Journal of Coastal Research	00	0	000–000	Coconut Creek, Florida	Month 0000

Toward Parsimony in Shoreline Change Prediction (III): B-Splines and Noise Handling

Tiffany R. Anderson and L. Neil Frazer

Department of Geology and Geophysics
School of Ocean and Earth Science and Technology
University of Hawaii
1680 East West Road
Honolulu, HI 96822, U.S.A.
tranders@hawaii.edu
neil@soest.hawaii.edu

www.cerf-jcr.org

www.JCRonline.org

ABSTRACT

Anderson, T.R. and Frazer, L.N., 0000. Toward parsimony in shoreline change prediction (III): B-splines and noise handling. *Journal of Coastal Research,* 00(0), 000–000. Coconut Creek (Florida), ISSN 0749-0208.

The traditional single-transect method for predicting long-term shoreline change uses far more parameters than necessary because it assumes that erosion/accretion (change) rates at adjacent alongshore positions (transects) are independent. Such overfitting can cause poor predictions of future shoreline location, so recent work has modeled change rates as linear sums of polynomials, or linear sums of principal components. Here we introduce an alternative method that uses linear sums of B-splines. As in earlier work, an information criterion is used to identify the optimal number of basis functions. The local nature of B-spline models makes them less susceptible to the Gibbs effect than polynomial models, and their smoothness makes them more robust to noise than principal components regression. We also compare three noise-handling techniques by examining their effects on the posterior probability density functions of rates. We find that noise handling affects both predicted rate and its uncertainty, and that correlated noise is best addressed by iteratively constructing a full covariance matrix from data residuals. We illustrate our procedure using synthetic data and shoreline data from Assateague Island and Ocean City, Maryland.

ADDITIONAL INDEX WORDS: *Shoreline change rates, inverse theory, coastal erosion, B-spline, Assateague Island, Ocean City, parsimony, covariance, correlated noise.*

INTRODUCTION

Quantification of trends in shoreline position is necessary for managing natural coastal environments and human communities. Models for shoreline change range from simple empirical relations, such as the model of this paper, to physics-based models. Empirical models gain parsimony (require fewer parameters) by ignoring waves, currents, and bathymetry. Physics-based forward modeling systems, such as Delft3D, developed by Delft Hydraulics (Roelvink and Van Banning, 1994), include those phenomena, but they must contend with limited availability of required data and large propagated errors. Semi-empirical models (*e.g.*, Davidson, Splinter, and Turner, 2013; Miller and Dean, 2004; Yates, Guza, and O'Reilly, 2009) must also contend with limited data availability and propagated errors, but to a lesser extent. Recently, Marghany, Hashim, and Cracknell (2011) incorporated wave spectra from airborne synthetic aperture radar (SAR) data into a long-term shoreline change model. Technical innovations such as global positioning systems (GPS) (*e.g.*, Dail, Merrifield, and Bevis, 2000), light detection and ranging (LIDAR) (*e.g.*, Stockdon *et al.*, 2002), and video imagery (*e.g.*, Becker *et al.*, 2007; Plant *et al.*, 2007) are greatly improving the quality and quantity of shoreline data, but available time series are not long enough to reveal trends over decades to centuries. In order to extract such trends, we are still largely dependent on historical shoreline surveys.

One of the simplest and most widely used methods of estimating long-term shoreline change is the single-transect (ST) method (Figure 1). It is an empirical method in which equally spaced cross-shore transects (*e.g.*, 20 to 50 m apart) are analyzed independently. The mathematical model for each transect is usually $y = rt + b$, in which y is shoreline position in the cross-shore (*i.e.* landward-seaward) direction, r is the rate of erosion or accretion, and b is an intercept that depends on the baseline relative to which y is measured. The data consist of a limited number of historical shorelines at each transect (Fletcher *et al.,* 2003; Genz *et al.*, 2007; Honeycutt, Crowell, and Douglas, 2001; Thieler *et al.,* 2009; *etc.*). The ST method is attractive because it is simple to understand and easy to implement, especially over large spatial regions. Also, it provides shoreline change statistics at a high spatial frequency along the shoreline. However, it ignores correlation between the rates at different transects, and thus it creates a model shoreline that can have large excursions that increase unrealistically as the model prediction is extended further into the future. Moreover, for shorelines along which rates change little from transect to transect, the ST method is highly

DOI: 10.2112/JCOASTRES-D-13-00032.1 received 7 February 2013; accepted in revision 19 August 2013; corrected proofs received 31 October 2013.
Published Pre-print online 5 December 2013.

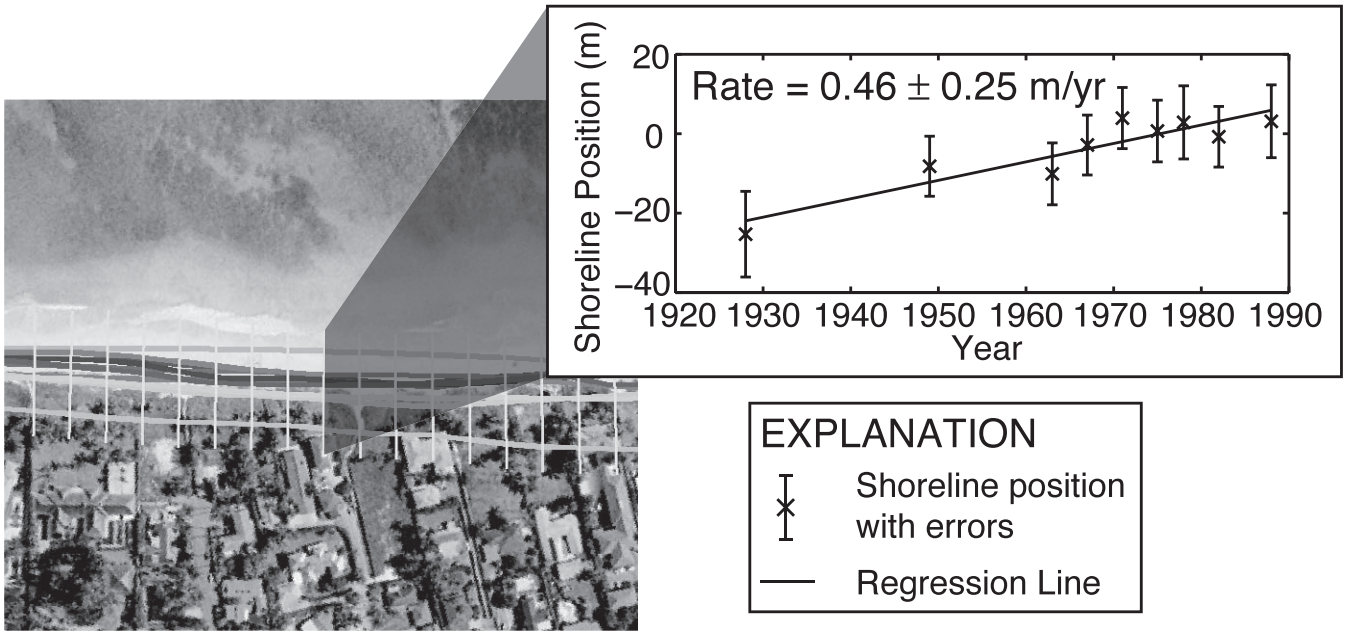

Figure 1. Weighted least squares linear regression of shoreline data along a single transect.

unparsimonious; it has far more parameters than are independent, as indicated by the wide autocorrelation of ST rate parameters along 40 km of shoreline centered on Ocean City inlet, Maryland, shown in Figure 2. Here, we focus on long-term (decades) and large-scale (kilometers) shoreline trends, attempting to improve on the ST method. Frazer, Genz, and Fletcher (2009), and Genz, Frazer, and Fletcher (2009) also addressed the overfitting issue of ST, and the procedure of this paper extends their work by the use of different basis functions and by an improved treatment of noise.

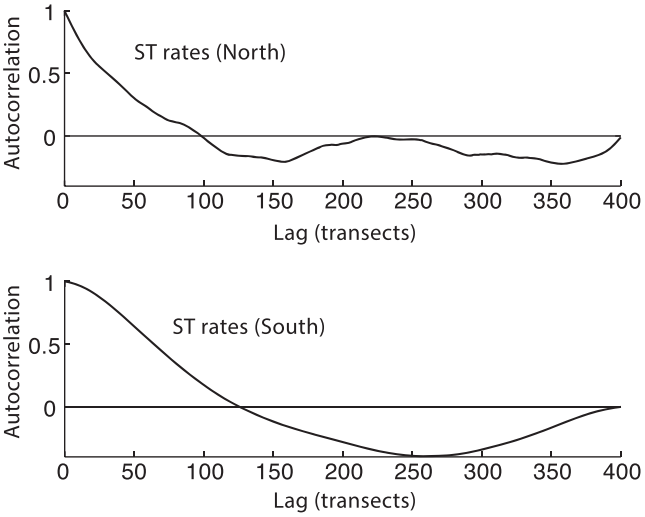

Figure 2. Autocorrelation of rates calculated by the ST method for the north (top) and south (bottom) study areas.

Frazer, Genz, and Fletcher (2009) examined three types of basis function: polynomials, eigenbeaches (principal components), and piecewise constant "bins." Examples of Legendre polynomials and eigenbeaches are shown in Figures 3d and e. Polynomials are subject to Gibbs effect (Bracewell, 2000) when modeling shorelines with sudden alongshore variations in rate (Figures 3b and c). The eigenbeaches method automatically does away with the Gibbs effect by using principal components of the shorelines themselves as the basis functions, but those basis functions are contaminated by process and measurement noise in the data (Figures 3b and c).

In this paper, we model cross-shore rate with a cubic spline defined by its value at alongshore locations called "knots" (de Boor, 1978). Splines avoid the Gibbs effect if extra knots are added at alongshore locations where rate changes rapidly, and they are not contaminated by noise (Figures 3b and c). We use B-splines (de Boor, 1978) as basis functions for the spline because of their simplicity. Following the methodology in Frazer, Genz, and Fletcher (2009), the parameters of our model are the coefficients of the basis functions; least squares regression is used to estimate those coefficients, and an information criterion (IC) is used to select the optimal number of basis functions. We improve on the noise methodology of Frazer, Genz, and Fletcher (2009) by iteratively estimating the spatial covariance of the data noise.

