# XGBOOST APPLIED TO MODEL AN AIRCRAFT ENGINE RPM-FUEL RELATIONSHIP THROUGH NONLINEAR BLAC-BOX SYSTEM IDENTIFICATION

ELCIO BARDELI JR., ROBERTO Z. FREIRE

*Industrial and Systems Engineering Graduate Program, Pontifical Catholic University of Paraná (PUCPR)*
*Imaculada Conceição St., 1155, Zip-code 80215-901, Curitiba, PR, Brazil*
*E-mails:* elcio.junior@qflux.com.br; roberto.freire@pucpr.br

ALIREZA R. TAVAKOLPOUR-SALEH

*Department of Mechanical and Aerospace Engineering Shiraz University of Technology*
*Modarres Blvd. P.O. Box,71555-313, Shiraz, Iran*
*E-mails:* tavakolpour@sutech.ac.ir

LEANDRO DOS S. COELHO

*Industrial and Systems Engineering Graduate Program, Pontifical Catholic University of Paraná (PUCPR)*
*Imaculada Conceição St., 1155, Zip-code 80215-901, Curitiba, PR, Brazil and*
*Department of Electrical Engineering, Federal University of Paraná (UFPR)*
*Coronel Francisco Heráclito dos Santos St., 100, Zip-code 81531-980, Curitiba, PR, Brazil*
*E-mails:* leandro.coelho@pucpr.br

HELON V. H. AYALA

*Department of Mechanical Engineering, Pontifical Catholic University of Rio de Janeiro (PUC-Rio)*
*Marques de São Vicente St., 255, Zip-code 22451-000, Rio de Janeiro, RJ, Brazil and*
*Industrial and Systems Engineering Graduate Program, Pontifical Catholic University of Paraná (PUCPR)*
*Imaculada Conceição St., 1155, Zip-code 80215-901, Curitiba, PR, Brazil*
*E-mails:* helon@puc-rio.br

**Abstract**— In the aerospace industry, it is important to understand the behavior of turbojet engines, which is considered a complex system and, in many cases, engine's performance data is not available, causing lack of knowledge about the system performance. In order to meet this requirement, black-box system identification strategies are been successfully applied complex engineering problems. Based on the previous assumptions, this paper presents a parametric black-box system identification approach based on data collected from a turbojet engine through an experimental investigation. In order to model the jet engine, Extreme Gradient Boosting (XGBoost) technique, which combines a series of regression trees, was associated to a nonlinear autoregressive exogenous (NARX) model structure. In order to evaluate the number of regressors associated to the model, the relative importance of each parameter was evaluated. Finally, the model was validated considering the coefficient of determination ($R^2$) and on a statistical test based on the statistical correlation of the residuals of the model.

**Keywords**— turbojet engine, system identification, parametric model, XGBoost technique.

**Resumo**— Na indústria aeroespacial, compreender o comportamento dos motores a jato, os quais são considerados sistemas de grande complexidade, não é uma tarefa simples, já que, em muitos casos, existem incertezas sobre o desempenho desses sistemas quando submetidos a diferentes condições de teste. Levando-se em conta o objetivo de entender o correto funcionamento das turbinas de aeronaves e projetar equipamentos mais eficientes, este trabalho propõe uma técnica de identificação paramétrica de sistemas baseada em uma abordagem caixa preta. Levando-se em conta um estudo prévio onde foram coletados dados de uma turbina em uma plataforma de testes, propõe-se a utilização da técnica *Extreme Gradient Boosting* (XGBoost), a qual combina uma série de árvores de regressão, com uma estrutura de modelo não linear autoregressivo e exógeno (NARX) para reproduzir o comportamento do motor a jato. Para avaliar o desempenho do modelo proposto, foi utilizado o coeficiente de determinação ($R^2$) e testes baseados na correlação estatística dos resíduos do modelo.

**Palavras-chave**— motor a jato, identificação de sistemas, modelo paramétrico, técnica XGBoost.

## 1 Introduction

Gas turbines have been widely used in industrial, marine, and especially in aerospace applications. This type of engine can be defined as a complex system comprising multiple subsystems that have dependent interactions. Additionally, gas turbines are highly valuable assets in aircraft, where large sums are spent in maintenance support and logistics (Zaidan, Harrison, Mills, & Fleming, 2015).

