# Wavelet Packet-Autocorrelation Function Method for Traffic Flow Pattern Analysis

Xiaomo Jiang & Hojjat Adeli*

*Department of Civil and Environmental Engineering and Geodetic Science, The Ohio State University, 470 Hitchcock Hall, 2070 Neil Avenue, Columbus, OH 43210, USA*

**Abstract:** *Accurate and timely forecasting of traffic flow is of paramount importance for effective management of traffic congestion in intelligent transportation systems. A detailed understanding of the properties of traffic flow is essential for building a reliable forecasting model. The discrete wavelet packet transform (DWPT) provides more coefficients than the conventional discrete wavelet transform (DWT), representing additional subtle details of a signal. In wavelet multiresolution analysis, an important decision is the selection of the decomposition level. In this research, the statistical autocorrelation function (ACF) is proposed for the selection of the decomposition level in wavelet multiresolution analysis of traffic flow time series. A hybrid wavelet packet-ACF method is proposed for analysis of traffic flow time series and determining its self-similar, singular, and fractal properties. A DWPT-based approach combined with a wavelet coefficients penalization scheme and soft thresholding is presented for denoising the traffic flow. The proposed methodology provides a powerful tool in removing the noise and identifying singularities in the traffic flow. The methods created in this research are of value in developing accurate traffic-forecasting models.*

## 1 INTRODUCTION

There has been a steady increase in both rural and urban freeway traffic in the recent years resulting in congestion in many freeway systems. The freeway traffic congestion can no longer be dealt with simply by extending more highways for economical and environmental reasons (Kerner, 1999). As a consequence, optimum use of

*To whom correspondence should be addressed. E-mail: *adeli.1@osu.edu*.

existing traffic network to manage the traffic congestion has increasingly become a more desirable alternative. Different types of freeway sections display various traffic flow characteristics. A number of researchers have investigated the characteristics of traffic flow in freeways (Disbro and Frame, 1989; Dendrinos, 1994; Smith and Demetsky, 1997; Williams et al., 1998; Kerner, 1999; Lee and Fambro, 1999).

Intelligent transportation systems (ITS) play a significant role in optimizing the existing traffic network. The traffic information collected in ITS is broadcasted through variable information display boards or radio and GPS systems. Accurate and timely forecasting of traffic flow is of paramount importance in ITS to manage the traffic congestion in the freeway network effectively. A detailed understanding of the properties of traffic flow is essential for building a reliable forecasting model.

Actual observations of traffic flow in fields are typically time-series measurements of a scalar quantity at a fixed spatial point. Traffic flow often demonstrates a strong periodicity or seasonality (e.g., per day or per week). On the other hand, the traffic flow displays an atypical pattern near major recreational areas, shopping centers, or sports stadiums. Also, imprecision and/or noise often exist in the observed traffic flow due to errors in recording equipments, discontinuous records, sensitivity of equipments, and counting methods.

Wavelet-based signal processing is a powerful tool for the analysis and synthesis of time series (Mallat, 1999; Percieval and Walden, 2000). Characteristics localization of time series in spatial (or time) and frequency (or scale) domains can be accomplished efficiently through wavelet decomposition. The power of wavelets for time-series analysis stems from three features. First,

wavelet analysis can determine the sharp transitions simultaneously in both frequency and time domains. Thus, wavelets can help identify nonlinear, chaotic or fractal behavior displayed in any signal. Second, wavelet analysis allows for an effective representation of discontinuities in the chaotic time series. The wavelet representation of information in the time series allows for its hierarchical decomposition. In this way, the information can be analyzed in components of desired characteristics and at various levels of details. Third, when the information in time series is transformed into the wavelet domain less storage is required for its effective representation, resulting in computational efficiency for large time series.

The goal of this research is to identify important characteristics of traffic flow using wavelets and statistical autocorrelation function (ACF) analysis. The ACF is used to judiciously choose the decomposition level of wavelets analysis. A hybrid wavelet packet-ACF method is proposed for analysis of traffic flow time series and determining its self-similar, singular, and fractal properties.

## 2 WAVELET TRANSFORM

Wavelets constitute a family of mathematical scaling functions used to represent data or other functions in both time and frequency. A "wavelet transform" decomposes a signal into two subsignals: "approximation" (the average of adjacent elements, denoted by A) and "details" (difference of adjacent elements, denoted by D). The approximations represent the high-scale, low-frequency components of the signal, whereas, the details represent the low-scale, high-frequency components. Wavelet transform is computationally similar to the Fast Fourier Transform (FFT). However, unlike the sine and cosine harmonics used in the FFT, the wavelet transform decomposes a signal into a set of orthogonal basis functions, called wavelets. The wavelet transform has several advantages over the FFT including a more effective representation of discontinuities in signals and transient functions.

Both "continuous wavelet transform" (CWT) and "discrete wavelet transform" (DWT) can be used to analyze the traffic flow series. If the traffic flow, $f(t)$, is considered to be a square-integrable function of a continuous time variable, $t$, its CWT is defined as (Chui, 1992)

$$W_f(a, b) = \int_{-\infty}^{\infty} f(t)\psi_{a,b}(t)\, dt \quad (1)$$

and the two-dimensional wavelet expansion functions (wavelets) $\psi_{a,b}(t)$ are obtained from the basic function

(also known as "mother" or generating wavelet) $\psi(t)$ by simple scaling and translation

$$\psi_{a,b}(t) = \frac{1}{\sqrt{|a|}}\psi\left(\frac{t-b}{a}\right), \quad a, b \in R, \psi \in L^2(R) \quad (2)$$

where the parameters $a \neq 0$ and $b$ denote the frequency (or scale) and the time (or space) location, respectively, and $R$ is the set of real numbers. The notation $L^2(R)$ represents the square-summable time-series space of the traffic flow $f(t)$, where the superscript 2 denotes the square of the modulus of the function.

To avoid intensive computations for every possible scale $a$ and dilation $b$, the dyadic values are used for both scaling and dilation in DWT as follows:

$$a_j = 2^j, b_{j,k} = k2^j, \quad j, k \in Z \quad (3)$$

where $k$ and $j$ denote the time and the frequency indices, respectively, and $Z$ is the set of all integers.

Substituting Equation (3) into Equation (2), the following wavelet expansion function is obtained:

$$\psi_{j,k}(t) = 2^{-j/2}\psi(2^{-j}t - k), \quad j, k \in Z, \psi \in L^2(R) \quad (4)$$

and Equation (1) is rewritten as

$$W_f(j, k) = 2^{-j/2}\int_{-\infty}^{\infty} f(t)\psi(2^{-j}t - k)\, dt,$$
$$j, k \in Z, \psi \in L^2(R) \quad (5)$$

which is the DWT of the traffic flow $f(t)$. The traffic flow can be reconstructed by inverse of the CWT of Equation (1) in the double-integral form or the DWT of Equation (5) in the double-summation form.

The DWT represented by Equation (5) can be regarded as an attempt to preserve the dominant features of the CWT represented by Equation (1) in a succinct manner. Conceptually, the DWT can be considered as a judicious subsampling of CWT coefficients with just dyadic scales (i.e., $2^{j-1}$, $j = 1, 2, \ldots, N$, where $N$ is the maximum number of the decomposition level to be discussed subsequently). Compared with the DWT, the CWT may represent the physical system more accurately as it makes very subtle information visible, but it requires more intensive computations for integrating over every possible scale, $a$, and dilation, $b$.

## 3 WAVELET MULTIRESOLUTION ANALYSIS

The "multiresolution" analysis (MRA) provides a powerful tool for time–frequency analysis of signals. In this analysis, the time series is decomposed in terms of different levels of details using scaling and wavelet functions

(Mallat, 1989, 1999; Holschneider, 1995; Burrus et al., 1998). A nesting relationship is defined for various scaling function subspaces, $\mathbf{V}_i$ ($i \in Z$) as

$$\cdots \mathbf{V}_{-2} \subset \mathbf{V}_{-1} \subset \mathbf{V}_0 \subset \mathbf{V}_1 \subset \mathbf{V}_2 \subset \cdots \subset L^2(R) \quad (6)$$

where the notation $\mathbf{V}_i \subset \mathbf{V}_{i+1}$ means that $\mathbf{V}_i$ is a subspace of $\mathbf{V}_{i+1}$. The subspace $\mathbf{V}_i$ is defined by spanning $L^2(R)$ by the scaling functions, $\varphi_k(t) = \varphi(k - t), k \in Z$ as:

$$\mathbf{V}_i = \overline{\underset{k}{\text{Span}}\{\varphi_k(t)\}} \quad (7)$$

in which the overbar indicates closure of the subspace.

