Monday, 30 August 2004

# Denoising Long-Wave Records

**Derek Goring, Mulgor Consulting Ltd.**

## *Background*

Long wave records are derived from sea-level records by removing the tide, then high-pass filtering. The resulting record contains not only the long waves, but also noise resulting from instrument error and from aliased short-period waves. Separating the noise from the long waves is the subject of this note.

## *Method*

A common assumption used in denoising is that all of the noise is contained in the first wavelet detail, but that the first detail also contains some valid data. The problem is to separate the valid data from the noise.

Katul & Vidakovic (1998) considered this problem for atmospheric turbulence and described two thresholding methods: universal threshold and Lorenz thresholding.

### 1. Universal threshold

The universal threshold was derived by Donoho & Johnstome (1994) who showed that for *n* independent, identically distributed, standard normal variables, the expected maximum is $\sqrt{2\log_e n}$ . This led to the universal threshold:

$$\lambda_u = \sqrt{2\log_e n}\hat{\sigma} \quad (1)$$

where $\hat{\sigma}$ is an estimate of the population's standard deviation (in practice, $\hat{\sigma}$ is calculated as the mean of the absolute difference, or MAD, which is more robust than the standard deviation of the sample). The universal threshold method assumes that all wavelet coefficients less than $\lambda_u$ are noise, and these are eliminated. An inherent assumption in this method is that the noise is Gaussian-distributed. We have no evidence that this is true for long-wave records.

### 2. Lorenz Thresholding

Lorenz thresholding is based on the Lorenz curve (shown in Figure 1) that relates the cumulative energy that is contained in the *p* smallest wavelet coefficients.

The principle of the method is to reject all coefficients up to the point where rejecting any more would increase the energy-loss more than the gain in the number of coefficients rejected (also called parsimony). This corresponds to the point where the slope of the tangent to the Lorenz curve matches the diagonal, as shown in Figure 1. Also shown in Figure 1, is the universal threshold. Clearly, applying the universal threshold results in a greater loss of energy than for the Lorenz threshold.

The advantage of the Lorenz threshold over the universal threshold is that it requires no assumption about the distribution of the noise.

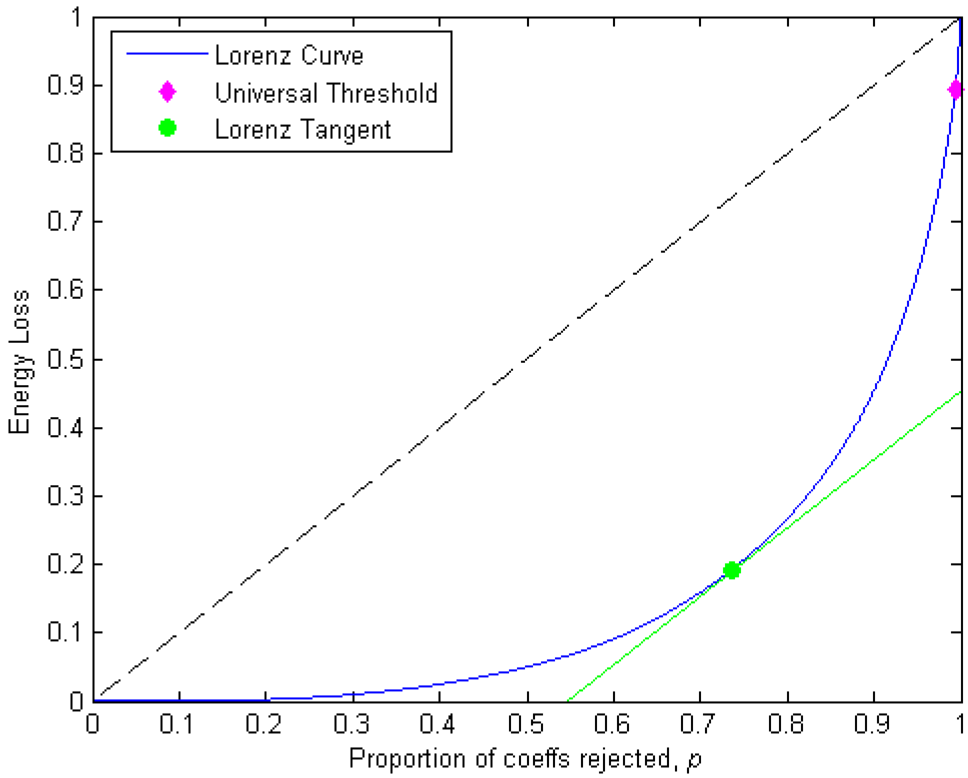


*Figure 1. Lorenz curve of energy loss from rejecting the $p^{th}$ smallest wavelet coefficients.*

## Results

In this section, we present the results from implementation of the thresholding methods described above to several different long-wave records. The amount of noise apparent in signals from different locations differs quite widely, depending upon the site and the instrument.

