Proc. of the International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME 2023)
19-21 July 2023, Tenerife, Canary Islands, Spain

2023 3rd International Conference on Electrical, Computer, Communications and Mechatronics Engineering (ICECCME) | 979-8-3503-2297-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICECCME57830.2023.10252975

Applied Anomaly Detection: a Bayesian Approach to Improve Robustness

Piero Danti *
Research Department
Yanmar R&D Europe
Firenze, Italy
piero_danti@yanmar.com

Alessandro Innocenti
Research Department
Yanmar R&D Europe
Firenze, Italy
alessandro_innocenti@yanmar.com

***Abstract*—The main drivers of the development and production of new energy devices are energy device efficiency and machinery maintenance strategies. The former to minimize pollutant emissions with a view to future carbon neutrality. Condition-Based Maintenance (CBM), on the other hand, can help improve machinery reliability and reduce downtime by monitoring equipment conditions and addressing potential problems before they become serious. It can also save companies money by reducing the number of unnecessary repairs, minimizing the need for spare parts, and optimizing maintenance schedules. In this paper, the authors propose a deep learning methodology to automatically detect anomalies on a real Combined Heat and Power (CHP) unit supplying a school in Germany. The core of the work is a convolutional autoencoder trained on the normal behavior of the energy generator. The autoencoder is enhanced with a Bayesian technique, the Monte Carlo dropout, used to add a stochastic component to the model to quantify the uncertainty degree of the detection. This information is crucial to determine if or when action is actually needed, optimizing the service and maintenance strategy. The proposed approach was applied to a real case study and was found to be effective, heat exchanger fouling was detected 5 weeks before the standard detection system. The algorithm returns high confidence in system anomalies and low detection confidence for minor alterations in behavior, less risky for the machine.**

***Index Terms*—Anomaly Detection, Uncertainty Quantification, Autoencoder, Monte Carlo dropout, CHP**

I. Introduction

Frequently, energy generators are operated with low productivity or, in worst cases, some negative trends appear weeks before a real breakdown of the machine; this lead to higher energy consumption and unplanned reparations that could be easily avoided through ad-hoc maintenance intervention. At present, the most promising methodologies for addressing anomaly detection issues are Machine Learning (ML) and Deep Learning (DL) models [1].

A prevalent scenario in this context is novelty detection, which is a technique utilized to handle partially unlabeled data. Industries frequently provide datasets in which they are confident that the machine under investigation is operating within the normal functioning range, rather than labeling a range of malfunctioning cases. Classically, novelty detection is applied to unbalanced datasets in which the majority of acquisitions describe the machine in a normal behavior while a small portion of data is representative of a faulty condition [2]. In contexts where machines are highly expensive and production cannot be interrupted, every detection is carefully considered to avoid costly breakdowns or performance losses. Conversely, for energy devices with productions below 100 kW, maintenance or interventions are typically carried out when the presence of a problem is certain. In this regard, it is important not only to provide anomaly detection but also a confidence level for the anomaly itself, and where possible, to point the sensors involved in the problem. This reinforces the concept of CBM.

Regarding adding a confidence level to the anomaly prediction, despite the success of standard DL methods in solving various real word problems, they cannot provide information about the reliability of their predictions [3]. Two main solutions have been proposed in recent years to introduce the information of uncertainty quantification: the Variational AutoEncoder (VAEs) [4] and the Monte Carlo dropout (MCD) [5].

VAEs are a type of generative model that learn a low-dimensional representation of the input data by encoding it into a latent space and then decoding it back to the original input space. In this process, VAEs minimize a reconstruction loss between the original input and its reconstruction, as well as a regularization term that encourages the latent space to follow a prior distribution. The resulting model can be used to encode-decode test samples and perform anomaly detection by measuring the reconstruction error of these samples [6] - [9].

