NASA/TM-2019-220420

Modeling Uncertainty in Time and Fuel Benefit Estimation for TASAR Operational Evaluation

Joseph Burris
Virginia Commonwealth University, Richmond, Virginia

Sara R. Wilson
Langley Research Center, Hampton, Virginia

David J. Edwards
Virginia Commonwealth University, Richmond, Virginia

Kathryn M. Ballard
Langley Research Center, Hampton, Virginia

November 2019

NASA STI Program . . . in Profile

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA scientific and technical information (STI) program plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NTRS Registered and its public interface, the NASA Technical Reports Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and by NASA in the NASA STI Report Series, which includes the following report types:

- TECHNICAL PUBLICATION. Reports of completed research or a major significant phase of research that present the results of NASA Programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA counter-part of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.

- TECHNICAL MEMORANDUM. Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.

- CONTRACTOR REPORT. Scientific and technical findings by NASA-sponsored contractors and grantees.

- CONFERENCE PUBLICATION. Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or co-sponsored by NASA.

- SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, often concerned with subjects having substantial public interest.

- TECHNICAL TRANSLATION. English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services also include organizing and publishing research results, distributing specialized research announcements and feeds, providing information desk and personal search support, and enabling data exchange services.

For more information about the NASA STI program, see the following:

- Access the NASA STI program home page at http://www.sti.nasa.gov

- E-mail your question to help@sti.nasa.gov

- Phone the NASA STI Information Desk at 757-864-9658

- Write to:
  NASA STI Information Desk
  Mail Stop 148
  NASA Langley Research Center
  Hampton, VA 23681-2199

NASA/TM-2019-220420

Modeling Uncertainty in Time and Fuel Benefit Estimation for TASAR Operational Evaluation

Joseph Burris
Virginia Commonwealth University, Richmond, Virginia

Sara R. Wilson
Langley Research Center, Hampton, Virginia

David J. Edwards
Virginia Commonwealth University, Richmond, Virginia

Kathryn M. Ballard
Langley Research Center, Hampton, Virginia

November 2019

Available from:

NASA STI Program / Mail Stop 148
NASA Langley Research Center
Hampton, VA 23681-2199
Fax: 757-864-6500

Table of Contents

List of Figures

List of Tables

Abbreviations and Acronyms

ATC	Air Traffic Control
CAR(1)	Continuous Autoregressive with One Time Lag
EFB	Electronic Flight Bag
REML	Restricted Maximum Likelihood
TAP	Traffic Aware Planner
TOC	Top of Climb
TOD	Top of Descent
TASAR	Traffic Aware Strategic Aircrew Requests

Abstract

The Traffic Aware Strategic Aircrew Requests concept aims to reduce weather-induced delays, improve route efficiency, and efficiently share route modification options by combining onboard avionics data, Automatic Dependent Surveillance-Broadcast data, and broadband internet data to generate optimal, traffic-compatible trajectory changes based on real-time traffic and weather data. Time and fuel benefits due to use of the Traffic Aware Planner (TAP) software can be estimated by taking the difference in predicted flight time and fuel usage before and after a TAP-inspired trajectory change is completed. Although TAP's optimization algorithm predicts flight and fuel usage based on the current flight route and weather data, it does not account for possible air traffic controller-initiated trajectory changes, reroutes due to sudden weather changes, or other pilot/controller actions that may occur during flight. This paper introduces an approach for quantifying the uncertainty in estimated time and fuel benefits.

1 Introduction

The FAA reports 70% of the total National Airspace System delays in U.S. operations are due to adverse weather [1]. NASA is developing new air traffic management concepts and technologies that focus on reducing weather-induced delays, improving route efficiency, and efficiently sharing route change options. The Traffic Aware Strategic Aircrew Requests (TASAR) concept involves the determination of optimal, traffic-compatible trajectory changes based on real-time traffic and weather data. In order to implement the TASAR concept, NASA developed the Traffic Aware Planner (TAP) software which includes optimization tools and a user interface [2]. TAP combines on-board avionics data, Automatic Dependent Surveillance-Broadcast data, and broadband internet data to generate the recommended trajectory changes. Each TAP instance operates individually and can be installed on a commercial-off-the-shelf Electronic Flight Bag (EFB) making implementation simple and low cost.

TAP generates optimal trajectory changes, on average, every ten seconds during flight. A pilot can set TAP's optimization objective to one of time savings, fuel savings, or a combination of both through the EFB interface. Once TAP generates a trajectory change that a pilot elects to utilize, they will make a standard pilot-initiated trajectory change request (i.e., a TAP-inspired request) to Air Traffic Control (ATC). If approved, the pilot records the approval in TAP and thus switches the active route to follow the trajectory change. TAP switches back to the original route once the aircraft reaches the first waypoint where the aircraft rejoins the original route (known as the *rejoin* waypoint).

TASAR was tested in two Human-in-the-Loop experiments and two flight trials before being installed on Alaska Airlines commercial aircraft in September 2017 [2], [3]. Three Alaska Airlines aircraft were equipped with TAP to test its capabilities outside a controlled environment. The primary goal of these tests was to quantify the time and fuel savings with TAP for individual flights, total by region, and total annually. Stage 1 of data collection took place from September 2017 to July 2018. For these flights, TAP ran in the background gathering data to be used as a control. Stage 2 of TAP data collection took place between July 2018 and April 2019 with a small group of Alaska Airlines pilots actively using TAP and requesting TAP-inspired trajectory changes. Stage 3 was designed to expand TAP usage to a larger group of pilots [3].

Although TAP's optimization algorithm predicts flight time and fuel usage based on the current flight route and weather data, it does not account for possible ATC-initiated trajectory changes, reroutes due to sudden weather changes, or other pilot/ATC actions that may occur during flight. This paper outlines a method for analyzing Stage 1 flight records to quantify uncertainty in estimated time and fuel benefits for both individual and aggregate flights. The rest of this paper is outlined as follows. In Section 2, we provide an overview of the Stage 1 data and the method used for estimating time and fuel benefits. The proposed approach for calculating uncertainty in estimated benefits is described in Sections 3 and 4. We apply the approach to individual and aggregate flights in Section 5. Section 6 presents concluding remarks.

2 Data

2.1 Defining Benefits

Defining uncertainty in estimated time and fuel benefits is predicated on defining the time and fuel benefit of an individual flight. The most natural definitions of time and fuel benefit are given by equation (1).

$$\begin{aligned} Benefit_{Time} &= FlightTime_{TAP} - FlightTime_{NoTAP} \\ Benefit_{Fuel} &= FuelUsage_{TAP} - FuelUsage_{NoTAP} \end{aligned} \tag{1}$$

In equation (1), $FlightTime_{TAP}$ refers to actual flight time when TAP is employed while $FlightTime_{NoTAP}$ is actual flight time when TAP is not employed and all other flight conditions are the same. Definitions are similar for $Benefit_{Fuel}$. Note that, as defined, a negative benefit is desirable and indicates time or fuel savings via TAP. Unfortunately, we cannot calculate actual benefits (i.e., equation (1)) since flight time and fuel usage both with and without TAP cannot be measured on the same flight. Since TAP predicts flight time and fuel usage in order to make optimal route determinations, we can make use of these predictions in order to *estimate* benefits. However, as mentioned previously, uncertainty in estimated benefits arises due to events (such as vectoring, unexpected weather patterns, etc.) that may occur during flight that are unaccounted for in TAP's predictions. Thus, we define estimated time and fuel benefits in equation (2) below.

$$\begin{aligned} \widehat{Benefit}_{Time} &= FlightTime^{*}_{TAP} - FlightTime^{*}_{NoTAP} + Uncertainty_{Time} \\ \widehat{Benefit}_{Fuel} &= FuelUsage^{*}_{TAP} - FuelUsage^{*}_{NoTAP} + Uncertainty_{Fuel} \end{aligned} \tag{2}$$

In equation (2), $FlightTime^{*}_{NoTAP}$ and $FuelUsage^{*}_{NoTAP}$ represent TAP's predictions of total flight time and fuel usage, respectively, taken just prior to the execution of the first TAP-inspired request. As a consequence, these predictions are based on the original flight route (i.e., before any TAP-inspired trajectory change has been executed). Similarly, $FlightTime^{*}_{TAP}$ and $FuelUsage^{*}_{TAP}$ are TAP predictions of total flight time and fuel usage, respectively, taken immediately after the final rejoin waypoint. $Uncertainty_{Time}$ and $Uncertainty_{Fuel}$ represent the uncertainty in estimated time and fuel benefits, respectively.

