Proceedings of the 32nd European Safety and Reliability Conference (ESREL 2022)
Edited by Maria Chiara Leva, Edoardo Patelli, Luca Podofillini, and Simon Wilson
 Published by Research Publishing, Singapore.
doi: 10.3850/978-981-18-5183-4_S02-03-080-cd

Wavelet-based Noise Extraction for Anomaly Detection Applied to Safety-critical Electronics at CERN

Felix Waldhauser

HSE/RP, European Organization for Nuclear Research CERN, Switzerland.
Reliability Department, Institute of Machine Components, University of Stuttgart, Germany.
E-mail: felix.johannes.waldhauser@cern.ch

Hamza Boukabache

HSE/RP, European Organization for Nuclear Research CERN, Switzerland.
E-mail: hamza.boukabache@cern.ch

Daniel Perrin

HSE/RP, European Organization for Nuclear Research CERN, Switzerland. E-mail: daniel.perrin@cern.ch

Martin Dazer

Head of Reliability Department, Institute of Machine Components, University of Stuttgart, Germany.
E-mail: martin.dazer@ima.uni-stuttgart.de

Due to the possible damage caused by unforeseen failures of safety-critical systems, it is crucial to maintain these systems appropriately to ensure high reliability and availability. If numerous units of a system are installed in various areas and permanent access is not guaranteed, remote, data-driven condition monitoring methods can be used to schedule maintenance actions and to prevent unexpected failures. Thereby, failure precursors identified by unsupervised anomaly detection algorithms can be used to detect system malfunctions or to assess the systems condition. The anomaly detection process presented in this paper proposes a novel integrative combination of noise extraction using wavelet transforms and unsupervised algorithms to improve the detectability of a broad variety of anomalies for safety-critical electronics. Here, the performance of this modular process is demonstrated by identifying outlying data samples in datasets generated by the CERN Radiation Monitoring Electronics (CROME) system.

Keywords: Anomaly detection, condition monitoring, predictive maintenance, machine learning, wavelet transform, safety-critical electronics.

1. Introduction and Motivation

In order for systems to fulfill safety-critical functions, it is crucial to keep their failure rates at an extremely low level through appropriate maintenance measures. Especially for occupational safety systems, e.g., radiation protection, unforeseen system failures could endanger people and the environment. Planning maintenance actions based on failure predictions helps to reduce unexpected failures and to increase system availability and reliability. If operational data is available, data-driven methods can be used to monitor the system's condition and to detect malfunctions. This requires the identification of characteristics related to the system's degradation and the detection of failure precursors. Data-driven methods are suitable for remote monitoring of the system's condition and can be executed automatically based on scheduled routines. Further, no intervention in system operation is required and methods are scalable and thus applicable to numerous devices in separated areas, if connected to a data infrastructure.

In this paper, an anomaly detection process is presented enabling remote, data-driven condition monitoring based on detected failure precursors and system malfunctions. The proposed process combines noise extraction using wavelet transforms and unsupervised anomaly detection algorithms to increase the detectability of anomalous system behavior of safety-critical electronics. This allows to include noise-related features in the analysis and is particularly appropriate when neither typical failure precursors nor the type of anomalies of interest are known. In this case, the process can be used to detect rare and deviant data events based on a calculated anomaly score.

The system used as demonstrator in this paper is the CERN Radiation Monitoring Electronics (CROME). It is an active system for reliable ambient dose rate monitoring used by the CERN radiation protection service to protect people and the environment from unjustified exposure to ionizing radiation. It is based on a reconfigurable System-on-Chip (SoC) architecture with an integrated Field Programmable Gate Array (FPGA) to process measurement calculations in real-time (Boukabache et al. (2017)). Many devices are located in restricted, high-radiation areas to which engineers do not have permanent access. All CROME devices are connected to a cloud-based data infrastructure with long-term storage, permanent data access and live data injection from all devices.

In the context of the proposed data processing for anomaly detection, useful features to distinguish between anomalous and normal observations are developed to improve the performance and reliability of anomaly detection algorithms. Here, it is assumed that quantifying the noise extracted from the raw signals using wavelet transforms will make anomalous noise in the signals detectable to the algorithms. Due to missing guidelines in the literature on how to configure the wavelet transform, a signal classification process is developed to select the most appropriate configuration of the wavelet transform for a noise extraction use case.

