Enhancing Aircraft Fuel Burn Modeling on the Airport Surface

Emily Clemons^{*} & Tom G. Reynolds[†] MIT Lincoln Laboratory, Lexington, MA 02421, USA

Yashovardhan Chati[‡] & Hamsa Balakrishnan[§] MIT Department of Aeronautics & Astronautics, Cambridge, MA 02139, USA

Aircraft fuel burn modeling is useful for a range of systems analysis activities. Models such as the FAA's Aviation Environmental Design Tool (AEDT) are widely used to support these activities. Previously, limited availability to thrust and movement data forced AEDT to make several simplifying assumptions, which adversely affect the accuracy of its current fuel burn estimates. This paper identifies first order airport surface fuel burn modeling enhancements in the areas of baseline taxi fuel flow modeling, taxi time estimation and pre-taxi fuel burn that may be suitable for inclusion in future versions of AEDT.

I. Introduction

A IRPORT surface fuel burn models are used for a variety of applications, including assessing impacts of changes to operational procedures, airport infrastructure and fleet mixes and conducting environmental studies. One of the standard tools to support these activities is the FAA's Aviation Environmental Design Tool (AEDT) [1]. In previous versions of AEDT, limited availability of key data required developers to make a variety of simplifying assumptions which may affect the accuracy of the fuel burn and emissions predictions [2,3]. Increased availability of data, coupled with appropriate analysis techniques, now provides an opportunity to make enhancements to the models, while also meeting the needs of a broader range of potential users, ranging from "basic" users seeking aggregate, pre-canned model options, through "advanced" users seeking control over many aspects of model parameters.

The main simplifying assumptions identified for consideration in this work are discussed below. Firstly, AEDT's current model assumes a constant engine-specific thrust level (and resulting fuel flow rate) during taxi, determined from engine manufacturer certification data [4]. However, this assumption can be significantly different than actual characteristics during operational conditions for a given aircraft [5]. This can be because of a variety of factors, such as the age of the engine (as the engine gets older the amount of fuel it burns changes), airline standard operating procedures (like single-engine taxi) and pilot technique (such as "riding the brakes" instead of throttling down the engines when coming to a stop). Secondly, default taxi times (which may be used by basic users) are either based on outdated empirical data or default to standard certification Landing and Take-Off (LTO) cycle which assumes 26 minutes of taxi time on the airport surface, typically broken into 19 min taxi-out and 7 min taxi-in. Different airports may have very different taxi times. Using empirical data to determine realistic taxi time distributions is effective, but these distributions need to be updated regularly to capture evolving airport conditions. Finally, the fuel

^{*} Assistant Staff, Air Traffic Control Systems Group.

[†] Associate Group Leader, Air Traffic Control Systems Group. AIAA Associate Fellow.

[‡] Research Engineer, Department of Aeronautics and Astronautics, MIT. AIAA Member.

[§] Associate Professor, Department of Aeronautics and Astronautics, MIT. AIAA Associate Fellow.

burn contribution in the non-movement area from the gate time, pushback and engine start events (including engine and Auxiliary Power Unit (APU) contributions) are typically neglected in current models but can be quite significant. This paper addresses these three issues in turn as illustrated in Figure 1. Empirical data from Flight Data Recorders (FDRs) are leveraged to build statistical and predictive models of fuel flow for a given airport and aircraft type. These analyses are designed to capture "first order" enhancements and provide recommendations for future development of tools such as AEDT.



Figure 1. Airport surface fuel modeling enhancements.

II. Enhanced Taxi Fuel Flow Modeling

The discussion in this section focuses on taxi-out conditions, but findings could be extended to taxi-in conditions too as needed. Figure 2 shows a typical fuel flow rate profile (post-pushback and engine start) during taxi-out. It can be seen that the fuel flow rate profile (red curve) can be divided into two distinct regions: a baseline region and a fuel flow spike region. The baseline fuel flow region is characterized by almost constant (low variation) fuel flow rate over extended time intervals under normal moving taxi conditions, and the fuel flow spike region corresponds to periods of increased thrust needed to re-start movement after a period of being stopped. Therefore, these two fuel flow rate regions need to be modeled separately [6].



Figure 2. Typical fuel flow rate profile in taxi-out (adapted from [6]).

