1-26-2016

Chapter 2: Time Series

Sara A. Myers

Time Series

Sara A. Myers

CONTENTS

> To see a world in a grain of sand, And a heaven in a wild flower, Hold infinity in the palm of your hand, And eternity in an hour.
>
> William Blake (1757–1827)

WHAT IS A TIME SERIES?

Time series is "simply a list of numbers assumed to measure some process sequentially in time" (Stergiou et al. 2004). Mathematicians have a more formal definition, that is, a set or a sequence of observations, with each one recorded at specific times, or at least sequentially (Brockwell and Davis 2002; Box et al. 2008). Time series are created from multiple sources for research purposes to understand various behaviors. For example, social scientists could collect graduation rates, physiologists record heart rates, economists study consumer spending, and climatologists examine weather patterns. Basically, any time observations are taken repeatedly over time, from any source or behavior, a time series is created.

A basic assumption of most time series analysis is that all time series inherently possess dependence between adjacent observations. In fact, this dependence is of interest because it reveals information about the source producing the behavior. In this way, time series analysis is essential for understanding human movement variability, because time series analysis reveals how the system evolves over multiple movement repetitions. To generate a time series, repeated measurements of some property of the system are made as the system varies in time. This may imply that time series data are essentially a list of numbers, but any list of numbers cannot be considered a time series. This chapter details what constitutes time series data and describes important specific considerations that should be kept in mind when working with time series data. Consider the following two lists of numbers:

List A: 1, 2, 3, 4, 5

List B: 3, 1, 4, 2, 5

The average value for list A is 3; the average value for list B is 3. The range for list A is 4; the range for list B is 4. The standard deviation for list A is 1.58, and the standard deviation for list B is 1.58. Thus, the statistical descriptors of the two lists are the same—we cannot tell the difference between these two lists by these statistical measures. However, by examining the lists, it is easy to observe that they are not the same. The difference between the two lists is the order in which the numbers appear in the sequence, that is, to understand the difference between the two lists, we must examine these lists as ordered lists of numbers. To indicate an ordered list, we will use the notation [1, 2, 3, 4, 5] for list A, and [3, 1, 4, 2, 5] for list B. Further, to investigate the difference between these two lists, we would need an analysis technique that is sensitive to the order of the numbers in the list. In fact, the remaining chapters of this book discuss various methods of data analysis that quantify various aspects of the patterns formed by lists of numbers, based on the order in which the numbers appear in the sequence.

Time series data are a specific example of an ordered list of numbers, where time is the parameter that gives order to the list. For example, say every year on your daughter's birthday you measure her height, starting at age 5 until age 15. Her height in centimeters is [108, 115, 121, 128, 134, 140, 146, 152, 157, 161, 163]. Her age in years at each measurement is [5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]. You can line them up, and each number in the top list is recorded at the same time as the corresponding one in the other list.

[108, 115, 121, 128, 134, 140, 146, 152, 157, 161, 163]

[5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15]

If you want to make a plot of height versus age, you will plot the pairs of corresponding numbers, that is, plot (5, 108), then plot (6, 115), then plot (7, 121), etc., because you will want to plot the first value, then plot the second value, then plot the third value, etc. (Figure 2.1).

Another way to record your daughter's age is to write the age at which she is 10 cm taller than the previous measurement. This ordered list would be age in years of [5.0, 6.5, 8, 9.5, 11.2, 13.1] and the corresponding height in cm is [108, 118, 128, 138, 148, 158]. This is also a time series. However, these readings were not taken at even time intervals. They are still time series data because they are an ordered list that is ordered based on time. Note that some analysis techniques assume equal time intervals between data points (such as the spectral analysis technique to be discussed later), so be aware of which type of data you are analyzing. This is where the concepts of discrete versus continuous time series (discussed in Chapter 1) become important. Discrete time series are series in which observations are made in discrete sets, such as at specific, fixed time intervals (Brockwell and Davis 2002). The term "continuous data or continuous time series" is described in various ways depending on the source. According to Brockwell and Davis, continuous series are those in which observations are recorded continuously for a specific amount of time (Brockwell and Davis 2002). However, Warner describes continuous data as those with a true interval/ratio level of measurement, or data in which the difference between two values is meaningful (Warner 1998). Discrete time series is also described as a sampling of continuous time series at certain intervals (Box et al. 2008). Most time series, which are analyzed using techniques described in this book, are discrete time series, with observations made at equal intervals (Warner 1998; Brockwell and Davis 2002; Box et al. 2008). As the examples that follow will illustrate, even time series that appear continuous, when examined closely, are composed of individual data points.

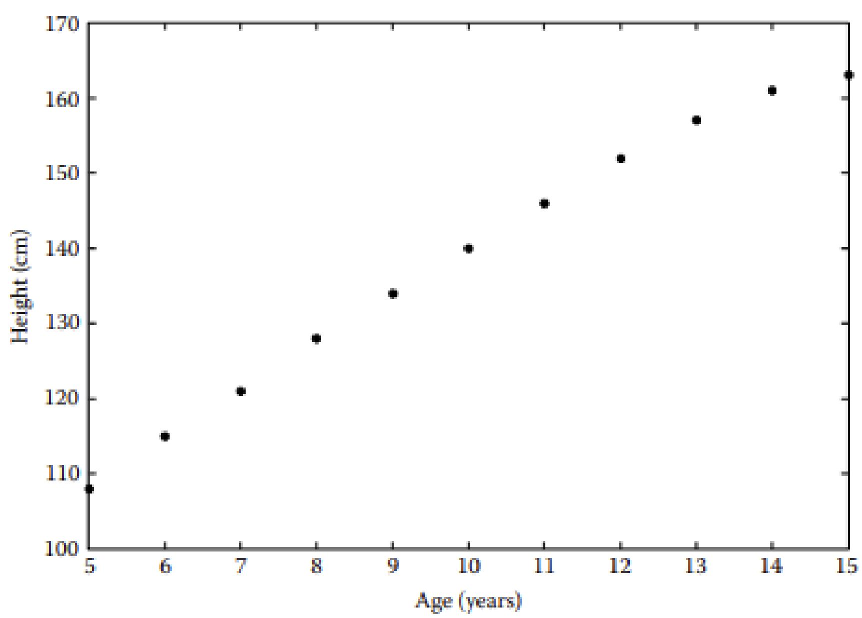

FIGURE 2.1 Plot of height versus age for a hypothetical girl. A time series is formed by the ordered list of how the girl's height changes with each year increase in age.

Now that we have a definition of what a time series is, let us examine some human movement data, that is, a plot of the knee angle versus time during walking (Myers et al. 2013, 1692–1702; Figure 2.2). These data were collected by applying several reflective markers to different body segments on a subject's body, making a video of the subject while walking, and then identifying where the markers were located in each frame of the video. Because the cameras had been calibrated to measure distance in three-dimensional space, the position of the markers could be used to calculate the position of the body segments in each frame, and knowing that, a knee angle could be calculated for each frame.

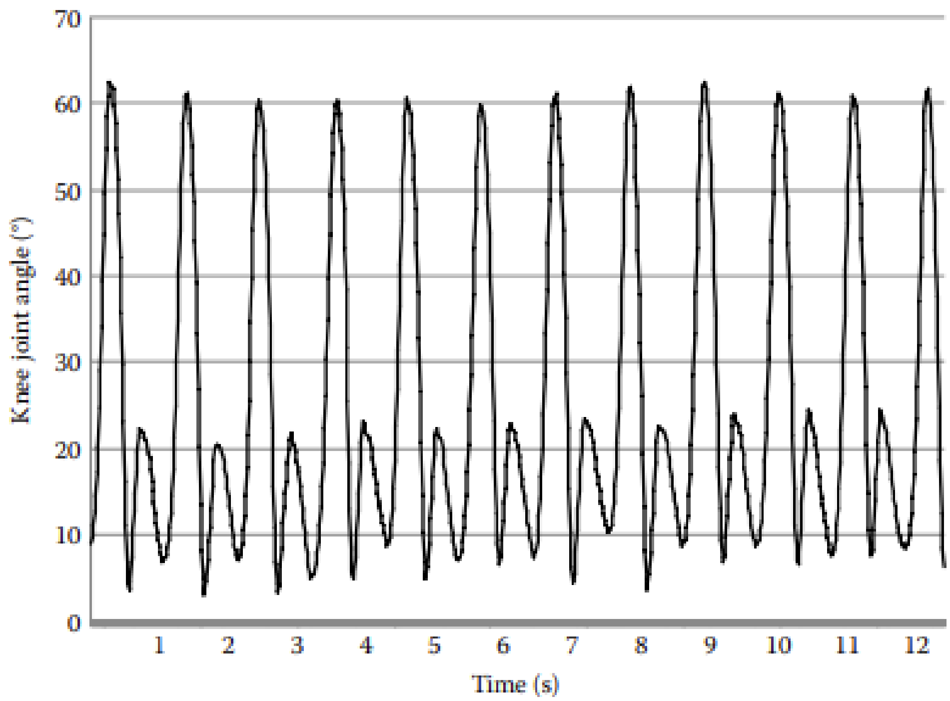

FIGURE 2.2 Plot of right knee flexion and extension angle versus time for a healthy older subject while walking.

The time series plotted in Figure 2.2 is the knee angle calculated from each frame in the video, with the time at which the frame was acquired plotted on the x axis. If you zoom in on just 1.5 s of the data (Figure 2.3), you can see that the data are not actually a continuous line, but rather it is composed of individual points. In other words, the time series is not continuous, but rather it is discrete. If you count the dots in the 1.5 s of data, you will find that there are 90 points. This is because the camera acquiring the video was taking a picture about every 0.0167 s, or 60 frames/s (60 frames/s times 1.5 s = 90 data points). The Hertz, or Hz for short, is a common unit of measure for how many data points are acquired every second. In this example, the data were collected at 60 Hz.

