Hard Landing Prediction System for Commercial Flights using Machine Learning

Chinmayi Sangaraju,
UG Student,
Department of CSE,
St. Martin's Engineering College,
Secunderabad, Telangana, India.
sangarajuchinmayi@gmail.com

Dr. B. Madhava Rao,
Associate Professor,
Department of CSE,
St. Martin's Engineering College,
Secunderabad, Telangana, India.
madhavaraocse@smec.ac.in

Abstract:

More than half of all commercial aircraft operation accidents could have been prevented by executing a go-around. Making a timely decision to execute a go-around manoeuvre can potentially reduce overall aviation industry accident rate. In this project, we describe a cockpit- deployable machine learning system to support flight crew 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 as inputs 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.

Keywords: Hard Landing, go-around point, cockpit-deployable model, and manoeuvre.

1.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 operation met the criteria of an unstable approach, 97% of them continued to be landing rather than executing a go-around [4]. A study conducted by Blajev and Curtis [5] found that 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 decisions can be made by either flight crew members and can be executed at any point from the final approach fix point to wheels touching down the runway (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 real time 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 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 be predicted, this project presents a cockpit deployable machine learning system to predict hard landings considering the aircraft dynamics and configuration. This project evaluates three main hypotheses. 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 automation factors).,noise, and loss of detail..

2. LITERATURE SURVEY

Although there is a lot of work addressing the prediction of flight landing incidents [9]–[12] and other unsafety situations [13]–[16], the prediction of hard landing accidents have been less researched. Furthermore, most of the existing works focus on the prediction of HL for unmanned aerial vehicles (UAV), which dynamical features and flying protocols are completely different from the ones of commercial flights. A Hard Landing (HL) is a phenomenon in which the airplane has an excessive impact on the ground now of landing. This impact is directly related to the vertical (or normal) acceleration; therefore, HL can be defined as flights where the vertical acceleration exceeds the limited value of the aircraft type during the landing phase. A threshold on such normal acceleration (Airbus uses vertical acceleration> 2G at Touch Down, TD) triggers maintenance requirement, so that can be considered as a criterion for HL detection. Under the former definition of HL, existing approaches for HL prediction can be split into two groups: those based on a classifier to discriminate flights with normal acceleration at TD above a given threshold from other flights and those based on a regressor that predicts the normal acceleration with the aim of using this predicted value as the HL detector.

Classifiers can be categorized into machine learning and deep learning approaches. Machine learning methods [17] [19] apply a classifier to UAV flight data recorder using the Quick Access Recorder (QAR) sampled at a discrete set of heights that define the feature space. Most methods [17], [19] use the values of variables describing aircraft dynamics sampled between 9 and 2 meters before TD. Others, like [18], use statistical descriptors (panel data) of such variables also sampled at the very last meters before TD. On one hand, it is not clear what is the capability of these approaches to capture time-sequence dependencies that variables might have across the approach phase. On the other hand, the temporal window (9-2 meters before landing) used for predictions in UAV flights might not be appropriate for HL predictions in commercial flights. The approximate limit altitude (known as Decision Height -DH-) in commercial flights to decide a go around is 100 feet (38 meters). Thus, regardless of their accuracy in predicting HL, these ML methods are not applicable for commercial flights due to the altitude range required.

Deep learning approaches are mainly based on Long Short-Term Memory Recurrent Neural Network (LSTM) architectures. Proposed by [20], these networks are a variant of Recurrent Neural Networks (RNN) [21] able to model long term dependencies within temporal data. In particular, the very recent work in [22] used the signals of 3 kinds of landing related features (aircraft dynamics, atmospheric environment, and pilot operations) as inputs to a LSTM network predicting HL. Their comparison to classic machine learning approaches in terms of precision and recall of HL events of A320 flights indicates a potentially higher performance in terms of HL recall with 70% of HL detection while keeping with a percentage (76%) of precision like the one obtained by classic machine learning approaches. Despite the promising results, we consider that the experimental design of [22] lacks in some aspects for properly assessing the potential for deployment in the cockpit.

