Combining complexity measures of EEG data: multiplying measures reveal previously hidden information

Thomas F Burns

Abstract

Many studies have noted significant differences among human EEG results when participants or patients are exposed to different stimuli, undertaking different tasks, or being affected by conditions such as epilepsy or Alzheimer's disease. Such studies often use only one or two measures of complexity and do not regularly justify their choice of measure beyond the fact that it has been used in previous studies. If more measures were added to such studies, however, more complete information might be found about these reported differences. Such information might be useful in confirming the existence or extent of such differences, or in understanding their physiological bases. In this study I analysed publically-available EEG data using a range of complexity measures to determine how well the measures correlated with one another. The complexity measures did not all significantly correlate, suggesting that different measures were measuring unique features of the EEG signals and thus revealing information which other measures were unable to detect. Therefore, the results from this analysis suggests that combinations of complexity measures reveal unique information which is in addition to the information captured by other measures of complexity in EEG data. For this reason, researchers using individual complexity measures for EEG data should consider using combinations of measures to more completely account for any differences they observe and to ensure the robustness of any relationships identified.

Keywords

electroencephalograph, EEG, complexity, complexity measure, sample entropy, permutation entropy, algorithmic complexity, Lemel-Ziv complexity, fractal dimension, Higuchi complexity, spectral flatness, Weiner entropy, spectral structure index, spectral structure variability, information theory, chaos theory

Introduction

Electroencephalography (EEG) is a common, relatively non-invasion research and diagnostic tool. Its one-dimensional signals from localised peripheral regions on the head make it attractive for its simplistic fidelity and has allowed high clinical and basic research throughput. When it comes to interpreting EEG data, investigators have a wide range of analytical tools at their disposal (Delorme & Makeig, 2004; Dauwels *et al*. 2010) and in recent years have explored a number of novel relationships between measures of complexity (Susmáková & Krakovská, 2008; Cao & Slobounov, 2011; Dauwels *et al*., 2011; Weiss *et al*., 2011; Jing *et al*. 2014; Sitt *et al*. 2014). Studies which have included complexity measures, however, do not regularly include more than one or two such measures. For example, Dauwels et al. (2011) include the Lempel-Ziv complexity measure (Lempel &

Ziv, 1976) - an algorithmic-based measure - and regularity measures, but ignore potential chaotic and fractal measures. This is not to suggest that the LZ complexity measure or that regularity measures are meaningless, nor that chaotic and fractal measures are more or less important than other measures of complexity, but that all may be measuring different features. Thus, for a more complete and robust picture of any relationships found for one complexity measure in EEG data, it might be useful for investigators to include other measures in their analyses.

This study therefore aims to determine whether different measures of complexity of EEG signals correlate, and (if so) to what degrees. To do this, a small battery of complexity measures were computed for publicly-available normative data and subsequently analysed for correlations. If some measures were found not to significantly correlate or correlate fully, this would suggest that these measures are detecting unique information which might otherwise have remained hidden to investigators who were computing only a single complexity measure from their data.

Methods

1100 EEG recordings of 1-sec duration from 13 healthy control subjects undergoing a basic psychophysics task were obtained from a publicly-available database created by Begleiter (1996) of the Neurodynamics Laboratory, State University of New York Health Center, Brooklyn, United States. Each recording had 64 channels and was sampled at 256 Hz (3.9-msec epoch). The following complexity measures were calculated for each recording: Lempel-Ziv algorithmic complexity (LZ) (Lempel & Ziv, 1976), fractal dimension estimation (FD) (Higuchi, 1988), permutation entropy (PE) (Bandt & Pompe, 2002), Wiener entropy (WE) (Wiener, 1954), and spectral structure variability (SSV) (Singh, 2011). These measures were chosen on the basis of their broad representation of different conceptions of 'complexity', including informational theoretic, chaotic/fractal, and computational informatic approaches. Many more measures exist than these, however as the principle aim of this paper was to determine if differences exist at all, any differences detected in this small cross-section of measures would sufficiently illustrate this. Results from the complexity measures were analysed by linear regression and significance for relationships between pairs of measures was calculated.

Results

Of the 10 pairs of measures, eight pairs exhibited highly significant ($p<0.0001$) correlations while two pairs - (i) PE and FD, (ii) WE and LZ - did not significantly correlate (Tables 1 and 2). High degrees of spread were noted among all correlations.

These relationships were visualised using scatter plots (Figures 1 and 2) to help determine if any of these relationships may be non-linear. Two such relationships - (i) LZ and FD, (ii) SSV and FD - appeared to follow a binomial trend (Figure 3), and binomial regression improved these relationships greatly.

Table 1. Pearson (r) correlation matrix for each pair of complexity measures computed for normative EEG recordings.

