A GENERALIZED LINEAR MODEL APPROACH FOR BEACH CHARACTERIZATION WITH MULTI-TEMPORAL LIDAR DATA

M. J. Starek[1], K.Clint Slatton[2]

1: National Research Council Postdoctoral Associate of the Army Research Office
Department of Marine, Earth, and Atmospheric Sciences, North Carolina State University USA
2: Department of Electrical and Computer Engineering, University of Florida USA
E-mail: mjstarek@ncsu.edu

1. INTRODUCTION

Monitoring beaches and studying the processes that govern their change are critical to the future sustainability of these valued environments and the economies that depend on them. Airborne light detection and ranging (lidar) systems have revolutionized beach monitoring enabling high resolution sampling of nearshore topography over long segments of coastline quickly, accurately, and economically. Small-footprint, discrete-return systems enable beach and upland mapping with average spatial resolutions greater than 1 point per m^2, vertical accuracy (z) of 5–10 cm, and horizontal accuracy (x,y) of 15-20 cm [1]. From the data, high resolution digital elevation models (DEMs) can be generated. By differencing DEMs or contours generated from repeat pass surveys, the change in volume or shoreline position for a beach can be measured respectively [2] [3]. Additionally, features can be extracted to measure changes in nearshore geomorphology [4] [5].

Data mining and pattern classification techniques offer great potential to move the analysis of lidar data beyond visual interpretation and simple (first order) measurements made from DEMs but to date have been relatively unexplored. This is particularly true for beach monitoring with lidar data because of the importance of subtle variation in topography and non-stationary processes along the beach, such as localized "hot spots" of anomalous erosion or accretion. When acquired with sufficient temporal coverage, the high spatial-resolution information in the lidar data can resolve non-stationary processes and reveal patterns in beach change otherwise unforeseen. Here, we present a systematic framework to mine morphologic features from time series lidar data acquired over a beach and characterize the joint effect of the features on the outcome of erosion. Our approach is stochastic in nature, and we use logistic regression to model the variation in morphology on probability of shoreline erosion along the beach (alongshore). Important features are methodically detected and the resultant models can be used for classification of high impact zones.

2. DATA PROCESSING

Airborne lidar surveys were acquired by University of Florida (UF) over the St. Augustine Beach region of Florida USA seven times between August 2003 and February 2007. This resulted in six sequential coverage epochs ranging in temporal spans of less than two months to over a year and a half. A 10 km stretch of beach was selected for investigation because it contains a historical accretion zone in the southern portion and a highly erosive pier zone to the north that requires periodic beach nourishment to maintain the usable beach area. The survey parameters coupled with the specific lidar system's scan properties resulted in a mean ground point density of 1.3 points/m^2. The raw point data underwent a series of pre-processing steps and were then interpolated to 1 m bare-earth elevation grids for each acquisition. The shoreline was delineated using the Mean Higher High Water (MHHW) tidal datum [3].

3. FEATURE EXTRACTION

The natural coordinate system traditionally used for studying shoreline change consists of a local 2D Cartesian system oriented with alongshore and cross-shore axes. The lidar-derived DEMs were sampled by an algorithm that auto-extracts elevation values along cross-shore profile lines oriented roughly orthogonal to the shoreline contour [5]. This provides the x,y-coordinates and elevation values along each profile at a user-defined spacing. The profiles were extracted every 5 m in the alongshore direction and they extend in the cross-shore from the dune toe line (shoreward edge of dune) to the MHHW shoreline contour. The temporal change in shoreline between the data acquisition periods was quantified using pairs of profiles, and the profiles were then segmented into binary *erosion tending* or *accretion tending* classes depending on whether the shoreline had moved landward or seaward during the epoch following a particular lidar acquisition. This approach allows us to extract several features per profile and directly relate to shoreline change for a given epoch. A total of ten features were extracted to describe the variation in morphology along the beach (Table 1).

