Wavelet Techniques for the Analysis and Synthesis of Rainfall Data

Andrew M.C. Chan

This chapter presents a relatively new time series analysis tool called wavelet analysis. The limited availability of fine time step rainfall data for use in long-term continuous modeling could be alleviated by synthetically generating rainfall data of fine time step increments from temporally coarse rainfall data. This study explores the spectral behaviour of rainfall of various temporal resolutions for dominant periodicities and presents a simple method of generating credible rainfall data that combines the approaches of stochastic modeling with a disaggregation goal. In other words, the goal is to get better temporal resolution rainfall data from existing data. Large- and small-scale periodic components are identified in the daily, hourly, and tipping bucket time-between-tips data for the City of Edmonton. A method of generating the desired rainfall data series is explained using the spectral behaviour of the rainfall data available. The desired rainfall data series would retain the record length of the most temporally coarse data (typically the longest period of record) and would have the desired fine time steps. The large-scale periodic components were extracted from three years of daily rainfall data and the small-scale components were extracted from one year of tipping bucket time-between-tips rainfall data for use in generating the synthetic rainfall data. The total rainfall volumes of the generated data compared well with observed values but tended to produce lower rainfall intensities and longer rainfall durations per event.

8.1 Introduction

Rainfall data is the fundamental input for almost all hydrologic models and has traditionally been acquired from artificial design storms, historical records, or synthetically generated data from mathematical models. Design storms (Keifer and Chu, 1957), such as those derived from intensity-duration-frequency (I-D-F) curves and tables, are still widely used in engineering design offices and water resource management institutions due to its simplicity and hence economic (cheap) design. Design storms, and its associated event-based methodology, were developed before the advent of computers to aid in computational and scientific expediency at a time when the availability and quality of historical rainfall data was poor (James, 1994). They are still being used despite the advances in computing ability, increases in the period of record of historical rainfall data, and our growing understanding of climatological processes.

In the realm of stormwater design, practitioners often use event hydrology to size pipes, channels, and other conveyance structures to some level of service such as the 12 h, 1 in 5 y storm. This use of event hydrology has been extended to the design of storage facilities such as dry and wet ponds, many of which are now also being designed to improve stormwater quality. The main drawback of using event hydrology in modeling is that dry-weather (or continuous) processes are not modeled and that assumptions about the antecedent conditions must be made at start up (James, 1994; Ormsbee, 1989; and Robinson and James, 1981). Event hydrology cannot provide adequate results when used for water quality modeling, or sizing water storage elements. Dry-weather, or inter-event, processes are very important for determining contaminant build-up from streets and gutters, pollutant loading in watercourses, agri-chemical build-up, storage levels of reservoirs at the start of storms, and antecedent soil moisture conditions, and so on, all of which vary with interevent duration and the intensity of drying processes.

Continuous modeling has long been advocated as an alternative to the design storm approach and associated event-based hydrology, as antecedent conditions and dry-weather processes can be accounted for, thus minimising start-up error and residual error. Today, long-term continuous modeling is increasingly necessary to infer the evolution of water quality and of the responses of the aquatic environment to anthropogenic inputs so that more informed water resource decisions can be made to protect the water environment (James, 1994; Yevjevich, 1991). In order for any long-term continuous modeling exercise to be credible, good long-term rainfall data must be used.

“All surface water ... modeling is wrong if the rain is wrong.” (James, 1994). Long-term historical rainfall data, in the order of one hundred years, is readily available for most locations in North America from sources such as the Atmospheric Environmental Services (AES) in Canada or from the National
Climate Data Centre (NCDC) in the USA. However, the temporal resolution of this data (generally at daily increments) is too coarse for use in small watershed hydrology. For the kind of long-term continuous modeling that is prudent for water quality modeling, rainfall data should ideally have time-steps small enough to properly gage the response of a watershed and should have a record long enough to consider long-term environmental impacts (James, 1994; Ormsbee, 1989; and Gregory and James, 1996). Rainfall data with smaller time-steps are available but as the time-step resolution increases, the period of record generally decreases. Ten to thirty years of hourly data is often available for Canadian urban centres but this time-step is still too coarse for urbanised watersheds. Data of even smaller time-steps are available (15 min and tipping bucket data) but the typical period of record is often 10 y or less. The desirable time-step for urban hydrologic modeling would be small enough to model the catchment response properly and yet not be too small to require enormous data storage requirements. Seventy-five years, or more, of 6-min rainfall data could be set as the desirable time-step resolution and record length to properly compare design alternatives over the lifespan of a water management facility by providing sufficient modeling duration for consideration of long-term environmental impacts (James, 1994). Unfortunately, such data will likely not be available until at least the middle of the next century.

In the meantime, numerous mathematical models have been pursued and developed in an attempt to remedy the paucity of useable rainfall data. Rainfall models have been developed over the last 30 y under the assumption that rainfall acts as a stochastic process. This assumption is made due to our limited knowledge of rainfall producing processes and the unforeseen factors that affect the apparent randomness of rainfall in both space and time. Until the actual physical processes that affect the occurrence, distribution, and variability of rainfall are fully known and can be modeled, rainfall can be said to be a stochastic or chaotic process.

