Optimising parameters in recurrence quantification analysis of smart energy systems

Conference or Workshop Item

Accepted Version

It is advisable to refer to the publisher's version if you intend to cite from the work. See Guidance on citing.

the End User Agreement.

www.reading.ac.uk/centaur

CentAUR

Central Archive at the University of Reading

Reading's research outputs online

Optimising Parameters in Recurrence Quantification Analysis of Smart Energy Systems

Georgios Giasemidis
CountingLab Ltd., UK
georgios@countinglab.co.uk

Danica Vukadinović Greetham
Knowledge Media Institute
Open University, UK
danica.vukadinovic-greetham@open.ac.uk

Abstract—Recurrence Quantification Analysis (RQA) can help to detect significant events and phase transitions of a dynamical system, but choosing a suitable set of parameters is crucial for the success. From recurrence plots different RQA variables can be obtained and analysed. Currently, most of the methods for RQA radius optimisation are focusing on a single RQA variable. In this work we are proposing two new methods for radius optimisation that look for an optimum in the higher dimensional space of the RQA variables, therefore synchronously optimising across several variables. We illustrate our approach using two case studies: a well known Lorenz dynamical system, and a time-series obtained from monitoring energy consumption of a small enterprise. Our case studies show that both methods result in plausible values and can be used to analyse energy data.

Index Terms—Recurrence Quantification Analysis, Smart grids, Energy disaggregation.

I. INTRODUCTION

Observing a particular dynamical system, defined by human behaviour, often results in a complex, non-linear time-series. Recurrence, one of its fundamental properties that captures the system's phase space, is a helpful tool to analyse the system. A recurrence plot [1], [2] can unfold the latent repeating patterns and is a visual tool that facilitates an investigation of the system. Recurrence quantification analysis (RQA) helps to quantify the recurrent structures revealed by these plots, by computing time dependent RQA variables [3], [4]. RQA variables can detect significant events and phase transitions, such as financial crashes, transitions in climate systems, etc.

There are many RQA variables or measures in literature that have been used to detect important insights about dynamical systems. The applications include heart rate [5], financial time-series [3], etc. Only recently, RQA was applied to energy data obtained from high resolution monitoring of electricity usage [6] and was used to map different appliances, to detect faulty devices and identify unexpected usage patterns. The benefits of RQA include relatively simple calculations, powerful investigative characteristics, and applicability to a wide range of systems. One of the main challenges is to choose a set of parameters that are suitable for the application. Therefore, parameters need to be optimised according to the observed system's characteristics and the application. In [7] several potential issues are recognised when choosing parameters. This is especially true with emerging areas of applications such as high resolution electric energy data.

The novelty of this work is twofold: (i) to the best of our knowledge, this is the first application of parameter optimisation for RQA to electricity data from Small and Medium Enterprises (SMEs), (ii) we introduce two practical methods for optimising the radius for RQA, which address the following open question in the literature. In existing work, the radius is optimised for each RQA variable, and different RQA variables might result in different optimal radius for the same data. In addition, a decision is made after visual inspection of the plots of the different RQA variables. To remedy this, our methods use all five main RQA variables to automatically identify the optimal radius without further visual assessment.

In this work, we start with a quick overview of energy disaggregation literature in Section II. Fundamentals of RQA are given in Section III-A and we list a selection of variables relevant for our energy case study in Section III-B. Our main contribution, two novel methods for radius optimisation are given in Section III-C. The methods are applied to two case studies, the Lorenz system and an energy data set in Section IV and we list the benefits and limitations of newly proposed methods in Section V.

II. PREVIOUS WORK

An area of extensive academic interest over the last decade, non-intrusive load monitoring (NILM) aims to identify electric load components (appliances) based on measured aggregated load.

An overview of main concepts, techniques and algorithms of NILM can be found in [8], [9]. The classical Hart's algorithm [10] starts by detecting significant changes in the signal and then clusters them in order to find different states' changes of each appliance.

In the last 25 years many different techniques were developed using unsupervised and supervised learning, and low and high frequency sampling. While high frequency sampling could detect fine features of the signal, most of smart meters work in low frequency mode - reporting usage at 1s or less frequently. Supervised learning algorithms contain a training phase, to create a map of all possible devices and their features. After training phase, the disaggregation is performed using optimisation or pattern recognition. The disaggregation can be formulated as an optimisation problem, more precisely error minimisation. In this problem, the known devices need to be identified, such that their sum is closest

to the actual consumption, usually requiring combinatorial optimisation techniques to find the solution. Alternatively, pattern recognition methods, classifiers, neural networks, hidden Markov Models and similar are used to identify devices [11]. In unsupervised learning, the training phase may happen in parallel with the disaggregation. Most of academic work in this area is concentrated on recognising household's appliances, and combinatorial optimisation and factorial Hidden Markov Models are currently regarded by most scientists as the state of art.

