Journal of Coastal Research	23	1	87–105	West Palm Beach, Florida	January 2007

The Predictive Accuracy of Shoreline Change Rate Methods and Alongshore Beach Variation on Maui, Hawaii

Ayesha S. Genz†, Charles H. Fletcher†, Robert A. Dunn†, L. Neil Frazer†, and John J. Rooney‡

†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.
agenz@hawaii.edu
fletcher@soest.hawaii.edu

‡Joint Institute of Marine and Atmospheric Research
University of Hawaii
Pacific Islands Fisheries Science Center
Kewalo Research Facility
1125-B Ala Moana Boulevard
Honolulu, HI 96814, U.S.A.

ABSTRACT

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. West Palm Beach (Florida), ISSN 0749-0208.

Beach erosion has direct consequences for Hawaii's tourist-based economy, which depends on the attraction of beautiful sandy beaches. Within the last century, however, beaches on Oahu and Maui have been narrowed or completely lost, threatening tourism and construction development. In order for the counties and state of Hawaii to implement coastal regulations to prevent infrastructure damage, it is necessary to find a statistically valid methodology that accurately delineates annual erosion hazard rates specific to Hawaii. We compare the following erosion rate methods: end point rate (EPR), average of rates (AOR), minimum description length (MDL), jackknifing (JK), ordinary least squares (OLS), reweighted least squares (RLS), weighted least squares (WLS), reweighted weighted least squares (RWLS), least absolute deviation (LAD), and weighted least absolute deviation (WLAD). To evaluate these statistical methods, this study determines the predictive accuracy of various calculated erosion rates, including the effects of *a priori* knowledge of storms, using (1) temporally truncated data to forecast and hindcast known shorelines and (2) synthetic beach time series that contain noise. This study also introduces binning of adjacent transects to identify segments of a beach that have erosion rates that are indistinguishable. If major uncertainties of the shoreline methodology and storm shorelines are known, WLS, RWLS, and WLAD better reflect the data; if storm shorelines are not known, RWLS and WLAD are preferred. If both uncertainties and storm shorelines are not known, RLS and LAD are preferred; if storm shorelines are known, OLS, RLS, JK, and LAD are recommended. MDL and AOR produce the most variable results. Hindcasting results show that early twentieth century topographic surveys are valuable in change rate analyses. Binning adjacent transects improves the signal-to-noise ratio by increasing the number of data points.

ADDITIONAL INDEX WORDS: *Coastal erosion, shoreline change rates, coastal management, Hawaii, erosion hazard area.*

INTRODUCTION

The coastal zone is one of the nation's greatest environmental and economic assets (OCEAN STUDIES BOARD, 1999). In Hawaii, for example, over 60% of all jobs are related to tourism, which depends on the appeal of sandy beaches. Yet widespread beach erosion in the Hawaiian Islands threatens sand-dependent ecosystems and abutting coastal owners (FLETCHER, MULLANE, and RICHMOND, 1997; NORCROSS-NU'U and ABBOTT, 2005; ROONEY *et al.,* 2003).

Recognition of beach value led Maui County to approve and adopt in October 2003 the only science-based setback rules in Hawaii. These rules are based on erosion rates that are calculated by the reweighted least squares (RLS) method, which identifies and removes outliers before modeling the shoreline change trend with a straight line (FLETCHER *et al.,* 2003; ROONEY, 2002; ROONEY *et al.,* 2003). The slope of the line represents the erosion (positive slope) and accretion (negative slope) rate of the beach.

Although comparisons of different shoreline change rate methods have been conducted along the continental East Coast of the United States, extensive research does not exist for beaches in Hawaii. Owing to fundamental differences between Hawaiian beaches and those of the continental main-

DOI:10.2112/05-0521.1 received 25 May 2005; accepted in revision 14 September 2005.

Funding for this study was provided by the U.S. Geological Survey National Shoreline Assessment Project, Maui County, the State of Hawaii Department of Land and Natural Resources, the Hawaii Coastal Zone Management Program, and a grant/cooperative agreement from the National Oceanic and Atmospheric Administration, Project #R/EP-22, which is sponsored by the University of Hawaii Sea Grant College Program, SOEST, under Institutional Grant No. NA16RG2254 from NOAA Office of Sea Grant, Department of Commerce. The views expressed herein are those of the authors and do not necessarily reflect the views of NOAA or any of its subagencies. UNIHI-SEAGRANT-JC-04-12.

land (*e.g.,* sediment composition, seasonal signal, storm frequency, and storm impact), there is a need for studies specific to Hawaii. Additionally, previous studies have not comprehensively tested shoreline change rate methods.

Our goal is to compare published statistical shoreline change rate methods, including three statistical methods previously not used in calculating shoreline change rates. These comparisons are made using shoreline change data from the island of Maui. We first investigate the effects of outliers based on *a priori* knowledge (*e.g.,* a devastating tsunami, hurricane, or storm event) on predictions that are based on shoreline change rates. We then compare different shoreline change rate methods using synthetically derived data. Finally, we examine the binning of adjacent transects to find beach segments that have indistinguishable change rates.

PREVIOUS WORK

Most studies of shoreline changes have been done on continental beaches of the eastern United States. We review these and other studies that determine the most appropriate method based on either a comparative approach or the prediction of known positions.

Dolan, Fenster, and Holme (1991) compared long-term and short-term erosion rates with methods such as end point rate (EPR), linear regression (hereafter ordinary least squares, OLS), jackknifing (JK), and average of rates (AOR). By plotting the rates from one method *vs.* the rates of another method, they concluded that AOR is most variable, while OLS and JK share a high degree of similarity. They stressed that the best method depends on the objective and the temporal variables of the research.

In discussing beach erosion at Rincón, Puerto Rico, Thieler, Rodriguez, and Carlo (1995) calculated erosion rates using EPR, OLS, JK, and AOR in the digital shoreline analysis system (DSAS, http://woodshole.er.usgs.gov/project-pages/dsas/). They divided their study site into four separate areas and calculated an average shoreline change rate at each section for each of the four methods. All four methods resulted in similar rates, but AOR was identified as the most appropriate shoreline change rate method at Rincón.

Dean and Malakar (1999) considered three shoreline change rate methods—OLS, EPR, and AOR—in mapping Florida's hazard zones. They calculated correlation coefficients to compare the three methods. All three methods agreed with each other; however, EPR and OLS correlated better with each other than either did with AOR. The authors chose OLS as their preferred method.

Fenster, Dolan, and Elder (1993) introduced a new method, minimum description length (MDL), as an alternative to existing methods. Based on the MDL modeling criterion of Rissanen (1989), this simplified version uses a complexity penalty to select the model that best fits the data (*e.g.,* constant, linear, quadratic, *etc.*) with the fewest number of parameters. Crowell, Douglas, and Leatherman (1997) compared predicted values from the MDL and OLS methods to actual values using sea-level data as a proxy for historical shoreline data. They concluded that OLS provides equal, if not better, results for shorelines without any physical modifications.

Differing from Crowell, Douglas, and Leatherman (1997), Honeycutt, Crowell, and Douglas (2001) compared EPR to OLS by predicting known historical shoreline data, not sea-level data, to determine the accuracy of the methods. Using *a priori* knowledge of major storms, they confirmed the findings of Galgano, Douglas, and Leatherman (1998) and Galgano and Douglas (2000), which showed that the accuracy of shoreline change rates improves without storm-influenced data points. They concluded that OLS better predicts shorelines than EPR. A good method to identify the best predictor involves using an earlier subset of shoreline positions to test forecasting of later positions (*i.e.,* cross-validating). Our forecasting and hindcasting procedures follow those of Honeycutt, Crowell, and Douglas (2001).

ESTABLISHED EROSION RATE METHODS

Dolan, Fenster, and Holme (1991) provided an excellent overview of some of the published shoreline change rate methods. We expand upon their study to include other published methods, describing the advantages and disadvantages of each (Table 1). The two most frequently cited methods are EPR and OLS, although most researchers now prefer OLS (Figure 1).

Most shoreline change rate methods assume shoreline change is linear through time, with any nonlinearity attributed to mapping and measurement errors. Shorelines do not recede or accrete in a uniform manner, which raises questions about the appropriateness of linear models (Douglas, Crowell, and Leatherman, 1998; Fenster and Dolan, 1994; Fenster, Dolan, and Elder, 1993; Morton, 1991).

End Point Rate (EPR)

The EPR method uses only two data points to delineate a change rate—the earliest and most recent shoreline positions. Given that only the end data points are used, the information contained in the other data points is entirely omitted, preventing the observation of variations in rate through time. The main disadvantage of this method is that if one or both end points are erroneous, the calculated erosion rate will be inaccurate (Crowell, Douglas, and Leatherman, 1997; Crowell, Honeycutt, and Hatheway, 1999; Dolan, Fenster, and Holme, 1991).

Average of Rates (AOR)

Shoreline positions are often defined from various sources (*e.g.,* topographic surveys, coastal monument and beach profiles, and aerial photographs), each with its own measurement uncertainty. For this reason, Foster and Savage (1989) developed the AOR method to average the long-term change, excluding changes due to measurement errors. To do this, they created a minimum time criterion that filters out any changes due to short time spans or measurement errors. EPRs are determined between all data point pairs and are removed if the time interval is less than a specified minimum. All EPRs that pass the criterion are averaged to de-

Table 1. *Advantages and disadvantages of shoreline change rate methods.*

Shoreline Change Rate Method	Advantages	Disadvantages
EPR	Simple computation	Only uses two end points; assumes linear trend
AOR	Uses measurement errors in identifying shoreline change rate	Assumes linear trend; minimum time criterion affected by large errors or small EPR; influenced by EPR rates of short time spans
MDL	Does not assume linear trend when identifying best model fit	Emphasis on recent data; analyst judgment is needed if model is nonlinear
OLS	Simple computation; uses statistical tests	Assumes linear trend; sensitive to statistical outliers
JK	Decreases influence of clustered data and extreme data points; uses statistical tests	Assumes linear trend
RLS	Robust to statistical outliers; uses statistical tests	Assumes linear trend; removes data points before identifying trend
WLS	Incorporates uncertainties into trend line; uses statistical tests	Assumes linear trend; sensitive to statistical outliers
RWLS	Incorporates uncertainties into trend line; robust to statistical outliers; uses statistical tests	Assumes linear trend; removes data points before identifying trend
LAD	Robust to statistical outliers	Assumes linear trend; analyst identifies range of slopes and intercepts
WLAD	Incorporates uncertainties into trend line; robust to statistical outliers	Assumes linear trend; analyst identifies range of slopes and intercepts

termine the shoreline change rate (DOLAN, FENSTER, and HOLME, 1991; FOSTER and SAVAGE, 1989). One drawback is that the minimum time criterion can be affected by large errors or small EPRs, resulting in potentially misleading results (DOLAN, FENSTER, and HOLME, 1991). AOR also gives more influence to EPR rates of short time spans (FENSTER, DOLAN, and ELDER, 1993). For these reasons, FOSTER and SAVAGE (1989) recommend confirming AOR results with other shoreline change rate methods such as OLS.

Minimum Description Length (MDL)

Since short-term changes may affect long-term trends, FENSTER, DOLAN, and ELDER (1993) proposed a simplified form of the MDL method to help identify influential short-term changes. Assuming Gaussian errors, MDL uses an error component and a complexity penalty to select the best model fit, whether it is a constant, line, quadratic, *etc.* If the resulting model is quadratic or higher, two lines are produced—the zero-weight line (MDL ZERO), which uses only recent data, and the low-weight line (MDL LOW), which assigns weights to older data. MDL rates based on nonlinear models tend to result in variable or highly inaccurate forecasts, though the MDL criterion can help identify physical changes within a beach (CROWELL, DOUGLAS, and LEATHERMAN, 1997).

