

# MACHINE LEARNING ENABLED TURBULENCE PREDICTION USING FLIGHT DATA FOR SAFETY ANALYSIS

Mariam Emara<sup>1</sup>, Marcos dos Santos<sup>1</sup>, Noah Chartier<sup>1</sup>, Jamey Ackley<sup>1</sup>, Tejas G. Puranik<sup>1</sup>, Alexia Payan<sup>1</sup>, Michelle Kirby<sup>1</sup>, Olivia J. Pinon<sup>1</sup> & Dimitri N. Mavris<sup>1</sup>

<sup>1</sup> Aerospace Systems Design Laboratory (ASDL), School of Aerospace Engineering, Georgia Institute of Technology, Atlanta, GA, 30332, U.S.A

## Abstract

The hazards posed by turbulence remain an important issue in commercial aviation safety analysis. Turbulence is among the leading cause of in-flight injury to passengers and flight attendants. Current methods of turbulence detection may suffer from sparse or inaccurate forecast data sets, low spatial and temporal resolution, and lack of in-situ reports. The increased availability of flight data records offers an opportunity to improve the state-of-the-art in turbulence detection. The Eddy Dissipation Rate (EDR) is consistently recognized as a reliable measure of turbulence and is widely used in the aviation industry. In this paper, both classification and regression supervised machine learning models are used in conjunction with flight operations quality assurance (FOQA) data collected from 6,000 routine flights to estimate the EDR (and thereby turbulence severity) in future time horizons. Data from routine airline operations that encountered different levels of turbulence is collected and analyzed for this purpose. Results indicate that the models are able to perform reasonably well in predicting the EDR and turbulence severity around 10 seconds prior to encountering a turbulence event. Continuous deployment of the model enables obtaining a near-continuous prediction of possible future turbulence events and builds the capability towards an early warning system for pilots and flight attendants.

**Keywords:** Safety, Turbulence, Machine Learning, Predictive Modeling, Risk

## 1. Introduction

Turbulence in flight is among the leading cause of injuries in non-fatal airline accidents. Yet, for unbelted passengers and flight attendants, the injuries could be fatal. For the commercial airline industry, encounters with turbulence cost in the millions of dollars each year in insurance premiums, workers compensation, and injury settlements [1].

Turbulence or eddies that affect aircraft are created by various larger-scale atmospheric forcing mechanisms, and the resulting turbulence is often classified according to its source [2]. At cruising altitudes of commercial aircraft, there are three common sources of turbulence: 1) Convective-Induced Turbulence (CIT) which is turbulence associated with the presence of convective clouds (either in-cloud or near-cloud) ; 2) Clear-Air Turbulence (CAT) which is turbulence associated with enhanced wind shears and reduced stability in the vicinity of jet streams, the tropopause, and upper-level fronts; it often occurs in clear air or sometimes in stratiform clouds but not in or near convective clouds; and 3) Mountain Wave Turbulence (MWT) which is turbulence associated with the breaking of mountain waves above mountainous terrain and which also often occurs in clear air.

Current turbulence detection/prediction algorithms are primarily physical in nature in that they are based on theories of how the causes of turbulence are represented in certain data sets (such as Doppler radar data) [3, 4]. Current methods of turbulence detection may suffer from sparse or inaccurate forecast data sets, low spatial and temporal resolution, lack of in-situ reports, etc. On the other hand, modern machine learning algorithms have become highly efficient at detecting patterns in data and using these patterns to make predictions. There is a wealth of data typically collected on-board routine flights that can be used for a variety of purposes. Thus, there is an opportunity to utilize in-flight data to improve the state of the art in turbulence detection. There have been previous attempts to identify EDR using quick access recorder (QAR) data recorded from flights [5].

While these efforts have focused on EDR estimation, this paper takes the work a few steps further by training a prediction model around EDR and using it to estimate turbulence severity. Considering the preceding observations, in this work, a methodology for building a model to predict the EDR and turbulence level at a future time horizon using data collected on-board is presented. This prediction can serve as an early warning system for pilots and flight attendants and help prevent injuries. The main objectives and contributions of this research are as follows:

1. Provides a novel predictive model of aircraft's future eddy dissipation rate and turbulence severity using data collected on-board an aircraft
2. Demonstrates the implementation of a supervised learning risk identification methodology for turbulence using large-scale flight data
3. Identifies critical parameters that potentially contribute towards prediction of moderate and severe turbulence

While this paper focuses on early prediction and detection of turbulence for alerting flight crews, the methods developed are also envisioned to eventually assist air traffic management (ATM) personnel in better managing airspace [6, 7].

The rest of the paper is organized as follows: Section 2 covers some of the background and prior work related to data-driven analysis in aviation in general and turbulence in particular. Section 3 provides an outline of the methodology used in this work. Section 4 contains the results of the implementation of the framework. Finally, Section 5 provides concluding remarks and avenues for future work.

