ELSEVIER

Marine Geology 191 (2002) 71–86

MARINE GEOLOGY
INTERNATIONAL JOURNAL OF MARINE GEOLOGY, GEOCHEMISTRY AND GEOPHYSICS

www.elsevier.com/locate/margeo

Analysis of the scale of errors in nearshore bathymetric data

Nathaniel G. Plant *, K. Todd Holland, Jack A. Puleo

Naval Research Laboratory, Code 7440.3, Stennis Space Center, MS 39529, USA

Received 26 June 2001; accepted 7 August 2002

Abstract

Most studies of nearshore hydrodynamics, sediment transport, and morphology focus on bathymetric variability within a narrow band of spatial and temporal scales. Typically, these studies rely on bathymetry estimates derived from field observations consisting of discrete samples in space and time with varying degrees of measurement error. Sampling limitations, which result in aliasing, and measurement errors can significantly contaminate variability at resolved scales, and may lead to large errors in the representation of the scales of interest. Using a spectral analysis, interpolation errors were analyzed for three different nearshore bathymetric data sets, each of which targeted a different range of spatial scales. Bathymetric features that were unresolved or poorly resolved (e.g. beach cusps) introduced the potential for contamination in two of the data sets. This contamination was significantly reduced using an appropriate scale-controlled interpolation method, leading to more accurate representations of the actual bathymetry. An additional benefit of using scale-controlled interpolation is that interpolation errors may be estimated independently of actual observations, which allows one to design bathymetric sampling strategies that ensure that dominant scales are either resolved or largely removed. Finally, interpolation errors corresponding to a particular sample design can be used to determine which interpolated values contribute usefully to a band-limited analysis of bathymetric variability.
Published by Elsevier Science B.V.

Keywords: spatial interpolation; field data; spectral analysis; error analysis; optimization

1. Introduction

Most studies of nearshore hydrodynamics, sediment transport, and morphology focus on bathymetric variability within a narrow band of spatial and temporal scales. However, nearshore bathymetry consists of a variety of morphologic patterns that span a broad range of spatial and temporal scales. The largest features include nearshore sandbars, with O(100 m) cross-shore wavelengths and O(100–1000 m) alongshore length scales (King and Williams, 1949; Bowen and Inman, 1971; Lippmann and Holman, 1990; Wijnberg and Terwindt, 1995). Intermediate features include beach cusps at the shoreline, which are typically characterized by alongshore wavelengths of O(1–10) m (Komar, 1971; Dean and Maurmeyer, 1981; Werner and Fink, 1993; Holland, 1998; Coco et al., 2000). The smallest features are ripples, which are characterized by O(0.1–1 m) wavelengths (Clifton et al., 1971; Hunter et al., 1979).

Typically, morphologic features with the largest

* Corresponding author.
E-mail address: nplant@nrlssc.navy.mil (N.G. Plant).

0025-3227/02/$ – see front matter Published by Elsevier Science B.V.
PII: S0025-3227(02)00497-8

spatial scale have the slowest response times, while short-scale features respond rapidly. For example, sandbars may exhibit significant response within an approximately seasonal cycle (Winant et al., 1975; Lippmann and Holman, 1990), and even over decades (Birkemeier, 1985; Lippmann et al., 1993; Ruessink and Kroon, 1994; Plant et al., 1999). Whereas beach cusps evolve over periods of minutes to hours (Antia, 1987; Miller et al., 1989; Holland, 1998; Coco et al., 2000), and ripples can evolve on time scales of seconds, or nearly the same time scale as orbital wave motions (Blondeaux, 1990; Hay and Bowen, 1993; Traykovski et al., 1999).

In spite of the broad range of scales that may be of interest to nearshore research, it is at present impossible for bathymetric measurements to simultaneously resolve all relevant scales. A consequence of this limited resolution is that attempts to adequately sample large-scale features will tend to suffer from aliasing of short-scale features. For example, there is an ongoing program to sample the long-term, large-scale bathymetry at the Army Corps of Engineer's Field Research Facility (FRF; Fig. 1; see Birkemeier and Mason, 1984; Lee et al., 1998; Birkemeier and Holland, 2001). A large amphibious vehicle (the CRAB) is used to survey the bathymetry with roughly 10-m resolution in the cross-shore, along-track direction. Cross-shore transects are spaced 50–200 m in the alongshore direction. This sampling pattern resolves features with cross-shore scales greater than several tens of meters (which is adequate for resolving sand bars), and it resolves alongshore scales greater than about 100 m. Beach cusps, which are commonly observed at this loca-

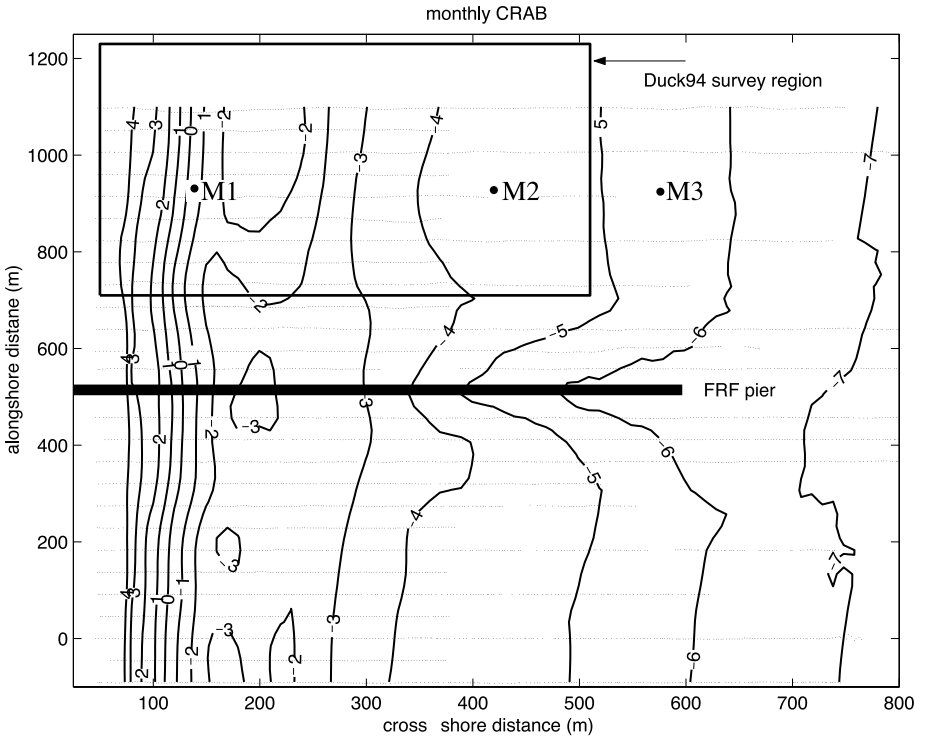

Fig. 1. Example of the monthly CRAB bathymetric survey data from the FRF, Duck, NC (20 September 1990). The survey transects are indicated with lines of densely spaced dots. The contours (solid lines) are drawn from interpolated bathymetry, computed using the quadratic loess method, with cross-shore and alongshore smoothing scales of $\lambda_x = 20$ m and $\lambda_y = 200$ m. Locations M1–M3 refer to analysis points, discussed in Section 3.

tion, have typical wavelengths of about 30 m (Holland, 1998) and are not resolved by the typical survey strategy. The variability associated with cusps aliases as variability at resolved scales and contributes to sampling errors.

