

FEASIBILITY OF USING HISTORICAL FLIGHT TRACK DATA TO NOWCAST UNSTABLE APPROACHES

Zhenming Wang (Ph.D. Candidate), Lance Sherry (Ph.D.), John Shortle (Ph.D.)

Center for Air Transportation Systems Research (CATSR), George Mason University, Fairfax, VA

Abstract

The approach and landing phase is one of the most complex procedures in airline operations. To mitigate the potential risks, airlines have established “stabilized approach” criteria that require the flight crew to check four conditions at 1000’ above ground level (AGL) and 500’ AGL. If any of these conditions are not satisfied, the flight crew is required to abort the approach adding operational costs to the flight.

Nowcasting unstable approaches prior to the stable approach altitudes (e.g. at 10 nm, 6 nm from the runway threshold) could provide lead time for flight crew to make adjustments to avoid a potential unstable approach. Kinematic models, already used in the Flight Management System (FMS) to predict future aircraft state are not practical as these models cannot account for events that will occur during flight progress (e.g. flap/slat and extension, ATC clearances ...).

This paper describes an analysis of massive amounts of surveillance track data for a given approach to assess the feasibility of using of historical flight track data to nowcast the likelihood of a stable approach given the situation at 10 nm, 6 nm, and 3 nm from the runway threshold. A logistic regression model was able to nowcast the 1000’ AGL aircraft state with accuracy of 61.7% at 10 nm, 73.6% at 6 nm, and 83.1% at 3 nm. Additional measures of prediction performances are summarized. The implications and limitations of this method are discussed.

1 Introduction

The approach and landing phase is one of the most complex procedures in airline flight operations [1]. During this phase of flight, the crew is required to monitor a rapidly evolving situation and make split second decisions based on a large number of factors. Further, decisions made early in the procedure have a big impact on the downstream trajectory.

To mitigate the potential risks, airlines and regulators have established “stabilized approach” criteria that require the flight crew to check four conditions at 1000’ above ground level (AGL) and 500’ AGL: (1) aligned with the extended runway center-line, (2) on the glide path, (3) within a tolerance of the landing airspeed, and (4) not an excessive rate of descent. If any of these conditions are not satisfied, the flight crew is required to abort the approach and perform a Go Around [2]. Go Arouns are expensive, estimated to add an average of \$1,200 to the cost of the flight [3]. In addition, they reduce safety margins by requiring another complex procedure – the Go Around.

Nowcasting unstable approaches prior to the stable approach altitudes could provide lead time to the flight crew to make adjustments to avoid a potential unstable approach.

Using kinematic models, such as those already used by the Flight Management System (FMS), to predict future aircraft state is not practical as these models cannot account for discrete events that may occur during the flight progress (e.g. flap/slat and extension, ATC clearances, etc.). An alternative approach is to use massive amounts of historical flight track data to assess the likelihood of a stable approach given the situation at 10 nm, 6 nm and 3 nm from the runway threshold.

This paper describes an analysis of historic surveillance track data for a given approach to assess the feasibility of using historical flight track data to nowcast the likelihood of a stable approach given the situation at 10 nm, 6 nm and 3 nm from the runway threshold.

An analysis of surveillance track data for 8,158 approaches to Newark International Airport (EWR) Runway 22L showed that 52.4% flights satisfy the stabilized approach criteria between 1000’ AGL and runway threshold, and 82.3% between 500’ AGL and runway threshold [4].

A simple method for nowcasting the likelihood of meeting the stabilized approach criteria at 1000’

and 500' is to assess whether or not the flight is following the desired approach profile at 10, 6 and 3 nm (i.e. on-speed, on lateral and vertical flight-path). Unfortunately, this method has a limited predictive power. There is only a 56% chance of correctly predicting a stable or unstable approach using this approach.

An improved method is to use a logistic regression model. Such a model using ten parameters (groundspeed, rate of descent, heading, altitude, lateral position, first glide path acquisition location, first runway center-line acquisition location, aircraft weight, runway centerline/glide path tracking, and level segment flight time) improves the accuracy of the nowcast.