The Lorenz curves for several sites are presented in Figure 2. The curves fall in three categories as follows:
Timaru and Gisborne have a large difference between energy loss using the Lorenz threshold and energy loss using the universal threshold.
Westgate has a relatively small difference between energy loss using the Lorenz threshold and energy loss using the universal threshold.
Queens Wharf and Marsden Point lie between these extremes.

Examples of the effect of these thresholds on significant wave heights and periods are presented in Figures 3 to 5. In these figures, the significant wave heights and periods have been calculated in 6-hour windows with 5-hour overlaps, resulting in hourly estimates.

Figure 3 shows that for Gisborne (and Timaru), using the universal threshold reduces the significant wave height substantially, whereas using the Lorenz threshold has little effect. Figure 4 shows that the effect is much smaller for Marsden Point (and Queens Wharf), and Figure 5 shows that for Westgate, significant wave height is not affected by thresholding. An explanation for this is that significant wave height, being the average of the highest third of the waves, is not overly sensitive to the waves at the low end of the distribution.

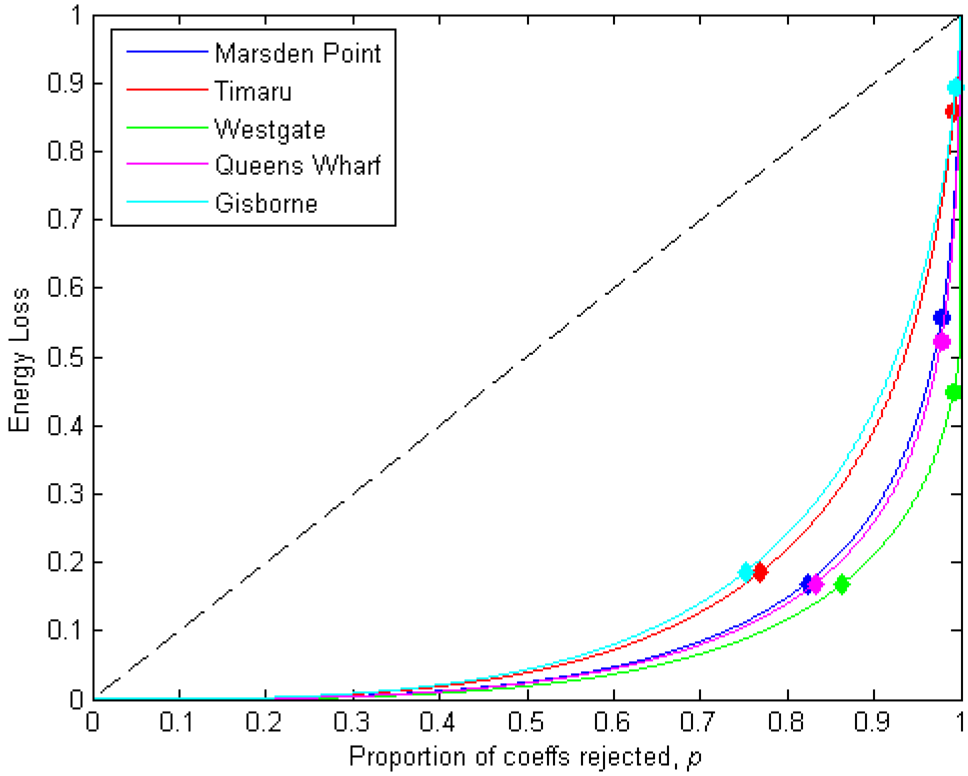


*Figure 2. Lorenz curves from several sites, with the Lorenz threshold marked by a diamond and the universal threshold marked by a circle.*