In contrast, intensive field experiments have included surveys over small regions with dense sampling in order to resolve short-term temporal variations. For example, during the Duck94 experiment at the FRF (Birkemeier and Thornton, 1994) a 500×500-m horizontal domain (see Fig. 1) was surveyed daily using the CRAB. The reduced size of the sample domain allowed the alongshore transect spacing to be reduced to 25 m (Fig. 2, top). As part of the same experiment, higher resolution surveys were obtained using a global positioning system (GPS) (Plant and Holman, 1997) within the subaerial portion of the CRAB's sample domain (see Fig. 2, top). The high-resolution system consisted of a GPS antenna mounted on one of several small vehicles, including a hand-pushed cart called DOLLY, which had footprints of about 1–2 m. (Below, we refer to the high-resolution surveys as the DOLLY data.) The samples were obtained every 5 m along track (which was oriented alongshore) and track spacing was typically 5 m in the cross-shore direction. Since the different survey approaches resolve different scales, these data sets may be compared to illustrate how sampling errors depend on scale.

For example, a comparison of the CRAB and DOLLY bathymetry interpolated using a nearest neighbor interpolation routine is shown in the lower two panels of Fig. 2. Cusps with a wavelength of 70 m appear in the interpolation of the CRAB bathymetry. However, the actual cusp spacing is revealed by the interpolation of the DOLLY bathymetry to be about 40 m. The apparent 70 m cusp spacing inferred from the CRAB data is an artifact of sparse sampling. This example shows that significant variations in the sampling error result from changes in resolution of the sampling method.

In addition to being used to characterize the dominant scales of nearshore morphology, bathymetric survey data are also used in most process models, which describe waves, currents, and sediment transport (e.g. Roelvink and Broker, 1993).

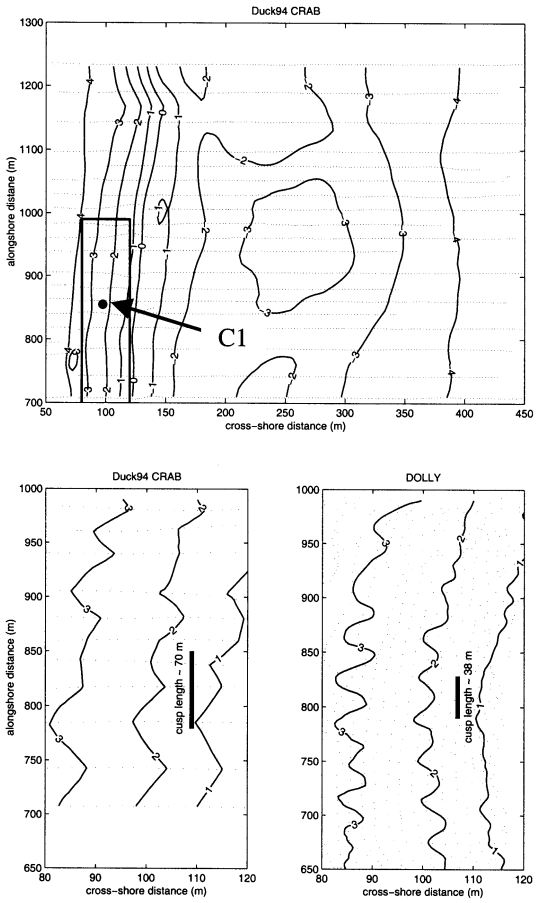

Fig. 2. Comparison of bathymetry surveyed on 21 October 1994. Bathymetry surveyed by CRAB, and interpolated with the quadratic loess method is shown at top, with a box delineating the sub-region surveyed synchronously by DOLLY. Location C1 is an analysis point, discussed in Section 3. The lower two panels compare the CRAB and DOLLY bathymetry, interpolated with a nearest neighbor method. A mean cusp wavelength (heavy line) has been identified in each data set.

Bathymetric data are used to prescribe the bottom boundary condition in the case of hydrodynamic models, and to prescribe initial conditions for seabed evolution (e.g. profile) models. These models typically require bathymetric data to be interpolated to regularly spaced grid nodes. The spacing of the grid nodes controls the range of scales that are resolved by the model. Thus, even results of nearshore hydrodynamic and sediment process models are, in practice, limited to the scales re-

solved by bathymetric data, and are affected by the errors at these scales.

In general, very few studies that utilize bathymetric survey data have bothered to report error estimates associated with sampling deficiencies (such as aliasing), even though these errors could seriously affect quantitative and qualitative interpretations of nearshore processes. The objective of this paper is to describe the magnitude and scale of errors associated with both sampling and utilization of typical nearshore bathymetric data sets. Our approach is based on the practical need to interpolate almost any sampled data set to a regular spatial/temporal grid. In Section 2, we describe a particular scale-controlled interpolation method and its associated error analysis. Interpolation errors can be quantitatively estimated only if the spectral density function of the true bathymetry is known. Section 3 presents a spectral analysis of the interpolation method, applied to the three bathymetric data sets discussed so far. In Section 4, the practical applications and implications of the interpolation error analyses are treated, and the discussion is summarized in Section 5.

2. Theory

2.1. Interpolation method

We consider linear interpolation methods consisting of a set of J observations, z_j, located at positions $\underline{x}_j$, and an interpolation location (or locations), $\underline{x}_i$, where we wish to obtain estimates, $\hat{z}_i$. (In the equations that follow, a subscript 'j' refers to an observation index, and subscript 'i' refers to an interpolation index.) The independent variable is a vector because it can include multi-dimensional location coordinates, such as two spatial coordinates and a time coordinate:

$$\vec{x} = (x, y, t, \ldots) \tag{1}$$

The observation locations may be randomly spaced, while the interpolation locations are usually, but not necessarily, regularly spaced.

Linear interpolation methods seek an elevation estimate of the form:

$$\hat{z}_i = \sum_{j=1}^{J} \hat{a}_{ij} z_j \tag{2}$$

which constructs the interpolated value as a linear combination of the observations. The parameters $\hat{a}_{ij}$ are a set of weights, which usually depend only on the location of the observations relative to the interpolation location. The interpolation weights may be computed using, for example, piecewise linear interpolation of nearest neighbors, inverse distance weighting, triangulation, and finite element methods (for a detailed review, see Franke, 1982), and may even depend on known, spatially-variable sampling errors.

Not all approaches to computing the interpolation weights control the smoothness of the interpolated output. Methods that control the scale (and hence are appropriate to typical band-limited analyses of bathymetry) include sub-optimal interpolation (Ooyama, 1987) and linear smoothing. We will use quadratic loess interpolation, which is a linear smoother (for a review, see Schlax and Chelton, 1992), and has filtering properties controlled by a parameter λ_x (in 1-D). For example, Fig. 3 shows a cross-section through the bathymetry presented in Fig. 1 (monthly CRAB survey). The data have been interpolated to a regularly spaced array using both quadratic loess and nearest neighbor methods. The quadratic loess method used observations within ± 20 m in the cross-shore direction (i.e. $\lambda_x = 20$ m) and ± 200 m alongshore ($\lambda_y = 200$ m) of each interpolation location, while the nearest neighbor method used only the nearest 3 observations to compute each estimate.

