Health Monitoring Framework for Aircraft Engine System using Deep Neural Network

Hyunseong Lee1, Guoyi Li1, Ashwin Rai2, and Aditi Chattopadhyay3

1Graduate Research Associate, School for Engineering of Matter, Transport, and Energy, Arizona State University, Tempe, AZ, 85287, USA
hlee269@asu.edu
guoyili@asu.edu

2Post-Doctoral Research Associate, School for Engineering of Matter, Transport, and Energy, Arizona State University, Tempe, AZ, 85287, USA
ashwin.rai@asu.edu

3Regents’ Professor and Ira A. Fulton Chair Professor of Mechanical and Aerospace Engineering, School for Engineering of Matter, Transport, and Energy, Arizona State University, Tempe, AZ, 85287, USA
aditi@asu.edu

ABSTRACT

A real-time monitoring framework is developed to detect operational anomalies in aircraft engine performance. A historical flight dataset recorded from commercial aircraft is utilized to perform the proposed method. Sampling frequency synchronization and denoise are performed on the flight dataset using signal processing techniques. A robust detection algorithm using the deep neural network is developed to capture flight performance anomalies that show significant off-nominal behavior in engine related and flight dynamic features. The accuracy and efficiency of the proposed monitoring method are validated through a demonstration of anomaly detection in the aircraft engine system associated with dynamic flight behavior.

1. INTRODUCTION

Flight passenger demand is expected by International Air Transport Association (IATA) to double by 2035, which may increase aviation safety issue due to the increase in air traffic density. Current air infrastructure heavily depends on human-dominated systems such as manual inspection, manual controller, and communications between airlines, which motivates the significant need for automated health management of the National Airspace System (NAS) to ensure aviation safety (IATA, 2018). However, the development of the monitoring framework is challenging since operational anomalies may occur in various degree of severity at various location (Kobayashi & Simon, 2003). Furthermore, the complexity in aircraft structure and various operating condition make the analysis of current flight condition difficult. To mitigate these issues, many methodologies have been developed for monitoring aircraft status and identifying anomalies. A model-based method has been widely utilized for anomaly detection, which requires an explicit mathematical model to characterize aircraft system response using system identification (de Loza, Cieslak, Henry, Dávila, & Zolghadri, 2015; Hajiyev & Soken, 2013; Kobayashi & Simon, 2003; Wang & Lum, 2007). An observer-based fault detection and isolation (FDI) technique is developed to capture actuator faults in the F-16 aircraft using the adaptive unknown input observers (Wang & Lum, 2007). Kobayashi and Simon (2003) propose a bank of Kalman filters for aircraft engine FDI, which detects faults in sensors and actuators based on health parameter estimation and fault factor. However, modeling of the whole aircraft system is excessively limited without exhaustive a-priori knowledge of each included subsystem in aircraft (Yoon & MacGregor, 2000). Furthermore, these models are strictly system dependent and cannot be applicable out of the specific application that they are designed for (Hou & Wang, 2013). Instead, the data-driven method utilizes historical information to model system behavior without exhaustive knowledge of physics-based modeling and provide detection capabilities (Qin, 2012). However, a vital issue in the data-driven method is the associated computational expense in training process, mainly due to the high dimensionality of training data. To mitigate this issue, the information fusion and feature extraction techniques have been developed such

---

Hyunseong Lee et al. This is an open-access article distributed under the terms of the Creative Commons Attribution 3.0 United States License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.
as principal component analysis (PCA), support vector machine (SVM), and neural networks (NN) (Fujimaki, Yairi, & Machida, 2005; Jung, Berges, Garrett Jr, & Poczos, 2015; Kromanis & Kripakaran, 2013; Lee, Li, Rai, & Chattopadhyay, 2019; Mounce, Mounce, & Boxall, 2011; Tayarani-Bathaie, Vanini, & Khorasani, 2014; Vanini, Khorasani, & Meskin, 2014). Fujimaki et al. (2005) proposed the monitoring framework using kernel principal component analysis (KPCA) to monitor engine performance of spacecraft by analyzing deviations of eigenvectors in feature space. Kromanis and Kripakaran (2013) utilized the support vector regression (SVR) model with moving fast Fourier transform (MFFT) to investigate the behavior of large structures under abnormal scenarios. Tayarani-Bathaie et al. (2014) proposed multiple dynamic neural networks (DNN) for aircraft turbine FDI by embedding memory term to enhance estimation accuracy in a nonlinear aircraft system, where classical NN structures cannot represent system behavior due to complexity in the aircraft structure. The developed framework was performed on engine performance data from the numerical model and showed detection capability with a coupled response under a fault situation. Furthermore, outlier detection algorithms are also utilized for anomaly detection, which identifies anomalies in flight performance as outliers (Das, Matthews, Srivastava, & Oza, 2010; Li, Rai, Lee, & Chattopadhyay, 2018). Das et al. (2010) developed multiple kernel anomaly detection (MKAD) algorithm using one class SVM model to capture anomalies in discrete and continuous flight features induced by off-nominal flight operations such as go-around operation.