Ordinary Least Squares (OLS)

Least squares regression assumes independent Gaussian errors and estimates the trend of shoreline data by minimizing the sum of the squared residuals between the data and line. The estimated parameters (b_0, or intercept; b_1, or slope) are those that minimize $\sigma_i(y_i - b_0 - b_1x_i)^2$. The assumption of Gaussian errors is usually valid, since the sum of many sources of error, as occur in these studies, tends toward a Gaussian distribution. However, outliers that violate the Gaussian assumption will bias the apparent trend (SEBER and LEE, 2003). *A priori* knowledge of non-Gaussian data points (*e.g.,* storm points) can be used to eliminate such points. Ordinary least squares assumes homoscedasticity (*e.g.,* KLEINBAUM *et al.,* 1998), which means that the variance of each Y component (shoreline position) is the same.

This method is easy to code, and many software companies include OLS as a tool in their spreadsheet programs. A number of statistical tests have been developed (*e.g.,* analysis of variance [ANOVA]) to determine the goodness of the fit and to calculate confidence intervals around the line, future position, and shoreline change rate. These tests require near Gaussian statistics, which are derived from data scatter rather than independent sources.

A linear fit provides a long-term trend over the years for which data are available, but shoreline change is not constant. Also, sediment supply and transport, presence of engineered structures, and storms may not result in Gaussian variations in the data (*e.g.,* FENSTER, DOLAN, and ELDER 1993; GALGANO and DOUGLAS, 2000; GALGANO, DOUGLAS, and LEATHERMAN, 1998; HONEYCUTT, CROWELL, and DOUGLAS, 2001). Clustering of data in time greatly affects the trend line by causing some points to have undue influence (DOLAN, FENSTER, and HOLME, 1991; FENSTER, DOLAN, and ELDER 1993). Since the line fit does not incorporate the uncertainty of each data point, the uncertainties of future shoreline positions may not reflect the data accurately. For example, according to ROUSSEEUW and LEROY (1987), this method is sensitive to outliers; often only one point is needed to distort the trend. If an outlier exists within a data set (*e.g.,* storm point that violates the Gaussian assumption), the resulting line may be highly influenced by that one point. *A priori* knowledge is therefore important, yet controversial (*e.g.,* FENSTER, DOLAN, and MORTON, 2001; GALGANO and DOUGLAS, 2000; GALGANO, DOUGLAS, and LEATHERMAN, 1998; HONEYCUTT, CROWELL, and DOUGLAS, 2001; ZHANG, DOUGLAS, and LEATHERMAN, 2002).

Jackknifing (JK)

The jackknifing method uses multiple OLS fits to determine the shoreline change rate. A different point for each line is omitted, resulting in a different slope for each line. The slopes are averaged to provide a shoreline change rate. Jack-

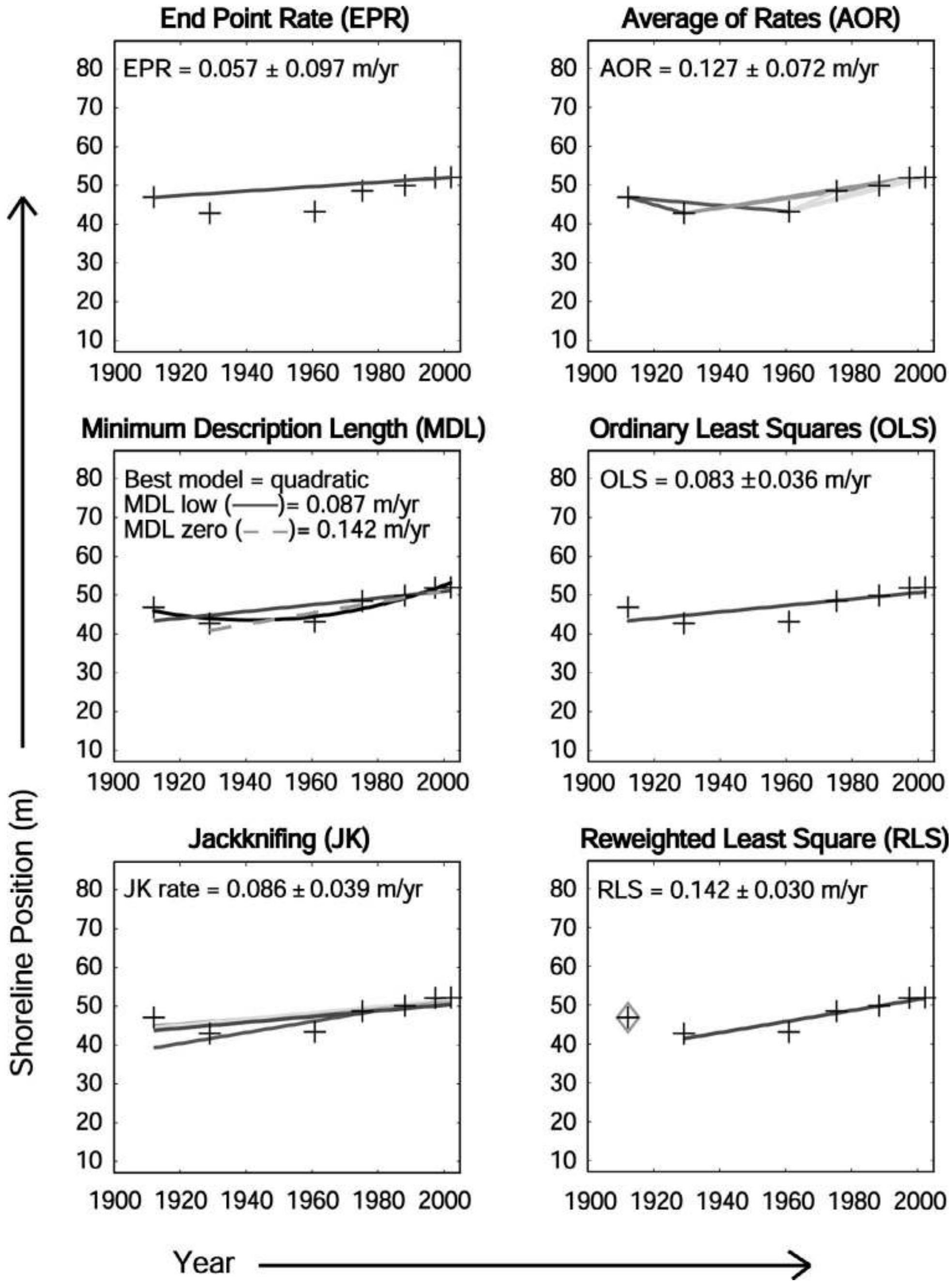

Figure 1. Published shoreline change rate methods applied to one data set of seven shoreline positions. A positive slope shows erosion, while a negative slope shows accretion. The shoreline positions are signified as crosses and the diamonds represent outliers.

knifing has the advantage of decreasing the influence of clustered data and extreme data points. However, computing all possible linear trends is not efficient (DOLAN, FENSTER, and HOLME, 1991).

Reweighted Least Squares (RLS)

RLS helps identify the true trend of shoreline change data by removing statistical outliers in the data. This two-step method first identifies outliers at a cutoff value ($\hat{\sigma}$) using the least median of squares (LMS) regression (ROUSSEEUW and LEROY, 1987). Points identified as statistical outliers are given a weight of 0, and all other points are assigned a weight of 1. An OLS fit then finds the trend with all data points of weights equal to 1. Unlike OLS, RLS is more robust and not as sensitive to outliers. RLS has a breakdown of 50% (that is, if 50% of the data are outliers, the trend of the data can

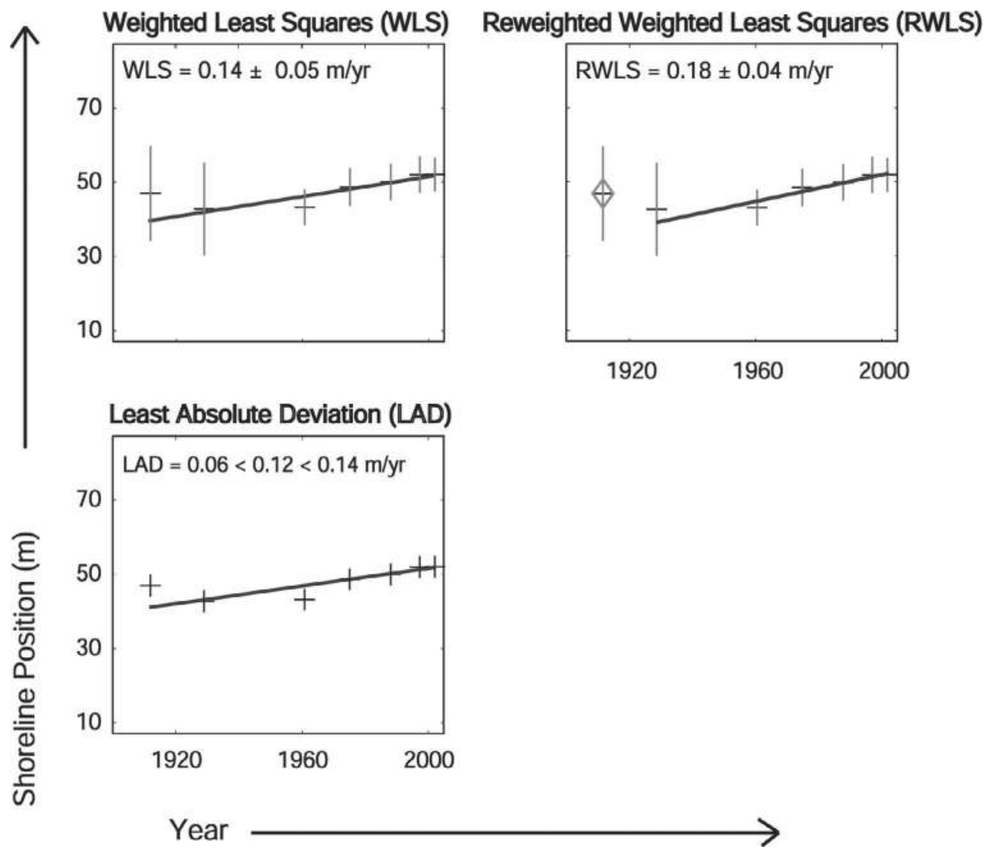

Figure 2. New shoreline change rates. The crosses represent shoreline positions. The vertical lines are $1 - \sigma$ error bars associated with each shoreline position, and the diamonds correspond to outliers.

still be identified accurately) (ROUSSEEUW and LEROY, 1987). Since most shoreline studies have a limited data set, the removal of two or three points without any prior knowledge runs an undesirable risk of discarding good, but noisy, data. Also, adjacent transects along a beach (spaced 20 m in our case) do not always identify the same year as an outlier. This can lead to significant alongshore variations in modeled shoreline rates that are inconsistent with natural beach dynamics. Thus, this method probably works best with a large amount of data or if data from adjacent transects are binned (discussed later) so that true outliers are more evident.

NEW EROSION RATE METHODS

In addition to the established methods discussed above, we explore three new methods to calculate shoreline change rates. These methods are based on well-established theoretical frameworks and are more robust than the least square methods described above (Figure 2).

Weighted Least Squares (WLS)

Unlike OLS, WLS assumes heteroscedastic uncertainties. This means that the variance associated with each Y component (shoreline position) is not necessarily the same at each X component (time) (*e.g.,* KLEINBAUM *et al.,* 1998). If the variances are the same, WLS reduces to OLS (GRAYBILL and IYER, 1994). In many studies it may be difficult to quantify the uncertainties for WLS; however, if the variance (σ^2) or standard deviation (σ) for each Y component is known, the weight (w) is equal to $1/\sigma^2$. In matrix form, solving for $\boldsymbol{b}$, a column vector with unknown parameters of intercept and slope, results in

$$\boldsymbol{b} = (\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{X})^{-1}\boldsymbol{X}^T\boldsymbol{W}\boldsymbol{Y}$$

in which $\boldsymbol{Y}$ is a column vector containing shoreline positions, $\boldsymbol{X}$ is a matrix composed of a column of ones and a column of time data, and $\boldsymbol{X}^T$ is the transpose of the matrix $\boldsymbol{X}$, (*e.g.,* DRAPER and SMITH, 1998). The weight matrix, $\boldsymbol{W}$ is

$$\mathbf{W} = \begin{pmatrix} w_1 & 0 & 0 & 0 \\ 0 & w_2 & 0 & 0 \\ 0 & 0 & w_3 & 0 \\ 0 & 0 & 0 & w_n \end{pmatrix}.$$

where $w_i = 1/\sigma_i^2$ and n is the total number of data points (*e.g.,* GRAYBILL and IYER, 1994).

