Thirteenth USA/Europe Air Traffic Management Research and Development Seminar (ATM2019)

Using Machine-Learning to Dynamically Generate Operationally Acceptable Strategic Reroute Options

Antony D. Evans, Ph.D.
Crown Consulting, Inc.
Moffett Field, CA, USA
antony.d.evans@nasa.gov

Paul U. Lee, Ph.D.
Human-Systems Integration Division
NASA Ames Research Center, Moffett Field, CA, USA
paul.u.lee@nasa.gov

Abstract— **The newly developed Trajectory Option Set (TOS), a preference-weighted set of alternative routes submitted by flight operators, is a capability in the U.S. traffic flow management system that enables automated trajectory negotiation between flight operators and Air Navigation Service Providers. The objective of this paper is to describe and demonstrate an approach for automatically generating pre-departure and airborne TOSs that have a high probability of operational acceptance. The approach uses hierarchical clustering of historical route data to identify route candidates. The probability of operational acceptance is then estimated using predictors trained on historical flight plan amendment data using supervised machine learning algorithms, allowing the routes with highest probability of operational acceptance to be selected for the TOS. Features used describe historical route usage, difference in flight time and downstream demand to capacity imbalance. A random forest was found to be the best performing algorithm for learning operational acceptability, with a model accuracy of 0.96. The approach is demonstrated for an historical pre-departure flight from Dallas/Fort Worth International Airport to Newark Liberty International Airport.**

Keywords- Air traffic management; trajectory option set; machine learning.

I. INTRODUCTION

A new Traffic Management Initiative (TMI) called the Collaborative Trajectory Options Program (CTOP) supports a more complex characterization of the reduced capacity problem than was previously possible in the U.S. National Airspace System (NAS), using multiple constraints, called Flow Constrained Areas (FCAs).[1] CTOP allows flight operators to submit a preference-weighted set of alternative routes through and around the FCAs called a Trajectory Option Set (TOS), from which the program can select. In the development of the CTOP concept, it was envisioned that TOSs would not only be generated for flights pre-departure, but also for airborne flights [1]. This would provide the U.S. Federal Aviation Administration (FAA) with user preferred routing alternatives from which to assign strategic airborne reroutes, in the event that demand or capacity constraints on downstream FCAs changed significantly, requiring adjustments to the air traffic flows through the downstream FCAs. Such a capability would allow flight operator preferences to be accommodated in assigning strategic airborne reroutes, enabling increased trajectory negotiation between the FAA and flight operators, which is a key component of the future air traffic control system envisioned by the U.S. National Aeronautics and Space Administration (NASA) [2]. It may also reduce congestion and increase throughput by distributing traffic across the available airspace to a greater extent than is currently done with strategic advisory reroutes or playbook routes that allocate impacted traffic to a small number of routes. Finally, it may also increase predictability for airborne flights, because TOS routes could be chosen to avoid constraints, resulting in less need for intervention from Air Traffic Control (ATC) during the flight. This is, however, contingent on the routes being operationally acceptable. For this paper, operationally acceptance refers to ATC being willing to implement the trajectory as a flight plan amendment, if requested, given the conditions (including downstream) at the time.

While CTOP is currently operational, it is not being used extensively. This is because relatively few flight operators have, as yet, invested in capabilities to generate TOSs beyond using Coded Departure Routes (CDRs) – formalized alternate routes by origin-destination pair provided by the FAA. This is in part because the business case is unclear, with TOS generation potentially increasing workload on dispatchers. Despite these challenges, some tools are under development by flight operators and third parties to aid in pre-departure TOS generation [3].

Most Air Route Traffic Control Centers (ARTCCs, or Centers) do not have a specific functionality to ensure that trajectories proposed pre-departure are operationally acceptable. However, if a filed trajectory is found to be operationally unacceptable while the flight is enroute, the trajectory would be amended tactically. Generating pre-departure TOSs that have a higher probability of operational acceptance would increase the predictability of routing and timing of flights, as these trajectories would have a higher

[1] An FCA is a volume of airspace, along with flight filters and a time interval, used to identify flights. It may be drawn graphically, around weather, or may be based on a NAS element. FCAs are used to evaluate demand on a resource, which may be constrained.

likelihood of being flown without amendment. In contrast to pre-departure reroute requests, airborne reroute requests are immediately reviewed by controllers, and, under certain conditions, by traffic managers. They would therefore be immediately rejected or amended were they found to be operationally unacceptable. For trajectory negotiation using an airborne TOS, this could significantly increase controller and Traffic Management Unit (TMU) workload. It would therefore be desirable for flight operators to automatically generate airborne TOSs that have a high probability of operational acceptance. The objective of this paper is to describe and demonstrate an approach for automatically generating TOSs, both pre-departure and airborne, that have high probability of operational acceptance for implementation as strategic reroutes, given the conditions at the time the TOS is generated. These operationally acceptable TOSs are generated using hierarchical clustering of historical route data to identify route candidates. The probability of operational acceptance is then estimated using predictors trained on historical flight plan amendment data using supervised machine learning algorithms, allowing the routes with highest probability of operational acceptance to be selected for the TOS by the flight operator.

Background literature is presented in Section II. The approach for generating operationally acceptable TOSs is described in Section III, followed by a sample application in Section IV. Section V discusses implications of the results, before conclusions and recommendations for future work are presented in Section VI.

II. Background

A number of tools and concepts are under development by NASA to generate airborne reroute advisories, particularly around weather. Tools under development include the NAS Constraint Evaluation and Notification Tool (NASCENT) [4], which extends the Dynamic Weather Routes (DWR) concept [5,6]; Multi-Flight Common Routes (MFCR) [7]; Dynamic Routes for Arrivals in Weather (DRAW) [8]; and Traffic Aware Strategic Aircrew Requests (TASAR) [9,10]. While these tools generate advisory reroutes for airborne flights to avoid near term constraints, particularly weather, these reroutes are typically tactical in nature (with a look-ahead of 60 to 90 minutes). The tools are not tailored to generate strategic reroutes (with a look-ahead of greater than 90 minutes) across multiple Centers, and are not designed to support TOS generation, whereby routes must avoid or intersect specified FCAs, sometimes far downstream. However, some of these tools do consider operational acceptability in the generation of reroute advisories. The TASAR concept [9] incorporates traffic, weather, and airspace information in the optimization of in-flight trajectory re-planning, increasing the likelihood of the resulting trajectory change request being operationally acceptable. NASCENT [4] incorporates historical usage data, similar to the approach described in [11], improving operational acceptance. Other algorithms and models have also been developed elsewhere to reroute traffic around constraints, e.g., using an autonomous agent approach [12], or by optimizing the traffic flow management problem [13,14,15].