SOME EXAMPLES OF TIME SERIES DATA

Time series analysis is useful in many different situations, whenever you are trying to describe the change of a system with time, that is, the dynamics of the system. For example, a psychologist had subjects fill out a survey once a day for 4 weeks, scoring the answers, and creating a time series that has 28 points (Fredrickson and Losada 2005). Astronomers have observed sunspot activity annually since 1700 by counting sunspots visible on the sun, creating a time series of length 314 (Figure 2.4; National Office of Oceanic and Atmospheric Administration 2013). Physiologists analyzed blood samples for luteinizing hormone every 10 min for 4 days, creating a time series 577 points long (Liu et al. 2007). Biomechanists interested in the gait function of patients with peripheral arterial disease analyzed the ankle flexion and extension angles captured by cameras taking 60 pictures/s for 45 s, resulting in a time series 2700 points long (Myers et al. 2009, 2011). Data from this experiment describe adjustments of the ankle with each step while walking on a treadmill. These adjustments are meaningful because they reflect the cooperative strategies of the locomotor system and are considered a marker of the system's health (Figure 2.5).

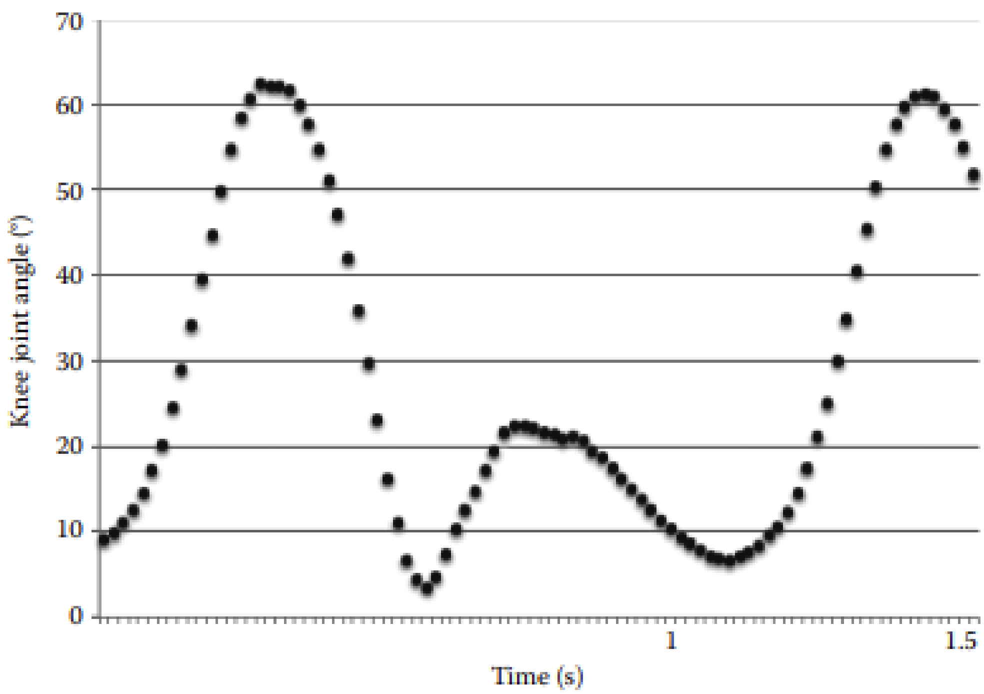

FIGURE 2.3 Expanded plot of right knee flexion and extension angle versus time for a healthy older subject while walking.

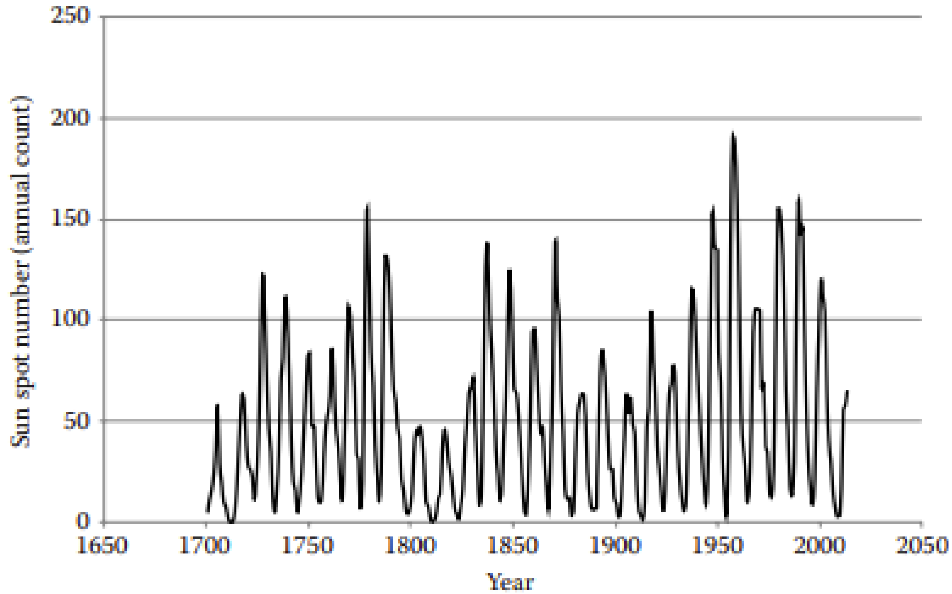

FIGURE 2.4 Plot of yearly sunspot activity data from National Office of Oceanic and Atmospheric Administration (2014) from 1700 until 2013. Analysis reveals that there is a cycle in sunspot activity, with activity maximizing about every 10.5 years.

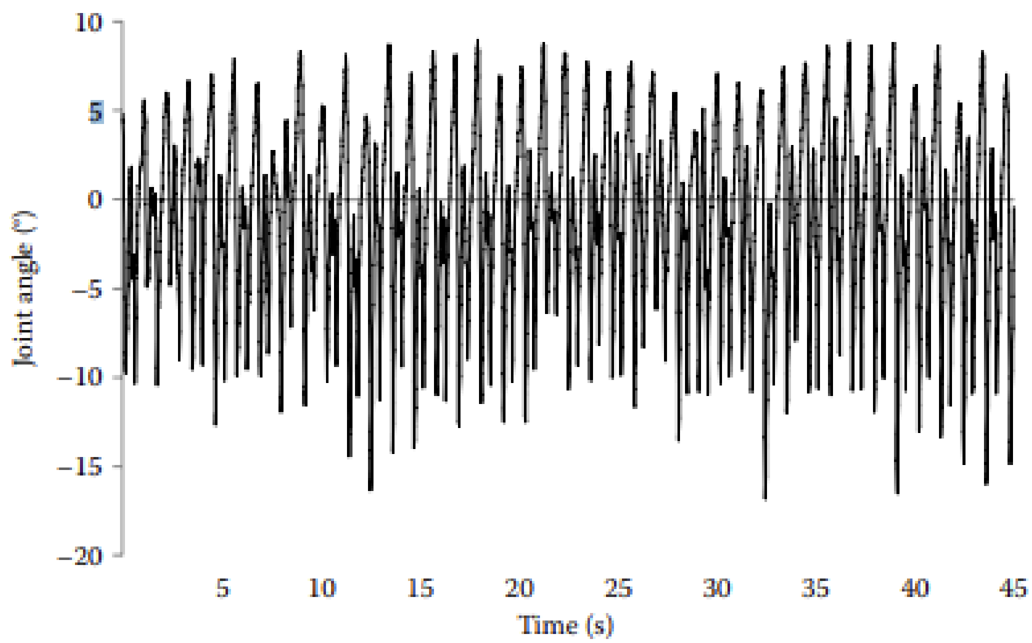

FIGURE 2.5 Ankle flexion and extension motion during walking in a patient with peripheral arterial disease. Plantarflexion is represented as negative values and dorsiflexion is shown with positive values. Data were collected at 60 Hz for 45 s, creating a time series that is 2700 data points long.

As you can see from this wide range of examples, there is no typical application of time series analysis. Scientists, engineers, and clinicians from many different disciplines use time series analysis to provide an understanding the dynamics of whatever system they are studying. While all these different applications may seem entirely unrelated, the methodology of the analysis overlaps considerably.

SEVERAL IMPORTANT ASPECTS OF TIME SERIES DATA ANALYSIS

While there are many varied uses for the analysis of time series data, there are some general aspects that are important regardless of the application, and these considerations are discussed in the following sections. Perhaps the most important initial consideration is the goal of the analysis. A general goal of all time series analysis is to understand the patterns in the data. How that understanding is to be used will influence the type of analyses that will be performed. The goals of time series analysis include forecasting, determining a transfer function for predictive purposes, describing a phenomenon or relationships between related time series, studying the effect of interventions on the time series, developing control schemes, and others (Warner 1998; Brockwell and Davis 2002; Box et al. 2008). Each of these purposes can be applied in varying domains from predicting the stock market, understanding weather patterns, or describing movement problems that result from pathology. Linear time series analysis methods assume observations are independent of past history. However, sequential observations of human movement are rarely independent of each other. This book acknowledges this fact and approaches human movement from this exact perspective that seems mostly lost in the current literature. Therefore, important features discussed here are centered on nonlinear algorithmical techniques to investigate human movement. However, regardless of the investigative approach, it is always essential to consider several fundamental issues of time series analysis. These include length of the time series, noise, resolution, stationarity and inherent periodicity, sampling frequency, spectral analysis, and filtering and smoothing. Each algorithm presented in this book has unique requirements, and many of these issues will be revisited in upcoming chapters. In the following section, we discuss these considerations as applicable for all algorithms.

