

Remotely-Sensed Oil and Shoreline Interaction ModelingZachary Nixon ¹, Jacqueline Michel ¹, and Scott Zengel ²¹ Research Planning, Inc. (RPI), 1121 Park St., Columbia, SC 29201 USA² Research Planning, Inc. (RPI), 247 E. 7th Avenue, Suite 200, Tallahassee, FL 32303**ABSTRACT No. 2017-233**

The broad adoption of remotely sensed data and derivative products from satellite and aerial platforms available to describe the distribution of spilled oil on the water surface was an important factor during *Deepwater Horizon* (DWH) oil spill both for tactical response and damage assessment. The availability and utility of these data in describing on-water oil distribution provide strong temptation to make estimates about on-shoreline oil distribution. The mechanisms by which floating oil interact with the shoreline, however, are extremely complex, heterogeneous at fine spatial scales, and generally not well described or quantified beyond broad conceptual or spill-specific empirical models. In short, oil on water does not necessarily lead to oil on adjacent shorelines. We combine data derived from NOAA's National Environmental Satellite, Data, and Information Service (NESDIS) using a variety of satellite platforms of opportunity describing the remotely-sensed, daily composite anomaly polygons representing oil on water over multiple months with ground observations made in the field, collocated in time and space extracted from a newly compiled database of ground survey data (SCAT, NRDA and others) from the northwestern Gulf of Mexico. Because this new compiled dataset is very large (100,000s of observations) and spans a wide range of habitats, geography, and time, it is particularly suitable for inference and predictive modeling. We use these combined datasets to make inference about the relative influence on shoreline oiling probability and loading of distance from on-water oil observation via multiple distance metrics, shoreline morphology,

water levels and ranges, wind direction and speed, wave energy, shoreline aspect and geometry. We also construct predictive models using machine-learning modeling methods to make predictions about shoreline oiling probability given observed distributions of on-water oil. The importance of this work is three part: firstly, the relationships between these parameters can assist hind-cast modeling of shoreline oiling probability for the *Deepwater Horizon* oil spill. Secondly, these data and models can permit similar modeling for future spills. Lastly, we propose that this dataset serve as a nucleus that can be expanded using data from subsequent or future spills to allow iteratively improvements in shoreline oil probability modeling using remotely sensed data, as well as an improved understanding of oil-shoreline interactions more generally.

INTRODUCTION

During the Mississippi Canyon 252 (MC252) /*Deepwater Horizon* (DWH) spill, apportioning Shoreline Cleanup Assessment Technique (SCAT) and other shoreline survey effort in time and space was a complicated and difficult task. During DWH, as after most spills, these field surveys captured a vast amount of information on shoreline oiling distribution and degree but left some areas un-surveyed or under-surveyed and were irregular in timing as conditions changed over the spill. By contrast, remote sensing approaches, primarily Synthetic Aperture Radar (SAR), can detect oil on the water surface with synoptic coverage and high temporal frequency. Data from satellite-based SAR instruments were broadly and routinely collected and employed operationally during DWH (Leifer *et al.*; Streett, 2011). Previous analyses of these data successfully demonstrated that SAR imagery could be used to detect oil on the water surface in proximity to the shoreline (Garcia-Pineda *et al.*, 2013a; 2013b). However, the physical mechanisms through which floating oil interacts with the shoreline at spatial scales finer than the

resolution of remotely sensed imagery (less than ~100 meters), including eddy diffusivity and water exchange with the surf zone, are very complex (e.g. Battjes, 1988; Grant *et al.* 2005; Reniers *et al.* 2009; Restrepo *et al.*, 2014). Though ongoing research efforts may facilitate better and faster physical modeling of these processes, our understanding of their relationship to oil transport is limited to broad conceptual models (see Etkin *et al.*, 2007 for a review). In short, while intuitively related, rigorous examinations of the relationship between nearshore oil-on-water proximity, timing, and duration to actual field-observed shoreline oiling data is rare. Recent advances in machine-learning and statistical modeling methods make it possible to bring more powerful non-physical modeling methods to bear quickly and efficiently where data are plentiful. Here, we attempt to leverage the very large datasets describing both synoptic estimates of on-water oil distribution derived from satellite-based SAR instruments and shoreline survey data collected during DWH for this purpose. We do this by constructing a machine learning-based model to estimate shoreline oiling probability via spatial analysis of satellite-derived surface oil analysis products, climatic summary data, and shoreline habitat type and geometry data. We use these data, and resulting models, to answer several questions about the relationship between the probability of shoreline oiling given distance to a remote sensing-derived potential oiling footprint, as well as water levels, and wind speeds and other potential predictor variables, given similar nearby on-water oil observations. We also investigate the amount of shoreline predicted to have been oiled but not observed using these models during DWH and compare with previous estimates.

