Chapter 2

Quantifying Movement Variability

2.1 Introduction

It has been stated in Chapter 1 that movement variability can be modelled and quantified with methods of nonlinear analysis mainly because (i) the structures of the human physiology (e.g. lungs, neurons, etc.) suggest that many of their dynamics are controlled with nonlinear dynamics (Goldberger et al., 1990) and (ii) data from human movement can be noisy, deterministic, stochastic, non-stationary or deterministic-chaotic (Caballero et al., 2014; Hatze, 1986; Newell and Slifkin, 1998; Preatoni et al., 2010, 2013; Stergiou and Decker, 2011; Stergiou et al., 2006). Therefore, in this chapter fundamentals of time series, methods of nonlinear analysis to quantify movement variability and nonlinear analysis with real-world data are reviewed.

2.2 Fundamentals of time-series analysis

Biosignals from living systems can typically be noisy, deterministic, stochastic, non-linear, non-stationary or deterministic-chaotic (Caballero et al., 2014; Gómez-García et al., 2014; Harbourne and Stergiou, 2009; Hatze, 1986; Klonowski, 2007; Newell and

Slifkin, 1998; Stergiou and Decker, 2011; Stergiou et al., 2006; Wijnants et al., 2009). That said, the following sections provide fundamental definitions of time series for this thesis.

2.2.1 Linear and non-linear systems

Linear systems are proportional or additive. For example, the interaction between variables of a linear system are negligible whereas for a nonlinear system such interaction of variables can produce emergent properties arising from the initial conditions of the system (Klonowski, 2007).

2.2.2 Stationary and non-stationary signals

Stationary signals have the same mean and variance as time progress (e.g. a sinusoidal signal), however such stationary signal can also be changeable (e.g. alternative sinusoidal signal). In contrast, when statistics of the time series change with time then such a signal is known as non-stationary signal. Non-stationary signals are therefore characterised by transients and drift over time. Examples of non-stationary signals are the time series of seasonal trends and changes (Kitagawa and Gersch, 1984) or Electroencephalography (EEG) signals which present different and changeable intensity over time (Klonowski, 2007).

2.2.3 Deterministic and stochastic systems

A deterministic system is predictable. Deterministic systems have a small number of variables of importance. Deterministic systems are hence modelled with linear ordinary differential equations and their initial conditions and constants. In contrast, stochastic systems are non-predictable and therefore have more variables of equal importance and are typically modelled with probability theory (Klonowski, 2007).

2.2.4 Deterministic-chaotic time series

Deterministic signals can dramatically change with a slight change of initial conditions and then after a long time-scale, the signal can appear to be stochastic (Amato, 1992). Klonowski 2007, p. 11 pointed out that "chaotic systems behave like they were stochastic but they are also deterministic", meaning that chaotic systems are predictable for a short time-scale but nonpredictable in a long time-scale because of the initial conditions of the systems. Preatoni et al. 2013, p. 78, in experiments in sport science, mentioned that "variability is likely to have both deterministic and a stochastic origin". It can then be assumed that time series for human body movement are neither independent nor stochastic but deterministic-chaotic (Harbourne and Stergiou, 2009; Stergiou and Decker, 2011; Stergiou et al., 2006).

2.3 Quantifying movement variability with nonlinear analysis

Previous studies have shown that movement variability is not considered as a undesired factor that creates errors but a signature for assessment of healthiness (associated with unhealthy pathological states) or skillfulness (associated with the functionality of movement) (Stergiou and Decker, 2011). That said, movement variability can fundamentally be either quantified based on (i) the magnitude of the variability or (ii) the dynamics of the variability (Caballero et al., 2014). However, finding the appropriate methods to quantify movement variability is still an open problem.

For instance, Preatoni et al. (2010, 2013) point out that conventional statistics (e.g. standard deviation, coefficient of variation, intra-class correlation coefficient) only quantify the overall variability. Also, Stergiou and Decker (2011) stated that statistical tools (e.g. mean, standard deviation and range) are a measure of centrality,

meaning such metrics are compared around a central point. Similarly, Coffey et al. (2011) pointed out that the use of means and standard deviations led to reduction of data and therefore information is discarded.

Additionally, one can apply frequency-domain tools to quantify movement variability. For example, Hatze (1986) proposed a measure of dispersion to quantify the deviation of motion from a certain reference using the Fourier series. However, deviations of motion are from angular coordinates (radians) and linear coordinates (meters) which made them an unacceptable fusion of variables. Vaillancourt et al. (2001) pointed out that it is rare for frequency and amplitude to differ in postural tremor of patients with Parkinson's disease but differences in time-dependent structures are apparent, and associated with a change of regularity of postural tremor. Klonowski (2002, 2007, 2009) stated that frequency-domain tools require stationary data, otherwise using other type of data might create misleading results.

