Stream flow characterization and feature detection using a discrete wavelet transform

Laurence C. Smith,* Donald L. Turcotte and Bryan L. Isacks

Department of Geological Sciences, Cornell University, Ithaca, NY, USA

Abstract:

An exploration of the wavelet transform as applied to daily river discharge records demonstrates its strong potential for quantifying stream flow variability. Both periodic and non-periodic features are detected equally, and their locations in time preserved. Wavelet scalograms often reveal structures that are obscure in raw discharge data. Integration of transform magnitude vectors over time yields wavelet spectra that reflect the characteristic time-scales of a river's flow, which in turn are controlled by the hydroclimatic regime. For example, snowmelt rivers in Colorado possess maximum wavelet spectral energy at time-scales on the order of 4 months owing to sustained high summer flows; Hawaiian streams display high energies at time-scales of a few days, reflecting the domination of brief rainstorm events. Wavelet spectral analyses of daily discharge records for 91 rivers in the US and on tropical islands indicate that this is a simple and robust way to characterize stream flow variability. Wavelet spectral shape is controlled by the distribution of event time-scales, which in turn reflects the timing, variability and often the mechanism of water delivery to the river. Five hydroclimatic regions, listed here in order of decreasing seasonality and increasing pulsatory nature, are described from the wavelet spectral analysis: (a) western snowmelt, (b) north-eastern snowmelt, (c) mid-central humid, (d) south-western arid and (e) 'rainstorm island'. Spectral shape is qualitatively diagnostic for three of these regions. While more work is needed to establish the use of wavelets for hydrograph analysis, our results suggest that river flows may be effectively classified into distinct hydroclimatic categories using this approach. © 1998 John Wiley & Sons, Ltd.

Hydrol. Process., Vol. **12**, 233–249 (1998)

KEY WORDS   stream flow variability; wavelet spectral analyses; hydroclimatic regions

INTRODUCTION

The wavelet transform is a recent advance in signal processing that has attracted much attention since its theoretical development in 1984 (Grossman and Morlet, 1984). Its use has increased rapidly in communications, image processing and optical engineering applications as an alternative to the Fourier transform in preserving local, non-periodic, multiscaled phenomena. Recently, wavelet analysis has been applied to a variety of geophysical signals, including a bathymetric profile near Hawaii (Little *et al.*, 1993), time-series of the Earth's length of day (Gambis, 1992; Chao and Naito, 1995) and solar irradiance (Kiang *et al.*, 1994), spatial and temporal variations in sea surface temperatures (Meyers and O'Brien, 1994) and radar altimeter profiles used for computation of a geoid surface model (Cazenave *et al.*, 1995). Wavelet transformations of meteorological time-series have revealed structures in boundary layer turbulence (Hagelburg and Gamage, 1994; Hayashi, 1994; Turner *et al.*, 1994; Katul and Parlange, 1995), winter thunderclouds (Takeuchi *et al.*,

---

1994), atmospheric gravity waves (Sato and Yamada, 1994) and the scale of fluctuations in spatial rainfall fields (Kumar and Foufoula-Georgiou, 1993). These and other studies have shown that the wavelet transform is an effective tool for precisely locating irregularly distributed, multiscaled features in time or space.

The problem of quantifying stream flow variability, particularly as it relates to climate, has been the subject of numerous studies (Meko and Stockton, 1984; Lins, 1985; Redmond and Koch, 1991; Webb and Betancourt, 1992; Cayan et al., 1993; Guetter and Georgakakos, 1993). Significant progress has been made using harmonic and multivariate statistical analysis to identify patterns of stream flow variation over large spatial areas. For example, variations in monthly and annual stream flow have now been linked to the La Niña (Dracup and Kahya, 1994) and El Niño–Southern Oscillation, or ENSO, events (Redmond and Koch, 1991; Kahya and Dracup, 1993; Simpson et al., 1993). Fourier spectral analyses have been used to reveal climate-related periodicities in stream flow (Entekhabi, 1984; Hameed, 1984). Kunhel et al. (1990) interpreted some low-frequency periodicities revealed by Fourier spectral analyses of rainfall and stream flow time-series in south-east Australia to be caused by ENSO events. Seasonal flow variations in glacial meltwater streams have been described by comparing Fourier power spectra for monthly blocks of daily discharge, temperature and precipitation (Gudmandsson, 1970). However, a shortcoming of standard Fourier analysis is that it does not retain the location of a particular event in time or space, nor does it perform well on irregularly spaced events or non-stationary signals. These problems are mitigated through the use of a windowed short-term Fourier transform (STFT) translated along a data series, such as the Gabor transform (Gabor, 1946). However, because the STFT frequency must be varied within a window function that is fixed in both length and form, time and frequency resolution cannot both be made small or large; one must be sacrificed for the other (Daubechies, 1990; Bentley and McDonnell, 1994; Kiang et al., 1994). Similarly, a wavelet basis function that has good localization in the time domain will have poor localization in the frequency domain, and vice versa. However, once a wavelet basis function is chosen, it may be rescaled to accommodate a large range of time-scales, without the need for a window of arbitrary length. The effects of the choice of wavelet on the transform's ability to detect the position and scale of hydrograph peaks is discussed further in Appendix A.

The next section describes the wavelet transform used in this study. Detection of transient stream flow features with wavelet scalograms is illustrated in the following section, using two typical rivers that display time-dependent discharge fluctuations. Finally, wavelet spectral analyses of daily discharge records for 91 rivers in the United States and some tropical islands are presented. Strong similarity between wavelet spectra from rivers of like climatic region reveals the potential value of wavelet analysis for identifying river hydroclimatic regime.

THE WAVELET TRANSFORM

Unlike the STFT, wavelet basis functions are not fixed in either length or form; successive wavelets are instead rescaled with a scaling factor a. Each wavelet is passed over the signal $s(t)$ with a translation in time b. In continuous form, the wavelet transform (WT) is expressed as:

$$WT(b, a) = \frac{1}{\sqrt{a}} \int g\left(\frac{t - b}{a}\right) s(t) \, \mathrm{d}t$$

where $g(t)$ is the so-called 'mother wavelet' at $b = 0$, $a = 1$. All subsequent wavelets are rescaled versions of the mother wavelet. The area of each wavelet must sum to zero, and wavelet amplitude decreases as scale a increases. When a is increased by a power of two, a suite of wavelets is generated that can accommodate a wide range of feature scales found in the signal $s(t)$. Even finer temporal sampling may be achieved using non-orthogonal wavelets. This study utilizes the non-orthogonal 'Mexican hat' wavelet, which is illustrated in Figure 1 and is the second derivative of the Gaussian function:

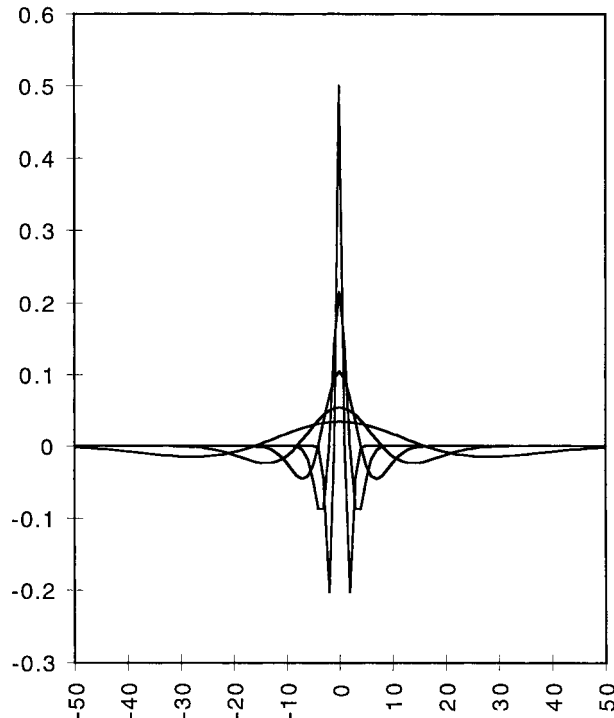$$g(x) = (1 - x^2)\exp(-x^2/2)$$

Figure 1. Discrete, normalized approximations of the 'Mexican hat' wavelet function. Amplitude decreases and wavelet width increases with increasing wavelet scale a

Substitution into the expression for the continuous WT, with $x = t - b$ yields:

$$WT(b, a) = \frac{1}{\sqrt{a}} \int \left[1 - \left(\frac{t - b}{a} \right)^2 \right] \exp \left[-\frac{1}{2} \left(\frac{t - b}{a} \right)^2 \right] s(t) \, \mathrm{d}t$$

which may be implemented numerically with a direct convolution procedure.