Table 1 shows different characteristics of the baseline fuel flow region for two example aircraft types available from the FDR data. It can be seen that more than 90% of the taxi-out fuel consumption occurs during the baseline fuel flow region on average. Therefore, in the current work where the objective is to identify first-order modeling enhancements, only the baseline fuel flow region is modeled and the fuel flow spikes are neglected. Figure 2 also shows a mean baseline fuel flow rate (in blue) obtained by averaging the baseline fuel flow rates for a particular taxi-out operation.

Table 1. Characteristics of the baseline fuel flow region. The table shows the mean and the range of time spent and fuel mass consumed in the baseline fuel flow region, as a percentage of the total time and fuel burn in taxi-out (reproduced from [5]).

Aircraft type		Time (%)	Fuel burn (%)	
	Mean	Range	Mean	Range
A330-343	94.1	76.1 - 100.0	91.0	68.0 - 100.0
B777-300ER	93.0	77.4 - 100.0	91.0	73.0 - 100.0

In this paper, the mean baseline fuel flow rate per engine during taxi is regressed against the mean values of the corrected ambient temperature (θ_{∞}) and pressure (δ_{∞}) during taxi for consistency with the Boeing Fuel Flow Method [1]. The mean baseline fuel flow rate is chosen as the output variable as a simplified approach to model the baseline fuel flow rate as a constant (as the baseline fuel flow shows little variation). The ambient values are corrected by the International Standard Atmosphere (ISA) sea-level static values of temperature (288.15 K) and pressure (101,325 Pa). This simplified model to estimate the mean baseline fuel flow rate is intended as an improvement over the equations used currently in the AEDT models to estimate taxi fuel flow. An Ordinary Least Squares (OLS) regression approach [7] is found to be sufficient to develop this simplified model. Table 2 shows the OLS-derived equations for modeling the fuel flow rate in taxi-out for the different aircraft types in the operational FDR dataset. All the regression coefficients are found to be statistically significant at a 5% significance level. The table also shows the number of flights for each aircraft type used in the FDR training sets.

А/С Туре	Engine Type	# Training Obs.	OLS Model Equation
A320-214	2 × CFMI CFM56-5B4/2	103	$0.812 \cdot \dot{m_{f_{ICAO}}} \cdot \delta_{\infty}^{-0.123} \cdot \theta_{\infty}^{-0.483}$
A321-111	2 × CFMI CFM56-5B1/2	46	$0.796 \cdot \dot{m_{f_{ICAO}}} \cdot \delta_{\infty} \cdot \theta_{\infty}^{0.209}$
A330-343	2 × RR Trent 772B-60	117	$0.779 \cdot \dot{m_{f_{ICAO}}} \cdot \delta_{\infty} \cdot \theta_{\infty}^{0.350}$
A340-313	$4 \times \text{CFMI CFM-56 5C4/P}$	37	$1.019 \cdot \dot{m_{f_{ICAO}}} \cdot \delta_{\infty}^{-6.690} \cdot \theta_{\infty}^{0.597}$
B777-300ER	2 × GE GE90-115BL	81	$0.753 \cdot \dot{m_{f_{ICAO}}} \cdot \delta_{\infty} \cdot \theta_{\infty}^{0.717}$
C Series-100 (RJ)	$2 \times PW PW1542G$	95	$0.966 \cdot \dot{m_{f_{ICAO}}} \cdot \delta_{\infty} \cdot heta_{\infty}^{0.186}$

Table 2. OLS regression equations to model fuel flow rate per engine during taxi-out. $m_{f_{ICAO}}$ is the ICAO Databank fuel burn index during taxi-out.

We also compare the predictions from such baseline fuel flow models with the estimates provided by AEDT, which uses the ICAO Aircraft Engine Emissions Databank [4] fuel burn indices multiplied by the Boeing Fuel Flow Correction factor of 1.1 during taxi-out [1]. The model evaluation is conducted using an independent test set comprising the entire taxi-out trajectory (i.e., the time elapsed between pushback and start of takeoff roll, including the time instants corresponding to fuel flow spikes). The metrics chosen for evaluation are the mean error and the mean absolute error. The mean error is the average of the relative mean prediction errors on unseen data in the test set and is an indicator of the bias in the model. The mean absolute error is the average of the absolute value of the relative mean prediction errors on unseen data in the test set and predictions. The model evaluation results are shown in Table 3.