Feature	Units	Feature	Units
Beach slope (S)	(m/m)	*Mean curvature* (C)	(m/m^2)
Near-shoreline slope (NS)	(m/m)	*Orientation* (O)	(degrees)
Beach width (W)	(m)	*Standard deviation of height* (SD)	(m)
Volume-per-width (V)	(m^3/m)	*Deviation-from-trend* (DT)	(m)
Mean gradient (G)	(m/m)	*Max gradient* (MG)	(m)

Table 1. Extracted features to characterize beach morphology alongshore.

4. GENERALIZED LOGISTIC MODEL

To model the joint effect of the features on the outcome of erosion, logistic regression is applied. Logistic regression is a member of the family of generalized linear models (GLMs) and can be applied to regress a binary class variable (0,1) on predictor variables [6]. Furthermore, logistic regression provides a rigorous statistical framework to evaluate the influence (importance) of each feature within the fitted model on determining class occurrence. Consider our case where we have binary observations, y_i for $i = 1$ to N profiles, indicating whether profile i belongs to class *erosion tending* or class *accretion tending* as defined previously. We arbitrarily set

$y_i = 1$ to indicate that a profile belongs to class erosion and $y_i = 0$ to indicate that a profile belongs to class accretion. Our objective then is to estimate the probability of class erosion, $y_i = 1$, for a specific profile given an input vector of morphologic features, $\mathbf{x}_i \rightarrow P_i(y_i = 1 \mid \mathbf{x}_i)$. The basic logistic model assumes the following parametric form for P_i [6] [7],

$$P_i(y_i = 1 \mid \mathbf{x}_i) = \frac{1}{1 + e^{-\mathbf{x}_i \beta}} = \frac{e^{\mathbf{x}_i \beta}}{1 + e^{\mathbf{x}_i \beta}} \tag{1}$$

where β is a k + 1 column vector of coefficients for our k = 10 features and an intercept term and $\mathbf{x}_i \beta = \beta_0 + \beta_1 x_{i1} + \ldots + \beta_k x_{ik}$ is called the linear predictor. To model our data effectively, a generalized estimating equation (GEE) is applied to handle spatial correlation in the binary responses [8]. To reduce model over fitting and address collinearity among the features themselves, the penalized Lasso criteria is employed providing more robust classification [9].

5. RESULTS

Logistic GEE models were fit to each data epoch to evaluate their impact on class occurrence (Table 2). The most influential positive and negative coefficients for each epoch are shown in bold. The features are standardized to provide ease of interpretation. Positive coefficients indicate that an increase in feature value results in an increase in probability of class erosion, and the converse is true for negative coefficients. For example, in Epoch 1 beach width (W) is most influential on the outcome of class erosion where a 1 σ increase in W produces, on average, a 1.70 increase in the log odds of class erosion ($\sigma_{width} \approx 33$ m). Note that Lasso penalization results in a reduction of feature coefficient magnitudes, with some non-significant terms reduced to zero; however, the relative importance of the features for a given epoch was generally maintained.

Feature	Epoch 1	Epoch 2	Epoch 3	Epoch 4	Epoch 5	Epoch 6
W	**1.70**	**2.76**	0.47 *	0.01	1.38	-0.60*
V	0.95	**-1.80**	0.21	-0.07*	-0.02	0.63
DT	-0.74	-0.93	**-0.49**	**-1.12***	**-0.90**	**1.55** *
MG	0.15	0.18	-0.08*	0.12 *	-0.82	-0.30*
NS	-0.09	0.13	-0.18*	0.01 *	**1.98**	-0.06*
O	0.29	-0.27	**1.38**	**1.18**	-0.17	-0.93*
G	0.71	-0.26*	0.09 *	-0.19*	0.66	0.41 *
C	0.12	-0.03*	0.05 *	0.03 *	-0.10	-0.74*
SD	0.14	0.07*	0.01	-0.01*	-0.01	0.03 *
S	**-1.47**	-0.12*	-0.49*	-0.40	-0.87	**-1.19***

Table 2. Fitted logistic GEE feature coefficients for each epoch. * indicate potentially non-significant features.

To provide an example of the potential utility of the logistic GEE model for beach characterization, Figure 1 shows estimated probability of erosion as a function of orientation (O). Four logistic plots are displayed: change in probability vs. O given the mean values of all other features, given a two deviation increase in S, given a two deviation decrease in S, and given a two deviation increase in S and DT. From the plot, we observe that as the beach approaches a more east to southeasterly orientation, its probability to erode increases.

