AIAA Guidance, Navigation and Control Conference and Exhibit
20 - 23 August 2007, Hilton Head, South Carolina

AIAA 2007-6449

Characterization of Days Based On Analysis of National Airspace System Performance Metrics

Gano B. Chatterji* and Bassam Musaffar†
University of California Santa Cruz, Moffett Field, CA 94035-1000

Motivated by the need to select days with distinct traffic characteristics for evaluating novel air traffic management concepts and validating simulations, 517 days of delay data from the Federal Aviation Administration's Air Traffic Operations Network database are analyzed. The daily total time delay in minutes is then used as a distance metric within the K-Means algorithm to organize the 517 days into ten groups. Convergence characteristics of the K-Means algorithm and summary statistics of the groups are presented. The single-metric K-Means algorithm is then extended to create a multiple-metric K-Means classifier. Two examples of multiple-metric classification are presented using two different sets of metrics to partition the 522 days, the original 517 days and five special days, of data into groups. Results show that this multiple-metric classifier is useful for creating sets of days that represent a variety of traffic conditions.

I. Introduction

This paper is motivated by the need for selecting days with particular air traffic flow patterns and operational characteristics, as encapsulated in the performance metrics, for validating simulation models and evaluating next generation air traffic system concepts. Evaluation of system-wide impacts in terms of cost and benefits with one or two days of data, or with several days of data with similar traffic conditions, is of limited utility. Such evaluations therefore, have to be conducted with a set containing days with distinct characteristics. In order to balance the effort required against the quality of results achieved for these types of simulations and evaluations, a small set of days that covers all the possible traffic conditions is desirable. The multiple-metric classification method proposed in this paper makes it possible to create such a set of days.

Prior effort on the selection of days for validating simulation models is described in Refs. 1 and 2. Reference 1 contains a detailed description of the data and the procedure used, while Ref. 2 is a summary of the same. The approach consisted of using the K-Means algorithm, first proposed in Ref. 3, to partition the set of days into six significant groups, each with at least 2% of the days, and one outlier group for days that could not be assigned to the six significant groups. Each group was separated from others in terms of a single Euclidean distance metric composed of the eight chosen metrics. Based on analysis of the metrics associated with each of the six significant groups, they concluded that Ground Delay Program (GDP) minutes and total operations count, a measure of traffic-volume, were the primary determinants of group membership. Threshold values were computed for these two metrics and used within a decision-tree for labeling a given day as a typical day characterized by one of the six groups. The main limitation of the method is that the Euclidean distance metric, constructed by adding quadratic terms corresponding to metrics with different scales and units, partitions days in the transformed domain of the combined metrics. This obscures the relation of the groups to the individual metrics. Thus, grouping with a finer level of granularity cannot be achieved with this method.

The method proposed in this paper overcomes the limitations of the previous approach by creating groups based on each metric individually using the K-Means algorithm. Each day is then tagged with a composite ID consisting of IDs of the groups it belongs to based on different metrics. For example, if a day is a member of Group 1 based on Metric 1, a member of Group 1 based on Metric 2, and a member of Group 3 based on Metric 3, it is tagged with the composite group ID of (1, 1, 3), where the first index indicates grouping based on Metric 1, the second based on Metric 2 and the third based on Metric 3. All days with the same tag are then placed in one group. A salient feature of the proposed algorithm is that the linguistic description of the group labels based on each metric is retained in the composite label of the final grouping. For example, if groups 1, 2 and 3 mean "low," "medium" and "high,"

* Principal Scientist and Task Manager, U. C. Santa Cruz, MS 210-8, Associate Fellow.
† Technical Area Manager, Aerospace Systems, U. C. Santa Cruz, MS 210-8.

Copyright © 2007 by the American Institute of Aeronautics and Astronautics, Inc.
The U.S. Government has a royalty-free license to exercise all rights under the copyright claimed herein for Governmental purposes.
All other rights are reserved by the copyright owner.

respectively, the composite label (1, 1, 3) means "low" based on the first metric, "low" based on the second metric and "high" based on the third metric. Thus, the fidelity of partitions of individual metrics is retained in the final grouping.

The rest of the paper is organized as follows. Major trends observed in the 517 days of NAS delay data are described in Section II. Total time delay in minutes is used as a distance metric within the K-Means algorithm to partition the set of 517 days into ten groups in Section III. Convergence characteristics of the K-Means algorithm and summary statistics of the ten groups are provided in this section. A multiple-metric classification technique that builds on the single-metric classification technique is then developed in Section IV. Two examples of grouping of days are provided in Section IV to highlight the salient features of the algorithm. Conclusions are discussed in Section V.

II. National Airspace System Delay and Traffic-Volume Characteristics

To keep track of operational efficiency of the air traffic system, the Federal Aviation Administration (FAA) and the Bureau of Transportation Statistics (BTS) keep records of a multitude of metrics including delay, number of operations, conditions at airports, and traffic management initiatives in databases. Several of the frequently used databases are: Aviation System Performance Metrics (ASPM), Air Traffic Control System Command Center (ATCSCC) Logs, BTS data, Enhanced Traffic Management System (ETMS) and OPSNET. Detailed descriptions of the contents of these databases are available in Refs. 1 and 4.

All the analysis and the results in this paper are based on OPSNET data, which are available via http://www.apo.data.faa.gov. OPSNET data only include delays of fifteen minutes or more experienced by Instrument Flight Rule (IFR) flights that are reported by the FAA facilities. These data do not include delays caused by mechanical or other aircraft operator problems. Speed reductions and pilot initiated deviations around weather are also not reported. Taxi times spent under non-FAA facilities, for example under company/airport ramp towers, are not included in delay reports.[5] ASPM also provides delay data that are computed based on the Out-Off-On-In (OOOI) data provided by nine commercial and cargo carriers, which can also be utilized for analyzing days via the methods discussed in this paper. Although the trends in ASPM and OPSNET data are similar, the two databases contain very different types of data that make comparisons between them difficult. They are both useful depending on the analysis desired.

OPSNET delay data are provided in a tabular form; numbers of aircraft delayed are reported by *category*, *by class* and *by cause*. Delays by category consist of numbers of aircraft that were subject to departure delays, arrival delays, enroute delays and traffic management system (TMS) delays. The distinction between the enroute and TMS delays is discussed later in this section. Delays by class consist of numbers of air carrier, air taxi, general aviation and military aircraft that were delayed. Delays by cause consist of numbers of aircraft that were delayed due to weather, terminal volume, center volume, equipment limitations, runway issues and "other" issues. International delays are included in the "other" category. In addition to these, average time delay in minutes and total time delay in minutes are included in the table.

Seventeen variables of OPSNET national delay data for two days are summarized in Table 1. This table shows that the numbers of aircraft delayed by category (departure + arrival + enroute + TMS) add up to the total number of aircraft delayed. Similarly, the numbers of aircraft delayed by class and by cause also add up to the total number of aircraft delayed during the day of operation. Observe that the average delay is obtained by dividing the total time delay in minutes by the total number of delayed aircraft.

Table 1. OPSNET NAS delay summary data.

Data Variable	4/10/2004	4/13/2004
Delays by Category		
Total # of Aircraft	391	2,312
Departure	257	651
Arrival	101	391
Enroute	0	12
TMS	33	1,258
Delays by Class		
Air Carrier	338	1,769
Air Taxi	26	474
General Aviation	27	69
Military	0	0
Delays by Cause		
Weather	235	2,049
Terminal Volume	59	27
Center Volume	41	13
Equipment	1	26
Runway	30	24
Other	25	173
Time Delay		
Average Time (min.)	32.27	53.51
Total Time (min.)	12,616	123,709

There are three significant trends that are easily seen in Table 1. First, the sum of departure and TMS delayed flights account for most of the delayed flights. Second, most aircraft are delayed due to weather. Third, as expected, total time delays are proportional to total numbers of aircraft delayed.

