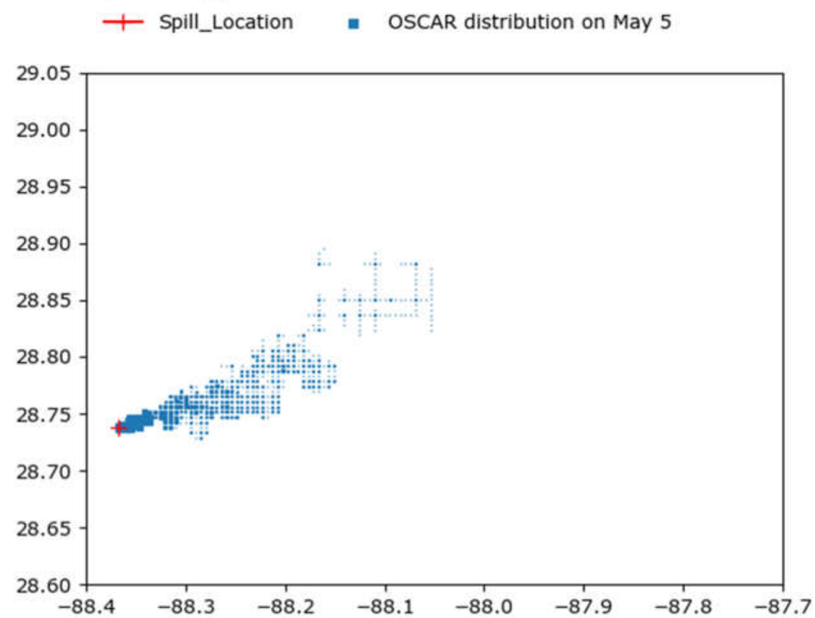
Supplement information

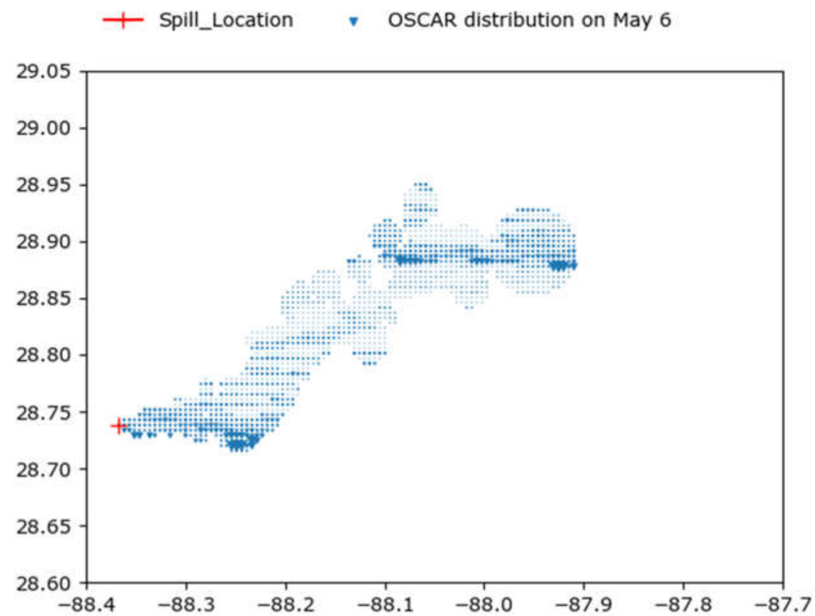
Appendix A: Sampling plan

1. OSCAR simulations on submerged oil

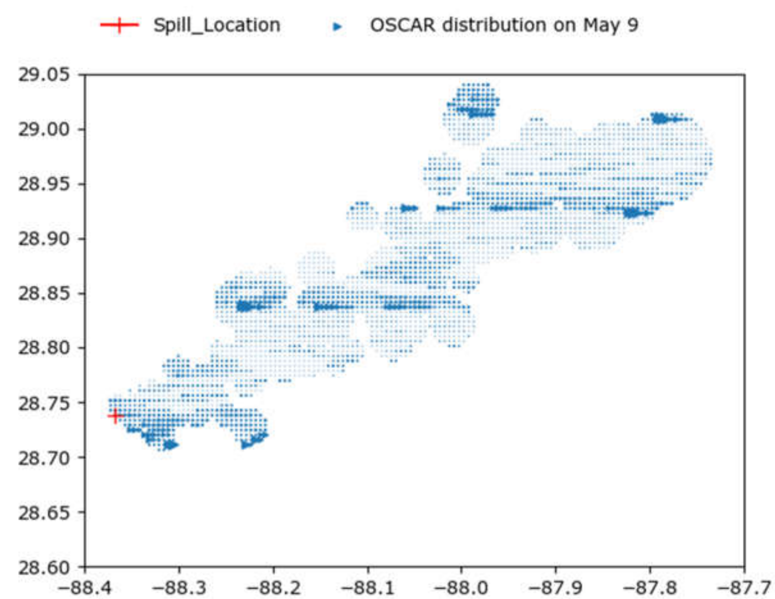
The simulation results of OSCAR on 5 May, 6 May and 9 May are plotted in Figure S1. The concentration of submerged oil on 5 May, 6 May, and 9 May predicted by OSCAR at the isopycnal layer ranges from 0.054 – 2620.21, 0.028-4820.55, 0.00030-3187.08 ppb respectively. The mean concentrations ± one standard deviation of submerged oil predicted by OSCAR on 5 May 5, 6 May and 9 May are 34.96 ± 146.94, 29.01 ± 173.31, 15.93 ± 121.13 ppb respectively. The simulated oil plume gradually disperses in the northeast direction. Considering the submerged oil is contained within an isopycnal layer, the average of all concentrations over the depths within the isopycnal layer (1027.62-1027.70 kg/m^3) are taken to obtain a concentration profile along the isopycnal.



