OPEN ACCESS
Saudi Journal of Engineering and Technology
Abbreviated Key Title: Saudi J Eng Technol
ISSN 2415-6272 (Print) |ISSN 2415-6264 (Online)
Scholars Middle East Publishers, Dubai, United Arab Emirates
Journal homepage: https://saudijournals.com

Original Research Article

Development of a Dynamic Neural Network Model for Multistep ahead Prediction of Exhaust Gas Temperature in Heavy Duty Gas Turbines

Attamah Chikaodili Stephnie[1*], Dr. Emmanuel E.C Osoka[2]

[1]MEng student researcher, Chemical engineering, Federal University of Technology Owerri
[2]Researcher, Chemical Engineering, Federal University of Technology Owerri

DOI: 10.36348/sjet.2022.v07i01.006 | **Received:** 13.12.2021 | **Accepted:** 22.01.2022 | **Published:** 30.01.2022

***Corresponding author:** Attamah Chikaodili S
MEng student researcher, Chemical engineering, Federal University of Technology Owerri

Abstract

Several studies has reported the use of neural networks in the dynamic modelling and simulation of heavy duty turbines. However; focus on exhaust gas temperature a key indicator of turbine thermal health is yet to be made. In this paper the modelling of exhaust gas temperature using the non-linear autoregressive network and subsequent multi step prediction with data collected from GT13E2 turbine was embarked upon. Features which were statistically significant for EGT prediction were selected through stepwise regression. One hidden layer was sufficient to approximate the function and The optimal architecture for training was achieved by training the network with a fixed hidden neuron and varying time delay at the inputs and output. It is observed that the optimal performance is realized when the prediction is regressed at tapped input delay of 1 in open loop. 7 hidden neuron and 1 tapped delay is selected for function approximation after series of neurons ranging from 4-15 was tested. The appropriate model was carefully selected by utilizing the method of holdout cross validation, corrected Akaike Informationon Criterion and Schwartz Bayesian information criterion. The final architecture was trained, and converted to close loop NARX network where 100 time steps ahead prediction of EGT was made. Although it was observed that accuracy diminishes as prediction horizon increases, the chosen optimised architecture successfully predicted EGT 100 steps ahead with MAE of 2.9665 and RMSE of 3.9675. Therefore; the dynamic NARX model can be utilized for multistep ahead prediction in incidence of sensor malfunction at the turbine outlet of heavy duty gas turbines.

Keywords: Multistep; Exhaust; Turbines; Prediction; Close-loop; Neural Network; Open Loop; Nonlinear Autoregressive Exogenous (NARX)

1. INTRODUCTION

The exhaust gas temperature (EGT) is an important gas path performance characteristic of a gas turbine (GT) which is an indicator of the gas turbine's thermal health (Wang *et al.*, 2015). This is because, there is a strong relationship between combustor temperature and EGT according to the Brayton thermodynamic cycle. As a result, EGT is frequently utilized as a measurable parameter for gas engine monitoring, fault diagnosis, and maintenance decisions (Wang *et al.*, 2015). It is often recommended that the maximum temperature during the transition phase should not exceed 80% of the reference temperature because overheating causes an increase in exhaust temperature, and beyond specific overheating limits, a risk of destruction exists for the hot part, particularly the gas turbine's internal combustion chamber, the nozzle, and possibly the moving blades. Furthermore, temperature monitoring in the combustion chamber is difficult to achieve, to prevent damage to the hot parts, a temperature control is applied, with the outlet temperature of the power turbine frequently employed as a control variable (Saadat *et al*., 2021).

Gas turbine supervision and control are complex roles that necessitate a high level of expertise coupled with the difficulty of developing a detailed mathematical model, as well as the randomness, fluctuating behaviour of the variable to be managed in this machine, and continuous availability of experts at all locations adds to the complexity of gas turbine supervision. Data-driven strategies has become more prevalent as artificial intelligence (AI) and big data techniques has proved to be useful in solving gas turbine engine problems. Many major turbine manufacturers, including GE, Rolls-Royce, and Pratt & Whitney, have implemented AI and machine learning technology for condition monitoring and predictive maintenance in their aero derivative engines and gas turbines (Bai *et al*., 2020).

Data-driven fault detection and diagnosis solutions use past data to extract information and do not rely on nonlinear model accuracy (Wang *et al*., 2019 & Pan 2020). Black box, grey box, and white box models are often used for modelling GT and predicting EGT dynamics behaviour during the part loading phase. Coupled and dynamic thermodynamic relations, energy balancing, and linearization approaches are frequently utilized with white box models. Moreover; when there is insufficient information on the mechanics of the system, black box models like the non-linear autoregressive network with exogenous input (NARX) are used (Bahlawan *et al*., 2017).

