RESEARCH ARTICLE

An Investigation of Exhaust Gas Temperature of Aircraft Engine Using LSTM

SHAFI ULLAH[1], SHUGUANG LI[1], KHALID KHAN[2], SHAHBAZ KHAN[3,4], ILYAS KHAN[5], AND SAYED M. ELDIN[6]

[1]Department of Traffic Information Engineering and Control, Chang'an University, Xi'an, Shaanxi 710064, China
[2]Department of Materials for Energy Storage and Conversion, Pakistan Center for Advanced Studies in Energy, University of Engineering and Technology Peshawar, Peshawar 25000, Pakistan
[3]Department of Engineering Management, University of Engineering and Applied Sciences, Swat, Swat, Khyber Pakhtunkhwa 01923, Pakistan
[4]Mechatronics Department, University of Engineering and Technology Peshawar, Peshawar, Khyber Pakhtunkhwa 25000, Pakistan
[5]Department of Mathematics, College of Science Al-Zulfi, Majmaah University, Al Majma'ah 11952, Saudi Arabia
[6]Center of Research, Faculty of Engineering, Future University in Egypt, New Cairo 11835, Egypt

Corresponding authors: Khalid Khan (engrkhaliduet@gmail.com) and Ilyas Khan (i.said@mu.edu.sa)

ABSTRACT A significant obstacle to creating efficient machine health monitoring systems is estimating performance degradation in dynamic systems, like aero plane engines. In exceedingly complex systems with many components, states, and parameters, conventional model-based and data-driven methods fall short of producing satisfactory results. While traditional methods had several drawbacks, deep learning has emerged as a viable computational tool for dynamic system prediction. In order to track system deterioration and estimate the EGT, a novel technique based on the Long Short-Term Memory (LSTM) network, (an architecture created to find the hidden patterns hidden in time series data) is provided in this research. The health monitoring information of aircraft turbofan engines is used to assess the effectiveness of the proposed strategy. As a result of this network's ability to recognize the input data as a real-time series, the output in the following step can be predicted. Results of the suggested study show a significant ability to anticipate the output in the following time step. Additionally, the proposed model has a shorter learning curve and is more accurate.

INDEX TERMS Predictive maintenance, artificial intelligence, exhaust gas temperature, condition-monitoring, machine learning.

I. INTRODUCTION

Modern-day manufacturing technology has taken a quantum leap into a new paradigm with the advent of technology, ensuing complexity in modern machinery. To necessitate the compelling demand for safety, reliability in operation, and productivity, it is indispensable having an intelligent system that would pro-actively anticipate any unforeseen situation and coordinate it with the operator/ supervisor, thus ensuring the top echelon of safety, yet mitigating the cost pertaining to maintenance [1], [2]. According to the International Organization for Standardization, Prognostics is ''a prediction of the time to failure and the risk of one or more potential failure modes, present or future'' [3]. These prognostic techniques

The associate editor coordinating the review of this manuscript and approving it for publication was Shadi Alawneh.

provide Real-time analysis of the system's health in a way to assess the performance degradation and future behavior of the system [4]. In the recent past decades, the field of prognostics has been developed as an active field of research. Recent advancements in the field of prognostics are based on two main approaches; the Machine Learning approach (ML) and the Bayesian approach [1]. The first approach (ML) enacts the basis for the predictive models using analysis and sensing of data, subsequently, numerically relating the learned/ discovered pattern from the data to a specific learning task. The second approach (Bayesian approach) establishes a relationship between the degradation of performance of a system (manufacturing) as a probability distribution, predicting its later behavior using some recursive steps of update and state prediction, using sensor measurements [5], [6]. Particle Filter (PF) and Kalman Filter are two main approaches in this field.

Engines have been studied over the years, using steady-state conditions in a very controlled environment [7]. Most of the engine data has been obtained using dynamic simulation tools previously. The lack of real-time data has been a major issue since most of the aircraft engine data is highly confidential, thus data acquisition is usually done using Dynamic simulation tools. The real-time data is crucial for training the model and subsequent prediction of the fault ahead of time, thus avoiding fiasco. Nevertheless, the downside to this is the complex nature of real-time data, wherein multiple variables interrelate in a highly convoluted manner, thus requiring deeper analysis and understanding of the real-time data. In the presented research work, real-time, in-flight data has been obtained from the Digital Electro- Electronic Control (DEEC) system of the aircraft engine, consisting of two sets of data i.e.

1. Healthy Data for training the Neural Network model the healthy data is obtained for the initial flight hours of the engine, wherein no fault is observed using routine inspections.
2. Faulty data for Testing and Validation of the Neural Network model whenever a fault is noticed in the engine e.g., cracks, burns, nick marks, etc. using borescope inspection.

Based on the transient behavior of data, Recurrent-Neural-Network (RNN) (a sub-type of Deep Learning) is considered is an effective technique for addressing time series, which is a frequently adopted technique [8]. However, the issue with vanishing and exploding gradients limits its potential usability [7]. Long Short-Term Memory (LSTM) is a variation of RNN that performs better than traditional RNNs, since they were designed to address the problem of vanishing or exploding gradients in traditional RNNs [7]. They are one of the most popular architectures used for modeling multi-variate time series sequential data in various fields, such as image processing, Natural Language Processing (NLP), Speech Recognition [7], [9]. A Long-Short-Term-Memory (LSTM), Neural Network is designed to recognize patterns in sequences of data and capture long term dependencies. They also allow previous outputs to be used as inputs while having hidden states. They have feedback memory with the ability to take decisions and output results based on sequences of data.

The aforestated literature reveals the implementation of various ANNs architectures pertaining to the prediction and condition monitoring of dynamic aero-engine behavior; nevertheless, based on author's knowledge in this field, LSTM has never been investigated previously. The proposed research work is aimed at training the neural network on the transient data taken from a low by-pass turbofan engine, with the Exhaust Gas Temperature (EGT) as the output parameter. After the successful training of the Neural Network, the estimated/predicted EGT of the engine is validated against the actual EGT obtained from the engine. The predicted values of the EGT provides a deeper insight of the incoming faults ahead of time, and enable the flight crew to take corrective measures, thus avoiding catastrophe.

