

Low-Parameter Critic-Based Multivariate WGAN Model for Clogging Detection in Drives

Special issue paper

Artur Dawid Surówka^{1,*}, Marcin Kosiba¹, Teemu Mikkilä², Asko Kavala³, Marcin Firla¹

¹Corporate Technology Center, ABB Sp. z o.o. Kraków, Poland

²Cadmatic Oy, Kouvola, Finland

³Motion Services, ABB Oy, Lappeenranta, Finland

Received: 26 February, 2025; Received in the revised form: 16 April, 2025; Accepted: 17 April, 2025

Abstract: Efficient detection of anomalies in the cooling system of variable frequency drives (VFDs) is crucial to minimise downtime costs from overheating. Smart condition monitoring tools, especially those using machine/deep learning, have proven effective for failure detection. Recent research has focussed on environmental factors, such as pollution and humidity, affecting VFDs. Clogging is particularly harmful as it can damage power electronics, leading to extended downtimes. This study explores the use of Wasserstein Generative Adversarial Networks (WGANs) for detecting clogging in drives, including inlet/outlet/heatsink clogging and fan blockage. WGANs are adept at recognising complex temporal patterns due to their feedback-driven training. Despite generative AI models being typically large and unsuitable for embedded systems, this work demonstrates the feasibility of a low-parameter WGAN critic-based model for detecting cooling issues in VFDs. Using temperature signals, the model can detect clogging as low as 20%–30% with high performance metrics, achieving up to 90% accuracy and an F1 score above 0.9 for heatsink clogging detection, using a lightweight 26-parameter critic model. This study shows the potential for developing low-parameter WGAN critic-based models for clogging detection in VFDs.

Keywords: variable frequency drives • generative adversarial networks • condition monitoring • machine learning • anomaly detection

1. Introduction

The role of environmental factors that impact operation of variable frequency drives (VFDs) has recently sparked a vibrant discussion in industry on extending the functionalities of their existing preventive maintenance systems (Al-Naseem and El-Sayed, 2013). This is due to the fact that frequent warranty interventions generate additional exploitation overheads both at the user and the producer/service ends (Yellamati et al., 2013). Among them, the influence of environmental factors, such as airborne pollution, humidity and ambient temperature, among others, were recently shortlisted common factors underlying most of their failures (Kennedy, 2021). In particular, the clogging of the cooling system, i.e., inlet/outlet filters, by airborne particles, seems to be particularly troublesome (Avor and Chang, 2019). It can lead to accumulation of metallic particles in boards, likely causing short circuits by contamination failure and increasing the risk of harmful corrosion (Avor and Chang, 2019). Therefore, detecting the requirement of filter replacement can be considered as a valuable maintenance action, so that anomaly/failure detection can be seen particularly beneficial to users to assure reduced exploitation costs (Avor and Chang, 2019). In the working environments with a high concentration of particulate matter, especially when combined with moisture and oil, the fan draws in all the debris, so that the layer of dust can be formed on the electronics, leading to the overheating of single elements in a VFD (Avor and Chang, 2019). Heat sinks are particularly vulnerable targets

* Email: artur.surowka@pl.abb.com

to dust accumulation. Their contamination results in reduced dissipation efficiency, so that the excessive heat is no longer efficiently drawn away to the exterior (Peterson, 2022). Excessive heat generation along with possible short circuits in a VFD can facilitate ageing of the power electronics system (Kang et al., 2023). For example, insulated gate bipolar transistors (IGBTs) can get damaged by increased temperatures resulting from overcurrent and overvoltage in the chips (Otsuki et al., 2003). One critical issue is that under prolonged thermal stress, the solder layer in IGBTs tends to develop cracks and voids, thus leading to delamination and subsequent failures (Ciappa and Fichtner, 2000). Inefficient operation of the cooling system in a VFD can also impact operation of the DC link circuit (Avor and Chang, 2019). This causes excessive capacitor wear, which can trigger transients, resulting in failures of input diodes and the capacitor itself. The latter can be due to an increased equivalent series resistance (ESR) value at high temperatures, gradually leading to DC link failure and time-consuming repairs in field conditions (Avor and Chang, 2019).

For both monitoring the conditions and screening failures in VFDs, anomaly detection algorithms are of increasing popularity in the power electronics of the drives (Zhang et al., 2023). Anomaly detection, in general, is a method that is aimed at detecting patterns that deviate from expected normal operation in the data and is typically used for outlier and novelty analysis (Chandola et al., 2009). There are two major types of anomalies: (a) point (abnormal parcels in the data) and (b) contextual anomalies (unusual temporal-wise patterns) (Chandola et al., 2009). Concerning time series data, often used in monitoring of VFDs using sensors, the latter type is more relevant, so that temporal correlations can be unravelled between consecutive timestamps (León-López et al., 2021). Model-based anomaly detection methods were recently found particularly reliable and flexible approaches for detecting faults in power electronics in VFDs, due to their flexibility, modularity and hence possibility for further using in embedded systems (Beattie et al., 2022). In particular, the usage of models utilising temperature signals was recently advised as it can yield models with better accuracy and training speed (Gómez et al., 2024). To this end, the use of machine/deep learning is often efficient (Surówka et al., 2023). We have recently demonstrated that anomaly detection systems using complex models, such as deep autoencoders (using either linear or long-short-term memory (LSTM) units), can be successfully applied to detect point anomalies in temperature signals from drives in wind turbine and simulated data (Surówka et al., 2024a,b). More recently, we showed that anomaly detection using hidden Markov models (HMM) with Gaussian mixture model initialisation is an effective method in detecting clogging in the cooling system of the drive: inlet, heatsink and outlet (Surówka et al., 2024a,b).

Recently, there has been growing interest in applying generative AI tools for anomaly detection due to their enormous potential to detect the most subtle and complex patterns in data, similar to a human expert (Tuli et al., 2022; Xu et al., 2022). Specifically, there has been an increasing number of applications of generative adversarial networks (GAN) for anomaly detection in time series data (Lüer and Böhm, 2021). The core idea behind GANs involves adversarial training of two models: a generator and a discriminator (or critic). The generator transforms random noise into data similar to the training set, while the discriminator aims to distinguish between real and generated data (Chollet, 2018). This adversarial process guides the generator to produce more realistic data while simultaneously improving the discriminator's ability to detect data that differ from the training examples (anomalies) (Yin and Kaynak, 2020). However, the training process of GANs is very difficult to converge due to the instability of the adversarial training dynamics and the sensitivity to hyperparameter settings (Xia et al., 2023). To address this issue, Wasserstein Generative Adversarial Networks (WGAN) networks were proposed to improve upon traditional GANs by providing more stable training and better quality of generated data. This is achieved through the utilisation of the Wasserstein loss function Eq. (1) (Chen et al., 2022):

$$L = E_{x \sim P_{\text{real}}} [D(x)] - E_{z \sim P_z} [D(G(z))] \quad (1)$$

where $D(x)$ denotes the critic score for real data x , $D(G(z))$, is the critic score for generated data $G(z)$, P_{real} is the distribution of real data, P_z is the distribution of the noise input to the generator and z represents the noise data. This equation encapsulates the zero-sum game nature of the adversarial training in WGANs. The generator model tries to minimise the loss L by generating data from noise ($G(z)$) that resembles the training data, while the discriminator (critic D) aims to maximise the loss by accurately distinguishing real data from generated data (Gulrajani et al., 2017). The noise data z are thus used to explore and adjust the latent space of the training data (Chollet, 2018).

WGAN networks are currently proposed as promising solutions for anomaly detection due to their enormous potential for detecting subtle and complex anomalies in the time domain (Qi et al., 2023). This capability stems from their unique adversarial training process, which mimics human-like learning by continuously improving

through feedback between the generator and the discriminator (Goodfellow et al., 2016). For example, recently, Dai et al. (2024) proposed an unsupervised anomaly detection model called Wave-GANomaly, which uses multi-feature fusion and selection, and was proved successful, as benchmarked against CIFAR-10 and MNIST datasets.

