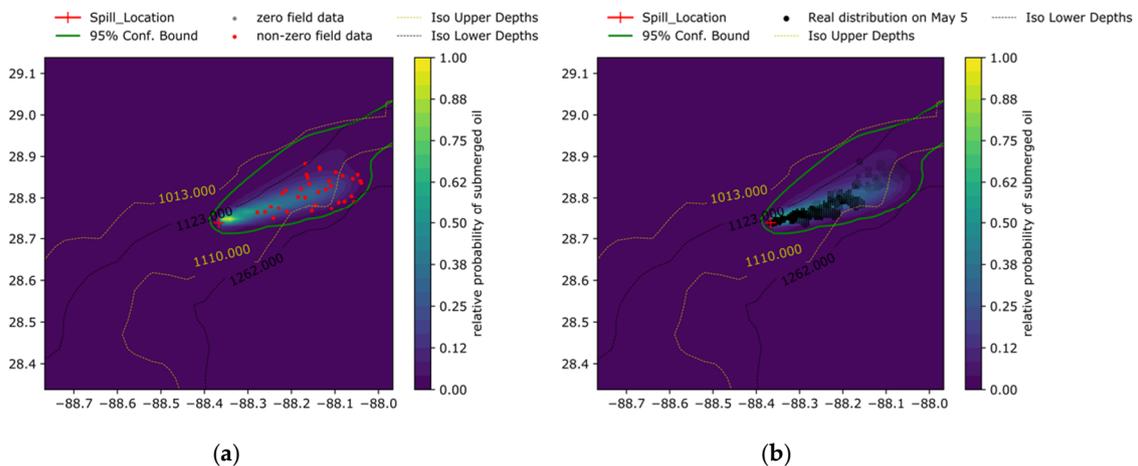


Figure 7. Left: SOSim prediction for 5 May following the naive zig-zag path sampling method (a); and SOSim prediction overlaid with the real distribution on 5 May (b).

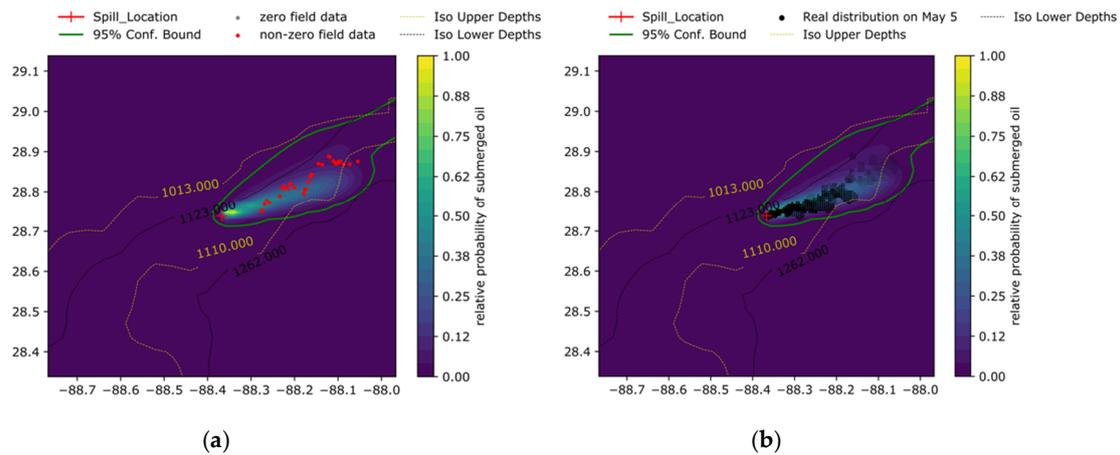


Figure 8. SOSim prediction for 5 May following the adaptive I zig-zag path sampling method (a); and SOSim prediction overlaid with the real distribution on 5 May (b).

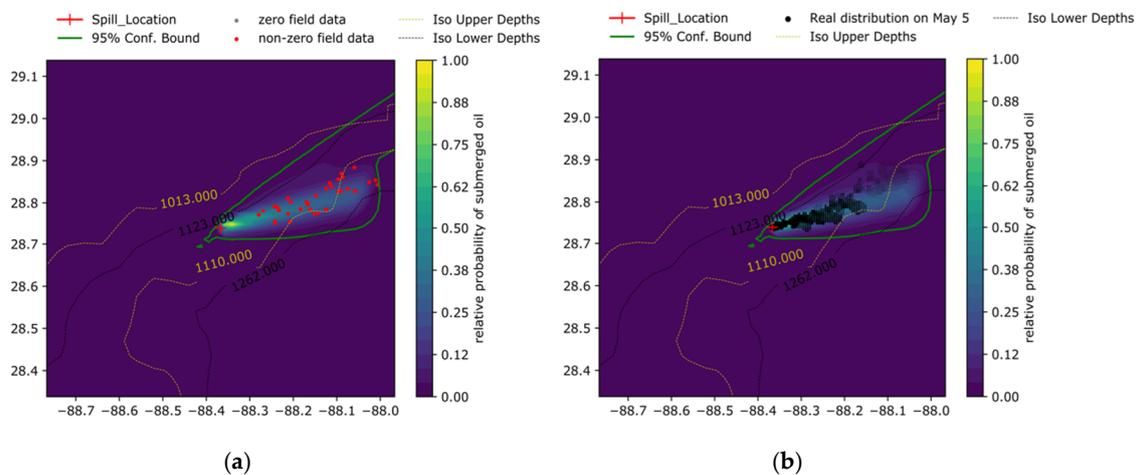


Figure 9. SOSim prediction for 5 May following the adaptive II tracking lines (a); and SOSim prediction overlaid with the real distribution on 5 May (b).

3.1.6. Comparisons of SOSim Predictions with Different Sampling Methods

Comparisons of SOSim predictions for 5 May using sampled data from the different proposed sampling methods are discussed here. In general, the predicted submerged oil area lies mainly northeast of the spill site, which agrees with OSCAR simulations. For more detailed comparisons, the relative error, the submerged oil extent distance and area, the success rate and the sampling effectiveness defined in Section 2.3 are computed, and the results are summarized in Table 1.

