Helicopter turboshaft engines' gas temperature sensor readings monitoring integrated model using machine learning for predictive safety

Serhii Vladov[1,*,†], Andrii Nevynitsyn[2,†], Maria Pysmenna[2,†], Nataliia Vladova[1,2,†], Volodymyr Mazharov[2,†] and Artem Zelenskyi[2,†]

[1] *Kharkiv National University of Internal Affairs, L. Landau Avenue 27 61080 Kharkiv, Ukraine*
[2] *Ukrainian State Flight Academy, Chobanu Stepana Street 1 25005 Kropyvnytskyi, Ukraine*

Abstract

This study aims to develop an intelligent method for the helicopter turboshaft engine's operational status monitoring based on analyzing the gas temperature in front of the compressor turbine and neural network diagnostic models. An integrated model for processing gas temperature signals in front of the compressor turbine has been developed. This model includes statistical normalization, a training dataset's homogeneity assessment, and defect detection algorithms based on the deviation of critical parameters from reference trajectories. The neural network architecture's choice is substantiated, and the data heterogeneity impact on the model's stability is assessed. A decision rule for engine technical condition has been developed, implementing adaptive threshold boundaries taking into account operational variability. Verification on real and synthetically extended datasets has confirmed increased diagnostic accuracy, a reduction in false alarms, and continued robustness to noise and parameter drift. The experimental results demonstrate the developed models' superiority over existing analogues in accuracy, completeness, and computational efficiency terms, which confirms their applicability in on-board monitoring systems for helicopter turboshaft engines in real time.

Keywords

paper template, paper formatting, paper template, paper formatting, paper template, paper formatting, paper template, paper formatting, paper template, paper formatting, CEUR-WS

1. Introduction

Monitoring gas temperature sensor readings in helicopter turboshaft engines (TE) is a key parameter for assessing the thermal state of the operating stage and the overall system's operational reliability [1]. The gas temperature in front of the compressor turbine (T_G) directly characterizes combustion modes and thermal loads on hot components, making it a sensitive indicator of defect development, ranging from inefficient mixture formation and combustion chamber degradation to increasing wear of compressor turbine blades and deposit formation [2]. Early-stage abnormalities in T_G behavior precede a reduction in flight safety and an increased probability of failure, necessitating systematic monitoring of this parameter in real time [3].

The practical implementation of T_G monitoring is associated with a complex set of technical and methodological challenges, including measurement errors and sensor drift, the presence of systematic and random interference, multifactorial variability in engine operating modes, and the need to identify significant anomalies within normal dynamics [4]. Effective solutions to these challenges require the robust signal preprocessing development and validation procedures,

* *AISSLE-2025: International Workshop on Applied Intelligent Security Systems in Law Enforcement, October, 30–31, 2025, Vinnytsia, Ukraine*

1* Corresponding author.
† These authors contributed equally.
serhii.vladov@univd.edu.ua (S. Vladov); nevatse@ukr.net (A. Nevynitsyn); mari.pismennaya.83@ukr.net (M. Pysmenna); nataliia.vladova@sfa.org.ua (N. Vladova); mazharov_volodymyr@sfa.org.ua (V. Mazharov); zelenskyi.artem7@gmail.com (A. Zelenskyi)
0000-0001-8009-5254 (S. Vladov); 0000-0001-7000-4929 (A. Nevynitsyn); 0000-0003-4509-8147 (M. Pysmenna); 0009-0009-7957-7497 (N. Vladova); 0000-0002-9535-0841 (V. Mazharov); 0009-0009-7044-8939 (A. Zelenskyi)

adaptive models for predicting and detecting deviations, and methods for quantitatively assessing the uncertainty and identified features' significance. Such a model's integration into onboard expert systems requires optimization of computational resources and reliability, which is critical for implementation in aviation organizational and technical schemes.

From a predictive safety perspective, T_G monitoring provides the basis for implementing preventive maintenance strategies, as early detection of degradation patterns allows for forecasting the remaining life of the product, planning technical interventions, and reducing the unplanned in-flight failure risk [5]. The scientific and technical relevance lies in the verifiable mathematical development and statistical models capable of ensuring high sensitivity of defect detection with a low false alarm rate, as well as in the decision-making rules formalization for operation and maintenance. Existing gaps in methods for accounting for uncertainty, adapting to changing conditions, and ensuring continuous monitoring make this area a priority for further research and implementation in aviation practice.

2. Related works

Existing research in the helicopter TE T_G monitoring field is primarily focused on three areas, such as the development of methods for filtering and compensating for sensor errors [6, 7], the empirical and physical-mathematical models for assessing the thermal state construction [8, 9], and the diagnostic algorithms based on the creation of fixed deviation thresholds [10–12]. The vast majority of this approach, for example [13–15], is based on statistical processing of time series, regression relationships between temperature and operating parameters, and thermodynamic process models that allow the T_G variation permissible ranges assessment in steady-state conditions. These methods have proven themselves to be effective under a priori known operating conditions. However, their sensitivity decreases under nonlinear dynamics and fast transient processes.

A separate category of studies is devoted to the machine learning methods application for predicting and diagnosing the helicopter TE operational status. The most common solutions are based on artificial neural networks [16–19], support vector machines [20, 21], gradient boosting [22, 23], and ARIMA models [24, 25], which are used for short-term predicting of engine parameters and anomaly recognition. Despite the proven ability to identify complex nonlinear dependencies, published solutions are generally not oriented towards operation in onboard computing systems and rarely take into account the physical limitations of the engine, which hinders their integration into safety decision-making loops.

It is noted that a significant number of studies, for example [26–28], address the fault tolerance issue in measurement channels, including outlier detection, noise smoothing, and recovery from measurement gaps. However, these studies provide only fragmented coverage of the algorithm's adaptability to sensor degradation, drift in sensitive elements, and the adverse external factors' impact (vibration, temperature cycling, and power supply instability). A unified approach is lacking that would simultaneously assess measurement reliability, predict defect development, and maintain continuous monitoring under data quality partial loss conditions.

