Learning Dominant Usage from Anomaly Patterns in Building Energy Traces

Cristina Nichiforov, Grigore Stamatescu, Iulia Stamatescu, and Ioana Făgărăşan

***Abstract*— Building energy usage is growing at a rapid pace under increasing urbanisation tendencies in both the developing and the developed world, at high environmental and social costs. Decentralised control architectures for local energy grids are seen as a key solution to optimise energy management at the local level. As the infrastructure for data collection, communication and embedded computing becomes more capable, new online algorithms can be deployed for forecasting and anomaly detection of large consumers. Fine grained tendencies and unusual artefacts can be thus exploited to improve local and grid level energy management. Our two-fold approach first leverages the Matrix Profile technique for time series data mining to build a dataset of anomaly patterns from public building energy traces and extract analytics information. Subsequently the labeled dataset is used in a supervised learning classification model to discriminate between various related dominant usage patterns. The case study is carried out on a public dataset of academic buildings. The approach can prove useful for exploiting complementary energy consumption patterns in a decentralised control structure towards grid balancing and economic operation.**

I. INTRODUCTION

An important challenge in handling large quantities of data generated by modern measurement and control systems within the Industrial Internet of Things (IIoT) and Cyber-physical Systems (CPS) is to efficiently extract useful and actionable information. This can be used in a timely manner for online modeling with potential for energy savings and technical benefits such as grid stabilisation. Several IIoT platforms have been developed and made available for buildings such as [1] that have the ability to transform both modern and legacy buildings into ”smart” entities. The main goal is to extend legacy building management systems with open hardware components and software libraries for improved data analytics. Extended data collection from more and diverse buildings leads to improved algorithm performance while some building-level test-beds are becoming available as living labs and simulation platforms [2].

Many techniques are currently available for extracting information from building time series data such as sensor measurements and energy meter readings. These range from basic statistical indicators, time domain features and frequency domain features e.g. Fast Fourier Transform (FFT) and wavelet coefficients. The main argument is that, by working on aggregated features instead of the full input dataset, the algorithms will be sped up and made suitable for online operation. In our case we use a time series search and indexing algorithm called the Matrix Profile (MP) which returns a distance time series as result of an all pair subsequence search across the initial data. The profile is computed in an efficient manner suitable both for online operation and incremental learning. Applications are documented for subsequence clustering and motif discovery as well as anomaly detection. We use the respective discord indices to discriminate between dominant usage in large commercial buildings based on their electrical energy consumption fingerprint. This can be useful for higher level energy management strategies at the grid and city level that can exploit complementary patterns, allow peak shifting or other load balancing strategies for improved control.

Matrix Profile algorithm is basically a dimension reduction approach, the anomalies are quantified by their distance from existing data. The main hyperparameter of MP is the subsequence length. In case of deep learning techniques, for example Recurrent Neural Networks (RNN) can be used as a predictor, wherein the difference between the predicted next value and the actual value is used for anomaly detection. These are more sensitive of the availability of a large and diverse input dataset. A combination of MP and learning algorithm can yield good classification outcomes while reducing the time it takes to preprocess the data and for initial training and retraining of the models.

The main contributions of the paper are two-fold:

- an application of the Matrix Profile time series data mining techniques to large commercial building energy datasets;
- using the resulting profile characteristics, in particular, discord instances as proxies for usage-specific anomaly modeling in a learning framework.

Section II describes the technical context of related work given that data-driven methods for time series modeling and information extraction have emerged as a feasible solution for online operation. Section III discusses the methodology which revolves around the Matrix Profile technique for efficient time series data mining. The focus is on discord identification and classification across the used building dataset with four classes of dominant usage: classrooms, offices, laboratories and dormitories. Relevant experimental results are presented in Section IV mainly related to analytical insights and classification results. Section V concludes the paper and lists directions for future work.

*The work of Cristina Nichiforov has been funded by the Operational Programme Human Capital of the Ministry of European Funds through the Financial Agreement 51675/09.07.2019, SMIS code 125125.

