Proc. of the European Safety and Reliability Conference (ESREL 2026)
Edited by José C. Matos, Paulo B. Lourenço, Michael Beer, Daniel Oliveira,
Edoardo Patelli, Hélder S. Sousa, Rui A. Silva and Monica Santamaria

doi: 10.3850/ESREL2026061419_esrel26-p30437-cd

Data Augmentation by Generative Adversarial Networks for Fault Detection in Power Production System

Alice Andolfi
Department of Energy, Politecnico di Milano, Milan, Italy. E-mail: alice.andolfi@polimi.it

Piero Baraldi
Department of Energy, Politecnico di Milano, Milan, Italy. E-mail: piero.baraldi@polimi.it

Giovanni Floreale
Department of Energy, Politecnico di Milano, Milan, Italy. E-mail: giovanni.floreale@polimi.it

Enrico Zio
MINES Paris-PSL, Centre de Recherche sur les Risques et les Crises (CRC), Sophia Antipolis, France
Department of Energy, Politecnico di Milano, Milan, Italy. E-mail:enrico.zio@polimi.it

Abstracts: We consider the problem of fault detection (FD) in power production systems (PPSs) working under variable conditions and the issue that, in practice, data from operating conditions are not available. To address this issue, a Generative Adversarial Network (GAN) is developed for generating synthetic data of normal system operation, leading to the augmentation of the available dataset. The augmented dataset is, then, processed by an Autoencoder (AE) for the reconstruction of the signal values expected in normal system operation to enable FD. The proposed modelling framework is applied to a high-fidelity model of a 4.8 MW wind turbine. Satisfactory FD performances are obtained under all considered wind regimes, including rare high-wind conditions. Specifically, the Area Under the Curve (AUC) increases from 0.84 to 0.86 when GAN generated data are added to the training set, and from 0.78 to 0.83 on high wind conditions.

Keywords: Power Production System, Wind Turbines, Fault Detection, Data Augmentation, GAN

1. Introduction

Failures in Power Production System (PPS) can cause large downtimes and costs (Baschel et al. 2018; Beganovic and Söffker 2016). To counteract this, Fault Detection (FD) methods leverage data collected by Supervisory Control and Data Acquisition (SCADA) systems to alert plant operators to the presence of anomalies. The early warning enables operators to intervene, ultimately reducing the risks and costs to an associated failure (Yan, Dunnett, and Jackson 2023).

Deep Learning (DL)-based approaches, such as Autoencoders (AEs) and Deep Support Vector Data Description (SVDD) have been applied for FD (Ruff et al. 2018) (Floreale et al. 2024). Considering AEs, they are used to reconstruct the signal values as expected in normal conditions of system operation and early-identify incident fault conditions by the difference between the original and the reconstructed data (Peng, Desmet, and Gryllias 2025). FD in PPSs such as Wind Plants (WPs) is challenging because operating conditions vary over time and the data available to train the model may not cover sufficiently all the conditions that are experienced by the system during operation (Liu, Hu, and Liu 2025; Zhang and Jiang 2024; Wu et al. 2023). As a result, the FD model may overfit the training data, failing to capture general relationships, resulting in excessive false alarms in those cases of previously unseen or rare operating conditions. A possible solution is to resort to Data Augmentation (DA) techniques (Perez and Wang 2017) to generate new synthetic patterns similar to the original data and, then, train the FD model using the union of original and generated data. To do this, the generated data should have two characteristics: fidelity, i.e. they should be indistinguishable from the original data, and diversity, i.e. they should provide meaningful variations rather than simple copies of the existing patterns (Shorten and Khoshgoftaar 2019; Olson, Wyner, and Berk 2018).

Although DA is a well-established practice in image classification, its use for tabular data is still limited. Simple techniques such as jittering (Fields, Hsieh, and Chenou 2019), window cropping or slicing (Le Guennec, Malinowski, and Tavenard 2016) can introduce synthetic diversity but may fail to capture the underlying data distribution. Alternatively, DL approaches based on adversarial learning have been proposed (Chu et al. 2020). Among them, we consider Generative Adversarial Networks (GANs), which are formed by a generator, whose objective is producing patterns from the same distribution of the original data, and a discriminator, whose objective is to provide the probability that a generated data comes from the same distribution (Goodfellow et al. 2014).