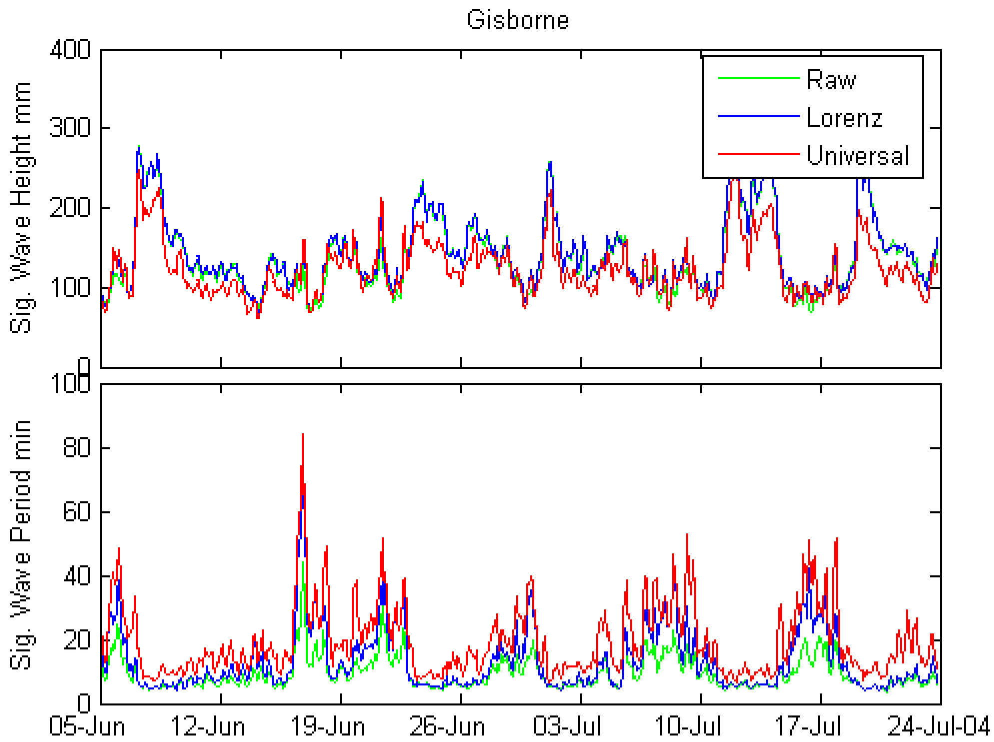


*Figure 3. Comparison of the effects of denoising with different thresholds on significant wave height (upper) and period (lower) for Gisborne.*

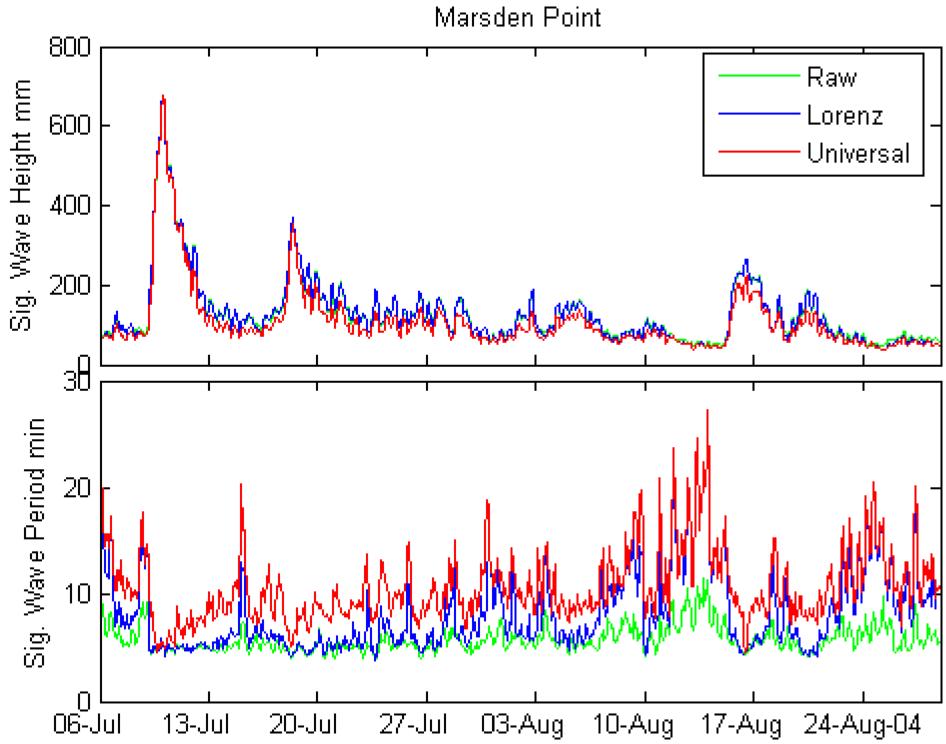


*Figure 4. Comparison of the effects of denoising with different thresholds on significant wave height (upper) and period (lower) for Marsden Point.*

