2009

Logistic Regression Analysis of Operational Errors and Routine Operations

Elaine M. Pfleiderer

Cheryl L. Scroggins

Carol A. Manning

LOGISTIC REGRESSION ANALYSIS OF OPERATIONAL ERRORS AND ROUTINE OPERATIONS

Elaine M. Pfleiderer
Cheryl L. Scroggins
Carol A. Manning
FAA Civil Aerospace Medical Institute
Oklahoma City, Oklahoma

Two separate logistic regression analyses were conducted for low- and high-altitude sectors to determine whether a set of dynamic sector characteristics variables could reliably discriminate between operational error (OE) and routine operation (RO) traffic samples. Dynamic sector characteristics submitted as predictors were: Average Control Duration, Number of Handoffs, Number of Heading Changes, Number of Intersecting Flight Paths, Number of Point Outs, and Number of Transitioning Aircraft. In the low-altitude sector model, the Number of Intersecting Flight Paths, the Number of Point Outs, and the Number of Handoffs produced a 75% overall classification accuracy. In the high-altitude sector model, the Number of Intersecting Flight Paths, the Number of Heading Changes, the Number of Transitioning Aircraft, and Average Control Duration produced a 79% overall classification accuracy. Classification rates achieved through the use of the selected sector characteristics support the assumption that elements of the sector environment contribute to the occurrence of OEs.

A considerable amount of research has focused on the relationship between sector characteristics and controller workload or perceived complexity. However, relatively few studies have examined the relationship between sector characteristics and the occurrence of OEs. In many early studies of OE causal factors, examinations of sector characteristics were limited to purely theoretical relationships (e.g., Arad, 1964) or to traffic counts and altitude transitions of the involved aircraft (e.g., Schroeder, 1982; Spahn, 1977). Grossberg (1989) expanded on this by collecting ratings from 97 controllers and supervisors regarding various aspects of the sector environment in the Chicago Air Route Traffic Control Center (ARTCC). Rodgers, Mogford, and Mogford (1998) evaluated the relationship between sector characteristics and the incidence of OEs at the Atlanta ARTCC. In both the Grossberg (1989) and Rodgers, Mogford, and Mogford (1998) studies, sector characteristics were evaluated without comparison with 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. Variables that correlate with sector OE frequency do not describe what was different about the sector at the time of the OE. 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.

Pfleiderer and Manning (2007) conducted an investigation to determine whether logistic regression analysis of objective sector characteristics could discriminate between OE and RO traffic samples. Two separate logistic regression analyses were performed for low- and high-altitude sector samples at the Indianapolis ARTCC (ZID). In the low-altitude sector sample, variables included in the final model were the Number of Point Outs, the Number of Handoffs, and the Number of Heading Changes. This model was able to accurately classify 79% of the low-altitude OE and RO traffic samples. In the high-altitude sector sample, a logistic regression model comprising the Number of Heading Changes, the Number of Transitioning Aircraft, and Average Control Duration was able to accurately classify 80% of the OE and RO traffic samples. Unfortunately, the study was flawed. Available traffic data consisted of OEs from 9/17/2001 to 12/10/2003 and ROs from 2/25/2005 to 3/3/2005. Clearly, the time differential between the OE and RO traffic samples was a confounding influence because it represented an uncontrolled, systematic difference between the two groups. A second problem with the design involved pairing OE and RO traffic samples (by sector, day of week, and time of day). Logistic regression analysis assumes that all cases are independent of one another. Only random selection of RO traffic samples would have guaranteed independence.

In the present study, OE and RO traffic samples are again compared using logistic regression analysis, but some important modifications were made to the design. OEs occurring in ZID airspace between 2001 and 2003 were compared with RO traffic samples from 2003, thereby reducing the time differential between the OE and RO groups. No attempt was made to match the RO traffic samples to the OE samples, thus meeting the assumption of independence.