II. LITERATURE REVIEW

For prediction, optimization, and condition monitoring, Exhaust Gas Temperature (EGT) is considered as one of the key parameters for predicting the health of an aero-engine [10]. EGT is the measure of the mean temperature at the exhaust section of engine. One significant sign of deterioration of an engine is the increase in the EGT, resulting in greater turbine wear and decline in remaining RUL [3], [11]. It is necessary to take note of the EGT during the take-off and keep it as low as possible because this is the point where EGT is higher and the exceeding of EGT from its normal limits are responsible for the failure of engine's components [3]. To determine the correlation between EGT and other factors, Multiple Linear Regression (MLR) was carried out for various gas path performance monitoring data. Strong linear correlations between the essential metrics EGT, a measure of gas performance, and other important factors are shown by Wang et al. research [12]. These characteristics produce similar outcomes at two dissimilar power settings, such as maximum continuous and takeoff [13].

Current prognostic methods are widely classified into two classes, namely data driven-based and model-based approaches. Model based methods are based on the system's underlying physics and mathematical model. This model incorporates physical models of the system and physical understanding of the system to predict the future health and condition of the engine components. A data driven model is based solely on current and historical data collected from the sensor's measurements. A model for fault propagation rate is developed using data from trials conducted under specified conditions and the component damage level [14]. The model learns from the historical data (that contains faults) and validates its results with the current data (predicts faults in this data). Detailed and accurate mathematical models of aeroengines are difficult to develop. Due to this reason data driven methods are of great interest in aero-engine modeling.

The Piece-Wise Linear Modeling for on-board dynamic model is the most popular for steady-state performance [15]. Khan et al. [16] investigated that data driven approaches are cost effective as compared to model based. Mathematical models are difficult to develop when system complexity increases (in EGT). This is where data driven models perform better. RNNs are perfect for health management systems because they consider previous states [2], [17]. In order to predict EGT parameters, a prognostic model based on the Gated Recurrent Unit (GRU) network was presented. EGT parameters are considered to be time series with non-linear features. GRU network is employed because it considers the data's non-linear and time-series features. The issues of vanishing gradients and expanding gradients that are present in RNN are addressed by GRU.

The uncertainty induced by the employment of steady state models in take-off conditions was measured by Zhong and

Verbist [18], [19]. By employing a medium-size, two-shaft turbofan engine output model with a high bypass ratio, they were able to show that the ambient temperature, the duration of the taxi-out, and the duration of the prior aircraft turn around all significantly affect the transient effects during takeoff. The authors suggested a strategy for correcting take-off snapshots to steady-state operational conditions based on the simulation results. A dynamic model has been created for the J-85 turbojet engine as the test case in the work of Andrei [20]. The design and off-design performances have been estimated for a single spool turbojet. A physics based dynamic model in [11] has been developed for a twin spool turbo-shaft engine with the aim of developing a tool to be used for designing novel control algorithms for controlling the engine, and for predicting off-design transient performance. The mathematical model is validated through integrating experimental data of engine components. Transient simulations are helpful in the design process of engine because they provide a mechanism of establishing new engine safety measures, and testbed for development of control systems. The engine also experiences overloads and over temperature during transient operations, so these simulations are also critical for diagnostic and prognostic health management systems.

The piece-wise linear model is not accurate enough for several on-line maneuvers, such as deceleration, acceleration, and after-burning on/off. In the work of Lu et al. [21] the issues of compensating for engine/model mismatch and adapting to performance deterioration are investigated. A novel Real-time life-cycle model titled ALPVM (Adaptive Linear Parameter Varying Model) has been developed to explain the dynamic behavior of a turbofan engine. Artificial neural networks are now frequently employed as data-driven models for simulating and modelling the performance of aviation engines. Some of the numerous techniques used in ANNs include backpropagation neural networks (BPNN), nonlinear input-output (NIO), adaptive network-based fuzzy inference system (ANFIS), nonlinear autoregressive with exogenous inputs (NARX), radial basis function (RBF), and feedforward multi-layer perceptron (MLP).

An algorithm for building self-organizing Radial Basis Function neural networks for estimating aeroengine thrust is presented by Li et al. [22]. The method has the capability to calculate the link weights and optimize the size of the neural network. The constructed networks are highly accurate, proving the usability and efficacy of the suggested strategy. A new Online Sequential Extreme Learning Machine (OS-ELM) algorithm and Logistic Regression were used to develop the aero-engine degradation prognostics-based approach created by Lu et al. [21]. KFOS-ELM, a novel OS-ELM training method based on the Kalman filter, outperforms OS-ELM in terms of stability and regression accuracy without requiring more computational work. The NARX model, a Recurrent Neural Network variation, is capable of capturing the dynamics of complex systems like gas turbines. For a heavy-duty single-shaft gas turbine, NARX models have been developed by Asgari et al. [23]. The results showed that NARX models are useful for forecasting gas turbine transient behavior. Agrawal and Yunis [24] have provided a generalized mathematical model of an aviation gas turbine to estimate engine performance. At whatever ambient temperature or height, the model does an excellent job of describing system dynamics. Bretschneider [25] describes the development of a turbofan engine model that can simulate the start-up, shutdown, or wind milling processes. The engine's physical impacts during startup were investigated, and modelling solutions were demonstrated. The model's ability to predict the proper trends has been demonstrated by results. By considering flight dynamic characteristics, Bagherzadeh [26] improved the ANN model for identifying aircraft systems. A feed-forward neural network is used with a variety of known flight dynamic modes as input. The simulated and actual flight data of the High Alpha Research Vehicle (HARV) aircraft performing at high angle of attack maneuvers served as the basis for training and assessing the neural network. Results show that the proposed method has higher accuracy when compared to traditional methods. A more practical solution for time series prognostic modelling is provided by the Deep Learning techniques described in Hochreiter [27] and Längkvist [28]. In order to extract hierarchical representations from the raw data, the deep learning model is composed of a variety of non-linear Recurrent Neural Networks [29]. Recurrent neural networks (RNNs) are a subset of deep learning techniques that are thought to be particularly useful for addressing the time series [30]. In an RNN, the recurrent connections communicate the time series features. The advantage of the conventional RNN is lessened in practice, though, by the issues of "vanishing gradients" and "exploding gradients." Long Short-Term Memory, one of the most effective RNN architectures for sequence learning, was suggested by Hochreiter [27] in order to address the above limitation. Hochreiter [27] continues to address the issue of Long-Term Memory capture using the Long Short-Term Memory (LSTM) network. Speech recognition, image processing, genomic analysis, and natural language processing are only a few of the applications that use RNNs, particularly LSTM, which were developed to address the issue of gradient exploding or vanishing in RNNs [31], [32], [33]. In order to determine a machine's Remaining Useful Life (RUL), LSTMs are used. Wu et al. [13] research demonstrates that the proposed approach is capable of achieving significant performance in one-step prediction tasks, long-term prediction tasks, or remaining useful life prediction tasks [13]. Additionally, according to Che et al. [34] research, the suggested LSTM model has good long and short-term prediction outcomes.

