Twelfth USA/Europe Air Traffic Management Research and Development Seminar (ATM2017)

Modeling Ground Delay Program Incidence using Convective and Local Weather Information

Yi Liu
Amazon
Seattle, USA
liuyi.feier@gmail.com

Mark Hansen, Danqing Zhang, Yulin Liu, Alexey Pozdnukhov
Department of Civil and Environmental Engineering
UC Berkeley
Berkeley, USA

Abstract—**In this work, we model the impact of weather condition on ground delay program (GDP) incidence using support vector machine (SVM) and logistic regression. We use SVM to analyze how spatial patterns of convective weather affect GDP occurrence and produce heatmaps to visualize the impact. Additionally, the SVM results are combined with local airport weather variables and airport traffic level indicator to yield a logistic model that considers both local conditions at the airport and convective weather in the surrounding area. We apply our methods to five airports: Newark Liberty International airport, John F. Kennedy International airport, LaGuardia airport, Philadelphia International airport, and Atlanta International airport. We find that the importance of convective weather depends on both its distance and direction from the airport. From the logistic regression we learn that both regional convective weather, as captured by the weights found in the SVM, and local weather are statistically significant. Convective weather is, however, the most important factor. Our models are found to have high accuracy and low false positive rates, but also low true positive rates because of the imbalance in our data.**

Keywords-Ground Delay Program; Convective Weather; Support Vector Machine; Logistic Regression

I. INTRODUCTION

Ground delay programs (GDPs) are commonly used in the United States when there are imbalances between flight demand and capacity at individual airports, or multi-airport metroplexes. When a ground delay program is implemented, flights bound for the affected airports are assigned controlled times of departure so that arrival demand at the destination does not exceed a specified rate. Most ground delay programs are the result of adverse weather that reduces airfield or terminal capacity. When instituted, ground delay programs severely impact flight operations, with delays as long as several hours as well as large numbers of flight cancellations. In 2011, 1065 GDPs were issued in the US, which imposes delay totaling 26.8 million minutes to 519,940 flights, an average of 52 minutes per impacted flight [1].

The purpose of this paper is to predict the incidence of GDPs based on weather conditions. The motivation is three-fold. First, flight operators would benefit from having greater foreknowledge of GDPs, so that they can plan their responses further in advance. This is an incidence of a broader desire to increase "predictability" in the national airspace system, which has been widely recognized [2; 3; 4; 5]. Second, FAA specialists responsible for GDP decision making would value a predictive model of GDP incidence both in order to anticipate when they may have to implement a GDP, and also to know when conditions are such that in the past a GDP would often have been implemented. In today's system, many FAA specialists have limited experience, and a model relating GDP incidence to weather conditions is an economical way of conveying to them GDP decisions made in the past. Finally, as will become clear later, the analytics used in developing model convey where weather matters for a particular airport, which can in turn be used to assess the similarity between weather conditions of a historical day to a given day of operation. Knowledge of what traffic management actions were taken on a similar historical day, and how well they worked, can also augment the personal experience of traffic management specialists [6].

While, as discussed in Section II, other researchers have also investigate the links between weather and GDP occurrence, our analysis is unique in its attention to the spatial patterns of convective weather. Using the support vector machine (SVM) technique applied to convective weather maps on a 1x1 square mile grid, we identify regions in which convective weather is associated with GDPs at a particular airport. In addition to confirming the importance of convective weather in the immediate area around the airport, our results identify more distant regions in which convective weather also has an influence. Additionally, the SVM results are combined with local airport weather variables such as wind and visibility to yield a model that considers both local conditions at the airport and convective weather in the surrounding area. In light of our focus on convective weather, we focus on attention on airports in the eastern US where convective weather is believed to be an important cause of GDPs.

Our work focuses on the relation between current realized weather and GDP occurrence. While for most practical uses weather forecasts would be required, our ability to accurately predict weather in future hours is continually increasing. Use of our models with weather forecasts, rather than realized weather, will reduce their reliability to some extent, but we will not address this problem in this research. Nor will we consider explicitly the fact that GDP decisions are themselves based on weather forecasts. The complex dynamics of "what did they know and when did they know it" are not readily observable; our working assumption is that the realized weather is a reasonable proxy for what was anticipated by GDP decision makers.

