Improving the Nowcast of Unstable Approaches

Zhenming Wang, Ph.D. candidate
Center for Air Transportation Systems Research
George Mason University
Fairfax, VA, USA
zwangk@gmu.edu

Lance Sherry, John Shortle
Center for Air Transportation Systems Research
George Mason University
Fairfax, VA, USA
lsherry@gmu.edu, jshortle@gmu.edu

Abstract*—The approach and landing is a complex phase of flight, in which flight crews need to conduct a series of procedures in a condensed time frame to make the approach stable. Unstable approaches can lead to events with reduced safety margins and additional operating cost. To ensure flights have stable approaches, airlines have established stabilized approach criteria to make sure the landing flight is on track, within appropriate range of speed and rate of descent. If the criteria are not satisfied at stabilization altitudes, a go-around must be executed. Nowcasting the likelihood of unstable approaches before reaching the stabilization altitudes may assist flight crews in taking actions to correct the flight trajectory to avoid a potential unstable approach. A previous study shows the feasibility of detecting unstable approaches using historical trajectory data and the feasibility of nowcasting unstable approaches using state variables. This paper describes improvements in the nowcast performance through the modification of existing features and the addition of more features. For the Nowcast at 6 nm for unstable approaches after reaching 1000' AGL, the accuracy was improved from 71.3% to 74.8%, the recall was improved from 60.8% to 64.7%, and the precision was improved from 74.0% to 78.4%. The improvements are discussed.

Keywords*—*unstable approach; prediction; supervised learning

I. INTRODUCTION

The approach and landing is one of the most complex phases of flight with procedures to be conducted by the airline flight crew in a short time frame. Unstable approaches can occur due to multiple goal reconciliation by the flight crew and can lead to significant consequences such as go-arounds or incident/accident such as runway excursions ([1], [2]). To ensure a safe approach and landing trajectory, the Standard Operating Procedures (SOPs) of airlines have established criteria for stabilized approach. In general, these criteria requires aircraft to be on runway centerline, on approach glide-path, within appropriate range of reference airspeed (e.g. $\pm$10 knots), and within appropriate range of required rate of descent (i.e. < 1000 ft/min). If any of these criteria is not satisfied at 1000' AGL (IMC) or 500' AGL (VMC), the procedure requires the flight crew to perform a go-around.

Nowcasting potential unstable approaches prior to the stabilization altitude 1000' AGL could provide lead time for flight crew to make adjustments to avoid a potential unstable approach. Kinematic models used in the Flight Management System (FMS) to predict future aircraft state are not practical for predicting unstable approaches because the events that will occur during flight progress (e.g. flap/slat and extension, ATC clearances …) cannot be considered. A previous study showed the feasibility of using surveillance track data to identify potential unstable approaches [3] and introduced a supervised learning methodology for unstable approach prediction [4]. The preliminary results showed the feasibility of nowcasting unstable approaches using flight track data.

In this paper, the prediction results are improved relative to previous work by incorporating a larger set of feature variables into the nowcast model. For example, at 6 nm from the runway threshold, the nowcast for 1000' AGL improves in accuracy (from 71.3% to 74.8%), recall (from 60.8% to 64.7%), and precision (from 74.0% to 78.4%) of predictions. At 9 nm, the nowcast for 1000' AGL improves in accuracy (from 61.3% to 65.4%), recall (from 43.6% to 55.7%), and precision (from 63.3% to 65.8%).

This paper describes refinement of the previous nowcast model to further improve its prediction performance. This paper is organized as follows. Section II introduces the methods to identify potential unstable approaches from historical flights data and wind data. Section III introduces the predictive modeling methods with introduction of modified and added features. Section IV discusses the nowcast results and shows the improvements based on previous results. Section V discusses the implications, limitations, and future work of this study.

II. IDENTIFICATION OF UNSTABLE APPROACHES

This section summarizes a methodology for identifying unstable approaches from historical track and wind data. A preliminary version of this methodology was published in [3]. However, that methodology did not account for wind conditions.

