

Application Of Data Fusion And Machine Learning To The Analysis Of The Relevance Of Recommended Flight Reroutes

Ghislain Dard*, Eugene Mangorley *, Olivia J. Pinon†, and Dimitri N. Mavris‡

Georgia Institute of Technology, Atlanta, GA, 30332

One of the missions of the Federal Aviation Administration (FAA) is to maintain the safety and efficiency of the National Airspace System (NAS). One way to do so is through Traffic Management Initiatives (TMIs). Traffic Management Initiatives, such as reroute advisories, are issued by Air Traffic Controllers whenever there is a need to balance demand with capacity in the National Airspace System. Indeed, rerouting flights ensures that aircraft operate with the flow of traffic, remain away from special use airspace, and avoid saturated areas of the airspace and areas of inclement weather. Reroute advisories are defined by their level of urgency i.e. Required, Recommended or For Your Information (FYI). While pilots almost always comply with required reroutes, their decisions to follow recommended reroutes vary. Understanding the efficiency and relevance of recommended reroutes is key to the identification and definition of future reroute options. In addition, because traffic in the National Airspace System can be forecasted through airline schedules and flight plans, it is also possible to predict the issuance of volume-related reroute advisories. Consequently, the objectives of this work is two-fold: 1) Assess the relevance of existing recommended reroutes, and 2) predict the issuance and the type of volume-related reroute advisories. This was achieved by 1) fusing data from relevant datasets, extracting statistics, and identifying trends and patterns within the data, and 2) developing models to predict the issuance of volume-related reroute advisories. It is expected that the capabilities developed may ultimately contribute to reducing unnecessary flight reroutes.

I. Nomenclature

<i>AAR</i>	=	Airport Acceptance Rate
<i>ARTCC</i>	=	Air Route Traffic Control Center
<i>ATCSCC</i>	=	Air Traffic Control System Command Center
<i>CASSIE</i>	=	Computing Analytics and Shared Services Integrated Environment
<i>FAA</i>	=	Federal Aviation Administration
<i>FIXM</i>	=	Flight Information Exchange Model
<i>GADV</i>	=	General Advisory
<i>CSV</i>	=	Comma Separated Values
<i>MAE</i>	=	Mean Absolute Error
<i>NAS</i>	=	National Airspace System
<i>NetCDF</i>	=	Network Common Data Form
<i>SFDPS</i>	=	System Wide Information (SWIM) Flight Data Publication Service
<i>SVM</i>	=	Support Vector Machines
<i>TFMS</i>	=	Traffic Flow Management System
<i>TMI</i>	=	Traffic Management Initiatives
<i>TRACON</i>	=	Terminal Radar Approach Control

*Graduate Research Assistant, Aerospace Systems Design Laboratory, School of Aerospace Engineering, AIAA Student Member

†Senior Research Engineer, Aerospace Systems Design Laboratory, School of Aerospace Engineering, AIAA Senior Member

‡Regents Professor for Advanced Systems Analysis, School of Aerospace Engineering, AIAA Fellow

II. Introduction

THE National Airspace System (NAS) is comprised of air navigation facilities, air traffic controllers, facilities, airports, technologies, rules, regulations and procedures that are needed to manage and ensure the safety of the United States airspace [1]. The US airspace itself is broken down into twenty-one sectors, as seen in Figure 1, with each one having precise characteristics in terms of capacity and traffic. On any given day, an area or sector in the National Airspace System can be impacted by a variety of events, such as weather, aircraft congestion, etc. Congestion, in particular, occurs when the demand for a specific sector exceeds its capacity. One way to mitigate the impact that such events may have on the efficiency and safety of the NAS is through the issuance of Traffic Management Initiatives (TMI).

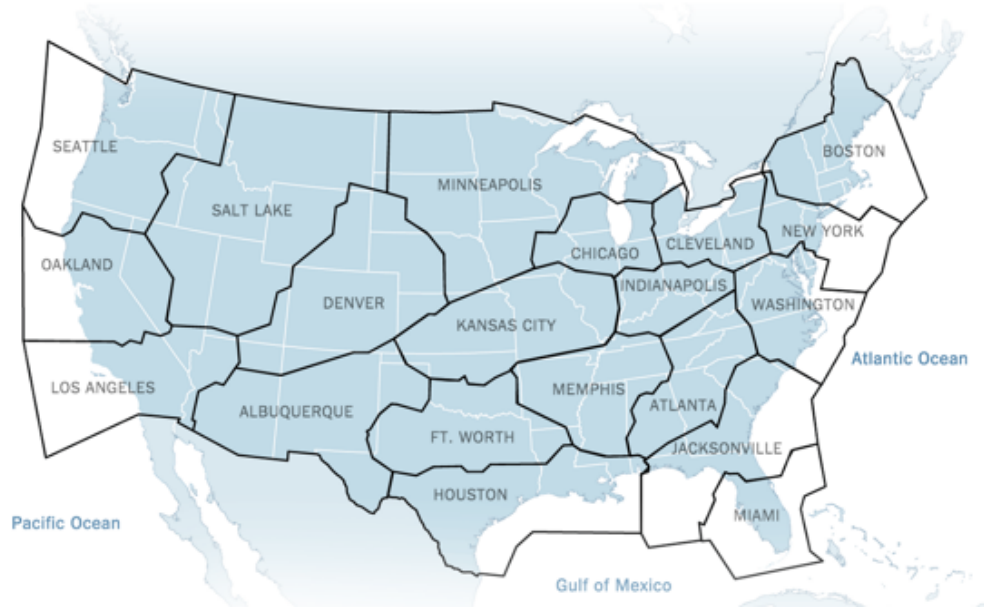


Fig. 1 Air Traffic Control Sectors of the NAS [2]

A. Traffic Management Initiatives (TMI)

Traffic Management Initiatives are implemented to balance demand with capacity either at an airport or in an area of the National Airspace System [3]. Traffic Management Initiatives are divided into two categories: Airport-Specific (Terminal) Traffic Management Initiatives and En Route Traffic Management Initiatives [4]. Airport-Specific (Terminal) Traffic Management Initiatives such as Ground Delay Programs (GDP) and Ground Stops (GS) are issued to deal with the flow of aircraft arriving at an airport. If the number of aircraft heading to an airport is above the airport's Airport Acceptance Rate (AAR), air traffic managers may implement these Traffic Management Initiatives to slow down air traffic and ensure that the airport's acceptance rate matches or exceeds aircraft demand [5]. En Route Traffic Management Initiatives on the other hand, may be issued to manage en route flights affected by constraints in the National Airspace System. These Traffic Management Initiatives include Airspace Flow Program (AFP), Miles-in-Trail (MIT), and Reroute Advisories.

B. Reroute Advisories

Whenever an area in the National Airspace System is constrained, traffic management personnel locate the constraint, assess which airport(s) and route(s) are affected, and evaluate the seriousness of the constraint and its duration. Once this information has been gathered, an ARTCC can decide to issue Traffic Management Initiatives such as reroutes to address the constraint. Rerouting flights ensures that aircraft operate with the "flow" of traffic, remain away from special use airspace such as those for military use, avoid overcrowded areas of the airspace, and avoid areas of inclement weather. Reroute advisories specify the constrained area, the effective period of the advisory, the nature of the incident, the probability of the reroute's extension, and the new routes for affected flights. Once flight operators receive a reroute advisory, they then have to either submit a flight plan amendment or submit an alternative route and check with Air

Traffic Controllers if it is accepted. Reroute advisories are characterized by their level of urgency. These are [4]:

- Required Reroutes: These routes are required to be followed by all aircraft captured in the scope of the reroute
- Recommended Reroutes: Air Traffic Controllers recommend flight operators to use these routes, but do not require them to use them
- For Your Information (FYI) Reroutes: Air Traffic Controllers issue these reroutes to let pilots know that these routes are available

Figure 2 shows the distribution of reroute advisories per urgency level for one day (April 21, 2017). In particular, it shows that even if a large majority of reroute advisories are required, recommended and FYI reroute advisories still represent more than one third of all reroute advisories.