Recent unsupervised approaches for households include using probabilistic graphical models to represent appliances, where the models' parameters are learned during training phase [12]. Both aggregate power reading and differences between two consecutive readings are used to create a version of difference Hidden Markov Model (HMM) for each appliance. Factorial HMMs, where several HMMs are evolved independently and some function of all the hidden states is the output, featured in [13] obtaining good disaggregation results on home appliances. Another application of factorial HMM is given in [11] where simultaneously running two devices was explicitly modelled as an interaction chain in HMM. Better disaggregation accuracy was obtained for devices with known interactions property while not much difference was observed for devices with unknown mutual interactions.

In [14] three deep neural network architectures were trained and tested on home appliances disaggregation. A recurrent neural network, so called long short term memory (LSTM) performed well on two-state appliances (on and off, e.g. toaster) but not on multi-state appliances that can be in many different discrete states (e.g. washing machine). While the preliminary results are promising, literally millions of trainable parameters represent a challenge. Google recently reported that the deployment of an ensemble of deep neural networks trained on diverse data obtained from thousands of sensors reduced cooling data centre energy costs by 40%[1].

Other recently developed complex techniques that are comparable with factorial HMMs and combinatorial optimisation include using spatiotemporal pattern networks between several variables such as indoor and outdoor temperature and time of the day, in addition to whole building electricity and its moving average to obtain disaggregation [15].

In [16] factorial HMMs are combined with particle filtering, to estimate disaggregated appliance states. Particle filtering is well suited to non-linear behaviour and non-Gaussian noise.

While most of methods are developed and tested for households or the largest enterprises where economy of scale helps to justify investments, small and medium companies are somewhat left out [17].

Recently, recurrence quantification analysis (RQA) was applied to supervised disaggregation of small business load in [6]. RQA, suitable for non-linearity and complex time-dependencies, was combined with principal component analy-

sis to create an easy visualisation of disaggregation for users. Here we explore parameter optimisation of such approach with the aim to investigate a possibility of unsupervised system based on RQA in the future.

III. PARAMETER OPTIMISATION

A. Recurrence Plot

Recurrence plots capture the recurrent states of a complex system [1]. A state of the system is considered recurrent when it is in a close neighbourhood of a previous state of the system in the phase space. Given a time-series of n observations $X = \{x_1, \ldots, x_n\}$, the phase space is defined by transforming the readings into time-delayed vectors at each time-step, i.e. $\mathbf{Y} = \{Y_1, \ldots, Y_m\}$, where $Y_i = (x_i, x_{i-\tau}, \ldots, x_{i-(D-1)\tau}) \in \mathbb{R}^D$, τ is the delay and D is the dimension of the phase-space, and $m = n - (D-1)\tau$.

The distance matrix $DM(i,j) = ||Y_i - Y_j||_2$, $i,j = 1, \ldots, m$, is the Euclidean distance between vectors Y_i and Y_j in the phase-space. The recurrence plot follows from the distance matrix, defined as

$$R(i,j) = H(\epsilon - DM(i,j)), \qquad (1)$$

where $H(x)$ is the Heaviside step function.

The entry $R(i,j)$ equals one and the states Y_i and Y_j are considered recurrent, when the distance between Y_i and Y_j is within an ϵ-radius in phase-space. As a state Y_i in phase space corresponds to a time-step of the original time-series X, recurrence plots inform us of recurring patterns within our current time-series.

From the aforementioned definitions, it becomes evident that the recurrence plot depends on three parameters, the delay τ, the embedding dimension D and the radius ϵ, to capture the correct dynamics of a system with noise.

B. RQA variables

Following [6], we concentrate on several RQA variables, all obtained from the recurrence plot matrix R:

- **REC**, the percentage of all points within the square window of size W that are recurrent.
- **DET**, the percentage of the recurrence points that form line segments parallel to the matrix diagonal.
- **ENT**, the Shannon entropy of the distribution of diagonal line segments.
- **LAM**, the percentage of the recurrence points that form vertical line segments.
- **TT**, the average length of the vertical line segments,

These variables are frequently used across different applications [3], [4], [18]. DET relates to repeating or deterministic patterns within the system, ENT represents complexity, where small and large ENT signify periodic and unpredictable behaviour respectively and LAM represents stationary behaviour.