By considering the intense competition that characterize both energy and aerospace sectors, gas turbine industry currently faces new challenges of increasing operational flexibility, reducing operating costs, improving reliability and availability while mitigating the environmental impact (Tahan, Tsoutsanis, Muhammad, & Abdul Karim, 2017).

According to (Filippone & Bojdo, 2018), from engine design to meeting more stringent targets that are agreed at the international level. In this way, proper reproduction of aircraft engine behavior and engine performance evaluation are important issues in aerospace engineering area.

Black-box system identification strategies are being applied to reproduce the operation of gas turbine engines, especially in the cases where data about the engine performance is not available, making it nearly impossible to make any predictions about its behavior. One example is the work presented by (Lazzaretto & Toffolo, 2001), where authors adopted an artificial neural network (NN) model in order to overcome the lack of knowledge about the system. Another application of NNs was presented by (Talaat, Gobran, & Wasfi, 2018), where the objective was the development of a diagnosis system for an electrical power plant gas turbine. A hybrid thermodynamic model has been used to simulate gas turbine performance as well as the deterioration of engine components, and posteriorly, the data generated by the thermodynamic model was adopted for training the neural network for fault detection. Finally, the black-box model presented promising results when tested in a real system. However, distinct black-box system identification strategies are available for system identification purposes. One technique which is attracting attention of academic and industrial researchers is the Extreme Gradient Boosting (XGBoost) (Chen, He, Benesty, Khotilovich, & Tang, 2018).

XGBoost is a library designed and optimized for boosting trees algorithms, based on gradient boosting trees model and has been originally proposed by (Friedman, 2002). It is used for supervised learning problems, the training data with multiple features is used to predict a target behavior (Mostakim, 2016). Some applications of XGBoost were described in the sequence of this introduction in order to emphasize the application potential and results reached by this technique.

In a first application, three machine learning techniques (regression tree, random forest and gradient boosting machine) were applied to predict the total flavonoid content in 22 red wine grape cultivars (Brillante et. al., 2015). Flavonoid are a class of bioactive compounds largely represented in grapevine and wine that affect the sensory quality of fruits and vegetables, and derived products contents. In this first analysed case, gradient boosting machine overcame the other two method when the coefficient of determination ($R^2$), and root mean squared error (RMSE) were evaluated as performance criteria in both training and validation phases.

In the work proposed by (Persson, Bacher, Shiga, & Madsen, 2017), multi-site (42 individual PV rooftop installations) prediction of solar power generation on a forecast horizon of one to six hours has been performed using single-site linear autoregressive analysis and gradient boosting method, which is a prior version of XGBoost. In this study, gradient boosted technique shows competitive results in terms of RMSE on all forecast horizons.

The comparative study presented by (Fan et al., 2018), Support Vector Machine (SVM) and XGBoost were compared in order to predict daily global solar radiation using temperature and precipitation in China. In this study, XGBoost overcame SVM in terms of accuracy, stability and computational speed.

On the other hand, system identification algorithms have been extensively used in the field of machine learning. Black-box modeling based on NNs have been thoroughly discussed in the literature since some decades (Juditsky et. al., 1995; Sjöberg et. al., 1995). Recently, Bagherzadeh (2018) proposes the identification of aircraft flight dynamics with NNs with specialized architecture. In (Tunjuni, 2016) the use of transfer functions to approximate NNs is proposed, where the NN weights are directly linked to transfer function parameters. As will be shown in the next sections, the vectorial mapping from the system's lagged inputs/outputs to predicted outputs can be directly treated with machine learning algorithms devised for regression. To bridge the methods conceived for machine learning to the field of dynamic systems modeling, however, is not straightforward and tests should be performed whenever e.g. corrupted measurements take place or free-run simulation is aimed.

Based on the impact of gradient boosting prediction mentioned on the previous studies, and the importance of modelling gas turbines in engineering applications, this work extends a previous study presented by (Tavakolpour-Saleh, Nasib, Sepasyan, & Hashemi, 2015). In the present paper, we propose a distinct system identification approach in order to reproduce an aircraft gas turbine behavior.