The relative error. The errors ε_b and ε_m for all the methods are all around 1.0 and the minimum error is given by the adaptive systematic sampling. This is because the adaptive systematic sampling has more data in the concentrated area with higher submerged oil concentrations and, thus, gives a better representativeness of the submerged oil distributions.

The extent of distance and area. The extent of distance and area using the maximum likelihood parameter and the 95% confidence bound are shown in the columns d_{0b} , S_b and d_{0m} , S_m in Table 1. The S_b of the systematic random and modified station plan combined with random sampling as well as the zig-zag II sampling are closest to the real extent area, S_0 . The S_b of the adaptive systematic sampling is the smallest and is significantly smaller than S_0 compared with other methods. As for the extent distances, the d_{0b} of all the methods are similar to d_0 except the systematic random sampling plan, whose d_{0b} is almost the double of d_0 and the adaptive systematic sampling plan, whose d_{0b} is almost half of d_0 . Combining the above comparisons with the figures in the previous section, we can see that

the prediction of the systematic random method has two separate patches, making the submerged oil extent and distance broader than the real distribution. The sampled points from the adaptive systematic method are more concentrated and thus the prediction has a smaller extent area and distance. Therefore, modified station methods and zig-zag pattern methods give better performances in determining the extent area and distance of submerged oil.

Table 1. SOSim prediction error for May 5 compared with the real distribution *.

Method	ϵ_b	ϵ_{ob} (%)	ϵ_m	ϵ_{om} (%)	S_b (km ²)	S_m (km ²)	ϵ_{sb} (%)	ϵ_{sm} (%)	d_{ob} (km)	d_{om} (km)	d_{db} (%)	d_{dm} (%)
Systematic random	1.01	32.47	1.01	32.79	643.51	1069.84	12.47	86.99	72.28	66.50	95.40	79.78
Adaptive systematic	0.73	15.04	0.72	16.17	179.38	314.78	-68.65	-44.98	20.44	31.44	-44.74	-15.00
Modified & systematic	0.94	33.84	0.91	15.89	694.88	1053.39	21.45	84.11	39.05	56.06	5.57	51.55
Modified & random	1.03	59.29	1.02	45.50	480.53	755.16	-16.01	31.99	41.38	65.45	11.87	76.94
Zig-zag I	0.98	45.02	0.98	45.38	389.34	856.61	-31.95	49.72	37.96	65.51	2.62	77.10
Zig-zag I	0.98	45.02	0.98	44.92	389.34	821.08	-31.95	43.51	37.96	65.36	2.62	76.70
Zig-zag II	1.03	59.27	1.04	59.21	480.53	964.67	-16.01	68.60	41.38	67.22	11.87	81.72

* Here, ϵ_b is the SOSim prediction error using the optimal parameter; ϵ_{ob} is the success rate using the optimal parameter; ϵ_m is the error of SOSim prediction using the 95% confidence bound; ϵ_{om} is the success rate using the 95% confidence bound; S_o is the extent of area predicted by OSCAR on 5 May, which is 572.15 km²; S_b is the extent of area predicted by SOSim for May 5 using the optimal parameter; S_m is the extent of area predicted by SOSim for 5 May using the 95% confidence bound; ϵ_{sb} is the extent of area relative error using the optimal parameter; ϵ_{sm} is the extent of area relative error using the 95% confidence bound; d_o is the extent of distance predicted by OSCAR on 5 May, which is 36.99 km; d_{ob} is the extent of distance predicted by SOSim for 5 May using the optimal parameter; d_{om} is the extent of distance predicted by SOSim for 5 May using the 95% confidence bound; d_{db} is the extent of distance relative error using the optimal parameter; d_{dm} is the extent of distance relative error using the 95% confidence bound.

To account for the uncertainty, the predictions using the 95% confidence bound are calculated. The extent distances and areas of these predictions are broader than those of these OSCAR data for most sampling methods, except for the adaptive systematic method. The broader predictions may be helpful to assist oil responders in finding the sensitive areas of low concentrations of oil that create high risks near particularly sensitive habitats, such as the coral reefs. Among these sampling methods, the systematic random method gives the broadest prediction results among the sampling methods proposed in the paper. The prediction using the systematic sampled data is 86.99% broader in area and 79.78% larger in distance, compared to the OSCAR data distribution. The corresponding values for the modified station plan combined with the systematic sampling plan are 84.11% in area and 76.94% in distance. The other sampling methods have 31.99% to 68.60% broader extent areas than the OSCAR data distribution. The systematic methods have a broader extent distance and area because the systematic sampling method is applied over the whole area of interest, and these sampled data are more spread over the whole area.

The success rate. The success rate is shown in Table 1 with columns ϵ_{ob} and ϵ_{om} for the prediction using the maximum likelihood estimate parameter and the 95% confidence bound. When using the maximum likelihood estimate parameter, the zig-zag II method has the highest value of 59.27%, and the adaptive systematic method has the lowest value of 15.04%. The values for the other methods are in the range 32–45%. When using the 95% confidence bound, most of the methods have a similar ϵ_{ob} value, except for the modified station plan combined with the systematic method. The zig-zag methods have high success rates, and the adaptive method has the smallest success rate because this method.

The sampling effectiveness. The sampling effectiveness of the zig-zag paths is the largest with 100%, because the sampling paths avoid the locations without oil. The adaptive systematic sampling has a sampling effectiveness similar to that of the modified station sampling method, which is around 40.46–44.54%. The systematic random method has the lowest efficiency, about $4.91 \pm 3.4\%$.

This is because the systematic random method has more widespread samples across the whole region of interest.

Accounting for the distance of the sampling path, the naive zig-zag path takes the longest distance at 236 km. The systematic random method takes the second longest distance with 204 km. The modified station plan combined with the random sampling method takes 199 km in total distance. The total distance for the adaptive systematic sampling is 151 km. The total distances for the other three methods (zig-zag I, zig-zag II, and modified station combined with the systematic sampling method) are shorter than 100 km. The zig-zag II path travels the shortest distance of 28 km.

