Journal of Coastal Research	31	2	464–476	Coconut Creek, Florida	March 2015

Long-Term Shoreline Change at Kailua, Hawaii, Using Regularized Single Transect

Tiffany R. Anderson*, L. Neil Frazer, and Charles H. Fletcher

Department of Geology and Geophysics
School of Ocean and Earth Science and Technology
University of Hawaii at Manoa
Honolulu, HI 96822, U.S.A.

www.cerf-jcr.org

www.JCRonline.org

ABSTRACT

Anderson, T.R.; Frazer, L.N., and Fletcher, C.H., 2015. Long-term shoreline change at Kailua, Hawaii, using regularized single transect. *Journal of Coastal Research,* 31(2), 464–476. Coconut Creek (Florida), ISSN 0749-0208.

Traditional long-term (decadal) and large-scale (hundreds of kilometers) shoreline change modeling techniques, known as single transect, or ST, often overfit the data because they calculate shoreline statistics at closely spaced intervals along the shore. To reduce overfitting, recent work has used spatial basis functions such as polynomials, B splines, and principal components. Here, we explore an alternative to such basis functions by using regularization to reduce the dimension of the ST model space. In our regularized-ST method, traditional ST is an end member of a continuous spectrum of models. We use an evidence information criterion (EIC = −2 times the log of the prior predictive distribution) to select the optimal value of the regularization parameter, instead of the usual L-curve method, because the EIC can also be used to evaluate basis function models yet does not require counting model parameters. To test the method, we apply it to historical shoreline data from Kailua, Hawaii, comparing the results with those from B splines (basis functions) and traditional ST. As expected, the regularized-ST and B-spline models both give shoreline change rates that vary more smoothly alongshore than the rates from ST. The regularized-ST model, along with the B-spline model, also shows significantly better predictive capability over the traditional ST model from a fivefold cross-validation. The regularized-ST model is more straightforward to implement than splines and may be attractive to users because of its continuous connection with the familiar ST method.

ADDITIONAL INDEX WORDS: *Shoreline change rates, coastal erosion, Tikhonov regularization, B splines, beaches, sediment processes*.

INTRODUCTION

Historical shoreline change studies provide coastal managers with data useful for building resilient communities. Many jurisdictions rely on quantitative measures of shoreline behavior, such as shoreline change rates, to implement building setback policies and other decision-making tools. Therefore, shoreline change statistics are often required over long stretches of coastline (hundreds of kilometers).

The complexity of coastal dynamics presents challenges to long-term shoreline prediction, and the temporal sparseness and short time windows of available data further complicate this problem (Hanson and Kraus, 1989; Miller and Dean, 2004; Morton, 1979; Stive *et al*., 2002). Although physics-based models can simulate shoreline behavior over short time spans, they are limited in their long-term predictive capabilities (De Vriend *et al*., 1993; Van Rijn *et al*., 2003) and often require quality data that are not available over large areas. Therefore, long-term and large-scale (hundreds of kilometers) shoreline change models are often simple empirical models that require a time series of shoreline data that is long enough to reveal trends over decades; data sets of this nature are typically sparse in time and contain large scatter because of short-term beach processes (*e.g*., Galgano and Douglas, 2000; Honeycutt, Crowell, and Douglas, 2001; *etc*.).

Least squares regression is commonly used to fit a simple empirical model (often a straight line) to historical shoreline data to identify a long-term trend (*e.g*., Crowell, Douglas, and Leatherman, 1997; Galgano and Douglas, 2000; Honeycutt, Crowell, and Douglas, 2001). Genz *et al*. (2007) investigated different forms of regression analysis, such as least absolute deviation, weighted least squares, and least median of squares. Frazer, Anderson, and Fletcher (2009) added a storm function to the typical rate-only regression equations, improving long-term shoreline change estimates. Improvements in quantifying error and bias in data derived from different sources (*e.g*., light detection and ranging, video imagery, and aerial photographs) have also improved the accuracy of shoreline change statistics (Douglas and Crowell, 2000; Fletcher *et al*., 2003; Plant *et al*., 2007; Ruggiero and List, 2009). As more data become available (*i.e*. wave height and wind speed), modeling efforts progress toward assimilating the various time resolutions (*e.g*., Davidson, Splinter, and Turner, 2013; Long and Plant, 2012) and combining multiple data inputs to provide more robust predictions (Gutierrez, Plant, and Thieler, 2011; Hapke and Plant, 2010; Yates and Le Cozannet, 2012).

The techniques mentioned earlier focus on shoreline change at independent locations along the shore. The widely used single-transect (ST) method, for example, calculates a shoreline change rate at each shore-normal transect (Figure 1a). ST assumes that the data and noise at each transect are

DOI: 10.2112/JCOASTRES-D-13-00202.1 received 29 October 2013; accepted in revision 8 July 2014; corrected proofs received 8 September 2014; published pre-print online 16 October 2014.
*Corresponding author: tranders@hawaii.edu

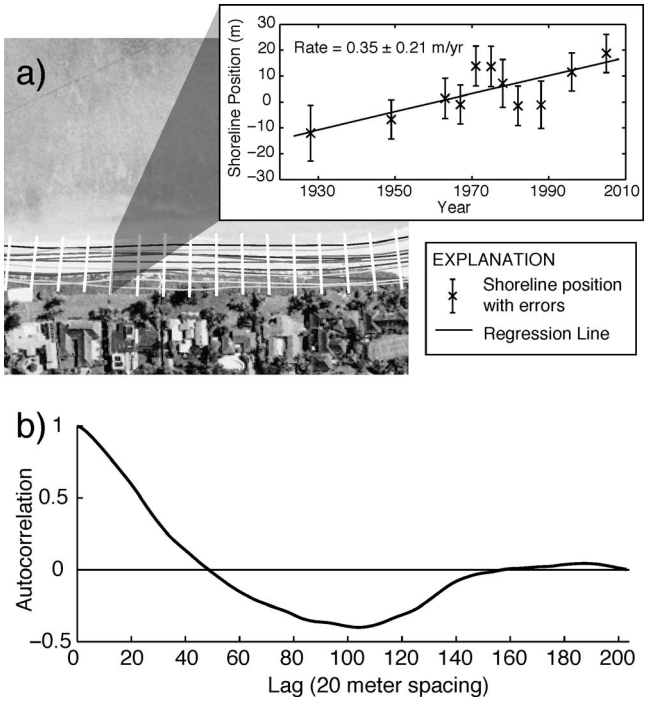

Figure 1. (a) Weighted least squares linear regression of shoreline data along one transect is used to produce a shoreline change rate, independent of shoreline behavior at adjacent transects. (b) The gradual decay of the autocorrelation of independently calculated shoreline change rates at 20-m increments along Kailua Beach shows that the rates are highly correlated in the alongshore direction.

independent, ignoring correlations between the rates at different transects. Figure 1b shows the autocorrelation of long-term rate parameters that were calculated independently at 20-m increments (transects) along Kailua Beach, Hawaii. The slow decay of the autocorrelation in the alongshore direction shows that rates are highly correlated. In this case, the ST method uses more parameters to fit the data than are necessary. Similarly, Tebbens, Burroughs, and Nelson (2002), who used wavelets to perform a spatial analysis of short-term shoreline change along the outer banks of North Carolina, found that the low-frequency portions of the signal along the shore dominated over the high-frequency portions and contained correlation at large spatial scales.

Frazer, Genz, and Fletcher (2009) and Genz, Frazer, and Fletcher (2009) reduced this overfitting by representing change rates as the sum of alongshore basis functions (Legendre polynomials, trigonometric functions, and principal components). They followed Fenster, Dolan, and Elder (1993) by using an information criterion to determine the number of basis functions to use. However, Frazer, Genz, and Fletcher (2009) found that smooth basis functions, such as Legendre polynomials and trigonometric functions, were problematic when they attempted to model rates that spiked or jumped along the shore, as in their example of Waihee, Maui, that contains a spike in local accretion. In general, the authors concluded that smooth basis functions produce ringing (Gibbs effect) in the alongshore direction when they are used to model parameters that vary suddenly along the shore. They added that the principal component basis function method does not produce ringing because alongshore discontinuities are represented in the lower-order basis functions. Anderson and Frazer (2014) noted that although principal component basis functions circumvent ringing, they are contaminated by noise from measurement errors and short-term processes.

As an alternative, Anderson and Frazer (2014) investigated B-spline basis functions as a way to avoid the Gibbs effect. B splines are locally occurring basis functions, each resembling a Gaussian curve; thus, they provide flexibility in the amount of variation allowed in parameters alongshore, including alongshore discontinuities. However, model selection can be awkward because adding one more spline in the obvious way (equally spaced splines) causes the center points of other splines to shift, which can increase the misfit rather than reducing it. The regularized-ST method of this paper avoids that problem by making the spectrum of models continuous rather than discrete. Anderson and Frazer (2014) further noted that selecting the basis functions to use in a shoreline analysis depends on the geology of a study area and the analysis objective and that no single type of basis function was superior in all situations. Here, we introduce an alternative to using basis functions, contributing to the shoreline modeler's toolbox of analysis methods. We also investigate an alternative model selection procedure that may have advantages compared with more widely known criteria such as the Akaike information criterion.