Data points with large variance will have less of an influence on the trend line than data points with smaller variance (GRAYBILL and IYER, 1994). For example, early shoreline data have larger uncertainties associated with them than re-

Table 2. *An example of a misfit function for LAD. The values of the calculated misfit function are in each box.* b_0 *= intercept and* b_1 *= slope. The minimum value is highlighted with a slope of 0.50 and intercept of 0.*

	b_1								
b_0	0.00	0.25	0.50	0.75	1.00	1.25	1.50	1.75	2.00
−2.00	22.36	16.86	11.56	7.72	6.51	7.62	10.77	16.14	21.64
−1.50	19.61	14.11	9.19	6.14	5.40	8.10	13.39	18.89	24.39
−1.00	16.86	11.36	7.24	4.69	5.55	10.64	16.14	21.64	27.14
−0.50	14.11	9.29	5.29	4.01	7.89	13.39	18.89	24.39	29.89
0.00	11.74	7.24	**3.65**	5.75	10.64	16.14	21.64	27.14	32.64
0.50	9.69	5.23	4.60	8.47	13.39	18.89	24.39	29.89	35.39
1.00	7.64	5.25	6.95	11.22	16.14	21.64	27.14	32.64	38.14
1.50	7.29	6.45	9.47	13.97	18.89	24.39	29.89	35.39	40.89
2.00	8.18	8.47	12.22	16.72	21.64	27.14	32.64	38.14	43.64

cent shorelines. WLS will put more weight on the recent data. The resulting trend line incorporates the uncertainty at each position as well as the uncertainty of the model. Since all other assumptions for WLS equate with OLS (*e.g.,* Gaussian errors), statistical tests and calculation of confidence intervals associated with OLS can also be performed on WLS (Kleinbaum *et al.,* 1998).

In order to apply this method, it is necessary for researchers to estimate all uncertainties associated with their study. As with OLS, this method is sensitive to outliers even if their weights are small. Hence, *a priori* knowledge is also important. The drawbacks of assuming a Gaussian distribution as discussed in the OLS section also apply to WLS. If the calculated uncertainties at each shoreline position do not accurately express the real deviations, then the resulting rate may underestimate or overestimate the true rate.

Reweighted Weighted Least Squares (RWLS)

RWLS is similar to RLS except that it takes into account the uncertainties of each shoreline position. After identifying and removing outliers using LMS, a WLS line is fit to the data. Like WLS, RWLS incorporates positional uncertainties into the rate.

Least Absolute Deviation (LAD)

Least absolute deviation is more robust with respect to outliers than least squares. Unlike the squared residuals of least squares, the sum of the absolute value residuals in LAD is minimized (*i.e.,* misfit function is $\Sigma_i |y_i - b_0 - b_1x_i|$). Since the residuals are not squared, an extreme value has less effect in LAD than in least squares. The assumed distribution of measurement errors is a Laplace, or two-sided exponential, distribution rather than a Gaussian distribution, and the Laplace distribution's longer tails make it less sensitive to outliers (Tarantola, 1987). In comparing robust estimators, Rousseeuw and Leroy (1987) stated that LAD is preferable to least squares methods when outliers are in the *y* direction, which is the case in nearly all historical shoreline analyses.

Calculating the LAD estimate is not as straightforward as it is with least squares. A grid search is performed to calculate a misfit over a range of slopes and intercepts. The best fitting line is the one whose slope and intercept minimize the misfit. For example, an intercept, b_0, that ranges from −2 to 2 with increments of 0.5, and slope, b_1, that ranges from 0 to 2 with increments of 0.25, identifies a minimum misfit of 3.65. The slope and intercept estimates are 0.50 and 0.00, respectively (Table 2).

The uncertainty calculation for the slope estimate is more difficult than that of least squares. To obtain a range of slopes at a certain percentile, we first calculate an estimator of standard deviation (analogous to the root mean square error in least squares) and use it to compute the likelihood function, which in this case is the joint probability density function (PDF) of both the slope and intercept. The marginal PDF of the slope is obtained by integrating the joint PDF over intercept. The marginal PDF then gives the slope range at the percentile of interest. Unlike least squares, the slope range is not necessarily symmetrical around the peak of the slope PDF.

The major advantage of LAD is its robustness with respect to outliers. Similar to WLS, weights can also be incorporated in LAD (weighted least absolute deviation, or WLAD). Since there are only two parameters, LAD is not difficult to code. Care must be taken in selecting a range of slopes and intercepts to search from when using the grid search. If the range is not broad enough, or point spacing not dense enough, the resulting estimates of slope and intercept might not reflect the data accurately; however, the algorithm can be made self-checking and self-adjusting to overcome this minor difficulty.

HAWAIIAN DATA AND UNCERTAINTIES

To calculate shoreline change rates in Hawaii, we digitize the toe of the beach as our shoreline position on images taken in different years (Fletcher *et al.,* 2003). The beach toe, which approximates the low water line, is a more accurate indicator for shoreline change analysis along Hawaiian coastlines when compared with the mean high water line (MHWL) (Coyne, Fletcher, and Richmond, 1999; Fletcher *et al.,* 2003; Rooney and Fletcher, 2000).

Several sources of error influence the delineation of shorelines. For example, aerial photographs taken at various tide levels influence the location of the digitized shoreline, which in turn influences the resulting shoreline change rate. For our data, Fletcher *et al.* (2003) made a special effort to identify and quantify all errors in order to assess the $1 - \sigma$

uncertainty of a shoreline position. The errors are squared and summed to get a total positional uncertainty. We assume the total uncertainty follows a Gaussian distribution, since the central limit theorem states that the sum of multiple sources of uncertainty of arbitrary distributions tends toward a normal distribution (DRAPER and SMITH, 1998).

We use two different types of images to generate our shoreline positions—topographic surveys (National Oceanic and Atmospheric Association T-sheets) and vertical aerial photographs. Only T-sheets that pass the National Map Accuracy Standards are used in this analysis (FLETCHER *et al.,* 2003). The original surveyors of these T-sheets designated MHWL as the shoreline position. We migrate the MHWL in T-sheets to the low water line based on seasonally collected profiles.

Following FLETCHER *et al.* (2003) and ROONEY *et al.* (2003), we calculate the total positional uncertainty (U_t) using the equation

$$U_t = \pm\sqrt{Er^2 + Ed^2 + Ep^2 + Ets^2 + Etd^2 + Es^2 + Ec^2},$$

where Er = rectification error, Ed = digitizing error, Ep = pixel error, Ets = error plotting on a T-sheet, Etd = tidal fluctuation error, Es = seasonal error, and Ec = error in field identification of MHWL and low water line, as evidenced by the beach toe. Errors for T-sheets include Ets and Ec, and exclude Er and Etd. Aerial photographs do not include Ets and Ec.

METHODS

Shoreline Change Rate Comparisons

Forecasting and Hindcasting (Cross-Validation)

Following HONEYCUTT, CROWELL, and DOUGLAS (2001), we compare the shoreline change rate methods discussed earlier by predicting known shoreline positions. For each prediction we calculate the difference between the actual and predicted position for all nine methods. HONEYCUTT, CROWELL, and DOUGLAS (2001) refer to this difference as the error in prediction (EIP) and describe the mean absolute EIP (or mean |EIP|) as a way of representing the magnitude of the error. In comparing the mean |EIP| for all the methods, we perform an ANOVA test at a 95% confidence interval to identify the differences. To compare the results of differing beach dynamics on Maui, we make forecasting predictions on two types of beaches—those with and without engineered (or hardened) structures. We make hindcasting predictions to check the validity of our earliest T-sheet points.

We compare EIPs of predictions that include storm-influenced points to EIPs of predictions exclusive of these points. Determining storm-influenced points is difficult in an island setting where different parts of the island are exposed to varying weather conditions (FLETCHER *et al.,* 2003; ROONEY, 2002; ROONEY *et al.,* 2003). We therefore classify three regions on Maui—Kihei, West Maui, and the North Shore—as each having its own distinct wave regime (FLETCHER *et al.,* 2003; ROONEY *et al.,* 2003) (Figure 3).

We use previous research and historical accounts to determine *a priori* knowledge of storms. Tide gauge data are used to confirm storm events for the North Shore only. From tide

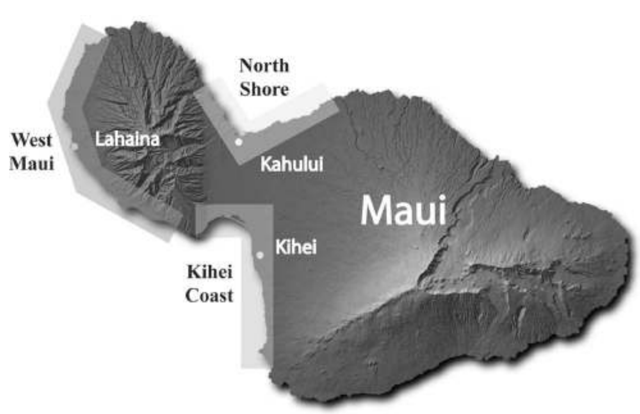

Figure 3. Three classified regions within Maui that experience different weather conditions and wave dynamics.

data and historical accounts of a devastating tsunami, we consider the 1960 shoreline to be an outlier for the North Shore. We similarly conclude that the 1963 shoreline is an outlier for West Maui because of kona storms based on the work of EVERSOLE (2002) and is also an outlier in the Kihei region because of kona storms based on the study of ROONEY (2002). Kona storms are low pressure systems that generate high winds and waves from the south and occur mostly in winter.

There are various difficulties associated with forecasting known positions. As HONEYCUTT, CROWELL, and DOUGLAS (2001) point out, predictions of the near future have lower uncertainty than predictions of the distant future. Also, if fewer data are available for a prediction, it will have large uncertainty. Another difficulty of forecasting analysis is that shoreline positions are not precisely known. Since all shoreline positions are subjected to errors inherent in aerial photogrammetry and T-sheets, the true positions are unknown. Instead, an estimate of the position is known and an estimate of the range of possible values is made. In comparing a measured position with a prediction, the uncertainties of each should be kept in mind. The predicted point might not equal the measured position but remains a good prediction because it falls within the uncertainty of the measured point. Hence, whether one method provides a better prediction over another is affected by random errors in the point being predicted.

Comparing Methods Using Synthetic Beach Time Series

Synthetic beach time series provide an alternative to hindcasting and forecasting. With synthetic time series, the calculated change rates of each method are compared with an assigned rate. To do this, we first assign a true slope and use this to calculate a database of synthetic shoreline positions at eight discrete years. We then introduce noise to the shoreline database and use this noisy data to calculate change rates using the various methods. We repeat this process 1000 times while keeping the true rate constant but varying the noise. Noise, in this case, is the scattering of the shoreline position based on the known uncertainties and an additional unknown factor. The unknown factor makes our knowledge of the shoreline position less certain. Hence, the generated noise in the synthetic analysis is greater than the uncertainty of the model in real data. Noise is created by sampling from a probability density function associated with each major uncertainty component. For example, we quantify a rectification uncertainty based on the aerial photo orthorectification process. This has several independent errors associated with it, such as lens distortion, camera tilt, Earth curvature, and terrain relief. We assume that a Gaussian distribution is the underlying distribution for the rectification error due to the central limit theorem (DRAPER and SMITH, 1998) and sample from this distribution to represent noise caused by the rectification process.