Table 3. Performance of the OLS-based baseline fuel flow rate models and the AEDT models to
predict fuel flow rates on unseen test data during taxi-out.

	# Test	Mean error (%)		Mean absolute error (%)	
А/С Туре	Observations from FDR	OLS Model	AEDT	OLS Model	AEDT
A320-214	34	1.0	36.3	13.3	39.4
A321-111	14	3.8	47.1	14.9	50.1
A330-343	37	-3.0	36.4	5.8	39.1
A340-313	12	-0.7	7.8	9.1	12.5
B777-300ER	25	-2.2	42.3	3.1	43.1
C Series-100 (RJ)	30	0.1	17.7	5.5	19.3

Each cell in the table reports the mean of the evaluation metric across all the flights in the unseen test dataset for each aircraft type. From the table, it can be observed that the median value of the fuel flow rate mean error across the different aircraft types is -0.3% for the OLS regression models. The median value of the fuel flow rate mean absolute error given by the OLS regression models across the different aircraft types is 7.5%. The OLS regression models give a statistically significant (at a 5% significance level) lower fuel flow rate mean absolute error than the AEDT models for each aircraft type in taxi-out. The reduction in mean absolute error for the OLS regression models is up to about 93% as compared to the AEDT models. The results thus suggest that significant benefits may be achieved over the current AEDT models through the data-driven methodology developed in this paper.

III. Enhanced Airport Taxi Time Estimates

Different AEDT users may have different needs with respect to taxi time estimates. Advanced users will likely use flight-specific taxi time estimates based on model outputs or surveillance data from actual operations (e.g., from ASDE-X). However, basic users will likely want to use aggregate taxi time estimates provided by AEDT and these users are the focus of this analysis. Airport-specific taxi out times are available in current versions of AEDT but these date from 2006 or earlier. For this part of the study, recent taxi-out data were collected from the FAA's Aviation System Performance Metrics (ASPM) [8]. This dataset contains flight-specific taxi-out times, available to the nearest minute. ASPM data from flights across 25 major US airports was aggregated for dates between October 2016 and September 2017, to provide a more recent model of the distribution of taxi out times at a given airport. This analysis could be extended to taxi-in operations and to other US or international airports as needed.

Figure 3 below shows the updated taxi out time distributions for three diverse sample airports: New York LaGuardia (LGA), Charlotte Douglas (CLT) and Washington Reagan (DCA). These distributions characterize the taxi-out time at each airport over the one year data period across all operating conditions (e.g., airport configuration, weather conditions, etc.). Such distributions are suitable for aggregate Monte Carlo type analyses. As expected, the times vary significantly within and between airports. For this particular set of airports, LGA is seen to have the longest peak and most varied taxi-out time; this is not surprising given the high congestion levels at LGA. The peak in the total taxi-out time distributions for LGA, CLT, and DCA are 18, 15, and 13 minutes, respectively. Compared to the standard 19 minutes of taxi-out time assumed from the LTO cycle (shown by the dashed magenta line in Figure 3), these correspond to errors of 5.6%, 26.7%, and 46.2% of the typical taxi out times for these particular airports.



Figure 3. Example ASPM taxi-out time distributions for LGA, CLT and DCA airports.

A similar comparison was conducted across all 25 airports, to get a more accurate representation of the true taxi-out times. However, sometimes it is beneficial to reduce the number of taxi time distributions to a few clusters which capture the spectrum of differences seen in the taxi-out distributions for all airports. To this end, each of the airport distributions were fitted to a cumulative distribution function (CDF) and the resulting curves were clustered. Since the CDF is a cumulative function, differences in the shape of the distributions will accumulate, and therefore better results are achieved by clustering on the CDFs instead of the probability density functions (PDFs). Clustering the data was accomplished through implementing a Mean Shift algorithm [9] but other techniques could also be applied. The resulting number of clusters can be modified by changing the bandwidth parameter for the Mean Shift algorithm. There is a trade-off

between the number of clusters and the error between each of the actual distributions and the corresponding cluster approximations. Fewer clusters allows for a simpler categorization of the different distributions, but results in higher error. A larger number of clusters will decrease this error, but at the cost of defeating the purpose of clustering. To determine the optimal number of clusters for the taxi-out distributions at these airports, the relationship between the error and the number of clusters was investigated. It was seen that the error continued to decrease steadily until around 6 clusters, at which point adding additional clusters had a reduced incremental impact on the total error, with zero error when there were as many clusters as airports. Figure 4 shows the CDF curves for the taxi-out distribution data for each of the 25 airports, as well as the percentage error when estimating the individual airport distributions with the resulting cluster centroids.