- At 10 nm, for approaches that end up unstable below 1000', 45.8% of them can be correctly predicted. The nowcast is accurate 62.7% of the time.
- At 6 nm, for approaches that end up unstable below 1000', 63.9% of them can be correctly predicted. The nowcast is accurate 76.0% of the time.
- At 3 nm, for approaches that end up unstable below 1000', 76.2% of them can be correctly predicted. The nowcast is accurate 86.7% of the time.

This analysis shows the degree of accuracy that can be achieved using historical flight track data to nowcast stabilized approaches before reaching 1000' AGL and 500' AGL. Whether this is useful on the flight deck and how this "probabilistic" information could be integrated into "deterministic" flight deck procedures are open research questions.

The paper is organized as follows: Section 2 provides an overview of the approach procedure and stabilized approach criteria. Section 3 describes the historical surveillance track data and how it is processed. Section 4 describes the logistic regression. Section 5 provides a case study for EWR 22 L. Section 6 discusses the limitations of this method and future work.

2 The Approach Phase of Flight

The approach phase of flight transitions the aircraft from a clean configuration (flaps and slats

retracted) at 250 knots to the flap/slat landing configuration at the reference landing speed for the aircraft. During this transition the aircraft must acquire the extended runway center-line and glide path (typically 3 degrees). A typical trajectory for the approach is shown in Figure 1.

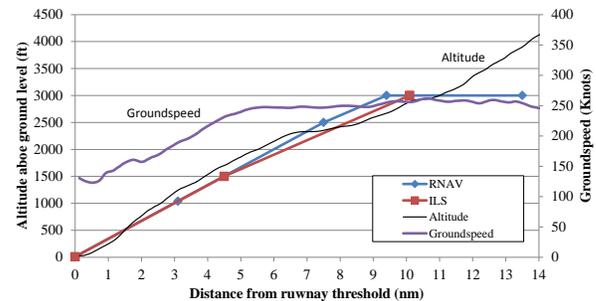


Figure 1: Typical approach trajectory

Setting up and flying the approach is complicated by close proximity of traffic on the approach and their associated wake vortices, as well as departure and surface traffic, weather phenomenon (i.e. wind, visibility, ceilings), aircraft performance, and aircraft system status.

To mitigate these risks, airline best-practices have established stable approach criteria that are applied in approach procedure. Criteria may vary between different airlines. A sample set of criteria is summarized below:

- At 1000'/500' AGL the aircraft ground speed should be within ± 10 knots of the landing ground speed;
- From 1000'/500' AGL to the runway threshold, the rate of descent should not be greater than 1000 feet per minute for more than 10 seconds;
- The aircraft should acquire the glide path by maintaining its position within the defined vertical boundary;
- The aircraft should acquire the runway centerline by maintaining its position within the defined lateral boundary.

If any of these four conditions is not satisfied, the approach is unstable, and the flight crew are required to abort the approach and perform a Go Around.

An analysis of 8,158 approaches to Newark International Airport (EWR) Runway 22L using

surveillance track data and change in ground speed as a proxy for change in airspeed, showed that 52.4% flights satisfy the stabilized approach criteria between 1000' AGL and runway threshold, and 82.3% between 500' AGL and runway threshold.

The objective of this nowcasting system is to increase the number of stable approaches by providing advance indications of not satisfying the stable approach criteria at 10 nm, 6 nm, and 3 nm from the runway threshold.

3 Surveillance Flight Track Data & Procedure Data

Flight Track Data

Ideally, Flight Data Recorder (FDR) or Flight Operational Quality Assurance (FOQA) data would be used for this analysis. However, without access to the FDR/FOQA data, surveillance track data is used to demonstrate the method.

Twenty eight days of surveillance track data from the FAA's National Offload Program for N90 TRACON are used. The key fields used are the track index, aircraft type, destination airport, seconds past midnight, latitude, longitude and altitude.

The sampling interval in the raw surveillance data is between 4 and 5 seconds, and the value is recorded to a precision of one second. The values of latitude and longitude are precise to the fifth decimal place. Altitude values are stored in numbers of flight levels, i.e. with precision level of 100 feet. To improve the fidelity of the data, linear interpolation between each pair of successive raw track points with the time interval of 1 second was added. The data was also smoothed by estimating each track point's horizontal/vertical position to be an averaged value of a series of track points centered at the current track point to reduce the potential variations in raw tracks. The average interval was set to be 10 seconds.