Applying either statistical tools or frequency-domain tools to quantify movement variability might create misleading results, specially when dealing with deterministic-chaotic signals (Amato, 1992; Dingwell and Cusumano, 2000; Dingwell and Kang, 2007; Miller et al., 2006). Hence, the properties of deterministic-chaotic signals are aligned with the subtle changes in the neuro-muscular-skeletal system are caused by influences of environmental changes, training or latent pathologies (Preatoni et al., 2010, 2013) and that movement variability involves evolution of human movement and the exploratory nature of movement (Caballero et al., 2014; Stergiou and Decker, 2011). That said, Caballero et al. (2014); Preatoni et al. (2010); Stergiou and Decker (2011) highlighted that movement variability can be better described and quantified with different methods of nonlinear analysis such as: largest Lyapunov exponent (Bruijn et al., 2009; Donker et al., 2007; Kurz et al., 2010; Yang and Wu, 2011), fractal analysis (Delignlères et al., 2003), entropy rate (Cavanaugh et al., 2010), Sample Entropy

(SampEn) (Donker et al., 2007; Liao et al., 2008; Richman and Moorman, 2000; Stins et al., 2009; Vaillancourt et al., 2004), Approximate Entropy (ApEn) (Cavanaugh et al., 2010; Kurz and Hou, 2010; Pincus, 1991; Sosnoff et al., 2006; Sosnoff and Voudrie, 2009), Fuzzy Entropy (FuzzyEn) (Chen et al., 2007), Multiscale Entropy (MSE) (Costa et al., 2002), Permutation Entropy (PE) (Bandt and Pompe, 2002; Vakharia et al., 2015), Quadratic Sample Entropy (QSampEn) (Lake and Moorman, 2011), Amplitude-aware permutation entropy (AAPE) (Azami and Escudero, 2016), Detrended Fluctuation Analysis (DFA) (Gates and Dingwell, 2007, 2008; Hausdorff, 2009) and Recurrence Quantification Analysis (RQA) (Marwan, 2008; Trulla et al., 1996; Zbilut and Webber, 1992).

Having so many nonlinear tools to measure movement variability (MV) led Caballero et al. 2014, p. 67 to raise the following question: "Is there a best tool to measure variability?" which lead me to ask two questions for this thesis: (i) what to quantify in movement variability? and (ii) which tools are appropriate to quantify movement variability?

2.3.1 What to quantify in movement variability?

Complexity for this thesis refers to the dynamics of joint biomechanical degrees of freedom of a person performing a task in a certain environment (Davids et al., 2003). That said, Vaillancourt and Newell (2002, 2003) stated that there is no universal increase or decrease in complexity for movement variability as a function of age or disease but a dependency with the task dynamics. For example, in a constant-force task (where the task dynamics is of low dimension), older adults present less complexity due their inability to introduce additional degrees of freedom in the neuromuscular system. However, when the task dynamic is oscillatory, older adults or unhealthy adults (having intrinsic low dimension dynamics of their resting state) present an increase of complexity

because these adults have more difficulty to reduce the dimension output to a lower dimension. In contrast, inspired by Tononi et al. (1998) who modelled complexity with the variables of complexity versus regularity of neural networks, Stergiou et al. (2006) proposed a model for optimal human movement variability with the variables of complexity versus predictability. The model of Stergiou et al. (2006) stated that higher values of complexity are associated with rich behaviour of movements, while lower values of complexity movements are associated with poor behaviours of movements. Hence, higher complexity of movements are characterised by chaotic systems, while lesser complexity of movement is characterised either as noisy systems or periodic systems (having either low or high amounts of predictability) (Stergiou et al., 2006).

Considering the works of Vaillancourt and Newell (2002, 2003), Tononi et al. (1998) and Stergiou et al. (2006), I assume that the quantification of movement variability can be based on the complexity and predictability of human movement.

2.3.2 Which methods of nonlinear analysis are appropriate to quantify movement variability?

Stergiou et al. (2006) proposed a model for movement variability which state that variables of complexity and predictability of a system can be used to characterise and quantify movement variability. With that in mind, this thesis has led me to understand other challenges such as (i) finding, (ii) understanding and (iii) applying the appropriate methods of nonlinear analysis that can measure such variables.