The wavelet subspace $\mathbf{W}_i$ is defined as the orthogonal complement of $\mathbf{V}_i$ in $\mathbf{V}_{i+1}$ such that

$$\mathbf{V}_1 = \mathbf{V}_0 \oplus \mathbf{W}_0, \quad \mathbf{V}_2 = \mathbf{V}_0 \oplus \mathbf{W}_0 \oplus \mathbf{W}_1 \quad (8)$$

and the entire time-series space of the square-summable integrals is represented as the summation of all subspaces:

$$L^2(R) = \mathbf{V}_{j0} \oplus \mathbf{W}_{j0} \oplus \mathbf{W}_{j0+1} \oplus \cdots \quad (9)$$

where $\oplus$ represents a direct space sum and the integer subscript $j_0$ is any general starting scaling parameter index from negative infinity to positive infinity including zero.

For the DWT, the subspace, $\mathbf{V}_i$, satisfies a natural scaling condition (Daubechies, 1988, 1992):

$$f(t) \in \mathbf{V}_i \Leftrightarrow f(2t) \in \mathbf{V}_{i+1} \quad (10)$$

where $\in$ denotes membership. Equation (10) means that the traffic flow series in a subspace is scaled by a dyadic factor in the next subspace.

The MRA equation is expressed in terms of a weighted sum of shifted basic scaling function $\varphi(2t)$ as follows:

$$\psi(t) = \sum_n h(n)\sqrt{2}\varphi(2t - n), \quad n \in Z \quad (11)$$

where $h(n)$ is a sequence of $n$ real or complex numbers called the scaling function coefficients (or scaling filters). The parameter $n$ is the order of the wavelet function. The values of $h(n)$ are obtained for any given type of wavelet function by satisfying two basic wavelet properties: the integral of $\psi(t)$ is zero and the integral of the squared $\psi(t)$ is unity (Daubechies, 1988, 1992). Equation (11) provides the prototype or mother wavelet $\psi(t)$ needed in Equation (4).

The traffic flow, $f(t)$, can then be represented by the inverse wavelet transform and the coefficients obtained from the MRA as follows:

$$f(t) = \sum_k s_{j_0}(k)\varphi_{j_0,k}(t) + \sum_k \sum_{j=j_0} w_j(k)\psi_{j,k}(t) \quad (12)$$

where the first term is a coarse resolution at scale $j_0$ and the second term provides the frequency and time breakdown of the signal.

Figure 1 shows the scaling, $\varphi(t)$, and wavelet, $\psi(t)$, functions for Daubechies wavelet of order 4 (Daubechies, 1992). It is a compactly supported wavelet function with an orthonormal basis. The order of the wavelet is the number of wavelet coefficients, $n$, used in Equation (11).

In Equation (12), the $j_0$ level scaling coefficients $s_{j_0}(k)$ and the $j$ level wavelet coefficients $w_j(k)$, in general, are expressed as (Burrus et al., 1998):
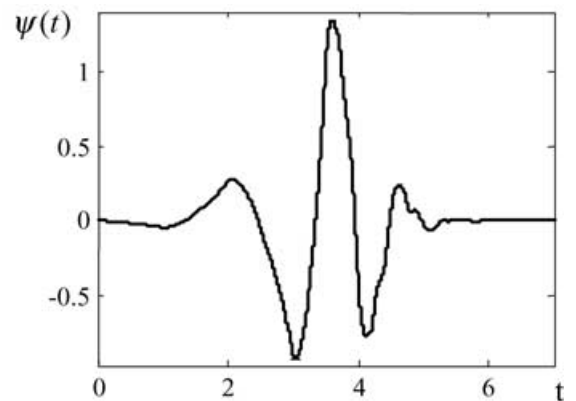
$$s_{j_0}(k) = \sum_m h(m - 2k)c_{j_0+1}(m) \quad (13)$$

$$w_j(k) = \sum_m h_1(m - 2k)c_{j+1}(m) \quad (14)$$

where $m = 2k + n$, and the sequences $h(\cdot)$ and $h_1(\cdot)$ are the scaling and wavelet filter coefficients, respectively,
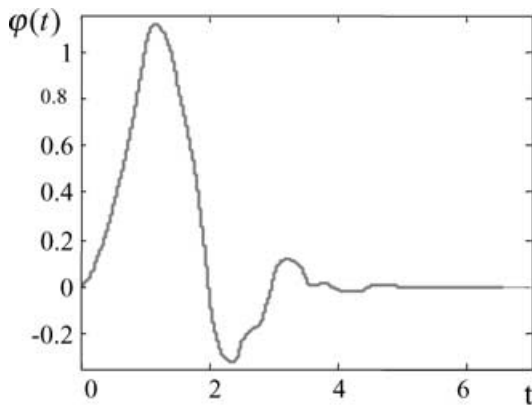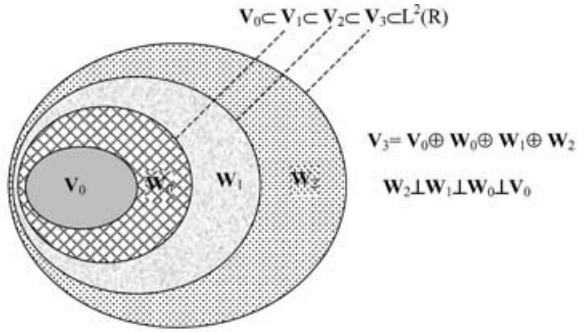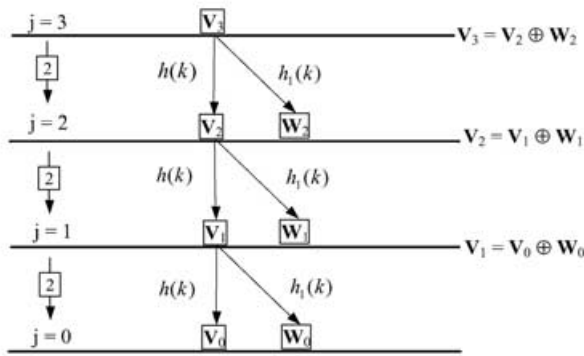


(a) Scaling function



(b) Wavelet function

**Fig. 1.** Daubechies wavelet of order 4.

(a) Nested vector spaces



(b) Decomposition tree

$\perp$: Orthogonal relationship    $\oplus$: Direct sum of space vectors
$\mathbf{W}_j$: Wavelet subspaces    $\mathbf{V}_j$: Scaling subspaces
$j$: Frequency (j=0,1,2,3)    k: Time or space location
$L^2(R)$: Square summable time series space of the traffic flow
$h(k)$: Scaling coefficients
$h_1(k)$: Wavelet coefficients    $\boxed{2}$ Dyadic scaling

**Fig. 2.** Three-level DWT decomposition of the scaling function space.

whose values are obtained for each type of wavelet by setting the integral of the wavelet equal to 0 and integral of its square equal to 1. The discrete values of the original signal are used for initial values of the scaling coefficients $c_j$. Equations (13) and (14) provide a recursive way to compute the DWT of the traffic flow series.

Figure 2 shows the decomposition of the scaling function space $\mathbf{V}_j$ in three-level DWT multiresolution analysis. The orthogonal nesting relationship of subspaces is shown in Figure 2a, where no subspace overlaps with another (nonoverlapping functions are always orthogonal). Figure 2b shows the decomposition tree for the three-level DWT wavelet MRA where a space vector is decomposed into two subspaces $\mathbf{V}_j$ (an approximation) with filter coefficients $h$ and $\mathbf{W}_j$ (a detail) with filter coefficient $h_1$. The approximation subspace $\mathbf{V}_j$ itself is then split into a second-level approximation $\mathbf{V}_{j-1}$ and detail

$\mathbf{W}_{j-1}$, and the process is repeated. The down-pointing arrows in Figure 2b denote downsampling by a factor of 2 (Equation (10)).

