Ninth USA/Europe Air Traffic Management Research and Development Seminar (ATM2011)
Berlin, Germany, 14 – 17 June 2011

New Method for Probabilistic Traffic Demand Predictions for En Route Sectors Based on Uncertain Predictions of Individual Flight Events

Eugene P. Gilbo and Scott B. Smith

J. A. Volpe National Transportation Systems Center
55 Broadway, Kendall Square, Cambridge, MA 02142, U.S.A.

{eugene.gilbo, scott.smith}@dot.gov

Abstract

Federal Aviation Administration (FAA) Air Traffic Flow Management (TFM) decision-making is based primarily on a comparison of predicted traffic demand and capacity (usually called Monitor/Alert Parameter, or MAP) at various elements of National Airspace System (NAS) such as airports, fixes and sectors to identify potential congestion. The current FAA Traffic Flow Management System (TFMS) and its decision-support tools operate with deterministic predictions and do not consider the stochastic nature of the predictions. Sector demand predictions appear to be less accurate and stable than predictions for airports and fixes. The major reason is that, unlike airports and fixes where flights are aggregated in 15-minute intervals, TFMS predicts sector demand by aggregating flights for each minute and using the one-minute peak demand as a measure for sector demand for entire 15-minute interval. This paper presents a novel analytical approach to and techniques for translating characteristics of uncertainty in predicting sector entry times and times in sector for individual flights into characteristics of uncertainty in predicting one-minute sector demand counts. The paper shows that expected one-minute sector demand predictions are determined by a probabilistically weighted average of one-minute sector entry demand predictions for several consecutive one-minute intervals within a sliding time window. The width of the window is determined depending on probability distributions of errors in flights' sector entry time predictions. Expected one-minute sector demands along with standard deviations of demand counts are expressed via probabilistic averaging of series of one-minute deterministic predictions of number of flights entering a sector. The results of the paper contribute to probabilistic predictions of congestion in airspace. These analytical results can also be used to evaluate the impact of improved accuracy in flight timing predictions on reducing uncertainty in traffic demand predictions, hence leading to better identification of congestion in airspace.

1. Introduction

Federal Aviation Administration (FAA) Air Traffic Flow Management (TFM) decision-making is based primarily on a comparison of traffic demand and capacity predictions at various National Airspace System (NAS) elements such as airports, fixes and en-route sectors. The current Traffic Flow Management System (TFMS) and its decision-support tools deal with deterministic predictions and neglect the stochastic nature of the predictions.

An important part of the Next Generation Air Transportation System (NextGen) program is the transition from deterministic to probabilistic TFM that would lead to more realistic and efficient TFM decisions.

Research on probabilistic TFM and related problems is underway in many organizations including government, private sector and academia. Recent results demonstrate the potential benefit of applying a stochastic approach to TFM (see, e.g., [1] – [13]).

A general concept of probabilistic TFM for managing congestion in the NAS under uncertain predictions of demand and capacity, as well as an incremental, probabilistic approach to TFM decision-making can be found in [10] and [11], respectively. Research results presented in [12] and [13] provide an important contribution to probabilistic TFM describing a constructive approach to incorporating probabilistic weather forecasts into probabilistic TFM, as well as design of the modeling tool for evaluating TFM strategies.

The practical value of decision-support tools used for probabilistic TFM will significantly depend on quality of data used in the tools to represent uncertainty in the aviation system, in particular, uncertainty in predicting traffic demand and capacity. Therefore thorough data analysis along with analytical tools are needed to examine the sources of uncertainty, to characterize uncertainty and to apply advanced statistical methods for reducing uncertainty in predictions.

Research on uncertainty, which is the core issue for probabilistic TFM, is conducted in two major directions: uncertainty in traffic demand predictions and uncertainty in predicting the capacity of NAS elements. Each of those directions addresses two major elements of the NAS: airports and airspace. Although airports and airspace are equally important in the NAS and are generally subject to similar sources of uncertainty in traffic demand and capacity predictions, there are substantial differences in measuring both demand and capacity in airports and airspace. For example, both demand and capacity are better defined and measured for airports than for en route sectors. Moreover, because of the differences in measuring traffic demand in airports and in sectors in the current TFMS, the TFM specialists noticed that demand predictions in sectors are more uncertain, more volatile and less reliable than those for airports. The TFMS measures traffic demand for airports and sectors for each 15-minute interval of the time period of interest. The principal differences in measuring traffic demands for sectors and airports are as follows:

1. Traffic demand in sectors is based on one-minute aggregate counts vs. 15-minute arrival or departure counts at airports.
2. Current TFMS determines traffic demand in sectors on a 15-minute basis and considers the maximum one-minute count within a 15-minute interval as the traffic demand for the sector for entire 15-minute interval, while a 15-minute traffic demand for airports includes all flights with ETAs within the 15-minute interval.
3. Significant fractions of one-minute demand counts in adjacent one-minute intervals in a sector might contain the same flights while, at airports, adjacent 15-minute intervals contain different flights.

Our research has been focused on developing a methodology that allows for a quantitative representation of uncertainty in air traffic demand predictions for NAS elements [1] - [6].

