Analysis of Excess Flying Time in the National Airspace System

George Solomos, James DeArmon, Gerald Dorfman, Ho Yi
Center for Advanced Aviation System Development (CAASD)
*The MITRE Corporation**
McLean, Virginia 22102
United States
gsolomos@mitre.org, jdearmon@mitre.org, gdorfman@mitre.org, yih@mitre.org

Abstract

Research on the United States (U.S.) National Airspace System (NAS) has sought answers to the following questions: Is there measurable excess flying time in the NAS? If so, where does it occur? Using aircraft track and flight information for multiple years, we have discovered a significant level of excess flying time when using a "best observed" flying time as a standard. In the en route regime, 4 to 5 minutes per flight, in good weather, can be detected. Drilling down to a fine-grained geographic mesh on a map of the U.S. enables detection of specific locations of significant delay. These locations can be associated with operational sectors. This information has allowed an examination of changes over time, and should aid in focusing the scarce Federal Aviation Administration (FAA) funding for congestion management.

Introduction

On 21 December 2004, 24 flights departed New York John F. Kennedy International airport and arrived at Los Angeles International airport (data obtained from [1]). Ten of these flights were operated by the same airline using an identical aircraft type. The average airborne time for these 10 flights was 338.8 minutes and ranged from a minimum of 333 minutes to a maximum of 349 minutes. Although the fastest and slowest flights had planned on essentially the same amount of airborne time (330 and 331 minutes, respectively) their actual performance was quite different. Why such different flying times? How much of the difference can be explained by factors such as differences in experienced winds, and how much by other factors such as airway congestion, traffic flow strategies (e.g., miles-in-trail and vectoring), differences in styles of piloting, or aircraft routings? Is it possible that improvements to the NAS could alleviate some of the excess amount of flying time?

Over the years the FAA, along with the aviation community as a whole, has strived to increase the number of flights the NAS can accommodate as well as to increase the efficiency of those flights. On a yearly basis the FAA releases its NAS Operational Evolution Plan (OEP) [2] that details specific capacity improvements planned over a rolling 10-year period. This plan contains much detail about improvements in the en route airspace that will take place to reduce flight delays as well as increase the efficiency of individual flights. How much improvement is possible in today's NAS? Do planned improvements address much of the current-day inefficiencies, or do they only scratch the surface? This study attempts to begin answering these questions.

Background

The MITRE Corporation's Center for Advanced Aviation System Development (CAASD) has been working with the FAA's Air Traffic Organization – Operational Planning (ATO-P) to evaluate the amount of potential benefits available from future enhancements. Every day across the United States many thousands of aircraft fly within the NAS. The amount of time they spend in the air varies from flight to flight. Some long-haul flights may be airborne for many hours, while short-haul flights may be in the air only for tens of minutes. As discussed above, individual flying times may vary widely even for identical aircraft types between the same airports. Having a better understanding of where inefficiencies may be taking place, and to what extent, is important for many reasons. In today's budget-constrained world it is necessary to make sure the right decisions are made regarding airspace, technological, and procedural enhancements. Knowing how much improvement is possible and how well a particular enhancement addresses this need is essential for setting expectations. Plans such as the OEP can then focus on areas of greatest need. This knowledge helps program evaluators understand if expected benefits by various programs are reasonable and achievable. Deciding the best time to implement improvements, and the locations to do so, helps ensure the most benefit is gained at the right times.

This study attempts to understand the amount of *excess flying time* that may exist in today's NAS, while accounting for some of the causes of flying time variation, such as wind effects. Its focus is strictly on the amount of time spent in the air compared to some minimal flying time. It was important in the performance of this analysis to measure the amount of excess flying time experienced in today's system compared to a *realistically achievable* flying time. To that end, a decision was made to use actual recorded data as opposed to simulated or modeled flight times in this study. The use of actual data allowed the selection of a minimal flying time to be based upon times *actually experienced* in today's system, as opposed to a theoretical optimal time that may not be achievable. Results are presented as absolute minutes of excess flying time as well as an excess percentage of flying time (to normalize for long versus short flights).

The authors are not aware of other studies that use flying time to measure *inefficiency*. Some other studies use flying time for other purposes. Bolczak et al. [3] used estimated time of arrival data to analyze a trend in flying time, and report some year-to-year changes. An early simulation of "Free Flight" (a set of industry and government programs that provide greater freedom for pilots and airlines to select planned and actual routes and take-off times) [4] analyzed flying times for flights, absent the constraints of route structure. Willemain [5] examined sources of variability in flying times for certain city pairs. The Bureau of Transportation Statistics [6] has collected flying time statistics and hosts a website that allows the public to access city pair flying time information.