MCD, on the other hand, is a dropout-based technique that uses dropout [10] during inference to estimate the model's uncertainty. Dropout randomly drops out units from the neural network during training, which acts as a regularization technique. During inference, dropout is applied multiple times with different dropout masks, and the resulting predictions are averaged to estimate the model's uncertainty. In anomaly detection, MCD can be used to estimate the uncertainty of the model's prediction for each test sample returning a rate of how many anomalies have been detected during the multiple inferences [11] - [13].

One key difference between VAEs and MCD is that VAEs are generative models, while MCD is a technique for uncertainty estimation. VAEs can be used not only for anomaly detection but also for tasks such as data generation. However, MCD is a simpler technique that can be easily applied to any existing neural network architecture.

Another difference is that VAEs require a prior distribution over the latent space, while MCDs do not. The choice of prior distribution can have a significant impact on the quality of the VAE's latent space representation, and finding an appropriate prior can be challenging. [14]

Finally, VAEs tend to be computationally more expensive than MCDs, as they require training a full generative model. MCD, on the other hand, only requires running the inference multiple times with different dropout masks.

Because of the above, an MCD method was chosen in this study. The subject of the work is a YANMAR cogenerator with a rated output of $20kW_e$ installed in a school in Germany. The main goals are the definition of an effective model for anomaly detection that can provide an uncertainty quantification of the anomaly and the identification of the sensors responsible for the anomaly. This last information can provide valuable insights into the underlying causes of the anomaly and can facilitate targeted maintenance and repair efforts. By isolating the specific sensor responsible for the anomaly, it is possible to focus resources on the relevant component and avoid unnecessary downtime or repairs.

This paper investigates an unexplored aspect in the current literature, namely the utilization of Bayesian quantification for anomaly detection in micro-CHP units. Specifically, the authors enhance an existing autoencoder architecture to accommodate the incorporation of MCD layers. The challenge lies in striking a balance that preserves the detection performance of the model despite the regularization introduced by the dropout layers, while also enabling the quantification of anomaly severity both from a system-level perspective and from the individual contribution of each signal.

II. Case study

The YANMAR micro-CHP under analysis provides electricity and thermal energy to a school facility located in Germany. Thermal energy is produced through a highly efficient process that recovers waste heat generated during electricity production and transfers it to warm water from the load that is then redirected back to the load. The core of the cogeneration unit is an Internal Combustion Engine (ICE) that converts natural gas into electrical energy. The heat extracted from the engine for cooling and the heat present in the hot exhaust gases are both recovered and transferred to a carrier liquid, usually water, that feeds the system. The system is part of a more complex system that includes a buffer tank and a gas boiler. The $20kWe$ CHP engine meets the electric load with backup from the grid. The recovered heat and the gas boiler are used to charge the buffer tank to cover thermal requests.

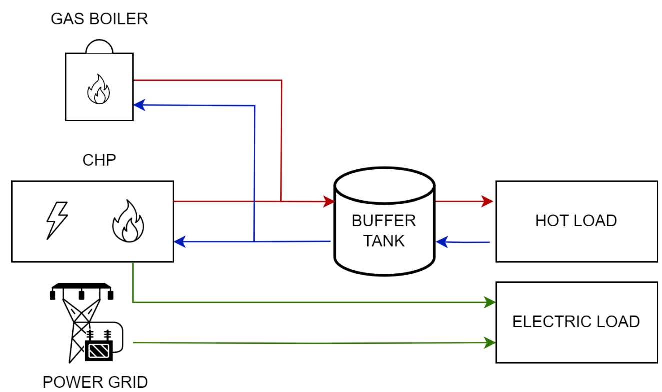

Fig. 1. Case study plant layout.

Fig. 1 offers an illustrative representation of the plant layout, emphasizing the interconnections of the generators. The red lines denote the piping system responsible for conveying hot water from the micro-CHP unit and the gas boiler to meet the thermal load, while the blue lines indicate the return flow of water after heat exchange with the school. The electrical connections, drawn in green, represent the wiring network that provides power supply to the facility.

