Hybrid method for restoring missing sensor data with adaptive control based on neuro-fuzzy networks*

Anatoliy Sachenko[1,†], Victoria Vysotska[2,†], Serhii Vladov[3,*,†], Viktor Vasylenko[3,†] and Maclej Dobrowolski[4,†]

[1] West Ukrainian National University, Lvivska Street 11 46009 Ternopil, Ukraine
[2] Lviv Polytechnic National University, Stepan Bandera Street 12 79013 Lviv, Ukraine
[3] Kharkiv National University of Internal Affairs, L. Landau Avenue 27 61080 Kharkiv, Ukraine
[4] Kazimierz Pulaski University of Radom, Department of Informatics, Jacek Malczewski Street 29 26-600 Radom, Poland

Abstract

This research presents a hybrid method for restoring missing sensor data using adaptive control, aimed at enhancing data recovery accuracy in complex situations. The proposed algorithm involves creating a training dataset, a control element, and training hybrid models. Key steps include identifying data subsets, randomly removing elements to simulate gaps, estimating their values, clustering for accuracy, and developing neuro-fuzzy models tailored to specific data groups. Implemented as a Takagi-Sugeno neuro-fuzzy network, the method effectively combines fuzzy rules with neural computations, allowing it to handle variability and noise in real measurements with high adaptability. Computational experiments focused on restoring gas temperature data from the TV3-117 turboshaft engine demonstrate high accuracy, with deviations ranging between 0.002 and 0.007, indicating robust performance even under sensor failures. The model achieves over 99% accuracy, a precision of 0.983, and an F1-score of 0.991, significantly outperforming traditional methods such as two-layer feedforward networks and ANFIS networks. This research offers a reliable solution for sensor data recovery and advances adaptive control systems, improving the reliability and performance of aviation technologies and other dynamic applications.

Keywords

restoring, missing sensor data, hybrid method, Takagi-Sugeno neuro-fuzzy network, helicopter turboshaft engine, fuzzy rules

1. Introduction

Restoring missing sensor data is one of the key tasks in technical monitoring and control systems for complex objects such as aviation engines [1], robotic systems [2], and production lines [3]. The completeness and reliability of information received from sensors are crucial for these systems maintaining safe and efficient operation. However, in practice, sensors may fail or transmit distorted data due to technical malfunctions, external influences, or communication failures [4]. These factors pose the risk of inaccurate assessments regarding the objects condition, potentially leading to undesirable outcomes such as accidents or inefficient management.

The research relevance methods for restoring sensor data arises from the need to enhance the information systems reliability and precision. Modern approaches, including neural networks and statistical models, enable not only the missing values restoring but also the systematic and random measurement errors correction. This is particularly important in environments with stringent safety

AdvAIT-2024: 1st International Workshop on Advanced Applied Information Technologies, December 5, 2024, Khmelnytskyi, Ukraine - Zilina, Slovakia

* Corresponding author.
† These authors contributed equally.
as@wunu.edu.ua (A. Sachenko); victoria.a.vysotska@lpnu.ua (V. Vysotska); serhii.vladov@univd.edu.ua (S. Vladov); klk.vonrgp@gmail.com (V. Vasylenko); m.dobrowolski@uthrad.pl (M. Dobrowolski)
0000-0002-0907-3682 (A. Sachenko); 0000-0001-6417-3689 (V. Vysotska); 0000-0001-8009-5254 (S. Vladov); 0000-0002-9313-861X (V. Vasylenko); 0000-0003-0296-9651 (M. Dobrowolski)

and cost-efficiency requirements, where timely detection and compensation for data anomalies can extend equipment lifespan and reduce maintenance costs.

2. Related works

Existing research on restoring missing sensor data covers a wide range of approaches, from traditional interpolation methods to advanced machine learning algorithms and neural networks. Classical methods like linear interpolation [5] or polynomial approximation [6] are effective for small gaps in data with predictable patterns but fall short when dealing with nonlinear dependencies, noise, or dynamic systems, making them less suitable for complex applications.

With the artificial intelligence and machine learning advancement, the neural networks use has gained significant attention in data restoring. Recurrent neural networks (RNN) [7], particularly long short-term memory (LSTM) models [8,9], are frequently used to account for temporal dependencies and can recover information despite extended gaps. Other approaches [10,11] involve autoencoders, which train data representations and restore missing values. These methods show strong performance with nonlinear time series and hold promise for real-time applications.

Neuro-fuzzy systems offer another effective approach, combining neural networks with fuzzy logic to handle uncertainty and noise [12]. Models like the Sugeno approach apply flexible rules to process fuzzy data and adapt to changing data characteristics, making them especially relevant for systems with high uncertainty or significant random errors [13].

However, key challenges remain unresolved. One issue is the need for adaptive control in data restore when system characteristics change [14]. Many models are fine-tuned to specific system parameters, but their performance degrades with changing conditions. Addressing this requires models capable of adjusting their parameters dynamically. Neuro-fuzzy systems with adaptive control elements [15] offer a promising solution for real-time adaptation.

Another challenge is managing multidimensional dependencies between sensors, especially when parameters are highly interdependent. Current studies [16] often focus on single-sensor data restore, neglecting correlations across channels. Developing multichannel models that consider sensor interactions and adapt to their evolving characteristics is critical for improving restore performance.

Thus, adaptive control and neuro-fuzzy systems represent a promising direction for addressing these challenges, enhancing data restore quality while adapting to changing operational conditions.

3. Proposed method

This research foundation is a general approach to developing a hybrid model, proposed in [29]. The original dataset portion, containing only complete data, is used for model construction, and certain elements are randomly removed from it. For each artificially excluded element, values are estimated using various methods. As a result, a restored set values are generated for each element, which is then used as the input vector to build the adaptation model. The known value of the removed element serves as the output:

$$\mathbf{u} = \{u_1, u_2, u_3, u_4, v\}, \tag{1}$$

where u_i is the value obtained using one method, and v is the element original value. The resulting training dataset (1) is applied for developing and tuning the adaptation model.

The missing element value estimate in the original dataset, each of the four methods is first applied. The results are input into the model, producing a final estimate for the missing element value. A critical aspect in model development is the architecture choice. A linear model in the form $v = w_1 \cdot u_1 + w_2 \cdot u_2 + w_3 \cdot u_3 + w_4 \cdot u_4 + w_0$ [17] was examined, along with various nonlinear models based on artificial neural networks [18], which implement a function in the form $v = f(u_1, u_2, u_3, u_4)$. None of the models analyzed demonstrated an improvement in accuracy compared to the individual methods included in the model.

