DUJASE Vol. 4(2) 117-124, 2018 (January & July)

Multiscale Complexity Analysis: A Novel Approach for Anomaly Detection in Multivariate Data

Mosabber Uddin Ahmed

Department of Electrical and Electronic Engineering, University of Dhaka, Dhaka-1000, Bangladesh

E-mail: mosabber.ahmed@du.ac.bd

Received on 13.08.18. Accepted for publication on 04.02.19

ABSTRACT

In this paper, a novel method is presented for anomaly detection in multivariate data. The proposed method is based on computing multivariate entropy of input data at multiple scales, via the MMSE method, a technique recently proposed for the dynamical complexity analysis of multivariate data. In the proposed methodology, the anomalous behaviour is assumed to be generated by a constrained system and thus is easily differentiated from the established normal behaviour, in accordance with the "complexity loss" hypothesis, traditionally used for physiological systems. Simulations are provided to demonstrate the effectiveness of the approach on real world data sets in terms of anomaly detection.

Keywords: Multivariate sample entropy (MSampEn), Multivariate multiscale entropy (MMSE), Multivariate system complexity, Multivariate embedding, Anomaly detection.

1. Introduction

Anomaly detection relates to the study of finding patterns in a given data set which deviate from the established normal behaviour. It has been used in different fields to detect and remove anomalous measurement that may arise from mechanical faults, human or instrumentation error, and changes in system behaviour etc. [1]. Isolating and characterizing anomalies is a prerequisite to improve the quality of the original data set, and aiding in the identification of system faults and frauds [2]. Anomaly detection techniques are relevant in a variety of domains, such as, fraud detection for credit cards, insurance or health care, intrusion detection for cyber-security [3], fault detection in safety critical systems, and military surveillance for enemy activities [1].

In literature, a number of anomaly detection techniques can be found; mainly, these can be classified into classification-based methods [4], nearest neighbour-based methods [5], clustering-based methods [6], [7], and statistical techniques [8], [9]. Moreover, methods based on information theoretic measures, and spectral estimates are also available [1], [10], [11]. Each of the above class of techniques makes some key assumptions to differentiate between normal and anomalous behaviours in input data. For instance, techniques employing information theoretic measures, such as, Kolomogorov complexity, entropy, and relative entropy, are based on the assumption that anomalies in data induce irregularities in the *information content* of input data. These techniques mainly involve dual optimization, attempting to minimize the "anomalous" subset size while concurrently maximizing the reduction in the complexity of the data set [1].

However, in techniques involving traditional entropy measures, such as, Kolomogorov complexity, and relative entropy etc, features related to the structure and the organization of patterns in input data over a range of time scales are not accounted for, as these measures are only computed over a one-step difference or a single time scale, that is, $H_{n+1} - H_n$ where H_n is the joint entropy for a time series with n variables. Costa *et al.* noticed this discrepancy and argued that the dynamics of a complex nonlinear system manifests in multiple inherent scales of the observed time series and, thus, entropy estimates calculated on a single scale are not sufficient descriptors for real world data. To that end, they proposed multiscale entropy (MSE) analysis method which aimed at quantifying the interdependence between entropy and scale. This was achieved by first extracting multiple temporal scales of input data using the so-called coarse graining method and sample entropy estimates were subsequently calculated for each scale separately [12], [13]. This facilitates the assessment of the dynamical complexity of a system, over a range of inherent scales residing in the data.

Recently, a multivariate extension of the MSE method known as the multivariate multiscale entropy (MMSE) technique [14], [15] has been proposed. MMSE method has shown to perform better than traditional information theoretic measures, since it operates on multiple scales of the signals and are thus able to extract information regarding inherent long range correlations in the data. In addition, MMSE can also quantify inter-channel correlations in multivariate data and is thus suited for signals containing multiple channels.

We propose to use the above multi-scale entropy measures for the anomaly detection in both univariate and multivariate real world data. Traditionally, the application of these techniques have been limited to physiological [15] and meteorological [14] data sets. However, they have tremendous potential in applications related to anomaly detection in input data which will be explored in this paper.

The basis behind using both univariate and multivariate techniques is that the (M)MSE measure of dynamical complexity has been shown to span a whole range of properties between the perfect regularity and total randomness: (M)MSE measures have smaller values for both deterministic (predictable) and uncorrelated random (unpredictable) signals, and comparatively large values for correlated (linear/nonlinear) stochastic processes. Moreover, in physiological data sets, (M)MSE is used along with the concept known as `complexity loss' hypothesis [16], which postulates that the complexity of a physiological or behavioural control system degrades with disease and aging.

