Modeling Weather Impact on Ground Delay Programs

Yao Wang, NASA Ames Research Center, Moffett Field, California

Deepak Kulkarni, NASA Ames Research Center, Moffett Field, California

Abstract

Scheduled arriving aircraft demand may exceed airport arrival capacity when there is abnormal weather at an airport. In such situations, Federal Aviation Administration (FAA) institutes ground-delay programs (GDP) to delay flights before they depart from their originating airports. Efficient GDP planning depends on the accuracy of prediction of airport capacity and demand in the presence of uncertainties in weather forecast. This paper presents a study of the impact of dynamic airport surface weather on GDPs. Using the National Traffic Management Log, effect of weather conditions on the characteristics of GDP events at selected busy airports is investigated. Two machine learning methods are used to generate models that map the airport operational conditions and weather information to issued GDP parameters and results of validation tests are described.

I.   Introduction

Air traffic congestion at the major commercial airports has been a serious problem in the National Airspace System (NAS), especially during bad weather [1]. FAA's Traffic Flow Management (TFM) manages air traffic flow to balance air traffic arrival demand against airport capacity, when the latter is reduced by bad weather or other circumstances. This results in airborne delays by holding some aircraft for landing. At major airports in the United States, when the expected demand for arrival air traffic flow exceeds the Airport Arrival Rate (AAR) for a significant period of time, GDP will be used as one of Traffic Management Initiatives (TMI) to smooth out the arrival flow and bring arrival demand in line with capacity. The most common reason for an overage of demand versus capacity is the reduction in airport acceptance rate due to adverse airport weather such as strong wind, low ceilings and visibility.

Air Traffic Control System Command Center (ATCSCC) implements GDPs after consultation with air traffic centers and airline operation centers to manage the arrival flow by holding some aircrafts at their departure airports for specified periods of time. As a result, airborne delays are reduced and less expensive and less risky ground delays are increased. If a GDP applies to a particular airport, the GDP start time and GDP stop time will be determined by the scheduled demand and forecasted weather profile at the time of the GDP planning. Given the nature of weather forecast uncertainty and the inaccurate weather translation models, the initial and final planned GDP durations are often different from the actual duration. This can result in unnecessary delays.

This paper studies the use of Ensemble Bagging Decision Tree (BDT) and Neural Networks (NN) methods to predict the GDP duration time during bad weather. Our approach is to develop predictive BDT and NN models using historical GDP and weather data, and then apply these models to forecast the GDP duration when GDP is being planned. The prediction outlooks are then discussed.

The remainder of the paper is organized as follows. In the next section, statistical analysis of GDP events at several major airports is provided. Section III details BDT and NN approaches and discusses factors affecting training, performance as well as model validation. Section IV describes the data used in this

study. Then section V presents computational results on the estimation of GDP duration using BDT and NN approach. Finally, section VI provides concluding remarks and future directions.

II.   Statistics of Ground Delay Programs

Traffic managers at the regional and national levels institute and modify TMIs in order to balance traffic demand with system capacity. The FAA developed the National Traffic Management Log (NTML) to provide a single system for automated coordination, logging, and communication of TMIs throughout the National Airspace System. Some aggregate statistics obtained by the examination of the NTML data have been published [2-3] recently. We collected more than 3400 GDP events from NTML data at all US airports over the years 2007 through 2009, and generated GDP event statistics as follows:
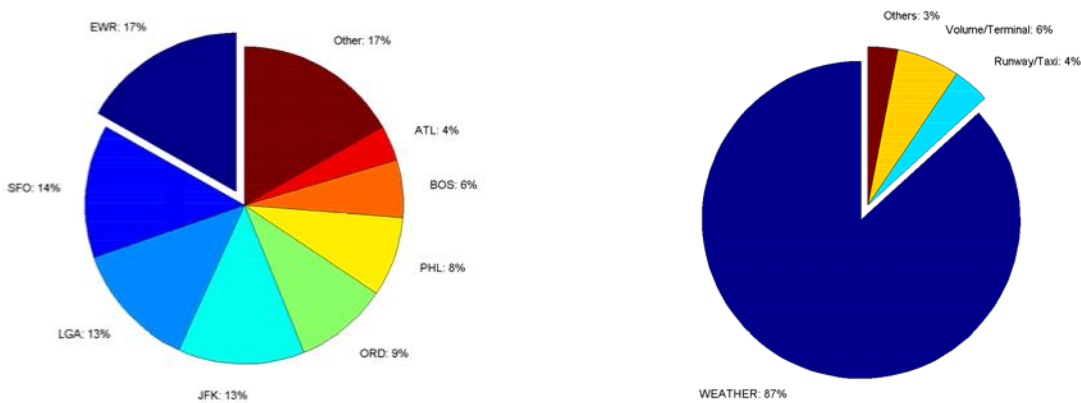

