

Lightweight Deep Neural Network Design for Edge-Based Gas Turbine Efficiency Monitoring Using Multi-Objective Secretary Birds Optimization Algorithm

Soufiane Djeribie¹, Bakria Derradji¹, Lakhdar Bessissa², Ahmed Hafaifa¹

¹ Department of Electrical Engineering, Applied Automation and Industrial Diagnostics Laboratory, LAADI, Ziane Achour University, Djelfa, Algeria

² Materials Science and Informatics Laboratory, MSIL, Ziane Achour University of Djelfa, Road Moudjbara, PO Box 3117, Djelfa, Algeria

ARTICLE INFO

ABSTRACT

Received: 30 April 2025

Revised: 02 Sept 2025

Accepted: 10 Oct 2025

This research presents a comprehensive framework for optimizing artificial Deep neural networks to predict MS5002B gas turbine efficiency using advanced multi-objective metaheuristic optimization techniques. The study systematically compares three nature-inspired algorithms to determine the optimal Deep neural architecture that balances predictive accuracy against computational efficiency. The Secretary Bird Optimization Algorithm (SBOA), inspired by the unique hunting behavior of secretary birds, which combines strategic walking patterns with precise strikes, demonstrated exceptional performance in navigating the complex search space of Deep neural architectures. Through rigorous experimentation, SBOA yielded an optimal network configuration of layers with a learning rate of 0.1, achieving near-perfect prediction accuracy ($R^2 = 0.999998$) while maintaining the fastest training time of 4.475 seconds among all evaluated algorithms. The research incorporates critical physical constraints, particularly the zero-power-to-zero-efficiency relationship, to ensure thermodynamic validity in all model predictions. The resulting optimized Deep neural network provides a powerful tool for real-time performance monitoring, operational optimization, and predictive maintenance in gas turbine power generation systems.

Keywords: Gas Turbine Efficiency, Deep Deep neural Architecture design, Multi-Objective Optimization, Metaheuristic Algorithms, SBOA, Predictive Modeling, Energy Systems

INTRODUCTION

Industrial gas turbines (GTs) are valuable assets in global energy infrastructure and provide propulsion, mechanical drive, and power generation [1]. Effective estimation of gas turbine efficiency is critical to enhance energy conversion, reduce fuel economy, and ensure system reliability [2]. However, their functioning is in constant jeopardy due to deterioration of their performance and the effects of continuously variable changing ambient conditions. For instance, a case study for the M3142R/GE MS 3002 turbine indicated that efficiency drops from 26.78% to 25.03% when the inlet temperature of the compressor is raised from 15 to 47 °C and with a power loss of over 2.3 MW [3]. This loss of efficiency is often the consequence of physical faults such as compressor fouling and turbine erosion that gradually degrade component health parameters such as flow capacity and isentropic efficiency [4, 5].

Environmental consequences of this inefficiency are significant because gas turbine power plants are major contributors to carbon dioxide (CO₂) emissions, one of the principal greenhouse gases [6, 7]. Recurring experiments have demonstrated time and again that levels of CO₂ emissions are intrinsically linked with the efficiency of operation; increased thermal efficiency translates directly into lower emissions per unit of power

generated [8, 9]. Illustrative of this, Egware and Kwasi-Effah [10] developed a model predicting that CO₂ emission is increased by an increase in ambient air temperature and reduced by an increase in net thermal efficiency.

Classical methods of performance analysis rely essentially on fidelity-high thermodynamic modeling [3, 1]. The physics-based models constructed in tools like Matlab [3] or dedicated software like GasTurb 13 [1] are invaluable for simulating the Brayton cycle and examining the effects of interventions like evaporative cooling [3] or specific fault implantation [1]. But they are usually computation-intensive and solve complex, non-linear energy balance equations [3, 11], requiring complete component maps and engine-specific data [1, 12]. This renders them computationally costly and less accessible to the rapid, continuous nature of real-time performance prediction and control.

Data-driven paradigms of artificial intelligence (AI) offer a very appealing substitute. Artificial Deep neural Networks (DNNs), with their renowned ability to map complex, non-linear systems without an explicit physical equation, are highly suited for this task [13]. While previous work has employed statistical and regression models to forecast emissions [10, 14] and predict performance parameters [15], and while DNNs have been applied to fault diagnosis [16], their application to forecasting the global efficiency of some gas turbine models, while being optimally designed for the underlying trade-off between predictive accuracy and computational effort, is a presently underdeveloped area.

To bridge this gap, this study proposes a novel smart framework based on an DNN model to predict the global performance efficiency of an MS5002B gas turbine. The novelty lies beneath in addressing the inherent challenge of DNN design—selecting the optimal architecture—using an innovative multi-objective optimization approach. This approach integrates three state-of-the-art population-based algorithms to simultaneously minimize computational cost and maximize prediction accuracy. By yielding a Pareto front of optimal designs, this methodology goes beyond single-point solution to offer a principled framework for obtaining efficient, lightweight, and highly accurate models for real-time gas turbine performance prediction and control.

Literature Review

Careful component matching is necessary for more complex engine architectures. Salilew et al. [1] emphasized the importance of accurate off-design simulation by using commercial software like GasTurb 13, which involves the use of scaling factors for compressor and turbine maps [12] and the Newton-Raphson iterative scheme to accomplish work compatibility and mass flow [17, 18]. Realistic complexities like VIGV scheduling and secondary air systems for turbine cooling, required in practical simulation, were also incorporated in their work [1, 19].