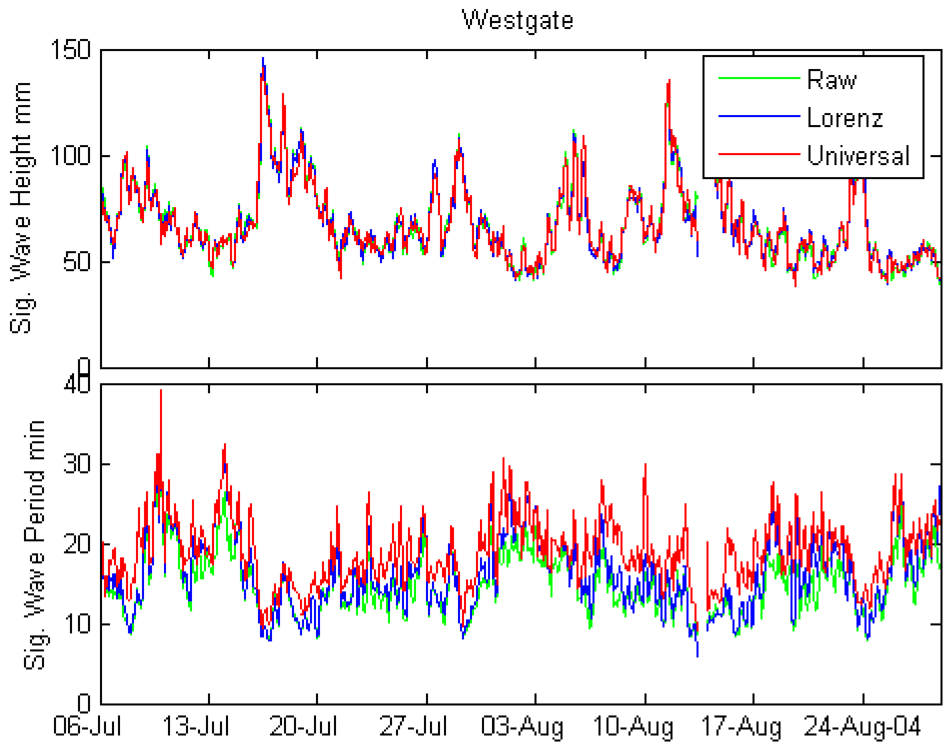


*Figure 5. Comparison of the effects of denoising with different thresholds on significant wave height (upper) and period (lower) for Westgate.*

Although the effect of denoising on significant wave height is negligible in some cases, the effect on the period of the significant waves is almost always substantial. The reason for this is easy to understand. Denoising removes the small waves consisting of only a few points and those points are accumulated into the adjacent, larger wave, thus increasing the period of the larger wave, though probably not its height.

The best way to view how denoising affects period is to plot the distribution of significant wave height with period and time after decomposing using continuous wavelets, as shown in Figures 6 and 7. The figures were prepared using noise thresholds of 0, 48, and 134 mm, corresponding to zero denoising, Lorenz threshold, and universal threshold, the latter two having been calculated from 6 weeks of record. Figure 6 is for a quiescent period. We would expect that in such a period noise would be a significant part of the signal, and indeed this is the case because applying either Lorenz or universal thresholding results in the significant wave height decreasing and the period increasing. In the contour plots, we see that with thresholding, the importance of the low period part of the signal diminishes. Figure 7 shows an event that was accompanied by strong swell, so the long waves are likely to have been infra gravity waves. There is almost no effect of thresholding on this period of record, and the contour maps are indistinguishable from each other, even though the significant period is small and hence must be contained in the first wavelet detail. Thus, we can say that for this site thresholding does not adversely affect events. Indeed, for this site the more severe universal thresholding is suitable.

## Conclusions

Long wave records can be denoised by applying a threshold to the first-level wavelet coefficients. The appropriate threshold varies from site to site, but ideally it should be set so that denoising affects the significant wave height and period in quiescent times, but has little effect during events. Two thresholds were considered, one based on rejecting wavelet coefficients until the energy loss exceeds the parsimony (Lorenz threshold), and the other based on the expected maximum of the noise (universal threshold). For Marsden Point, it appears that most of the energy in the first wavelet detail is noise, so the universal threshold is the appropriate one to use. However, this result is not necessarily applicable to other sites, and each site needs to be considered in detail on its merits.