In this work, we develop a novel FD method based on: 1) a GAN for generating synthetic normal conditions data; 2) an AE, trained using the real and GAN generated normal conditions data, for reconstructing the signal values as expected in normal conditions of system operation. We evaluate the performance of the proposed method considering a Wind Turbines (WT) dataset provided by a high-fidelity Simulink model (Odgaard, Stoustrup, and Kinnaert 2013). In the dataset, the wind speed is imbalanced, with high-wind regimes occurring rarely. The structure of the work is as follows. Section 2 introduces the FD method based on signal reconstruction. Section 3 describes the developed method. Section 4 presents the wind turbine case study and discusses the obtained results. Finally, Section 5 summarizes the main conclusions and outlines future research directions.

2. Fault Detection by Signal Reconstruction

In general terms, we consider a generic PPS working under variable operating conditions and monitored by M sensors measuring signals at regular frequency. The vector of measured signal values at time step t is $\boldsymbol{x}^t = [x_1^t, \dots, x_m^t, \dots, x_M^t] \in \mathbb{R}^M$, where x_m^t is the value of the m-th signal measured at time t. The matrix $\boldsymbol{X} \in \mathbb{R}^{N \times M}$ contains N M-dimensional vectors $\boldsymbol{x}^\tau$ of measurements collected at times $\tau \in [1, N]$ while the system was operating in normal conditions. We realistically assume that $\boldsymbol{X}$ contains only few patterns collected in rare operating conditions and that signal measurements collected in fault conditions are not available.

The objective of this work is to develop a FD method to promptly identify the occurrence of a fault. This problem is addressed by developing a data-driven signal reconstruction model that receives in input the vector $\boldsymbol{x}^t$ of the current signal measurements and produces in output the vector $\hat{\boldsymbol{x}}^t$ of the signal values as expected in normal conditions of operation:

$$\hat{\boldsymbol{x}}^t = f_{\boldsymbol{\theta}}^{A}(\boldsymbol{x}^t) \tag{1}$$

where $\boldsymbol{\theta}$ is the vector of the model internal parameters and $\boldsymbol{A}$ denotes the model architecture. The residuals between the signal measurements and their reconstructions at time t:

$$\boldsymbol{r}^t = \boldsymbol{x}^t - \hat{\boldsymbol{x}}^t, \boldsymbol{r}^t \in \mathbb{R}^M \tag{2}$$

are, then, used to construct a Health Index (*HI*):

$$HI^t = \boldsymbol{r}^t (\boldsymbol{r}^t)^\top \tag{3}$$

Finally, the binary health state σ^t of the system at time *t*:

$$\sigma^t = \begin{cases} 0 & Normal\ Condition \\ 1 & Fault\ Condition \end{cases} \tag{4}$$

is identified by comparing the *HI* at time *t*, with a preset threshold th :

$$\sigma^t = \begin{cases} 0 & if\ \ HI^t < th \\ 1 & if\ \ HI^t > th \end{cases} \tag{5}$$

3. Proposed Method

This work utilizes a Generative Adversarial Network (GAN) (Goodfellow et al. 2014) for data augmentation and an Autoencoder (AE) for signal reconstruction.