Uncertainty in estimated benefits should, intuitively, depend on how far the first TAP prediction is made in advance of the final rejoin waypoint. If the time between the first TAP-inspired trajectory change and crossing the last rejoin waypoint is small, weather-related issues, vectoring, and other sources of variability (which, again, TAP does not take into account) have less time to occur. We define this time difference as Δt (shown in equation (3) below).

$$\Delta t = Time@FinalRejoin - Time@FirstRequest \tag{3}$$

$Time@FirstRequest$ and $Time@FinalRejoin$ represent the timestamps at the first TAP-inspired trajectory change and final rejoin waypoint, respectively.

Assuming that TAP can only impact estimated benefits while a TAP-inspired trajectory change is being executed, Figure 1 illustrates the calculation of estimated benefits. In Figure 1, the blue line represents the original flight path while the green line indicates TAP inspired trajectory changes.

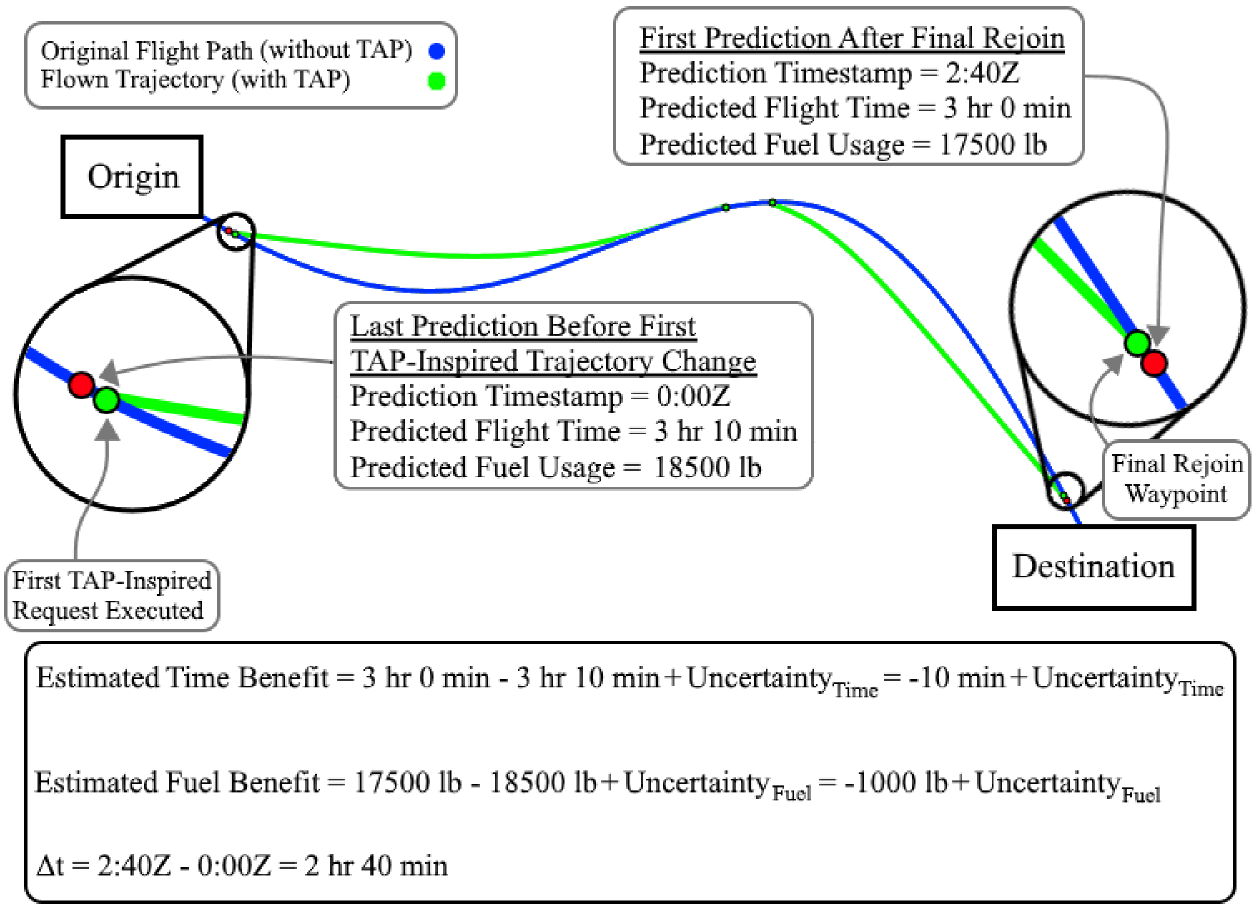


Figure 1: Illustration of Estimated Benefits Calculation.

2.2 Measuring Uncertainty in Estimated Benefits

The main goal of this analysis is to model how uncertainty in estimated benefits change with Δt using data from Stage 1. However, since no Stage 1 flights executed any TAP-inspired requests, one must select two timestamps (say, t_0 and t_1, where $t_0 < t_1$) along the flight route to represent $Time@FirstRequest$ and $Time@FinalRejoin$ in equation (3).

For the analysis, we fix the final rejoin waypoint at the top of descent (TOD). Denote this timestamp as t_{TOD}. The TOD is the point in time when the aircraft begins its final descent from cruise altitude towards the destination and is a reasonable choice for the final rejoin waypoint. The timestamp where a theoretical TAP-inspired request could be executed (i.e., t_0) is allowed to vary between the top of climb (TOC), when the aircraft first reaches cruising altitude, and t_{TOD} (see Section 2.4.3 for more details regarding this choice). Thus, Δt can be redefined as in equation (4).

$$\Delta t = t_{TOD} - t_0 \tag{4}$$

Since no Stage 1 flights executed any TAP-inspired requests, we know that the estimated time and fuel benefit for these flights should be zero. As a consequence, we can set $\widehat{Benefit}_{Time} = 0$ and $\widehat{Benefit}_{Fuel} = 0$ in equation (2). With this framework, uncertainty in estimated benefits is defined for Stage 1 flights as the difference in TAP predictions made at t_0 and t_{TOD}. That is, with some simple algebra, equation (5) below follows directly from equation (2).

$$\begin{aligned} Uncertainty_{Time} &= FlightTime^*_{t_0} - FlightTime^*_{t_{TOD}} \\ Uncertainty_{Fuel} &= FuelUsage^*_{t_0} - FuelUsage^*_{t_{TOD}} \end{aligned} \tag{5}$$

For the remainder of this paper, we refer to uncertainty in estimated time benefit as *time uncertainty* and uncertainty in estimated fuel benefit as *fuel uncertainty.* Note that a single flight will have thousands of values for time and fuel uncertainty corresponding to different values of Δt.

2.3 Data Sources

The proposed analysis makes use of TAP prediction data as well as more general flight characteristics data. These two types of data arise from different sources. Flight characteristic data such as airport of origin cannot be reliably obtained through TAP data recordings as TAP is often turned on by pilots after departure. Instead, we obtain this data through the air traffic tracking service flightaware.com. Although almost all the flights in the dataset we consider occur in or out of the Pacific Northwest (as shown in Figure 2), there are a variety of flight lengths represented in the final dataset (see Figure 3). Overall, we consider the dataset to be representative of Alaska Airlines operations.