Even though many approaches have been proposed, the development of a robust monitoring framework is required to mitigate drawbacks in current methods: (i) lack of the integrated aircraft health monitoring which can monitor the flight performance associated with aircraft subsystems; (ii) demonstration of the developed method to commercial flight data such as onboard sensor data in aircraft to validate detection ability in realistic flight situation.

Motivated by these challenges, a real-time monitoring framework for aircraft engine systems using the deep neural network (DNN) is presented. A historical flight dataset recorded from onboard sensors in commercial aircraft is utilized to conduct the proposed method and detect performance anomalies in the aircraft engine system. Signal processing using decimation and Savitzky-Golay filter method is conducted to preprocess the flight dataset. Subsequently, the DNN model is trained by the processed flight features to estimate normal engine system behavior of commercial aircraft associated with flight dynamics. The potential anomalies are detected when the estimation error exceeds pre-defined rational bound due to significant deviation in monitoring features. Demonstration of anomaly detection is performed with simulated historical flight dataset in real-time to validate the real-time detection capability of the proposed monitoring method.

2. HISTORICAL FLIGHT INFORMATION

Sanitized commercial flight data recorder (FDR) datasets, obtained from the National Aeronautics and Space Administration (NASA) Discovery in Aeronautics Systems Health (DASH) link network (Monroe, Freeman, & Jones, 2012), is utilized to perform the proposed monitoring method. The FDR dataset is comprised of flight features recorded from the onboard sensors in commercial aircraft during in-flight and on-ground operation and is generally used for flight analytics after flight accidents. The FDR dataset includes 186 flight features that can be categorized into continuous and discrete features. The continuous features generally indicate aircraft states and performance such as fan speed, engine gas temperature, and velocity and angle components of aircraft coordinate system while the discrete features denote binary information such as system alarms and landing gear status. Besides, the sampling frequency of flight features varies between 0.25 Hz to 16 Hz, which restricts the direct application of the dataset to the proposed framework. In this study, a total of 2044 flights are investigated, where the flight times in the dataset vary from 56 minutes to 3 hours and 22 minutes and flight phases are labeled in to 7 digits, parking (1), taxi (2), takeoff (3), ascent (4), cruise (5), descent (6), and landing (7) as shown in Figure 1. Among the flight phases, the proposed monitoring work focuses on detecting engine performance anomalies in cruise phase. The magnitude of engine-related features such as fan speed and exhaust gas temperature are ideally maintained during cruise phase, while fluctuations are observed during multiple ascent and descent phases. Therefore, off-nominal behavior of flight performance is relatively distinguishable compared to other flight phases. However, health metric and standard operation ranges are not presented in the FDR dataset, so it prevents cross-validation of captured operational anomalies from flight accident reports and literature because of unknown information about aircraft type, flight plan, and weather condition. With this limitation, a decision boundary was formed under the assumption that the anomalous flight behavior is uncommon patterns in dataset and presents significant deviation from common patterns due to non-standard flight operations or mechanical issues in the aircraft systems.

![Figure 1. Flight phases of scheduled flight in FDR dataset](image-url)
3. Monitoring Framework

3.1. Preprocessing

In order to estimate engine performance with dynamic flight behavior, the number of 47 features, including engine related features and flight dynamic features, are selected to be monitored as presented in Table 1. Subsequently, a downsampling technique called decimation is implemented to synchronize the various sampling frequencies of the selected features, which leads to consistency of features in the time domain to train the DNN model. In general, a process of decimation is divided by two steps: (i) eliminate high-frequency elements that provoke aliasing using a digital low pass filter; (ii) implement down-sampling by an integer factor (Milic, 2009). In this work, finite impulse response (FIR) filter is utilized for anti-aliasing method to take advantage of beneficial properties of the FIR filter; no iterations are required due to inherent stability (Milic, 2009). In each flight p, where p = 1, 2, ..., Pmax (Pmax = 2044 in this study), the flight features are denoted as $x_i^j$, where feature index $i = 1, 2, ..., n$, n is the number of features (n = 47), and time step $j = 1, 2, ..., l_i$, $l_i$ is the length of $i^{th}$ feature. Therefore, the decimated $i^{th}$ feature for the synchronized time step $k$ can be described by