The aim of this study is (1) to present an alternative approach to estimating long-term trends in shoreline change, an approach that takes into account spatial correlation and shows practical advantages over basis function methods, and (2) to compare the new method with the traditional ST method and with the B-spline basis function method presented by Anderson and Frazer (2014). As an additional point of interest, we use the evidence (sometimes called the prior predictive distribution or marginal likelihood) to select the best model within each model family. Evidence provides a particularly convenient method of model selection when the spectrum of models is continuous yet retains the ability to compare models within countable model families such as splines.

METHODS

Here, we use a technique called regularization, which has long been used in regression analysis to avoid overfitting (*e.g.*, Aster, Borchers, and Thurber, 2012; Press *et al.*, 2007). Second-order Tikhonov regularization in particular involves penalizing solutions for roughness but not for locally linear behavior. As an example, consider Figure 2, in which the circles show rate parameters calculated using the ST method at transects spaced 20 m apart for Kailua Beach. Although rates from the ST model show detail at high spatial frequencies along the shore, much of that detail is suspect because it is not repeated in surveys at other times. A simpler relationship results when the difference between ST rates at adjacent transects is constrained to be constant, giving the smoothest representation possible: a straight line (Figure 2, dashed line). That alongshore-linear rate model does not fit the data nearly as well as the ST model,

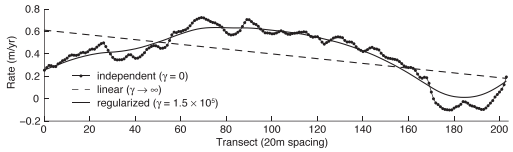

Figure 2. Rates from ST (dots) fit the data well but capture noise within the data. Conversely, rates from regularized ST with a very large γ (dashed line) limit alongshore variability but give a poor fit. Rates from regularized ST with a moderate γ (solid line) allow alongshore variability while limiting shorter wavelength variations. Positive rates indicate accretion, and negative rates indicate erosion.

but it does not overfit the data. The preferred model, in this example, lies somewhere between the two extremes of fitting the data well at each transect and limiting the alongshore variability. In Figure 2, the solid line represents the rate in which alongshore variation is constrained by a fairly large penalty. Determining the appropriate amount of allowable alongshore variability can be achieved using a model selection criterion, which is addressed later. In this paper, we use second-order Tikhonov regularization to make the traditional ST method more parsimonious by forcing model parameters to vary more smoothly alongshore. Kailua Beach, Hawaii, is used as a case study.

Physical Setting

Kailua Beach is located on the windward side of the Hawaiian island of Oahu (Figures 3a and b). The 4-km carbonate sand beach is bounded by limestone at Kapoho Point to the north and basalt at Alala Point to the south. A wide fringing reef platform provides moderate protection to the beach from year-round NE trade wind waves and winter (October–March) N swells. The reef platform is bisected by a winding 200-m-wide sand-floored channel that widens toward the shore into a broad sand field at the center of the beach. The residential area of Kailua sits upon a low-lying expansion of Holocene-age carbonate dune ridges and terrestrial lagoon deposits (Harney and Fletcher, 2003). Low, vegetated dunes front many oceanfront homes (Figure 4a). Kaelepulu Stream empties into the ocean near the south end of Kailua Beach. Episodic removal and occasional redistribution of sand near the stream mouth began in the 1980s. A boat ramp, constructed between 1949 and 1963 at the south end of Kailua Beach, generally inhibits sediment movement toward the north, as evidenced by sediment accumulation on the south side and deprivation on the north of the boat ramp (Figure 4c) in historical aerial photos (available online from University of Hawaii Coastal Geology Group, 2014).

Historical Shoreline Data

The data used in the study are cross-shore distances relative to a user-defined baseline (Figure 1a). Historical shorelines extracted from aerial photographs and one topographic (T) sheet dating between 1928 and 2005 were obtained from the University of Hawaii Coastal Geology Group as geographic information system shapefiles (see Romine *et al.*, 2009, for shoreline extraction procedures). Total position errors, as calculated by Romine *et al.* (2009), range from 7.35 to 9.22 m for shorelines that were derived from aerial photos and 10.78 m for the shoreline derived from the T sheet. Approximately shore-normal transects, spaced 20 m apart, are cast off of a smooth baseline that follows the general shape of the shoreline (Figures 3b and d). Here, the baseline is a spline fitted to all shoreline positions simultaneously. At each transect location x_i along the baseline, the relative distance from the shoreline to the baseline y_{ij} is calculated for shoreline times t_j. Thus, $y_{ij} = y(x_i, t_j)$, where transect index i ranges from zero to $I - 1$, and time index j ranges from one to J. The Kailua Beach data set is the collection of cross-shore positions relative to the baseline over all transects (Figure 3c).

Regularized-ST Model

The basic procedure of the new regularized-ST model is as follows (also see Figure 5):

(1) Define the forward shoreline change model at each transect: *e.g.*, $y = b + r\,t$.
(2) Assign an array of candidate values for each regularization parameter (*e.g.*, rate and intercept), regularly spaced over a search interval.
(3) For each combination of regularization parameters:
  (a) Fit the forward model to the data using linear regression.
  (b) Calculate the evidence information criterion (EIC).
(4) Determine the optimal model (indexed by regularization parameters) by identifying the model with the smallest EIC.
(5) Extrapolate from the optimal model to predict future shoreline positions.

Details of the procedure in the following sections include application to the Kailua Beach study area.

Forward Model

We use the following simple forward model for shoreline change over time:

$$y_{ij} = b_i + r_i(t_j - \bar{t}) + n_{ij}, \quad (1)$$

in which b_i and r_i are the intercept and rate, respectively, at alongshore location x_i, $\bar{t}$ is the mean of shoreline survey dates, and n_{ij} is noise. It is straightforward to include additional terms in the mathematical forward model, such as acceleration (*e.g.*, Frazer, Genz, and Fletcher, 2009; Romine *et al.*, 2009) or a storm function (Anderson, Frazer, and Fletcher, 2010). We opted not to include acceleration here because it can result in unstable long-term (decades) predictions. Likewise, we found no storm signal in the Kailua Beach data set.

Second-Order Tikhonov Regularization (Linear Regression)

Least squares regression is used to estimate parameter values (rates and intercepts) within the forward model. To regularize the problem, constraint equations are added to the system of shoreline equations in Equation (1). The influence that the constraint equations have on the estimated rates and intercepts is governed by two regularization parameters: one for smoothing the rates and the other for smoothing the intercepts. Because there is no

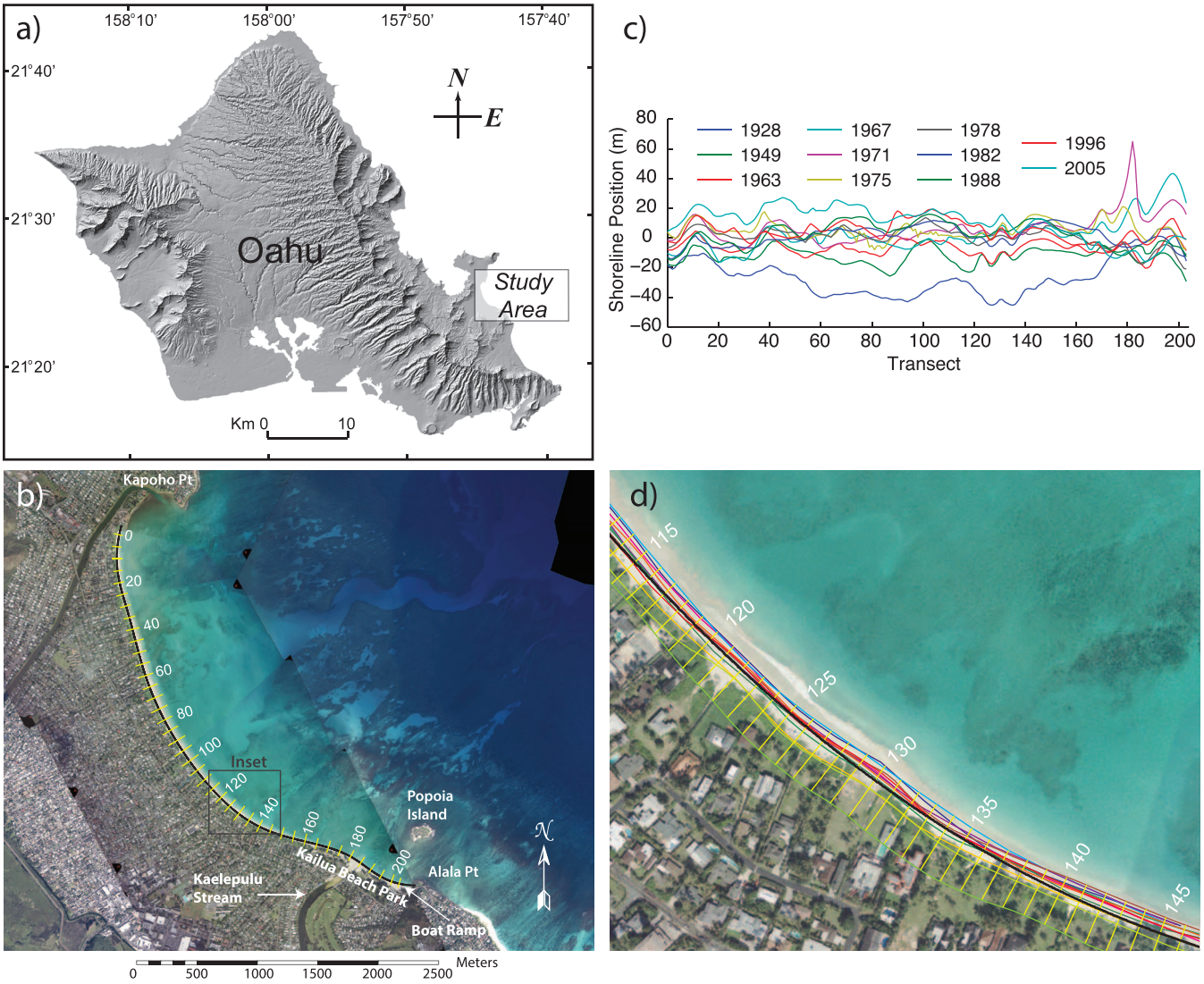

Figure 3. (a) The Kailua beach study area is located on the Hawaiian island of Oahu. (b) A sand-filled paleochannel bisects the shallow fringing reef as seen in a 2005 aerial photomosaic (from University of Hawaii Coastal Geology Group, 2014). Shore-normal transects (thin yellow lines) are cast off a baseline (black line) that follows the general shape of the coast. Only every fifth transect is shown for clarity. (c) Shoreline data from Kailua, Hawaii, are shown relative to the baseline. Transects are spaced 20 m apart. (d) Historic shoreline vectors display small-scale deviation from the baseline, following the nearshore topography.

simple formula for selecting regularization parameter values, we test a range of values.