Unfortunately, WGAN models often pose a challenge in productization due to generalisation/training convergence issues caused by limited data availability, especially that of high accuracy and diversity (Tang et al., 2025). Not to mention that in case of huge model architectures, there are complex problems with their deployment in embedded systems, with limited memory, thus constraining their operation to Clouds (Qi et al., 2023). To address these issues, herein, we propose to use a WGAN model with a low parameter critic (for anomaly detection) and a substantially complex generator (for training only).

2. Methods

2.1. Experiments and data

The data were collected at the Power Electronics Laboratory of Lappeenranta University of Technology (Lappeenranta, Finland) using the equipment shown in Figures 1a,b. First, a set of additional temperature sensors (PT100, RS PRO, RTD Sensor, 4.5mm Dia, 50mm Long, 4 Wire, Class B) was installed inside the ACS880 (low voltage drive, ABB) inverter of the device under test (DUT). The temperature measurements were done using the PT100 sensors, located in the cooling system inside the DUT assembly (Figures 1c,d). The data from the PT100 sensors were collected to NI 9216 analogue temperature input modules and sent from the NI cDAQ-0184 via Ethernet to a computer. For the duration of the experiments, the DUT was enclosed in a calorimeter where the ambient temperature was controlled. The DUT drive controlled the DUT motor (ABB IE3 M3BP, ABB) according to the programmed load sequence. The DUT motor was mechanically coupled by a shaft to a load motor (ABB M2BA, ABB), which was also controlled by a load actuator (ACS850, ABB). During the tests, the DUT drive was speed controlled, and the load motor was torque controlled. This resulted in the load motor acting as a controlled variable load, with the DUT variable speed drive changing the rotational speed according to a given reference. The experiments collected data under normal and failure conditions (DUT jamming) with different severities. The sampling time of the data was 1 s. The normal operation data were acquired at various ambient temperatures including 25°C, 35°C and 45°C. For more information on the raw data, please see our recent study (Surówka et al., 2024a,b). Failure modes, simulated in the drive, included partially blocking the ventilation system of the DUT drive: inlets and outlets, the heat sink's fins and introducing blockage (speed reduction) to the input fan. All the blockages involved 50% or 75% of active areas of vents/fins or fan speed restriction. Each equally repeated cycle of operation consisted of three phases, as presented in our previous work: (a) start-up (30 min) was the step response stage, then (b) stop (40 min) was the cooling stage and finally active operation (22 min) (see [Surówka et al., 2024a,b]). The start-up process was configured as a step response, during which nominal speed and torque were applied following a brief ramp. This allowed sufficient time for temperature measurements to stabilise. For this study, the following variables were considered relevant to clogging anomalies and incorporated to the modelling phase: inverter temperature, IGBT junction temperature and the IGBT module's case temperature. For training purposes, a series of healthy operation data collected at 25°C, 35°C and 45°C were used. The constant ambient temperature was first subtracted from each raw profile to avoid abrupt changes in the data upon further stacking. To facilitate the training process, the data were augmented by interpolating between the respective datasets to obtain profiles for intermediate ambient conditions, thereby generating more profiles from the initial three. To ensure the stability and effectiveness of the training process, we augmented the data using 1D interpolation between the given (entire) temperature profiles at 25°C, 35°C and 45°C (using the 1D interpolation function from Python's SciPy library). This technique allowed us to create intermediate profiles, facilitating smoother transitions between different cases for the model, thereby making the cost space more regular and easier to navigate during training. The 1D interpolation method is robust and computationally stable, ensuring that the augmented data did not introduce new information but rather smoothed the cost space for the model. After that, the generated profiles were stacked to create one series of data. Additionally, we applied the Robust Scaler to the augmented stacked dataset after the interpolation procedure, which normalised the data in a way that was robust to outliers and further stabilised the training process. Figure 2a shows the final augmented and scaled normal operation data used in the training of WGAN. Notably, the ambient-subtracted training data

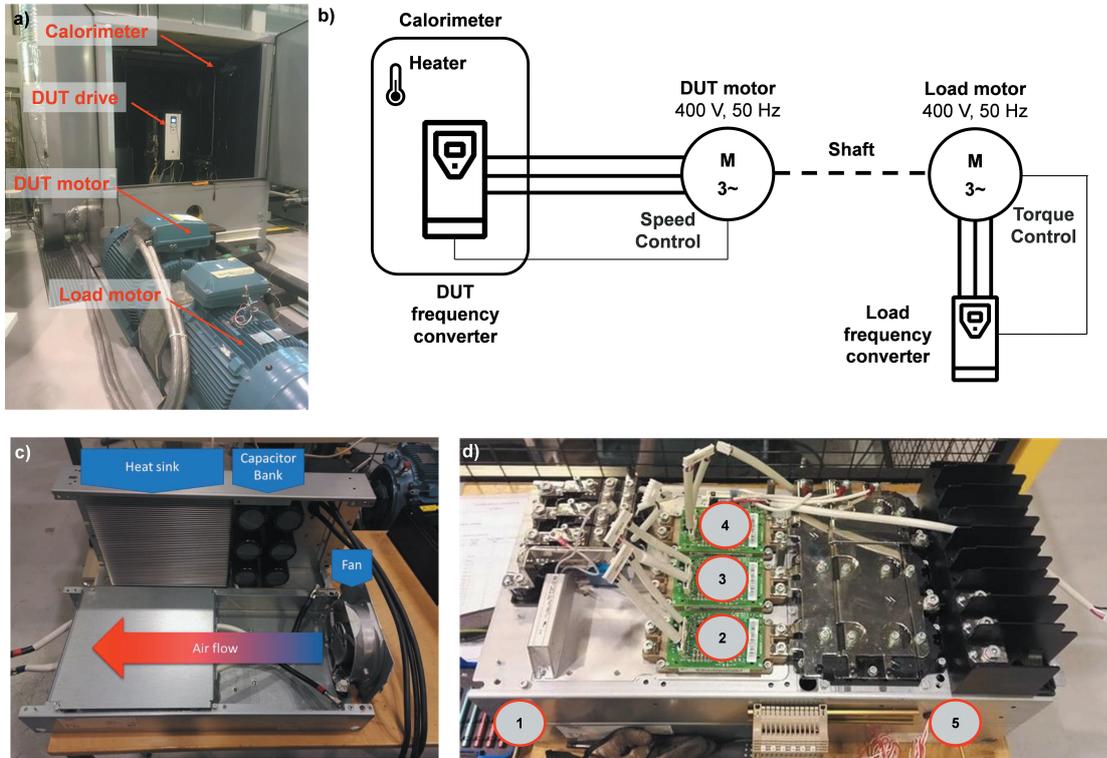


Figure 1. (a) Layout of the experimental bench; (b) graphical sketch of the setup; (c) localisation of the drive's parts that were the focus of the study: heatsink, capacitor bank and fan with the airflow indicated and (d) localisation of PT100 temperature sensors inside the drive's assembly: (1) heatsink temperature (outlet temperature), (2, 3, 4)—IGBT base (in the heatsink under IGBT) and (5)—heatsink temperature (inlet, fan side). DUT, device under test; IGBT, insulated gate bipolar transistor.

profiles in Figure 2a show distinct maximum values for each ambient temperature dataset: 25°C, 35°C and 45°C. The maximum temperature values (such as inverter temperature values shown in Figure 2) increase for the 35°C dataset compared to the 25°C one but are lower again for the 45°C data. This behaviour is expected because, for higher ambient temperatures, the constant baseline in the time series data shifts upward. However, the maximum temperature is constrained by the ventilation system's efficacy. Therefore, upon ambient subtraction, a lower data gap (max–min) is expected in the 45°C data. Testing operation datasets were prepared by joining some of the healthy and failure profiles. The failure profiles were acquired with different clogging conditions and acquired at 25°C, 35°C and 45°C. The testing data are shown in Figures 2b–e for the inlet, outlet, heatsink and fan clogging experiments, respectively. The data were scaled using the quantile-based Robust Scaler with the cut-off thresholds set to 25% and 75% percentiles. The data were then down sampled, so that each 10th point in the data was used. As for the preparation for the training, the data were subjected to segmentation that included generation of sliding window overlapping samples with the buffer size of 25 points ($25 \text{ s} \times 10 \text{ s} = 250 \text{ s}$ data buffer), and the stride (the overlap factor) was 1 point. All data preparation and visualisation procedures were implemented in Python 3.11 using SciPy, NumPy, matplotlib and arviz libraries (Harris et al., 2020; Hunter, 2007; Kumar et al., 2019; Bokeh, 2018).

