Influence of embedding parameters and noise in center of pressure recurrence quantification analysis

Christopher J. Hasson [a,*], Richard E.A. Van Emmerik [a], Graham E. Caldwell [a],
Jeffrey M. Haddad [a,1], Jeff L. Gagnon [b], Joseph Hamill [a]

[a] *Biomechanics and Motor Control Laboratories, Department of Kinesiology, University of Massachusetts Amherst, Amherst, MA 01003, United States*
[b] *Math, Physics, and Computer Science Department, Springfield College, Springfield, MA 01109, United States*

Abstract

Recurrence quantification analysis (RQA) can extract the dynamics of postural control from center of pressure (CoP) data by quantifying the system's repeatability, complexity, and local dynamic stability through several variables. Computation of these variables requires the selection of suitable embedding parameters for state space reconstruction (i.e. time delay and embedding dimension); however, it is unclear how the parameters influence RQA variables when examining noisy CoP data. This study evaluated the sensitivity of RQA variables to embedding parameter values and noise level, and assessed methods of selecting embedding parameters for CoP data. Five healthy male subjects maintained quiet stance for 30 s while the anterior–posterior CoP was measured. The effect of noise was evaluated by adding uniform white noise of increasing amplitude to the raw CoP signal. The magnitude of all RQA variables decreased with increasing noise amplitude for all subjects. A sensitivity analysis was performed by systematically altering the embedding parameters for the raw data with and without a selected level of added noise. The key result was that, for all subjects, the RQA variables were sensitive to the embedding parameter values and the level of noise in the CoP data. Finally, the performance of false nearest neighbors and average displacement algorithms for choosing embedding parameters was evaluated. Both methods gave clear and consistent results for all subjects with either raw or noisy data. The results suggest that careful selection of embedding parameters is essential when using RQA to examine postural control based on noisy CoP data.
© 2007 Elsevier B.V. All rights reserved.

Keywords: Recurrence quantification analysis; Postural control; Center of pressure; State space reconstruction; Embedding parameters

Information about the state of the postural control system can be gained from the ground reaction force center of pressure (CoP). Traditional analysis methods often utilize scalar measures of stability, such as the mean and variance of different CoP measures (e.g. path length, velocity, range, etc.). These measures offer somewhat limited information about the postural control system, particularly when the non-

stationarity of the CoP time series is considered, as the mean and variance of the signal have been shown to change over time [1].

Newer techniques, based on nonlinear dynamical systems models, provide information regarding the patterns and structure of postural fluctuations. These methods do not assume stationarity, and are increasingly being applied to assess changes in the CoP signal. One such technique, recurrence quantification analysis (RQA) [2], evaluates the structure present in recurrence plots (Fig. 1), which are visual representations of the recurrent patterns present in time series data [3]. RQA quantifies the repeatability, complexity, and local dynamic stability of dynamical systems. Various biophysical phenomena have been studied with RQA, including the dynamics of postural control

* Corresponding author at: Department of Kinesiology, 110 Totman Building, University of Massachusetts Amherst, 30 Eastman Lane, Amherst, MA 01003-9258, United States. Tel.: +1 413 545 4421; fax: +1 413 545 2906.

 E-mail address: cjhasson@kin.umass.edu (C.J. Hasson).

