

Article

System Identification Methodology of a Gas Turbine Based on Artificial Recurrent Neural Networks

Rubén Aquize ¹, Armando Cajahuaringa ¹, José Machuca ¹, David Mauricio ^{2,*} and Juan M. Mauricio Villanueva ³ ¹ Universidad Nacional de Ingeniería, Rimac 150101, Peru² Universidad Nacional Mayor de San Marcos, Lima 15081, Peru³ Universidade Federal da Paraíba Campus I, Joao Pessoa, Paraíba 58051-900, PB, Brazil

* Correspondence: dmauricios@unmsm.edu.pe

Abstract: The application of identification techniques using artificial intelligence to the gas turbine (GT), whose nonlinear dynamic behavior is difficult to describe through differential equations and the laws of physics, has begun to gain importance for a little more than a decade. NARX (Nonlinear autoregressive network with exogenous inputs) is one of the models used to identify GT because it provides good results. However, existing studies need to show a systematic method to generate robust NARX models that can identify a GT with satisfactory accuracy. In this sense, a systematic method is proposed to design NARX models for identifying a GT, which consists of nine precise steps that go from identifying GT variables to obtaining the optimized NARX model. To validate the method, it was applied to a case study of a 215 MW SIEMENS TG, model SGT6-5000F, using a set of 2305 real-time series data records, obtaining a NARX model with an MSE of 1.945×10^{-5} , RMSE of 0.4411% and a MAPE of 0.0643.

Keywords: method; gas turbine; neural network; NARX; identification system; prediction performance



Citation: Aquize, R.; Cajahuaringa, A.; Machuca, J.; Mauricio, D.; Mauricio Villanueva, J.M. System Identification Methodology of a Gas Turbine Based on Artificial Recurrent Neural Networks. *Sensors* **2023**, *23*, 2231. <https://doi.org/10.3390/s23042231>

Academic Editors: Fangqing Wen, Wei Liu, Jin He and Veerendra Dakulagi

Received: 21 December 2022

Revised: 25 January 2023

Accepted: 26 January 2023

Published: 16 February 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

1. Introduction

For the most part, power production plants use fossil fuels or nuclear energy. However, with the depletion of fossil fuels and environmental pollution problems due to greenhouse gas emissions, the proportion of power generated from renewable energy has gradually increased to replace other power generation systems [1]. The production of energy from those that are renewable varies according to natural conditions, such as wind speed and solar radiation. For this reason, it has been studied how to overcome the problem of the intermittency of renewable energies and their storage [2]. Faced with this problem, attention is being paid to the Gas Turbine (GT), which has the fastest response among conventional power generators, which is convenient so as not to affect the stability of electrical networks [3].

A GT has a high specific power and emits much fewer pollutants because it uses natural gas as fuel. In addition, it can achieve fast starts and stops compared to other power generation systems, such as coal and nuclear power generation, and fast load-following operation. Recently developed GTs use the combustion of hydrogen and natural gas mixtures, making greener operations possible [4].

For reasonable control and monitoring of the operation of the GT, a model is required that represents with excellent approximation its real dynamic behavior. There are two approaches to modeling a GT: the white box and the black box.

White box models. The white box models of the GT describe its behavior using physical equations based on engineering principles and its dynamics. These models are used when there is sufficient knowledge about the physics of the system [5], and they can be nonlinear [6,7] or linearized [8,9].

Several investigations have been carried out in the field of identification of the behavior of a GT using white-box models, among them, for adaptive control [10], for a low-power

TG [11], for a heavy-duty power plant [12], for a micro GT [13], for a GT of a power plant based on the Rowen model [14], for a robust control system of a GT, etc.

Nonlinear GT models are usually in the form of large simulation codes and many details. Due to their complexity, they are not generally used for controller design and system stability analysis. On the other hand, the linearized models only consider the input and output variables, so the intermediate variables are not included in the model, losing quality for the control of the GT [15].

Although it is true that, to model a GT, physical equations that govern the real dynamics of the TG can be used with excellent approximation, these are pretty complex. Then, there will always be inaccuracies due to the unmodeled dynamics and parametric uncertainties [15]. In addition, detailed information on its components is needed to develop a physical model, and usually, this information is kept reserved for manufacturers and limited to outsiders.

Black box models. Its objective is to model the behavior of a GT for a set of input data $(u_1, u_2, \dots, u_n)$ and output $(y_1, y_2, \dots, y_n)$ when little or no data is available on information about the physical system [16,17]. One of the most important methods in this type of modeling is the artificial neural network (ANN) since they have demonstrated high precision in industrial applications [6], good operation and control systems [18], and, in addition, allow the optimization of the fuel consumption of the GT [19].

A type of ANN that has shown excellent results is the Nonlinear AutoRegressive network with eXogenous inputs (NARX) [20] because it contemplates feedback of the output variable and delays in the input variables. NARX applications for GT with a layer are found in turboshaft engines [5], aircraft engines [21], and in improving starting operation [22]. An application of NARX with two layers for modeling a GT is found in [17].

Studies show NARX models for specific cases of a GT, whose efficiency depends on the number of layers, number of nodes in each layer, input variables, number of delays, and data. However, no studies have been identified that present procedures to systematically design a NARX for any case of a GT, this being the objective of this paper.

The main contributions of this article are: (a) Provide a general review of the NARX black box models for identifying a GT, its concept, and modeling. (b) Provide a systematic method for designing a NARX model to identify a GT. (c) Show the method's usability through its application to a real case of electric power generation through a heavy-duty single-axis GT.

This article is organized into six sections, and Section 2 reviews the operation of the GT. Section 3 reviews the NARX model applied to a GT. Section 4 proposes the systematic method for the design of the NARX model. Section 5 describes a case study to test the proposed method. Finally, in Section 6, the conclusions are presented.

2. Operation of the GT

The GT is an internal combustion engine that uses the air's gaseous energy to convert the fuel's chemical energy into mechanical energy and works according to the Brayton cycle [23].

Figure 1 shows a typical schematic of a GT system made up of an air compressor, a combustor, and a turbine. Air at atmospheric pressure enters the compressor (1), which is increased by the compressor at its outlet (2), and enters the combustion chamber to mix with the fuel and is ignited to produce hot expanding gas that enters (3) and drives the turbine to generate mechanical energy on its axis that rotates at a certain angular speed, in order to drive electric generators, pumps, compressors, among others. Finally, the gases leave the turbine (4).

The ideal Brayton cycle consists of two isobaric and two isentropic processes [24]. The two isobaric processes consist of the combustion system of the gas turbine and the gas side of the HRSG (Heat Recovery Steam Generating). The two isentropic processes represent compression and expansion processes.

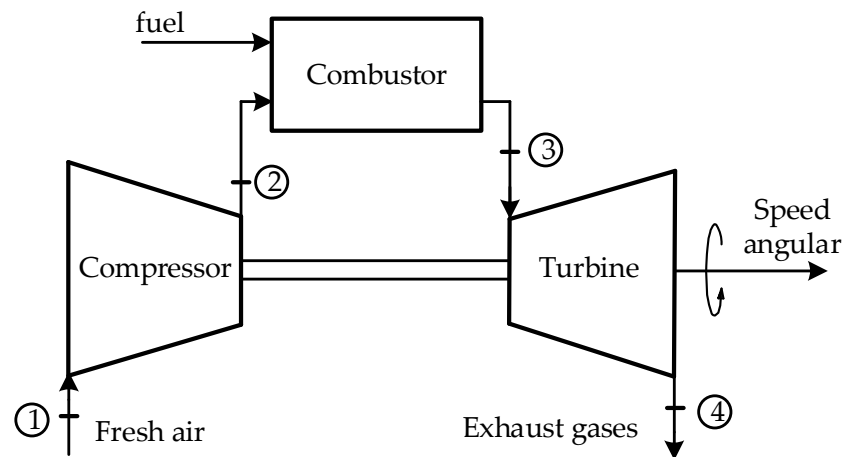


Figure 1. Typical schematic of an open-loop GT system.

