1-2018

Introducing Statistical Persistence Decay – A Quantification of Stride-to-Stride Time Interval Dependency in Human Gait

P. C. Raffalt

J. M. Yentes
University of Nebraska at Omaha, jyentes@unomaha.edu

Recommended Citation
Raffalt, P.C. & Yentes, J.M. Ann Biomed Eng (2018) 46: 60. https://doi.org/10.1007/s10439-017-1934-1

ORIGINAL RESEARCH ARTICLE

Title:

Introducing Statistical Persistence Decay – A Quantification of Stride-to-Stride Time Interval Dependency in Human Gait

Authors:

Raffalt, PC.[1,2] and Yentes, JM.[3]

Affiliations:

1) Julius Wolff Institute for Biomechanics and Musculoskeletal Regeneration, Charité – Universitätsmedizin Berlin, Berlin, Germany.

2) Department of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark.

3) Center for Research in Human Movement Variability, Department of Biomechanics, University of Nebraska Omaha, Omaha, NE, USA.

Corresponding Author:

Jennifer M. Yentes

Center for Research in Human Movement Variability

University of Nebraska Omaha

Email: jyentes@gmail.com

Office phone: 402 554 3251

Abstract

Stride-to-stride time intervals during human walking is characterised by predictability and statistical persistence quantified by sample entropy (SaEn) and detrended fluctuation analysis (DFA) which indicates a time dependency in the gait pattern. However, neither analyses quantify time dependency in a physical or physiological interpretable time scale. Recently, entropic half-life (ENT½) has been introduced as a measure of the time dependency on an interpretable time scale. A novel measure of time dependency, based on DFA, statistical persistence decay (SPD), was introduced. The present study applied SaEn, DFA, ENT½, and SPD in known theoretical signals (periodic, chaotic, and random) and stride-to-stride time intervals during overground and treadmill walking in healthy subjects. The analyses confirmed known properties of the theoretical signals. There was a significant lower predictability ($p=0.033$) and lower statistical persistence ($p=0.012$) during treadmill walking compared to overground walking. No significant difference was observed for ENT½ and SPD between walking condition, and they were exhibited a low correlation. ENT½ showed that predictability in stride time intervals was halved after 11–14 strides and SPD indicated that the statistical persistency was deteriorated to uncorrelated noise after ~50 strides. This indicated a substantial time memory, where information from previous strides affected the future strides.

Keywords: walking, dynamics, nonlinear behaviour, entropy, DFA, stride time fluctuations

Introduction

The stride time pattern during continuous walking in healthy individuals has been shown to include stride-to-stride fluctuations exhibiting statistical persistence[6,14]. Thus, each stride depends on many previous strides with a stride-to-stride dependency that "*decay in a scale-free (fractal-like), power-law fashion*"[12]. Equally, stride length fluctuations during both treadmill and overground walking have been shown to exhibiting statistical persistence[6,32]. While stride speed fluctuations during treadmill walking have been shown to exhibit statistical anti-persistence[6], statistical persistence has been observed during overground walking[32].

Reduced stride-to-stride persistence has been interpreted differently in relation to the function of the underlying motor control system. Loss of statistical persistence in stride time fluctuations has been observed in older adults[13,16], different neurological patients[11,13], and has been suggested to reflect a degraded motor control function[10]. However, reduced persistence observed in frail individuals has been suggested to potentially implicate an increased control effort to achieve a more cautious gait pattern[6].

These contradicting observations call for caution when linking loss of stride-to-stride persistence to either an impaired motor control function or enhanced motor control effort[6]. The aforementioned studies have applied detrended fluctuation analysis (DFA) to assess the presence and strength of statistical persistence in the investigated time series. However, DFA does not quantify the time dependency on an interpretable physiological or physical time scale. This means that even though existence of statistical persistence can be confirmed by DFA, the stride-to-stride dependency cannot be quantified in terms of a specific number of previous strides that influences the current stride. Equally, DFA does not quantify for how long into the future in terms of seconds or minutes a completed stride will influence new strides.

In addition to DFA, sample entropy (SaEn) has been applied to both kinematic signals and stride time interval time series recorded during gait in order to quantify the predictability of the gait pattern[1,9,18]. SaEn is high in random white noise signals where no point-to-point dependency exists and the predictability is low. In contrast, in both chaotic and periodic signals, point-to-point dependency does exist and they are characterized by relative low SaEn indicating high predictability (e.g. demonstrated in [33]). However, the predictability is reported on a relative scale and cannot be translated into physical or physiological terms. Both DFA and SaEn quantify the time dependency of a time series but return outcome values not easily interpretable in relation to other physiological measurements (e.g. duration of muscle activity, latency in reflex measurements, reaction time).

Inspired by multiscale entropy (MSE) and to overcome the aforementioned methodological limitation, entropic half-life (ENT½) was proposed by Zandiyeh and Von Tscharner[34]. ENT½ estimates the time until the predictability in a time series is halved. This is also a measure of how long data points remain related to one another. Applied to movement related variables, ENT½ could quantify how long time elapses before previously performed movements have substantially reduced their influence on future movements[2].

While DFA, SaEn, and MSE previously have been applied to characterize the stride-to-stride time dependency in human gait, ENT½ has to the best of our knowledge not previously been used to quantify time dependency in stride-to-stride time intervals. When applied to stride-to-stride time interval time series, ENT½ will estimate how long (in terms of number of strides) it takes to deteriorate the predictability of the stride time intervals by 50 %. DFA could be used to quantify decay in time persistence of time series through an application similar in manner to ENT½.

In many gait experiments, the treadmill has been used instead of overground walking due to its advantages with respect to continuous data collection of motion capture, ground reaction forces, etc. The results and interpretation have often been extrapolated to overground walking even though substantial biomechanical differences have been observed between treadmill and overground walking[20,31]. The constraints of the constant speed and limited space of the treadmill have been suggested to induce a less persistent and more unpredictable walking pattern compared to overground walking[31].

The present study aimed at introducing two novel tools, ENT½ and SPD as methods to quantify time dependency in stride-to-stride time intervals during human gait. To validate the use of these methods, the present study included known theoretical signals (i.e., periodic, chaotic, and random) and stride time data recorded during overground and treadmill walking. With respect to the theoretical signals, it was hypothesized that the periodic signals would be characterized by high ENT½ and SPD, the random signal would be characterized by low ENT½ and SPD and the chaotic signals would be characterized by intermediate ENT½ and SPD.

Further, experimental data from one-hour of walking overground and on a treadmill were used to verify the use of the methods. Based on previous observations of less statistical persistency and more unpredictability during treadmill walking[31], we hypothesized that the time dependency during treadmill will be less pronounced indicated by lower ENT½ and SPD values compared to overground walking.

As a secondary aim, the association between statistical persistence and predictability were investigated. If changes in the statistical persistence in stride-to-stride time intervals observed during human walking[6,14] could explain changes in the predictability; it was hypothesized that the scaling exponent and SaEn in such time series would correlate. Furthermore, if changes in

the deterioration in statistical persistence in stride-to-stride time intervals could explain changes in the deterioration in the predictability, ENT½ and SPD values would also correlate. Such correlations could indicate that the underlying mechanisms for creating statistical persistence and predictability are regulated simultaneously.

Materials and Methods

Theoretical Procedures

To verify the interpretation of the outcome measure, twenty time series of four different mathematical signals of 2500 data points with different characteristics (periodic, chaotic, and random) were created using the *colored noise generator* function in Matlab (MathWorks R2011b). These signals included a brown noise signal (power spectrum of $1/f^2$) which was considered to be periodic, a signal derived from the second Lorenz differential equation ($\sigma = 10$, $\rho = 28$, and $\beta = 8/3$) which was considered to be chaotic, a pink noise signal (power spectrum of $1/f$) which also was considered to be chaotic, and a white Gaussian noise signal (constant power spectrum) which was considered to be random (figure 1). The periodic brown noise signal and the random Gaussian noise signal represented two extremes for each of the four applied analyses with the chaotic Lorenz attractor and pink noise as intermediate signals.

Subjects

Fourteen volunteers (seven males and seven females) with a mean (± SD) age of 25.0 years (± 4.2), height of 170.8 cm (± 11.9) and body mass of 69.4 kg (± 16.9) participated in the present study. The participants had no diagnosed lower limb injuries within the past year. They were informed of the experimental conditions and gave their written consent to participate in the study. The study was approved by the by the Institutional Review Board of the University of Nebraska Medical Center, and it was carried out in accordance with the approved guidelines.

Protocol

The study consisted of two experimental sessions. During the first session the subjects completed a one hour overground walking trial on an elliptical indoor track (circumference ~ 201m) at their self-selected walking speed. The walking speed was not recorded and was allowed to fluctuate.

At the second session, the subjects completed a one hour treadmill walking trial at a constant self-selected walking speed. The walking speeds were not registered. During both walking trials, footswitches (Trigno™ 4-channel FSR Sensor, Delsys Inc., Natick, MA) placed under both heels recorded heel strikes at a sampling rate of 148 Hz. No objective measurement of fatigue was obtained during the trials. However, none of the subjects reported fatigue to influence their gait.

Analysis

The right heel strike data series from each walking trial was processed in Matlab (MathWorks R2011b) in order to create stride time interval time series (figure 1 and supplementary material 1). Stride time was defined as the time from heel strike of one foot until the subsequent heel strike of the same foot. Each time series was cut to contain 2500 strides and were subjected to four different analyses: 1) SaEn, 2) ENT½, 3) DFA, and 4) SPD.

Sample entropy

SaEn was based on the algorithm by Richman and Moorman[27] (equation 1). SaEn was defined as the negative logarithm for conditional properties that a series of data points within a certain distance, m, would be repeated within the distance m+1[27].