Some tools for generating pre-departure TOSs are under development by flight operators and third parties. These typically apply existing techniques developed for flight planning systems to identify wind optimal routings through and around specified FCAs. These may be constrained to use an underlying "clearable route network" that is considered operationally acceptable, calculated based on historical usage of flight plan segments, such as that described in [3]. The approaches used for such pre-departure TOS generation may be adaptable to airborne TOS generation.

Reference [16] describes an approach that dynamically creates optimized flight specific reroutes to aid traffic managers in efficiently maneuvering flights. It is specifically designed for situations in which weather requires traffic managers to reroute flights that plan to pass through the weather, while balancing demand through sectors with reduced capacity or increased traffic volume (resulting from other flights deviating from their original routes). Routes are optimized using simulated annealing, given an operationally acceptable routing network to which they must conform.

In generating the optimized reroutes in [16], a number of factors are considered, including route deviation distance, conformance of the reroute to historically flown routes, weather impact on the current route, sector congestion, and factors including required point-outs[2] and inter-facility coordination. The routing network used for the optimization was generated by segmenting historically flown routes into fix-pair segments. Thus, all arcs in the modeled network consist of previously-flown connections between fixes, so each individual arc in the network has some level, depending on usage, of operational acceptability. Reroutes are constructed from these arcs using the optimization algorithm, and the reroutes that best meet a set of metrics of operational acceptability are presented as potential alternatives to users [17]. The Advanced Flight Specific Trajectory (AFST) tool developed by MITRE [18,19] incorporates many of the capabilities described in [16] and [17].

The capabilities described above generate optimized trajectories that comply with an underlying routing network that has some degree of operational acceptability, and therefore have applications in TOS generation. Historical flight plan and flight plan amendment data is used in the generation of the underlying routing network, while dynamic conditions impacting a flight, such as downstream demand and capacity, are accounted for in the optimization of the route. An alternative approach is to use supervised machine learning algorithms to train predictors in the operational acceptability of trajectories, based on historical usage as well as dynamic conditions impacting the flight. This allows routes to be generated based explicitly on amendments that traffic

[2] A point-out refers to the need for one controller to request that the controller of an adjacent sector also monitors a flight that is close to the sector boundary.

managers have issued in the past. The approach may therefore capture nuances in the way traffic managers allocate routes that are not captured by the list of factors explicitly considered in an optimization. Using machine learning may therefore increase the operational acceptability of the chosen routes. This alternative approach is the focus of the present paper. Integration with optimization approaches such as those described in [16] and those used in some commercial TOS generators may be explored in future research.

The present paper extends past NASA work on predicting route operational acceptability. Reference [11] analyzed the historical usage of different flight routings in order to improve route acceptance for the NASA developed DWR algorithm [5,6]. The results suggest that historical usage is a key requirement for a route's acceptance by ATC, but that requesting a reroute that was observed in historical data does not guarantee ATC acceptance. Reference [20] extended this work to develop a predictor of a proposed route's operational acceptability based on a number of features, including historical usage, demand to capacity ratio in the sector in which the maneuver was to begin, how close to hand-off the flight was at the time of the request, whether the proposed route was direct or not, and how long the reroute was. The predictor was trained on data from a DWR trial at American Airlines [5], so is specific to tactical reroute requests from the pilot, which, unlike airborne TOSs, are not typically coordinated through the TMU.

Further relevant NASA research includes [21] and [22], which take initial steps towards providing recommendations of available strategic routing options in response to convective weather, by examining historical data to determine which previous reroute options were used in similar weather and traffic conditions. Dominant routing structures were identified using hierarchical clustering, and methods were described to extract relevant features from the large volume of weather data to quantify the convective weather scenario during a particular time range.

III. Approach

This section describes an approach for a flight operator to dynamically generate lateral trajectories for a TOS (pre-departure or airborne) that have high probability of operational acceptance by ATC. The specification of altitude profiles for each trajectory is left for future work, as is the estimation of relative trajectory preference weightings. The approach is to use clustering of historical route amendments as a static foundation for the operationally acceptable routes, and then to use a model trained on historical amendment data using machine learning to predict the operational acceptability of the routes in that set. This allows the TOS to be selected based on probability of operational acceptance. The approach comprises four steps, as follows:

A. Identify the full set of available trajectory options from the origin airport or sector to the destination airport, based on historical flight plan amendments;

B. Down-select the available trajectory options using route clustering, to define a set of geographically distinct route options;

C. Use machine learning algorithms trained on historical flight plan amendments from across the NAS, including static and dynamic features, to predict the operational acceptance of the down-selected trajectory options; and

D. Select the TOS based on the location of FCAs and the probability of trajectory acceptance by ATC.

The approach followed for each of these steps is described in detail below. A sample application is described in the Section IV.

A. Identify Available Trajectory Options

The approach used for identifying available trajectory options for any flight is similar to that described in [11], which defined routing alternatives based on all historical flight plans and flight plan amendments from the flight's maneuver start point to the destination. For the present paper, historical flight plans and amendments were extracted from historical Aircraft Situation Display to Industry (ASDI) flight data for the period from April to June 2015. This period was used because it is close in time to the sample application used, so trajectories are likely to use the same waypoints, and because it includes significant convective weather activity, like the sample application. The period is therefore likely to contain a diverse set of trajectory options, increasing the likelihood that options relevant to any specific convective weather scenario are discovered. All flight plans and flight plan amendments, from the point at which the amendment was made in each case, were recorded from the historical data, and used to generate a table of unique trajectories. A full set of available trajectory options can be extracted from this table given any flight's maneuver start point and destination. For pre-departure TOS generation, the maneuver start point was the flight's origin airport, while for airborne TOS generation it was the sector in which the aircraft was located at the start of the flight plan amendment. In all TOS generation cases, the end of the maneuver was defined as the destination airport.