We illustrate the new procedures using barrier island shoreline data from Assateague Island and Ocean City, Maryland. We examine 40 km of shoreline centered on Ocean City (OC) inlet, as seen in Figure 4a. Inlet jetties disrupt alongshore sediment transport from the north, resulting in

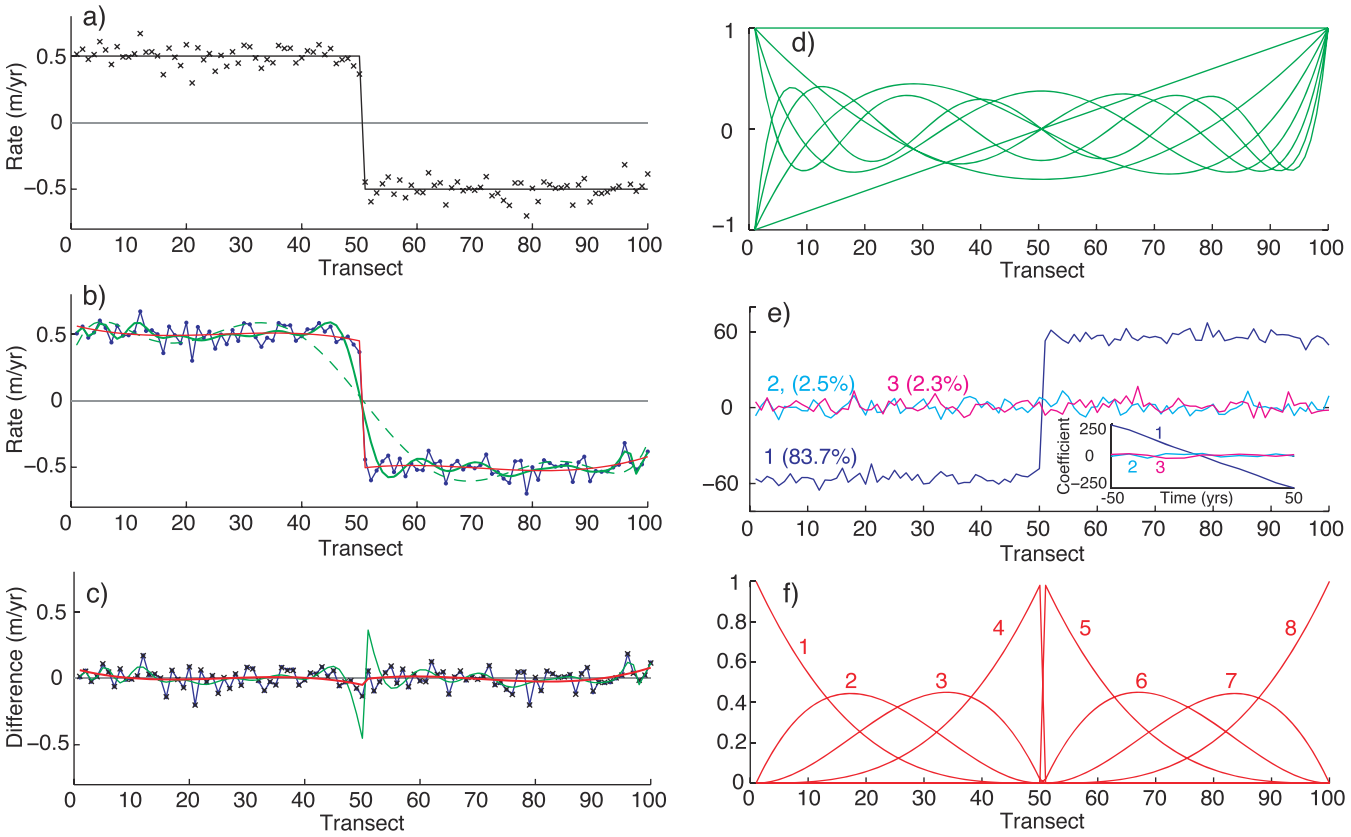

Figure 3. Results from modeling a synthetic data set consisting of 11 shorelines 10 years apart. Additive noise in the data is from $N(0, [7.5\ \text{m}]^2)$. (a) Solid line shows the true rates, constant on each side of a jump between transects 50 and 51; black crosses are the rates estimated by ST. (b) Rates estimated using 30 Legendre polynomials (green, solid), eight Legendre polynomials (green, dashed), eight B-splines (red), and one eigenbeach (blue). (c) Errors in rate (estimated – true) for ST (black crosses), 30 Legendre polynomials (green), 8 splines (red), and one eigenbeach (blue). (d) The Legendre polynomials. (e) The three most dominant eigenbeaches with percent variance. (f) The eight B-splines. Note Gibbs effect and noise in (b) and (c). The location of the rate jump was assumed known prior to modeling.

episodes of sand bypassing (Kraus, 2000; Schupp, Bass, and Grosskopf, 2007). The coast is mostly developed north of OC inlet, and it is undeveloped south of the inlet where Assateague Island National Park is located.

METHODOLOGY

Historical Shoreline Data

Our data consist of 12 historical shoreline surveys (1849–2000) obtained online as geographic information system (GIS) shape files from the U.S. Geological Survey (USGS) National Assessment of Shoreline Change for the New England and Mid-Atlantic Coasts (Himmelstoss *et al*., 2010). Table 1 lists the range of years for USGS shorelines derived from different data sources, and their average uncertainty values calculated from quantifiable sources of error (Hapke *et al*., 2010). Shorelines were extracted from T-sheets and air photos by digitally identifying the high water line (HWL) from georeferenced maps and photos (Hapke *et al*., 2010). Mean high water (MHW) indicates shoreline position in LIDAR-derived shorelines. Details of data extraction and error quantification procedures are available in Hapke *et al*. (2010). To correct for the horizontal difference between the MHW and HWL, we shifted the LIDAR shoreline landward by the proxy-datum bias values in Himmelstoss *et al*. (2010). The baseline, a proxy for a mean shoreline shape, was also obtained from the USGS online. We cast cross-shore transects perpendicular to the baseline and recorded the distance between each shoreline and the baseline at each transect using the Digital Shoreline Analysis System (DSAS) Version 4.2 (Thieler *et al*., 2009). There is a gap in the shoreline where OC inlet cuts through the barrier island, so we modeled the 20 km sections north and south of the inlet separately (Figures 4b and c).

Summary of Procedure

Spline models are distinguished by the number and location of their knots. For the Assateague data analyzed here, it is sufficient to count basis functions because we used knots that were regularly spaced, except for coincident knots at end points. For each model, we created a spline matrix, where each column of the matrix is a B-spline evaluated at each transect. We used the matrix in a generalized least squares regression model that includes iterative estimation of the data covariance matrix. From the residuals and the number of model

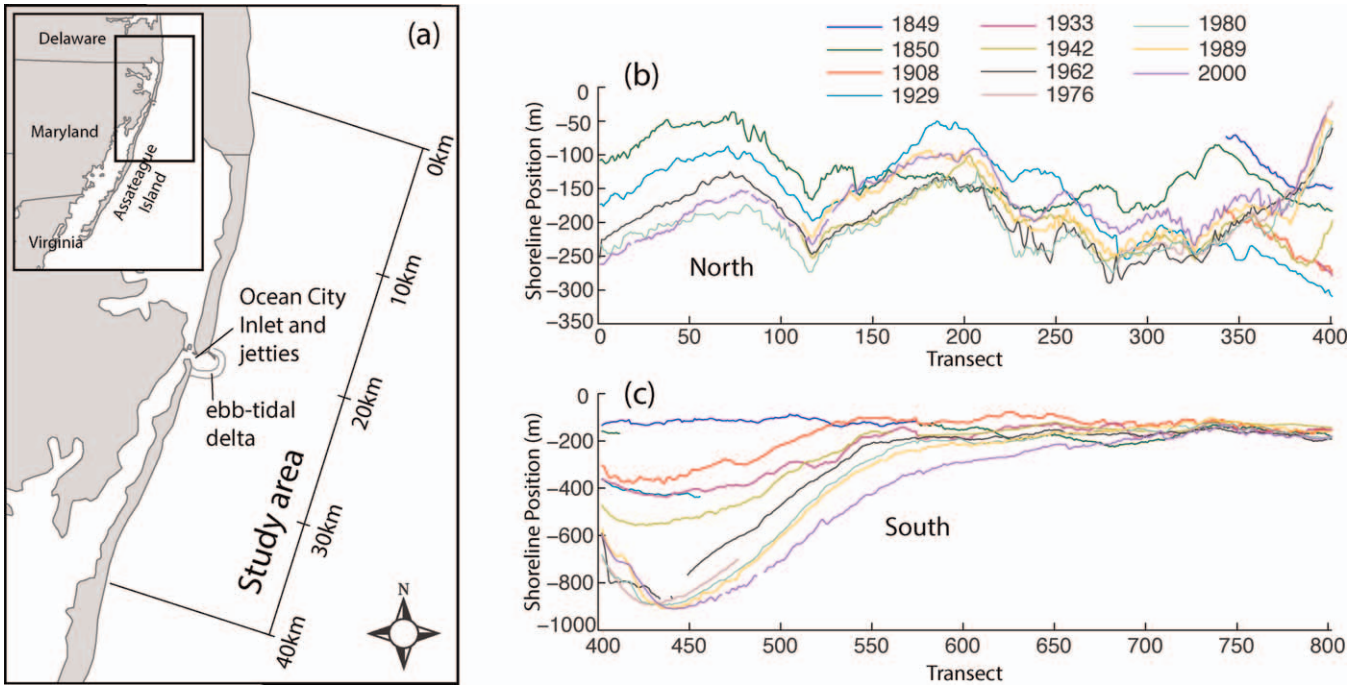

Figure 4. (a) Study area and (b–c) shoreline data relative to offshore baseline for sections north and south of Ocean City inlet. Transects are spaced 50 m apart.

parameters, including the parameters required for covariance, we calculated an information criterion (IC) statistic. We took the model with the lowest IC to be the best model, and we used it to predict shoreline positions at specified times within the study area. We also generated the posterior probability density function for rate at each transect.

B-Splines as Basis Functions

A cubic spline is a linear sum of cubic B-splines (de Boor, 1978). Each B-spline is defined everywhere, but it vanishes outside an interval that spans five knots in a prescribed knot sequence. Its first and second derivatives vanish at the end points of that interval, so each B-spline has a continuous second derivative, and thus the linear sum also has a continuous second derivative. Figure 5a shows 11 B-splines (solid lines) generated from a knot sequence consisting of 15 knots: 9 evenly spaced knots (dashed lines), 50 transects apart, between transects 0 and 400, with 4 knots at each end point. (The three extra knots at each end point cause the spline to vanish outside the interval.) The heavy line in Figure 5a is the B-spline with support from transect 50 to transect 250. The shape of the B-spline shows how it acts as a local weighted average. The B-splines near the ends have smaller apertures than the interior B-splines because of the duplicate knots at the ends. Although the B-splines in the figure were generated using evenly spaced interior knots, the spacing can be allowed to vary if there is compelling *a priori* information that warrants it. For example, if there were a geologic or structural feature, such as a stream mouth or pier, in the middle of an otherwise uniform sandy shoreline, one could add additional knots in the vicinity of the feature, allowing more basis functions to capture the higher spatial frequencies, as in Figure 3f.

Table 1. *Shorelines used in the study (HWL = high water line; MHW = mean high water; # Shore = number of shorelines; Unc. = average uncertainty from quantifiable sources).*

Date	Source	# Shore	Indicator	Unc. (m)
1849–1942	T-sheet	6	HWL	10.8
1962–1976	T-sheet	2	HWL	5.1
1980–1989	Air photo	3	HWL	3.2
2000	LIDAR	1	MHW	5.3

The number of knots at a knot location determines the smoothness at that point. More knots allow less continuity. For example, if there are two knots at the same location, the resulting spline will have a continuous first derivative at that point, but not necessarily a continuous second derivative. Four knots at the same location allow a zero-order discontinuity there.

Suppose for a moment that we are modeling only rate. We use B-splines to create a spline matrix V that linearly transforms a short column vector of spline coefficients $\tilde{m}$ to a much longer $I \times 1$ column vector of rates r, with one rate for each transect. (The tilde notation is called for because, in order to be consistent with single-transect models, we refer to r as the rate model, and later we will include rates and intercepts in a model vector m.) Each column of V is a B-spline generated using the recursion relation of de Boor (1978, p. 131), and the modeled rate at transect $1 \le i \le I$ is

$$r_i = V_{i,}\tilde{m}, \tag{1}$$

in which $V_{i,}$ denotes the ith row of V.

