Accepted manuscript

Johannesen, N. J., Kolhe, M. L. & Goodwin, M. (2023). Vertical Approach Anomaly Detection Using Local Outlier Factor. In H. Haes Alhelou, N. Hatziargyriou & Z. Y. Dong (Eds.), Power Systems Cybersecurity (pp. 297–310). Springer. https://doi.org/10.1007/978-3-031-20360-2_12.

Is a part of: Power Systems Cybersecurity

DOI: https://doi.org/10.1007/978-3-031-20360-2_12

AURA: https://hdl.handle.net/11250/3126243

Copyright: © 2023 The Author(s)

License: Under exclusive license to Springer Nature Switzerland AG

Vertical approach Anomaly Detection using Local Outlier Factor

Nils Jakob Johannesen, Mohan Lal Kolhe, and Morten Goodwin

Department of Engineering and Science, Department of ICT,
University of Agder, PO Box 422, N-4604 Kristiansand, Norway
Department of Electrical Engineering, IT and Cybernetics
University of Southest-Norway, PO Box 4, N-3199 Borre, Norway
{nils.j.johannesen}@usn.no
{mohan.l.kolhe,morten.goodwin,nils.j.johannesen}@uia.no
http://www.usn.no http://www.uia.no

Abstract. Detection of anomalies based on smart meter data is crucial to identify potential risks and unusual events at an early stage. In addition anomaly detection can be used as a tool to detect unwanted outliers, caused by operational failures and technical faults, for the preprocessing of data for machine learning, to detect concept drift as well as enhancing cyber-security in smart electrical grid operations. It is known that anomalies are defined through their contextual appearance. Hence, anomalies are divided into point, conceptual and contextual anomalies. In this work the contextual anomaly detection is examined, through a novel type of load forecasting known as vertical approach. This chapter explores the use of anomaly detection in the relevant learning systems for machine learning in smart electrical grid operation and management through data from New South Wales region in Australia. The presented vertical time approach uses seasonal data for training and inference, as opposed to continuous time approach that utilizes all data in a continuum from the start of the dataset until the time used for inference. It is observed that Local Outlier Factor identifies different local outliers given different vertical approaches. In addition, the local outlier factor score vary vertically. An anomaly is defined as a deviation from an established normal pattern. Spotting an anomaly depends on the ability to defy what is normal. Anomaly detection systems aim at finding these anomalies. Anomaly detection systems are in high demand, despite the fact that there is no clear validation approach. These systems rely on deep domain expertise.

Keywords: cybersecurity, anomaly detection, smart grid, local outlier factor

1 Introduction

MACHINE learning (ML) can provide electrical load demand forecasting, giving information about future loads, which provides essential input to other applications such as Demand Response, Topology Optimization and Anomaly Detection, facilitating the integration of intermittent clean energy sources.

An anomaly is defined as a deviation from an established normal pattern. Spotting an anomaly depends on the ability to defy what is normal. Anomaly detection systems aim at finding these anomalies. Anomaly detection systems are in high demand, despite the fact that there is no clear validation approach. These systems rely on deep domain expertise.

In the safe operation and management of the smart grid there is a need for efficient and reliable detection of anomalies. The data used in grid operation is of such an amount, that it is not possible to do so manually or by visual inspection alone, and there is a need for efficient, automated and accurate anomaly detection methods [1]

The available advanced information and communicating platform and computational capability renders smart grid prone to attacks with extreme social, financial and physical effects. The smart grid concept enables the utilization of smart appliances in homes and electric vehicles for providing support for frequency regulation and voltage regulation. Cyber threats could affect the ancillary services that are being delivered from the aggregators, which might lead to stability and security issues resulting in brownout or massive blackouts [2].

To coordinate and manage the increase of renewable energy sources, such as wind, sun and hydro, they can be operated using gried-tied voltage source converters (VSCs) [3]. VSCs regulate voltage and frequency locally. The VSCs enable the operation of intelligent microgrids (MGs), and vulnerable for attack. In the distributed power network the attack can disrupt the frequency regulation, voltage stability and the power flow management [4].

The implementation of two way communication by the use of sensors and intelligent agents such as advanced metering infrastructure (AMI) as well as load aggregation, make these attractive objects for cyber attacks. Sensors can be penetrated using a Trojan Horse, to manipulate the adversary inside the control platform, and change reference inputs in both outer and secondary control for VSCs. The attacker can here change acquisition gains, that create bias in the measurements report.

The enhanced agent topology of a smart electrical grid, reveals the mentioned vulnerabilities. The valuable question is; what is the price of the smartness of the smart electrical grid, in terms of the security of the supply?

