



Modeling storms improves estimates of long-term shoreline change

L. Neil Frazer,¹ Tiffany R. Anderson,¹ and Charles H. Fletcher¹

Received 14 July 2009; revised 11 September 2009; accepted 29 September 2009; published 31 October 2009.

[1] Large storms make it difficult to extract the long-term trend of erosion or accretion from shoreline position data. Here we make storms part of the shoreline change model by means of a storm function. The data determine storm amplitudes and the rate at which the shoreline recovers from storms. Historical shoreline data are temporally sparse, and inclusion of all storms in one model over-fits the data, but a probability-weighted average model shows effects from all storms, illustrating how model averaging incorporates information from good models that might otherwise have been discarded as un-parsimonious. Data from Cotton Patch Hill, DE, yield a long-term shoreline loss rate of 0.49 ± 0.01 m/yr, about 16% less than published estimates. A minimum loss rate of 0.34 ± 0.01 m/yr is given by a model containing the 1929, 1962 and 1992 storms.

Citation: Frazer, L. N., T. R. Anderson, and C. H. Fletcher (2009), Modeling storms improves estimates of long-term shoreline change, *Geophys. Res. Lett.*, 36, L20404, doi:10.1029/2009GL040061.

1. Introduction

[2] Shoreline change models use historical shorelines to estimate the rate of change, or to predict a future shoreline, as in Figure 1a. Storm influenced shorelines raise difficult questions in this regard because shorelines tend to recover from storms [Birkemeier, 1979; Kriebel, 1987; Morton, 1988; Morton *et al.*, 1994]. After a storm, shoreline change rate may return to its long-term trend [Galvano and Douglas, 2000; Zhang *et al.*, 2002] over 5–15 years, depending on the magnitude of the storm. Should one therefore remove storm-influenced shorelines from the data when attempting to estimate the long-term trend? If so, how can one tell whether a shoreline is truly storm influenced? Here we address such questions by explicitly modeling the transient part of the storm-induced change. The storm-driven permanent change is implicitly modeled as part of the long term trend.

[3] In modeling long-term shoreline change, data contaminated by large storms violate the least-squares assumption that errors are normally distributed. One alternative to least squares is to objectively edit storm effects from the data using methods such as least median of squares (LMS), sometimes referred to as re-weighted least squares (RLS) [Rousseuw, 1984; Rousseuw and Leroy, 1987; Genz *et al.*, 2007]. Another alternative is to leave storms in the data and fit the model by minimizing absolute differences rather than squared differences (LAD) [Rousseuw and Leroy, 1987;

Genz *et al.*, 2007]. Douglas and Crowell [2000] removed storm points from the data until the misfit (average residual) was comparable to the standard error of the measurements plus 20% of beach width. They illustrated their procedure using data from Cotton Patch Hill, Delaware, which was hit by large storms in 1929, 1962, 1991 and 1992. We illustrate our procedure with the same Cotton Patch Hill data analyzed by Douglas *et al.* [1998] and Douglas and Crowell [2000, Table 1]. The 1991 and 1992 storms occurred between surveys in 1990 and 1993. The 1991 storm was smaller than the 1992 storm, and we model the effects of these two storms using a single storm function with onset in 1992.

2. One-Dimensional Models With Storms

[4] A storm is defined here as any event that changes the position of the shoreline suddenly, with subsequent slow recovery toward the long term trajectory. Although seasonal changes in shoreline can be rapid, recovery usually occurs within a time much shorter than the time between historical shoreline surveys, so we regard uncorrected seasonal effects as part of the noise.

