# Diagnosis of Concurrent Sensor Faults in Structural Health Monitoring Systems

CARLOS CHILLÓN GECK, THAMER AL-ZURIQAT, KOSMAS DRAGOS and KAY SMARSLY

## ABSTRACT

Sensor faults in structural health monitoring (SHM) systems may occur due to aging, exposure to harsh weather conditions, manufacturing defects in hardware components, damage during installation or operation, and issues with data transmission. If undetected, sensors faults may result in inaccurate or incomplete sensor readings, which may significantly impact the accuracy, reliability, and performance of SHM systems. As a result, fault diagnosis in SHM systems may help improve the accuracy, reliability, and performance of SHM systems. However, most fault diagnosis approaches for SHM only consider single-fault occurrence, which may oversimplify actual fault occurrences in real-world SHM systems, where sensor faults may occur concurrently in multiple sensors. To extend fault diagnosis in SHM towards concurrent sensor faults in multiple sensors, this paper presents an adaptive fault diagnosis approach based on analytical redundancy. The approach encompasses four steps, (i) initialization (ii) fault detection, (iii) fault isolation and (iv) fault accommodation, using correlated data from multiple sensors of an SHM system. The proposed fault diagnosis approach is validated using data recorded using a real-world SHM system. The results show the high accuracy, reliability, and performance of the proposed approach in detecting concurrent sensor faults in real-world SHM systems.

## INTRODUCTION

Structural health monitoring (SHM) aims to assess the condition and to estimate the lifetime of civil infrastructure through non-destructive evaluation based on data recorded by sensors ("sensor data") [1]. The goal of SHM, representing a testing strategy that complements traditional nondestructive testing, is to reduce maintenance expenses through providing insights into the structural condition of civil infrastructure [2]. Sensors in SHM systems may encounter faults that may affect the accuracy, reliability, and performance of the monitoring systems. Faults may result from various causes, such as hardware or software malfunctions, power outages, environmental factors, or signal interferences [3]. The most common sensor faults are bias, complete failure, complete failure with noise, gain, drift, and outliers [4].

Carlos Chillón Geck, Thamer Al-Zuriqat, Kosmas Dragos, Kay Smarsly
Hamburg University of Technology, Institute of Digital and Autonomous Construction, Blohmstraße 15, 21079 Hamburg, Germany

Fault diagnosis (FD) approaches have been developed to detect, isolate, identify, and accommodate sensor faults in monitoring systems, including SHM systems [5]. FD for SHM has been based on either physical or analytical redundancy. Physical redundancy involves installing multiple sensors on infrastructure and using majority-voting logic to determine whether a sensor is faulty or not [6]. However, the high cost, power consumption, and maintenance requirements associated with physical redundancy approaches have fueled analytical redundancy approaches [7]. Analytical redundancy relies on mathematical models to describe a system by making use of redundant information in sensor data. In analytical redundancy approaches, fault detection relies on residuals between sensor data and corresponding "virtual outputs", estimated by mathematical models [3]. The residuals are evaluated using threshold logic or hypothesis testing for fault detection [8].

Mathematical models for analytical redundancy in FD are frequently based on artificial intelligence. Examples include multilayer neural networks that have been used for fault detection in mechanical components of wind turbines [9] as well as artificial neural network (ANN) models embedded in wireless sensor nodes for decentralized detection and isolation of sensor faults both in the time domain and in the frequency domain [7,10]. ANN models have also been combined with convolutional neural networks, performing fault identification, thus achieving full FD [11]. However, the analytical redundancy approaches mentioned above for SHM are limited to diagnosing sensor faults that occur in individual sensors at different times [12]. This limitation restricts the applicability in real-world SHM systems where simultaneous sensor faults in multiple sensors may occur ("concurrent sensor faults").

