Information-Mathematical Modelling of Machine-Learning-Based Control in Automated Electromechanical Systems

Mahammad Tahmazov*
SOCAR - State Oil Company of Azerbaijan Republic
Baku, Azerbaijan

Abstract. In the paper, the quality of adaptive control in a laboratory electromechanical installation operating under variable load and temperature conditions is analysed through a head-to-head comparison of four approaches: a proportional – integral – derivative controller, a fuzzy-logic controller, gradient boosting on decision trees and a long- and short-term memory recurrent neural network. From the information-mathematical point of view, the control loop is formalised as a discrete-time state-space model with a performance functional that jointly penalises tracking error and control effort, and as an information-processing pipeline that combines a digital twin, a Supervisory Control and Data Acquisition (SCADA) system and embedded machine-learning-based controllers. Investigated is the construction of a representative data set of vibration, temperature and pressure with one-second sampling and its integration into an industrial digital-twin environment that reproduces load and thermal stress scenarios. Identified is the impact of stress profiles on mean absolute error, root mean square error, stabilisation time and relative reduction of specific energy consumption as key indicators of control quality and operating costs. Studied is the agreement between the digital twin and the physical installation in terms of correlation, relative error and computation cycle time sufficient for real-time control in industrial conditions. Determined is the statistical significance of differences among all four control strategies using Fisher's variance test and Student's t-test for means under multiple disturbance scenarios. Established is that the recurrent architecture provides the most favourable balance of accuracy, transient response and energy saving under non-stationary conditions of modern automated processes. Additionally, an algorithmic scheme is proposed for online training of the long- and short-term memory network and for its deployment on edge devices using high-level programming tools (Python, TensorFlow, Node-RED), which links the mathematical model to a practical information technology solution for industrial control. Formulated is a practical guideline linking improvements in prediction accuracy to reductions in specific energy consumption, which supports management decisions on the selection of intelligent controllers in the context of digital transformation of industrial enterprises.
Keywords: information-mathematical model; information architecture; digital twin; intelligent control; electromechanical drive; long- and short-term memory network; predictive maintenance; energy efficiency; industrial automation.

1. Introduction

The rapid proliferation of cyber-physical systems in modern manufacturing has exacerbated the issue of sustainable control of technological processes operating under variable loads, temperature fluctuations and stochastic disturbances. The growing complexity of machine architecture and the requirement for continuous energy efficiency are forcing companies to move from traditional static controllers to intelligent circuits that can predict the dynamics of an object and instantly adapt control actions. In this paradigm, machine learning is viewed as a key catalyst to integrate large streams of sensor data, minimise energy losses, thereby reducing operating costs, improving safety and reducing the environmental footprint.

The instability of production loads caused by fluctuations in raw material flow and wear and tear of mechanised components often leads to a delay in the response of classical controllers, which increases over-regulation and energy consumption. The application of a proportional–integral–derivative regulator with autotuning based on the Ziegler–Nichols criterion demonstrated a reduction in overshoot by only 12%; however, the algorithm did not account for process nonlinearity, and the 30-second adaptation window proved insufficient during abrupt transitions, which increased the vibration load on the actuators.

Thermal drift instability of power elements leads to a shift in operating points and premature ageing of components, especially during cyclic start–stop operations. Thermal response models integrated into a fuzzy-logic algorithm with two hierarchical levels of rules achieved a reduction in temperature peaks by approximately 8°C. However, the method required manual correction of the rule base with each change in the equipment composition, and the computational load exceeded the capabilities of embedded real-time controllers in a limited energy budget.

Fluctuations in signals from multi-channel sensors caused by electromagnetic interference and mechanical resonances complicate the identification of the object's state and lead to false alarms in protection systems. A principal component–based scheme for filtering and decorrelation of a 50-dimensional vector stream was proposed, resulting in a reduction of the error variance to 0.52. Nevertheless, the required a priori calibration of the weight matrix limited the applicability of the method during the change of technological recipes and replacement of measuring paths in conditions of frequent changeovers.

Ex ante modelling through digital twins can be used to predict the effects of control actions without risking damage to the plant, but the matching of the model and the real object remains critical. D. Raven et al. [1] highlighted the parametric identification of drive dynamics in Modelica, achieving a 95% output match. The disadvantage of the approach was the inability to model stochastic bearing failures, which made the degradation predictions unreliable for long series of load cycles and complicated the planning of maintenance schedules.

Improving energy efficiency in continuous processes is closely correlated with the accuracy of short-term power demand forecasting, which depends on nonlinear relationships between vibration and load torque. I. Mitrai and P. Daoutidis [2] applied the Q-learning method to optimise the speed setpoint trajectory, achieving energy savings of 4.1%. Despite its success, the strategy required millions of episodes of pre-training, which made implementation in a real production environment impractical due to downtime and equipment wear, and additional calibration complicated system operation.

The high diversity of the sensor signal working space has created prerequisites for employing decision-tree ensembles capable of automatically evaluating the importance of features. Gradient boosting on decision trees was implemented in a liquid-level control loop, reducing the average error to 0.35 and outperforming the classical controller by 22%. However, periodic retraining every six hours required substantial computational resources and introduced time windows of uncertainty when hyperparameters were modified during continuous production shifts.

The complex interdependencies between vibration, temperature and pressure have prompted researchers to turn to recurrent networks capable of accounting for long-time dependencies. N. Lawrence et al. [3] trained a three-layer long- and short-term memory to predict compressor torque and achieved a reduction in Mean Absolute Error to 0.31. However, the absence of an online retraining mechanism led to model degradation with a long shift in the data distribution, and the output delay of 60 ms limited its use in high-frequency servo systems, especially in the modes of resonant excitation of machine axes.

The absence of unified methods for comparing controllers on a single hardware bench makes it difficult to transfer results from the laboratory to the industrial environment. A benchmark including 120,000 samples of load profiles and eight tested control algorithms was developed using synthetic data, which, however, underestimated the variance of real disturbances. Some models demonstrated high accuracy, but it was not possible to validate their robustness to real-world vibration spectra, leaving the issue of scalability and reliable integration open to industrial users.