C. Methodology

Several studies have focused on the optimisation of the three parameters [7], [19]. Common methods for optimising the delay include the first minimum of either the autocorrelation function or the mutual information of the time-series [7], [20]. Here, we choose the latter as it captures non-linear characteristics of the time-series. The embedding dimension is often determined using the false nearest neighbours parameters [21]. In this study, we use a modification of this algorithm, introduced by Cao [22], in order to be parameter-free and efficient.

Several methods have been proposed for choosing the radius. A too small value allows no recurrent patterns, a too large values may result in false recurrences [19, references therein]. Several rules of thumb have been proposed, such as a value corresponding to 1% of REC [23], $\epsilon = 0.1\sigma$, where σ is the standard deviation of the time-series, a value that does not exceed the 10% of the mean or the maximum of the phase-space [19], [24, references therein].

Other more sophisticated methods have been proposed in [19], [25], [26]. Although these methods are distinct, they share some common features. They create a surrogate signal, which is the original signal with additive noise. Then they run RQA for several segments of the original and the surrogate signals. The objective is to identify the radius that is the best in discriminating between the original and the surrogates dynamics. This is achieved by optimising several measures such as the average loss of decision action [19], the area under the ROC [24] or the quality loss function [25]. However, these scores are applied per RQA variable, and in most cases, the score is optimised at different radii values for different RQA variables. Therefore, the question remains, how can one decide in a practical way which radius to choose, as further analysis is required to decide which RQA variable is more significant [24]. Here, we propose two practical methods to identify the optimal radius without having to choose 'the most important' variable.

1) Method 1: In the first method, a surrogate signal is produced. A number N of segments are randomly drawn from the original time-series and its surrogate. For several values of the radius, RQA variables are computed for each segment of the original and surrogate signals. These correspond to points $Q = (REC, DET, ENT, LAM, TT) \in \mathbb{R}^5$ in the 5-dimensional RQA space (the space of the RQA variables defined in the previous section). If we label the points corresponding to the original signal as "Cluster 1" and those corresponding to the surrogate as "Cluster 2", as the radius is varied, these points in each cluster spread or cluster together depending the quality of the clustering. To assess the quality of the clustering we use scores from the Machine Learning clustering literature. The inertia, the within-cluster distance, is defined as

$$I = \sum_i \min_{\mu_j \in C} ||Q_i - \mu_j||_2, \tag{2}$$

where $\mu_j \in \mathbb{R}^5$ is the centroid of the cluster j. The inter-cluster distance is defined

$$d_C = ||\mu_1 - \mu_2||_2 \tag{3}$$

The maximum discrimination between the original and surrogate signals is achieved when the inertia is minimised and inter-cluster distance of the centres of the clusters is maximised. In order for the two measures to be of comparable sizes when combined, they are normalised to take values in the interval $[0, 1]$. The optimal radius is given by

$$\epsilon^* = \arg\min_\epsilon (\hat{I}(\epsilon) - \hat{d}_C(\epsilon)), \tag{4}$$

where $\hat{I}$ and $\hat{d}_C$ are the normalised inertia and inter-cluster distance respectively.

2) Method 2: In the second method, the RQA variables of the original signal are computed for different values of the radius, $Q_{or}(\epsilon) \in \mathbb{R}^5$. These points form a curve in the 5-dimensional phase-space. A sample of surrogates is also produced for each value of the radius, their RQA variables are extracted and their mean value is computed, $Q_{su}(\epsilon) \in \mathbb{R}^5$. The maximum discrimination occurs when the Euclidean distance between the RQA variables of the original and mean surrogate signals is maximal, i.e.

$$\epsilon^* = \arg\max_\epsilon ||Q_{or}(\epsilon) - Q_{su}(\epsilon)||_2. \tag{5}$$

Several surrogate functions have been used, such as shuffling or Fourier based surrogates [19]. Here, we use an additive Gaussian noise to produce the surrogate [24], i.e. $X_{wn} = X + \mathcal{N}(0, \alpha\sigma)$, where σ is the standard deviation of the original time-series X and α controls the amount of noise.

IV. Applications

We apply the proposed methods to two systems, the Lorenz attractor and to real data from electricity demand from an SME [6]. For all systems, the time-delay is identified through the first local minimum of the mutual information and the embedding dimension using Cao's algorithm [22]. Then, we identify the radii values that correspond to a range of REC values of the system from 0% to 100% in increments of 0.5%. For the white-noise surrogate we use $\alpha = 0.2$. In the following examples, for comparison, we also plot the values of the radius that correspond to 1% REC, 0.1σ, 10% of maximum and mean distance in the phase space.
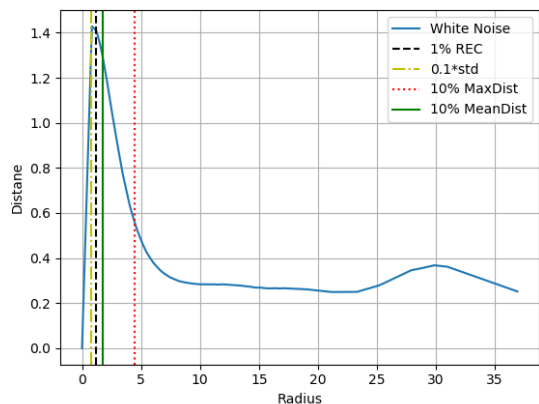
A. Lorenz Attractor