Several authors have adopted white box mathematical model, grey box and black box methodology in predicting EGT in GT and Internal combustion engines (ICE). Shabakhti *et al*., (2010), adopted white box mathematical model and predicted exhaust gas temperature on a homogenously charged compression ignition (HCCI) engine and spark ignition (SI) engine at both steady and transient state. Their research showed that the predicted EGT is independent on engine load for HCCI engines but; highly load dependent for SI engines. The thermal health condition of an industrial gas turbine was substituted with predicted EGT by Wang *et al*., (2014). In the proposed report a fusion technique was adopted by fusing fuzzy C means clustering and support vector machine. Open loop dynamic NARX network model was utilized by Asgari *et al*., (2016), for simulating the start-up of heavy duty gas turbines (HDGT). The developed model was able to predict turbine EGT one step ahead at a time in addition to three other features from the turbine. Hadroug *et al*., (2017), made use of grey box model to integrate Rowen mathematical model with artificial neuro fuzzy inference system (ANFIS) in predicting, controlling EGT and speed in HDGT. Olubiwe *et al*., (2018), identified EGT as a capstone on the thermal health state of turbines. In the research the authors utilized white box technique to model EGT output for a low power gas turbine (LPGT). The researchers attempted to improve the performance and control strategy that regulates EGT from ramping beyond its critical temperature. Bahlawan *et al*., (2018), developed an open loop NARX network for the modelling and simulation of transient behaviour of EGT and other selected features in HDGT with three start up scenarios which include: hot, warm, and cold start-up. Moreover; the authors were limited in the availability of independent features, although only one step ahead prediction of EGT was reported. Research work on multistep ahead prediction using black box NARX network has been investigated by a handful of researchers in different research areas. Sholahudeen *et al*., (2019), utilized an optimized close loop NARX network of seven time delays and 4 hidden neurons in making 20 steps ahead prediction of superheat temperature and cooling capacity in vapour compression air conditioning systems. Intermediate exhaust gas temperature was predicted in the modelling of Siemens aero derivative gas turbine engine (SGT-ADGTE) by Ibrahim *et al*., (2019), in addition to four other features. In the research Series parallel NARX network was compared with FFNN in the modelling of the Aero Derivative GT engine (ADGTE) and the dynamic open loop NARX network outperformed the FFNN in the representation of the dynamic response of the intermediate EGT in the Siemens ADGTE and a circumferential temperature distribution model to simulate EGT at varying defective scenarios was proposed by Liu *et al*., (2021). They tried to improve the sensitivity of combustor fault detection through the simulated EGT.

In order to monitor the turbine EGT effectively a nonlinear machine learning approach with the dynamic NARX network is therefore proposed as turbine transient performance has a direct impact on component life and performance. The use of this modelling approach which utilize information at antecedent time step allows for the setting up of a powerful simulation tool that can also be used for the real time control and diagnosis of GT thermal sensors. The main contribution of this paper is a general methodology for feature and model selection and close loop multistep prediction of EGT at part loading. Cross validation and probabilistic criterion was utilized for model selection while stepwise regression is adopted for feature selection. The trained open loop NARX network was converted to close loop network for many steps ahead prediction of the EGT. The approach utilised in this paper for feature, model selection and eventual multistep ahead prediction can be applied in similar time series processes.

2. METHODOLOGY

2.1 MODEL DESCRIPTION

The turbine chosen for this experiment is the GT13E2 single shaft power gas turbine cited in Afam power station in Rivers State, Nigeria. The model specification by design of the General Electric (formerly Alston) GT13E2 power gas turbine is as follows:

Table 1: Model specification of GT13E2 gas turbine (Afam Power Plant, 2021)

Specification	Minimum Value	Maximum Value
Efficiency	-	38%
Turbine Outlet Temperature	$0^{O}C$	$700^{O}C$
Compressor Outlet Temperature	$0^{O}C$	$600^{O}C$
Load Generator Active Power	-18MW	198MW
Flow Rate	0 Kg/s	15 Kg/s
Rotational Speed	0 RPM	3000 RPM
Compressor Inlet Temperature	$-40^{O}C$	$80^{O}C$
Compressor Inlet Pressure	60 KPA	110 KPA
Compressor Outlet Pressure	0 KPA	KPA

2.2 Data Acquisition

Power generation facilities are often designed to have a capacity sufficient for the rated load. But often times they are made to run under a partial load; a load lower than its generation capacity (IHI Corporation, 2017). GT13E2 turbines system is able to change its loading levels to either full load or part load depending on the external energy demand. The turbine is ramped up from 49.941 MW to 79.288MW during the dynamic partial load operation when the acquisition was made. During this period it was observed that the turbine has attained full speed but was transitioning from part load to its rated load. The data was retrieved at 1s interval after which 1608 data points were retrieved at a duration of 26.80 minutes.