Length of the Time Series

The length of the time series is often seen as a major limitation for the utilization of a certain types of analyses for time series data. In particular, for nonlinear analysis, the individuals (usually mathematicians) who have derived the formulas suggest that a certain number of data points are critical for performing the analysis. It is not that the nonlinear calculation with fewer data points than what they suggest cannot be done, but the problem is whether the answer received from using a shorter time series is really an accurate characterization of the dynamics of the system.

From a more intuitive perspective, you need a time series long enough to capture the essential dynamics of the system. As an example, consider the periodic behavior of the sunspot cycle. It takes over 10 years for the cycles to repeat. If you only record sunspot activity for 5 years, you would not be able to discover from your data that there is a 10-year cycle. If you collect 10 years, you will observe a complete cycle, but you would not know that the same pattern will be repeated. The more complete the cycles you have, the better your ability to characterize the dynamics, just as in estimating the population mean, the more samples you have from that population the better.

So how long is long enough? This depends on which analytical technique you want to use—some require more data than others. One mathematical rule of thumb is that you want a time series that is at least 10D data points in length, where D is the dimensionality of the system. The dimensionality will be discussed later in the book, but for many systems of interest, it is probably at least three, and likely a bit higher, depending on the system that you are studying. So if the dimension of your system happens to be 6, you need 106 or 1,000,000 data points. However, whether or not it is practical to obtain such a time series is another important consideration. Let us say you are studying ecological time series data, such as the annual population of a certain species. You can only collect one data point per year, so a technique that requires 1,000,000 data points is clearly out of the question. Less stringent guidelines suggest having at least 5 and if possible 10 repetitions of the cycle to be able to understand the underlying dynamics (Warner 1998).

The problem with the discussion is that, while the inventor of certain technique can only guarantee the results if there are this large number of data points, the technique may be somewhat robust with respect to sample size and thus able to give reasonable results for a smaller number of data points. The big question, then, is how robust are the techniques for nonlinear analysis with respect to the length of the time series. This varies from one technique to the next, and even for a given technique, different algorithms can have different requirements. Also see the following discussion about sampling frequency as it relates to the length of the time series.

Sampling Frequency and Spectral Analysis

Sampling frequency is a critical consideration when dealing with time series data. It is a measure of how often you acquire a data sample, and thus, sampling frequency multiplied by the length of time that you sampled gives the number of data points in your time series. The sampling frequency needs to be high enough to capture the dynamics of the quickest changes in your system. For example, with the sunspot data (Figure 2.4), there is a cycle about every 10.5 years. So if you only sampled every 20 years, you would miss it. Even if you sampled every 10 years, that would not be enough to catch the cycle. The minimum frequency at which you need to sample in order to have a chance of obtaining periodic dynamics is twice the frequency of the fastest dynamics. This principle is known as the Nyquist sample theory. But this gives you only two data points for every cycle, so going up to a sampling frequency of about five times the fastest frequency is a good rule of thumb for periodic data. In human locomotion, the highest frequencies that occur during walking are less than 12 Hz. Thus, a 24 Hz sampling rate should be satisfactory; however, in reality, biomechanists usually sample at 5–10 times the highest frequency in the signal. Remember, if data are undersampled, the entire signal is not captured. If data are oversampled, more measurement noise could be introduced.

So how do you decipher what frequencies are in your data? One method is spectral analysis, which entails breaking down the biological signal into simple signals. The typical approach to analyzing data is to describe the data in terms of how they change over time, which is known as the time domain. An alternative approach arises in the form of data analysis in the frequency domain. This type of analysis presents data as a function of frequencies contained in the signal rather than a function of amplitudes over time. Frequency-domain analysis is used extensively to provide additional insights into healthy and pathological movement (Giakas et al. 1996; Giakas and Baltzopoulos 1997; Stergiou et al. 2002; Giakas 2004; Wurdeman et al. 2011; McGrath et al. 2012). More specifically, spectral analysis is a numerical technique to write your data as the sum of multiple discrete sine and cosine functions of different frequencies. There are many different frequency transforms available, but the most commonly used transform is the Fourier transform. This transform uses sums of sine and cosine functions to represent the more complex functions, in our case, signals of human movement. The plot of the power at each frequency is referred to as the power spectral density plot or simply the power spectrum.

Calculating the power spectrum is like shining light through a prism. Just as the prism shows the light has many component colors, the spectral analysis shows your data as many component wave functions. There are four characteristics of signals included in the sine and cosine waves that represent the signal. These are frequency, amplitude, vertical offset or shift off the baseline, and the phase angle, which indicates where the signal starts and shifts from right to left. Any signal, h(t) is made up of these four characteristics. Equation 2.1 incorporates each of these variables:

$$h(t) = A_0 + A\sin(2\pi ft + \theta) \tag{2.1}$$

but $2\pi f = \omega$, so another way to write Equation 2.1 is Equation 2.2:

$$h(t) = A_0 + A\sin(\omega t + \theta) \tag{2.2}$$

where

A_0 is the offset

A is the amplitude

f is the frequency

θ is the phase shift

ω is the angular velocity

Thus, if you know these characteristics, then you can write the equation for the signal.

To demonstrate how different frequencies contribute to a signal consider the following: a sine wave of 3 Hz, a cosine wave of 13 Hz, a sine wave of 30 Hz, and the sum of these three wave functions. The peak in the power spectrum corresponds to the frequency of the wave function in the time series (Figure 2.6a through c), and the time series that is a sum of three wave functions has three peaks, corresponding to the three frequencies which were added together (Figure 2.6d). The reason for wanting to divide your data into sine and cosine functions is because then you can see which frequencies are contributing the most to your data just by examining the peak positions. If the peak corresponding to 3 Hz is very high, then you likely have a 3 Hz component in your signal. If the 3 Hz component of your signal is the highest frequency that is significant, then you know you would need a sampling frequency of at least 6 Hz (2 × 3 Hz = 6 Hz) to be able to see it, but something more like 15 Hz (5 × 3 Hz = 15 Hz) would be best to define it better. For real data, determining the highest frequency can be a bit tricky, because often the signal approaches the baseline gradually without a clear cutoff point (Figure 2.7).

To examine the sampling frequency issue from another perspective, let us examine sampling a known signal at various frequencies (Figure 2.8). Each frame on the left side of Figure 2.8 shows the signal we are trying to sample in a solid black line, and the data that we actually sample in black dots, with a line connecting the dots. Each frame on the right side is the power spectrum calculated from the data that we sampled (from the black dots). The signal that we are trying to sample is the sum of two wave functions, one at 2 Hz and one at 3 Hz; it is the same data as in Figure 2.6d. In Figure 2.8a, the sampling frequency is 30 Hz—10 times the highest frequency in the signal we are trying to sample. The curve fits the data so well that the plot of the actual data can barely be seen. Moving down in Figure 2.8, the sampling frequency is reduced. At 6 Hz, the sampling rate is twice the highest frequency in the signal, and one can see from the power spectrum that the spectral analysis is just barely able to capture the highest frequency peak. However, examining the time series plot on the left, we see that the there are significant gaps where the sampled data do not match the signal of interest very well. The power spectrum contains all peaks that it should based on the original data. So, based on this spectral analysis, 6 Hz is the lowest sampling rate that is acceptable. At even lower sampling rates (Figure 2.8e), peaks show up in the power spectra that do not belong based on the known dynamics of the time series. This phenomenon is called aliasing and is the reason that you always need to have a sampling rate at least twice the highest frequency of interest in your time series.

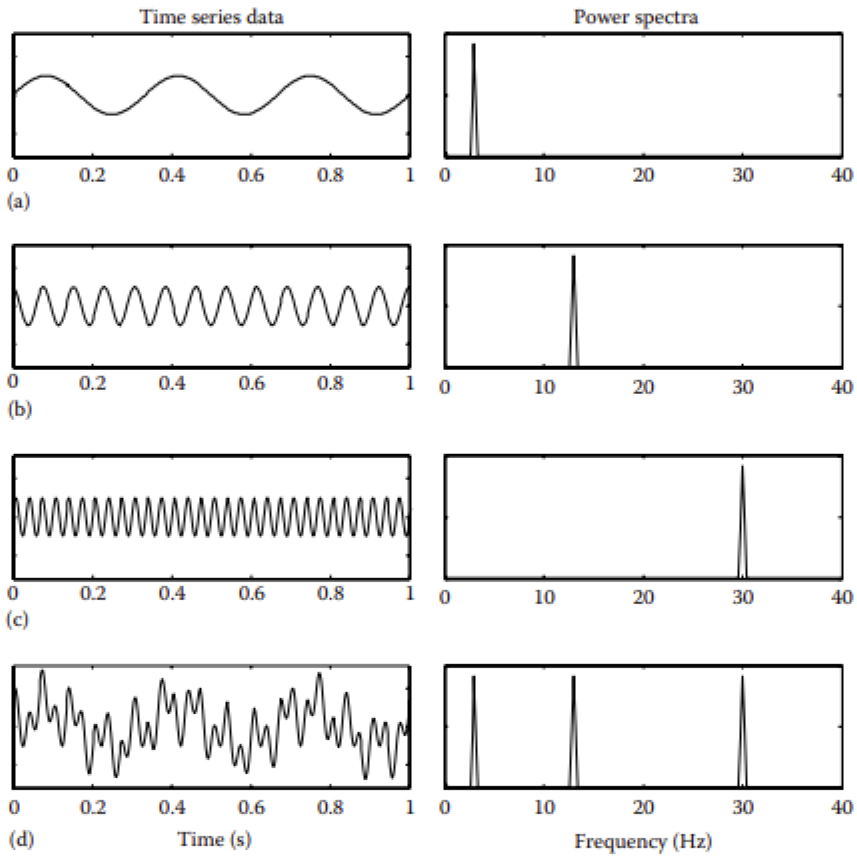

FIGURE 2.6 Time series data (left) and corresponding power spectra (right) for a sine wave of 3 Hz (a), a cosine wave of 13 Hz (b), a sin wave of 30 Hz (c), the sum of these three wave functions (d).