We first treat the case of a single regularization parameter for rates; the regularization parameter for intercepts is handled similarly. Consider the system of equations in Equation (1) where all b_i are zero (shoreline positions at each transect are centered about their mean). In matrix form, this system is $d = Gm + \eta$, where d is an $N \times 1$ column vector of N observed shoreline positions, m is an $M \times 1$ column vector of M model parameters, G is an $N \times M$ system matrix, and η is an $N \times 1$ column vector of errors with zero mean and covariance matrix C_d. Model roughness is controlled by augmenting the system with equations that set the second alongshore derivative of the rate parameters to zero. The second derivative is approximated by the centered second difference matrix L:

$$L = \frac{1}{(\Delta x)^2}\begin{bmatrix} 1 & -2 & 1 & & & \\ & 1 & -2 & 1 & & \\ & & & \cdots & & \\ & & & 1 & -2 & 1 \end{bmatrix}, \qquad (2)$$

where Δx is the spacing between transects. Thus, Lm approximates the second derivative of the parameter vector m. Although the second derivative is not calculated at endpoints, the endpoints are used in calculating the second derivative of adjacent points, thus forcing the endpoints to

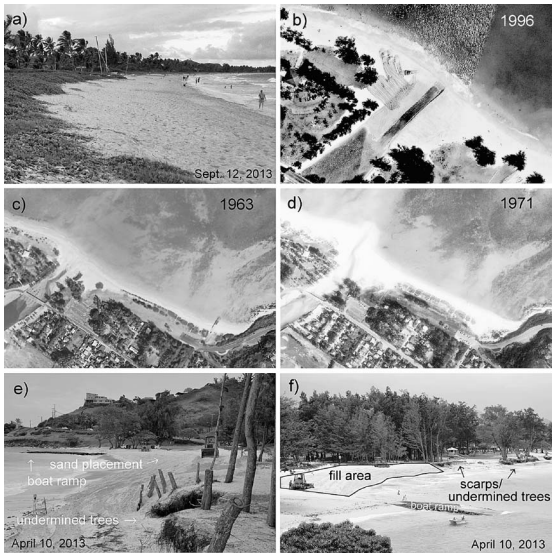

Figure 4. (a) Low, vegetated dunes front homes along Kailua Beach. Photo is looking north from transect 78. (b) Grading and sand redistribution at the Kaelepulu Stream mouth in 1996. (c) Kailua boat ramp (bottom, right) obstructs sediment movement toward the NW. (c) and (d) Variability in Kaelepulu Stream flow alters the shoreline. (e) and (f) Recent erosion at Kailua Beach Park causes undermining of trees and scarping of the beach face; subsequent mitigation attempts include nourishing the beach with excess sand that has accumulated near the Kaelepulu Stream mouth.

adhere to prescribed smoothing. In coding, the equation $Lm = 0$ becomes $\mathbf{0} = \gamma Lm + \varepsilon$, where $\mathbf{0}$ is an $(M-2)\times 1$ column vector of zeros, ε is an $(M-2)\times 1$ column vector of independent and identically distributed errors with zero mean and unit variance (*i.e.* the covariance matrix is the identity matrix), and γ is the regularization parameter.

Together, the forward shoreline model equations and the roughness constraints are

$$\begin{aligned} d &= Gm + \eta \\ \mathbf{0} &= \gamma Lm + \varepsilon, \end{aligned} \tag{3}$$

with $\eta \sim N(0, C_d)$ and $\varepsilon \sim N(0, I)$. It follows that the least squares solution to the augmented system of equations in Equation (3) is found (*e.g.*, Hastie, Tibshirani, and Friedman, 2009) by minimizing

$$\phi(m) = (d - Gm)^T C_d^{-1}(d - Gm) + \gamma^2 m^T L^T L m. \tag{4}$$

It can be seen that when $\gamma = 0$, the regularization terms disappear, leaving only the sum of squared residuals typically seen in least squares regression. As γ approaches infinity, the second derivative penalty term is given so much weight compared to the residuals that the parameters are forced to vary linearly in the alongshore direction.

Taking the differential of (4) and setting it to zero, we find that the model vector m that minimizes (4) for a given γ is

$$\widehat{m} = (G^T C_d^{-1} G + \gamma^2 L^T L)^{-1} G^T C_d^{-1} d \tag{5}$$

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

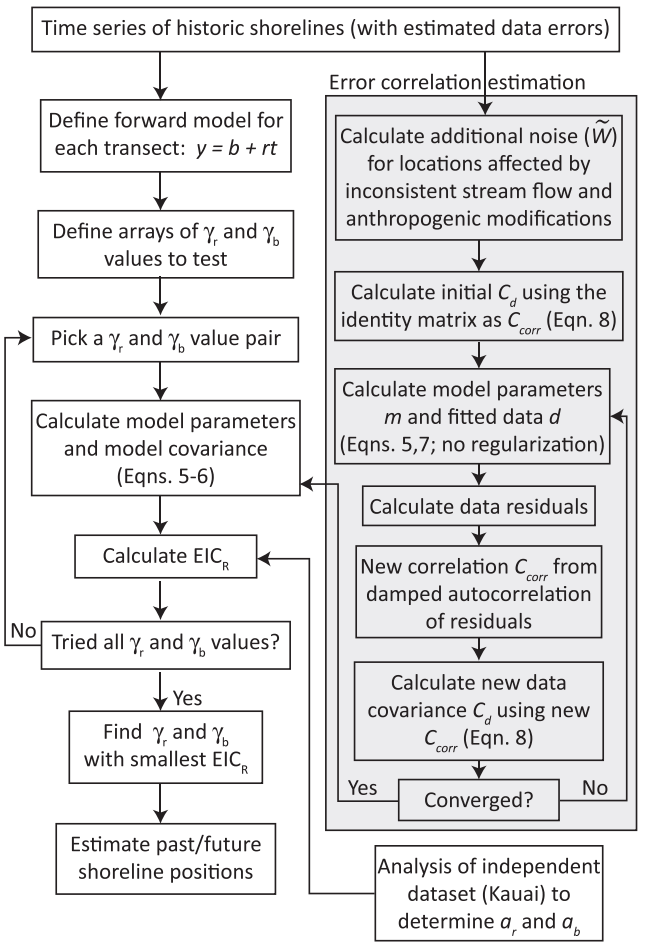

Figure 5. The flowchart for the regularized-ST model procedure includes estimating data error correlation (shaded region).

$$\widehat{C}_m = (G^T C_d^{-1} G + \gamma^2 L^T L)^{-1}. \tag{6}$$

The vector of predicted data is then

$$\begin{aligned} \widehat{d} &= G(G^T C_d^{-1} G + \gamma^2 L^T L)^{-1} G^T C_d^{-1} d \\ &= Hd \end{aligned} \tag{7}$$

where H is the data resolution matrix, also known as the "hat" matrix.

As mentioned earlier, we use two regularization parameters because we smooth rates and intercepts independently. Including the second regularization parameter for intercepts is straightforward, resulting in matrix equations similar to Equation (3), except that the system of equation $d = Gm + \eta$ now includes intercept parameters b_i, and the expression γL in the second line is replaced by a block diagonal matrix with blocks $\gamma_r L$ and $\gamma_b L$, where γ_r and γ_b are the rate and intercept regularization parameters, respectively. It can be shown that Equations (5) to (7) are altered only by replacing the matrix $\gamma^2 L^T L$ with the block diagonal matrix containing blocks $\gamma_r^2 L^T L$ and $\gamma_b^2 L^T L$.

Correlated Noise

Although data errors are assumed *a priori* for each data value, we estimate the correlations between these errors. Our procedure for this is summarized in the shaded region of our procedural flowchart (Figure 5). Because of the large time intervals between surveys (decadal), data are not considered correlated in time. However, data correlation in the alongshore direction is present because of the small spacing (20 m) between transects. Treating correlated data as independent "can cause dramatic differences in the inferences which may be legitimately drawn from a set of observations, Box (1954), Zellner and Tiao (1964)" (from Box and Tiao, 1973, p. 81). Tests in Anderson and Frazer (2014) confirm that incorrect assumptions about data independence lead to poor estimates of parameters and their uncertainties in the shoreline problem.