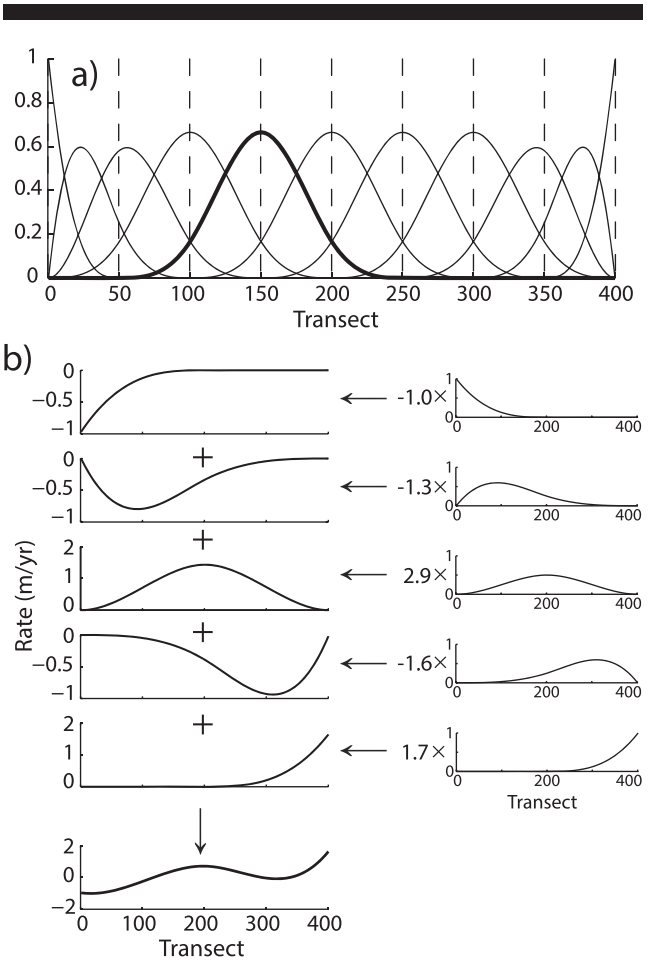

Figure 5. (a) Cubic B-splines. (b) Five B-splines multiplied by coefficients (right column) form a linear combination of B-splines to create a spline function (left column at bottom).

Shoreline Change Model

Here we combine B-splines with the simple, time-linear model $y = rt + b$. Let $y(x_i, t_j)$ represent the cross-shore shoreline position at alongshore location x_i and time t_j, for $i = 1:I$ transects and $j = 1:J$ shoreline surveys. Our shoreline equation is $y(x_i, t_j) = b_i + (t_j - \bar{t})r_i + n_{ij}$, where r_i and b_i are the rate and intercept, respectively, at the ith transect; n_{ij} is the noise at shoreline position $y(x_i, t_j)$; and $\bar{t}$ is the mean of shoreline survey years. Shifting the time origin to $\bar{t}$ helps condition the system matrix for regression, without affecting rate parameters.

Frazer, Genz, and Fletcher (2009) and Genz, Frazer, and Fletcher (2009) used an additional acceleration term in some of their models. Frazer, Anderson, and Fletcher (2009) and Anderson, Frazer, and Fletcher (2010) included a storm function in the time component of their shoreline model. We do not include acceleration or storm functions here because preliminary testing showed no significant acceleration or storm signals in this data set. However, both storm functions and acceleration terms may also be expressed as linear combinations of spline basis functions.

Here we use splines to model the alongshore variation in intercept as well as rate, so we construct the matrix V as a block-diagonal matrix with two blocks of splines, one for rates and one for intercepts. The two blocks have different numbers of columns if the number of splines needed for intercept differs from the number needed for rate. The vector $\tilde{m}$ contains the spline coefficients for both rate and intercept. Combining the time-linear model with spline modeling of rate and intercept gives

$$\begin{aligned} y(x_i, t_j) &= b_i + (t_j - \bar{t})r_i + n_{ij} \\ &= V_{i,}\tilde{m} + (t_j - \bar{t})V_{I+i,}\tilde{m} + n_{ij}. \end{aligned} \tag{2}$$

Basic Generalized Least Squares

Our generalized least squares (GLS) model is similar to that given in Frazer, Genz, and Fletcher (2009), but it is presented here in a slightly different way so that the matrix of spline basis functions is explicit. In the following section, and in the appendix, we present our method for handling spatially correlated noise. The GLS model for the single-transect method is written as

$$d = Gm + \eta, \tag{3}$$

in which $N \le I \times J$ is the number of data points, d is an $N \times 1$ vector of shoreline positions relative to the baseline; G is an $N \times 2I$ system matrix; m is a $2I \times 1$ vector of parameters (I rates and I intercepts); η is an $N \times 1$ vector of noise with zero mean and covariance matrix C_{dd}. To reduce the number of parameters in the model, we write

$$m = V\tilde{m} \tag{4}$$

in which V is the $2I \times M$ spline matrix described earlier, and $\tilde{m}$ is an $M \times 1$ vector of spline coefficients with $M \ll 2I$. Substituting Equation (4) into (3), and introducing the notation $\tilde{G} = GV$, gives the familiar GLS form $d = \tilde{G}\tilde{m} + \eta$, still with noise covariance C_{dd}.

In basic GLS, there is only one noise parameter. One begins with an *a priori* estimate of the covariance $\tilde{C}_{dd}$ and then scales it using a scaling factor that maximizes the fit of the model to the data. The basic GLS estimator of $\tilde{m}$ is the usual relation

$$\hat{\tilde{m}} = (\tilde{G}^T\tilde{C}_{dd}^{-1}\tilde{G})^{-1}\tilde{G}^T\tilde{C}_{dd}^{-1}d \tag{5}$$

with parameter covariance matrix (*e.g.*, Menke, 2012)

$$\hat{C}_{\tilde{m}\tilde{m}} = (\tilde{G}^T\hat{C}_{dd}^{-1}\tilde{G})^{-1} \tag{6}$$

in which the estimated data covariance $\hat{C}_{dd} = \hat{\alpha}\tilde{C}_{dd}$ contains the best-estimate constant of proportionality

$$\hat{\alpha} = (N - M)^{-1}(d - \tilde{G}\hat{\tilde{m}})^T\tilde{C}_{dd}^{-1}(d - \tilde{G}\hat{\tilde{m}}). \tag{7}$$

The estimated parameter vector $\hat{m}$ containing rates and intercepts at each transect is now

$$\hat{m} = V\hat{\tilde{m}} \tag{8}$$

with associated covariance matrix

$$\hat{C}_{mm} = V\hat{C}_{\tilde{m}\tilde{m}}V^T. \tag{9}$$

Shoreline positions are predicted using the best-fit shoreline model. The predicted shoreline position $\hat{y}(x_i, t)$ at desired time t

and transect location x_i is

$$\hat{y}(x_i,t) = q_i^T V\hat{\tilde{m}} = q_i^T \hat{m} \quad (10)$$

in which $q_i = q_i(t)$ is a $2I \times 1$ column vector, which we refer to as a prediction kernel. (If t is one of the data times t_j, then q_i is just the ith row of the system matrix G.) The variance of the prediction is

$$\hat{\sigma}^2(x_i,t) = q_i^T V\hat{C}_{\tilde{m}\tilde{m}}V^T q_i = q_i^T \hat{C}_{mm} q_i \quad (11)$$

in which $\hat{C}_{\tilde{m}\tilde{m}}$ (Eq. [6]) is the parameter covariance matrix, and $\hat{C}_{mm}$ (Eq. [9]) is the covariance matrix for modeled rates and intercepts (Eq. [8]). Confidence intervals are calculated by multiplying the square root of the above variance by the Student's t-statistic $t_{\nu,1-\varepsilon/2}$, where ν is the degrees of freedom, and $100(1-\varepsilon)$ is the percent confidence level.

Data Covariance Estimation

Ideally, the GLS model outlined in the previous section would contain a data covariance matrix C_{dd} that is known. In reality, the covariance matrix is rarely known *a priori*, even to within a scaling factor, and must be estimated. In the problem of this paper, the structure of the covariance matrix may be simplified by noting that shoreline surveys are typically dense in space but ~10 years apart in time. Therefore, it is a good approximation to assume that data errors are correlated in the alongshore direction but weakly correlated in time.

Frazer, Genz, and Fletcher (2009) and Genz, Frazer, and Fletcher (2009) addressed alongshore correlation by fitting the data residuals from the ST method to a decaying exponential and then using that exponential to generate the rows of the covariance matrix. Another method for handling correlated noise is to inflate a diagonal covariance matrix by using $N^* - M$ instead of $N - M$ in Equation (7), where N^* is the effective number of independent observations (Bayley and Hammersley, 1946). That is essentially what others (*e.g.*, Hapke *et al.*, 2010; Kane *et al.*, 2012) have done to estimate the variance in the average of correlated rates; we refer to that method as W_{N^*}.

Here, in what we refer to as the C_{full} method, we estimate the full covariance matrix for each spline model by an iterative process in which each successive covariance matrix is calculated from the residuals of the previous iteration (*e.g.*, Dosso, Neilsen, and Wilmut, 2006). Later, in the context of the spline model, we compare differences in the posterior probability density function (pdf) for all three methods of noise estimation: uncorrelated (W_{diag}), inflated variance based on an effective number of data (W_{N^*}), and full covariance (C_{full}).

Each of the noise handling methods incorporates in some way the shoreline position uncertainties provided by the USGS (Himmelstoss *et al.*, 2010). We condition the USGS LIDAR-derived shoreline uncertainty estimates so that uncertainty values are no less than their geometric mean. This conservative approach follows assumptions about noise within our C_{full} model given in Appendix A. For consistency, we use the conditioned errors as *a priori* estimates for all noise models. Let $\tilde{C}_{dd}^{(0)} = \text{diag}(w_{ij}^2)$ be the diagonal matrix containing conditioned USGS uncertainty estimates w_{ij} at locations i and times j. In the uncorrelated noise method (W_{diag}), we use $\tilde{C}_{dd}^{(0)}$ as the *a priori* estimate of $\tilde{C}_{dd}$. In the noise estimation method (W_{N^*}), we also use $\tilde{C}_{dd}^{(0)}$ as the *a priori* estimate $\tilde{C}_{dd}$, but we increase the constant of proportionality $\hat{\alpha}$ by substituting the effective number of independent data N^* for the actual number of data N in Equation (7). In the full-covariance (C_{full}) method, we iteratively construct $\tilde{C}_{dd}$ from the data according to the flowchart in Figure 6, and then multiply by the best-estimate constant of proportionality $\hat{\alpha}$. We limit the number of noise parameters associated with the scaled autocorrelation by assuming that it is independent of time and alongshore distance. The flowchart procedure for the C_{full} method is given in Appendix A, and the details of the convergence are given in Table 2 and explained in Appendix A.

