AN OPTIMIZATION MODEL FOR SAMPLE DAY SELECTION IN NAS-WIDE MODELING STUDIES

Feng Cheng, John Gulding, Bryan Baszczewski, Ruth Galaviz
Federal Aviation Administration, Washington, DC

Abstract

Future flight Schedules are generated based on air traffic demand forecast for the purpose of aviation planning and performance analysis studies. A selection process needs to be designed and implemented by sampling historical operational data for each fiscal quarter and choosing representative days that best reflect seasonality in terms of a given set of performance metrics.

We propose an optimization based solution method for the sample day selection problem, which is formulated as a Mixed Integer Program (MIP). The objective of the MIP is to minimize the weighted difference between the true population and the sample to be selected in terms of the defined metrics subject to a set of constraints including the sample size limit, coverage requirements and other desired properties. An efficient solution algorithm has been implemented using the CPLEX MIP solver.

Experiments have been conducted with a wide range of flight data from the recent years. The results from the MIP method provided robust solutions for the sample day selection problem. It is also shown that the method is quite flexible to incorporate additional constraints based on expert knowledge.

Introduction

FAA's Air Traffic Organization produces Future Schedules for the purpose of aviation planning and performance analysis studies [1-3]. A major design criterion for the ATO future schedules is to reflect trends in the National Airspace System (NAS) that occur by season over the course of the year. To accomplish this goal, a selection process needs to be designed and implemented by sampling historical operational data for each fiscal quarter and choosing representative days that best reflect seasonality in terms of selected performance metrics.

For practical purposes, the existing implementation of the selection process focused on analyzing 4 peak seasonal days and 4 off-peak seasonal days. A uniform busy day, based on a 90% Day Score is calculated by computing the percent difference between the 90% day traffic volume and the traffic volume of any given day. The selection of the peak days is based on the 90% Day Score with the days closest to the 90% day traffic volume chosen. The off-peak days are selected in such a way that the overall difference (typically arithmetic averages or weighted averages) between the selected sample days and the entire population of a given quarter are minimized in terms of the defined metrics. However, the existing procedure is largely a manual process performed mostly through inspection and not though an optimized computer program. If there are a large number of dimensions to be measured in the defined metrics, the process becomes computationally intensive and even intractable.

Random sampling technique has also been used to generate an improved and even coverage across the distribution of the selected metrics with a larger sample size of 36 days. However, the random sampling technique provides no guarantee to produce the "best" selection. Without using advanced optimization techniques, the "best" selection would require a complete enumeration of all possible selections which is known to be computationally intractable.

To overcome the computational difficulty that is inherent in the existing process we propose an optimization based solution method for the sample day selection problem. We formulate the problem as a Mixed Integer Program with the objective to minimize the weighted difference between the true population and the sample in terms of the defined metrics and subject to a number of constraints. The variables are binary variables representing the days in a given fiscal year. The Mixed Integer Program then minimizes the difference between the true average and the sample average for all 4 parameters over all airport and center facilities.

The computation results showed that the MIP solution outperformed the other two methods in terms of the objective function defined by the MIP formulation. With more sample days selected, the performance of MIP solutions are consistently better. Based on the experiments that we have conducted, the MIP method proves to be a robust approach for the sample day selection problem. It is also quite flexible in the sense that any expert domain knowledge can be incorporated into the formulation as long as it can be expressed mathematically.

Methodology

We formulate the problem as a Mixed-Integer Program with the objective to minimize the weighted difference between the true population and the to-be-selected sample in terms of the defined metrics and subject to a number of constraints including:

- The sample size (number of days to be selected) constraint
- Coverage requirements
    - same number of days per quarter or month
    - minimum number of Peak or off-peak days
- Inclusion of the 90% planning days given by the current FAA process
- Any other ad-hoc conditions

The variables are binary variables representing the days in a given fiscal year. The value of the variable is either 1 if the corresponding day is included in the sample or 0 if it is not. The data required in the problem include OPSNET counts and ASPM delay for the Core 30 airports and ATALAB counts and flight hours for both the 20 CONUS, Alaska and oceanic centers. The mixed integer program then minimizes the difference between the true average and the sample average for all 4 parameters over all airport and center facilities.