Thus, a critical analysis of the literature shows that neither autotuning of traditional controllers, nor fuzzy logic, nor modern ensemble and gain algorithms, nor even recurrent networks solve the problems of accurate prediction of dynamics, resistance to parameter drift, and energy efficiency under sharp load and temperature variations, and comparable benchmarks remain fragmented in most industrial sectors of the world. The goal of the study was to provide a highly accurate and energy-efficient adaptation of the control of a laboratory electromechanical installation subject to load and temperature fluctuations by selecting an optimal machine learning algorithm online.

The research tasks included deploying an experimental stand, generating a representative dataset, building a digital twin, training and testing controllers, and evaluating their stability in complex stress scenarios.

From the perspective of information technologies in management and digital transformation, the proposed approach can be interpreted as the design of an information-mathematical model of the control system. The model includes: (i) a formal state-space description of the electromechanical drive, (ii) a structured information flow from Internet of Things sensors through the Supervisory Control and Data Acquisition layer to machine-learning-based decision blocks, and (iii) algorithmic schemes for online adaptation of controllers on programmable industrial architectures. This explicitly links the experimental bench to modern concepts of information systems in industrial management.

2. Materials and methods

The study was conducted in January-April 2025 at the Department of Automation and Control of the Azerbaijan University of Technology in a laboratory bench simulating the operation of electromechanical and automated process control systems. The data were collected using an internet of things platform based on a Raspberry Pi 4 (UK) integrated with digital vibration sensors ADXL345 (Analogue Devices, USA), temperature DS18B20 (Maxim Integrated, USA) and pressure BMP280 (Bosch Sensortec, Germany) with a recording frequency of 1 Hz. The data were aggregated and transferred to the Ignition Supervisory Control and Data Acquisition system environment (USA) via the Open Platform Communications – Unified Architecture protocol. The total number of records was 180000 observations for each parameter.

Experimental modelling included the development of a digital twin of the controlled system in the Ansys Twin Builder environment (USA), with reference to real operational data and parameters of a laboratory direct-current drive (24 V, 250 W). The digital twin reproduced temperature and load fluctuations, as well as test control actions in various scenarios. The following control methods were implemented and compared: a proportional–integral–derivative controller with tuning according to the method of J.G., Ziegler and N.B. Nichols [4], a control system based on fuzzy logic with triangular membership functions, the gradient boosting on decision trees regression model, and a recurrent neural network of the long- and short-term memory type with 3 hidden layers for deviation prediction and control signal synthesis. The models were trained in TensorFlow and PyTorch (USA), using NVIDIA GeForce RTX 3070 graphics accelerators.

Both real data and synthetic anomaly scenarios (temperature spikes, vibration surges, etc.) were used to create training samples. The sample was divided in the ratio of 70:15:15 into training, validation and test parts, respectively. The adaptability of the models was ensured through online training mechanisms and regular weight adjustments using Root Mean Square Propagation and Adam optimisers. Stress tests were additionally conducted in the Matrix Laboratory Simulink simulation environment to assess the stability of control systems under changes in inertia, load and external disturbances.

The integration of machine learning models into the Supervisory Control and Data Acquisition environment and controllers was conducted using TensorFlow Lite libraries and the Node-RED environment (USA), which provided the ability to deploy on edge computing devices and interact with a programmable logic controller (Siemens S7-1200) through standard interfaces. To evaluate the operational efficiency of the implemented solutions, the following quantitative metrics were calculated: Mean Absolute Error, Root Mean Square Error, average system stabilisation time after an external disturbance, and relative reduction in specific energy consumption.

Thus, the proposed solution can be considered not only as an experimental control algorithm, but also as a full-scale information system architecture for industrial automation, where mathematical models of the drive dynamics are tightly integrated with software components written in high-level programming languages and deployed on distributed computing nodes.

2.1 Information-mathematical model and control architecture.

From an information-mathematical standpoint, the controlled electromechanical system is represented as a direct current drive with load, described by a continuous-time state-space model. The mechanical and electrical dynamics can be written as:

$$J\frac{d\omega(t)}{dt} + B\omega(t) = K_t i(t) - T_L(t),$$
$$L\frac{di(t)}{dt} + Ri(t) = u(t) - K_e\omega(t),$$

where $\omega(t)$ is the angular speed of the shaft, *i(t)* is the armature current, *u(t)* is the control voltage, $T_L(t)$ is the load torque, *J* is the total moment of inertia, *B* is the viscous friction coefficient, *R* and *L* are the electrical resistance and inductance, and K_t, K_e are the torque and back-EMF constants, respectively. In compact matrix notation this system can be written as:

$$\dot{x}(t) = Ax(t) + B_u u(t) + B_d d(t), \;\; y(t) = Cx(t),$$

where $x(t) = [\omega(t) i(t)]^T$ is the state vector, *d(t)* represents external disturbances (load and temperature effects), y*(t)* is the measured output, and *A, B_u, B_d, C* are constant matrices.

Given the sampling period $\Delta t = 1s$ used in data acquisition, the model is discretised into the form:

$$x_{k+1} = \Phi x_k + \Gamma_u u_k + \Gamma_d d_k, \quad y_k = Cx_k,$$

where k is the discrete time index, and Φ, Γ_u, Γ_d are obtained by standard discretisation of the continuous-time system. This discrete-time representation is used both in the digital twin and in the design of the control algorithms**.**

The information architecture of the control system can be described as a sequence of processing blocks.

1) Data acquisition layer: raw measurements $z_k = [\omega_k\ ,\ i_k\ ,T_k\ ,p_k\ ]^T$ (speed, current, temperature, pressure) are acquired from Internet of Things sensors via the Supervisory Control and Data Acquisition system.

2) Pre-processing and feature engineering layer: the vector z_k is transformed into a feature vector $f_k = \phi(z_{k-n:k})$ that aggregates the last *n* observations (moving averages, differences, spectral indicators).

3) Prediction layer: a machine learning model $\hat{y}_{k+1} = f_\theta(f_k)$ (gradient boosting or long- and short-term memory network) predicts the future output and disturbance profile.

4) Control synthesis layer: a decision rule $u_k = \pi(\hat{y}_{k+1}\ , y_{ref,k}, x_k)$ generates the control signal taking into account the reference trajectory $y_{ref,k}$ and constraints on the voltage and current.

