Twelfth USA/Europe Air Traffic Management Research and Development Seminar (ATM2017)

Causal Analysis of En Route Flight Inefficiency – the US Experience

Yulin Liu, Mark Hansen
Institute of Transportation Studies
University of California, Berkeley
Berkeley, CA, USA
liuyulin101@berkeley.edu, mhansen@ce.berkeley.edu

Michael O. Ball
Robert H. Smith School of Business and Institute for Systems Research
University of Maryland
College Park, USA
mball@rhsmith.umd.edu

David J. Lovell, Cara Chuang
Department of Civil and Environmental Engineering and Institute for Systems Research
University of Maryland
College Park, USA
lovell@umd.edu, cychuang@umd.edu

John M. Gulding
Federal Aviation Administration
Washington D.C., USA
john.gulding@faa.gov

Abstract— En route inefficiency is measured in terms of extra distance flown by an aircraft, above what would have been optimal under perfect conditions. Three sources of inefficiency are explored: convective weather, horizontal miles-in-trail restrictions, and winds. Historical flight records are projected onto a small set of nominal trajectories clustered from historical data, and compared against the history of the potential causal factors. Statistical models reveal the estimated influence of the factors. In this case, convective weather was the most influential factor in seeming to cause flights to deviate from what would have been a less costly trajectory. Winds and miles-in trail restrictions are also important for some origin-destination pairs, but less significant than convective weather.

Keywords- enroute inefficiency; trajectory clustering; miles-in-trail; climatological condition; linear regression; multinomial logit model

I. INTRODUCTION

Horizontal en route inefficiency, which evaluates actual flight trajectories against a benchmark trajectory, has received considerable attention in the literature. It is a fundamental measure in the European performance scheme for air navigation services [1]. In 2016, the most recent edition of the ICAO Global Air Navigation Plan (GANP) listed an en route trajectory metric as a recommended indicator for assessing Aviation System Block Upgrades (ASBUs) [2]. Although the US does not employ targets for trajectory measures, they are still used in many applications. They are core to many fuel efficiency benefit assessments and the Air Traffic Organization records Airborne Holding as part of its reportable delay in OPSNET. For 2017, it is expected that the Federal Aviation Administration (FAA), with airlines, will assess filed vs. actual flown trajectories or assess routings and flows to determine if the best filed options are available.

For all of these projects, metrics can be defined and calculated and trends assessed. However, work on linking inefficiencies to causal reasons and then determining if they can be mitigated through investment in technology, staffing or training is more difficult. As a first step, FAA performance databases would require improved methods for assigning causal factors to trajectory efficiency metrics similar to the manner in which causal factors are assigned to reportable delay in OPSNET [3] or airline reported delay in ASQP [4].

The en route phase of a flight is defined as the portion between the 40 nautical-mile circular boundary around the departure airport (D40) and the 100 nautical-mile circular boundary around the arrival airport (A100). This definition is intended to exclude the portions of the flight path that are strongly influenced by terminal operations. Ref. [5] calculates the horizontal inefficiency based on the extra distance flown in the en route phase with respect to an ideal distance known as "achieved distance", which represents the average of how much further the flight has gotten from the origin and how much closer it has gotten to the destination over the en route portion of the flight ([6], [7]). This method accounts for inefficiency resulting from entry and exit points that are not along the great circle path between origin and destination, as well as excess distance between the entry and exit points, but does not include excess distance flown between the runway and the entry/exit points. Equation (1) defines the horizontal inefficiency of a flight (HIE) where, A is the actual flown distance and H is the achieved distance.

$$HIE = \frac{A-H}{H} \quad (1)$$

We use the same methodology for calculating en route inefficiency as the US-Europe Performance report [8]. In 2015, the average horizontal en route inefficiency for flights to and from the main 34 airports was 2.92% in Europe compared to 2.83% in the US. Europe experienced continuous improvement of en route efficiency between 2011 and 2014; however, this trend reversed in 2015. The US had an overall increase in flight inefficiency from 2013 – 2015, which can be linked to airports with increased traffic levels (LAX, SEA, DAL). A fire at the

This project is sponsored by the Federal Aviation Administration.

Chicago Air Route Traffic Control Center (ARTCC) caused a peak in inefficiency during the fall months of 2014.

For our period of analysis – 2013 – the en route inefficiency was 2.91% and 2.71% for Europe and the US, respectively [5]. The 2013 report also includes statistics about the inefficiency of different flight distances. In both regions, inefficiency decreases as flight distance increases, mainly because the excess distance due to suboptimal entry and exit points is independent of the achieved distance. In Europe, the implementation of free route airspace (FRA) has improved en route efficiency significantly, especially for flights through those areas.

Current literature exists on horizontal flight inefficiency and its causal factors, but none focuses on a combination of convective weather, wind, and Miles-in-Trail (MIT) restrictions. Ref. [9] provides a look at the methodology for comparing flight en route efficiency between the US and Europe before the FAA and EUROCONTROL generated regular US-Europe performance reports. That paper utilizes an earlier version of our approach, including inefficiency indicators based on the comparison of radar trajectories and theoretically optimum trajectories, and the separation of terminal and en route flight portions using a radial distance surrounding the origin and destination airports. In more recent literature, [10] develops inefficiency metrics that quantify the discrepancy between the shortest lateral trajectory and the actual flight trajectory, but does not consider the terminal extension inefficiency of the en route portion of the flight. That work, which includes terminal airspace analysis, suggests that the biggest contributor to inefficiency at the time of publication was the airspace restrictions observed by standard routes. Congested airspace and adverse weather contributed about 13% of the extra track distance flown in the US.

A major criticism of "achieved distance" as the benchmark distance for efficiency is that the Great Circle Distance (GCD) is not an "optimum" trajectory for most flight operations when considering meteorological conditions. Ref. [11] examines the benefit of wind-optimal trajectories in transatlantic routes as measured by fuel efficiency. In many cases, a flight taking the shortest ground path would burn more fuel and sometimes take longer, thus underperforming in those standards of efficiency. Based on such knowledge, we expect that wind has a strong impact on en route inefficiency.