To understand NAS delay characteristics, OPSNET delay data covering a period of 517 days (17 months) spanning the period from January 1, 2003 through May 31, 2004 were analyzed. Figure 1 shows a scatter plot of the percentages of aircraft delayed due to weather as a function of days. The mean percentage of aircraft delayed due to weather was found to be 66% and the standard deviation was found to be 21% for this dataset. Additional statistical characteristics are summarized in Table 2. These results show that the number of aircraft delayed due to weather represents a large fraction of the number of aircraft that experience delay in the NAS, a finding consistent with Ref. 6, which states that weather is responsible for approximately 70% of NAS delays.

The data shown earlier in Fig. 1 was reorganized as a function of total number of aircraft that experienced delay. These transformed data are shown in Fig. 2. Figure 2 shows that on days when a large number of aircraft are delayed, weather is the dominant cause of delays. Percentages of aircraft delayed due to weather are widely scattered when fewer aircraft are delayed, which indicates that factors other than weather are also responsible for delays on those days.

Figure 3 shows the number of aircraft delayed due to weather versus the total number of aircraft delayed. Viewing the sample points with respect to the diagonal line across the figure, it is clear that a high degree of correlation exists between the number of aircraft delayed due to weather and the total number of aircraft delayed in the NAS. Assuming both the number of aircraft delayed due to weather and the total number of aircraft delayed are random variables, the correlation coefficient was computationally determined to be 0.95. Correlation between the number of aircraft delayed due to weather and the total time delay in minutes due to all reportable causes (see the last row of Table 1 for an example) was found to be 0.94.

The causes of delay other than weather were also studied. Their statistics are summarized in Table 3 along with those of weather delays. Correlation coefficients ρ_1 in Table 3 are defined with respect to the total number of aircraft delayed, and correlation coefficients ρ_2 are defined with respect to the total accrued time delay in minutes. As evident from the correlation coefficient value of 0.21 in this table, the number of aircraft delayed due to volume has a weak linear correlation with the total number of aircraft delayed in the NAS. Correlation is even lower, 0.11, with respect to the total time delay in minutes. Similarly, the value of the correlation coefficient between the number of aircraft delayed due to equipment, runway and other non-US facilities, and the total number

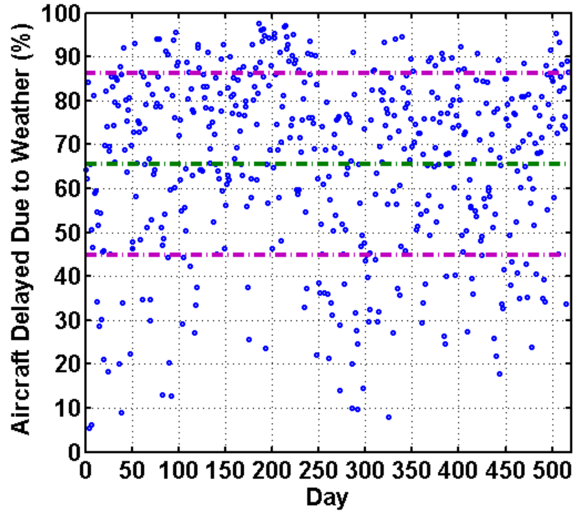

Figure 1. Percentage of aircraft delayed due to weather.

Table 2. Statistical characteristics of percentages of aircraft delayed due to weather.

Characteristic	Aircraft delayed by weather
Mean	66%
Standard deviation	21%
Minimum	5%
Maximum	98%
Median	70%

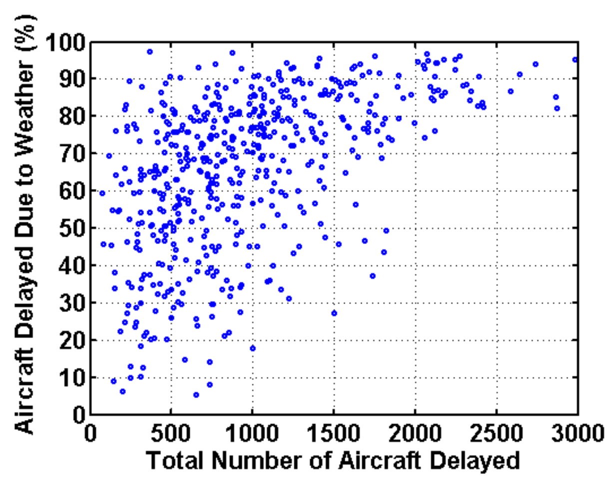

Figure 2. Percentage of aircraft delayed due to weather as a function of total number of aircraft delayed.

of aircraft delayed in the NAS was found to be 0.27. It was found to be 0.16 with respect to the total time delay in minutes. In the hierarchy of prime causal factors for delays, equipment and runway conditions follow weather. Results presented in this section suggest that delay metrics that encapsulate weather characteristics are likely to be useful in the classification of days.

Delays attributed to weather, volume, and equipment, runway and other causes are realized via departure, arrival, enroute and TMS restrictions. Departure delays incur by holding aircraft at the gate, on the taxiway, short of the runway, and on the runway. Arrival delays accrue when aircraft are delayed in the arrival Center's airspace or in Terminal Radar Approach Control airspace due to restrictions at arrival airports. Enroute delays occur when aircraft are held as a result of initiatives imposed by a facility for traffic management reasons such as volume regulation, frequency outage and weather. The other major category of delays in the OPSNET data is TMS delays, which result from national or local Center (coordinated with Air Traffic Control System Command Center) traffic flow management initiatives such as Ground Delay Programs, local Ground Stops (GS), Departure Sequencing Programs, Enroute Spacing Programs, Arrival Sequencing Programs, airborne holding, rerouting, Miles-in-Trail, Minutes-in-Trail and Fuel Advisory.

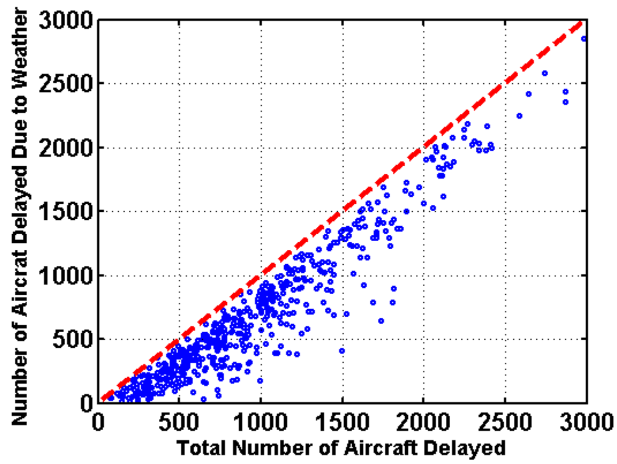

Figure 3. Proportion of number of aircraft delayed due to weather to the total number of aircraft delayed in the NAS.

Table 3. Summary of weather, volume, and equipment, runway and other delay characteristics.

Characteristic	Weather	Equip., Runway & Other	Volume
Mean	66%	20%	14%
Standard deviation	21%	16%	11%
ρ_1	0.95	0.27	0.21
ρ_2	0.94	0.16	0.11

Table 4. Summary of departure, enroute, arrival and TMS delay characteristics.