Spectral analysis is a very powerful tool for finding periodic components in your data, and thus, it is widely used. But you should bear in mind that it is a mathematical technique to write your time series data as a sum of sine and cosine functions.
The interpretation of the periodicity in your data depends on your understanding of the underlying mechanism, and what periodicity or lack thereof means in terms of the system you are studying. There are a number of software programs that will allow you to calculate the power spectrum of your data, but there are many subtleties, not discussed here (such as windowing, detrending, and zero-padding (Percival and Walden 1993; Stoica and Moses 1997; Huang et al. 1998; Beard 2003; Prabhu 2013), that can help you do a better job with the analysis. See Percival and Walden (1993), or any good digital signal processing textbook for a more in-depth discussion of spectral analysis. Or, for a more lighthearted introduction, see “Who Is Fourier? A Mathematical Adventure” (Sakakibara 1995).

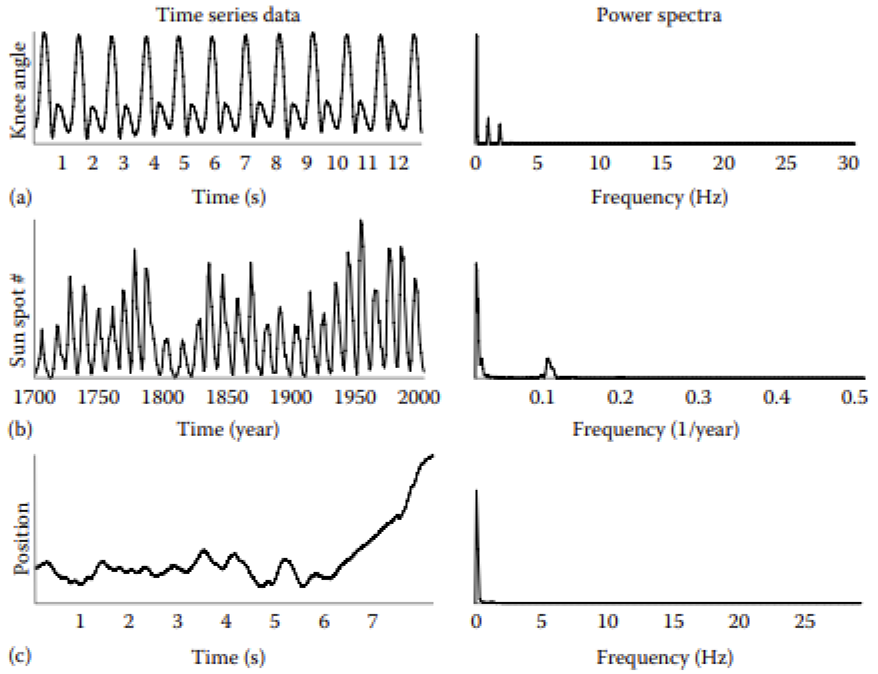

FIGURE 2.7 Time series data (left) and corresponding power spectra (right) for knee joint angle (a), sun spot number (b), and the front-to-back postural sway from infant sitting (c). Unlike the sine wave, determining the highest occurring frequency for actual data is not clear-cut. Here, the highest frequency is approximately 3 Hz for the knee joint angle, 0.1 Hz for sun spot number, and 1 Hz for the front-to-back postural sway.

Much of the discussion so far has focused on acquiring sufficient data—the algorithms require long time series, and higher sampling frequencies ensure that higher frequency components are captured in the data. However, there are reasons not to just sample at the highest frequency possible. One limitation may just be storage space. If you sample at 10,000 Hz instead of 100 Hz, you have 100 times as much data to store. With computer memory being cheaper and cheaper, this is not the problem once was. But do not be confused by the requirement of many of the nonlinear analysis algorithms for having a large number of data points, and think that you can crank up the sampling frequency to get the required number of data points. Remember that the reason you need all those data points is because you need to track the system over a long enough time that the dynamics of the system can be observed. For example, the sunspot data (Figure 2.6b) have a cycle of about 10.5 years. The data were acquired every year for over 307 years. One could gather data every day for 1 year and have an even longer time series (365 data points). But even though the time series is longer, it would not allow you to determine the 10.5-year cycle. This is because you need to let the system evolve for long enough time that the essential dynamics can be captured. Sampling every day does not tell you much about the 10.5-year cycle, because the system has not evolved significantly from one data sample to the next. In other words, the result you get today will be highly correlated with the result you get tomorrow. This concept will be elaborated further in the chapter that contains the discussion of the autocorrelation function (Chapter 8).

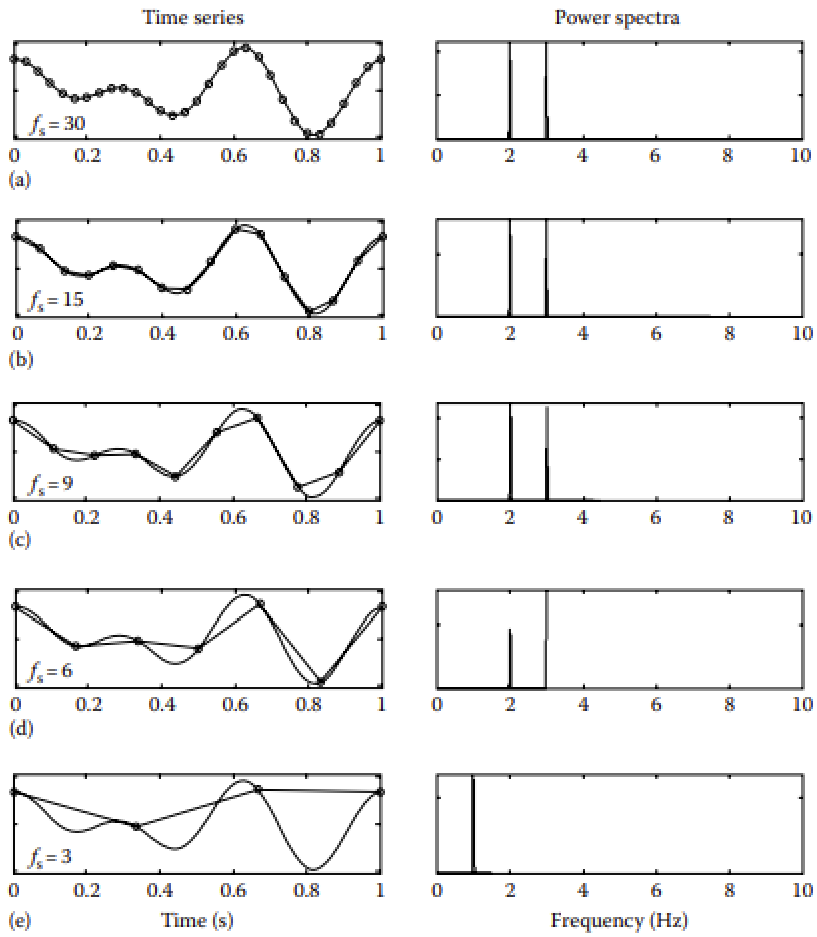

FIGURE 2.8 Time series data (left) and corresponding power spectra (right) for sum of a 2 and a 3 Hz wave function. Smooth plot of left side is the underlying signal; the black circles are data points sampled from the underlying signal, and the sampled data points are connected by black lines to show the time series that one would obtain if sampling at the frequencies indicated (a) 30 Hz, (b) 15 Hz, (c) 9 Hz, (d) 6 Hz, and (e) 3 Hz. The power spectra on the right side were calculated using the sampled data points. As the sampling frequency is reduced, the actual dynamics of the sample are lost in the signal, but until 3 Hz, all frequency peaks are still present. When sampled at 3 Hz, peaks at frequencies that are not part of the actual data begin to show up, which is called aliasing.

Noise

In any experimental measurement, there are always concerns about measurement error or contamination of what you are trying to measure with other information that you are not trying to measure. For example, let us assume that I am measuring the position of a simple pendulum as a function of time. The time series in Figure 2.9a represents the actual position of a pendulum bob. Random noise, representing measurement error, is plotted in Figure 2.9b. Noise is particularly important for the nonlinear analyses. Often the assumption is made that the noise is “random” meaning that there is no correlation between noise at one data point and the noise at another data point. You may recall from physics that white light contains all colors of visible light and that the prism separates the light into component colors. The spectral analysis described here plays on the analogy between separating light into component colors based on frequency and separating time series data into wave functions (sine and cosine functions) with different frequencies. When random noise is broken down into a sum of sine and cosine functions, it has all frequencies that the time series can have represented. Thus, it is called “white” noise to emphasize that it has all frequencies, just as white light has all colors in it. The power spectrum of white noise is flat, as can be seen by looking carefully at the baseline in Figure 2.9b and comparing it to the zero baseline in Figure 2.9a. When the signal is contaminated by white noise, this is also seen in the power spectrum (Figure 2.9c).