Our previous research, reported in [1] – [3], was focused on the accuracy of TFMS 15-minute aggregate traffic demand predictions without considering individual flights that comprise those aggregate counts. We also developed a regression model aimed at improving the accuracy and stability of those aggregate predictions. The regression model included demand predictions at three consecutive 15-minute intervals with the interval of interest in the middle of the 45-minute time window. Including demand predictions in adjacent intervals implicitly takes into account the effect of uncertainty in predictions of arrival times for individual flights. The regression improved both the accuracy of demand predictions and the stability and accuracy of TFMS Monitor/Alert.

The next step in our research, reported in [4] and [5], was focused on analyzing uncertainty in predictions of airport arrival times for individual flights and developing a methodology for translating the uncertainty in estimated time of arrival (ETA) for individual flights into uncertainty in aggregate 15-minute traffic demand predictions for arrival airports. The result was a methodology for probabilistic traffic demand predictions at airports with quantitative characteristics of uncertainty in the predictions.

The motivation for using uncertainty in predicting times for individual flights in the characterization of uncertainty in predicting aggregate traffic demand is as follows.

Current TFMS provides deterministic demand predictions by aggregating flights whose estimated times of arrival or departure (ETAs or ETDs) fall within a time interval of interest without considering uncertainty (random errors) in flight's ETA or ETD. Those errors are the major contributors into uncertainty in aggregate demand count predictions. The "physical" mechanism that causes the errors in aggregate demand is migration of some flights' ETAs from one time interval to another during flight updates due to random errors, so that a flight that counted in one interval can be counted in another (earlier or later) interval after updating its ETA. The question is how to formalize the translation of characteristics of uncertainty in individual flight timing predictions into characteristics of uncertainty in aggregate traffic demand predictions.

A method that allowed for this translation was first proposed in [7]. The method provided the analytical means for obtaining a probability distribution of aggregate demand at a specific time interval of interest through probability distributions of errors in times of arrivals of individual flights predicted to arrive at this specific interval of interest. The method, however, did not consider an extended set of flights that also includes the flights predicted to arrive in several adjacent intervals. Considering the probabilities for those flights to arrive within the interval of interest will affect the probability distribution of aggregate demand for the interval of interest.

Another approach was proposed in [8] for probabilistic prediction of aggregate traffic demand in en route sectors based on probabilistic characteristics of uncertainty for individual flights' times to be in a sector. In particular, the paper focused on considering the probability distributions of flight departure times from origin airports for estimating the expected number of flights in a sector.

The paper presents a methodology and analytical results that demonstrate how characteristics of uncertainty in prediction of times for individual flights translates into characteristics of uncertainty in prediction of aggregate one-minute traffic demand counts in sectors. Like our previous work, it is based on a statistical analysis of current TFMS data. The translation of the characteristics of uncertainty in TFMS predictions of times for individual flights into characteristics of uncertainty in predictions for aggregate demand counts is a challenging problem. However, it is much more complicated for sectors than for airports because of the differences in measuring traffic demands for these NAS elements.

The paper is organized as follows.

Section 2 presents characteristics of uncertainty in TFMS predictions of times for individual flights to cross sector boundary and be in a sector

Section 3 presents the methodology for recalculating characteristics of uncertainty in individual flights' timing predictions into characteristics of uncertainty in aggregate one-minute demand counts for both entering a sector and being in a sector.

Conclusions are given in Section 4.

The Appendix contains a more detailed derivation of mathematical expressions necessary for probabilistic sector demand predictions.

2. Characterization of uncertainty in TFMS flights' sector entry and sector occupancy time predictions

TFMS estimates and periodically updates sector entry times for individual flights. To estimate the accuracy of those predictions, the TFMS data was collected during the days of moderate demand when there were no TFM initiatives. Without such interference by TFM control, the errors in predictions can be measured by the difference between predicted and actual times. The analysis was conducted on the TFMS data for sixteen en route sectors: ZBW02, ZBW09, ZBW17, ZBW20, ZBW46, ZID82, ZID83, ZID86, ZLC06, ZLC16, ZMP20, ZOB57, ZOB67, ZOB77, ZSE14, and ZTL43. The data included repeated updates for flight sector entry and sector exit times. Altogether, during April 10 through April 16 of 2009, there were approximately 834,000 time predictions for 39,000 flights analyzed. The look-ahead times (LAT) for predictions varied from 0 (for actual times) to 3 hours.

The data analysis was performed on the above mentioned data and on additional data set collected in April and June of 2007 and reported in [4]. The detailed results of analysis of accuracy in sector entry time predictions can be found in the Volpe report [6]. The distribution of prediction errors and their parameters (average and standard deviation) were estimated separately for active (airborne) flights and for proposed (not yet departed) flights.

Figure 2-1 illustrates the typical probability density functions of prediction errors in sector entry times for active and proposed flights.