[1] Currently with the Motor Development Laboratory, Department of Health and Kinesiology, Purdue University, West Lafayette, IN 47907, United States.
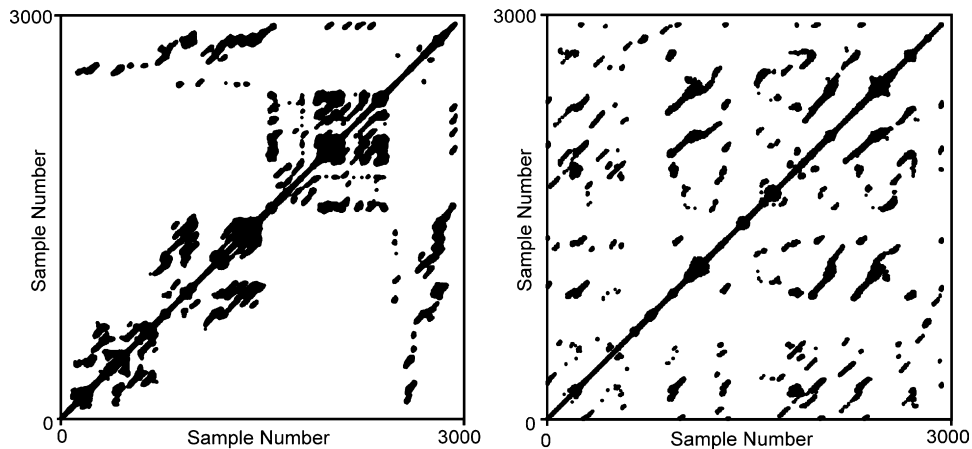
Fig. 1. Plots of the recurrence matrix generated for the anterior–posterior center of pressure time series in the hard-surface-eyes-open quiet stance condition for the older subject (left) and a young subject (right). Darkened areas represent recurrent points.

reflected in the CoP [4–8]. For example, CoP complexity increases during quiet stance with eyes closed compared to eyes open [4].

RQA requires reconstruction of a dynamical system's state space from a single scalar time series through the process of embedding [9,10]. A common method of embedding is to use time-delayed versions of the original time series to reconstruct the state space, which entails selection of a time delay (τ) and embedding dimension (m). The mathematics behind this technique were detailed by Takens [9]. However, this work was based on noise-free data, where practically any τ and m could be chosen if they did not violate the embedding theorem. In the presence of noise, the choice of embedding parameters becomes important.

If τ is too small the coordinates of the vectors in the reconstructed state space will be almost identical, with the trajectory flattened along the main diagonal, which yields little information about the system (i.e. "redundance") [11]. However, if τ is too large the delay vectors become "causally disconnected" in time (i.e. "irrelevance") [11]. If m is too low the trajectory will not be completely unfolded and therefore will not faithfully represent the dynamics of the system. If m is increased too much, noise will begin to dominate the embedded space and computation cost may become excessive [12]. These issues are important because proper state space reconstruction is a fundamental requirement for meaningful RQA results.

Studies employing RQA to analyze CoP records have used different methods for embedding parameter selection. Riley, Balasubramaniam, and Turvey [4] used a combination of techniques for determining τ and m, including autocorrelation and average mutual information (for τ), as well as random shuffling of the data and sensitivity analyses (for τ and m); however, the rationale behind the final parameter selection was not clear. This study has since been cited as a basis for RQA embedding parameter selection by others [6–8].

CoP time series contain small-amplitude fluctuations that may reflect subtle control of the center of mass during quiet

stance, which could be masked by the presence of noise. None of the CoP time series examined in the literature were reported as filtered [4–8], even though some are qualitatively quite noisy (e.g. Fig. 1 in [5]). It has been demonstrated analytically that even a small amount of noise has the potential to decrease the reliability of RQA calculations [13]. However, the combined effects of noise and embedding parameter values on RQA CoP outcome variables are unknown.

The primary purpose of this study was to evaluate the sensitivity of the RQA procedure to the choice of embedding parameters and noise amplitude when reconstructing postural control dynamics from the CoP during quiet stance. A secondary purpose was to evaluate and recommend methods for selecting embedding parameters that give clear and consistent results for quiet stance CoP data.

1. Methods

1.1. Experiment details

During quiet stance, ground reaction forces and moments were measured from four healthy, young male subjects (age: 30 ± 4.8 years, mass: 82.9 ± 9.0 kg, height: 177 ± 5.6 cm; mean $\pm$ S.D.) and a healthy, older male subject (52 years, 86.5 kg, 194 cm) using a force platform (AMTI, Watertown, MA). Data were amplified (gain = 4000, anti-alias filtered at 1000 Hz cutoff), and sampled with a 16-bit analog-to-digital converter at 100 Hz. The position of the CoP in the anterior–posterior direction was calculated from these force data.