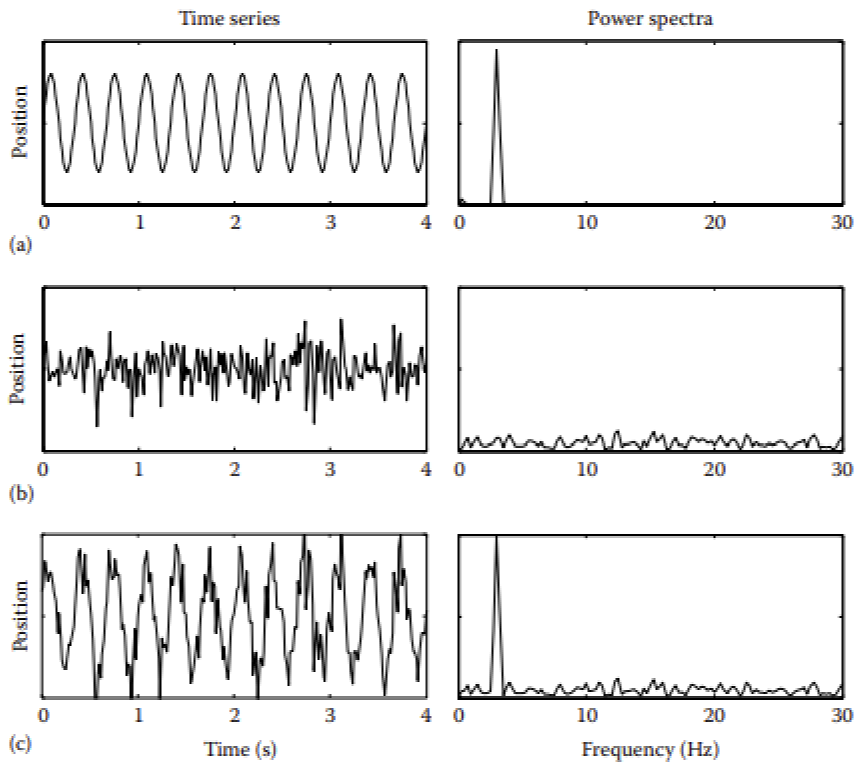

FIGURE 2.9 Sine function (a), white noise (b), and the sum of the sine function and white noise (c). The left side is the time series, and the right side is the power spectrum. The sine function represents the position of a pendulum bob.

However, there is no guarantee that the noise in an experiment will be white noise. For example, the AC power outlets into which data acquisition equipment is plugged are 60 Hz power in North America and many other parts of the world and 50 Hz in Europe. The equipment is designed to minimize the contamination of the output from this power line noise, but the result is not perfect, and so 60 Hz (or 50 Hz in Europe) is a common type of noise that can be seen in data acquired at 120 Hz (or 100 Hz in Europe) or higher sampling frequency (recall the requirement to sample at twice any given frequency to detect it). Thus, the problem with noise in nonlinear dynamics analysis is that you are trying to detect the dynamics of the system of interest, and the experimental data may be contaminated by a signal with unknown dynamics. As in any experiment, anything that can optimize the signaltonoise ratio is beneficial. Additionally, some of the algorithms to be discussed in other chapters may be more robust to experimental noise than others.

Filtering and Smoothing

Experimental noise is a problem in time series analysis. As highlighted in the previous section, interference from other biological signals, movement artifacts or high frequency noise in our acquisition system contaminate the dynamics of our signal of interest. One approach commonly used to deal with these issues in time series analysis is filtering the data. “Filtering” comes from the analogy to light, where a color filter allows only certain wavelengths of light to pass. For example, a red filter absorbs green light, and allows the red light to pass through, so when you look through the filter, everything appears as shades of red. One might call a red filter a “red pass” filter, since red light is allowed to pass through it. Any operation that changes the data by reducing or amplifying components in either the time or frequency domain should be considered filtering. Another common term for filtering is “smoothing,” which comes from the idea that data become smoother when fit to an equation such as polynomials or splines (more about this later) (Woltring 1985; Dohrmann et al. 1988; Vaughan et al. 1999; Giakas 2004). Whether the data should be filtered or not depends on the research question. When asking questions about movement variability, filtering and smoothing are avoided as much as possible. However, these operations are commonly performed in discrete point analyses, especially those that use differentiation with multiple calculation steps such as joint torque and power calculations. Differentiation in this case is the calculation of velocity and acceleration from displacement data. This process preferentially amplifies higher frequencies. Thus, every time you differentiate, the noise becomes larger relative to the signal, so the measurement noise must be filtered to maintain the biological phenomenon. Velocities are noisier than positions and accelerations are noisier than velocities. Position data must be smoothed before calculating velocities and accelerations. The first central difference method (of differentiation) is a type of smoothing: averaging. To check whether the proper smoothing was performed, the position, velocity, and acceleration should be graphed. The more filtering that is performed, or the more frequencies that are removed, the “smoother” the signal will be. However, this is exactly when you need to consider if you have missed any important true biological phenomena, especially high-frequency impact phenomena. This is notoriously the case with running related biomechanical data in the literature where kinematics are filtered with a cutoff below 6 Hz even though important high-frequency phenomena exist between 12 and 20 Hz (Giakas 2004). We believe that one of the reasons for the lack of a true understanding of the mechanisms behind running injuries is actually the massive contamination of the literature with inaccurate results which stifle scientific progress (Stergiou et al. 1999; Giakas 2004).

A common type of smoothing is using polynomials. In this method, the data are forced to fit a certain mathematical model. This is a poor method because there is limited control over what data are included or excluded, so it is likely that true data will be removed. A better method that is an extension of the idea of using polynomials to smooth is the spline. A spline function consists of a number of low-order polynomials that are pieced together. Cubic, or third order, and quartic, or fourth order, splines are the most popular for biomechanics applications. Another type of smoothing is digital filtering. Again, the basic premise of filtering in time series analysis is based on the idea of breaking up the measured signal into its various frequency components using the spectral analysis techniques discussed earlier, and then removing frequencies that are not of interest. Filtering is performed by setting a key frequency known as the cutoff frequency. The cutoff informs the filter to keep or remove subsequent or remaining frequencies. Filters also differ based on the operation used to change the frequency components, which is called the transfer function. The type of transfer function determines which frequencies are kept and which are filtered. Three different bands can be used in a transfer function, including a pass, transition, and stop. A pass preserves the specified frequency components, a transition band progressively decreases the power of the frequency components covered, and the stop band removes all remaining frequencies.

There are also different algorithms that can be implemented to correctly filter the data. Two of the most common implementations of this technique are the Butterworth filter, the critically damped filter, and the Jackson filter (Smith 2002). The algorithms will use the selected cutoff and contain different bands depending on what type of data are being smoothed. One can select whether to remove frequencies above a certain cutoff frequency, called a “low-pass” filter, because lower frequencies are allowed to pass through the filter. A “high-pass” filter would block low frequencies and allow high frequencies to pass through. A “band pass” or “notch pass” filter passes through frequencies in an intermediate range, while rejecting higher and lower frequencies. Because random noise is a high-frequency component of the measured signal, a low-pass filter is used to remove it. The cutoff frequency must be carefully selected to remove noise without removing the signal of interest. Figure 2.10b shows that a cutoff frequency above the three features of interest leaves them intact, but if the cutoff frequency is too low (Figure 2.10c and d), the peaks of interest are removed. It is important to select a cutoff frequency that will preserve most of the data of interest. Typically, a cutoff frequency that maintains 99% of the data is chosen. The roll-off is the transition between areas to maintain or smooth. The slope of the cutoff is the roll-off and this changes with the order of the polynomial. As the order is increased, the sharpness of the slope is increased and vice versa.

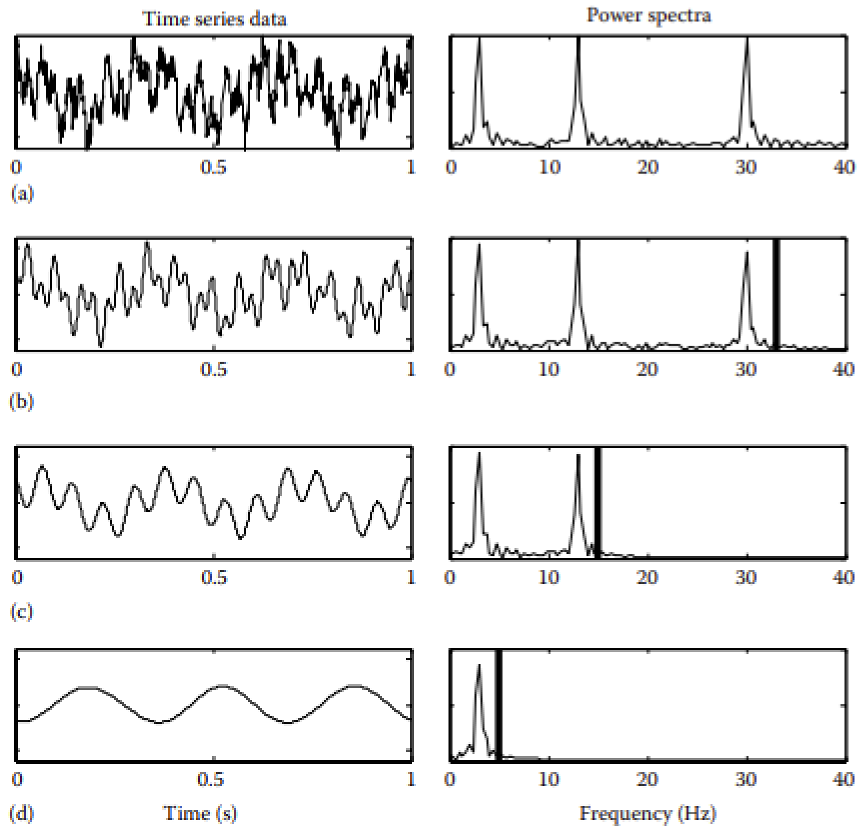

FIGURE 2.10 Time series left and power spectrum right: (a) unfiltered, (b) filtered using a cutoff frequency of 33 Hz, (c) filtered using a cutoff frequency of 15 Hz, and (d) filtered using a cutoff frequency of 5 Hz. The filter used was a "brick wall" filter (very high-order Butterworth filter), and cutoff frequency is indicated by vertical bar.