Anomaly detection is an important first step in supervised machine learning processing to search out erroneous data, where the data was recorded due to an error or disrupted by other causes. Examples of such erroneous data could be the product of a sensor fault, downtime due to maintenance or when time series data is recorded in the changes from daylight saving time. This first step in data cleaning process has been known to enhance any forecasting algorithm [5][6].

From the field of Time Series Analysis and different correlation studies, it is known that seasonality affects the electrical load consumption. Time series analysis includes this knowledge in parametric methods such as Seasonal AutoRegressive Moving Average [7] [8]. The authors of this book chapter has also included drift due to seasonal variations, in the method vertical axis approach [9][10]

In the Section 2 of this book chapter the different learning systems in anomaly detection are explained, as well as the main categories of anomalies. Section 3 covers the literature review. Section 4 gives the mathematical foundation for LOF. Section 5 presents the methodology used, and finally in Section 6 are given the results and conclusion.

2 Learning systems in Anomaly Detection

Machine learning algorithms are divided into unsupervised, supervised and semi-supervised learning. Unsupervised learning is when the algorithm is learning without knowing the target for its learning. The algorithm is trying to make sense amongst the features of the data, as in discovering natural groups within the data by clustering techniques. In supervised learning the algorithm know the target of prediction during the training stage of the algorithm. When training the algorithm features are divided into dependent and independent features. The independent features involves the basis of the decision making, the input vector, and the dependent variable is the target vector. In supervised learning the patters between input and output is learned in the training phase and based on this result the models parameters are identified. Semi-supervised learning method is a hybrid between the mentioned supervised and unsupervised learning, and it involves elements from both supervised and unsupervised learning-Semi-supervised learning uses partially labeled and unlabeled instances to detect anomalies. One instance is in autoencoders where only the data that depicts normal behavior are labeled and used to train the algorithm. Based on the assumption of what is normal it will flag anomalous datapoints when exposed to data that differs from normal data [11]

To learn patterns across timescales, sliding window technique is introduced. Sliding window method is used when doing time series forecasting, and it introduces a shift in the input and output variable relation. This can be done through dynamic or static modeling. For this work we will explain the static modeling of sliding window method.

In the Sliding window method the training and test set input and target vectors timestamp is shifted using a sliding technique that allows the number of timesteps shifted, to equal the predictive window. In practical terms it means that the algorithm reads a features input for a certain timeframe and then matches it with the selected predictive window. In the training phase of the modelling, the algorithm will then learn the patterns that suggest the following behaviour ahead in time. In sliding window adaptation the 'Feature Data' columns is the input (time of day, dry-bulb temperature, humidity, previous loads, etc.) and the target vector is 'Load Data', and the 'Load Data' is shifted 48 timesteps back (half-hourly values), and this is how the algorithm is fed the information when doing day ahead forecasting.

The sliding window method explained here is used in supervised learning for time series. Alternative versions of sliding window method have been adopted to deal with concept drift. In machine learning concept drift occurs when the underlying distribution of the data changes over time changes over time, making the model unfit to predict for future events [12]. An adaptive sliding window method, that calculates an adoptive window-size on the fly is has been proposed to deal with concept drift [13]. Another adaptation raises the complexity by using multi sliding window detecting growth length over several windows detecting the drift length by adjoining several windows finding the optimal window length useful for online learning [14].

In Recurrent Neural Networks like Long Short-term Memory networks (LSTM) different gates are used to remember and forget time-occurrence over different time windows. To deal with concept drift in LSTMs it has been proposed a novel forgetting mechanism for anomaly detection [15]. It has also been proposed to narrow down the scope, by critical lines detecting distribution change. The first step in the proposed method is to reduced dimensionality through an orthogonal transformation of the data reducing the feature space to its principal components [16]. After reducing the feature space the distributions are compared by two-sample Kolmogorov–Smirnov Test (KS test) [17].

Anomalies depend on the structure, distribution and type of data, as well as carnality of relationship of the data [18]. Basically there are 3 main types of anomalies: Point, Contextual and Collective anomalies. Point anomaly is an instant, that can be regarded anomalous, amongst other instances in the data. They are extreme, and caused instantly. In electric load demand point anomalies when other sources of energy (such as district heating) are curtailed [19]. Collective Anomaly is an instance that is defined an anomaly based on the context. Meaning that, in some context (e.g based on the temperature of a season), a particular behaviour is normal, but put in a different context (a different season) it qualifies as abnormal behaviour. Collective anomalies are data instances that deviate significantly from the entire dataset, but individual data amongst the collective instances may not qualify as an outlier.

3 Literature Review

Anomaly detection is done on any time series data. Robust statistical methods have been known to determine deviation from normal electrical load patterns. Simple statistical methods, such as Box-plots have been used in the preprocessing of electrical consumption load profiles to determine daily usage patterns [19]. Box-plot have also been used to identify anomalies on regression forecast errors in order to improve the prediction [20].

