Journal of Sound and Vibration (2002) **256**(5), 980–988
doi:10.1006/jsvi.2001.4227, available online at http://www.idealibrary.com on IDEAL®

ON THE PRESENCE OF END EFFECTS AND THEIR MELIORATION IN WAVELET-BASED ANALYSIS

T. KIJEWSKI AND A. KAREEM

NatHaz Modeling Laboratory, Department of Civil Engineering and Geological Sciences, University of Notre Dame, 156 *Fitzpatrick Hall, Notre Dame, IN* 46556, *U.S.A. E-mail: tkijewsk@nd.edu*

(*Received* 29 *November* 2001)

1. INTRODUCTION

While the Fourier transform has reshaped the manner in which engineers interpret signals, it becomes evident that by utilizing a series of infinite basis functions, time-varying features cannot be captured. The realization that non-stationary features often characterize processes of interest led to the definition of alternative transforms that rely on bases of finite length, one of the most popular of which is the wavelet transform.

1.1. WAVELET THEORY

The wavelet is a linear transform that decomposes an arbitrary signal $x(t)$ via basis functions that are simply dilations and translations of the parent wavelet $g(t)$ through a convolution operation,

$$W(a,t) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} x(\tau) g^* \left(\frac{t-\tau}{a} \right) \mathrm{d}\tau. \tag{1}$$

Dilation by the scale a, inversely proportional to frequency, represents the periodic nature of the signal. By this approach, time–frequency localization is possible, since the parent wavelet serves as a window function. Just as in Fourier analysis, an indicator of the signal's time-varying energy content over a range of frequencies can be generated by plotting the squared modulus of the wavelet transform as a function of time and frequency to generate a *scalogram* [1].

1.2. MORLET WAVELET

As it is quite natural to view information in terms of harmonics instead of scales, the Morlet wavelet [2] has become a popular choice for analysis, as given by

$$g(t) = \mathrm{e}^{\mathrm{i}2\pi f_o t} \mathrm{e}^{-t^2/2} = \mathrm{e}^{-t^2/2} (\cos(2\pi f_o t) + \mathrm{i} \sin(2\pi f_o t)). \tag{2}$$

The dilations of this temporally localized parent wavelet then allow the effective frequency of this sine–cosine pair, oscillating at central frequency f_o, to change in order to match

0022-460X/02/$35.00

harmonic components within the signal. As a result of obvious analogs, the wavelet scale is uniquely related to f, the Fourier frequency: $a = f_o/f$.

2. END EFFECTS THEORY

The presence of end effects in wavelet-transformed data has been noted in a number of applications, e.g., reference [3]. In many cases, the *a priori* knowledge of the signal characteristics allows anomalies to be qualitatively distinguished and neglected in subsequent analyses. However, this is, in general, not possible as it requires a quantitative guideline to establish as to what portions of the wavelet-transformed signal are accurate.

By examining the convolution operation in equation (1) in light of the parent wavelet in equation (2), it is evident that, although the wavelet is focused at a given time and represents the signal content in that vicinity, the window extends equally into the past and future. Though the span of this analysis window is dependent on both the parent wavelet and scale being analyzed, near the ends of the signal, the wavelet's analysis window may extend significantly beyond the length of data. Thus, the resulting wavelet coefficients in these end-effects regions are based on incomplete information and have questionable accuracy.

2.1. END EFFECTS CRITERION FOR MORLET WAVELET

For the Morlet wavelet, the use of a Gaussian window on the Fourier basis functions makes the precise definition of temporal duration impractical. Instead, Gabor [4] proposed a mean square definition to establish effective durations in time and frequency. Using this approach, an effective temporal duration Δt_i for a scaled Morlet wavelet at frequency f_i can be defined as the product of that scale and the duration of the Morlet's Gaussian window [5], given by

$$\Delta t_i = \frac{1}{\sqrt{2}} \frac{f_o}{f_i}. \tag{3}$$

By this definition, the Morlet wavelet is assumed to effectively span $2\Delta t_i$ in the time domain, or one standard deviation of the Guassian window. As shown in Figure 1, there is a considerable portion of the window beyond one standard deviation of the mean value. A stricter interpretation would define the effective temporal duration of this wavelet as several standard deviations of the Gaussian window. Dependent on the desired level of