Stochastic rainfall models take statistical properties of historical rainfall data and probabilities to generate long-term rainfall data records. Most stochastic rainfall models attempt to model the occurrence and duration of both wet and dry events (intermittency) coupled with another algorithm to determine rainfall amounts for the wet events. These increasingly complex models do not, however, provide any explanations relating to processes that drive the rainfall phenomenon (Yevjevich, 1991). Rainfall is a complex atmospheric process that is affected by incident solar radiation, lunar cycles, wind circulation patterns, global ocean current patterns, and so on. Models based on probabilistic outcomes and parameter fitting without any correlation to physical processes have been described as insufficient and dishonest (Klemes, 1986).

Some deterministic or physically-based rainfall models have been developed that attempt to predict precipitation by following the principles of conservation
of mass, momentum, and energy and incorporating meteorological information as input (Foufoula-Georgiou and Krajewski, 1995). However, because of the sophistication and complexity of atmospheric physics and the breadth of associated parameters, these physically-based precipitation models have been limited to mesoscale and general circulation models (GCM) and are of little use in small-watershed hydrology. In addition to the stochastic and deterministic approaches, rainfall modeling has been approached in the purely temporal domain, purely spatial, and combinations of both.

Rainfall disaggregation methods do not generate new rainfall data but rather, the methods attempt to enhance the temporal resolution of existing rainfall data. Disaggregation methods take existing rainfall data of a certain time step and period of record and breaks down the series into smaller time step increments while spanning the same period of record. Ormsbee (1989) proposed a method to generate sub-hourly rainfall from existing hourly rainfall by using Monte Carlo processes within the hour being examined. Unfortunately, disaggregation models do not generate new information and the intermittent behaviour of rainfall is not reflected through these methods.

There is also an interest in developing a general theory within the stochastic rainfall modeling approach to characterise the internal structure of rainfall. By examining the internal structure of rainfall, hydrologists attempt to describe the link between observed rainfall and the mechanisms responsible for its occurrence. Estimating the properties of the underlying processes that lead to an observed series is termed time series analysis (Bras and Rodriguez-Iturbe, 1985). Time series analysis models of rainfall attempt to analyse trends, periodicities, and frequencies in order to describe, explain, and simulate rainfall rates. Fourier methods have been used in the past in an attempt to describe, explain and predict some climatic signals such as rainfall. However, a Fourier analysis can only describe the dominant scale features that exist over the entire time scale of the signal. Any time-localised or non-stationary behaviour is masked by the Fourier analysis, which can only represent the global behaviour over the entire signal record. Wavelet analysis is a relatively new time-series analysis tool that is similar to a Fourier analysis but the temporal variations of the signal’s scale features are retained such that any time-localised behaviour can be represented locally.

The purpose of this chapter is to investigate wavelet analysis and to explore the frequency-analysing capabilities of wavelets on rainfall data. This study used wavelet analysis to examine point-rainfall data of various temporal resolutions at various periods of record and proposes a simple method to produce a synthetic series of 6-min point-rainfall data extracted from historic rainfall data sets of various temporal resolutions. The synthetic series has the same temporal resolution as the data set with the smallest time step and has a period of record equal to the longest data set record.
8.2 Description of Data

The rainfall data used in this study was recorded at the Edmonton Municipal Airport located within Edmonton, Alberta. The data includes daily, hourly, and tipping bucket (0.2mm, 0.08 in) formats collected over different years with a common overlapping period of 1980 through 1993 for all three data sets. All rainfall data were reported in mm depths.

The hourly and tipping bucket data were sampled only during the non-freezing months and thus did not constitute a continuous record of precipitation. The daily data was sampled continuously for twelve months of the year but only includes rainfall measurements. Data pre-processing was performed prior to any analysis in order to convert the raw data into a continuous time series format that could be processed by wavelet transforms. Missing or absent data were replaced by zero values to maintain a time series continuity. The tipping bucket data could be converted to any time-step resolution limited only by the size of the bucket. The tipping bucket data was converted from the time-between-tip (TBT) format to a continuous time series with time-step increments 6-min in order to give a fine time-scale that corresponds with the goal of generating 6-min time-step data of long record duration without requiring excessive data storage. General characteristics of the three rainfall data sets are listed in Table 8.1.