## 4 WAVELET PACKETS

For the conventional DWT and wavelet MRA described in the previous section, the resulting time–frequency resolution has narrow bandwidths in the low frequencies and wide bandwidths in the high frequencies. It is not sufficient for analyzing the time series displaying fractals. Coifman and Wickerhauser (1992) proposed the "wavelet packet" analysis to allow for a finer and adjustable resolution of frequencies at high frequencies (details). In the conventional wavelet transform, only the scaling functions or approximations are decomposed into subspaces. In wavelet packet analysis, both scaling functions representing the approximations and the wavelet functions representing the details are decomposed into subspaces. In that case, Equations (8) and (9) are rewritten as

$$\mathbf{V}_1 = \mathbf{V}_0 \oplus \mathbf{W}_0, \quad \mathbf{V}_2 = \mathbf{V}_0 \oplus \mathbf{W}_0 \oplus \mathbf{W}_{10} \oplus \mathbf{W}_{11} \quad (15)$$

and

$$L^2(R) = \mathbf{V}_{j0} \oplus \mathbf{W}_{j0} \oplus \mathbf{W}_{j10} \oplus \mathbf{W}_{j11} \oplus \cdots \quad (16)$$

The discrete wavelet packet transform (DWPT) provides greater flexibility for detecting the oscillatory or periodic behavior and the fractal properties of time series than the conventional DWT. Figure 3 shows the three-level DWPT decomposition of the scaling function space $\mathbf{V}_j$. The orthogonal nesting relationship of subspaces is shown in Figure 3a. In contrast to the conventional wavelet MRA shown in Figure 2a, the wavelet subspaces in DWPT are also split in the successive decompositions. Figure 3b shows the decomposition tree for the three-level DWPT, where a space vector is decomposed into two subspaces $\mathbf{V}_j$ (approximation) and $\mathbf{W}_j$ (details). The scaling and wavelet subspaces are simultaneously split into second-level subspaces, and the process is repeated until the desirable decomposition level is reached.

## 5 DWPT-BASED DENOISING

A dilemma in the signal processing of traffic signals is that it is not possible to know by any measure of certainty whether and how the measured traffic flow is corrupted by noise. There exists no mathematical time-series expression to represent the traffic flow dynamics or the noise in the traffic flow. The Fourier transform has been used in the past to reduce the noise in measured time series. This approach works reasonably well when the

$$V_0 \subset V_1 \subset V_2 \subset V_3 \subset L^2$$



$$V_0 \perp W_0 \perp W_{10} \perp W_{11} \perp W_{200} \perp W_{201} \perp W_{210} \perp W_{211}$$

$$V_3 = V_0 \oplus W_0 \oplus W_{10} \oplus W_{11} \oplus W_{200} \oplus W_{201} \oplus W_{210} \oplus W_{211}$$

(a) Nested vector spaces



(b) Decomposition tree

$\perp$: Orthogonal relationship        $\oplus$: Direct sum of space vectors
$W_j$: Wavelet subspaces        $V_j$: Scaling subspaces
$j$: Frequency ($j=0,1,2,3$)        $k$: Time or space location
$L^2(R)$: Square summable time series space of the traffic flow
$h(k)$: Scaling coefficients
$h_1(k)$: Wavelet coefficients        $\downarrow \boxed{2}$ Dyadic scaling

**Fig. 3.** Three-level DWPT decomposition of the scaling and wavelet function spaces.

signal and the noise are located in different bands of the spectrum. It does not work when the time series is chaotic, which is the case for certain traffic flow situations (Disbro and Frame, 1989; Dendrinos, 1994). In such situations, the Fourier transform cannot be used effectively to separate the noise from the traffic flow. Another disadvantage of Fourier-based denoising methods is that abrupt changes in the frequency content of a signal are spread out over the entire spectrum. As a result, transient events cannot be properly isolated from the Fourier spectrum.

Recently, Adeli and his associates developed wavelet-based denoising approaches for ITS applications and used them to create accurate and robust incident detection algorithms (Adeli and Karim, 2000; Adeli and Samant, 2000; Samant and Adeli, 2000; Karim and Adeli, 2002, 2003). They demonstrate that the wavelet-based denoising approach can remove the low-amplitude and high-frequency noise effectively. The noise-free signal can then be retrieved with little loss of details through inversing wavelet transform. This denoising approach provides a powerful tool in eliminating undesirable

fluctuations in observed data while at the same time preserving sharp transients. The DWPT described in the previous section provides more coefficients representing additional subtle details of the signal and therefore can be used to denoise the signal even more effectively than the conventional wavelet transform. Not only the DWPT can provide greater flexibility for detecting the oscillatory or periodic behavior and the fractal properties of time series, but also it can be used to denoise the signal even more effectively than the DWT.

Suppose that the traffic flow is contaminated by an additive white Gaussian noise, $e(t)$, as follows:

$$f(t) = x(t) + \sigma e(t) \qquad (17)$$

where $f(t)$ and $x(t)$ represent the noisy and denoised traffic flow, respectively, and $\sigma$ represents the standard deviation of the noise.

The proposed DWPT-based denoising approach for the traffic flow involves three steps: DWPT-based decomposition of the noisy signal, thresholding the details coefficients, and reconstructing the denoised signal (Figure 4). The DWPT-based decomposition and reconstruction of the signal were described earlier. In the thresholding step, two operations are performed to remove the noise from the details coefficients. First, the threshold parameter for de-noising at a given moment $j$, $\delta_j(t)$, is obtained by the wavelet coefficients penalization method (Barron et al., 1999):

$$\delta_j(t) = \sum_{i=0}^{t} [w_j(i)]^2 + 2\hat{\sigma}^2 t \left[ \alpha + \log\left(\frac{N}{t}\right) \right] \qquad (18)$$

where the standard deviation $\hat{\sigma}_j$ is estimated from the wavelet coefficients of the $j$th level decomposition and the second term is a penalization function, in which the parameter $\alpha$, a real number in the range 1–5, is used to modify the penalization term (A value of $\alpha = 2$ is used in this study.). Next, the denoised wavelet coefficients at

a given moment, $t$, are obtained using soft thresholding (Donoho, 1995):

$$\overline{w}_j(t) = \begin{cases} \mathrm{sgn}[w_j(t)][|w_j(t)| - \delta_j(t)], & |w_j(t)| \geq \delta_j(t) \\ 0, & \text{otherwise} \end{cases} \qquad (19)$$

where $\mathrm{sgn}(w)$ is the signum function (it returns 1 if $w$ is positive and $-1$ if $w$ is negative).

## 6 AUTOCORRELATION FUNCTION

In wavelet MRA of a given time series, an important decision is selection of the decomposition level. To the best of the authors' knowledge, currently there is no theoretical criterion to select the decomposition level. A try-and-error process is most commonly used to select the level for wavelet decomposition (Percieval and Walden, 2000). This approach, however, does not provide any rational basis for the selection and does not guarantee that the wavelet decomposition has effectively identified the desirable characteristics of the time series.

In this work, the statistical ACF is used for the selection of the decomposition level in wavelet MRA of a given time series. The motivation for the use of ACF is its capability in characterizing "self-similarity" or periodicity (i.e., the scale-invariance property of the time series). The ACF has been used widely to detect the trend and seasonality in a time series. In this research, the invariant characteristic is used to verify whether the wavelet decomposition level is sufficient or not. When the trend and/or seasonality shown in the ACF of the wavelet decomposition coefficients is the same as that of the original time series, we can reasonably conclude that the decomposition level used for DWPT provides sufficient accuracy.