To detect an unstable approach, surveillance track data and wind data are used together with aircraft parameters and information about the navigational procedures. First, the flight track data and wind data are preprocessed to reduce noise and to obtain a satisfactory sampling rate. The runway centerline, glide-path, and approach procedures are obtained from navigational procedures. From these, a geometric region is drawn around the approach procedure to identify outliers (Figure 1). The cross-sectional position distributions at 6 nm and at the runway threshold are studied for determining the dimension of lateral and vertical deviation for a majority of flights (e.g. with 95% flights are inside). Then the boundaries are linearly connected between 6 nm and 0 nm. Beyond 6 nm, the boundary sizes are fixed. Any tracks going out of this boundary are considered not acquiring the runway centerline (laterally) or glide-path (vertically). Based on the published navigation procedures of EWR 22L, the level flight altitude is at 3,000' AGL which is at about 10 nm.

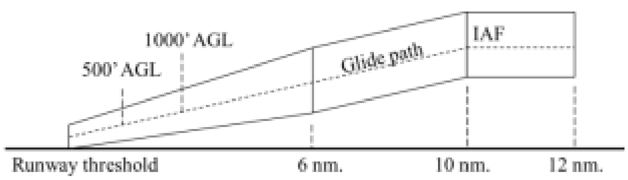

Figure 1 Side View of the Geometric Region for Detecting Unstable Approaches

Given the stabilized approach criteria from different sources, the following criteria are used for detecting unstable approaches:

- An aircraft stays inside the boundaries of the wireframe after reaching 1000' AGL (or 500') until touchdown.
- The change in airspeed from 1000'AGL (or 500') to landing is less than 10 knots.
- The rate of descent is not greater than 1000 ft/min for more than 10 seconds after reaching 1000' AGL (or 500').

A flight not satisfying any of these criteria is marked as unstable for corresponding stabilization altitudes.

The airspeed of an aircraft is derived using groundspeed and wind speed. The aircraft groundspeed is derived from flight track data, and the wind speed is obtained from METAR data. The relationship between the three speed vectors are shown in Figure 2. The groundspeed is the vector sum of airspeed and wind speed.

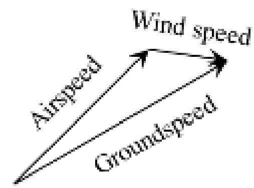

Figure 2 Relationship between Airspeed, Wind speed, and Groundspeed

Further details about the methods to detect unstable approaches can be found in [3] and [4].

III. PREDICTIVE MODELING

A. *Feature selection*

In previous work [4], ten basic features were selected as inputs to the prediction model, including aircraft current state variables at the nowcast location and several historical state variables prior to the nowcast location. In this paper, 3 previous features are modified and 12 new features are added to enrich the input information and to improve the prediction performance. Of particular importance are added features related to the wind conditions (Table III).

Overall, the 22 basic features capture the state of an aircraft at the nowcasting location. These features are summarized in Tables I, II, and II. Features related to the recent *historical* behavior of a flight are shown in Table I. These features include: average airspeed, altitude, and course at 12 nm (x_1 through x_3), level flight time before descending (x_4), distance traveled after lateral/vertical acquisition (x_5 and x_6), distance traveled after the deceleration point (x_7), and number of times penetrating the wireframe (x_8). The level flight time is defined as the time a flight is between 2700' and 3300' AGL between 12 nm and 10 nm from the runway threshold. This is modified from previous work, where this feature was the time an aircraft spent inside the level flight segment of the geometric region. The lateral (or vertical) acquisition point is defined to be the first point the flight is laterally (or vertically) inside the wireframe zone. The deceleration point is the point where the aircraft airspeed first drops 10 knots or more below its airspeed at 12 nm. For lateral/vertical acquisition and deceleration points, if the flight has not acquired these locations, the relative distances between the current location and these points are set to zero. The number of times penetrating the geometric region before the current nowcasting location is a measure of variability of flight trajectory.