Figure 2 presents a simplified application of the first law of thermodynamics to the Brayton cycle of standard air in the absence of changes in kinetic and potential energy.

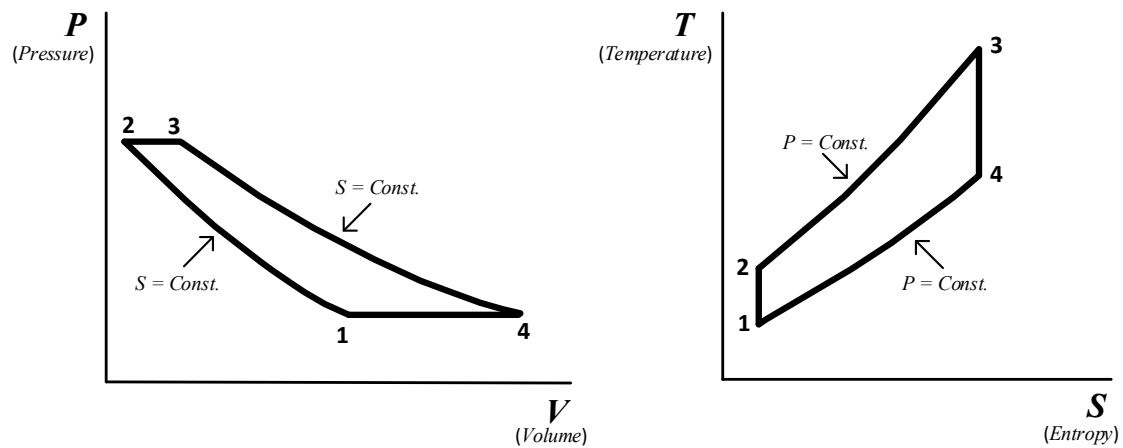


Figure 2. The air-standard Brayton cycle.

The expressions for compressor work (W_c), Turbine work (W_t), total output work (W_{cyc}), heat added to the system ($Q_{2,3}$) and efficiency total cycle ($\eta_{2,3}$), are formulated with Equations (1) to (5), respectively [7].

$$W_c = \dot{m}_a(h_2 - h_1) \quad (1)$$

$$W_t = (\dot{m}_a + \dot{m}_f)(h_3 - h_4) \quad (2)$$

$$W_{cyc} = W_t - W_c \quad (3)$$

$$Q_{2,3} = \dot{m}_f \times LHV_{fuel} = (\dot{m}_a + \dot{m}_f)h_3 - \dot{m}_a h_2 \quad (4)$$

$$\eta_{2,3} = W_{cyc} / Q_{2,3} \quad (5)$$

where $\dot{m}_a$ is the air mass, $\dot{m}_f$ is the mass of fuel, LHV_{fuel} is the lower heating of fuel and h_1, h_2, h_3, h_4 are the enthalpies of states 1 through 4, respectively.

The higher the turbine ignition pressure-temperature ratio, the greater the efficiency of the Brayton cycle. The total cycle efficiency relationship is based on the simplifying assumptions: (1) $\dot{m}_a \gg \dot{m}_f$, (2) the gas is calorically and thermally perfect, that is, the specific heat at constant pressure (c_p) and the specific heat at constant volume (c_v) are constant, and the ratio of specific heat γ remains constant throughout the cycle, (3) the pressure ratio (r_p) of the compressor and the turbine are the same, and (4) components

work at 100% efficiency. Under these assumptions, the efficiency of the ideal Brayton cycle, operating between room temperature and ignition temperature, is given by Equation (6):

$$\eta_{ideal} = \left(1 - \frac{1}{r_p^{\frac{\gamma-1}{\gamma}}} \right) \quad (6)$$

where r_p = Pressure ratio; and γ is the ratio of the specific heats.

Supposing that r_p is the same in the compressor and the turbine, the following relationships hold.

$$\eta_{ideal} = 1 - \frac{T_1}{T_2} \quad (7)$$

with r_p in the turbine.

$$\eta_{ideal} = 1 - \frac{T_4}{T_3} \quad (8)$$

In a real cycle, taking into account the efficiencies η_c of the turbine compressor and the efficiency η_t of the expander, the efficiency of the entire cycle between the ignition temperature T_f and room temperature T_{amb} of the turbine is given by Equation (9).

$$\eta_{cycle} = \left(\frac{\eta_t T_f - \frac{T_{amb} r_p^{\frac{\gamma-1}{\gamma}}}{\eta_c}}{T_f - T_{amb} - T_{amb} \left(\frac{r_p^{\frac{\gamma-1}{\gamma}} - 1}{\eta_c} \right)} \right) \left(1 - \frac{1}{r_p^{\frac{\gamma-1}{\gamma}}} \right) \quad (9)$$

The maximum total efficiency of the thermal cycle, considering the optimal pressure ratio for fixed inlet temperatures and efficiencies for the compressor and the turbine, is obtained with the following relationship.

$$(r_p)_{copt} = \left\{ \frac{1}{T_1 T_3 \eta_1 - T_1 T_3 + T_1^2} \left[T_1 T_3 \eta_t - \sqrt{(T_1 T_3 \eta_t)^2 - (T_1 T_3 \eta_t - T_1 T_3 + T_1^2)(T_3^2 \eta_c \eta_t - T_1 T_3 \eta_c \eta_t + T_1 T_3 \eta_t)} \right] \right\}^{\frac{\gamma}{\gamma-1}} \quad (10)$$

So that there are no losses in the compressor and the turbine, in Equation (10), it must be considered $\eta_c = \eta_t = 1$, with which the optimal pressure ratio is reduced to:

$$(r_p)_{copt} = \left(\frac{T_1 T_3}{T_1^2} \right)^{\frac{\gamma}{\gamma-1}} \quad (11)$$

The optimum pressure ratio for the maximum output work of a turbine, taking into account the efficiencies of the compressor and the turbine expander, is given by Equation (12).

$$r_{p_{wopt}} = \left[\left(\frac{T_3 \eta_c \eta_t}{2 T_1} \right) + \frac{1}{2} \right]^{\frac{\gamma}{\gamma-1}} \quad (12)$$

3. The NARX Model

The NARX model is a type of ANN with feedback suitable for non-linear modeling systems, especially time series, that use past measurements to predict future values [25].

In Figure 3, the NARX model is illustrated, where the “System” represents a real or artificial process to be approximated, and the “Model” represents the system that allows its behavior to be simulated. NARX can be used in series-parallel mode or parallel mode. The series-parallel mode predicts one or more future steps based on past exogenous inputs (u) and outputs (y) of the system. On the contrary, in parallel mode, the model’s output is considered instead of the output of the system [17]. Considering that the model output

may have an error, these could negatively influence the model outputs in relation to the system, so in this work, we will consider the series-parallel mode.

$$U(N, n) = (U_1(n), \dots, U_R(n)) \quad (13)$$

$$U_i(n) = (u_i(t-1), \dots, u_i(t-n)), \forall i = 1, \dots, n \quad (14)$$

The future value of the output signal $\hat{y}(t)$ depends on its m past values $Y(t, m)$ and on $R \times n$ past exogenous inputs $U(R, n)$. The function f approximates the behavior of the TG. NARX can be implemented using a feedforward neural network to approximate the function f [26].