Statistically, autocorrelation measures the degree of association between data in a time series separated by different time lags. The value of ACF for a time series $y$
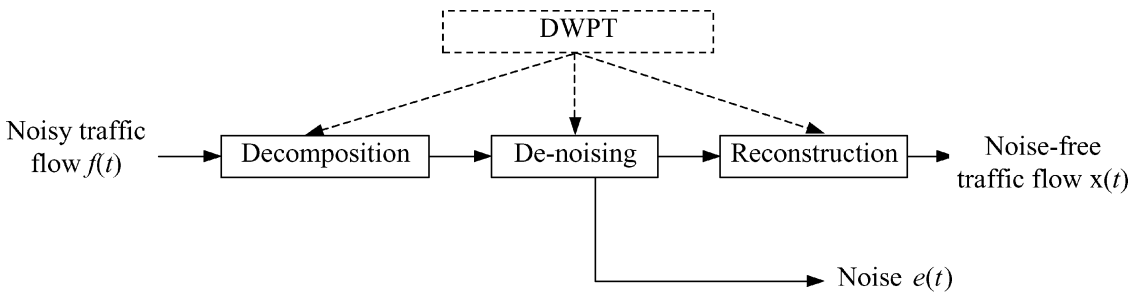


DWPT = Discrete wavelet package transform

**Fig. 4.** DWPT-based signal denoising procedure for traffic flow.

with $M$ data points at any lag time index $\tau$ is estimated as follows (Brockwell and Davis, 2002):

$$\hat{\rho}(\tau) = \frac{\sum_{m=1}^{M-\tau} [y(m+\tau) - \bar{y}][y(m) - \bar{y}]}{\sum_{m=1}^{M-\tau} [y(m) - \bar{y}]^2} \qquad (20)$$

where $\bar{y}$ is the average of the time series. In this work, the time series can be either the traffic flow or its wavelet decomposition coefficients. The value of ACF can range from $-1$ to 1.

# 7 ILLUSTRATIVE EXAMPLE OF TRAFFIC FLOW ANALYSIS

## 7.1 Traffic flow and ACF

Figure 5 illustrates the hourly average traffic flow along a two-lane freeway in the state of North Carolina with one lane closure obtained from the North Carolina Department of Transportation (NCDOT). It contains 1,824 hourly traffic flow data, continuously recorded



(a) Time series plot



(b) Three dimensional graph
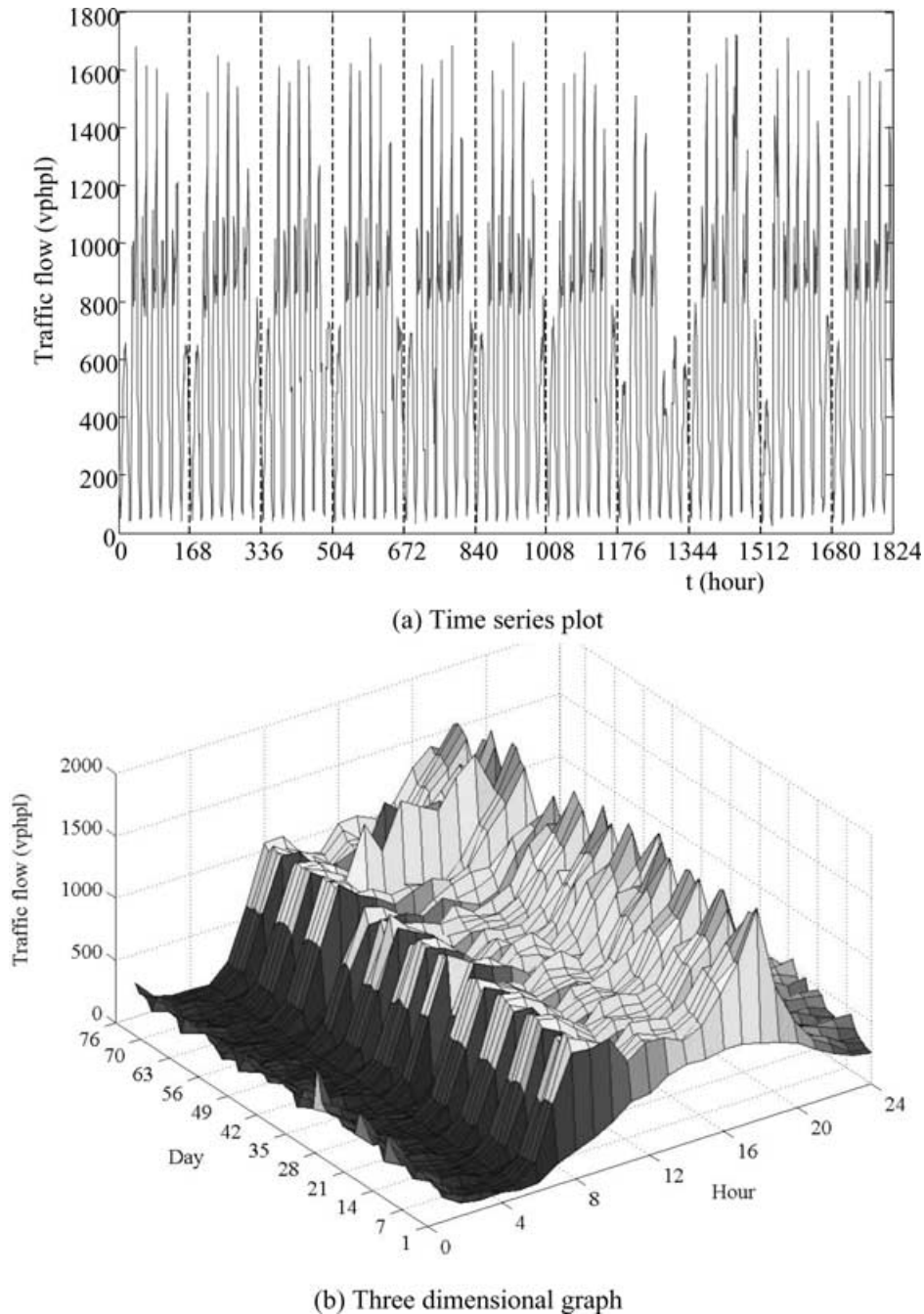
**Fig. 5.** Traffic flow measured in a two-lane freeway in the state of North Carolina with one lane closure.

over a period of 76 days from October 1 to December 15, 2000 (i.e., $M = 1,824$). Figure 5a shows the time-series plot displaying a strong seasonal periodical pattern of 168 hours (1 week) as expected. The three-dimensional graph of the time series with the plan axes of 24 hours and 76 days is shown in Figure 5b. This graph shows the periodical pattern of 24 hours (1 day) with two peak periods at around 8:00 A.M. and 6:00 P.M. everyday. The 3D graph also indicates an atypical pattern in the 61st day (November 30) from 10:00 A.M. (1,450th data point) to 4:00 P.M. (1,456th data point) (this is not readily visible in Figure 5a because of the small scale of the horizontal time axis).

Figure 6 shows the histogram of the ACF for the traffic flow data of Figure 5. Figure 6a displays the ACF values



(a) Large lags ($\tau_{max} = 600$)



(b) Small lags ($\tau_{max} = 32$)
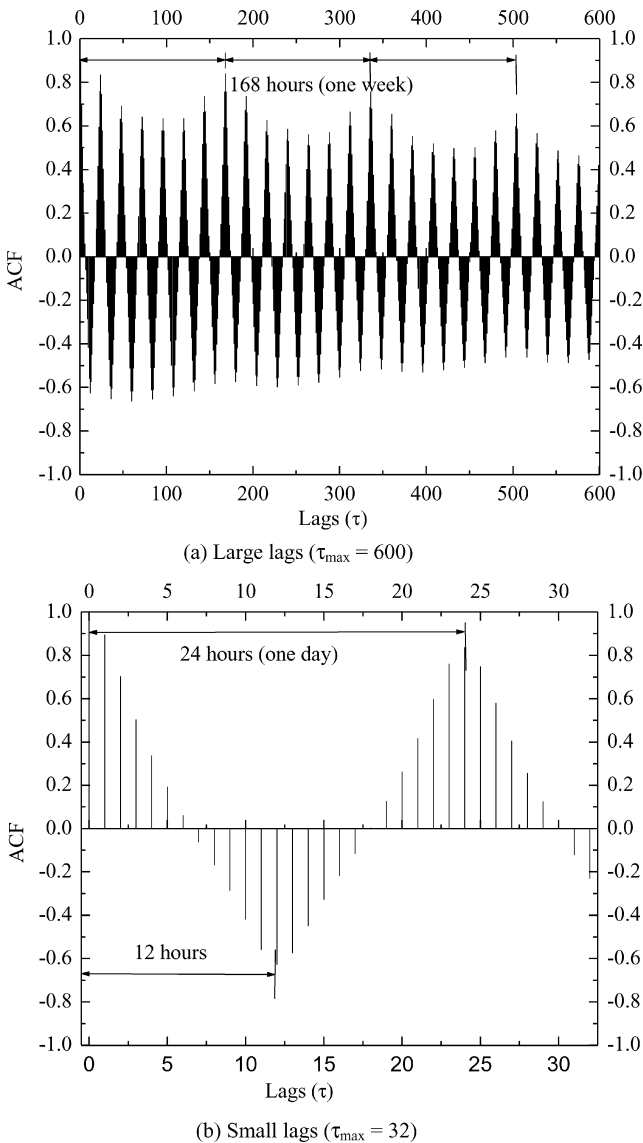
**Fig. 6.** Histogram of the ACF for the traffic flow data of Figure 5.

