Prediction of Weather Impacts on Airport Arrival Meter Fix Capacity

Author, co-author (Do NOT enter this information. It will be pulled from participant tab in MyTechZone)

Affiliation (Do NOT enter this information. It will be pulled from participant tab in MyTechZone)

Abstract

This paper introduces a data driven model for predicting airport arrival capacity with a look-ahead time 2-8 hour forecast. The model is suitable for air traffic flow management by explicitly investigating the impact of convective weather on airport arrival meter fix throughput. Estimation of the arrival airport capacity under arrival meter fix flow constraints due to severe weather is an important part of Air Traffic Management (ATM). Airport arrival capacity can be reduced if one or more airport arrival meter fixes are partially or completely blocked by convective weather. When the predicted airport arrival demands exceed the predicted available airport's arrival capacity for a sustained period, Ground Delay Program (GDP) operations will be triggered by ATM system. Serious imbalances between demand and capacity occur most frequently when the airport capacity is severely degraded due to either bad airport terminal surface weather or inclement convective weather around airport arrival fixes. A model that predicts the weather-impacted airport arrival meter fix throughput may help ATM personnel to plan GDP operations more efficiently. This paper identifies the characteristics of air traffic flow across arrival meter fixes at Newark Liberty International Airport (EWR). The proposed approach, based on machine-learning methods, is developed to predict the weather impacted EWR arrival Meter Fix (MF) throughput. Sector forecast coverage is used to envision the weather impact on airport arrival MF flow, and the validation is accomplished by using Convective Weather Avoidance Model (CWAM) 0.5 to 2-hour and Collaborative Convective Forecast Product (CCFP) 4 to 8-hour look-ahead forecast data for the period of April-September in 2014. Furthermore, the regression tree ensemble learning of random forests approach for translating a sector forecast coverage model to an EWR arrival meter fix throughput model is examined. The results suggest that ATM decision makers in charge of MF flow control and GDP planning may benefit from adopting the airport arrival meter capacity prediction models to estimate the inclement weather impacts.

I.   Introduction

In today's airport arrival operations, as aircraft transition for landing, air traffic controllers guide the aircraft from cruise altitude to the runway. In order to ensure the aircraft are at appropriate altitudes, speeds, and separation, arrival aircraft are metered over the airport arrival meter fixes prior to entering the airport terminal area. It is crucial for air traffic controllers to manage airport arrival traffic to ensure efficiency and safety. Directing aircraft over a specified meter fix is to maintain minimum safety separation. During bad convective weather, the airport arrival meter fix maximum throughput is decreasing due to the increasing aircraft separation. The reduced arrival meter fix maximum permissible throughput, i.e. the reduced airport arrival meter fix capacity, can sometimes cause unnecessary arrival airborne delay and airborne holding; even through the airport runway capacity is still underutilized. Moreover, air traffic flows at some airport's arrival meter fix can be blocked, and the airport's capacity can be severely degraded during severe convective weather.

Airport arrival meter fix throughput affected by convective weather constraints has both direct (e.g. meter fix blockage) and indirect (e.g. separation standards at meter fixes) impacts on airport operations [1-3]. Therefore, estimating airport capacity under the convective weather forecast constraints at the airport arrival meter fixes becomes more and more challenging for long look-ahead times (2-hour and beyond) as the uncertainty inherent in weather forecast is increased significantly.

As an example, Figs. 1 and 2 illustrate New York Center ATM operations impacted by severe convective weather on July 2 of 2014 (Wednesday). Fig. 1 is a multiple graph of top, medium, and bottom vertical bar charts for the three airports EWR, JFK, and LGA, respectively. Each chart includes the airport total arrival Meter Fix (MF) throughput (bar), airport OAG (Official Airline Guides) scheduled arrival rate (cyan line), and the airport arrival rate (capacity), AAR (red line). The total MF flow rates for these three airports were consistently much less than the airport OAG scheduled arrival rate and below their airport runway capacity during 13:00-21:00 EDT on that day.
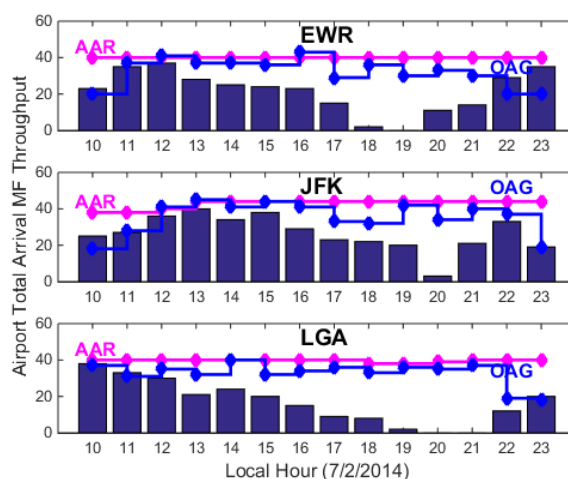

Figure 1. EWR, JFK and LGA airport total arrival Meter Fix (MF) hourly flow rate (throughput), OAG Arrival Rate, and AAR on 7/2/2014