To apply a correlated noise model to a single sample of data, it is necessary to make some stationarity assumptions about the noise. Accordingly, in the correlated noise model for the shoreline problem, errors are assumed to be stationary in the alongshore direction. Shoreline behavior fronting Kailua Beach Park, however, challenges this assumption, where inconsistent stream flow (Figures 4c and d), beach grading (Figure 4b), and nourishment fronting Kailua Beach Park (transects 170–203) cause variability in shoreline data that is not explicitly modeled. We incorporate some of these influences into the modeling procedure as additional noise that is independent of the correlated noise (see Anderson and Frazer, 2014) in the following way. The matrix containing data errors is defined as

$$C_d = \bar{W}^{1/2} C_{corr} \bar{W}^{1/2} + \tilde{W} \tag{8}$$

in which C_{corr} is an $N \times N$ matrix containing the estimated data error correlations, $\bar{W}^{1/2}$ is a diagonal matrix containing the data errors as estimated by Romine *et al*. (2009), and $\tilde{W}^{1/2}$ is a diagonal matrix of additional data errors. The term $\bar{W}^{1/2} C_{corr} \bar{W}^{1/2}$ represents correlated data errors, while $\tilde{W}$ represents spatially independent noise. We define $\tilde{W} = diag(e_{ij}^2 [\hat{\alpha}_i - 1]_+)$, in which e_{ij} are the data error estimates from Romine *et al*. (2009) for locations i and times j, $\hat{\alpha}_i$ represents the scaling factors for locations i (Equation (16), shown later), and the subscript + denotes positive values, indicating that additional errors are only included if the scaling factor exceeds one (*i.e*. if data errors estimated from regression are larger than *a priori* estimates).

The spatial correlations C_{corr} are derived using a technique that is fully described in Anderson and Frazer (2014); only a brief description is given here. The iterative procedure begins by calculating the first covariance matrix (Equation [8]) using the identity matrix as the initial C_{corr}. The new covariance matrix is then used in the least squares procedure to find the model parameter vector (Equation [5]), and subsequently the vector of fitted data (Equation [7]), to produce residuals, defined as the difference between the fitted data and the observed data. The autocorrelation function of the residuals is estimated, damped with a cosine taper function, and then used as the new data error correlations (see Appendix A in Anderson and Frazer, 2014). In all cases, the new correlation matrix C_{corr} has ones along the diagonal and entries that decline in size with distance from the diagonal. The correlation matrix from the previous iteration is then used to provide new residuals that produce a new correlation matrix, and so on, until correlation matrices from successive iterations are sufficiently close. Our stopping criterion is $||C_{corr}^{(u)} - C_{corr}^{(u-1)}||_2 < 10^{-3} ||C_{corr}^{(u-1)}||_2$, where u is the number of iterations and $||\ldots||_2$ denotes the matrix 2-norm, also called the spectral norm.

Our model selection criterion (see the next section) favors models with correlated data errors because if data errors are more correlated, then effectively fewer parameters are used in representing the error part of the model while more parameters are used for the rate and intercept part of the model; thus, a model with more variable rates and intercepts may give a lower EIC than a simple model. To favor simple models, we estimate error correlations from the residuals in ST (regularization parameters are zero) and use those correlations in all subsequent calculations for regularized-ST and spline models.

Selecting the Regularization Parameters

The appropriate regularization parameters are determined by means of the EIC given in the appendix. As with other information criteria, the lower the EIC, the better the model. We cannot minimize the EIC analytically, so we find its minimum by searching over a range of regularization parameter values that give results similar to ST at the lower limit and alongshore-linear behavior at the upper limit (*e.g*., Figure 2). For the Kailua Beach study area, regularization parameters for rates range from 1.8×10^3 to 3.3×10^6 m·y and from 55 to 2.1×10^5 m for intercepts.

The EIC is defined as $\mathrm{EIC} = -2\ln(p(d|\psi))$, where $p(d|\psi)$ is the evidence, the probability of the data vector d given model ψ. Using marginalization and the product rule, the evidence can be written in the form (Sivia and Skilling, 2006, p. 79)

$$p(d|\psi) = \int p(d|m,\psi) p(m|\psi) dm, \tag{9}$$

in which $p(d|m, \psi)$ is the likelihood of parameter vector m and $p(m|\psi)$ is the parameter prior for model ψ. Because the formula for the likelihood is different for the regularized-ST and spline model families, we provide an EIC formula for each.

The EIC is applied to regularized-ST models by substituting the likelihood for regularized-ST models into Equation (9) and simplifying (see the appendix), giving the explicit formula for regularized-ST models:

$$\begin{aligned} \mathrm{EIC_R} &= (d - G\widehat{m})^T C_d^{-1} (d - G\widehat{m}) + \gamma_r^2 \widehat{r}^T L^T L \widehat{r} + \gamma_b^2 \widehat{b}^T L^T L \widehat{b} \\ &\quad + \ln(|C_m^{-1}|) + (N-4)\ln(2\pi) + M\ln(a_r a_b) \\ &\quad + (4-M)\ln(\gamma_r \gamma_b) + \ln(|C_d|). \end{aligned} \tag{10}$$

Here, $|\ldots|$ denotes the matrix determinant, C_m is the model covariance matrix (Equation [6]), L is the second derivative operator (Equation [2]) with $M/2 - 2$ rows, γ_r is the regularization parameter for rate, γ_b is the regularization parameter for intercept, M is the length of the parameter vector m, N is the length of the data vector d, and a_r and a_b represent lengths related to rate and intercept, respectively, that are sufficiently large to cover the range of potential rates and intercepts (see the appendix). For Kailua, 4 m/y is used for a_r

and 300 m is used for a_b because these values exceed the range of rates and intercepts, respectively, calculated using the ST method for 3701 locations spanning all sandy beaches on the Hawaiian island of Kauai, excluding the Na Pali coastline.

Substituting the likelihood for basis function (*e.g.*, B splines) models into Equation (9) and simplifying (see the appendix) gives the explicit equation for basis functions models:

$$\begin{aligned}\mathrm{EIC_B} &= (d - G\widehat{m})^T C_d^{-1}(d - G\widehat{m}) - \ln(|C_m|) + (N - M)\ln(2\pi) \\ &\quad + 2\,[Mr\ln(a_r) + Mb\ln(a_b)] + \ln(|C_d|), \end{aligned} \tag{11}$$

in which the system matrix G has Mr basis functions for rate and Mb basis functions for intercept, one in each column. For purposes of model selection, the covariance matrix C_d is assumed constant for all regularized-ST and spline models, so the last term can be omitted. The ST model has a covariance matrix that is diagonal (assumes temporal and spatial independence) for consistency with traditional methodology.

Because the numerical values of the regularization parameters have no intuitive value, it is helpful for understanding, although unnecessary for model selection, to estimate the effective number of model parameters associated with each regularization parameter. In regression methodology, this number is called the regression degrees of freedom. For basis function methods, this number is simply the number of basis functions, *i.e.* the number of components in the parameter vector m. For regularization, the length of the parameter vector is fixed at the number of transects, but as the regularization parameters increase in value, the effective number of parameters declines. The regression degrees of freedom df_γ is defined as the trace of the data resolution matrix (hat matrix) given in Equation (7) (*e.g.*, Hastie, Tibshirani, and Friedman, 2009), or

$$df_\gamma = trace(H). \tag{12}$$

Shoreline Prediction

Future shoreline location is predicted by extrapolating from the model whose parameters we have just estimated. For future time t_f and location x_i, let $q_{x_i} = q_{x_i}(t_f)$ be an $M \times 1$ column vector, which we refer to as the prediction kernel. The vector $q_{x_i}^T$ resembles a row of the system matrix G, except that the time used is $t_f - \bar{t}$, following Equation (1). The predicted position of the shoreline at location x_i and future time t_f is then

$$\widehat{y}(x_i, t_f) = q_{x_i}^T \widehat{m} \tag{13}$$

with estimated variance

$$\widehat{\sigma}_{x_i}^2 = q_{x_i}^T \widehat{C}_m q_{x_i}. \tag{14}$$

A $100(1 - \varepsilon)\%$ confidence interval for $\hat{y}$ is given by

$$\widehat{y}(x_i, t_f) = q_{x_i}^T \widehat{m} \pm z_{1-\varepsilon/2}\widehat{\sigma}_{x_i} \tag{15}$$

in which $z_{1-\varepsilon/2} = \Phi^{-1}(1 - \varepsilon/2)$, where Φ is the cumulative distribution function of the standard normal. When the data variance is estimated, Student's t distribution is used to estimate confidence intervals for parameters. Here, we assume a known variance, using data error estimates calculated by Romine *et al.* (2009), so the posterior distribution is normal. If the data uncertainties are regarded as process errors rather than measurement errors, then the data uncertainty should be added to the right-hand side of Equation (14) for Equation (15) to represent the mean and standard deviation (SD) of the posterior predictive distribution (*e.g.*, Gelman *et al.*, 2014, p. 357).