To create a more sophisticated model, a cluster analysis is performed on the training datasets (1) for all employed data arrays. As an analytical tool, presented in [17], it is recommended to utilize the Fuzzy C-means method with the Xie-Beni criterion [19] for the clusters number assessing. For the clustering procedure, it is proposed to use modified samples that include not the results generated by the methods, but the absolute errors related to their performance:

$$\mathbf{u} = \{|u_1 - v|, |u_2 - v|, |u_3 - v|, |u_4 - v|\}. \tag{2}$$

Conducting cluster analysis on the training datasets allows the several main groups identification within the original data arrays, where a similar pattern is observed regarding the original methods accuracy. Analysis [19,20] demonstrated that constructing a separate hybrid model for each of these groups based on a feedforward neural network significantly improved the hybrid method accuracy compared to the original approaches. The feedforward neural network key limitation lies in its inability to effectively handle uncertainty and variability in data, particularly under conditions characterized by complex nonlinear dependencies. This necessity highlights the neuro-fuzzy networks application, which, unlike standard neural networks, can account for the information fuzziness and adapt to changing conditions, ensuring higher accuracy in tasks such as clustering and complex data analysis.

This research proposes a method for restoring missing sensor data, featuring control based on neuro-fuzzy networks, with the developed scheme illustrated in Figure 1.

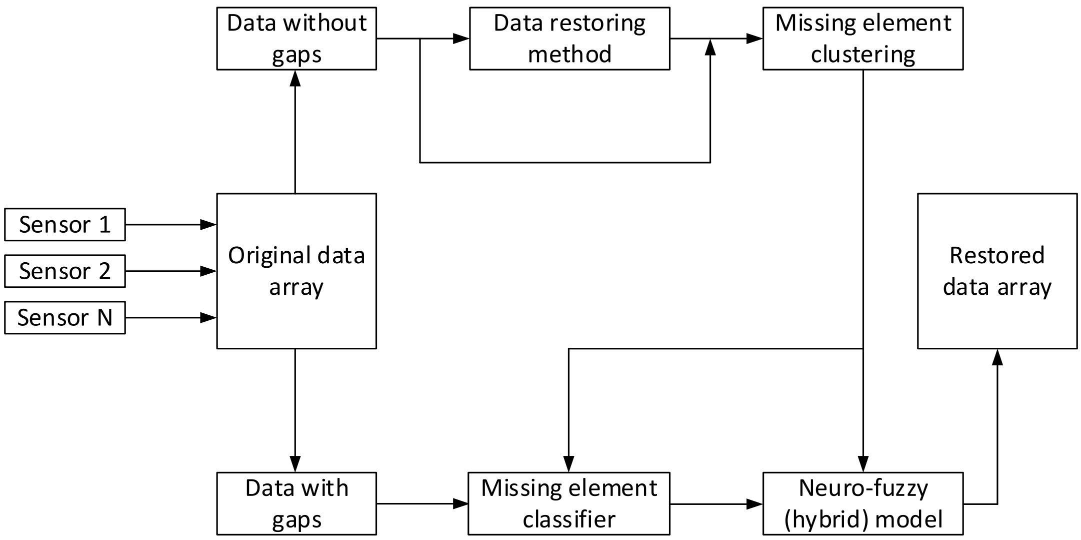

Figure 1: The scheme for the proposed hybrid method for restoring missing sensor data, featuring control based on neuro-fuzzy networks. (author's research).

Based on this scheme, the following algorithm has been developed:

Step 1. The training dataset formation.

1.1. The subset $S \subseteq D$ identification from the array containing complete data to create a reliable foundation for training, where $|S| = n$ represent instances number.

1.2. The elements' specified number m is randomly removed from the selected subset S, creating gaps in the data, resulting in a new dataset S' such that:

$$S' = S - \{x_1, x_2, \ldots, x_m\}, \tag{3}$$

where x_i denotes removed elements.

1.3. The missing element values y_i estimation is performed using all available methods M, such as regression, interpolation, or machine learning techniques.

$$y_i = f\left(x_{j_1}, x_{j_2}, \dots, x_{j_k}\right), \tag{4}$$

with $j_1, j_2, \dots, j_k$ being indices of available elements in S'.

1.4. The training dataset T is formed based on the obtained estimates and missing elements known values, creating input-output pairs for model training, i.e.:

$$T = \left((x_i, y_i) \middle| x_i \in S', y_i \text{ is estimated or known}\right) \tag{5}$$

1.5. Steps 1.2–1.4 are repeated until the training dataset T amount reaches a sufficient size, defined as the instances number exceeding a predetermined threshold $T_{\min}$ (e.g., 500 instances, i.e., $|T| > 500$).

Step 2. The control element creation.

2.1. A clustering method C (for instance, fuzzy C-means) is applied to determine elements groups $G_k \subseteq S$ in the array exhibiting similar accuracy in method performance.
2.2. A control element is developed and trained is a classifier $C_{classifier}$ for missing elements based on known portions x_i and assigned groups G_k. The classifier primary idea is to the elements group identify (from step 2.1) to which a particular missing value belongs, based on the instance with gaps known part. Instances from the array with randomly removed elements (from step 1.2) are used for training. Thus,

$$C_{classifier}\colon x_i \mapsto G_k\,. \tag{6}$$

Step 3. Training hybrid models.

For each identified group G_k, a hybrid model H_k is constructed based on a neuro-fuzzy network utilizing a feedforward structure. Training is conducted on the dataset formed in step 1 portion T, corresponding to the elements specific group G_k, ensuring that the model accounts for each group characteristics. In this context, it is advisable to employ a Takagi-Sugeno neuro-fuzzy network architecture [21–23] since it effectively combines fuzzy rules with neural computations, providing a high degree of adaptability to changing data conditions. This architecture includes fuzzy rules for managing uncertainty in data, which is crucial in the variability presence and real measurements noise typical [21, 22]. Moreover, the networks' neural component enables modeling complex nonlinear relations between input and output data, leading to the missing values more accurate restoring [23]. This architecture use also facilitates adaptive learning mechanisms [22, 24] that can account for each group characteristics, forming flexible and highly effective models tailored to the data specifics, significantly enhancing prediction accuracy in missing value restore tasks. The Takagi-Sugeno-type architecture employs fuzzy rules R_j defined as:

$$R_j\text{: IF “}x\text{” IS “}A_j\text{” THEN “}y = f_j(x)\text{”}, \tag{7}$$

where A_j denotes fuzzy sets and $f_j(x)$ represents linear combinations for outputs.

Step 4. The missing values restoring.