To account for any additional uncertainties that are not part of the shoreline analysis, we also generate noise from a Laplace distribution and add it to each shoreline position. We sample from each uncertainty distribution and add the samples together to get a total noise value for each synthetic data point. We calculate a change rate for every synthetic time series and make a histogram of all the calculated rates to see which method consistently is closest to the true value.

To compare the methods, we use the Kolmogorov-Smirnov test (hereafter K-S test) to determine whether two distributions of calculated change rates are significantly different (the Gaussian assumption is not needed to use the K-S test). The K-S test is sensitive to the mean, standard deviation, and shape of each distribution (SIEGEL, 1956). Thus, this test shows whether different rate methods give statistically indistinguishable predictions. If more than one method has indistinguishable predictions, the choice of method is left to the discretion of the analyst. To test the methods at different noise levels, we calculate rates with synthetic data twice—one set of data with less noise and the other with more noise. In the less noise calculation, we sample from each uncertainty distribution that has a standard deviation that is equal to the average of the source of error plus a Laplace distribution with a standard deviation of 10 m. For the calculation of rates with more noise, we use the maximum value of each source of error component and the Laplace distribution with a standard deviation of 20 m.

We also compare methods when an extreme outlier exists, such as a storm-influenced shoreline. Since it is difficult to identify storm-influenced data points at specific beaches, we want to identify methods that best determine the actual rate with the inclusion of an extreme non-Gaussian point. We first add noise from a Laplace distribution with a standard deviation of 100 m to a middle data point (specifically, the fourth position) and then repeat the above process of calculating a change rate 1000 times. We then add extreme noise to the last point and repeat the above process. A K-S test is also performed on the resulting distribution of rates. As with the synthetic data without a storm-influenced point, we compute two different data sets—one with less noise and one with more noise.

Results of Rate Method Comparisons

The prediction analysis (forecasting and hindcasting) use real data to compare shoreline change rate methods. Although many caveats are associated with this analysis, forecasts and hindcasts suggest how well each rate method predicts future real data. On the other hand, the advantage of using the synthetic analysis is that the true synthetic rate is known, so the error of the predicted rate is also known. However, this analysis uses manufactured rather than real noise and assumes that errors are additive. Both prediction and synthetic results agree that MDL and AOR provide the least desirable results and OLS, WLS, RLS, RWLS, JK, and LAD are valid methods under certain conditions.

Forecasts

Excluding the North Shore, forecasts of positions with hardened structures have lower |EIP| than forecasts of positions free of hardened structures. The North Shore predictions (Table 3) with the 1960 storm position included in the analysis have considerably higher |EIP|, which can be attributed to one beach where predictions made from T-sheets and 1960 positions did not reflect actual positions (Figure 4). Generally, predictions from OLS, WLS, RLS, RWLS, JK, EPR, and LAD are statistically not different and have the smallest |EIP|. AOR and MDL are continuously singled out as methods that are significantly different from other methods because they show appreciably higher |EIP| (Table 3).

Predictions improve for all methods when storm-influenced shorelines are removed from the dataset. This improvement ranges from $<$0.1 m to 15 m. Geographically, the most improvement occurs on the North Shore with hardened structures. MDL and AOR improve more than other methods when outliers are removed (average not including North Shore = 3.6 m). All least squares methods, JK and LAD, on the other hand, have minimal improvements when storm-influenced shorelines are removed for all regions (average improvement excluding the North Shore = 0.7 m).

Hindcasts

The mean |EIP| for hindcasts is slightly larger than that of forecasts (Table 3). This increase is due to a lack of detailed information that surrounds T-sheets. A shoreline position from a T-sheet is not as well constrained and has a larger uncertainty associated with it than a shoreline position from an aerial photo. Hindcasts of beaches with hardened structures have lower mean |EIP| than beaches without hardened structures (excluding the North Shore). The hardened structures consist of seawalls, groins, and revetments, which may stabilize the beach on a short-term basis. Similar to forecasts, MDL and AOR hindcasts have high mean |EIP| and generally do not reflect the results of all other methods.

Removing storm-influenced outliers improves hindcasts minimally. Hindcasts of West Maui improve by only 1 m when storms are removed. Hindcasts of the North Shore with hardened structures are less accurate by an average of 2.4 m without storm shorelines (individual beaches have even a greater negative difference). This decrease in accuracy could be attributed to the storm shoreline having no negative influence on the trend; hence, it is not a true storm outlier for that area. All other areas, however, show marked improvement when storm-influenced outliers are excluded. AOR has

Table 3. *Forecast and hindcast results. EIP = error in prediction. Mean |EIP| is the average magnitude difference between predicted and known shoreline positions.*

		Forecasts				Hindcasts			
		Natural Beaches Mean |EIP| (m)		Hardened Beaches Mean |EIP| (m)		Natural Beaches Mean |EIP| (m)		Hardened Beaches Mean |EIP| (m)	
Region	Method	All Points	W/o Storms	All Points	W/o Storms	All Points	W/o Storms	All Points	W/o Storms
Kihei	OLS	11.7	9.3	7.4	7.0	20.3	16.4	11.9	11.5
	WLS	12.6	9.9	7.7	7.2	20.1	16.4	11.9	11.6
	RLS	11.8	9.3	7.7	7.0	20.3	16.4	11.9	11.5
	RWLS	12.6	9.9	8.0	7.2	21.1	16.4	12.0	11.6
	EPR	12.5	9.7	7.7	7.1	18.1	16.8	15.1	12.9
	AOR	19.0	12.7	14.1	8.2	27.6	17.0	30.0	14.1
	JK	13.6	10.8	8.5	7.4	21.0	15.6	11.8	10.6
	MDL LOW	20.3	12.4	14.1	8.2	17.0	15.2	13.3	11.9
	MDL ZERO	29.8	16.8	18.6	10.1	22.1	17.6	16.8	15.5
	LAD	11.7	9.5	7.2	6.8	22.9	21.9	12.5	15.0
West Maui	OLS	8.8	8.5	6.0	6.1	13.4	13.1	9.0	8.7
	WLS	9.2	8.7	6.2	6.2	13.4	13.1	8.7	8.5
	RLS	8.9	8.6	6.1	6.1	13.2	13.0	8.7	8.2
	RWLS	9.3	8.7	6.3	6.2	13.2	13.0	8.6	8.0
	EPR	9.6	8.8	6.2	6.2	14.5	12.9	10.2	10.0
	AOR	14.7	10.4	8.1	8.1	17.1	14.0	17.2	11.3
	JK	9.6	9.3	6.7	6.8	13.2	13.0	8.9	8.3
	MDL LOW	9.8	9.0	6.3	6.6	13.1	13.1	9.6	8.7
	MDL ZERO	11.6	10.5	8.3	8.5	13.6	13.2	10.8	10.0
	LAD	8.8	8.7	6.0	6.1	13.8	13.0	9.6	9.5
North Shore	OLS	5.2	4.9	20.4	10.0	13.3	9.1	26.0	31.2
	WLS	5.1	5.0	20.7	9.8	13.2	9.0	27.5	31.0
	RLS	5.2	4.9	20.4	10.0	13.4	9.1	26.1	30.8
	RWLS	5.2	5.0	20.6	9.8	13.3	9.0	27.6	30.7
	EPR	5.0	4.9	20.7	10.0	12.2	8.9	24.5	31.8
	AOR	5.5	5.3	21.9	11.6	13.5	9.3	27.5	35.5
	JK	4.9	4.8	20.8	10.4	13.5	9.1	26.4	33.0
	MDL LOW	5.1	4.9	24.0	10.2	12.4	8.0	41.5	35.2
	MDL ZERO	5.6	5.4	26.7	11.4	13.9	8.9	53.7	41.7
	LAD	5.0	4.8	20.1	9.6	13.1	9.3	26.4	29.7

the biggest improvement in hindcasts when shorelines affected by storms are removed.

Synthetic Data

Three sets of synthetic data were generated—one with no storm shoreline, one with a storm shoreline in a middle position, and one with a storm shoreline at the last position. For each set, two runs were made—one with less noise and one with more noise. The significance of two standard deviations was also calculated for all rate methods using a Siegel-Tukey test. A method with a low standard deviation better identifies the true slope than a method with a higher standard deviation. The results of the Siegel-Tukey test concur with the K-S test results.

Time Series without Storm Shoreline. K-S test results of time series with less noise show that WLS, RWLS, and WLAD predictions are not statistically different at the 95% confidence interval but are statistically different from all other methods (Table 4, column A). WLS, RWLS, and WLAD also have the smallest standard deviations, or data spread, and thus provide better predictions. K-S test results of time series with more noise show that all methods except for AOR and MDL perform equally well and are not significantly different from each other (Table 4, column B). AOR and MDL distributions have high spreads compared with all other methods.

Time Series with Storm Shoreline in Middle Position. K-S test results of data with less noise demonstrate that methods other than AOR and MDL are not statistically different (Table 4, column C). The spreads of MDL and AOR are higher than all other methods. For data with more noise, the K-S test results show that OLS, WLS, RLS, RWLS, JK, and WLAD are not statistically different (Table 4, column D). EPR and LAD are statistically different from other methods, but not statistically different from each other. EPR and LAD also have higher spreads than other methods, excluding AOR and MDL. AOR and MDL are statistically different from all other methods and have the highest spreads.

Time Series with Storm Shoreline in Last Position. K-S test results of data with less noise show that RLS, RWLS, LAD, and WLAD are not statistically different and have the lowest spreads (Table 4, column E). OLS, WLS, and JK do not have statistically different distributions from each other but have higher spreads than RLS, RWLS, LAD, and WLAD. EPR, AOR, and MDL have much higher spreads. When more noise is added to the time series, K-S test results show that all

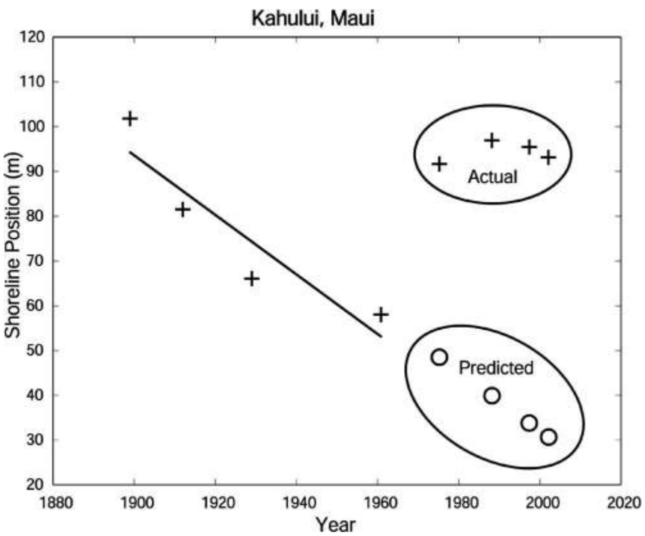

Figure 4. Predictions from a transect in Kahului. Predictions were made using the T-sheets and 1960 aerial photograph positions. Although a tsunami affected the area 5 mo before the 1960 photograph, accretion is predicted. Erosion at the rate of 3 m/y occurred between 1960 and the next position (1975).

methods except for EPR, AOR, and MDL are not statistically different (Table 4, column F). EPR, AOR, and MDL have higher spreads than all other methods.

Hawaiian Beaches and Binning of Shoreline Data

We calculate change rates from shore-normal transects spaced 20 m alongshore. Individual transects contain between five and nine unevenly spaced points spanning *ca.* 100 years. These datasets are typically limited and characterized by large data scatter. To reduce the effect of natural data variation from transect to transect and measurement errors, we bin data onto one plot from adjacent transects belonging to a contiguous stretch of beach and calculate change rates using the RWLS method. Since the whole beach does not necessarily behave in the same manner, we need to identify sections of the beach that are indistinguishable. Importantly, transects that have engineered structures and no beach fronting them are removed from this analysis because most coastal managers define these areas as having no erosion or accretion, and their inclusion will unduly influence rates calculated along the rest of the beach.