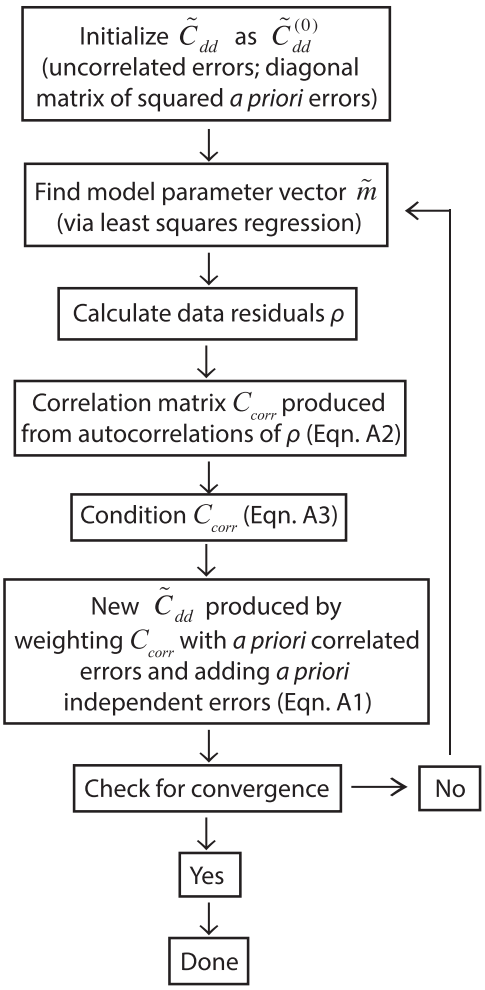

Figure 6. Flowchart for estimating the unscaled covariance matrix $\tilde{C}_{dd}$ (before scaling by $\hat{\alpha}$).

To test the goodness of our covariance matrix estimate, we compare the autocorrelation function for raw data residuals $\rho = d - \tilde{G}\tilde{m}$ with the autocorrelation function for standardized residuals defined as $\tilde{\rho} = \hat{C}_{dd}^{-1/2}\rho$. In GLS regression, the standardized residuals are minimized, and the autocorrelation function of the standardized residuals indicates the accuracy of the data covariance matrix estimate. Figure 7 shows the autocorrelations of raw residuals (left) and standardized

Table 2. *The decline in* LL *during iteration shows how an improved estimate of the covariance matrix gives a model with higher likelihood.* ΔLL *is the change in* LL *between successive iterations.* $\hat{\alpha}$ *is the covariance scaling factor needed to calculate* LL*, but is not used in the iteration process.* $||\cdots||_1$ *is the matrix* L_1 *norm of the difference in estimated covariance matrices between iterations. The iteration process on the left is for the best model in the north region, composed of seven rate and 72 intercept basis functions. The best model in the south region (right) has six rate and 61 intercept basis functions. Convergence requires about three iterations.*

	North Section			South Section		
Iteration	$(\Delta LL)/LL_{it-1}$	$\hat{\alpha}$	$\|\|\tilde{C}_{dd}^{(it)} - \tilde{C}_{dd}^{(it-1)}\|\|_1 \|\|\tilde{C}_{dd}^{(it-1)}\|\|_1^{-1}$	$(\Delta LL)/LL_{it-1}$	$\hat{\alpha}$	$\|\|\tilde{C}_{dd}^{(it)} - \tilde{C}_{dd}^{(it-1)}\|\|_1 \|\|\tilde{C}_{dd}^{(it-1)}\|\|_1^{-1}$
0	—	23.0	—	—	45.4	—
1	-8.25×10^{-1}	15.8	1.13×10^{3}	-9.14×10^{-1}	25.6	2.04×10^{3}
2	-1.30×10^{-3}	16.0	2.15×10^{-2}	-3.75×10^{-3}	27.4	7.30×10^{-2}
3	-6.78×10^{-6}	16.0	5.71×10^{-4}	-3.17×10^{-6}	27.4	1.07×10^{-3}
4	-1.07×10^{-7}	16.0	2.12×10^{-6}	-1.57×10^{-7}	27.4	3.05×10^{-5}
5	-2.32×10^{-9}	16.0	7.37×10^{-8}	9.24×10^{-9}	27.4	1.81×10^{-6}
6	-5.02×10^{-11}	16.0	1.09×10^{-9}	1.14×10^{-10}	27.4	8.66×10^{-8}
7	—	—	—	-5.75×10^{-12}	27.4	4.61×10^{-9}
8	—	—	—	-1.36×10^{-12}	27.4	2.46×10^{-10}

residuals (right) for two survey years. The delta-like autocorrelation of the standardized residuals indicates considerable improvement over the diagonal covariance assumption.

Information Criterion

Within a given family of models, the best model is the one with the minimum information criterion (IC) statistic (score, value). An IC statistic is the sum of two penalty terms: one term, the misfit, increases with the data residuals, and the other term increases with the number of parameters. Increasing the number of parameters reduces the first term and increases the second term. The IC score thus helps prevent one from fitting the noise instead of the data. For further reading on IC use in regression, see McQuarrie and Tsai (1998) or Frazer, Genz, and Fletcher (2009). This paper uses the corrected Akaike information criterion (AICc) (Hurvich and Tsai, 1989; Sugiura, 1978). Our formula for AICc includes additional terms that were unnecessary in Frazer, Genz, and Fletcher (2009) because they were constant across all models. Since our method of estimating the covariance matrix produces matrices with different off-diagonal behavior, we use the more general expression

$$\text{AICc} = \log|\hat{C}_{dd}| + (N - M) + 2KN/(N - K - 1), \qquad (12)$$

in which $\log(\ldots) + (N - M)$ represents the misfit and $2KN/(N - K - 1)$ penalizes the model based on parameter count. Here K is the number of parameters used in the modeling process; it is the sum of M and the number of parameters used to model variance. Table 3 summarizes parameter counts for different methods. For uncorrelated noise (W_{diag} method), only one parameter (the covariance scaling factor $\hat{\alpha}$) is associated with variance. The inflated diagonal covariance technique (W_{N^*}) uses two parameters for variance, one for the covariance scaling factor, and one associated with calculating the effective number of independent data, N^*. For model selection in the W_{N^*} method, we substitute N^* for N in Equation (12).

Because the C_{full} procedure uses an estimated autocorrelation function, one may first think to count the number of autocorrelation lags as the number of parameters associated with correlation. However, correlations for large lags contribute little to the covariance estimation, and they are poorly estimated, so we damp them in estimating the autocorrelation (Box, Jenkins, and Reinsel, 1994). The effective number of parameters associated with spatial correlation is thus the integral of the damping function given in Appendix A.

When comparing IC scores for models with different noise handling methods, it is helpful to understand how the misfit term is related to the likelihood of a model. As we assume Gaussian noise, the likelihood function for parameter vector m, given data d, is

$$L(m|d) = (2\pi)^{-N/2}|C_{dd}|^{-1/2}\exp\left[-\frac{1}{2}(d - Gm)^T C_{dd}^{-1}(d - Gm)\right]. \qquad (13)$$

As noted above, the misfit LL, is defined as –2 times the logarithm of the likelihood function. Therefore, smaller LL values correspond to models with higher likelihood. The $(\Delta LL)/LL$ values in Table 2 show LL decreasing (model likelihood increasing) during iteration toward the full covariance matrix, with convergence within three iterations.

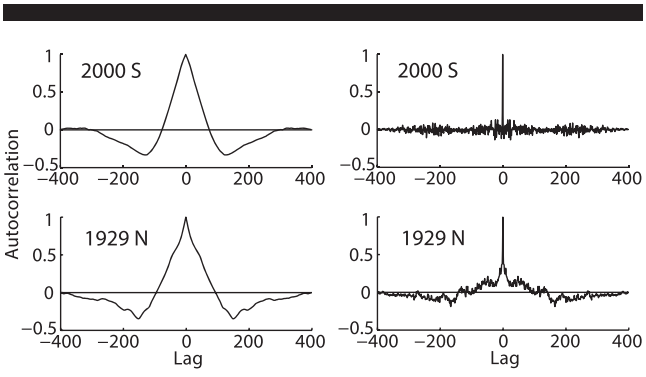

Figure 7. Selected autocorrelation functions for raw residuals (left column) and standardized residuals (right column) show the improvement due to a full covariance matrix.

RESULTS AND DISCUSSION

Model Selection: AICc

Figure 8 shows the ΔAICc scores, the difference in AICc from the lowest-scoring model, for each combination of rate and intercept parameters (up to 40 rate and 90 intercept parameters). The lowest AICc score in the north section corresponds

Table 3. *Model parameter counts and relative AICc scores of methods with differing covariance estimators. (ST_{ind} = single transect with variance calculated independently at each transect; W_{diag} = spline method with weighted diagonal covariance matrix; W_{N^*} = spline method with inflated weighted diagonal covariance calculated using the effective number of data; C_{full} = spline method with full covariance matrix; n_r = number of parameters used to model rates; n_b = number of parameters used to model intercepts; n_{var} = number of parameters used to estimate variance; K = IC parameter count ($n_r + n_b + n_{var}$); N = number of data; N* = effective number of data; ΔAICc = the AICc score minus the lowest AICc score.)*

	North Section						South Section					
Method	n_r	n_b	n_{var}	K	N^* (N = 2819)	ΔAICc	n_r	n_b	n_{var}	K	N^* (N = 3321)	ΔAICc
ST_{ind}	401	401	401	1203	—	13,053	401	401	401	1203	—	17,174
W_{diag}	8	34	1	43	—	8889	16	11	1	28	—	14,692
W_{N^*}	8	20	2	30	276	13,244	7	6	2	15	196	21,368
C_{full}	7	72	101	180	—	0	6	61	101	168	—	0

to the model with seven basis functions for rate and 72 basis functions for intercepts (see Table 3). Many more basis functions are required for modeling alongshore variations in intercept because the mean shoreline location depends on the shape of the baseline, and the baselines used in the USGS study are straight lines over large geographic regions, lines that do not mimic the mean shoreline shape on smaller scales. Alongshore variations in rate, however, are insensitive to cross-shore shifts in mean shoreline position. In fact, if only rates are desired, one could simply remove the mean of each transect, as was done in Frazer, Genz, and Fletcher (2009) and Genz, Frazer, and Fletcher (2009). Calculated shoreline positions, however, depend on intercepts, so reducing the number of intercept parameters is desirable for parsimony in shoreline prediction models. The same phenomenon is reflected in the AICc scores for the southern study region, where six rate basis functions and 61 intercept basis functions were found optimal.

Modeled Rates and Intercepts

Modeled rates and intercepts produced from the optimal number of spline basis functions are shown in Figure 9, along with rates and intercepts calculated from the ST method. The two methods generally agree, indicating long-term erosion from the northernmost location (transect 0), turning to accretion, which is greatest just north of the OC inlet. In the south section, erosion is most severe near the inlet, in agreement with previous studies, which found that the OC inlet jetties have disrupted the natural long-term sediment transport to the south, causing sand to accumulate just north of the inlet while retreating landward south of the inlet (Buttolph *et al.*, 2006; Dean and Perlin, 1977; Kraus, 2000; Leatherman, 1984).

For the north and south regions, we also computed an average change rate with 95% confidence interval by assuming a constant rate over each region, and inverting for that rate directly. For that calculation, we shifted the y-baseline to zero, so the intercept is zero, and the spline matrix was reduced to an $I \times 1$ column vector of ones. A large difference is apparent between the constant rates calculated for the beaches north (–0.12 ± 0.25 m/y) and south (–1.55 ± 0.45 m/y) of OC inlet, due to disruption of longshore transport by OC inlet jetties.

The spline and ST rates have similar long-wavelength shapes, but they deviate at shorter length scales, as expected. The spline intercepts closely resemble the ST intercepts because many spline basis functions were needed to model the intercepts. If one continues to increase the number of basis functions used to model intercept and rate, the spline results would approach ST estimates because the single transect method is the end member of basis function saturation—equivalent to having delta-like spline basis functions at each transect. If we also then applied the correlated noise handling technique to ST, the estimates would be identical.

