Federal Aviation Administration

DOT/FAA/AM-07/18
Office of Aerospace Medicine
Washington, DC 20591

Prediction and Classification of Operational Errors and Routine Operations Using Sector Characteristics Variables

Elaine M. Pfleiderer
Carol A. Manning
Civil Aerospace Medical Institute
Federal Aviation Administration
Oklahoma City, OK 73125

July 2007

Final Report

Technical Report Documentation Page

1. Report No.	2. Government Accession No.	3. Recipient's Catalog No.
DOT/FAA/AM-07/18		

4. Title and Subtitle
Prediction and Classification of Operational Errors and Routine Operations Using Sector Characteristics Variables

5. Report Date
July 2007

6. Performing Organization Code

7. Author(s)
Pfleiderer EM, Manning CA

8. Performing Organization Report No.

9. Performing Organization Name and Address
FAA Civil Aerospace Medical Institute
P.O. Box 25082
Oklahoma City, OK 73125

10. Work Unit No. (TRAIS)

11. Contract or Grant No.

12. Sponsoring Agency name and Address
Office of Aerospace Medicine
Federal Aviation Administration
800 Independence Ave., S.W.
Washington, DC 20591

13. Type of Report and Period Covered

14. Sponsoring Agency Code

15. Supplemental Notes
Work was accomplished under approved task AM-B07-HRR-522

16. Abstract
This study examined prediction and classification of operational errors (OEs) and routine operations (ROs) using sector characteristics variables. Average Control Duration, Aircraft Mix Index, Average Lateral Distance, Average Vertical Distance, Number of Handoffs, Number of Point Outs, Number of Transitioning Aircraft, and Number of Heading Changes were used as predictors in two stepwise logistic regression analyses conducted for the high-altitude and low-altitude sectors. In the high-altitude sample, variables included in the final model (Number of Heading Changes, Number of Transitioning Aircraft, and Average Control Duration) accurately classified OE and RO samples for 80% of the cases. In the low-altitude sample, variables included in the final model (Number of Point Outs, the Number of Handoffs, and the Number of Heading Changes) accurately classified OE and RO samples for 79% of the cases. Although logistic regression cannot be used to determine causation, it effectively identified variables that predicted the occurrence of OEs.

17. Key Words	18. Distribution Statement
Air Traffic Control, Sector Characteristics, Complexity, Operational Errors, Logistic Regression Analysis	Document is available to the public through the Defense Technical Information Center, Ft. Belvior, VA 22060; and the National Technical Information Service, Springfield, VA 22161

19. Security Classif. (of this report)	20. Security Classif. (of this page)	21. No. of Pages	22. Price
Unclassified	Unclassified	16	

Form DOT F 1700.7 (8-72) Reproduction of completed page authorized

Contents

Executive Summary

This study examined the ability of selected Performance and Objective Workload Evaluation Research (POWER) measures to predict and classify operational errors (OEs) and routine operations (ROs) using logistic regression analysis. Logistic regression is a statistical technique designed to predict whether an event of interest will or will not occur. Thus, logistic regression is well suited for identifying factors that discriminate between OEs and ROs. An OE database was derived from Systematic Air Traffic Operations Research Initiative (SATORI; Rodgers & Duke, 1993) re-creations of OEs occurring at the Indianapolis Air Route Traffic Control Center (ZID) between 9/17/2001 and 12/10/2003. POWER variables were computed in 5-minute intervals from each OE in the sample (i.e., 4 minutes prior, 1 minute after). Whenever possible, OEs were matched (by sector, day of the week, and time interval) with ROs extracted from System Analysis Recordings (SARs) taped between 2/25/2005 and 3/3/2005. This produced a total of 229 observations (120 OEs, 109 ROs). Variables based on complexity factors identified in a previous analysis (Pfleiderer, Manning, & Goldman, 2007) were computed for both the OE and RO traffic samples. These variables (i.e., Average Control Duration, Aircraft Mix Index, Average Lateral Distance, Average Vertical Distance, Number of Handoffs, Number of Point Outs, Number of Transitioning Aircraft, and Number of Heading Changes) were then used as predictors in two stepwise logistic regression analyses conducted for the high-altitude (N = 161) and low-altitude (N = 68) sector samples.

In the high-altitude sample, variables included in the final model were the Number of Heading Changes, the Number of Transitioning Aircraft, and Average Control Duration. Every heading change that occurred increased the likelihood of an OE by 128%, every transitioning aircraft increased the likelihood of an OE by 26%, and every one-second increase in Average Control Duration increased OE likelihood by 2%. The model was able to accurately classify OE and RO samples for 80% of the high-altitude traffic samples.

In the low-altitude sample, variables included in the final model were the Number of Point Outs, the Number of Handoffs, and the Number of Heading Changes. Every point out that occurred increased the likelihood of an OE by 230%, every handoff increased OE likelihood by 54%, and every heading change increased OE likelihood by 49%. The model was able to accurately classify OE and RO samples for 79% of the low-altitude cases.

Overall, the logistic regression technique was very effective in identifying variables that predicted the occurrence of OEs. Although logistic regression cannot be used to determine causation, the results provide direction for further study. Continued investigations along these lines may highlight complexity factors that should be addressed before adopting changes to the National Airspace System (NAS) to ensure that safety is maintained.

Prediction and Classification of Operational Errors and Routine Operations Using Sector Characteristics Variables

In the current Air Traffic Control (ATC) system, an Operational Error (OE) occurs whenever there is a violation of aircraft separation minima, as defined by the applicable version of Federal Aviation Administration (FAA) Order 7110.65 (FAA, 2006). Although the FAA has called for a reinvention of the National Airspace System (NAS) to keep pace with projected growth in the aviation industry, it is reasonable to assume that future systems will also include some kind of separation standards. Thus, it becomes increasingly important to understand the environmental and contextual factors that contribute to the loss of separation to ensure that increases in capacity are not purchased at the price of safety.