2.3 Feature Selection

Models can have reduced performance due to the irrelevant predictors causing excessive model variation. Feature selection techniques improves models by reducing the unwanted noise of extra variables (Kuhn & Johnson, 2020). Moreover to ensure that variables which are statistically significant are introduced for neural network modelling, feature selection through stepwise regression was chosen. In this study stepwise regression was adopted as an attribute selection methodology. Stepwise regression involves adding and removing components from a multilinear model according to their statistical significance. Step wise fit estimates the model coefficients using the least squares approach in each stage of the procedure. Following the addition of a term to the model at an earlier stage, the algorithm may eventually remove the item if it is no longer useful when combined with other terms added later. When no single step improves the model, the procedure ends (Mathworks, 2019). Stepwise regression has been adopted by Guercio (2017), and Qi *et al.*, (2016) for feature selection. In this study the algorithm for stepwise regression is presented.

2.3.1 Algorithm

As reported by Qi *et al.*, 2016 the process is explained in three steps.

Step 1

Initial linear regression equation

$y = b0 + b1x_1 + b2\,x2 + ... + bn\,xn$ …………………. 1

Suppose there are *n* candidate variables for *m* samples. The raw data matrix is *x* = *(xij)n×m.*
Where *x*j is the *i*-th variable.

Step 2

Wrapper/Stepwise selection

$$Sij = Sji = \sum^{m}{}_{k=1} X_{jk}X_{ik} - [\sum_{k=1}^{m} X_{jk} \sum_{k=1}^{m} X_{ik}]/m \quad \text{…………………. 2}$$

$$Siy = \sum_{k=1}^{m} X_{jk}X_{ik} - [\sum_{k=1}^{m} X_{jk} \sum_{k=1}^{m} X_{ij}]/m \quad \text{………………. 3}$$

i, j=1,2,...,n i,j are arrays starting from 1 to *n* and *S* is the number of variables included in the equation.

Correlation coefficients between independent variables and dependent variable y are:
zij = sij / sii sqq)0.5 ………………………………. 4
ziy=siy / (lii lyy)0.5 ……………………………….. 5

The variance contribution of each variable is $v_i = z^2_{iy} / z^2_{ii}$

Set v_r=max v_i, then solve $F = (m\text{-}s\text{-}1)v_i/e$, where, *e* is the square of residuals. For first screening, *e=vr*. Where *r* is the $r_{\text{-th}}$ included or removed variable. If $F \geq 0.05$, include the variable x_r in the equation, or else remove x_r.

Calculate $v_r(s+1)$ = max $v_i(s+1)$, and $F = (m\text{-}s\text{-}2)\ v_r(s+1)/(e(s)\text{-}v_r(s+1))$. If $F(s+1) \geq 0.05$, include the variable x_r in the equation, and change the correlation matrix. Let *vr*=max *vi*, where x_r is the variable already in the equation, $Fr = (m\text{-}s\text{-}1)\ vr(s)/e(s)$. If $F_r \leq 0.05$, remove the variable x_r from the equation, otherwise include the variable.

Step 3

If any of the model's terms have p-values higher than the exit tolerance, eliminate the one with the highest p-value and proceed to step 2; otherwise, stop. Resume the process until no more variables can be added or subtracted from the equation. As a result of this, the linear regression equation is as follows:

$y = \tilde{b}0 + \tilde{b}i\ x_i + \ldots\ \tilde{b}y\ x_j + \ldots + \tilde{b}r\ x_r$ and the variables left in the equation are the qualified variables.

2.3.2 Final Feature Selection

The parameters to be measured directly from the IPGT includes the independent and dependent features after feature selection include:

Independents features

1. Network Load (MW),
2. Compressor inlet stagnation pressure (CIP KPA),
3. Compressor inlet temperature (CIT K),
4. Fuel flow rate (Kg/s

Dependent Features

1. Exhaust gas temperature (*EGT K*)

2.4 Training Methodology

The software utilized for training is MATLAB™ 2019a and the NARX model was implemented with the help of the feed forward neural network for function approximation (*f.*). However the NARX neural network model is established by introducing the time delay module and the output feedback to establish the dynamic recurrent network (Xui & Zang, 2017). The data is sectioned into three partitions where 70 percent is used for training, 15% is used to validate generalization and to ensure training stops before overfiting while the remaining 15% is used as a wholly non-independent test of network generalization. A tansigmoid transfer function is set in the hidden layer while a liner transfer function is present in the output layer to determine layers output from net input (Boussada *et al.*, 2018). Trainlm is introduced as training function because it is regarded as the fastest BPP algorithm in Statistics and Machine Learning Toolbox (2019a) and often recommended as a first choice for supervised learning. The momentum parameter ‘mu’ is set as 0.001 and at an incremental interval of 10, this is to provide a good balance so performance can be improved. Furthermore validation checks was introduced as a stopping criterion when the validation vectors fails to improve or remains the same (Khamis and Nabilah, 2014). The number of validation checks when error rate begins to increase is set to 6; this check is set to prevent overfiting of the network. Tapped delay lines at the input u(t) and output y(t) was experimented with by varying number of delays from 1 to 5 using 9 fixed neuron in the hidden layer to determine the delay that would yield the optimum performance in the time delay range. Figure 1 shows the flowchart for training the NARX model in both open and closed loop.

