Decision Support Tool for Predicting Aircraft Arrival Rates, Ground Delay Programs, and Airport Delays from Weather Forecasts

David A. Smith
Center for Air Transportation Systems Research
George Mason University
Fairfax, VA 22030
Telephone: (703) 975-1386
Email: dsmit4@gmu.edu

Dr. Lance Sherry
Center for Air Transportation Systems Research
George Mason University
Fairfax, VA 22030
Telephone: (703) 993-1711
Email: lsherry@gmu.edu

***Abstract*—The principle "bottlenecks" of the air traffic control system are the major commercial airports. Atlanta, Detroit, St. Louis, Minneapolis, Newark, Philadelphia, and LaGuardia all expect to be at least 98% capacity by 2012. Due to their cost and the environmental and noise issues associated with construction, it is unlikely that any new airports will be built in the near future. Therefore to make the National Airspace System run more efficiently, techniques to more effectively use the limited airport capacity must be developed.**

Air Traffic Management has always been a tactical exercise, with decisions being made to counter near term problems. Since decisions are made quickly, limited time is available to plan out alternate options that may better alleviate arrival flow problems at airports. Extra time means nothing when there is no way to anticipate future operations, therefore predictive tools are required to provide advance notice of future air traffic delays. This research describes how to use Support Vector Machines (SVM) to predict future airport capacity. The Terminal Aerodrome Forecast (TAF) is used as an independent variable within the SVM to predict Aircraft Arrival Rates (AAR) which depict airport capacity. Within a decision support tool, the AAR can be derived to determine Ground Delay Program (GDP) program rate and duration and passenger delay. The introduction of this decision support tool will expand the amount of time available to make decisions and move resources to implement plans.

I. Problem Statement

Air traffic congestion has become a widespread phenomenon in the United States. The principle bottlenecks of the air traffic control system are the major commercial airports, of which at least a dozen currently operate near or above their point of saturation under even moderately adverse weather conditions [1]. The Macroscopic Capacity Model (MCM) analyzed 16 airports within a 1000 nmi. triangle from Boston, Massachusettes, to Minneapolis, Minnesota, to Tallahassee, Florida. Based on this analysis, the MCM showed that in 1997 these airports were operating at 74% of maximum capacity. The model further went on to predict that the these airports will be at 89% capacity by 2012 [2].

The congestion problem is made worse because most airline schedules are optimized without any consideration for unexpected irregularities. When irregularities occur, the primary goal of the airlines is to get back to the original schedule as soon as possible, while minimizing flight cancellations and delays [4]. When trying to get back on schedule, sometimes it is the complexity of the situation, coupled with time pressure, which results in results in quick decisions that may be less than optimal [5]. Therefore, it would be advantageous to develop techniques to lessen the complexity of the situation and increase the time available.

One way to increase the time available is to create a tool that can predict the impact of weather on future inbound flight operations. Weather reports such as the TAF, Aviation Routine Weather Report (METAR), and the Collaborative Convective Forecast Product (CCFP) all provide raw weather forecast information. None of these forecasts though inform National Airspace System (NAS) stakeholders what the effect of that weather will be on flight operations. This research intends to fill this void by developing a process from which a forecast can be entered to produce estimate of the delay and capacity of the airport within the forecast area. Capacity estimates, in the form of AARs are produced for four time periods of the operational day. Ground Delay Program estimates of duration and program AARs along with expected delays can be derived from the predicted AARs. Now the forecast will not only provide the winds and ceiling, but also the AARs, GDPs, and expected delay.

II. Background

For efficient operation of the NAS, there is a need for the weather forecasting services and TFM products to estimate the reduction in capacity due to adverse weather. Weather forecast products are uncertain and the uncertainty increases with lead-time. Useful applications of weather forecasts requires either refinement, consultation, and application of the weather forecast to estimate air traffic capacity or decision support tools that take forecasts and make predictions based on past

 ISBN: 978-0-615-20720-9

forecasts and those forecasts connections to NAS capacity [6]. This paper describes a methodology used to create one such decision support tool known as the Weather Delay Prediction Tool. With this tool, the user enters the TAF for a given day and airport and the tool provides AAR predictions which can be derived to estimate delay and GDP time and duration.

Initially, this research focused on the CCFP as the weather forecast. The CCFP is a thunderstorm forecast for the entire United States and Canada and the research focused on its use as a predictive tool. After conversations with traffic management personnel and airline management, it was concluded that they rarely used the CCFP for any weather planning and relied on the TAF instead. The TAF has a good collection of available archived forecasts, so it was a good fit for the research objectives. To measure delays, a tool to predict GDPs was first considered. Over the course of the research it was determined that measuring delays may be more appropriate and then derive GDPs from the results. However, after presenting the work to air traffic management experts at the National Airspace System Performance Workshop, it was determined that it was better to use the AAR, since that was a common used factor to measure degraded airport capacity due to irregular operations. Also, GDPs and delays can be derived easily if the AAR is known.