	FD	LZ	WE	PE	SSV
FD	-	0.5402	0.4155	-0.0255	0.6517
LZ	0.5402	-	-0.0472	-0.1273	0.5983
WE	0.4155	0.4155	-	0.3469	0.5977
PE	-0.0255	-0.1273	0.3469	-	0.1672
SSV	0.6517	0.5983	0.5977	0.1672	-

Table 2. Significance (p) of correlations for each pair of complexity measures computed for normative EEG recordings.

	FD	LZ	WE	PE	SSV
FD	-	<0.0001	<0.0001	0.3990	<0.0001
LZ	<0.0001	-	0.1174	<0.0001	<0.0001
WE	<0.0001	0.1174	-	<0.0001	<0.0001
PE	0.3990	<0.0001	<0.0001	-	<0.0001
SSV	<0.0001	<0.0001	<0.0001	<0.0001	-

Discussion

Some - but not all - measures of complexity of EEG signals correlate, and to varying degrees. Of the many complexity measures available to researchers investigating EEG data, overreliance or overconfidence in any single measure therefore seems misplaced. As research groups who have attempted to classify or predict sleep stages or conscious states from EEG data have implicitly noted (Susmáková & Krakovská, 2008; Weiss *et al*., 2011; Sitt *et al*., 2014), no individual measurement can reliably predict all possibly relevant physiology. Instead, combinations of measures are needed. In the same way, no individual measurement of complexity can reliably predict all possibly relevant complexity.

In part, the results from this study might reflect on a more generalised ambiguity of the concept of 'complexity'. Who is to say, after all, that more is revealed about 'complexity' by FD than LZ? It seems that it cannot be said that either reveal any more or less, since both ultimately treat complexity in a different way. Perhaps this only further reiterates the conclusion we can reach from this study: by multiplying measures we can reveal information which was previously hidden or unknown to us.

It would be interesting to analyse previously-noted complexity differences - between, for example, patients with and without Alzheimer's disease (Dauwels *et al.*, 2010) - to determine if these differences were all measuring the same difference. The current study suggests they may not have been. And, if not, perhaps more can be learned from the available data; it could even be possible that there exists entirely separate complexity dimensions, along which patients progress at different rates. Such information could contain physiological, clinical, or other significance.

Acknowledgements

Thank you to MATLAB community members for their assistance with some of the functions used for calculating complexity measures in this paper.

References

Bandt, C. and Pompe, B., 'Permutation entropy: a natural complexity measure for time series', *Physical Review Letters*, vol. 88, i. 17, 2002, e174102.

Begleiter, H. (1996) EEG data set of healthy and alcoholic adults completing psychophysics tasks [data-set]. Machine Learning Repository, University of California, Irvine [distributor] <http://archive.ics.uci.edu/ml/datasets/EEG+Database>

Cao, C., and Slobounov S., 'Application of a novel measure of EEG non-stationarity as 'Shannon-entropy of the peak frequency shifting' for detecting residual abnormalities in concussed individuals', *Clinical Neurophysiology*, vol. 122, i. 7, 2011, pp. 1314-21.

Dauwels, J., Srinivasan, K., Ramasubba Reddy, M., Musha, T., Vialatte, F.B., Latchoumane, C., Jeong, J., and Cichocki, A., 'Slowing and Loss of Complexity in Alzheimer's EEG: Two Sides of the Same Coin?', *International Journal of Alzheimer's Disease*, vol. 2011, 2011, e539621.

Delorme, A., and Makeig, S., 'EEGLAB: an open source toolbox for analysis of single-trial EEG dynamics including independent component analysis', *Journal of Neuroscience Methods*, vol. 15, i. 134, no. 1, 2004, pp. 9-21.

Dauwels, J., Vialatte, F., and Cichocki, A., 'Diagnosis of Alzheimer's Disease from EEG Signals: Where Are We Standing?', *Current Alzheimer Research*, vol. 7, i. 6, 2010, pp. 487-505.

Higuchi, T., 'Approach to an irregular time series on the basis of the fractal theory', *Physica D: Nonlinear Phenomena*, vol. 31, i. 2, 1988, pp. 277-83.

Jing, L., Jiaqing, Y., Xianzeng, L., and Gaoxiang, O., 'Using Permutation Entropy to Measure the Changes in EEG Signals During Absence Seizures', *Entropy*, vol. 16, i. 6, 2014, pp. 3049-61.

Lempel, A., and Ziv, J., 'On the complexity of finite sequences', *IEEE Transactions on Information Theory*, vol. IT-22, i. 1, 1976, pp. 75-81.

Singh, N.C., 'Measuring the 'complexity' of sound', *Pramana - Journal of Physics*, vol. 77, i. 5, 2011, pp. 811-6.