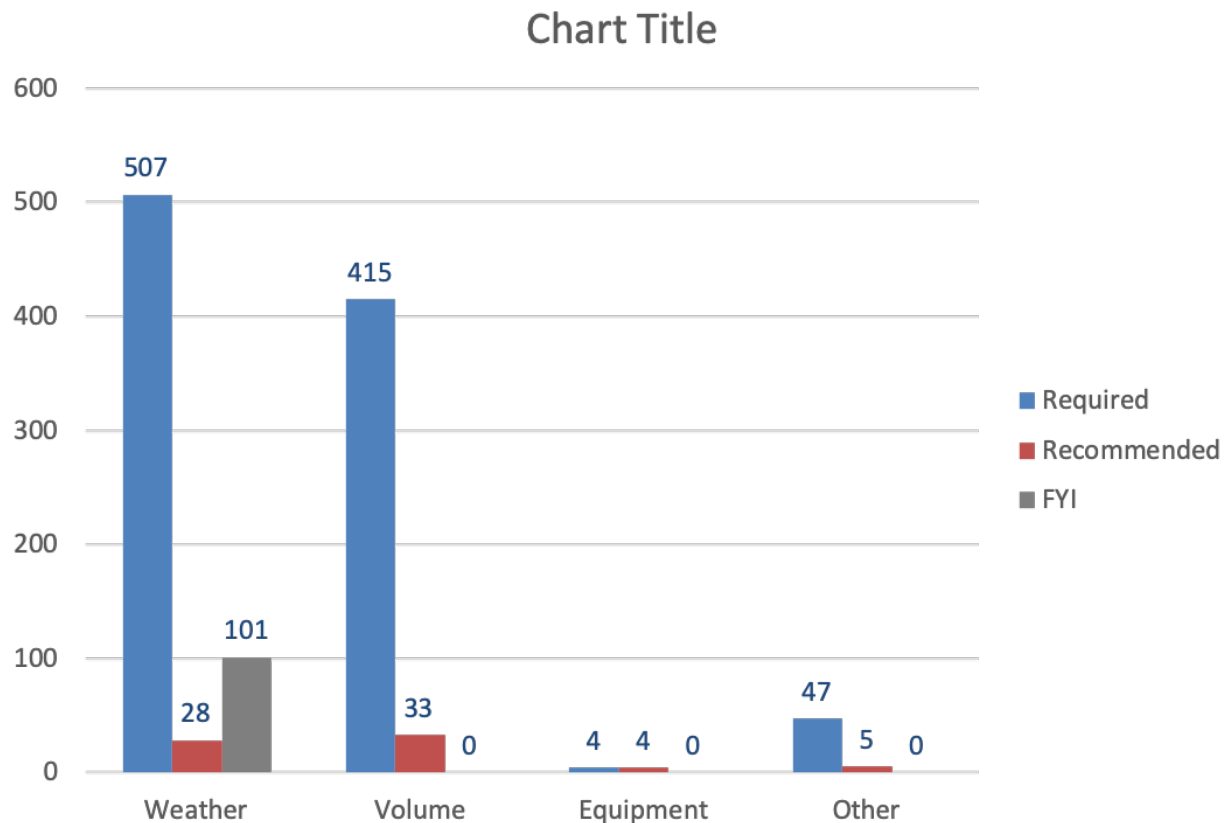


Fig. 2 Reroute messages by levels of urgency on April 21, 2017

In order to analyze reroutes, there is a need to understand the format of routes used in the National Airspace System. Flight routes follow an internationally standardized format to ensure uniformity across the world. Their format is a sequence of elements that belong to a catalogue of points and routes accredited by the aviation administration they belong to. Below is a sample route extracted from a flight plan:

KEWR..ELVAE..COL..WHITE.J209.SBY..KEMPR..ILM.AR21.CRANS.FISEL6.KFLL

Flight routes between two airports always start and end with the origin airport and the destination airport: in the example provided in Figure 3, KEWR (Newark Liberty International Airport) is the origin and KFLA (Fort Lauderdale–Hollywood International Airport) is the destination.

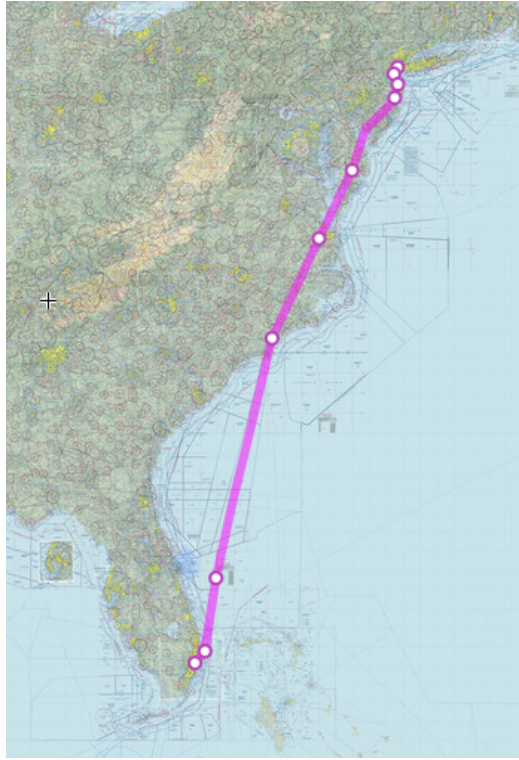


Fig. 3 Visualization of a route from KEWR to KFLA

ELVAE, WHITE, KEMPR, CRANS are FAA Fix Waypoints. These waypoints are geographical points on the Earth's surface. COL, SBY, ILM are Nav aids, which are physical devices on the ground that transmit radio signals that aircraft can detect and follow. Finally, J209 and AR21 are air route names. These are alphanumeric codes that define corridors connecting specified locations to each other at specified altitudes.

Presently, the Federal Aviation Administration does not have a means to assess the relevance of recommended reroute advisories. Consequently, the present research proposes to provide FAA analysts with an approach for assessing the relevance of recommended reroutes.

C. Research Scope

The objective of this research is two-fold:

1. Assess the relevance of recommended reroutes

As mentioned previously, reroute advisories are defined by their level of urgency (Required, Recommended or For Your Information (FYI)). According to Federal Air Regulation §91.123 [6], flight operators are allowed to refuse a specific route as long as they submit an alternative one that is validated by Air Traffic Controllers. The Federal Aviation Administration's analysts are thus interested in analyzing how often pilots follow recommended reroutes in order to assess their relevance, and to identify optimal routes for future events.

2. *Predict the issuance of volume-related reroute advisories*

Reroute advisories due to volume constraints are implemented by Air Traffic Controllers when air traffic demand in an area of the airspace exceeds its capacity. Since air traffic flow can be predicted with flight plans and airline schedules, so can volume-related reroutes. Being able to predict the issuance of volume-related reroutes and eventually the type of the reroute advisory (Required, Recommended, FYI) will assist flight operators and traffic flow management personnel to plan routes more efficiently to improve the general efficiency of the National Airspace System.

D. Review of prior research related to rerouting advisories

Few studies have been conducted that assess the efficiency of recommended reroutes. Most of the prior research related to rerouting advisories focused on developing algorithms that created reroutes, and improved, or created, new processes to affected reroutes.

1. *Rerouting algorithms*

- “ARTCC Initiated Rerouting”, 2006 [7]: This effort analyzed the process of reroute planning and execution. The ultimate goal was to develop a new rerouting process for ARTCC in order to increase the common situational awareness for potential local reroutes. This would allow NAS users to submit reroute alternatives and increase the automation support for ARTCC to identify, assess and execute a reroute. Even though the experimental results validated the model, it is interesting to note that this work was completed before the implementation of the Traffic Flow Management System (TFMS), which is currently used by the FAA.
- “Robust Air Traffic Control Using Ground Delays and Rerouting of flights”, 2009 [8]: This effort assessed the efficiency of ground delays and rerouting advisories for three scenarios that may affect the performance of the National Airspace System: loss of ARTCC, loss of a link between two ARTCCs, and isolation of an ARTCC over a period of time. The main metric used to assess the efficiency was the number of aircraft that needed to be diverted in order to restore the performance of the National Airspace System. This work involved developing two models comprised of seven and twenty ARTCCs, respectively. In both cases, it appeared that ground delays and rerouting were efficient at mitigating the effect of an ARTCC going down in the NAS. However, from a computational point of view, a ground delay-based optimization approach is significantly less complex than a rerouting optimization approach, which generally results in a nonlinear programming problem.
- “Pilot Convective Weather Decision Making in En Route Airspace”, 2012 [9]: This effort examined the strategic aspects of pilots’ decisions during a weather avoidance process in the tactical time frame (0 – 2 hours from the incident). The goal of this study was to implement an algorithm to increase the automation level of a rerouting process involving pilots and Air Traffic Controllers. This work involved asking eighteen transport pilots to participate in lab studies where they were presented with a weather encounter scenario in an en-route environment. In particular, these pilots were asked to modify the planned trajectory in the event that they found it unsafe given the weather forecast. Results of these simulations showed that pilots were more willing to trade safety for flight efficiency even if it implied not respecting FAA guidelines on separation assurance. The scope of this study can be extended to non-weather related reroutes.

2. *Traffic Management Initiatives Statistics*

- “Aggregate Statistics of National Traffic Management Initiatives”, 2010 [10]: This effort aggregated and analyzed data from the National Traffic Management Log to provide a set of statistics on the implementation of Traffic Management Initiatives. Extracted statistics on reroutes focused on ranking points such as airports, waypoints and nav aids according to the number of reroutes they are affected by. Results showed that the five points most affected by reroutes were airports (Denver, Newark, Dallas, JFK and La Guardia). Denver and New York area airports (EWR and JFK) were affected by roughly two reroutes per day for instance. This effort also consisted in developing a visualization tool aimed at increasing both understanding and situational awareness of Traffic Management Initiatives. While this effort statistically described Traffic Management Initiatives, it did not assess the efficiency or the relevance of Traffic Management Initiatives.

E. Research Objectives

The review of prior research conducted on Traffic Management Initiatives, and reroutes in particular, highlights some limitations and gaps. First, most of the research conducted so far focused on developing and/or improving the

traffic management rerouting process. This was achieved by either developing a new reroute creation algorithm or determining the most efficient reroute using an optimization algorithm. Assessing the relevance of current reroutes has not been explored yet. This research aims to address this limitation by extracting statistics from up-to-date datasets, and identifying trends and patterns that may lead to the assessment of the relevance of reroutes.

Second, most of previous efforts have only focused on weather-related reroutes, ignoring other causes such as air traffic constraints. This research aims to address this limitation by analyzing the relevance of all recommended reroute advisories.