In summary, the zig-zag patterns are better than most environmental sampling methods. Among the environmental sampling methods, the modified station combined with the random method performs the best. The systematic random method has a low sampling effectiveness, and the adaptive systematic method is occasionally too concentrated and lacks spatial representativeness. The modified station plan combined with systematic sampling improves the efficiency of the systematic random method but may also have a broader predicted submerged oil area.

3.1.7. Comparisons on 6 May and 9 May

The parameters inferred by assimilating the samples on 5 May in the first phase are further applied for submerged oil distribution predictions on 6 May and 9 May. The prediction errors of all the proposed first phase sampling methods for 6 May and 9 May are summarized in Table 2. The errors increase as SOSim forecasts the submerged oil distribution into the future. The systematic method predicts the broadest area. Still, the zig-zag patterns are better than the other environmental sampling methods in terms of the success rate and the errors. Hence, in this case study, the prediction result using the naive sampling path is used for the subsequent sampling plan design in the second phase sampling. The predictions using the naive path method for 6 May and 9 May are plotted in Figures S11 and S12, respectively. The predictions agree well with the simulation result of OSCAR.

Table 2. SOSim prediction error for 6 May and 9 May compared with the real distribution by the first phase sampling plan *.

Dates	6 May				9 May			
	ϵ_b	ϵ_{ob} (%)	ϵ_m	ϵ_{om} (%)	ϵ_b	ϵ_{ob} (%)	ϵ_m	ϵ_{om} (%)
Systematic random	1.30	27.82	1.36	60.42	1.53	38.62	1.52	45.73
Adaptive systematic	1.02	18.52	1.02	20.72	1.34	19.36	1.32	19.76
Modified & systematic	1.20	29.33	1.16	33.30	1.53	42.70	1.58	54.21
Modified & random	1.25	47.61	1.54	17.21	1.57	43.99	0.98	2.01
Zig-zag naive	1.26	46.98	1.27	44.18	1.57	50.65	1.55	50.40
Zig-zag I	1.25	46.98	1.27	50.86	1.57	50.65	1.57	51.59
Zig-zag II	1.25	47.17	1.27	47.61	1.57	43.97	1.55	35.38

* Here ϵ_b is the SOSim prediction error using the optimal parameter; ϵ_{ob} is the success rate using the optimal parameter; ϵ_m is the error of SOSim prediction using the 95% confidence bound; ϵ_{om} is the success rate using the 95% confidence bound.

3.2. Second Phase Sampling Plan

In this section, we consider the sampling plans after the first phase samples are collected. We compare the design of sampling plans within the region determined by SOSim predictions with those within the whole region without the considerations of SOSim predictions. Two sampling methods are applied for the use of SOSim predictions using the 95% confidence bound parameter: a systematic random method combined with the kriging-based sampling and a naive zig-zag path.

3.2.1. Sampling Plans without Using the SOSim Prediction

The systematic random sampling method is used to sample data on 6 May and 9 May. No information on the potential polluted area is known, so the sampling is carried out over the whole region of interest, which is around 0.2° and 0.4° in both latitude and longitude around the spill location for 6 May and 9 May, respectively. The sampling results are plotted in Figure S13. The average sampling efficiencies of the 30 repeated sampling using the systematic random sampling method are $6.57 \pm 2.87\%$ and $5.56 \pm 3.18\%$ on 6 May and 9 May, respectively. These sampled data are input to SOSim and the predictions for 6 May and 9 May are plotted in Figures S12 and S13, respectively. The predictions show that the submerged oil is distributed along separated directions. On 6 May, the oil is distributed in the northeast and southwest directions and on 9 May, the oil is distributed in the west, east and south directions.

3.2.2. Sampling Plan with SOSim

After the first phase samples are collected, SOSim makes a prediction on the submerged oil distribution for May 6. In this section, two sampling methods are used to design the sampling plans within the area with submerged oil predicted by SOSim.

The systematic random combined with the kriging-based sampling. These sampled data are plotted in Figure S14. The sampling efficiency is 22.22%. The updated SOSim prediction using these sampled data on 6 May is plotted in Figure 10.

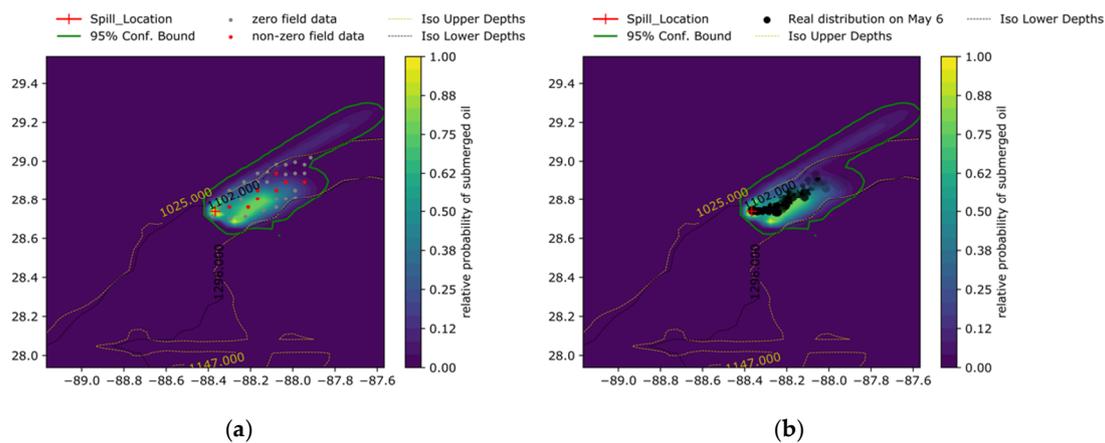


Figure 10. The SOSim prediction for 6 May using systematic random combined with the kriging-based sampling method (a); and the SOSim prediction overlaid with ‘real’ distribution on 6 May from oil spill contingency and response (OSCAR) (b).

These sampled data from 6 May are input to SOSim to make a prediction for May 9. Samples are drawn using the same method on 9 May within the SOSim predicted submerged oil area. The sampling efficiency for this case is 19.44%. The updated prediction for 9 May using these sampled data is plotted in Figure 11. The result reasonably predicts the diffusion of the submerged oil, as well as the moving tendency. The prediction covers the real submerged oil area from the OSCAR simulation well.