Equation 1: $$SaEn(m,r,N) = -\ln[\frac{A^{m+1}(r)}{B^{m}(r)}]$$

Where A is the number of similar vector lengths (m+1) falling within a relative tolerance limit (r times standard deviation of the stride time intervals) and B is the number of similar vector lengths (m) falling within the tolerance limit[33]. The three parameters m, r, and N (time series length) should be selected prior to calculating SaEn and have been shown to have crucial importance for the SaEn value[33]. In order to control for parameter consistency, SaEn was calculated using m of 2 and 3 and r of 0.05, 0.1, 0.15, 0.2, 0.25, and 0.3. Based on this analysis

(see supplementary material 2) m of 2 and r of 0.2 were used for both the SaEn and ENT½ analyses.

Entropic half-life

ENT½ is based on consecutive calculations of SaEn with increasing randomization of the stride time interval time series as described briefly below and in detail elsewhere[2,34]. Firstly, SaEn is calculated on the original time series. Secondly, the original time series is gradually randomised through successive reshaping according to the principle described in figure 2. In the present study, each reshaping resulted in an increased distance between two subsequent strides. The stride time interval time series was reshaped 100 times and SaEn was calculated for each of the reshaped time series. These reshaped time series would, in all cases, exhibit SaEn values between the SaEn of the original time series (lowest SaEn value) and the SaEn of a complete random time series (highest SaEn value). The SaEn of the reshaped time series were normalized to the difference between these two extremes according to equation 2:

Equation 2: $$Normalized\ SaEn = \frac{SaEn_{RS} - SaEn_{OR}}{SaEn_{RAN} - SaEn_{OR}}$$

where $SaEn_{RS}$ is the SaEn of the reshaped time series, $SaEn_{OR}$ is the SaEn of the original time series, and $SaEn_{RAN}$ is the average SaEn of 100 randomized time series created by a random permutation of the data points in the original time series.

The normalized SaEn values were then plotted in a semi logarithmic plot as a function of the stride number (figure 2). The stride number corresponding to the first SaEn value above 0.5 was considered the stride number indicating a change in the characteristics of the reshaped time series from predictability to unpredictability and termed entropic half-life[2,34].

Detrended fluctuation analysis

The presence of statistical persistence or anti-persistence in the stride time interval time series was assessed using DFA. The correlations related to persistence or anti-persistence are part of the multifractal cascades that exist over a wide range of time scales[14]. DFA has the advantage of enabling detection of statistical persistence within noisy signals with embedded polynomial trends[24]. The applied DFA algorithm has been described in details elsewhere[24] and briefly below. Using equation 3, the time series x(i) is first integrated by calculating the cumulated sum of the deviations of the mean

Equation 3: $$y(k) = \sum_{i}^{k}[x(i) - x_{ave}]$$

Next, the time series is divided into boxes of equal length, *n* and a least square line is fitted to each box. The *y* coordinate of the straight-line segments is designated by $y_n(k)$ and used to detrend the time series y(k) before the root mean square is calculated (equation 4).

Equation 4: $$F(n) = \sqrt{\frac{1}{N}\sum_{k=1}^{N}[y(k) - y_n(k)]^2}$$

This procedure is repeated across the entire time series in order to establish a relationship between the average fluctuation, $F(n)$, as a function of box size *n.* The fluctuations can be characterized by the scaling exponent, which is determined by finding the slope of the line relating $\log F(n)$ to $\log n$[24]. In the current study, a box size range of [2,N] and a scaling region of 10 – 30 were used for the DFA as this range represented a linear section of the log*F(n)*-log*n* plot[14]. A scaling exponent greater than 0.5 indicated a presence of statistical persistence meaning that a deviation in stride time from the mean in one direction is more likely to be followed by a deviation in stride time in the same direction. A scaling exponent less than 0.5 indicated a presence of statistical anti-persistence, meaning that a deviation in stride time from the mean in one direction is more likely to be followed by a deviation in stride time in the opposite direction.

If the scaling exponent is 0.5, this indicated an uncorrelated white noise like pattern of the stride time interval time series[6,13,14].

Statistical persistence decay

The reshaped time series used for ENT½ analysis were also used for SPD. For each reshaped time series, DFA was performed as described above with a box size range of [2,N] and a scaling region of 10 – 30. A critical limit was calculated following equation 5 (figure 3).

Equation 5: $Critical\ limit = \mu_{\alpha Ran} + 2 \cdot \sigma_{\alpha Ran}$

Where $\mu_{\alpha Ran}$ is the averaged scaling exponent of 100 random time series created by a random permutation of the data points in the original time series and $\sigma_{\alpha Ran}$ is the corresponding standard deviation. Thus, the critical limit is based on the upper 95% confidence limit ($\mu + 2\sigma$) of the scaling exponent of the randomised time series. As the order of data points is changed with every rescaled time series, the statistical persistence is changed towards the critical limit. Any scaling exponent below the critical limit will not be significantly different from that of randomised patterns in the original time series. The number of strides required to reduce the scaling exponent below the critical limit is considered the SPD. Thus, SPD indicates a change in the stride time interval fluctuation from statistical persistence towards uncorrelated noise.

Statistics

Statistical difference in the analysis outcome measures (scaling exponent, SaEn, ENT½, and SPD) between the four mathematical signals were assessed using a one way ANOVA on ranks with signal types as independent factor and the outcome measures as the dependent measure. In case of an overall significant effect, a Turkey post hoc test was applied.

Paired Student's t-tests were applied to SaEn, ENT½, scaling exponent, and SPD to investigate if there was a statistical difference between the two walking conditions. To determine the nature of

the linear relationship between scaling exponent and SaEn and between ENT½ and SPD during both overground and treadmill walking, linear regression analyses were applied. For all statistical analysis the level of significance was set at 5%. All statistical calculations were performed in Sigmaplot (Systat Software, Inc. 2014, version 13.0, Germany).

Results

Theoretic signals

The scaling exponent, SaEn, ENT½ and SPD of the brown noise, Lorenz attractor, pink noise, and white Gaussian noise signals are presented in table 1. As known signals, the brown and Gaussian noise signals represented the two extremes (periodic and random, respectively) with the chaotic Lorenz attractor and pink noise signals as intermediate signals. The four analyses confirmed the known properties of the four theoretical signals. For all four analyses, there was a significant overall effect of the type of signal ($p < 0.001$ in all cases). The post hoc analyses revealed a significant difference between each signal type for the scaling exponent ($p \leq 0.034$ in all cases). SaEn did not differ between brown noise and the Lorenz attractor but was significantly higher for pink and Gaussian noise ($p \leq 0.033$ for all comparisons). ENT½ differed significantly between all signals ($p \leq 0.033$ for all comparisons) except between pink and Gaussian noise. The SPD differed significantly between all signals ($p \leq 0.049$ for all comparisons) except between Lorenz attractor and pink noise. While the brown noise signal had the strongest statistical persistence, lowest sample entropy, and the highest ENT½ and SPD, the random signal had the highest SaEn, and the lowest ENT½ and SPD. The Lorenz attractor and pink noise returned intermediate values for all four analyses.

Treadmill and overground walking

The SaEn was significantly greater ($p=0.033$) and the scaling exponent was significantly lower ($p=0.012$) during treadmill walking compared to overground walking (figure 4A and 4B). There was no significant difference between the two walking conditions for ENT½ (for overground walking: mean = 11.3 strides and median = 9 strides and for treadmill walking: mean = 14.6 strides and median = 8 strides) and SPD (for overground walking: mean = 51.6 strides and

median = 65 strides and for treadmill walking: mean = 53.0 strides and median = 59.5 strides). It should be noted that large inter-subject variations existed for ENT½ and SPD (figure 4C and 4D). Low non-significant correlations were observed both between SaEn and scaling exponent (for overground walking: R = 0.421, R^2 = 0.177, Adj. R^2 =0.109, p = 0.134; for treadmill walking: R = 0.103, R^2 = 0.011, Adj. R^2 = -0.072, p = 0.726) and between ENT½ and SPD (for overground walking: R = -0.064, R^2 = 0.004, Adj. R^2 = -0.079, p = 0.829; for treadmill walking: R = 0.502, R^2 = 0.252, Adj. R^2 = 0.190, p = 0.067) (figure 5).

Discussion

The present study aimed at introducing two novel tools, ENT½ and SPD, as methods to quantify time dependency in stride-to-stride time intervals during human gait. To validate the use of these methods, the dynamic characteristics of known theoretical signals (periodic, chaotic, and random) were assessed by DFA and SaEn in addition to ENT½ and SPD. It was expected that the periodic signal (brown noise) with low SaEn (high predictability) and with a high scaling exponent above 0.5 (high statistical persistency) would be characterized with high ENT½ and SPD, while the random signal (Gaussian noise) with high SaEn (low predictability) and a scaling exponent close to 0.5 (uncorrelated pattern) would be characterized with low ENT½ and SPD. Chaotic signals (Lorenz attractor and pink noise) would be characterized by intermediate SaEn, scaling exponent, ENT½, and SPD values. This expectation was confirmed which indicated that the presented methods were able to assess the number of data points involved in creating the potential time dependency in time series.

Furthermore, the present study applied ENT½ and SPD to the stride-to-stride time intervals recorded during treadmill and overground walking in healthy young adults. Significant lower predictability and lower statistical persistence were observed during treadmill walking compared to overground walking. However, no significant difference was observed for ENT½ and SPD between walking condition.

The present study is the first to estimate the time dependency of human gait in an interpretable scale. The predictability was halved within 11 and 14 consecutive strides during overground and treadmill walking, respectively, and the statistical persistence was deteriorated within ~50 strides during both conditions. This indicates a substantial time memory, where information from previous strides was included in the formation of future strides.

