

Wavelet Analysis of Tide-affected Low Streamflows Series

Yeo-Howe Lim¹ and Leonard M. Lye²

¹*University of North Dakota and* ²*Memorial University of Newfoundland*

Abstract: In certain rivers that drain very flat terrains in coastal areas, the streamflow series observed at a flow-gauging station may come under the direct influence of the backwater effects of tides. The phenomena may be negligible under conditions of high flows but can be critical under some extreme low-flow conditions. The errors in low flow estimation are large if a proper de-noising is not implemented to remove the effects of the tidal effects. Scrutinizing the hydrologic time series using a standard time-frequency domain based Fourier transform methodology cannot resolve conclusively the sources of the noise. However, a new perspective can be obtained by using a wavelet transformation to analyze the time series in the time-scale domain. By using this approach, a case study involving a streamflow series observed at Kapit, Sarawak, Malaysia yielded conclusive evidence of the influence of tides at the flow-gauging site during the low flow period. Upon confirmation that the noise is indeed of tidal origin, the observed water level series was subjected to an appropriate wavelet-based de-noising procedure to derive a smoothed series. Then, together with an established rating curve, a de-noised discharge series could also be approximated. Low-flow quantiles were subsequently derived by fitting a suitable frequency distribution to the annual minimum series abstracted from the de-noised discharge series. The methodology presented illustrates the potential of using wavelet analysis methods in solving other similar problems.

Key words: Fourier power spectrum, hydrology, low-flow, low flow quantiles, river flow-gauging, streamflow series, tides, wavelets, wavelet de-noising, wavelet power spectrum.

1. Introduction

Wavelet analysis is a major development in the methods of data analysis in the last decade. It evolved from fields including signal processing, physics, and mathematics. The application of wavelet analysis in analysing time based data, particularly those with non-stationary characteristics, has been found to be very successful (Nason and Sachs, 1999; Percival, 2000). The property of being localized in time or space as well as scale or frequency provides a time-scale map of a signal, enabling the extraction of features that vary in time. This

makes wavelets an ideal tool for analyzing signals of a transient or non-stationary nature.

The literature about the application of wavelets in the field of hydrology and water resources engineering is still relatively scarce. Jay and Flinchem (1995, 1997, 1999) have demonstrated the interaction of fluctuating river flows with a barotropic tide using wavelet analysis. Smith et al. (1998) were able to characterize streamflows from different climatic regions in the United States using the wavelet transform. Applications in the meteorological field are also useful as the processes involved are closely linked to hydrology. An excellent example is the paper presented by Torrence and Compo (1998). Gan (2001) showed the evidence of the droughts in the Canadian Prairies being related to the current circulation patterns in the Pacific Ocean using a wavelet analysis.

Many aspects of hydrology, in one way or another, involve an analysis of the hydrologic time series. Hence, if there is any tangible improvement in the technique of time series analysis, a significant impact on the understanding of hydrological processes may be uncovered. Motivated by this potential benefit, a common flow-measurement problem in a river is reassessed and analyzed using the new tool of wavelets. A particular river flow series is suspected of being influenced by tides when the river discharge has dropped well below the normal flow range, or *low-flows*, in the terminology of hydrology. It is not possible to ascertain the presence of the tides when the fluctuation of river levels can be due to other dynamic processes, such as the local runoff effects due to localized rainfall events. If the presence of the tide-effect is proven, the flow series at low-flow periods can to be smoothed using a robust wavelet de-noising method. An extreme-event frequency analysis can then be performed on the abstracted annual low-flow series to determine the quantiles of low-flows associated with certain non-exceedance probabilities which can be used for the design of water resources projects.

2. Wavelet Theory and Methods

All wavelets are based on one mother wavelet, $\psi_0(u)$, a function with some special properties. u is the dimensionless time parameter for time series analysis. The function must oscillate and decay, preferably rapidly, giving

$$\int \Psi_0(u) du = 0 \quad (2.1)$$

Typical examples of mother wavelets are Morlet, Mexican hat, and Daubechies. In harmonic analysis, frequency is the single parameter used. By contrast, wavelets are controlled by two parameters, dilations λ (or “scale”), and translations, t , of the mother wavelet.

$$\Psi_{\lambda,t} = \frac{1}{\sqrt{\lambda}} \Psi_0 \left(\frac{u-t}{\lambda} \right) \quad (2.2)$$

Physically, λ is used to stretch or shrink the wavelet and t is for translating the wavelet along the time axis. For most applications a restricted set is used. A unique power of 2 is specified for the scale $\lambda = 2^j$ and $t = k$ where j and k are integers. For a specially chosen mother wavelet, the collection of $\{\Psi_{\lambda,t}(u)\}$ for all integers j, k form an orthonormal basis for various function spaces. This is similar to the Fourier series where the cosine and sine functions form an orthonormal basis.