(a)



(b)



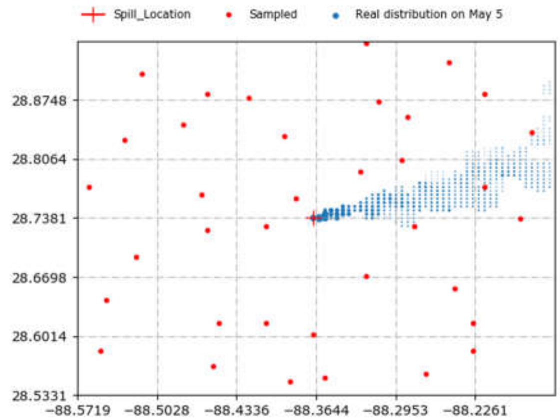
(c)

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).

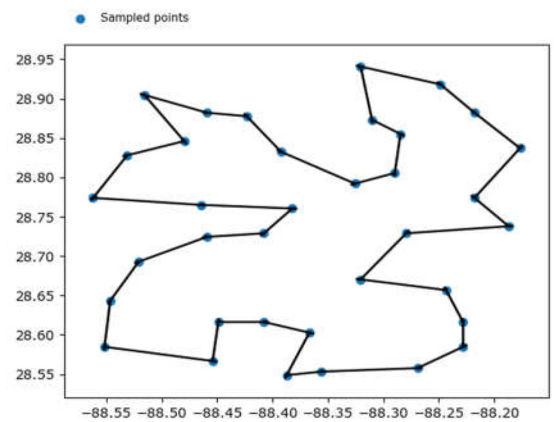
2. First phase sampling

2.1 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 using the systematic random sampling plan are plotted as red dots in Figure S2 (a) and the designed sampling path is plotted in Figure S2 (b). The distance of the path is 204 km.



(**a**)

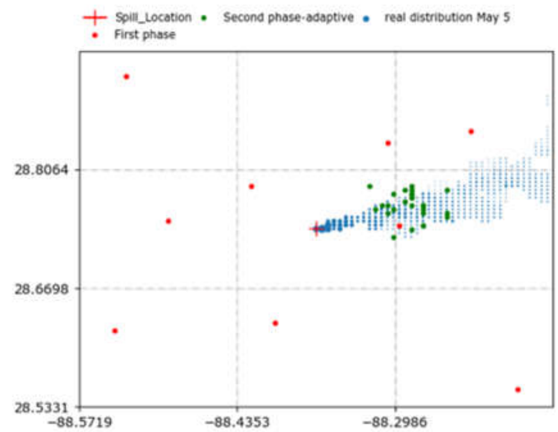


(**b**)

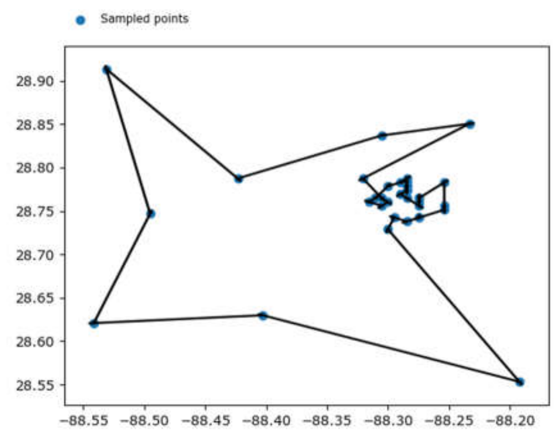
Figure S2. Sampled data using the systematic random sampling plan (a) and the designed path (b).

2.2 Adaptive systematic sampling plan

The sampled points using the adaptive systematic sampling plan are plotted as red dots in Figure S3. Among the sampled locations marked by red dots, oil is detected at only one location at −88.29° W, 28.74° N. Further, 26 data points are sampled around this location. These new samples are marked in green in Figure S3. The distance of the path is 151 km.