## 2. Background

In recent years, there has been significant interest in using flight data collected on-board for various safety analysis tasks. The number of machine learning techniques for solving complex problems and enabling predictive approaches in the transportation domain have increased in recent years. Flight data is typically used during retrospective analysis to identify anomalies during routine operations using various machine learning techniques [8, 9, 10, 11]. While many of the previous applications have focused on unsupervised learning, supervised learning has also been applied on flight data in recent years. Supervised learning has been used to predict future states of the aircraft [12, 13, 14], probability of hard landings [15], prediction of unstable approaches [16], etc. For a more comprehensive overview of various applications of flight data the readers are referred to recent review papers [17, 18]. The current work focuses on the application of supervised learning algorithms for turbulence prediction using recorded flight data.

A methodology for predicting turbulence in future time horizons based on recorded flight data is proposed in this work to address limitations observed in literature.

## 3. Methodology

The methodology and steps followed in this paper are outlined in Figure 1. The elements of the methodology follow those of a general machine learning pipeline. The main difference and innovation lies in the way flight data is processed to produce sliding window based features for the ML model to enable a prediction in the future based on collected data.

### 3.1 Flight Data

The first step consists of data collection from routine flights conducted by commercial airlines. This data is in the form of time-series measurements of hundreds of parameters at a 1 Hz frequency. Each flight in the data set can contain hundreds of rows that represent unique time stamps and columns that represent recorded parameters. The parameters recorded can be of various types such as continuous, discrete, boolean, text, etc. The parameters in the flight data can be divided into different categories and levels based on their source system/sub-system in the aircraft. Atmospheric data refers to data gathered from pitot tubes, barometers, thermometers, etc. It includes airspeed, wind speeds, pressure altitude, atmospheric temperature, etc. Attitude data refers to roll, pitch, yaw angles and their corresponding rates and accelerations. GPS data contains the latitude, longitude, altitude, and related rates. Engine data contains RPM, Exhaust Gas Temperatures (EGT), Cylinder

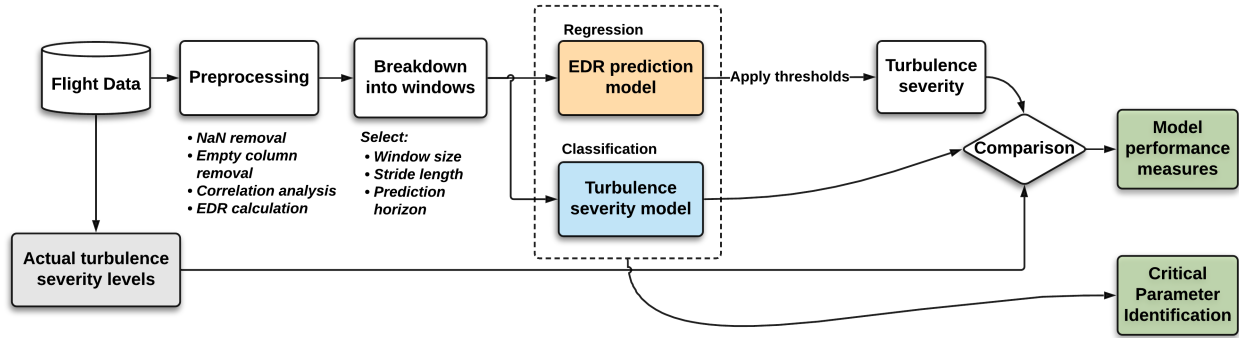


Figure 1 – Flow of data and steps in the prediction model generation framework

Head Temperatures (CHT), oil temperature and pressure, fuel flow rates, fuel quantities, etc. Control data contains the deflection of flaps, elevator, aileron, rudder, etc. Communications data includes details about the communication status of the vehicle, such as the common frequency. Navigation data includes information on any way-point guidance or autopilot features. These are among the numerous categories of parameters typically present in flight data. For the purpose of this paper, the segments of flight that contained turbulence are isolated, and the data from two minutes before and after the recorded event is obtained. The data consists of multiple airframe types and phases of flight. Figure 2 shows a notional hierarchical breakdown of the flight data used in this work.

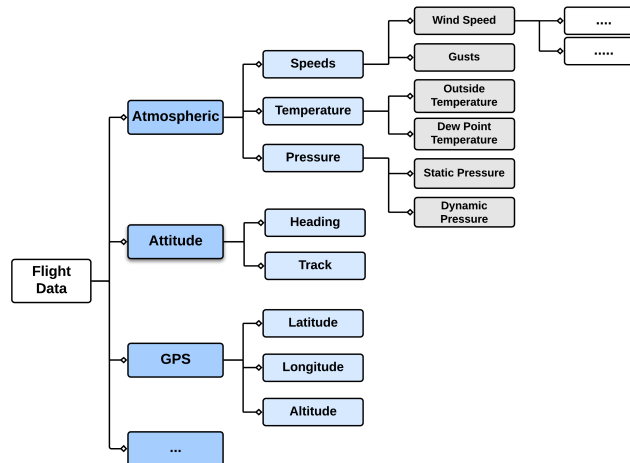


Figure 2 – Notional hierarchical representation of the data

### 3.2 Preprocessing

The data is pre-processed to remove empty and corrupted columns. This results in the reduction of the total number of columns but is an essential step to ensure the machine learning models can be trained and used properly. Correlation analysis is then conducted to identify highly correlated columns and only retain the most representative columns for further analysis. The correlation analysis is typically able to remove redundant columns. The data is then broken down into windows of a specific length (e.g., 10 seconds). The windows are then slid across the duration of the flight under consideration to create a continuous source of data. The parameter values within each window are then flattened into a single long feature vector for machine learning analysis.