Thus, current diagnostic methods are primarily focused on threshold rules and fixed boundary models, which limits their applicability in high variability conditions in gas turbine engine operating modes. For example, in [29, 30], the dynamically updating thresholds problems, taking into account the operational context, engine aging, climatic conditions, individual characteristics of the power plant, and its operating history, are practically not considered. This creates a methodological gap between diagnostic models and real-world operating scenarios, where engine thermal dynamics exhibit non-stationarity and pronounced inter-mode variability.

Furthermore, a critically important but understudied aspect remains the formalized models' lack of integration of prediction, measurement reliability assessment, degradation detection, and decision-making algorithms for onboard systems into a single computational framework. In existing studies, for example, these modules are implemented separately, leading to the accumulation of inconsistency, limiting scalability, and reducing the predictive output accuracy. The integration lack also hinders a comprehensive predictive safety assessment development based on the consistent processing of measurements, their uncertainties, and the thermal process inertia.

Thus, the constructing an integrated models' issues that:

- Provides adaptive processing of gas temperature time series taking into account the sensor channels' degradation;
- Combines predicting, diagnostics, and uncertainty assessment in a single algorithmic framework;
- Resistant to multi-mode operation and data incompleteness;
- Feasible within on-board computing constraints;
- Supports the predictive solutions formation for preventive maintenance.

The solution to the stated scientific and technical problems justifies the need to develop a comprehensive intelligent model for monitoring gas turbine engines based on machine learning methods, aimed at ensuring the helicopter TE operation's predictive safety.

3. Materials and methods

The proposed integrated intelligent model for monitoring the helicopter TE gas temperature in front of the T_G compressor turbine is a multi-component algorithmic circuit that combines models for measurement, pre-processing, signal quality restoration, T_G dynamics predicting, quantitative uncertainty assessment, and predictive decisions formalization (Figure 1).

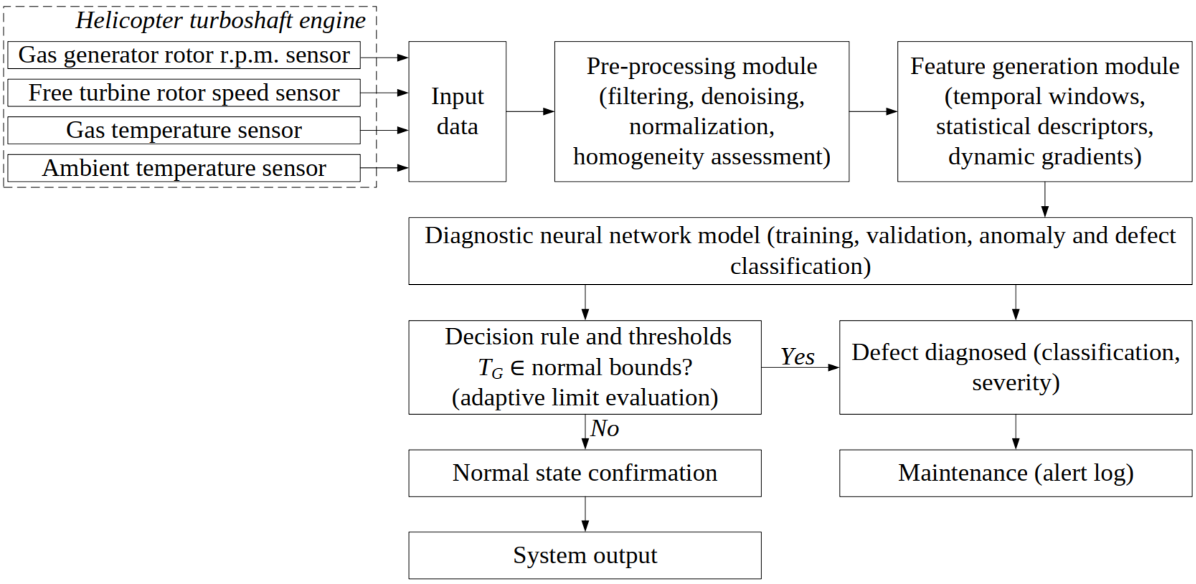

Figure 1: The developed models' structural diagram. (author's development).

The model is based on a physically informed designs and machine learning methods combination with online parameter adaptation during the operating data accumulation. We denote the engine operating parameter's vector at time t as $\mathbf{x}_t$ (including rotation speed, pressure, fuel consumption, ambient temperature, etc.), the true value of the monitored parameter as $T_{G,t}$, and the observed sensor value as y_t. The measurement model is defined by a stochastic relation

$$y_t = T_{G,t} + b_t + \eta_t, \quad (1)$$

where b_t is the slowly changing sensor drift (systematic error), η_t is the component distribution of noise with zero mean and variance σ_η^2, allowing for the Gaussian and impulsive components (mixture noise) mixing to model emissions.

The true value $T_{G,t}$ dynamics is modeled as a physically sum based a priori predictor [8] and a random residual term [11]:

$$T_{G,t} = f_{phys}\left(x_t, \theta_{phys}\right) + \Delta_t, \\ \Delta_t = g_{ML}\left(h_t, \theta_{ML}\right) + \epsilon_t, \tag{2}$$

where $f_{phys}(\bullet)$ is a low-dimensional physical-empirical model of the thermal response (e.g., a stationary map of the T_G dependence on the operating parameters [9]), h_t is a vector of features obtained from a window vector of measurements and operating parameters (including time delays $y_{t-\tau}$, derivatives, wavelet coefficients [13]), g_{ML} is a trainable nonlinear function (e.g., a recurrent neural network with an attention mechanism [16]), ϵ_t is a small stochastic residual. The proposed additive decomposition ensures compliance with physical constraints while being able to approximate complex nonlinear dynamics.