The first part of this study, the flight-based method, analyzes flying time between city pairs for similar aircraft types. Adjustments are made to actual flying times in order to account for wind effects upon the flights. This part of the study discusses the segments of flights analyzed, the methodology used for adjusting flying times for winds, how minimal flying times were selected, the time-period studied, and the overall findings of this analysis. The second part of this study focuses on understanding *where* the excess flying time is actually taking place. It shifts from a flight-based approach to a geographic cell-based approach. The Cell-Based Approach measures excess flying time over small areas while accounting for differences in equipment type, wind effects, and flight altitude. Follow-on work is also discussed at the end of this report.

Flight-Based Method

The flight-based method estimates excess flying time in the NAS on a per-flight basis. This excess flying time is the amount by which flights exceed minimum flying times. In order to compare flying times on different days, wind effects on flying times are estimated. Then, excess flying times are computed after adjustments for wind effects. The ideas and methods were first described in [7].

The following sections present the methodology and results for the flight-based method.

Methodology

We selected 15 good-weather days from 2001 and from 2002, which we studied as a population of 30 days. Days were selected using a scoring scheme, called the *misery index*. The misery index is the sum of percent of flight cancellations, twice the percent of flight diversions (landing at an airport different from the one seen in the original flight plan), and percent of flights with more than 30 minutes of departure delay. These three percentage measures are based on Airline Service Quality Performance (ASQP) data [8]. Previous studies such as [9] showed a good correlation between this scoring scheme and general weather conditions in the NAS. Based on their scores, the days in each year were ranked from best to worst. Then, the 15 best days in each year were selected for the study.

As illustrated in Figure 1, this analysis examined flying time for three segments of flight:

- Ten nautical miles (nmi) from origin airport to ten nmi from destination airport
- Forty nmi from origin airport to forty nmi from destination airport
- One hundred nmi from origin airport to one hundred nmi from destination airport

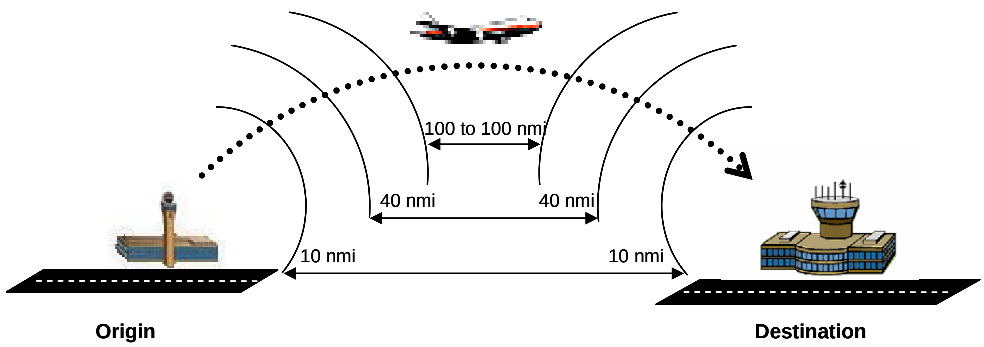

Figure 1. The Three Segments of Flight Analyzed

The three data sets corresponding to these three segments of flight, respectively, provide a rough representation of flying times between airports, between terminal radar approach control airspaces (TRACONs), and within en route airspace. The datasets for these segments of flight are referred to as the 10/10, 40/40, and 100/100 datasets, respectively.

Using Enhanced Traffic Management System (ETMS) data, flying times were calculated for each flight in the three datasets. (See [10] for an overview of ETMS.) The datasets were filtered to include only flights whose origin and destination airports are both in the Conterminous U.S. (CONUS).

Adjusting for Winds

Winds, especially winds aloft—which are high velocity winds that impact flights at cruise altitude—have a major impact on flying times. Faced with strong headwinds, to adjust their arrival times, pilots may "throttle forward"; faced with strong tailwinds, they may "throttle back." Because pilots compensate for winds, it is inappropriate to simply apply vector algebra to adjust flying times for wind speed and direction.

We computed wind effects using an *opposing traffic* calculation. Although winds do vary during the day, aircraft traveling in opposite directions on the same day tend to experience opposite wind effects. For each aircraft equipment type (as specified in ETMS), we calculated average speed (ground flight track distance divided by time) for flights from airport A to airport B, and also for flights from airport B to airport A. We assume that, if winds were not a factor, traffic in each direction (for a given aircraft equipment type) would want to travel at about the same speed. Therefore, we computed the wind effect as half the difference between the speeds of aircraft flying in opposing directions. For example, if the average speed from A to B were 400 knots, and the average speed from B to A were 600 knots, then the wind effect would be a 100-knot *headwind* from A to B, and a 100-knot *tailwind* from B to A. This opposing traffic method is affected by flow control impedance, which is the increase in flying time due to Air Traffic Management (ATM) requirements in response to congestion. However, we found this effect to be minor—on the order of about 4 knots.