III. Method

The general procedure used to determine a connection between weather forecast and airport cpapcity was:

- Collect data from the various available data sources,
- using assorted tools, format the data into a usable layout,
- use a classification tool to connect the two sets, and
- test the data to ensure there is a correlation.

A. Data Collection

FAA officials, airlines, air traffic controllers and others say Philadelphia plays a major role in delays up and down the coast thanks to poor airport design, bad weather, heavy traffic and close proximity to New York. Through September 2007, 68% of departures were on time in Philadelphia, better only than New York's JFK International, Chicago's O'Hare International and Liberty International in Newark, N.J. Fewer than two-thirds of arrivals were on time in Philadelphia during that period. The FAA has deemed Philadelphia a "pacing" airport that, because it sits in the middle of the busy East Coast air corridor, causes delays nationwide. Because of these facts, Philadelphia was chosen as the airport to evaluate the weather prediction tool [7]. The data used in this paper came from three areas:

- The TAF data was collected from a website provided by the National Climatic Data Center (NCDC).
- The Aircraft Arrival Rate data was collected from the Aviation System Performance Metrics (ASPM) database based maintained by the FAA.
- The delay data was found on the Bureau of Transportation Statistics website for summary statistics for destination airports.

1) Terminal Aerodrome Forecast: The TAF is an operational forecast consisting of the expected meteorological conditions significant to a given airport or terminal. TAFs always include a forecast of surface wind speed and direction, visibility, and clouds. Weather type, obstructions to vision, and low level wind shear are included as needed. The National Weather Service (NWS) produces over 570 TAFs. A TAF is a report established for the 5 statute mile radius around an airport. In the U.S., TAFs are produced four times a day starting at approximately 30 minutes before each main synoptic hour (00Z, 06Z, 12Z, and 18Z). All the forecasts produced starting one hour before the main synoptic hour up to four hours past the main synoptic hour are considered to be for the same cycle [8]. NWS is responsible for providing terminal forecasts to commercial and general aviation pilots for the protection of life and property and in response to requirements levied by International Civil Aviation Organization (ICAO) via the FAA in order to promote the safety and efficiency of the NAS.

2) Aircraft Arrival Rates: A Strategic Plan of Operations for managing flows during severe weather events in the NAS takes into account reduced AARs due to weather constraints. If the predicted capacity (number of aircraft that the airport can safely land in a given time period) falls short of scheduled demand (number of aircraft that wish to land at an airport in a given time period), traffic flow managers may implement a GDP [9]. GDPs are implemented by the Air Traffic Control System Command Center (ATCSCC) after consultation with regional Federal Aviation Administration (FAA) centers and with airline operations centers. A GDP applies to a particular airport, has specified start and stop times, and sets an allowable arrival rate.

Originally this research we focused on predicting GDPs by using the SVM. However, after discussions with air traffic managers, it was decided that it was more appropriate to predict AARs. AARs offer several advantages. First, each airport tends to revert to a finite set of AAR rates when airport capacity had to be reduced due to weather. This allowed grouping the possible outcomes into only a few distinct bins. Then a value was chosen between each bin and tested whether the day was $\geq$ to the in between value or $<$ the between value. Finally, a predictor function was developed for each of these values and from the results we were able to predict the future AAR.

The second advantage of the AAR was that GDPs could be predicted based on the conclusions of the predictor function. GDPs occur when the AAR is below the rate for a normal operations when the weather is favorable. AAR predictions are made for four times during the day based on the demand level of the airport. This generated a graph found in Figure 1. For this airport, the greatest demand hours were at 0700, 1100, 1500, and 2000 local time. Table I shows the demand hour and the assumed coverage hours for the airport. This airports normal AAR was 44, so Figure 1 predicts a GDP from 1300 to 2400.