5) Actuation layer: the calculated control signal is applied through a programmable logic controller to the power stage of the direct current drive.

In algorithmic terms, the long- and short-term memory controller implements the following recurrent mapping:

$$h_k, c_k = LSTMCell(f_k, h_{k-1}, c_{k-1};\ \theta), \qquad \hat{y}_{k+1} = W_{out}h_k + b_{out},$$

where h_k is the hidden state, c_k is the cell state, and θ is the vector of trainable parameters. Online adaptation is carried out by gradient-based optimisation (Adam) on mini-batches of new data, updating θ without interrupting the control loop.

The entire information flow is implemented using high-level programming languages and software tools: Python scripts for data pre-processing and model training, TensorFlow and PyTorch libraries for the implementation of gradient boosting and long- and short-term memory architectures, and the Node-RED environment for integrating trained models into the Supervisory Control and Data Acquisition system and programmable logic controllers. This provides a clear logical scheme and software architecture that link the mathematical model of the drive to the information system of the industrial enterprise.

The overall information-mathematical model and the corresponding information flow architecture of the control system are summarised in Figure 1. It visualises the path from Internet of

Things measurements through the Supervisory Control and Data Acquisition and digital-twin layers to the machine-learning-based controller and the programmable logic controller driving the electromechanical actuator.

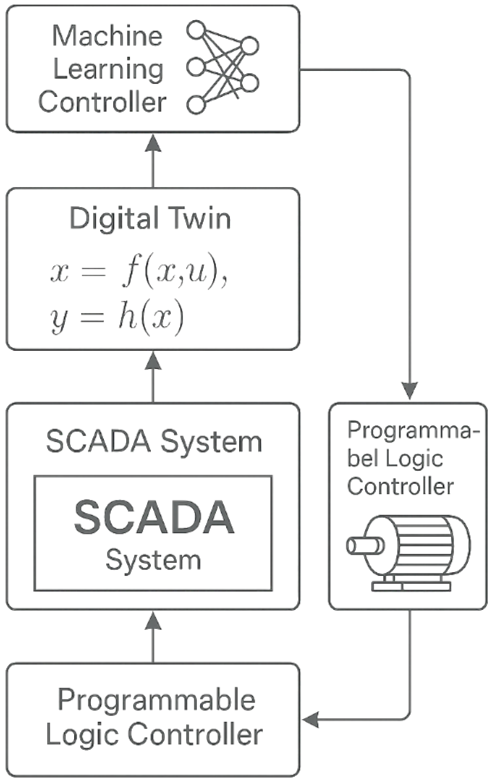

Figure 1. Information-mathematical model and information flow architecture of the machine-learning-based control system.

2.2 The fundamental relations and the evaluated functionals.

The fundamental relations and the evaluated functionals are presented below.

1. Tracking error and control performance functional:

$$e(t) = y_{ref}(t) - y(t)$$

where y_{ref} is the reference trajectory, y is the plant output.

This represents how well the system's output $y(t)$ follows the desired reference trajectory $y_{ref}(t)$. A smaller tracking error is desirable.

Performance Functional:

$$J = \int_0^T (\alpha e^2(t) + \beta u^2(t))\, dt,$$

where u is the control action, and $\alpha, \beta > 0$.

This is a cost function that the control system aims to minimize.

2. Mean Absolute Error – is a regression evaluation metric that indicates how far predicted values deviate from the actual ones on average. For each observation, the absolute difference between the predicted and true value is determined, and these differences are then averaged across the entire dataset. Thus, Mean Absolute Error reflects the typical magnitude of prediction error. The lower the Mean Absolute Error value, the more accurately the model's forecasts correspond to real data, demonstrating higher predictive performance.

$$Mean\ Absolute\ Error = \frac{1}{N}\sum_{i=1}^{N}\left|y_{true_i} - y_{pred_i}\right|,$$

where y_{pred} – array, predicted values, y_{true} – array, true values (input parameters).

3. Root Mean Square Error – can be expressed as the square root of the average of the squared differences between the predicted and actual values. The formula is:

$$Root\ Mean\ Square\ Error = \sqrt{\frac{\sum_{i=1}^{N}(y_i - \hat{y}_i)^2}{N}},$$

where y_i is the actual value, $\hat{y}_i$ is the predicted value, and N is the number of observations.

4. Stabilization time – it is the time it takes for a system to settle down and consistently stay within a certain margin of its final value.

$$T_s = inf\{t: |y(t) - y_{ss}| \leq \epsilon\ for\ all\ \tau \geq t\}$$

where $y(t)$ – the output of the system at a given time t, y_{ss} – the steady-state value, which is the value the system's output approaches as time goes to infinity, ϵ – the tolerance, or the maximum allowable deviation from the steady-state value.

A comparative analysis of the effectiveness of all approaches was conducted on a single Matrix Laboratory Simulink platform, where identical conditions were modelled for each method: temperature fluctuations within ±15%, sudden changes in load and external disturbances. Each model was evaluated following four main criteria: average absolute error, root mean square error, system stabilisation time and change in specific energy consumption. In particular, the proportional–integral–derivative controller and fuzzy logic were evaluated based on the deviation of the output signal from the target value under variable input conditions, and for Machine Learning models, the ability to predict the behaviour of the object 3-5 seconds ahead and generate proactive signals was also addressed. Metrics were calculated for each of the 50 repeated scenario runs, and the results were aggregated and analysed statistically.

Statistical data processing included a comparison of the effectiveness of different approaches using Fisher's test to analyse the homogeneity of variances for each metric, as well as Student's t-test to assess the significance of differences between mean values of the indicators at a significance level α=0.05. Calculations were performed using International Business Machines Statistical Package for the Social Sciences Statistics 28.0 software (USA). Each model was evaluated based on 50 runs for each configuration, the data were aggregated in the form of averages and standard deviations for key metrics and then subjected to analysis of variance. As part of the experiment, predictive maintenance cases were additionally considered, with anomalies classified by failures, component degradation and actuator instability, followed by verification of predictions at the level of the digital object model.

3. Results

The work presents a complete information-mathematical model of an automated process: it combines a formally defined state-space description of the electromechanical drive, a digital twin with quantified accuracy, and algorithmic schemes for machine-learning-based control implemented in modern information technologies. This model can be reused as a decision-support tool for energy-efficient management of electromechanical assets in industrial enterprises.