(**a**)

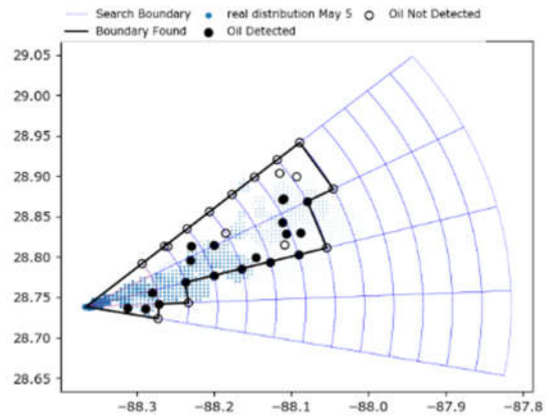


(**b**)

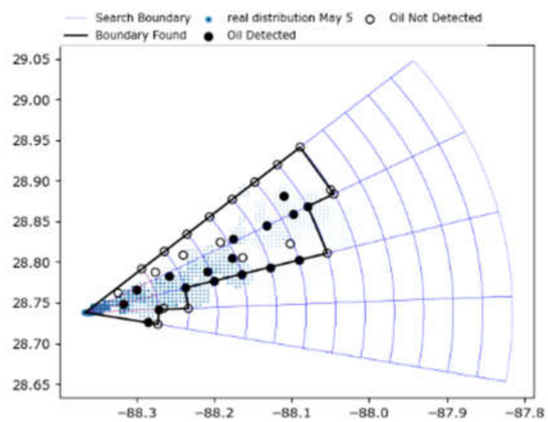
Figure S3. Sampled data using the adaptive sampling plan (a) and the designed path (b).

2.3 Modified station sampling plan

The modified station sampling plans are described here. The designed path and sampled data using the modified sampling plan combined with random sampling and systematic random sampling method are shown in Figure S4. In Figure S4, the detected extent area is the solid black line, and the sampled locations are the dark solid and hollow points. The solid points are locations where oil is detected and the hollow points are locations where oil is not detected. The vessel moves to these points to collect samples. The path is designed by solving the traveling salesman problem. The path distance for the modified random sampling is 199 km, and for the modified systematic sampling is 87 km.



(**a**)



(**b**)

Figure S4. Modified sampling plan combined with: (a) random sampling method (distance 199 km); (b) systematic random sampling method (distance 87 km).

2.4 Zig-zag paths

2.4.1 Naive zig-zag pattern

The designed path and sampled data using the naive zig-zag pattern are in Figure S5. The first sampling point is placed 9 km away from the spill site along the southern boundary. The sampling path then connects the first sampling point with the leftmost point on the northern boundary which is 11 km away from the spill site along the northern boundary. The path then connects the point on the northern boundary to the next point on the southern boundary which is 13 km away from the spill location and then connects it to the point on the northern boundary which is 15 km away. The points on the search boundary are 5 km apart and the path connects all these points until out of the region of interest. The path distance is 355.88 km.

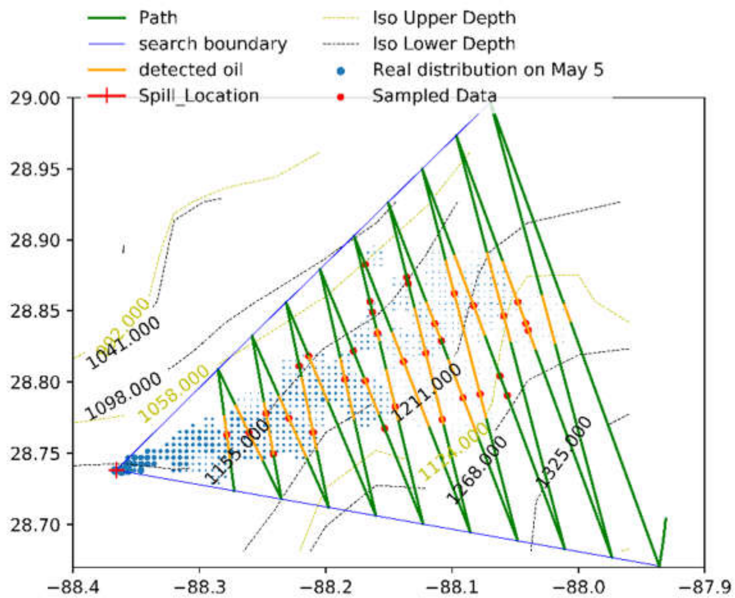
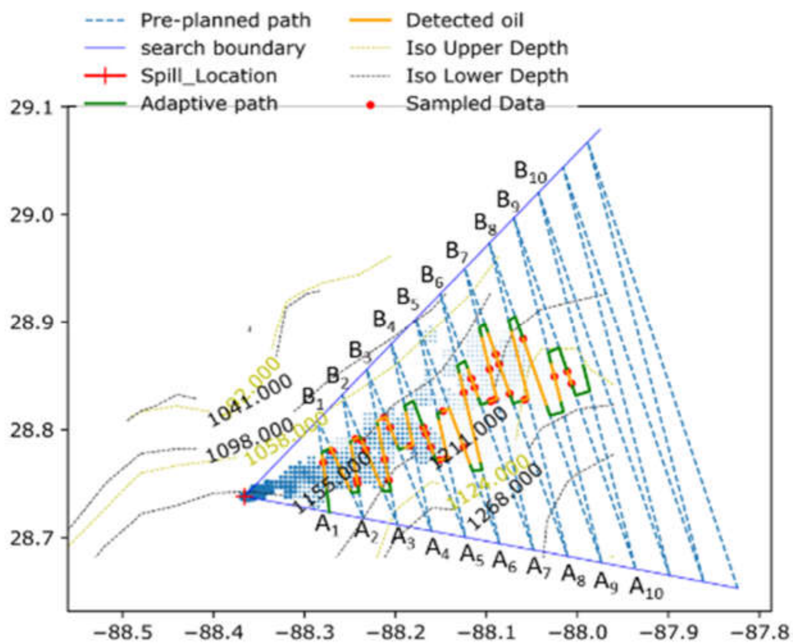


Figure S5. The naive zig-zag pattern (the green and orange lines are the 335.88 km path).