From a research standpoint, OEs have the advantage of being very clearly operationally defined. However, they are problematic when studied from the perspective of the individual controller because "OEs are rare events, relative to the number of controllers working and operations conducted in any given day or year" (Broach & Schroeder, 2005, p. 4). On the other hand, multiple OEs occur in many sectors within any given year. The exact number depends on the characteristics of the particular sector. Thus, studying OEs from a sector perspective not only provides a statistical advantage, but it has a certain logical appeal as well. If airspace characteristics did not contribute to at least a portion of their occurrence, OE frequency would be relatively equal in all sectors. Yet, some sectors are more prone to OEs than others.

Numerous studies have investigated the relationship between sector characteristics and the occurrence of OEs (e.g., Grossberg, 1989; Kershner, 1968; Lowry, MacWilliams, Still, & Walker, 2005; Rodgers, Mogford, & Mogford, 1998; Schroeder, 1982; Spahn, 1977). Most of this work has been done without reference to routine operations (ROs). Yet, for every OE that occurs in a sector, there are hundreds (possibly thousands) of hours in which an OE did not occur. To truly understand the environmental and contextual factors that contribute to OEs, it is necessary to identify what was different about the sector environment at the time the OE occurred. In short, the factors must be able to discriminate between OEs and ROs.

Logistic regression is a statistical tool that develops models to predict whether an event of interest will or will not occur. Although logistic regression is a non-causal analysis (i.e., the extracted predictors are not necessarily the cause of increased odds of an event), the variables included in the model are those that uniquely identify conditions in which the event is likely to occur. Hence, logistic regression is well suited for the task of discriminating between OEs and ROs.

Of course, the first question to be answered is whether a model constructed solely of sector characteristics variables can predict OEs with sufficient accuracy to be useful for practical applications. After all, other factors (e.g., human elements, organizational influences) also contribute to the occurrence of OEs. On the other hand, if such a model could be identified, the included variables would be ideal for incorporation into automation tools that warn of sector conditions in which the risk of an OE is higher.

The second question is whether it is possible to identify a parsimonious set of routinely recorded, easily measured predictors. Obviously, predictor variables that cannot be routinely recorded or those requiring substantial post-processing are impractical for most applications. The quest for parsimony, however, is motivated by statistical concerns. In a recent review, Hilburn (2004) listed in excess of 100 complexity factors purported to be associated with sector complexity, controller workload, or OEs. Viewing the literature, one is left with the impression that sector complexity research is at risk of being crushed under the sheer weight of the potential predictors. Even within individual studies, the number of predictors submitted for analysis often exceeds the limitations of the statistical procedure. For example, in an investigation of the relationship between sector characteristics and OEs at the Atlanta Air Route Traffic Control Center (ARTCC), Rodgers et al. (1998) conducted a linear regression analysis to predict the number of OEs per sector. More than 20 predictors were submitted for analysis with only 45 sectors in the airspace. The authors relied on tests for tolerance in the forward stepwise procedure to exclude variables that failed to contribute a significant amount of unique information to the model. Unfortunately, stepwise procedures require a much higher case-to-predictor ratio to ensure that the solution will generalize beyond the model-building sample (Tabachnick & Fidell, 2006, p. 129). This is certainly not meant as a condemnation of the Rodgers et al. report. (Indeed, the present study is part of the Sector Characteristics and Operational Errors [SCOpE] project, which is an extension of their investigation.) Rather, it illustrates one of the many challenges faced by researchers in this area: to provide a sufficient number of predictors to describe the data without overstepping the limitations of the analysis.

In a previous phase of the SCOpE project (Pfleiderer, Manning, & Goldman, 2007), a sample of 32 controllers and 4 supervisors from Indianapolis ARTCC (ZID) provided ratings for a set of 22 sector complexity factors. Principal Components Analysis (PCA) consolidated the 22 individual complexity factors into four groups (i.e., components). Scores were computed from the four components and used as predictors in a linear multiple regression analysis of the number of OEs in the ZID sectors, thereby circumventing the case-to-predictor ratio problem presented by 22 predictors and 37 sectors. Only Component 1 (climbing and descending aircraft in the vicinity of major airports) and Component 2 (services provided to non-towered airports) contributed significantly to the total proportion of variance explained by the model ($R = .78$, $R^2 = .61$).

The present study evaluates the prediction and classification accuracy of OE and RO traffic samples using objective measures related to the complexity factors identified in Pfleiderer et al. (2007). The basic experimental design concept is simple: Collect variables that objectively describe dynamic elements of the sector environment during OEs and a corresponding set of variables that represent the same elements during ROs and submit them to statistical analysis. The challenge was to develop a parsimonious set of predictors. The list of potential predictors had already been reduced in the previous analysis using PCA and multiple regression. Further reduction of the list could be accomplished by employing a standard of usability. Because Component 2 comprised static variables (e.g., Terrain/Obstructions) and complexity factors that were extremely difficult to collect (e.g., Adequacy of Radio or Radar Coverage), we focused on factors that were most closely associated with Component 1. These factors and their component loadings (i.e., the correlation between the variable and the component) are listed in Table 1.

Performance and Objective Workload Evaluation Research (POWER; Mills, Pfleiderer, & Manning, 2002) refers to a set of objective measures that are calculated from routinely recorded air traffic data. As such, POWER variables met our usability criteria. However, some of the complexity factors associated with Component 1 are not directly measurable (e.g., Amount of Coordination/ Interfacing Required, Number of Multiple Functions, and Number of Required Procedures). Others could not be extracted without a considerable amount of post-processing (e.g., Number of Intersecting Flight Paths, Amount of Radio Frequency Congestion). Still others were static sector characteristics that are unsuitable because of their lack of variability (e.g., Number of Major Airports, Size of Sector Airspace). Nevertheless, many of the factors that were associated with Component 1 could at least be partially measured by one of the POWER variables. These are discussed in the following paragraphs.