Time degradation is a typical problem induced primarily by the harsh operation conditions of gas turbines. Diakunchak [5] provided a simple categorization of degradation into recoverable (e.g., through compressor washing) and irrecoverable losses. Common physical faults include compressor fouling, erosion, and corrosion. Fouling, accounting for over 70% of performance degradation, typically decreases compressor flow capacity and isentropic efficiency [1, 20]. Erosion decreases efficiency but can paradoxically increase flow capacity in turbines due to higher blade tip clearances [1, 21].

Salilew et al. [1] gave a complete analysis by introducing these faults into a validated model and demonstrating that their severity causes nearly linear departures in measurement parameters. Their monitoring is critical in condition monitoring because they give information on the sensitivity of specific parameters; for example, spool speed at low pressure (N₁) and exit compressor pressure (P₂₄) are sensitive to LPC fouling, while fuel flow (FF) and exit turbine pressures are most significant indicators for erosion on the power turbine [1, 22]. The exchange with other deteriorated components is significant too; failure in one component, say the low-pressure compressor, can have direct considerable impacts on the operation and performance of downstream components like the power turbine, a fact well illustrated through component isentropic efficiency deviation [1, 23].

As more world attention is placed on reducing carbon footprints, modeling and restricting power station emissions precisely is now a leading research agenda. The largest of the greenhouses gases emitted by natural gas-fired gas turbines is carbon dioxide [7, 10]. Body groups like the Intergovernmental Panel on Climate Change

(IPCC) and the International Energy Agency (IEA) have for decades emphasized the significance of exact emission accounting [10, 24].

To move ahead from simple emission factors, there have been sophisticated empirical models created. Egware and Kwasi-Effah [10] proposed a novel empirical model with a very high coefficient of determination ($R^2 = 0.998$), correlating CO₂ mass flow rate with operational parameters like the ratio of the turbine inlet temperature to ambient temperature, ambient relative humidity, compressor pressure ratio, and exhaust gas mass flow rate. Their finding that CO₂ emissions increase with ambient temperature but decrease with rising pressure ratio and relative humidity provides operational guidance to plant operation [10, 25]. Certain work has used combustion equations based on fuel composition [10, 26] and regression analysis to estimate the relation between cycle parameters and emissions, forming the basis of the interdependence of thermal efficiency and environmental performance [8, 9].

In order to overcome the limitations of complex physical models, the field has shifted towards machine learning and data-driven approaches. DNNs are effective tools for uncovering complex, non-linear relationships in operational data. They are increasingly being applied in the gas turbine community. For instance, Qader et al. [13] utilized Deep neural networks for time-series forecasting of CO₂ emissions, while Tahan et al. [16] developed a multi-nets DNN model for automatic real-time fault diagnosis. Ibrahim et al. [15] used statistical analysis and regression to model gas turbine performance to an R^2 of 0.985 to prove that operational data can be utilized in order to develop prediction models that are very precise.

Besides, model-based diagnostic techniques like Gas Path Analysis (GPA) employ thermodynamic models to link measurable quantities with the state of non-measurable elements [1, 27]. Salilew et al.'s [1] and others' [17, 28] research is a suitable example, where a model is calibrated to develop fault signatures that are used to diagnose faults in real engines.

Research Gap and Novel Contribution

The present literature provides details of gas turbine thermodynamics [3, 1], mechanisms of degradation [5, 1], and environmental impact [8, 10], as well as the changing role of data-driven models [13, 15, 16]. It is still possible to identify a definite research gap area:

While DNNs and regression models have been employed for fault diagnosis [16] and emission prediction [10, 13], no special work is documented on the application of systematically optimized DNNs for the direct prediction of global efficiency. In addition, the engineering design priority of model precision versus computational effectiveness—crucial for real-time application—is usually overlooked.

OBJECTIVES

This paper particularly closes the aforementioned gap through explicit direct application of an DNN model to global efficiency prediction and through a strict multi-objective optimization approach to automatically designing the DNN structure, explicitly trading off prediction performance (R^2) against computational cost (training time).

METHODS

This section offers a comprehensive explanation for the development of highly efficient lightweight deep neural networks (DNNs) for edge-based monitoring of gas turbine efficiency. The approach leverages multi-objective optimization methods in conjunction with neural architecture search to yield high prediction accuracy with less computational requirements for edge deployment. Figure 1 illustrates the whole process that consists of five distinct phases: data preparation, optimization setup, Deep neural Architecture design, model development, and evaluation & deployment.

