Calculating wavelet variance associated with discrete wavelet transform (DWT)\textsuperscript{1}

Xuejun Dong\textsuperscript{a} and Haitao Li\textsuperscript{b}

\textsuperscript{a}North Dakota State University
Central Grasslands Research Extension Center
Streeter, ND 58483

\textsuperscript{b}University of Toledo
Department of Environmental Sciences
Toledo, OH 43606

e-mail: xuejun.dong@ndsu.edu

1 Introduction

Wavelet transform is defined as a collection of convolutions of the data function, $f(x_i)$, with an appropriately shifted and stretched windowing function (wavelet) along a given time or space range (Bradshaw and Spies [1992], Walker [1999]):

$$W(a, x_j) = \frac{1}{\sqrt{a}} \sum_{i=1}^{n} f(x_i) \phi\left(\frac{x_i - x_j}{a}\right),$$

(1)

in which $\phi()$ is the wavelet function, $a$ is the scale factor determining the extent the wavelet is stretched or compressed, and $x_j$ is the extent of the shift with which the wavelet is moved along the time or space range. The above equation (1) usually refers to a continuous wavelet transform (CWT), in which the scale factor, $a$, is chosen such that a redundant time-scale portrait is provided as seen from a typical 2-D view of a CWT graph (Walker [1999]).

Understanding the dynamics of natural phenomena, such as earthquake signals, soil spatial variations and a randomly selected piece of the music by Wolfgang Amadeus Mozart (1756–1791), requires a consideration of scale. While wavelet analysis is a rapidly developing research area in statistics (see Serroukh et al. [2000]), the fundamental importance of the related methodology to numerous branches of natural science seems to be well recognized by many mathematicians. As a result, some of the clearly written articles/books, such as those by Percival and Guttrop [1994], Walker [1999], Percival and Walden [2000] and Fugal [2007], encourages non-mathematicians (such as ecologists and biologists) to utilize some of the basic tools, as can be seen in an article of Dong et al. [2008], who used minimal mathematics in the application of elementary wavelet

\textsuperscript{1} Created: August 11, 2008 / Last modified: June 1, 2010. ©2008, 2010 Xuejun Dong & Haitao Li
All Rights Reserved
Wavelet variance and DWT

Dong & Li

analysis in biology.

The concept of decomposing the total variance of an experiment into components due to different factors and their interactions has been widely used in biology. However, breaking down the total variance of observed signals into components due to the variations occurring at different spatial/temporal scales is receiving increased attention only recently (Bradshaw and Spies, 1992; Keitt and Urban, 2005). Because wavelet transform can break down the original signal into components of different scales, it provides scientists with a powerful tool to detect the pattern of variations across scales in observed data. In particular, it is interesting to calculate the wavelet variance (Bradshaw and Spies, 1992), so that different data sets may be compared at different scales:

\[
V(a) = \frac{1}{n} \sum_{j=1}^{n} W^2(a, x_j),
\]

(2)

where \(W^2(a,x_j)\) is the squared wavelet coefficient associated with scale \(a\) at data point \(x_j\), and \(n\) is the number of data points. From this definition, wavelet variance is a function of scale. Bradshaw and Spies (1992) pointed out that “High values of wavelet variance at a given scale reflect the presence of a greater number of peaks and a greater intensity of the signal, or both”. As for computer implementation, we realized that Dr. Donald B. Percival and colleagues provided an interesting discussion on the DWT-based wavelet variance analysis, some basic elements of which may be directly applied by biologists (Percival and Guttorp, 1994; Percival and Mofjeld, 1997; Percival and Walden, 2000). In the following, we review a few key concepts on wavelet variance based on above-mentioned papers by Percival et al.

Assume we have an observed time series \(X_t\), with \(t = 1, 2, ..., N\) and \(N = 2^p\) for some positive integer \(p\). Following \(J\) levels of DWT (\(1 \leq J \leq p\)) using the Haar wavelet filter (Press et al., 1992; Percival and Guttorp, 1994), the total variance of the original signal can be decomposed as (Percival and Guttorp, 1994; Percival and Mofjeld, 1997; Percival and Walden, 2000):

\[
\hat{\sigma}_X^2 = (1/N) \sum_{t=1}^{N} (x_t - \bar{x})^2
= (1/N) \sum_{t=1}^{N} x_t^2 - \bar{x}^2
= \sum_{k=1}^{J} \left( \frac{1}{N} \right) \sum_{j=1}^{N/2^k} d_{j,k}^2 + \left( \frac{1}{N} \right) \sum_{j=1}^{N/2^J} V_j^2 - \bar{x}^2,
\]