TABLE I. FEATURES FOR HISTORICAL AIRCRAFT PERFORMANCE

Feature	Description	Update Status
x_1	Airspeed at 12 nm (knot)	New
x_2	Altitude at 12 nm (ft)	New

x_3	Angle difference between course and centerline at 12nm (degree)	New
x_4	Time spent in level flight segment (second)	Modified
x_5	Distance from lateral acquisition position to current position (nm)	Existing
x_6	Distance from vertical acquisition position to current position (nm)	Existing
x_7	Distance from deceleration point to current position (nm)	New
x_8	Number of times penetrating wireframe before current location	Existing

Features reflecting the *current* state of an aircraft are summarized in Table II. These include an aircraft's basic information at current location such as weight, speed, position, and course. In addition to the existing features in previous work such as airspeed and rate of descent, the deviation of airspeed and rate of descent from the baseline values for current aircraft weight class at the current location are included in this study. The previous features reflect the magnitude of speed/rate of decent, while the added features represent the deviations of these feature values from the baseline values they are expected to be at the location. These baseline values of airspeed/rate of descent for each weight class at each location is estimated by calculating the mean airspeed/rate of descent of all flights with their aircraft type belong to the weight class. For positional features, the spirit is quite the same. Both lateral and vertical position relative to the centerline and the absolute value of positional deviations are considered. Finally, the current angle difference between the course and the runway centerline is included.

TABLE II. FEATURES FOR CURRENT AIRCRAFT STATES

Feature	Description	Update Status
x_9	Maximum takeoff weight (lbs)	Existing
x_{10}	Current lateral deviation (negative if left of centerline) (ft)	Existing
x_{11}	Current lateral deviation (absolute value) (ft)	New
x_{12}	Current vertical deviation (negative for below glide-path) (ft)	Existing
x_{13}	Current vertical deviation (absolute value) (ft)	New
x_{14}	Current average airspeed (knots)	Modified
x_{15}	Current average airspeed deviation from baseline of corresponding weight class (absolute value) (knots)	New
x_{16}	Current average rate of descent (ft/min)	Existing
x_{17}	Current average rate of descent deviation from baseline of corresponding weight class (absolute value)	New
x_{18}	Current angle difference between course and centerline (degrees)	Existing

Finally, the features related to wind conditions are included in Table III. The four wind features used are crosswind speed, crosswind magnitude, headwind, and gust, all derived from METAR data. The original wind data set contains wind speed, direction and gust with a non-constant sampling rate. The wind speed and direction records are interpolated to estimate the wind condition at a nowcasting current time. The crosswind and headwind are derived given the direction of current wind and the alignment of the runway. In these features, feature x_{19} is the signed crosswind speed which can be either positive or negative, and feature x_{20} is the absolute value of magnitudes, i.e. $x_{20} = |x_{19}|$.

TABLE III. FEATURES FOR CURRENT WIND CONDITION

Feature	Description	Update Status
x_{19}	Crosswind speed (negative for northwest wind) (knots)	New
x_{20}	Crosswind speed magnitude (absolute value) (knots)	New
x_{21}	Headwind (knots)	New
x_{22}	Gust (knots)	New

B. Logistic Regression

A logistic regression model is applied to predict the unstable approach events using the input feature values. The model is trained using 5,000 randomly chosen flights out of the 8,158 available tracks, and is tested using the remaining 3,158 tracks. Several measures of prediction performance are reported and analyzed.

The mathematical expression of logistic regression is given in the following equations (1)-(4). First, a hypothesis function is expressed as follows:

$$h_\theta(x) = \frac{1}{1+e^{-\theta x}}, \tag{1}$$

where x is a column vector containing all the feature values and θ is a row vector containing all the regression coefficients. $h_\theta(x)$ can be treated as the predicted probability that a flight with feature-vector x experiences an unstable approach after reaching 1000' AGL. The binary output of this model is 1 (i.e., an unstable approach is predicted) if $h_\theta(x)$ is greater than 0.5 or 0 otherwise.

To measure the deviation of the predicted outputs from the actual results, a cost function $J(\theta)$ is defined as follows:

$$J(\theta) = -\frac{1}{m}\sum_{i=1}^{m}\left[y^{(i)}\log h_\theta\left(x^{(i)}\right) + \left(1-y^{(i)}\right)\log\left(1-h_\theta\left(x^{(i)}\right)\right)\right] \tag{2}$$

where m is the number of training samples, $y^{(i)}$ is the actual outcome of sample i, which is equal to 1 if an unstable approach event is detected for flight i, otherwise $y^{(i)} = 0$. The goal is to minimize the cost function so that the model is trained to best fit the relationships between features and outputs.