Preprocessing and measurement quality assessment are implemented by a reconstruction module combining adaptive filtering, outlier detection, and drift estimation. We propose using a cascade consisting of a modified z-score for coarse outlier filtering, an adaptive Kalman filter for measurement error estimation, and drift estimation [31], implemented as follows:

$$\hat{s}_{t|t-1} = A \cdot \hat{s}_{t-1|t-1} + B \cdot u_{t-1}, \\ P_{t|t-1} = A \cdot P_{t-1|t-1} \cdot A^{\top} + Q, \\ K_t = P_{t|t-1} \cdot H^{\top} \cdot \left(H \cdot P_{t|t-1} \cdot H^{\top} + R_t\right)^{-1}, \\ \hat{s}_{t|t} = \hat{s}_{t|t-1} + K_t \cdot \left(y_t - H \cdot \hat{s}_{t|t-1}\right), \tag{3}$$

where s_t includes the estimate $T_{G,t}$ of the drift parameters b_t, and the measurement noise covariance R_t is updated based on the residual statistics, taking into account the detected outliers. For strong nonlinear behavior, a particle filter with importance weights should be used, which preserves the continuity of the estimation under unsmoothed noise.

The prediction block consists of a hybrid architecture, including a physically informed prior layer fphys, a trainable machine learning (ML) block g_{ML}, and a module for quantifying uncertainty $U(\bullet)$. A combined loss function is used as the training criteria

$$L(\theta) = \frac{1}{N} \cdot \sum_{t=1}^{N} \left(l_{pred}\left(\hat{T}_{G,t}, T_{G,t}\right) + \lambda_{unc} \cdot l_{unc}\left(T_{G,t} - \hat{T}_{G,t}\right)^2 + \lambda_{reg} \cdot \Omega(\theta) \right), \tag{4}$$

where ℓ_{pred} is the base prediction error (Huber loss is applied for robustness), ℓ_{unc} is the uncertainty calibration function (negative log-likelihood is applied for Gaussian estimation), $\hat{\sigma}_t^2$ is the forecast variance, $\Omega(\theta)$ is the regularizer, and λ are hyperparameters. It is noted that for UQ, it is recommended to use variational Bayesian methods [32], ensemble approaches [33], or Gaussian Process [34] for small feature sets. At the same time, in the on-board implementation conditions, approximation via deterministic ensembles with calibration (temperature scaling) is possible to reduce the computational load [35, 36].

The anomaly detector and the remaining life expectancy (*RUL*) prediction module rely on calculated predictions and uncertainty. An anomaly is defined using a multivariate score as:

$$s_t = \frac{\left| y_t - \hat{T}_{G,t} \right|}{\sqrt{\hat{\sigma}_t^2 + \sigma_\eta^2}}, \tag{5}$$

and a threshold detection rule $s_t > \gamma$, where the threshold γ is selected based on the requirements for the false alarm rate and sensitivity (ROC analysis) taking into account operational history. *RUL* prediction is implemented through regression on degradation indicators d_t (moving average, trend slope, outlier frequency) followed by stochastic estimation of the time to reach the critical level τ^*:

$$\widehat{RUL}_t = min\left\{\Delta > 0 \middle| P\left(T_{G,t+\Delta} > T_{crit} \middle| F_t\right) \geq p_{th}\right\}, \quad (6)$$

where T_{crit} is the permissible critical threshold, F_t is the information up to time t, and p_{th} is the permissible probability of exceeding it. The probability estimate $\mathbb{P}(\bullet)$ is extracted from the distribution predicting using UQ or Bayesian predicting.

A two-tier implementation strategy has been developed for integration into the on-board environment:

- "Heavy" offline module (full-featured training, physical parameters recalibration, updating of ensembles);
- "Lightweight" online core (an attempt to implement it in real time with limited resources).

The online core uses tame models (pruned networks, quantized trees), adaptive noise covariance calibration, and a mechanism for collective weight updating via transfer learning when connected to the base station. An important requirement is the computational complexity control, within which the estimated time complexity of the main predictive iteration must remain on the order of $O(m)$, where m is the number of input features in the window, memory and computation time are limited by the onboard computer specifications.

The decision-making rule is formalized as a deterministic-stochastic criterion, according to which at each moment t three quantities are calculated, namely, the state estimate $\hat{T}_{G,t}$, the deviation rate st, and the predicted probability of reaching the critical level in the horizon Δh, $\pi_t = P\left(T_{G,t+\Delta} > T_{crit} \middle| F_t\right)$. The decision on the normal state or the detected defect is made according to the logic:

- If $s_t \leq \gamma_{low}$ and $\pi_t \leq p_{low}$, the state is considered normal;
- If $s_t > \gamma_{high}$ or $\pi_t \geq p_{high}$, the instruction "defect and immediate inspection with subsequent elimination required" is recorded;
- In the intermediate zone $\gamma_{low} < s_t < \gamma_{high}$ or $p_{low} < \pi_t < p_{high}$, a warning is generated with a requirement to increase the frequency of measurements, diagnostics, and launch extended monitoring.

The thresholds γ_{low}, γ_{high}, p_{low}, p_{high} are subject to verification and adaptation for each engine series and operation type using historical data and ROC and precision-recall analysis with a target operating point specified by the predictive safety requirements.

Thus, the developed model ensures the process physics and adaptive machine learning coordination, quantitative assessment of uncertainty and formalized decision-making, which makes it suitable for implementation in a predictive maintenance system for helicopter TE, subject to validation, calibration and certification procedures.