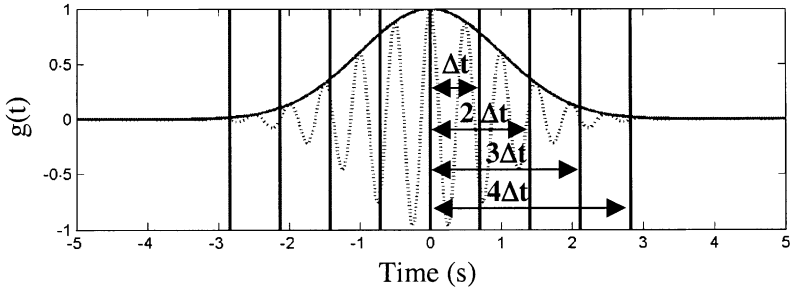

Figure 1. Real component of Morlet wavelet enveloped by the Gaussian window. Various temporal duration measures are marked by vertical bars.

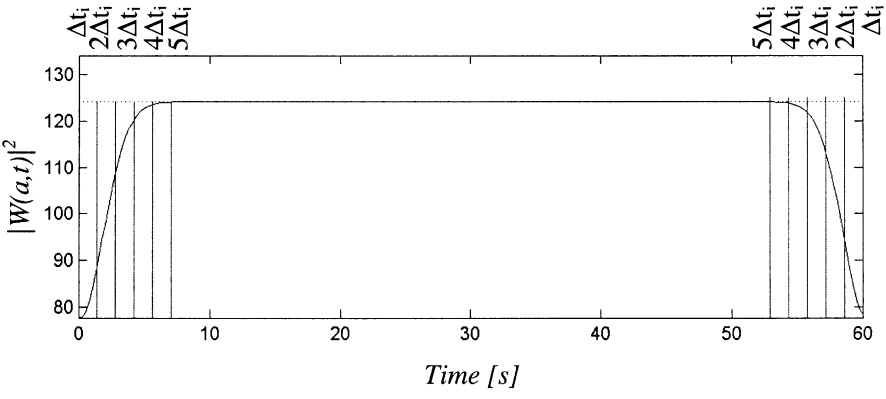

Figure 2. Scalogram of sine wave at scale associated with 1 Hz: ····· denotes theory and —— is the calculated result. Vertical bars demarcate end-effects regions, $\beta \Delta t_i$, for $\beta = 1$–5.

accuracy, an integer multiple β of the measure in equation (3) can be imposed to quantify the usable region within a set of wavelet-transformed data of length T, according to

$$\beta \Delta t_i \leqslant t_j \leqslant T - \beta \Delta t_i. \tag{4}$$

2.2. EXAMPLE 1: INFLUENCE OF END EFFECTS ON SPECTRAL AMPLITUDE

For a simple illustration of the implications of end effects on spectral amplitude, consider a sine wave with frequency f_n (taken as 1 Hz). In theory, the Morlet wavelet transform of this signal yields a scalogram that is constant with time. At each time ordinate, this yields an instantaneous power spectrum, according to

$$|W(a, t)|^2 = 2\pi^3 a e^{-4\pi^2(af_n - f_o)^2}. \tag{5}$$

In contrast, the Morlet wavelet transform ($f_o = 2$ Hz) of this signal yields a time-varying scalogram, as evidenced by plotting the skeleton or wavelet maxima at each time. This result, shown in Figure 2, displays a rounding of what should be a constant scalogram coefficient. The degree of deviation from the theoretical result, shown as the dotted line, becomes less marked in the interior of the signal. The vertical bars denote the end-effects regions defined in equation (4) for various values of β and indicate that the calculated wavelet transform will more accurately approach the theoretical result for $\beta > 2$.