Pincus (1995, 1991) proposed Approximate Entropy (ApEn) to quantify regularity of time series. Then, Richman and Moorman (2000) due to self-matching found that the algorithm of ApEn could evoke the occurrence of ln(0) which made ApEn dependant on the available data for which Sample Entropy (SampEn) was proposed as an algorithm that does not consider self-matching. Hence, SampEn values are

independent of the length of time series and its algorithm is simpler than ApEn. Then, instead of using single statistics, Costa et al. (2002) proposed Multiscale Entropy (MSE) which computes SampEn of consecutive coarse-grained time series of the original time series defined by the scale factor. With MSE algorithm, (Costa et al., 2002) noted that pathology dynamics for time series of heartbeat intervals are associated with reduction of complexity. Therefore, Costa et al. 2002, p. 3 concluded that physiological complexity is associated with the adaptive capacity of the organism, disease states and aging which "may be defined by a sustained breakdown of long-term correlations and loss of information". Essentially, entropy measures (AppEn and SampEn), quantify regularity and complexity of time series (Preatoni et al., 2013). However, Goldberger (1996) mentioned that the increase of irregularity in time series is not synonymous of increase with physiological complexity. Similarly, an increase of ApEn or SampEn, "implying increase of irregularity and decrease in predictability" (Goldberger et al., 2002, p. 25), is not synonymous with an increase of dynamical complexity when analysing physiology signals (Costa et al., 2002). Hence, Goldberger et al. (2002) demonstrated that ApEn as a regularity statistic is not a direct index of physiological complexity where, for example, a randomised time series of an healthy heartbeat with multi-scale and complex patterns of variability show a higher value of ApEn even though the time series is less complex. Therefore, Goldberger et al. 2002, p. 24 concluded that the loss of physiological complexity can be "better assessed using other measures which can detect and quantify the presence of long-term correlations in non-stationary series." Hence, Costa et al. (2002); Goldberger et al. (2002); Vaillancourt and Newell (2002) concluded that ApEn and SampEn do not necessary show the right representation of what they intend to measure.

Therefore, considering the previous cons of ApEn, SampEn and MSE, Detrended Fluctuation Analysis (DFA) can tackle the problem of quantifying long-term correlations

of time series (Peng et al., 1995). DFA is based on analysing fractal features and is calculated as the root mean square fluctuation of an integrated and detrended time series and it is represented by a scaling exponent, α, which is an indicator for roughness of time series, e.g. "the larger the value of α, the smoother the time series (Peng et al., 1995, p. 83). However, DFA can result in a false conclusion for long-term correlations in the time series (Rangarajan and Ding, 2000, p. 5001), therefore DFA "can falsely classify certain type of time series as fractals" (Wijnants et al., 2009, p. 80). With that in mind, Wijnants et al. (2009) proposed the use of Recurrence Quantification Analysis (RQA) as a technique that does not present constraints with regards to length size, stationary or statistical distribution of the time series. Wijnants et al. (2009) also highlighted that SampEn index is computed over the sequential values of the time series, whereas Shannon entropy with RQA, RQAEn, is computed over the distribution of deterministic lines in the Recurrence Plots (RP) (Marwan, 2008; Trulla et al., 1996; Zbilut and Webber, 1992). Similarly, Rhea et al. (2011) highlighted that algorithms to compute entropy measures are different since ApEn and SampEn are approximations of the Kolmogorov-Sinai Entropy computing the likelihood that a template pattern repeats in the time series while RQAEn is derived from Shannon entropy and is computing number of line segments of varying length in the RP. Even with those differences in the algorithms, smaller values of recurrence percentage of the RQA show the increase on practice of movement dynamics, concluding that such recurrence percentage indicate an increase of system stability (Wijnants et al., 2009).

Another method to measure variability is the largest Lyapunov exponent (LyE) which is used to "quantify the rate at which nearby trajectories converge or diverge" (Stergiou, 2016, p. 85). For instance, "LyE from a stable system with little to no divergence will be zero (e.g. sine wave)" and "LyE for an unstable system that has highest amount of divergence will be positive and relative high in value (e.g. 0.469 for

random noise)" and for chaotic systems like the Lorenz system, LyE is in between the two of the previous extremes (LyE$\approx$ 0.1) (Miller et al., 2006, p. 2874).