Separate logistic regression analyses were conducted for the low- and high-altitude sector samples because there was reason to suspect they represent heterogeneous sub-samples (Pfleiderer & Manning, 2007). Therefore, combining sector strata would probably produce a model that fit the high-altitude sectors poorly and the low-altitude sectors not at all. Predictor variables were restricted to dynamic sector characteristics. The variance of static variables would be seriously limited because multiple OEs occurred in many of the same sectors in the sample. Consequently, even if static sector characteristics were related to OEs, it is unlikely this relationship would be detected. The dynamic sector characteristics variables Average Control Duration, Number of Handoffs, Number of Heading Changes, Number of Intersecting Flight Paths, Number of Point Outs, and Number of Transitioning Aircraft (described in detail in the Method section) were submitted to logistic regression analysis to determine the degree to which they could discriminate between OE and RO traffic samples.

Method

Traffic Samples

All traffic samples were initially derived from System Analysis Recordings (SAR) generated by en route Host Computer Systems. The Host 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 reports produced by one such program, the Data Analysis and Reduction Tool (DART).

OE traffic samples were derived from reconfigured DART information from Systematic Air Traffic Operations Research Initiative (SATORI) files. SAR data require a prohibitive amount of storage space. SATORI re-creations require less space and so these files are often the only traffic data saved after an OE. Therefore, the primary constraint on the size and range of the data set was the availability of SATORI re-creations. SATORI data meeting processing criteria (i.e., five minutes prior to the initial loss of separation) were only available for 119 OEs occurring in the ZID airspace from 9/17/2001 through 12/10/2003. Of these, 40 occurred in low-altitude sectors and 79 occurred in the high-altitude sectors.

RO traffic samples were derived from ZID SAR data recorded on 5/8/2003 (15:55 to 17:05, 18:55 to 20:10, and 20:50 to 22:15 ZULU), 5/9/2003 (0:00 to 1:10 ZULU) and 5/10/2003 (11:20 to 12:40 ZULU). DART text reports were first encoded into database files and then processed in 5-minute intervals using custom software designed to calculate objective measures from routinely recorded NAS data. This produced a total of 2644 RO traffic samples. Of these, 992 occurred in low-altitude sectors and 1652 occurred in the high-altitude sectors.

The 40 low-altitude OE traffic samples were combined with 40 randomly-selected low-altitude RO traffic samples to produce a total of 80 traffic samples for the low-altitude sector analysis. The 79 high-altitude OE traffic samples were combined with 79 randomly-selected high-altitude RO traffic samples to produce a total of 158 traffic samples for the high-altitude sector analysis. The number of traffic samples in the RO and OE groups was kept equal because widely disparate group size produces logistic regression models that favor the largest group. Equal group size also ensures that classification accuracy in excess of 50% represents improvement over chance.

Predictor Variables

Average Control Duration. Aircraft control duration is influenced by a number of factors, including aircraft performance characteristics, Traffic Management Initiatives (TMI), and sector size – all of which have been associated with sector workload or complexity (Grossberg, 1989; Mogford, Murphy, & Guttman, 1994; Pfleiderer, Manning, & Goldman, 2007). Average Control Duration is the mean of the durations (in seconds) of all aircraft controlled by the sector within a processing interval. Control time occurring before or after the interval was not included in the calculations.

Number of Handoffs. Although traffic count remains the best single predictor of the number of OEs on a national level, previous research suggests that it is not an effective predictor of OEs at the sector level (Schroeder, 1982; Schroeder, Bailey, Pounds, & Manning, 2006; Spahn, 1977). Perhaps the biggest drawback to traffic count is that it tends to be highly correlated with other traffic-related measures, thereby creating redundancies that may overshadow more effective predictors. Handoffs are correlated with the number of aircraft in the sector, but may also capture elements of communication workload, coordination, and required procedures. The Number of Handoffs is the total number of handoff initiates and handoff accepts occurring within the 5-minute processing interval.