Using similar argument, we extend the analysis for anomaly detection, hypothesizing that an anomaly in data would be 'fickle' in a sense that it will generally not reveal complex variability associated with long-range (fractal) correlations [16] due to its constrained nature, and is thus expected to yield lower complexity estimates over a range of inherent data scales. On the other hand, data without anomalies is expected to retain its inherent structures (or correlations) and will typically yield higher complexity estimates, allowing us to make a distinction between the two regimes, that is, normality and anomaly.

The organization of the paper is as follows: Section 2 illustrates the concepts of multivariate sample entropy and associated multivariate multiscale entropy method (MMSE). This is followed by a theoretical analysis of MMSE on multivariate correlated and uncorrelated noises in Section 3 & Section 4. The simulation results on real world multivariate data for the purpose of anomaly detection are given in Section 5; uterine EMG data, financial data and 3D motion sensor data are analyzed in that section and finally conclusion is drawn.

2. Multivariate Multiscale Entropy

The recently introduced multivariate multiscale entropy (MMSE) analysis has the following steps [14] [15]:

1) Temporal scales are defined by averaging a p-variate time series $\{x_{k,i}\}_{i=1}^{N}, k = 1,2,\dots,p$ over non-overlapping time segments of increasing length (coarse graining), where N is the number of samples in every channel. This way, for scale ϵ, a coarse grained multivariate time series is obtained as:

$$y_{k,j}^{\epsilon} = \frac{1}{\epsilon}\sum_{i=(j-1)\epsilon+1}^{j\epsilon} x_{k,i}, \tag{1}$$

where $1 \leq j \leq \frac{N}{\epsilon}$ and the channel index $k = 1,\dots,p$.

2) Multivariate sample entropy, $MSampEn$, is evaluated as in Algorithm 1 for each coarse-grained multivariate $y_{k,j}^{\epsilon}$, and then $MSampEn$ is plotted as a function of the scale factor ϵ

Algorithm 1: Multivariate sample entropy (MSampEn)

1: For a p-variate time series $\{x_{k,i}\}_{i=1}^{N}, k = 1,2,\dots,p$, observed through p measurement functions $h_k(y_i)$, the multivariate embedded reconstruction is based on a composite delay vector

$$X_m(i) = [x_{1,i}, x_{1,i+\tau_1},\dots,x_{1,i+(m_1-1)\tau_1}, x_{2,i}, x_{2,i+\tau_2}, \dots, x_{2,i+(m_2-1)\tau_2},\dots, x_{p,i}, x_{p,i+\tau_p},\dots,x_{p,i+(m_p-1)\tau_p}],$$

where $M = [m_1, m_2,\dots,m_p] \in \mathbb{R}^p$ is the embedding vector, $\tau = [\tau_1,\tau_2,\dots,\tau_p]$ is the time lag vector, and the composite delay vector $X_m(i) \in \mathbb{R}^m$, where $m = \sum_{k=1}^{p} m_k$.

2: Form $(N-\delta)$ composite delay vectors $X_m(i) \in \mathbb{R}^m$, where $i = 1,2,\dots,N-\delta$ and $\delta = \max\{M\} \times \max\{\tau\}$.

3: Define the distance between any two composite delay vectors $X_m(i)$ and $X_m(j)$ as the maximum norm, that is, $d[X_m(i), X_m(j)] = \max_{l=1,\dots,m}\{|x(i+l-1) - x(j+l-1)|\}$.

4: For a given composite delay vector $X_m(i)$ and a threshold r, count the number of instances P_i for which $d[X_m(i), X_m(j)] \leq r$, $j \neq i$, then calculate the frequency of occurrence, $B_i^m(r) = \frac{1}{N-\delta-1}P_i$, and define $B^m(r) = \frac{1}{N-\delta}\sum_{i=1}^{N-\delta} B_i^m(r)$.

5: Extend the dimensionality of the multivariate delay vector in (2) from m to $(m+1)$. This can be performed in p different ways, as from a space with embedding vector $M = [m_1, m_2,\dots,m_k,\dots,m_p]$, the system can evolve to any space for which the embedding vector is $[m_1, m_2,\dots,m_k + 1,\dots,m_p]$ (k=1, 2, ..., p). Thus, a total of $p \times (N-\delta)$ vectors $X_{m+1}(i)$ in $\mathbb{R}^{m+1}$ are obtained, where $X_{m+1}(i)$ denotes any embedded vector upon increasing the embedding dimension from m_k to (m_k+1) for a specific variable *k.*

6: For a given $X_{m+1}(i)$, calculate the number of vectors Q_i, such that $d[X_{m+1}(i), X_{m+1}(j)] \leq r$, where $j \neq i$, then calculate the frequency of occurrence, $B_i^{m+1}(r) = \frac{1}{p(N-\delta)-1}Q_i$, and define $B^{m+1}(r) = \frac{1}{p(N-\delta)}\sum_{i=1}^{p(N-\delta)} B_i^{m+1}(r)$.