Further, Figure 3 illustrates the ramifications of analyzing instantaneous power spectra taken from end-effects regions. The calculated instantaneous spectra at each time are plotted one atop the other, essentially collapsing the scalogram in time. The theoretical result in equation (5) is also plotted atop these data for reference. According to equation (5), there should be no variation among them; however, by including the spectra from end-effects regions ($\beta = 0$), there is considerable variance in the plot. Note that the deviations are more marked on the high-frequency side of the spectrum, a result of the lessened frequency resolution at lower scales. Through a more stringent condition, increasing β in equation (4), the neglected regions lengthen and the variance among the spectra is reduced. Note that even the commonly used definition of wavelet temporal duration ($\beta = 1$) is an insufficient measure of the end-effects region producing these deviant spectra. Unfortunately, the use of larger β values reduces the amount of usable transformed

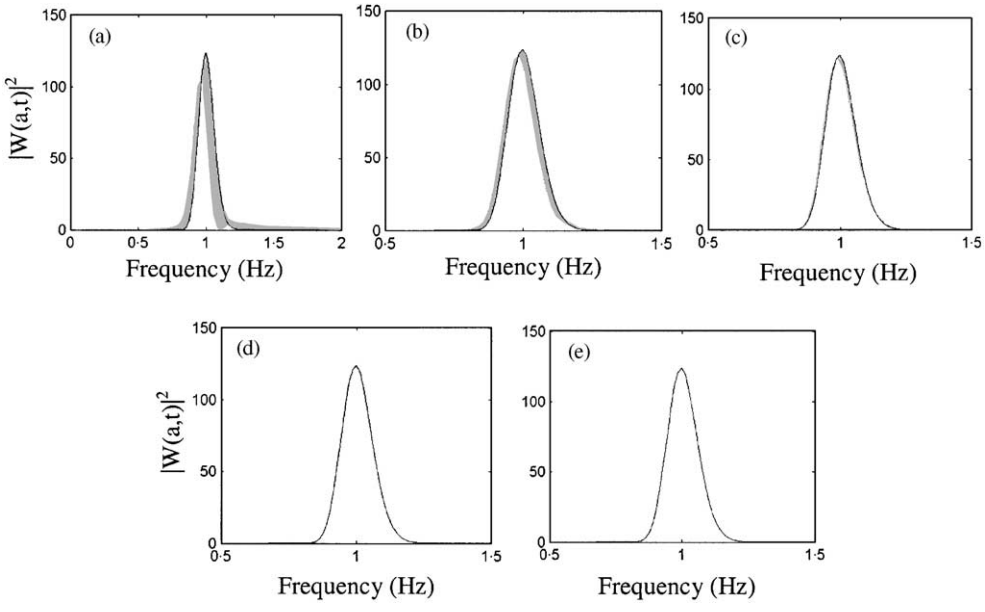

Figure 3. Deviations of simulated instantaneous spectra (gray) from theoretical result (black) as end-effects regions are progressively neglected: (a) $\beta = 0$, (b) 1, (c) 2, (d) 3, (e) 4.

data. Thus, while $\beta = 4$ produces a sufficiently accurate means to quantify end-effects regions and separate deviant spectra, $\beta = 3$ was shown to be sufficient for most analyses in terms of capturing accurately the spectral amplitude [6].

2.3. EXAMPLE 2: INFLUENCE OF END EFFECTS ON BANDWIDTH ESTIMATION

As a consequence of the windowing applied by the Gaussian function in the Morlet wavelet, the bandwidths of the resulting wavelet instantaneous spectra are larger than their Fourier equivalent. This can be shown by considering the Morlet wavelet expression in the Fourier domain, given by

$$\hat{G}(f) = \sqrt{2\pi}\mathrm{e}^{-2\pi^2(f-f_o)^2}. \tag{6}$$

The half-power bandwidth (HPBW) is then used to provide a simple measure of the bandwidth contributed by the Morlet wavelet. Since the resolutions of the wavelet transform are merely scaled versions of the parent wavelet, the half-power bandwidth of the Morlet wavelet in equation (6) can be similarly scaled to determine the wavelet bandwidth contributions to a simple sine wave, evaluated at the ridge or instantaneous frequency of the system. This operation yields

$$HPBW_{WT} = \frac{\sqrt{\ln(2)}}{2\pi}\frac{f_n}{f_o}. \tag{7}$$

As discussed in the previous example, deviations in terms of amplitude of the instantaneous spectra were sufficiently mitigated by neglecting those spectra which were generated in the