There is little recent work on the contribution of MIT restrictions to en route inefficiency. Ref. [12] concludes that MIT restrictions did not appear to have a significant impact on airborne delay or flight spacing during the period of analysis (May 1-14, 2004). They found that 39% of restrictions involved fewer than 6 flights and thus may have been unnecessary. Ref. [13] outlines the capability of an MIT Impact Assessment (MIA) tool that would allow traffic managers to evaluate MITs prior to implementation and prevent the initiation of redundant restrictions. Due to the age of these reports and the increasing number of MITs implemented in more recent years, it is fair to say that very little is known about whether or to what degree MIT restrictions contribute to en route inefficiency in the present system.

Comparative assessments typically emphasize macro comparisons, however, from the standpoint of senior decision makers, such high-level conclusions represent the ultimate payoff from vast amounts of data collection and analysis. To researchers and scholars, however, they beg more detailed questions. Are taxi-out times in the US high everywhere or are the results skewed by a few highly-congested airports? Similarly, is airborne inefficiency fairly constant across space and time, or are there pronounced patterns of variation and, if so, what are they? Most importantly, what are the more important causes of flight inefficiency, and how can their contributions be quantified? Answers to such questions may be equally or more relevant to the practical project of improving system performance than the macro comparisons. It is for this reason that we investigate causal factors for the en route inefficiencies observed in the present system.

The remainder of this paper is organized as follows. In Section II, we introduce the data sources and review our previous work on micro variations of en route inefficiency. Section III describes the methodologies and algorithms. Section IV applies statistical models to quantify the impacts of wind, convection and MIT on en route inefficiency, and Section V offers the conclusions.

II. PRELIMINARIES

A. Data Sources and Summary Statistics

In this project, we used five datasets from different sources. The flight event data, which come from the FAA Traffic Flow Management System (TFMS) and from the Aviation System Performance Metrics (ASPM), contain flight-level records with variables including arrival airport, departure airport, aircraft type, and D40 to A100 Actual/Great Circle/Achieved distances. We obtained flight event data for the calendar years 2013 and 2014 and found an average inefficiency of approximately 3.4% for both years. This data set includes around 12 million records, 87% of which are domestic flights and less than 1% are diverted flights or have missing records. We focused on flights arriving at and departing form the main 34 US airports listed in the 2013 US-Europe performance report. This list includes the 30 core US airports and an additional 4 high-traffic airports. The 34 main airports handle around 3 million flights per year, which accounts for about 50% of the total number of flights in the US.

The flight tracks dataset, which also comes from TFMS, contains information about the position and movement of each aircraft throughout its flight. Some fields of interest include: latitude, longitude, altitude, groundspeed, and time, with an interval of approximately one minute between each track point. For the purposes of our current research, we only obtained flight tracks between the 8 airport pairs listed in TABLE I for the 2013 calendar year. During the process to remove erroneous trajectories, we excluded tracks where spatial or temporal discontinuities were detected and ones that started or ended outside of the selected terminal areas.

Our convective weather data are derived from the Quality Controlled Local Climatological Data (QCLCD), which is obtained from the National Oceanic and Atmospheric Administration (NOAA). The dataset includes hourly summaries for convective weather conditions like thunderstorms, rain, hail, etc. at 1,600 locations in the US. Each record is a vector of binary variables indicating if there was a certain type of weather occurring at a specific time and location. QCLCD was collected for the same time period as the 8 airport pairs – CY2013. While more precise and higher resolution weather data are available from other sources, we considered the QCLCD data acceptable, as well as easier to work with, for performance analysis purposes.

TABLE I SUMMARY STATISTICS FOR REPRESENTATIVE CITY PAIR INEFFICIENCY ANALYSIS

Airport Pair	Original No. of Records	% of Records Removed	Average Inefficiency $((A-H)/H)$
IAH → BOS	1817	7.59%	4.37%
BOS → IAH	1883	7.49%	2.58%
FLL → JFK	4267	6.00%	3.59%
JFK → FLL	4273	5.38%	3.06%
ORD → DCA	7574	2.97%	4.11%
DCA → ORD	7557	2.30%	4.20%
JFK → LAX	11586	7.43%	2.04%
LAX → JFK	11543	9.49%	2.43%

The forecast wind data used in this paper are obtained from the North American Mesoscale Forecast System (NAM) of the National Centers for Environmental Prediction (NCEP). It predicts the high-resolution wind field four times a day; each prediction covers a six-hour period. Within each prediction cycle, the system produces a forecast for every hour for the first three hours, and a forecast for the last hour of the cycle. The horizontal resolution is roughly 0.1×0.1 degree latitude/ longitude. The data also have 39 equally spaced vertical isobaric pressure levels ranging from 50 mbar to 1000 mbar. By parsing all CY2013 wind field data from the NAM website, we obtained approximately 8000 data files that include the vertical/horizontal wind speed, the position grid, and time.

The MIT dataset comes from the National Traffic Management Log (NTML), and contains information about where, when, and why MIT restrictions were implemented. Some fields of importance include: providing facility, requesting facility, NAS element, start/end time, altitude, and the actual MIT spacing parameter. For our study, we excluded restrictions that were cancelled before initiation and those whose MIT values were less than 0. Similar to the cleaning of the flight tracks data set, the MIT restrictions were pre-processed to remove erroneous records and entries that we were unable to match with trajectories. We only included restrictions with known facility and NAS element geometries (ARTCC, TRACONS, fixes, and airways). TABLE II provides a few summary statistics about MIT restrictions during CY2013.

TABLE II SUMMARY STATISTICS FOR MIT RESTRICTIONS

	TOTAL	ENROUTE
Cancelled Before Initiation Rate	2.8%	2.6%
Number Initiated	186,184	81,294
Modification Rate (Extended or Modified)	21.3%	19.0%
Average Duration	2.007 hours	1.861 hours
Average MIT Value	18.67 MIT	20.52 MIT
Total Number of MIT after cleaning	151,545	63,142
Percent lost by cleaning	18.60%	22.33%

B. Movitation from Previous Work

In previous work [14], we apply fixed effects regression techniques to a two-year flight-level performance dataset to quantify how departure/arrival airports and seasons impact en route inefficiency. Consistent with the literature, our estimation results suggested that long-haul flights are more efficient than short-haul flights. Additionally, we find that flights occurring in the summer season, when convective weather is more frequent, are less efficient than flights during other seasons.

