Monitoring and Alerting Congestion at Airports and Sectors under Uncertainty in Traffic Demand Predictions

Eugene Gilbo and Scott Smith

Important functions of the Traffic Flow Management System (TFMS) include predicting air traffic demand for National Air Space (NAS) elements (airports, fixes and en route sectors) for several hours into the future, and using these predictions to alert traffic flow management (TFM) specialists to potential congestion when predicted demand exceeds available capacity. The current TFMS Monitor/Alert functionality uses deterministic predictions, neglecting their stochastic nature. This paper focuses on improving the accuracy and stability of traffic demand predictions for airports and sectors by considering the uncertainty in aggregate demand count predictions. The emphasis is on uncertainty caused by errors inherent in TFMS during processing flight data not affected by future air traffic control. We propose a constructive approach for improving aggregate demand predictions under uncertainty based on linear regression that includes TFMS demand counts for several adjacent time intervals within a sliding time window. Numerical examples based on TFMS data showed that the regression models produce more accurate (up to 22% reduction in the standard deviation of errors in demand predictions) and more stable (fewer crossings of the Monitor/Alert threshold) predictions than current TFMS predictions. For airports, regression significantly reduced the total number of missed alerts (21%) with a small increase in the total number of false alerts (3%). For sectors, the reduction in missed alerts was 22%, with an increase in false alerts of 8%.

The authors are with the Volpe National Transportation Systems Center, Cambridge, MA.

Received January 29, 2010; accepted March 26, 2011.

Air Traffic Control Quarterly, Vol. 19(2) 83–113 (2011) CCC 1064-3818/95/030163-20

INTRODUCTION

The key functions of the Traffic Flow Management System (TFMS) include predicting air traffic demand for National Air Space (NAS) elements (airports, fixes and en route sectors) for several hours into the future, monitoring predicted demands, and alerting traffic flow management (TFM) specialists to potential congestion at NAS elements wherever and whenever predicted demand exceeds available capacity. The information on the magnitude and duration of congestion helps a TFM specialist in the decision making process for resolving congestion problems.

The TFMS currently makes deterministic predictions of aggregate traffic demand counts per 15-minute interval and does not consider uncertainty in the predictions. A general methodology for traffic demand predictions in TFMS is based on processing the flight plan data for individual flights to predict times for flight events along the flight plan routes from origin to destination airport and aggregating the flights within specific time buckets. For the flights still on the ground, the prediction of future flight events starts from the estimated time of departure (ETD) at the origin airport. As soon as a flight departs, the TFMS uses the flight tracking radar data (e.g., en route TZ messages) that include coordinates of the flight in the airspace along with the time stamp to predict flight events starting from this point along the remaining route to the destination airport. The prediction algorithm uses aircraft speed and considers predictions for wind along the route. Many factors contribute to errors in predicting times of flight events. The major contributor to prediction errors for individual flights is uncertainty in flight departure time for those flights that are still on the ground. After a flight has departed, there is often some deviation from the flight plan route, so that the flight coordinates in TZ messages usually deviate from the plan. TFMS projects the coordinates of the last TZ message to the flight plan route and continues predicting flight events from this point on the route further along the planned route up to the destination airport. There are also differences between flight speeds used in predicting flight event times and actual flight speed as well as errors in wind predictions. Another contributor to the errors is inaccuracy in modeling flight ascent and descent profiles in TFMS. TFMS uses several profiles customized by airports that usually differ from actual flights' accent and descent profiles. All those factors contribute to uncertainty in flight event predictions and the errors are considered as "internal", inherent TFMS errors that absorb errors in TFMS flight modeling as well as in processing flight tracking messages. When TFMS predicts aggregate demand counts for NAS elements for specific time intervals (15-minute intervals at airports and one-minute interval in sectors) it deterministically aggregates the flights whose estimated times for being in the NAS elements fall within the time intervals. The predicted aggregate demand counts are compared with capacities (or Monitor Alert Parameter, called MAP) to determine whether the NAS element is potentially congested or not. Those aggregate demand count predictions play a crucial role (along with MAP values) in identifying potential congestion and subsequent TFM decision making process on triggering remedial Traffic Management Initiatives (TMIs) to avoid congestion. The question is how accurate are the deterministic aggregate demand predictions? In this paper, we are interested in analyzing the prediction errors inherent in TFMS without considering the sources of errors outside TFMS, such as unexpected flight cancellations or pop up flights.

The errors in traffic demand predictions make the deterministic predictions not only inaccurate but also unstable, with the possibility of significant variation during consecutive demand updates. The instability of demand predictions may lead to instability of TFMS Monitor/Alert. Alerts will flicker on and off when demand exceeds capacity at a specific 15-minute interval and, the successive, updated demand for the same interval is below capacity due to prediction error.

High accuracy and stability of traffic demand predictions should play an especially important role in predicting congestion and its severity when traffic management initiatives, such as ground delay programs (GDP), rerouting or miles in trail (MIT), are contemplated. Inaccurate predictions may be costly because

- Over-predicted demands may lead to excessively conservative strategies by over-controlling traffic with the possibility of unnecessary ground delays. Excessively conservative strategies may lead to the loss of valuable arrival slots during collaborative decision making (CDM) GDP slot allocation procedures.
- Under-predicted demands may lead to excessively aggressive strategies that under-control traffic with the possibility of excessive airborne delays.
- Instability of traffic demand predictions may cause frequent flickering of alerts during consecutive traffic updates (i.e., alerts being turned on and off frequently). This reduces the credibility of the system, and may lead to unnecessary TFM actions.

Currently, there is a significant effort in aviation research community to develop a probabilistic TFM within the NextGen program that would improve the TFM decision making process by bringing it closer to reality. Probabilistic TFM acknowledges the stochastic nature of predictions and incorporates characteristics of uncertainty in predictions into the decision-making process.

Several publications have presented the concepts and potential applications of probabilistic TFM as well as modeling and benefit analysis. A general concept of probabilistic TFM and representation of uncertainty in air traffic demand and capacity predictions for identifying and managing congestion in NAS elements are described in Wanke et al. [1, 2, 3], Mueller et al. [4] and Ramamoorthy et al. [5]. Ball et al. [6] presented a methodology for analyzing effects of uncertainty in traffic demand predictions on Ground Delay Programs. A method for relating the uncertainty in individual flight time predictions to probabilistic characteristics of aggregate demand counts was presented in Meyn [7]. A sequential decision-making approach to probabilistic TFM that makes it possible to update TFM strategies in accordance with updated probabilistic forecasts on demand and capacity was described in Wanke and Greenbaum [8].

Gilbo and Smith [9, 10, 11] proposed a different approach to dealing with uncertainty in aggregate demand predictions and improving the accuracy of the predictions. The approach recognizes the stochastic nature of predictions and, for estimating traffic demand prediction at each 15-minute interval of interest, it uses a linear regression of deterministic predictions for several consecutive 15-minute intervals surrounding the 15-minute interval of interest, both preceding and following ones, along with the deterministic prediction for this interval. As a result, the demand estimations for each 15-minute interval of a time period are represented by a linear regression of a set of deterministic predictions within a sliding time window. For example, in the case when regression includes deterministic predictions at three consecutive 15-minute intervals, the sliding window has the 45-minute width with the interval of interest in the middle of the window, and the regression comprises a linear combination of deterministic demand predictions for an interval of interest (in the middle of the 45-minute window) along with demand predictions for two immediately adjacent intervals, the preceding and the following ones. The major reasoning behind using such a regression model is as follows. An aggregate demand prediction for an interval comprises those predictions for individual flights with ETAs in that interval. Because ETA predictions for flights are uncertain and contain random errors, it is possible (indeed, quite likely) that a flight's predicted ETA will move from one interval to an adjacent interval during flight updates. Therefore, the aggregate count predictions for adjacent intervals may provide useful information for improving the aggregate count prediction for the interval of interest. A prediction model that considers aggregate counts in adjacent intervals thus provides a mechanism for implicitly transferring uncertainty in individual flights' predictions into aggregate demand predictions making those predictions more accurate and more stable than the deterministic predictions. More detailed material can be found in [12] where the authors showed how the use of uncertainty in predicting flights' arrival times resulted in predicting aggregate traffic demand at airports via the weighted averages of deterministic demand predictions in adjacent 15-minute intervals.