B. Down-Select Trajectory Options using Clustering

The full set of unique historical trajectories was down selected to a set of geographically distinct route options using clustering. The trajectories with the highest historical usage in each route cluster were chosen to define the down selected route options. Clustering was required to ensure that the trajectory options considered were suitably spaced geographically. References [23] and [24] describe the methods used, which apply hierarchical clustering with the dissimilarity metric between routes calculated as the Euclidean distance between routes, with each route defined by a fixed number (N=200) of evenly spaced points, as in (1).

$$d_{ij} = \sqrt{\sum_{n=1}^{N}\left(x_{in} - x_{jn}\right)^2 + \left(y_{in} - y_{jn}\right)^2} \quad (1)$$

where d is the dissimilarity metric between trajectories i and j; and x_{in} and y_{in} are the Lambert conformal projection coordinates of the n'th point on trajectory i.

The number of clusters was chosen by maximizing average Silhouette score $\bar{S}$, defined in (2) [23].

$$\bar{S} = \frac{1}{N_r}\sum_{1}^{N_r}\frac{b_i - a_i}{max\{a_i, b_i\}} \quad (2)$$

where a_i is the average Euclidean distance between trajectory i and all other trajectories within the same cluster; b_i is the smallest average Euclidean distance of trajectory i to all trajectories in any other cluster, of which trajectory i is not a member; and N_r is the number of trajectories. To ensure that a sufficiently large number of trajectory options was evaluated, a minimum number of clusters was also set to 15. This could be adjusted in future work.

Reference [20] showed that high historical usage is key to operational acceptance by ATC. For this reason, trajectory options were identified from each route cluster based on historical usage. This required that historical usage was quantified for each historical route. All flight plans and flight plan amendments in the historical ASDI data from April to June 2015 were tracked as they were used to generate the table of unique trajectories described in Section III-A. This allowed a count to be generated of how often each trajectory was used in the historical data, similar to the approach used in [11] and [20] (described in Section II). Also recorded was how often the trajectories were used as amendments specifically, as opposed to an original flight plan, since an amendment matches the application in this paper, which identifies reroute options from the original flight plan.

Because amendments are rarely repeated exactly, historical counts for many full trajectories were found to be very low. Hence a table was also generated of waypoint pair counts, in which counts of every waypoint pair used in the historical flight plan and flight plan amendment data from April to June 2015 was recorded. By identifying the waypoint pairs in any trajectory, a series of waypoint pair counts could then be identified. An alternative count was then generated for each trajectory, calculated as the minimum of all the waypoint pair counts for the trajectory. A similar count was also generated for historical use in flight plan amendments only. In summary, four historical counts were generated for each trajectory from the historical data:

1. A count of the full trajectory as a flight plan or flight plan amendment;
2. A count of the full trajectory as a flight plan amendment only;
3. A minimum of all the waypoint pair counts for the trajectory, with waypoint pairs counted for both flight plans and flight plan amendments; and
4. A minimum of all the waypoint pair counts for the trajectory, with waypoint pairs counted for flight plan amendments only.

All four of these historical counts were used as features for learning route operational acceptability. However, for down-selecting the trajectory options, only the trajectory with highest minimum waypoint pair count (3 in the list above) in each cluster was extracted as a trajectory option for further analysis.

C. *Predict Operational Acceptability using Machine Learning*

The approach used in [20] (described in Section II) was adapted to build a model that predicts the operational acceptability of the most commonly flown route in each trajectory option cluster, based on a number of features describing its historical usage, downstream demand to capacity imbalance, and increase in flight duration relative to the original flight plan. For this paper, a single model was developed for the whole NAS.

Training data is required to develop a predictor of operational acceptability. Training data was generated by extracting appropriate flight plan amendments from historical data from July to September 2015. This represents a late summer period when there was significant convective weather activity, but does not overlap with the period used to identify historical usage. Flight plan amendments are made in response to both strategic decisions by the TMU and more tactical decisions by the controller.

TMU initiated flight plan amendments are typically for flow management, including pre-departure route changes for weather, playbook and other mandatory reroutes, reroutes in response to sector volume, and reroutes for fix balancing [25]. Controller initiated flight plan amendments include reroutes for spacing and separation, and pilot requests for deviation to avoid weather [20]. TOSs are typically used for flow management, requiring TMU decision making. It is therefore likely that airborne TOSs must be acceptable to traffic managers. Traffic managers have access to more information than controllers, particularly with regard to downstream demand and capacity. To ensure that the chosen airborne TOS is acceptable to traffic managers, it is desirable to train the predictor on flight plan amendments that implement TMU decision making specifically.

Unfortunately, no information is recorded on who makes the decision leading to a flight plan amendment. To work around this issue, flight plan amendments were filtered to exclude direct routing (direct routings are typically pilot requests), and any routing changes that do not extend across multiple Centers (and are therefore likely used for tactical avoidance of weather or for spacing and separation) [25]. While this approach does not filter out all controller initiated

reroutes, and may filter out some TMU initiated reroutes (especially intra-facility TMU route amendments), the remaining reroutes are likely to be predominantly TMU-initiated.

Two-class classification, which produces better performing predictors than one-class classification, requires both 'positive' and 'negative' training data, for which reroutes were either operationally acceptable (positive data) or unacceptable (negative data). A total of 3,443 historical flight plan amendments from July to September 2015, with the controller initiated reroutes filtered out, represent the positive training data for the development of the predictor. Unfortunately, negative training data – operationally unacceptable reroutes that were rejected by the TMU – are not recorded by the FAA, and were not available. Negative training data was therefore generated artificially. This was done by identifying potential alternative amendments, for each flight plan amendment made, that were not implemented, as described below. These alternative amendments could be operationally acceptable in general but not for the given weather, traffic and airspace situation in question.