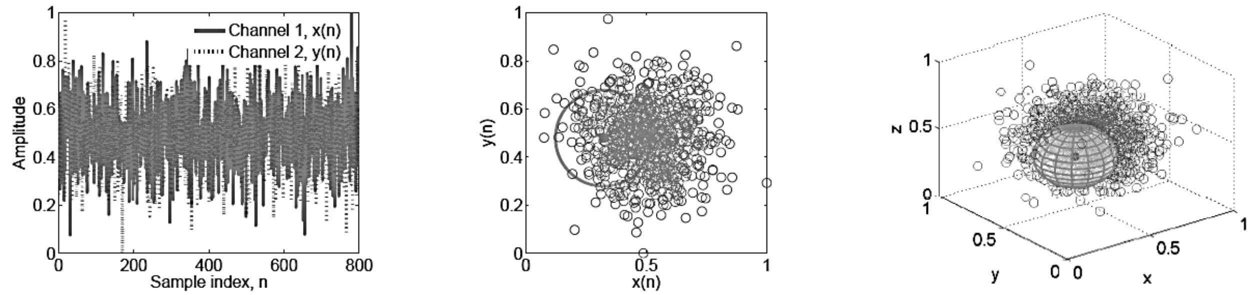
7: Finally, for a tolerance level *r*, estimate $MSampEn$ as

$$MSampEn(M,\tau,r,N) = -ln\left[\frac{B^{m+1}(r)}{B^m(r)}\right]. \tag{2}$$

3. Geometric Interpretation Of Multivariate Sample Entropy

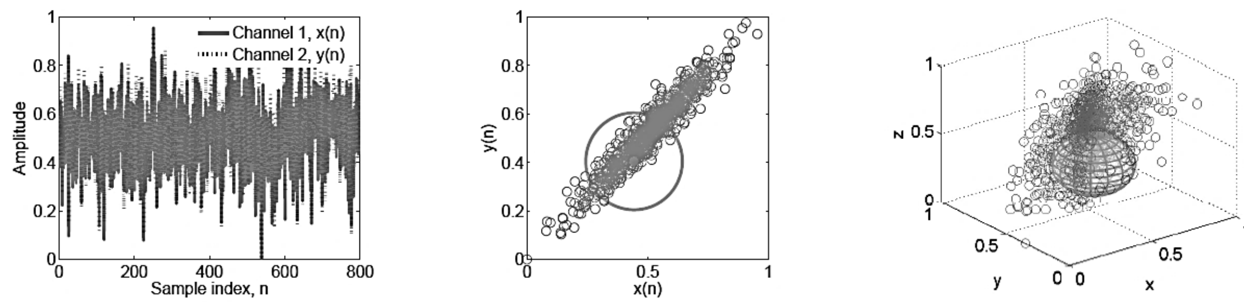
Underpinning the multivariate sample entropy method is the estimation of the conditional probability that two similar sequences will remain similar when the next data point is included. This is achieved by calculating the average number of neighbouring delay vectors for a given tolerance level (r) and repeating the process after increasing the embedding dimension, from m to $m+1$, a geometric interpretation of which is shown in Fig. 1 for uncorrelated bivariate white noise and in Fig. 2 for correlated bivariate white noise, respectively. Fig. 1(b) shows the set of delay vectors for a 2-channel uncorrelated white noise ([x(n), y(n)] with $\tau = [1,1]$ & $M = [1,1]$) and illustrates the neighbors as neighboring vectors at a point in an m-dimensional space that can be represented by the points enclosed by an m-sphere or an m-cube, for the Euclidean and maximum norm respectively. Upon increasing the embedding dimension from $(m=2)$ to $(m=3)$, we have two different subspaces spanning: (i) the vectors [x(n), x(n+1), y(n)] and (ii) the vectors [x(n), y(n), y(n+1)]. The combined 3D-plot is shown in Fig. 1(c). The $MSampEn$ algorithm accounts fully for both within- and cross-channel correlations by comparing the composite delay vectors of all such subspaces.

