

Predictive Models of Duration of Ground Delay Programs in New York Area Airports

Deepak Kulkarni

1. Introduction

Initially planned GDP duration often turns out to be an underestimate or an overestimate of the actual GDP duration. This, in turn, results in avoidable airborne or ground delays in the system. Therefore, better models of actual duration have the potential of reducing delays in the system. The overall objective of this study is to develop such models based on logs of GDPs. In a previous report, we described descriptive models of Ground Delay Programs. These models were defined in terms of initial planned duration and in terms of categorical variables. These descriptive models are good at characterizing the historical errors in planned GDP durations. This paper focuses on developing predictive models of GDP duration.

Traffic Management Initiatives (TMI) are logged by Air Traffic Control facilities with The National Traffic Management Log (NTML) which is a single system for automated recoding, coordination, and distribution of relevant information about TMIs throughout the National Airspace System. (Brickman, 2004; Yuditsky, 2007) We use 2008-2009 GDP data from the NTML database for the study reported in this paper. NTML information about a GDP includes the initial specification, possibly one or more revisions, and the cancellation. In the next section, we describe general characteristics of Ground Delay Programs. In the third section, we develop models of actual duration. In the fourth section, we compare predictive performance of these models. The final section is a conclusion.

2. Characterization of Ground Delay Programs

A GDP can be characterized by a number of important factors including the following:

Initial Planned Duration: Duration of the GDP specified in the initial announcement of the GDP.

Overall Planned Duration: Overall duration for which the GDP was planned.

Actual Duration: Actual duration of the GDP.

Lead Time: The duration between initial time of announcement of the GDP and the time of the start of the GDP.

Early Cancel Time: The duration between the planned time for ending the GDP and the actual time when it ended.

Affected Flights: Number of Flights affected by the GDP.

Start Time Of Day: Time of the day when the GDP starts.

Decision Time (Model Time) Of Day: Time of the day when the GDP planners make

decisions about initial parameters of the GDP
GDP Cause: Cause of the GDP

3. Actual Duration Models

In this section, we will examine three different models of GDP duration: (1) models in terms of GDP start time alone, (2) models in terms of GDP start time and the season, and (3) models in terms of GDP start time and the GDP cause.

3.1 Models in Terms of GDP Start Time

GDP start time and GDP Decision Time are represented in GMT time in NTML database. For the purpose of model development, we represent these in terms of number of minutes from midnight. Table 3 lists coefficients of correlation of actual duration with GDP start time and GDP Decision Time. Figures 2 and 3 show corresponding scatter-plots in the case of LGA. Both GDP Decision Time and GDP start time have a strong correlation with GDP duration at these airports. However, correlation coefficient is higher for GDP start time than for GDP Decision Time. Correlation coefficient in the case of multiple linear regression with both GDP start time and GDP Decision Time is about the same as the correlation coefficient with GDP start time.

Airport	EWR	LGA	JFK
GDP Model Time	-.61	-.87	-.58
GDP Start Time	-.72	-.90	-.76
Both	-.72	-.90	-.76

