Marine Geology 278 (2010) 140–149

Contents lists available at ScienceDirect

Marine Geology

journal homepage: www.elsevier.com/locate/margeo

Predicting coastal cliff erosion using a Bayesian probabilistic model

Cheryl Hapke [a,*], Nathaniel Plant [b]

[a] US Geological Survey, Woods Hole Coastal and Marine Science Center, Woods Hole, MA 02543, United States
[b] US Geological Survey, St. Petersburg Coastal and Marine Science Center, St. Petersburg, FL 33701, United States

ARTICLE INFO

Article history:
Received 1 June 2010
Received in revised form 22 September 2010
Accepted 3 October 2010
Available online 30 October 2010

Communicated by J.T. Wells

Keywords:
coastal erosion
coastal cliffs
Bayesian
predictive model
Southern California

ABSTRACT

Regional coastal cliff retreat is difficult to model due to the episodic nature of failures and the along-shore variability of retreat events. There is a growing demand, however, for predictive models that can be used to forecast areas vulnerable to coastal erosion hazards. Increasingly, probabilistic models are being employed that require data sets of high temporal density to define the joint probability density function that relates forcing variables (e.g. wave conditions) and initial conditions (e.g. cliff geometry) to erosion events. In this study we use a multi-parameter Bayesian network to investigate correlations between key variables that control and influence variations in cliff retreat processes. The network uses Bayesian statistical methods to estimate event probabilities using existing observations. Within this framework, we forecast the spatial distribution of cliff retreat along two stretches of cliffed coast in Southern California. The input parameters are the height and slope of the cliff, a descriptor of material strength based on the dominant cliff-forming lithology, and the long-term cliff erosion rate that represents prior behavior. The model is forced using predicted wave impact hours. Results demonstrate that the Bayesian approach is well-suited to the forward modeling of coastal cliff retreat, with the correct outcomes forecast in 70–90% of the modeled transects. The model also performs well in identifying specific locations of high cliff erosion, thus providing a foundation for hazard mapping. This approach can be employed to predict cliff erosion at time-scales ranging from storm events to the impacts of sea-level rise at the century-scale.

Published by Elsevier B.V.

1. Introduction

Coastal erosion is a worldwide societal issue and problems associated with it are expected to increase with rising in sea levels. With the recognition of the hazards facing coastal development, there is an increasing interest in models that forecast or predict where the highest hazards will be, whether from the gradual sea-level rise or large storm events. Existing models that forecast coastal response to sea-level rise or high water levels (i.e. storm surge or swell) are typically geometric models that focus mainly on sandy and dune-backed coasts (Bruun, 1962; Komar et al., 1999). Along geologically and geomorphically variable coasts such as are found along the U.S. west coast, there is a need to model both sandy beaches and coastal cliff systems. Existing geometric models for predicting erosion rates on bluffed or cliffed coasts typically apply a modified Bruun Rule (Bray and Hooke, 1997), whereas similar but empirically driven models rely on historical water level data in conjunction with cliff toe elevations (Ruggiero et al., 2001) to establish potential erosion hazard zones. Sunamura (1982) combined empirically derived data from wave tank experiments with measurements of cliff geometry, material strength and wave frequency distributions to predict cliff erosion rates. More simplistically, hazard zones are delineated by a linear forward projection of historic rates (Moore et al., 1999; Priest, 1999). All of these models and methods focus on climatologically averaged predictions but generally do not account well for the spatial and temporal variability and uncertainty of cliff retreat processes.

Recently, there has been a focus on the development of statistical and probabilistic models of coastal cliff retreat of which Lee et al. (2001) provide detailed examples. A conclusion they reach is that cliff retreat is not amenable to statistical forecasting models because each retreat event is not independent (i.e. each event is influenced by previous events) but that a probabilistic model can accommodate the spatial and temporal uncertainty inherent to the process of cliff retreat. Hall et al. (2002) utilized probabilistic models to predict the maximum likelihood distributions of cliff failure based on a time series of historic cliff retreat data, and additionally developed a Bayesian probabilistic model based on historical retreat rates, recent observations and expert assessments of the expected recession frequencies. They generated a probability density function of cliff retreat for various future time periods (22 to 84 yr), but the output was not in the form of geospatial data suitable for generating hazard maps, nor do they provide a verification of the model output.

The Bayesian approach is well-suited to the prediction of cliff retreat due to its ability to include prior (e.g., historic) information and to address the complexity of the feedback mechanisms inherent

* Corresponding author.
E-mail addresses: chapke@usgs.gov (C. Hapke), nplant@usgs.gov (N. Plant).

0025-3227/$ – see front matter. Published by Elsevier B.V.
doi:10.1016/j.margeo.2010.10.001

in cliff failure processes. In addition it addresses the existing difficulties in correlating the multiple variables that influence the process and is able to infer the relationships between them. As noted previously, cliff retreat is not a continuous process; it occurs episodically and is difficult to predict. A model must track the cliff evolution to account for the fact that a cliff must have an unstable geometry to fail, but after failure it may be in a stable configuration for a long period of time. In a Bayesian network, variables such as cliff geometry can be updated as changes occur, and more information can be added to the model as it becomes available.

Bayesian statistics have been used in the geosciences for several decades, mainly in the seismic and landslide hazard communities. The heavy computational requirements of calculating joint probability distributions using Bayesian methods limited its application until recently with the accessibility of fast and powerful computers and availability of software compatible with most standard desktop computers. Examples of seismic applications include the use of Bayesian statistics to generate extreme-value distributions of earthquake occurrences (Cambell, 1982; Stavrakakis and Drakopoulos, 1995) in California and Greece, respectively. Additionally, Oh et al. (2008) apply Bayesian methods to develop early warning systems for earthquakes, and Amiri and Tabatabaei (2008) use a Bayesian approach in earthquake risk assessment studies. Gritzner et al. (2001) employ a Bayesian probability model to help identify which variables are most important in watershed-scale landslide risk assessments. Lee et al. (2002), Demoulin and Chung (2007), and Miller et al. (2007) develop geospatial landslide susceptibility models based on Bayesian statistics. Dahal et al. (2008) assessed the predictive performance of Bayesian statistics to map landslide hazards and found a prediction accuracy greater than 88%.

Coastal cliff failure and accompanying cliff edge retreat are similar in behavior to earthquakes and landslides, exhibiting nonlinear behavior with occurrences that are episodic in both time and space. Few studies have been conducted that explore the use of Bayesian methods to predict future behavior of coastal systems (Hall et al., 2002; Plant and Holland, in press). In this study, we apply a Bayesian network to model the probability of coastal cliff retreat in two study areas located in Southern California. Cliff retreat within a study area along the San Diego coast (Fig. 1) is modeled over a 3-yr time period and a study area in Santa Barbara is modeled over a 7-yr time period.

2. The Bayesian model

An advantage of a Bayesian inference approach is that it can be used to combine multiple parameters to make statistically robust forecasts. Additionally, unlike more classical inferential models, Bayesian models permit the incorporation of prior knowledge, and a Bayesian network allows the utilization of conditional probabilities in the predictive model.