10/19/2016

Fig. 2 displays the aircraft flight tracks arriving EWR, JFK, and LGA in green, blue, and magenta colored lines during 16:00-16:59 EDT on 7/2/2014. The red line in the figure represents the CWAM weather polygons for the 80% avoidance probability threshold at 16:00 EDT on that day. The plot shows that the airport northern and western arrival meter fixes for all three airports were blocked by the severe weather. Take a look at EWR arrivals (green lines) during the hour. The arrival aircraft from west had to be rerouted to the south, then through the EWR south arrival meter fix to arrive the airport.
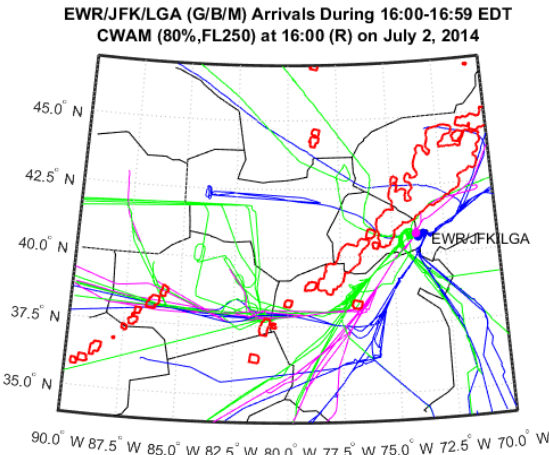

Figure 2. EWR, JFK, and LGA Arrival Flights during 16:00-16:59 EDT and CWAM weather polygons at 16:00 EDT on 7/2/2014 (Legend: Red polygons show CWAM weather at FL250 with 80% threshold. Green, blue, and magenta colored lines display the flight tracks arriving EWR, JFK, and LGA airport, respectively.)

Since airport arrival demands (OAG scheduled rate) had exceeded the available airport's arrival meter fix capacity for a sustained period (Fig. 1), Ground Delay Program (GDP) and Ground Stop (GS) operations were implemented at all three airports to smooth out the arrival flow and bring arrival demand in line with the airport capacity. The top and bottom tables in Table 1 show the GDPs and multiple GSs applied at these airports during the day, respectively. Both tables list the actual operating period, as well as the number of updates (due to the changes in program rate, maximum or average delay, cancellation, etc.) at each of these airports. The workload for airports and controllers can increase dramatically when too many modifications are involved.

Table 1. GDPs (top table) and GSs (bottom table) implemented on 7/2/2014 for EWR, JFK, and LGA

GDP		
Airport	Time Ranges (EDT)	#Updates
EWR	13:30-23.49	4
JFK	13:00-21:16	3
LGA	13:00-21:13	3

GS		
Airport	Time Ranges (EDT)	#Updates
EWR	12:07-14:30	1
	14:43-16:00	0
	16:36-19:15	7
JFK	12:13-15:30	2
	18:42-21:00	1
LGA	12:13-21:47	12

When the forecasted airport arrival demands exceed the available predicted airport's arrival capacity for a sustained period, GDP

10/19/2016

operations are implemented by ATM. A GDP intends to balance arrival demand and airport capacity by delaying aircraft departures at origin airports to achieve reduced arrival demand at the constrained destination airport. The major cause of GDP is weather; either inclement weather around meter fixes (thunderstorm as the GDP cause) or a bad terminal weather at the airport (such as low ceilings, low visibility, wind, etc. as the GDP cause) would have a great influence on GDP. The airport runway capacity, AAR, is degraded under severe terminal weather. Airport arrival meter fix throughput is reduced during the inclement convective weather, thunderstorms, which may not have impacts on airport runway capacity as indicated in Fig. 1. The causes for all GDPs and GSs listed in table 1 were due to thunderstorms.

GSs are usually used for a relative short term imbalance of airport demand and capacity, such as unexpected events or events with questionable duration time. In particular, GSs are often used to preclude additional extended periods of airborne holding for the arrivals destined for these airports. GSs are considered as being one of the most restrictive Traffic Management Initiatives (TMIs) and they override all other TMIs used to manage air traffic flows in the National Airspace System (NAS). However, the interaction of GDPs and GSs may create more scheduled arrival delays and airborne delays [4-6]. GS implementations during GDPs also indicate that the GDP was ineffective.

Efforts have been made during the past decade to understand the connection between weather and capacity both at the airport and airspace level. A common approach in many studies is to develop weather transition models using the historical distribution of throughput as targets during convective weather [7-14]. To date, there is little work on weather-related capacity investigation for non-standard airspace regions such as airport arrival meter fixes. To model the connection between weather and capacity (maximum flow rate) for the region, one may rely on the inspection of historical airport arrival meter fix throughput affected by convective weather.

This paper studies the use of a powerful machine learning ensemble method, Random Forecasts (RF), for projecting Newark Liberty International Airport (EWR) arrival meter fix throughput. The approach is to develop and evaluate predictive RF models using historical sector forecast weather coverages and airport arrival demands (schedule arrivals) as inputs and observed EWR arrival meter fix throughput as targets by the cross-validation method. Multiple sectors around the EWR airport arrival meter fixes and terminal arrival routes were selected for this experimental study.

Modeling of weather-impacted meter fix flow was designed to assist controllers and ATM personnel in determining the degree of arrival flow reduction caused by weather for operating and planning purposes. In post analysis, such a model can be used to check if the recorded operation is within the range of safely and efficiently controlled operations, or not, under similar circumstances.