### 3.3 Turbulence Label Generation

For any supervised learning effort, the true value of the prediction, i.e., a continuous value for regression and a class label for classification, is required. For the turbulence prediction problem, Eddy Dissipation Rate (EDR) is used. The calculation for EDR stems from a combination of literature review and available commercial airline turbulence flights. In the literature, three different methods

are found for predicting EDR from Quick Access Recorder (QAR) data. QAR data consists of over 1000 parameters and is functionally similar to the available FOQA data. Then, the parameters used in each method are compared to available turbulence data in this work to determine feasibility of implementing one of the methods.

A 2019 paper by Huang [5] used 14 QAR properties and 7 overarching equations to develop a method to estimate EDR. A couple of concepts used in this methodology include estimating the acceleration response function and estimating the acceleration response energy. Upon comparison of the parameters needed, parameters available, and uncertainties between the two this method quickly becomes a complicated option to model in the current work.

The second paper by Haverdings [4], also from 2019, looks at predicting EDR using Air-to-Ground (ATG) technology. A comparison between the parameters used in the paper and parameters available for the current work reveals a lack of all the necessary information to implement this method.

The final paper analyzed by Chen [19], from 2010, uses a WINDGRAD algorithm to predict the EDR. The comparisons of parameters show that some of the parameters used are already pre-calculated in the available FOQA data. In turn, to implement this method, all that was needed was one equation and two parameters. These two parameters are vertical wind speed and the true airspeed. Due to the simplicity and having all the needed parameters, this method is implemented in this paper. Besides the two parameters in the dataset, the other symbols in the equation represent constants that were either determined through sensitivity analysis or recommended from the paper. Specifically, the cutoff frequencies were given in the paper as constants and the running standard deviation for vertical wind was implemented through multiple window sizes for sensitivity.

$$\varepsilon^{1/3} = \frac{\hat{\sigma}_w}{\sqrt{1.05 \times V_a^{2/3} (\omega_1^{-2/3} - \omega_2^{-2/3})}} \quad (1)$$

Where ( $\sigma_w$ ) is running standard deviation in vertical wind (either using 10 or 20-second window sizes, as suggested by the paper),  $V_a$  is the airspeed, and  $\omega_1$  and  $\omega_2$  are the cutoff frequencies for EDR calculation (0.15 Hz and 2 Hz are used, respectively, as suggested by the paper). Both window sizes to calculate EDR are used to produce the feature vectors and test sensitivity.

Using equation1, the model is developed in Python. Various window sizes for the standard deviation in the calculation are explored to identify the most suitable ones. Although the paper recommended a window size between 10 and 20, the team looked at a wide range of window sizes from 5 to 50 seconds. It was concluded that the smaller window sizes led to undesirable quantity of noise and a window size of 20 seconds was used. The severity of turbulence is categorized depending on the EDR. Through literature search, four different thresholds were found. Based on typical commercial airline and aviation weather center's EDR definitions, the following three labels are used: light turbulence (EDR = 0.00 - 0.15 m<sup>2/3</sup>/s), moderate (0.16 - 0.45 m<sup>2/3</sup>/s), and severe turbulence (> 0.46 m<sup>2/3</sup>/s).

### 3.4 Machine Learning Models

The feature vectors generated along with true turbulence labels (EDR value for regression and actual turbulence level for classification) are used to train the models. In this work, regression and classification models are both built using the Gradient Boosting algorithm [22]. Gradient Boosting classification<sup>1</sup> and Gradient Boosting regression<sup>2</sup> are respectively used for building the classification and regression models using publicly available implementations in Python. While there are several algorithms available for building the classification and regression models, the focus of this paper is not on the model type itself but rather on the application for turbulence prediction. Each model utilizes the same feature vectors that are generated by the sliding-window approach but differs in the output prediction (turbulence severity for classification and EDR value for regression). The output from the regression model can be converted to a turbulence severity using the thresholds identified earlier.

<sup>1</sup><https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingClassifier.html>

<sup>2</sup><https://scikit-learn.org/stable/modules/generated/sklearn.ensemble.GradientBoostingRegressor.html>

### 3.5 Post-processing

The trained, validated models are then compared against the actual data to determine the performance of the models and identify the best performing model. Additionally, parameters that are critical to the successful prediction of turbulence for both the regression and classification models are identified using the safety analysis for flight events (SAFE) methodology [21].

