International Conference on Recent Trends in Engineering & Technology- 2023 (ICRTET-3)
Organised by: VSM College of Engineering, Ramachandrapuram

E-Pilots: A System to Predict Hard Landing During theApproach Phase of Commercial Flights

Ch. Anuradha[1], S.N.Uday Kiran[2], M.Lakshman[3], M.Poojitha[4], M.Gayathri[5], Ch.Chandini[6]

[1]Assistant Professor, [2-6]Students B.Tech Computer Science Engineering, V.S.M College of Engineering, Ramachandrapuram, A.P, India

---***---

Abstract

More than half of all commercial aircraft operation accidents could have been prevented by executing a go- around. Making timely decision to execute a go-around manoeuvre can potentially reduce overall aviation industry accident rate. In this paper, we describe a cockpit-deployable machine learning system to support flightcrew go-around decision-making based on the prediction of a hard landing event. This work presents a hybrid approach for hard landing prediction that uses features modelling temporal dependencies of aircraft variables asinputs to a neural network. Based on a large dataset of 58177 commercial flights, the results show that our approach has 85% of average sensitivity with 74% of average specificity at the go-around point. It follows that our approach is a cockpit-deployable recommendation system that outperforms existing approaches.

Key words - Decision support systems, Hard landing prediction, Machine learning, Neural networks

I. Introduction

Between 2008-2017, 49% of fatal accidents involving commercial jet worldwide occurred during final approach and landing, and this statistic has not changed in several decades [1]. A considerable proportion of approach and landing accidents/incidents involved runway excursions, which has been identified as one of the top safety concerns shared by European Union Aviation Safety Agency (EASA) member states [2], as well as US National Transportation Safety Board and US Federal Aviation Administration [3].

According to EASA [2], there are several known precursors to runway excursions during landing. These include unstable approach, hard landing, abnormal attitude or bounce at landing, aircraft lateral deviations at high speed on the ground, and short rolling distance at landing. Some precursors can occur in isolation, but they can also cause the other precursors, with unstable approach being the predominant one. Boeing reported that whilst only 3% of approaches in commercial aircraft operationmet the criteria of an unstable approach, 97% of them continued to landing rather than executing ago-around [4].

A study conducted by Blajev and Curtis [5] foundthat 83% of runway excursion accidents in their 16-year analysis period could have been avoided by a go-around decision. Therefore, making timely decision to execute a go-around manoeuvre could therefore potentially reduce the overall aviation industry accident rate [4]. A go-around occurs when the flight crew makes the decision not to continue an approach or a landing, and follows procedures to conduct another approach or to divert to another airport. Go-around decision can be made by either flight crew members, andcan be executed at any point from the final approach fix point to wheels touching down on therunway (but prior to activation of brakes, spoilers, or thrust reversers). In addition to unstable approaches, traffic, blocked runway, or adverse weather conditions are other reasons for a go- around. Despite a clear policy and training on go-around policies in most airlines, operational data show that flight crew decision-making process in deciding for a go-around could be influenced by many other factors. These include fatigue, flight schedule pressure, time pressure, excessive a head-down work, incorrect anticipation of aircraft deceleration, visual illusions, organizational policy/culture, inadequate training or practice, excessive confidence in the ability to stabilize approach, and Crew Resource Management issues [5]. It is for these reasons that on-board realtime performance monitoring and alerting systems that could assist the flight crew with the landing/go-around decision are needed [5], [6].

Such on-board systems could utilize the huge and ever increasing amount of data collected from aircraft systems and the exponential advances in machine learning methods and artificial intelligence. EASA is anticipating a huge impact of machine learning on aviation, including helping the crew to take decisions in particular in high workload circumstances (e.g. go-around, or diversion [7]. Artificial Intelligence in aviation is considered one of the strategic priorities in the European Plan for Aviation Safety 2020-2024 [8].

Under the hypothesis that a hard-landing (HL) occurrence has precursors and, thus, it can bepredicted, this paper presents a cockpit deployablemachine learning system to predict hard landings considering the aircraft dynamics and configuration. In particular, this paper evaluates three main hypothesis. A primary hypothesis is to assess to what extend HL can be predicted at DH for go-around recommendation from the analysis of the variables recorded from FMS. A second hypothesis is to analyze if precursors are particular to aircraft types. A third hypothesis is to validate if the variability on the aircraft state variables can provide enough information to predict a HL regardless of the operational context (like environmental conditions and automationfactors).