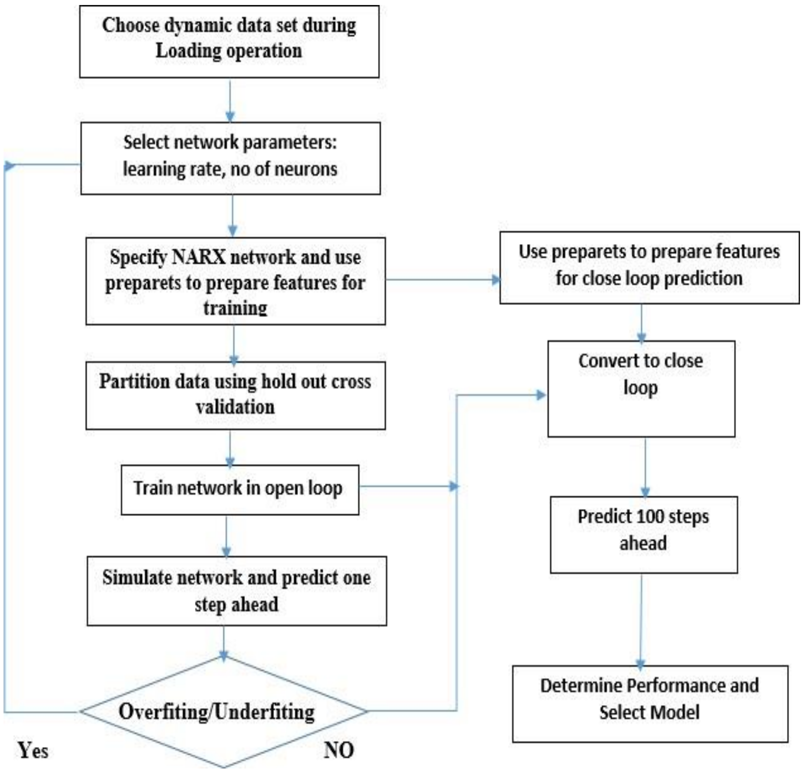

Figure 1: methodology for NARX model multistep prediction

2.5 Model Evaluation

In the regression setting, the most commonly-used measure is the mean squared error (MSE), root mean squared error (RMSE) and mean of absolute error (James *et al*., 2016), given by

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(y_i - \hat{f}(x_i))^2 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots 6$$

The RMSE is given by:

$$\text{RMSE} = \sqrt{MSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - f(x_i))^2} \quad \ldots\ldots\ldots\ldots\ldots 7$$

While the MAE calculates the absolute difference between actual and predicted values. It is given thus:

$$\text{MAE} = \sqrt{\sum_{i=1}^{n} |y - \hat{y}|} \quad \ldots\ldots\ldots\ldots\ldots\ldots 8$$

2.6 Model Selection

For model selection to be done appropriately the following selection criterion is introduced:

Hold Out Cross Validation

The holdout method splits the dataset into three parts, a training set, validation and an independent test (Raschka, 2020).

Probalistic Model Selection

Corrected Akaike's Information Criterion (AICc), and the Schwarz Bayesian Information Criterion (BIC) was also utilized for model selection. They are made up of a goodness-of-fit term and a penalty term to prevent overfitting, and they give a standardized method for balancing sensitivity and specificity (Dziak *et al*., 2017). The formula for calculating the scwartz BIC and corrected AIC can be found in equations 9 and 10.

SBC = n * log(RSS/n) + k * log(n) ………………… 9

AICc = n * log(RSS/n) + (n + p) / (1 - (p + 2) / n) … 10

Given that n is the number of data points and k is number of parameters or degrees of freedom, and RSS is the residual sum of squares.

3. RESULT AND DISCUSSION

Four independent features is selected as input variables which include: fuel flow rate in kg/s, compressor inlet temperature in kelvin, compressor inlet pressure and network load is introduced as input variables while turbine exhaust temperature was chosen as the dependent feature. Subset selection was done with the Statistical toolbox in Matlab 2019a. At the end the algorithm suggested that the four independent features were statistically significant given that their values did not exceed the maximum threshold of $p < 0.05$.

Table 2: Stepwise regression report for feature selection

Features	Coeff	Std.Error	P value	Status
CIP	-126.465	18.3415	7.72E-12	in
FR	-10.4858	1.26	1.82E-16	in
CIT	-4.2132	0.7877	1.05E-07	in
Load	2.509	0.0507	0.00E+00	in