To define the formulation more precisely, we introduce the following notation.

$$d^i = \frac{1}{\sum_{k \in K_{OEP}} \overline{d}^i_k} \sum_{k \in K_{OEP}} \left| \overline{d}^i_k - \overline{d}^i_{k,n} \right|, \qquad (1)$$

where

- d^i = weighted average of absolute differences by airport/center for metric *i*,
- $d^i_{k,t}$ = performance metric *i* for airport/center *k* on day *t*, $k \in K_{OEP}$, $t \in T, |T| = 365$
- x_t = binary variable with x_t=1 if day *t* is included in the sample, x_t=0 otherwise
- U = set of pre-selected 90% uniform days,
- Q_m = set of peak days selected in quarter *m*,
- N = sample size,
- N_m = maximum number of peak days selected in quarter *m*,
- K_{OEP} = set of airports,
- T = set of days in a fiscal year, $T_n \subset T$
- T_n = set of sample days, $T_n = \{x_t \mid x_t = 1\}$,
- $\overline{d}^i_k = \sum_{t \in T} d^i_{k,t} / |T|$,
- $\overline{d}^i_{k,n} = \sum_{t \in T_n} d^i_{k,t} / |T_n|$

With the above notations defined, we can write the problem formulation as follows.

$$\min \sum_{i=1}^{4} d^i \qquad (2)$$

subject to:

$$|T_n| = N,$$

$$x_t = 1, \quad \forall t \in U,$$

$$\sum_{t \in Q_m} x_t \le N_m.$$

Note that mathematically this is a linearly constrained minimization problem, with objective function being a sum of absolute values of linear functions. However, the absolute values $|\overline{d}_k^i - \overline{d}_{k,n}^i|$ in the objective function can be replaced by breaking down each item into a combination of positive differences and negative differences [4]. Thus the model above can be converted to a mixed integer linear programming problem as defined below.

$$\min \sum_{i=1}^{4} \sum_{k \in K_{OEP}} \left(d_k^{i+} + d_k^{i-} \right) \qquad (3)$$

subject to:

$$\sum_{t \in T} x_t = N,$$

$$\overline{d}_k^i - \sum_{t \in T} d_{k,t}^i x_t / N + d_k^{i+} \geq 0, \quad \forall k \in K_{OEP}$$

$$\overline{d}_k^i - \sum_{t \in T} d_{k,t}^i x_t / N - d_k^{i-} \leq 0, \quad \forall k \in K_{OEP}$$

$$x_t = 1, \quad \forall t \in U,$$

$$\sum_{t \in Q_m} x_t \leq N_m.$$

Furthermore, additional terms in the objective function and more constraints can be added to the base formulation as desired.

Experiments and Results

We have tested the MIP model with data from fiscal years 2008-2010 with target sample sizes of 8, 12 and 16. The four performance metrics used for the analysis are the daily flight counts for each of the 30 Core airports, the average delays associated with the flights at these airports, the daily flight counts for each of the Air Route Traffic Control Centers, and the average flight hours for the flights tracked by these centers. With the daily records of these metrics provided as input to the MIP model, the goal of the optimization is to pick a subset (sample) of the days such that the average values of the selected sample will be closest to the average values of the entire population for any given fiscal year. In the following, we report the results under three scenarios.

The Base Scenario (S1) is defined exactly as the formulation given in the previous section where the pre-selected 90% uniform days are fixed as part of the sample. Specifically, the 4 peak days determined by the scoring method for the 90% uniform busy days will be fixed in the MIP model. The rest of the days in the sample will be picked by the optimization algorithm. Table 1 shows the selected days for the cases of 8, 12 and 16 sample days, and Table 2 provides the corresponding performance metrics in comparison with the other methods.

The MIP programs were coded in OPL scripts using IBM/ILOG CPLEX Optimization Studio [5]. Our testing results showed that the computation to obtain the optimal solution was reasonably fast. It typically took a few minutes to solve the 8-day problem and up to a few hours for the 16-day problem. Furthermore, a decomposition technique has been investigated by exploring the structure of the MIP formulation to reduce the computation time substantially.