For orthogonal wavelets, total wavelet energy must be conserved for wavelets of different scale. We have also imposed this restriction on the non-orthogonal wavelet used here. With direct convolution, this may be accomplished by normalizing the energy of the kernel wavelet for each scale to 1·0. Alternatively, each generated transform vector may be divided by its own standard deviation.

A common signal processing technique is to permit the wavelet to 'wrap around' to the beginning of $s(t)$ when the end is reached. For stream flow time-series, this can lead to erroneous transform magnitudes near the beginning and end of the record. A 'wrap around' algorithm permits events near the end of the time-series to be detected at the beginning of the transform magnitude vectors, and vice versa. For this reason, the convolution procedure used in this study follows the approach of Schiff (1992). Each wavelet is translated from one end of $s(t)$ to the other, starting and stopping within the confines of the data series.

It should be noted that hierarchical or *pyramidal* algorithms have been devised (e.g. Mallat, 1989) that are utilized for many orthogonal wavelets. These algorithms are superior to direct convolution in terms of computational efficiency and also form a complete and orthogonal basis function; thus no redundancy or artificial correlation between wavelet coefficients is introduced to the transformed data vector. However, these procedures do not generate a full suite of transform magnitudes at all wavelet scales and values of t.

Direct convolution of a non-orthogonal wavelet basis function may be used to provide finer sampling in the time-scale plane, but transformed data vectors should be tested for redundancy introduced by the convolution procedure (see Appendix B). The direct convolution algorithm used in this study was determined to be too slow to generate transform magnitudes at all time scales of t. Wavelets were therefore increased as powers of two. Direct convolution may be greatly expedited through multiplication in the Fourier domain, but this method was not used because of the prevalence of harmonic 'ringing' found in the Fourier spectra of analysed river flows (see section on analysis of stream flow records). For discussion of hierarchical algorithms that form a complete and orthogonal basis function (e.g. the Daubechies wavelets) the reader is referred to Daubechies (1988, 1990), Mallat (1989), Press *et al.* (1993), Newland (1993), Lee and Yamamoto (1994) and Bentley and McDonnell (1994).

WAVELET SCALOGRAMS OF DISCHARGE TIME-SERIES

Direct convolution of a suite of j scaled wavelets with a one-dimensional signal of length n produces a corresponding array of j transform magnitude vectors, also of length n. For daily discharge observations n is the record length in days. The width or time duration of the positive portion of the Mexican hat wavelet is approximately equal to the duration of an optimally detected flood event or other stream flow structure, a duration hereafter referred to as 'feature scale'. Since wavelet scale a is one-half of the width of the positive portion of the function $g(x) = (1 - x^2)\exp(-x^2/2)$, the feature scale is approximately $2a$ and reflects the time-scale of flood events a particular wavelet will optimally detect in the discharge record. While it is common practice to plot transform magnitudes as a function of wavelet scale a, conversion to feature scale $2a$ is convenient: for the Mexican hat wavelet, a four-day flood event will have a wavelet scale a of 2 days, but a feature scale of 4 days.

Arrangement of the j transform magnitude vectors on a common axis permits them to be observed as a function of feature scale and n, yielding a wavelet scalogram. The transform magnitude rises out of the page, and may be contoured to provide a plot of transform magnitude vs. time (or translation b) and feature scale. The choice of contour threshold enables features of differing magnitude to be revealed in the scalogram. A high contour threshold will permit only the dominant features in the signal to be observed; lower thresholds reveal more subtle features. The utility of wavelet scalograms for assessing time-dependent variations in river discharge are illustrated next, using two examples: (1) spring snowmelt events in a north-eastern US river; and (2) diurnal trends in a glacial meltwater stream.

Spring snowmelt in a north-eastern US river

Rivers in the north-eastern US tend to experience a significant increase in flow each spring from the melting of accumulated winter snow. However, this structure is commonly obscured by rainstorm events throughout the year, as illustrated by a seven-year subset of the daily discharge record for the Ammonoosuc River, Maine (Figure 2c). Transform magnitude vectors for wavelets of scale $a = \{1, 2, 4, 8, 16, 32, 64, 128\}$ days, corresponding to feature scales of approximately $\{2, 4, 8, 16, 32, 64, 128$ and $256\}$ days are shown in Figure 2a. These vectors are contoured with a single threshold of 5.5×10^{-6} m^3/s in Figure 2b: intensities above this level appear in black, lower values are not contoured. The wavelet scalogram of this subset seven-year discharge record reveals seven spring snowmelt events (which are of longer duration than rainstorms) as large blobs centred around feature scales of 64–128 days (Figure 2b). Note that time is preserved on the x-axis, feature scale (event duration) increases along the y-axis and transform intensity rises out of the page.

Numerous studies have utilized the scale and timing of snowmelt features to identify the hydrological regimes of northern watersheds (see Woo, 1990). Wavelet transformation of daily discharge time-series permits clear identification of diagnostic structures that may be obscure in the raw data, while wavelet spectral analysis (described in the section on analysis of stream flow records) provides easy assimilation and comparison of long records for numerous rivers.
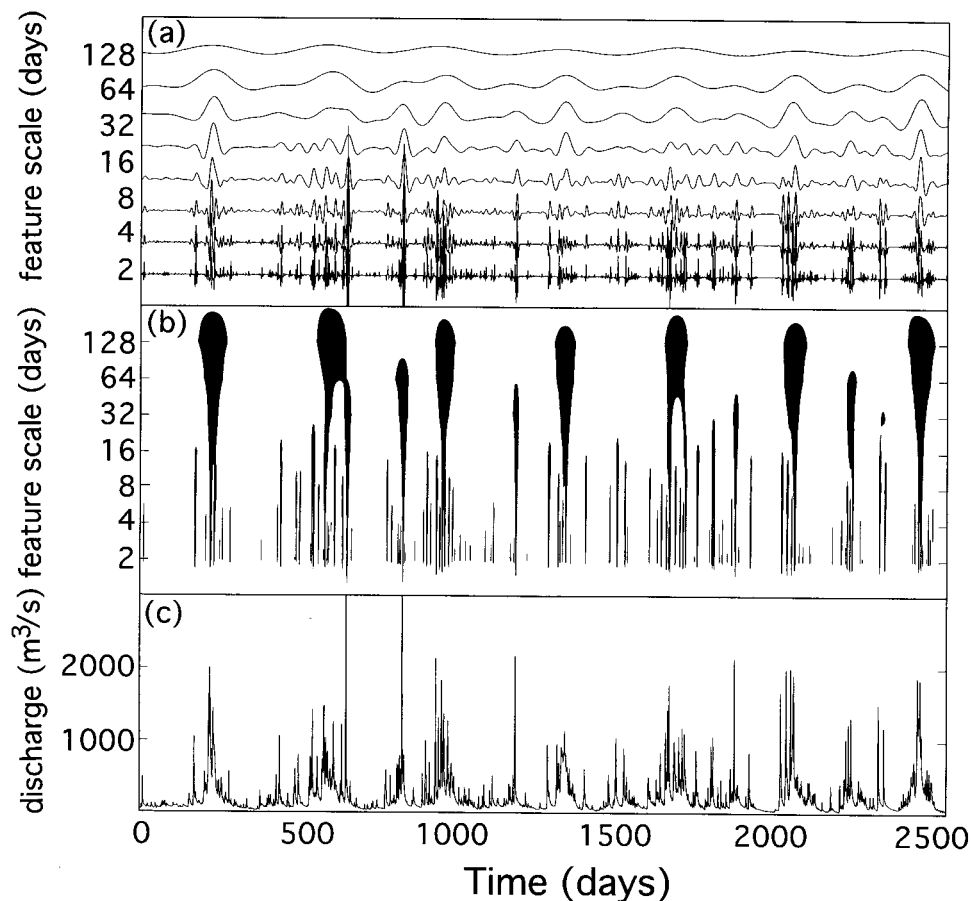
Figure 2. (a) Suite of normalized wavelet transform magnitude vectors generated from a seven-year daily discharge record, Ammonoosuc River, Maine. Time is preserved on the *x*-axis; feature time scale (2*a*) increases on the *y*-axis. (b) Wavelet scalogram contoured from (a). Feature scale and time are still preserved, but transform intensity rises out of the page. Application of an arbitrary contour threshold (5.5×10^{-6} m^3/s) reveals locations of annual spring snowmelt events as large blobs around feature time-scales of 64 and 128 days. (c) Raw daily discharge data used to produce (a) and (b)