The Lorentz system is a dynamic system of three differential equations [27] resulting in chaotic behaviour for some choices of parameters and initial conditions:

$$\left(\frac{dx}{dt}, \frac{dy}{dt}, \frac{dz}{dt} \right) = (\sigma y - x, x(\rho - z) - y, xy - \beta z) \tag{6}$$

We create $N = 3000$ times-steps of the system with step-size 0.01 and parameters $\sigma = 10, \rho = 28, \beta = 8/3$ that result in chaotic behaviour. The time-delay is determined to be $\tau =$
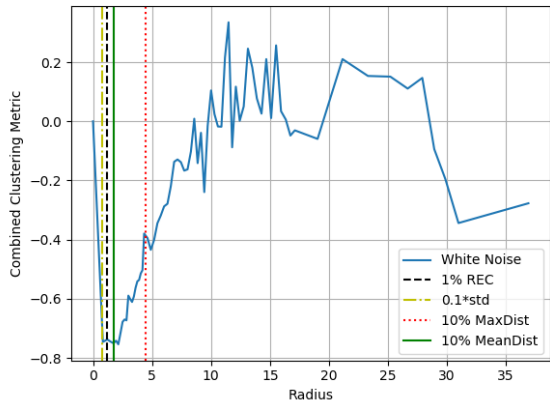
(a) The metric based on clustering of the RQA values of the segments of the original signal and the white noise (blue).

(b) The distance measure in the 5-dimensional space between the original signal and the white noise (blue).

Fig. 1: Lorenz attractor

Fig. 2: A weekly profile of the total current from a dry cleaners business. The business is closed on Sundays.

17 and the embedding dimension $D = 3$, consistent to the literature [21], [22].

Figure 1a shows the measure (4) which assesses the quality of the clustering of the RQA values between the original and the surrogate signals. The maximum discrimination (i.e. minimum value) to the white-noise surrogate is achieved at the value $\epsilon^* = 2.14$, which is comparable to the optimal value in [28], [29], and corresponds to 3% REC, a conclusion similar to [19], [25] too.

The distance in the RQA-space between the original signal and its surrogate as a function of the radius is presented in Figure 1b. The optimal value is at at $\epsilon^* = 0.84$, corresponding to 0.5% REC, which also coincides with 0.1σ rule of thumb.

B. Case study: energy data

The energy data corresponds to total electric current at five minute resolution from a dry cleaners, a small UK business [6], for a period of 6 weeks from 11-th of September to 22-nd of October 2017. Figure 2 shows a typical weekly profile.
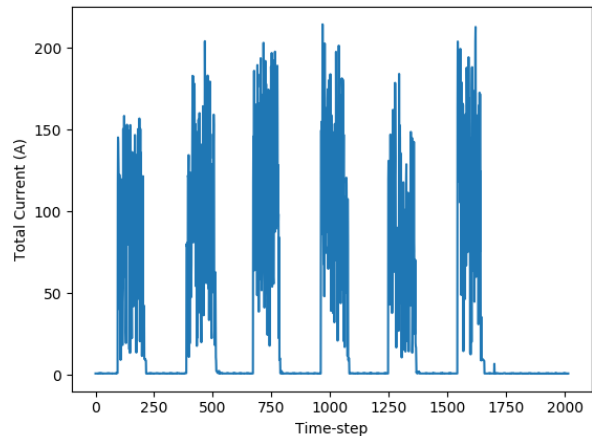
As we are interested in the dynamics while the business is open, we include only the operational times in our analysis. To identify the operational times, we searched for the maximum reading in the data between 10pm and 6am at nigh, where the business is closed, see Figure 2. The maximum value is $1.4A$ and any values greater than this are treated as "on" states, i.e. operational times.

In Figure 3, we plot the distribution of readings for each day of the week. We observe that all days have similar medians and quartiles. In Figure 4, the distribution of readings for each hour of the day is plotted. The distributions of the readings are similar for the hours between 8am to 5pm, the current usage drops between 5pm and 6pm, however after 6pm the median and quartiles of the electrical current are much lower, but higher than the "off-state". This is because most devices have been turned off and not in use, whereas one or two devices still operate until the closure of the business.