During the last twenty years, extensive research has addressed the time dependency and nonlinear dynamics of the inherent variability in human movements[21,29]. Both basic and applied research has acknowledged the functional role and importance of the observed movement variability[22,25,29]. In relation to human gait, variability has been discussed in relation to the fundamental motor control of walking[7,14,31], the development of a mature walking pattern[3,15], and the impairment following aging[3,13,16], and pathology[1,9,11,13,18]. The increasing interest in gait variability has led to a number of different nonlinear tools applied to either kinematic data or stride characteristics quantifying different characteristics of the time series in question (e.g. largest Lyapunov exponent quantifying the rate of trajectory divergence or convergence in state space, approximate and sample entropy quantifying predictability, correlation dimension quantifying dimensionality, and detrended fluctuation analysis quantifying statistical persistence or anti-persistence)[28]. While these different tools acknowledge the time dependency in the investigated time series, they do not quantify this time dependency on an interpretable physical or physiological time scale.

The characteristics and strength of this time dependency in movements observed during walking could be interpreted as the reliance of the motor control system on previous strides in order to perform future strides. The present study confirmed previous observations that stride time intervals during both overground and treadmill walking are characterised by statistical persistence, meaning that deviations in one direction are statistically likely to be followed by deviations in the same direction[31]. It has been emphasised that the statistical persistence should be interpreted within the context of the control process of the parameter in question, the influence of biomechanical, anatomical, and neuro-muscular redundancy and the task constraints[6]. Thus, the constraints imposed by the treadmill on the motor control system cause a

reduction of the statistical persistence which has been suggested to be linked to a tighter control[6,31]. Furthermore, Terrier and Deriaz[31] observed that a reduction in statistical persistence in stride time intervals was accompanied by a reduction in largest Lyapunov exponent of the centre of mass accelerations during treadmill walking. The authors interpreted this as an increase in gait stability during treadmill walking compared to overground walking. Interestingly, this proposed tighter control is accompanied by a lower predictability in the stride time intervals. Inducing further constraints through use of virtual reality environments with different optic flow, Katsavelis and colleagues[19] observed an additional decrease in predictability (quantified by approximate entropy).

Two alternative interpretations could be made based on this. It could be speculated that the tighter control during constraint walking (e.g. treadmill or virtual reality environment) increases the gait stability (decrease in largest Lyapunov exponent) through more usage of the available degrees of freedom (increase in sample/approximate entropy). This interpretation suggests that treadmill walking constitutes an optimal walking condition compared to overground walking. Alternatively, it could be speculated that the unconstrained overground walking is successfully performed through a more flexible control which relies on the self-organized interplay of the degrees of freedom within the body. This interaction creates a movement solution characterized by higher statistical persistence and a sufficient level of predictability enabling an adaptable walking pattern. This would furthermore indicate that the human locomotor control during unconstrained walking is more complex compared to constrained walking as suggested by Costa et al.[5]. Accordingly, constrained walking could be considered a more challenging task compared to unconstrained walking for the motor control system to solve which induces a more random-like pattern in the stride-time intervals. As a consequent, tighter motor control reduces the rate of

divergence of the centre of mass accelerations to ensure a stable upper body motion. Although both explanations are valid, we consider the latter to be the most likely. In support of this explanation, Dingwell et al.[7] observed that during treadmill walking stride speed fluctuations exhibited anti-persistence while stride time and stride length exhibited statistical persistence. It was suggested that while stride speed required a tighter motor control in order to stay in the middle of the treadmill belt, stride time and stride length was allowed to fluctuate more freely with more flexible control due to the redundancy of these two parameters[7]. Furthermore, Terrier et al.[32] observed that during overground walking the stride speed as well as stride time and stride length exhibited statistical persistence and suggested that these parameters were allowed to fluctuate freely.

In contrast to the scaling exponent and SaEn, the two walking conditions did not induce differences in ENT½ and SPD. This could indicate that while the constraints imposed by the treadmill may affect the observed statistical persistence and predictability, it does not seem to affect the rate at which this deteriorates. Furthermore, the present study investigated the relationship between SaEn and scaling exponent and between ENT½ and SPD, to assess the potential shared mechanisms behind the generation of statistical persistence and predictability observed in stride time intervals. The result showed low non-significant correlations between SaEn and scaling exponent and between ENT½ and SPD for both walking conditions indicating no relationship between the generation of statistical persistence predictability and between the rates of deterioration of the statistical persistence and gradually reduction in predictability. Based on this, it could be speculated that the sources of these characteristics in the gait patterns does not changes these parameter synchronously. It is well established that the motor cortex, corticospinal tract, and spinal cord are involved in the motor control of human locomotion[23]. Although

potentially involved in generating the observed time dependency in gait, the actual contribution from the nervous system is unknown. Furthermore, the need for involvement of higher cortical-spinal structures in the formation of the statistical persistence observed in human gait has been questioned[8]. In addition, it has previously been established that the statistical persistence in stride time intervals is lower in patients of neurological diseases, elderly, and fall prone individuals[10,13,16]. Thus, pathological and age-related changes in both neurological musculoskeletal structures could contribute to altered temporal structure of the gait pattern[8]. However, whether the time dependency is affected in these types of individuals is a topic for future research.

The present study only included walking at the preferred walking speed of the included subjects. Thus, the motor control was not challenged beyond what could be considered the less demanding walking task. Thus, it remains unknown if the time dependency of stride time intervals quantified with ENT½ and SPD exhibits the same walking speed relationship as previously shown for scaling exponent at walking speeds beyond and below the preferred walking speed[17]. Time dependency was only quantified in the stride-to-stride time intervals. Thus, it remains unknown if the observed results also apply to other gait characteristics (i.e., stride length, stride speed). Additionally, future studies should explore the potential of using ENT½ and SPD on lower limb joint angle trajectory data obtained during walking. Previous studies have applied various nonlinear analyses (e.g. largest Lyapunov exponent, correlation dimension[1,4,26]) to assess joint angles dynamics. Thus, difference the knee angle joint dynamics has been observed between young and elderly individuals[4] and between the injured and healthy knee in ACL deficient patients[30]. It is likely that differences in the time dependency in joint angles during various locomotion tasks could be detected by ENT½ and SPD. The gait pattern of elderly individuals

has been observed to be more random compared to younger individuals[4,10]. This would potentially be characterized by a shorter ENT½ and SPD in the elderly group. However, future studies should investigate this topic.

The present study introduced ENT½ and SPD as novel methods to quantify time dependency during overground and treadmill walking and were able to show that predictability in stride time intervals was halved after approximate 14 strides and that the statistical persistence was deteriorated to uncorrelated noise after approximately 50 strides. These observations were accompanied by a lower statistical persistence and lower predictability during treadmill walking compared to overground walking.

Acknowledgements

Funding for this project was provided by NASA Nebraska EPSCoR and the National Institutes of Health (P20 GM109090).

References

1. Alkjaer T., P. C. Raffalt, H. Dalsgaard, E. B. Simonsen, N. C. Petersen, H. Bliddal and M. Henriksen. Gait variability and motor control in people with knee osteoarthritis. *Gait Posture* 42: 479-484, 2015.
2. Baltich J., V. Von Tscharner, P. Zandiyeh and B. M. Nigg. Quantification and reliability of center of pressure movement during balance tasks of varying difficulty. *Gait Posture* 40: 327-332, 2014.
3. Bisi M. C. and R. Stagni. Complexity of human gait pattern at different ages assessed using multiscale entropy: From development to decline. *Gait Posture* 47: 37-42, 2016.
4. Buzzi U. H., N. Stergiou, M. J. Kurz, P. A. Hageman and J. Heidel. Nonlinear dynamics indicates aging affects variability during gait. *Clin Biomech (Bristol, Avon)* 18: 435-443, 2003.
5. Costa M., C. K. Peng, A. L. Goldberger and J. M. Hausdorff. Multiscale entropy analysis of human gait dynamics. *Physica A* 330: 53-60, 2003.
6. Dingwell J. B. and J. P. Cusumano. Re-interpreting detrended fluctuation analyses of stride-to-stride variability in human walking. *Gait Posture* 32: 348-353, 2010.
7. Dingwell J. B., J. John and J. P. Cusumano. Do humans optimally exploit redundancy to control step variability in walking? *PLoS Comput Biol* 6: e1000856, 2010.
8. Gates D. H., J. L. Su and J. B. Dingwell. Possible Biomechanical Origins of the Long-Range Correlations in Stride Intervals of Walking. *Physica A* 380: 259-270, 2007.
9. Georgoulis A. D., C. Moraiti, S. Ristanis and N. Stergiou. A novel approach to measure variability in the anterior cruciate ligament deficient knee during walking: the use of the approximate entropy in orthopaedics. *J Clin Monit Comput* 20: 11-18, 2006.
10. Goldberger A. L., L. A. Amaral, J. M. Hausdorff, P. Ivanov, C. K. Peng and H. E. Stanley. Fractal dynamics in physiology: alterations with disease and aging. *Proc Natl Acad Sci U S A* 99 Suppl 1: 2466-2472, 2002.
11. Hausdorff J. M. Gait dynamics in Parkinson's disease: common and distinct behavior among stride length, gait variability, and fractal-like scaling. *Chaos* 19: 026113, 2009.
12. Hausdorff J. M., Y. Ashkenazy, C. K. Peng, P. C. Ivanov, H. E. Stanley and A. L. Goldberger. When human walking becomes random walking: fractal analysis and modeling of gait rhythm fluctuations. *Physica A* 302: 138-147, 2001.
13. Hausdorff J. M., S. L. Mitchell, R. Firtion, C. K. Peng, M. E. Cudkowicz, J. Y. Wei and A. L. Goldberger. Altered fractal dynamics of gait: reduced stride-interval correlations with aging and Huntington's disease. *J Appl Physiol (1985)* 82: 262-269, 1997.
14. Hausdorff J. M., P. L. Purdon, C. K. Peng, Z. Ladin, J. Y. Wei and A. L. Goldberger. Fractal dynamics of human gait: stability of long-range correlations in stride interval fluctuations. *J Appl Physiol (1985)* 80: 1448-1457, 1996.
15. Hausdorff J. M., L. Zemany, C. Peng and A. L. Goldberger. Maturation of gait dynamics: stride-to-stride variability and its temporal organization in children. *J Appl Physiol (1985)* 86: 1040-1047, 1999.
16. Herman T., N. Giladi, T. Gurevich and J. M. Hausdorff. Gait instability and fractal dynamics of older adults with a "cautious" gait: why do certain older adults walk fearfully? *Gait Posture* 21: 178-185, 2005.
17. Jordan K., J. H. Challis and K. M. Newell. Walking speed influences on gait cycle variability. *Gait Posture* 26: 128-134, 2007.
18. Kaipust J. P., J. M. Huisinga, M. Filipi and N. Stergiou. Gait variability measures reveal differences between multiple sclerosis patients and healthy controls. *Motor Control* 16: 229-244, 2012.
19. Katsavelis D., M. Mukherjee, L. Decker and N. Stergiou. The effect of virtual reality on gait variability. *Nonlinear Dynamics Psychol Life Sci* 14: 239-256, 2010.
20. Lee S. J. and J. Hidler. Biomechanics of overground vs. treadmill walking in healthy individuals. *J Appl Physiol (1985)* 104: 747-755, 2008.