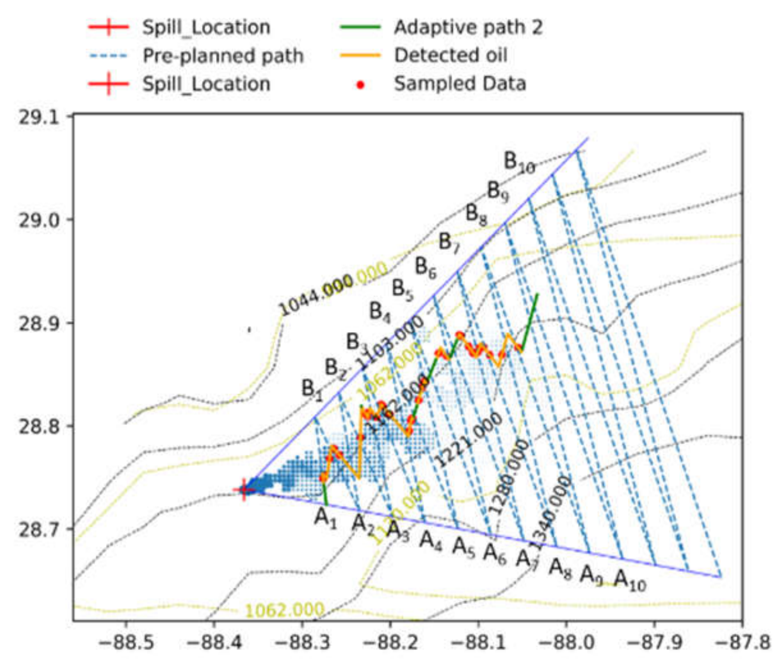
2.4.2 Adaptive zig-zag pattern

The adaptive zig-zag patterns are shown in Figure S6. In Figure S6 (a), the points A1-A10, B1-B10 are placed every 1 mile along the southern and northern boundaries and A1, B1 are 9 km away from the spill location. The pre-planned path is defined by connecting A1-B1-A2-B2-… until reaching A10. The device first moves along the pre-planned path (e.g. A1-B1) until no oil is detected 0.2 km from the last coordinate where oil is detected along that pre-planned path; and at that point the sampling path changes direction and moves along the respective radial line connecting the spill location to the current location until touching the next 'pre-planned path'. This whole process is repeated until no oil is detected for ~2 km along the last searched 'pre-planned path'. The 2 km distance is chosen empirically, and this distance should be neither too small nor too large. When the distance is small, the gaps inside the oil are ignored and some areas with submerged oil area will be missed. When the distance is large, the path will be similar as the 'naive path' sampling method. In Figure S6 (b), the device moves along the first 'pre-planned path' A1-B1 until no oil is detected along the path; then the sampling path changes direction pointing to the location A2 on the southern boundary until no oil is found again within 0.2 km of the previously detected oil location along the path. Then the device changes to the direction pointing to the next intersection point B2 of the 'pre-planned path' with the northern boundary. When no oil is detected within 0.2 km of the previously detected oil on the path, the device then turns to the direction pointing the next intersection point A3 of the 'pre-planned path' on the southern boundary. This process is repeated until no oil is detected along the path for 2 km. The difference between the two patterns is that the device changes direction pointing to the next boundary line in Figure S6b, and pointing to the next pre-planned path along radials in Figure S6a. The distance of the adaptive zig-zag pattern I is 66.61 km, and the distance of adaptive zig-zag pattern II is 27.90 km.

To get these data for SOSim input, we uniformly selected the samples within each straight track line segment along the sampling path with oil, which are the 25 orange straight line segments in Figure S6 (a) and the 22 orange straight line segments in Figure S6 (b). In order to get the same number of data points as the other methods, we additionally random select 11 (for Figure S6 (a)) and 14 (for Figure S6 (b)) orange line segments respectively and uniformly select one point on each line segment.



(a)



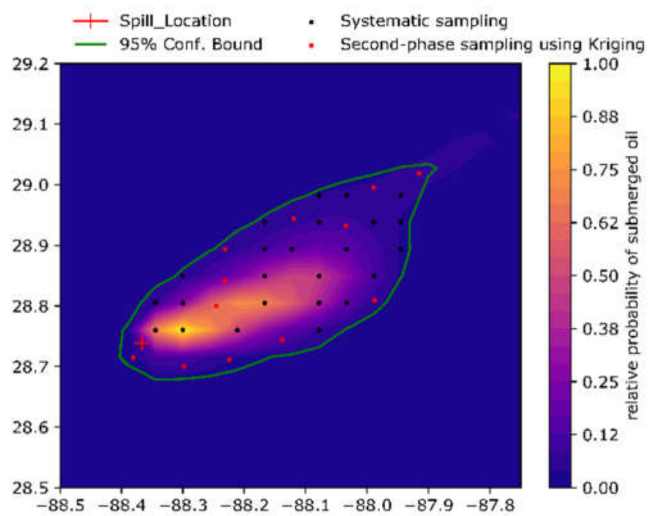
(b)

Figure 6. Two adaptive zig-zag paths (a: 66.61 km zig-zag pattern I path; b: 27.90 km zig-zag pattern II path).