(3)

where \(d_{j,k}\) represents the \(j\)th wavelet coefficient at \(k\)th level of decomposition, \(V_j\) is the \(j\)th scaling coefficient following \(J\) levels of decomposition, and \(\bar{x}\) is the sample average. Based on formulas provided by Percival and Guttorp (1994) for computing Haar wavelet
variance\footnote{They call it “non-overlapped” wavelet variance, as compared to the variance based on the maximal-overlapped discrete wavelet transform (MODWT). Essentially, DWT is a “judiciously” sampled version of MODWT, but interestingly it retains all the information of the original signal because the perfect reconstruction can be obtained from the coefficients of DWT\cite{Percival and Walden 2000}. For simplicity, we confine ourselves only to the non-overlapped variance in this paper.} Equation \ref{eq:3} can be rewritten as

\begin{align}
\hat{\sigma}_X^2 &= \sum_{K=1}^{2^{J-1}} \frac{v_X^2(K)}{2K} + \left(\frac{1}{N}\right) \sum_{j=1}^{N/2^j} V_j^2 - \bar{x}^2, \tag{4}
\end{align}

where $v_X^2(K)$ is the Haar wavelet variance at scale $K$ ($K$ only takes values of 1, 2, 4, 8, ..., $2^{J-1}$) under $J$ levels of decomposition, and the other terms are the same as in Equation \ref{eq:3}. The contribution to the total signal variance from the variations at scale $K$ is $v_X^2(K)/(2K)$. In the meantime, the contribution of changes in scaling coefficients, representing the smooth subsignal at $J$th level of decomposition, or higher, is characterized by \( (1/N) \sum_{j=1}^{N/2^j} V_j^2 - \bar{x}^2 \). As a special case, a full decomposition (to the maximum scale of $K = 2^{p-1}$) would allow a partition of the signal variance to different components at scales $K = 1, 2, 4, ..., 2^{p-1}$, because, as shown in Percival and Guttorp (1994)\footnote{There is a small typo in the original article of Percival and Guttorp (1994), because the single smooth coefficient following a full decomposition should be expressed as $s_1 = \sum_{t=1}^{N} X_t/(\sqrt{2})^p$, instead of the grand average $\sum_{t=1}^{N} X_t/N$.}, the final scaling coefficient is $V_1 = \sum_{t=1}^{N} X_t/(\sqrt{2})^p$, which cancels out the 2nd and 3rd terms in Equations \ref{eq:3} and \ref{eq:4}. Thus, under such a situation, Equation \ref{eq:3} becomes

\begin{align}
\hat{\sigma}_X^2 &= \sum_{k=1}^{p} \left(\frac{1}{N}\right) \sum_{j=1}^{N/2^k} d_{j,k}^2 \\
&= \sum_{K=1}^{2^{p-1}} \frac{v_X^2(K)}{2K}, \tag{5}
\end{align}

where $k$ and $K$ represent, respectively, the $k$th level of decomposition and the decomposition at scale $K$, $d_{j,k}$ is the $j$th wavelet coefficient at $k$th decomposition level and $v_X^2(K)$ the wavelet variance at scale $K$. As a result, in Equations \ref{eq:3}, \ref{eq:4}, and \ref{eq:5} the total variance of the original signal is completely decomposed into contributions from variations occurring at different dyadic scales of $K = 1, 2, 4, ..., 2^{J-1}$.

\subsection*{2 Computer implementation}

The stand-alone program \texttt{NDHaar} now has been slightly changed so that wavelet variances corresponding to different scales are also computed. These may provide useful information and new insights into our data set\footnote{The Haar wavelet, instead of modified Haar, should be used in order to get a correct wavelet variance decomposition, because in the former, the factor $\frac{1}{\sqrt{2}}$, instead of $\frac{1}{2}$, is used in Equation \ref{eq:1} which ensures that the contributions to the total variance from the variations at different scales cancel out, thus allowing the correct decomposition of the signal variance.}. The purpose of this program is to generate
Table 1: Scale-specific decomposition of the total signal variance for the sample data set pdsi.txt using a discrete wavelet transform based on the Haar wavelet.

<table>
<thead>
<tr>
<th>Scale</th>
<th>Allan variance</th>
<th>Haar variance</th>
<th>Re-scaled Haar variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>128</td>
<td>1.21059</td>
<td>154.956</td>
<td>0.6053</td>
</tr>
<tr>
<td>64</td>
<td>1.68193</td>
<td>107.644</td>
<td>0.84097</td>
</tr>
<tr>
<td>32</td>
<td>3.92032</td>
<td>125.45</td>
<td>1.96016</td>
</tr>
<tr>
<td>16</td>
<td>1.84356</td>
<td>29.497</td>
<td>0.92178</td>
</tr>
<tr>
<td>8</td>
<td>2.01862</td>
<td>16.149</td>
<td>1.00931</td>
</tr>
<tr>
<td>4</td>
<td>1.25792</td>
<td>5.032</td>
<td>0.62896</td>
</tr>
<tr>
<td>2</td>
<td>0.54744</td>
<td>1.095</td>
<td>0.27372</td>
</tr>
<tr>
<td>1</td>
<td>0.51059</td>
<td>0.511</td>
<td>0.2553</td>
</tr>
</tbody>
</table>