21. Newell K. M. and D. M. Corcos. Issues in variability and motor control. In: *Variability and motor control*, edited by K. M. Newell and D. M. Corcos. Champagne, IL: Human Kinetics, 1993.
22. Newell K. M. and D. E. Vaillancourt. Dimensional change in motor learning. *Hum Mov Sci* 20: 695-715, 2001.
23. Nielsen J. B. How we walk: central control of muscle activity during human walking. *Neuroscientist* 9: 195-204, 2003.
24. Peng C. K., S. Havlin, H. E. Stanley and A. L. Goldberger. Quantification of scaling exponents and crossover phenomena in nonstationary heartbeat time series. *Chaos* 5: 82-87, 1995.
25. Preatoni E., J. Hamill, A. J. Harrison, K. Hayes, R. E. Van Emmerik, C. Wilson and R. Rodano. Movement variability and skills monitoring in sports. *Sports Biomech* 12: 69-92, 2013.
26. Raffalt P. C., M. K. Guul, A. N. Nielsen, S. Puthusserypady and T. Alkjaer. Economy, Movement Dynamics, and Muscle Activity of Human Walking at Different Speeds. *Sci Rep* 7: 43986, 2017.
27. Richman J. S. and J. R. Moorman. Physiological time-series analysis using approximate entropy and sample entropy. *Am J Physiol Heart Circ Physiol* 278: H2039-2049, 2000.
28. Stergiou N. *Nonlinear analysis for human movement variability*. Boca Raton, FL: CRC Press, 2016.
29. Stergiou N. and L. M. Decker. Human movement variability, nonlinear dynamics, and pathology: is there a connection? *Hum Mov Sci* 30: 869-888, 2011.
30. Stergiou N., C. Moraiti, G. Giakas, S. Ristanis and A. D. Georgoulis. The effect of the walking speed on the stability of the anterior cruciate ligament deficient knee. *Clin Biomech (Bristol, Avon)* 19: 957-963, 2004.
31. Terrier P. and O. Deriaz. Kinematic variability, fractal dynamics and local dynamic stability of treadmill walking. *J Neuroeng Rehabil* 8: 12, 2011.
32. Terrier P., V. Turner and Y. Schutz. GPS analysis of human locomotion: further evidence for long-range correlations in stride-to-stride fluctuations of gait parameters. *Hum Mov Sci* 24: 97-115, 2005.
33. Yentes J. M., N. Hunt, K. K. Schmid, J. P. Kaipust, D. McGrath and N. Stergiou. The appropriate use of approximate entropy and sample entropy with short data sets. *Ann Biomed Eng* 41: 349-365, 2013.
34. Zandiyeh P. and V. Von Tscharner. Reshape scale method: a novel multi scale entropic analysis approach. *Physica A* 392: 6265-6272, 2013.

Legends

Table 1: Mean ± SD of scaling exponent, sample entropy, entropic half-life and statistical persistence decay of twenty iteration of brown noise, Lorenz attractor, pink noise signal and white Gaussian noise. NOTE: a) indicates significant difference from brown noise signal, b) indicates significance different from Lorenz attractor signal, c) indicates significant difference from pink noise signal ($p<0.05$).

Figure 1: Examples of the brown noise signal, Lorenz attractor signal, pink noise signal, and the white Gaussian noise signal (top four signals) and an example of the stride time interval time series for overground and treadmill walking (bottom two signals). One hundred data points or stride time intervals are depicted.

Figure 2: Calculation procedure for ENT½ (see text for details).

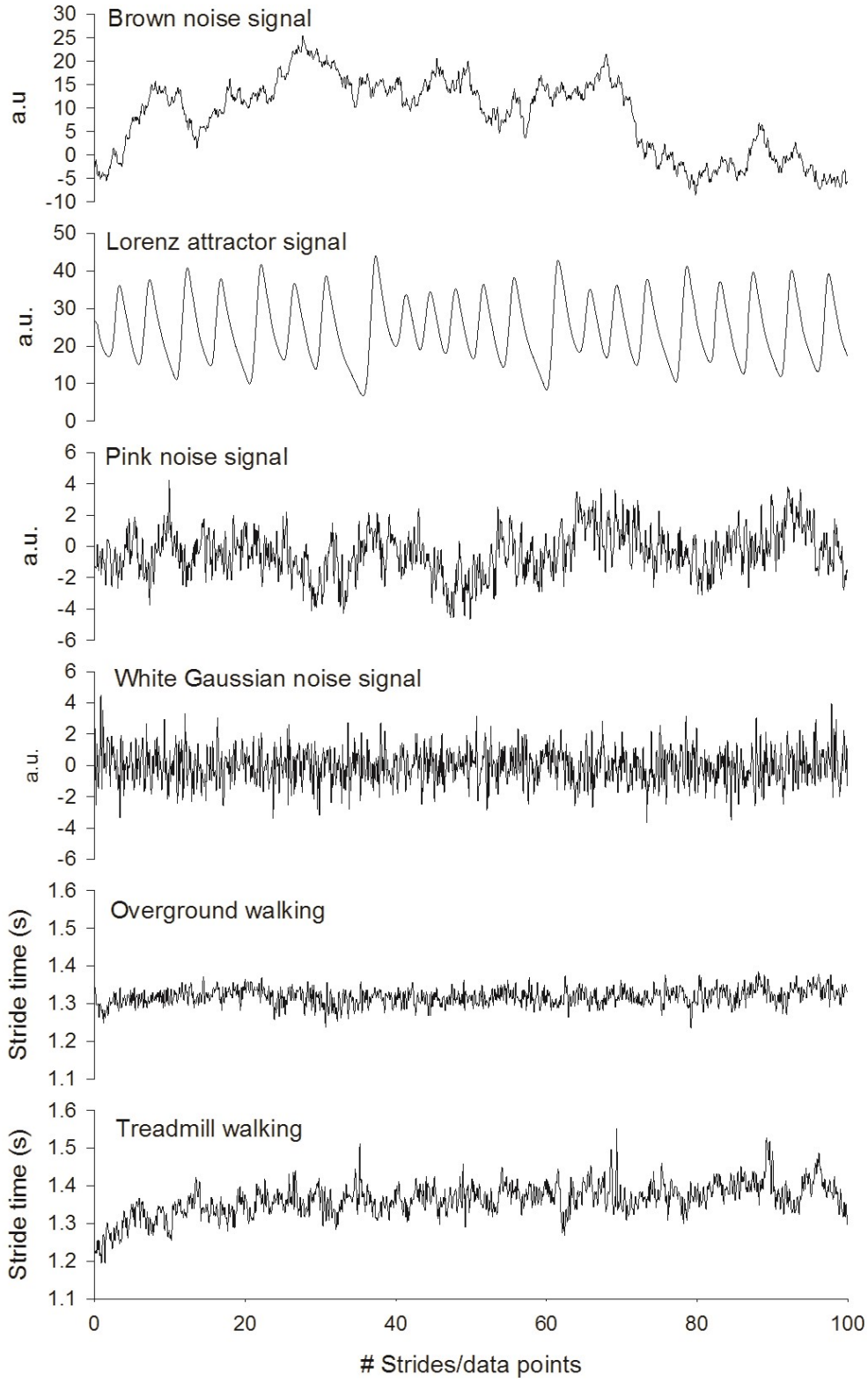
Figure 3: Calculation procedure for SPD (see text for details).

Figure 4A-D: Boxplot including group mean (dashed line) and median (solid line) for SaEn (A), scaling exponent (B), ENT½ (C), and SPD (D) for overground walking (OG) and treadmill walking (TM).