The variable with the highest Component 1 loading (.79) was Climbing and Descending Traffic. This complexity factor has long been recognized as a contributor to the difficulty of working a sector (e.g., Arad, 1964). Grossberg (1989) observed that one of the factors most often identified as being responsible for sector complexity in the Chicago ARTCC was climbing and descending flight paths. In his report, he cites an analysis conducted by the FAA Office of Aviation Safety in which 85% of the OEs in 1986 involved at least one transitioning aircraft, whereas most air carrier flights only spent an average of 42% of their flight time in a climb or descent. Kopardekar and Magyarits (2003) found that the number of descending

Table 1. ***Component 1 Factors and Component Loadings from Pfleiderer et al. (in press)***

Complexity Factor	Component Loading
Climbing and Descending Traffic	.79
Coordination/ Interfacing Required	.74
Number of Multiple Functions	.73
Mix of Aircraft Types	.69
Number of Major Airports	.68
Number of Required Procedures	.66
Number of Intersecting Flight Paths	.63
Traffic Volume	.59
Amount of Radio Frequency Congestion	.58
Size of Sector Airspace	.51

aircraft and the number of altitude changes greater than 750 feet per minute both contributed significantly to the explanation of variance in a linear regression model of subjective complexity ratings collected at the Fort Worth, Atlanta, Cleveland, and Denver ARTCCs. On the other hand, a measure of the number of aircraft with altitude changes greater than 500 feet per minute failed to contribute significantly to prediction in a proportional odds logistic regression model of the same sample (Masalonis, Callaham, & Wanke, 2003). Despite that, Masalonis and coworkers concluded that an altitude change metric "may still be operationally desirable" (p. 5) based on the results of structured interviews with Traffic Flow Management personnel. The POWER measure corresponding to Climbing and Descending Traffic is the Number of Transitioning Aircraft. This variable represents the number of aircraft making one or more altitude changes during a given processing interval. To be counted as a change, altitude must increase or decrease by a minimum of 200 feet per 12-second radar update and must continue to change in the same direction for at least three updates.

Coordination/ Interfacing Required had the second highest loading (.74) on Component 1. Coordination between controllers was one of the events selected by Schmidt (1976) for his Control Difficulty Index, even though he considered it to be one of the most difficult to process – with good reason. Coordination is not often recorded. The POWER variable Number of Point Outs (i.e., the total number of point out entries made by the R-side and RA-side controllers during a given processing interval) represents a small part of a far greater whole. Still, it is one of the few instances in which coordination between sectors is documented. Thus, the Number of Point Outs constitutes a partial measure of the Coordination/ Interfacing Required factor.

Mix of Aircraft Types had a relatively high loading (.69) on Component 1. Aircraft mix has often been proposed as one of the traffic characteristics contributing to sector complexity in en route air traffic control (Robertson, Grossberg, & Richards, 1979; FAA, 1984; Grossberg, 1989; Mogford, Murphy, & Guttman, 1994). Although the Aircraft Mix Index (Pfleiderer, 2003) failed to contribute significantly to the explained variance in a linear regression analysis of subjective complexity ratings (Pfleiderer, 2005), it may prove to be a useful predictor of OEs. For calculation of the Aircraft Mix Index, controlled aircraft are assigned aircraft type codes based on designator information recorded by the en route computer system. A half matrix of aircraft type differences is then calculated for all aircraft pairs in the sector at approximately 12-second intervals. The Base1 aircraft mix index is computed by summing the items in these matrices. The Base2 index is the mean of the Base1 values for each minute of data. The Aircraft Mix Index is the mean of the Base2 indexes.

The Number of Major Airports (.68) is a static variable, but there are some aspects of dynamic activity that may be related to the presence of underlying airports. Aircraft tend to converge as they approach a major airport. Therefore, average distances between aircraft should be inversely related to this factor. For the POWER measures Average Lateral Distance and Average Vertical Distance, the lateral and vertical distance between all aircraft pairs is calculated at approximately 12-second intervals. Average Lateral Distance is the mean of the lateral distances for all pairs (in nautical miles), and the Average Vertical distance is the mean of the vertical distances for all pairs (in feet) for any given processing interval.

As mentioned previously, the Number of Required Procedures (.66) is not directly measurable. However, the Number of Heading Changes may be a partial measure of this factor. Heading changes are involved with a number of procedures such as merging and spacing, Standard Terminal Arrival Routes (STARs), Standard Instrument Departure Routes (SIDs), and holds. In addition to its putative relationship with the Number of Required Procedures, this variable has much to recommend it as a predictor. Heading changes have demonstrated a relationship with controller ratings of activity (e.g., Laudeman, Shelden, Branstrom, & Brasil, 1998) workload (e.g., Stein, 1985), and complexity (e.g., Kopardekar & Magyarits, 2003). The POWER variable Number of Heading Changes is a count of all turns in excess of 10° per 12-second radar update that continue in the same direction for at least three updates.

Traffic Volume (.59) is related to two POWER variables, the Number of Controlled Aircraft and the Maximum Number of Aircraft Controlled Simultaneously. However, previous research suggests that traffic volume is not an effective predictor of OEs at the sector level (e.g., Lowry, et al. 2005; Schroeder, 1982; Spahn, 1977). In addition, so many aspects of complexity are correlated with traffic volume that inclusion of either POWER variable might overshadow the contributions of more effective predictors. Therefore, the POWER variable, Number of Handoffs (i.e., handoffs with complete initiate and accept message pairs), was selected as a substitute for traffic volume. Handoffs are highly correlated with the number of aircraft in the sector, plus they provide other information as well. While aircraft counts only describe the number of aircraft in the sector, handoff counts describe the number of aircraft entering and exiting the sector. As a result, handoffs capture an element of movement that is missing from aircraft counts. Handoffs may also reflect the impact of sector geography. According to Couluris and Schmidt (1973), the number of handoffs, coordination, and point outs "result from, or are influenced by, the existence and design (shape) of the sectors. The additional work created can be thought of as the cost of sectorization" (p. 657). Although most handoffs are fairly automatic, some require coordination with other sectors. In some instances, aircraft must comply with altitude or other restrictions before they can be handed off. Thus, the Number of Handoffs may provide supplemental information about coordination and required procedures.