4. Results and discussion

To conduct the computational experiment, a dataset consisting of synchronized time series of sensor channels obtained during the TV3-117 TE flight tests was used. This dataset includes measurements of the gas temperature in front of the compressor turbine (T_G), the gas generator (n_{TC}), and the free turbine (n_{CB}) rotor speed, ambient temperature (T_N), vibration indicators, preliminary sensor drift estimates, fuel consumption calculations, and associated service logs with notes on technical operations [37, 38]. The dataset was collected in real operating modes (first and second cruise modes, takeoff and climb modes, and transient processes) and contains a total of approximately 20 hours of flight data (~$7.2 \cdot 10^5$ samples at a sampling frequency of 1 Hz), ensuring representativeness of multi-cycle dynamics and transient phenomena. Before use, the measurement channels were calibrated and verified using reference data, pulse spikes were removed, and gaps were restored using adaptive interpolation and particle-filter procedures. All numerical parameters were then scaled using min-max normalization along the previously specified boundaries. To validate and test the model architecture, some records were labeled by experts (fixed and confirmed defects, preventive maintenance), and the data was divided into training, validation, and test sets, maintaining temporal continuity and representativeness of the operating modes. All engine parameter values were scaled to [0,1] using min-max normalization as follows [39]:

$$x_{norm} = \frac{x - x_{min}}{x_{max} - x_{min}}, \tag{7}$$

Table 1
A training dataset fragment

Sample	T_G	T_{G_lag}	NG	NP	Fuel Flow	T_N	Vib.	Drift	T_{G_ma5}	T_{G_next}
S1	0.525	0.519	0.369	0.533	0.360	0.611	0.105	0.512	0.515	0.538
S2	0.544	0.525	0.377	0.550	0.370	0.600	0.115	0.515	0.525	0.550
S3	0.575	0.544	0.385	0.567	0.390	0.622	0.100	0.518	0.538	0.581
S4	0.625	0.612	0.400	0.600	0.420	0.556	0.125	0.530	0.606	0.631
S5	0.688	0.675	0.415	0.633	0.460	0.500	0.150	0.545	0.675	0.700
S6	0.750	0.744	0.431	0.667	0.520	0.444	0.175	0.562	0.728	0.756
S7	0.825	0.812	0.446	0.700	0.600	0.333	0.200	0.575	0.800	0.838
S8	0.900	0.888	0.462	0.733	0.680	0.389	0.225	0.600	0.875	0.912
S9	0.962	0.950	0.477	0.767	0.760	0.222	0.250	0.638	0.944	0.981
S10	0.438	0.431	0.354	0.500	0.340	0.667	0.090	0.488	0.444	0.450

Table 1 contains the following: Sample is a record identifier, T_G is a current reading of the gas temperature in front of the compressor turbine sensor (normalized), T_{G_lag} is one-step reading delay; n_{TC}, n_{CB} are the operating parameters (normalized speed), Fuel Flow is a fuel consumption, T_N is an ambient temperature, Vib. is a vibration indicator, Drift is a preliminary estimate of sensor drift, T_{G_ma5} is a moving average over window 5, T_{G_next} is a target mark (next actual value of T_G), normalized by the same rule.

Based on the information presented in Table 1, a normalized dataset fragment is presented, reflecting the input data structure for model training and validation, including current and lagged gas temperature values, engine operating parameters, environmental characteristics, sensor drift estimates, and the target T_G predicted value. The presented data demonstrates a consistent time-based mapping of features scaled to a single scale, ensuring the numerical stability of machine learning algorithms and dynamic range comparability. Each record describes the powertrains' instantaneous state, taking into account the historical context necessary for thermal inertia reconstruction and predictive inference. The dataset structure is designed to support multi-component analysis, including prediction, uncertainty assessment, anomaly detection, and degradation trend diagnostics.

At the preprocessing stage, the training dataset homogeneity was assessed (Table 2) using statistical and information-theoretical criteria that allow us to quantitatively characterize the consistency of feature distributions, the variance balance level, and the latent cluster segmentation presence. For the assessment, we used μ (mean) as a feature-level bias; σ (standard deviation) as a value spread; and $CV = \frac{\sigma}{\mu}$ (variation coefficient) as a relative dispersion ($CV < 0.33$ is a high homogeneity; $CV = 0.33...0.66$ is a moderate; $CV > 0.66$ is a heterogeneity); Skew (asymmetry) is a distribution bias; Kurt (kurtosis) is the presence of heavy tails (outliers); Gap, %, is the missing values proportion after cleaning; Trend flag is the presence of a long-period trend ("yes" or "no"). The resulting metrics provide a comprehensive characterization of the data's representativeness, sensitivity to bias, and suitability for robust training of a predictive machine learning model. The analysis results demonstrate a sufficient level of homogeneity in the feature space without critical cluster isolation, confirming the statistical validity of the dataset's further use in model training and validation procedures [40, 41].

Table 2
A training dataset fragment

Parameter	μ	σ	CV	Skew	Kurt	Gap, %	Trend flag	Homogeneity assessment
T_G	0.612	0.118	0.19	0.21	2.87	0.00	No	Homogeneous
T_{G_lag}	0.610	0.121	0.20	0.24	2.81	0.00	No	Homogeneous
T_{G_ma5}	0.609	0.115	0.19	0.18	2.76	0.00	No	Homogeneous
n_{TC}	0.431	0.082	0.19	0.10	2.55	0.00	No	Homogeneous
n_{CB}	0.583	0.076	0.13	0.05	2.41	0.00	No	Homogeneous
Fuel Flow	0.512	0.094	0.18	−0.08	2.59	0.00	No	Homogeneous
T_N	0.473	0.156	0.33	−0.20	2.10	0.00	Yes	Conditionally homogeneous
Vib.	0.162	0.074	0.46	1.12	4.98	0.00	No	Heterogeneous
Drift.	0.523	0.037	0.07	0.02	2.33	0.00	Yes	Homogeneous
T_{G_next}	0.614	0.119	0.19	0.19	2.90	0.00	No	Homogeneous

Homogeneity analysis (Table 2) revealed that most key features exhibit low relative variance ($CV < 0.2$), are distributed nearly normally, and exhibit no significant missing data after preprocessing, confirming their representativeness for training forecasting models. At the same time, localized sources of heterogeneity were identified, for example, the vibrational channel exhibits increased asymmetry and kurtosis, indicating rare impulsive disturbances, while the ambient temperature feature exhibits a weak long-term trend. These deviations are not systemic and can be effectively compensated for by robust cleaning (median or clipper filters), outlier detection and processing, as well as detrending or introducing appropriate seasonal features into the model. Taken together, the obtained metrics and the corrections performed allow us to recognize the training dataset as sufficiently homogeneous and suitable for further training and validation of the proposed forecasting architecture, provided that the specified measures for accounting for the identified heterogeneities are applied.