[5] The traditional 1D model for shoreline change is

$$y(t) = b + rt + n(t), \quad (1)$$

in which y is the cross-shore coordinate (shoreline location), b is the intercept, r is the long-term shoreline change rate, and n is uncorrelated noise with zero mean. The inverse problem for shoreline change is to infer the change rate from a time series of shoreline data $y(t_1), y(t_2), \dots, y(t_j)$. The intercept depends on the baseline used to measure y , and on the time origin, both of which can be adjusted to condition the solution of the inverse problem. In mathematical parlance, the linear 1D model is a linear sum of the basis functions 1 and t . If acceleration is included in the model, there is a third basis function $\frac{1}{2}at^2$, and Fenster *et al.* [1993] showed that models with acceleration can be more parsimonious than simple linear models. However, Crowell *et al.* [1997] showed that the acceleration term is good at fitting noise and that its inclusion in a model can lead to inaccurate predictions. This situation is a reminder of the importance of prior information in any inversion problem and that parsimony by itself does not select a good model unless candidate models are suitably chosen.

[6] Suppose that the data area is struck by a storm at time t_s . We use a basis function called the storm function, given by

$$e_+^{-\gamma(t-t_s)} = \begin{cases} e^{-\gamma(t-t_s)} & \text{if } t \geq t_s \\ 0 & \text{if } t < t_s, \end{cases} \quad (2)$$

¹Department of Geology and Geophysics, School of Ocean and Earth Science and Technology, University of Hawaii, Honolulu, Hawaii, USA.