3.1. *GAN for generation of normal conditions Data*

To augment the normal conditions dataset $\boldsymbol{X}$, we develop a Generative Adversarial Network (GAN). It is composed of two multilayer perceptron models: a Generator G and a Discriminator D. G aims to produce patterns $\boldsymbol{x_g}$ from the distribution p_x of the normal conditions data $\boldsymbol{X}$, whereas D provides in output the probability that a generated pattern $\boldsymbol{x_g}$ indeed comes from the distribution p_x (Goodfellow et al. 2014). The input of G consist of independent and identically distributed (i.i.d.) vectors $\boldsymbol{z} \in \mathbb{R}^{M_z}$ drawn from a standard Gaussian distribution, $p_{z(z)}$. During training, G is optimized to learn a distribution p_g that approximates p_x, whereas D is optimized to distinguish between original and generated data. Training proceeds iteratively alternating steps between G and D. First, the generator G is held fixed and the discriminator D internal parameters are updated to maximise the log-likelihood of assigning the correct label $D(\boldsymbol{x}) = 1$ to original patterns $\boldsymbol{x} \sim p_x$ and $D(\boldsymbol{x_g}) = 0$ to generated patterns $x_g = G(\boldsymbol{z})$, with $\boldsymbol{z} \sim p_z$; then, D is fixed and G is updated with the objective of generating data that are misclassified by D as original. This interaction is formalized as a zero-sum game represented by a minmax optimization problem, whose dynamics aim reaching an equilibrium maintaining a balance between losses of G and D. The Value function $V(G, D)$ of G and D is:

$$\min_G \max_D V(D,G) = E_{\boldsymbol{x}\sim p_x}[\log D(\boldsymbol{x})] + E_{\boldsymbol{z}\sim p_z(\boldsymbol{z})}\left[\log\left(1 - D\left(G(\boldsymbol{z})\right)\right)\right] \quad (6)$$

Following the successful training of the GAN, the Generator G can generate new samples pertaining to the learned distribution p_g from noise vectors $\boldsymbol{z}$.

3.2. *Autoencoder for signal reconstruction*

The signal reconstruction model, $\widehat{\boldsymbol{x}}^t = f_{\boldsymbol{\theta}}^A(\boldsymbol{x}^t)$, used in this work is an Autoencoder (AE), which is a neural network trained to reconstruct in output the vector $\boldsymbol{x}^t$ received in input. The internal parameters, $\boldsymbol{\theta}$, are set by means of the gradient descent algorithm, where the gradient is computed using the error backpropagation algorithm with the objective of minimizing the reconstruction error $(\widehat{\boldsymbol{x}^\tau} - \boldsymbol{x}^\tau)$ over a training dataset of N_{Train} patterns $\boldsymbol{x}^\tau \in \boldsymbol{X_{Train}}$ of normal conditions data. Specifically, in this work, the normal conditions dataset $\boldsymbol{X_{Train}}$ is made of the union of the real data $\boldsymbol{X}$ and the data $\boldsymbol{X_g}$ generated by the GAN. The loss function is the Mean Square Error (MSE):

$$\mathcal{L}_{rec} = \frac{1}{N_{train}} \cdot \sum_{\tau=1}^{N_{train}} \left(\left\|\widehat{\boldsymbol{x}^\tau} - \boldsymbol{x}^\tau\right\|\right)^2 \quad (7)$$

4. Case Study

The case study concerns a three-bladed, pitch-controlled, variable-speed wind turbine of nominal power $P = 4.8\ MW$. The WT behaviour is simulated by using a Simulink model (Odgaard, Stoustrup, and Kinnaert 2013). The input is a vector of wind speed values, and the output is the 4400s timeseries of the plant signals reported in Table 1, measured at a frequency of 3 Hz.

Table 1 List of signals and location of the sensors

Signal ID	**Quantity**	**Acronym**	**Location**
1	Wind Speed	v_w	Blades
2	Rotational speed	ω_{r_1}	Rotor
3	Rotational speed	ω_{r_2}	Rotor
4	Rotational speed	ω_{g_1}	Generator
5	Rotational speed	ω_{g_2}	Generator
6	Generator Torque	τ_g	Generator
7	Produced Power	P_g	Generator
8	Pitch position	$\beta_{1,1}$	Blade 1
9	Pitch position	$\beta_{1,2}$	Blade 1
10	Pitch position	$\beta_{2,1}$	Blade 2
11	Pitch position	$\beta_{2,2}$	Blade 2
12	Pitch position	$\beta_{3,1}$	Blade 3
13	Pitch position	$\beta_{3,2}$	Blade 3