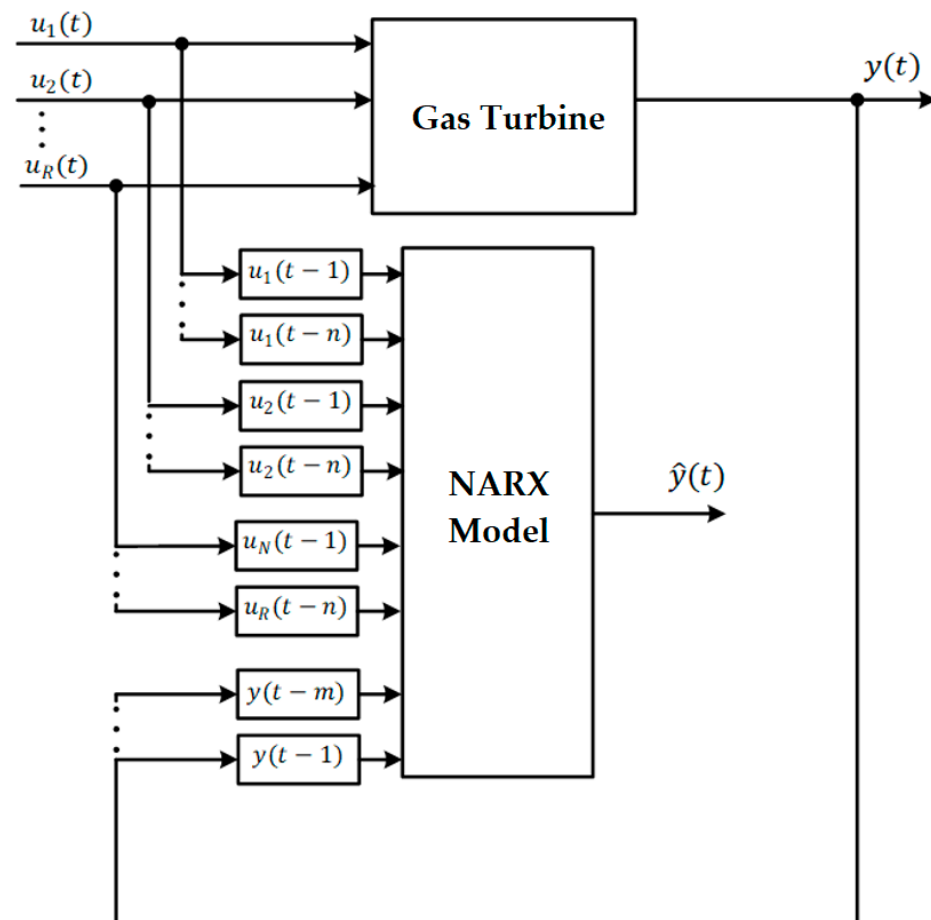


Figure 3. Structure of the NARX Serial-Parallel model for the GT.

4. Systematic Method to Build the NARX Model

A method for the systematic design of a NARX model with high precision for a GT is proposed, based on nine steps (Figure 4), which go from identifying variables to obtaining a GT's identification model. Each of the steps is described below:

4.1. GT Variables (P1)

The first step in any design of a model for a GT is the definition of its input and output variables. The output variables refer to the behavior of the GT to be identified, such as the angular velocity [27], the outlet temperature [20], the output power, and the thermal efficiency [28]. The input variables, also called exogenous, are the variables that influence the behavior of the GT, such as the compressor inlet temperature [10], the compressor inlet pressure, the fuel flow, and the Inlet Guide. Vane (IGV) [22]. All these variables allow for

knowing the dynamic behavior of a GT. Depending on the TG to be identified, the input and output variables of the GT must be identified.

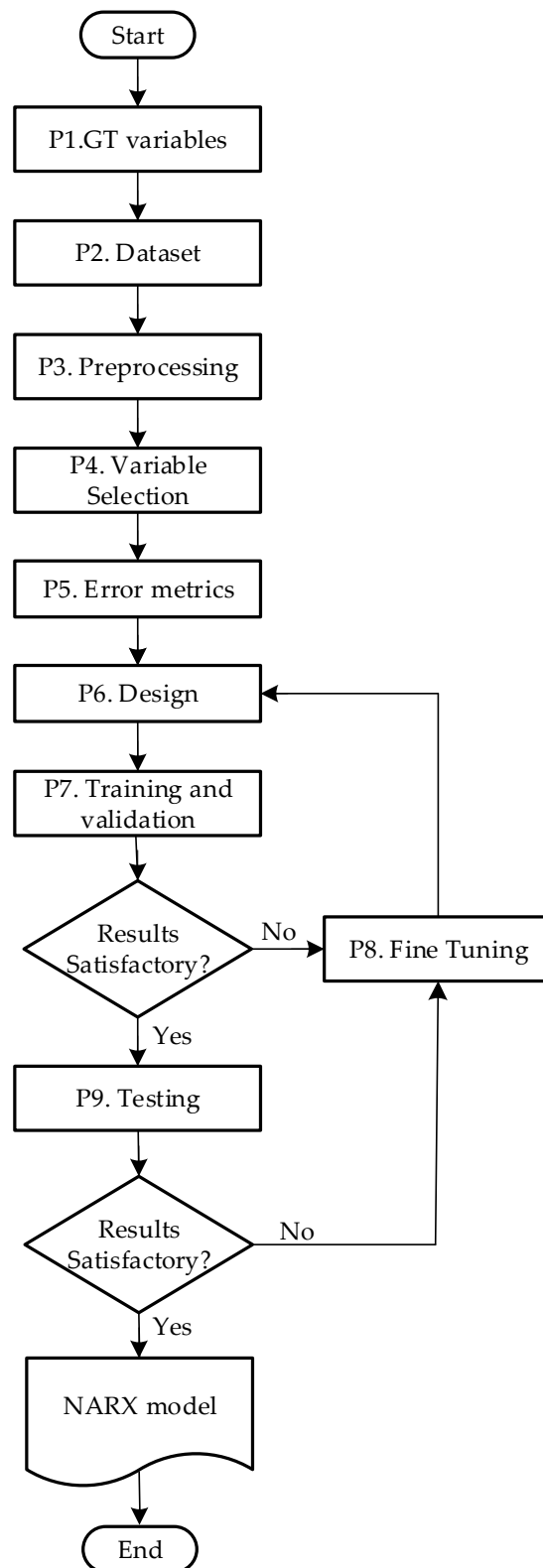


Figure 4. Method for the systematic design of NARX.

4.2. Dataset (P2)

In this step, a dataset is determined on the input and output variables identified in step P1. The dataset can be obtained by simulation [29] or measuring the GT under study [30]. In either case, it is required that data cover one operating cycle of the GT, that the period between data recordings is constant and as small as a few minutes or seconds, and that data are reliable. The dataset, in turn, will be divided into three datasets, one for training, another for validation, and one for testing, with some authors considering proportions of 70–15–15% [16] or 60–15–25% [31], where a–b–c% are the proportions of the dataset for training (a%), validation (b%) and testing (c%).

4.3. Preprocessing (P3)

This process aims to place the dataset ready for processing, that is, to be used in the training, validation, and testing process. For this, the following activities [32,33] are carried out if the case warrants: integration of data in various repositories, cleaning, imputation, and normalization.