Diurnal variability in a glacial stream

The 311 km^2 drainage area of Forrest Kerr Creek in north-western British Columbia contains several large glaciers. A 100-day subset (17 May to 25 August) of its 1992 hourly discharge record is plotted in Figure 3c. Transform magnitude vectors for wavelets with scale $a = \{1, 2, 4, 8, 16, 32, 64, 128\}$ hours, corresponding to feature scales of approximately $\{2, 4, 8, 16, 32, 128$ and $256\}$ hours, are plotted versus time in Figure 3a. The contoured scalogram is shown in Figure 3b. The wavelet transformation was carried out on a longer time-series, permitting transform magnitudes to be plotted up the edges of the 100-day subset.

The amplitude of diurnal rhythms in glacial streams is believed to be a function of the area of bare glacial ice and the development of internal drainage throughout the melt season (Young, 1990). The overall trend of increasing transform magnitudes at feature scales 8 and 16 hours in Figure 3a and b supports this explanation. For Forrest Kerr Creek, this general trend is interrupted irregularly by quiescent periods lasting a few days, followed by rapid recovery to previous wavelet magnitudes. These temporary effects are presumably independent of either bare ice area or glacier internal drainage development, and are most likely to be attributable to decreases in meltwater supply owing to cloudiness or cold weather. Such non-stationarity in the diurnal response has posed difficulties in interpreting standard Fourier power spectra of
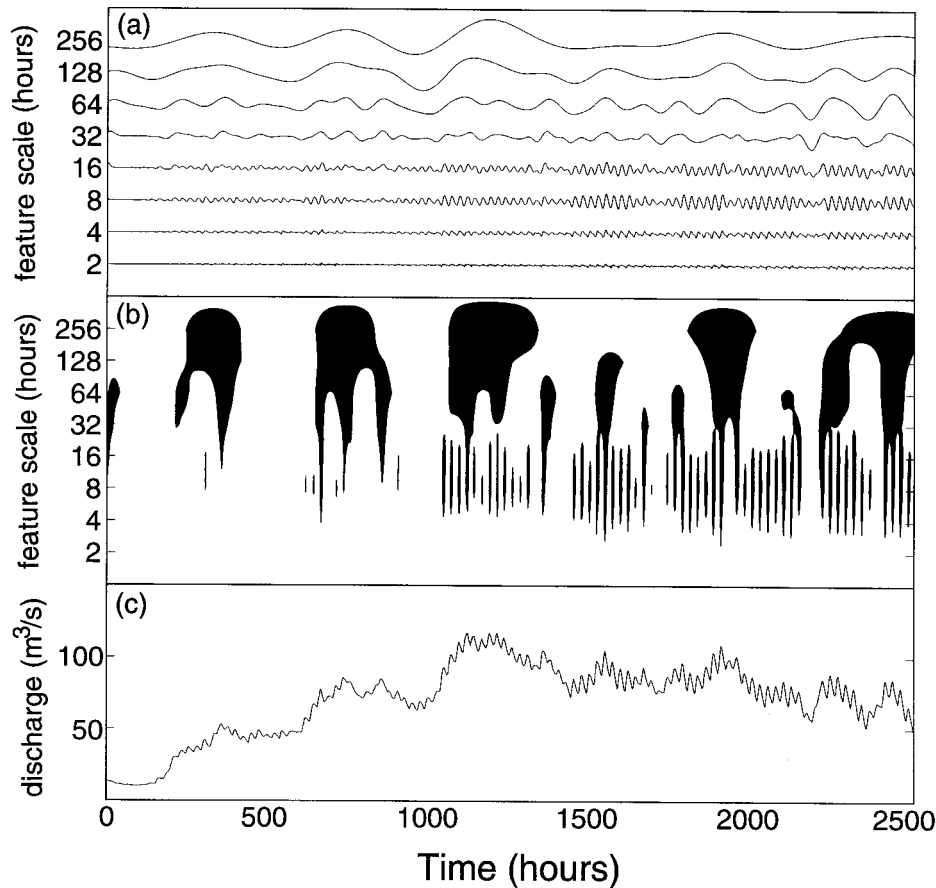
Figure 3. (a) Suite of normalized wavelet transform magnitude vectors generated from a 100-day hourly discharge record (May–August 1992) for a glacier-fed stream in British Columbia. (b) Wavelet scalogram contoured from (a). As the melt season progresses, diurnal structures intensify (at feature scales of 8 and 16 hours) with increasing glacier internal drainage development and surface area of bare ice (contour threshold = 9.0×10^{-6} m^3/s). (c) Raw hourly discharge data used to produce (a) and (b)

discharge time-series from glacier-fed streams (Gudmandsson and Sigbjarnarson, 1972). This problem is reduced through the use of wavelet scalograms.

WAVELET SPECTRAL ANALYSIS OF DAILY STREAM FLOW RECORDS

Integrating each transform magnitude vector over time to obtain the total energy at each wavelet scale a permits construction of spectral curves that reflect the distribution of event time-scales contained in the signal. Wavelet spectra are similar to Fourier spectra in that the temporal location of specific features is lost; the distribution of event scales is instead summarized for the entire input signal. Wavelet scalograms, which preserve the temporal locations of events, should be used to describe flow variability in non-stationary discharge hydrographs. For rivers that do not display long-term changes in stream flow structure, wavelet spectra are useful for summarizing a river's temporal variability and comparing it with flows in other rivers.

Normalized wavelet spectral curves for 91 rivers in the US and some tropical islands are presented in Figure 4. Twenty-year daily discharge records were arbitrarily selected from five geographic regions with significantly different climates. Records for non-regulated rivers complete between 1969–1988 were obtained from the US Geological Survey HCDN (Hydro-Climatic Data Network) stream flow dataset (Slack *et al.*,

1993). River names, station numbers, drainage basin areas and locations are listed in Table I. Each wavelet spectral curve in Figure 4 has been normalized by its total energy to facilitate comparison between rivers. A sample five-year daily discharge record from one river in each region is also shown in Figure 4, along with a complete Fourier power spectrum for one 20-year record from the region. The wavelet spectral shape is controlled primarily by the distribution of feature scales, and appears diagnostic of the hydroclimatic regime despite a large range in watershed sizes. A description and interpretation of the wavelet spectra are presented for each region in Table II. For the 'western snowmelt', 'mid-central humid' and 'rainstorm island' regions, wavelet spectra are sufficiently distinct to permit classification of the discharge records based on the wavelet spectral shape alone. Wavelet spectra from the 'north-eastern snowmelt' and 'south-western arid' regions are too similar to permit hydrograph classification using only spectral analysis. However, the temporal information contained in the wavelet scalograms (see previous section) permits the separation of these records. For example, the regular spring snowmelt event presented in Figure 2 is found in discharge records throughout the 'north-eastern snowmelt' region. This feature is notably lacking in discharge records from the 'south-western arid' region; the scalograms can thus be used to distinguish these two regions even though their wavelet spectra are similar.

The complete Fourier power spectra presented in Figure 4 also reflect differences in stream flow variability between the five climatic regions presented here, but the regional contrasts are quite subtle when compared with the wavelet spectra. For example, the greatest contrast in Fourier spectra is found between the 'western snowmelt' and 'rainstorm island' rivers: the Fourier spectral slope is steeper for the former. The wavelet spectra show a more striking difference between the two regions. Harmonic artefacts are also found in the Fourier spectra of stream flows containing a strong annual signal. Fourier power spectra for both snowmelt regions in Figure 4 display erroneous periodicities at 6, 3 and 1·5 months. These artifacts result from the transform's attempt to decompose the extremely sharp annual snowmelt event contained in hydrographs from these regions.