Size of Sector Airspace (.51) has demonstrated a relationship with OEs in several studies (e.g., Goldman et al., 2006; Lowry et al., 2005; Rodgers et al., 1998). Unfortunately, sector size is a static variable. Fortunately, it is related to the dynamic variable control duration. Although the amount of time an aircraft remains within a sector is influenced by other factors, such as aircraft characteristics, Average Control Duration reflects dynamics associated with sector size. In a ZID airspace sample consisting of 168 hours of continuous data, the POWER variable Average Control Duration was significantly correlated with sector size in 23 high-altitude ($r = .60, p<.01$) and 9 low-altitude sectors ($r = .82, p<.01$). Thus, Average Control Duration is strongly associated with sector size. The POWER variable Average Control Duration is the mean of durations (in seconds) of all aircraft controlled by the sector within a processing interval. Control time occurring before or after the interval is not included in the calculations.

The POWER variables just described were submitted as possible predictors in two stepwise logistic regression analyses conducted for the high-altitude and low-altitude sector samples. The results of these analyses will determine whether a parsimonious set of routinely recorded sector characteristics variables can discriminate between OE and RO traffic samples.

Method

Data Extraction

Both the OE and RO data used in the analyses were initially derived from System Analysis Recordings (SAR) generated by en route Host computer systems. SAR data are written in Jovial, a binary computer language. Because humans find it difficult to read binary output, the Host system features data reduction programs that generate text reports of selected subsets of SAR data. The information used to calculate the predictor variables was extracted from log and track reports produced by one of these programs, the Data Analysis and Reduction Tool (DART). Log reports include controller entries and information sent to the radar display and the auxiliary text display (e.g., data blocks and list items). Track reports contain detailed information from the Host computer's internal radar track database (e.g., altitude, heading, ground speed, and position).

The DART log and track text reports of the RO data were first encoded into database files by the NAS Data Management System (NDMS) and then predictor variables were computed by the POWER software system. For more information about the NDMS and POWER programs, see Mills, Pfleiderer, and Manning (2002).

OE data were derived from Systematic Air Traffic Operations Research Initiative (SATORI; Rodgers & Duke, 1993) files. The primary constraint on the size and range of the data set was the availability of SATORI re-creations. SATORI data meeting interval processing criteria (i.e., four minutes prior and one minute after the OE) were only available for 120 OEs that occurred in the ZID airspace from 9/17/2001 through 12/10/2003. SATORI samples were matched (by sector, day of the week, and time interval) with RO data recorded from 2/25/2005 to 3/3/2005. Sector characteristics were extracted from ZID's Adaptation Control Environmental Systems (ACES) files using the OpenCreate software package (part of the SATORI system). Examination of the extracted characteristics (e.g., cubic area, number of shelves, number of VORTACs, number of miles of jet ways and airways, number of intersections) verified that the basic structure of the sectors remained consistent between the OE and RO sample time frames. Nine of the OEs could not be matched because the sector was closed or combined with another sector during the target interval. Two cases had to be excluded because only one aircraft was controlled by the sector, resulting in missing data for distances and the Aircraft Mix Index. The final dataset consisted of 229 observations (120 OEs, 109 ROs).

Materials

Although SATORI files are formatted to optimize graphical re-creations of incidents, all the information necessary to calculate the POWER measures is still accessible. The POWERsatori program was developed to compute objective measures from this information. The calculations for POWER and corresponding POWERsatori measures are identical, with the exception of lateral distances. The SATORI system converts the stereographic coordinates recorded in the DART track report to latitudes and longitudes. The distance equations in the POWERsatori program were adapted to accommodate this difference.

Results and Discussion

Descriptive statistics listed in Table 2 clearly demonstrate that several of the variables were not normally distributed. For example, the distribution of the Number of Point Outs differed from a normal distribution by 14 standard deviations in skewness and 22 standard deviations in kurtosis. The Aircraft Mix Index deviated by 34 standard deviations in skewness and more than 129 standard deviations in kurtosis. Beyond theoretical issues surrounding the use of frequency data in parametric statistics, such extreme departures from normality necessitated the use of a non-parametric analysis.

Logistic Regression Analysis

Although other techniques are available for group prediction (e.g., discriminant function analysis) and for identifying predictive variables (e.g., linear regression), the assumptions required for reliable results limit their usefulness. Logistic regression is a "distribution-free" alternative to these procedures. Predicted group membership in logistic regression is comparable to case classification in discriminant function analysis. Logistic regression also provides measures of multiple associations and significance tests for the contribution of the predictor variables similar to linear regression models. However, the logistic model is expressed in terms of odds (i.e., the ratio of the probability of an event to the probability the event will not occur).

Stepwise methods are extremely useful in exploratory analysis, particularly when the focus is on identifying predictors. Unfortunately, forward stepwise procedures are notorious for their tendency to omit useful variables. Backward elimination is less prone to these effects because all variables are in the model at the beginning of the process (Menard, 1995; see also Agresti & Finlay, 1986). Therefore, backward stepwise elimination was selected for the high- and low-altitude sector analyses. The likelihood-ratio test was used as the selection criterion because it is more rigorous than other methods (Menard, 1995; Norušis, 1990; Pampel, 2000).

Two caveats must be considered when viewing the results. First and foremost, prediction is not the same as causation. Variables in the logistic regression model may be excellent predictors of OEs, but this does not mean they cause or even contribute to the occurrence of OEs. In other words, just because the model includes the Number of Transitioning Aircraft as a predictor does not imply that eliminating all altitude changes could reduce OEs. The results merely indicate that the predictor variables in the logistic regression model are associated, in varying degrees, with the *likelihood* of an OE. Second, logistic regression weights are not standardized. This makes comparisons between predictors problematic. A one-unit change in Control Duration (one second) is not necessarily comparable to a one-unit change in the Number of Transitioning Aircraft. The logistic regression weights and odds ratios do not adjust for such differences.

Logistic Regression Analysis of High-Altitude Sectors. Logistic regression is a "distribution-free" statistic, but that does not mean it is assumption free. As with other forms of regression, multicollinearity among the predictors can lead to biased estimates (Menard, 1995). Several of the Spearman's correlations shown in Table 3 are statistically significant, but it is doubtful that any are of sufficient magnitude to have a negative impact on the analysis.