CONTRIBUTIONS

This paper presents an analysis of approaches for early prediction of hard-landing events in commercial flights. Unlike previous works, experiments are designed to analyse to what extend methods can be deployable in the cockpit as go-around recommendation systems. With thisfinal goal, we contribute to the following aspects:

1) **Hybrid model with optimized net architecture.** We propose a hybrid approach that uses features modelling temporal dependencies of aircraft variables as input to a neural network with an optimized architecture. In order to avoid any bias caused by a lack of convergence of complex models (like LSTM), we use a standard network and model potential temporal dependencies associated with unstable approaches as the variability of different types of aircraft variables at a selected set of altitudes. The concatenation of such variability for variables categorized into 4 main types (physical, actuator, pilot operations and all of them) are the input features of different architectures in order to determine the optimal subset.

2) **Exhaustive comparison to SoA in a large database of commercial flights.** A main contribution compared to existing works is that our models have been tested and compared to SoA methods on a large database of Flight Management System (FMS) recorded data of an airline no longer in operation that includes 3 different aircraft models (A319, A320, A321). Results show that the optimal classification network when all variable types are considered achieves an average recall of HL events of 85% with a specificity of 75% in average, which outperforms current LSTM methods found in the literature. Regarding regression networks, our hybrid model performs similarly to LSMT methods with an average MSE of the order of 10−3 in accelerations estimated at TD.

3) **Analysis of the performance of classifiers and regressors.** With the final goal of developing a cockpit deployable recommendation system we have conducted a study of the performance of classification and regression models in terms of the flight height and different aircraft variables including the impact of automation and pilot manoeuvres. Results on our large dataset of commercial flights, show that although our regression networks performs similarly to SoA methods (with MSE of 10−3 in estimations at TD), the accuracy for detecting HL is very poor (46% of sensitivity). This indicates that regression models might not be the mostappropriate for the detection of HL events in a cockpit deployable support system.

4) **Sources of errors and capability for go-around recommendation.** Unlike previous approaches, we analyse the capability of networks for the detection of HL before the decision height, as well as, the influence of the operational context. We have also performed an analysis of the sources of errors, including selection of the best variable type, optimal altitude range used for predictions, biases due to aircraft type and capability of regressors for HL prediction.

The paper is organized as follows. Section 2 describes the methodology, including the description of variables, analysis of automation factors and network models. Section 3 reports the experiments conducted to assess the performance of models and error analysis. Section 4 discusses the results obtained and compares them to existing methods.

II. METHODS

A. DATASET DESCRIPTION

The authors have access to a large database of Flight Monitoring System (FMS) recorded data of an airline no longer in operation. This database has the following information:

- Fleet: A319/A320/A321.
- Various airports.
- 377,446 flights.
- 370 parameters available at various sampling frequencies.

Several primary criteria were defined to limit the data to what is considered meaningful for the hard landing predictions and the evaluation of the 3 hypothesis posed in this paper:

- All (A319/A320/A321).
- LHR - Heathrow Airport.

- Start of data: Final Approach Fix (FAF).
- End of data: 20 seconds after touch down.
- 58 parameters selected.

Heathrow airport was chosen as the sole airport to ease flight comparison and training of ML. Moreover, aircraft landing at Heathrow must follow a straight corridor further easing the landing comparison. This drops the number of available flights to 178,654. The data retrieved from the FMS starts at the FAF defined as 3 minutes before touching down and ends 20 seconds after touching down to capture the maximum G, labelled maxG, at touch down. A binary variable, labelled Wheel_on_Ground, was added to indicate the time of touch down when set to 1. Then, maxG was computed as the maximum value of Normal_acc_g in a window of +/- 5 seconds around Touch Down (TD) time as the maximum time Wheel_on_Ground equals1.

Parameters linked to characterizing unstable approaches are selected for the study. These parameters are linked to the aircraft dynamics (e.g. accelerations, rates, angle of attack), the position relative to the runway (glideslope and localizer), the aircraft configuration (landing gear state, control surfaces position) and the cockpit activity with the stick and throttle inputs. This reduces the number of raw parameters from 370 to 58. Additionally, dropouts and a significant amount of noise and data quantisation were identified. The poor data quality led to a reduction in the number of flights to approximately 58,177. Flights with maxG higher than the Mean plus 2x Standard Deviation of the normal acceleration at TD are classified as HL. This defines the threshold at 1.4037g and 2673 flights are flagged as HL. This represents approximately 4,6% of the total number of flights, which is consistent with the numbersreported [26].