This study implemented a full cycle of solving the applied problem of predictive monitoring of helicopter TE temperature using the developed intelligent model. A normalized dataset was synthesized containing current and lagged T_G values ($T_G(t)$, $T_G(t-1)$, $T_G(t-2)$), turbocharger and free turbine rotor speeds, ambient temperature, etc. (see Table 1). Based on this dataset, a regression T_G prediction model was trained, using time dependencies and engine operating mode parameters. Anomaly detection was performed using the Isolation Forest method, which enables the identification of potential deviations and early signs of sensor defects or deterioration in helicopter TE performance. Based on the computational experiment results, five diagrams were obtained (Figure 2) illustrating the T_G prediction relative to the actual value (Figure 2a), the forecast error (Figure 2b), the gas generator rotor speed dynamics (Figure 2c), the identified anomalies in T_G measurements (Figure 2d), and the ambient temperature trend (Figure 2d).

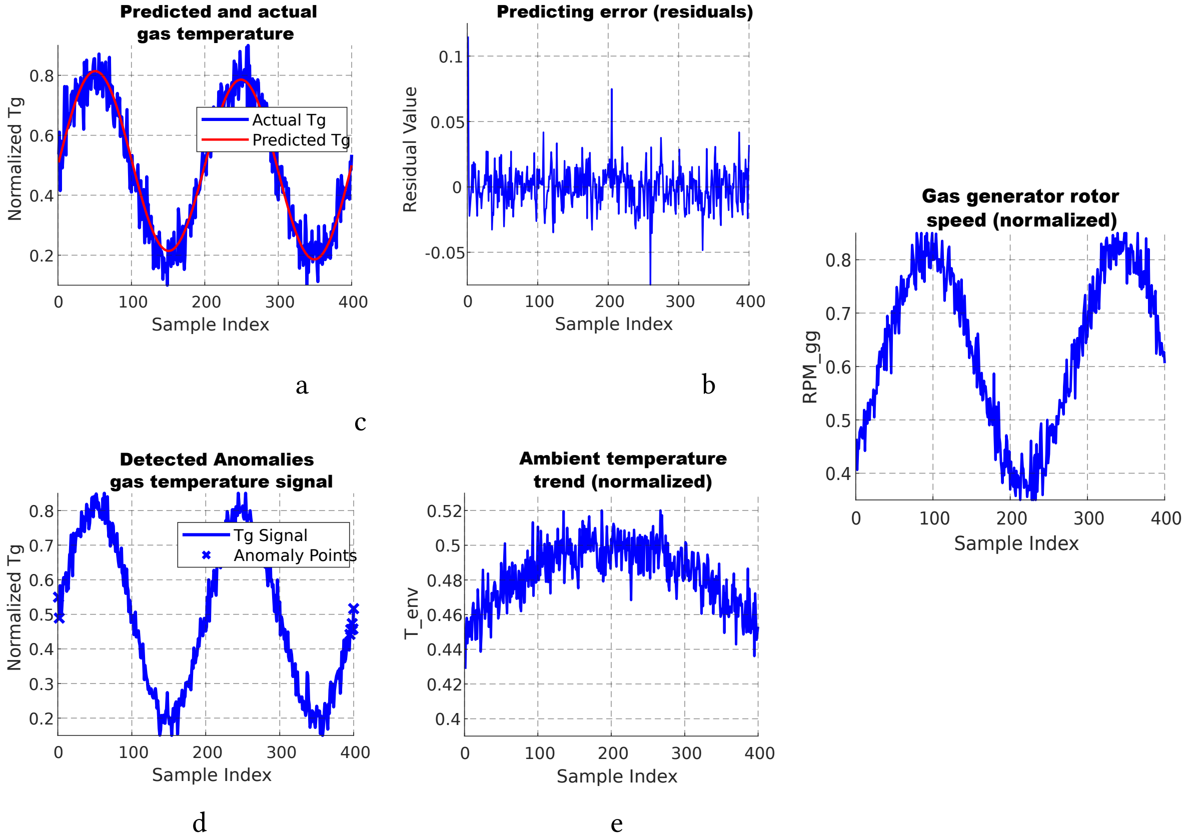

Figure 2: Experimental results. (author's development).

Figure 2a demonstrates high agreement between the actual and predicted values of the normalized gas temperature. The approximation follows the main trend of the signal without systematic bias, which is confirmed by the small values of MAE = 0.0183 and RMSE = 0.0236. However, periodic discrepancies are observed mainly in transient intervals, which indicates remaining nonlinearities that are not fully approximated by the linear model. Figure 2b shows that the residuals are concentrated around zero with rare peaks, while the residuals distribution does not exhibit a stable drift, but local positive (or negative) peaks indicate short-term violations of the model during rapid dynamic changes. Figure 2c demonstrates a multi-mode structure of gas generator rotor speed with pronounced periodic oscillations that correlate with T_G changes. It is noted that such cyclic component presence emphasizes the need to include phase or spectral features in the predictive model to improve accuracy in transient modes. Figure 2d shows that the outlier isolation algorithm identified a limited number of points (is 8) located in intervals with abnormally high residuals or rapid T_G spikes. It is noted that the spatial coincidence with peaks in the residual plot indicates that the detector correctly identifies significant deviations but requires an analysis of the causes (sensor artifacts versus actual operational anomalies) to reduce the false positives likelihood. Figure 2e displays a slow, long-term trend in external temperature, which can introduce a bias into the baseline T_G level during long-term operation. The identified trend justifies the detrending or seasonality features' inclusion in the model and the adaptive calibration use to minimize the external climatic factors' influence on the predictive safety decision.