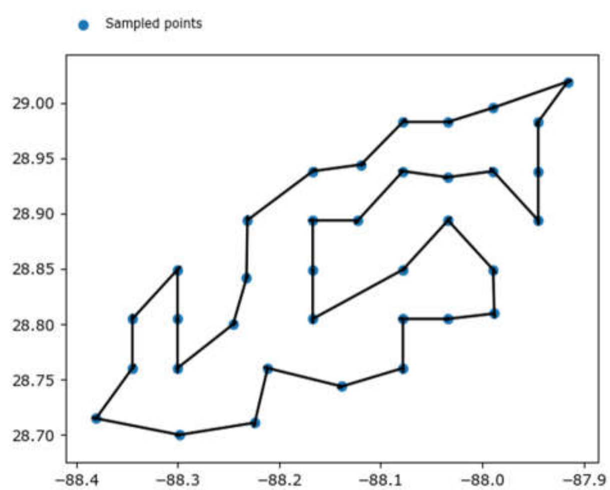
3. Second phase sampling

3.1 Systematic random combined with the kriging-based sampling

The systematic random combined with the kriging-based sampling method is applied to sample within the region predicted by SOSim. In Figure S7(a), the filled 2-D contour plot represents the SOSim one-day forecast of relative submerged oil concentrations on 6 May. The region inside the green circle is the SOSim predicted region of submerged oil. The black points represent the sampling locations according to systematic sampling and the red points represent the locations by the kriging-based sampling. The designed path is in Figure S7(b).



(a)



(b)

Figure S7. a) Sampled data using the systematic random sampling combined with the kriging-based sampling method and (b) the 184 km designed path.

3.2 Naive sampling

The designed path using naive sampling for 6 May, and the resulting pattern of sample collection, are shown in Figure S8. In Figure S8, the black lines are the zig-zag path and the segments with yellow color are where submerged oil is detected. To input these data to SOSim, 1-2 points are randomly collected at each yellow segment, which are represented by the red points.

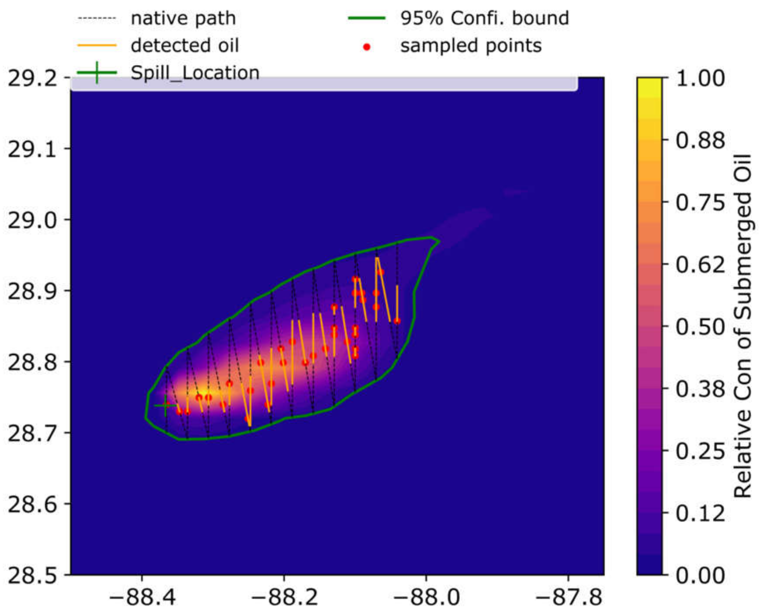
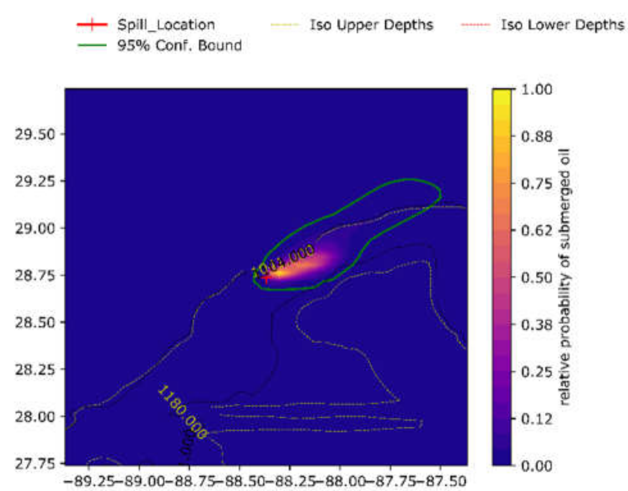