Table 3. Correlation of Actual Duration with Various Parameters

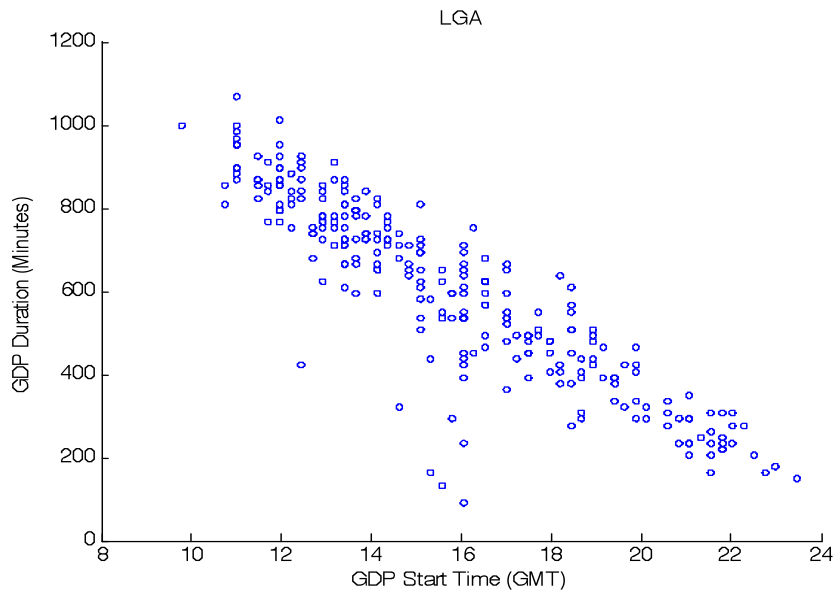


Figure 2. Actual Duration vs. GDP Start Time at LGA

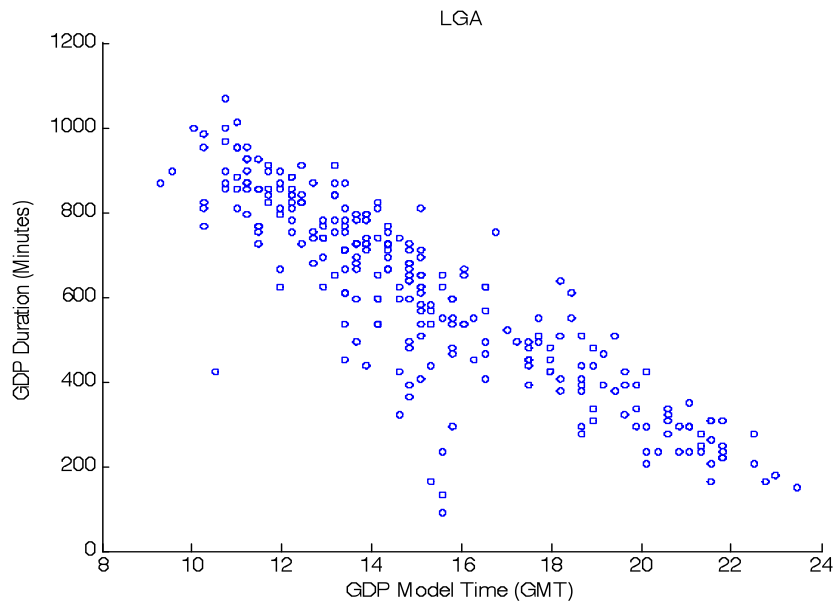


Figure 3. Actual Duration vs. GDP Decision Time at LGA

Therefore, a linear model in terms GDP start time would be almost as good as a linear model in terms of both the start time and the decision Time. Table 4 shows linear models of actual duration in terms of GDP start time for New York airports.

Airport	Model in Terms of GDP Start Time
EWR	$-1.2x + 1768$
LGA	$-x + 1576$
JFK	$-.9x + 1431$
All NY	$-1.1x + 1617$

Table 4. Models of Actual Duration in Terms of GDP start time

3.2 Models in Terms of GDP Start Time and GDP Cause

One may expect that GDP duration would be affected by the cause of a GDP. Under ANOVA test, GDP cause is relevant to GDP duration for JFK, LGA, EWR as well as for all NY airports as a group. Figure 4 shows box-plots of GDP Duration at New York area airports for different GDP causes. X-axis in Figure 4 represents the different GDP causes: wind, low ceilings, thunderstorms, non-weather causes, rain, low visibility and snow. In the box plot, the bottom and top of the box are the 25th and 75th percentile, and the band near the middle of the box is the median. The box-plots show the variation in GDP Duration for different GDP causes. As the average duration of disruptive events as well as predictability of duration is different for different GDP causes, one would expect that models of GDP duration in terms of GDP start time could be different if we develop these using only the data corresponding to a particular GDP cause.

