



Full length article

Effect of parameter selection on entropy calculation for long walking trials

Jennifer M. Yentes^{a,*}, William Denton^a, John McCamley^b, Peter C. Raffalt^{c,d}, Kendra K. Schmid^e^a Department of Biomechanics, Center for Research in Human Movement Variability, University of Nebraska at Omaha, 6160 University Drive, Omaha, NE 68182-0860, USA^b MORE Foundation, 18444 N 25th Ave Ste 110, Phoenix, AZ 85023, USA^c Julius Wolff Institute for Biomechanics and Musculoskeletal Regeneration, Charité – Universitätsmedizin Berlin, Augustenburger Platz 1, 13353 Berlin, Germany^d Department of Biomedical Sciences, University of Copenhagen, Blegdamsvej 3B, 2200 Copenhagen N, Denmark^e Department of Biostatistics, University of Nebraska Medical Center, 984355 Medical Center, Omaha, NE 68198-4375, USA

ARTICLE INFO

Keywords:

Locomotion
Gait
Treadmill
Predictability
Regularity
Complexity

ABSTRACT

It is sometimes difficult to obtain uninterrupted data sets that are long enough to perform nonlinear analysis, especially in pathological populations. It is currently unclear as to how many data points are needed for reliable entropy analysis. The aims of this study were to determine the effect of changing parameter values of m , r , and N on entropy calculations for long gait data sets using two different modes of walking (i.e., overground versus treadmill). Fourteen young adults walked overground and on a treadmill at their preferred walking speed for one-hour while step time was collected via heel switches. Approximate (ApEn) and sample entropy (SampEn) were calculated using multiple parameter combinations of m , N , and r . Further, r was tested under two cases r *standard deviation and r constant. ApEn differed depending on the combination of r , m , and N . ApEn demonstrated relative consistency except when $m = 2$ and the smallest r values used ($r_{SD} = 0.015*SD$, $0.20*SD$; $r_{Constant} = 0$ and 0.003). For SampEn, as r increased, SampEn decreased. When r was constant, SampEn demonstrated excellent relative consistency for all combinations of r , m , and N . When r constant was used, overground walking was more regular than treadmill. However, treadmill walking was found to be more regular when using r_{SD} for both ApEn and SampEn. For greatest relative consistency of step time data, it was best to use a constant r value and SampEn. When using entropy, several r values must be examined and reported to ensure that results are not an artifact of parameter choice.

1. Introduction

The use of entropy methods to calculate regularity or predictability in a time series has vastly increased over the last twenty years. Claude Shannon was the first scientist to introduce an algorithm for calculating information entropy within a time series [1]. Modifications to the original Shannon entropy algorithm have been made, including approximate [2,3], sample [4,5], multiscale [6,7], increment [8], permutation [9], and multiscale permutation [10] entropy, just to name a few. The use of entropy analysis has been wide ranging from investigations in financial markets to weather patterns to thermodynamics to biology.

Each entropy algorithm requires parameter selection that may or may not affect the calculation of entropy [11,12]. The m parameter, is the number of data points that are to be compared, sometimes called the window or vector of data. The parameter r , is the set tolerance, sometimes called radius, that is utilized to determine if two vectors are considered similar. The last parameter, N , is the length of the entire data set.

The two most popular entropy algorithms utilized for human movement analysis are approximate entropy (ApEn) and sample entropy (SampEn). ApEn was originally developed as a method to measure regularity within a time series [2,3,13]. However, a limitation of ApEn is that it is biased toward regularity because it includes a self-match count to avoid taking the logarithm of zero [4,5]. Furthermore, it requires fixed parameters when comparing data [13] and lacks relative consistency [13], a term used to describe how stable the output of the algorithm is when input parameter selection is changed slightly. Our previous work has demonstrated that ApEn is more prone to inconsistent output as compared to SampEn when utilized for time series of 200 data points or less [12]. However, both algorithms were sensitive to certain combinations of parameters. When reporting entropy results, it is of highest priority that the results are not an artifact of parameter choice. In addition, due to the sensitivity of ApEn to parameter choice, comparisons between studies and data can only be done if the parameter choices are fixed. This is very difficult, as each data set requires careful selection of parameters based upon that unique set of data.

* Corresponding author.

E-mail address: [jyentes@gmail.com](mailto: jyentes@gmail.com) (J.M. Yentes).