Table 1. Results for Base Scenario: Sample Days

Uniform Busy Days		MIP Solutions					
4 Peak Days		8 Sample Days		12 Sample Days		16 Sample Days	
11/20/09	Friday	10/17/09	Saturday	10/6/09	Tuesday	10/27/09	Tuesday
3/25/10	Thursday	11/20/09	Friday	10/17/09	Saturday	10/31/09	Saturday
5/6/10	Thursday	1/10/10	Sunday	11/20/09	Friday	11/20/09	Friday
7/22/10	Thursday	3/25/10	Thursday	1/10/10	Sunday	12/24/09	Thursday
		5/6/10	Thursday	3/9/10	Tuesday	1/9/10	Saturday
		6/13/10	Sunday	3/25/10	Thursday	1/25/10	Monday
		7/22/10	Thursday	5/6/10	Thursday	2/3/10	Wednesday
		8/21/10	Saturday	5/18/10	Tuesday	3/25/10	Thursday
				6/5/10	Saturday	4/18/10	Sunday
				7/3/10	Saturday	5/6/10	Thursday
				7/13/10	Tuesday	5/19/10	Wednesday
				7/22/10	Thursday	6/21/10	Monday
						7/13/10	Tuesday
						7/22/10	Thursday
						8/9/10	Monday
						8/29/10	Sunday

Table 2. Results for Base Scenario: Comparison of Performance Measures

Year	Method	# of Days	Airport Counts	% diff1	% diff2	Delays	% diff2	Center Counts	% diff1	% diff2	Flight Hours	% diff2
FY10	Actual	365	34,836			259,963		105,575			67,355	
FY10	MIP	8	34,843	0.74	1.18	265,000	7.31	105,753	1.61	0.97	67,795	1.12
FY10	MIP	12	34,885	0.72	0.87	258,463	2.82	105,256	2.79	0.47	67,391	0.78
FY10	MIP	16	34,835	0.16	0.57	260,404	2.01	105,459	1.15	0.40	67,415	0.59
FY09	Actual	365	34,880			294,394		109,008			69,020	
FY09	MSM	8	34,268	2.26	2.05	288,725	18.44	106,836	4.21	2.46	67,221	3.10
FY09	RSM	36	35,124	0.40	0.83	302,244	4.56	109,341	0.24	0.44	69,217	0.61
FY09	MIP	8	34,473	1.61	1.63	295,913	6.90	109,475	0.77	0.82	69,530	1.04
FY09	MIP	12	34,656	0.34	1.06	295,721	2.92	109,514	0.57	0.66	68,997	0.53
FY09	MIP	16	34,774	0.20	0.79	294,583	2.04	108,993	0.52	0.58	69,141	0.51
FY08	Actual	366	36,689			376,904		116,227			73,556	
FY08	MSM	8	36,633	2.34	1.70	338,153	16.30	116,743	3.73	1.01	73,859	1.39
FY08	RSM	36	36,710	0.18	0.55	367,825	5.41	115,207	0.44	0.92	72,882	1.02
FY08	MIP	8	36,302	2.10	1.88	385,939	6.77	116,202	4.89	1.53	73,756	1.66
FY08	MIP	12	36,494	0.58	0.94	377,857	2.80	116,761	1.91	0.55	73,805	0.60
FY08	MIP	16	36,702	0.25	0.62	378,062	1.21	116,260	0.85	0.49	73,601	0.53

Table 1 shows the selected days obtained for the model for the base scenario with the sample size of 8, 12 or 16. In Table 2, the results from the optimization solutions are compared against the corresponding results using either the existing manual scoring method (MSM) with 8 sample days or a random sampling method (RSM) with 36 sample days. The columns labeled *%diff1* in Table 2 are the average percentage errors, and the columns labeled *%diff2* are the average absolute value percentage errors as defined in (1). The comparison shows the MIP solution generally outperformed the other two methods in terms of the average performance metrics which are closer to the actual averages. Particularly, with more sample days selected, the performance of MIP solutions is also consistently better.

Figures 1-4 provide additional details to show the differences in the results of MIP, MSM, and RSM (labeled MIP 12, MSM and RSM respectively), where the percentage performance differences between the sample averages and the averages of the actual data are plotted at the airport/center level. The FY09 actual averages by airport/center are also plotted in the same figures against the second y-axis.

The objective function of the MIP method implies that higher weights are given to the airports with higher numbers of flight counts (or delays, etc.). As a result, the percentage performance differences for those airports will be minimized with a higher priority relatively to that of airports with small numbers of flight counts (or delays, etc.) For example this can be seen in Fig.1, in which the airports near the right-hand side of the figure typically have smaller percentage errors. Some percentage errors from the MIP solution for the airports near the left-hand side of the figure may appear to be larger than that of MSM or RSM. However, they carry less weight given the smaller number of flight counts at those airports; therefore, do not contribute to the objective function as much as those airports near the right-hand side of the figure. Another observation from Figures 1-4 is that either MSM or RSM method produced a reasonably good fit in terms of one performance measure but not for other measures, while the MIP results performed well and more balanced in all four measures. This demonstrates the powerfulness of the optimization technique that is generally superior to a manual scoring method or random sampling technique.