The dataset available contains information from the whole plant, indeed measurements acquired by 49 sensors are recorded and stored in a relational database every 15 minutes. Only 12 measurements were considered interesting to develop an anomaly detection routine able to catch deviations from the normal behavior of the CHP unit:

- CHP active power;
- CHP pump rotation rate;
- ICE inlet temperature;
- ICE outlet temperature;
- CHP cabin temperature;
- ICE oil temperature;
- ICE exhaust gas temperature;
- heat exchanger exhaust temperature;
- heating circuit pump rotation rate;
- ICE lambda sensor value;
- ICE gas mixer valve position;
- ICE throttle valve position.

In addition, a manipulated signal has been created. This signal is a collection of boolean digits that keeps track of the reliability of the lambda sensor's records. The lambda sensor can provide unreal behaviors when the CHP is starting or shutting down.

A. Dataset split

To train the model, a 6-week training period is required [15] where the CHP behaves normally. Two test-sets are also required: the first one where the CHP does not present particular deviation from the learned behavior to verify that the model does not present false positives and the second one where a certain quantity of abnormalities happens to check that the model correctly detects the presence of anomalies.

Every YANMAR CHP has a diagnostic onboard system that at each acquisition reports the CHP status. Unluckily faulty statuses are created utilizing coarse thresholds and, consequently, the presence of an alarm is a sufficient condition to say that the CHP has an anomaly but it is not a necessary condition. Many faults or degradation trends go unnoticed by the diagnostic system. The analyzed dataset presents the following fault' codes with the respective message:

- 0 - CHP stopped but ready to start;
- 2 - CHP stopped. Maintenance needed;
- 11 - Shutdown: generator protection;
- 12 - Low hydraulic pressure;
- 16 - Shutdown: water pressure too low;
- 18 - Shutdown: overheat interior;
- 19 - Shutdown: overheat engine oil.

Furthermore, the CHP can operate in four different operating modes. If the anomaly detection model is trained on a specific mode then also the test-sets must be selected accordingly. Five operation modes are present:

- Mode 0, The CHP is turned off;
- Mode 1, the CHP is ready to charge an electric car;
- Mode 2, the CHP is optimized for summer operation;
- Mode 3, the CHP is optimized for heat production;
- Mode 4, the CHP is optimized for power production.

However, CHP under analysis mainly operates to optimize heat production; As depicted in Fig. 2, except for Mode 0 when the plant is idle, the CHP primarily functions in either Mode 2 or Mode 3. When the CHP operates at maximum capacity, it is consistently configured to operate in Mode 3, which may be attributed to the high demand for thermal power within the facility. However, at partial load, the CHP primarily operates in Mode 2.

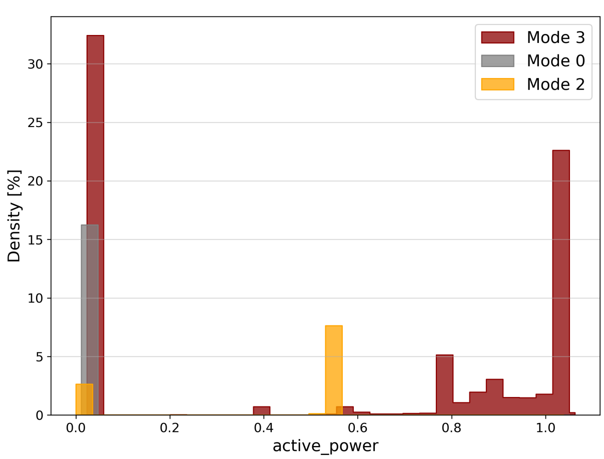

Fig. 2. Distribution of the active power in relation to the CHP operating mode.