$$\tilde{x}_i^k = \sum_{v=0}^{V-1} x_i^{[jD-v]} \cdot h[v] \quad (1)$$

where $\tilde{x}_i^k$ and D indicate the decimated features and integer factor while $h$ and $V$ represent impulse response and its length, respectively. With this approach, the sampling frequencies of the selected features are synchronized to 1 Hz, which results in the synchronization of selected feature length. As an example, the sampling frequency of original VS (16 Hz) and EGT (4Hz) are synchronized to 1Hz as shown in Figure 2 (a) and (b) respectively.

For the features that are recorded initially under 1Hz of the sampling frequency, the Savitzky-Golay filter (Hargittai, 2005) is implemented to reduce their inherent noise while maintaining useful information. The filtered $i^{th}$ feature for the time step $k$ through the Savitzky-Golay filter is described as follows

$$\tilde{x}_i^k = \sum_{u=-\frac{f}{2}}^{\frac{f}{2}} C_u \cdot x_i^{k+u} \quad (2)$$

where $C_u$ and $f$ indicate the convolution coefficient and a set of $C_u$.

### Table 1. Description of selected features

<table>
<thead>
<tr>
<th>Feature name</th>
<th>Description</th>
<th>Unit</th>
</tr>
</thead>
<tbody>
<tr>
<td>ALT</td>
<td>Altitude</td>
<td>ft</td>
</tr>
<tr>
<td>VS</td>
<td>Vertical speed</td>
<td>ft/min</td>
</tr>
<tr>
<td>TAS</td>
<td>True airspeed</td>
<td>knots</td>
</tr>
<tr>
<td>GS</td>
<td>Ground speed</td>
<td>knots</td>
</tr>
<tr>
<td>LGA</td>
<td>Longitudinal acceleration</td>
<td>G</td>
</tr>
<tr>
<td>VTA</td>
<td>Vertical acceleration</td>
<td>G</td>
</tr>
<tr>
<td>LTA</td>
<td>Lateral acceleration</td>
<td>G</td>
</tr>
<tr>
<td>FPAC</td>
<td>Flight path acceleration</td>
<td>G</td>
</tr>
<tr>
<td>AOA</td>
<td>Angle of attack</td>
<td>deg</td>
</tr>
<tr>
<td>TH</td>
<td>True heading</td>
<td>deg</td>
</tr>
<tr>
<td>ROLL</td>
<td>Roll angle</td>
<td>deg</td>
</tr>
<tr>
<td>DA</td>
<td>Drift angle</td>
<td>deg</td>
</tr>
<tr>
<td>WS</td>
<td>Wind speed</td>
<td>knots</td>
</tr>
<tr>
<td>WD</td>
<td>Wind direction</td>
<td>deg</td>
</tr>
<tr>
<td>FSC</td>
<td>Fan speed command</td>
<td>% rpm</td>
</tr>
<tr>
<td>PLA 1~4</td>
<td>Power lever angle input 1~4</td>
<td>deg</td>
</tr>
<tr>
<td>FF 1~4</td>
<td>Fuel flow rate in engine 1~4</td>
<td>lbs/hr</td>
</tr>
</tbody>
</table>
### 3.2. Anomaly Detection using Deep Neural Network

The DNN model is a feedforward NN that includes multiple hidden layers between its inputs and output layer. It is capable of modeling complex and highly nonlinear relationships between input and outputs (Hinton et al., 2012). Main issue in the DNN model is expensive computational cost during the training phase, but computational expense in testing phase is relatively low, which indicates potential applicability of the DNN model to real-time monitoring framework. Therefore, the DNN is utilized in this work to estimate aircraft engine and flight dynamic features for anomaly detection. Initially, the selected features are normalized between 0 and 1 to avoid optimization issue in loss function during the training process. The normalized features are then introduced in the input layer that is formed with 47 neurons to train the DNN model. In this case, each selected feature is estimated with coupled responses in other features for each time step as follows

$$\hat{x}_i = f(\hat{x}_1, \hat{x}_2, \ldots, \hat{x}_n)$$  \hspace{1cm} (3)

