A Comparative Analysis of Coronary and Aortic Flow Waveforms

CHARLES Q. DANG,[1] ZORAN NENADIC,[2] and GHASSAN S. KASSAB[1,3,4]

[1]Department of Biomedical Engineering, Indiana University Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202, USA; [2]Department of Biomedical Engineering, University of California, Irvine, CA, USA; [3]Department of Surgery, Indiana University Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202, USA; and [4]Department of Cellular and Integrative Physiology, Indiana University Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202, USA

Abstract—Empirical Mode decomposition (EMD) is a mathematical tool designed to analyze non-stationary, non-linear stochastic waves. EMD separates a waveform into its constituent modes of oscillations or intrinsic mode functions (IMFs) and provides meaningful definitions of instantaneous frequency, instantaneous energy, mean trends and oscillation about the mean trends. This study provides a detailed mathematical analysis of blood flow waveforms in the porcine left anterior descending artery and aorta using EMD. Flow data with non-stationary and non-linear characteristics were obtained for several hours using an implanted wireless biotelemetry device. EMD was validated against modern numerical techniques of principal component analysis (PCA) and wavelet analysis by comparing their predicted mean trends and energy distribution. EMD has an advantage over both techniques since it combines the strengths of both: it is adaptive (similar to PCA), and it can define instantaneous frequencies (similar to wavelet analysis). Because of the iterative nature, however, calculations using EMD can be computationally intensive. Sampling rate reduction was used to reduce computation time, without significantly effecting accuracy of IMF calculations. It was found that IMFs calculated at a sampling rate as low as 20 Hz were not significantly different ($<6\%$) from those obtained at the original sampling rate (200 Hz). Our findings suggest that EMD may be a powerful mathematical tool to characterize flow waveforms.

Keywords—Fourier analysis, Empirical mode decomposition, Principal component analysis, Wavelet analysis, Shear stress.

INTRODUCTION

Blood flow is an important determinant of the shear stress acting on the endothelium surface of the vessel and exhibits temporal and spatial variation. Flow and shear stress influence the homeostasis of endothelial cells, mechanotransduction, tone of blood vessel, growth and remodeling of blood vessel wall, and initiation and development of cardiovascular diseases. They have been found to play a critical role in atherogenesis[2,4,5,6,18,22,25] with numerous biomechanical connections to molecular and cellular studies of the endothelial and vascular smooth muscle cells.[9,10,12] Despite these advances, the dynamics of blood flow remain incompletely understood.

Several mathematical methods may be used to analyze blood flow. The periodic nature of blood flow waveforms invites Fourier analysis and Fourier-based techniques as a first approach. Most Fourier-related methods have shortcomings, however, due to their inability to address the nonlinear and nonstationary properties (the heart rate is a variable parameter) of blood flow waveforms. Using examples from either blood pressure or flow data, it can be verified that these waves vary somewhat in frequency and in shape from one heartbeat to the next. Since these variations are not necessarily cyclic, they appear as spurious high frequency components that are difficult to separate from high frequency noise in the Fourier frequency spectrum. Also, when the Fourier method is used to analyze frequency, time information is lost. In summary, events may be identified by frequency using Fourier analysis, but it is impossible to determine the time course of those events. The windowed or Short-Time Fourier transform (STFT) was developed to address this shortcoming. This technique involves application of the Fourier transform locally to short time segments, or time windows, instead of globally to the entire data.[11] This approach provides both frequency and time information, but with precision limited by the size of the time window. Shorter time windows provide greater resolution of higher frequencies, while longer time windows improve resolution of lower frequencies. Hence, for physiologic data, STFT must compromise between resolving mean trends (lower frequencies) and resolving oscillations (higher frequencies).

Address correspondence to Ghassan S. Kassab, Department of Biomedical Engineering, Indiana University Purdue University Indianapolis (IUPUI), Indianapolis, IN 46202, USA. Electronic mail: gkassab@iupui.edu

Current popular numerical techniques such as principal component analysis (PCA) and wavelet analysis can handle nonlinear and nonstationary data, but these and other techniques including the Wigner-Ville[7] distribution and the evolutionary spectrum[23] have some shortcomings in comparison with Empirical Mode Decomposition (EMD).[14] EMD represents a departure from the conventional Fourier analysis and is powerful in its ability to analyze both nonlinear and non-stationary waveforms. It is not limited by a dependence on fixed frequencies or fixed time windows, such as with Fourier analysis and STFT. Furthermore, the two criteria for the sifting of intrinsic mode functions (IMFs) (see Methods) allow for the application of the Hilbert transform to the IMFs, leading to a mathematically meaningful definition of instantaneous frequency.[14] The application of EMD to flow waves in the coronary artery and aorta is explored in this study.

EMD is a numerical method designed to analyze non-stationary and non-linear data.[14] EMD is used to parse wave data into characteristic components, or IMFs. These IMFs describe the modes of oscillation inherent in a wave. EMD was initially applied to the study of ocean waves.[13] The first application of this method to biological waves was in the study of blood pressure waves in rats.[15] Recently, it has been used to study EMG[26] and ECG[24] waves. To our knowledge, there are no known applications of EMD to blood flow waves, nor has there been any analysis of blood flow waves in the awake, free ranging condition.

We recently developed a fully implantable telemetric system in chronically instrumented awake free ranging swine for measurement of aortic and coronary pressures and flows.[3] The measurements were recorded over several hours or days and flow waveforms were found to be non-stationary and non-linear stochastic functions. The conventional analyses that assume stationary random oscillations are not suitable to analyze the non-stationary data acquired. In this study, IMFs were determined for left anterior descending (LAD) artery and aortic flow velocity data. IMFs were used to derive mean trends and oscillations about the mean trends. The results from EMD were compared with principal component analysis and wavelet analysis as validation.

We found that EMD has several advantages over conventional methods and provides a powerful tool to describe flow waveforms with mathematical accuracy. The major shortcoming of EMD is currently the computation expense. Analysis of lengthy data sets (5 h or more at 200 samples/s) requires downsampling of data in order to achieve acceptable computation times. The effect of downsampling to 40 samples/s was found to have insignificant effect on accuracy of the EMD technique.

METHODS

The flow measurements were made by a recently developed fully implantable, wireless telemetric system in awake, free ranging animals.[3] Briefly, the telemetry implant consisted of up to 9 independent measurement channels: 1 ECG, 1 temperature, 3 pressure (Millar transducers, Millar Instruments Inc., Houston) and 4 Doppler based flow channels (PZT5-A piezoelectric transducer, Iowa Doppler Products, Iowa City). Five male swine weighing 32 ± 3 kg were sedated with ketamine (20 mg/kg) and atropine (0.05 mg/kg). The animals were intubated, and anesthesia was induced with 1–2% isoflurane. An opening was made in the fourth left lateral intercostal space. A pressure probe was inserted in the descending aorta to monitor systemic pressure. A Doppler-based ultrasonic blood velocity flow probe was inserted at the aortic arch, facing the ascending aorta. Blood flow velocity in the LAD artery was measured by a Doppler based ultrasonic probe secured to the outer wall of the vessel. A transmitter device connecting all the probes was positioned 15 cm away from the chest wound and secured subcutaneously. Measurements from the Doppler and pressure probes were transmitted wirelessly to a base station located nearby. Measurements were recorded at 200 samples/s using a data acquisition system (Biopac Systems, Goleta, CA). The telemetry device was implanted in five animals and recorded intermittent data on a daily basis. Continuous recordings exceeding 10 h were only obtained in one animal due to concerns that extended recordings would significantly reduce battery life and possibly prevent telemetric measurement at later periods. Telemetry data from a representative animal within the first week post-implantation is reported in this study.

IMFs of the flow waves were determined by an iterative process called sifting. The details of the sifting process are described by Huang *et al.*,[14] and implementation details are discussed in the Appendix. Briefly, a flow wave data was fitted with upper and lower envelope functions. The upper envelope function was found by fitting the local maxima of the wave with a cubic spline. The lower envelope function was similarly found by fitting the local minima. The average of the upper and lower envelope functions was then subtracted from the original wave; i.e., the original wave was "sifted". The wave resulting from the sifting process may be considered an IMF only if two criteria are met: (1) the wave has a local mean of zero; i.e., its upper and lower envelopes are symmetric and (2) the total number of extrema (maxima and minima) differ from the total number of zero crossings by no more than one. If the IMF criteria are not met, the wave is iteratively sifted. Once an IMF was found, it is subtracted from

the original wave and the resultant wave was sifted to produce the next IMF. In this manner, each successive IMF will have fewer oscillations than the preceding IMF. Hence, the IMFs are generated in increasing order of smoothness. When sifting is no longer possible, the result is generally a monotonic signal called a residual. Consequently, the residual is often interpreted as the mean trend in the data. Sorting IMFs according to smoothness is one possible way of ranking. An alternative sorting method is to rank IMFs according to their energy content. The details and the rationale for this approach are given in the Appendix.