Third, previous efforts have not distinguished the urgency level of reroutes: required, recommended or FYI. This may have occurred because previous efforts did not have access to the data needed for such a study. This research aims to address this limitation by extracting reroute advisories from Traffic Flow Management System (TFMS) datasets and analyzing recommended reroutes.

Finally, no work has been conducted on the prediction of the issuance of volume-related reroute advisories. This research aims to fill this gap by benchmarking Machine Learning algorithms and developing a model to predict the issuance and eventually the type of reroute advisories, caused by volume constraints. The remainder of this paper will focus on highlighting the methodology used for this research, results, and concluding remarks.

III. Methodology

This section highlights the steps identified, and eventually implemented, to achieve the objectives of this research.

A. Assessment of the compliance of recommended reroute advisories by flight operators

The methodology used to assess the compliance of recommended reroute advisories by flight operators is as follows:

1. Data identification and acquisition

System Wide Information Management Flight Data Publication Service (SFDPS)

The System Wide Information Management Flight Data Publication Service (SFDPS) dataset provides en-route flight data to National Airspace System stakeholders. It allows stakeholders to receive and process real-time data for informational, analytics, research or any other purposes related to air traffic over the NAS. SFDPS provides Service-Oriented (SOA) message patterns for publishing data from the En-Route Automation Modernization (ERAM) system. ERAM data is transmitted through the Host Air Traffic Management (ATM) Data Distribution System (HADDS), which is part of the En-route Data Distribution System (EDDS). These systems are located at each of the 21 Air Route Traffic Control Centers (ARTCC) in the contiguous United States [11]. SFDPS messages are divided into three subsets in FIXM format, which captures flight and flow information that is globally standardized [12]:

- SFDPS Derived Messages: These messages are created by SFDPS to provide answers to customer requests or to provide system status information
- Reconstitution Messages: Reconstitution Messages are received from the Host Air Traffic Management (ATM) Data Distribution System (HADDS) and provide information about the Database Record Transfer (DRBT)
- Flight Data Messages: These messages include any relevant information about each individual flight in the National Airspace System such as flight plans, track data for active flights, arrival and departure information, etc.

Each of these message groups consists of different messages containing different information. Flight Data Messages such as Track Information (TH_FIXM) messages were used for this research as they contained data specific to individual flights such as tracking positions, altitude, speed, flight plans, etc [11]. The following fields were extracted from the Track Information (TH_FIXM) messages [11]:

- propMessageType: Specifies the message type received from the HADDS
- propFlightId: Specifies flight numbers
- propOrigin: Specifies the origin of flights
- propDestination: Specifies the destination of flights
- propSentTime: Specifies the time at which the message was sent from SFDPS to the NAS Enterprise Messaging System (NEMS)
- arrivalTime: Specifies the expected arrival time of flights

- departureTime: Specifies the departure time of flights
- flightState: Specifies the status of the flight. This can be either Active, Cancelled, Dropped, Landed or Proposed
- trackPosition_23d: specifies the real-time position of the flight as a latitude/longitude pair
- reportedAlt_54a: specifies the reported altitude of the aircraft in hundreds of feet

Traffic Flow Management System (TFMS)

The Traffic Flow Management System (TFMS) predicts, at both national and local scales, traffic surges, gaps, and volume based on current and anticipated airborne aircraft. Traffic management specialists evaluate the projected flow of traffic into airports, sectors, and fixes, and then implement the least restrictive action necessary to ensure that traffic demand does not exceed system capacity [13]. TFMS also provides Aircraft Situation Display (ASDI) data such as aircraft scheduling, routing and positional information. TFMS is comprised of two subsets: TFMS Flight and TFMS Flow [14].

- **TFMS Flight** provides data related to flights being managed by TFMS and is made up of the following elements: Flight plan data and potential updates and amendments, departure & arrival time notifications, flight cancellations, boundary crossings, track position records, flight management, NAS common situational model data, and flight table manager deltas
- **TFMS Flow** provides the definition of Traffic Management Initiatives, changes to the definitions, and their cancellations. TFMS Flow is comprised of the following messages: Traffic Management Initiative definitions, Ground Delay Program / Unified Delay Program, Airspace Flow Program, Collaborative Trajectory Options Program, Flow Constrained Area / Flow Evaluation Area definitions, ATCSCC advisories, restrictions, airport runway configuration and rates, airport deicing status, and route availability planning tool timeline data

For the purpose of this research, General Advisory (GADV) messages, which are contained in the TFMS Flow subset, were extracted and used because they contain recommended reroute advisories. The following fields contained in GADV messages provide relevant information regarding recommended reroutes:

- fcm:advisoryNumber: Specifies the advisory ID number
- fce:startTime: Specifies the start time of the advisory
- fce:endTime: Specifies the end time of the advisory
- fcm:advisoryTitle: Coded sentence that summarizes the advisory
- fcm:advisoryText: Contains extensive information on the advisory including the constrained area, the reason, the probability of extension, the affected airports, the new routes, and any relevant remark

Recommended reroutes are detailed in the 'advisoryText' section of the advisory and can appear in two formats. The first format, as seen in Figure 4 is mostly used for required reroute advisories and occasionally for recommended reroute advisories.

ROUTES		
ORIG	DEST	ROUTE
KMSP	KJFK	DLL HASTE DAFLU J70 LVZ LENDY6
ATL	DTW	VXV J91 HNN DJB GEMNI4
LAS	ATL	INW J86 ELP ABI J4 MEI DUUCK PRICI RAGGZ1

Fig. 4 Example of reroutes defined in the first format with origin, destination and route

The second format, which is mostly used for recommended advisories, defines route origin segments and route destination segments that need to be compiled in order to form a complete route from origin to destination, as seen in Figure 5. There was a need to further analyze and process reroutes in this format because occasionally, an airport is listed in both origin and destination segment routes in the advisories. Also, reroutes between two close airports were identified as unnecessary as they led to the creation of long and inefficient routes, as seen in Figure 6. In order to avoid this, reroutes between two airports of distance five times more than the direct distance between the two airports were not

used for this research.

ROUTES	
FROM ORIG	ROUTE ORIGIN SEGMENTS
EWR	DIXIE PREPI UNYAD OWENZ POPPN OHRYN BEHHR WEBBB HOBHOH PAEPR M201 HANRI
FLL FXE	ZAPPA PERMT AR16 EMCEE M201 PAEPR
MIA TMB	VALLY PERMT AR16 EMCEE M201 PAEPR
PBI BCT SUA	PBI A699 PERMT AR16 EMCEE M201 PAEPR
TO DEST	ROUTE DESTINATION SEGMENTS
BCT	HANRI M201 JENKS AR19 AYBID CAYSL4
CDW MMU	PAEPR HOBHOH SILLY STINK YAALE YETTI MOUGH DONAA OWENZ CYN GXU RBV V249 METRO
FLL	HANRI M201 BAHAA AR21 CRANS FISEL7

Fig. 5 Example of reroutes defined in the second format with origin, origin segments and destination, destination segments

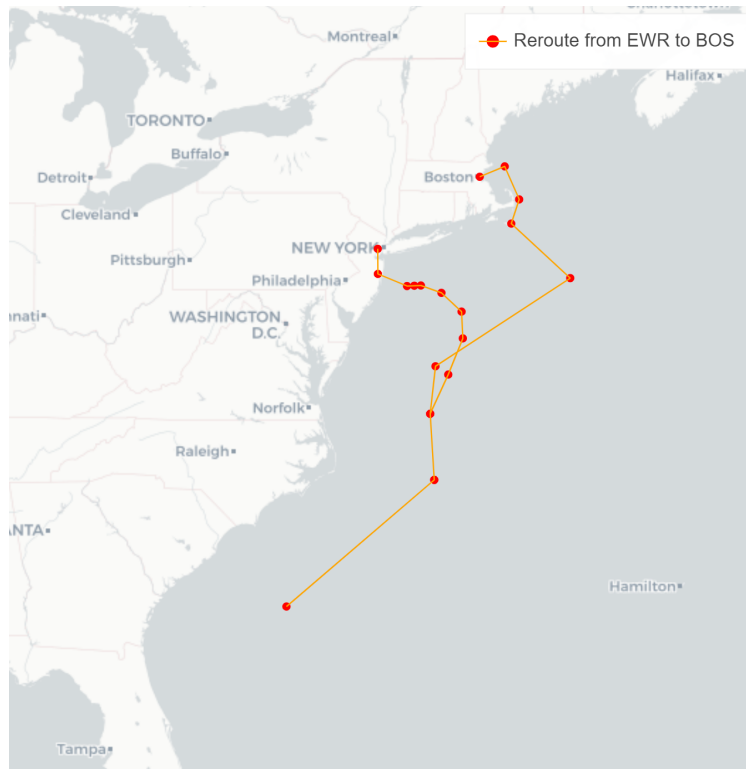


Fig. 6 Example of non-sensical reroute defined with second reroute format

Fix/Reporting Point/Waypoint Dataset

This dataset was extracted from the National Airspace System Resource (NASR) website [15], and is updated every 28 days. It lists all fix waypoints and provides their record identifier, ICAO region code, geographical coordinates, and other relevant information.

Nav aids Dataset

This dataset was extracted from the National Airspace System Resource (NASR) website [15], and is updated every 28 days. It lists all Navigational Aids (Nav aids) and provides their record identifier, facility type, geographical coordinates, and other relevant information.