The models presented in this paper were calibrated for nine airports and thirteen en route sectors using data from the summer of 2005 and the winter and spring of 2006. The models were then validated using 7 days of data (from the summer of 2005 and the winter of 2006) that were not in the original calibration set. The analysis showed that the regression model provided a reduction in demand prediction errors of nearly the same magnitude as reported in Meyn [7] but the effect was achieved by a simpler prediction algorithm.

Characteristics of accuracy in demand predictions provided by linear regression can be directly applied to probabilistic representation of traffic demand, which is a core part of probabilistic TFM. In particular, the regression provides the expected values of demand and the distribution and/or standard deviation of prediction errors can be used for determining the area of uncertainty around expected values restricted by specific percentiles (e.g., between 25^{th} and 75^{th} percentiles).

The next question is how the improvements in accuracy of demand predictions provided by regression model would improve the quality of Monitor/Alert. This paper presents the results of comparative analysis of various Monitor/Alert characteristics at several US airports and en route sectors under current TFMS deterministic demand predictions and under regression model predictions. The Monitor/Alert characteristics included both stability (rate of flickering in alert status) and accuracy in predicting alert status. It appeared that the regression model provided improvements in Monitor/Alert characteristics such as stability and the number of missed alerts.

The remainder of this paper is in the following sections:

- Accuracy of Current TFMS Traffic Demand Predictions
- Linear Regression Approach for Traffic Demand Predictions at Airports and Sectors
- Model Validation
- Impact on Monitor/Alert
- Conclusions

ACCURACY OF CURRENT TFMS TRAFFIC DEMAND PREDICTIONS

This section describes the results of analysis of the quality of current TFMS predictions of traffic demand in a 15-minute interval. This analysis was performed on selected airports and sectors using historical data from the following 34 days:

- January 1–19, 2006
- April 10–16, 2006 and
- May 5–12, 2006.

The results for the three months were similar. Therefore, this section presents only the January analysis. In 2009, additional data was gathered for the sectors during the week of April 10-16, 2009.

The data was collected and analyzed for nine airports (ATL, BOS, DFW, LAX, MCI, MIA, ORD, SFO, and STL), and thirteen sectors (ZBW02, ZBW17, ZID82, ZID83, ZID86, ZLC06, ZLC16, ZMP20, ZOB57, ZOB67, ZOB77, ZSE14, and ZTL43).

For airports, the data includes the number of arrivals in a 15-minute interval. For sectors, the data includes the peak number of flights within a one-minute bucket of a 15-minute interval. TFMS predicts both of these quantities.

Figure 1 illustrates the distributions of errors in 15-minute traffic demand predictions at ATL airport for various look-ahead times (LAT). In this figure, the distributions are nearly symmetric. The mean errors are close to zero and standard deviations increase with

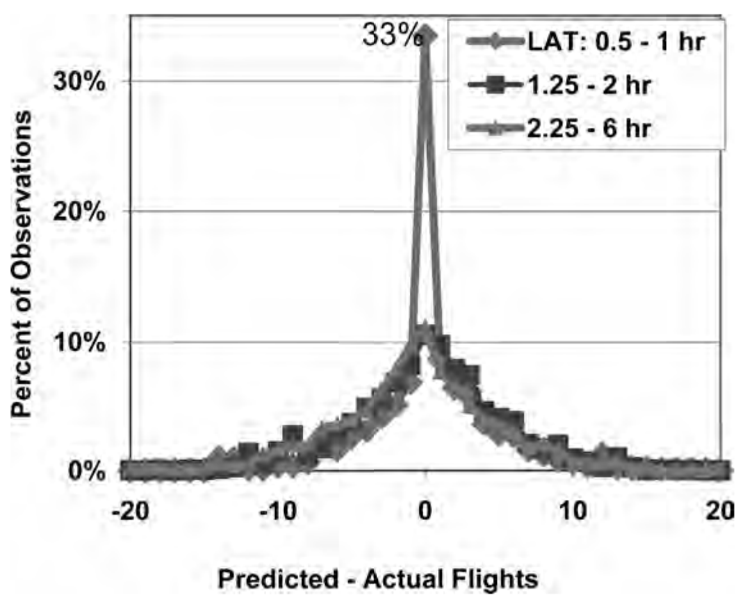

Figure 1. Histogram of Prediction Errors for ATL, January Data.

increasing LAT. The curve for LAT < 1 hour has a standard deviation of 4.5 flights; while the curve for LAT > 2 hours has a standard deviation of 6 flights.

Analysis of accuracy of demand predictions for en route sectors showed that the prediction errors have slightly asymmetric distributions with heavier right-hand tails that indicate that the over-prediction of traffic demand is more likely than under-prediction. Figure 2 illustrates the distributions of demand prediction errors for sector ZBW02 for various LAT. In this example, the average error ranges from 1 to 2 flights, and the standard deviation ranges from 3 to 4.5.

Rather than showing separate histograms of the error for every airport and sector, Tables 1 and 2 display the average and standard

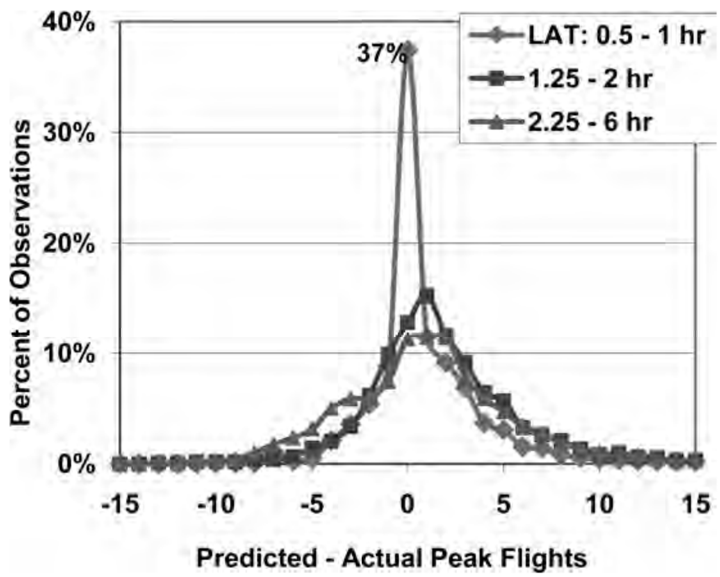

Figure 2. Histogram of Prediction Errors for ZBW02, January Data.

Table 1. Errors in TFMS Demand Predictions for Airports

Airport	Average # of Flights per 15 min	Average Error (Predicted–Actual Flights) LAT: 0.5 – 1 hr	1.25 – 2 hr	2.25 – 6 hr	Standard Deviation of Error LAT: 0.5 – 1 hr	1.25 – 2 hr	2.25 – 6 hr
ORD	18	−1.57	−2.27	−2.59	4.26	4.93	6.32
ATL	16	0.36	0.05	−0.81	5.15	5.26	6.00
DFW	13	−0.27	−0.39	−0.76	2.70	3.15	4.16
LAX	11	0.36	−0.31	−0.65	2.71	3.16	3.79
MIA	7	0.00	0.10	−0.34	2.28	2.66	3.07
BOS	6	−0.25	−0.45	−0.60	2.80	2.93	3.39
SFO	6	0.04	−0.27	−0.49	2.16	2.41	2.96
STL	5	−0.01	−0.06	−0.41	1.66	2.25	2.60
MCI	3	0.10	0.32	0.14	1.32	1.58	1.92

Table 2. Errors in TFMS Demand Predictions for Sectors

Sector	Average peak occupancy in 15 min	Average Error (Predicted – Actual Flights)			Standard Deviation of Error		
		LAT: 0.5 – 1 hr	1.25 – 2 hr	2.25 – 6 hr	LAT: 0.5 – 1 hr	1.25 – 2 hr	2.25 – 6 hr
ZBW02	8	0.83	1.82	1.02	3.45	4.06	4.48
ZBW17	3	1.21	2.23	1.53	2.49	2.39	2.99
ZID82	9	−0.12	−0.20	−1.73	3.34	3.92	4.13
ZID83	7	0.27	0.03	−1.86	3.47	4.20	4.24
ZID86	9	0.08	−0.12	−1.29	3.71	4.37	4.61
ZLC06	9	0.40	−0.02	−2.29	2.85	3.09	3.72
ZLC16	8	0.07	−0.71	−2.81	3.57	4.29	4.57
ZMP20	8	1.12	1.52	−0.30	2.51	3.17	3.84
ZOB57	5	−0.18	−0.66	−0.82	2.67	2.99	3.07
ZOB67	7	0.73	0.65	0.21	3.08	3.41	3.42
ZOB77	7	1.15	2.09	3.08	3.41	4.94	6.41
ZSE14	6	0.03	0.04	−0.66	2.36	3.01	3.07
ZTL43	7	−0.18	−0.51	−2.10	2.56	3.57	3.31

deviation of errors in 15-minute traffic demand predictions for various LAT at selected airports and sectors, respectively.