## 4. Implementation and Results

This section contains the details of the implementation of the developed method on a real-world set of approximately 6,000 flights that contained some level of turbulence.

### 4.1 Model Considerations

While building and improving the models, the three subsequently discussed approaches are considered and used to improve typical challenges faced while building the predictive models based on the nature of the problem.

#### 4.1.1 Stratified Sampling

Considering the high imbalance of the data (evidenced by the discrepant amount of light turbulence events as opposed to moderate and severe events), a stratified sampling approach is used. Simple random sampling randomly selects data from the entire population so that each possible sample is likely to occur. Stratified sampling, on the other hand, creates a test set with a population that best represents the studied population, and it is used to eliminate sample bias (i.e., when certain values are under-represented). The advantages of using this approach are that it accurately represents the population being studied and it ensures each subgroup is properly represented, while the disadvantages are that it should not be used when confidently classifying every member of the population is not possible [10].

#### 4.1.2 Hyperparameter Optimization

Model hyperparameters can be arbitrarily set by user before starting the training, and they determine how the model is structured. In running the hyperparameter optimization exercise, the goal is to find the right combination of parameters to find either the minimum (e.g., loss) or maximum (e.g., accuracy) of a function. A search over a space of two main hyperparameters (number of estimators and maximum depth) is conducted to find the optimum combination whose learning and model performance were optimum. In addition, a design of experiments was used to tune the rest of the model hyperparameters and reduce overfitting. The final values for hyperparameters used is shown in Table 1.

Parameter	Regression	Classification
Random state	42	42
Number of estimators	139	86
Maximum depth of estimators	7	9
Fraction of samples used for fitting (Subsample)	0.9	0.7
Minimum number of samples required to split internal node	5	5
Minimum number of samples required to be at a leaf node	3	3
Learning rate	0.1	0.05

Table 1 – Optimal hyperparameter values

#### 4.1.3 Error Types

In this paper, the confusion matrix is a central part of the evaluation criteria used to define the effectiveness of the models in predicting turbulence. This is because, rather than attempting to predict the exact value of eddy dissipation rates (EDR), the main goal is to predict the turbulence intensity level,



which can be divided into light, moderate, and severe. Each of these labels have a range of EDR values and therefore predicting the exact value of EDR may not be necessary.

Two error types are typically encountered in the models built for turbulence prediction. Type 1 error occurs, for instance, when severe turbulence is predicted when actual turbulence is moderate or light. Type 2 error happens when light or moderate turbulence is predicted when actual turbulence is severe. Type 2 error is more serious in this problem since it can give the false impression of safety when the risk is high.

## 4.2 Results

The model performance metrics discussed below are for the final models built for models built using regression and classification algorithms. The main performance metrics used are overall accuracy, f1 score, EDR or severity predictions on specific flights.

### 4.2.1 Regression

The regression models continuously predicts the values of EDR, and the severity levels are determined from the commercial airline EDR thresholds. In this section, the results for the best tuned regression model are discussed.

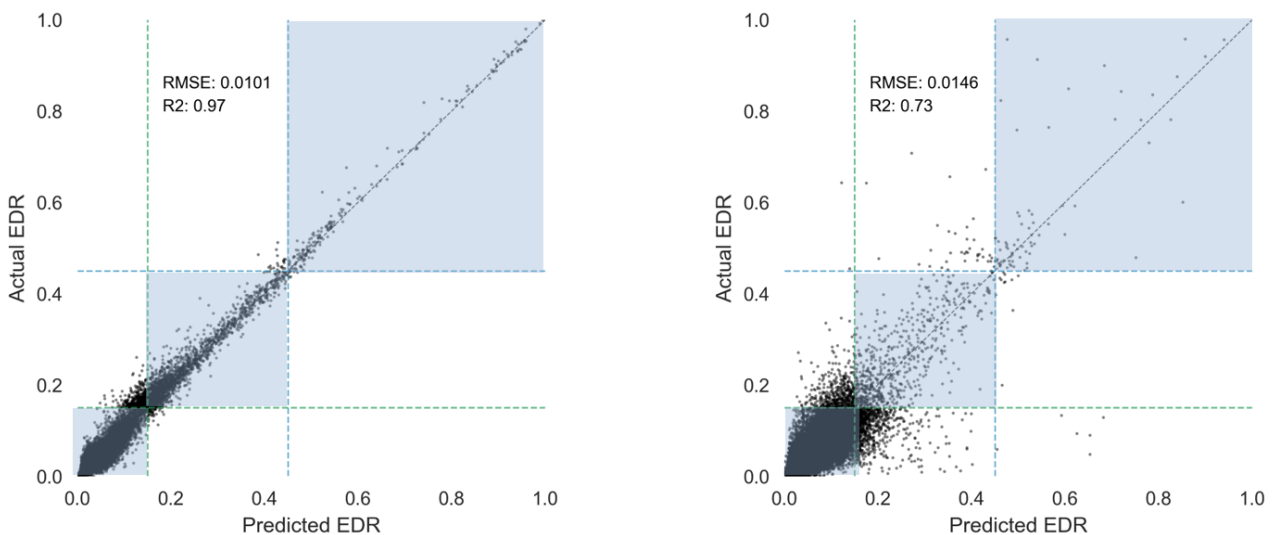


Figure 3 – Actual vs predicted for training data (left) and test data (right)

The actual vs. predicted plots (Figure 3) show that the error is slightly higher for the test data, but regression results in overall good fitting model. Highlighted regions which indicate correct severity level prediction contain most of the data which means regression creates a reliable EDR prediction model. Figure 4 shows the confusion matrix for the regression model after comparing the severity labels with the actual labels.