This paper presents an adaptive FD approach based on analytical redundancy (AFDAR). The AFDAR approach builds upon previous work, in which artificial neural networks and signal processing have been proposed for FD in SHM systems [7, 11, 13]. The AFDAR approach achieves FD through a combination of ANN models and moving averages of individual sensor data to detect, isolate, and accommodate sensor faults in multiple sensors. Fault identification, being independent of single-fault or multiple-fault occurrence and having been effectively addressed in previous work [11], is excluded from this study. The rest of the paper is organized as follows: First, the AFDAR approach is illuminated. Then, the validation of the AFDAR approach is presented, and the results of the validation test are analyzed. The paper concludes with a summary and an outlook on future work.

## DIAGNOSIS OF CONCURRENT SENSOR FAULTS IN SHM SYSTEMS

This section introduces the design and implementation of the AFDAR approach, which comprises four steps: (i) initialization, (ii) fault detection, (iii) fault isolation, and (iv) fault accommodation. Figure 1 illustrates a flowchart of the workflow of the AFDAR approach. In what follows, the model used for analytical redundancy is briefly explained, and the four steps of the AFDAR approach are discussed.

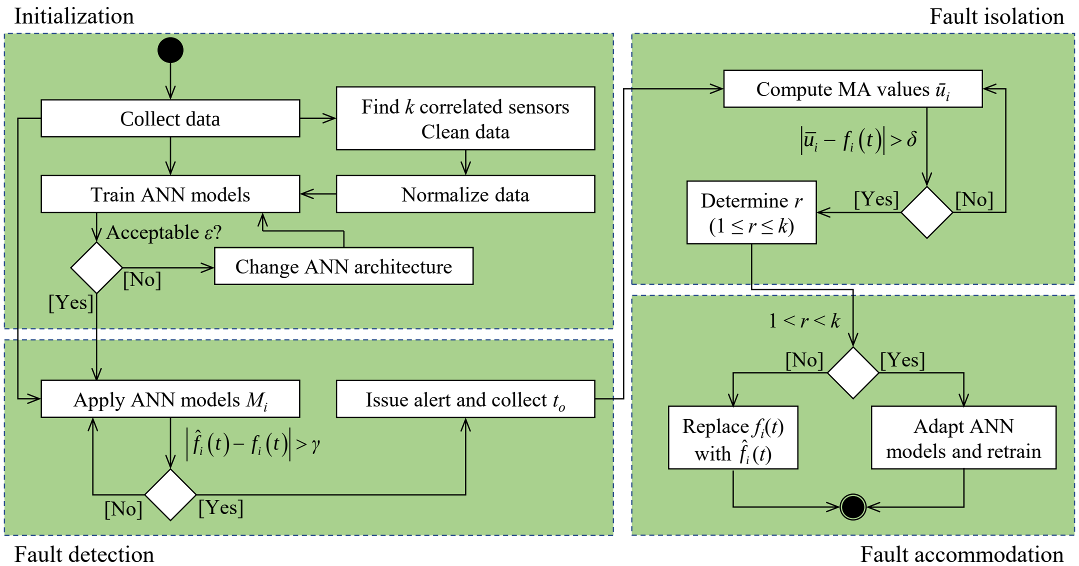


Figure 1. Flowchart of the AFDAR approach.

Artificial neural networks represent a class of artificial intelligence models, used for mapping input data with output data, both known a priori and collectively referred to as "labeled data". The layout of a typical ANN includes an input layer, one or several hidden layers, and an output layer, as shown in Figure 2. Each layer comprises nodes, referred to as "neurons", which accept input data from previous neurons via connections ("synapses") and produce outputs, termed "activations", using the input data and an activation function. The mapping result of the ANN ("prediction") is provided in the output layer. Training the ANN models involves adjusting the weights of the synapses until the prediction error (difference between prediction and output data) drops below a predefined threshold, the weights being adjusted according to the "learning rate".