Number of Heading Changes. Heading changes have demonstrated a relationship with controller ratings of activity (e.g., Laudeman et al., 1998), workload (e.g., Stein, 1985), and complexity (e.g., Kopardekar & Magyarits, 2003). 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 holding. The 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. Heading changes made in an attempt to avoid an imminent OE were excluded.

Number of Intersecting Flight Paths. This was one of the highest rated complexity factors in the high-altitude and super high-altitude sectors in the Pfleiderer, Manning, and Goldman (2007) study. A similar factor (several traffic flows converging at the same point) was highly rated in an investigation of Maastricht airspace conducted by Eurocontrol (2006). The Number of Intersecting Flight Paths is the maximum number of flight paths that might be expected to intersect, irrespective of altitude, within a 10-minute projected time frame given the aircraft's current speed and trajectory. Projections were calculated for every 12-second radar update within each minute of data. The length and slope of the projected paths were based on the distance and angle of the current and previous radar position coordinates.

Number of Point Outs. Point out entries represent one of the few instances in which coordination between sectors is recorded. The Number of Point Outs is the total number of point out entries made by the Radar and Radar Associate controllers during the 5-minute processing interval.

Number of Transitioning Aircraft. The amount of climbing and descending traffic has long been recognized as a contributor to the difficulty of working a sector (e.g., Arad, 1964; Grossberg, 1989; Kopardekar & Magyarits, 2003). The Number of Transitioning Aircraft represents the number of aircraft making one or more altitude changes during the 5-minute 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. Altitude changes resulting from last-minute clearances made in an attempt to avoid an OE were excluded.

Results

Stepwise elimination was employed for the logistic regression analyses because such methods are extremely valuable in exploratory research. Backward elimination was used because it is less prone to omit useful variables, since all variables are in the model at the beginning of the process. The likelihood-ratio test, which compares the fit of the model with and without each predictor at every step, was the selection criterion because it is more rigorous than other methods ((Menard, 1995). A criterion level of .10 was used to ensure that all relevant variables are included in the logistic regression model.

Low-Altitude Sector Sample

Tolerance values were >.45 for all predictors, far in excess of the <.20 that would indicate multicollinearity in the low-altitude sector sample. The logistic regression model for the low-altitude sample generated a Model $X^2(3, N=80)=23.82, p<.01$, indicating significantly improved prediction over the model with the constant only. The non-significant Hosmer-Lemeshow $X^2(8, N=80)=1.61$, $p=.99$ suggests that the model fit the data well. Logistic regression coefficients (*B*), standard errors (*S.E.*), estimated odds ratios (Odds), and significance values for the likelihood-ratio tests for the low-altitude sector sample are provided in Table 1. Note that neither the logistic regression coefficients nor standard errors are inflated, indicating a sufficient ratio of cases to predictors.

Table 1. *Logistic Regression Summary: Low-Altitude Sector Sample (N = 80).*

	B	***S.E.***	**Odds**	***p***
Number of Intersecting Flight Paths	1.06	.36	2.89	.00
Number of Point Outs	.45	.24	1.57	.04
Number of Handoffs	.17	.11	1.19	.10
Constant	-1.94	.57	.14	

In the low-altitude sample model, the Number of Intersecting Flight Paths had the highest odds ratio (2.89), followed by the Number of Point Outs (1.57), and the Number of Handoffs (1.19). In other words, each intersecting flight path increased the likelihood that the traffic sample was an OE by 189%, each point out increased the likelihood by 57%, and each handoff increased OE likelihood by 19%. Classification accuracy in the low-altitude sample is shown in Table 2. Of the 40 ROs in the low-altitude sample, 32 (80%) were correctly classified and 8 (20%) were misclassified as OEs. Of the 40 OEs in the sample, 28 (70%) were correctly classified and 12 (30%) were misclassified as ROs. Overall, the low-altitude model had a 75% classification accuracy. This represents 25% improvement over prior probabilities (i.e., the number that would be correctly classified by chance).