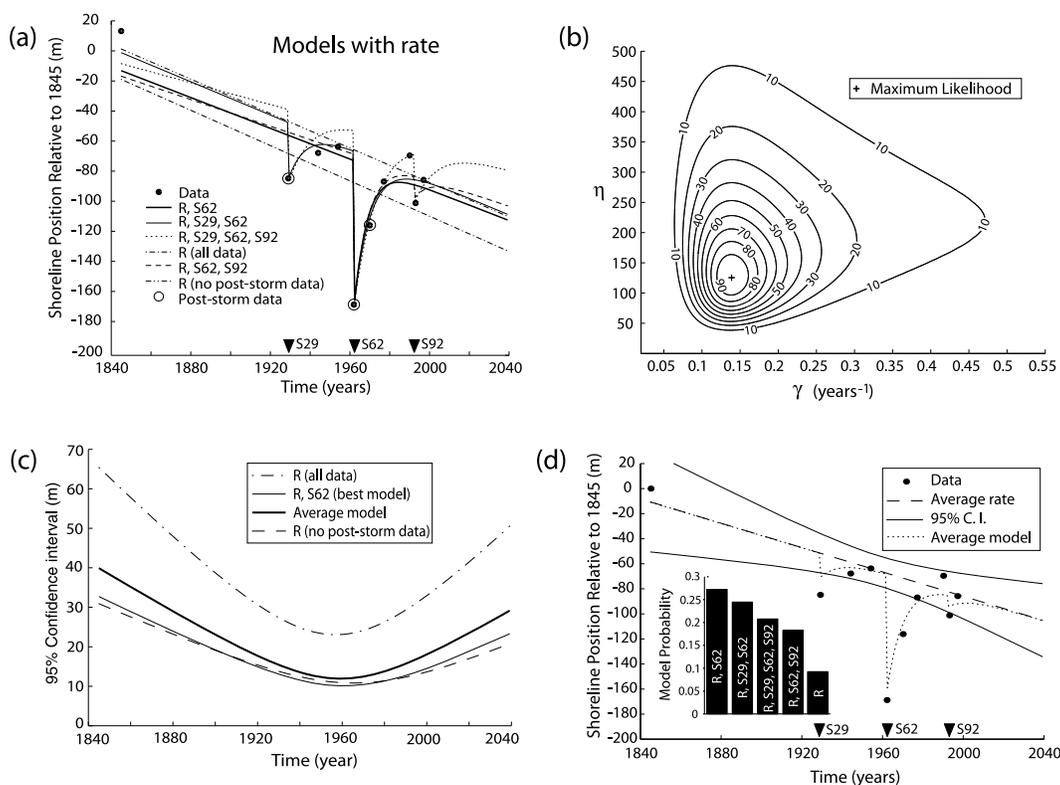


Figure 1. (a) Shoreline data from Cotton Patch Hill, DE, with fits from various models. The circled data points are post-storm shorelines. All models had an intercept, a long-term rate (R), and some models had one or more storms (1929, 1962, 1992). The model with lowest residuals (dotted line) had all three storms. The model with the lowest AICc value (solid line) had only the 1962 storm. The $---$ model had only rate and intercept, but post-storm data were not used. (b) Profile likelihood of storm recovery rate and process noise for the shoreline model with only the 1962 storm. The likelihood is scaled to have maximum value 100, and the MLE is at $\eta = (11.2 \text{ m})^2$, $\gamma^{-1} = 7.2 \text{ y}$. (c) Confidence intervals (95%) for shoreline position based on long-term rate for various models. Except for model R , the confidence intervals are conditioned on the absence of storms. (d) The probability-weighted model average (dotted line), the predicted shoreline position (dashed) using only the rate and intercept from the average model, and 95% confidence interval (solid lines). The inset shows the probabilities of the component models.

in which $\gamma \geq 0$ is the recovery rate (the inverse time scale of recovery). The augmented model for shoreline change is now

$$y(t) = b + rt + se_+^{-\gamma(t-t_s)} + n(t) \quad (3)$$

in which s is the storm amplitude parameter. Since the time of the storm is known, model (3) has two more parameters than model (1). It is linear in b , r and s , but nonlinear in γ . The fact that the storm function is discontinuous at t_s does not make it more difficult to use, because basis functions are not required to be continuous, only independent of each other. Versions of (3) implemented with different storms or combinations of storms are regarded as different models, and we use an information criterion (below) to rate the relative goodness of models. Most historical data sets contain only 6–10 shorelines, and the information criterion usually excludes models in which each storm has its own recovery rate; therefore we use the same recovery rate parameter for all storms.

[7] If some storm-induced shoreline change is permanent, an appropriate model would be (3) plus a function $s_p H(t - t_s)$ in which H is the unit step function and s_p is the amplitude of the permanent change. In temporally sparse historical data

with multiple storms we find that such step functions can trade off with both our storm function and the rate function to such an extent that a model with only step functions fits the data fairly well. Admitting models consisting only of step functions replaces the problem of estimating a long-term rate parameter with the problem of estimating frequencies and amplitudes of storms. Moreover, models consisting only of step functions ignore the abundant evidence that storm-altered shorelines do recover to a large extent. Accordingly, for historical data we explicitly model only the transient part of the storm, leaving the permanent part as a component of long-term trend. Although it is not needed for this paper, beach nourishment can be modeled like a storm. For nourishment that alters a shoreline the storm function is used, but for offshore nourishment, which does not immediately alter a shoreline, we use the function $e_+^{-\gamma(t-t_n)}(1 - e_+^{-\gamma(t-t_n)})$, where t_n is the time of nourishment.

3. Linearization

[8] As the model is nonlinear in γ , we find γ by maximizing the profile likelihood [e.g., Coles, 2001, p. 34]. In order to include the effects of uncertainty in γ we then linearize the model in a neighborhood of the maximum

Table 1. Process Error for Models

Model	Process Error (m)
Averaged model	12.5
R	30.9
R, S29	30.4
R, S62	11.2
R, S92	30.2
R, S29, S62	8.2
R, S29, S92	29.9
R, S62, S92	9.8
R, S29, S62, S92	0.1
R (no post-storm data)	9.5

likelihood estimate (MLE) $\hat{\gamma}$. Since recovery rate is necessarily positive, and our noise model is Gaussian, we use $\mu = \ln \gamma - \ln \hat{\gamma}$ as a parameter in the linearized model. As $\partial e^{\gamma t} / \partial \ln \gamma = \gamma t e^{\gamma t}$, the linearized model is

$$y(t) = b + rt + \sum_{k=1}^{K_s} s_k e_+^{-\hat{\gamma}(t-t_{sk})} + \mu \sum_{k=1}^{K_s} (t - t_{sk}) \hat{\gamma} \hat{s}_k e_+^{-\hat{\gamma}(t-t_{sk})} + n(t), \quad (4)$$

in which s_k is the amplitude of the k^{th} storm, K_s is the number of storms, and the coefficient of μ is a single basis function.