 ISBN: 978-0-615-20720-9

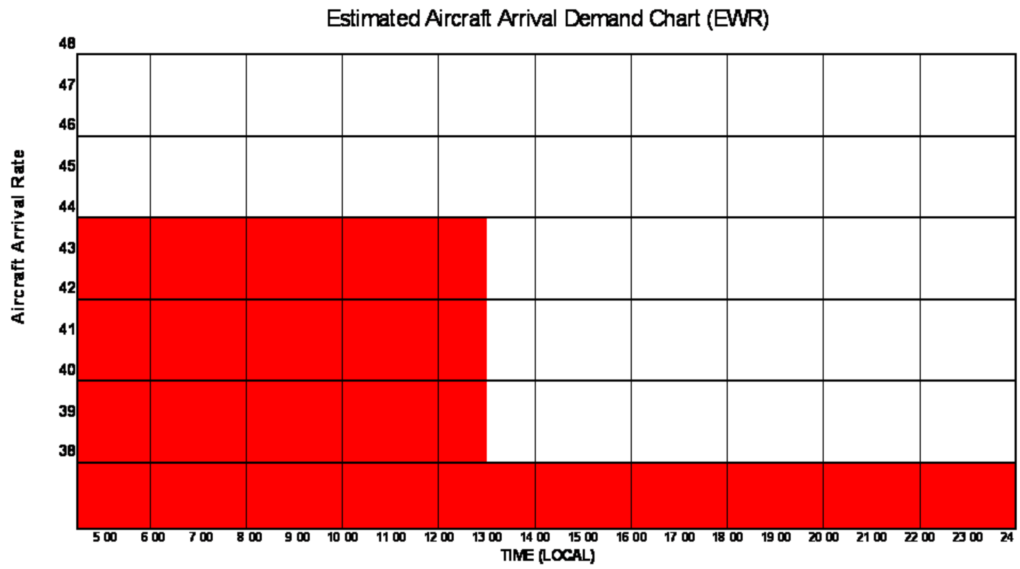

Fig. 1. Aircraft Arrival Demand Chart

Demand Hour	Assumed Time Block
0700	0500-0900
1100	0900-1300
1500	1300-1730
2000	1730-2400

TABLE I
DEMAND HOUR AND ASSUMED TIME BLOCK

B. Support Vector Machines

The Support Vector Machine (SVM) is a supervised learning method that generates input-output mapping functions from a set of labeled training data. In our case we are using the mapping function as a classification function. In addition to its solid mathematical foundation in statistical learning theory, SVMs have demonstrated highly competitive performance in numerous real-world applications, such as bioinformatics, text mining, face recognition, and image processing [10]. SVMs are based on the concept of decision planes that define decision boundaries. A decision plane is one that separates between a set of objects having different class memberships. A schematic example is shown in the Figure 2. In this example, the objects belong either to class square or circle. The separating line defines a boundary on the right side of which all objects are squares and to the left of which all objects are circles.

Figure 2 is a classic example of a linear classifier, i.e., a classifier that separates a set of objects into their respective groups (square and circle in this case) with a line. Most classification tasks, however, are not that simple, and often more complex structures are needed in order to make an optimal separation, i.e., correctly classify new objects (test cases) on the basis of the examples that are available (training cases). This situation is depicted in Figure 3. Compared to Figure 2, it is clear that a full separation of the square and circle objects would require a curve (which is more complex than a line). Classification

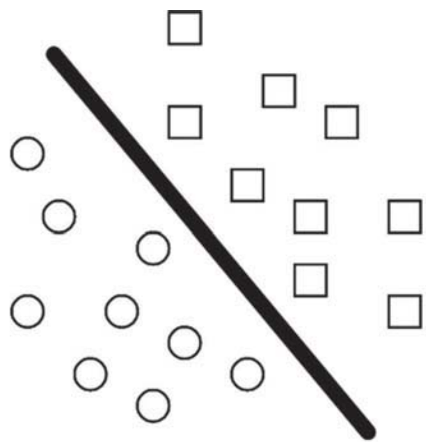

Fig. 2. Separating line defines a boundary

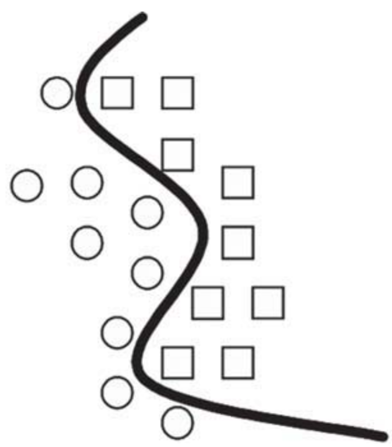

Fig. 3. Full Separation requires a curve.

 ISBN: 978-0-615-20720-9

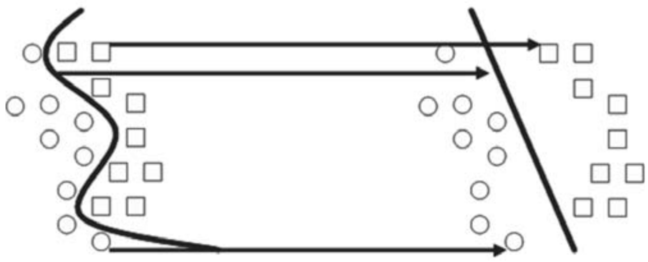

Fig. 4. Objects are mapped using a set of mathematical functions.

tasks based on drawing separating lines to distinguish between objects of different class memberships are known as hyper-plane classifiers. Support Vector Machines are particularly suited to handle such tasks.