for a lag time of up to 600 hours, showing the periodicity over a long period of 25 days. Figure 6b displays the ACF values for a lag time of up to 32 hours, showing the periodicity over a short period of time. The ACF plots demonstrate strong seasonality and display both weekly and daily periodicity.
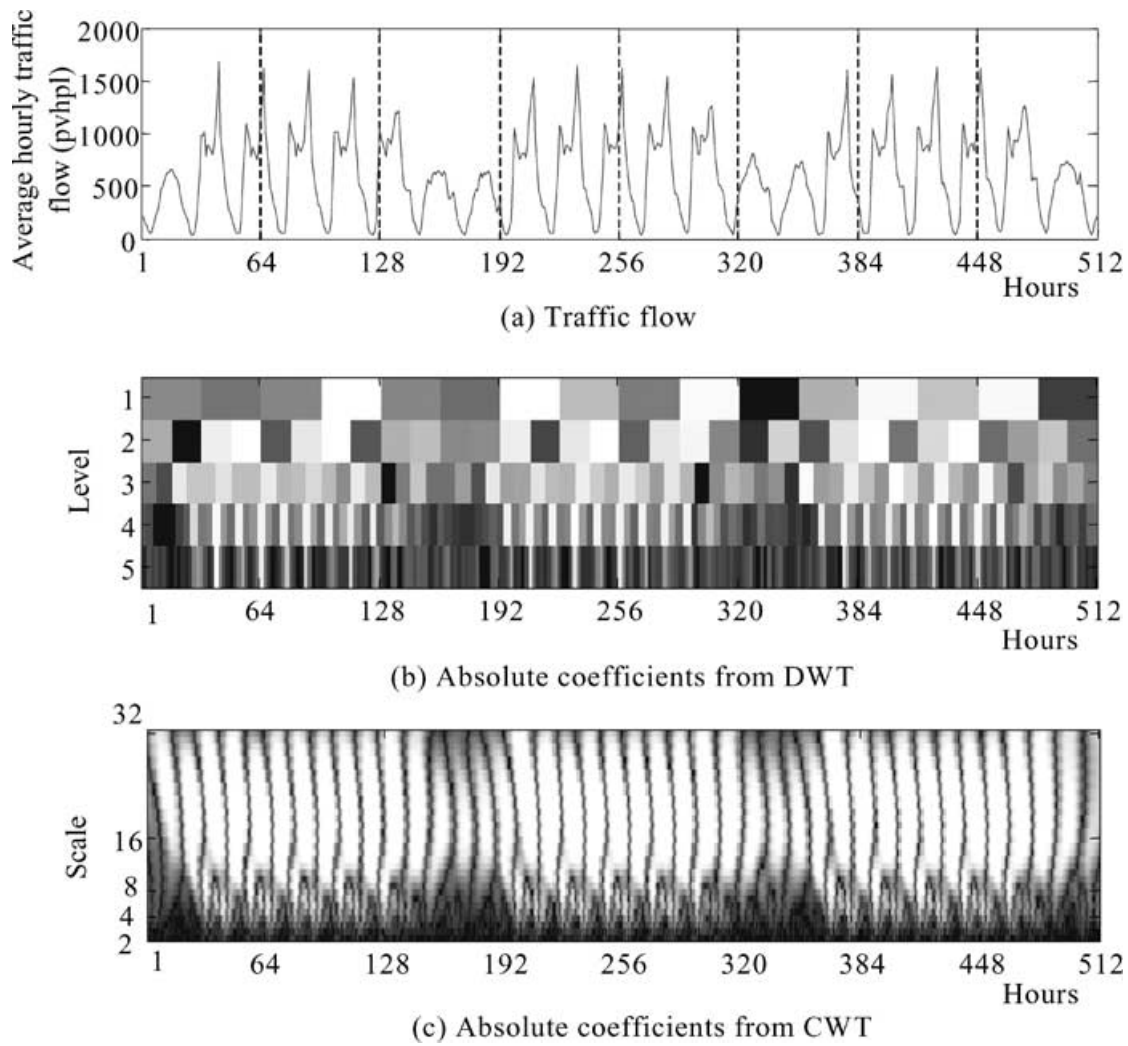
## 7.2 Wavelet multiresolution-ACF analysis

Figure 7 presents a comparison of the DWT and CWT of the first 512 ($2^9$) data points of the traffic flow series shown in Figure 5 using Daubechies wavelet of order 4 (Figure 1). Figures 7b and c show the gray-scaled graphs of MRA for DWT and CWT, respectively. Five decomposition levels are used in the DWT analysis and 32 (equal to $2^5$) scales in CWT. In Figures 7b and c the bright (white) spots represent large traffic flow. The traffic flow decreases as the darkness increases. The gray-scaled graphs in Figure 7 indicate scales and hours at which the traffic flow is large or low.

Figure 7b consists of a number of discrete gray-scaled blocks. It cannot indicate the self-similarity properties in the traffic flow. On the contrary, the CWT coefficients in Figure 7c indicate continuous and subtle information about the traffic flow signals. The CWT multiresolution coefficients map provides a powerful tool for identifying self-similarity and fractal patterns in the traffic flow.

Figure 8 shows the DWT MRA coefficients of the first 512 points of the traffic flow series shown in Figure 5 using Daubechies wavelet of order 4. The DWT coefficients are organized into six series. Five of these series denoted by $D_1$ to $D_5$ are the wavelet coefficients representing the details of five levels of decomposition. They are equivalent to the CWT coefficients with scales of 2 to $2^5 = 32$ with the dyadic factor of 2. It should be noted that in Figure 8, for the sake of visibility of the illustration, the scale of the vertical axis is chosen differently for different levels of details. The details identify subtle characteristics of the traffic flow at different levels. For example, $D_1$ and $D_2$ show the high-frequency and low-amplitude components, roughly reflecting the long-term periodicity of 168 hours (1 week), but $D_4$ characterizes a contour with the periodicity of 24 hours (1 day). The sixth series, denoted by $A_5$, represents the scaling coefficients and the decomposition approximation. Because the average of adjacent elements is used to approximate the traffic flow in the wavelet analysis, the scaling coefficient $A_5$ has the same sample mean as the traffic flow. The approximation series, $A_5$, reflects the long-term periodicity of the traffic flow (1 week). Based on Equation (12), the traffic flow can be numerically represented by:

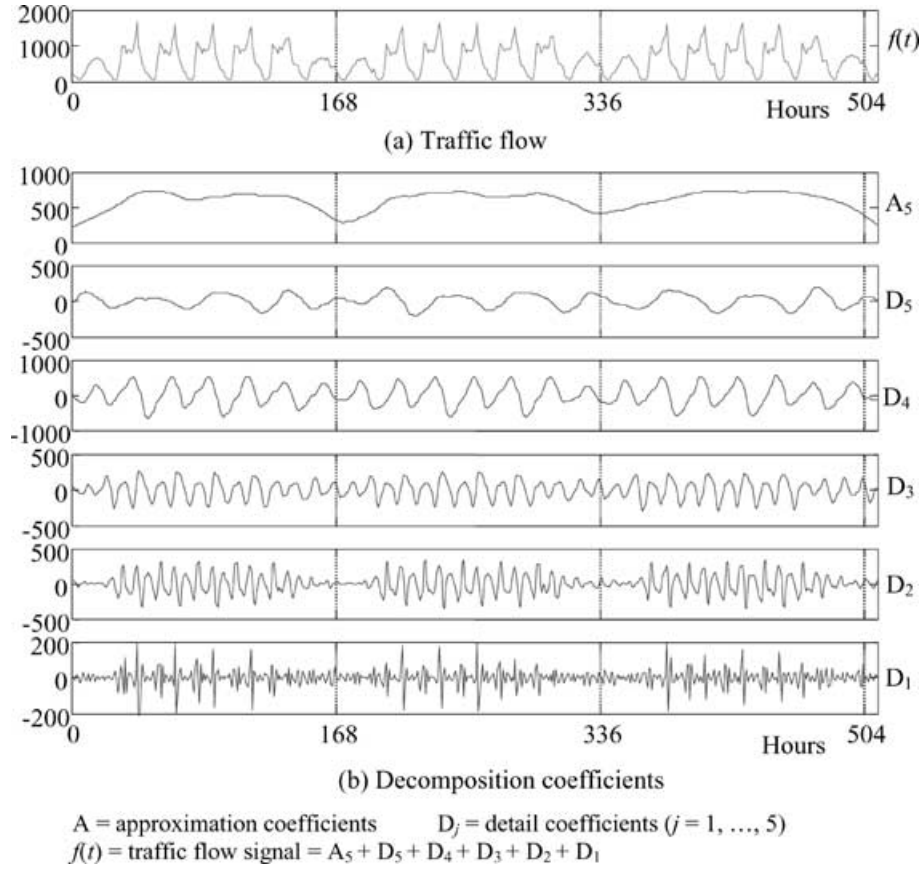$$f(t) = A_5 + D_5 + D_4 + D_3 + D_2 + D_1 \qquad (21)$$