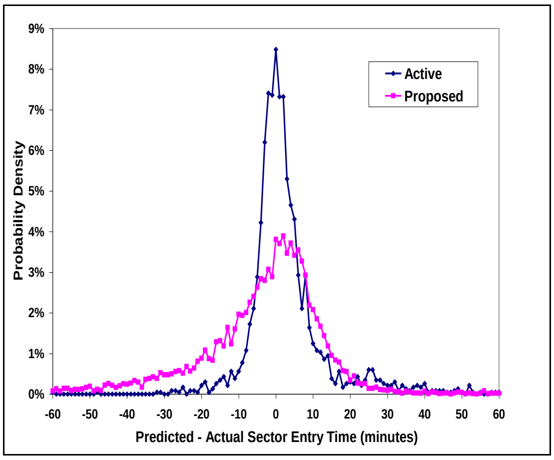

Figure 2-1 Probability density of errors in sector entry time predictions

The results of analysis show and Figure 2-1 illustrates that

- For active flights, the distributions of prediction errors are somewhat asymmetric with a heavier right-hand tail, indicating a tendency for flights to enter sectors earlier than predicted
- For the proposed flights, the distributions of prediction errors are asymmetric with heavy left-hand tails, indicating a

tendency for flights on the ground to, on average, enter sectors later than predicted.

- The standard deviation of error is substantially lower for active than for proposed flights. It was in the 4 to 12 minute range for active flights, and in the 15 to 22 minute range for proposed flights.

Sectors vary in size and in the manner flights traverse them. As a result, there are significant differences among the sectors in time-in-sector for a flight. As for the accuracy of time-in-sector predictions, it is much better than for flight's sector entry times and is not much different for active and proposed flights and for different LAT. For example, the standard deviation of time-in-sector error was 4 minutes or less for both active and proposed flights [6].

The probability distributions of flight timing predictions are fundamental for aggregating flights in probabilistic one-minute sector demand predictions.

3. Probabilistic predictions of sector demand counts

The number of flights in a sector during a specific one-minute interval includes the flights that entered the sector during this interval and the flights that entered the sector earlier and are still in a sector. As standard deviations of errors in sector entry time predictions are significantly greater than one minute, there are significant probabilities that a flight would enter a sector earlier or later than the flight's ETA. Therefore, a probabilistic demand prediction at a specific one-minute interval would need to consider TFMS one-minute demand predictions for several consecutive one-minute intervals surrounding the interval of interest.

The probabilistic predictions of sector one-minute demand counts require the following steps:

1. Translate a flight's time predictions and associated prediction errors into the probability for the flight to enter a sector during a particular minute.

2. Develop a probabilistic characterization for the number of flights entering a sector, based on the probabilities from step 1.

3. Develop probabilistic count predictions for the number of flights present in a sector during a particular minute (a one-minute traffic demand counts for a sector).

Each step is considered below.

3.1 Probability for a flight to enter a sector during a one-minute interval.

Our report [6] presents a detailed analytical approach for determining probabilities for the flights to enter a sector during a one-minute interval depending on the flights' estimated sector entry times (ETAs) and probability distributions of errors in predicting sector entry times. The status of individual flights (active or proposed) is taken into account by using differences in accuracy of predicting sector entry time for active and proposed flights. This section presents an example of the flight probability to enter a sector during a one-minute interval.

Following the notation of [6],

i is a one-minute interval of interest that is between the beginning of minute i and the beginning of minute $(i+1)$,

k is a one-minute interval for a flight's estimated sector entry time (ETA),

$F(x)$ is a cumulative distribution function (CDF) of prediction error for sector entry time,

$P_{i,k}$ is a probability for a flight deterministically predicted to enter a sector during a one-minute interval k to enter a sector during a one-minute interval i,

$$P_{i,k} \approx 0.5\,[F(i-k+1) - F(i-k-1)].$$

It should be noted that, since the CDFs are different for active and proposed flights, separate calculations of probabilities are performed for active and proposed flights.

These formulas permit the calculation of probabilities for various intervals of interest, including a series of consecutive intervals, e.g., i, $i+1$, $i+2$, etc.

For the sake of simplicity, in the examples presented below, we assume that the prediction errors are symmetrically distributed.

Table 3-1 shows several values of probabilities for different k, surrounding the interval of interest i. This table illustrates the probabilities for a flight to cross a sector boundary at an interval of interest i if it deterministically predicted to enter a sector earlier ($k < i$) or later ($k > i$) or on time ($k = i$). It also illustrates the relative significance of the probabilities depending on the time difference $|i - k|$.

The probabilities have been calculated for the Gaussian distribution $F(x)$ with zero mean and standard deviations of σ = 4 minutes (which might correspond to the accuracy of predictions for active flights), and σ = 15 minutes (which might correspond to the accuracy of predictions for proposed flights).