Bayes rule is expressed as:

$$p\left(F_i|O_j\right) = p\left(O_j|F_i\right)p\left(F_i\right) / p\left(O_j\right), \tag{1}$$

where the left-hand term is the updated conditional probability (or 'posterior probability') of a forecast, F_i, given a particular set of observations, O_j. In the case of this study, we are forecasting the probability distributions of cliff erosion on a number of transects

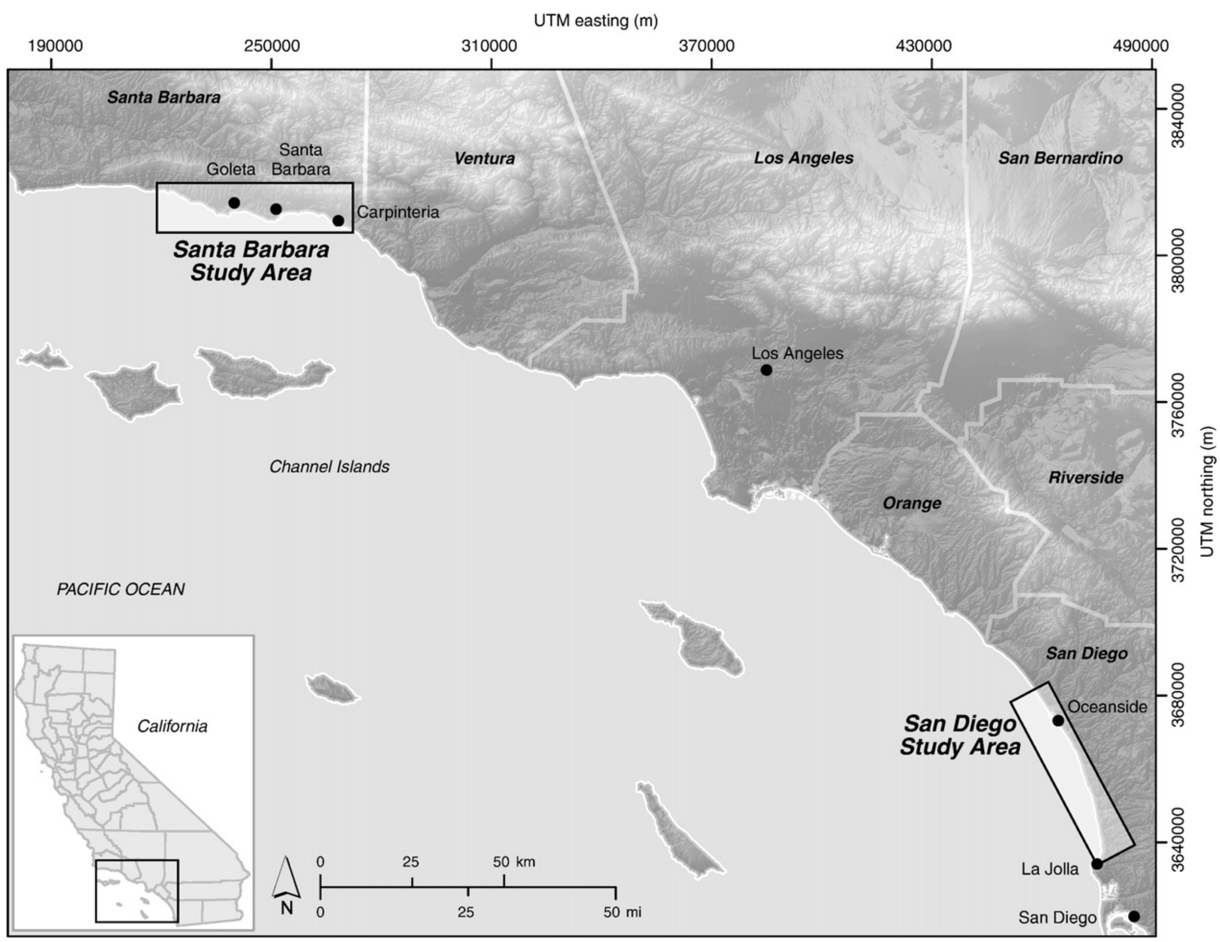

Fig. 1. Location map of study areas in Southern California.

along the coast. The term $p(O_j|F_i)$ is the likelihood of the observation if the forecast is known and is the inverse of the left-hand term—this term contains all the information relating cliff retreat to the input variables, and includes uncertainties due to measurement or process errors. The term $p(F_i)$ is the prior probability of F and is what is known about the problem before new or additional information is available. Finally, the term $p(O_j)$ is the sum of the joint probabilities of all the observations and is a normalizing or scaling factor that is independent of F.

The Bayesian network approximates a generalized probability density function for a domain containing a number of variables, each having either discrete (e.g., a lithologic description) or continuous (e.g., cliff height) states. The network consists of a set of conditional probability density functions combined with a set of conditional independence assertions (indicating which parameters are related to one another and which are not) that together tell the domain how the individual variables are causally related. Bayesian networks are often represented using a schematic diagram to represent the variables and their conditional probabilities. Processes of cliff erosion, including the feedback and interplay between the variables, are shown in a simplified schematic in Fig. 2.

To develop a Bayesian network, the data for each parameter (described in detail in Section 3) were binned (Table 1), with the size of the bins optimized to provide as wide a distribution as possible. Therefore, the bin widths are not equal for the distribution of an individual parameter. In the case of the long-term erosion rate, the bin widths also attempt to represent the uncertainty limits of the data, and therefore the bins could not be divided in increments smaller than 0.05 m/yr. As a result, the input distribution for the long-term erosion rate is skewed towards the lower values. Figs. 3 and 4 show the distribution of the input data for each parameter.

The Bayesian approach allows for the incorporation of existing knowledge and conditional relationships into a system. In our network, this is indicated by unidirectionally linked parameters (i.e. geology, cliff slope, etc.). The links imply a causal relationship in the direction indicated by the vector. For example, we infer that the cliff slope is directly influenced by wave impact hours, geology and long-term erosion rate and that the only parameter in our model that directly influences cliff height is the geology. All parameters are inferred to contribute to the short-term retreat.

3. Application to Southern California cliffs

3.1. Model data

The study areas are an approximately 52 km stretch of coast in northern San Diego County, California, extending from the southern portion of Camp Pendleton to La Jolla, and an approximately 60 km portion of the Santa Barbara County coast (Fig. 1). The geomorphology along this coastline is variable and consists of cliffed coastline interrupted by coastal lagoons, river outlets and harbors.

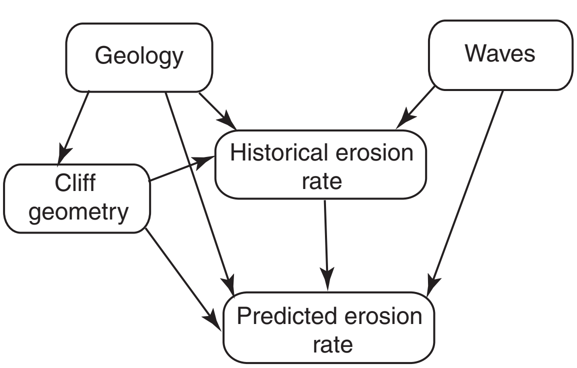

Fig. 2. Simplified schematic diagram of a Bayesian network in the context of modeling coastal cliff retreat.

Table 1
Bin boundary values for model parameters.

Model parameter	Bin boundary values	
	San Diego	Santa Barbara
Long-term retreat rate (m/yr)	−0.05 −0.15 −0.45 −0.065 −1.0	0 −0.1 −0.2 −0.4 −0.8
Cliff slope (°)	0 15 30 45 70	0 10 30 50 70
Cliff height (m)	0 10 20 30 60 120	0 10 20 30 60
Geologic ranking	1 2 3 4	1 2
Impact hours (h)	0 100 1000 10000 100000	0 50 500 2500 5000 10000
Short-term retreat (m)	0–0.5 −0.75 − 1.0 −2.0 −5.0	0–1.0 −2.0 −3.0 −5.0 −30.0

Long-term retreat rate = 70 yr; short-term retreat is 3 yr (San Diego) and 7 yr (Santa Barbara).

The model data were generated for 100-m spaced transects, in coincident locations to Coastal Data Information Program (CDIP) nearshore wave transformation sites consisting of predictions along profiles being developed as part of a U.S. Geological Survey multi-hazards evaluation project (Barnard et al., 2009). Transects coincident with cliffs and data coverage results in model development for 194 transects in the San Diego study area and 157 transects in the Santa Barbara area.

The input geospatial database consists of five parameters including long-term historical retreat rate, cliff slope and height, geology, and wave impact hours. The selected parameters represent prior behavior, initial state and forcing, respectively, and were chosen as being critical for forecasting retreat. Additionally, datasets of 3- and 7-yr (San Diego and Santa Barbara, respectively) cliff retreat were generated to independently validate the results. In the current iteration of the model presented herein, terrestrial forcing in the form of precipitation and associated groundwater forcing are not incorporated into this model. We recognize that rainfall-associated processes are important in coastal cliff retreat (Hampton and Griggs, 2004; Collins and Sitar, 2008; Young et al., 2009). However, we assume that the spatial variation in rainfall along the ~50–60 km study areas over the modeled durations was negligible. Data ranges and means for the cliff slope, height and short-term retreat are shown in Table 2.