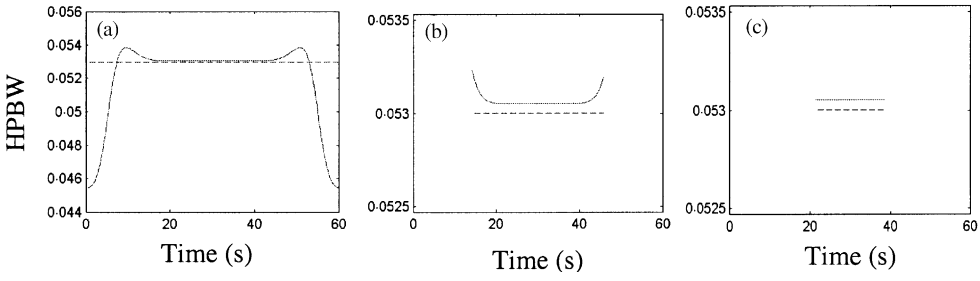

Figure 4. Improvements made in half-power bandwidth estimates by successively neglecting larger end-effects regions. Theoretical prediction (– – –) and calculated result (·····): (a) $\beta = 0$, (b) 4, (c) 6.

end-effects regions, defined by assuming $\beta = 4$. However, the implication of end effects on more sensitive spectral measures such as bandwidth is not completely remedied by neglecting this region, as shown in Figure 4. The calculated half-power bandwidth deviates significantly at the ends of the signal from the theoretical result denoted by the dashed lines. Using the criteria of $\beta = 4$ to neglect regions defined by equation (4) improves the result, though the deviations from theory are still quite evident in the tails. By selecting more stringent conditions on β, the deviations from theory are minimized and the bandwidth estimated in the simulation takes on a constant value. For $\beta = 6$, the deviation between theory ($HPBW_{WT} = 0{\cdot}0530$) and simulation ($HPBW_{WT} = 0{\cdot}0531$) is a mere 0·1887% and arises from a number of approximations made in determining equation (7). Though the deviations in Figure 4 are easily explained by the end-effects phenomenon, simply neglecting these regions in analysis yields to a considerable loss of data, especially in the case of bandwidth estimation, where for $\beta = 6$, only one-third of the transformed signal is deemed reliable.

3. END-EFFECTS MELIORATION: SIGNAL PADDING

The loss of considerable regions of a signal is the unfortunate consequence of end effects. Particularly for more sensitive measures like bandwidth, the loss of usable transformed signal can be quite significant: approximately 10 times the effective duration of the lowest frequency component of interest. One possible solution to this problem would be to pad the beginning and end of the signal with surrogate values. This elongation places the true signal of interest at the center of the transformed vector and leaves the virtual values at the tails to be corrupted by the end-effects phenomenon. It should be reiterated that the wavelet considers both past and present information at each time step. Though the greatest contribution to a wavelet coefficient at that point in time comes from the signal immediately surrounding that point, data displaced further in time are also considered to an extent. Therefore, the regions should locally preserve the frequency and bandwidth characteristics of the signal. This local preservation can be achieved by merely reflecting the signal's negative about its beginning and end.

Figure 5 illustrates an arbitrary signal and the shaded regions are those potentially corrupted following the wavelet transform. Depending on the level of β and f_o selected, these regions can consume two-thirds or more of the signal. In the padding operation, the signal is elongated by $2\beta\Delta t$ as the signal is negatively reflected about the start and end of the signal. Now, the two shaded regions envelop the virtual reflections of the signal, while

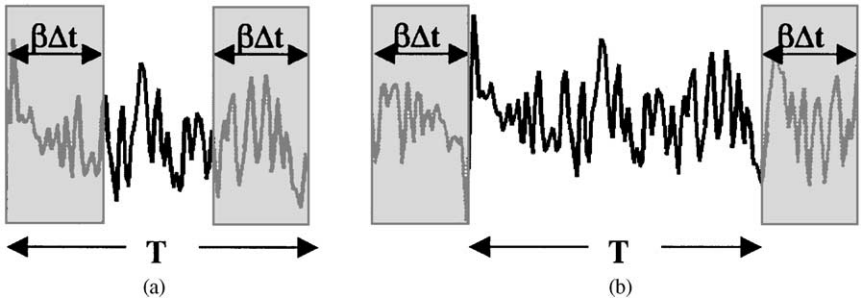

Figure 5. Signal padding concept: (a) original signal, (b) padded signal.

the entire duration of the true signal is conserved and can be analyzed with little contamination from end effects.