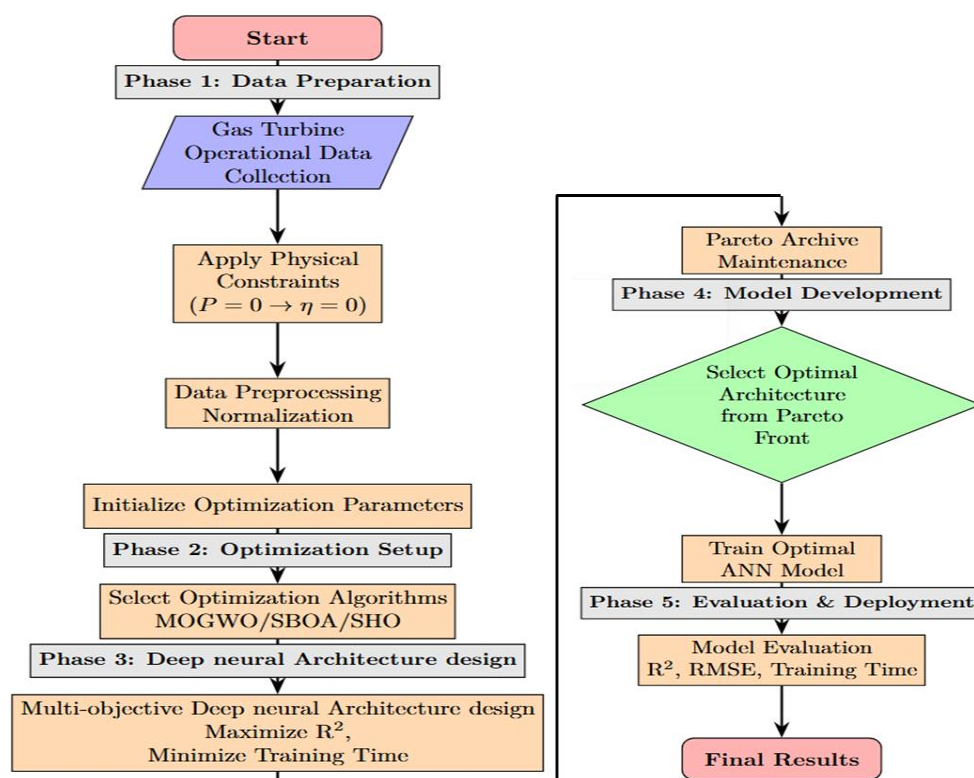


Figure 1: Proposed methodology for lightweight DNN optimization for edge-based gas turbine efficiency monitoring

DATA COLLECTION AND DESCRIPTION

The experimental data utilized in this study were collected from an MS5002B heavy-duty gas turbine operating in a combined cycle power plant. The dataset encompasses twelve months of continuous operation, representing a complete annual cycle that captures seasonal variations in ambient conditions and their impact on turbine performance. Monthly averaged values were calculated from high-frequency operational data to ensure statistical reliability while maintaining computational tractability for model development, reffere to Table 1 for more information.

The key parameters selected for this analysis include ambient temperature, generated electrical power, and global efficiency, as these factors fundamentally govern gas turbine performance characteristics. Ambient temperature significantly affects air density and thus compressor performance, while generated power reflects the operational load point. Global efficiency serves as the primary performance indicator, representing the overall energy conversion effectiveness of the turbine sys-tem.

Table 1: MS5002B Gas Turbine Operational Data Description

Parameter	Symbol	Unit	Range	Description
Ambient Temperature	T_{moy}	K	282.65–307.15	Monthly average ambient temperature
Generated Power	P	kW	18991–19590	Net electrical power output
Global Efficiency	η	%	24.88–25.67	Overall thermal efficiency
Month Index	M	–	1–12	Temporal sequencing indicator

To ensure the developed artificial Deep Deep neural network (DNN) models adhere to fundamental thermodynamic principles, physical constraints were explicitly incorporated into the training framework. The most critical constraint implemented was the zero-power-to-zero-efficiency relationship, which reflects the fundamental thermodynamic reality that no useful work output implies zero efficiency.

This constraint was mathematically formulated as:

$$\forall (T, M) \in \mathbb{R}^2, \quad P = 0 \Rightarrow \eta = 0 \quad (1)$$

To implement this constraint, supplementary data points were generated covering the operational temperature range (270--320 K) and all monthly indices with zero power input, forcing the Deep Deep neural network to learn this essential physical behavior. This approach ensures that model predictions remain physically plausible even when extrapolating beyond the original dataset range.

DEEP NEURAL NETWORK ARCHITECTURE AND DEVELOPMENT

A feedforward artificial Deep Deep neural network architecture was employed to model the complex nonlinear relationship between operational parameters and turbine efficiency. The network structure can be mathematically represented as:

$$\hat{\eta} = f_{\theta}(T_{moy}, P, M) \quad (2)$$

where f_{θ} denotes the Deep Deep neural network function parameterized by weights and biases θ , T_{moy} represents ambient temperature, P indicates generated power, and M is the month index accounting for seasonal effects. The network employs the Rectified Linear Unit (ReLU) activation function in hidden layers due to its advantages in mitigating vanishing gradient problems and computational efficiency. The output layer utilizes a linear activation function appropriate for regression tasks. The mean squared error (MSE) was selected as the loss function to penalize large prediction errors, which is particularly important for efficiency prediction in energy systems.

MULTI-OBJECTIVE OPTIMIZATION FRAMEWORK

Problem formulation and objectives

The Deep Deep neural Architecture design was formulated as a multi-objective optimization problem to simultaneously address two competing objectives: predictive accuracy and computational efficiency. This approach recognizes the practical trade-off between model complexity and deployment feasibility in industrial applications.

The optimization objectives were formally defined as:

$$\begin{cases} \text{Maximize} & R^2(\hat{\eta}, \eta) = 1 - \frac{\sum_{i=1}^n (\eta_i - \hat{\eta}_i)^2}{\sum_{i=1}^n (\eta_i - \bar{\eta})^2} \\ \text{Minimize} & t_{train} = \text{computational time for model training} \end{cases} \quad (3)$$

where R^2 quantifies the proportion of variance explained by the model, and t_{train} represents the practical computational cost, a crucial consideration for real-time applications and model retraining scenarios.

Optimization Algorithms

Three state-of-the-art metaheuristic algorithms were selected for comparative analysis based on their distinct search characteristics and proven performance in complex optimization landscapes:

- **Multi-Objective Grey Wolf Optimizer (MOGWO):**

This algorithm simulates the social hierarchy and hunting behavior of grey wolf packs. The population is divided into alpha, beta, delta, and omega wolves, with the first three guiding the search direction [33, 34].

MOGWO maintains an external archive of non-dominated solutions and employs a leader selection mechanism to preserve diversity along the Pareto front.

