JOURNAL OF INFORMATION SCIENCE AND ENGINEERING 37, 1405-1418 (2021)
DOI: 10.6688/JISE.202111_37(6).0012

# Using Deep Learning Approach in Flight Exceedance Event Analysis

HUAN-JYH SHYUR[1], CHI-BIN CHENG[1] AND YU-LIN HSIAO[2]
[1]*Department of Information Management*
*Tamkang University*
*New Taipei City, 251 Taiwan*
*E-mail: {110620; cbcheng}@mail.ntust.edu.tw*
[2]*Department of Industrial and Systems Engineering*
*Chung Yuan Christian University*
*Taoyuan, 320 Taiwan*
*E-mail: yhsiao@cycu.edu.tw*

Causal analysis of flight exceedance events, *e.g.* hard-landing, is a key task for modern airlines performing Flight Operation Quality Assurance (FOQA) programs. The main objective of the program is to learn from experience: detect early signs of major problems and correct them before accidents occur. It has been found that flare operation would greatly influence the landing performance. According to the finding, we proposed a deep learning approach to assist airlines performing causal analysis for hard landing events. Experimental results confirm that compared with the other state-of-the-art techniques, the proposed approach provides a more reliable results. The technique can be the basis of developing advanced models for further revealing the relationships between pilot operations and flight exceedance events.

***Keywords:*** hard landing, quick access recorder, deep learning, BLSTM, RNN

## 1. INTRODUCTION

Quick Access Recorder (QAR) is part of the Aircraft Monitoring System (ACMS) to record the raw data of major aircraft system parameters, and has become an important flight safety management tool today. Civil aviation companies have been using QAR data to assist in the establishment of a Flight Data Monitoring (FDM)/Flight Operations Quality Assurance (FOQA) system. A modern QAR equipment can record more than 2000 airborne parameters in a sampling rate as high as 16 Hz. QAR data provides a way for the airlines to monitor the deviations of flight operations from their normal ranges, and acknowledge the operations by the cockpit crews. Such information is useful for fault analysis, or to discover potential problems of the aircraft. In addition, the analysis results can provide feedbacks to the training unit to formulate necessary training programs for pilots, and thus improve the crew's operational quality to enhance aviation safety. The airline industry has confirmed the merit of QAR in improving flight safety and quality [1].

Current FDM software, such as Aerobytes FDM [2] or AirFASE [3], is able to interpret the QAR data and recreate and animate the flight dynamics of each mission, such as aircraft attitude, instrument indicators, and positions of the control devices. This analysis automatically identifies the flight parameters that overrun their tolerable ranges and deter-

Received December 1, 2019; revised March 31 & May 18, 2020; accepted June 23, 2020.
Communicated by Chien Chin Chen.

mines the event type associated with the flight. A follow-up analysis is carried out to identify the causal operations behind the event when an event waring is triggered. Such a causal analysis is beyond the capacity of the FDM software and has to be performed by human experts, usually the senior pilots or the flight safety supervisors. The way human experts justify the causes of the event are generally to examine the readings of various parameters recorded during the approach-and-landing stage. The high frequency readings and interactions among multiple parameters make this analysis very time-consuming and inefficient, and the interpretation of the data sometimes is often not consistent between different experts due to human subjectivity or experience difference.

In recent years, the advance of computer and information technologies has boosted the development of artificial intelligence (AI) and enabled the applications of the technique to real-world problems. Among the AI techniques, deep learning that roots in artificial neural networks has been proved to be powerful in image recognition, speech recognition, language translation, *etc.* Many tasks that traditionally rely on human judgment can be replaced by deep learning models with an extensive amount of training data. Jasra *et al.* [4] provide a good literature review of applying machine learning techniques to analyze FDM data. Oehling and Barry [5] used an unsupervised learning method called "Local Outlier Probability" to generate safety-relevant knowledge from existing FDM data. Nanduri and Sherry [6] applied the Recurrent Neural Network (RNN) with Long Short-Term Memory (LSTM) to detect unknown and unusual patterns by using FDM data. Li *et al.* [7] proposed a Cluster based Anomaly Detection (ClusterAD) approach to analyze the flight phases of takeoff and final approach by transferring FDM data into high dimensional vectors to capture the multivariate and temporal features of flights. In this study, aiming to improve the efficiency of QAR-based event causal analysis, we employ deep learning models to identify the causes by using QAR data as inputs.

The critical limitation of applying machine learning techniques to flight event casual analysis is the event records are relatively scarce. This study serves as an initial investigation of the feasibility of using such techniques in this problem domain and to examine how well the techniques can perform and how big the gap is if the techniques fail to achieve the goal. This study particularly chooses the event of hard landing as the subject for it accounts for a large share of all events. The event of hard landing refers to the impact on the aircraft when landing with a large vertical velocity and force. For example, according to the Boeing Flight Crew Training Manual, Boeing commercial aircraft are designed with a maximum of 600 FPM (feet per min) for landings, with 60-180 FPM as an ideal speed, while exceeding 240 FPM will be triggered as a hard landing. With a hard landing event, even the pilot can still control or at least partially control the aircraft without creating significant damage to the aircraft, the impact may still create incidences varying from minor passenger discomfort to structural failure, or even injury and/or loss of lives.