3.1.1. Cliff slope and height

Many researchers have documented the relationship between slope angle, cliff height, and the processes governing cliff retreat (Edil and Vallejo, 1980; Emery and Kuhn, 1982; Sunamura, 1982, 1983; among many others) and these studies are rooted in the basic principles of rock and soil mechanics. In general, steeper slopes are in less stable configuration than less steep slopes and are more likely to fail, resulting in cliff retreat. In addition, because slope failure is a gravity-driven process, the higher a cliff section, the more prone it is to failure given an identical section of a cliff that is not as high. For this analysis, we derive the cliff slope and height from the lidar data for the initial date of the short-term assessment (1998 for Santa Barbara and 2003 for San Diego).

For the San Diego County study area, lidar data were acquired for 2003 and 2006 from the NOAA Coastal Services Center website (http://www.csc.noaa.gov/digitalcoast/data/coastallidar/index.html). For Santa Barbara, the available lidar data are 1998 and 2005. The data for both study areas were chosen based on availability and coverage, and to allow for a long enough duration between surveys to ensure that sufficient retreat of the cliff occurred with which to evaluate the model results.

All lidar datasets were gridded to 1 m cell size using a natural neighbors interpolation technique (Hapke, 2005). The slope of the cliff face was measured from the earliest lidar dataset in both study areas

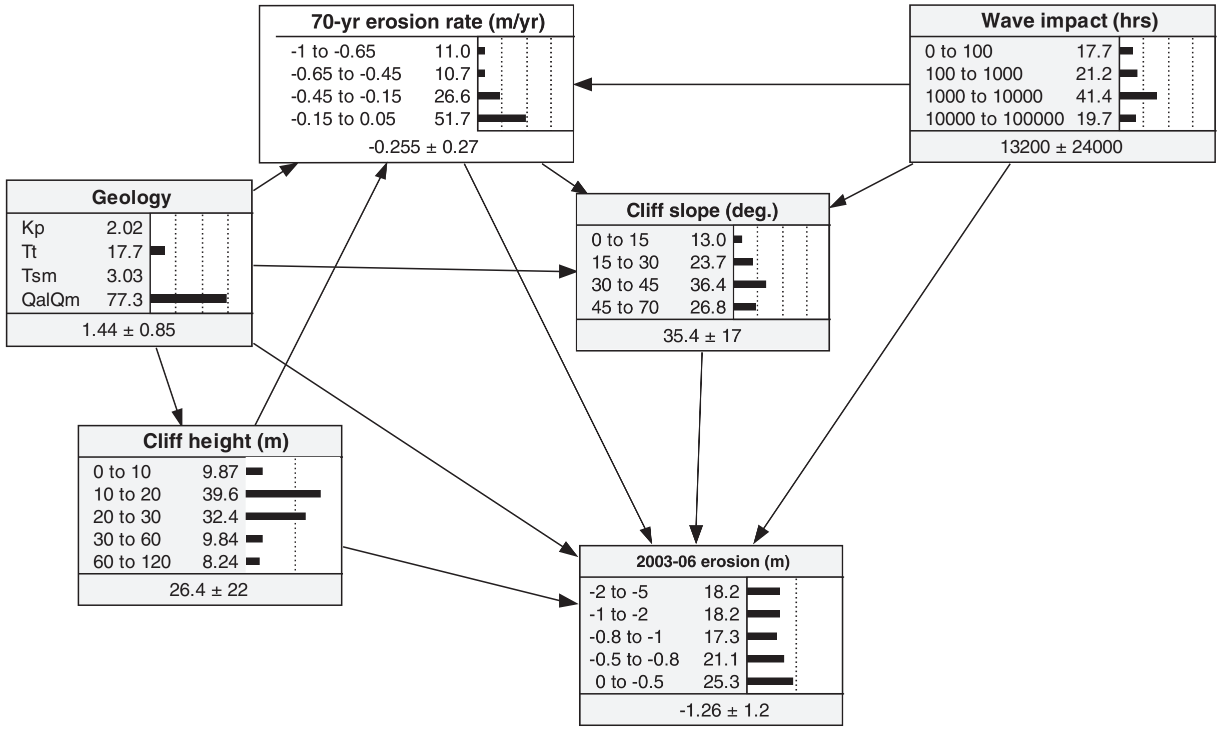

Fig. 3. Bayesian network for the San Diego study area. The bin divisions and the prior distributions ($p(F_i)$) are shown for each parameter. The first column in each box contains the bin boundary ranges and the second column (adjacent to the histogram) is the percentage of the prior distribution found in each bin. The value at the bottom of each box is the mean of the prior distribution values ± one standard deviation. The probability distribution of erosion from 2003 to 2006 is what the model is predicting ($p(F_i|O_j)$). K_p, T_t, T_{sm} and $Q_{al}Q_m$ in the geology parameter are the varying lithologies in the study area and are described in more detail in the text.

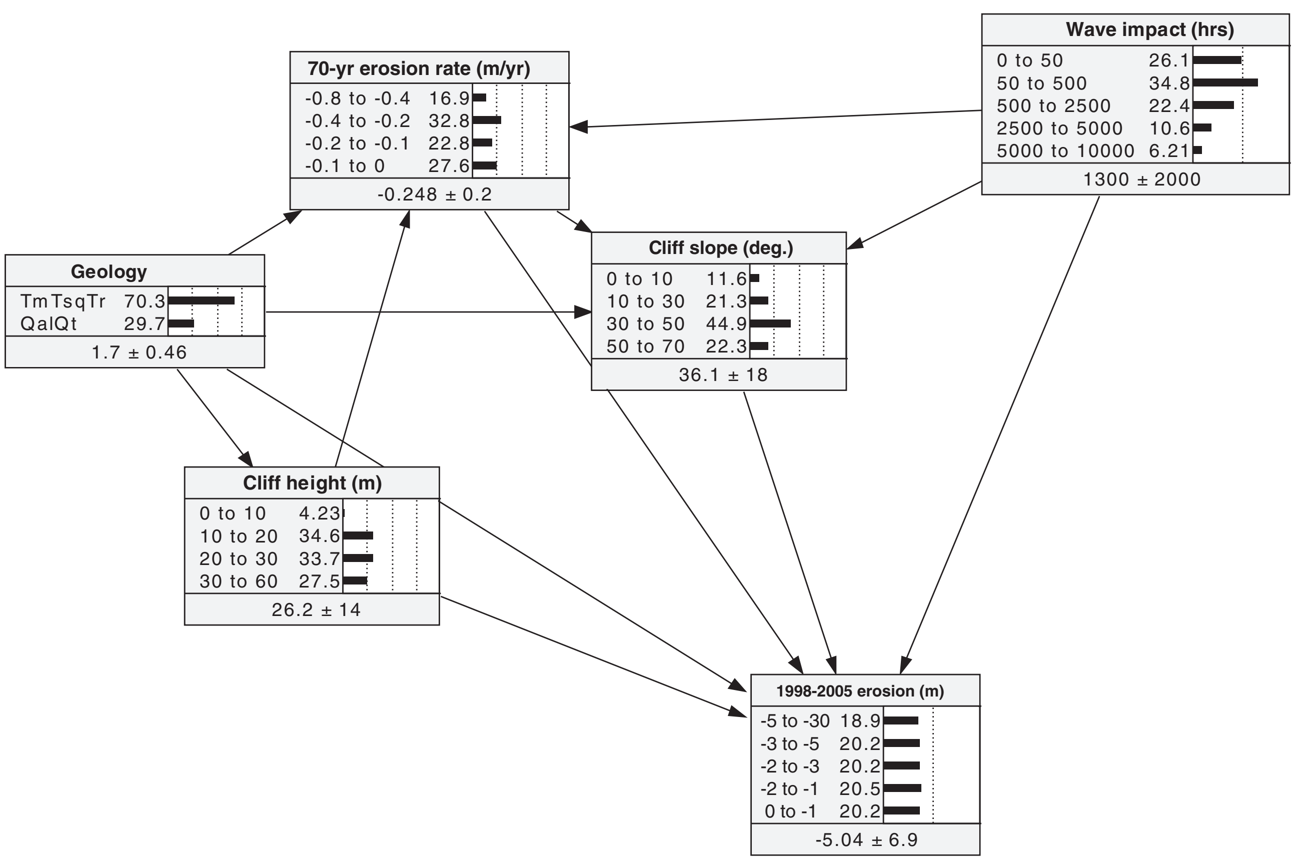

Fig. 4. Bayesian network for the Santa Barbara study area. The bin divisions and the prior distributions ($p(F_i)$_) are shown for each parameter. The first column in each box contains the bin boundary ranges and the second column (adjacent to the histogram) is the percentage of the prior distribution found in each bin. The value at the bottom of each box is the mean of the prior distribution values ± one standard deviation. The probability distribution of erosion from 1998 to 2005 is what the model is predicting ($p(F_i|O_j)$).T_t $T_{sq}T_{sm,}$ and $Q_{al}Q_t$ in the geology parameter are the varying lithologies in the study area and are described in more detail in the text.