Zig-zag sampling. The samples obtained by the naive zig-zag path are discretized to provide inputs to SOSim. These data are plotted in Figure S15. The updated SOSim prediction using these sampled data for 6 May is plotted in Figure 12. In Figure 12, the predicted submerged oil area is narrower in the northwestern-southeastern direction. Compared to Figure 9, the prediction in Figure 12 is more accurate. A detailed comparison of the two sampling methods are given later in this section.

Applying the same methods to design a sampling plan for 9 May, the updated prediction using these sampled data is plotted in Figure 13. The result also reasonably predicts the moving direction of the submerged oil. Compared with Figure 11, the result here is more accurate and fits the real distribution better.

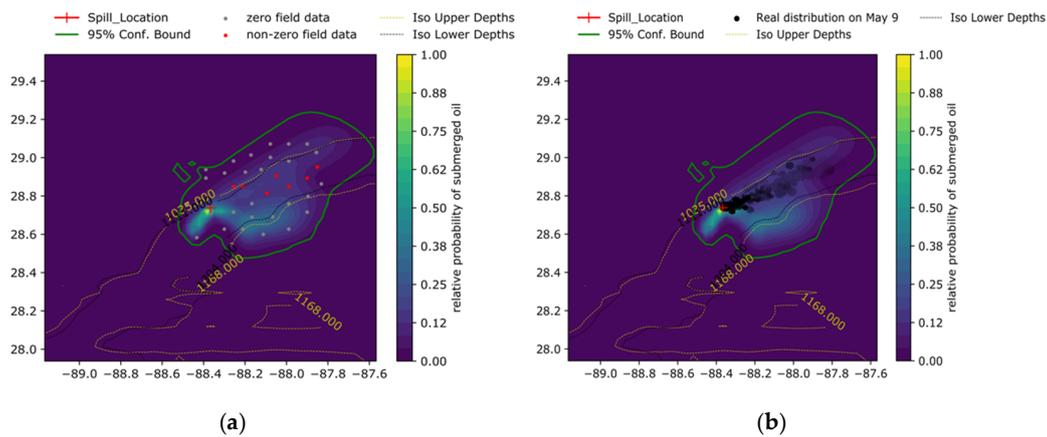


Figure 11. The SOSim prediction for 9 May using systematic random combined with the kriging-based sampling method (a); the SOSim prediction overlaid with ‘real’ distribution on 9 May from OSCAR (b).

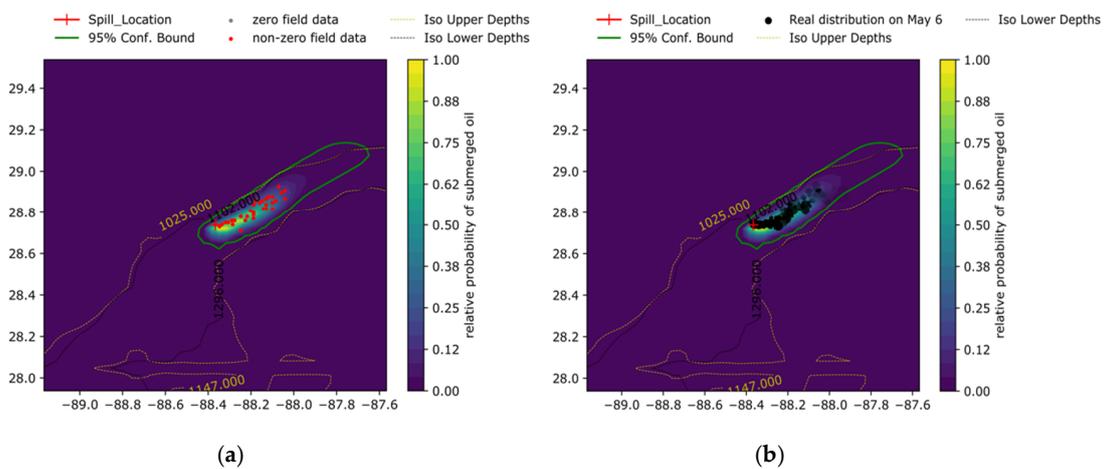


Figure 12. The SOSim prediction for 6 May following the naive zig-zag path sampling method (a); and the SOSim prediction overlaid with ‘real’ distribution on 6 May from OSCAR (b).

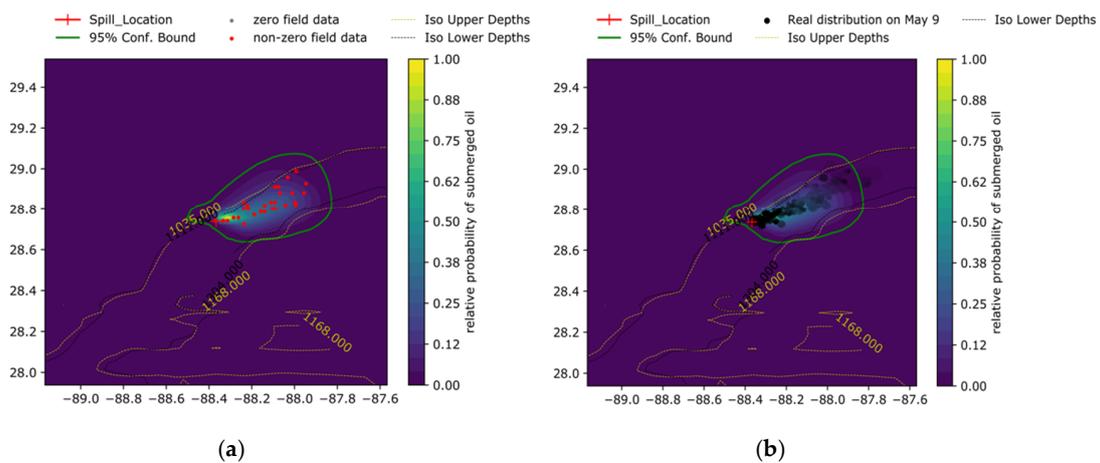


Figure 13. The SOSim prediction for 9 May following the naive zig-zag path sampling method (a); and the SOSim prediction overlaid with ‘real’ distribution on 9 May from OSCAR (b).