METHODS

Remote-Sensing Derived On-water Footprint Data

We obtained daily composite surface oil analysis products generated from satellite-derived SAR imagery data collected during the Mississippi Canyon 252 (MC252) /*Deepwater Horizon* oil spill by NOAA's National Environmental Satellite, Data, and Information Service (NESDIS) using several satellite platforms of opportunity (NOAA, 2013). These data were subsequently re-analyzed by Garcia-Pineda *et al.* (2013a; 2013b) via textural classifier neural network algorithms (TCNNA) to delineate potential oiling footprints from imagery using edge-detection filters, descriptors of texture, collection information, and other environmental data. We used two different but similar datasets, compiled from two different batches of SAR imagery for this purpose termed "potential oiling footprints" and "nearshore potential oiling footprints" (ERMA, 2015). We merged the resulting polygons from each dataset for each given date of satellite observation. These resulting merged daily composite anomaly data consist of vector polygons, each describing composite potential on-water oil footprints for 76 individual dates between April 23, 2010 and August 10, 2010 across the northwestern Gulf of Mexico.

Shoreline Oil Stranding Training Data

The model was trained using survey-derived, shoreline surface oiling presence-absence data from SCAT and Rapid Assessment (RA) surface oiling data collected during the spill (Nixon *et al.*, 2016; Michel *et al.*, 2013). We obtained all linear surface oiling zones that were collected over the duration of the spill from the data sets compiled by Nixon *et al.* (2016) and extracted only those within 10 km of any given cumulative anomaly polygon. For each individual date, we then extracted any candidate shoreline observations made within the 15 days after the date of the remote sensing observation. We then extracted only survey zones where no oil was observed

(NOO), or where oiling was observed for the first time at that location. We excluded all survey zones where oiling had already been observed because it cannot be determined if oiling observations at any location subsequent to initial oiling were due to re-observation of the same stranded oil, or due to new stranding. Any candidate shoreline oil observation that met these temporal and spatial matching criteria for multiple preceding potential oiling footprint polygons was only matched with the closest potential oiling footprint polygon to its location. Thus, any shoreline observations of oil were only matched with a single potential oiling footprint polygon in time and space, but observations of no oil could be matched with multiple potential oiling footprint polygons given the temporal and spatial matching criteria. Because these individual observations consisted of linear segments of shoreline of highly variable length, we converted these linear observations to a raster using a 25-meter (m) by 25-meter cell size, generated points at the centroids of those raster cells, and considered each point as a unique training sample observation. Individual linear observations shorter than 25 meters in length were assigned a unique training sample location at the centroid of that line segment. This resulted in 896,461 individual observations at 54,656 unique locations.

Shoreline Oil Stranding Predictor Variables

First, a common raster grid was created for the study area using a 25-meter (m) by 25-meter cell size, and a land-water raster and accompanying shoreline raster were then created using this standard grid. For this purpose, we used the data described by Nixon *et al.* (2016) to represent the land-water interface location. Using this same common grid, we then generated a series of raster grids representing variables thought to predict shoreline oiling probability. All spatially explicit predictor variables were prepared and processed in ArcGIS 10.4.1 (ESRI, 2016).

For each date with a known daily on-water oil footprint polygon within 10 km of the shoreline, we computed the Euclidean distance and direction of all observation sample locations to that polygon. We also computed the overwater distance using a cost raster wherein all open water cells were assigned a unit traversal cost of 1 and all non-water cells were assigned an infinite traversal cost. For each date with a known daily on-water oil footprint polygon within 10 km of the shoreline, we computed a set of indices of the relative quantity of on-water oiling within neighborhoods at different spatial scales. These indices were calculated, for each cell, as the sum of all cells within a given radius in a coarse grid with a 250-meter (m) by 250-meter cell size wherein cells were coded as 1 for on-water oiling observation and 0 for no observation. These indices were calculated using multiple radii: 1 km, 3 km, and 10 km. We then computed shoreline aspect for all shoreline cells in this 25-m raster by converting land-water to raster with land having a value of 1, water having a value of -1, and then computing aspect of this raster. We also computed an index of the relative angle (ra) between the direction of the closest on-water oil footprint polygon and shoreline aspect as:

$$ra = \frac{\cos(a - n) + 1}{2}$$

where a is the Euclidean direction to the closest on-water oil footprint polygon and n is the shoreline aspect.