Characteristic	Departure	TMS	Arrival	Enroute
Mean	47%	39%	14%	1%
Standard deviation	17%	17%	8%	1%
ρ_1	0.82	0.86	0.55	0.45
ρ_2	0.73	0.86	0.51	0.49

To determine relative contributions of delays attributed to departure, enroute and arrival phases of flights, and to TMS restrictions, percentages of aircraft delays by category were calculated for the 517-day dataset. It was determined that on an average, on any given day, 47% of the aircraft that are delayed in the NAS, are delayed during departure, 1% during enroute and 14% during arrival phases of flight. The average percentage of aircraft delayed due to TMS was 39%. Analysis of the data showed that, on average, departure delayed flights and TMS delayed flights roughly account for 86% of the flights that are delayed. Additional statistics that characterize these delays are summarized in Table 4. Note that the values of the correlation coefficients ρ_1 and ρ_2 listed in the table are with respect to the total number of aircraft delayed and the total accrued time delay in minutes, respectively.

Since traffic management initiatives such as GDP and GS are applied to aircraft while they are on the ground, and rerouting and holding while they are airborne, TMS delays include both ground and airborne delay components. To separate TMS delay into ground delay and airborne delay components, analysis of GDP and GS data, that are also available via OPSNET, is needed. Like data in Table 1, these data are also provided in a tabular format with the following items: 1) date, 2) number of aircraft delayed, 3) total delay in minutes, and 4) average time delay in minutes. For example, GDP and GS data for two days, 10 April 2004 and 13 April 2004, are shown in Table 5. NAS delay data for the same two days were previously itemized in Table 1.

Ground delay and airborne delay components can be computed using NAS delay data (see delays by category in Table 1) and the GDP and GS delays (see Table 5) as follows. Let, n_d, n_{GDP}, and n_{GS} be the numbers of aircraft delayed during departure, due to GDP, and due to GS, respectively. The total number of aircraft delayed on the ground is then

Table 5. OPSNET GS and GDP delay data.

Data Variable	4/10/2004	4/13/2004
Ground Stops		
# of Aircraft Delayed	3	98
Minutes of Delay	292	7,079
Average Delay	97.33	72.23
Ground Delay Program		
# of Aircraft Delayed	6	1,044
Minutes of Delay	244	85,166
Average Delay	40.66	81.57
Total Delays Due to GS and GDP		
# of Aircraft Delayed	9	1,142
Minutes of Delay	536	92,245
Average Delay	59.55	80.77

$$n_G = n_d + n_{GDP} + n_{GS} \tag{1}$$

Number of aircraft delayed during the airborne phase can be obtained after subtracting GDP and GS components from TMS delays as,

$$n_A = n_a + n_e + (n_{TMS} - n_{GDP} - n_{GS}) \tag{2}$$

where n_a, n_e, and n_{TMS} are the numbers of aircraft delayed in arrival phase, in enroute phase, and due to TMS, respectively. Note that

$$n_T = n_G + n_A = n_d + n_a + n_e + n_{TMS} \tag{3}$$

where n_T is the total number of aircraft delayed in the NAS. Results obtained using Eqs. (1) through (3) with 517 days of OPSNET data are shown in Fig. 4; n_G and n_A values are plotted against n_T values. This figure shows that when more aircraft are delayed, a significantly higher number of them are delayed on the ground compared to in the air.

Statistical trends summarized in Table 6 show that on average 74% of the aircraft that experience delay are delayed on the ground, compared to an average of 26% that are delayed while airborne. The last row of Table 6 shows that on some days NAS conditions are unusual in that a large percentage of delayed aircraft experience airborne

Table 6. Summary of aircraft delayed on ground versus aircraft delayed in the air.

Characteristic	Delayed on Ground	Delayed in Air
Mean	74%	26%
Median	76%	24%
Standard Deviation	11%	11%
Minimum	23%	7%
Maximum	93%	77%

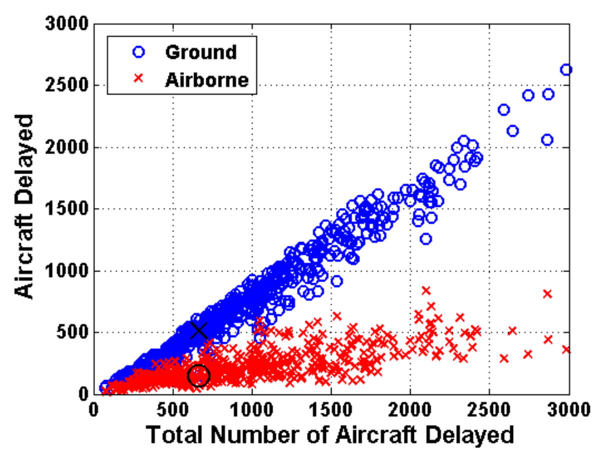

Figure 4. Proportion of number of aircraft delayed on the ground and in the air to the total number of aircraft delayed in the NAS.

delay. Of the 665 aircraft that were delayed on 2/22/2003, 513 aircraft (77%) were delayed during their airborne phase of flight. The airborne and ground delay values for this day are marked with a large 'X' and a large 'O' in Fig. 4.

In addition to the NAS delay metrics discussed in this section, past studies such as Ref. 1, 2 and 7 have used metrics of traffic volume to select days for analysis. There is consensus in the literature that the "traffic volume" and the delay taken together characterize NAS operations, therefore traffic volume metrics are discussed next. OPSNET database includes *Towers: Summary Report*, *Instrument Operations: Summary Report* and *Centers: Summary of Domestic Operations Report* that contain traffic volume data. These three reports count traffic from different perspectives. One is unable to separate departure counts from arrival counts in the Towers: Summary Report and in the Instrument Operations: Summary Report. A departure at one facility is counted as a departure at that facility and as an arrival at a different facility. Since departures and arrivals are counted together twice in these reports, the total number of operations excluding the overflight operations have to be halved to estimate the number of departures or the number of arrivals. The Centers: Summary of Domestic Operations Report directly provides a count of the number of departures from airports within the ARTCCs. Since departure count eventually drives the overflight count and the arrival count, it represents the traffic demand. Due to this reason, departure count from the Centers: Summary of Domestic Count Report has been used in this paper. Table 7 lists the departure counts and the overflight counts for the two days. Departure counts excluding military flights for the two days obtained by summing the air carrier, air taxi and general aviation departures are 31,959 and 42,062.

A histogram of the total domestic departure counts for the 517 days of data is shown in Fig. 5. The minimum and the maximum numbers of departures were found to be 25,677 on 11/27/2003 (Thanksgiving holiday) and 51,399 on 5/27/2004 (Thursday). Observe that the histogram is bimodal which indicates that days can be classified into two categories - low departure count day and high departure count day. Reference 1 noted similar observations and offered evidence that the bimodal distribution is primarily due to the weekday versus weekend traffic levels.

This section described several delay and traffic volume metrics that are available in OPSNET data. Summary statistics described in the tables and the patterns observed in the figures suggest that these metrics can be used for distinguishing one day of NAS operations from another day of NAS operations. To illustrate the use of a metric for classifying days of operations, total time delay in minutes, n_T in Eq. (3), is used as a distance metric within the K-Means method in the next section.

Table 7. Centers: Summary of Domestic Operations Report.