Figure 1: (a) GDP Airport Counts and (b) GDP Causes

Figure 1a displays percentage share of airport GDP counts for the top 8 airports. It shows that most GDPs are implemented at the airports in the northeast region of the United States, including the three New York-area airports (EWR, LGA, and JFK), Philadelphia (PHL), and Boston Logan International Airport (BOS). Figure 1b shows that the major cause of Ground Delay Programs is inclement weather. The diverse weather causes are shown in Figure 2. Details of weather causes for the top 8 airports are provided in Table 1 and 2. Altogether, these data illustrate that the dominant weather causes for GDPs are different at different airports. For example, while close to 90% GDPs at SFO are caused by low ceilings due to marine stratus, wind accounts for about 50% of GDPs at the three New York-area airports, and thunder storms are the major sources of GDPs at ATL.
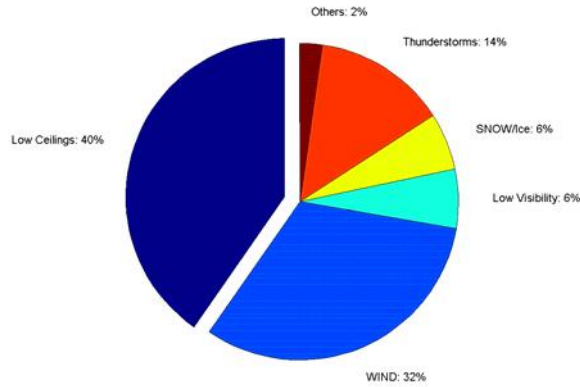
Figure 2: Ratios of the Counts between Weather Subcategories and the Total Weather GDPs

Table 1 Category Percentage Ratio for the Top 8 airports

Airport	Weather	Equipment	Center Volume	Terminal Volume	Runway Taxi	Others
EWR	92%			4%	3%	1%
SFO	96%				3%	1%
LGA	88%	1%		9%	2%	
JFK	78%		1%	17%	2%	2%
ORD	98%	2%				
PHL	91%			1%	8%	
BOS	95%	2%			2%	1%
ATL	96%	4%				

Table 2 Weather Cause Percentage Ratio for the Top 8 Airports

Airport	Wind	Low Ceilings	Low Visibility	Rain	Fog	Snow/Ice	Thunder Storms
EWR	52%	27%	9%	1%		3%	7%
SFO	8%	88%	3%		1%		
LGA	51%	26%	5%	1%		3%	13%
JFK	50%	29%	4%	1%		3%	14%
ORD	29%	25%	8%	6%		14%	15%
PHL	17%	57%	4%	1%	1%	6%	14%
BOS	15%	58%	8%	2%	2%	6%	9%
ATL	5%	37%	9%	1%		3%	45%

The GDP start time, end time, scale, and planned AAR (PAAR) are defined by TFM when a GDP is issued initially. The initial planned duration is the difference between the planned end time and start time at GDP issued time. During a GDP, these program parameters might need to be revised because of changing forecast and operation conditions. GDP revisions may lead to further GDP end time substitutions and the final planned duration is the time duration between the final revised GDP end time and start time. GDP actually ends either

when the final planned GDP end time is reached or when it is cancelled. Table 3 shows average initial planned duration, average final planned duration and average actual duration for the top eight airports. The time durations for diverse weather causes are shown in Table 4. For most of these airports, the differences between averages of Final and Initial Planned Durations are much less than an hour with the exceptions of PHL and SFO, where the differences are about 1 hour. The average of the actual time duration is less than the averages of both initial planned and final planned durations. Such conservative planning of GDP events results in unnecessary delays for many GDP events. For the three New York-area airports where wind is a prominent cause of GDPs, the average time differences between the average final planned durations and the average actual durations are within one to two hours; for other airports, the time differences are more than two hours.