Figure 4 shows the basic idea behind Support Vector Machines. Here we see the original objects (left side of the schematic) mapped, i.e., rearranged, using a set of mathematical functions, known as kernels. The process of rearranging the objects is known as mapping (transformation). Note that in this new setting, the mapped objects (right side of the schematic) is linearly separable and, thus, instead of constructing the complex curve (left schematic), all one has to do is to find an optimal line that can separate the square and the circle objects. Support Vector Machine (SVM) is a method that performs classification tasks by constructing hyper-planes in a multidimensional space that separates cases of different class labels. SVM supports both regression and classification tasks and can handle multiple continuous and categorical variables. For categorical variables a dummy variable is created with case values as either -1 or 1. For this type of SVM, training involves the minimization of the error function:

$$\begin{aligned} &\tfrac{1}{2}w^T w + C\sum_{i=1}^{N}\xi_i \\ s.t. \quad & y_i\left(w^T x + b\right) \geq 1 - \xi_i \\ & \xi_i \geq 0, i = 1 \ldots N \end{aligned}$$

Where C is the capacity constant, x is the vector of coefficients, b a constant, y the dummy variable, and ξ_i are parameters for handling non-separable data (inputs). The index i labels the N training cases. The kernel w is used to transform data from the input (independent) to the feature space. It should be noted that the larger the C, the more the error is penalized. Thus, C should be chosen with care to avoid over fitting [11].

C. Predictor Function

After collecting the TAF data as the independent variable matrix and the ASPM AAR as the dependent variable vector the SVM was applied to determine a function to predict future AAR's. The quadratic program introduced earlier was coded into AMPL. AMPL is a comprehensive and powerful algebraic modeling language for linear and nonlinear optimization problems, in discrete or continuous variables. After coding, the program was submitted and the associated data to the NEOS Server for Optimization.

1) Creating the TAF Vector: The x in the quadratic program represented the 57 character long vector from the TAF weather data collected from 2002 through 2006. To create the vector, TAF data was collected from a website provided by the National Climatic Data Center (NCDC). These files tend to be long, up to 100 pages of text data, because all reports received are placed in these files as they are received and they are updated approximately every five minutes as data becomes available. Also, forecasts may be duplicated within the files and multiple forecasts received from a station may appear in a file [12]. To transform the raw TAF data into usable vector form, data was pasted into an Excel Spreadsheet. Then the text to column function was used to put each part of the data into a separate cell. After the data was transformed into a linear format, it was then parsed down to include only the 0600 Zulu TAF reports. It was assumed that planning would take place early in the morning and the 0600 Zulu TAF, which equates to 0100 EST, was the first of the day.

2) Support Vector Machine Method: The first step in the process was to find the common AARs for each airport in the study. Using the ASPM database, AARs were collected for each of the four peak hours for the 1826 days in the dataset. Airports tend to have a set of common AARs that they use, so there are a consistent set of values to perform the classification algorithm.

In the quadratic program, y represents a binary variable that indicates whether or not an AAR was set at a certain numerical rate for a given airport. Values equal to -1 indicate that day was greater than or equal to the numerical rate while values equal to 1 indicate that day was less than the numerical rate. One advantage of the SVM method is the way it deals with data outliers. For most methods, statistical techniques are used to eliminate values that are considered abnormalities. The SVM has an error function in the objective function, were the C variable is set to a value that increases or decreases the number of incorrect classifications within the data. A high C allows fewer outliers, while a smaller C allows more. For our analysis C was set at 1000 after experimenting with other values. This helped to determine a ξ vector, which was only used to relax the function, so a solution was possible. The ξ vector was not used in the final prediction function.

For the independent variables, the five years worth of data included 1826 days so this created an 1826×57 data matrix for the independent variable. The AMPL code was run on the NEOS Server and found a solution vector w and variable b for each airport. After determining the w vector and the y variable the current TAF forecast could be used to develop an x vector using that data and then use Equation 1 to develop a prediction value.

$$w^T x_i + b \tag{1}$$

If the prediction value was greater than 0, then the algorithm predicts that less than an AAR will occur on that day.

 ISBN: 978-0-615-20720-9

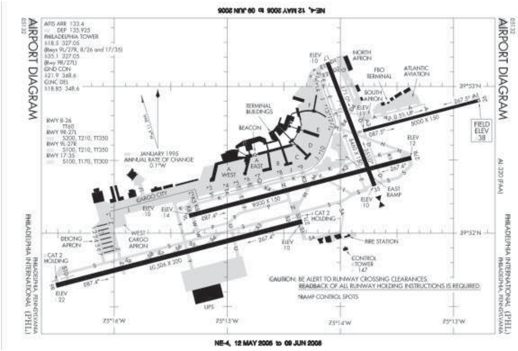

Fig. 5. Philadelphia Airport Map
[13]

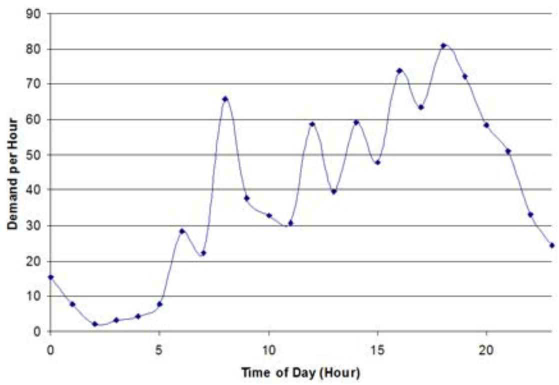

Fig. 6. PHL Hourly Demand

Conversely, if the value is less than zero then the algorithm predicts greater than an AAR value on that day.