We next illustrate why MSampEn yields higher complexity for correlated multichannel signals than for uncorrelated signals. In Fig. 1(b), we see that there is no correlation between the channels of uncorrelated white noise, whereas in Fig. 2(b), the correlation between the channels of a correlated bivariate noise is clearly seen. In Fig. 1(c), there is no perceived structure in the 3-dimensional space and as a result both the quantity $B^m(r)$ and $B^{m+1}(r)$ are very similar. In other words, the probability of finding any two composite delay vectors which are similar within a tolerance level r in 2-dimensional space and in 3-dimensional space are higher and of the same order. As a result, MSampEn is lower. For correlated bivariate noise (Fig. 2), some structures (non-circular shape) are seen both in the 2-D space (Fig. 2(b)) and in the 3-D space (Fig. 2(c)). Moreover, this time the quantity $B^{m+1}(r)$ was much smaller than the quantity $B^m(r)$. In other words, the probability of finding any two composite delay vectors which are similar within a tolerance level r in 2-dimensional space is much higher than in the 3-dimensional space. As a result, MSampEn estimate is relatively higher.



(a) Uncorrelated bivariate white noise (b) 2D scatter plot (c) Combined 3D-plot for MMSE

Fig. 1: Geometric interpretation of MSampEn calculation from uncorrelated bivariate white noise



(a) Correlated bivariate white noise (b) 2D scatter plot (c) Combined 3D-plot for MMSE

Fig. 2: Geometric interpretation of MSampEn calculation from correlated bivariate white noise.

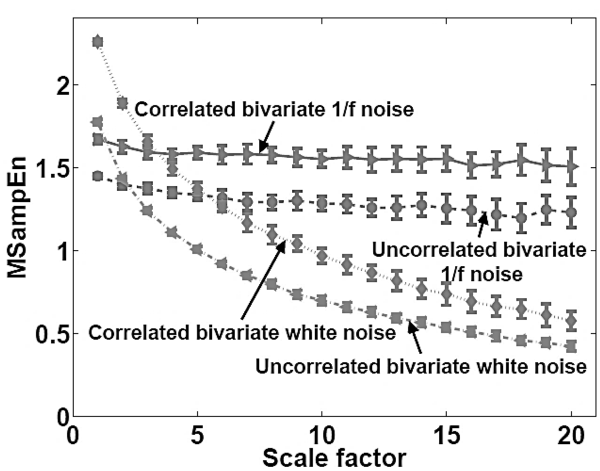


Fig. 3: Multivariate multiscale entropy (MMSE) analysis for bivariate white and 1/F noise, each with 10,000 data points. The curves represent an average of 20 independent realizations and error bars the standard deviation (SD)

4. Simulation Using Correlated Vs Uncorrelated Noises

MMSE is designed for multivariate data and caters for both within- and cross-channel correlations. To illustrate this, we first generated independent realizations of white and $1/f$ noise, and the two channels of bivariate white and $1/f$ noise were constructed using combinations of those independent realizations, thus making the channels correlated. Fig. 3 shows that, as desired, the proposed multivariate MSE fully caters for both within- and cross-channel correlations. Indeed, based on MMSE curves, the complexity of the correlated bivariate $1/f$ noise at large scales was the highest, followed by the uncorrelated $1/f$ noise, and correlated and uncorrelated white noise. This conforms to the underlying physics and validates the proposed MMSE method, as the complexity of the considered multivariate processes exhibiting both within- and cross-channel correlations is higher than that of uncorrelated multivariate white noise and uncorrelated multivariate $1/f$ noise (where long range correlations only exist within single channels).

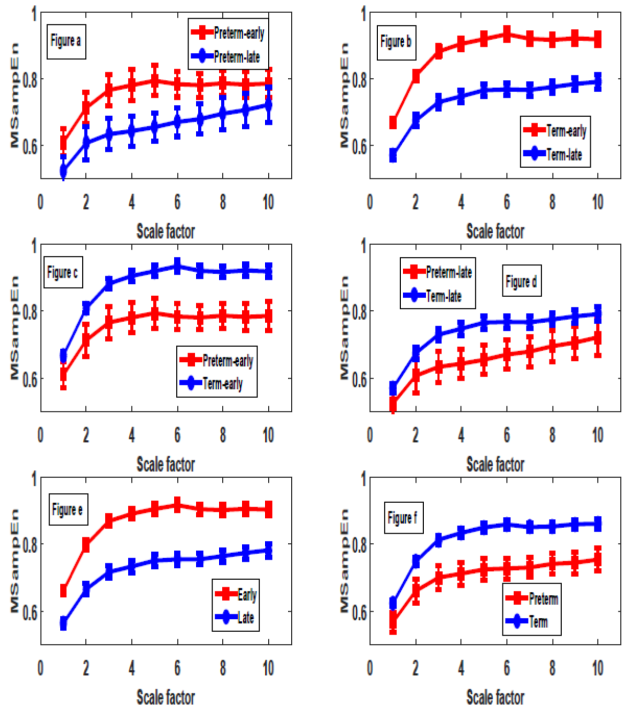


Fig. 4: MMSE analysis of filtered 3-channel Uterine EMG (UEMG) signal. The curves represent average complexity results, where the error bars denote the standard error (SE).