For the purposes of tracking system performance degradation and RUL prediction, a bi-directional LSTM is suggested. The theoretical foundation of the RNN and LSTM is presented first, followed by the mathematical formulation

of the system deterioration tracking and RUL prediction. The bi-directional LSTM network's structure and training procedure are then reported in Zhang, 2018 #157. In EGT modelling, LSTM performs significantly better than RNN (Chen, 2019 #194). The raw data was initially normalized during data pre-processing in RUL in prognostics utilizing deep convolutional neural networks, and MSE was employed as a loss function (Lipton, 2015 #196; Yuan, 2016 #195). System health management related deep RNN publications have been widely published. For instance, LSTMs for RUL estimation are studied by Zheng et al. in Zhang, 2018 #157. The simple RNN, LSTM, and GRU have all been effectively compared by Yuan et al. in Yuan, 2016 #195 for the diagnosis and prognostics of aeroengine faults. The authors of Zhao et al. empirical evaluation of a health monitoring system utilizing LSTMs (Zhao, 2017 #197) predicted the tool wearing (Zhao, 2016 #198). Inspired by prior research, this paper presents an LSTM-based approach for estimating EGT.

The analysis executed here differs from the past studies in a few ways. Maximum of the previous studies conducted in the past is based on simulator and nearly steady state or test bench data. However, we have transient data to train the model on a data driven basis using a neural network AI approach such as RNN and LSTM. As from literature study and conceptual point of view the key performance parameter for monitoring the health and RUL of an engine are exhaust gas temperature (EGT) also referred as Turbine Outlet Temperature (TOT) [10], [11]. But, as it is clear that from the analysis made that engine is non-linear complex system so selection of optimized input parameter such inlet pressure P1, inlet temperature T1, fan speed N1, RPM, compressor pressure ratio, turbine pressure ratio, pressure difference of fuel flow, Mach no, bleed ratio, altitude etc. from the Digital electronic engine control (DEEC) parameter is a central task. Initially, the engine was analyzed by dividing the engine into four modules that are inlet section, the compressor section, the combustor section, and the turbine section for better understanding of DEEC parameters. And then with help of past studies and literature review and regression analysis, thirteen 13 features were selected overall including the output parameter as EGT.

Various RUL prediction models for aviation engines are reviewed, evaluated, and their performance is contrasted with a suggested Long-Short Term Memory (LSTM) method based on a data-driven machine learning approach. The results collected demonstrate that the modified LSTM approach with Attention mechanism enhances and gives higher performance for RUL prediction for aircraft engines [35]. Long Short-Term Memory (LSTM) neural networks are being used in another study to achieve good diagnosis and prediction performance in the presence of complex procedures, hybrid errors, and significant noise. Testing is done on a NASA-provided dataset for aircraft turbofan engine health monitoring to illustrate and debate the entire notion. Tests and comparisons were done on the effectiveness

of the LSTM and a few of its modifications. The conventional LSTM performed better than others, according to experiment data [36]. Improved multi-stage Long Short-Term Memory (LSTM) network for RUL prediction model, is suggested in order to increase the RUL of the aero-prediction engine's accuracy. Using this model as a foundation, they investigated a comparable multi-stage RUL prediction algorithm that combines the benefits of clustering analysis and LSTM modelling. The dataset from the National Aeronautics and Space Administration (NASA) is used for validation. The experimental findings demonstrate the effectiveness of the strategy suggested in this research in reducing the prediction error of the aero-engine RUL. [37]

The underlying literature reveals the implementation of various ANNs architectures pertaining to the prediction and condition monitoring of dynamic aero-engine behavior; nevertheless, based on author's knowledge in this field, LSTM has never been investigated previously. The proposed research work is aimed at training the neural network on the transient data taken from a low by-pass turbofan engine, with the Exhaust Gas Temperature (EGT) as the output parameter. After the successful training of the Neural Network, the estimated/predicted EGT of the engine is validated against the actual EGT obtained from the engine. The predicted values of the EGT provides a deeper insight of the incoming faults ahead of time, and enable the flight crew to take corrective measures, thus avoiding catastrophe.

The remainder of this paper is divided into 4 sections as follows. Section II includes the literature review. In section III the Data set description is provided and discussed. While in section IV Data pre-processing has been considered: including 1) separate air ground data, 2) Feature selection, and 3) normalization of data. In section V the model has been trained using neural networking RNN approach showing the results of the training model and finally interpretation on the results has been made.

III. METHODOLOGY
The methodology for modeling the above-mentioned parameter (EGT) includes the following steps

    A. Data Preparation
    B. Modeling Neural Networks

        i. Validating the Neural Networks
       ii. Testing the Model against other Engines

A. DATA PREPARATION
To be able to predict labels with higher accuracy, the input data to our neural networks need to be refined and free of any inconsistencies. The following steps were taken to achieve the preparation of data for the neural networks:

1) Separating air and ground data: The data is separated in air and ground data based on the ''Signal Afterburner Event'' (when signal turns 1 for the first time, it shows the start of data when the aircraft is airborne). The data starting from the first afterburner event till the

aircraft touching the ground was considered as air data and the rest as ground data and was dropped from the analysis since we are only interested in the air data.

2) Feature Selection: In this step, we identify parameters in the dataset that potentially affect the performance of the model. After being left with a substantial chunk of data, the next step we performed was to remove the redundant and unnecessary feature columns. These included time features such as seconds, minutes, and hours etc. as well as other redundant features (features that were recorded by two sensors simultaneously for accuracy). We also removed some features that help start the engine of the aircraft such as "turbine starter" etc.