2. Anomaly Detection for Predictive Maintenance

Generally, anomaly detection follows the principle of finding patterns or instances in datasets that are different from the rest of the dataset and do not conform to the expected normal behavior (Pecht and Kang (2018)). Hence, anomalies are data points which are sufficiently far away from normal patterns and can be caused by errors in the data acquisition or new, previously unknown, processes indicating changes in the system behavior or malfunctions. Data-driven predictive maintenance requires the monitoring of failure precursors, e.g., by anomaly detection algorithms, to schedule appropriate maintenance actions in order to achieve maximum system availability and reliability (Niu (2017)).

CROME datasets are discrete time-series generated with a device-specific sampling frequency. At the beginning of the anomaly detection process, datasets are split into samples of equal duration, more precisely one hour. This allows to account for different sampling rates and hence the calculation of comparable anomaly scores. Each sample is assigned a unique identifier containing all necessary information related to the corresponding device and time period. Subsequently, statistical features are calculated for each sample describing the data distribution of the raw signals of all measurement variables. The selection of these features is listed in Table 1 and is based on time-domain features that are commonly used for signal classification, e.g., in the field of brainwave signals (see, e.g., EL Menshawy et al. (2015) or Djordjevic et al. (2009)), or to detect failure precursors in data-driven Prognostics and Health Management (PHM) approaches (see Tsui et al. (2015)). Additionally, since anomalies related to noise in measurement signals are of special interest, certain statistical measures calculated for extracted noise signals are included in the feature list (see Table 1). Here, wavelet transform is used to extract noise from raw signals, similar to the noise extraction explained in section 3.

The first part of the anomaly detection process, data processing, handles the data import from the database and subdivides the data into hourly samples. Moreover, data cleaning and the calculation of statistical features, as listed in Table 1, is performed. As a result, two datasets are generated: one dataset contains the raw measurement data of 21 measurement variables whereas the other dataset contains all calculated features (see dataset schematic in Figure 1). Since the entire feature set

Table 1. Statistical features for anomaly detection calculated for raw signals and extracted noise signals. Std/Mean is the ratio of standard deviation and mean.

Feature	Raw Signal	Noise signal
Variance	X	X
Standard Deviation	X	
Mean	X	X
Std/Mean	X	
Median	X	X
Root Mean Square (RMS)	X	X
Number of Zero Crossings	X	X
Shannon Entropy	X	X
Line Length	X	X
Skewness	X	X
Kurtosis	X	X
Signal-to-Noise Ratio (SNR)	X	

is calculated for each measurement variable, the feature dataset consists of a total of 441 features.

Subsequent to data processing, two anomaly detection algorithms are employed to identify outlying data samples. Since neither the type of anomalies nor a definition of failure precursors is known, only unsupervised algorithms are applicable. This also means that no pre-existing knowledge can be used to link detected anomalies to failure precursors. Therefore, manual analysis of detected samples is required to gain information on anomalous system behavior and to identify possible failure precursors. Furthermore, CROME is a newly installed system with no known critical failures as of now, which is why anomalies are

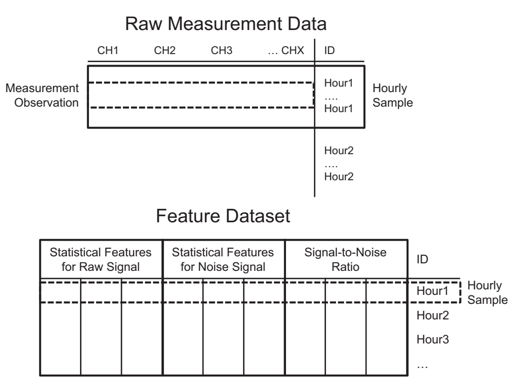

Fig. 1. Schematic of datasets resulting from the data processing routine.

expected to be rare in datasets.