Assessment of Models' Predictive Capabilities

Having selected the appropriate regularization parameters and number of spline basis functions to use in regularized-ST and spline models, we us fivefold cross-validation (CV) to assess the predictive capabilities (Geisser, 1975; Kohavi, 1995; Hastie, Tibshirani, and Friedman, 2009, Section 7.10) of these models and the traditional-ST model. The data are split into five approximately equal parts. For each one-fifth part, we fit the model to the remaining four-fifths of the data (training data) and then calculate the mean of the squared difference (the mean-square error, or MSE) between the actual and the predicted one-fifth parts (test data). The CV measure of the prediction error is the mean of the five MSEs (Hastie, Tibshirani, and Friedman, 2009, Section 7.10; Markatou *et al.*, 2005; Nadeau and Bengio, 2003). We use the formula $(1/5 + n_t/(N - n_t)) \cdot s^2$ for the variance of the CV measure, which somewhat accounts for the correlation in MSEs because of overlapping data in training data sets (Nadeau and Bengio, 2003). Here, N is the total number of data, n_t is the number of data in each test data set, and s^2 is the sample variance of the five MSEs.

RESULTS

The ΔEIC scores, the EIC scores relative to the lowest (best) score, for all pairs of regularization parameters tested are shown in Figure 6a. The regularized-ST model with the lowest EIC score has a rate regularization parameter of 4.8×10^5 m·y and an intercept regularization parameter of 1.1×10^3 m. Figure 6b shows the ΔEIC values from regularized-ST models as in Figure 6a, but they are plotted against corresponding regression degrees of freedom for rate and intercept parameters. The regularized-ST model with the lowest EIC score for Kailua Beach has roughly 4 (4.3) regression degrees of freedom for rate parameters and 64 regression degrees of freedom for intercept parameters. The most parsimonious spline model (Figure 6c) has 3 rate basis functions and 18 intercept basis functions. In contrast, the traditional ST model has 204 rate parameters and 204 intercept parameters: 1 rate and 1 intercept at each transect.

The CV measures of the prediction error (prediction error is defined as the squared difference between model predictions and independent test data) are shown in Table 1 for the traditional-ST model (103.0 ± 11.1 m^2), the best regularized-ST model (64.4 ± 5.2 m^2), and the best spline model (65.4 ± 5.9 m^2). For comparison, squared shoreline data *a priori* errors had a median value of 59 m^2 and ranged from 54 to 85 m^2 for the 10 shorelines derived from aerial photos and 116 m^2 for the National Oceanic and Atmospheric Administration T sheet–derived shoreline. Both the regularized-ST and the spline model show significant improvement in prediction over the traditional-ST method at the 95% confidence level using a Student's t test. There is no significant difference between the regularized-ST and the spline model CV scores.

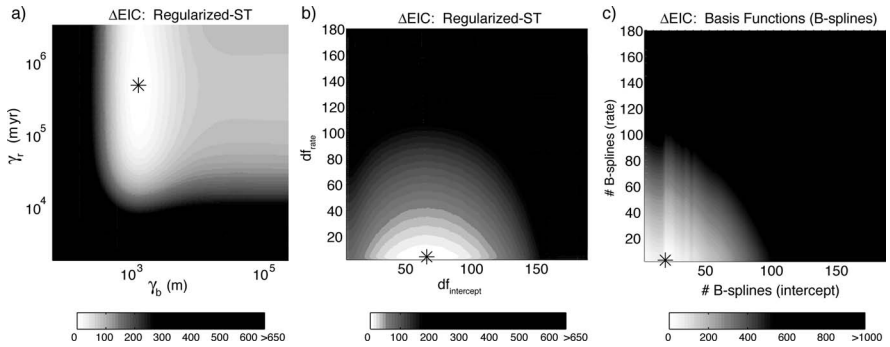

Figure 6. The ΔEIC values for regularized-ST models are shown plotted against (a) the regularization parameters and (b) the regression degrees of freedom for comparison with B-spline models. Asterisks indicate minimum EIC scores (best model) in that model family. (c) The ΔEIC values for the B-spline models are plotted against the number of B splines used for intercept and rate.

Predicted shoreline positions for the time span 1928–2075 are shown with historic data for the ST model (Figure 7a) and regularized-ST model (Figure 7b). The plots reveal the relatively high spatial frequency of the ST model compared to the more parsimonious regularized-ST model. Shoreline positions for the spline model, not shown, look similar to those for regularized ST because of the similar alongshore frequencies of their modeled parameters.

There is general agreement, as shown in Figure 8a, among the shoreline change rates in the alongshore direction, calculated by the three methods that we tested. Maximum accretion rates for each of the three methods are between 0.59 and 0.72 m/y (Table 1) and are located near the center of the beach. Romine *et al*. (2009) analyzed Kailua Beach data using polynomial basis functions with similar results. Areas of accretion, indicated by all three models, are also consistent with the occurrence of seaward-growing vegetated dunes that have formed on the ocean side of many coastal properties and with the documented seaward migration of the vegetation line from 1949 to 1978 (Hwang, 1981).

Of the three modeling techniques that we tested, rates calculated using the ST method have the most high-frequency variation alongshore, as expected, while the regularized rates and spline rates both have a significantly smoother, long wavelength signal (Figure 8a). The similar "smoothness" of the modeled regularized rates and spline rates is a reflection of the roughly equivalent regression degrees of freedom (four) and number of basis functions (three), respectively, for the model rates.

The intercept parameters (Figure 8b), which represent mean shoreline positions relative to the baseline, are more resolved in the alongshore direction than are rates in all three methods. Although the spline model uses fewer parameters for the intercept (18) compared to the regularized-ST model (64), the spatial signals of intercepts for these models are similar alongshore (Figure 8b).

Table 1. *For each method, the number of parameters* Np *or regression degrees of freedom* df *are given for rate and intercept. For the regularized-ST method, the regularization parameter* γ *is given in brackets. Minimum and maximum rates calculated by the different methods are given, along with corresponding transect locations, in brackets. As expected, the difference between maximum and minimum rates is largest with ST because of the alongshore smoothing inherent in the other methods. The CV measures of prediction error and their estimated SDs* σ_{cv} *for each method show the considerable prediction improvement of the regularized-ST and spline models over the traditional-ST model.*

	ST	Regularized ST	Spline
Np_r (rate)/df_r [γ_r (m·y)]	204	4 [4.8×10^5]	3
Np_b (intercept)/df_b [γ_b (m)]	204	64 [1.1×10^3]	18
Max. rate (m/y) [transect]	0.72 [69]	0.59 [85]	0.59 [83]
Min. rate (m/y) [transect]	−0.10 [178]	−0.04 [203]	−0.12 [203]
CV ± σ_{cv} (m^2)	103.0 ± 11.1	64.4 ± 5.2	65.4 ± 5.9

DISCUSSION

Alongshore variations in parameter values can sometimes be linked to geologic features. For example, between transects 120

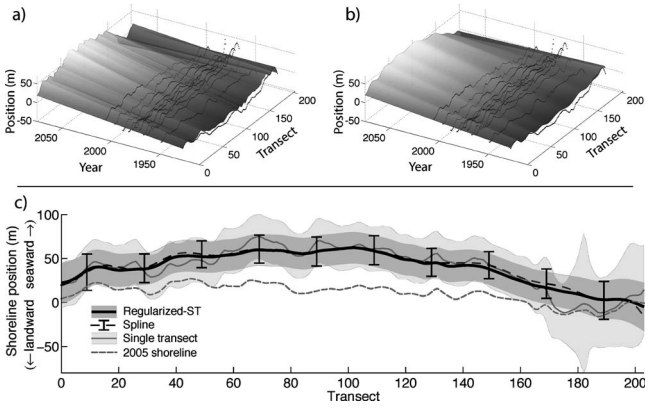

Figure 7. Estimated shoreline positions for (a) the ST method and (b) the regularized-ST method are shown, along with the data (black circles). As expected, none of the models fit the data well near the Kaelepulu Stream mouth (higher-numbered transects) because of variable stream flow. Predictions from the spline method, not shown, are similar to those from the regularized-ST method. (b) Seventy-year shoreline predictions for ST (thin gray), spline (dashed black), and regularized ST (heavy black) all show accretion focused near the center of the bay. Ninety-five percent confidence bands are shown. The most recent (2005) shoreline (low water mark) is shown in dashed gray for reference.

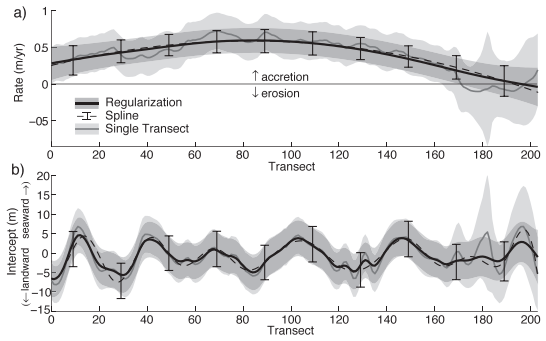

Figure 8. Shoreline change (a) rates and (b) intercepts calculated *via* three methods: (1) regularized ST, (2) B-spline basis function, and (3) ST are shown surrounded by their 95% confidence bands. The three models show general agreement, although the ST model has the most alongshore variability.

and 135, intercept values (representing mean shoreline positions) follow a sinuous pattern that mimics the shape of the landward extent of the offshore reef (Figure 3d). Other deviations from the baseline may be because of persistent physical conditions in combination with reef topography that are not apparent in aerial photos, or they may simply represent noise inherent in the data.