Predicting Future Shorelines

Figure 10 shows predicted spline and ST shorelines for the northern section of the study site, from 1849 to 2100, a 100 year extension to the time span of the original data. With time, the high spatial frequencies in the ST method are amplified, producing unrealistic variations between neighboring transects, inconsistent with the smoothing processes observed over time in natural coastal environments. The spline methods, however, give smooth shoreline predictions (Fig. 10c). Figure 11 shows the corresponding results for the southern section. Figure 12 shows confidence intervals for the northern section predictions for the year 2100, using ST and all three spline methods, and it is interesting that the C_{full} spline method allows more alongshore detail than the W_{diag} and W_{N^*} spline methods because it fits the data better (next section). Moreover,

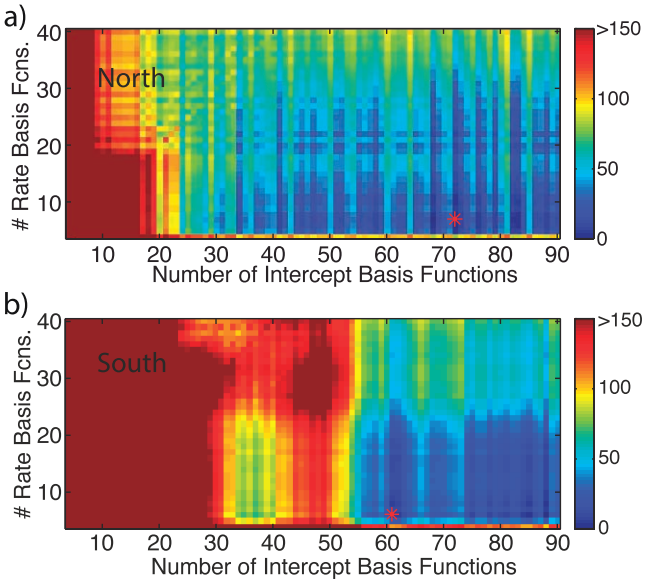

Figure 8. ΔAICc scores for each combination of rate and intercept basis functions north (a) and south (b) of OC inlet. Red stars indicate the minimum AICc values.

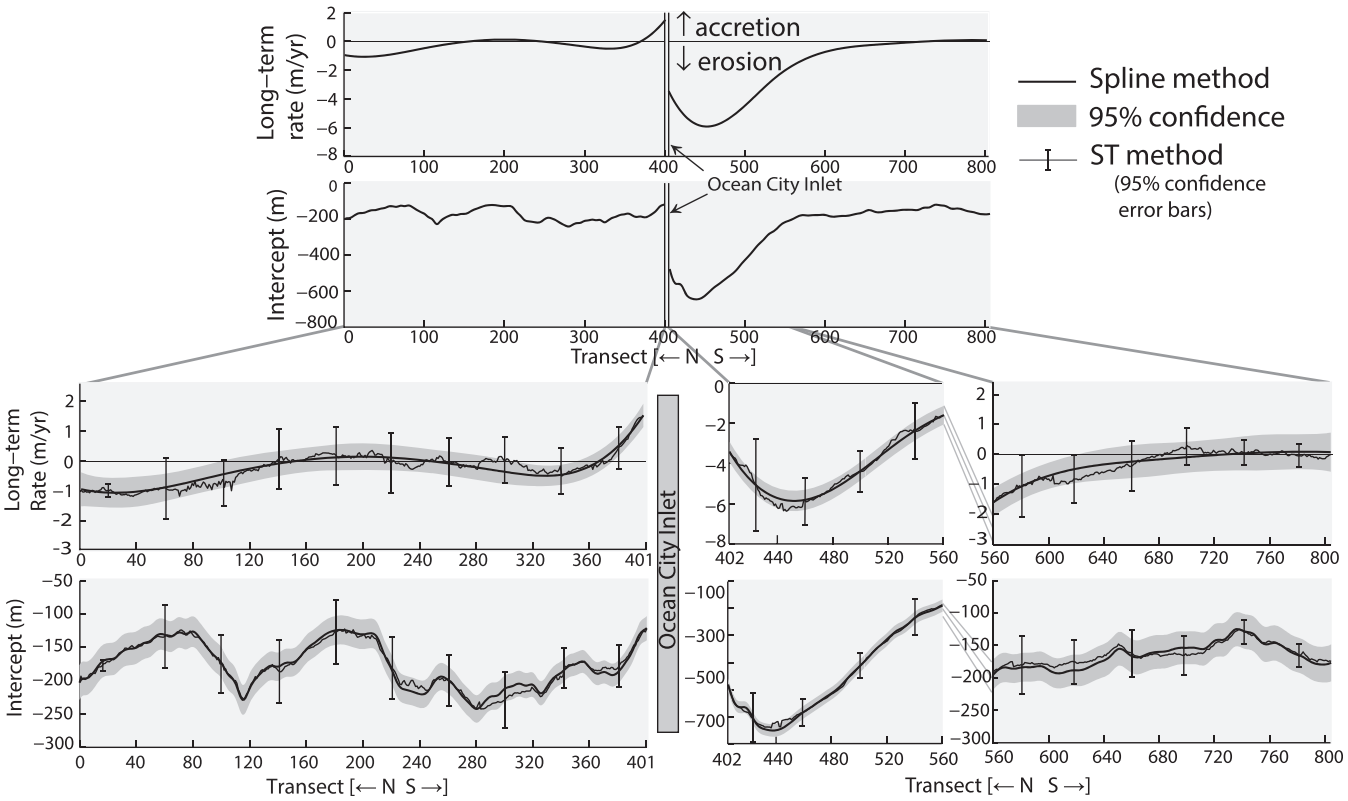

Figure 9. Modeled rates and intercepts along the coast of Ocean City and Assateague Island, Maryland, using the spline model and ST model. For clarity, the lower middle panels are plotted with a different vertical scale.

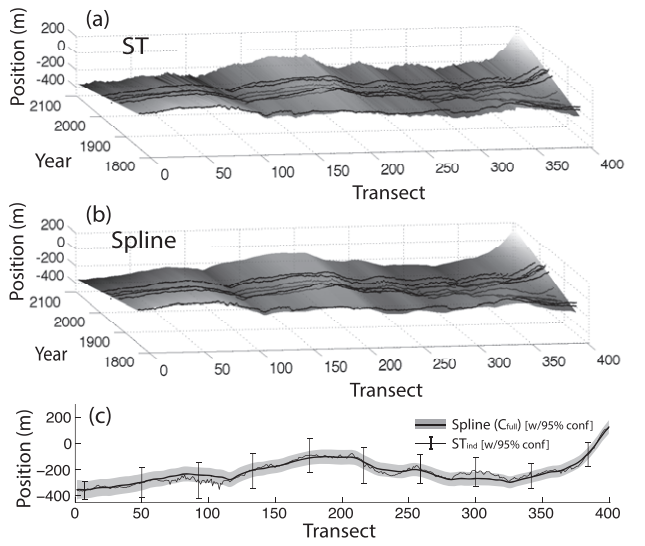

Figure 10. North section predicted shoreline positions (grayscale) for (a) ST and (b) C_{full} spline methods. Data are shown in black. Predictions are for the time span 1849–2100, a 100 year extension from the latest data. (c) Predicted shoreline positions in year 2100 for the ST and C_{full} spline methods reveal the unrealistic amplification of high spatial frequencies by ST.

the W_{diag} result shows that if correlations in the noise are not properly handled, the erroneously low estimates of parameter variance result in erroneously narrow confidence intervals. For example, the average standard error of W_{diag}-predicted positions in 2100 ($\sim$6 m) is significantly lower than the measurement-based estimate of $\sim$11 m by Douglas and Crowell (2000). They quantified several sources of shoreline position uncertainty at nearby Cotton Patch Hill, Delaware, including variability in the shoreline indicator (high water line) due to seasonal and tidal influence.

Comparing Noise Handling Techniques

As mentioned previously, we processed the data using our spline method in conjunction with several different noise handling techniques. It is worth noting that the issue of spatially correlated noise spans all modeling techniques, no matter how unparsimonious they may be. One could apply all of the noise handling techniques we present (W_{diag}, W_{N^*}, C_{full}) to the ST method because ST is just a spline method with delta-like splines at each transect. Since ST is so widely employed, we present only the standard ST noise methodology of estimating uncertainty independently at each transect, comparing that with our spline models. When comparing the ST method to our spline models, we denote the method as ST_{ind} as a reminder of the noise handling in the ST method.

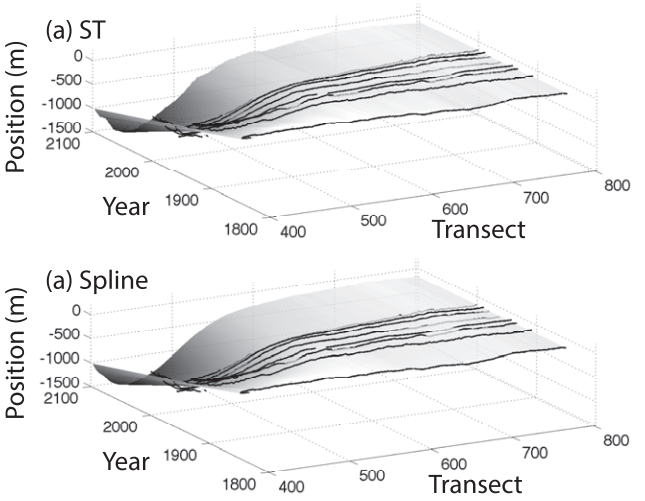

Figure 11. South section predicted shoreline positions for the (a) ST and (b) C_{full} spline methods during the time span 1849–2100.

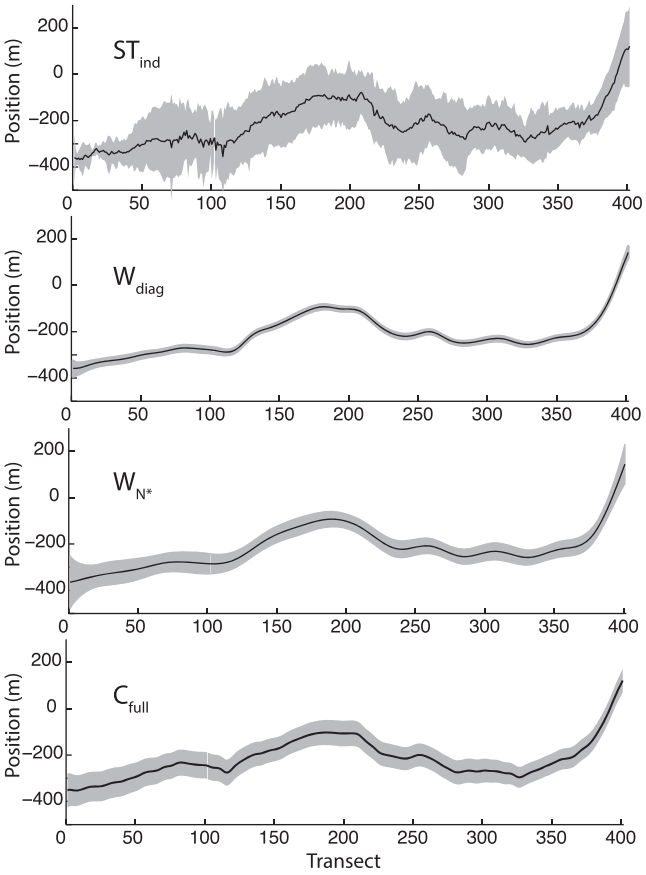

Figure 12. Northern section predictions for the year 2100, with 95% confidence intervals.