We were left with the following 13 features:

1) Inlet Air Temperature
2) Inlet Air Pressure
3) Turbo Compressor rotor rotation frequency
4) Fan Rotation Frequency
5) Gas Temperature after the fan turbine
6) Air Pressure after the Compressor
7) Fuel Feeding
8) Nozzle Throat
9) Gas Pressure after the Fan Turbine
10) Throttle position
11) Fuel Temperature after the fuel filter
12) Oil Temperature going out the engine
13) Pressure Difference on the Fuel Filter

B. NORMALIZATION

Different parameters in our data have different ranges; temperatures are in the range of hundreds, whereas pressures are both in fractional and negative values. Therefore, it is necessary to convert all parameters on a similar scale. Without normalization, the neural network has a challenging time to move the weight vectors towards a satisfactory solution. Nevertheless, with normalization, the data will be more concentrated, and the neural network will be able to find a helpful solution for the problem. This method of normalization, however, could cause relative errors in the system when new data is presented that is yet to be normalized.

For example, if the EGT of an aircraft engine ranges from 0 to 900 °C, however, the data provided could contain values only from 0 to 840 °C. Addition of a new data that contains EGT values of greater than 840 °C could cause a disturbance in the data. The thresholds of these input parameters are taken from the manual, which came with the engine. The thresholds of these input parameters are taken from the manual, which came with the engine. A better solution, therefore, is to obtain the highest and lowest values of each parameter, which acts like some sort of benchmark limits and then normalize the data according to these limits rather than through their relative values. The following table provides an overview of all the parameters and their minimum and maximum values.

TABLE 1. EGT parameters and their thresholds.

S.no.	Parameter	Minimum value	Maximum value
1	Inlet Air Temperature	-10	60
2	Inlet Pressure	-0.5	3.0
3	Turbo Compressor Rotor Rotation Frequency	0	105
4	Fan Rotation Frequency	0	1000
6	Fuel Feeding	0	100
7	Nozzle Throat	0	100
8	Air Pressure after the Compressor	-5	50
9	Gas Pressure after the Fan Turbine	-40	110
12	Oil Temperature going out the Engine	-60	195
13	Pressure Difference on the Fuel Filter	-1	1

C. PROGNOSTIC MODEL BASED ON LSTM NETWORK

LSTMs, short for Long Short-Term Memory networks, are one of the special kinds of recurrent neural network (RNN) that are used for modeling sequential data which gives them an advantage over simple neural networks. The data in the proposed work was evaluated with many models, such as Recurrent neural network, which also worked. But in actual practiced, two major problems were anticipated, i.e., exploding gradients and vanishing gradients, thus making it unusable. Later on, LSTM was opted by introducing a memory unit, also called as cell in the network. The block diagram of a LSTM model is shown in Figure 1.

They are designed with a goal of addressing the vanishing gradients and exploding gradient problems of traditional RNNs. They have a wide range of applications in various fields such as speech recognition, time-series prediction, and human action recognition. The architecture of LSTMs was first proposed in 1997 by Sepp Hochreiter and Jürgen Schmidhuber [27]. LSTMs are explicitly designed to learn long-range dependencies in temporal sequence prediction problems. It remembers information for extended periods of time and performs better than RCNN in several time-series tasks. The network structure of LSTM consists of a memory block in a recurrent hidden layer which is responsible for storing the temporal state of the network and the flow of information is controlled through three major mechanisms,
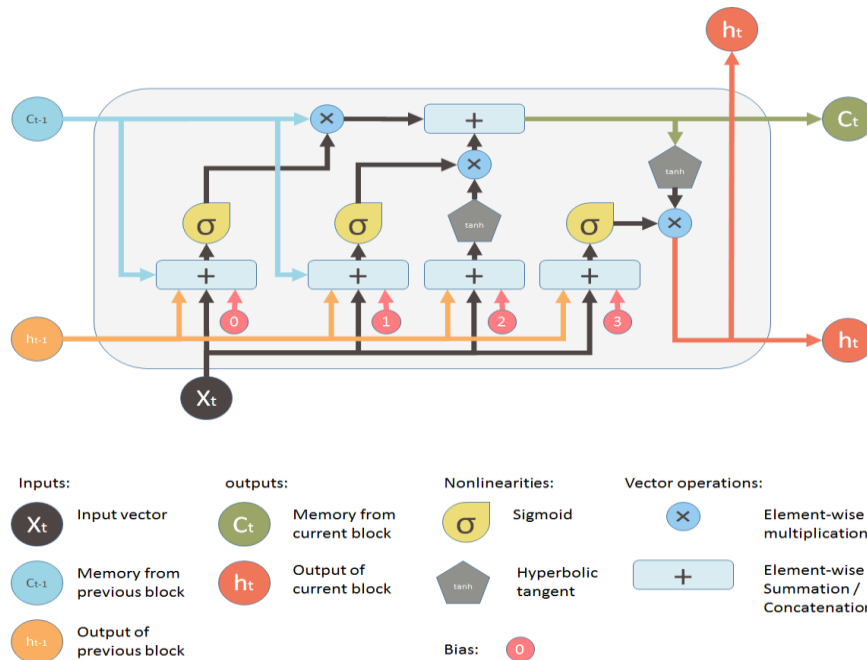
FIGURE 1. Block diagram of an LSTM.

TABLE 2. Flights of different engines for measuring performance.

Training MSE	Testing MSE Engine 2	Testing MSE Engine 3	Testing MSE Engine 8	Testing MSE Engine 9
$2.8*10^{-4}$	$2.2*10^{-3}$	$2.0*10^{-3}$	$3.4*10^{-3}$	$1.3*10^{-3}$

called gates. It consists of three gates mainly input gate, forget gate, and output gate. The sigmoid activation function is used in these gates which outputs a value in an interval [0, 1]. A value of 0 indicates that the gates are discarding all the information, while a value of 1 indicates the allowing of all the information to pass. Similarly, the status information is also recorded after each moment. While the forget gate instructs the information to be removed from the cell state, the input gate controls the added information that is stored there. The output gate's responsibility is to activate the final output at the time of activation.