4.4. Variable Selection (P4)

The purpose of this step is to select the input variables that are correlated with the output variable. In this way, we can understand the behavior of the GT. For this purpose, there are various techniques, such as the Principal Component Analysis evaluator [34,35] and the Pearson correlation [36], some of them implemented in libraries such as Python, Pytorch, Keras, and in tools such as MatLab and Waikato Environment for Knowledge Analysis (Weka).

4.5. Error Metrics (P5)

The error metrics that are commonly used to evaluate and report the performance of a regression model are, among others: the mean absolute error (MAE) [28], the mean square error (MSE) [37,38], the error mean absolute percentage (MAPE) [39], and root mean square error (RMSE) [40]. Let $\hat{y}(t)$ be the output of the model, $y(t)$ the output of the GT, and N the number of records to evaluate, then these metrics are determined as:

$$MAE = \frac{1}{N} \sum_{t=1}^{t=N} |y(t) - \hat{y}(t)| \quad (15)$$

$$MSE = \frac{1}{N} \sum_{t=1}^N (y(t) - \hat{y}(t))^2 \quad (16)$$

$$MAPE = \frac{1}{N} \sum_{t=1}^N \left| \frac{y(t) - \hat{y}(t)}{y(t)} \right| \quad (17)$$

$$RMSE = \sqrt{\frac{\sum_{t=1}^N (y(t) - \hat{y}(t))^2}{N}} \quad (18)$$

4.6. Design (P6)

The design of a NARX implies a network structure that includes all the elements for it to learn to identify a GT [41,42]. For what should be considered: input variables, output variables, an input layer, and an output layer, one or more hidden layers, number of neurons in each hidden layer, transfer functions, such as logsig, purelin, hardlim, satlin, and poslin, activation function, propagation function, training function, such as trainlm and trainbfg, and number of delays of the output signal and input signals. In addition, values for the hyperparameters, the number of epochs, the learning rate, the momentum rate, and the desired final error must be considered.

Some tools allow configuring a NARX topology quickly, such as MATLAB and Python, in their current versions.

4.7. Training and Validation (P7)

In this step, the NARX model designed in step 6 is trained using the training dataset and learning algorithms, such as backpropagation [43,44] and Q-Learning [45]. The model was validated based on cross-validation techniques. These algorithms were implemented in MATLAB [46,47], Weka [44,48], and Python, as well as its frameworks, such as Tensor Flow and Keras [49].

4.8. Fine Tuning (P8)

This step is applied only if the model obtained in step 7 presents unsatisfactory results (metrics).

Fine-tuning is the procedure that aims to determine which hyperparameter values (number of layers, number of nodes per layer, types of activation functions for each layer, initial weights, etc.) provide the model with better results in the learning process [50]. In general, fine-tuning is implemented in various tools that provide machine learning algorithms, some fine-tuning code being GridSearch or RandomizedSearch.

4.9. Testing (P9)

In this step, the successful NARX model obtained in step 7 is tested against the validation dataset, and the error metrics given in step 5 are measured. If the results of this step are satisfactory, that is, an acceptable error, then an adequate NARX model is obtained, and the process is complete; otherwise, it must return to step 6 to build a satisfactory new model. For testing, the same tools are used for training.

Satisfactory NARX model. The final model will be obtained from step 9 if the result of the metrics used is satisfactory. In this way, the model will allow the representation of a GT's behavior with the desired accuracy.

Next, a case study is developed to illustrate the application of the proposed method.

5. Case Study: Single Shaft Open Cycle Gas Turbine

The GT of Lima (GTL) is from 2009 and is a single-axle, heavy-duty, whose characteristics and a view of it are shown in Table 1 and Figure 5, respectively. It works five days a week, 24 h a day, Monday through Friday, and is used by a private electricity generation company in Peru. In the present case study, the GTL output variable of interest is the rotational speed of its axis.

Table 1. GTL Specifications (SIEMENS, SGT6-5000F).

Characteristic	Value
Number of axes	1
Rotational speed	3600 rpm
Compression ratio	15.8
Inlet temperature	599 °C
Outlet temperature	1327 °C
Airflow range	571 kg/s
Power	215 MW
Heat ratio	9643 kJ/kWh
Efficiency	39.5 %

Next, the application of the proposed method step by step for the GTL is shown. The input and output variables of the GTL are shown in Table 2. Data were obtained through transmitters of all the variables, through a data logging system of the control and supervision system SPPA-T3000 control system of SIEMENS, during eight days, every 5 min, considering the cycle of intermittent starts and stops, obtaining 2305 records in Excel,

which are available upon request. Figure 6 shows a sample of the dataset. It has been considered to divide the dataset for training, validation, and testing into 70%, 15%, and 15%, respectively.



Figure 5. A view of the SIEMENS brand gas turbine, model SGT6-5000F [51].

Table 2. GTL input and output variables.

Input Variables	Output Variables
$u_1(t)$: Gas fuel flow (kg/s)	$y_1(t)$: Angular speed (RPM)
$u_2(t)$: Inlet air temperatura ($^{\circ}\text{C}$)	
$u_3(t)$: Barometric pressure (INH ₂ O)	
$u_4(t)$: Megawatt selected (MW)	
$u_5(t)$: Inlet pressure (kPa)	

Since data were correctly obtained by the SCADA data logging system, no integration, cleaning, or data imputation was required. However, to avoid biases of one variable, data were normalized, proportionally passing all data to a scale from 0 to 1. The variation of the values of the variables u_1 , $u_2(t)$, $u_3(t)$, $u_4(t)$, $u_5(t)$ and $y_1(t)$ before normalization is from 10^{-9} to 10.879533 kg/s, from 14.758274 to 21.932644 $^{\circ}\text{C}$, from -0.0604022 to 5.960286 INH₂O, -2.258088 at 198.91522 MW, 98.76779 at 99.30859 kPa, and 3.4272156 at 3634.8547 RPM, respectively.

For the selection of the input and output variables, the InfoGain AttributeEval program of the Weka version 3.4.8 tool was applied, through which the input variables $u_4(t)$ and $u_5(t)$ were discarded. Thus, the selected variables are $u_1(t)$, $u_2(t)$ and $u_3(t)$. In addition, to confirm the selected variables, the Pearson correlation coefficient was used, which shows a value greater than 0.95. Figure 7 shows part of the behavior of the normalized data for the three input variables.

	Inputs			Output
	$u_1(t)$	$u_2(t)$	$u_3(t)$	$y_1(t)$
0	1×10^{-9}	21.0765	−0.024957	3.45883
1	1×10^{-9}	21.0645	−0.030744	3.45978
2	1×10^{-9}	21.0574	−0.030744	3.45628
3	1×10^{-9}	21.0414	−0.030744	3.46483
4	1×10^{-9}	21.0464	−0.030744	3.45628
...	...	...	...	...
2300	10.66150	18.6797	6.720200	3596.88
2301	10.67660	18.4779	6.816040	3600.78
2302	10.65520	18.4930	6.795790	3597.54
2303	10.66810	18.5602	6.720200	3599.22
2304	10.67490	18.5813	6.736110	3599.82

Figure 6. A sample of the dataset.