The remainder of the paper is organized as follow. A description of the characteristics of arrival meter fix flows for EWR is presented in Section II. Section III depicts weather data and, sector weather coverage index (WCI), adopted to calculate the weather impact on sector capacity for the experimental setup. The correlation analysis for selected WCI sectors with different look-ahead forecasts and the model predictions for EWR meter fix throughput are presented in section IV. Finally, concluding remarks are presented in section V.

II. EWR Arrival Meter Fix Flow

Air traffic controllers meter aircraft over arrival MF to manage the traffic arrival flows prior to entering airport terminal airspace safely and efficiently. There are four MFs about 40 nautical miles from EWR: SHAFF at the north of EWR, PENNS at the west, and DYLIN and RBV at the south (see Fig. 3). The EWR meter fix ring is defined as a ring with a 40nm radius from the center at EWR. The arrival aircraft directions at the EWR MF ring to EWR (see Fig. 4) show that EWR arrivals were concentrated in the following three directions: from the south (ZDC center) across DYLIN (52° to EWR), from the east (ZOB center) across PENNS (105°), and from the north (ZBW center) across SHAFF (161°) in general.
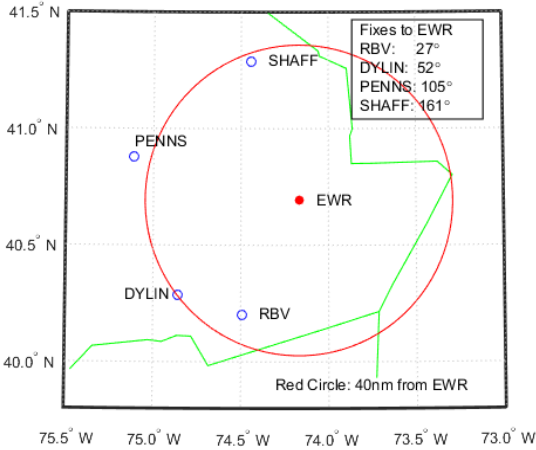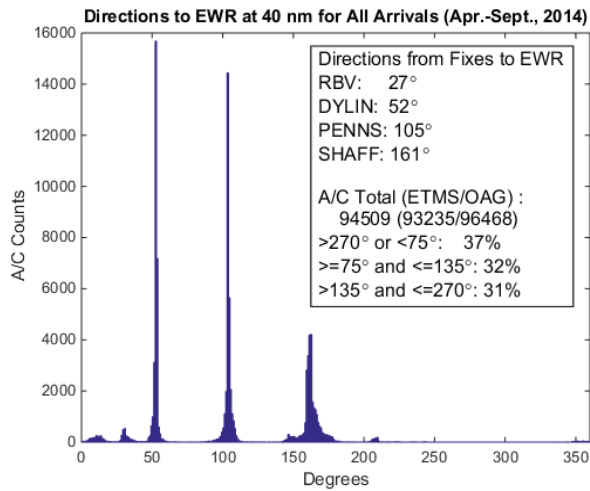

Figure 3. EWR Arrival Meter Fix Positions

Figure 4. Directions from Arrival Aircraft at 40 nm to EWR

The EWR DYLIN, PENNS, and SHAFF MF arrival throughput was calculated as the number of arrival aircraft passing through the MF ring with the directions to EWR within ($270\text{-}75^o$), ($75^o\text{-}135^o$), and ($135^o\text{-}270^o$), respectively. The sum of calculated daily EWR DYLIN, PENNS, and SHAFF MF throughput was consistent with the Aviation System Performance Metrics (ASPM) daily arrival rates, such as the counts for ASPM efficiency computation, ASPM metrics, and Enhanced Traffic Management System (ETMS) arrivals.

The arrival aircraft ground speeds and Flight Levels (FL) of EWR arrival aircraft at the MF ring for the three MFs are displayed in Fig. 5 and 6, respectively. The average ground speeds at the MF ring for three EWR MF flows were similar, at about 300 knots. The aircraft FLs at the MF ring for the three EWR MF flows were consistent with those required by EWR Standard Terminal Arrival Routes (STARs), which specify an altitude range of 7000-10000 ft. The hourly EWR arrival flow rates at the MF are shown in Fig. 7. Assuming a rate of 20 arrival aircraft per hour as the EWR MF operational acceptance rate (capacity), the arrow displayed in the figure points to the 98th, 99.7^{th}, and 99th percentiles for DYLIN, PENNS, and SHAFF throughput for the year 2014, respectively. Most MF throughput values shown in Fig. 7 were restricted by MF upstream demand and/or the EWR airport capacity. In the case where one MF flow exceeded 20 aircraft per hour, it was usually because other MFs were blocked by bad weather [2]. Using the average ground speed of 300 knots at MF, the average permitted lateral separation between two aircraft in an hour was about 15 nautical miles for 20 aircraft.
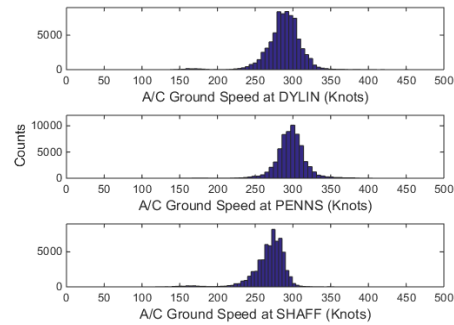