IV. Results

Figure 5 shows that the primary runways at Philadelphia are 9L/R and 27 L/R. Poor runway arrangement limits the number of planes that can take off from the airport at once, especially during bad weather. Although a small runway was added in 1999, most of the layout dates back to the 1970s or earlier [7].

The SVM only classifies the data for a given AAR value. In order to create a useful tool, several SVM operations had to be done for one airport. The first step in the process was to find the average demand rates for each hour during the day. These peaks are highlighted in Figure 6.

Figure 6 shows six peaks, but to reduce the dimensionality of the problem we chose only 0800, 1200, 1600, and 1800. For those time periods the most common AAR was 52, indicating normal operations, which occurred 60% of the time. During times of irregular operations the AARs are reduced to 48 for 20% of the time and 36 for 9% of the time. Because these three AARs constitute 89% of the possible AARs, these were set as the only possible solutions that the model will predict. Two SVMs were solved for each time period. The first SVM will classify whether or not the day had an AAR less than 52 or greater than or equal to 52. If the SVM classifies a given

Time	Divider	Sensitivity	Specificity	PPV	NPV	Correct
0800	48	0.38	1.00	1.00	0.84	0.86
	52	0.61	0.90	0.79	0.78	0.78
1200	48	0.35	0.96	0.64	0.88	0.86
	52	0.50	0.91	0.79	0.74	0.75
1600	48	0.31	0.98	0.75	0.89	0.88
	52	0.49	0.91	0.74	0.76	0.75
1800	48	0.32	0.98	0.75	0.89	0.88
	52	0.48	0.90	0.72	0.75	0.75
	Combined	0.46	0.95	0.77	0.82	0.81

TABLE II
PHILADELPHIA TRAINING DATA

Time	Divider	Sensitivity	Specificity	PPV	NPV	Correct
800	48	0.40	1.00	1.00	0.83	0.85
	52	0.49	0.94	0.81	0.79	0.80
1200	48	0.44	0.99	0.85	0.92	0.91
	52	0.40	0.92	0.70	0.77	0.76
1600	48	0.36	0.98	0.77	0.89	0.88
	52	0.39	0.92	0.66	0.80	0.78
1800	48	0.28	0.98	0.73	0.88	0.87
	52	0.35	0.92	0.61	0.79	0.76
	Combined	0.40	0.96	0.75	0.84	0.83

TABLE III
PHILADELPHIA TESTING DATA

day and time period as greater than or equal to 52, then the tool will show a AAR of 52. This AAR would also indicate no GDP during this period. If the SVM predicts less than 52, then we would develop an SVM to test to see if the given day and time period is less than 48 or greater than or equal to 48. Again, if the SVM indicates greater than 48, then the tool sets the AAR to 48 and indicates a GDP. If the SVM indicates less than 48, then we set the AAR is set to 36 and also a GDP is predicted during this period. The duration of a predicted GDP is based on what time periods have GDP AARs.

A. Philadelphia Results

To evaluate how the SVM worked for Philadelphia, two methods were applied. The first method observed the success rate of the SVM prediction functions for the two test points for each time period. Data was also separated between training data, which was the data from January 2002 through December 2006, and testing data which is data from January 2007 through June 2007. The results for the training data are found in Table II and the results for the testing data are found in Table III.

Table II and Table III indicate that the SVM algorithm was correct 81% of the time for the training data and 83% for the testing data. To create a meaningful tool containing these algorithms a set of rules was established to estimate the AAR. The tool only considers three possible AARs, one associated with normal operations, one associated with a slight reduction in capacity, and one associated with a large reduction of capacity.

The first rule tested whether or not the point, that represents a day, was below 48. If it was below 48, then the AAR was determined based on a weighted average of the observed AARs

 ISBN: 978-0-615-20720-9

			Predicted AAR		
Time	Actual	Accuracy	36	48	52
0800	36	0.718	158	41	21
	48	0.409	102	135	93
	52	0.778	96	187	993
1200	36	0.640	105	32	27
	48	0.464	87	141	76
	52	0.736	109	249	1000
1600	36	0.746	85	21	8
	48	0.396	89	126	103
	52	0.736	94	238	1026
1800	36	0.754	89	17	12
	48	0.389	88	126	110
	52	0.754	98	242	1044