Indices of shoreline convexity/concavity may influence the wave and current energy incident upon a shoreline and, thus, may affect the deposition and persistence of stranded oil. A continuous index of concavity/convexity was calculated for all observation sample locations. This index was calculated, for each cell, as the arithmetic mean of all cells within a given radius in a land/water raster grid wherein cells were coded as 0 for water and 1 for land. This yielded a unitless index ranging from near zero (extremely convex) to near one (extremely concave). This

index was calculated using multiple radii: 100 m, 500 m, 1 km, and 2.5 km. We assigned a simplified habitat and geomorphic setting classification to all observation sample location using the data and methods described in Nixon *et al.* (2016). Each location was classified as beach, wetland, or other. We similarly assigned a geomorphic setting to all shoreline locations as either mainland, barrier island, or delta. These variables were converted to a set of binary presence/absence variables, one for each habitat or geomorphic class, for the purposes of the modeling procedure described below.

We generated estimates of shoreline fetch from a specific direction for the wider study area from eight cardinal directions for observation sample locations. Fetch was estimated as a raster operation in ArcGIS according to the modified effective fetch calculation methodology to a maximum fetch of 50 km in any given direction via code for ArcGIS made available by Rohweder, *et al.* (2008). We then computed mean and maximum fetch in meters for all observation sample locations. Similarly, we generated two different Relative Exposure Index (REI) values calculated using methods modified from Keddy (1982) to summarize relative differences in wind-wave exposure across the area of interest. The first, or mean based, index was based on fetch, average wind velocities, and proportion of all winds in each evaluated direction. The second, or exceedance-based, index was based on fetch and the proportion of all winds that exceed a specific velocity in each evaluated direction. We used the 80th percentile of wind speeds at each location as a cutoff for calculating an exceedance-based REI. We collected wind data as 6-minute averages from Climactic Summaries for all of 2010 through 2012 from four NOAA National Data Buoy Center (NDBC) stations (NOAA, 2016c) across the study area: Apalachicola Bay, Florida, Ft. Morgan, Alabama, Grand Isle, Louisiana, and Lake Calcasieu, Louisiana. Missing wind direction data for the Grand Isle station were imputed with hourly data

from the NOAA National Climate Data Center (NCDC) Boothville station (NOAA, 2016b) for a limited number of small time periods. We then compiled average wind speed and frequency of occurrence for each direction, as well as the 80th percentile wind speed exceedance value and frequency of occurrence above this value for each station over this period. We computed REI values were for each observation sample locations using the formulae presented in Keddy (1982).

To investigate the effect of wind speed and direction relative to shoreline orientation on probability of shoreline oiling we calculated a set of variables related to winds on the day of each on-water satellite observation. Using the compiled wind data described above, we computed average and maximum of wind speed in meters per second (m/s) and percentages of all observations were computed in 8 directional bins each centered about cardinal directional headings representing 45° intervals for each day. We then computed, for all observation sample locations, a relative directional wind-loading index at time of potential oil stranding (i_d) for each shoreline cell for each available daily on-water oil polygon (d) as:

$$i_d = \sum_{a=0^{\circ}}^{315^{\circ}} \left(\frac{\cos(a - n) + 1}{2} * p_{ad} \right)$$

where d is a given date of on-water oil polygon, a is one of eight cardinal wind direction radials (0° – 315°), n is the shoreline aspect, and p_{ad} is the percentage of time the wind blew from within a 45° cone centered about angle a on day d . We used the nearest meteorological station to each observation sample location.

For each date with a compiled daily composite anomaly polygon within 10 km of the shoreline, we computed the average, minimum, and maximum water level above MLW for the four selected meteorological stations using NOAA CO-OPS data (NOAA, 2016a). We then computed

for all observation sample locations a normalized index of relative water level by scaling all elevations by the diurnal tidal range at that station, where 0 is equivalent to MLW and 1 is equivalent to MHW. We used the nearest meteorological station to each observation sample location.

Model Parameterization and Evaluation

We modeled of the probability of shoreline oiling as predicted by the variables described above using boosted regression trees (BRT). BRT models attempt to classify observations, here as oiled or unoled, by repeated recursive binary partitioning of explanatory variable values. The trees are applied sequentially to classify incorrectly classified observations from previous tree models, and the final output is equivalent to a weighted ensemble of hundreds or thousands of tree models (Friedman *et al.*, 2000; Friedman, 2001). See Elith *et al.* (2008) for a review of these models and their applications. All model fitting and evaluation was conducted using the R statistical computing language (R Development Core Team. 2016). The model was constructed using the training data set using the *xgboost* package (Chen and Guestrin, 2016) as implemented for the R language.

The observation sample described above was randomly separated into a 60% training sample, a 20% test sample and a 20% sample for calibration of probabilities. We parameterized the model with a learning rate of 0.05, and a maximum tree complexity of 6, and set as our learning metric the minimization of the negative log-likelihood. A 10-fold cross validation was used to determine the optimal number of iterations for the final model. These folds were not constructed randomly, but instead assembled using a variation of h-block cross validation (Burman *et al.*, 1994) wherein each of the 76 individual dates of a composite remote sensing observation was assigned randomly to one of 10 folds. Using this approach, the effect of any spatial

autocorrelation present in the training data would be lessened. All other parameters were set to default software settings. We used a final number of 84 iterations selected from this cross-validation.