The cost function is convex, indicating that a local minimum will be the global minimum. A gradient descent algorithm is applied to minimize the cost function to find the optimal set of coefficients θ. The coefficients are updated using the following equation:

$$\theta_j := \theta_j - \alpha\frac{1}{m}\sum_{i=1}^{m}\left(h_\theta(x^{(i)}) - y^{(i)}\right)x_j^{(i)} \tag{3}$$

where j is the index for features and the parameter α controls the step length. The term following α is the gradient $\frac{\partial J(\theta)}{\partial \theta_j}$.

Based on the basic features, some dependent features can be derived for inputs. In this research, the squared terms of the basic features and the interaction terms between basic features are included as derived features. That is, the degree of the model polynomial 2. This results in a total number of 275 features (with 22 basic features, 22 squared features, and 22-choose-2 interaction features, which is 231).

To use the algorithm for prediction, the data set containing all the feature values are split into a training set and a test set. The training set is used to train the model parameters (i.e. the coefficients for features) by minimizing the cost function $J(\theta)$ via the gradient descent method. All features are mean normalized before being fed into the model using (4), where x is the original feature value, μ is the mean value of all samples of the feature in training set, and σ is the standard deviation of the feature. The purpose of normalization is to improve the efficiency of the algorithm.

$$z = \frac{x-\mu}{\sigma}. \tag{4}$$

IV. PREDICTION PERFORMANCE

With added features based on the previous study, improvements in the prediction performance are expected. To verity the improvement, the same case study of EWR Airport Runway 22L is conducted. 28 days of operations with 8,158 detected landing flights to the target runway are used. Historical data including flight trajectories and wind conditions are processed and the stabilized approach criteria are implemented using C++. The logistic regression is implemented in MATLAB. A random set of 5,000 flights out of 8,158 flights is chosen for training the model. The model is tested using the remaining 3,158 flights. Ten replicates of experiments are conducted. Several measures are applied to quantify the performance of the model, including accuracy, recall, precision, and F1 Score.

The *accuracy* measure is defined as the fraction of correct predictions made by the model (e.g., in Table IV, the accuracy of the model is (962+1397)/(962+281+518+1397) = 74.7%). Accuracy is an overall measure reflecting the model performance for all samples of stable and unstable flights. When the classes are skewed, this measure can be misleading. For example, when a very small portion of flights experience an unstable approach, the model may perform poorly on predicting unstable events, yet the total accuracy can be high due to a large numbers of stable flights which might be predicted well. Therefore, some other measures are needed.

The *recall* is the measure of the proportion of correctly identified unstable approaches within the actual unstable ones. The *precision* represents the proportion of correctly identified unstable approaches within all predicted unstable approaches. It can be shown that the percentages of unstable approach events form the baseline for precisions. A prediction model should always produce a higher value of precision compared to the original proportion of unstable flights. *F1 Score* is an aggregate measure of recall and precision. The definition of F1 Score is given in (5), where P and R stand for precision and recall.

$$F1 = \frac{2PR}{P + R}. \tag{5}$$

To demonstrate the calculation of these measures, a sample nowcast outcome at 6 nm is shown in Table IV. In this case, the calculations of these measures are demonstrated below:

- Total accuracy of the model:

  (962+1397)/(962+281+518+1397) = 74.7%;

- The recall of predicting unstable approaches:

  962/(962+518) = 65.0%;

- The precision of predicting unstable approaches:

  962/(962+281) = 77.4%;

- F1 Score: 2*0.65*0.774/(0.65+0.774) = 70.7%.

TABLE IV. SAMPLE NOWCAST PERFORMANCE AT 6 NM

	Actual Unstable	Actual Stable
Predicted Unstable	962	281
Predicted Stable	518	1397

The performances of nowcasting unstable approaches for 1000’ AGL at different locations are shown in Figure 3. The figure shows the mean values of precision, recall, and F1 Score based on ten replicates of experiments at each nowcasting location (i.e. distance from runway threshold). It shows that the nowcast performance improves as the flights progresses from 10 nm to 3.5 nm (the stabilization altitude 1000’ AGL is near 3 nm). The maximum F1 Score achieved is 80.3% (at 3.5 nm), with an average recall of 75.3% and precision 85.9%.