TABLE IV
TOOL RESULTS FOR PHILADELPHIA TRAINING DATA

		Predicted AAR		
Time	Delay	36	48	52
0800	Low	0	0	0
	Mean	9	0	0
	High	19	8	1
1200	Low	0	0	0
	Mean	15	0	0
	High	37	9	3
1600	Low	1	0	0
	Mean	25	6	0
	High	48	19	9
1800	Low	13	0	0
	Mean	54	19	7
	High	96	42	22

TABLE VI
PHILADELPHIA DELAY PREDICTIONS

			Predicted AAR		
Time	Actual	Accuracy	36	48	52
0800	36	0.760	19	3	3
	48	0.250	5	3	4
	52	0.793	15	15	115
1200	36	0.923	12	0	1
	48	0.300	6	6	8
	52	0.772	8	26	115
1600	36	0.769	10	2	1
	48	0.294	7	5	5
	52	0.829	12	14	126
1800	36	0.727	8	1	2
	48	0.278	7	5	6
	52	0.810	15	14	124

TABLE V
TOOL RESULTS FOR PHILADELPHIA TESTING DATA

below the tested rate, which for all four time periods was 36. If the SVM indicated the point was equal to or greater than 48 or less than 52, then we assumed the AAR was 48. All other results were assumed to be 52. Table IV and Table V show the tool performance for the training and testing data.

Table IV and Table V show that the accuracy is better at the extreme points then the points in the middle. This shows that the SVM method is better at finding extreme points on the edge instead of points inside.

B. Delay Prediction

Within the airline industry and air traffic management the AAR determines the airport capacity and is used to highlight the severity of a GDP, therefore it is the preferred prediction variable. Most flying consumers do not understand what AARs are and prefer to know what are the potential delays. The Weather Channel uses a Red, Amber, or Green rating system to highlight the airport impact. Although, the website does not explain the rating system, one would assume that Red impact means the most delays and Green impact means little to no delays. Amber is somewhere in the middle. To make the Weather Delay Prediction Tool applicable to the traveler, a delay prediction needed to be added.

1) Delay Prediction Method: Since the SVM model only predicts three potential AAR outcomes, then the average delays during the time of those AARs would provide not only a mean value, but also a range. Using the AAR data from ASPM and the delay data from Bureau of Transportation Statistics website, the average delay was calculated for each corresponding AAR. To provide a range of values, the standard deviation was calculated and added and subtracted to the mean value to provide a range.

2) Delay Results: Delay values are negative if the average arrival is early. Therefore, many of the average values are negative. If this was true then the value was changed to zero. Delay values and ranges are rounded to the nearest minute in Table VI. Table VI indicates the delay mean and the high and low range for each time period and predicted AAR. For instance, for the 1800 time period, if the model predicts an AAR of 36, then the delay mean is 54 minutes with a range as high as 96 and as low as 13 minutes. Now the flying consumer has information that they can use to plan their travel day.

C. Strategies for the Weather Delay Tool

Hub and spoke networks have become the most popular type of airline scheduling. In this type of scheduling, several points of departure feed into a single hub airport from which connecting flights carry passengers to their final destination. The advantage of the cross-connections is the multiplier effect as to the number of city pair that can be served. However, airports that are designated as the “hub” are subjected to increased congestion that are exasperated by irregular operations [4].

Meyer et al. (1998) [14] introduces the reliever and floating hub concept in a 1998 paper. Since most airline schedules are made without regard to unexpected daily changes due to severe weather conditions, there is very little slack time which means that any delay early in the day is likely to affect the schedule for the rest of the day unless the airline can take effective steps to correct the problem. Most carriers have developed procedures to follow in the event of unexpected disruptions in operations. However, most of these procedures are implemented manually, with little or no reliance on automated decision support systems [4].

The reliever hub is a strategy to reassign and optimize airport and airline schedules when experiencing a disruptive disturbance at a major hub airport and still maintain reasonable service. Figure 7 shows an example of a hub and spoke system.