A deficiency of this work was that our models explained a low percentage of the variance in flight inefficiency. We observed that variations for a specific airport pair within one month were quite pronounced, suggesting the need for a model that considers circumstances specific to each flight. Extensive literature also suggests that wind and convection have strong effects on strategic routing, while en route traffic management initiatives affect tactical rerouting significantly. In this paper, we investigate inefficiency variation that arises from strategic routing choices made by flight operators (often with considerable guidance from the FAA). We also analyze the impact of convective weather and MIT restrictions on this route selection and on flight efficiency.

III. METHODOLOGIES AND ALGORITHMS

Our analysis is based on the concept of nominal route. Flight trajectories are unique, but most flights for a given origin-destination (OD) pair fall into one of several groups, or clusters. Trajectories within a cluster are quite similar to one another. Among the trajectories that form a given cluster, one can identify a specific trajectory that is the “center” of the cluster, and which we term the nominal route. Much of the variation in flight inefficiency can be attributed to variation in the inefficiency of the nominal routes. Moreover, one can use nominal routes as a basis for determining cluster attributes (for example, exposure to convective weather) that help explain to which cluster a given flight belongs. As will be shown, this can provide an approach to the main problem addressed by this paper: relating en route inefficiency to causal factors. To implement this idea, we must (a) apply a trajectory clustering algorithm to classify flight tracks; (b) identify the nominal route for each cluster; and (c) determine, for the departure time of a given flight, the relevant attributes of each nominal route.

A. Trajectory Clustering

There is an extensive literature discussing the application of clustering algorithms to flight trajectories. Ref. [15] proposed a three-step framework – linear interpolation, Principal Component Analysis (PCA) and k-means [16] – to efficiently classify flight tracks into unique groups. Ref. [17] improved the work of [15] by constructing a Gaussian kernel dissimilarity matrix and applying a variant spectral clustering algorithm. However, such a method is computationally expensive. Other methods such as way-point-based clustering [18] and partition-and-group framework for trajectory clustering ([19], [20]) can be applied to some circumstances requiring higher granularity (e.g., terminal trajectory clustering). Refer to [21] and [22] for further discussions on trajectory clustering.

For purposes of simplicity and efficiency, we mainly follow the framework of [15], but switch from k-means to DBSCAN

[23] in the final step for three reasons: a) it does not need us to predetermine the number of clusters; b) it allows trajectories to be classified as outliers (i.e., not belonging to any cluster); and c) it can limit the variation within each cluster. After the clusters are determined, we further calculate the nominal route for each cluster by solving a 1-median problem [24], and we define the set of nominal routes as a route choice set. Notice that each nominal route is a specific 4D trajectory.

We apply our clustering framework to the eight OD pairs listed in TABLE I. To save space, we only show the clustering for flights from IAH to BOS in Fig. 1 and Fig. 2. Other clustering results are summarized in TABLE III.

In Fig. 1, different colors represent different clusters. The black trajectories are classified as outliers, and the white curves within each cluster are the associated nominal routes. Fig. 2 shows the boxplot of en route inefficiency for flights within the different clusters. While the distributions for natural clusters (r - b) are compact, that for outlier group (k) has more variation. TABLE III indicates that for most airport pairs, individual flight trajectories can be divided into 3 – 5 natural clusters whose members are very similar to each other, and at most 21% of the trajectories are outliers. However, for traffic between DCA and ORD, we find that no matter how we tweak the parameters, there is always one dominant cluster that includes more than 95% of the total trajectories. In this case, it is inappropriate to model the assignment of flights to clusters based on attributes of nominal routes. Therefore, we exclude the flights between DCA and ORD in our further analysis.

TABLE III SUMMARY STATISTICS FOR CLUSTERING RESULTS

City Pair	Cluster ID Average Inefficiency/ (Size of cluster in percentage)					
	r	**g**	**m**	**c**	**b**	**k (outlier)**
IAH BOS	3.80% (36.21%)	6.14% (25.25%)	2.84% (30.55%)	1.76% (0.95%)	3.26% (1.13%)	8.78% (5.90%)
BOS IAH	1.65% (51.49%)	6.48% (4.71%)	2.28% (26.41%)	1.66% (4.94%)	5.63% (1.55%)	6.00% (10.91%)
JFK LAX	0.94% (42.45%)	4.72% (2.73%)	2.50% (18.36%)	2.73% (8.37%)	1.44% (9.56%)	3.74% (18.53%)
LAX JFK	2.25% (12.32%)	1.62% (20.64%)	2.59% (3.99%)	1.96% (42.16%)	-	4.24% (20.89%)
FLL JFK	2.54% (86.49%)	8.61% (10.72%)	8.10% (0.92%)	-	-	21.12% (1.87%)
JFK FLL	1.85% (84.15%)	11.59% (1.11%)	8.05% (9.57%)	12.28% (2.05%)	15.09% (0.64%)	10.00% (2.47%)
DCA ORD	3.55% (95.27%)	4.01% (1.64%)	14.06% (0.28%)	12.32% (0.28%)	51.94% (0.11%)	25.75% (2.41%)
ORD DCA	3.5% (96.43%)	20.73% (1.36%)	12.75% (0.95%)	8.50% (0.14%)	-	28.31% (1.12%)

B. Mapping Framework

We start this section by introducing the general ideas behind the algorithms that connect individual trajectories to instances of the proposed causal factors behind inefficiency. For each flight, we construct a scenario whereby it needs to choose a route from a choice set obtained from section III.A. The nominal routes in the choice set differ only in the route-specific characteristics, such as convective weather exposure, wind and MIT occurrence. While the nominal routes as originally identified have unique departure times, here we assume they have a common departure time which is specified by the flight whose cluster assignment we are modelling. Based on this assumption, we determine the route-specific features for all nominal routes faced by each designated flight. Essentially, we are asking, for a given flight and for each of the nominal routes, what would likely have happened to that flight, with respect to convective weather, miles-in-trail restrictions, and wind, if it had used that route, with its original departure time? Each of these impacts is measured using a specific metric, described in the following sections.