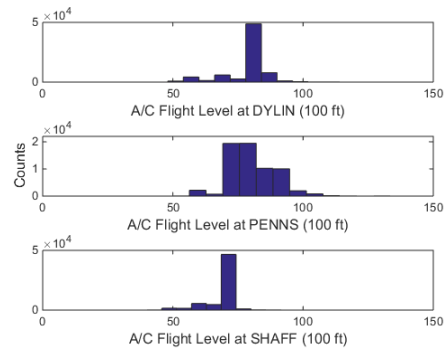Figure 5. EWR Arrival Aircraft Ground Speeds at MFs
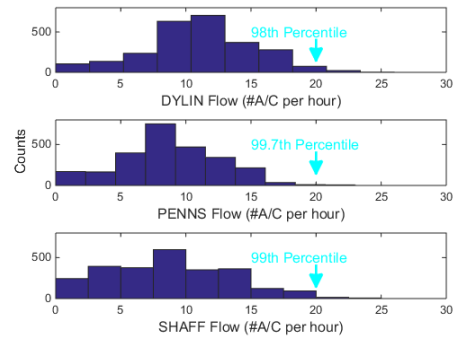

Figure 6. EWR Arrival Flight Levels at MFs

Figure 7. EWR Arrival Hourly Flow Rate at MFs

In Fig. 8, the normalized flows, defined as the percentages of the MF flows divided by the total MF flow, for the three EWR MFs are displayed. Even though the median of normalized DYLIN flow was higher than that for PENNS and SHAFF, they were all very close to one third of the total EWR MF flows. This indicates the EWR arrival flows were usually distributed evenly over three-meter fixes under normal operation conditions. With a maximum EWR AAR of 48 aircraft per hour from ASPM for the year 2014, even for the high demand cases, each MF throughput would not reach the capacity of 20 aircraft per hour.
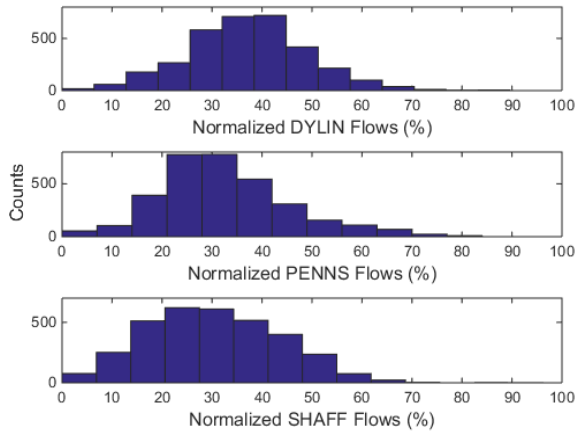

Figure 8. EWR Arrival MF Normalized Flow Rates

It is clear that unforeseen weather influence on landings at EWR can lead to large delays and ultimately be very costly to the airlines and the travelling public. If weather impacts are either short-lived or local, they can be mitigated effectively by using the available airspace. All the airborne and scheduled flights can be best handled using reroutes. Flights rerouted from one MF to the other MFs within a distance less than 400 nm between the reroute start point and EWR airport were selected from 2014 data. Almost all these reroutes were caused by weather. The flight rerouting percentage was calculated as the ratio between the number of rerouted flights and all flights. The results are listed in table 2.

It can be seen from Table 2 that rerouting happened more frequently from PENNS to SHAFF as well as from DYLIN to PENNS than that for the rest of the cases.

Table 2. Statistics of EWR Reroutes

Planned MF	Reroute to MF	Percentages
DYLIN	PENNS	1.06%
	SHAFF	0.02%
PENNS	SHAFF	1.39%
	DYLIN	0.04%
SHAFF	PENNS	0.05%
	DYLIN	0.00%

III.   Experimental Setup

This section describes the weather forecast data, sector weather coverage index, and the evaluation methods used for the meter fix throughput prediction study.

Weather Data
The basic source of weather information for this study was the Collaborative Convective Forecast Product (CCFP) and Convective Weather Avoidance Model (CWAM) forecast data.

The CCFP is a forecast for intense convection activity that is updated every 2 hours and made for 4, 6, and 8-hour periods by a group consisting of the National Weather Service (NWS), airline customers, Air Route Traffic Control Center (ARTCC) weather units, and the Canadian meteorological service. The CCFP is discussed, critiqued, and adjusted to develop a forecast based on many different inputs. It is the primary weather planning tool for Air Traffic Management personnel during severe convective weather periods. It consists of a defined area (polygon) and describes maximum cloud tops, growth and decay tendencies, the direction and speed of movement, and the forecaster's confidence in the forecast. The CCFP used in the study offers the polygon convective coverage rate at 25-49% (sparse) and 50-74% (medium) and the confidence level at high and low.