The authors are with the Department of Automatic Control and Industrial Informatics, University Politehnica of Bucharest, 060042 Bucharest, Romania. {cristina.nichiforov, grigore.stamatescu, iulia.stamatescu, ioana.fagarasan}@upb.ro

II. RELATED WORK

Time series forecasting for energy applications is thoroughly reviewed in [3]. Three dominant methods are emphasised ranging from conventional (S)ARIMA models that offer good performance with extensive expert knowledge and limited generalisability, machine learning (ML) models with dedicated feature engineering that capture domain expertise and end-to-end deep learning models (DL), resulting in highly complex black-boxes. An example of DL approach is a two-level system with both RNNs for temporal level features, and CNNs for extracting spatial level features. To improve forecasting performance state of the art methods rely on hybrid models have emerged that try to combine the advantages of the most accurate methods and compensate their disadvantages in terms of generalisability and periodic retraining, mostly in multi-level data processing and modeling pipelines.

Unsupervised learning problems of building energy traces [4] are focused on discovering spurious and recurring patterns on historical data that can be used to improve diagnosis and control tasks. Energy analytics clustering and motif discovery is also an important at the system and grid level [5]. Large consumer entities can thus be logically grouped together to improve on energy management schemes such as demand response (DR) and load shaping (LS). At the subsystem level, data-driven AHU energy models are developed by [6]. As the heating/cooling subsystem is largely responsible for most of the energy consumption of a building [7], narrowing down the analysis to such units can improve the understanding of the underlying patterns with a small impact from neglecting the rest of the local consumers.

Availability of public benchmarking datasets on which ML/DL techniques are evaluated is highly important for replicability and scalability evaluation of the algorithms [8]. In the recent times many such data sets have been publicly disseminated by public authorities and building technology companies, either directly or through online competitions [9]. Other comparison in the security domain using MP derived features is presented in [10]. The relevance here stems from the periodic input time series that are analysed comparatively against a labelled data set of attacks, events or anomalies in our case, resulting in performance metrics for three types of methods on three different datasets.

Some previous related work has been targeted at building ARIMA [11] and NN [12] models for large commercial building energy forecasting. This has used conventional performance metrics such as MSE, RMSE, CV-RMSE and MAPE to establish an own baseline against which future improvements can be critically evaluated. NN and DL models for building energy forecasting are compared in [13]. These rely on new software libraries for efficient algorithm implementation and integration modules to retrieve the data from online repositories. We now extend our focus to anomaly detection and using the derived features for higher level tasks such as mapping the dominant usage of building clusters within a campus or a smart city.

III. METHODOLOGY

A. Matrix Profile

For feature extraction one promising approach is an online implementation of the Matrix Profile [14]. This represents a novel method for time series data processing. It consists of efficient search algorithms that use a sliding window mechanism to create a minimum distance profile of subsequences of length m over a time series of length n. The key speed-up that is argued for this stems from the usage of the Fast Fourier Transform for the z-normalised Euclidean distance computation in the form of the following formula [14]:

$$D[i] = \sqrt{(2m(1-(QT[i]-m\mu_Q\mu_T[i])/(m\sigma_Q\sigma_T[i])))} \tag{1}$$

Where D is the distance between two subsequences Q and T, using their dot product QT, m is the subsequence length, μ_Q is the mean of Q, μ_T the mean of T, σ_Q and σ_T are the standard deviations of Q and T respectively. In this sense the algorithm is useful for both static data and incremental modeling of streaming values with limited slow-down on even very large and multi-variate time series.