2.2. Model architecture

In this study, at first, different WGAN architectures were tested using a trial-and-error approach. The model consisted of a generator and a critic as shown in Figure 3. The generator was used during the training phase, but for anomaly detection, only the critic was utilised. The key goal was to find the most stable topologies for the generator and critic where the training of both is balanced. It is important that the final setting of the parameters was achieved using a gradient tracking strategy, so that average per-layer gradient values were of a similar order of magnitude. This assured balanced training and prevented both vanishing and exploding gradients. Figure 3 shows an example WGAN model with generator (left side) and critic (right side), for the latter with a simple topology with one filter

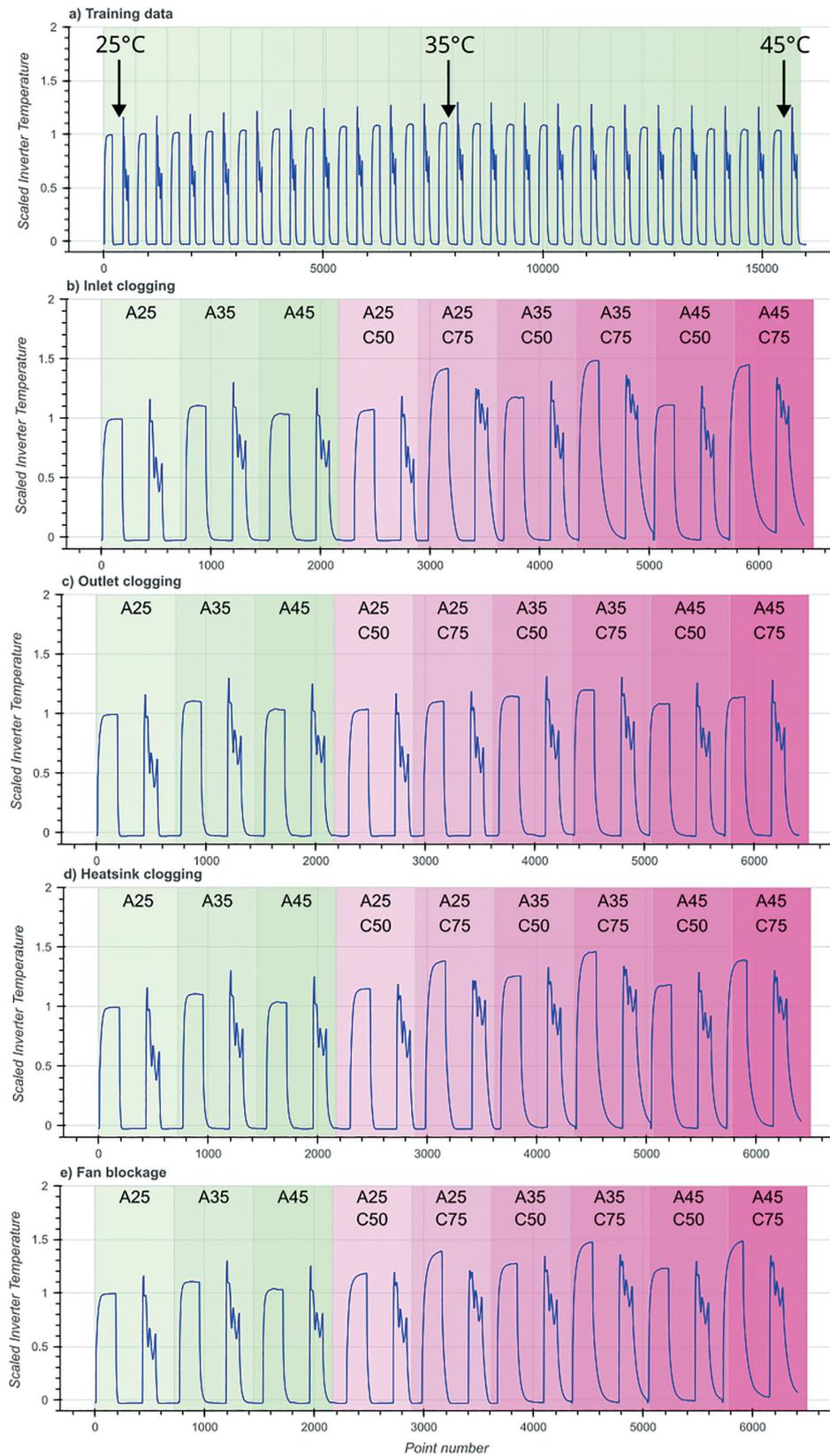


Figure 2. Ambient-subtracted and scaled: (a) training (upon augmentation by 1D interpolation) and testing data involving (b) inlet, (c) outlet, (d) heatsink and (e) fan clogging scenarios.

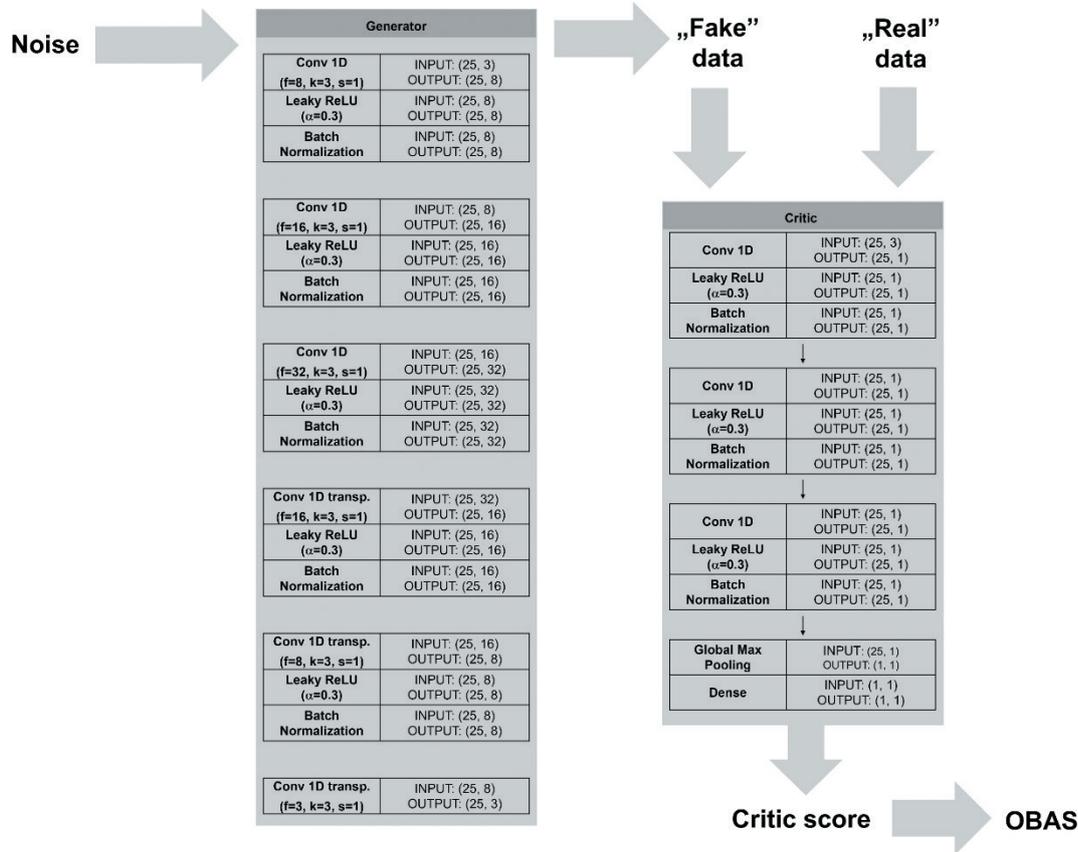


Figure 3. Model architecture and training scheme (the generator was consistent across all tested model architectures, and the critic, varied); the critic presented in this scheme has one filter only. OBAS, out of bounds anomaly scores.