Table 2
Range and means of data variables incorporated into the Bayesian network.

Model parameter	Max.	Min.	Mean
San Diego (n = 194)			
Long-term retreat rate	−1.0 m/yr	0.0 m/yr	−0.1 m/yr
Cliff slope	66.5°	10.3°	36.1°
Cliff height	96.5 m	4.3 m	24.2 m
Geologic ranking	4	1	–
Impact hours	22,546 h	0 h	4697 h
Short-term retreat	−4.8 m	0.0 m	−0.8 m
Santa Barbara (n = 156)			
Long-term retreat rate	−0.8 m/yr	0.0 m/yr	−0.2 m/yr
Cliff slope	63.6°	14.9°	40.5°
Cliff height	50.0 m	6.2 m	24.5 m
Geologic ranking	2	1	–
Impact hours	9271 h	0 h	1142 h
Short-term retreat	−26.0 m	−0.2 m	−3.1 m

to represent an initial condition slope. A straight-line slope was calculated at each transect by connecting the points representing the cliff top edge and base. The 'end-point slope' method does not account for complex cliff face geometries. However, it is a good indicator of whether the cliff is in a more or less stable geometry.

The top edge of the cliff was interpreted and digitized from both lidar survey dates for each study area using methods developed by Hapke and Reid (2007) and Hapke et al. (2009) and the elevation of the earliest edge was determined from the grid cell elevation value where the each transect intersects the cliff edge. Table 2 contains the ranges and means of the cliff height and slope datasets incorporated into the model.

3.1.2. Historical cliff retreat rates

Long-term erosion rates are input as a descriptor of the prior behavior of the cliffs. The rates are approximately 70-yr historic rates calculated using the cliff top edge data generated by Hapke and Reid (2007). The cliff edges were originally derived from the 1930s-era historical maps and modern (1998) lidar data, and the rates of cliff retreat calculated on 50-m spaced transects. For this analysis, rates were re-generated at each transect using DSAS (Digital Shoreline Analysis System; Thieler et al., 2009) so that the erosion rates corresponded to the CDIP transects. The uncertainties in the historical rate assessments are taken to be the same as those detailed in Hapke and Reid (2007), ± 0.1 m/yr for the average annualized rates. Long-term rates in each study area ranged from zero to 0.8 m/yr (Santa Barbara) and zero to 1.0 m/yr (San Diego) (Table 2).

3.1.3. Geology

A simplified geologic parameter, intended to represent a relative 'erodibility ranking', was assigned at each transect by identifying the cliff-forming geologic unit. Geologic units were initially determined from a digital statewide coastal geology GIS database (Griggs, 2002) and subsequently refined using available geologic maps of the area: Minor et al. (2002) for the Santa Barbara study area, and Kennedy and Tan (2008) and Kennedy et al. (2007) for the San Diego study area. The erodibility ranking is broadly based on the lithology and age of the cliff-forming material and is not empirically derived from field data or geotechnical reports.

In the San Diego study area, there are four geologic units that we rank from 1 to 4 for their relative erodibility (1 being the weakest, 4 being the most resistant). (1) Quaternary shallow marine and alluvial deposits (Q_{al} and Q_m, respectively); (2) shallow marine sandstones, siltstones and conglomerates of the late Pliocene San Mateo Formation (T_{sm}); (3) middle Eocene sandstones including the Torrey Sandstone and the Scripps Formation (T_t); and (4) the Point Loma Formation, an upper Cretaceous interbedded sandstone and shale (K_p).

For Santa Barbara, the geologic units were grouped into two relative erodibility categories: (1) Quaternary shallow marine and alluvial deposits (Q_{al} and Q_m, respectively); and (2) various Pliocene and Miocene-age shales and mudstones of the Monterey Formation, and the Sisquoc Formation, a lower Pliocene and upper Miocene mudstone, shale and conglomerate (T_m, T_{sq}, and T_r). The geologic units were broadly categorized, partly for simplification, but also to establish a more regional database in which the parameters from different areas can be integrated into one model without revising the categories.

3.1.4. Wave impact hours

To estimate the amount of time that total water levels exceeded the elevation of the base of the cliff, wave impact hours were predicted using the methods developed by Ruggiero et al. (1996, 2001). For the total water level predictions, local wave and tide data were obtained from NOAA (http://www.ndbc.noaa.gov/) for the 3- and 7-yr periods over which cliff retreat predictions are being evaluated. Beach slope and cliff base elevation for the run-up calculations at each transect were extracted from the earliest date of lidar data. The wave run-up equation of Stockdon et al. (2006) was then used in conjunction with tide data to predict the total number of hours that waves exceeded the elevation of the cliff base over the 3- and 7-yr durations of the cliff retreat model. The results indicate that water levels exceed cliff base elevations much more frequently in San Diego than that in the Santa Barbara study area (Table 2) indicating that the elevation of the shoreline angle (the intersection of the cliff and the beach) is in general lower in San Diego.

3.1.5. Short-term cliff retreat

DSAS was used to calculate the amount of landward retreat of the top edge of the cliff over the model period: 3 yr for San Diego and seven for Santa Barbara. The cliff top edges were generated from the lidar data using the approach described earlier in this section. Retreat was highly variable along the coast in both study areas, and maximum values of −4.8 m and −26.0 m were measured in San Diego and Santa Barbara, respectively (Table 1). The short-term retreat is only used to independently validate the model predictions.

4. Results

An initial run of the model used three of the five parameters to assess how well the model performed with minimal inputs. This parsimonious approach serves as a baseline to determine if the other inputs add any real predictive value. The parameters included in the initial run were (1) long-term erosion rate, (2) cliff slope and (3) wave impact hours. The Bayesian network was used to determine the probability distribution of the cliff retreat over 3 yr in San Diego and 7 yr in Santa Barbara, based on the distributions of the input parameters and the conditional probability relationships established between them. To evaluate the model results, we produced a second model output that included the actual measured outcome (i.e. cliff retreat). The evaluation indicated that the three-parameter model was able to accurately predict the correct cliff retreat amount only ~50% of the time, essentially the same as chance.

The model was assessed again using all five parameters. Including the ranked geology parameter and cliff height noticeably improved the outcome. At approximately 30% of transects, the model predicted an equally high probability value of short-term erosion in more than one bin. In these cases, if the observed output matched one of the most probable bins, it is considered an agreement in the cross-evaluation, but is noted to be a transect with higher uncertainty than those where the most probable output had high confidence in one bin.