Table 1 shows that average errors in demand predictions for airports are close to zero (except for ORD) in a wide range of LAT. For ORD, however, TFMS under-predicts demand by approximately 2 flights per 15-minute.

Table 2 shows no clear pattern in average error for sector demand. In most sectors, except ZBW02, ZBW17, ZMP20, and ZOB77, the average errors are close to zero for look-ahead times within two hours, and at some of them the error significantly increased for longer LAT with the tendency of under-prediction (e.g., ZLC06, ZLC16, ZID82 – 86 and ZLT43). Some sectors demonstrated over-prediction between one and three flights per 15-minute (e.g., ZBW02, ZBW17, and ZOB77).

For both airports and sectors, Tables 1 and 2 showed that standard deviations of TFMS demand prediction errors increase with increasing look-ahead times. The busier airports (e.g., ORD, ATL) also have a larger standard deviation.

LINEAR REGRESSION APPROACH FOR TRAFFIC DEMAND PREDICTIONS

Suppose that at time n TFMS made a deterministic prediction F(t, n) of traffic demand for a 15-minute interval that starts at time t > n (i.e., for interval [t, t + 15]). The Look Ahead Time of the prediction is (t − n). The estimation of traffic demand prediction $\hat{A}(t,n)$ for interval t can then be obtained by a regression model that includes the deterministic predictions F(t, n) for this interval and the preceding and following 15-minute intervals. In the case when regression includes predictions F(t − 15, n) and F(t + 15, n) for one interval earlier and one interval later than 15-minute interval of interest t, respectively, the estimated traffic demand $\hat{A}(t,n)$ for interval t predicted at time n is

$$\begin{aligned}\hat{A}(t,n) &= a\,\mathrm{F}(\mathrm{t}-15,\ \mathrm{n}) + b\,\mathrm{F}(\mathrm{t},\ \mathrm{n}) + c\,\mathrm{F}(\mathrm{t}+15,\ \mathrm{n}) + k \\ &= A(t,n) + \varepsilon(t,n),\end{aligned} \tag{1}$$

where
a, b, c, and *k* are regression parameters to be determined,
A(t, n) – correct prediction (with zero error) of number of flights at interval t when prediction was made at time n,
ε(t, n) – a random error term that reflects uncertainty in predictions (currently ignored by TFMS).

With this notation, regression model (1) with $a = c = k = 0$ and $b = 1$ represents the current TFMS demand prediction model:

$$\hat{A}(t,n) = \mathrm{F}(\mathrm{t},\mathrm{n}) = \mathrm{A}(\mathrm{t},\mathrm{n}) + \varepsilon(\mathrm{t},\mathrm{n}).$$

The idea of using a linear regression over deterministic predictions for several consecutive 15-minute intervals to improve the accuracy of 15-minute demand predictions at each interval was first proposed by the authors in 2005 [9].

The motivation for involving predictions for adjacent intervals in the regression model is reduction of uncertainty in traffic demand predictions caused by errors in prediction of flights' estimated time of arrival (ETA) at airports or sectors and, as a result, the potential random migration of ETA for a flight from one 15-minute interval to another during consecutive updates. As a result, the same flight can contribute to aggregate counts in different, most likely adjacent, 15-minute intervals during consecutive demand updates.

The motivation can be best illustrated by Figures 3 a) and b) that present the probability density function (pdf) of the prediction error in Estimated Time of Arrival (ETA) at an airport for an individual flight. For illustration only, the pdf represents a random error with standard deviation of 15 minutes that might correspond to predictions with LAT of more than two hours. On Fig. 3 a), the ETA is within the 15-minute interval t, and the flights should be counted in the aggregate demand for this interval. However, due to random error in ETA predictions, the flight has the probability P_{t} to be counted in a 15-minute interval t, and noticeable probabilities $\mathrm{P}_{\mathrm{t}-15}$ and $\mathrm{P}_{\mathrm{t}+15}$ to be counted in adjacent 15-minute intervals t – 15 and t + 15, respectively (see the corresponded striped areas in Fig. 3 a)).

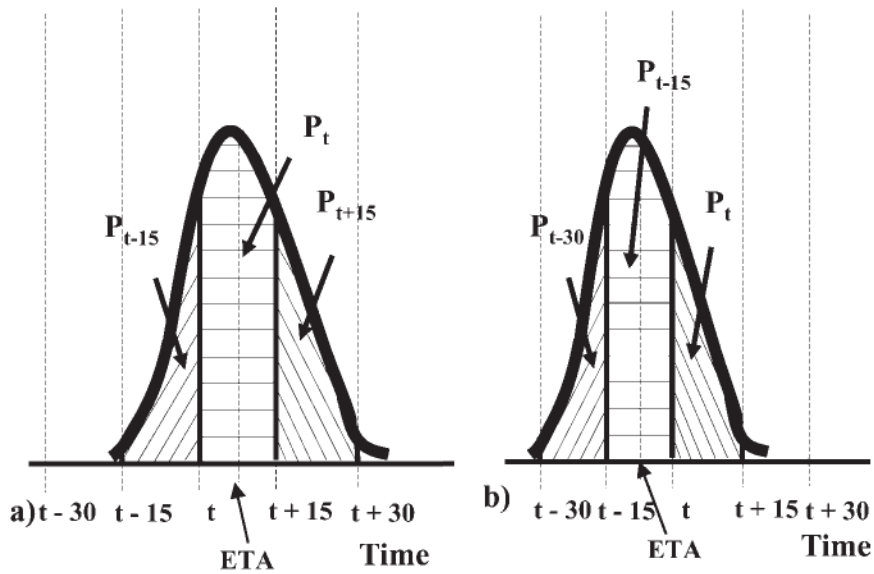

Figure 3. Illustration of probabilities for a flight to arrive at adjacent intervals.

Fig. 3 b) illustrates the case when the flight's ETA falls within a t – 15 interval and, according to deterministic demand prediction, this flight should contribute to aggregate demand for t – 15 interval. However, there are noticeable probabilities P_t and P_{t-30} that the flight would contribute to the adjacent intervals t and t – 30, respectively. That is why one could expect that considering deterministic demand count predictions at the preceding and following adjacent intervals, along with the demand count for the interval of interest, and including them in the regression model, might take into account uncertainty in aggregate demand predictions in current TFMS and improve the accuracy of demand predictions.

Figure 4 illustrates an example of possible migration of ETAs from one 15-minute interval to another due to errors in ETA predictions. At 1200, a demand prediction was made for a 45-minute period from 1500 to 1545. Nine flights were predicted with an ETA for each flight: two flights predicted for the first 15-minute interval (the ETAs of two flights, Flight 111 and Flight 112, are within (1500 – 1514) interval), four flights for the second 15-minute interval and three flights for the third interval. Five minutes later, at 1205, the predictions were updated, and the ETAs for some flights appeared in another 15-minute interval (see the thick lines in Figure 4). In particular, the ETA of Flight 112 moved from (1500 – 1514) interval to the adjacent (1515 – 1529) interval; Flight 113's ETA moved to an earlier, (1500 – 1514) interval; Flight 116's ETA moved from (1515 – 1529) to (1530 – 1544) interval. As a result, the aggregate demand predictions changed from 2, 4 and 3 to 2, 3 and 4 flights for each 15-minute interval between 1500 and 1545.