One of the main purposes of this analysis was to determine how accurately the logistic regression model can differentiate between OEs and ROs. Of the 79 ROs in the sample, 64 (81%) were correctly classified and 15 were misclassified as OEs. Of the 82 OEs in the sample, 65 (79%) were correctly classified and 17 were misclassified as ROs. The final model had an 80% overall classification accuracy, which represents an improvement

Table 2. ***Descriptive Statistics (N = 229)***

Variable	Mean	*SD*	Skew.[1]	Kurtosis[2]
Average Control Duration (seconds)	194.86	38.83	-.31	.26
Aircraft Mix Index	3.77	9.87	5.50	41.28
Average Lateral Distance (nm)	50.23	14.91	.25	.43
Average Vertical Distance (ft)	44.02	24.36	1.65	3.43
Number of Handoffs	4.83	2.87	.62	.02
Number of Point Outs	.86	1.28	2.23	6.94
Number of Transitioning Aircraft	3.76	2.25	.74	.45
Number of Heading Changes	2.14	1.96	1.36	2.61

[1]SE Skew. = .161; [2]SE Kurt. = .320

of approximately 30% over prior probabilities (i.e., the number of cases that might be correctly classified by chance).

Perhaps the most important information to be gained from the logistic regression analysis is a list of variables that distinguish between OE and RO traffic samples. The logistic regression coefficient *(B)*, its standard error, the estimated odds ratio (Odds), and likelihood-ratio test output (Model if Removed) for each variable are provided in Table 4. The estimated odds ratio is interpreted as the change in odds associated with a one-unit change in the predictor variable. The logistic regression coefficients are the natural logs of the odds ratios. The likelihood-ratio is computed by comparing the fit of the model with and without each predictor.

The first variable excluded from the model was the Number of Handoffs, followed by Average Vertical Distance, Average Lateral Distance, Aircraft Mix Index, and the Number of Point Outs. Variables included in the final model were the Number of Heading Changes, the Number of Transitioning Aircraft, and Average Control Duration. It is clear from the odds ratios in Table 4 that the Number of Heading Changes was the most influential variable in the model. For every heading change that occurred, the probability of an OE was increased by a multiplicative factor of 2.28 (i.e., increased the odds by $e^{2.28}$). In simpler terms, every heading change increased the likelihood of an OE by 128%. However, it is important to remember that this is almost certainly an overestimate. Heading changes issued in an attempt to avoid the OE could not be excluded from the sample. The same is true of the Number of Transitioning Aircraft, though to a lesser degree because, as a summary measure, it is not as sensitive as the Number of Heading Changes.

The 1.02 odds ratio associated with Average Control Duration might seem small when compared with the other predictors in the model, but keep in mind that this measure is based on the average number of seconds each

Table 3. ***Correlation* Matrix: High-Altitude Sectors (N=161)***

	1	2	3	4	5	6	7
1 Control Duration							
2 Aircraft Mix Index	-.04						
3 Lateral Distance	-.18*	-.18*					
4 Vertical Distance	-.09	.05	-.10				
5 Number of Handoffs	-.41**	.16	-.06	.04			
6 Number of Point Outs	.03	.12	-.11	.01	.08		
7 Transitioning Aircraft	.02	.30**	-.32**	.28**	.40**	.20*	
8 Heading Changes	.22**	.16*	-.34**	.04	.18*	.18*	.43**

* Spearman's rho; ***p* < .01; **p* < .05

Table 4. ***Logistic Regression Variable Summary: High-Altitude Sectors (N = 161)***

Variable	*B*	*S.E.*	Odds	Model if Removed: Model Log Likelihood	Model if Removed: Change in -2LL
Control Duration	.02	.01	1.02	-80.37	11.51**
Transitioning Aircraft	.23	.12	1.26	-76.68	4.13*
Heading Changes	.82	.18	2.28	-91.77	34.33**
Constant	-5.94	1.33	.00		

** *p* < .01; * *p* < .05

Table 5. ***High-Altitude Sample Cases***

Case Number	Avg. Control Duration (secs.)	Transitioning Aircraft	Heading Changes	Predicted Probability
1	90	0	0	.02
2	239	4	1	.59
3	198	3	5	.94
4	252	7	8	.99

aircraft was controlled by the sector. Thus, a minimal change in Average Control Duration produced a relatively large 2% change in the likelihood of an OE.

To demonstrate how the logistic regression weights and their associated odds ratios translate to actual events, consider examples from the data shown in Table 5. Case 1 is a traffic sample in which there were low values for every variable in the model. Consequently, the predicted probability (.02) was also very low. On the other hand, Case 4 had high values for all predictors. Thus, Case 4 had a very high predicted probability (.99). Heading Changes increase the probability of an OE at a greater rate than the other predictors and so Case 2, with only one Heading Change, had a moderate predicted probability (.59), even though the values of the other predictors were high. Similarly, Case 3 had a very high predicted probability (.94), even though the values of most predictors were comparatively low due to the Number of Heading Changes associated with this traffic sample.

Because there were considerably fewer OEs in the super high-altitude sectors (n = 34), they were combined with the high-altitude sectors (n = 127). However, the fact that there *were* fewer OEs in the super high-altitude sectors suggests that there might be fundamental differences between the two sector types. A logistic regression analysis was performed to test the effects of removing cases involving super high-altitude sectors from the sample. The results were identical to those of the combined sample except for a slight improvement in classification accuracy (86% RO; 78% OE; 82% overall). This does not necessarily mean that super high-altitude sectors were identical to high-altitude sectors with regard to predictors of OEs. (In fact, the improvement in classification suggests the opposite.) There simply were not enough super high-altitude cases to have a significant impact on the results.

Logistic Regression Analysis: Low-Altitude Sectors. As shown in Table 6, the correlation between Number of Handoffs and the Number of Transitioning Aircraft (r_s=.64) is of sufficient magnitude to potentially impact the analysis. The correlation between the Number of Transitioning Aircraft and the Number of Heading Changes (r_s=.54) is also fairly high. These relationships will need to be addressed when interpreting the results.

Classification accuracy in the low-altitude sample was comparable to that of the high-altitude sample. Of the 30 ROs in the low-altitude sample, 23 (77%) were correctly classified and 7 were misclassified as OEs. Of the 38 OEs in the sample, 31 (82%) were correctly classified and 7 were misclassified as ROs. The 79% overall classification accuracy of the final model represents approximately a 30% improvement over chance.