Data Variable	4/10/2004	4/13/2004
Departures		
Air Carrier	17,122	20,231
Air Taxi	9,132	12,831
General Aviation	5,705	9,000
Overflights		
Air Carrier	23,021	23,753
Air Taxi	3,556	4,318
General Aviation	2,763	4,763

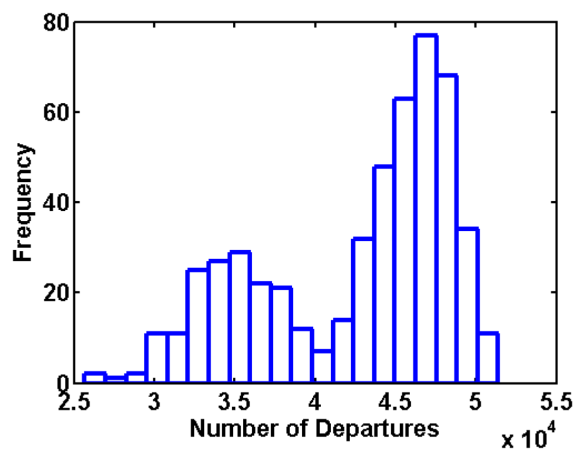

Figure 5. Histogram of 517 days of total domestic departure counts.

III. Single-Metric Classification

The motivation for assigning or labeling days into groups with associated properties, such as mean delay values, is to aid selection of prototype days for analysis. For example, a few days from a group of days with large delays and from a group of days with small delays can be selected for system-wide studies using the National Aeronautics and Space Administration's air traffic simulation, concept evaluation, and decision support tools such as the Airspace Concept Evaluation System (ACES), the Center TRACON Automation System (CTAS) and the Future ATM Concepts Evaluation Tool (FACET).[8-11]

All classification processes use metrics, or features, of the data to partition it into groups. A popular classification method, known as the K-Means method, partitions data such that the means associated with the groups are as widely separated as possible.[3] The method labels the data elements based on their closeness to the group means for reducing the group variance. The K-Means algorithm consists of two steps: 1) the initialization step, and 2) the iterative step. Data elements that are far apart from each other are chosen as the initial means of the groups during the initialization step. Each element is then assigned to the group that it is closest to, based on its distance with respect to the group's initial mean value. Group means are then recomputed based on the elements assigned. Each element is then reassigned to its closest group based on its distance with respect to the recomputed mean values. This process of computation of the means and reassignment of elements to groups is continued in subsequent iterative steps. Convergence is achieved when the numerical values of the group means do not change with reassignment of the elements. Iterations are halted once convergence is achieved.

To further clarify the initialization and the iterative steps of the K-Means algorithm, consider a vector with the following elements [0, 0.5, 0.8, 1.2, 5, 7, 12, 15, 20, 25]. If two groups are desired, the elements with values closer to 0 are assigned to the first group and the elements with values closer to 25 are assigned to the second group. Thus, elements one through seven are assigned to Group One and elements eight through ten are assigned to the Group Two in the initialization step. With this assignment of the elements to the groups, average values of the first group and the second group are 3.79 and 20 and the standard deviation values are 4.47 and 5. Reassignment of the elements based on the recomputed means results in the first six elements being assigned to Group One and the last four elements being assigned to Group Two in the first iterative step. Group means are recomputed in the next iterative step. These means are 2.42 and 18 and the standard deviations are 2.87 and 5.72. The next iterative step results in the same means and the standard deviations as those in the prior step; final grouping is therefore achieved in the previous step. For this example, the K-Means algorithm partitions the data into two groups in three steps.

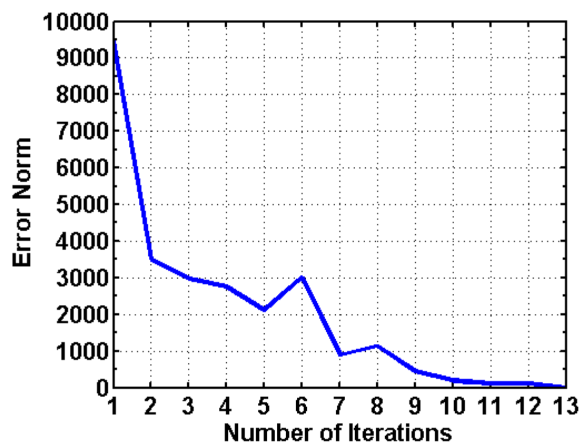

Figure 6. Convergence characteristics of the K-Means algorithm as it partitions the 517 days into ten groups.

If three groups are desired for the above example, a value from the vector that is far away from both 0 and 25 needs to be selected as the initial value for the third group. Observe that this value is 12 since its minimum distance to 0 (12 units) and 25 (13 units) is a maximum compared to the minimum distances of other elements to 0 and 25. Other values in the vector are less than 12 units with respect to either 0 or 25. Once these initial group means are chosen, the subsequent iterative steps are the same as those described in the previous paragraph. It should be noted that good initial conditions are needed because the K-Means algorithm is sensitive to initial conditions.

The K-Means algorithm was used for classifying 517 days into ten groups using total time delay in minutes as the discriminating metric. The choice of ten groups was arbitrary. The algorithm converged in thirteen iterations. Its convergence characteristics are shown in Figure 6. Properties of the ten groups, based on total time delay statistics of the elements assigned to the groups, are summarized in Table 8. The second column of the table shows the number of days in the group, with the group ID given in the first column. Columns three and four show the average delay and the standard deviation of the delays in minutes. Columns five and six show the minimum and maximum delays in minutes of the days belonging to the particular groups. The data in this table

Table 8. Summary of properties of the ten groups.

Group ID	Number of Days	Mean Delay (min.)	Standard Deviation (min.)	Minimum Delay (min.)	Maximum Delay (min.)
1	145	11,302	4,310	1,686	18,834
2	126	26,626	4,489	19,107	34,628
3	92	43,026	5,095	34,947	52,113
4	57	61,239	5,198	52,562	70,011
5	38	80,090	5,421	71,186	90,464
6	32	102,820	5,944	94,023	112,031
7	11	124,652	3,612	119,692	131,172
8	7	141,633	5,854	133,884	148,341
9	7	163,501	5,887	156,717	172,211
10	2	183,426	4,083	180,539	186,313

show that there are fewer days in groups associated with large delays. For example, group number ten consists of only two days compared to group number one with 145 days. Observe that the mean values associated with the groups are approximately equally spaced and that the standard deviation values are fairly close to each other. Standard deviation values can be expected to increase with fewer groups.

Probability density functions corresponding to Gaussian distributions with the group means and standard deviations listed in Table 8 are shown in Figure 7. Note that the extent of the abscissa is limited to the range of the delay data.

Sixteen days (seven days in Group 8, seven days in Group 9 and two days in Group 10) that experienced large delays are listed in Table 9. Of the 517 days grouped by the K-Means algorithm, the least delay of 1,686 minutes occurred on 1/11/2003 (a Saturday) and most delay of 186,313 minutes occurred on 5/13/2004 (a Thursday).

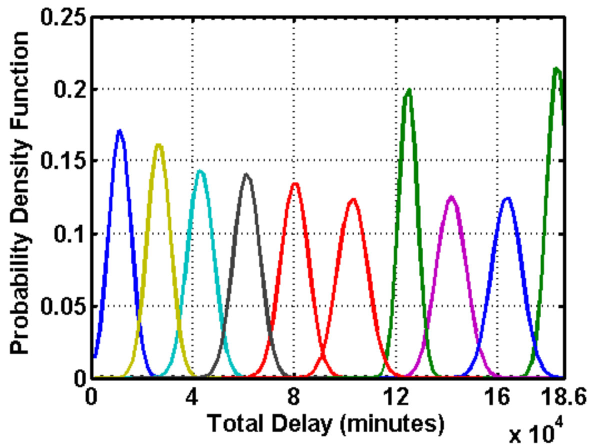

Figure 7. Gaussians representing the ten groups.

Table 9. Days with large delays.