Experimental Results

Next we performed anomaly detection in different multivariate data through multiscale complexity analysis. The data sets included:

a) Uterine EMG data from term and preterm labour;

b) Financial data from different historical period;

c) Gait pattern form normal and constrained walking;

5.1 Anomaly detection in physiology of parturition

First, multivariate multiscale entropy analysis was applied on the uterine EMG records to detect anomaly by quantifying the degree of complexity underlying the term and pre-term deliveries. The Electrohysterogram records (uterine EMG records) used in this study are included in the Term-Preterm Electrohysterogram Database (TPEHG DB) and publicly available from PhysioNet [17]. The records were obtained during regular check-ups either around the 22nd week of gestation or around the 32nd week of gestation. The database contains 300 uterine EMG records from 300 pregnancies (one record per pregnancy) of which:

a) 262 records were obtained during pregnancies where delivery was on term (duration of gestation at delivery > 37 weeks): among them, 143 records were obtained before the 26th week of gestation and 119 were obtained later during pregnancy, during or after the 26th week of gestation;

b) 38 records were obtained during pregnancies which ended prematurely (pregnancy duration ≤ 37 weeks), of which: 19 records were obtained before the 26th week of gestation and 19 records were obtained during or after the 26th week of gestation.

Each record is composed of three channels, recorded from 4 electrodes, sampled at 20 Hz and 30 minutes in duration. Besides, each signal was digitally filtered using 3 different 4-pole digital Butterworth filters with a double-pass filtering scheme to ensure zero phase shift. A detail discussion about the database is given in Reference [18]. We had chosen two different m values, that is m=3 and m=2; and found that m=3 yield slightly better separation result than m=2 for each cases. The result also showed better separation when the band-pass filter cut-off frequencies were from 0.08Hz to 4Hz. As a result, only the results for m=3 and of cut-off frequency 0.08Hz-4Hz are reported here. In all the cases, r was taken as 0.15 times the total variation of the 3-channel UEMG signal.

There are significant differences between the UEMG signals recorded early and late (a, b and e panel of Fig. 4). This means as the time of gestation progresses, the average multivariate sample entropy values for both term and pre-term delivery records drop indicating higher predictability or less complexity of the signals as the delivery approaches. On the other hand, the MSampEn values are lower for pre-term delivery records (c, d and f panel of Fig. 4) regardless of the gestation duration at the time of recording which confirms that the pre-term delivery records are less complex or more predictable than the signals of term delivery records. Besides, in all the cases, the separation is better if we consider the multiscale MMSE curves than the measures in scale 1. This also confirms that the original signal not only contains information in the smallest scale but also reveals new information at all scales. Thus we can differentiate between term and preterm labour and also between EMG recorded earlier vs late in pregnancies in agreement with the earlier studies [19].

5.2 Anomaly detection in financial data

Next, the anomaly in financial time series was detected by analyzing the underlying complexity of the world economy at large. The severity of the major events and their relative impacts on the global economy is well understood and analyzed, thus placing us in a better position to interpret and analyze the results from the MMSE analysis. Four major indices in the United States economy were selected for our study. The choice of using the economy of the United States for analysis was attributed to the fact that it is recognized as one of the most influential financial markets around in the world, with deep and intricate connections and interdependent relationships with many other markets such as Europe, Japan, and China. It can therefore be assumed that the market trends of these major indices are indicative of the world economy by and large. The 4 indices chosen were Dow Jones Industrial Average, S&P 500, Nasdaq Composite, and Russell 2000.

They were chosen for our study because they are amongst the most widely used benchmark indices in tracking the market performance in the United States [20].

5.2.1 Segmentation of market trend history

For the purpose of our study, we zoomed into approximately two decades of market trend between 1991 and 2011. Before putting the time series of the 4 indexes through MMSE algorithm, we must first have a good understanding of some of the historical events which shaped the progression of these time series. To that end, the data between 1991 and 2011 was divided into 2 different periods, with the intention of classifying the different eras that the world economy went through based on the various historical events that took place. It is difficult to clearly demarcate the various periods according to the different times of stress because it is hard to define clear start and end dates for an economic crisis. In addition, the lead up to a crisis and the duration of the impact after the crisis may vary and is usually subjective.