Table 3 Weather GDP Time Durations for the Top 8 Airports

Airport	*Average Initial Planned Duration*	*Average Final Planned Duration*	*Average Actual Duration*
EWR	9:52	10:23	8:53
SFO	6:05	6:57	4:29
LGA	12:04	12:23	10:33
JFK	7:21	7:34	5:57
ORD	10:04	10:32	7:56
PHL	9:11	10:15	7:41
BOS	8:25	8:50	6:22
ATL	8:01	8:34	6:13

Table 4 Weather GDP Time durations of Diverse Causes for the Top 8 Airports

Weather Subcategory Causes	*Average Initial Planned Duration*	*Average Final Planned Duration*	*Average Actual Duration*
Wind	9:25	9:45	8:00
Low Ceilings	8:34	9:18	7:09
Low Visibility	9:06	9:43	7:28
Rain	8:21	8:58	7:24
Fog	6:57	8:12	8:25
Snow or Ice	11:06	11:51	7:41
Thunder Storms	8:32	8:53	6:42

As shown in Tables 3 and 4, these GDP programs usually last for several hours and the average durations depend on weather conditions as well as the airports. Table 5 lists the average total EDCT (Expect Departure Clearance Time) delay minutes (in 1000 minutes) and the total numbers of aircraft assigned with EDCT delays for the top 8 airports due to different weather causes. As shown in Table 5, GDPs generated more delays for ORD and ATL airports than for other airports. At any one airport, the delays can be quite different due to different weather causes.

Table 5 Average GDP Delay (1000 Minutes/Aircraft Numbers) due to Different Weather Causes

Airport	*Wind*	*Low Ceilings*	*Low Visibility*	*Rain*	*Fog*	*Snow/Ice*	*Thunder Storms*
EWR	14/240	20/258	14/218	16/226		16/199	18/212
SFO	14/223	4.5/95	5.0/118		12/34		
LGA	12/288	17/325	15/292	9.2/302		12/192	11/162
JFK	5.3/119	9.6/171	8.8/163	6.7/144		7.8/132	11/152
ORD	20/425	31/517	21/373	18/387		26/431	27/412
PHL	8.8/193	14/248	13/197	13/249	13/264	13/199	9.4/156
BOS	7.0/133	7.9/160	6.1/127	4.3/108	8.0/142	6.6/101	9.3/127
ATL	30/491	31/577	16/375	47/796		11/130	25/331

Reducing the number of GDP revisions and the time difference between GDP initial planned duration and actual GDP duration will reduce air traffic management work load as well as unnecessary aircraft delays. With more accurate air traffic demand prediction, less uncertainty in weather forecast, and better airport capacity models, the initial issued GDP can be improved [4-5]. In this paper, we present machine learning methods to reduce the number of GDP revisions and to improve initial planned GDP duration in order to support TFM.

III.   Approach and Modeling Methodology

We propose to use two machine learning techniques to predict characteristics of a GDP. Ensemble learning with Bagging Decision Tree (BDT) model was used here to classify the difference between GDP initial planned duration and final planned duration. Furthermore, Neural Networks (NN) regression model was applied to predict GDP duration. Supervised machine learning was used to train models, i.e., BDT and NN, and the models were validated using data cross validation methods.

Ensemble Bagging Decision Tree:
Ensemble methods use multiple machine learning models to obtain better predictive performance than what any of its individual constituent members can produce. Bagging is an ensemble method that uses random resampling of a dataset to construct models [6]. In classification scenarios, the random resampling procedure in bagging induces some classification margin over the dataset. Additionally, when bagging is performed in different feature subspaces, resulting classification margins are likely to be diverse, which is essential for an ensemble to be accurate. This method takes into account the diversity of classification margins in feature subspaces to improve the performance of bagging. First, it studies the average error rate of bagging, converts the task into an optimization problem for determining some weights for feature subspaces. Then, it assigns the weights to the subspaces via a randomized technique in classifier construction. Experimental results demonstrate that the ensemble method is robust to classification noise and often generates improved predictions than any single classifier [7-8].