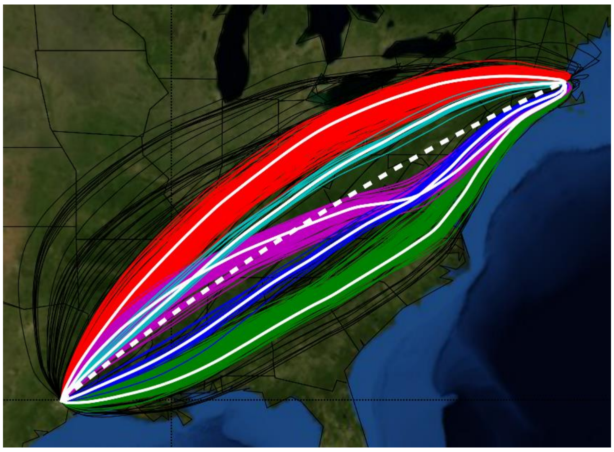

Fig. 1 Clustering results for flights from IAH to BOS

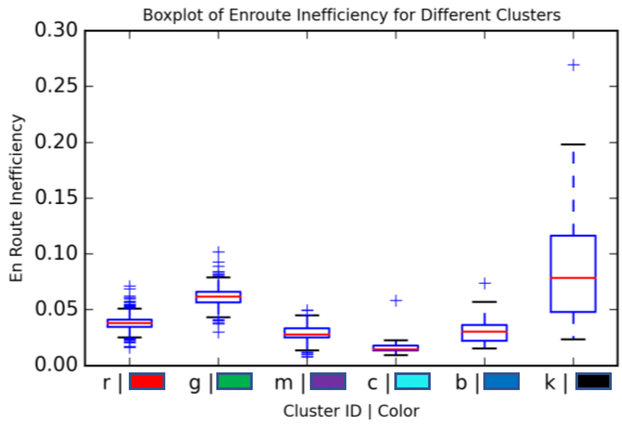

Fig. 2 Clustering statistics for flights from IAH to BOS

C. Convective Weather Mapping Algorithm

The convective weather records obtained from NOAA are recorded at individual weather stations, and are represented as time-stamped vectors of the form $[I_t, I_r, I_s]$, where the binary indicator variables $\{I_i\}$ represent the presence or absence of thunderstorms, rain, and squalls, respectively. Each weather station is assigned membership in its geographically appropriate ARTCC.

For a given flight and a given nominal route, we compute the weather-related impacts using the following steps:

1. Determine a hypothetical 4-D trajectory for that flight along that route by translating in time the 4-D trajectory originally associated with that route to the actual departure time of the flight.

2. From that hypothetical trajectory, determine the time-stamped sequence of ARTCCs that flight would have encountered.

3. For each track point along the entire route, collect the binary weather vectors from stations that are members of the subject ARTCCs, with time stamps that are within a window Δt (default = 1 hour) of the 4-D trajectory time stamps.

4. Average the collected weather vectors across the entire route, to give the percentage of relevant weather stations along the nominal route reporting weather phenomena of each of the three types: thunderstorms, rain, and squalls

D. Miles-in-trail (MIT) Mapping Algorithm

The miles-in-trail data from NTML are assembled as individual data records with the structure $\text{MIT} = \{st, et, reqfac, provfac, nas\_elem, alt, mit\_value\}$, where st and et are the start and end time of the restriction, respectively; $reqfac$ and $provfac$ are the requesting and providing facilities; nas_elem represents the NAS element where the MIT is enforced (it can be a jet route, TRACON or fix); alt is the altitude restriction; and mit_value is the actual mileage between subsequent aircraft imposed by the restriction.

For a given flight and a given nominal route, we compute the MIT-related impacts using the following steps:

1. Follow step 1 from the weather algorithm above.

2. Determine the set of MIT restrictions that would have been within sufficient geographic proximity of that flight, and applying in the same overall direction of travel. This required an exhaustive search of all of the MIT records, although in a real implementation one might also pre-compute membership in ARTCCs so as to take advantage of the results of step 2 of the weather algorithm. In any case, each MIT restriction considered must be assigned an influence area:

    a. If the NAS element was a fix, then the influence area we used is a circle of radius ~15 nmi around the fix. This corresponds to about 0.25 degree of latitude in the continental U.S. The trajectory was considered to have been a candidate for this MIT if it intersected this circle in the same direction of travel for which the MIT was specified.

    b. If the NAS element was a jet route, then we defined a swath centered on the jet route with overall width 30 nmi. The trajectory was considered to have been a candidate for the MIT if it intersected this swath for at least 60 nmi, in the appropriate direction of travel for the MIT.

    c. If the NAS element was a center / TRACON, we used its existing polygonal boundary. The trajectory was considered to have been a candidate for the MIT if it intersected this polygon in the MIT restriction's direction of travel.

3. The list of candidate MIT restrictions was filtered for timestamps. If the time of intersection of the trajectory and the boundary of the MIT restriction was within the interval $[st_i, et_i]$ for MIT_i, then that restriction was retained as a candidate; otherwise it was discarded.

4. The remaining list of candidate MIT restrictions was filtered for altitude. An MIT restriction is assigned one of three possible altitude requirements: a) above flight level *x*, b) below flight level *x*, or c) at flight level *x*. If a particular MIT involved cases a) or b), then we would filter out that MIT if the trajectory altitude at the time of intersection with the MIT influence area did not satisfy the requirement. In case c), we allowed a 1000-foot buffer around *x*, and discarded MITs if the trajectory altitude did not fall within that interval. It should be noted that the filtering steps 2-4 can be done in any order, and it would be wisest to do them in the most efficient order.

5. We consider that the flight in question, had it flown the route in question, would have been impacted by whatever MIT restrictions remain after steps 2-4. The magnitude of that impact is measured as the sum of the miles in trail distances, the sum of the hours in durations, and the sum of the MIT stringency imposed by all of the restrictions. This additive metric may be refined in subsequent research, as aircraft that are well-spaced from one restriction may not need much additional spacing to comply with others, unless these involve additional traffic.