We worked with an international airline to collect the historical QAR data that are triggered as hard landing events, as well as the safety reports associated with these events. Through in-depth interviews with safety supervisors and senior pilots, five major causes of hard landing events were identified and used to label the QAR data. The multiple parameters collected over time by QAR make the inputs in the form of multi-dimensional time series data. To deal with the time series nature, a particular deep learning method, end to end Bi-directional Long Short-Term Memory (BLSTM), was introduced to establish the models. The proposed model was evaluated by empirical data. The results confirm the proposed approach provides a more reliable results then the other state-of-the-art methods.

The rest of the paper is structured as follows. Sections 2 discusses the causes of flight hard landing. Section 3 presents the proposed machine leaning model. Performance evaluation is carried out in Section 4. Concluding remarks are given in Section 5.

## 2. BACKGROUND

Pilot's operation will have a direct impact on flight safety, because the pilot may make mistakes in an emergency or unexpected situation [8, 9]. The occurrence of a hard landing event could be traced to many complicated and interactive factors, such as the weather conditions, the mechanical problems of the aircraft, the overload of the aircraft, the improper allocation of aircraft cargo loading, and the pilot's skills or mental state [10-14]. These factors continuously change over time during the entire journey of the flight, and such changes are reflected in the readings of aircraft attitude and kinematic parameters by QAR device. Though such high-frequency recording of a thorough set of parameters possibly contains all clues we need to discover the causes of event, such massive amount of data hinders the analysis by human. To make the analysis doable, human experts (*i.e.* safety supervisors and senior pilots) generally concentrate on a much smaller set of data beginning at the approaching and landing phase, which is considered to be more relevant to the hard landing event. The present study adopts the same strategy, and collects the data after the flight entering the approach and landing phase, with a focus on the same set of parameters that the human experts take into account.

Each event comes with a safety report that was prepared by safety supervisor and the fleet crews. The reports are the important source for us to categorize the causes. However, the writing of the report is generally in a free style, and the causes analyzed by the human experts are not consistently defined; it is commonly to see different terms being used to describe the same cause. We have performed many in-depth interviews with the senior pilots to clarify the definitions of causes and attribute the event to a fixed set of causes.

### 2.1 Data Collection

The QAR data we collected spanning from 2006 to 2017, during which the number of parameters, the names of parameters and the data format had been changed a few times. We unified the names of the parameters that are used for the causal analysis of hard landing events. A few QAR records are not in a digital format and have to be discarded due to the failure of the optical character recognition (OCR) operation. The resulting numbers of records for different models are: 161 instances for B777, 32 for B747, 43 for A321, and 5 for A330. Besides B777, the data are very scarce for the other models, thus, this study establishes analysis models for B777 only.

Parameters are sampled at different frequencies by the QAR device, and hence resulting in many blanks in the data table. The way we process such data is to discard the time instances with blank readings. As a result, the sampling rate of QAR data prepared for analysis is 1 second.

Safety reports of hard landing events were all kept as paper copies stored by the airline. We scanned every report and used OCR software to convert the scanned files to computer-readable text, so we could identify the keywords in the reports by programming tools, as shown in Fig. 1, in which, the most frequent keywords are: "*insufficient flare*". These keywords assist us in discussing with senior pilots the major causes of hard landing.

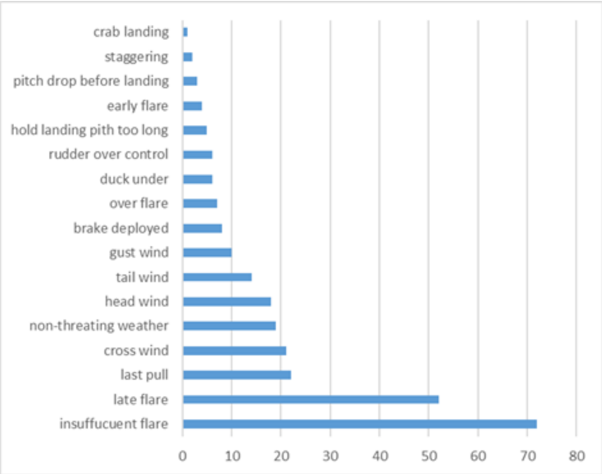


Fig. 1. Frequencies of keywords from the safety reports of hard landing events.

## 2.2 Grouping of Major Causes of Hard Landing

As mentioned earlier, the occurrence of hard landing events could involve many factors. In the interviews with senior pilots, they pointed out that the events were mainly due to improper maneuvers at the flaring phase. Pilot's inappropriate operation of the aircraft or the pilot's insufficient alertness may lead to a hard landing event. The purpose of the flare maneuver is to reduce the vertical speed to a safety range and enable the main landing gears touch the ground first [12, 15]. At the end of the interview process, we reach an agreement on the grouping of hard landing causes, which are all associated with the flare maneuver. These causes are presented and defined in Table 1.