only. Importantly, herein, the generator architecture was fixed, as presented in Figure 3, and critics with different complexities (filter numbers) were tested. The WGAN model developed here was built upon 1D Convolutional layers to allow fast extraction of features from multivariate time series data with fewer parameters compared to recurrent networks. For example, we also tested LSTM layers, but this incurred significant computational overheads both at the training and testing phases. Herein, the generator had an autoencoder-like architecture composed of a series of 1D Convolutional Layers (with 8, 16 and 32 filters with a kernel size of 3 and stride of 1) followed by a series of 1D transposed convolutional layers (with 16, 8 and 3 filters with a kernel size of 3 and stride of 1). In between these, batch normalisation layers were added as they greatly stabilised the training process in terms of less noisy loss values. No activation was applied in the last convolutional layer of the generator, as adding these seemed to impose too much constraint on the final generated data regardless of the activation function (ReLU, LeakyReLU, SeLU, Tanh and Sigmoid). Concerning the critic architectures, we tested three proposed models with varying levels of feature extraction, reflected in the number of filters in the convolutional layers. Table 1 compares the complexity of these three critic architectures, including their number of trainable and non-trainable parameters, as well as the size of the resulting lightweight model encapsulation. Trainable parameters are those that the model learns and updates during the training process, while non-trainable parameters are fixed hyperparameters, such as the number of filters, kernel size and stride in all layers. The proposed critic architectures followed the same topology as presented in Figure 3 (right side) but differed in the number of filters in the convolutional layers. More specifically, they consisted of a tandem of three layers: one 1D convolutional layer, leaky ReLU activation and batch normalisation. In the last layer, a global max pooling layer was applied to take the maximum value from each feature map produced by the convolutional layer. The output was then sent to a dense layer with a single neuron (no activation applied), producing a single scalar value referred to as the raw discriminator/critic output. This numerical output served as a proxy for anomaly strength, expected to be high for anomalies and low for normal data (i.e., data presented during the training process). The proposed M1 critic model had only 1 filter in all convolutional layers,

Table 1. Comparison of complexities of the critic models.

Architecture	No. all parameters	No. of trainable parameters	Size in Kb (tflite)
M1	26	32	5.97
M4	173	197	6.46
M8	537	585	7.79

designed to be lightweight and efficient, yielding only 26 trainable parameters. The M4 and M8 critic models had 4 and 8 filters in all convolutional layers, respectively. These architectures build upon the M1 design by increasing the number of filters to enhance feature extraction capabilities. By increasing the number of filters in the M4 and M8 architectures, we aimed to determine whether a higher number of trainable parameters improves the model's ability to capture more complex patterns. The resultant M4 and M8 models had significantly more trainable parameters: 173 and 573, respectively.

2.3. Model training

The training data were split into batches of 256 samples, with input data formatted as 3D cubes of size (256, 25, 3). The training process spanned $N = 60$ epochs. Models were implemented and trained using Tensorflow (Google Brain Team) without Graphics Processing Unit (GPU) acceleration. To ensure stable training, the learning rates of both the generator and the critic were decayed by a factor of 0.5 every $N = 10$ epochs. The Wasserstein loss function was employed to train both the generator and the critic. The RMSprop optimiser was used for both models, with the generator having a learning rate of $1.0e-6$ and weight decay of $1.0e-2$, and the critic having a learning rate of $1.0e-5$ and weight decay of $1.0e-1$. To enforce the Lipschitz constraint, a gradient penalty term was added to the critic's loss (Bear et al., 2024). For each batch of real data, the critic was trained for $N = 10$ iterations as follows: (a) generate synthetic data from random noise; (b) compute the critic's loss using the Wasserstein loss and the gradient penalty and (c) apply gradient clipping to the critic's gradients and update the critic's weights. The generator was trained for one iteration (in a single epoch) in a similar fashion without the gradient penalty. The model was trained using a personal computer Lenovo ThinkPad (RAM 32 Gb, 11th Gen Intel(R) Core (TM) i7-11850H @ 2.50GHz, NVIDIA T1200). The training time took max 1.5 h.

2.4. Model calibration

To determine whether a specific point or data parcel was anomalous, threshold values were initially derived using critic score values computed from the training data. The 95% quantile values were used as a cut-off. Since it was necessary to analyse the response strength of the critic model to detected anomalies, the scores were converted to a common 'anomaly severity' scale, enabling more robust comparisons. In this study, the out of bounds anomaly scores (OBAS) were applied that we recently proposed (Surówka et al., 2024a,b). OBAS coefficients indicate the strength of an anomaly, specifically how far a numerical value deviates from the normal operation interval, expressed as a multiple of the width of the non-anomalous data range. In this study, we proposed a modified equation for the asymmetric OBAS factor:

$$OBAS(x) = \alpha OBAS(x_i) = \begin{cases} \frac{x_i - x_{t, \min}}{x_{t, \max} - x_{t, \min}} & \text{if } x \leq x_{t, \max} \\ \frac{x_i - x_{t, \max}}{x_{t, \max} - x_{t, \min}} & \text{if } x > x_{t, \max} \end{cases} \quad (2)$$

where x_i denotes a raw anomaly score value (here, the critic score), $x_{t, \min}$ is the minimum threshold value of an anomaly score for normal/training data and $x_{t, \max}$ is the maximum threshold value an anomaly score for normal/training data (here, the 95th percentile). For normal data, $OBAS < 1.0$, while for failure data, $OBAS \gg 1.0$. OBAS here has a minimum value of 0 and increases monotonically with the anomaly score. In previous implementations, the minimum value was set to 0.5, regardless of whether the value was at the minimum or maximum of the healthy data interval.

2.5. Performance analysis of the anomaly detection models

Performance assessment of the critic-based anomaly detection models included impact analysis of both operation mode (start-up vs. cooling vs. active operation) and models' complexity (M1 vs. either M4 or M8). Overall predictive power of the model towards detecting the clogging of a certain level was in focus. The analysis was performed by computing numerical values of accuracy (ACC), true positive rate/recall/sensitivity (TPR), true negative rate/specificity (TNR), false negative rate (FNR), positive predictive value/precision (PPV) and F1 score, as defined in Surówka et al. (2023, 2024a,b). In addition, the PPV and TPR indicators were subjected to harmonic scaling to correct against the imbalance between the number of true positives and true negatives, which impacts on other metrics given as (Flach and Kull, 2015):

$$PPV_s = \frac{PPV - \delta}{(1 - \delta) \cdot PPV} \quad (3a)$$

$$TPR_s = \frac{TPR - \delta}{(1 - \delta) \cdot TPR} \quad (3b)$$

where δ is a scaling factor (here $\delta = 0.5$) to yield balanced proportion between true positives and true negatives (Flach and Kull, 2015).

3. Results and Discussion

Figure 2a shows the stacked, augmented, ambient-subtracted training data (no clogging) corresponding to ambient temperature changes between 25°C and 45°C. It is important to note that we originally had only three profiles for 25°C, 35°C and 45°C. However, the training of the WGANs with such limited data was inefficient with frequent gradient explosions. To somewhat smooth the cost hyperspace, the new data were interpolated from the original profiles, thus introducing data augmentation. By analysing these data, we can see that the net peak temperature values (upon ambient subtraction) were growing until approximately 35 C and then started to slightly decrease (<5%), which was due to the cooling efficiency of the ventilation system in the drive. Figures 2b–d show the applied testing data that were a blend of healthy profiles (25°C, 35°C and 45°C) and those for the 50% and 75% clogging for the inlet (Figure 2b), outlet (Figure 2c), heatsink (Figure 2d) and fan blockage (Figure 2e). Compared with healthy data, the most dramatic changes in the profiles were observed for the heatsink clogging case (Figure 2d); whereas, the least changes were noted for the outlet clogging case (Figure 2c), including both startup/cooling and active operation cycles. Therefore, we could conclude that air flow restrictions in the outlet caused no substantial harm to the performance of the ventilation system up to the clogging limit of 75% (cf. Figure 2c). Above this point, pressure losses and air resistance could become more significant. This could make the air flow turbulent, reducing overall cooling efficiency by the fan (Manaserh et al., 2021; Zmrhal and Boháč, 2021). At the same time, clogging of the inlet restricts the system from cool air with more dramatic impact. The data from Figure 2a data were then used for training and testing of WGAN networks. First of all, we capitalised on developing a WGAN architecture with a small critic model. The goal was to produce the critic model with the smallest number of trainable parameters possible using the 1D convolutional layers, which helped to save on parameter counts. As shown in Table 1, the M1 architecture with a one-filter critic turned out to be the smallest possible. We also tested the M4 and M8 architectures to compare their performance with the M1 model, so that we could assess if higher information capacity of a model brings better performance. By comparing final implementations of these models, their size was < 10 kb, which seems to be sufficient to be handled in low memory systems. With the M1 architecture in place, the major challenge in this study was to perform stable and convergent training of the M1 model while generating data that resembled the training dataset. Figure 4a shows the training curve for the WGAN model of the M1 architecture. These data indicate that the training was convergent for both the generator and the critic, with neither significant vanishing gradients nor gradient explosions. The generator's quality was assessed by generating example data using input normal noise, which were then visualised using the t-distributed stochastic neighbor embedding (t-SNE) transform with two components, as shown in Figure 4b. The distributions of the generated and real training data were found