For the San Diego study area, the model was run and cross-validated on 194 transects. The results shown in Fig. 5a indicate that there is a very good agreement between the predicted most probable bin value and the

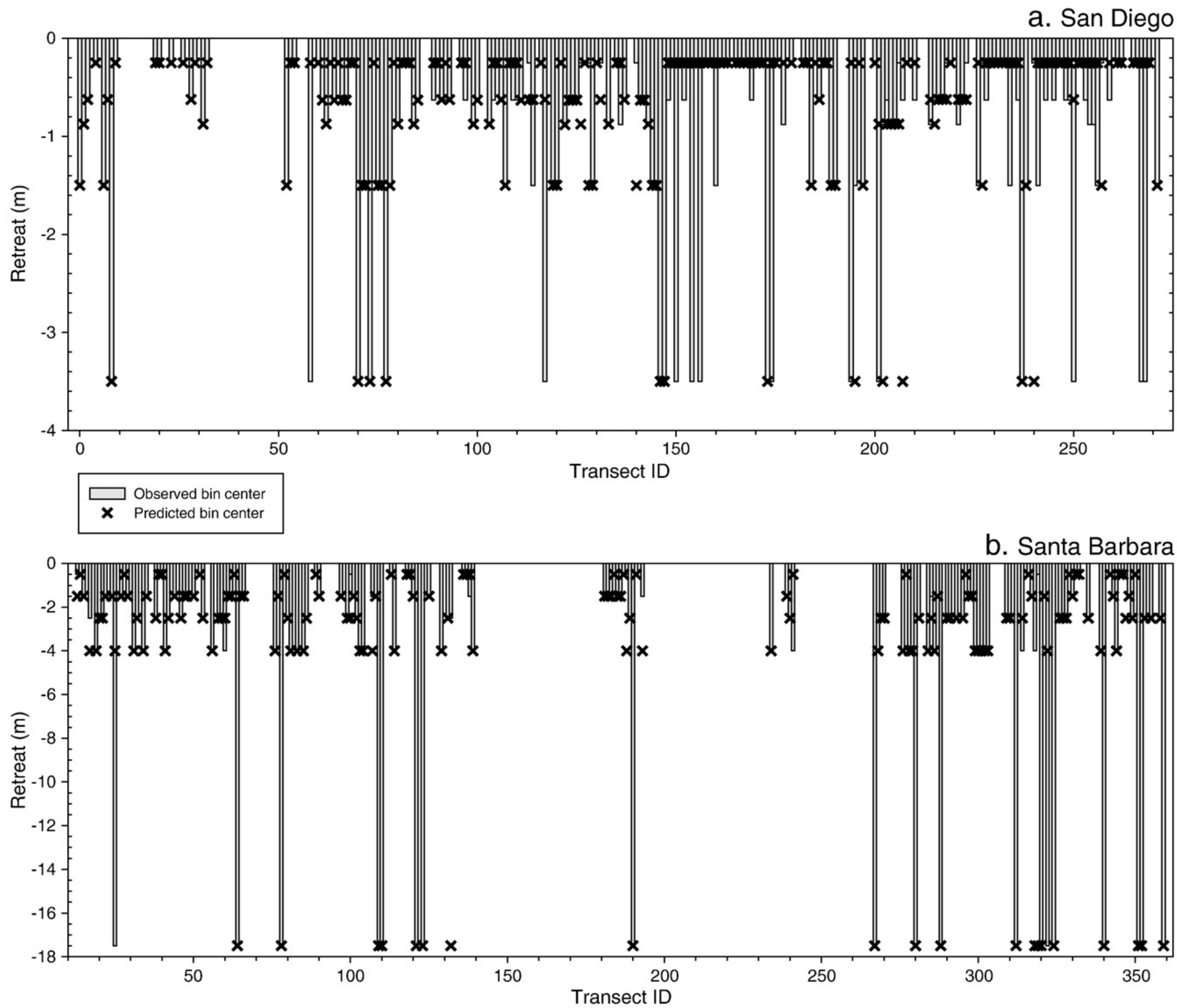

Fig. 5. Observed (bars) and predicted (x) values of cliff retreat for the (a) San Diego and (b) Santa Barbara study areas. The plotted values of the model output are the center values of the bin with the highest probability at each transect location.

observed, with correct erosion predictions at 71% (138) of the modeled locations. Uncertainties were higher at 31% of transects wherein the output predicted equally high probability in more than one bin (worst case was 3). The confidence in the majority of the predictions, however, was high.

In addition to the cross-validation of the posterior model predictions, we also examined how well the model was able to accurately predict specific locations of extreme erosion events or "hotspots". For the San Diego study area, extreme erosion was considered in any location where the cliff retreated more than 2 m in the 3-yr period of the analysis. The actual measured amount of the cliff retreat is compared with the along coast distribution of the modeled highest probabilities of the high retreat (Fig. 6a). The model predicts thirteen locations where the probability of the cliff retreat greater than 2 m is high. High probability is defined by the central 50% of the predicted probability distribution. The cliff retreat measured over the 3-yr study period indicates that 16 extreme events were observed, and that the observed and predicted were coincident at 10 locations. Since the cliff retreat was observed to exceed 2 m at only about 10% of the transects, it was a relatively rare event. Thus, a prediction of 50% probability (i.e., as likely as not) is a five-fold increase in predicted vulnerability. The observed and predicted extreme events were coincident at 10 locations. This indicates at 63% success rate of correctly identifying specific locations of extreme erosion. There were three occurrences of false positive predictions—that is, the model predicted a high probability of an extreme event but none was observed at that location. Oppositely, there were six false negatives wherein extreme erosion was observed but the model was unable to accurately predict their occurrence. The 95% confidence bounds (Fig. 6) include the false negative predictions, indicating that the potential for increased prediction error was known to be high at these locations.

In the Santa Barbara study area, the model was evaluated on 156 transects. The cross-validation of the posterior prediction compared to the observed for the 7-yr study period shows a very good agreement (Fig. 5b). The model prediction of the most probable bin is correct (as compared with the observed most probable bin) on 89% of transects. The uncertainties are high on 28% of transects on which the model predicted equally high probability in more than one bin (worst case was 2).

The assessment of the model to identify locations of the highest probability of extreme erosion resulted in the prediction of 21 extreme events (Fig. 6b), which for the Santa Barbara study area was defined as greater than 5 m of erosion over the 7-yr period of the study. Of the measured 7-yr cliff retreat, there are thirteen occurrences of extreme erosion, and all are coincident with a predicted location of high