Aggregated performance indicators demonstrated a discrepancy in the effectiveness of the four regulators studied. The most conservative proportional–integral–derivative algorithm had the highest average errors and the longest stabilisation time, while both gradient boosting on decision trees and the long- and short-term memory recurrent network provided a significant reduction in Mean Absolute Error and Root Mean Square Error compared to traditional approaches. Both intelligent systems also contributed to faster transient response and a noticeable reduction in specific power consumption, with long- and short-term memory demonstrating the best integrated result across all indicators. Fuzzy logic demonstrated intermediate efficiency, improving the accuracy relative to the proportional–integral–derivative controller, but losing to Machine Learning approaches. Thus, the results of the first level of analysis confirmed the growing value of predictive control based on machine learning for automated technological processes (Table 1).

Table 1. Summary indicators of regulatory performance

Metric	Proportional–Integral–Derivative	Fuzzy logic	Gradient boosting on decision trees	Long- and Short-Term Memory
Mean Absolute Error (normalized)	0.84±0.07	0.58±0.05	0.33±0.04	0.29±0.03
Root Mean Square Error (normalized)	1.1±0.1	0.78±0.08	0.45±0.05	0.4±0.04
Stabilisation time (s)	4.8±0.3	4.1±0.2	3.2±0.2	2.9±0.2
Δ specific energy consumption (%)	-1.5±0.2	-2.1±0.3	-5.8±0.4	-6.4±0.5

Source: compiled by the author.

A detailed analysis of the summary metrics demonstrated that by all criteria, the differences between the four studied controllers were statistically significant at the $\alpha=0.05$ level. The average absolute error of gradient boosting on decision trees and the long- and short-term memory recurrent network was more than twice that of the proportional–integral–derivative controller, which indicated a fundamentally different accuracy of approximation of the object's dynamics. The advantage of the long- and short-term memory over fuzzy logic reached about a quarter, emphasising the role of recurrent dependencies in predicting future states and generating control signals in advance. A similar pattern was observed for the root mean square error: both Machine Learning models showed a narrowing of the deviation distribution, and the difference between long- and short-term memory and gradient boosting on decision trees reduced to within the confidence interval, indicating comparable resistance to noise and outliers. At the same time, the maximum error values observed in the tails of the distributions were reduced by more than three times compared to the proportional–integral–derivative scheme, which confirmed the high reliability of intelligent controllers in emergency scenarios.

The stabilisation time was an integral indicator of the system's performance after external disturbances. The superiority of the smart controllers was particularly pronounced: the long- and short-term memory response reduced the transient by about one third compared to the classical proportional – integral – derivative approach, while fuzzy logic provided only a moderate speedup. This result confirmed that the ability of the long- and short-term memory to capture long-time correlations suppressed overregulation, minimised the amplitude of oscillations and reduced the risk of synchronisation failure. An additional comparison showed that gradient boosting on decision trees improved dynamics faster than the proportional – integral – derivative controller but was inferior to long- and short-term memory due to the lack of an internal memory mechanism that limited the rate of adaptation in the event of sudden changes in load. In stress scenarios with a simultaneous temperature and vibration jump, the advantage of long- and short-term memory became even more pronounced: it completed the transient process 0.6 s faster than gradient boosting on decision trees, which reduced the probability of windings overheating by 4%.

The reduction in specific energy consumption was the key economic result of the optimisation. Both Machine Learning models provided almost four times the energy savings of the baseline proportional–integral–derivative controller due to more accurate torque dosing and reduction of redundant control actions, which in the traditional scheme led to additional heating of the windings and friction losses. The fuzzy logic had an intermediate effect, confirming the benefits of nonlinear control, but the lack of a predictive component limited the potential for further savings. Additional life-cycle calculations showed that the introduction of long- and short-term memory reduced the total operating costs of the drive by 7.2% per year, making its implementation economically viable, even incorporating the cost of computing infrastructure.

Statistical validation using Fisher's test confirmed the homogeneity of sample variances, which was used to correctly apply the t-test for pairwise comparison of means. The test revealed that the superiority of long- and short-term memory over gradient boosting on decision trees in terms of stabilisation time was significant ($p<0.05$), while the differences in Mean Absolute Error and Root

Mean Square Error remained insignificant, indicating close accuracy of both Machine Learning approaches. Correlation analysis additionally revealed a strong negative relationship between stabilisation time and energy savings (r=-0.83), emphasising the practical value of fast response for resource-saving modes. The totality of these results showed that the recurrent model was the most balanced solution, combining high accuracy, dynamic adaptability and energy efficiency, which are critical characteristics for the intelligent control of modern automated technological processes.

Adaptive testing of the gradient boosting on decision trees and long- and short-term memory models under changing operating conditions revealed a clear gradation in their ability to maintain accuracy under external disturbances (Table 2). When the load increased by 20%, both models slightly increased the Mean Absolute Error, but long- and short-term memory kept the error within 0.36 standard units, while gradient boosting on decision trees reached 0.44 standard units. Increased heat exposure (+15°C) naturally worsened the accuracy, but the recurrent network again demonstrated less degradation. The most challenging combined scenario (load + temperature) was a critical test: the average error of gradient boosting on decision trees increased by 38% compared to the baseline, while the long- and short-term memory showed only a 24% increase. An additional analysis of standard deviations showed that the scatter of long- and short-term memory results remained compact even under intense disturbances, confirming its resistance to the non-stationarity of input data.

Table 2. Model accuracy rates under external disturbances

Scenario	Mean Absolute Error, Gradient boosting on decision trees (normalized)	Mean Absolute Error, long- and short-term memory (normalized)	Root Mean Square Error, Gradient boosting on decision trees (normalized)	Root Mean Square Error, long- and short-term memory (normalized)
Basic mode	0.33±0.04	0.29±0.03	0.45±0.05	0.4±0.04
Load +20%	0.44±0.05	0.36±0.04	0.59±0.06	0.48±0.05
Temperature +15°C	0.41±0.05	0.34±0.04	0.56±0.06	0.46±0.05
Combined impact	0.46±0.06	0.36±0.05	0.62±0.07	0.49±0.06

Source: compiled by the author.