Sitt, J.D., King, J.R., El Karoui, I., Rohaut, B., Faugeras, F., Gramfort, A., Cohen, L., Sigman, M., Dehaene, S., and Naccache, L., 'Large scale screening of neural signatures of consciousness in patients in a vegetative or minimally conscious state', *Brain*, vol. 137, pt. 8, 2014, pp. 2258-70.

Susmáková, K., and Krakovská, A., 'Discrimination ability of individual measures used in sleep stages classification', *Artificial Intelligence in Medicine,* vol. 44, i. 3, 2008, pp. 261-77.

Weiss, B., Clemens, Z., Bódizs, R., and Halász, P., 'Comparison of fractal and power spectral EEG features: effects of topography and sleep stages', *Brain Research Bulletin*, vol. 5, i. 84, no. 6, 2011, pp. 359-75.

Wiener, N., *In The human use of human beings: Cybernetics and society*, Boston, Houghton Mifflin, 1954, pp. 15-27.

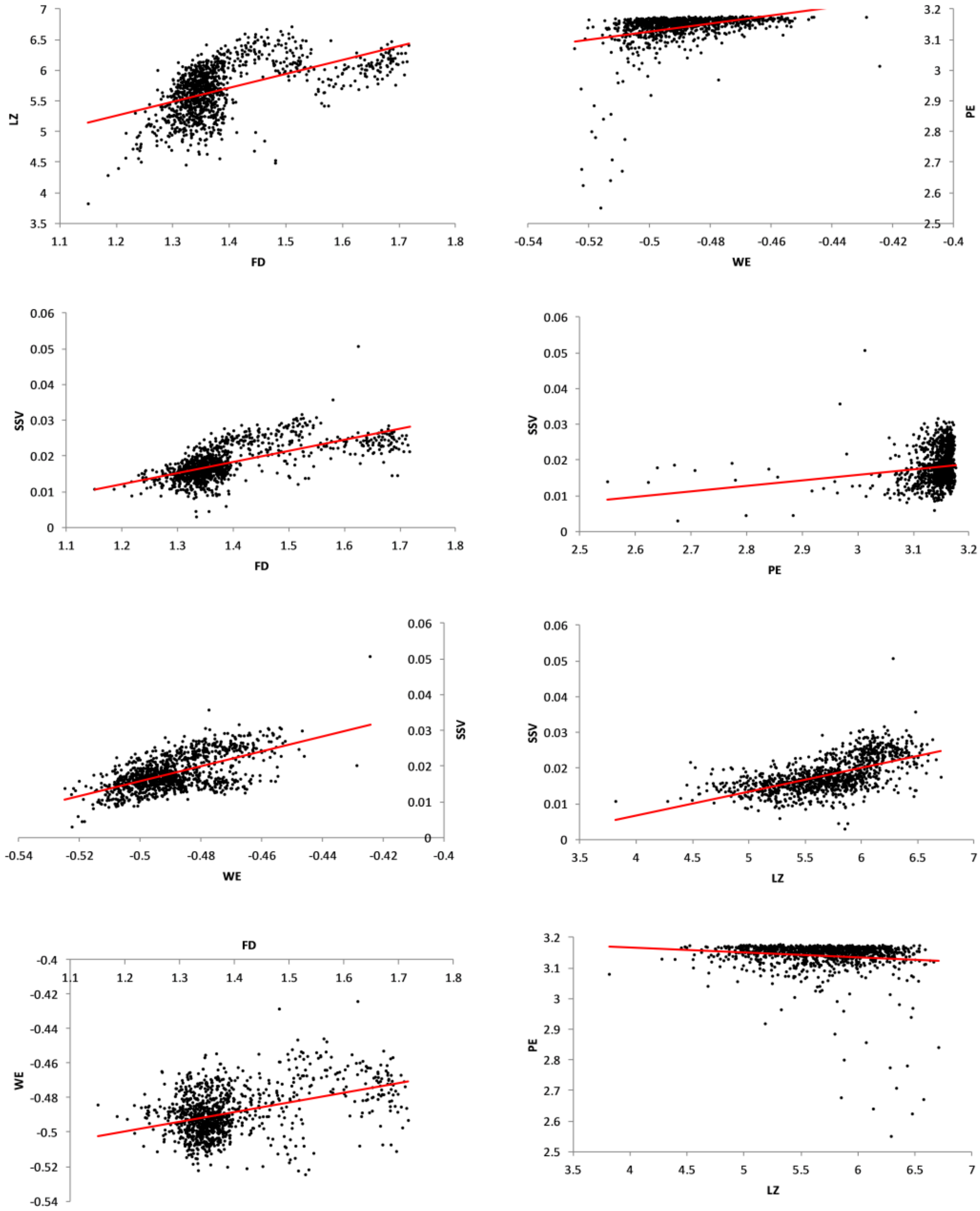

Figure 1. Scatter plots with linear trendlines for pairs of significantly-correlated complexity measures. Eight pairs of complexity measures of the EEG signals had a significant (p<0.0001) correlation. Although the relationships are significant, high degrees of spread are noticeable and some of the relationships may have non-linear components.