The remainder of this paper is organized as follows. In Section II, we review previous work on GDPs and in particular efforts to relate GDP occurrence to weather and traffic conditions. Our data and methodology are described in Section III. In Section IV we present our results. Conclusions and directions for future research are presented in Section V.

II. Literature Review

Most of the research on GDPs has focused on setting GDP parameters, such as GDP start time and GDP scope, using simulation approach or mathematical modeling assuming a GDP is needed [7; 8; 9; 10]. Recently, the aviation community has begun applying machine learning methods to understand how and why GDP decisions were made using historical data [11; 12; 13; 14].

Reference 11 used ensemble bagging decision tree (BDT) to predict GDP revision event and Neural Network (NN) to predict the GDP duration. The predictions were made at actual GDP initiation time. Their models employed the following variables: actual airport weather data from aviation system performance metrics, forecast of weather impacted traffic indexes (WITIs), and air traffic data such as scheduled arrivals. For convective weather, they considered two WITIs: en-route convective weather WITI with a scope of approximately 500NM range and local convective weather WITI with a scope less than 100 NM.

Reference 12 used logistic regression and decision tree to predict GDP occurrence at an hourly level for Newark Liberty International airport (EWR) and San Francisco International airport (SFO). Predictor variables in the models include actual weather condition variables, such as visibility and ceiling, and variables reflecting traffic condition at the airport, such as nominal queueing delay and demand capacity ratio. Convective weather in the airport belonged air route traffic control center is considered by using WITI. They found that logistic regression model outperforms the decision tree model in terms of prediction performance on the test data set where they used area under receiver operating characteristic curve as the performance metric. They also found that while WITI of New York Center (ZNY) is an important factor impacting GDP at EWR, the Oakland Center (ZOA) WITI does not have such a strong influence on GDPs at SFO.

Reference 13 used random forest and inverse reinforcement learning (IRL) to predict GDP initiation, cancellation and GDP parameters if a GDP is predicted to be in place. They modeled the GDP decisions in two submodels: first they predicted whether or not a GDP would be implemented for a given hour; then they predict GDP parameters such as GDP scope and GDP start time if a GDP is predicted as needed. The model is a simplification of reality in that it requires that a GDP plan either progress as planned or be canceled (no modifications or extensions are permitted). The predictable variables are the same for the two submodels except that the GDP parameter model does not have previous GDP plan as a variable: actual and predicted weather condition, traffic schedule, actual and predicted airport arrival rate, runway configuration, departure queue, reroute variables and ground and air buffers. The models left out convective weather variables and thus did not consider the impact of convective weather on GDP decisions. They applied their methods to EWR and SFO airports. They found that while random forest was better than IRL in predicting hourly GDP implementations at the two airports, both models struggled to predict the initiation and cancellation of GDP.

Reference 14 used random forest to predict hourly GDP incidences for the three airports in the New York area: EWR, John F. Kennedy International airport (JFK), and LaGuardia airport (LGA). The features considered in the model included scheduled arrivals, weather forecast for crosswind speed, visibility, thunderstorm, rain and snow, and distance from New York city to three levels of precipitation: very high level, high level, and moderate level. The three distance variables, which to some degree reflected convective weather, turned out to be the most important factors.

III. Data and Methodology

We model hourly GDP incidences using three types of information: actual convective weather condition, actual airport local weather condition and airport busy hour indicator. We use two data mining techniques: SVM and logistic regression. We use SVM to understand and quantify the impact of convective weather on GDP incidences where convective weather is represented using geo-referenced image. We use logistic regression to model GDP incidences using SVM output, airport local weather variables and airport busy hour indicator variable. We apply our analysis to five airports: EWR, JFK, LGA, Philadelphia International airport (PHL), and Atlanta International airport (ATL). The data range covers 2012 to 2014 with occasional gaps due to missing or incomplete data in one of the data sources for any of the airports. Below, we will introduce data used in this study, how data is processed to generate our variables including the dependent variable, and how SVM and logistic regression are employed in our analysis. We will introduce SVM algorithm together with convective weather.

A. GDP Label

The dependent variable in both SVM and logistic regression is a GDP occurrence indicator taking value 1 if GDP is in effect in a given hour for a given airport and 0 otherwise. We collect the GDP occurrence data from FAA's National Traffic Management Log.