From the interpolated and smoothed tracks, several basic variables can be derived: ground speed, rate of descent and heading angle.

For further details on the processing of this track data see Wang et. al. [3]

Procedure Data

Parameters for the approach procedure are obtained from the Approach Chart and Airport diagram provided by the FAA's National Flight Data Center. The key fields include the runway elevation, glide path angle, and threshold crossing height. These parameters are used for constructing a wire-frame for the approach segment for each runway (Figure 2).

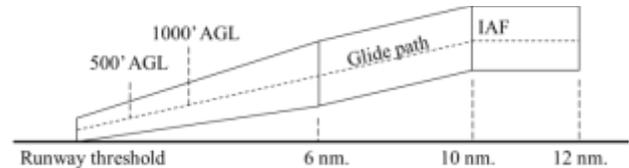


Figure 2: Side View of the Wireframe Zone for Approach

The dimensions of the wire-frame were derived from the approach chart and the statistical distributions. For example, Figure 3 shows the distribution of crossing positions at 6 nm. The red box represents the lateral and vertical size of the wire-frame boundary, which contains 95% of crossing flight in each dimension. The Initial Approach Fix of runway 22L is at about 10 nm from runway threshold. The 6 nm boundary was extended up to 10 nm. Beyond 10 nm, the altitude of the wire-frame is unchanged to model the level flight segment.

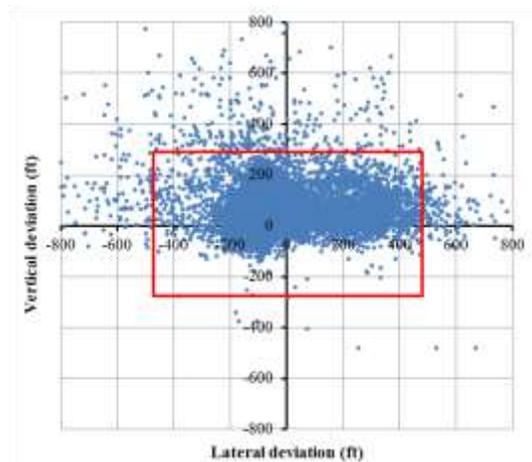


Figure 3 Scatterplot of Aircraft Crossing Positions at 6 nm.

Stable Approach Criteria

Based on the published stabilization criteria from different sources, the stabilized approach criteria used in this algorithm are as follows:

- The aircraft should acquire the glide-path by maintaining its position within the defined *vertical boundary of the wire-frame*
- The aircraft should acquire the runway centerline by maintaining its position within the defined *lateral boundary of the wire-frame*.
- At 1000'/500' AGL the aircraft ground speed should remain within ± 10 knots of the groundspeed at the runway threshold (i.e. landing ground speed). Since this analysis is intended to measure the change in airspeed, the change in groundspeed is used as a proxy.
- From 1000'/500' AGL to the runway threshold, the rate of descent should not be greater than 1000 feet per minute for more than 10 seconds;

If any of these four conditions is not satisfied, the unstable approach occurs at corresponding altitude.

4 Methodology for Nowcasting Unstable Approaches

The general idea for nowcasting unstable approaches is to select several aircraft features from flight track data to build a model to map the values of these features to the occurrence of unstable events. These features can either be state variables (e.g. speed) at the specific nowcasting location, or historical variables prior to the current location. The state variable and historical performance features used for the nowcast are listed in Table 1. In this table, x_2 to x_6 represent the state variables at the nowcasting location, and x_7 to x_{10} represent the historical performance variables before the nowcasting location.