Firstly, the Isolation Forest (iForest) algorithm, introduced by Liu et al. (2008), is applied only to the feature dataset (Figure 1). Since reducing the number of input dimensions can help to improve the performance of the iForest algorithm, the input dataset is reduced to a subspace of 50 features based on the kurtosis of each feature, as shown in Liu et al. (2008). In contrast to most of the common anomaly detection algorithms, this algorithm does not construct a profile of the healthy system state but identifies anomalies as rare and distinctly different. Anomalies are thus more likely to be isolated in binary isolation trees. It does therefore not require the definition of the system's healthy state and is applicable to high-dimensional datasets (Liu et al. (2008)), making this algorithm particularly suitable for this use case. The configuration of the iForest algorithm requires only the selection of the sub-sampling size n and the number of trees t. The configuration used in this study is shown in Table 2. Liu et al. (2008) and Ahmed et al. (2019) found that the number of trees can usually be set to $t = 100$. Furthermore, Liu et al. (2008) state that $n = 256$ is sufficient for a wide range of anomaly detection applications.

Table 2. Configuration of the Isolation Forest algorithm.

Parameter	Value
Number of trees t	100
Sub-sampling size n	256

Additionally, a Long Short-Term Memory Autoencoder (LSTM-AE) model is applied to the dataset containing raw measurements to identify time-dependent anomalies, e.g., unusual correlations or inconsistencies between variables. LSTM-AEs are commonly used for anomaly detection and to learn temporal-dependencies in time-series data (see, e.g., Said Elsayed et al. (2020) or Sharma et al. (2020)). Autoencoders represent the input data on a low-dimensional bottleneck layer and reconstruct it to the original dimensionality (Pawar and Attar (2020)). If used for anomaly detection, the reconstruction error between input and output serves as anomaly score. The configuration of the autoencoder model is mainly based on Sharma et al. (2020). The selected hyperparameters are shown in Table 3. The hyperbolic tangent function is selected as activation function as it is the preferable choice according to Pawar and Attar (2020). Two hidden layers for each encoder and decoder result in four hidden layers with 16, 4, 4 and 16 neurons. The input dataset with 21 measurement variables is therefore projected to a four-dimensional hidden representation.

Table 3. Configuration of the Long Short-Term Memory Autoencoder (LSTM-AE).

Parameter	Value
Hidden layers	2
Hidden layer size	16, 4
Activation function	tanh
Regularization	L2
Learning rate	0.0001
Loss function	Mean Squared Error
Optimizer	Adam
Batch size	256
Validation split	0.1

The entire anomaly detection process consists of modular elements that allow the routine to be adapted to case-specific requirements. All routines are executed on distributed cluster computing resources with GPU support and are implemented in the high-level programming language Python. The Python packages mainly used for anomaly detection are scikit-learn (Pedregosa et al. (2011)) for the iForest algorithm and tensorflow (Abadi et al. (2016)) for LSTM-AE models.

3. Configuration of Wavelet Transforms for Noise Extraction

Wavelet transform is used in the course of data processing for anomaly detection to extract noise signals from raw signals (as shown in Boukabache et al. (2013)), allowing the quantification and evaluation of possible noise in the signals by calculating statistical measures for the extracted signals. Signals generated by CROME devices are non-stationary which is why wavelet transform is selected as the most suitable time-frequency transform promising high resolution in both time and frequency domain, as found by Al-Qazzaz et al. (2015). More specifically, Multi-Resolution Analysis Wavelet Transform (MRA-WT), first introduced by Mallat (1989), which decomposes a signal into m decomposition levels by applying the wavelet transform in a hierarchical routine with approximation coefficients cA_i and detail coefficients cD_i as output, $i = 1, ..., m$, is used. Wavelet decomposition is a common method to separate noise from signals by modifying the output coefficients (Al-Qazzaz et al. (2015)). To obtain correct representations of the extracted noise signals, it is crucial to define the best combination of mother wavelet and level of decomposition for this use case (Al-Qazzaz et al. (2015)). However, there are currently no clear guidelines in the literature for selecting the most appropriate configuration for wavelet transforms. To overcome this issue, a novel signal classification process is introduced to make this selection based on the performance and the feature importance of a classification algorithm. Similarly to the method proposed by Al-Qazzaz et al. (2015), this process also uses wavelet transform for noise extraction, but evaluates the extracted noise signal instead of comparing the denoised signal to the raw signal. In general, the proposed approach is based on the following idea: take an input dataset with signal samples generated by CROME devices and modify some of these signals by adding a synthetically generated noise signal. Then, extract the noise signal for each sample by using the wavelet transform and calculate statistical measures describing these noise signals (similar to the measures shown in Table 1). Finally, train a Gradient Boosted Decision Trees (GBDT) classifier on this binary signal classification task using the calculated measures to predict whether the synthetic noise signal is present in a sample or not. Using the model performance to select the best configuration assumes that the noise signal extracted by the wavelet transform is more suitable for noise quantification if the model precision is higher and the feature importance more balanced. Figure 2 shows the process schematics.