4. Noise and Recovery Rate

[9] We model the noise $n(t)$ as a zero-mean Gaussian. We assume the noise consists of observational noise (measurement noise) and process noise, and that the two noise processes are unrelated. Observational noise is estimated prior to modeling [Crowell *et al.*, 1991; Douglas and Crowell, 2000; Fletcher *et al.*, 2003; Genz *et al.*, 2007]. Douglas and Crowell [2000] estimated the uncertainty in the high water line at Cotton Patch Hill as 6.5m, for a process noise variance of $(6.5 \text{ m})^2$, but here we estimate process noise from the data. Our data covariance matrix has the form

$$C_y = \langle nm^T \rangle = C_o + \eta C_p, \quad (5)$$

in which C_o is the covariance matrix of measurement noise, C_p is an unscaled covariance matrix of process noise, and η is a scaling parameter to be estimated from the data. We assume that observational noise at one time is uncorrelated with observational noise at other times, so C_o is diagonal. Observational error ranged from 2.6 m in 1997 to 8.9 m in 1845 (Table A1 of section A in Text S1 of the auxiliary material¹).

[10] Process noise should be correlated, as white noise convolved with a storm function gives a covariance matrix $C_p(i, j) = (\gamma/2)\exp(-\gamma|t_i - t_j|)$, but our experiments suggest that γ is poorly resolved by the residuals in historical data sets because the data are too sparse. In the numerical calculations presented here we take C_p to be the identity matrix (as did Douglas and Crowell [2000]). Table 1 gives

the process error for the models of this paper. The best-fit model (R,S62) has a process error of 11.2 m that is roughly 35% of the beach width. The three-storm model with the 0.1 m process error over-fits the data.

[11] Our likelihood function is the usual Gaussian,

$$\mathcal{L}(m|y) = (2\pi)^{-N/2} |C_y|^{-1/2} \exp\left[-\frac{1}{2}(y - Gm)^T C_y^{-1} (y - Gm)\right], \quad (6)$$

in which G is the system matrix (design matrix, configuration matrix), and m is the column vector of model coefficients. For the model of equation (4), the parameter vector is $m = [b, r, s_1, s_2, \dots, \mu]^T$, and the columns of G are the basis functions evaluated at each survey time. The first column of G is all ones, and the second column is $[t_1, t_2, \dots, t_N]^T$, and we condition G by removing the mean from all columns after the first. Maximizing the likelihood with respect to the parameter vector m gives the usual relation

$$0 = G^T \hat{C}_y^{-1} (y - G\hat{m}), \quad (7)$$

and maximizing the likelihood with respect to the noise parameter η gives the nonlinear relation

$$0 = \text{tr}\left\{\hat{C}_y^{-1} C_p\right\} - (y - G\hat{m})^T \hat{C}_y^{-1} C_p \hat{C}_y^{-1} (y - G\hat{m}). \quad (8)$$

The noise parameter enters both these equations through the definition $C_y = C_o + \eta C_p$. We find the MLE $\hat{\eta}$ by a 1D search: pick a value of η ; compute $m(\eta)$ as the solution of (7) and substitute it into (8). The value of η satisfying (8) is the MLE $\hat{\eta}$, and $m(\hat{\eta})$ is the MLE \hat{m} . In practice, since the recovery rate parameter also requires a search, we find both parameters by maximizing the profile likelihood with respect to γ and η , as shown in Figure 1b. If a prior distribution were available for γ and η we would multiply it times the profile likelihood to obtain a posterior profile. In temporally sparse data sets with early storms we find that storm recovery rate γ can trade off with long-term rate r . To minimize this effect, we estimate γ separately for each model, then fix γ at its model probability-weighted average, $\hat{\gamma}$, then use that average γ with every model. For the Cotton Patch Hill data, γ^{-1} ranges from 7.2–12.1 y, and $(\hat{\gamma})^{-1} = 8.4$ y.