Wind effects were applied to each flight to obtain an adjusted flight speed. For each flight, using actual ground track distance flown, adjusted speed was converted to an *adjusted flying time*. For each of the study days, each origin/destination group (defined in the next section), and each equipment type, a wind effect adjustment was computed and used to determine adjusted flying times. For example, on a given day, assume an individual flight from A to B had a speed of 410 knots, a flight (of the same aircraft equipment type) from B to A had a speed of 588 knots, and the wind effect (between A and B for that aircraft equipment type) was 100 knots. The A-to-B flight would get an adjusted speed of 510 knots, and the B-to-A flight would get an adjusted speed of 488 knots. Then, adjusted *flying time* is calculated as

adjusted flying time = track distance flown / adjusted speed

Grouping Airports

Because winds differ from day to day, wind effect adjustments must be computed and applied on a daily basis. In order to apply the opposing traffic method described here, enough flights for each origin/destination pair and aircraft equipment type are needed to compute reasonably accurate daily wind effect adjustments.

In calculating each daily wind adjustment value for a given day, we required at least three flights in each direction for each aircraft equipment type. (There are approximately 200 equipment type designators in this data.) We discarded any flight that had no available wind adjustment value. This minimum sample size requirement for wind adjustment calculations forces airports with few flights to be ignored in the computations, biasing our overall result towards larger, busier airports. To reduce this bias, we grouped airports by proximity using a clustering algorithm known as K-means clustering [11].

Wind adjustment values were computed between groups of airports, called *clusters* instead of between pairs of individual airports. Based on a tradeoff analysis of fineness of wind adjustment versus number of flights available for wind adjustment, we computed wind adjustments using 25 clusters of airports. In ancillary studies, we found that cluster-based wind adjustments yield consistent flying time distributions of opposing traffic between cluster pairs, and the number of clusters has little effect on flying time metrics.

Computing Excess Flying Time

We calculated a *minimum flying time* for each origin/destination airport and aircraft equipment type combination. We analyzed 30 good-weather days from 2001 and 2002. For a subset of days considered, a single minimum (wind-adjusted) flying time was computed for each origin/destination/aircraft equipment type combination. Since the minimum flying times, and thence the excess flying times, can vary with the number and specific choices of days, we analyzed the dependence of excess flying times on the subsets of days used for analysis. We examined 30 samples of 1 day, 15 samples of 2 days (randomly ordered), 10 samples of 3 days, …, and 1 sample of 30 days.

Analysis Results

Figure 2 shows for the 40/40 dataset the average amount of excess flying time (in minutes) for samples of 1 day, 2 days, …, 30 days, both with and without wind adjustment. The body of the figure contains *box and whisker plots*, which feature key points of a distribution of values: a horizontal median line across the box, box upper and lower edges at approximately the third and first quartiles, "whiskers" (i.e., extended vertical lines with short horizontal head and foot) at about 1.5 times the interquartile range beyond the ends of the box, and finally, individual outlier observations beyond the whiskers. (See [12] for precise definitions.)

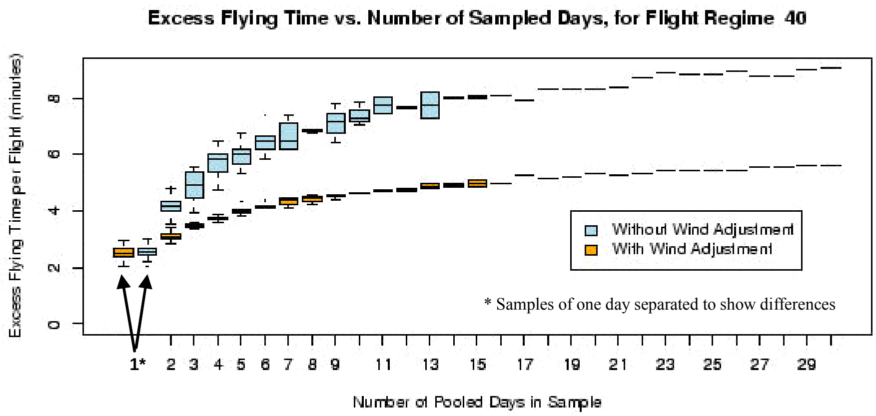

Figure 2. Excess Flying Time Versus Number of Sampled Days, for Dataset 40/40, With and Without Wind-Adjustment

Because wind adjustments are necessary only for comparison *across* days, starting with 2-day samples, we see a large divergence in the level and variability of the distributions of wind-adjusted and non-wind-adjusted excess flying times. The results for the 10/10 and 100/100 datasets are similar in form, with 10/10 having greater excess time, and 100/100 having less excess time than the displayed 40/40 dataset. The wind-adjusted results show about 5 minutes excess flying time; the curve flattens at about 15 pooled days in a sample. (For 16- through 30-day samples, we have only a single observation, and hence a box flattened to a single horizontal line segment.)

Figure 3 shows the same information as Figure 2 except that excess flying time is expressed as a percentage. Percentages are not sensitive to distance flown. Results show a flattening at about 10 percent savings (15-day sample point).