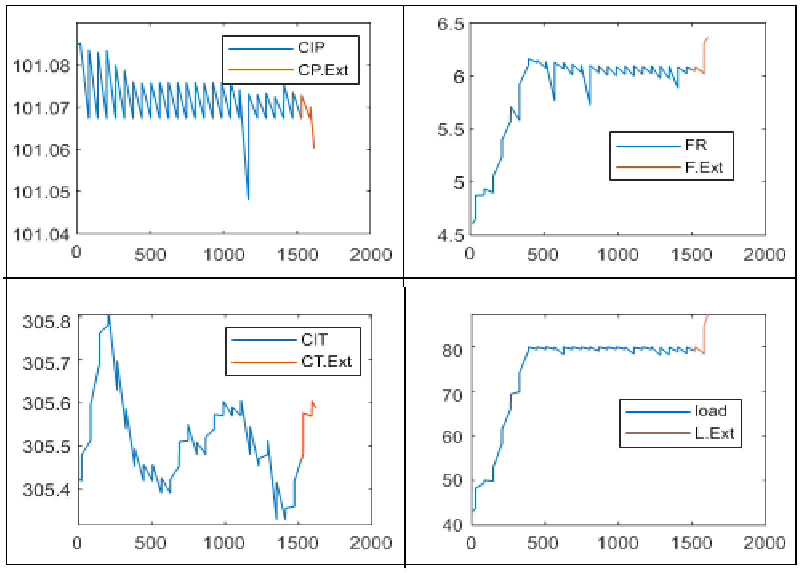

Figure 2: Plot of four independent variables with the extended data points (red)

Table 3: result for delays with 9 fixed neurons

DELAY	T MSE	Train MSE	SBC	AICc	SSE	P
1	0.1596	0.14898	-1563.15	-814.99	157.17	6.40E+01
2	0.2337	0.14249	-1295.01	-753.79	150.19	109.00
3	0.2439	0.16452	-828.61	-483.45	173.24	154.00
4	0.6195	0.18457	-392.85	-231.10	194.17	199.00
5	0.3381	0.16215	-217.53	-223.25	170.75	244.00

The performance metric used in the assessment of the models are the holdout cross validation with MSE as performance metrics, corrected Akaike information criterion (AICc) and the Schwartz Bayesian Interpolation Criterion (BIC). The training was done in open loop with nine fixed neurons and the delays were varied from 1 to 5. The training for five of the time delay architectures was carried out several (7) times in order to obtain the optimal networks that generalises unseen data well. Moreover the validation performance is not reported as the validation data set is usually used for hyper parameter tuning and weight adjustment therefore making its result biased. During the process of training they were extreme overftting cases and instances where the fit is unknown. The unknown fit was as a result of the test performance error being overwhelmingly smaller than the training error. As a result of these issues models restrictions were placed in the recording of the result. Training to test error ratio in the range higher than four and test to training performance ratio lower than 2 is rejected and were not reported.

The hold out cross validation approach suggested that although the minimum training MSE is at delay 2 it did not succeed in outperforming the model with 1 delay because; the 1 time delay model is able to generalize better than the 2 time delay model with a MSE test performance of 0.1596. Moreover; the two interpolation criterions suggested the network with 1 delay should be selected because relative to other models it has the least BIC value of -1563.15 and corrected AIC of -814.99.

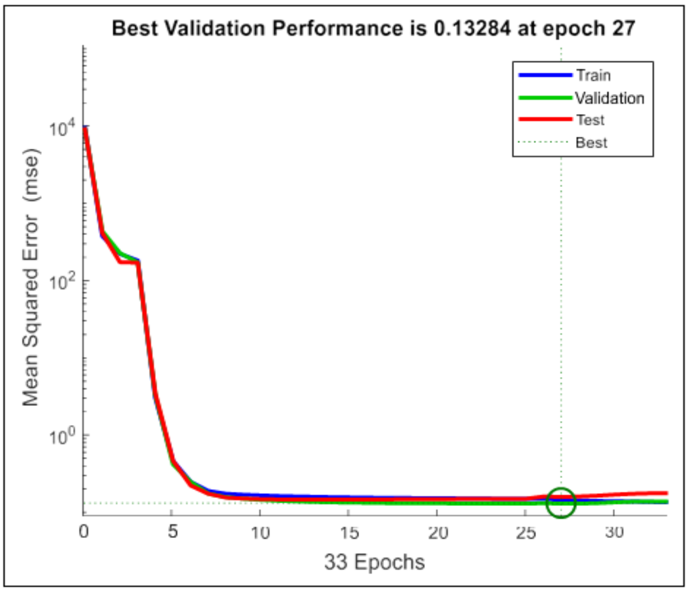

Figure 3: Plot of training epochs and MSE performance for the 1 time delay architecture

Table 4: Result for neurons with fixed delay

Neurons	Test	Train	SBC	AICc	SSE	P	
4	0.2282	0.1423	-1855.05	-940.11	150.14		29.00
5	0.1584	0.1357	-1856.45	-975.29	143.18		36.00
6	0.2806	0.1578	-1648.83	-801.24	166.45		43.00
7	0.1734	0.1203	-1885.94	-1071.72	126.94		50.00
8	0.1499	0.1741	-1447.70	-666.63	183.64		57.00
9	0.4040	0.1303	-1704.45	-972.08	137.47		64.00
10	0.1817	0.1620	-1426.05	-710.60	170.90		71.00
11	0.1929	0.1271	-1633.30	-950.32	134.08		78.00
12	0.1984	0.1623	-1326.57	-675.82	171.23		85.00