The calculations were performed on a Pentium 4 Dual Core 3.0 GHz computer equipped with 3 GB of RAM. The EMD software used was HHT-DPS version 1.4 provided by NASA. A brute force computation of IMFs of large data sets is time consuming and can run for days or weeks. Sampling rate reduction was used to reduce computation time. To assess the effectiveness of this strategy, a comparison was made between the residual wave obtained from sampling rate reduction and the residual obtained from a brute force computation. To alleviate the computational burden, a 1-h subset of LAD blood flow velocity data, acquired at 200 Hz, was used as a reference wave. The sampling rate of the reference wave was reduced to 100, 40, and 20 Hz. IMFs were determined at each sampling rate, along with mean trends and oscillations about the mean trends. The effect of sampling rate on the magnitude and appearance of the IMFs was determined by comparing the residuals obtained at sub-sampled frequencies with those obtained at the original sampling rate.

Decomposition of the data was also carried out using PCA and wavelet analysis. PCA was performed by dividing the data into non-overlapping segments and by removing the mean of the segments. The covariance matrix of the population of segments was then calculated. Principal eigenvalues of the covariance matrix represent the energy of signal modes obtained by projecting the data to the corresponding eigenvectors (principal components). By design, the signal modes are uncorrelated, and so their partial sum approaches the original signal as the number of modes is increased (similar to the EMD technique). The implementation details are given in the Appendix and the analysis was performed with custom-written MATLAB^{TM} (The MathWorks, Inc., Natick, MA) scripts.

Wavelet analysis was performed by applying the discrete wavelet transform of the family coif5 (MATLAB^{TM} Wavelet Toolbox) to the data. The scales were ranked according to the total energy carried by the corresponding wavelet coefficients. Therefore, the signal was decomposed into constituent modes, obtained by the inverse transform at these scales. By increasing the number of scales in the inverse transformation, the

accuracy of signal reconstruction improves (similar to EMD and PCA). The implementation details are given in the Appendix.

RESULTS

The time required to compute IMFs was seen to increase exponentially with increasing size of data. If we consider only data at the same sampling rate of 40 Hz, it takes approximately 1 s to calculate IMFs for 10-s data, 3 min to process 10-min data, 40 min to process 1-h data and 48 h to process 10-h data. Hence, the effect of sampling rate on the accuracy of IMF calculation was determined. As explained in the previous section, calculating IMFs from full data was time-consuming, and therefore a 1-h long subset of LAD flow data was used. Figure 1 shows the percent difference in residuals obtained at lower sampling rates compared to residuals obtained at the original sampling rate of 200 Hz. The differences are within 6%. A sampling rate reduction as low as 40 Hz can still reproduce the residuals at higher sampling rates, with a difference within 3%. This result may be extended to a 24-h data set, where the data may be resampled at 40 Hz to significantly reduce computation time. Similar results were observed with aortic flow data, but were not presented in the interest of space.

EMD was performed on 10-h LAD and aortic flow velocity data down-sampled at 40 Hz. Figure 2a shows the resulting IMFs and residual obtained for 10-h LAD flow data. These EMD components were sorted in decreasing order of energy content for comparison with PCA and wavelet analysis. To appreciate the relationships between various signal components and the original waveform, the insets showing the representative 6 s of data are also provided. Note that the residual (mean trend) exhibits a steady flow increase in
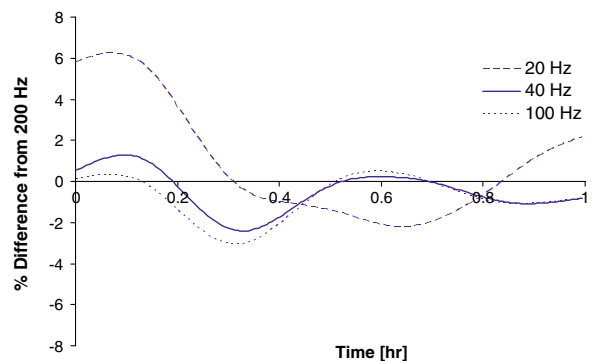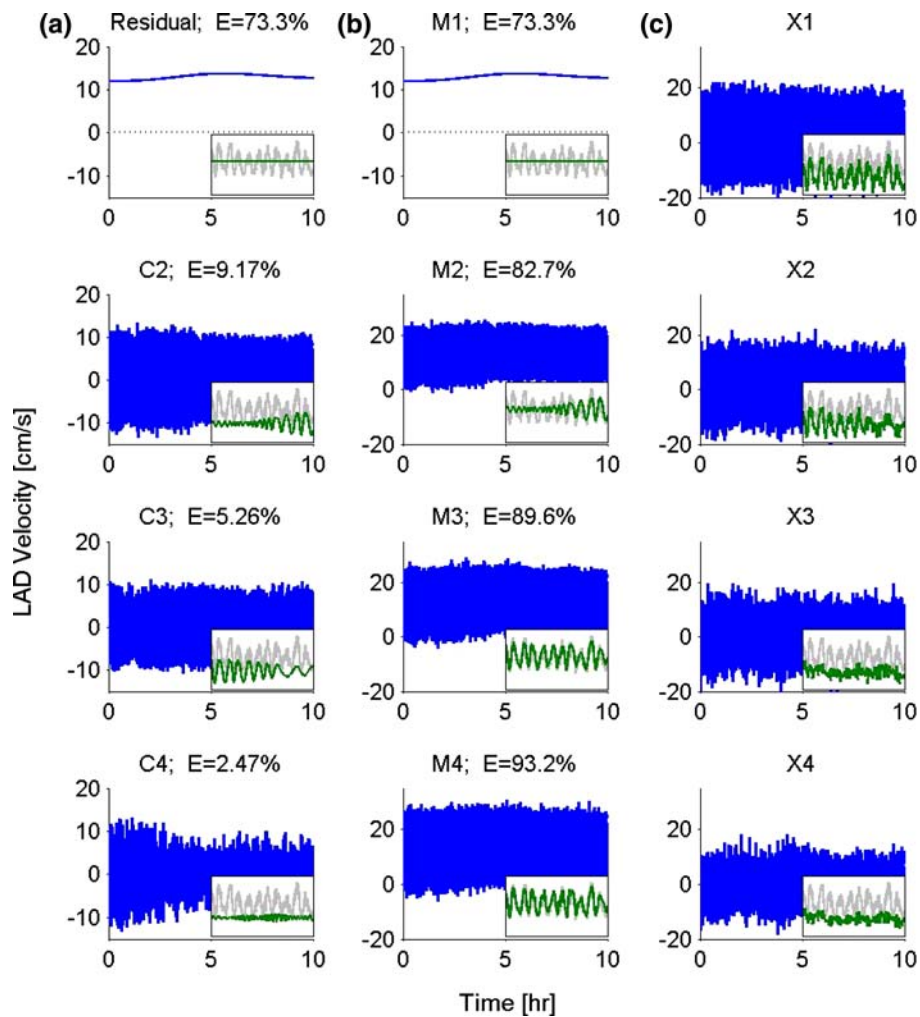

FIGURE 1. Effect of decreased sampling rate. Percent difference between subsampled residuals and residuals obtained at reference sampling frequency (200 Hz) for 1 h of LAD flow data.

FIGURE 2. EMD of LAD data. (a) EMD components comprising IMFs and the residual; (b) partial reconstructions (mean trends); and (c) oscillations around the mean trends of 10-h LAD flow velocity data subsampled at 40 Hz. *E* is the fraction of total energy contained in the mode of oscillation or mean trend. Insets correspond to 6 s of data, with the original signal given in gray.

the first 5 h, followed by a steady decrease towards initial values. Also note that the residual carries 73.3% of signal's energy and is by far the most energetic component. The mean trends and oscillations around the mean trends (see Appendix for detailed explanation) are shown in Figs. 2b and 2c, respectively. The 4th order mean trend M4 captures 93.2% of the original waveform energy, and provides a very good reconstruction of the original waveform (see inset). Consistent with this observation is the appearance of the oscillation X4, containing mostly high frequency details of the original wave (inset).