It should be stressed that the wavelet spectral curves in Figure 4 reflect the distribution of event time-scales contained in a stream flow record, while the Fourier spectra are most sensitive to periodicities. While it so happens that many of the diagnostic features are in fact periodic (e.g. annual snowmelt cycles and seasonal clustering of rainfall events), wavelet transform magnitude depends only upon the wavelet scale and local signal intensity. For example, in the 'south-western arid' region, the dispersion of wavelet spectral energy at all wavelet scales reflects true variability in stream flows for this region rather than the lack of a distinct periodic structure.

DISCUSSION

The results from the section on analysis of stream flow records demonstrate that a clear qualitative difference is found in the wavelet spectra of river flows from different climatic regimes. This is an important first step in developing a methodology for using wavelet analysis to identify, characterize and classify distinct hydroclimatic regimes of stream flow. However, to establish the wavelet transform as an effective tool for hydrograph analysis, more work is needed. It is not known which wavelet scales are most sensitive to climatic inputs. Optimal or idealized spectral curves need to be defined for each hydroclimatic category, by parameter fitting or constructing indices that utilize critical parts of the wavelet spectra. The amount of hydroclimatic variation required to create statistically significant changes in spectral shape must be determined, since this will define the limit to which the method may resolve differences between regions. While hydroclimatic classifications based on robust spectral signatures (e.g. the 'western snowmelt') will probably be most successful, the problem of distinguishing flows from different regimes that yield similar spectra (without *a priori* knowledge of location) still remains.

It is encouraging that the wavelet transform identifies some hydroclimatic regimes also detected with established methods. The present study considers stream flow variability at time-scales of 2–256 days. However, several of the hydroclimatic regions presented here are in general agreement with the findings of a
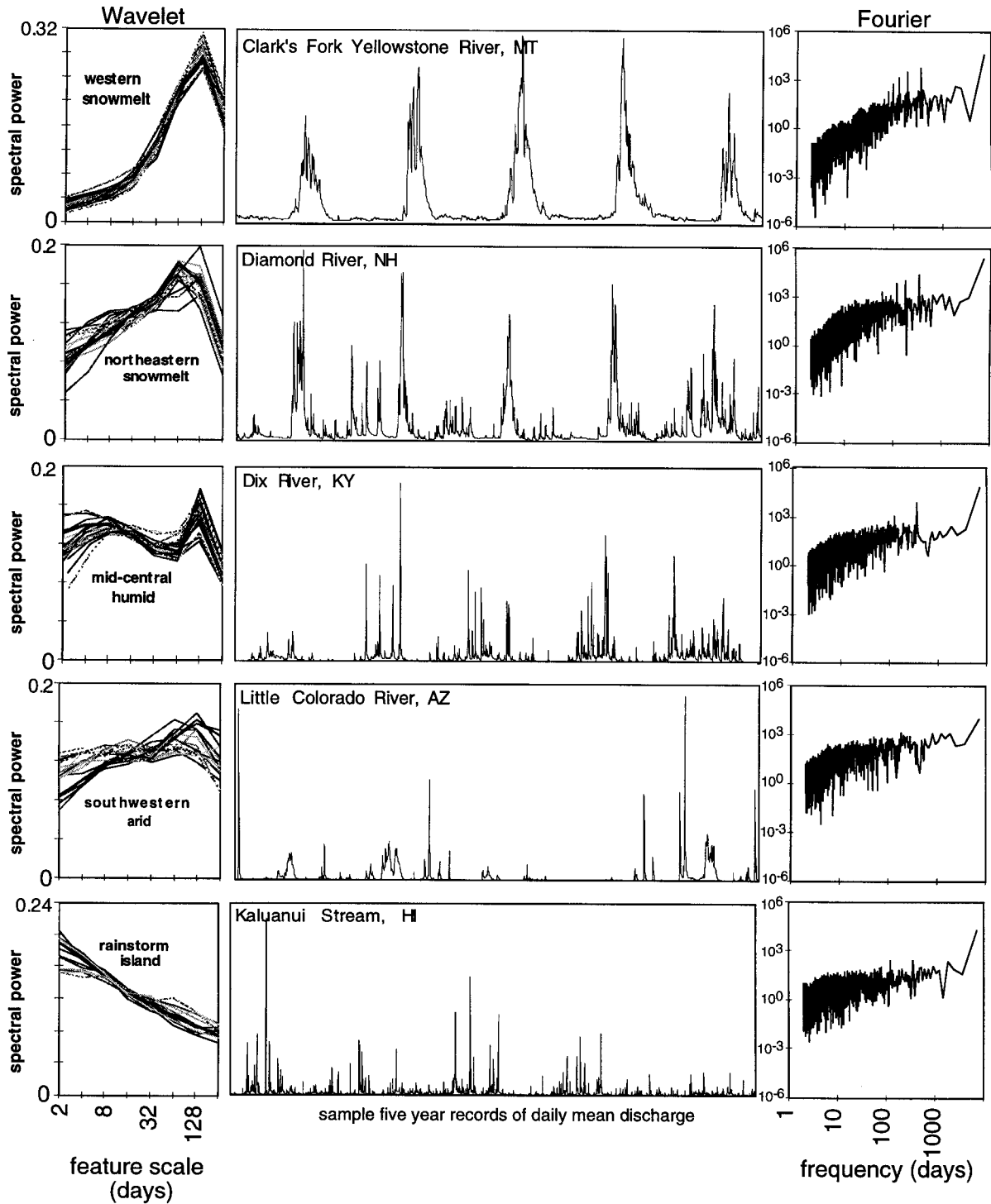
Figure 4. Wavelet and Fourier spectra for 20-year daily discharge records from 91 rivers in the US and on tropical islands. Normalized wavelet spectra for all rivers and complete Fourier power spectra for one river in each hydroclimatic region are shown. Wavelet 'feature scale' approximates the duration (in days) of random or periodic structures contained in the discharge record. A sample five-year daily discharge record is shown for each of the five regions examined in this study

Table 1. US Geological Survey Water Resources Division gauging stations used to produce wavelet spectra in Figure 4. River Drainage basin areas and locations are also shown