In its basic form a time series motif is the closest (non-trivial) pair of subsequences. The locations of the two minimum values of the matrix profile are identified through locations of the first motif pair. The subsequence that has the maximum distance to its nearest neighbor represents the time series discord. The discords basically capture the most unusual subsequence within a time series. They have several uses for data mining, but are particularly used as anomaly detectors because they only require a single parameter which is the subsequence length, unlike most algorithms used for anomaly detection that usually require more parameters. According to [15], a time series discord is defined as follows: considering a time series T, the subsequence Q of length n starting at position p represents the discord of T if Q has the largest distance to its nearest non-self match; this is ∀ subsequence C of a time series, non-self match M_Q of Q, and non-self match M_C of C, min(distance(Q, M_Q)) $>$ min(distance(C,M_C)). Also, in the current research it is also used the concept of K-th time series discord which has the following definition: the subsequence D of length n starting at position p of a time series T, is the K-th discord if D has the K-th largest distance to its nearest non-self match, with no overlapping region to the m-th discord starting at position p_m, for all 1≤m<K; this is $|p-p_m|\geq$ n. [14]

Besides the actual matrix profile, the matrix profile index is also constructed to return the actual position of the nearest neighbor subsequence. The variance of the profile can result in the complexity while their histogram returns the time series density estimation, also relevant for anomaly detection tasks. We use the time series discord indices to extract local anomaly energy consumption patterns that are input to the classification algorithm.

B. *Classification algorithms*

The goal of a classification model is to correctly assign test examples to predefined labels or classes. In our case the classes are defined as dominant usage patterns from the original dataset. These are: 1 - classrooms, 2 - offices, 3 - laboratories, 4 - dormitories. Through the classification model we aim at correctly mapping new energy traces to one of the dominant usage classes, using a limited input set of examples associated to the energy time series discords. The usual performance metrics for evaluating classification performance are: accuracy, precision i.e. accuracy of the positive predictions, recall i.e. the sensitivity, true positive rate, or the ratio of positive instances that are correctly detected by the classifier, f1-score as the harmonic mean of precision and recall. The confusion matrix is a tabular representation to assess the classifier performance and derive the metrics listed before [16]. Table I presents the structure of the confusion matrix for a binary (two-class) classification problem. The approach can be extended to a multi-class problem by replacing the positive and negative classes with the respective labels.

TABLE I: Confusion matrix

	Predicted-Positives	Predicted-Negatives
Actual-Positives	$true\ positives(TP)$	$false\ positives(FP)$
Actual-Negatives	$false\ negatives(FN)$	$true\ negatives(TN)$

1) Accuracy: represents the ratio of number of correct predictions to the total number of input samples:

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{2}$$

Accuracy is not a good metric for evaluating classifier performance in imbalanced datasets as by just predicting the dominant class with yield a floor performance value. Better classifiers can be however compared to this as a baseline.

2) Precision(P): represents the number of correct positive results divided by the number of positive results predicted by the classifier:

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

3) Recall(R): represents the number of correct positive results divided by the number of all relevant samples (all samples identified as actual positive):

$$Recall = \frac{TP}{TP + FN} \tag{4}$$

4) F1-score: represents the harmonic mean of precision and recall; the score is a number between 0 and 1:

$$F1 = 2 * \frac{P * R}{P + R} \tag{5}$$

The F1-score is a way to balance the trade-off between correctly predicting all instances of the positive class while limiting the amount of misclassifications.

IV. RESULTS

A. *Matrix Profile*

The paper presents the experimental results of the described approach on a reference building energy dataset. The dataset used contains the energy consumption for 507 university buildings from Europe and USA over one year period. The sampling time for each data-set is one hour and the they consists of 8.760 data points. Four types of dominant energy usage patterns were identified: classrooms, offices, laboratories and dormrooms, for a subset of 422 buildings that met out criteria. The data is slightly imbalanced as we count 177 classrooms ($\sim 42\%$), as the dominant class, 98 offices ($\sim 23\%$), 86 laboratories ($\sim 20\%$) and 61 dormitories ($\sim 15\%$). The input data is publicly available through the Building Data Genome repository [17].