To identify which transects should be binned together and thus represent a section of beach where erosion rates are indistinguishable, we group adjacent transects and compare their combined rate to the combined rate of all other transects on that beach. The reader is referred to Figures 5–9 for a graphic illustration of the binning process. We start with a window spacing of four transects (Figure 5) and group the first four adjacent transects together and then calculate the shoreline change rate. We compare this rate with the rate of a bin of the remaining transects using a Student's t-test (Kleinbaum *et al.,* 1998) at a 95% confidence interval to determine any disagreements (Figure 5, bottom). The window is then shifted over by one transect and a new t-test is performed. The window continues shifting by one transect until the last four transects are grouped together. Each time the window shifts, a t-test is calculated to compare the grouped transects within the window with the binned rate of the rest of the transects. The window size is then increased from four transects to six and the process of calculating a t-test is repeated. The window size is increased and the binning procedure is repeated until we reach a window size equal to (n transects)/2. When the binning procedure is complete, clusters of transects are identified by executing a Student's t-test on groups of transects that are found to be statistically different from the rest of the beach. Within each window size, a t-test is performed on any overlapping transects (Figure 6). If the overlapping transects are statistically not different,

Table 4. *Kolmogorov-Smirnov results of synthetic time series. Each grouping contains methods that are not statistically significant from each other.*

Standard Deviation (data spread)	Without Storm Position		Storm—Middle Position		Storm—End Position	
	Less Noise (A)	More Noise (B)	Less Noise (C)	More Noise (D)	Less Noise (E)	More Noise (F)
Low	WLS	WLS	RWLS	JK	WLAD	RLS
	RWLS	JK	WLAD	OLS	RLS	JK
	WLAD	RWLS	JK	RLS	RWLS	OLS
		OLS	RLS	WLS	LAD	RWLS
		RLS	OLS	RWLS		WLAD
		WLAD	LAD	WLAD		WLS
		EPR	EPR			LAD
		LAD	WLS			
	RLS	AOR	MDL LOW	EPR	OLS	EPR
	JK			LAD	JK	
	OLS				WLS	
	LAD					
	EPR					
	AOR	MDL LOW	AOR	AOR	EPR	AOR
	MDL LOW	MDL ZERO	MDL ZERO	MDL LOW	AOR	MDL LOW
↓	MDL ZERO			MDL ZERO	MDL LOW	MDL ZERO
High					MDL ZERO	

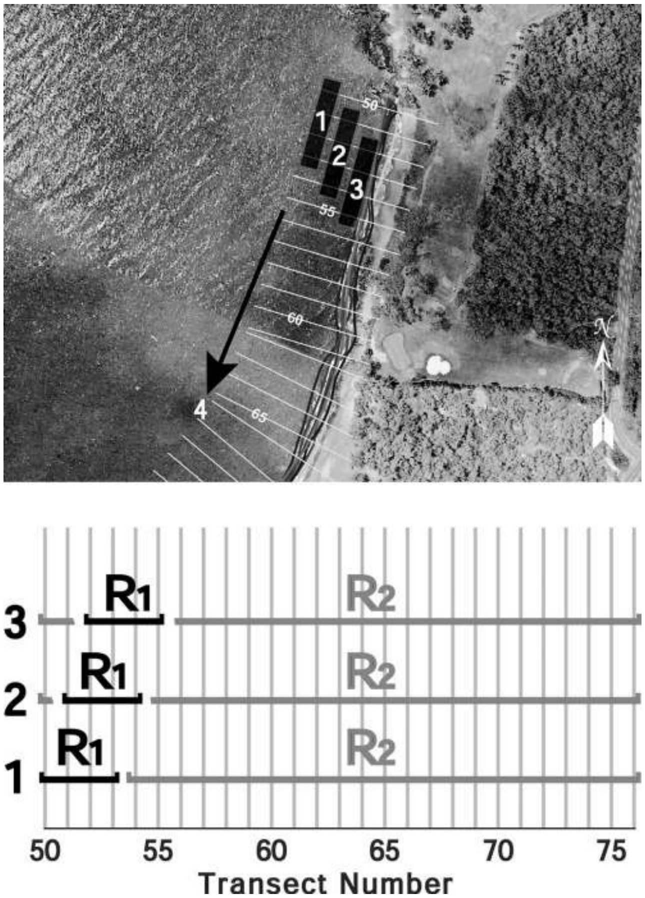

Figure 5. Top—Shore-normal transects are spaced 20 m alongshore. In this example, transect window length is four. The erosion rate of the first four adjacent transects (1) is compared with the rate of all the other transects in this beach. Next, the erosion rate of the next four transects (2) is compared with the rate of all other transects. This continues throughout the beach (3 and 4) until the rate of the last four transects is compared with the rest of the beach. The window spacing then increases to 6, 8, 10, *etc.* transects. Bottom—Another visualization of transect binning at a window size of 4. The numbers 1, 2, and 3 correspond to the numbers in the figure above. R_1 is the binned erosion rate of four transects. R_2 is the binned erosion rate of the rest of the beach. A t-test examines whether R_1 is significantly different from R_2.

then they are grouped together as one bin (Figure 7A). If they are different, then they are grouped separately (Figure 7B). Another t-test is performed to determine whether the bins at each window size are statistically different from bins of other window sizes that have overlapping transects. Bins that are found to be statistically not different are clustered together, and a rate is calculated for that region (Figure 8). For visual purposes, each cluster is assigned a color and each transect within a cluster is allotted a shade (Figure 9). The shade depends on the frequency of windows that intersect a given transect—transects that contain a higher frequency of windows that belong to the same cluster will be a darker shade than transects that have a lower frequency of windows. For example, in Figure 8, transect 57 has 27 windows that intersect it, while transect 51 has only three intersecting windows.

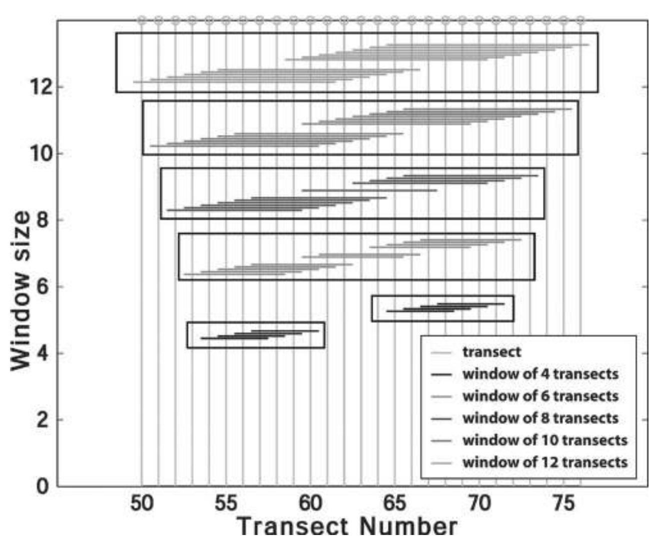

Figure 6. Vertical lines represent transects. Y axis is the window size of the transect grouping. The first window size is four, meaning four adjacent transects are grouped for each rate calculation. Horizontal lines represent groups of transects whose rates did not agree with the binned rates of all other transects (t-test). A t-test is calculated for any overlapping transects within a window size (represented by boxes).

As a result, transect 57 will be darker than transect 51 (Figure 9). Transects that encompass more than one cluster will have a mixed color value.

Results of Binning

Based on both the forecasting and synthetic time series results, we chose to bin data using the RWLS method. RWLS

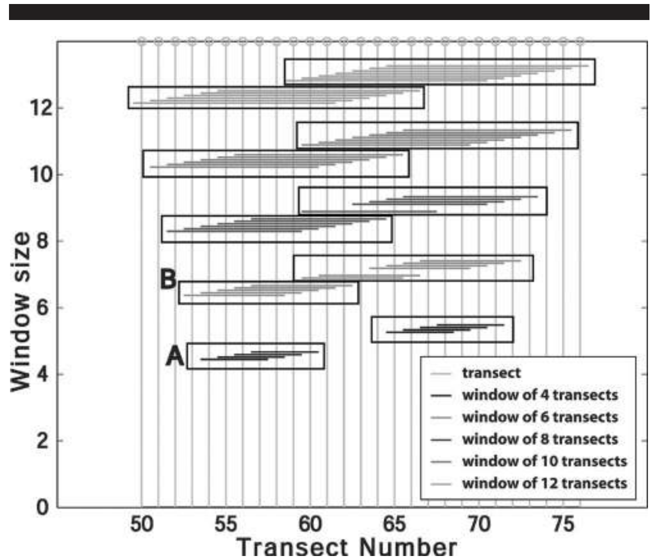

Figure 7. T-test results within each window size. At window size four, the overlapping transects are not significantly different from each other (A). At window size six, two groups were identified that were significantly different from each other (B).

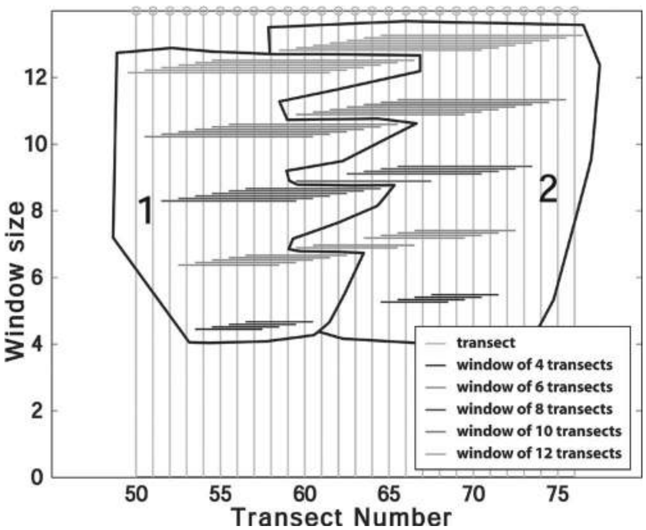

Figure 8. Clusters of transects (or cells) are identified based on t-tests of all windows with common transects. For this beach, there are two distinct clusters (labeled 1 and 2). One side of the beach has significantly different long-term rates compared with the other side of the beach. Transects that are shared by both clusters are considered transitional zones.

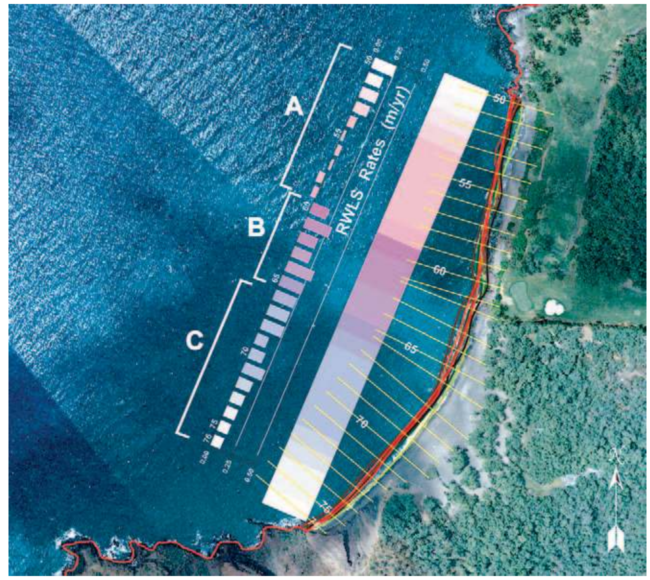

Figure 9. Each cluster is assigned a color, and each transect within a cluster is given a shade of that color. Transect shade corresponds to the frequency of windows that intersect the transect. Transects with a higher frequency of windows will be darker than transects with a lower frequency of windows. Groups A and C represent two distinct clusters. Group B represents the overlap between groups A and C.

is chosen because we are confident in our knowledge of the major uncertainties in our shoreline methodology and are not confident of the identification of storm-influenced shorelines. Binned rates and their uncertainties are better resolved than unbinned rates and their uncertainties. Eighty-four percent of the trends of binned rates are significant, whereas only 38% of the trends of unbinned rates are significant. There is a 0.1-m decrease in uncertainties with binned rates compared with unbinned rates. We perform the binning analysis on 15 beaches—eight from Kihei, four from West Maui, and three from the North Shore. The eight beaches of Kihei display two different patterns and are further categorized into two geographical groups within this study site—four central beaches and four southern beaches (Table 5).