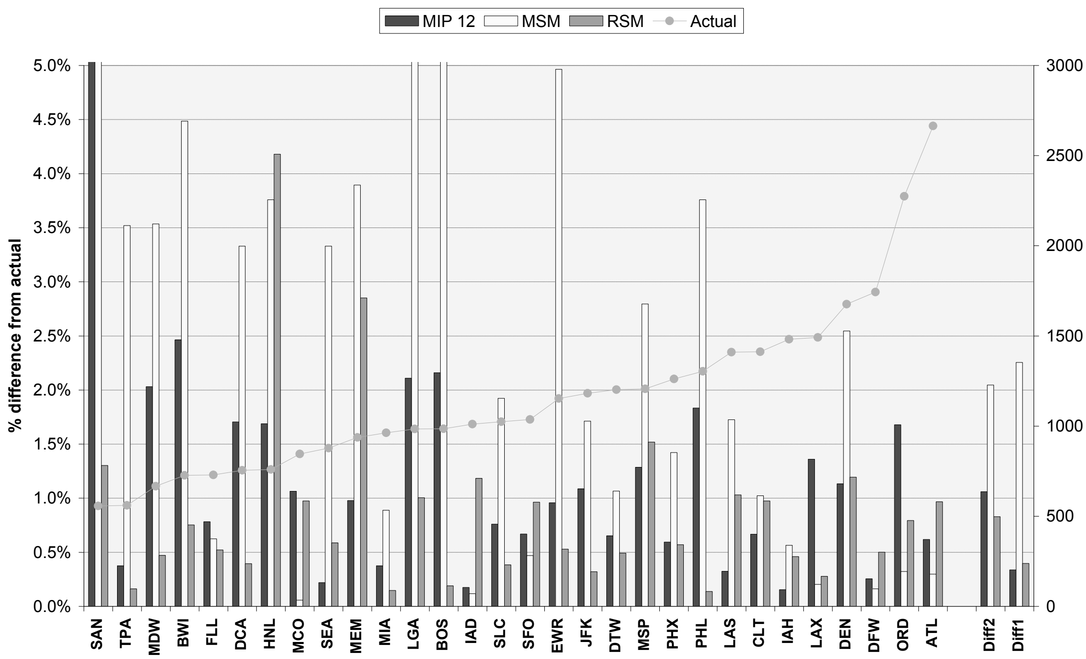

Figure 1. Flight Hours by Center for FY09

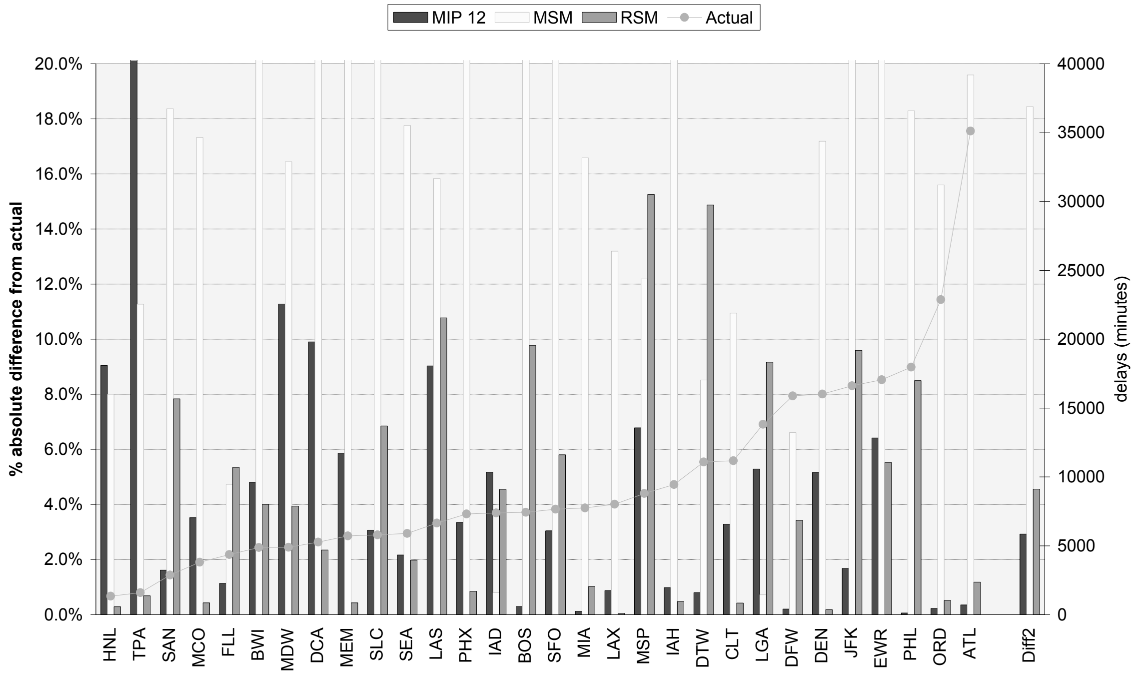

Figure 2. Airport Delays for FY09

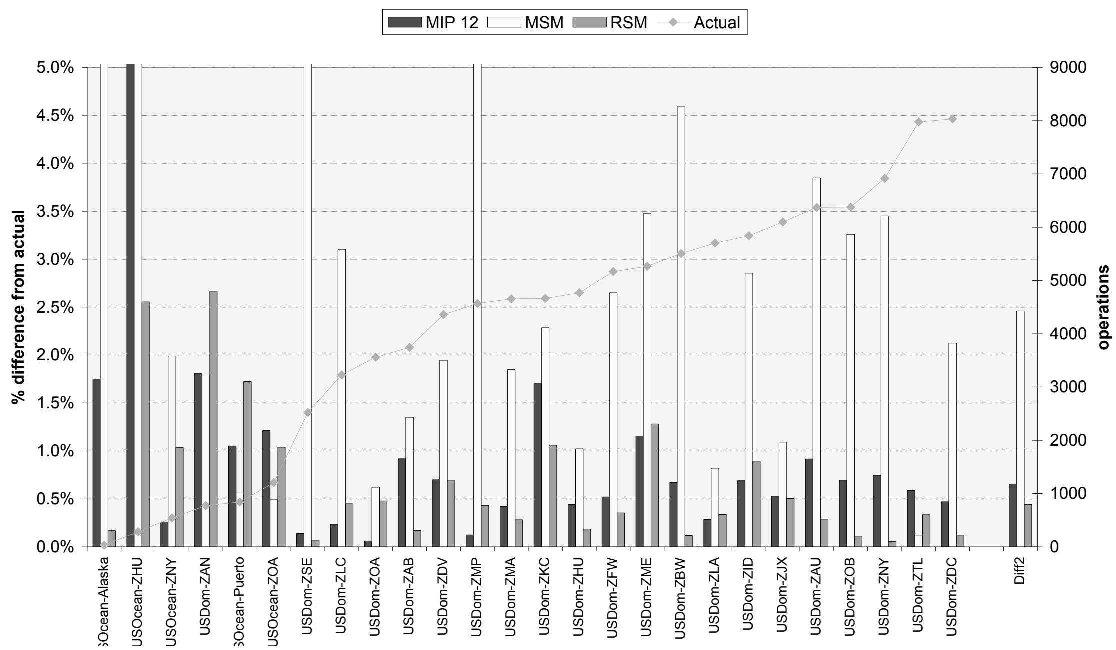

Figure 3. Flight Counts by Center for FY09

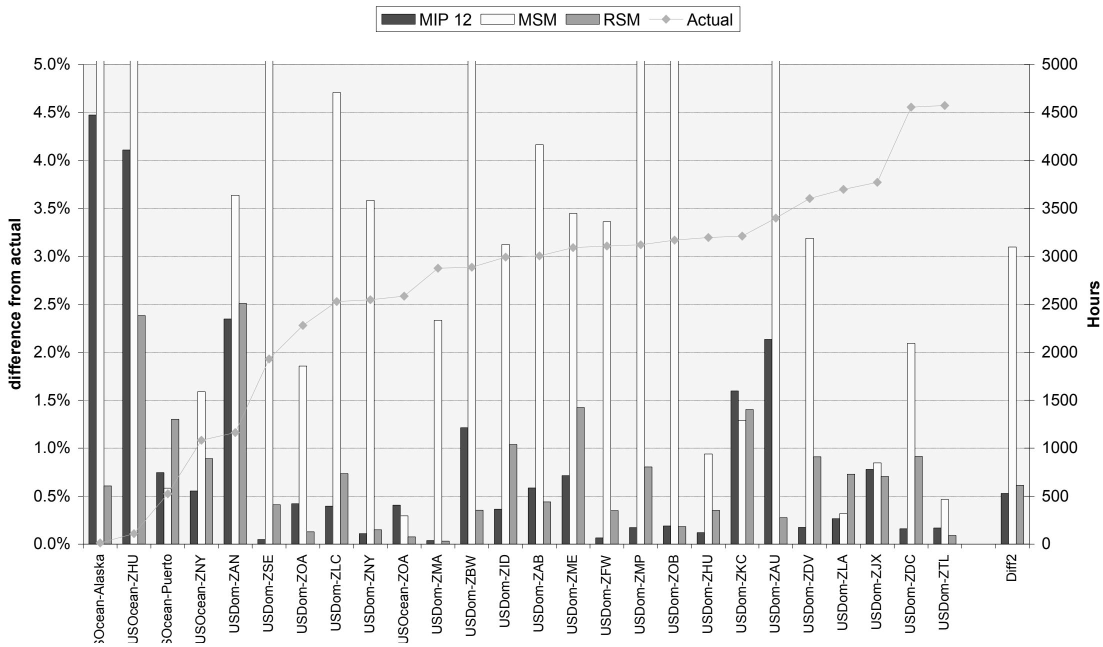

Figure 4. Flight Hours by Center for FY09

Two additional testing scenarios are created to demonstrate the flexibility of the MIP modeling capability. **The Unconstrained Scenario (S2)** is the second scenario in which the constraints of using the pre-selected 90 uniform days and the restriction on the number of days that can be selected per quarter are removed. The third scenario, **the Two-peak Day Scenario (S3),** is same as the Base Scenario except that only two peak days (the spring and summer peak days) are fixed. These two scenarios are designed as part of the "what-if" experiments that would enable analysts to explore the potential of the optimization solutions. The second scenario allows the model to pick the sample days based on the objective function only and free of any constraints. The third one demonstrates the flexibility of the optimization