Table 3 contains, for each noise method, the optimal number of basis functions used to model rate (n_r), intercept (n_b), and variance (n_{var}). The total number of parameters used in the IC calculations, K, is also shown. The ST_{ind} method naturally has the largest number of model parameters—a rate, intercept, and variance at each transect. The W_{diag} method requires far fewer parameters, with the W_{N^*} method requiring even fewer parameters due to a much reduced effective number of data N^*. In the north section, N^* is 276, compared to the actual number of data $N = 2819$. In the south section, N^* is 196, out of $N = 3321$ actual data. The W_{diag} method also has reduced AICc values compared to ST in both the northern and southern portions, mainly due to the large reduction in parameters. The W_{N^*} method, on the other hand, actually has the highest AICc score of all due to its high LL. The high AICc score is mainly due to the unlikely large variance in combination with assumed uncorrelated errors, and it may not reflect actual accuracy (see Synthetic Data section). The C_{full} method requires about as many rate and intercept basis functions as the other two spline methods, but it needs extra parameters to estimate the covariance matrix; even with the extra parameters, its lower LL gives it the lowest AICc score.

Notably, the different noise handling methods change not only the variances of the parameter estimates, but also their means. The posterior pdfs of rate at four locations are shown in Figure 13. The pdfs of the ST_{ind} method are typically the broadest because of the relatively small amount of data at each transect. The pdfs for the W_{diag} method are more localized due to the increased number of data used to estimate each rate in the spline method. In the W_{diag} method each B-spline coefficient is essentially a weighted average, and since the data are assumed independent and identically distributed, the variance of the average is smaller than that of any individual datum. This illustrates how ignoring correlation causes estimates of parameter uncertainty to be erroneously low. On the other hand, the W_{N^*} noise model gives rates with broader pdfs than the W_{diag} noise model because of the inflated variances; using N^* in lieu of N reduces the number of parameters allowed by the IC because the fit is degraded, and that changes the mean of the rate pdf while increasing the variance. The pdfs for the C_{full} spline method are also broader than those in the W_{diag} method because covariance terms are included. Nevertheless, as Figure 12 shows, the C_{full} method reveals more alongshore detail because it is allowed more model parameters by the IC—because it fits the data better.

Comparing Basis Functions

The choice of basis function used in an analysis depends on both the analysis objective and the geology of the study area. No single type of basis function is superior in all situations, as summarized in Table 4. The ST method, although not truly a basis function method, is included in the table because it is equivalent to using delta-like basis functions at each transect.

All of the basis function methods produce more parsimonious models than ST, but they require careful handling of spatially correlated data errors. If the geology of the region warrants an alongshore discontinuity in rate (or intercept), B-splines and eigenbeaches (principal components) address the Gibbs effect (Figure 3) inherent in polynomial (Legendre, trigonometric)

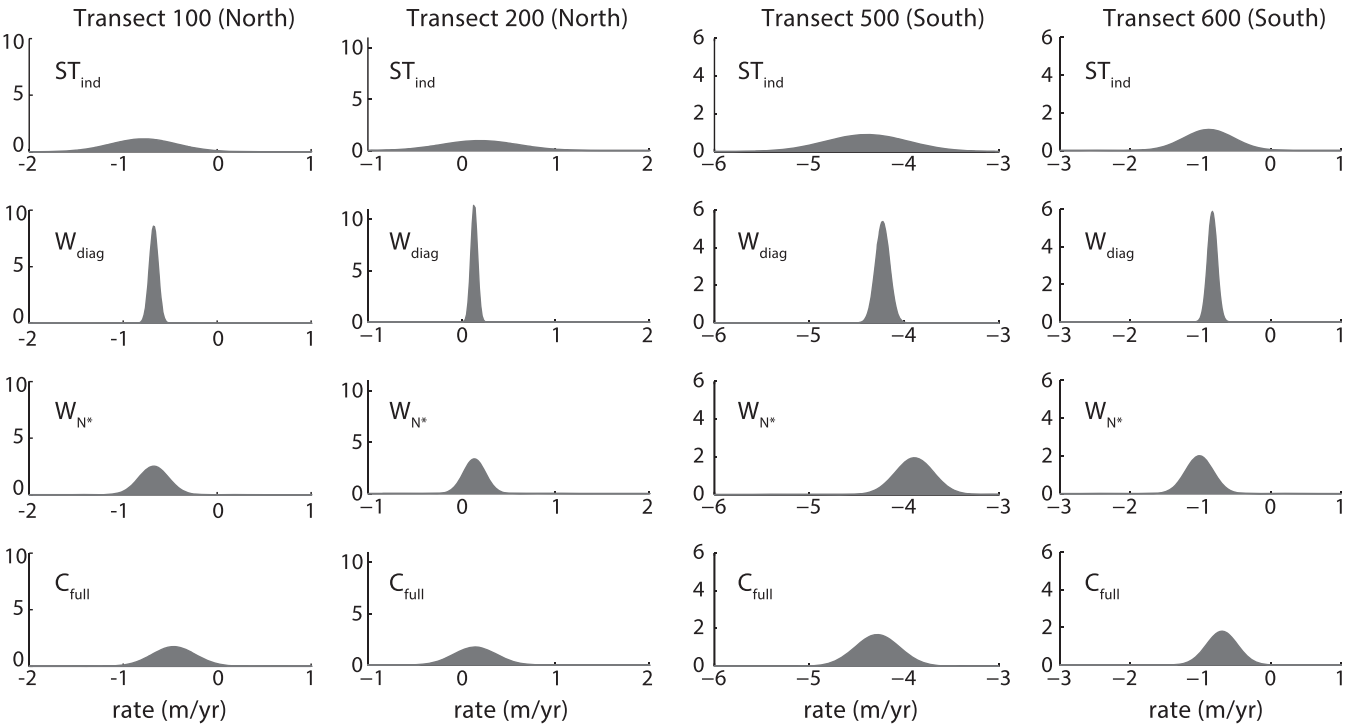

Figure 13. Posterior probability density functions for rate at four locations using: the single transect method with noise calculated independently at each transect (ST_{ind}); the spline method with uncorrelated noise (W_{diag}); the spline method with inflated diagonal (W_{N^*}); and the spline method with full covariance (C_{full}). The change in mean rate between ST_{ind} and W_{N^*} occurs because the IC allows the W_{N^*} method fewer basis splines.

basis functions, but not without added complication. The spline requires careful knot placement alongshore to circumvent the Gibbs effect. Eigenbeaches are contaminated by noise because they are derived from the data, requiring additional care in model selection and error estimation. Only the polynomial and spline basis functions are defined at every alongshore location, not just at each transect, a property that could be important if transects are widely spaced. If estimates are needed between transect locations, interpolation is required in the ST and eigenbeaches methods.

Limitations of the Time-Linear Model

All models tested here have the same shoreline change assumption in time—that shorelines erode or accrete at rates that are constant over decade to century time scales. This limitation affects predictions on either side of the OC inlet jetties because shoreline behavior has changed, perhaps even reversed, over the time span of our data. The predicted shoreline locations in the year 2100 for all models just north of the inlet actually exceed the seaward extent of the northern jetty. Since the jetty is the cause of sand accumulation to the north of the inlet, it is obviously incorrect to predict that the shoreline will continue to accrete past the seaward extent of the jetty. Given persistent coastal conditions, shorelines tend to move only when perturbed by a disruption, and then they gradually re-equilibrate to a stable state. Accordingly, the time component of our shoreline change model might be improved by incorporating relaxation in the time domain, following work done by Miller and Dean (2004) or Yates, Guza, and O'Reilly (2009). Incorporating our alongshore spline methods and noise

Table 4. *The effects of using different basis functions on model estimates.*

Method	Parsimonious	Defined Everywhere?	Gibbs Effect	Error Estimation Assumptions
ST	N	N	N	Independent
Spline (B-splines)	Y	Y	N[b]	Correlated errors[c]
Polynomial (Legendre, trigonometric)	Y	Y	Y	Correlated errors
Eigenbeaches (principal components)	Y[a]	N	N	Correlated errors; nonparametric[d]

[a]Eigenbeaches basis functions are not independent of the data; additional care is required for selecting a parsimonious number of basis functions (*e.g.*, cross-validation).
[b]Requires careful knot placement to circumvent Gibbs effect,
[c]Correlated errors = covariance matrix must account for spatially correlated errors.
[d]Nonparametric = requires nonparametric estimation method (*e.g.*, bootstrap).

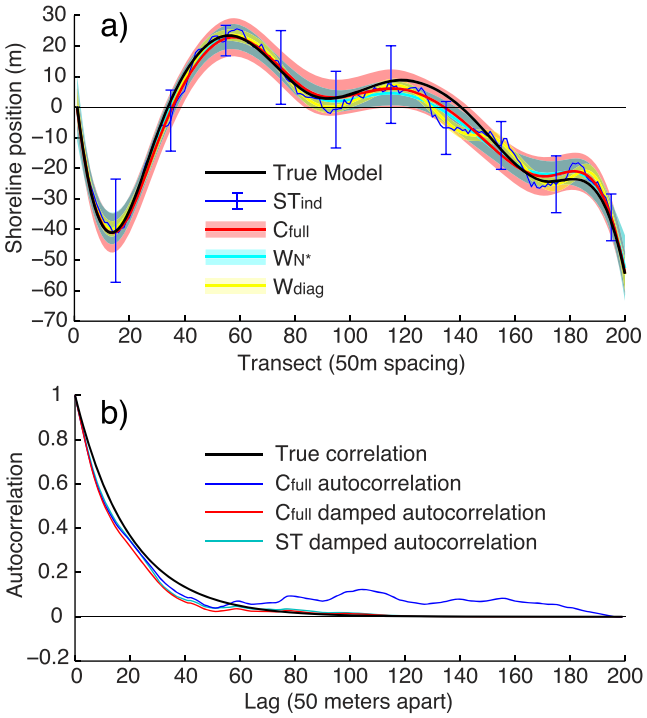

Figure 14. (a) True (black line) and modeled (colored) most-recent synthetic shoreline. (b) True alongshore correlation and modeled correlation. Damped ST autocorrelation is not used in modeling, but it has a similar shape to the C_{full} model autocorrelation.

handling into the Kalman Filter shoreline evolution framework described in Long and Plant (2012) might also provide additional improvement.