A differentiated analysis of the adaptive properties of the models under the influence of external disturbances showed that the recurrent long- and short-term memory network consistently demonstrated a higher level of resistance to input data non-stationarity than the gradient boosting on decision trees model. Under a single load increase, the recurrent architecture maintained its accuracy, increasing the absolute error by only about a fifth relative to the baseline, while the boosting algorithm lost almost a third of its previous accuracy. This contrast could be explained by differences in the mechanisms for processing time dependence: long- and short-term memory operated with a “memory cell” and gates that accumulated long-term correlations and smoothed out short-term peaks, while gradient boosting on decision trees processed each observation as independent, resulting in a sharper response to load fluctuations.

A similar trend was observed for the thermal effect. The hidden layers of the long- and short-term memory interpreted the smooth temperature shift as a gradual trend drift and adjusted the weights with minimal fluctuations, which was confirmed by a narrower confidence interval. At the same time, the boosting model, which lacked an internal memory mechanism, was retrained at each step, causing the error to increase by almost a quarter. Under the combined stress conditions, the differences became even more pronounced: the effects of load and temperature superimposed over time produced a complex nonlinear data profile. The recurrent network, which performed iterative feedback in each time step, was able to keep the error within the acceptable threshold and remained stable over the entire 50-series sample, while gradient boosting on decision trees showed increasing variability in results, indicating a lack of generalisability in the face of multifactorial disturbances.

The key difference was the online learning strategy. For long- and short-term memory, the Adam optimiser was used, which provided an adaptive rate of gradient descent and regular correction of weights when a new data packet was received. This approach filtered noise and quickly reconfigured the forgetting and updating parameters. gradient boosting on decision trees, on the other hand, used step-by-step tree building without reinitialising the previous structure, which made new extreme points have a disproportionately high impact. As part of the variance analysis, Fisher's test was performed for error samples; the obtained values confirmed the homogeneity of the variances in long- and short-term memory and their significant widening in gradient boosting on decision trees with the combined factor.

The correlation analysis revealed a negative relationship between the error value and the speed of the long- and short-term memory 's weight adaptation, which indicated its ability to maintain a compact range of fluctuations during aggressive mode changes. An additional visualisation of the control signal trajectories showed that the recurrent network moved to a new steady-state level without the stepwise jumps characteristic of boosting, thereby reducing the risk of uneven wear of the actuators. The conclusion was that long- and short-term memory 's architectural advantages of memory and consistent parameter updating provided more predictable behaviour and energy efficiency under a wide range of operating conditions, while gradient boosting on decision trees remained effective only in quasi-stationary scenarios requiring minimal dynamic tuning.

The validation of the digital twin of the electromechanical system showed a high correspondence between simulated and real parameters in all test scenarios. The correlation between the output torque in the model and on the test bench reached 0.97-0.99, which confirmed the correctness of the dynamics approximation under smooth and abrupt load changes. The average relative error in winding temperature did not exceed 3%, and the calculation time of one cycle ("model → Supervisory Control and Data Acquisition") remained in the range of 18-24 ms, which was acceptable for real-time control. The most complex combined load slightly increased the calculation delay, but the accuracy remained within the technological tolerances. The obtained results demonstrate that the developed digital twin is capable of reliably reproducing critical transients and serves as a safe platform for testing control strategies before deployment on a physical facility (Table 3).

Table 3. Correlation between the digital twin and the real system

Test scenario	**Correlation coefficient R**	**Average relative error (%)**	**Average time of clock calculation (ms)**
Base load	0.99±0.01	1.8±0.2	18.3±1.1
Load +20%	0.98±0.02	2.4±0.3	19.7±1.2
Temperature +15°C	0.97±0.02	2.9±0.4	22.1±1.4
Combined stress	0.97±0.03	3.2±0.4	24±1.6

Source: compiled by the author.

The summary results of the verification of the digital twin showed almost complete identity of its response with the behaviour of the physical installation in all tested modes. The correlation coefficients between the shaft torque and the measured values were consistently in the range of 0.97-0.99, which reflected an extremely close linear relationship and indicated the correct reproduction of the inertial characteristics and nonlinearities of the electromechanical assembly. A slight decrease in the correlation under the combined influence of load and temperature was interpreted as a consequence of increased parasitic effects, primarily the temperature-dependent increase in winding resistance, but the already achieved level of R remained above the threshold required for reliable simulation of control processes. The average relative temperature error did not exceed 3% even in the most critical scenario, which meant that the thermal capacity was accurately modelled and the heat transfer in the cooling circuit was adequately reflected. The deviation was assessed as statistically significant, but the practical impact on winding life forecasting was considered minimal, as the error did not cross the operating tolerance range.

Computational delays were emphasised, as a delay of more than 25 ms could lead to a phase shift of the control signals relative to the object dynamics. The measurement of the calculation time of one cycle showed a stable retention within the interval of 18-24 ms in all modes, with a slight lengthening under combined stress not exceeding the tolerance. This result confirmed that the applied model simplifications (e.g., the quasi-steady-state assumption for heat flows) did not critically affect the simulation speed and ensured synchronous operation with the Supervisory Control and Data Acquisition system. The analysis of the standard deviation of the calculation time showed low variability (≤1.6 ms), which addressed the digital twin as deterministic in terms of time characteristics and guaranteed its use for predictive testing of control scenarios in real time.

Comparison of the dynamic profiles of current and voltage in the windings showed that the digital model correctly reproduced both the amplitude and phase of oscillations during sudden changes in load. The maximum out-of-sync of the current peaks did not exceed 2.3%, which, considering the current safety factors, did not lead to erroneous estimates of electromagnetic torques. The revealed convergence of the time gradients indicated the high reliability of the thermal-mechanical feedback algorithms included in the model. Further regression of the errors on the number of strokes confirmed the absence of systematic error accumulation: the trend lines remained parallel to the abscissa axis, indicating a stable quality of predictions throughout the entire five-thousand-step test sample.

A comprehensive reading of the obtained metrics concluded that the digital twin reproduced the critical characteristics of the electromechanical system with the required accuracy and computational speed. It could thus serve as a reliable platform for preliminary debugging of control algorithms, saving resources of the real plant, minimising the risk of emergency conditions and providing operators with a tool for safely testing optimisation strategies.