In the following, the signal modification and the calculation of statistical measures quantifying the extracted noise signals (steps 4 and 5 in Figure 2) will be explained in more detail. As stated above, the proposed method relies on the detection of synthetic noise signals. During step 4, the signal modification, a synthetically generated noise signal is added to a proportion of the signal samples. More precisely, a sine wave signal with a certain frequency is generated and an additional random value drawn from a normal distribution with $\mu = 0$ is added to each value of the sine wave to create random deviations and to increase variability of the synthetic noise signal. It is important to say that this signal is added only to randomly selected parts of a sample to simulate spontaneously occurring noise. Moreover, both the frequency and amplitude of the noise signal are randomly drawn from corresponding lists of allowed values to generate a variety of synthetic noise signals in both training and testing dataset. This is expected to reduce overfitting of the classification algorithm and to ensure that the model quantifies the noise in a signal, rather than being trained on a certain unique signal characteristic arising from the synthetic noise signal.

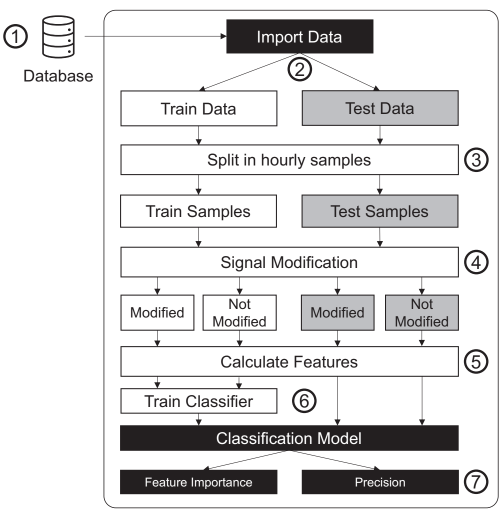

Fig. 2. Schematic of the proposed classification process for the selection of the best wavelet configuration for noise extraction.

The classification algorithm, a GBDT algorithm, is then trained on statistical measures calculated for the noise signals extracted by MRA-WT. Precisely, the raw signal is decomposed in sets of detail and approximation coefficients at each level of decomposition. The noise signal is then obtained by recomposing the detail coefficients only. Hence, no specific thresholding technique to modify the coefficients is applied since neither the frequency bands nor other properties of the raw signals are known. Instead, since noise is separable from the information of the signal with respect to frequency, approximation and detail coefficients are considered as outputs of lowpass and highpass filters respectively, as stated in Ben Abbes et al. (2018).

GBDT algorithms have the advantage of providing possibilities to analyze the feature importance that describes the impact of features on the model decision. Therefore, the best suitable configuration of the wavelet transform is selected considering the model performance, more precisely the model precision, and the balance of the feature importance. Figure 3 shows the resulting model precision for various configurations of the wavelet transform obtained by executing the proposed process (Figure 2) for measurement data of several CROME devices. Although the combination of db1 as mother wavelet and one level of decomposition (marked with solid line in Figure 3) leads to the highest model precision, db2 with 2 levels of decomposition (marked with dashed line in Figure 3) is selected as the best configuration as this comes with the best combination of high model precision and balanced feature importance (see Figure 4).

4. Demonstration of Anomaly Detection in Operational Data

To demonstrate the anomaly detection capabilities of the proposed process, operational data from CROME devices is processed and injected to the algorithms with the configurations shown in section 2.