For the seconds phase of the model selection process the optimal number of hidden neuron for the model was determined by training (4-12 hidden neurons) seven times with the input and target series to the network regressed on the first(1) time delay as suggested from table and figure. The result shows that the optimal architecture for the model is the model with 7 neurons in the hidden layer and 1 time delay. Similar to table models that were significantly overfit and whose fit were unknown were not reported. The best architecture had the least training performance error and the corresponding test error is not significantly higher than the training error. Furthermore; the architecture had the least AICc and BIC values. This shows that the chosen model has a good bias variance trade off and the network generalizes well even when unseen data is being introduced.

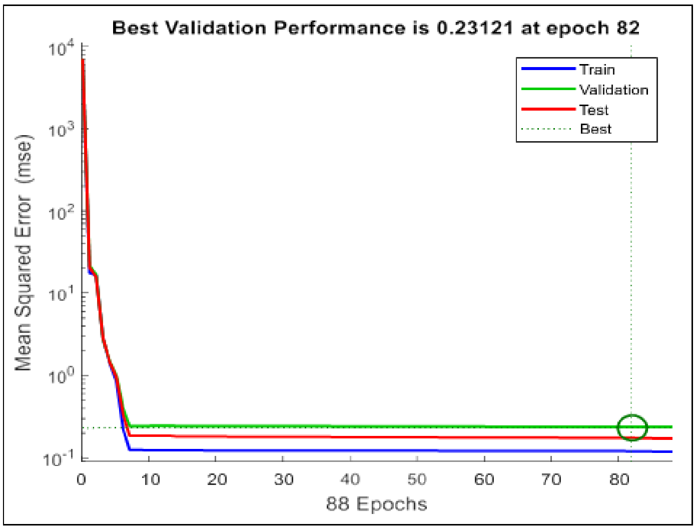

Figure 4: Best performance with 7 hidden neurons & 1 time delay

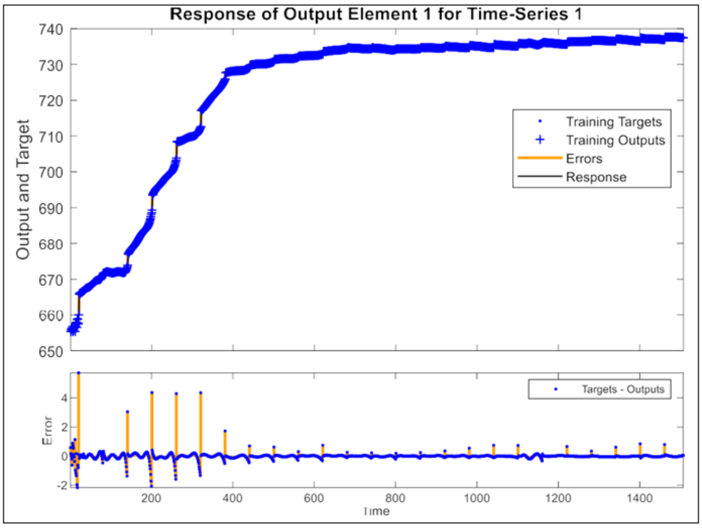

Figure 5: Time - series response of predicted EGT

3.1 Multistep Ahead Prediction

Table 5: Architecture and performance of multistep ahead prediction

Description	Value
Neuron	7
Time Del	1
MSE	15.7414
MAE	2.9665
RMSE	3.9675
Training	0.1066

The network is retrained using the best model architecture but the entire data set is used for training (1:1508). To make 100 steps ahead prediction the network was converted from open loop to closed loop where 1509:1608 of EGT data points was predicted from past predicted values of output ($\hat{y}(t)$) time series fed back as input to the network, in addition to the use of present and past values of the input data. It is observed as the prediction interval increases or as time progresses into the future, model accuracy or ability to predict into the future decreases. The close loop model is able to predict 100 steps ahead with RMSE of 3.9675 of exhaust temperature in kelvin and MAE of 2.9665 kelvin in exhaust gas temperature deviation.