**Table 1. Categorization of hard landing causes.**

| Group | Cause | Definition |
|---|---|---|
| Improper flare attitude | Insufficient flare | The flare maneuver fails to maintain a proper pitch attitude (*i.e.* insufficient pitch angle), and hence unable to reduce the vertical speed to an acceptable range and cause a firm landing. |
| | Over flare | An over pull of the control column and hence producing an exceeding pitch angle, which results in insufficient power and cause a firm landing. |
| Improper flare timing | Late flare | Flare begins below a normal flare height, and hence unable to reduce the vertical speed to an acceptable range and cause a firm landing. |
| | Early flare | Flare begins above a normal flare height. As a result, the speed is insufficient to provide enough power and hence cause a firm landing. |
| Improper control input | Last moment input | At the last moment of landing, a sudden pull of control column results in a change in the pitch attitude. |

To map the QAR records to the causes in Table 1, we reviewed all safety reports associated with events during the sampling period. These safety reports were filed by safety supervisors and the fleet crews after the events to comment on the causes of events. However, the writing of the report was generally in a free style, and the causes analyzed by the human experts are not consistently defined. We had performed many in-depth interviews with the senior pilots to clarify the definitions of causes and attribute the causes to each event instance.

### 2.3 Parameters Selection

The pilot controls the aircraft by adjusting engine thrust and aircraft attitude. In the final landing stage, pilots maintain the aircraft to fly within the profile of a landing glide path. Any slight backward pulling of the control column can change the pitch attitude and airspeed, and hence change the vertical velocity of the aircraft. Wang *et al.* [14] used 3 parameters, touchdown distance, vertical acceleration, and pitch angle, to evaluate landing operation performance. The QAR data we obtained from the case airline contain 48 parameters. Our feature selection process began with an initial screening to exclude apparently non-influential parameters. We then consulted a group of experts for confirming the screening process and recommending the final features from the remainder candidate parameters. We also performed paired sample *t*-tests on the data to justify the experts' decision. 50 normal flight data and 50 hard landing flight data were random selected to perform the test. The QAR data with a radio height of less than 50 feet before the final landing was selected for statistical analysis. Taking every 10 feet as a unit, calculate the mean value of each parameter for each selected QAR record. We verify that when different parameters are at different flight altitudes, there are significant differences between hard landing flights and normal flights. Under 95% confidence interval, results show only the pitch attitude, ground speed, and vertical speed have significant difference in more than one group of data.

After the interviews with 5 safety supervisor and senior pilots, we selected 7 parameters as independent variables in our model (as presented and described in Table 2). Among which, radio height and control column position provide information regarding the initiation of flaring. Pitch attitude, calibrated air speed, and N1 indicator are associated with the flare attitude, *i.e.* insufficient or over. The vertical speed is a resulting indicator of the flight maneuver, and is the direct cause of hard landing. The correlations between the above parameters are exemplified by the cases presented in Fig. 2. The case in Fig. 2 (a) shows that at radio height 25feet, the pitch angle is around 2.5 degrees with the vertical speed greater than 500feet/min. According to the standard maneuver of Boeing B777, the pitch attitude ought to be maintained within 3-4 degrees at this height, and thus the hard landing event is blamed on insufficient flare. In contrast, the second case shown in Fig. 2 (b) is a landing event due to over flare, where the pitch angle is sharply raised to 4.4 degrees at radio height 12 feets. Fig. 2 (c) is a case of early flare compounded with a last moment input, where the flare initiates at the radio height 55feets (evidenced by the pitch change), followed by a rapid decrease of pitch at 30feet, and at the last moment the pilot decided to raise the aircraft nose that caused a hard landing when the pilot acknowledged a high vertical speed of 540 feet/min at radio height 12 feet. At last, Fig. 2 (d) shows a late flare case, where the pitch angle was raised very late to 2.33 degrees when the radio height is 13 feet left to landing.

**Table 2. The set of QAR parameters for hard landing causal analysis.**

| Parameter | Description |
|---|---|
| Control column position | The position of captain's control column. |
| Ground speed | The horizontal speed of an *aircraft* relative to the ground |
| Pitch attitude | The angle between the longitudinal axis of an aircraft and the local horizontal. |
| Radio height | The altitude of the aircraft above the terrain presently beneath the aircraft. |
| Vertical speed | The rate of climb or descent of an aircraft |
| Calibrated air speed | The corrected air speed of an aircraft for instrument errors, position errors, and installation errors. |
| N1 indicator | A cockpit gauge which presents the rotational speed of the low speed engine spool. The gauge is usually calibrated in percent RPM based on an engine manufacturer defined rotational speed that corresponds to 100%. |