2.1 Continuous wavelet transform W

Let the observed time series x_n with equal time spacing Δt and $n = 0, 1, \dots, N-1$. Continuous wavelet transform is the convolution of the discrete series with a scaled and translated version of a wavelet $\psi(u)$:

$$W_n(\lambda) = \sum_{m=0}^{N-1} x_m \psi^* \left(\frac{(m-n)\Delta t}{\lambda} \right) \quad (2.3)$$

where $*$ indicates complex conjugate. This is similar to the convolution found in the classical Fourier transformation. By varying the scale λ and translating along the localized time index n , a picture (scalogram) can be constructed showing both the amplitude of any features versus the scale and how this amplitude varies with time.

The continuous wavelet transform can be approximated by performing N convolutions in the Fourier space using a discrete Fourier transform given by:

$$\hat{x}_f = \frac{1}{N} \sum_{m=0}^{N-1} x_m e^{-2\pi I f n / N} \quad (2.4)$$

where \hat{x}_f is the discrete Fourier transform of x_n at frequency index f . The wavelet transform can be obtained from the inverse Fourier transform of the product:

$$W_n(\lambda) = \sum_{f=0}^{N-1} \hat{x}_f \psi^*(\lambda \omega_f) e^{i \omega_f n \Delta t}, \quad (2.5)$$

where ω_f is the angular frequency.

2.2 Function approximation

Any function g can be represented in terms of a superposition of wavelets of different dilations and translations, as

$$g(u) = \sum_{j=-\infty}^{\infty} \sum_{k=-\infty}^{\infty} d_{j,k} \psi_{j,k}(u) \quad (2.6)$$

where $d_{j,k}$ coefficients of expansion which can be found by

$$d_{j,k} = \int_{-\infty}^{\infty} f(u) \psi_{j,k}(u) du \quad (2.7)$$

Wavelets can represent some functions using far fewer coefficients than can Fourier and this is the advantage used in image processing to achieve data compression.

2.3 Wavelet power spectrum

The wavelet approach in deriving the spectrum is similar to the Fourier approach but a time/scale instead of a time/frequency decomposition. Because the wavelet function $\psi(\cdot)$ is in general complex, the wavelet transform $W_n(\lambda)$ is also complex.

The wavelet power spectrum of the series with scale λ is defined as

$$P_W(\tau) = \{|W_n(\lambda)|^2 : \lambda \in \tau\} \quad (2.8)$$

where τ is a set of scales. For plotting of the wavelet power spectrum, it is convenient to express the scales as fractional powers of two (Torrence and Compo, 1998):

$$\lambda_h = \lambda_0 2^{h\Delta h}, h = 0, 1, \dots, H \quad (2.9)$$

$$H = \frac{1}{\Delta h} \log_2 \left(\frac{N\Delta t}{\lambda_0} \right) \quad (2.10)$$

where λ_0 is the smallest resolvable scale and H is the largest scale. An equivalent Fourier period $1/f$ can be related to the wavelet scale for a particular wavelet function. This is done by substituting a cosine wave of a known frequency into equation (2.5) and computing the scale that matches the maximum wavelet power spectrum. For wavelet such as Morlet wavelet,

$$\psi_0(u) = \pi^{-1/4} e^{i\omega_0 u} e^{-u^2/2} \quad (2.11)$$

with the nondimensional frequency ω_0 set to 6 to satisfy admissibility condition (Farge, 1992), $1/f = 1.03\lambda$. This allows the association of any specific periodic pattern found in the wavelet power spectrum with the real time period.

2.4 Normalization and expectation

A common normalization for the wavelet spectrums is used to make it easier to compare among them. The wavelet function at each scale λ is normalized to have unit energy. The expectation for the wavelet power spectrum $|W_n(\lambda)|^2$ is

$$E|W_n(\lambda)|^2 = N \times E|\hat{x}_f|^2 \quad (2.12)$$

The expectation of $|\hat{x}_f|^2$ is σ^2/N where σ^2 is the variance. For a white-noise process, the expectation value for the wavelet power spectrum at all λ and n is

$$E|W_n(\lambda)|^2 = \sigma^2. \quad (2.13)$$

Since the square of a normally distributed variable is chi-square distributed with one degree of freedom, the normalized Fourier spectrum $|\hat{x}_f|^2$ is chi-square distributed (χ_2^2) with two degree of freedoms. If the Fourier components are normally distributed, then the wavelet coefficients are also normally distributed. It follows that the wavelet power spectrum are also chi-square distributed and the 95% confidence level can be estimated (Torrence and Compo, 1998).