[12] Equation (7) leads to the definition of a generalized inverse matrix G^{-g} such that $\hat{m}G^{-g}y$. Thus $G^{-g} = (G^T C_y^{-1} G)^{-1} G^T C_y^{-1}$. If C_o were zero, the noise parameter $\hat{\eta}$ would cancel out of the expression for G^{-g} , and the parameter covariance matrix would be given by the usual formula $C_m = G^{-g} C_y (G^{-g})^T$. For our noise model, the noise parameter does not cancel out of the expression for G^{-g} , and thus a data variation δy causes a corresponding variation in G^{-g} as well as in \hat{m} . The parameter covariance matrix is derived in the auxiliary material.

5. Prediction

[13] To predict the shoreline location at time t , we use the linearized model formula (4). It is helpful to express this as $y(t) = q^T \hat{m}$ where q is a column vector—we refer to it as a prediction kernel—containing the value of each basis func-

¹Auxiliary materials are available in the HTML. doi:10.1029/2009GL040061.

tion at time t . For a prediction of the long-term rate, the prediction kernel is just $q = [0, 1, 0, \dots, 0]^T$. For a prediction of the actual rate at time t , the elements of q are the time derivatives of the basis functions in (4). For example, suppose we want only the component of shoreline displacement due to the first storm, at time t . The first storm involves the parameters s_1 and μ , and the prediction kernel is

$$q = \left[0, 0, e_+^{-\hat{\gamma}(t-t_1)}, 0, \dots, 0, (t-t_1)\hat{s}_1\hat{\gamma}e_+^{-\hat{\gamma}(t-t_1)} \right]^T. \quad (9)$$

Here one might guess that the last element of q could be replaced by zero, since $\hat{\mu}$ is always zero at the MLE. However, the last element in q contributes to the variance by coupling the uncertainty in μ to the uncertainty of the prediction. In each case, the variance of the prediction is given by the scalar $q^T C_m q$ where C_m is the parameter covariance matrix given in the auxiliary material. Figure 1c shows the 95% confidence interval for shoreline predicted with several models.

6. Information Criteria

[14] An information criterion (IC) is a function whose value increases with the sum of squared residuals and with the number of model parameters (model complexity). The best model is the one with the lowest IC value. The use of an IC prevents over-fitting data with too many storms, but it is not a panacea, since the choice of basis functions affects the performance of the IC [McQuarrie and Tsai, 1998]. In a case where the true basis functions are included in the candidate basis functions, an IC that picks the true basis functions with probability 1 as the number of data approaches infinity is said to be consistent. In the case where at least some of the true basis functions are missing from the set of candidate basis functions, an IC that picks the combination of basis functions that best approximates the true model is said to be asymptotically efficient. The corrected Akaike Information Criterion (AICc) used here is asymptotically efficient [McQuarrie and Tsai, 1998].

[15] An important feature of any information criterion I is that for any positive numbers a and b , the quantity $a + bI$ takes its minimum at the same model as I and is therefore an equally good information criterion. The AICc formula of this paper is

$$\text{AICc} = \hat{\phi} + 2K/(N - K - 1) \quad (10)$$

in which N is the number of data points, K is the number of model parameters, and $\hat{\phi} = \hat{L}/N - 1 - \log(2\pi)$, where \hat{L} is -2 times the logarithm of the maximum likelihood. The second term in equation (10) is sometimes referred to as the complexity penalty or simply the penalty. The constant addends in the definition of $\hat{\phi}$ make our AICc formula agree with the formula found in most books when our noise model is simplified to the usual noise model. For the noise model of this paper (equation (5)), $\hat{\phi}$ is given by