Period 1 (01/01/1991 to 31/12/1999). The economy in the 1990s went through what is popularly termed by many as the 'dot-com boom'. The decade saw exceptional levels of growth in the markets, particularly in the technology industry, as market sentiments were buoyed by fast moving technological advancements. Several incidences of major political and civil unrest around the world such as the Collapse of the Soviet Union (1991), Persian Gulf War (1990-1991), Yugoslav Wars (1991-1995 and 1998-1999), Rwanda Genocide (1994) and the Second Congo War (1998) failed to dampen the exceedingly positive sentiments surrounding the Information Age of the market. The only significant dent inflicted on the markets during this decade was the Asian Financial Crisis (1997) which saw sizeable losses across all major indexes due to the severe turbulence that originated from the economies within South-east Asia [21].

Period 2 (01/01/2000 to 31/12/2011). The early 21st century brought about high levels of uncertainty as the 'dot-com bubble' burst. A lethal combination of over-enthusiasm of the information era and bad financial practices [22] which developed towards the end of the 20th century culminated in one of the biggest financial market meltdown recorded since the Great Depression. The crisis was further compounded by the terrorist attacks on 11th September 2001 which saw retaliation in the form of the declaration of the War on Terror by the United States and the subsequent invasions of Afghanistan (2001) and Iraq (2003).

The middle of the new decade of the 21st century witnessed renewed enthusiasm amongst investors who saw huge investment opportunities in stocks which were believed by many to be undervalued after the turbulent start to the new century. There was rapid recovery amongst all markets as major indexes retraced steadily and regained large amounts of ground that were lost during the recession of the early 2000s. Just less than a decade on from the recession of the early 21st century, the world economy was to plunge back into yet another deep recession owing to the collapse in the sub-prime mortgage market in the United States. The severity of the crisis deepened through 2008 which eventually led to the infamous and iconic collapse of Lehman Brothers as they filed for bankruptcy on 15th September 2008, leading to what is known as the Global Financial Crisis of the late 2000s. In a rather similar fashion to 'dot-com bubble' burst, the recession was brought about by the bursting of a bubble which in this case was termed as the 'housing bubble'.

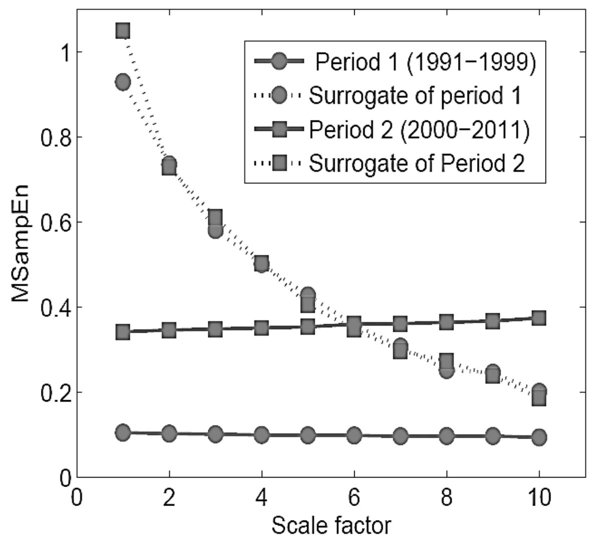


Fig. 5: MMSE analysis for quadrivariate financial time series

5.2.2 MMSE analysis

Each two period of the time series for the 4 indexes were grouped together to give a quadrivariate input to the MMSE. The values of the parameters used to calculate MSampEn were $m_k = 2$, $\tau_k = 1$ and $r = 0.15 \times$(standard deviation of the normalized time series) for each data channel. For rigor, we also performed the complexity analysis on a set of multivariate surrogates generated from the above data sets via random shuffling; this provided a reference for a suitable comparison of complexity estimates obtained from different physical systems. Randomized shuffling of the input data channels effectively destroyed temporal and cross-channel correlations among their samples, while preserving their first and second order statistical properties. This way, significant difference between observed complexity estimates from input data sets and their respective (randomly shuffled) surrogates, over a range of scales, would reject the null hypothesis of both temporal and cross-channel independence, implying a higher complexity and nonlinear coupling in the considered data sets.