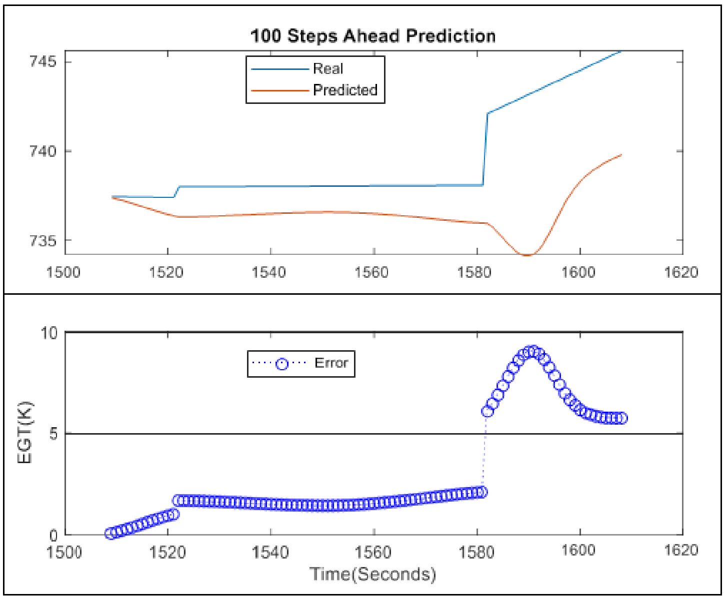

Figure 6: 100 steps ahead closed loop prediction in addition with prediction error

4. CONCLUSION

The Modelling and multistep ahead prediction of EGT from GT13E2 turbine was successfully achieved by the use of NARX network. Stepwise regression was adopted as feature selection strategy to check for irrelevant predictors and unwanted noise from irrelevant variables that could lead to excessive model variation. Training was carried out in three stages. In the first phase, the best time delay response architecture was selected on a range of 1-5 and it was observed that the network performed best with 1 time delay response. Furthermore; since the neuron at the output and hidden layer remains fixed the number of hidden neuron was selected at the second stage by exhaustively training 4-15 neurons at the hidden layer seven times to obtain the best result. 7 hidden neuron and 1 time delay response was chosen as the best NARX architecture from 4-15 hidden neuron, because it had a good trade off of bias and variance and moreover they had the least corrected AIC and BIC values. The chosen architecture was retrained with the whole data set starting from 1 to 1508 time series data points in seconds. However; the multistep ahead prediction was achieved by converting the final trained model to closed loop and predicting 100 steps ahead from 1509 to 1608 data points, where the performance is 2.9665 in MAE and 3.9675 in RMSE. It was observed that as the prediction interval increases the error rate tends to increase. Therefore; it is recommended that further research should carried out to combat the challenge of extrapolation, and approaches to uncertainties of neural networks to improve the forecasting horizon of time series networks.

Declaration of Competing Interest

The authors assert that they have no known competing financial interests or personal ties that could have influenced the research presented in this research.

Acknowledgment

The authors express immense appreciation to First Independent Power Limited (FIPL), Afam Power Station, Rivers State for their support in the provision of historical data needed for the research.

Declaration of Competing Interest

The authors assert that they have no known competing financial interests or personal ties that could have influenced the research presented in this research.

REFERNCES

- Asgari, H., Chen, X., Morini, M., Pinelli, M., Sainudiin, R., Spina, P. R., & Venturini, M. (2016). NARX models for simulation of the start-up operation of a single-shaft gas turbine. *Applied Thermal Engineering*, 93, pp. 368-376
- Bahlawan, H., Morini, M., Pinelli, M., Ruggero Spina, P., & Venturini, M. (2018). Development of reliable narx models of gas turbine cold, warm, and hot start-up. *Journal of Engineering for Gas Turbines and Power*, *140*(7). https://doi.org/10.1115/1.4038838
- Bahlawan, H., Morini, M., Pinelli, M., Spina, P. R., & Venturini, M. (2017). Development of reliable narx models of gas turbine cold, warm and hot start-up. *Proceedings of the ASME Turbo Expo 2017: Turbomachinery Technical Conference and Exposition*. *Volume 9: Oil and Gas Applications; Supercritical CO2 Power Cycles; Wind Energy*. https://doi.org/10.1115/GT2017-63332
- Bai, M., Liu, J., Ma, Y., Zhao, X., Long, Z., & Yu, D.

(2020). Long Short-Term Memory Network- Based Normal Pattern Group for Fault Detection of Three-Shaft Marine Gas Turbine. *Energies*, *14*(1), 13. MDPI AG. Retrieved from http://dx.doi.org/10.3390/en14010013.