CWAM was developed based on Corridor Integrated Weather System (CIWS), which was developed by MIT Lincoln Laboratory (MIT-LL) [15]. This weather product combines data from dozens of weather radar with satellite data, surface observations, and numerical weather models in order to improve the accuracy and timeliness of the storm severity information. It provides automated, real-time, high spatial resolution data at a 5-minute update rate, and three-dimensional storm forecast information. CIWS also offers precipitation measured by vertically integrated liquid (VIL) and the Echo Tops forecast. CWAM models the flight deviation behavior of pilots around severe weather as a function of reflectivity level and echo-tops and translates convective weather information from CIWS data into Weather Avoidance Fields (WAFs) at each of the flight altitude levels. The WAF provides an estimate of the probability of aircraft deviation around severe weather in en route airspace as a function of horizontal location. For each of the WAF files, the data include polygons of regions of airspace that aircraft are likely to deviate around with corresponding avoidance probability thresholds, such as 60%, 70% and 80%. This data is provided for each of the flight levels from 25,000 ft (FL250) up to 45,000 ft (FL450) in 1,000 ft increments. The CWAM data include forecasted WAFs from zero to two hours with a 15 minutes forecast interval and an update rate of 5 minutes. The CWAM actual weather (zero hours) and 0.5 to 2-hour forecasts were used in this paper [16].

The sector weather coverage indexes (WCI) were used to indicate weather impacts on sector capacity. The sector three- and two-dimensional WCIs were computed using CWAM and CCFP forecast data, respectively. A description of the sector weather coverage model that was used to estimate the sector-level weather impacts are provided in the following subsections.

Sector Weather Coverage Index (WCI)
The sector two-dimensional WCI is the simplest sector weather impacted capacity model. This model estimates the sector-level weather impacts by calculating the extent to which a sector is covered by severe weather. The three-dimensional (3D) model enhances the two-dimensional weather coverage models by considering the weather at each vertical level plus storm height (e.g., echo top) data that is available in weather observations and forecasts. The 3D WCI is defined as follow:

$$WCI = \sum_{k=1}^{N_{FL}} w_k \times S_k$$

Here the summation is over the total number of flight levels, N_{FL}, within a sector, and flight levels are assumed to be separated by 1,000 ft. The flight-level dependent weighting factor, w_k, can be used to account for varying air traffic densities within a sector. However, for simplicity, w_k was set to a constant value of $1/N_{FL}$. Finally, S_k is

equal to the area of the weather coverage at the kth flight level divided by the sector area at the kth flight level. The weather coverage index ranges from zero to one. A value of zero indicates that no weather is present in a sector, while a value of one indicates that the sector is completely covered by weather.

Sector WCI Selection

Air traffic data for EWR arrivals from April 1, 2014 through September 30, 2014 were obtained from historical archives of the Federal Aviation Administration's (FAA's) Aircraft Situation Display to Industry (ASDI). Using this data, the sectors passed through for EWR arrival trajectories during good and bad weather days were selected for sector WCI calculations. Fig. 9 shows the sectors surrounding the meter fix circle used by most EWR arrivals during normal good weather days. WCIs of all sectors used by EWR arrivals during good and bad days were calculated for every hour. Both CWAM and CCFP forecast sector WCI data was used to estimate the weather impacted EWR meter fix throughput.
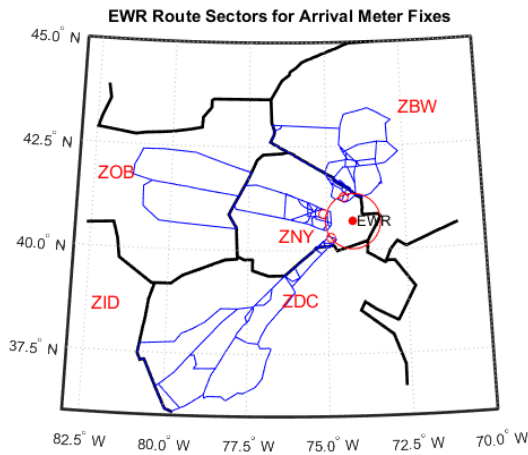

Figure 9. Sectors used by EWR arrivals in normal good weather days

Regression Tree Ensemble Learning

Regression Tree Ensemble learning is based on a predictive model composed of a weighted combination of multiple regression trees using machine learning algorithms [17]. The regression tree is a decision tree when the predicted outcome is a real number. A decision tree is built through a binary recursive partitioning process using a decision-tree algorithm. The tree is a flow-chart-like structure, where each internal node denotes a test of an input variable by the algorithm, each branch represents the outcome of a test, and each leaf or terminal node holds output variable values. A decision tree is a classic weak learner for which its predictive performance is better than random guessing and the training and prediction processes are fast by limiting the maximum depth of the tree.

Ensemble methods adopt multiple weak learners (regression trees) to obtain a better predictive performance than any of its individual constituent members can produce. The popular and powerful machine learning ensemble method, Random Forest (RF), was applied in this study [18].

The RF method works by training multiple weak regression trees using a fixed number of randomly selected features (one third of the number of features), then takes the average value for the weak learners and assigns that value to the predictor. Typically, the number

of weak trees generated could range from several hundred to several thousand depending on the size and difficulty of the training set. The method was implemented using the MATLAB ***Treebagger*** function [19].

Model Validation and Statistical Correlation Analysis

A cross-validation approach was implemented for model validation. In cross-validation, a series of RF models were constructed, each time by dropping a different part of the data from the training set and applying the resulting model to predict the target. The merged series of predictions for dropped or tested data were checked for accuracy against the observations. In one version of the cross-validation approach, called the group cross-validation approach, data are divided into N groups. A total of N models is then constructed with each using N-1 data groups for model training, and the Nth one for testing. Tenfold cross-validation was used in this paper.