An instance $x_{missing}$ containing missing values is input into the control element, realized based on the classifier. The control element decides d_k which hybrid model should be applied for the given instance, i.e.:

$$d_k = C_{classifier}(x_{missing}). \tag{8}$$

Apply the selected hybrid model H_k to restore missing values $y_{restored}$:

$$y_{restored} = H_k(x_{missing}). \tag{9}$$

Additionally, a confidence assessment mechanism $C_{confidence}$ is implemented, which analyzes the uncertainty level when selecting a model, considering factors such as the initial data quality and the outcomes from previous models. This will improve the value restore accuracy, particularly in complex or atypical situations. Thus,

$$C_{confidence}(d_k) = \text{quality}(x_{missing}) + \text{performance}(H_k). \quad (10)$$

The proposed algorithm scientific novelty lies in its comprehensive approach to restoring missing values, which includes stages for forming a training dataset, clustering elements, and adaptively selecting hybrid models. Unlike similar methods [17, 19, 20], the algorithm does not simply apply a single universal model to all data; it first elements groups identifies with similar performance accuracy, allowing for the specialized hybrid models optimized construction for each group. Furthermore, the control element implementation in the classifier form, which makes decisions regarding model selection based on the instance specific characteristics, ensures higher accuracy and reliability in restoring missing values under data uncertainty and variability conditions.

The Takagi-Sugeno neuro-fuzzy network architecture (Figure 2) for missing value restore is designed to combine the fuzzy logic strengths for handling uncertainty and neural networks for modeling complex, nonlinear relations between inputs and outputs.

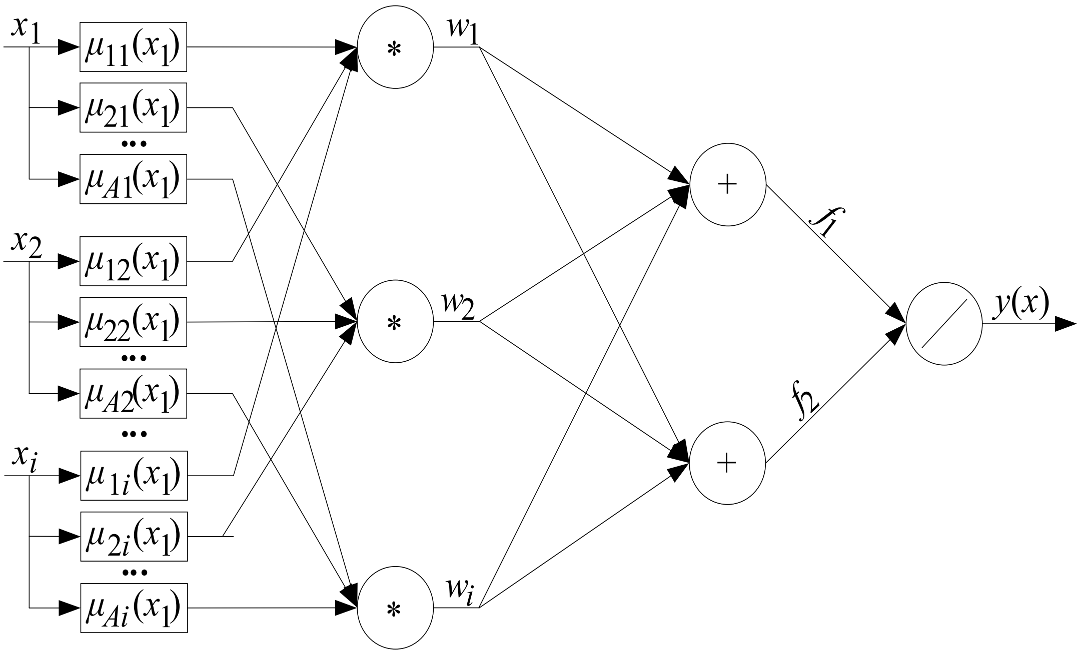

Figure 2: The Takagi-Sugeno neuro-fuzzy network proposed architecture [22, 23].

The input layer receives the available data from an instance $x \in R$, which contains some missing elements. The input layer goal is to feed the known values of the instance into the subsequent layers, where $x_1, x_2, \ldots, x_k$ are known elements and $y_1, y_2, \ldots, y_m$ are the missing elements. For each instance, the input is:

$$x = [x_1, x_2, \ldots, x_k, \_, \_, \ldots, \_]. \quad (11)$$

In fuzzification layer, each input is mapped to fuzzy sets. The fuzzification process uses membership functions (typically Gaussian or triangular) to represent uncertainty in the input values. Each input x_i is associated with a fuzzy set A_j, where the membership degree is computed:

$$\mu_{A_j}(x_i) = \exp\left(-\frac{(x_i - c_j)^2}{2 \cdot \sigma_j^2}\right), \quad (12)$$

where c_j is the fuzzy set center and σ_j is its width.

The rule layer consists of fuzzy IF-THEN rules in the Takagi-Sugeno form [21, 23]. Each rule R_j takes the form:

$$R_j\text{: IF } (x_1 \text{ is } A_{j1}) \text{ AND } (x_2 \text{ is } A_{j2}) \text{ AND} \ldots \text{ THEN } y_j = f_j(x), \tag{13}$$

where $f_j(x)$ is the inputs linear function:

$$f_j(x) = w_1 \cdot x_1 + w_2 \cdot x_2 + \ldots + w_k \cdot x_k + b_j. \tag{14}$$

The rule layer aggregates multiple rules, each corresponding to different fuzzy regions of the input space.

In the inference layer, the network evaluates the degree to which each rule applies to the given input. This is done by calculating the product of the membership degrees for each rule:

$$\mu_j = \prod_{i=1}^{k} \mu_{A_j}(x_i). \tag{15}$$

This layer output is the firing strength μ_j of each rule, indicating how well the input matches the rule fuzzy conditions.

In the defuzzification layer, the network merges the results from all rules to generate the final output by applying a modified center of gravity equation [25]:

$$y = \frac{\sum_j w_j \cdot f_j(x) \cdot \mu_j}{\sum_j w_j \cdot \mu_j}. \tag{16}$$

Here, $f_j(x)$ represents the linear function output corresponding to the j-th rule, and μ_j is the firing strength of the j-th rule.

The output layer produces the restored missing values $y_{restored}$. The result is the estimated values set for the missing elements based on the weighted contribution from the fuzzy rules. Table 1 presents the Takagi-Sugeno neuro-fuzzy network training algorithm.

The Takagi-Sugeno neuro-fuzzy network's architecture and training process allows for flexible, adaptive modeling of complex relations in data, making it highly effective for restoring missing values in datasets with uncertainty and variability. By clustering elements and assigning specialized models, this approach enhances accuracy and robustness in challenging restoring tasks.

Table 1
The Takagi-Sugeno neuro-fuzzy network training algorithm (author's research).