Table 3-1 Probabilities for Flights to Enter a Sector during Interval i

Probability	σ = 4	σ = 15
$P_{i,i}$	0.099	0.027
$P_{i,i+1} = P_{i,i-1}$	0.096	0.027
$P_{i,i+2} = P_{i,i-2}$	0.087	0.026
$P_{i,i+3} = P_{i,i-3}$	0.075	0.026
$P_{i,i+4} = P_{i,i-4}$	0.060	0.026
$P_{i,i+5} = P_{i,i-5}$	0.046	0.025
$P_{i,i+6} = P_{i,i-6}$	0.033	0.025
$P_{i,i+7} = P_{i,i-7}$	0.022	0.024
$P_{i,i+8} = P_{i,i-8}$	0.014	0.023
$P_{i,i+9} = P_{i,i-9}$	0.008	0.022
$P_{i,i+10} = P_{i,i-10}$	0.005	0.021
$P_{i,i+15} = P_{i,i-15}$	0.000	0.016
$P_{i,i+20} = P_{i,i-20}$	0.000	0.009
$P_{i,i+25} = P_{i,i-25}$	0.000	0.007
$P_{i,i+30} = P_{i,i-30}$	0.000	0.004
$P_{i,i+35} = P_{i,i-35}$	0.000	0.002
$P_{i,i+40} = P_{i,i-40}$	0.000	0.001

These probabilities tend to be small (less than 0.10), since they correspond to single minutes. For example, if a flight is deterministically predicted to enter a sector two minutes earlier ($k = i$-2) or two minute later ($k = i$+2) than the interval of interest i, the probability for the flight to enter a sector during interval i is smaller: $P_{i,\, i-2} = P_{i,\, i+2} \approx 0.087$.

For the active flights (σ = 4 minutes) with ETAs at least nine minutes from the interval of interest i (earlier or later) the probabilities to enter a sector during interval i are negligibly small (less than 0.01).

It is important to note that when predictions of sector entry times are less accurate, the probability for a flight to enter a sector in a particular minute becomes smaller, even in the minute that it was forecast to enter the sector (approximately 0.027 when σ = 15).

Another important observation from Table 3-1 is that, since the standard deviations of prediction errors are much larger than one-minute, the probabilities for the flights to enter a sector do not vary much from one minute to the next. In particular, for σ = 4 min, the probabilities for flights with ETAs in intervals i, i+1 and i -1 to arrive to a sector in interval i are, respectively equal to 0.099, 0.096 and 0.096, i.e., they are nearly identical. For σ = 15, the corresponding probabilities a practically the same for the flights with ETAs in *thirteen (!)* intervals around the interval of interest i (including interval i).

3-2 Number of Flights Entering a Sector During a One-minute Interval

The results of the previous section can be used for characterization of uncertainty in prediction of number of flights entering a sector.

Our report [6] describes a detailed analytical approach to the probabilistic prediction of aggregate one-minute sector entry counts based on probabilities for individual flights to enter a sector and on deterministic predictions of sector entry counts at various one-minute intervals. In this analysis, independence of errors in predicting times among the flights is assumed.

The main result of taking into consideration characteristics of uncertainty in flights' sector entry time predictions is that the flights with ETAs in several adjacent one-minute intervals to the interval of interest will be considered in calculating the aggregate demand in the interval of interest. For example, if d_k flights deterministically predicted to enter a sector during one-minute interval k, there is the probability $P_{i,k}$ for each of the d_k flights to enter the sector during the one-minute interval of interest i. Assuming that the errors in flight arrival predictions are independent, the probability distribution of number of flights from d_k that can be counted in interval i is a binomial distribution. The total random number of flights predicted for a specific one-minute interval i is equal to the sum of the random numbers of flights from several adjacent intervals k that can be counted in the interval of interest i. The number of adjacent intervals β that should be taken into account depends on relative values of probabilities $P_{i,k}$. β = max | $i - k$ |, where max | $i - k$ | is the distance beyond which the probabilities $P_{i,k}$ become too small and should be neglected.

According to the properties of the binomial distribution, the expected one-minute count $\tilde{\bar{d}}_i$ and the standard deviation of the one-minute count are equal to, respectively [14]:

$$\tilde{\tilde{d}}_i = \sum_{k=i-\beta}^{i+\beta} P_{i,k} d_k , \qquad (3.1)$$

$$\sigma_i = \sqrt{\sum_{k=i-\beta}^{i+\beta} P_{i,k} (1 - P_{i,k}) d_k} , \qquad (3.2)$$

where d_k is a deterministic prediction of sector entry counts for a one-minute interval k.

The probabilistic prediction of one-minute flight counts entering a sector can be represented by the expected number $\tilde{\tilde{d}}_i$ and the uncertainty area around the expected number: $\tilde{\tilde{d}}_i \pm j\sigma_i$, where j determines a size of the uncertainty area restricted by certain percentiles. For j = 1, $\tilde{\tilde{d}}_i + \sigma_i$ corresponds to the 84th percentile, and $\tilde{\tilde{d}}_i - \sigma_i$ corresponds to the 16th percentile. For j = 2, the percentiles are 2.3% and 97.7%.

Consider an example of probabilistic prediction of number of flights entering a sector during a one-minute interval when the errors in predicting sector entry times for individual flights are normally distributed with zero average and with standard deviation of 4 minutes (σ = 4 min).

The probabilities $P_{i,k}$ in this case are shown in Table 3-1 for $k = i$-10, i-9, i-8, ..., i-1, i, i+1, …, i+8, i+9, i+10, i.e., for ten one-minute intervals from both sides of the interval i of interest. The values of probabilities $P_{i,k}$ for $k \le i - 9$ and $k \ge i + 9$, become too small and can be neglected. As a result, $\beta = 8$.

Figure 3-1 illustrates an example where the predicted one-minute count for the interval of interest (for i = 1200) is much higher than predictions in the adjacent intervals.