ID	Name and Location	Drainage Area		
		km^2	Latitude	Longitude
Southwest Arid:				
08053500	Denton Creek near Justin, TX	1,036	33:07:08N	097:17:25W
08080500	DMF Brazos River near Aspermont, TX	22,782	33:00:29N	100:10:49W
08080950	Duck Creek near Girard, TX	1,116	33:21:22N	100:42:17W
08082700	Millers Creek near Munday, TX	269	33:19:45N	099:27:53W
08083100	Clear Fork Brazos River near Roby, TX	591	32:47:15N	100:23:18W
08084800	California Creek near Stamford, TX	1,238	32:55:51N	099:38:32W
08086212	Hubbard Creek below Albany, TX	1,588	32:43:58N	099:08:25W
08129300	Spring Creek above Tankersley, TX	1,101	31:19:48N	100:38:24W
08144500	San Saba River at Menard, TX	2,940	30:55:08N	099:47:07W
09402000	Little Colorado River near Cameron, AZ	68,635	35:55:35N	111:34:00W
09430500	Gila River near Gila, NM	4,828	33:03:40N	108:32:12W
09430600	Mogollon Creek near Cliff, NM	179	33:10:00N	108:38:57W
09442692	Tularosa River above Aragon, NM	244	33:53:29N	108:30:54W
09444500	San Francisco River at Clifton, AZ	7,164	33:02:58N	109:17:43W
09471000	San Pedro River at Charleston, AZ	3,157	31:37:33N	110:10:26W
09480000	Santa Cruz River near Lochiel, AZ	212.9	31:21:19N	110:35:20W
09508300	Wet Bottom Creek near Childs, AZ	94.3	34:09:39N	111:41:32W
09508500	Verde River below Tangle Creek above Horseshoe Dam, AZ	15,208	34:04:23N	111:42:56W
Western snowmelt:				
06019500	Ruby River above Reservoir near Alder, MT	1,393	45:10:31N	112:08:52W
06191500	Yellowstone River at Corwin Springs, MT	6,794	45:06:43N	110:47:37W
06207500	Clarks Fork Yellowstone River near Belfry, MT	2,989	45:00:37N	109:03:53W
06218500	Wind River near Dubois, WY	601	43:34:43N	109:45:33W
06222700	Crow Creek near Tipperary, WY	78.2	43:34:37N	109:15:42W
06224000	Bull Lake Creek above Bull Lake, WY	484	43:10:37N	109:12:08W
06278300	Shell Creek above Shell Creek Reservoir, WY	59.8	44:30:29N	107:24:11W
06280300	South Fork Shoshone River near Valley, WY	769	44:12:30N	109:33:15W
06289000	Little Bighorn River at State Line near Wyola, MT	500	45:00:25N	107:36:52W
06311000	North Fork Powder River near Hazelton, WY	63.5	44:01:40N	107:04:49W
09241000	Elk River at Clark, CO	559	40:43:03N	106:54:55W
09245000	Elkhead Creek near Elkhead, CO	166.3	40:40:11N	107:17:04W
09404450	East Fork Virgin River near Glendale, UT	192.2	37:20:19N	112:36:13W
10015700	Sulphur Creek above Reservoir near Evanston, WY	166.3	41:07:45N	110:48:21W
10032000	Smiths Fork near Border, WY	427	42:17:36N	110:52:05W
10131000	Chalk Creek at Coalville, UT	648	40:55:14N	111:24:03W
10174500	Sevier River at Hatch, UT	881	37:39:04N	112:25:46W
10205030	Salina Creek near Emery, UT	134	38:54:43N	111:31:47W
10234500	Beaver River near Beaver, UT	235	38:16:50N	112:34:25W
12358500	Middle Fork Flathead River near West Glacier, MT	2,922	48:29:43N	114:00:33W
13120000	North Fork Big Lost River at Wild Horse near Chilly, ID	295	43:55:59N	114:06:47W
Northeast snowmelt:				
01022500	Narraguagus River at Cherryfield, ME	588	44:36:29N	067:56:10W
01031500	Piscataquis River near Dover-Foxcraft, ME	772	45:10:31N	069:18:55W
01038000	Sheepscot River at North Whitefield, ME	376	44:13:23N	069:35:38W
01047000	Carrabassett River near North Anson, ME	914	44:52:09N	069:57:20W
01052500	Diamond River near Wentworth Location, NH	394	44:52:40N	071:03:25W
01054200	Wild River at Gilead, ME	180	44:23:25N	070:58:47W
01055000	Swift River near Roxbury, ME	251	44:38:32N	070:35:17W
01055500	Nezinscot River at Turner Center, ME	438	44:16:10N	070:13:49W
01076500	Pemigewasset River at Plymouth, NH	1,611	43:45:33N	071:41:10W

Table continues on next page

Table 1. Continued

ID	Name and Location	Drainage Area km²	Latitude	Longitude
01064500	Saco River near Conway, NH	997	43:59:27N	071:05:29W
01078000	Smith River near Bristol, NH	222	43:34:04N	071:44:54W
01137500	Ammoonoosuc River at Bethlehem Junction, NH	226	44:16:08N	071:37:52W
01144000	White River at West Hartford, VT	1,787	43:42:51N	072:25:07W
01333500	Little Hoosic River at Petersburg, NY	145	42:45:50N	073:20:16W
01350000	Schoharie Creek at Prattsville, NY	614	42:19:10N	074:26:13W
04256000	Independence River at Donnattsburg, NY	230	43:44:50N	075:20:05W
04262500	West Branch Oswegatchie River near Harrisville, NY	632	44:11:08N	075:19:52W
04287000	Dog River at Northfield Falls, VT	197	44:10:58N	072:38:27W
04293500	Missisquoi River near East Berkshire, VT	1,240	44:57:30N	072:41:55W
04296000	Black River at Coventry, VT	316	44:52:08N	072:16:14W

Mid-central humid:

ID	Name and Location	Drainage Area km²	Latitude	Longitude
03080000	Laurel Hill Creek at Ursina, PA	313	39:49:13N	070:19:18W
03114500	Middle Island Creek at Little, WV	1,186	39:28:30N	080:59:50W
03179000	Bluestone River near Pipestem, WV	1,020	37:32:38N	081:00:38W
03180500	Greenbrier River at Durbin, WV	344	38:32:37N	079:50:00W
03186500	Williams River at Dyer, WV	332	38:22:44N	080:29:03W
03198500	Big Coal River at Ashford, WV	1,013	38:10:47N	081:42:42W
03217000	Tygarts Creek near Greenup, KY	627	38:33:51N	082:57:08W
03219500	Scioto River near Prospect, OH	1,469	40:25:10N	083:11:50W
03248500	Licking River near Salyersville, KY	363	37:45:03N	083:05:04W
03275000	Whitewater River near Alpine, IN	1,352	39:34:46N	085:09:29W
03281500	South Fork Kentucky River at Booneville, KY	1,870	37:28:45N	083:40:38W
03285000	Dix River near Danville, KY	824	37:38:31N	084:39:39W
03294000	Silver Creek near Sellersburg, IN	490	38:22:15N	085:43:35W
03298000	Floyds Fork at Fisherville, KY	357	38:11:18N	085:27:37W
03301500	Rolling Fork near Boston, KY	3,364	37:46:02N	085:42:14W
03416000	Wolf River near Byrdstown, TN	275	36:33:37N	085:04:23W
03434500	Harpeth River near Kingston Springs, TN	1,764	36:07:19N	087:05:56W
03436000	Sulphur Fork Red River near Adams, TN	482	36:30:55N	087:03:32W
03465500	Nolichucky River at Embreeville, TN	2,085	36:10:35N	082:27:27W
03528000	Clinch River above Tazewell, TN	3,818	36:25:30N	083:23:54W

Rainstorm island:

ID	Name and Location	Drainage Area km²	Latitude	Longitude
16108000	Wainiha River near Hanalei, Kauai, HI	26.4	22:08:20N	159:33:38W
16200000	North Fork of Kaukaonahua Stream near Wahiawa, Oahu, HI	3.57	21:31:09N	157:56:53W
16212800	Kipapa Stream near Wahiawa, Oahu, HI	11.1	21:28:13N	157:57:40W
16304200	Kaluanui Stream near Punaluu, Oahu, HI	2.87	21:35:22N	157:54:38W
16330000	Kamananui Stream at Maunawai, Oahu, HI	32.1	21:38:20N	158:03:27W
16345000	Opaeula Stream near Wahiawa, Oahu, HI	7.72	21:33:55N	158:00:10W
16400000	Halawa Stream near Halawa, Molokai, HI	11.9	21:09:31N	156:45:53W
16508000	Hanawi Stream near Nahiku, Maui, HI	9.04	20:48:37N	156:07:00W
16518000	West Wailuaiki Stream near Keanae, Maui, HI	9.48	20:49:16N	156:08:37W
16620000	Honokohau Stream near Honokohau, Maui, HI	10.6	20:57:48N	156:35:22W
16897900	Lewi River, Ponape, TT	1.20	06:55:32N	158:12:18E
50071000	Rio Fajardo near Fajardo, PR	38.6	18:17:56N	065:41:42W

study that applied eigenvector analysis to aggregated monthly river outflows measured at the perimeter of the conterminous United States (Guetter and Georgakakos, 1993). Eigenvector factor analysis (FA) of these monthly outflow anomalies aggregated within a three-degree latitude or longitude range identified eight zones by their interannual and intraseasonal variability. Several of these zones, called 'modes', are approximately analogous to the 'north-eastern snowmelt, 'south-western arid' and 'mid-central humid'

Table II. Description and interpretation of wavelet spectral curves for the five hydroclimatic regions shown in Figure 4