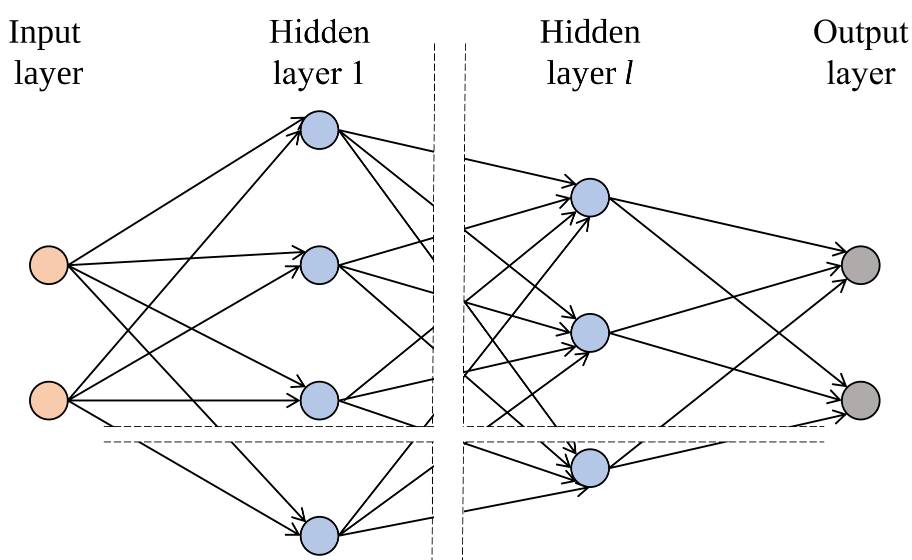


Figure 2. Layout of a typical artificial neural network.

1. **Initialization:** The initialization step starts with exploring the correlations in the sensor data to identify a set of $k$ "correlated sensors". Then, sensor data, recorded by the correlated sensors, $f_{1\rightarrow k}(t)$ is "cleaned", i.e. if sensor data from an individual sensor is missing at a specific time window, the same time window is neglected in all correlated sensors. Next, sensor data from the correlated sensors data is normalized to avoid extreme values in activations that would hinder the training process, using a minimum-maximum normalization. Upon being normalized, the sensor data is used to train ANN models. One ANN model $M_i$ is designed and trained for each correlated sensor $i$ ($i$ = 1…$k$). For training the ANN model $M_i$, sensor data from the correlated sensors (excluding sensor $i$) is used as input data, while sensor data $f_i(t)$ from sensor $i$ is used as output data. Upon completing training, the model $M_i$ is capable of yielding predictions $\hat{f}_i(t)$ for sensor $i$. The training threshold $\gamma$ is established by the root mean squared error (RMSE) value $\varepsilon$ between the predictions $\hat{f}_i(t)$ and the sensor data $f_i(t)$, as described in Equation 1.

$$\varepsilon = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[\hat{f}_i(t) - f_i(t)\right]^2} \tag{1}$$

2. **Fault detection:** Upon completing the initialization, newly recorded sensor data is input into all ANN models. If faults occur in $r$ sensors (where $1 < r < k$), the residuals between the actual sensor data and predictions of models $M_n$ (where $n = 2...r$) are expected to exceed $\gamma$, which triggers a fault detection. The time at which $\gamma$ is violated is noted as the "fault time stamp" $t_o$. Since the faulty sensor data $f_n(t)$ is also utilized as input to the $s$ models $M_v$ (where $v = 1...s$ and $r + s = k$) of the unaffected sensors, the virtual outputs of the $M_v$ models are contaminated, yielding $\varepsilon$ values that exceed $\gamma$ as well. Therefore, fault isolation requires further analysis of the sensor data on an individual sensor level, using the fault time stamp $t_o$, which represents the knowledge transferred to the next step.
3. **Fault isolation:** Following fault detection, a time window $N$ is determined based on $t_o$. Then, the moving average (MA) values $\bar{u}_i$ of $p$ data points $u_{ij}$ ($j$ = 1…$p$, $p < N$) are computed for sensor $i$. The window must have adequate length $N$ ($t_o - N/2$, $t_o + N/2$) around $t_o$ to ensure reliable tracking of the MA. Gradual or abrupt changes in the $\bar{u}_i$ values indicate sensor faults. As a result, discrepancies between MA values $\bar{u}_i$ and a fault isolation threshold $\delta$ from time $t_o$ forward indicate faulty sensor data of sensor $i$. Once fault isolation is completed and the faulty sensors are identified, the ANN models adapt to the new conditions of the SHM system, as described in the next step.
4. **Fault accommodation:** Following fault isolation, the ANN models adapt to the new conditions of the SHM system as follows:
   a. Sensor data of the $r$ correlated sensors that have been diagnosed as faulty are removed from the ANN input layers of all models. As a result, the architectures of the ANN models are modified and retrained to yield predictions for the faulty sensors.
   b. Retraining uses sensor data prior to $t_o$. Thereupon, the predictions of the $M_n$ ($n$ = 2...$r$) models are used to replace the faulty sensor data.