Confirmation with Synthetic Data

To be sure our method is self-consistent, we generated a synthetic data set with splines and then analyzed it. Ten cubic B-splines were used to generate alongshore rates; intercepts were defined as zero everywhere. From the defined rates and intercepts, 11 synthetic shorelines at 10 year intervals were sampled at 50 m increments alongshore. Correlated, zero-mean

Table 5. *Model parameter counts and relative AICc scores of methods with differing covariance estimators. (True = true model;* ST_{ind} *= single transect with variance calculated independently at each transect;* W_{diag} *= spline method with weighted diagonal covariance matrix;* W_{N^*} *= spline method with inflated weighted diagonal covariance calculated using the effective number of data;* C_{full} *= spline method with full covariance matrix;* n_r *= number of parameters used to model rates;* n_b *= number of parameters used to model intercepts; ΔAICc = the AICc score minus the lowest AICc score).*

Method	n_r	n_b	ΔAICc
True	10	1	—
ST_{ind}	401	401	5674
W_{diag}	12	20	5036
W_{N^*}	10	5	7133
C_{full}	10	1	0

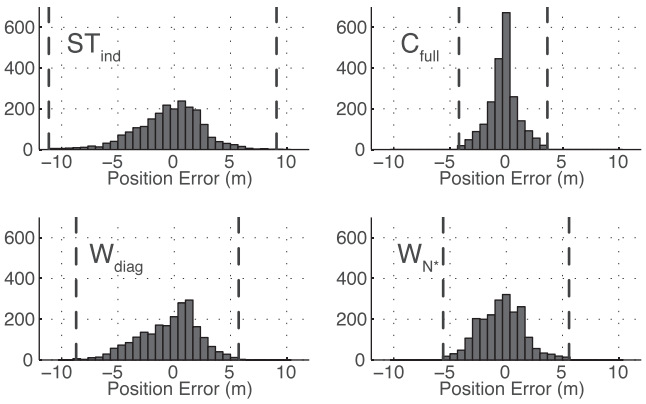

Figure 15. Histograms of shoreline positions errors (estimated – true) for ST_{ind}, W_{diag}, C_{full}, and W_{N^*} methods. Minimum and maximum error values are indicated by vertical dashed lines.

noise (Figure 14b) was weighted by *a priori* uncertainty estimates of 10 m for the four oldest shorelines and 7.5 m for the remaining seven shorelines, and then added to the synthetic data. We modeled the data using the ST_{ind}, W_{diag}, W_{N^*}, and C_{full} methods.

The C_{full} model is unique in identifying the true number of rate (10) and intercept (one) B-splines used to create the synthetic data (Table 5). Its low relative AICc score indicates that the C_{full} model is also the most probable. The W_{N^*} method also selects 10 rate B-splines, but it uses four B-splines to model intercepts alongshore, indicating that W_{N^*} overfits the data compared to C_{full}. As in the real data example, the W_{N^*} method has the highest AICc due to its inflated variance. The W_{diag} and ST_{ind} methods use the most B-splines (Table 5) to model alongshore parameters, indicating that they overfit the data compared to both the C_{full} and W_{N^*} methods.

Histograms of shoreline position errors (estimated – true) (Figure 15) show the C_{full} model to be the most accurate, followed closely by the W_{N^*} model, and then the W_{diag} and ST_{ind} models. The modeled noise correlation in the C_{full} method resembles the true correlation (Figure 14, inset). The autocorrelation of ST residuals is also similar to the true correlation. Model predictions of the most recent shoreline are shown in Figure 14; the synthetic data test also confirms that the uncertainty in the W_{diag} model is significantly underestimated, as seen in Figure 14, where the estimate does not overlap the true signal.

CONCLUSION

Modeling of real and synthetic data shows that B-splines do a good job of representing spatial variations of model parameters. B-splines can avoid the Gibbs effect at abrupt discontinuities in rates (if knots are properly chosen), a problem that arises with Legendre polynomial and trigonometric basis functions, and they avoid problems with noise contamination in eigenbeaches. It is well known that model parsimony is important to avoid overfitting, but we also show that differences in noise handling affect mean parameter estimates,

as well as estimated uncertainties. Treating correlated noise as if it were uncorrelated results in erroneously low variance estimates, but trying to get around this phenomenon by inflating variance estimates *via* an effective number of data limits the number of model parameters allowed, causing an increase in the misfit. Correlated noise is best handled by including a full covariance matrix, constructed iteratively from data residuals. Although our methods address both parsimony and noise handling, the shoreline near Ocean City Inlet reveals the usual limitations of the underlying time-linear assumption.

ACKNOWLEDGMENTS

We thank Brad Romine and Matt Barbee of the University of Hawaii Coastal Geology Group for their technical expertise, and Emily Himmelstoss from the USGS for help with shoreline data. Coastal Group Leader, Chip Fletcher, supplied encouragement, support, and good advice, as always. Janet Becker and Paul Wessel provided helpful comments on the manuscript. We are grateful for the comments and suggestions provided by three anonymous reviewers. This study was funded by the U.S. Geological Survey National Shoreline Assessment, the U.S. Army Corps of Engineers Regional Sediment Management Program, and the Department of Interior Pacific Islands Climate Science Center through Cooperative Agreement G12AC00003.

LITERATURE CITED

Anderson, T.R.; Frazer, L.N., and Fletcher, C.H., 2010. Transient and persistent shoreline change from a storm. *Geophysical Research Letters*, 37(8), L08401. doi:10.1029/2009GL042252

Bayley, G.V. and Hammersley, J.M., 1946. The "effective" number of independent observations in an autocorrelated time series. *Supplement to the Journal of the Royal Statistical Society*, 8(2), 184–197.

Becker, J.M.; Firing, Y.L.; Aucan, J.; Holman, R.; Merrifield, M., and Pawlak, G., 2007. Video-based observations of nearshore sand ripples and ripple migration. *Journal of Geophysical Research*, 112(C1), C01007. doi:10.1029/2005JC003451

Box, G.E.P.; Jenkins, G.M., and Reinsel, G.C., 1994. *Time Series Analysis: Forecasting and Control*, 3rd edition. Upper Saddle River, New Jersey: Prentice-Hall, 598p.

Bracewell, R.N., 2000. *The Fourier Transform and its Applications*, 3rd edition. New York: McGraw-Hill, 616p

Buttolph, A.M.; Grosskopf, W.G.; Bass, G.P., and Kraus, N.C., 2006. Natural sand bypassing and response of ebb shoal to jetty rehabilitation, Ocean City Inlet, MD, USA. *In: Proceedings of the 30th Coastal Engineering Conference*. Hackensack, New Jersey: World Scientific Press, pp. 3.344–3.356.

Dail, H.J.; Merrifield, M.A., and Bevis, M., 2000. Steep beach morphology changes due to energetic wave forcing. *Marine Geology*, 162(2–4), 443–458.

Davidson, M.A.; Lewis, R.P., and Turner, I.L., 2010. Forecasting seasonal to multi-year shoreline change. *Coastal Engineering*, 57(6), 620–629.

Davidson, M.A.; Splinter, K.D., and Turner, I.L., 2013. A simple equilibrium model for predicting shoreline change. *Coastal Engineering,* 73, 191–202.

de Boor, C., 1978. *A Practical Guide to Splines*. New York: Spinger-Verlag, 392p.

Dean, R.G. and Perlin, M., 1977. Coastal engineering study of Ocean City Inlet, Maryland. *In: Proceedings Coastal Sediments '77* (Reston, Virginia, ASCE), pp. 520–540.

Dosso, S.E.; Neilsen, P.L., and Wilmut, M.J., 2006. Data error covariance in matched-field geoacoustic inversion. *Journal of the Acoustical Society of America*, 119(1), 208–219.

Douglas, B.C. and Crowell, M., 2000. Long-term shoreline position prediction and error propagation. *Journal of Coastal Research,* 16(1), 145–152.

Fletcher, C.H.; Rooney, J.J.B.; Barbee, M.; Lim, S.-C., and Richmond, B.M., 2003. Mapping shoreline change using digital ortho-photogrammetry on Maui, Hawaii. *In:* Byrnes, M.; Crowell, M., and Fowler, C. (eds.), *Shoreline Mapping and Change Analysis: Technical Considerations & Management Implications*. Journal of Coastal Research, Special Issue No. 38, pp. 106–124.

Frazer, L.N.; Anderson, T.R., and Fletcher, C.H., 2009. Modeling storms improves estimates of long-term shoreline change. *Geophysical Research Letters*, 36(20), L20404. doi: 10.1029/2009GL040061.

Frazer, L.N.; Genz, A.S., and Fletcher, C.H., 2009. Toward parsimony in shoreline change prediction (I): Methods. *Journal of Coastal Research,* 25(2), 366–379.

Genz, A.S.; Fletcher, C.H.; Dunn, R.A.; Frazer, L.N., and Rooney, J.J., 2007. The predictive accuracy of shoreline change rate methods and alongshore beach variation on Maui, Hawaii. *Journal of Coastal Research,* 23(1), 87–105.

Genz, A.S.; Frazer, L.N., and Fletcher, C.H., 2009. Toward parsimony in shoreline change prediction (II): applying basis function methods to real and synthetic data. *Journal of Coastal Research*, 25(2), 380–392.

Hapke, C.J.; Himmelstoss, E.A.; Kratzman, M.; List, J.H., and Thieler, E.R., 2010. National Assessment of Shoreline Change; Historical Shoreline Change along the New England and Mid-Atlantic Coasts: *U.S. Geological Survey Open-File Report 2010-1118*, 57 p., available at http://pubs.usgs.gov/of/2010/1118/.

Himmelstoss, E.A.; Kratzmann, M.; Hapke, C.; Thieler, E.R., and List, J., 2010. The National Assessment of Shoreline Change: A GIS Compilation of Vector Shorelines and Associated Shoreline Change Data for the New England and Mid-Atlantic Coasts: *U.S. Geological Survey Open-File Report 2010-1119*, available at http://pubs.usgs.gov/of/2010/1119/.

Honeycutt, M. G.; Crowell, M., and Douglas, B.C., 2001. Shoreline-position forecasting: impact of storms, rate-calculation methodologies, and temporal scales. *Journal of Coastal Research,* 17(3), 721–730.

Hurvich, C.M. and Tsai, C.-L., 1989. Regression and time series model selection in small samples. *Biometrika*, 76(2), 297–307.

Hwang, D.J., 2005. *Hawaii Coastal Hazard Mitigation Guidebook*. Honolulu, Hawaii: University of Hawaii Sea Grant College Program, 216p.

Kane, H.H.; Fletcher, C.H.; Romine, B.M.; Anderson, T.R.; Frazer, N.L., and Barbee, M.M., 2012. Vulnerability assessment of Hawai'i's cultural assets attributable to erosion using shoreline trend analysis techniques. *Journal of Coastal Research*, 28(3), 533–539.

Kraus, N.C., 2000. Reservoir model of ebb-tidal shoal evolution and sand bypassing. *Journal of Waterways, Port, Coastal and Ocean Engineering*, 126, 305–313.

Leatherman, S.P., 1984. Shoreline evolution of North Assateague Island, Maryland. *Shore and Beach*, 52(3), 3–10.

Long, J.W. and Plant, N.G., 2012. Extended Kalman Filter framework for forecasting shoreline evolution. *Geophysical Research Letters*, 39(13), L13603. doi:10.1029/2012GL052180

Marghany, M.; Hashim, M., and Cracknell, A., 2011. Simulation of shoreline change using AIRSAR and POLSAR C-band data. *Environmental Earth Sciences,* 64(4), 1177–1189.

McQuarrie, A.D.R. and Tsai, C.-L., 1998. *Regression and Time Series Model Selection*. Singapore: World Scientific.

Menke, W., 2012. *Geophysical Data Analysis: Discrete Inverse Theory*, 3rd Edition. San Diego, California: Academic, 293p.

Miller, J.K. and Dean, R.G., 2004. A simple new shoreline change model. *Coastal Engineering*, 17(3), 531–556.

Plant, N.G.; Aarninkhof, S.G.J.; Turner, I.L., and Kingston, K.S., 2007. The performance of shoreline detection models applied to video imagery. *Journal of Coastal Research*, 23(3), 658–670.

Roelvink, J.A. and Van Banning, G.K.F.M., 1994. Design and development of Delft3D and application to coastal morphodynamics. *In:* Verwey, A.; Minns, A.W.; Babovic, V., and Maksimovic, C. (eds.), *Hydroinformatics* '94. Rotterdam, The Netherlands: Balkema, pp. 451–455.