**Fig. 7.** DWT and CWT of the first 512 data points of the traffic flow series shown in Figure 5 using Daubechies wavelet of order 4.

Figure 9 shows the ACFs for details coefficients of the seven-level DWT for the first 512 points of the traffic data (Figure 8a) using Daubechies wavelet of order 4 and a maximum time lag of 50 hours ($\tau_{max} = 50$). This figure is used to determine what level of wavelet decomposition is necessary to be able to identify short-term periodicity. Short-term periodicity can be observed in the ACF of $D_1$ to $D_4$ within the maximum time lag of 50 hours. The period of periodicity is 24 hours, the same as that seen in the original flow data of Figure 5b. It is concluded that four levels of wavelet decomposition are necessary and sufficient to capture and represent the short-term characteristics and periodicity in the traffic flow.
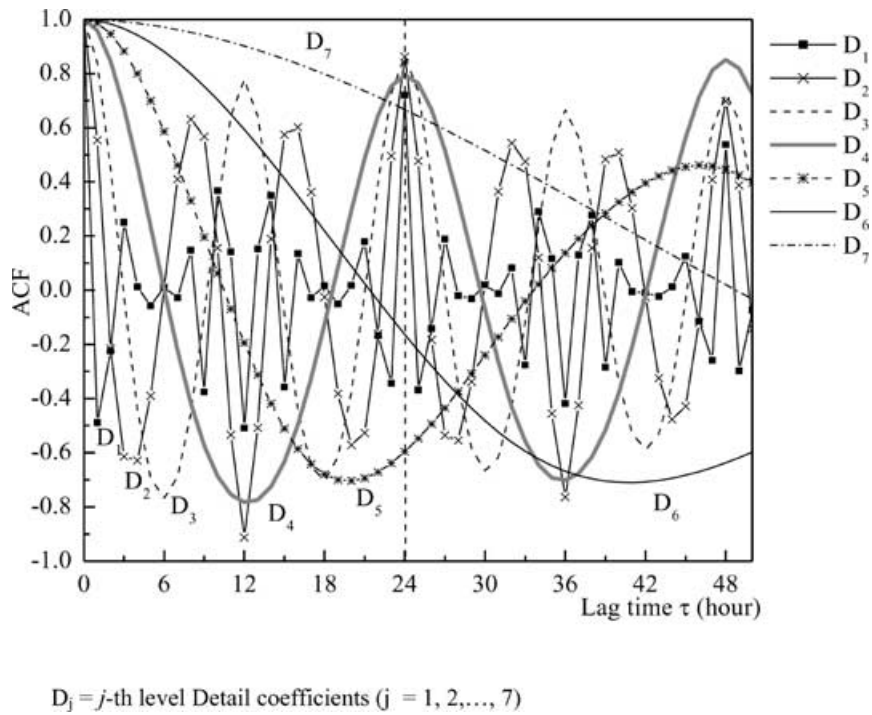
The DWT approximation coefficients for decomposition levels 4 to 7 are shown in Figure 10a. This figure shows that the approximation coefficients can identify the long-term periodicity of the traffic flow when the decomposition level is 4 or 5. The approximation coefficients for decomposition levels 6 and 7 do not display any periodicity and consequently are not necessary for accurate traffic flow signal representation. The ACFs for approximations $A_4$, $A_5$, $A_6$, and $A_7$ with a maximum time lag of $\tau_{max} = 350$ hours are shown in Figure 10b. The long-term periodicity or self-similarity of the traffic flow with a period of 168 hours (1 week) is clearly identified in the ACF of $A_5$ and $A_6$ (Figure 10b). It is concluded that five levels of wavelet decomposition are necessary and sufficient to capture and represent the long-term characteristics and periodicity in the traffic flow.

Based on the ACF analysis of the DWT multiresolution coefficients (details in Figure 9 and approximations in Figure 10), we can conclude that five decomposition

(b) Decomposition coefficients

A = approximation coefficients    $D_j$ = detail coefficients ($j$ = 1, ..., 5)
$f(t)$ = traffic flow signal = $A_5 + D_5 + D_4 + D_3 + D_2 + D_1$

**Fig. 8.** DWT multiresolution analysis coefficients of traffic flow series with five-level details and 5th approximation.



$D_j$ = $j$-th level Detail coefficients ($j$ = 1, 2,..., 7)

**Fig. 9.** Autocorrelation functions for details coefficients of the seven-level DWT using Daubechies wavelet of order 4 ($\tau_{max}$ = 50).

(a) Approximation coefficients



(b) ACFs of approximation coefficients with
maximum lag $\tau_{max} = 350$
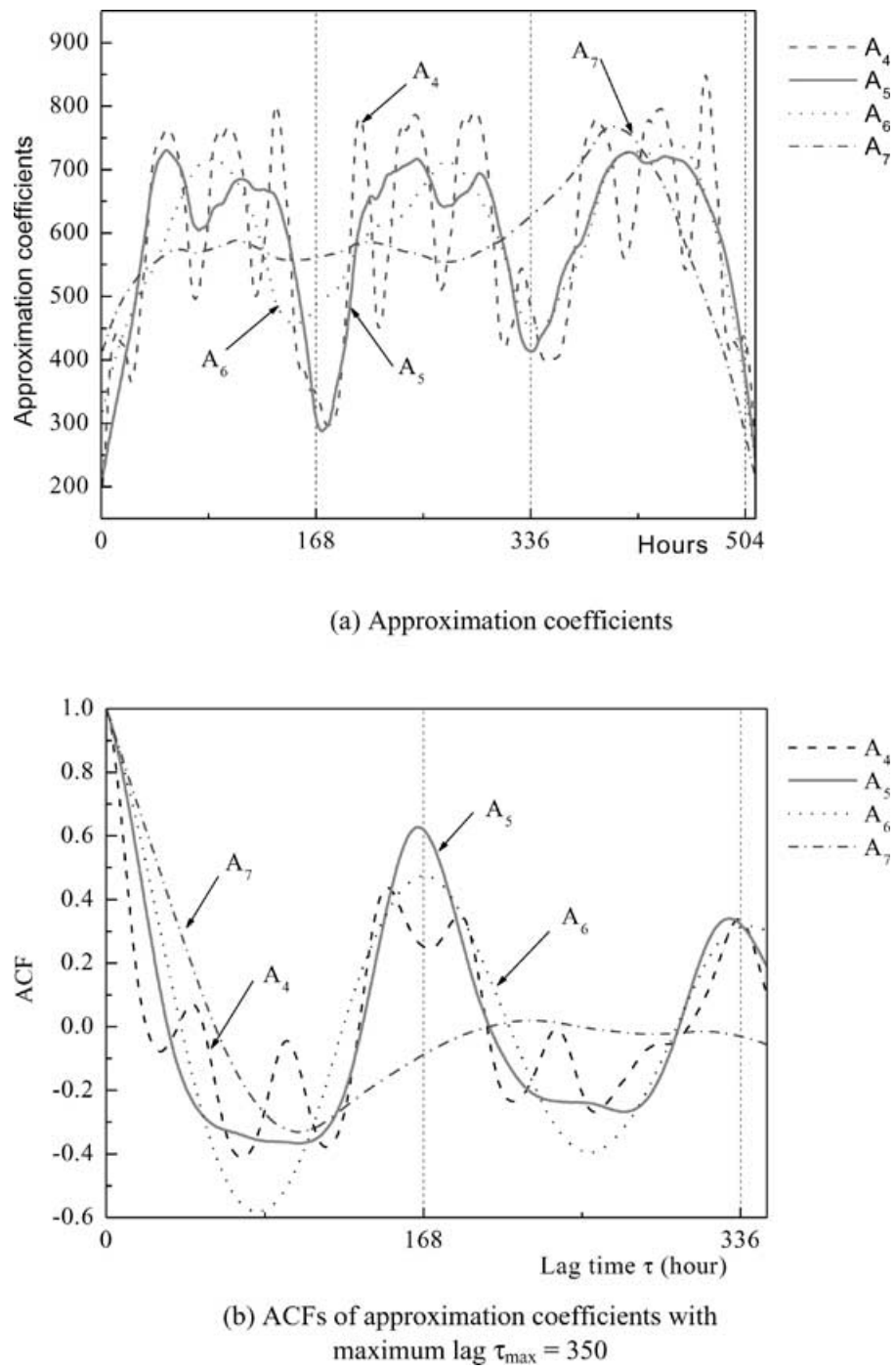
**Fig. 10.** Approximation coefficients with various decomposition levels and their ACFs.