Fig. 5 shows the results of the MMSE analysis. The quadrivariate financial time series in both periods exhibit long-range correlation. This is reflected in the approximately horizontal plots of the various time series. Besides, period 2 demonstrate a higher measure of complexity as during this period there was relatively high volatility and financial distress and conversely a lower measure of complexity during period 1 of relative financial stability. The intuition here is that when the degree of volatility and stress in the market increases, the progression of the time series becomes more responsive to the events and factors that shape market movement, thus giving rise to an increase in the complexity measure. Conversely, when there is financial stability in the Market, the progression of the time series becomes more regular. Although the financial time series all tend to generally trend upwards during periods of stability, we can argue that the business cycles become much more predictable, which intuitively makes the financial time series less complex under such a scenario. The MMSE analysis therefore provides us with results that are intuitively agreeable with the behaviour of the financial time series. This affirms the accuracy and precision of the MMSE algorithm in characterizing multivariate time series in terms of complexity, with applicability to financial time series. This result presents us with a novel approach to detect anomaly in financial time series by using complexity measure and thus characterize the degree of volatility and stress in the financial markets.

5.3 Anomaly detection in gait pattern

Next, MMSE was used to detect anomaly in gait patterns. 3D acceleration was recorded from a subject who walked continuously for two minutes on level ground around an obstacle free, long, approximately oval path using motion tracker sensors from Xsens Technologies. Six motion tracker sensors are placed on left and right ankle, thigh and wrist. The sampling frequency was 75 Hz. Two anomaly walking pattern was also simulated. In one case, one backpack with some weights was carried by the subject. In other case, some weights were tied with both left and right thighs of the subject and the subject walked blind-folded. For each condition, five trials were recorded.

The values of the parameters used to calculate MSampEn were $m_k = 2$, $\tau_k = 1$ and $r = 0.15\ \times$(standard deviation of the normalized time series) for each data channel; these parameters were chosen on the basis of previous studies indicating good statistical reproducibility for SampEn [23]. For MMSE, the length of each coarse-grained sequence was ϵ (scale factor) times shorter than the length of the original series, so the highest scale factor considered in the analysis was ϵ =10.

Fig. 6 shows the results obtained by the MMSE analysis. As expected, the anomaly conditions (weighted and blind-folded or backpacked) have lower complexity for most of the scale factors than the normal condition (normal walking). This is because the normal walking condition represents unconstrained system whereas by wearing a backpack or blind-folding, we have actually constrained the gait dynamics. On the other hand, the surrogates show the MMSE profiles similar to that of the random noises.

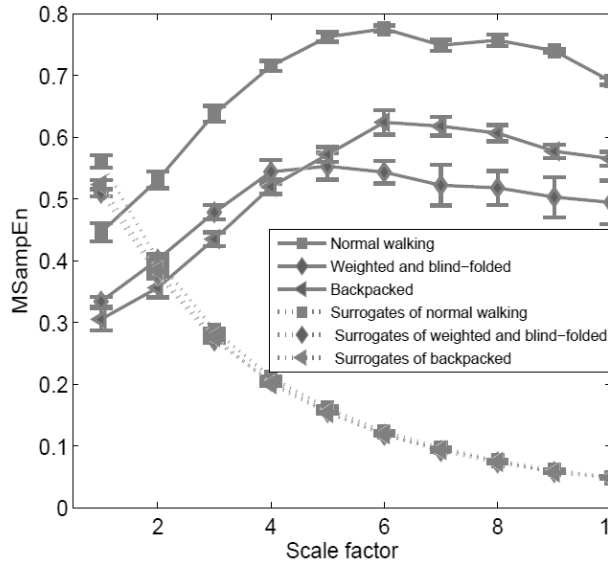


Fig. 6: Multivariate multiscale entropy (MMSE) analysis for anomaly detection in different gait patterns. The curves represent an average of 5 trials and error bars the standard deviation (SD).

5. Conclusion

This work has introduced a novel method for anomaly detection from multivariate data based on multiscale complexity analysis. The proposed method assumes the anomalies as generated by a constrained system and thus have different degree of complexity than the normal, unconstrained system. The proposed multivariate multiscale entropy (MMSE) method has been shown to be naturally suited to reveal the long range within- and cross-channel correlations present in multichannel data and thus can detect anomalies as the breakdown of this sort of spatio-temporal correlations. The approach has been validated on synthetic data and on real world multivariate financial and biomedical datasets. This method works on systems or signals which shows long range correlation as the method is based on complexity measures which estimate this type of correlation. This is the only limitation of this method. In future, we will compare our methods with the existing anomaly detection techniques.