The subjects stood barefoot with arms hanging at their sides. The feet were positioned hip-width apart, and turned outwards $10°$ from the sagittal plane (with the lateral malleoli and anterior–superior iliac spines in the same sagittal plane). Data were collected for 30 s for comparison with other studies [4–8]. All subjects performed the postural experiment on the bare force platform (hard surface) with the eyes open. The older subject also performed quiet stance with the eyes closed and on a foam cushion (soft surface).

1.2. RQA procedure

All calculations were performed in MATLAB® (Mathworks, Inc., Natick, MA). Details regarding the RQA procedure, parameters, and variables can be found in Webber and Zbilut [14]. Briefly, the CoP data were embedded in multiple dimensions using m copies of the original time series, with each copy shifted in time by integer multiples of τ samples. A distance matrix was created by determining the Euclidean distances between all embedded vectors. The distance matrix was scaled by dividing all elements by the mean distance. A threshold (radius) of 20% of the mean distance was applied, where all cells in the scaled distance matrix with values below this threshold were identified as recurrent points to create a recurrence matrix (Fig. 1).

Several variables were used to quantify the structure present in the recurrence matrix. The percent recurrence (%REC) signifies how often a trajectory visits similar locations in state space (time-independent), computed as the percentage of recurrent points in the recurrence matrix. The percent determinism (%DET) relates to how often the trajectory repeatedly re-visits similar state space locations (time-dependent), quantified as the percentage of recurrent points in diagonal line structures (at least three consecutive points in length) parallel to the main diagonal. The ratio %DET/%REC was also computed, as it may be related to changes in physiological states [15]. Entropy (ENT), a measure of system complexity, was quantified by the distribution of the lengths of diagonal line segments parallel to the main diagonal. TREND, a measure of drift and non-stationarity, was calculated from the slope of the line-of-best-fit between the percentage of recurrent points in each diagonal line and the distance of the points from the main diagonal. MAX LINE, a measure of dynamic stability, is the length of the longest diagonal line in the recurrence plot and is inversely proportional to the magnitude of the largest Lyapunov exponent [3].

1.3. Effect of noise on RQA variables

To assess the effects of different amplitudes of noise on RQA variables, increasing amounts of uniformly distributed white noise were added to the raw CoP data for each subject (Fig. 2A and B). White noise was used to emulate the background noise level from the electronics used to collect the data (e.g. strain gauges, amplifiers, analog-to-digital conversion process), which is present in most CoP data. The magnitude of the added noise was equal to the smallest difference between points in the CoP time series, multiplied by a scaling factor (μ), which ranged from 0 to 200. For each noise scaling level, a new white noise time series was generated and added to the raw data. The embedding parameters were computed on an individual basis using the raw data sets, and then were kept constant for each of the noisy data sets. The raw and noisy data for each subject were then evaluated using the RQA procedure.