A statistical correlation analysis was used to evaluate and compare the sector weather coverage indexes. Relations between two variables were described by the Pearson correlation coefficient r and r^2. A Pearson correlation greater than .8 is generally considered as strong whereas a correlation of less than .5 is generally treated as weak. As an example, if $r = .80$, then $r^2 = .64$, which means that 64% of the total variation in the actual weather can be explained by the forecast using a linear relationship. The other 36% of the total variation of actual weather remains unexplained by forecast.

Normalized Root Mean Squared Error (*NRMSE*) facilitates the comparison between models with different scale. *NRMSE*s were applied for assessment of model predictions for total and each EWR MF throughput (at different scales). *RMSE* of EWR MF throughput estimates for n test events is defined as

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(X_{obs,i} - X_{model,i})^2}{n}}$$
.

And *NRMSE* is defined as RMSE divided by the average of $X_{obs,i}$. where $X_{model,i}$ and $X_{obs,i}$ represent the MF throughput model estimate and observation measurement for the ith event, respectively.

The Pearson correlation coefficient r, r^2 and *NRMSE* have been used to evaluate the forecast "error" qualitatively in this paper.

IV.   Experimental Results

The experimental results were divided into two groups: the evaluation of uncertainties in forecasting sector WCIs using CWAM and CCFP forecast data, and the validation of EWR airport arrival meter fix throughput predictions against the actual throughput. A comparison of forecast averages of sector WCIs using the CWAM method compared with the actual (using CWAM zero-hour forecast) sector WCIs is provided in the first subsection. The evaluation of 4 to 8-hour CCFP forecasted sector WCIs is described in second subsection. Finally, the validation of predicted EWR arrival meter fix throughput is presented in the last subsection.

As described earlier, the correlation analysis was applied to evaluate the uncertainties of forecast sector WCIs. An accurately forecast sector WCI would have high correlation coefficients; on the other hand, lower correlation coefficients may indicate larger forecast errors. In the study validating EWR arrival meter fix throughput predictions using the RF modeling algorithm, NRMSE was also used to estimate the relative errors in throughput predictions, where lower values indicate less residual variance.

Evaluation of CWAM forecasted Sector WCIs

The accuracy in forecasting sector CWAM WCI varies with weather avoidance probabilities, as described by some previous studies [12, 13]. Lower weather avoidance probability used in sector WCI calculations usually results in better WCI predictions. Therefore, the different avoidance probability thresholds are used in the CWAM forecast validation study.

The correlation between the average forecasted sector WCIs calculated using 0.5 to 2-hour CWAM weather forecasts and the actual observations (zero-hour CWAM forecast) with 60%, 70%, and 80% weather avoidance probability are listed in Table 3. The correlation between forecast and observed WCIs varies considerably with the different forecast look-ahead times and the correlation difference with different avoidance probability for the same forecast look-ahead time was insignificant. This suggests that the ability to forecast WCI depends mainly on look-ahead time.

The CWAM WCI correlation coefficients r between 0.5-1hour forecasts and observations ranged from 0.84 to 0.95. This suggests that the correlations between the actual WCIs and short time (0.5-1 hour) forecast WCIs are quite strong. The correlations for more than one hour look-ahead times were not that strong. For example, consider the weakest correlation cases of the two-hour weather forecast WCI with 70% avoidance probability thresholds. In this case, the r-value of 0.67 indicates that only 45% (r^2) of the variation of the actual sector WCIs could be explained by the forecasted sector WCIs.

Table 3. Correlations between CWAM forecast sector WCIs and the actual sector WCIs

CWAM Forecast Sector WCI	CWAM Zero-hour Sector WCI		
	60%	70%	80%
0.5-hour	0.95	0.94	0.93
1-hour	0.88	0.86	0.84
1.5-hour	0.79	0.77	0.74
2-hour	0.69	0.67	0.65

Evaluation of CCFP forecasted sector WCIs

The correlation between the 6 or 8-hours forecast and 4-hour CCFP forecast sector WCIs was not weak, as presented in Table 4. The CCFP forecast sector WCI accuracy is also dependent on weather polygon coverage rate and the confidence level; therefore these two parameters are listed in the table.

Table 4. Correlation between the CCFP forecast 6 or 8-hour sector WCIs and 4-hour sector WCIs

CCFP Forecast Sector WCI	CCFP Four-hour Sector WCI		
	High, 50-74%	High, 25-49%	Low, 25-49%
Six-hour	0.80	0.84	0.83
Eight-hour	0.63	0.70	0.70

Table 5. Correlation between the CCFP 4, 6, or 8-hour forecasted sector WCIs and CWAM actual sector WCIs

CCFP Forecast Sector WCI (High, 25-49%)	CWAM Zero-hour Sector WCI		
	60%	70%	80%
Four-hour	0.39	0.36	0.33
Six-hour	0.37	0.34	0.32
Eight-hour	0.33	0.31	0.30

While the CCFP polygons with a 25-49% coverage rate took advantage of larger weather coverage area, as a result, the correlation values were in general slightly higher than that for the 50-74% coverage rate. But the difference is relatively insignificant comparing with different forecast look-ahead times.

The correlation between the CCFP 4, 6, or 8-hour forecast sector WCIs and the actual CWAM sector WCIs were quite weak -about 30-40% as displayed in table 5- meaning the likelihood of getting a good quality forecast is reduced as the forecast time becomes longer.