Step number	Step name	Description
Step 1	Initialization	The membership function parameters c_j, σ_j and the linear function weights w_1, w_2, ..., w_k, and biases b_j are initialized randomly.
Step 2	Forward Propagation	For each training instance x the membership degree in the fuzzification layer is computed. The rule activations (firing strengths) in the inference layer are computed. The network output is computed by aggregating the rule outputs using a weighted average in the defuzzification layer.
Step 3	Error Calculation	The error as the difference between the network output $y_{restored}$ and the true value y_{true} for the missing element is defined as: $$E = \frac{1}{2} \cdot \sum_{i=1}^{m} \left(y_{restored,i} - y_{true,i}\right)^2.$$

Step 4	Backpropagation	The error gradient with respect to the parameters w_j, b_j of the linear functions in the rule layer is computed. For each weight w_j, the gradient is calculated as: $$\frac{\partial E}{\partial w_j} = \frac{\partial E}{\partial y_{restored}} \cdot \frac{\partial y_{restored}}{\partial w_j}.$$ The gradients for the membership function parameters w_j, b_j are also computed using the chain rule, adjusting the centers and widths of the fuzzy sets: $$\frac{\partial E}{\partial c_j} = \frac{\partial E}{\partial y_{restored}} \cdot \frac{\partial y_{restored}}{\partial \mu_j} \cdot \frac{\partial \mu_j}{\partial c_j}.$$
Step 5	Parameter Update	The parameters using RMSProp optimization method is updated as: $$\theta_j^{(t+1)} = \theta_j^{(t)} - \frac{\eta}{\sqrt{G_j^{(t)} + \epsilon}} \cdot g_j^{(t)},$$ where $g_j^{(t)} = \frac{\partial E}{\partial \theta_j^{(t)}}$, $G_j^{(t)} = \rho \cdot G_j^{(t-1)} + (1-\rho) \cdot \left(g_j^{(t)}\right)^2$, ρ is the decay rate (usually set to 0.9), and $G_j^{(t)}$ is the squared gradients running average, η is the training rate, ϵ is a small value (e.g., 10^{-8}) added for numerical stability, θ_j represent each parameter (which can be w_j, b_j, c_j, or σ_j).
Step 6	Repeat	The steps 2–5 are repeated until the network converges, i.e., the error falls below a predefined threshold

4. Results

The research presents the results from a computational experiment focused on restoring missing gas temperature before the compressor turbine T_G^* data in the TV3-117 TE [26], which is part of the Mi-8MTV helicopter's power plant. The parameter values are recorded by a dual sensor comprising 14 T-101 thermocouples [27] at 1-second intervals (Figure 3). Data recording was carried out over a 256-second time interval. The application of temporal discretization and adaptive quantization with a dynamic range to the reconstructed T_G^* signal diagram (Figure 3) resulted in its discrete form (Figure 4).

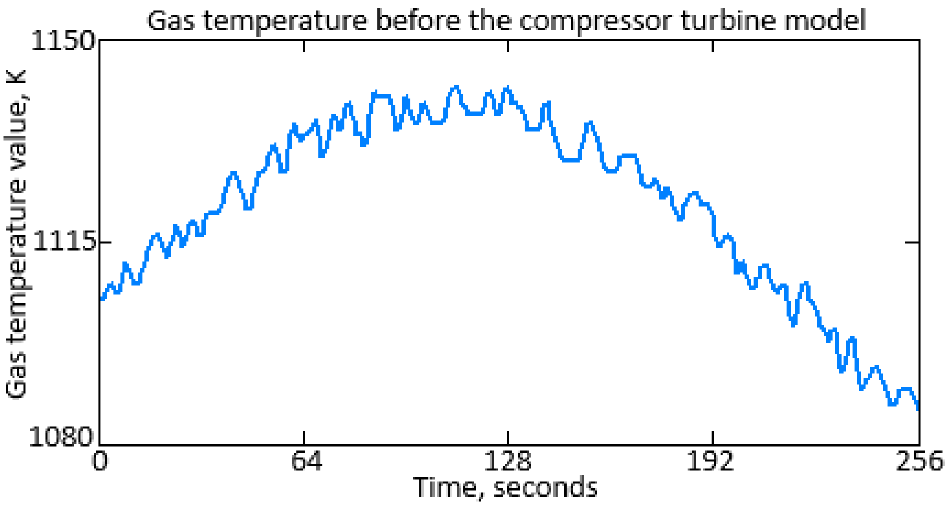

Figure 3: Diagrams of the gas temperature before the compressor turbine T_G^* parameter recorded onboard the helicopter during the 256-second study interval. (author's research).

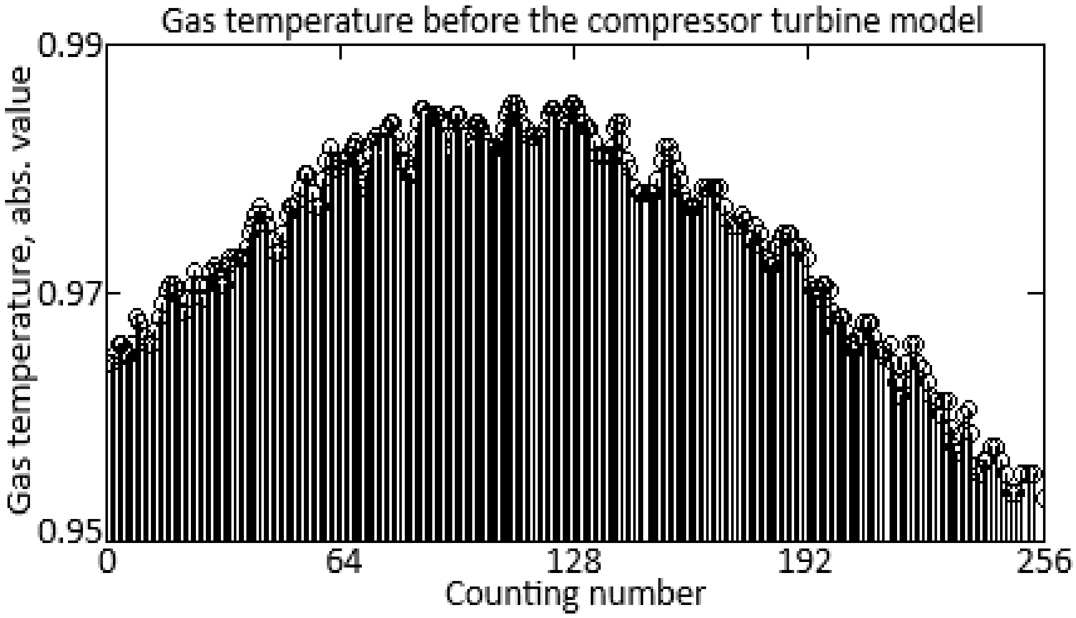

Figure 4: Diagrams of the gas temperature before the compressor turbine T_G^* reconstructed parameter recorded onboard the helicopter during the 256-second study interval. (author's research).