levels capture both short-term and long-term periodicity and, therefore, can adequately characterize the details of the given traffic flow. The same analysis was performed using the DWPT. Results similar to those presented in Figures 8–10 are obtained but not presented in the article for the sake of brevity. However, it is found that three decomposition levels of DWPT can adequately characterize the details of the given traffic flow as opposed to five decomposition levels necessary for DWT. DWPT-based results for denoising and singularity identification using only three decomposition levels are presented in the following section.

## 7.3 DWPT-based denoising and singularity identification

Figures 11b and c show the denoised average hourly traffic flow for the first 512 points of the original traffic flow shown in Figure 11a using Daubechies wavelet of order 4 and DWT and DWPT methods, respectively. Considering the details of traffic flow in part of Figures 11a–c enlarged and enclosed in a rectangular box, it is observed that the DWPT-based denoising approach smoothens the traffic flow more effectively than the DWT-based denoising approach, while at the same time capturing the subtle features in the signal.

The DWPT is used to identify singularity in the traffic flow time series in addition to denoising the data. The three-level wavelet packet decomposition using Daubechies wavelet of order 4 is adopted to identify the singularity in the last 512 traffic flow data points (shown in Figure 5a) in the range 1,312–1,824 (shown in Figure 12a). The data include an atypical pattern, as shown in Figure 5b. The three-level DWPT decomposition for this traffic flow data series is shown in Figure 13. The data are resolved into eight series as shown in Figure 12b and identified at the third level of the de-

composition tree of Figure 13 as $AAA_3$, which represents the third-level approximation coefficients ($A_3$) resulting from the second level approximation ($AA_2$) to $DDD_3$, which represents the third-level details coefficients ($D_3$) resulting from the second level details ($DD_2$), too. Thus, the denoised traffic flow can be represented mathematically by the third-level DWPT decomposition coefficients as follows:

$$x(t) = AAA_3 + AAD_3 + ADA_3 + ADD_3 \\ + DAA_3 + DAD_3 + DDA_3 + DDD_3 \tag{22}$$

Again, in Figure 12b, for the sake of visibility of the illustration, the scale of the vertical axis is chosen differently for different levels of details.

In Figure 12b, the singular oscillatory behavior is clearly identified in the coefficients $AAD_3$ (as well as in $DDD_3$ but not as conspicuously as in $AAD_3$) in the range 1,350–1,357, where there is an abnormal enlargement of the localized oscillatory amplitude compared to other typical traffic patterns. The singularity in the decomposition coefficients results from an atypical increase of the traffic flow (Figure 5b). A similar analysis was performed using the DWT with five levels of decomposition. It is



(a) Original traffic flow

(b) Denoised traffic flow using DWT

(c) Denoised traffic flow using DWPT
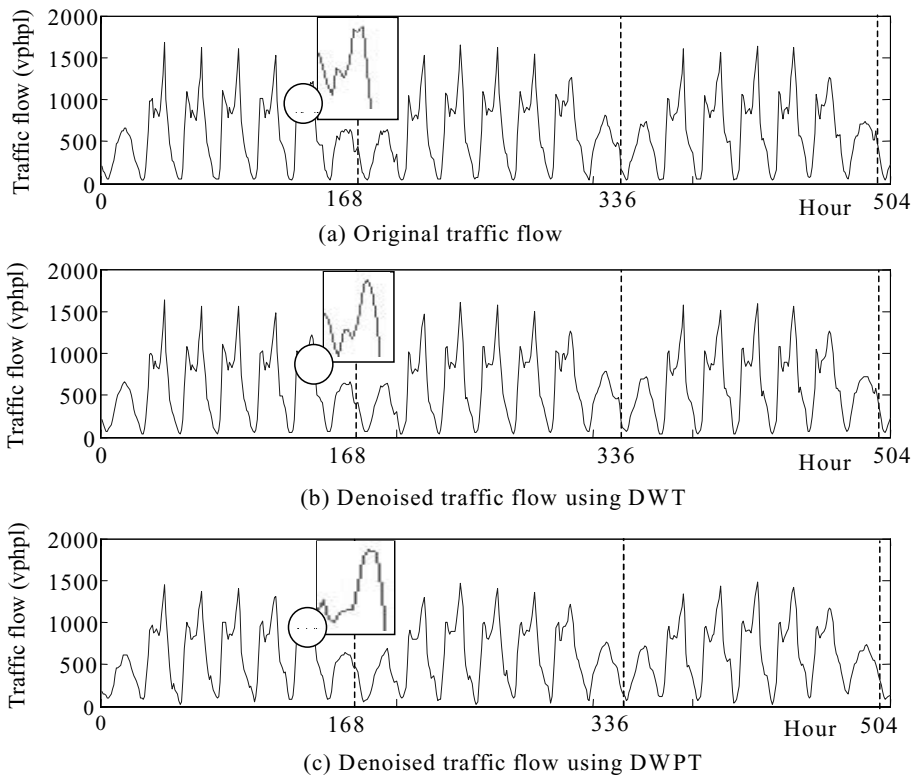
DWT = Discrete wavelet transform          DWPT = Discrete wavelet package transform