1.4. Embedding parameter sensitivity analysis

RQA was performed on the raw CoP data from each subject using different combinations of embedding parameters. This process was repeated with noise added ($\mu = 120$) to approximate the
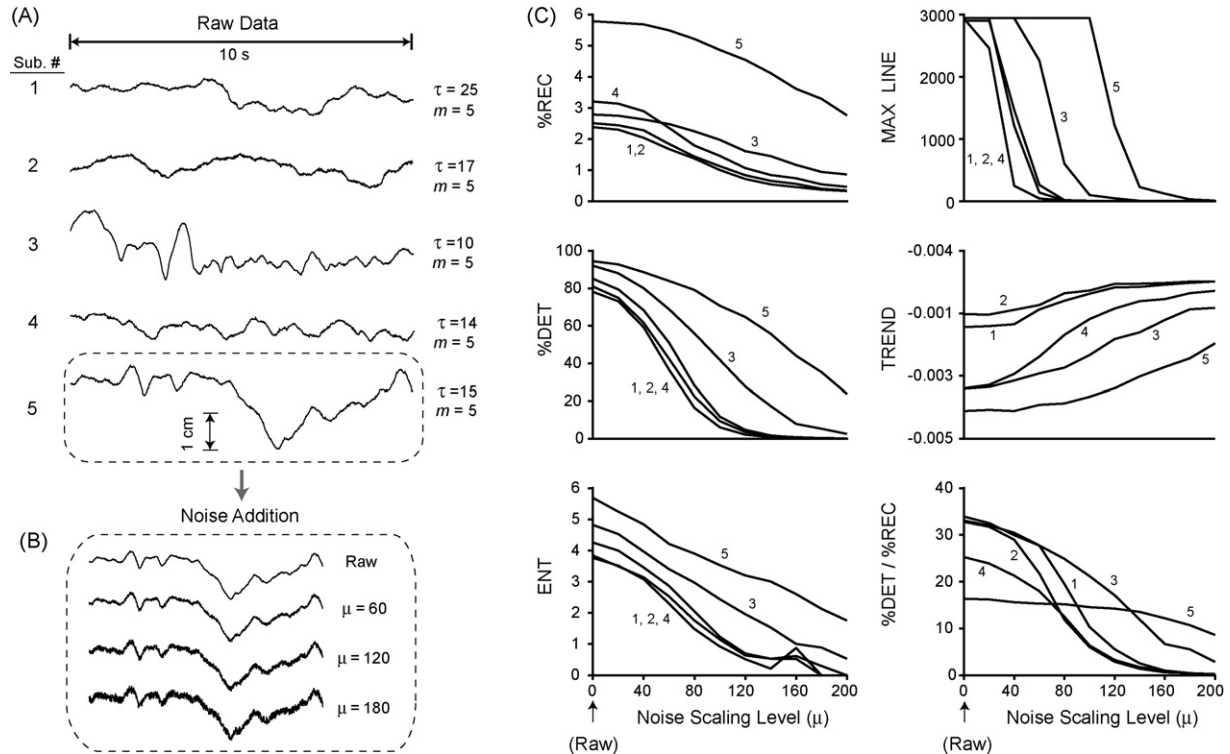

Fig. 2. (A) Raw anterior–posterior center of pressure time series measured during quiet stance on a hard surface with the eyes open for the four young subjects (1–4) and an older subject (5). Note that only the last 10 s of each time series is shown for clarity. The time delay (τ) and embedding dimension (m) computed for each of the time series is indicated. (B) For the older subject, the raw data with and without selected levels of added white noise are shown (μ = noise scaling factor). (C) Changes in recurrence quantification analysis variables with the addition of increasing amounts of noise to the raw time series shown in A (numbers beside data reference individual subjects).

noise levels of CoP time series reported in the literature [4–8]. Values for τ were varied from 1 to 30 samples (0.01–0.3 s) and m from 2 to 20.

1.5. Embedding parameter selection

Similar to Riley et al. [4] we found that autocorrelation and mutual information techniques for choosing τ values were inconsistent, and sometimes gave exceedingly high values (τ ranged from 59 to 831 samples). Therefore, we used the average displacement (AVD) method, which was developed specifically for noisy data sets by Rosenstein, Collins, and De Luca [16]. This method quantifies the expansion of the reconstructed trajectory from the line of identity as a function of τ, and attempts to strike a balance between error associated with redundance and irrelevance. A value for τ was selected where the slope of the AVD curve decreased by 40% of the initial slope, as recommended by Rosenstein et al. [16].

A false nearest neighbors (FNN) algorithm was used to select an appropriate m, with R_{tol} and A_{tol} set to 15 and 2, respectively (see [12] for details). Both AVD and FNN computations were performed over a range of different embedding parameters (Figs. 4 and 5).