Historical data and clustering were used to generate potential alternatives for each amendment in the same way that trajectory options were generated, described above. In order to ensure that the alternative amendments generated were not only geographically distinct from each other, but also from the original flight plan, the original flight plan was also assigned to a cluster, so that the cluster containing it could be dropped. The trajectory with highest historical usage (based on minimum waypoint pair count in Section III-B) was then extracted from the remaining clusters, to define the alternative amendments. If any of these alternative trajectories matched the actual reroute implemented, it was also dropped. Because none of the remaining alternative amendments were implemented by the traffic manager, they were considered to be operationally unacceptable, and were used as negative training data. In reality, it is possible that these routes would have been acceptable, but were either not requested or not chosen for implementation. The model trained on this data therefore captures the TMUs decision making on both operational acceptance, and choice of which operationally acceptable route to implement. Between 1 and 5 alternative amendments were generated for each historic flight plan amendment in July 2015, with a total of 5,913 alternative amendments generated. A total of 9,356 observations (positive and negative) were therefore available.

A number of features were calculated for each amendment (actual and alternative) in the training data. These features describe historical usage, downstream demand to capacity imbalance, and increase in flight duration relative to the original flight plan. They are described in detail below.

Historical usage was calculated as described in Section III-B, with all four metrics included as features (full trajectory count, full trajectory count as a reroute, minimum waypoint pair count, and minimum waypoint pair count as a reroute). Also included as features were the difference in count between the original route and the amendment (for each of the four metrics). Eight features were therefore included describing historical usage. Because historical usage was also used to identify the trajectory options for which operational acceptability is to be predicted, these features were not expected to dominate in the way they did in past work [20].

For each amendment, demand was calculated for each sector downstream of the maneuver start point. Demand was defined as the number of flights predicted to be in each sector, estimated as described by [20] using the Future ATM Concepts Evaluation Tool (FACET). This required that flight demand in each downstream sector be predicted for each amendment in the training set based on what was known when the flight plan amendment was implemented. Demand was therefore predicted based on the position, speed and flight plan of airborne aircraft and departure times and flight plans for flights still on the ground, at the time the flight plan amendment was made.

Capacity was calculated for all downstream sectors on each amendment in two separate ways – firstly based on the Monitor Alert Parameter (MAP), which is an estimate of how many aircraft can be reasonably controlled in a sector in clear weather, and secondly by the weather impact on the sector. The latter was estimated using the percentage overlap between weather polygons generated using the Convective Weather Avoidance Model (CWAM) [26] and the sector. CWAM polygons are based on probabilistic weather avoidance fields, which represent regions of airspace that pilots are likely to avoid due to the presence of convective weather. While traffic managers do not always use CWAM when making decisions about trajectories, it provides a good proxy for the weather products and other factors that are considered, which may include how fast the weather is moving, what direction it is moving, how it is developing, and the extent to which pilots may be able to vector through the weather etc. CWAM includes forecasts in 15-minute increments, for up to 2-hrs into the future, and includes polygons for different probabilities of deviation – 60%, 70% and 80% – across a range of altitudes. For each amendment, the predicted sector entry times were used to extract the appropriate CWAM polygons for up to 2-hrs into the future, over an altitude range from 30,000ft to 40,000ft. Capacities were calculated using each of the 60%, 70% and 80% polygons, and included as separate features. For sectors further than 2-hrs in the future, the capacity was assumed to be unaffected by weather. In future work, more strategic weather forecasts, such as the CDM (Collaborative Decision Making) Convective Forecast Planning (CCFP) forecast, which extends up to 6-hrs into the future, will be included. However, it is unclear to what extend these highly strategic forecasts are considered by traffic managers when rerouting aircraft, because they are so far in the future, and therefore have high uncertainty associated with them.

A number of features were calculated describing the downstream demand to capacity imbalance for each amendment. Because the reroute options are likely in response to capacity overload downstream, either created by traffic demand or weather blockage, these features were defined as follows:

1. The maximum ratio of predicted sector demand to capacity across all sectors downstream;
2. The average ratio of predicted sector demand to capacity across all sectors downstream;
3. The total number of downstream sectors with predicted demand exceeding capacity;
4. Whether or not any downstream sector had a ratio of predicted demand to capacity greater than unity; and finally
5. A metric comparing the sum of predicted demand to capacity ratios for the amendment to that of the original route, calculated as in (3):

$$\Delta_{SumD/C} = \sum_{S_O} \frac{D_s}{C_s} - \sum_{S_A} \frac{D_s}{C_s} \quad (3)$$

where S_O and S_A represents the set of sectors on the original route and amendment, respectively; D_s represents the predicted demand in sector s; and C_s represents the capacity of sector s. Each of these features was calculated with capacity defined by the sector MAP, and with capacity defined by the sector MAP reduced by the degree to which the CWAM polygon overlaps the sector, as described in (4).

$$C_{CWAM} = C_{MAP} \cdot (1 - Sector\ CWAM\ Overlap) \quad (4)$$

Because there are three different definitions of sector CWAM overlap, depending on probabilities of deviation, a total of 20 features (5 based on C_{MAP} and 15 based on C_{CWAM}) describe demand to capacity imbalance.

Downstream weather impacts were also included explicitly as features. This was done by including the following features:

1. The maximum predicted sector CWAM overlap downstream;
2. The average predicted sector CWAM overlap downstream;
3. The total number of sectors with predicted sector CWAM overlap greater than zero downstream;
4. Whether or not any downstream sector had a predicted sector CWAM overlap greater than zero downstream; and
5. The difference between the sum of predicted sector CWAM overlaps downstream for the amendment and for the original route.

These five features were calculated with CWAM polygons at each of the three values of probability of pilot deviation, resulting in a total of 15 downstream weather impact features.

For many amendments, the weather impacting the route is well downstream, requiring forecasts of weather impact hours in advance. However, forecasts of this type are not typically very accurate at the sector level, because of the small size of a sector. They are, however, more accurate at the Center level. Therefore, the features listed above were also included describing CWAM Center overlap, as opposed to CWAM sector overlap. In the available data, Center CWAM overlap was only calculated for a probability of pilot deviation of 60%, so only 5 features were added describing weather impact at the Center level.