14	Reference Torque	τ_r	Generator
15	Reference pitch position	β_r	Blades

The turbine operates at variable speed, by either controlling the angle of its blades or by changing its rotational speed through the drivetrain, efficiently capturing wind energy in varying wind conditions.
The simulator has been used to produce the dataset $\boldsymbol{X}$ of 7951 patterns in normal conditions, used to develop the FD method. Figure 1 shows the distribution of wind speed in the dataset $\boldsymbol{X}$, which can be split at $v^* = 15\left[\frac{m}{s}\right]$ into two wind regimes:

- Low-wind regime: $v_{wind} < v^*$, corresponding to wind conditions well covered in the available data distribution.
- High-wind regime: $v_{wind} > v^*$ corresponding to rare high-wind conditions.

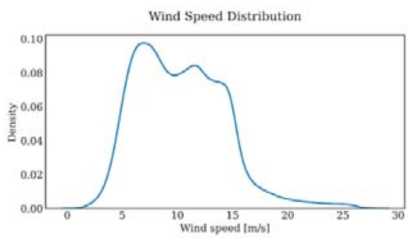

Figure 1 Wind Speed Distribution

The test set consists of 112 simulated timeseries sequences. They include 12 simulations in normal conditions and 100 which contain fault conditions caused by F=10 failure modes (Pinciroli Luca, Baraldi Piero, and Zio Enrico 2025). Starting time and length of the time windows in which the fault happens have been sampled from uniform distributions $U(0{,}4000)s$ and $U(134{,}400)s$ respectively. Figure 2 shows the architecture of the developed GAN, which is characterized by the inclusion of a Batch Normalization layer to enhance gradient flow and mitigate mode collapse (Radford, Metz, and Chintala 2016).

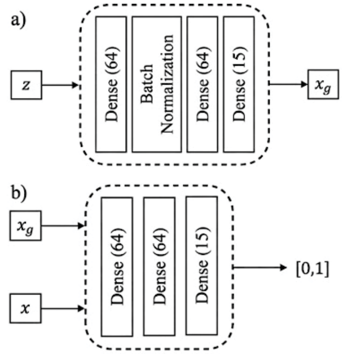

Figure 2 Generator and Discriminator architectures: (a) Generator network; (b) Discriminator network

The hyperparameters employed for the GAN training are reported in Table 2. To ensure training stability, G is optimized using an adaptively decreasing learning rate, whereas D is trained with a constant and larger learning rate. This follows the Two Time-Scale Update Rule (TTUR) introduced by (Heusel et al. 2018), which proves that GANs converge when G and D are trained with separate learning rates.
Training the final model required approximately 12 hours on a Google Colab Pro environment equipped with an NVIDIA Tesla T4 GPU (16 GB GDDR6 memory, 2560 CUDA cores, and 320 Turing Tensor Cores optimized for deep learning tasks). While this demonstrates the feasibility of the approach, the extensive training time highlights potential scalability concerns.

Table 2 Hyperparameters used for GAN training

Hyperparameter	**Value**
Batch size	128
Optimizer	Adam
G learning rate l_{r_G}	$8 \cdot 10^{-5}$ reduced by a factor of 0.98 every 30 epochs
D learning rate l_{r_D}	$5 \cdot 10^{-4}$, constant
Final Epoch	2417
Loss	Binary Cross Entropy

We consider a two-layer AE, where the numbers of neurons in the layers are optimized through a grid search performed over a set of validation data randomly sampled from $\boldsymbol{X}$ and not used to train the AE, with a search space {3, 6, 9, 12, 15, 18, 21, 24, 27} for the first hidden layer and {0, 3, 6, 9, 12} for the second hidden layer. The second hidden layer is constrained to remain smaller than the number of input signals to ensure a bottleneck design, which is a well-established approach for learning compact representations. ReLU activations are employed in the first layer, and Linear in the second layer. We assess the quality of the normal conditions data generated by the GAN by qualitatively evaluating their diversity and coverage in the 2-dimensional spaces mapped by PCA and t-SNE (Figures 3a and 3b).