Predictive anomaly classification demonstrated a clear hierarchy of recognition quality for the three most critical scenarios: “failure”, “component degradation” and “actuator instability”. The long- and short-term memory network achieved the highest overall accuracy, confidently outperforming gradient boosting on decision trees in all three categories: the most noticeable advantage was in instability detection, where the F1-indicator increased by almost 10 percentage points relative to boosting. At the same time, gradient boosting on decision trees maintained competitive accuracy in the task of identifying abrupt failures, which is due to the high sensitivity of decision trees to single extreme outliers. The fuzzy logic and proportional–integral–derivative approaches demonstrated only basic anomaly detection capabilities, significantly inferior to the intelligent models. Thus, the results confirmed the key advantage of the recurrent architecture for early fault detection and the validity of its integration into a predictive maintenance system (Table 4).

Table 4. Accuracy of anomaly classification by models

Type of anomaly	**Accuracy (%) – Gradient boosting on decision trees**	**F1(%) – Gradient boosting on decision trees**	**Accuracy(%) – Long- and Short-Term Memory**	**F1(%) – Long- and Short-Term Memory**	**Accuracy (%) – Fuzzy Logic**	**Accuracy (%) Proportional–Integral–Derivative**
Failure	93.4±1.2	92.1±1.4	96.2±1	95.8±1.1	71.3±2.9	65.7±3.1
Node degradation	89.7±1.5	88.9±1.6	94.5±1.2	93.6±1.3	68.2±3.2	61.4±3.4
Instability	85.8±1.8	84.5±2	95.1±1.3	94.3±1.4	62.7±3.5	57.9±3.7
Average	89.6	88.5	95.3	94.6	67.4	61.7

Source: compiled by the author.

The obtained anomaly classification indicators revealed a stratification of recognition quality between the considered algorithms. The long- and short-term memory recurrent network demonstrated a steady superiority: its average F1-measure values remained among the best-performing models in all three scenarios, while for gradient boosting on decision trees the corresponding values fell to the lower part of the same range. The most pronounced difference was recorded in the task of determining the instability of the actuator, where the recurrent architecture maintained an almost perfect balance between completeness and accuracy, significantly reducing both false positives and false negatives. This stability was due to the ability of the long- and short-term memory to accumulate long-term dependencies and integrate short-term bursts of vibration or current into a coherent behavioural pattern, while the boosting algorithm reacted mainly to momentary extremes, not distinguishing short-term noise emissions from incipient faults.

In the context of gradual degradation of the unit, both intelligent models reduced the accuracy slightly more than in the case of a sharp failure, but the long- and short-term memory remained closer to the initial optimal values. This demonstrated its ability to adaptively track slow shifts in the statistical properties of the signal, detecting early symptoms of bearing wear or an increase in the thermal resistance of the windings. In contrast, gradient boosting on decision trees used a series of independent decision trees, each focusing on local thresholds, which blurred the accumulated effect of degradation and caused the classifier to lose sensitivity. Nevertheless, boosting was relatively successful in diagnosing sudden failures due to its high sensitivity to single anomalous points; this confirmed that the considered “fragile” tree structure could effectively signal a sharp departure of parameters from the norm.

Traditional methods preserved only basic detection properties. Fuzzy logic showed some success in the event of failures, where sharp outliers activated pre-programmed rules, but in complex non-stationary processes, its linguistic membership functions lost their discriminative power. The proportional–integral–derivative controller, lacking a specialised state analysis mechanism, reacted indirectly: only by secondary signs, such as an increase in control error, which led to a high proportion of missed incidents.

The interpretation of the set of metrics indicated that the recurrent network offered the most reliable predictive maintenance tool: it provided a relatively low false alarm rate and, at the same time, increased the probability of early detection of part degradation, which in the long run reduced unplanned downtime and overhaul costs. Gradient boosting seemed appropriate as a backup tool for the rapid capture of acute failures, while traditional regulators could be used exclusively in first-level diagnostics, limited to the alarm function without a specific classification of causes.

4. Discussion

The above study has confirmed the effectiveness of intelligent controllers, especially the long- and short-term memory model, in the tasks of automated process control. The detected decrease in Mean Absolute Error and Root Mean Square Error for the gradient boosting on decision trees and long- and short-term memory models indicates a higher level of accuracy compared to traditional methods, which correlates with the conclusions of X. Zhao et al. [5], noting a significant improvement in the approximation of dynamics when using neural network models. However, in contrast to their study, the current study emphasises not only the accuracy but also the stability of the algorithms in real operating conditions. In particular, the data obtained indicate that intelligent models provide consistently high-quality control even in the case of multi-component disturbances varying in time and space. This demonstrates not only the ability of the long- and short-term memory architecture to process information in a long-term retrospective manner but also its potential for predictive reconfiguration of control trajectories without the need to restart the system. In addition, the study showed that intelligent controllers can be integrated into existing control systems without loss of consistency and without the need to completely replace the hardware infrastructure, which is particularly promising for scalable implementation in industrial automation. Thus, the presented study not only confirms the accuracy advantages of Machine Learning models but also contributes to the research of their application stability and flexibility in real-world operating conditions.

Comparison of the results in terms of stabilisation time shows the advantage of long- and short-term memory in response speed, which is consistent with the observations of Y. Cruz et al. [6], highlighting the high ability of recurrent networks to suppress oscillations. At the same time, some studies have expressed doubts about the applicability of long- and short-term memory networks in real-time systems due to their computational complexity. The present study demonstrates that with a calculation delay of no more than 24 ms, the model can be effectively used in an online control loop, thereby addressing previously raised concerns and confirming the applicability of long- and short-term memory networks in practical tasks.