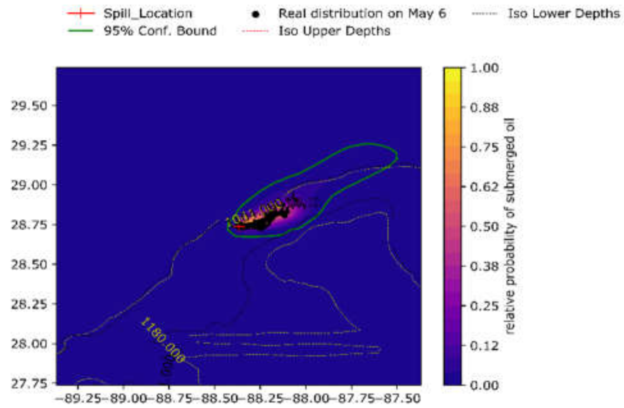
Figure S8. The 442 km naive zig zag path within the SOSim prediction for 6 May.

4. SOSim predictions for May 6 and 9

The predictions of SOSim using the naive path method in Figure S5 for 6 May and 9 May are plotted in Figure S9 and S10.

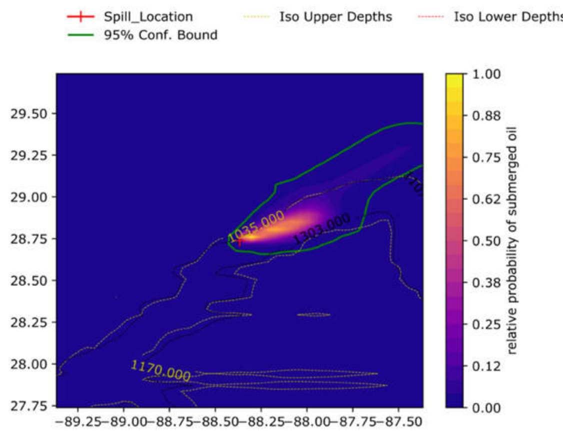


(**a**)

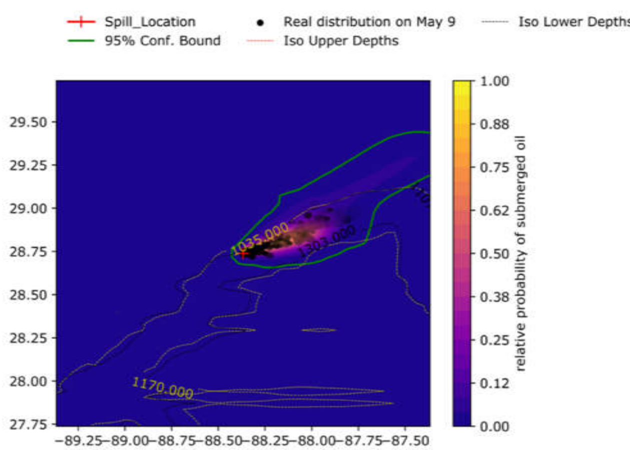


(**b**)

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).



(a)



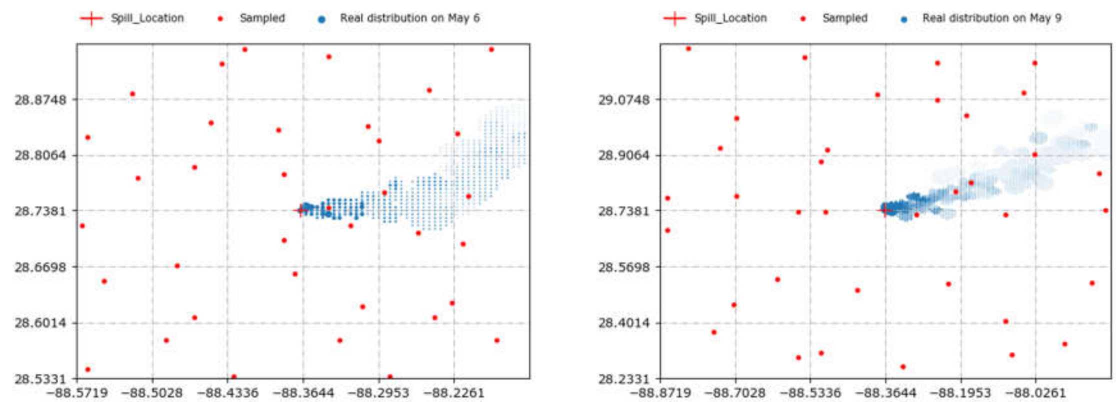
(b)

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).

5. Sampling plans without using the SOSim prediction

5.1 Data

The sampling results for 6 May and 9 May using the systematic random sampling method are plotted in Figure S11.