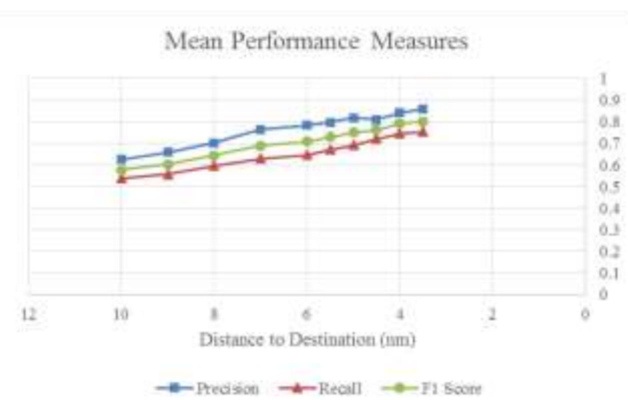

Figure 3 Precision, Recall, and F1 Score at Each Nowcast Location

These measures are improved compared with the results in the previous work. The comparisons of the performance measures is shown in Figure 4. It can be seen that at all locations the current nowcasting model with more features performs better than the previous model with fewer input features. Specifically, up to 10% increase in the recall are observed for nowcast at further distances from runway. The improvements in the precision is less compared to recall, with a maximum increase of approximately 8% at 7 nm. The precisions at nowcast locations that are closer to the runway threshold have not improved significantly. For the overall performance indicator F1 Score, the result shows that the improvements are more significant at further locations (with an approximate 10% increase).

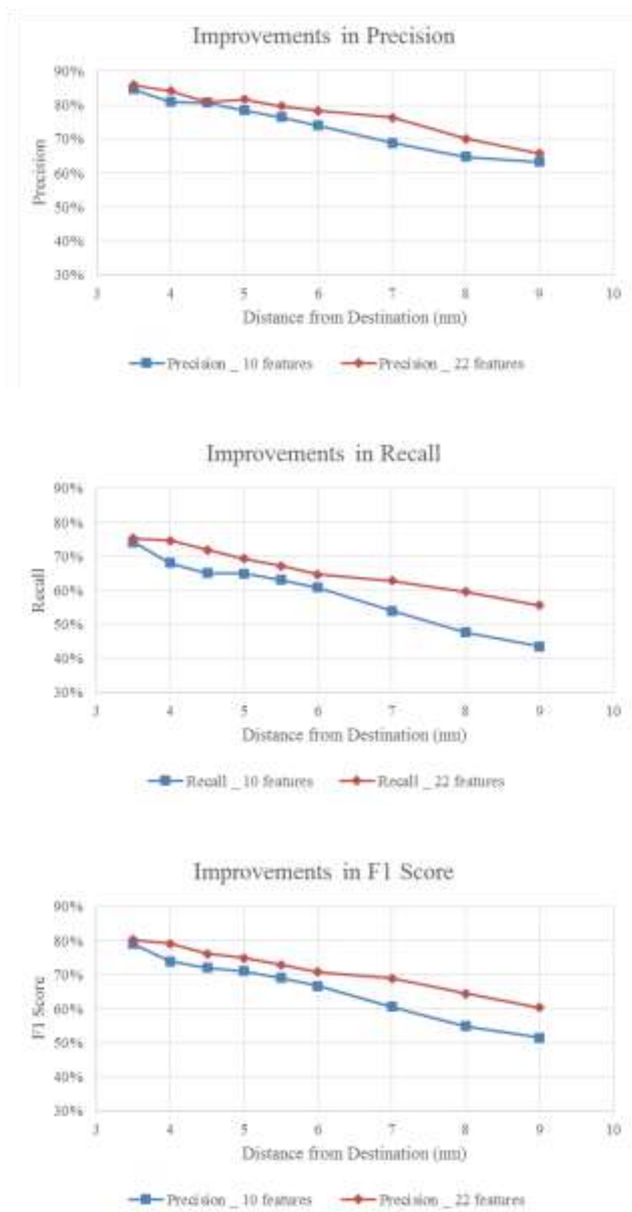

Figure 4 Improvements Based on Previous Work

To use the nowcast results, decision makers may need to predefine a threshold value for F1 Score, e.g. 70%. Then, the nowcast results of unstable approach starting from 6 nm will be considered given the performance of current model. This distance is reduced to 5 nm with previous model. In this sense, the prediction power is increased using current model by adding 1 nm for the nowcast starting point based on the previous model.