To do so, the temporal duration of the analyzing wavelet is then determined using equation (4) for all frequencies being analyzed. As the lowest frequency being considered in the analysis (f_1) will yield the largest duration Δt_1, it dictates the maximum end effects anticipated. β is then selected based on the desired accuracy of the resulting spectra, and the time ordinates of the sampled time vector $t = [t_1 \ldots t_N]$ closest to the termination of the end-effects regions are then identified by

$$t_n = \min [t > \beta \Delta t_1] \quad \text{and} \quad t_m = \max [t < (t_N - \beta \Delta t_1)]. \tag{8}$$

The modified signal x_{MOD} is constructed by reflecting the negative of the signal for the duration of $\beta \Delta t_1$ about t_1 and t_N, according to

$$x_{MOD} = [-x_n \ -x_{n-1} \ \cdots \ x_1 \ \cdots \ x_N \ -x_{N-1} \ \cdots \ -x_m], \tag{9}$$

where x_n and x_m are the values of the sampled data at t_n and t_m. Following the wavelet transform of x_{MOD}, the coefficients calculated from the padded regions are simply neglected, and only the coefficients of the true signal are retained.

3.1. PADDING EXAMPLE

As sine waves are represented by wavelets in a very simplistic form, they are now used to illustrate the efficacy of the proposed padding scheme. Recall that the wavelet instantaneous spectrum of a simple sine wave does not vary in time. Deviations from that constant were shown to be the hallmark of end effects in the wavelet transform. Thus, any signal composed of a series of M sine waves should yield a scalogram containing M constant ridges in the time–frequency domain. The values of the scalogram along each ridge can be individually examined for deviations from the theoretical result. Fortunately, though sine waves are simply analyzed in the wavelet domain by virtue of its inherent bandpass filtering, this same summation of sines is capable of generating complicated time series, which will be subjected to the proposed padding operation to examine the validity of this remedy.

A summation of sine waves at frequencies of 0·28, 0·5, 0·7, 1·0, 1·4, 1·65, 1·9, 2·25, 2·7 and 3·25 Hz was simulated for 10 min, sampled at 10 Hz. To enhance the frequency resolution and separate closer spaced higher frequency content, a central frequency of 10 Hz is chosen for the analysis. Following the calculation of the wavelet transform, 10 ridges are extracted