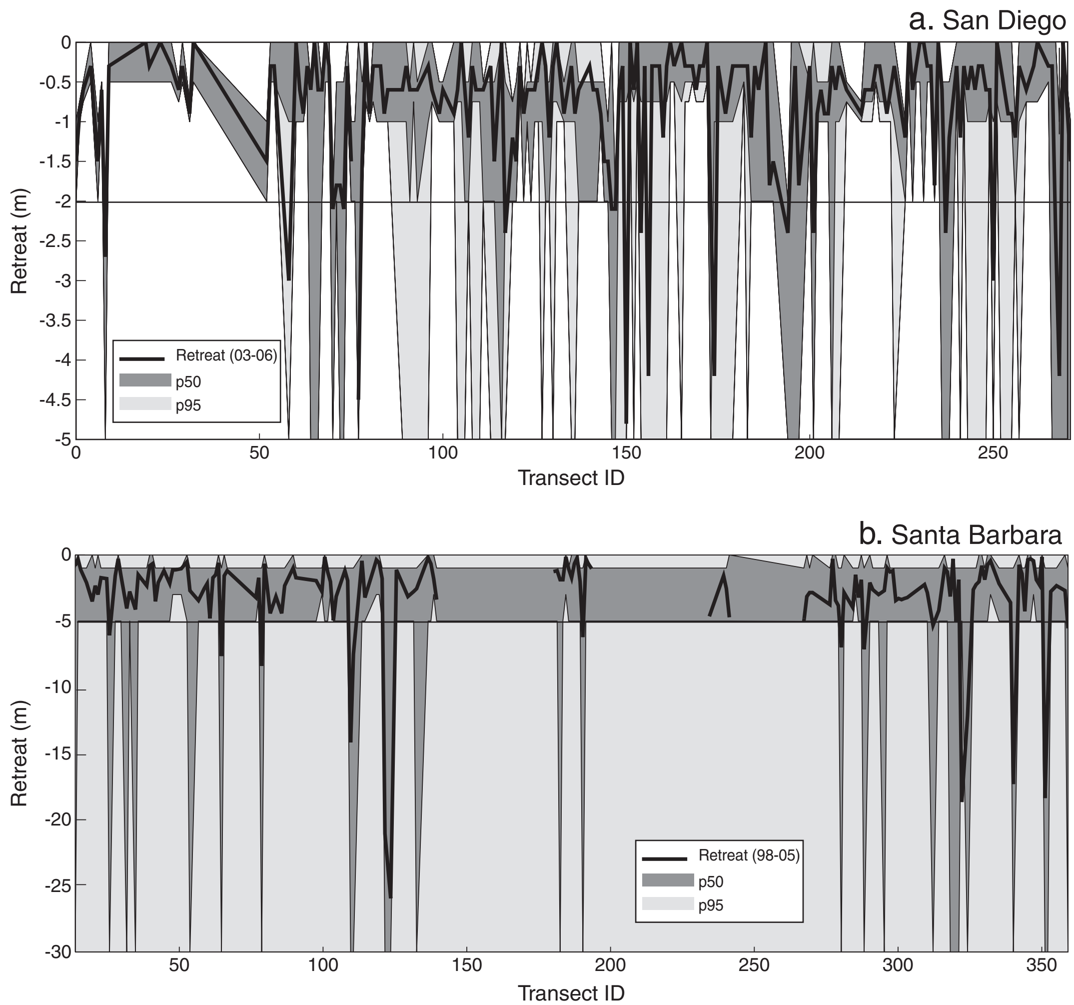

Fig. 6. Locations where model results indicate the highest probability of extreme erosion (gray areas) and measured cliff retreat (variable black line) for the a) San Diego and b) Santa Barbara study areas. Extreme erosion is defined as >2 m for San Diego and >5 m for Santa Barbara. High probability is defined by the central 50% (p50) of the predicted probability distribution. The 95% confidence bounds (p95) include the false negative predictions in the San Diego area, indicating that the potential for increased prediction error was high at these locations. The 95% confidence bounds for the Santa Barbara region span the entire range of cliff retreat scenarios, indicating that overall prediction uncertainty is higher in this region as compared to the San Diego region.

probability of extreme erosion, yielding a result of 62% accuracy in prediction. Seven of the extreme erosion predictions were not observed. However, the model did predict the correct location of every transect on which extreme erosion was measured. The 95% confidence bounds for the Santa Barbara region (Fig. 6) spanned the entire range of cliff retreat scenarios, indicating that prediction uncertainty was higher in this region as compared to the San Diego region.

5. Discussion

The very episodic nature of the cliff retreat coupled with a typical lack of long-term high-resolution time series data makes predicting cliff erosion very difficult. Additionally, the complex interplay between cliff geometry, lithology of cliff-forming materials, and long-term erosion rates and trends requires a multivariate approach for accurate forecasting. The Bayesian approach provides a statistical framework that describes what is known about the physical parameters of a system and, perhaps more importantly, what is known about the causal relationships between them.

Based on the results presented in this study, the Bayesian approach is particularly well-suited to forecasting the coastal cliff erosion. The compilation of the model parameters as geospatial data also subsequently allows for the creation of probability distribution maps to provide visual hazard information, especially useful for the management and planning communities. The Bayesian learning process, incorporating prior information and likelihood functions into a decision framework, is parallel to the adaptive management approach of managing natural resources (Ellison, 1996), which is a mode of decision-making in which results are assessed and decisions or actions are modified based on the new information that has been learned. It stands to reason then that the Bayesian approach is ideal for producing the science needed to address numerous issues associated with natural resource management in the coastal zone, especially in the context of sea-level rise and climate change scenarios. Although the model presented herein does not include extreme events for the short-term validation such as El Niño seasons or earthquakes, which have been shown to correlate to the increased cliff retreat in California (Hapke and Richmond, 2002), it is assumed that these events are recorded in the 70-yr historic cliff retreat rates that are incorporated into the model. The original design of this study is to test the application of the Bayes method to predict cliff response to moderate to severe storms, and is validated over too short a time period to incorporate sea-level rise or increased storm intensity data, both of which will need to be considered and incorporated in longer-term predictions (decades to centuries).

In this discussion, we wish to understand which processes and variables are most influential in the model predictions and why the predictions succeed at some locations and fail at others. To address these issues, a sensitivity study that varies the input variables was performed. The spatial variations in the predictive skill were inspected to identify missing information that may be responsible for poor predictions.

5.1. Sensitivity analysis

A sensitivity analysis was conducted to identify which input parameters most affect the model results, whether this varies from one study area to the other, and to understand the relative importance of the parameters to the predicted outcomes. The sensitivity values are based on a ratio of the change in the model outcome if the model input is varied (Hamby, 1994). The results (Fig. 7) are unitless ratio values and the relative values between them demonstrate that for both study areas the historical erosion rate is the most influential parameter. The relatively low sensitivity of the Santa Barbara study area to the geology parameter is likely a function of the low variation in the distribution of the input data (only 2 categories).

Tests of conditional independence were conducted to assess the dependence of the likelihood ratios on the updated distributions of individual parameters and combinations of parameters, to gauge the importance of the relationships that were user-established. The likelihood ratio (LR) is determined by:

$$LR = \sum_{i=1}^{n} \{\log\left(p\left[F_i \mid O_j\right]\right) - \log(p[F_i]) \qquad (2)$$

where $LR = 1$, the updated probability ($p[F_i|O_j]$) is not an improvement over the prior probability distribution ($p[F_i]$). LR values that fall below one indicate that the update is worse than the prior and the model has failed to make an improved prediction. LR values greater than one indicate that the updated probability is an improvement over the prior. Table 3 shows the calculated values of the LR for both individual and combined parameters. The likelihood ratios for the updated probabilities of each single parameter are either no better or just slightly better than the prior probability, demonstrating the significant limitations of developing forecasts based on single parameter updates. This approach is somewhat common in more simplistic coastal hazard forecasts, specifically the forward projection of historical erosion rates (Moore et al., 1999; Priest, 1999). The results shown in Table 3 indicate that the success of the prediction is dictated by the inclusion of multiple parameters. The likelihood ratios increase exponentially as parameters are combined and updates are generated. The likelihood ratios of the probabilities updated using all

Table 3
Likelihood ratios for the updated (posterior) probability distributions of cliff retreat for single parameter updates and multiple (combined) updates for the San Diego and Santa Barbara study areas. The values listed are exponents of 10.

	Likelihood ratio (10^x)	
	San Diego	Santa Barbara
Single parameter update		
Short-term retreat	121[a]	77[a]
Long-term retreat rate	1	1
Cliff slope	1	0
Impact hours	2	0
Geology	2	0
Cliff height	3	1
Multiple parameter update		
Long-term retreat rate + slope	4	3
Long-term retreat rate + slope + impact hours	15	9
Long-term retreat rate + slope + impact hours + height	48	22
All parameters	59	32

[a] Maximum obtainable for model.

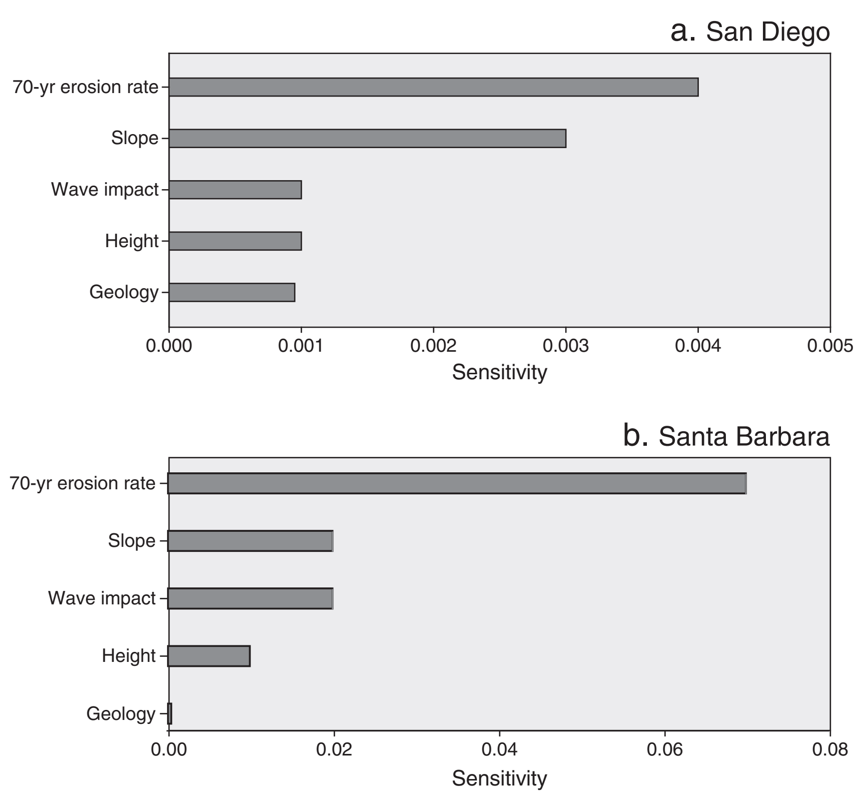

Fig. 7. Sensitivity analysis results for the two study sites (a) San Diego and (b) Santa Barbara, show that the Bayesian model is most sensitive to the long-term (70-yr) erosion rate. The analysis is used to determine how sensitive the model is to changes in the values of the prior distribution parameters and is a unitless number.

five parameters provide a prediction that is 10^{59} times better than the prior in the San Diego study area and 10^{32} times better than the prior in Santa Barbara.