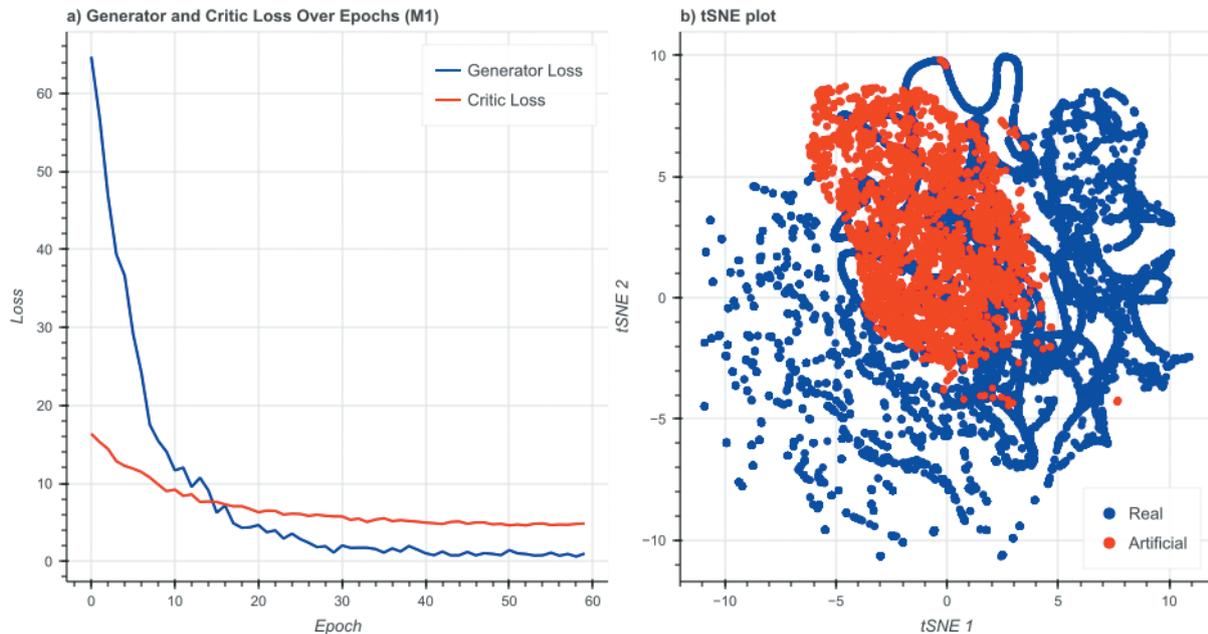


Figure 4. (a) Training curve and (b) tSNE plot for the generated vs. real data for the M1 WGAN model.

to overlap, although the spread of the artificially generated data was smaller than that of the original training data. This suggests that the WGAN captured only the most essential part of the variance in the training data, implying that the M1 model was not overfitted. Figure 5 shows the outcome of the critic-based anomaly detection procedure for the testing data for: the heatsink (cf. Figure 5a), inlet (cf. Figure 5b), outlet (cf. Figure 5c) and fan blockage (cf. Figure 5d) cases. In these figures, given a specific clogging case, critic scores were computed for the data shown in Figures 5b–e and then subjected to the asymmetric OBAS transform (see Eq. 2). The parameters for the OBAS transform were derived from the critic scores computed for the training data (see Figure 2a). Figure 5 also presents the outcomes for more complex architectures, M4 and M8, to evaluate whether adding more complexity enhances performance. These data indicate that the outcomes for the M4 and M8 models were nearly four times weaker than those for M1. This seems to be due to their excessive information capacity and the relatively low volume and complexity of the data, which could lead to undertraining. Undertraining is challenging to conclude definitively, as it depends on the available testing data and can be unmasked by any new datasets. If any anomaly detection model cannot distinguish between healthy and failure (unseen) data (similar anomaly scores assigned to healthy and failure data), likewise presented herein, it indicates a failure to capture essential features. This could be caused by, among others, insufficient training or a too complex structure of a specific model, making generalisation too difficult to achieve, given available data. Despite further attempts at parameter tuning and further training, the M4 and M8 models faced vanishing gradient issues due to their higher number of parameters, making training difficult. The M1 model, with lower information capacity, performed well, showing convergent training and appropriate responses to anomalous data. In contrast, the M4 and M8 models, with 7× and 20× more parameters respectively, struggled due to likely numerical instability, despite attempting to train them at the limits of regularisation. Conversely, the OBAS values for the M1 model were generally higher for clogging data compared to healthy cases, regardless of ambient temperature. The critic-based OBAS scores generally reflected the patterns observed in the data in Figure 2b–e, with the highest OBAS scores for inlet clogging and the lowest for outlet clogging. The highest peak values for OBAS were observed during the start-up and cooling operation phases, while responses during active operation were just slightly above the threshold. Figure 5 shows that the M1 model generally remained silent ($OBAS < 1$) for no-clogging data, regardless of ambient temperature. However, for inlet clogging, as presented in Figure 5a, the anomaly indicator far exceeded the threshold starting from the 75% clogging data acquired at 25°C during both the startup and active operation phases, although the responses were borderline for the latter. As shown in Figure 5b, the M1 model indicated weak/boundary ($OBAS \sim 1$) outlet clogging anomalies starting from the 75% clogging case at 25°C, with maximum indications for the 75% clogging case at 35°C. Figure 5c demonstrates that