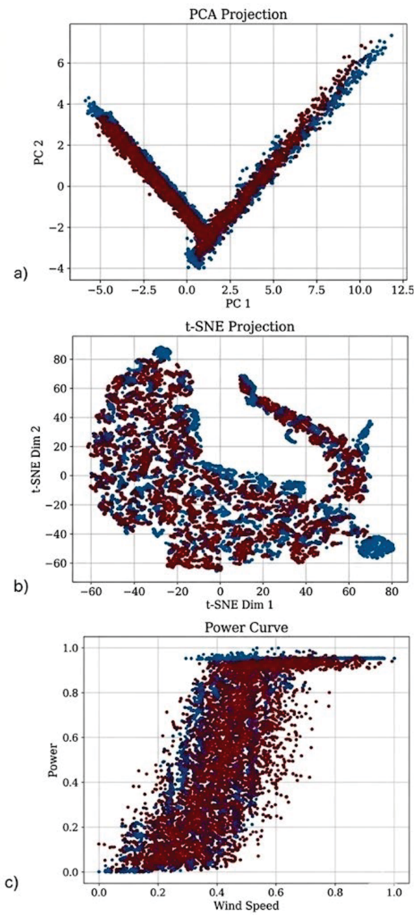

Figure 3 2D projections of original (blue) and generated (red) data: a) PCA, b) t-SNE, c) power curve.

The generated data (red dots) exhibit considerable overlap with the original data, demonstrating that the GAN effectively captures the training data distribution (coverage). Furthermore, rather than simply replicating exact copies of the training samples, the GAN synthesizes novel data points. This indicates that the model has learned a generalized representation of the wind turbine behavior, thereby ensuring data diversity. The anomaly detection performance is estimated on the whole test set and considering separately the low and high-wind speed regimes defined in Section 4. We compare the proposed method with an AE trained on the original (non-augmented) data. Figure 4 shows the obtained Receiver Operating Characteristic (ROC) curve, which is used for evaluating the performance of FD methods independently from the setting of the threshold th (Narkhede 2018).

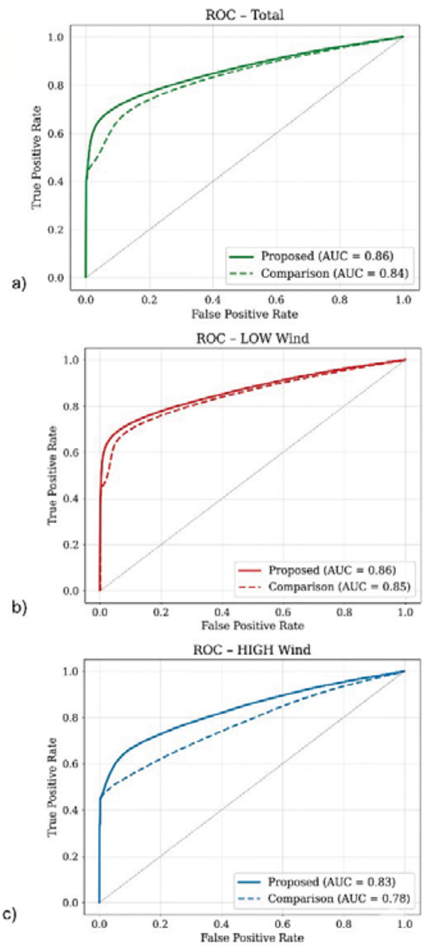

Figure 4 ROC curves on a) all wind regimes; b) low-wind regime; c) high-wind regime. The dashed line represents the comparison method, whereas the solid line represents the proposed one.