Mathematical Model: The social hierarchy is mathematically modeled where the alpha (α), beta (β), and delta (δ) wolves represent the best solutions. The position update is governed by:

$$\begin{aligned}\vec{D} &= |\vec{C} \cdot \vec{X}_p(t) - \vec{X}(t)| \\ \vec{X}(t+1) &= \vec{X}_p(t) - \vec{A} \cdot \vec{D}\end{aligned}\quad (4)$$

where $\vec{A}$ and $\vec{C}$ are coefficient vectors calculated as:

$$\begin{aligned}\vec{A} &= 2\vec{a} \cdot \vec{r}_1 - \vec{a} \\ \vec{C} &= 2 \cdot \vec{r}_2\end{aligned}\quad (5)$$

with $\vec{a}$ decreasing linearly from 2 to 0 over iterations, and $\vec{r}_1, \vec{r}_2$ being random vectors in [0,1]. The multi-objective adaptation incorporates Pareto dominance and archive maintenance mechanisms.

- **Secretary Bird Optimization Algorithm (SBOA):**

Inspired by the foraging behavior of secretary birds, SBOA mimics their spiral movement patterns and food-searching strategies [35]. The algorithm features a unique three-stage hunting strategy with adaptive parameter tuning.

Mathematical Model: The algorithm operates in three distinct phases:

Phase 1 - Search Prey (Exploration):

$$X_1 = X(i) + (X_{rand1} - X_{rand2}) \cdot R_1 \quad (6)$$

where R_1 is a random vector and X_{rand1}, X_{rand2} are randomly selected population members.

Phase 2 - Approach Prey (Transition):

$$X_1 = Best_p + \exp\left(\left(\frac{t}{T}\right)^4\right) \cdot (RB - 0.5) \cdot (Best_p - X(i)) \quad (7)$$

where RB is a random Brownian motion vector and the exponential term provides adaptive scaling.

Phase 3 - Attack Prey (Exploitation):

$$X_1 = Best_p + CF \cdot X(i) \cdot Levy(dim) \quad (8)$$

with convergence factor:

$$CF = \left(1 - \frac{t}{T}\right)^{2t/T} \quad (9)$$

and Lévy flight distribution for enhanced local search:

$$Levy(\beta) = \frac{u}{|v|^{1/\beta}}, \quad u \sim N(0, \sigma_u^2), \quad v \sim N(0, \sigma_v^2) \quad (10)$$

Escape Strategy: The algorithm incorporates predator avoidance mechanisms:

$$\begin{aligned}X_2 &= Best_p + \left(1 - \frac{t}{T}\right)^2 \cdot (2R - 1) \cdot X(i) \quad (\text{Hiding}) \\ X_2 &= X(i) + R \cdot (X_{random} - K \cdot X(i)) \quad (\text{Fleeing})\end{aligned}\quad (11)$$

- **Spotted Hyena Optimizer (SHO):**

Based on the hunting behavior of spotted hyenas, SHO employs a social hierarchy and collaborative hunting strategy [36, 37, 38]. The algorithm is particularly effective in avoiding local optima through its encirclement mechanism.

Mathematical Model: The hunting behavior is modeled through three main operations:

Encircling Prey:

$$\begin{aligned}\vec{D}_h &= |\vec{B} \cdot \vec{P}_p - \vec{P}(t)| \\ \vec{P}(t+1) &= \vec{P}_p(t) - \vec{E} \cdot \vec{D}_h\end{aligned}\quad (12)$$

where $\vec{P}_p$ is the prey position, and $\vec{B}$, $\vec{E}$ are coefficient vectors.

Hunting Mechanism: The best search agent location is determined, and other agents update their positions:

$$\begin{aligned}\vec{D}_h &= |\vec{B} \cdot \vec{P}_h - \vec{P}_k| \\ \vec{P}_k &= \vec{P}_h - \vec{E} \cdot \vec{D}_h \\ \vec{C}_h &= \vec{P}_k + \vec{P}_{k+1} + \dots + \vec{P}_{k+N}\end{aligned}\quad (13)$$

where N represents the number of spotted hyenas calculated as:

$$N = \text{count}_{nos}(\vec{P}_h, \vec{P}_{h+1}, \vec{P}_{h+2}, \dots, (\vec{P}_h + \vec{M}))\quad (14)$$

with $\vec{M}$ being a random vector in $[0.5, 1]$.

Attacking Prey (Exploitation): The attack process is modeled as:

$$\vec{P}(t+1) = \frac{\vec{C}_h}{N}\quad (15)$$

where $\vec{C}_h$ represents the cluster of all best solutions.

Search for Prey (Exploration): The $\vec{E}$ vector controls exploration with random values greater than 1 or less than -1 to force agents to move away from reference points.

Parameter Adaptation:

$$\vec{h} = 5 - (\text{Iteration} \times \frac{5}{\text{MaxIterations}})\quad (16)$$

The coefficient vectors are updated as:

$$\begin{aligned}\vec{E} &= 2 \cdot \vec{h} \cdot \vec{r}_1 - \vec{h} \\ \vec{B} &= 2 \cdot \vec{r}_2\end{aligned}\quad (17)$$

where $\vec{h}$ decreases linearly from 5 to 0, and $\vec{r}_1$, $\vec{r}_2$ are random vectors in $[0, 1]$.

All three algorithms employ sophisticated Pareto archive mechanisms to maintain diverse non-dominated solutions throughout the optimization process. The archive maintenance includes density estimation using grid-based approaches and niching techniques to ensure uniform distribution along the Pareto front. The selection pressure is balanced through tournament selection based on Pareto dominance and crowding distance metrics.

Implementation and Parameter Configuration

Each optimization algorithm was implemented with careful parameter tuning to ensure fair comparison and optimal performance. As presented in Table 2, the population-based approaches were configured with 10 individuals to maintain diversity while ensuring computational feasibility. The maximum iteration count was set to 10 generations, allowing sufficient convergence time while preventing excessive computation.