Region	Spectral characteristics	Interpretation
Western snowmelt	Consistent energy maxima at 128 day feature scale; very low energies at small wavelet scales	Snowmelt streams in Montana, Wyoming, Utah, Colorado and Idaho are dominated by a highly regular annual snowmelt cycle characterized by a high summer flow period approximately four months in duration. Low spectral energies at small wavelet scales reflect relatively minor contributions from rainstorm events.
North-eastern snowmelt	Maximum energies around 64 days; moderately high energies at small wavelet scales	Spring snowmelt features are consistently present but are of shorter duration than found in the western snowmelt region. Higher spectral energies at small wavelet scales reflect effects of rainstorm events during the summer and autumn.
Mid-central humid	High energies at small wavelet scales; sharp feature scale maxima at 128 days	Stream flows in this region are dominated by short-duration rainstorm events, which are randomly clustered around an annual precipitation maximum of the order of 4 months in duration. Spring snowmelt events are relatively minor, as reflected by low spectral energies at the 32 and 64 day feature scales.
South-western arid	Spectral energies dispersed at all scales; some maxima at 64 and 128 days	Stream flow records in New Mexico, Arizona and Texas contain irregularly distributed spiky flood events, sustained flows of varying duration and intervals of low flow. Wavelet spectra among rivers in this region display the least coherence of the five regions presented here.
Rainstorm island	Highest spectral energies at the smallest wavelet scales	Stream flows in Hawaii and two tropical islands consist almost entirely of brief rainstorm events, with little seasonal clustering. Most of the spectral energy is contributed at feature scales of 2 to 8 days in duration.

hydroclimatic regions identified here from wavelet analysis of daily records. However, the 'western snowmelt' and part of the 'mid-central humid' regions described in the present study are not distinguished in the eigenvector analysis. This may result from the fact that the input aggregated monthly stream flows were only obtained at or near the US border, in order to characterize total outflows from the conterminous US without redundancy. More work is needed to determine the similarities and differences in results between eigenvector and wavelet analysis of hydrological time-series.

The five hydroclimatic regimes presented here are by no means comprehensive; others are certain to exist within the US and elsewhere. Construction of wavelet spectral curves from daily stream flow records in other countries and parts of the US is required to ascertain whether global hydrograph data can be grouped into hydroclimatic classifications using wavelet spectral and scalogram analyses. Classification of global stream flow records using such an approach would permit mapping of stream flow hydroclimatic regime based on hydrograph properties rather than *a priori* assumptions about geographical location.

Several applications in hydrology would benefit from such a uniform global mapping of river hydrological regime. Attempts to model the effects of changing climate on stream flow require knowledge of the flow characteristics of the river in question. For example, it is clear from Figure 4 that some watersheds are sensitive to temperature while others are dominated by the amount and timing of liquid precipitation. Such constraints are well known where climate change modelling is carried out using a runoff model calibrated for a single watershed. However, assumptions must be made about the hydroclimatic regime of a large number of unknown rivers in order to model the effects of climate change on stream flow at the continental or global scale.

A better understanding of stream flow properties around the world may also aid interpretation of a global synchronicity in flood responses recently detected by Burn and Arnell (1993). Their statistical analysis of annual discharge maxima for 200 rivers found spatial and temporal patterns in flooding that are not well understood, although a weak connection with ENSO events was found. It is conceivable that such patterns may be correlated with river hydroclimatic regime. Again using Figure 4 as an example, increased summer cyclonic storm activity would presumably exert more influence in the rivers of the 'mid-central humid' type than the 'western snowmelt' type.

Erosion models would also benefit from a quantitative breakdown of the temporal components of stream flow, particularly for watersheds where sediment is supply limited. Water can pulse through a channel in days, weeks or months, and may follow interannual patterns as well; wavelet decomposition of the discharge record reveals the proportions of these different time-scales, and is limited only by record length and computational capacity. Wavelet spectral analyses may also be useful in determining optimal sampling strategies for monitoring river discharge. Measurements should be made more frequently for rivers whose wavelet spectra display high energies at small wavelet scales rather than large wavelet scales. This is not an issue where ground stations are currently providing hourly or daily discharge measurements. However, as new techniques emerge for estimating river stage and discharge from space (Koblinsky *et al.*, 1993; Brakenridge *et al.*, 1994; Smith *et al.*, 1996), optimal satellite return-time requirements will need to be determined for different hydroclimatic regions since high-resolution satellites rarely provide daily coverage.

CONCLUSIONS

The wavelet transform is a powerful tool for hydrograph analysis, both for identifying transient features and quantifying the temporal variability of stream flow. Wavelet scalograms permit precise location of both stochastic and periodic events in time and may reveal subtle structures not easily seen in the raw discharge data. Integration of scalograms over time permits construction of wavelet spectral curves that reflect the distribution of event time-scales and are diagnostic of hydroclimatic regime. Wavelet spectral curves for 91 rivers from five different climatic regions in the US and on tropical islands are strongly similar within the same region, but differ between regions.

More work is needed to establish wavelet transformation as a viable and consistent tool for hydrological time-series analysis. Remaining issues include determining which wavelet scales are most sensitive to climatic signatures, how to quantify the relationship between spectral shape and hydroclimatic regime and how the choice of wavelet and sampling rate affects results. However, our first exploration of the topic suggests that river flows may be effectively classified by hydroclimatic regime from the shape of their spectra, or by combining spectral curves with structural information from wavelet scalograms.

REFERENCES

Bentley, P. M. and McDonnell, J. T. E. 1994. 'Wavelet transforms: an introduction', *Electron. Commun. Eng. J.*, **6**(4), 175–186.
Brakenridge, G. R., Knox, J. C., Paylor, E. D., and Magilligan, F. J. 1994. 'Radar remote sensing aids study of the great flood of 1993', *Eos Trans., AGU*, **75**, 521–527.
Burn, D. H. and Arnell, N. W. 1993. 'Synchronicity in global flood responses', *J. Hydrol.*, **144**, 381–404.
Cayan, D. R., Riddle, L. G., and Aguado, E. 1993. 'The influence of precipitation and temperature on seasonal streamflow in California', *Wat. Resour. Res.*, **29**, 1127–1140.

Cazenave, A., Parsons, B., and Calcagno, P. 1995. 'Geoid lineations of 1000 km wavelength over the central pacific', *Geophys. Res. Lett.*, **22**, 97–100.

Chao, B. F. and Naito, I. 1995. 'Wavelet analysis provides a new tool for studying Earth's rotation', *Eos Trans., AGU*, **76**, 161, 164–165.

Daubechies, I. 1988. 'Orthonormal bases of compactly supported wavelets', *Commun. Pure Appl. Math.*, **41**, 909–996.

Daubechies, I. 1990. 'The wavelet transform, time-frequency localization and signal analysis', *IEEE Trans. Info. Theory*, **36**, 961–1005.

Dracup, J. A. and Kahya, E. 1994. 'The relationships between US streamflow and La Niña events', *Wat. Resour. Res.*, **30**, 2133–2141.

Entekhabi, D. 1984. 'Spectral analysis and stochastic modeling of streamflow using the Fourier series', *Eos Trans., AGU*, **65**, 885.

Gabor, D. 1946. 'Theory of communication', *J. Inst. Electr. Eng. (London)*, **93**(III), 429–457.

Gambis, D. 1992. 'Wavelet transform analysis of the length of the day and the El-Nino/Southern Oscillation variations at intraseasonal and interannual time scales', *Ann. Geophysic.*, **10**, 429–437.

Grossman, A. and Morlet, J. 1984. 'Decomposition of Hardy functions into square integrable wavelets of constant shape', *SIAM J. Math. Anal.*, **15**, 723–736.

Gudmandsson, G. 1970. 'Short term variations of a glacier-fed river', *Tellus*, **XXII**(3), 341–353.

Gudmandsson, G. and Sigbjarnarson, G. 1972. 'Analysis of glacier run-off and meteorological observations', *J. Glaciol.*, **11**, 303–318.

Guetter, A. K. and Georgakakos, K. P. 1993. 'River outflow of the conterminous United States, 1939–1988', *Bull. Am. Meteor. Soc.*, **74**, 1873–1891.

Hagelberg, C. R. and Gamage, N. K. 1994. 'Structure-preserving wavelet decompositions of intermittent turbulence', *Boundary-Layer Meteorol.*, **70**, 217–246.

Hameed, S. 1984. 'Fourier analysis of Nile flood levels', *Geophys. Res. Lett.*, **11**, 843–845.

Hayashi, T. 1994. 'An analysis of wind velocity fluctuations in the atmospheric surface layer using an orthonormal wavelet transform', *Boundary-Layer Meteorol.*, **70**, 307–326.

Kahya, E. and Dracup, J. A. 1993. 'US Streamflow patterns in relation to the El Niño/Southern Oscillation', *Wat. Resour. Res.*, **29**, 2491–2503.