Measuring human movement variability requires a combination of the pros and cons of the previous methods that analyse either (i) the dynamic complexity or (ii) the degree of regularly, stability or predicability in a system (Goldberger et al., 2002; Harbourne and Stergiou, 2009; Stergiou and Decker, 2011). For instance, Rangarajan and Ding (2000) stated the use of both spectral analysis and random walk analysis, the base of DFA, is a better approach than only using one method because, for instance, using only DFA can lead to false conclusion for long-term correlations in the time series. Similarly, Wijnants et al. (2009) selected different methods (e.g. spectral analysis, standard dispersion analysis, DFA, RQA and SampEn) to quantify movement variability that can complement the strengths of some of them and compensate the weakness of others. Recently, Caballero et al. (2014) proposed the unification of different methods of nonlinear analysis to address every aspect of the dynamics of a systems and the characterisation of movement variability. Although, there is no best method to measure movement variability and an unification of methods to quantify human movement variability is still an open question (Caballero et al., 2014), finding the appropriate method of nonlinear analysis to measure movement variability for a specific problem, and knowing its strengths and weakness of such appropriate method is one of the research questions for this thesis.

2.4 Nonlinear analysis with real-world data

Recently, Huffaker et al. (2017) pointed out that one of the caveats when applying methods of nonlinear time series analysis is its unreliability when the estimated metrics come from real-world data which are generally short, noisy and non-stationary. Similarly, Preatoni et al. (2013) mentioned the limitations of methods of nonlinear analysis in

sport activities where data required to be large (e.g. number of trials, duration of the experiment and sampling frequency). Caballero et al. (2014) argued that there are weaknesses of different methods of nonlinear analysis regarding the characteristics of the time series such as non-stationarity, length data size, noise, sampling rate. However, in the work of Huffaker et al. (2017), Preatoni et al. (2013) and Caballero et al. (2014) no further exploration of the metrics of nonlinear analysis with real-world data is presented.

2.4.1 Non-stationarity

Non-stationarity of time series signals might create spurious increase or decrease in methods of nonlinear analysis. For instance, Costa et al. (2007) noted that non-stationarity in the signals might alter the increase of irregularity of signals for the shortest scales when applying MSE. Also, Dingwell and Cusumano (2000) reported non-stationarity in time series when using LyE, where LyE requires to be validated using surrogation (Dingwell and Cusumano, 2000; Miller et al., 2006) to ensure the robustness of the metric. Caballero et al. (2014) reported three options when dealing with non-stationary data: (i) remove non-stationary data, (ii) use empirical mode decomposition (EMD), or (iii) apply nonlinear tools, such as DFT and RQA, which are less sensitive to non-stationary data.

To remove non-stationary data, Carroll and Freedman (1993) suggested to remove the trends or to eliminate the initial data (e.g. first 20 seconds of samples) to ignore the trend of time series. For instance, van Dieën et al. (2010), in experiments with center of pressure movements in seated balancing, discarded the first 5 seconds of the time series in the start of the measurement.

Non-stationary time series can also be treated with Empirical Mode Decomposition (EMD) method which decompose nonlinear, non-stationary signals into their intrinsic

frequency components (Huang et al., 1998; Wu and Huang, 2004, 2009). Hence, Costa et al. (2007); Flandrin et al. (2004) tested whether EMD is a robust method for detrending and denoising time series and noted that EMD does not require selection of input parameters. However, the reliability of EMD methods is still an open problem. For instance, an extension of EMD called Multivariate Empirical Mode Decomposition (MEMD) has been proposed to analyse multiple time series (Mandic et al., 2013; Rehman and Mandic, 2010). See (Bonnet et al., 2014; Costa et al., 2007; Daubechies et al., 2011; Mert and Akan, 2018; Wu and Hu, 2006) for applications of EMD.

Finally, one can use methods of nonlinear analysis that are unaffected by non-stationarity of time series such as Detrended Fluctuation Analys (DFA) (Hausdorff et al., 1995) and Recurrence Quantification Analysis (RQA) (Marwan, 2008; Trulla et al., 1996; Zbilut and Webber, 1992). However, Bryce and Sprague (2012) reported negatives of DFA such as the introduction of uncontrolled bias, computational expensiveness and highlighted that DFA cannot provide a generic protection against the non-stationarities of the signals. The implication of this review is that RQA remains a promising approach.

2.4.2 Data length

Many of the methods of nonlinear analysis are sensitive to the length of time series (Caballero et al., 2014). For example, given that Multiscale Entropy (MSE) is a statistical measure, the data lengths when using MSE are recommended to be large (e.g. up to the scale of 6×10^3 data points) to ensure enough samples for the analysis (Costa et al., 2007). Also, the methods of LyE (Wolf et al., 1985), DFA (Peng et al., 1995), SampEn (Rhea et al., 2011) and ApEn (Richman and Moorman, 2000) require a minimum of data length whereas FuzzyEn (Chen et al., 2007) is more robust for data length. However, the methods of RQA (Riley et al., 1999; Webber and Zbilut, 1994;

Wijnants et al., 2009) and Permutation Entropy (Zunino et al., 2009) are less sensitive to the length of time series.