E. Wind Mapping Algorithm

The wind data from NAM are stored in raster files, each of approximate size $648 \times 428 \times 34 \cong 10^7$. Each raster file corresponds to a single geographical snapshot of wind vectors at a specific time. For a given flight and a given nominal route, we determine the estimated wind-related impacts using the following steps:

1. Construct the 4-D trajectory as in step 1 from Section III.C.

2. Determine which wind raster file is closest in time to the departure time for the subject flight.

3. Using that raster file, build altitude-specific kd-trees [25] as an efficient queryable spatial data structure.

4. For each point in the prospective trajectory, query the kd-trees based on latitude, longitude, and altitude, and find the wind speed of the nearest raster point.

5. For each track point, calculate the azimuth θ with respect to the previous adjacent point, then calculate the headwind/tailwind by $HTW = v_{\perp} cos\theta + v_{\parallel} sin\theta$, where HTW is the headwind/tailwind, $v_{\perp}$ and $v_{\parallel}$ are the horizontal and vertical wind speed, respectively.

6. Use the ground speed in the flight track data and headwind/tailwind speed to calculate airspeed, then calculate the cumulative summation of the product of airspeed and time along the route.

IV. Statistical Analysis

In this section, we will investigate how the causal factors impact en route inefficiency.

A. Framework

By the fact that en route inefficiency is mainly caused by the route selection (strategic) and reroute (tactical), we establish two separate models, i.e., linear regression (LR) model and multinomial logit (MNL) model, for each OD pair to understand the mechanism behind the inefficiency.

The MNL model constructs a scenario whereby each designated flight, before departure, chooses one nominal route (assignment of cluster) from the route choice set determined in III.A. Different nominal routes at a specific designated departure time vary in climatological condition and MIT restrictions, and a flight is expected to tend toward paths where overall conditions are best. Therefore, by including those route specific features in the MNL model, we can capture their impacts on en route inefficiency in the strategic routing phase.

We observe from Fig. 2 that flights vary in en route inefficiency even though they belong to the same cluster (meaning they choose the same nominal route). This might be attributed to tactical reroute, such as avoiding unpredictable convection activities or traffic, during cruising phase. Therefore, we use a linear regression model, in which flight level en route inefficiency is the dependent variable, to further capture such effects.

We end this section by introducing our metrics for causal factors in the models:

a) Convection activities Wx: In the MNL model, this metric agrees with what we have described in section III.C. However, since we don't produce a nominal route for the outlier cluster, we set $\mathrm{Wx} = 0$ for all alternatives in the outlier group. In the LR model, we first exclude the outlier alternative for each flight, then use the weighted average of Wx for all alternatives as an indicator of weather condition in the airspace, where the weights are determined by the cluster size in TABLE III, but excluding the outlier group.

b) Wind WD: We use wind distance, which is defined as the cumulative summation of the product of airspeed and time, to evaluate the wind effect. Notice that we only include the wind distance in the MNL model, due to the fact that wind affects the inefficiency mainly through strategic routing.

c) MIT MIT_STR: We use MIT stringency, which is defined as the inner product of the MIT value (in miles) and MIT duration (in hours), as the metric. Similar to the convective weather, we include this metric both in the MNL model and the linear regression model.

B. Model Specifications

In the linear regression model, the dependent variable is the flight en route inefficiency. There are four categories of independent variables in the model. The first category includes variables related to convective weather activities. Among the variables are thunderstorm, rain, squall, hail, and ice. Since a flight may extend its route to avoid convection, we expect that those variables will have positive effects on the en route inefficiency. The second category of variables includes MIT stringency. Due to the fact that aircrafts may be vectored to comply with the MIT restrictions, we expect that this variable will also have positive sign. The third category pertains to the route structure. Fig. 2 indicates that while the within-cluster en route inefficiency maintains a small variation, the cross-cluster inefficiencies tend to have significant difference. Therefore, we include the cluster memberships as the fixed effects in the model, and set the outlier cluster membership as the baseline. We expect that the order of estimates for different clusters should match the order of the averages of within-cluster inefficiencies. The final set of variables relate to the departure time. We include both the departure seasons and departure busy hour indicator (10 am – 8 pm) as control variables to further capture the seasonal patterns and traffic effect, respectively. As indicated in [14], flights in the summer season tend to be more inefficient, and a higher volume of traffic in the airspace will also induce unnecessary maneuvers; thus, we expect that the estimate for summer season indicators should be systematically higher than the other four seasons, and the estimate for the busy hour should also be positive. The formula for the LR model is shown below, where the β's are the coefficients that need to be estimated. The explanatory variables and their notations are listed in TABLE IV.

$$InEff(\%)=\beta_0 + \beta_1 \cdot TS + \beta_2 \cdot R + \beta_3 \cdot SQ + \beta_4 \cdot Ice + \beta_5 \cdot Hail + \beta_6 \cdot \mathrm{MIT} + \beta_7 \cdot \mathrm{BH} + \sum_{\mathrm{CL}} \beta_{\mathrm{CL}} \cdot \mathrm{Dep\_Arr}_{\mathrm{CL}} + \sum_{\mathrm{k}} \beta_k^s \cdot \mathrm{Season}_{\mathrm{k}} \quad (2)$$

TABLE IV DESCRIPTION OF EXPLANATORY VARIABLES

Category	Explanatory variable notation	Variable description
Convective weather activities	**TS**	Thunderstorm exposure
	R	Rain exposure
	SQ	Squall exposure
	Ice	Ice exposure
	Hail	Hail exposure
Route structure	**Dep_Arr$_{\mathrm{CL}}$**	**Des** represents departure airport, **Arr** represents for arrival airport, and **CL** represents for cluster membership, which can be found in TABLE III
Departure time	**BH**	Busy hour indicator, equals to 1 if local departure hour is between 10 am and 8 pm
	Seaon	Seasonal indicator. Equals to "winter" if local departure month is from December to February; "spring" if from March to May; "summer" if from June to August; "fall" otherwise
Wind	**WD**	Wind distance (1000 nautical miles)
MIT	**MIT**	MIT stringency (1000 mile×hour)