In the original case study, an experimental procedure was carried out to collect data from a turbojet engine, and two models, focusing on parametric and nonparametric techniques, were proposed to reproduce the engine behavior, using as input fuel flow rate (kg/h) and rotation speed of the engine (RPM). In the present paper, Extreme Gradient Boosting (XGBoost), which combines several different XGBoost tree models, was adopted in the system identification procedure. The contribution of this work consists on testing the XGBoost algorithm for the purpose of dynamic systems modeling to the same set of data used in (Tavakolpour-Saleh, Nasib, Sepasyan, & Hashemi, 2015), with linear models and NN based-NARX models.

The remainder of the paper is organized as follows. Section 2 describes the mathematics behind XGBoost algorithm. Section 3 presents the model structure adopted in this work, followed by the metrics assumed to validate the proposed strategy, and system identification procedures. Section 4 describes the case study based on the turbo jet engine and also presents the results of the system identification procedures. Finally, section 5 addresses the conclusions and future works of this research.

## 2 Extreme Gradient Boosting (XGBoost)

Boosting is one of the most powerful learning ideas introduced in the last twenty years. It was originally designed for classification problems, but it can profitably be extended to regression and system identification as well (Hastie, Tibshirani, & Friedman, 2009).

XGBoost combines a series of regression trees to constitute a model. These structures can capture complex interaction in the dataset. It has both linear model solver and tree learning algorithms. This makes XGBoost at least 10 times faster than existing gradient boosting implementations. It supports various objective functions, including regression, classification and ranking.

As presented in (Chen, Jiang, Zheng, & Chen, 2018), in most cases, a single regression tree is inadequate for a good regression model. In order to improve the method performance, the idea is to combine a number of regression trees into an ensemble.

Separate trees can be added together in the same way those individual predictors can be added together in a regression model. In XGBoost, $K$ additive functions are used to predict the output as presented in the sequence:

$$\hat{y}_i = \sum_{k=1}^{K} f_k(x_i) \tag{1}$$

where $\hat{y}_i$ is the i-*th* estimated output, $f_k \in \mathcal{F}$, and $\mathcal{F} = \{f(x) = w_{q(x)}\}$ $(q : \Re^m \to T, w \in \Re^T)$ is the space of regression trees. Each $f_k$ refers to an independent tree structure $q$ and leaf weights $w$. Besides, $q$ denotes the structure of each tree that maps an example to the corresponding leaf index, where $T$ represents the number of leaves in the tree.

Generally, the regression tree $T(x)$ can be expressed as (J. Hastie, Tibshirani, & Friedman, 2009):

$$T(x) = \sum_{m=1}^{M} c_m I \tag{2}$$

In Eq. (2), $x \in R_m$, and the data is split into $M$ regions $R_1, R_2, \ldots, R_M$. The parameter $c_m$ is the response in region $m$. Under the mean squared error (MSE) loss function, $c_m$ can be estimated by:

$$\hat{c}_m = E[y_i | x_i \in R_m] \tag{3}$$

The final optimization target can be represented as:

$$\mathcal{L}_t = \sum_i l(\hat{y}_i, y_i) \sum_k \Omega(f_k) \tag{4}$$

where

$$\Omega(f_k) = \gamma T + \frac{1}{2}\lambda \|w\|^2 \tag{5}$$

In the previous equations, $l$ is a differentiable convex loss function that measures the difference between the predicted output $\hat{y}_i$ and the measured output $y_i$. In the regression task, $l$ is normally set to squared-error function. $\Omega(f)$ is the penalty term, which avoids the problem of over-fitting in the model, where $\gamma$ is a threshold for the gain and $\lambda$ is a regularization parameter. By assuming this penalty term, a model with simple and predictive functions is more likely to be selected.

In order to train all trees at the same time, an additive strategy is introduced. The trees that have been trained are fixed, then add one new tree at a time. Assume the prediction value at step $t$ is denoted as $\hat{y}_i{}^{(t)}$, then Eq. (4) can be written as:

$$\mathcal{L}_t = \sum_i^n \left[ g_i f_k(x_i) + \frac{1}{2} h_i f_k{}^2(x_i) \right] \Omega(f_k) \tag{6}$$

where $g_i = \partial_{\hat{y}_i^{(t-1)}} l\big(y_i, \hat{y}_i(t-1)\big)$, and the term $h_i = \partial^2_{\hat{y}_i^{(t-1)}} l\big(y_i, \hat{y}_i(t-1)\big)$. Equation (6) can be reformulated as:

$$\mathcal{L}_t = \sum_{j=1}^{T} \left[ \begin{array}{c} \left(\sum_{i \in I_j} g_i\right) w_j \\ + \frac{1}{2}\left(\sum_{i \in I_j} h_i + \lambda\right) w_j^2 \end{array} \right] + \gamma T \tag{7}$$

where $I_j = \{i | q(x_i) = j\}$. Given a fixed tree structure $q(x)$, the optimal leaf weight scores on each leaf node ${}^*j$, and the extreme value of $\mathcal{L}_t$ can be solved as:

$$w_j^* = \frac{\sum_{i \in I_j} g_i}{\sum_{i \in I_j} h_i + \lambda} \tag{8}$$

$$\mathcal{L}_t(q) = -\frac{1}{2} \sum_{j=1}^{T} \left[ \frac{\left(\sum_{i \in I_j} g_i\right)^2}{\sum_{i \in I_j} h_i + \lambda} \right] + \gamma T \tag{9}$$

## 3 Black-box System Identification

Black-box models such as nonlinear autoregressive exogenous (NARX) models are employed when access to the complicated dynamic equations of the system is not possible or struggling with them is difficult and undesirable (Jelali & Kroll, 2012). This type of model structure has a recurrent dynamic nature and is commonly used in time-series modeling (Asgari et al., 2016).

NARX includes feedback connections enclosing several layers of the network. A nonlinear ARX model (NARX) can be defined as an extension of the linear model found in (Ljung, 1987), which structure can be modified to create a nonlinear form presented in Eq. (10):

$$\hat{y}(t) = \phi\begin{pmatrix} y(t-1), y(t-2), \dots, \\ y(t-n_a), u(t-n_k-1), \\ u(t-n_k-2), \dots, \\ u(t-n_k-n_b) \end{pmatrix}. \quad (10)$$

The nonlinear function $\phi$ can be expressed in terms of the model regressors, and the nonlinear mapping can be performed using nonlinear estimators. The model structure is entirely defined by the integers $n_a$, $n_b$, and $n_k$, where $n_a$ represents the number of lagged outputs, $n_b$ is the number of lagged exogenous inputs, and $n_k$ is the time delay of the systems. The above mentioned orders define the complexity of the model with respect to the dynamics of the system to be represented.

### *3.1 Model Validation Metrics*

Taking into account that the autocorrelation of the residuals and the cross-correlation of the residuals and inputs are not sufficient to validate nonlinear models (Billings & Voon, 1983). This work assumes a series of tests that indicate the validity of a model based on the calculation of the correlation function coefficients, for systems on the form presented in Eq. (10). The set of tests are given by:

$$\begin{cases} \varphi_{\xi\xi}(\tau) = \delta(\tau), \\ \varphi_{\xi\xi}(\tau) = 0, \forall\tau, \\ \varphi_{\xi(\xi u)}(\tau) = 0, \tau \geq 0, \\ \varphi_{(u^2)'\xi}(\tau) = 0, \forall\tau, \\ \varphi_{(u^2)'\xi^2}(\tau) = 0, \forall\tau, \end{cases} \quad (11)$$

where $\xi(t) = \hat{y}(t) - y(t)$, $\delta(.)$ Is the Kronecker delta function, $(u^2)'(t) = u(t)^2 - \bar{u}^2$, $\xi u = \xi(t+1)u(t+1)$, and $\varphi_{ab}$ is the normalized cross-correlation function between two sequences $\{a\}$ and $\{b\}$, which can be described as (Billings, 2013):

$$\varphi_{ab}(\tau) = \frac{\sum_{t=1}^{N-\tau}[a(t)-\bar{a}]\left[b(t+\tau)-\bar{b}\right]}{\left[\sum_{t=1}^{N}[a(t)-\bar{a}]^2 \sum_{t=1}^{N}\left[b(t)-\bar{b}\right]^2\right]^{1/2}} \quad (12)$$

In order to evaluate the model performance, the coefficient of determination, which was described in Eq. (13), was considered for both training ($R_{tr}^2$) and validation ($R_{val}^2$) phases.

$$\mathrm{R}^2 = 1 - \frac{\sum_{t=1}^{Ns}[y(t) - \hat{y}(t)]^2}{\sum_{t=1}^{Ns}[y(t) - \bar{y}]^2} \quad (13)$$

### *3.2 XGBoost for System Identification*

In the present subsection we detail how the ensemble learning approach devised with XGBoost can be applied to the identification of dynamic systems.