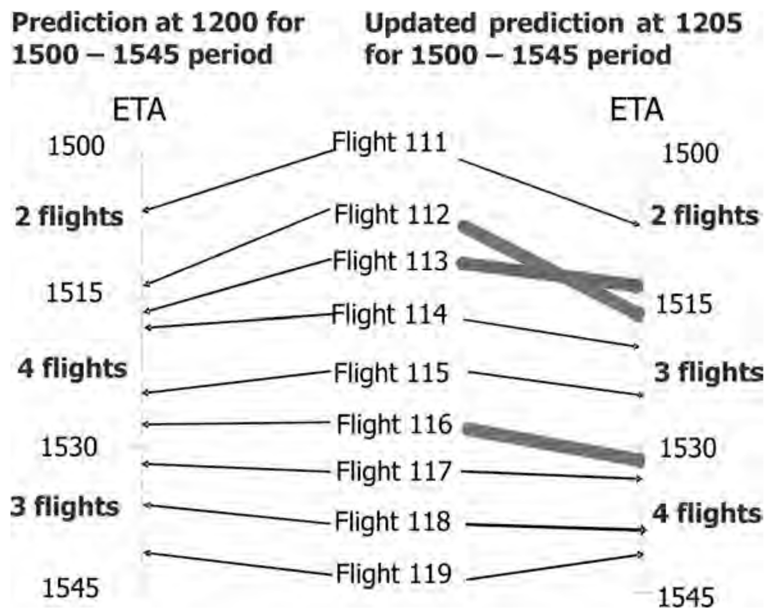

Figure 4. Two consecutive updates of traffic demand predictions.

To generalize, if the random errors in predictions of time of flight arrival at a NAS element can put the flight to an earlier or later 15-minute interval during consecutive flight updates, it is possible that the predicted demand counts for those adjacent intervals could provide useful information for improving predictions for the 15-minute interval of interest.

So far, we discussed the reasoning for using deterministic predictions for several consecutive 15-minute intervals in a linear regression model for improving demand predictions by taking into account uncertainty that accompanies current TFMS deterministic demand predictions. There are other factors, however, that could be considered in the regression model with subsequent analysis to estimate their impact on demand prediction properties, such as

- Look Ahead Time (LAT) that affects the accuracy of deterministic predictions, and, hence, the accuracy of regression model
- The width of a sliding window and its location relative to the 15-minute interval of interest that determines the number of preceding and following intervals for which deterministic predictions are included in the regression model
- Active (airborne or landed flights) and proposed (flights that have not yet departed from the origin airport) components of deterministic TFMS predictions of traffic demand

The following regression models were analyzed:

Model 0: $\hat{A}(t,n) = \mathrm{F(t, n)}$ (current TFMS deterministic predictions)

Model 1: $\hat{A}(t,n) = a$ F(t − 15, n) + b F(t, n) + c F(t+15, n) + k (regression includes TFMS predictions for (t − 15) and (t + 15) adjacent intervals)

Model 2: $\hat{A}(t,n) = g$ F(t − 30, n) + a F(t − 15, n) + b F(t, n) + c F(t + 15, n) + h F(t + 30, n) + k (regression includes TFMS predictions for two adjacent intervals on each side of interval of interest)

Models 1 and 2 were considered to estimate the effects of including deterministic predictions in adjacent intervals in the regression model as well as impact of predictions in more distant intervals (Model 2) on accuracy of demand predictions. The above regression models were analyzed from the following viewpoints:

- Determining parameters (coefficients) of regressions for airports and sectors
- Evaluating impact of various variables in the regression
- Comparing demand prediction accuracy provided by various regression models
- Determining benefits that regression models provide relative to current TFMS predictions.

The models were calibrated on TFMS datasets from

- June 21–25 and July 13–17, 2005
- January 1–19, 2006
- April 10–16, 2006 and May 5–12, 2006

The calibration was performed separately on

– integrated datasets that combined the data for all airports and sectors, respectively, so that regression coefficients were not airport- or sector- specific;
– datasets for individual airports and sectors so that the regression coefficients could vary by airport and sector.

Table 3 presents a summary of results of analysis of regression models 0, 1, and 2 performed on combined datasets of TFMS data for all nine airports and thirteen sectors, respectively, for two ranges of look-ahead times: LAT < 1hr and LAT between 1 and 2 hrs. The table includes regression coefficients, correlation coefficients R^2, average and standard deviation of demand prediction errors as well as percent reduction of standard deviation provided by regression models relative to TFMS demand predictions.

Table 3 shows that, among the weight coefficients a, b and c in both Model 1 and Model 2, coefficient b is the largest one, and coefficient a is greater than coefficient c. Recall that coefficient b determines the weight of deterministic prediction for a 15-minute interval of interest in the regression models, while coefficients a and c determine the

Table 3. Comparison of Regression Models Based on Combined Datasets

NAS Element	LAT	Model	Coefficients	R^2	Standard deviation σ (flights per 15-min)	Percent reduction of σ relative to TFMS
Airport	< 1 hr	0	Current TFMS deterministic prediction	0.81	3.45	
		1	$k = -0.5, a = 0.27, b = 0.61, c = 0.21$	0.87	2.83	18%
		2	$k = -0.64, g = 0.15, a = 0.21, b = 0.55, c = 0.14, h = 0.05$	0.88	2.75	20%
	$1 - 2$ hr	0	Current TFMS deterministic prediction	0.78	3.75	
		1	$k = -0.2, a = 0.31, b = 0.54, c = 0.20$	0.85	3.12	17%
		2	$k = -0.26, g = 0.12, a = 0.24, b = 0.49, c = 0.14, h = 0.08$	0.86	3.04	19%
Sector	< 1 hr	0	Current TFMS deterministic prediction	0.61	3.75	
		1	$k = 1.64, a = 0.25, b = 0.48, c = 0.01$	0.67	3.43	9%
		2	$k = 1.26, g = 0.07, a = 0.23, b = 0.5, c = 0.09, h = -0.006$	0.69	3.33	11%
	$1 - 2$ hr	0	Current TFMS deterministic prediction	0.49	4.32	
		1	$k = 2.94, a = 0.18, b = 0.42, c = 0.14$	0.59	3.86	11%
		2	$k = 2.91, g = 0.05, a = 0.14, b = 0.42, c = 0.16, h = -0.03$	0.6	3.86	11%

weights for deterministic predictions in immediate earlier and later adjacent intervals, respectively. As for coefficients g and h in Model 2 that determine the contributions of deterministic predictions for more distant 15-minute intervals into the regression, they are much smaller than coefficients a and c, respectively. It means that deterministic predictions for more distant intervals relative to the interval of interest have much smaller weights in the regression model, and, hence, make a very small contribution in the predicted demand.

Table 3 also demonstrates that regression models that include deterministic predictions for adjacent 15-minute intervals provide noticeable improvements in accuracy of demand predictions in comparison with current TFMS deterministic predictions. The major improvement provides the regression Model 1 that includes TFMS predictions for two immediately adjacent 15-minute intervals. For airports, the Model 1 reduces standard deviation of prediction error of current TFMS by 18%. For sectors, the reduction is smaller (approximately 10%).

Another observation from Table 3 shows that including additional predictions for more distant intervals in the regression (Model 2) does not contribute significantly to improving demand prediction accuracy in comparison with Model 1 for both airports and sectors.

To examine the possible impact of separating active and proposed flights in total demand count predictions, a third regression model (Model 3) was considered, which included active and proposed components of TFMS demand predictions within 45-minute sliding window:

$$\text{Model 3: } \hat{A}(t,n) = a_1\,\mathrm{Active}(\mathrm{t}-15,\,\mathrm{n}) + a_2\,\mathrm{Proposed}(\mathrm{t}-15,\,\mathrm{n}) + b_1\,\mathrm{Active}(\mathrm{t},\mathrm{n}) + b_2\,\mathrm{Proposed}(\mathrm{t},\mathrm{n}) + c_1\,\mathrm{Active}(\mathrm{t}+15,\,\mathrm{n}) + c_2\,\mathrm{Proposed}(\mathrm{t}+15,\,\mathrm{n}) + k$$

where Active (t, n) and Proposed (t, n) are active and proposed components of TFMS demand prediction F(t, n) for 15-minute interval t made at time n, respectively: F(t, n) = Active (t, n) + Proposed (t, n).