The flight duration of the amendment was also included as a feature, along with the change in flight duration relative to the original flight plan. The number of downstream sectors between the maneuver start point and the destination was also included, along with the difference in number of downstream sectors between the amendment and the original route. This added 4 features to the total feature list, resulting in a total of 52 features.

A number of machine learning algorithms were trained on the developed feature set using the *Python sklearn* library [27]:

- Logistic regression – with C, the inverse of regularization strength, set to 2.5;
- Multi-layer perception neural network – with two hidden layers, each with a depth of 100 neurons, logistic activation functions, and the alpha L2 penalty regularization term set to 0.001;
- Support Vector Machine (SVM) – with linear kernel, and penalty parameter C set to 1.0;
- SVM – with sigmoid kernel, and penalty parameter C set to 0.5;
- Random forest – with 100 estimators; and
- Adaptive Boost – with 200 estimators.

All model parameters were set as default in the *sklearn Python* library, with the exception of those listed above, which were set based on best model performance, training and testing each algorithm on a range of values for each parameter. Because more negative training data was used than positive training data (between 1 and 5 negative alternative amendments were generated for each positive observed amendment), the dataset was imbalanced, with 36.8% of observations positive and 63.2% negative. The Synthetic Minority Over-Sampling Technique (SMOTE) [28], which artificially generates observations for the minority class (in this case the positive class) based on the existing observations, was therefore used to balance the dataset, for improved model performance. K-fold cross-validation, with K=10, was used to estimate the performance of the different models, which are presented in **Table I**.

TABLE I. MACHINE LEARNING ALGORITHM PERFORMANCE IN PREDICTING OPERATIONAL ACCEPTANCE OF TRAJECTORIES FROM JUNE TO SEPTEMBER 2015.

	Logistic Regression	***Multi-Layer Perceptron***	***SVM-Linear Kernel***	***SVM-Sigmoid Kernel***	***Random Forest***	***Ada Boost***
Accuracy	0.95	0.94	0.96	0.88	0.96	0.95
Recall	0.97	0.94	0.97	0.87	0.95	0.96
Precision	0.91	0.91	0.92	0.82	0.93	0.92
F_1-Score	0.94	0.92	0.94	0.84	0.94	0.94
AUC	0.99	0.99	0.98	0.93	0.99	0.99

A number of metrics were considered. *Accuracy*, which measures the fraction of correct predictions from all predictions made, is the most intuitive, but can be misleading when datasets are imbalanced. An alternative metric is *F_1-Score* (also called *F-Score* or *F-Measure*), which is the harmonic mean of precision and recall, calculated as in (5).

$$F_1 = 2 \cdot \frac{precision \cdot recall}{precision + recall} \tag{5}$$

Here *precision* refers to the number of elements correctly labeled by the model as belonging to the positive class divided by the total number of elements labeled by the model as belonging to the positive class (i.e., the fraction of retrieved instances that are relevant). *Recall* refers to the number of elements correctly labeled by the model as belonging to the positive class divided by the total number of elements that actually belong to the positive class in the data (also called *Sensitivity* or *True Positive Rate*). Accuracy, recall, precision and F_1-Score vary from 0 to 1, with 1 being best. The discrimination threshold for all these metrics is set to 0.5, giving equal importance to both classes.

A Receiver Operating Characteristic (ROC) curve plots recall against false positive rate across varying discrimination thresholds. The area under the curve (AUC) provides a metric of model performance ranging from 0 to 1 (1 being best) that is not a function of the chosen discrimination threshold. When using normalized units, AUC indicates the probability that a classifier will rank a randomly chosen positive instance higher than a randomly chosen negative one (assuming positive ranks higher than negative) [29].

The results in **Table I** indicate that the feature set developed is very effective at predicting the observed data. The model with highest accuracy (0.96), precision (0.93), F-Score (0.94) and AUC (0.99) was the random forest, so this model was used for evaluating the trajectory options in the sample application. The other models also performed well across all metrics.

The relative importance of each feature can be evaluated using the average rank of the features used as decision nodes in each predictor tree in the random forest. These average ranks are shown for the top 10 features in **Table II**. The most important features were identified accordingly to be the differences in flight duration, demand to capacity imbalance (particularly accounting for the impact of downstream weather on sector capacity), and number of sectors traversed, between the amendment and original route. Trajectories are therefore predicted to have high probability of acceptance when they deviate relatively little from the original routing, but pass through sectors that have less CWAM polygon overlap than the original routing. Features that have lower importance include the historical usage of the specific routes, and the downstream weather impact on capacity, without consideration of demand (in the form of sector and center CWAM polygon overlap), neither of which appear in the list in **Table II**.

The best performing of the applied machine learning algorithms was used to categorize each of the available trajectory options identified in Section III-A as operationally acceptable or unacceptable, outputting a percentage probability of acceptance.

D. Select the TOS based on Probability of Acceptance

CTOP currently limits the number of trajectory options that can be submitted within a TOS to five. While this may change in the future, and may be different for airborne TOSs, for this paper one trajectory option was generated through each active FCA, which define the constrained region of airspace, and one around all active FCAs, on either side. Therefore, for a single FCA, three trajectory options were chosen. If the original trajectory routed through the FCA, this was included as one trajectory option, leaving two to be identified either side of the FCA.

A trade-off must typically be made between cost efficiency and operational acceptability. In this paper, the focus is on maximizing operational acceptability, under the conditions at the time the TOS was generated. Hence, the final TOS was selected based on maximizing probability of operational acceptance, predicted in Step C. This is in contrast to the approach in [16] and in some commercial TOS generators, which optimize for cost, subject to a pre-calculated clearable route network. In future work, these approaches could be combined.

IV. ANALYSIS OF PRE-DEPARTURE SAMPLE APPLICATION

An operationally acceptable TOS was generated for an historic pre-departure flight scheduled from Dallas/Fort Worth International Airport (DFW) to Newark Liberty International Airport (EWR), on July 12, 2015 at 13:06Z. For the example application, one FCA was assumed to be set, at the boundary of Memphis and Indianapolis Centers (ZME

TABLE II. RELATIVE IMPORTANCE OF FEATURES, DERIVED FROM A RANDOM FOREST.