Efforts to find data-mining framework to typical electricity load patterns (TELP). TELP has been proven successful for anomaly detection in builduing electricity consumption data. The first step in TELP is to cluster data in temporal segments, such as daily electricity load profiles (DELP) using density based spatial clustering application with noise (DBSCAN). The framework is aimed to identify typical electricity load patterns and gain knowledge hidden in the patterns and to potentially be used in an early fault detection of anomalous electricity load profiles [21]. Also to detect anomalies of electricity consumption in office buildings an improved kNN is proposed, ikNN, to automatically classify consumption footprints as normal or abnormal [22].

Other techniques for identifying patterns in electricity consumption, are determining distinctive clusters within seasons, and the obtained clusters are used to create seasonal curves for the four seasons, with each having two optimal clusters representing the load demand [23]. An autoregressive integrated moving average with exogenous inputs (ARIMAX) model is used to extract weather dependency to find the residuals, then through hypothesis testing the extremities, maximum and minimums are found [24]. This procedure was reproduced, with linear regression finding the residuals and a Bayesian maximum likelihood classifier to identify anomalies [1]. Dynamic Bayesian Networks and Restricted Boltzman Machine has been proposed for anomaly detection in large-scale smart grids. Simulated on the IEEE 39, 118, and 2848 bus systems the results were verified [25]. Real-Time Mechanism for detecting false data injection attacks analyzed the change of correlation between two phasor measurement units parameters using Pearson correlation coefficient on IEEE 118 and 300-bus systems [26].

Machine learning techniques have been highlighted for their ability to differentiate between cyber-attacks and natural disturbances. By a simulating a variety of scenarios the ability for One R, Random Forest, Naive Bayes and J-Ripper to recognize attacks was investigated: Short Circuit faults; location is represented by the percentage range, Line maintenance; identified through remote relay trip command, Remote tripping command injection; the attacker operates the relay remotely that causes a breaker to open, Relay setting change; the attacker misconfigures the relay settings to cause maloperation of relays, FDIA; attacker manipulates measurements sensors. The simulated scenarios was grouped into classes; natural events, attack events, and no events [27].

4 Local Outlier Factor for Anomaly Detection

Local Outlier Factor (LOF) is a density based unsupervised anomaly detection algorithm introduced in 2000 [28]. LOF compares the local density of a point to the local density of k of its neighbors. By comparing the local density of a point to the local density of its neighbors one can identify point that have substantially lower density than its neighbors. These points are considered to be outliers.

The first step in LOF is to compute all distances. As shown for in (Equation 1) for an m-dimensional Euclidean space:

$$d(x_i, x_j) = \sqrt{\sum_{i,j=1}^{n} (x_i - x_j)^2} \tag{1}$$

In an example dataset consisting of n=7 points, as shown in Fig. 1, and Table 1.

The distance x_i to x_j between B and D is:

$BD^2 = (B_x - D_x)^2 + (B_y - D_y)^2$

$BD^2 = (1 - 4)^2 + (5 - 1)^2$

$BD^2 = 25$

$BD = \sqrt{25} = 5$

The distances k to all point in the neighborhood of x, $N_{k(x_i)}$. All distances are computed in order to define the nearest neighbors distances N_k:

- AB: 6.08 AC: 6.32 AD: 5.83 AE: 6.4 AF: 7.21 AG: 6.71
- BA: 6.08 BC: 1.0 BD: 5.0 BE 4.47 BF: 5.39 BG: 5.83
- CA: 6:32 CB: 1.0 CD: 4.24 CE 3.61 CF: 4.47 CG: 5.0
- DA: 5.83 DB: 5.0 DC: 4.24 DE 1.0 DF: 1.41 DG: 1.0
- EA: 6.4 EB: 4.47 CC: 3.61 ED: 1.0 EF: 1.0 EG: 1.41
- FA: 7.21 FB: 5.39 FC: 4.47 FD: 1.41 FE 1.0 FG: 1.0
- GA: 6.71 GB: 5.83 GC: 5.0 GD: 1.0 GE 1.41 GF: 1.0

After computing all the distances the values are sorted from nearest in row 1, see Table 2, in ascending order to the 6th nearest neighbor in row 6. From the N_k-matrix in Table 2 reveal some intuition for the k-parameterisation. The value of k is also controlling the smoothing strength of LOF [29]. It is recommended choosing a minimum k and a maximum k, and for each point, taking the maximum LOF value over each k in that range. Several guidelines for choosing the bounds are given. For the minimum value, the LOF values fluctuate the points in a uniform distribution for k<10, with points in a uniform distribution sometimes