In the multinomial logit model, the clusters (nominal routes) are the choice set, and a flight chooses a cluster to which it belongs. Apart from the weather activities and departure time variables, we also include two additional covariates – wind distance and MIT stringency – in the utility functions. Due to the nature of the MNL model, we expect that the coefficients of the weather, wind distance, and MIT variables should all have negative signs, since higher values of those covariates will always tend to degrade the attractiveness of a route. The utility formulas are shown from (3) to (5) below, where $V_j, j \in \{1,2,\dots,N-1\}$ is the utility for the j^{th} alternative, and V_N is the utility for the outlier group. The $ASC_j, j \in \{1,2,\dots,N-1\}$ is the alternative specific constant, and $W_{j,i}, j \in \{1,2,\dots,N\}$ represents the i^{th} weather activity listed in TABLE IV.

$$V_1 = ASC_1 + \beta_{1,0} \cdot Season + \sum_{i=1}^{5} \beta_{w,i} \cdot W_{1,i} + \beta_{\text{wd},6} \cdot WD_1 + \beta_{\text{mit},7} \cdot MIT_1 \quad (3)$$

$$V_2 = ASC_2 + \beta_{2,0} \cdot Season + \sum_{i=1}^{5} \beta_{w,i} \cdot \mathrm{W}_{2,\mathrm{i}} + \beta_{\text{wd},6} \cdot WD_2 + \beta_{\text{mit},7} \cdot MIT_2 \quad (4)$$

$$V_N = \sum_{i=1}^{5} \beta_{w,i} \cdot W_{N,i} + \beta_{\text{wd},6} \cdot WD_N + \beta_{\text{mit},7} \cdot MIT_N \quad (5)$$

C. Estimation Results

TABLE V reports the complete estimation results for the linear regression model. To save space, TABLE VI only reports the coefficients of factors that interest us the most for the multinomial logit model; other covariates, while most of which are significant, are not presented in the paper. For both tables, each column presents the estimations for the airport pair shown in the header.

TABLE V ESTIMATION RESULTS FOR LR MODEL

Var.	Airport Pair (Est./ Std.)					
	IAH BOS	BOS IAH	FLL JFK	JFK FLL	JFK LAX	LAX JFK
TS	25.76*** (2.87)	11.61*** (2.68)	5.49** (2.30)	7.27*** (1.27)	30.72*** (1.66)	41.35*** (2.24)
R	1.35* (0.81)				1.55*** (0.45)	3.70*** (0.52)
SQ					226.72*** (65.32)	
MIT		1.06*** (0.32)	0.86** (0.36)	1.29*** (0.25)		4.10*** (0.30)
BH		0.15** (0.06)	1.05*** (0.10)	0.13** (0.05)	0.04* (0.02)	0.54*** (0.03)
CL_r	-4.56*** (0.15)	-4.14*** (0.10)	-18.04*** (0.35)	-8.03*** (0.17)	-2.69*** (0.03)	-1.85*** (0.06)
CL_g	-2.22*** (0.15)	0.58*** (0.16)	-12.14*** (0.37)	1.42*** (0.30)	1.13*** (0.07)	-2.31*** (0.05)
CL_n	-5.42*** (0.16)	-3.49*** (0.11)	-12.82*** (0.59)	-1.92*** (0.19)	-1.15*** (0.04)	-1.64*** (0.09)
CL_c	-7.03*** (0.36)	-4.18*** (0.16)		2.21*** (0.25)	-1.08*** (0.05)	-2.31*** (0.04)
CL_b	-5.14*** (0.35)	-0.4 (0.25)		5.11*** (0.36)	-2.23*** (0.04)	
Spring	-0.04 (0.10)	0.09 (0.09)	0.26** (0.13)	0.14** (0.07)	-0.03 (0.03)	0.24*** (0.05)
Summer	0.19* (0.10)	-0.04 (0.09)	0.21 (0.15)	-0.08 (0.09)	-0.21*** (0.04)	-0.20*** (0.05)
Fall	0.05 (0.10)	0.08 (0.09)	-0.04 (0.14)	-0.03 (0.08)	0.01 (0.03)	0.01 (0.05)
Const.	8.00*** (0.17)	5.56*** (0.12)	19.74*** (0.37)	9.65*** (0.18)	3.48*** (0.04)	3.39*** (0.05)
R^2	0.63	0.65	0.54	0.76	0.53	0.33
Obs.	1664	1732	3988	4021	10637	10367

*Standard errors in parentheses; * p<.1, ** p<.05, ***p<.01*

TABLE VI ESTIMATION RESULTS FOR MNL MODEL

Var.	Airport Pair (Est./ Std.)					
	IAH BOS	BOS IAH	FLL JFK	JFK FLL	JFK LAX	LAX JFK
TS	-31.52*** (4.64)	-43.10*** (5.00)	-17.92*** (3.12)	-9.92*** (3.078)	-22.72*** (2.38)	-39.64*** (2.89)
R	-9.25*** (1.36)	- 11.04*** (1.41)	-3.92*** (1.49)	-6.67*** (1.15)	-5.86*** (0.71)	-5.62*** (0.68)
SQ					-211.03** (81.95)	-146.69* (79.04)
WD	-16.43*** (1.096)	-11.76*** (1.38)			-13.64*** (0.52)	-12.78*** (0.64)
MIT			-1.38*** (0.47)	-1.39*** (0.35)		

*Standard errors in parentheses; * p<.1, ** p<.05, ***p<.01*

The vast majority of the coefficients in the LR models are significant, and their signs match our expectations in general. First of all, the estimates for thunderstorm confirm that it increases en route inefficiency for all six pairs considered, and it has the largest impact on flights from LAX to JFK. However, rain only has a minor positive impact on IAH to BOS, JFK to LAX and LAX to JFK flights, and squall only influences flights from JFK to LAX. The large coefficient of squall is mainly caused by the considerably low occurrence of squall activities. Second, MIT restrictions increase the inefficiency for BOS to IAH, FLL to JFK, JFK to FLL and LAX to JFK flights, which could be explained by the vectoring of flights subject to MITs. Third, the estimates for cluster membership fixed effects are all significant, and their order matches the statistics shown in TABLE III, which reemphasis the effects of route selection. For instance, the estimate for CL_r in IAH to BOS case implies that if a flight decides to follow the path in cluster “r”, then its inefficiency will be systematically 4.56% lower than flying an outlier path. Busy hour flights tend to be less efficient for most pairs except IAH to BOS case, which can be attributed to the higher volume of traffic in the airspace. Finally, estimates for summer season fixed effects are somewhat surprising. While we expect that their estimates to be positive, JFK to LAX and LAX to JFK both report significant and negative sign coefficients. This is probably because the effect of summertime convective weather is explicitly accounted for in these models, and thus does not contribute to the summer fixed effect.