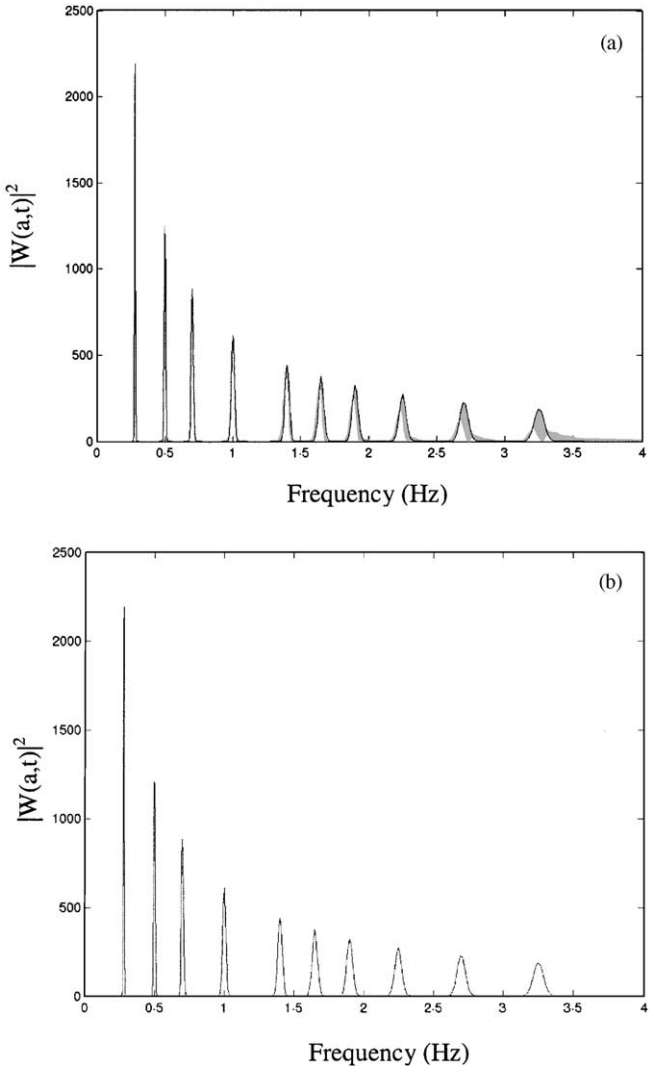

Figure 6. Superposition of instantaneous spectra over all time calculated using wavelet and theoretical prediction: (a) without padding operation, (b) with padding operation.

from the resulting scalogram. Though omitted for brevity, these plots display a characteristic rounding as previously observed in Figure 2. It is evident that the end-effects regions, even for such a large central frequency, decrease significantly with increasing frequency, as indicated by equation (4). This may be the reason that previous studies did not encounter significant manifestations of end effects, as most wavelet analyses have been concerned with higher frequency mechanical systems and not low-frequency oscillations common to many civil engineering structures. The overlaying of the theoretical result and the calculated instantaneous spectra at each time demonstrates the deviations that occur in these end regions, as shown in Figure 6(a). First note that the spectra again have a variable bandwidth that decreases with frequency, as discussed previously. Also note that the deviations are more marked for the high-frequency components due to the lack of frequency resolution. Although the end-effects regions for the higher frequency modes are

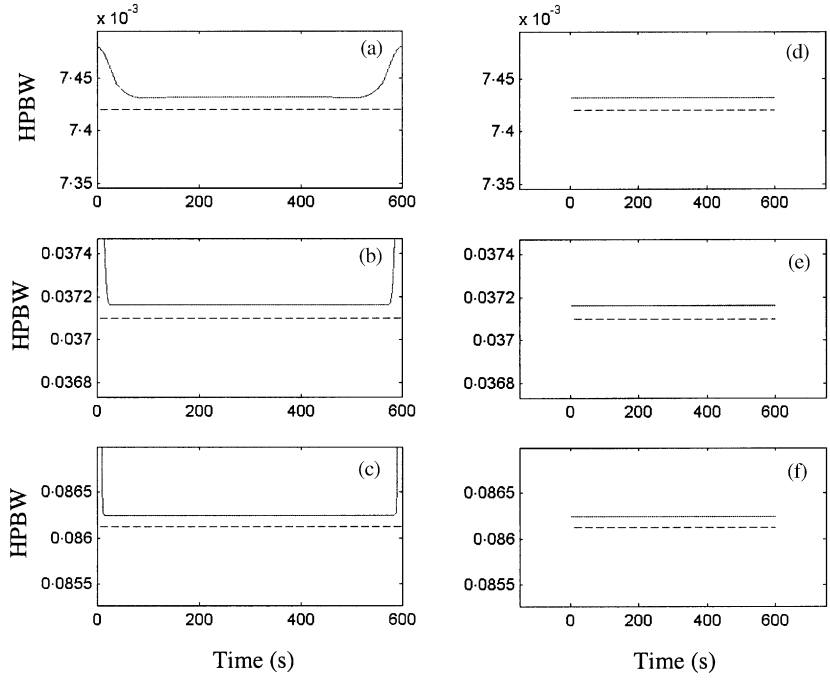

Figure 7. Efficacy of padding operation to reduce end effects in wavelet bandwidth measures, theoretical prediction (– – –) and simulation (——): (a) first mode half-power bandwidth without padding, (b) fifth mode half-power bandwidth without padding, (c) tenth mode half-power bandwidth without padding, (d) first mode half-power bandwidth with padding and $\beta = 6$, (e) fifth mode half-power bandwidth with padding and $\beta = 6$ (f) tenth mode half-power bandwidth with padding and $\beta = 6$.

not as lengthy, the deviations of the few spectra taken from these regions are considerable, especially in the case of the 10th component. Note that the quality of Figure 6(a) could have been enhanced by simply neglecting the spectra that were derived from end-effects regions, as shown in Figure 3, however that results in a loss of a significant amount of data.