Katul, G. G. and Parlange, M. B. 1995. 'Analysis of land surface heat fluxes using the orthonormal wavelet approach', *Wat. Resour. Res.*, **31**, 2743–2749.

Kiang, R. K., Kyle, H. L., Telfer, B. A., and Szu, H. H. 1994. 'Analyses of long-term solar irradiance data with wavelet transforms', *Wavelet Applications, Proc. SPIE-The Int. Soc. for Optical Eng.*, **2242**, 454–461.

Koblinsky, C. J., Clarke, R. T., Brenner, A. C., and Frey, H. 1993. 'Measurement of river level variations with satellite altimetry', *Wat. Resour. Res.*, **29**, 1839–1848.

Kumar, P. and Foufoula-Georgiou, E. 1993. 'A multicomponent decomposition of spatial rainfall fields 1. Segregation of large- and small-scale features using wavelet transforms', *Wat. Resour. Res.*, **29**, 2515–2532.

Kunhel, I., McMahon, T. A., Finlayson, B. L., Haines, A., Whetton, P. H., and Gibson, T. T. 1990. 'Climatic influences on streamflow variability: a comparison between southeastern Australia and southeastern United States of America', *Wat. Resour. Res.*, **26**, 2483–2496.

Lee, D. T. L. and Yamamoto, A. 1994. 'Wavelet analysis: theory and applications', *Hewlett-Packard J.*, **45**, 44–54.

Lins, H. F. 1985. 'Interannual streamflow variability in the United States based on principal components', *Wat. Resour. Res.*, **21**, 691–701.

Little, S. A., Carter, P. H., and Smith, D. K. 1993. 'Wavelet analysis of a bathymetric profile reveals anomalous crust', *Geophys. Res. Lett.*, **20**, 1915–1918.

Mallat, S. 1989. 'A theory for multiresolution signal decomposition: the wavelet representation', *IEEE Trans. Pattern Anal. Mach. Intell.*, **31**, 674–693.

Meko, D. M. and Stockton, C. W. 1984. 'Secular variations in streamflow in the western United States', *J. Clim. Appl. Meteorol.*, **23**, 889–897.

Meyers, S. D. and O'Brien, J. 1994. 'Spatial and temporal 26-day SST variations in the equatorial Indian Ocean using wavelet analysis', *Geophys. Res. Lett.*, **21**, 777–780.

Newland, D. E. 1993. *An Introduction to Random Vibrations, Spectral and Wavelet Analysis*, 3rd edn. Longman Scientific and Technical/Wiley & Sons, New York. 477 pp.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P. 1993. *Numerical Recipes in FORTRAN*, 2nd edn. Cambridge University Press, Cambridge. 963 pp.

Redmond, K. T. and Koch, R. W. 1991. 'Surface climate and streamflow variability in the western United States and their relationship to large-scale circulation indices', *Wat. Resour. Res.*, **27**, 2381–2399.

Sato, K. and Yamada, M. 1994. 'Vertical structure of atmospheric gravity waves revealed by the wavelet analysis', *J. Geophys. Res.*, **99**, 20623–20631.

Schiff, S. J. 1992. 'Resolving time-series structure with a controlled wavelet transform', *Optical Eng.*, **31**, 2492–2495.

Simpson, H. J., Cane, M. A., Herczeg, A. L., Zebiak, S. E., and Simpson, J. H. 1993. 'Annual river discharge in southeastern Australia related to El Nino–Southern Oscillation forecasts of sea surface temperatures', *Wat. Resour. Res.*, **29**, 3671–3680.

Slack, J. R., Lumb, A. M., and Landwehr, J. M. 1993. 'Hydro-Climatic Data Network (HCDN) streamflow data set', *US Geol. Survey Water-Resour. Invest. Rep. 93-4076* (CD-ROM).

Smith, L. C., Isacks, B. L., Bloom, A. L., and Murray, A. B. 1996. 'Estimation of discharge from three braided rivers using synthetic aperture radar (SAR) satellite imagery: potential application to ungaged basins', *Wat. Resour. Res.*, **32**, 2021–2034.

Takeuchi, N., Narita, K., and Goto, Y. 1994. 'Wavelet analysis of meteorological variables under winter thunderclouds over the Japan sea', *J. Geophys. Res.*, **99**, 10751–10757.

Taswell, C., WavBox 2.0 (c) 1992, (MATLAB 4.1 m-files for discrete wavelet transforms using compact orthogonal and biorthoganal wavelets) available over the World-Wide Web at http://www.wavbox.com

Turner, B. J., Leclerc, M. Y., Gauthier, M., Morre, K. E., and Fitzjarrald, D. R. 1994. 'Identification of turbulence structures above a forest canopy using a wavelet transform', *J. Geophys. Res.*, **99**, 1919–1926.

Webb, R. H. and Betancourt, J. L. 1992. 'Climatic variability and flood frequency of the Santa Cruz River, Pima County, Arizona', *US Geol. Surv. Water-Supply Paper 2379*. 40 pp.

Woo, M. K. 1990. 'Permafrost hydrology', in Prowse, and Ommanney, (Eds), *Northern Hydrology: Canadian Perspectives*, Environment Canada Nat. Hydrol. Res. Inst. Science Report No. 1. pp. 63–76.

Young, G. J. 1990. 'Glacier hydrology', in Prowse, and Ommanney, (Eds), *Northern Hydrology: Canadian Perspectives*, Environment Canada Nat. Hydrol. Res. Inst. Science Report No. 1. pp. 135–162.

APPENDIX A: CHOICE OF WAVELET BASIS FUNCTION

A wavelet basis function that has good localization in the time domain will have poor localization in the frequency domain, and vice versa. Therefore, the choice of wavelet should be dictated by the nature of the application. Many hydrograph peaks can be approximated as a single-peaked event of varying duration. For this reason, single-peaked wavelets such as the Mexican hat and Haar provide a good approximation of flood events contained in daily discharge records. Figure A1 shows discrete wavelet approximations of a sample flood event from the Little Colorado River, Arizona, using 10 different wavelets at the scale $a = 2$. The Mexican hat, Morlet and Haar wavelets were implemented directly following the procedure described in the section on the wavelet transform. The Morlet wavelet is complex; the real part has been used as noted by Gambis (1992). The DAUB (Daubechies), Coiflet and DMP (Daubechies orthogonal minimum phase) wavelets were generated from the WavBox 2.0 software package (Taswell, 1992), available over the World-Wide Web. Time differences, in days, between the time of maximum discharge and the time of maximum wavelet transform magnitude were 0 days (perfect match) for the Mexican hat, Morlet, DAUB8, Coiflet 3 and Coiflet 5 wavelets. Transform magnitude maxima lagged the actual flood peak by one day for the Haar and DMP2 wavelets. The DAUB4, DAUB10 and DMP8 transform magnitude maxima all preceded the actual flood peak by 3 days. Inspection of Figure A1 suggests that the Mexican hat wavelet provides a reasonable approximation of a typical flood hydrograph, and the day of maximum flow is precisely preserved in time.

The shapes of wavelet spectral curves are also influenced by the choice of wavelet basis function. Haar, Morlet and Mexican hat wavelet spectra for the first river listed in each hydroclimatic subsection of Table I are shown in Figure A2. Spectral powers have been normalized as described in the section on analysis. Although the generated spectral curves are in general agreement, significant differences in spectral shape and maximum power are found. Haar and Mexican hat wavelet spectra are more similar to each other than to Morlet wavelet spectra. Of the three wavelets used here, the Mexican Hat generates spectra that best differentiate between hydroclimatic regions. A particularly important feature is its sensitivity to decreasing spectral energy at wavelet scale $a = 128$ days, which corresponds to a feature scale of 256 days for the Mexican hat. The sensitivity of the Mexican hat wavelet to the scale and timing of flood peaks in daily discharge time-series is likely to be a result of (1) its central maxima and (2) its low modulation, as compared with the (1) Haar and (2) Morlet wavelets, respectively.