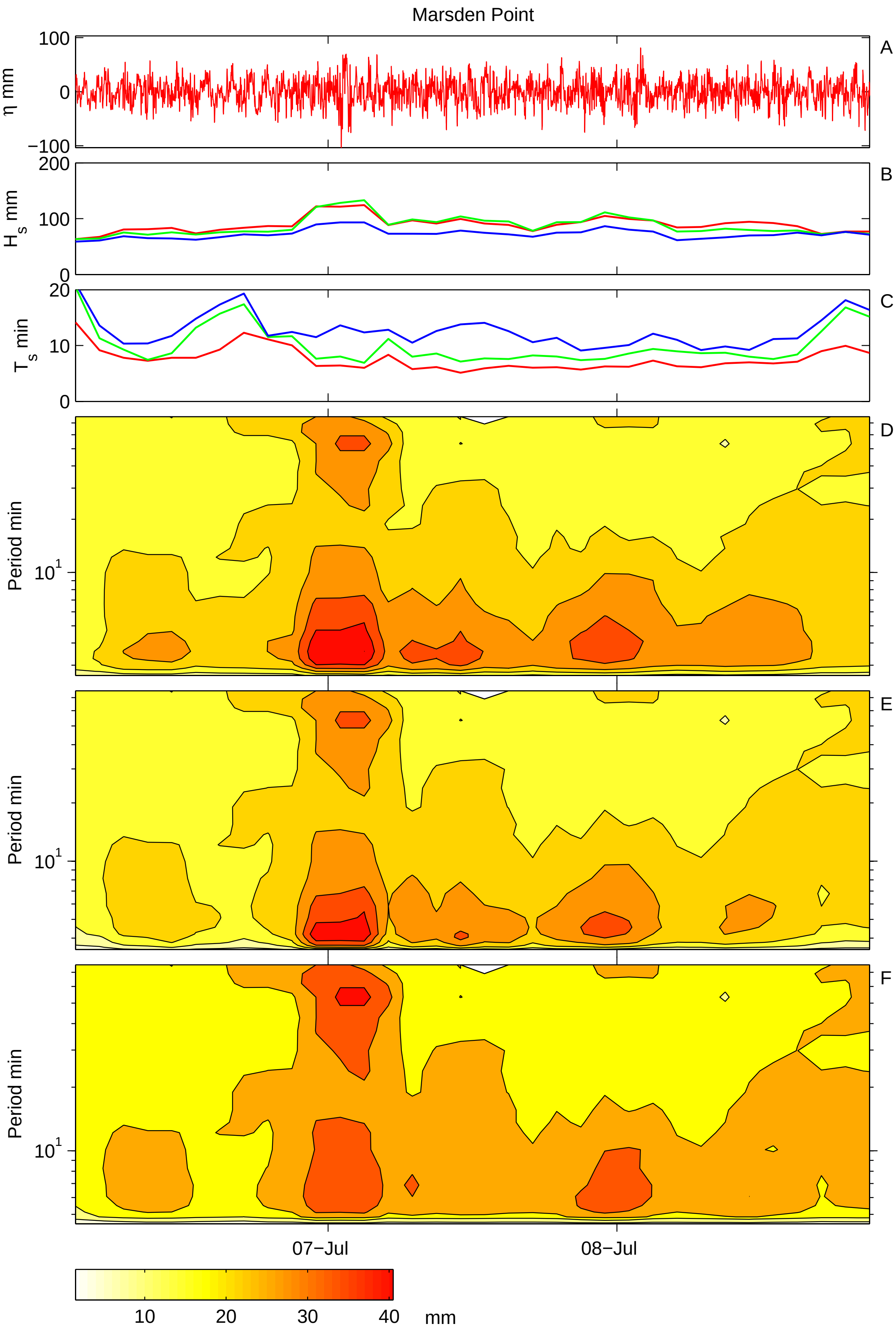


*Figure 6. The effects of denoising on a quiescent period of record, showing A. the long wave record, B. significant wave heights, C. period of significant waves, D. to F. distribution of significant wave height with time and period for zero denoising (blue curve in B and C), denoising using the Lorenz threshold (green curve), and denoising using the universal threshold (red curve) respectively.*