Group ID	Date	Total Delay (min.)
8	7/7/2003	133,884
8	8/3/2003	148,341
8	8/4/2003	147,888
8	11/12/2003	134,602
8	2/6/2004	143,857
8	3/5/2004	139,558
8	5/14/2004	143,298
9	3/20/2003	160,764
9	7/22/2003	164,069
9	8/1/2003	159,471
9	5/12/2004	170,777
9	5/17/2004	160,496
9	5/18/2004	172,211
9	5/31/2004	156,717
10	5/13/2004	186,313
10	5/21/2004	180,539

Results presented in this section demonstrate the use of the K-Means algorithm for partitioning a set of days into groups organized in order of a single metric like total time delays. The next section describes a labeling technique that enables use of the single-metric K-Means classification technique for achieving classification based on multiple metrics.

IV. Multiple-Metric Classification

Motivation for multiple-metric classification stems from the desire for finer levels of partitioning. For example, a group with large mean delay contains days when aircraft were delayed due to weather and also days when aircraft were delayed due to runway conditions. In order to discern which ones were affected by weather and which ones were affected by runway conditions, one would need metrics such as numbers of aircraft delayed due to weather and due to runway conditions, in addition to total delays. Fidality

Classification based on multiple metrics has been traditionally accomplished by weighing and combining several metrics into a single metric, and then using it in a K-Means algorithm. For example, if day 'q' was characterized by f metrics: $u_{q,1}$, $u_{q,2}$, …, $u_{q,f}$ and if day 'r' was similarly characterized by $u_{r,1}$, $u_{r,2}$, …, $u_{r,f}$, a weighted quadratic function of the form

$$d_{q,r} = \sum_{1 \le l \le f} w_l (u_{q,l} - u_{r,l})^2 \tag{4}$$

can be defined as the distance metric between days q and r. Note that w_1 through w_f are the weights corresponding to the different metrics. Interpretation of the distance metric, Eq. (4), for grouping days with the K-Means algorithm is straightforward with $p_{r,l}$ as the mean of measure l in group r. Distance to each group can then be written as:

$$d_{k,j} = \sum_{1 \le l \le f} w_l (u_{k,l} - p_{j,l})^2; \quad 1 \le j \le n; \quad 1 \le k \le m \tag{5}$$

where f is the number of metrics, m is the number of days and n is the number of groups.

Although the distance metric, Eq. (5), enables transformation of a multiple-metric classification problem into a single-metric classification problem, its deficiencies are noteworthy. Limitations stem from the fact that metrics have different scales and units, and that only their combined contribution is available in the distance metric; classification is insensitive to individual contribution. For example, consider the two metrics in Table 1: 1) number of aircraft affected by departure delays and 2) total time delay. Units of the two metrics are quite different, number of aircraft and minutes. The scales are also different by an order of magnitude; 257 aircraft impacted by departure delays versus 12,616 minutes of total time delay on 10 April 2004. To compensate for these differences, the associated weights have to be scaled correctly, and their units have to be chosen appropriately to enable summation of quantities with disparate units. References 1 and 2 suggest that the inverse of the statistical variance of the metric should be used to weigh its contribution. Even with this scaling, a meaning cannot be ascribed to the grouping in the native domain of the metrics.

In order to overcome the limitations of the weighted quadratic distance function used in the prior approach of Refs. 1 and 2, a multiple metric classification technique that treats each metric independently of others in an f-dimensional metric space is proposed. Since each metric is treated independently, the single metric K-Means algorithm described earlier in Section III can be used for assigning days to groups based solely on each metric. IDs of these groups are then coordinates in the f-dimensional metric space. For the sake of discussion, consider the problem of classifying days into four groups using two metrics. Using the K-Means algorithm twice, days are first assigned to four groups based on Metric One, and then to four groups based on Metric Two. The resulting sixteen possible groups are labeled using two indices as follows: (1, 1), (1, 2), (1, 3), (1, 4), (2, 1), (2, 2), (2, 3), (2, 4), (3, 1), (3, 2), (3, 3), (3, 4), (4, 1), (4, 2), (4, 3) and (4, 4). The first index denotes group ID based on Metric One and the second index denotes group ID based on Metric Two. Thus, a day which is assigned to the second group based on Metric One and to the first group based on Metric Two is a member of group (2, 1). Since a unique group is labeled using two indices in this two metric classification problem, the combined group IDs are coordinates in a two-dimensional metric space.

Generalization of the technique to f metrics such that days are classified into n_l groups using metric l results in

$$n_g = \prod_{1 \le l \le f} n_l \tag{6}$$

number of possible groups. Equation (6) shows that if the same number of groups is desired for each metric, the number of possible groups is given in terms of the power of f. For example if n groups are desired with each metric, the number of possible groups is n^f. Thus, it is seen that the growth in the number of groups is explosive with increasing number of metrics. Should one conclude that the growth is unbounded based on this observation, or is there an upper bound on the number of groups? The answer is provided by the following. If each day is classified into its own group, one would have the same number of groups as the number of days; hence, number of days is the upper bound. This fact also implies that if m is the number of days and $n_g > m$, there are at least $n_g - m$ number of empty groups, groups without any members. Removing these empty groups, the number of possible groups, n_g, is given as

$$n_g = \min(m, \prod_{1 \le l \le f} n_l) \tag{7}$$

where each $n_l \le m$.

Is it possible that several of the groups counted in Eq. (7) are empty? One can demonstrate this to be true by constructing the following examples. Consider the problem of classifying ten days into two sets of five groups using two metrics. Following the nomenclature of Eq. (7), $f = 2$ for the two metrics, $n_1 = 5$ using Metric One, $n_2 = 5$ using Metric Two, and $m = 10$ for the ten days. Substituting these numerical values in Eq. (7) it is seen that $n_g = 10$. Assume that the first two days are assigned to Group 1, the third and the fourth to Group 2, the fifth and the sixth to Group 3, the seventh and the eighth to Group 4, and the final two to Group 5 based on Metric One, and also based on Metric Two. In this scenario, groups with members are (1, 1), (2, 2), (3, 3), (4, 4) and (5, 5). All other groups labeled with two different indices, such as (1, 2), (1, 3) and (5, 1), are empty.

A similar example can be constructed to show that empty groups are possible even when $n_g < m$. If $f = 2$ for the two metrics, $n_1 = 2$ using Metric One, $n_2 = 2$ using Metric Two, and $m = 10$ for the ten days, $n_g = 4$ from Eq. (7). If the first five days are assigned to Group 1 and the next five to Group 2 based on both the metrics, groups (1, 1) and (2, 2) are non-empty while groups (1, 2) and (2, 1) are empty. These two examples clearly show that it is always possible to have empty groups.

An aspect of multiple-metric classification that has not been discussed so far is the semantics associated with the group IDs. Without a linguistic meaning, it is difficult to interpret what do group IDs such as (1, 1) and (1, 2) mean. One of the ways of attributing a meaning to the indices is to order them according to the increasing values of the group means. For example if total time delay in minutes was the metric being considered, the index with the least value would correspond to the group with the minimum mean total time delay while the index with the highest value would correspond to the group with the maximum mean total time delay. From an implementation perspective, once classification into specified number of groups is accomplished with the K-Means algorithm using a single metric, and group means are computed based on the metric values of the assigned members, the group means are sorted in an increasing order. Indices of the groups are then re-labeled to reflect the sorted order. This procedure is repeated for each metric to obtain the complete set of indices required for labeling the groups.

To illustrate the utility of the multiple-metric classification technique, an example of classifying 522 days, which included the 517 days discussed previously and five special days used in earlier studies, into groups is presented next. These five special days are 5/17/2002, 4/17/2005, 4/21/2005, 6/5/2005 and 7/15/2005. 5/17/2002 is the ACES baseline day. The other four days were used earlier in Ref. 7. They were categorized as a low-volume low-weather day, high-volume low-weather day, low-volume high-weather day and high-volume high-weather day, respectively in Ref. 7.

