Application of Discrete and Continuous Wavelets for Saudi Arabian Meteorological Data Analysis

Shafiqur Rehman,1* Abul Hasan Siddiqui2 and Naif M. Al-Abbadi3
1Engineering Analysis Section, Centre for Engineering Research, Research Institute
King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia
2Department of Mathematical Sciences, King Fahd University of Petroleum and Minerals Dhahran,
Saudi Arabia
3Energy Research Institute, King Abdul Aziz City for Science and Technology, Riyadh, Saudi Arabia

Abstract: This paper presents the wavelet-based analysis of meteorological parameters for nine meteorological stations in Saudi Arabia. The study utilised the daily mean values of pressure, temperature, relative humidity and wind speed data over a period of 16 years between 1990 and 2005. In case of rainfall data, daily total values were used. The time series of meteorological parameters has so much noise that their overall shape is not visible upon visual inspection. It has been observed in the present investigation that trends become more and more readable at decomposition at different scales. The wavelet tools have been used to reveal the trends of the meteorological data time series. If the signal itself includes sharp changes then successive approximation look less and less similar to the original signal. This paper deals with the meteorological data visualisation, such as pressure, temperature, relative humidity, rainfall and the wind speed time series of the Kingdom of Saudi Arabia. The present analysis is based on discrete and continuous wavelet transforms.

Keywords: Wavelets, Meteorology, Denoising, Compression, Pressure, Temperature, Relative humidity, Precipitation, Wind speed.

Introduction
Wavelet analysis is a tool for analysing localised variations in power by decomposing a trace into time frequency space to determine both the dominant modes of variability and how these modes vary in time. This method is appropriate for analysis of non-stationary traces, i.e. where the variance does not remain constant with increasing length of the data set. Fractal properties are present where the wavelet power spectrum is a power law function of frequency. The wavelet method is based on the property that wavelet transforms of the self-affine traces have self-affine

* Author for Correspondence. srehman@kfupm.edu.sa
Wavelet analysis has been systematically developed in the last two decades and its applications to diverse fields of human knowledge are being investigated. A plot of wavelet coefficients clearly shows the exact location of the discontinuity in time domain. Wavelet analysis is capable of revealing certain aspects of data, which are missed by other signal analysis techniques, viz., trends, breakdown points, discontinuities in higher derivatives and self similarities. Further, wavelet analysis can compress or denoise a signal without appreciable degradation of the original signal. Local features of a signal can be described better with wavelets that have a local extent.

According to Warren et al. [1], major environmental changes are likely to be experienced in the arid and semi-arid regions of the world in the coming few decades. Furthermore, these changes are expected to be induced by increasing human pressure like exponentially growing population, increasing usage of fossil fuel and ambitious water development projects. Climatic changes strongly affect the process of desertification due to their impact on the vegetation, soil and hydrological cycle, as stated by Pimenta et al. [2]. According to Schwallm and Ek [3], climate changes directly change the temperature, wind speed, barometric pressure and rainfall. The climate variability and climate changes have been studied based on the analysis of different climatic variables. Temperature and rainfall time-series are commonly used to study different climatic changes [4-7].

Investigators from around the world have studied climatic change, such as Balling and Brazel [8] (1987), as well as Conmee and Broyles [9] for USA; Jose et al. [10] for Philippines; Elagib and Abdu [11] for the Kingdom of Bahrain; Velichko et al. [7] for Europe; Kipkorir [12] for Kenya; Abahussain et al. [13] for the Arab Region; Yu et al. [14] for Taiwan; and Moonen et al. [15] for Italy. There are a number of studies that discuss the anthropogenic impacts on the environment in arid regions such as Kuwait [16], Riyadh and Saudi Arabia [17], as well as Tehran and Iran [18]. Qureshi and Khan [19] have documented localised impacts resulting from urban and agriculture development on the environment, including the modification of local climate. Böer [20] used existing meteorological data of the United Arab Emirates (UAE) to study the country's climatic conditions. The climate of UAE is classified as hyper-arid with different bioclimatic zones within the country. There are higher mean precipitation rates in the northeastern part and low temperatures as compared to those in the southern and western regions [20].

Abbry and Veitch [21] studied the long-range dependence of physical processes using wavelet technique. Katherine and Guy [22] used wavelets to predict wind conditions using short-term data collected at a site and referred to long-term data from other meteorological stations. They modelled the response time-series in terms of a multi-scale wavelet decomposition of the explanatory time-series. Wavelet analysis was used to investigate the frequency structure of the air temperature variability in the Aral Sea region by Valentina et al. [23]. In order to investigate the time-scale structure of natural wind, Kitagawa and Nomura [24] applied the wavelet transform to the time history of a measured wind velocity data and constructed the wavelet-based method in such way that the created time history possesses characteristics similar to those of the natural wind data.