V. DISCUSSION

This research uses surveillance track data and wind data to build and improve a nowcasting model for predicting unstable approaches. A case study is conducted at EWR 22L runway. More features are added and improvements in performance over the previous model are found. The current prediction performance measures provide a baseline for future nowcasting performance. Real time onboard data and weather data with much richer features will have more accurate information of conditions, and improvements in the prediction performance are expected. Increasing the amount of historical data used for training and testing the model can also help in improving the prediction performance.

This study applies a logistic regression model for predictive modeling comparing a 22 feature set with a 10 feature set. The results are summarized in Table V, and VI.

TABLE V. SUMMARY OF NOWCASTING PERFORMANCE AT 9 NM

Nowcast at 9 nm	For Unstable Approaches after Reaching 1000' AGL		For Unstable Approaches after Reaching 500' AGL	
	10 Features	22 Features	10 Features	22 Features
Accuracy	61.3%	65.4%	81.9%	81.1%
Precision	63.3%	65.8%	38.0%	34.7%
Recall	43.6%	55.7%	3.1%	9.6%
F1 Score	51.6%	60.3%	5.8%	14.9%

TABLE VI. SUMMARY OF NOWCASTING PERFORMANCE AT 6 NM

Nowcast at 6 nm	For Unstable Approaches after Reaching 1000' AGL		For Unstable Approaches after Reaching 500' AGL	
	10 Features	22 Features	10 Features	22 Features
Accuracy	71.3%	74.8%	82.1%	82.3%
Precision	74.0%	78.4%	47.9%	48.8%
Recall	60.8%	64.7%	12.8%	27.5%
F1 Score	66.8%	70.9%	20.1%	35.2%

There are some limitations for this study. First, the input data have limited precision levels which can limit the performance of the prediction. Second, the causes of unstable approaches can be complex. Many potential causal factors are beyond the current data set, which can explain the gap of current performance measures from a perfect accuracy. This work is more of proof-of-concept analysis. There is still a lot of research to be conducted before it can be applied in the real world. For example, the design and integration of this system into the FMS, whether the level of accuracy is useful on the flight deck and how the nowcast information could be integrated into a flight deck are open questions for future research. Other techniques such as support vector machines and artificial neural networks can be applied in the future work to study their performance using the current data set.

This method can be applied to approaches at other airports and runways. Different runways/airports may correspond to different model parameters. Depending on the approach track patterns, the performance of the model may vary.

ACKNOWLEDGMENTS

The authors acknowledge technical contributions from Anvardh Nanduri (CATSR/GMU), Houda Kourdali (CATSR/GMU), George Donohue (CATSR/GMU), Robert Mauro (Decision Research, Inc.), Julia Trippe (University of Oregon), Michael Feary (NASA), Immanuel Barshi (NASA), Bryan Mathews (NASA), Steve Young (NASA). This research was funded through internal research from the George Mason University Research Foundation.

REFERENCES

[1] Moriarty, D., S. Jarvis (2014) A Systems Perspective on the Unstable Approach in Commercial Aviation. Reliability Engineering and System Safety 131 (2014) 197–202

[2] Sherry, L., Z. Wang, H. Kourdali, J. Shortle. 2013. Big Data Analysis of Irregular Operations: Aborted Approaches and Their Underlying Factors. Integrated Communication, Navigation, and Surveillance Conference, Herndon, VA.

[3] Wang, Z., L. Sherry, J. Shortle (2015) Airspace Risk Management using Surveillance Track Data: Stabilized Approaches. In Proceedings 8th Integrated Communications, Navigation, Surveillance (ICNS) Conference, Dulles, Va. April 21-23, 2015

[4] Wang, ., L. Sherry, J. Shortle (2016) Feasibility of Using Historical Flight Track Data to Nowcast Unstable Approaches. In Proceedings 9th