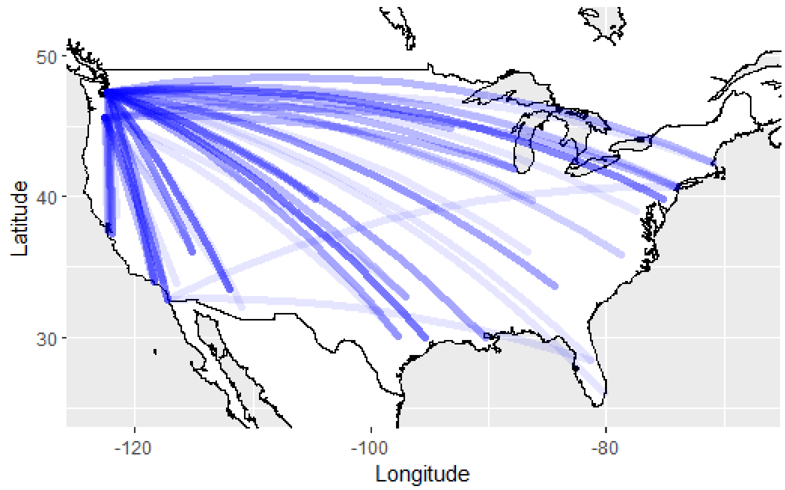


Figure 2: Origin-Destination Pairs for Flights Used in the Analysis.

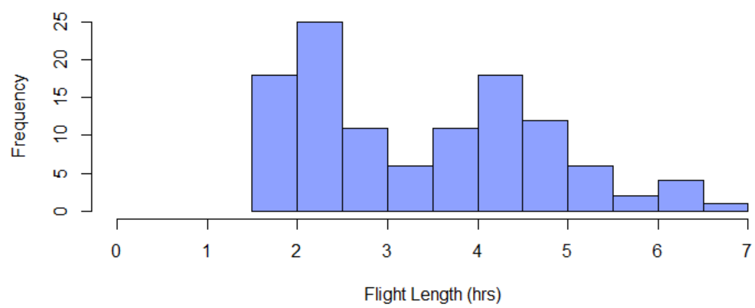


Figure 3: Observed Flight Times for Flights Included in the Final Dataset.

Following guidelines set in other analyses of the Alaska Airlines flight trials, the following criteria must be met for a flight to be included in the dataset:

- Departure and arrival airports must be within the continental US,
- Flight time above flight level 180 within the continental US must make up at least 75% of total flight time,
- At least one of the following:
  - Cumulative 30 minutes of data collection
  - Data collected for at least 50% of the flight time above flight level 180.

Note that one transcontinental flight whose data was not representative of the entire flight was removed manually (outside the scope of the guidelines) due to an error in the initial input flight route. This was the only flight removed from the dataset outside the scope of the above criteria. With these guidelines in place, a total of 114 flights were included in the dataset.

2.4 Data Quality

2.4.1 Stage 1a, Stage 1b, and Stage 1a Rerun Datasets

Stage 1a flights lacked in-flight internet connectivity which led to TAP's predictions being based solely on on-board avionics. This led to less suitable predictions for analysis [3]. TAP was had internet connectivity and algorithm updates for Stage 1b. All internet-based data feeds (such as weather data) were recorded on the ground at NASA throughout Stage 1a. This data was later combined with the recorded avionics data under the new algorithm to produce a set of data with updated predictions (denoted as the Stage 1a *rerun* data). A comparison of the time and fuel uncertainty for a single flight before and after the rerun can be seen in Figures 4a and 4b.

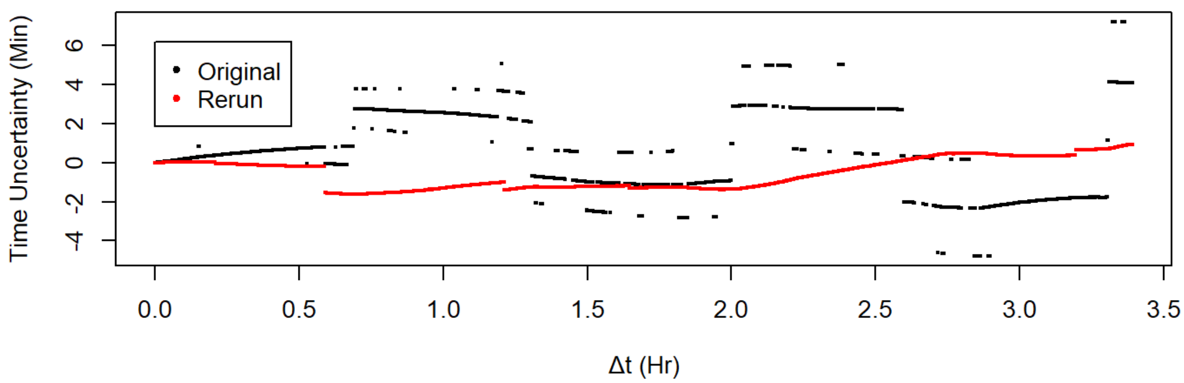


Figure 4a: Comparison of Time Uncertainty Plots for a Stage 1 Flight before and after Being Rerun with Updated TAP Software.

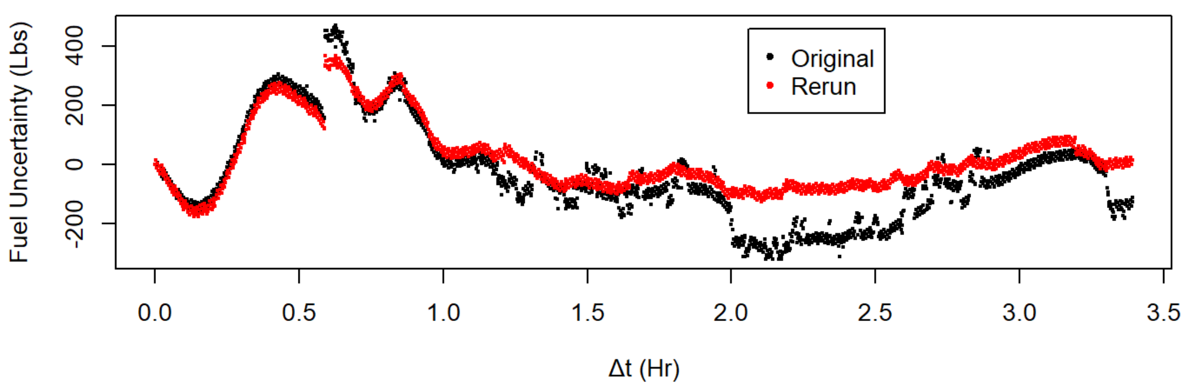


Figure 4b: Comparison of Fuel Uncertainty Plots for a Stage 1 flight before and after Being Rerun with Updated TAP Software.

The discrete jumps and overall noisiness in Figure 4a for the original data indicate TAP transitioning between multiple optimal routes. In contrast, with the addition of internet-based data feeds, optimal flight routes generated by TAP during the rerun simulations tended to be more similar over time. Figure

4b indicates that fuel uncertainty values were more similar for the original and rerun data. Due to the improvement in Stage 1a rerun data, our analysis uses Stage 1 data comprised of Stage 1a rerun and Stage 1b data.

2.4.2 Bias in Fuel Uncertainty

TAP's fuel predictions rely heavily on fuel readings from on-board instruments. These predictions overestimate the weight of the aircraft at the destination leading to a negative bias in fuel uncertainty as seen in Figure 5. We conjecture this is due to fuel readings received by TAP on aircraft with multiple fuel tanks and that problems arise when the active fuel tank is switched midflight.

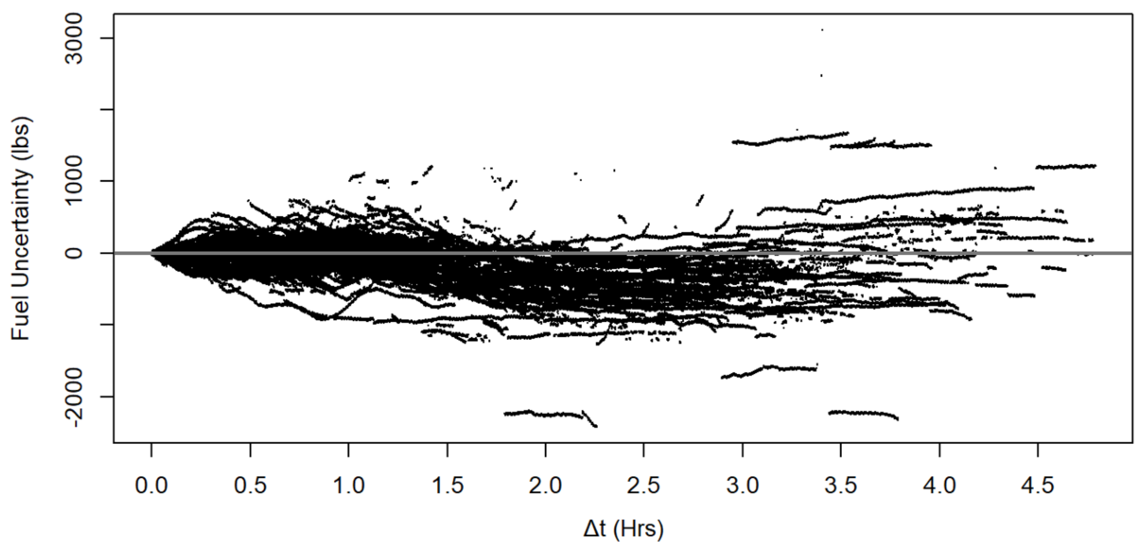


Figure 5: Fuel Uncertainty Negative Bias Illustration.