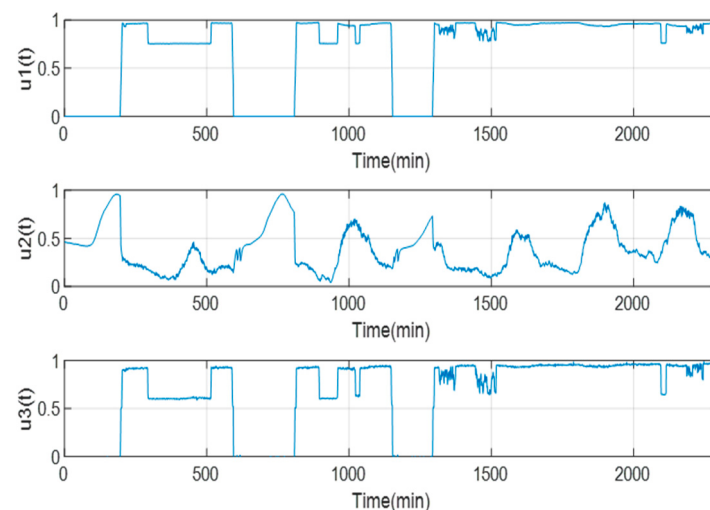


Figure 7. Behavior of the normalized data of the input variables and the output data of the GTL.

Figure 8 shows part of the spin-holds of the angular velocity, which constitute short periods of rotation of the GTL, driven by a prime motor and not by gas, with the purpose of its cooling. Therefore, when it keeps rotating, the fins of the GTL produce cooling. These spin-holds occur between the time intervals of 600–650 min and 1150–1200 min.

The metrics used to evaluate the performance of the NARX models are MSE and MAPE, whose calculations are given in Equations (4) and (5), respectively, and where $\hat{y}(t)$ is the model output and N is the number of records of the dataset to be evaluated.

For the design of the NARX model of the TGL, the MATLAB tool (v. 20–22a) was used with the three selected input variables and the output variable; likewise, one or more hidden layers were considered with a variation of neurons per layer, plus various backpropagation training functions, transfer functions, and one or more lags on the input and output variables. On the other hand, the number of epochs was set equal to 100, the learning rate at 0.01, the momentum rate at 0.9, and the desired error of no more than 1%.

Next, the NARX model established in the design is trained using the training dataset. An iterative process in two loops has been considered to determine the values of the hyperparameters that provide the best results for the model. The first is to tune the hyperparameters in the training and validation process (Step P7), and the second for the testing process (Step P9). In each loop, the number of neurons has been varied from 1 to 30 for each hidden layer. Tests have been carried out with different backpropagation training functions and a combination of other transfer functions for the hidden and output layers. In addition, 1 to 4 delays have been considered.

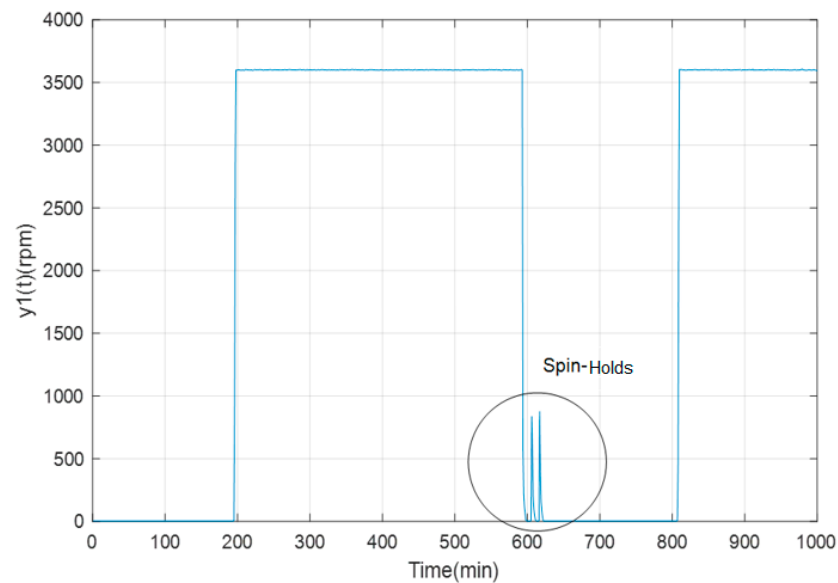


Figure 8. Spin-Holds of the angular velocity of the TGL.

The NARX model, obtained by the training and validation processes and fine-tuning, is tested with the Testing dataset to know the model's performance through the metrics (Step P9). Steps P6–P9 were repeated until a model was obtained that provided satisfactory results in the validation process.

After carrying out all the experiments, the NARX model was finally obtained (Figure 9), which we will call optimal NARX, with the hyperparameters: Number of input and output delays = 2; Number of hidden layers = 2; Number of neurons in the first layer = 15; Number of neurons in the second layer = 3; Transfer function = Trainlm; Activation function = Logsig with a value between 0 and 1; Number of epochs = 100; Desired final error = 10^{-4} ; Learning rate = 0.01; momentum rate = 0.9; Minimum Relative Absolute Error = 1%.

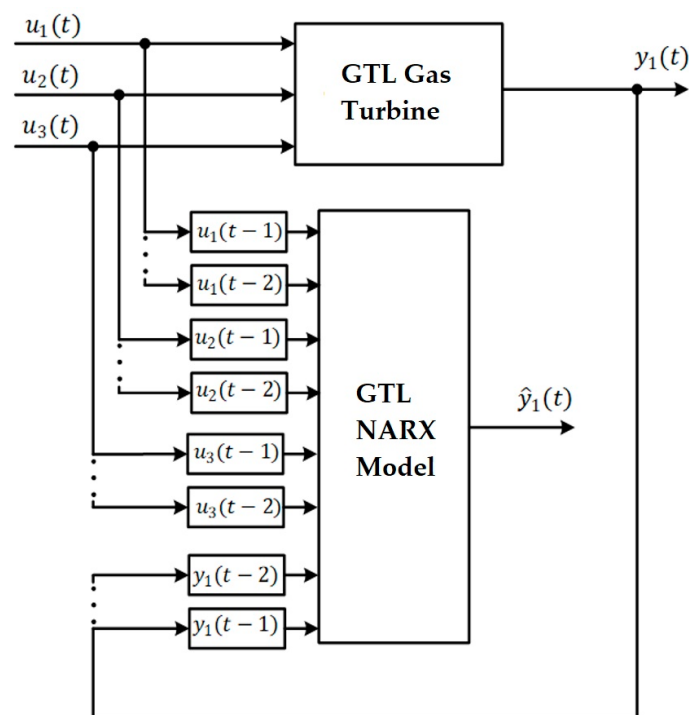


Figure 9. NARX model for the GTL.

Figure 10 indicates the performance of the optimal NARX model (obtained with the proposed method), reaching an MSE of 1.9459×10^{-5} in epoch 17, an RMSE of 0.4411%, and a MAPE of 0.0643.

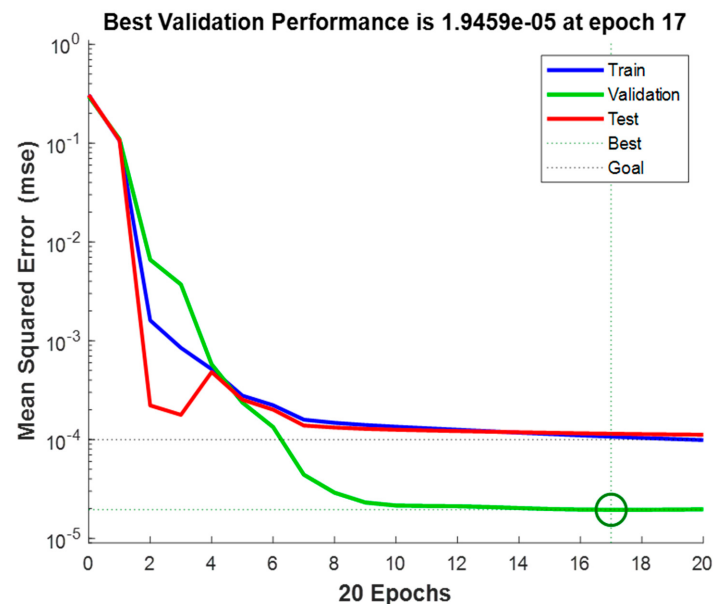


Figure 10. Optimal NARX network performance.