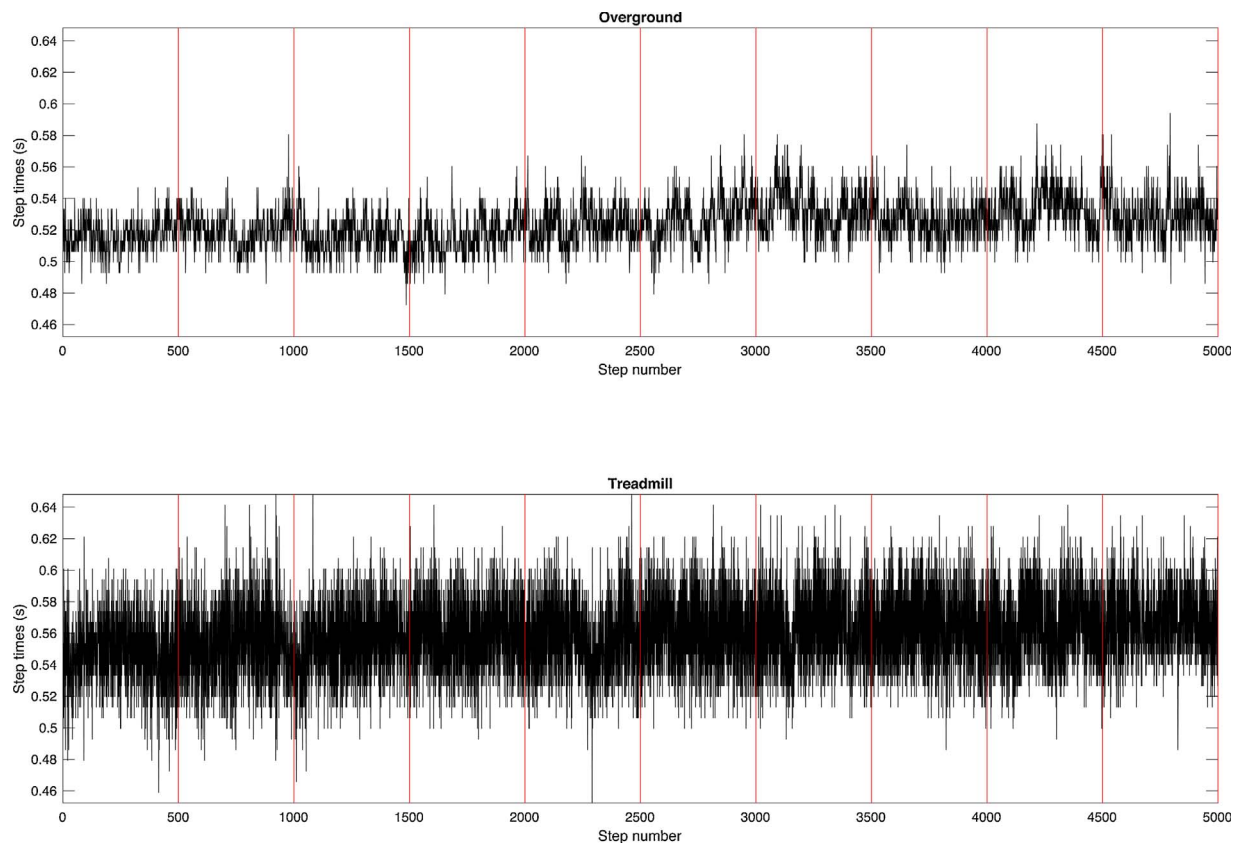


Fig. 1. Data from a representative subject for overground (top) and treadmill (bottom) step times.

SampEn was developed to overcome the limitations of ApEn [4,5]. The calculation of SampEn does not include a self-match and the logarithm of the sum of conditional probabilities is taken as compared to ApEn. As stated by Richman and Moorman [5], ApEn calculates probabilities “in a template-wise fashion” whereas, SampEn calculates “the negative logarithm of a probability associated with the time series as a whole” (pp H2042).

One of the most difficult parameters to select is that of the tolerance, r , value. If the tolerance level is large compared to differences in values of sequential points, the probability values used to calculate entropy will be high and vice versa. Therefore, selection of the tolerance has a crucial impact on the outcome. There have been many proposed methods to determine r , including using the standard deviation (SD) of the entire time series [2,13], standard error of the entropy values [4], and using fixed tolerance values [14,15].

The length of the data set is a parameter of concern for human movement scientists. When dealing with pathological populations, it is sometimes difficult to obtain long enough data sets in order to perform analysis. Data sets less than 200 points appear to be too short for entropy analysis, yet, it may take up to 2000 data points for stabilization of entropy values [12]. It is currently unclear as to how many data points are needed for reliable entropy analysis.

Another limitation of entropy analysis is the need for collection of uninterrupted data [2,3]. To collect uninterrupted data, continuous or discrete, and the amount of data required for analysis, many researchers have subjects walk on a treadmill versus walking overground. This allows the researcher to collect uninterrupted steps without the concern of space and/or equipment constraints. However, the treadmill could be considered a constraint as it limits speed fluctuations in walking that are normally present in overground walking. In addition, there are physiological and biomechanical differences between overground and treadmill walking [16–25]. Nonlinear measures have shown conflicting results regarding the difference in the structure of

variability between treadmill and overground walking [26–30].

Thus, the purpose of this research was to determine the effect of changing parameter values on entropy calculations for long gait data sets (i.e., step time) using two different modes of walking. In addition, to understand the effect of changing parameters on entropy calculations, an examination of tolerance, r , was completed. It was hypothesized that SampEn would maintain relative consistency across all data lengths and be resistant to changes in parameter values.

2. Materials and methods

Twenty-one subjects participated in this research study. Foot switch data was collected from subjects, but subjects whose foot switch data contained any signal dropout were excluded. Therefore, 14 subjects' data from the original cohort were included in analysis (7 males; 24.9 ± 4.2 years; 1.71 ± 0.12 m; 69.3 ± 16.8 kg). All participants were in excellent health, had no conditions that would inhibit their ability to walk for one hour, and reported physical activity at or above the currently recommended level [31]. The University's Institutional Review Board approved all procedures and subjects provided consent prior to participation.

Participants were instrumented with force-sensitive resistors on each heel and second metatarsal head (Trigno 4-channel footswitch sensor, Delsys Inc., Natick, MA; 148 Hz). The foot switches were secured to the skin of the foot using tape before the participant placed their own sock and shoe on their foot. Cables were then secured to the top of the participants' shoes with tape. Data streamed wirelessly to a data acquisition unit (Delsys Inc., Natick, MA).

During visit one, participants walked for one-hour at a self-selected speed around an indoor track (201.2 m/loop). Speed was not recorded or monitored. Study personnel followed them within five meters to keep the wireless unit within signal range. Participants returned for a second visit and were asked to walk for one-hour at a self-selected speed

Download English Version:

<https://daneshyari.com/en/article/8798646>

Download Persian Version:

<https://daneshyari.com/article/8798646>

[Daneshyari.com](https://daneshyari.com)