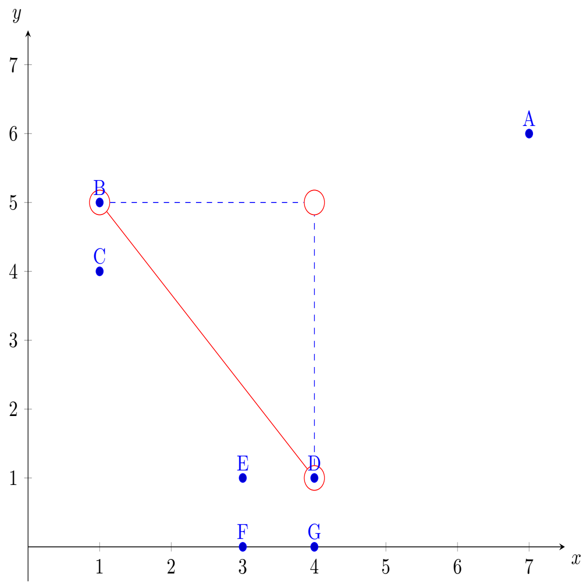

Fig. 1: An example dataset

x	y	Letter
7	6	A
1	5	B
1	4	C
4	1	D
3	1	E
3	0	F
4	0	G

Table 1: An example dataset

A	B	C	D	E	F	G	k
5.83	1	1	1	1	1	1	1
6.08	4.47	3.61	1	1	1	1	2
6.32	5	4.24	1.41	1.41	1.41	1.41	3
6.4	5.39	4.47	4.24	3.61	4.47	5	4
6.71	5.83	5	4.47	4.47	5.39	5.83	5
7.21	6.08	6.32	5	6.4	7.21	6.71	6

Table 2: All nearest distances in ascending order

showing up as outliers. Secondly, the minimum k-value serves as a minimum size to be considered a "cluster", and points can be outliers relative to that cluster. If k=20, and you have a group of 10 points and a point X, each point in the group will include X in its nearest neighbors, and x will include those points, leading them to have very similar LOFs.

For the maximum value, a similar criteria applies, in that it should be the maximum number of objects that you want to be considered outliers if clustered together. A group of N objects isolated from the main set can either be a cluster, or N outliers; for k<N, they will be the first; for k>N, they will be the second.

Once the optimal k-value is found the LOF uses the reachability distance (RD) to compute the LOF of a point x:

$$\text{Reachability Distance}(x_i, x_j) = max(k - distance(x_j), dist(x_i, x_j)) \tag{2}$$

If x_i is further away from x_j, than x_j's k^{th} nearest neighbor, then x_i is the RD, if the opposite, that x_j's k^{th} nearest neighbor the actual distance between x_i and x_j then $dist(x_i, x_j)$ is used as RD, this is known as the smoothing factor.

When all RD's are computed the next step is to find the Local Reachability Density (LRD). This is found by taking the inverse of the average RD:

$$LRD(x_i) = \left(\sum_{x_j \epsilon N(x_i)} \left\{ \frac{RD(x_i, x_j}{|N(x_i)|} \right\} \right)^{-1} \tag{3}$$

If the $LRD_{(x_i)}$ is high then the point x_i is in a dense neighborhood, and opposite, when $LRD_{(x_i)}$ is low, then x_i is in a sparse neighborhood. The LRD's of the points in $N_k x_i$, multiplied by the LRD of x_j is used to compute the LOF of point x_i:

$$LOF(x_i) = \frac{\sum_{x_j} \epsilon N(x_i) LRD(x_j)}{|N(x_i)|} * \frac{1}{LRD(x_i)} \tag{4}$$

If the average value of the LRD's in $N_k x_i$ is large and the LRD of x_i is small, then x_i is an outlier.

5 Methodology

In this work data from New South Wales, Sydney region, is considered. New South Wales, Sydney region electrical load profile data set [30] includes meteorological parameters (e.g. DryBulb and WetBulb Temperature, Humidity, Electricity price and time of use) [31]. Data is gathered from 2006-2011. The overall energy mix in New South Wales consists mainly of Coal, Natural Gas, Hydro and other renewable energy sources as shown in Table 3.