The variable summary for the low-altitude sectors is provided in Table 7. In the low-altitude sample model, the Number of Point Outs had the highest odds ratio (3.30), followed by the Number of Handoffs (1.54) and the Number of Heading Changes (1.49). In other words,

Table 6. ***Correlation* Matrix: Low-Altitude Sectors (N = 68)***

	1	2	3	4	5	6	7
1 Control Duration							
2 Aircraft Mix Index	.32*						
3 Lateral Distance	.05	.08					
4 Vertical Distance	-.01	.05	.37**				
5 Number of Handoffs	-.37**	.33**	-.09	-.21			
6 Number of Point Outs	.24	.25*	-.35**	-.24	.10		
7 Transitioning Aircraft	.06	.44**	-.15	-.34**	.64**	.24	
8 Heading Changes	.30*	.36**	-.10	-.35**	.19	.32**	.54**

* Spearman's rho; **p < .01; *p < .05

Table 7. ***Logistic Regression Variable Summary: Low-Altitude Sectors (N = 68)***

Variable	*B*	*S.E.*	Odds	Model if Removed: Model Log Likelihood	Model if Removed: Change in -2LL
Number of Handoffs	.43	.16	1.54	-37.26	9.07**
Number of Point Outs	1.19	.46	3.30	-37.61	9.76**
Heading Changes	.40	.20	1.49	-34.77	4.08*
Constant	-2.90	.88	.06		

** p < .01; * p < .05

each point out increases the likelihood of an OE in the low-altitude sectors by 230%, each handoff increases OE likelihood by 54%, and each heading change increases OE likelihood by 49%. Variables excluded from the model (in order of their removal) were: the Number of Transitioning Aircraft, Average Control Duration, Average Vertical Distance, Average Lateral Distance, and the Aircraft Mix Index.

Because the Number of Handoffs was strongly associated with the Number of Transitioning Aircraft in the low-altitude sample (r_s .64, $p < .01$), a diagnostic stepwise analysis was conducted to evaluate the effects of removing the Number of Handoffs as a predictor. As expected, the Number of Transitioning Aircraft was included in the final logistic regression model when its correlate was eliminated from the variable set. The Number of Point Outs remained the dominant predictor, but the Number of Transitioning Aircraft made a significant contribution (Odds = 1.53). The Number of Heading Changes was excluded from the predictor set, possibly because it is also significantly correlated with the Number of Transitioning Aircraft (r_s=.54, $p < .01$). It should be noted that classification accuracy was considerably reduced in the test model.

Both the Number of Handoffs and the Number of Transitioning Aircraft were correlated with the Total Number of Controlled Aircraft at the <.01 level of significance (r_s =.72 and r_s =.74, respectively). Therefore, a second diagnostic stepwise analysis was conducted to evaluate the effects of including the Total Number of Controlled Aircraft in the set of potential predictors. The model remained identical to the original shown in Table 7. Apparently, the relationship between the Number of Handoffs and the Number of Transitioning Aircraft is not due to their shared association with the Total Number of Controlled Aircraft. Moreover, both contributed significantly to the prediction of OEs, whereas the Total Number of Controlled Aircraft did not.

In spite of the reputed relationship between aircraft mix and sector complexity, the Aircraft Mix Index failed to be included in the final model. However, the Aircraft Mix Index *was reliably different* between OEs and ROs. In sharp contrast to the high-altitude sample, very few Aircraft Mix Index values were zero. Differences between the performance characteristics of aircraft in the low-altitude sectors were generally larger and more varied than in high-altitude sectors. This was particularly true in the OE traffic samples. The mean of the Aircraft Mix Index for the RO traffic samples was 5.16 (*S.D.*=7.74). The mean of the Aircraft Mix Index for the OE traffic samples was 13.81 (*S.D.*=18.29). Results of a t-test for independent samples showed this difference to be statistically significant, $t(51.52) = -2.57$, $p < .05$.[1] Therefore, it might be useful to consider aircraft mix as a possible complexity factor when description, rather than prediction, is the goal.

Conclusions

The results of logistic regression analyses of high- and low-altitude sectors were beneficial because they identified variables that reliably discriminate between OEs and ROs. However, it must be re-emphasized that logistic regression is a non-causal analysis. The predictors included in the high- and low-altitude logistic regression models are most certainly associated with OEs, but that does not mean they cause them. Nevertheless, the results represent an important step toward uncovering elements of the sector environment that contribute to the occurrence of OEs. The next step involves examining the cases for which the logistic regression model was not an adequate fit (i.e., misclassified cases). OEs that failed to be identified as such might have been due to weather influences, interactions with static sector characteristics, or the impact of traffic characteristics that were not included in the set of predictors. ROs that were mistakenly classified as OEs might represent instances where controller experience or other human factors overcame environmental conditions. These cases represent an opportunity to recognize and learn from exceptional performance.

It is important to remember that stepwise selection methods tend to be somewhat restrictive. Depending on their intended function, it may be practical to incorporate some of the variables that were excluded from the models. If the intention is development of an automation tool to alert supervisors that an OE is likely to occur, then exclusive use of the variables included in the final model is almost certainly the best choice. On the other hand, a broader list (i.e. one that incorporates variables excluded because they were extremely correlated with others in the set of predictors, or those that were retained until the final step) might be beneficial when seeking to identify the static sector characteristics that interact with dynamic predictors. Highly correlated variables may be inadvisable for prediction algorithms but extremely valuable for identifying airspace design elements that could be changed to reduce the risk of OEs.

One fact made evident by the results is that altitude strata differ with regard to factors related to OEs. Accordingly, the conclusions drawn from our analyses are addressed separately in the following sections.

[1] Levene's test for equality of variances was significant (F= 5.50, p<.05), so the adjusted t is reported.

High-Altitude Sectors

Classification accuracy of the high-altitude sector sample was mildly disappointing, but not entirely unexpected. There will always be a proportion of OEs that cannot be explained using contextual measures alone. Nevertheless, only 30% improvement over chance suggests that essential sector characteristics may have been missing from the analysis. In our recent examination of controllers' subjective ratings (Pfleiderer, Manning, & Goldman, 2007), Intersecting Flight Paths was one of the complexity factors most closely associated with the number of OEs in the ZID sectors. It is unfortunate that we were unable to incorporate this variable into the POWER measures for these analyses, but it certainly will be in future analyses.