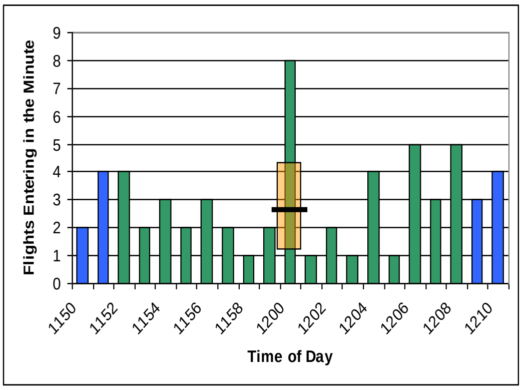

Figure 3-1 Sector Entry Counts

The figure shows expected values for i = 1200 and uncertainty areas around the expected values restricted by $\pm\ \sigma_i$. The expected values and standard deviations in those cases were calculated by using formulas (3.1) and (3.2) with numerical values of probabilities from Table 3-1 and demand values shown in the figure.

The expected number of flights entering a sector in 1200 one-minute interval is equal to 2.7 flights, which **is much smaller than deterministic prediction** of 8 flights. The standard deviation is equal to 1.6. The uncertainty are in the figure indicates that with the 0.68 probability the number of flights entering the sector during a one-minute interval i = 1200 is between 1.1 and 4.2 flights with the expected number of 2.7 flights.

This example illustrates how the probabilistic prediction for a particular one-minute interval depends on the predictions for several adjacent intervals. If the deterministic prediction for the minute is unusually high (as it was in Figure 3-1), the probabilistic prediction will be lower. Conversely, in cases where the deterministic prediction for the minute is exceptionally low, the probabilistic prediction will be higher (see Figure 3-2).

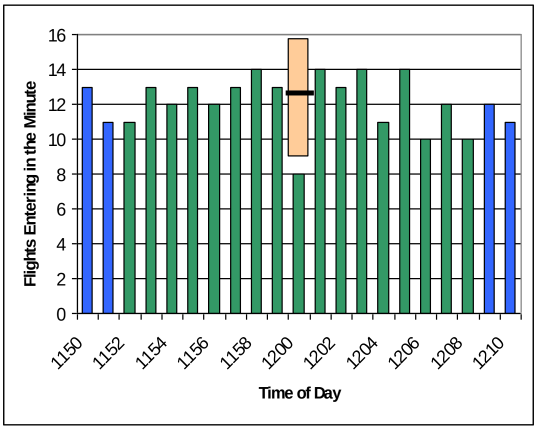

Figure 3-2 Sector Entry Counts

3-3 Number of Flights in a Sector During a One-minute Interval

In the previous section, we developed a probabilistic representation of the number of flights ***entering*** a sector at a particular 1-minute time interval i. To develop a representation of the number of flights ***in*** a sector at interval i, we need to consider an additional factor: the time-in-sector for individual flights.

The Appendix presents the mathematical expressions and analytical approach needed for probabilistic predictions of one-minute demand counts of the flights in a sector. The analytical formulas presented in the Appendix allow for determining probability distributions of one-minute demand predictions in a sector, calculating average and standard deviation of predicted number of flights in a sector at any one-minute interval by using deterministic predictions of sector demand and characteristics of uncertainty in predicting times of entering a sector for individual flights. The analytical results consider important factors for traffic demand predictions, such as status of individual flights (active or proposed) and difference in times in sector for various flights. What is important is that characteristics of probabilistic predictions of one-minute traffic demand in a sector depends heavily on probabilistic predictions of one-minute sector entry counts, considered in the previous section.

The starting point for calculating one-minute sector demand is constructing the relationship between number of flights in a sector and the number of flights entering the sector.

Let τ be the amount of time that a flight spends in a sector. If the flight's predicted entry time is k, its predicted sector exit time is $k+\tau$. In this analysis, τ is assumed to be known and non-probabilistic.[1]

The flight will be in the sector during minute i if the flight enters at or before minute i, and exits after minute i. In inequalities, this is

$k \leq i$ and $k+\tau > i$.

Hence, if d_k is a number of flights deterministically predicted to enter a sector during a one-minute interval k and time in sector for each flight is τ, then the deterministic prediction of number of flights in a sector $D_{i,\tau}$ during one-minute interval i is

$$D_{i,\tau} = d_i + d_{i-1} + d_{i-2} + \ldots + d_{i-(\tau-2)} + d_{i-(\tau-1)} = \sum_{j=i-\tau+1}^{i} d_j \;. \qquad (3.3)$$

For example, if k = 1158, i = 1200 and τ = 5 minutes, the flight will be in the sector at 1200 if the actual entry time is between 1156 and 1200. If the entry time is before 1156, the flight is too early, and will have left the sector before 1200. If it is after 1200, the flight is too late. This flight will be in the sector at 1200 if it enters the sector at any of the following times: 1156, 1157, 1158, 1159, and 1200.

[1] Recall from Section 2 that the error in time-in-sector is substantially lower than the error in sector entry time.

The numbers of predicted flights entering a sector are not deterministic because of random errors in predicting times of entering the sector for individual flights. This case was considered in the previous section.