ISBN: 978-0-615-20720-9

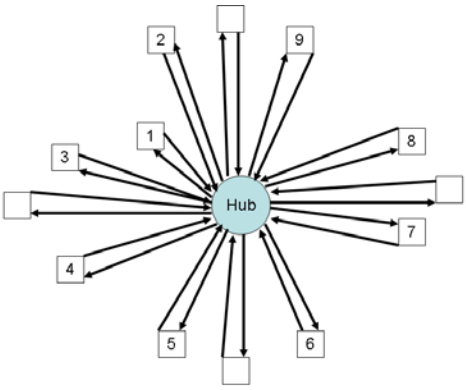

Fig. 7. Hub and Spoke System

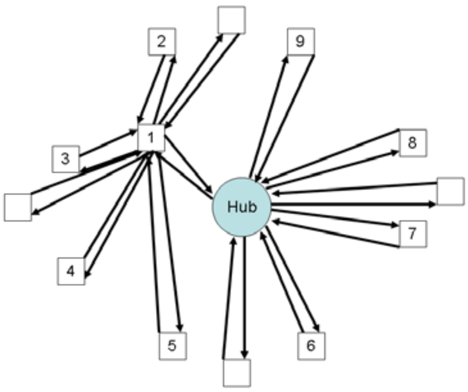

Fig. 8. Reliever Hub Option

The strategy is to temporarily use a nearby airport to act as a connecting hub, which can reduce the delays caused by a capacity reduction at the major hub. Figure 8 shows city 1 as a reliever hub. All cities to the west of the hub are sent to the reliever hub while all cities to the east continue to go to the main hub which reduces the demand on the main hub and decreases delays within the system. Service from city 1 to the hub would probably have to use a larger plane or more flights to insure passengers that need to get to the eastern cities or to the hub city arrive at their final destination.

Determining the best location for a reliever hub is a complex task. Even if the alternate hub is chosen ahead of time, the airline still must reconfigure schedule and passenger itinerary to minimize the total delay. Whether this reconfiguration is done by hand or is automated, it still requires time and employee manpower. It also requires a decision to use this manpower. Manpower has a cost associated with it, so the airline manager has to reasonably sure that the labor cost will help reduce any future loss due to extended delays. Because the TAF Delay Prediction Tool predict future delays it provides time and justification. Since the tool has been trained by historical data, it makes an AAR prediction based on what happened with similar TAFs in the past. Now the manager has justification to begin planning for the reliever hub.

Because the manager can enter the TAF a day in advance, there is now time to implement the plan by rescheduling flights and even informing passengers of any changes. Larger or smaller aircraft can be swapped to account for the change in passenger. This time also allows ground crews at the reliever hub more time to ready themselves for unusually high activity. The reliever hub may not have the permanent infrastructure to support an increase in passengers, therefore temporary solutions may have to be implemented. Since there may not be enough gates, planes may have to be serviced on the apron. This may require the use of buses and bus drivers to drive passengers from the planes to the terminal. Temporary shelter may have to be set up to shield passengers from the heat or cold.

V. Conclusion and Future Work

A. Weather Delay Prediction

The paper shows the possibilities of a Weather Delay Prediction Tool and what it can do to help NAS stakeholders. The algorithm is capable of classifying weather forecasts into three sets, where each set represents a specific AAR. Typically the highest AAR represents the airport during normal operations, while the two lower values represent reduced capacity due to weather or other congestion issues. Analysis showed that the SVM was more effective at predicting the normal AAR and the lower reduced capacity AAR. Therefore, for the weather delay prediction tool, it is appropriate to set a red, amber, or green scale. If the tool indicates green operations, then it is likely that the capacity at the airport will be at the maximum AAR and delays will be minimal. If the tool indicates red operations, then it is likely that the capacity at the airport will be significantly reduced and delays may be excessive. The Amber response indicates that the prediction is more uncertain, however, planners should prepared to have reduced operations at that airport. This appears to be the same rating system employed by the Weather Channel website, but it is also used by the military to rate progress of projects, describe the suitableness of terrain for armored vehicles, or any other situation that requires a general rating. Since the tool provides only a general assessment of airport capacity through AARs, then a general prediction of delays is also included based on AAR and time period. This tool helps the flying public know how long they can expect to be delayed due to weather.

GDPs will be predicted based on the tool prediction for each time period. The time periods for each airport were determined based on the demand peaks during the operational day. The two important pieces that came from a GDP are the programmed AAR and the duration. Programmed AAR is predicted based on the tool's prediction for each time period. The length of the GDP is determined based on which time period are below the normal rate. For instance, the peak time periods at Philadelphia were 0800, 1200, 1600, and 1800. Therefore, if the AAR prediction at 1600 and 1800 were below normal and the AARs at 0800 and 1200, then we assume that the GDP begins half way between 1200 and 1600 at 1400 and lasts until the the end of the operational day at 2400.

ISBN: 978-0-615-20720-9

B. SVM Disadvantages

A disadvantage of the SVM is that it does not show if any factor has more influence on the outcome then another. For each individual prediction equation developed, there were same factors that were weighted higher than others. The prediction equation is not an intuitive answer. However, across all of the prediction equations, there was not a value that consistently had more influence than another. By the nature of the algorithm, recursive partitioning searches for the value that best divides the data, so if determining which factors have the most influence on the final solution, the recursive partitioning method is more appropriate.

It is difficult to determine the effect of some of the data sets on the SVM. For instance, construction and airport upgrades at an airport can create inconsistent data. Analysts can attempt to normalize the data to try to maintain a consistent data set. However there was no way to determine how this affected the actual results. Results from airports that required normalization may not be as accurate as airports that had a consistent data set. Also, the SVM can not predict rare occurrences. For instance if a AAR rarely happened and the SVM tried to separate it from the rest of the data, a prediction vector of all zeros was the output with a b value of either 1 or -1.