The Matrix Profile was applied on all 422 datasets and Figure 1 and Figure 2 present the Matrix Profile result with a monthly window length for two particular time series. The first stems from a university laboratory in London, Europe and the second one from a university office in London, Europe. The readings cover the full year 2015. The figures show also the top three discords which represent the highest relative peaks on the Matrix Profile graph. For example, the first discord in Figure 1 is correlated with the period of Summer holiday spanning between 20 July and 31 August, the second one is correlated with Summer half term holiday between 23 May – and 31 May while the third discord is correlated with Easter holiday in the period 28 March – 19 April.

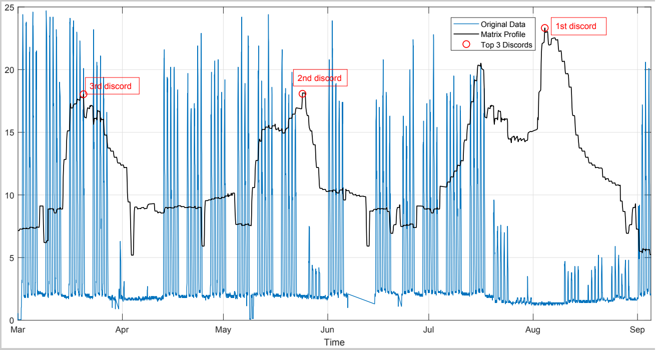

Fig. 1: Original Dataset versus Matrix Profile with top Three Discords (Laboratory)

Figures 3 and 4 present the number of cases when the first discord is correlated with the Summer vacation period and the situations when one of the three discords corresponds to the summer vacation, respectively. For this simulation there were considered all 422 datasets grouped by each type of dominant energy usage pattern: classroom, office, laboratory and dormroom. From both figures it can be seen that in the cases of ”Classroom” and ”Dormroom” bigger percentages values in terms of discords associated with Summer vacation were obtained than ”Office” and ”Laboratory” cases. One potential explanation for this behavior is that during the summer vacation there is not didactic activity but the offices

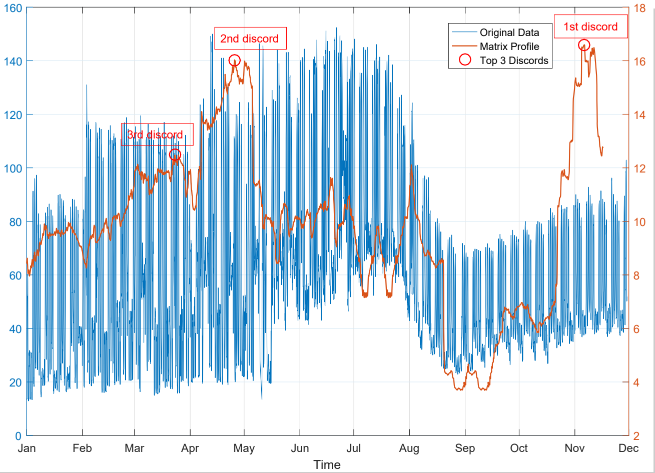

Fig. 2: Original Dataset versus Matrix Profile with top Three Discords (Office)

are still used by the professors and other technical staff and also laboratories still have activity during this period because they are used for research. We have run the MP algorithm with daily, weekly and monthly windows lengths for comparison purposes.

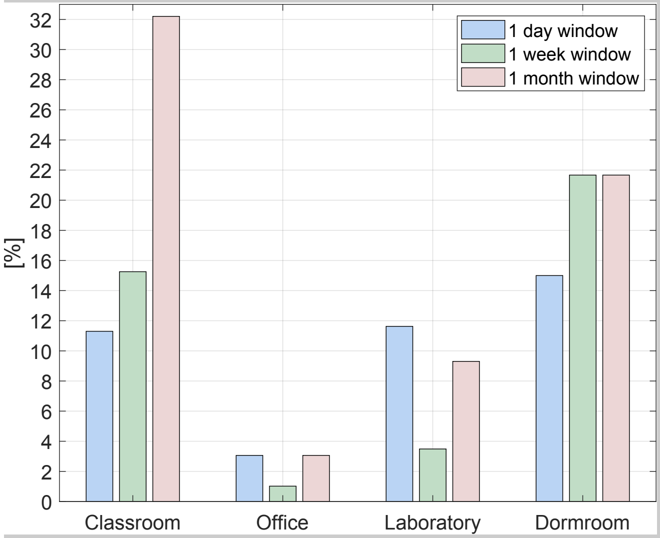

Fig. 3: Number of cases when the *top* Discord is correlated with the Summer holiday