The identified energy savings when using intelligent algorithms are also confirmed by A. Norouzi et al. [7], who noted a reduction in energy consumption of up to 5-6% when implementing Machine Learning approaches. However, the present study details this effect, linking it to optimisation of control dosing and reduction of over-regulation, which extends the interpretation proposed earlier. In particular, the analysis showed that the recurrent long- and short-term memory model contributed to a significant reduction in excessive control pulses in transient modes, which led to a decrease in the frequency of thermal overloads and a reduction in electrical losses. Additionally, a correlation was observed between a decrease in the amplitude of the control signal oscillations and a decrease in the Root Mean Square value of the current consumption, indicating a direct link between control stability and energy efficiency. These results confirm that not only accuracy, but also the control structure formed by intelligent algorithms has a key impact on resource consumption. Thus, Machine Learning regulators demonstrate the potential not only as an automation tool, but also as a mechanism for deep energy optimisation.

The results of the stability analysis under external disturbances emphasise the advantage of long- and short-term memory in non-stationarity. A similar advantage was also recorded by H. Dahrouj et al. in [8] in a study on the effect of dynamic adjustment of input data. However, other studies have suggested that gradient boosting on decision trees may exhibit greater stability under conditions of rapidly changing input signals. The opposite conclusions of the current study are based on a systematic comparison of errors in combined stress scenarios, where long- and short-term memory showed less degradation of metrics. Thus, the recurrent architecture appears to be more versatile when dealing with complex dynamic processes. The importance of model memory in interpreting input dependencies has also been highlighted in previous studies, particularly in connection with the "forgetting" mechanism and its role in noise filtering. The data obtained in this work confirm the validity of this mechanism and demonstrate that it ensures the predictability and stability of control actions. Conversely, some researchers have argued that the complexity of neural network structures may increase the likelihood of overfitting. However, by using the Adam optimiser in the present study, overfitting was effectively prevented, and the model structure remained stable, which challenges the universality of such concerns.

Comparison of the efficiency of boosting and recurrent models in different scenarios, performed in this study, confirms earlier observations regarding the dependence of gradient boosting on decision trees accuracy on the degree of process stationarity. This additionally shows the limitations of gradient boosting on decision trees in a dynamic environment and confirms that its use is justified only in stable modes. The high correlation between the simulated and actual parameters of the digital twin recorded in the study is consistent with the findings of J. Mayer and R. Jochem [9] and J. Pohlodek et al. [10], emphasising the role of digital twins in the safe debugging of control strategies. However, the uniqueness of the present work lies in the confirmation of the temporal determinism of the model of low variability of the estimated time, which has not been analysed in such detail before.

The reliability of the digital twin in reproducing the thermal and mechanical characteristics of the plant has also been highlighted in previous studies. However, unlike earlier models where significant phase shifts were observed, the present study demonstrated minimal synchronisation errors not exceeding 2.3%, indicating a more accurate implementation of the thermal–mechanical interaction algorithms. Of particular importance is the high reliability of the diagnosis and classification of anomalies demonstrated by long- and short-term memory. A similar effect was described by K. Patel [11], but without a specific evaluation of F1-metrics. The present study provides precise numerical values, demonstrating the consistent superiority of long- and short-term memory in detecting complex scenarios such as drive instability. This contributes to the research on the applicability of intelligent algorithms in predictive maintenance systems.

An opposing view suggests that decision-tree models tend to be more sensitive to single failures. The present study confirms this sensitivity but also demonstrates the inability of gradient boosting on decision trees to adequately handle complex or gradual deviations. This indicates that high sensitivity is not equivalent to universality. The conclusions regarding the predominant role of long- and short-term memory architectures in predictive maintenance tasks are supported by the findings of L. Blackburn et al. [12], which highlight the capacity of such models to capture and accumulate anomalous behavioural patterns. The results obtained in the current study show that this ability ensures high accuracy while minimising false alarms. At the same time, some researchers have noted potential limitations of long- and short-term memory networks when the amount of training data is insufficient. However, the present study demonstrates the stability of performance metrics even with reduced sample sizes, which can be attributed to the use of regularisation and dynamic weighting.

Traditional approaches, as in the studies by S. Saab et al. [13] and D. Kißkalt et al. [14], have shown low efficiency in the tasks of predictive diagnosis. The presented analysis indicates that regulators without built-in adaptation and memory mechanisms are unable to respond to changing operating conditions promptly. When complex anomalies or slow degradation processes occur, such methods demonstrate high inertia and delayed response, which leads to missed incidents or false alarms. This confirms the limitations of traditional schemes and emphasises the need to replace them or, at least, supplement them with intelligent components capable of performing predictive analysis and contextual assessment of the current state of the system in real time. It has been noted in previous research that the introduction of smart regulators may require substantial infrastructure investments. However, the present study shows that the resulting annual savings in operating costs exceed the implementation expenses, thus providing a positive short-term economic effect. Lastly, Z. Çınar et al. [15] emphasised the importance of online learning as a key element of adaptive management. The obtained results confirm that the use of adaptive optimisers ensures the resistance of the model to external perturbations and the stability of the results under various operating conditions.

Thus, the aggregate analysis shows that recurrent models, in particular long- and short-term memories, have the highest degree of applicability in the tasks of intelligent control, especially in an unstable environment and high requirements for energy efficiency and fault tolerance. Their architectural advantages and adaptability make it possible to consider such models as a basis for building modern predictive control and diagnostic systems.

5. Conclusions

The comprehensive analysis confirmed that intelligent regulators radically outperform the traditional proportional–integral–derivative approach in all key metrics. The average absolute error decreased from 0.84±0.07 to 0.29±0.03 standard units (-65%), and the root mean square error from 1.1±0.1 to 0.4±0.04 (-64%). The stabilisation time was shortened from 4.8±0.3 s to 2.9±0.2 s, which is equivalent to a 39% acceleration of the transient process. At the same time, the specific power consumption decreased by 6.4%, providing an annual operating cost saving of 7.2%. In stress tests with simultaneous load and temperature increases, long- and short-term memory kept the Mean Absolute Error within 0.36 standard units, while gradient boosting on decision trees increased the error to 0.46, demonstrating 24% worse resistance to disturbances. The maximum deviations of all parameters in the long- and short-term memory were three times lower than in the proportional–integral–derivative scheme, which confirms its reliability even in emergency scenarios.