The x-axis represents the False Positive Rate (FPR), i.e. the proportion of negatives incorrectly identified as positives, whereas the y-axis shows the True Positive Rate (TPR), which is the Recall

(Forman 2005). A curve closer to the top-left corner indicates better performance.
Quantitively, we can consider as performance metric the Area Under the Curve (AUC) metric, with larger values indicating better performances. The proposed method overperforms the comparison method in terms of AUC (0.86 vs 0.84) when considering the entire test set. A similar improvement (AUC 0.86 vs 0.85) is obtained in the low-wind speed regime, whereas the improvement in the high-wind speed regime is more significant (AUC 0.83 vs 0.78)

5. Conclusion

In this work we develop a novel method for FD in PPSs, leveraging data augmentation to overcome the challenges of variable operating conditions in PPSs. The method is based on the generation by GAN of synthetic data of normal operating conditions spanning all wind conditions. FD is, then, performed using AE to reconstruct the signals as they are expected in normal operating conditions.
The results obtained in the case study considered demonstrate that the proposed data generation strategy significantly enhances the performance of an AE trained without using the GAN-generated data, especially in the high-wind regime.
Future work will focus on: *i*) extending the proposed framework to model temporal dynamics by considering specialized architectures, such as recurrent neural networks and LSTM cells. This is expected to allow for earlier fault detection; *ii*) investigating the use of Conditional GANs (C-GANs), where the generation process can be explicitly guided by manually setting operational variables (e.g., wind speed). This is expected to allow enriching the original set in underrepresented conditions such as high-winds, potentially improving both coverage and fault detection performanced.

References

Baschel, Stefan, Elena Koubli, Jyotirmoy Roy, and Ralph Gottschalg. 2018. “Impact of Component Reliability on Large Scale Photovoltaic Systems’ Performance.” *Energies* 11 (6): 1579. doi:10.3390/en11061579.

Beganovic, Nejra, and Dirk Söffker. 2016. “Structural Health Management Utilization for Lifetime Prognosis and Advanced Control Strategy Deployment of Wind Turbines: An Overview and Outlook Concerning Actual Methods, Tools, and Obtained Results.” *Renewable and Sustainable Energy Reviews*. Elsevier Ltd. doi:10.1016/j.rser.2016.05.083.

Chu, Mengyu, You Xie, Jonas Mayer, Laura Leal-Taixé, and Nils Thuerey. 2020. “Learning Temporal Coherence via Self-Supervision for GAN-Based Video Generation.” *ACM Transactions on Graphics* 39 (4). Association for Computing Machinery. doi:10.1145/3386569.3392457.

Fields, Tonya, George Hsieh, and Jules Chenou. 2019. “Mitigating Drift in Time Series Data with Noise Augmentation.” In *2019 International Conference on Computational Science and Computational Intelligence (CSCI)*, 227–30. doi:10.1109/CSCI49370.2019.00046.

Floreale, Giovanni, Piero Baraldi, Xuefei Lu, Paolo Rossetti, and Enrico Zio. 2024. “Sensitivity Analysis by Differential Importance Measure for Unsupervised Fault Diagnostics.” *Reliability Engineering and System Safety* 243 (March). Elsevier Ltd. doi:10.1016/j.ress.2023.109846.

Forman, George. 2005. “Counting Positives Accurately Despite Inaccurate Classification.” In *Machine Learning: ECML 2005*, edited by Rui and Brazdil Pavel B. and Jorge Alípio Mário and Torgo Luís Gama João and Camacho, 564–75. Berlin, Heidelberg: Springer Berlin Heidelberg.

Goodfellow, Ian J., Jean Pouget-Abadie, Mehdi Mirza, Bing Xu, David Warde-Farley, Sherjil Ozair, Aaron Courville, and Yoshua Bengio. 2014. “Generative Adversarial Networks,” June. http://arxiv.org/abs/1406.2661.

Guennec, Arthur Le, Simon Malinowski, and Romain Tavenard. 2016. *Data Augmentation for Time Series Classification Using Convolutional Neural Networks*. https://shs.hal.science/halshs-01357973v1.