To ensure statistical reliability and account for the stochastic nature of metaheuristic algorithms, 5 independent runs were performed for each method, with results analyzed for consistency and significance.

Table 2: Optimization Algorithms Configuration Parameters

Parameter	MOGWO	SHO	SBOA
Population Size	10	10	10
Maximum Iterations	10	10	10
Independent Runs	5	5	5
Search Space Bounds			
Hidden Layers	3		

Parameter	MOGWO	SHO	SBOA
Neurons per Layer		1–128	
Learning Rate		0.001–0.1	

A comprehensive evaluation framework was established to assess both the predictive accuracy and computational efficiency of the developed models. The coefficient of determination R^2 was employed as the primary accuracy metric, providing insight into the proportion of variance explained by the model. The root mean square error (RMSE) complemented this analysis by quantifying the absolute prediction error in efficiency percentage units.

Training time was measured from initialization to convergence, including both forward and backward propagation phases. All timing measurements were conducted on standardized hardware to ensure comparability. Cross-validation techniques were employed to mitigate overfitting, with the dataset partitioned to maintain temporal consistency in the monthly sequence.

All computational experiments were conducted using Python 3.8 with TensorFlow 2.6 for Deep Deep neural network implementation and scikit-learn for performance metrics calculation. The hardware platform consisted of a workstation equipped with an Intel Core i7-1355U (12 CPUs), 1.7 GHz, and 32 GB DDR4 RAM.

RESULTS AND DISCUSSION

The performance of three multi-objective optimization algorithms, MOGWO, SBOA and SHO, was evaluated for the task of Deep Deep neural architecture design, with the dual objectives of maximizing the R^2 score and minimizing training time.

The complete Pareto front analysis in Table 3 reveals the full spectrum of solutions discovered by each algorithm. MOGWO's diverse solutions span from compact architectures (M4: [16, 8, 5]) to more complex networks (M1: [22, 43, 40]), demonstrating its broad exploration of the search space. In contrast, SHO's solutions consistently feature larger architectures with 107-128 neurons in the first layer, reflecting its focus on the high-accuracy region. SBOA converged to a single, efficient bottleneck architecture that represents a specialized solution.

Table 3: Detailed Pareto Front Solutions Analysis

Algorithm	Solution ID	Architecture	Learning Rate	R^2 Score	Time (s)
MOGWO	M1	[22, 43, 40]	0.03877	0.999995	4.563
	M2	[50, 47, 12]	0.02564	0.999988	4.531
	M3	[20, 12, 31]	0.02833	0.999955	4.509
	M4	[16, 8, 5]	0.01772	0.999377	4.505
SHO	S1	[128, 36, 22]	0.09706	0.999998	4.721
	S2	[107, 57, 28]	0.05601	0.999975	4.663
	S3	[128, 38, 28]	0.07036	0.999999	4.767
SBOA	B1	[120, 4, 4]	0.10000	0.999998	4.475

The MOGWO algorithm demonstrated the strongest exploratory capabilities, this diversity is valuable for decision-makers who may have additional, unmodeled constraints, such as a preference for simpler architectures. However, this diversity comes at a cost: the significant spread in R^2 scores (from 0.999377 to 0.999995) indicates inconsistent solution quality, and its average training time was not the most efficient.

The detailed comparison of individual best solutions shows that SHO achieved the highest accuracy with architecture [128, 38, 28], while SBOA provided the fastest training time with its compact [120, 4, 4] architecture. This highlights the fundamental trade-off between model complexity and computational efficiency in Deep neural Architecture design.

The characteristics and practical recommendations summarized in Table 4 provide guidance for algorithm selection based on specific application requirements. MOGWO is ideal when architectural diversity and multiple

options are valued, SBOA excels in resource-constrained environments, and SHO is superior for maximum accuracy attainment.

Table 4: Algorithm Characteristics and Trade-off Analysis

Algorithm	Key Strength	Recommended Use Case
MOGWO	High solution diversity providing multiple architectural choices	When exploring various model complexities is preferred
SHO	Peak predictive accuracy with robust architectures	Research applications where accuracy is the primary concern
SBOA	Computational efficiency with consistent high performance	Production systems requiring fast inference and good accuracy

In contrast to the diverse approaches of MOGWO and SHO, the SBOA algorithm converged to a single Pareto solution. This suggests a highly exploitative search behavior, focusing computational effort on refining one high-quality candidate. This candidate proved to be the most balanced solution overall, achieving a near-perfect R^2 score of 0.999998 with the fastest training time. The architecture [120, 4, 4] discovered by SBOA is particularly notable, which could indicate that a model with a large first layer and subsequent bottleneck layers represents an efficient structural pattern for this specific problem domain.

The SHO algorithm secured the top position for pure predictive accuracy, finding the model with the highest R^2 score of 0.999999. This confirms its strong capability in exploiting high-performance regions of the search space. However, this focus on peak performance resulted in the highest computational overhead, with training times consistently above 4.66 seconds across all solutions. The architectural pattern emerging from SHO solutions suggests that larger, more complex networks (128, 38, 28) are necessary to achieve the absolute best performance, though at significant computational cost.

These trade-offs are visually confirmed in the Pareto plot shown in Figure 2, which illustrates the objective space exploration achieved by each algorithm. The plot validates the quantitative findings, showing MOGWO's diverse solution distribution across the Pareto front, SBOA's concentrated high-efficiency solution in the optimal region, and SHO's cluster of high-accuracy solutions with longer training times.