2.4.3 Sampling rate

One possible solution for the sensitivity of nonlinear methods to data length is the increase or decrease of sampling rate (Caballero et al., 2014). However, Duarte and Sternad 2008, p. 267 stated "the increase of sampling rate frequency would only increase artificially the data points without adding information" which raises the problem of oversamping signals. Then, Rhea et al. (2011) investigated the influence of sampling rate in three entropy measures (ApEn, SampEn and RQAEn) concluding that ApEn and RQAEn were robust across to the increase of sampling frequency, while SampEn presented significant difference across all sampling frequencies. Rhea et al. (2011) noted that SampEn is more sensitive to coliniarities than ApEn and RQAEn at higher frequencies which lead to a decrease of SampEn. Rhea et al. (2011) then concluded that signals at higher frequencies appear to be more regular due to the increase of data, therefore producing erroneous entropy results. Caballero et al. (2013) stated that for short length time series for SampEn and DFA, the decrease of sampling rate frequency is recommended because it presents less consumption of computational power. Additionally, Caballero et al. (2013) showed the robustness of the methods of SampEn and DFA when using different sampling rate frequencies, stating that frequencies near the dynamics of the activity create a more reliable analysis of the dynamics.

2.4.4 Noise

Caballero et al. (2014) reviewed methods of nonlinear analysis that are affected by noise. Rosenstein et al. (1993), for instance, tested the robustness of LyE against three levels of noise (lowest, moderate and highest) in order to note the unreliability of LyE

exponents in high-noise environments. However, such case of unreliability of LyE is unreal as the reported values of signal-to-noise ratios are substantially lower than those used at the experiments of Rosenstein et al. (1993). Bandt and Pompe (2002) proposed Permutation Entropy (PeEn) which is an appropriate method for chaotic time series in the present of observational and dynamical noise. Another example is the work of Chen et al. (2009) who compared the robustness of FuzzEn, ApEn and SampEn metrics against different levels of noise, concluding that for a large value of the parameter r of ApEn and SampEn, these two metrics can work well with high levels of noise, however when noise increases, ApEn and SampEn fail to distinguish time series with different levels of noise, whereas FuzzEn is robust to such highest levels of noise.

Regardless of the source of noise which can either be mechanical (due to recording equipment) or physiological (due to different neural noise), Rhea et al. (2011) highlighted the importance of the effects on noise in three entropy measures (ApEn, SampEn and RQAEn) which produce different results. Values for AnEn and SampEn, for instance, tended to increase as noise was added to the signals, while RQAEn showed an inverse effect, e.g. RQAEn values decreased as noise in the signal was increased. Similar results for synthetic data were also reported by Pellecchia and Shockley (2015) where RQAEn values decreased from ($RQAEn \approx 5$) for Lorenz system to a ($RQAEn \approx 2$) for a periodic signal with a further decrease ($RQAEn \approx 0.3$) for a sinusoid signal with superimposed noise. Therefore, RQA can be affected by noise (Rhea et al., 2011). However, the effects of noise and non-stationarity can be mitigated with the selection of the right parameters to perform RQA, particularly, using embedding dimensions from 10 to 20 (Webber and Zbilut, 2005).

Another solution for noisy time series is the use of traditional filtering methods. However, the attenuation of all frequencies of the signal along the with the noise, given a cutoff frequency, can cut out information that might be useful for nonlinear

time-series. Another option is apply DFA, which additionally to the remove of local trends, it also reduces the noise of the signal (Hausdorff et al., 1995). Alternatively, filtering strategies for nonlinear time-series data can be applied which tailor in a more effective way the properties of nonlinear dynamics (see Bradley and Kantz 2015 and references therein).

2.5 Final remarks

In this chapter, literature has been reviewed based on the questions of: (i) what to quantify in movement variability, (ii) which non-linear tools are appropriate to quantify movement variability, and (iii) what are the strengths and weaknesses of nonlinear analysis methods with real-world data. It can then be concluded that little research has been done on the effects with Reconstructed State Spaces (RSSs), Recurrent Plots (RPs), and Recurrence Quantification Analysis (RQA) metrics for different embedding parameters, different recurrence thresholds and different characteristics of time series (window length size, smoothness and structure). That said, nonlinear analysis methods such as estimation of embedding parameters, RSSs, RPs, and RQAs are reviewed in the following chapter.