Fig. 3 depicts the CHP behavior in terms of adimensionalized produced electric power (orange scatter), alarms reported by the onboard diagnostic system (red scatter), and the human detection made by the plant supervisor (purple vertical line). Grey-shaded areas emphasize the dataset split used for this work and in particular datasets ranges are as below:

- *train-set*, 1st July 2019 - 12th August 2019 (6 weeks);
- *test-set with heat exchanger fouling*, 1st July 2021 - 31st December 2021;
- *test-set in healthy conditions*, 15th July 2022 - 7th September 2022;

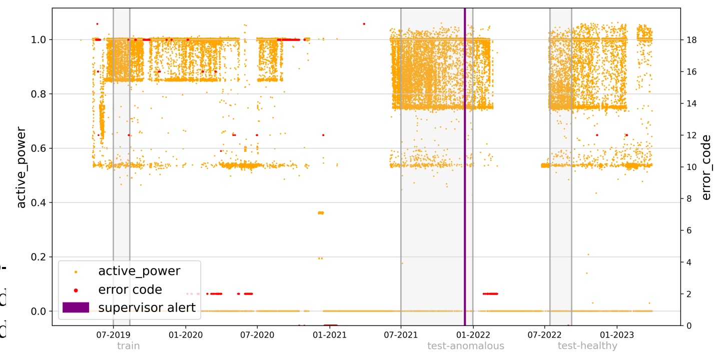

Fig. 3. Active power describing CHP behavior.

The observations presented in Fig. 2 are supported by the findings in Fig. 3, which indicate that the CHP system primarily operates either at full load (between 80% and 100% of the rated power) or at 50% partial load. It is worth noting that the training period for the model corresponds to a period where the CHP management logic is slightly distinct from the two test sets. Specifically, the range of powers covered during the training phase is comparatively smaller than that observed after a plant shutdown period of approximately 5 months, where the power range has increased.

It is important to remark that the errors reported by the onboard CHP system (red dots in the picture) are not correlated to the heat exchanger fouling acknowledged by the plant supervisor. Furthermore, these errors are of minor interest in terms of research as trivial faults can be recognized more easily due to their disruptive consequences often leading to a machine shutdown. On the other hand, the heat exchanger deterioration studied in this work is very interesting due to its slow degradation trend (low negative) over time which is not easily detectable by naive thresholds or by the human naked eye.

III. Proposed Model

The authors addressed the same problem in [16] using an autoencoder with convolutional layers (ConvAE) that consider temporal information. In this paper, a new feature is introduced by using a Bayesian approach to estimate the confidence with which the autoencoder detects an anomaly.

As delineated in Section I, compared to VAE, the MCD approach facilitates the introduction of stochasticity in the model without necessitating any modification in the original

architecture of the deep learning model, if a dropout layer has already been incorporated.
In the reference architecture, there are no dropout layers but it has produced excellent results ([16]) and therefore we would like to keep it similar. To compensate, changes were made to the original ConvAE model by incorporating MCD to maintain the same level of performance when it comes to reconstruction errors. In fact, since MCD functions as a regularization technique during the training process, the original architecture required additional complexity to maintain the same level of information retention during learning.

Hyperparameters employed for both models are presented in Table I, which highlights the augmentation of convolutional filter numbers, from 10 to 18, and an expansion of the latent space from 3 to 5 dense neurons. Additionally, a MCD of 5% was incorporated, and the inference was conducted 50 times.

TABLE I.
HYPERPARAMETERS OF CONFIGURATIONS FOR THE TWO MODELS.

hyper-parameter	**ConvAE**	**ConvAE MCD**
sliding window	10	10
input dimension	(10,13)	(10,13)
latent dimension	3	5
output dimension	(10,13)	(10,13)
activation function	*relu*	*relu*
batch size	32	32
learning rate	10^{-4}	10^{-4}
l1 regularization	0	0
dropout	0	-
padding	same	same
strides	1	1
filters number	10	18
kernel size	5	5
optimizer	Adam	Adam
loss function	MSE	MSE
validation split	0.1	0.1
Montecarlo dropout	-	0.05
Montecarlo samples	-	50

The architecture of the Convolutional autoencoder with Monte Carlo Dropout (ConvAE MCD) is reported in Fig. 4. The ConvAE MCD exhibits an asymmetrical arrangement, where the encoder is marginally bigger than the decoder. The input data comprising 13 features are initially segmented into 10-sample moving windows and then processed using a 1-D convolutional layer composed of 18 filters. The output of this layer then proceeds to the bottleneck of the autoencoder through a fully connected layer. On the decoding side, the architecture is identical, except for the convolutional layer, which employs 13 filters to align with the number of feature reconstructions.