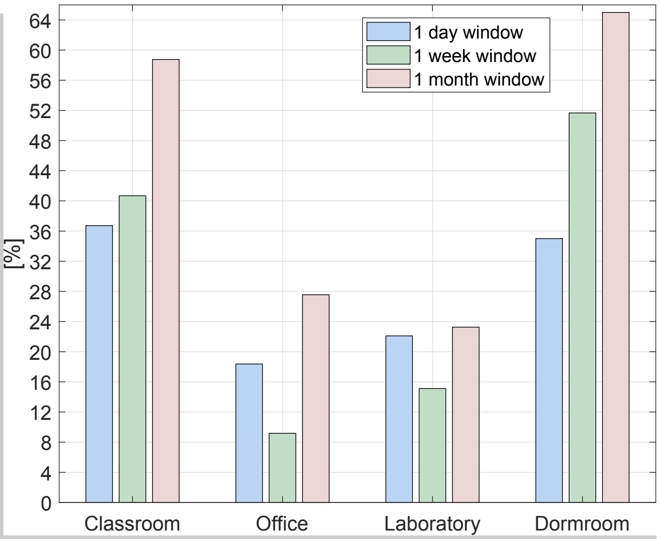

Fig. 4: Number of cases when *any* Discord is correlated with the Summer holiday

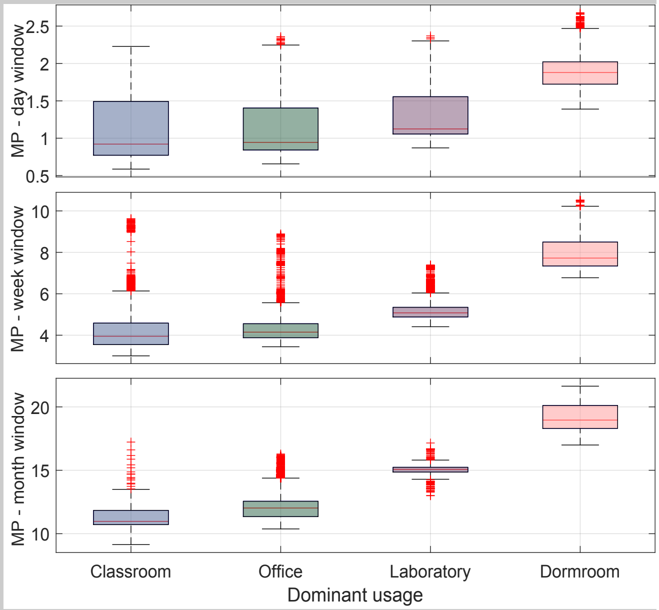

Fig. 5: Boxplot with MP values grouped by 4 types of building dominant usage

Figure 5 presents a boxplot graphic of the matrix profile values of all 422 datasets grouped by each type of dominant usage of the buildings. It can be noticed that in every case, either the one-day window MP or the one-week or one-month window the MP is on average similar in case of ”Classroom” and ”Office” groups. Significant differences occur in case of ”Dormroom” group where MP values are higher. This can be because the energy consumption pattern is based on occupant behavior and is quite unique. It can be noticed a slightly difference also in case of the ”Laboratory” group, the explanation being correlated with the laboratory type, the laboratory equipment and the activity that is held there presents aperiodic usage variations, reflected in the behavior of the energy use. The observations hold for each of the three subsequence lengths: daily, weekly and monthly.

A histogram of the first discord occurrences across the whole building energy dataset of 8760 samples is illustrated in Figure 6. It confirms that the most unusual yearly behaviour occurs during December and is associated with the Winter break and holiday period. Excepting this period, the top discords seems to present an uniform distribution during the rest of the year for all the studied buildings.