Although the large-scale trends are similar, there are significant differences between the interpolated profiles. In particular, seaward of about $x = 400$ m, the data are relatively sparse and the nearest neighbor interpolation is poorly constrained. Also, where the bathymetry is relatively rough (Fig. 3, inset), the nearest neighbor method reproduces short-scale, but not necessarily accurate, variability present in the data. The quadratic loess method removes much of this variability to produce a more reliable, albeit smoother, estimate of the bathymetry. Clearly, the impact of sam-

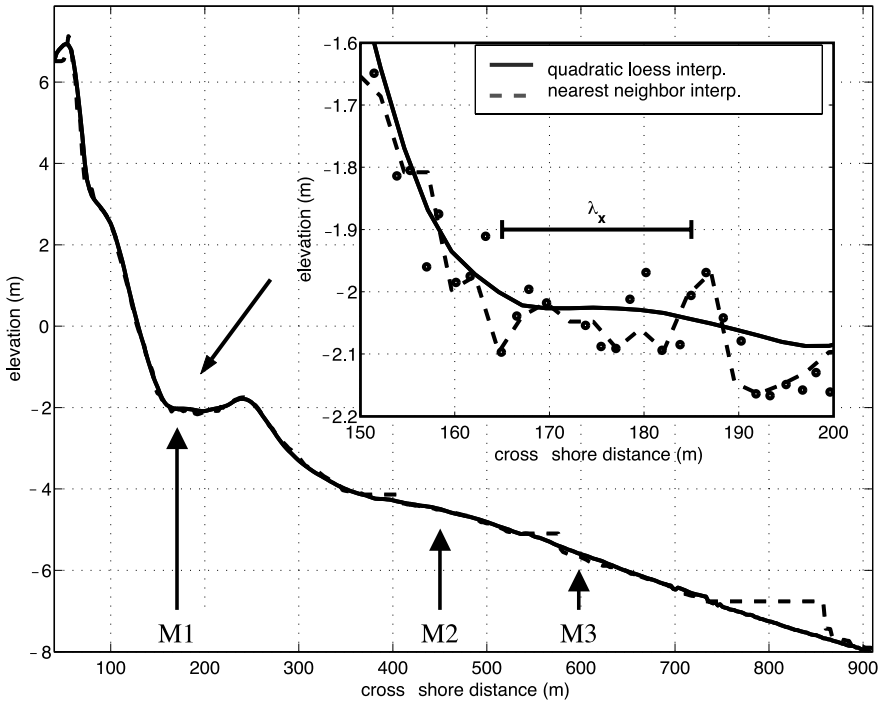

Fig. 3. Cross-section through monthly CRAB bathymetry, comparing quadratic loess (solid line) and nearest neighbor (dashed line) interpolation methods. Inset includes data (circles) lying within 12.5 m of the cross-section (y = 875 m). The cross-shore scale, λ_x, governs the smoothness of the quadratic loess interpolation methods.

pling errors depends on the interpolation method as well as the sampling resolution. The next section describes several methods of quantifying interpolation reliability.

2.2. Mean square interpolation error

Discrete sampling of bathymetric data introduces two kinds of sampling errors. The first is aliasing of short-scale, unresolved features. For instance, the 10-m wheelbase of the CRAB does not resolve features that are shorter than about 20 m. The unresolved features will be misrepresented by the discrete observations. In addition to sample-resolution errors, there are measurement errors, which are typically assumed to be represented as Gaussian-distributed white noise. Thus, we may decompose the observations into components representing a part of the true bathymetry that is resolved, $\tilde{z}_j$, and an unresolved part, z_j', plus measurement error, e_j. Then, the interpolated estimate described by Eq. 2 may be rewritten as

$$\hat{z}_i = \sum_{j=1}^{J} \hat{a}_{ij}\left(\tilde{z}_j + z'_j + e_j\right) = \underbrace{\sum_{j=1}^{J} \hat{a}_{ij}\tilde{z}_j}_{\text{signal}} + \underbrace{\sum_{j=1}^{J}\left(\hat{a}_{ij}z'_j + \hat{a}_{ij}e_j\right)}_{\text{noise}}. \tag{3}$$

Interpolation errors result from each of the terms in Eq. 3. If the convolution of the interpolation weights with the first term in parentheses does not yield the resolved scales of the true bathymetry at the point x_i, then the interpolation procedure damages the resolvable scales (i.e. the 'signal') of bathymetry, introducing interpolation error. For instance, the interpolator might overly smooth the observations, or even amplify them (also called overshoot by Thiebaux and Pedder, 1987). The remaining two terms, the unresolved component of the true bathymetry and the measurement errors, are of no value to an accurate interpolation of the true bathymetry and both are,

effectively, 'noise'. Ideally, convolution with the interpolation weights would remove the noise.

Interpolation errors associated with each of the terms in Eq. 3 may be estimated. For example, if the measurement errors have identical variances and are spatially uncorrelated, then a normalized mean square interpolation error (which describes the fraction of the measurement error variance that passes unchecked through the interpolation method) is

$$\hat{\varepsilon}_i^2 = \sum_j \hat{a}_{ij}^2 \tag{4}$$

We will refer to this normalized mean square error as the 'sampling error' since it can be computed independent of any actual data values, and depends only on the choice of interpolation method and the distribution of the observation locations. For instance, Fig. 4 shows the sampling error for the quadratic loess interpolation method along the same cross-section presented in Fig. 3. Near the shore, where the data were most densely sampled, the root mean square (rms) sampling error was about 0.2 (corresponding to 4% error variance transmission), suggesting that the scales of interest (as indicated by the smoothing scales, λ_x,λ_y) were adequately sampled. Sharp increases in the sampling error correspond to progressive doubling of the alongshore sample spacing at $x \sim 400$ m, 500 m, and 700 m, indicating that the scales of interest may be poorly sampled. The sampling error reaches a value of 1 (no removal of measurement error) at locations $x > 700$ m, where single observations from distant alongshore locations contributed to the interpolation estimate. The sampling error is useful for analyzing the error properties of a particular surveying strategy.

Another measure of error is the weighted mean square residual,

$$\hat{q}_i^2 = \frac{1}{\hat{\varepsilon}_i^2} \sum_j (\hat{z}_i - z_j)^2 \hat{a}_{ij}^2 \tag{5}$$

which describes the spatially varying misfit between the smooth, interpolated surface and the observations. This is the variability that is filtered from the observations. This error statistic may reflect some of the damage done to the resolved scales of the true bathymetry (first term in Eq. 3), in addition to reflecting removal of damaging noise (last two terms in Eq. 3). In the example shown in Fig. 4, the residual errors are high along the front of the dunes ($x < 100$ m) and along the shoreline ($x \sim 110$ m), and generally decrease in the offshore direction. It is not possible to determine whether the spatial variability in the residuals corresponds to spatial variation in the resolved or unresolved scales. If, however, the residuals are due to measurement errors alone, then an estimate of the mean square interpolation error, is (again assuming white noise, Priestley, 1981, p. 368),

$$\hat{s}_i^2 = \frac{\varepsilon_i^2}{1-\varepsilon_i^2} (\hat{q}_i)^2 \tag{6}$$

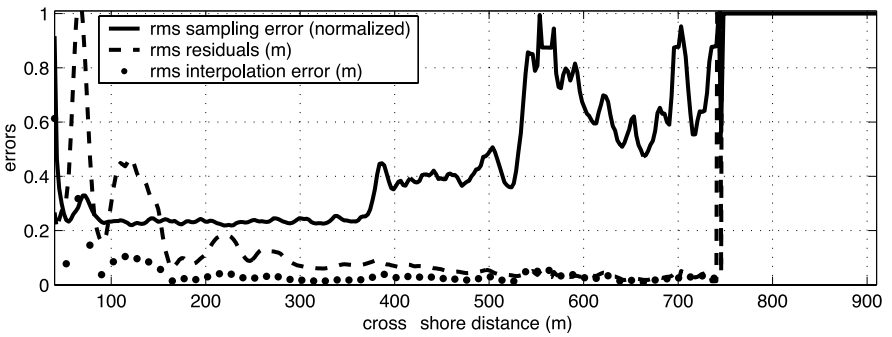

Fig. 4. Interpolation error corresponding to the profile shown in Fig. 3. The rms sampling error describes the fraction of white noise that would be expected to contaminate the interpolation. The rms residuals describe the deviations between the interpolated surface and the observations, and the rms interpolation error describes the estimated deviation between the interpolated surface and the true bathymetry.