Figure 11 shows the regression graph, which indicates the relationship between the NARX network's output and the system's output (objective). The R-value is an indication of the relationship between outputs and goals. As the figure shows, the R values for all plots are very close to 1. Therefore, the result for each training, validation, and testing data set indicates a good fit. Figure 12 shows the behavior of the real angular velocity (blue) and that obtained by the optimal NARX model (red) for the GTL.

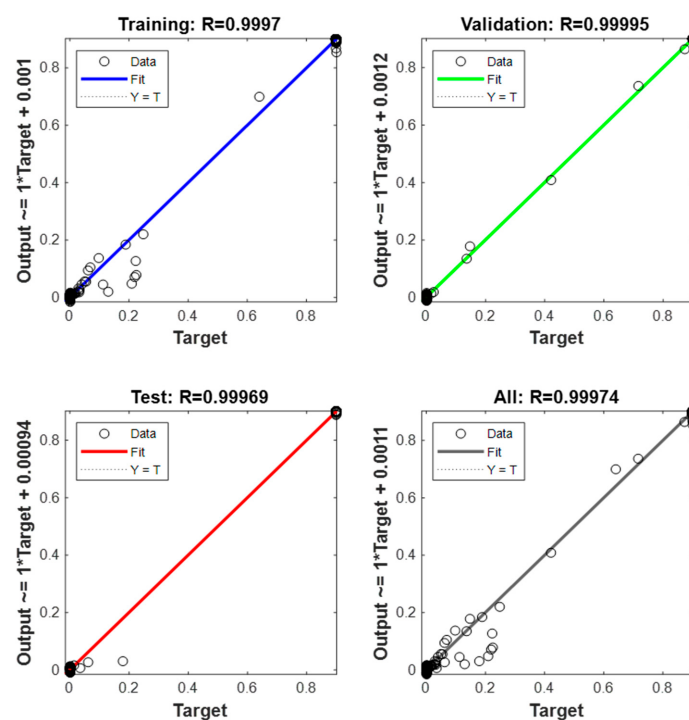


Figure 11. Regression of the optimal NARX network.

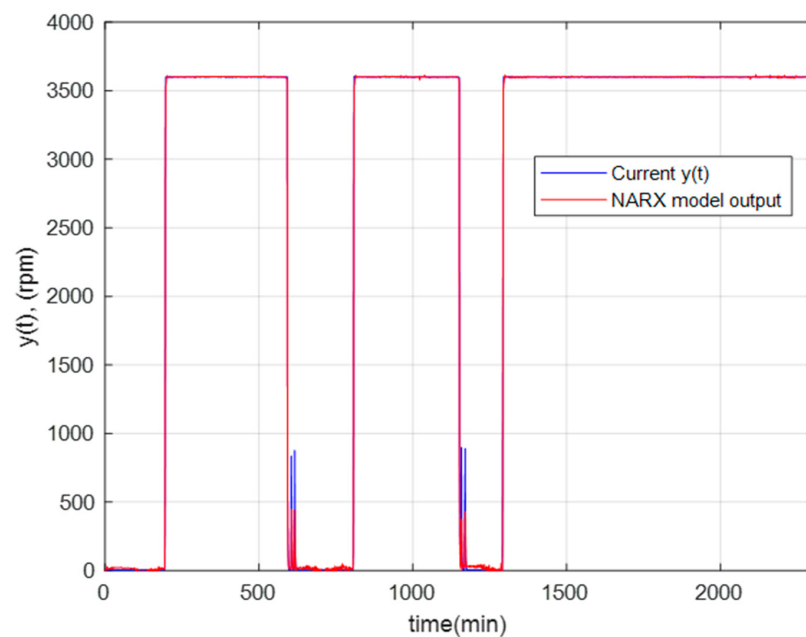


Figure 12. Behavior of the real angular velocity and the optimal NARX model for the GTL.

Figure 13a shows the angular velocity variations for the Testing and the optimal NARX between 3595 and 3605 rpm. It should be noted that the expected angular speed is 3600 rpm. Figure 13a shows that between 2100 and 2250 min, a transient occurs in the turbine load, which was controlled automatically. Likewise, Figure 13b shows the relative error (RE).

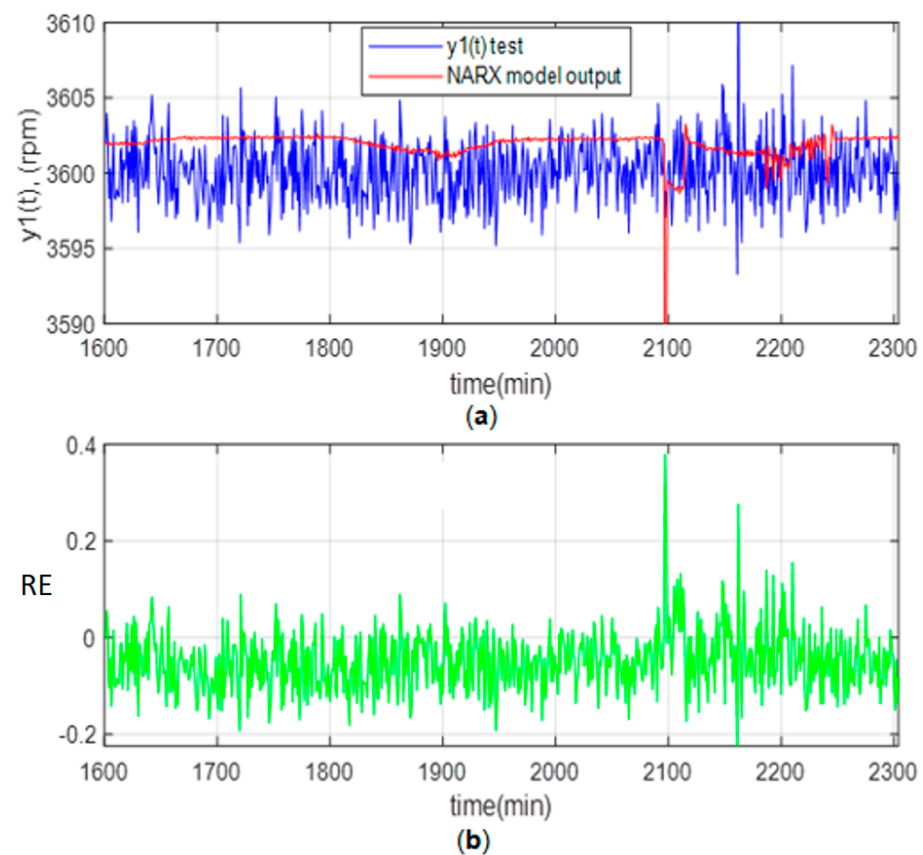


Figure 13. Test results: (a) Angular velocity; (b) Relative Error (RE).