2. Data Processing

The SFDPS and TFMS datasets are in Flight Information Exchange Model (FIXM) format, which is appropriate for storing and sharing large amounts of aviation data. However, it is not appropriate for data analytics purposes. To use these datasets, for analytical purposes, there was a need to parse them into a much more usable and appropriate format such as Comma-Separated Values (CSV). The main advantages of the CSV format over FIXM are its compatibility with Python modules used for data analytics purposes such as Pandas. The datasets are stored as hourly files by the FAA and are comprised of all messages generated within the hour. Furthermore, SFDPS and TFMS datasets have schemas which dictate the datasets' structure. Using schemas is critical to make sure that all required fields are extracted in their correct formats. A python parser developed by Mangortey et al. [16] for the TFMS dataset, and one developed for the SFDPS dataset were used to achieve the objectives of this research.

As stated previously, reroute advisories are issued whenever an ARTCC identifies constraint(s) and assigns new routes to affected flights. However, constraints such as bad weather conditions, aircraft congestion or equipment failures may change, leading to updates to the scope of reroute advisories. These updates are issued as new reroute messages with new advisory numbers. However, occasionally, reroute advisories are still generated after being updated, making them invalid. Therefore, there was a need to exclude initial reroute advisory messages after they had been updated. In order to do this, the end time of the initial reroute advisory was set as the start time of the updated advisory. This ensured that the effective periods of both the initial and updated advisories did not overlap and were treated as two separate reroute advisories.

The Fix/Reporting Point/Waypoint and Nav aids datasets were extracted in text format from a FAA database, hence eliminating the need to process them into a different format.

3. Data Fusion

The data fusion process involved understanding how the different datasets and their features are related to each other. The first common feature between the two datasets (SFDPS, TFMS) is time. The duration of recommended reroute advisories was extracted from TFMS while flight departure and arrival times was extracted from SFDPS. Another common field between the two datasets is the airport. Recommended reroute advisories from TFMS state the affected airport(s) and/or area(s) of the airspace. The origin and destination airport(s) of flights were extracted from SFDPS. The SFDPS and TFMS datasets were fused using the following steps:

- 1) From TFMS, extract the affected airports (departure and arrival), the effective period of the advisory and the suggested routes from recommended reroute advisories messages
- 2) From SFDPS, extract messages of all flights flying to and from the airport(s) affected by the reroute advisory
- 3) Extract and order TH_FIXM messages by generation time for each affected flight
- 4) Create a list of the path taken by the flight from origin to destination using flight coordinates, using data from the extracted TH_FIXM messages

4. Data Analysis

To analyze the relevance of recommended reroute advisories, actual flight paths were compared to recommended reroute advisories. This was done in two ways: 1) Using a Flight Plan Approach by comparing flight plans to recommended reroutes, and 2) using a Tracking Flights Approach by tracking flights using their coordinates and comparing them to the path of recommended reroutes.

A. Flight Plan Approach

The following steps were taken to compare flight plans and recommended routes:

- Extract and order all waypoints, nav aids and airways chronologically from the flight plan

- Extract and order all waypoints, navaids and airways chronologically from the recommended route
- Identify common points in both routes
- Compare both routes by examining if the points in the flight plan are in the same sequence as those in the recommended route

B. Tracking Flights Approach

The following steps were taken to compare flight tracks and recommended routes:

- Extract all fixes from the recommended route
- Extract, from a FAA dataset gathering positions of all American fixes, the geographical coordinates of the fixes extracted at the previous step
- Store these coordinates in a list in the same order than the recommended route
- Implement an algorithm and a set of metrics to compare the path of a flight and the path of the recommended route

According to FAA Subject Matter Experts (SMEs), pilots do not always keep their flight plan updated which might invalidate results obtained from the flight plan approach. Thus, the second approach based on tracking flights was implemented and used for this research. This approach involved comparing the trajectory flown by a flight with the path of the recommended reroute. This was done using two different methods:

1) Polygon Method

This method is based on the generation of a polygon that can also be described as a corridor around the recommended route. The width of each polygon was 10 nautical miles (18.52 kilometers) based on the recommendation of Subject Matter Experts at the FAA. A Python algorithm was developed to automatically create a corridor around each reroute by generating rectangles of 10 nm width and 1nm long around each element of the reroute. Each of the rectangles was oriented to the next element such that the polygon was generated by linking all of the individual rectangles. Figure 7 is an example of a reroute between Newark Liberty International Airport and Fort Lauderdale-Hollywood International Airport with a polygon around the reroute.

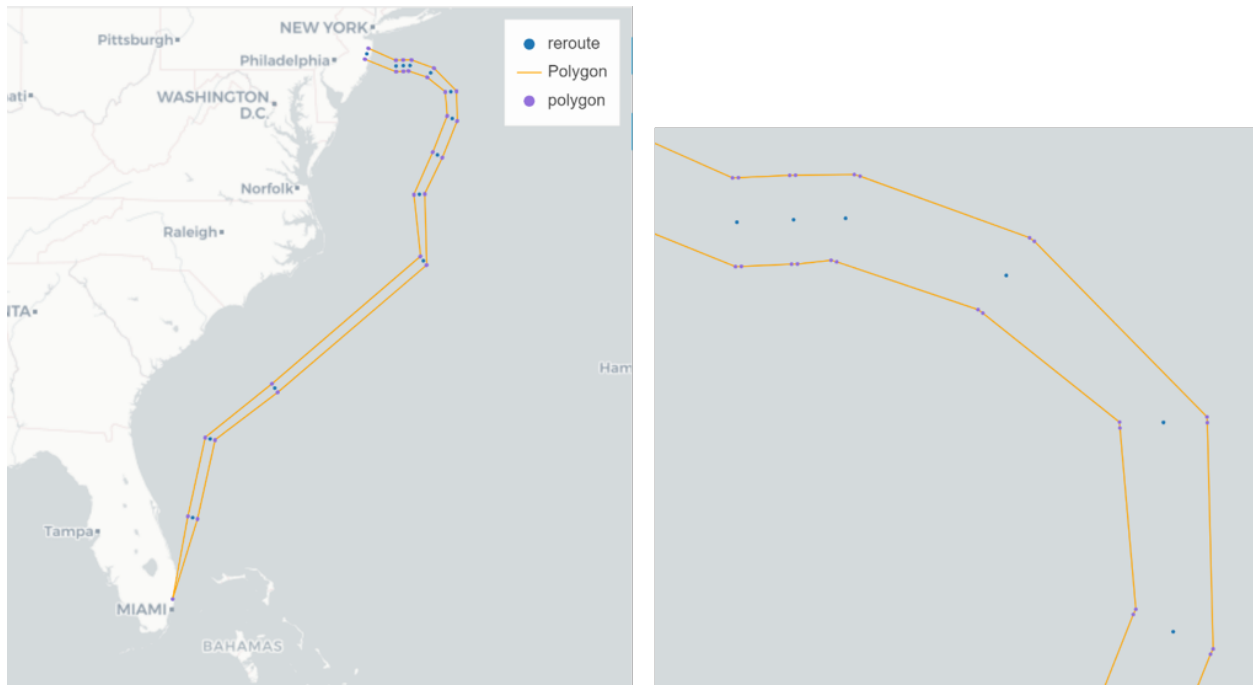


Fig. 7 Polygon around recommended reroute EWR-FLL, November 6th, 2018

Three metrics were developed and eventually used to assess the compliance of flights to recommended reroutes using the polygon method:

- Metric 1: Polygon Metric

$$\text{Metric 1} = \frac{\text{Number of flight positions in the reroute polygon}}{\text{Total number of flight positions}} \quad (1)$$

- Metric 2: Flight Distance Metric

$$\text{Metric 2} = \frac{\text{Distance flown by the plane within the polygon}}{\text{Total distance flown by the plane}} \quad (2)$$

- Metric 3: Reroute Distance Metric

$$\text{Metric 3} = \frac{\text{Distance flown by the plane within the polygon}}{\text{Length of the reroute}} \quad (3)$$

2) Circle Method

This method is based on the generation of geographical circles of 5nm-radius around each fix and navaid of the reroute, as seen in Figure 8. The circles were defined with a 5nm radius to be coherent with the polygon width selected in the Polygon Method. A waypoint is then said to be "validated" when at least one position of the flight is found within the circle generated around the waypoint.