Figure 4 shows excess flying time for all three datasets analyzed in this study—10/10, 40/40, and 100/100. Each curve in Figure 4 appears to flatten at an x-axis value between 14 and 18 pooled days (the shaded yellow rectangle on the figure). This is the region where increasing the “Number of Pooled Days in Sample” ceases to have much effect on “Excess Flying Time per Flight (Minutes).” This is most likely caused by limiting our dataset to good weather days.

Therefore, in Table 1 we summarize the study results using the means of two 15-day pools for each of the datasets. Average minutes of excess flying time for the 10/10 dataset is larger than those for the 40/40 and 100/100 datasets, while those for the 40/40 and 100/100 datasets are about the same. Average percent excess flying time is largest for the 40/40 dataset and smallest for the 100/100 dataset. The results indicate that beyond the terminal areas there is a potential pool of benefits of about 4.9 minutes reduced flying time, amounting to a potential reduction in flying time of 8 to 10 percent.

An extension of the analysis (not shown here in tables or figures) examined excess flying time for the 100/100 dataset for all of the days of 2003 (i.e., all available days, a total of 355). This sample included all the good and bad kinds of weather one would see in an entire year. A plot of overall average excess flying time per flight vs. pooled days, 1 to 355, showed a monotonic increase—steep at first, then rather gradual. Even beyond 300 pooled days, there was a slight increase in average excess flying time as each single day was added: 301, 302, …, 355. This indicates that each additional day added to a sample pool supplies either new minimum flying time flights to reset a baseline, or new high-delay flights. The average excess flying time for 355 days of 2003, for the 100/100 dataset, was about 8.5 minutes.

Cell-Based Approach

Given the results of the flight-based method, a question immediately arises: Where are the excess flying times taking place? Knowledge about geographic location would allow informed decision-making on deployment of procedures and automation to ameliorate the condition, if that is possible. The data used for this analysis

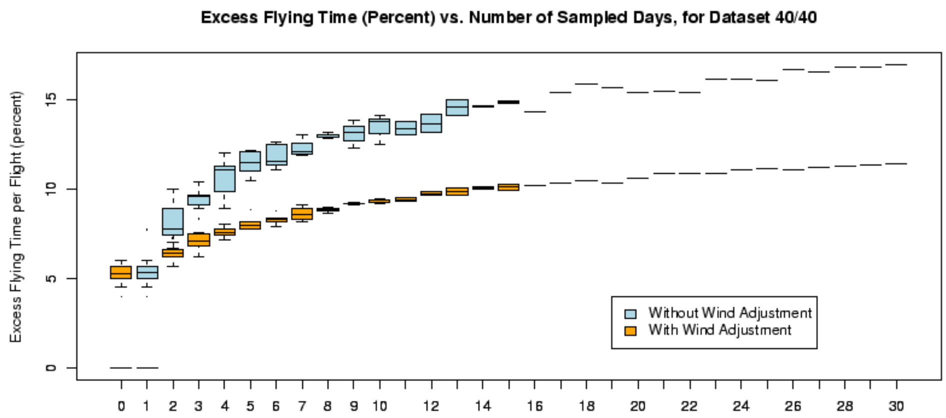

Figure 3. Excess Flying Time (Percent) Versus Number of Sampled Days, Dataset 40/40, With and Without Wind Adjustment