The iForest algorithm does not require a training dataset and is directly applied to the feature dataset resulting in an anomaly score for each

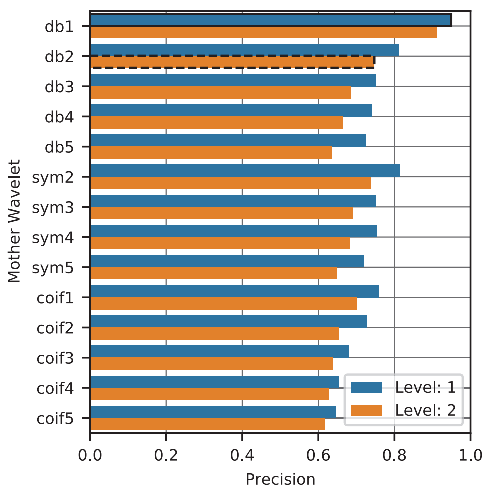

Fig. 3. Model precision for various configurations of the MRA-WT using measurement data from CROME devices. Two configurations marked in solid and dashed line.

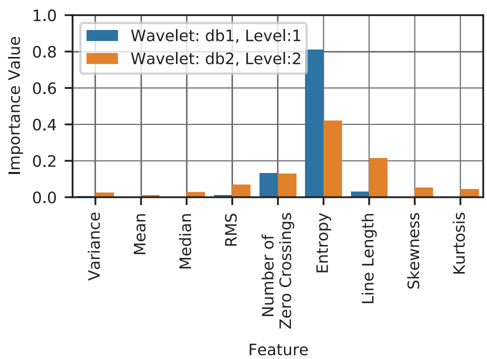

Fig. 4. Comparison of feature importances for two configurations of MRA-WT.

hourly sample according to the respective path length in the ensemble of isolation trees. Here, a shorter path length means that the sample was sorted out earlier and is thus more likely to be an outlying sample (Liu et al. (2008)).

The LSTM-AE, however, needs to be trained with data from healthy devices. During the model training, the networks weights and biases are consecutively optimized to minimize the reconstruction error for training samples. As there is currently no specific data from healthy devices, the autoencoder model is initially trained on pseudo-healthy data that is selected based on the results of the iForest algorithm. More precisely, the top five percent of samples with the lowest anomaly score calculated by the iForest algorithm are considered as healthy and used to train the autoencoder model. It should be mentioned that this is only the proposed solution for initial anomaly detection and is not recommended for operational applications. Analysis of the model training shows that the learning curve converges well after 15 epochs. Consequently, the number of training epochs will be adjusted to 10. This serves as a stopping criterion to prevent overfitting. The remaining 95 percent of the samples are used as testing set and an anomaly score is calculated for each sample based on the corresponding reconstruction error. In addition, due to the lack of healthy data, no explicit threshold can be set to determine which samples are truly anomalous. Hence, the samples are ranked by anomaly score in descending order and the top 20 samples are analyzed in detail.

It has been found that the iForest algorithm is able to detect samples with rare characteristics based on the calculated statistical features. Samples with high anomaly score are often related to implausible measurement values (e.g., negative dose rates), spikes in signals (see Sample 1 in Figure 5) or other unusual behaviors. The detected characteristics are in many cases related to extraordinary values of the statistical features, e.g., high kurtosis values for spikes in signals or high numbers of zero crossings for signals with negative dose rates. It should be noted that the calculated anomaly scores are generally relatively close to the value 0.5 which means, according to the definitions given in Liu et al. (2008), that no sample in the dataset is clearly considered as anomalous or normal by the iForest algorithm.

Compared to the results of the iForest algorithm, samples detected by the LSTM-AE model are rather related to correlations of variables or unusual deviations of signals. Thus, it can be assumed that the model is able to learn temporal dependencies in the input dataset. This leads to higher reconstruction errors if, e.g., the correlation between two variables is not as expected (see sample 3 in Figure 5: direct correlation between temperature and humidity, most samples show inverse correlation). Generally, the calculated reconstruction errors are relatively low and there is no sample with extraordinary high reconstruction error.