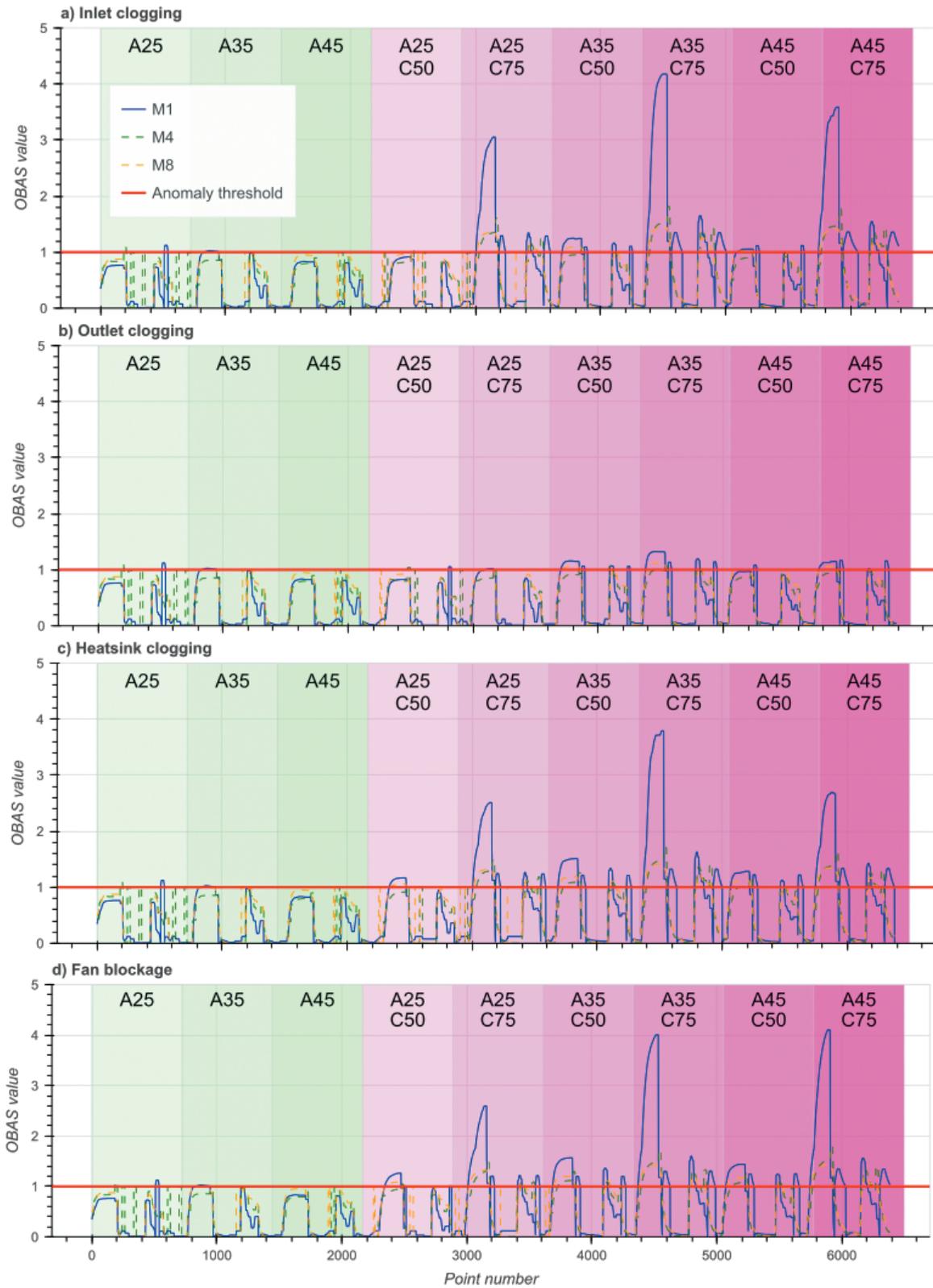


Figure 5. Values of critic-based OBAS for the VFD system affected by (a) heatsink; (b) outlet; (c) heatsink clogging and (d) fan blockage. The time periods during which the system operates at ambient temperatures of 25 C, 35 C and 45 C are shaded in green. Anomalous operations due to clogging are highlighted with magenta boxes. The letter 'A' indicates ambient temperature, while 'C' denotes the clogging/blockage level (50% vs. 75%). OBAS, out of bounds anomaly scores; VFD, variable frequency drive.

for the heatsink clogging case, the M1 model responded to all clogging cases, with the strongest response during the startup/early cooling operation phases. Again, the responses to anomalous cases were just above the threshold during active operation (OBAS in the range [1, 2]). The results for fan blockage, as shown in Figure 5d, were generally similar to those for inlet and heatsink clogging scenarios, with OBAS values increasing for the 75% clogging case as the ambient temperature increased. Overall, the results from Figure 5 demonstrate that the M1 model can generally detect all clogging cases; whereas, the responses for the M4 and M8 architectures were at most borderline. This indicates that increasing the number of parameters (filters) in the critic model does not improve performance. By comparing these outcomes with our previous study, we can conclude that the critic response to anomalies was weaker than that of HMM (Surówka et al., 2024a,b). The latter produced very strong responses up to $OBAS \sim 10^2$ for 75% clogging cases in the inlet and heatsink. However, HMMs fell short in detecting some clogging scenarios: the outlet clogging was almost undetectable, and they struggled to respond to 50% clogging cases (Surówka et al., 2024a,b). In contrast, the critic-based WGAN model was more balanced in its response: OBAS indicators were significantly smaller ($OBAS < 4$), but they responded to almost all clogging cases, including the outlet clogging and 50% clogging cases missed by HMMs (Surówka et al., 2024a,b). Additionally, the critic scores were more reflective of the mutual difference in temperature readouts between anomalous and non-anomalous data, ranging between 5% and 10%, compared to HMMs where the anomaly scores were exaggerated (Surówka et al., 2024a,b). This reflects the nature of the feedback-based training, so that the 'critic' level, as the name suggests, is proportional to the effect the same way a human expert will do. Another question we posed was the minimal detectable clogging/blockage level that could be identified using our proposed model architecture (M1). This turned out to be a difficult question due to limited data: we only had profiles acquired for 25°C, 35°C and 45°C including no clogging and clogging data. However, we found out that we can interpolate between these boundary profiles with three anchors: 0% clogging (just ambient data), 50% clogging and 75% clogging data, given a specific ambient temperature. The interpolation covered ranging from 0% to 75% with 5% increments. The M1 critic was then applied to these data, including the OBAS transform, as shown in Figure 6 for ambient temperature values of 25°C (cf. Figure 6a), 35°C (cf. Figure 6b) and 45°C (cf. Figure 6c). Generally, the response to anomalies at given clogging levels varied with different ambient temperatures. From Figure 6a, we can see that at 25°C, the anomaly threshold was exceeded at 25% fan clogging/blockage, 30% heatsink clogging and 55% inlet clogging. However, the outlet clogging was almost undetectable at 25°C, with weak anomalies appearing from 70% on. At 35°C (cf. Figure 6b), stronger responses to clogging were observed, and all clogging scenarios could be detected starting from the 5% level. However, caution is needed as the indications for 0% clogging just hit the anomaly threshold. At 45°C (cf. Figure 6c), fan blockage could be detected from 25%, heatsink clogging from 30%, inlet clogging from 45% and outlet clogging from 40% (with a borderline outcome). In total, the data in Figure 6 indicate that the critic model is reactive from roughly 20% to 30% clogging levels, with the response being influenced to some extent by ambient conditions. To further understand this behaviour, we quantitatively assessed the anomaly detection capabilities of our low-parameter critic-based WGAN model (M1) by comparing performance metrics, such as ACC, TPR, FNR, PPV, and F1, as shown in Table 2 (input data from Figure 5). The metrics were computed for each clogging scenario separately. As the non-anomalous data ranges and the proportion of positive and negative data were the same in these plots, therefore computing FPR would not indicate any meaningful differences between different scenarios and was therefore skipped. For inlet, heatsink and fan clogging cases, the best outcomes were noted for startup and cooling phases. These phases exhibited 10-20% higher ACC and significantly higher TPR compared to the active operation phase. It can also be drawn from these data that the performance of outlet clogging detection was low, although slightly higher than a random classifier with 50% accuracy. Additionally, there was good confidence in positive predictions, as reflected by PPV values ranging from 0.7 to 1.0 for all cases except outlet clogging. Starting with the inlet clogging prediction of the heatsink, as shown in Table 2, the models achieved good ACC of 76%, TPR between 82% and 84% and FPR of 10%–30% for both startup and cooling phases. Regarding the outlet clogging case, the performance figures indicated that the model had only a weak detection capability, reflected by low F1 scores between 0.2 and 0.6. For heatsink detection during startup and cooling phases, the performance was excellent, with an accuracy of nearly 90% and an F1 score above 0.9. Similarly, for fan blockage detection, the model achieved an F1 score of 0.87 for both startup and cooling phases. Overall, Table 2 shows that the developed clogging detection model had higher accuracy and TPR during startup and cooling phases; whereas, the active phase exhibited lower performance metrics, particularly for outlet clogging detection. These results align with our previous study on clogging detection in VFDs using HMMs, where higher performance figures were also concluded for the startup phase compared to active operation. The only difference is that HMMs