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Time Span</th>
<th>Max. Recorded Value (mm)</th>
<th>Annual Average or Annual Total (mm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Daily</td>
<td>Aug. 1, 1880</td>
<td>Dec. 31, 1996</td>
<td>114.0</td>
</tr>
<tr>
<td>Hourly</td>
<td>Apr. 1, 1960</td>
<td>Nov. 1, 1993</td>
<td>37.8</td>
</tr>
<tr>
<td>TBT</td>
<td>May 1, 1980</td>
<td>Oct. 29, 1980</td>
<td>5.6</td>
</tr>
<tr>
<td>(converted to 6-min time steps)</td>
<td>May 3, 1981</td>
<td>Oct. 31, 1981</td>
<td>7.6</td>
</tr>
<tr>
<td></td>
<td>May 18, 1982</td>
<td>Oct. 31, 1982</td>
<td>18.0</td>
</tr>
<tr>
<td></td>
<td>May 4, 1983</td>
<td>Oct. 25, 1983</td>
<td>13.0</td>
</tr>
<tr>
<td></td>
<td>Apr. 27, 1984</td>
<td>Oct. 29, 1984</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td>Apr. 28, 1985</td>
<td>Sept. 27, 1985</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>Apr. 15, 1986</td>
<td>Sept. 30, 1986</td>
<td>6.6</td>
</tr>
<tr>
<td></td>
<td>May 6, 1988</td>
<td>Nov. 4, 1988</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td>May 23, 1989</td>
<td>Oct. 26, 1989</td>
<td>11.0</td>
</tr>
<tr>
<td></td>
<td>May 6, 1992</td>
<td>Oct. 5, 1992</td>
<td>4.2</td>
</tr>
<tr>
<td></td>
<td>May 7, 1993</td>
<td>Nov. 3, 1993</td>
<td>7.8</td>
</tr>
</tbody>
</table>

Average = 280.57

* Denotes removal of year 1975 data from the calculation of annual average rainfall from hourly data
The daily rainfall data spans a total of 117y, commencing on July 11, 1880 and ceasing on December 31, 1996. The daily rainfall data set was missing 548 d out of the total 42,552 data elements. There was sparse data in the first three years of record (1880 to 1882, inclusive) for the daily data set. The hourly rainfall data set spans a total of 34 y of data starting on April 1, 1960 and ending on November 1, 1993. Due to the interruption of the winter months, approximately 40% of the year was unrecorded for the hourly and tipping bucket records. During 1975, hourly rainfall data was recorded but there appeared to be an error in recording as the rainfall was unusually low and sparse. That year was treated as an outlier, as the annual total rainfall for that year was well below the annual average by one order of magnitude. The tipping bucket data collected from the Edmonton Municipal Airport spans 15 y from May 1, 1980 to October 23, 1995 with missing data for the year of 1994.

8.3 Wavelet Analysis

Wavelet analysis is a relatively new time-series analysis tool that can be used to analyse multi-frequency, multi-scale signals that may exhibit non-stationary or non-sinusoidal behaviour (Bradshaw and McIntosh, 1994). In this chapter, the term frequency refers to the notion of the number of repetitions that a periodic process completes in a unit of time, also known as the inverse of a period. Accordingly, frequency analysis refers to the analysis of the frequency-domain component of a particular time series and it should not be confused with the calculation of a hydrologic occurrence for a particular return period (e.g., the 1-in-100 year flood).

Frequency-domain analysis (frequency analysis) is well developed in the study of electrical signals, voice recognition, and geophysics. Many authors have used frequency-domain methods to seek frequency components of various types of geophysical, oceanographic, and climate data records such as major river discharges, flood and drought indices, air and sea-surface temperatures, sea-level pressures, fish catches, tree-ring indices, and rainfall (Andel and Balek, 1971; Hartmann and Michelsen, 1989; Currie and O’Brien, 1990; Currie, 1993; Suppiah, 1993; and Yiou et al., 1996, among others). Many large temporal-scale climatic signals have been linked to astronomic influences such as sunspot activity and lunar and terrestrial orbital rotation (e.g., angular velocity, tilt, and eccentricity) (Lau and Weng, 1995; and Palutikof et al., 1999). These types of large-scale processes can influence the occurrence, distribution, and intensities of rainfall over large temporal and spatial scales (i.e., decades-long periods of increased rainfall variance over different geographical regions). Variances of rainfall at smaller temporal or spatial scales, as well as at the larger scales, can be detected and simulated through spectral analysis using wavelets.
8.3 Wavelet Analysis

Fourier methods have been used in the past to perform spectral analysis of time series signals. The biggest drawback to Fourier methods, and indeed an obstacle to the analysis of any time series, is its inability to properly analyse or represent non-stationary signals. Climate signals are generally non-stationary and rainfall in particular exhibits extreme variability and intermittency over many temporal scales. This time-localised behaviour must be examined along with its global time behaviour (Venugopal and Foufoula-Georgiou, 1996) since two climate signals can have very different localised information that when averaged over some time period (as is done with Fourier methods) can give very similar mean information (Lau and Weng, 1995). Fourier methods rely on basis functions (sine and cosines) that are uniform in scale and are unbounded in time such that any singularities in the time series will be represented or spread out over the entire spectrum (Yiou et al., 1996). Wavelet analysis methods, on the other hand, use basis functions that have compact support (are well localised) in both time and frequency while also being capable of analysing or decomposing several scales in the time series. The time-localisation property of wavelet methods tolerates local errors in the time series and ensures that any errors do not propagate through the wavelet spectrum as it would in a Fourier spectrum.