We can attempt to account for the fuel bias by splitting the dataset based on Δt. However, since the bias likely occurs due to switching fuel tanks, this is not an exact method to handle fuel bias. The development of a methodology to appropriately account for fuel bias is a recommended area of future research.

2.4.3 Fuel Uncertainty during the Initial Climb

As mentioned in Section 2.1, only flight data taken between the TOC and TOD was used in this analysis. During the initial climb, uncertainty in estimated fuel benefit exhibits "behaviors" that are distinctly different than at cruising altitude (see Figure 6 for an example flight). We hypothesize that this is due to differences between the as-flown altitude and the planned altitude during the initial climb. The behavior is non-linear in nature and depends heavily on the time the aircraft reaches its TOC. As a consequence, we elected to utilize uncertainty values between the TOD and TOC.

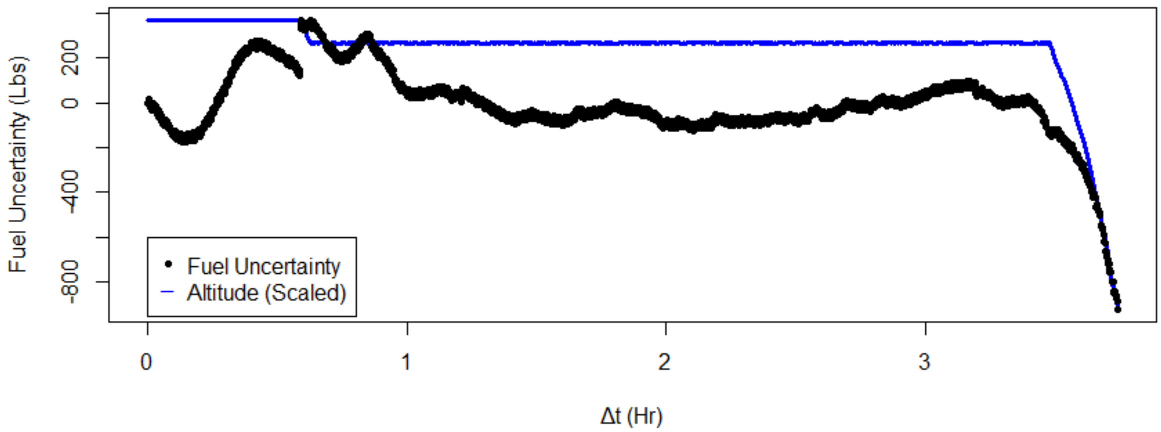


Figure 6: Fuel Uncertainty Plot with Scaled Altitude

3 Methodology

3.1 Modeling Uncertainty

For a given Stage 1 flight, the uncertainty in estimated benefits at each prediction will slowly approach zero as the flight nears the TOD and Δt reaches zero. Pilots vectoring off route and/or unexpected weather conditions will cause the flight's overall duration to increase and the uncertainty to trend downwards. On the other hand, ATC approved trajectory changes and/or early expiration of weather hazards will cause the flight's overall duration to decrease and the uncertainty to trend upwards. Figure 7 shows two example flights of how time uncertainty can change with Δt .

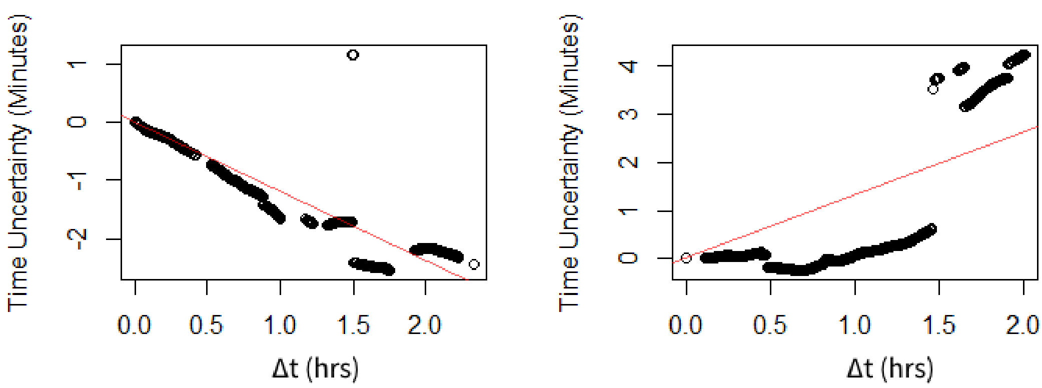


Figure 7: Plots of Uncertainty$_{\text{Time}}$ vs Δt for Two Different Flights.

In general, a description of how uncertainty in estimated benefits approaches zero as Δt decreases can be roughly summarized by a trend line through the origin. Ultimately, a trend line can be computed for each Stage 1 flight creating a distribution of slopes. The uncertainty in estimated benefits for a future flight can then be computed via this distribution of slopes. If the slopes are assumed to follow a normal distribution, the resulting model is known as a linear mixed model [4].

3.1.1 Linear Mixed Model

Linear mixed models are a class of flexible linear models that have widespread use due to their ability to model clustered (i.e., hierarchical) data. Linear mixed models contain both fixed effects and random effects. A fixed effect is an unknown constant that does not vary, whereas a random effect is a parameter that is also a random variable. Random effects are often used to measure variability between clusters of observations [4]. In our case, values of uncertainty in estimated benefits are clustered by flight. Therefore, the use of a linear mixed model allows us to treat a flight as a random sample from a larger population of flights and account for the correlation of uncertainty values within an individual flight.

As stated above, we can model how uncertainty in estimated benefits changes with Δt for each flight via a linear trend line through the origin. Let B be a random variable representing these linear trend lines and let β_f be an observed value of B corresponding to a specific flight. Let B have a mean of β_{Mean} and variance σ_β^2 . In what follows, bolded values represent matrices or vectors. For instance, $\boldsymbol{Uncertainty}|B = \beta_f$ represents a vector of time or fuel uncertainty values for a specific flight f. In general, linear mixed models are defined conditionally on outcomes of their random effects. Thus, the linear mixed model used to model uncertainty in estimated benefits is specified in equation (6):

$$
\begin{gathered}
\mathrm{E}\left[\boldsymbol{Uncertainty}|B = \beta_f\right] = \Delta \boldsymbol{t} \cdot \beta_f \\
Var\left(\boldsymbol{Uncertainty}|B = \beta_f\right) = \boldsymbol{R}_f \\
\mathrm{E}[B] = \beta_{Mean} \\
Var(B) = \sigma_\beta^2
\end{gathered}
\tag{6}
$$

In practice, β_{Mean} is often subtracted from B (i.e., $\tilde{B} = B - \beta_{Mean}$) so that $\mathrm{E}[\tilde{B}] = 0$. This results in the model formulation as shown in equation (7).

$$
\begin{gathered}
\mathrm{E}\left[\boldsymbol{Uncertainty}|\tilde{B} = \tilde{\beta}_f\right] = \Delta \boldsymbol{t} \cdot \beta_{Mean} + \Delta \boldsymbol{t} \cdot \tilde{\beta}_f \\
Var\left(\boldsymbol{Uncertainty}|\tilde{B} = \tilde{\beta}_f\right) = \boldsymbol{R}_f \\
\mathrm{E}[\tilde{B}] = 0 \\
Var\left(\tilde{B}\right) = \sigma_\beta^2
\end{gathered}
\tag{7}
$$