During the training stage, the model strives to minimize the reconstruction error of the inputs. A threshold is set using the 99^{th} percentile of the training error to discern between normal and abnormal samples during the testing phase. Two types of anomalies are generated. The first type is generated for the entire system if the average error over the 13 features exceeds the 99^{th} percentile threshold calculated on the train-set reconstructed inputs. The second type of anomaly is generated

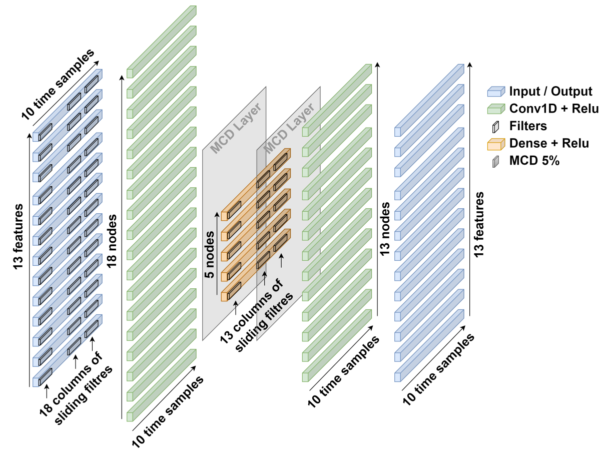

Fig. 4. Architecture of the Convolutional Autoencoder with MCD.

for each signal: a specific reconstruction error is calculated and compared with a 99^{th} percentile threshold obtained by considering the error committed during the training phase inference of the particular signal itself. The authors refer to the first type of anomaly as *system anomaly* and the second type as *signal anomaly*.
Therefore, the autoencoder generates anomaly signals that may be susceptible to yielding false positives. To mitigate this challenge, a post-processing low pass filter has been devised to disregard any alarms that persist for less than 12 hours. The methodology proposed in the article is illustrated in Fig. 5, which presents a flowchart outlining the steps followed in the study.

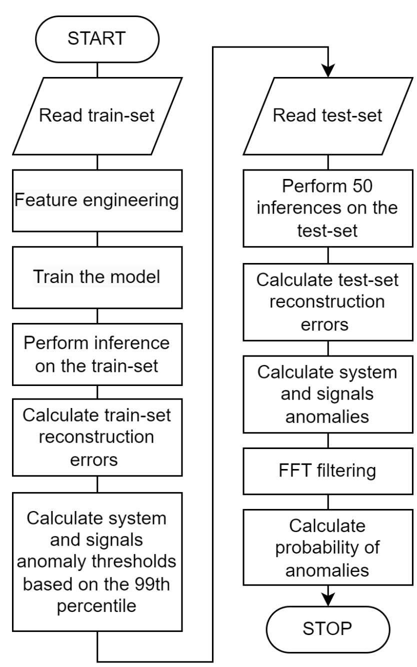

Fig. 5. Flowchart of the proposed methodology.

IV. Results

As mentioned in Section III, the autoencoder model has been designed to reconstruct 13 features but, as a matter of simplicity, in the following the authors describe only the most interesting signals recorded during the *test-set with heat exchanger fouling* and the *test-set in healthy conditions*.

During periods where anomalies are present the heat exchanger becomes dirty, and the efficiency of the CHP has a negative trend, which leads to a significant increase in costs. Fig. 6 shows how the anomaly was discovered only on 10^{th} December 2021 (vertical purple line). However, the proposed algorithm detects the first system anomaly (black line) on October the 10^{th}, due to an unexpected shutdown. Subsequently, a second general anomaly is detected on November the 7^{th}. In this case, the anomaly depends on the fouling of the heat exchanger, since signals related to the heat exchanger produce specific alarms (red lines): in particular, it can be seen how the exhaust gas temperature decreases significantly on average.