From Figure 3 and 4, we observe no daily or hourly patterns. This is expected as this is data from an SME, which operates in full demand during its operational hours, in contrast to household data which has daily and weekly behavioural patterns. The lack of strong periodicity patterns and the high standard deviation of the distributions indicate that the readings have a strong stochastic component, which is also unveiled from our analysis below.

For the optimisation of the RQA parameters we focus on a period of two weeks from 11-th to 25-th of September 2017. In Figure 5a we plot the mutual information for several lags. There is a steep drop in the mutual information for lags greater than 1, with a constant plateau, which indicates that the optimal value is $\tau = 1$. Figure 5b demonstrates how Cao's measures E_1 and E_2 [22] change with the embedding dimension. The fluctuation of E_2 around the value 1 and the slow increase of measure E_1 to high values of embedding dimension are characteristic of signals with a dominant stochastic component [21], [22], which is indeed the case for
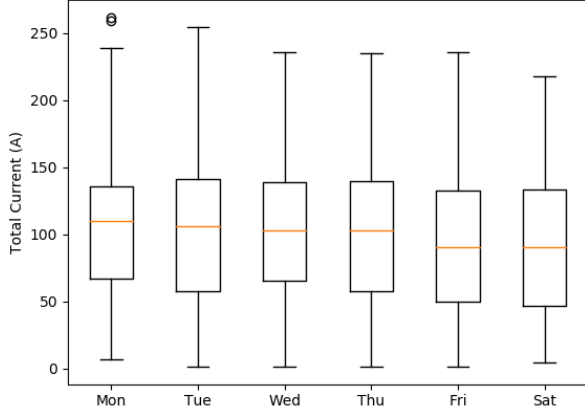
Fig. 3: Distribution of electrical current readings for days of the week (the business is closed on Sundays)
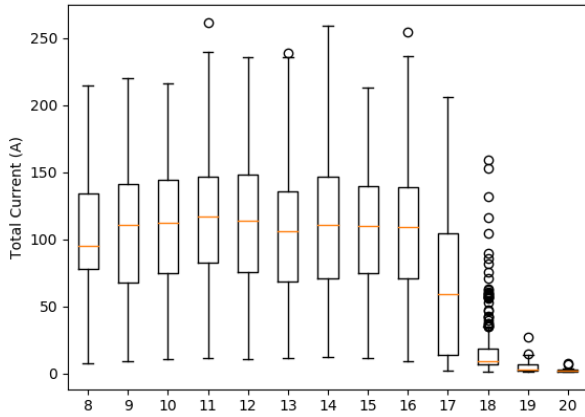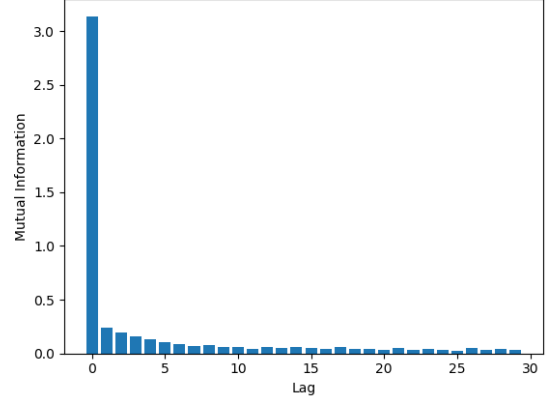

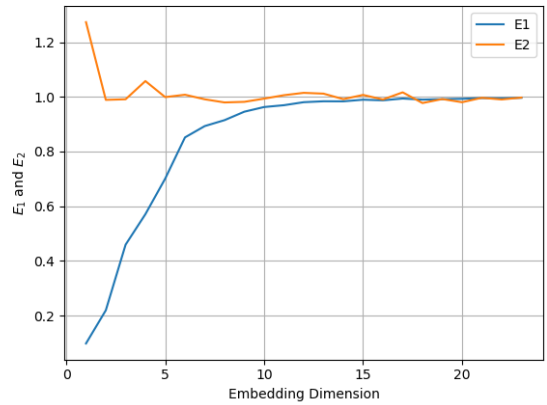
Fig. 4: Distribution of electrical current readings for each hour of the day (business is closed before 8am and after 9pm).

(a) Mutual information for dry-cleaners data.

(b) Cao's metrics for dry-cleaners.

Fig. 5: Mutual information and embedding dimension for the dry-cleaners.

real load data. We conclude that the embedding dimension should be $D = 10$.