2. Results and discussion

2.1. Effect of noise on RQA variables

For all subjects, the magnitude of all RQA variables decreased with increasing noise level (Fig. 2C), supporting the hypothesis that RQA variables are sensitive to the amount of noise present in CoP time series. The effect of noise on %DET values was comparable to those reported by Thiel and colleagues for experimental laser data [13]. These
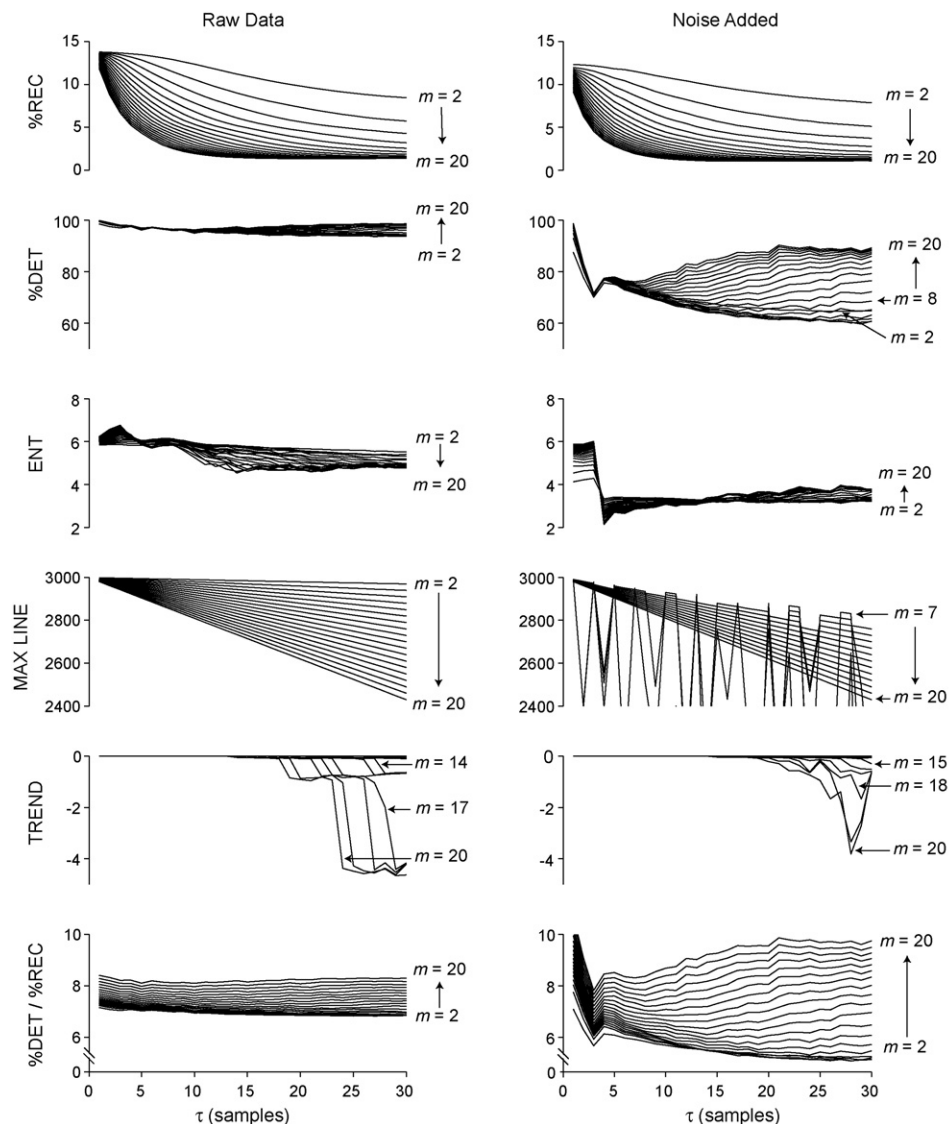

Fig. 3. Sensitivity analysis showing changes in recurrence quantification analysis variables as a function of time delay (τ) and embedding dimension (m) for the anterior–posterior center of pressure measured during quiet stance on a hard surface with the eyes open for the older subject. The left column displays the results from the analysis of the raw data; the right column shows the results from the same data set with white noise added at a scaling level of $\mu = 120$. Similar results were found for all subjects.

results stress the importance of understanding the noise characteristics of sampled CoP data when using RQA, as different conclusions may be reached based on differences in background noise level rather than physiological changes. The effect of noise is variable-specific; note that noise amplitude had a discontinuous effect on MAX LINE, which decreased rapidly, as only a single break at the middle of the longest line could reduce the MAX LINE value by half.