The IMFs of aortic flow velocity data are shown in Fig. 3a. Note that unlike in the case of LAD flow data, the residual is not the component with the highest energy. Rather the mode of oscillations, C1, matching the frequency of the cardiac cycle is the highest ranking mode. The residual in this case exhibits roughly a constant value for the first 5 h followed by a gradual

decrease in the subsequent 5 h. Successive partial reconstructions and their corresponding oscillations are given in Figs. 3b and 3c, respectively. While M4 carries only 80.4% of the total signal energy, it nevertheless provides a reasonable approximation of the original wave (see inset).

Short IMF segments from aortic and LAD data are compared in Fig. 4a. Their C5 IMF components were selected for comparison because of their resemblance to the heartbeat. Figure 4a confirms that the oscillations in aortic and coronary flow velocities are out of phase, as expected. A frequency-energy-time plot or Hilbert spectrum calculated for 5-min LAD flow velocity data is shown in Fig. 4b. The *x*-axis is time, the *y*-axis is frequency and the intensity corresponds to the energy of oscillation. It can be inferred that the dark band (indicated by arrow) near the bottom of Fig. 4b is the heart rate. The Hilbert spectrum for aortic data is similar (not shown).
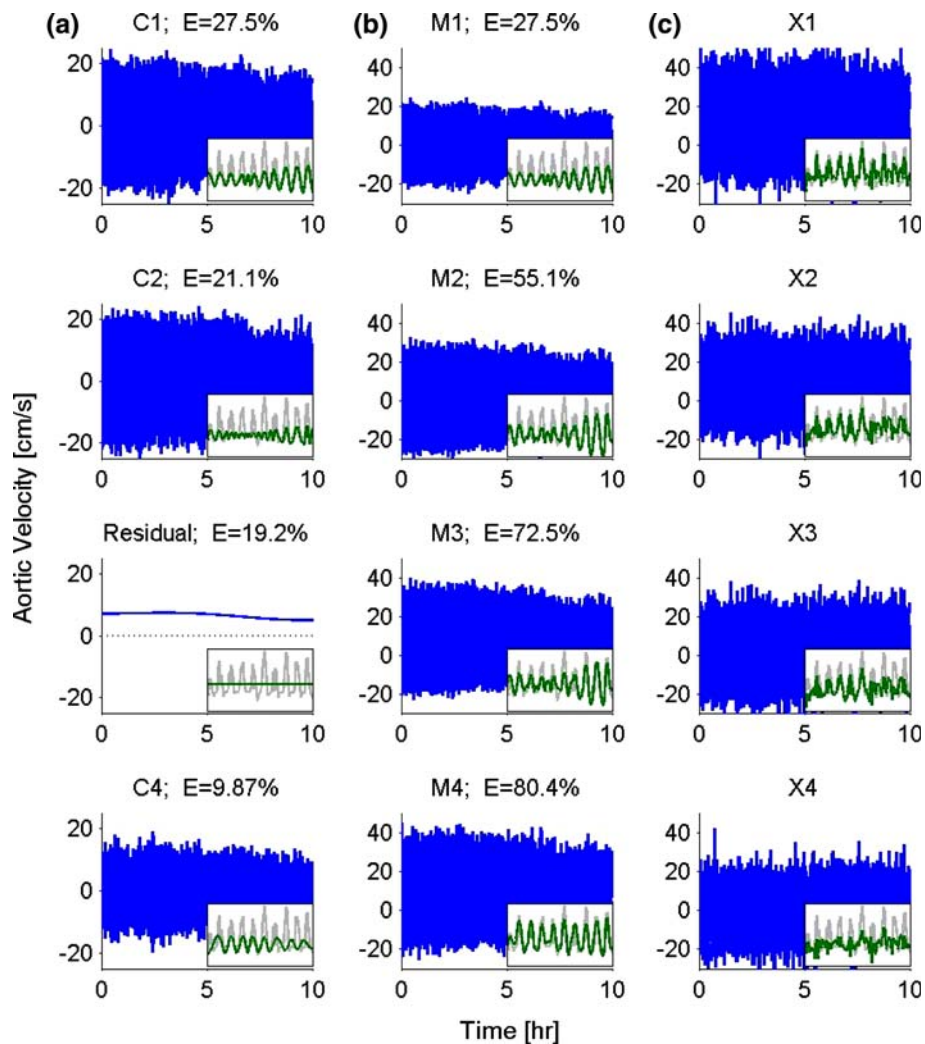
FIGURE 3. EMD of aortic data. (a) IMFs and the residual; (b) partial reconstructions; and (c) the corresponding oscillations of 10-h aortic flow velocity data subsampled at 40 Hz. *E* is the fraction of total energy contained in the mode of oscillation or partial reconstruction. Insets correspond to 6 s of data, with the original signal given in gray.

PCA of LAD data is shown in Fig. 5. The first 3 principal components, ranked by relative energy content, are shown in Fig. 5a. Similar to the residual in EMD analysis, the mean, shown at the bottom of Fig. 5a, has the highest energy contribution (~73.8%), but was placed out of order from the principal components. It is not a true principal component because it was subtracted from the original data, before the principal components were calculated. Figure 5b shows the energy content of the mean trends, which are calculated as partial sums of the mean and modes of oscillation (see Appendix for details). Oscillations around the mean trends are shown in Fig. 5c. Similarly, PCA of aortic flow velocity data is shown in Fig. 6. It is interesting to note that the mean with ~18% of energy is the 3rd highest component, which is quite similar to the results of the EMD analysis (see Fig. 3 for comparisons). Mean trends and oscillations

around the mean trends are shown in Figs. 6b and 6c, respectively. From the insets, it can be seen that the most energetic modes of oscillation for both the LAD and aortic data have frequencies corresponding to the heart rate (1.6 Hz or 96 bpm).

Wavelet analysis of LAD and aortic data reveal modes of oscillation similar to those found with PCA. These modes of oscillation are classified according to scale, a quantity that is closely related to the frequency (see Appendix), and are shown in Figs. 7 and 8. The scales are ordered according to relative energy contribution. Unlike the PCA mean, which was just an arithmetic mean subtracted from the data before analysis, the mean component was determined as a direct result of wavelet analysis, e.g. by reconstructing the signal at a very coarse scale. Therefore, the wavelet mean component can be considered a genuine mode of the signal, much like the residuals in the EMD analysis,

(a)

(b)
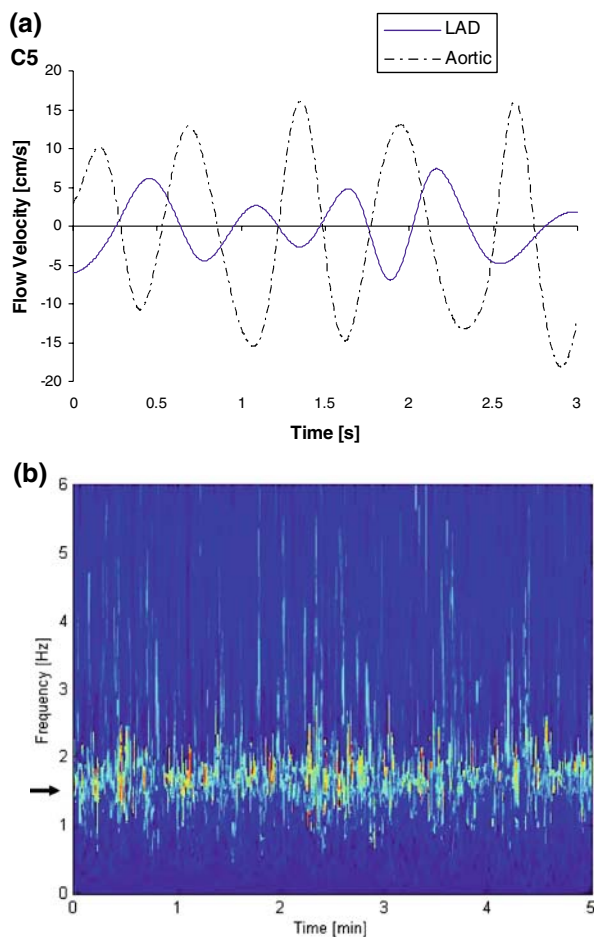
FIGURE 4. (a) Comparison of the intrinsic mode function C5 of aortic and LAD flow velocities; (b) Hilbert spectrum of 5-min LAD flow velocity data. Arrow indicates instantaneous frequencies concentrate about the heart rate (1.6 Hz).