8.3.1 Introduction to Wavelets

Many families of wavelets exist in literature and a few are shown in Figure 8.1. Wavelet functions can be orthogonal or non-orthogonal, complex or real, narrow or wide, have many vanishing moments or oscillations like the Morlet wavelet, or can have sharp discontinuities like the ‘box-car’ shape of the Haar wavelet (Torrence and Compo, 1998). Each wavelet family has advantages and disadvantages in performing data analysis.

While a spectral analysis performed on a time series using the Fourier transform can give a representation of the signal’s frequencies and amplitudes over the entire signal, a wavelet transform of the same time series can give a representation of the signal’s frequencies and amplitudes that evolve over time. A wavelet transform is essentially a generalised form of the Fourier transform where the wavelet transform decomposes a signal into a series of daughter wavelets that are scaled and shifted versions of a mother wavelet. Not only does the wavelet transform shift or translate the mother wavelet, but it also dilates the mother wavelet in the frequency-domain for each analysing scale. This dilation in the frequency direction is accompanied by a contraction of the wavelet in the time-direction. This enables the frequency information to be localised in the time domain.

An example of the time-frequency localisation ability of wavelet analysis is shown in the contour plots in Figures 8.2 and 8.3 of two different combinations of regular sinusoidal signals (Sin 5x and Sin 20x). The wavelet power spectrum
Wavelet Techniques for Analysis and Synthesis of Rainfall Data

Figure 8.1 Sample of various wavelet families.

In the middle panel in Figure 8.2 confirms that two frequency components are present throughout the signal record by having two bright contour swathes. The lighter the contour shade, the higher the wavelet power (variance). The wavelet power spectrum in Figure 8.3 shows the abrupt change in frequency of the signal midway through the record. Interestingly, the global wavelet power spectrums (analogous to the Fourier power spectrum) of both signals are identical.

A wavelet transform is the integral convolution of a function, $h(t)$, with a family of functions, $\psi_{b,a}(t)$, derived from translations and dilations of an analyzing wavelet, or mother wavelet, $\psi(t)$. The family of functions, $\psi_{b,a}(t)$, are called daughter (or analyzing) wavelets, and is given by:

$$
\psi_{b,a}(t) = \frac{1}{a^{\frac{1}{2}}} \psi \left( \frac{t - b}{a} \right)
$$

(8.1)

where:

- $\psi$ is the mother wavelet
- $b$ is the position or translation parameter of the wavelet
- $a$ is the scale or dilation parameter of the wavelet ($a>0$)

When $a>1$, the wavelet is contracted in the frequency direction and dilated in the time direction. When $a<1$, the wavelet is dilated in the frequency direction and contracted in the time direction. These translations and dilations are performed
for each signal segment at each of the analysing scales. Wavelet coefficients are calculated to represent the correlation between the translated/dilated wavelet and the analysed signal segment. This is akin to Fourier series representation, where the wavelet basis function is used to represent the signal by superposition. However, unlike the Fourier series representation that uses only sine and cosine functions, wavelets use an infinite set of possible basis functions that are all dilations and contractions of the mother wavelet.

The Morlet wavelet, with a non-dimensional frequency ($\omega_0$) of 6, was used in this analysis due to its simplicity, its ability to convey information on the amplitude and phase of the data, its compact support in time, and its many oscillations that mimic the oscillatory behaviour of rainfall. The Morlet wavelet can be approximated with the non-dimensional time parameter, $\eta$, to give:

$$\psi(\eta) = \pi^{-1/4} e^{-i\omega_0 \eta} e^{-\eta^2/2}$$  \hspace{1cm} (8.2)

A continuous wavelet transform (WT) of the discrete rainfall data was used for this analysis. The continuous WT provides a smooth transition of wavelet power between the scales (or frequencies) being analysed (Torrence and Compo, 1998). The continuous WT for a discrete series can be approximated by:

$$Wh(b, a) = \sum_{b=0}^{N-1} \frac{1}{a^{\frac{N}{2}}} \psi^\ast \left( \frac{\eta - b}{a} \right) h(t) dt$$ \hspace{1cm} (8.3)

where: $\psi^\ast$ is the complex conjugate of the wavelet function $\psi$.
The rainfall data was padded with sufficient zeros at the ends in order to remove any edge effects (Torrence and Compo, 1998; Lau and Weng, 1995; and Meyers et al., 1993). In order to reduce the influence of edge effects at the ends of the data on the analysis, various authors suggest padding the ends of the data with sufficient zeros, such that the WT of this region does not influence the desired data. Typically the data can be padded with enough zeros to bring the total number of data elements to the next power of 2. This padded region was then discarded after the WT was performed (Meyers et al., 1993; and Torrence and Compo, 1998).

The results obtained by a wavelet transform are not as intuitive as the output from a Fourier transform. While the Fourier transform of a one-dimensional time series can be returned as a one-dimensional representation of the relative power of the frequency spectrum across the entire data record, the wavelet transform of a time series can be returned as a two-dimensional time-frequency image, commonly called the wavelet power spectrum (Torrence and Compo, 1998). The wavelet power spectrum is similar to the Fourier power spectrum but with an added time-dimension. In comparison, where the Fourier power spectrum represents the strengths of frequencies present in a signal for the entire time-domain, the wavelet power spectrum represents the relative strengths (variance) of the frequencies (scales) at each time-step.