B. Classification

The hypothesis that we aim to investigate in this subsection is whether a reduced input data set, based on the anomaly

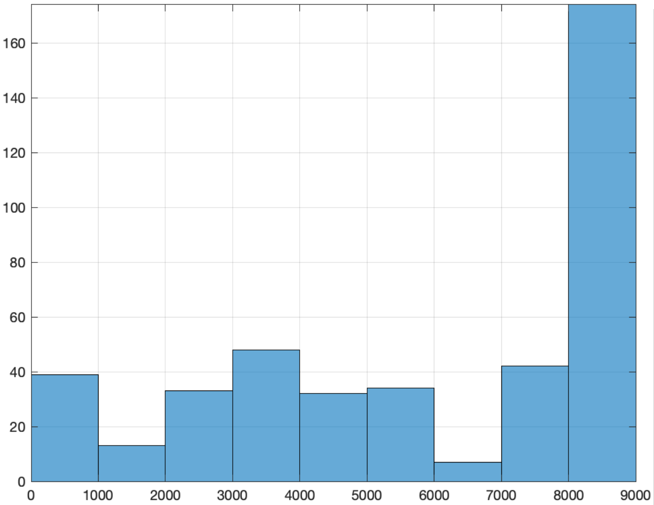

Fig. 6: Top discords distributions for weekly MP

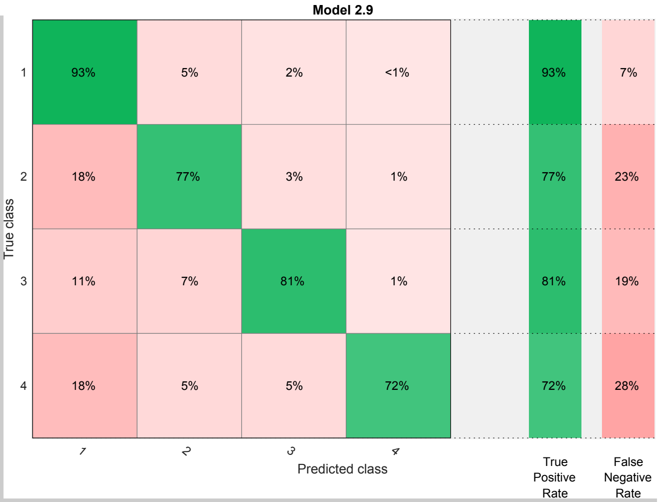

Fig. 7: Confusion Matrix for SVM Model

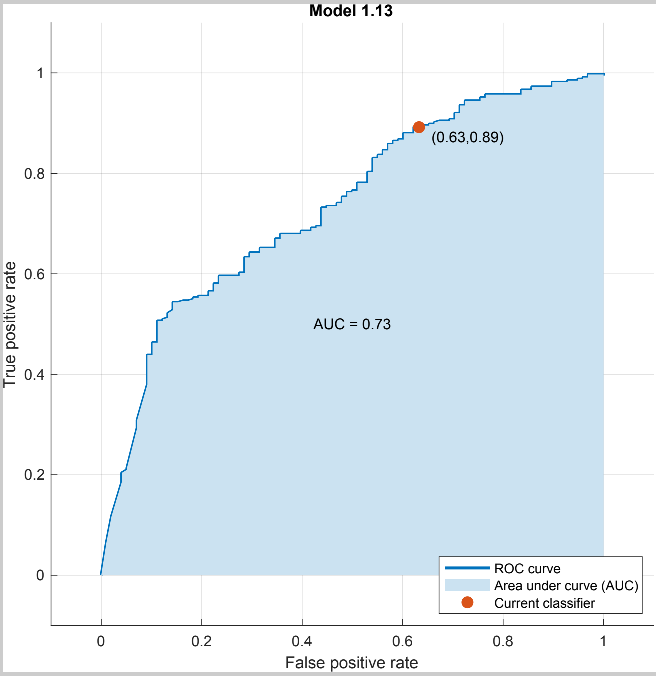

Fig. 8: Binary classification ROC diagram for Ensemble Model

instances detected through the MP, can provide useful discrimination between the dominant usage patterns of each building. For the classification task we build the training frame around the top three discords for each of the 422 buildings thus resulting in 1266 examples. The size of the dataset does not justify highly complex black-box models such as deep neural networks. The variables used for training include: the day of the week, month of the year and if it was a weekend when the anomaly occurred, the momentary power, the average, minimum, maximum, and standard deviation of the power over the past 24 hours, the normalised power per square meter of the building and the overall surface of the building. All numerical variables are z-score normalised to mitigate differences in absolute scale. The response variable is a categorical feature with the values 1/2/3/4 corresponding to the respective usage types. Several types of models are trained based on decision trees, nearest neighbours, support vector machines and ensemble models using classification and regression trees (CART).

The result in Figure 7 show the confusion matrix for a trained SVM model with Gaussian kernel. The ratios of misclassified instances are illustrated and it can be seen what are the mistakes that the classifier makes. The results show a satisfactory improvement over random guessing as well as baseline classifiers such as Naive Bayes.

The second example presents a restricted two-class binary problem where the positive class is the classroom category and the other three classes are group together. The receiver operating characteristic for an ensemble classifier model with $\sim 81\%$ accuracy and 0.73 area under curve (AUC) is shown in Figure 8. The AUC quantifies the improvement over a random choice denoted as the area over 45 degrees diagonal that starts from the origin.

V. CONCLUSIONS

The paper investigated whether MP based discord features are a good predictor for dominant usage patterns in large academic buildings. The insight is that the unusual behaviour in terms of energy consumption patterns provides sufficient discriminative information for reliable differentiation. The results show an improvement in classification performance using these features when accounting for a considerable decrease in training time compared to using the full year energy readings in the modelling stage.