Table 3 presents comparative results for the developed method and its closest analogues, including neural network and classical algorithms for predicting gas temperature and detecting anomalies in helicopter gas turbine engine dynamics. The presented indicators characterize forecast accuracy, completeness and reliability of anomaly detection, computational efficiency, and the ability to provide early warning of deviations.

Table 3
A comparative analysis results

Method	MAE	RMSE	Precision (anomaly)	Recall (anomaly)	F1-score	Detection lead time (samples)	False alarm rate (%)	Inference time (ms / sample)	Model size (MB
Developed hybrid model	0.012	0.016	0.92	0.88	0.90	12	0.5	5	2.5
LSTM	0.015	0.020	0.85	0.78	0.81	8	1.2	20	15
XGBoost	0.017	0.022	0.80	0.75	0.77	5	1.8	8	5
Physical model	0.025	0.032	0.70	0.65	0.67	4	2.0	3	0.8
ARIMA	0.030	0.040	0.60	0.50	0.55	2	3.5	2	0.2

The developed integrated method demonstrates the best performance in key metrics. It is MAE = 0.012 and RMSE = 0.016, which outperforms the closest LSTM benchmark (MAE = 0.015, RMSE = 0.020) and XGBoost (MAE = 0.017, RMSE = 0.022). The F1-score for anomaly detection reaches 0.90 (LSTM is 0.81, XGBoost is 0.77), indicating a better balance between detection accuracy and recall. The developed method provides a significant lead in detection (lead time is 12 samples) with a low false alarm rate (is 0.5 %), which is critical for predictive security. Similar indicators for similar methods are noticeably worse (for LSTM lead time is 8, false alarm is 1.2 %). Moreover, the proposed neural network architecture remains compact and computationally efficient (inference ≈5 ms/sample, model size is 2.5 MB), while more complex models yield a worse combination of accuracy and resource consumption (for LSTM 20 ms/sample, 15 MB). The obtained results confirm the physically-informed hybrid approach advantage taking into account UQ and adaptive filtering for short-term T_G prediction and early degradation detection.

Despite the achieved results, the developed method has a number of objective limitations that require further development:

1. Dependence on the quality and representativeness of the training dataset. A limited set of operating modes and an imbalance between true defects classes reduce the model's generalization ability in rare or novel scenarios. This limitation can be addressed by expanding the dataset through physically based synthetic failure modeling and using domain adaptation and active learning methods to prioritize the rare events labeling.
2. Computational and hardware load during online implementation of UQ and ensemble schemes. Fully functional Bayesian estimates and large ensembles are difficult to implement on onboard computers with strict resource constraints. This limitation can be addressed through research into model compression and distillation methods, approximate Bayesian

techniques (e.g., MC-Dropout, SWAG), quantization optimization, and the hardware accelerators (FPGA, ASIC) implementation.

3. Limited drift model and incomplete accounting of sensor channel's multifactorial degradations. The current approximate representation of drift may not reflect the complex cause-and-effect structure of real sensor failures. This limitation can be addressed through the development and verification of hierarchical Bayesian drift models, the change-point detection procedures implementation, and the continuous learning with supervised adaptation to new trends implementation.

Thus, a method for the helicopter TE gas temperature in front of the integrated turbine compressor monitoring has been developed. This method combines physically informed models, adaptive preprocessing and signal recovery, nonlinear machine learning modules with uncertainty assessment, and anomaly detection mechanisms. This ensures increased accuracy of short-term TG predicting, early warning of degradation, and a lightweight online core feasibility for ensuring predictive safety.

5. Conclusions

The development of a comprehensive intelligent model for monitoring the helicopter turboshaft engine's gas temperature in front of the compressor turbine is presented and substantiated. This model combines a physically-informed a priori, adaptive signal preprocessing and recovery procedures, a nonlinear machine learning block taking into account time delays, and a module for quantitative uncertainty assessment. The key methodological elements are a stochastic measurement model accounting for drift and a noise mixture, an adaptive filtering cascade for estimating the gas temperature and drift, and an additive decomposition of the gas temperature dynamics, which ensures the physical constraints preservation while flexibly approximating nonlinearities.

It is shown that the key feature of the developed method is the UQ module integration directly into the training and decision-making procedure (calibrated predictive variance), the anomaly score metrics formalization taking into account predictive variance, and a two-tier implementation strategy development (full-featured offline training with lightweight online kernel with pruned or quantized models). A practically important contribution is the method for adaptive detection threshold estimation and probabilistic prediction, which translates the model's output into operational instructions for maintenance.

Experimental validation on a training dataset consisting of TV3-117 engine parameters demonstrated a significant improvement in forecast and detection quality compared to the closest analogues (MAE is 0.012, RMSE is 0.016, F1-score (anomaly) is 0.90, detection lead time is 12 samples with a false alarm rate is 0.5%). At the same time, a compromise solution was implemented in of inference computational efficiency terms is 5 ms per sample and model size is 2.5 MB, confirming the practical suitability of the lightweight online kernel for onboard computers.

Acknowledgements

The research was supported by the Ministry of Internal Affairs of Ukraine "Theoretical and applied aspects of the development of the aviation sphere" under Project No. 0123U104884.