Figure 5 shows the results of the clustering. The left plot shows the fitted CDFs seen in the previous figure, with the color corresponding to the cluster group for that curve. The centroid for each cluster is shown as the solid, thicker lines. The right plot shows the PDF representative of each group of airports, obtained by differentiating the CDF centroids. The results show that the majority of airports have taxi-out distributions which can be approximated with a few distributions, while a few airports have different enough characteristics that they have their own clusters.



Figure 4. CDFs of the taxi-out times for the 25 airports (left); Average error per number of resulting clusters (right).



Figure 5. Resulting cluster centroids: CDFs (left) and PDFs (right).

The boxplot in Figure 6 allows a side-by-side comparison of all the airport taxi-out distributions across the 25 airports studied, as well as the differences between the clusters. The 19-minute taxi-out simplification is provided as a reference, along with the error between this assumption and median of each of the distributions. It is seen that for some clusters (such as cluster 1), the 19-minute approximation is relatively accurate. However for other clusters (such as cluster 5), such an estimate would introduce significant error compared to the actual taxi-out time at the airports in this group. It is also interesting to note which airports are paired together. For example, JFK and LGA have been identified as airports with similar taxi-out characteristics. Both distributions have a higher spread and higher mean compared to the other airports. In contrast, Chicago Midway (MDW) is dissimilar enough from any of the other airports that it has been placed into a cluster by itself.



Figure 6. Boxplot of airport distributions by clustering group (outliers removed).

The 19-minute default taxi-out time assumption is intended to represent average airport taxi time. This chart shows that the errors in this estimate vary from 0% to 72.7% for these particular airports, which is one reason why users typically do not use the 19 minute default taxi time. By using recent historical data at an airport, the error resulting from predicting the taxi-out time for a given flight can be decreased drastically. In addition, if it is better to have fewer groups of airports, much of the error can still be reduced by clustering airport taxi-out data and using the respective cluster centroid as an estimation for the flight's taxi-out time.

IV. Adding Pre-Taxi (Gate, Push-back & Engine Start) Fuel Estimates

In order to establish a more accurate model of the fuel burn at a given airport, the fuel consumed by both the engine and APU during the "pre-taxi" phases at the gate, push-back and engine startup up processes has also been investigated in this study. Differences in fuel burn for these phases between the 25 major US airports studied above were found to be negligible, however the fuel burn distributions were found to vary significantly between different aircraft types. For this analysis, data from the FDR set used for the baseline fuel flow modeling was used again. This contains a record over time of information specific to a flight, including fuel burn and velocity. Figure 7 below shows the raw FDR data for a sample flight during pushback. For this part of the analysis, the flight was broken up into multiple segments, as the gate, pushback, and engine start have different APU and engine fuel burn settings. APU fuel burn rates were obtained from an Airport Cooperative Research Program (ACRP) guidance document [10], which groups aircraft into categories (Narrow Body, Wide Body, Jumbo Wide Body, Regional Jet, and Turbo Prop) and gives the

APU settings for the "no load" (gate), "environmental control systems" (pushback), and "main engine start" conditions for each aircraft category. The APU is turned on while still at the gate in the "no load" condition. Through discussion with an experienced commercial pilot, it was determined that the APU is first turned on typically between 10-15 minutes before pushing back from the gate at large US airports. Therefore, for all aircraft, the gate time was assumed to be 12.5 minutes in this study, although different assumptions may be appropriate at other airports, for example where off-gate stands are more common. Pushback was defined from the point at which the aircraft began to move back from the gate, until the point at which one of the engines began burning fuel. As can be seen in Figure 7, most aircraft begin starting the first engine while still in the process of pushback by the tug from the gate, before halting and completing engine startup with the remaining engines. Engine startup was defined from the end of pushback to when the aircraft begins to move for taxi after all engines have started up and post-engine checklists are complete.