References

1. V. Chandola, A. Banerjee, and V. Kumar, “Anomaly detection: A survey,” *ACM Computing Surveys*, vol. 41, no. 3, pp. 15:1–15:58, Jul. 2009.
2. F. Esponda, S. Forrest, and P. Helman, “A formal framework for positive and negative detection schemes,” *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, vol. 34, no. 1, pp. 357 – 373, feb. 2004.
3. G. Fernandes, J. J. P. C. Rodrigues, L. F. Carvalho, J. F. Al-Muhtadi, and M. L. Proenca, “A comprehensive survey on network anomaly detection,” *Telecommunication Systems*, Jul 2018. [Online]. Available: https://doi.org/10.1007/s11235-018-0475-8
4. R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern Classification* , 2nd ed. Wiley-Interscience, Nov. 2001.

5. P.-N. Tan, M. Steinbach, and V. Kumar, *Introduction to Data Mining*, 1st ed. Addison Wesley, May 2005.

6. A. K. Jain and R. C. Dubes, *Algorithms for clustering data*. Prentice-Hall, Inc., 1988.

7. A. Abid, A. Kachouri, and A. Mahfoudhi, “Outlier detection for wireless sensor networks using density-based clustering approach,” *IET Wireless Sensor Systems*, vol. 7, no. 4, pp. 83–90, 2017.

8. D. Agarwal, “Detecting anomalies in cross-classified streams: A Bayesian approach,” *Knowledge and Information Systems*, vol. 11, pp. 29–44, 2007.

9. H. Ozkan, F. Ozkan, and S. S. Kozat, “Online anomaly detection under markov statistics with controllable type-i error,” *IEEE Transactions on Signal Processing*, vol. 64, no. 6, pp. 1435–1445, March 2016.

10. Z. He, X. Xu, and S. Deng, “Discovering cluster-based local outliers,” *Pattern Recognition Letters*, vol. 24, no. 910, pp. 1641 – 1650, 2003.

11. V. Hodge and J. Austin, “A survey of outlier detection methodologies,” *Artificial Intelligence Review*, vol. 22, pp. 85–126, 2004.

12. M. Costa, A. L. Goldberger, and C. K. Peng, “Multiscale entropy analysis of complex physiologic time series,” *Physical Review Letters*, vol. 89, no. 6, p. 068102, 2002.

13. ——, “Multiscale entropy analysis of biological signals,” *Physical Review E*, vol. 71, p. 021906, 2005.

14. M. U. Ahmed and D. P. Mandic, “Multivariate multiscale entropy: A tool for complexity analysis of multichannel data,” *Physical Review E*, vol. 84, pp. 061 918–1 061 918–10, 2011.

15. ——, “Multivariate multiscale entropy analysis,” *IEEE Signal Processing Letters*, vol. 19, no. 2, pp. 91–94, 2012.

16. A. L. Goldberger, L. A. N. Amaral, J. M. Hausdorff, P. C. Ivanov, C. K. Peng, and H. E. Stanley, “Fractal dynamics in physiology: Alterations with disease and aging,” *Proceedings of National Academy of Sciences USA*, vol. 99, no. Suppl 1, pp. 2466–2472, 2002.

17. A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C.-K. Peng, and H. E. Stanley, “Physiobank, Physiotoolkit, and Physionet : Components of a new research resource for complex physiologic signals,” *Circulation*, vol. 101, no. 23, pp. e215–e220, 2000.

18. G. Fele-Zorz, G. Kavsek, Z. Novak-Antolic, and F. Jager, “A comparison of various linear and non-linear signal processing techniques to separate uterine EMG records of term and pre-term delivery groups,” *Med. Biol. Engineering and Computing*, vol. 46, no. 9, pp. 911–922, 2008.

19. M. U. Ahmed, T. Chanwimalueang, S. Thayyil, and D. P. Mandic, “A multivariate multiscale fuzzy entropy algorithm with application to uterine emg complexity analysis,” *Entropy*, vol. 19, no. 1, 2017.

20. About.com, Part of The New York Times Company. What market indexes tell us, the dow and other market indexes explained. [Online]. Available: http://stocks.about.com/od/understandingstocks/a/Indexes102704.htm.

21. G. L. Kaminsky and S. L. Schmukler, “What triggers market jitters?: A chronicle of the asian crisis,” *Journal of International Money and Finance*, vol. 18, no. 4, pp. 537 – 560, 1999.

22. J. J. Morris and P. Alam. (2008, June 27) Analysis of the dot-com bubble of 1990s. [Online]. Available: http://ssrn.com/abstract=1152412

23. J. S. Richman and J. R. Moorman, “Physiological time-series analysis using approximate entropy and sample entropy,” *American Journal of Physiology - Heart and Circulatory Physiology*, vol. 278, no. 6, pp. H2039–2049, 2000.