In general, continuous uncomplicated sandy shorelines can be parsimoniously represented using any of the basis functions or regularization. However, shorelines reflecting uneven variation in rates (or intercepts) alongshore, typically because of complex offshore topography, complicate the application of parsimonious methods. Pacific island shorelines in particular tend to be naturally complex because of the intricate fringing reefs that surround most of them. Shorelines everywhere are susceptible to construction of jetties, piers, and other features that alter long-term shoreline trends on small spatial scales. We find that some methods are more successful in dealing with particular morphologic circumstances than others. Table 2 summarizes some differences among regularized-ST, traditional ST, and recently used basis function methods, such as Legendre polynomials, trigonometric functions, principal components (Frazer, Genz, and Fletcher, 2009) and cubic B splines (Anderson and Frazer, 2014). The table is an extension of Table 4 found in Anderson and Frazer (2014) except that it includes an entry for the regularized-ST method.

Frazer, Genz, and Fletcher (2009), as mentioned in the introduction, illustrate how both Legendre polynomials and trigonometric functions cause alongshore ringing in calculated shoreline change rates because of a spike in actual trends across an accretionary point. In regularized ST, a discontinuity in the alongshore derivative of rate (or intercept) is allowed by removing a single row of the second derivative matrix L; to allow a jump in rate, the rows corresponding to the transects that bracket the discontinuity are removed. In general, the regularization parameter for, say, rates can be allowed to vary to either increase or decrease the relative amount of prescribed variation in rates alongshore. Removing a row in the second derivative matrix or allowing a regularization parameter to vary creates a separate class of models. The EIC can be used to compare models between these classes, but the model space will be greatly enlarged.

As pointed out by Anderson and Frazer (2014), principal components regression (eigenbeaches) and B splines can also allow alongshore discontinuities of parameters, but each presents further complication. Principal components regression requires a nonparametric method, such as the bootstrap, to estimate model parameter uncertainty, which can be computationally intensive. For B splines, additional knots must be placed at the discontinuity (de Boor, 1978): one extra knot for a discontinuity in the second derivative, two extra knots for a discontinuity in the first derivative, and three extra knots for a discontinuity in the rate. In general, searching for the most parsimonious B-spline model is complex, because both the number and the locations of the knots should be allowed to vary. By comparison, searching for the most parsimonious regularization model may be simpler, because the number of regularization parameters cannot exceed the number of transects. In this paper, we constrained our spline models to equally spaced splines and our regularized-ST models to equal smoothing parameters at each row, so the search over model space was greatly simplified. Additional spatial analyses, such as wavelet analysis (Li, Lark, and Reeve, 2005; Tebbens, Burroughs, and Nelson, 2002), might provide a way to identify the spatial resolution of shoreline change rates alongshore, linking them with local geology, engineered structures, *etc*.

Table 2. *This table summarizes the advantages and disadvantages of selected long-term shoreline change methods. The eigenbeaches method is the only parsimonious method that is totally robust to rapid alongshore changes in rate, but because such jumps typically correspond to easily identifiable geologic features, regularized ST may be best for long-term prediction.*

Method	Parsimonious	Defined Everywhere	Gibbs Effect	Error Estimation Assumptions
Traditional ST[1]	N	N	N	Independent
Regularized ST	Y	N	N[2]	Correlated errors[3]
Basis functions				
Spline (B splines)	Y	Y	N[4]	Correlated errors
Polynomial (Legendre, trigonometric)	Y	Y	Y	Correlated errors
Eigenbeaches (principal components)	Y[5]	N	N	Correlated errors; nonparametric[6]

[1] Traditional ST assumes independence at each transect.
[2] Must modify L to avoid Gibbs effect.
[3] For correlated errors, the covariance matrix must account for spatially correlated errors.
[4] Careful knot placement is needed to avoid Gibbs effect.
[5] Eigenbeach basis functions are not independent of the data.
[6] Nonparametric requires the nonparametric estimation method (*e.g.*, bootstrap).
Y = yes, N = no.

This may be a way to generalize the relative regularization parameter values alongshore or the spline knot density based on local setting.

Given the relatively large number of parameters allowed by the EIC in determining intercepts alongshore, reducing intercept parameters in the alongshore direction may provide only a small benefit as opposed to calculating an intercept independently at each transect. Anderson and Frazer (2014) found the same disproportionally high number of intercept parameters compared to rate parameters using B splines when they analyzed shoreline data from Assateague Island and Ocean City, Maryland. This suggests that calculating the mean shoreline position at each transect and using this as a new baseline, as done by Frazer, Genz, and Fletcher (2009), simplifies the modeling process without greatly compromising model parsimony.

As mentioned in the introduction, our focus is on methods that provide information about long-term trends in shoreline change for long-term planning purposes. An obvious, but often overlooked, aspect of shoreline change modeling is the practicality of applying the results of scientific studies to planning initiatives. The efficacy of scientific results adopted by nonscientific arenas is not the focus of this paper. Still, the simplicity of a model influences whether it will be adopted. When shoreline change statistics are used to determine the setback for building structures, coastal decision makers and coastal property owners naturally take great interest in how that setback is determined, especially given the relatively high market value of coastal properties. The continuing wide application of the ST method is therefore not surprising, because it is relatively easy to understand and explain, whereas the concept of basis function expansions can be difficult to explain. Its connection with ST may thus make regularized ST the most palatable parsimonious method for planning applications.

Thoughts on Model Selection

Our first inclination was to use to EIC to compare models between the model families (splines and regularized models). We were surprised to find that the best regularized-ST model and the best spline model give comparable CV estimates of general prediction capability, yet the regularized-ST model has a significantly lower (more parsimonious) EIC score. This occurs despite the spline model having fewer basis functions than effective number of parameters in the regularized-ST model. So we ask: How can a model with effectively fewer parameters that is roughly equally as good at predicting independent data be more parsimonious? To investigate this, we looked at the CV scores for all models tested as an alternative means of model selection (Breiman and Spector, 1992; Hastie, Tibshirani, and Friedman, 2009, Section 7.10; Kohavi, 1995; Shao, 1996). Hastie, Tibshirani, and Friedman (2009) suggest selecting the most parsimonious model whose CV score does not exceed a 1-SD distance from the lowest CV score. Because we have two types of model parameters that we are searching over (rate and intercept), the 1-SD height above the lowest CV score creates a contour on our two-dimensional grid of CV scores. For example, adding the smallest spline model CV score (60.2 m^2) to its SD (4.1 m^2) gives 64.3 m^2. This corresponds to roughly the CV scores of a spline model with 4 rate and 34 intercept basis functions and a spline model with 20 rate and 18 intercept basis functions; each of these models has 38 basis functions. The spline model with the fewest overall number of basis functions within 1 SD of the lowest CV score has 6 rate and 18 intercept basis functions. The EIC-selected spline model (3 rate and 18 intercept basis functions) has a CV score (64.4 m^2) that is arguably negligibly above the 1-SD value (64.3 m^2).

The minimum CV score plus 1-SD value for regularization models (64.2 m^2) is remarkably similar to that for spline models (64.3 m^2). The CV score for the EIC-selected regularization model is 64.4 m^2, which is, as with splines, arguably within the 1-SD range. Regularization models are indexed by their regularization parameters, which do not provide a means of selecting the most parsimonious model for CV, so we look at the regression degrees of freedom as a measure of the effective number of model parameters. The regularized model with effectively the fewest number of parameters had 10 rate and 35 intercept parameters. This is quite different from the EIC-selected model with 4 rate and 64 intercept regression degrees of freedom. This leads to another question: Can model complexity be adequately represented by the regression degrees of freedom? Furthermore, does a model with three B-spline basis functions of degree 1 (piecewise linear function) have the same model complexity as a model with three trigonometric (sine and cosine) basis functions? Perhaps the EIC is taking into account some measure of model complexity that finds regularization models less complex than cubic spline models. This would explain the similarity in the alongshore signal of intercepts (Figure 8b) modeled by the regularized-ST model and the spline model, even though the number of spline basis functions was much smaller than the effective number of regularized-ST intercept parameters.

The Vapnik–Chervonenkis (VC) dimension (Vapnik, 1995) is an alternative measure of the complexity of a class of functions that measures how wavy the functions can be. Model selection methods that incorporate the VC dimension have been applied to regression models with varying degrees of success (Cherkassky *et al.*, 1999; Hastie, Tibshirani, and Friedman, 2009; Vapnik, Levin, and Le Cun, 1994).

In the end, we could not justify using the EIC to compare spline and regularized-ST models because the EIC appears to be biased toward regularized-ST models. Yet these questions on comparing models from different model families and measures of model complexities give credence as to why model selection is such an active area of research in which consensus is rare. Recent reviews are given by Claeskens and Hjort (2008) and Ando (2010).

Unknown Variance

For some shoreline data sets, the data errors are unknown or uncertain. In these situations, a best-estimate scaling factor is typically used to estimate the amplitude of the predicted error, based on the data residuals. For example, Anderson and Frazer (2014) modeled the data covariance matrix as $C_d = \alpha \tilde{C}_d$, where α is the best-estimate constant of proportionality that scales $\tilde{C}_d$, the estimated covariance structure matrix, weighted by *a priori* uncertainty estimates (the covariance matrix used in

this paper). For unregularized models, such as the spline model or ST model, the best estimate of α is

$$\hat{\alpha} = (d - G\hat{m})^T \tilde{C}_d^{-1} (d - G\hat{m})/(N - df), \quad (16)$$

where df is the regression degrees of freedom (Equation [12]), which is equal to the number of basis functions. It is possible to estimate α when using regularization, but the calculation is not straightforward because the best-fit model parameters depend on α, and one must iterate to find both the value of α and the value of the parameters.