3.2.3. Second Phase Sampling Plans Comparisons

Comparing Figures 10–13 (where SOSim’s predictions are used to guide sampling) with Figures S14 and S15 (where the predictions of SOSim are not used), the use of the SOSim’s predictions improves the accuracy of the predictions significantly. In Figures S14 and S15, the predicted distributions of submerged oil are different from the real distributions in both trajectory, area and shape. The results shown in Figures 10–13 are closer to real distributions.

In general, both the systematic random combined with the kriging-based and the naive zig-zag pattern produce reasonable results on predicting submerged oil distributions. The area and trajectory of submerged oil are well predicted. The major difference of the two methods is that the systematic sampling combined with the kriging-based sampling method has a broader predicted area of submerged oil. This is due to the larger spatial variety of the collected samples compared with the zig-zag pattern method. For the same reason, the sampling efficiency of the systematic random combined with the kriging-based sampling method is lower than the zig-zag pattern method. Overall, there are two results: (1) the zig-zag pattern method is a more efficient sampling plan; and (2) the systematic random combined with the kriging-based sampling method produces broader distributed samples, which leads to a broader predicted submerged oil area.

In detail, the relative errors and the success rates between the three predictions and the real distribution on 6 May and 9 May are summarized in Table 3. Similar to the results of the first phase sampling plan, the model prediction following the naive path has the smallest relative errors and the narrowest area of submerged oil among the three methods. As for the success rate, the values using maximum likelihood parameters on 6 May and 9 May are 0%, 0.28% for the systematic random method, 24.67% and 34.69% for the systematic random combined with the kriging-based sampling method, and 56.96% and 56.86% for the naive zig-zag pattern. The zig-zag naive path has the highest success rate on 6 May and 9 May.

Table 3. SOSim prediction error compared with the real distribution by an autonomous underwater vehicle (AUV) for May 6 and May 9 for subsequent sampling plan *.

Method	6 May				9 May			
	ϵ_b	ϵ_{ob} (%)	ϵ_m	ϵ_{om} (%)	ϵ_b	ϵ_{ob} (%)	ϵ_m	ϵ_{om} (%)
Systematic random combined with the kriging-based method	1.56	24.67	1.57	34.69	1.75	8.29	1.72	7.38
Naive zig-zag path	1.38	56.96	1.38	56.86	1.75	35.20	1.47	36.24
Systematic across the area	1.69	0	1.69	0.28	1.75	3.80	1.65	4.82

* Here ϵ_b is the SOSim prediction error using the optimal parameter; ϵ_{ob} is the success rate using the optimal parameter; ϵ_m is the error of SOSim prediction using the 95% confidence bound; ϵ_{om} is the success rate using the 95% confidence bound.

4. Conclusions

During an oil or chemical spill, knowing where the substance is located is important for responders, and some of the material could be subsurface. Though we do not have technology to remove subsurface oil and dissolved chemicals, we may need to provide warnings to the public, such as closing fishing areas. Thus, both modeling and sampling are needed to identify the locations and make further predictions for improved warnings. Our novel Bayesian implementation method directed towards the environmental and geospatial sampling plans are applied to sample submerged oil based on SOSim predictions. These proposed plans are evaluated using a case study on the DWH oil spill,

a spill of national significance (SONS)¹. OSCAR model simulations of the DWH oil spill are regarded as the real distributions for this exercise, and are used to examine the performance of the sampling plans calculated by SOSim. Here, the sampling plans are designed for submarine devices (e.g., AUV) and ship-based sampling. We propose several sampling plans during the DWH oil spill submerged oil response period for both first phase and second phase sampling and use SOSim predictions to guide the second phase sampling. A first phase sampling procedure is designed to collect the first field observations and seven environmental sampling methods are compared, including the systematic random method, adaptive systematic method, two modified station plans, and three zig-zag patterns. After the first phase samples are collected, SOSim makes predictions using these sampled data set in the first phase and gives an estimate of the submerged oil concentrated region. Inside this region, two methods are used for the second phase sampling, including the systematic random combined with the kriging-based sampling method and the naive zig-zag pattern. The evaluations of the different sampling methods are performed by comparing the sampling efficiency and the relative error, success rate, the extent area, and distance of SOSim predictions using these sampled data. The recommendations of the sampling plans under different circumstances are summarized in Table 4.

Table 4. Recommendations of sampling plans for the two phases sampling procedures.

Phase 1			
Equipment			
Ship		AUVs	
Methods	Circumstances	Methods	Circumstances
Modified station plan	Having prior knowledge on the possible moving directions of the oil plume	Naive zig-zag	<ol style="list-style-type: none"> 1. The plume separates into different patches; 2. High battery duration; 3. Caring about the low concentration but high risk areas.
Adaptive systematic	<ol style="list-style-type: none"> 1. Caring about the concentrated area; 2. Having no prior knowledge on the possible moving directions; 3. The plume is clustered. 	Adaptive I	<ol style="list-style-type: none"> 1. Moderate battery duration; 2. Caring about the concentrated area.
Systematic random	<ol style="list-style-type: none"> 1. Having no prior knowledge on the possible moving directions; 2. The plume is dispersed; 3. Caring about the low concentration but high risk areas. 	Adaptive II	<ol style="list-style-type: none"> 1. Low battery duration; 2. Caring about the concentrated area; 3. The plume is clustered.
Phase 2			
Equipment			
Ship		AUVs	
Methods	Circumstances	Methods	Circumstances
With the Bayesian model SOSim, using systematic sampling combined with the kriging -based sampling method	<ol style="list-style-type: none"> 1. Caring about the low concentration but high-risk areas. 	With the Bayesian model SOSim, using naive zig-zag	<ol style="list-style-type: none"> 1. In common situations

¹ <https://nrt.org/sites/2/files/SONS.pdf>.