The success of the proposed padding operation is gauged in Figure 6(b). In this case, an overlay of theoretical and simulated wavelet spectra at each time is considered for the modified signal in equation (9). Only spectra obtained from the true signal are plotted and all those determined from the virtual values of the padded signal are discarded. Note that there is no discernable difference between the predicted and calculated results when assuming $\beta = 4$. In terms of bandwidth estimates, the half-power bandwidth's accuracy is enhanced in these end regions when the padding scheme is employed. For brevity, a demonstration is provided using only three of the modes. Figure 7(a)–(c) displays the unpadded bandwidth measures, demonstrating the characteristic trademark of end effects. Note again that the portions of the signal lost due to end effects is more marked at lower frequencies, where nearly one-third of the values have been compromised. As demonstrated in Figure 4, $\beta = 6$ is a necessary condition to remove all traces of end effects in the bandwidth measure. Using this condition in conjunction with the padding operation, a precise definition of the half-power bandwidth is maintained over the entire duration of the signal, as shown in Figure 7(d)–(f). The stimulated bandwidth measure is $< 0{\cdot}2\%$ of the theoretical prediction for all of the modes in this example.

Although padding the signal with itself insures that the spectral content of the surrogate regions locally matches that of the true signal, this should not be viewed as a way to defeat the Heisenberg uncertainty principle. It should be reiterated that the end effects are merely

a physical manifestation of the wavelet's inherent analysis windows, which lengthen as f_o is increased. Although the end effects can be repaired, the larger temporal analysis windows imply that changes in the system that occur in shorter time intervals than this window may be completely obscured. Thus, the central frequency should be kept to the smallest value possible to provide the required frequency resolution without compromising the ability of the wavelet to detect non-linear and non-stationary phenomenon.

4. CONCLUSIONS

Although the presence of end effects had been previously noted, these regions were customarily neglected in an *ad hoc* manner. However, as shown in this study, these effect can compromise the accuracy of wavelet scalograms and have even more marked effects on bandwidth measures. In light of the mean square definition of the duration of the Morlet wavelet, the span of end-effects regions was quantified through a flexible criterion that balances the desired quality of the resulting scalogram with the amount of data lost. Recognizing the significant losses possible for low-frequency systems, a simple padding scheme was proposed to extend the length of the signal at both ends. The extended region, being a reflection of the actual signal in that local, preserves the locale spectral content, permitting the end effects to consume the surrogate values while leaving the actual signal unscathed. The wavelet coefficients obtained from these augmentations can then simply be neglected in analysis and the true signal is maximally analyzed by the wavelet. It should be stressed that despite the ability to repair these end effects, the central frequency should still be kept to the smallest value possible to minimize temporal analysis windows. As a result of these larger windows, changes in the system that occur over shorter time intervals may be completely obscured, compromising the wavelet's ability to track non-linear and non-stationary characteristics.

ACKNOWLEDGMENTS

The authors would like to thank NSF Grant CMS 00-85109, the Center for Applied Mathematics at the University of Notre Dame, and the NASA Indiana Space grant for their support.

REFERENCES

1. K. Gurley and A. Kareem 1999 *Engineering Structures* **21**, 149–167. Applications of wavelet transforms in earthquake, wind and ocean engineering.
2. A. Grossman and J. Morlet 1985 in *Mathematics and Physics, Lecture on Recent Results* (L. Streit, editor), 131–157. Singapore: World Scientific. Decompositions of functions into wavelets of constant shape and related transforms.
3. W. J. Staszewski 1998 *Journal of Sound and Vibration* **214**, 639–658. Identification of non-linear systems using multi-scale ridges and skeletons of the wavelet transform. doi: 10.1006/jsvi.1998.1616
4. D. Gabor 1946 *Proceedings of IEEE* **93**, 429–457. Theory of communication.
5. C. K. Chui 1992 *Wavelet Analysis and Applications*: *An Introduction to Wavelets*. San Diego: Academic Press.
6. T. Kijewski and A. Kareem 2002 *Computer-Aided Civil and Infrastructure Engineering* Wavelet transforms for system identification: considerations for civil engineering applications (in press).