where $\hat{x}$ and $\bar{x}$ denote the estimated and normalized feature. Therefore, the number of neurons in the input layer and the output layer are set to be the same. This process enables an integrated estimation of flight features with dependencies between engine system and flight performance, which permits a comprehensive health monitoring with analysis of coupled response under anomalous flight situation. The number of 12 hidden layers are defined with 64 neurons in each hidden layer. In general, rectified linear unit (ReLU) function is used for the activation function to mitigate gradient vanishing or exploding problem in DNN due to the multiplication of too small or large gradient during backpropagation process. However, the ReLU function may result in the deactivation of neurons under certain condition (dead ReLU) (Lau & Lim, 2017). Therefore, leaky ReLU function is utilized in this work to take advantage of ReLU property while avoids dead ReLU issue. The formulation of leaky ReLU function is defined by

$$R(z) = \begin{cases} z, & z > 0 \\ \alpha z, & z \leq 0 \end{cases}$$ \hspace{1cm} (4)

and

$$z = \sum_{i=1}^{K} w_i \hat{x}_i + b$$ \hspace{1cm} (5)

where $\hat{R}(\cdot)$ and $\alpha$ represent the leaky ReLU function and its coefficient while $w$ and $b$ denote weight vector and bias term in hidden layer neuron. Subsequently, the optimized model parameters $w$ and $b$ are obtained by minimizing the loss function. In this work, mean square error (MSE) function that measures the error between ground truth feature and the estimated feature is used as follows

$$\epsilon = \frac{1}{n} \sum_{i=1}^{n} (\bar{x}_i - \hat{x}_i)^2$$ \hspace{1cm} (6)

where $\epsilon$ and $n$ are the MSE and number of measured features. In addition, a strategy called dropout is implemented in this work to enhance the generalization capability of the DNN model, which enables the DNN model to detect off nominal behavior in the FDR dataset. The dropout randomly excludes a certain percentage of neurons in input and hidden layer during training the DNN model, which leads to model averaging to mitigate overfitting problem (Xu, Du, Dai, & Lee, 2014). In this case, flight performance anomalies that significantly deviates from standard behavior may not contribute to train the DNN model. Because the significant off nominal response in the selected feature is a minor portion of the FDR dataset and therefore these trends are most likely ignored during the training process. Therefore, the DNN model parameters are mainly determined by the majority of the FDR data, which is considered as normal response in engine related and flight dynamic features. To fulfill the estimation accuracy and generalization capability, the percentage of dropout is set to be 5%. With this assumption, a large estimation error between features from the FDR and trained model is expected to be observed under the presence of operational anomalies. The estimation error is measured by root mean square error (RMSE) and the coefficient of determination ($R^2$). In particular, the RMSE are assumed to be normally distributed. A normal error bound $\delta_{\epsilon}^{\text{L}}$ is utilized for a safety baseline and defined by the mean and variance of the observed RMSE in the $i^{th}$ features for $k^{th}$ time step as

$$\delta_{\epsilon}^{K} = [\mu_{\epsilon_i}^{K} - \tau \sigma_{\epsilon_i}^{K} \mu_{\epsilon_i} + \tau \sigma_{\epsilon_i}^{K}].$$ \hspace{1cm} (7)

where $\mu_{\epsilon_i}^{K}$ and $\sigma_{\epsilon_i}^{K}$ denote mean and variance of RMSE while $\tau$ is the coefficient to determine distribution bound. In this work, $\tau = 3$ is used to define the normal error bound with the assumption of 0.3% anomaly ratio. Therefore, operational anomalies are detected when the value of $\delta_{\epsilon}^{K}$ violates normal error bound. The anomaly label $\theta$ in observed RMSE in $i^{th}$ feature at $j^{th}$ time step is defined by the binary variable as
\[
\theta_i^j = \begin{cases} 
1, & \text{if } |e_i^k| \leq \delta_i^k \\
0, & \text{otherwise}
\end{cases}
\]  

where 1 and 0 indicate the anomaly and normal flag. In particular, the anomalies are flagged when these off-nominal errors are continued for at least 5 seconds to exclude data acquisition error in this analysis.