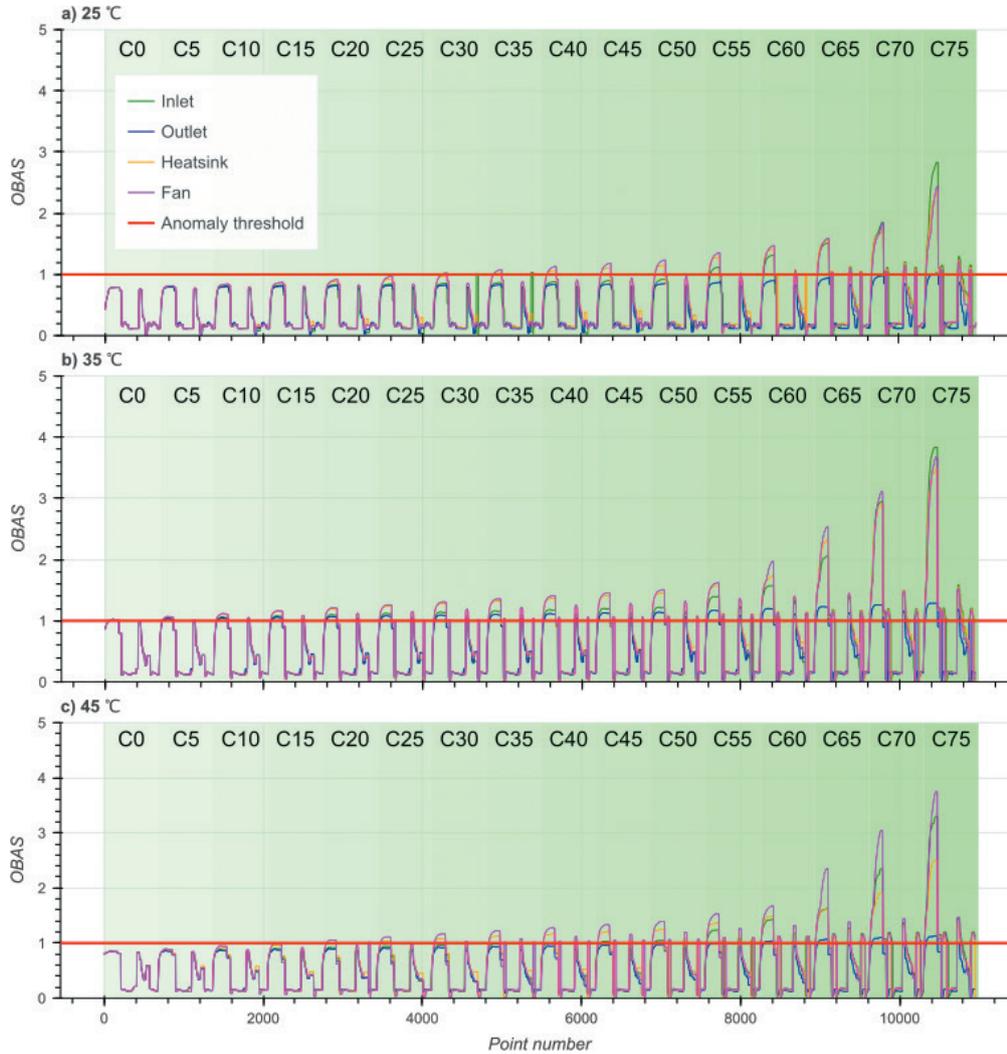


Figure 6. Analysis of the minimal clogging/blockage level detectable in the analysed VFD under different ambient conditions: (a) 25 C, (b) 35 C, and (c) 45 C. The letter 'C' indicates the clogging/blockage level for which the data are presented. OBAS, out of bounds anomaly scores; VFD, variable frequency drive.

Table 2. Outcomes of the performance assessment of the M1 model.

Type	Phase	ACC	TPR	FNR	PPV	F1
Inlet	Start	0.78	0.79	0.21	0.86	0.83
	Cooling	0.78	0.84	0.16	0.83	0.84
	Active	0.57	0.36	0.64	0.91	0.51
Outlet	Start	0.65	0.55	0.45	0.81	0.65
	Cooling	0.67	0.67	0.32	0.80	0.73
	Active	0.44	0.14	0.85	0.82	0.24
Heatsink	Start	0.90	0.98	0.02	0.88	0.93
	Cooling	0.88	1.00	0.00	0.84	0.91
	Active	0.53	0.95	0.69	0.91	0.46
Fan	Start	0.87	1.00	0.00	0.79	0.86
	Cooling	0.84	1.00	0.00	0.76	0.86
	Active	0.66	0.32	0.68	1.00	0.48

FNR, false negative rate; PPV, positive predictive value/precision; TPR, true positive rate/recall/sensitivity.

were more efficient in detecting inlet clogging, while in this study, substantially better performance was reported for heatsink clogging detection. It shall also be mentioned that the previously applied 10–40 states HMMs had substantially higher number of parameters ranging from 100 to 1,500 (Surówka et al., 2024a,b). Concerning other studies, to our knowledge, there is little evidence of the application of critic-based WGAN models for condition monitoring/anomaly detection of power electronics in VFDs. Nevertheless, Zhang and Yang (2023) have recently presented a prototype Long Short-Term Memory-Based Variational Autoencoder Wasserstein Generation Adversarial Network for anomaly detection in wind turbines based on supervisory control and data acquisition (SCADA) data involving wind speed, output power and generator input/output shaft temperature among others. The authors achieved TPR above 90%, PPV between 50% and 70% and F1 score between 0.7 and 0.8, all of which comparable with the data presented in this study (Zhang and Yang, 2023). Recently, Dai et al. (2024) presented a novel WaveGAN architecture that scored TPR, ACC and F1 of over 90% using open-source datasets. It should be noted that Zhang and Yang (2023) and Dai et al. (2024) used a rather large GAN architectures with tens of thousands of parameters, utilising 2D convolutional layers and intermediate transformation of data to images using, for example, the wavelet transform. In contrast, in this study, we managed to prepare a model with less than 30 parameters in the critic, without intermediate transformation of the time series data, and with memory-efficient 1D convolutional layers.

4. Conclusions

In this study, we demonstrated the feasibility of developing a 26-parameter critic-based model for clogging detection in drives. The key feature of this model is not its complexity, but the feedback-based training approach, which mimics a human-like learning and inference process. The model proved effective in detecting fan failure and heatsink/inlet clogging, although it was less successful in detecting outlet clogging. The performance metrics achieved were comparable to other developed solutions, with top scores obtained for both fan failure and heatsink clogging detection. Additionally, we showed that the minimum detectable clogging level ranged between 10% and 30%, depending on the scenario. It was also shown that clogging detection is successful during the startup or cooling phases of drive operations, with moderate outcomes noted during active operation. Overall, this study demonstrates that it is possible to develop a low-parameter WGAN model for clogging detection in drives and paves the way for generative AI models for anomaly detection in electric machines.

References

- Al-Naseem, O. A. and El-Sayed, M. A. (2013). Analysis of electrical and non-electrical causes of variable frequency drive failures. In 2013 9th IEEE International Symposium on Diagnostics for Electric Machines, *Power Electronics and Drives (SDEMPED)* (pp. 221-226). IEEE.
- Avor, J. K. and Chang, C.-K. (2019). Reliability Analysis of Application of Variable Frequency Drive on Condensate Pump in Nuclear Power Plant. *Journal of International Council on Electrical Engineering*, 9(1), pp. 8–14. doi: 10.1080/22348972.2018.1564548
- Bear, J., Prügel-Bennett, A. and Hare, J. (2024). Rethinking Deep Thinking: Stable Learning of Algorithms using Lipschitz Constraints. Proceedings of 38th Conference on Neural Information Processing Systems (NeurIPS 2024), arXiv:2410.23451. <https://doi.org/10.48550/arXiv.2410.23451>
- Beattie, A., Mulink, P., Sahoo, S., Christou, I. T., Kalalas, C., Gutierrez-Rojas, D., and Nardelli, P. H. (2022). A Robust and Explainable Data-Driven Anomaly Detection Approach for Power Electronics. In 2022 IEEE International Conference on Communications, Control, and Computing Technologies for Smart Grids (SmartGridComm) (pp. 296-301). IEEE.
- Bokeh Development Team (BDT) (2018). Bokeh: Python library for interactive visualization. <http://www.bokeh.pydata.org>.
- Chandola, V., Banerjee, A. and Kumar, V. (2009). Anomaly Detection: A Survey. *ACM Computing Surveys (CSUR)*, 41(3), pp. 1–58. doi: 10.1145/1541880.1541882
- Chen, Y., Gao, Q. and Wang, X. (2022). Inferential Wasserstein Generative Adversarial Networks. *Journal of the Royal Statistical Society Series*