Three metrics- 1) total domestic departure counts, 2) number of aircraft delayed on the ground, and 3) number of aircraft delayed in the air were used as the basis for classification in the multiple-metric K-Means algorithm described in this paper. Recollect that the total domestic departure counts were obtained from the *Centers: Summary of Domestic Counts Report* discussed in Section II. Numbers of aircraft delayed on the ground and delayed in the air were computed using Eqs. (1) and (2). Days were initially organized into three groups using the single-metric K-Means algorithm. When number of aircraft delayed in the air was used as a metric, 393 days were assigned to Group 1, 128 to Group 2 and one to Group 3. The mean and the standard deviation values derived from the number of aircraft delayed in the airborne phase metric of the assigned members for these groups are listed in Table 10. The sole member of Group 3, 7/15/2005, had excessive amount of airborne delay of 1891 minutes. 1661 aircraft were delayed on the ground on this day. Since this day is an outlier, it has the effect of increasing the standard deviation of the other groups. Due to this reason, it was removed from the dataset.

Table 10. Three groups based on number of aircraft delayed in the airborne phase.

Group Number	Number of Days	Mean Number Delayed	Standard Deviation	Minimum Number Delayed	Maximum Number Delayed
1	393	124	49	7	218
2	128	314	100	222	970
3	1	1891	0	1891	1891

The analysis was repeated with the remaining 521 days. The resulting grouping showed that 6/5/2005 became the sole member of Group 3 with 970 aircraft delayed in the airborne phase and 753 aircraft delayed on the ground. Table 11 summarizes these results. Comparing Table 10 to Table 11, it is seen that the removal of 7/15/2005 data lowers the standard deviation values of the groups.

Since 6/5/2005 is an outlier day, it was also removed from the dataset.

Classification based on number of aircraft delayed in the airborne phase for the remaining 520 days are summarized in Table 12. Observe that the standard deviation values decrease further for groups 1 and 2. It increases for Group 3. The mean values decrease for all three groups.

Days belonging to Group 1 can be thought of as days with low number of aircraft delayed while airborne; in Group 2 as days with medium number of aircraft delayed and the ones in Group 3 as days with large number of aircraft delayed. Similar categorization based on number of aircraft delayed on the ground partitions the days in Groups 1 through 3, whose statistics are summarized in Table 13. Results obtained using total domestic departure counts are provided in Table 14. Note that the days were classified into two groups based on the bimodal distribution seen in Fig. 5.

Comparing Tables 12 and 13, it is seen that the trends are similar with the largest number of days assigned to groups with lower mean and lower standard deviation values. The trends are different in Table 14. More days are assigned to Group 2 with higher mean departure counts.

Table 11. Three groups based on number of aircraft delayed in the airborne phase (excluding 7/15/2005).

Group Number	Number of Days	Mean Number Delayed	Standard Deviation	Minimum Number Delayed	Maximum Number Delayed
1	386	123	47	7	213
2	134	304	81	214	604
3	1	970	0	970	970

Table 12. Groups based on number of aircraft delayed in the airborne phase (excluding 7/15/2005 and 6/5/2005).

Group Number	Number of Days	Mean Number Delayed	Standard Deviation	Minimum Number Delayed	Maximum Number Delayed
1	291	103	36	7	158
2	174	214	38	159	295
3	55	382	71	304	604

Table 13. Three groups based on number of aircraft delayed on the ground.

Group Number	Number of Days	Mean Number Delayed	Standard Deviation	Minimum Number Delayed	Maximum Number Delayed
1	249	379	160	48	646
2	185	913	173	651	1,293
3	86	1,689	328	1,309	2,778

Table 14. Two groups based on total domestic departure counts.

Group Number	Number of Days	Mean Delay (min.)	Standard Deviation (min.)	Minimum Delay (min.)	Maximum Delay (min.)
1	168	34,792	2,920	25,677	40,528
2	352	46,335	2,243	40,596	51,759

Given that three groups were created using two metrics and two groups using one metric, the total number of possible composite groups, determined using Eq. (7), is 18. The range of IDs for these groups is (1, 1, 1) to (2, 3, 3). With the first index being the group number associated with the total domestic departure counts metric, the second index being the group number associated with the number of aircraft delayed on the ground metric, and the third index being the group number associated with the number of aircraft delayed during the airborne phase metric, each day in the set of 520 days has a three index group ID associated with it.

Organizing the resulting triple index group IDs using a "dictionary sort" algorithm, each unique group and its members are determined. The values of the metrics of the members are then used for determination of minimum, maximum, mean and standard deviation values associated with the groups. Results of this process for the three-metric classification problem, being discussed here, are outlined in Table 15. Group means for the three metrics are listed in the columns labeled μ_1, μ_2 and μ_3; standard deviation values are listed in the columns labeled as σ_1, σ_2 and σ_3.

Table 15. Final grouping with three-metric classification.

Group Number	Group ID	Number of Days	μ_1	μ_2	μ_3	σ_1	σ_2	σ_3
1	(1, 1, 1)	94	33,970	310	83	2,718	158	37
2	(1, 1, 2)	14	33,698	444	224	3,179	149	41
3	(1, 1, 3)	2	34,650	373	348	6,588	120	41
4	(1, 2, 1)	27	36,060	880	110	2,565	152	33
5	(1, 2, 2)	19	36,816	951	221	1,916	159	37
6	(1, 2, 3)	2	36,583	951	390	4,185	305	95
7	(1, 3, 1)	3	35,219	1,713	131	4,418	301	8
8	(1, 3, 2)	6	37,067	1,586	213	2,296	204	30
9	(1, 3, 3)	1	36,485	1,462	355	0	0	0
10	(2, 1, 1)	89	45,495	383	110	2,200	151	31
11	(2, 1, 2)	47	46,195	483	204	2,008	117	33
12	(2, 1, 3)	3	43,944	505	447	2,494	148	52
13	(2, 2, 1)	60	47,330	898	113	1,961	188	31
14	(2, 2, 2)	50	46,393	911	205	2,383	156	36
15	(2, 2, 3)	27	45,934	953	359	2,201	196	58
16	(2, 3, 1)	18	46,686	1,614	116	2,190	326	31
17	(2, 3, 2)	38	47,027	1,728	231	1,947	380	41
18	(2, 3, 3)	20	46,539	1,721	408	2,537	261	81

Many of the salient features of multiple-metric classification algorithm are apparent from data in Table 15. A majority of days, 94 and 89, are assigned to Group (1, 1, 1) and Group (2, 1, 1). These groups represent days on which few aircraft were delayed. The difference between them is that Group (1, 1, 1) represents low-volume days while Group (2, 1, 1) represents high-volume days. Table 15 shows that a large number of days were assigned to groups (2, 2, 1) and (2, 2, 2) that represent high-volume days on which many aircraft were delayed on the ground. There are several groups with few days assigned to them; five groups had three or less than three days assigned to them. Table 16 lists the corresponding dates. Member days belonging to small groups can be considered to be special. Days in Group 3 with Group ID (1, 1, 3) experienced relatively low total departure counts, fewer aircraft affected by ground delay, and higher number of aircraft affected by airborne delay. Two members of Group 6 had more aircraft delayed on the ground compared to members of Group 3. Three member days of Group 7 had many more aircraft delayed on the ground and few during the airborne phase. The sole member of Group 9 had many aircraft delayed while on ground and while in the air. The member days of Group 12 experienced high departure counts, relatively few aircraft delayed on the ground, and a large number of aircraft delayed during the airborne phase.

Table 16. Days in small groups.