The discussed alarms are considered extremely reliable by the model as the percentage of Monte Carlo experiments generating an anomaly is 100%. However, the behavior of the alarm linked to the heating circuit pump is different. The pump starts operating at lower speeds, initially generating an anomaly confidence level of around 60%, which then decreases ranging between 20% and 10% when the pump speeds up. In contrast, the alarm signal of the CHP cabin temperature shows a confidence level of around 40%, which then increases to 100%.

Ultimately, the algorithm detected the heat exchanger fouling 5 weeks before the actual detection. Unluckily, the system was not yet embedded in an online routine, otherwise, it would have saved costs linked to the CHP inefficiencies. On the other hand, if the heat exchanger anomaly was a late detection by the engineer supervising the plant, the algorithm revealed another problem of the cogeneration unit that went totally untracked but then acknowledged by the YANMAR maintenance department: the engine cooling system exhibited abnormal functionality, resulting in a lower range of oil temperature. As a result, the algorithm yielded a probability of anomaly with complete certainty, i.e., 100%.

In order to demonstrate that the algorithm is not only sensitive to anomalies but is also robust and does not produce a large number of false positives, it is necessary to have a dataset where we are confident that the CHP is functioning properly and without malfunctions. At the beginning of July 2022, a global maintenance intervention was planned and also the heat exchanger was substituted.
Fig. 7 shows how all signals recovered and no anomalous trend is present: indeed, the algorithm proved to be well-performing also when the CHP is healthy and to be not prone to false positives. The presented data demonstrate that the five signals, which are directly related to the heat exchanger, have regained a normal behavior, devoid of any discontinuities. Fig. 6 illustrates that all temperature signals (cabin, oil, and gas exhaust) exhibited a negative temporal trend, which has now vanished in Fig. 7. Furthermore, the pump of the heating circuit has resumed its healthy pattern of modulation, operating within the range of 0% to 100%.
By comparing the plots of these five signals before and after the maintenance intervention, the authors have reached the conclusion that the proposed routine exhibits sensitivity towards the anomaly presented and effectively distinguishes when a healthy condition has been restored.

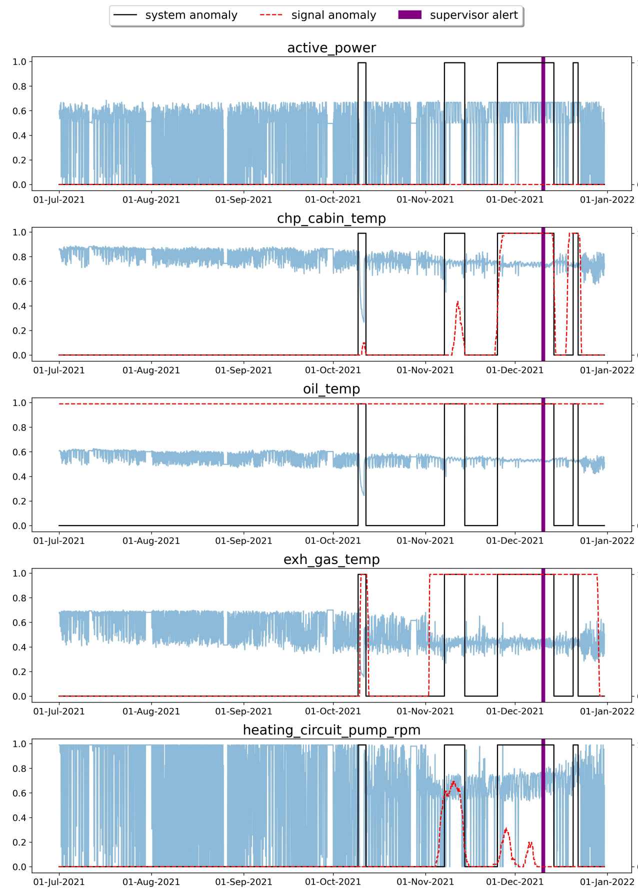

Fig. 6. Time-series of the normalized features used to test the ConvAE MCD in the presence of heat exchanger fouling.