- Boussaada, Z., Curea, O., Remaci, A., Camblong, H., & Mrabet Bellaaj, N. (2018). A nonlinear autoregressive exogenous neural network model for the prediction of the daily direct solar radiation. *Energies*, *11*(3), 620. MDPI AG. Retrieved from http://dx.doi.org/10.3390/en11030620
- Buitrago, J., & Asfour, S. (2017). Short-term forecasting of electric loads using nonlinear autoregressive artificial neural networks with exogenous vector inputs. *Energies*, *10*(1), 40. https://doi.org/10.3390/en10010040
- Dziak, J. J., Coffman, D. L., Lanza, S. T., Li, R., & Jermiin, L. S. (2020). Sensitivity and specificity of information criteria. *Briefings in Bioinformatics*, *21*(2), 553–565. https://doi.org/10.1093/bib/bbz016
- Guercio, G. P. V. (2017). Stepwise regression: an application in earthquake localization. *International Journal of Contemporary Research and Review*, *8*(09), Article 09. https://doi.org/10.15520/ijcrr/2017/8/09/319
- Hadroug, N., Hafaifa, A., Guemana, M., Kouzou, A., Salam, A., & Chaibet, A. (2017). Heavy duty gas turbine monitoring based on adaptive neuro-fuzzy inference system: Speed and exhaust temperature control. *Mathematics-in-Industry Case Studies*, *8*(1), 8. https://doi.org/10.1186/s40929-017-0017-8
- Hastie, T., Tibshirani, R., & Friedman, J. (2017). Elements of statistical learning: data mining, inference, and prediction. 2nd Edition. Retrieved August 2, 2021, from https://web.stanford.edu/~hastie/ElemStatLearn/
- IHI Corporation. (2017). *Operating Gas Turbines at Maximum Efficiency*. Retrieved September 18, 2021 from https://www.ihi.co.jp/var/ezwebin_site/storage/original/application/01274f53c7132171a31cb 267bdcd75eb.pdf
- Ibrahem, I. M. A., Akhrif, O., Moustapha, H., & Staniszewski, M. (2019). Neural networks modelling of aero-derivative gas turbine engine: A comparison study. *1*, 738–745. https://doi.org/10.5220/0007928907380745
- James, G, Witten, D., & Hastie, T. (2016). An Introduction to Statistical Learning—With Applications in R. Retrieved August 2, 2021, from https://www.springer.com/gp/book/9781071614174
- Kuhn, M., & Johnson, K. (2020). Feature Engineering and Selection: A Practical Approach for Predictive Models. *Routledge & CRC Press*. Pg 12Retrieved August 2, 2021, from https://www.routledge.com/Feature-Engineering-and-Selection-A-Practical-Approach-for- Predictive-Models/Kuhn-Johnson/p/book/9781032090856
- Liu, J., Long, Z., Bai, M., Zhu, L., & Yu, D. (2021). A Comparative Study on Fault Detection Methods for Gas Turbine Combustion Systems. *Energies*, *14*(2), 389. MDPI AG. Retrieved from http://dx.doi.org/10.3390/en14020389
- Mathworks, (2019). Statistics toolbox user's guide. Retrieved from https://www.mathworks.com/help/stats/stepwisefit.html
- Pan, Z., Meng, Z., Chen, Z., Gao, W., & Shi, Y. (2020). A two-stage method based on extreme learning machine for predicting the remaining useful life of rolling-element bearings. *Mechanical Systems and Signal Processing*, *144*, 106899. https://doi.org/10.1016/j.ymssp.2020.106899
- Panchal, G., Ganatra, A., Kosta, Y. P., & Panchal, D. (2010). Searching most efficient neural network architecture using akaike's information criterion (AIC). *International Journal of Computer Applications*, *1*(5), 54–57. https://doi.org/10.5120/126-242
- Qi, Y., Liu, G., & Zhang, W. (2016). A Matlab program for stepwise regression. Retrieved from https://www.semanticscholar.org/paper/A-Matlab-program-for-stepwise-regression-Qi-Liu/e251487b62adaa26aad9f6e01ab18318845976c1
- Raschka, S. (2020). Model evaluation, model selection, and algorithm selection in machine learning. Retrieved from http://arxiv.org/abs/1811.12808
- Saadat, B., Hafaifa, A., & Kouzou, A. (2021), Supervised gas turbine speed and axial load control exhaust temperature. *Gulf publishing Holdings*.
- Shahbakhti, M., Ghazimirsaied, A., & Koch, C. R. (2010). Experimental study of exhaust temperature variation in a homogeneous charge compression ignition engine. *Proceedings of the Institution of Mechanical Engineers*, Part D: Journal of Automobile Engineering, 224(9), 1177–1197. https://doi.org/10.1243/09544070JAUTO1473
- Sholahudin, S., Ohno, K., Yamaguchi, S., & Saito, K. (2019). Multi-step ahead prediction of vapor compression air conditioning system behaviour using neural networks. *IOP Conference Series: Materials Science and Engineering*, *539*(1), 012003. https://doi.org/10.1088/1757- 899X/539/1/012003
- Ugoh C. A., Olubiwe M., D. S., & Inaibo D. S. (2018). Improving the response time of load – frequency control of gas turbine in a nigerian refinery. *International Journal of Engineering Research & Technology (ijert)* Volume 07, Issue 01.
- Wang, B., Lei, Y., Li, N., & Yan, T. (2019). Deep separable convolutional network for remaining useful life prediction of machinery. *Mechanical Systems and Signal Processing*, *134*, 106330.
- Xiu, Y., & Zhang, W. (2017). Multivariate chaotic time series prediction based on NARX neural networks. 164–167. https://doi.org/10.2991/eame-17.2017.40
- Wang, Z. T., Zhao, N. B., Wang, W. Y., Tang, R., & Li, S. Y. (2015). A fault diagnosis approach for gas turbine exhaust gas temperature based on fuzzy c-means clustering and support vector machine. *Mathematical Problems in Engineering*, *2015*.