Analysis showed that Model 3 did not improve demand prediction accuracy in comparison with Model 1.

Including predictions for adjacent intervals in the regression models resulted in improving demand prediction quality for both airports and sectors in terms of correlation coefficient R^2. The regression models tended to have better correlations R^2 for the airports than for the sectors. Airport models had R^2 in the 0.8 to 0.9 range, while the sector models had R^2 in the 0.6 to 0.7 range. It is worth noting that the current TFMS predictions provide a high correlation for airports and lower correlation for sectors (see Table 3). Lower values of R^2 for the regression model for sectors can be explained by the fact that the measures for sector and airport demand are different. Sector demand considers maximum one-minute counts applied to an entire 15-minute interval. Therefore, sector demands in adjacent 15-minute intervals are actually the maximum one-minute counts in adjacent intervals and, hence those relatively distant in time one-minute peaks might make much weaker contribution to the regression model than the aggregate 15-minute counts in adjacent intervals at airports.

More detailed analysis of above regression models confirmed that the regression Model 1 provided the major improvements in TFMS traffic demand predictions and more complex models, Model 2 and Model 3, did not add significant additional improvement. Therefore the rest of the paper is concentrated on more detailed analysis of Model 1

$$\hat{A}(t,n) = a\,\mathrm{F}(\mathrm{t}-15,\mathrm{n}) + b\,\mathrm{F}(\mathrm{t},\mathrm{n}) + c\,\mathrm{F}(\mathrm{t}+15,\mathrm{n}) + k$$

and its application for improving traffic demand prediction accuracy and monitor/alert function.

REGRESSION MODEL FOR AIRPORT DEMAND PREDICTIONS

In this section, we will analyze coefficients of the model for both non-airport-specific and airport-specific cases. In each case, the accuracy of the predictions will be analyzed.

Table 4 shows the coefficients of the regression model for various look-ahead times estimated on the datasets that integrated the historical TFMS demand prediction data for all nine airports considered in this study (non-airport-specific case)

Table 5 shows the coefficients of regression model estimated for each airport on the TFMS prediction data for look-ahead times between one and two hours (airport-specific case).

Both tables show (similar to Table 3) that, among the regression weight coefficients a, b, and c, coefficient b is consistently the largest and coefficient c is the smallest. This means that the predicted demand counts at the 15-minute interval of interest (F(t, n)

Table 4. Non-Airport-Specific Model: Regression Coefficients for Various Look-ahead Times

	Look-ahead Time (hrs.)				
Regression Coefficients	<1	$1-2$	$2-3$	$3-4$	$4-5$
a	0.25	0.30	0.31	0.29	0.28
b	0.60	0.54	0.47	0.45	0.41
c	0.22	0.20	0.25	0.28	0.28
k	−0.30	−0.10	0.30	0.68	1.17

Table 5. Airport-Specific Regression Coefficients

	Airports								
Regression Coefficients	ATL	BOS	DFW	LAX	MCI	MIA	ORD	SFO	STL
a	0.24	0.19	0.19	0.30	0.13	0.16	0.37	0.19	0.18
b	0.42	0.46	0.67	0.59	0.61	0.59	0.48	0.62	0.57
c	0.18	0.17	0.18	0.10	0.13	0.20	0.18	0.14	0.22
k	1.79	1.29	−0.43	0.45	0.02	0.26	1.15	0.28	−.008

component) have the highest weight (coefficient b) in the regression model, while the demand counts at the preceding and the following 15-minute intervals (F(t − 15, n) and F(t + 15, n) components, respectively) have smaller weights. Additionally, coefficient a is consistently greater than coefficient c (in one case, when LAT > 4 hours, they are equal). In other words, the preceding counts F(t − 15, n) contribute to the regression with a higher weight than the following counts F(t + 15, n). The latter reflects the common observation that flights are more often late than early, so that the flights predicted to arrive at (t − 15) interval are more likely to arrival later at interval t than the flights predicted to arrive at (t + 15) interval would arrive earlier at interval t. The regression coefficients tended to change with increasing look-ahead time. In particular, coefficient b becomes smaller for the longer look-ahead times, which might be explained by reduction in accuracy of flight ETAs predictions for longer LATs, while coefficients a and c insignificantly increase, still remaining much smaller than b.

In Table 4, coefficient k (constant term of the regression) varies around 0 while slightly increasing with LAT. It is close to 1 flight per 15-minute for the LAT exceeding 4 hours. For airport-specific regressions (see Table 5), coefficients k are close to 0 at six of nine airports. Table 5 demonstrates the correlation between behavior of coefficients b and k: the smaller b the greater k. Moreover, when b is smaller than 0.5 (in Table 5 it is between 0.42 and 0.48) coefficient k is greater than 1 (between 1.15 and 1.79), but it is reduced and becomes close to 0 when b is greater than 0.5.

Tables 4 and 5 show that regression coefficients a, b and c vary depending on airport and LAT. However, each of them varies within a narrow range.

The regressions with airport-specific coefficients were used for both validation testing and comparison with current TFMS.

Given a limited variability of regression coefficients, the following simplified regression model with the single set of coefficients $a = 0.25$, $b = 0.55$, $c = 0.2$ and $k = 0$ was also used for airport validation testing:

$$\hat{A}(t,n) = 0.25\,\mathrm{F}(\mathrm{t}-15,\mathrm{n}) + 0.55\ \mathrm{F}(\mathrm{t},\mathrm{n}) + 0.20\,\mathrm{F}(\mathrm{t}+15,\mathrm{n}). \quad (2)$$

REGRESSION MODEL FOR SECTOR DEMAND PREDICTIONS

Regression analysis was also performed for sectors, where traffic demand is measured by the peak number of flights within a one-minute bucket of a 15-minute interval. Regression Model 1, which includes demand predictions for two adjacent intervals, was used for sector demand prediction, and its parameters were determined on the historical datasets collected at the same dates as for airports. Table 6 shows the coefficients of this model for various look-ahead times obtained for the datasets with combined data for all sectors considered. Again, coefficients were as expected, with b being the largest, and the value of b decreased as look-ahead times increased. However, the constant term was substantially larger for the sector model than for airports (compare Table 6 with Table 4).

Over the course of the analysis, a number of regression models were considered. They included, along with the number of flights in two immediate adjacent intervals, additional variables such as the numbers of flights in two more distant intervals, active and proposed flights, and the look-ahead time. Those additions did not provide any significant benefits and showed very small improvements in the basic model [9]. Accordingly, a model [10] was developed where some of the coefficients depend on a precise look-ahead time. Table 7 shows the values of these coefficients for several specific look-ahead times estimated from the sets of data collected for all sectors considered

From this table, coefficients b and k can be extrapolated as linear functions of LAT for LAT ≤ 2 hrs, while coefficients a and c remained constant.