The generated diagrams (Figures 3 and 4) play a key role in creating the training dataset. According to the Nyquist theorem, the sampling rate must be at least twice the maximum frequency for accurate signal reconstruction [28]. With 256 readings over 256 seconds, the sampling rate is 256 / 256 = 1.0 readings per second, which corresponds to a time interval between readings of approximately 7 seconds. This frequency is well-suited for tracking slowly varying parameters and provides the necessary detail without excessive data amount [27, 28]. Thus, selecting 256 readings represents an optimal compromise between sampling rate, data volume, and computational costs, ensuring a balance between accuracy and data processing efficiency. As shown in Figure 4, the training dataset for the T_G^* parameter has been created and is presented in Table 2. After normalization, it appears as shown in Table 3.

Table 2
The training dataset fragment of the gas temperature before the compressor turbine T_G^* parameter values (author's research).

Number	**1**	**…**	**38**	**…**	**84**	**…**	**173**	**…**	**256**
Value	1106	…	1118	…	1130	…	1122	…	1104

Table 3
The training dataset fragment of the gas temperature before the compressor turbine T_G^* parameter normalized values (author's research).

Number	**1**	**…**	**38**	**…**	**84**	**…**	**173**	**…**	**256**
Value	0.971	…	0.979	…	0.989	…	0.984	…	0.968

Thus, the normalized values for the T_G^* parameter range from 0.95 to 0.99. Let's assume there are missing values for the T_G^* parameter between 56 and 65 seconds, 137 and 144 seconds, and 215 and 223 seconds.

According to the developed algorithm, at the initial stage, we select a subset $S \subseteq D$ from the complete dataset, containing all T_G^* parameter values over the entire period without missing data. Let's assume, for example, that $|S|$ = 256 records.

Next, we remove data from S corresponding to the periods from 56 to 65 seconds, 137 to 144 seconds, and 215 to 223 seconds. Thus, S' is the subset S with the removed values, and its size becomes $|S'|$ = 256

– 28 = 228 values. The missing values will be restored using various methods (regression, interpolation, machine learning). It is assumed that the implementation of fuzzy rules by a neuro-fuzzy network (Figure 2) for restoring the missing values is carried out according to expressions such as:

$$y_{56} = \frac{T_{55} + T_{66}}{2}, y_{57} = \frac{T_{56} + T_{67}}{2}, \dots, y_{65} = \frac{T_{64} + T_{66}}{2}. \tag{17}$$

Similarly, interpolation is performed for the other intervals with missing data. At the next stage, a training set T is created, including pairs of inputs x_i and their corresponding restored values y_i, using the missing and known data. The dataset size is $|T| > 500$. After that, a fuzzy clustering method (such as fuzzy C-means) is applied to divide the data into groups G_k with similar restoring characteristics. For example, let the data be divided into 3 groups:

$$G_1: [0.95 \dots 0.96], G_2: [0.96 \dots 0.97], G_3: [0.97 \dots 0.99]. \tag{18}$$

The classifier is trained on the data to determine which group G_k the missing values belong to, based on the data known portion. For example, if the missing T_G^* parameter value is 0.955, the classifier will assign it to group G_1. For each group G_k, a hybrid model H_k is constructed based on the Takagi-Sugeno neuro-fuzzy network (Figure 2). A separate model H_1 will be built for group G_1, model H_2 for group G_2, and so on. These models are trained on the corresponding data from the training set T according to the proposed training algorithm (see Table 1).

For the missing value at the 56th second, using the known data, for example, T_{55} = 0.956, the classifier $C_{classifier}$ determines that this value belongs to group G_1. To restore the temperature value at the 56th second, model H_1 (corresponding to group G_1) is applied, which reconstructs the missing value $y_{56} = H_1(T_{55}, T_{57}, \dots)$. The confidence score $C_{confidence}$ is calculated based on the quality of the input data and the performance of model H_1, helping to improve the restoration accuracy. Let the neighboring parameter values for the adjacent seconds appear as follows (in normalized form): T_{54} = 0.960, T_{55} = 0.955, T_{66} = 0.957, T_{67} = 0.958. For the 56th second, using the neuro-fuzzy network (Figure 2), the missing value is restored as:

$$y_{56} = \frac{T_{55} + T_{66}}{2} = \frac{0.955 + 0.957}{2} = 0.956. \tag{19}$$

Similarly, other missing values are restored by applying hybrid models and regression methods for more accurate restoring. Table 4 presents the results evaluating the quality of solving the missing T_G^* parameter values restoring task. The experimental results show that for most values, the deviations are only 0.002...0.005, confirming the high accuracy achieved by the restoration models. The maximum deviation of 0.007 also falls within acceptable error limits, proving the stability and reliability of the proposed approach. Thus, hybrid models and regression methods demonstrate excellent performance in restoring critical parameters, ensuring accurate system operation even in the event of sensor failures.

Table 4
The results evaluating the quality of solving the missing T_G^* parameter values restoring task

Number	The true value with a functioning sensor	The restored value in case of a sensor failure	Discrepancies between the true and restored values
56	0.959	0.956	0.003
57	0.960	0.957	0.003
58	0.960	0.955	0.005
59	0.961	0.956	0.005
60	0.960	0.956	0.004
61	0.960	0.957	0.003
62	0.960	0.958	0.002

63	0.961	0.958	0.003
64	0.961	0.957	0.004
65	0.961	0.957	0.004
137	0.985	0.982	0.003
138	0.987	0.984	0.003
139	0.988	0.984	0.004
140	0.987	0.983	0.004
141	0.987	0.982	0.005
142	0.985	0.979	0.006
143	0.985	0.980	0.005
144	0.985	0.978	0.007
215	0.973	0.969	0.004
216	0.971	0.966	0.005
217	0.972	0.970	0.002
218	0.970	0.969	0.001
219	0.973	0.969	0.004
220	0.975	0.970	0.005
221	0.974	0.970	0.004
222	0.976	0.971	0.005
223	0.978	0.974	0.004

The neuro-fuzzy network's performance (Figure 2) evaluation was conducted using traditional metrics, including Accuracy, Loss, Precision, Recall, and F1-score [29]. In the context of restoring missing sensor values on helicopters, Accuracy reflects the proportion of correctly restored values compared to the actual sensor data. Loss measures the prediction error, indicating how far the restored values deviate from the true ones [30]. Precision and Recall evaluate the model's effectiveness in correctly identifying and recovering missing values, while F1-score provides a balance between Precision and Recall, ensuring both completeness and accuracy in the restoration process [29, 30]. These metrics are determined using the formulas provided in [29]:

$$
\begin{gathered}
Accuracy = \frac{1}{N} \cdot \sum_{i=1}^{N} \mathbf{1}(y_i = \hat{y}_i)\,, Precision = \frac{TP}{TP + FP}, Recall = \frac{TP}{TP + FN},\\
F1 - score = \frac{2 \cdot Precision \cdot Recall}{Precision + Recall}, AUC - ROC = \int_{0}^{1} TPR \cdot \left(FPR^{-1}(t)\right) dt.
\end{gathered}
\tag{20}
$$

In the context of restoring missing sensor-registered parameters on board the helicopter, *TP* (True Positives) refers to correctly restored values that match the actual sensor data when the sensor is functioning. *TN* (True Negatives) represents cases where missing data were correctly identified as non-recoverable or not requiring restoration. *FP* (False Positives) indicates instances where data were incorrectly restored, leading to discrepancies between the restored and actual values. *FN* (False Negatives) refers to cases where the system failed to restore missing data correctly, leaving a gap in the sensor readings.