The principal findings of our research are:

- 1) The sampling locations determined using the systematic sampling methods are more broadly distributed than those determined by the other methods;
- 2) The zig-zag pattern methods for both phases sampling perform the best in terms of sampling efficiency, predicted submerged oil area, and predicted errors;
- 3) Through the SOSim's forecast comparisons in Sections 3.2.1 and 3.2.2, the use of SOSim's predictions to narrow down the area for sampling plan designs significantly improves the accuracy of the SOSim forecast on submerged oil distributions;
- 4) For the first phase sampling, the adaptive systematic, modified station, and zig-zag pattern methods generate representative samples of the submerged oil distribution, and the corresponding SOSim predictions fit these real data well; and
- 5) Both systematic random combined with the kriging-based sampling and zig-zag pattern methods for subsequent sampling are able to generate samples that represent the entire distribution of submerged oil.

For practical applications, the zig-zag pattern methods are more efficient than other environmental sampling methods, but require a continuous operation of the devices, and are, thus, more suitable for AUVs. The other methods are more suitable for ship-based sampling.

Based on the results, the three proposed zig-zag patterns are suggested to be adopted based on the sampling capability and the characteristics of the oil plume in practice. The assumption of the naive zig-zag pattern is that each coordinate along the track path has an equal probability of finding submerged oil. Naive zig-zag patterns can successfully manifest all submerged oil locations within the area of interest. The adaptive zig-zag patterns are based on the assumption that if no oil is detected 0.2 km away from the last coordinate where oil is detected along one direction, then there is no oil along that direction. When the assumption fails and there is a separate patch located further away from the main oil patch, the adaptive zig-zag patterns may miss the separated oil patch. If the oil is separated into different patches far away from each other, for example, in high latitude conditions such that the oil plume sheds from the oil plume source periodically [1], the naive zig-zag patterns would have a better performance than the adaptive zig-zag patterns. However, the naive zig-zag patterns have a longer sampling path than the adaptive zig-zag patterns, and the adaptive zig-zag patterns are more cost-effective.

Overall, the proposed sampling plans in this paper based on SOSim predictions provide a comprehensive guidance on designing sampling plans during oil responses to locate and track the submerged oil. With the collected information from the field observations, SOSim is able to make accurate predictions and provide the oil responders with a real time-updated sampling plan, and also a time-updated forecast on submerged oil distributions. The sampling approaches proposed in this paper can also find be applied to other marine sampling efforts.

Supplementary Materials: The following are available online at <http://www.mdpi.com/2077-1312/8/12/984/s1>, Figure S1: OSCAR predictions for submerged oil distributions at ~1000–1300 m depths on 5 May (a), 6 May (b), and 9 May, 2010 (c) (The shades of the dots are proportional to its concentration), Figure S2: Sampled data using the systematic random sampling plan (a) and the designed path (b), Figure S3: Sampled data using the adaptive sampling plan (a) and the designed path (b), Figure S4: Modified sampling plan combined with: (a) random sampling method (distance 199 km); (b) systematic random sampling method (distance 87 km), Figure S5: The naive zig-zag pattern (the green and orange lines are the 335.88 km path), Figure S6: Two adaptive zig-zag paths (a: 66.61 km zig-zag pattern I path; b: 27.90 km zig-zag pattern II path), Figure S7: (a) Sampled data using the systematic random sampling combined with the kriging-based sampling method and (b) the 184 km designed path, Figure S8: The 442 km naive zig zag path within the SOSim prediction for 6 May, Figure S9: SOSim prediction for 6 May following the naive zig-zag path sampling method (a); and SOSim prediction overlaid with the real distribution on 6 May (b), Figure S10: SOSim prediction for 9 May following the naive zig-zag path sampling (a); and SOSim prediction overlaid with the real distribution on 9 May (b), Figure S11: Sampled data on May 6 (a) and May 9 (b) using systematic random sampling method, Figure S12: SOSim prediction for 6 May using the systematic random sampling method (a), and SOSim prediction overlaid with the real distribution on 6 May (b), Figure S13: SOSim prediction for 9 May inputting the systematic random sampled data (a), and SOSim prediction overlaid

with the real distribution on 9 May (b), Figure S14: (a) The sampled data based on SOSim prediction for 9 May search using the systematic random combined with the kriging-based method; (b) The designed 334 km path for data collection, Figure S15: The sampled data using the naive zig-zag path method based on SOSim prediction for 9 May. The final version of SOSim is publicly available through the Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC) at (<https://data.gulfresearchinitiative.org/data/R6.x812.000:0005>). The sampling plans used in this case study are publicly available through the Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC) at (<https://data.gulfresearchinitiative.org/data/R6.x812.000:0007>).

Author Contributions: Conceptualization, C.J., J.D.E. and C.J.B.-K.; Methodology, C.J., J.D.E. and C.J.B.-K.; software, C.J. and C.J.B.-K.; validation, C.J.; writing—original draft preparation, C.J.; writing—review and editing, J.D.E. and C.J.B.-K.; supervision, J.D.E. and C.J.B.-K.; project administration, J.D.E. and C.J.B.-K.; funding acquisition, J.D.E. and C.J.B.-K. All authors have read and agreed to the published version of the manuscript.

Funding: This research and the APC were funded by the Gulf of Mexico Research Initiative (Grant ID: GR009472).

Acknowledgments: This research is supported by the Gulf of Mexico Research Initiative (Grant ID: GR009472).