2.5 Wavelet de-noising

In general, a smoothed signal $s(t)$ can be recovered from a signal $y(t)$ by

$$s(t) = y(t) - \sigma x(t) \quad (2.14)$$

where $x(t)$ represents noise and σ is the noise level. The wavelet de-noising method is a nonparametric estimation of the function s using orthogonal basis. There are several issues to be addressed in wavelet de-noising; among them are the choices of wavelet, decomposition levels, and thresholding methods. The descriptions of these choices are found in recent papers by Donoho and Johnstone (1994, 1995) among others. However, some heuristic judgments are still required in selecting the optimal combination.

3. Problem of Tides at Stream Gauging Stations

Typically at a river or stream discharge station, only a limited number of discharge measurements are carried out at some randomly chosen water levels. A relationship between discharge Q and water level H is subsequently established by fitting a suitable curve through the set of Q - H data, known as the discharge rating curve for the station. An automatic recorder operated on a permanent basis then yields a continuous water level series at the site. A continuous long-term discharge series at the site can be derived based on the established rating

curve. Some parameters for engineering designs are derived by performing various time series analyses using the discharge series. One major task is to perform a frequency analysis on the extreme event abstracted from the series such that the probability of occurrence associated with the magnitude of floods or droughts (low-flows) can be rationalized. The design benchmarks of many civil engineering projects are dependent on the derivation of the frequency analysis. Owing to the critical outcomes of the analysis, the accuracy of the observed water level series used cannot be compromised. The key lies in observing and quantifying the extreme discharges well.

An established river gauging station is located at Kapit, a riverside town on the mid course of Batang Rejang, the largest river system in Malaysia. The watershed area consists of about 18,100 km² of rainforest on Borneo Island. Although the station is located at about 190 km upstream of the river mouth, the technicians who frequently measured the discharge at the site reported that the water level recorded at the station might be affected by tidal intrusion from the sea during certain low-flow periods. It is indeed difficult to confirm the validity of the claim based on a casual inspection of the fluctuating water levels. This kind of problem is quite common in places where tidal influence reaches far inland. To investigate the problem using a traditional approach would involve a spectral analysis of the Fourier transforms, which may provide some insights about the occurrence of tides. However, as shown in the later sections, wavelet analysis method presents a compelling different perspective and offers a much more solid conclusion.

4. Confirming the Presence of Tides

Figure 1 shows the plot of a segment of the recorded hourly water level series and the corresponding discharge series at Kapit discharge measurement station, which started on 23rd September 1986 and lasted 193 days. Some high frequency fluctuations of water levels can be spotted at certain low-flow periods.

Figure 2 shows the Fourier power spectrum for the hourly levels. The spectrum shows two distinct peaks, at periods of about 12.5 and 6 hours (or frequency of 0.08 and 0.16). Tidal phenomenon is mainly caused by lunar excitation, which has a major periodic cycle of about 12.5 hours. Other periodic intervals are also observed which account for the other lunar-solar effects. Despite the coincidence of one of the period at 12.5 hours, the actual presence of tidal effects on the water level series still remains a suspicion. Since no further information can be obtained from the spectrum, no conclusive remark can be made.

Wavelet analysis is a relatively new, powerful, analytical tool that is increasingly used as an alternative to Fourier analysis. It follows naturally that an alternative analysis of the data series using the wavelet analysis may provide a