Kihei is sheltered from waves by Kahoolawe and Lanai islands but is subjected to south swells, refracted north swells, and kona storm waves. No fringing reef is present in our study area (Fletcher *et al.,* 2002; Makai Ocean Engineering and Sea Engineering, 1991; Rooney and Fletcher, 2005). The central beaches in Kihei behave uniformly. Only one of these central beaches contains a seawall, but all four have significant erosion. The southern beaches depict distinct behaviors between their northern and southern sections (Table 5).

West Maui beaches are affected by North Pacific swells,

Table 5. *Binning trends.*

Region	Beach	Wave Climate of Region	Offshore Bottom	Structures	Overall Bin Trends
Kihei	Southern beaches (Big Beach, Little Beach, Maluaka, Onuli)	South swell, refracted north swell, and tradewind waves	No fringing reef	None	Northern portions of a beach are distinct from southern portions.
Kihei	Central beaches (Kam 1, Kam 2, Kam 3, Ulua)	Minimal south swell and refracted north swell, occasional kona storm waves	No fringing reef, sandy bottom	Seawall (on Ulua)	Uniformity within beaches
West Maui	Kaanapali, North Kaanapali, Keonenui, Kapalua	North Pacific swell, south swell, and kona storm waves	Portions have fringing reefs, rocky and sandy bottoms	Seawall (on Keonenui)	Erosion on one end of the beach and accretion on the other end of the beach
North Shore	Kaehu, Kanaha, Spreckelsville	North Pacific swell and tradewind waves	Fringing reefs	Five groins throughout beach (on Kanaha), offshore rock platform, and revetment (on Spreckelsville)	Eastern and western sections of the beach are distinct

south swells, and kona storm waves. North Pacific swells do not severely affect this area because it is located in the shadow of Molokai Island (Eversole and Fletcher, 2003; Fletcher *et al.,* 2002). Segments of the shoreline in this area contain fringing reefs, and only one beach has a seawall (Eversole and Fletcher, 2003; Makai Ocean Engineering and Sea Engineering, 1991). West Maui study sites depict a pattern of erosion at one end of each beach and either minimal erosion or accretion at the other end (Table 5).

North Shore beaches are influenced by North Pacific swells and tradewind waves. Fringing reefs are widespread in this area (Fletcher *et al.,* 2002; Makai Ocean Engineering and Sea Engineering, 1991). The North Shore study area illustrates distinct behaviors between eastern and western sections of each beach. Two of the three beaches have engineered structures—one has five groins and one has an offshore rock platform with an onshore revetment. These two beaches exhibit more erosion than the beach without any structures (Table 5).

DISCUSSION

Comparisons of Rate Methods

Based on our synthetic work and forecasting analysis, we advise analysts to determine how well they understand the uncertainty in their methodology before picking the best erosion rate method. When uncertainties are mostly understood, weighted methods are recommended; conversely, if uncertainties are poorly understood, OLS, RLS, JK, and LAD are recommended. If doubt exists on the Gaussian nature of the uncertainty, LAD and WLAD are recommended. We are confident in our understanding of the major uncertainty components in our shoreline assessment study; however, we are not confident in our assessment of storm-influenced shorelines. As stated earlier, Hawaii is an island state and different beaches are affected differently by storms. The difficulty in identifying storm-influenced shorelines, together with the results from the synthetic storm analysis and the prediction analysis, is the basis for choosing RWLS as the most robust method for Hawaii.

Based on the K-S test results of synthetic time series with less noise, weighted methods are far superior to all other methods (Table 4, columns A, C, and E). This superiority disappears as noise increases. As more noise is introduced to the data, the weights no longer reflect the uncertainty, and other methods that put more emphasis on the uncertainty in the model achieve better results (Table 4, columns B, D, and F). OLS, RLS, JK, and LAD outperform weighted methods when noise is sampled solely from a Laplace distribution. If, however, the majority of uncertainty components are known, weighted methods reflect the true process better (Table 4, column A).

Synthetic time series results vary when storm shorelines are added to middle and end positions. The bias created by the storm shoreline at the middle position is not as detrimental in finding the long-term trend as the bias with a storm shoreline at the end position. This is because an end point is a leverage point that can influence the trend of the model more than interior points. With the storm at the end position, WLS, OLS, and JK do not perform as well as RLS, RWLS, WLAD, and LAD (Table 4, column E). Since WLS, OLS, and JK do not identify statistical outliers and the storm shoreline is at a leverage position, the storm shoreline unduly influences the results. RWLS and RLS remove statistical outliers, and LAD and WLAD are less susceptible to outliers in the y direction (Rousseeuw and Leroy, 1987); hence, they perform better than all other methods when a storm shoreline exists at the end position. When noise increases, however, all methods perform similarly, except for EPR, AOR, and MDL (Table 4, column F).

Forecasting results show that in most cases, methods with the highest mean $|EIP|$ are statistically different from methods with lower mean $|EIP|$. OLS, WLS, RLS, RWLS, EPR, and LAD consistently have the lowest mean $|EIP|$ and are insignificantly different from each other.

MDL and AOR have the highest mean $|EIP|$. Their predictions also have the most improvement when storm shorelines are removed. Both methods remove data from the rate calculation. The end point rate combinations that incorporate storm-influenced shorelines strongly affect the AOR rate. Once the storm is removed from the data, AOR predictions improve. These are still less effective than the least squares methods, since AOR depends on the minimum time criterion, which decreases the number of EPRs available for averaging. The MDL ZERO line often discards early shoreline positions that do not fit the pattern of the most recent trend of the data. Predictions made with the most recent trend are more variable than predictions made with a longer term trend. Removing the storm-influenced shorelines influences the more recent trends and results in improved, yet still variable, predictions. The MDL LOW line has better predictions than the MDL ZERO line because it does incorporate the early shoreline positions, which agrees with the results of Crowell, Douglas, and Leatherman (1997). The predictions with the low-weight line do not always perform as well as the least squares because they still give more weight to the more recent data points.

In all forecasts, except for the North Shore with hardened structures, predictions exclusive of storm-influenced shorelines improve only slightly and are insignificant in most cases (Table 3). For the North Shore with hardened structures, the mean $|EIP|$ is extremely large ($>$20 m) when the 1960 storm (in this case, tsunami) shoreline is present. In examining individual beaches on the North Shore, the mean $|EIP|$ for Kanaha does not show improvement when the 1960 tsunami position is removed; rather the predictions are worse (Table 6). Kahului, on the other hand, is responsible for the large mean $|EIP|$. The Kahului analysis is composed of three early T-sheets (1899, 1912, and 1929), and five later aerial photo positions. Forecasts made with three T-sheets and the 1960 aerial photo account for the large mean $|EIP|$. The shoreline change rates calculated with these four points indicate accretion or minor erosion. However, erosion averaged 2.80 m/y between 1960 and 1975. This significant increase in erosion resulted in forecasts that do not reflect the actual positions (Figure 4). Comparing the Kahului results without this set of forecasts reveals that all methods, excluding MDL, do not improve significantly when the 1960 position is taken out

Table 6. *Forecasts at Kanaha (North Shore). Removing the 1960 tsunami position worsens the prediction.*

Method	Mean |EIP| (m) All Points	Mean |EIP| (m) W/o Storms
OLS	11.3	12.9
WLS	11.4	12.8
RLS	11.2	12.9
RWLS	11.4	12.8
EPR	11.4	13.0
AOR	11.5	13.1
JK	11.3	12.9
MDL LOW	11.3	12.5
MDL ZERO	11.3	12.5
LAD	11.3	13.1

Table 7. *Forecasts at Kahului (North Shore) after removing predictions from data using T-sheet and the 1960 shorelines.*

Method	Mean |EIP| (m) All Points	Mean |EIP| (m) W/o Storms
OLS	8.9	9.2
WLS	10.1	9.1
RLS	9.0	9.2
RWLS	10.1	9.1
EPR	9.2	9.2
AOR	11.0	11.4
JK	8.9	9.8
MDL LOW	21.6	9.5
MDL ZERO	28.8	10.8
LAD	10.2	8.7

(Table 7). When segmenting the study area into individual beaches, some beaches have no significant improvement at the 95% confidence interval when storm-influenced shorelines are removed, while other beaches do. This could be due to differing storm or tsunami effects. Some beaches are protected from the full force of the waves by surrounding islands or fringing reefs, while others are more exposed.

The EIPs for hindcasts are somewhat greater than those of forecasts. This is because T-sheet positions usually have greater uncertainties than aerial photos (±7–10 m). Hindcast predictions that are less than 10 m away from the true position are still within the uncertainty bounds of the true positions. We calculate a 95% confidence interval uncertainty around the predicted positions (Table 8). The minimum average uncertainty of the predicted positions is 15 m. By incorporating the uncertainties of both the true and predicted positions, even the North Shore predictions fall within the uncertainty bands. Also, removing storm-influenced shorelines does not improve the hindcasts. Thus, early twentieth century T-sheet positions are valuable in the shoreline change rate analyses.

Outliers

We investigated two types of outliers and their influence on the accuracy of predicting shoreline behavior: (1) *a priori* knowledge based on historical data, such as a tsunami, hurricane, or storm event and (2) outliers based on residual statistics.

The time necessary for a shoreline to recover from a major erosional event can vary (ZHANG, DOUGLAS, and LEATHERMAN, 2002), resulting in non-Gaussian behavior. ZHANG, DOUGLAS, and LEATHERMAN (2002) argue that storms are independent of any long-term trend and should be considered separately because beaches eventually recover to their prestorm positions. With limited data sets, such as individual transects that have only five to eight points, a storm-influenced shoreline may unduly bias any calculation of a long-term trend. ZHANG, DOUGLAS, and LEATHERMAN (2002) support the assertion of DOUGLAS and CROWELL (2000) that the most practical option is to remove these points. In our study, however, we identify two storm-influenced shorelines—the 1960 tsunami that affected the north shore of Maui and the 1963 kona storms. When we use the dataset to predict the position of a known shoreline at each beach, removing storm shorelines improves our prediction by an average of 1.1 m (least squares), with the exception of one beach that experienced accretion during a storm event. This improvement is minimal when compared with the results of HONEYCUTT, CROWELL, and DOUGLAS (2001) from U.S. East Coast beaches, which demonstrate an improvement of 15–30 m when storm shorelines are removed. We note, however, that carbonate beaches in general and Hawaiian beaches specifically tend to be much narrower than East Coast beaches. An improvement of 1 m may represent 5% of the dry beach width in many cases. The cost of removing outliers from small datasets, typically used in erosion analysis, is usually an increase in the uncertainty of the calculated long-term trend. In the end, an analyst must weigh the cost of increased un-

Table 8. *Uncertainties of known positions and hindcasted positions with and without storm data.*

Area	Year	Known Position Uncertainty (m)	95% C.I. All Points Average Predicted Position Uncertainty (m)	95% C.I. W/o Storms Average Predicted Position Uncertainty (m)
Natural Beaches				
Maluaka	1931	8.81	15.71	23.67
Onuoli	1931	8.81	23.31	38.67
Kam 1	1912	10.57	114.63	132.41
Kam 2	1912	10.57	64.79	117.75
Kam 3	1912	10.57	88.43	105.44
N. Kaanapali	1912	9.17	28.03	27.16
	1932	7.78	22.71	21.90
Waiehu	1899	6.55	28.04	62.11
	1912	6.76	23.00	49.90
Waihee	1912	6.44	22.73	87.38
	1929	6.25	18.52	69.88
Engineered Beaches				
Honokowai	1912	8.39	87.62	173.37
Kahului	1899	8.45	68.69	52.57
	1912	8.61	61.24	46.37
Kanaha	1912	9.17	56.76	131.00
	1929	7.62	46.54	105.26
N. Kihei	1900	7.45	42.67	63.03
Ulua	1912	10.42	15.16	35.11
Sprecklesville	1912	10.18	141.39	919.97

certainty against the benefit of improved predictive accuracy. We conclude that storm-influenced shorelines need to be investigated at each study site before deciding on their treatment.