Note that β_{Mean} is a fixed effect shared among all flights and $\tilde{\beta}_f$ is an observed value of the random variable $\tilde{B}$. For any particular flight, the model is much the same as a standard linear regression model. The residual error variance for a particular flight (i.e., $Var(\boldsymbol{Uncertainty}|\tilde{B} = \tilde{\beta}_f)$) is given by the matrix $\boldsymbol{R}_f$, a symmetric matrix with dimension equal to the number of observations for that flight. Instead of assuming observations are independent, the uncertainty in estimated benefit values are assumed to be correlated (i.e., they are not independent (recall Figure 7)). More specifically, observations within a flight are assumed to be correlated based on how far apart they are observed. Observations from different flights are assumed to be uncorrelated. This assumed structure for $\boldsymbol{R}_f$ is known as a CAR(1) covariance structure (i.e., continuous autoregressive with one time lag) and is expressed in Equation (8).

$$\boldsymbol{R}_f = \sigma_e^2 \begin{bmatrix} 1 & \rho^{|t_1-t_2|} & \cdots & \rho^{|t_1-t_n|} \\ \rho^{|t_1-t_2|} & 1 & \cdots & \rho^{|t_2-t_n|} \\ \vdots & \vdots & \ddots & \vdots \\ \rho^{|t_1-t_n|} & \rho^{|t_2-t_n|} & \cdots & 1 \end{bmatrix} \quad (8)$$

Here, σ_e^2 is the constant error variance, and t_i ($i = 1, \dots, n$) is the timestamp of observation i measured in minutes. ρ represents the correlation between two observations one minute apart. From equation (8), we see that two observations from the same flight that are farther away from each other will have less correlation than observations from the same flight that are closer together. Note the choice of time units for t_i and t_j has no impact on the model fit, just the interpretation of ρ. The "one time lag" in CAR(1) means that the current value of uncertainty is assumed to depend on the immediately preceding value. Assuming B is normally distributed, the marginal distribution for an unknown uncertainty value is given in equation (9) [4].

$$Uncertainty \sim N(\Delta t \cdot \beta_{Mean}, \Delta t^2 \cdot \sigma_\beta^2 + \sigma_e^2) \quad (9)$$

This marginal model of the response can be used to generate prediction intervals for uncertainty in estimated benefits for future flights. The details of the different prediction intervals that can be generated using equation (9) can be found in Section 5.1.

3.2 Preparing the Data

Before computing estimates for the linear mixed models for time and fuel uncertainty, observations were removed from the dataset to give each flight an equal number of observations. As all observations from all flights equally contribute to the estimate of the mean slope β_{Mean}, longer flights with more observations will have a larger impact on the estimate of β_{Mean}. This is known as "informative cluster sizes" [4]. By removing observations from the dataset, we force all clusters to have the same number of observations and thus mitigate this potential issue. Heuristics suggest approximately 10 to 20 observations per predictor is enough to precisely estimate all parameters in a linear regression model [5]. With this in mind, taking a representative sample of the TAP predictions from each flight will prevent longer flights from having more influence on the estimate of β_{Mean}. To generate the samples, each flight was divided into 50 bins based on the quantiles of the observed Δt. Then, one observation was randomly selected from each bin. Note that neither the assumed linear relationship of Δt with uncertainty in estimated benefits nor the CAR(1) structure require the observations to be equally spaced by Δt. We chose to use 50 bins to guarantee enough observations to estimate β_{Mean}, the individual slopes ($\tilde{\beta}_f$), the autocorrelation coefficient (ρ), and σ_e^2 while keeping the model computationally tractable.

4 Analysis

Estimates of model parameters were calculated using Restricted Maximum Likelihood (REML) [4]. REML is often the default choice for fitting linear mixed models in many statistical software suites and has many nice properties including asymptotic normality of parameter estimates and less sensitivity to outliers compared to traditional maximum likelihood [4], [6]–[8]. The linear mixed models for time and fuel uncertainty were fitted using the *nlme* R package [9]. In what follows, we refer to the linear mixed models for time and fuel uncertainty simply as the model for time uncertainty and the model for fuel uncertainty, respectively.

4.1 Analysis of Fitted Parameters

4.1.1 Fitted Parameters for Time Uncertainty

The fitted values and 95% confidence intervals for the model for time uncertainty are listed in Table 1. For this model, β_{Mean} was not statistically significantly different from 0 ($p\text{-value} = 0.23$, standard error $=$ 0.17). Recalculating the parameter estimates after removing β_{Mean} from the model changes their values very slightly. Note that the estimate of σ_β^2 is almost twice that of σ_e^2. The estimate of ρ (0.93) indicates that two observations observed one minute apart are strongly correlated. This accounts for much of the deviation from the linear trend.

Table 1: Time Uncertainty Parameter Estimates

Parameter	Full Model Estimate	Reduced Model Estimate	Full Model 95% CI
$\hat{\beta}_{Mean}$	-0.21	---	(-0.54, 0.13)
$\widehat{\sigma_\beta^2}$	2.92	2.93	(2.13, 4.01)
$\widehat{\sigma_e^2}$	1.61	1.61	(1.46,1.78)
$\hat{\rho}$	0.93	0.93	(0.92,0.94)

4.1.2 Fitted Parameters for Fuel Uncertainty

The model for fuel uncertainty yields the estimates and intervals given in Table 2. As expected, a mean slope of -122.63 indicates significant bias in fuel uncertainty ($p\text{-value} < 0.0001$, standard error $=$

15.70). For every minute of difference between the two predictions' timestamps, the estimated fuel benefit is on average underestimated by 122.63 lbs of fuel. Similar to the uncertainty in estimated time benefit model, autocorrelation is extremely high.

Table 2: Fuel Uncertainty Parameter Estimates

Parameter	Estimate	95% CI
$\hat{\beta}_{Mean}$	-122.63	(-153.43, -91.84)
$\widehat{\sigma_{\beta}^{2}}$	13203.50	(8399.51, 20755.07)
$\widehat{\sigma_{e}^{2}}$	64778.14	(56767.02, 73919.82)
$\hat{\rho}$	0.97	(0.96,0.97)

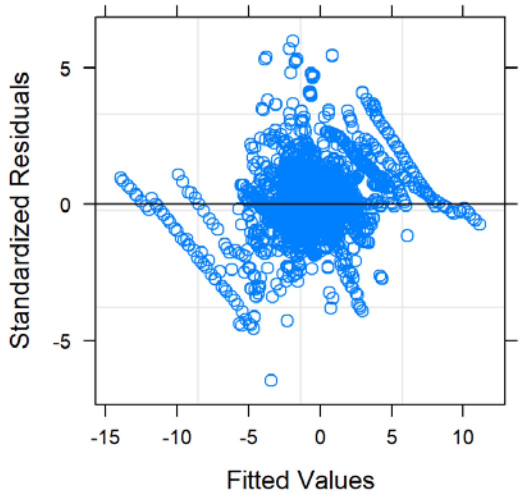
4.2 Checking Model Assumptions

The linear mixed model requires four assumptions: normally distributed random errors (normality), errors that have the same variance (homoscedasticity), a linear model is correct (linearity), and normally distributed random effects. This section discusses diagnosing violations of these assumptions, the consequences of their violation, and possible solutions to violation. The assumptions of homoscedasticity, linearity, and normality can be checked using various diagnostic plots as seen in the following sections.

4.2.1 Model for Time Uncertainty Assumptions

The assumption of linearity can be checked with a residual versus fitted value plot. If the fitted model is correct, the mean of the residuals should be approximately constant at zero for all fitted values. Departures from zero indicate areas where the model fits poorly. The assumption of normality for both the residuals and the random effects can be checked using normal quantile-quantile plots. Departures from the linear trend indicate departures from normality.

For the model for time uncertainty, assumptions of linearity/homoscedasticity and normality of the residuals are examined in Figure 8a and Figure 8b. Figure 8a does not indicate any unusual patterns in the residuals and thus we would conclude that the assumptions of linearity and homoscedasticity are reasonable.