Feature	***Importance***
Change in flight duration	0.21
Change in sum of sector demand/reduced capacity – 60%	0.12
Change in number of sectors traversed	0.11
Change in sum of sector demand/reduced capacity – 80%	0.11
Change in sum of sector demand/MAP	0.07
Change in sum of sector demand/reduced capacity – 70%	0.06
Number of sectors in amendment	0.05
Amendment duration	0.03
Change in sum of Center CWAM overlap – 60%	0.03
Maximum sector demand/MAP of amendment	0.02

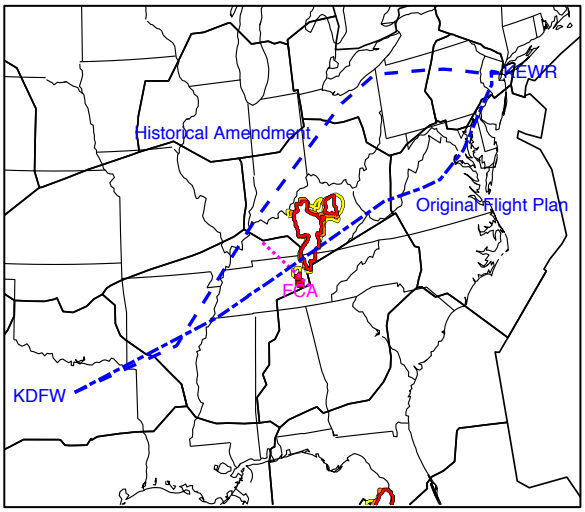


Figure 1. Sample application for generating operationally acceptable TOS for pre-departure flight from DFW to EWR on July 12, 2015, at 13:06Z.

and ZID, respectively), because of convective weather. The original flight plan, FCA and convective weather before departure (in the form of CWAM polygons, with yellow boundaries representing 60% probability of deviation, orange 70%, and red 80%) are shown in **Figure 1**, along with the flight plan amendment implemented on July 12, 2015, to avoid the weather.

A. Identify Available Trajectory Options

A total of 73 trajectory options were extracted from the historical data processed (April to June 2015) for the sample application.

B. Down-Select Trajectory Options Using Clustering

As described in Section III-B, clustering was used to ensure that the trajectory options selected for the TOS were suitably spaced geographically, and to reduce the number of

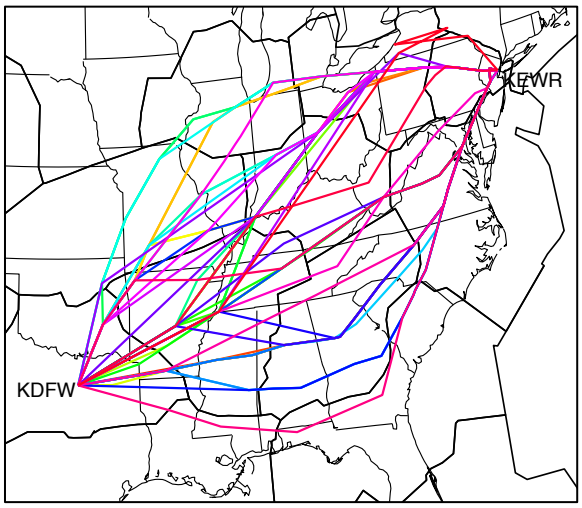
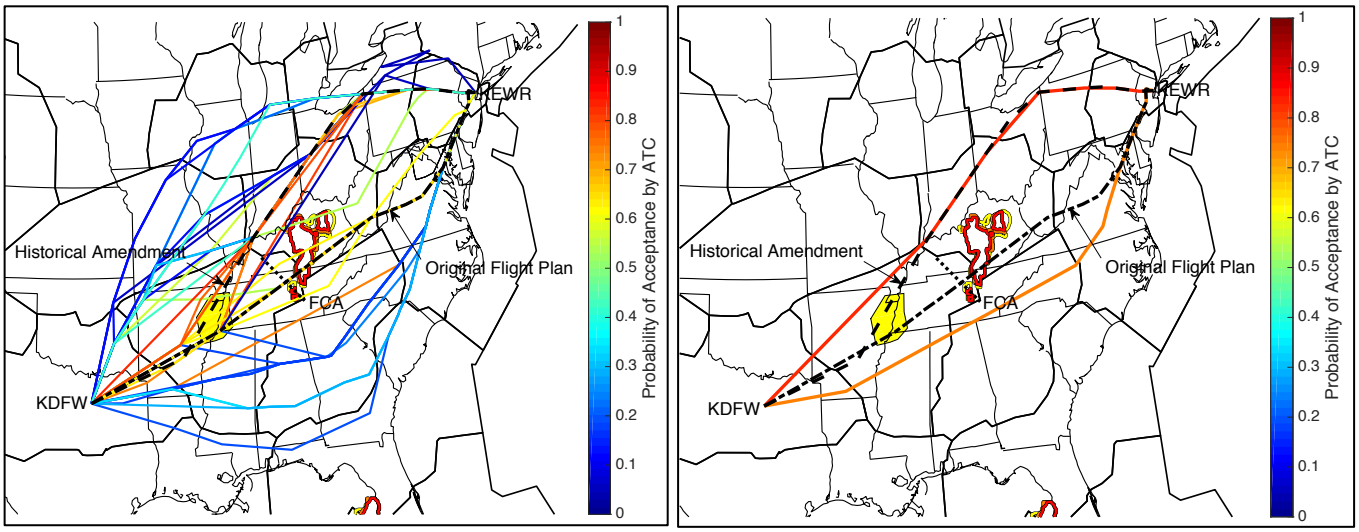


Figure 2. Available trajectory options for pre-departure TOS from DFW to EWR.

trajectory options for which operational acceptability must be predicted. All trajectory options identified in Section IV-A for the sample application were therefore clustered as described in Section III-B. Forty-seven clusters were extracted. The most commonly flown route in each cluster is shown in **Figure 2** for the clusters identified.

C. Predict Operational Acceptability using Machine Learning

Operational acceptability was predicted for only the most commonly flown routes in each cluster. The probability of each of the trajectory options identified for the sample application being operationally acceptable was calculated using the trained random forest model. The results are shown in **Figure 3(a)**, with the probability of acceptance shown by color. Sectors with demand predicted to exceed the MAP value are shown in yellow, while CWAM polygons when the amendment was implemented historically are also shown.