EEG = electroencephalogram; LZ = Lempel-Ziv algorithmic complexity; FD = fractal dimension estimate (Higuchi method); PE = permutation entropy; SSV = spectral structure variability; WE = Wiener entropy (also known as spectral flatness)

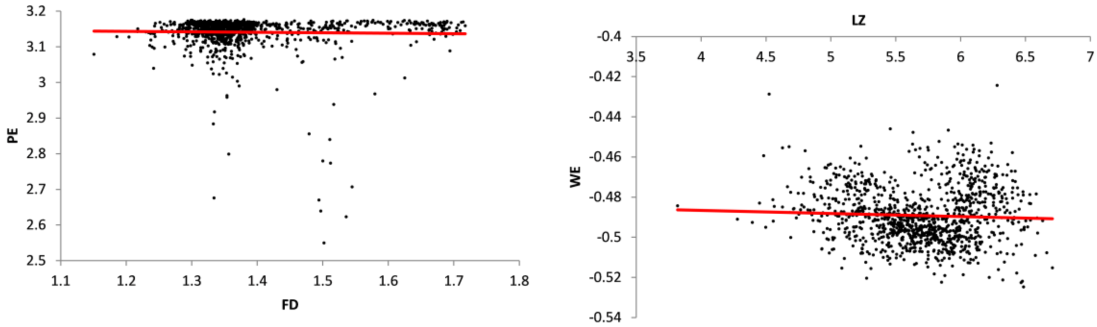

Figure 2. Scatter plots with linear trendlines for two pairs of insignificant, uncorrelated complexity measures. Two pairs of complexity measures of the EEG signals were insignificant and uncorrelated - PE & FD (r=-0.0255, p=0.3990) and WE & LZ (r=-0.0472, p=0.1174). There appears to be no non-linear components or any evidence of a clear relationship between these pairs of measures.

EEG = electroencephalogram; LZ = Lempel-Ziv algorithmic complexity; FD = fractal dimension estimate (Higuchi method); PE = permutation entropy; WE = Wiener entropy (also known as spectral flatness)

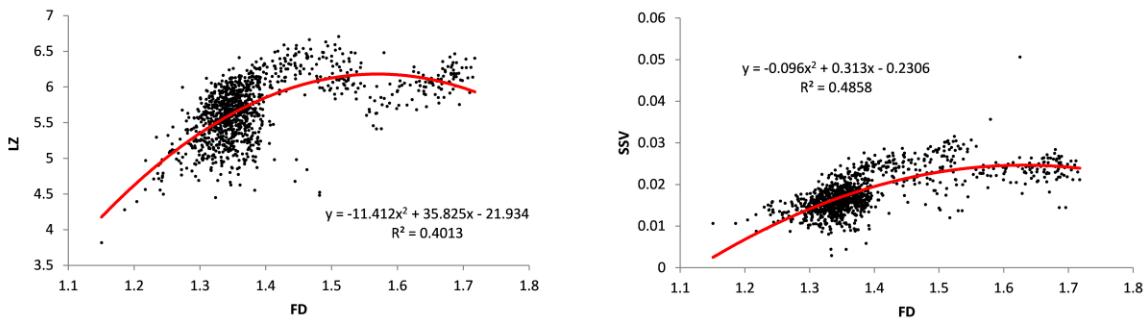

Figure 3. Scatter plots with binomial regression lines for potential non-linearly-related pairs of complexity measures. Two pairs of complexity measures of the EEG signals appeared to have noticeable non-linear relationships: (i) LZ and FD; and (ii) SSV and FD. Although these binomial relationships were - like their linear relationships - significant ($p<0.0001$), the binomial regressions produced less spread and appear to be truer representations of the relationships.

EEG = electroencephalogram; LZ = Lempel-Ziv algorithmic complexity; FD = fractal dimension estimate (Higuchi method); SSV = spectral structure variability

PeerJ PrePrints | https://dx.doi.org/10.7287/peerj.preprints.1121v1 | CC-BY 4.0 Open Access | rec: 26 May 2015, publ: 26 May 2015

[supplementary file]

Data spreadsheet 1. Calculated complexity measures for 1100 EEG recordings. The following data are the results from MATLAB functions which calculated complexity measures for each EEG recording.

ID = identification code as per Begleiter (1996); LZ = Lempel-Ziv algorithmic complexity; FD = fractal dimension estimate (Higuchi method); PE = permutation entropy; WE = Wiener entropy (also known as spectral flatness)

PeerJ PrePrints | https://dx.doi.org/10.7287/peerj.preprints.1121v1 | CC-BY 4.0 Open Access | rec: 26 May 2015, publ: 26 May 2015