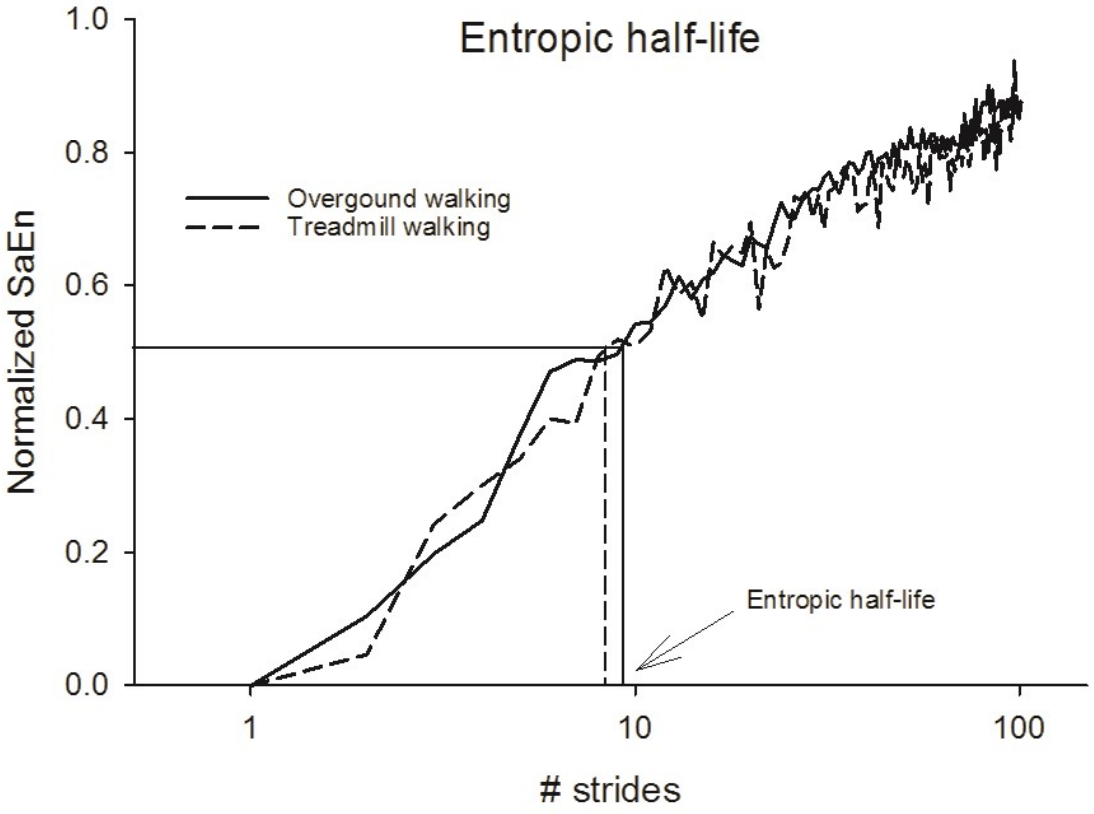
Figure 5: Linear regression of SaEn and scaling exponent and of ENT½ and SPD for both overground and treadmill walking.

Table 1: Mean ± SD of scaling exponent, sample entropy, entropic half-life and statistical persistence decay of twenty iteration of brown noise, Lorenz attractor, pink noise signal and white Gaussian noise. NOTE: a) indicates significant difference from brown noise signal, b) indicates significance different from Lorenz attractor signal, c) indicates significant difference from pink noise signal ($p<0.05$).

	Brown noise	Lorenz attractor	Pink noise	Gaussian noise
Alpha	1.40 ± 0.05	1.08 ± 0.03[a]	0.93 ± 0.06[a,b]	0.47 ± 0.03[a,b,c]
SaEn	0.18 ± 0.08	0.22 ± 0.01	1.80 ± 0.05[a,b]	2.18 ± 0.01[a,b,c]
ENT½	43.30 ± 29.88	8.00 ± 0.00[a]	4.20 ± 0.83[a,b]	2.60 ± 0.82[a,b]
SPD	68.65 ± 11.14	16.95 ± 15.10[a]	45.05 ± 18.59[a]	1.05 ± 0.22[a,b,c]

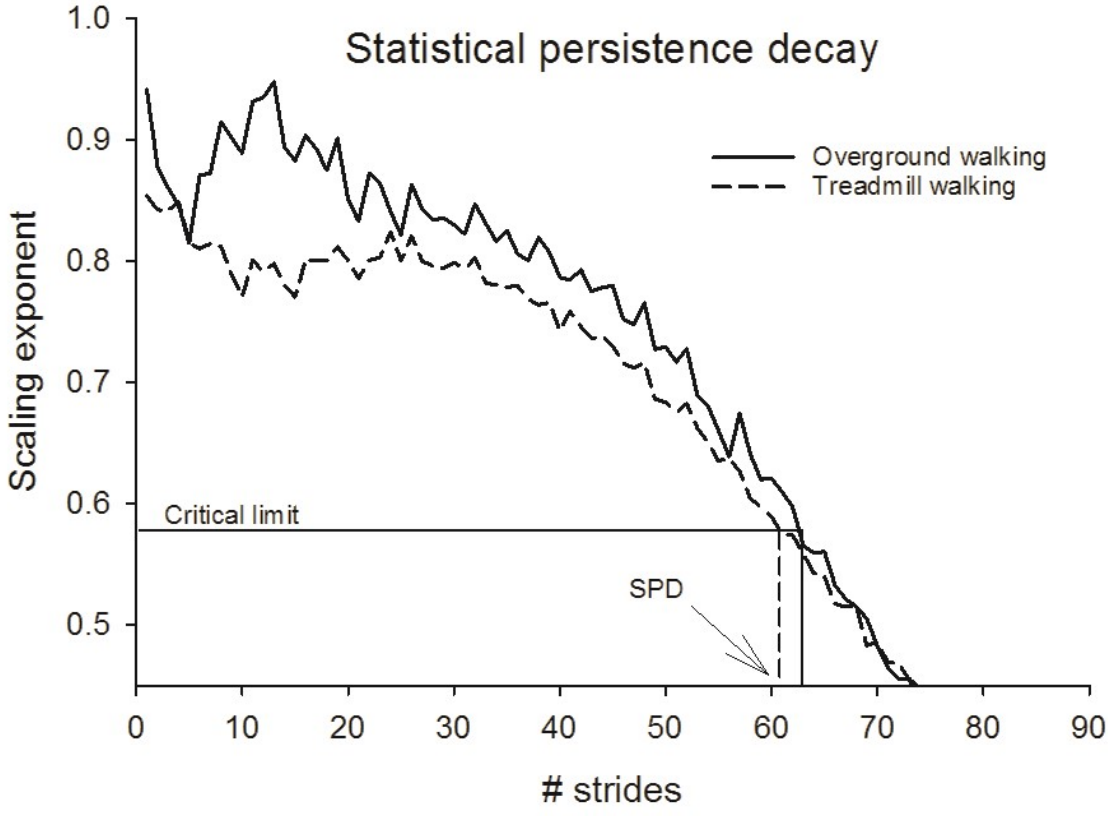

Brown noise signal
a.u
Lorenz attractor signal
a.u.
Pink noise signal
a.u.
White Gaussian noise signal
a.u.
Overground walking
Stride time (s)
Treadmill walking
Stride time (s)
Strides/data points

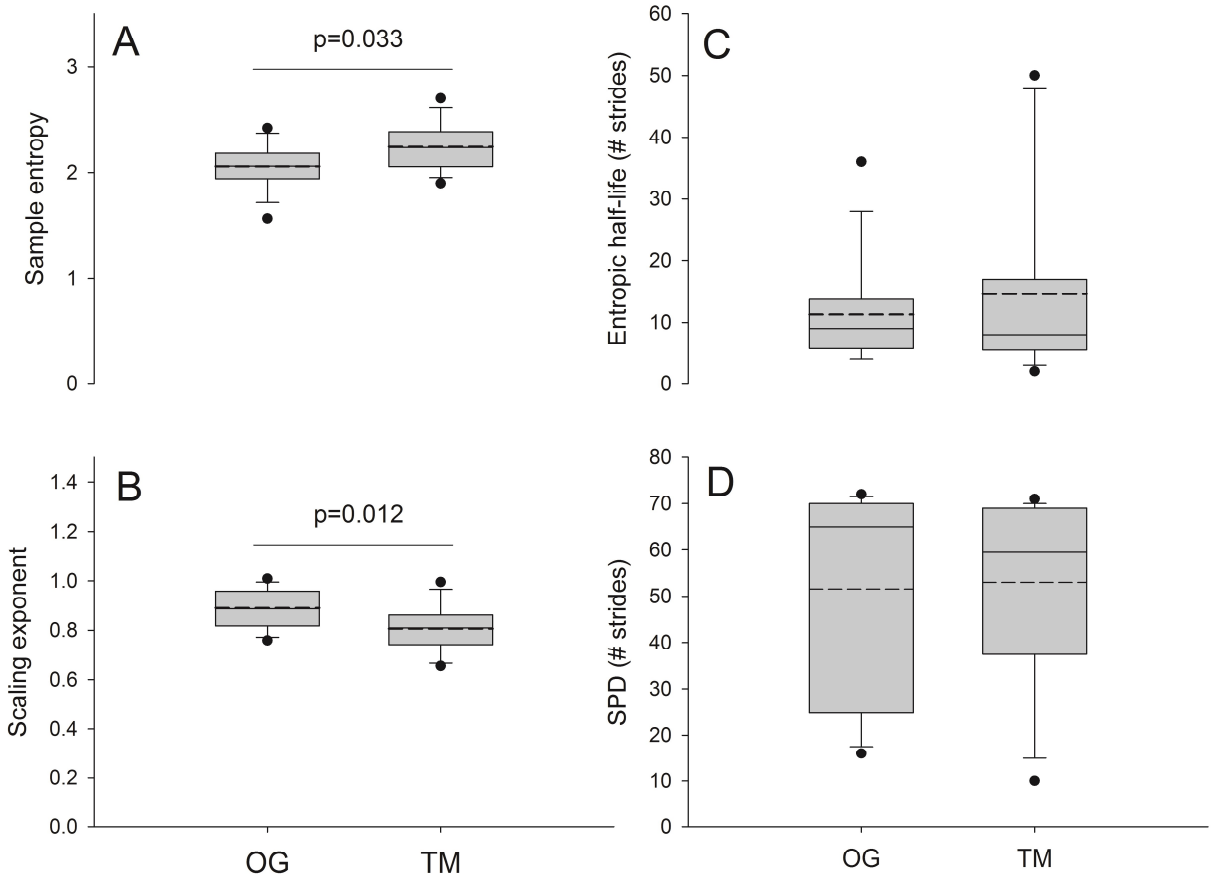
1) Calculate SaEn for the original time series ($SaEn_{OR}$)
2) Create one hundred reshaped time series through successively reshaping the original time series which gradually randomised each new reshaped time series following the principle described below:
   Original time series: [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16]
   Reshaped time series 2: [1 3 5 7 9 11 13 15 2 4 6 8 10 12 14 16]
   Reshaped time series 3: [1 4 7 10 13 16 2 5 8 11 14 3 6 9 12 15]
   Reshaped time series 4: [1 5 9 13 2 6 10 14 3 7 11 15 4 8 12 16]
3) Calculate SaEn for each reshape time series ($SaEn_{RS}$)
4) Calculate the average SaEn for one hundred randomized time series created by a random permutation of the data points in the original time series ($SaEn_{RAN}$)
5) Normalize the $SaEn_{RS}$ to the difference between the $SaEn_{OR}$ and $SaEn_{RAN}$ according to equation 2 (see text)
6) Plot the $SaEn_{RS}$ in as a function of the stride number (see below)
7) The stride number at which $SaEn_{RS}$ exceeds 0.5 is defined as the entropic half-life