Most researchers currently do not remove outliers unless they can assign the points to some meteorological or geological factor. Statistical outliers, however, are critical components to consider when using least squares because of this method's susceptibility to outliers, especially for small datasets. A large deviation at a point causes a bias in the trend if there are few data points, and this can be amplified if the point is at a leverage position. Some studies have attempted to relate statistics, such as residuals, to outliers. Focusing on reducing the root mean squared error (RMSE), GALGANO, DOUGLAS, and LEATHERMAN (1998) state that storm-influenced residuals increase the error in the model fit and thus invalidate the model. They choose to use *a priori* information to remove these points, but the erosion rates with and without these points are not significantly different. FENSTER, DOLAN, and MORTON (2001) identify outliers by calculating studentized residuals and compare them with known storm dates. They find that none of the statistical outliers correspond to any known storms and advise not to remove them as outliers.

FLETCHER *et al.* (2003) and ROONEY *et al.* (2003) identify and remove statistical outliers differently. As mentioned earlier, the least median of squares (LMS) method is part of a two-step process involving RLS or RWLS that calculates residuals (ROUSSEEUW and LEROY, 1987). They disregard a data point if a residual is greater than an assigned cutoff value. The cutoff value ($\hat{\sigma}$) is an estimate of the true standard deviation of a population (σ), which is dependent on the sample size. For small sample sizes, there is less certainty in any estimate of the true σ, making the cutoff boundary less exact and causing the removal or retention of too many outliers. If the outlier analysis at adjacent transects identifies different points as outliers, the resulting erosion rates are also likely to differ, leading to the case where physically adjacent beach segments are assigned inconsistent long-term trends. In our dataset with an alongshore spacing of 20 m, we find that adjacent transects do not behave independently of each other.

One way to use LMS is to increase the number of points used in the calculation of a trend, which can be done by binning data from adjacent groupings of transects and calculating a trend. Binning will reduce the spread, or uncertainty, around the cutoff value, which will improve the identification of outliers. Therefore, we recommend using a relatively large sample size to increase the signal-to-noise ratio and improve the estimate of the spread of the data ($\hat{\sigma}$) when removing statistical outliers.

Binning Analysis

Setbacks on Maui currently are based on erosion rates from transects spaced 20 m alongshore. Some adjacent transects have differing rates, which affect the setback location. Because only five to nine historical shoreline positions are available, the noise in the data can mask the signal. One advantage of binning is that by spatially increasing our points, we decrease the noise by averaging out the random errors. This decrease in noise allows us to better identify a region of a beach that has indistinguishable rates of change (*i.e.,* sublittoral cells) and assign it one rate of shoreline change. Temporally, our data are unchanged—we can only increase the number of temporal positions by adding more photographs or T-sheets. Coastal planners will then be able to use one rate to determine the setback for that subcell of beach.

When comparing trends of all beaches, the four central beaches in Kihei behave similarly. All four beaches are relatively small pocket beaches with highly developed backshores. Their location, in the shadow of the islands of Molokai, Lanai, and Kahoolawe, protects them from large swells; however, kona storms have a history of inflicting great damage to this area (FLETCHER *et al.,* 2003; ROONEY and FLETCHER, 2000, 2005). In a study of net sediment transport on a stretch of armored beach just north of these beaches, ROONEY and FLETCHER (2000) conclude that tradewind waves cause southward movement of sediment, though a northward movement of sediment predominates due to kona storm activity. In actuality, we do not see such movement of sediment. Rather, a uniform manner of erosion is characteristic throughout each beach. This uniformity could be due to the relatively small size of the beach that is evenly affected by both kona storm waves and tradewind waves.

The southern beaches in Kihei are more susceptible to south swells and have more variability in binning results. These beaches are less developed and less eroded than their central counterparts. A cinder cone divides these four beaches into two northwest facing and two southwest facing beaches. A very small pocket beach situated within the south end of the cinder cone accretes uniformly. The cinder cone protects it from both north and south waves. The cinder cone interferes with sediment transport on the beaches directly north and south of it. The pocket beach directly south of the cinder cone exhibits extensive erosion near the headland. The pocket beach directly north of the cinder cone exhibits the most erosion in the central portion of the beach, while the northern section shows either minimal erosion or accretion. This follows the conclusion of ROONEY and FLETCHER (2000) of net sediment transfer to the north.

Two of the three beaches on the North Shore have engineered structures that influence the results. The North Shore is affected by strong North Pacific swells in the winter and strong, consistent tradewind waves throughout the year. One beach contains five groins—four successive groins at the western end and one groin at the easternmost point of the beach. The western end is less erosive than the eastern end. The groin at the easternmost point of the beach reduces the amount of sediment delivery to the beach directly west of it. The groins were installed to slow the alongshore sediment transport to the west (MAKAI OCEAN ENGINEERING AND SEA ENGINEERING, 1991) but have caused extensive erosion. Another beach has an offshore rock platform, and a revetment in the center of the beach that has caused considerable erosion.

We compare our binning results (Figure 10) to the sediment transport study of EVERSOLE and FLETCHER (2003) at Kaanapali Beach in Maui in order to relate annual transport

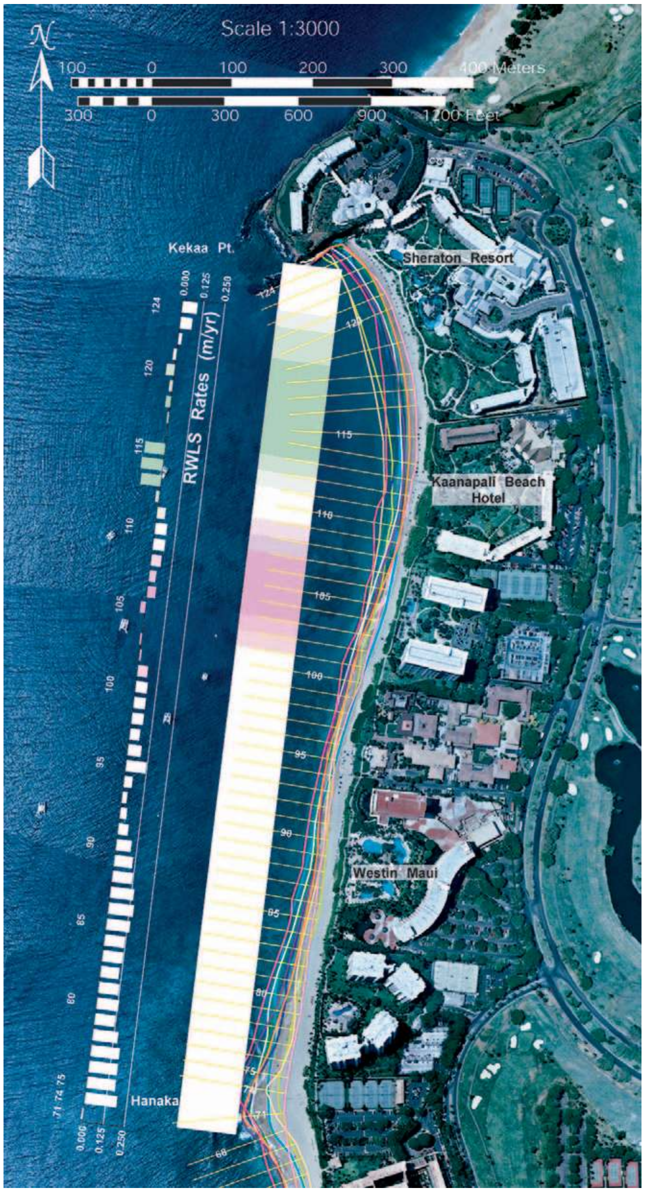

Figure 10. Kaanapali, Maui. Results from binning. Eversole and Fletcher (2003) identify transect 105 as an inflection point.

to multidecadal observations. EVERSOLE and FLETCHER (2003) examine sediment transfer by longshore sand transport at the Kaanapali Beach littoral cell and conclude that transport of sand is northward in summer and southward in winter, although the net annual transport is northward. They identify an inflection point within the littoral system and demonstrate a seasonal volume shift at this position, which also approximates the net annual transport in this location. In our binned analysis, the same region is clustered together and identified as a group that is accreting. We find that this group differs from the southern portion of the beach, which is eroding long term at a much faster rate. Although we are not able to identify seasonal fluxes, we do observe erosive south and accretive north sections that agree with the observation of EVERSOLE and FLETCHER (2003) of net annual transport to the north.

Erosion Hazard Maps

Erosion hazard maps are important in identifying setbacks that are used by coastal planners. Based on the results of our study, we have produced erosion hazard maps that reflect the trend of shoreline movements more accurately than previously used maps (Figure 11). These new maps incorporate uncertainties of the shoreline change rate method and identify binned regions of a beach. The drawbacks of these maps include assumptions of a linear shoreline behavior, with no change in long-term effects of storms.

There are three major steps coastal managers need to take in order to produce erosion hazard maps (Figure 12). The first step is to identify a change rate method based on their data. The next step is to bin the data. For example, two bins are identified in Figure 11—one on the eastern end of the beach and one on the western end. Finally, transects of each bin are grouped together and a 50-year predicted position with a $1 - \sigma$ uncertainty is identified by projecting the regression line into the future. The setback is calculated with $1 - \sigma$ confidence bands placed on either side of the setback, which creates a hazard zone. Maui County measures the setback from the certified shoreline. The vegetation line is most often used as a proxy for the shoreline and includes a 6.1-m (20-ft) buffer that is designed to partially compensate for method errors, storm and tsunami hazards, and nonlinear shoreline change. In our analysis, the 50-year predicted position is calculated from low water line data; consequently, a vegetation line offset and buffer are added to the future position before the setback is projected onto the map.

CONCLUSIONS

By comparing the shoreline change rate methods and investigating outliers, we make the following conclusions. (1) OLS, RLS, WLS, RWLS, JK, LAD, and WLAD are preferred methods based on results from synthetic time series and forecasts. If major uncertainties in a methodology are known and quantifiable, WLS, RWLS, and WLAD are preferred. If uncertainties are unknown or not quantifiable, LAD is preferred, although OLS, RLS, and JK can be considered. If effects of storms are unknown or storm-influenced shorelines are hard to identify, RLS, RWLS, LAD, and WLAD are preferred. (2) We choose to use RWLS on Maui as our method based on our knowledge of uncertainties and our lack of confidence in identifying storm-influenced shorelines. (3) MDL and AOR produce the most variable results. (4) Early twentieth century T-sheets are valuable in shoreline change rate analysis. (5) Hardened shorelines reduce variability of beach behavior. (6) Increasing the number of data points via binning neighboring transects in a RWLS or RLS analysis improves the estimate of spread in data when identifying statistical outliers.