The selected dataset allows to validate the 3 hypothesis posed in this paper. The temporal window always includes the decision height in order to validate to what extend the analysis of the aircraft dynamic state variables is enough for a go-around recommendation. The inclusion of the 3 types of aircraft allows to evaluate if HL precursors are particular to aircraft types, which is the second hypothesis of the paper. Finally, in order to validate the impact of environmental conditions (third hypothesis) data did not included the weather measurements rather its impact on the aircraft parameter features. The selected parameters were recorded at sampling frequencies between 0.25 and 8 Hz. However, since pilots make decisions according to altitude, we resampled all numerical variables as a function of altitude. To do such a change of variables, we used a linear interpolation of the values sampled at the frequencies to obtain values sampled at a uniform sampling of altitudes.

The final set of selecte0.82 cmd parameters were split into four different categories:

1) **actuators**, linked to actuators states, 2) **pilot**, related to pilot activity in the cockpit, 3) **physical**, as those parameters related to physical magnitudes as well as other factors such as 4)**automation factors**, as those binary parameters indicate whether an automatic system or guidance is engaged.

B. IMPACT OF AUTOMATION FACTORS IN HL

In order to explore the impact of automation in HL, the correlation between maxG and the following pilot decision making variables: autopilot, flight director, speed break, landing gear, and autothrust is evaluated. Autopilot, autothrust and flight director are computed as the last time/ altitude they are engaged. Landing gear and speed break are computed as the time/altitude they are first engaged. To better explore the impact of the above factors in HL, thedata has been split into hard landing (labelled HL) and non-hard landing (labelled NHL) eventsto detect any bias in the factors associated with HL.

Figure 1 shows the boxplots for the factors grouped according to their label. Notice that there are no significant differences between the values obtained in HL and NHL. Therefore automation factors do not seem to have an impact on the maxG and do not favour HL. Consequently, they will not be included in prediction models.

C. HL PREDICTION MODELS

A hard landing (HL) is defined as an event where vertical (or normal) acceleration exceeds a threshold value specific to the airplane type during the landing phase. A threshold on such normal acceleration (Airbus uses vertical acceleration > 2G at touch down, TD) triggers maintenance requirement and, thus, can be considered as a criterion for HL detection. Under this criterion, a Machine Learning System (ML) for HL prediction could be a classifier to discriminate flights with normal acceleration at TD above a given threshold from other flights. However, the values of the normal acceleration at TD follow a continuous unimodal probabilistic distribution. This fact also suggests using a regressor to predict the normal acceleration at TD and use either its value or a threshold on it as the HL predictor. In this work, we have considered both approaches:

- **Regressors.** The dependent variable to be predicted is the maximum normal acceleration (labelled maxG) at TD. This variable is computed as the maximum value ofNormal_acc_g in a window of ±5 seconds around TD time set as the maximum time Wheel_on_Ground = 1.

- **Classifiers.** We have considered a binary problem to classify hard landing (labelled HL) from non-hardlanding (labelled NHL). In our dataset flights with maxG > 1.4037 at TD are classified as HL.

For all ML methods (both regressors and classifiers) the input features are the concatenation of the variability of the continuous variables described in subsection III-A at a discrete set of flight altitudes which include the decision height, DH. The discrete sampling altitudes are [1500,1000,500,400,300,200,150,100,50,40,30] and the decision height was set to 100 feet. The lower altitude of 30 feet was selected as the limitpoint the pilot can safely avoid a HL event.

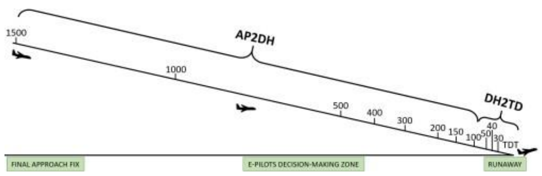

Fig. Altitude Sampling

III. EXPERIMENTS

A. EXPERIMENTAL DESIGN

The performance of the different approaches fordetection of HL events was assessed using sensitivity and specificity measures, which are common metrics in classification assessment. The sensitivity measures the capability of the system to detect HL events, while the specificitymeasures the capability for detection NHL.

Let us note TP the number of true positives (i.e. HL correctly detected by the system), FP, the number of false positives (NHL detected as HL by the IA system), TN the number of true negatives (NHL detected by the system) and FNthe number of false negatives (HL missed by thesystem), then sensitivity and specificity are given by equations in (3) and (4).

$$\text{Sensitivity} = TP / TP + FN \quad (3)$$

$$\text{Specificity} = TN / TN + FP \quad (4)$$

The following experiments have been conducted:

1) **Predictive Power of Models.** Optimal architectures were chosen as the ones that achieved better quality scores (sensitivity, specificity for classifiers and MSE for regressors) in training. The optimal regression neural network is compared to the optimal classification nets in terms of sensitivity, specificity in testing.