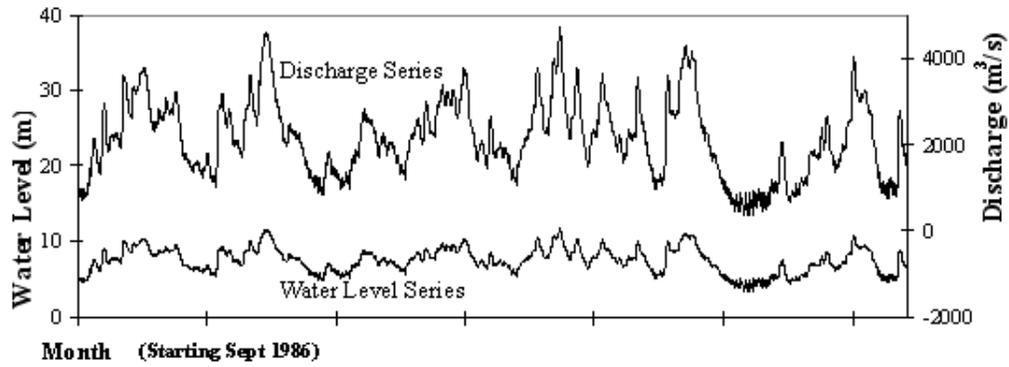


Figure 1: Kapit hourly water level and discharge series

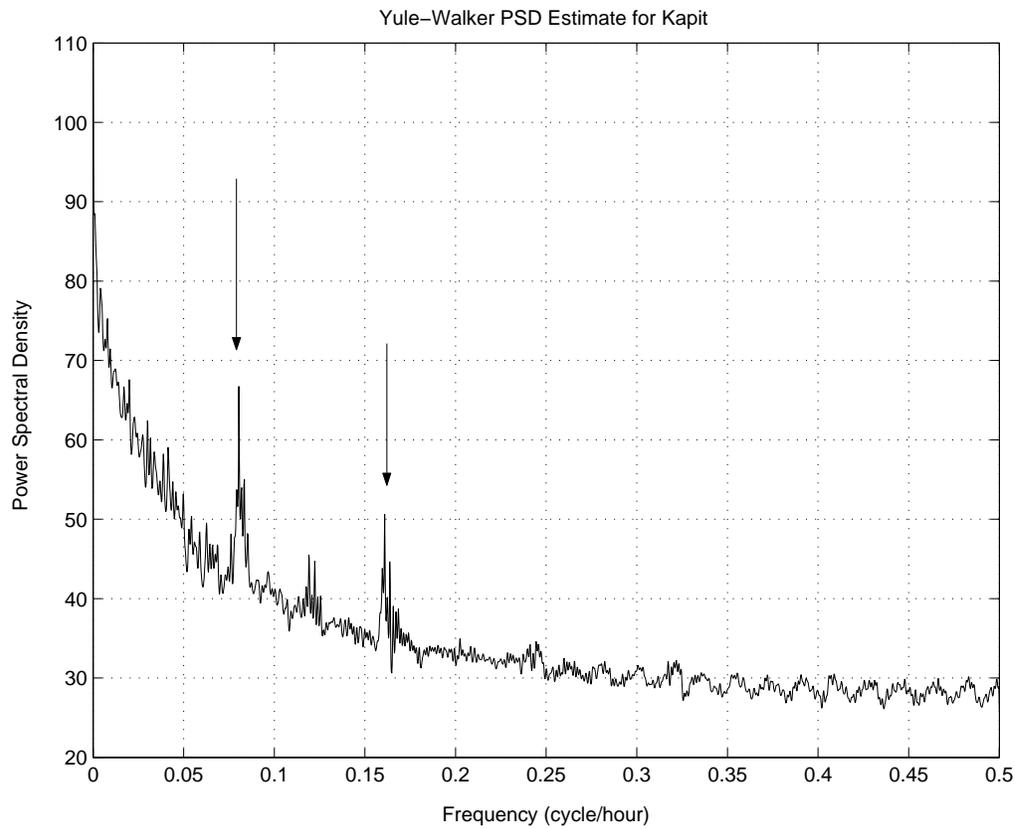


Figure 2: Fourier power spectrum of Kapit hourly flows series

better perspective.

A major portion of Figure 3 is the plot of wavelet power spectrum $P_W(\tau)$ as defined in equation (2.8) and is commonly referred as a scalogram. The x axis is the real time space in hour and the y axis is the equivalent Fourier period $1/f$ in hour. The Morlet wavelet as defined in equation (2.11) is used in deriving the wavelet power spectrum. From a casual inspection of the spectrum, it is confirmed that the spectral power is relatively weak at the short periods or at high frequencies. However, it is not difficult to locate some distinct dual-patches of relatively higher contours at around scale factor (or equivalent Fourier period) of around 12 and 24 hours (e.g. at about 3750th hours from the beginning of the time series). A dual-pointing arrow is shown in Figure 3 to indicate that the timing actually corresponds to a low-flow period, typified by a trough on the time series plot. Hence it can be seen that wavelet power spectrum has an added dimension of information on the time-scale domain, not available in the traditional Fourier spectra. Now it becomes transparent that the sinusoids or noises that occur at the time of low-flow, e.g., at 3750th hour, have a periodicity of around 12.5 and 25 hours. These specific periodicities certainly match the traceable signatures of the semi-diurnal tides which dominates the coastal waters of the study area.

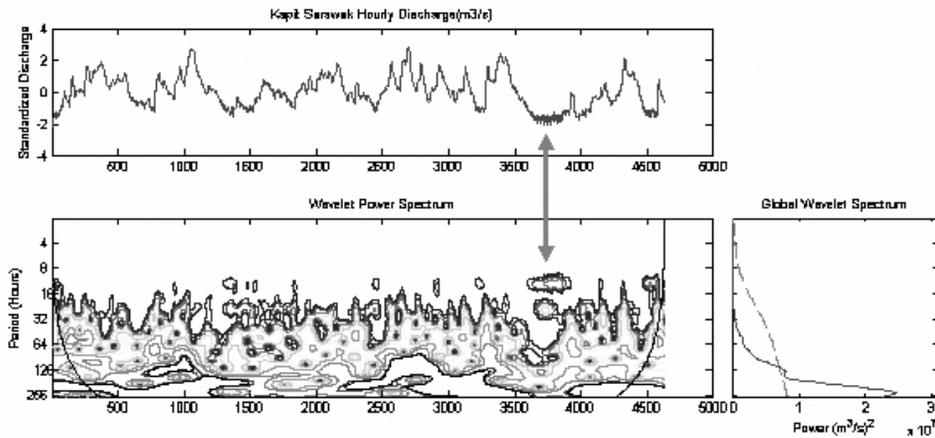


Figure 3: Local and global wavelet spectra of hourly flows at Kapit (193 days).