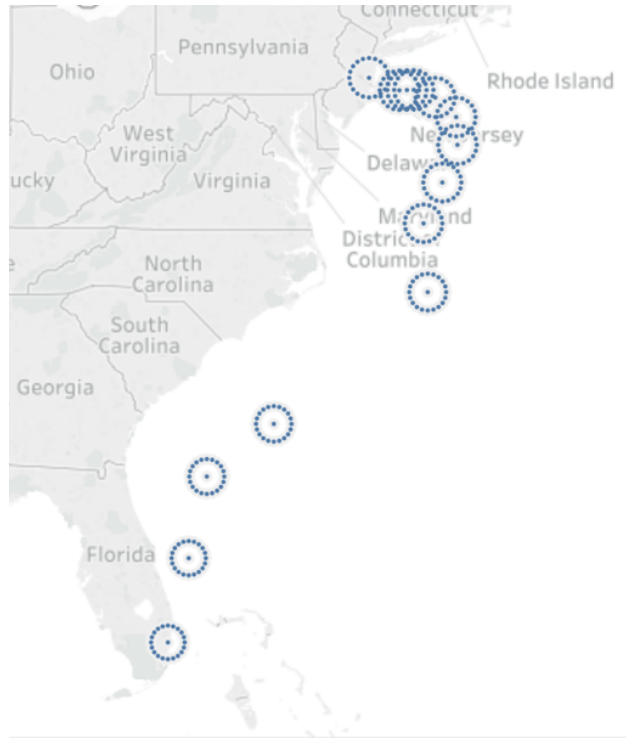


Fig. 8 Circles around recommended reroute EWR-FLL, November 6th, 2018

Based on the notion of validated waypoints, a fourth metric was developed to assess the compliance of flights to reroutes:

$$\text{Metric 4} = \frac{\text{Number of waypoints validated by the flight}}{\text{Total number of waypoints of the reroute}} \quad (4)$$

B. Predicting the issuance of volume-related reroutes

The methodology used to predict the issuance of volume-related reroutes is as follows:

1. Data Generation

Volume-related reroute advisories are issued by Air Traffic Controllers when air traffic demand is expected to exceed the capacity of an affected area in the airspace. Being able to predict the issuance of reroute advisories may help airlines and air traffic controllers to better plan their schedules, and potentially reduce flight delays and durations. All reroute advisory types (Required, Recommended, FYI) issued between January and April 2017 were extracted from the Traffic Flow Management System (TFMS) for the development of the prediction models. Hourly traffic count at the different ARTCCs was not available for this research. Consequently, the hourly traffic count data was generated using daily traffic counts extracted from the FAA's Air Traffic Activity Data System (ATADS) database [17]. Figure 9 shows the distribution of flights per hour across the NAS on July 27th, 2017, which was used to compute the percentage of flights per hour. Hourly traffic count per facility was then generated by multiplying the daily traffic count per facility with the hourly percentage shown in Figure 10.

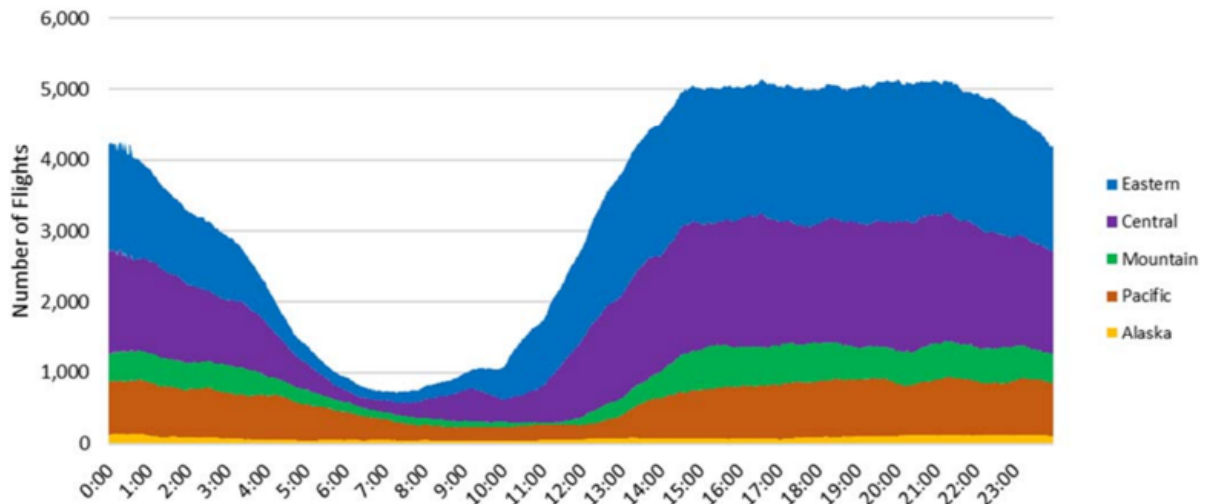


Fig. 9 Number of flights in the U.S. airspace per hour (GMT) and time zone on July 27th, 2017 [18]

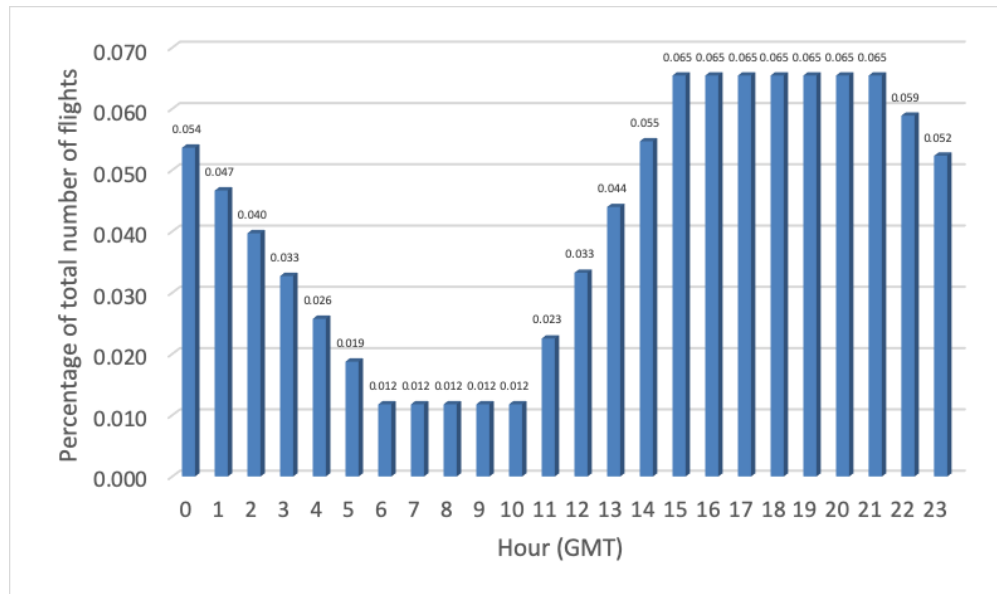


Fig. 10 Distribution of the number of flights per hour (GMT) in the US airspace

In order to develop the prediction models, the TFMS and hourly traffic count datasets were fused using this process:

- 1) Extract all reroute advisories due to volume issued between January and April 2017
- 2) Collect hourly traffic count per facility between January and April 2017
- 3) For each hour and facility, check from step 1 if a volume-related reroute has been issued
- 4) Build a data matrix stipulating for each hour and each facility if a reroute has been issued and if yes which type

Since actual traffic count data was not used for this research, the prediction models presented in this work serve as a proof of concept that can be validated with actual traffic count data.

2. Prediction Model Development

The Decision Trees, Naive-Bayes, Nearest Neighbor, Random Forests, Support Vector Machine, Bagging Ensembles, and Boosting Ensembles were benchmarked to identify the best suited technique for predicting:

- 1) The issuance of volume-related reroute advisories without specifying the reroute type
- 2) The type of volume-related reroute advisory

In order to develop the prediction models, the data was randomly split into three sets. Half of the data was assigned to a training set which was used to train the models, one-fourth of the data was assigned to a validation set which was used to tune and refine the models, and one-fourth of the data was used to test and evaluate model performance.

3. Prediction Model Evaluation

Many different metrics exist to evaluate a Machine Learning model. Classification problems are typically evaluated using results of a confusion matrix, as seen in Table 1. A confusion matrix is a table summarizing four different combinations of predicted and actual values that are used in the computation of most performance metrics such as Recall, Precision, Specificity, Accuracy, etc. [19].

Table 1 Confusion Matrix

	Predicted: No	Predicted: Yes
Actual: No	True Negative (TN)	False Positive (FP)
Actual: Yes	False Negative (FN)	True Positive (TP)

Based on this Confusion Matrix, many performance metrics can be defined. The following metrics were used to assess the performance of each Machine Learning technique:

- 1) Sensitivity: This measures the number of correct positive predictions over the total number of actual positive values that should have been predicted if the model was perfect

$$Sensitivity = \frac{TP}{TP + FN} \quad (5)$$

- 2) Specificity: This measures the proportion of correct negative predictions over the actual number of negative values. Specificity is the exact opposite of Sensitivity.

$$Specificity = \frac{TN}{TN + FP} \quad (6)$$

- 3) Matthew's Correlation Coefficient: This metric has a maximum value of 1 corresponding to perfect predictions, and a minimum value of -1 corresponding to total contradiction.

$$MCC = \frac{TP * TN - FP * FN}{\sqrt{((TP + TN) * (TP + FP) * (FP + TN) * (TN + FN))}} \quad (7)$$

The MCC was used to evaluate the performance of the techniques for the first prediction because this prediction has only two classes: Reroute and No Reroute.

- 4) Kappa statistic: This is very often used to measure the performance of algorithms on multi-class and imbalanced datasets.

$$K = \frac{P_0 - P_E}{1 - P_E} \quad (8)$$

P_0 is the observed value and P_E is the expected value. The Kappa statistic was used to evaluate both prediction models.

IV. Results

A. Assessment of the compliance of recommended reroute advisories by flight operators

Results presented below are based on the analysis of data from January to April 2017. 4,974 recommended reroutes and 22,016 affected flights were used to assess the compliance of recommended reroute advisories by flight operators during this period.

1. Evaluation of Metrics

The metrics highlighted in the previous section were evaluated to identify the best suited metric for assessing the relevance of recommended reroutes.