In this case, the expected number of flights $\tilde{\tilde{D}}_{i,\tau}$ in a sector in a one-minute interval i is, thus the sum of the expected numbers of flights entering the sector over a series of times (see formula (A5) in the Appendix):

$$\tilde{\tilde{D}}_{i,\tau} = \sum_{j=i-\tau+1}^{i} \; \sum_{k=j-\beta}^{j+\beta} P_{j,k} d_k \qquad (3.4)$$

where $P_{j,k}$ is a probability for a flight to enter a sector during interval j if it is predicted to enter the sector during interval k.

The formula (3.4) can be modified and presented in a more convenient form so that the expected number of flights in a sector $\tilde{\tilde{D}}_{i,\tau}$ is equal to (see formula (A6) in the Appendix)

$$\tilde{\tilde{D}}_{i,\tau} = \sum_{k=i-\tau-\beta+1}^{i+\beta} P_{i,k,\tau}\, d_k \qquad (3.5)$$

where $P_{i,k,\tau}$ can be interpreted as a probability for a flight to be in a sector during minute i if it is deterministically predicted to enter a sector during minute k.

The standard deviation of one-minute sector demand counts is (see (A9) in the Appendix)

$$\sigma\,(\tilde{D}_{i,\tau}) = \sqrt{\sum_{k=i-\beta-\tau+1}^{i+\beta} P_{i,k,\tau}(1-P_{i,k,\tau})\, d_k} \qquad (3.6)$$

The formula for calculating probabilities $P_{i,k,\tau}$ is given in the Appendix (see formula (A7)).

The analytical expression for probabilities $P_{i,k,\tau}$, selects the probabilities $P_{i,k,\tau}$ for certain combinations of i and k minutes only, for which the flights deterministically predicted to arrive in a sector at minute k satisfy specific constraints to be in a sector during minute i.

The above results can be easily expanded to the case where flights have different time in sector (τ), since the probabilistic prediction is simply a matter of summing all of the predictions across the various in-sector times. The resulting formulas can be found in [6].

4. Conclusions

The paper introduced a new analytical method for probabilistic prediction of traffic demand in en route sectors via transferring probabilistic characteristics of uncertainty in prediction times for individual flights into characteristics of uncertainty in aggregate one-minute demand predictions for sectors.

The results of the analysis of uncertainty in individual flight predictions were used to develop a new methodology for probabilistic predictions of aggregate, one-minute traffic demand counts for sector boundary crossing and for the number of flights in a sector.

Because the errors in sector entry time predictions are usually much greater than one minute, the probabilities for an individual flight with an ETA at a specific one-minute intervals to enter a sector at that interval is small (often less than 0.1). Moreover, the flights with ETAs close to (but not equal to) the one-minute interval of interest have nearly the same probabilities to enter a sector during the interval of interest. **This justifies a fundamental result of the paper that, if the prediction errors of arrival time for individual flights is much greater than the time interval of interest for aggregate count predictions, then probabilistic predictions of aggregate number of flights for this interval should also consider the flights with ETAs in several neighboring intervals.**

The paper presented an analytical approach and methodology for translating characteristics of uncertainty in individual flight's predictions into probabilistic predictions for sector demand counts. It was shown that probabilistic predictions of the number of flights in a sector are expressed through probabilistic predictions of number of flights entering a sector during several consecutive one-minute intervals that, in turn, depend on probabilities for individual flights to enter a sector during various consecutive one-minute intervals. The number of intervals involved in the probabilistic predictions depends on both the accuracy of predicting times for individual flights and the predicted time-in sector for the flights.

The results of the paper contribute to probabilistic predictions of congestion in airspace. These results can also be used to *analytically* evaluate the potential impact of improved accuracy in flight timing predictions on reducing uncertainty in traffic demand predictions, hence leading to better identification of congestion.

Acknowledgements

This research was sponsored by the Federal Aviation Administration. We thank John McCarron and Mike McKinney of the FAA for encouragement and support; we also thank our colleague Rick Oiesen for fruitful discussions and valuable suggestions.

References

[1] Smith, Scott and Eugene Gilbo, "Analysis of Uncertainty in ETMS Aggregate Demand Predictions", Volpe National Transportation Systems Center, Report no. VNTSC-ATMS-05-05, November 2005.

[2] Gilbo, Eugene and Scott Smith, "Reducing Uncertainty in ETMS Aggregate Traffic Demand Predictions," Volpe National Transportation Systems Center, Report no. VNTSC-CE-07-01, March 2007.

[3] Gilbo, Eugene and Scott Smith "A New Model to Improve Aggregate Air Traffic Demand Predictions," Paper number AIAA 2007-6450, AIAA Guidance, Navigation, and Control Conference and Exhibit, Hilton Head, SC, 20 -23 August 2007.

[4] Gilbo, Eugene and Scott Smith, "Characterization of Uncertainty in ETMS Flight Events Predictions and its Effect on Traffic Demand Predictions," Volpe National Transportation Systems Center, Report no. VNTSC-TFM-08-08, July 2008.