The threshold values $\gamma$ and $\delta$ depend on the type of data recorded by the SHM system and are thus application-specific. The validation of the AFDAR approach using sensor data from a real-world SHM system is presented in the next section.

## VALIDATION OF THE AFDAR APPROACH

In this section, the validation test of the AFDAR approach is presented along with the results of the test. The validation test is conducted using sensor data recorded by a real-world SHM system. The SHM system is installed on a double-track composite railway bridge located in Germany. The bridge consists of two parallel steel trusses supporting a 45 cm thick reinforced concrete (RC) slab. The bridge consists of 15 spans, each 58 m long – except the edge spans, which are 57 m long – and has a total length of 868 m. The deck width is 14.1 m, and the distance between the centroids of the steel truss girders is 6.2 m. The SHM system comprises temperature sensors, embedded in the RC slab, of type Pt100, measuring within a range of –35 °C to 105 °C with a sensitivity of ± 0.5 °C. Sensor data from 10 temperature sensors (S1...S10) is used for the validation test, the positions of the sensors being shown in Figure 3.

The temperature measurements used for validation have been recorded over almost five years with a sampling rate of 1.7 mHz, i.e. one temperature measurement has been recorded every 10 minutes, with a total of 256,000 measurements recorded by each sensor. In the initialization step, correlations between the temperature measurements, recorded over a period of two years, are investigated via correlation analysis. A strong positive correlation is found, based on the Pearson correlation coefficient, among all 10 temperature sensors ($k$ = 10). Next, the temperature measurements from the correlated sensors in the SHM system are cleaned and normalized. The number of ANN models is set equal to the number of correlated sensors ($k$ = 10). Each model predicts the virtual outputs of one sensor, using temperature measurements from the other nine correlated sensors in the SHM system as input data. As a result, each ANN model has nine input neurons and one output neuron; the number of hidden layers and neurons per hidden layer is determined with different ANN architectures. Before training the ANN models for FD, the temperature measurements are split into training (80 %) and testing sets (20 %). The architecture determined for all ANN models is 9-32-64-256-256-1, based on the lowest RMSE values $\varepsilon$ (0.09-0.15), with a total training time of approximately 680 s for each ANN model. The fault detection threshold is set to $\gamma$ = 0.15 equal to the highest RMSE value from training, erring on the side of safety regarding fault detection. The remaining steps of the AFDAR approach, i.e. fault detection, fault isolation and fault accommodation, are executed separately.