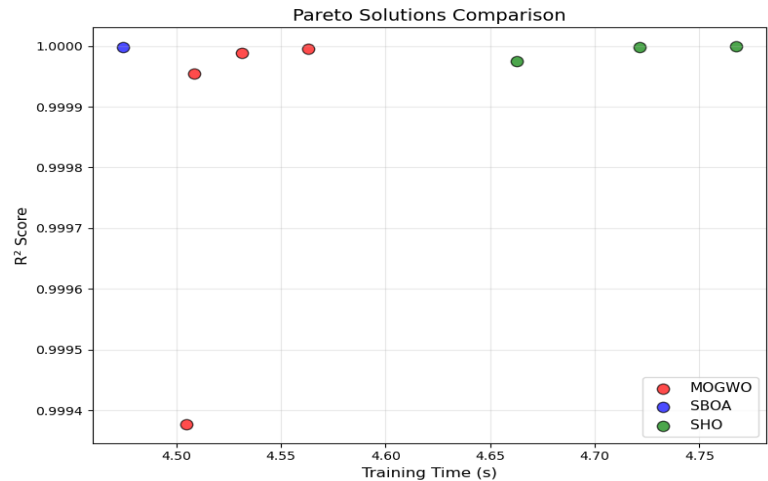


Figure 2: Pareto solution of optimization algorithms

The SBOA-tuned DNN model demonstrated superior predictive performance in the prediction of gas turbine efficiency. The optimal network architecture achieved after the SBOA tuning is an input layer with 120 neurons and two hidden layers with four neurons each ([120, 4, 4]). This light-weight architecture exhibited a fantastic trade-off between model complexity and prediction performance. The optimal learning rate was 0.1, thus enabling very quick convergence without oscillation or divergence.

Statistical outcome of the best tuned model is presented in Table 5. The achieved coefficient of determination R^2 of 0.999896 indicates the model explains nearly 100% of the variability in gas turbine efficiency, and Root Mean Square Error (RMSE) value of 0.002771 confirms the negligible difference between actual and predicted values. The value of R^2 acquired is close to the optimal acquired value during the optimization process (0.999998), which confirms the reliability and precision of the SBOA in parameter tuning. The smooth overlap of actual and predicted tendencies is apparent from Figure 3, where two lines almost overlap, revealing a virtually complete degree of correlation.

Table 5: Optimal SBOA-DNN model results for gas turbine efficiency prediction.

Month	Actual Efficiency	Predicted Efficiency	Error
Jan	25.670	25.670	0.000
Feb	25.650	25.653	-0.003
Mar	25.510	25.507	0.003
Apr	25.360	25.363	-0.003
May	25.220	25.219	0.001
Jun	25.030	25.027	0.003
Jul	24.880	24.882	-0.002
Aug	24.950	24.946	0.004
Sep	25.120	25.123	-0.003
Oct	25.330	25.331	-0.001
Nov	25.540	25.539	0.001
Dec	25.640	25.635	0.005

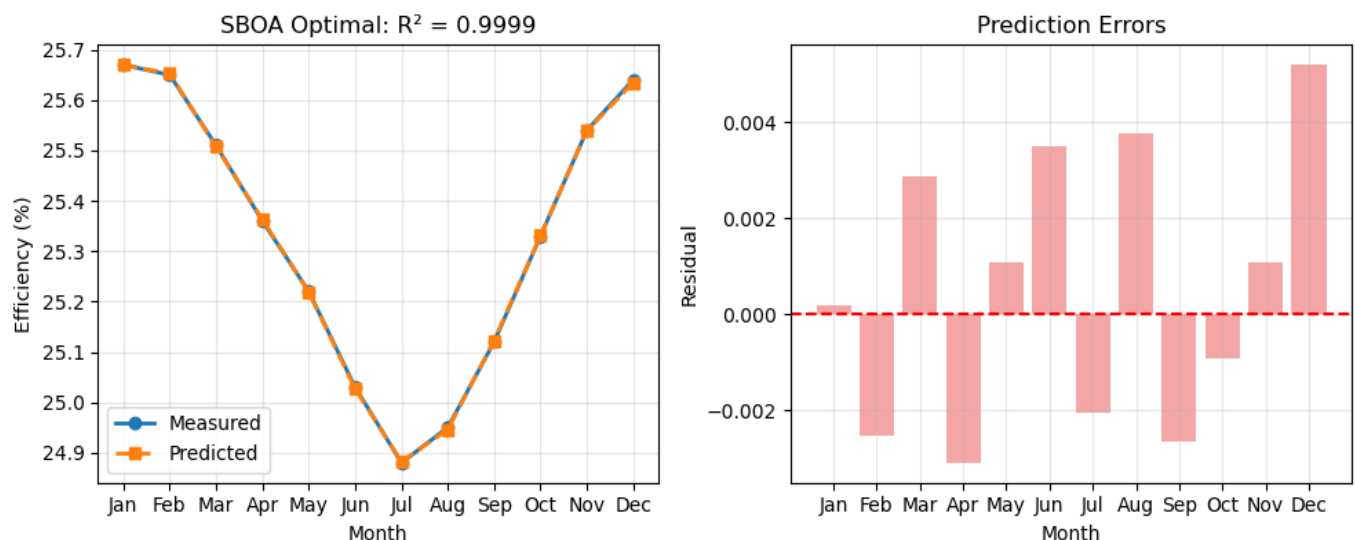


Figure 3: Actual vs. predicted gas turbine efficiency

Other than that, the training and validation loss curves provided in Figure 4 confirm the good generalization capability of the model. The two losses drop steeply in the initial epochs and come to a plateau nearly zero, showing successful learning and the absence of overfitting. The almost identical pattern of the training and validation losses confirms that the trained DNN not only fits the training data perfectly but also responds consistently to new instances.