This study utilises daily average values of meteorological data for a period of 16 years between 1990 and 2005 to study the climatic dynamics of nine locations through modern wavelet and wavelet based techniques. The wavelet-based analysis of meteorological parameters was performed in terms of decomposition, approximation, compression and denoising of the original signals.
Meteorology of Saudi Arabia and Description of Meteorological Stations Used

Saudi Arabia is one of the hottest and driest countries in the world. It is located, approximately, between 17° and 31° North latitudes and 37° and 56° East longitudes. The land elevation varies from a few metres to a few thousand meters above mean sea level. Most of the country consists of desert and semi-desert with oasis. Almost 50% of the total surface area is uninhabitable. The eastern part of Saudi Arabia is lowland with highly humid and hot climate while most of the western part is a plateau. The Southwest part of the country is surrounded by mountains, almost 3,000 m high. The annual average precipitation in the country varies from 80 to 140 mm, as reported by Alkolibi [25]. The summer extreme temperature reaches as high as 45°C while the relative humidity is very low and the skies are clear most of the time during the year. In the middle part of Saudi Arabia, known as Empty Quarter or Rub Al-Khali desert, the rainfall is minimal, almost nonexistent. In highland and mountainous areas of the country, there occurs relatively more precipitation. According to Shepered [26], the precipitation increases towards the Southwest of the country in the Atlas Mountains and towards Northeast in the Zagros Mountains.

In the present study, data between 1990 and 2005 – available at nine selected meteorological stations, keeping in view their geographical location and climatic conditions – has been considered. The meteorological data includes the daily minimum, maximum and mean values of wind speed (WS), wind direction, dry bulb temperature (referred as temperature (TP) here), wet bulb temperature, station pressure (PR), relative humidity (RH), sea level pressure, cloud parameters and visibility. The rain was given as the daily total rain (Rain) in mm for all the stations. Table 1 summarises the latitudes, longitudes, altitudes, data period and number of daily records for each location. Of these nine stations, Gizan, Jeddah and Yanbu are located on the Red Sea while Dhahran is situated on the Arabian Gulf coast. Turaif and Guriat are located in the North, close to the border with Jordan while Riyadh is in the central part and represents dry as well as hot climate. Abha is located in the South. It is surrounded by high hills as well as complex topography and is classified as cold climatic region. All these stations (Table 1), have national and international airports of the Kingdom. The data was manually collected and recorded by an attendant on an hourly basis. Then it was fed to a computer as daily mean, minimum and maximum values for each parameter.

The meteorological data describes the climatic conditions of a location and, hence, the range of each parameter needed to be checked for accuracy. The data was also checked for completeness.

<table>
<thead>
<tr>
<th>Station</th>
<th>Latitude</th>
<th>Longitude</th>
<th>Elevation (m)</th>
<th>Period</th>
<th>Records</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abha</td>
<td>18°13'</td>
<td>42°31'</td>
<td>2,200</td>
<td>1990-2005</td>
<td>5964</td>
</tr>
<tr>
<td>Dhahran</td>
<td>26°06'</td>
<td>50°10'</td>
<td>22</td>
<td>1990-2005</td>
<td>5964</td>
</tr>
<tr>
<td>Gizan</td>
<td>16°52'</td>
<td>42°35'</td>
<td>5</td>
<td>1990-2005</td>
<td>5964</td>
</tr>
<tr>
<td>Guriat</td>
<td>30°54'</td>
<td>41°08'</td>
<td>542</td>
<td>1990-2005</td>
<td>5964</td>
</tr>
<tr>
<td>Hail</td>
<td>27°31'</td>
<td>41°44'</td>
<td>992</td>
<td>1990-2005</td>
<td>5964</td>
</tr>
<tr>
<td>Jeddah</td>
<td>21°30'</td>
<td>39°12'</td>
<td>17</td>
<td>1990-2005</td>
<td>5964</td>
</tr>
<tr>
<td>Riyadh</td>
<td>24°42'</td>
<td>46°44'</td>
<td>624</td>
<td>1990-2005</td>
<td>5964</td>
</tr>
<tr>
<td>Turaif</td>
<td>31°41'</td>
<td>38°40'</td>
<td>827</td>
<td>1990-2005</td>
<td>5964</td>
</tr>
<tr>
<td>Yanbu</td>
<td>24°07'</td>
<td>38°03'</td>
<td>6</td>
<td>1990-2005</td>
<td>5964</td>
</tr>
</tbody>
</table>
and erroneous values in the light of standard data validation practices of the World Meteorological Organisation (WMO). In the present case, WS, PR, TP, RH and Rain values were checked for accurate range, erroneous values and missing data. The range of each station is dictated by the long-term behaviour of the data as well as the experience and expertise of the meteorologist familiar with the meteorology of the region. The present data was corrected for out of range values of WS and PR. The TP and RH values were found to be correct with few exceptions. PR and WS were found to be out of normal range for most of the locations and were corrected.