and is ranked along with the other components. The mean component of the LAD data, C1, (see Fig. 7) is the most energetic component, and provides a similar energy contribution (~74.5%) as the residual in the EMD analysis (see Fig. 2), and the mean value in the PCA analysis (see Fig. 5). The mean component of aortic data, C2, (Fig. 8) is only a secondary component, and its energy contribution (~19.9%) is similar to those of the residual (see Fig. 3) and the mean value (see Fig. 6). This drop in the ranking of the mean trend in the aortic velocity data may be attributed to the higher overall mean of LAD velocity (~13 cm/s) compared to aortic velocity (~6.3 cm/s). Wavelet analysis confirms that the most energetically significant oscillations are coupled to the heart rate, in agreement with EMD and PCA.

A comparison of the mean trends of coronary and aortic flows obtained through EMD, wavelet analysis and PCA are shown in Figs. 9a and 9b, respectively. The mean trend from wavelet analysis of LAD data (dotted line) is almost identical to the mean trend from EMD (solid line) as shown in Fig. 9a. Figure 9b shows deviations between the mean trends of EMD (solid line) and wavelet analysis (dotted line) of aortic flow data. In both figures, the PCA mean trend (dashed line) is a simple arithmetic mean, and is insensitive to local variations.

One representative period of LAD velocity is shown in Fig. 10. The corresponding IMFs are shown in Fig. 11a (left panel), mean trends in Fig. 11b (middle), and oscillations about mean trends in Fig. 11c (right).

DISCUSSION

Data from this study were acquired continuously for several hours at a time. EMD analysis required significant computation time, which made it necessary to employ sampling rate reduction to achieve acceptable computation times. It was determined that IMFs calculated at a sampling rate as low as 40 Hz were not significantly different (<3%) from those obtained at the original sampling rate (200 Hz). Improvements in computer processing power in the future will enable IMF calculation without sampling rate reduction.

Currently, the most widely used technique to analyze non-stationary data is wavelet analysis. Wavelet analysis decomposes a flow wave into its constituent basis functions, called wavelets. Unlike sines and cosines of Fourier analysis, the wavelets of compact support are well-suited to analysis of nonstationary data. This feature allows time-frequency events to be resolved simultaneously (subject to Heisenberg uncertainty principle). While wavelet analysis provides better (more uniform) resolution across frequencies than does STFT, the resolution is inferior to that of EMD.[14] Another difference between EMD and wavelet analysis is that EMD gives rise to data dependent representation basis, whereas wavelet analysis uses *a priori* chosen set of basis functions. While the results in the present study were not significantly dependent upon the choice of wavelet basis (see Appendix), the *a priori* chosen wavelet basis may in principle cause a bias, and represents a potential weakness compared with the adaptive basis generated by EMD.

Another useful numerical decomposition technique is the principal component analysis (PCA). Like EMD, it does not make any assumptions about the shape of the basis functions, i.e., the basis is data dependent. Therefore, PCA is fully adaptive as it relies on the data alone to describe its modes of oscillation. It is often used for data compression and denoising, where it extracts the modes (principal components) that account for most of the variance in the data.
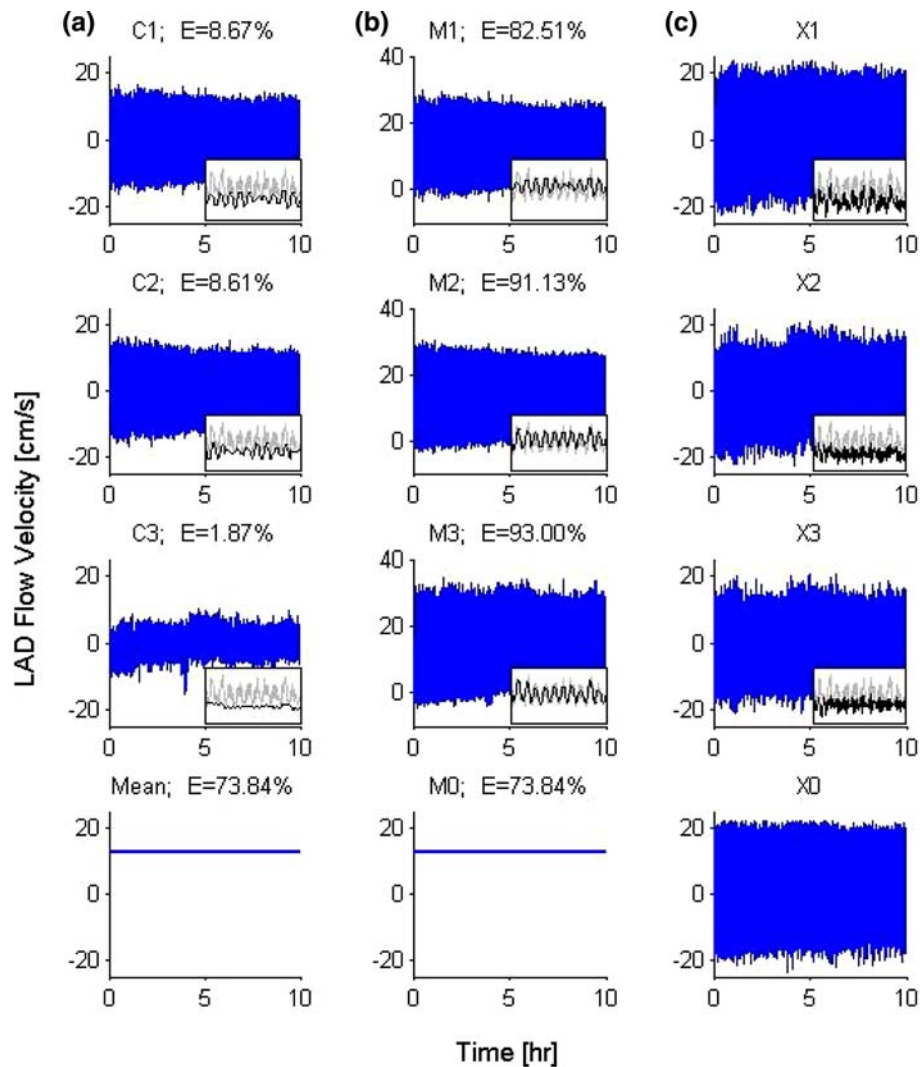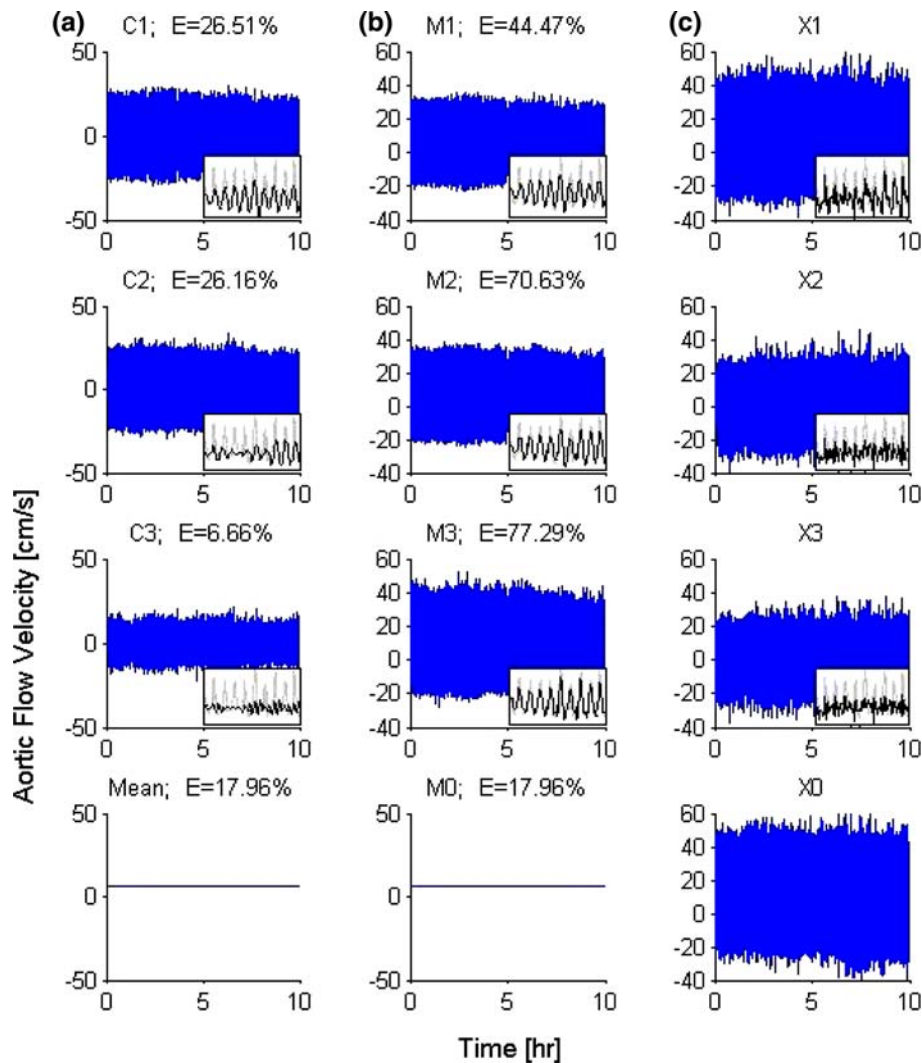
FIGURE 5. PCA of LAD flow velocity data. (a) The first 3 modes of oscillation and the mean value; (b) mean trends; and (c) oscillations about the mean trends. E is the fraction of total energy contained in the mode of oscillation or mean trend. Insets correspond to 6 s of data, with the original signal given in gray.