It should also be noted that the constant radius parameter used in the present study was chosen based on methods used in the literature [4]. Thiel and colleagues [13] demonstrated that the disruption of recurrent structures caused by measurement noise could be minimized by choosing an optimal radius. However, if noise levels are too high (>20% of the standard deviation of the underlying process), the RQA variables may be affected even with an optimal radius [13].

2.2. Embedding parameter sensitivity analysis

The results of the sensitivity analysis were similar across subjects, and across the different quiet stance conditions examined in the older subject. Consequently, results for a representative subject will be examined (Fig. 3). In both the raw and noisy data, %REC decreased with increasing m and τ. This is because as m increases the trajectory is unfolded and points become farther apart in space, and as τ increases successive state space vectors become farther apart in time.

In contrast, %DET was much more sensitive to the selection of embedding parameters in the presence of added noise. For example, at $\tau = 15$ the range of %DET values for different embedding dimensions is only a few percent for the raw data; however, the range is ~20% for the noisy data. With noise, values for ENT shift abruptly at $\tau > 3$. The MAX LINE variable changed linearly with changes in embedding parameters in the raw data, due to the number of samples included in the reconstruction, defined by

$$M = N - [\tau(m - 1)]$$

where M is the number of reconstructed state space vectors and N is the number of samples in the time series ($N = 3000$). However, with a noisy signal, MAX LINE was very erratic when $m < 11$. In both the raw and noisy data sets, TREND was highly variable when $\tau > 16$ and $m > 14$. Such high volatility suggests that caution is needed when interpreting MAX LINE and TREND values. Although similar trends were found from the sensitivity analysis of the data from all subjects, the %DET values decreased more rapidly as τ was increased for the younger subjects.

2.3. Embedding parameter selection

Because of the increased sensitivity of RQA parameters when noise is present, selection of embedding parameters is important. The FNN analysis for the selection of m yielded consistent results for both raw and noisy data for all subjects
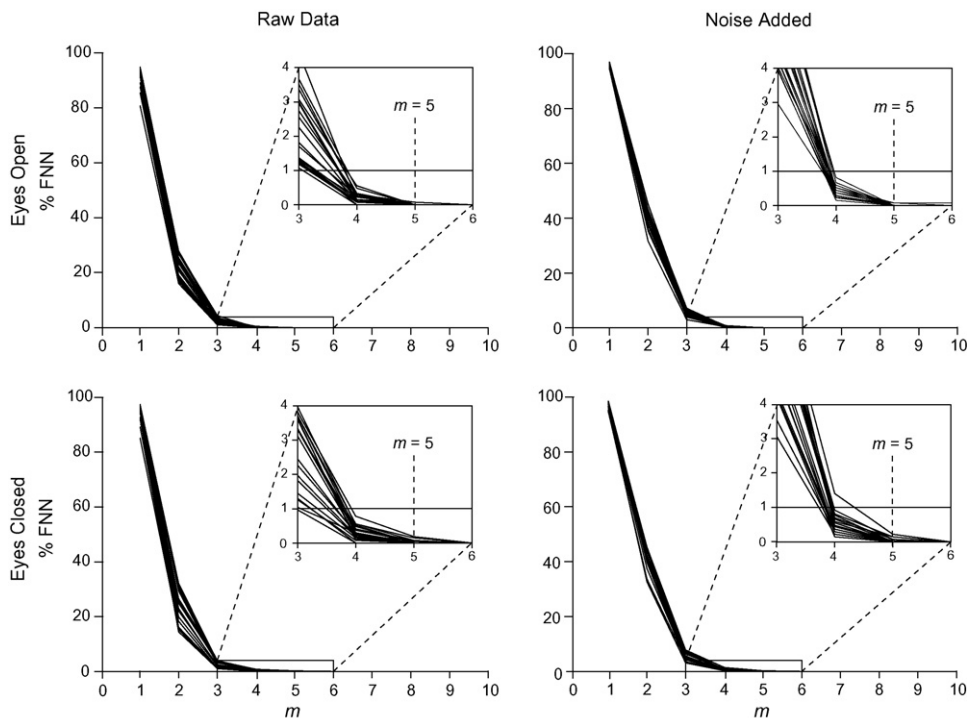