V. Conclusions

Major companies are striving to migrate their business from the production and sale of products to the provision of services. In this context, data-driven approaches can flourish and provide significant added value. Specifically, with growing attention to climate change and energy and economic savings, the transition from time-based maintenance techniques to CBM plays a primary role.
In this paper, the authors propose a DL-based technique that, when trained on 6 months of data under normal conditions of a YANMAR micro-cogenerator, can provide both qualitative and quantitative indications of the CHP's behavior. After a description of the case study, the authors explain how the data were selected to train and test the proposed algorithm.

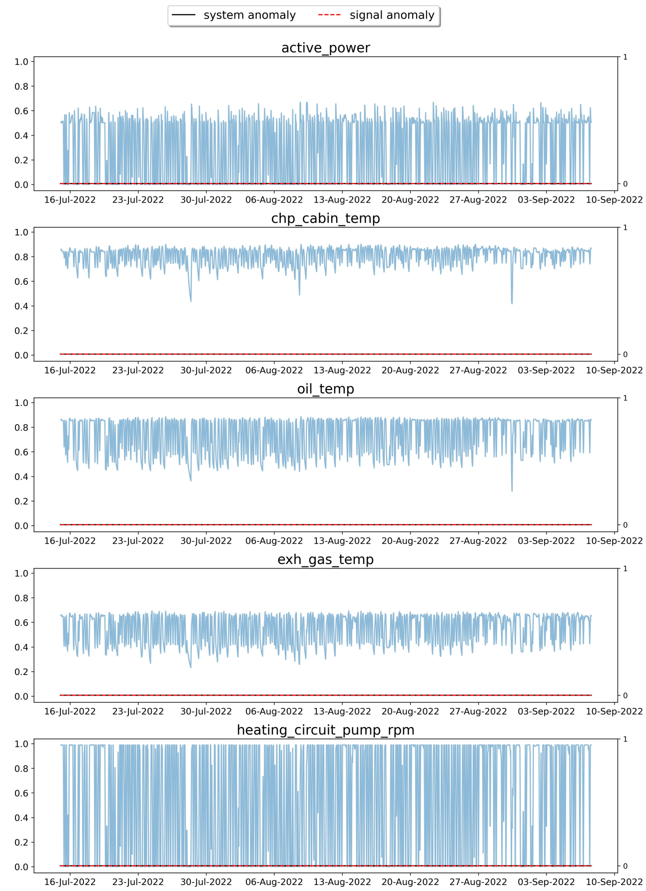

Fig. 7. Time-series of the normalized features used to test the ConvAE MCD in the absence of anomalies.

In particular, two test-sets were chosen: the first where deterioration of the heat exchanger was known a-priori, and the second where maintenance and replacement of faulty parts had just occurred. The DL model consists of an asymmetric autoencoder embedded with 1D-convolutional layers, Monte Carlo dropout layers, and a fully connected bottleneck. At the end of the detection pipeline, a frequency-based filter is used to reduce false positive alarms and increase robustness. The proposed algorithm demonstrated predictive capabilities by detecting heat exchanger fouling five weeks before the plant supervisor noticed and by revealing an anomaly in the oil temperature measure that had gone disregarded. Furthermore, the algorithm demonstrated not to be prone to false positives by not detecting any anomalies during periods of normality.

The MCD layers allow the introduction of stochasticity in the diagnosis process and by performing a certain amount of inferences it is possible to get a quantification of the uncertainty in detecting the anomalies adding a piece of important information to decide if the maintenance intervention is urgent or must be planned in the short future.

References

[1] O. Surucu, S. A. Gadsden, and J. Yawney, 'Condition Monitoring using Machine Learning: A Review of Theory, Applications, and Recent Advances', Expert Systems with Applications, vol. 221, p. 119738, Jul. 2023, doi: 10.1016/j.eswa.2023.119738.