The applicability of the approach can lead to considerable impact on the control loop performance and in decision support systems [18]. Further on, we aim to exploit the potential for application of the state-of-the art time series deep learning models such as HIVE-COTE and InceptionTime [19] on densly sampled energy measurement for anomaly detection and classification.

REFERENCES

[1] E. Brümmendorf, J. H. Ziegeldorf, and J. P. Fütterer, “IoT platform and infrastructure for data-driven optimization and control of building energy system operation,” *Journal of Physics: Conference Series*, vol. 1343, p. 012040, nov 2019. [Online]. Available: https://doi.org/10.1088%2F1742-6596%2F1343%2F1%2F012040

[2] J. Tang and Q. Jia, “A simulation platform for sensing system selection for occupant distribution estimation in smart buildings,” in *2019 IEEE 15th International Conference on Automation Science and Engineering (CASE)*, Aug 2019, pp. 985–990.

[3] C. Deb, F. Zhang, J. Yang, S. E. Lee, and K. W. Shah, “A review on time series forecasting techniques for building energy consumption,” *Renewable and Sustainable Energy Reviews*, vol. 74, pp. 902 – 924, 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S1364032117303155

[4] A. Capozzoli, M. S. Piscitelli, and S. Brandi, “Mining typical load profiles in buildings to support energy management in the smart city context,” *Energy Procedia*, vol. 134, pp. 865 – 874, 2017, sustainability in Energy and Buildings 2017: Proceedings of the Ninth KES International Conference, Chania, Greece, 5-7 July 2017. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S187661021734674X

[5] N. Ludwig, S. Waczowicz, R. Mikut, and V. Hagenmeyer, “Assessment of unsupervised standard pattern recognition methods for industrial energy time series,” in *Proceedings of the Ninth International Conference on Future Energy Systems*, ser. e-Energy ’18. New York, NY, USA: Association for Computing Machinery, 2018, pp. 434–435. [Online]. Available: https://doi.org/10.1145/3208903.3212051

[6] X. Yu, S. Ergan, and G. Dedemen, “A data-driven approach to extract operational signatures of hvac systems and analyze impact on electricity consumption,” *Applied Energy*, vol. 253, p. 113497, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261919311717

[7] Q. Jia, J. Wu, Z. Wu, and X. Guan, “Event-based hvac control—a complexity-based approach,” *IEEE Transactions on Automation Science and Engineering*, vol. 15, no. 4, pp. 1909–1919, Oct 2018.