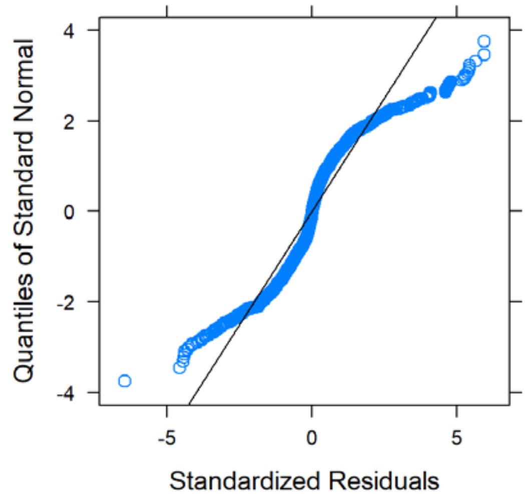


Figure 8a: Time Uncertainty Residual vs Fitted Plot.

Figure 8b: Time Uncertainty Residual Normal Quantile-Quantile Plot.

The normal quantile-quantile plot in Figure 8b indicates the residuals are more peaked with fatter tails than a normal distribution. However, given the large sample size of this data, misspecification of the distribution of the random errors should have minimal impact—especially since the departure from normality is not severe [10]. With this in mind, we proceed with the analysis of the fitted model.

The assumption of normality for the random effects is also tested via the normal quantile-quantile plot in Figure 9. With the exception of a few outliers, the distribution of the random effects looks sufficiently normal.

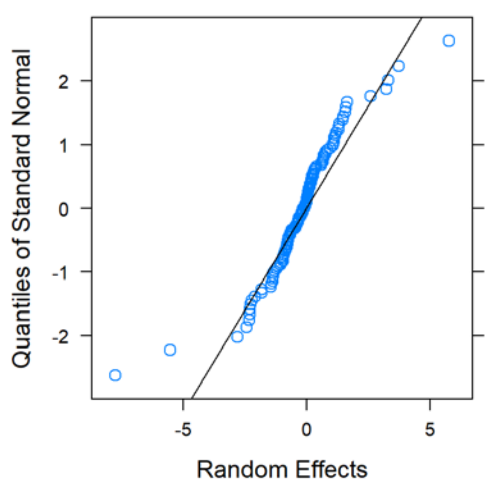


Figure 9: Normal Quantile-Quantile Plot of Fitted Slopes for Model for Time Uncertainty.

The assumption of homoscedasticity of the random effects is examined via the flight length versus fitted slope plot in Figure 10. Two clear groups of outliers can be seen. The negative outliers (shown in blue in

Figure 10) were short flights along the West Coast which were vectored due to congestion at their destination. This was determined by examining changes in the filed routes. These vectored flights do not generalize well to the longer flights as time deviations due to vectoring do not depend on flight length or Δt. Positive outliers (shown in red in Figure 10) occurred in cases of unexpected time savings. In most cases, we suspect this was due to expiring weather advisories. This was determined by cross-referencing the filed routes with historical radar data [11]. First, the filed route would avoid an area of convective weather. Once the aircraft got close to this deviation from the normal route, the convective weather would have cleared, and the pilot would return to the more optimal route. These outliers aside, variability in the random effects appears consistent for flights of all recorded lengths.

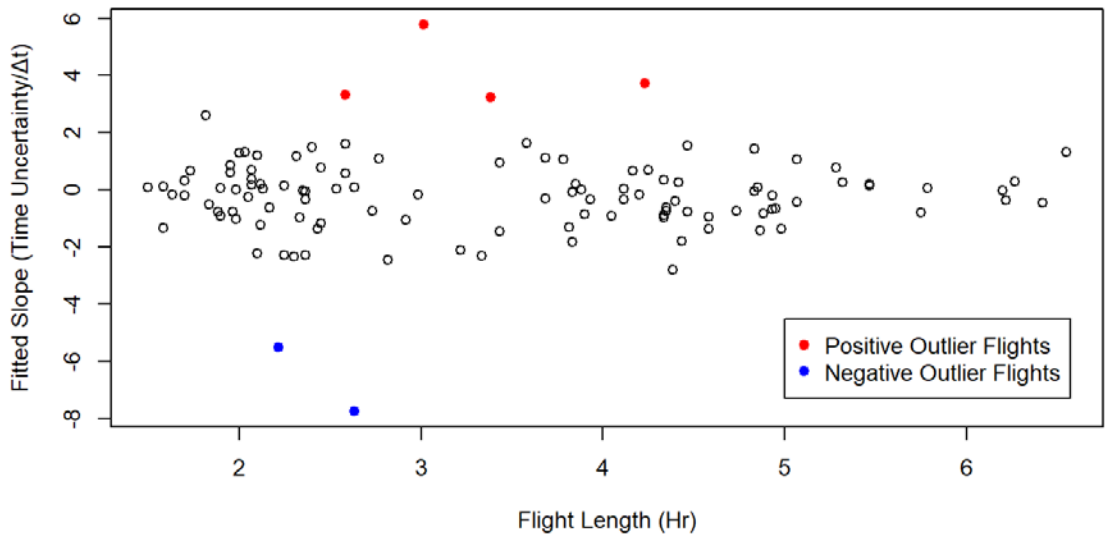


Figure 10: Fitted Slopes vs Flight Length Plot for Model for Time Uncertainty.

4.2.2 Model for Fuel Uncertainty Assumptions

Based on the residuals versus fitted plot in Figure 11b and the normal quantile-quantile plot in Figure 11b, both the assumptions of homoscedasticity and normality appear reasonable (although we note slight departure from normality in Figure 11b).

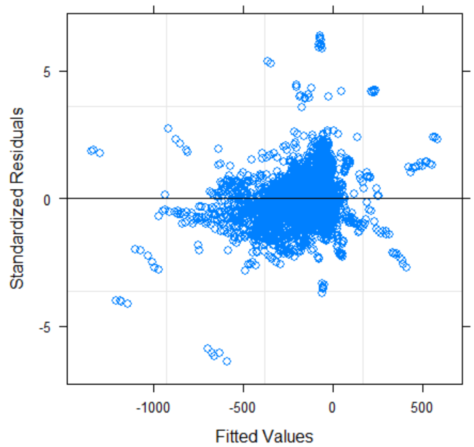


Figure 11a: Fuel Uncertainty Residual vs Fitted

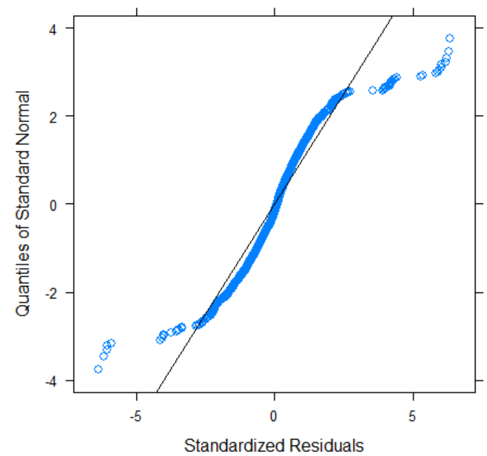


Figure 11b: Fuel Uncertainty Residual Normal Quantile-Quantile Plot

The assumptions of normality for the random effects are tested and confirmed by the normal quantile-quantile plot in Figure 12a. The Fitted Slopes vs Flight Length plot in Figure 12b shows outliers similar to the uncertainty in estimated time benefits case, but which observations are outliers is wholly different.

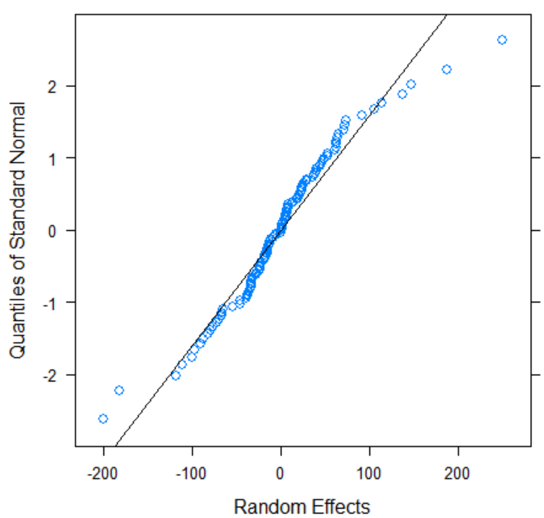


Figure 12a: Normal Quantile-Quantile Plot of Fitted Slopes for Model for Fuel Uncertainty.