When WTs are used for time series analysis the graphical representation is typically a contour plot showing the wavelet variances over various scales over the time span of the data. Unlike the FT that maps a one-dimensional time series (e.g. precipitation as a function of time) onto a one-dimensional Fourier spectrum (amplitude versus frequency), the WT maps a one-dimensional time series onto a two-dimensional image of time, scale, and variance. The ‘x’-axis is the time axis, the ‘y’-axis is the frequency axis (the analysing scale), and the ‘z’-axis (perpendicular to the two-dimensional plane) is the relative wavelet power axis. The scale axis displays contour intervals as a linear progression of shades from black to white with the higher values of wavelet power spectra represented as lighter shades of grey. Values on the contour scale bar represent the existence of wavelet power above each contour level. These contour threshold values are arbitrary and are chosen to give a meaningful contour image. Torrence and Compo (1998) introduced statistical significance testing to qualify the wavelet analysis results.

The significance of different peaks in the wavelet power spectrum are determined by using either a white noise or a red noise background spectrum to establish a null hypothesis. In this chapter, regions that exceed the 95% confidence levels corresponding to a red-noise spectrum are outlined by a thick black contour line on the wavelet power spectrum. Spectral peaks may also be significant in the absence of any dominant periodicities such as the annual and semi-annual periods. A spectral peak below the 95% confidence limit may not necessarily be regarded as noise.
8.3 Wavelet Analysis

A cone of influence (COI) is also specified as the region surrounded by the crosshatched area. Results within the crosshatched area are questionable as edge effects within this region become important. The plot of the global wavelet power spectrum is the wavelet power spectrum averaged over the time-span and is represented as a function of wavelet power variance versus Fourier period. The dotted line in the global wavelet power spectrum represents the 95% confidence limit with the red-noise spectrum as background.

The global wavelet power spectrum can be calculated to determine the time-averaged wavelet spectrum over a period of the entire time series (analogous to the Fourier power spectrum). The global wavelet power spectra, $W^2 h(b,a)$, over the entire times series, $h(t)$, is given by:

$$
\bar{W}^2 h(b,a) = \frac{1}{N} \sum_{b=0}^{N-1} \left| Wh(b,a) \right|^2
$$

(8.4)

and is akin to taking vertical slices of the wavelet power spectrum for each $b$ and time-averaging all the slices. This gives a one-dimensional plot of wavelet power versus frequency.


8.3.2 Current Rainfall Studies Using Wavelets

Research in the analysis of rainfall using wavelets is relatively new and hence there is little work available. Much more research exists in other studies such as seismology, oceanography, and climatology. (Lau and Weng, 1995; Meyers et al., 1993; Yiou et al., 1996; Jiang et al., 1997; and Torrence and Compo, 1998). Studies of rainfall using wavelet analysis have generally focused on the spatial variability and scaling properties of rainfall. Kumar and Foufoula-Georgiou (1993a and 1993b) examined the spatial variations of rainfall intensities of spatial radar data of a single summer squall storm during 1987 in Oklahoma. Large-scale fluctuations were examined for mean behaviour, while small-scale fluctuations were examined for self-similarity. The purpose of their study was to infer statistical properties of spatial rainfall at various scales so that the information can be used to characterise rainfall fields in mesoscale climate models and smaller scale rainfall models and to provide links between statistical and physically based models. More recently, Venugopal and Foufoula-Georgiou (1996) examined high-resolution temporal rainfall data (sampled at a rate of 5 to 10 s) of some select individual summer storms between 1989 and 1992 at the Iowa Institute of
Hydraulic Research. Venugopal and Foufoula-Georgiou (1996) found that for the summer storms they examined the active periodicities, or scales, were 20 to 160s, 200 to 640s, 4 to 5min, and 4 to 6h using wavelet packet transforms.

Kumar (1996) examined six storms during 1990 at the same recording station in Iowa used by Venugopal and Foufoula-Georgiou (1996) at the same sampling rate. Kumar (1996) tested the hypothesis of scale invariance of temporal rainfall for that data set and found that there were two or more active scales present. Small periods, in the order of 5 to 320 s were attributed to the scales of the individual rain cells. Larger periods, in the order of 2,560 to 5,120s (43 to 85 min), were attributed to extra-terrestrial influences (Kumar, 1996).

8.4 Results

The city of Edmonton rainfall data described earlier was analysed using wavelet analysis. This section contains plots of the wavelet power spectra of a sample of the rainfall data, their corresponding global wavelet power spectra, and a discussion of the spectral behaviour.

8.4.1 Daily Rainfall Data

Figure 8.4 shows a plot of the daily rainfall data for the Edmonton Municipal Airport for the period between July 1,1880, and December 31,1996 (top panel), the corresponding local wavelet power spectrum (middle panel), and its global wavelet spectrum (bottom panel).