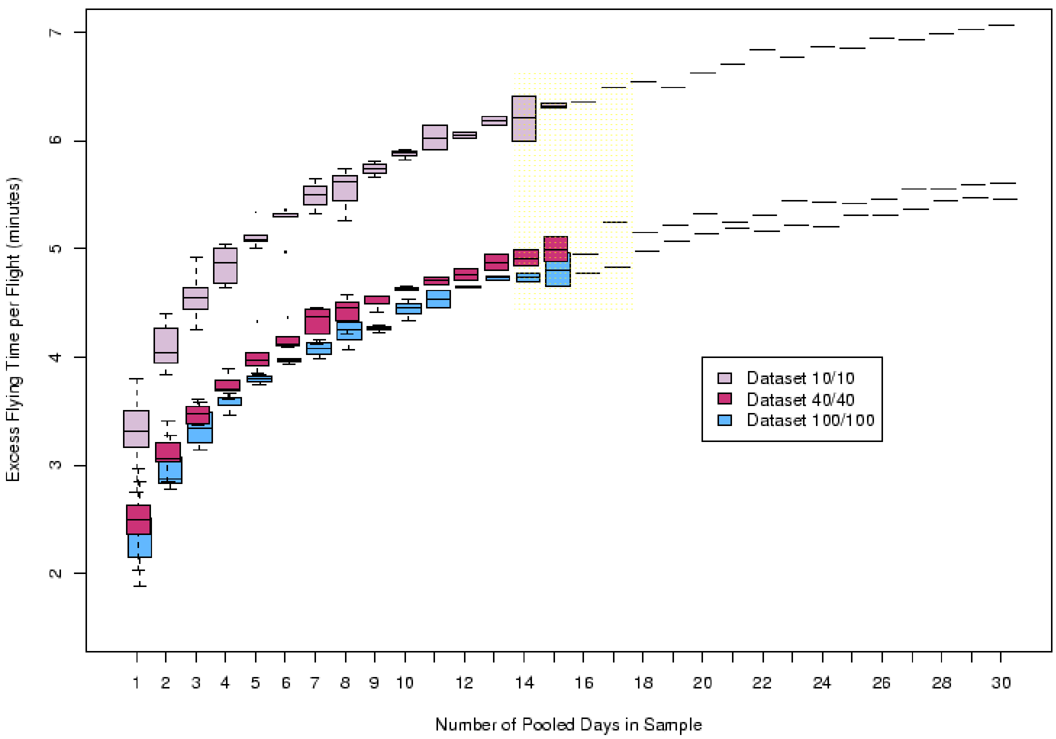

Figure 4. Excess Flying Time (Minutes) Versus Number of Sampled Days, All Datasets

Table 1. Average Excess Flying Times for 15-Day Pools With Wind Adjustment

Dataset	Average Minutes Excess	Average Percent Excess
10/10	6.4	9.5
40/40	4.9	10.0
100/100	4.7	8.1

were the same as that used with the flight-based method, ETMS data for the CONUS.

The method relies on the overlay of an imaginary grid, a collection of equi-spaced horizontal and vertical lines which define cells, on the CONUS. Excess flying time is assessed for each individual grid cell, with color displays of the results. A first consideration was the cell size. After some analysis, we decided on cells with 50 nmi on a side. However, flights with less than 25 nmi flight distance through a cell were discarded, since we wanted to avoid the possible increased variance associated with transient flights such as “corner cutters.” In addition, if very short traversals were allowed, then interpolation error would become a larger percentage of the estimated traversal time.

The steps of the methodology were as follows:

Step 1. Compute wind-adjustment value for each combination of: cell, altitude, aircraft type, and direction of flight (referred hereafter as “combination”). Just as with the flight-based method, we employed the “opposing flight” logic to discern wind effects, and enforced a minimum sample size rule. Altitude data was stratified to three levels: 0-18K′ MSL, 18K′ MSL-FL290, above FL290, (where MSL=Mean Sea Level and FL=Flight Level). Direction of flight was divided into eight points of the compass.

Step 2. Create a look-up table of wind-adjustment values for each day of interest.

Step 3. Analyze entire dataset to find the best speed (after wind adjustment) for each cell/combination. Cell traversal speed is the arbiter of excess flying time, since flights have differing cell traversal distances. The set of best speeds per cell/combination establishes a benchmark against which computed excess flying times will necessarily be positive (or zero).

Step 4. Select one or more days for analysis. For each cell, for its set of flight traversals, compare flight speeds (after wind adjustment) to the best speeds of Step 3. Flights with speeds slower than the benchmark are considered to have deficit speed, from which can be computed an excess flying time. Combinations are aggregated as appropriate for analysis.

An Application of the Cell-Based Approach

Recently, some questions have arisen at the FAA regarding the en route airspace: Are current traffic patterns different from those seen in 2000? Is there more traffic, with more delay? (Following the traffic decrease after the September 11, 2001 (9/11) tragedy, traffic levels in 2004 approached, and in some quarters exceeded, the very busy traffic year 2000.) We selected 22 weekdays in 2000 and 22 weekdays in 2004 with generally good weather, and with similar *misery index* values for comparison. A benchmark best-speeds matrix was computed by examining all of the available data for 2000 and 2004, guaranteeing that excess flying times in each cell/combination would be non-negative.

Figures 5, 6, and 7 show the results, each being a difference of the daily average of 2004 minus 2000. Figure 5 shows total flight count difference, i.e., each cell is the difference, 2004 minus 2000, for the average daily (averaged over 22 subject days) flight count through the cell (summed over all altitudes and all directions of flight). Flight count is generally higher in 2004 than 2000 for most of the cells. Areas of greatest growth are ATL (Atlanta), CVG (Cincinatti), and the Washington to New York corridor. Figure 6 shows total excess flying time minutes difference, 2004 minus 2000, and highlights