4. RESULTS AND DISCUSSIONS

4.1. Evaluation of trained model

In this section, the training process of the DNN model is described and the performance of the trained model is evaluated. Initially, the monitoring features (47 features) are introduced as normalized form (between 0 and 1) to train the DNN model. The normalized features are then scaled back to the original scale in the testing phase for anomaly detection. The adaptive moment estimation (Adam) method is used for the training algorithm to utilize its performance on sparse gradient problem (Kingma & Ba, 2014). Among the 2044 flights investigated in this work, 1635 flights (80% of data) are utilized for training, and 409 flights (20% of data) are used for testing. To obtain the generalized model, the DNN model is trained with a five-fold cross-validation method. In particular, the preprocessed dataset is randomly divided by five segments; four segments are used to train the model, and one segment is utilized to validate the model. The overall estimation errors are measured by 0.0148 RMSE and 0.8432 \( R^2 \) in training phase, and 0.0195 RMSE and 0.7313 \( R^2 \) in testing phase, respectively, which shows a high estimation accuracy of the trained model. An example is presented in Figure 3, where the estimated GS (Figure 3 (a)) and FS (Figure 3 (b)) in testing phase are compared with the recorded GS and FS. In general, the estimated GS and FS show good agreements with the recorded GS and FS in the FDR dataset. The behavior of GS and FS are estimated within 3% error.

4.2. Anomaly detection in historical flight data

Before conducting anomaly detection, the estimation time of the proposed monitoring work is measured to validate computational efficiency for real-time monitoring. The time measurement is conducted by introducing 50 seconds time step of 47 features into the trained DNN model using an Intel Core i5-7200U mobile processor (dual cores with 2.5 GHz base core clock) with 8GB memory. A total of 10 measurements is performed, and the averaged calculation time shows 0.1538 seconds, which demonstrates the real-time applicability of the proposed method.

Subsequently, the proposed monitoring framework is validated through a demonstration of anomaly detection in the FDR dataset. In particular, the selected features during cruise phase are simulated for real-time analysis. The detection time with a duration of anomalies in engine performance and flight dynamic features are analyzed. As shown in Figure 4, sudden changes in engine performance features are detected during cruise phase. It should be note that four features in each engine performance feature are averaged for visualization. For instance, four features in PLA (i.e., PLA 1-4) are averaged and then denoted as PLA in this work. In general, the value of engine performance features such as PLA, FS, CS, and OIT are relatively stable during cruise phase compared to the other flight phases. However, a sudden drop in PLA, FS, and CS is detected at 1749 seconds with a duration of 22 seconds (Figure 4 (a)-(c)) while a sudden increase in OIT is captured at 1769 seconds with same anomaly duration (Figure 4 (d)).
Figure 4. Detected anomalies in engine performance features: (a) PLA, (b) FS, (c) CS, and (d) OIT

These anomalies in engine-related features result in rapid changes in flight dynamic features, as shown in Figure 5. The degree of TAS and AOA is ideally maintained during cruise phase while sudden drop and increase are observed in TAS and AOA respectively. In particular, the anomalies in TAS is detected at 1785 seconds and continued for 223 seconds (Figure 5 (a)) while the anomalies in AOA is captured at 1807 seconds with a duration of 176 seconds (Figure 5 (b)). It is hypothesized that the potential cause of this event may be due to a temporary malfunction of the autopilot system or control mistake by pilots.
Figure 5. Detected anomalies in flight dynamic features: (a) TAS and (b) AOA

5. CONCLUSION

This paper presents a real-time aircraft engine health monitoring framework using DNN, which detects flight performance anomalies during cruise flight operation. A historical flight dataset recorded from multiple onboard sensors in aircraft was utilized to perform the proposed method. Signal processing technique called decimation, and Savitzky-Golay filter was used to synchronize different sampling frequencies and reduce inherent noise in the FDR dataset. The DNN model was adopted to estimate normal engine system behavior related to flight dynamics and trained by the preprocessed features, including engine performance and flight dynamic features. The accuracy of the trained DNN model was evaluated by error measurements (0.0148 RMSE and 0.8432 $R^2$) and the estimation showed good agreements with the features in the dataset. In order to perform anomaly detection, the trained model was combined with a normal estimation error bound for a safety baseline. The real-time detection capability of the proposed monitoring framework was validated through the FDR dataset simulated for real-time and the results showed promising ability in detecting anomalies in engine performance and flight dynamic behavior. Future work will include a development of a monitoring framework that can monitor the whole aircraft system in all flight phases to improve aircraft safety management.

ACKNOWLEDGEMENT

The research reported in this paper was supported by funds from NASA University Leadership Initiative program (Contract No. NNX17AJ86A, Project Officer: Dr. Anupa Bajwa). The support is gratefully acknowledged.

REFERENCES