Figure 7. Example FDR data for a single flight gate, engine and push-back events.

Much of the work incorporated pre-processing the data before performing the statistical analysis, as many of the flights had corrupted data, such as non-zero fuel or velocity at the beginning of the track. In order to get an accurate representation of the total fuel burned from gate to start of taxi, corrections had to be made for the missing data. A significant portion of flights had tracks beginning with the aircraft already in the pushback phase. During pushback, the aircraft's engines have not yet powered on, therefore the only source of fuel is from the APU. Since APU fuel burn during pushback is calculated using a fixed rate multiplied by the duration of the pushback phase, only the time missing from the start of the track is needed to recover the missing fuel burn. By approximating the velocity curve of the flight, this elapsed time can be calculated. This works as during pushback all flights move away from the gate before slowing down to a halt during the engine start process, and thus exhibit similar velocity profiles. Figure 8 shows a few sample flights with the actual velocity (blue curve) and the fitted velocity curve (red), and depicts how the missing time was estimated. The additional fuel burn over this time interval was added to the total fuel burn for that flight to get the corrected overall fuel burn.



Figure 8. Calculating missing time from start of pushback using velocity profiles.

In addition to these corrections, further pre-processing had to be completed on the B757 flights, since the start of all the tracks were missing up to midway through the engine start process. Unlike previously, the velocity profiles could not be approximated and used to find the elapsed time, as the entire pushback phase was missing. Therefore, the fuel burn during pushback was estimated by looking at the average fuel burn during pushback over all other flights in the FDR dataset. Determining the amount of fuel burn missing from the engine start-up procedure was more complicated, as during this phase there is fuel burn from the engine in addition to that from the APU. An analysis was done on all other twin engine flights in the database, and it was predicted that the typical time between the start of both engines for the B757 was about 40 seconds. It was then assumed that engines on a particular aircraft behaved similarly, such that the engine startup profiles are identical. Combining both of these results, the fuel burn during engine startup was calculated using the fuel flow profile of the second engine for the first, and assuming the first engine idled for 40 seconds before the start of the second engine. Figure 9 provides an illustration of this process.



Figure 9. Example flight with estimated fuel flow.

Once tracks had been corrected for these issues, the fuel burn totals for the gate/pushback/engine start processes were aggregated over all the flights of a given aircraft type available in the FDR data as a statistical approach to building the fuel burn histograms from historical data. The resulting fuel burn distributions for the types studied are shown in Figure 10, as the solid curves on the left plot.



Figure 10. PDF curves for gate, pushback, and engine start fuel burn by aircraft type (left); Regression line comparing total fuel burn and aircraft weight (right).

The relationship between fuel burn and aircraft weight was then investigated as a means to predict the pretaxi fuel burn of aircraft types not within the FDR dataset. The maximum gross takeoff weight (MGTOW) was used for each data type, pulled from the Eurocontrol Base of Aircraft Data (BADA) Version 3.6 dataset [11]. The total fuel burned during gate/pushback/engine start was seen to be linearly related to the weight of the aircraft type, and this correlation was used to then predict the approximate fuel burn for aircraft types not available in the FDR data set. This is shown in the right plot of Figure 10, and estimates for some example types using the observed correlation are presented as the dashed lines on the left plot.



Figure 11. Pre-taxi fuel burn as a percentage of total fuel burned during departure from gate to wheels-off, per aircraft type (left) and aggregate distribution (right).

The total fuel burn between the gate up to the point at which an aircraft begins to taxi is seen to vary significantly between aircraft types, and can be a significant fraction of total surface fuel burn. Figure 11 shows the percentage of pre-taxi fuel burn relative to total surface fuel burn from the types analyzed from the FDR data. For narrow body aircraft, this percentage of total fuel burn is seen to be higher than the larger body aircraft, as seen in the left plot of Figure 11. The right plot of the figure shows that the typical

contribution of the pre-taxi fuel in comparison to the total fuel burned on the aircraft surface is between 10-40% on average, reinforcing the need to carefully consider this component of surface fuel burn.