We tested this procedure on our Kailua Beach data and found that α was always less than (but close to) one, regardless of how much smoothing was imposed on the model. This indicates that the estimated *a priori* errors slightly exceed the data errors estimated from the data residuals, so it is conservative to use the *a priori* errors. Seven difference sources of uncertainty were used to estimate *a priori* data errors for Kailua (see Romine *et al*., 2009), including measurement errors (digitization of shoreline vectors, orthorectification of coastal imagery, *etc*.) and errors because of short-term physical processes (*e.g*., tidal fluctuations and seasonal uncertainties). A scaling factor may be more appropriate for use in locations in which quantification of only a limited fraction of all potential error sources is available.

CONCLUSIONS

A regularization technique for modeling long-term shoreline change in a parsimonious manner is presented and compared with traditional ST and with the B-spline method. The technique is demonstrated on historic shoreline data from Kailua, Hawaii. There is general agreement among the three methods tested, which all indicate long-term accretion of Kailua Bay focused toward the center of the beach, consistent with previous studies of the area (Hwang, 1981; Edward K. Noda and Associates, Inc., Staff, 1989; Norcross, Fletcher, and Merrifield, 2002; Romine *et al*., 2009). Both the spline and the regularized-ST methods (parsimony methods) produce shoreline change rates that are smooth alongshore compared to ST. The parsimony methods used between three and four parameters to characterize the variation in rates alongshore. However, intercept parameters estimated by the parsimony methods showed much higher alongshore variation, compared to rates, and closely resembled those calculated by ST. This phenomenon, also seen at Assateague Island and Ocean City, Maryland (Anderson and Frazer, 2014), suggests that there is little benefit in attempting to reduce the number of intercept parameters, even though uncertainty values will be slightly reduced. Improved *a priori* data error estimates might produce more accurate models.

An information criterion based on maximizing the evidence is used as an objective model selection criterion for data with correlated noise. The CV estimates of prediction error indicate that both the regularized-ST and the spline models have significantly better predictive capabilities than the ST model, while the parsimonious models are comparable to each other.

In applying basis function and regularization methods, spline basis functions (B splines), principal component regression, and regularization methodologies provide the most flexibility in handling alongshore discontinuities often seen in shoreline behavior. Principal component regression was not favored because its basis functions include noise from the data and it requires nonparametric methods of estimating uncertainty. For shorelines with discontinuities or varying spatial scales, the model selection process for regularized ST is easier to code than the one for splines. Finally, the way that regularization imparts varying degrees of smoothness to alongshore rates may be easier to explain to clients than the action of basis functions, making it a practical choice for use in long-term planning.

ACKNOWLEDGMENTS

This work was supported in part by the Pacific Islands Climate Science Center *via* U.S. Geological Survey Cooperative Agreement G12AC00003. We thank four anonymous reviewers for critiquing our paper and providing creative suggestions toward a clearer and more engaging manuscript.

LITERATURE CITED

Anderson, T.R. and Frazer, L.N., 2014. Toward parsimony in shoreline change prediction (III): B-splines and noise handling. *Journal of Coastal Research*, 30(4), 729–742.

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.

Ando, T., 2010. *Bayesian Model Selection and Statistical Modeling*. Boca Raton, Florida: Chapman and Hall/CRC, 286p.

Aster, R.C.; Borchers, B., and Thurber, C.H., 2012. *Parameter Estimation and Inverse Problems,* 2nd ed. Oxford, United Kingdom: Academic Press/Elsevier, 360p.

Box, G.E.P., 1954. The exploration and exploitation of response surfaces: Some general considerations and examples. *Biometrics,* 10(1), 16–60.

Box, G.E.P. and Tiao, G.C., 1973. *Bayesian Inference in Statistical Analysis*. Redding, Massachusetts: Addison-Wesley, 588p.

Breiman, L. and Spector, P., 1992. Submodel selection and evaluation in regression. The X-random case. *International Statistical Review*, 60(3), 291–319.

Cherkassky, V.; Xuhui, S.; Mulier, F.M., and Vapnik, V.N., 1999. Model complexity control for regression using VC generalization bounds. *IEEE Transactions on Neural Networks,* 10(5), 1075–1089.

Claeskens, G. and Hjort, N.L., 2008. *Model Selection and Model Averaging*. New York: Cambridge University Press, 312p.

Crowell, M.; Douglas, B.C., and Leatherman, S.P., 1997. On forecasting future U.S. shoreline positions: A test of algorithms. *Journal of Coastal Research,* 13(4), 1245–1255.

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: Springer-Verlag, 392p.

De Vriend, H.J.; Zyserman, J.; Nicholson, J.; Roelvink, J.A.; Pechon, P., and Southgate, H.N., 1993. Medium-term 2DH coastal area modeling. *Coastal Engineering,* 21(1), 193–224.

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

Edward K. Noda and Associates, Inc., Staff, 1989. *Hawaii Shoreline Erosion Management Study, Overview and Case Study Sites—Makaha, Oahu; Kailua–Lanikai, Oahu; Kukuiula–Poipu, Kauai*. Report for the Hawaii Coastal Zone Management Program, Volume 1. Honolulu, Hawaii: Edward K. Noda and Associates, Inc., pp. 4-34–4-74.

Fenster, M.S.; Dolan, R., and Elder, J.F., 1993. A new method for predicting shoreline positions from historical data. *Journal of Coastal Research,* 9(1), 147–171.

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.

Galgano, F.A. and Douglas, B.C., 2000. Shoreline position prediction: Methods and errors. *Environmental Geosciences,* 7(1), 23–31.

Geisser, S., 1975. The predictive sample reuse method with applications. *Journal of the American Statistical Association,* 70(350), 320–328.

Gelman, A.; Carlin, J.B.; Stern, H.S.; Dunson, D.B.; Vehtari, A., and Rubin, D.B., 2014. *Bayesian Data Analysis,* 3rd ed. Boca Raton, Florida: CRC Press, 675p.

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.

Gutierrez, B.T.; Plant, N.G., and Thieler, E.R., 2011. A Bayesian network to predict coastal vulnerability to sea level rise. *Journal of Geophysical Research: Earth Surface,* 116(F2), F02009. doi:10.1029/2010JF001891.

Hanson, H. and Kraus, N.C., 1989. *GENESIS: Generalized Model for Simulating Shoreline Change. Vol. 1: Reference Manual and Users Guide*. Technical Report *CERC-89-19*. Vicksburg, Mississippi: Coastal Engineering Research Center, U.S. Army Corps of Engineers, 247p.

Hapke, C. and Plant, N., 2010. Predicting coastal cliff erosion using a Bayesian probabilistic model. *Marine Geology,* 278(1–4), 140–149.

Harney, J.N. and Fletcher, C.H., 2003. A budget of carbonate framework and sediment production, Kailua Bay, Oahu, Hawaii. *Journal of Sedimentary Research,* 73(6), 856–868.

Hastie, T.; Tibshirani, R., and Friedman, J., 2009. *The Elements of Statistical Learning,* 2nd ed. New York: Springer-Verlag, 745p.

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.

Hwang, D., 1981. *Beach Changes on Oahu as Revealed by Aerial Photographs*. Technical Report *HIG-81-3*. Honolulu, Hawaii: University of Hawaii, Hawaii Institute of Geophysics, pp. 66–75.

Kohavi, R., 1995. A study of cross-validation and bootstrap for accuracy estimation and model selection. *Proceedings of the 14th International Joint Conference on Artificial Intelligence* (Montreal, Canada), pp. 1137–1143.

Li, Y.; Lark, M., and Reeve, D., 2005. Multi-scale variability of beach profiles at Duck: A wavelet analysis. *Coastal Engineering,* 52(12), 1133–1153.

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.

Markatou, M.; Tian, H.; Biswas, S., and Hripcsak, G., 2005. Analysis of variance of cross-validation estimators of the generalization error. *Journal of Machine Learning Research,* 6(Jul), 1127–1168.

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

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

Morton, R.A., 1979. Temporal and spatial variations in shoreline changes and their implications, examples from the Texas Gulf Coast. *Journal of Sedimentary Research,* 49(4), 1101–1112.

Norcross, Z.M.; Fletcher, C.H., and Merrifield, M., 2002. Annual and interannual changes on a reef-fringed pocket beach: Kailua Bay, Hawaii. *Marine Geology,* 190(3–4), 553–580.

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.

Press, W.H.; Teukolsky, S.A.; Vetterling, W.T., and Flannery, B.P., 2007. *Numerical Recipes: The Art of Scientific Computing,* 3rd ed. New York: Cambridge University Press, 1235p.

Romine, B.M.; Fletcher, C.H.; Frazer, L.N.; Genz, A.S.; Barbee, M.M., and Lim, S.-C., 2009. Historical shoreline change, southeast Oahu, Hawaii: Applying polynomial models to calculate shoreline change rates. *Journal of Coastal Research,* 25(6), 1236–1253.

Ruggiero, P. and List, J.H., 2009. Improving accuracy and statistical reliability of shoreline position and change rate estimates. *Journal of Coastal Research,* 25(5), 1069–1081.

Shao, J., 1993. Linear model selection by cross-validation. *Journal of the American Statistical Association*, 88(422), 486–494.

Sivia, D.S., and Skilling, J., 2006. *Data Analysis: A Bayesian Tutorial,* 2nd ed. New York: Oxford University Press, 246p.