**Methodology Used**

The name wavelet means small waves – the sinusoids used in Fourier analysis are big waves. In brief, a wavelet is an oscillation that decays quickly. The equivalent mathematical conditions are

\[ \int_{-\infty}^{\infty} |\psi(t)|^2 \, dt < \infty \]

(1)

\[ \int_{-\infty}^{\infty} |\psi(t)| \, dt = 0 \]

(2)

The admissibility condition is given as

\[ \int_{-\infty}^{\infty} \frac{|\psi(\xi)|^2}{|\xi|} \, d\xi < \infty \]

In wavelet theory, a function is represented by infinite series expansion in terms of dilated and translated versions of a basis function, and is called *mother wavelet*, satisfying Eqs. (1) and (2)

\[ \psi_{a,b}(t) = a^{-1/2} \psi \left( \frac{t-b}{a} \right) \text{, where } a > 0 \]

\[ T_{\psi}f(a, b) = a^{-1/2} \int_{-\infty}^{\infty} f(t) \psi \left( \frac{t-b}{a} \right) \, dt \]

\[ = \langle f, \psi_{a,b} \rangle \]

where \( T_{\psi}f(a, b) \) is called the wavelet transform of \( f(t) \).

A wavelet transform, \( T_{\psi} \) decomposes a signal into several groups of coefficients. The vectors of different coefficients contain information about characteristics of the sequence at different scales. It may be observed that a wavelet transform is a prism, which exhibits properties of a signal, such as points of abrupt changes, seasonality or periodicity. The wavelet transform is a function of \( a \) and \( b \), where \( a \) is the scale of frequency and \( b \) the spatial position or time. The plane defined by the variables \( (a, b) \) is called the scale-space or time frequency plane. The wavelet transform \( T_{\psi}(a, b) \) measures the variation of \( f \) in a neighborhood of \( b \).

At small scales, \( b \), \( T_{\psi}f(a, b) \) provides localised information, such as localised regularity of \( f \). The local regularity of a function is often measured with Lipschitz exponent. The global and local
Lipschitz regularity can be characterised by the asymptotic decay of wavelet transformation at small scales. For example, if \( f \) is differentiable at \( b \), \( T_{\psi} f(a, b) \) has the order \( a^{3/2} \) as \( a \to 0 \).

\[
\psi_{j, k}(t) = 2^{j/2} \psi(2^j t - k) \quad \text{is a discretisation version of } \psi_{a, b}(t)
\]

where \( j \) and \( k \) are integers.

\( \psi(t) \) is called an orthonormal wavelet if

\[
\langle \psi_{j, k}(t), \psi_{m, n}(t) \rangle = \int_{-\infty}^{\infty} \psi_{j, k}(t) \psi_{m, n}(t) \, dt = 0 \quad \text{for } j \neq m, \text{ or } k \neq n
\]

\[= 1 \quad \text{for } j = m, \text{ and } k = n
\]

\( d_{j, k} = \int_{-\infty}^{\infty} f(t) \psi_{j, k}(t) \, dt \) are called wavelet coefficients of \( f(t) \) and \( \sum_{j, k} d_{j, k} \psi_{j, k}(t) \) is called the wavelet series of \( f \).

With every orthonormal wavelet there is an associated function \( \phi(t) \), called scaling function, such that \( \phi_{j, k}(t) = 2^{j/2} \phi(2^j t - k) \).

\( c_{j, k} = \int_{-\infty}^{\infty} f(t) \phi_{j, k}(t) \, dt \) are called scaling coefficients and \( \sum_{j, k} c_{j, k} \phi_{j, k}(t) \) is called the scaling series.