In general, the estimates for the MNL model agree with the LR model. Thunderstorm and rain have largely negative impacts on route utility, indicating that in the strategic planning phase, a flight will tend to fly a route that has less convective weather. Squall only has significant effects for long haul flights between JFK and LAX, mainly due to its low frequency. The estimates for wind distance are only significant for flights between IAH and BOS, JFK and LAX, and their negative signs imply that flights have a tendency to choose routes with favorable winds. After investigating the wind field map in the US, we find that there is a strong jet stream in the lower 48 US states, which covers most of the routes for the four pairs mentioned above. However, for flights between FLL and JFK, possibly because of the comparably shorter distance and lower wind activities, wind is not an important factor on the route selection. Finally, MIT restrictions have negative and significant coefficients only for traffic between JFK and FLL, indicating that they are important in strategic route choice for some but not all OD pairs.

D. Contributions of Causal Factors

We further use the models described in IV.B to quantify the contributions of wind, convection and MIT restrictions. We first use the MNL models to estimate the probabilities of different routes in the choice set, given the current wind, weather and MIT conditions in the original dataset. Then we plug the probabilities into the LR model, and further calculate the expected en route inefficiency for each record in the original dataset. Then we construct a counter factual case in which the convection variables and MIT stringency are set to 0 (i.e., no convective activities and no MITs), and wind distances are set to the great circle distance of nominal routes (i.e., no wind), respectively. Finally, we estimate the route choice probabilities and expected en route inefficiency by the new data. The percentage change in the expected inefficiency is defined as the contribution of the corresponding factor. The framework above is described in the formula below, where $\mathcal{E}(\mathrm{Ineff}|\mathrm{F},\mathrm{C_L})$ is the expected inefficiency given factor and cluster membership, $\mathrm{P}(\mathrm{C_L}|\mathrm{F})$ is the probability of cluster $\mathrm{C_L}$ given factor condition.

$$E_{withFactor} = \sum_{CL} \mathcal{E}(Ineff|F, C_L)P(C_L|F) \quad (6)$$

$$E_{noFactor} = \sum_{CL} \mathcal{E}(Ineff|F = 0, C_L)P(C_L|F = 0) \quad (7)$$

$$Contribution = \%\Delta = \left|\frac{E_{withFactor} - E_{noFactor}}{E_{noFactor}}\right| \quad (8)$$

The contributions for wind, convection (including thunderstorm, rain and squall) and MIT restrictions are summarized in TABLE VII. The first column is the expected en route inefficiency in the current condition. The two columns of each of the following panels are the expected en route inefficiency without a specific factor and its corresponding contribution, respectively. The table shows that if there were no weather impact, flights would be 7-18% less inefficient, depending on the OD pair. If winds did not influence flight routing, inefficiency would decrease from 0-11%, while the impact of eliminating MITs would be to decrease inefficiency from 0-5%.

TABLE VII CONTRIBUTIONS OF WEATHER AND WIND

City Pair	Base-line	Weather		Wind		MIT	
		With-out	Contri-bution	With-out	Contri-bution	With-out	Contri-bution
IAH BOS	4.36%	3.93%	9.99%	4.13%	5.26%	4.36%	0.00%
BOS IAH	2.58%	2.21%	14.28%	2.51%	2.75%	2.53%	1.87%
FLL JFK	3.59%	3.30%	8.08%	3.59%	0.00%	3.46%	3.54%
JFK FLL	3.06%	2.83%	7.43%	3.06%	0.00%	2.92%	4.66%
JFK LAX	2.43%	1.73%	15.42%	1.82%	11.26%	2.05%	0.00%
LAX JFK	2.05%	2.00%	17.74%	2.30%	5.21%	2.31%	4.97%

V. CONCLUSIONS

This research demonstrates a practical means for fusing FAA trajectory performance databases with thunderstorm, wind and Miles-In-Trail (MIT) sources for the purpose of obtaining an improved estimate of flight trajectory efficiency. Trajectory clustering methods are presented which can be used for any application where it is useful to group traffic into manageable flows.

The results in this paper present a first step in ascribing en-route trajectory inefficiencies to causal factors. In our framework, the inefficiency of a trajectory is modeled as resulting from an assignment process through which the trajectory "chooses" among several trajectory clusters based on attributes of the nominal routes for each cluster, and a tactical process in which flight inefficiency for trajectories a given cluster is related to composite conditions. Using this approach, the contributions of wind, connective weather, and MIT restrictions to flight inefficiency are estimated. Results vary across OD pairs, but in general convective weather makes the greatest contribution, followed by winds and MIT restrictions.

In addition to its scientific contribution, this paper has a number of practical applications. First, it helps target interventions to reduce flight inefficiency, by identifying which causes of inefficiency are the most important for different OD pairs. Equally important, it estimates the extent to which flight inefficiency, as defined in this paper, is an illusion arising from the choice of routes with favorable winds but longer ground distances. Finally, it reveals the degree to which inefficiency is not the result of any of any identifiable cause. This inefficiency "dark matter" is presumably attributable to features of the NAS route structure, which may also be targets for intervention.

ACKNOWLEDGMENT

This work was supported by the FAA through the NEXTOR-II Consortium.

REFERENCES

[1] European Union, "Regulation 390/2013 – Laying Down a Performance Scheme for Air Navigation Services and Network Functions," 2013.

[2] International Civil Aviation Organization, "Global Air Navigation Plan Resources," 2016. [Online]. Available: http://www.icao.int/airnavigation/Pages/GANP-Resources.aspx. [Accessed 28 01 2017].

[3] ATO, "Order JO 7210.3W Facility Operation and Administration," Federal Aviation Administration, 2010.

[4] USDOT, "14 CFR Part 234 - Airline Service Quality Performance Reports," U.S. Department of Transportation, 1987.

[5] Performance Review Commission and FAA-ATO, "2013 Comparison of air traffic management-related operational performance: U.S./ Europe.," European Commission, U.S. Department of Transportation, 2014.