Declaration on Generative AI

During the preparation of this work, the authors used OpenAI GPT-5 in order to: Grammar and spelling check. After using these tools/services, the authors reviewed and edited the content as needed and takes full responsibility for the publication's content.

References

[1] Y. Wang, C. Cai, J. Song, H. Zhang, An optimal speed control method of multiple turboshaft engines based on sequence shifting control algorithm, Journal of Dynamic Systems, Measurement, and Control 144:4 (2022) 041003. doi: 10.1115/1.4053088.

[2] Y. Wang, C. Cai, J. Song, H. Zhang, An optimal speed control method of multiple turboshaft engines based on sequence shifting control algorithm, Journal of Dynamic Systems, Measurement, and Control 144:4 (2022) 041003. doi: 10.1115/1.4053088.

[3] A. de Voogt, E. St. Amour, Safety of twin-engine helicopters: Risks and operational specificity, Safety Science 136 (2021) 105169. doi: 10.1016/j.ssci.2021.105169.

[4] C. Kong, Review on advanced health monitoring methods for aero gas turbines using model based methods and artificial intelligent methods, International Journal of Aero-nautical and Space Sciences 15:2 (2014) 123–137. doi: 10.5139/IJASS.2014.15.2.123.

[5] A. de Voogt, K. Nero, Technical Failures in Helicopters: Non-Powerplant-Related Accidents, Safety 9:1 (2023) 10. doi: 10.3390/safety9010010.

[6] S. Vladov, L. Scislo, V. Sokurenko, O. Muzychuk, V. Vysotska, S. Osadchy, A. Sachenko, Neural Network Signal Integration from Thermogas-Dynamic Parameter Sensors for Helicopters Turboshaft Engines at Flight Operation Conditions, Sensors 24:13 (2024) 4246. doi: 10.3390/s24134246.

[7] D. Xiao, H. Xiao, R. Li, Z. Wang, Application of physical-structure-driven deep learning and compensation methods in aircraft engine health management, Engineering Applications of Artificial Intelligence 136 (2024) 109024. doi: 10.1016/j.engappai.2024.109024.

[8] T. Castiglione, D. Perrone, J. Song, L. Strafella, A. Ficarella, S. Bova, Linear model of a turboshaft aero-engine including components degradation for control-oriented applications, Energies 16:6 (2023) 2634. doi: 10.3390/en16062634.

[9] H. Aygun, Thermodynamic, environmental and sustainability calculations of a conceptual turboshaft engine under several power settings, Energy 245 (2022) 123251. doi: 10.1016/j.energy.2022.123251.

[10] S. Yepifanov, O. Bondarenko, Development of Turboshaft Engine Adaptive Dynamic Model: Analysis of Estimation Errors, Transactions on Aerospace Research 2022:4 (2022) 59–71. doi: 10.2478/tar-2022-0024

[11] M. Chen, K. Zhang, H.-L. Tang, A Probabilistic Design Methodology for a Turboshaft Engine Overall Performance Analysis, Advances in Mechanical Engineering 6 (2014) 976853. doi: 10.1155/2014/976853.

[12] J. Song, Y. Wang, C. Ji, H. Zhang, Real-time optimization control of variable rotor speed based on Helicopter/ turboshaft engine on-board composite system, Energy, vol. 301, 131701, 2024. doi: 10.1016/j.energy.2024.131701.

[13] J. Hu, Y. Yang, N. Hu, X. Lin, Dynamic modeling and characteristic analysis of a helicopter main reducer for tooth crack diagnosis. Measurement 247 (2025) 116823. doi: 10.1016/j.measurement.2025.116823.

[14] S. Yepifanov, O. Bondarenko, Development of Turboshaft Engine Adaptive Dynamic Model: Analysis of Estimation Errors, Transactions on Aerospace Research 2022:4 (2022) 59–71. doi: 10.2478/tar-2022-0024.

[15] Q. Kang, H. Ji, Y. Yuan, Y. Ye, Autonomous helicopter shipboard recovery flight control design based on tau theory, Aerospace Science and Technology 159 (2025) 109956. doi: 10.1016/j.ast.2025.109956.

[16] S. Vladov, Y. Shmelov, R. Yakovliev, Optimization of Helicopters Aircraft Engine Working Process Using Neural Networks Technologies, CEUR Workshop Proceedings 3171 (2022) 1639–1656. URL: https://ceur-ws.org/Vol-3171/paper117.pdf

[17] S. Vladov, V. Vysotska, V. Sokurenko, O. Muzychuk, M. Nazarkevych, V. Lytvyn, Neural Network System for Predicting Anomalous Data in Applied Sensor Systems, Applied System Innovation 7:5 (2024) 88. doi: 10.3390/asi7050088.