Power plant	Number	Installed Power (MW)
Hydro	24	4794
Wind	14	1250
Solar	9	228
Coal	8	11730
Biogas	11	56
Natural Gas	19	3766

Table 3: Energy Mix in New South Wales, Australia

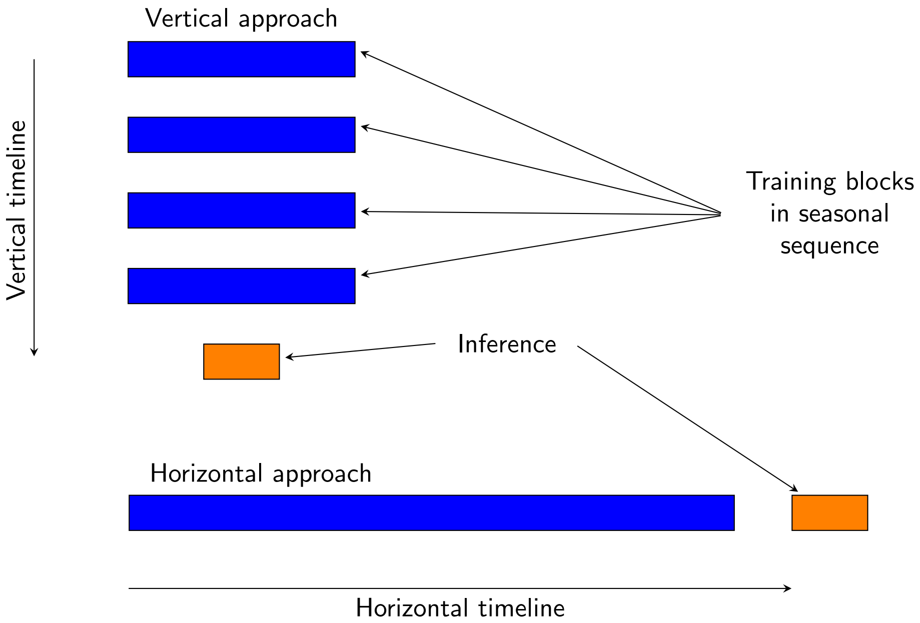

Fig. 2: Vertical and Horizontal Approach

The methodology in this work is based on a comprehensive correlation analysis on the impact of external parameters on electrical load demand [7][8][9][10]. It is observed from these analyses that that meteorological temperatures are highly correlating to the electrical load demand. The vertical time approach uses seasonal data for training and inference. The horizontal approach uses continuous datasets, i.e., it utilizes all data in a continuum from the start of the dataset until the time period used for inference. The illustration of horizontal and vertical approaches is presented in Fig. 2.

Vertical approach can be performed with minimum amount of data compared to continuous approach. Also, the vertical time approach predictive results are compared with prediction based on continuous time series data. In vertical approach, the training set, $D = \{x_i\}_{i=1}^{N}$, is partitioned into subsets by each season

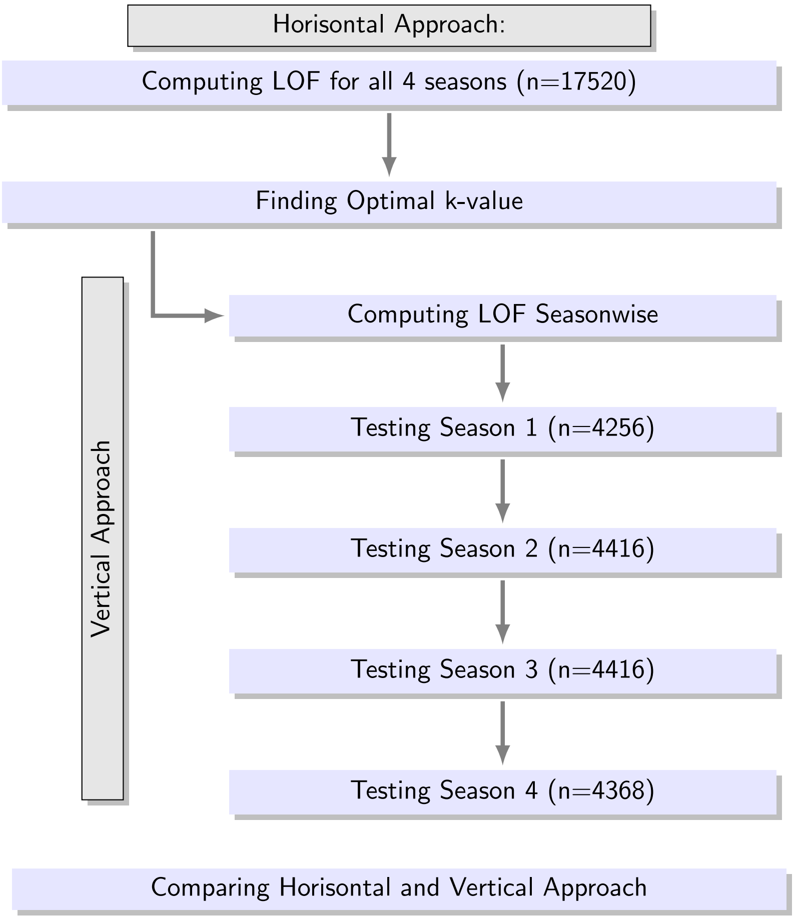
Horisontal Approach:
Computing LOF for all 4 seasons (n=17520)
Finding Optimal k-value
Vertical Approach
Computing LOF Seasonwise
Testing Season 1 (n=4256)
Testing Season 2 (n=4416)
Testing Season 3 (n=4416)
Testing Season 4 (n=4368)
Comparing Horisontal and Vertical Approach

Fig. 3: Flowchart of the proposed model

of the year, and then are merged together only containing seasonally information about the load pattern. In a dataset containing time observation for five years (e.g., 2016–2020), time is separately selected season-wise, and then merged to contain only the specific season for training, $D = \{x_{spring_i}\}_{i=2016}^{2019}$.