Fig. 4. False nearest neighbors (FNN) analysis performed using different combinations of time delays and embedding dimensions (m). Each line represents a unique value for the time delay, which ranged from 1 to 20 samples. Results are shown for the raw data and raw data with added noise (noise scaling level $\mu = 120$) for the older subject under different visual conditions. Values for m that give less than 1% FNN are considered acceptable; a value of $m = 5$ was chosen for this data set (see insets). Similar results were found for all subjects.
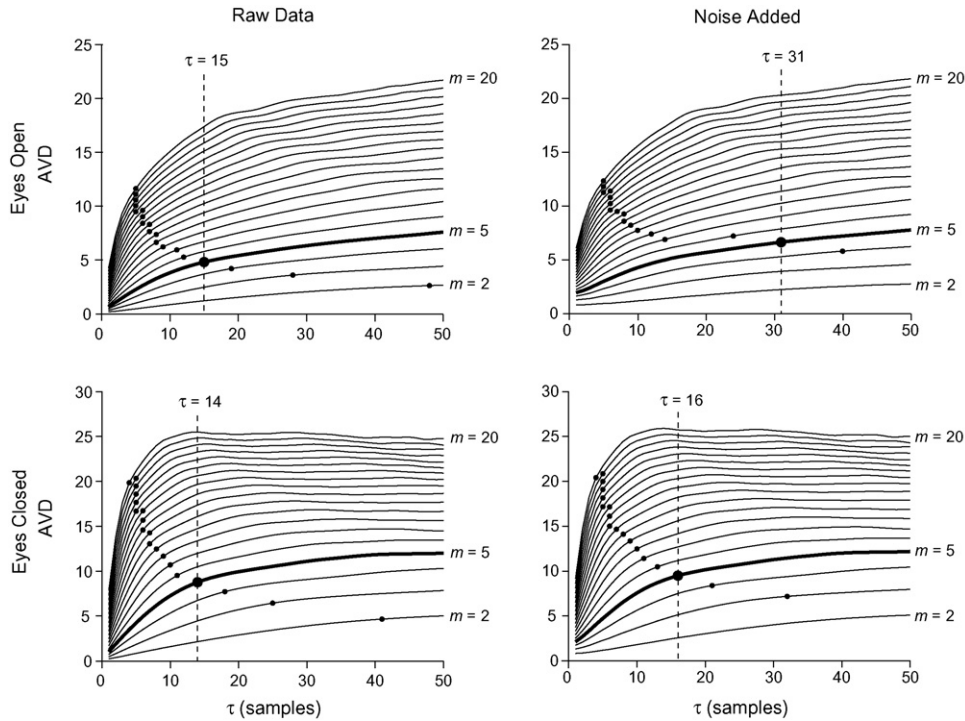
Fig. 5. Results of the average displacement (AVD) analysis using different time delays (τ) and embedding dimensions (m) for the older subject under different quiet stance conditions. Based on a false nearest neighbors analysis, an m of 5 was chosen (thick line). The time delays recommended for $m = 5$ are indicated by vertical dashed lines, which represent a 40% decrease from the initial slope. Similar results were found for all subjects.

and conditions (Fig. 4). Kennel, Brown, and Abarbanel [12] suggest that any m yielding a FNN percentage below 1% is acceptable. In our data set, an m of 4 met the 1% criterion for most values of τ; we chose $m = 5$ to be conservative.

For the raw data of the older subject, the AVD analysis yielded consistent results for selecting τ across all conditions (Fig. 5). Adding noise increased values for τ, particularly for the eyes-open condition (Table 1). In the eyes-closed condition, τ only increased slightly with the addition of noise, while values for the eyes-open condition increased more. This may be due to the differing CoP frequency content in the eyes-closed and eyes-open conditions. Median CoP frequencies were higher for eyes-closed than eyes-open. Because the added noise scaling was based on the smallest difference between two points in each time series, the relative amount of noise was lower for the eyes-closed condition (with its higher frequency CoP movements).