The equations for the gates in LSTM are:

$$i = bi + \sigma Wi(ht-1, xt) \tag{1}$$
$$f = \sigma bf + Wf(ht-1, xt)\} \tag{2}$$
$$o = \sigma \{b_o + W_o(h_{t-1}, x_t)\} \tag{3}$$

Input gate, forget gate, and output gate are represented, respectively, by equations 1, 2, and 3. The sigmoid function is denoted by the symbol σ, where W stands for the weight for the corresponding gate x neurons, ht-1 for the output of the preceding block at timestamp (ht-1), xt for the input at the current timestamp, and bx for the biases for the corresponding gates x. Currently, the LSTM equations for cell state, candidate state, and final output are as follows;

$$z = tanh\{W_z(h_{t-1}, x_t)\} + b_z \tag{4}$$
$$ct = f \times ct - 1 + i \times z \tag{5}$$
$$ht = o \times tanh(ct) \tag{6}$$

where z represents candidate for cell state, ht is the final output at time-stamp t, ct represents the cell state at current moment and ct−1 at the last moment.

The EGT has been considered as the primary parameter in this study to be estimated using LSTM; hence, the EGT value is the output y. (t). It is necessitated to provide as input a few factors that significantly affect the EGT. Twelve variables that are considered for predicting the EGT have been chosen.

D. MODEL CONFIGURATION

Three layers make up the proposed LSTM network: an input layer, a hidden layer, and an output layer. We have twelve input parameters; hence the input layer has twelve nodes, and the hidden layer has twenty-four nodes. A dense layer with just one node serves as the output layer for IEGT. Each neuron in the dense layer receives information from every neuron in
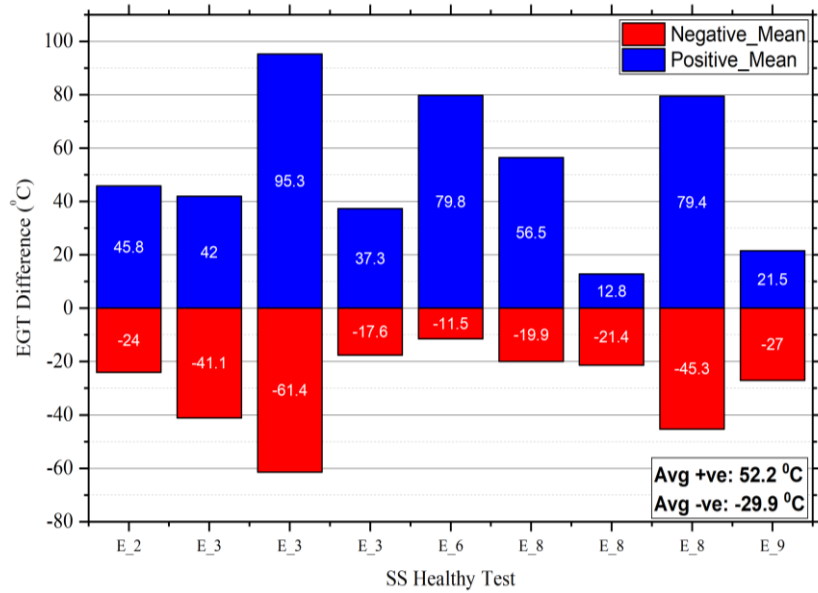
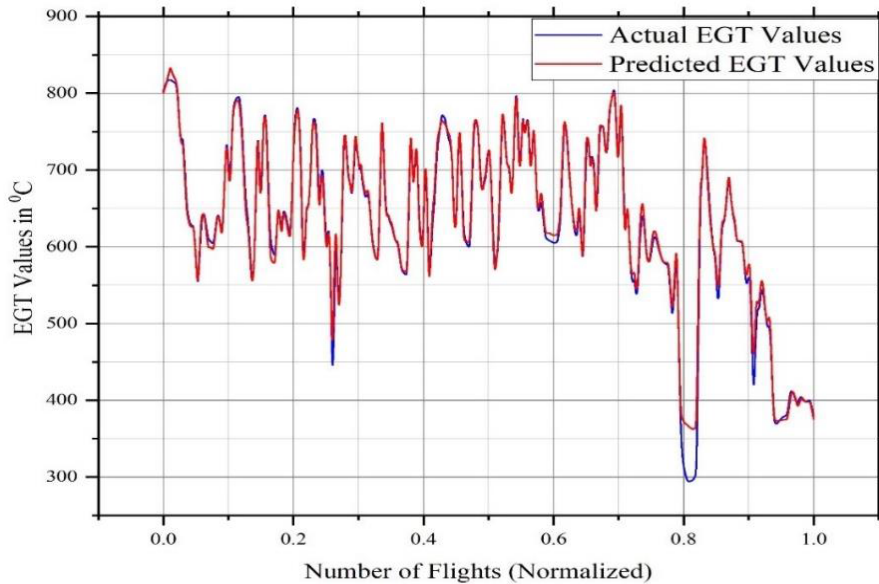**FIGURE 2.** EGT results for different engines.

FIGURE 3. Real Vs Estimated EGT.

the preceding layer, making it densely connected. The layer has the activation of the preceding layer a, a bias vector b, and a weight matrix W. We used batch sizes of 10, 50 and 200 in order to find the optimum batch size and an Adam optimizer [41] with a learning rate of 0.01. Number of epochs were set at 10, and look-back data points were 50 (10 seconds because our data is recorded at 5 Hz).

IV. RESULTS AND DISCUSSIONS

The output of the proposed EGT prognostic model is shown in this section. As previously noted, we used 12 parameters as input that are thought to be correlated with EGT and used

EGT as our main parameter to be estimated. We have nine engines with a set of flight data, as was covered in the section on the data-set description. Each engine differs from the others due to the setup and operating circumstances that are unique to it. A prognostic model is therefore created for each engine. Furthermore, for comparison two models were developed and trained for every engine; one for single flight data, and one where data from different flights was concatenated and a model was trained on it. Examining the residuals allows one to gauge the model's final performance. The discrepancy between the quantity's observed and estimated values, in this case, EGT, is known as the residual value.

TABLE 3. NN training results, batch size 200.

MSE	Variance	Std Dev	Positive Mean	Negative Mean
0.00012344	0.00012343	0.01110972	0.00722085	-0.0084785

TABLE 4. NN training results, batch size-50.