In this study, seasons are divided by months, where Season 1 is December, January, February, and Season 4 is September, October, November. The LOF, is compared for the Horizontal and Vertical Approach, as seen in the flowchart presented in Fig. 3. First the LOF is computed on all 4 seasons, finding the optimal value of k, to be compared seasonwise, independently on each season, and finally comparing the percentage of detected anomalies.

6 Results, Discussion and Conclusion

The first results from step one in the flowchart for the proposed model, in Fig. 3, are illustrated in Fig. 4: The results show all the outliers from the horizontal approach, the red rings encircling these LOFs in, Fig. 4, are the radius used to describe their individual LOFs, from the threshold value of LOF <-1.5 to the maximum LOF=-2.93. This maximum LOF is recorded on the 15 th of september 2010 at 03:30 hours, and recorded a system load of 6882.9 MW. From these results the upper and lower bounds are chosen.

The data is computed from LOF using k-nearest neighbor of 6, 8 and 10, and showng 3 different LOF's (LOF <-1.5, LOF <-2 and LOF <-2.5), shown in Table 4. The results show that as the search space widens the fraction of outliers detected decreases. In the first row in Table 4 The fractions continues to decrease as for the lower level of LOF. For k=10, the fraction of LOF is halfed when regarding the -1.5 threshold compared to the -2.5 threshold.

Observing the results from the LOF's for the different k-values; when comparing for horizontal and vertical approach most of the results show that vertical approach detects a bigger fraction of anomalies. This is valid for 25 out of 32 results a fraction of 78 %. For LOF <-1.5 and k=6, k=8, k=10, for LOF <-2 and k=6, K=8, k=10 and finally for LOF <-2.5 and k=6, vertical approach detects more anomalies than horizontal approach.

Observing the results from horizontal approach to vertical approach on Season One, in 7 out of 9 test results the fraction of detected anomalies are higher using vertical approach. In Season Two this fraction is reduced to 4 out of 8 (since the k=8 for LOF<-2 has equal values). In Season Three vertical approach has a higher fraction on 100 % of the test results, compared to horizontal approach. In Season Four in 6 out of 9 test results the fraction of detected anomalies are higher using vertical approach. These results verify observations from analysing the seasonal impact on the load electrical load demand.

Season One has a relatively higher electrical load demand as compared to Season Three. These two seasons stand out in the detection of outlier and the explanation for this is found in the correlation to external parameters. In Season Three (June-July-August) dry bulb temperatures recorded are so low that the demand for heating increases the electrical load demand. This might be an explanation to the many outliers found in Season Three. Season One has a lower load demand, but the visual inspection shows higher occasional peaks in the demand. By visual inspection, Fig 3, non of these peaks are detected by horizontal approach.

This work reflects the previous extensive correlational analysis as well as algorithmic development designed to option for changing seasonal behaviour due to the impact of external parameters. It is shown that the number and fraction of detected outliers are higher when using vertical approach, as this work proposes,

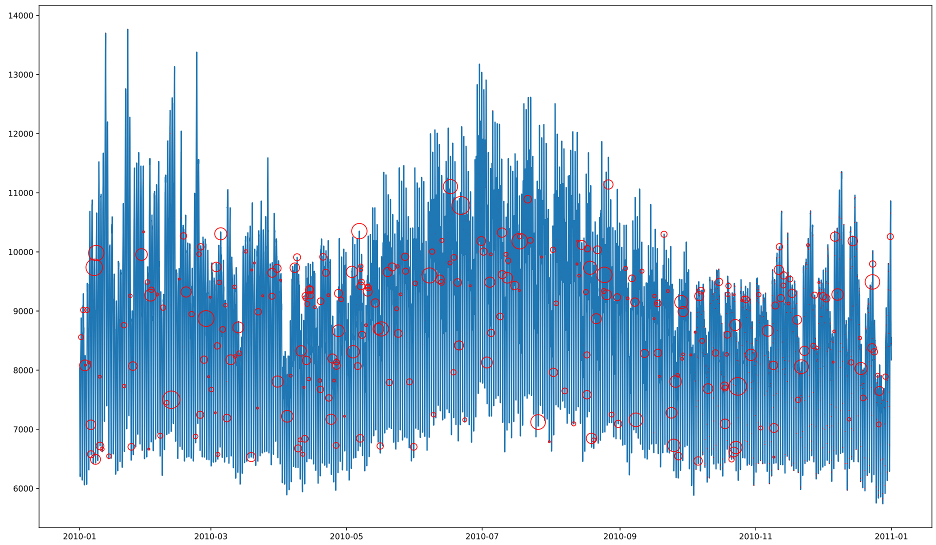

Fig. 4: Outliers detected varying radius of the red encirclement show their individual LOFs ranging from, LOF<-1.5 to the maximum LOF=-2.93.