B. Convective Weather Data and SVM Algorithm

The convective weather data come from the national convective weather forecast (NCWF) product, designed and implemented by the National Center for Atmospheric Research. NCWF provides current convective hazards, with locational information and a 6-level intensity scale of detected hazards, updated every 5 minutes, with the direction of movement and storm tops. The forecast of the evolution of convective hazard areas based on the thunderstorm identification tracking and nowcasting [15] algorithm are provided as well for a 1-hour forecasting horizon. Level 3 and higher levels are considered as significant hazard intensity justifying in-flight rerouting and other TMIs including GDPs. NCWF provides coverage at national airspace system level. Polygonal representations of convective areas provided in NCWF are discretized into geo-referenced image spanning the areas around the selected five airports. A range of area sizes and resolutions were evaluated experimentally, resulting in a squared area of 200x200 square miles centered at the airport of interest, with a resolution of 200x200 pixels and one pixel per 1x1 square mile. In this study, we use realized convective weather data at a beginning of an hour to predict GDP in the next hour. While we have observed that the forecasting performance drops with increasing the lead time (see Fig. 1 and a discussion below), the use of the longer-term forecast and different lead times would not impact the methodology and we are analyzing this in our on-going research.

A binary SVM classifier is used and NCWF data are transformed into indicator variables signifying the presence intensity level of 3 and higher for each pixel. The resulting hourly feature vectors **x** of binary covariates are used as inputs of a linear soft-margin SVM [16]. The number of samples is 25310. GDP label data are used as outputs. SVM learns a set of weights **w** and an offset b that lead to the decision function f(**x**)=**wx**+b that minimizes misclassification rate of the decision rule *sign*(f(**x**)) under a maximum margin separation criterion and a given trade-off hyper-parameter C. In the selected feature representation, the component of the weight vector **w** corresponds to the weight that the presence of a convective weather within a given 1x1 square mile location carries on the incidence of a GDP. The analysis of the spatial variability of **w** also carries valuable information for GDP decisions. In addition, the convective weather score **wx**+b reflects the likelihood of GDP at a given airport and can be used as a predictor in the logit models.

SVM scores (denoted as ***wx***) were used as covariates along with the airport-level variables within a logistic regression model presented below. The choice to use SVM exclusively for convective-weather imagery data with no extra airport-level variables is due to the imbalance in the dimensionality of the feature space that such addition creates. By adding an additional variable of potentially different levels of importance as compared to the convective weather binary variables would require tuning an extra hyper-parameter to account for this difference. It would also complicate the usability and interpretability of the weights from SVM.

C. Local Weather Variables

We consider six local weather variables: visibility (in statute miles), ceiling (in feet), instrument meteorological conditions (IMC) dummy, crosswind (in knots), tailwind (in knots), and headwind (in knots). We find values for the local weather variables from aviation system performance metrics (ASPM) database. More specifically, we pulled actual local weather information from the daily weather by hour report of ASPM Efficiency. The relevant columns in this report are visibility, ceiling, wind magnitude, wind direction and IMC indicator. We decompose wind to crosswind and tailwind/headwind according to the main runway direction at the airport. For the five airports: EWR, LGA, JFK, ATL and PHL, the main runways are selected as 22, 31, 31, 27 and 27. When the wind is tail wind, we set headwind to zero, and vice versa.

D. Busy Hour Indicator Variable

At an airport, traffic level tends to have strong correlation with hour of the day and adverse weather can be recurrent. Considering this and the focus of our paper—modeling impact of weather conditions on GDP, we employ only a dummy variable to reflect the busy level of an airport. The variable is set as 1 when the hour is between 7 am and 10 pm [6].

E. Model Specification and Estimation

To quantify how local weather and regional convection activities impact on GDPs, we construct a set of airport-specific binary logit models where the dependent variable is whether a GDP was implemented.