We conclude from the binning analysis that (7) binning adjacent transects improves the signal-to-noise ratio. The re-

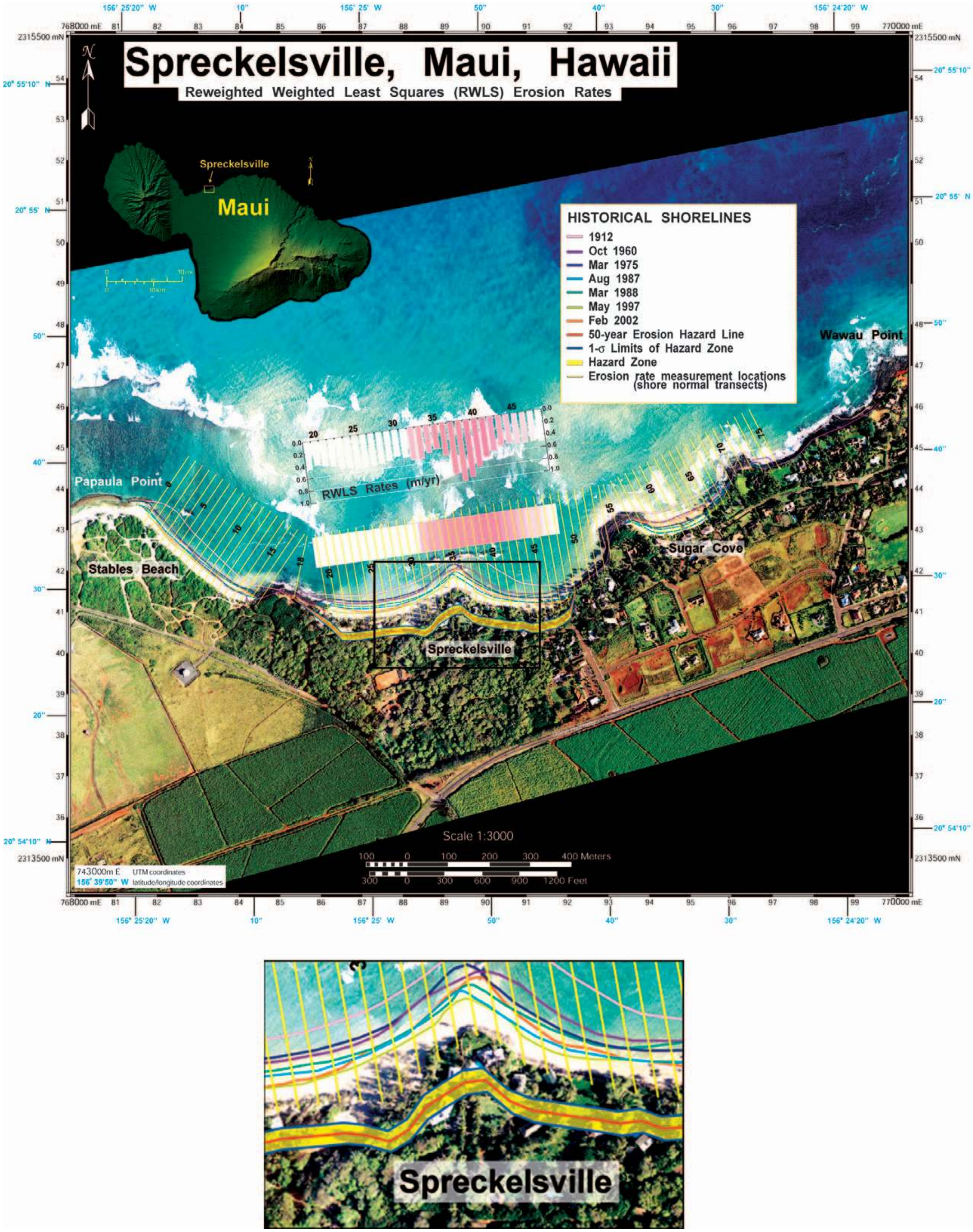

Figure 11. An example of an Erosion Hazard Map. The setback (red line) is surrounded by $1 - \sigma$ confidence bands (blue lines). The hazard zone is highlighted in yellow.

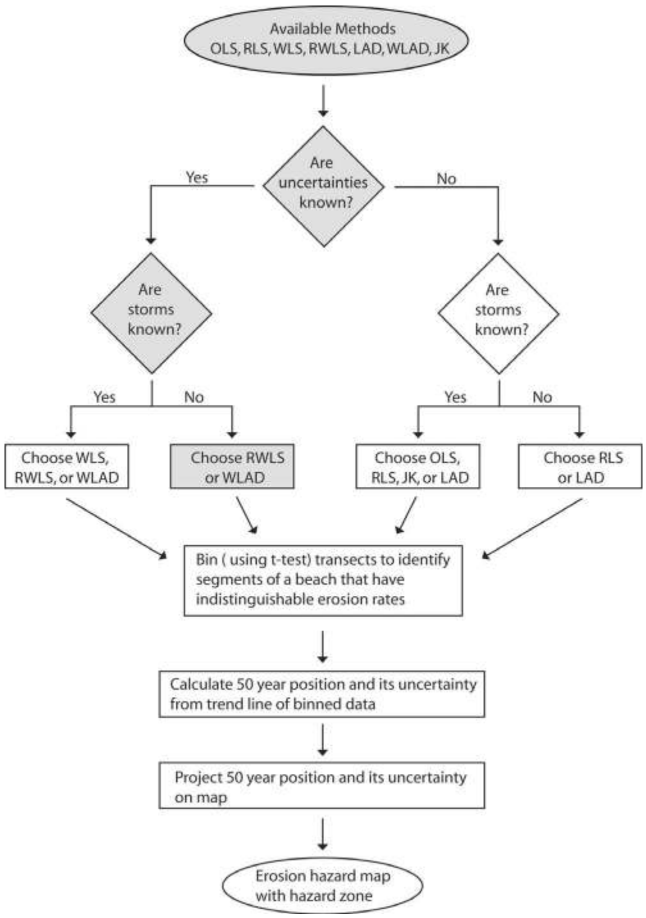

Figure 12. Flow chart showing steps in creating a hazard map. Methods appropriate for Hawaii are highlighted in gray. We choose RWLS over WLAD because of the simplicity of RWLS.

sulting binned rates reflect long-term sand transport within a littoral cell.

ACKNOWLEDGMENTS

The authors would like to thank Christopher Conger, Siang-Chyn Lim, Mathew Barbee, and William Morrison of the University of Hawaii Coastal Geology Group for their advice and support. The authors would like to acknowledge Dr. Michael Fenster for graciously providing us the MDL executable program.

LITERATURE CITED

Coyne, M.A.; Fletcher, C.H., and Richmond, B.M., 1999. Mapping coastal erosion hazard areas in Hawaii: observation and errors. *Journal of Coastal Research,* Special Issue No. 28, pp. 171–184.

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.

Crowell, M.; Honeycutt, M., and Hatheway, D., 1999. Coastal erosion hazards study: phase one mapping. *Journal of Coastal Research,* Special Issue No. 28, pp. 10–20.

Dean, R.G. and Malakar, S.B., 1999. Projected flood hazard zones in Florida. *Journal of Coastal Research,* Special Issue No. 28, pp. 85–94.

Dolan, R.; Fenster, M.S., and Holme, S.J., 1991. Temporal analysis of shoreline recession and accretion. *Journal of Coastal Research,* 7(3), 723–744.

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

Douglas, B.C.; Crowell, M., and Leatherman, S. P., 1998. Considerations for shoreline position prediction. *Journal of Coastal Research,* 14(3), 1025–1033.

Draper, N.R. and Smith, H., 1998. *Applied Regression Analysis,* 3rd edition. New York: John Wiley and Sons, Inc.

Eversole, D., 2002. Large-Scale Beach Change: Kaanapali, Hawaii. Manoa: University of Hawaii, Master's thesis, 63p.

Eversole, D. and Fletcher, C.H., 2003. Longshore sediment transport rates on a reef-fronted beach: field data and empirical models Kaanapali Beach, Hawaii. *Journal of Coastal Research,* 19(3), 649–663.

Fenster, M. and Dolan, R., 1994. Large-scale reversals in shoreline trends along the U.S. mid-Atlantic coast. *Geology,* 22, 543–546.

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.

Fenster, M.S.; Dolan, R., and Morton, R.A., 2001. Coastal storms and shoreline change: signal or noise? *Journal of Coastal Research,* 17(3), 714–720.

Fletcher, C.H.; Grossman, E.E.; Richmond, B.M., and Gibbs, A., 2002. *Atlas of Natural Hazards in the Hawaiian Costal Zone.* U.S. Geologic Survey, 182p.

Fletcher, C.H.; Mullane, R.A., and Richmond, B.M., 1997. Beach loss along armored shorelines on Oahu, Hawaiian Islands. *Journal of Coastal Research,* 13(1), 209–215.

Fletcher, C.H.; Rooney, J.J.B.; Barbee, M.; Lim, S.-C., and Richmond, B.M., 2003. Mapping shoreline change using digital orthophotogrammetry on Maui, Hawaii. *Journal of Coastal Research,* Special Issue No. 38, pp. 106–124.

Foster, E.R. and Savage, R.J., 1989. Methods of historical shoreline analysis. *In: Coastal Zone '89.* New York: American Society of Civil Engineers, pp. 4434–4448.

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

Galgano, F.A.; Douglas, B.C., and Leatherman, S.P., 1998. Trends and variability of shoreline position. *Journal of Coastal Research,* Special Issue No. 26, pp. 282–291.

Graybill, F.A. and Iyer, H.K., 1994. *Regression Analysis: Concepts and Applications.* Belmont, California: Duxbury Press, 701p.

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.

Kleinbaum, D.G.; Kupper, L.L.; Muller, K.E., and Nizam, A., 1998. *Applied Regression Analysis and Other Multivariable Methods,* 3rd edition. Pacific Grove, California: Duxbury Press, 798p.

Makai Ocean Engineering and Sea Engineering, 1991. *Aerial Photograph Analysis of Coastal Erosion on the Islands of Kauai, Molokai, Lanai, Maui and Hawaii.* Honolulu: Office of State Planning, Coastal Zone Management Program, 200 p.

Morton, R.A., 1991. Accurate Shoreline Mapping: Past, Present, and Future. *In: Coastal Sediments '91.* (ASCE) pp. 997–1010.

Norcross-Nu'u, Z.M. and Abbott, T., 2005. Adoption of erosion rate-based setbacks in Maui, Hawaii: observations and lessons learned. Solutions to Coastal Disasters Conference (Charleston, S.C.).

Ocean Studies Board, 1999. *Science for Decision-Making.* Washington D.C.: Commission on Geosciences, Environment and Resources, 113p.

RISSANEN, J., 1989. *Stochastic Complexity in Statistical Inquiry.* Singapore: World Scientific, 178p.

ROONEY, J.J.B., 2002. A Century of Shoreline Change along the Kihei Coast of Maui, Hawaii. Manoa: University of Hawaii, Ph.D. dissertation, 174p.

ROONEY, J.J.B. and FLETCHER, C.H., 2000. A high resolution, digital, aerial photogrammetric analysis of historical shoreline change and net sediment transport along the Kihei coast of Maui, Hawaii. Thirteenth Annual National Conference on Beach Preservation Technology (Melbourne, Florida).

ROONEY, J.J.B. and FLETCHER, C.H., 2005. Shoreline change and pacific climatic oscillations in Kihei, Maui, Hawaii. *Journal of Coastal Research,* 21(3), 535–547.

ROONEY, J.J.B.; FLETCHER, C.H.; BARBEE, M.; EVERSOLE, D.; LIM, S.-C.; RICHMOND, B.M., and GIBBS, A., 2003. Dynamics of sandy shorelines in Maui, Hawaii: consequences and causes. Coastal Sediments 2003 Proceedings, (Clearwater Beach, Florida).

ROUSSEEUW, P.J. and LEROY, A.M., 1987. *Robust Regression and Outlier Detection.* New York: John Wiley and Sons, Inc., 329p.

SEBER, G.A.F. and LEE, A.J., 2003. *Linear Regression Analysis,* 2nd edition. Hoboken, New Jersey: John Wiley and Sons, Inc., 557p.

SIEGEL, S., 1956. *Nonparametric Statistics for the Behavioral Sciences.* New York: McGraw-Hill Book Company, Inc., 312p.

TARANTOLA, A., 1987. *Inverse Problem Theory Methods for Data Fitting and Model Parameter Estimation.* New York: Elsevier, 613p.

THIELER, E.R.; RODRIGUEZ, R.W., and CARLO, M., 1995. Beach erosion and coastal development at Rincón, Puerto Rico. *Shore and Beach,* 63(4), 18–28.

ZHANG, K.; DOUGLAS, B., and LEATHERMAN, S., 2002. Do storms cause long-term beach erosion along the U.S. East Barrier Coast? *Journal of Geology,* 110(4), 493–502.