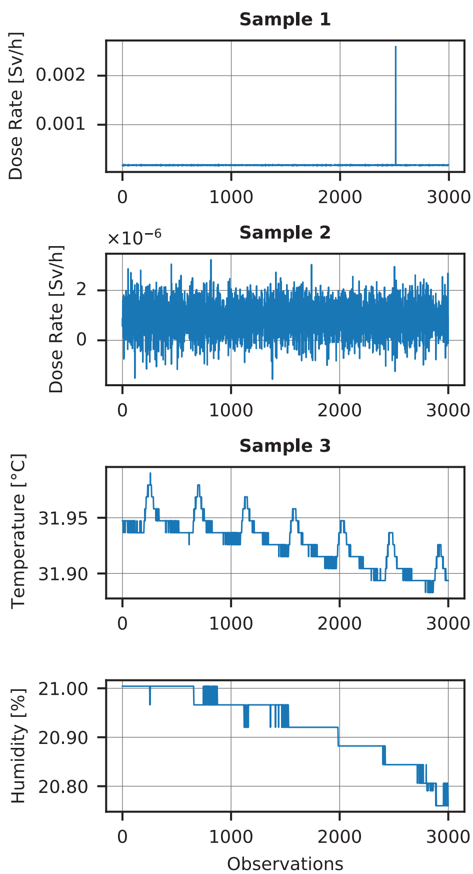

Fig. 5. Overview of detected, potentially anomalous, samples. Sample 1 (upper plot) shows an unusual spike in the dose rate measurement. Sample 2 shows negative dose rate measurements (physically incorrect). Temperature and Humidity have a direct correlation in data from Sample 3 (two lower plots, most samples show inverse correlation). All samples have a consistent duration of one hour.

Despite promising results and the demonstrated anomaly detection capabilities of the two algorithms, it needs to be verified by system experts whether the samples detected by iForest or LSTM-AE show actual anomalous system behavior or if they are rare and different but compliant with the expected behavior. One way to integrate expert knowledge in the process by analyzing samples manually could be a feedback loop, as shown in Figure 6. Samples classified as normal by a system expert can thus be used to retrain the model, while true anomalies can be used to further optimize the system or plan maintenance actions.

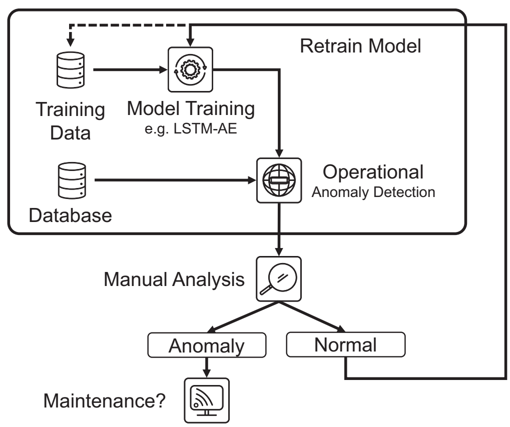

Fig. 6. Proposed feedback loop to integrate system knowledge in the anomaly detection process.

5. Discussion and Conclusion

As demonstrated for the use case of the safety-critical radiation monitoring electronics CROME, it has been shown that the proposed combination of unsupervised anomaly detection algorithms and wavelet-based noise extraction is able to identify data samples with a variety of characteristics that possibly represent anomalous behavior and/or failure precursors. The presented routine allows to include noise-related features in the analysis and is particularly appropriate when neither typical failure precursors nor the type of anomalies of interest are known. In this case, the process can be used to detect rare and deviant data events and is applicable to a variety of use cases due to its adaptability. In general, it is highly recommended to include knowledge from end-of-life tests or knowledge about the physical behavior of the system before using the results of the anomaly detection algorithms to plan maintenance actions.

Both data-driven maintenance strategies and wavelet transforms for noise extraction are often applied in a variety of research fields, although separately. Admittedly, anomaly detection methods are commonly used for condition monitoring but rarely for safety-critical electronics especially in combination with wavelet-based noise extraction. Moreover, features used in this paper describing statistical signal characteristics are widely used for signal characterization and specification, e.g., in the field of brain wave signals. Hence, the presented combination of these methods shows a novel approach to identify anomalies related to noise in signals. Additionally, the proposed routine for selecting the best configuration of the wavelet transform proved to be useful. However, it still needs to be verified whether this routine is beneficial for other noise extraction use cases.