We evaluated model performance computing the Area Under the Curve (AUC) statistic of Receiver Operating Characteristic (ROC) curves (Sing *et al.*, 2005) for both the training data and the independent test data. We evaluated the importance and effect of individual predictors by generating measures of relative importance of all predictors and individual conditional expectation plots (Goldstein *et al.*, 2015). Boosted classification tree algorithms often perform better as classifiers than as calibrated estimators of the true class probability (Mease *et al.* 2007; Flach and Matsubara 2008). As such, we calibrated the BRT model output to yield class membership probabilities using the scaling method of Platt (1999) as suggested in Niculescu-Mizil and Caruana (2005). We used the 20% calibration sample described above for this purpose and evaluated calibration via calibration plots and calculated model Brier scores. An initial model was fit using only the 60% training data to estimate model performance and generate the Platt scaling calibration from independent data sets. The final model was fit using all sample data, but using the parameters and Platt scaling calibration from the earlier training run.

RESULTS AND DISCUSSION

The performance of the model was very good, with nearly identical AUC values for both training and testing data (Figure 1). There was no significant decline in AUC evaluated using the independent test and both AUC values were above 0.92. The 10-fold h-block cross validation AUC values were noticeably lower (though still implying reasonable model performance) at 0.838. These estimates were made using samples from remote sensing observation dates that were not used for model construction, so this estimate of model performance is likely much more

reflective of the accuracy and specificity of the model when applied to dates and locations not used to train the model. Calibration plots indicated that the model was initially reasonably calibrated, but that Platt scaling improved overall calibration of predicted shoreline oiling probability. Relative contribution of predictor variables as model gain and conditional expectation plots for the top 12 predictors, which depicts the effect of that variable with other variables integrated out, were plotted and inspected (Figure 2).

Of the two distance metrics, overwater distance is a significantly better predictor than simple Euclidean distance in terms of predicting shoreline oiling probability, and is in fact the best predictor in the model contributing 15% of the total explanatory power as evaluated by gain. There is a clear declining relationship between overwater distance and probability of shoreline oiling, with noticeable drop-offs in oiling probability at 1-2 km and 3-4 km overwater from oil footprint.

There appears to be only a weak relationship wind direction relative to shoreline aspect (2% of model power) but the relationship increases monotonically, implying increased probability of shoreline oiling when winds are on-shore. Maximum wind speed also appears to control shoreline oiling, with a generally declining probability of shoreline oiling with increasing wind speeds, except for a noticeable increase between 12-14 m·s⁻¹, or roughly the average wind speeds of the passage of winter low pressure storm systems in the region. Water levels are also correlated with shoreline oiling. Mean, minimum, and maximum water level elevation, scaled to local diurnal tidal range, are all in the top 12 predictors and together contribute 15% of the explanatory power of the model. To some extent, all of these metrics display a relationship indicating increasing probability of oiling with increasing water levels.

Other predictors that were significantly correlated with shoreline oiling probability include classification of an observation as a beach habitat, rather than wetland or other shoreline type. Presence of sand beach was the second most significant predictor variable contributing 14% of the total explanatory power. As evaluated with conditional expectation plots, beaches are 30% more likely have been observed as oiled than other shoreline types, given similar other characteristics and on-water oil distributions. We speculate that this is due to both the fact that beaches are generally more physically energetic environments where transportation of on-water oil is faster, as well as the fact that low levels of oil is more difficult to detect in wetlands. Similarly, both exposure indices were related to shoreline oiling probability with higher wave exposures generally related to slowly decreasing oiling probability except at very low exposures such as the interior of tidal creeks where oil was less likely to have penetrated.

The sample we used represents 1,298 km of shoreline. Because we know the length of shoreline each individual observation represents, we used the final model results to estimate the total length of shoreline that was not observed as oiled in the datasets compiled by Nixon *et al.* (2016) that may have been oiled. We extracted all sample observation locations where no oil ever was observed (representing 873 km), and predicted the probability of shoreline oiling on each date when an on-water polygon was within 10 km. We then computed the conditional probability of any oiling at each location, given multiple potential oiling events (Figure 3). We finally conducted a Monte Carlo analysis by drawing 10,000 random samples from these conditional probabilities, summing the representative lengths of predicted oiled shoreline at each draw, and computing the empirical 95% confidence interval of predicted oiled length. These results imply that 176 km (95% 173-178 km) of surveyed shoreline, where no oil was observed during field surveys, may have been oiled at some point. This represents an increase of 8% over the 2,113