There are many decisions to be made when it comes to whether and how to filter data. Ultimately, the hypotheses and research questions should determine which filter should be implemented. One way to make sure that filtering is not altering the data in a way that changes the phenomenon is to filter interactively. This is important because a computer has no way to know when the filter being implemented is affecting the data. As the filter is being applied, the raw data should be graphed with the filtered data to make sure that phenomena are not being smoothed out of the data. A good starting point to know what not to filter is to think back to sampling theorem. Based on Nyquist sample theory, data under the Nyquist frequency should not be filtered. The discussion later regarding the potential biological importance of nonstationarity further emphasizes this point.

Filtering data is a very common method of data manipulation for many types of linear analyses. The problem is that most of the filtering techniques are based on statistically preserving the linear features of the data. These methods, such as Butterworth filtering, were not designed to be used on data that will subsequently be analyzed using nonlinear techniques. Thus, there is no reason to believe that the nonlinear dynamics of the time series would still be intact after filtering, and in fact, one would expect that these methods would be counterproductive for nonlinear analysis (Rapp et al. 1993; Theiler and Eubank 1993). Some nonlinear filters have been developed, but again testing them on experimental data is problematic because the nature of the underlying attractor, if there is one, is unknown. Thus, it is likely the filter is distorting the underlying dynamics (Kantz and Schreiber 2003). Data previously described or seen as “noisy” have been shown in recent literature to have deterministic patterns and provide important information about the dynamics of the system. Filtering the time series can alter the embedding dimension needed to properly reconstruct the state space and can influence results of calculations of the time lag and others. In nearly all situations presented in this book, it is recommended that filtering be avoided (or at least be considered with extreme care) to capture the true dynamics of the system.

Resolution

The concept of significant figures is probably familiar from high school. If a person’s height is measured by comparing with the door frame, then the person might be measured to be about 72 in. (two significant figures). If a tape measure is used, the person’s height might be measured as 72.75 in. (four significant figures). The more significant figures used, the more precision the measurement has, and the greater the number of figures that need to be recorded in the lab notebook. A similar issue arises when using computers for data acquisition in that different measurement techniques have different levels of precision associated with them. For example, if one is using video cameras to record knee angles, the resolution of the camera will limit the precision with which the knee position can be measured. If the knee position stored in the computer is examined, then after seeing 13.67485362842 stored as knee position for a given frame, it might be concluded that there are 13 significant figures. The problem is that typically the computer can store more significant figures than your measurement technique can provide. If one divides one (one significant figure) by three (one significant figure), your calculator will show 0.333333333, with as many three’s as the screen will allow it to show. Doing division does not increase the number of significant figures in the result. Similarly, knee position of 13.67485362842 is limited by the pixel resolution of the camera, not by how many digits the computer can store. Some measurement techniques have better precision than others, so you need to be familiar with the limitations of the equipment being used.

The algorithms that have been developed for performing nonlinear analysis are often generated and tested by people who are not working with experimental data, but rather work with computer-generated time series. Computer programs are written to produce a time series based on known equations, the Lorenz equations, for example. The time series created in this manner has no noise, and precision as high as the computer is capable of storing. Experimental data, on the other hand, will likely be contaminated with some measurement noise and will only be able to measure the parameter of interest to a limited degree of precision. Why is a little round-off error a problem? A chaotic-behaving system is sensitive to small perturbations, frequently discussed in terms of sensitivity to initial conditions. Some of the algorithms used to study such systems may give different results depending on the precision of the data used. In fact, Lorenz discovered this property of chaotic systems when he entered the result of a previous calculation by hand into his computer and rounded it off as he entered it, only to find that the rounded off number gave dramatically different results in the calculation than the higher precision number that was not entered by hand on another run (Kerry 2008). In general, the significant digits should be used, as determined by the equipment measuring the phenomenon. It is recommended that you test the effect of rounding the data on the outcome value of the particular algorithm used to see how rounding may change the results. A detailed guide for testing the limitations of any particular algorithm can be found at the end of the chapter, which will guide decisions on the number of digits to use for time series analysis.

Stationarity

The concept of stationarity is, in vague terms, the requirement that there is statistical similarity of successive parts of a time series. It indicates that the mean and the variance should not change as a function of time in the time series. In Figure 2.11a, there is a stationary white noise time series, in Figure 2.11b, the time series is not stationary because the mean is different in the first half compared to the second half, and in Figure 2.11c, the time series is not stationary because the variance is different in the first half compared to second half. The sound of a cymbal clashing, if hit only once, is not stationary because the acoustic power of the clash, and hence, its variance diminishes with time.

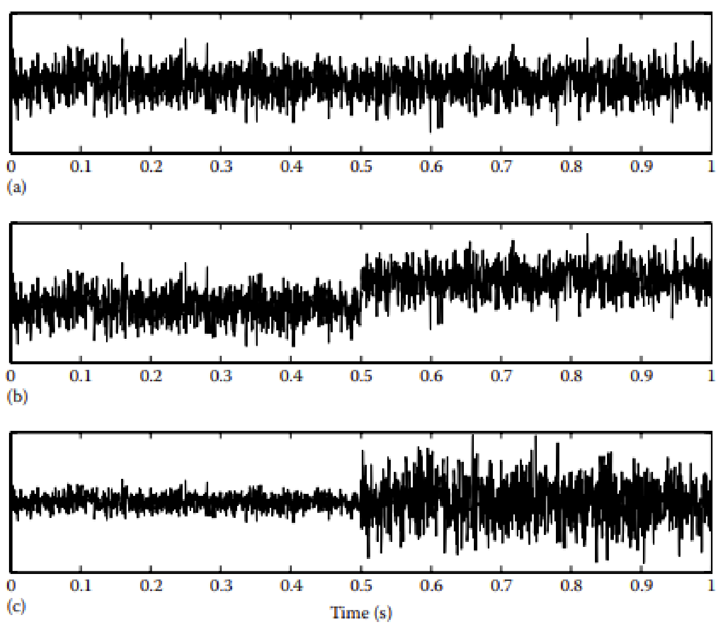

FIGURE 2.11 White noise time series data showing (a) stationarity, (b) nonstationarity due to change in mean, and (c) nonstationarity due to change in variance.

The stationarity issue is perplexing because the exact nature of the stationarity required for nonlinear analysis is not clear. The stationarity requirement for the experimental data comes from an assumption commonly made by mathematicians in the derivation of the mathematical algorithms that are used for the nonlinear analysis. For example, Pincus assumes stationarity in the derivation of the algorithm for approximate entropy (Pincus 1991), and Wolf et al. assume stationarity in the derivation of the algorithm for maximum Lyapunov exponent (Wolf et al. 1985). There are nominally two types of stationarity for mathematicians, strong stationarity and weak stationarity. Strong stationarity requires all possible moments of ensembles (or shorter subsets) of time series data to be time invariant. In mathematics, a moment is, loosely speaking, a quantitative measure of the shape of a set of points. The first moment of the distribution of the random variable is the population mean. The second moment measures the width or distribution of a set of points in one dimension or in higher dimensions and can be represented as the shape of a cloud of points as it could be fit by an ellipsoid. Weak stationarity only makes this requirement for the mean and the autocorrelation function of time series data to be time invariant (Bendat and Piersol 2000). However, even weak stationarity is a difficult requirement to meet in experimental time-series data. For example, a sine function is not a stationary signal, since the mean for data that are near the bottom of a trough is not the same as the mean for data points that are near the top of a crest. Thus, none of the time series presented in this chapter, except for Figures 2.9b and 2.11a (white noise), would meet the requirement of stationarity.

Lack of stationarity or nonstationarity is often discussed as a limitation to other forms of nonlinear analysis. However, some authors have attempted to quantify nonstationarity of a time series as a useful measure in itself (Rieke et al. 2003; Cao et al. 2004; Chau et al. 2005; Makinen et al. 2005; Gourevitch and Eggermont 2007; Tong et al. 2007). For example, some of the debate about global climate change centers around the stationarity, or lack thereof, of climate-related variables (Karner 2002; Gagan et al. 2004; Mailhot et al. 2007). More specifically, nonstationarity may be inherent to biological systems and should actually be embraced and studied and not be considered as a limitation. Further, multiple researchers suggested that nonstationarities in physiological data may be the result of fractal properties, which has led to investigations of studying these long-term fluctuations in the data (Stanley 1971; Goldberger and West 1987; Meesman et al. 1993; Peng et al. 1993; Turcott and Teich 1993, 1996; Goldberger 1996; Viswanathan et al. 1997). Nonstationarity as a true physiological phenomenon can occur because such processes are more complex than the concept of homeostasis; instability or nonstationarity in the behavior demonstrate that the dynamics can occur over multiple timescales (Viswanathan et al. 1997; Pavlov et al. 2006). For example, in an electroencephalogram, the interspike intervals formed three distinct orbits during the hour before the occurrence of an epileptic seizure. In this case, nonstationarity in the data indicated the onset of a seizure (So et al. 1998). In the analysis of interbeat intervals of healthy individuals and individuals with congestive heart failure, the diseased population demonstrated greater inconsistency across multiple time scales. Even though the interbeat intervals of healthy individuals did change time scales, the behaviors displayed similar dynamics across those multiple scales, whereas individuals with congestive heart failure (CHF) were unable to regulate the interbeat intervals over particular time scales (Viswanathan et al. 1997). The concept of nonstationarity as a true biological phenomenon is supported by the fact that systems are under neurophysiological control and those control mechanisms, in the absence of disease, can regulate activity in many varying situations (i.e., time scales).