Figures 5 and 6 display the Accuracy and Loss metric diagrams, demonstrating the network's high effectiveness. Specifically, the Accuracy exceeds 99 %, while the Loss metric decreases from 2.5% to 0.5% over 180 training epochs, indicating significant improvement in model quality and training efficiency. This confirms the network's capability for high-precision classification tasks with minimal error.

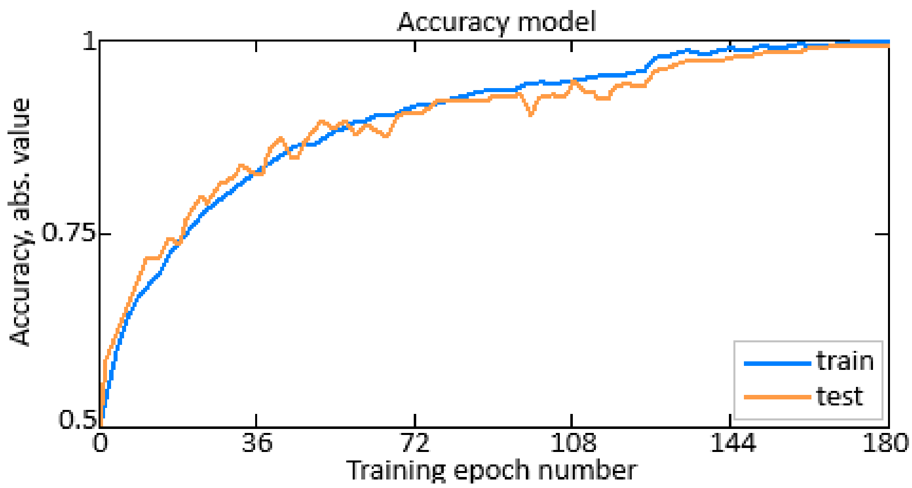

Figure 5: The accuracy metric diagram. (author's research).

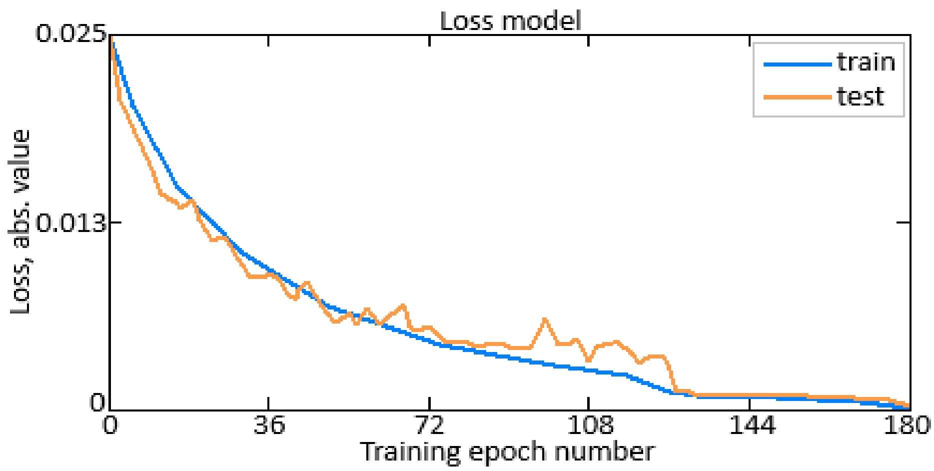

Figure 6: The loss metric diagram. (author's research).

Specifically, the Accuracy (see Figure 5) exceeds 99 %, while the Loss (see Figure 6) metric decreases from 2.5 to 0.5 % over 180 training epochs, indicating significant improvement in model quality and training efficiency. This confirms the network's capability for high-precision classification tasks with minimal error.

These results highlight the robustness and reliability of the neuro-fuzzy network in achieving high classification accuracy with minimal loss. The steady reduction in the Loss metric demonstrates the model's ability to generalize well to new data, further supporting its applicability in real-world scenarios requiring precise decision-making and control.

In the context of restoring missing sensor-registered parameters on board the helicopter, obtained Precision is 0.983 means that 98.3 % of the values identified as successfully restored are indeed correct, indicating a very low false positive rate. Recall is 1.0 means that the method was able to correctly restore all the missing values (no false negatives), achieving 100 % completeness in the restoration process. The F1-score is 0.991 combines Precision and Recall, indicating a balance between accuracy and completeness, with a near-perfect performance in restoring the missing sensor data.

A comparison (Table 5) of the proposed approaches for restoring missing sensor-registered parameters on board the helicopter was conducted by replacing the developed neuro-fuzzy network with a two-layer feedforward network (alternative approach 1) and the ANFIS neuro-fuzzy network (alternative approach 2), using a traditional training algorithm. This comparison aimed to evaluate the performance of different neural network structures in the context of the restoration task.

Table 5
The comparative analysis results (author’s research).

Number	Proposed approach	Alternative approach 1	Alternative approach 2
Accuracy	0.991	0.932	0.976
Precision	0.983	0.924	0.971
Recall	1.0	0.817	1.0
F1-score	0.991	0.867	0.985

The comparative analysis of the proposed neuro-fuzzy network against alternative approaches revealed superior performance in restoring missing sensor-registered parameters on board the helicopter. The proposed approach achieved an accuracy of 0.991, significantly higher than the 0.932 obtained by the two-layer feedforward network and the 0.976 from the ANFIS network. Furthermore, the proposed method demonstrated exceptional precision (0.983) and an F1-score (0.991), indicating its effectiveness in accurately identifying and restoring missing values compared to the alternative methods, which showed lower precision and recall rates.

5. Discussion

This research lies foundation in the development of a hybrid method for restoring missing sensor data with adaptive control (see Figure 1). The proposed method is based on an algorithm that involves forming a training dataset, creating a control element, and training hybrid models. The restoration of missing values includes identifying a data subset, randomly removing elements to create gaps, estimating their values using various methods, clustering to identify groups with similar accuracy, developing and training a classifier, constructing neuro-fuzzy models for each group, and utilizing a control element to select the model and assess the classifier's confidence. This collective approach enhances the accuracy of restoring missing values in complex situations. Unlike similar methods, this algorithm does not use a single universal model for all data; instead, it first identifies groups of elements with similar accuracy, allowing for the development of specialized hybrid models optimized for each specific group.