The results from both wavelet analysis and PCA were compared with EMD of LAD and aortic flow data. In general, there was agreement in the energy distribution across the first 4 components of all 3 techniques. For LAD flow data, all the techniques pick the mean trend as the most energetic mode, accounting for ~73–74% of the total signal energy (Figs. 2, 5, and 7). The distribution of energy among oscillatory components is somewhat different, though. PCA yields 2 components with nearly identical energy contribution of ~8%, whereas in EMD and wavelet analyses the energy of one component is nearly twice as high as that of the next component (~9% vs. ~5% and ~12% vs. ~7%, respectively). In any case, these components seem to capture the heart rate frequency, as can be seen in the insets of Figs. 2, 5, and 7, and in more details in Fig. 4b. Finally, the last components for all 3

techniques seem to carry very little energy (~2%), and are mostly responsible for high-frequency details (see insets). Based on this observation, we conclude that low energy components are mostly representing the noise in the waveform, and are therefore excluded from the present analysis.

Similar conclusions can be drawn based on the analysis of aortic flow data. EMD and PCA rank the mean trend in the 3rd place with ~19 and ~18% of the total energy, respectively; while wavelet analysis ranks the mean trend in the 2nd place, with ~20% contribution (see Figs. 3, 6, and 8). Likewise, the oscillatory components in EMD analysis and PCA are fairly balanced (~27% vs. ~21% and ~26% vs. ~26%, respectively), whereas those of wavelet analysis are largely disproportionate (~40% vs. ~20%). Similar to LAD flow data, the first two oscillatory components

FIGURE 6. PCA of aortic flow velocity data. (a) The first 3 modes of oscillation and the mean value; (b) mean trends; and (c) oscillations about the mean trends. *E* is the fraction of total energy contained in the mode of oscillation or mean trend. Insets correspond to 6 s of data, with the original signal given in gray.

are sufficient to capture the majority of the energy at the frequency of the heart beat (see insets). Finally, note that unlike in the case of LAD flow data, the last components in all 3 techniques are quite significant with energy contributions of ~10, ~7, and ~12%, for EMD, PCA and wavelet analysis, respectively. This is consistent with a lack of a dominant mean trend in the aortic data, meaning that a large portion of energy has to be accounted for by oscillatory components.

A detailed comparison of the mean trends from all three techniques revealed further similarities between EMD and wavelet analysis (Figs. 9a and 9b). Since the PCA mean trend is a simple arithmetic mean of the total data, it does not capture the general increase in LAD flow velocity (Fig. 9a) or the general decrease in aortic velocity (Fig. 9b) that both EMD and wavelet

analysis predict. A drawback of the wavelet mean trend is an edge effect where the mean has abrupt changes at 0 and 10 h (Fig. 9). These abrupt changes do not appear to be physiologic, but are present in the wavelet means of both LAD and aortic data. The EMD mean trend does not show any apparent edge effects.

This study shows that EMD may be applied to lengthy recordings of flow data. The results of EMD analysis are consistent with those of well-known (statistical) signal processing techniques, such as PCA and wavelet analysis. EMD method seems to combine the best qualities of the two techniques: it yields a data-dependent (adaptive) basis, reminiscent of the PCA basis, while being able to account for non-stationary trends in the data, which is the strength of compactly supported wavelets. When applied to flow waves in a
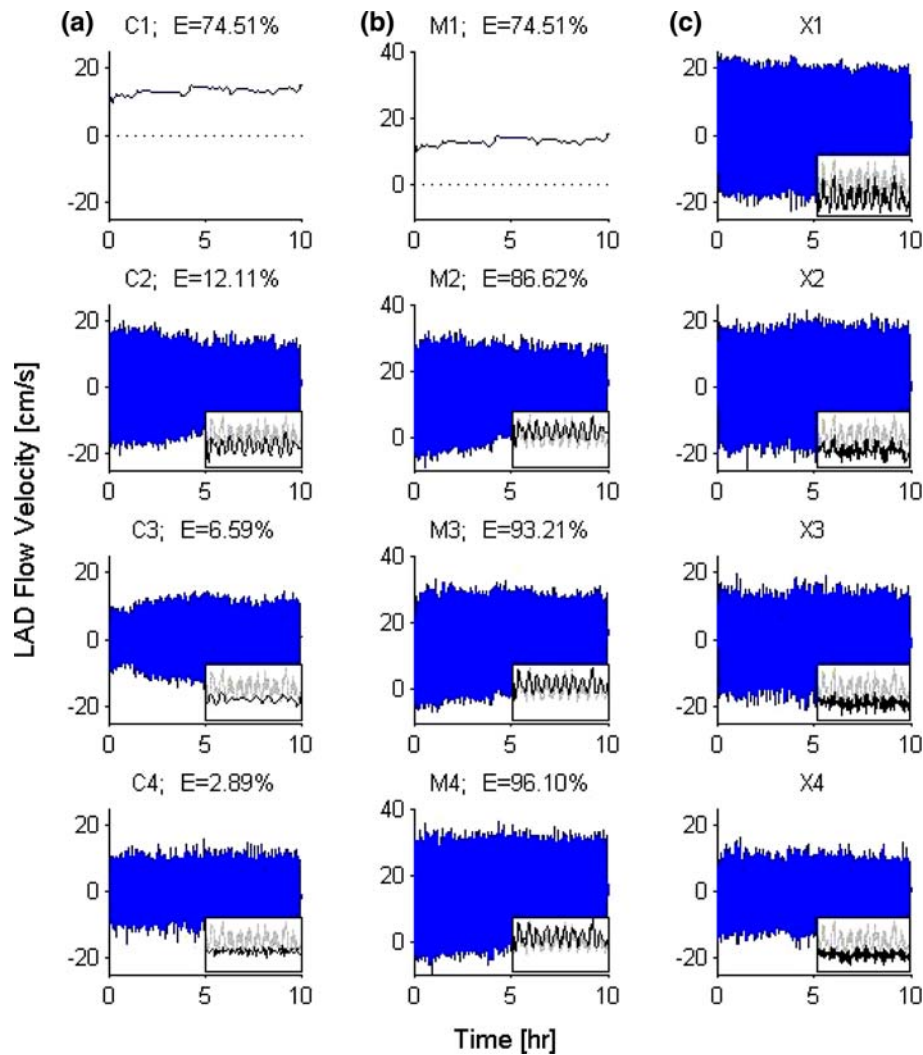
FIGURE 7. Wavelet analysis of LAD flow velocity data. (a) The first 4 modes of oscillation; (b) partial reconstructions; and (c) the corresponding oscillations. *E* is the fraction of total energy contained in the mode of oscillation or partial reconstruction. Insets correspond to 6 s of data, with the original signal given in gray.

single heartbeat (Fig. 10), EMD produces a low number of components (4 total shown in Fig. 11), as opposed to the high number of components generated in longer time periods (25 total, only a few are shown in Fig. 2). Future studies must assess the reproducibility of these single heartbeat IMFs. Once these IMFs become well characterized and assigned physiological meaning, EMD may then be used as a tool to differentiate between healthy and diseased flows. This premise is based on the well-known observation that atherogenic and atheroprotective flows are different. Hence, it is essential to quantify which feature(s) of the flow profile are responsible for activation of atherogenic pathways. For example, flow reversal, high frequency components, time averaged flow magnitude, peak flow magnitude and oscillatory features may be implicated.[5,8,18,20,22,25] The proposed method provides a mathematical tool to quantify such features. Future studies must focus on the features of waveform that are essential for mechanotransduction, growth, remodeling and atherosclerosis.