Table 1 Selected Features for Logistic Regression Model

Features	Description
x_1	Maximum takeoff weight (lb.)
x_2	Ground speed at prediction location (knot)

x_3	Rate of descent at prediction location (feet per minute)
x_4	Lateral deviation from runway centerline at prediction location (ft.)
x_5	Vertical deviation from glidepath at prediction location (ft.)
x_6	Heading angle difference from runway centerline at prediction location (degree)
x_7	Distance from lateral acquisition point to current location (zero if not acquired) (nm.)
x_8	Distance from vertical acquisition point to current location (zero if not acquired) (nm.)
x_9	Level segment flight time (s)
x_{10}	Measure of runway centerline/glidepath tracking (number of times an aircraft penetrating the modeled wire-frame zone)

A logistic regression model is developed to establish the relationship between these features and the outcome of stable/unstable approaches. The mathematical expression is given in the following equation:

Equation 1

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

In this equation, $h_{\theta}(x)$ is the probability that a flight with features x ends up in an unstable approach, where x is a vector containing the feature values and θ is a vector containing the regression coefficients. The degree of the polynomial θx is set to 2, so the potential interactions between features are considered in the model. This results in a total of 65 features. The optimal coefficients can be found by minimizing the cost function $J(\theta)$ given below:

Equation 2

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

In this formula, m is the number of training samples and i is the index for samples. $y^{(i)}$ is the actual outcome of sample i , which is equal to 1 if sample flight i ends up with unstable approach, otherwise $y^{(i)} = 0$. \log is the natural logarithm. The Gradient Descent method is applied to find the optimal coefficient θ . The coefficients updating procedure is shown below:

Equation 3

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

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

5 A Case Study of EWR Runway 22L

The nowcast was developed for flights landing at Runway 22L at EWR airport. The data processing and criteria are implemented in a C++ algorithm, and logistic regression is implemented in a MATLAB algorithm. The 28 days of available records, included 18,426 landing flights to EWR airport. Among these flights, 8,158 flights landed at Runway 22L. The majority of these flights were Large aircraft (73.8%).

To evaluate the performance of the nowcast, a random set of 5,000 flights out of 8,158 flights is chosen for training the coefficients of the logistic regression model. The trained model is tested using the remaining 3,158 flights.

Results from the logistic regression nowcasting at 10 nm, 6 nm, and 3 nm from runway threshold for unstable approach events at 1000'/500' are shown in Table 2 to 4. The percentages show the accuracy of the nowcast for 1000' and 500' AGL. The shaded cells show the accuracy of the nowcasts, the unshaded cells show the errors in the nowcasts. The columns for 1000' and 500' add up to 100%.

In the test set of data, at 1000' AGL 46.99% of the flights were unstable. The nowcast at 10 nm

correctly identified 21.5%, and was not correct for 25.49%. At 500' AGL 17.63% of the flights were unstable. The nowcast from 10 nm for unstable approaches was correct for 0.63% and was not correct for 17%. This a consequence of the small number of unstable approaches in the test set as well as the distance from the 500' point.

The nowcast improved as the flights progressed from 10 nm to 6 nm to 3 nm. For the 1000' point the correct nowcast for unstable approaches improved from 21.5% to 30.05% to 35.81. The incorrect nowcast decreased from 25.49% to 16.94% to 11.18%.

Table 2: Nowcast at 10 nm

Nowcast	Actual	1000'	500'
Unstable	Unstable	679 (21.50%)	20 (0.63%)
	Stable	404 (12.79%)	30 (0.95%)
Stable	Unstable	805 (25.49%)	537 (17.00%)
	Stable	1270 (40.22%)	2571 (81.41%)

Table 3: Nowcast at 6 nm

Nowcast	Actual	1000'	500'
Unstable	Unstable	949 (30.05%)	92 (2.91%)
	Stable	299 (9.47%)	83 (2.63%)
Stable	Unstable	535 (16.94%)	465 (14.72%)
	Stable	1375 (43.54%)	2518 (79.73%)

Table 4: Nowcast at 3 nm

Nowcast	Actual	1000'	500'
Unstable	Unstable	1131 (35.81%)	175 (5.54%)
	Stable	182 (5.76%)	155 (4.91%)
Stable	Unstable	353 (11.18%)	382 (12.10%)
	Stable	1492 (47.25%)	2446 (77.45%)

From these tables several performance metrics can be derived: overall accuracy, precision, and recall.

Overall Accuracy

The overall accuracy of the model can be measured by the proportion of the total number of correctly nowcast stable/unstable approaches among all flights (Table 5). The accuracy for the 1000' AGL nowcast improves as the location is closer. The accuracy for the 500' AGL nowcast remains constant.