By taking an averaged wavelet spectrum over all the local wavelet spectra, a global wavelet power spectrum can be derived. This plot is shown at the lower-right portion of Figure 3. There is no significant spike at the shorter periods indicating that the higher frequency component attributed by tides is not dominant throughout the run of the series. Since the tides were present in the

predominantly fluvial discharge series, it could only be felt at a certain segment of the series. For this investigation, the local wavelet power spectrum provides much more insights.

5. Wavelet De-noising of Data Series

The Kapit data series was then subjected to a wavelet de-noising process. There are two possible ways of using the observed data, i.e., either using the water level or the discharge series. Since it is reasonable to assume that there is a physical maximum tidal level H_{\max} unique to the site, the use of water level series has an advantage in relating the de-noised series to the H_{\max} . If deriving an overall streamflow series is the prime objective, the following algorithm is proposed:

1. Adapt a satisfactory combination of wavelet de-noising option and estimate H_{\max} .
2. Perform wavelet de-noising on the water level $y(t)$ and obtain a de-noised series $s(t)$.
3. For $s(t) < H_{\max}$, $s(t)$ is assigned to the final series $z(t)$, and for $s(t) > H_{\max}$, the segment of original $y(t)$ series is assigned to $z(t)$.
4. Check for continuity and smoothness in the transition around H_{\max} .
5. Adjust H_{\max} and repeat step 2 to 4 until the continuity and smoothness criterion are satisfied.
6. Convert the de-noised water level series $z(t)$ to discharge series using the stage-discharge rating equation developed after many field discharge measurements.
7. Perform low-flow frequency analysis based on the de-noised discharge series.

Several combinations involving various options of wavelet type, decomposition level, and thresholding methods are considered. A few combinations are adopted to further the de-noising process after an initial elimination process by visual inspection of the de-noised series. Figure 4 shows the detailed results of applying various de-noising options to a short water level series influenced by tides at Kapit. A brief review of the observed river level series at a station called Entawa located 35km upstream of the Kapit station, shows that the average rate of change in water level is about 0.05 metre/hour, while the maximum rate ever recorded is 0.9 metre/hour. These physical values can be used to eliminate some poor combinations of wavelet type and decomposition level. However, it is a heuristic decision in deciding the best among the better combinations. The Daubechies-4 wavelet at decomposition level 4 is assumed to be the best in terms of smoothness and continuity. Table 1 shows the result of performing de-noising on the data series based on the best choice of Daubechies-4 wavelet, at decomposition level 4

Table 1: Results of wavelet de-noising on low-flows at Kapit.

| Year | Minimum Water Level (m) | | Discharge (m ³ /s) | | Change in Discharge | |
|------|-------------------------|-----------|-------------------------------|-----------|---------------------|---------|
| | Original | De-noised | Original | De-nosied | Absolute | Percent |
| 1986 | 3.42 | 4.05 | 352 | 547 | 195 | 55.2 |
| 1987 | 3.72 | 3.83 | 441 | 475 | 34 | 7.8 |
| 1988 | 2.89 | 3.63 | 214 | 413 | 200 | 93.6 |
| 1989 | 2.83 | 3.49 | 200 | 372 | 173 | 86.5 |
| 1990 | 4.04 | 4.55 | 543 | 722 | 178 | 32.8 |
| 1991 | 2.85 | 3.12 | 204 | 271 | 66 | 32.5 |
| 1992 | 3.16 | 3.77 | 281 | 456 | 175 | 62.4 |
| 1993 | 3.32 | 4.03 | 324 | 540 | 216 | 66.6 |

Note: Option used in wavelet de-noising: Daubechies-4 wavelet, level 4, soft and universal threshold, H_{\max} of 7.75m.