This error estimate is the expected variance of the difference between true and interpolated bathymetry values, and takes into account both the sampling pattern and the scale of variations in the observations.

To summarize this section, interpolated bathymetry estimates consist of (1) a component due to the resolvable part of the true bathymetry, plus (2) a component due to the unresolved portion of the true bathymetry, plus (3) a component due to measurement errors. Interpolation errors from any of these three sources may appear as (I) damage done by the interpolation weights to the resolved component of the true bathymetry (e.g. over-smoothing), (II) misrepresentation of true, but unresolved, features (e.g. aliasing, see Fig. 2), and (III) transmission (including possible amplification) of measurement errors. Because of the explicit dependence of interpolation error on scale, quantitative estimates of the magnitude of interpolation errors can be shown through spectral analysis of the interpolation weights, the true bathymetry, and the measurement errors.

2.3. Interpolation errors – spectral representation

Schlax and Chelton (1992) describe the distribution by scale of interpolation errors using the 'equivalent transfer function' (ETF). The ETF is the Fourier transform of the interpolation weights (Fig. 5). A detailed description of the computation of the ETF in practice can be found in Schlax and Chelton (1992). Since the interpolation weights are convolved with the observations, the spectral representation of the interpolated bathymetry is obtained by multiplying the spectral density function of the interpolation weights ($\hat{A}_i = (ETF)_i{}^*(ETF)_i$, where the asterisk denotes complex conjugate) against the spectral density function of each component of the observations:

$$\hat{Z}_i(k) = \underbrace{\hat{A}_i(|k| < k_N)Z(|k| < k_N)_i}_{\text{signal}} + \underbrace{\hat{A}_i(|k| > k_N)Z(|k| > k_N)_i + \hat{A}_i(k)E_i(k)}_{\text{noise}}. \quad (7)$$

The capitalized variables denote the spectral density function of the corresponding lower-case variables in Eq. 3. The term on the left hand side of Eq. 7 is the (1-D) local wave number spectrum

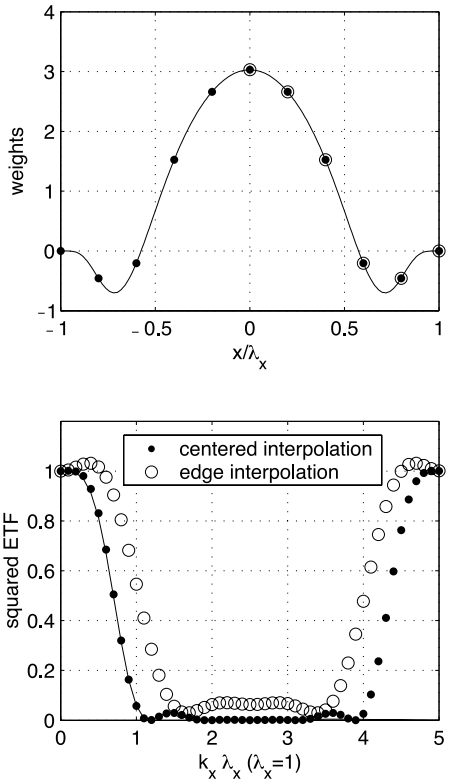

Fig. 5. Distribution of quadratic loess interpolation weights (top) and ETF (bottom). The weighting function (top, solid line) is centered on the interpolation location ($x/\lambda_x = 0$). Observations arranged symmetrically about a centered interpolation location yield a symmetric pattern of discrete interpolation weights (dots), while observations that lie to one side of the interpolation location (circles) yield an asymmetric pattern of weights. The two sampling patterns give rise to different ETF shapes. The Nyquist wave number is $k_N\lambda_x = 2.5$ and the cutoff wave number (for centered sampling) is $k_c\lambda_x = 0.7$.

of the interpolation estimate (k is the wave number, defined as the reciprocal wavelength). The first term on the right hand side of Eq. 7 describes the spectral transmission of the resolved scales, $|k| > k_N$, where k_N is the Nyquist wave number ($k_N = 0.5(\Delta x)^{-1}$ for equally spaced observations in one dimension). The second term describes the transmission of the unresolved part of the true bathymetry spectrum, $|k| > k_N$. The third term describes the transmission of the measurement errors, which may contaminate both the resolved and unresolved wave numbers.

Ideally, the ETF would pass all of the variability below a certain wave number cutoff value, and

it would pass no variability above this cutoff. These properties are not achieved in practice. For the case of quadratic loess interpolation (Fig. 5), the ETF is rounded near a cutoff wave number that can be characterized by the half-power wave number, $k_c \sim 0.7/\lambda_x$. At the lowest wave numbers, $k < k_c$, where the ETF is nearly 1, variability is passed unchanged. Variations in the resolved band are preserved in this wave number range and contribute to an accurate interpolation, while measurement error variance in this band contributes to interpolation error.

Variance is typically attenuated at wave numbers above the cutoff. An exception to this generalization occurs when interpolating at the edge of a sample region, which can lead to amplification of variance at some wave numbers (Fig. 5, open circles), increasing the interpolation error. Attenuation of the unresolved wave numbers and the measurement error is desirable, since this reduces the interpolation error. Attenuation of resolved wave numbers results in an overly smooth representation of the true bathymetry, which adds to the total interpolation error. Bathymetric analyses are impacted adversely only if the scales of interest are attenuated, and this impact may be controlled by choosing appropriate interpolation smoothing scales.

Because of discrete sampling, the ETF contains a series of side bands (only part of the first sideband is shown in Fig. 5). In the case of equally spaced samples, the sidebands are centered on wave numbers equal to $k_{SB} = 2\ n\ k_N$ (where n is any positive integer). Variance in the side bands is unresolved and is aliased into the pass band, resulting in additional interpolation error. However, variance is removed at the unresolved scales lying between side bands, where the ETF is nearly zero (i.e. in the region $2.5 < k_x\lambda_x < 4.5$ in Fig. 5). The important point to note here is that scale-controlled interpolation removes some of the noise variance unlike some other interpolation methods, which may pass all noise variance!

3. Error analysis

Analysis of the scale-dependent distribution of errors can be applied to any linear interpolation method described by Eq. 2 and to arbitrary sampling patterns. In this section, an analysis of the three survey data sets obtained at Duck illustrates the influence of changes in sampling patterns and interpolation methods on interpolation errors. It is quite possible to perform the interpolation and analysis using both spatially and temporally distributed data. For simplicity, the analysis is restricted to the spatial domain and a 2-D (i.e. cross-shore and alongshore coordinates) quadratic loess filter is used to compute interpolation weights. Alongshore smoothing scales are chosen such that the sampling strategy adequately resolved the variability of the interpolation weights. Since 99% of the spectral density in the main band of the quadratic loess ETF is recovered for wave numbers less than $2/\lambda$ (Fig. 5), placing the Nyquist wave number at this limit yields a constraint $\lambda_y \geq 4\Delta y$. A similar constraint for the cross-shore smoothing scale was applied to the DOLLY data set. However, the along-track (cross-shore) sampling of the CRAB is very dense ($\Delta x \sim 1$ m), while the 10-m footprint of the vehicle limits the actual resolution of the data. Accordingly, cross-shore and alongshore smoothing scales (λ_x, λ_y) were (20 m, 200 m) for the monthly CRAB survey, (20 m, 100 m) for the Duck94 CRAB survey, and (20 m, 20 m) for the DOLLY survey.