MSE	Variance	Std Dev	Positive Mean	Negative Mean
$1.1*10{-4}$	$9.8*10{-5}$	$9.9*10{-3}$	$5.6*10{-3}$	$-8.4*10{-3}$
$0.1\ ^{0}C$	–	$9.87\ ^{0}C$	$5.55\ ^{0}C$	$-8.40\ ^{0}C$

A. TRAINING MODEL ON SINGLE FLIGHT DATA

This section describes how the training and testing outcomes of a single fight from the healthy data. An individual flight from this dataset is chosen, a model is trained on it, and its performance is evaluated using data from several flights powered by various engines. The residual errors for each flight in our data set are shown in Figure 2 below. Nine flights were chosen at random from the data, yielding a total of nine unique findings. The bars in this graph display the discrepancy between the measured values and the actual values derived from our model.

Positive and negative differences are further separated from one other. The graph makes it very evident that there are large disparities between the actual and anticipated numbers (except E8). This is a bad strategy since the model cannot accurately predict from the training data.

B. TRAINING MODEL ON MULTIPLE FLIGHTS DATA

Before training the model, the dataset is divided into training and testing sets. The first 20 flights are concatenated and used as training data. The remaining flights data is used to evaluate the model. The same model configuration is used but trained with three different batch sizes. Results of these batch sizes are compared to determine the best option. The performance of the model is measured by the residual errors. Less values of residuals indicate better performance.

1) TRAINING RESULTS (ENGINE 3 B200)

The following Table 3 shows the training results on a batch size of 200 of engine 1.

The data consists of 200 flights of this engine.

Only the first 20 flights are selected as training data so that the model does not over-fit. These 20 flights were concatenated and then fed to the model as training data. For testing purposes, single flights from the remaining flights are used. Table 3 shows the results. The following table gives the MSE, variance, the standard deviation of data, and mean of points miss-classified as positive, and mean of points miss classified as negative are given.

2) NN TRAINING RESULTS (ENGINE F 3 – B200)

The following graph 4 plots the result of all the single flight testing of the first engine. The model is evaluated on 23 different flights spread across the dataset to ensure the model is tested on different fights operating in different conditions. From the graph it is evident that the errors are low in the first 5 flights and the model estimates the EGT very well here. In the next flights, there are small cracks appearing in the data, which causes the error rate to go up slightly as the cracks are increasing. Towards the end the error is going up as there are more cracks appearing. The errors increase with the increase in crack values on the turbine blades. An explanation for this is that the function of turbine blades is to extract energy from the gas. When cracks appear in the blades it affects their ability to extract energy. This extra energy manifests in the high exhaust gas temperature which gives rise to higher values of EGT. Overall, the performance of the model is good as the predicted values and actual values are only 9 °C apart on average.

3) NN TRAINING RESULTS (ENGINE F 3–B50)

The results in Table 4 show the results of training the model with a batch size of 50. The EGT values follow the same trend as was observed in the preceding sections. The error differences remain small when there are no cracks, and they start to increase when cracks are introduced in the data. The average positive and negative temperatures here are slightly smaller than the previous model where we evaluated on a batch of size 200. Reducing batch size from 200 to 50 shows slightly better results but the difference from the previous model to this is not significant, just 2 °C.

4) NN TRAINING RESULTS (ENGINE F 3 - B10)

From the results of evaluating the following flights shown in Figure 3 and Table 5, it is clear that reducing the batch size from 50 to 10 has produced better results. The average positive and negative temperatures are similar to when the batch size was 50 and the model also performs slightly better. This is evident from the value of negative mean. For flight

TABLE 5. NN training results, batch size 10.

MSE	Variance	Std Dev	Positive Mean	Negative Mean
$7.9*10{-}5$	$7.6*10{-}5$	$8.7*10{-}3$	$5.7*10{-}3$	$-7.2*10{-}3$
$0.07\ {}^{0}$C	-	$8.76\ {}^{0}$C	$5.7\ {}^{0}$C	$-7.16\ {}^{0}$C

numbers F-31 to F-230 the residual error values are smaller. From flights F-240 to F-247, the residual error values have increased. The value for the positive mean remains almost similar but there is a decrease in the negative mean value from 10.76 °C to 8.43 °C. The reason for this increase is the introduction of cracks in the dataset. As the cracks have intensified (4.5 mm in F-240 to 45 mm in F-247) the model is unable to accurately estimate the EGT.

V. CONCLUSION

In this paper, an approach for the estimation of EGT is proposed, based on LSTM neural networks. LSTM neural networks are amazingly effective for the estimation of EGT because of the sequential nature of the data. An LSTM network with 12 nodes in the input layer, 24 nodes in the hidden layer, and 1 node in the output layer is constructed. The network was trained for three different batch sizes. The model was then validated by a real engine dataset that contains both steady and transient conditions and the MSE is analyzed for all configurations. The results show that the change in batch size from 200 to 50 and from 50 to 10 produces better EGT estimated values. The MSE for batch size of 10 are $7:9 \times 10({-}5)$. The average difference in the estimated EGT and the real EGT was $5.7\ {}^{0}$C which is the acceptable range. The errors in the prediction increase when large cracks on the turbine blade appear in the dataset, which gives rise to abrupt changes in EGT and subsequently the model under-performs. In future work, we will study how to extract features more effectively to better estimate the EGT when cracks appear.

REFERENCES

[1] J. Zhang, P. Wang, R. Yan, and R. X. Gao, "Long short-term memory for machine remaining life prediction," *J. Manuf. Syst.*, vol. 48, pp. 78–86, Jul. 2018.

[2] K. Khan, M. Sohaib, A. Rashid, S. Ali, and H. Akbar, "Recent trends and challenges in predictive maintenance of aircraft's engine and hydraulic system," *J. Brazilian Soc. Mech. Sci. Eng.*, vol. 43, no. 8, p. 403, 2021.

[3] M. T. Yildirim and B. Kurt, "Aircraft gas turbine engine health monitoring system by real flight data," *Int. J. Aerosp. Eng.*, vol. 2018, Mar. 2018, Art. no. 9570873.