Stive, M.J.; Aarninkhof, S.G.; Hamm, L.; Hanson, H.; Larson, M.; Wijnberg, K.M.; Nicholls, R.J., and Capobianco, M., 2002. Variability of shore and shoreline evolution. *Coastal Engineering,* 47(2), 211–235.

Tebbens, S.F.; Burrows, S.M., and Nelson, E.E., 2002. Wavelet analysis of shoreline change on the Outer Banks of North Carolina: An example of complexity in the marine sciences. *Proceedings of the National Academy of Sciences Colloquium on "Self-Organized Complexity in the Physical, Biological, and Social Sciences"* 99(Suppl. 1), 2554–2560.

University of Hawaii Coastal Geology Group, 2014. Hawaii Coastal Erosion Website: Mosaics. http://www.soest.hawaii.edu/coasts/erosion/mosaics.php?sArea=kailua.

Van Rijn, L.C.; Walstra, D.J.R.; Grasmeijer, B.; Sutherland, J.; Pan, S., and Sierra, J.P., 2003. The predictability of cross-shore bed evolution of sandy beaches at the time scale of storms and seasons using process-based profile models. *Coastal Engineering,* 47(3), 295–327.

Vapnik, V., 1995. *The Nature of Statistical Learning Theory,* New York: Springer-Verlag, 188p.

Vapnik, V.; Levin, E., and Le Cun, Y., 1994. Measuring the VC-dimension of a learning machine. *Neural Computation,* 6(5), 851–876.

Yates, M.L. and Le Cozannet, G., 2012. Brief communication: Evaluating European coastal evolution using Bayesian networks. *Natural Hazards Earth System Science,* 12(4), 1173–1177.

Zellner, A. and Tiao, G.C., 1964. Bayesian analysis of the regression model with autocorrelated errors. *Journal of the American Statistical Association,* 59(307), 763–778.

APPENDIX. EVIDENCE INFORMATION CRITERIA

We determine the appropriate model by maximizing the evidence, which is the prior probability of the data given a model. Our EIC is defined as

$$\mathrm{EIC} = -2\ln(p(d|\psi)), \tag{A1}$$

where $p(d|\psi)$ is the evidence and ψ represents the model. Thus, the model with the maximum evidence is the one whose EIC score is the minimum of all candidate model scores. Roughly speaking, in regularized ST ψ represents the system matrix and its regularization parameter or parameters, and in the spline method it represents the system matrix and the number of splines.

Using marginalization and the product rule, the evidence can be written in the form (Sivia and Skilling, 2006, p. 79)

$$\begin{aligned} p(d|\psi) &= \int p(m,d|\psi)dm \\ &= \int p(d|m,\psi)p(m|\psi)dm \end{aligned} \tag{A2}$$

in which $p(d|m, \psi)$ is the likelihood of parameter vector m and $p(m|\psi)$ is the parameter prior for model ψ.

We first consider regularized ST. Following Sivia and Skilling (2006, pp. 79–83), our parameter prior is the uniform probability density function on a hypercube. Thus, the prior for the parameters is given by

$$p(m|\psi) = a_r^{-Mr} a_b^{-Mb} \, [-a_r/2 < m_k < a_r/2 \ : \ 1 \le k \le Mr \text{ and } -a_b/2 < m_l < a_b/2 \ : \ 1 \le l \le Mb]. \tag{A3}$$

Here, [...] is an indicator function, equal to unity if condition ... is satisfied and zero if it is not; a_r and a_b are the hypercube side lengths for rate and intercept parameters, respectively; and Mr and Mb are the respective numbers of rate and intercept parameters. Following Sivia and Skilling (2006), we assume that the hypercube side lengths a are sufficiently large that the indicator function can be neglected in the integral over m needed to calculate the evidence. For all regularized-ST models, the prior is the same because Mr and Mb do not vary; for basis function models, such as splines, Mr and Mb vary and cause the prior to affect the evidence of models within the same model family.

We maximize the evidence by varying the model in the likelihood. The likelihood is

$$p(d|m, \psi) = \frac{\exp(-2^{-1}(\tilde{G}m - \tilde{d})^T \tilde{C}_d^{-1} (\tilde{G}m - \tilde{d}))}{(2\pi)^{D/2} |\tilde{C}_d|^{1/2}}, \tag{A4}$$

in which the terms $\tilde{G}$, $\tilde{d}$, and $\tilde{C}_d$ are the generalized system matrix, data vector of length D, and data covariance matrix, respectively, for each model. Explicit values of the preceding terms, which differ for regularized models and basis function models, are provided later.

Evaluating the integral in Equation (A2) in the Sivia and Skilling (2006) approximation is a simple matter of completing the square in the exponent of the likelihood and using the formula for the integral of a multivariate Gaussian. Alternatively, one can simply divide the product of the likelihood and the prior, $p(d|m, \psi)p(m|\psi)$, by the posterior $p(m|d, \psi)$ and (recognizing that this quotient is independent of m) set m equal to a convenient value φ such as $\hat{m}$, the mean of the posterior. The result is

$$p(d|\psi) \approx \frac{(2\pi)^{M/2} |C_m|^{1/2} \exp(-2^{-1}\hat{\varphi})}{(2\pi)^{D/2} |\tilde{C}_d|^{1/2} |a_r|^{Mr} |a_b|^{Mb}}, \tag{A5}$$

in which

$$\hat{\varphi} = (\tilde{G}\hat{m} - \tilde{d})^T \tilde{C}_d^{-1} (\tilde{G}\hat{m} - \tilde{d}). \tag{A6}$$

Multiplying –2 times the log of $p(d|\psi)$ gives the EIC

$$\text{EIC} = (D - M)\ln(2\pi) + \ln(|\tilde{C}_d|) + \hat{\varphi} + \ln(|C_m^{-1}|) + 2\,[Mr\ln(a_r) + Mb\ln(a_b)]. \tag{A7}$$

Here, M is the length of parameter vector m, in which $M = Mr + Mb$, or Mr and Mb as defined earlier; C_m is the model covariance matrix; and D, $\tilde{C}_d$, $\hat{\varphi}$, a_r, and a_b were defined earlier.

For regularized-ST models, $\tilde{G}$, $\tilde{d}$, and $\tilde{C}_d$ in Equations (A4) to (A7) are defined as

$$\tilde{G} = \begin{bmatrix} G \\ L \end{bmatrix}, \tilde{d} = \begin{bmatrix} d \\ \mathbf{0} \end{bmatrix}, \text{and } \tilde{C}_d = \begin{bmatrix} C_d & 0 \\ 0 & C_{reg} \end{bmatrix}, \tag{A8}$$

in which G is the system matrix that appears in Equation (3); L is a block diagonal matrix whose blocks are the roughening operators, usually the second derivative linear operators L_r and L_b (Equation [2]); d is the vector of data; $\mathbf{0}$ is a column vector of zeros of length $(M - 4)$, in which M is the length of parameter vector m; C_d is the covariance matrix; C_{reg} is a block diagonal matrix with blocks $\gamma_r^{-2} \cdot I_{M/2-2}$ and $\gamma_b^{-2} \cdot I_{M/2-2}$, in which γ_r and γ_b are the rate and the intercept regularization parameters, respectively; and $I_{M/2-2}$ is the identity matrix of the size $(M/2– 2) \times (M/2 - 2)$.

Substituting Equation (A8) into (A4) and then taking –2 times the natural log of the resulting expression leads to the explicit EIC formula for regularized-ST models:

$$\begin{aligned} \text{EIC}_\text{R} = {} & (d - G\hat{m})^T C_d^{-1} (d - G\hat{m}) + \gamma_r^2 \hat{r}^T L_r^T L_r \hat{r} + \gamma_b^2 \hat{b}^T L_b^T L_b \hat{b} \\ & + \ln(|C_m^{-1}|) + (N - 4)\ln(2\pi) + M\ln(a_r a_b) \\ & + (4 - M)\ln(\gamma_r \gamma_b) + \ln(|C_d|), \end{aligned} \tag{A9}$$

in which $C_m^{-1} = G^T C_d^{-1} G + \gamma^2 L^T L$ and N is the length of data vector d. In this paper, we are comparing regularized-ST models that have the same data covariance C_d, so the term on the last line does not change and may be discarded.

When using basis functions, such as splines, the model system matrix G has M basis functions, one in each column. We identify a model ψ by its system matrix G. In Equations (A4) to (A7), $\tilde{G}$ is simply the system matrix G, $\tilde{d}$ is the vector of data d, and $\tilde{C}_d$ is the data covariance matrix C_d. The parameter prior is a hypercube, as for the regularization model (Equation [A3]), but the dimension of the hypercube changes with the number of splines, *i.e.* with the number of columns of G. It follows that the explicit EIC formula for basis function models is

$$\begin{aligned} \text{EIC}_\text{B} = {} & (d - G\hat{m})^T C_d^{-1} (d - G\hat{m}) + \ln(|C_m^{-1}|) + (N - M)\ln(2\pi) \\ & + 2\,[Mr\ln(a_r) + Mb\ln(a_b)] + \ln(|C_d|), \end{aligned} \tag{A10}$$

in which $C_m^{-1} = G^T C_d^{-1} G$ and $\hat{m} = C_m G^T C_d^{-1} d$. Our procedure assumes that the data covariance matrix is constant, so the term in the last line may be discarded, as with EIC_R.