As described before, three types of variables are included: local weather variables, regional convection weather variables and traffic demand proxies (i.e., busy hour indicator). The description on these variables is summarized in Table I. The specification can be written as below, where V is the utility where a GDP is observed, $\beta's$ are the corresponding coefficients:

$$V = \beta_0 + \beta_1 \cdot Wx + \beta_2 \cdot IMC + \beta_3 \cdot C + \beta_4 \cdot Vis + \beta_5 \cdot TW + \beta_6 \cdot HW + \beta_7 \cdot CW + \beta_8 \cdot BH \quad (1)$$

The probability of a GDP for airport j is then quantified as:

$$Pr_j(Y_i = 1) = \frac{1}{1+exp\,(-V_i)} \quad (2)$$

TABLE I. DESCRIPTION OF EXPLANATORY VARIABLES IN THE LOGIT MODEL

Category	Explanatory variable notation	Variable description
Local weather activities	$\boldsymbol{IMC}$	Dummy variable. = 1 if instrumental meteorological condition
	$\boldsymbol{C}$	Ceiling (in 1000 ft)
	$\boldsymbol{Vis}$	Visibility (in mile). = 10 if very good visibility condition
	$\boldsymbol{TW}$	Tailwind speed (knot), = 0 if the wind is headwind
	$\boldsymbol{HW}$	Headwind speed (knot), = 0 if the wind is tailwind
	$\boldsymbol{CW}$	Crosswind speed (knot)
Traffic demand	$\boldsymbol{BH}$	Busy hour indicator, equals to 1 if local departure hour is between 7 am and 10 pm
Regional convection activities	$\boldsymbol{Wx}$	Convective weather score obtained from SVM model

IV. RESULTS

A. Parameter Selection in SVM

Selecting the parameter C in the SVM model involves using multiple criteria. First, note that the dimensionality of the feature space (200x200) is higher than the number of data samples (25310) meaning that there exists a linear binary classifier that produces perfect separation for any random designation of GDP labels. It means that the metrics based on accuracy alone can be misleading and should not be applied in order to avoid over-fitting. In addition to accuracy, precision and recall, we have used the F1-score, and the area under the receiving operating characteristic curve (AUC) in our analysis for selecting C [17]. Sample 5-fold cross-validation surface of AUC metric for EWR for a range of C and lead times is presented in Fig. 1. To estimate 5-fold cross-validation AUC performance, we partition the sample into five equal sized subsamples and each time we use four subsamples to train the model and record the AUC performance on the fifth fold. We repeat the model training and testing five times where each time the AUC performance is reported on a different subsample. The 5-fold cross-validation AUC is then the average AUC over different subsamples. The plot suggests an optimal choice of C in the range of 1e-4, as well as degrading performance with increasing lead time (as mentioned before, lead time impact is not a focus of this paper.).

A second consideration in setting C is that any near-identical weather pattern can result in GDP or no GDP given other factors. This means that the samples of the two classes can overlap heavily in the feature space, favoring use of lower values of C, and, hence models that tolerate misclassification [16]. Also, the problem is imbalanced as there is a higher number of non-GDP hours. Finally, the analysis of the spatial pattern of **w** (for instance, Fig. 2 for EWR) provides useful insights on the choice of C. Considering all these reasons, we use a combination of metrics and criteria in selecting C [18] and resulting in the values of C=1e-4 for all 5 airports.

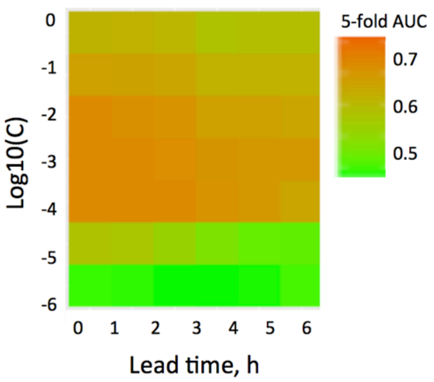

Figure 1. Average Area under ROC Curve over 5-fold Cross-validation for the EWR, as a function of hyper-parameter value C (in log scale) and GDP lead time (hour)

B. Convective Weather Impact on GDP Incidences

We evaluate the impact of convective weather on GDP incidence based on the weight vector **w** from SVM models. Spatial variability of the components of the weight vector **w** for EWR at C = 1e-4 is shown in a heatmap in Fig. 2. Higher values of w indicate a positive impact of the presence of the convective weather of intensity 3 and higher at a given location on the likelihood of GDP activation at a given airport. One can observe spatially contingent pattern centered on the EWR and spread along the main east coast corridors. For the range of parameters C from 1e-4 to 1e-5, we have observed a similar pattern for all of the five airports.