3.1. Spatial variation of interpolation errors

Due to expected spatial variation in the scale of nearshore morphology, the large-scale, monthly CRAB surveys were conducted with variable spatial resolution, which decreased offshore. This practice minimized the cost of sampling resources, but an additional price may be paid in sampling error. Fig. 6 shows the ETF analysis of the monthly CRAB survey along the cross-section $y = 875$. The band limits of the analysis are set to $(k_x\lambda, k_y\lambda) \leq \pm 6$, which includes most morphologically relevant scales. Near the shoreline, where the alongshore sample spacing is about 50 m (Fig. 1; location M1), the ETF is non-zero within the pass band. As the ETF is essentially zero at all cross-shore wave numbers outside the pass band,

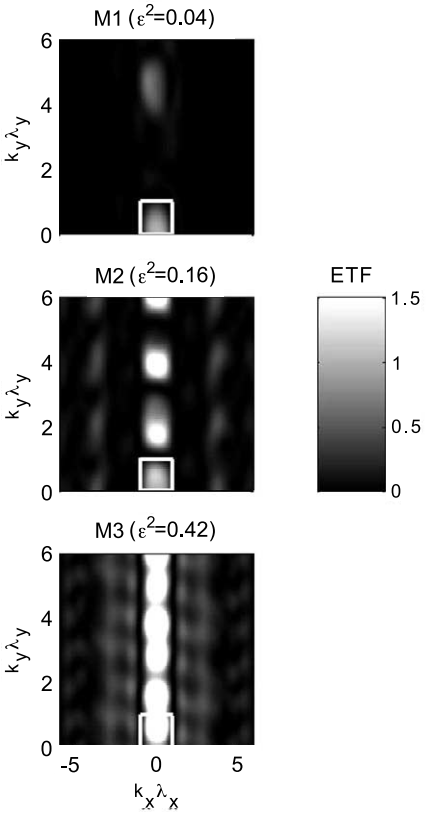

Fig. 6. ETF analysis of the monthly CRAB survey along a cross-section (y = 875 m) centered at x = 175 m (M1), 450 m (M2), and 600 m (M3). Each panel shows the squared ETF value as a function of normalized cross-shore and alongshore wave numbers. Medium gray shades correspond to ETF values of 1, bright white regions correspond to amplification, and black corresponds to attenuation. The white box delineates the approximate bounds of the pass band of the interpolator. In all cases $\lambda_x = 20$ m and $\lambda_y = 200$ m.

the cross-shore sample spacing is sufficient to suppress aliasing out to cross-shore wave numbers at the limit of this analysis. Aliasing is indicated by a side band, located along the zero cross-shore wave number and centered at an alongshore wave number of about $(40 \text{ m/cycle})^{-1}$. This wave number band happens to correspond to beach cusps at Duck, which have wavelengths between 10 and 40 m (Holland, 1998). Because the side bands repeat, variability corresponding to cusps with 20-m wavelengths will be aliased as well. Fortunately, the quadratic loess interpolation suppresses intermediate scale cusps (e.g. 30-m wavelength), which is clearly preferable to passing all unresolved variability.

At the second location ($x \sim 400$ m; Fig. 1, location M2) the alongshore sample spacing increased to 100 m. This results in side bands appearing at alongshore wave numbers of $(100 \text{ m/cycle})^{-1}$ and $(40 \text{ m/cycle})^{-1}$. In addition to an increased number of side bands, there is strong amplification within the side band and pass band. Cusps are not typically present at this sample location, but potentially aliased features include transverse bar with alongshore wavelengths between 25 and 150 m (Konicki and Holman, 2000) and mega cusps with wavelengths between 150 and 400 m (Lippmann and Holman, 1990). Some of this variability will be misrepresented through a combination of amplification and aliasing.

Finally, at the most seaward location (M3), nearly all alongshore wave numbers are capable of aliasing into the pass band and variance at these wave numbers is significantly amplified. This occurs because the effective alongshore sample spacing is infinitely long compared to the interpolation smoothing scales, and the spacing of side bands is nearly infinitely close together in the wave number space. The scales of interest are poorly resolved at this location and the performance of the interpolator is unreliable. As indicated in the profile shown in Fig. 4, the normalized sampling error increases from about 5% at M1 (where data were densely sampled), to 20% at B, and to 50% at M3. The spatial variation of the ETF means that, at different locations, different spatial scales contribute to the interpolation errors differently. Such variability in the interpolation error could ruin an analysis that attempts to describe real variability at a particular spatial scale.

3.2. *Short-scale features: removed or resolved?*

The next analysis focuses on the region near the shoreline where cusps were observed (Fig. 2). In this case, the ETF (Fig. 7) was computed at a fixed location C1 ($x = 100$, $y = 875$, see Fig. 2) for each of the three data sets. In all three examples, the normalized sampling error is less than

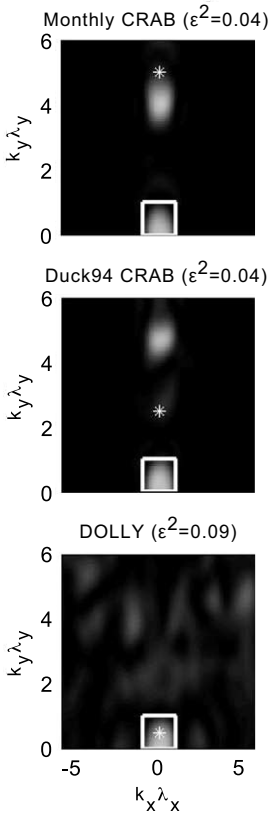

Fig. 7. ETF analysis at point C1 (Fig. 2). The smoothing scales of the quadratic loess interpolation applied to the monthly CRAB survey (top), Duck94 CRAB survey (middle), and DOLLY survey (bottom) were ($\lambda_x = 20$ m, $\lambda_y = 200$ m), ($\lambda_x = 20$ m, $\lambda_y = 100$ m), and ($\lambda_x = 20$ m, $\lambda_y = 20$ m), respectively. The spectral location of the beach cusps with an alongshore wavelength of 40 m is marked with an asterisk in each panel. Shading and pass-band boundaries (white box) are described in Fig. 6.

0.1, indicating that less than 10% error variance transmission is expected if uncorrelated measurement errors are the only noise contribution.

However, unresolved scales associated with beach cusps must also contribute to the noise in two of the data sets. Using the highest resolution DOLLY survey data, we determined that the cusp wavelength was approximately 40 m near the analysis location. This corresponds to normalized alongshore wave numbers presented in Fig. 7 of 5 (Monthly CRAB), 2.5 (Duck94 CRAB), and 0.5 (DOLLY). These cusps are not resolved by the monthly CRAB survey and the cusp wavelength falls near a side band of the corresponding ETF (Fig. 7), indicating the potential for harmful aliasing. In this case, about 30% of the cusp variance aliases and masquerades as resolved variability in the interpolated bathymetry. While the cusps are not resolved by the Duck94 CRAB survey, they correspond to an almost completely attenuated region of the ETF such that only 10% of the cusp variability is transmitted. Indeed, the aliased cusp variations, which appear with an erroneous 70-m wavelength when using nearest neighbor interpolation (Fig. 2, bottom), do not appear with the scale-controlled, quadratic loess interpolation (Fig. 2, top). Lastly, the cusp wavelength falls well within the pass band of the DOLLY survey, and less than 10% attenuation of these features is expected. Interestingly, the somewhat random sampling pattern of the DOLLY survey suppresses well-defined side bands in the ETF, but admits roughly twice the sampling error as the other two interpolation examples. A lack of sidebands reduces the danger of resonance with a particular, unresolved wave number band, which could be particularly advantageous whenever a sampling strategy does not resolve rhythmic features.