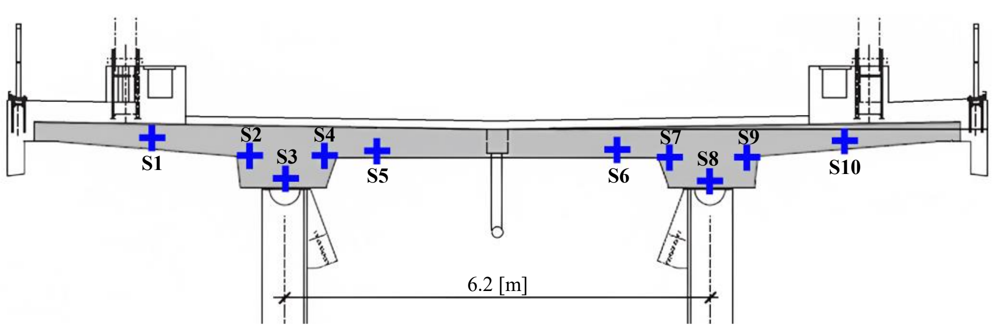


Figure 3. Cross section of the bridge with embedded temperature sensors.

TABLE I. SENSORS AND NUMBER OF FAULTS DETECTED PER SENSOR

| **S1** | **S2** | **S3** | **S4** | **S5** | **S6** | **S7** | **S8** | **S9** | **S10** |
|---|---|---|---|---|---|---|---|---|---|
| 0 | 18 | 274 | 0 | 1 | 0 | 0 | 4,339 | 0 | 0 |

Representing temperature measurements "newly recorded" by the SHM system, sensor data corresponding to a period of one year, i.e. 52,560 temperature measurements per sensor, are used as input data to the 10 ANN models. Upon applying the AFDAR approach, 4,632 faults are diagnosed in the sensor data, as summarized in Table 1. Figure 4 exemplarily shows the data recorded by the correlated sensors in the SHM system, focusing on faults detected in sensor S3 and sensor S8. As can be seen from Figure 4, concurrent faults of sensors S3 and S8 are detected by the AFDAR approach because the recorded temperature measurements exceed the fault detection threshold. It should be noted that a total of 274 simultaneous faults occurred in sensors S3 and S8 between December 19 and December 20; however, the focus is only on the simultaneous faults that occurred in the aforementioned period.

To illustrate the results of the approach for concurrent real-world sensor faults, the sensors S3 and S8 are analyzed in more detail between December 19 and December 20. Fault detection is performed when the residuals between the temperature measurements of sensors S3 and S8 and the virtual outputs of models $M_3$ and $M_8$ exceed the fault detection threshold $\gamma = 0.15$. Fault isolation is then performed using the concurrent fault time stamp $t_o$ in both sensors S3 and S8, determined at $t_o$ = 01:20 on December 19. The fault isolation threshold is set to the accuracy of the temperature sensors $\delta = \pm\ 0.5$ °C. Because the residuals between the MA values of sensors S3 and S8 and the corresponding temperature measurements at $t_o$ exceed the fault isolation threshold $\delta$, the faulty sensors are isolated. Finally, fault accommodation is performed: Since both sensors S3 and S8 are faulty, the models $M_3$ and $M_8$ are adapted by modifying the architecture of the ANN models through moving sensors S3 and S8 from the input layer to the output layer; the data recorded before $t_o$ is used to train the "adapted" ANN model $M_{3,8}$. Figure 5 shows the architecture of the adapted ANN model $M_{3,8}$, which predicts the virtual outputs $\hat{f}_3(t)$ and $\hat{f}_8(t)$ for both sensors S3 and S8.