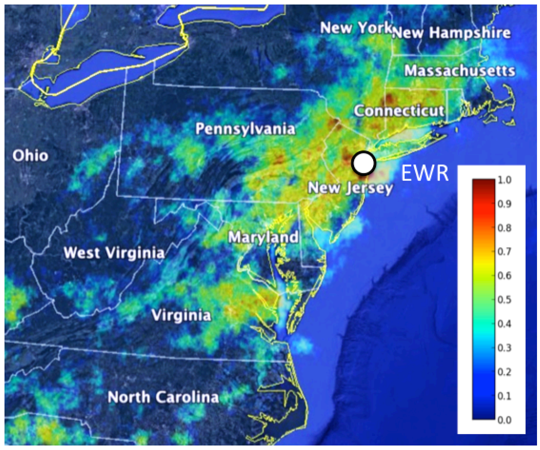

Figure 2. Map overlay of the convective weather weight vector w for the EWR model at C=1e-4

Fig. 3 illustrates the consistency of this pattern for all the airports of New York metroplex. While Fig. 3 confirms that weather closer to the airport (located at the center of the figure) is more important, it also reveals that the area of high impact is elongated along the northeast/northwest axis. This belies the assumption that weather impact depends only on the distance from the airport.

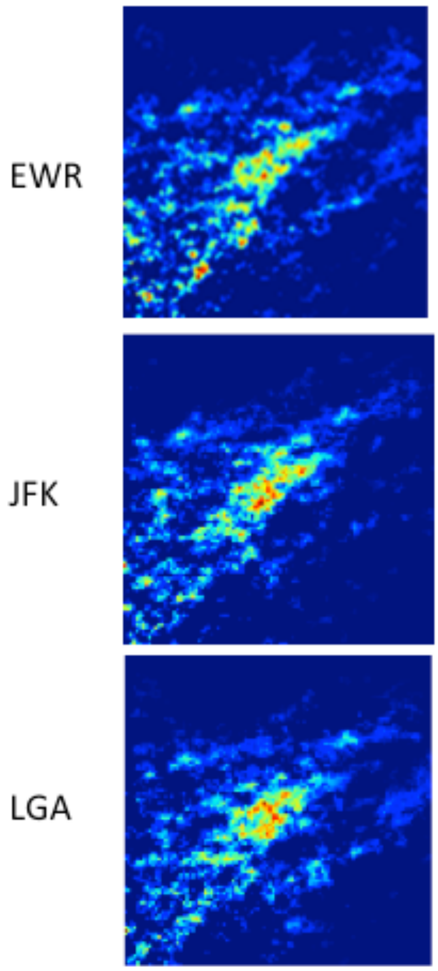

Figure 3. Weight vectors for the three airports of the NY metroplex demonstrating consistency in the spatial pattern

C. *Logistic Rregression Estimation Results*

Table II summarizes numbers of GDP and no GDP hours at the five airports and Table III reports the estimation results for the airport-specific logit models. Coefficients for regional convective weather are highly significant with positive signs for all five airports considered, meaning positive impact of convection activities on GDP incidences. For the three airports within New York area, the magnitude of the convective weather variable is comparable. ATL has the largest coefficient for convective weather variable indicating strong impact of convective weather on GDP at ATL. Also, as expected, increased ceiling and visibility reduces the likelihood of GDP incidence, while the presence of IMC increases the likelihood of GDP. All the effects of local weather variables are as expected. The tailwind seems to have mixed effects on GDP for different airports, but headwind and crosswind both increase the likelihood of GDP, except for ATL. The estimates for busy hour indicator, which is a proxy for traffic demand, are positive and significant, indicating that GDPs are more likely during higher demand period.