5.2. Site-specific prediction skill

The results of the cross-evaluation of the model predictions for two independent stretches of cliffed coastline investigated in this study support the strength and reliability of a Bayesian approach to forecasting cliff erosion. The lower ability to accurately forecast erosion in San Diego (71%) versus Santa Barbara (89%) is likely related to the relative amount of cliff top development and engineering structures in the study areas. In general, development and structural reinforcement of the cliffs are lower in Santa Barbara than in San Diego. In a less natural (or more modified) state, the cliffs are less likely to behave based on the parameters that were input into the network, which assume a natural system. For example, a seawall may result in a section of a cliff responding as if it is composed of a stronger material than it is, in which case the model would tend to over-predict the erosion. This suggests that the addition of a parameter describing the level of modification/reinforcement is likely to be useful in improving the model results.

The Bayesian network performed relatively well in its ability to accurately identify areas with a higher probability of extreme erosion, indicating that it may be a useful tool for management applications. In locations where the observed and predicted erosion extremes were not coincident, the model either over-predicted (false positive) or under-predicted (false negative). For the cases of false positive predictions, the models forecasted an extreme event where none was observed. For hazard assessments, this situation may result in an over-preparation for an expected erosion event. Oppositely, in the locations of false negatives, the probability of extreme erosion was low but a large erosion event was observed. This situation is the "worst case" for management applications in that communities or property owners would not expect high erosion at a specific location and therefore could be caught off-guard in terms of preparation or mitigation procedures. For the Santa Barbara study area, there were no false negative predictions, which support the potential value of this method for localized hazard assessments. Overall, the predictive capability yielded similar results in both study areas (62–63%) when considering both false negative and false positive predictions, suggesting that this may be the limitation of the model given the current parameters being incorporated.

To assess whether there is an identifiable process-related or physical reason for higher uncertainties and/or poor predictions at some locations as compared with others, specific sites were examined using a digital photograph database (http://www.californiacoastline.org/). Locations were examined to look for consistency in the presence or absence of seawalls, cliff top development, basal notching, and evidence of erosion processes (block falls, rilling/gullying of cliff face, etc.). The hypothesis for this exercise was that areas with high uncertainty and poor prediction have been modified to the extent that the parameters input into the model no longer accurately represent the state of the cliff. Additionally, a site with consistent evidence of the cliff retreat related to terrestrial processes rather than the marine processes might not generate results consistent with a model forced with wave impact hours. More than seventy individual locations were inspected but no consistent patterns emerged. It is beyond the scope of this study to develop a quantitative database and include a thorough analysis of the spatial distribution of the various influencing parameters and how they correlate to the model results. This will be investigated as part of future efforts.

Other explanations for the high uncertainty/poor prediction at certain locations include inaccuracies in the construction of the network, unaccounted for uncertainties in the data, and the exclusion or over-simplified representation of an influencing input parameter. The network construction was aided by published literature on cliff erosion processes and is unlikely to contribute substantially to the levels of higher uncertainties observed in the model outputs. Unaccounted for uncertainties in the input data could result in translation of uncertainties to the model output and/or poor posterior predictions. For example, in the estimate of wave impact hours, the run-up is estimated for a specific slope and back beach elevation which do not evolve through the time period of the model; rather they are fixed in the initial state. It would require a fairly complex model to incorporate hour by hour changes in the beach geometry over periods of several years. This approach is being incorporated into an application of the model for a single storm event (Barnard et al., 2009); however, it is not practical for longer-term modeling at this time.

Some of the uncertainty in the model prediction is also likely related to the simplified input geology parameter and the lack of a terrestrial forcing parameter (rainfall/groundwater/slope wash). The geology parameter could be refined by including data on rock strength and fracture density. Some data exist and could be derived from existing technical reports and building permits, but the compilation of a complete database would require detailed and regionally extensive field mapping. A pilot study to assess how much the model uncertainty improves with the inclusion of rock strength characterization would be useful for determining if a regional mapping effort was warranted.

The contribution of terrestrial forcing in processes of coastal cliff retreat is widely recognized and, in combination with marine forcing, is held to be fundamental to the development and evolution of coastal cliffs (Trenhaile, 1987; Sunamura, 1992; Hampton and Griggs, 2004). Recent studies by Collins and Sitar (2008) and Young et al. (2009) highlight the importance of terrestrial processes. The exploration of the performance of the Bayesian model presented in this study requires consideration of the influence that a lack of a terrestrial forcing component may have on the model outcome. With recognition that some descriptor of variation in terrestrial forcing along the coast would improve the model results, our results suggest that it is unlikely that this can be the main culprit to explain the areas of poor performance. If the model output was highly sensitive to terrestrial forcing, it would be expected that it would consistently under-predict cliff erosion and the occurrence of false negative predictions for extreme erosion would be high. This is not consistently manifested in the results; for in the Santa Barbara study area there were no false negative extreme erosion predictions. In the San Diego study area, 37% of the extreme events were false negative predictions, possibly indicating that terrestrial forcing is more important in San Diego than in Santa Barbara. It is also likely that the dominant retreat process (marine versus terrestrial) is manifested in the cliff slope parameter. In general, cliffs that are dominated by terrestrial processes tend to have lower slopes and thus would be associated with lower historical cliff retreat rates and have a lower probability of failure.

6. Conclusions

The Bayesian application to forecasting coastal cliff retreat is advantageous over other predictive methods because a user-defined network is used to combine multiple parameters with information on how the parameters influence or are related to one another. This is critical for cliffed coastlines where erosion is dependent on a complex interplay of slope, height, material properties, historical behavior and driving forces. The Bayesian prediction also incorporates the uncertainties of the prior probabilities and can be updated as additional information is obtained or learned. In this study, we use a multi-parameter Bayesian network to investigate correlations between key variables that control and influence variations in cliff retreat processes.

The Bayesian network predicted the most likely outcome of cliff failure with a high degree of confidence. In terms of ranked probabilities, the method accurately predicted the correct outcome (location and amount of cliff erosion) in approximately 70–90% of the study areas. The model was also highly successful in correctly predicting where extreme erosion events (cliff retreat >2 m or >5 m) occur.

Model evaluations and sensitivity analyses indicate that information on prior behavior is crucial for accurately predicting the most likely outcome. In this study, the model performed poorly in locations lacking information on long-term erosion rate of the cliff that represents prior behavior. The assessment of likelihood functions clearly shows that a single parameter approach (i.e. projection of historical rates) is insufficient for predicting coastal cliff erosion, and the strongest improvement of the prediction is achieved when all parameters are combined.

The formulation of a Bayesian model is conceptually similar to the framework of adaptive management strategies that are gaining favor for natural resource management and as such, Bayesian modeling of the physical system is compatible with planning for the future. The Bayesian approach is ideal for producing the science needed to address numerous issues associated with natural resource management in the coastal zone, especially in the context of sea-level rise and climate change scenarios.

