Brain and Cognition 123 (2018) 126–135

Contents lists available at ScienceDirect

Brain and Cognition

journal homepage: www.elsevier.com/locate/b&c

A practical comparison of algorithms for the measurement of multiscale entropy in neural time series data

Karl Kuntzelman[a,b], L. Jack Rhodes[a], Lillian N. Harrington[a], Vladimir Miskovic[a,*]

[a] Department of Psychology, State University of New York at Binghamton, USA
[b] Center for Brain, Biology and Behavior, University of Nebraska-Lincoln, USA

ABSTRACT

There is a broad family of statistical methods for capturing time series regularity, with increasingly widespread adoption by the neuroscientific community. A common feature of these methods is that they permit investigators to quantify the entropy of brain signals – an index of unpredictability/complexity. Despite the proliferation of algorithms for computing entropy from neural time series data there is scant evidence concerning their relative stability and efficiency. Here we evaluated several different algorithmic implementations (sample, fuzzy, dispersion and permutation) of multiscale entropy in terms of their stability across sessions, internal consistency and computational speed, accuracy and precision using a combination of electroencephalogram (EEG) and synthetic $1/f$ noise signals. Overall, we report fair to excellent internal consistency and longitudinal stability over a one-week period for the majority of entropy estimates, with several caveats. Computational timing estimates suggest distinct advantages for dispersion and permutation entropy over other entropy estimates. Considered alongside the psychometric evidence, we suggest several ways in which researchers can maximize computational resources (without sacrificing reliability), especially when working with high-density M/EEG data or multivoxel BOLD time series signals.

1. Introduction

Recordings of population-level brain activity exhibit frequent state transitions and substantial moment-to-moment variability across a broad array of measurement modalities (e.g., electrocorticography [ECoG], electroencephalography [EEG], magnetoencephalography [MEG], functional magnetic resonance imaging [fMRI]) (Deco, Jirsa, & McIntosh, 2013; Friston, 1997; Garrett et al., 2013). In conventional terms, intraindividual brain signal variability has often been treated as a nuisance factor arising from exogenous or endogenous sources ("brain noise") that investigators attempt to minimize by using estimates of central tendency (e.g., averaged amplitudes of brain signals). An alternative framework, with increasing theoretical and empirical support, focuses on modeling variability as a means of facilitating the exploration of a broader manifold of functional network states (Deco et al., 2013; Deco & Corbetta, 2011; Tognoli & Kelso, 2014). Viewed from this perspective, it is not only that nervous systems have evolved to be able to tolerate noise, but under some circumstances the nervous system may actively exploit noise to optimize adaptability in uncertain environments (Faisal, Selen, & Wolpert, 2008; McDonnell & Ward, 2011; Sejdić & Lipsitz, 2013). By comparison, system pathology can be conceptualized as a reduction in degrees of freedom, generating activity that is more regular over numerous time scales as suggested by the complexity loss theory of disease and aging (Pincus & Goldberger, 1994; Yang & Tsai, 2013). Recent findings demonstrate that this loss of large-scale brain signal variability in cognitive aging is partially explained by reductions of dopaminergic neurotransmission (Garrett et al., 2015; Guitart-Masip et al., 2016). Quantifying the temporal structure of variability that is inherent in a system's output can therefore provide a useful estimate concerning the size of its dynamical repertoire (Garrett et al., 2017; McDonough & Nishira, 2014; Mišić, Vakorin, Paus, & McIntosh, 2011).

Practical advances in information theory and signal processing have enabled neuroscientists to quantify the diversity of patterns present in physiological signals using a metric of uncertainty known as entropy (Peng, Costa, & Goldberger, 2009; Richman & Moorman, 2000; Vakorin & McIntosh, 2012). In this context, entropy captures the unpredictability of time series data – patterns with low predictability are assigned high entropy, while ordered, regular signals (e.g., pure sine waves) contain very little entropy. One consequence of quantifying the temporal complexity of neural time series data in this manner is that the procedure always assigns higher entropic values to signals with greater

* Corresponding author at: Department of Psychology, State University of New York at Binghamton, Binghamton, NY 13902-6000, USA.
E-mail address: miskovic@binghamton.edu (V. Miskovic).

https://doi.org/10.1016/j.bandc.2018.03.010
Received 14 February 2017; Received in revised form 13 March 2018; Accepted 13 March 2018

randomness. A hallmark feature of brains as complex systems, however, is that they operate in a middle zone between the extremes of order and randomness, reflecting dynamics that are constrained by a structural backbone with a mixture of deterministic and stochastic elements (Pedersen, Omidvarnia, Walz, Zalesky, & Jackson, 2017; Sporns, 2013; Tononi, Edelman, & Sporns, 1998). The major advantage of multiscale measures of brain signal variability is that they can reveal system dynamics across a range of time scales, in contrast to conventional estimates of sample entropy at a single timescale factor or related techniques (e.g., PCA dimensionality).

To capture the fact that information content in physiological signals is expressed over multiple time scales, Costa et al. (2002, 2005) proposed a multiscale extension of entropy that is sensitive for long-range correlation structure in time series data and assigns higher temporal complexity estimates for signals that are neither completely ordered nor completely random. Application of multiscale methods produces a profile of system dynamics as revealed in entropy curves extending over a range of transient time scales. Since their introduction, multiscale entropy analyses have found growing popularity within the neuroscientific community with application to signals collected from various recording modalities and subject populations (see Garrett et al., 2013, for a review). An advantage of multiscale entropy over more conventional power spectral density (PSD) analyses is that the latter reflect only the linear stochastic properties of signals while the former permit investigators to also index non-linear deterministic correlations present in neural time series (Courtiol et al., 2016; Vakorin & McIntosh, 2012).