TABLE II. NUMBERS OF GDP AND NO GDP HOURS

	EWR	JFK	LGA	PHL	ATL
NO GDP	20305 (82.1%)	23160 (91.9%)	21588 (86.6%)	22521 (90.3%)	24355 (98.6%)
GDP	4418 (17.9%)	2055 (8.1%)	3354 (13.4%)	2426 (9.7%)	355 (1.4%)

TABLE III. ESTIMATION RESULTS FOR AIRPORT-SPECIFIC LOGIT MODELS

Variable	Airport Est./ (Std.)				
	EWR	JFK	LGA	PHL	ATL
Wx	3.893^{***}	4.511^{***}	3.369^{***}	2.773^{***}	9.660^{***}
	(0.12)	(0.13)	(0.11)	(0.10)	(0.41)
IMC	0.789^{***}	0.798^{***}	1.366^{***}	1.526^{***}	0.23
	(0.07)	(0.09)	(0.06)	(0.08)	(0.20)
C	-0.280^{***}	-0.750^{***}	-0.716^{***}	-1.627^{***}	-0.903^{***}
	(0.05)	(0.08)	(0.07)	(0.09)	(0.23)
Vis	-0.104^{***}	-0.044^{***}	-0.090^{***}	-0.178^{***}	0.061
	(0.01)	(0.01)	(0.01)	(0.01)	(0.03)
TW	0.017^{***}	0.096^{***}	0.024^{***}	-0.034^{***}	-0.058^{**}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
HW	0.046^{***}	0.065^{***}	0.056^{***}	0.044^{***}	-0.038^{*}
	(0.00)	(0.01)	(0.00)	(0.01)	(0.02)
CW	0.088^{***}	0.033^{***}	0.054^{***}	0.071^{***}	-0.047^{**}
	(0.00)	(0.01)	(0.01)	(0.01)	(0.02)
BH	2.284^{***}	2.267^{***}	2.466^{***}	3.767^{***}	2.546^{***}
	(0.07)	(0.12)	(0.08)	(0.16)	(0.37)
Constant	-0.942^{***}	-1.354^{***}	-1.746^{***}	-2.489^{***}	2.399^{***}
	(0.14)	(0.20)	(0.15)	(0.20)	(0.60)
Observations	24723	25215	24942	24947	24710

Note: $^{*}p<0.1$; $^{**}p<0.05$; $^{***}p<0.01$

We further quantify the contributions for different factors considered in the logit model. We construct a counter factual scenario in which each factor is set to the value in the dataset that would minimize the GDP probability, and use the logit model to predict the corresponding GDP probability for each record. The percentage change between the expected and current GDP probability is defined as the factor contribution. Mathematically, the percentage change is formulated as:

$$\%\Delta = \frac{\overline{Pr}\,(GDP) - E[Pr(GDP)|NewFactor]}{\overline{Pr}\,(GDP)} \tag{3}$$

Table IV and Fig. 4 show the GDP predictions and factor contributions, respectively. The first row of Table IV is the baseline predictions, which agree with the percentage of GDP incidences in our dataset, and the rest of the table is the logit model predictions based on the scenarios described above. From Fig. 4, we find that regional convection activities are by far the most important factor to the GDPs, especially for New York area airports, which agrees with [12]. Other factors, such as ceiling and wind, play important roles in GDP incidences for PHL; while the wind seems to also have a significant effect for ATL. The general conclusion is that convective weather information is indispensable to predicting GDPs, while local weather variables improve predictive performance marginally.

TABLE IV. PREDICTION OF GDP INCIDENCES

Variable	Airport GDP Prediction				
	EWR	JFK	LGA	PHL	ATL
Baseline	17.87%	8.15%	13.45%	9.72%	1.44%
Wx	0.10%	0.23%	1.29%	2.48%	0.00%
IMC	15.81%	7.11%	9.84%	6.92%	1.39%
C	16.32%	5.69%	9.62%	4.14%	1.04%
Vis	16.51%	7.75%	12.31%	7.89%	1.15%
TW	17.44%	6.43%	13.03%	5.64%	0.68%
HW	16.27%	7.00%	11.73%	8.86%	0.89%
CW	12.63%	6.87%	10.49%	7.35%	0.73%