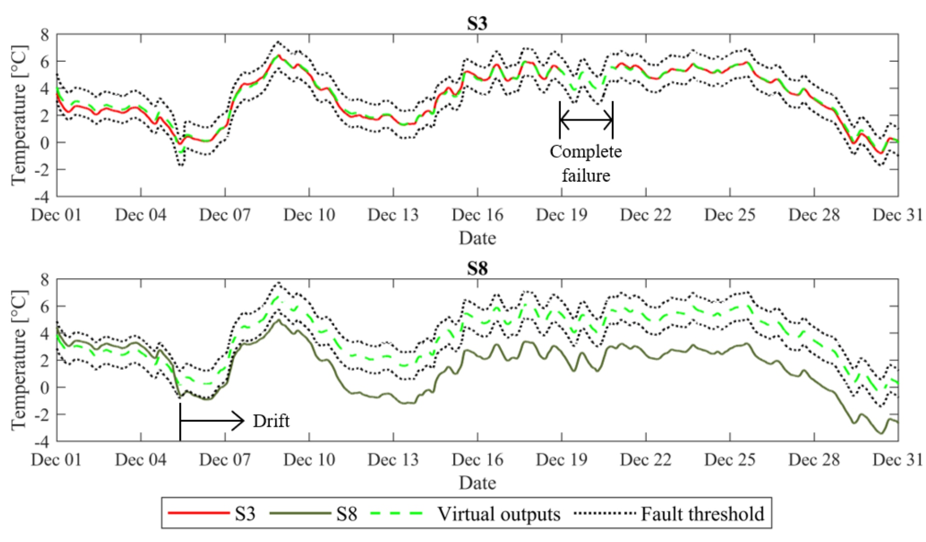


Figure 4: Comparison of temperature measurements and virtual outputs with the fault threshold.

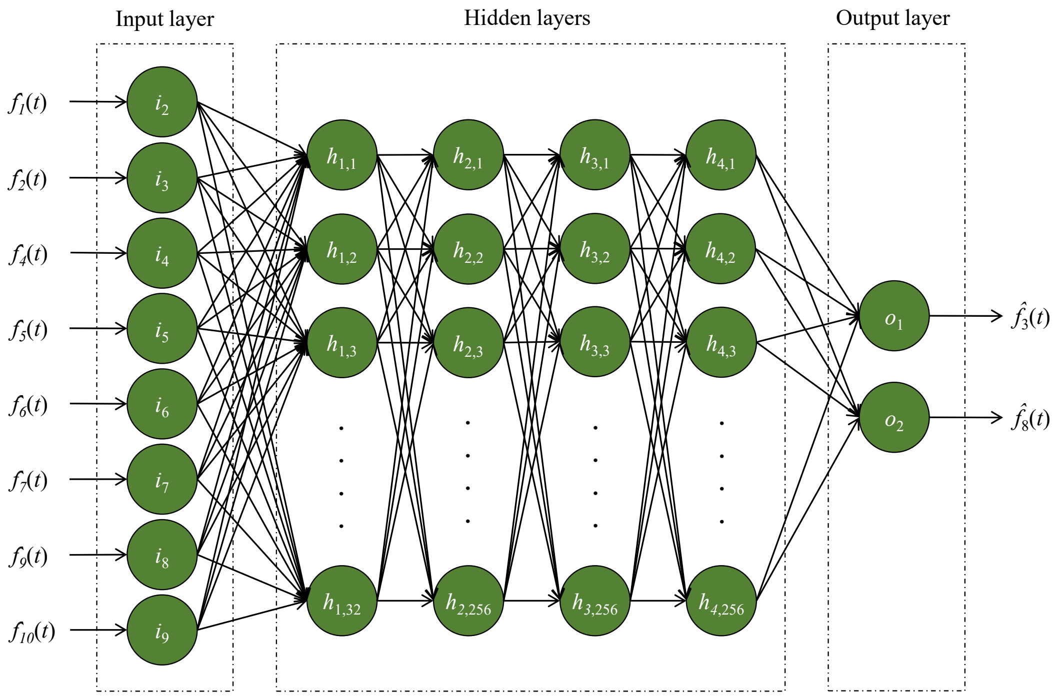


Figure 5: Adapted model $M_{3,8}$ for sensors S3 and S8.

## SUMMARY AND CONCLUSIONS

This paper has presented an adaptive FD approach based on analytical redundancy (AFDAR) that can reliably diagnose simultaneous sensor faults in multiple sensors of SHM systems. The approach combines ANN models with moving averages of individual sensor data to detect, isolate, and compensate for simultaneous sensor faults. The ANN models are used to predict virtual outputs for each sensor of an SHM system. To validate the proposed approach, data collected from a real-world SHM system have been used, showcasing the accuracy, reliability, and performance of FD in detecting, isolating, and accommodating sensor faults. Furthermore, the AFDAR approach has proven capable of adapting to the state of the SHM system regardless of the number of faulty sensors. In summary, the AFDAR approach can be used to ensure the accuracy of sensor data and thus to maintain the reliability and performance of SHM systems installed on civil infrastructure. Future work may focus on extending the AFDAR approach to distinguish between sensor faults and structural damage, as well as on improving the computational efficiency of the approach.