[2] M. A. F. Pimentel, D. A. Clifton, L. Clifton, and L. Tarassenko, 'A review of novelty detection', Signal Processing, vol. 99, pp. 215–249, Jun. 2014, doi: 10.1016/j.sigpro.2013.12.026.

[3] M. Abdar et al., 'A review of uncertainty quantification in deep learning: Techniques, applications and challenges', Information Fusion, vol. 76, pp. 243–297, Dec. 2021, doi: 10.1016/j.inffus.2021.05.008.

[4] D. P. Kingma and M. Welling, 'Auto-Encoding Variational Bayes', arXiv:1312.6114 [cs, stat], May 2014, Accessed: Mar. 17, 2022. [Online]. Available: http://arxiv.org/abs/1312.6114

[5] Y. Gal and Z. Ghahramani, 'Dropout as a Bayesian Approximation: Representing Model Uncertainty in Deep Learning'. arXiv, Oct. 04, 2016. Accessed: Mar. 30, 2023. [Online]. Available: http://arxiv.org/abs/1506.02142

[6] C. Sun, Z. He, H. Lin, L. Cai, H. Cai, and M. Gao, 'Anomaly detection of power battery pack using gated recurrent units based variational autoencoder', Applied Soft Computing, vol. 132, p. 109903, Jan. 2023, doi: 10.1016/j.asoc.2022.109903.

[7] M. Castangia, R. Sappa, A. A. Girmay, C. Camarda, E. Macii, and E. Patti, 'Anomaly detection on household appliances based on variational autoencoders', Sustainable Energy, Grids and Networks, vol. 32, p. 100823, Dec. 2022, doi: 10.1016/j.segan.2022.100823.

[8] A. González-Muñiz, I. Díaz, A. A. Cuadrado, D. García-Pérez, and D. Pérez, 'Two-step residual-error based approach for anomaly detection in engineering systems using variational autoencoders', Computers and Electrical Engineering, vol. 101, p. 108065, Jul. 2022, doi: 10.1016/j.compeleceng.2022.108065.

[9] M. Cozzatti, F. Simonetta, and S. Ntalampiras, 'Variational Autoencoders for Anomaly Detection in Respiratory Sounds'. arXiv, Aug. 05, 2022. Accessed: Mar. 30, 2023. [Online]. Available: http://arxiv.org/abs/2208.03326

[10] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, 'Dropout: A Simple Way to Prevent Neural Networks from Overfitting'.

[11] M. Carletti, C. Masiero, A. Beghi, and G. A. Susto, 'A deep learning approach for anomaly detection with industrial time series data: a refrigerators manufacturing case study', Procedia Manufacturing, vol. 38, pp. 233–240, 2019, doi: 10.1016/j.promfg.2020.01.031.

[12] M. A. M. Sadr, Y. Zhu, and P. Hu, 'An Anomaly Detection Method for Satellites Using Monte Carlo Dropout', IEEE Trans. Aerosp. Electron. Syst., pp. 1–9, 2022, doi: 10.1109/TAES.2022.3206257.

[13] A. Legrand, H. Trannois, and A. Cournier, 'Use of Uncertainty with Autoencoder Neural Networks for Anomaly Detection', in 2019 IEEE Second International Conference on Artificial Intelligence and Knowledge Engineering (AIKE), Sardinia, Italy: IEEE, Jun. 2019, pp. 32–35. doi: 10.1109/AIKE.2019.00014.

[14] D. P. Kingma and M. Welling, 'An Introduction to Variational Autoencoders', FNT in Machine Learning, vol. 12, no. 4, pp. 307–392, 2019, doi: 10.1561/2200000056.

[15] P. Danti, A. Innocenti, 'A methodology to determine the optimal train-set size for autoencoders applied to energy systems', unpublished.

[16] P. Danti, A. Innocenti, and S. Sandomier, 'Anomaly Detection and Root Cause Analysis using Convolutional Autoencoders: a real case study', unpublished.