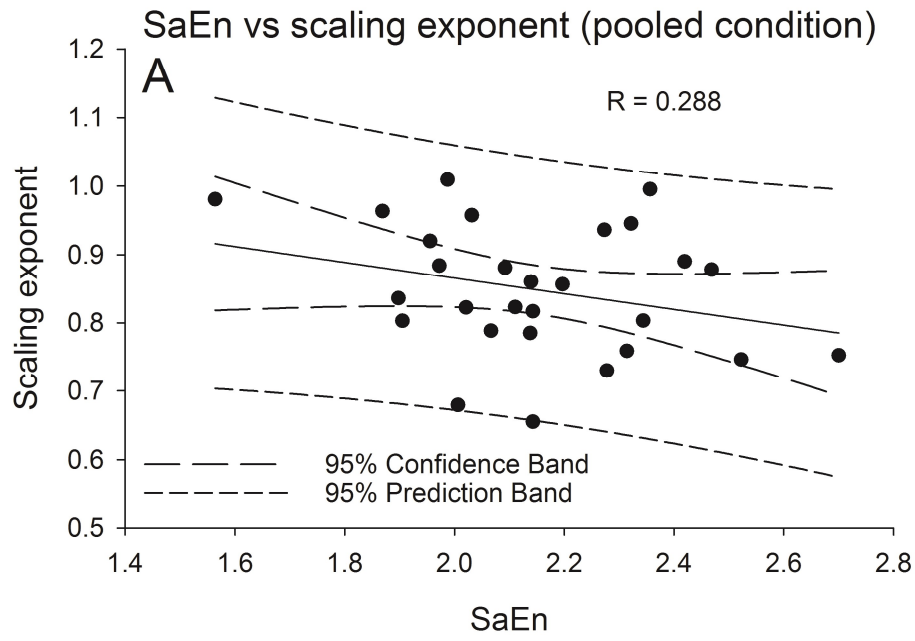


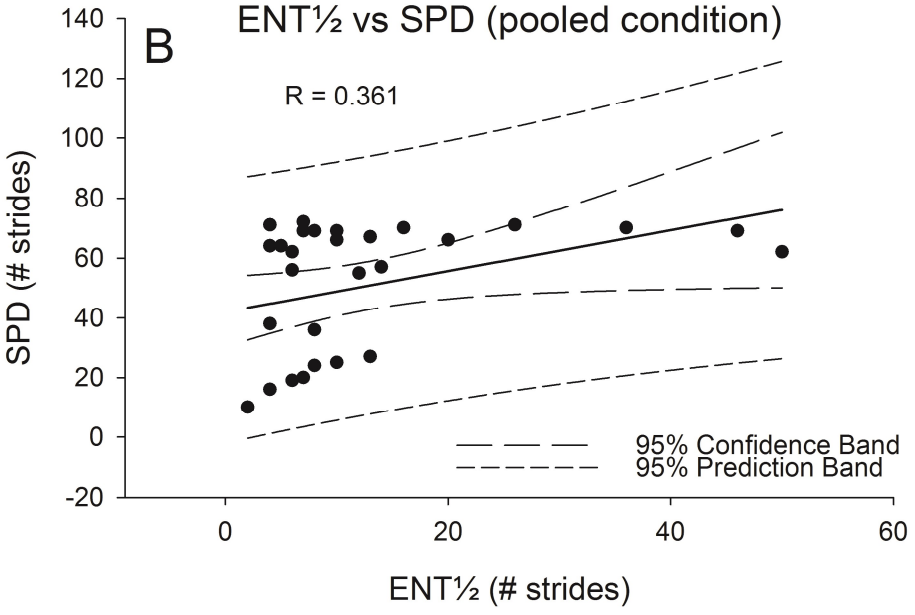
1) Calculate scaling exponent for the original time series
2) Create one hundred reshaped time series through successively reshaping the original time series which gradually randomised each new reshaped time series following the principle described below:

   Original time series: [1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16]
   Reshaped time series 2: [1 3 5 7 9 11 13 15 2 4 6 8 10 12 14 16]
   Reshaped time series 3: [1 4 7 10 13 16 2 5 8 11 14 3 6 9 12 15]
   Reshaped time series 4: [1 5 9 13 2 6 10 14 3 7 11 15 4 8 12 16]
3) Calculate scaling exponent for each reshaped time series
4) Calculate the average and standard deviation of the scaling exponents of one hundred randomized time series created by a random permutation of the data points in the original time series ($\mu_{\alpha RAN}$ and $\sigma_{\alpha RAN}$)
5) Calculate a critical limit according to equation 5 (see text)
6) Plot the scaling exponents of the reshaped time series as function of the stride number (see below)
7) The stride number at which the scaling exponent is below the critical limit is defined as the statistical persistence decay (SPD)




A
p=0.033
Sample entropy
0
1
2
3
C
Entropic half-life (# strides)
0
10
20
30
40
50
60
B
p=0.012
Scaling exponent
0.0
0.2
0.4
0.6
0.8
1.0
1.2
1.4
D
SPD (# strides)
0
10
20
30
40
50
60
70
80
OG
TM
OG
TM


SaEn vs scaling exponent (pooled condition)
A
R = 0.288
Scaling exponent
SaEn
95% Confidence Band
95% Prediction Band
1.2
1.1
1.0
0.9
0.8
0.7
0.6
0.5
1.4
1.6
1.8
2.0
2.2
2.4
2.6
2.8


ENT½ vs SPD (pooled condition)
B
R = 0.361
SPD (# strides)
ENT½ (# strides)
95% Confidence Band
95% Prediction Band
140
120
100
80
60
40
20
0
-20
0
20
40
60

SUPPLEMENTARY MATERIAL

Title:

Introducing Statistical Persistence Decay – A Quantification of Stride-to-Stride Time Interval Dependency in Human Gait

Authors:

Raffalt, PC.[1] and Yentes, JM.[2]

Affiliations:

1) Julius Wolff Institute for Biomechanics and Musculoskeletal Regeneration, Charité – Universitätsmedizin Berlin, Berlin, Germany.

2) Department of Biomedical Sciences, University of Copenhagen, Copenhagen, Denmark.

3) Center for Research in Human Movement Variability, Department of Biomechanics, University of Nebraska Omaha, Omaha, NE, USA.

Corresponding Author:

Jennifer M. Yentes

Center for Research in Human Movement Variability

University of Nebraska Omaha

Email: jyentes@gmail.com

Mobil phone: 402 554 3251

Supplementary material 1

In this material the stride time interval time series for each subject during overground and treadmill walking is presented. The trials included 2500 strides and are presented in four figures.