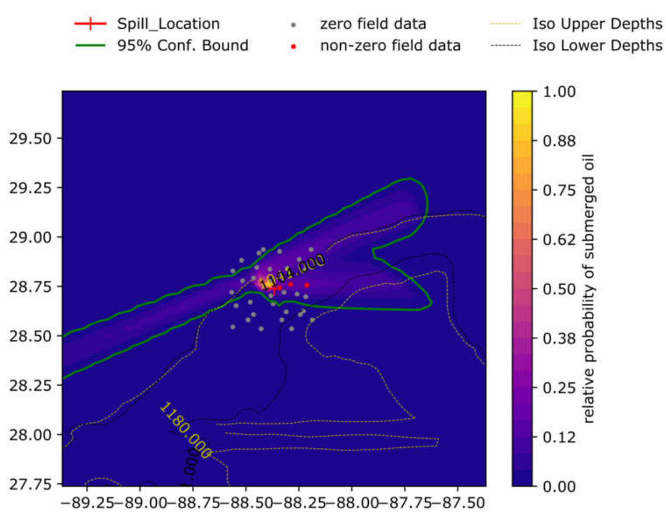


(**a**) (**b**)

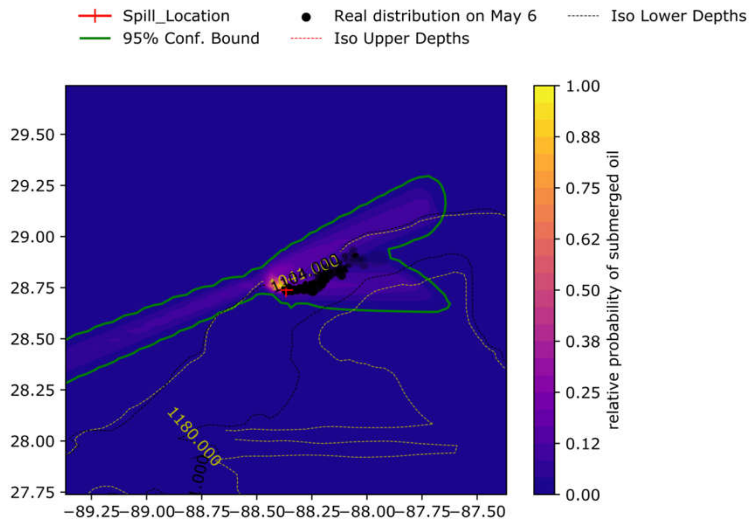
Figure S11. Sampled data on May 6 (a) and May 9 (b) using systematic random sampling method.

5.2 SOSim predictions

SOSim predictions for 6 May and 9 May using data in Figure S11 (a) and (b) are plotted in Figures S12 and S13.

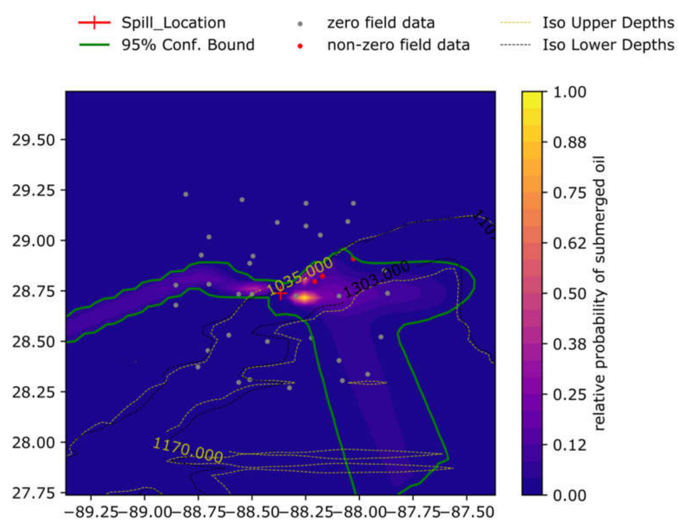


(a)

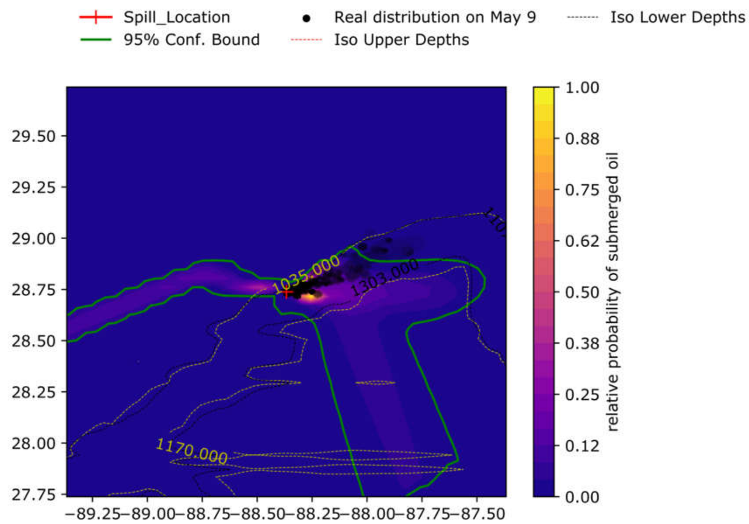


(b)

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).