## ACKNOWLEDGMENTS

The authors of this research endeavor gratefully acknowledge the support offered by the German Research Foundation (DFG) under grant SM 281/20-1 as well as by the German Federal Ministry for Digital and Transport (BMDV) within the mFUND program under grant 19FS2013B. Any opinions, findings, conclusions, or recommendations expressed in this paper are those of the authors and do not necessarily reflect the views of DFG or BMDV.

## REFERENCES


1. Law, K.H., Smarsly, K., & Wang, Y., 2014. Sensor data management technologies for infrastructure asset management. In: Wang, M.L., Lynch, J.P., Sohn, H. (eds.). Sensor Technologies for Civil Infrastructures. Pp. 3-32. Sawston, UK: Woodhead Publishing.
2. Liu, Y. & Nayak, S., 2012. Structural Health Monitoring: State of the Art and Perspectives. The Journal of The Minerals, Metals & Materials Society, 64(7), pp. 789-792.
3. Li, L., Liu, G., Zhang, L., & Li, Q., 2019. Sensor fault detection with generalized likelihood ratio and correlation coefficient for bridge SHM. Journal of Sound and Vibration, 442, pp. 445-458.
4. Zhang, Z., Mehmood, A., Shu, L., Huo, Z., Zhang, Y., & Mukherjee, M., 2018. A survey on fault diagnosis in wireless sensor networks. IEEE Access, 6, pp. 11349-11364.
5. Patton, R.J., 1990. Fault detection and diagnosis in aerospace systems using analytical redundancy. In Proceedings of the IEE Colloquium on Condition Monitoring and Fault Tolerance. London, UK, 11/06/1990.
6. Frank, P.M., 1990. Fault diagnosis in dynamic systems using analytical and knowledge-based redundancy: A survey and some new results. Automatica (Oxf.), 26, pp. 459-474.
7. Smarsly, K. & Law, K.H., 2014. Decentralized fault detection and isolation in wireless structural health monitoring systems using analytical redundancy. Advances in Engineering Software, 73, pp. 1-10.
8. Isermann, R. & Ballé, P., 1997. Trends in the application of model-based fault detection and diagnosis of technical processes. Control Engineering Practice, 5(5), pp. 709-719.
9. Zaher, A., McArthur, S.D.J., Infield, D.G., & Patel, Y., 2009. Online wind turbine fault detection through automated SCADA data analysis. Wind Energy, 12(6), pp. 574-593.
10. Dragos, K. & Smarsly K., 2016. Distributed adaptive diagnosis of sensor faults using structural response data. Smart Materials and Structures, 25(10), 105019.
11. Fritz, H., Peralta Abadía, J. J., Legatiuk, D., Steiner, M., Dragos, K., & Smarsly, K., 2022. Fault diagnosis in structural health monitoring systems using signal processing and machine learning techniques. In: Cury, A., Ribeiro, D., Ubertini, F. & Todd, M.D. (eds.) Structural Health Monitoring Based on Data Science Techniques. Pp. 143-164. Cham, Switzerland: Springer International Publishing.
12. Samy, I., Postlethwaite, I., & Gu, D.-W., 2011. Survey and application of sensor fault detection and isolation schemes. Control Engineering Practice, 19(7), pp. 658-674.
13. Al-Zuriqat, T., Chillón Geck, C., Dragos, K., & Smarsly, K., 2023. Adaptive fault diagnosis for simultaneous sensor faults in structural health monitoring systems. Infrastructures, 8(3), 39.