APPENDIX

Empirical Mode Decomposition (EMD)

EMD is a numerical technique which yields a decomposition of a signal into intrinsic modes through the sifting process. Normally, IMFs are ordered by their smoothness, as determined by the sifting process, with the residual being the last (smoothest) mode in the sequence. In the present study, IMFs are ordered according to their energy content. This renders EMD
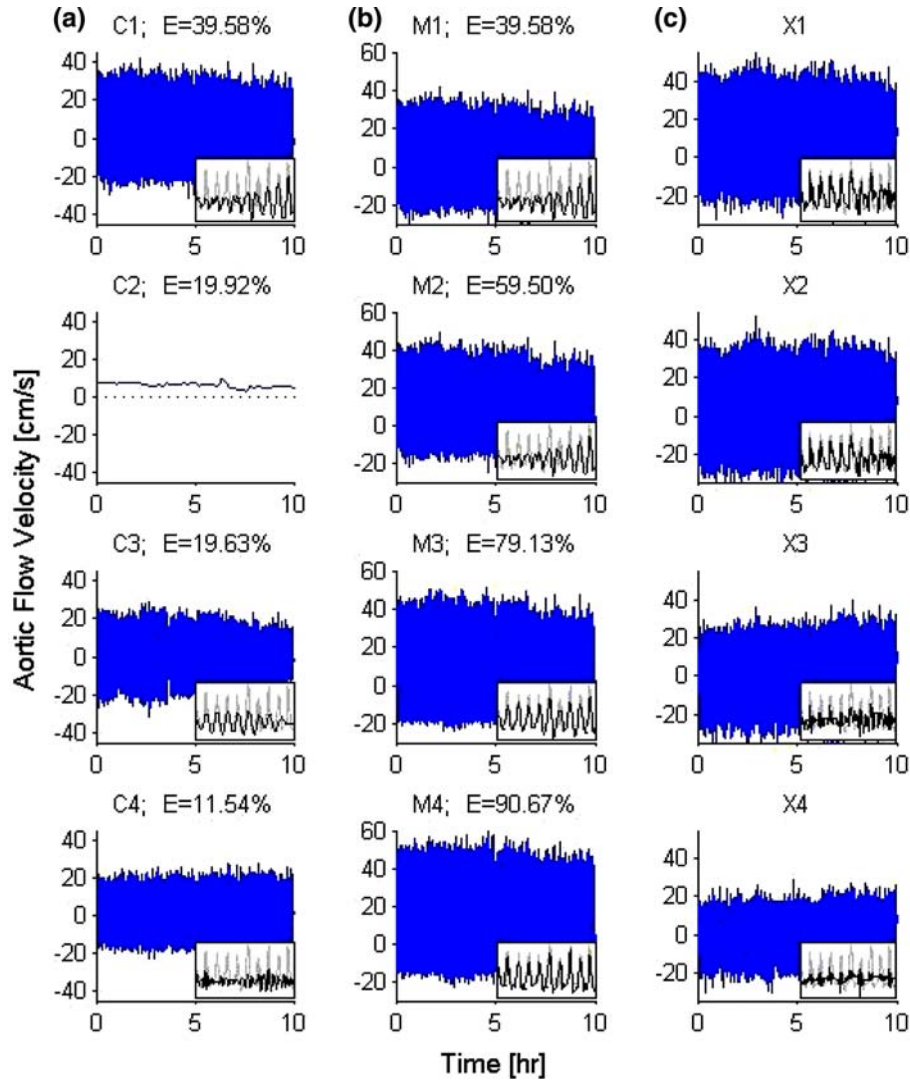
FIGURE 8. Wavelet analysis of aortic flow velocity data. (a) The first 4 modes of oscillation, (b) partial reconstructions; and (c) the corresponding oscillations. *E* is the fraction of total energy contained in the mode of oscillation or partial reconstruction. Insets correspond to 6 s of data, with the original signal given in gray.

comparable to other classical methods, such as PCA and wavelet analysis. The energy, *E*, of a time series is related to the square of its amplitude, i.e. $E = \sum a_i^2$, where a_i is the amplitude of a signal at a discrete time i, and the summation runs over all discrete times. Let C_j ($j = 1,2,\ldots,n$), denote EMD components (IMFs and the residual) of the original wave, X, ranked in order of decreasing energy. Note that in this sorting scheme, the residual is not necessarily the last component, C_n. The first 4 components of X are shown in Figs. 2a (LAD data) and 3a (aortic data).

By design, the EMD components, C_j, are additive; i.e., the sum of the residual and all IMFs is equal to the original signal: $X = \sum_{j=1}^{n} C_j$, where n is the total number of components, determined through the sifting process. With the EMD components ranked according to energy,

we may define the *m*th order partial reconstruction of X as a partial sum of the m most energetic components:

$$\hat{X}(m) = M_m = \sum_{j=1}^{m} C_j \qquad (A.1)$$

Note that $\hat{X}(m) \to X$ as $m \to n$; i.e., increasing the order of partial sum will add more features to the reconstruction, and in the limit the original wave will be reconstituted. The partial reconstructions, M_m, in Figs. 2b and 3b are shown for $m \leq 4$. Finally, Figs. 2c and 3c show the oscillations, X_m, around the partial reconstructions, defined by $X_m = X - M_m$ ($m = 1, 2, 3, 4$). If the residual happens to be one of the components, C_j, in Eq. (A.1), then the partial reconstructions M_m can be viewed as higher order mean trends and X_m,
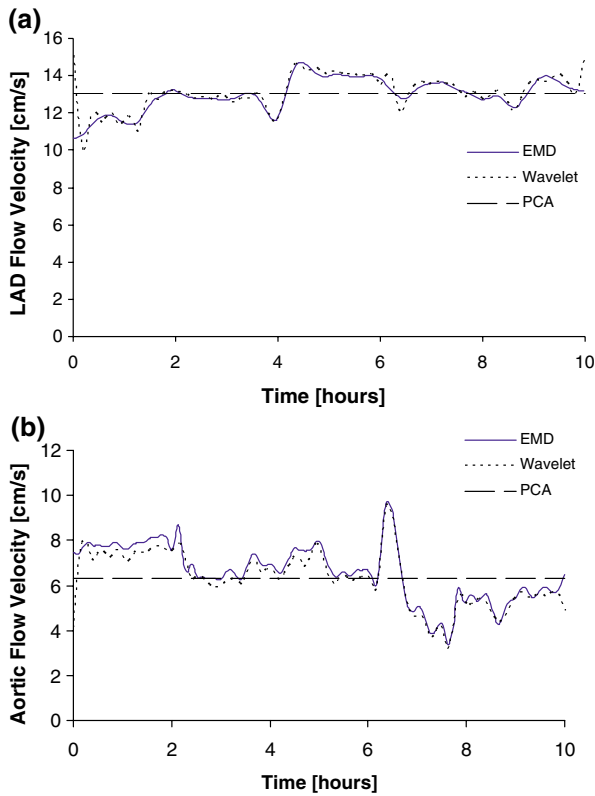
(a)

(b)

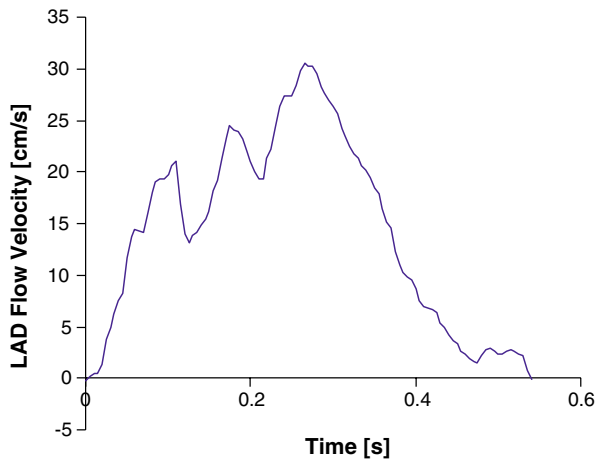FIGURE 9. Comparison of mean trends from EMD, wavelet analysis and PCA.