Neural Networks:
A feed-forward neural network consists of input, hidden and output layers and provides a general framework for representing non-linear functional mapping between a set of input variables and a set of output variables. A feed-forward neural network with 8 input neurons, a hidden layer with 5 neurons and a single output neuron is used in this paper. The output from each layer is connected to the next layer by modifiable

weights represented by links between the layers. The weighted outputs from one layer will go through nonlinear sigmoid functions to form the input to the neuron in the next layer. A bias unit is connected to all neurons except the neurons in the input layer. The back-propagation algorithm based on minimizing the output error using a gradient descent method is used for training neural networks [9].

For a NN to have good generalization properties and to avoid over-fitting, the training data should have 5 to 10 times training cases as the weights in NN and it should be statistically representative [10]. A common approach to reduce the number of weights or the complexity of the NN is to reduce the redundant information in a large number of inputs. The input reduction is achieved by Principal Component Analysis (PCA) in this paper. PCA is a technique whose goal is to reduce the dimensionality of data, but retain most of the significant variability in the data for further analysis [11]. PCA creates a summary of attributes that are weighted averages of original attributes, and are uncorrelated to each other. Most of the variability in the data is concentrated in the first a few components. With PCA analysis, we reduce the number of inputs from 48 variables of airport operational conditions and airport weather, METAR and forecast WITI values to a linear combination consisting of 8 inputs. The number of weights in the NN is much smaller with 8 inputs. The test results were found to be stable regardless of when the training was stopped.

Model Validation Methods:
Machine learning models are data driven and therefore resist analytical or theoretical validation. The models are constructed from an initial random state to a trained state using training data sets and must be tested or validated using a different data set. Several validation approaches are available. Among them, the very popular one which has been used frequently by researchers is cross-validation.

In cross-validation, a series of BDT or NN models are constructed, each time by dropping a different part of the data from the training set and applying the resulting model to the dropped data to predict the target. The merged series of predictions for dropped or test data are checked for accuracy against the observations. In one version of the cross-validation approach, called group cross-validation approach, data are divided into N groups. A total of N models are then constructed one by one using N-1 data groups for model training, and the remained group is used for testing. Normally, N can be chosen as 3, 5, or 10. "Leave-one-out" cross-validation is an extreme case of the group cross-validation procedure where N equals to the number of data points. At the end of this procedure, N predictions assembled from the dropped cases are compared with the observed targets to compute validation of model error for the cross-validation result. This cross-validation is used here in this paper.

The Spearman's rank correlation coefficient is used to compare the dependence between two variables. A correlation greater than .8 is generally considered as strong whereas a correlation of less than .5 is generally treated as weak.

IV.   Data Used in the Study

This section describes the GDP data, weather data, and air traffic data that were used in this analysis. The data sources are National Traffic Management Log (NTML) [12], Airport surface Terminal Forecast Weather Impacted Traffic Index, T-WITI-FA [13-15], and the Aviation System Performance Metrics (ASPM) database [16]. All data at GDP issue time over years 2008 through 2009 were derived from these data sources.

GDP data
GDP start hours, initial planned GDP durations, final planned durations, and actual durations were calculated from NTML data. GDP start hours and initial planned GDP durations are inputs to the machine learning models.

For classification, we converted the GDP revisions of GDP events into a categorical attribute with values "YES" and "NO" and used them as the targets. For NN regression, the GDP actual durations were selected as targets.

Current Weather Data at GDP Issued Time:

   At each GDP issue time, actual hourly airport surface weather observations (METAR), such as wind, ceiling, visibility, meteorological condition flags and so on, were selected from ASPM database. These data were preprocessed to convert character records to numerical values and filter out the missing ones. The processed METAR data were used as inputs to the machine learning methods.