discrete wavelet transform and wavelet variance for one-dimensional data series with a length of one of the following choices: 64, 128, 256, 512, 1024, 2048, 4096, 8192, 16384, 32768, or 65536. The input data file must have the same format as test series pdsi.txt and needs to be in the same folder as the executable program file. The outputs include (a) results of the multi-resolution analysis (MRA) based on different scales (the first three output files); (b) wavelet coefficients and scaling coefficients (the forth output file); and (c) the wavelet and Allan variances.

As this modified program (NDHaar_wva.exe) was bound using the True Basic Silver, the three dll files must be installed in the \C:\Windows\System directory of a PC using the Windows XP Operating System.

Checklist of files:

1. “NDHaar_wva.exe” is the updated program for calculating DWT and wavelet variance;
2. “wva.pdf” is the documentation;
3. The five output files (01.txt, 02.txt, 03.txt, 04.txt, 05.txt) from the test series “pdsi.txt” using the Haar wavelet.
4. The three dll files that are needed in order to run this program on a computer not installed with True Basic.

“energy” conservation [Walker, 1999]. In the meantime, the interested reader might run this program using the “modified” Haar wavelet so that to actually see that it does not provide a correct scale-specific decomposition of the signal variance, although the scale-specific synthesized signals can be used to perfectly reconstruct the original signal. See discussion in the next section for detail.

Available from http://www.ag.ndsu.nodak.edu/streeter/wavelets/wavelets.htm

X. Dong thanks Professor Na Zhang at the Graduate School of the Chinese Academy of Sciences, Beijing, for her interest and for providing literature in wavelet variance.
Table 2: The effect of using different wavelet filters (Haar or modified Haar wavelet) on variance decomposition of the sample data series pdsi.txt. With this current True Basic program, the decomposition stops when there are four elements left in the trend sub-signal. In this particular example (using data series pdsi.txt), there are four scaling coefficients, which are the first four elements of the array of wavelet coefficients, as can be found in the forth output file (O4.txt). So, “Sum of variances” and “Averaged sum of squared scaling coefficients” refer to the first and second term, respectively, of Equation 4. The “Re-constructed total variance” refers to the total variance obtained by adding together the contributions from all scales used in the DWT. As a result, the fact that “Re-constructed total variance” equals the “Original total variance” signifies a correct decomposition of the original signal variance, which is what we want for variance decomposition.

<table>
<thead>
<tr>
<th>Wavelet type</th>
<th>Original total variance</th>
<th>Sum of variances</th>
<th>Averaged sum of squared scaling coefficients</th>
<th>Re-constructed total variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Haar</td>
<td>6.67999</td>
<td>6.49549</td>
<td>0.312219</td>
<td>6.67999</td>
</tr>
<tr>
<td>Modified Haar</td>
<td>6.67999</td>
<td>0.406147</td>
<td>0.00121961</td>
<td>0.279645</td>
</tr>
</tbody>
</table>

3 Some testing results

Interested readers can run the program on a PC using the sample data set, or any other one dimensional data series with a length as specified in this note. In the following, we give some comments on two important issues: (a) the interpretation of the variance decomposition results and (b) the importance why $1/\sqrt{2}$, but not $1/2$, is used in the Haar wavelet filter:

- In a study by Keitt and Urban (2005), the Allan variance (Percival and Guttrop, 1994) was used to signify the existence of pattern at a particular scale. In the meantime, we note that the computed Haar wavelet variances for different scales do not directly indicate the “contributions” from specific scales (Table 1). Instead, as shown in Equation 4 the Haar wavelet variance at a particular scale needs to be divided by $2^K$, so as to signify that scale’s “contribution” to the total signal variance (see the re-scaled Haar variance in Table 1). Also, note that the re-scaled Haar wavelet variance is half of the corresponding Allan variance.

- Dong et al. (2008) used a “modified” Haar wavelet filter for an easy computation of the major steps of the DWT. However, although using $1/2$, instead of $1/\sqrt{2}$, retains some of the features of a DWT, such as the effect of “moving averaging” and “moving differencing”, as well as the possibility of a perfect reconstruction of the original signal using the scale-specific synthesized signals through MRA, it does not give a correct decomposition of the total signal variance, as shown in Table 2. The main reason is that, transforming the original signal using our modified Haar wavelet filter (containing $1/2$) does not meet the orthonormality condition in wavelet transform (Press et al., 1992; Percival and Walden, 2000).

---

8This can be verified easily using the current True Basic program.
References