The developed algorithm is implemented as a Takagi-Sugeno neuro-fuzzy network (see Figure 2), as it effectively combines fuzzy rules and neural computations (see Table 1), ensuring high adaptability to changing data conditions and allowing for the processing of uncertainty. This is particularly important in environments with variability and noise in real measurements, as well as modeling complex nonlinear dependencies between input and output data, facilitating more accurate restoration of missing values. A computational experiment has been conducted to restore missing gas temperature data before the compressor turbine (see Figures 3 and 4, Tables 2 and 3) for the TV3-117 TE. The results confirm the high accuracy of the restoration models, with deviations for most values ranging from 0.002 to 0.005, and the maximum deviation of 0.007 falling within acceptable error limits (see Table 4). This indicates the stability and reliability of the approach, ensuring precise system operation even in cases of sensor failures.

The quality of the proposed Takagi-Sugeno neuro-fuzzy network (see Figure 2) has been assessed based on traditional quality metrics: Accuracy, Loss, Precision, Recall, and F1-score. The model's accuracy exceeds 99 % (see Figure 5), while the Loss value decreases from 2.5 to 0.5 % (see Figure 6) over 180 training epochs, demonstrating a significant improvement in model quality and training efficiency. These results confirm the capability of the neuro-fuzzy network to perform high-precision classification tasks with minimal errors, highlighting its reliability and robustness. In the context of restoring missing values recorded by sensors on board the helicopter, the achieved Precision of 0.983 indicates that 98.3 % of the values identified as successfully restored are indeed correct. The Recall of 1.0 confirms the complete restoration of all missing values, while the F1-score of 0.991 indicates a balance between accuracy and completeness, demonstrating nearly perfect performance in restoring missing sensor data.

The comparative analysis of the proposed approaches for restoring missing sensor-registered values on board the helicopter showed superior results for the developed neuro-fuzzy network compared to alternative methods, such as a two-layer feedforward network and the ANFIS neuro-fuzzy network (see Table 6). The proposed approach achieved an accuracy of 0.991, significantly higher than 0.932 for the two-layer network and 0.976 for the ANFIS network. Furthermore, the method demonstrated exceptional precision (0.983) and an F1-score (0.991), indicating its high efficiency in identifying and restoring missing values compared to alternative methods, which exhibited lower accuracy and completeness metrics.

Despite the successes achieved in developing the hybrid method for restoring missing data, the research faces certain limitations. The effectiveness of the algorithm depends on the quality and completeness of the original data, as insufficient representation in the training dataset may negatively impact the accuracy of the restored values. The method may encounter challenges when working with data containing high levels of noise or anomalies, which could lead to a decrease in overall reliability. Although the proposed algorithm has demonstrated high accuracy, its performance may vary based on the specifics of different sensor types and helicopter operating conditions.

Prospects for further research include expanding the developed method to process data from other sensor types and in various operating conditions. It would also be beneficial to consider integrating the algorithm with machine learning methods, such as deep neural networks, to enhance adaptability and improve restoration quality. Additionally, developing more complex strategies for handling noise and anomalies in the data could significantly increase the algorithm's reliability, allowing for effective application in real-world scenarios that require high precision and resilience to failures.

6. Conclusions

This research successfully presents a hybrid method for restoring missing sensor data through adaptive control, emphasizing the development of specialized hybrid models tailored to specific groups of data. The algorithm incorporates multiple stages, including data subset identification, random removal of elements to simulate missing values, and various estimation methods, culminating in the construction of a Takagi-Sugeno neuro-fuzzy network. This approach not only enhances the accuracy of missing value restoration but also demonstrates adaptability to changing data conditions and the ability to handle uncertainty in real-world measurements.

The computational experiments conducted on the gas temperature data before the compressor turbine of the TV3-117 turboshaft engine confirm the efficacy of the proposed method. The results show deviations in the restored values ranging from 0.002 to 0.005, with a maximum deviation of 0.007, all within acceptable error limits. This indicates the method's stability and reliability in operational scenarios, even when sensor failures occur. The ability to maintain precision within these limits underscores the method's potential for practical applications in critical systems, such as helicopter avionics.

Furthermore, the performance metrics indicate that the model's accuracy exceeds 99%, with a Precision of 0.983 and an F1-score of 0.991, validating the robustness of the Takagi-Sugeno neuro-fuzzy network. These metrics demonstrate that the proposed method effectively balances accuracy and completeness, achieving near-perfect performance in the restoration of missing sensor data. The comparative analysis with alternative methods, such as a two-layer feedforward network with an accuracy of 0.932 and an ANFIS network with an accuracy of 0.976, illustrates the superiority of the developed approach.

This research contributes significantly to the field by addressing the challenges of missing data recovery in sensor systems. The proposed hybrid method, achieving an accuracy of 0.991, demonstrates substantial improvements over traditional approaches, paving the way for further advancements in the adaptive restoration of missing values. Future research could explore the extension of this method to different sensor types and operational conditions, potentially integrating advanced machine learning techniques to enhance adaptability and further improve restoration quality.

Acknowledgements

This paper is supported by the EU Erasmus+ programme within the Capacity Building Project “WORK4CE” (619034-EPP-1–2020-1-UA-EPPKA2-CFHE-JP).

Declaration on Generative AI

The authors have not employed any Generative AI tools.

References

[1] M. Pasieka, N. Grzesik, K. Kuźma, Simulation modeling of fuzzy logic controller for aircraft engines, International Journal of Computing. Research Institute for Intelligent Computer Systems 16:1 (2017), 27–33. doi: 10.47839/ijc.16.1.868.

[2] Y. Zheng, J. Zheng, K. Shao, H. Zhao, Z. Man, Z. Sun, Adaptive fuzzy sliding mode control of uncertain nonholonomic wheeled mobile robot with external disturbance and actuator saturation, Information Sciences 663 (2024) 120303. doi: 10.1016/j.ins.2024.120303.

[3] K. Zhang, J. Dai, X. Yu, G. Zhang, Missing data recovery based on temporal smoothness and time-varying similarity for wireless sensor network, Internet of Things 28 (2024) 101349.

[4] E. M. Cherrat, R. Alaoui, H. Bouzahir, Score fusion of finger vein and face for human recognition based on convolutional neural network model, International Journal of Computing 19:1 (2020) 11–19. doi: 10.47839/ijc.19.1.1688.

[5] L. Zhao, Z. Wang, T. Chen, S. Lv, C. Yuan, X. Shen, Y. Liu, Missing interpolation model for wind power data based on the improved CEEMDAN method and generative adversarial interpolation network, Global Energy Interconnection 6(5) (2023), 517–529.