Forecast Weather Data:

   The forecast airport Terminal Weather Impacted Traffic Index, T-WITI-FA is provided by Alexander Klein from Air Traffic Analysis, Inc. It was computed based on airport Terminal Area Forecast (TAF) and Collaborative Convective Forecast Product (CCFP) data and other air traffic information. Hourly computed data includes 2-hour, 4-hour, and 6-hour forecast WITI data. Each forecast has seven factors. These are en-route convective WITI, local convective WITI, wind WITI, snow WITI, IMC WITI, Volume/ripple effects WITI, and Other WITI factor values. All factors and the sum of seven factors for each forecast were used as inputs for the modeling process. More details of these factors can be described below.

- En-route convective weather WITI: The convective weather impact on an airport's inbound/outbound flows within approximately 500-NM range is used in this study. This component does not affect queuing delay at the airport.

- Local convective weather WITI: It reflects how convective weather in the vicinity (<= 100 NM) or directly over the airport reduces airport's capacity. It may affect queuing delay.

- Wind WITI: Any time there is a wind greater than 20 Kt, or there is precipitation and wind greater than 15 Kt, the corresponding impact is recorded. Airport capacity may decrease, i.e. queuing delays may increase.

- Snow WITI: It also includes freezing rain, ice etc. The corresponding impact is recorded. Airport capacity may decrease, i.e. queuing delays may increase.

- IMC WITI: This term indicates ceiling or visibility that is below airport specific minima, fog or heavy rain. The corresponding FAA capacity benchmarks for IMC are used. Queuing delays may increase.

- Volume plus Ripple Effects WITI: This can be simply due to high volume of traffic demand or in an aftermath of a major weather event when queuing delays linger on (even as the weather has moved out). Additionally, Ripple Effects are recorded in this component. For example, if ORD experiences departure queuing delays, its corresponding destination airports will get some additional arrival queuing delay.

- Other WITI: It includes other minor impacts due to light/moderate rain or drizzle but ceilings/visibility above VFR minima; also unfavorable RWY configuration usually due to light-to-moderate winds (15-20 Kt or even 10 Kt) that prevent optimum-capacity runway configurations from being used.

Air Traffic Data:

   Air traffic data at GDP issue time, such as aircraft count of scheduled and ETMS arrivals are collected from ASPM database. These data are also used as inputs to our models.

V.   Results

This section presents computational results using two different modeling techniques outlined in the earlier sections to identify GDP revision events and to estimate GDP initial duration at GDP issue time. All the models are trained using the 2008-2009 data and tested by "Leave-one-out" cross-validation. The input data for the BDT and NN models are GDP initial information, air traffic, METAR weather data, and weather forecast (T-WITI-FA) at GDP issue time. These inputs do not include other uncertainty factors, such as airport convective weather observations, airline operations and unscheduled demand, which are used in GDP planning.

For prediction of GDP revision, the GDP training data were grouped into two classes: "Yes" for GDP events having the final planned stop time different from the initial planned stop time; "No" for those with the planned final stop time same as the initial stop time. Using the data described in section IV as inputs and the binary indicator responses of GDP revisions as targets, the BDT classification models were trained first. These models were then applied to make predictions on the test data. The proposed GDP revision event by this BDT model can be utilized for a comparative evaluation of the initial planned stop time by TFM at GDP issue time to reduce the possibility of GDP revision later during GDP procedure.

An accurate GDP planning means that the initial duration estimates are close to the actual GDP time durations and also a better correlation exists between the two. The accuracy is especially important in cases where the planned duration is less than the actual duration because airborne delay not only costs more but also requires more work for the traffic flow manager to safely control the flights to avoid any collision.

Using actual GDP durations as targets and the same input dataset as that for developing BDT classifier, the NN regression models were derived and trained. The prediction of initial GDP time duration on the test data was computed for several airports. The correlations between the predicted durations and actual durations were compared with the correlations between the initial planned durations and actual durations. The number of cases where the initial or predicted time durations are less than the actual durations is also examined. A good prediction NN model would result in good correlation between the time durations proposed by the model and the actual time durations. And it would result in reduction in the cases where the initial planned time duration is less than the actual one if the number of events of this type produced by our NN model is less than those in actual operations.