Heusel, Martin, Hubert Ramsauer, Thomas Unterthiner, Bernhard Nessler, and Sepp Hochreiter. 2018. "GANs Trained by a Two Time-Scale Update Rule Converge to a Local Nash Equilibrium," January. http://arxiv.org/abs/1706.08500.

Liu, Zhaohui, Enhong Hu, and Hua Liu. 2025. "A Fault Diagnosis Framework for Pressurized Water Reactor Nuclear Power Plants Based on an Improved Deep Subdomain Adaptation Network." *Energies* 18 (9): 2334. doi:10.3390/en18092334.

Narkhede, Sarang. 2018. "Understanding Auc-Roc Curve." *Towards Data Science* 26 (1): 220–27.

Odgaard, Peter Fogh, Jakob Stoustrup, and Michel Kinnaert. 2013. "Fault-Tolerant Control of Wind Turbines: A Benchmark Model." *IEEE Transactions on Control Systems Technology* 21 (4): 1168–82. doi:10.1109/TCST.2013.2259235.

Olson, Matthew, Abraham Wyner, and Richard Berk. 2018. "Modern Neural Networks Generalize on Small Data Sets." In *Advances in Neural Information Processing Systems*, edited by S Bengio, H Wallach, H Larochelle, K Grauman, N Cesa-Bianchi, and R Garnett. Vol. 31. Curran Associates, Inc. https://proceedings.neurips.cc/paper_files/paper/2018/file/fface8385abbf94b4593a0ed53a0c70f-Paper.pdf.

Peng, Dandan, Wim Desmet, and Konstantinos Gryllias. 2025. "Reconstruction-Based Deep Unsupervised Adaptive Threshold Support Vector Data Description for Wind Turbine Anomaly Detection." *Reliability Engineering and System Safety* 260 (August). Elsevier Ltd. doi:10.1016/j.ress.2025.110995.

Perez, Luis, and Jason Wang. 2017. "The Effectiveness of Data Augmentation in Image Classification Using Learning." *CoRR* abs/1712.04621. http://arxiv.org/abs/1712.04621.

Pinciroli Luca, Baraldi Piero, and Zio Enrico. 2025. "Graph Neural Networks for Anomaly Detection in Wind Turbines." In *In Proceedings of the 35th European Safety and Reliability Conference (ESREL 2025) and the 33rd Society for Risk Analysis Europe Conference (SRA-E 2025), Stavanger, Norway, 2025*. Stavanger, Norway. doi: 10.3850/978-981-94-3281-3_ESREL-SRA-E2025-P9555-cd.

Radford, Alec, Luke Metz, and Soumith Chintala. 2016. "Unsupervised Representation Learning with Deep Convolutional Generative Adversarial Networks," January. http://arxiv.org/abs/1511.06434.

Ruff, Lukas, Robert Vandermeulen, Nico Görnitz, Lucas Deecke, Shoaib Siddiqui, Alexander Binder, Emmanuel Müller, and Marius Kloft. 2018. "Deep One-Class Classification." In .

Shorten, Connor, and Taghi Khoshgoftaar. 2019. "A Survey on Image Data Augmentation for Deep Learning." *Journal of Big Data* 6 (July). doi:10.1186/s40537-019-0197-0.

Wu, Shifa, Xiaolong Ma, Junfeng Liu, Jiashuang Wan, Pengfei Wang, and G. H. Su. 2023. "A Load Following Control Strategy for Chinese Modular High-Temperature Gas-Cooled Reactor HTR-PM." *Energy* 263 (January). Elsevier Ltd. doi:10.1016/j.energy.2022.125459.

Yan, Rundong, Sarah Dunnett, and Lisa Jackson. 2023. "Impact of Condition Monitoring on the Maintenance and Economic Viability of Offshore Wind Turbines." *Reliability Engineering and System Safety* 238 (October). Elsevier Ltd. doi:10.1016/j.ress.2023.109475.

Zhang, Zhibo, and Jin Jiang. 2024. "On Load-Following Operations of Small Modular Reactors." *Progress in Nuclear Energy* 173 (August): 105274. doi:10.1016/j.pnucene.2024.105274.