First, the test set used is balanced with almost the same number of HL and non-HL cases. However, in a real situation, HL cases are rare events that represent only 3-4% of flights [23]. By balancing the test set, precision might be too optimistic and even unrealistic. To guarantee a useful decision support system, the number of false alarms should be properly estimated. Second, the authors conducted an analysis that showed that the optimal temporal window for doing predictions was between 10 and 2 seconds before landing. This temporal window corresponds to heights between 164 and 13 feet, which are below the decision height (100 feet) of commercial flights. Finally, the data only include a single aircraft type (A320). Given that aircraft aerodynamics are strongly related to aircraft design, the generalization of the approach remains unknown. Regression approaches predicting normal acceleration are also mostly based on deep learning LSTM strategies.

This might be limiting the capability of the system for fully exploring time dependencies and might discard discriminative features. Although both works obtain accurate predictions with an average Mean Squared Error (MSE) of the order of 10^{-3}, LSTM is not trained to predict the vertical acceleration at TD at the next time interval after the current observation. In fact, a recurrent network can only predict acceleration at the immediate time interval from the current observation and its capability for long term predictions is not clear. Since HL depends on the values of such vertical acceleration in a tight temporal window at the time of TD, this limits the deploy ability of system in a cockpit.

3. PROPOSED METHODOLOGY

This proposed methodology aims to develop a system called E-Pilots that uses machine learning algorithms to predict hard landings during the approach phase of commercial flights. The system will analyze flight data to precede hard landings. The goal is to provide pilots with real-time warnings and guidance to prevent accidents and improve safety. The research methodology includes the collection and analysis of flight data, the development and testing of machine learning algorithms, and the integration of the E-Pilots system with existing flight systems. The findings of this paper are expected to contribute to the improvement of aviation safety and reduce the occurrence of hard landings. The implications of this research may also extend to other areas of aviation safety and flight automation.

The system would use machine learning algorithms, such as Naive Bayes, Logistic Regression, SVM, Decision Tree, SGD Classifier to analyze the flight data and identify patterns that precede hard landings. These patterns might include changes in altitude, airspeed, or rate of descent. Once the system detects these patterns, it would provide real-time warnings and guidance to pilots to prevent hard landings advantages are: Enhanced safety, Accurate Predictions, Cost Savings, and Continuous Improvement

Advantages:

- Enhanced Safety
- Accurate Predictions
- Cost Saving
- Continuous Improvements

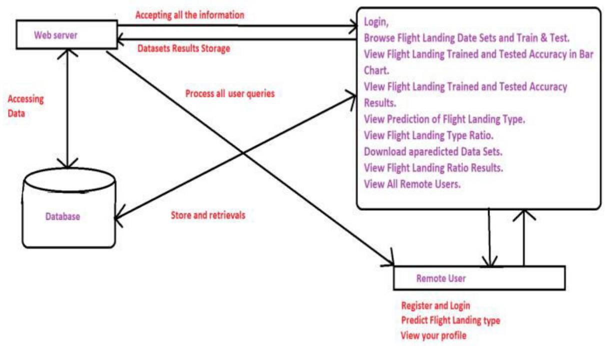


Figure 1: Overall Design of proposed system

Dataset Description:

We 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.

2

- 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 the3 hypothesis posed in this project:

- 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 quantization was 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 numbers reported [26].

HL Prediction Methods:

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 predictioncould 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 nonhardlanding (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 continuousvariables 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.

4. EXPERIMENTAL ANALYSIS

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). Sensitivity = T P /T P + F N (3)

Specificity = T N /T N + F P (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.

In Figure 2, a Graphical User Interface is created using python and the dataset is divided into categories and uploaded. In Figure 3, user uploads the details and registers and Figure 4, gives the details of overall viewers or users. In Figure 5, we give the details of the flight and it's landing to predict the type of flight landing (hard or soft landing). Figure 6, gives the accuracy with which the various types of algorithms predict the type of landing in percentages and also in bar and pie charts.