[6] EUROCONTROL, "Horizontal flight efficiency, achieved distances," EUROCONTROL, 2013. [Online]. Available: https://www.eurocontrol.int/sites/default/files/content/documents/single-sky/pru/news-related/2013-05-08-slides-workshop-achieved-distance.pdf. [Accessed 28 01 2017].

[7] I. Fuller, H.-C. Hustache and K. Tarja, "Enhanced Flight Efficiency Indicators," Eurocontrol Experimental Centre, EEC/SEE/2004/011, 2004.

[8] Performance Review Commission and FAA-ATO, "2015 Comparison of Air Traffic Management-Related Operational Performance: U.S./Europe. European Commission," U.S. Department of Transportation, 2016.

[9] T. Kettunen, J.-C. Hustache, I. Fuller, D. Howell, J. Bonn and D. Knorr, "Flight efficiency studies in europe and the united states," in *5th NASA Integrated Communications Navigation and Surveillance Conference (ICNS)*, Fairfax, Virginia, USA, 2005.

[10] T. G. Reynolds, "Analysis of Lateral Flight Inefficiency in Global Air Traffic Management," in *8th AIAA Aviation Technology, Integration and Operations Conference*, Anchorage, Alaska, USA, 2008.

[11] B. Sridhar, N. Y. Chen, H. K. Ng, O. Rodionova, D. Delahaye and F. Linke, "Strategic Planning of Efficient Oceanic Flights," in *11th USA/Europe Air Traffic Management Research and Development Seminar (ATM2015)*, Lisbon, Portugal, 2015.

[12] T. A. Myers, M. Klopfenstein, J. Mintzer and G. Wilmouth, "A preliminary analysis of the impact of miles-in-trail restrictions on NAS flight operations," in *6th USA/Europe Air Traffic Management Research and Development Seminar (ATM2005)*, Baltimore, MD, USA, 2005.

[13] P. Ostwald, Tejal and J. DeArmon, "The miles-in-trail impact assessment capability," in *6th AIAA Aviation Technology, Integration and Operations Conference (ATIO), Aviation Technology, Integration, and Operations (ATIO) Conferences*, Wichita, KS, USA, 2006.

[14] M. Hansen, Y. Liu, C. Chuang, D. Lovell and M. O. Ball, "En route performance in the national airspace system," in *96th Transportation Research Board Annual Meeting*, Washington D.C., USA, 2017.

[15] A. Eckstein, "Automated flight track taxonomy for measuring benefits from performance based navigation," in *2009 Integrated Communications, Navigation and Surveillance Conference (ICNS)*, Crystal City, VA, USA, 2009.

[16] J. MacQueen, "Some methods for classification and analysis of multivariate observations," in *Proceedings of the fifth Berkeley symposium on mathematical statistics and probability*, Oakland, CA, USA, 1967.

[17] M. Enriquez and C. Kurcz, "A simple and robust flow detection algorithm based on spectral clustering," in *5th International Conference on Research and Air Transportation (ICRAT 2012)*, Berkeley, CA, USA, 2012.

[18] M. Gariel, A. N. Srivastava and E. Feron, "Trajectory clustering and an application to airspace monitoring," *IEEE Transactions on Intelligent Transportation Systems,* vol. 12, no. 4, pp. 1511-1524, 2011.

[19] J.-G. Lee, J. Han and K.-Y. Whang, "Trajectory clustering: a partition-and-group framework," in *International Conference on Management of Data*, Beijing, China, 2007.

[20] J.-G. Lee, J. Han and X. Li, "Trajectory outlier detection: A partition-and-detect framework," in *Proceedings of 24th International Conference on Data Engineering*, 2008.

[21] M. Enriquez, "Identifying temporally persistent flows in the terminal airspace via spectral clustering," in *10th USA/Europe Air Traffic Management Research and Development Seminar (ATM2013)*, Chicago, IL, USA, 2013.

[22] M. a. K. G. a. G. D. Vlachos, "Discovering similar multidimensional trajectories," in *Proceedings 18th International Conference on Data Engineering*, 2002.

[23] M. Ester, H.-P. Kriegel, J. Sander and X. Xu, "A density-based algorithm for discovering clusters in large spatial databases with noise," in *Proceedings of 2nd International Conference on Knowledge Discovery and Data Mining* , 1996.

[24] N. Megiddo, "The weighted Euclidean 1-center problem," *Mathematics of Operations Research,* vol. 8, no. 4, pp. 498 -- 504, 1983.

[25] J. L. Bentley, "K-d trees for semidynamic point sets," in *Proceedings of the sixth annual symposium on Computational geometry*, 1990.

AUTHOR BIOGRAPHIES

Yulin Liu is a Ph.D. student of Civil and Environmental Engineering at the University of California, Berkeley. His research interests include air traffic flow management, trajectory optimization, and machine learning algorithms. He received his B.S. degree in Civil Engineering in 2015 from Tsinghua University.

Mark Hansen is a Professor of Civil and Environmental Engineering at UC Berkeley, and a co-director of the National Center of Excellence for Aviation Operations Research.

David Lovell is a Professor of Civil and Environmental Engineering at the University of Maryland. He holds a joint appointment with the Institute for Systems Research. Dr. Lovell received his Ph.D. in Civil Engineering in 1997 from the University of California, Berkeley.

Cara Chuang is a M.S. student in the Systems Engineering program at the University of Maryland. She is specializing in Reliability with a research interest in Operations Analysis. She received her B.S. in Bioengineering from the University of Illinois in 2014.

Michael Ball holds the Dean’s Chair in Management Science in the Robert H. Smith School of Business at the University of Maryland. He also has a joint appointment within the Institute for Systems Research in the Clark School of Engineering, and is co-Director of NEXTOR-II, an FAA consortium in aviation operations research. Dr. Ball received his Ph.D. in Operations Research in 1977 from Cornell University.

John Gulding currently serves as the manager for the Strategic Analysis group within the FAA/ATO Office of Performance Analysis. The group produces performance reporting tools as well as the metrics used by senior ATO management. His education includes a B.A. in Mathematics from the University of Virginia and Masters in Operations Research from George Mason University