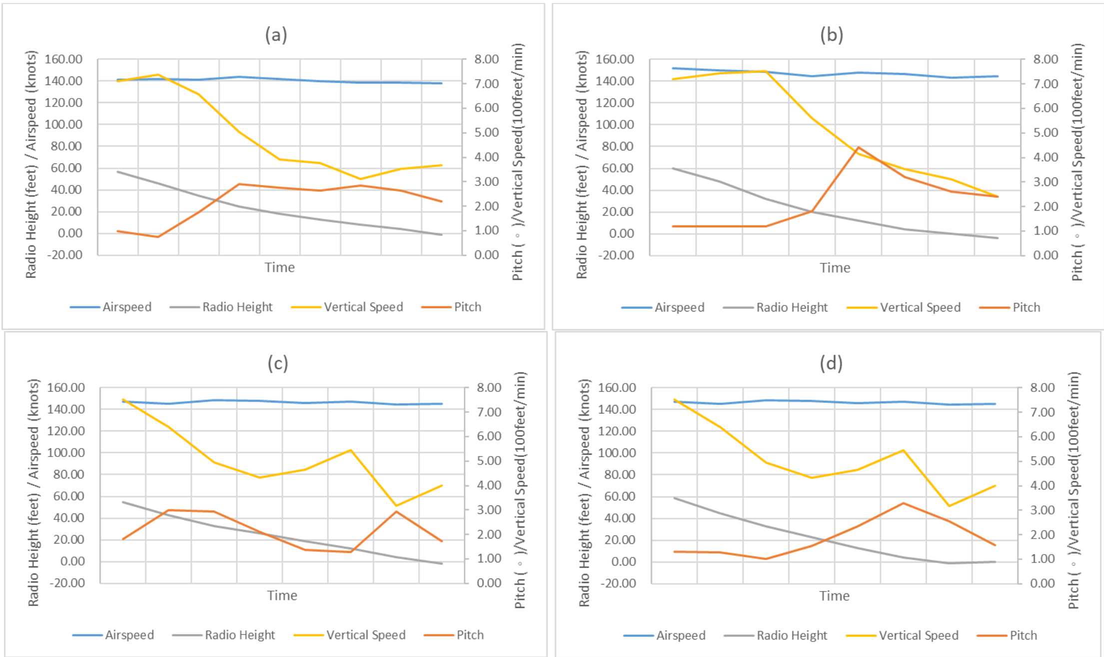



Fig. 2. The relationship between selected features and events.

## 3. PROPOSED APPROACH

The proposed approach is an extension of RNNs that connects the network outputs back to the input end, so that the output value of the last time point can be transmitted back to the neuron. With such a recurrent feature, the network is not only able to memorize and record the inputs at the previous point in time, but also store the information of the chronological order of the inputs. RNNs are powerful tools for modeling sequential data since they are capable of learning and maintaining a set of memory cells overtime [16]. However, training them by back-propagation through time can be difficult [17]. The gradient vanishing problem that is usually found with the traditional RNN [18].

### 3.1 End to End BLSTM

LSTM introduced by Hochreiter & Schmidhuber [19] is a special kind of RNN. It is currently the most popular RNN. The neuron of LSTM consists of three gates to control the learning of memory overtime: a forget gate, an input gate and an output gate as shown in Fig. 3.

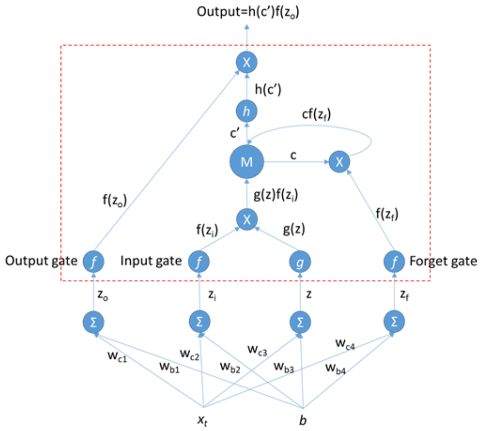


Fig. 3. The memory cell of LSTM.

The input gate controls the admission of an input to enter the memory by multiplying the activation of input $z$, *i.e.* $g(z)$, with the input gate activation $f(z_i)$; the forget gate determines the degree that a memory is kept by multiplying the old memory $c$ with the forget gate activation $f(z_f)$, and the memory is updated by $c' = g(z)f(z_i) + cf(z_f)$; and the output gate controls how much of the memory is passed to other cells. The above operations allow the previous inputs being kept in the memory cell until the forget gate is closed, and enable the network to learn and determine how long to hold the old memory, and how to associate the old memory with the new inputs. In the past few years, LSTMs have been applied to a variety of problems and received incredible successes in domains such as: speech recognition, language modeling, translation, image captioning, *etc.* [20-23]

Box *et al.* [24] argued that the performance of time-series-based prediction can be enhance if the system is modeled with both forward and backward temporal perspectives. Based on this idea, Schuster and Paliwal [25] proposed the bidirectional RNN to improve the prediction performance of RNN. Using the same idea, bidirectional LSTM (BLSTM) extends the standard LSTM network to enhance the prediction performance [20]. The structure of a BLSTM network is presented in Fig. 4. In the network, two hidden layers from the two opposite direction connect to the same output, where one feeds forward and another one backwards in time. It learns the representation of data from previous time steps and future time steps simultaneously to increase the amount of input information available to the network. A specified BLSTM, end to end BLSTM, was proposed to construct out deep learning model, where only the last outputs of the forward and backward output sequences are concatenated together and returned to the next network layer (see Fig. 5). It simples the learning process and provides the benefit of more input information.