Figure 2: GUI Homepage

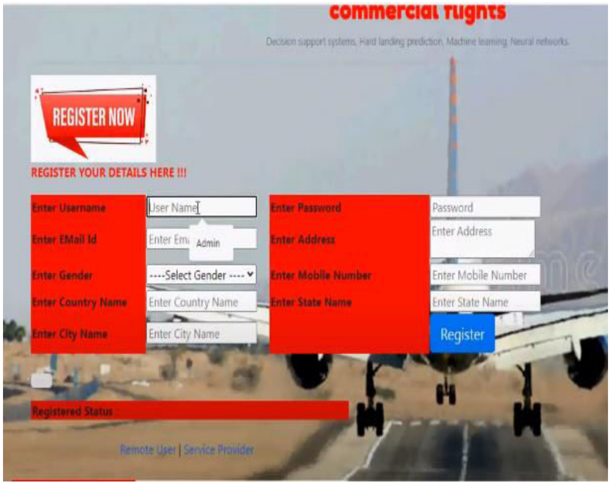
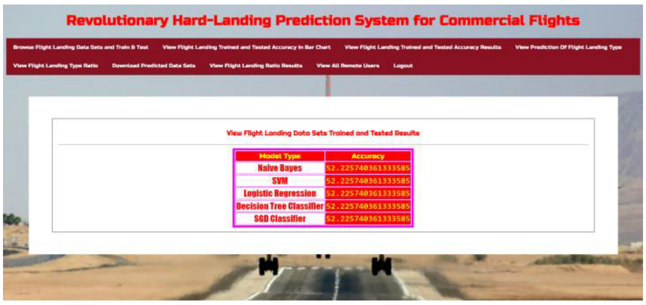


Figure 3: User Details

Figure 4: Overall viewers

Figure 5: Prediction details



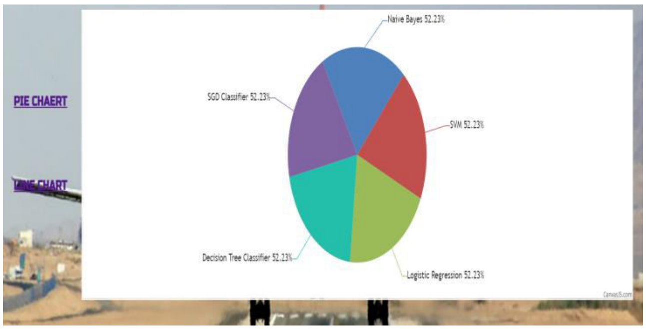


Figure 6: Accuracy Details

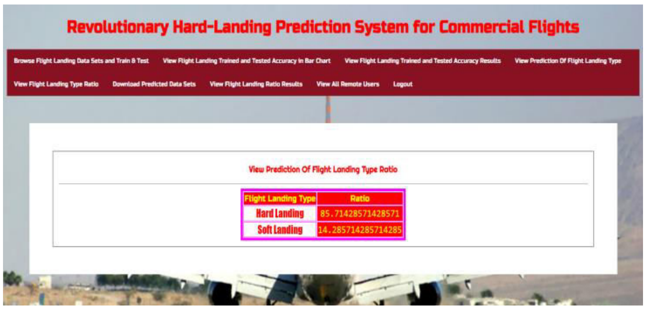


Figure 7: Results

5. CONCLUSION

The analysis of automation factors (autopilot, flight direct-tor and auto-thrust) suggests that these factors do not have any influence on the probability of a HL event and, thus, it might not be 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 increasing the number of layers and neurons does not improve the performance of either classifiers nor regressors. Models using only Physical variables achieve an average recall of 94% with a specificity of 86% and outperform state- of- 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.

4

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 can accurately predict HL before DH. This is not the case of regressors, which predict maxG more accurately if data close to TD is considered into the model. The study suggests that classifiers are a better approach for early prediction of hard landing. Neural networks performance could be increased if they were used to extract deep learning features from continuous signals by using one dimensional convolutional networks and different architectures for a better combination of the three categories of variables.Also, models should incorporate additional parameters such as aircraft mass and center of gravity position which are known to impact vehicle dynamics.