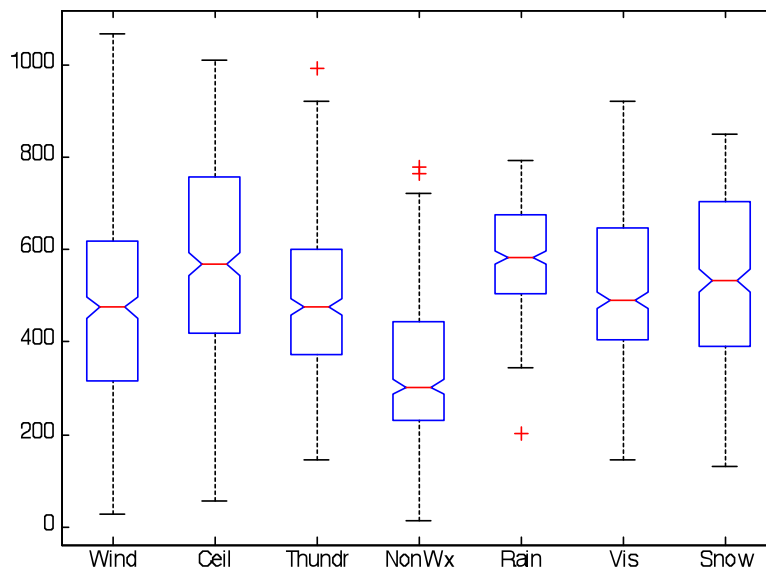


Figure 4. Impact of GDP Cause on New York area GDP Duration

	EWR	LGA	JFK	All NY
Wind	$-1.2x+1794$	$-x+1613$	$-x+1492$	$-1.1 x +1634$
Low Ceiling	$-1.4x+2033$	$-x+1572$	$-0.8x+1269$	$-1.1 x + 1685$
Non-weather	$-0.6x+1035$	$-0.9 x+1344$	$-0.8x+1198$	$-0.8 x + 1257$
Thunderstorm	$-x+1585$	$-0.9x+1445$	$-0.8x+1302$	$-0.9 x +1473$

Table 6. GDP Duration Models in Terms of GDP Start Time and Selected GDP Causes

Table 6 shows such models. Depending on the weather cause, the slope and intercept terms in the models are different. Models for rain, snow and low visibility are not shown as the number of cases with these GDP causes was small.

3.3 Interpretation of the Intercept Term in the Models

Most of the models we identified in the previous section have an intercept close to -1. Therefore, the intercept term in these models is close to the sum of Start Time and GDP Duration Time. Latter corresponds to GDP End Time. Thus, the intercept term corresponds to GDP End Time. This is owing to the low correlation between GDP End Time and GDP Start Time. The scatter plots in Figure 7 illustrate this in the case of EWR. Correlation coefficient of GDP End Time with GDP start time is -.17, -.06 and -.09 for EWR, LGA and JFK. Similarly, correlation coefficient of GDP End Time with GDP Decision Time is -.16, -.08 and -.08 for EWR, LGA and JFK.

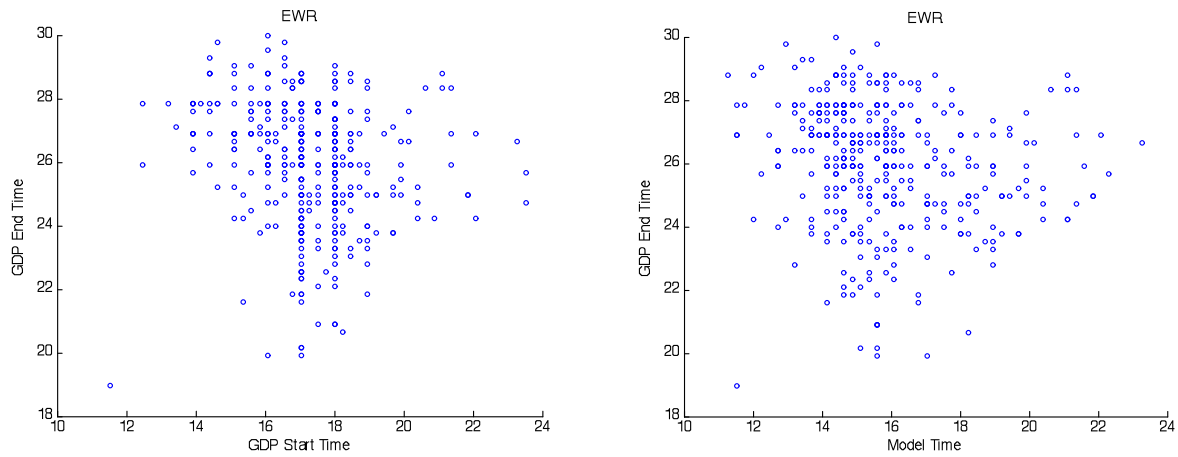


Table 7. Scatter Plots of GDP End Time at EWR

4. Comparison of Performance of Models

	EWR	LGA	JFK
Model in Terms of Start Time	117	97	92
Model in Terms of Start Time and Season	114	97	91
Model in Terms of Start Time and Cause	108	94	88
Initial Planned Duration	87	92	92

Table 7. Standard Deviation in Error of Different Predictors

Table 7 shows standard deviation in error in predicting GDP duration for different predictors. First three rows correspond to the three models we discussed in the previous section. The last row corresponds to the initial planned duration. A model of GDP duration that is a better predictor of GDP duration as compared to the initial planned duration can be used to reduce avoidable delays in the system. Models at LGA and JFK in Table 7 have similar performance to initial planned duration whereas those at EWR are worse than initial planned duration.