It is possible to see that Eq. (1) we have the regression performed by the sum of a vectorial mapping from the model's input vector to the predicted output. On the other hand, by inspecting Eq. (10), it can be observed that, for the system identification procedure, the model inputs are defined with past measured input/output values. Thus by setting the features of the machine learning model as the lagged inputs/outputs and the targets as the current outputs we can create the data-driven dynamic system model.

In the following section we depict the results of the application of the aforementioned strategy to dynamic systems modeling.

## 4 Results

This section starts presenting the case study adopted for the system identification approach. Posteriorly, results were reported and the proposed model was evaluated.

### *4.1 Case study description*

The turbo jet tester, which generated data for this study was originally presented in (Tavakolpour-Saleh et al., 2015). Figure 1 presents the experimental plant, while Figs. 2 and 3 show both input and output datasets that were acquired from the turbo jet engine.

### *4.2 Numerical experiments*

The modelling procedure described in Sections 2 and 3, has been applied to the aforementioned case study. Specifically, the XGBoost algorithm has been used to proceed the dynamic modeling of the aircraft engine relation to RMP and fuel input. $Y(t)$ and $u(t)$ are used as the model lagged inputs for the purpose of modeling following the NARX structure.

The XGBoost version implemented in R programing language (Chen et al., 2018) was used. We employed 10 lags in both system's input and output to build the regression matrix for model construction. The parameters for the XGBoost are stated in Table 1, where the cross-validation function has been used to determine the number of rounds.

The variable importance is plotted in Fig. 4. It is possible to see that the most important variables are the first lagged input and output of the system, which were assumed in the model structure.

Figure 1. Turbo jet engine tester (Tavakolpour-Saleh et al., 2015).

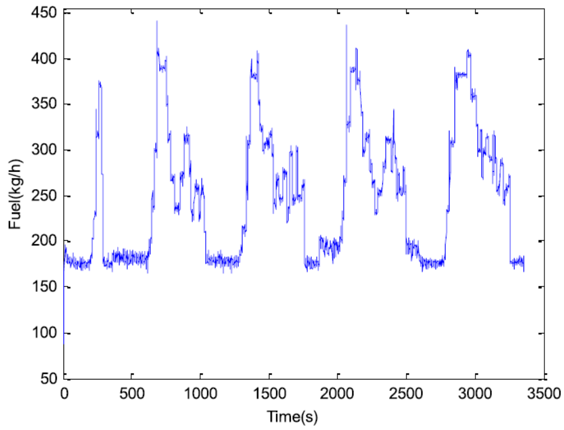


Figure 2. Input of the system: fuel value vs. time (Tavakolpour-Saleh et al., 2015).

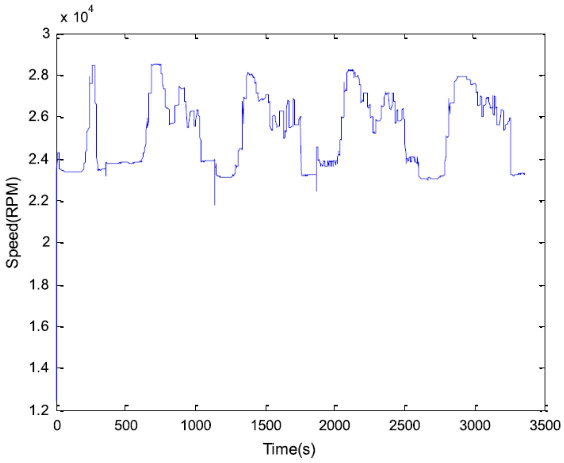


Figure 3. Output of the system: resolutions per minute vs. time (Tavakolpour-Saleh et al., 2015).

Table 1. Parameters for the XGBoost algorithm.

| Parameter | Value |
|---|---|
| Booster | gbtree |
| Objective | reg:linear |
| eta | 0.01 |
| Max_depth | 10 |
| Min_child_weight | 1 |
| Subsample | 1 |
| Colsample_bytree | 1 |
| nrounds | 1076 |

Values for the coefficient of determination of $R^2_{tr} = 0.9999672$ and $R^2_{val} = 0.9948271$ were obtained for training and validation phases, respectively. According to (Schaible, Xie, & Lee, 1997), values for $R^2$ higher than 0.9 are enough to express a model in system identification field.