2) **Cockpit Deployable Potential.** In order to assess to what extent models can be effectively deployed in the cockpit, we have analyzed their performance according to the categorization of variables to determine the minimum set of variables and according to the altitude ranges to assess their capability for early detection of HL and for recommending a go-around.

B. RESULTS

Cockpit Deployable Potential

The analysis indicates that the performance of models (both, classifiers and regressors) depends on the type of aircraft variable used to trainmodels.

In the case of the regressor detected significant differences between the range AP2DH and the ones that used data until TD. This indicates that regressors might only accurately predict maxG ifdata close to TD is taken into account. This together with their poor performance for actuallydetecting HL events, discards regression modelsas the approach to use in a cockpit deployable system for early detection of HL.

IV. DISCUSSION IMPROVEMENTS

The proposed models only use the selected parameters on the final approach (below 2000 ft above ground). Wind direction and amplitude, the level of turbulence and the risk of gusts can have a significant impact on the possibility of an unstable approach resulting into a hard landing. The current analysis capture some of these effects as the variability the aircraft physical parameters are directly linked to the aforementioned Accuracy results of different models weather conditions.

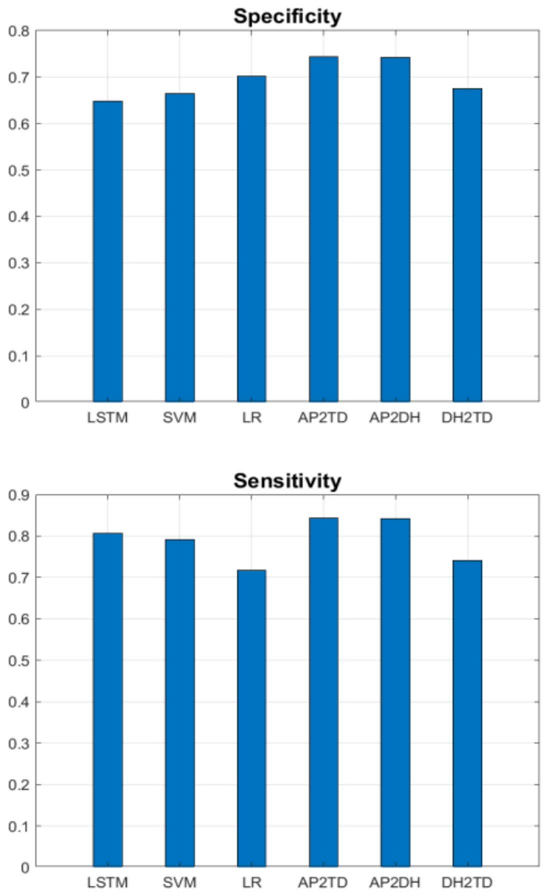

Fig. Accuracy results of different models

Low visibility conditions, such as fog, and icingconditions can also impair the quality of thelanding thus increasing the risk of hard landings but they are not considered in the study. The distance of flight has no specific influence on the results but the quality of the approach before reaching the final straight will have a significant impact. The aircraft might not be correctly configured or still have a significant level of energy to dissipate. This can be the result of ATC commands such as delayed descent instructions.

Another issue to be considered is that models did not include some key parameters that could have an impact in predictions. First, aircraft weight (mass) was missing in the study because the aircraft mass dataset from the flight databaseis unreliable. However, aircraft weight has several potential impacts:

- The potential energy needed to be dissipated to land is directly proportional to aircraft weight. In other words, the heavier the aircraft, the more energy needs to be dissipated to land at an acceptable speed and descent rate.
- The aircraft weight will have an impact on the aircraft dynamics. The aircraft inertias are directly linked to the aircraft weight. The aircraft centre of gravity is another parameter impacting aircraft stability and controllability and thus takeoff and landing performance. That is the reason why an aircraft will have load planner as well as in-flight centre of gravity target system with a trim tank to maintain it within operational ranges. The more forward the centre of gravity is, the higher the minimum speed, the higher thelanding speed.

Secondly, the aircraft centre of gravity, or the point at which the total weight of the aircraft is centred, is also a key parameter in the vehicle stability and control margins calculation which can greatly impact aircraft take-off and landing performance. Although commercial aircraft have trim-tank with centre of gravity target systems, the centre of gravity can vary within a range of certified positions within its airworthiness requirements. Therefore, including the centre of gravity and mass within the model could substantially improve the accuracy of the hard landing prediction.