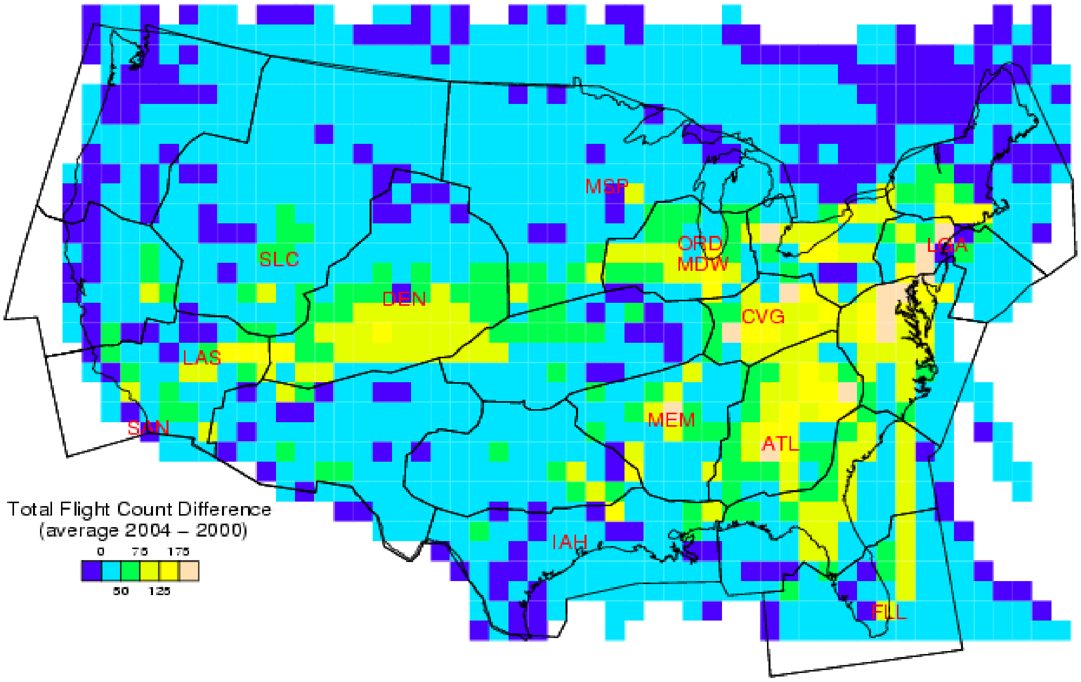

Figure 5: Total Flight Count Difference, Daily Average 2004 Minus 2000 (Units are Flights)

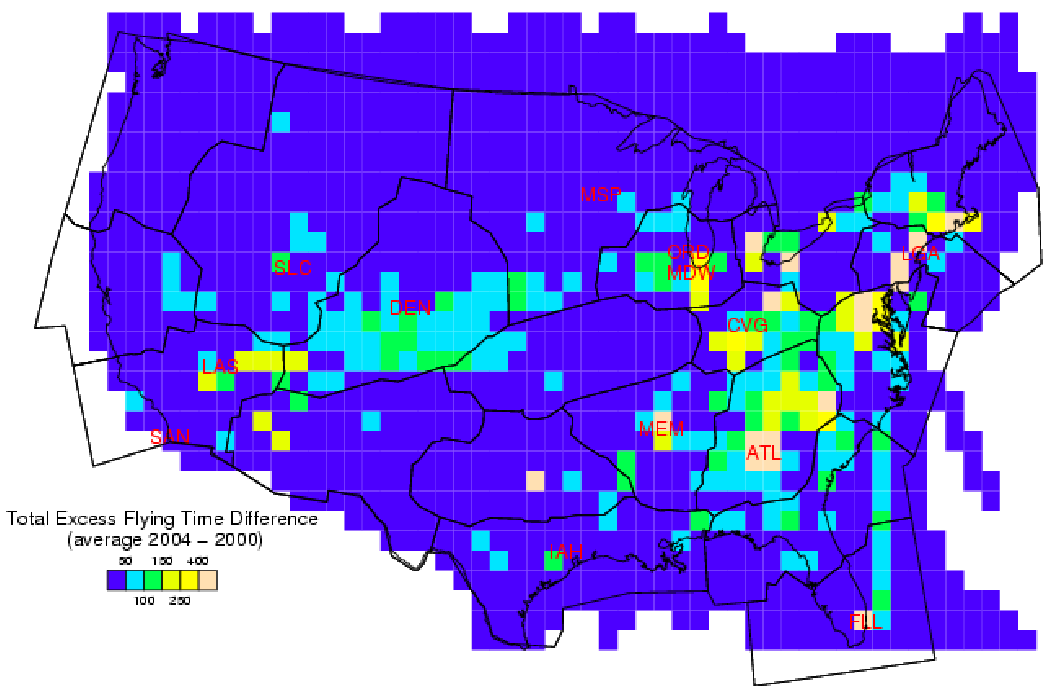

Figure 6: Total Excess Flying Time Difference, Daily Average 2004 Minus 2000 (Units are Minutes)