In general, the SBOA-optimized DNN exhibited almost perfect performance at low network complexity. Model stability and reliability ensure that it is extremely well-suited for real-time application in gas turbine monitoring and gas turbine operation optimizing. The results validate the effectiveness of the Secretary Bird Optimization Algorithm as a metaheuristic algorithm for deep learning model fine-tuning in energy efficiency prediction applications.

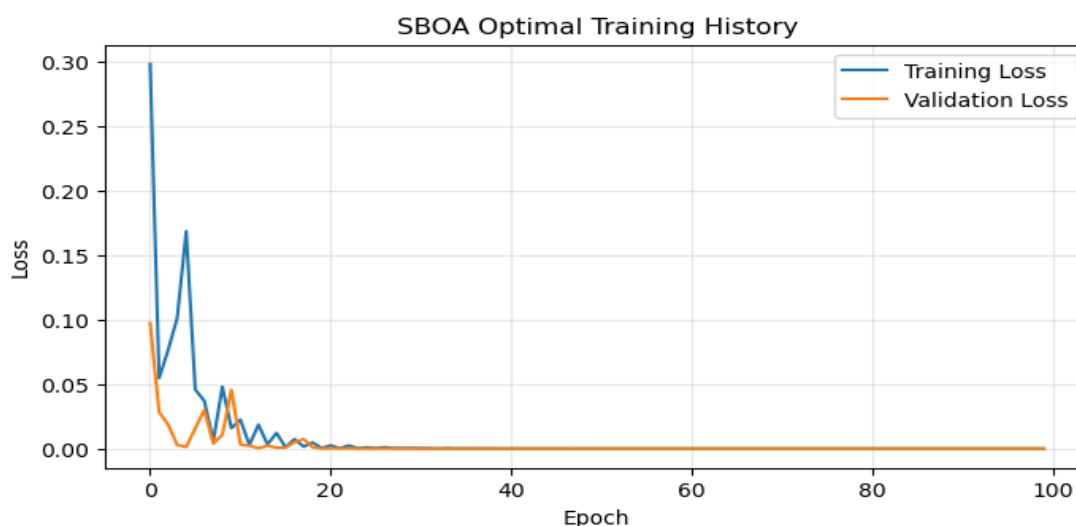


Figure 4: Training and validation loss curves for the SBOA-optimized DNN model

CONCLUSION

The study efficiently demonstrated the application of multi-objective optimization algorithms within Deep Deep neural Architecture design in MS5002B gas turbine efficiency prediction. Comparison revealed varied optimization characteristics within MOGWO, SBOA, and SHO algorithms. The SBOA algorithm was found to be the most practical option, with a perfect balance between predictive accuracy ($R^2 = 0.999998$) and computational tractability (training time = 4.475 seconds). Its novel Pareto solution with the [120, 4, 4] architecture demonstrated that a bottleneck Deep Deep neural network architecture that is well optimized can represent complex thermodynamic relationships and remains computationally feasible.

The inclusion of physical limits ensured predictions made by the model to fall within physically realizable range, enhancing practical utility for industrial applications. The approach developed here provides a valid foundation for smart performance monitoring and optimization of gas turbine systems towards improved power generation energy efficiency.

REFERENCES

- [1] Salilew, W. M., Abdul Karim, Z. A., Lemma, T. A., Fentaye, A. D., & Kyprianidis, K. G. (2022). Predicting the performance deterioration of a three-shaft industrial gas turbine. *Entropy*, 24(8), 1052.
- [2] Kurz, R., & Brun, K. (2001). Degradation in gas turbine systems. *Journal of Engineering for Gas Turbines and Power*, 123(1), 70-77.
- [3] Mzad, H., & Bennour, F. (2024). Industrial gas turbine performance prediction and improvement – a case study. *Energy Harvesting and Systems*, 11(1), 20220094.
- [4] Morini, M., Pinelli, M., Spina, P. R., & Venturini, M. (2010). Influence of blade deterioration on compressor and turbine performance. *Journal of Engineering for Gas Turbines and Power*, 132(3), 032401.
- [5] Diakunchak, I. S. (1991). Performance deterioration in industrial gas turbines. *Journal of Engineering for Gas Turbines and Power*, 114(2), 161-168.