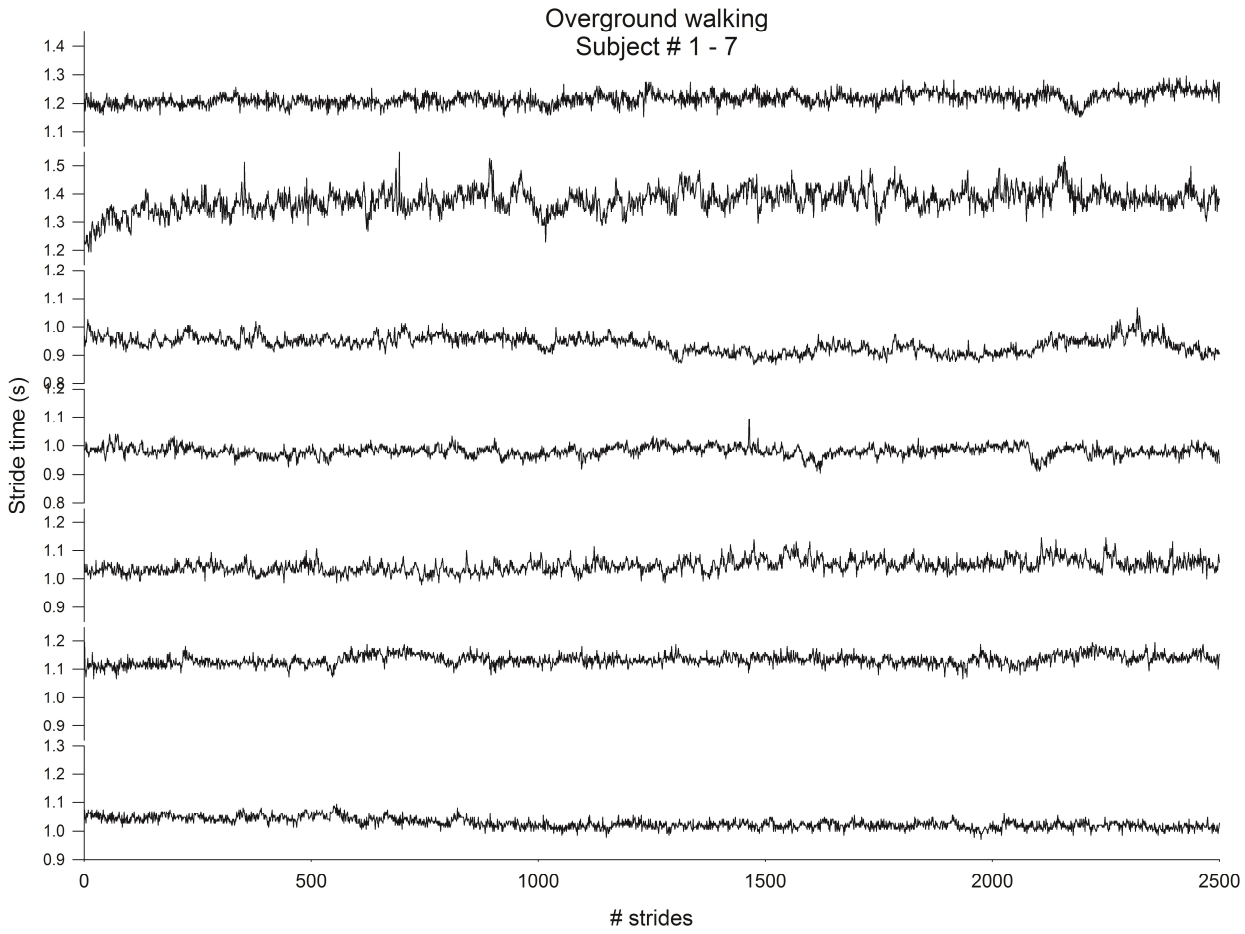


Figure S1.1: Stride time interval time series for subjects 1 – 7 during overground walking.

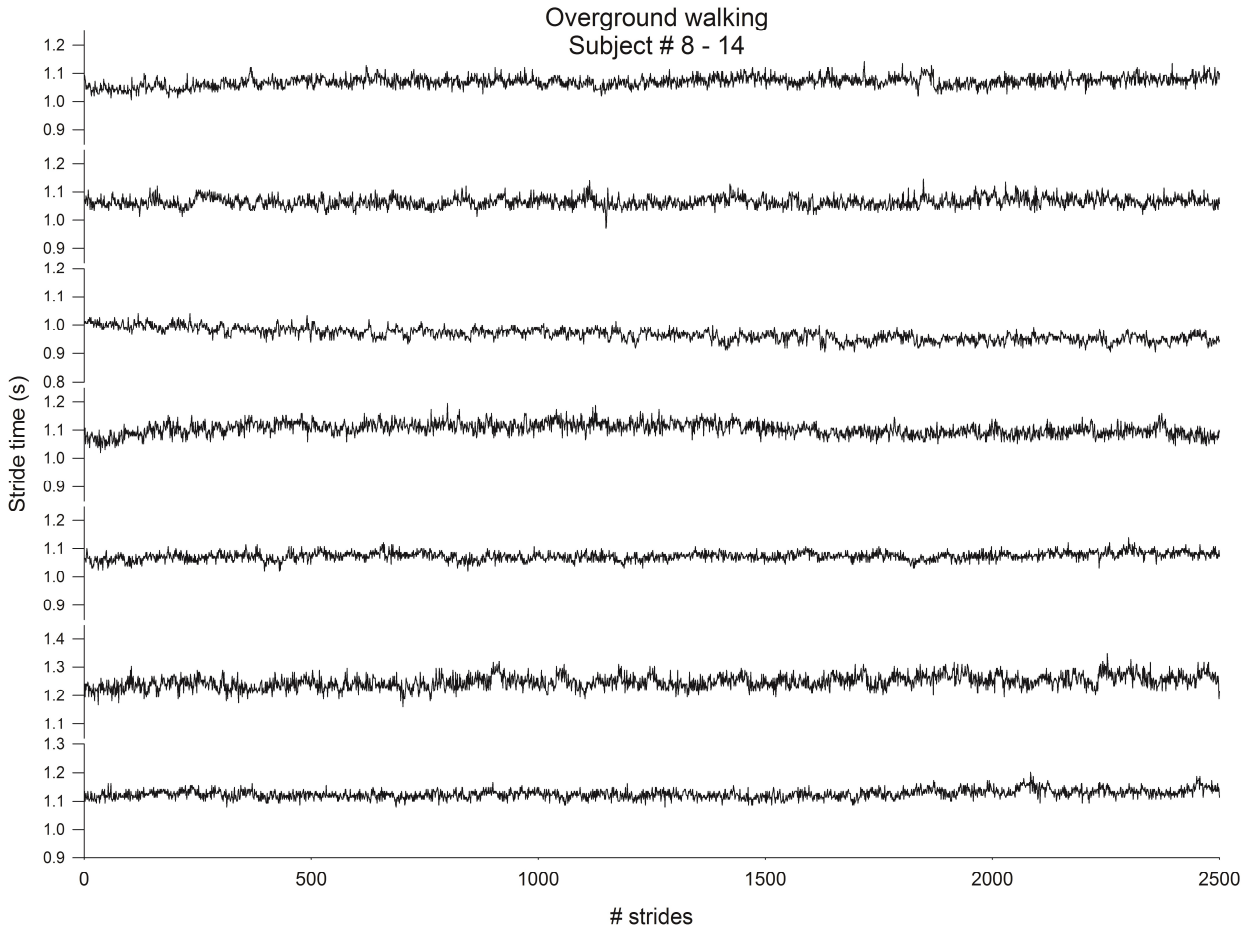


Figure S1.2: Stride time interval time series for subjects 8 – 14 during overground walking.

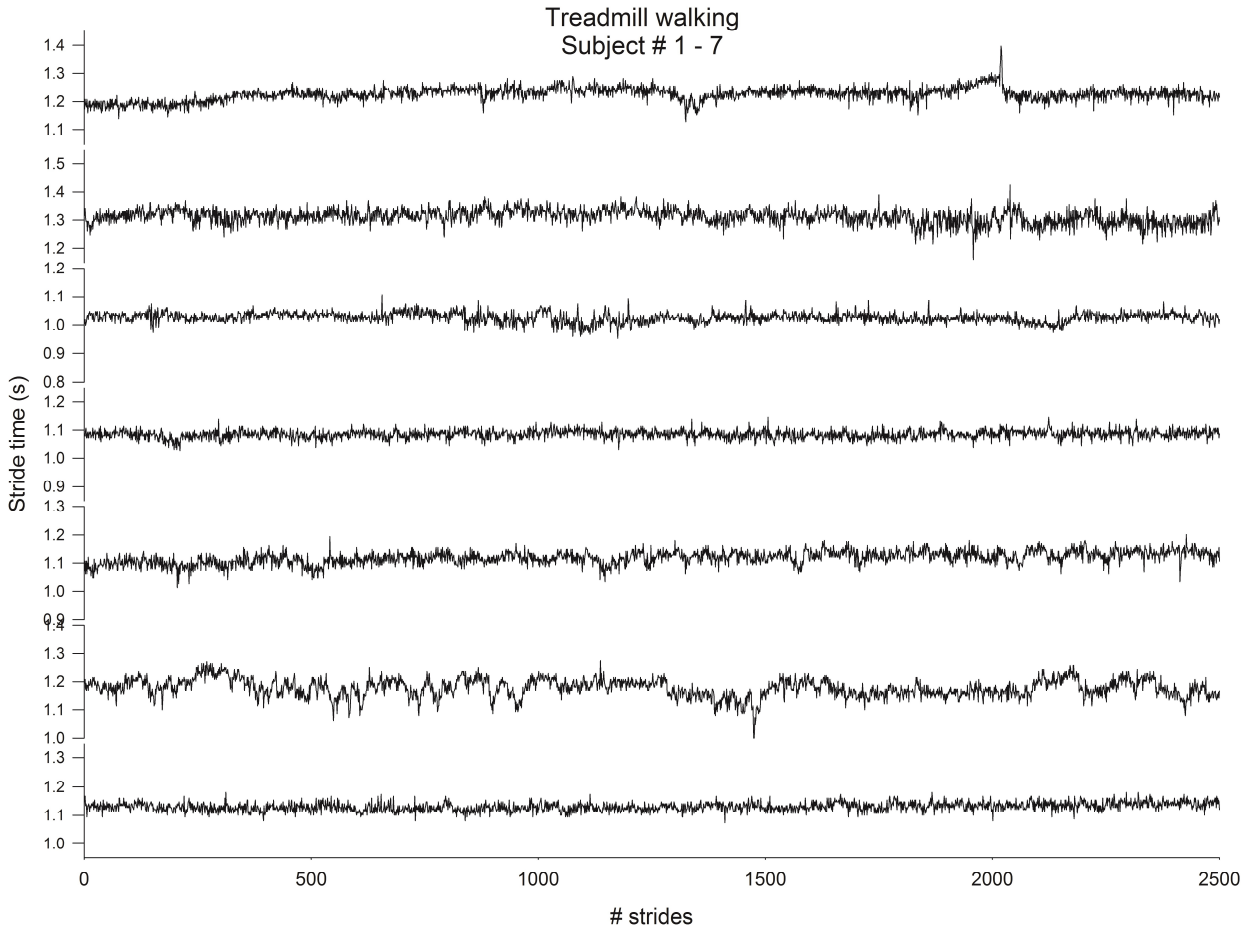


Figure S1.3: Stride time interval time series for subjects 1 – 7 during treadmill walking.