Season	LOF <-1.5			LOF <-2			LOF <-2.5		
	k=6	k=8	k=10	k=6	k=8	k=10	k=6	k=8	k=10
All seasons	3.39	0.52	0.12	1.28	0.11	0.01	0.51	0.03	0.01
Season One	3.40	0.44	0.16	1.50	0.12	0.05	0.77	0.05	0
Season Two	3.54	0.59	0.09	1.26	0.11	0.02	0.46	0	0
Season Three	3.88	0.73	0.14	1.52	0.18	0.05	0.75	0.05	0.05
Season Four	3.34	0.60	0.21	1.09	1.09	0.07	0.46	0.21	0.12

Table 4: Results

compared to the traditional continuous method using horizontal approach. There is a need for evolving the impact of external parameters and seasonal behavioural patterns on electrical load demand to enhance the outlier detection for smart energy systems.

References

1. H. N. Akouemo and R. J. Povinelli, "Probabilistic anomaly detection in natural gas time series data," *International Journal of Forecasting*, vol. 32, no. 3,

pp. 948–956, 2016. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S016920701500076X

2. K. Huang, C. Zhou, Y.-C. Tian, S. Yang, and Y. Qin, "Assessing the physical impact of cyberattacks on industrial cyber-physical systems," *IEEE Transactions on Industrial Electronics*, vol. 65, no. 10, pp. 8153–8162, 2018.
3. F. Blaabjerg, R. Teodorescu, M. Liserre, and A. V. Timbus, "Overview of control and grid synchronization for distributed power generation systems," *IEEE Transactions on Industrial Electronics*, vol. 53, no. 5, pp. 1398–1409, 2006.
4. S. Sahoo, T. Dragičević, and F. Blaabjerg, "Cyber security in control of grid-tied power electronic converters–challenges and vulnerabilities," *IEEE Journal of Emerging and Selected Topics in Power Electronics*, pp. 1–1, 2019.
5. I. Chang, G. C. Tiao, and C. Chen, "Estimation of time series parameters in the presence of outliers," *Technometrics*, vol. 30, no. 2, pp. 193–204, 1988. [Online]. Available: http://www.jstor.org/stable/1270165
6. N. J. Johannesen, M. Kolhe, and M. Goodwin, "Deregulated electric energy price forecasting in nordpool market using regression techniques," in *2019 IEEE Sustainable Power and Energy Conference (iSPEC)*, 2019, pp. 1932–1938.
7. ——, "Comparison of regression tools for regional electric load forecasting," in *2018 3rd International Conference on Smart and Sustainable Technologies (SpliTech)*. IEEE, 2018, pp. 1–6.
8. N. J. Johannesen and M. L. Kolhe, "Application of regression tools for load prediction in distributed network for flexible analysis," in *Flexibility in Electric Power Distribution Networks*. CRC Press, 2021.
9. N. J. Johannesen, M. Kolhe, and M. Goodwin, "Relative evaluation of regression tools for urban area electrical energy demand forecasting," *Journal of Cleaner Production*, vol. 218, pp. 555–564, 2019. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652619301192
10. N. J. Johannesen, M. L. Kolhe, and M. Goodwin, "Smart load prediction analysis for distributed power network of holiday cabins in norwegian rural area," *Journal of Cleaner Production*, vol. 266, p. 121423, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0959652620314700
11. S. Alla and S. K. Adari, *Beginning anomaly detection using python-based deep learning*. Springer, 2019.
12. J. a. Gama, I. Žliobaitundefined, A. Bifet, M. Pechenizkiy, and A. Bouchachia, "A survey on concept drift adaptation," *ACM Comput. Surv.*, vol. 46, no. 4, mar 2014. [Online]. Available: https://doi.org/10.1145/2523813
13. S. Q. Du, Lei and X. Jia, "'detecting concept drift: An information entropy based method using an adaptive sliding window'." *Intelligent Data Analysis*, vol. 18, no. 3, pp. pp. 337–364, june 2014.
14. H. Guo, H. Li, Q. Ren, and W. Wang, "Concept drift type identification based on multi-sliding windows," *Information Sciences*, vol. 585, pp. 1–23, 2022. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0020025521011439
15. R. Xu, Y. Cheng, Z. Liu, Y. Xie, and Y. Yang, "Improved long short-term memory based anomaly detection with concept drift adaptive method for supporting iot services," *Future Generation Computer Systems*, vol. 112, pp. 228–242, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0167739X20302235
16. N. J. Johannesen, M. L. Kolhe, and M. Goodwin, "Comparing recurrent neural networks using principal component analysis for electrical load predictions," in *2021 6th International Conference on Smart and Sustainable Technologies (SpliTech)*. IEEE, 2021, pp. 1–6.