[5] Gilbo, Eugene and Scott Smith "Probabilistic Prediction of Aggregate Traffic Demand Using Uncertainty in Individual Flight Predictions," Paper number AIAA 2009-6494, AIAA Guidance, Navigation, and Control Conference and Exhibit, Chicago, IL, 10 – 13 August, 2009

[6] Gilbo, Eugene and Scott Smith, "Probabilistic Predictions of Traffic Demand for En Route Sectors Based on Individual Flight Data," Volpe National Transportation Systems Center, Report no. VNTSC-TFM-10-01, January 2010.

[7] Meyn, Larry A., "Probabilistic Methods for Air Traffic Demand Forecasting," Paper number 2002-4276, AIAA Guidance, Navigation, and Control Conference and Exhibit, Monterey, CA, 5 - 8 August 2002.

[8] Tu, Y., M.O. Ball and W. Jank, "Estimating Flight Departure Delay Distributions: a Statistical Approach with Long-term Trend and Short-term Pattern" Journal of the American Statistical Association, Vol. 103, pp. 112-125, 2008.

[9] Wanke, Craig R., Michael B. Callaham, Daniel P. Greenbaum, Anthony J. Masalonis, "Measuring Uncertainty in Airspace Demand Predictions for Traffic Flow Management Applications," AIAA Guidance, Navigation, and Control Conference and Exhibit, Paper # AIAA-2003-5708, Austin, TX, 11 – 14 August, 2003.

[10] Wanke, C., Song, L., Zobell, S., Greenbaum, D. and Mulgund, S., "Probabilistic Congestion Management," 6th USA/Europe Air Traffic Management R&D Seminar, Baltimore, MD, June 27-30, 2005

[11] Wanke, C., and Greenbaum, D., "Incremental, Probabilistic Decision-Making for En Route Traffic Management", 7th USA/Europe ATM R&D Seminar, Barcelona, Spain, July 2007

[12] Hunter, George, and Kris Ramamoorthy, "A Trajectory-Based Probabilistic TFM Evaluation Tool and Experiment", ICNS Conference, Baltimore, MD, May 2006

[13] Ramamoorthy, K., Boisvert, B., Hunter, G., "A Real-Time Probabilistic Traffic Flow Management Evaluation Tool", 25th Digital Avionics Systems Conference, 2006 IEEE/AIAA, pp. 1 – 13, Portland, OR, 15-19 Oct. 2006.

[14] R. C. Pfaffenberger and J. H. Patterson, "Statistical Methods," Third Edition, IRWIN, 1987

Appendix. Probabilistic characterization of number of flights in a sector during a one minute interval

In order to determine whether a flight is within a sector during a certain one minute interval, one needs to know the flight's entry and exit time (the difference gives the total time for a flight to be in the sector, or the time required to traverse the sector). Aggregating all the flights that are predicted to be within the sector during a specific one-minute interval would provide the aggregate sector traffic demand for this minute.

Suppose that the time to traverse the sector for each flight entering the sector is the same and equal to τ minutes. It means that each flight spends τ minutes within the sector since its entry before leaving the sector. Therefore, for each one-minute interval of interest, for example, interval *i*, the flights will be in the sector during this interval if they entered the sector during the interval *i* and $(\tau - 1)$ preceding one-minute intervals, i.e., during the one-minute intervals $i, i - 1, i - 2, \ldots, i - (\tau - 2), i - (\tau - 1)$.This provides the opportunity to determine the sector traffic demand count predictions at any one-minute interval through the series of one-minute aggregate number of flights *entering* the sector during several consecutive one-minute intervals. In particular, deterministically predicted traffic demand D_i in a sector for a one-minute interval *i* is equal to

$$D_{i,\tau} = d_i + d_{i-1} + d_{i-2} + \ldots + d_{i-(\tau-2)} + d_{i-(\tau-1)} = \sum_{j=i-\tau+1}^{i} d_j \,, \qquad \text{(A1)}$$

where d_k is deterministically predicted number of flights entering a sector during a one-minute interval *k*.

It is important to notice that each component d_k in (A1) consists of different flights.

The deterministically predicted number of flights in a sector during s one-minute intervals immediately following the interval *i* can be expressed recursively as follows

$$D_{i+s,\tau} = d_{i+s} + D_{i+s-1,\tau} - d_{i-\tau+s}\,, \quad s = 0, 1, 2, 3, \ldots \qquad \text{(A2)}$$

Formula (A2) shows that the one-minute sector demand for the $(i + s)$ minute is equal to the demand prediction in the previous $(i + s - 1)$ one-minute interval plus the number of flights entering a sector during the $(i + s)$ minute minus the number of flights $d_{i-\tau+s}$ that left the sector during the $(i + s)$ minute (those flights are the ones that entered the sector τ minutes prior to the $(i + s)$ minute).

Due to random errors in predicting number of flights that enter a sector, the sector demand for a one-minute interval is actually a random number that can be represented by a formula like (A1), where each component of the sum is a random number of flights entering a sector during the corresponding one-minute intervals. If $\tilde{D}_{i,\tau}$ is the random number of flights predicted to be in a sector during one-minute interval *i* then it can be determined as follows:

$$\tilde{D}_{i,\tau} = \tilde{d}_i + \tilde{d}_{i-1} + \tilde{d}_{i-2} + \ldots + \tilde{d}_{i-(\tau-2)} + \tilde{d}_{i-(\tau-1)} = \sum_{j=i-\tau+1}^{i} \tilde{d}_j \,, \qquad \text{(A3)}$$

where $\tilde{d}_j$ is a random number of flights predicted to enter a sector during a one-minute interval *j*.