- **Metric 1: Polygon Metric**

Metric 1 is defined as the ratio of the number of flight positions in the reroute polygon to the total number of flight positions.

$$\text{Metric 1} = \frac{\text{Number of flight positions in the reroute polygon}}{\text{Total number of flight positions}} \quad (9)$$

Metric 1 relies largely on the distribution and frequency of flight track data. Usually, reroutes between airports affect areas in the route other than the takeoff and landing airports. Thus, if the frequency of flight track data is higher during the takeoff and landing segments of flight compared to the other segments of flight, the metric might incorrectly indicate metric compliance because many of the flight's track positions were located in the polygon. Figure 11 shows the flight path of JBU83 and the polygon representing the recommended reroute. It can be clearly seen that the flight did not follow the recommended reroute. However, Metric 1 provides a compliance score of 40% which is largely due to the large number of flight track data being generated during the takeoff and landing segments of the flight. Thus, it can be concluded that Metric 1 is not an appropriate metric for assessing compliance of recommended reroute advisories by flight operators.

- **Metric 2: Flight Distance Metric**

Metric 2 is defined as the ratio of the distance flown by the plane within the polygon to the total distance flown by the plane.

$$\text{Metric 2} = \frac{\text{Distance flown by the plane within the polygon}}{\text{Total distance flown by the plane}} \quad (10)$$

This metric compares the distance flown inside the polygon, representing a recommended reroute to the total distance flown during the flight. From Figure 11, it can be seen that the flight did not follow the recommended reroute. This observation is validated using Metric 2 as it provides a compliance score of 7%. Metric 2 was used to evaluate the compliance of over 22,000 flights and it validated observations from diagrams such as Figure 11.

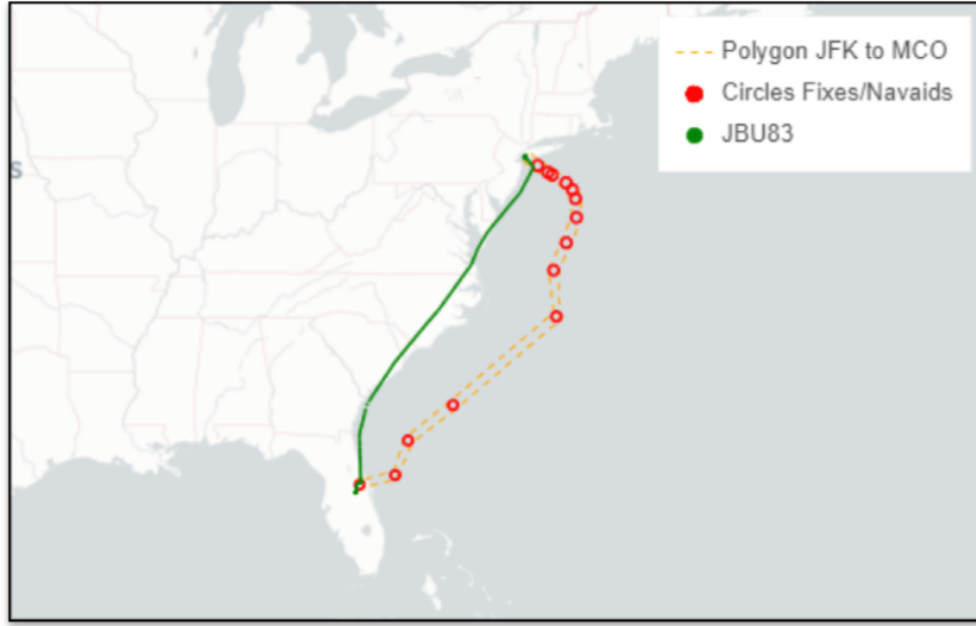


Fig. 11 Comparison of flight path of JBU83 and a recommended reroute

- **Metric 3: Reroute Distance Metric**

Metric 3 is defined as the ratio of the distance flown by the plane within the polygon to the length of the reroute.

$$\text{Metric 3} = \frac{\text{Distance flown by the plane within the polygon}}{\text{Length of the reroute}} \quad (11)$$

This metric compares the distance flown inside the polygon to the entire length of the recommended reroute. Similarly to Metric 2, Metric 3 validated observations from diagrams such as Figure 11 after the compliance of over 22,000 flights was assessed.

- **Metric 4: Circle Metric**

Metric 4 is defined as the ratio of the number of waypoints validated by the flight to the total number of waypoints of the reroute.

$$\text{Metric 4} = \frac{\text{Number of waypoints validated by the flight}}{\text{Total number of waypoints of the reroute}} \quad (12)$$

Metric 4 largely depends on the distribution of waypoints in a recommended reroute. Indeed, Metric 4 indicates poor compliance when the number of waypoints in a recommended reroute are concentrated in an area not flown by the flight. Figure 12 compares the flight path of JBU2106 and a recommended reroute from FLL to EWR. It can be seen that most of the waypoints of the recommended reroute are concentrated close to EWR. It can also be seen that the flight largely followed the recommended reroute. However, Metric 4 indicated a compliance of 44% particularly because the flight did not go through the waypoints close to EWR. Thus, it can be concluded that metric 4 is not an appropriate metric for assessing compliance of recommended reroute advisories by flight operators.

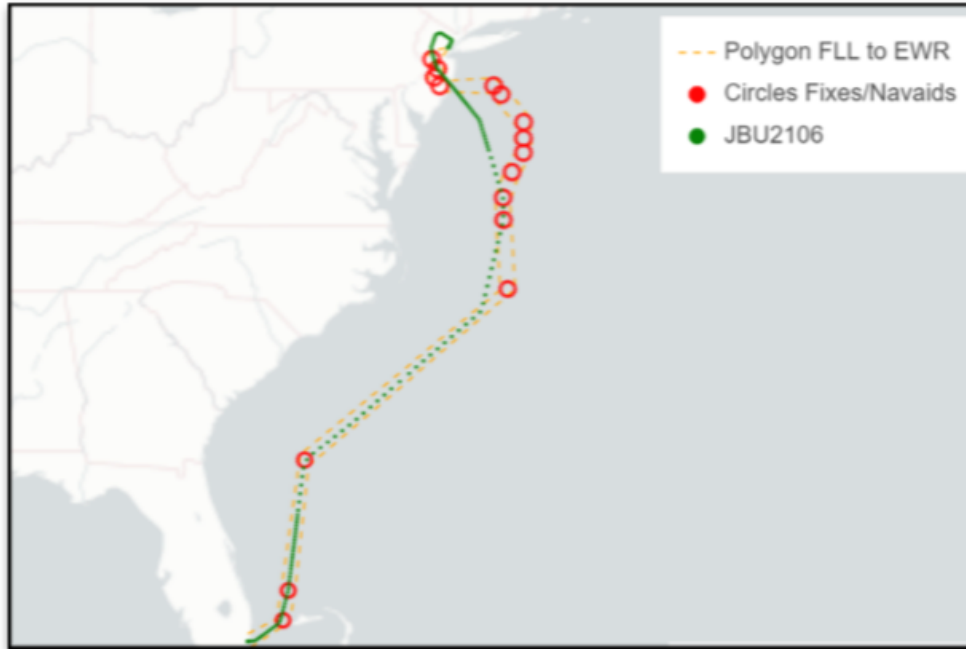


Fig. 12 Comparison of flight path of JBU2106 and a recommended reroute

2. Summary Of Observations

This section provides a summary of observations made from the analysis of recommended reroute advisories.

- **Analysis of compliance of recommended reroute advisories by flight operators**

Metric 2 was identified as the best suited metric for assessing the compliance of recommended reroute advisories by flight operators. As mentioned, 4,974 recommended reroutes affecting 22,016 flights were issued between January and April 2017. The compliance of each of these flights to their recommended reroutes was computed using Metric 2 and analyzed. Figure 13 provides a summary of the compliance of recommended reroutes by flight operators using Metric 2. It can be seen that 53.26% of affected flights had compliance scores between 0 and 1%. Only 2.63% of affected flights had a compliance score over 50%. Both of these observations indicate that in general, recommended reroutes are not followed by flight operators. There was, however, a need to identify a threshold for assessing compliance of recommended reroutes by flight operators. The threshold for assessing compliance of recommended reroutes by flight operators was identified by analyzing scores obtained using Metric 2, and analyzing the comparison between flight paths and recommended reroutes, as seen in Figure 14. The analysis of the 22,016 flights affected by recommended reroutes revealed that the majority of flights with scores between 50% and 60% followed most, if not all, of their recommended reroutes.

Further analysis of flights with scores between 50% and 60%, as seen in Figure 15 helped determine that an appropriate score for determining the compliance of recommended reroutes by flight operators is **55%**. Table 2 shows that 97.9% of affected flights had a compliance score less than 55% while 2.1% of flights had a compliance score greater than 55%. Thus, it was concluded that 2.1% of flights followed recommended reroutes between January and April 2017.