Table 5 Overall Accuracy

Nowcast location	1000' AGL	500' AGL
10 nm	61.7%	82.1%
6 nm	73.6%	82.7%
3 nm	83.1%	83.0%

The accuracy results for 500' AGL events are counter-intuitively higher than the 1000' AGL results. This is because there are more stable approaches at 500' than stable approaches at 1000'. This biases the measure of accuracy. Additional measures are needed to evaluate the performance of predicting unstable approaches.

Precision

Precision is a measure that answers the question: Given a nowcast result, what is the probability it is correct? For example, when nowcasting 1000' AGL events at 6 nm, the precision of nowcasting unstable approaches can be calculated as $949/(949+299) = 76\%$. Table 6 summarizes the results.

Table 6 Precisions for Nowcasting Unstable Approaches

Nowcast location	1000' AGL	500' AGL
10 nm	62.7%	40.0%
6 nm	76.0%	52.6%
3 nm	86.7%	53.0%

Recall

Recall is another metric for nowcast performance, which answers the question: For a specific event, given that it occurs, what is the probability that it is correctly predicted to occur?

For example, when predicting 1000' AGL events at 6 nm, the recall of nowcasting unstable approaches can be calculated as $949/(949+535) = 63.9\%$. That is for the actual unstable approaches, 63.9% can be correctly predicted at 6 nm using current model. Table 7 summarizes the Recall metrics.

Table 7 Recalls for Nowcasting Unstable Approaches

Nowcast location	1000' AGL	500' AGL
10 nm	45.8%	3.6%
6 nm	63.9%	16.5%
3 nm	76.2%	31.4%

The goal in tuning the logistic regression is to overlap the predicted unstable flights with actual unstable flights (Figure 4).

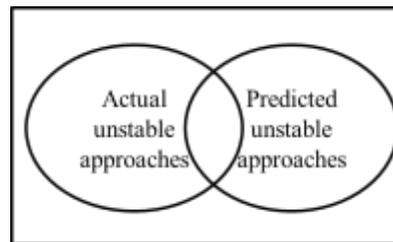


Figure 4: Venn Diagram for Actual and Predicted Unstable Approaches

5 Conclusions

This paper describes logistic regression method for nowcasting unstable approaches using historic flight track data. A case study is conducted at EWR 22L runway to test the nowcast performance of this model. These accuracy, precision and recall metrics provide as baseline for nowcast performance. Adding other features and applying with more/better data sources can improve the prediction performance.

Whether this level of accuracy for a nowcast is useful on the flight deck and how this “probabilistic” information could be integrated into a “deterministic” flight deck are open research questions.

Future Work

The logistic regression model can be extended by inclusion of additional parameters including

wind conditions, traffic, aircraft configuration, and instructions from ATC. In addition, these results were based on a single randomly selected test set. The results should be calculated based on several more randomly selected test sets.

This and future logistic regression models shall be applied to other approaches at other airports.

Other methods could also be applied to the data such as Support Vector Machines, Bayesian Network models. Bayesian Network models could backward-chain and allow the analysis of the factors that lead to an unstable approach at 500' due to an over speed condition.

This logistic regression model does not provide a recommendation to the flight crew to adjust the trajectory to avoid an unstable approach. For example, more specific information on likelihood of not meeting the airspeed criteria would allow the flight crew to take mitigative actions. For some conditions, preliminary results show improved nowcast performance.

This research demonstrated the method and the accuracy using surveillance flight track data. Future work should use FDR/FOQA data that has better fidelity and additional parameters (e.g. flap/slat position). Preliminary results show improved nowcast performance with this type of data.

Acknowledgements

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] FAA. (2003). Advisory Circular. 120-71a Standard Operating Procedures for Flight Deck Crewmembers, Washington, DC.

[3] Sherry, L., Z. Wang, H. Kourdali, J. Shortle (2013). Big Data Analysis of Irregular Operations: Aborted Approaches and their Underlying Factors. In Proceedings 6th Integrated Communications Navigation and Surveillance (ICNS) Conference, Dulles, Va. April, 2013

[4] 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