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

References

1. Helfrich, K.R.; Battisti, T.M. Experiments on baroclinic vortex shedding from hydrothermal plumes. *J. Geophys. Res. Ocean.* **1991**, *96*, 12511–12518. [\[CrossRef\]](#)
2. Joint Analysis Group. *Review of R/V Brooks McCall Data to Examine Subsurface Oil*; NOAA: Silver Spring, MD, USA, 2010.
3. Operational Science Advisory Team. Summary report for sub-sea and sub-surface oil and dispersant detection: Sampling and monitoring. *December 2010*, 17.
4. Beegle-Krause, C.; Daae, R.L.; Skancke, J.; Brakstad, O.G.; Christos Stefanakos Valentine, D.L. Deepwater Wells and the Subsurface Dissolved Oxygen Minimum: A Tale of Two Sides of the Atlantic Ocean. In Proceedings of the Environment Canada's Arctic and Marine Oil Pollution Conference, Halifax, NS, Canada, 7–9 June 2016.
5. White, H.K.; Hsing, P.-Y.; Cho, W.; Shank, T.M.; Cordes, E.E.; Quattrini, A.M.; Nelson, R.K.; Camilli, R.; Demopoulos, A.W.; German, C.R. Impact of the Deepwater Horizon oil spill on a deep-water coral community in the Gulf of Mexico. *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 20303–20308. [\[CrossRef\]](#) [\[PubMed\]](#)
6. American Petroleum Institute. Industry Recommended Subsea Dispersant Monitoring Plan. *API Tech. Rep.* **2013**, 1152.
7. Landry, M.E.; Adams, E.; Bejarano, A.; Boufadel, M.; White, H.K. *The Use of Dispersants in Marine Oil Spill Response*; The National Academies Press: Washington, DC, USA, 2019. [\[CrossRef\]](#)
8. French-McCay, D.P.; Horn, M.; Li, Z.; Jayko, K.; Spaulding, M.L.; Crowley, D.; Mendelsohn, D. Modeling distribution, fate, and concentrations of Deepwater Horizon oil in subsurface waters of the Gulf of Mexico. In *Oil Spill Environmental Forensics Case Studies*; Elsevier: Amsterdam, The Netherlands, 2018; pp. 683–735.
9. United States. Coast Guard, and National Response Team (US). *On Scene Coordinator Report: Deepwater Horizon Oil Spill*; US Department of Homeland Security, US Coast Guard: Washington, DC, USA, 2011.
10. Ji, C.; Beegle-Krause, C.J.; Englehardt, J.D. Formation, Detection, and Modeling of Submerged Oil: A Review. *J. Mar. Sci. Eng.* **2020**, *8*, 642. [\[CrossRef\]](#)
11. Englehardt, J.D. *Final Version of SOSim (Subsurface Oil Simulator)*; Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC), Harte Research Institute, Texas A&M University: Corpus Christi, TX, USA, 2020. [\[CrossRef\]](#)
12. Ji, C.; Englehardt, J.D. *SOSim Case Study: Sampling Plans for Oil Response*; Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC), Harte Research Institute, Texas A&M University: Corpus Christi, TX, USA, 2020. [\[CrossRef\]](#)
13. Delmelle, E. Spatial sampling. In *The SAGE Handbook of Spatial Analysis*; SAGE Publications Ltd.: Newbury Park, CA, USA, 2009; Volume 183, p. 206.
14. Zhang, C. *Fundamentals of Environmental Sampling and Analysis*; John Wiley & Sons: Hoboken, NJ, USA, 2007.
15. Agency, U.E.P. *Guidance on Choosing a Sampling Design for Environmental Data Collection*; United States Environmental Protection Agency: Washington, DC, USA, 2002.
16. Hook, S.; Batley, G.; Holloway, M.; Ross, A.; Irving, P. *Oil Spill Monitoring Handbook*; Csiro Publishing: Collingwood, Australia; Clayton, Australia, 2016.

17. Delmelle, E.M.; Goovaerts, P. Second-phase sampling designs for non-stationary spatial variables. *Geoderma* **2009**, *153*, 205–216. [CrossRef]
18. Turk, P.; Borkowski, J.J. A review of adaptive cluster sampling: 1990–2003. *Environ. Ecol. Stat.* **2005**, *12*, 55–94. [CrossRef]
19. Thompson, S.K. Adaptive cluster sampling. *J. Am. Stat. Assoc.* **1990**, *85*, 1050–1059. [CrossRef]
20. Thompson, S.K. Stratified adaptive cluster sampling. *Biometrika* **1991**, *78*, 389–397. [CrossRef]
21. Salehi, M.; Seber, G.A. Two-stage adaptive cluster sampling. *Biometrics* **1997**, *53*, 959–970.
22. Cox, L.A. Adaptive spatial sampling of contaminated soil. *Risk Anal.* **1999**, *19*, 1059–1069. [CrossRef] [PubMed]
23. Hsiao, C.K.; Juang, K.-W.; Lee, D.-Y. Estimating the second-stage sample size and the most probable number of hot spots from a first-stage sample of heavy-metal contaminated soil. *Geoderma* **2000**, *95*, 73–88. [CrossRef]
24. Szatmári, G.; László, P.; Takács, K.; Szabó, J.; Bakacsi, Z.; Koós, S.; Pásztor, L. Optimization of second-phase sampling for multivariate soil mapping purposes: Case study from a wine region, Hungary. *Geoderma* **2019**, *352*, 373–384. [CrossRef]
25. Smith, D.R.; Vilella, R.F.; Lemarié, D.P. Application of adaptive cluster sampling to low-density populations of freshwater mussels. *Environ. Ecol. Stat.* **2003**, *10*, 7–15. [CrossRef]
26. Hanselman, D.H.; Quinn II, T.J.; Lunsford, C.; Heifetz, J.; Clausen, D. Applications in adaptive cluster sampling of Gulf of Alaska rockfish. *Fish. Bull.* **2003**, *101*, 501–513.
27. Payne, J.; Driskell, W. Offshore adaptive sampling strategies. US DOI, DWH Response & Restoration, Admin, DWH-AR0023786. 2015. Available online: <https://www.fws.gov/doiddata/dwh-ar-documents/946/DWH-AR0023786.pdf> (accessed on 1 December 2020).
28. Diercks, A.R.; Highsmith, R.C.; Asper, V.L.; Joung, D.; Zhou, Z.; Guo, L.; Shiller, A.M.; Joye, S.B.; Teske, A.P.; Guinasso, N. Characterization of subsurface polycyclic aromatic hydrocarbons at the Deepwater Horizon site. *Geophys. Res. Lett.* **2010**, *37*. [CrossRef]
29. Operational Science Advisory Team. *Summary Report for Sub-Sea and Sub-Surface Oil and Dispersant Detection: Sampling and Monitoring*; US Coast Guard, Unified Area Command: New Orleans, LA, USA, 2010.
30. Ryan, J.; Zhang, Y.; Thomas, H.; Rienecker, E.; Nelson, R.; Cummings, S. A high-resolution survey of a deep hydrocarbon plume in the Gulf of Mexico during the 2010 Macondo blowout. *Geophys. Monogr. Ser.* **2011**, *195*, 63–75.
31. Payne, J.; Driskell, W. Deepwater Horizon Oil Spill NRDA Offshore Adaptive Sampling Strategies and Field Observations, DWH-AR0023786. PEI Technical Report to the Trustees in Support of the PDARP; 2015. Available online: <https://www.fws.gov/doiddata/dwh-ar-documents/946/DWH-AR0023786.pdf> (accessed on 1 December 2020).
32. Diercks, A.-R.; Asper, V.; Highsmith, R.; Woolsey, M.; Lohrenz, S.; McLetchie, K.; Gossett, A.; Lowe, M.; Joung, D.; McKay, L. NIUST-Deepwater horizon oil spill response cruise. In Proceedings of the OCEANS 2010 MTS/IEEE SEATTLE, Seattle, WA, USA, 20–23 September 2010; pp. 1–7.
33. Ledwell, J.; He, R.; Xue, Z.; DiMarco, S.F.; Spencer, L.J.; Chapman, P. Dispersion of a tracer in the deep Gulf of Mexico. *J. Geophys. Res. Ocean.* **2016**, *121*, 1110–1132. [CrossRef]
34. Jacketti, M.; Ji, C.; Englehardt, J.D.; Beegle-Krause, C. Development of the SOSIM model for inferential tracking of subsurface oil. In Proceedings of the 42nd Arctic and Marine Oilspill Program-Technical Seminar on Environmental Contamination and Response, AMOP, Halifax, NS, Canada, 4–6 June 2019.
35. Englehardt, J.; Echavarria-Gregory, A.; Avellaneda, P. Development of a Predictive Bayesian Data-Derived Multi-Modal Gaussian Maximum Likelihood Model of Sunken Oil Mass. Draft Final Report. 2010. Available online: https://crrc.unh.edu/sites/crrc.unh.edu/files/final_report_univ_miami_sosim_development.pdf (accessed on 1 December 2020).
36. Echavarria-Gregory, M.A.; Englehardt, J.D. A predictive Bayesian data-derived multi-modal Gaussian model of sunken oil mass. *Environ. Model. Softw.* **2015**, *69*, 1–13. [CrossRef]
37. Ji, C.; Englehardt, J.; Beegle-Krause, C. SOSim: A Probabilistic Bayesian Model for Submerged Oil Tracking (Poster Presentation). In Proceedings of the Gulf of Mexico Oil Spill & Ecosystem Science Conference, Tampa, FL, USA, 3–6 February 2020.
38. Ji, C.; Englehardt, J.D.; Beegle-Krause, C. Bayesian Submerged Oil Tracking with SOSim: Inference from Field Reconnaissance Data and Fate-Transport Model Output. *Mar. Pollut. Bull.* **2020**. Unpublished work.