Finally, there are some issues that have not been covered 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 behavior under a drifting data environment. In a safety demanding environment as aviation, it surely be needed to investigate such issues and we expect to 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 to anticipate the likely scenario and optimize runway use.

REFERENCES

[1] Guo, Chunle, Cho, J., Choi, Y. J., Sohn, J., Suh, M., Cho, S. K., Ha, K. H., ... & Shin, D. C. (2015). Ambient ozone concentration and emergency department visits for panic attacks. Journal of Psychiatric Research, 62, 130-135.

[2] Power, M. C., Kioumourtzoglou, M. A., Hart, J. E., Okereke, O. I., Laden, F., & Weisskopf, M. G. (2015). The relation between past exposure to fine particulate air pollution and prevalent anxiety: observational cohort study. bmj, 350.

[3] Rotko, T., Oglesby, L., Künzli, N., Carrer, P., Nieuwenhuijsen, M. J., & Jantunen, M. (2002). Determinants of perceived air pollution annoyance and association between annoyance scores and air pollution (PM2. 5, NO2) concentrations in the European EXPOLIS study. Atmospheric Environment, 36(29), 4593-4602.

[4] Llop, S., Ballester, F., Estarlich, M., Esplugues, A., Fernández-Patier, R., Ramón, R., ... & Iñiguez, C. (2008). Ambient air pollution and annoyance responses from pregnant women. Atmospheric Environment, 42(13), 2982-2992.

[5] Brook, R. D., Rajagopalan, S., Pope III, C. A., Brook, J. R., Bhatnagar, A., Diez-Roux, A. V., ... & Kaufman, J. D. (2010). Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. Circulation, 121(21), 2331-2378.

[6] Cho, J., Choi, Y. J., Sohn, J., Suh, M., Cho, S. K., Ha, K. H., & Shin, D. C. (2015). Ambient ozone concentration and emergency department visits for panic attacks. Journal of Psychiatric Research, 62, 130-135.

[7] Sass, V., Kravitz-Wirtz, N., Karceski, S. M., Hajat, A., Crowder, K., & Takeuchi, D. (2017). The effects of air pollution on individual psychological distress. Health & place, 48, 72-79.

[8] Rotko, T., Oglesby, L., Künzli, N., Carrer, P., Nieuwenhuijsen, M. J., & Jantunen, M. (2002). Determinants of perceived air pollution annoyance and association between annoyance scores and air pollution (PM2. 5, NO2) concentrations in the European EXPOLIS study. Atmospheric Environment, 36(29), 4593-4602.

[9] Xu, W., Ding, X., Zhuang, Y., Yuan, G., An, Y., Shi, Z., & Hwa Goh, P. (2020). Perceived haze, stress, and negative emotions: An ecological momentary assessment study of the affective responses to haze. Journal of health psychology, 25(4), 450-458.

[10] Lu, J. G., Lee, J. J., Gino, F., & Galinsky, A. D. (2018). Polluted morality: Air pollution predicts criminal activity and unethical behavior. Psychological science, 29(3), 340-355.

[11] C. Haipeng, S. Ping, H. Shengguo,- “Study of aircraft hard landing diagnosis based on neural network [j], Computer Measurement & Control 7” (2008).

[12] L. Yi, S. Zhang, L. Xueqing- “A hazard analysis-based approach to improve the landing safety of a bwb remotely piloted vehicle, Chinese Journal of 325 Aeronautics 25 (6)” (2012).

[13] L. Witte, R. Roll, J. Biele, S. Ulamec, E. Jurado, Rosetta lander philae– “310 landing performance and touchdown safety assessment, Acta Astronautica 125” (2016).

[14] L. Yanhui, Z. Shuguang, G. Lei- “A safety strategy for high-speed uav land330 ing taxiing control, Procedia Engineering 17” (2011).

[15] T. Patterson, S. McClean, P. Morrow, G. Parr- “ Modelling safe landing zone detection options to assist in safety critical uav decision making, Procedia Computer Science“ (2012).