Table 1
Results for average displacement analysis for choosing time delay (τ) for the older subject

Variable	Eyes open		Eyes closed	
	Hard surface	Soft surface	Hard surface	Soft surface
τ (samples)				
Raw data	15	15	14	13
with added noise	31	24	16	14
Median frequency (Hz)[a]	0.03	0.03	0.30	0.53

[a] Median frequency values were the same for the raw and noisy data.

Therefore, τ should be chosen with caution under high noise conditions. Values computed for τ using the AVD method for the raw data from the younger subjects ranged from 10 to 25 samples (Fig. 2A).

3. Conclusions

When using RQA to quantify the dynamic structure present in quiet stance from a noisy CoP time series, the embedding parameters should be selected carefully because of the demonstrated greater sensitivity of RQA variables to noise. This may be due to the folding back of high-dimensional noise into the lower-dimensional embedded space during the reconstruction process, causing changes in the structure of the trajectory. As the reconstructed state space forms the basis of RQA, this distortion may be responsible for the increased sensitivity of RQA variables to noise. The measured CoP time series will always contain a certain amount of noise; therefore, researchers who study postural control should be aware of the effect of noise when computing and interpreting RQA variables. We also found that the FNN and AVD techniques for selecting embedding parameters gave clear and consistent results for the subjects and conditions investigated in this study. The AVD method may be a preferred technique for choosing time delays for noisy CoP data; however, the sensitivity of the AVD method to non-stationarity is currently unknown.

References

[1] Carroll JP, Freedman W. Nonstationarity properties of postural sway. J Biomech 1993;26:409–16.

[2] Zbilut JP, Webber Jr CL. Embeddings and delays as derived from quantification of recurrence plots. Phys Lett A 1992;171:199–203.

[3] Eckmann JP, Kamphorst S, Ruelle D. Recurrence plots of dynamical systems. Europhys Lett 1987;4:973–7.

[4] Riley MA, Balasubramaniam R, Turvey MT. Recurrence quantification analysis of postural fluctuations. Gait Posture 1999;9:65–78.

[5] Balasubramaniam R, Riley MA, Turvey MT. Specificity of postural sway to the demands of a precision task. Gait Posture 2000;11:12–24.

[6] Schmit JM, Regis DI, Riley MA. Dynamic patterns of postural sway in ballet dancers and track athletes. Exp Brain Res 2005;163:370–8.

[7] Schmit JM, Riley MA, Dalvi A, Sahay A, Shear PK, Shockley KD, et al. Deterministic center of pressure patterns characterize postural instability in Parkinson's disease. Exp Brain Res 2006;168: 357–67.

[8] Riley MA, Clark S. Recurrence analysis of human postural sway during the sensory organization test. Neurosci Lett 2003;342: 45–8.

[9] Takens F. Detecting strange attractors in fluid turbulence. In: Rand D, Young LS, editors. Symposium on Dynamical Systems and Turbulence, vol. 898 of Lecture Notes in Mathematics. Berlin: Springer-Verlag; 1981. p. 366.

[10] Packard NH, Crutchfield JP, Farmer JD, Shaw RS. Geometry from a time series. Phys Rev Lett 1980;45:712–6.

[11] Casdagli M, Eubank S, Farmer JD, Gibson J. State space reconstruction in the presence of noise. Physica D 1991;51:52–98.

[12] Kennel MB, Brown R, Abarbanel HDI. Determining embedding dimension for phase-space reconstruction using a geometrical construction. Phys Rev A 1992;45:3403–11.

[13] Thiel M, Romano MC, Kurths J, Meucci R, Allaria E, Arecchi FT. Influence of observational noise on the recurrence quantification analysis. Physica D 2002;171:138–52.

[14] Webber Jr CL, Zbilut JP. Dynamical assessment of physiological systems and states using recurrence plot strategies. J Appl Physiol 1994;76:965–73.

[15] Webber Jr CL. Rhythmogenesis of deterministic breathing patterns. In: Haken H, Koepchen HP, editors. Rhythms in Physiological Systems. Berlin: Springer-Verlag; 1991. p. 171–91.

[16] Rosenstein MT, Collins JJ, De Luca CJ. Reconstruction expansion as a geometry-based framework for choosing proper delay times. Physica D 1994;73:82–98.