4. Discussion

4.1. Interpolation errors in practice

While the ETF analysis provides a very complete means of analyzing the spectral distribution of interpolation errors due to arbitrarily spaced observations, it may not be practical to inspect the ETF at all interpolation locations every time a bathymetric data set is interpolated. Instead, a practical application of the theory described so far would present the expected error in terms of bulk estimates, such as the normalized sampling error, ε, the residual error, $\hat{q}$, and the estimated interpolation error, $\hat{s}$, all of which may be obtained without spectral computations (see Eqs. 4–6). Maps of these bulk error estimates may be used to determine the effectiveness of a particular sampling–interpolation strategy, make inferences about the nature of the true bathymetry, and appropriately utilize the interpolated bathymetry es-

timates in subsequent analyses of nearshore processes.

Fig. 8 shows spatial maps of the three types of interpolation errors, which were computed for the Duck94 CRAB data set. The mean value of ε, averaged over the entire map, was 0.36. The expected measurement error of the CRAB, due to survey errors alone, was 0.1 m (rms) (Birkemeier and Mason, 1984), so the present interpolation scheme is expected to yield bathymetry estimates with about 0.04 m rms error due to this contribution. Errors will be highest where sampling is sparse, such as along the dune line (x ~ 75 m) and far offshore. Depending on research objectives, this error map may be used to re-design the interpolation method (to give smaller errors everywhere) or restrict analysis to only a portion of the interpolated field (where errors are tolerably low).

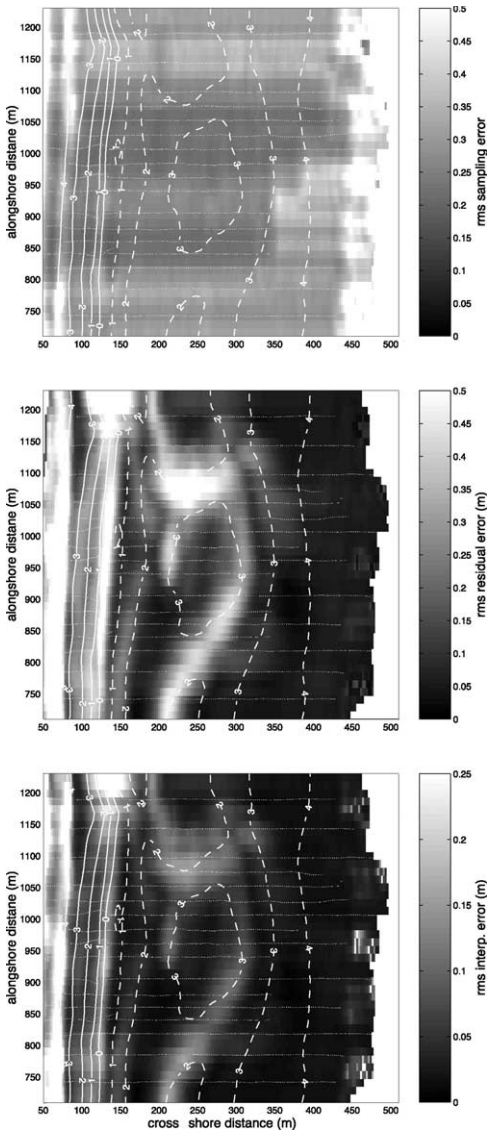

Fig. 8. Rms sampling (top), residual (middle), and, estimated interpolation (bottom) errors corresponding to the Duck94 CRAB survey, with smoothing scales $\lambda_x = 20$ m, $\lambda_y = 100$ m.

The pattern of the sampling error, ε, corresponds quantitatively to actual interpolation errors only if (1) the observation errors were described by band-limited white noise, (2) observation error variance was spatially constant, and (3) the true bathymetric surface contained no features outside the pass band (because these would be removed). In practice, none of these conditions will hold exactly. In particular, error sources may be expected to vary spatially (such as with distance from the shoreline). The most serious source of interpolation error in the Duck94 CRAB example may result from real bathymetric features that were not well resolved by the data (such as beach cusps), and were removed by the interpolation method. This error depends on the choice of interpolation smoothing scale as well as on the sampling pattern.

The residual error map (Fig. 8, middle) quantifies the spatial variability of those features that were removed by the interpolation. Aside from the poorly sampled regions near the dune line and far offshore there are several regions with relatively large rms residuals. (Note: in regions where the interpolation weights, â, are zero for all observations, the mean square residual was set to the total variance of the data set.) One region is a shore-parallel band along the shoreline (x ~ 125 m), where beach cusps were removed by the interpolation method (Fig. 7, middle panel). Assuming that the residual errors were dominated by cusp features, an estimate of the rms cusp amplitude from the residual map is a maximum of 0.6 m. Another region of relatively large residual error lay farther offshore, along the landward face of a sandbar (Fig. 9, top and middle). In this case, the residual error was due to short-scale along-

shore variability. The observed bar crest elevations alternate above and then below the interpolated bathymetry at adjacent cross-shore transects. Without additional information, the true scale of this variability is unresolvable. In addition, the steep landward face of the sandbar was flattened by the combined alongshore and cross-shore smoothing effect of the interpolation. Finally, there was a patch of large residual error, which corresponded to features with cross-shore scales that were well resolved by the observations (Fig. 9, bottom), but were removed by the interpolator due to the choice of smoothing scale ($\lambda_x = 20$ m). However, these features likely had relatively short, and unresolved alongshore scales as well, so there is little justification for trying to preserve them.

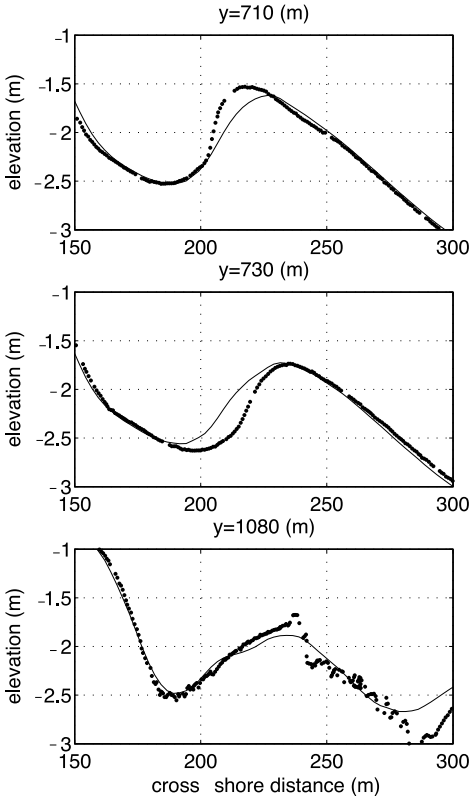

Fig. 9. Cross-shore profiles showing the mismatch between interpolated (solid line) and observed (dots) bathymetry, corresponding to the Duck94 CRAB data set.

Scale-controlled interpolation yields estimates of bathymetric components that fall within the filter's pass band. If we want to describe how well we have estimated these components, the sampling (ε) and residual errors ($\hat{q}$) are combined to estimate the interpolation error ($\hat{s}$), which accounts for spatial variation in the noise level, as well as the sampling pattern's ability to remove this noise. The map of the interpolation error (Fig. 8, bottom) looks very much like that of the residual error. Over much of the interpolated region, the interpolation errors were less than 0.05 m and the median value was 0.04 m, which is the expected reduction of the 0.1 m survey error. Aside from the poorly sampled regions, even the regions with relatively large residual errors show much-reduced interpolation errors.