The machine learning approach can also be improved in several aspects. Although results appear superior to existing methods, our modelswould benefit from a more complex analysis of temporal dependencies using a convolutional neural network to extract deep dependencies. The impact in predictions of meteorological conditions affecting visibility or aircraft aerodynamics should also be investigated to assess the benefits of their incorporation into our models. Given that the combination of all categories by straight concatenation of features does not significantly improve the performance of models trained with any single category, alternative architectures for their combination should be further investigated. Finally, the percentage of HL due to condition changes at TD should be determined to properly assess the capability of systems for early prediction of HL.

Finally, for a cockpit-deployable machine learning system to support flight crew go-around decision, some results regarding the hardware and software requirements, especially for the speed of networks should be investigated. The deployment of fully connected networks is already available even for low resource microcontrollers [27] and the latency in such cases [28], and with similar models as ours, is below 50 ms o 1 s, which are our main samplingrates. Hence, deployment software and latency are not considered as strong impediments for the future deployment in a cockpit.

V. CONCLUSION

The following conclusions can be extracted from the analysis carried out in this paper.

The analysis of automation factors (autopilot, flight director and auto-thrust) suggests that these factors do not have any influence on the probability of a HL event and, thus, it might notbe necessary to incorporate them into models.

Experiments for the optimization of architectures show that the configurations that achieve higher sensitivity are the ones with the lowest number of neurons. As reported in the literature [24] increasing the number of layers and neurons does not improve the performance of neither classifiers nor regressors.

Models using only Physical variables achieve an average recall of 94% with a specificity of 86% and outperform stateof-the-art LSTM methods.

This brings confidence into the model for early prediction of HL in a cockpit deployable system. Regarding capability for go-around recommendation before DH, even if we perform better than existing methods, there is a significant drop in recall and specificity due to the dynamic nature of a landing approach and factors influencing HL close to TD.

Comparing classifiers and regression approaches, experiments show that a low MSE error in estimation of maxG does not guarantee accurate HL predictions. Experiments for assessing the capability of models for early detection of HL show that classifiers are able to accurately predict HL before DH. The studysuggests that classifiers are a better approach for early prediction of hard landing.

Finally, there are some issues that have not beencovered in this work, that remain as future work,and should be further developed. Among such cases, stand out the robustness of the classifier (regressor) to unseen cases and its behaviour under a drifting data environment. In a safety demanding environment as aviation, it surely be needed to investigate such issues and we expectto do in further works. In the future, such a system could be expanded to also include Air Traffic Management in which the information is shared with the Air Traffic Controller in order to anticipate the likely scenario and optimize runway use.

ACKNOWLEDGEMENT

We are appreciative to our division of Computer Science Engineering for their help and giving usa chance to make things simpler. While looking about different significant and fascinating realities. The tools and the web services providedare of much help.

REFERENCES

[1] Boeing Commercial Airplanes. Statistical summary of commercial jet airplane accidents– worldwide operations| 1959–2017. Aviation Saf., Seattle, WA, USA, 2018

[2] European Aviation Safety Agency. Developing standardised fdm-based indicators. Technical report, European Aviation Safety Agency, 2016.

[3] Federal Aviation Administration. Advisory circular ac no: 91-79a mitigating the risks of a runway overrun upon landing. Technical report, Federal Aviation Administration, 2016.

[4] Michael Coker and Lead Safety Pilot. Why and when to perform a go around maneuver. Boeing Edge, 2014:5–11, 2014.

[5] Tzvetomir Blajev and W Curtis. Go-around decision making and execution project: Finalreport to flight safety foundation. Flight Safety Foundation, March, 2017.

[6] Eurocontrol. European action plan for the prevention of runway excursions. Report, Eurocontrol, 2013.

[7] European Union Aviation Safety Agency. Artificial intelligence roadmap - a human- centric approach to ai in aviation. Technical report, European Union Aviation Safety Agency, 2020.

[8] European Union Aviation Safety Agency. The european plan for aviation safety (epas 2020-2024). Technical report, European Union Aviation Safety Agency, 2019.

BIOGRAPHIES

Team Guide,

Ch. Anuradha ,Assistant Professor,

Computer Science Engineering in

V.S.M College of Engineering,

Ramachandrapuram.

Team Leader,

S.N.Uday Kiran ,Computer Science Engineering Student in V.S.M College of Engineering, Ramachandrapuram.

Team member,

M. Lakshman ,Computer Science Engineering Student in V.S.M College of Engineering, Ramachandrapuram.

Team member,

M. Poojitha ,Computer Science Engineering Student in V.S.M College of Engineering, Ramachandrapuram.

Team member,

M. Gayathri ,Computer Science Engineering Student in V.S.M College of Engineering, Ramachandrapuram.

Team member,

Ch. Chandini, Computer Science Engineering Student in V.S.M College of Engineering, Ramachandrapuram.