- B: Statistical Methodology*, 84(1), pp. 83–113. doi: 10.1111/rssb.12476
- Chollet, F. (2018). *Deep Learning with Python*. New York, Manning Publications.
- Ciappa, M. and Fichtner, W. (2000). Lifetime Prediction of IGBT Modules for Traction Applications. *IEEE International Reliability Physics Symposium Proceedings*, San Jose, CA, USA.
- Dai, H., Wang, J., Zhong, Q., Chen, T., Liu, H., Zhang, X. and Lu, R. (2024). A GAN-Based Anomaly Detector Using Multi-Feature Fusion and Selection. *Scientific Reports*, 14, p. 52378. <https://doi.org/10.1038/s41598-024-52378-9>
- Flach, P. and Kull, M. (2015). Precision-recall-gain curves: PR analysis done right. In: *Proceedings of Advances in Neural Information Processing Systems 28 (NIPS 2015)*, Montreal, Quebec, Canada.
- Gómez, P. I., López, G. M. E., Mijatovic, N. and Dragičević, T. (2024). A Self-Commissioning Edge Computing Method for Data-Driven Anomaly Detection in Power Electronic Systems. *IEEE Transactions on Industrial Electronics*, vol. 71, no. 10, pp. 13319–13330. doi: 10.1109/TIE.2023.3347839
- Goodfellow, I., Bengio, Y. and Courville, A. (2016). *Deep Learning*. Cambridge, US, MIT Press.
- Gulrajani, I., Ahmed, F., Arjovsky, M., Dumoulin, V., Courville, A. (2017). *Improved Training of Wasserstein GANs*. arXiv, arXiv:1704.00028, doi: 10.48550/arXiv.1704.00028
- Harris, C., Millman, K., van der Walt, S., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M. H., Brett, M., Haldane, A., Del Río, J. F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C. and Oliphant, T. E. (2020). Array Programming with NumPy. *Nature*, 585, pp. 357–362. doi: 10.1038/s41586-020-2649-2
- Hunter, J. D. (2007). Matplotlib: A 2D Graphics Environment. *Computing in Science & Engineering*, 9(3), pp. 90–95. doi: 10.1109/MCSE.2007.55
- Kang, Y., Dang, L., Yang, L., Wang, Z., Meng, Y., Li, S., Sun, Y., Wang, Y., and Dong, H. (2023). Research Progress in Failure Mechanism and Health State Evaluation Index System of Welded IGBT Power Modules. *Electronics*, 12(15), p. 3248. doi: 10.3390/electronics12153248
- Kennedy, D. (2021). *Common Causes of VFD Failure* [online]. Available at: <https://goemc.com/2021/05/20/common-causes-of-vfd-failure/> [Accessed 12 Feb, 2025].
- Kumar, R., Carroll, C., Hartikainen, A. and Martin, O. (2019). ArviZ a Unified Library for Exploratory Analysis of Bayesian Models in Python. *Journal of Open Source Software*, 4(33), p. 1143. doi: 10.21105/joss.01143
- León-López, K. M., Mouret, F., Arguello, H. and Tourneret, J.-Y. (2021). Anomaly Detection and Classification in Multispectral Time Series Based on Hidden Markov Models. *IEEE Transactions on Geoscience and Remote Sensing*, 60, p. 5402311. doi: 10.1109/TGRS.2021.3101127
- Lüer, F. and Böhm, C. (2021). Anomaly Detection using Generative Adversarial Networks: A Review. In: *Proceedings of the 27th ACM SIGKDD Conference on Knowledge Discovery & Data Mining*, Singapore.
- Manaserh, Y. M., Tradat, M. I., Hoang, C. H., Sammakia, B., Ortega, A., Nemat, K., Seymour, M. J. (2021). Degradation of Fan Performance in Cooling Electronics: Experimental Investigation and Evaluating Numerical Techniques. *International Journal of Heat and Mass Transfer*, 174, p. 121291. doi: 10.1016/j.ijheatmasstransfer.2021.121291
- Otsuki, M., Onozawa, Y., Kanemaru, H., Seki, Y., Matsumoto, T. (2003). A Study on the Short-Circuit Capability of Field-Stop IGBTs. *IEEE Transactions on Electron Devices*, 50, pp. 1525–1531. doi: 10.1109/TED.2003.813505
- Peterson, D. (2022). Teardown: What's Inside a Variable Frequency Drive (VFD)? *Control Automation*, technical article, website: <https://control.com/technical-articles/teardown-whats-inside-a-vfd/>, accessed on 05-05-2025
- Qi, S., Chen, J., Chen, P., Wen, P., Shan, W. and Xiong, L. (2023). *An Effective WGAN-Based Anomaly Detection Model for IoT Multivariate Time Series*. Kyoto, Japan: Pacific-Asia Conference on Knowledge Discovery and Data Mining.
- Surówka, A., Tan, R., Saberi, A., and Firla, M. (2023). Performance of machine-learning-based algorithms for anomaly detection in variable frequency drives using temperature signals. In: *2023 IEEE 14th International Symposium on Diagnostics for Electrical Machines, Power Electronics and Drives (SDEMPED)*, Chania, Greece, 28–31 August 2023.
- Surówka, A., Mikkilä, T., Kavala, A., and Firla, M. (2024a). Dual-mode hidden Markov models for smart detection of clogging in variable frequency drives. In: *IEEE 21st International Power Electronics and Motion Control Conference*, Pilsen, Czech Republic, 2024.
- Surówka, A., Tan, R., Saberi, A. and Firla, M. (2024b). Out of Bounds Anomaly Scores in Anomaly

- Detection in Variable Frequency Drives Using Temperature Signals. *IEEE Transactions on Industry Applications*, 60(5), pp. 6988–7000. doi: 10.1109/TIA.2024.3427712
- Tang, S., Shi, H., Song, B., Tao, Y., and Tan, S. (2025). Physically-Consistent-WGAN Based Small Sample Fault Diagnosis for Industrial Processes. *Chinese Journal of Chemical Engineering*, 78, pp. 163–174. doi: 10.1016/j.cjche.2024.10.028
- Tuli, S., Casale, G. and Jennings, N. (2022). TranAD: Deep Transformer Networks for Anomaly Detection in Multivariate Time Series Data. *arXiv preprint*, p. 2201.07284. doi: 10.48550/arXiv.2201.07284
- Xia, X., Pan, X., Li, N., He, X., Ma, L., Zhang, X., and Ding, N. (2023). GAN-Based Anomaly Detection: A Review. *Neurocomputing*, 493, pp. 497–535. doi: 10.1016/j.neucom.2021.12.093
- Xu, L., Xu, K., Qin, Y., Li, Y., Huang, X., Lin, Z., and Ji, X. (2022). TGAN-AD: Transformer-Based GAN for Anomaly Detection of Time Series Data. *Applied Sciences*, 12(16), 8085. doi: 10.3390/app12168085
- Yellamati, D., Arthu, E., James, S., Morris, G., Heydt, T., and Graf, E. (2013). Predictive Reliability Models for Variable Frequency Drives Based on Application Profiles. Orlando, FL: USA.
- Yinka-Banjo, C., and Ugot, O.A. (2020). A review of generative adversarial networks and its application in cybersecurity. *Artif Intell Rev* 53, 1721–1736. doi: 10.1007/s10462-019-09717-4
- Zhang, C. and Yang, T. (2023). Anomaly Detection for Wind Turbines Using Long Short-Term Memory-Based Variational Autoencoder Wasserstein Generation Adversarial Network under Semi-Supervised Training. *Energies*, 16(19), p. 7008. doi: 10.3390/en16197008
- Zhang, M., Gómez, P. I., Xu, Q. and Dragicevic, T. (2023). Review of Online Learning for Control and Diagnostics of Power Converters and Drives: Algorithms, Implementations and Applications. *Renewable and Sustainable Energy Reviews*, 186, p. 113627. doi: 10.1016/j.rser.2023.113627
- Zmrhal, V. and Boháč, J. (2021). Pressure Loss of Flexible Ventilation Ducts for Residential Ventilation: Absolute Roughness and Compression Effect. *Journal of Building Engineering*, 44, p. 103320. doi: 10.1016/j.jobe.2021.103320