V. Conclusions and Next Steps

This paper has identified first order enhancements to airport surface fuel burn modeling in the areas of baseline taxi fuel flow modeling, taxi time estimation and pre-taxi fuel burn that could improve the modeling accuracy of AEDT when analyzing environmental impacts of operations at a given airport. Discussions with AEDT developers are now underway to identify the suitability of these enhancements for inclusion in future releases of the tool. Work is also on-going to identify and evaluate implementation options for future versions of AEDT in a number of areas, including:

- Refine first order AEDT surface fuel burn improvements including extending baseline fuel estimates to a wider range of aircraft types, more refined modeling of APU fuel burn rates and usage times to reflect differences between airports and airlines; extending taxi time distributions to a wider set of airports (potentially the Core 30) and operating conditions, such as taxi-out and in time distributions conditioned on factors such as weather (VMC, IMC), airport congestion level, airport runway configuration; and extending pre-taxi fuel estimates and correlation techniques to a wider range of aircraft types, operating conditions, etc.
- Extend surface fuel burn analyses to align with planned additional AEDT enhancements based around different users, such as "basic users" wanting the ability to select pre-canned options representative of typical operating conditions, e.g., based on ASPM-derived empirical distributions; "intermediate users" wanting the ability to modify behaviors based on appropriate modeled parameters; and "advanced users" wanting complete control over all aspects of aircraft and airport dynamics, e.g., based on ASDE-X data.
- Undertake studies to explore how the first order fuel modeling enhancements detailed in this paper allow refinements of surface noise and emissions estimates, as well as their temporal and spatial assignments.
- Coordinate with AEDT developers to transition key findings into the development plan for the tool.

References

- M. Ahearn et al., "Aviation Environmental Design Tool (AEDT) Technical Manual", Version 2b, Service Pack 3, U.S. Department of Transportation John A. Volpe National Transportation Systems Center, Report No. DOT-VNTSC-FAA-16-11, 2016.
- [2] ACRP 02-45, "Methodology to Improve EDMS/AEDT Quantification of Aircraft Taxi/Idle Emissions", Transportation Research Board, 2016.
- [3] S. Herndon, E. Wood, J. Franklin, R. Miake-Lye, W. B. Knighton, M. Babb, A. Nakahara, T. Reynolds and H. Balakrishnan. "Measurement of Gaseous HAP Emissions from Idling Aircraft as a Function of Engine and Ambient Conditions," ACRP 02-03A, May 2012.
- [4] International Civil Aviation Organization (ICAO), "ICAO aircraft engine emissions databank." [Online database], cited 12 February 2014.
- [5] Nikoleris, T., Gupta, G., Kistler, M., "Detailed Estimation of Fuel Consumption and Emissions During Aircraft Taxi Operations at Dallas/Fort Worth International Airport", *Transportation Research Part D*, Vol. 16, No. 4, pp. 302–308, 2011.
- [6] Chati, Y.S., Statistical modeling of aircraft engine fuel burn, Ph.D. Thesis, Department of Aeronautics and Astronautics, Massachusetts Institute of Technology, Cambridge, MA, USA, 2018.
- [7] Kutner, M.H., Nachtsheim, C.J., Neter, J., Li, W., *Applied Linear Statistical Models*, 5th Edition, McGraw Hill Education (India) Private Limited, New Delhi, 2013.

[8] FAA, Aviation System Performance Metrics (ASPM) database, https://aspm.faa.gov/, 2018.

- [9] Fukunaga, K., & Hostetler, L, "The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition", *IEEE Transactions on Information Theory*, Vol. 21, No. 1, pp. 32-40, 1975.
- [10] ACRP 02-25, "Handbook for Evaluating Emissions and Costs of APUs and Alternative Systems", Transportation Research Board, 2012.
- [11] Nuic, A., "User Manual for the Base of Aircraft Data (BADA)", EUROCONTROL, EEC Technical/Scientific Report, 2004.

Acknowledgments

Many thanks to Dr. Hua (Bill) He and Dr. Mohammed Majeed of the FAA Office of Environment and Energy and the AEDT development team for their support of this activity.

Disclaimer

DISTRIBUTION STATEMENT A. Approved for public release: distribution unlimited. This material is based upon work supported by the Federal Aviation Administration under Air Force Contract No. FA8721-05-C-0002 and/or FA8702-15-D-0001. Any opinions, findings, conclusions or recommendations expressed in this material are those of the author(s) and do not necessarily reflect the views of the Federal Aviation Administration.