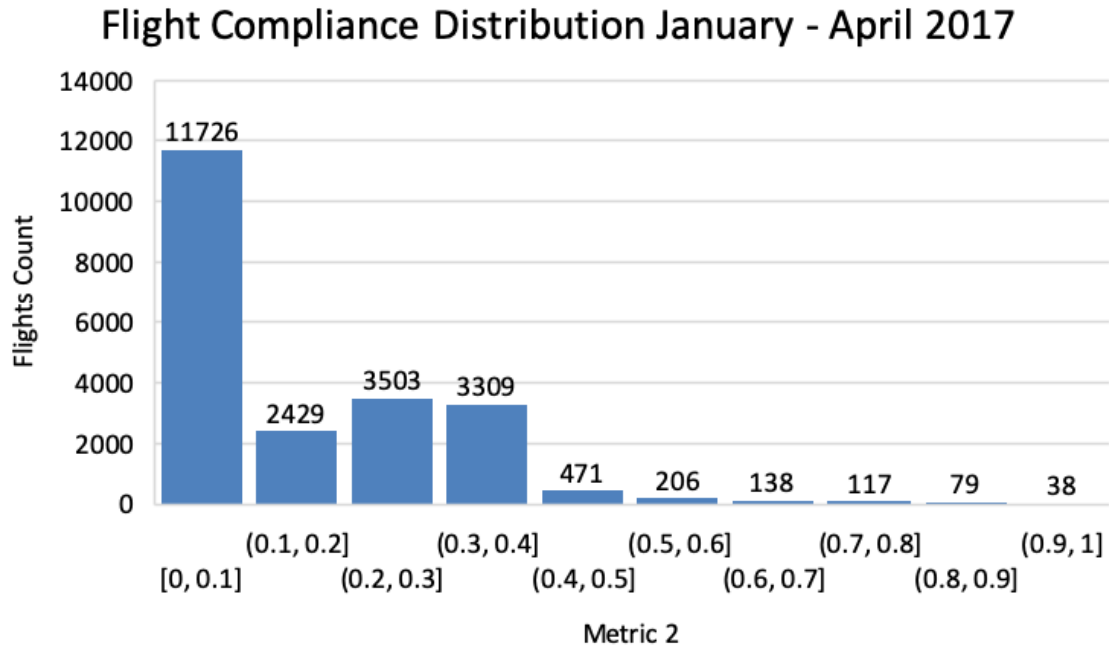


Fig. 13 Summary of the compliance of recommended reroutes by flight operators using Metric 2

Table 2 Breakdown of flights with by compliance threshold

Metric 2	Number of flights
< 55%	21,559
> 55%	457

- **Analysis of flight compliance analysis by distance flown**

Flights impacted by recommended reroutes between January and April 2017 were analyzed to identify trends and patterns. Table 3 shows that over 64% of flights affected by recommended reroutes during this time period flew distances between 810 - 1080nm. It also shows that these flights had the highest average and median Metric 2 scores. Consequently, mid-distance ranges (810 - 1080nm) was identified to be the most common range for recommended reroutes. Mid-distance ranges were also identified to have the highest average and median compliance scores. Finally, analysis of the data revealed that recommended reroutes over long distances had poor compliance.

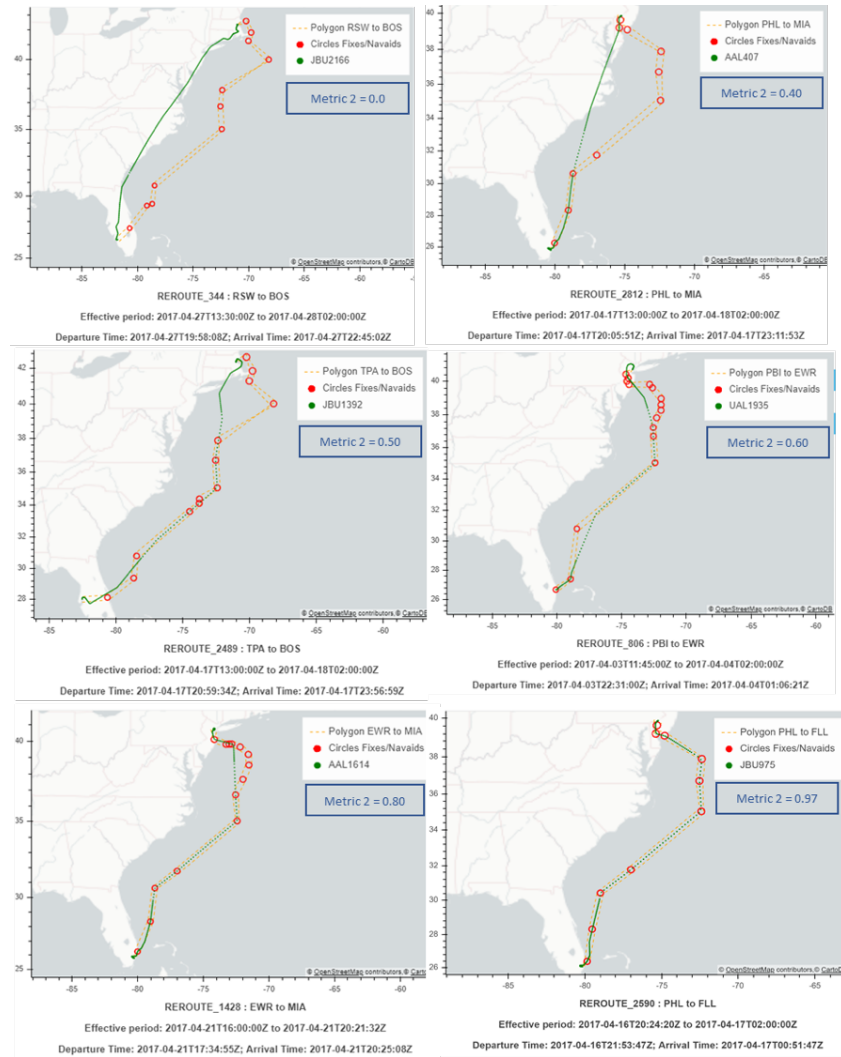


Fig. 14 Comparison of flight paths and recommended reroutes using Metric 2

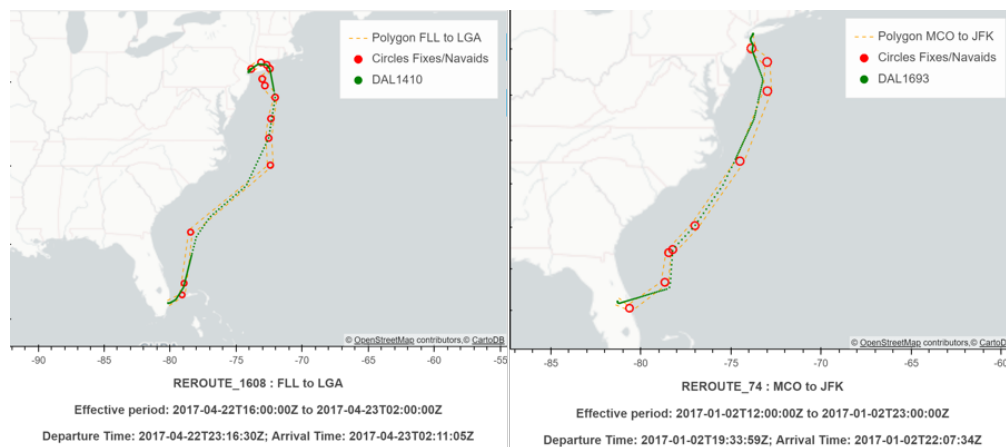


Fig. 15 Comparison of flight paths and recommended reroutes with Metric 2 scores between 50% and 60%

Table 3 Number of flights affected by Recommended Reroutes with average and median Metric 2 scores per distance flown

Distance (nm)	Number of affected flights	Average compliance score (%)	Median compliance score (%)
0 - 270	35	17	15
270 - 540	2473	6	3
540 - 810	2996	9	7
810 - 1080	14165	18	15
1080 - 1340	2259	18	7
1340 - 1620	59	16	2
1620 - 1890	18	3	2

B. Predicting the issuance of volume-related reroutes

As mentioned previously, Machine Learning algorithms were benchmarked to identify the best suited algorithm for predicting 1) the issuance of volume-related reroute advisories without specifying the reroute type, and 2) the type of volume-related reroute advisory. The benchmarked algorithms were: Decision Tree, Nearest Neighbor, Naïve-Bayes, Support Vector Machines, Bagging Ensembles, Boosting Ensembles, and Random Forest. Both models were developed using R.

1. Prediction of the issuance of volume-related reroute advisories without specifying the reroute type

The predictors for this prediction problem are hourly traffic count, the facility, the month, the day, and the hour. Reroute and No Reroute were the targets of this prediction model. Figure 16 provides a comparison of the performance of various Machine Learning algorithms using Matthew's Correlation Coefficient. Figure 16 shows that the Random Forest algorithm is the best suited algorithm for predicting the issuance of volume-related reroute advisories using Matthew's Correlation Coefficient.