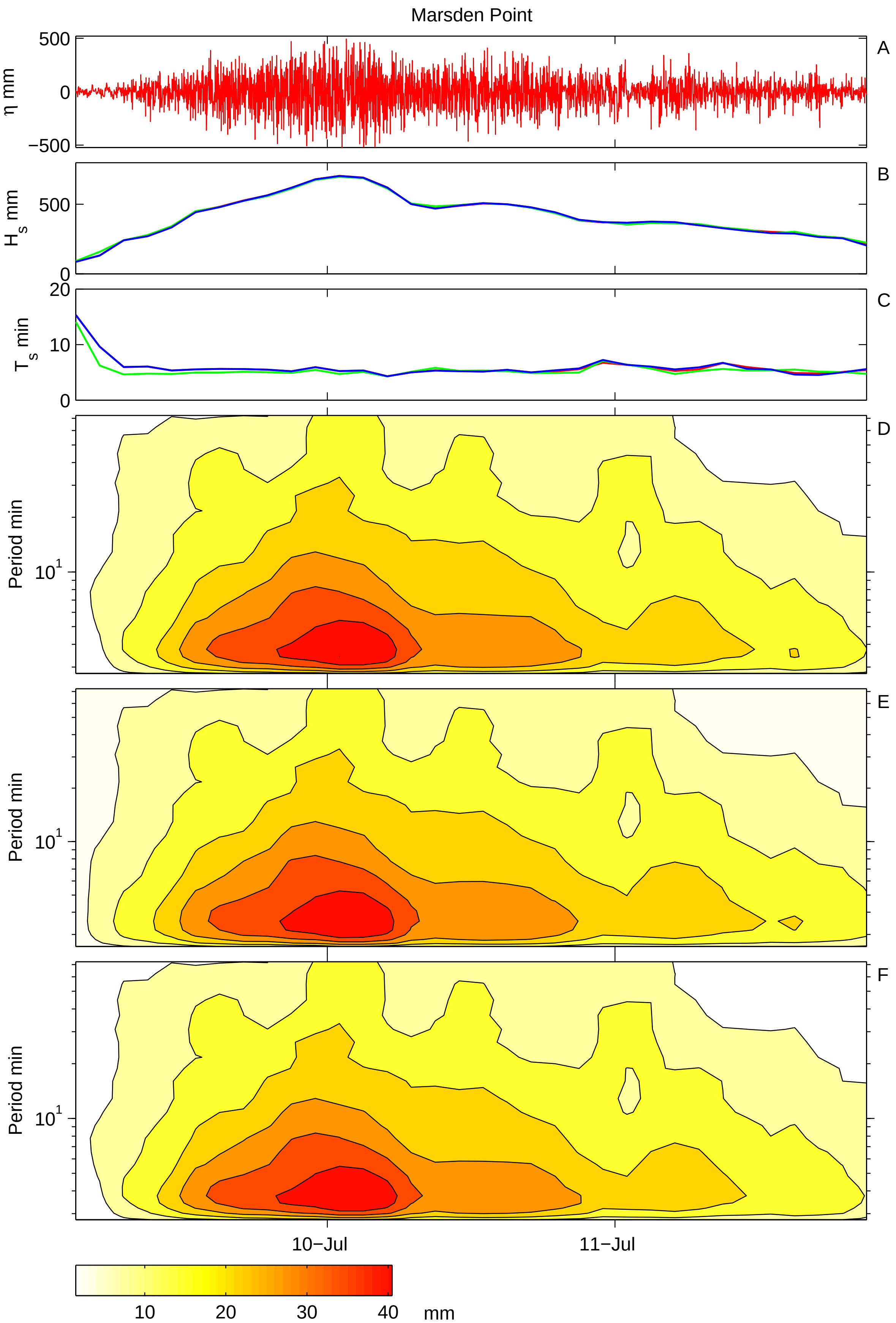


*Figure 7. The effects of denoising on an event (see caption of Figure 6 for details).*

## *References*

Donoho, D. L.; Johnstone, I. M. 1994: Ideal spatial adaptation by wavelet shrinkage. *Biometrika*, 81(3): 425-455.

Katul, G.; Vidakovic, B. 1998: Identification of low-dimensional energy containing/ flux transporting eddy motion in the atmospheric surface layer using wavelet thresholding techniques. *Journal of the Atmospheric Sciences*, 55: 377-389.