km of oiled shoreline reported by Nixon *et al.* (2016) compiled from ground survey data alone. The bulk of this estimated length (82%) is along wetland shorelines in Louisiana. This result is surprising, but consider that over 10,000 km of unique shoreline was surveyed during the DWH response, surveys in some areas only occurred months after probable initial oiling, and repeated surveys outside of heavily oiled areas often either were conducted once, or were separated by many months or years. We expect the vast majority of this shoreline oiling to be light, very light or trace, as defined by Michel *et al.*, (2013) because this oiling was most prevalent and most ephemeral. Note that we do not make predictions for shoreline locations more than 10 km from any on-water oil observation because such predictions would be significant extrapolation beyond model training data. Similarly, we make no predictions to shorelines that were never surveyed, because it is likely that there is substantial correlation between survey locations and oiling, and any model predictions in these locations would thusly be biased. While some shoreline oiling undoubtedly occurred further than 10km from observed on-water oiling or along shorelines that were never surveyed, it is likely that the majority of unsurveyed oiling occurred along shorelines that are represented in these estimates.

CONCLUSION

This effort is the first of its kind in attempting synoptic hind-cast shoreline oil probability modeling using estimates of on-water oil distribution derived from satellite-based SAR instruments via machine learning-based modeling methodologies. In general, the model presented here is able to reproduce calibrated estimates of shoreline oiling probability with good accuracy. The majority of the findings about individual predictor variables such as wind speed and direction and water levels comport with both intuition and previous findings regarding these factors and oil stranding probability or retention (see Etkin *et al.*, 2007). The most important

result is that the model suggests that 176 km of shoreline that was not observed as oiled in the field may have been oiled. This is a substantial increase over previous estimates, though it is likely that this represents only very light and ephemeral oiling. Regardless, we feel that these and similar approaches constitute valuable tools, together with robust field survey and hydrodynamic modeling efforts, in operational modeling and hind-casting shoreline exposure during and after large and complex spills.

REFERENCES

- Battjes, J. A. 1988. Surf-zone dynamics. *Annual Review of Fluid Mechanics* 20:257–291 doi: <https://doi.org/10.1146/annurev.fl.20.010188.001353>.
- Burman, P., Chow, E. and Nolan, D. 1994. A cross-validatory method for dependent data. *Biometrika* 81(2):351-358 doi: 10.1093/biomet/81.2.351
- Chen, T. and C. Guestrin. 2016. XGBoost: A scalable tree boosting system. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD '16)*. ACM, New York, NY, USA, 785-794. doi: <https://doi.org/10.1145/2939672.2939785>
- Elith, J., Leathwick, J. and T. Hastie. 2009. A working guide to boosted regression trees. *Journal of Animal Ecology*, 77:802-813. doi: 10.1111/j.1365-2656.2008.01390.x
- Environmental Response Management Application (ERMA). 2015. Web Application: Deepwater Gulf Response Environmental Response Management Application, National Oceanic and Atmospheric Administration. Available online at: <https://gomex.erma.noaa.gov/> (accessed on 12/1/2016).

- Environmental System Research Institute (ESRI). 2016. ArcGIS Desktop, v10.4.1. Redlands CA, Environmental Systems Research Institute.
- Etkin, D., French-McCay, D., and J. Michel. 2007. Review of the state-of-the-art on modeling interactions between Spilled Oil and Shorelines for the Development of Algorithms for Oil Spill Risk Analysis Modeling. Mineral Management Service Outer Continental Shelf Study Report 2007-063.
- Flach, P. and E. Matsubara. 2008. On classification, ranking, and probability estimation. In: de Raedt, L. T. Dietterich, L. Getoor, K. Kersting and S. Muggleton (eds.). Probabilistic, Logical and Relational Learning – A Further Synthesis. Dagstuhl Seminar Proceedings.
- Friedman, J., Hastie, T., and R. Tibshirani. 2000. Additive logistic regression: a statistical view of boosting. *Annals of Statistics* 28(2): 337-407. doi: <https://doi.org/10.1214/aos/1016218223>
- Friedman, J. 2001. Greedy Function Approximation: A Gradient Boosting Machine. *Annals of Statistics* 29(5):1189-1232. doi:10.1214/aos/1013203451
- Garcia-Pineda O, MacDonald I.R., Li X., Jackson C.R., Pichel W.G. 2013a. Oil spill mapping and measurement in the Gulf of Mexico with textural classifier neural network algorithm (TCNNA). *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 6(6):2517-2525. doi: 10.1109/JSTARS.2013.2244061
- Garcia-Pineda, O., MacDonald, I., Hu, C., Svejksky, J., Hess, M., Dukhovskoy, D., and S.L. Morey. 2013b. Detection of floating oil anomalies from the Deepwater Horizon oil spill with synthetic aperture radar. *Oceanography*, 26(2), 124-137. doi: <http://dx.doi.org/10.5670/oceanog.2013.38>