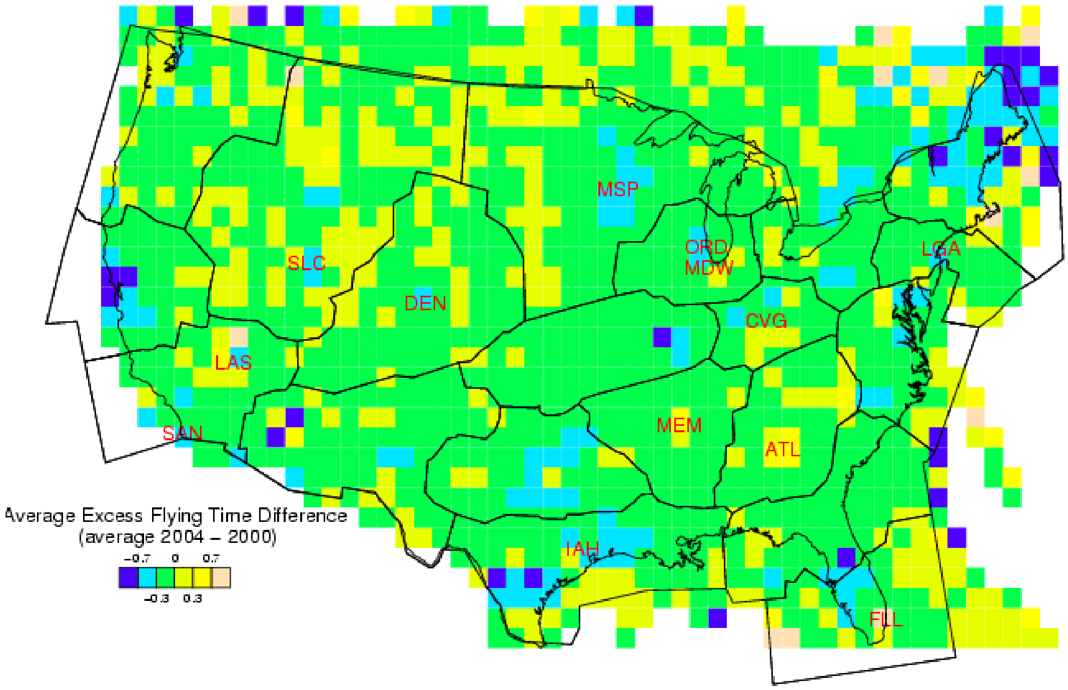

Figure 7: Average per Flight Excess Flying Time Difference, Daily Average 2004 Minus 2000 (Units are Minutes)

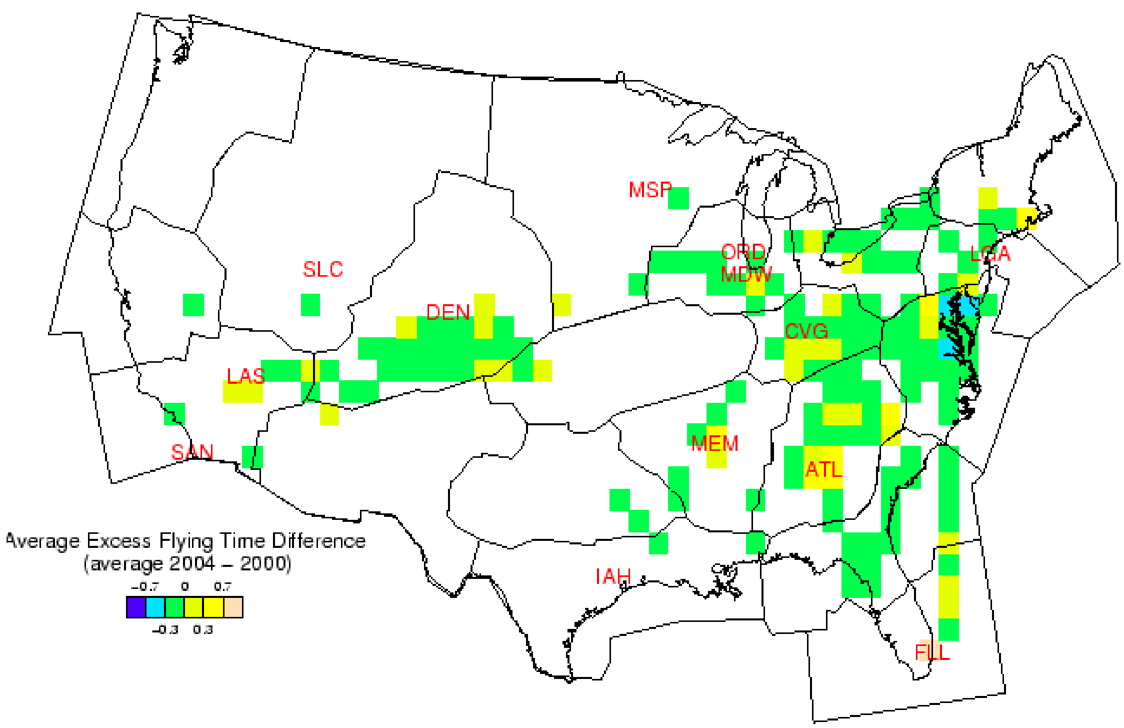

Figure 8: Average per Flight Excess Flying Time Difference for High-Difference Traffic Count Cells, Daily Average 2004 Minus 2000 (Units are Minutes)

many of the same regions as Figure 5 does. That is, not surprisingly, excess total flying time difference is greatest where flight count difference is greatest. Note the vertical bar on the lower right, which is the over-ocean route between Florida and cities in the northeast. Figure 7 shows average excess flying time per cell difference, 2004 to 2000. Here the color scale is symmetric, and values are mostly less than zero, meaning that 2004 saw generally less excess flying time per flight than did 2000. Figure 8 focuses the results of Figure 7, displaying only the high-difference-cell-count cells (those cells of Figure 5 with an increase of >75 average flight traversals per day). Among these displayed cells, there are more than three times more cells with *lower* average excess flying time in 2004 than with higher.