[6] D. Okrushko, A. Kashtalian, System of distribution and evaluation of tasks in the software development process, Computer Systems and Information Technologies 2 (2023), 86–97.

[7] J. Loy-Benitez, S. Heo, C. Yoo, Soft sensor validation for monitoring and resilient control of sequential subway indoor air quality through memory-gated recurrent neural networks-based autoencoders, Control Engineering Practice 97 (2020), 104330.

[8] Q. Wang, J. Li, S. Zhang, H. Tian, S. Jie, C. Qu, Z. Liu, A novel temperature drift compensation method based on LSTM for NMR sensor, Measurement, 240 (2025), 115573.

[9] T. J. Maginga, E. Masabo, P. Bakunzibake, K. S. Kim, J. Nsenga, Using wavelet transform and hybrid CNN – LSTM models on VOC & ultrasound IoT sensor data for non-visual maize disease detection, Heliyon 10:4 (2024), e26647. doi: 10.1016/j.heliyon.2024.e26647.

[10] S. Vladov, R. Yakovliev, O. Hubachov, J. Rud, Neuro-Fuzzy System for Detection Fuel Consumption of Helicopters Turboshaft Engines, CEUR Workshop Proceedings 3628 (2023) 55–72. URL: https://ceur-ws.org/Vol-3628/paper5.pdf

[11] H. Zhao, W. Guo, Coordinated control method of multiple hybrid energy storage systems based on distributed event-triggered mechanism, International Journal of Electrical Power & Energy Systems 127 (2024) 106637. doi: 10.1016/j.prime.2024.100771.

[12] H. M. Omar, S. M. S. Mukras, Developing geno-fuzzy controller for suppressing quadrotor slung-load oscillations, Ain Shams Engineering Journal 14(8) (2023) 102051.

[13] X. Ban, H. Zhang, F. Wu, Advanced controller synthesis for fuzzy parameter varying systems, Journal of the Franklin Institute 359:2 (2022) 762–785. doi: 10.1016/j.jfranklin.2020.12.038.

[14] H. Diao, L. Yin, B. Liang, and Y. Chen, An intelligent system control method based on visual sensor, Measurement: Sensors 29 (2023), 100857. doi: 10.1016/j.measen.2023.100857.

[15] Z. Wu, W. Huang, K. Huang, Z. Liu, G. Lai, H. Xiao, C.L. Philip Chen, Inversion-based fuzzy adaptive control with prespecifiable tracking accuracy for uncertain hysteretic systems, Fuzzy Sets and Systems 498 (2025), 109141. doi: 10.1016/j.fss.2024.109141.

[16] K. Sun, A. Yin, Multi-sensor temporal-spatial graph network fusion empirical mode decomposition convolution for machine fault diagnosis, Information Fusion 114 (2025) 102708.

[17] F. Yang, R. Kang, Q. Liu, C. Shen, R. Du, A new active learning method for reliability analysis based on local optimization and adaptive parallelization strategy, Probabilistic Engineering Mechanics 75 (2024) 103572. doi: 10.1016/j.probengmech.2023.103572.

[18] K. El Hamidi, M. Mjahed, A. El Kari, H. Ayad, Adaptive Control Using Neural Networks and Approximate Models for Nonlinear Dynamic Systems, Modelling and Simulation in Engineering 2020 (2020), 1–13. doi: 10.1155/2020/8642915.

[19] B. Yu, Z. Zheng, M. Cai, W. Pedrycz, W. Ding, FRCM: A fuzzy rough c-means clustering method, Fuzzy Sets and Systems 480 (2024), 108860. doi: 10.1016/j.fss.2024.108860.

[20] Z. Yu, Z. Zhang, Q. Jiang, X. Yan, Neural network-based hybrid modeling approach incorporating Bayesian optimization with industrial soft sensor application, Knowledge-Based Systems 301 (2024), 112341. doi: 10.1016/j.knosys.2024.112341.

[21] L. H. Tran, A modified Takagi-Sugeno-Kang neuro-fuzzy network in estimation of shortage faults location and resistance on transmission lines, Engineering Applications of Artificial Intelligence 133 (2024), 108358. doi: 10.1016/j.engappai.2024.108358.

[22] E.-J. Pérez-Pérez, V. Puig, F.-R. López-Estrada, G. Valencia-Palomo, I. Santos-Ruiz, Neuro-fuzzy Takagi Sugeno observer for fault diagnosis in wind turbines, IFAC-PapersOnLine 56:2 (2023), 3522–3527. doi: 10.1016/j.ifacol.2023.10.1508.

[23] Y. Zhang, G. Wang, T. Zhou, X. Huang, S. Lam, K. S. Choi, J. Cai, W. Ding, Takagi-Sugeno-Kang fuzzy system fusion: A survey at hierarchical, wide and stacked levels, Information Fusion 101 (2024), 101977. doi: 10.1016/j.inffus.2023.101977.

[24] U. K. Gupta, D. Sethi, P. K. Goswami, Adaptive TS-ANFIS neuro-fuzzy controller based single phase shunt active power filter to mitigate sensitive power quality issues in IoT devices, e-Prime - Advances in Electrical Engineering, Electronics and Energy 8 (2024), 100542.

[25] S. Vladov, L. Scislo, V. Sokurenko, O. Muzychuk, V. Vysotska, A. Sachenko, A. Yurko, Helicopter Turboshaft Engines' Gas Generator Rotor R.P.M. Neuro-Fuzzy On-Board Controller Development, Energies, 17:16 (2024), 4033. doi: 10.3390/en17164033.

[26] R. M. Catana, G. Dediu, Analytical Calculation Model of the TV3-117 Turboshaft Working Regimes Based on Experimental Data, Applied Sciences 13:19 (2023) 10720.

[27] J. Liu, Sensor fault analysis of aero-engine using ensemble SCNN and Bayesian interval estimation, Engineering Applications of Artificial Intelligence 125 (2023) 106675.

[28] V. D. Deepak, N. K. Arun, K. V. Shihabudheen, An investigation into the stability of helicopter control system under the influence of time delay using an improved LKF, Journal of the Franklin Institute 361 (10) 106913. doi: 10.1016/j.jfranklin.2024.106913.

[29] S. Vladov, R. Yakovliev, V. Vysotska, M. Nazarkevych, V. Lytvyn, The Method of Restoring Lost Information from Sensors Based on Auto-Associative Neural Networks, Applied System Innovation 7:3 (2024) 53. doi: 10.3390/asi7030053.

[30] C. Wang, N. Shakhovska, A. Sachenko, M. Komar, A New Approach for Missing Data Imputation in Big Data Interface, Information Technology and Control 49:4 (2020) 541–555.