FIGURE 10. LAD flow velocity in a single heartbeat.

represent the oscillations around the mean trends, common to traditional EMD analyses.[15]

Principal Component Analysis (PCA)

PCA, also known as a proper orthogonal decomposition, or Karhunen-Loeve decomposition,[16,17] is a statistical technique that relies on a rotation of the data coordinate system so that in the new coordinate system the data are maximally uncorrelated. The method has numerous desirable properties and has been extensively used in biomedical applications for analysis of various wave phenomena.

PCA is typically applied to a population of signals, which can be regarded as realizations of a common random process.[20,21] To analyze a single time series $X = [x_1, x_2, ..., x_d]$ in this fashion, it is necessary to divide the time series into a number of non-overlapping segments, say $X_1, X_2, ..., X_N$, where each segment is an n-sample long time series ($Nn = d$). The objective of choosing N and n, is to maximize the correlation in the population $\{X_i\}_{i=1}^{N}$, which is crucial for the success of PCA. To this end, we calculated the auto-covariance sequence of the aortic flow velocity data, X, and we consider a lag at which the sequence attains a maximum. The first non-trivial peak[1] appears at the lag $\tau = 121$, which at the sampling rate of 200 Hz corresponds to ~0.6 s. Not surprisingly, this period roughly matches the duration of the cardiac cycle. Therefore, by choosing $n = 121$, the time series, X, is to split into segments $X_1, X_2, ..., X_N$, which appear as maximally correlated random samples of the same random vector $X \in \Re^n$.

Once the population $\{X_i\}_{i=1}^{N}$ is defined, PCA proceeds by finding the eigenvalues of the sample covariance matrix, Σ_{XX}, defined by

$$\Sigma_{XX} \propto \sum_{i=1}^{N} [X_i - \mu_X][X_i - \mu_X]^T \in \Re^{n \times n} \quad (A.2)$$

where $\mu_X = \frac{1}{N}\sum_{i=1}^{N} X_i$ is the sample mean of the population. Since the matrix Σ_{XX} is symmetric and positive definite, its eigenvectors are orthogonal and form a basis in $\Re^n$.[1] The strength of PCA is that data can be faithfully represented in a low-dimensional subspace, spanned by the principal eigenvectors of Σ_{XX}, while the complementary subspace can be discarded as noise. The principal eigenvectors, also known as the principal components, are the eigenvectors of Σ_{XX} that correspond to a few of its largest eigenvalues. While the number of chosen principal components is application specific, more correlated populations generally allow choosing a smaller number of principal components.

If $\{v_j\}_{j=1}^{m}$ are the m principal components, then any vector, $X_i \in \Re^n$, can be represented in the new basis as $c_i = [c_i^1, c_i^2, ..., c_i^m]$, where the coefficients, c_i^j, are defined as: $c_i^j = (X_i - \mu_X)^T v_j$. Similarly, each X_i from the population can be recovered using the mth order partial reconstruction:

$$\hat{X}_i(m) = \mu_X + \sum_{j=1}^{m} c_i^j v_j \quad (A.3)$$

---

[1]Autocovariance sequence of a time series always has a trivial peak at the lag $\tau = 0$ since each sample is in perfect correlation with itself.
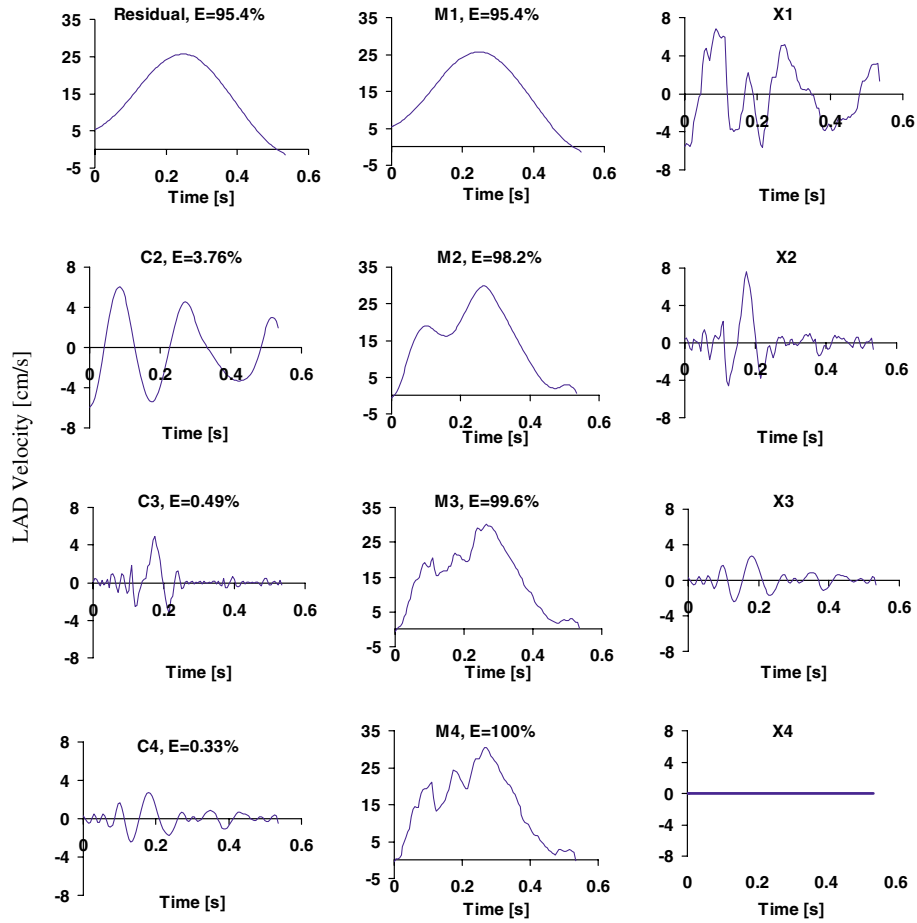
FIGURE 11. (a) IMFs and residual; (b) mean trends; and (c) oscillations about the mean trends of LAD flow velocity in a single heartbeat. *E* is the fraction of total energy contained in the mode of oscillation or mean trend.

In this study we use $m \leq 3$ principal components. Note that $\hat{X}_i(m) \rightarrow X_i$ as $m \rightarrow n$. The modes C_j ($j = 1, 2, 3$) of the time series, X (see Figs. 5a and 6a) are constructed by concatenating the modes of individual population vectors, X_i, i.e.

$$C_j = \left[c_1^j v_j^T, c_2^j v_j^T, \ldots, c_N^j v_j^T \right] \quad (A.4)$$

Similarly, jth order partial reconstruction of X (see Figs. 5b and 6b) is obtained by concatenating the partial reconstructions of individual vectors, X_i, i.e.,

$$M_j = \left[\hat{X}_1^T(j), \hat{X}_2^T(j), \ldots, \hat{X}_N^T(j) \right], \quad \forall_j = 0, 1, 2, 3 \quad (A.5)$$

where $\hat{X}_i(0) = \mu_X$. Analogous to EMD, these partial reconstructions are referred to as mean trends. Finally, oscillations, X_j, around the mean trends (see Figs. 5c and 6c) are defined as: $X_j = X - M_j$ ($j = 0, 1, 2, 3$).

Wavelet Analysis

Wavelets represent an important tool for signal analysis, data compression, statistical pattern recognition, and other signal processing applications. The mathematical theory of wavelets is beyond the scope of this article, and a comprehensive exposition can be found in Mallat.[19] Historically, the development of wavelets was inspired by the limitations of Fourier-based techniques, and wavelets quickly attained popularity because of their ability to handle non-stationary signals.