The variables included in the high-altitude logistic regression model were the Number of Heading Changes, the Number of Transitioning Aircraft, and Average Control Duration. That longer Average Control Durations were associated with the occurrence of OEs in high-altitude sectors was initially confusing. In the high-altitude sectors at ZID, Average Control Duration is significantly positively associated with sector size ($r = .60$, $p < .01$). This contradicts the results of Goldman, Manning, and Pfleiderer (2006), Lowry et al. (2005), and Rodgers et al. (1998) in which smaller sectors were associated with a higher incidence of OEs. It is not possible to determine why the Lowry et al. and Rodgers et al. results contradict the findings of this study, but the Goldman et al. effect was caused by combining high-altitude and super high-altitude sectors into a single group. Super high-altitude sectors in the ZID airspace tend to be larger than high-altitude sectors and have significantly fewer OEs. Unfortunately, this led to the misperception that smaller high-altitude sectors were more prone to have OEs than larger ones. When sector strata were recoded into three separate groups, the sector size effect was no longer statistically significant. Fortunately, excluding super high-altitude sector cases had no impact on the logistic regression model in the present analysis, although prediction accuracy increased to 82%. This suggests that the model does not fit the super-high altitude cases as well as the high-altitude ones. It is regrettable that the limited sample size of the super high-altitude sectors precluded conducting a separate analysis to identify a unique set of predictors for this sector type.

The Number of Heading Changes was undoubtedly the strongest predictor of OEs in the high-altitude sample. Granted, a portion of its predictive strength may be an artifact of the OE, but if it were a substantial proportion, similarly high odds ratios should have been observed in the low-altitude sector model as well. This simply was not the case. Therefore, it is safe to assume that the Number of Heading Changes genuinely is an effective predictor of OEs in the high-altitude sector stratum. This is appealing on theoretical grounds because Kopardekar and Magyarits (2003) found heading changes to be associated with controllers' subjective sector complexity ratings. Thus, the relationship between the Number of Heading Changes and OEs in this analysis provides indirect support for the hypothesis that OEs are related to sector complexity.

The variables included in the high-altitude logistic regression model (i.e., Heading Changes, Transitioning Aircraft, and Control Duration) are consistent with controller and supervisor ratings collected at ZID (Pfleiderer et al., 2007) The top-ranked factors for the high-altitude sectors were Climbing and Descending Traffic, Traffic Volume, Traffic Management Initiatives (TMI), the Number of Intersecting Flight Paths, and Major Airports. Although the POWER variable Number of Transitioning Aircraft corresponds directly to the complexity factor Climbing and Descending Traffic, other communalities are more subtle. TMI relates to Average Control Duration because initiatives to keep aircraft spaced, either by minutes or miles-in-trail, would increase the amount of time an aircraft remains in the sector. For high-altitude sectors adjacent to low-altitude arrival sectors, adherence to TMI directives might require holds that would, in turn, relate to the Number of Heading Changes.

Low-Altitude Sectors

As with high-altitude sectors, classification accuracy for the low-altitude sector sample was less than ideal. Perhaps low-altitude sectors are too heterogeneous to be analyzed as a single group. Case in point: A component based on controllers' ratings of complexity factors associated with sectors providing services to non-towered airports shared an inverse relationship with the number of OEs (Pfleiderer et al, 2007). If one type of low-altitude sector is this distinctive, then others may also have unique characteristics that should be addressed independently.

The low-altitude sector logistic regression model comprised the Number of Point Outs, the Number of Handoffs, and the Number of Heading Changes. Heading changes were far less influential in the low-altitude sector model: the likelihood of an OE increased by only 49% (as compared to the 138% increase observed in the high-altitude sector model). The fact that the Number of Heading Changes was excluded when the Number of Transitioning Aircraft was included in the test model casts doubt on its validity as a predictor in low-altitude sectors. When controllers become aware that an OE is developing, they may issue altitude and/or heading clearances in an attempt to resolve the situation. The portion of variance these measures share may reflect these actions.

The Number of Point Outs was clearly the strongest predictor of OEs in the low-altitude sector model. The Number of Handoffs, the second most influential predictor, also has an element of coordination. However, there is another aspect of the sector environment to which both variables relate — sector boundaries. As Couluris and Schmidt (1973) observed, handoffs and point outs "result from, or are influenced by, the existence and design (shape) of the sectors. The additional work created can be thought of as the cost of sectorization" (p. 657). Their study was conducted in response to the rapid increase in the number of en route sectors between 1962 and 1972 and the corresponding increase in the controller workforce. The results of the logistic regression analysis indicate that the cost/benefit tradeoff associated with point outs and handoffs at sector boundaries may have been even more expensive than Couluris and Schmidt assumed. Every point out increased the likelihood of an OE in low-altitude sectors by 230%. Every handoff increased OE likelihood by 54%. This does not mean that OEs are caused by point outs or handoffs, nor does it imply that point outs, handoffs, or even sector boundaries should be eliminated. In the Pfleiderer et al. (2007) study, controllers and supervisors at ZID rated coordination as one of the primary sources of complexity in low-altitude sectors, but this factor was ranked third (after the Number of Transitioning Aircraft and Aircraft Mix). The results of the logistic regression suggest that certain types of coordination may have a greater impact on the occurrence of OEs in low-altitude sectors than the ratings indicated.

In closing, because of the research that remains to be accomplished (e.g., analysis of misclassified cases), these results must be viewed as preliminary. However, the methodology employed is promising. The usability of the predictor variables makes them suitable for a number of applications within the Indianapolis airspace. Similar analyses must be conducted at other facilities before they can be recommended for general use. Although logistic regression cannot be used to determine causation, the results provide direction for further study. Continued investigations along these lines may highlight complexity factors that should be addressed before adopting changes to the NAS to ensure that safety will be maintained.