References

Abadi, M., A. Agarwal, P. Barham, et al. (2016, 3). TensorFlow: Large-Scale Machine Learning on Heterogeneous Distributed Systems.

Ahmed, S., Y. Lee, S. H. Hyun, and I. Koo (2019, 10). Unsupervised Machine Learning-Based Detection of Covert Data Integrity Assault in Smart Grid Networks Utilizing Isolation Forest. *IEEE Transactions on Information Forensics and Security 14*(10), 2765–2777.

Al-Qazzaz, N. K., S. H. B. Mohd Ali, S. A. Ahmad, et al. (2015). Selection of mother wavelet functions for multi-channel EEG signal analysis during a working memory task. *Sensors 15*(11), 29015–29035.

Ben Abbes, A., O. Bounouh, I. R. Farah, et al. (2018). Comparative study of three satellite image time-series decomposition methods for vegetation change detection. *European Journal of Remote Sensing 51*(1), 607–615.

Boukabache, H., C. Escriba, S. Zedek, and J.-Y. Fourniols (2013, 2). Wavlet Decomposition based Diagnostic for Structural Health Monitoring on Metallic Aircrafts: Case of Crack Triangulation and Corrosion Detection. *International Journal of Prognostics and Health Management 4*(3).

Boukabache, H., M. Pangallo, G. Ducos, et al. (2017, 4). Toward a Novel Modular Architecture for CERN Radiation Monitoring. *Radiation Protection Dosimetry 173*(1-3), 240–244.

Djordjevic, V., N. Reljin, V. Gerla, et al. (2009). Feature extraction and classification of EEG sleep recordings in newborns. In *2009 9th International Conference on Information Technology and Applications in Biomedicine*, pp. 1–4.

EL Menshawy, M., A. Benharref, and M. Serhani (2015). An automatic mobile-health based approach for EEG epileptic seizures detection. *Expert Systems with Applications 42*(20), 7157–7174.

Liu, F. T., K. M. Ting, and Z. H. Zhou (2008). Isolation forest. *2008 Eighth IEEE International Conference on Data Mining*, 413–422.

Mallat, S. G. (1989). A theory for multiresolution signal decomposition: the wavelet representation. *IEEE Transactions on Pattern Analysis and Machine Intelligence 11*(7), 674–693.

Niu, G. (2017, 1). *Data-Driven Technology for Engineering Systems Health Management*. Singapore: Springer Singapore.

Pawar, K. and V. Z. Attar (2020). Assessment of Autoencoder Architectures for Data Representation. In W. Pedrycz and S.-M. Chen (Eds.), *Deep Learning: Concepts and Architectures*, pp. 101–132. Cham: Springer International Publishing.

Pecht, M. G. and M. Kang (Eds.) (2018, 9). *Prognostics and Health Management of Electronics*. Chichester, UK: John Wiley and Sons Ltd.

Pedregosa, F., G. Varoquaux, A. Gramfort, et al. (2011, 1). Scikit-learn: Machine learning in Python. *Journal of Machine Learning Research 12*, 2825–2830.

Said Elsayed, M., N.-A. Le-Khac, S. Dev, and A. D. Jurcut (2020, 11). Network Anomaly Detection Using LSTM Based Autoencoder. In *Proceedings of the 16th ACM Symposium on QoS and Security for Wireless and Mobile Networks*, New York, NY, USA, pp. 37–45. ACM.

Sharma, B., P. Pokharel, and B. Joshi (2020, 7). User Behavior Analytics for Anomaly Detection Using LSTM Autoencoder - Insider Threat Detection. In *Proceedings of the 11th International Conference on Advances in Information Technology*, New York, NY, USA, pp. 1–9. ACM.

Tsui, K.-L., N. Chen, Q. Zhou, et al. (2015, 5). Prognostics and Health Management: A Review on Data Driven Approaches. *Mathematical Problems in Engineering 2015*, 1–17.