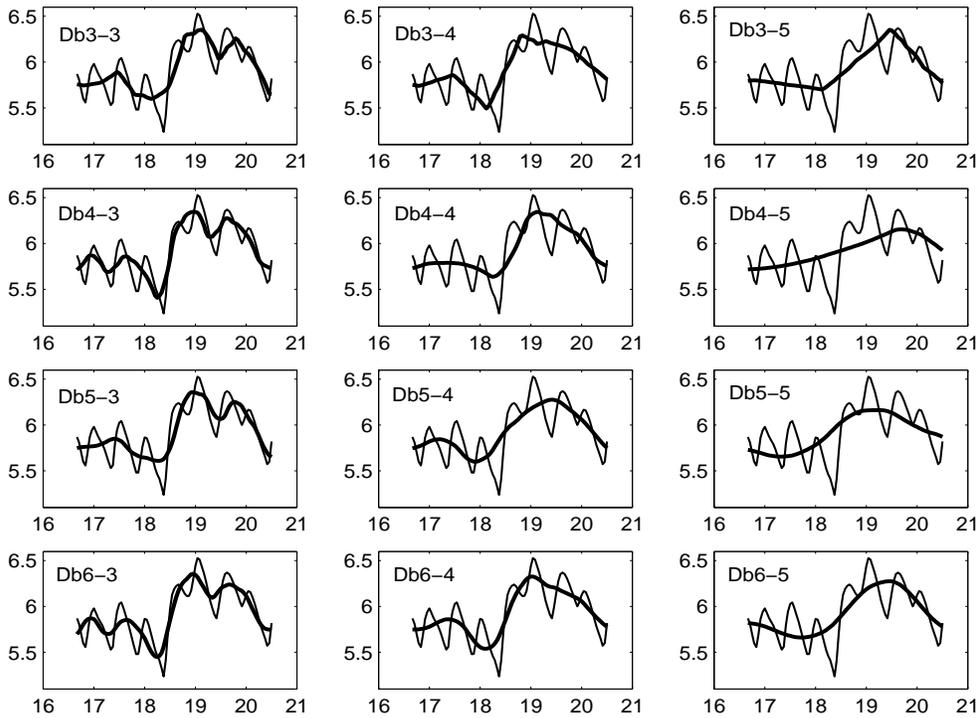


Figure 4: Comparing de-noising options of using various wavelet types and decomposition levels. The vertical axes are the water level elevations (metres) and the horizontal axes are the days in April 1992.