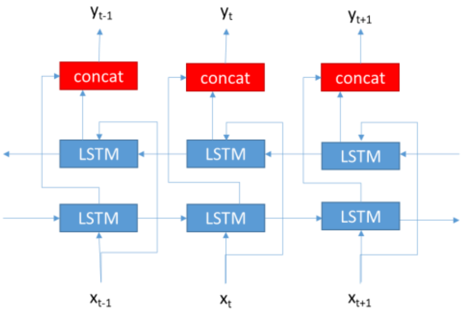



Fig. 4. Architecture of BLSTM.

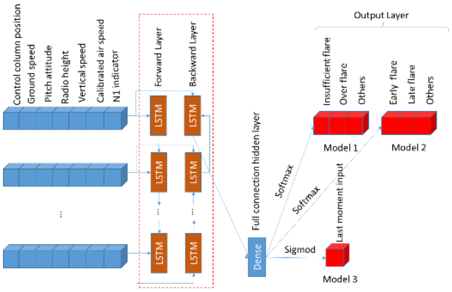



Fig. 5. The proposed network model.

### 3.2 Proposed Network Model

With QAR data as inputs, the models classify the inputs to one of the causes presented in Table 1. It is common that a hard landing event is attributed to multiple causes by the human experts. For example, in some safety reports, the cause of insufficient flare is followed by a last moment input; such a combination of causes often occurred with fresh pilots who attempted to pull up the craft just before landing when they found the pitch degree is not enough. Combinatorial causes would complicate the network outputs and make it difficult to learn, especially with the limit amount of training data. Thus, instead of using a single model to model all possible cause combinations, we use three separate subnetworks, each corresponding to a cause group in Table 1, to synthesize the causes of a hard landing events. The same set of QAR data of an event is input to the three respective networks, and the resulting outputs from the three subnetworks are synthesized as the causes of the event.

The complete network model we used to perform the causal analysis are depicted in Fig. 5, which consists of an end to end BLSTM layer, a fully-connected hidden layer, and an output layer. To reduce the model complexity and provide additional context to the network, end to end BLSTM networks are employed to model the hard landing causal analysis. As mentioned above only the last outputs of the forward and backward output sequences are concatenated together and returned to the next network layer. The function of the end to end BLSTM layer is to capture the patterns of QAR time series data. It holds both the final information from forward and backward layers. The hidden layer is to facilitate the nonlinear mapping between the input and the output, and the output layer is to produce causal categories. Inside the end to end BLSTM layer, the hyperbolic tangent function is used as the activation functions for all gates of the memory cells. To facilitate the generalization of our models to future analysis, a dropout operation [26] is applied during the learning of weights. To prevent the network from overfitting, the dropout operation gives up the updating of some connection weights during the learning process.

The causes of hard landing in Table 1 are encoded with an operation named one-hot encoding [27] to facilitate the classification by the above network models. One-hot encoding is a process by which categorical variables are converted into a form of binary vector consisting of 0s in all cells with the exception of a single 1 in a cell used uniquely to indicate a certain category.

The output nodes of the network are embedded with a softmax activation function. The softmax function assigns decimal probabilities to each cause, and these decimal probabilities must add up to 1.0.

## 4. PERFORMANCE EVALUATION WITH EMPIRICAL DATA

All models are implemented and run on a desktop computer with 6 cores Intel i7 CPU and 32G RAM. The models are implemented with TensorFlow.

### 4.1 Training and Testing Data

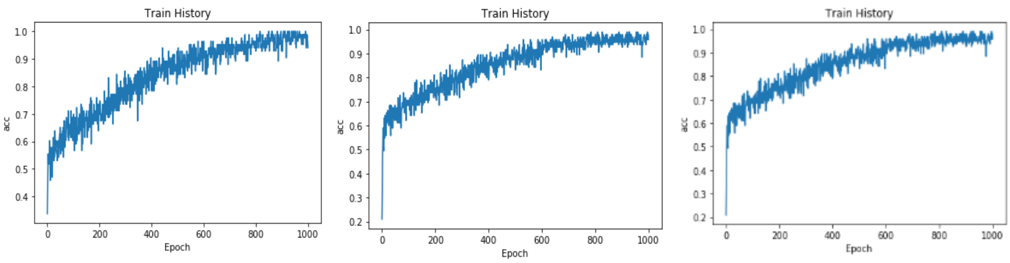
QAR records of 202 flights of model B777 were collected and processed as described in Section 2.1, and used as the dataset to build the causal analysis models. Due to the limited data sample, the 10-fold cross-validation procedure is used to evaluate predictive models by partitioning the original data into a training set to train the model, a validation set to evaluate it, and a test set to provide an unbiased evaluation. In which, 20% of the dataset are selected randomly in each iteration as the testing data and the remaining 80% are used as the training and validation data. The last 45 seconds of the QAR records before aircraft touched the ground were extracted as the analysis data. To prevent distorting differences in the ranges of values or losing information in the learning process, three columns of the raw data are rescaled. Radar Height is rescaled by using 100 feet as the measuring unit, and Vertical Speed and Calibrated Air Speed are measured in a unit of 100 FPM.