Having determined the time delay and embedding dimension, we use these values to estimate the radius of RQA. Figures 6a and 6b show how the radius changes for methods 1 and 2, respectively. We observe that the radius extends to high values, which is expected as the embedding dimension of the system is relatively large. The optimal values are at $\epsilon^* = 264.5$ (75% REC) and $\epsilon^* = 200.5$ (35% REC). These optimal values occur at high percentages of the REC, results that are consistent with the findings in [19]. In this study, the authors studied heart rate dynamics data (which consists of linear, chaotic and stochastic components) at different embedding dimensions and they found that as the embedding dimension increases, the optimal radius (according to their method) occurs at higher levels of REC. Particularly, at $D = 10$, their

optimal radius occurs at 68% REC, which is consistent with our findings.

In order to further understand the dependence of the optimal radius with the embedding dimension, we apply the proposed methods for $\tau = 1$ and $D = 1$. Figure 7a shows how the clustering measure eq. (4) varies with the radius. We observe that the radius extends to very high values in order to span REC values from 0% to 100%. The white-noise surrogate achieves a maximum discrimination at $\epsilon^* = 1.08$ which corresponds to 1.5% REC. Figure 7b shows that the distance measure eq. (5) is maximised at $\epsilon^* = 0.69$ (1% REC). The discrimination between the original signal and the surrogate reduces with the increase of the radius as expected, due to the identification of false recurrences in the surrogate. For $D = 1$, we observe that optimal values occur at lower values of the REC levels as in [19].

V. DISCUSSION

Optimising the parameters in RQA is a non-trivial task and the success of the proposed methods depends on the applied
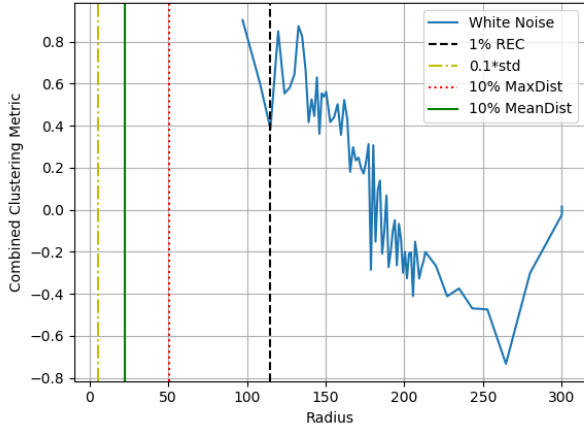
(a) The metric based on clustering of the RQA values of the segments of the original signal and the white-noise surrogate (blue).
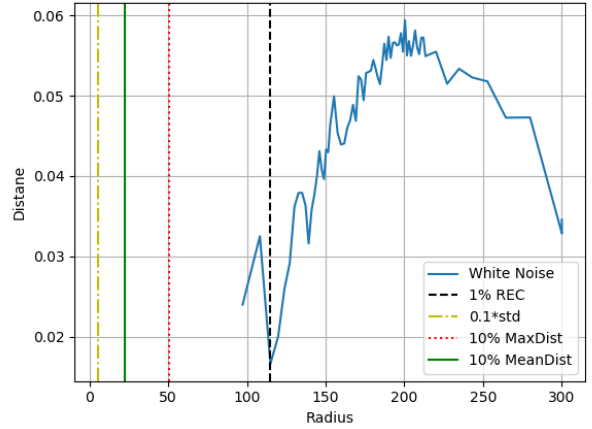
(b) The distance measure in the 5-dimensional space between the original signal and the white-noise surrogate (blue).

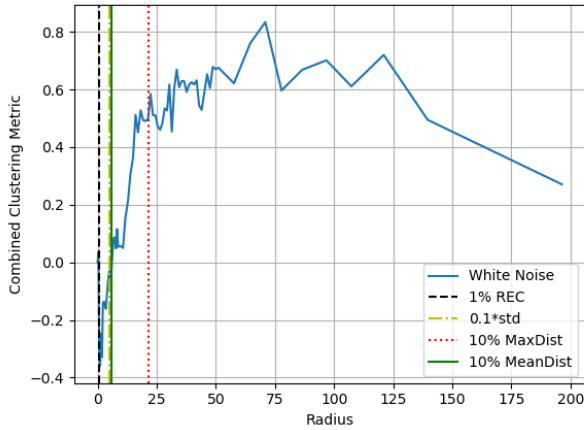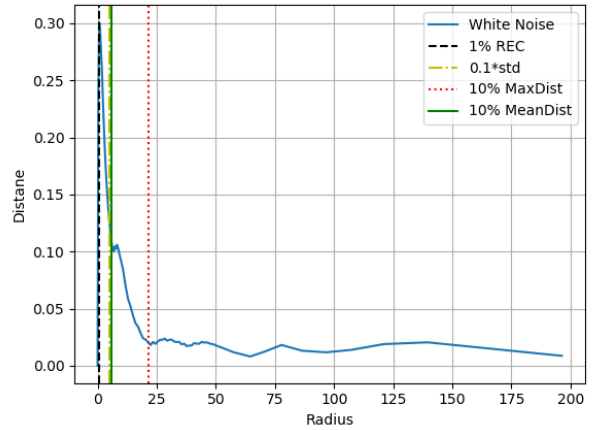Fig. 6: Dry cleaners ($\tau = 1$ and $D = 10$).