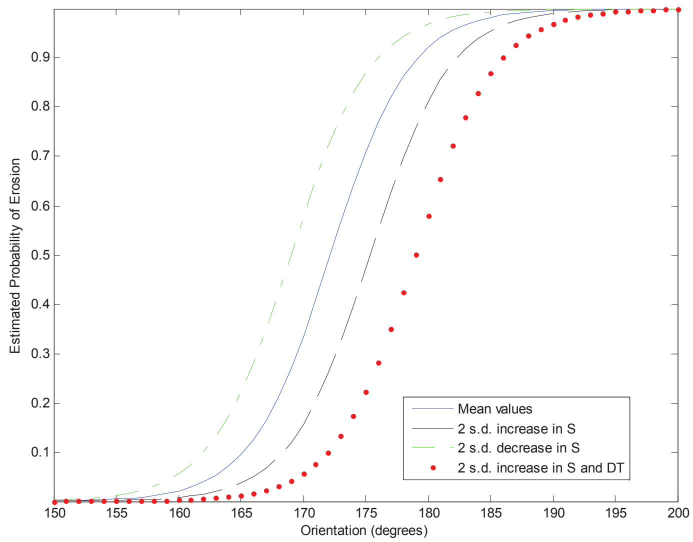

Figure 1. Estimated probability of erosion vs. orientation for Epoch 3 (s.d. is standard deviation).

Finally, we evaluated the ability of the Lasso penalized GEE model to classify profiles more or less prone to erosion given the morphologic features. For each epoch a random subset of the profiles (< 30%) was selected for model training and the remaining majority of profiles reserved for testing. The Lasso model had a mean classification accuracy of 80% across all epochs with a high classification rate of 86%. Overall, results are encouraging and demonstrate the potential of the developed logistic framework to effectively model and characterize patterns in beach change measured within multi-temporal lidar data.

6. REFERENCES

[1] Slatton, K.C., Carter, W.E., Shrestha, R.L., Dietrich, W., 2007. Airborne laser swath mapping: achieving the resolution and accuracy required for geosurficial research. Geophysical Research Letters 34: L23S10, doi:10.1029/2007GL031939.
[2] Sallenger, A.H, et al., 2003. Evaluation of airborne topographic lidar for quantifying beach changes. Journal of Coastal Research 19 (1),125-133.
[3] Starek, M.J., Slatton, K.C., Shrestha, R.L., Carter, W.E., 2009. Airborne lidar measurements to quantify change in sandy beaches. In: Laser Scanning for the Environmental Sciences, Blackwell Publishing, Oxford (in press).
[4] Saye, S.E., van der Wal, D., Pye, K., and Blott, S.J., "Beach–dune morphological relationships and erosion/accretion: An investigation at five sites in England and Wales using LIDAR data", Geomorphology, 72, pp. 128– 155, 2005.
[5] Starek, M.J., Vemula, R.K., Slatton, K.C., Shrestha, R.L., Carter, W.E., 2007. Automatic feature extraction from airborne lidar measurements to identify cross-shore morphologies indicative of beach erosion. Proceedings of IEEE International Geoscience and Remote Sensing Symposium, IGARSS 07, 2511-2514.
[6] Myers, R.H., Montgomery, D.C., Vining, C.G., 2002. Generalized Linear Models with Applications in Engineering and the Sciences. John Wiley & Sons, New York.
[7] Hastie, T., Tibshirani, R., Friedman, J., 2001. The Elements of Statistical Learning: Data Mining, Inference, and Prediction. Springer, New York.
[8] Fu, W.J., 2003. Penalized estimating equations. Biometrics 59, 126-132.
[9] Tibshirani, R., 1996. Regression shrinkage and selection via the Lasso. Journal of Royal Statistical Society B (58), 267-288.A.B. Smith, C.D. Jones, and E.F. Roberts, "Article Title," *Journal*, Publisher, Location, pp. 1-10, Date.