Group Number	Group ID	Date
3	(1, 1, 3)	2/22/2003
3	(1, 1, 3)	5/16/2004
6	(1, 2, 3)	6/14/2003
6	(1, 2, 3)	3/28/2004
7	(1, 3, 1)	9/14/2003
7	(1, 3, 1)	5/23/2004
7	(1, 3, 1)	5/30/2004
9	(1, 3, 3)	1/4/2004
12	(2, 1, 3)	5/2/2003
12	(2, 1, 3)	5/5/2003
12	(2, 1, 3)	9/30/2003

Another example of multiple-metric classification using total domestic departure counts and delays by cause: 1) weather, 2) volume, 3) equipment, runway and other, as metrics for partitioning 522 days into groups is summarized in Table 17. Numbers of aircraft delayed due to terminal volume and due to center volume (see "Delays by Cause" in Table 1) were combined to obtain the number of aircraft delayed due to volume. Similarly, numbers of aircraft delayed due equipment, runway and other issues were added together for obtaining the number of aircraft impacted due to these causes. As in the previous example, days were categorized into groups with the K-Means algorithm using each of these four metrics one at a time. Observe that in this example, days are partitioned into 37 groups out of 54 possible groups.

Table 17. Final grouping obtained using departure counts and delays by cause metrics.

Group Number	Group ID	Number of Days	μ_1	μ_2	μ_3	μ_4	σ_1	σ_2	σ_3	σ_4
1	(1, 1, 1, 1)	86	33,648	268	49	72	2,932	160	36	44
2	(1, 1, 1, 2)	22	35,243	357	60	252	2,172	174	39	64
3	(1, 1, 1, 3)	2	33,691	154	79	498	175	168	58	4
4	(1, 1, 2, 1)	7	35,796	355	212	89	1,339	159	61	34
5	(1, 1, 2, 2)	3	37,448	402	198	259	794	251	28	61
6	(1, 1, 2, 3)	2	36,120	352	229	495	889	40	20	144
7	(1, 1, 3, 1)	1	35,484	296	401	173	0	0	0	0
8	(1, 1, 3, 2)	1	39,454	394	541	238	0	0	0	0
9	(1, 2, 1, 1)	25	35,770	963	50	90	2,920	209	35	46
10	(1, 2, 1, 2)	7	36,854	960	80	238	2,653	302	38	68
11	(1, 2, 1, 3)	2	37,021	962	73	587	378	385	30	170
12	(1, 2, 2, 1)	2	37,118	960	291	105	3,430	353	1	28
13	(1, 2, 2, 2)	2	38,540	741	306	249	659	41	100	164
14	(1, 2, 3, 2)	1	36,485	898	710	209	0	0	0	0
15	(1, 3, 1, 1)	4	37,143	1,567	40	148	3,049	132	32	25
16	(1, 3, 2, 1)	2	37,814	1,867	177	129	580	27	11	1
17	(2, 1, 1, 1)	78	45,508	325	80	104	1,774	171	26	36
18	(2, 1, 1, 2)	21	45,740	330	85	234	2,349	173	21	47
19	(2, 1, 1, 3)	8	45,690	370	91	458	1,609	210	24	83
20	(2, 1, 2, 1)	29	46,356	337	194	98	2,204	181	46	36
21	(2, 1, 2, 2)	21	47,496	319	215	274	2,293	129	51	65
22	(2, 1, 2, 3)	7	47,610	314	195	491	1,867	148	36	58
23	(2, 1, 3, 1)	2	45,004	411	412	147	6,233	270	33	13
24	(2, 1, 3, 2)	2	50,358	192	382	265	1,981	85	18	11
25	(2, 1, 3, 3)	1	48,983	407	551	540	0	0	0	0
26	(2, 2, 1, 1)	64	46,090	940	80	96	2,318	217	30	39
27	(2, 2, 1, 2)	27	46,392	977	96	250	2,481	220	28	48
28	(2, 2, 1, 3)	3	47,117	1,172	106	555	2,916	331	24	293
29	(2, 2, 2, 1)	19	46,817	898	205	96	1,910	216	51	38
30	(2, 2, 2, 2)	17	47,587	1,021	205	262	2,533	241	53	58
31	(2, 2, 2, 3)	3	47,360	765	424	407	4,315	142	337	113
32	(2, 2, 3, 3)	2	45,890	744	429	434	1,870	64	20	71
33	(2, 3, 1, 1)	26	46,221	1,940	66	80	1,646	341	30	42
34	(2, 3, 1, 2)	13	47,142	1,897	88	252	2,020	279	26	61
35	(2, 3, 1, 3)	2	45,527	1,943	109	410	3,341	586	8	3
36	(2, 3, 2, 1)	5	47,467	1,735	209	120	482	227	34	37
37	(2, 3, 2, 2)	3	49,992	2.055	160	280	1,271	915	17	70

Table 18 lists the group membership of holidays and special days - ACES baseline day, Joint Planning and Development Office (JPDO) baseline day and days studied in Ref. 7. Group IDs from Table 15 are listed under Group ID 1 heading and from Table 17 under Group ID 2 heading. Results in this table show that the domestic departure counts are generally lower on holidays. Group ID 1 (2, 2, 2) indicates that the ACES baseline day has high departure counts, moderate number of aircraft delayed on the ground, and moderate number of aircraft delayed in the air; Group ID 2 (2, 2, 1, 1) indicates high departure counts, moderate number of aircraft delayed due to weather, low number of aircraft delayed due to volume and low number of aircraft delayed due to equipment, runway and other issues. The two group IDs for the JPDO baseline day indicate high departure counts, moderate number of aircraft delayed on ground, low number of aircraft delayed in the air, low number of aircraft delayed due to weather, moderate number of aircraft delayed due to volume and high number of aircraft delayed due to equipment, runway and other conditions. Ref. 7 considered 4/17/2005 to be a low departure count, low-delay due to weather day. The group IDs in Table 18 label this day as a high departure count, low-delay due to weather day. The departure count of 40,653 on this day is at the lower end of the high departure count group can be inferred from the statistics given in

Table 18. Classification of holidays and special days.

Number	Date	Significance	Day of Week	Group ID 1	Group ID 2
1	5/17/2002	ACES Baseline Day	Friday	(2, 2, 2)	(2, 2, 1, 1)
2	1/1/2003	New Year's Day	Wednesday	(1, 1, 1)	(1, 1, 1, 1)
3	1/20/2003	Martin Luther King Day	Monday	(1, 1, 1)	(1, 1, 1, 1)
4	2/17/2003	President's Day	Monday	(1, 1, 1)	(1, 1, 1, 1)
5	5/26/2003	Memorial Day	Monday	(1, 1, 1)	(1, 1, 1, 1)
6	7/4/2003	Independence Day	Friday	(1, 1, 1)	(1, 1, 1, 1)
7	9/1/2003	Labor Day	Monday	(1, 2, 1)	(1, 2, 1, 1)
8	10/13/2003	Columbus Day	Monday	(2, 1, 1)	(2, 1, 2, 1)
9	11/11/2003	Veterans Day	Tuesday	(2, 2, 1)	(2, 1, 2, 1)
10	11/25/2003	Two Days Before Thanksgiving	Tuesday	(2, 3, 1)	(2, 1, 3, 3)
11	11/27/2003	Thanksgiving Day	Thursday	(1, 1, 2)	(1, 1, 1, 1)
12	12/25/2003	Christmas Day	Thursday	(1, 1, 1)	(1, 1, 1, 1)
13	1/1/2004	New Year's Day	Thursday	(1, 1, 1)	(1, 1, 1, 1)
14	1/19/2004	Martin Luther King Day	Monday	(2, 1, 1)	(2, 1, 2, 2)
15	2/16/2004	President's Day	Monday	(2, 2, 2)	(2, 2, 2, 2)
16	2/19/2004	JPDO Baseline Day	Thursday	(2, 2, 1)	(2, 1, 2, 3)
17	5/31/2004	Memorial Day	Monday	(1, 3, 2)	(1, 3, 2, 1)
18	4/17/2005	Ref. 7 L/L Day	Sunday	(2, 1, 1)	(2, 1, 1, 2)
19	4/21/2005	Ref. 7 H/L Day	Thursday	(2, 1, 2)	(2, 1, 3, 2)
20	6/5/2005	Ref. 7 L/H Day	Sunday	Not included	(1, 3, 1, 1)
21	7/15/2005	Ref. 7 H/H Day	Friday	Not included	(2, 3, 2, 2)