BDT Classification:

The BDT prediction of GDP revision events for EWR airport is listed in Table 6. There are 363 total GDP events with 171 events having revised GDP stop times. The ratio of 0.47 (171/363) was defined as Initial Revision Rate. The accuracy of BDT binary classifier is the proportion of correct results, (96+128)/363 = 0.62 [17]. Out of a total of 160 revision events predicted by BDT models, the number of correctly forecasted is 96. The precision 0.60 (96/160) is the predicted revision event rate by the BDT classifier. For these revision events (160) proposed by models, a review at GDP issue time may help to reduce the number of late revisions, so as to reduce TFM work load and unnecessary delays.

Table 6 Confusion Matrix to Compare BDT Predictions against Actual Revised GDP Events for EWR

Revised GDP events		BDT Prediction		Sum of Actual
		Yes	No	
Actual	Yes	96	75	171
	No	64	128	192
Sum of Predicted		160	203	363

The initial revision rates and BDT classifier performances for SFO, ORD, PHL, and BOS airports are listed in Table 7. It's clear that the prediction of corrected GDP revised event rate (precision) is higher than the actual GDP initial revision rate. This method offers GDP events with higher revision rate to help evaluation of the GDP parameters at GDP issue time.

Table 7 Initial Revision Rate, BDT Accuracy and BDT Precision

Airport	Initial Revision Rate	BDT Accuracy	BDT Precision
SFO	0.32	0.71	0.59
ORD	0.35	0.67	0.55
PHL	0.45	0.58	0.54
BOS	0.36	0.64	0.50

NN Regression:

Figure 3 shows the histograms of two time differences for EWR airport: (a) the time differences between GDP initial durations and actual durations, and (b) the time differences between NN predictions and actual durations. The two distributions are similar; the NN model is not much better than GDP initial parameters assigned by TFM. For EWR, the correlation between initial GDP duration and actual duration, 0.8 as listed in Table 8, is less but close to the correlation between NN predicted duration and actual duration, 0.82. Table 8 also shows that the number of events with the initial planned duration less than actual duration is the same as for the number of events with the predicted duration less than actual duration. For EWR airport, durations predicted by NN model are only slightly better than initial planned GDP time durations; the differences are not statistical significant.

Tables 9-12 show relevant performance metrics for SFO, ORD, PHL, and BOS airports, respectively. The results for these four airports are similar to those for EWR airport.
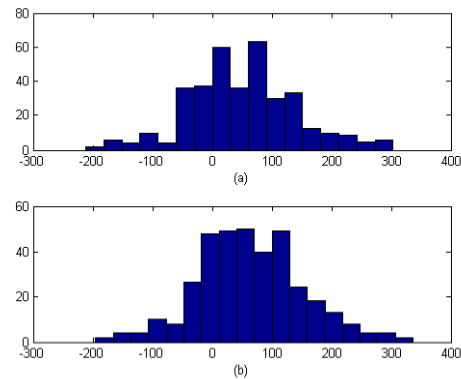

Figure 3: (a) The time difference (minutes) between GDP initial duration and actual duration. (b) The time difference (minutes) between NN GDP duration predictions at GDP issue time and actual GDP duration.

Table 8 Prediction performance of NN model for EWR

GDP Durations	Correlation with actual durations	GDP events with initial or prediction durations less than actual durations
Initial Planned	0.80	83
NN predicted	0.82	83

Table 9 Prediction performance of NN model for SFO

GDP Durations	Correlation with actual durations	GDP events with initial or prediction durations less than actual durations
Initial Planned	0.61	48
NN predicted	0.69	42

Table 10 Prediction performance of NN model for ORD

GDP Durations	Correlation with actual durations	GDP events with initial or prediction durations less than actual durations
Initial Planned	0.71	21
NN predicted	0.74	21

Table 11 Prediction performance of NN model for PHL

GDP Durations	Correlation with actual durations	GDP events with initial or prediction durations less than actual durations
Initial Planned	0.70	31
NN predicted	0.73	27

Table 12 Prediction performance of NN model for BOS

GDP Durations	Correlation with actual durations	GDP events with initial or prediction durations less than actual durations
Initial Planned	0.74	11
NN predicted	0.77	10

VI.   Concluding Remarks

This paper presents machine learning methods for predicting GDP parameter from a set of weather and air traffic data drawn at GDP issue time. The classification method is used to predict GDP Revision events with ensemble Bagging Decision Tree model. The model proposed predictions give better GDP revision event rate than what is currently done. Traffic flow managers can use these predictions of revision events to conduct a further review of their GDP plans, thereby improving initial planning of GDPs.