### 4.2 Parameter Settings for Model Training

The selection of the number of memory cells in the recurrent neural network model is usually application-dependent. Using a greater number of memory cells at a hidden layer generally produces a greater prediction accuracy on the training data, but it would increase the computation load and lose model generality, *i.e.* unable to predict with future unseen data. We experimented different sizes of memory cells (from5 to 25) for the RNN layer, and determined to set the number to 10, which best balanced the computation load and the predict accuracy in our experiment. The RNN layer is connected to the output layer with a dropout rate of 0.1. The connection weights of the models were randomly initialized before training, and ‘adam’ was chosen as the optimizer. The training is performed for 1000 epochs with a batch size of 8.

### 4.3 Model Performance Evaluation

Fig. 6 presents the training accuracies of the end to end BLSTM models for flare attitude, flare timing and control input, respectively. The training accuracies of the three models all converge after 600 epochs. For comparison purpose, we consider LSTM and another two non-time-series based techniques, multinomial logistic regression (MLR) and back-propagation neural network (BPN). The time series data was converted to block-type input. Table 3 presents the average prediction results by the LSTM, end to end BLSTM, MLR and BPN. Macro-average method was adopted for combining the measures for individual models. The BPN models in this experiment have an input layer with 15×7 nodes, two hidden layers with 25 and 30 hidden nodes respectively, and an output layer with the same structure of LSTM and end to end BLSTM models. It is seen from Table 3 that LSTM and end to end BLSTM outperform the other two models. Evaluation of the testing results, only the average accuracies of the LSTM and BLSTM models are more than 70% in the ten trials. It also can be found that the BLSTM is an effective mechanism to learn how to identify the flare timing related causes because of the 80% testing accuracy. We consider it is due to the time sequence features accommodated by RNN type of models can potentially improve the effectiveness of cause identification by QAR data. It is also noted that end to end BLSTM models outperform the LSTM models, because end to end BLSTM provides two times of input information to the learning model. In addition, we have performed experiments with the traditional SVM method. Their performances were very closed to that of the BPN. The average training accuracy for flare attitude, flare timing and control input models are 0.72, 0.74, 0.70, respectively. The following metrics were employed to evaluate the performance of the proposed localization scheme.




Flare attitude model　　Flare Timing model　　Control input model

Fig. 6. Plot of training accuracy of end to end BLSTM models.

**Table 3. Accuracy comparisons between different models.**

| Models | | MLR | | | BPN | | | LSTM | | | end to end BLSTM | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | | A | T | L | A | T | L | A | T | L | A | T | L |
| Training Data | Accuracy | 0.73 | 0.75 | 0.70 | 0.71 | 0.74 | 0.72 | 0.83 | 0.83 | 0.80 | 0.88 | 0.90 | 0.81 |
| | Precision | 0.62 | 0.58 | 0.64 | 0.60 | 0.58 | 0.65 | 0.76 | 0.68 | 0.80 | 0.88 | 0.85 | 0.82 |
| | Recall | 0.71 | 0.72 | 0.63 | 0.66 | 0.72 | 0.64 | 0.79 | 0.77 | 0.69 | 0.91 | 0.93 | 0.70 |
| | F1 | 0.64 | 0.60 | 0.63 | 0.62 | 0.60 | 0.65 | 0.77 | 0.71 | 0.71 | 0.89 | 0.88 | 0.72 |
| Testing Data | Accuracy | 0.56 | 0.62 | 0.61 | 0.64 | 0.63 | 0.68 | 0.71 | 0.73 | 0.78 | 0.71 | 0.80 | 0.79 |
| | Precision | 0.75 | 0.55 | 0.52 | 0.75 | 0.56 | 0.60 | 0.77 | 0.63 | 0.77 | 0.79 | 0.70 | 0.71 |
| | Recall | 0.58 | 0.57 | 0.52 | 0.60 | 0.58 | 0.58 | 0.77 | 0.66 | 0.67 | 0.65 | 0.72 | 0.65 |
| | F1 | 0.60 | 0.54 | 0.52 | 0.64 | 0.55 | 0.58 | 0.77 | 0.63 | 0.69 | 0.69 | 0.70 | 0.67 |