The purpose of computing interpolation errors is to ensure that the interpolated bathymetry is appropriately utilized in subsequent analyses. For instance, the calibration of remotely sensed bathymetry estimates (e.g. Plant and Holman, 1997; Stockdon and Holman, 2000) relies on the interpolation of directly surveyed bathymetry, such as are obtained with CRAB and DOLLY surveys. Remote sensing and direct survey methods resolve different spatial scales and observations from each method are not generally co-located. Using scale-controlled interpolation, the remotely sensed and directly surveyed data should be interpolated to a common set of co-located coordinates (e.g. a regular grid), preserving a common set of resolved scales. There may be spatial variation in the difference between the two observation methods (related to the distance between survey vehicle and theodolite, for instance) and, due to spatial variations in the sampling patterns, there will be spatial variability in the reliability of interpolated values. Fortunately, the impact of spatial variation in the interpolation error may be removed by using the estimated interpolation errors as weights, such that large errors correspond to small weights. The inter-comparison can be presented in terms of a weighted mean difference or weighted squared difference between the two observation methods (Priestley, 1981, p. 315).

4.2. Rules of thumb

To take advantage of scale-controlled interpolation, the interpolation smoothing scales must be chosen to minimize interpolation errors. Large smoothing scales lead to smaller and more homogeneous sampling error, ensuring a consistent interpolation at the scales of interest. However, increasing the smoothing scale may lead to over-smoothing. To prevent over-smoothing, the cutoff wave number of the interpolator's ETF needs to be higher than the scales of interest. This suggests a constraint $k_c < (1/L_0)$, where L_0 is the shortest length scale of interest. In addition, there is the constraint that the pass band should include at least a portion of the resolved scales, which led to the constraint $k_c < k_N$. This suggests (in one dimension)

$$(4\Delta x) < \lambda_x < (0.5L_0)$$

(quadratic loess interpolation weights) (8a)

If both these limits cannot be satisfied, then the data have been sampled too sparsely to adequately resolve the scales of interest. Similar rules of thumb can be established for other interpolation methods by applying the ETF analysis. For instance:

$$(2\Delta x) < \lambda_x < (0.2L_0)$$

(Hanning interpolation weights) (8b)

$$(1\Delta x) < \lambda_x < (0.1L_0)$$

(Boxcar interpolation weights) (8c)

The ratio of the upper and lower bounds of these rules-of-thumb indicates the minimum number of observations, J, that are needed to resolve a single wavelength of a feature of interest. For the quadratic loess weights, J = 8, J = 10 for the Hanning weights, and J = 10 for the Boxcar. So, the quadratic loess method provides the most efficient use of the observations, although it comes at a cost of increased sensitivity to sparse sample spacing (Schlax and Chelton, 1992). This sensitivity is primarily due to the negative lobes of the interpolation weights (Fig. 5, top), which are not present in the Hanning or Boxcar filters.

4.3. Optimization

The use of scale-controlled interpolation comes at some computational cost compared to simpler methods. For instance, using nearest neighbor interpolation (provided within subroutine 'griddata' as part of the MATLAB programming package), the monthly CRAB data set (with about 10^4 observations, which were interpolated to 10^4 grid points) was interpolated in 2 s, while it took 150 s using the quadratic loess routine, including computing the error fields. (All computations were performed on a personal computer with 1300-MHz Pentium processor.) Since the computational effort of interpolation increases with both the number of input data (which affects the speed of performing the convolution in Eq. 2) and the number of output interpolation points (which affects the number of times Eq. 2 must be implemented), both over-sampling the input and over-resolving the output are inefficient.

4.3.1. Optimal resolution of interpolated output

The spacing of the interpolation output should be as coarse as possible, without aliasing variability at wave numbers within the pass band of the interpolator. This can be accomplished if the Nyquist wave number of the output exceeds the cutoff wave number of the ETF. Thus, the pass band will be resolved if $\Delta x_{out} = \lambda_x/4$, where Δx_{out} is the resolution of the output grid. Interpolated values that are required at a finer resolution may be produced with simple and much faster nearest neighbor interpolation without loss of information or increased error. The error estimates can be similarly interpolated in most cases, although their smoothness is not guaranteed.

4.3.2. Optimal resolution of input data

Over-sampling results when the Nyquist wave number associated with the sample spacing is higher than the wave numbers that contain significant variability. Over-sampled data should be sub-sampled, both for the computational savings

and for accurate estimates of interpolation errors. Sub-sampling can be implemented efficiently with a running average, which removes variance at the over-sampled scales. A running average with a half width of λ_{box} is a form of scale-controlled interpolation and has its own wave number cutoff of $0.25/\lambda_{box}$. The sub-sampled data must be obtained at a spacing $\Delta x = \lambda_x/4$, which implies that $\lambda_{box} = 0.5\Delta x = \lambda_x/8$. Implementation of this sub-sampling scheme on the monthly CRAB data set required an additional 0.5 s of processing time, and reduced the input data (and hence interpolation time) by a factor of two. The performance gains associated with sub-sampling will be larger for larger data sets and larger smoothing scales.

4.3.3. *Accuracy of interpolation errors*

Over-sampling affects the accuracy of interpolation error estimates when short scales that contribute negligibly to the residual error are resolved. As suggested in Section 4.1, residual errors typically encountered in field data probably result from attenuation of partially resolved features. This situation is shown schematically in Fig. 10, in which the spectral representation of the residual error is concentrated near to the pass band of an interpolator's ETF. The estimated interpolation error, $\hat{s}$, is computed with the assumption that the residuals represent a portion of a white noise spectrum, which is spread uniformly across the resolved scales, and that only the portion of that spectrum falling within the pass band (arbitrarily chosen as $k < 1/2$) will contribute to interpolation error. In the schematic example, only one third of the apparent noise variance falls in the pass band. However, one half of the residual error lies within the pass band, so the interpolation error is underestimated.

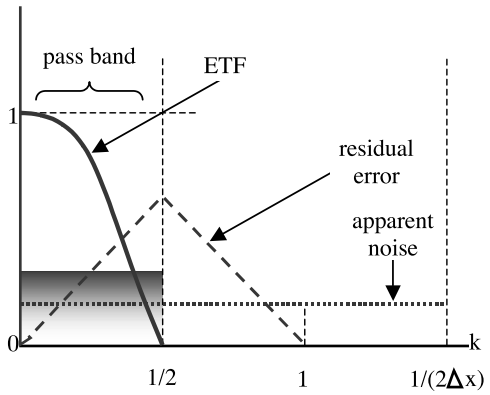

Fig. 10. Schematic spectra of interpolation errors. The residual error spectrum (dashed line) is triangular shaped and its variance is equal to that of the apparent noise spectrum (dotted line). The area of the shaded region represents the true interpolation error variance.

5. Conclusions

Most studies of nearshore hydrodynamics, sediment transport, and morphology focus on bathymetric variability within a narrow band of spatial and temporal scales. These studies rely on bathymetry estimates that are derived from field observations consisting of discrete samples with varying degrees of measurement error. Usually, the bathymetry must be interpolated to produce estimates over a regular grid, and the estimates must accurately represent the scales of interest. However, aliasing of short-scale, unresolved bathymetric features may significantly contaminate variability at larger scales, sometimes leading to serious errors in the representations of the scales of interest. Using a spectral analysis of interpolation errors, we have demonstrated that a scale-controlled interpolation method can minimize the adverse affects of measurement errors and aliasing.

Interpolation errors were analyzed for three different nearshore bathymetric data sets, each of which targeted a different range of spatial scales. All data sets were obtained at Duck, NC, USA. Bathymetric features that were unresolved (beach cusps) or poorly resolved (mega-cusps) introduced serious potential for contamination in two of the data sets. Using a typically applied nearest neighbor interpolation method the unresolved beach cusps were aliased into a wavelength that was about twice the true beach cusp wavelength. While an appropriate scale-controlled interpolation method (a quadratic loess smoother) could not recover the aliased variability, it could filter this variability yielding a smooth estimate of the true bathymetry.