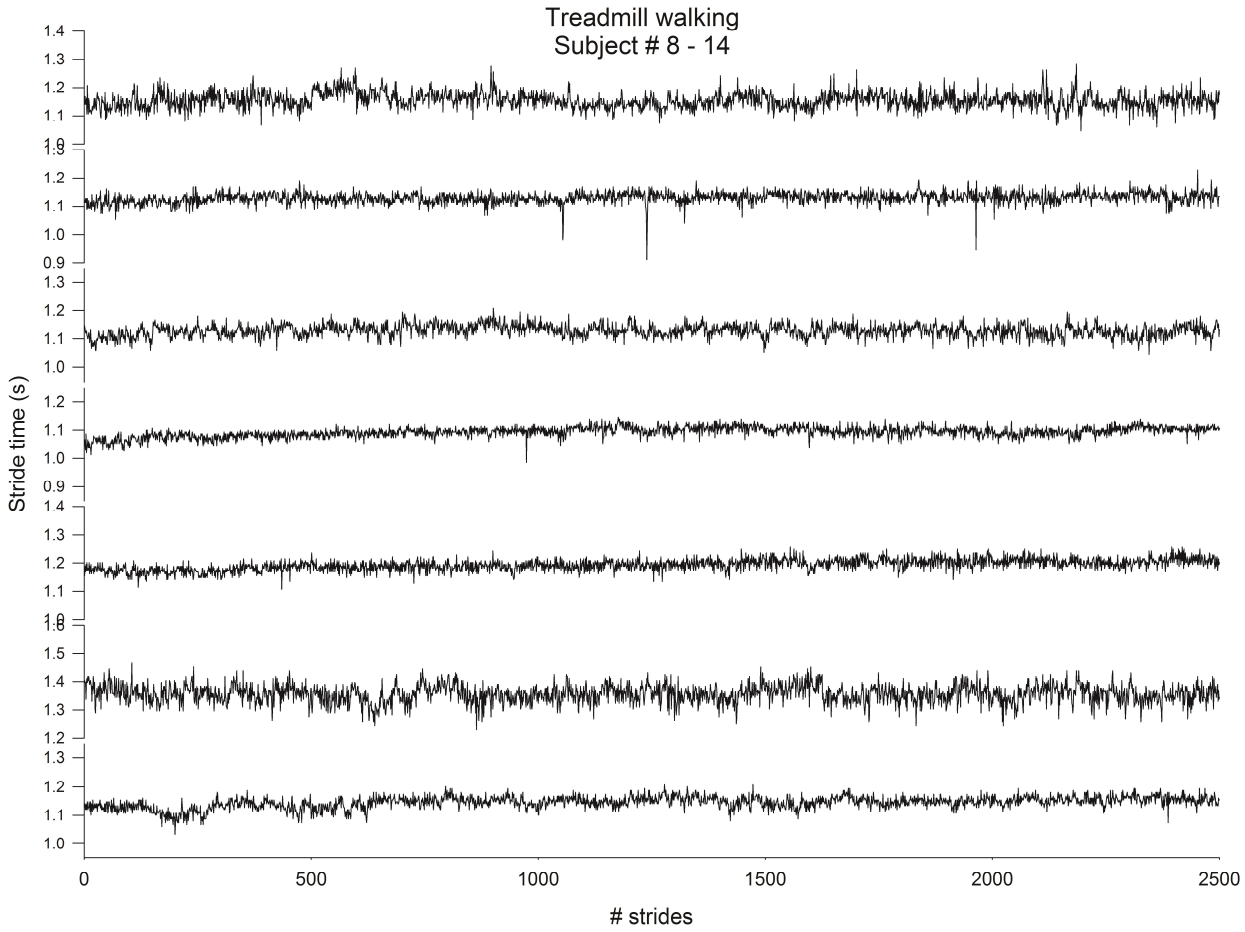


Figure S1.4: Stride time interval time series for subjects 8 – 14 during treadmill walking.

Supplementary material 2

Introduction

Sample entropy (SaEn) analysis of a time series require the determination of three parameters 1) r, the similarity criterion, 2) m, the length of data that will be compared and 3) N, the length of the time series prior to the calculation. As shown by Yentes and colleagues (2013), the selected parameters are crucial to the outcome of sample entropy (SaEn) calculations. Applying different parameters to the same time series will potentially change the outcome. To avoid this, the use of SaEn to time series should be accompanied by a test for parameter consistency. The purpose of the supplementary material was to investigate the parameter consistency for the time series in the present study.

Method

The present study used 2500 data point in each time series which is well beyond what is needed to achieve a consistent SaEn outcome (Yentes et al. 2013). Therefore, parameter consistency analysis was only performed for parameters r and m. SaEn was calculated for the stride time interval time series during both treadmill and overground walking using r = 0.5, 1, 1.5, 2, 2.5 and 3 and m = 2 and 3. A two-way ANOVA for repeated measures was applied to evaluate the effect of different r and m values for both the overground and treadmill data. In case of significant effects, a Holm-Sidak post hoc test was completed. Level of significance was set at 5 %. Calculations were conducted in SigmaPlot (Systat Software, Inc. 2014, version 13.0, Germany).

Results and Conclusion

There was no significant effect of the different r values on the SaEn values (figure S1). The non-significant decrease in SaEn at higher r values was similar for both the overground and treadmill walking data. There was a significant main effect of walking condition ($p=0.029$ for $m=2$ and $p=0.036$ for $m=3$) indicating that across r values the SaEn was significantly higher during treadmill walking compared to overground walking. There was a significant effect of the parameter m for both overground walking ($p<0.001$) and treadmill walking ($p<0.001$) (figure S1). Thus, SaEn was significantly higher with $m=2$ compared to $m=3$ for both walking conditions. It is noteworthy that the inter-condition relationship did not change with changes in r. Equally, the inter-parameter relationship of m did not change with changes in r.

Based on these results (figure S2.1), it could be concluded that there in general was a satisfying parameter consistency for the tested parameters and that r = 0.2 and m = 2 was well suited for SaEn calculations of the data in the present study.

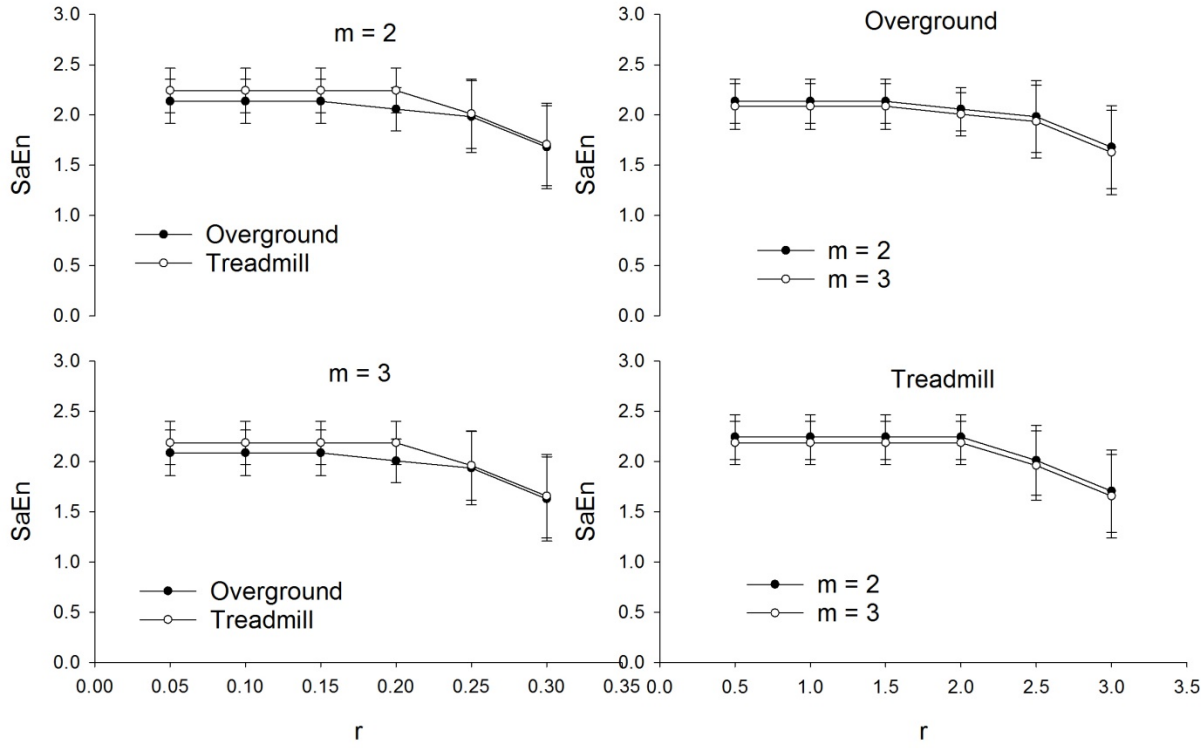


Figure S2.1: Group mean ± SD of SaEn calculated for overground and treadmill walking with r values of 0.05, 0.1, 0.15, 0.2, 0.25 and 0.3 and m values of 2 and 3.

References

Yentes, J. M., Hunt, N., Schmid, K. K., Kaipust, J. P., McGrath, D., & Stergiou, N. (2013). The appropriate use of approximate entropy and sample entropy with short data sets. *Ann Biomed Eng, 41*, 349-365.