The statistical tests as in (Chen & Billings, 1992) have been used in order to validate the model statistical properties. In Fig. 5 we can see the good adherence to the model validation tests. In Fig. 6, it is possible to inspect the model output versus the measured data. As it can be verified in Fig. 6, the model reasonably represents the system under study during both training and validation phases.

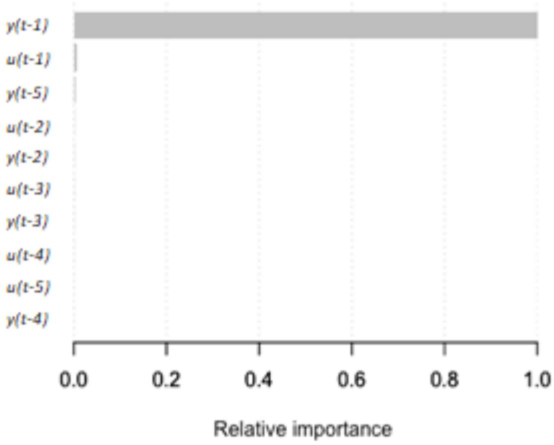


Figure 4. Relative importance of model input variables.

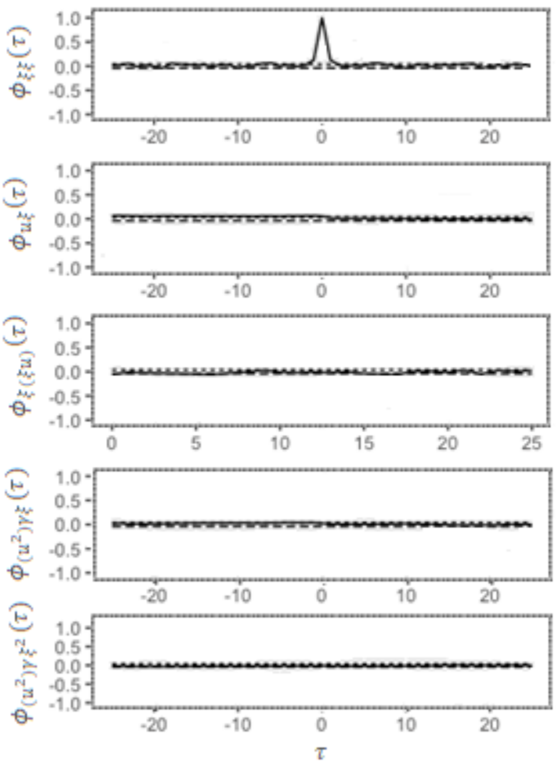


Figure 5. Statistical tests based on correlation for the model.

## 5 Conclusion

In the present work, it was presented a system identification procedure solely based on measured data obtained from an aircraft engine with an ensemble learning method called Extreme Gradient Boosting (XGBoost). Results indicated that the tool is able to capture and adequately represent the system dynamics as the error quantitative and statistical tests have indicated.

Future work will aim at the adaptation of the workflow possible in ensemble learning for automating the procedure of model construction. For example, we can use the variable importance which is automatically extracted from the trees, to perform input selection for black-box nonlinear models. In addition, other operating envelopes of the aircraft model should be explored so that a switching logic can be implemented to simulate the engine in various operating conditions such as different values altitude and flight speed.

## References


Asgari, H., Chen, X., Morini, M., Pinelli, M., Sainudiin, R., Ruggero, P., & Venturini, M. (2016). NARX models for simulation of the start-up operation of a single- shaft gas turbine, *93*, 368–376. https://doi.org/10.1016/j.applthermaleng.2015.09.074

Billings, S. A. (2013). *Nonlinear system identification: NARMAX methods in the time, frequency, and spatio-temporal domains. Nonlinear System Identification: NARMAX Methods in the Time, Frequency, and Spatio-Temporal Domains*. https://doi.org/10.1002/9781118535561