A: Improper flare attitude　T: Improper flare timing　L: Improper last moment pull

Although the end to end BLSTM models outperform the others, their prediction accuracies are not high enough to replace a human expert. More than 70% (11 out of 15) of the prediction failure by the proposed models have similar patterns. A closer look reveals that the trained models have difficulty to distinguish between *late flare* and *insufficient flare* events from the QAR data. A common case occurs when the hard landing is due to *late flare*, but the model will determine the cause due to both *insufficient flare* and *late flare*. During the interview phase with the pilots and safety supervisors, it was discovered that the boundary between the causes of *insufficient flare* and *late flare* can be very confusing sometimes. Thus, we suspect that such an undesired prediction result might be due to the inconsistent interpretation of QAR data when safety supervisors making judgement on the causes of hard landing. To confirm our suspicion, we invited safety supervisors to double check the review reports, and they agreed that the cause identification in some reports were not consistent. Such a finding is not surprising for inconsistency and vagueness being common natures in human's subjective judgement. The mislabeled training data hindered the learning of our models especially when the data amount was limited.

## 5. CONCLUSIONS AND DISCUSSIONS

Fewer studies paid attention to the crucial flare operation in landing. Hard landing affects the flight safety seriously. With the analyze of QAR data, and perform many in-depth interviews with the senior pilots, this study aims to automate the process of causal analysis of flight hard landing events by deep learning approaches. The proposed causal analysis model has been implemented on a web-based Smart FOQA system by the case airline. By accepting sequences of multivariate QAR data points as input, pre-trained BLSTM models can be used for real-time to assist safety supervisors in identifying the possible causes of hard landing and preparing safety report. This research successfully shows that the chosen approach is able to detect the possible causes of hard landing event by analyze QAR data than traditional neural network models. The work can be the fundamental of developing advanced models for further revealing the relationships between pilot operations and flight exceedance events.

Two main limitations of our proposed models are the amount of available cases for training, and the labeling quality of the training data. The limited amount of training cases makes the models difficult to learn patterns of cause identification, and the inconsistent judgement of hard-landing causes doubles the difficulty from limited training cases. The limited training cases also restrict the size of the proposed model. To avoid over-fitting, we constrained the number of nodes in the end to end BLSTM layer to 10 or less. Despite the above limitations, we consider that the proposed model provides a significant advance in QAR data analysis. They have the potential to serve as a basis for the future establishment of an automatic flying event analysis system.

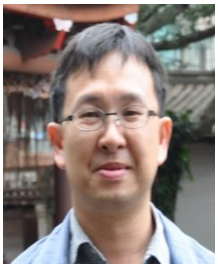
Class imbalance was also found in our training dataset, where in the first causal category, 55% of the hard landing events were attributed to *insufficient flare*, while only 6% were due to *over flare*. Similarly, in the second causal category, 40% of the events were attributed to *late flare* and only 3% were considered as *early flare*. The imbalance of classes generally makes the models over-fitting to the major classes.

In this paper, an efficient RGL scheme was presented for locating sensor nodes and achieving fine-grained accuracy by employing range-free and RSSI-based localization schemes. All computations were performed locally; thus, the mechanism is distributed, scalable, effective, and energy-efficient. The proposed technique is supported by a mobile beacon with a GPS. The beacon moves along a specific trajectory and periodically broadcasts information. Each sensor node receives the packets without interacting with other nodes for localization information. The distributed computation of node position involving elementary operations allows the system to operate while consuming minimal power. The experimental results were based on various parameters (*e.g.*, localization error, execution time, system throughput, and energy consumption), showing that the performance of the RGL scheme outperforms well-known RSSI-based schemes.

## REFERENCES

1. Flight Safety Foundation, *Guide to Methods & Tools for Airline Flight Safety Analysis*, 2003.
2. Aerobytes, https://www.aerobytes.co.uk/aerobytes_fdm.htm, 2015
3. AirFASE, https://www.navblue.aero/product/airfase, 2017
4. S. Jasra, J. Gauci, A. Muscat, and G. Valentino, "Literature review of machine learning techniques to analyze flight data," *Advanced Aircraft Efficiency in a Global Air Transport System*, At Toulouse, France, 2018.
5. J. Oehling and D. J. Barry, "Using machine learning methods in airline flight data monitoring to generate new operational safety knowledge from existing data," *Safety Science*, Vol. 114, 2019, pp. 89-104.
6. A. Nanduri and L. Sherry, "Anomaly detection in aircraft data using Recurrent Neural Networks (RNN)," in *Proceedings of Integrated Communications, Navigation and Surveillance Conference*, 2016, pp. 5C2-1-5C2-8.
7. L. Li, S. Das, R. J. Hansman, R. Palacios, and A. Srivastava, "Analysis of flight data using clustering techniques for detecting abnormal operations," *Journal of Aerospace Information Systems*, Vol. 12, 2015, pp. 587-598.
8. M. Ebbatson, D. Harris, J. Huddlestone, and R. Sears, "Combining control input with flight path data to evaluate pilot performance in transport aircraft," *Aviation, Space, and Environmental Medicine*, Vol. 79, 2008, pp. 1061-1064.
9. D. Harris, *Human Performance on the Flight Deck*, Farnham, Ashgate Publishing, UK, 2011.
10. D. Benbassat, C. I. Abramson, and K. W. Williams, "Comparative approach to pilot error and effective landing flare instructions," *International Journal of Comparative Psychology*, Vol. 15, 2002, pp. 249-255.
11. D. Benbassat and C. I. Abramson, "Landing flare accident reports and pilot perception analysis," *The International Journal of Aviation Psychology*, Vol. 12, 2002, pp. 137-152.
12. D. Benbassat and C. I. Abramson, "General aviation landing flare instructions," *The Journal of Aviation/Aerospace Education & Research*, Vol. 11, 2002, pp. 31-38.
13. W. Lei, C. Wu, and R. Sun, "An analysis of flight quick access recorder (QAR) data and its applications in preventing landing incidents," *Reliability Engineering and System Safety*, Vol. 127, 2014, pp. 86-96.