(a) The metric based on clustering of the RQA values of the segments of the original signal and the white-noise surrogate (blue).

(b) The distance measure in the 5-dimensional space between the original signal and the white-noise surrogate (blue).

Fig. 7: Dry cleaners ($\tau = 1$ and $D = 1$).

system [7]. In this study, we apply existing and novel methods of RQA parameter optimisation to smart energy systems. The time delay and embedding dimension are successfully determined through the mutual information and Cao's algorithm respectively. For the optimal radius, we propose two new methods, based on the simultaneous discrimination power of several RQA variables between the original signal and a noise surrogate.

We validate our results using a well-studied Lorenz system, where our findings are consistent with the previous literature. Our main interest is in the new application, energy usage data, where both methods 1 and 2 behave qualitatively as expected, producing similar results. The discrimination power reaches its maximum value at a relative low value of the radius and the REC percentage and then decreases as the radius increases

due to the occurrence of false recurrences in the surrogates. Note that the optimal radius value depends on the embedding dimension, as expected. We leave further research to better understand this dependence for future studies.

REFERENCES

[1] J.-P. Eckmann, S. O. Kamphorst, and D. Ruelle, "Recurrence plots of dynamical systems," *EPL (Europhysics Letters)*, vol. 4, no. 9, p. 973, 1987.
[2] J. Zbilut and C. Webber, "Embeddings and delays as derived from quantification of recurrence plots," *Physics Letters A*, vol. 171, no. 3, pp. 199–203, 1992.

[3] A. Fabretti and M. Ausloos, "Recurrence plot and recurrence quantification analysis techniques for detecting a critical regime," *International Journal of Modern Physics C*, vol. 16, no. 05, pp. 671–706, 2005.

[4] N. Marwan, S. Schinkel, and J. Kurths, "Recurrence plots 25 years later-gaining confidence in dynamical transitions," *EPL (Europhysics Letters)*, vol. 101, no. 2, p. 20007, 2013.

[5] J. Zbilut and C. Webber, "Recurrence quantification analysis," in *Wiley Encyclopedia of Biomedical Engineering*. John Wiley & Sons, Inc., 2006.

[6] L. Hattam and D. Vukadinović Greetham, "Energy disaggregation for SMEs using Recurrence Quantification Analysis," in *Proceedings of the Ninth International Conference on Future Energy Systems*, ser. e-Energy '18. New York, NY, USA: ACM, 2018, pp. 610–617. [Online]. Available: http://doi.acm.org/10.1145/3208903.3210280

[7] N. Marwan, "How to avoid potential pitfalls in recurrence plot based data analysis," *International Journal of Bifurcation and Chaos*, vol. 21, no. 04, pp. 1003–1017, 2011. [Online]. Available: https://www.worldscientific.com/doi/abs/10.1142/S0218127411029008

[8] M. Zeifman and K. Roth, "Nonintrusive appliance load monitoring: Review and outlook," *IEEE Trans. on Consum. Electron.*, vol. 57, no. 1, pp. 76–84, Feb. 2011. [Online]. Available: http://dx.doi.org/10.1109/TCE.2011.5735484

[9] P. Klemenjak, C.and Goldsborough, "Non-intrusive load monitoring: A review and outlook," in *INFORMATIK 2016*. Gesellschaft fr Informatik e.V., 2016, pp. 2199–2210.

[10] G. W. Hart, "Nonintrusive appliance load monitoring," *Proceedings of the IEEE*, vol. 80, no. 12, pp. 1870–1891, Dec 1992.