References

Agresti, A., & Finlay, B. (1986). *Statistical Methods for the Social Sciences* (2nd ed.). San Francisco: Dellen.

Arad, B.A. (1964, May). The control load and sector design. *Journal of Air Traffic Control*, 12-31.

Broach, D., & Schroeder, D.J. (2005). *Relationship of air traffic control specialist age to en route operational errors.* (Report No. DOT/FAA/AM-05/22). Washington, DC: FAA Office of Aerospace Medicine.

Couluris, G.J., & Schmidt, D.K. (1973). Air traffic control jurisdictions of responsibility and airspace structure. *Conference on Decision and Control, 4th Annual Symposium on Adaptive Processes,* 655-658. New York: Institute of Electrical and Electronics Engineers.

Federal Aviation Administration (2006). *Air Traffic Control.* (Order No. 7110.65R). Available online at: www.faa.gov/ATpubs.

Federal Aviation Administration (1984). *Establishment and validation of en route sectors.* (FAA Order No. 7210.46). Washington, DC: Author.

Goldman, S., Manning, C., & Pfleiderer, E. (2006). *Static sector characteristics and operational errors.* (Report No. DOT/FAA/AM-06/1). Washington, DC: FAA Office of Aerospace Medicine.

Grossberg, M. (1989, April). Relation of sector complexity to operational errors. *Quarterly report of the Federal Aviation Administration's Office of Air Traffic Evaluations and Analysis.* Washington, DC: Federal Aviation Administration.

Hilburn, B. (2004). *Cognitive complexity in air traffic control – A literature review.* (EEC Note No. 04/04). Brétigny-sur-Orge, France: EUROCONTROL.

Kershner, A.M. (1968). *Air traffic control system error data for 1965 and 1966 as related to age, workload, and time-on-shift of involved controller personnel* (Report No. BA-68-32). Atlantic City, NJ: FAA NAFED.

Kopardekar, P. & Magyarits, S. (2003, June). *Measurement and prediction of dynamic density.* Paper presented at the 5th Eurocontrol/FAA Air Traffic Management Research and Development Conference. Budapest, Hungary.

Laudeman, I.V., Shelden, S.G., Branstrom, R., & Brasil, C.L. (1998). *Dynamic density: An air traffic management metric*. (Report No. NAS/TM-1998-112226.) Moffett Field, CA: NASA Ames.

Lowry, N.M., MacWilliams, K.J., Still, R.J., & Walker, M.G. (2005). *Analysis of operational errors.* Unpublished Report (MP05W0000025, Rev. 1), MITRE Corporation.

Masalonis, A.J., Callaham, M.B., & Wanke, C.R. (2003, June). *Dynamic density and complexity metrics for realtime traffic flow management.* Paper presented at the 5th Eurocontrol/FAA Air Traffic Management Research and Development Conference. Budapest, Hungary.

Menard, S. (1995). *Applied logistic regression.* (Sage University Papers Series on Quantitative Applications in the Social Sciences, No. 106). Thousand Oaks, CA: Sage.

Mills, S.H., Pfleiderer, E.M., & Manning, C.A. (2002). *POWER: Objective activity and taskload assessment in en route air traffic control.* (Report No. DOT/FAA/AM-02/2). Washington, DC: FAA Office of Aerospace Medicine.

Mogford, R.H., Murphy, E.D., & Guttman, J.A. (1994). Using knowledge exploration tools to study airspace complexity in air traffic control. *International Journal of Aviation Psychology,* 4(1), 29-45.

Norušis, M.J. (1990). *SPSS advanced statistics student guide.* Chicago: SPSS.

Pampel, F.C. (2000). *Logistic Regression: A Primer.* (Sage University Papers Series on Quantitative Applications in the Social Sciences, No. 132). Thousand Oaks, CA: Sage.

Pfleiderer, E.M. (2003). *Development of an empirically-based index of aircraft mix.* (Report No. DOT/FAA/AM-03/8). Washington, DC: FAA Office of Aerospace Medicine.

Pfleiderer, E.M. (2005). *Relationship of the aircraft mix index with performance and objective workload evaluation research measures and controllers' subjective complexity ratings.* (Report No. DOT/FAA/AM-05/16). Washington, DC: FAA Office of Aerospace Medicine.

Pfleiderer, E.M., Manning, C.A., & Goldman, S.M. (2007). *Relationship of complexity factor ratings and operational errors.* (Report No. DOT/FAA/AM-07/11). Washington, DC: FAA Office of Aerospace Medicine.

Robertson, A., Grossberg, M., & Richards, J. (1979). *Validation of air traffic controller workload models.* (Report No. DOT/FAA/RD-79/83). Cambridge, MA: U.S. Department of Transportation Research and Special Programs Administration – Volpe National Transportation System Center.

Rodgers, M.D., & Duke D.D. (1993). *SATORI: Situation assessment through the re-creation of incidents.* (Report No. DOT/FAA/AM-92/12). Washington, DC: Office of Aviation Medicine.

Rodgers, M., Mogford, R., & Mogford, L. (1998). *The relationship of sector characteristics to operational errors* (Report No. DOT/FAA/AM-98/14). Washington, DC: FAA Office of Aviation Medicine.

Schmidt, D.K. (1976). On modeling ATC workload and sector capacity. *Journal of Aircraft*, 13(7), 531-537.

Schroeder, D.J. (1982, October). The loss of prescribed separation between aircraft: How does it occur? *Proceedings of the Behavioral Objectives in Aviation Automated Systems Symposium, Society of Automotive Engineers (SAE), P-114,* 257-269.

Spahn, M.J. (1977). *The human element in air traffic control: Observations and analyses of the performance of controllers and supervisors in providing ATC separation services. Supplement 2: Analysis of the System Effectiveness Information System (SEIS) data base.* (Report No. MTR-7655, Supplement 2). McLean, VA: METREK Division of the MITRE Corporation.

Stein, E. (1985). *Air traffic controller workload: An examination of workload probe.* (Report No. DOT/FAA/CT-TN84/24). Atlantic City, NJ: Federal Aviation Administration.

Tabachnick, B.G., & Fidell, L.S. (2006). *Using multivariate statistics* (5th ed.). New York: Harper Collins.