- Grant, S. B., Kim, J. H., Jones, B. H., Jenkins, S. A., Wasyl, J., and C. Cudaback. 2005. Surf zone entrainment, along-shore transport, and human health implications of pollution from tidal outlets. *Journal of Geophysical Research: Oceans*, 110(C10). doi: 10.1029/2004JC002401
- Goldstein, A., Kapelner, A., Bleich, J., and E. Pitkin. 2015 Peeking inside the black box: Visualizing statistical learning with plots of Individual Conditional Expectation. *Journal of Computational and Graphical Statistics*, 24(1): 44-65 doi: <https://doi.org/10.1080/10618600.2014.907095>
- Keddy, P.A. 1982. Quantifying within-lake gradient of wave energy: interrelationships of wave energy a, substrate, particle size and shoreline plants in Axe Lake, Ontario. *Aquatic Botany* 14(1982): 41-58. doi:10.1016/0304-3770(82)90085-7
- Leifer I., Lehr WJ, Simecek-Beatty D, Bradley E, Clark R, Dennison P, Hu Y, Matheson S, Jones CE, Holt B, Reif, M., Roberts, D., Svejksky, J., Swayze, G., and J. Wozencraft. 2012. State of the art satellite and airborne marine oil spill remote sensing: Application to the BP Deepwater Horizon oil spill. *Remote Sensing of Environment*, 124:185-209. doi: <http://dx.doi.org/10.1016/j.rse.2012.03.024>
- Michel J., Owens, E., Zengel S., Graham, A., Nixon, Z., Allard, T., Holton W., Reimer, P., LaMarche, A., White, M., Rutherford, N., Childs, C., Mauseth, G., Challenger, G., Taylor, E. 2013. Extent and Degree of Shoreline Oiling: *Deepwater Horizon* Oil Spill, Gulf of Mexico, USA. *PLoS ONE* 8(6): e65087 doi: <https://doi.org/10.1371/journal.pone.0065087>

- National Oceanic and Atmospheric Administration (NOAA) Center for Operational Oceanographic Products and Services (CO-OPS). 2016a. Verified Observed Hourly Water Levels. Available online at: <https://tidesandcurrents.noaa.gov/> (accessed on 12/1/2016).
- National Oceanic and Atmospheric Administration (NOAA) National Climate Data Center (NCDC). 2016b. Hourly Global Surface Data. Available online at: <http://cdo.ncdc.noaa.gov/> (accessed on 12/1/2016).
- National Oceanic and Atmospheric Administration (NOAA) National Data Buoy Center (NDBC). 2016c. Climatic Summary Tables. Available online at: <http://www.ndbc.noaa.gov/> (accessed on 12/1/2016).
- National Oceanic and Atmospheric Administration (NOAA) National Environmental Satellite, Data, and Information Service (NESDIS). 2013. Satellite Derived Surface Oil Analysis Products - Deepwater Horizon. Available online at: <http://www.ssd.noaa.gov/PS/MPS/deepwater.html> (accessed on 12/1/2016).
- Niculescu-Mizil, A., and R. Caruana. 2005. Predicting good probabilities with supervised learning. In: L. De Raedt and S. Wrobel, (eds.). ICML '05: Proceedings of the 22nd International Conference on Machine Learning. ACM, New York, New York, pp. 625-632. doi: <https://doi.org/10.1145/1102351.1102430>
- Nixon, Z., Zengel, S., Baker, M., Steinhoff, M., Fricano, G., Rouhani, S., and J. Michel, J. 2016. Shoreline oiling from the Deepwater Horizon oil spill. *Marine Pollution Bulletin* 107(1):170-178. doi: <http://doi.org/10.1016/j.marpolbul.2016.04.003>

- Platt, J. 2000. Probabilistic outputs for support vector machines and comparison to regularized likelihood methods . In: A. Smola, P. Bartlett, B. Scholkopf, and D. Schuurmans (eds.): Advances in Large Margin Classifiers. Cambridge, MA. doi: <http://doi.org/10.1.1.41.1639>
- R Development Core Team. 2016. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. Available online at: <http://www.R-project.org>. (accessed on 12/1/2016).
- Reniers, A. J. H. M., J. H. MacMahan, E. B. Thornton, T. P. Stanton, M. Henriquez, J. W. Brown, J. A. Brown, and E. Gallagher, 2009. Surf zone surface retention on a ripchanneled beach. *Journal of Geophysical Research*, 114, C10010, doi:10.1029/2008JC005153.
- Restrepo, J. M., Venkataramani, S. C., and C. Dawson. 2014. Nearshore sticky waters. *Ocean Modelling*, 80 49-58. doi: <http://dx.doi.org/10.1016/j.ocemod.2014.06.003>
- Rohweder, J., J. Rogala, B. Johnson, D. Anderson, S. Clark, F. Chamberlin, and K. Runyon. 2008. Application of wind fetch and wave models for habitat rehabilitation and enhancement projects. U.S. Geological Survey Open-File Report 2008–1200.
- Sing, T., Sander, O., Beerenwinkel, N., T, Lengauer. 2005. ROCR: visualizing classifier performance in R. *Bioinformatics* 21(20):3940-3941. doi: <http://doi.org/10.1093/bioinformatics/bti623>
- Streett D. 2011. NOAA's Satellite Monitoring of Marine Oil. In: Liu Y, MacFadyen A, Ji Z-G, Weisberg RH, editors. *Monitoring and Modeling the Deepwater Horizon Oil Spill: A Record-Breaking Enterprise*. Washington DC: AGU. p. 9-18. doi: 10.1029/2011GM001104