39. Beegle-Krause, C.J. *The Oil Spill Contingency and Response (OSCAR) Example Source Data for the Subsurface Oil Simulation (SoSIM) Model of the Deepwater Horizon Oil Spill, 2010-05-21 to 2010-06-17*; Gulf of Mexico Research Initiative Information and Data Cooperative (GRIIDC), Harte Research Institute, Texas A&M University: Corpus Christi, TX, USA, 2020. [CrossRef]
40. Reed, M.; Daling, P.S.; Brakstad, O.G.; Singaas, I.; Faksness, L.-G.; Hetland, B.; Ekrol, N. OSCAR2000: A multi-component 3-dimensional oil spill contingency and response model. In Proceedings of the Arctic and Marine Oilspill Program Technical Seminar, Environment Canada, Ottawa, ON, Canada, 14 June 2020; Volume 2, pp. 663–680.
41. Institute, A.P. Industry Guidelines on Requesting Regulatory Concurrence for Subsea Dispersant Use. In *API TECHNICAL REPORT 2017*; API Publishing Services: Washington, DC, USA, 2017; Volume 1152.
42. Wang, K.-P.; Huang, L.; Zhou, C.-G.; Pang, W. Particle swarm optimization for traveling salesman problem. In Proceedings of the 2003 International Conference on Machine Learning and Cybernetics (IEEE cat. no. 03ex693), Xi'an, China, 5 November 2003; pp. 1583–1585.
43. Pebesma, E.J. Multivariable geostatistics in S: The gstat package. *Comput. Geosci.* **2004**, *30*, 683–691. [CrossRef]
44. Pebesma, E. The Meuse Data Set: A Brief Tutorial for the Gstat R Package. 19 February 2013. Available online: <https://training.fws.gov/courses/references/tutorials/geospatial/CSP7304/documents/gstat.pdf> (accessed on 1 December 2020).
45. Skoien, M.J. Package ‘Intamapinteractive’ Documentation. 23 February 2018. Available online: <http://cran.uni-muenster.de/web/packages/intamapInteractive/intamapInteractive.pdf> (accessed on 1 December 2020).
46. Joint Analysis Group. *Deepwater Horizon Oil Spill. Review of the Subsurface Dispersed Oil and Oxygen Levels Associated with the Deepwater Horizon MC 252 Spill of National Significance*; NOAA Technical Report; NOS OR&R: Seattle, WA, USA, 2012.
47. Belogay, E.; Cabrelli, C.; Molter, U.; Shonkwiler, R. Calculating the Hausdorff distance between curves. *Inf. Process. Lett.* **1997**, *64*, 17–22. Available online: <https://repository.library.noaa.gov/view/noaa/390> (accessed on 1 December 2020). [CrossRef]
48. Valentine, D.L.; Fisher, G.B.; Bagby, S.C.; Nelson, R.K.; Reddy, C.M.; Sylva, S.P.; Woo, M.A. Fallout plume of submerged oil from Deepwater Horizon. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 15906–15911. [CrossRef] [PubMed]
49. Peirce, C.S. The numerical measure of the success of predictions. *Science* **1884**, *93*, 453–454. [CrossRef] [PubMed]
50. Jolliffe, I.T.; Stephenson, D.B. (Eds.) *Forecast Verification: A Practitioner’s Guide in Atmospheric Science*; John Wiley & Sons: Hoboken, NJ, USA, 2012.

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



© 2020 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 (<http://creativecommons.org/licenses/by/4.0/>).