Most obvious and intuitive, was the significantly high wavelet power through the one-year period for the entire length of the record. This is illustrated by the predominant contour band of white located at the one-year period on the local wavelet power spectrum (Figure 8.4, lower left panel). The one-year period also illustrated by the major peak in the global wavelet spectrum (Figure 8.4, lower right panel). Other notable periods of wavelet power are the half-year, 2- to 4-y, 5- to 7-y, 8- to 16-y, and 32- to 48-y components shown by light contours. The half-year period showed variable wavelet power for most of the record and low wavelet power during 1880-1886. The 2- to 4-y component showed an oscillating wavelet power with high power during 1885-1895, 1910-1915, 1932-1935, and 1987-1996. The 5- to 7-y component showed high wavelet power during 1900-1910 and 1940-1965 and a diminished wavelet power during 1890-1900, 1930-1945, and 1990-1996. The 8- to 16-y component showed a substantially higher wavelet power during 1880-1925, 1940-1960, and 1975-1994 and lower wavelet power during 1920-1940, and 1960-1975. The high wavelet power for the 8- to 16-y period during 1880-1925 is comparable to a similar high wavelet power for the 2-to 8-y period found during 1880-1920.
8.4 Results

Figure 8.4 Local and global wavelet power spectra of Edmonton daily rainfall (1880-1996).

in the annual Niño3 sea-surface temperature (SST) data reported by Torrence and Compo (1998). The Niño3 SST data, recorded between 1871 to 1996, was used as a surrogate for the ElNiño-Southern Oscillation (ENSO) index.

8.4.2 Hourly Rainfall Data

An analysis of the hourly rainfall showed similar dominant scales at the periods larger than a half-year to the analysis of the daily rainfall. For the sake of brevity, only the wavelet analysis results for rainfall data from 1983 to 1985 inclusive, is shown in Figure 8.5. An examination of all wavelet analysis results for the hourly data set revealed the following periods smaller than the annual periodic component: half-year, 91.3-d, 28.5- to 40-d, and 14.3- to 30-d periods. The half-year component showed significant power during 1965-1966 (inclusive), 1968-1969 (inclusive), 1970-1975 (inclusive), 1979-1980 (inclusive), 1982-1984 (inclusive), and 1987-1988 (inclusive). The 91.3-d (0.25-y) component generally followed the oscillation of the half-year component with the addition of high wavelet power during 1984, 1986, and 1991. The 28.5- to 40-d (0.08-to 0.10-y) periodic component showed high power during 1960-1961 (inclusive), 1963, 1965, 1969-1970 (inclusive), 1972-1974 (inclusive), 1976-1978 (inclusive), 1982-1983 (inclusive), 1987-1988 (inclusive), 1990-1991 (inclusive), and
Wavelet Techniques for Analysis and Synthesis of Rainfall Data

Figure 8.5 Local and global wavelet power spectra of Edmonton hourly rainfall (1983-1985 inclusive).

1993. The wavelet power for the 14.3- to 30-d periodic component generally followed the oscillation of the wavelet power for the 28.5-to 40-d component. A better determination of this behaviour can be discerned using finer increment data, such as with tipping bucket data.

8.4.3 TBT Rainfall Data

A sample of the wavelet analysis results for the converted tipping bucket data set is shown for 1983 in Figure 8.6.

The variations in rainfall storm density, intermittency, and intensities were much more evident using this high time-resolution data. There was less evidence of any kind of normal, or Gaussian, distribution for the TBT rainfall, as there was in the daily and hourly rainfall data. Individual storms may still exhibit a Gaussian shape depending on the storm type. The local wavelet power spectra for all analysed years show broad spectrum-peaks that spanned several scales. This indicated the presence of several high wavelet power variances that occurred at scales adjacent to one another to create a smoothed spectrum peak across a range of scales. Generally, most of the global wavelet power-spectral peaks for the TBT data analysed showed many wavelet power variances above 95% confidence-limit line for a red-noise spectrum.
There were several spectral peaks that exceeded the 95% confidence limit. The global wavelet power spectra provides the easiest method to discern these peaks from the broad peaks shown in the local wavelet power spectra. The most notable spectral peaks, smaller than 1024-h (42.7-d or 0.17-y), occurred approximately at the following periods: 768-h, 384-h, 96-to192-h, 64-to96-h, and 48-to 64-h. The 768-h period was significant in the years 1980, 1982-1986, and 1988 and low in the years 1981, 1987, and 1989. The 384-h period is significant in the years 1980, 1982, 1984, 1986, and 1988. This period is possibly showing a regular cycle of high and low variance for every two years. The 192-h period was significant in the years 1981, 1985-1986, and 1988, and low in the years 1980, 1982, 1987, and 1989. The 64-to 96-h period exhibited high global wavelet power variance in almost every analysed year with the exception of 1987, and 1989. The 48-to 64-h period was significant in 1980, 1981, 1985, 1986, 1987, and 1988. There were also smaller periods present that did not appear to be as significant. However, if the larger periodicities such as the half-year, 2048-h, and 1024-h were removed, the smaller scale components such as the 32-to 48-h, 24-h, 12-h, and 7-h periods would be more prominent. The 32-to48-h period showed higher wavelet power variance in 1980-1982, 1984-1987, and 1990. The 24-h period was higher in 1980-1982, 1984, 1985, 1987, 1988, and 1990. The 12-h period was higher in 1980-1981, 1983- 1984, and 1987. The 7-h period was high in 1981, 1985-1987, and 1990. It should be noted that scales
described in this section are likely to be influenced by local-level and small-scale climate processes that are much more volatile than the scales influenced by large-scale climate processes or extra-terrestrial periodic forcings. Periodicities smaller than a few hours are likely to exist but are difficult to discern even though the smallest resolvable period was approximately 12 min. These small-scale periodicities are likely to be caused by a myriad of hydrologic and atmospheric processes on a regional level.