$$\hat{\phi} = N^{-1} \left\{ \log |\hat{C}_y| + (d - G\hat{m})^T \hat{C}_y^{-1} (d - G\hat{m}) \right\} - 1, \quad (11)$$

in which \hat{m} is the MLE of the parameter vector, and $\hat{C}_y = \hat{\eta}C_p + C_o$ where $\hat{\eta}$ is the MLE of the noise parameter η ,

and $|\cdot|$ indicates a determinant. If $C_o = 0$, the expression for $\hat{\phi}$ simplifies to

$$\hat{\phi} = \log \left[|C_p|^{1/N} N^{-1} (d - G\hat{m})^T C_p^{-1} (d - G\hat{m}) \right] \quad (12)$$

which is independent of $\hat{\eta}$. If $C_o = 0$ and C_p is proportional to the identity matrix, $\hat{\phi}$ simplifies to

$$\hat{\phi} = \log \left[N^{-1} (d - G\hat{m})^T (d - G\hat{m}) \right] \quad (13)$$

which is the expression found in the book by Burnham and Anderson [2002], and many papers.

[16] Usually the parameter count K is equal to the number of basis functions plus one (for the variance of the noise), but if one or more basis functions contain the recovery rate γ , it must be included in the count. As we are interested mainly in long-term rate r , we do not count intercept as a parameter. (Notice that shifting all the data points by a fixed amount does not change the estimated long-term rate.)

7. Model Likelihood and Model Average

[17] The number of possible models included in equation (4) is 2^{3+K_s} , but we exclude all models without rate or intercept. As several models have similar IC scores we average models based on their prior probability and IC weights [e.g., Burnham and Anderson, 2002, p. 75], referred to here as IC likelihood. We omit model selection error [Buckland et al., 1997] for consistency with methods utilizing only rate and intercept. Our method is related to Bayesian model averaging [Hoeting et al., 1999], but is thought to be less computationally intensive.

[18] In order to define an IC likelihood, it is numerically prudent to first subtract the IC score of the best model from all the other models. Each model thus has a delta-IC given by $\Delta_j = \text{IC}_j - \min_i(\text{IC}_i)$. The IC likelihood of the j^{th} model is then given by

$$w_j = \frac{\exp\left(-\frac{1}{2}\Delta_j\right)}{\sum_j \exp\left(-\frac{1}{2}\Delta_j\right)}. \quad (14)$$

The IC-likelihoods sum to 1, and are interpreted as model likelihoods conditional on the data used to compute the IC scores. We incorporate prior information about model probabilities using the probability calculus of Tarantola and Valette [1982]. Let π_j be the prior probability of model j and μ_j be the non-informative probability of model j . The posterior probability of model j is then

$$p_j = \frac{w_j \pi_j / \mu_j}{\sum_j w_j \pi_j / \mu_j}. \quad (15)$$

If one has no prior information about various models, $\pi_j = \mu_j$, and so $p_j = w_j$. Here we take the non-informative probability to be uniform, so μ_j is the same for each j . (Even when prior probabilities are uniform, the π_j are useful. For example, to compute the average of models that do not include a particular storm, set $\pi_j = 0$ for each model

Table 2. Rates of Shoreline Loss for Various Models, With Their Prior Probabilities, IC-Likelihoods and Posterior Probabilities^a

Model	Rate \pm std (m/yr)	Δ AICc	IC Likelihood	Model Prior	Model Probability
Averaged model	-0.49 ± 0.01	-	-	-	-
R	-0.59 ± 0.06	0.78	0.13	0.11	0.09
R, S29	-0.62 ± 0.06	1.77	0.08	0.00	0.00
R, S62	-0.51 ± 0.01	0.00	0.20	0.22	0.27
R, S92	-0.69 ± 0.07	1.72	0.08	0.00	0.00
R, S29, S62	-0.55 ± 0.01	0.22	0.18	0.22	0.24
R, S29, S92	-0.69 ± 0.07	2.62	0.05	0.00	0.00
R, S62, S92	-0.44 ± 0.02	0.80	0.13	0.22	0.18
R, S29, S62, S92	-0.34 ± 0.01	0.54	0.15	0.22	0.21
R (no post-storm data)	-0.57 ± 0.01	-	-	-	-