Billings, S. A., & Voon, W. S. F. (1983). Structure detection and model validity tests in the identification of nonlinear systems. *IEE Proceedings D - Control Theory and Applications*, *130*(4), 193–199. https://doi.org/10.1049/ip-d.1983.0034

Chen, K., Jiang, J., Zheng, F., & Chen, K. (2018). A novel data-driven approach for residential electricity consumption prediction based on ensemble learning. *Energy*, *150*, 49–60. https://doi.org/S0360544218302561

Chen, S., & Billings, S. A. (1992). Neural networks for nonlinear dynamic system modelling and identification. *International Journal of Control*, *56*(2), 319–346. https://doi.org/10.1080/00207179208934317

Chen, T., He, T., Benesty, M., Khotilovich, V., & Tang, Y. (2018). CRAN - Package XGBoost: Extreme Gradient Boosting. Retrieved from https://cran.r-project.org/package=XGBoost

Fan, J., Wang, X., Wu, L., Zhou, H., Zhang, F., Yu, X., … Xiang, Y. (2018). Comparison of Support Vector Machine and Extreme Gradient Boosting for predicting daily global solar radiation using temperature and precipitation in humid subtropical climates: A case study in China. *Energy Conversion and Management*, *164*(February), 102–111. https://doi.org/10.1016/j.enconman.2018.02.087

Filippone, A., & Bojdo, N. (2018). Statistical model for gas turbine engines exhaust emissions. *Transportation Research Part D*, *59*, 451–463. https://doi.org/10.1016/j.trd.2018.01.019

Friedman, J. H. (2002). Stochastic gradient boosting. *Computational Statistics and Data Analysis*, *38*(4), 367–378. https://doi.org/10.1016/S0167-9473(01)00065-2

Hastie, J., Tibshirani, R., & Friedman, J. (2009). *The Elements of Statistical Learning - Data Mining, Inference, and Prediction, Second Edition. Springer series in statistics*. https://doi.org/10.1007/978-0-387-84858-7

Hastie, T., Tibshirani, R., & Friedman, J. (2009). The Elements of Statistical Learning. *Elements*, *1*, 337–387. https://doi.org/10.1007/b94608

Jelali, M., & Kroll, A. (2012). *Hydraulic servo-systems: modelling, identification and control*. (S. S. & B. Media, Ed.).

Brillante, L., Gaiotti, F., Lovat, L., Vincenzi, S. Giacosa, S., Torchio, F., Segade, S.R., Rolle, L., Tomasi, D. (2015) Investigating the use of gradient boosting machine, random forest and their ensemble to predict skin flavonoid content from berry physical–mechanical characteristics in wine grapes, *Computers and Electronics in Agriculture*, 117, 186-193.

Lazzaretto, A., & Toffolo, A. (2001). Analytical and neural network models for gas turbine design and off-design simulation. *International Journal of Applied Thermodynamics*, *4*(4), 173–182. https://doi.org/10.5541/IJOT.1034000078

Ljung, L. (1987). Ljung L System Identification Theory for User.pdf. *PTR Prentice Hall Upper Saddle River NJ*. https://doi.org/10.1016/0005-1098(89)90019-8

Mostakim, M. (2016). *Prediction on Large Scale Data Using Extreme Gradient Boosting*. BRAC University, Dhaka, Bangladesh.

Persson, C., Bacher, P., Shiga, T., & Madsen, H. (2017). Multi-site solar power forecasting using gradient boosted regression trees. *Solar Energy*, *150*, 423–436. https://doi.org/10.1016/j.solener.2017.04.066

Schaible, B., Xie, H., & Lee, Y.-C. (1997). Fuzzy logic models for ranking process effects. *IEEE Transactions on Fuzzy Systems*, *5*(4), 545–556.