Within the past few years, a number of novel entropy-based measures of time series variability have appeared in the literature: fuzzy entropy (Azami & Escudero, 2017), dispersion entropy (Rostaghi & Azami, 2016) and permutation entropy (Bandt & Pompe, 2002), along with their multiscale versions, are a few of the commonly used metrics. Each of these algorithms differs in its computational details (for a comprehensive technical background, readers are referred to the original sources) and purports to provide some refinement over conventional measures of sample entropy, either in terms of improved processing efficiency, accuracy, precision or some combination of factors. It is important to note that, although all of these various measures share the goal of quantifying time series irregularity/unpredictability, their use reflects different assumptions and they may have different properties in applied settings – for instance, a specific quantitative index of entropy might be more or less sensitive to non-linearities present in brain signals.

Given the high throughput nature of large-scale neuronal recordings, establishing benchmark data on relative performance across several commonly used multiscale entropy metrics could provide practical guidance to researchers looking to optimize limited computing resources. The computational complexity of the traditional sample entropy algorithm ($O(N^2)$),[1] for example, is substantially higher when compared to permutation or dispersion entropy ($O(N)$). Such considerations have non-trivial consequences for informing data analysis strategies – at 1 kHz acquisition rates, one minute of continuous high-density EEG data can contain up to $> 1.53 \times 10^6$ sample points, an amount that could quickly impose a resource allocation bottleneck in a typical laboratory. However, computational burden is not the only relevant factor to consider when designing a sensible data processing pipeline, especially if distinct metrics exhibit markedly divergent results when applied to neural time series data, or if one has unacceptable psychometric properties such as poor stability, internal consistency or low precision.

Despite the growing popularity of brain signal variability measures in the study of learning and memory (Heisz, Shedden, & McIntosh, 2012, 2014; Lafontaine et al., 2016), development/aging (Heisz, Gould, & McIntosh, 2015; Lippé, Kovacevic, & McIntosh, 2009; McIntosh, Kovacevic, & Itier, 2008, 2014; Mišić, Mills, Taylor, & McIntosh, 2010; Miskovic, Owens, Kuntzelman, & Gibb, 2016; Takahashi et al., 2009; Vakorin, Lippé, & McIntosh, 2011; Wang, McIntosh, Kovacevic, Karachalios, & Protzner, 2016; Yang et al., 2013), neuropsychiatric pathology (Bosl, Tierney, Tager-Flusberg, & Nelson, 2011; Catarino, Churches, Baron-Cohen, Andrade, & Ring, 2011; Escudero, Abasalo, Hornero, Espino, & Lopez, 2006; Mišić, Dunkley et al., 2016; Takahashi et al., 2016; Yang et al., 2015), neurostimulation (Farzan, Pascual-Leone, Schmahmann, & Halko, 2016; Okazaki et al., 2013) and resting state *f*MRI networks (McDonough & Nishira, 2014; Smith, Yan, & Wang, 2014; Yang et al., 2014), very little is known about the stability of entropy derived from neural time series recordings.

Our goal here was to provide benchmark information that can guide the implementation of reliable and efficient data processing strategies for investigators who wish to quantify the multiscale variability of brain signals. We used a combination of empirical scalp EEG recordings and synthetic EEG-like time series signals to compare and contrast several commonly used metrics of multiscale entropy along several different desiderata: reliability (one-week stability), internal consistency, computational speed, accuracy, and precision. Ultimately, a researcher's decision to use one or another of these metrics ought to be guided by the nature of the research question and ideally convergence across metrics will be sought – here, we simply wished to provide some relevant information that might inform the design of data analytic strategies. Our analyses focused on recordings of intrinsic brain activity, which is likely to provide a more stringent evaluation of metric stability, since we are capturing spontaneous cortical dynamics that are unconstrained by external inputs or tasks. In addition, the recording of ongoing, endogenous brain signals allowed us to evaluate the stability of entropy estimates over a somewhat wider range of temporal scales than would be possible with shorter, event- or response-locked segments of neural time series.

2. Method

2.1. Participants

A total of eighteen undergraduate students at SUNY Binghamton (ages 18–22, mean = 18.93, SD = 1.28, 11 female) participated in at least the first of two EEG recording sessions separated by exactly one week, such that an individual's session began at the same time on each day in order to control for possible circadian fluctuations (Ly et al., 2016). All recordings took place roughly between the hours of 12:00 and 4:00P.M. One participant did not return for the second session and data from two participants' second sessions were not properly recorded, leaving a total of fifteen participants who provided usable data at both time points. The participants included here were part of a larger experimental study, the results of which are not further described in this paper. The resting state recordings that are the focus of analysis here preceded the experimental procedures.

2.1.1. Procedures and electrophysiological recording

Participants were greeted and informed about the experimental procedures. Following the completion of a written consent and preliminary paperwork associated with the behavioral experiment, participants were ushered into a dimly lit EEG recording room equipped with a white noise generating sound screen. Nasion-inion and preauricular anatomical measurements were made to locate and mark each individual's vertex site. Participants were then fitted with a 128 electrode (plus Cz reference) Electrical Geodesics, Inc. (EGI) HydroCel Geodesic Sensor Net. The target electrode impedances were at or below 60 kOhm at recording onset. Each session consisted of a four-minute resting period with eyes closed. Participants were seated in front of a computer monitor and instructed to close their eyes when a fixation

[1] This notation (referred to as either Landau or Big O notation) refers to the way the time and/or memory required to perform a calculation scales with the size of the input.