[8] C. Miller, “More buildings make more generalizable models—benchmarking prediction methods on open electrical meter data,” *Machine Learning and Knowledge Extraction*, vol. 1, no. 3, pp. 974–993, Aug 2019. [Online]. Available: http://dx.doi.org/10.3390/make1030056

[9] J. Y. Park, X. Yang, C. Miller, P. Arjunan, and Z. Nagy, “Apples or oranges? identification of fundamental load shape profiles for benchmarking buildings using a large and diverse dataset,” *Applied Energy*, vol. 236, pp. 1280 – 1295, 2019. [Online]. Available: http://www.sciencedirect.com/science/article/pii/S0306261918318439

[10] S. Duque Anton, L. Ahrens, D. Fraunholz, and H. D. Schotten, “Time is of the essence: Machine learning-based intrusion detection in industrial time series data,” in *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, Nov 2018, pp. 1–6.

[11] C. Nichiforov, I. Stamatescu, I. Făgărăşan, and G. Stamatescu, “Energy consumption forecasting using arima and neural network models,” in *2017 5th International Symposium on Electrical and Electronics Engineering (ISEEE)*, Oct 2017, pp. 1–4.

[12] C. Nichiforov, G. Stamatescu, I. Stamatescu, I. Făgărăşan, and S. S. Iliescu, “Intelligent load forecasting for building energy management systems,” in *2018 IEEE 14th International Conference on Control and Automation (ICCA)*, June 2018, pp. 896–901.

[13] C. Nichiforov, G. Stamatescu, I. Stamatescu, V. Calofir, I. Fagarasan, and S. S. Iliescu, “Deep learning techniques for load forecasting in large commercial buildings,” in *2018 22nd International Conference on System Theory, Control and Computing (ICSTCC)*, Oct 2018, pp. 492–497.

[14] C. M. Yeh, Y. Zhu, L. Ulanova, N. Begum, Y. Ding, H. A. Dau, D. F. Silva, A. Mueen, and E. Keogh, “Matrix profile i: All pairs similarity joins for time series: A unifying view that includes motifs, discords and shapelets,” in *2016 IEEE 16th International Conference on Data Mining (ICDM)*, Dec 2016, pp. 1317–1322.

[15] E. Keogh, J. Lin, S.-H. Lee, and H. van herle, “Finding the most unusual time series subsequence: Algorithms and applications,” *Knowl. Inf. Syst.*, vol. 11, pp. 1–27, 01 2007.

[16] M. Hossin and S. M.N, “A review on evaluation metrics for data classification evaluations,” *International Journal of Data Mining and Knowledge Management Process*, vol. 5, pp. 01–11, 03 2015.

[17] C. Miller and F. Meggers, “The building data genome project: An open, public data set from non-residential building electrical meters,” *Energy Procedia*, vol. 122, pp. 439 – 444, 2017.

[18] I. Stamatescu, N. Arghira, I. Făgărăşan, G. Stamatescu, S. Iliescu, and V. Calofir, “Decision support system for a low voltage renewable energy system,” *Energies*, vol. 10, no. 1, p. 118, Jan 2017. [Online]. Available: http://dx.doi.org/10.3390/en10010118

[19] H. Ismail Fawaz, B. Lucas, G. Forestier, C. Pelletier, D. F. Schmidt, J. Weber, G. I. Webb, L. Idoumghar, P.-A. Muller, and F. Petitjean, “InceptionTime: Finding AlexNet for Time Series Classification,” *arXiv e-prints*, p. arXiv:1909.04939, Sep 2019.