[11] M. Aiad and P. H. Lee, "Unsupervised approach for load disaggregation with devices interactions," *Energy and Buildings*, vol. 116, pp. 96 – 103, 2016. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0378778815304795

[12] O. Parson, S. Ghosh, M. Weal, and A. Rogers, "Non-intrusive load monitoring using prior models of general appliance types," in *Proceedings of the Twenty-Sixth AAAI Conference on Artificial Intelligence*, ser. AAAI'12. AAAI Press, 2012, pp. 356–362. [Online]. Available: http://dl.acm.org/citation.cfm?id=2900728.2900780

[13] J. Z. Kolter and T. Jaakkola, "Approximate inference in additive factorial hmms with application to energy disaggregation," in *Proceedings of the Fifteenth International Conference on Artificial Intelligence and Statistics*, ser. Proceedings of Machine Learning Research, N. D. Lawrence and M. Girolami, Eds., vol. 22. La Palma, Canary Islands: PMLR, 21–23 Apr 2012, pp. 1472–1482. [Online]. Available: http://proceedings.mlr.press/v22/zico12.html

[14] J. Kelly and W. Knottenbelt, "Neural NILM: Deep Neural Networks applied to Energy Disaggregation," in *Proceedings of the 2Nd ACM International Conference on Embedded Systems for Energy-Efficient Built Environments*, ser. BuildSys '15. New York, NY, USA: ACM, 2015, pp. 55–64. [Online]. Available: http://doi.acm.org/10.1145/2821650.2821672

[15] C. Liu, A. Akintayo, Z. Jiang, G. P. Henze, and S. Sarkar, "Multivariate exploration of non-intrusive load monitoring via spatiotemporal pattern network," *Applied Energy*, vol. 211, pp. 1106 – 1122, 2018. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261917317397

[16] D. Egarter, V. P. Bhuvana, and W. Elmenreich, "Paldi: Online load disaggregation via particle filtering," *IEEE Transactions on Instrumentation and Measurement*, vol. 64, no. 2, pp. 467–477, Feb 2015.

[17] M. Schulze, H. Nehler, M. Ottosson, and P. Thollander, "Energy management in industry a systematic review of previous findings and an integrative conceptual framework," *Journal of Cleaner Production*, vol. 112, pp. 3692–3708, 2016.

[18] E. I. Vlahogianni and M. G. Karlaftis, "Comparing traffic flow time-series under fine and adverse weather conditions using recurrence-based complexity measures," *Nonlinear Dynamics*, vol. 69, no. 4, 2012.

[19] H. Ding, S. Crozier, and S. Wilson, "Optimization of Euclidean distance threshold in the application of recurrence quantification analysis to heart rate variability studies," *Chaos, Solitons & Fractals*, vol. 38, no. 5, pp. 1457 – 1467, 2008. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S096007790800146X

[20] H. Kantz and T. Schreiber, *Nonlinear Time Series Analysis*. Cambridge, UK: Cambridge University Press, 2003.

[21] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geomet-rical construction," *Phys. Rev. A*, vol. 45, pp. 3403–3411, Mar 1992. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevA.45.3403

[22] L. Cao, "Practical method for determining the minimum embedding dimension of a scalar time series," *Physica D: Nonlinear Phenomena*, vol. 110, no. 1, pp. 43 – 50, 1997. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0167278997001188

[23] J. P. Zbilut, J.-M. Zaldivar-Comenges, and F. Strozzi, "Recurrence quantification based Lyapunov exponents for monitoring divergence in experimental data," *Physics Letters A*, vol. 297, no. 3, pp. 173 – 181, 2002. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S037596010200436X

[24] S. Schinkel, O. Dimigen, and N. Marwan, "Selection of recurrence threshold for signal detection," *The European Physical Journal Special Topics*, vol. 164, no. 1, pp. 45–53, Oct 2008. [Online]. Available: https://doi.org/10.1140/epjst/e2008-00833-5

[25] D. Yang, W.-X. Ren, Y.-D. Hu, and D. Li, "Selection of optimal threshold to construct recurrence plot for structural operational vibration measurements," *Journal of Sound and Vibration*, vol. 349, pp. 361 – 374, 2015. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0022460X15002850

[26] Z. Gao and N. Jin, "Flow-pattern identification and nonlinear dynamics of gas-liquid two-phase flow in complex networks," *Phys. Rev. E*, vol. 79, p. 066303, Jun 2009. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.79.066303

[27] E. Lorenz, "Deterministic non-periodic flow," *Journal of Atmospheric Science*, vol. 20, pp. 130–141, 1963.

[28] P. beim Graben and A. Hutt, "Detecting recurrence domains of dynamical systems by symbolic dynamics," *Phys. Rev. Lett.*, vol. 110, p. 154101, Apr 2013. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevLett.110.154101

[29] P. beim Graben, K. K. Sellers, F. Frählich, and A. Hutt, "Optimal estimation of recurrence structures from time series," *EPL (Europhysics Letters)*, vol. 114, no. 3, p. 38003, 2016. [Online]. Available: http://stacks.iop.org/0295-5075/114/i=3/a=38003