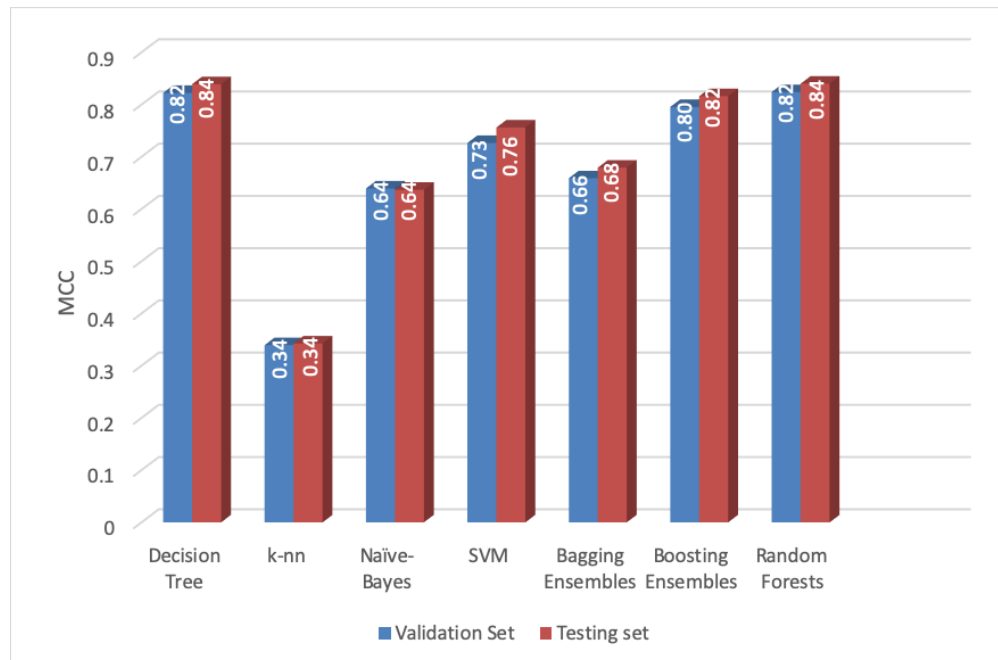


Fig. 16 Comparison of Machine Learning algorithms using Matthew's Correlation Coefficient

2. Prediction of the type of volume-related reroute advisory

The predictors for this prediction problem are hourly traffic count, the facility, the month, the day, and the hour. The analysis of reroute advisories from January - April 2017 revealed that only recommended and required volume-related reroutes were implemented. Thus, No Reroute, Recommended Reroute, and Required Reroute were the targets of this prediction model. Figure 17 provides a comparison of the performance of various Machine Learning algorithms using Kappa's Statistic. Figure 17 shows that the Decision Tree algorithm is the best suited algorithm for predicting the issuance of volume-related reroute advisories using Kappa's Statistic.

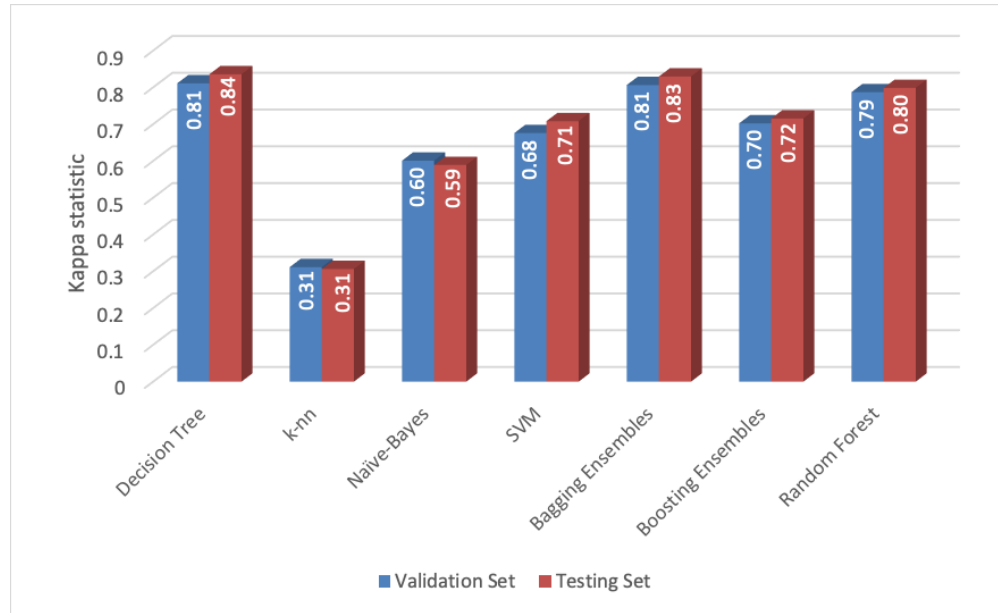


Fig. 17 Comparison of Machine Learning algorithms using Kappa's Statistic

V. Conclusion

Reroute advisories are Traffic Management Initiatives issued when Air Traffic Controllers (ATC) identify constraints in the National Airspace System, and assign new routes to affected flights. These new routes are defined by their level of urgency: Required, Recommended or For Your Information (FYI). Over the years, efforts have been made to reduce the impact of reroutes on flight operations. Previous studies have focused on the definition and optimization of reroutes, but have not analyzed the compliance of recommended reroute advisories by flight operators. Furthermore, little to no work has been done to predict the issuance of reroutes or the type of reroute. Consequently, the objectives of this work was two-fold: 1) Assess the relevance of existing recommended reroutes, and 2) predict the issuance and type of volume-related reroute advisories. This was achieved by 1) fusing data from relevant datasets, extracting statistics, and identifying trends and patterns within the data, and 2) developing models to predict the issuance of volume-related reroute advisories. Two approaches and four metrics were developed and analyzed to assess the compliance of recommended reroutes by flight operators. The best suited approach and metric were then identified, revealing that only 2.1% of the 22,016 flights considered followed recommended reroutes between January and April 2017. Further analyses also revealed that flights covering distances between 810 and 1080nm were impacted the most by recommended reroutes. In addition, these flights also had the highest average and median compliance scores per distance flown. Finally, the benchmarking of Machine Learning techniques revealed that the best suited algorithms for predicting the issuance and type of volume-related reroute advisories is Random Forests and Decision Tree algorithms, respectively.

Acknowledgments

The authors wish to acknowledge the support of FAA analysts and researchers for spearheading this research. Particularly, they wish to thank Mike Paglione, Anya Berges, Nelson Brown, and the rest of the Big Data Analytics Working Group (Big DAWG) team at the FAA Tech Center for helping shape this research and providing valuable feedback. The views and findings expressed in this document are those of the authors only, and do not represent those of the FAA.

References

- [1] Federal Aviation Administration, “National Airspace System Overview,” 2015. URL https://www.faa.gov/air_traffic/nas/nynjphl_redesign/documentation/feis/media/Appendix_A-National_Airspace_System_Overview.pdf.
- [2] Vanhoenacker, M., “In Flight, Image from Navtech,” 2015. URL <https://www.nytimes.com/interactive/2015/05/14/opinion/14-in-flight-mark-vanhoenacker.html>, imagefromNavtech.
- [3] National Business Aviation Association Air Traffic Services (ATS), “Traffic Management Initiatives,” 2016. URL <https://filesmart.org/resources/traffic-management-initiatives/>.
- [4] National Business Aviation Association (NBAA), “Tools used for Traffic Flow Management - Reroutes,” 2018. URL <https://nbaa.org/aircraft-operations/airspace/tfm/tools-used-for-traffic-flow-management/reroutes/>.
- [5] National Business Aviation Association (NBAA), “Tools used for Traffic Flow Management,” 2018. URL <https://www.nbaa.org/ops/airspace/tfm/tools/#terminal-tmi>.
- [6] Federal Aviation Administration, “Electronic Code of Federal Regulations,” 2018. URL https://www.ecfr.gov/cgi-bin/text-idx?c=ecfr&sid=3efaad1b0a259d4e48f1150a34d1aa77&rgn=div5&view=text&node=14:2.0.1.3.10&idno=14#se14.2.91_1123.
- [7] Lynne Fellman and Tejal Topiwala, “ARTCC-Initiated Rerouting,” *6th AIAA Aviation Technology, Integration and Operations Conference (ATIO)*, 2006. doi:10.2514/6.2006-7829.
- [8] Dutta, P. and Bhattacharya, R. and Chakravorty, S., “Robust Air traffic Control Using Ground Delays and Rerouting of Flights,” *Aerospace Computing, Information and Communication*, Vol. 6, No. 12, 2009, pp. 649–674.
- [9] Gooding, C. and Shelley, A. and Wu, S. and Johnson, W. and Duong, C., “Pilot Convective Weather Decision Making in En Route Airspace,” *ICAS 2012 - 28th Congress of the International Council of the aeronautical Science*, 2012.
- [10] Rios, J., “Aggregate Statistics of National Traffic Management Initiatives,” *10th AIAA Aviation Technology, Integration and Operations (ATIO) Conference*, 2010. doi:10.2514/6.2010-9382.
- [11] National Transportation Systems Center, “Swim Flight Data Publication Service (SFDPS) Reference Manual,” *SFDPS*, 2015.

- [12] Lepori, H., “Introduction to FIXM,” , 2017. URL <https://www.icao.int/MID/Documents/2017/SWIM>.
- [13] Federal Aviation Administration, “Traffic Flow Management Operational Support and Maintenance Statement of Work (SOW),” *TFMS*, 2010.
- [14] Federal Aviation Administration, “Java Messaging Service Description Document Traffic Flow Management Data Service (TFMData),” *TFMS*, 2016.
- [15] Federal Aviation Administration, “28 Day NASR Subscription,” , 2019. URL https://www.faa.gov/air_traffic/flight_info/aeronav/Aero_Data/NASR_Subscription/.
- [16] Mangortey, E. and Gilleron, J. and Dard, G. and Pinon, O. and Mavris, D., “Development of a Data Fusion Framework to support the Analysis of Aviation Big Data,” *AIAA Science and Technology Forum (AIAA Scitech)*, 2019. doi:10.2514/6.2019-1538.
- [17] Federal Aviation Administration, “Air Traffic Activity System (ATADS),” , 2019. URL <https://aspm.faa.gov/opsnet/sys/main.asp>.
- [18] Federal Aviation Administration, “Traffic Flow Management System,” *Office of Performance Analysis (AJR-G)*, 2018.
- [19] Sokolova, M., and Lapalme, G., “A systematic analysis of performance measures for classification tasks,” *Information Processing and Management*, Vol. 45, No. 4, 2009, pp. 427–437.