8.5 Generation of a Sample Duration

Synthetically generated rainfall data should resemble historic rainfall for the same geographic/hydrologic region, share some similar statistics of a similar magnitude such as event volume, event duration, and interevent duration. In other words, the generated rainfall data should look and feel the same as actual rainfall in a hydrologic sense, but does not have to be exactly the same as actual rainfall that occurred for that period of record and location.

The large and small scale periodic components exhibited in the daily rainfall data were reconstructed using selected periodic components. A year of TBT rainfall data was used in conjunction with the reconstructed periodic components in order to incorporate the intermittancies and variability found at that time-scale. The result was the generation of three years of 6-min rainfall data derived from 1-y of TBT data and 3-y of daily rainfall data.

8.5.1 Reconstruction Technique

Reconstruction from a wavelet transform for a known mother wavelet usually involves the inverse wavelet transform or a deconvolution of the wavelet coefficients with the daughter wavelets at each scale increment. However, while the continuous wavelet transform is useful for analysing periodicities by developing smooth contours of the local wavelet transform, it accomplishes this at a cost: by performing redundant analysis over time and scale. This redundancy inhibits a straightforward reconstruction using the inverse wavelet transform. One method to perform a reconstruction of data from a continuous wavelet transform is presented by Torrence and Compo (1998). They proposed that for complex wavelets such as the Morlet wavelet, the real component, \( \Re\hat{W}(b,a) \), can be summed over all scales, \( a \), using the Dirac delta function as a basis function. To reconstruct the original \( h(t) \) time series, the formula is as follows:

\[
h(t) = \frac{d_j \sqrt{dt}}{C_\delta \psi_\delta(0)} \sum_{j=0}^j \Re\hat{W}(b,a) \sqrt{a_j}
\]  

(8.5)
8.5 Generation of a Sample Duration

where: $C_\delta$ is the reconstruction factor,

$$= 0.776 \text{ for the Morlet wavelet; and}$$

$\psi_0(0)$ is used to the remove the normalising constant,

$$= \pi^{1/2} \text{ for the Morlet wavelet.}$$

The reconstruction formula described above was applied to three years of daily rainfall data (1983 to 1985 inclusive) and one year of TBT data (1983). Three years of generated data was chosen to provide a manageable series that can be reasonably computed and reconstructed. The reconstructed data consisted of three years of 6-min interval rainfall data, giving a total of 262,980 data points.

The large-scale periodic components were extracted using the reconstruction formula. These periodic components were aggregated to get a large-scale periodic representation of the rainfall over the three year period. The 1983 tipping bucket rainfall data set was used to obtain representative small-scale characteristics for a typical year such as the intermittancy of rainfall, typical magnitudes of rainfall measurements over the time interval, and typical distributions of the rain-storm shape. The TBT data was then modulated by the large-scale components for each year to produce the reconstructed rainfall of 3-y of 6-min data (See Figure 8.7).

The reconstructed and actual observed rainfall volumes for 1983, 1984, and 1985 are given in Table 8.2. The measured TBT rainfall data is shown below in Figure 8.8.

---

**Figure 8.7** Reconstructed 6-min rainfall from TBT (6-min) and daily rainfall data (1983-1985).

![Reconstructed 6-Minute Rainfall from TBT (6-Minute) and Daily Data](image)

**Figure 8.8** Observed 6-min rainfall converted from TBT data (1983-1985).

![Actual TBT (6-minute) Rainfall (1983–1985)](image)
Table 8.2 Comparison of actual and reconstructed total annual rainfall volumes.

<table>
<thead>
<tr>
<th>Year</th>
<th>Total Annual Rainfall Volumes, mm (in)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual</td>
</tr>
<tr>
<td>1983</td>
<td>400.0 mm (15.7 in)</td>
</tr>
<tr>
<td>1984</td>
<td>288.0 mm (11.3 in)</td>
</tr>
<tr>
<td>1985</td>
<td>266.4 mm (10.5 in)</td>
</tr>
<tr>
<td></td>
<td>Reconstructed</td>
</tr>
<tr>
<td>1983</td>
<td>451.8 mm (17.8 in)</td>
</tr>
<tr>
<td>1984</td>
<td>331.5 mm (13 in)</td>
</tr>
<tr>
<td>1985</td>
<td>261.7 mm (10.3 in)</td>
</tr>
</tbody>
</table>

In addition to comparing total annual volumes, the reconstructed rainfall and observed rainfall series were processed by the RAIN and STATISTICS blocks of XP-SWMM to compare their statistical behaviours. Both series were analysed using a minimum interevent time duration of 1 hour and an event threshold value of 10 mm.