It is encouraging that, per Figures 7 and 8, 2004 is generally better than 2000. It may be that the operations in the NAS are becoming more efficient, although this is conjecture and would require much more evidence to substantiate. However, one can also spot higher delay regions, for example, near ATL (Atlanta), FLL (Ft. Lauderdale, Florida), and LAS (Las Vegas). Obviously, these new high delay regions are good targets for instigation of procedures and tools which would ameliorate the situations.

Although preliminary, this example demonstrates the ability of the cell-based method to identify geographic areas of interest with respect to delay and congestion.

Conclusion and Next Steps

An examination of excess flying time in the U.S. has been described in this paper using two different but compatible methods. The flight-based method computes flight-specific excess flying time, and finds 4 to 5 minutes average per flight in the en route regime, with an additional 1.5 minutes if the terminal area is also considered. The cell-based method computes excess flying time using an imaginary grid overlaying the U.S., resulting in cell-specific results. Both approaches rely on a minimum-observed flying time as the basis of computation. The second technique allows a focus on geographic region of interest, and should help decision makers focus scarce resources.

This analysis is proceeding by further examination of differences between traffic levels and traffic delay in 2000 (i.e., pre-9/11) and 2004 (when traffic levels approach or exceed pre-9/11 levels).

References

1. Federal Aviation Administration, 2004, Aviation System Performance Metrics (ASPM), see http://www.apo.data.faa.gov/

2. Federal Aviation Administration, National Airspace System Operational Evolution Plan (OEP), A Foundation for Capacity Enhancement 2004 – 2014, Version 6.0.

3. Bolczak, C., L. Brown, J. Hoffman, C. Schellenberg, and W. Voss, 1997, "Flexibility, Predictability, and Access: An Expanded View of System Performance," ATC Quarterly, ATCA Institute, Arlington, VA.

4. Ball, M., J. DeArmon, J. Pyburn, 1995, "Is Free Flight Feasible? Results from Initial Simulations," Journal of Air Traffic Control, ATCA, Arlington, VA.

5. Willemain, T., 2001, "Estimating Components of Variation in Flying Times," http://www.its/berkeley.edu/nextor/pub/WP-01-2.pdf.

6. Bureau of Transportation Statistics, 2002, Airborne Flying Time Analysis, http://www.bts.gov.ntda/oai/DetailedStatistics/.

7. Solomos, G., D. Knorr, M. Blucher, J. DeArmon, G. Dorfman, H. Yi, 2003, "Analysis of Excess Flying Time in the National Airspace System," submitted for publication.

8. Bureau of Transportation Statistics, Office of Airline Information, 2002, http://www.bts.gov/oai/.

9. Callaham, M. et al., 2001, “Assessing NAS Performance: Normalizing for the Effects of Weather,” Fourth USA/Europe Air Traffic Management R&D Symposium, Santa Fe, NM.

10. Oiesen, R., 2000, “The ETMS Hub Site and Weather,” http://www.caasd.org/library/presentations/weather_conference/2_Oiesen.pdf.

11. Hartigan, J., 1975, *Clustering Algorithms,* Wiley and Sons, Inc., New York.

12. Tukey, J., 1977, *Exploratory Data Analysis*, Addison-Wesley, Reading, PA.

Key Words

Excess flying time, air traffic, en route airspace, grid overlay

Author Biographies

James DeArmon is a Principal Engineer with The MITRE Corporation. He has worked for the last several years in research and modeling of air traffic flow dynamics. He holds a Master’s degree in Operations Research, and has taught graduate courses at the University of Maryland. He has written numerous papers in aviation and applied math.

Dr. Gerald Dorfman is a Lead Scientist at The MITRE Corporation's Center for Advanced Aviation System Development. His current focus is on statistical analysis of the U.S. National Airspace System (NAS). He views the NAS as a system whose performance can be analyzed using industrial quality control methods.

George H. Solomos is an Associate Program Manager at The MITRE Corporation’s Center for Advanced Aviation System Development, working closely with the Federal Aviation Administration (FAA). Most recently, Mr. Solomos has worked on assessing the benefits of system improvements around the National Airspace System. He is experienced in computer modeling and simulation, pre- and post-implementation assessment, and future performance forecasting.

Ho Yong Yi is senior staff at The MITRE Corporation's Center for Advanced Aviation System Development. He received a B.S. in aviation computer science and M.S. in software engineering from the Embry-Riddle Aeronautical University in Florida. His work includes software development, system analysis, and system testing.