GDP duration predicted by neural network method has a better correlation with actual duration than the correlation between initially planned GDP duration and the actual duration. In comparison with initially planned duration, NN model results have a smaller or at least same number of events where the actual time duration is underestimated. The differences between GDP initial estimates and models predictions are not statistical significant.

There is a room for improving models described in this paper. For example, the estimates for other uncertainty factors used in GDP plans, such as airport convective weather observations, airline operations and unscheduled demand, at GDP initial issue time would be useful in developing better models. GDP planning can be improved if models that capture the impact of airport surface weather forecast on capacity in detail can be developed.

Acknowledgments

References

1. U.S. Department of Transportation, Federal Aviation Administration, and the MITRE Corporation "*Capacity Needs in the National Airspace System*", June, 2004.
2. Yuditsky, T., and Brickman, B., *"Benefits Analysis for the National Traffic Management Log"*, DOT/FAA/TC-07/24, 2007.
3. Rios J., *"Aggregate Statistics of National Traffic Management Initiatives", 10th AIAA Aviation Technology, Integration, and Operations Conference,* Fort Worth, Texas, Sept. 13-15, 2010.
4. Cook, L., Wood, B., *"A Model for Determining Ground Delay Program Parameters Using a Probabilistic Forecast of Stratus Clearing", USA/Europe Air Traffic Management R&D Seminar*, Napa, CA, June, 2009.
5. Clark, D., *"Investigating A New Ground Delay Program Strategy for Coping with SFO Stratus", 89th AMS Annual Meeting*, Phoenix, AZ, 11-15 January 2009.
6. L. Breiman, "*Bagging Predictors,*" *Machine Learning,* vol. 24, no. 2, pp. 123-140, 1996
7. T.G. Dietterich, "*Ensemble Methods in Machine Learning,*" *Proc. Conf. Multiple Classifier Systems,* pp. 1-15, 2000.
8. P. Melville, N. Shah, L. Mihalkova, and R.J. Mooney, "*Experiments with Ensembles with Missing and Noisy Data,*" *Proc Fifth Int'l Workshop Multiple Classifier Systems,* pp. 293-302, 2004.
9. Foresee, F.D. and Hagan M.T., "*Gauss-Newton approximation to Bayesian regularization,*" Proceedings of the 1997 International Joint Conference on Neural Networks, pp. 1930-1935, 1997.
10. Duda, R.O., Hart, P.E., and Stork, D.G., "*Pattern Classification*", John Wiley and Sons, Inc., New York, 2001.
11. Fukunaga, K., "*Introduction to Statistical Pattern Recognition*", Academic Press, 1990.
12. "*Air Traffic Organization Policy.*" FAA, Feb 11, 2010. Web. July 26, 2011. <http://www.faa.gov/air_traffic/publications/atpubs/FAC/1705.html>.
13. Klein, A., MacPhail, T., Etc. *"NAS Weather Index: Quantifying Impact of Actual and Forecast En-route and Surface Weather on Air Traffic", 89th AMS Annual Meeting*, Phoenix, AZ, 11-15 January 2009
14. Klein, A., Cook, L., Wood, B., and Simenauer, D., "*Airspace Capacity Estimation Using Flows and Weather-Impacted Traffic Index*", Integrated Communications, Navigation and Surveillance Conference, 2008.
15. Klein, A., Kavoussi, S. and Lee, R., *"Weather Forecast Accuracy: Study of Impact on Airport Capacity and Estimation of Avoidable Costs", 8th USA/Europe Air Traffic Management Research and Development Seminar (ATM2009)*, Napa, California, June-July 2009.
16. "*FAA Operations and Performance Data.*" FAA, n.d. Web. July 26, 2011 <http://www.apo.data.faa.gov/>.
17. "*Accuracy and Precision.*" n.p., n.d. Web. July 26, 2011. <http://en.wikipedia.org/wiki/Accuracy_and_precision>.