[4] D. Wu, S. Liu, L. Zhang, J. Terpenny, R. X. Gao, T. Kurfess, and J. A. Guzzo, "A fog computing-based framework for process monitoring and prognosis in cyber-manufacturing," *J. Manuf. Syst.*, vol. 43, pp. 25–34, Apr. 2017.

[5] A. Doucet and A. M. Johansen, "A tutorial on particle filtering and smoothing: Fifteen years later," in *Handbook of Nonlinear Filtering*, vol. 12, nos. 656–704. New York, NY, USA: Oxford Univ. Press, 2009, p. 3.

[6] C. Nemeth, P. Fearnhead, and L. Mihaylova, "Sequential Monte Carlo methods for state and parameter estimation in abruptly changing environments," *IEEE Trans. Signal Process.*, vol. 62, no. 5, pp. 1245–1255, Mar. 2014.

[7] M. Grazia De Giorgi, S. Campilongo, and A. Ficarella, "Predictions of operational degradation of the fan stage of an aircraft engine due to particulate ingestion," *J. Eng. Gas Turbines Power*, vol. 137, no. 5, May 2015, Art. no. 052603.

[8] Q. Zheng, H. B. Zhang, Y. Li, and Z. Hu, "Aero-engine on-board dynamic adaptive MGD neural network model within a large flight envelope," *IEEE Access*, vol. 6, pp. 45755–45761, 2018.

[9] S. Kiakojoori and K. Khorasani, "Dynamic neural networks for gas turbine engine degradation prediction, health monitoring and prognosis," *Neural Comput. Appl.*, vol. 27, no. 8, pp. 2157–2192, Nov. 2016.

[10] A. Von Moll, A. R. Behbahani, G. C. Fralick, J. D. Wrbanek, and G. W. Hunter, "A review of exhaust gas temperature sensing techniques for modern turbine engine controls," in *Proc. 50th AIAA/ASME/SAE/ASEE Joint Propuls. Conf.*, Jul. 2014, p. 3977.

[11] L. Fedele, L. Di Vito, and F. E. Ramundo, "Increasing efficiency in an aeronautical engine through maintenance evaluation and upgrades: Analysis of the reliability and performance improvements under financial issues," *Energies*, vol. 13, no. 12, p. 3059, Jun. 2020.

[12] Z.-T. Wang, N.-B. Zhao, W.-Y. Wang, R. Tang, and S.-Y. Li, "A fault diagnosis approach for gas turbine exhaust gas temperature based on fuzzy C-Means clustering and support vector machine," *Math. Problems Eng.*, vol. 2015, pp. 1–11, 2015.

[13] Y. T. Wu, M. Yuan, S. Dong, L. Li, and Y. Liu, "Remaining useful life estimation of engineered systems using vanilla LSTM neural networks," *Neurocomputing*, vol. 275, pp. 167–179, Jan. 2018.

[14] K. Goebel and N. Eklund, "Prognostic fusion for uncertainty reduction," in *Proc. AIAA Infotech@Aerospace Conf. Exhibit*, May 2007, p. 2843.

[15] T. Brotherton, A. Volponi, R. Luppold, and D. L. Simon, "eSTORM: Enhanced self tuning on-board real-time engine model," Intell. Automat. Corp, Poway, CA, USA, Tech. Rep. NASA/CR-2003-215272, 2003.

[16] S. Khan and T. Yairi, "A review on the application of deep learning in system health management," *Mech. Syst. Signal Process.*, vol. 107, pp. 241–265, Jul. 2018.

[17] S. Khan, P. Phillips, C. Hockley, and I. Jennions, "No fault found events in maintenance engineering Part 2: Root causes, technical developments and future research," *Rel. Eng. Syst. Saf.*, vol. 123, pp. 196–208, Mar. 2014.

[18] S. Zhong, Z. Li, L. Lin, and Y. Zhang, "Aero-engine exhaust gas temperature prognostic model based on gated recurrent unit network," in *Proc. IEEE Int. Conf. Prognostics Health Manage. (ICPHM)*, Jun. 2018, pp. 1–5.

[19] M. L. Verbist, W. P. Visser, and J. P. van Buijtenen, "Experience with gas path analysis for on-wing turbofan condition monitoring," in *Turbo Expo: Power for Land, Sea, Air*, vol. 55188. New York, NY, USA: American Society of Mechanical Engineers, 2013, Art. no. V004T06A023.

[20] I.-C. Andrei, A. Toader, G. Stroe, and F. Frunzulica, "Performance analysis and dynamic modeling of a single-spool turbojet engine," in *Proc. AIP Conf.*, 2017, vol. 1798, no. 1, Art. no. 020005.

[21] F. Lu, J. Wu, J. Huang, and X. Qiu, "Aircraft engine degradation prognostics based on logistic regression and novel OS-ELM algorithm," *Aerosp. Sci. Technol.*, vol. 84, pp. 661–671, Jan. 2019.

[22] Z.-Q. Li, Y.-P. Zhao, Z.-Y. Cai, P.-P. Xi, Y.-T. Pan, G. Huang, and T.-H. Zhang, "A proposed self-organizing radial basis function network for aero-engine thrust estimation," *Aerosp. Sci. Technol.*, vol. 87, pp. 167–177, Apr. 2019.

[23] H. Asgari, X. Chen, M. Morini, M. Pinelli, R. Sainudiin, P. R. Spina, and M. Venturini, "NARX models for simulation of the start-up operation of a single-shaft gas turbine," *Appl. Thermal Eng.*, vol. 93, pp. 368–376, Jan. 2016.

[24] R. K. Agrawal and M. Yunis, "Generalized mathematical model to estimate gas turbine starting characteristics," *J. Eng. Power*, vol. 104, no. 1, pp. 244–267, 1982.

[25] S. Bretschneider and J. Reed, "Modeling of start-up from engine-off conditions using high fidelity turbofan engine simulations," in *Turbo Expo: Power for Land, Sea, Air*, vol. 56628. New York, NY, USA: American Society of Mechanical Engineers, 2015, Art. no. V001T01A026.

[26] S. A. Bagherzadeh, "Nonlinear aircraft system identification using artificial neural networks enhanced by empirical mode decomposition," *Aerosp. Sci. Technol.*, vol. 75, pp. 155–171, Apr. 2018.

[27] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997.