solution by making changes to the constraints. Table 3 shows the comparison of the results from the three scenarios. Clearly, by removing or relaxing some or all constraints, the optimization model is able to obtain solutions with improved performance metrics overall.

Table 3. Comparison for Scenarios S1-S3

Year	Scenario	# of Days	Airport Counts	% diff1	% diff2	Delays	% diff2	Center Counts	% diff1	% diff2	Flight Hours	% diff2
FY10	Actual	365	34,836			259,963		105,575			67,355	
FY10	S1	8	34,843	0.02	1.18	265,000	7.31	105,753	0.17	0.97	67,795	0.17
FY10	S1	12	34,885	0.14	0.87	258,463	2.82	105,256	0.30	0.47	67,391	0.30
FY10	S2	12	34,824	0.04	0.58	259,100	2.08	105,550	0.02	0.38	67,196	0.57
FY10	S3	12	34,871	0.10	0.64	257,335	2.13	105,723	0.14	0.58	67,687	0.69

Conclusion

Based on the experiments that we have conducted, the MIP method proves to be a robust approach for solving the sample day selection problem. The MIP formulation captured all of the practical requirements for the sample day selection process. The optimization approach produced the sample days that best matched with the actual performance measures at the daily average level. The results obtained from the MIP method provided the basis for the recommendations on sample days for fiscal years 2008-2010, which had been made available to support FAA's current NAS system-wide modeling efforts and initiatives.

The MIP method is also quite flexible in the sense that additional knowledge or requirements from subject matter experts can be incorporated into the formulation to facilitate future enhancements of the NAS modeling capabilities.

References

[1] Barnhart, C., A Cohn. 2004. Airline Schedule Planning: Accomplishments and Opportunities. Manufacturing & Service Operations Management, 6(1), pp.3-22.

[2] U.S. Department of Transportation, Federal Aviation Administration. 2010. Federal Aviation Administration Aerospace Forecast, Fiscal Years 2010-2030.

[3] Chen, X., Gulding, J. Assessment of System Constraints for Producing Constrained Feasible Schedules. 2008 ATRS Conference.

[4] Vanderbei, R. Linear Programming: Foundations and Extensions. Springer-Verlag, 2001.

[5] IBM ILOG CPLEX Optimization Studio V12.2, 2010.

2011 Integrated Communications Navigation and Surveillance (ICNS) Conference

May 10-12, 2011