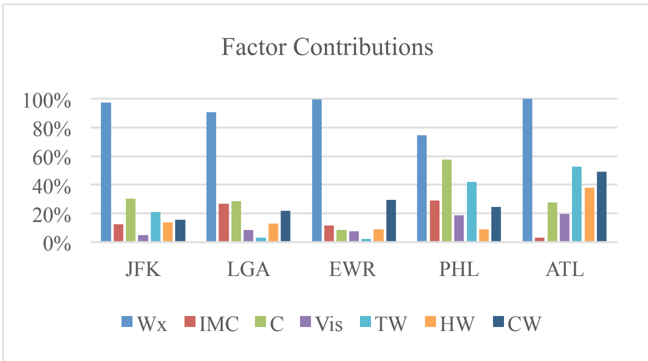

Figure 4. Factor contributions

D. *Prediction Performance of SVM and Logistc Regression*

We compare prediction performance from SVM and logistic regression in Table V. To provide further insights into the performance of the final logistic regression models, we also report the true positive and false positive rates. While there is no much difference in accuracy from SVM and logistic regression, F1-score and AUC performances are considerably better in the logistic regression. Lower scores for SVM partly reflects the tuning of the SVM model, which was based on multiple criteria specified to produce the convective index as an output of the SVM model rather than to solely maximize its predictive performance. This tuning results in an interpretable convective weather impact maps and scores, but the results are biased toward no-GDP if used as raw predictions. The AUC disparity, however, is considerably less, implying that SVM scores track GDP probabilities almost as well as the logistic regression.

TABLE V. 5-FOLD CROSS-VALIDATION RESULTS FROM SVM AND LOGISTIC REGRESSION

		EWR	JFK	LGA	PHL	ATL
SVM	Accuracy	0.83	0.92	0.86	0.90	0.99
	F1-score	0.08	0.07	0.09	0.05	0
	AUC	0.61	0.64	0.62	0.63	0.59
Logistic Regression	Accuracy	0.84	0.94	0.89	0.92	0.99
	F1-score	0.38	0.41	0.45	0.48	0.60
	AUC	0.83	0.86	0.86	0.90	0.90
	True positive rate	0.26	0.28	0.33	0.38	0.46
	False positive rate	0.16	0.06	0.11	0.06	0.01

V. CONCLUSIONS AND FUTURE RESEARCH

This paper has attempted to predict the incidence of GDPs combining spatially detailed convective weather information in the region surrounding the airport with airport local weather data. Our two-stage method combines using SVM to compute a weather score based on the location of convective weather in the airport region with a logistic regression that combines that score with local weather variables. We apply our method to five airports: EWR, JFK, LGA, PHL and ATL.

We find that the SVM, when properly tuned, provides a reasonable and spatially coherent picture of where convective weather is important. While, as expected, the area near the airport has the highest weights, the pattern is not purely a radial one. Thus the importance of convective weather depends on both its distance and direction from the airport.

From the logistic regression we learn that both regional convective weather, as captured by the weights found in the SVM, and local weather are statistically significant and have the expected signs. Convective weather is, however, the most important factor. Without it, our models predict that the vast majority of GDPs would not occur.

The model is able to predict GDP incidence with high accuracy and a low false positive rate. On the other hand, true positive rates are low. If the model predicts a GDP, it is likely to happen, but it does not predict the majority of GDPs that actually occur. This results from the unbalanced nature of the data set, in which the vast majority of hours do not have a GDP. In practice, a user of the model could choose some lower threshold than 50% in order to increase the true positive rate, albeit at the cost of also increasing the false positive one. The high AUC scores for the models indicate the combinations of true and false positive rates they can provide. These scores are in the 0.83-0.9 range, which in academic terms, correspond to a grade of B to A-.

Future research should be geared to improving the model and testing it in a more real-world setting. One basic improvement is to couple the SVM tuning and logistic regression estimation in order to obtain the optimal tuning parameters for the application. Secondly, it is desirable to incorporate air traffic demand more explicitly into the model, rather than simply using a dummy variable for busy hours. Another important step, already underway for the SVM, is to determine how to use forecast rather than realized weather in the analysis, and find how use of forecasts, which will be unavoidable in real world application, affects model performance. Together with this, we are also testing the impact of lead time on GDP prediction performance. With these enhancements, the model presented here has great potential to enable flight operators when to expect GDPs, and FAA specialists when to consider implementing them.

ACKNOWLEDGMENT

This research is funded by NASA under Award No NNX14AJ79A.

REFERENCES

[1] Liu, Y., Hansen, M., “Evaluation of performance of ground delay program,” Transportation Research Record 2400, 2013, pp. 54-64.

[2] Liu Y., Hansen, M., Gupta, G., Malik, W., Jung, Y., “Predictability impacts of airport surface automation,” Transportation Ressearch Part C 44, 2014,pp. 128–145.