Table 2. *Classification: Low-Altitude Sector Sample (N = 80).*

	Predicted		
Observed	Routine Operations	Operational Errors	Total
Routine Operations	32 (80%)	8 (20%)	40
Operational Errors	12 (30%)	28 (70%)	40

High-Altitude Sector Sample

As with the low-altitude sample, Tolerance values were high (.56 and above) for all predictors. The logistic regression model for the high-altitude sample generated a Model $X^2(4, N=158) = 73.01$, $p<.01$, indicating significantly improved prediction over the model with the constant only. The non-significant Hosmer-Lemeshow X^2 $(8, N=158) = 3.33$, $p=.91$ for the high-altitude sample verified that the model fit the data. Logistic regression coefficients (*B*), standard errors (*S.E.*), estimated odds ratios (Odds), and significance values for the likelihood-ratio tests for the high-altitude sector sample are provided in Table 3. Note that neither the logistic regression coefficients nor standard errors are inordinately large, indicating a sufficient ratio of cases to predictors.

Table 3. *Logistic Regression Summary: High-Altitude Sector Sample (N = 158).*

	B	***S.E.***	**Odds**	***p***
Number of Intersecting Flight Paths	.69	.28	2.00	.01
Number of Heading Changes	.31	.15	1.36	.03
Number of Transitioning Aircraft	.24	.11	1.27	.02
Average Control Duration	.01	.01	1.01	.01
Constant	-4.43	1.11	.01	

In the high-altitude sample model, the Number of Intersecting Flight Paths had the highest odds ratio (2.00), followed by the Number of Heading Changes (1.36), the Number of Transitioning Aircraft (1.27), and Average Control Duration (1.01). In other words, each one-unit increase in the Number of Intersecting Flight Paths increased the likelihood that a traffic sample was an OE by 100%, each one-unit increase in the Number of Heading Changes increased the likelihood by 36%, every Transitioning Aircraft increased the likelihood by 27%, and each one-second increase in Average Control Duration increased the likelihood by 1%. Classification accuracy in the high-altitude sample, shown in Table 4, was slightly better than that of the low-altitude sample. Of the 79 ROs in the high-altitude sample, 64 (81%) were correctly classified and 15 (19%) were misclassified as OEs. Of the 79 OEs in the sample, 60 (76%) were correctly classified and 19 (24%) were misclassified as ROs. Overall, the high-altitude model had 79% classification accuracy. This represents 29% improvement over prior probabilities (i.e., the number that would be correctly classified by chance).

Table 4. *Classification: High-Altitude Sector Sample (N = 158).*

Observed	Predicted: Routine Operations	Predicted: Operational Errors	Total
Routine Operations	64 (81%)	15 (19%)	79
Operational Errors	19 (24%)	60 (76%)	79

Discussion

The results of the logistic regression analyses indicate that a sufficient model may be constructed from sector characteristics variables. Overall classification accuracy between 75-79% is remarkable for models constructed solely of environmental and contextual factors. After all, other factors (e.g., human elements, organizational influences) also contribute to the occurrence of OEs. Unfortunately, all the logistic regression models were better at classifying ROs than OEs. Classification of OEs ranged from as low as 70% in the low-altitude sector sample, to 76% in the high-altitude sample.

Low-Altitude Sector Model

The most influential variable in the low-altitude sector model was the Number of Intersecting Flight Paths (Odds=2.89), followed by the Number of Point Outs (Odds=1.57), and the Number of Handoffs (Odds = 1.19). In Pfleiderer and Manning (2007), the most influential predictor was the Number of Point Outs (Odds=3.30), followed by the Number of Handoffs (Odds=1.54), and the Number of Heading Changes (Odds=1.49). The predictive strength of the Number of Point Outs and the Number of Handoffs in the Pfleiderer and Manning (2007) results suggested that coordination played a primary role in the development of OEs in the ZID low-altitude sectors. This impression was bolstered by the Pfleiderer et al. (2007) data, in which controllers and supervisors at ZID rated coordination as one of the primary sources of complexity in low-altitude sectors. Consequently, the emergence of the Number of Intersecting Flight Paths as the most influential predictor in the current low-altitude logistic regression model was surprising, because ratings for this complexity factor were moderate in the low-altitude sectors. The results of the logistic regression analysis suggest that coordination may be a contributing factor, but converging traffic patterns are of greater consequence.