[28] M. Längkvist, L. Karlsson, and A. Loutfi, "A review of unsupervised feature learning and deep learning for time-series modeling," *Pattern Recognit. Lett.*, vol. 42, pp. 11–24, Jun. 2014.

[29] G. Hinton, L. Deng, D. Yu, and G. E. Dahl, "Deep neural networks for acoustic modeling in speech recognition: The shared views of four research groups," *IEEE Signal Process. Mag.*, vol. 29, no. 6, pp. 82–97, Nov. 2012.

[30] J. Z. Sikorska, M. Hodkiewicz, and L. Ma, "Prognostic modelling options for remaining useful life estimation by industry," *Mech. Syst. Signal Process.*, vol. 25, no. 5, pp. 1803–1836, Jul. 2011.

[31] F. Lu, J. Qian, J. Huang, and X. Qiu, "In-flight adaptive modeling using polynomial LPV approach for turbofan engine dynamic behavior," *Aerosp. Sci. Technol.*, vol. 64, pp. 223–236, May 2017.

[32] M. Auli, M. Galley, C. Quirk, and G. Zweig, "Joint language and translation modeling with recurrent neural networks," in *Proc. EMNLP*, 2013, pp. 1–11.

[33] A. Karpathy and L. Fei-Fei, "Deep visual-semantic alignments for generating image descriptions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 3128–3137.

[34] C. Che, H. Wang, Q. Fu, and X. Ni, "Combining multiple deep learning algorithms for prognostic and health management of aircraft," *Aerosp. Sci. Technol.*, vol. 94, Nov. 2019, Art. no. 105423.

[35] A. Boujamza and S. Lissane Elhaq, "Attention-based LSTM for remaining useful life estimation of aircraft engines," *IFAC-PapersOnLine*, vol. 55, no. 12, pp. 450–455, 2022.

[36] M. Yuan, Y. Wu, and L. Lin, "Fault diagnosis and remaining useful life estimation of aero engine using LSTM neural network," in *Proc. IEEE Int. Conf. Aircr. Utility Syst. (AUS)*, Oct. 2016, pp. 135–140.

[37] J. Liu, F. Lei, C. Pan, D. Hu, and H. Zuo, "Prediction of remaining useful life of multi-stage aero-engine based on clustering and LSTM fusion," *Rel. Eng. Syst. Saf.*, vol. 214, Oct. 2021, Art. no. 107807.

SHAFI ULLAH was born in Khyber Pakhtunkhwa, Pakistan, in 1994. He received the B.Sc. degree in science from the University of Swat, and the M.Sc. degree from the Xi'an University of Science and Technology. He is currently pursuing the Ph.D. degree with the Department of Traffic Information and Control Engineering, School of Electronic and Control Engineering, Chang'an University, China.

SHUGUANG LI received the Ph.D. degree in electrical engineering from Xi'an Jiaotong University. He has been employed with the Department of Traffic Information and Control Engineering, School of Electronic and Control Engineering, Chang'an University, China, since 2005. His research interests include intelligent transportation systems, traffic management and control, expressway electromechanics, dynamic network traffic flow theory, and machine learning, deep learning, and AI applications in transportation.

KHALID KHAN received the bachelor's degree in mechanical engineering from the University of Engineering and TechnologyPeshawar, Pakistan, in March 2015, and the master's degree from Arizona State University, Tempe, AZ, USA, in 2018. He joined the U.S. Pakistan Center for Advances Studies in Energy, in 2016. He did his research work in fuel cell with Arizona State University, Tempe, in 2017. He is currently working as a Mechanical Team Lead in a joint research project of the University of Engineering and Technology in collaboration with TDF and Pakistan Airforce (Technology Development Fund), funded by the Higher Education Commission of Pakistan. His research interests include renewable energy, fuel cell, batteries, solar cell, industrial simulation, ergonomics, robotics, artificial intelligence, and manufacturing.

SHAHBAZ KHAN received the B.Sc. degree in mechatronics engineering from the University of Engineering and Technology (UET) Peshawar, Pakistan, in 2012, the M.Sc. degree in mechanical engineering from Deakin University, Australia, and the Ph.D. degree from the Department of Mechatronics Engineering, UET Peshawar, in 2022. From 2017 to 2019, he was a Lecturer at the Department of Mechanical Engineering, HITEC University, Taxila. From 2019 to 2022, he worked as a Laboratory Engineer at the Department of Mechatronics Engineering, UET Peshawar. He is currently a Lecturer with the Institute of Manufacturing, University of Engineering and Applied Sciences (UEAS). He is also a Research Team Member with the Advanced Robotics and Automation Laboratory, Department of Mechatronics Engineering, UET Peshawar. His research interests include agricultural robotics, control, and machine vision.

ILYAS KHAN received the Ph.D. degree in applied mathematics from the Universiti Teknologi Malaysia (UTM), in 2012. He held a postdoctoral position at UTM, in 2013. He received several distinctions, such as a Senior Visiting Research Fellow and a Visiting Researcher. He has over 15 years of academic experience in different reputed institutions of the world. He is currently working as an Associate Professor with the Department of Mathematics, College of Science, Majmaah University, Saudi Arabia. He has published more than 900 research articles in different well-reputed international journals. His research interests include mathematical modeling, analytical and computational fluid dynamics, biomathematics, and numerical computing. He has received twice Distinguish Researcher Award from Majmaah University. According to AD-Scientific Index 2022, he received the First Place as a Best Scientist at Majmaah University and the Third Place as a Best Scientist in Saudi Arabia. He is an Editorial Board Member of different journals, like *Applied Nanoscience*, *Scientific Reports*, and *Advances in Mechanical Engineering*. He is a reviewer of several journals and remained guest editor of some special issues in reputed journals.

SAYED M. ELDIN was the Dean of the Faculty of Engineering, Cairo University, where he achieved many unique signs of progress in both academia and research on the impact of emerging technologies in electrical engineering. He is currently with the Faculty of Engineering and Technology, Future University in Egypt, on leave from Cairo University after 30 years of service to the Faculty of Engineering, Cairo University. He was a PI of several nationally and internationally funded projects. He has many publications in highly refereed international journals and specialized conferences in the applications of artificial intelligence on protection of electrical power networks. In addition, he is in the editorial boards of several international journals.