(a) (b)

Figure 3. Estimated probability of acceptance for trajectory options in pre-departure sample application from DFW to EWR on July 12, 2015 at 13:06Z: (a) all trajectory options and (b) chosen trajectory option set.

The FCA is also shown, on the boundary of ZME and ZID in **Figure 3**. There is a wide range in plotted probability of acceptance, varying from 0.1 to 0.85.

The longest trajectories, to the north and south, have probabilities of acceptance of 0.4 or lower. The shortest trajectory options, which deviate least from the original routing but route through the forecast convective weather, have probabilities of acceptance of 0.6. The trajectories with highest probability of acceptance – between 0.7 and 0.85 – lie in between, routing close to the original route, but not through the forecast convective weather. These results indicate the dominant effect of flight duration and demand to capacity imbalance, accounting for the impact of downstream weather on sector capacity, in the trained algorithm.

D. Select the TOS based on Probability of Acceptance

The trajectories with highest estimated probability of acceptance either side of the FCA were chosen for the TOS, and are shown in **Figure 3(b)**. The northerly trajectory has a probability of acceptance of 0.85, and very closely matches the reroute given to this flight historically. The southerly trajectory has a probability of acceptance of 0.76.

V. Implications of Results

The machine learning test results in **Table I** indicate that, given the limited testing completed to date, operational acceptability may be predictable with high accuracy. This suggests that a tool such as that developed could be useful in TOS generation.

The feature importance results in **Table II** indicate that the features describing differences in flight duration, demand to capacity imbalance (particularly accounting for the impact of downstream weather on sector capacity), and number of sectors traversed, between the amendment and original route, are the most important. In fact, with only the top four features, the model accuracy is 0.95 (compared to 0.96 with all features). It makes sense that these features dominate, because traffic managers are expected to minimize any increase in flight duration, while avoiding sectors with high capacity to demand imbalance. The implication, however, is that a heuristic that accounts for only these factors may be sufficient to improve operational acceptability of TOSs.

Downstream weather impact on capacity, without consideration of demand, and historical usage do not show high importance in **Table II**. The former indicates that traffic managers consider demand to capacity imbalance. The latter is expected because the underlying set of trajectory options considered already filter for routes with high historical usage. This result does not therefore indicate that historical usage is unimportant.

In the sample application presented in **Figure 3**, the TOS selection is clear, with large differences in probability of acceptance between trajectories. However, this is a relatively simple problem, with clearly alternatives around the weather. In more difficult weather problems that were tested, with less clear alternatives around the weather, there were fewer trajectories with high probability of acceptance, which is to be expected.

VI. Conclusions and Recommendations for Future Work

Trajectory negotiation between the ANSP and flight operators is likely to be a key component of future air traffic control systems. The objective of this paper was to describe and demonstrate an approach for automatically generating TOSs, both pre-departure and airborne, that have high probability of operational acceptance as strategic reroutes, given the conditions at the time the TOS was generated.

An approach was developed that uses hierarchical clustering of historical route data to identify route candidates, for which operational acceptability can then be predicted using models trained on historical flight plan amendment data using supervised machine learning.

Features used to classify trajectories as operationally acceptable or not described historical route usage, change in flight duration relative to the original route, downstream demand to capacity imbalance, and changes in these conditions relative to the original route. These data are not readily available, and had to be generated using simulation and historical data. Key challenges were the identification of relevant TMU initiated reroutes for positive training data, and the generation of artificial negative training data, which are not otherwise available.

While all models tested performed well, a random forest with synthetic minority class oversampling was found to be the best performing algorithm for learning operational acceptability from three months of historical flight amendment data (July to September 2015) and one month of artificially generated alternative amendments that were not flown (July 2015). Model accuracy was 0.96, F_1-Score 0.94, and AUC 0.99. This indicates that the operationally acceptability of strategic reroutes is predictable. The most important features were identified to be differences in flight time, demand to capacity imbalance (particularly accounting for the impact of downstream weather on sector capacity), and number of sectors traversed, between the amendment and original route.

The approach was demonstrated for an historical pre-departure flight from DFW to EWR. The approach was able to identify routes either side of the FCA with the highest probability of operational acceptance, and the chosen trajectory options can be seen to limit increases in flight time relative to the original route, while reducing the overlap with forecast convective weather. In future work, the models and approach will be refined and can be expanded to optimize flight trajectories based on operational conditions, such as wind, to produce efficient routes that have a high probability of operational acceptance.

Acknowledgments

This research was funded by NASA under contract number NNA16BD14C. The authors would like to acknowledge helpful comments from Heather Arneson,

Deepak Kulkarni and Yao Wang of NASA Ames Research Center, and the provision of CWAM sector and Center overlap data by *ATAC Corporation*.

AUTHOR BIOGRPAHIES

Dr. Antony Evans is a senior scientist with Crown Consulting, Inc., in the Aviation Systems Division at NASA Ames Research Center. He has two Master's degrees – in aeronautics and astronautics, and technology policy – from MIT and a Ph.D. in aviation and the environment from Cambridge University.

Dr. Paul Lee is a sub-project manager for Integrated Demand Management research in the Human-Systems Integration Division at NASA Ames Research Center. He has a Master's degree in engineering and a Ph.D. in cognitive psychology, both from Stanford University.

REFERENCES

[1] Arora, N., *et al.*, "Collaborative Air Traffic Management Technologies (CATMT) Functional Analysis for Extending Collaborative Trajectory Options Procedures (CTOP) into Tactical the Time-frame," Version 1.0, Metron Aviation, ARCE-F810-0910-037, Dulles, VA, September 30, 2010.

[2] Fernandes, A.D., *et al.*, "Initial Concept of Operations for Full Management by Trajectory," NASA/CR-2017-219674, Washington, DC, 2017.

[3] Hall, W., and Hunter, G., "Trajectory Optimization and the Clearable Route Network," 18th AIAA Aviation Technology, Integration, and Operations Conference, Atlanta, GA, 24-29 June, 2018.