APPENDIX B. TEST FOR SIGNAL DISTORTION BY THE WAVELET TRANSFORMATION

Transformation of a data vector with N points into a domain that requires N^2 coefficients for complete description can introduce a certain redundancy to the transformed data vector. In some cases this

Figure A1. Discrete wavelet approximations of a flood hydrograph from the Little Colorado River, Arizona using the (1) Mexican hat, (2) Morlet, (3) Haar, (4) DAUB4, (5) DAUB8, (6) DAUB 10, (7) Coiflet3, (8) Coiflet 5, (9) DMP2 and (10) DMP8 wavelets. Discharge and transform magnitudes have been arbitrarily scaled to improve their visibility

Figure A2. Comparison of Haar, Mexican hat and Morlet wavelet spectra for one 20-year daily discharge record from each hydroclimatic region presented in Figure 4: (a) western snowmelt, (b) north-eastern snowmelt, (c) mid-central humid, (d) south-western arid and (e) 'rainstorm island'
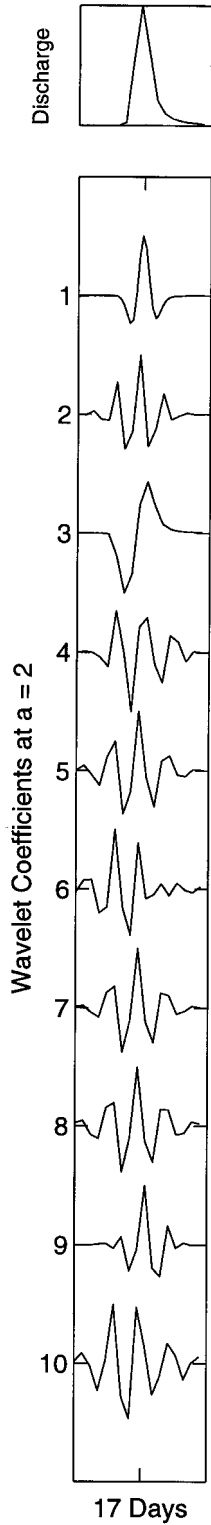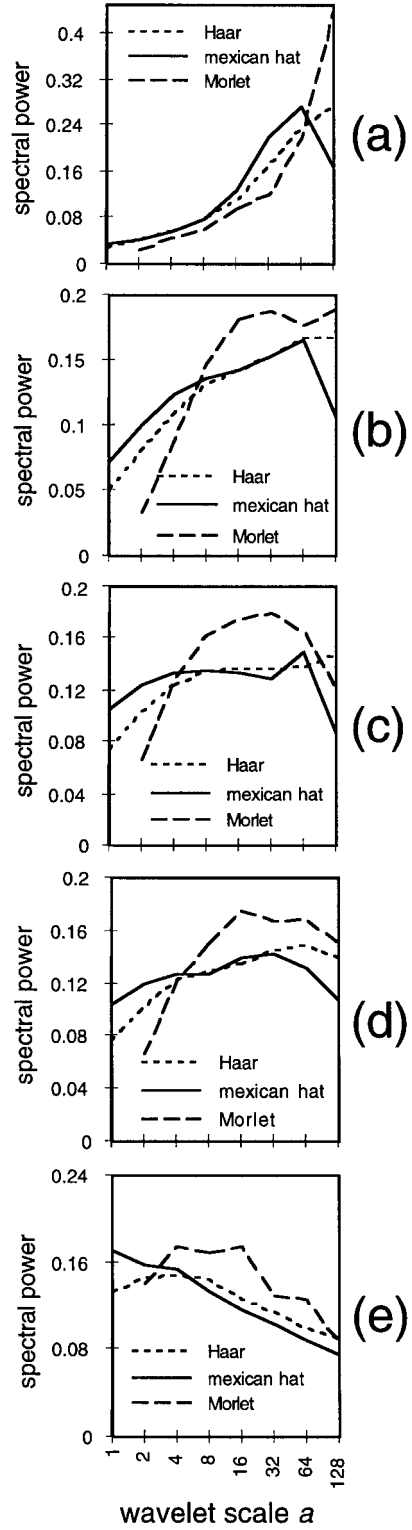
Figure A1. Caption opposite
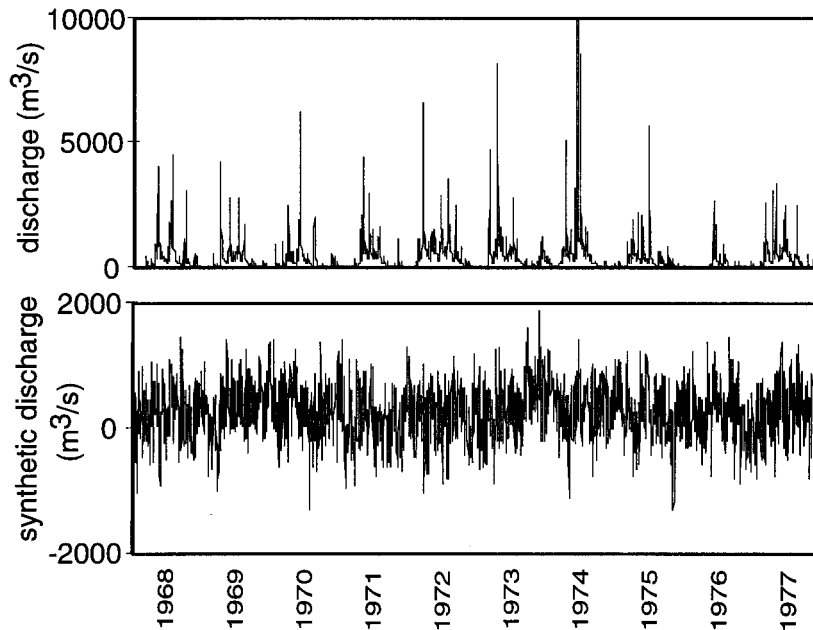
Figure A2. Caption opposite

Figure B1. Ten-year daily discharge record and its phase-randomized equivalent, Nolichucky River, Tennessee

redundancy may create artificial correlations between wavelet coefficients in the time–frequency domain, producing false detections and erroneous conclusions. The direct convolution procedure used in this paper was tested for such distortion using the following test.

Consider an input signal, here a 20-year daily discharge record $Q(t)$. The real and imaginary Fourier coefficients of $Q(t)$ are first determined using the Cooley–Tukey FFT algorithm. Randomization of the phase angle and inverse transformation of the modified real and imaginary Fourier coefficients produces a new time-series $Q_s(t)$, which has a Fourier power spectrum identical to that of the original time series $Q(t)$, but with a purely random phase (the phase preserves the locations of events). Figure B1 shows a 10-year daily discharge record [$Q(t)$] and its phase-randomized equivalent [$Q_s(t)$] from the Nolichucky River, Tennessee. The third- and fourth-order moments of skew and kurtosis may be used to establish the presence or absence of Gaussian statistics in wavelet transform coefficients (Katul and Parlange, 1995). From the central limit theorem, wavelet coefficients of $Q_s(t)$ should be normally distributed unless a distortion has been introduced by the wavelet transformation. By definition, the skew and kurtosis of a normal distribution are 0 and 3, respectively. If the wavelet transformation has distorted the phase-randomized signal, then its skew and kurtosis will not be 0 and 3. Table BI presents skew and kurtosis values for the wavelet coefficients of both

Table BI. Skew and kurtosis values for the wavelet coefficients of original [$Q(t)$] and phase-randomized [$Q_s(t)$] 20-year daily discharge record for the Nolichucky River, Tennessee

Wavelet scale, a	Skew, $Q(t)$	Kurt, $Q(t)$	Skew, $Q_s(t)$	Kurt, $Q_s(t)$
1	5.14	92.66	0.03	3.09
2	4.61	63.49	0.03	2.97
3	2.99	32.04	−0.06	3.12
4	2.21	17.09	−0.01	3.06
5	1.67	7.24	0.01	2.89
6	1.25	3.09	−0.16	2.74
7	0.37	0.25	−0.14	2.94
8	0.49	0.54	0.13	2.64

the original $[Q(t)]$, and phase-randomized $[Q_s(t)]$ time-series for the Nolichucky River. Values of skew and kurtosis are near 0 and 3 at all wavelet scales, providing one indication that the direct convolution of the Mexican hat function $g(x) = (1 - x^2)\exp(-x^2/2)$ with daily hydrograph time-series does not introduce significant distortion to the transformed data vectors. This finding, combined with results from Appendix A, suggests that the Mexican hat is an appropriate choice of wavelet basis function for time-series analysis of daily discharge records.