Both nearest neighbor and quadratic loess in-

terpolation methods are forms of linear interpolation, in which the estimates are a linear combination of the observations. All linear interpolation methods allow for straightforward estimation of unavoidable interpolation errors. Additionally an analysis of interpolation errors can be performed independently of actual observations, which allows for optimization of bathymetric sampling strategies by ensuring that dominant scales are either resolved or largely removed. Once computed for a particular sample design, interpolation errors can be used to determine which estimates contribute usefully to a band-limited analysis of bathymetric variability.

The method presented here makes optimal use of the observations. To ensure statistical robustness and optimal computational efficiency, some statistical constraints yield the following guidelines for quadratic loess interpolation.

(1) Choose a smoothing scale (or scales), λ_x, which satisfies sampling constraints: $(4\Delta x) < \lambda_x$. Or, if there are sufficient data, preserve a length scale of interest: $\lambda_x < (L_0/2)$.

(2) Design an output grid that does not over-resolve the pass band: $\Delta x_{out} = \lambda_x/4$.

(3) Use a running average filter to decimate large data sets: $\lambda_{box} = \lambda_x/8$.

(4) Apply scale-controlled interpolation.

(5) Examine error fields and adjust λ_x to suit particular needs:

– increase λ_x if interpolation errors are large (or spatially variable);

– decrease λ_x if residual errors are large and normalized errors are small.

Acknowledgements

This work was supported by Office of Naval Research base funding, PE#0601153N to the Naval Research Laboratory. N.G.P. contributed to this paper while he held a CORE/NRL Postdoctoral Fellowship. He is also thankful for the teachings of – and conversations with – Dudley Chelton, who pointed out the benefits of linear smoothers. We are grateful for the excellent data collected at Duck by the FRF staff and by Rob Holman and his students.

References

Antia, E.E., 1987. Preliminary field observations on beach cusp formation and characteristics on tidally and morphodynamically distinct beaches on the Nigerian coast. Mar. Geol. 78, 23–33.

Birkemeier, W.A., 1985. Time scales of nearshore profile change, Proc. 19th International Conference on Coastal Engineering. ASCE, New York, pp. 1507–1521.

Birkemeier, W.A., Holland, K.T., 2001. The corps of engineers' field research facility: More than two decades of coastal research. Shore Beach 69, 3–12.

Birkemeier, W.A., Mason, C., 1984. The CRAB: A unique nearshore surveying vehicle. J. Surv. Eng. 110, 1–7.

Birkemeier, W.A., Thornton, E.B., 1994. The DUCK94 nearshore field experiment. In: A.S.-Arcilla, M.J.F. Stive, N.C. Kraus (Eds.), Proc. Coastal Dynamics '94. ASCE, Barcelona, pp. 815–821.

Blondeaux, P., 1990. Sand ripples under sea waves, Part 1. Ripple formation. J. Fluid Mech. 218, 1–17.

Bowen, A.J., Inman, D.L., 1971. Edge waves and crescentic bars. J. Geophys. Res. 76, 8662–8671.

Clifton, H.E., Hunter, R.E., Phillips, R.L., 1971. Depositional structures and processes in the non-barred high-energy nearshore. J. Sediment. Petrol. 41, 651–670.

Coco, G., Huntley, D.A., O'Hare, T.J., 2000. Investigation of a self-organization model for beach cusp formation and development. J. Geophys. Res. 105, 21991–22002.

Dean, R.G., Maurmeyer, E.M., 1981. Beach cusps at Point Reyes and Drakes Bay Beaches, California, Proc. 17th Conference on Coastal Engineering. ASCE, Sydney, pp. 863–884.

Franke, R., 1982. Scattered data interpolation: Test of some method. Math. Comput. 38, 181–200.

Hay, A.E., Bowen, A.J., 1993. Spatially correlated depth changes in the nearshore zone during autumn storms. J. Geophys. Res. 98, 12387–12404.

Holland, K.T., 1998. Beach cusp formation and spacing at Duck, USA. Cont. Shelf Res. 18, 1081–1098.

Hunter, R.E., Clifton, H.E., Phillips, R.L., 1979. Depositional processes, sedimentary structures, and predicted vertical sequences in barred nearshore systems, Southern Oregon Coast. J. Sediment. Petrol. 49, 711–726.

King, C.A.M., Williams, W.W., 1949. The formation and movement of sand bars by wave action. Geogr. J. 113, 70–85.

Komar, P.D., 1971. Nearshore cell circulation and the formation of giant cusps. Geol. Soc. Am. Bull. 82, 2643–2650.

Konicki, K., Holman, R.A., 2000. The statistics and kinematics of transverse sand bars on an open coast. Mar. Geol. 169, 69–101.

Lee, G.H., Nicholls, R.J., Birkemeier, W.A., 1998. Storm-driven variability of the beach–nearshore profile at Duck, North Carolina, USA, 1981–1991 (vol 148, 163, 1998). Mar. Geol. 151, 155–156.

Lippmann, T.C., Holman, R.A., 1990. The spatial and temporal variability of sand bar morphology. J. Geophys. Res. 95, 11575–11590.

Lippmann, T.C., Holman, R.A., Hathaway, K.K., 1993. Episodic, non-stationary behavior of a two sand bar system at Duck, NC. USA. J. Coast. Res. SI, 49–75.

Miller, J.R., Miller, S.M.O., Torzynski, C.A., Kochel, R.C., 1989. Beach cusp destruction, formation, and evolution during and subsequent to an extratropical storm, Duck, North Carolina. J. Geol. 97, 749–760.

Ooyama, K.V., 1987. Scale-controlled objective analysis. Mon. Weather Rev. 115, 2479–2506.

Plant, N.G., Holman, R.A., 1997. Intertidal beach profile estimation using video images. Mar. Geol. 140, 1–24.

Plant, N.G., Holman, R.A., Freilich, M.H., Birkemeier, W.A., 1999. A simple model for interannual sandbar behavior. J. Geophys. Res. 104, 15755–15776.

Priestley, M.B., 1981. Spectral Analysis and Time Series. Academic Press, London, 890 pp.

Roelvink, J.A., Broker, I., 1993. Cross-Shore Profile Models. Coast. Eng. 21, 163–191.

Ruessink, B.G., Kroon, A., 1994. The behavior of a multiple bar system in the nearshore zone of Terschelling, the Netherlands, 1965–1993. Mar. Geol. 121, 187–197.

Schlax, M.G., Chelton, D.B., 1992. Frequency domain diagnostics for linear smoothers. J. Am. Stat. Assoc. 87, 1070–1081.

Stockdon, H.F., Holman, R.A., 2000. Estimation of wave phase speed and nearshore bathymetry from video imagery. J. Geophys. Res. 105, 22015–22033.

Thiebaux, H.J., Pedder, M.A., 1987. Spatial Objective Analysis: With Applications to Atmospheric Science. Academic Press, Orlando, FL, 299 pp.

Traykovski, P., Hay, A.E., Irish, J.D., Lynch, J.F., 1999. Geometry, migration, and evolution of wave orbital ripples at LEO-15. J. Geophys. Res. 104, 1505–1524.

Werner, B.T., Fink, T.M., 1993. Beach cusps as self-organized patterns. Science 260, 968–971.

Wijnberg, K.M., Terwindt, J.H.J., 1995. Extracting decadal morphological behavior from high-resolution, long-term bathymetric surveys along the Holland coast using eigenfunction analysis. Mar. Geol. 126, 301–330.

Winant, C.D., Inman, D.L., Nordstrom, C.E., 1975. Description of seasonal beach changes using empirical eigenfunctions. J. Geophys. Res. 80, 1979–1986.