Therefore, according to (A3), the average predicted traffic demand $\overline{\tilde{D}}_i$ in a sector for a one-minute interval i is equal to the sum of average numbers of flights entering the sector during τ consecutive one-minute intervals: during the interval *i* and *(τ – 1)* intervals preceding interval *i*:

$$\overline{\tilde{D}}_{i,\tau} = \sum_{j=i-\tau+1}^{i} \overline{\tilde{d}}_j , \qquad \text{(A4)}$$

where the average number of flights $\overline{\tilde{d}}_j$ predicted to cross the sector boundaries during one-minute interval *j* (for $j = i, i-1, i-2, \dots, i-\tau+1$) are determined by formula (3.1).

After substituting $\overline{\tilde{d}}_j$ in (A4) with (3.1), the equation (A4) is as follows:

$$\overline{\tilde{D}}_{i,\tau} = \sum_{j=i-\tau+1}^{i} \sum_{k=j-\beta}^{j+\beta} P_{j,k} d_k . \qquad \text{(A5)}$$

After several simple transformations, equation (A5) can be rewritten as follows:

$$\overline{\tilde{D}}_{i,\tau} = \sum_{k=i-\tau-\beta+1}^{i+\beta} P_{i,k,\tau}\, d_k , \qquad \text{(A6)}$$

where $P_{i,k,\tau}$ is the probability for a flight to be in a sector during a one-minute interval *i* if it is deterministically predicted to enter a sector during one-minute interval *k* and be in a sector during time interval τ.

The probabilities $P_{i,k,\tau}$ are determined by the following summation, for various values of *k*:

$$P_{i,k,\tau} = \sum_{j=\max(k-\beta,\, i-\tau+1)}^{\min(i,\, k+\beta)} P_{j,k} \qquad \text{(A7),}$$

where $i-\beta-\tau+1 \le k \le i+\beta$

where $P_{j,k}$ is a probability for a flight to enter a sector during a one-minute interval *j* if its ETA to enter a sector is in the one-minute interval *k*. In other words, this summation states that:

- the relevant values of *j* are those ranging from *i-τ+1* to *i*, and

- we are only interested in those values of *j* that are within *β* of *k*.

If a flight deterministically predicted to enter a sector during interval k has the probability $P_{i,k,\tau}$ to be in a sector during interval *i*, then the predicted number of flights in the sector during the one-minute interval *i* is a binomially distributed random number.

The variance Var ($\tilde{D}_{i,\tau}$) and standard deviation σ($\tilde{D}_{i,\tau}$) of one-minute counts in a sector during minute i are, respectively, equal to

$$\mathrm{Var}(\tilde{D}_{i,\tau}) = \sum_{k=i-\beta-\tau+1}^{i+\beta} P_{i,k,\tau}(1-P_{i,k,\tau})\, d_k \qquad \text{(A8)}$$

$$\sigma(\tilde{D}_{i,\tau}) = \sqrt{\sum_{k=i-\beta-\tau+1}^{i+\beta} P_{i,k,\tau}(1-P_{i,k,\tau})\, d_k} \,. \qquad \text{(A9)}$$

The expressions (A6) and (A9) can be easily extended in the case when traffic demand contains both active and proposed flights. The corresponding formulas can be found in [6].

This section presented the basic results necessary for probabilistic characterization of one-minute sector traffic demand that consists of flights with the same time for traversing a sector.

Analytical results for the case when the flights have different times in sector are presented by the authors in [6].

,

Eugene P. Gilbo is a senior operations research analyst of the Traffic Flow Management Division at the Volpe National Transportation Systems Center, Cambridge, MA. For more than 20 years, he has been a senior researcher for the FAA's operational Traffic Flow Management System. His research is focused on optimization of air traffic flow management (TFM) strategies with an emphasis on airport arrival and departure operations ,airport and airspace capacity issues, TFM decision-making under uncertainty as well as on statistical data analysis and system performance evaluation. He is the author and co-author of more than 90 publications including the book (co-authored with I. B. Chelpanov) "Signal Processing Based on Order Statistics" (1975, in Russian). Dr. Gilbo holds the M.S., Ph.D. and Doctor of Sciences degrees in Mechanical Engineering, Electrical Engineering and Applied Mathematics, respectively, from Leningrad Polytechnic Institute (now St. Petersburg Technical University).

Scott B. Smith is a senior level operations research analyst with wide-ranging experience in the analysis and modeling of transportation systems. His work had included analysis and improvement of air traffic management predictions, an examination of the potential impact of Intelligent Transportation Systems on regional planning models, and the development of several traffic assignment and simulation models. Project sponsors have include most of the modal administrations in U.S. DOT as well as local agencies. Dr. Smith holds a Doctor of Philosophy and a MS in Transportation from the Massachusetts Institute of Technology. He is a member of the Institute for Operations Research and the Management Sciences (INFORMS), the American Institute of Aeronautics and Astronautics (AIAA) and the Project Management Institute.