**Mallat Algorithm (Pyramid or Cascade Algorithm)**

\[
C_{j, k} = \sum_{l \in \mathbb{Z}} h_{l-2k} C_{j+1, l}
\]

(3)

\[
d_{j, k} = \sum_{l \in \mathbb{Z}} (-1)^l h_{-l+2k+1} C_{j+1, l}
\]

(4)

where \( \{h_j\} \) is a sequence characteristic of the associated wavelet, e.g., for Haar Wavelet

\[
h_j = \begin{cases} 
1, & j = 0, 1 \\
\frac{1}{\sqrt{2}}, & j = 0 \text{, otherwise} \\
0, & \text{otherwise}
\end{cases}
\]

It is important to note that for given scaling coefficients, at any level \( j \), all lower-level scaling function coefficients \( i < j \) can be computed recursively using Eq. (3) and all lower-level wavelet coefficients \( i < j \) can be computed from the scaling function coefficients, applying Eq. (4).

**Reconstruction Algorithm**

\[
C_{j, k} = \sum_{l \in \mathbb{Z}} h_{l-2k} C_{j+1, l} + (-1)^k h_{2l-k+1} d_{j-1, l}
\]

(5)

The scaling function coefficient at any level can be computed from only one set of low-level scaling function coefficients and all the intermediate wavelet coefficients by applying Eq. (5), recursively. For wavelets with compact support, the sequences \( \{d_{j, k}\} \) and \( \{c_{j, k}\} \) will only contain finitely many non-zero elements. For wavelets with support on all of \( R \), each sequence element \( h_k \) is, in general, non-zero, but elements decay exponentially. Discrete wavelet transform (DWT) is
commonly introduced using a matrix or a computation form. In matrix form, the DWT can be represented through an orthogonal matrix

\[ W = \begin{bmatrix} w_1^t, w_2^t, ..., w_j^t, v_j^t \end{bmatrix} \]  

(6)

where \( v_j \) is a scaling function, \( j \) the largest level of the transform and \( t \) indicates the transpose. A DWT is applied to a vector \( X \) of observations as \( d = WX \) and decomposes the data into sets of wavelets coefficients

\[ d = \begin{bmatrix} d_j^1, d_j^2, ..., d_j^j, c_j^j \end{bmatrix} \]  

(7)

where \( d_j = w_j X \) and \( c_j = v_j X \).

**Multi-resolution Analysis**

A wavelet transform leads to an additive decomposition of a signal into a series of different components describing smooth and rough features of the signal. In fact, we have

\[ X = W'd = \sum_{j=1}^{J} W_j^t d_j + V_j^t c_j = \sum_{j=1}^{J} D_j + C_j \]  

(8)

where \( D_j \) is the detail of the signal that describes changes at scale \( \tau_j \) and \( C_j \) is the smooth component associated with variations, such as \( \tau_{j+1} \) and higher. Mallat algorithms are used for computation of DWT.

**Results and Discussion**

The results for discrete and continuous wavelets are presented as follows:

**Meteorological Data Through Discrete Wavelet Lens**

This section deals with discrete wavelet analysis of meteorological parameters, such as PR, TP, RH, Rain and WS for Abha, Dahran, Giza, Guryat, Hail, Jeddah, Riyadh, Turayf and Yanbu meteorological data collection stations. The data covers the period between 1990 and 2005. In this analysis, daily average values of PR, TP, RH and WS were used for the above period. For rainfall data analysis, Rain values were used.

The discrete wavelet analysis of meteorological parameters was performed in terms of the decomposition, approximation, compression and denoising of the original signal. The decomposition analysis of surface pressure data for the nine stations was performed using discrete wavelets. For example, Fig. 1 shows the analysis for Abha only. In Fig. 1, the x-axis shows the number of days of the entire data period used in this study. Figures 1 to 5 have seven parts. The first part, \( S \) represents the signal or raw data; the second part, \( a_5 \) corresponds to the amplitude of the signal for wavelet Daubechies (db) at level 5. The other five parts – \( d_1, d_2, d_3, d_4 \) and \( d_5 \) – represent details of the signal or raw data at five different levels.

The comparison of signal strength \( S \), or \( S/d \), in the first part of Figs. 1 to 5 shows that the minimum value of 788-798 corresponds to Abha while the maximum value of 1000-1020 corresponds to Dhahran, Hail and Turayf.. The corresponding amplitudes, i.e. \( a_5 \), are 790-795, 995-1015, 1000-
Fig. 1. Decomposition of PR time series data for Abha using db5.

Fig. 2. Decomposition of TP time series data for Abha using db5.
Fig. 3. Decomposition of RH time series data for Abha using db5.