Table 6. Sector Model: Coefficients for Various Look-ahead Times

	Look-ahead Time (hrs.)				
Regression Coefficients	< 1	$1-2$	$2-3$	$3-4$	$4-5$
a	0.27	0.21	0.20	0.19	0.20
b	0.44	0.36	0.25	0.23	0.21
c	0.08	0.14	0.14	0.16	0.16
k	1.23	2.54	3.26	3.75	3.89

Table 7. Sector Model Coefficients

	Look-ahead Time		
Regression Coefficients	0	1 hr	2 hrs
a: F(t−15,n)	0.25	0.25	0.25
b: F(t,n)	0.51	0.38	0.25
c: F(t+15,n)	0.14	0.14	0.14
k: Constant	−0.28	1.21	2.69

Expressed in terms of an equation, the sector regression model with coefficients as shown in Table 7 is as follows,:

$$\begin{aligned}\hat{A}(t,n) &= 0.25\,\mathrm{F}(\mathrm{t}-15,\mathrm{n}) + (0.51 - 0.13\,LAT)\,\mathrm{F}(\mathrm{t},\mathrm{n}) \\ &\quad + 0.14\,\mathrm{F}(\mathrm{t}+15,\mathrm{n}) + 1.49\,LAT - 0.28 \\ &= \mathrm{A}(\mathrm{t},\mathrm{n}) + \varepsilon(\mathrm{t},\mathrm{n}),\end{aligned} \tag{3}$$

where LAT is the look-ahead time in hours, and $0 < LAT \leq 2$

MODEL VALIDATION

The accuracy and volatility of predictions from the regression models were tested using 7 days of data that were not in the calibration set. The data included some 27,000 predictions for various airports, time of prediction (n), and event time (t) combinations, and some 33,000 predictions for various sectors, time of prediction (n), and event time (t) combinations. The days included three days in July 2005 (Friday July 22, Saturday July 23, Wednesday July 27) and four days in February 2006 (Tuesday February 14 – Friday February 17). The look-ahead time was between 30 minutes and two hours.

Airport Model

To assess the impact of regression models on accuracy of demand predictions relative to current TFMS predictions, the accuracy analysis was performed for each airport using both airport-specific regressions and a single, non-airport specific regression model from equation (2). Parameters for airport-specific regressions were taken from Table 5.

Table 8 shows the average and standard deviation of demand prediction errors for TFMS (the existing model), the new single regression model and the new airport-specific regression models. In the Table, "Average Error" refers to Predicted minus Actual number of flights per 15-minute, while the "Correlation" is the correlation coefficient between predicted and actual numbers of flights.

Table 8 shows that the TFMS and the single regression model applied to individual airports provide similar average errors in traffic demand predictions: a small fraction of one flight at SFO, MIA, STL, LAX and MCI, close to one flight at ATL DFW and BOS, and more than two flights at ORD. Not surprisingly, the airport specific regression significantly reduced the average errors at many airports: at ORD, the error reduced from –2.63 in TFMS to 0.98, at ATL it reduced from 1.02 in TFMS to 0.52. There was, however, an increase in

Table 8. Accuracy of Demand Predictions by Airport

	Average Error (flights per 15-min)			Standard Deviation of Prediction Error			Correlation		
Airport	TFMS	Single Regression	Airport-Specific Regression	TFMS	Single Regression	Airport-Specific Regression	TFMS	Single Regression	Airport-Specific Regression
ATL	1.02	0.84	−0.52	4.76	3.88	3.75	0.80	0.85	0.85
BOS	−1.03	−1.03	−1.17	2.92	2.58	2.59	0.66	0.70	0.70
DFW	−0.78	−0.86	−0.61	3.52	3.44	3.32	0.85	0.85	0.86
LAX	0.29	0.24	0.48	3.22	2.57	2.57	0.73	0.79	0.80
MCI	0.41	0.41	−0.05	1.69	1.47	1.43	0.65	0.66	0.68
MIA	0.11	0.11	−0.05	2.74	2.42	2.41	0.79	0.82	0.83
ORD	−2.63	−2.64	−0.98	5.83	5.13	5.09	0.70	0.75	0.76
SFO	0.09	0.08	−0.03	2.18	1.93	1.90	0.78	0.81	0.81
STL	0.22	0.23	−0.09	2.17	1.85	1.81	0.69	0.72	0.72
All Airports Combined	−0.26	−0.29	−0.34	3.60	3.17	3.00	0.74	0.77	0.78

average error at LAX from 0.24 to 0.48, but the error remained within a fraction of one flight.

The standard deviations of prediction error (Table 8) show a more noticeable improvement in prediction accuracy, with all airports showing an improvement. Linear regression reduced the standard deviation at each airports ranging from 11% to 21% (except for DFW, where the reduction was smaller). The single regression (2) and the airport-specific regressions are very close in terms of accuracy of traffic demand predictions at airports.

Table 8 shows that the correlation between predicted and actual demand counts varies by airport and in most cases is high. The regression models provide slightly higher correlation than TFMS predictions with 4% – 10% increase in correlation coefficients (except DFW where a high correlation of 0.85 remained unchanged). It is worth noticing that a single regression model (2) provided nearly the same correlation as airport-specific models.

To assess the sensitivity of accuracy of the regression model to its coefficients, several values of the coefficients were applied to the validation data sets. For values of coefficient b ranging from 0.45 to 0.6, coefficient a ranging from 0.2 to 0.35, coefficient **c** ranging from 0.2 to 0.25, and all coefficients adding to one, there was very little change (less than 3%) in the standard deviation.

Sector Model

To assess the impact of regression model on accuracy of sector demand predictions, an analysis of the current TFMS and regression model (equation 3) predictions was performed using historical TFMS data that was not in the calibration data set for the same 13 sectors considered for the model calibration.

Table 9 shows characteristics of demand prediction accuracy at specific sectors. In the table, Average Error refers to Predicted minus Actual number of flights per 15-minute, while the Correlation is the correlation coefficient between predicted and actual numbers of flights.

The regression model reduced the standard deviation of demand prediction errors at all sectors considered (except ZLC06 and ZMP20): the reduction ranged from 5% at ZBW17 to 22% at ZID83 and ZOB77. The results for average errors were mixed but mainly without significant difference in absolute values of errors (except ZMP20 and ZLC16). The regression model did not make any noticeable difference in increasing correlation coefficients for sector demand predictions. For majority of sectors considered, the correlation coefficient is lower than for airports, which can be explained by the differences in airport and sector demand definitions: aggregate number of flights per 15-minute vs. one-minute peak demand within a 15-minute interval.

Table 9. Accuracy of Demand Predictions by Sector

Sectors	Average Error (flights per 15-min)		Standard Deviation of Prediction Error		Correlation	
	TFMS	Regression	TFMS	Regression	TFMS	Regression
ZBW02	1.42	−0.16	4.28	3.69	0.79	0.81
ZBW17	2.00	1.84	2.69	2.55	0.89	0.88
ZID82	−0.22	−1.07	3.52	2.92	0.64	0.63
ZID83	0.10	−0.66	4.00	3.12	0.54	0.52
ZID86	−0.09	−0.69	3.46	2.88	0.56	0.59
ZLC06	0.24	−1.09	3.09	3.10	0.83	0.84
ZLC16	−0.72	−1.79	4.39	4.08	0.61	0.62
ZMP20	2.39	0.48	3.09	3.02	0.88	0.89
ZOB57	−0.79	−0.64	2.82	2.60	0.58	0.62
ZOB67	0.81	0.02	3.52	3.10	0.59	0.61
ZOB77	1.05	0.41	3.91	3.05	0.58	0.58
ZSE14	0.40	0.17	2.68	2.26	0.62	0.65
ZTL43	−1.19	−1.28	3.44	2.98	0.47	0.47
All Sectors Combined	0.42	−0.34	3.65	3.19	0.66	0.67

IMPACT ON MONITOR / ALERT

For traffic flow managers, unstable predictions can manifest themselves as frequent changes in alert status (flickering) where Monitor/Alert switches on and off as the prediction crosses the capacity threshold. The problem is most evident when demand is near capacity. When demand remains far below capacity, there might be no alert, even though the volatility of the demand predictions is still high. Similarly, if demand is far above capacity, there is always an alert. In this study, in order to avoid predictions being overly influenced by TFM actions, we focused on non-congested days, with demand usually less than capacity (alert thresholds) represented in TFMS.

It is difficult to verify the accuracy of predictions when predicted demand is far above capacity. A simple comparison of predicted demand with actual traffic may be misleading. Since actual traffic, generally speaking, cannot exceed capacity, such a prediction would generally result in a Traffic Management Initiative (TMI) that would bring the actual demand to the capacity level by delaying or rerouting some flights. At the time the prediction was made, it may have been perfectly accurate; however, the subsequent TMI then changed the actual demand. Therefore, in this study the analysis was conducted on the data from low demand days when TMIs are unlikely and the errors in predictions are likely not affected by control decisions of traffic management specialists. In the absence of air traffic control, the inaccuracy in demand predictions is primarily caused by “internal” TFMS errors that accompany the processing of flight data. That is why in the low demand periods it is reasonable to expect that predicted demand is close to what would actually happen. However, the rarity of alerts in low demand periods make, it difficult to analyze the benefits of improved accuracy of demand predictions for monitor/alert improvements. Therefore, this study proposed to reduce capacity thresholds in order to increase the number of alerts.