17. M. Mohammadpourfard, Y. Weng, M. Pechenizkiy, M. Tajdinian, and B. Mohammadi-Ivatloo, "Ensuring cybersecurity of smart grid against data integrity attacks under concept drift," *International Journal of Electrical Power Energy Systems*, vol. 119, p. 105947, 2020. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0142061519331904
18. R. Foorthuis, "On the nature and types of anomalies: a review of deviations in data," *International Journal of Data Science and Analytics*, vol. 12, pp. 461–478, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S1364032113007211
19. J. E. Seem, "Using intelligent data analysis to detect abnormal energy consumption in buildings," *Energy and Buildings*, vol. 39, no. 1, pp. 52–58, 2007. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378778806001514
20. G. F. Martin Nascimento, F. Wurtz, P. Kuo-Peng, B. Delinchant, and N. Jhoe Batistela, "Outlier detection in buildingsrsquo; power consumption data using forecast error," *Energies*, vol. 14, no. 24, 2021. [Online]. Available: https://www.mdpi.com/1996-1073/14/24/8325
21. X. Liu, Y. Ding, H. Tang, and F. Xiao, "A data mining-based framework for the identification of daily electricity usage patterns and anomaly detection in building electricity consumption data," *Energy and Buildings*, vol. 231, p. 110601, 2021. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0378778820333879
22. Y. Himeur, A. Alsalemi, F. Bensaali, and A. Amira, "Smart power consumption abnormality detection in buildings using micromoments and improved k-nearest neighbors," *International Journal of Intelligent Systems*, vol. 36, no. 6, pp. 2865–2894, 2021. [Online]. Available: https://onlinelibrary.wiley.com/doi/abs/10.1002/int.22404
23. J. D. Rhodes, W. J. Cole, C. R. Upshaw, T. F. Edgar, and M. E. Webber, "Clustering analysis of residential electricity demand profiles," *Applied Energy*, vol. 135, pp. 461–471, 2014. [Online]. Available: https://www.sciencedirect.com/science/article/pii/S0306261914009349
24. H. N. Akouemo and R. J. Povinelli, "Time series outlier detection and imputation," *2014 IEEE PES General Meeting — Conference & Exposition*, pp. 1–5, 2014.
25. H. Karimipour, S. Geris, A. Dehghantanha, and H. Leung, "Intelligent anomaly detection for large-scale smart grids," in *2019 IEEE Canadian Conference of Electrical and Computer Engineering (CCECE)*, 2019, pp. 1–4.
26. Y. He, G. J. Mendis, and J. Wei, "Real-time detection of false data injection attacks in smart grid: A deep learning-based intelligent mechanism," *IEEE Transactions on Smart Grid*, vol. 8, no. 5, pp. 2505–2516, 2017.
27. M. Panthi, "Anomaly detection in smart grids using machine learning techniques," in *2020 First International Conference on Power, Control and Computing Technologies (ICPC2T)*, 2020, pp. 220–222.
28. M. M. Breunig, H.-P. Kriegel, R. T. Ng, and J. Sander, "Lof: Identifying density-based local outliers," in *Proceedings of the 2000 ACM SIGMOD International Conference on Management of Data*, ser. SIGMOD '00. New York, NY, USA: Association for Computing Machinery, 2000, p. 93–104. [Online]. Available: https://doi.org/10.1145/342009.335388
29. O. Alghushairy, R. Alsini, T. Soule, and X. Ma, "A review of local outlier factor algorithms for outlier detection in big data streams," *Big Data and Cognitive Computing*, vol. 5, no. 1, 2021. [Online]. Available: https://www.mdpi.com/2504-2289/5/1/1

30. AEMO. (2021) National electricity market data - nem. [Online]. Available: https://aemo.com.au/energy-systems/electricity/national-electricity-market-nem/data-nem/aggregated-data
31. V. Dehalwar, A. Kalam, M. L. Kolhe, and A. Zayegh, "Electricity load forecasting for urban area using weather forecast information," in *2016 IEEE International Conference on Power and Renewable Energy (ICPRE)*, 2016, pp. 355–359.