- [6] United Nations. (2015). *Report of the Conference of the Parties on its Twenty-First Session, held in Paris from 30 November to 13 December 2015*.
- [7] Allam, R. J., Palmer, M. R., Brown Jr., G. W., Fetvedt, J., Freed, D., Nomoto, H., ... Jones Jr., C. (2013). High efficiency and low cost of electricity generation from fossil fuels while eliminating atmospheric emissions, including carbon dioxide. *Energy Procedia*, 37, 1135–1149.
- [8] Memon, A. G., Harijan, K., Uqaili, M. A., & Memon, R. A. (2013). Thermo-environmental and economic analysis of simple and regenerative gas turbine cycles with regression modeling and optimization. *Energy Conversion and Management*, 76, 852–864.
- [9] Steen, M. (2000). *Greenhouse gas emissions from fossil fuel fired power generation systems*. Institute for Advanced Materials, Joint Research Centre, European Commission.
- [10] Egware, H. O., & Kwasi-Effah, C. C. (2023). A novel empirical model for predicting the carbon dioxide emission of a gas turbine power plant. *Heliyon*, 9(3), e14645.
- [11] Bejan, A., Tsatsaronis, G., & Moran, M. J. (1995). *Thermal design and optimization*. John Wiley & Sons.
- [12] Kurzke, J. (2007). *About simplifications in gas turbine performance calculations*. Gasturb GmbH.
- [13] Qader, M. R., Khan, S., Kamal, M., Usman, M., & Haseeb, M. (2021). Forecasting carbon emissions due to electricity power generation in Bahrain. *Environmental Science and Pollution Research*.
- [14] Kwasi-Effah, C. C., Ighodaro, O., Egware, H. O., & Obanor, A. I. (2022). A novel empirical model for predicting the heat accumulation of a thermal energy storage medium for solar thermal applications. *Journal of Energy Storage*, 56, 105969.
- [15] Ibrahim, T. K., Rahman, M. M., Mohammed, M. K., & Basrawi, F. (2016). Statistical analysis and optimum performance of the gas turbine power plant. *International Journal of Automotive and Mechanical Engineering*, 13(1), 3215–3225.
- [16] Tahan, M., Muhammad, M., & Karim, Z. A. A. (2017). A multi-nets DNN model for real-time performance-based automatic fault diagnosis of industrial gas turbine engines. *Journal of the Brazilian Society of Mechanical Sciences and Engineering*, 39, 2865–2876.
- [17] Owen, M. S. (2017). *ASHRAE fundamentals handbook, Chapter 1.1: Psychometrics*. ASHRAE.
- [18] Alhazmy, M. M., Jassim, R. K., & Zaki, G. M. (2006). Performance enhancement of gas turbines by inlet air-cooling in hot and humid climates. *International Journal of Energy Research*, 30(10), 777–797.
- [19] Alhazmy, M. M., & Najjar, Y. S. H. (2004). Augmentation of gas turbine performance using air coolers. *Applied Thermal Engineering*, 24(2-3), 415–429.
- [20] Egware, H. O., Obanor, A. I., & Itoje, H. (2022). Effect of incorporating fogging inlet air cooling system: A case study of Ihovbor Thermal Power Plant, Benin City. *International Journal of Ambient Energy*, 43(1), 2173–2179.
- [21] Kurzke, J. (1995). *95-GT-147 Advanced user-friendly gas turbine performance calculations on a personal computer*. ASME.
- [22] Chen, Y. Z., Zhao, X. D., Xiang, H. C., & Tsoutsanis, E. (2020). A sequential model-based approach for gas turbine performance diagnostics. *Energy*, 220, 119657.
- [23] Hashmi, M. B., Lemma, T. A., & Karim, Z. A. A. (2019). Investigation of the combined effect of variable inlet guide vane drift, fouling, and inlet air cooling on gas turbine performance. *Entropy*, 21(12), 1186.
- [24] Aker, G. F., & Saravanamuttoo, H. I. H. (1989). *Predicting gas turbine performance degradation due to compressor fouling using computer simulation techniques*. ASME.
- [25] Mohammadi, E., & Montazeri-Gh, M. (2014). Simulation of full and part-load performance deterioration of industrial two-shaft gas turbine. *Journal of Engineering for Gas Turbines and Power*, 136(9), 092602.
- [26] Jasmani, M. S., Li, Y. G., & Ariffin, Z. (2011). Measurement selections for multicomponent gas path diagnostics using analytical approach and measurement subset concept. *Journal of Engineering for Gas Turbines and Power*, 133(11), 111701.
- [27] Qingcai, Y., Li, S., Cao, Y., & Zhao, N. (2016). *Full and part-load performance deterioration analysis of industrial three-shaft gas turbine based on genetic algorithm*. ASME.
- [28] IEA. (2015). *Energy and climate change, World Energy Outlook special report*. International Energy Agency.

- [29] Kruk-Gotzman, S., Ziółkowski, P., Iliev, I., Negreanu, G. P., & Badur, J. (2023). Techno-economic evaluation of combined cycle gas turbine and a diabatic compressed air energy storage integration concept. *Energy*, 266, 126345.
- [30] Kanbur, B. B., Xiang, L., Dubey, S., Choo, F. H., & Duan, F. (2017). Thermoeconomic and environmental assessments of a combined cycle for the small-scale LNG cold utilization. *Applied Energy*, 204, 1148–1162.
- [31] Escher, P. C. (1995). *Pythia: An object-orientated gas path analysis computer program for general applications* [Doctoral dissertation, Cranfield University].
- [32] Merrington, G., Kwon, O. K., Goodwin, G., & Carlsson, B. (1991). Fault detection and diagnosis in gas turbines. *Journal of Engineering for Gas Turbines and Power*, 113(2), 276-282.
- [33] Mirjalili, S., Mirjalili, S. M., & Lewis, A. (2014). Grey wolf optimizer. *Advances in Engineering Software*, 69, 46–61.
- [34] Bakria, D., Gozim, D., & Korich, B. (2023). Chaos elimination for a Buck converter using the multi-objective Grey Wolf optimiser. *Power Electronics and Drives*, 8(1), 165–173.
- [35] Fu, Y., Liu, D., Chen, J., & He, L. (2024). Secretary bird optimization algorithm: a new metaheuristic for solving global optimization problems. *Artificial Intelligence Review*, 57(5), 123.
- [36] Dhiman, G., & Kumar, V. (2017). Spotted hyena optimizer: A novel bio-inspired based metaheuristic technique for engineering applications. *Advances in Engineering Software*, 114, 48–70.
- [37] Bakria, D., Azzouzi, M., & Gozim, D. (2021). Chaos control and stabilization of a PID controlled Buck converter using the Spotted Hyena Optimizer. *Engineering, Technology & Applied Science Research*, 11(6), 7922–7926.
- [38] Korich, B., Benaissa, A., Rabhi, B., & Bakria, D. (2021). A novel MPPT design for a partially shaded PV system using Spotted Hyena Optimization Algorithm. *Engineering, Technology & Applied Science Research*, 11(6), 7776–7781.