The confusion matrices show that the f1 score is higher for the training set, but both show good performance for the regression model. The extreme diagonals (light turbulence predicted as severe or the opposite) are either very low or zero. Finally, the model accurately predicts EDR throughout most of the sample flight shown in Figure 5, and almost all the turbulence severity levels are correctly predicted.

### 4.2.2 Classification

The classification models are used to directly predict the discrete severity level labels (light, moderate, and severe), as opposed to predicting EDR values over time as the regression model does.

The results discussed in this section use optimized hyperparameters for building the classification model. Compared to regression, the classification model shows higher accuracy and f1 score for the training set, but lower f1 score for the test set (approximately 7% lower) as seen from Figure 6. While

## TURBULENCE PREDICTIVE MODELING USING MACHINE LEARNING

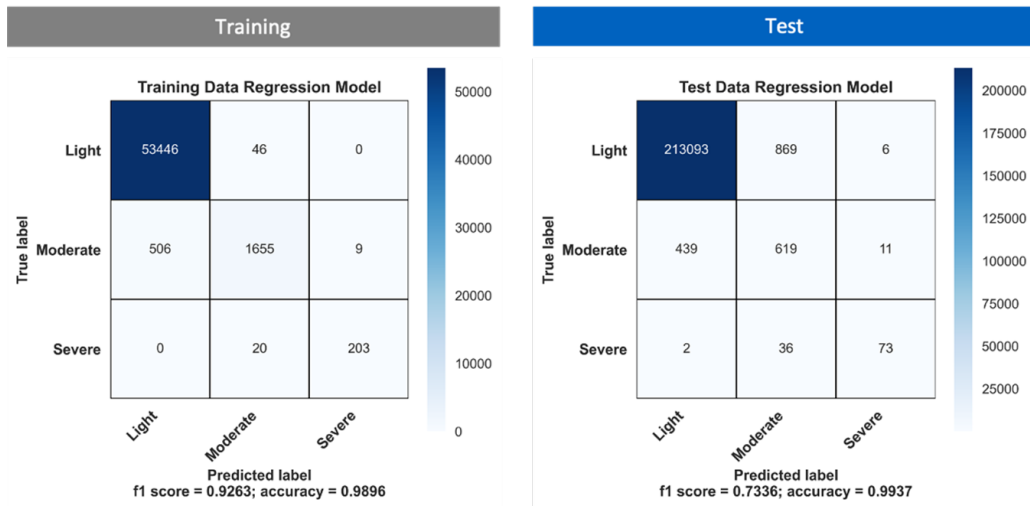


Figure 4 – Confusion matrix for the labels obtained based on regression model

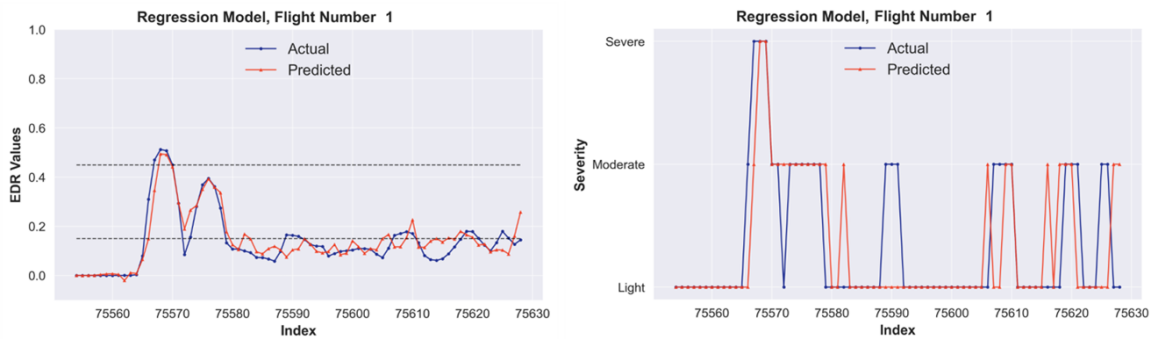


Figure 5 – Prediction of EDR and severity for a sample flight using regression model

the classifier is not able to capture EDR values that the regressor can predict, it is able to accurately predict turbulence severity levels and captures severe turbulence particularly well.