[3] Hao, L., Hansen, M., “Flight predictability: concepts, metrics and impacts,” NEXTOR Final Report, 2014.

[4] Coppenbarger, R., Jung, Y., Kozon, T., Farrahi, A., Malik, W., Lee, H., Chevalley, E., and Kistler, M., “Benefit opportunities for integrated surface and airspace departure scheduling,” 35th Digital Avionics Systems Conference (DASC), Sacramento, CA, 25-29 Sep. 2016.

[5] Kang, L. Liu, Y., Hoffman R., Hansen, M., “Ground delay program decision-making based on utility maximization,” 2017, Under review.

[6] Gorripaty, S., Liu, Y., Pozdnukhov, A, Hansen, M., “Identifying similar days for air traffic management,” World Conference on Transport Researc, Shanghai, China, 10-15 July 2016.

[7] Cook, L.S., Wood, B., “A model for determining ground delay program parameters using a probabilistic forecast of stratus clearing”, 8th USA/Europe Air Traffic Management R&D Seminar, 2009.

[8] Ball, M.O., Hoffman, R., Mukherjee, A., “Ground delay program planning under uncertainty based on the ration-by-distance principle”, Transportation Science 44 (1), 2010, pp. 1-14.

[9] Mukherjee, A., Hansen, M., Grabbe, S., “Ground delay program planning under uncertainty in airport capacity,” Transporation Planning Technology 35(6), 2012, pp. 611-628.

[10] Liu, Y., Hansen, M., “Incorporating predictability into cost optimization for ground delay programs,” Transportation Science 50(1), 2015, pp.132-149.

[11] Wang, Y., Kulkarni, D., “Modeling weather impact on ground delay programs,” SAE International Journal of Aerospace 4 (2), Nov. 2011, pp. 1207–1215.

[12] Mukherjee, A., Grabbe, S., and Sridhar, B., “Predicting ground delay program at an airport based on meteorological conditions,” AIAA Aviation Technology, Integration, and Operations Conference, AIAA Paper 2014-2713, June 2014.

[13] Bloem, M., Bambos, N., “Ground delay program analytics with behavioral cloning and inverse reinforcement learning,” Journal of Aerospace Information Systems 12(3), 2015, pp. 299-313.

[14] Kuhn, K.D., “A methodology for indetifying similar days in air traffic flow management initiative planning,” Transportation Research Part C 69, 2016, pp. 1-15.

[15] Dixon, M., Wiener, G., “TITAN: Thunderstorm identification, tracking, analysis, and nowcasting—a radar-based methodology,” Journal of Atmospheric and Oceanic Technology 10(6), 1993, pp. 785-797.

[16] Cortes, C., Vapnik, V., "Support vector networks". Machine Learning 20, 1995,pp. 273-297.

[17] Wilks, D, Chapter 8: Forecast Verification, Statistical Methods in the Atmospheric Sciences (3rd ed.). Elsevier, 2011.

[18] Pozdnoukhov, A., Matasci, G., Kanevski, M., and Purves, R.S., "Spatio-temporal avalanche forecasting with Support Vector Machines," National Hazards Earth System Sciences 11, 2011, pp. 367-382.

AUTHOR BIOGRAPHY

Yi Liu is a research scientist at Amazon. Her research interests are in air traffic management, air transportation system performance, airline economics, machine learning and data-driven decision making. Dr. Liu received her Ph.D. in transportation in 2016 from University of California, Berkeley.

Mark Hansen is a Professor of Civil and Environmental Engineering at UC Berkeley, and a co-director of the National Center of Excellence for Aviation Operations Research.

Danqing Zhang is a PhD student in Civil & Environmental Engineering at University of California, Berkeley. Her research interest is in developing machine learning models for spatial data, with applications in spatial-embedded social network, human mobility and activity analysis and environmental geo-statistics.

Yulin Liu is a Ph.D. student of Civil and Environmental Engineering at the University of California, Berkeley. His research interests include air traffic flow management, trajectory optimization, and machine learning algorithms. He received his B.S. degree in Civil Engineering in 2015 from Tsinghua University.

Alexei Pozdnoukhov is an assistant professor with the University of California, Berkeley. His research focuses on the development of data analysis methodologies and machine learning methods for spatial data, with applications in urban infrastructure and transportation systems planning and operation.