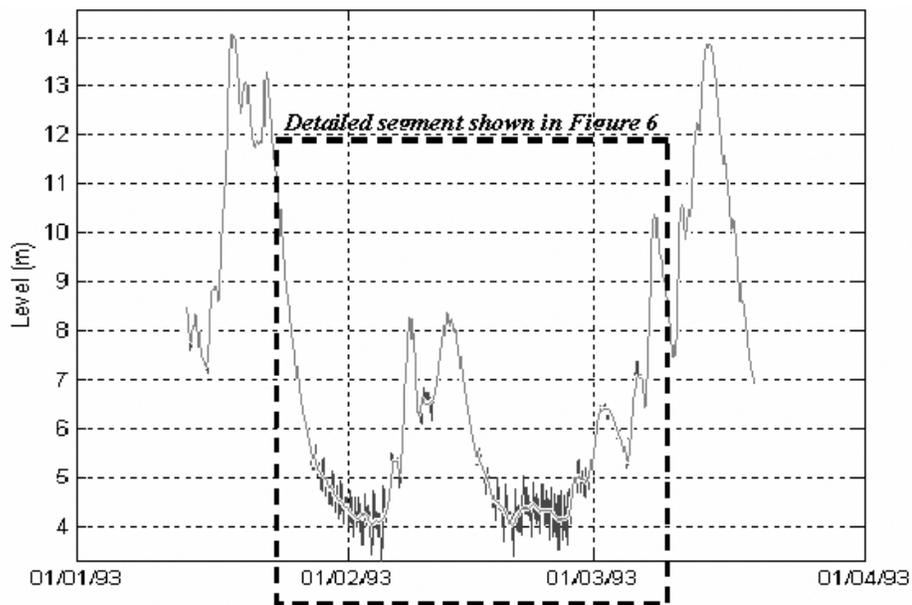


Figure 5: A segment of wavelet de-noised hourly water levels at Kapit

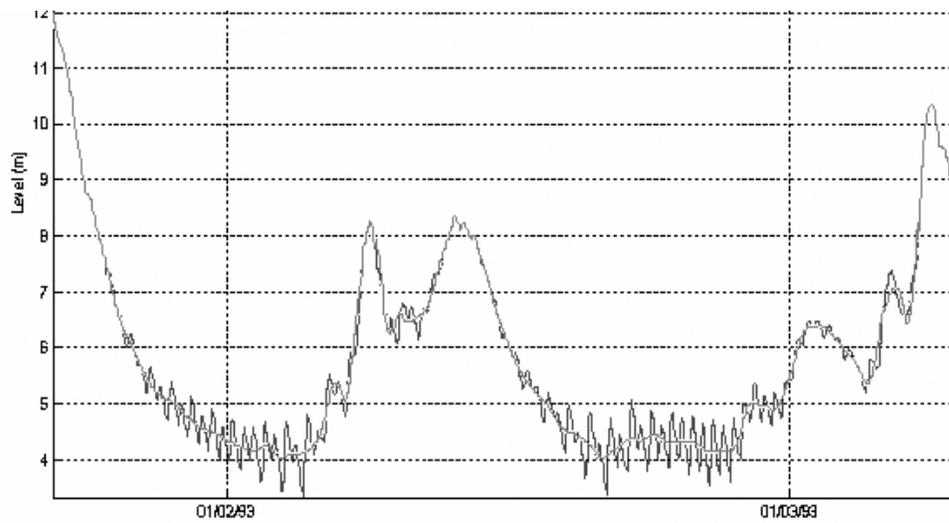


Figure 6: Details of de-noised low-flow levels at Kapit

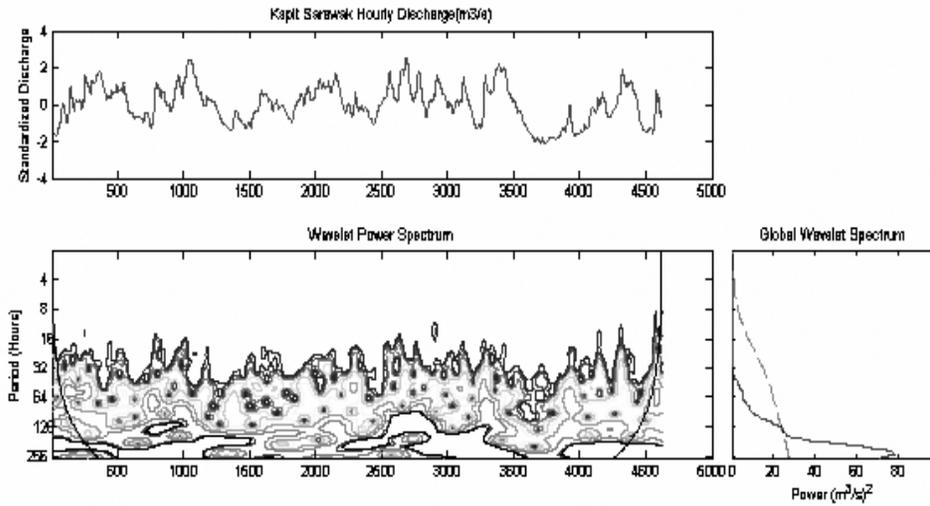


Figure 7: Wavelet Power Spectra of De-noised Kapit Hourly Discharge Series

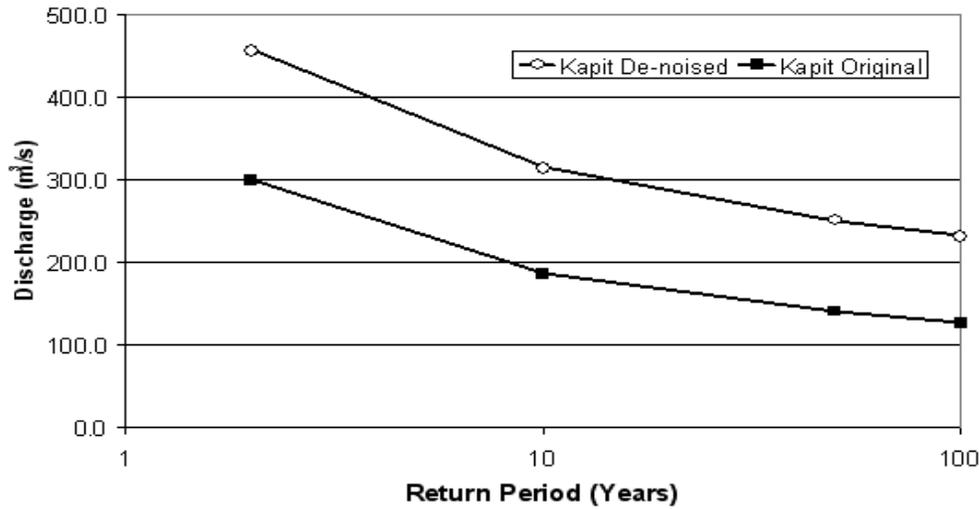


Figure 8: Comparison of low-flow quantiles of original and de-noised series (low-flow frequency curve fitted with log-normal distribution)

using soft and universal thresholding.

A segment of the de-noised water level series is shown in Figure 5 with a detailed view shown in Figure 6. The smoothing of the series is quite satisfactory considering the varying patterns of tidal noise. For checking purposes, the de-

noised discharge series was subjected to the same wavelet decomposition as before, and the corresponding wavelet spectra are shown in Figure 7. In comparison with Figure 3, the tidal noises at low-flow periods are obviously removed as shown in the time series plot. These changes are clearly reflected in the local wavelet spectrum; the isolated patches of relatively high local spectral energies associated with tides at the low-flow periods are removed.