^a‘R’ indicates that the model has a long-term rate. ‘S62’ indicates that the 1962 storm is included in the model, and similarly for other storms. The low-IC model (R,S62) is also the preferred model in an F-test.

containing that storm.) The 1962 storm has a storm erosion potential index three times greater than other storms [Zhang *et al.*, 2001, Figure 8], and it is prominent in the Cotton Patch Hill data. Accordingly, we give a prior probability of zero to each model that does not include the 1962 storm. We give the model with no storms, only long-term rate, a prior probability half as large as that of models that include the 1962 storm. This is conservative with regard to the uncertainty because the model with no storm has the highest residuals, and one could reasonably exclude all models with no storm. The model priors, likelihoods and posterior probabilities are given in Table 2.

[19] To see how model-averaging affects predictions, let ϕ be the quantity whose value is to be predicted. The model-averaged ϕ is given by

$$\hat{\phi} = \sum_j p_j \phi_j = \sum_j p_j q_j^T \hat{m}_j, \quad (16)$$

where q_j^T is the prediction kernel and \hat{m}_j is the MLE of the parameter vector for model j .

[20] As the model probabilities p_j depend on the data, the formula for $\sigma_{\hat{\phi}}^2$ requires some care and is derived in the auxiliary material. Figure 1d shows the model average and the probabilities of its component models. Although the model with three storms is not the model with the highest probability, it gives by far the best fit to the data, as shown by its low process error in Table 1; it is interesting and desirable that the average model also shows the effects of all three storms.

8. Discussion and Conclusions

[21] As the times of large storms are known, their effects can be incorporated into models of historical shoreline change by use of the storm function, a one-sided exponential with delay. Parsimony in the form of an IC prevents over-fitting the data by inclusion of too many storms, and model averaging is an objective way of reconciling competing models. Subjectivity is explicit in the form of a prior probability for models. The method may have some advantages over other methods, such as least absolute deviations, because it gives a more precise estimate of long-term rate, as well as information about the magnitude of storms (auxiliary material). At Cotton Patch Hill, DE, the minimum long-term rate of shoreline loss is 0.34 ± 0.01 m/y (from a model with all three storms). The model-averaged rate,

0.49 ± 0.01 m/y, is about 16% lower than earlier estimates. The sudden shoreline loss associated with the 1929, 1962 and 1992 storms was 19.4 ± 7.9 m, 94.8 ± 11.7 m and 9.6 ± 6.5 m, respectively. Here we outlined and solved the 1D problem, which is fundamental in shoreline change studies. Our solution to the 2D problem uses the methods of this paper to model the temporal coefficients of alongshore basis functions [Frazer *et al.*, 2009] and will be presented separately.

[22] **Acknowledgments.** Funding for this study was provided by the U.S. Geological Survey, the University of Hawaii Sea Grant College and the Hawaii Department of Land and Natural Resources.