FIGURES

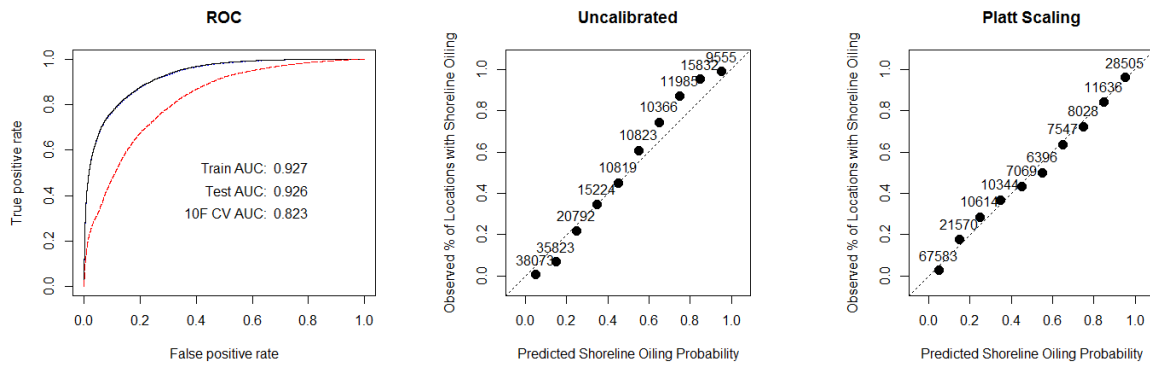


Figure 1. ROC plots and AUC values for training sample in black, independent test sample in blue, and 10-fold h-block cross validation in red (left), and calibration plots of predicted and actual empirical probabilities in 10 bins from independent calibration data before (middle), and after (right) Platt scaling