Table 14. Classification of 4/21/2005 as high departure count, low-delay due to weather day is in agreement with Ref. 7 except that many aircraft were delayed due to volume on this day. The results in Table 18 for 6/5/2005 and 7/15/2005 are in agreement with Ref. 7. Both these days experienced an inordinate amount of airborne and ground delays due to weather. A number of aircraft were also delayed due to volume, and equipment, runway and other issues on 7/15/2005.

The results of the two examples considered here show that 1) days can be classified into the specified number of groups based on each individual metric, 2) the individual metric group labels can be used for creating multiple-metric group labels, and 3) linguistic description of the individual metric grouping is retained in the composite group label. These examples also illustrate that the multiple-metric classification method does not require that the number of groups be the same based on each metric for creating composite group IDs. In the first example, days were organized in two groups using the total number of departure counts metric and in three groups using the number of aircraft delayed on the ground and the number of aircraft delayed in the air. This technique of maintaining different numbers of groups along different axes of the metric space is in contrast with the method described in Ref. 1 and 2 that only partitions data along the single distance metric.

Results demonstrate that the multiple-metric classification method generates groups with several members and also groups with few members; thus, identifying both nominal and off-nominal days. By selecting a typical day from each group, and then using traffic data corresponding to those days, enough data diversity can be assured for validation of simulations and for Monte Carlo type of benefits analysis of novel air traffic management concepts. Resulting benefits metrics can be weighed with number of members in the group that each day is associated with for estimating overall benefits.

V. Conclusions

Consistent with other studies, analysis of 517 days of National Airspace System (NAS) delay data, which were obtained from the Federal Aviation Administration's Air Traffic Operations Network (OPSNET) database, showed that weather is the predominant causal factor for delays; equipment and runway conditions, and traffic-volume are the other major causal factors. It was also determined the departure and traffic management system delays account for about 86% of the aircraft that are delayed. Ground Delay Program and Ground Stop delay data, also obtained from OPSNET, were combined with the NAS delay data to obtain the number of aircraft delayed on the ground and in the air. The results obtained indicate that on an average 74% of the delayed aircraft are delayed on the ground while only 24% are delayed in the air.

The daily total time delay in minutes was used as a discriminating metric within the K-Means algorithm for partitioning 517 days into ten groups. Mean time delay values associated with the resulting groups, computed using time delay values of the member days, arranged in increasing order were found to be approximately equidistant from the preceding and succeeding mean values. Differences between the standard deviation values associated the groups were also found to be small. Most of the days were assigned to groups with smaller mean time delays. Days with large delays were also identified by the algorithm.

A multiple-metric algorithm was synthesized with the single-metric K-Means algorithm at its core. The technique consists of creating groups using each metric individually as a distance metric within the single-metric K-Means algorithm. Member days are labeled with the group numbers associated with the metrics. Final grouping is achieved by assigning days with a common label to the same group, such that groups are labeled by the same number of indices as the number of metrics. The multiple-metric algorithm was applied to the problem of organizing the 522 days into groups using a) total domestic departure counts, b) number of aircraft delayed on the ground and c) number of aircraft delayed in the air as the three metrics. Two days that were found to be outliers were removed and the remaining 520 days were classified into 18 groups. Six of the 18 groups had six or fewer days as members. Although these groups represent unusual days, their inclusion in a set of days that represents diverse air traffic conditions is essential for evaluating concepts and validating simulations. The other 12 groups had 14 or more days as members. Another example of multiple-metric classification of 522 days into groups with a) total domestic departure counts, b) weather delays, c) volume delays and d) equipment, runway and other delays as the chosen metrics was presented. In this instance, 37 groups out of 54 possible groups had member days. Of the 37 groups, 20 groups had five or fewer days as members. Comparing the results obtained via the two examples, it was seen that different sets of days can be created and certain unusual days can be identified based on the choice of metrics. The two examples serve as illustrations of the ability of the multiple-metric algorithm to create datasets, consisting of days classified into groups, with enough data diversity for concept evaluation and simulation validation.

References

[1]Krozel, J., Hoffman, J., Penny, S., and Butler, T., "Selection of Datasets for NAS-Wide Simulation Validations," Metron Aviation, Inc., 131 Elden St., Suite 200, Herndon, VA 20170, October 2002.

[2]Hoffman, B., Krozel, J., Roy, A.., and Roth, K., "A Cluster Analysis to Classify Days in the National Airspace System," AIAA-2003-5711, *Proceedings of AIAA Guidance, Navigation, and Control Conference,* Austin, TX, August 11-14, 2003.

[3]MacQueen, J. B., "Some Methods for Classification and Analysis of Multivariate Observations," *Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability,* University of California Press, Berkeley, 1967, pp. 281-297.

[4]Krozel, J., Hoffman, B., Penny, S., and Butler, T., "Aggregate Statistics of the National Airspace System," AIAA-2003-5710, *Proceedings of AIAA Guidance, Navigation, and Control Conference,* Austin, TX, August 11-14, 2003.

[5]Federal Aviation Administration, "Order 7210.55C: Operational Data Reporting Requirements," U. S. Department of Transportation, October 1, 2004.

[6]Krozel, J., Capozzi, B., Andre, A. D., and Smith, P., "The Future National Airspace System: Design Requirements Imposed By Weather Constraints," AIAA-2003-5769, *Proceedings of AIAA Guidance, Navigation, and Control Conference,* Austin, TX, August 11-14, 2003.

[7]Zelinski, S., and Meyn, L., "validating The Airspace Concept Evaluation System For Different Weather Days," AIAA 2006-6115, *Proceedings of AIAA Modeling and Simulation Technologies Conference and Exhibit,* Keystone, CO, August 21-24, 2006.

[8]Meyn, L., et al, "Build 4 of the Airspace Concept Evaluation System," AIAA-2006-6110, *Proceedings of AIAA Modeling and Simulation Technologies Conference and Exhibit*, Keystone, Colorado, August 21-24, 2006.

[9]Erzberger, H., Davis, T. J., and Green, S., "Design of Center-TRACON Automation System," *AGARD Conference Proceedings 538: Guidance and Control Symposium on Machine Intelligence in Air Traffic Management*, Berlin, Germany, May 11-14, 1993.

[10]Murphy, J. R., Reisman, R., and Savoye, R., "A Data-Centric Air Traffic Management Decision Support Tool Model," AIAA-2006-7830, *Proceedings of 6th AIAA Aviation Technology, Integration and Operations Conference (ATIO),* Wichita, Kansas, September 25-27, 2006.

[11]Bilimoria, K. D., Sridhar, B., Chatterji, G. B., Sheth, K. S., and Grabbe, S. R., "FACET: Future ATM Concepts Evaluation Tool," *Air Traffic Control Quarterly*, Vol. 9, No. 1, 2001, pp. 1-20.