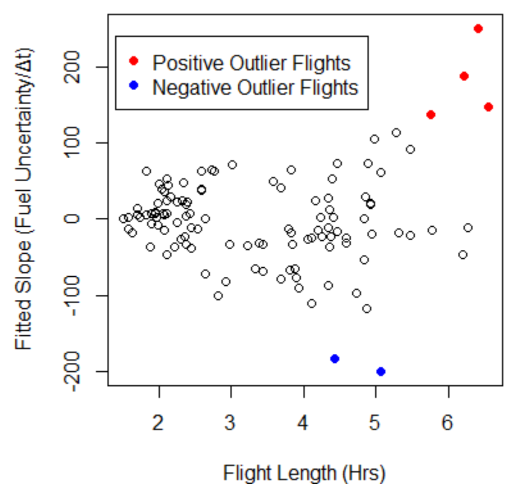


Figure 12b: Fitted Slopes vs Flight Length Plot for Model for Fuel Uncertainty.

Positive fuel uncertainty outliers are all transcontinental flights with unexpected fuel costs, likely due to an altitude change. Negative fuel uncertainty outliers likely occurred for similar reasons as the outliers in Figure 10.

4.3 Model Fit

Variance explained by a linear mixed model can be partitioned into three components: variance explained by the fixed effects, variance explained by the random effects, and variance due to random error. Using these partitions, a statistic similar to the usual R^2 statistic can be calculated via the implementation in the *MuMin* R package [11]. For more details, see [10].

For the model for time uncertainty, the R^2 statistic is 0.78 indicating the model explains a large portion of the variation in uncertainty. Time uncertainty likely follows a more complex pattern than a linear trendline. Additionally, this model does not explicitly account for different sources of variability such as weather hazards and pilots vectoring off route.

R^2 for the model for fuel uncertainty was calculated to be only 0.37 indicating the model explains less than half of the variation in fuel uncertainty. We conjecture that this poorer fit is due to the previously mentioned bias in the fuel estimates. Given the bias in fuel uncertainty is not properly addressed, we currently consider the model for fuel uncertainty unfit for use.

5 Application

The primary goal of analyzing time uncertainty and fuel uncertainty is to quantify the uncertainty in the estimated TAP benefits calculations outlined in Section 2.1. By specifying a certain level of confidence, we can develop a prediction interval for use in estimated time benefit and estimated fuel benefit calculations. We develop this idea in the following section and then apply the prediction interval in an estimated benefits analysis for a single flight and for aggregate flights.

5.1 Prediction Intervals

Recall for a given Stage 2 or Stage 3 flight, the estimated benefit is calculated via equation (2). Again, note negative benefit indicates savings with respect to time or fuel. Following the method used in the *predictInterval* package for R, we assume the estimate of the random effect variance to be constant without error [12]. Following the form of the marginal model in equation (9), we can compute the prediction interval for time uncertainty by replacing the true values with their estimates and accounting for the standard error of $\hat{\beta}_{Mean}$.

$$\begin{aligned} &\hat{\beta}_{Mean}\Delta t \pm z_{\alpha/2} \cdot \sqrt{\Delta t^2 \cdot \left(\widehat{\sigma_\beta^2} + \sigma^2_{\hat{\beta}_{Mean}}\right) + \widehat{\sigma_e^2}} \\ &= 0 \cdot \Delta t \pm z_{\alpha/2} \cdot \sqrt{\Delta t^2 \cdot (2.92 + 0) + 1.61} \\ &= \pm z_{\alpha/2} \cdot \sqrt{2.92 \cdot \Delta t^2 + 1.61} \end{aligned} \tag{10}$$

Here, Δt is the Δt value of a flight of interest, $z_{\alpha/2}$ is the $\left(1 - {}^{\alpha}/_{2}\right)$ quantile of the standard normal distribution, and $\sigma^2_{\hat{\beta}_{Mean}}$ is the standard error of $\hat{\beta}_{Mean}$. Note $\sigma^2_{\hat{\beta}_{Mean}} = 0$ since β_{Mean} was removed from the model in Section 4.1.1. Figure 13 shows how the 95% and 90% prediction intervals change as Δt increases.

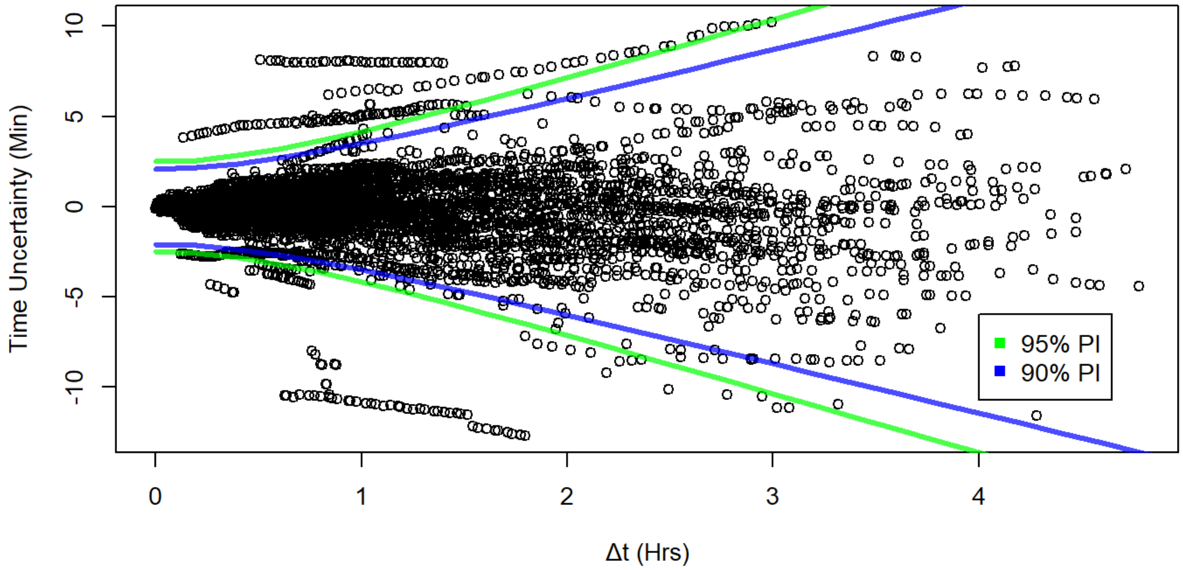


Figure 13: Time Uncertainty Data with Prediction Intervals

The prediction interval for an aggregate of n flights is derived similarly.

$$\hat{\beta}_{Mean} \sum_{i=1}^{n} \Delta t_i \pm z_{\alpha/2} \cdot \sqrt{\sum_{i=1}^{n} \Delta t_i^2 \cdot \left(\widehat{\sigma_\beta^2} + \sigma_{\beta_{Mean}}^2\right) + n \cdot \widehat{\sigma_e^2}}$$
$$= \pm z_{\alpha/2} \cdot \sqrt{2.92 \cdot \sum_{i=1}^{n} \Delta t_i^2 + n \cdot 1.61} \qquad (11)$$

5.2 Estimated Benefits Analysis Example

Table 3 contains the estimated time benefit in minutes (min) and seconds (s) for three different flights along with the associated Δt values of the predictions.