**Fig. 11.** Denoising the traffic flow using DWT and DWPT approaches.

(a) Denoised traffic flow with singularity



(b) 3-level decomposition of DWPT
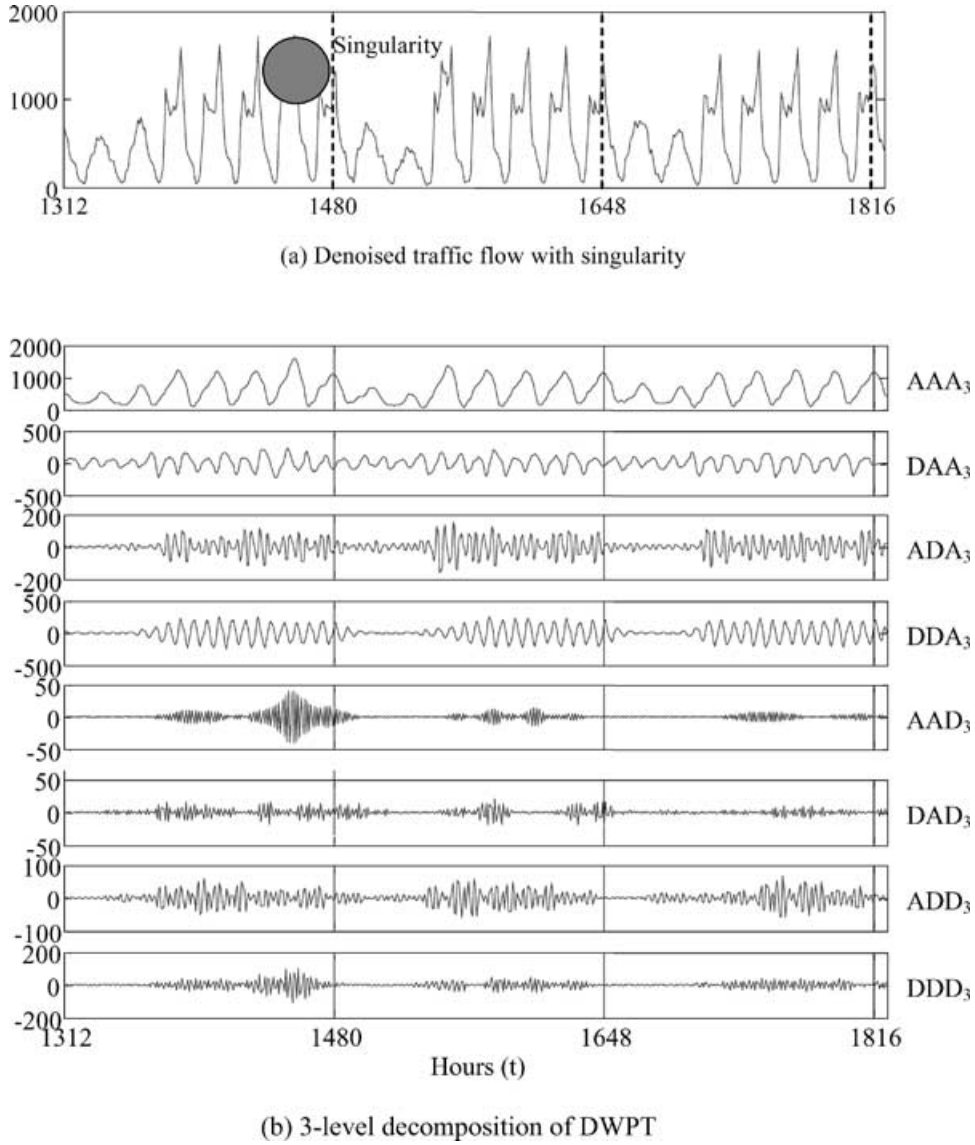
**Fig. 12.** Singularity identification using DWPT and Daubechies wavelet of order 4.
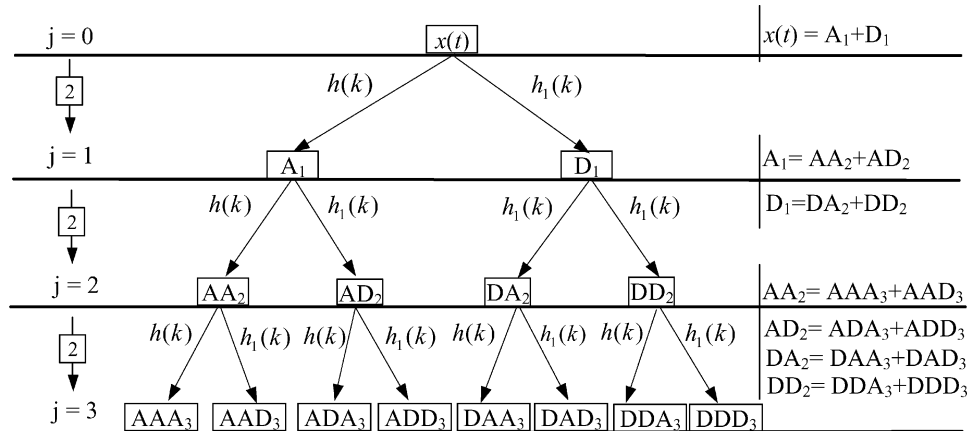


**Fig. 13.** Three-level DWPT decomposition tree of the last 512 traffic flow data points shown in Figure 5a.

found that the DWT with five levels of decomposition cannot identify the singularity. Further, increasing the number of decomposition levels does not help identify singularities.

## 8 CONCLUSION

A hybrid wavelet packet-ACF method is proposed for analysis of traffic flow time series and determining its self-similar, singular, and fractal properties. The proposed method provides a powerful tool in removing the noise and identifying the singularity in the traffic flow. The DWPT analysis provides more coefficients representing additional subtle details of the signal and, therefore, can be used to denoise the signal even more effectively than the conventional wavelet transform. The DWPT also provides greater flexibility for detecting the oscillatory or periodic behavior and the fractal properties of time series. The three-level DWPT analysis can capture the subtle properties of the traffic flow including singularities without any loss of details. The methods created in this research can be used to develop accurate traffic-forecasting models.

## ACKNOWLEDGMENT

## REFERENCES

Adeli, H. & Karim, A. (2000), Fuzzy-wavelet RBFNN model for freeway incident detection, *Journal of Transportation Engineering, ASCE*, **126**, 464–71.

Adeli, H. & Samant, A. (2000), An adaptive conjugate gradient neural network: Wavelet model for traffic incident detection, *Computer-Aided Civil and Infrastructure Engineering*, **13**(4), 251–60.

Barron, A., Birgé, L. & Massart, P. (1999), Risk bounds for model selection via penalization, *Probability Theory and Related Fields*, **113**, 301–413.

Brockwell, P. J. & Davis, R. A. (2002), *Introduction to Time Series and Forecasting*, Springer-Verlag, New York.

Burrus, C. S., Gopinath, R. A. & Guo, H. (1998), *Introduction to Wavelets and Wavelet Transforms: A Primer*, Prentice Hall, Inc., New Jersey.

Chui, C. K. (1992), *An Introduction to Wavelets*, Academic Press, Inc., San Diego, CA.

Coifman, R. R. & Wickerhauser, M. V. (1992), Entropy-based algorithms for best basis selection, *IEEE Transaction on Information Theory*, **38**(2), 713–18.

Daubechies, I. (1988), Orthonormal bases of compactly supported wavelets, *Communication on Pure and Applied Mathematics*, **41**, 909–96.

Daubechies, I. (1992), *Ten Lectures on Wavelets*, Society for Industrial and Applied Mathematics, Phildelphia, PA.

Dendrinos, D. S. (1994), Traffic-flow dynamics: A search for chaos, *Chaos, Solitons and Fractals*, **4**(4), 605–17.

Disbro, J. E. & Frame, M. (1989), Traffic flow theory and chaotic behavior. *Transportation Research Record, TRB*, vol. 1225. National Research Council, Washington DC, pp. 109–15.

Donoho, D. L. (1995), De-noising by soft-thresholding, *IEEE Transactions on Information Theory*, **413**(3), 613–27.

Holschneider, M. (1995), *Wavelets: An Analysis Tool*, Oxford University Press, Oxford, UK.

Karim, A. & Adeli, H. (2002), Incident detection algorithm using wavelet energy representation of traffic patterns, *Journal of Transportation Engineering, ASCE*, **128**(3), 232–42.

Karim, A. & Adeli, H. (2003), Fast automatic incident detection on urban and rural freeways using wavelet energy algorithm, *Journal of Transportation Engineering, ASCE*, **129**(1), 57–68.

Kerner, B. S. (1999), Theory of congested traffic flow: Observation and theory, *Transportation Research Record, TRB*, vol. 1678. National Research Council, Washington DC, pp. 160–67.

Lee, W. & Fambro, D. B. (1999), Application of subset autoregressive integrated moving average model for short-term freeway traffic volume forecasting, *Transportation Research Record, TRB*, vol. 1678. National Research Council, Washington DC, pp. 160–67.

Mallat, S. G. (1989), A theory for multiresolution signal decomposition: The wavelet representation, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, **11**(7), 674–93.

Mallat, S. G. (1999), *A Wavelet Tour of Signal Processing*, Academic Press, San Diego.

Percieval, D. B. & Walden, A. T. (2000), *Wavelet Methods for Time Series Analysis*, Cambridge University Press, Cambridge.

Samant, A. & Adeli, H. (2000), Feature extraction for traffic incident detection using wavelet transform and linear discriminant analysis, *Computer-Aided Civil and Infrastructure Engineering*, **15**(4), 241–50.

Smith, B. L. & Demetsky, M. J. (1997), Traffic flow forecasting: Comparison of modeling approaches, *Journal of Transportation Engineering, ASCE*, **123**(4), 261–66.

Williams, B. M., Durvasula, P. K. & Brown, D. E. (1998), Urban freeway traffic flow prediction: Application of seasonal autoregressive integrated moving average and exponential smoothing models, *Transportation Research Record, TRB*, vol. 1644. National Research Council, Washington DC, pp. 132–41.