To create a more meaningful comparison between current TFMS and the new regression model with respect to stability of traffic demand predictions, an artificial alert threshold equal to the average number of flights in a 15-minute interval was created for airports. For sectors, a threshold was created that approximated the average number of flights present in the sector. The artificial thresholds are shown in Table 10 and Table 12. In most cases, this reduced the alert threshold.

By artificially creating alerts during what were, in reality, non-congested conditions, we could analyze alerts during times where TFM actions are ***not*** being taken and predicted demand is likely to be close to what actually would occur subject to "internal" TFMS prediction errors. Several measures were examined:

- The total number of times the demand prediction exceeds the threshold (number of alerts).
- The number of changes in alert status (crossing the threshold) of airport or sector. This is the number of times the demand prediction crosses the capacity threshold during consecutive traffic updates. To minimize flickering of alerts, a lower value is better.
- Number or probability of false alerts (Type I errors) and failures to alert, or missed alerts (Type II errors). False alerts happen when the predicted demand exceeds capacity due to prediction errors while the correctly predicted demand does not. Missed alerts happen when the predicted demand does not exceed capacity while the correctly predicted demand exceeds capacity.

Figure 5 illustrates the concepts. The bar (a) on the left half of Figure 5 represents a true demand prediction below MAP so that there should be no alert. However, the prediction has some uncertainty associated with it, and there is a chance that the prediction might be higher than the MAP; hence the possibility of a false alert. In the right half (b), the true predicted demand is higher than the MAP, so there should be an alert. However, the uncertainty in the prediction might result in a failure to alert, or missed alert.

Note that the smoothing of consecutive demands (e.g., via regression) can cause a change in the total number of alerts. This

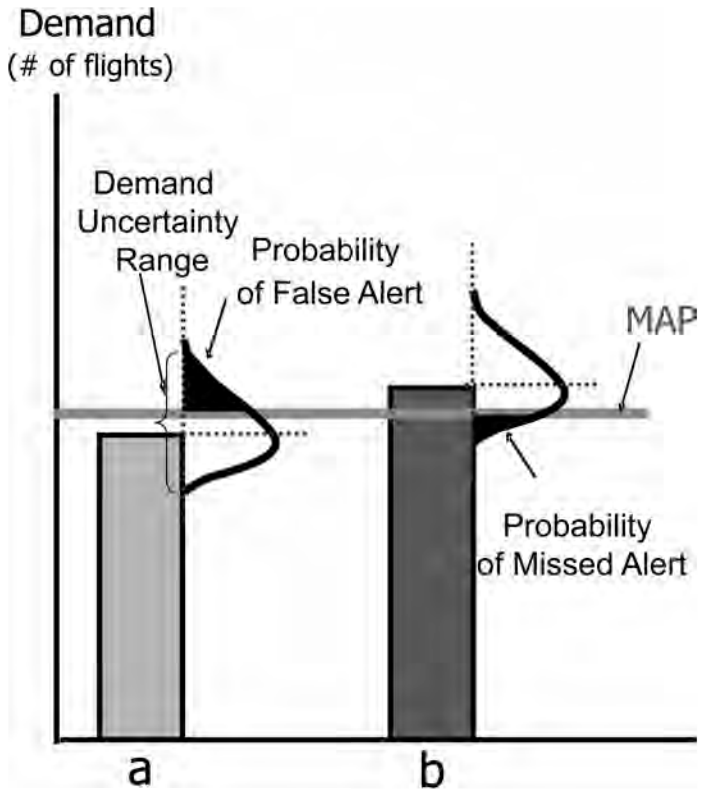

Figure 5. Illustration of probabilities of false and missed alerts.

change can go in either direction depending on the demand patterns. It is reasonable to expect a change in the number of alerts. For instance, when predicted demand exceeds MAP (and TFMS would alert this interval) but demands in adjacent intervals are much lower than MAP, the smoothing process would result in estimated expected demand that does not exceed MAP and there would be no alert. In this case the predicted demand in a single interval was not big enough to make expected demand higher than MAP.

Airport Demand Prediction Stability and Impact on Monitor/Alert

Table 10 shows the number of alerts, crossings of the alert threshold, false and missed alerts at airports. The numbers in the Regression columns correspond to airport-specific regression models.

The "Grand Total" row in Table 10 shows results aggregated for all airports. The last row of the table shows percents of differences in the total numbers relative to TFMS numbers.

Table 10. Monitor/Alert Measures for Airports

		Number of Alerts		Crossing the MAP		False Alerts		Missed Alerts	
Airport	MAP	TFMS	Regression	TFMS	Regression	TFMS	Regression	TFMS	Regression
ATL	16	2045	1967	231	164	301	230	216	223
BOS	6	1763	1948	412	235	157	175	655	488
DFW	13	1676	1924	298	186	218	311	374	219
LAX	11	1720	2011	369	218	385	528	387	239
MCI	3	1445	1110	412	322	493	296	262	400
MIA	7	1328	1339	301	214	338	331	276	258
ORD	18	1569	2052	397	192	220	341	734	372
SFO	6	1606	1688	355	235	321	344	301	242
STL	5	1564	1533	365	264	390	358	362	361
Grand Total		**14716**	**15572**	**3140**	**2030**	**2823**	**2914**	**3567**	**2802**
Pct. Difference Relative to TFMS		**6%**		**−35%**		**3%**		**−21%**	

Table 11. Summarized Monitor / Alert Measures for Airports

	TFMS	Single Regression	Airport-Specific Regression
Number of Alerts	14716	16216	15572
Crossing the Threshold	3146	1850	2030
False Alerts	2823	3374	2914
Missed Alerts	3567	2618	2802

The single-regression model, applied to airports, provided results similar to those shown in Table 10.

Table 11 shows the summarized Monitor/Alert measures for TFMS, the single regression model and the airport-specific regression model.

For these models, there were a substantial number of instances where predicted demand exceeded the artificial threshold. However, regression significantly improved stability in identifying alerts: for demands estimated by regression models, the number of crossings of this threshold has been reduced by an average of 35%. Regression also provided significant reduction in the number of missed alerts in comparison with the TFMS Monitor/Alert: 21% reduction by airport-specific regression, and 27% reduction by single regression model. The airport-specific regressions insignificantly increased the number of false alerts (3% higher than under TFMS) while the single regression gave a higher increase (around 19%). This table illustrates potential benefits of using regression models, especially airport-specific models, for improving Monitor/Alert at airports. It also should be noted that the numbers of false and missed alerts are often in a tradeoff relationship: as one increases the other tends to decrease.

Sector Demand Prediction Stability and Impact on Monitor/Alert

A similar analysis was performed for sectors using the model shown in Equation 3. There were 39,236 observations. The summarized results are shown in Table 12.

The "Grand Total" row in Table 12 shows results aggregated for all sectors. The last row of the table shows percents of differences in the total numbers relative to TFMS numbers.

Here, similar to the airports, the regression model for predicting demand at sectors showed much fewer instances of crossing the alert threshold than current TFMS, which indicates improved stability. The total number of missed alerts decreased by 22%, while the number of false alerts increased by 8%.