Schupp, C.A.; Bass, G.P., and Grosskopf, W.G., 2007. Sand bypassing restores natural processes to Assateague Island, Maryland. *In:* Kraus, N.C. and Rosati, J.D. (eds.), *Coastal Sediments '07: Proceedings of the Sixth International Symposium on Coastal Engineering and Science of Coastal Sediment Processes*, May 13–17, 2007 (New Orleans, Louisiana, ASCE), pp. 1340–1353.

Stockdon, H.F.; Sallenger, A.H., Jr.; List, J.H., and Holman, R.A., 2002. Estimation of shoreline position and change using airborne topographic LIDAR data. *Journal of Coastal Research*, 18(3), 502–513.

Sugiura, N., 1978. Further analysis of the data by Akaike's information criterion and the finite corrections. *Communications in Statistics, Theory and Methods*, A7, 13–26.

Thieler, E.R.; Himmelstoss, E.A.; Zichichi, J.L., and 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*. http://pubs.usgs.gov/of/2008/1278/

Yates, M.L.; Guza, R.T., and O'Reilly, W.C., 2009. Equilibrium shoreline response: observations and modeling. *Journal of Geophysical Research: Oceans*, 114(C9), C09014. doi:10.1029/2009JC005359

APPENDIX A: ESTIMATING THE CORRELATED DATA COVARIANCE MATRIX

We reduce the number of parameters needed to model noise and increase model parsimony by assuming stationary errors in the alongshore direction of our correlated noise model. This assumption requires special handling of LIDAR-derived shoreline uncertainty estimates because they are not constant alongshore; notably, adjacent uncertainty estimates can vary significantly (Fig. A1d, solid line; dashed line shows geometric mean of uncertainty values). One option is to use a constant, such as the mean or median, to represent all LIDAR uncertainty values. We opted instead to retain some variability in LIDAR uncertainties by assuming that noise is a mixture of spatially correlated and uncorrelated errors that are independent of each other; this simplifies the model while still allowing incorporation of larger LIDAR uncertainties. Our covariance matrix $\tilde{C}_{dd}$ (prior to scaling by the best-estimate constant of proportionality, Eq. [7]) has the form

$$\tilde{C}_{dd} = \bar{E}^{1/2}\ C_{corr}\,\bar{E}^{1/2} + \tilde{E} \tag{A1}$$

in which $\bar{E}^{1/2}C_{\text{corr}}\bar{E}^{1/2}$ is a block diagonal matrix representing correlated noise and $\tilde{E}$ is a diagonal matrix representing additional spatially independent noise. Matrix $\bar{E}$ consists of values $\bar{E}_{ii'jj'} = \bar{e}_j{}^2\delta_{jj'}\delta_{ii'}$ where $\bar{e}_j$ are the geometric means of uncertainty values at times j, and $\delta_{ii'}$ and $\delta_{jj'}$ are Kronecker deltas. Simply put, matrix $\bar{E}$ is used to weight each block (J blocks—one for each shoreline) within the block diagonal correlation matrix C_{corr} (defined below) by the geometric mean of uncertainty estimates for the corresponding shoreline. Matrix $\tilde{E}$ is defined as $\tilde{E} = diag([e_{ij}^2 - \bar{e}_j^2]_+)$, where e_{ij} are the raw USGS uncertainty estimates at locations i and times j, and $\bar{e}_j$ are the geometric means of uncertainty values at times j. The '+' subscript denotes positive values, indicating that only e_{ij} values exceeding means $\bar{e}_j$ are used to determine additive, independent errors. Shorelines not derived from LIDAR have constant uncertainty estimates alongshore, so values along the diagonal of matrix $\bar{E}$ corresponding to shorelines not derived from LIDAR will be zero; only the LIDAR-derived shorelines have additional spatially independent uncertainty.

For consistency across noise models, we use conditioned USGS error estimates in which LIDAR-derived shoreline uncertainties are given a lower bound equal to their geometric mean; uncertainty values less than their geometric mean are replaced with the mean. We define, in the main text, $\tilde{C}_{dd}^{(0)} = \text{diag}(w_{ij}^2)$ as the diagonal matrix containing conditioned USGS uncertainty estimates w_{ij} at locations i and times j. Matrix $\tilde{C}_{dd}^{(0)}$ can also be written as $\tilde{C}_{dd}^{(0)} = \bar{E} + \tilde{E}$.

Following the flowchart in Figure 6, we begin our iterative process of finding $\tilde{C}_{dd}$ by initializing $\tilde{C}_{dd}$ as the diagonal matrix $\tilde{C}_{dd}^{(0)}$ defined in the preceding paragraph. A diagonal covariance matrix represents uncorrelated data errors. Using $\tilde{C}_{dd}^{(0)}$, we invert for model parameter vector $\hat{\tilde{m}}$ (Eq. [5]), and subsequently calculate residuals $\rho = d - \tilde{G}\hat{\tilde{m}}$. The correlation matrix C_{corr} is then constructed using the following equation:

$$c_{ii'jj'} = \left(\sum_j \sum_i \rho_{i,j}^2\right)^{-1} \sum_{j=1}^{J} \sum_{k=1}^{I-|i'-i|} \rho_{k,j}\rho_{k+|i'-i|,j} \cdot \delta_{jj'} \tag{A2}$$

in which i and i' are location indices ranging from 1 to I, j and j' are time indices ranging from 1 to J, and $\delta_{jj'}$ is the Kronecker delta, which enforces our assumption that data are not correlated in time. If the data vector in the GLS problem is arranged such that data are grouped by time, and ordered spatially, the correlation matrix is block diagonal with each block equal and in Toeplitz form. Figure A1, left column, shows the spatial autocorrelation of shoreline residuals (first row of one block) for the north and south study sections.

To condition the matrix and ensure that it is positive definite, each autocorrelation in the correlation matrix is damped with the taper function (Fig. A1c)

$$T(|i' - i|) = \cos^l \frac{\pi|i' - i|}{2(I-1)}, \tag{A3}$$

in which the exponent l controls the rate of decay as data are spatially farther apart. We tested the values $l = 3$, 6, 10, and 20 for selected combinations of basis functions, and we found that they all produced nearly identical parameter vectors. We used $l = 6$ in our study, which causes the correlation to go nearly to zero by about lag $3/4I$. Figures A1a and b show the damped autocorrelation of residuals *versus* lag. Figure A1e shows one of the $I \times I$ blocks within the block-diagonal C_{corr} matrix after damping with the taper function.

A new $\tilde{C}_{dd}^{(1)}$ is calculated from correlation matrix C_{corr} following Eq. (A1). Selected rows of the $I \times I$ block within the final iteration of $\tilde{C}_{dd}$ relating to the LIDAR shoreline are shown in Figure A1f. We then use the new full covariance matrix to invert for an updated parameter vector $\hat{\tilde{m}}$, and we use that parameter vector to estimate yet another covariance matrix and model pair, and so on, until the covariance matrix is sufficiently similar to the one from the previous iteration. Our convergence criterion is $||\tilde{C}_{dd}^{(u)} - \tilde{C}_{dd}^{(u-1)}||_1 < 10^{-10}||\tilde{C}_{dd}^{(u-1)}||_1$, where u is the number of iterations, and $||\cdots||_1$ denotes the matrix L_1-norm. Notice that it is not necessary to calculate the constant of proportionality $\hat{\alpha}$ until the iterations have converged, because it drops out of Eq. (5), the GLS estimator of parameter vector $\tilde{m}$.

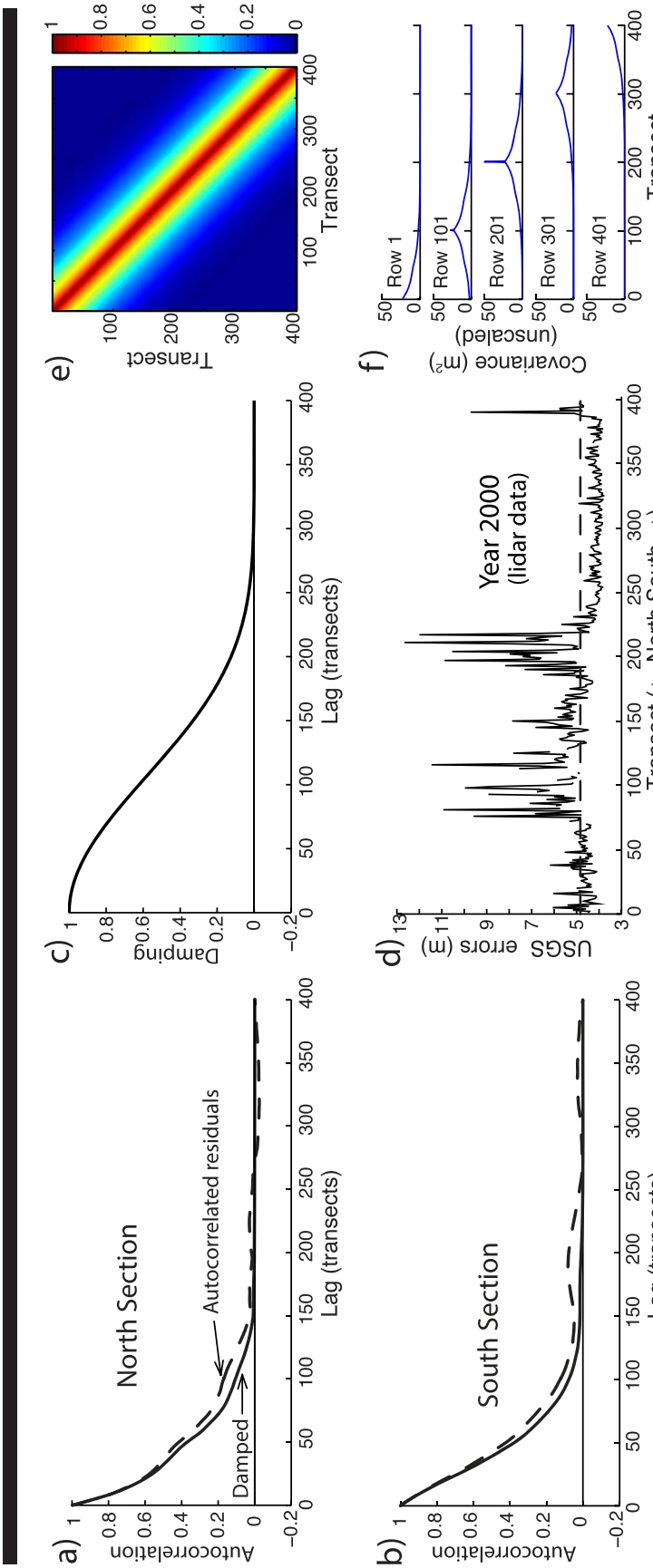

Figure A1. (a-b) Autocorrelation (dashed) and damped autocorrelation (solid) of residuals for north (a) and south (b) portions of the study area, (c) the cosine damping function ($l = 6$), and (e) the resulting autocorrelation matrix for each year in the north section. (d) The USGS error estimates for the 2000 LIDAR shoreline (north section) fluctuate alongshore (solid line). The dashed line indicates their geometric mean of 4.8 m. (f) Selected rows of the correlation matrix in (e) after weighting by the square of the geometric mean in (d) and addition of spatially independent errors that exceed the mean in (d). For years in which *a priori* shoreline error estimates do not vary alongshore (shorelines not derived from LIDAR data), the entire block is simply scaled by the constant. (Color for this figure is available in the online version of this paper.)