A time series, $X \in \Re^d$, can be represented in the wavelet basis as

$$X = \sum_{a,b} c_{a,b} \psi_{a,b} \quad (A.6)$$

where $\psi_{a,b}$ are wavelet (basis) functions and $c_{a,b}$ are the expansion coefficients. The basis functions, $\psi_{a,b}$, are obtained by scaling and translating a common wave pattern, ψ, called the *mother wavelet*. In particular

$$\psi_{a,b}(t) = \frac{1}{\sqrt{a}} \psi\left(\frac{t-b}{a}\right), \quad a, b \in \Re \quad (A.7)$$

where, a and b, represent the scale and translation parameters, respectively. The major strength of

wavelets is that the scales and translations can be varied systematically; thus, capturing analyzed function features at various scales (frequencies) and occurrence times. Since X is a discrete signal, the basis functions in Eq. (A.2) are obtained by discretization of the wavelets in Eq. (A.3). The expansion coefficients, $c_{a,b}$, are obtained by projecting the signal, X, onto the wavelet basis; i.e., $c_{a,b} = X^T \psi_{a,b}$. If the dyadic scales and translations are chosen, that is $\{a = 2^j; b = k2^j; j,k \in \mathbf{Z}\}$, it can be shown that the basis in Eq. (A.2) is orthogonal.

While there are a number of wavelet families, most of them have similar properties. In this study, we use a wavelet from the family of coiflets.[8] Coiflets belong to the family of compactly supported wavelets, which makes them a suitable tool for analysis of transient and non-stationary phenomena. The choice of wavelet family did not significantly influence the results of our analysis. Coiflets are parameterized by the number of vanishing moments; and here, we use the family coif5[2] which has 10 vanishing moments. A change in the number of vanishing moments does not significantly affect the results of our study.

The modes C_j ($j = 1, 2, 3, 4$) of the time series, X, (see Figs. 7a and 8a) are obtained by reconstructing the signal at various scales:

$$C_j = \sum_b c_{a_j,b} \psi_{a_j,b} \qquad (A.8)$$

where a_1, a_2, a_3, and a_4 are the 4 scales that carry the most of signal energy. Similar to EMD and PCA analyses, the jth order partial reconstructions M_j ($j = 1, 2, 3, 4$), (see Figs. 7b and 8b) are defined as partial sums $M_j = \sum_{k=0}^{j} C_k, \forall_j = 1, 2, 3, 4$, and their respective oscillations, X_j (see Figs. 7c and 8c) are defined as: $X_j = X - M_j$ ($j = 1, 2, 3, 4$).

ACKNOWLEDGMENT

REFERENCES

[1]Anton, H. Elementary Linear Algebra. 6th ed. New York: Wiley, 1991.

[2]Asakura, T., and T. Karino. Flow patterns and spatial distribution of atherosclerotic lesions in human coronary arteries. *Circ. Res.* 66:1045–1066, 1990.

[3]Axelsson, M., Q. Dang, K. Pitsillides, S. Munns, J. Hicks, and G. S. Kassab. A novel, fully implantable, multichannel biotelemetry system for measurement of blood flow, pressure, ECG, and temperature. *J. Appl. Physiol.* 102:1220–1228, 2007.

[4]Bassiouny, H. S., C. K. Zarins, M. H. Kadowaki, and S. Glagov. Hemodynamic stress and experimental aortoiliac atherosclerosis. *J. Vasc. Surg.* 19:426–434, 1994.

[5]Caro, C. G., J. M. Fitz-Gerald, and R. C. Schroter. Arterial wall shear and distribution of early atheroma in man. *Nature* 223:1159–1160, 1969.

[6]Cheng, C., D. Tempel, R. van Haperen, A. van der Baan, F. Grosveld, M. J. Daemen, R. Krams, and R. de Crom. Atherosclerotic lesion size and vulnerability are determined by patterns of fluid shear stress. *Circulation* 113:2744–2753, 2006.

[7]Cohen, L. Time-Frequency Analysis. Englewood Cliffs, NJ: Prentice-Hall, 1995.

[8]Daubechies, I. Ten lectures on wavelets. *CBMS, SIAM* 61:258–261, 1994.

[9]Davies, P. F., D. C. Polacek, J. S. Handen, B. P. Helmke, and N. DePaola. A spatial approach to transcriptional profiling: mechanotransduction and the focal origin of atherosclerosis. *Trends Biotechnol.* 17:347–351, 1999.

[10]Dimmeler, S., and A. M. Zeiher. Endothelial cell apoptosis in angiogenesis and vessel regression. *Circ. Res.* 87:434–439, 2000.

[11]Gabor, D. Theory of communication. *J. Inst. Elect. Eng. (Lond.)* 93:429–460, 1946.

[12]Gimbrone, M. A., Jr., J. N. Topper, T. Nagel, K. R. Anderson, and G. Garcia-Cardena. Endothelial dysfunction, hemodynamic forces, and atherogenesis. *Ann. N.Y. Acad. Sci.* 902:230–239, 2000; discussion 239–240.

[13]Huang, N. E., Z. Shen, and S. R. Long. A new view of nonlinear water waves: the Hilbert spectrum. *Annu. Rev. Fluid Mech.* 31:417–457, 1999.

[14]Huang, N. E., Z. Shen, S. R. Long, M. L. C. Wu, H. H. Shih, Q. N. Zheng, N. C. Yen, C. C. Tung, and H. H. Liu. The empirical mode decomposition and the Hilbert spectrum for nonlinear and non-stationary time series analysis. *Proc. R. Soc. Lond. Ser. A: Math. Phys. Eng. Sci.* 454:903–995, 1998.

[15]Huang, W., Z. Shen, N. E. Huang, and Y. C. Fung. Engineering analysis of biological variables: an example of blood pressure over 1 day. *Proc. Natl. Acad. Sci. U.S.A.* 95:4816–4821, 1998.

[16]Jolliffe, I. T. Principal Component Analysis. New York: Springer Verlag, 1986.

[17]Ku, D. N. Blood flow in arteries. *Annu. Rev. Fluid Mech.* 29:399–434, 1997.

[18]Malek, A. M., S. L. Alper, and S. Izumo. Hemodynamic shear stress and its role in atherosclerosis. *JAMA* 282:2035–2042, 1999.

[19]Mallat, S. G. A Wavelet Tour of Signal Processing. San Diego, CA: Academic Press, 1999.

[20]Mitra, P. P., S. Ogawa, X. Hu, and K. Ugurbil. The nature of spatiotemporal changes in cerebral hemodynamics as manifested in functional magnetic resonance imaging. *Magn. Reson. Med.* 37:511–518, 1997.

[21]Nenadic, Z., B. K. Ghosh, and P. S. Ulinski. Modeling and estimation problems in the turtle visual cortex. *IEEE Trans. Biomed. Eng.* 49:753–762, 2002.

[22]Pedersen, E. M., M. Agerbaek, I. B. Kristensen, and A. P. Yoganathan. Wall shear stress and early atherosclerotic lesions in the abdominal aorta in young adults. *Eur. J. Vasc. Endovasc. Surg.* 13:443–451, 1997.

---

[2]The nomenclature for the wavelet names used throughout this article is consistent with that of MATLAB™ Wavelet Toolbox.

[23]Priestley, M. B. Evolutionary spectra and non-stationary processes. *J. Roy. Statist. Soc. B* 27:204–237, 1965.

[24]Salisbury, J. I., and Y. Sun. Assessment of chaotic parameters in nonstationary electrocardiograms by use of empirical mode decomposition. *Ann. Biomed. Eng.* 32:1348–1354, 2004.

[25]Wentzel, J. J., R. Corti, Z. A. Fayad, P. Wisdom, F. Macaluso, M. O. Winkelman, V. Fuster, and J. J. Badimon. Does shear stress modulate both plaque progression and regression in the thoracic aorta? Human study using serial magnetic resonance imaging *J. Am. Coll. Cardiol.* 45: 846–854, 2005.

[26]Xie, H., and Z. Wang. Mean frequency derived via Hilbert-Huang transform with application to fatigue EMG signal analysis. *Comput. Methods Programs Biomed.* 82:114–120, 2006.