EWR Arrival Meter Fix Throughput Prediction

Weather impacted airport arrival meter fix throughput is affected not only by the inclement weather around the meter fix but also by arrival demand, airport terminal runway capacity (AAR), and airport arrival meter fix flows from previous hour [2].

To model hourly throughput for EWR arrival meter fixes altered by forecast, the data were filtered to select those events where there was a convective forecast within a distance of 400 nm between the weather and EWR airport in a one-hour time period. The hourly throughput observations for each EWR meter fix were defined as the model output, or target; while the hourly forecast WCIs for the underlining sectors described in section III were computed as model inputs to indicate how "bad" the forecast around meter fixes was. The input variables also include hourly EWR arrival demand (ASPM scheduled arrivals), EWR airport capacity, AAR (ASPM), and EWR meter fix flows at the previous hour from the forecast issue hour.

The EWR DYLIN, PENNS, SHAFF and total (sum of three fixes) forecast meter fix throughput ensemble learning models using the RF algorithm were trained and tested by a tenfold cross-validation using the input and output data in the time frame of April-September in 2014.

The predictive RF model performance for EWR total, DYLIN, PENNS, and SHAFF arrival MF throughput are listed in Table 6, 7, 8, and 9, respectively. Each table shows the RF model predictions for EWR arrival meter fix throughput using CWAM 0.5 to 2-hour look-ahead times and CCFP forecast 4 to 8-hour look-ahead times. The linear Pearson's correlation coefficient (r) between predictions and actual throughput, squared correlation coefficient (r^2), and the Normalized Root Mean Squared Error ($NRMSE$) of the meter fix throughput predictions for each look-ahead time are displayed in these tables.

The model predictions for the short look-ahead time forecast matching the corresponding actual throughput were better than those for the longer ones. The correlation between EWR total arrival throughput predictions and observations was quite strong with the coefficient varying from 0.9 to 0.7 for predictions at 1-hour to 8-hour (see Table 6).

The model prediction performance for each EWR individual meter fix throughput is not that strong, with correlation coefficients ranging from 0.8 to 0.5, see Table 7-9. However, for EWR DYLIN arrival MF throughput predictions, the correlation was better than other MFs with coefficients varying from 0.8 to 0.6 for 1-hour to 8-hour forecast (see Table 7).

The EWR total arrival MF throughput prediction performance for look-ahead time up to 4-hour (see Table 6) is quite good with NRMSE less than 25%. The model prediction performance for each of the EWR individual meter fix throughput values was not that great with NRMSE between 30 and 50% only.

The model performance for prediction of EWR total arrival MF throughput was better than the estimations of any individual meter fix throughput. This can be explained by the fact that the airport arrival demand (scheduled arrival) played an important role in improving model predictions.

For DYLIN, the model produced better throughput predictions than for the other two MFs. A possible explanation may be due to the fact that the DYLIN flow was the most dominant one (see Fig. 8) and the impact on DYLIN flow due to rerouted aircraft from other MFs was the smallest (see Table 2) among EWR arrivals coming from three directions.

The degradations of EWR MF throughput RF model prediction performance with the look-ahead time were better than that shown in Table 2 and 4 with the weather forecasted sector WCIs. This could be the case because the input features for RF models included not only forecast sector WCIs, but also air traffic demand predictions (scheduled EWR arrivals) and each MF actual throughput at the previous hours from the forecast issue hour.

Table 6. Prediction for EWR Total Arrival MF Throughput

Weather Forecast	r	r²	NRMSE
CWAM 30-minute Forecast	0.91	0.82	20%
CWAM 1-hour Forecast	0.89	0.80	22%
CWAM 2-hour Forecast	0.85	0.73	24%
CCFP 4-hour Forecast	0.78	0.66	25%
CCFP 6-hour Forecast	0.77	0.58	29%
CCFP 8-hour Forecast	0.73	0.54	32%

Table 7. Results for DYLIN Arrival MF Throughput Modeling

Weather Forecast	r	r²	NRMSE
CWAM 30-minute Forecast	0.82	0.67	32%
CWAM 1-hour Forecast	0.81	0.65	33%
CWAM 2-hour Forecast	0.77	0.58	34%
CCFP 4-hour Forecast	0.68	0.46	37%
CCFP 6-hour Forecast	0.67	0.45	39%
CCFP 8-hour Forecast	0.63	0.40	41%

Table 8 Validation Results for PENNS Throughput Modeling

Weather Forecast	r	r²	NRMSE
CWAM 30-minute Forecast	0.70	0.49	39%
CWAM 1-hour Forecast	0.68	0.46	40%
CWAM 2-hour Forecast	0.63	0.40	41%
CCFP 4-hour Forecast	0.60	0.36	43%
CCFP 6-hour Forecast	0.57	0.33	44%
CCFP 8-hour Forecast	0.52	0.27	45%

Table 9. Validation Results for SHAFF Arrival Modeling

Weather Forecast	r	r²	NRMSE
CWAM 30-minute Forecast	0.80	0.64	40%
CWAM 1-hour Forecast	0.78	0.61	42%
CWAM 2-hour Forecast	0.76	0.58	44%
CCFP 4-hour Forecast	0.73	0.53	45%
CCFP 6-hour Forecast	0.72	0.52	47%
CCFP 8-hour Forecast	0.70	0.49	49%

V. Summary

This paper begins by providing an extensive analysis of EWR arrival meter fix throughput impacted by severe convective weather. Widespread severe weather nearby or over the meter fix could prevent arrival aircraft from landing at the airport even though the airport capacity (AAR) is being underutilized. Airport arrival meter fix capability estimates are among the key inputs critical for implementing GDP operations. Achieving accurate airport arrival meter fix throughput predictions, however, is difficult due to convective weather forecast uncertainty, as described in the example from section I.