14. L. Wang, J. Zhang, C. Dong, H. Sun, and Y. Ren, "A method of applying flight data to evaluate landing operation performance," *Ergonomics*, Vol. 62, 2019, pp. 171-180.
15. Flight Safety Foundation, "Stabilized approach and flare are keys to avoiding hard landings," *FSF Digest*, Vol. 23, 2004.
16. A. Graves, A. R. Mohamed, and G. Hinton, "Speech recognition with deep recurrent neural networks," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, 2013, pp. 6645-6649.
17. S. Hochreiter, Y. Bengio, P. Frasconi, and J. Schmidhuber, "Gradient flow in recurrent nets: the difficulty of learning long-term dependencies," *A Field Guide to Dynamical Recurrent Neural Networks*, IEEE Press, 2001.
18. R. Pascanu, T. Mikolov, and Y. Bengio, "On the difficulty of training recurrent neural networks," in *Proceedings of the 30th International Conference on Machine Learning*, 2013, pp. 1310-1318.
19. S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Computation*, Vol. 9, 1997, pp. 1735-1780.
20. A. Graves and J. Schmidhuber, "Framewise phoneme classification with bidirectional lstm and other neural network architectures," *Neural Networks*, Vol. 18, 2005, pp. 602-610.
21. X. Li and X. Wu, "Constructing long short-term memory based deep recurrent neural networks for large vocabulary speech recognition," in *Proceedings of IEEE International Conference on Acoustics, Speech and Signal Processing*, 2014, pp. 4520-4524.
22. N. Tax, I. Verenich, M. La Rosa, and M. Dumas, "Predictive business process monitoring with LSTM neural networks," in *Proceedings of the International Conference on Advanced Information Systems Engineering*, 2017, pp. 477-492.
23. T. Yao, Y. W. Pan, Y. Li, Z. F. Qiu, and T. Mei, "Boosting image captioning with attributes," in *Proceedings of IEEE International Conference on Computer Vision*, 2017, pp. 4894-4902.
24. S. Padmanaban and S. M. E. Mahendran, "Enhancing flight data monitoring and analysis can increase flight safety," *Journal of Aeronautics & Aerospace Engineering*, Vol. 4, 2015, pp. 1-5.
25. M. Schuster and K. K. Paliwal, "Bidirectional recurrent neural networks," *IEEE Transactions on Signal Processing*, Vol. 45, 1997, pp. 2673-2681.
26. N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: a simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, Vol. 15, 2014, pp. 1929-1958.
27. S. L. Harris and D. M. Harris, *Digital Design and Computer Architecture*, Morgan Kaufmann, San Francisco, 2013.
28. M. Fortunato, C. Blundell, and O. Vinyals, "Bayesian recurrent neural networks," in *Proceedings of the 31st Conference on Neural Information Processing Systems*, 2017, arXiv:1704.02798.



**Huan-Jyh Shyur (徐煥智)** received his Ph.D. degrees in Industrial and Systems Engineering from Rutgers University, USA. He is currently a Professor of Information Management at Tamkang University, Taiwan. His research interests include machine learning, decision theory, and software reliability.

**Chi-Bin Cheng (鄭啟斌)** received his Ph.D. degree in Industrial and Manufacturing Systems Engineering from Kansas State University, USA. He is a Professor of Information Management with Tamkang Univeristy, Taiwan. His main research interests are in the fields of soft computing, swarm intelligence, group decision-making, multi-agent systems, and supply chain management.

**Yu-Lin Hsiao (蕭育霖)** received his Ph.D. degree in Industrial and Systems Engineering from State University of New York, USA. He is currently an Assistant Professor of Industrial and Systems Engineering at Chung Yuan Christian University, Taiwan. His research interests include ergonomics, risk analysis, and machine learning.