Figures 8.9a through 8.13b show comparisons of generated (reconstructed) and recorded rainfall volume, average intensity, peak intensity, event duration, and interevent duration versus return period (in log base 10 months) and event frequency.

![Figure 8.9](image_url)  
(a) Rainfall volume vs. return period and rainfall volume vs. frequency.

![Figure 8.10](image_url)  
(a) Rainfall average intensity vs. return period and rainfall average intensity vs. frequency.
8.5 Generation of a Sample Duration

Figure 8.11 Rainfall peak intensity vs. return period and rainfall peak intensity vs. frequency.

Figure 8.12 Rainfall event duration vs. return period and rainfall event duration vs. frequency.

Figure 8.13 Rainfall interevent duration vs. return period and rainfall interevent duration vs. frequency.

percentage of occurrence less than or equal to the given magnitude. While the reconstructed and actual total rainfall volumes compare well, the reconstructed rainfall series generally produced lower overall rainfall intensities spread over longer durations, and that interevent durations between rainfall events were significantly shorter.
The reconstructed rainfall volumes closely resembled the actual rainfall volumes across all events. For smaller storm events (less than 1 month return period), the average and peak intensities of the reconstructed and actual rainfall matched closely. Larger storm events exhibited lower reconstructed average and peak intensities compared with the actual values. The reconstructed rainfall event durations were higher than the recorded rainfall event durations for storms with return periods greater than about 0.25 months. Conversely, the interevent durations of the actual rainfall record exceeded the interevent durations of the reconstructed rainfall. Variation of the minimum interevent time and event threshold value did not affect results significantly.

Future refinements of this technique will (hopefully) improve similarities of storm event/interevent and rainfall intensity distributions between generated and observed series. The current disparities, despite the relatively good match in rainfall volumes, would create very different runoff responses regardless of the subcatchment characteristics. This may become very important depending on the goal(s) of the hydrologic exercise.

8.6 Conclusions

The wavelet analysis of the City of Edmonton rainfall data provided some valuable insights into the rainfall processes and patterns. This relatively new time series analysis tool allowed the examination of spectral behaviour, present in the rainfall data series on a time-localisation basis. Wavelet analysis as a time series analysis tool is superior to the more traditional methods such as Fourier analysis as not only are frequencies retained and localised in time, but that the non-stationary and variability of rainfall is preserved and depicted in the wavelet power spectrum.

Various large-scale and small-scale spectral components were revealed from the three separate rainfall data types. The periods identified ranged from a few hours in the case of the tipping bucket time-between-tip (TBT) data to several years in the case of the daily data. These periods showed variable and oscillating variances over time, which may be linked to large-scale global climatic inputs and extra-terrestrial periodic forcings to smaller regional influences such as the El Niño phenomenon. Linking the periodic components of the observed rainfall to observed climatic signals or to extra-terrestrial inputs could be accomplished by performing a wavelet cross-spectra of the two signals. One such issue that might be worth exploring is the impact of local climate change and variability as a result of increased development in urban centres thereby increasing the areal coverage and intensity of convective heat islands and thus intensifying summer convective storms.

A data reconstruction technique was described and applied to 3-y of daily data and 1-y of TBT data to generate 3-y of 6-min rainfall data. Large-scale
periodic components were extracted from the daily data and the small-scale components extracted from the TBT data. While the reconstruction does not produce exact results, the total annual volumes compared well to the observed rainfall volumes. The intensities of the generated rainfall were lower by 50% or more for large storms. The storm durations of the generated rainfall were an order of magnitude greater than the observed series. The interevent durations of the generated rainfall were also an order of magnitude lower than the observed rainfall. In the interim, an extension of this method could be to 'fill-in' missing data for a series of fine time-increment rainfall, precipitation, or streamflow data if a complete and concurrent record of some other coarse time-step data is available.

Future studies are planned to refine the reconstruction approach, to use total precipitation data instead of rainfall, to analyse other geographic regions, and to infer a correlation to other physical processes that may influence the behaviour of rainfall.

Acknowledgements

The author acknowledges and thanks Dr. Christopher Torrence and Gilbert Compo of the University of Colorado for providing valuable wavelet theory and code that helped to ease the analyses; to Rob James of CHI for providing programming help and rainfall data. This chapter is based on work presented for an MSc degree at the University of Guelph. Dr. James provided the original research concept and funds for data collection and also proved to be an enjoyable advisor.

References


Klemes, V. 1986. Dilettantism in Hydrology: Transition or Destiny? Water Resources Research. 22(9): 177S-188S.