6. Frequency Analysis of Low-flow Series

Once the issue of tidal influence was confirmed, the water level series was appropriately de-noised. A de-noised discharge series is subsequently derived from the de-noised water level series using the discharge rating curve established for the station. An extrapolation of the curve is often required for the upper and lower tails of the rating curve. It is clear that the river fluvial discharges released from a large watershed should not fluctuate rapidly as exhibited by the recorded water levels. The use of the de-noised series is a logical approximation solution.

A low-flow frequency analysis is often performed on an abstracted annual minimum flow series such that the low-flow quantiles, the magnitude associated with a probability of occurrence can be estimated. The method is illustrated using the de-noised series, albeit the length of data is far from ideal. By fitting the extreme series to a few distributions, it is found that the log-normal distribution fit best. The distribution can be used to derive the low-flow quantiles. In general the de-noising exercise results in higher low-flow quantiles as shown in Figure 8. The neap tide has a drawdown effect on the water level. Using various options available to de-noising exercise can lead to a few possible outcomes. Since there is no absolute way in confirming the characteristic values of the de-noised series, one method of exploring the error limits is by sensitivity analysis. The options available in wavelet de-noising are systematically adjusted and the de-noised series are subjected to frequency analysis.

7. Conclusion

Wavelet de-noising has been shown capable of locating the periodic features of tides and confirms that they occur during the low-flow time series. The traditional Fourier analysis does not provide any information on the time of occurrence of tides at low-flow periods. On the contrary, for the wavelet method, it is explicitly shown on the local wavelet spectrum.

The low-flow series has been de-noised and conventional low-flow frequency analyses are performed. The derived low-flow quantiles are much higher using the de-noised series indicating that the errors are large without de-noising by the wavelet approach.

Comparison of the various de-noised series shows that Daubechies-4 at decomposition level 4 is the best option in performing de-noising related to tides.

Acknowledgement

The authors are thankful to the Department of Irrigation and Drainage, Sarawak, Malaysia for supplying the hydrological data. A suite of MATLAB software developed by Torrence and Compo(1998) are adapted and used in deriving the local and global wavelet power spectra.

References

- Donoho, D. L. and Johnstone, I. M. (1994). Ideal Spatial Adaptation via Wavelet shrinkage. *Biometrika* **81**, 425-455.
- Donoho, D. L. and Johnstone, I. M. (1995). Adapting to unknown smoothness via wavelet shrinkage. *Journal of American Statistical Association* **90**, 1200-1224.
- Farge, M. (1992). Wavelet transforms and their applications to turbulence. *Annual Review of Fluid Mechanics* **24**, 395-457.
- Gan, T. Y. (2001). Precipitation of Western Canada – Wavelet, scaling, and multifractal analysis and teleconnection to large-scale climate anomalies. *Proceedings of 15th Hydrotechnical Specialty Conference*. Canadian Society of Civil Engineers, Victoria, 2001.
- Jay, D. A. and Flinchem, E. P. (1997). Interaction of fluctuating river flow with a barotropic tide: A demonstration of wavelet tidal analysis methods. *Journal of Geophysical Research* **102**, 5705-5720.
- Jay, D. A. and Flinchem, E. P. (1999). A comparison of methods for analysis of tidal records containing multi-scale non-tidal background energy. *Continental Shelf Research* **19**, 1695-1732.
- Jay, D. A. and Flinchem, E. P. (1995). Wavelet transform analyses of non-stationary tidal currents. *Proceedings of The IEEE Fifth Working Conference on Current Measurement* (Edited by A. E. Williams), 101-106.
- Nason, G. P. and Sachs, R. V. (1999). Wavelets in time series analysis. *Phil. Trans. R. Soc. Lond. A* **357**, 2511-2526.
- Percival, D. B. (2000). *Wavelet Methods for Time Series Analysis*. Cambridge University Press.
- Smith, L. C., Turcott, D. L. and Isacks, B. L. (1998). Stream flow characterization and feature detection using a discrete wavelet transform. *Hydrological Processes* **12**, 233-249.
- Torrence, C. and Compo, G. P. (1998). A practical guide to wavelet analysis. *Bulletin of The American Meteorological Society* **79**, 61-78.

Received October 28, 2002; accepted April 1, 2003.

Yeo-Howe Lim
Department of Civil Engineering
University of North Dakota
Grand Forks, ND 58202-8115, USA
howe.lim@mail.und.nodak.edu

Leonard M. Lye
Faculty of Engineering and Applied Science
Memorial University of Newfoundland
St. John's, NL, A1B 3X5, Canada
llye@engr.mun.ca