6. Conclusions

This article proposes a systematic method for designing and developing black box models based on NARX neural networks to identify a robust, accurate, and reliable GT model. The proposal consists of 9 well-defined steps that go from identifying GT variables to obtaining a satisfactory NARX model.

In order to validate the proposed method, a case study was carried out using a dataset of 2305 records obtained by measuring the input and output variables of a SIEMENS brand gas turbine, model SGT6-5000F of 215 MW, located in Peru, called GTL. A computer program code was generated and executed in the MATLAB environment (v. 20–22a), following the proposed method step by step, systematically achieving a NARX model for the GTL. The training, validation, and testing results of the generated NARX model did not present significant deviations between the simulated and measured data. The values of the MSE, RMSE, and MAPE metrics during the training, validation, and testing phases were satisfactory, less than 10^{-5} ; 0.5% and 0.06, respectively, values that are very competitive compared to other similar cases.

The case study shows that the proposed method for the development of a NARX model is adequate to identify a GT and to predict with high precision its output parameters based on the changes in the system inputs. Furthermore, since the method is indifferent to the type of gas turbine, it can be applied to predict the behavior of similar gas turbine systems with high accuracy.

This applied research work has as future activities the development of advanced control systems using artificial intelligence techniques. In addition, new neural network architectures can be considered, such as LSTM.

Author Contributions: Conceptualization, R.A. and D.M.; methodology, R.A., D.M. and J.M.M.V. Formal analysis, R.A., D.M. and J.M.M.V. Software, R.A., J.M.M.V. and J.M. Validation, R.A., D.M. and J.M.M.V. Data curation, A.C. and J.M.; writing—original draft preparation, R.A., D.M. and J.M.M.V. Supervision and funding acquisition, R.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially financed by Universidad Nacional de Ingeniería, Lima-Perú.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The private data presented in this study are available on request from the first author.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Singh, R.; Setiawan, A.D. Biomass energy policies and strategies: Harvesting potential in India and Indonesia. *Renew. Sustain. Energy Rev.* **2013**, *22*, 332–345. [CrossRef]
2. Aslanidou, I.; Rahman, M.; Zaccaria, V.; Kyprianidis, K.G. Micro Gas Turbines in the Future Smart Energy System: Fleet Monitoring, Diagnostics, and System Level Requirements. *Front. Mech. Eng.* **2021**, *7*, 676853. [CrossRef]
3. Giampaolo, T. *Gas Turbine Handbook: Principles and Practices*, 3rd ed.; Fairmont Press/CRC Press: Lilburn, GA, USA, 2006. Available online: https://www.academia.edu/8075953/Gas_Turbine_Handbook_Principles_and_Practices (accessed on 14 May 2022).
4. Rodriguez, A.; Garduno, R.; Gerardo, L. PI Fuzzy Gain-Scheduling Speed Control at Startup of a Gas-Turbine Power Plant. *IEEE Trans. Energy Convers.* **2011**, *26*, 310–317. [CrossRef]
5. Bahlawan, H.; Morini, M.; Pinelli, M.; Spina, P.R.; Venturini, M. Development of Reliable NARX Models of Gas Turbine Cold, Warm, and Hot Start-Up. *J. Eng. Gas Turbines Power* **2018**, *140*, 071202-1. [CrossRef]
6. Yoru, Y.; Karakoc, T.H.; Hepbasli, A. Application of Artificial Neural Network (ANN) Method to Exergetic Analysis of Gas Turbines. In Proceedings of the International Symposium on Heat Transfer in Gas Turbine Systems, Antalya, Turkey, 9–14 August 2009. Available online: https://www.researchgate.net/publication/256002552_APPLICATION_OF_ARTIFICIAL_NEURAL_NETWORK_ANN_METHOD_TO_EXERGETIC_ANALYSES_OF_GAS_TURBINES (accessed on 24 November 2022).
7. Boyce, M.P. *Gas Turbine Engineering Handbook*, 4th ed.; Gulf Professional Publishing: Houston, TX, USA, 2012. Available online: <https://www.elsevier.com/books/gas-turbine-engineering-handbook/boyce/978-0-12-383842-1> (accessed on 24 November 2022).