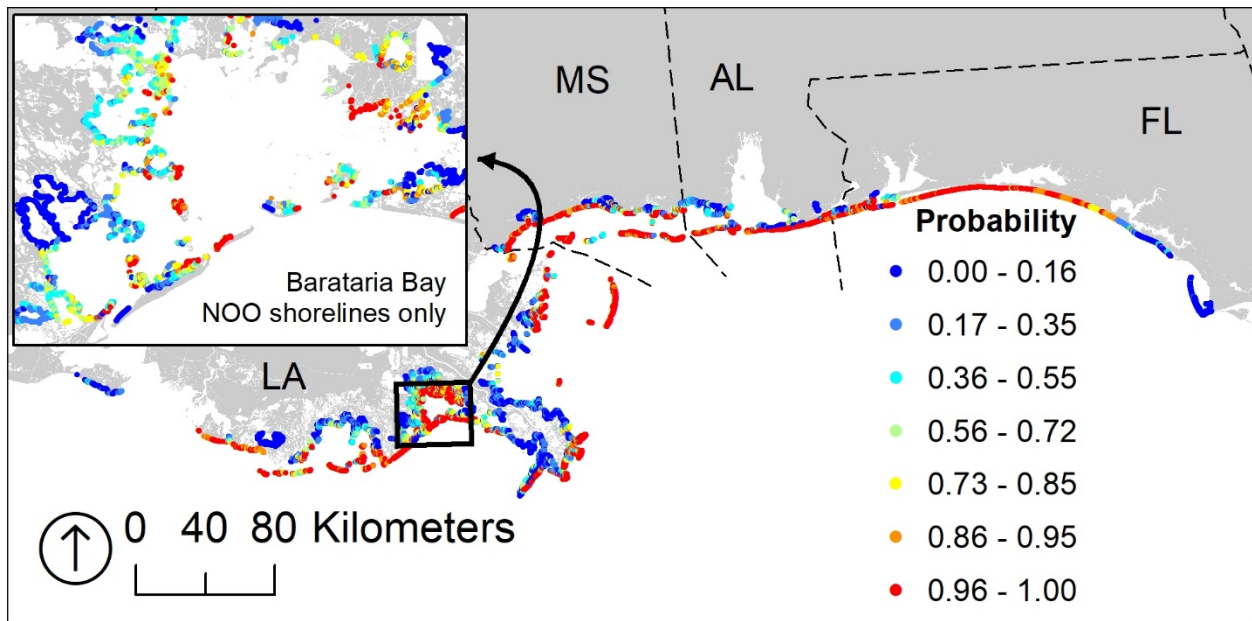


Figure 3. Map of cumulative predicted oiling probability for all observation locations. Inset shows detail of cumulative predicted oiling probability in Barataria Bay for only locations where no oil was observed (NOO) via field survey.

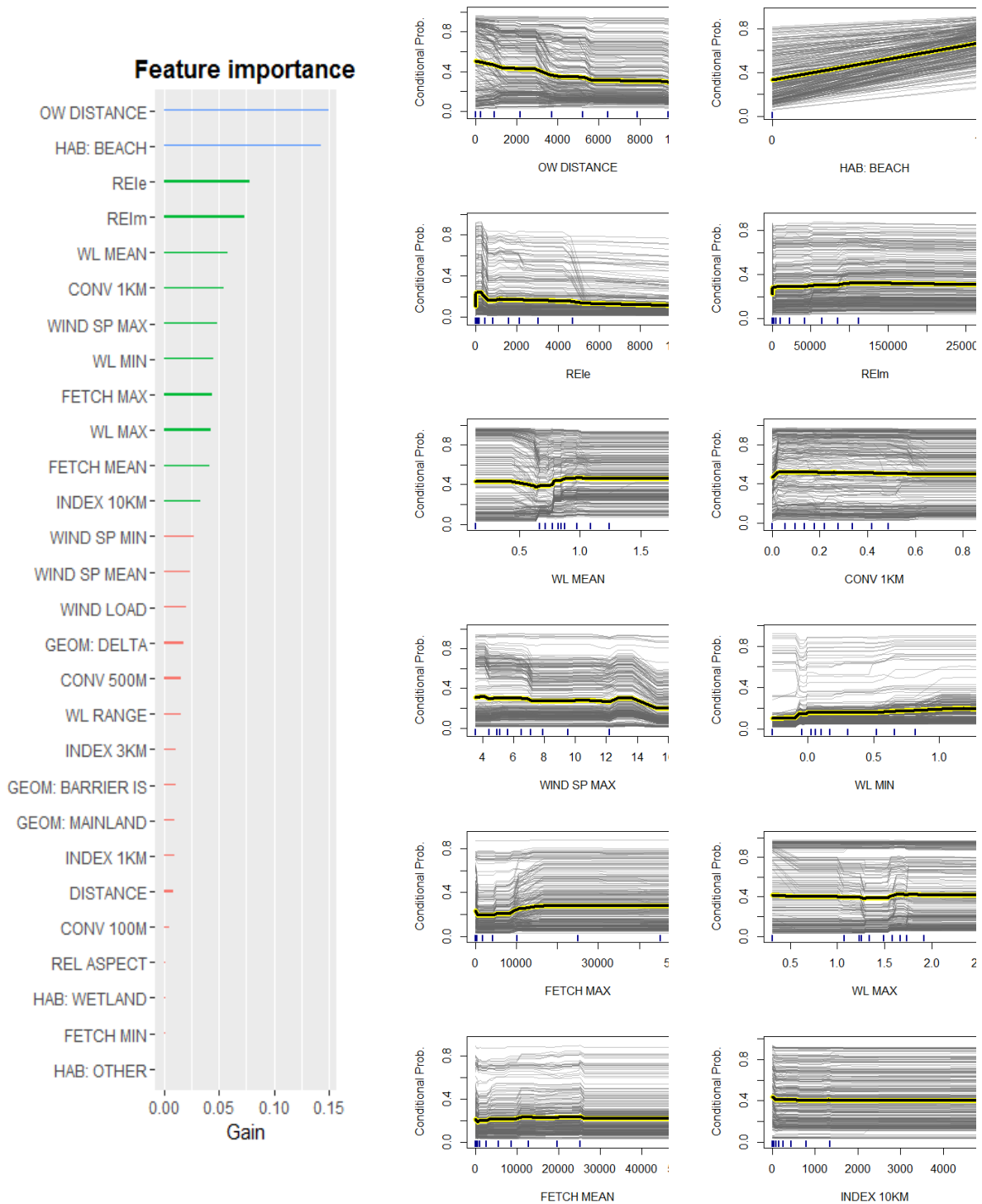


Figure 2. Relative contribution of predictor variables as model gain (left) and conditional expectation plots (right) of the top 12 variables included in the model by importance.