A commonly used technique to remove nonstationarity of time series data is to difference the data (Chatfield 2003). Differencing is the process of subtracting values between two data points to create a new point in a "differenced" series. Thus, a new time series is created, where the value at each point is found by subtracting the data at that time point from the next data point, resulting in a time series shorter by one than in the original time series. This procedure may be repeated multiple times, as necessary, until the data appear to be stationary. For example, investigators in our laboratory applied this technique to time series data from infant sitting postural control, and the approximate entropy was calculated before and after differencing. The results showed no benefit to differencing the data, as the approximate entropy was better correlated with developmental variables if no differencing was done. Our conclusion was that approximate entropy is robust to the requirement of mathematical nonstationarity (Deffeyes et al. 2007).

Another technique that can address the issue of nonstationarity is detrending. Briefly, detrending occurs before the application of a nonlinear algorithm, usually as a first step in the calculation process. A good example is detrended fluctuation analysis (DFA Chapter 7) that fits a power law to the series' average fluctuations across different scales (Viswanathan et al. 1997). These scales come in the form of box sizes and the least square line is fitted to the data in each box. The original data are detrended by subtracting the least square line of each window, which makes the time series stationary. This technique is described in more detail in Chapter 7. A note of caution, detrending must only be done if you are certain that the trend is not part of the dynamics of the signal. This would occur in the case of errors in measurement due to calibration, or drift from the equipment signal (Kantz and Schreiber 2003).

Some experimentalists have a more pragmatic concept of stationarity. The need for requiring a stationary time signal is that the underlying system dynamics must not change over the course of the data acquisition. If the system jumps from a chaotic behavior to a limit cycle behavior in the middle of the time series, clearly the system is nonstationary, and application of any of the nonlinear analysis algorithms discussed in this book would be problematic. However, the time series could potentially be divided into multiple stationary signals, evaluated, and compared. In this way, stationarity within signals is maintained and nonlinear algorithms can be implemented and used to describe differences between signals. There is also precedence for interpreting stationarity or lack thereof as a motor control technique to understand the dynamics of a system (Newell 1997; Stergiou et al. 2004).

An interesting description of stationarity is also provided by Small (2005). Small notices that the definition of stationarity is not the same for all systems. For linear systems, there is stationarity if all its moments (statistical descriptors) remain the same over time. A nonstationary system is then one that has some type of temporal dependence that arises from an external source. Thus, if we extend the definition of the system to include all such external sources, then the system is stationary. This reminds us of dynamical systems theory in which a system cannot be considered without its constraints, namely morphological, environmental, and biomechanical constraints.

Small (2005) also provides an interesting example to illustrate the earlier discussion, as he describes someone standing on the beach watching the ups and downs of the tide. This individual will describe this system as nonstationary. However, if the same individual considers the relative positions of the earth, the moon, and the sun, then he could determine that if all are studied together, he will have an approximately stationary system (Small 2005).

A GENERAL NOTE ON EXPERIMENTAL LIMITATIONS OF TIME SERIES DATA

One approach to understanding the effect of these limitations on your data as you are using any particular algorithm is to use computer-generated data from several systems with known dynamics, such as a sine function, the Lorenz equations, the Henon map, pseudorandom noise, and so forth, and manipulate them to have the same limitations (data length, resolution, etc.) as your experimental data. For example, you can round the data off to have a precision similar to the experimental data, add noise, and/or add a linear trend to increase nonstationarity, and then run the algorithm of interest on these different time series and observe if the result is consistent with expectations based on the system dynamics that generated the time series. One way to add noise is to use a pseudorandom number generator, such as the “randn” function in MATLAB®, to generate a random time series and then add that to clean computer-generated time series from a system with known dynamics. One issue is that the dynamics of the noise that may be contaminating the experimental data are not known. It may not be random. It could be periodic, as in the case of contamination with 60 Hz (or 50 Hz in Europe), or it could be chaotic, generated by a process that appears to be random but actually has some structure to it. Dealing with experimental limitations of time series data is one of the greatest challenges to successful application of nonlinear analysis algorithms to systems with unknown dynamics (Rapp 1994). However, by exploring the effect of each limitation, the dynamics of the time series can be understood much better and the benefits could outweigh the time expense. In closing, remember the following considerations when conducting time series analysis:

- *Length*: Make sure data are long enough to capture a minimum of 5–10 cycles of the phenomenon being studied; the more data present, the more likely the true dynamics have been captured in the signal.

- *Sampling frequency*: Sample at the correct rate of frequency at least 2 times, but not more than 10 times the highest frequency. This will make sure the entire signal is captured without adding noise to the signal. The frequency in the data is determined with a spectral analysis.

- *Noise*: Optimize the signal-to-noise ratio and be aware of how the nonlinear algorithm used is affected by noise.

- *Filtering and smoothing*: Be very cautious in filtering data intended to use for nonlinear analysis. If filtering must be done, ensure the cutoff is above the Nyquist frequency and perform the filtering interactively.

- *Resolution*: Determine how many significant digits to be used based on the precision of the equipment collecting the data. Test the effect of rounding the data on the outcome value of the nonlinear algorithm of choice.

- *Stationarity*: Nonstationarity may be an important biological measure in itself. Various techniques can be used to remove nonstationarity; however, these techniques may impact the dynamics of the signal. Some algorithms are affected by stationarity less than others.

EXERCISES

1. Define a time series.

2. What is a common unit of measure for how many data points are acquired every second?

3. How long does your data have to be?

4. Let {1, 5, 20, 2, 35} be the original time series. Create a differenced time series.

5. Competitive cyclists pedal with a cadence of about 90 cycles/min. What is the minimum sampling frequency you would need to estimate the cadence from knee flexion time series data? What would be a preferred sampling rate?

6. A physiologist measures the electrical signal from a research subject’s muscle as he flexes his elbow using electromyography (EMG). Spectral analysis shows frequencies lower and higher than the range of interest. What sort of filter might the physiologist want to apply to the data to remove the unwanted low and high frequencies? Another researcher only wants to remove the high frequencies. What sort of filter should the second researcher apply?

7. For a period of 15 years an ornithologist measured the annual population of the white-faced scops owl living in a patch of forest, and found the following result: [11 13

15 16 15 17 19 22 21 21 24 27 26 28 29]. Make a time series plot of the data. Does it appear to be stationary? Difference the data and make a plot of the differenced time series. Does the differenced time series appear more stationary than the original time series?

8. If you have access to MATLAB, make a plot of a sine function from time 0 to 10 s. Sample at a frequency of 100 Hz, and make the sine function with a with a frequency of 0.5 Hz and amplitude of 5. Restrict the axes to 3–4 s, and −10 to +10 in sine amplitude.

9. Make a plot of a sine function from time 0 to 10 s, with a frequency of 0.5 Hz and add in Gaussian distributed random noise, with a mean of 0 and a standard deviation of 1.

10. Calculate the power spectra of both of the plots you made in questions #4 and #5.

REFERENCES

Beard, J. 2003. The FFT in the 21st Century: Eigenspace Processing. New York: Springer.

Bendat, J.S. and A.G. Piersol. 2000. Random Data: Analysis and Measurement Procedures, 3rd edn. New York: John Wiley & Sons.

Box, G.E.P., G.M. Jenkins, and G.C. Reinsel. 2008. Time Series Analysis: Forecasting and Control, 4th edn. Hoboken, NJ: John Wiley & Sons, Inc.

Brockwell, P.J. and R.A. Davis. 2002. Introduction to Time Series and Forecasting, 2nd edn. New York: Springer.

Cao, H.Q., D.E. Lake, M.P. Griffin, and J.R. Moorman. 2004. Increased nonstationarity of neonatal heart rate before the clinical diagnosis of sepsis. Annals of Biomedical Engineering 32(2): 233–244.

Chatfield, C. 2003. The Analysis of Time Series: An Introduction, 6th edn. Boca Raton, FL: Chapman & Hall/CRC Press. Chau, T., D.

Chau, M. Casas, G. Berall, and D.J. Kenny. 2005. Investigating the stationarity of pediatric aspiration signals. IEEE Transactions on Neural Systems and Rehabilitation Engineering 13(1): 99–105.

Deffeyes, J.E., R.T. Harbourne, S.L. DeJong, W.A. Stuberg, A. Kyvelidou, and N. Stergiou. 2007. Approximate entropy is robust to non-stationarity in analysis of infant sitting postural sway. Paper presented at Proceedings of the 2007 American Society of Biomechanics Annual Meeting, Stanford, CA.

Dohrmann, C.R., H.R. Busby, and D.M. Trujillo. 1988. Smoothing noisy data using dynamic programming and generalized cross-validation. Journal of Biomechanical Engineering 110(1): 37–41.

Fredrickson, B.L. and M.F. Losada. 2005. Positive affect and the complex dynamics of human flourishing. American Psychologist 60(7): 678–686.

Gagan, M.K., E.J. Hendy, S.G. Haberle, and W.S. Hantoro. 2004. Post-glacial evolution of the Indo-Pacific Warm Pool and El Niño-Southern Oscillation. Quaternary International 118: 127–143.