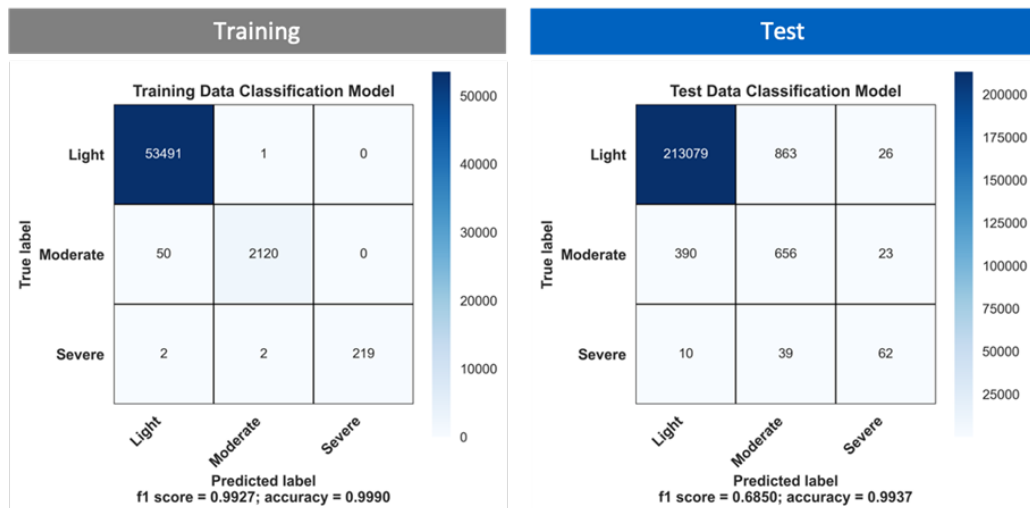


Figure 6 – Confusion matrix for the labels obtained based on classification model

Figure 7 shows the prediction of turbulence severity for the same flight for which regression model was demonstrated earlier. A similar type of performance is observed for this flight, with most of the severities predicted correctly by the model with occasional misclassifications.

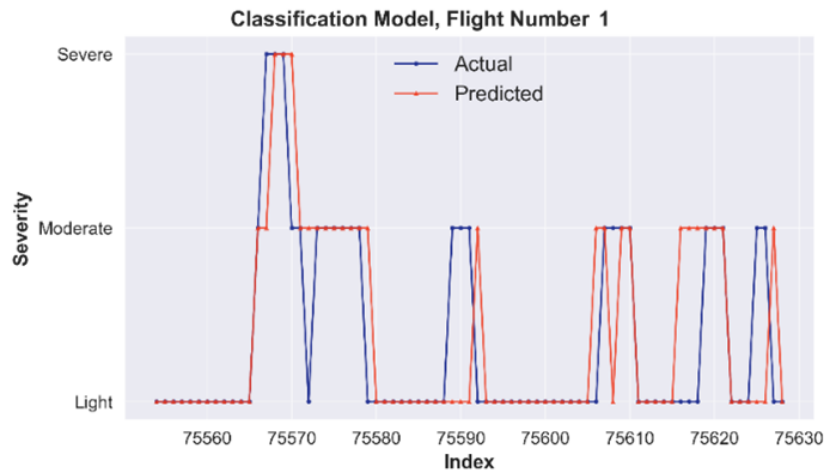


Figure 7 – Prediction of turbulence severity for a sample flight using classification model

### 4.3 Critical Parameters

The critical parameter identification is an important component of this research effort since they are an indication of the potential precursors that lead to initiating events during an encounter with turbulence. Effectively, they are a list of the most important parameters in predicting the turbulence severities, as found by the regression and classification supervised machine learning models.