The paper subsequently presents machine-learning methods for predicting weather impacted EWR arrival meter fix throughput. These predictions were accomplished by using regression tree ensemble learning. The Random Forest algorithm was employed to train the throughput prediction models. The models are validated using data cross validation methods. When predicting the EWR total meter fix throughput under 4-hour predictions, the model was able to achieve relative errors (NRMSE) better than 25% and correlation between predictions and actual observations higher than 80%. For EWR individual meter fix throughput predictions, the worst relative error was better than 45% and the correlation was higher than 60% for a less than 4-hour forecast.

In summary, the predictions proposed here by the Random Forest models provide an approach to understanding and accounting for the uncertainty in weather impacted airport meter fix capacity and demonstrate how it is possible to learn from the past experience. The study provides information and may render aid in improving FAA TFM operations. Air traffic scheduling using flight track based operations could be enhanced significantly if the superior performance of airport arrival meter fix throughput prediction models were incorporated.

References

1. Clifford, S.F., et al., "*Weather Forecasting Accuracy for FAA Traffic Flow Management*", the National Academies Press500 Fifth Street, N.W. Washington, D.C., 2003.
2. Wang, Y., "Weather Impact on Airport Arrival Meter Fix Throughput", 36th Digital Avionics Systems Conference, 2017.
3. http://www.fly.faa.gov/Products/Training/Traffic_Management_for_Pilots/TFM_in_the_NAS_Booklet_ca10.pdf, "Traffic Flow Management in the National Airspace System" (Report). Federal Aviation Administration. October 2009.
4. https://www.nbaa.org/ops/airspace/tfm/tools/FAA-TMI-Interaction.pdf, "Traffic Management Initiative Interaction", NBAA, 2014.
5. Wang, Y., "Analysis and Prediction of Weather Impacted Ground Stop Operations", 33rd Digital Avionics Systems Conference, 2014.
6. Wang, Y. and D. Kulkarni, 2011, "Modeling Weather Impact on Ground Delay Programs", SAE International Journal of Aerospace, vol. 4 no. 2, 1207-1215; doi:10.4271/2011-01-2680.
7. Callaham, M., DeArmon, J., Cooper, A., Goodfriend, J., Moch-Mooney, D., and Solomos, G., "Assessing NAS Performance: Normalizing for the Effects of Weather", 4th USA/Europe ATM R&D Seminar, Santa Fe, NM, Dec. 4-7, 2001.
8. Chatterji, G. B., and Sridhar, B., "National Airspace System Delay Estimation Using Weather Weighted Traffic Counts", AIAA Guidance, Navigation and Control Conference, San Francisco, CA, August 15-18, 2004.
9. Krozel, J., Mitchell, S., Polishchuk, V., and Prete, J., "Capacity Estimation for Airspaces with Convective Weather Constraints", AIAA Guidance, Navigation, and Control Conf., Hilton Head, SC, 2007.
10. 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.
11. Sridhar, B., Wang, Y., Klein, A., and Jehlen, R., "Modeling Flight Delays and Cancellations at the National, Regional and Airport Levels in the United States", 8th USA/Europe ATM R&D Seminar, Napa, California, June 29 – July 2, 2009.

12. Wang, Y. and Grabbe, S. "*Comparison of Sector Capacity Weather Translation Models*". NASA Tech Report TM-2010-216401, 2010.
13. Wang, Y and Sridhar, B. "Convective Weather Forecast Accuracy Analysis at Center and Sector Levels", Digital Avionics Systems Conference, 2012.
14. Matthews, M.P., Veillette, M.S., Venuti, J.C., DeLaura, R.A. and Kuchar, J.K., "Heterogeneous Convective Weather Forecast Translation into Airspace Permeability with Prediction Intervals",Journal of Air Transportation, pp.41-54, 2016.
15. Evans, J. and Ducot, E., "Corridor Integrated Weather System"*,* MIT Lincoln Laboratory Journal, Volume 16, Number 1, 2006.
16. Matthews, M., Wolfson, M., Delaura, R., Evans, J., Reiche, C., Balakrishnan, H., and Michalek, D., "Measuring The Uncertainty of Weather Forecasts Specific to Air Traffic Management Operations", The 89th American Meteorological Society Annual Meeting, January 2009.
17. Breiman, L. *Bagging Predictors.* Machine Learning 26, pp. 123–140, 1996.
18. Breiman, L. *Random Forests.* Machine Learning 45, pp. 5–32, 2001.
19. MATLAB R2017a, Statistics Toolbox, TreeBagger Class, 2017.

Contact Information

Yao Wang, Ames Research Center Mail Stop 269-2, Moffett Field, CA 94035-0001 (yao.x.wang@nasa.gov)