Acknowledgements

We are indebted to Peter Ruggiero for calculating the wave impact hour estimates used in the model. We thank Meredith Kratzmann, who spent many hours compiling critical baseline data. We are grateful to Patrick Barnard for his continued support of this effort, and to the USGS for funding. Fruitful discussions with Ben Gutierrez, Erika Lentz, Emily Himmelstoss and Elizabeth Pendleton greatly improved the design and approach of this study.

References

Amiri, A., Tabatabaei, R., 2008. Earthquake risk management strategy plan using nonparametric estimation of hazard rate. American Journal of Applied Sciences 5, 581–585.
Barnard, P.L., O'Reilly, B., van Ormondt, M., Elias, E., Ruggiero, P., Erikson, L.H., Hapke, C., Collins, B.D., Guza, R.T., Adams, P.N., Thomas, J.T., 2009. The framework of a coastal hazards model; a tool for predicting the impact of severe storms. U.S. Geological Survey Open-File Report 2009-1073, 21 pp.
Bray, M.J., Hooke, J.M., 1997. Prediction of soft-cliff retreat with accelerating sea-level rise. Journal of Coastal Research 13, 453–467.
Bruun, P., 1962. Sea level rise as a cause of shore erosion. Journal of the Waterways and Harbors Division 88, 117–130.
Cambell, K.W., 1982. Bayesian analysis of extreme earthquake occurrences, Part I. Probabilistic hazard model. Bulletin of the Seismological Society of America 72, 1689–1705.
Collins, B.D., Sitar, N., 2008. Processes of coastal bluff erosion in weakly lithified sands, Pacifica, California, CA. Geomorphology 97, 483–501.
Dahal, R.K., Hasegawa, S., Nonomura, A., Yamanaka, M., Dhakal, S., Paudyal, P., 2008. Predictive modeling of rainfall-induced landslide hazard in the Lesser Himalaya of Nepal based on weights-of-evidence. Geomorphology 102, 496–510.
Demoulin, A., Chung, C.F., 2007. Mapping landslide susceptibility from small datasets: A case study in the Pays de Herve (E Belgium). Geomorphology 89, 391–404.
Edil, T.B., Vallejo, L.E., 1980. Mechanics of coastal landslides and the influence of slope parameters. Engineering Geology 16, 83–96.
Ellison, A.M., 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. Ecological Applications 6, 1036–1046.
Emery, K.O., Kuhn, G.G., 1982. Sea cliffs: their processes, profiles and classifications. Geological Society of America Bulletin 93, 644–654.
Gritzner, M.L., Marcus, W.A., Aspinall, R., Custer, S.G., 2001. Assessing landslide potential using GIS, soil wetness modeling and topographic attributes, Payette River Idaho. Geomorphology 37, 149–165.
Hall, J.W., Meadowcroft, I.C., Lee, E.M., van Gelder, P.H., 2002. Stochastic simulation of episodic soft cliff recession. Coastal Engineering 46, 159–174.
Hamby, D.M., 1994. Environmental Monitoring and Assessment. Kluwer Academic Publishers, Netherlands.
Hampton, Griggs, 2004. Formation, evolution, and stability of coastal cliffs—status and trends. USGS Professional Paper 1683, 123 pp.
Hapke, C.J., 2005. Estimated material yield from coastal landslides based on historical digital terrain modeling, Big Sur, California. Earth Surface Landforms and Processes 30, 679–697.
Hapke, C.J., Reid, D., 2007. The National assessment of shoreline change: Part 4, historical coastal cliff retreat along the California coast. U.S. Geological Survey Open-file Report 2007-1133.
Hapke, C.J., Richmond, B., 2002. The impact of climatic and seismic events on the short-term evolution of seacliffs based on 3-D mapping: Northern Monterey Bay, California. Marine Geology 187 (3–4), 259–278.
Hapke, C.J., Malone, S., Kratzmann, M., 2009. National assessment of historical shoreline change: a pilot study of historical coastal bluff retreat in the Great Lakes, Erie, Pennsylvania. U.S. Geological Survey Open-File Report 2009-1042, 25 pp.
Kennedy, M.P., Tan, S.S., 2008. Geologic map of the San Diego 30′×60′ quadrangle, California: California Geological Survey, Regional Geologic Map No. 3, scale 1:100000.
Kennedy, M.P., Tan, S.S., Bovard, K.R., Alvarez, R.M., Watson, M.J., Gutierrez, C.I., 2007. Geologic map of the Oceanside 30×60-minute quadrangle, California. California Geological Survey, Regional Geologic Map No. 2, scale 1:100000.
Komar, P.D., McDougal, W.G., Marra, J.J., Ruggiero, P., 1999. The rational analysis of setback distances: applications to the Oregon Coast. Shore and Beach 67, 41–49.
Lee, E.M., Hall, J.W., Meadowcroft, 2001. Coastal cliff recession: the use of probabilistic prediction methods. Geomorphology 40, 253–269.
Lee, S., Choi, J., Min, K., 2002. Landslide susceptibility analysis and verification using the Bayesian probability model. Environmental Geology 43, 120–131.
Miller, S., Harris, N., Williams, L., Bhalai, S., 2007. Landslide susceptibility assessment for St. Thomas, Jamaica, using geographical information system and remote sensing methods. In: Teeuw, R.M. (Ed.), Mapping Hazardous Terrain using Remote Sensing. Geological Society, London, pp. 77–91.
Minor, S.A., Kellog, K.S., Stanley, R.G., Stone, P., Powell, II, C.L. Gurrola, L.D., Setling, A.J., Brandt, T.R., 2002. Preliminary geologic map of the Santa Barbara Coastal Plain Area. U.S. Geological Survey Scientific Investigations Map 3001, scale 1:25,000.
Moore, L.J., Benumof, B.T., Griggs, G.B., 1999. Coastal erosion hazards in Santa Cruz and San Diego Counties, California. Journal Coastal Research Special Issue 28, 121–139.
Oh, C.K., Beck, J.L., Yamada, M., 2008. Bayesian learning using automatic relevance determination prior with an application to earthquake early warning. Journal of Engineering Mechanics 134, 1013–1020.
Plant, N.G., Holland, K.T., in press. Prediction and assimilation of surfzone processes using a Bayesian network — part I: forward models. Coastal Engineering.
Priest, G.R., 1999. Coastal shoreline change study Northern and Central Lincoln County, Oregon. Journal Coastal Research Special Issue 28, 140–157.
Ruggiero, P., Komar, P.D., McDougal, W.G., Beach, R.A., 1996. Extreme water levels, wave runup and coastal erosion. Proceedings of the 25th International Conference on Coastal Engineering. ASCE, pp. 2793–2805.
Ruggiero, P., Komar, P.D., McDougal, W.G., Marra, J.J., Beach, R.A., 2001. Wave runup, extreme water levels, and the erosion of properties backing beaches. Journal of Coastal Research 17, 401–419.
Stavrakakis, G.N., Drakopoulos, J., 1995. Bayesian probabilities of earthquake occurrences in Greece and surrounding areas. Pure and Applied Geophysics 144, 307–319.
Stockdon, H.F., Holman, R.A., Howd, P.A., Sallenger, A.H., 2006. Empirical parameterization of setup, swash, and runup. Coastal Engineering 53, 573–588.
Sunamura, T., 1982. A predictive model for wave-induced cliff erosion, with application to Pacific Coasts of Japan. Journal of Geology 90, 167–178.
Sunamura, T., 1983. Processes of sea cliff and platform erosion. In: Komar, P.D. (Ed.), CRC Handbook of Coastal Processes and Erosion. CRC Press, Boca Raton, FL, pp. 233–265.
Sunamura, T., 1992. Geomorphology of Rocky Coasts. John Wiley and Sons, New York.
Thieler, E.R., Himmelstoss, E.A., Zichichi, J.L., Ergul, A., 2009. Digital Shoreline Analysis System (DSAS) version 4.0—an ArcGIS extension for calculating shoreline change. U.S. Geological Survey Open-File Report 2008-1278.
Trenhaile, A.S., 1987. The Geomorphology of Rock Coasts. Oxford University Press, New York.
Young, A.P., Guza, R.T., Flick, R.E., O'Reilly, W.C., Gutierrez, R., 2009. Rain, waves, and short-term evolution of composite seacliffs in southern California. Marine Geology 267, 1–7.