Fig. 4. Decomposition of Rain time series data for Abha using db5.
1010 and 1000-1010 for Abha, Dhahran, Hail and Turaij, respectively. The values of $d_5$ for all of these locations were found to be between $-2$ and $+2$, as can be seen from Fig. 1. A close look of $S$ (signal), $a_5$ (amplitude), and $d_5$ (detail at level 5) indicates that Dhahran, Gizeh, Hail and Turaij exhibit similar surface pressure characteristics. Similarly, the study of $S, a_5$ and $d_5$ show that Guryat, Riyadh and Yanbu have similar surface pressure characteristics. Abha, being a different station, both from the meteorological and geographical point of view, has unique values of different parameters as compared to the other stations. Jeddah and Yanbu, both coastal locations on the Red Sea, show similar surface pressure characteristics. The decomposition for TP, RH, Rain and WS is shown in Figs. 2 to 5 for Abha and it is self-explainable.

**Meteorological Data Through Continuous Wavelet Lens**

One dimensional continuous wavelet analysis of PR, TP, RH, Rain and WS data time series is performed using Mexican Wavelet. The visual outputs of this analysis for PR time series from Abha, Dhahran, Gizeh, Guryat, Hail, Jeddah, Riyadh and Yanbu were obtained for all the locations, but are included for Abha only (Figs. 6), as an example, in this paper. The visual outputs for TP, RH, Rain and WS for Abha time series data are shown in Figs. 7 to 10, respectively. The variation and changes can be explained in a similar fashion as that for PR. Figures 6 to 10 have four parts each. The first (top) part represents the analysed signal, or the raw data. In the present case, it is PR variation. The second part shows the scalogram of pressure values. The colour-scale changes from minimum (dark colours) to maximum (light colours). The vertical axis shows the frequency values while the horizontal axis represents the number of days. The third part shows the daily variation of coefficients. The fourth part explains the local maximums of related parameters.
Fig. 6. Continuous wavelet analysis of PR data using Mexican hat wavelet for Abha.

Fig. 7. Continuous wavelet analysis of TP data using Mexican hat wavelet for Abha.
Fig. 8. Continuous wavelet analysis of RH data using Mexican hat wavelet for Abha.

Fig. 9. Continuous wavelet analysis of Rain using Mexican hat wavelet for Abha.
Fig. 10. Continuous wavelet analysis of WS data using Mexican hat wavelet for Abha.

As seen from Fig. 6 (and others, not included here), the wavelet coefficient lies between 1 and 61 for all the locations (second part), the coefficient lines lie between –5000 and +5000 (third part) and the lines of local maxima lie between 1 and 61 (last part). The colour transition from dark to light in the second part of Fig. 6 shows that PR changes from minimum to maximum during different days of the year. PR is higher during winters and lower during summers. The scalograms of the second part of Fig. 6 are in good agreement with the pattern shown in their first part.

Wavelet Approximation, Compression and Denoising
We have studied the approximation, compression and denoising of PR, TP, Rain, RH and WS time series data for nine meteorological data collection stations. Figures related to approximation include reconstruction approximation at level 5 for db5 and statistical details. Figures related to wavelet compression of PR, TP, Rain, RH and WS time series for the nine stations contain original and compressed signals, original coefficients, thresholded coefficients, retained energy in percentage and number of zeros in percentage. Figures related to wavelet denoising of PR, TP, Rain, RH and WS time series for the nine stations contain original details coefficients, original and denoised signals, original coefficients and thresholded coefficients. Due to limitation of space, only one sample related to approximation, compression and denoising for PR time series from Abha has been shown in Figs. 11 to 13, respectively.
Fig. 11. Approximation of PR time series data for Abha using db5.

Fig. 12. Compression of PR time series data for Abha using db5.
Fig. 13. Denoising of PR time series data for Abha using db5.

Conclusion
During the investigation, it was observed that the wavelet method is a powerful tool to analyse and understand the complex structure of meteorological parameters. Through this methodology we can have a general view of overall changes in meteorological parameters with time. Discontinuities giving abnormal events can be easily visualised using wavelet tools. This study does not attempt to predict the behaviour of meteorological parameters in the future time domain, but it seems to be workable.

Acknowledgement
The authors wish to thank the King Fahd University of Petroleum and Minerals for providing the financial and administrative support vide project number INT-MS/FRACTAL/310/2006 for carrying out this research work.

References


22. Katherine, H. and Guy, P.N., “Wind Speed Modeling and Short-Term Prediction Using Wavelets”, Department of Mathematics, University Walk, Bristol, BS8 1TW, UK.