References

- Birkemeier, W. A. (1979), The effects of the 19 December 1977 coastal storm on beaches in North Carolina and New Jersey, *Shore Beach*, *47*, 7–15.
- Buckland, S. T., K. P. Burnham, and N. H. Augustin (1997), Model selection: An integral part of inference, *Biometrics*, *53*, 603–618, doi:10.2307/2533961.
- Burnham, K. P., and D. R. Anderson (2002), *Model Selection and Multi-model Inference*, 2nd ed., Springer, New York.
- Coles, S. (2001), *An Introduction to Statistical Modeling of Extreme Values*, Springer, New York.
- Crowell, M., S. P. Leatherman, and M. K. Buckley (1991), Historical shoreline change: Error analysis and mapping accuracy, *J. Coastal Res.*, *7*, 839–852.
- Crowell, M., B. C. Douglas, and S. P. Leatherman (1997), On forecasting future U.S. shoreline positions: A test of algorithms, *J. Coastal Res.*, *13*, 1245–1255.
- Douglas, B. C., and M. Crowell (2000), Long-term shoreline position prediction and error propagation, *J. Coastal Res.*, *16*, 145–152.
- Douglas, B. C., M. Crowell, and S. P. Leatherman (1998), Considerations for shoreline position prediction, *J. Coastal Res.*, *14*, 1025–1033.
- Fenster, M. S., R. Dolan, and J. F. Elder (1993), A new method of predicting shoreline positions from historical data, *J. Coastal Res.*, *9*, 147–171.
- Fletcher, C. H., J. J. B. Rooney, M. Barbee, S.-C. Lim, and B. M. Richmond (2003), Mapping shoreline change using digital orthophotogrammetry on Maui, Hawaii, *J. Coastal Res.*, *38*, 106–124.
- Frazer, L. N., A. S. Genz, and C. H. Fletcher (2009), Toward parsimony in shoreline change prediction (I): Methods, *J. Coastal Res.*, *25*, 366–379, doi:10.2112/06-0756.1.
- Galgano, F. A., and B. C. Douglas (2000), Shoreline position prediction: Methods and errors, *Environ. Geosci.*, *7*, 23–31, doi:10.1046/j.1526-0984.2000.71006.x.
- Genz, A. S., C. H. Fletcher, R. A. Dunn, L. N. Frazer, and J. J. Rooney (2007), The predictive accuracy of shoreline change rate methods and alongshore beach variation on Maui, Hawaii, *J. Coastal Res.*, *23*, 87–105, doi:10.2112/05-0521.1.
- Hoeting, J. A., D. Madigan, A. E. Raftery, and C. T. Volinsky (1999), Bayesian model averaging: A tutorial, *Stat. Sci.*, *14*, 382–417, doi:10.1214/ss/1009212519.
- Kriebel, D. L. (1987), Beach recovery following Hurricane Elena, *Coastal Sediments*, *87*, 990–1005.

- McQuarrie, A. D. R., and C.-L. Tsai (1998), *Regression and Time Series Model Selection*, World Sci., Singapore.
- Morton, R. A. (1988), Nearshore responses to great storms, in *Sedimentologic Consequences of Convulsive Geologic Events*, edited by H. E. Clifton, pp. 7–22, Geol. Soc. Am., Boulder, Colo.
- Morton, R. A., J. G. Paine, and J. G. Gibeau (1994), Stages and durations of post-storm beach recovery, southeastern Texas coast, USA, *J. Coastal Res.*, *10*, 884–908.
- Rousseuw, P. J. (1984), Least median of squares regression, *J. Am. Stat. Assoc.*, *79*, 871–880, doi:10.2307/2288718.
- Rousseuw, P. J., and A. M. Leroy (1987), *Robust Regression and Outlier Detection*, John Wiley, New York.
- Tarantola, A., and B. Valette (1982), Inverse problems = quest for information, *J. Geophys.*, *50*, 159–170.
- Zhang, K., B. Douglas, and S. Leatherman (2001), Beach erosion potential for severe nor'easters, *J. Coastal Res.*, *17*, 309–321.
- Zhang, K., B. Douglas, and S. Leatherman (2002), Do storms cause long-term beach erosion along the U.S. east barrier coast?, *J. Geol.*, *110*, 493–502, doi:10.1086/340633.
-
- T. R. Anderson, C. H. Fletcher, and L. N. Frazer, Department of Geology and Geophysics, School of Ocean and Earth Science and Technology, University of Hawaii, 1680 East West Rd., Honolulu, HI 96822, USA. (neil@hawaii.edu)