

Editorial

Entropy Measures for Data Analysis: Theory, Algorithms and Applications

Karsten Keller 

Institut für Mathematik, Universität zu Lübeck, D-23562 Lübeck, Germany; keller@math.uni-luebeck.de

Received: 11 September 2019; Accepted: 22 September 2019; Published: 25 September 2019



Keywords: data analysis; entropy; entropy measures; automatic learning

Entropies and entropy-like quantities are playing an increasing role in modern non-linear data analysis and beyond. Fields that benefit from their application reach from diagnostics in physiology, for instance, electroencephalography (EEG), magnetoencephalography (MEG) and electrocardiography (ECG), to econophysics and engineering. During the last few years, classical concepts as approximate entropy and sample entropy have been supplemented by new entropy measures, like permutation entropy and its variants. Recent developments are focused on multidimensional generalizations of the concepts with a special emphasis on the quantification of coupling and the similarity between time series and system components behind them. One of the main future challenges in the field include finding a better understanding of the nature of the various entropy measures and their relationships, with the aim of adequate application including good parameter choices. The utilization of entropy measures as features in automatic learning and their application to large and complex data for such tasks as classification, discrimination and finding structural changes requires fast and well-founded algorithms. This issue is facing a different aspect of the use of entropy measures for data analysis in a wide sense, including those described.

Papers 1–3 discuss the problem of parameter choice mentioned and aspects related to it. Ahmadi et al. [1] investigate the sensitivity of sample entropy with respect to different parameters like, for example, tolerance size and sampling rate for gait data. Cuesta-Frau et al. [2] study parameter choice for permutation entropy, particularly embedding dimension and time series length, in the context of a lot of synthetic and real data sets. Here special emphasis is put on practical aspects of data analysis. In particular, the authors point out that in many cases permutation entropy can be used for shorter data sets than reported by other authors. In a certain sense complementary to the paper of Cuesta-Frau et al. [2], Piek and Keller [3] investigate parameter choice for permutation entropy and, more generally, ordinal pattern-based entropies from a computational and theoretical viewpoint. Fast algorithms are presented, possibilities and limits of the estimation of the Kolmogorov–Sinai entropy are discussed. A further aspect of the paper is the generation of artificial data for testing ordinal pattern methods.

The main objective of papers 4 and 5 is entropy-based feature extraction. Lu et al. [4] utilize approximate entropy, sample entropy, composite multiscale entropy and fuzzy entropy for identifying auditory object-specific attention from single-trial EEG signals by support vector machine (SVM)-based learning. For circuit fault diagnosis, He et al. [5] propose a new feature extraction method, which is mainly based on a measure called joint cross-wavelet singular entropy and a special dimension reduction technique. The obtained features are entered into a support vector machine classifier in order to locate faults. Besides feature extraction, direct applications of entropy for automatic learning is also addressed in this issue. Bukovsky et al. [6] discuss and further develop the recently introduced concept of learning entropy (LE) as a learning-based information measure, which is targeted at real-time novelty detection based on unusual learning efforts. For assessing the quality of data transformations

in machine learning, Valverde-Albacete et al. [7] introduce an information-theoretic tool. They analyze performance of the tool for different types of data transformation, among them principal component analysis and independent component analysis.

Papers 8–10 are devoted to the aspect of coupling and similarity analysis. For studying the Chinese stock market around the 2015 crash, Wang and Hui [8] utilize effective transfer entropy (ETE), which is an adaptation of transfer entropy to limited and noisy data. From this base, they discuss and compare dependencies of 10 Chinese stock sectors during four characteristic time periods near the crash. In [9], Craciunescu et al. introduce a new measure for describing coupling in interconnected dynamical systems and test it for different system interactions. Besides such in model systems, real-life system interactions like between the El Niño Southern Oscillation, the Indian Ocean Dipole, and influenza pandemic occurrence are considered. Here, coupling strength is quantified by entropies of adjacency matrices associated to networks constructed. Wang et al. [10] use entropy-based similarity and synchronization indices for relating postural stability and lower-limb muscle activity. Their study is based on two types of signals, one measuring the centre of pressure (COP) in dependence on time and one being an electromyogram (EMG). The authors show high correlation of COP and the low frequency EMG and that the cheaper COP contains much information on the EMG.