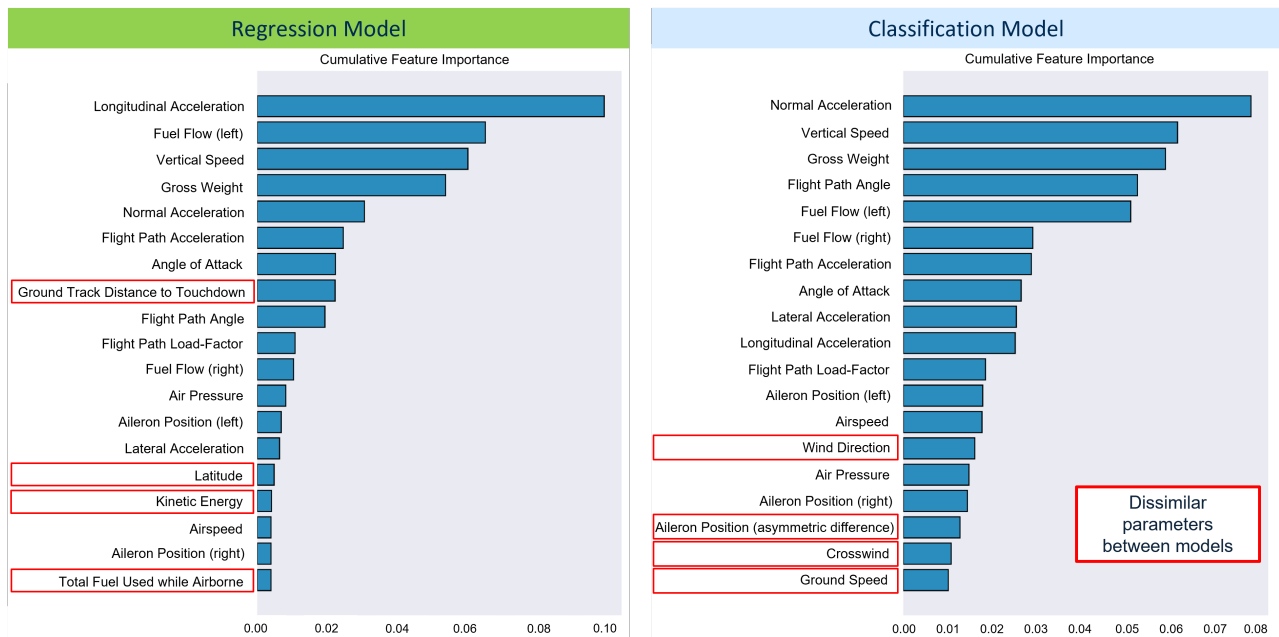


Figure 8 – Critical parameters for turbulence prediction from regression and classification models

Figure 8 shows top 20 parameters against their proportional contribution to the prediction. Note that vertical wind, as expected, is the most important parameter in all the models and is therefore removed from the figure. The plots show that most of the turbulence critical parameters are common for both models. Out of the 40 top parameters, only 8 are different among them (80% of the critical parameters are similar between the two models). Acceleration, vertical speed, and fuel flow are among the top parameters for turbulence prediction. Although this sample feature importance rank seems consistent across the ML model types, future research efforts should focus on characterizing the consistent top features as more feature vector sets are deployed. The features that rank high among both models can be further monitored to help identify early-warning systems for turbulence.



## 5. Concluding Remarks

In this paper, initially, a literature review to better understand turbulence as a phenomenon as well as understand the different approaches to predict turbulence for improving flight safety was undertaken. A methodology to estimate EDR from the airline FOQA data was successfully implemented, leveraging the approach for labeling by Chen et al. The EDR values were calculated and added to the feature vector with 10 seconds of previous data. Once predicted by the ML algorithm, the EDR values were used along with turbulence level thresholds to create turbulence labels. Model performance metrics, such as f1 score and accuracy of confusion matrices as well as R2 and RMSE were used to evaluate the effectiveness of the models. The model performance metrics were improved by implementing stratified sampling, hyperparameter optimization, and reducing overfitting by changing key parameters. The regression and classification models were built for predicting turbulence 10 seconds in advance of the actual occurrence with a reasonably high level of accuracy. Critical parameter identification enabled isolating the parameters that have the highest contribution to successfully predicting turbulence. The results show that both the classification and regression models are effective in predicting the severity levels of turbulence 10 seconds prior to the event. Although additional analysis needs to be made to demonstrate its effectiveness, commercial airlines can leverage the methodology here presented and develop prototypes that test the approach for turbulence prediction. In a comparison between regression and classification models, 80% of the top critical parameters are similar across both model types. For the set of feature vectors and model hyperparameters used, regression had 7% higher f1 score than classification. The goal of this effort was to predict turbulence severity levels with high accuracy and with enough time for a pilot to safety precautions. Ideally, flight safety analysts at commercial airlines will be able to use the methodology to record predicted turbulence severities. In this way, the predictive models to be created will be progressively improved and the prediction horizons made longer time intervals in advance. The main beneficiaries of the developed methodology will be both flight crew and passengers (since they will not experience discomfort and will have lower risk of being injured) as well as the commercial airlines (since lower injury cases are likely to be reported as well as structural damage during turbulence).

## 6. Copyright Statement

The authors confirm that they hold copyright on all of the original material included in this paper. The authors also confirm that they have obtained permission, from the copyright holder of any third party material included in this paper, to publish it as part of their paper. The authors confirm that they give permission, for the publication and distribution of this paper as part of the ICAS 2021 proceedings or as individual off-prints from the proceedings.

## 7. Contact Author Email Address

Contact author (Tejas Puranik): [tpuranik3@gatech.edu](mailto:tpuranik3@gatech.edu)

## References

- [1] Golding, W.L., 2000. Turbulence and its impact on commercial aviation. *Journal of Aviation/Aerospace Education & Research*, 11(2), p.8.
- [2] Krozel, J., Klimenko, V. and Sharman, R., 2011. Analysis of clear-air turbulence avoidance maneuvers. *Air Traffic Control Quarterly*, 19(2), pp.147-168.
- [3] Cotter, A., Williams, J.K., Goodrich, R.K. and Craig, J., 2007. A Random Forest Turbulence Prediction Algorithm. In *5th AMS Conference on Artificial Intelligence Applications to Environmental Science* (Vol. 1).
- [4] Haverdings, H. and Chan, P.W., 2010. Quick access recorder data analysis software for windshear and turbulence studies. *Journal of Aircraft*, 47(4), pp.1443-1447.
- [5] Huang, R., Sun, H., Wu, C., Wang, C. and Lu, B., 2019. Estimating eddy dissipation rate with QAR flight big data. *Applied Sciences*, 9(23), p.5192.
- [6] Klimenko, V. and Krozel, J., 2009. Impact Analysis of Clear Air Turbulence Hazards. In *AIAA Guidance, Navigation, and Control Conference* (p. 6067).
- [7] Krozel, J.A. and Sharman, R., 2015. Remote Detection of Turbulence via ADS-B. In *AIAA Guidance, Navigation, and Control Conference* (p. 1547).