8. Schobeiri, M.T.; Attia, M.; Lippke, C. GETRAN: A generic, modularly structured computer code for simulation of dynamic behavior of aero- and power generation gas turbine engines. *J. Eng. Gas Turbine Power* **1994**, *116*, 483–494. [\[CrossRef\]](#)
9. Agarwal, M. A systematic classification of neural-network-based control. *Control Syst.* **1997**, *17*, 75–93.
10. Agüero, J.L.; Beroqui, M.; Molina, R. Combined cycle plants: Models and in-situ reliability tests. In Proceedings of the Power Engineering Society Summer Meeting. Conference Proceedings, Vancouver, BC, Canada, 15–19 July 2001; pp. 1788–1793, Cat. No. 01CH37262.
11. Bahrami, S.; Ghaffari, A.; Sadati, S.; Thern, M. Identifying a simplified model for heavy duty gas turbine. *J. Mech. Sci. Technol.* **2014**, *28*, 2399–2408. [\[CrossRef\]](#)
12. Ghorbani, H.; Ghaffari, A.; Rahnema, M. Constrained model predictive control implementation for a heavy-duty gas turbine power plant. *WSEAS Trans. Syst. Control* **2008**, *3*, 507–516.
13. Reale, F.; Sannino, R. Numerical Modeling of Energy Systems Based on Micro Gas Turbine: A Review. *Energies* **2022**, *15*, 900. [\[CrossRef\]](#)
14. Verda, V.; Borchellini, R. Exergetic and economic evaluation of control strategies for a gas turbine plant. *J. Energy* **2004**, *29*, 2253–2271. [\[CrossRef\]](#)
15. Rowen, W.I. Simplified mathematical representations of single-shaft gas turbines in mechanical drive service. *Turbomach. Int.* **1992**, *33*, 26–32.
16. Sajad, A.; Mokrova, N.V. Black-Box Identification of Gas Turbine Output Parameters Using Dynamic Neural Network. In Proceedings of the International Multi-Conference on Industrial Engineering and Modern Technologies (FarEastCon), Vladivostok, Russia, 6–7 October 2020. [\[CrossRef\]](#)
17. Salehi, A.; Montazeri-Gh, M. Black box modeling of a turboshaft gas turbine engine fuel control unit based on neural NARX. *J. Eng. Marit. Environ.* **2018**, *233*, 1–8. [\[CrossRef\]](#)
18. Hanachi, H.; Liu, J.; Mechefske, C. Multi-mode diagnosis of a gas turbine engine using an adaptive neuro-fuzzy system. *Chin. J. Aeronaut.* **2017**, *31*, 1–9. [\[CrossRef\]](#)
19. Nikpey, H.; Assadi, M.; Breuhaus, P.; Mørkved, P.T. Experimental evaluation and ANN modeling of a recuperative micro gas turbine burning mixtures of natural gas and biogas. *Appl. Energy* **2014**, *117*, 30–41. [\[CrossRef\]](#)
20. De Giorgi, M.G.; Strafella, L.; Ficarella, A. Neural Nonlinear Autoregressive Model with Exogenous Input (NARX) for Turboshaft Aeroengine Fuel Control Unit Model. *Aerospace* **2021**, *8*, 206. [\[CrossRef\]](#)
21. Ruanoa, A.E.; Fleming, P.J.; Teixeira, C.; Rodríguez, K.; Fonseca, C.M. Nonlinear identification of aircraft gas-turbine dynamics. *Neurocomputing* **2003**, *55*, 551–579. [\[CrossRef\]](#)
22. Asgari, H.; Chen, X.; Morini, M.; Pinelli, M.; Sainudiin, R.; Ruggero, P.; Venturini, M. NARX Models for Simulation of the Start-up Operation of a Single-Shaft Gas Turbine. *Appl. Therm. Eng.* **2015**, *93*, 368–376. [\[CrossRef\]](#)
23. Han, J.C. Fundamental Gas Turbine Heat Transfer. *J. Therm. Sci. Eng. Appl.* **2013**, *5*, 021007–1. [\[CrossRef\]](#)
24. Asgari, H.; Chen, X. *Gas Turbines Modeling, Simulation, and Control Using Artificial Neural Networks*; CRC Press: Boca Raton, FL, USA; Taylor & Francis Group: Abingdon, UK, 2016. Available online: <https://www.taylorfrancis.com/books/mono/10.1201/b18956/gas-turbines-modeling-simulation-control-hamid-asgari-xiaoqi-chen> (accessed on 23 November 2022).
25. Ibrahim, M.A.; Akhrif, O.; Moustapha, H.; Staniszewski, M. Nonlinear generalized predictive controller based on ensemble of NARX models for industrial gas turbine engine. *Energy* **2021**, *230*, 120700. [\[CrossRef\]](#)
26. Moursi, A.S.A.; El-Fishawy, N.; Djahel, S.; Shouman, M.A. Enhancing PM2.5 Prediction Using NARX-Based Combined CNN and LSTM Hybrid Model. *Sensors* **2022**, *22*, 4418. [\[CrossRef\]](#)
27. Bartolini, C.M.; Caresana, F.; Comodi, G.; Pelagalli, L.; Renzi, M.; Vagni, S. Application of artificial neural networks to micro gas turbines. *Energy Convers. Manag.* **2011**, *52*, 781–788. [\[CrossRef\]](#)
28. Rahmounea, M.B.; Hafaifaa, A.; Kouzoua, A.; Chenc, X.; Chaibetd, A. Gas turbine monitoring using neural network dynamic nonlinear autoregressive with external exogenous input modeling. *Math. Comput. Simul.* **2021**, *179*, 23–47. [\[CrossRef\]](#)
29. Asgari, H.; Chen, X.; Mohammad, B.; Sainudiin, R. Artificial Neural Network Based System Identification for a Single-Shaft Gas Turbine. *J. Eng. Gas Turbines Power* **2013**, *135*, 092601–1. [\[CrossRef\]](#)
30. Asgari, H.; Chen, X.; Menhaj, M.B.; Sainudiin, R. Modeling and Simulation of the Start-Up Operation of a Heavy-Duty Gas Turbine by Using NARX Models. In Proceedings of the ASME Turbo Expo, American Society of Mechanical Engineers. New York, NY, USA, 2014; Volume 3A. [\[CrossRef\]](#)
31. Nikpey, H.; Assadi, M.; Breuhaus, P. Development of an optimized artificial neural network model for combined heat and power micro gas turbines. *Appl. Energy* **2013**, *108*, 137–148. [\[CrossRef\]](#)
32. Diaconescu, E. The use of NARX Neural Networks to predict Chaotic Time Series. *WSEAS Trans. Comput. Res.* **2008**, *3*, 182–191.
33. McKinney, W. *Python for Data Analysis, Data Wrangling with Pandas, NumPy, and IPython*, 2nd ed.; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2018.
34. Zaki, M.J.; Meira, W. *Data Mining and Analysis, Foundations and Algorithms*; Cambridge University Press: Cambridge, UK, 2014; pp. 202–208. ISBN 978-0-521-76633-3.
35. García, S.; Luengo, J.; Herrera, F. *Data Preprocessing in Data Mining*; Springer International Publishing: Cham, Switzerland, 2015; pp. 140–151. [\[CrossRef\]](#)
36. Spiegel, M.R. *Theory and Problems of Statistics*; McGraw Hill: New York, NY, USA, 1999; pp. 311–318. ISBN -0-07-060282-6.

37. Moon, S.W.; Kim, T.S. Advanced Gas Turbine Control Logic Using Black Box Models for Enhancing Operational Flexibility and Stability. *Energies* **2020**, *13*, 5703. [CrossRef]
38. Alsarayreh, M.; Mohamed, O.; Matar, M. Modeling a Practical Dual-Fuel Gas Turbine Power Generation System Using Dynamic Neural Network and Deep Learning. *Sustainability* **2022**, *14*, 870. [CrossRef]
39. PAwełczyk, M.; FularA, S.; SePe, M.; De lucha, A.; Badora, M. Industrial gas turbine operating parameters monitoring and data-driven prediction. *Eksploat. I NiEzawodNosc–MaiNtENaNcE REliability* **2020**, *22*, 2020. [CrossRef]
40. De Giorgi, M.G.; Quarta, M. Hybrid MultiGene Genetic Programming-Artificial neural networks approach for dynamic performance prediction of an aeroengine. *Aerosp. Sci. Technol.* **2020**, *103*, 105902. [CrossRef]
41. Mohamed, O.; Khalil, A. Progress in Modeling and Control of Gas Turbine Power Generation Systems: A Survey. *Energies* **2020**, *13*, 2358. [CrossRef]
42. Hamid, A.; Ory, E. Prediction of Dynamic Behavior of a Single Shaft Gas Turbine Using NARX Models. In Proceedings of the ASME Turbo Expo 2021: Turbomachinery Technical Conference and Exposition, Virtual, Online, 7–11 June 2021. [CrossRef]
43. Lazzaretto, A.; Toffolo, A. Analytical and Neural Network Models for Gas Turbine Design and Off-Design Simulation. *J. Appl. Thermodyn.* **2001**, *4*, 173–182.
44. Witten, I.H.; Frank, E.; Hall, M.A. *Data Mining: Practical Machine Learning Tools and Techniques*, 3rd ed.; Morgan Kaufmann Publishers: Burlington, MA, USA, 2011; Chapter 6; pp. 235–241.
45. Géron, A. *Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow*, 2nd ed.; O'Reilly Media, Inc.: Sebastopol, CA, USA, 2019; Chapter 18.
46. Tsogkas, S. *matlab-utils/crossValidation.m*; GitHub, Inc.: San Francisco, CA, USA, 2014. Available online: <https://github.com/tsogkas/matlab-utils/blob/master/crossValidation.m> (accessed on 26 November 2022).
47. MATLAB R2022b. Cross-Validate Machine Learning Model. MathWorks Help Center. 2022. Available online: <https://es.mathworks.com/help/stats/classificationsvm.crossval.html> (accessed on 26 November 2022).
48. Stack Overflow. Cross Validation in Weka. Available online: <https://stackoverflow.com/questions/10437677/cross-validation-in-weka> (accessed on 26 November 2022).
49. Unpingco, J. *Python for Probability, Statistics, and Machine Learning*; Springer International Publishing: Cham, Switzerland, 2016; Chapter 4; pp. 215–222.
50. Dreyfus, G. *Neural Networks, Methodology and Applications*, 2nd ed.; Springer: Berlin/Heidelberg, Germany, 2004; Chapter 4; p. 260.
51. Foust, A. *Siemens Expands Footprint in the Middle East with SGT6-5000F Power Plant Solutions and New Gas Turbine Manufacturing Facility*; Siemens AG Energy Sector: Munich, Germany, 2012. Available online: <https://www.siemens-energy.com/global/en/offerings/power-generation/gas-turbines/sgt6-5000f.html> (accessed on 26 November 2022).

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.