The other four papers touch further interesting aspects of entropy measure use. Selvachandran et al. [11] consider complex vague soft sets (CCVS), defined as a hybrid model of vague soft sets and complex fuzzy sets, which is, for example, useful for the description of images. Some distance and entropy measures for CCVSs are axiomatically defined and relations between them are investigated. The work [12] of Pan et al. focuses on Dempster–Shafer evidence theory, which can be considered as a generalization of probability theory. A new belief entropy, measuring uncertainty in this framework, and its performance are discussed on the base of numerical experiments. García-Gutiérrez et al. [13] introduce a new model for the particle size distribution (PSD) of granular media, which relates two models known for a long time. For this purpose, a differential equation involving the information entropy is used. The interesting point is that experimental data can be considered as an initial condition for simulating a PSD. Last but not least, Liu et al. [14] demonstrate that entropy methods also can be helpful in solving nonlinear and multimodal optimization problems. They develop an algorithm based on the firefly algorithm and the cross-entropy method and report its good performance, especially powerful global search capacity precision and robustness for numerical optimization problems.

Acknowledgments: I express my thanks to the authors of the above contributions, and to the journal Entropy and MDPI for their support during this work.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Ahmadi, S.; Sepehri, N.; Wu, C.; Szturm, T. Sample Entropy of Human Gait Center of Pressure Displacement: A Systematic Methodological Analysis. *Entropy* **2018**, *20*, 579. [[CrossRef](#)]
2. Cuesta-Frau, D.; Murillo-Escobar, J.P.; Orrego, D.A.; Delgado-Trejos, E. Embedded Dimension and Time Series Length. Practical Influence on Permutation Entropy and Its Applications. *Entropy* **2019**, *21*, 385. [[CrossRef](#)]
3. Piek, A.B.; Stolz, I.; Keller, K. Algorithmics, Possibilities and Limits of Ordinal Pattern Based Entropies. *Entropy* **2019**, *21*, 547. [[CrossRef](#)]
4. Lu, Y.; Wang, M.; Zhang, Q.; Han, Y. Identification of Auditory Object-Specific Attention from Single-Trial Electroencephalogram Signals via Entropy Measures and Machine Learning. *Entropy* **2018**, *20*, 386. [[CrossRef](#)]
5. He, W.; He, Y.; Li, B.; Zhang, C. Analog Circuit Fault Diagnosis via Joint Cross-Wavelet Singular Entropy and Parametric t-SNE. *Entropy* **2018**, *20*, 604. [[CrossRef](#)]
6. Bukovsky, I.; Kinsner, W.; Homma, N. Learning Entropy as a Learning-Based Information Concept. *Entropy* **2019**, *21*, 166. [[CrossRef](#)]
7. Valverde-Albacete, F.J.; Peláez-Moreno, C. Assessing Information Transmission in Data Transformations with the Channel Multivariate Entropy Triangle. *Entropy* **2018**, *20*, 498. [[CrossRef](#)]

8. Wang, X.; Hui, X. Cross-Sectoral Information Transfer in the Chinese Stock Market around Its Crash in 2015. *Entropy* **2018**, *20*, 663. [[CrossRef](#)]
9. Craciunescu, T.; Murari, A.; Gelfusa, M. Improving Entropy Estimates of Complex Network Topology for the Characterization of Coupling in Dynamical Systems. *Entropy* **2018**, *20*, 891. [[CrossRef](#)]
10. Wang, C.; Jiang, B.C.; Huang, P. The Relationship between Postural Stability and Lower-Limb Muscle Activity Using an Entropy-Based Similarity Index. *Entropy* **2018**, *20*, 320. [[CrossRef](#)]
11. Selvachandran, G.; Garg, H.; Quek, S.G. Vague Entropy Measure for Complex Vague Soft Sets. *Entropy* **2018**, *20*, 403. [[CrossRef](#)]
12. Pan, Q.; Zhou, D.; Tang, Y.; Li, X.; Huang, J. A Novel Belief Entropy for Measuring Uncertainty in Dempster-Shafer Evidence Theory Framework Based on Plausibility Transformation and Weighted Hartley Entropy. *Entropy* **2019**, *21*, 163. [[CrossRef](#)]
13. García-Gutiérrez, C.; Martín, M.Á.; Pachepsky, Y. On the Information Content of Coarse Data with Respect to the Particle Size Distribution of Complex Granular Media: Rationale Approach and Testing. *Entropy* **2019**, *21*, 601. [[CrossRef](#)]
14. Li, G.; Liu, P.; Le, C.; Zhou, B. A Novel Hybrid Meta-Heuristic Algorithm Based on the Cross-Entropy Method and Firefly Algorithm for Global Optimization. *Entropy* **2019**, *21*, 494. [[CrossRef](#)]



© 2019 by the author. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).