Table 12. Monitor/Alert Measures for Sectors

		Number of Alerts		Crossing the MAP		False Alerts		Missed Alerts	
Sector	MAP	TFMS	Regression	TFMS	Regression	TFMS	Regression	TFMS	Regression
ZBW02	8	1974	1845	335	218	268	191	109	161
ZBW17	4	1621	1885	506	341	701	918	85	38
ZID82	7	2007	1953	608	343	139	114	522	551
ZID83	6	2208	2348	607	323	214	253	419	318
ZID86	6	2134	2244	559	301	258	277	448	357
ZLC06	7	2398	2385	343	209	187	138	148	112
ZLC16	7	2196	2178	395	229	200	176	354	348
ZMP20	9	2169	2037	287	209	319	242	60	115
ZOB57	4	1972	2347	586	213	150	221	494	190
ZOB67	6	2369	2459	429	181	363	388	233	168
ZOB77	6	1933	1937	556	318	499	506	360	363
ZSE14	4	2368	2583	388	111	279	391	169	66
ZTL43	4	2262	2549	510	130	103	173	421	204
Grand Total		27611	28750	6109	3126	3680	3988	3822	2991
Pct. Difference Relative to TFMS		**4%**		**–49%**		**8%**		**–22%**	

Discussion

The TFMS deterministic predictions of traffic demand at each 15-minute interval can be considered as a time series, each element of which contains random errors. The linear regression of TFMS deterministic traffic demand predictions for immediately adjacent time intervals performs a weighted average of the time series within a 45-minute moving time window which contains three consecutive 15-minute predictions. The moving weighted average provided by linear regression smoothes the time series of deterministic TFMS predictions and reduces prediction errors in comparison with the current TFMS prediction errors. The question is in what cases the regression approach for traffic demand predictions is most beneficial to the Monitor/Alert function. When predicted demand at a 15-minuite interval is far above or far below the MAP, this particular 15-minute interval will or will not be alerted with or without a moderate reduction of prediction error. However, when traffic demand is in a closer proximity to the MAP, the higher level of prediction errors could cause more frequent flickering of predicted demand around the MAP during periodic updates of demand predictions. Because of the flickering, alert status of airport or sector would change from on to off and vice versa. Reducing prediction errors in these cases would increase stability in alert identification. Additionally, smoothing of the TFMS traffic demand time series by linear regression would result in substantially fewer crossings of the alert threshold for both airports and sectors, providing the TFM specialists with more stable and more reliable information about potential alerts.

The impact of using regression models on the number of false and missed alerts depends on the severity of congestion, namely, the magnitude and duration that the predicted demand exceeds the MAP. In cases of sustained high demand, significantly exceeding the MAP over a period of time, both current TFMS and regression model predictions will provide small numbers of false and missed alerts. The relationship between the number of false and missed alerts changes as traffic demand becomes closer to the MAP. When demand only slightly (e.g., by 1 or 2 flights) exceeds the MAP in a few time intervals, TFMS will alert the entire 15-minute interval in a sector regardless of the magnitude of a single minute overload. The regression model smoothes those TFMS demands and will likely bring them below the MAP. This could lead to an increase in the number of missed alerts in comparison with TFMS. In a situation like this, the TFM specialist might ignore the TFMS alert, recognizing that it is not significant. The most positive impact of demand predictions via linear regression on Monitor/Alert reliability (in terms of missed and false alerts) occurs in the cases of moderate congestion, when predicted demands are frequently higher than the MAP. In these cases, the number of missed alerts can be substantially reduced in comparison with the current TFMS Monitor/Alert. This reduction might be accompanied by a smaller increase in the number of false alerts, as missed and false alerts are generally in a tradeoff relationship. Tables 10 and 12 illustrate this effect.

CONCLUSIONS

TFMS currently makes its aggregate traffic demand predictions based on deterministic projections of traffic and neglects random errors in predictions. The TFMS predictions can be improved by including in the calculation a factor for uncertainty. The new prediction models that were proposed and analyzed in this study took into account uncertainty in traffic demand predictions and showed improvements in both accuracy and stability of demand predictions compared with current TFMS.

The research was focused on analysis and characterization of errors in aggregate demand predictions inherent in TFMS and on improvements of TFMS predictions by using a new prediction approach based on linear regression that includes, along with deterministic predictions for the 15-minute interval of interest, the predictions for two immediate adjacent 15-minute intervals $(t - 15)$ and $(t + 15)$. Including demand predictions for adjacent intervals would take into account possible random migration of some flights from one 15-minute interval to another during consecutive demand updates due to errors in predictions of flight arrival times. The proposed new models were used for predicting demands at airports and sectors.

The results of the study can be summarized as follows:

- Statistical analysis of TFMS historical data provided a characterization of the uncertainty in current TFMS aggregate demand predictions. Average errors and standard deviation of prediction errors were estimated at nine airports and thirteen en route sectors for various look-ahead times (LAT) ranging from 30 minutes to 6 hours.
- Linear regression was proposed and used as a new prediction model that improves accuracy of aggregate demand predictions.
- The parameters of regression models were calibrated on historical TFMS demand data for selected airports and sectors for various look-ahead times.
- The regression models were tested on data not in the original calibration set for look ahead times ranging from 30 minutes to 2 hours. It showed an improvement in accuracy of demand prediction in comparison with the current TFMS predictions.

  - For airports, the reduction in standard deviation over the current TFMS predictions averaged 13 percent, and ranged between 2 and 21 percent, depending on the airport.
  - For en route sectors, the reduction in standard deviation over the current TFMS predictions averaged 12 percent, and ranged between 0 and 22 percent, depending on the sector.
- Improvements in accuracy of demand predictions by the new regression models provide significant benefits for the stability of the Monitor/Alert function mainly in cases when predicted demand counts are in the vicinity of airport or sector capacity, when the instability and high fluctuations of successive demand predictions may therefore cause significant instability and fluctuations (flickering) in the display of alert status. When capacity was set equal to average demand, the number of changes in alert status decreased by 35 to 49%, depending on the airport or sector.
- For airports and sectors, the new model significantly reduced the number of missed alerts, at the cost of a smaller increase in false alerts.
- The material presented in the paper has a direct connection with probabilistic representation of traffic demand: the expected demand provided by regression and dispersion around it can be easily translated to an uncertainty band around expected demand that covers the area where demand values could be predicted with a certain probability.

ACKNOWLEDGMENTS

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

The authors would like to thank all three anonymous reviewers for their extremely valuable and insightful comments and suggestions that helped significantly improve the paper.

ACRONYMS

CDM	Collaborative Decision Making
ETA	Estimated Time of Arrival
ETD	Estimated Time of Departure
GDP	Ground Delay Program
LAT	Look-Ahead Time
MAP	Monitor / Alert Parameter
MIT	Miles in Trail
NAS	National Airspace System
NextGen	Next Generation Air Transportation System

TFM Traffic Flow Management
TFMS Traffic Flow Management System

REFERENCES

[1] Wanke, Craig, 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.
[2] Wanke, C., Mulgund, S., Greenbaum, D., and Song, L., “Modeling Traffic Prediction Uncertainty for Traffic Flow Management Decision Support,”, AIAA Guidance, Navigation, and Control Conference and Exhibit, Providence, RI, 16–19 August, 2004.
[3] 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, USA, June 27–30, 2005.
[4] Mueller, K. Tysen, John A. Sorensen, and George J. Couluris, “Strategic Aircraft Trajectory Prediction Uncertainty and Statistical Sector Traffic Load Modeling,” Paper number AIAA 2002-4765, AIAA Guidance, Navigation, and Control Conference and Exhibit, Monterey, CA, 5–8 August 2002.
[5] 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.
[6] Ball, Michael, Thomas Vossen, and Robert Hoffman, “Analysis of Demand Uncertainty Effects in Ground Delay Programs,” 4th USA/Europe Air Traffic Management R&D Seminar, Santa Fe, NM, USA, 4–7 December 2001.
[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] Wanke, Craig and Daniel Greenbaum, “Incremental, Probabilistic Decision Making for En Route Traffic Management,” Air Traffic Control Quarterly, vol. 15, no.4, pp. 299–319, 2007.
[9] 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.
[10] 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.
[11] 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.
[12] 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.

BIOGRAPHIES

Eugene Gilbo is a senior operations research analyst of the Volpe National Transportation Systems Center, Cambridge, MA. For more than 20 years, he has been conducting research for the FAA 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 M.S, Ph.D. and Doctor of Sciences degrees in Mechanical Engineering, Electrical Engineering and Applied Mathematics, respectively, from Leningrad Polytechnic Institute.

Scott 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 an 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.