Giakas, G. 2004. Power spectrum analysis and filtering. In Innovative Analysis of Human Movement, 1st edn. N. Stergiou, ed., pp. 223–258. Champaign, IL: Human Kinetics.

Giakas, G. and V. Baltzopoulos. 1997. Time and frequency domain analysis of ground reaction forces during walking: An investigation of variability and symmetry. Gait and Posture 5(189): 197.

Giakas, G., V. Baltzopoulos, P.H. Dangerfield, J.C. Dorgan, and S. Dalmira. 1996. Comparison of gait patterns between healthy and scoliotic patients using time and frequency domain analysis of ground reaction forces. Spine 21(19): 2235–2242.

Goldberger, A.L. 1996. Non-linear dynamics for clinicians: Chaos theory, fractals, and complexity at the bedside. Lancet 347(9011): 1312–1314.

Goldberger, A.L. and B.J. West. 1987. Applications of nonlinear dynamics to clinical cardiology. Annals of the New York Academy of Sciences 504: 195–213.

Gourevitch, B. and J.J. Eggermont. 2007. A simple indicator of nonstationarity of firing rate in spike trains. Journal of Neuroscience Methods 163(1): 181–187.

Huang, N.E., Z. Shen, S. Long, M.C. Wu, H.H. Shih, Q. Zheng, N. Yen, C.C. Tung, and H.H. Liu. 1998. The empirical mode decomposition and the Hilbert spectrum for nonlinear and nonstationary time series analysis. Proceedings of the Royal Society London A 454: 903–995.

Kantz, H. and T. Schreiber. 2003. Nonlinear Time Series Analysis, 2nd edn. Cambridge, U.K.: Cambridge University Press.

Karner, O. 2002. On nonstationarity and antipersistency in global temperature series. Journal of Geophysical Research-Atmospheres 107(D20): 4415.

Kerry, E. May 23, 2008. Retrospective—Edward N. Lorenz (1917–2008). Science 320(5879): 1025.

Liu, P.Y., A. Iranmanesh, D.M. Keenan, S.M. Pincus, and J.D. Veldhuis. 2007. A noninvasive measure of negative-feedback strength, approximate entropy, unmasks strong diurnal variations in the regularity of LH secretion. American Journal of Physiology Endocrinology and Metabolism 293(5): E1409–E1415.

Mailhot, A., S. Duchesne, D. Caya, and G. Talbot. 2007. Assessment of future change in intensity-duration-frequency (IDF) curves for southern Quebec using the Canadian regional climate model (CRCM). Journal of Hydrology 347(1–2): 197–210.

Makinen, V.T., P.J.C. May, and H. Tiitinen. 2005. The use of stationarity and nonstationarity in the detection and analysis of neural oscillations. NeuroImage 28(2): 389–400.

McGrath, D., T.N. Judkins, I.I. Pipinos, J.M. Johanning, and S.A. Myers. 2012. Peripheral arterial disease affects the frequency response of ground reaction forces during walking. Clinical Biomechanics 27(10): 1058–1063.

Meesmann, M., F. Gruneis, P. Flachenecker, and K.D. Kniffki. 1993. A new method for analysis of heart-rate-variability counting statistics of 1/f fluctuations. Biological Cybernetics 68(4): 299–306.

Myers, S.A., J.M. Johanning, I.I. Pipinos, K.K. Schmid, and N. Stergiou. 2013. Vascular occlusion affects gait variability patterns of healthy younger and older individuals. Annals of Biomedical Engineering 41(8): 1692–1702.

Myers, S.A., J.M. Johanning, N. Stergiou, R.I. Celis, L. Robinson, and I.I. Pipinos. 2009. Gait variability is altered in patients with peripheral arterial disease. Journal of Vascular Surgery 49(4): 924–931.

Myers, S.A., I.I. Pipinos, J.M. Johanning, and N. Stergiou. 2011. Gait variability of patients with intermittent claudication is similar before and after the onset of claudication pain. Clinical Biomechanics 26(7): 729–734.

National Office of Oceanic and Atmospheric Administration. 2013. Available from ftp://ftp.ngdc. noaa.gov/STP/SOLAR_DATA/SUNSPOT_NUMBERS/YEARLY, cited March 2008.

Newell, K.M. 1997. Degrees of freedom and the development of center of pressure profiles. In Applications of Nonlinear Dynamics to Developmental Process Modeling. K.M. Newell and P.M.C. Molenar, eds., pp. 63–84. Hillsdale, NJ: Erlbaum.

Pavlov, A.N., V.A. Makarov, E. Mosekilde, and O.V. Sosnovtseva. 2006. Application of wavelet-based tools to study the dynamics of biological processes. Briefings in Bioinformatics 7(4): 375–389.

Peng, C.K., J. Mietus, J.M. Hausdorff, S. Havlin, H.E. Stanley, and A.L. Goldberger. 1993. Long-range anticorrelations and non-Gaussian behavior of the heartbeat. Physical Review Letters 70(9): 1343–1346.

Percival, D.B. and A.T. Walden. 1993. Spectral Analysis for Physical Applications, 1st edn. Cambridge, MA: Cambridge University Press.

Pincus, S.M. March 15, 1991. Approximate entropy as a measure of system complexity. Proceedings of the National Academy of Sciences of the United States of America 88(6): 2297–2301.

Prabhu, K.M.M. 2013. Window Functions and Their Applications in Signal Processing, 1st edn. Boca Raton, FL: CRC Press.

Rapp, P.E. July–September 1994. A guide to dynamical analysis. Integrative Physiological and Behavioral Science 29(3): 311–327.

Rapp, P.E., A.M. Albano, T.I. Schmah, and L.A. Farwell. April 1993. Filtered noise can mimic low-dimensional chaotic attractors. Physical Review E: Statistical Physics, Plasmas, Fluids, and Related Interdisciplinary Topics 47(4): 2289–2297.

Rieke, C., F. Mormann, R.G. Andrzejak, T. Kreuz, P. David, C.E. Elger, and K. Lehnertz. May 2003. Discerning nonstationarity from nonlinearity in seizure-free and preseizure EEG recordings from epilepsy patients. IEEE Transactions on Bio-Medical Engineering 50(5): 634–639.

Sakakibara, Y. 1995. Introduction. In Who is Fourier? A Mathematical Adventure. Boston, MA: Language Research Foundation.

Small, M. 2005. Applied Nonlinear Time Series Analysis, 1st edn. Singapore: World Scientific Publishing.

Smith, S.W. 2002. Digital Signal Processing: A Practical Guide for Engineers and Scientists, 1st edn. Boston, MA: Newnes.

So, P., J. Francis, T. Netoff, B.J. Gluckman, and S.J. Schiff. 1998. Periodic orbits: A new language for neuronal dynamics. Biophysical Journal 74: 2776–2785.

Stanley, H.E. 1971. Introduction to Phase Transitions and Critical Phenomena. International Series of Monographs on Physics. New York, NY: Oxford University Press.

Stergiou, N., B.T. Bates, and S.L. James. 1999. Asynchrony between subtalar and knee joint function during running. Medicine and Science in Sports and Exercise 31(11): 1645–1655.

Stergiou, N., U.H. Buzzi, M.J. Kurz, and J. Heidel. 2004. Nonlinear tools in human movement. In Innovative Analysis of Human Movement. N. Stergiou, ed., pp. 63–90. Champaign, IL: Human Kinetics.

Stergiou, N., G. Giakas, J.B. Byrne, and V. Pomeroy. 2002. Frequency domain characteristics of ground reaction forces during walking of young and elderly females. Clinical Biomechanics 17(8): 615–617.

Stoica, P. and R.L. Moses. 1997. Introduction to Spectral Analysis, 1st edn. Upper Saddle River, NJ: Prentice Hall.

Theiler, J. and S. Eubank. October 1993. Don’t bleach chaotic data. Chaos 3(4): 771–782.

Tong, S., Z. Li, Y. Zhu, and N.V. Thakor. 2007. Describing the nonstationarity level of neurological signals based on quantifications of time-frequency representation. IEEE Transactions on Biomedical Engineering 54(10): 1780–1785.

Turcott, R.G. and M.C. Teich. 1993. Long-duration correlation and attractor topology of the heartbeat rate differ for healthy patients and those with heart failure. Proceedings of the Society of Photo-Optical Instrumentation Engineers, Bellingham, WA.

Turcott, R.G. and M.C. Teich. 1996. Fractal character of the electrocardiogram: Distinguishing heart-failure and normal patients. Annals of Biomedical Engineering 24(2): 269–293.

Vaughan, C., B. Davis, and J. O'Connor. 1999. Dynamics of Human Gait. Cape Town, South Africa: Kiboho Publishers.

Viswanathan, G.M., C.K. Peng, H.E. Stanley, and A.L. Goldberger. 1997. Deviations from uniform power law scaling in nonstationary time series. Physical Review E 55(1): 845–849.

Warner, R.M. 1998. Spectral Analysis of Time-Series Data. New York: Guilford Press.

Wolf, A., J.B. Swift, H.L. Swinney, and J.A. Vastano. 1985. Determining Lyapunov exponents from a time series. Physica 16D: 285–317.

Woltring, H.J. September 1985. On optimal smoothing and derivative estimation from noisy displacement data in biomechanics. Human Movement Science 4(3): 229–245.

Wurdeman, S.R., J.M. Huisinga, M. Filipi, and N. Stergiou. 2011. Multiple sclerosis affects the frequency content in the vertical ground reaction forces during walking. Clinical Biomechanics 26(2): 207–212.