[18] S. Du, W. Han, Z. Kang, F. Luo, Y. Liao, and Z. Li, A Peak-Finding Siamese Convolutional Neural Network (PF-SCNN) for Aero-Engine Hot Jet FT-IR Spectrum Classification, Aerospace 11:9 (2024) 703. doi: 10.3390/aerospace11090703.
[19] G. Li, L. Zhang, G. Xia, S. Li, Structural Reliability Analysis of Aero-Engine Turbine Components Based on Particle Swarm Optimization Back Propagation Neural Network, Applied Sciences 15:6 (2025) 3160. doi: 10.3390/app15063160.
[20] Y.-P. Zhao, G. Huang, Q.-K. Hu, B. Li, An improved weighted one class support vector machine for turboshaft engine fault detection, Engineering Applications of Artificial Intelligence 94 (2020) 103796. doi: 10.1016/j.engappai.2020.103796.
[21] J. Grehan, D. Ignatyev, A. Zolotas, Fault Detection in Aircraft Flight Control Actuators Using Support Vector Machines, Machines 11:2 (2023) 211. doi: 10.3390/machines11020211.
[22] S. Vladov, A. Banasik, A. Sachenko, W. M. Kempa, V. Sokurenko, O. Muzychuk, P. Pikiewicz, A. Molga, V. Vysotska, Intelligent Method of Identifying the Nonlinear Dynamic Model for Helicopter Turboshaft Engines, Sensors 24:19 (2024) 6488. doi: 10.3390/s24196488
[23] L. Shang, H. Wang, H. Si, Y. Wang, T. Pan, H. Liu, Y. Li, Flight Trainee Performance Evaluation Using Gradient Boosting Decision Tree, Particle Swarm Optimization, and Convolutional Neural Network (GBDT-PSO-CNN) in Simulated Flights, Aerospace 11:5 (2024) 343. doi: 10.3390/aerospace11050343.
[24] M. Wang, H. Wang, L. Cui, G. Xiang, X. Han, Q. Zhang, J. Chen, Remaining Useful Life Prediction for Aero-Engines Based on Time-Series Decomposition Modeling and Similarity Comparisons, Aerospace 9:10 (2022) 609. doi: 10.3390/aerospace9100609.
[25] T. Berghout, L.-H. Mouss, O. Kadri, L. Saïdi, M. Benbouzid, Aircraft Engines Remaining Useful Life Prediction with an Improved Online Sequential Extreme Learning Machine, Applied Sciences 10:3 (2020) 1062. doi: 10.3390/app10031062.
[26] C. Raikar, R. Ganguli, Denoising Signals Used in Gas Turbine Diagnostics with Ant Colony Optimized Weighted Recursive Median Filters, INAE Letters 2:3 (2017) 133–143. doi: 10.1007/s41403-017-0023-y.
[27] S. J. Mohammadi, S. A. M. Fashandi, S. Jafari, T. Nikolaidis, A scientometric analysis and critical review of gas turbine aero-engines control: From Whittle engine to more-electric propulsion, Measurement and control 54:5–6 (2021) 935–966. doi: 10.1177/0020294020956675.
[28] X. Zhang, H. Liang, J. Feng, H. Tan, Kalman Filter Based High Precision Temperature Data Processing Method, Frontiers in Energy Research 10 (2022). doi: 10.3389/fenrg.2022.832346.
[29] C. Hu, K. Miao, M. Zhou, Y. Shen, J. Sun, Intelligent Performance Degradation Prediction of Light-Duty Gas Turbine Engine Based on Limited Data, Symmetry 17:2 (2025) 277. doi: 10.3390/sym17020277.
[30] F. Lu, Y. Wang, J. Huang, Y. Huang, Gas Turbine Transient Performance Tracking Using Data Fusion Based on an Adaptive Particle Filter, Energies 8:12 (2015) 13911–13927. doi: 10.3390/en81212403.
[31] H. Li, Y. Guo, X. Ren, New Method for Improving Tracking Accuracy of Aero-Engine On-Board Model Based on Separability Index and Reverse Searching, Aerospace 12:3 (2025) 175. doi: 10.3390/aerospace12030175.
[32] X. Dang, Y. Shao, H. Liu, Z. Yang, M. Zhong, M. Sun, W. Deng, Risk Analysis Method of Aviation Critical System Based on Bayesian Networks and Empirical Information Fusion, Electronics 14:12 (2025) 2496. doi: 10.3390/electronics14122496.
[33] G. Xiong, X. Tan, G. Cao, X. Hong, X. Lu, J. Zhu, Ensemble Modeling Method for Aero-Engines Based on Automatic Neural Network Architecture Search Under Sparse Data. Aerospace 12:9 (2025) 804. doi: 10.3390/aerospace12090804
[34] X. Liu, H. Tang, X. Zhang, M. Chen, Gaussian Process Model-Based Performance Uncertainty Quantification of a Typical Turboshaft Engine, Applied Sciences 11:18 (2021) 8333. doi: 10.3390/app11188333.

[35] W. Gao, M. Pan, W. Zhou, F. Lu, J.-Q. Huang, Aero-Engine Modeling and Control Method with Model-Based Deep Reinforcement Learning, Aerospace 10:3 (2023) 209. doi: 10.3390/aerospace10030209.

[36] K. V. Santhosh, B. K. Roy, An intelligent temperature measurement technique using J type thermocouple with an optimal neural network, Sensors and Transducers, 147:12 (2012), 6–14.

[37] S. Vladov, M. Petchenko, Y. Shmelov, S. Drozdova, R. Yakovliev, Helicopters Turboshaft Engines Parameters Identification at Flight Modes Using Neural Networks, In Proceedings of the 2022 IEEE 17th International Conference on Computer Sciences and Information Technologies (CSIT), Lviv, Ukraine, 10–12 November 2022, pp. 5–8. doi: 10.1109/csit56902.2022.10000444.

[38] S. Vladov, Y. Shmelov, R. Yakovliev, Method for Forecasting of Helicopters Aircraft Engines Technical State in Flight Modes Using Neural Networks, CEUR Workshop Proceedings 3171 (2022) 974–985. URL: https://ceur-ws.org/Vol-3171/paper70.pdf

[39] C. M. Stefanovic, A. G. Armada, X. Costa-Perez, Second Order Statistics of Fisher-Snedecor Distribution and Their Application to Burst Error Rate Analysis of Multi-Hop Communications, IEEE Open Journal of the Communications Society 3 (2022) 2407–2424. doi: 10.1109/ojcoms.2022.3224835.

[40] H.-Y. Kim, Statistical notes for clinical researchers: Chi-squared test and Fisher's exact test, Restorative Dentistry & Endodontics 42:2 (2017) 152. doi: 10.5395/rde.2017.42.2.152.

[41] N. Shakhovska, V. Yakovyna, N. Kryvinska, An improved software defect prediction algorithm using self-organizing maps combined with hierarchical clustering and data preprocessing. Lecture Notes in Computer Science 12391 (2020) 414–424. doi: 10.1007/978-3-030-59003-1_27.