Tahan, M., Tsoutsanis, E., Muhammad, M., & Abdul Karim, Z. A. (2017). Performance-based health monitoring, diagnostics and prognostics for condition-based maintenance of gas turbines: A review. *Applied Energy*, *198*, 122–144. https://doi.org/10.1016/j.apenergy.2017.04.048

Talaat, M., Gobran, M. H., & Wasfi, M. (2018). A hybrid model of an artificial neural network with thermodynamic model for system diagnosis of electrical power plant gas turbine. *Engineering Applications of Artificial Intelligence*, *68*(September 2017), 222–235. https://doi.org/10.1016/j.engappai.2017.10.014

Tavakolpour-Saleh, A. R., Nasib, S. A. R., Sepasyan, A., & Hashemi, S. M. (2015). Parametric and nonparametric system identification of an experimental turbojet engine. *Aerospace Science and Technology*, *43*, 21–29. https://doi.org/10.1016/j.ast.2015.02.013

Zaidan, M. A., Harrison, R. F., Mills, A. R., & Fleming, P. J. (2015). Bayesian Hierarchical Models for aerospace gas turbine engine prognostics. *Expert Systems with Applications*, *42*(1), 539–553. https://doi.org/10.1016/j.eswa.2014.08.007

Bagherzadeh , S. A. (2018). Nonlinear aircraft system identification using artificial neural networks enhanced by empirical mode decomposition. *Aerospace Science and Technology*, 75, 155-171. https://doi.org/10.1016/j.ast.2018.01.004

Tutunji, T. A. (2016). Parametric system identification using neural networks. *Applied Soft Computing*, 47, 251-261. https://doi.org/10.1016/j.asoc.2016.05.012.

Juditsky, A., Hjalmarsson, H., Benveniste, A., Delyon, B., Ljung, L. Sjöberg, Zhang, J. Q. (1995) Nonlinear black-box models in system identification: Mathematical foundations, *Automatica*, 31(12), 1725-1750. https://doi.org/10.1016/0005-1098(95)00119-1.

Sjöberg, J., Zhang, Q., Ljung, L., Benveniste, A., Delyon, B., Glorennec, P., Hjalmarsson, H., Juditsky, A. (1995) Nonlinear black-box modeling in system identification: a unified overview, *Automatica*, 31(12), 1691-1724. https://doi.org/10.1016/0005-1098(95)00120-8.

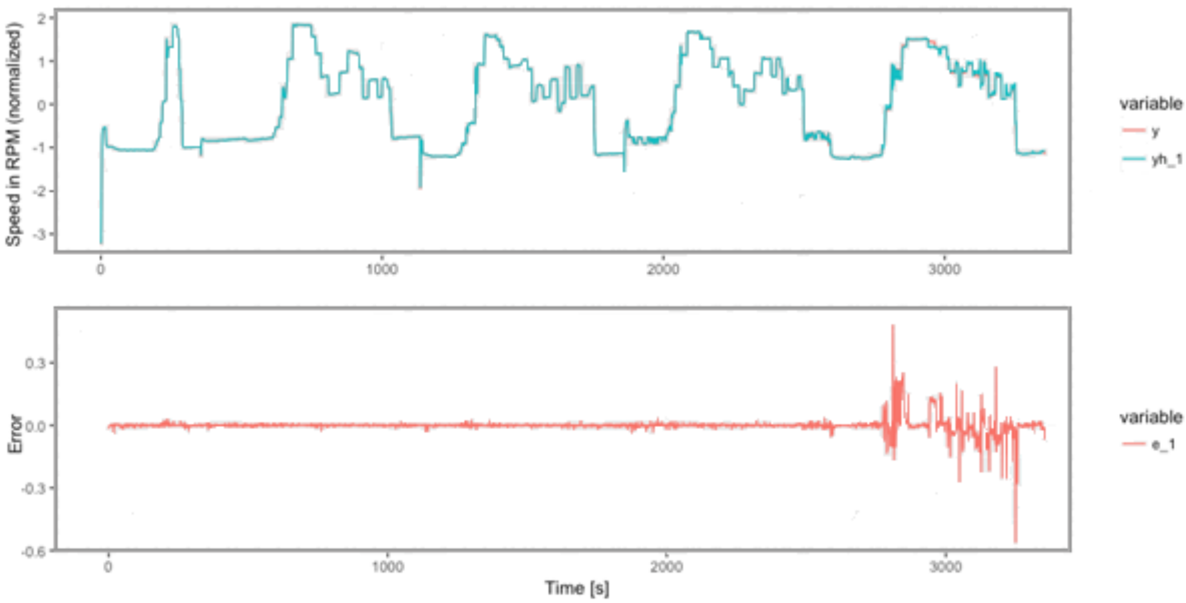


Figure 6. Output of the model and measured data (upper) and residuals (lower). Note the increase on the magnitude of the residuals in the test phase.