High-Altitude Sector Model

The Number of Intersecting Flight Paths was the most influential predictor in the high-altitude sector model, followed by the Number of Heading Changes. This is consistent with previous research. Controllers and supervisors rated the Number of Intersecting Flight Paths as one of the most influential complexity factors in the high- and super high-altitude sectors (Pfleiderer et al., 2007), and the Number of Heading Changes received the highest beta weight in a linear multiple regression analysis of controller ratings of activity in four sectors at the Denver ARTCC. Laudeman et al. (1998) attributed the influence of heading changes to the “significant arrival traffic in all the sectors that were observed” (p. 7). Arrival and departure traffic complexity is generally considered to be a low-altitude phenomenon, but this perception may be inaccurate. In the present study, the Number of Heading Changes was only influential in the high-altitude sector model. The third most influential factor in the high-altitude logistic regression analysis was the Number of Transitioning Aircraft, which has long been recognized as a contributor to the difficulty of working a sector (e.g., Arad, 1964; Grossberg, 1989; Kopardekar & Magyarits, 2003). This finding is also consistent with Pfleiderer et al. (2007) in which the complexity factor Climbing and Descending Traffic received the highest complexity rating for the high-and super high-altitude sectors.

Future Research

Logistic regression cannot be used to directly identify causal factors (i.e., prediction is not the same as causation), but elements of the models reveal aspects of the sector environment that might be altered to reduce the number of OEs. For example, the combination of the Number of Point Outs and the Number of Handoffs in the low-altitude sector model may indicate that the location of sector boundaries increases coordination workload and complexity. On the other hand, the combination of the Number of Point Outs and the Number of Intersecting Flight Paths may point to problems with the orientation of traffic paths relative to those boundaries.

Because of the research that remains to be accomplished, these results must be viewed as preliminary. Multiple studies must be conducted at a number of facilities before such models might be viable for practical applications. Nevertheless, the methodology of comparing OE and RO traffic samples is promising. Continued investigations along these lines may highlight complexity factors that should be addressed to ensure that safety is maintained.

References

Arad, B.A. (1964, May). The control load and sector design. *Journal of Air Traffic Control,* 5, 12-31.

Eurocontrol (2006). *A complexity study of the Maastricht upper airspace centre.* (EEC Report No. 403). Brétigny-sur-Orge, France: Eurocontrol Experimental Centre.

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.

Kopardekar, P., & Magyarits, S. (2003, June), Measurement and prediction of dynamic density. *Proceedings of the 5th Eurocontrol/FAA Air Traffic Management Research and Development Conference*, Budapest, Hungary. Available online at: http://atm2003.eurocontrol.fr.

Laudeman, I.V., Shelden, S.G., Branstrom, R., & Brasil, C.L. (1998). *Dynamic density: An air traffic management metric.* (NAS/TM-1998-112226). Moffett Field, CA: NASA Ames.

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

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.

Pfleiderer, E.M., & Manning, C.A. (2007). *Prediction and classification of operational errors and routine operations using sector characteristics variables.* (Report No. DOT/FAA/AM-07/18). Washington, DC: FAA Office of Aerospace Medicine.

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

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.

Schroeder, D., Bailey, L., Pounds, J., & Manning, C. (2006). *A human factors review of the operational error literature.* (Report No. DOT/FAA/AM-06/21). Washington, DC: FAA Office of Aerospace Medicine.

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.

Stein, E. (1985), *Air traffic controller workload: An examination of workload probe.* (Report No. DOT/FAA/CT-TN84/24). Atlantic City, NJ: FAA William J. Hughes Technical Center.