- [8] Sheridan, K., Puranik, T.G., Mangortey, E., Pinon-Fischer, O.J., Kirby, M. and Mavris, D.N., 2020. An application of DBSCAN clustering for flight anomaly detection during the approach phase. In AIAA Scitech 2020 Forum (p. 1851).
- [9] Puranik, T.G. and Mavris, D.N., 2018. Anomaly detection in general-aviation operations using energy metrics and flight-data records. *Journal of Aerospace Information Systems*, 15(1), pp.22-36.
- [10] Das, S., Matthews, B.L., Srivastava, A.N. and Oza, N.C., 2010, July. Multiple kernel learning for heterogeneous anomaly detection: algorithm and aviation safety case study. In *Proceedings of the 16th ACM SIGKDD international conference on Knowledge discovery and data mining* (pp. 47-56).
- [11] Matthews, B., Das, S., Bhaduri, K., Das, K., Martin, R. and Oza, N., 2013. Discovering anomalous aviation safety events using scalable data mining algorithms. *Journal of Aerospace Information Systems*, 10(10), pp.467-475.
- [12] Lee, H., Puranik, T.G. and Mavris, D.N., 2021. Deep spatio-temporal neural networks for risk prediction and decision support in aviation operations. *Journal of Computing and Information Science in Engineering*, 21(4), p.041013.
- [13] Puranik, T.G., Rodriguez, N. and Mavris, D.N., 2020. Towards online prediction of safety-critical landing metrics in aviation using supervised machine learning. *Transportation Research Part C: Emerging Technologies*, 120, p.102819.
- [14] Diallo, O.N., 2012, October. A predictive aircraft landing speed model using neural network. In *2012 IEEE/AIAA 31st Digital Avionics Systems Conference (DASC)* (pp. 3D2-1). IEEE.
- [15] Tong, C., Yin, X., Li, J., Zhu, T., Lv, R., Sun, L. and Rodrigues, J.J., 2018. An innovative deep architecture for aircraft hard landing prediction based on time-series sensor data. *Applied Soft Computing*, 73, pp.344-349.
- [16] Martinez, D., Fernández, A., Hernández, P., Cristóbal, S., Schwaiger, F., Nunez, J.M. and Ruiz, J.M., 2019. Forecasting Unstable Approaches with Boosting Frameworks and LSTM Networks.
- [17] Basora, L., Olive, X. and Dubot, T., 2019. Recent advances in anomaly detection methods applied to aviation. *Aerospace*, 6(11), p.117.
- [18] Hegde, J. and Rokseth, B., 2020. Applications of machine learning methods for engineering risk assessment—A review. *Safety science*, 122, p.104492.
- [19] Chen, Q., Wang, Z., Wan, J., Feng, T., Chen, P., Wang, C. and Zhang, C., 2019, October. Design of a Turbulence Prevention System Based on ATG. In *2019 IEEE 1st International Conference on Civil Aviation Safety and Information Technology (ICCASIT)* (pp. 445-449). IEEE.
- [20] Ackley, J.L., Puranik, T.G. and Mavris, D., 2020. A Supervised Learning Approach for Safety Event Precursor Identification in Commercial Aviation. In *AIAA AVIATION 2020 FORUM* (p. 2880).
- [21] Lee, H., Madar, S., Sairam, S., Puranik, T.G., Payan, A.P., Kirby, M., Pinon, O.J. and Mavris, D.N., 2020. Critical Parameter Identification for Safety Events in Commercial Aviation Using Machine Learning. *Aerospace*, 7(6), p.73.
- [22] Breiman, L., 2001. Random forests. *Machine learning*, 45(1), pp.5-32.
- [23] Gill, P.G. and Buchanan, P., 2014. An ensemble based turbulence forecasting system. *Meteorological Applications*, 21(1), pp.12-19.
- [24] Kim, S.H., Chun, H.Y. and Chan, P.W., 2017. Comparison of turbulence indicators obtained from in situ flight data. *Journal of Applied Meteorology and Climatology*, 56(6), pp.1609-1623.
- [25] Krozel, J., Robinson, P., Buck, B. and Wang, D., 2009. Modeling and feedback of turbulence hazards based on automated real-time pilot reports. In *AIAA Guidance, Navigation, and Control Conference* (p. 6066).