(a)



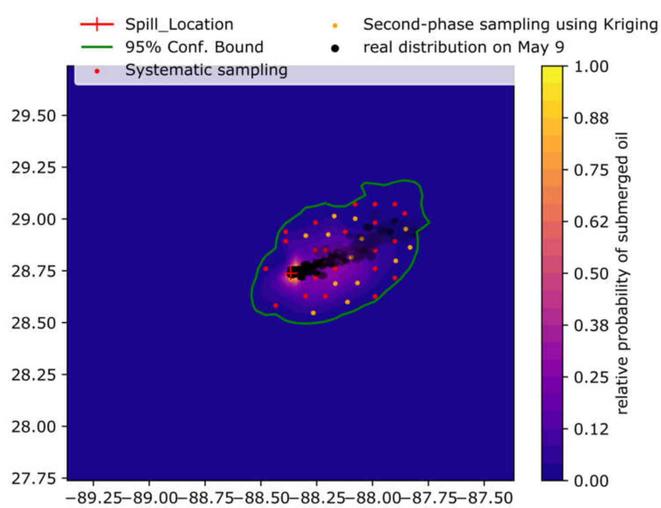
(b)

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).

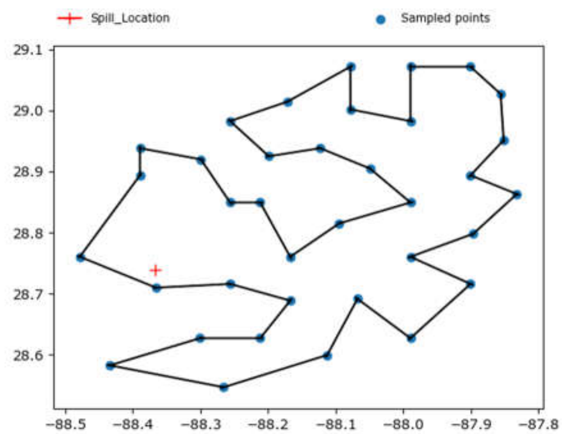
6. Sampling plans with SOSim

6.1 The systematic random combined with the kriging-based method

Using the systematic random combined with the kriging-based method to sample from SOSim predictions, these sampled data are plotted in Figure S14.



(a)



(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.

6.2 Zig-zag sampling

The naive zig-zag path is applied to sample from the SOSim prediction for 9 May. Data are discretized to input to SOSim. The path and discretized data are plotted in Figure S15.

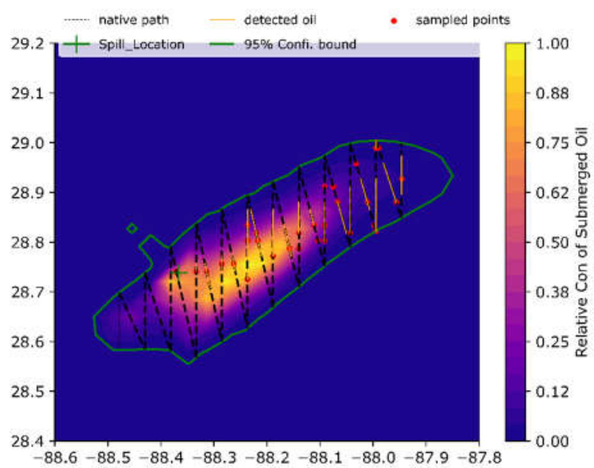


Figure S15. The sampled data using the naive zig-zag path method based on SOSim prediction for 9 May.

Appendix B: Kriging

Kriging is a method that accounts for the spatial dependence of data. The spatially dependent data are interpreted as the result of a random process in the kriging model and kriging is an interpolation method based on the Gaussian process governed by the covariance. Here, the oil concentration $C(x)$ at location x is interpreted as a Gaussian process. Kriging assumes that the correlation between two random variables $C(x_i)$ and $C(x_j)$ only depends on the spatial distance between them, which is $x_i - x_j$. The variance γ of the Gaussian process is defined as the variance of $C(x)$ at different locations x and is related to the semi-variogram γ_s by

$$\gamma(x_1, x_2) = Var(C(x_1) - C(x_2) = \gamma_s(x_1 - x_2),$$

where the semi-variogram γ_s is defined by

$$\gamma_s(h) = \frac{1}{2V} \iint_V [C(x) - C(x+h)]^2 \, dh,$$

in which V is the area of interest, and h is the distance between point $(x + h)$ and point x .

The sample semi-variogram $\hat{\gamma}_s$ is calculated using these sample data $\{c(x_i), i = 1, \cdots, M\}$ by

$$\hat{\gamma}_s = \frac{1}{2N} \sum_{x_i - x_j \approx h} \left(c(x_i) - c(x_j) \right)^2.$$

The semi-variogram γ_s is estimated by fitting an exponential semi-variogram model

$$\gamma_s(h) \approx (s - n)\left(1 - e^{-h/ra}\right),$$

in which s, n, r, a are the parameters to be fitted by using the sampled semi-variogram. The kriging method then makes an inference on $C(x_0)$ at unobserved location x_0 based on the sampled data $\{c(x_i), i = 1, \cdots, M\}$ by taking $C(x_0) = \sum_{i=1}^{M} w_i(x) c(x_i)$.

Here the weights (w_i), which sum to 1 to make the kriging model unbiased, are determined by minimizing the variance

$$Var(\epsilon(x_0)) = Var(\hat{C}(x_0) - C(x_0)).$$

The variance is usually called kriging variance. The semi-variogram model is used to compute the kriging variance.