[4] Sheth, K., McNally, D., Somersall, P., Morando, A., Clymer, A., Shih, F.T., "Assessment of a National Airspace System Airborne Rerouting Tool," Eleventh USA/Europe Air Traffic Management Research and Development Seminar (ATM2015), Lisbon, Portugal, 2015.

[5] McNally, D., *et al.*, "Dynamic Weather Routes: A Weather Avoidance System for Near-Term Trajectory-Based Operations," 28th International Congress of the Aeronautical Sciences, Brisbane Australia, September 2012.

[6] McNally, D., *et al.*, "Dynamic Weather Routes: Two Years of Operational Testing at American Airlines," Air Traffic Control Quarterly, Vol. 23, No. 1, pp. 55-81, 2015.

[7] Sheth, K., Clymer, A., Morando, A., and Shih, F., "Analysis of Multi-Flight Common Routes for Traffic Flow Management," 16th AIAA Aviation Technology, Integration, and Operations Conference, Washington, DC, p. 13. 2016.

[8] Gong, C., McNally D., and Lee C-H., "Dynamic arrival routes: A trajectory-based weather avoidance system for merging arrivals and metering." In AIAA Aviation Technology, Integration, and Operations Conference, p. 3394, 2015.

[9] Ballin, M.G., and Wing, D.J., "Traffic Aware Strategic Aircrew Requests (TASAR)," 12th AIAA Aviation Technology, Integration, and Operations Conference, Indianapolis, IN, 2012.

[10] Idris H., Enea G., Burke K., and Wing D., "Assessment of Air Traffic Controller Acceptability of Aircrew Route Change Requests", FAA Eurocontrol, ATM R&D Seminar, Seattle, WA 2017.

[11] Evans, A.D., Sridhar, B., and McNally, D., "Improving Operational Acceptability of Dynamic Weather Routes Through Analysis of Commonly Used Routings," In 16th AIAA Aviation Technology, Integration, and Operations Conference, p. 3600, 2016.

[12] Adacher, L., M. Flamini, and E. Romano, "Rerouting algorithms solving the air traffic congestion," AIP Conference Proceedings, vol. 1836, no. 1, p. 020053, AIP Publishing, 2017.

[13] Agustı, A., A. Alonso-Ayuso, L.F. Escudero, and C. Pizarro, "On air traffic flow management with rerouting. Part I: Deterministic case," European Journal of Operational Research 219, no. 1, pp. 156-166, 2012.

[14] Agustı, A., A. Alonso-Ayuso, L.F. Escudero, and C. Pizarro, "On air traffic flow management with rerouting. Part II: Stochastic case," European Journal of Operational Research 219, no. 1, pp. 167-177, 2012.

[15] Ayo, B.S., Y.F. Hu, and J.-P. Li. "Flight Parout: A simulation platform for intelligent flight path reroutes for adverse weather." IEEE/AIAA 37th Digital Avionics Systems Conference, London, UK, pp. 1-9. IEEE, 2018.

[16] Taylor, C., and Wanke, C., "Dynamically Generating Operationally Acceptable Route Alternatives Using Simulated Annealing," Air Traffic Control Quarterly, Vol. 20, No. 1, pp. 97-121, 2012.

[17] Taylor, C., Wanke, C., "Dynamic Generation of Operationally Acceptable Reroutes," 9th AIAA Aviation Technology, Integration, and Operations Conference, Hilton Head, SC, 21-23 September 2009.

[18] Stewart, T., Askey, L., and Hokit, M., "A Concept for Tactical Reroute Generation, Evaluation and Coordination," 12th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, AIAA 2012-5586, Indianapolis, IN, 2012.

[19] DeArmon, J.S., Chaloux, D.A., Hokit, M.E., Schafer, T.M., Conroy, J., "Metering During Severe En Route Weather Via Advanced Flight-Specific Trajectories (AFST)," 17th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, AIAA 2017-3591, Denver, CO, 2017.

[20] Evans, A.D., Lee, P., "Predicting the Operational Acceptance of Route Advisories," 17th AIAA Aviation Technology, Integration, and Operations (ATIO) Conference, AIAA.2017-3078, Denver, CO, 2017.

[21] Arneson, H., "Initial Analysis of and Predictive Model Development for Weather Reroute Advisory Use," 15th AIAA Aviation Technology, Integration, and Operations Conference, AIAA 2015-3395, Dallas, TX, 2015.

[22] Arneson, H., Bombelli, A., Segarra-Torné, A., Tse, E., "Analysis of convective weather impact on pre-departure routing of flights from Fort Worth Center to New York Center," 17th AIAA Aviation Technology, Integration, and Operations Conference, AIAA 2017-3593, Denver, CO, 2017.

[23] Bombelli, A., Torné, A.S., Trumbauer, E. and Mease, K.D., "Automated Route Clustering for Air Traffic Modeling," AIAA Modeling and Simulation Technologies Conference, Grapevine, TX, p. 1318, 2017.

[24] Conde Rocha Murca, M., DeLaura, R., Hansman, R.J., Jordan, R., Reynolds, T. and Balakrishnan, H., "Trajectory clustering and classification for characterization of air traffic flows," 16th AIAA Aviation Technology, Integration, and Operations Conference, Washington DC, p. 3760, 2016.

[25] Personal correspondence with Subject Matter Experts Mark Evans, Danny Vincent and Bob Staudenmeier.

[26] Chan, W., Refai, M., DeLaura, R., "Validation of a Model to Predict Pilot Penetrations of Convective Weather," 7th AIAA Conference on Aviation, Technology, Integration and Operations, Belfast, Ireland, 2007.

[27] Pedregosa *et al.*, "Scikit-learn, Machine Learning in Python," Journal of Machine Learning Research 12, pp. 2825-2830, 2011.

[28] Chawla, N.V., Bowyer, K.W., Hall, L.O. and Kegelmeyer, W.P., 2002. SMOTE: synthetic minority over-sampling technique. Journal of artificial intelligence research, 16, pp.321-357.

[29] Fawcett, T., "An introduction to ROC analysis," Pattern Recognition Letters, (2006) 27, 861–874.