Standard deviation of error is a good statistical measure of predictive ability of models. However, from practical perspective, one may want to characterize error in terms that can be considered easily in the decision-making process. Average overestimate and average underestimate of a predictor of GDP duration are two such measures with different consequences on operations. However, one model may produce smaller overestimates but higher underestimates than another model. In such cases, it is not obvious which model is better. To enable easy comparison between such models, we can add a bias term to the second model that increases its average overestimate of the model and decreases its average underestimate. We use this technique so that both the model and initial planned duration have

the same average underestimate, so that the average overestimate can be used to judge if a model is better than the initial planned duration.

GDP Cause		EWR		LGA		JFK	
		Std	OE	Std	OE	Std	OE
Ceiling	Initial Planned Duration	88	52	64	93	106	125
	Model Prediction	109	74	75	104	100	123
Non-weather	Initial Planned Duration	55	56	115	110	74	80
	Model Prediction	124	192	111	112	96	140
Thunderstorm	Initial Planned Duration	84	75	95	118	110	81
	Model Prediction	98	72	75	83	77	60
Wind	Initial Planned Duration	90	83	96	105	73	67
	Model Prediction	107	101	97	104	84	88

Table 8. Standard Deviation in Prediction Error and Overestimates for GDP Causes

Tables 8 and 9 shows standard deviation in prediction error and overestimates for various GDP causes and for various seasons. The cases where model prediction has lower standard deviation of error or average overestimate as compared to the initial planned duration is shown in bold. For example, when GDP cause is Thunderstorm at JFK, the model prediction has standard deviation of just 77 whereas initial planned duration has standard deviation of 110.

		EW R		LG A		JF K		NY Group	
		Std	OE	Std	OE	Std	OE	Std	OE
Fall	Initial Planned Duration	84	58	88	100	116	92	97	80
	Model Prediction	131	121	91	103	138	116	125	112
Spring	Initial Planned Duration	82	70	115	103	92	100	90	90
	Model Prediction	110	110	92	95	97	121	106	119
Summer	Initial Planned Duration	80	62	93	96	81	75	81	76
	Model Prediction	105	86	63	77	71	90	84	84
Winter	Initial Planned Duration	99	87	109	122	96	87	103	98
	Model Prediction	115	94	127	118	74	63	111	94

Table 9. Standard Deviation in Prediction Error and Overestimates for Different Seasons

5. Conclusion

GDP is an important traffic flow initiative that is used by Traffic Flow Managers to reduce the impact of disruptions. Inaccurate estimation of actual duration results in significant avoidable delays in the system. Therefore, better models of actual duration have a potential of reducing delays in the system. We use 2008-2009 GDP data from the NTML database to develop such models. Actual duration was found to have a strong correlation with GDP start time and GDP decision Time. Furthermore, GDP cause and season are important factors influencing the GDP duration. Therefore, the actual duration can be modeled in terms of GDP start time separately for each GDP cause. Overall, such models perform only slightly worse predictors of GDP duration than the initial planned duration. However, for certain GDP causes and for specific seasons, such models can be better predictors as compared to the initial planned duration.

References

Brickman, B. and Yuditsky, T., "Improving the Usability of an Automated Tool for the Recording, Coordination and Communication of Traffic Management Initiatives," Proceedings of the Human Factors and Ergonomics Society 48th Annual Meeting, 2004, pp 46-50.

Cook, L., "Translating Weather Information (Convective and Non-convective) Into TFM Constraints," NASA NRA Report, August 2010.

FAA, Order JO 7110.65S. Air Traffic Control, February 2008.

FAA, Order 7210.3W. Facility Operation and Administration, February 2010.

Rios, J., "Aggregate Statistics of National Traffic Management Initiatives," ATIO Conference, Toronto, Canada, 2010.

Yuditsky, T. and Brickman, B., "Benefits Analysis for the National Traffic Management Log," Technical Report, Federal Aviation Administration, July 2007.