C. Future Work

1) Proper Data Set Size: One of the issues with the SVM is what is the right amount of training data needed to produce a prediction equation that produces the most accurate predictions without overfitting the data. The process in this paper was to develop a prediction equation and then compare it to the training data and then the testing data. The analysis used 57 factors in the prediction function. The number was based on the factors found in the TAF for four time periods. If more factors were added to the data, then the performance of the predictor function applied to the training data improved. Unfortunately this did not improve the performance of the testing data which indicates that the data was overfitted. This situation is similar to using a 8th order polynomial to predict a line with 8 variables. It performs well with the 8 variables, but has little prediction value for any new independent variable. An optimization algorithm could be developed to determine the proper mix of training data and factors. Additional data factors could be added by adding time periods and the size of the training data could be reduced, although not increased since TAF data only goes back as far as January 2002.

2) Factors Other than Weather: This research focused on weather and using a forecast product to predict reduced airport capacity. Within the paper, it was discussed that other factor such as schedule congestion and runway construction also affect airport capacity. Further study should determine if there are any factors, besides weather factors, that can be added to the set of independent variables to produce a better predictive model. Techniques such as linear regression, have methods available to add or remove variables to the equation. At present though, there is no such standard process for SVMs other than adding the variable and testing the results to see if there is improvement. Developing this technique in itself would entail extensive analysis.

ACKNOWLEDGMENT

We would like to thank the Center for Air Transportation System Research (CATSR) at George Mason University for providing support to this study. We would also like to thank Dr. George Donohue from George Mason University and Dr. David Rodenhuis from the University of Victoria, British Columbia. Earlier conversations with both were the genesis of this work and their advice on air traffic management and weather forecasting were vital. We would also like to thank Ved Sud from the FAA, Mike Brennan from Metron Aviation, and Mike Wambsgans from Flight Explorer for their technical contributions. Danyi Wang and Bengi Manley from the CATSR lab provided valuable information on delays and ground delay programs and provided input to early study results. I would also like to thank Dr. Rajesh Ganesan, Dr. Andrew Loerch, and Dr. Aimee Flannery for the time and effort they set aside to help us complete this research.

REFERENCES

[1] M. Terrab and S. Paulose, “Dynamic strategic and tactical air traffic flow control,” Rensselaer Polytechnic Institute, Tech. Rep., August 1992.

[2] G. Donohue and W. Laska, *United States and European Airport Capacity Assessment Using the GMU Macroscopic Capacity Model*. Lexington, MA: American Institute of Aeronautics and Astronautics, 2001, vol. 193, ch. 5, pp. 61–73.

[3] Y. Ageeva, “Approaches to incorporating robustness into airline scheduling,” Masters Thesis, Massachusetts Institute of Technology, August 2000.

[4] F. Durso, T. Truitt, C. Hackworth, D. Ohrt, J. Hamic, and C. Manning, “Factors characterizing en route operational errors: Do they tell us anything about situation awareness?” in *Proceedings of the International Conference on Experimental Analysis and Measurement of Situation Awareness*, D. Garland and M. Endsley, Eds. Daytona Beach, Fl: Embry-Riddle Aeronautical University Press, June 1996, pp. 189–196.

[5] D. Rodenhuis, “Hub forecast prototype test,” in *Paper J3.9, Proc. Aviation, Range, and Aerospace Meteor (ARAM), American Meteor. Soc*, June 2006.

[6] P. Walters, “Delays at Philly airport caused by poor design, bad weather,” *Aviation*, p. 20, December 2007.

[7] *National Air Traffic Training Program, Air Traffic Guide, Aviation Routine, Weather Report (METAR), Aerodrome Forecast (TAF)*, Aviation Weather Center, Washington, D.C., May 2007.

[8] B. Hoffman, J. Krozel, and R. Jakobavitis, “Potential benefits of fix-based ground delay programs to address weather constraints,” Metron Aviation, Inc. Herndon, VA 20170, Tech. Rep., August 2004.

[9] V. Kecman, “Studies in fuzziness and soft computing,” in *Support Vector Machines: Theory and Applications*, L. Wang, Ed. Berlin: Springer, January 2005, ch. Support Vector Machines - An Introduction, pp. 1–47.

[10] T. Hill and P. Lewicki, *STATISTICS Methods and Applications*. StatSoft, 2006.

[11] “HDSS access system, national climatic data center,” http://Hurricane.ncdc.noaa.gov, June 2007.

[12] *U.S. Terminal Procedure Publications*, Federal Aviation Administration, Washington, D.C., August 2007.

[13] E. Meyer, C. Rice, P. Jaillet, and M. McNerney, “Evaluating the feasiblility of reliever and floating hub concepts when a primary airline hub experiences excessive delays,” University of Texas at Austin, Tech. Rep., 1998.

ISBN: 978-0-615-20720-9