Table 3: Estimated Time Benefit Example Dataset

Flight Number	Observed Time Benefit	Δt	95% Prediction Interval for Time Benefit
FLIGHT1	-9 min 26 s	3 hr 10 min 26 s = 3.17 hr	(-20 min 7 s , 1 min 15 s)
FLIGHT2	-2 min 41 s	1 hr 36 min 53 s = 1.61 hr	(-8 min 7 s , 2 min 45 s)
FLIGHT3	-0 min 28 s	0 hr 32 min 30 s = 0.54 hr	(-2 min 17 s , 1 min 21 s)

5.2.1 Single Flight Benefit

To calculate a 95% prediction interval for the time benefit for FLIGHT1, we apply equation (11) to generate the prediction interval for time uncertainty. For $\alpha = 0.05$, $z_{\alpha/2} = 1.96$. Then,

$$\pm z_{\alpha/2} \cdot \sqrt{2.92 \cdot \Delta t^2 + 1.61} = \pm 1.96 \cdot \sqrt{2.92 \cdot 3.17^2 + 1.61} = \pm 10 \text{ min } 54 \text{ s.}$$

We then subtract the time uncertainty interval from the estimated benefit to generate the prediction interval for true time benefit for FLIGHT1.

$$-9 \text{ min } 26 \text{ s} \pm 10 \text{ min } 54 \text{ s} = (-20 \text{ min } 20 \text{ s}, 1 \text{ min } 28 \text{ s}).$$

For FLIGHT1 we can say with 95% confidence that TAP saved between -1 min 28 s and 20 min 20 s of time.

5.2.2 Aggregate Flights Benefit

To calculate a 95% prediction interval for the combined estimated time benefit for FLIGHT1, FLIGHT2, and FLIGHT3, we apply equation (11). Again, for $\alpha = 0.05$, $z_{\alpha/2} = 1.96$. We have three flights, so $n = 3$.

$$\pm z_{\alpha/2} \cdot \sqrt{2.92 \cdot \sum_{i=1}^{n} \Delta t_i^2 + n \cdot 1.61} = \pm 1.96 \cdot \sqrt{2.92 \cdot 12.93 + 4.83}$$
$$= \pm 12 \text{ min } 47 \text{ s}$$

We then subtract the time uncertainty from the total observed time benefit to generate the prediction interval for true time benefit for FLIGHT1.

$$[-9 \text{ min } 26 \text{ s} - 2 \text{ min } 41 \text{ s} - 28 \text{ s}] \pm 12 \text{ min } 47 \text{ s}$$
$$= -12 \text{ min } 35 \text{ s} \pm 12 \text{ min } 47 \text{ s} = (-24 \text{ min } 43 \text{ s}, 0 \text{ min } 12 \text{ s})$$

We can say with 95% confidence the total time savings from using TAP across all three flights is between -12 s and 24 min 43 s .

6 Conclusion

The TAP software developed by NASA calculates optimal flight routes in real time using live traffic and weather data. The Stage 1 series of flight trials by Alaska Airlines consisted of TAP making optimal flight route suggestions and flight time predictions on board routine flights. Our research attempted to quantify the uncertainty in estimated time and fuel benefits analysis using Stage 1 data. The linear mixed model offers a relatively simple way to model the clustered, longitudinal time and fuel uncertainty data observed in the Stage 1 flight trials. We find the model for time uncertainty to fit well. While analysis of estimated time benefit for a single flight is unlikely to produce significant results, aggregating flights greatly improves the precision of the estimated time benefits analysis. On the other hand, the model for fuel uncertainty fails to capture a majority of the variation largely due to biased data. Additionally, the fixed slope effect of this model fails to capture the stepwise nature of the fuel bias. Recalculating the model for fuel uncertainty using unbiased data should provide for a usable model.

Future work incorporating effects such as weather and pilot vectoring in the model definition may be able to account for significantly more variability while maintaining the simplicity of the linear mixed model. This is the most practical direction to take future work. More complex models are more interesting from a statistical point of view but will lose the interpretability of a simple linear mixed model.

7 Works Cited

[1] Federal Aviation Administration, “FAA OPSNET Database: Delays by Cause,” 01-Jan-2008. [Online]. Available: https://aspm.faa.gov/opsnet/sys/Delays.asp. [Accessed: 02-Mar-2018].

[2] S. Woods, R. A. Vivona, D. J. Wing, and K. A. Burke, “Traffic Aware Planner for Cockpit-based Trajectory Optimization,” in *16th AIAA Aviation Technology, Integration, and Operations Conference*, 0 vols., American Institute of Aeronautics and Astronautics, 2016.

[3] D. J. B. Wing, “Initial Implementation and Operational Use of TASAR in Alaska Airlines Flight Operations,” presented at the AVIATION 2018, Atlanta, GA, United States, 2018.

[4] C. E. McCulloch and S. R. Searle, *Generalized, Linear, and Mixed Models*. John Wiley & Sons, 2004.

[5] F. E. Harrell, “Regression modeling strategies,” *Implemented R Package ‘rms’ Version*, vol. 3, no. 3, 2014.

[6] “SPSS Mixed Command.” [Online]. Available: https://stats.idre.ucla.edu/spss/seminars/spss-mixed-command/. [Accessed: 17-Sep-2018].

[7] SAS Institute, “PROC MIXED: The MIXED Procedure :: SAS/STAT(R) 9.2 User’s Guide, Second Edition.” [Online]. Available: https://support.sas.com/documentation/cdl/en/statug/63033/HTML/default/viewer.htm#mixed_toc.htm. [Accessed: 17-Sep-2018].

[8] D. Bates, M. Mächler, B. Bolker, and S. Walker, “Fitting Linear Mixed-Effects Models using lme4,” *ArXiv Prepr. ArXiv14065823*, Jun. 2014.

[9] “CRAN - Package nlme.” [Online]. Available: https://cran.r-project.org/web/packages/nlme/index.html. [Accessed: 03-Aug-2019].

[10] H. Jacqmin-Gadda, S. Sibillot, C. Proust, J.-M. Molina, and R. Thiébaut, “Robustness of the linear mixed model to misspecified error distribution,” *Comput. Stat. Data Anal.*, vol. 51, no. 10, pp. 5142–5154, Jun. 2007.

[11] National Centers for Environmental Information, “Radar Data Map | GIS Maps,” *Radar data map, National Reflectivity Mosaic*. [Online]. Available: https://gis.ncdc.noaa.gov/maps/ncei/radar. [Accessed: 03-Jul-2019].

[12] “Prediction Intervals from merMod Objects.” [Online]. Available: https://cran.r-project.org/web/packages/merTools/vignettes/Using_predictInterval.html. [Accessed: 13-Nov-2018].

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
1-11-2019

2. REPORT TYPE
Technical Memorandum

3. DATES COVERED *(From - To)*

4. TITLE AND SUBTITLE
Modeling Uncertainty in Time and Fuel Benefit Estimation for TASAR Operational Evaluation

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER
330693.04.30.07.06

6. AUTHOR(S)
Burris, Joseph; Wilson, Sara R.; Edwards, David J.; Ballard, Kathryn M.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
NASA Langley Research Center
Hampton, VA 23681-2199

8. PERFORMING ORGANIZATION REPORT NUMBER
L-21073

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
National Aeronautics and Space Administration
Washington, DC 20546-0001

10. SPONSOR/MONITOR'S ACRONYM(S)
NASA

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
NASA-TM-2019-220420

12. DISTRIBUTION/AVAILABILITY STATEMENT
Unclassified-
Subject Category 65
Availability: NASA STI Program (757) 864-9658

13. SUPPLEMENTARY NOTES

14. ABSTRACT The Traffic Aware Strategic Aircrew Requests concept aims to reduce weather-induced delays, improve route efficiency, and efficiently share route modification options by combining onboard avionics data, Automatic Dependent Surveillance-Broadcast data, and broadband internet data to generate optimal, traffic-compatible trajectory changes based on real-time traffic and weather data. Time and fuel benefits due to use of the Traffic Aware Planner (TAP) software can be estimated by taking the difference in predicted flight time and fuel usage before and after a TAP-inspired trajectory change is completed. Although TAP's optimization algorithm predicts flight and fuel usage based on the current flight route and weather data, it does not account for possible air traffic controller-initiated trajectory changes, reroutes due to sudden weather changes, or other pilot/controller actions that may occur during flight. This paper introduces an approach for quantifying the uncertainty in estimated time and fuel benefits.

15. SUBJECT TERMS
TASAR; Traffic Aware Planner; Linear mixed model; Uncertainty estimation

16. SECURITY CLASSIFICATION OF: a. REPORT	b. ABSTRACT	c. THIS PAGE	17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES
U	U	U	UU	29

19a. NAME OF RESPONSIBLE PERSON
STI Help Desk (email: help@sti.nasa.gov)

19b. TELEPHONE NUMBER *(Include area code)*
(757) 864-9658