A statistical t-test at $\alpha=0.05$ showed a significant difference between the methods: the superiority of long- and short-term memory over gradient boosting on decision trees in terms of stabilisation time was significant ($p<0.05$), while the differences in Mean Absolute Error and Root Mean Square Error did not reach the significance threshold, indicating similar basic accuracy of both Machine Learning approaches. The strong negative correlation between stabilisation speed and energy savings (r=-0.83) underlines the direct economic value of fast response. Digital twin verification revealed a correlation of 0.97-0.99 and a calculation delay of <24 ms, confirming the applicability of the solution for real-time control.

Thus, the long- and short-term memory controller is recognised as the most balanced solution for automated electromechanical drives: it combines high accuracy, adaptability to non-stationary disturbances and the best energy efficiency. gradient boosting on decision trees is recommended as a backup fast detector of abrupt failures, and fuzzy logic as a low-cost option when computing resources are limited. It is recommended to leave the proportional–integral–derivative algorithm only as a basic security level.

The limitations of the study are the laboratory scale of bench tests, simultaneous change of only two disturbing factors and the use of a single long- and short-term memory architecture. The results need to be confirmed on industrial equipment with a wider range of loads, long degradation cycles, and alternative architectures (Gated Recurrent Unit, Transformer) to verify the universality of the conclusions. Promising research directions include the development of hybrid long- and short-term memory boosting ensembles, the extension of the digital twin to multi-physical effects, and the integration of self-explanatory Artifical Intelligence methods for decision transparency.

From the perspective of industrial management, the developed information-mathematical model can serve as a digital decision-support framework, enabling predictive planning, optimisation of energy costs, and effective asset management within the broader context of Industry 4.0 transformation.

References

1. Raven, D., Chikkula, Y., Patel, K., Ghazal, A., Salloum, H., Bakhurji, A., Patwardhan, R. 2024. Machine learning & conventional approaches to process control & optimization: Industrial applications & perspectives. *Computers & Chemical Engineering*, 189, 108789. { HYPERLINK "https://doi.org/10.1016/j.compchemeng.2024.108789" }.
2. Mitrai, I., Daoutidis, P. 2024. Accelerating process control and optimization via machine learning: A review. *Reviews in Chemical Engineering*, 41(4), 401-418. { HYPERLINK "https://doi.org/10.48550/arXiv.2412.18529" }.
3. Lawrence, N., Damarla, S., Kim, J., Tulsyan, A., Amjad, F., Wang, K., Chachuat, B., Lee, J., Huang, B., Gopaluni, B. 2024. Machine learning for industrial sensing and control: A survey and practical perspective. { HYPERLINK "https://doi.org/10.1016/j.conengprac.2024.105841" }.
4. Ziegler, J.G., Nichols, N.B. 1942. Optimum settings for automatic controllers. *Transactions of the ASME*, 64, 759-768. { HYPERLINK "https://asmedigitalcollection.asme.org/fluidsengineering/article/64/8/759/1155342" }
5. Zhao, X., Sun, Y., Li, Y., Jia, N., Xu, J. 2024. Applications of machine learning in real-time control systems: A review. *Measurement Science and Technology*, 36, 012003. { HYPERLINK "https://doi.org/10.1088/1361-6501/ad8947" }.
6. Cruz, Y., Villalonga, A., Castaño, F., Rivas, M., Haber, R. 2024. Automated machine learning methodology for optimizing production processes in small and medium-sized enterprises. *Operations Research Perspectives*, 12, 100308. { HYPERLINK "https://doi.org/10.1016/j.orp.2024.100308" }.
7. Norouzi, A., Heidarifar, H., Borhan, H., Shahbakhti, M., Koch, C. 2023. Integrating machine learning and model predictive control for automotive applications: A review and future directions. *Engineering Applications of Artificial Intelligence*, 120, 105878. { HYPERLINK "https://doi.org/10.1016/j.engappai.2023.105878" }.
8. Dahrouj, H., Alghamdi, R., Alwazani, H., Bahanshal, S., Ahmad, A., Faisal, A., Shalabi, R., Alhadrami, R., Subasi, A., Al-Nory, M., Kittaneh, O., Shamma, J. 2021. An overview of machine learning-based techniques for solving optimization problems in communications and signal processing. *IEEE Access*, 9, 74908-74938. { HYPERLINK "https://doi.org/10.1109/ACCESS.2021.3079639" }.

9. Mayer, J., Jochem, R. 2024. Capability indices for digitized industries: A review and outlook of machine learning applications for predictive process control. *Processes*, 12(8), 1730. { HYPERLINK "https://doi.org/10.3390/pr12081730" }.
10. Pohlodek, J., Morabito, B., Schlauch, C., Zometa, P., Findeisen, R. 2022. Flexible development and evaluation of machine-learning-supported optimal control and estimation methods via HILO-MPC. *International Journal of Robust and Nonlinear Control*, 35(7), 2835-2859. { HYPERLINK "https://doi.org/10.1002/rnc.7275" }.
11. Patel, K. 2023. A practical Reinforcement Learning implementation approach for continuous process control. *Computers & Chemical Engineering.*, 174, 108232. { HYPERLINK "https://doi.org/10.1016/j.compchemeng.2023.108232" }.
12. Blackburn, L., Tuttle, J., Andersson, K., Hedengren, J., Powell, K. 2022. Dynamic machine learning-based optimization algorithm to improve boiler efficiency. *Journal of Process Control*, 120, 129-149. { HYPERLINK "https://doi.org/10.1016/j.jprocont.2022.11.002" }.
13. Saab, S., Shen, D., Orabi, M., Kors, D., Jaafar, R. 2022. Iterative learning control: Practical implementation and automation. *IEEE Transactions on Industrial Electronics*, 69(2), 1858-1866. { HYPERLINK "https://doi.org/10.1109/TIE.2021.3063866" }.
14. Kißkalt, D., Mayr, A., Lutz, B., Rögele, A., Franke, J. 2020. Streamlining the development of data-driven industrial applications by automated machine learning. *Procedia CIRP*, 93, 401-406. { HYPERLINK "https://doi.org/10.1016/j.procir.2020.04.009" }.
15. Çınar, Z., Nuhu, A., Zeeshan, Q., Korhan, O., Asmael, M., Safaei, B. 2020. Machine learning in predictive maintenance towards sustainable smart manufacturing in Industry 4.0. *Sustainability*, 12(19), 8211. { HYPERLINK "https://doi.org/10.3390/su12198211" }.